Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3132374 A1 2020/09/10

(21) 3 132 374

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

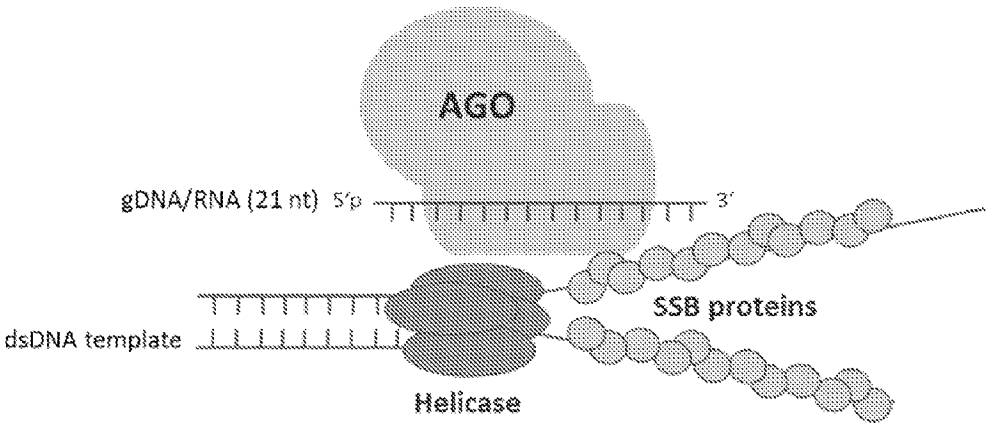
(86) Date de dépôt PCT/PCT Filing Date: 2020/03/05
(87) Date publication PCT/PCT Publication Date: 2020/09/10
(85) Entrée phase nationale/National Entry: 2021/09/01
(86) N° demande PCT/PCT Application No.: US 2020/021163
(87) N° publication PCT/PCT Publication No.: 2020/181072
(30) Priorité/Priority: 2019/03/06 (US62/814,787)

(51) Cl.Int./Int.Cl. *C12N 15/10* (2006.01), *C07K 14/33* (2006.01), *C12N 15/09* (2006.01), *C12N 15/63* (2006.01), *C12N 9/22* (2006.01), *C12P 19/34* (2006.01)

(71) Demandeurs/Applicants:
THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, US;
INTIMA BIOSCIENCE, INC., US

(72) Inventeurs/Inventors:
QI, LEI S., US;
CHOUDHRY, MODASSIR, US;
LIN, XUEQIU, US;
COLLINGWOOD, TREVOR N., US; ...

(54) Titre : SYSTEMES POUR ARGONAUTES MESOPHILES ET LEURS UTILISATIONS
(54) Title: MESOPHILIC ARGONAUTE SYSTEMS AND USES THEREOF



dsDNA nicking/cleavage

FIG. 44

(57) Abrégé/Abstract:
Constructs comprising Argonautes and neighboring genes are disclosed for use in gene editing. Disclosed are also compositions and methods utilizing these Argonautes and neighboring genes. Also disclosed are the methods of making and using the Argonautes and neighboring genes in treating various diseases, conditions, and cancer.

(21) 3 132 374

(13) A1

(72) Inventeurs(suite)/Inventors(continued): HENLEY, THOMAS, GB; KLAPHOLZ, BENJAMIN, GB; BUERCKSTUEMMER, TILMANN, AT; SALIC, SEJLA, AT

(74) Agent: SMART & BIGGAR LLP

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 September 2020 (10.09.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/181072 A1

(51) International Patent Classification:
C12N 15/10 (2006.01) *C12N 15/63* (2006.01)
C12N 9/22 (2006.01) *C12N 15/09* (2006.01)
C07K 14/33 (2006.01) *C12P 19/34* (2006.01)

(21) International Application Number:
PCT/US2020/021163

(22) International Filing Date:
05 March 2020 (05.03.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/814,787 06 March 2019 (06.03.2019) US

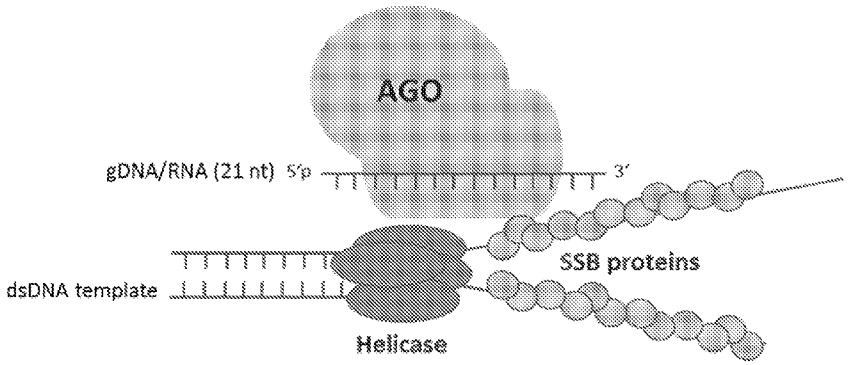
(71) Applicants: **THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY** [US/US]; Office of the General Counsel Building 170, Third Floor, Main Quad P.O. Box 20386, Stanford, California 94305-2038 (US). **INTIMA BIOSCIENCE, INC.** [US/US]; 3 Columbus Circle, New York, New York 10019 (US).

(72) Inventors: **QI, Lei S.**; 515 Olmstead Road, Stanford, California 94305 (US). **CHOUDHRY, Modassir**; 106 Central Park South, #29A, New York, New York 10019 (US). **LIN, Xueqiu**; 12 Pacific Bay Circle, #302, San Bruno, California 94066 (US). **COLLINGWOOD, Trevor N.**; 1111 Cambridge Street, Novato, California 94947 (US). **HENLEY, Thomas**; 75 Cyprian Rust Way, Soham Cambridgeshire CB7 5ZE (GB). **KLAPHOLZ, Benjamin**; 71 Perne Road, Cambridge Cambridgeshire CB1 3RX (GB). **BUERCK-STUEMMER, Tilmann**; Corneliusgasse 3/26, 1060 Vienna (AT). **SALIC, Sejla**; Beingasse 5-9/1/4, 1150 Vienna (AT).

(74) Agent: **HAO, Joe C.** et al.; Kilpatrick Townsend & Stockton LLP, Mailstop: IP Docketing - 22, 1100 Peachtree Street, Suite 2800, Atlanta, GA 30309 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,

(54) Title: MESOPHILIC ARGONAUTE SYSTEMS AND USES THEREOF



dsDNA nicking/cleavage

FIG. 44

(57) Abstract: Constructs comprising Argonautes and neighboring genes are disclosed for use in gene editing. Disclosed are also compositions and methods utilizing these Argonautes and neighboring genes. Also disclosed are the methods of making and using the Argonautes and neighboring genes in treating various diseases, conditions, and cancer.

[Continued on next page]

WO 2020/181072 A1

WO 2020/181072 A1

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).

Declarations under Rule 4.17:

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

MESOPHILIC ARGONAUTE SYSTEMS AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 62/814,787 filed on March 6, 2019, the disclosure of which is hereby incorporated by reference in its entirety.

BACKGROUND

[0002] With the rapid progress being made in genome sciences, effective genome engineering holds great promise both in understanding the molecular bases of human diseases and in treating human disorders with identifiable alterations in the genome. The past few years have witnessed a rapid rise of the RNA-guided CRISPR/Cas9 technology from obscurity. Significant efforts are being devoted to optimizing the current CRISPR/Cas9 system and to identifying more Cas9-like nucleases with better efficiency and specificity. Similarly, significant efforts are being employed to identify new systems that can be harnessed for genome editing with improved specificity and efficiency.

INCORPORATION BY REFERENCE

[0003] All publications, patents, and patent applications herein are incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. In the event of a conflict between a term herein and a term in an incorporated reference, the term herein controls.

SUMMARY

[0004] In one aspect, provided herein are systems comprising: a. an Argonaute (Ago) polypeptide, or a polynucleic acid encoding the same, wherein said Ago polypeptide is a Clostridia Ago polypeptide, or a functional fragment or functional variant thereof; and b. a non-naturally occurring guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence.

[0005] In some embodiments, the Ago polypeptide is a mesophilic Clostridia Ago polypeptide.

[0006] In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid.

[0007] In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 40 °C, 19 °C to about 50 °C, 19 °C to about 60 °C, 19 °C to about 70 °C, 19 °C to about 80 °C, 20 °C to about 40 °C, 20 °C to about 30 °C, 20 °C to about 50 °C, 20 °C to about 60 °C, 20 °C to about 70 °C, 20 °C to about 80 °C, 25 °C to about 40 °C, 25 °C to about 30 °C, or 25 °C to about 50 °C. In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 19 °C, 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, or 40°C. In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 37 °C. In some embodiments, the Ago polypeptide demonstrates a maximal nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 45 °C, 19 °C to about 40 °C, 20 °C to about 45 °C, 25 °C to about 45 °C, 30 °C to about 45 °C, or 30 °C to about 40 °C, as compared to nucleic acid-cleaving activity at a different temperature.

[0008] In some embodiments, the nucleic acid-cleaving activity of the target polynucleic acid is directed by the guiding polynucleic acid.

[0009] In some embodiments, the Ago polypeptide demonstrates one, two, three, or four of: single stranded DNA (ssDNA) cleaving activity, double stranded DNA (dsDNA) cleaving activity, single stranded RNA (ssRNA) cleaving activity, or double stranded RNA (dsRNA) cleaving activity. In some embodiments, the Ago polypeptide demonstrates single stranded DNA (ssDNA) cleaving activity In some embodiments, the target polynucleic acid is a single stranded DNA (ssDNA) sequence, a double stranded DNA (dsDNA) sequence, a single stranded RNA (ssRNA) sequence, or a double stranded RNA (dsRNA) sequence. In some embodiments, the target polynucleic acid is a single stranded DNA (ssDNA) sequence.

[0010] In some embodiments, the target polynucleic acid is DNA.

[0011] In some embodiments, a region of the target DNA sequence that the Ago polypeptide cleaves is about at least 50%, 60%, 70%, 80%, or 90% deoxyadenosine and deoxythymidine.

[0012] In some embodiments, said target polynucleic acid comprises a gene sequence. In some embodiments, said Ago polypeptide produces a disruption in said gene sequence when introduced into a cell. In some embodiments, said disruption comprises a double strand break or a single strand break.

[0013] In some embodiments, said guiding polynucleic acid is capable of interacting with said Ago polypeptide and directing said Ago polypeptide to said target polynucleic acid. In some embodiments, the guiding polynucleic acid is a guide DNA or a guide RNA. In some embodiments, said guiding polynucleic acid is from about 1 nucleotide to about 30 nucleotides in length.

[0014] In some embodiments, said system comprises a complex, and wherein said complex comprises said Ago polypeptide and said guiding polynucleic acid.

[0015] In some embodiments, the Ago polypeptide comprises a PIWI-like domain. In some embodiments, the Ago polypeptide comprises a PIWI domain. In some embodiments, the Ago polypeptide comprises a PAZ domain. In some embodiments, the Ago polypeptide comprises a PAZ-like domain.

[0016] In some embodiments, the Ago polypeptide is an Ago polypeptide, or a functional fragment or a functional variant thereof, from: Candidatus Comantemales, Clostridiales, Halanaerobiales, Natranaerobiales, Thermoanaerobacterales, or Negativicutes,.

[0017] In some embodiments, the Ago polypeptide is an Ago polypeptide, or a functional fragment or a functional variant thereof, from: Caldicoprobacteraceae, Christensenellaceae, Clostridiaceae, Defluviitaleaceae, Eubacteriaceae, Graciibacteraceae, Heliobacteriaceae, Lachnospiraceae, Oscillospiraceae, Peptococcaceae, Peptostreptococcaceae, Ruminococcaceae, Syntrophomonadaceae, Halanaerobiaceae, Halobacteroidaceae, Natranaerobiaceae, Thermoanaerobacteraceae, or Thermodesulfobiaceae.

[0018] In some embodiments, the Ago polypeptide is a Clostridiaceae Ago polypeptide, or a functional fragment or a functional variant thereof.

[0019] In some embodiments, the Ago polypeptide is a Clostridium, Acetanaerobacterium, Acetivibrio, Acidaminobacter, Alkaliphilus, Anaerobacter, Anaerostipes, Anaerotruncus, Anoxynatronum, Bryantella, Butyricicoccus, Caldanaerocella, Caldisalinibacter, Caloramator, Caloranaerobacter, Caminicella, Candidatus Arthromitus, Cellulosibacter, Coprobacillus, Crassaminicella, Dorea, Ethanologenbacterium, Faecalibacterium, Garciella, Guggenheimella, Hespellia, Linmingia, Natronincola, Oxobacter, Parasporobacterium, Sarcina, Soehngenia, Sporobacter, Subdoligranulum, Tepidibacter, Tepidimicrobium, Thermobrachium, Thermohalobacter, or Tindallia Ago polypeptide, or a functional fragment or a functional variant thereof.

[0020] In some embodiments, the Ago polypeptide is a Clostridium Ago polypeptide, or a functional fragment or a functional variant thereof.

[0021] In some embodiments, the Ago polypeptide is a Clostridium absonum, Clostridium aceticum, Clostridium acetireducens, Clostridium acetobutylicum, Clostridium acidisoli, Clostridium aciditolerans, Clostridium acidurici, Clostridium aerotolerans, Clostridium aestuarii, Clostridium akagii, Clostridium aldenense, Clostridium aldrichii, Clostridium algidicarnis, Clostridium algidixylanolyticum, Clostridium algifaecis, Clostridium algoriphilum, Clostridium alkalicellulosi, Clostridium amazonense, Clostridium aminophilum, Clostridium aminovalericum, Clostridium amygdalinum, Clostridium amylolyticum, Clostridium arbusti, Clostridium arcticum, Clostridium argentinense, Clostridium asparagiforme, Clostridium aurantibutyricum, Clostridium baratii, Clostridium barkeri, Clostridium bartlettii, Clostridium beijerinckii, Clostridium bifermentans, Clostridium bolteae, Clostridium bornimense, Clostridium botulinum, Clostridium bowmanii, Clostridium bryantii, Clostridium budayi, Clostridium butyricum, Clostridium cadaveris, Clostridium caenicola, Clostridium caminithermale, Clostridium carboxidivorans, Clostridium carnis, Clostridium cavendishii, Clostridium celatum, Clostridium celerecrescens, Clostridium cellobioparum, Clostridium cellulofermentans, Clostridium cellulolyticum, Clostridium cellulosi, Clostridium cellulovorans, Clostridium chartatabidum, Clostridium chauvoei, Clostridium chromiireducens, Clostridium citroniae, Clostridium clariflavum, Clostridium clostridioforme, Clostridium coccoides, Clostridium cochlearium, Clostridium cocleatum, Clostridium colicanis, Clostridium colinum, Clostridium collagenovorans, Clostridium combesii, Clostridium cylindrosporum, Clostridium difficile, Clostridium diolis, Clostridium disporicum, Clostridium drakei, Clostridium durum, Clostridium estertheticum, Clostridium estertheticum subsp. Estertheticum, Clostridium estertheticum subsp. Laramiense, Clostridium fallax, Clostridium felsineum, Clostridium fervidum, Clostridium fimetarium, Clostridium formicaceticum, Clostridium frigidicarnis, Clostridium frigoris, Clostridium ganghwense, Clostridium gasigenes, Clostridium ghonii, Clostridium glycolicum, Clostridium glycyrrhizinilyticum, Clostridium grantii, Clostridium guangxiense, Clostridium haemolyticum, Clostridium halophilum, Clostridium hastiforme, Clostridium hathewayi, Clostridium herbivorans, Clostridium hiranonis, Clostridium histolyticum, Clostridium homopropionicum, Clostridium huakuii, Clostridium hungatei, Clostridium hydrogeniformans, Clostridium hydroxybenzoicum, Clostridium hylemonae, Clostridium indolis, Clostridium innocuum, Clostridium intestinale, Clostridium irregulare,

Clostridium isatidis, Clostridium jeddahense, Clostridium jejuense, Clostridium josui, Clostridium kluyveri, Clostridium lactatifermentans, Clostridium lacusfryxellense, Clostridium laramiense, Clostridium lavalense, Clostridium lentocellum, Clostridium lentoputrescens, Clostridium leptum, Clostridium limosum, Clostridium liquoris, Clostridium litorale, Clostridium lituseburense, Clostridium ljungdahlii, Clostridium lortetii, Clostridium lundense, Clostridium luticellarii, Clostridium magnum, Clostridium malenominatum, Clostridium mangenotii, Clostridium maximum, Clostridium mayombei, Clostridium methoxybenzovorans, Clostridium methylpentosum, Clostridium moniliforme, Clostridium neonatale, Clostridium neopropionicum, Clostridium neuense, Clostridium nexile, Clostridium nitritogenes, Clostridium nitrophenolicum, Clostridium novyi, Clostridium oceanicum, Clostridium orbiscindens, Clostridium oroticum, Clostridium oryzae, Clostridium oxalicum, Clostridium pabulibutyricum, Clostridium papyrosolvens, Clostridium paradoxum, Clostridium paraperfringens, Clostridium paraputrificum, Clostridium pascui, Clostridium pasteurianum, Clostridium peptidivorans, Clostridium perenne, Clostridium perfringens, Clostridium pfennigii, Clostridium phytofermentans, Clostridium piliforme, Clostridium polyendosporum, Clostridium polysaccharolyticum, Clostridium populeti, Clostridium propionicum, Clostridium proteoclasticum, Clostridium proteolyticum, Clostridium psychrophilum, Clostridium punense, Clostridium puniceum, Clostridium purinilyticum, Clostridium putrefaciens, Clostridium putrificum, Clostridium quercicolum, Clostridium quinii, Clostridium ramosum, Clostridium rectum, Clostridium roseum, Clostridium saccharobutylicum, Clostridium saccharogumia, Clostridium saccharolyticum, Clostridium saccharoperbutylacetonicum, Clostridium sardiniense, Clostridium sartagoforme, Clostridium saudiense, Clostridium scatologenes, Clostridium schirmacherense, Clostridium scindens, Clostridium senegalense, Clostridium septicum, Clostridium sordellii, Clostridium sphenoides, Clostridium spiroforme, Clostridium sporogenes, Clostridium sporosphaeroides, Clostridium stercorarium, Clostridium stercorarium subsp. Leptospartum, Clostridium stercorarium subsp. Stercorarium, Clostridium stercorarium subsp. Thermolacticum, Clostridium sticklandii, Clostridium straminisolvens, Clostridium subterminale, Clostridium sufflavum, Clostridium sulfidigenes, Clostridium swellfunianum, Clostridium symbiosum, Clostridium tarantellae, Clostridium tagluense, Clostridium tepidiprofundi, Clostridium tepidum, Clostridium termitidis, Clostridium tertium, Clostridium tetani, Clostridium tetanomorphum, Clostridium thermaceticum, Clostridium thermautotrophicum, Clostridium thermoalcaliphilum, Clostridium thermobutyricum, Clostridium thermocellum, Clostridium

thermocopriae, Clostridium thermohydrosulfuricum, Clostridium thermolacticum, Clostridium thermopalmarium, Clostridium thermopapyrolyticum, Clostridium thermosaccharolyticum, Clostridium thermosuccinogenes, Clostridium thermosulfurigenes, Clostridium thiosulfatireducens, Clostridium tyrobutyricum, Clostridium uliginosum, Clostridium ultunense, Clostridium ventriculi, Clostridium villosum, Clostridium vincentii, Clostridium viride, Clostridium vulturis, and Clostridium xylanolyticum, or Clostridium xylanovorans Ago polypeptide, or a functional fragment or a functional variant thereof.

[0022] In some embodiments, the Ago polypeptide is a Clostridium perfringens, Clostridium butyricum, Clostridium saudiense, or Clostridium disporicum Ago polypeptide, or a functional fragment or a functional variant thereof.

[0023] In some embodiments, said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 1-3 or 134-136.

[0024] In some embodiments, said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity with one of SEQ ID NOs: 11-14 or 137-139.

[0025] In some embodiments, said system comprises a nucleic acid unwinding polypeptide or a polynucleic acid encoding the same.

[0026] In some embodiments, said nucleic acid unwinding polypeptide is a helicase, a single strand DNA binding (SSB) protein, or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

[0027] In some embodiments, said nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

[0028] In some embodiments, said nucleic acid unwinding polypeptide is a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain. In some embodiments, said Cas protein domain is a catalytically dead Cas polypeptide.

[0029] In some embodiments, said Ago polypeptide is fused either directly or indirectly to a nuclear localization signal (NLS). In some embodiments, said nucleic acid unwinding polypeptide is fused either directly or indirectly to a NLS.

[0030] In some embodiments, said Ago polypeptide and said nucleic acid unwinding polypeptide are fused either directly or indirectly. In some embodiments, said Argonaute polypeptide and said nucleic acid unwinding polypeptide are fused and a NLS is in between said Ago polypeptide and said nucleic acid unwinding polypeptide.

[0031] In some embodiments, said Ago polypeptide is encoded by a gene located in an adjacent operon to at least one of a gene involved in defense, stress response, gene editing, CRISPR, DNA replication, DNA recombination, DNA repair, and transcription.

[0032] In some embodiments, said system comprises one or more recombinant expression vectors. In some embodiments, said one or more recombinant expression vectors comprise an adeno-associated virus vector, a plasmid vector, a retroviral vector, a lentiviral vector, an adenovirus vectors, a poxvirus vectors, a herpesvirus vector, or a split-intron vector.

[0033] In some embodiments, said Ago polypeptide, or functional fragment or variant thereof, comprises a DEDX motif sequence. In some embodiments, said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

[0034] In one aspect, provided herein is ex vivo cell (or population of cells) comprising a system described herein. In some embodiments, the cell is a human cell. In some embodiments, the cell is an immune cell, a stem cell, or a germ cell.

[0035] In one aspect, provided herein is a recombinant expression vector encoding a system described herein.

[0036] In one aspect, provided herein is a pharmaceutical composition comprising a system described herein, and at least one of: an excipient, a diluent, or a carrier. In some embodiments, said pharmaceutical composition is in a form of intravenous, subcutaneous, or intramuscular administration formulation.

[0037] In one aspect, provided herein is a kit comprising: (a) a system described herein (b) instructions for use thereof, and optionally (c) a container.

[0038] In one aspect, provided herein are polypeptide constructs, wherein said constructs comprise a mesophilic Clostridia Ago (C-Ago) polypeptide sequence, or a functional fragment or a functional variant thereof, wherein said C-Ago polypeptide sequence cleaves a nucleic acid in a target polynucleic acid sequence at a mesophilic temperature, wherein said target polynucleic acid sequence is bound by a non-naturally occurring guide polynucleic acid sequence.

[0039] In some embodiments, said C-Ago polypeptide sequence or functional fragment or variant thereof comprises a DEDX motif sequence. In some embodiments, said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said C-Ago polypeptide as compared to a corresponding C-Ago polypeptide without said mutation in said DEDX motif sequence.

[0040] In one aspect, provided herein is a nucleic acid molecule encoding a polypeptide construct described herein.

[0041] In one aspect, provided herein are recombinant fusion polypeptides, wherein said fusion polypeptides comprise: (a) an Argonaute (Ago) polypeptide, wherein said Ago polypeptide is a Clostridia Ago (C-Ago) polypeptide; and (b) a nucleic acid unwinding polypeptide.

[0042] In some embodiments, the nucleic acid unwinding polypeptide comprises a helicase, a single strand DNA binding protein (SSB) polypeptide, or a Cas protein domain.

[0043] In some embodiments, the nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

[0044] In some embodiments, said nucleic acid unwinding polypeptide is a Cas protein domain. In some embodiments, said Cas protein domain is a catalytically dead Cas polypeptide.

[0045] In some embodiments, said fusion polypeptide comprises at least one nuclear localization signal (NLS) polypeptide. In some embodiments, said fusion polypeptide comprises at least two, three, or four NLSs polypeptides. In some embodiments, said fusion polypeptide comprises a nuclear localization signal between said nucleic acid unwinding polypeptide and said C-Ago.

[0046] In some embodiments, said C-Ago polypeptide comprises a DEDX motif sequence. In some embodiments, said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said C-Ago polypeptide as compared to a corresponding C-Ago polypeptide without said mutation in said DEDX motif sequence.

[0047] In one aspect, provided herein is a nucleic acid encoding a recombinant fusion polypeptide described herein.

[0048] In one aspect, provided herein are methods of modifying a target polynucleic acid, said methods comprising: introducing into a cell a system described herein; or a polypeptide construct described herein; or a recombinant fusion polypeptide described herein and a non-naturally occurring guiding polynucleic acid that is complementary to said target polynucleic acid; and modifying said target polynucleic acid.

[0049] In one aspect, provided herein are methods of treating a disease or disorder in a subject in need thereof, said method comprising administering to the subject: system described herein, a polypeptide construct described herein, a recombinant fusion polypeptide described herein, a cell described herein, a vector described herein, or a pharmaceutical composition described herein. In some embodiments, said disease is cancer, an autoimmune disease, a genetic disease, or an infection. In some embodiments, said disease is cancer.

[0050] In one aspect, provided herein are systems comprising: a mesophilic Argonaute (Ago) polypeptide, or a polynucleic acid encoding the same, or a functional fragment or variant thereof; and an exogenous non-naturally occurring guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence.

[0051] In some embodiments, said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%sequence identity with one of SEQ ID NOs: 4-10 or 134-136. In some embodiments, said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 15-21.

[0052] In some embodiments, said Ago polypeptide comprises a DEDX motif sequence. In some embodiments, said DEDX motif sequence comprises a mutation, wherein said mutation reduces

catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

[0053] In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid.

[0054] In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 40 °C, 19 °C to about 50 °C, 19 °C to about 60 °C, 19 °C to about 70 °C, 19 °C to about 80 °C, 20 °C to about 40 °C, 20 °C to about 30 °C, 20 °C to about 50 °C, 20 °C to about 60 °C, 20 °C to about 70 °C, 20 °C to about 80 °C, 25 °C to about 40 °C, 25 °C to about 30 °C, or 25 °C to about 50 °C. In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 19 °C, 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, or 40°C. In some embodiments, the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 37 °C. In some embodiments, the Ago polypeptide demonstrates a maximal nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 45 °C, 19 °C to about 40 °C, 20 °C to about 45 °C, 25 °C to about 45 °C, 30 °C to about 45 °C, or 30 °C to about 40 °C, as compared to nucleic acid-cleaving activity at a different temperature.

[0055] In some embodiments, the nucleic acid-cleaving activity of the target polynucleic acid is directed by the guiding polynucleic acid. In some embodiments, the Ago polypeptide demonstrates one, two, three, or four of: single stranded DNA (ssDNA) cleaving activity, double stranded DNA (dsDNA) cleaving activity, single stranded RNA (ssRNA) cleaving activity, or double stranded RNA (dsRNA) cleaving activity. In some embodiments, the Ago polypeptide demonstrates single stranded DNA (ssDNA) cleaving activity. In some embodiments, the target polynucleic acid is a single stranded DNA (ssDNA) sequence, a double stranded DNA (dsDNA) sequence, a single stranded RNA (ssRNA) sequence, or a double stranded RNA (dsRNA) sequence. In some embodiments, the target polynucleic acid is a single stranded DNA (ssDNA) sequence.

[0056] In some embodiments, the target polynucleic acid is DNA. In some embodiments, a region of the target DNA sequence that the C-Ago polypeptide cleaves is about at least 50%, 60%, 70%, 80%, or 90% deoxyadenosine and deoxythymidine.

[0057] In some embodiments, said target polynucleic acid comprises a gene sequence. In some embodiments, said Ago polypeptide sequence produces a disruption in said gene sequence when introduced into a cell. In some embodiments, said disruption comprises a double strand break or a single strand break.

[0058] In some embodiments, said guiding polynucleic acid is capable of interacting with said Ago polypeptide and directing said Ago polypeptide to said target polynucleic acid.

[0059] In some embodiments, the guiding polynucleic acid is a guide DNA or a guide RNA.

[0060] In some embodiments, said guiding polynucleic acid is from about 1 nucleotide to about 30 nucleotides in length.

[0061] In some embodiments, said system comprises a complex, and wherein said complex comprises said Ago polypeptide and said guiding polynucleic acid.

[0062] In some embodiments, the Ago polypeptide comprises a PIWI-like domain. In some embodiments, the Ago polypeptide comprises a PIWI domain. In some embodiments, the Ago polypeptide comprises a PAZ domain. In some embodiments, the Ago polypeptide comprises a PAZ-like domain.

[0063] In some embodiments, said system comprises a nucleic acid unwinding polypeptide or a polynucleic acid encoding the same. In some embodiments, said nucleic acid unwinding polypeptide is a helicase, a single strand DNA binding (SSB) protein, or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

[0064] In some embodiments, said nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

[0065] In some embodiments, said nucleic acid unwinding polypeptide is a Cas protein domain. In some embodiments, said Cas protein domain is a catalytically dead Cas polypeptide.

[0066] In some embodiments, said Ago polypeptide is fused either directly or indirectly to a NLS. In some embodiments, said nucleic acid unwinding polypeptide is fused either directly or indirectly to a NLS. In some embodiments, said Ago polypeptide and said nucleic acid unwinding polypeptide are fused either directly or indirectly. In some embodiments, said Ago polypeptide and said nucleic acid unwinding polypeptide are fused and a NLS is in between said Ago polypeptide and said nucleic acid unwinding polypeptide.

[0067] In some embodiments, said Ago polypeptide is encoded by a gene located in an adjacent operon to at least one of a gene involved in defense, stress response, gene editing, CRISPR, DNA replication, DNA recombination, DNA repair, and transcription.

[0068] In some embodiments, said system comprises one or more recombinant expression vectors. In some embodiments, said one or more recombinant expression vectors comprise an adeno-associated virus vector, a plasmid vector, a retroviral vector, a lentiviral vector, an adenovirus vectors, a poxvirus vectors, a herpesvirus vector, or a split-intron vector.

[0069] In one aspect, provided herein is an ex vivo cell (or population thereof) comprising a system described herein. In some embodiments, the cell is a human cell. In some embodiments, the cell is an immune cell, a stem cell, or a germ cell.

[0070] In one aspect, provided herein is a recombinant expression vector encoding a system described herein.

[0071] In one aspect, provided herein is a pharmaceutical composition comprising a system described herein, and at least one of: an excipient, a diluent, or a carrier.

[0072] In some embodiments, said pharmaceutical composition is in a form of intravenous, subcutaneous, or intramuscular administration formulation.

[0073] In one aspect, provided herein is a kit comprising: (a) a system described herein; and (b) instructions for use thereof, and optionally (c) a container.

[0074] In one aspect, provided herein are polypeptide constructs, wherein said constructs comprise a mesophilic Ago polypeptide sequence, or a functional fragment or a functional variant thereof, wherein said Ago polypeptide sequence cleaves a nucleic acid in a target polynucleic acid sequence at a mesophilic temperature, wherein said target polynucleic acid sequence is bound by a non-naturally occurring guide polynucleic acid sequence.

[0075] In some embodiments, said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 4-10. In some embodiments, said Ago polypeptide is encoded by a polynucleic acid

sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 15-21.

[0076] In some embodiments, said Ago polypeptide comprises a DEDX motif sequence. In some embodiments, said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

[0077] In one aspect, provided herein is a nucleic acid sequence encoding a polypeptide described herein.

[0078] In one aspect, provided herein are recombinant fusion polypeptides, said fusion polypeptides comprising: a mesophilic Argonaute (Ago) polypeptide; and a nucleic acid unwinding polypeptide.

[0079] In some embodiments, the nucleic acid unwinding polypeptide comprises a helicase, a single strand DNA binding protein (SSB), or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

[0080] In some embodiments, said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 4-10. In some embodiments, said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity with one of SEQ ID NOs: 15-21.

[0081] In some embodiments, the nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49. In some embodiments, said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22. In some embodiments, said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

[0082] In some embodiments, said nucleic acid unwinding polypeptide is a Cas protein domain. In some embodiments, said Cas protein domain is a catalytically dead Cas polypeptide.

[0083] In some embodiments, said fusion polypeptide comprises at least one nuclear localization signal (NLS) polypeptide. In some embodiments, said fusion polypeptide comprises at least two, three, or four NLS polypeptides. In some embodiments, said fusion polypeptide comprises a NLS between said nucleic acid unwinding polypeptide and said Ago polypeptide.

[0084] In some embodiments, said Ago polypeptide comprises a DEDX motif sequence. In some embodiments, said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

[0085] In one aspect, provided herein is a nucleic acid encoding a recombinant fusion polypeptide described herein.

[0086] In one aspect, provided herein are methods of modifying a target polynucleic acid, said methods comprising: introducing into a cell a system described herein; or a polypeptide construct described herein; or a recombinant fusion polypeptide described herein, and a non-naturally occurring guiding polynucleic acid that is complementary to said target polynucleic acid; and modifying said target polynucleic acid.

[0087] In one aspect, provided herein are methods of treating a disease or disorder in a subject in need thereof, said method comprising administering to the subject: a system described herein, a polypeptide construct described herein, a recombinant fusion polypeptide described herein, a cell described herein, a vector described herein, or a pharmaceutical composition described herein. In some embodiments, said disease is cancer, an autoimmune disease, a genetic disease, or an infection. In some embodiments, said disease is cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

[0088] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0089] **FIG. 1** shows the argonaute phylogenetic tree (1,091 Agos; NCBI marked with in-vitro validated Agos). Of the branch representatives 80 were selected and 8/8 (10%) were validated in vitro. A refined selection of 7 Agos was made, 2 of which (28.5%) were validated in vitro.

[0090] **FIG. 2** shows the argonaute 41/69/70 branch of 13 Agos.

[0091] **FIG. 3** shows the taxonomy information of bacteria of the Ago 41/69/70 branch; this includes NCBI ID number, the organism, and the taxonomy. Each of the thirteen are domain:

bacteria, Phylum: Firmicutes, Class: Clostridia, Order: Clostridiales, Family: Clostridiaceae, and Genus: Clostridium.

[0092] **FIG. 4** shows the host and environmental information of bacteria in the Ago 41/69/70 branch.

[0093] **FIG. 5** shows the representative taxonomy-specificity, including Kingdom, Phylum, Class, Order, Family, Genus, and Species) of the Ago41 branch.

[0094] **FIG. 6** shows the taxonomy-specificity of the Ago41 branch, showing Clostridiaceae family associated Agos are enriched in Ago41 branch.

[0095] **FIG. 7** shows the sequence-specificity for the Ago41 branch, based on a Needleman-Wunsch algorithm for global sequence pairwise comparison.

[0096] **FIG. 8** shows an image of an electrophoresis gel showing a time course of the cleavage of single stranded DNA (ssDNA) by Ago41 with guide DNA (gDNA). Time course ranged from 5-240 minutes. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0097] **FIG. 9** shows an image of an electrophoresis gel showing a time course of the cleavage of single stranded DNA (ssDNA) by Ago69 with guide DNA (gDNA). Time course ranged from 5-240 minutes. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0098] **FIG. 10** shows an image of an electrophoresis gel showing a time course of cleavage of single stranded DNA (ssDNA) by Ago69 with guide DNA (gDNA). Time course ranged from 0-10 minutes. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0099] **FIG. 11** is a graphic depiction showing the effect of temperature on single stranded DNA (ssDNA) template structure (NUPAK), with temperatures of 37°C, 55°C, 65°C, and 75°C.

[0100] **FIG. 12** is a graphic depiction showing the effect of temperature on single stranded DNA (ssDNA) guide structure (NUPAK), with temperatures of 37°C, 55°C, 65°C, and 75°C.

[0101] **FIG. 13** shows an image of an electrophoresis gel showing the single stranded DNA (ssDNA) cleavage by Ago69 at different temperatures with ssDNA guide. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0102] **FIG. 14** shows an image of an electrophoresis gel showing single stranded DNA (ssDNA) cleavage by Ago69 at different temperatures with target (D) and non-target (NT) ssDNA guide. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0103] **FIG. 15A** shows an image of an electrophoresis gel showing single strand DNA (ssDNA) cleavage by Ago69 using different ssDNA guides. "$gDNA^P$" indicates the 5' most nucleotide of

the gDNA is phosphorylated. **FIG. 15B** shows the location of the ssDNA guides relative to ssDNA target sequence and secondary structure.

[0104] **FIG. 16** shows an image of an electrophoresis gel showing single stranded DNA (ssDNA) cleavage by Ago69 after denaturation before ssDNA guide binding. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0105] **FIG. 17** shows an image of an electrophoresis gel showing single stranded DNA (ssDNA) cleavage by Ago69 after denaturation after ssDNA guide binding. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0106] **FIG. 18** shows a sequence comparison of the amino acid sequence of Ago41, Ago69, and Ago70.

[0107] **FIG. 19** shows an image of an electrophoresis gel showing single stranded DNA (ssDNA) cleavage by Ago41, Ago69, and Ago70 with ssDNA guide (D1) and ssRNA guide (R1). "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0108] **FIG. 20** shows an image of an electrophoresis gel showing single stranded DNA (ssDNA) cleavage by Ago69 with guide RNA (gRNA). "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0109] **FIG. 21A** shows an image of an electrophoresis gel showing optimization of NaCl concentration during cleavage by Ago 41 with guide DNA (gDNA). **FIG. 21B** shows an image of an electrophoresis gel showing optimization of NaCl concentration during cleavage by Ago69 with guide DNA (gDNA).

[0110] **FIG. 22A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago02 with guide DNA (gDNA) with different levels of Ago02. The level of Ago02 added to each reaction is 150ng, 300ng, 600ng, 900ng, 1200ng, and 1500ng. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated. **FIG. 22B** shows an image of an electrophoresis gel showing cleavage of single stranded DNA template (90 nucleotides) by Ago02 with guide DNA (gDNA) ranging in length from 13-30 nucleotides. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0111] **FIG. 23A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago02 with guide DNA (gDNA) and a Mg^{2+} titration of 1mM $MgCl_2$, 5mM $MgCl_2$, 10mM $MgCl_2$, and 20mM $MgCl_2$. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated. **FIG. 23B** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago02 with guide

DNA (gDNA) and a Mn^{2+} titration of 1mM $MnCl_2$, 5mM $MnCl_2$, 10mM $MnCl_2$, and 20mM $MnCl_2$. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0112] FIG. 24 shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago02 with guide DNA (gDNA) and a $NaCl_2$ titration of 50mM $NaCl_2$, 125mM $NaCl_2$, 250mM $NaCl_2$, and 500mM $NaCl_2$. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0113] FIG. 25A shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago70 with guide DNA (gDNA) ranging in amount from 150ng – 1500ng. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated. **FIG. 25B** shows an image of an electrophoresis gel showing cleavage of single stranded DNA template (90 nucleotides) by Ago70 with guide DNA (gDNA) ranging in length from 13-30 nucleotides. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0114] FIG. 26A shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago70 with guide DNA (gDNA) and a Mg^{2+} titration of 1mM $MgCl_2$, 5mM $MgCl_2$, 10mM $MgCl_2$, and 20mM $MgCl_2$. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated. **FIG. 26B** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago70 with guide DNA (gDNA) and a Mn^{2+} titration of 1mM $MnCl_2$, 5mM $MnCl_2$, 10mM $MnCl_2$, and 20mM $MnCl_2$. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0115] FIG. 27 shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago70 with guide DNA (gDNA) and a $NaCl_2$ titration of 50mM $NaCl_2$, 125mM $NaCl_2$, 250mM $NaCl_2$, and 500mM $NaCl_2$. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0116] FIG. 28 shows an image of an electrophoresis gel showing the stability of guide RNA (gRNA) during Ago23, Ago29, and Ago51 cleavage. RNase inhibition was mediated by the addition of RNasin as indicated (40U/reaction). For the Ago29 experiments, 125ng of Ago29 was used per reaction. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0117] FIG. 29A shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago23 with guide RNA (gRNA) ranging in amount from 150ng – 1500ng. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated. **FIG. 29B** shows an image of an electrophoresis gel showing cleavage of single stranded DNA

template (90 nucleotides) by Ago23 with guide RNA (gRNA) ranging in length from 13-30 nucleotides. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0118] **FIG. 30A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago23 with guide RNA (gRNA) and a Mg^{2+} titration of 1mM $MgCl_2$, 5mM $MgCl_2$, 10mM $MgCl_2$, and 20mM $MgCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated. **FIG. 30B** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago23 with guide RNA (gRNA) and a Mn^{2+} titration of 1mM $MnCl_2$, 5mM $MnCl_2$, 10mM $MnCl_2$, and 20mM $MnCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0119] **FIG. 31** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago23 with guide RNA (gRNA) and a $NaCl_2$ titration of 50mM $NaCl_2$, 125mM $NaCl_2$, 250mM $NaCl_2$, and 500mM $NaCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0120] **FIG. 32A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago29 with guide RNA (gRNA) ranging in amount from 150ng – 1500ng. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated. **FIG. 32B** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago29 with guide RNA (gRNA) ranging in length from 13-30 nucleotides. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0121] **FIG. 33A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago29 with guide RNA (gRNA) and a Mg^{2+} titration of 1mM $MgCl_2$, 5mM $MgCl_2$, 10mM $MgCl_2$, and 20mM $MgCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated. **FIG. 33B** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago29 with guide RNA (gRNA) and a Mn^{2+} titration of 1mM $MnCl_2$, 5mM $MnCl_2$, 10mM $MnCl_2$, and 20mM $MnCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0122] **FIG. 34** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago29 with guide RNA (gRNA) and a $NaCl_2$ titration of 50mM $NaCl_2$, 125mM $NaCl_2$, 250mM $NaCl_2$, and 500mM $NaCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0123] **FIG. 35A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago51 with guide RNA (gRNA) ranging in amount

from 150ng – 1500ng. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated. **FIG. 35B** shows an image of an electrophoresis gel showing cleavage of single stranded DNA template (90 nucleotides) by Ago51 with guide RNA (gRNA) ranging in length from 13-30 nucleotides. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0124] **FIG. 36A** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago51 with guide RNA (gRNA) and a Mg^{2+} titration of 1mM $MgCl_2$, 5mM $MgCl_2$, 10mM $MgCl_2$, and 20mM $MgCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated. **FIG. 36B** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago51 with guide RNA (gRNA) and a Mn^{2+} titration of 1mM $MnCl_2$, 5mM $MnCl_2$, 10mM $MnCl_2$, and 20mM $MnCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0125] **FIG. 37** shows an image of an electrophoresis gel showing the cleavage of single stranded DNA template (90 nucleotides) by Ago51 with guide RNA (gRNA) and a $NaCl_2$ titration of 50mM $NaCl_2$, 125mM $NaCl_2$, 250mM $NaCl_2$, and 500mM $NaCl_2$. "R1(p)" indicates the 5' most nucleotide of the gRNA is phosphorylated.

[0126] **FIG. 38** shows a schematic of the double strand DNA "bubble" nicking assay. *Bubble template:* ssDNA oligo with complementary regions to assure that no ssDNA is present. *3'overhangs:* RecQ Helicase unwinds substrates with 3'overhangs. *Nt.AlwI site:* positive control. *ssDNA template:* gDNA/cleavage control.

[0127] **FIG. 39** shows an image of an electrophoresis gel showing single stranded DNA (ssDNA) guide dependent nicking of double stranded DNA (dsDNA) bubble template of Ago69. "D^P" indicates the 5' most nucleotide of the gDNA is phosphorylated. "NT^P" indicates the gDNA is a non-target guide DNA (negative control); and the 5' most nucleotide of the gDNA is phosphorylated.

[0128] **FIG. 40** shows an image of an electrophoresis gel showing the effect of GC content of guide DNA (gDNA) on the cleavage activity of Ago69. "D^P" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0129] **FIG. 41** shows an image of an electrophoresis gel showing the effect of GC content of guide DNA (gDNA) on the cleavage activity of Ago02. "D^P" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0130] **FIG. 42** shows an image of an electrophoresis gel showing the effect of GC content of guide DNA (gDNA) on the cleavage activity of Ago41. "D^P" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0131] **FIG. 43** shows an image of an electrophoresis gel showing the effect of GC content of guide DNA (gDNA) on the cleavage activity of Ago70. "D^P" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0132] **FIG. 44** shows the testing impact of SSB proteins on the processivitity of DNA unwinding by RecQ helicase.

[0133] **FIG. 45** shows a line a graph showing the effect of ET-SSB on RecQ mediated DNA unwinding using 3' overhang long substrate.

[0134] **FIG. 46** shows a line a graph showing the effect of ET-SSB on RecQ mediated DNA unwinding using 3' overhang short substrate.

[0135] **FIG. 47** shows a line a graph showing the effect of Eco-SSB on RecQ mediated DNA unwinding using 3' overhang short substrate.

[0136] **FIG. 48** shows an image of an electrophoresis gel showing the elimination of cleavage activity of Ago41 with guide DNA (gDNA) when the DEDX catalytic domain of Ago41 is mutated. Mutations D559A, E595A, and D629A result in an inhibition of Ago41 cleavage activity on gDNA. "D1(p)" indicates the 5' most nucleotide of the gDNA is phosphorylated. "DNT(p)" indicates the gDNA is a non-target guide DNA (negative control); and the 5' most nucleotide of the gDNA is phosphorylated.

[0137] **FIG. 49** shows the amino acid sequence of Ago69 with the comparable mutations in Ago69 to those of the DEDX motif in Ago41 (see FIG. 48).

[0138] **FIG. 50** shows the amino acid sequence of Ago69 with the conserved lysine residues highlighted that are putatively involved in DNA binding specificity are potential sites for mutagenesis.

[0139] **FIG. 51** shows a depiction of the location of the eight guide DNAs (gDNAs) used in the dsDNA cleavage assay described in Example 21. The depiction further includes the GC content and Tm for each gDNA.

[0140] **FIG. 52A** shows a depiction of the location of the eight guide DNAs (gDNAs) and expected cleavage products used in the dsDNA cleavage assay described in Example 21.

FIG. 52B shows an image of an electrophoresis gel showing double stranded DNA (dsDNA) cleavage by Ago69. "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated.

[0141] **FIG. 53A** shows a map of plasmid #56. The M1uI digestion sites are marked with scissors. The expected cleavage products produced by cleavage of plasmid #56 with M1uI are 4487bp and 1827bp fragments. **FIG. 53B** shows a map of plasmid #56. The M1uI digestion sites are marked with scissors; as well as the Ago69 cleavage site. The expected cleavage products produced by cleavage of plasmid #56 with M1uI and Ago69 are 3816bp, 1827bp, and 671bp fragments.

[0142] **FIG. 54** shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 with or without preincubation of plasmid at 75 °C; with and without ET-SSB; and with and without gDNAs 54 and 55. The cleavage was conducted at both 37 °C (left) and 39 °C (right). Stars mark the expected cleavage products.

[0143] **FIG. 55** shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 with or without preincubation of plasmid at 75 °C; with and without ET-SSB; and with and without gDNAs 54 and 55. The cleavage was conducted at both 41.5 °C (left) and 44.9 °C (right).

[0144] **FIG. 56** shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 with or without preincubation of plasmid at 75 °C; with and without ET-SSB; and with and without gDNAs 54 and 55. The cleavage was conducted at both 49.1 °C (left) and 67 °C (right).

[0145] **FIG. 57A** shows a map of plasmid #56. The BsmI digestion sites are marked with scissors. The expected cleavage products produced by cleavage of plasmid #56 with BsmI are 4596bp, 1641bp, and 77bp fragments. **FIG. 57B** shows a map of plasmid #56. The BsmI digestion sites are marked with scissors; as well as the Ago69 cleavage site. The expected cleavage products produced by cleavage of plasmid #56 with BsmI and Ago69 are 4596bp, 1081bp, 552bp, and 77bp fragments.

[0146] **FIG. 58** shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 and M1uI (left) or BsmI (right); with and without ET-SSB; and with gDNA 54 alone, 55 alone, 55 and 54, or no gDNA. The cleavage was conducted at both 41.5 °C (left) and 44.9 °C (right). "$gDNA^P$" indicates the 5' most nucleotide of the gDNA is phosphorylated. Stars indicate the expected cleavage products.

[0147] **FIG. 59** shows an image of a high exposure electrophoresis gel showing cleavage of plasmid #56 by Ago69 and M1uI (left) or BsmI (right); with and without ET-SSB; and with

gDNA 54 alone, 55 alone, 55 and 54, or no gDNA. "gDNA^{P}" indicates the 5' most nucleotide of the gDNA is phosphorylated. Stars indicate the expected cleavage products.

[0148] FIG. 60 shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 and BsmI; with and without ET-SSB; and with or without gDNAs 54 and 55. "gDNA^{P}" indicates the 5' most nucleotide of the gDNA is phosphorylated. Stars indicate the expected cleavage products. "gDNA^{P}" indicates the 5' most nucleotide of the gDNA is phosphorylated. Stars indicate the expected cleavage products.

[0149] FIG. 61 shows an image of a high exposure electrophoresis gel showing cleavage of plasmid #56 by Ago69 and BsmI; with and without ET-SSB; and with or without gDNAs 54 and 55. "gDNA^{P}" indicates the 5' most nucleotide of the gDNA is phosphorylated. Stars indicate the expected cleavage products. "gDNA^{P}" indicates the 5' most nucleotide of the gDNA is phosphorylated. Stars indicate the expected cleavage products.

[0150] FIG. 62 shows a graphical depiction of a protocol of plasmid DNA cleavage assay.

[0151] FIG. 63 shows an image of an electrophoresis gel showing cleavage of plasmid 56 by Ago69 and BsaI-HF. The expected cleavage products produced by cleavage of plasmid #56 with BsaII are 4973bp and 1341bp fragments. C1: preloading of AGO with D54^{P}/D55^{P} + SSB + UvrD + plasmid (standard condition. C2: preloading of AGO with D54P/D55P in presence of SSB + plasmid preincubated with Tte UvrD. C3: preloading of AGO with D54P/D55P in presence of SSB and UvrD + plasmid. C4: preloading of AGO with D54P/D55P + plasmid preincubated with SSB and UvrD. Ctrl: preloading of AGO with no gDNA + SSB + UvrD + plasmid. X: pipetting mistake. Stars indicate the expected cleavage products.

[0152] FIG. 64A shows an image of an electrophoresis gel showing the expression and purification of SSBs, including TneSSB, TthSSB, NeqSSB; and helicases including HEL#100, and EcoRecQ. **FIG. 64B** shows an image of an electrophoresis gel showing the expression and purification of SSBs, including TaqSSB, TmaSSB, SsoSSB, EcoSSB; and helicases including EcoUvrD and TthUvrD.

[0153] FIG. 65 shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 with the indicated SSB and helicase. The left gel is a short exposure. The right gel is a high exposure. 1: Tte UvrD; 2: HEL#65; 3: HEL#71, 4: HEL#78, 5: HEL#92, 6: No helicase, Ctrl1: Plasmid 56 with M1uI-HF, Ctrl2: Plasmid #56 + Ago69 + M1uI-HF. The expected cleavage products of M1uI only: 4487bp and 1827 bp fragments. The expected cleavage products of M1uI + Ago69: 3816bp, 1827bp, and 671 bp fragments.

[0154] **FIG. 66** shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 with the indicated SSB and helicase. The left gel is a short exposure. The right gel is a high exposure. 1: Tte UvrD; 2: HEL#65; 3: HEL#71, 4: HEL#78, 5: HEL#92, 6: No helicase, Ctrl1: Plasmid 56 with M1uI-HF, Ctrl2: Plasmid #56 + Ago69 + M1uI-HF. The expected cleavage products of M1uI only: 4487bp and 1827 bp fragments. The expected cleavage products of M1uI + Ago69: 3816bp, 1827bp, and 671 bp fragments.

[0155] **FIG. 67** shows an image of an electrophoresis gel showing cleavage of plasmid #56 by Ago69 with the indicated SSB and helicase. The left gel is a short exposure. The right gel is a high exposure. 1: Tte UvrD; 2: HEL#65; 3: HEL#71, 4: HEL#78, 5: HEL#92, 6: No helicase, Ctrl1: Plasmid 56 with M1uI-HF, Ctrl2: Plasmid #56 + Ago69 + M1uI-HF. The expected cleavage products of M1uI only: 4487bp and 1827 bp fragments. The expected cleavage products of M1uI + Ago69: 3816bp, 1827bp, and 671 bp fragments.

[0156] **FIG. 68** shows a graphical depiction of Ago69 containing fusion proteins. L: linker; SV40NLS: SV40 nuclear localization signal.

[0157] **FIG. 69A** shows an image of an electrophoresis gel showing expression and purification of the indicated fusion protein. **FIG. 69B** shows an image of an electrophoresis gel showing expression and purification of the indicated fusion protein.

[0158] **FIG. 70** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB, guides D54 and D55, and helicase Tte UvrD were included as indicated. The expected cleavage products were 4604, 1388, and 35bp fragments.

[0159] **FIG. 71** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB, guides D54 and D55, and helicase Tte UvrD were included as indicated. The expected cleavage products were 4604, 1388, and 35bp fragments.

[0160] **FIG. 72A** shows an image of an electrophoresis gel showing expression and purification of the indicated fusion protein. **FIG. 72B** shows western blot of the indicated fusion protein using an anti-6XHis tag antibody for detection of each fusion protein.

[0161] **FIG. 73** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB, guides D54 and D55, and helicase Tte UvrD were included as indicated. The expected cleavage products were 4604, 1388, and 35bp fragments.

[0162] **FIG. 74** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB, guides D54 and D55, and helicase Tte UvrD were included as indicated. The expected cleavage products were 4604, 1388, and 35bp fragments.

[0163] **FIG. 75** shows a graphical depiction of Ago69 and SsoSSB containing fusion proteins.

[0164] **FIG. 76A** shows an image of an electrophoresis gel showing expression and purification of the indicated fusion protein. **FIG. 76B** shows an image of an electrophoresis gel showing expression and purification of the indicated fusion protein.

[0165] **FIG. 77** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB and guide AE1 (gDNA 54 and 55) were included as indicated. The expected cleavage products were 4723bp and 159bp fragments. Cleavage reactions were carried out at 37 °C.

[0166] **FIG. 78** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB and guide AE1 (gDNA 54 and 55) were included as indicated. The expected cleavage products were 4723bp and 159bp fragments. Cleavage reactions were carried out at 37 °C.

[0167] **FIG. 79** shows an image of an electrophoresis gel showing cleavage of plasmid 56 by the indicated Ago69 containing fusion protein. ET-SSB and guide AE1 (gDNA 54 and 55) were included as indicated. The expected cleavage products were 4723bp and 159bp fragments. Cleavage reactions were carried out at 75 °C.

[0168] **FIG. 80** shows a schematic of Ago69 fusion constructs containing two SV40 nuclear localization signals.

[0169] **FIG. 81** shows a series of microscopy images showing nuclear localization of construct AP109.

[0170] **FIG. 82** shows a series of microscopy images showing nuclear localization of construct AP109.

[0171] **FIG. 83** shows a series of microscopy images showing cytosol localization of construct AP110.

[0172] **FIG. 84** shows a series of microscopy images showing nuclear localization of construct SPL0398.

[0173] **FIG. 85** shows a series of microscopy images showing nuclear localization of construct SPL0389.

[0174] **FIG. 86** shows a series of microscopy images showing nuclear localization of construct SPL0390.

[0175] **FIG. 87** shows the GC content of the guide DNAs and cleavage of plasmid 70 or plasmid 56 by Ago69 utilizing the indicated guide DNA.

[0176] **FIG. 88A** shows standard plasmid construct wherein the indicated regions have the indicated GC content. **FIG. 88B** shows a guide swapping construct wherein the indicated regions have the indicated GC content.

[0177] **FIG. 89** shows a schematic of plasmid 56, plasmid 114, and plasmid 115.

[0178] **FIG. 90** shows cleavage of plasmid 56, plasmid 114, and plasmid 115 in the presence or absence of the indicated guide, ETSSB, and ClaI restriction enzyme.

[0179] **FIG. 91** shows cleavage of plasmid 56, plasmid 114, and plasmid 115 in the presence or absence of the indicated guide, ETSSB, and PspOMI restriction enzyme.

[0180] **FIG. 92** is a schematic showing where the indicated DNA guides bind within the HAT region of a HAT plasmid generated according to Example 34.

[0181] **FIG. 93** shows an image of an electrophoresis gel showing cleavage of plasmid 70-HAT by Ago69 with or without ET SSB and with the indicated guide DNA.

[0182] **FIG. 94** shows an image of an electrophoresis gel showing cleavage of plasmid 70-HAT by Ago69 with or without ET SSB and with the indicated guide DNA.

[0183] **FIG. 95** shows an image of an electrophoresis gel showing cleavage of plasmid 70-HAT by Ago69 or the indicated Ago69 homologue (HG2, HG4, HG5) with or without ET SSB and with the indicated guide DNA.

[0184] **FIG. 96A** is a schematic showing sequence identity between Ago69, HG2, and HG4, including the PAZ, MID, and PIWI. **FIG. 96B** is a table showing the percent Percent sequence identity between Ago69, HG2, and HG4.

[0185] **FIG. 97** shows the Ago69 homologues identified, expressed, and purified.

[0186] **FIG. 98A** shows an image of an electrophoresis gel showing purified Ago69 homologues HG1, HG2, HG3, and HG4. **FIG. 98B** shows an image of an electrophoresis gel showing purified Ago69 homologues HG6 and HG7.

[0187] **FIG. 99** shows an image of an electrophoresis gel showing purified Ago69 homologues HG5 and HG9.

[0188] **FIG. 100** shows an image of an electrophoresis gel showing plasmid DNA cleavage by Ago69 homologues HG2, HG4, and HG6.

[0189] **FIG. 101** shows an image of an electrophoresis gel showing plasmid DNA cleavage by Ago69 homologues HG2, HG4, and HG6.

[0190] **FIG. 102** shows a sequence alignment and indicates homology of Ago69, HG2, and HG4.

[0191] **FIG. 103A** shows a first (N terminal) part of a sequence alignment and homology of Ago69, HG2, and HG4 along with an indication of the PAZ, MID, and PIWI domains. **FIG. 103B** shows a second part of a sequence alignment and homology of Ago69, HG2, and HG4 along with an indication of the PAZ, MID, and PIWI domains. **FIG. 103C** shows a third part of a sequence alignment and homology of Ago69, HG2, and HG4 along with an indication of the PAZ, MID, and PIWI domains. **FIG. 103D** shows a fourth (C terminal) part of a sequence alignment and homology of Ago69, HG2, and HG4 along with an indication of the PAZ, MID, and PIWI domains.

[0192] **FIG. 104** shows microscopy image of cells transfected with the SPL0390 construct, indicated guide DNA, and treatment (6-TG or DSMO control).

[0193] **FIG. 105** shows microscopy image of cells transfected with the AP109 contract, indicated guide DNA, and 6-TG.

[0194] **FIG. 106** shows microscopy image of cells transfected with the SPL0398 construct, indicated guide DNA, and 6-TG.

DETAILED DESCRIPTION

[0195] The following description and examples illustrate embodiments of the invention in detail. It is to be understood that this invention is not limited to the particular embodiments described herein and as such can vary. Those of skill in the art will recognize that there are numerous variations and modifications of this invention, which are encompassed within its scope.

DEFINITIONS

[0196] The term "activation" and its grammatical equivalents as used herein refers to a process whereby a cell transitions from a resting state to an active state. This process can comprise a response to an antigen, migration, and/or a phenotypic or genetic change to a functionally active state. For example, the term "activation" can refer to the stepwise process of T cell activation. For example, a T cell can require at least two signals to become fully activated. The first signal can occur after engagement of a TCR by the antigen-MHC complex, and the second signal can occur by engagement of co-stimulatory molecules. Anti-CD3 can mimic the first signal and anti-CD28 can mimic the second signal *in vitro*.

[0197] The term "adjacent" and its grammatical equivalents as used herein refers to right next to the object of reference. For example, the term adjacent in the context of a nucleotide sequence can mean without any nucleotides in between. For instance, polynucleotide A adjacent to polynucleotide B can mean AB without any nucleotides in between A and B.

[0198] The term "Argonaute," "Ago," and its grammatical equivalents as used herein refer to a naturally occurring or engineered domain or protein having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100% sequence identity to a wild type Argonaute polypeptide as measured by protein-protein BLAST algorithm. Some Ago domains or proteins, also referred to herein as "Argonaute nucleases" have endonuclease activity, *e.g.*, the ability to cleave an internal phosphodiester bond in a target nucleic acid.

[0199] A "Clostridia argonaute" or "C-Ago" as used interchangeably herein refers to a naturally occurring or engineered domain or protein having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100% sequence identity to a wild type Argonaute polypeptide derived from a bacterium of the class Clostridia as measured by protein-protein BLAST algorithm.

[0200] The term "autologous" and its grammatical equivalents as used herein refers to as originating from the same being. For example, a sample (*e.g.*, cells) can be removed, processed, and given back to the same subject (*e.g.*, subject) at a later time. An autologous process is distinguished from an allogenic process where the donor and the recipient are different subjects.

[0201] The term "cancer" or "tumor," used interchangeably herein, and their grammatical equivalents as used herein refers to a hyperproliferation of cells whose unique trait—loss of normal controls—results in unregulated growth, lack of differentiation, local tissue invasion, and/or metastasis. With respect to the inventive methods, the cancer can be any cancer, including any of acute lymphocytic cancer, acute myeloid leukemia, alveolar rhabdomyosarcoma, bladder cancer, bone cancer, brain cancer, breast cancer, cancer of the anus, anal canal, rectum, cancer of the eye, cancer of the intrahepatic bile duct, cancer of the joints, cancer of the neck, gallbladder, or pleura, cancer of the nose, nasal cavity, or middle ear, cancer of the oral cavity, cancer of the vulva, chronic lymphocytic leukemia, chronic myeloid cancer, colon cancer, esophageal cancer, cervical cancer, fibrosarcoma, gastrointestinal carcinoid tumor, Hodgkin lymphoma, hypopharynx cancer, kidney cancer, larynx cancer, leukemia, liquid tumors, liver cancer, lung cancer, lymphoma, malignant mesothelioma, mastocytoma, melanoma, multiple myeloma,

nasopharynx cancer, non-Hodgkin lymphoma, ovarian cancer, pancreatic cancer, peritoneum, omentum, and mesentery cancer, pharynx cancer, prostate cancer, rectal cancer, renal cancer, skin cancer, small intestine cancer, soft tissue cancer, solid tumors, stomach cancer, testicular cancer, thyroid cancer, ureter cancer, and/or urinary bladder cancer.

[0202] The term "engineered" and its grammatical equivalents as used herein refers to one or more alterations of a nucleic acid, *e.g.*, the nucleic acid within an organism's genome. The term "engineered" can refer to alterations, additions, and/or deletion of genes. An engineered cell can also refer to a cell with an added, deleted and/or altered gene.

[0203] The term "checkpoint gene" and its grammatical equivalents as used herein refers to any gene that is involved in an inhibitory process (*e.g.*, feedback loop) that acts to regulate the amplitude of an immune response, for example, an immune inhibitory feedback loop that mitigates uncontrolled propagation of harmful responses. These responses can include contributing to a molecular shield that protects against collateral tissue damage that might occur during immune responses to infections and/or maintenance of peripheral self-tolerance. Non-limiting examples of checkpoint genes can include members of the extended CD28 family of receptors and their ligands as well as genes involved in co-inhibitory pathways (*e.g.*, CTLA-4 and PD-1). The term checkpoint gene, in some embodiments, refers to an immune checkpoint gene.

[0204] A "CRISPR," "CRISPR system," or "CRISPR nuclease system" and their grammatical equivalents refer to a system that comprises an RNA molecule (*e.g.*, guide RNA) that binds to DNA and a Cas protein (*e.g.*, Cas9) with nuclease functionality (*e.g.*, two nuclease domains). *See, e.g.*, Sander, J.D., *et al.*, "CRISPR-Cas systems for editing, regulating and targeting genomes," Nature Biotechnology, 32:347–355 (2014); *see also e.g.*, Hsu, P.D., *et al.*, "Development and applications of CRISPR-Cas9 for genome engineering," Cell 157(6):1262-1278 (2014). In some embodiments, a CRISPR system includes a Cas protein with nickase functionality (*e.g.*, one catalytically dead nuclease domain and one catalytically active nuclease domain). A Cas can be partially catalytically dead.

[0205] The term "disrupting" and its grammatical equivalents as used herein refers to a process of altering a gene, *e.g.*, by deletion, insertion, mutation, rearrangement, or any combination thereof. For example, a gene can be disrupted by knockout. Disrupting a gene can, for example, partially or completely suppress expression of the gene. Disrupting a gene can also cause activation of a different gene, for example, a downstream gene.

[0206] The term "function" and its grammatical equivalents as used herein refers to the capability of operating, having, or serving an intended purpose. Functional can comprise any percent from baseline to 100% of normal function. For example, functional can comprise or comprise about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50,55, 60, 65, 70, 75, 80, 85, 90, 95, and/or 100% of normal function. In some cases, the term functional can mean over or over about 100% of normal function, for example, 125, 150, 175, 200, 250, 300% and/or above normal function.

[0207] The term "gene editing," "genome editing," and their grammatical equivalents as used herein refers to genetic engineering in which one or more nucleotides are inserted, replaced, or removed from a genome. Gene editing can be performed using a nuclease (*e.g.*, a natural-existing nuclease or an artificially engineered nuclease).

[0208] The term "mutation" and its grammatical equivalents as used herein include the substitution, deletion, and insertion of at least one nucleotide in a polynucleotide. For example, up to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 40, 50, or more nucleotides/amino acids in a polynucleotide (cDNA, gene) or a polypeptide sequence can be substituted, deleted, and/or inserted. A mutation can affect the coding sequence of a gene or its regulatory sequence. A mutation can also affect the structure of the genomic sequence or the structure/stability of the encoded mRNA.

[0209] The term "non-human animal" and its grammatical equivalents as used herein includes all animal species other than humans, including non-human mammals, which can be a native animal or a genetically modified non-human animal.

[0210] The terms "nucleic acid," "polynucleotide," "polynucleic acid," and "oligonucleotide" and their grammatical equivalents are used interchangeably herein and refer to a deoxyribonucleotide or ribonucleotide polymer, in linear or circular conformation, and in either single- or double-stranded form. For the purposes of the present disclosure, these terms should not to be construed as limiting with respect to length, unless the context clearly indicates otherwise. The terms can also encompass analogues of natural nucleotides, as well as nucleotides that are modified in the base, sugar and/or phosphate moieties (*e.g.*, phosphorothioate backbones). Modifications of the terms can also encompass demethylation, addition of CpG methylation, removal of bacterial methylation, and/or addition of mammalian methylation. In general, an analogue of a particular nucleotide can have the same base-pairing specificity, *e.g.*, an analogue of A can base-pair with T.

[0211] The term "construct" refers to an artificial or synthetic construct. For example, a polypeptide construct can refer to an artificial or synthetic polypeptide, e.g., comprising one or more polypeptide sequences. Similarly, a nucleic acid construct can refer to an artificial or synthetic nucleic acid, e.g., comprising one or more nucleic acid sequences.

[0212] The term "percent (%) identity" can be readily determined for nucleic acid or amino acid sequences, over the full-length of a sequence, or a fragment thereof. Generally, when referring to "identity", "homology", or "similarity" between two different sequences (e.g., nucleotide or amino acid sequences), "identity", "homology" or "similarity" is determined in reference to "aligned" sequences. "Aligned" sequences or "alignments" refer to multiple nucleic acid sequences or protein (amino acids) sequences, often containing corrections for missing or additional bases or amino acids as compared to a reference sequence.

[0213] The term "phenotype" and its grammatical equivalents as used herein refer to a composite of an organism's observable characteristics or traits, such as its morphology, development, biochemical or physiological properties, phenology, behavior, and/or products of behavior. Depending on the context, the term "phenotype" can sometimes refer to a composite of a population's observable characteristics or traits.

[0214] "Polypeptide," "peptide," and their grammatical equivalents as used herein refer to a polymer of amino acid residues. A "mature protein" is a protein which is full-length and which, optionally, includes glycosylation or other modifications typical for the protein in a given cellular environment. Polypeptides and proteins disclosed herein (including functional portions and functional variants thereof) can comprise synthetic amino acids in place of one or more naturally-occurring amino acids. Such synthetic amino acids are known in the art, and include, for example, aminocyclohexane carboxylic acid, norleucine, α-amino n-decanoic acid, homoserine, S-acetylaminomethyl-cysteine, trans-3-and trans-4-hydroxyproline, 4-aminophenylalanine, 4-nitrophenylalanine, 4-chlorophenylalanine, 4-carboxyphenylalanine, β-phenylserine β-hydroxyphenylalanine, phenylglycine, α-naphthylalanine, cyclohexylalanine, cyclohexylglycine, indoline-2-carboxylic acid, 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, aminomalonic acid, aminomalonic acid monoamide, N'-benzyl-N'-methyl-lysine, N',N'-dibenzyl-lysine, 6-hydroxylysine, ornithine, α-aminocyclopentane carboxylic acid, α-aminocyclohexane carboxylic acid, α-aminocycloheptane carboxylic acid, α-(2-amino-2-norbornane)-carboxylic acid, α,γ-diaminobutyric acid, α,β-diaminopropionic acid, homophenylalanine, and α-tert-butylglycine. The present disclosure further contemplates that expression of polypeptides described herein in

an engineered cell can be associated with post-translational modifications of one or more amino acids of the polypeptide constructs. Non-limiting examples of post-translational modifications include phosphorylation, acylation including acetylation and formylation, glycosylation (including N-linked and O-linked), amidation, hydroxylation, alkylation including methylation and ethylation, ubiquitination, addition of pyrrolidone carboxylic acid, formation of disulfide bridges, sulfation, myristoylation, palmitoylation, isoprenylation, farnesylation, geranylation, glypiation, lipoylation and iodination. The term polypeptide includes a polypeptide that has been separated from components that naturally accompany it. Typically, the polypeptide is isolated when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. In some embodiments, the preparation is at least 75%, at least 90%, or at least 99%, by weight, a polypeptide. An isolated polypeptide may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

[0215] The term "protospacer" and its grammatical equivalents as used herein refers to a PAM-adjacent nucleic acid sequence capable to hybridizing to a portion of a guide RNA, such as the spacer sequence or engineered targeting portion of the guide RNA. A protospacer can be a nucleotide sequence within gene, genome, or chromosome that is targeted by a guide RNA. In the native state, a protospacer is adjacent to a PAM (protospacer adjacent motif). The site of cleavage by an RNA-guided nuclease is within a protospacer sequence. For example, when a guide RNA targets a specific protospacer, the Cas protein will generate a double strand break within the protospacer sequence, thereby cleaving the protospacer. Following cleavage, disruption of the protospacer can result though non-homologous end joining (NHEJ) or homology-directed repair (HDR). Disruption of the protospacer can result in the deletion of the protospacer. Additionally or alternatively, disruption of the protospacer can result in an exogenous nucleic acid sequence being inserted into or replacing the protospacer.

[0216] The term "recipient" and their grammatical equivalents as used herein refers to a human or non-human animal. The recipient can also be in need thereof.

[0217] The term "recombination" and its grammatical equivalents as used herein refers to a process of exchange of genetic information between two polynucleic acids. For the purposes of this disclosure, "homologous recombination" or "HR" can refer to a specialized form of such

genetic exchange that can take place, for example, during repair of double-strand breaks. This process can require nucleotide sequence homology, for example, using a donor molecule to template repair of a target molecule (*e.g.*, a molecule that experienced the double-strand break), and is sometimes known as non-crossover gene conversion or short tract gene conversion. Such transfer can also involve mismatch correction of heteroduplex DNA that forms between the broken target and the donor, and/or synthesis-dependent strand annealing, in which the donor can be used to resynthesize genetic information that can become part of the target, and/or related processes. Such specialized HR can often result in an alteration of the sequence of the target molecule such that part or all of the sequence of the donor polynucleotide can be incorporated into the target polynucleotide. In some cases, the terms "recombination arms" and "homology arms" can be used interchangeably.

[0218] The term "transgene" and its grammatical equivalents as used herein refer to a gene or genetic material that is transferred into an organism. For example, a transgene can be a stretch or segment of DNA containing a gene that is introduced into an organism. When a transgene is transferred into an organism, the organism is then referred to as a transgenic organism. A transgene can retain its ability to produce RNA or polypeptides (*e.g.*, proteins) in a transgenic organism. A transgene can be composed of different nucleic acids, for example RNA or DNA. A transgene can encode for an engineered T cell receptor, for example a TCR transgene. A transgene can be a TCR sequence. A transgene can be a receptor. A transgene can comprise recombination arms. A transgene can comprise engineered sites.

[0219] A "therapeutic effect" occurs if there is a change in the condition being treated. The change can be positive or negative. For example, a 'positive effect' can correspond to an increase in the number of activated T-cells in a subject. In another example, a 'negative effect' can correspond to a decrease in the amount or size of a tumor in a subject. There is a "change" in the condition being treated if there is at least 10% improvement, preferably at least 25%, more preferably at least 50%, even more preferably at least 75%, and most preferably 100%. The change can be based on improvements in the severity of the treated condition in an individual, or on a difference in the frequency of improved conditions in populations of individuals with and without treatment with the therapeutic compositions with which the compositions of the present invention are administered in combination. Similarly, a method of the present disclosure can comprise administering to a subject an amount of cells that is "therapeutically effective." The

term "therapeutically effective" should be understood to have a definition corresponding to 'having a therapeutic effect.'

[0220] The term "sequence" and its grammatical equivalents as used herein refers to a nucleotide sequence, which can be DNA or RNA; can be linear, circular or branched; and can be either single-stranded or double stranded. A sequence can be mutated. A sequence can be of any length, for example, between 2 and 1,000,000 or more nucleotides in length (or any integer value there between or there above), *e.g.*, between about 100 and about 10,000 nucleotides or between about 200 and about 500 nucleotides.

OVERVIEW

[0221] The present disclosure provides methods, systems, compositions, and kits for modifying a target polynucleic acid using a system comprising an Argonaute (Ago) polypeptide. The present disclosure also provides methods of treating a disease or disorder using the herein described systems, compositions, or kits. In some embodiments, the systems described herein comprise, for example, a nuclease and a helicase. These systems overcome technical challenges associated with argonaute proteins including, for example, a lack of activity at temperatures that are conducive for gene editing in human cells. The methods, systems, compositions and kits described herein allow for this physiologically-relevant gene editing by providing an argonaute system from a bacterium. In some embodiments, the argonaute is a mesophilic argonaute or a mesothermic argonaute. Without wishing to be bound by theory, such systems are able to induce single- or double-stranded polynucleic acid breaks at physiological temperatures. In some embodiments, the herein described systems comprise a fragment of a mesophilic Ago polypeptide gene or protein. In some embodiments, the system comprises one or more associated genes. In some embodiments, the one or more associated genes are found in proximity to the argonaute gene in its genome of origin. In some embodiments, a herein described Ago polypeptide and a protein encoded by an associated gene are provided as a fusion protein.

I. Argonaute Proteins

[0222] Provided herein are a gene editing systems comprising Ago polypeptides, or functional fragments or functional variants thereof. Provided herein are also compositions, constructs, systems, and methods for disrupting a genomic sequence in a subject (*e.g.* mammal, non-mammal, or plant). Also provided herein are compositions, constructs, systems, and methods of treating or inhibiting a condition caused by a defect in a target sequence in a genomic locus of

interest in a subject (*e.g.*, mammal or human) or a non-human subject (*e.g.*, mammal) in need thereof. In some embodiments, a method comprises modifying a subject or a non-human subject by manipulation of a target sequence and wherein a condition is susceptible to treatment or inhibition by manipulation of a target sequence.

[0223] Disclosed herein is a system comprising a Clostridia Argonaute (Ago) polypeptide, or a polynucleic acid encoding the same, and an exogenous guiding polynucleic acid. The Ago polypeptide is a prokaryotic Ago (p-Ago) polypeptide. In some embodiments the Ago polypeptide comprises an amino acid sequence having at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 1-10 or 134-136 as measured by protein-protein BLAST algorithm. In some cases, the system comprises an Ago polypeptide. In some cases, the system comprises a polynucleic acid encoding the Ago polypeptide. In some embodiments, the polynucleic acid encoding the Ago polypeptide comprises a nucleic acid sequence having at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 11-21 or 137-139 as measured by nucleotide-nucleotide BLAST algorithm.

[0224] In one aspect, disclosed herein is a system comprising (a) an Ago polypeptide, or a polynucleic acid encoding the same; and (b) an exogenous guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence. In one aspect, disclosed herein is a system comprising (a) an Ago polypeptide, or a polynucleic acid encoding the same, wherein said Ago polypeptide comprises an amino acid sequence having at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 1-10 or 134-136 as measured by protein-protein BLAST algorithm; and (b) an exogenous guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence. In another aspect, disclosed herein is a system comprising (a) an Ago polypeptide, or a polynucleic acid encoding the same, wherein said Ago polypeptide is a mesophilic Ago; and (b) an exogenous guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence.

[0225] Examples of an Ago include, but are not limited to, Ago polypeptides comprising an amino acid sequence having 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100% identity with one of SEQ ID NOs: 1-10 or 134-136. Percent sequence identity can be determined by BLAST (basic local alignment search tool) algorithm, specifically protein-protein BLAST

(BLASTP). BLAST is provided by National Center for Biotechnology Information (NCBI) for aligning query sequences against those present in databases. The parameters of BLASTP can be set as Matrix BLOSUM62, Gap Costs Existence:11, Extension:1, and Compositional Adjustments Conditional Compositional Score Matrix Adjustment, with applying any filters or masks. In some embodiments, alignment is determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is disclosed in Smith & Waterman (1981) Adv. Appl. Math. 2: 482-489.

[0226] In some cases, the Ago may be an argonaute polypeptide or a protein with sequence similarity to a known Argonaute. Examples of known Argonautes include, but are not limited to, Clostridia Agos. In some cases, the Ago may be an argonaute polypeptide or a protein with at least 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100% positive scoring matches relative to a known Argonaute.

[0227] In some cases, the Ago comprises one or more domains or motifs commonly found in argonaute polypeptide. In some cases, the Ago comprises a PAZ domain. In some cases, the Ago lacks a PAZ domain. In some cases, the Ago comprises a domain with sequence similarity to a PAZ domain. In some cases, the Ago comprises a Sir2 domain. In some cases, the Ago comprises a Sir2-like domain. In some cases, the Ago comprises an additional Sir2 or Sir2-like domain. In some cases, the Ago comprises a Sir2 and a Sir2-like domain. In some cases, the Ago lacks a Sir2 domain. In some cases, the Sir2 domain is an N-terminus Sir2 domain. In some cases, the Sir2-like domain is an N-terminus Sir2-like domain. In some cases, the Ago lacks a Sir2-like domain. In some cases, the Ago comprises a functional DEDX motif. In other cases, the Ago lacks a functional DEDX motif. A DEDX motif is a catalytic tetrad in the PIWI domain, wherein the "X" can vary. In some embodiments, a polypeptide as described herein comprises an RNAse H-like domain with a DEDX motif, or a functional variant thereof. In some cases, the Ago comprises a PIWI domain. In other cases, the Ago lacks a PIWI domain. In some cases, the Ago

comprises a PIWI-like domain. In other cases, the Ago lacks a PIWI-like domain. In some cases, the PIWI domain or the PIWI-like domain is at a C-terminus of the Ago.

[0228] In some embodiments, the Ago described herein, or a fragment thereof, is a polypeptide or a protein with nucleic acid-cleaving activity. In some embodiments, he protein or polypeptide with nucleic acid-cleaving activity (*e.g.*, a nuclease) is an enzyme (i.e., enzymatic protein or polypeptide) that cleaves a chain of nucleotides in a nucleic acid into smaller units. In some embodiments, the protein or polypeptide with nucleic acid-cleaving activity is from a eukaryote or a prokaryote. In some embodiments, the protein or polypeptide with nucleic acid-cleaving activity is from a eukaryote. In some embodiments, the protein or polypeptide with nucleic acid-cleaving activity is from a prokaryote. In some embodiments, the protein or polypeptide with nucleic acid-cleaving activity is from archaea. In some embodiments, the protein or polypeptide with nucleic acid-cleaving activity is from bacteria. In some embodiments, a nuclease is a protein that is located in proximity to the Ago gene in a microbiome genome.

[0229] In some embodiments, the enzymatic polypeptide is an RNA-dependent DNase editor, an RNA-dependent RNase editor, a DNA-dependent DNase editor, or a DNA-dependent RNase editor. Examples of an RNA-dependent DNase editor are Cas9 and Cpf1 to name a couple. An example of an RNA-dependent RNase editor is Cas13. An enzymatic protein can contain multiple domains. For example, in some embodiments, an enzymatic polypeptide contains domains that can bind to a duplex of DNA-RNA, DNA-DNA, or RNA-RNA. For example, RuvC can bind Cas9 and Cpf1; HNH can bind Cas9, RNase-H can bind ribonuclease, and PIWI can bind Ago.

[0230] In some cases, the nuclease activity is double stranded polynucleic acid cleaving activity. In some cases, nuclease activity is single stranded polynucleic acid cleaving activity. In some cases, the Ago polypeptide or Ago polypeptide fragment has nickase activity. In some embodiments, the Nickase activity is single stranded DNA or RNA cleaving activity. In some cases, the Ago polypeptide or Ago polypeptide fragment has RNase activity. In some cases, RNase activity is double stranded RNA cleaving activity. In some cases, RNase activity is RNA cleaving activity. In some cases, the Ago polypeptide or Ago polypeptide fragment or polypeptide has RNase-H activity. In some cases, RNase-H activity is RNA cleaving activity. In some cases, the Ago polypeptide or Ago polypeptide fragment has recombinase activity. In some embodiments, the Ago polypeptide or Ago polypeptide fragment also has DNA-flipping activity. In some cases, the Ago polypeptide or Ago polypeptide fragment has transposase activity.

[0231] In some cases, the Ago polypeptide or Ago polypeptide fragment demonstrates nucleic acid-cleaving activity in a range of temperatures of from about 19 °C to about 41 °C. In some cases, the Ago polypeptide or Ago polypeptide fragment has nucleic acid-cleaving activity at temperatures of about 17 °C, about 18 °C, 19 °C, about 20 °C, about 21 °C, about 22 °C, about 23 °C, about 24 °C, about 25 °C, about 26 °C, about 27 °C, about 28 °C, about 29 °C, about 30 °C, about 31 °C, about 32 °C, about 33 °C, about 34 °C, about 35 °C, about 36 °C, about 37 °C, about 38 °C, about 39 °C, or up to 40 °C. In some embodiments, the Ago polypeptide or Ago polypeptide fragment has nucleic acid-cleaving activity at temperatures from about 17 °C to 40 °C. In some cases, the Ago polypeptide or Ago polypeptide fragment has nucleic acid-cleaving activity at temperatures of about 37 °C. In some cases, a mesophilic Ago can be active at temperatures of at least about 17 °C. In some cases, when the Ago polypeptide is a mesophilic. In some cases, the Ago polypeptide is derived from a mesophilic Clostridia bacterium.

[0232] In some cases, the Ago polypeptide is expressed by a gene located adjacent to an operon of at least one of DNA replication, recombination or repair gene. In some cases, the Ago polypeptide is expressed by a gene located adjacent to an operon of at least one of a defense mechanism related gene, or a transcription related gene. In some cases, the Ago polypeptide is derived from a polypeptide encoded by a gene located in an adjacent operon to at least one of a P-element induced Wimpy testis (PIWI) gene, RuvC, Cas, Sir2, Mrr, TIR, PLD, REase, restriction endonuclease, DExD/H, superfamily II helicase, RRXRR, DUF460, DUF3010, DUF429, DUF1092, COG5558, OrfB_IS605, Peptidase_A17, Ribonuclease H-like domain, 3'-5' exonuclease domain, 3'-5' exoribonuclease Rv2179c-like domain, Bacteriophage Mu, transposase, DNA-directed DNA polymerase, family B, exonuclease domain, Exonuclease, RNase T/DNA polymerase III, yqgF gene, HEPN, RNase LS domain, LsoA catalytic domain, KEN domain, RNaseL, Irel, RNase domain, RloC, or PrrC. In some cases, the Ago polypeptide is derived from a polypeptide encoded by a gene located in an adjacent operon to at least one of a gene involved in defense, stress response, a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), or DNA repair.

[0233] In some cases, the Ago polypeptide or Ago polypeptide fragment is chosen based on proximity to a secondary gene in a genome. For example, in some embodiments, the Ago polypeptide or Ago polypeptide fragment is chosen based on proximity to DNA repair associated genes. In some cases, the Ago polypeptide or Ago polypeptide fragment is chosen based on a predicted alignment (*e.g.*, structural analysis) or phylogenetic analysis. For example, the Ago

polypeptide or Ago polypeptide fragment may have homology or be conserved in relation to a gene sequence of a secondary gene. In some embodiments, conservation refers to a sequence or structure. In some embodiments, the structural conservation refers to the presence or absence of structural features. A structural feature can be a secondary structural feature such as an alpha helix or beta pleated sheet. An Ago polypeptide can be screened or chosen based on a secondary structure.

[0234] In some cases, the Ago polypeptide or portion thereof is a naturally-occurring Ago polypeptide (*e.g.*, naturally occurs in a Clostridia bacterial cell). In other cases, the Ago polypeptide may not be a naturally-occurring polypeptide (*e.g.*, the Ago polypeptide is a variant, chimeric, or fusion). In some cases, the Ago polypeptide has nuclease activity. In some cases, the Ago polypeptide may not have nuclease activity.

[0235] In some cases, the Ago is a type I prokaryotic Argonaute. In some cases, a type I prokaryotic Argonaute carries a DNA nucleic acid-targeting nucleic acid. In some cases, a DNA nucleic acid-targeting nucleic acid targets one strand of a double stranded DNA (dsDNA) to produce a nick or a break of the dsDNA. In some embodiments, a nick or break triggers host DNA repair. In some cases, a host DNA repair is nonhomologous end joining (NHEJ) or homologous directed recombination (HDR). In some cases, a dsDNA is selected from a genome, a chromosome, and a plasmid. In some embodiments, a type I prokaryotic Argonaute is a long type I prokaryotic Argonaute, which may possess an N-PAZ-MID-PIWI domain architecture. In some cases, a long type I prokaryotic Argonaute possesses a catalytically active PIWI domain. In some embodiments, the long type I prokaryotic Argonaute possesses a catalytic tetrad encoded by aspartate-glutamate-aspartate- aspartate/histidine (DEDX). In some embodiments, a DEDX motif is mutated at any of the positions, which can suppress catalytic activity. In some embodiments, the catalytic tetrad can bind one or more magnesium ions or manganese ions. In some cases, the type I prokaryotic Argonaute anchors the 5' phosphate end of a DNA guide. In some cases, a DNA guide has a deoxy-cytosine at its 5' end.

[0236] In some embodiments, the Ago is a type II Ago, for instance a type II prokaryotic Argonaute A type II prokaryotic Argonaute carries an RNA nucleic acid-targeting nucleic acid. In some embodiments, an RNA nucleic acid-targeting nucleic acid targets one strand of a double stranded DNA (dsDNA) to produce a nick or a break of the dsDNA which may trigger host DNA repair; the host DNA repair can be non-homologous end joining (NHEJ) or homologous directed recombination (HDR). In some cases, a dsDNA is selected from a genome, a chromosome and a

plasmid. A type II prokaryotic Argonaute may be a long type II prokaryotic Argonaute or a short type II prokaryotic Argonaute. A long type II prokaryotic Argonaute may have an N- PAZ-MID-PIWI domain architecture. A short type II prokaryotic Argonaute may have a MID and PrWI domain, but may not have a PAZ domain. In some cases, a short type II Ago has an analog of a PAZ domain. In some cases a type II Ago may not have a catalytically active PIWI domain. A type II Ago may lack a catalytic tetrad encoded by aspartate- glutamate-aspartate-aspartate/histidine (DEDX). In some cases, a gene encoding a type II prokaryotic Argonaute clusters with one or more genes encoding a nuclease, a helicase or a combination thereof. A nuclease may be natural, designed or a domain thereof. In some cases, the nuclease is selected from a Sir2, RE1 and TIR. The type II Ago may anchor the 5' phosphate end of an RNA guide. In some cases, the RNA guide has a uracil at its 5' end. In some cases, the type II prokaryotic Argonaute is a Rhodobacter sphaeroides Argonaute. In some cases, it may be desirable to use an Argonaute nuclease that has lost its ability to cleave a nucleic acid, such as in applications where the Argonaute: guide molecule complex is used as a probe. In some cases, a dead Argonaute system may utilize secondary nucleases to perform a genomic disruption. In such cases, one or more of the amino acid residues in a catalytic domain are substituted or deleted, such that catalytic activity is abolished, or diminished. In other cases, using a cleavage temperature-inducible Argonaute may be desired to control the timing of cleavage, or if cleavage should be inhibited at non-inducible temperatures.

[0237] In some cases, the Ago has at least one active domain. For example, in some embodiments, the Ago's active domain is a PIWI domain. In some embodiments, in addition to a catalytic PIWI domain the Ago contains non-catalytic domains such as PAZ (PIWI-Argonaute-Zwille), MID (Middle) and N domain, along with two domain linkers, L1 and L2. A MID domain can be utilized for binding the 5′-end of a guiding polynucleic acid and can be present in an Ago protein. A PAZ domain can contain an OB-fold core. An OB-fold core can be involved in stabilizing a guiding polynucleic acid from a 3'end. An N domain may contribute to a dissociation of the second, passenger strand of a loaded double stranded genome and to a target cleavage. In some cases, an Argonaute family may contain PIWI and MID domains. In some cases, an Argonaute family may or may not contain PAZ and N domains.

[0238] In some cases, the Ago is or comprises a naturally-occurring polypeptide (*e.g.*, naturally occurs in Clostridia bacterial cell), such as a nuclease. In other cases, the Ago is or comprises a non-naturally-occurring polypeptide. A non-naturally occurring polypeptide can be engineered.

In some embodiments, an engineered Ago polypeptide is a chimeric nuclease, mutated, conjugated, or otherwise modified version thereof. In some cases, the Ago comprises a sequence encoded by any one of the sequences of Table 1 (SEQ ID NOs: 1-10), modified versions thereof, derivatives thereof, or truncations thereof. In some cases, the Ago polypeptide or portion thereof comprises a percent identity to any one of SEQ ID NOs: 1-10 from at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100%. In some embodiments, the Ago comprises an amino acid sequence 100% identical to SEQ ID NO: 1. In some embodiments, the Ago comprises an amino acid sequence 100% identical to SEQ ID NO: 1, except there is a non-lysine amino acid residue at one or more of (e.g., 1, 2, 3, 4, or 5) positions 479, 522, 563, 581, 642 of SEQ ID NO: 1. In some embodiments, the Ago has a mutation in one or more residue of the DEDX domain. In some embodiments, these one or more mutations reduce catalytic activity of the Ago as compared to a corresponding Ago without the one or more mutations.

Table 1: Mesophilic Ago Amino Acid Sequences

SEQ ID NO	Ago/Genus Species	Amino Acid Sequence
1	69 (Referred to herein as Ago69 or Argonaute 69) (Clostridium perfringens WAL-14572; Accession No. NZ_JH594533.1) Underlined: PIWI Domain	MVGGYKVSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKII KSARYKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVK SEVLSIEDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRINYSKE ILSNTLLTRELKDEFKKSNKGFNLKRKFRISPVVNKMGKVILYL SCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVIESVL ETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRG NKKIYMYIPHALKPIITREYLAKNDPEFSKEIEQLIKMNMNYRY ETLKSFVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGKIDEPVL MGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAI RAIYDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGDI TEYKKAALKLNNYNNVDFVIAIVPNMSDEEIENSYNPFKKIWAE LNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGILGKIGGIPW VVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDKYGKLINYYK PNIPQNGEKINTEILQEIFDKVLISYEEENGAYPKNIVIHRDGF SREDLDWYENYFGKKNIKFNIIEVKKSTPLKIASINEGNITNPE KGSYILRGNKAYMVTTDIKENLGSPKPLKIEKSYGDIDMLTALS QIYALTQIHVGATKSLRLPITTGYADKICKAIEFIPQGRVDNRL FFL
2	41 *Clostridium disporicum*	MNNLTLEAFRGIGTIKPLLFYRYKLIGKGKIENTYKTIRNAQNR MSFNNKFKATFSKDEIIYTLEKFEIIPTLDDVTIIFDGEEVLPI KDNNKIYSEVIEFYINNNLRNVKFNYKYPKYRAANTREITGNVI LDKDMNEKYKKSNKGFELKRKFIISPKVDDEGKVTLFLDLNASF DYDKNIYQMIKAGIDVVGEEVINIWSNKKQRGKIKEISDIKINE PCNFGQSLIDYYISSNQASRVNGFTEEEKNTNVIIVESGKSRLS

		YIPHALKPIITREYIAKNDEVFSKEIEGLIKINMNYRYEILKRF VSDIGTIKELNNLRFEKIYMDNIESLGYEQGQLKDPVLIGGKGI LKDKIHVFKSGFYKSPNDEIKFGVIYPRGYIKDTQSVIRAIYDF CTEGKYQGKDNIFINNKLMNIKFSNKECVFEEYELNDITEYKRA ANKLKKNENIKFVIAIIPTINESDIENPYNPFKRVCAEINLPSQ MISLKTAKRFSTSRGQSELYFLHNISLGILGKIGGVPWVIKDMP GEVDCFVGLDVGTKEKGIHYPACSVLFDKYGKLINYYKPTIPQS GEIIKTDVLQEIFDKVLLSYEEENGQYPRNIVIHRDGFSREDLE WYKNYFLKKNIEFSIVEVRKNFATRLVNNFNDEVSNPSKGSFIL RDNEAIVVTTDINDNMGAPKPIKVEKTYGDIDMLTIINQIYALT QIHVGSAKSLRLPITTGYADKICKAIDYIPSGQVDNRLFFL
3	70 *Clostridium saudiense*	MNNLMLEAFKGIGTIKPLVFYRYKLIGKGKIENTYKTISNAKNK MSFNNKFKATFSKGETIYTLEKFEVMPNLNDVTIEFDGEEVLPI KDNNEIYSEVVQFYINNNLRKIKLDNKYQKYRATNTREITGNVI LDKDFKEKYKKSKSGFQLKRKFIISPKVNDEGKVTLFLDLNSSF DYDKNIYQMIKAGMDVVGQEVINTWNNKKQKGKIKKISELTISE PCNFGQSLIDYYVSLNQAVRVKNFTEEEKNTNVIVVQVGKGEVE YIPHALKPIITREYIKKYDEAFSKEVENLIKINMSYRYEILKKF IDDIGSITELNNLKFENTYIDNIESLGYQQGKLNDPVLIGGKGI LKDKIHVFKSGFYKSPIDEVKFGVIYPKGHTNDSKSTIRAIYDF CTDGKYQGKDNIFINNKLMNIKFSNQDCVFEEYELNDITEYKRA ANKLKNNENIKFVIAIIPAIDESDIENPYNPFKRVCAELNLPSQ MVSLKTAKRFGTSKGNNELYFLHNISLGILGKIGGVPWVIKDMP GEVDCFVGLDVGTKEKGIHYPACSVLFDKYGKLINYYKPTIPQS GEIIKTDVLQEIFDKVLLSYEEENGQYPRNIVIHRDGFSREDLE WYKNYFIKKNINFTIVEIKKNFATRVANNINNEVSNPFKGSFIL RENEAIVVTTDIKDNIGAPKPIKVEKTYGDIDMMTIINQIYALT QIHVGSAKSMRLPITTGYADKICKSIEYIPSGRVDNRLFFL
4	51 *Rhodopirellula maiorica*	MLQLNGFSIEIAGGSLTVLKSKIAPTDVKETRRSLEDDWFTMYH EGHLYSLAKNSNASGGLGETELLVLSDHLGLRFVKAMLDQAMRG VFEAYDPVRDRPFTFLARNVDLVALAAENLESKPSLLSKFEIRP KYELEAKVVEFRPGELELMLALNLTTRWICNASVDELIEKNIPV RGMHLIRRNREPGQRSLVGTFDRMEGDNALLQDAYDGQDKIAAS QVRIEGSKEVFATSLRRLLGNRYTSFMHSVDNEYGKLCGGLGFD GELRKMQGFLAKKSPIQLHGGVEVSVGQRVQLTNQPGYKTTVEL LQSKYCFDRSRTKLHPYAWDGLARFGPFDRGSFPTRSPRILLVT PDSASGKVSQALKKFRDGFGSSQSSMYDGFLDTFHLSNAPFFPL PVKLDGVQRSDVGKAYRKAIEDKLARDDDFDAAFNILLDEHANL PDSHNPYLVAKSILLSHGIPVQEARVSTLTANEYSLQHTFRNVA TALYAKMGGVPWTVDHGETVDDELVVGIGNAELSGSRFEKRQRH IGITTVFRGDGNYLLSNLSKECRYEDYPDVLRESTIAVLREVKQ RNNWLPGQTVRIVFHAFKPLKNVEIADIIASSVKEVGSEQTIEF AFLNVSLDHSFTLLDMAQRGITKKNQTKGIYVPRRGMTVQVGRY TRLVTSIGPHMVKRANLALPRPLLIHLHKQSTYRDLSYLSEQVL NFTTLSWRSTLPSEKPVTILYSSLIADLLGRLKSVDDWSPAVLN TKLRNSKWFL
5	02	MNTPLTHYVLTEWESDTNTNVLHIHLYTLPVRNVFEQHKENGNA CFDLRKLNRSLIIDFYDQYIVSWQPIENWGEYTFTQHEYRSINP TILAERAILERLLLRTIESVQPKKEIAAGSRKFTWLKAEKVVEN

	Paenibacillus odorifer	ISIHRVIQCDVTVDYAGKISVGFDLNHSYRTNESVYDLMKSNAI FKGDRVIDIYNNLHYEFVEISNSTINDSIPELNQSVVNYFTKER KQAWKVDKLEQSMPVVYLKAFNGSRIAYAPAMLQKELTFESLPT NVVRQTSEIFKQNANQKIKTLLDEIQKILARTDKIKFNKQKLLV QQAGYEILELSNPNLQFGKNVTQTQLKYGLDKGGVVASKPLSIN LLVYPELIDTKLDVINDFNDKLNALSHKWGVPLSILKKSGAYRN RPIDFTNPHQLAILLKELTKNLFQELTLVIIPEKISGMWYDLVK KEFGGNSSVPTQFITIETLQKANDYILGNLLLGLYSKSGIQPWI LNSPLSSDCFIGLDVSHEAGRHSTGIVQVVGKDGRVLSSKANTS NEAGEKIRHETMCQIVYSAIDQYQQHYNERPKHVTFHRDGFCRE DLLSLDEVMNSLDVQYDMVEIIKKTNRRMALTVGKQGWETKPGL CYLKDESAYLIATNPHPRVGTAQPIKIIKKKGSLPIEAIIQDIY HLSFMHIGSLLKCRLPITTYYADLSSTFFNRQWLPIDSGEALHF V
6	29 *Hyphomonas*	MPHTSLLLNFLPVSLSGDTRIHVGYRPYNEDVLRELREEFGESH VFKRDYQEDTISEIPVIPGAEPLSDKSTGVDLAEARWLWKPLLN AALLRLFSGSREITSDYPVSVLGNPKNNFISHANLPDWVRILPL LEFESRTLFGGKSGPQFGLVCNARTRHQVLAGCDHLIERGISPI GRYVQIDQPQRDSRLAPRGLTVGKVSSIDGDTLILEDHRKGYER VKASDARLTGNRADFDWCVNALLPGQGQATLSRAWDAMSALNQG PGRLQMINQTAEYLRTVNLEAVPGVAFEIGEWLSSTDAQFPVTE TIDRPTLVFHPSGRPNDTWNERGIKDNGPHDQRTFTPKQLNIAV ICQGRFEGQVDRFVGKLLDGIPDFQLRNGRKPYDDGFLSRFRLE RANVQTFQANSASREAYEAACEDALKHAADNGFGWDLAIVQIEE DFKALPGPQNPYYATKAMLLRNNVAVQNIRIETMSEPDKSLVYT MNQVSLACYAKLGGRPWLLGAQQSVAHELVIGLGSHTEQQSRFD QSVRYVGITTVFSSDGGYHLSERTGVVPFEDYAKELTDTLTRTI ERVRREDNWKNTDRVRLVFHAFKQIKDIEAEAIKQAVESLDLEN VVFAFVHVAEHHPYLIFDQNQEGLPHWEKNRSKRKGVLGPSRGV HIKLADSESLVVFAGASELKQAAHGMPRACLLKLHRNSTFRDMT YLARQAFDFTAHSWRVMTPEPFPITIKYSDLIAERLAGLKQIET WDDDAVRFRNIGKAPWFL
7	23 *Calothrix sp. PCC 7103*	MIMSLESNIFTFSNLGTLTTQYRLYEIRGLQKRHQEYYQNRQIL IHRLSYLLKNAVTIIERDEKLYLVVAADAPEPPNSYPIVRGVIY FKPTGQILTLDYSLRTPQNEEICQRFLHFMVQSALFQNANLWQP SAGKAFFEKKPSFEFGSILLFQGFSVRPIFTKDKIGLCVDIHHK FVSKEPLPSYLNFNEFQKYRGVSCIYHFGHQWYEIQLSELSELN ATEAMVPIENKFVTLINYITQQARKPIPEELANVSQDAAVVHYF NNQNQDRMAVTSLCYQVYDNSYPEIRKYHQHTILKPHIRRSAIH GIVQKYLAELRFGDITLKVSTIPELVPQEMFNLPDYCFGNDYKL SVKGSEGTAQISLDQVGKQRLELLSKAEAGIYVQEKFDRQYILL PQTVGDSFGSRFIDDLKKTVDKLYPAGGGYDPKIIYYPDRGLRT YIEQGRAILKTVEENELQPGYGIVMLHDSPDRLLRQHDKLAALV IRELKDYDLYVAVIHSKTGRECYELRYNNQGEPFYAVIHEKRGK LYGYMRGVALNKVLLTNERWPFVLSTPLNADVVIGIDVKHHTAG YIVVNKNGSRIWTLPTITSKQKERLPSIQIKASLIEIITKEAEQ TVDQLHNIVIHRDGRIHESEIEGAKQAMAELISRCTLPVNATLT ILEVAKSSPVSFRLFDVSNTNSKDPFVQNPQVGCYYIANSTDAY LCSTGRAFLKFGTVNPLHIRYVEGTLPLKLCLEDVYYLTALPWT

		KPDGCIRYPITVKINDRRLGEDASEYDEDALRFELFESLESEDD FDEMTDSDFNQEETMV
8	04 *Deinococcus sp. YIM 77859*	LKLNHFPLNPDLPLYITEYAHRNPRALLGFVRGQGFWAQQVGEQ VQVYHGRPQPTFRGVQVISHTRLDPDHPAFDQGVLSLIRQALVR AGYVLTYRERMAIHPRLERVVLRPPDRHPAELTVHAHLRWEWEL ERHSGQRWLVLRPGRRHLSALPWPAEAVQMWSAALPATCQKLHA LCLDRGQQMALLRQEDGWHFANPGAATQGRWHLSFSPQALHELG LAQAAHHAAAFRWDEVQRLVQLTDLWKPFVTSLEPLEVAAPIIA GKRLRFGRGLGRDVTEVHKRGILEPPPLPVRLAVVSPHLPDEHA NAQLRRELLAHLLPRHQVLRSAESRQGLHEHLRRQDQDDTLYTF WSGGEYRKLGLPPFDLARGLHTYDPASGQLQQPAALAPAPAQAT QAGRQLIALVVLPDDLTRSVRDTLFQQLQQLGLRCLFSVSRTLL HRPRTEYMAWVNMAVKLARTAGAVPWDLADLPGVTEQTFFVGVD LGHDHTHQQSLPAFTLHDHRGRPLQSWTPPRRTNNERLSLAELK KGLHRLLARRSVDQVIVHRDGRFLAGEVDDFTLALHDLGIPQFS LLAIKKSNHSVAVQAEEGSVLSLDERRCLLVTNTQAALPRPTEL ELVHSDRLSLATLTEQVFWLTRVFMNNAQHAGSDPATIEWANGI ARTGQRVPLAGWRL
9	37 *Chroococcidiopsis thermalis*	MNNVMQEFPVASFPTFLSEISLLDITPKNFICFRLTPEIERKTG NSFSWRFSQKFPDAVVIWHNKFFWVLAKPNRPMPSQEQWREKLL EICEELKKDIGDRTYAIQWVSQPQITPEILSQLAVRVLKINCRF SSPSVISVNQVEVKREIDFWAETIEIQTQIQPALTITVHSSFFY QRHLEEFYNNHPYRQNPEQLLIGLKVRDIERNSFATITDIVGTI ADHRQKLLEDATGAISKQALIEAPEEQPVVAVQFGKNQQPFYYA MAALRPCITAETARKFDVDYGKLLSATKIPYLERKELLALYKKE AGQSLATYGFQLKISINSRRHPELFFSPSVKLSETKLVFGKNQI GVQGQILSGLSKGGVYRRHEDFSDLSRPIRIAALKLCDYPANSF LQETRQRLKRYGFETLLPVENKKTLLVDDLSGVEARAKAEEAVD ELMVNHPDIVLTFLPTSDRHSDNTEGGSLYSWIYSRLLRRGIAS QVIYEDTLKSVEAKYLLNQVIPGILAKLGNLPFVLAEPLGIADY FIGLDISRSAKKRGSGTMNACASVRLYGRKGEFIRYRLEDALIE GEEIPQRILESFLPAAQLKGKVVLIYRDGRFCGDEVQHLKERAK AIGSEFILVECYKSGIPRLYNWEEEVIKAPTLGLALRLSAREVI LVTTELNSAKIGLPLPLRLRIHEAGHQVSLESLVEATLKLTLLH HGSLNEPRLPIPLFGSDRMAYRRLQGIYPGLLEGDRQFWL
10	27 *Thermosynechococcus elongates*	MPTQFQEVEVILNRFFVKKLSRPDLTFHEYQCQFTQVPEQGSEQ KAISSVCYKLGVTAVRLGSCIITREPIDPERMRTKDWQLQLIGC RELSCQNYRERQALETFERKILEEKLKETFKKTIIEKDYELGLI WWISGEEGLEKTGHGWEVHRGRQIDLKIETDEKLYLEIDIHHRF YTPFKLEWWLSEYPNIQIKYVRNTYKDKKKWILENFADKSPNEI QIEALGISLAEYHRQEGATQQEIDESRVVIVKKISDYKAKPVYH LSQRLSPILTMETLAQIAEQGREKKEIQGVFDYIRKNIGTRLQE SQKIAQVIFKNVYNLSSQPEIMKVNGFVMPRAKLLARNNKEVNQ TARIKSFGCAKIGETKFGCLNLFDNKPEYPEEVHKCLLAIARSS GVQIKIDSYFTGSDYPKDDLAQQRFWQQWAAQGIKTVLVVMPWS PHEEKTRLRIQALKAGIATQFMIPTPQDNPYKALNVALGLLCKA KWQPVYLKPLDDPQAADLIIGFDTSTNRRLYYGTSAFAILANGQ SLGWELPDIQRGETFSGQSIWQVVSKLVLKFQDNYDSYPKKILL MRDGLVQDGEFEQTIRELTHQGIDVDILSVRKSGSGRMGRELTS

		GNTAITYDDAEVGTVIFYSATDSFILQTTEVIKTKTGPLGSARP LRVVRHYGNTPLELLALQTYHLTQLHPASGFRSCRLPWVLHLAD RSSKEFQRIGQISLLQNVDREKLIAV

Table 2: Mesophilic Ago Nucleic Acid Sequences

SEQ ID NO	Ago/Genus Species	Nucleic Acid Sequence
11	69 (Referred to herein as Ago69 or Argonaute 69) *Clostridium perfringens WAL-14572* Accession No. NZ_JH594533.1	ATGGTAGGAGGCTATAAAGTGTCAAATTTAACAGTTGAAGCATT TGAAGGAATTGGAAGTGTAAACCCTATGTTGTTTTATCAATATA AGGTAACTGGAAAAGGCAAATATGATAATGTGTATAAAATCATT AAGAGTGCAAGGTATAAAATGCATTCAAAAAATAGATTTAAGCC TGTATTTATAAAAGATGATAAGCTTTATACATTAGAAAAACTTC CAGATATAGAAGATTTAGATTTTGCAAATATAAACTTTGTTAAA AGCGAGGTTTTGAGTATAGAGGACAATATGAGTATATATGGGGA GGTTGTTGAGTATTATATAAATTTAAAGCTTAAGAAAGTTAAAG TTTTAGGTAAGTATCCTAAATATAGAATTAATTATTCTAAGGAG ATACTTAGCAATACTTTGTTAACAAGAGAGTTAAAGGATGAATT TAAGAAAAGTAATAAAGGATTTAATTTAAAGCGTAAATTTCGAA TTTCACCAGTAGTTAATAAAATGGGTAAAGTGATTTTATATTTA AGCTGTTCAGCTGATTTTTCAACTAATAAGAATATTTATGAAAT GCTTAAAGAAGGATTAGAAGTAGAAGGGTTAGCTGTAAAAAGTG AATGGTCAAATATAAGTGGAAACTTAGTTATAGAAAGTGTATTA GAAACAAAAATAAGTGAGCCAACAAGTTTAGGGCAATCTTTGAT AGATTACTATAAAAATAATAATCAAGGGTATAGAGTTAAAGATT TTACTGATGAGGATTTAAATGCAAACATAGTAAATGTAAGGGGC AATAAGAAAATATATATGTACATACCACATGCATTAAAACCTAT TATAACTAGGGAGTATTTAGCTAAAAATGATCCAGAATTTTCTA AAGAGATAGAACAATTAATAAAAATGAATATGAATTATAGATAT GAGACCTTAAAGTCATTTGTGAATGATATTGGAGTTATTGAAGA ACTTAATAACTTAAGTTTTAAAAATAAATATTATGAAGATGTTA AATTATTAGGTTATAGCAGTGGGAAAATAGATGAACCAGTACTT ATGGGAGCAAAAGGGATTATAAAAAATAAGATGCAAATCTTTTC TAATGGATTTTATAAGTTACCAGAGGGGAAAGTTAGGTTTGGAG TTTTATATCCTAAAGAGTTTGATGGAGTAAGTAGAAAAGCTATA AGAGCTATATATGATTTTTCTAAAGAGGGAAAATATCATGGCGA AAGTAATAAATACATAGCAGAGCATTTAATAAATGTAGAATTTA ATCCTAAAGAATGTATCTTTGAAGGATATGAACTAGGAGATATT ACTGAATATAAAAAGGCTGCATTAAAGTTAAATAATTATAATAA TGTAGATTTTGTAATAGCTATTGTACCTAATATGAGTGATGAAG AGATAGAAAATTCATATAATCCTTTTAAGAAGATATGGGCTGAA TTGAATTTACCATCTCAAATGATATCTGTAAAGACAGCAGAAAT CTTTGCAAATAGTAGAGATAATACAGCATTATATTATTTACATA ATATAGTCTTAGGCATTTTGGGAAAAATAGGAGGAATACCATGG GTTGTTAAGGATATGAAGGGGGATGTAGATTGTTTTGTTGGATT AGATGTTGGAACTAGGGAGAAGGGAATACACTATCCAGCGTGTT CAGTTGTTTTTGATAAATATGGGAAGCTTATAAATTACTATAAA CCTAATATACCTCAAAATGGTGAAAAGATTAATACTGAAATACT

		TCAAGAGATTTTTGATAAAGTATTAATTTCATATGAAGAGGAAA ATGGAGCTTATCCTAAAAATATTGTTATACATAGGGATGGCTTT TCAAGAGAAGATTTAGATTGGTATGAAAATTATTTTGGAAAAAA GAATATAAAGTTTAATATAATAGAAGTTAAAAAAAGTACACCAT TAAAAATTGCATCTATTAATGAGGGCAATATAACAAACCCAGAA AAGGGAAGTTATATATTAAGAGGAAATAAGGCATATATGGTTAC TACTGATATTAAAGAAAATTTAGGATCACCAAAACCATTAAAGA TAGAAAAATCTTATGGGGATATAGATATGTTAACTGCATTGAGT CAGATATATGCACTAACTCAAATACATGTTGGAGCAACAAAGAG TTTAAGACTTCCAATAACTACAGGATATGCAGATAAGATTTGTA AAGCAATAGAGTTTATTCCTCAAGGAAGAGTTGATAATAGATTG TTCTTTTTATGA
12	69 *Clostridium perfringens WAL-14572* NZ_JH594533.1 Human codon optimized nucleic acid sequence	ATGGTCGGCGGCTATAAAGTCAGCAATTTGACAGTGGAAGCGTT CGAAGGTATCGGGAGTGTCAACCCGATGCTGTTTTACCAATACA AAGTCACCGGAAAGGGAAAGTACGATAATGTGTATAAGATTATC AAAAGCGCACGGTACAAGATGCATTCTAAGAACCGATTCAAGCC CGTGTTCATCAAGGACGACAAACTGTACACCCTCGAGAAGCTCC CGGATATAGAAGACCTGGATTTCGCAAACATTAACTTCGTGAAA AGCGAGGTTCTCAGCATAGAGGATAATATGTCAATTTATGGCGA GGTGGTGGAATACTATATCAATCTCAAGCTGAAAAAAGTGAAGG TGTTGGGAAAATACCCCAAGTACAGGATCAATTACAGCAAAGAG ATTCTCAGTAATACGCTGCTGACACGAGAGCTCAAAGACGAGTT TAAGAAATCAAATAAGGGTTTTAACCTGAAACGGAAGTTTAGAA TTTCCCCCGTGGTGAATAAGATGGGCAAAGTGATACTCTATTTG TCCTGCAGTGCTGATTTCAGCACCAACAAGAACATTTACGAAAT GTTGAAAGAGGGCTTGGAGGTTGAGGGGCTGGCCGTTAAGAGCG AGTGGAGCAATATCAGTGGCAACCTGGTGATCGAGAGCGTACTG GAAACCAAGATATCCGAGCCCACTAGCCTGGGCCAATCCCTGAT AGACTACTATAAGAATAACAACCAGGGCTATAGGGTGAAGGATT TCACCGATGAGGATCTGAATGCCAACATTGTCAACGTGAGAGGA AATAAGAAGATCTATATGTATATTCCGCACGCGTTGAAGCCGAT AATCACCCGGGAGTACCTGGCCAAGAACGATCCAGAGTTTTCTA AGGAGATCGAGCAGCTTATCAAGATGAATATGAACTACCGATAT GAAACCCTCAAGTCATTTGTGAATGACATCGGGGTCATTGAAGA GCTGAACAACCTGAGCTTCAAAAACAAATACTACGAAGATGTGA AACTGCTGGGTTACTCCAGCGGCAAAATAGACGAACCCGTCCTG ATGGGGGCAAAAGGGATCATAAAGAACAAAATGCAGATTTTTTC CAATGGATTCTACAAACTCCCCGAAGGCAAGGTACGATTTGGCG TTCTGTACCCAAAAGAATTTGATGGCGTGTCAAGGAAAGCTATC CGCGCCATTTATGACTTCAGTAAGGAGGGCAAATACCACGGCGA AAGCAACAAGTATATCGCGGAACACCTGATAAACGTGGAGTTCA ATCCAAAGGAGTGCATATTTGAGGGATACGAACTGGGCGATATC ACCGAATACAAGAAGGCGGCTCTGAAACTTAATAACTACAACAA TGTCGACTTCGTAATCGCAATAGTCCCGAACATGTCCGACGAAG AGATAGAGAACAGCTACAATCCGTTCAAGAAAATATGGGCCGAA CTGAATCTGCCCAGCCAGATGATTAGCGTCAAGACGGCCGAAAT CTTTGCCAATAGCAGGGATAACACGGCGCTTTACTACCTGCATA ACATCGTCCTCGGTATCCTGGGTAAGATAGGAGGGATTCCCTGG

		GTGGTTAAAGACATGAAGGGCGACGTGGATTGCTTCGTTGGACT CGATGTCGGCACCAGGGAGAAGGGCATACATTACCCCGCCTGCA GCGTTGTGTTTGACAAGTACGGCAAGCTTATTAACTATTACAAG CCTAACATCCCGCAGAACGGAGAGAAGATTAACACAGAAATACT TCAGGAAATTTTCGACAAGGTGCTCATAAGCTATGAGGAGGAGA ATGGAGCCTACCCGAAGAATATCGTGATCCACAGGGACGGCTTT AGCCGAGAGGACCTTGACTGGTATGAGAACTACTTCGGTAAGAA AAACATAAAGTTTAACATCATCGAAGTCAAAAAGTCAACTCCGT TGAAAATCGCCAGTATAAACGAGGGAAATATCACGAATCCTGAA AAGGGTTCCTACATCCTGCGCGGCAACAAAGCCTACATGGTGAC CACAGATATTAAGGAAAACCTGGGAAGCCCAAAGCCCCTGAAGA TAGAAAAGAGCTACGGCGACATAGACATGCTCACAGCTCTCAGC CAAATATACGCACTCACGCAAATCCATGTGGGGGCGACCAAAAG CCTGCGCCTCCCAATCACCACCGGCTACGCCGACAAGATTTGCA AGGCGATCGAGTTCATCCCCCAAGGGCGCGTGGACAACCGCCTT TTCTTTCTG
13	70 *Clostridium saudiense*	ATGAATAATTTAATGTTAGAAGCTTTTAAAGGAATAGGAACAAT AAAACCATTGGTTTTTTATAGATACAAATTAATAGGAAAAGGTA AAATAGAAAATACATATAAAACCATAAGCAATGCTAAAAATAAG ATGAGTTTTAATAATAAATTTAAAGCAACATTTAGTAAAGGAGA AACAATATATACATTAGAAAAGTTTGAAGTAATGCCAAATTTAA ATGATGTAACAATTGAATTTGATGGTGAGGAAGTATTACCTATA AAAGATAATAATGAAATTTATTCTGAAGTTGTTCAATTTTATAT TAATAATAATTTACGTAAGATTAAGTTAGATAATAAATATCAAA AGTATAGAGCTACAAATACAAGGGAAATAACAGGTAATGTTATA TTAGATAAAGATTTTAAGGAAAAATATAAAAAGAGTAAAAGTGG ATTTCAATTAAAAAGAAAATTTATAATTTCTCCTAAGGTAAATG ATGAGGGAAAAGTAACTTTATTTTTAGATTTAAATTCTAGTTTT GATTATGATAAAAATATTTACCAAATGATAAAGGCTGGAATGGA TGTAGTAGGTCAAGAGGTAATTAATACATGGAACAATAAAAAAC AAAAAGGGAAGATCAAGAAAATATCAGAATTAACAATAAGTGAG CCATGTAACTTTGGACAATCCTTAATTGATTACTATGTTAGTTT AAATCAAGCTGTCAGGGTTAAGAACTTCACAGAAGAAGAGAAGA ATACAAATGTTATAGTAGTTCAAGTTGGAAAAGGTGAAGTAGAA TATATTCCACATGCATTAAAACCAATTATAACTAGGGAGTATAT TAAAAAATATGATGAAGCTTTTTCAAAAGAAGTAGAAAATCTAA TCAAAATAAATATGAGTTATAGGTATGAAATACTTAAGAAATTT ATTGATGATATAGGAAGTATAACTGAGTTAAATAATTTAAAGTT TGAAAATACATATATAGATAATATTGAAAGTTTGGGGTACCAGC AAGGTAAATTGAATGATCCAGTATTAATTGGGGGTAAAGGGATA CTAAAAGATAAGATTCATGTATTTAAAAGTGGATTTTATAAATC TCCAATTGATGAGGTGAAGTTTGGAGTTATTTATCCAAAAGGAC ATACTAATGATAGCAAAAGTACAATTAGAGCTATATATGATTTT TGTACTGATGGAAAATATCAGGGAAAAGATAATATATTTATAAA TAATAAATTAATGAATATAAAATTTAGTAATCAAGATTGTGTTT TTGAAGAGTATGAATTAAATGATATTACAGAGTATAAAAGGGCT GCTAATAAGCTTAAAAATAATGAAAATATTAAATTTGTCATTGC TATTATACCAGCAATAGATGAAAGTGATATTGAAAATCCATATA ATCCTTTTAAGAGAGTTTGTGCAGAATTAAATTTACCATCACAA

		ATGGTTTCTTTAAAAACTGCAAAAAGATTTGGGACAAGTAAAGG TAATAATGAACTTTATTTTTTACATAATATTTCTTTAGGTATAT TAGGTAAGATAGGAGGAGTTCCATGGGTAATTAAAGATATGCCT GGAGAAGTAGATTGTTTTGTTGGATTAGATGTGGGAACAAAGGA GAAAGGAATACATTATCCCGCTTGTTCAGTGCTTTTTGATAAGT ATGGTAAGTTAATAAATTATTATAAGCCTACAATACCTCAAAGT GGAGAAATAATTAAAACGGATGTTTTACAAGAGATATTTGATAA GGTACTACTTTCTTATGAAGAGGAAAATGGGCAGTATCCAAGAA ATATTGTTATTCATAGGGATGGTTTTTCTAGGGAAGATTTAGAG TGGTATAAGAATTATTTTATTAAAAAGAATATTAATTTTACAAT AGTAGAAATTAAAAAGAACTTTGCAACTAGGGTAGCAAATAATA TTAATAATGAAGTTAGTAATCCTTTTAAGGGAAGTTTTATTTTA AGGGAAAATGAAGCAATAGTAGTTACAACGGATATTAAGGATAA TATTGGAGCACCTAAGCCAATTAAGGTGGAAAAGACATATGGAG ATATTGATATGATGACTATAATAAATCAAATATATGCATTAACT CAAATTCATGTTGGATCTGCAAAGAGTATGAGATTGCCAATAAC AACAGGTTATGCAGATAAGATTTGTAAATCTATAGAGTATATAC CTTCAGGGCGAGTTGATAATAGATTGTTCTTTTTATAG
14	41 *Clostridium disporicum*	ATGAATAACTTGACACTAGAAGCATTTAGAGGGATAGGAACAAT AAAGCCACTACTTTTTTATAGATATAAATTGATAGGAAAAGGTA AAATAGAAAATACATATAAAACAATAAGAAATGCTCAAAATAGA ATGAGTTTTAATAATAAATTTAAAGCAACATTTAGTAAAGATGA AATAATATATACATTAGAAAAATTTGAAATAATACCAACTTTAG ATGATGTAACAATCATTTTTGATGGAGAAGAGGTTTTACCTATA AAAGATAATAATAAAATTTATTCTGAAGTAATTGAGTTTTATAT TAATAATAATTTACGTAATGTCAAATTTAATTATAAATATCCAA AGTATAGGGCAGCAAATACAAGAGAAATAACAGGGAATGTTATA TTAGATAAAGATATGAATGAGAAATACAAAAAGAGTAATAAGGG GTTTGAGTTAAAAAGAAAATTTATAATTTCTCCTAAGGTAGATG ATGAGGGAAAAGTAACTTTATTTTTAGATTTAAATGCAAGTTTT GATTATGATAAAAATATTTATCAAATGATAAAAGCTGGAATAGA TGTAGTAGGAGAAGAGGTTATTAATATCTGGAGTAATAAAAAAC AAAGAGGAAAAATTAAAGAGATATCAGATATAAAAATAAACGAA CCGTGTAATTTTGGACAATCATTAATTGATTATTATATTAGTTC TAATCAAGCTTCTAGAGTTAATGGATTTACTGAAGAAGAAAAAA ATACAAATGTAATAATAGTTGAATCAGGAAAAAGTCGCTTAAGT TATATTCCACATGCATTAAAACCAATTATAACGAGAGAATATAT TGCGAAAAATGATGAAGTTTTTTCAAAAGAAATTGAAGGTTTAA TTAAAATTAATATGAACTATAGATATGAAATACTTAAGAGATTT GTTAGTGATATAGGAACTATAAAAGAATTAAATAATTTGAGGTT TGAAAAAATATATATGGATAATATCGAAAGTTTAGGGTATGAGC AAGGTCAATTAAAGGATCCAGTATTGATTGGGGGTAAAGGAATA CTAAAAGATAAAATTCATGTTTTTAAAAGTGGATTTTATAAATC ACCAAACGATGAAATAAAGTTTGGAGTTATTTATCCTAGGGGAT ATATTAAGGATACTCAAAGTGTGATCAGAGCTATATATGACTTT TGTACTGAAGGAAAATATCAGGGAAAAGATAATATATTTATAAA TAATAAATTAATGAATATAAAGTTTAGTAATAAAGAGTGTGTTT TTGAAGAGTATGAATTAAATGATATTACTGAGTATAAAAGAGCT GCAAATAAACTTAAAAAGAATGAAAATATTAAATTTGTTATTGC

		TATTATACCAACAATAAATGAAAGTGATATTGAAAATCCATATA ATCCTTTTAAGAGAGTTTGTGCTGAAATAAATTTACCATCACAA ATGATTTCTTTGAAAACAGCAAAAAGATTTAGTACAAGCAGGGG ACAGAGTGAACTTTATTTTTTACATAATATTTCCTTAGGTATAT TAGGAAAGATAGGAGGAGTTCCATGGGTAATTAAAGATATGCCT GGAGAAGTAGATTGTTTTGTTGGATTAGATGTGGGAACAAAGGA GAAAGGAATACATTATCCAGCTTGTTCAGTACTTTTTGATAAGT ATGGTAAGTTAATAAATTATTATAAACCTACAATACCTCAAAGT GGAGAAATAATTAAAACGGATGTTTTACAAGAGATATTTGATAA GGTATTACTTTCTTATGAAGAGGAAAATGGTCAGTATCCAAGAA ATATTGTTATTCATAGGGATGGTTTTTCTAGGGAAGATTTAGAA TGGTATAAGAACTATTTTCTTAAAAAGAACATTGAATTCTCTAT TGTAGAAGTAAGAAAGAATTTCGCAACTAGGTTAGTGAATAATT TTAATGATGAAGTTAGCAATCCTAGTAAAGGAAGTTTTATTTTA AGAGATAATGAAGCAATAGTAGTTACAACTGATATTAATGATAA TATGGGGGCACCTAAGCCAATTAAGGTGGAAAAGACATATGGAG ATATTGATATGTTAACTATAATAAATCAAATATACGCATTAACT CAAATTCATGTTGGTTCTGCTAAGAGTTTAAGGTTGCCTATAAC AACTGGATATGCAGATAAGATTTGTAAGGCTATAGATTATATAC CTTCTGGACAAGTTGATAATAGGTTATTCTTTTTATAG
15	51 *Rhodopirellula maiorica*	ATGCTGCAACTGAACGGATTTTCAATCGAGATTGCCGGCGGGTC GTTGACGGTACTGAAGTCGAAGATCGCACCGACGGACGTCAAGG AAACGCGACGTTCGCTCGAGGACGATTGGTTTACGATGTATCAC GAAGGGCACCTCTATTCCCTTGCAAAGAACTCGAACGCATCGGG CGGGCTTGGTGAGACGGAACTCTTGGTGCTCTCCGACCACCTCG GGCTGCGTTTTGTAAAAGCCATGCTCGATCAGGCGATGCGAGGC GTCTTTGAAGCGTACGATCCTGTACGCGATCGCCCGTTTACCTT TCTGGCTCGCAATGTCGATCTTGTCGCGTTAGCTGCCGAGAATT TGGAATCAAAGCCCAGTTTGCTTTCTAAGTTTGAGATTCGACCT AAGTATGAACTAGAAGCAAAAGTGGTTGAGTTCCGGCCGGGCGA GCTGGAATTGATGCTCGCACTCAATCTGACCACTCGTTGGATCT GCAACGCCAGCGTGGATGAATTGATCGAAAAGAACATTCCAGTC CGGGGAATGCATCTGATTCGCAGGAATCGTGAGCCAGGACAACG AAGCTTGGTCGGGACTTTCGACCGAATGGAAGGAGACAACGCTC TACTCCAGGATGCGTACGACGGCCAGGACAAGATCGCTGCATCG CAAGTCCGAATCGAGGGATCGAAGGAGGTCTTCGCGACAAGTCT CCGGCGTCTGCTTGGCAATCGGTACACCAGCTTTATGCACTCAG TGGACAATGAGTATGGGAAGTTGTGTGGCGGTCTTGGGTTTGAC GGTGAGCTTCGCAAAATGCAAGGATTTCTTGCGAAGAAGAGCCC GATTCAATTGCATGGCGGTGTGGAGGTGTCGGTCGGACAGCGAG TTCAGCTAACCAATCAGCCGGGGTACAAAACGACTGTCGAACTG CTGCAAAGCAAATACTGCTTCGACCGATCTCGAACGAAACTACA TCCATACGCTTGGGACGGTTTAGCTAGATTCGGGCCGTTTGACC GCGGAAGCTTTCCCACACGATCTCCGCGCATTTTGCTTGTCACA CCCGATTCGGCATCCGGCAAGGTCAGCCAAGCTTTGAAGAAATT TCGCGACGGTTTCGGGTCAAGCCAATCGAGCATGTACGACGGAT TTCTCGATACTTTCCATCTTTCGAACGCACCCTTTTTCCCGCTC CCTGTCAAATTGGATGGCGTCCAGCGATCGGATGTTGGCAAGGC ATATCGAAAAGCAATCGAAGACAAGTTGGCCCGTGATGATGATT

		TCGATGCTGCGTTTAACATTCTGCTGGATGAACACGCCAATCTC CCCGACTCGCACAACCCGTATCTGGTTGCCAAATCAATCCTGCT CTCGCATGGAATCCCCGTGCAAGAAGCCAGGGTGTCGACGCTTA CTGCGAACGAGTATTCGCTACAGCACACATTCAGAAATGTTGCC ACCGCACTGTACGCAAAAATGGGCGGAGTCCCTTGGACGGTCGA TCACGGCGAAACGGTTGACGATGAACTGGTGGTGGGAATTGGGA ACGCTGAACTGTCCGGCAGCCGCTTCGAGAAACGACAGCGTCAC ATCGGAATCACCACGGTATTCCGTGGCGATGGCAACTATCTTCT CTCAAATCTGTCCAAAGAGTGCAGATACGAAGACTACCCCGACG TGCTTCGCGAATCGACGATCGCAGTCCTTCGCGAAGTCAAACAA CGAAACAACTGGCTGCCCGGACAAACAGTGAGAATTGTTTTCCA TGCGTTCAAACCGCTGAAGAATGTCGAGATCGCGGACATCATCG CGTCAAGCGTCAAAGAAGTCGGCAGCGAACAGACGATTGAATTC GCTTTCTTGAATGTCTCGCTGGACCATTCGTTCACGTTGCTGGA TATGGCCCAGCGAGGAATAACGAAGAAGAACCAGACGAAAGGAA TTTACGTTCCTCGCCGAGGAATGACCGTTCAGGTTGGACGGTAC ACACGGCTCGTCACGTCAATCGGTCCTCACATGGTCAAGCGTGC AAACCTTGCGTTGCCCAGGCCCCTGTTGATCCACCTGCACAAGC AATCGACTTACCGCGACCTTTCCTACCTTTCTGAGCAGGTGCTG AACTTCACGACGCTGTCGTGGCGATCAACGCTGCCGTCGGAGAA GCCGGTCACGATTCTTTACTCATCACTTATCGCTGACTTGCTGG GCCGTCTCAAGTCCGTGGACGATTGGTCGCCAGCGGTCCTCAAT ACCAAACTTCGCAACAGCAAGTGGTTTCTGTGA
16	02 *Paenibacillus odorifer*	ATGAATACACCACTAACACACTACGTACTTACAGAATGGGAATC AGACACGAACACAAATGTTCTACATATTCATTTATATACGCTAC CAGTGCGTAATGTATTTGAACAGCATAAAGAAAACGGAAATGCT TGTTTTGACCTTAGAAAATTAAACAGATCGCTCATTATCGATTT TTATGATCAATATATTGTAAGTTGGCAACCTATTGAAAATTGGG GTGAATACACATTCACTCAGCATGAATATCGCTCAATTAATCCC ACCATCTTAGCAGAAAGAGCTATCTTGGAACGATTACTTTTAAG AACTATAGAAAGCGTTCAGCCAAAAAAAGAAATTGCTGCTGGAA GTCGAAAATTCACGTGGTTAAAAGCTGAAAAAGTCGTAGAAAAT ATTTCAATACACAGGGTTATACAATGTGATGTTACTGTAGATTA TGCGGGCAAAATTTCCGTTGGCTTTGATTTAAATCATAGTTATC GTACAAATGAATCGGTATATGATCTCATGAAATCAAATGCTATT TTTAAAGGCGATCGAGTAATTGATATATATAACAATCTTCATTA TGAGTTTGTGGAAATCTCCAATTCCACAATCAATGATTCAATTC CAGAACTTAATCAATCTGTTGTTAATTATTTTACTAAAGAACGA AAGCAAGCTTGGAAAGTTGATAAACTTGAACAGAGTATGCCTGT TGTCTATCTAAAAGCGTTTAATGGATCTCGTATTGCTTATGCAC CTGCTATGCTACAAAAAGAACTTACTTTTGAGTCGCTTCCTACT AATGTAGTACGTCAAACATCAGAAATTTTTAAACAAAACGCAAA TCAAAAAATTAAGACATTGCTGGATGAAATACAAAAAATATTAG CACGCACTGATAAAATTAAATTTAATAAACAAAAGCTCCTAGTC CAGCAAGCTGGGTATGAAATATTGGAGTTGTCAAATCCTAATCT TCAATTCGGAAAAAATGTTACTCAAACACAGCTTAAATATGGAC TTGATAAGGGCGGTGTTGTAGCCTCAAAACCTTTATCAATTAAT TTGCTTGTATACCCAGAACTGATAGATACAAAACTCGATGTCAT AAATGACTTTAATGATAAATTAAATGCGCTATCCCATAAATGGG

		GTGTACCATTATCAATCTTAAAAAAATCAGGAGCATATCGGAAC AGACCGATAGATTTCACAAATCCACACCAACTTGCCATCCTACT GAAAGAACTTACAAAAAATTTATTTCAAGAGTTGACGCTCGTGA TAATTCCGGAGAAAATTTCAGGAATGTGGTATGACTTGGTAAAG AAGGAATTCGGAGGAAATTCCTCAGTACCAACTCAGTTCATAAC TATAGAAACTCTTCAAAAAGCTAATGACTACATTCTAGGCAATC TATTATTAGGACTCTATTCTAAATCTGGCATTCAACCGTGGATC TTAAATTCTCCTCTGTCATCTGATTGTTTTATTGGGTTGGATGT ATCTCATGAAGCTGGTAGACACTCTACAGGAATTGTACAGGTCG TAGGAAAAGACGGCAGAGTATTGTCTAGCAAGGCAAACACCTCT AATGAGGCCGGTGAAAAAATCCGTCATGAAACTATGTGTCAAAT TGTATACTCAGCTATAGACCAATATCAGCAACATTACAATGAAA GACCAAAGCATGTTACTTTTCATCGGGATGGTTTTTGTAGAGAA GATTTACTTAGTCTAGACGAAGTAATGAATAGTTTGGATGTACA ATATGACATGGTTGAAATCATCAAGAAAACGAATCGTCGCATGG CTCTAACCGTTGGTAAGCAAGGTTGGGAGACCAAGCCAGGATTG TGTTATCTAAAAGATGAGTCGGCATATCTTATCGCCACAAATCC TCATCCACGTGTCGGCACAGCACAACCGATTAAAATTATCAAGA AAAAGGGATCACTACCGATTGAAGCAATAATTCAAGATATTTAT CATCTTTCGTTCATGCACATTGGTTCATTATTAAAATGTCGCCT ACCTATTACTACGTATTATGCTGATTTAAGCTCTACATTCTTCA ACCGGCAATGGCTCCCGATCGATTCTGGCGAAGCCCTACATTTT GTATAA
17	29 *Hyphomonas*	ATGCCGCACACATCTCTTCTCTTGAACTTTTTGCCCGTATCGCT CTCCGGCGATACGCGAATTCACGTTGGCTATCGGCCCTACAACG AAGACGTTTTGCGGGAACTACGAGAGGAGTTCGGTGAAAGCCAC GTTTTCAAACGCGACTATCAAGAAGACACAATATCGGAGATTCC AGTAATCCCCGGTGCCGAACCACTCTCCGACAAATCGACCGGAG TAGACCTCGCCGAAGCACGGTGGCTTTGGAAGCCGCTTCTGAAC GCGGCTCTCCTGCGGCTGTTTTCAGGGTCCAGAGAGATTACAAG CGATTATCCTGTCAGCGTTTTGGGCAATCCGAAGAACAATTTCA TTTCGCATGCTAACTTACCTGATTGGGTAAGGATCTTGCCGTTA CTGGAATTTGAGTCTCGCACGTTGTTCGGTGGCAAGTCGGGTCC GCAGTTCGGACTGGTGTGTAATGCGCGGACCCGTCACCAGGTAT TGGCAGGTTGCGATCATCTCATTGAGCGGGGCATTTCTCCGATT GGGCGTTATGTTCAAATCGATCAACCGCAACGAGACTCCCGATT GGCGCCGCGCGGCCTTACAGTGGGAAAGGTATCATCGATCGACG GTGACACACTCATTTTGGAAGACCATCGCAAAGGCTATGAACGG GTCAAAGCATCTGATGCGAGGTTGACCGGCAATCGCGCTGATTT CGACTGGTGCGTCAATGCGCTTTTGCCGGGACAAGGGCAGGCTA CCCTGTCGAGGGCGTGGGACGCCATGAGCGCATTGAACCAGGGC CCAGGTCGCCTGCAAATGATCAACCAAACAGCCGAGTATTTGAG GACAGTCAATCTGGAAGCTGTTCCAGGCGTTGCTTTTGAAATCG GCGAATGGCTGTCGAGCACGGACGCTCAGTTTCCGGTGACGGAA ACGATCGATCGCCCCACGCTTGTATTTCACCCTTCGGGAAGGCC GAACGATACGTGGAACGAACGCGGCATCAAAGACAATGGACCGC ACGACCAACGCACGTTCACGCCAAAGCAACTCAACATCGCTGTC ATTTGCCAAGGCCGGTTCGAAGGACAAGTCGATCGTTTCGTGGG AAAACTTCTCGATGGCATCCCCGACTTTCAACTCCGAAATGGGC

		GCAAGCCTTATGATGACGGCTTCCTTAGTCGATTTCGTCTGGAA CGAGCCAATGTGCAGACATTTCAGGCAAACTCAGCCTCGCGCGA AGCATACGAAGCAGCTTGCGAAGACGCACTGAAACATGCTGCTG ACAATGGCTTTGGCTGGGACTTGGCGATTGTCCAGATTGAAGAA GACTTCAAGGCGTTGCCTGGGCCTCAGAACCCTTACTACGCCAC GAAGGCCATGCTGCTCCGCAACAATGTAGCGGTACAGAATATTC GTATCGAGACGATGAGCGAGCCGGACAAAAGTCTCGTGTATACA ATGAATCAGGTGAGTCTTGCCTGCTATGCGAAGCTGGGCGGTCG TCCGTGGCTTTTAGGGGCGCAGCAATCGGTTGCCCATGAACTCG TCATTGGCCTCGGGTCGCACACAGAACAACAATCTCGATTTGAT CAGAGCGTTCGGTATGTCGGCATTACCACTGTGTTTTCGAGCGA TGGCGGGTATCACCTCAGTGAACGCACCGGCGTCGTGCCATTTG AGGATTACGCCAAGGAATTAACCGACACGCTCACGCGCACCATC GAACGGGTCCGCCGGGAAGATAACTGGAAAAACACAGACCGGGT GCGGCTGGTCTTCCATGCGTTCAAGCAAATCAAGGACATTGAGG CTGAGGCCATCAAGCAGGCGGTCGAATCTCTCGACCTCGAAAAC GTGGTTTTCGCTTTCGTTCATGTTGCTGAACATCATCCGTATTT GATCTTCGACCAGAACCAAGAGGGATTGCCGCATTGGGAGAAAA ATCGGTCTAAACGAAAAGGCGTATTGGGCCCATCCCGAGGCGTG CACATCAAACTGGCTGATTCTGAGTCGCTGGTTGTGTTTGCGGG CGCAAGTGAACTTAAGCAAGCGGCCCACGGGATGCCGAGGGCCT GCCTTCTAAAACTGCATCGCAATTCGACCTTCCGCGATATGACT TATCTCGCACGCCAGGCGTTCGACTTCACCGCCCATTCTTGGCG AGTTATGACGCCGGAACCGTTTCCGATCACAATTAAGTATTCCG ACCTAATCGCTGAACGTCTTGCCGGCCTGAAGCAGATCGAGACG TGGGATGATGACGCCGTCCGGTTTCGCAACATCGGCAAGGCGCC CTGGTTCCTGTGA
18	23 *Calothrix sp. PCC 7103*	ATGATTATGAGTTTAGAAAGTAATATTTTCACCTTTTCCAATCT CGGAACGCTTACAACTCAATATCGTTTGTATGAAATACGAGGAC TTCAAAAGCGTCATCAAGAATACTATCAAAATCGACAAATTTTG ATACATCGGTTGAGCTATCTTCTGAAAAACGCTGTCACAATTAT AGAACGTGATGAAAAACTGTATCTTGTTGTTGCAGCAGATGCAC CAGAACCTCCTAATTCTTATCCAATTGTTCGAGGAGTAATTTAT TTTAAGCCAACTGGGCAAATTTTGACTTTAGACTATTCGTTACG TACACCACAGAATGAAGAAATTTGCCAGCGATTCTTACATTTTA TGGTACAGTCTGCATTATTCCAAAATGCAAATTTATGGCAACCA TCTGCAGGAAAAGCTTTCTTTGAGAAAAAACCGTCTTTTGAATT TGGCTCTATTCTATTGTTTCAAGGTTTTTCTGTTCGCCCTATAT TTACAAAAGATAAAATTGGATTATGTGTAGATATCCATCATAAA TTTGTAAGCAAAGAACCTCTCCCAAGTTATTTAAATTTTAACGA GTTCCAGAAATATCGCGGTGTTAGCTGTATTTATCACTTTGGTC ATCAATGGTATGAGATTCAACTTAGTGAACTTTCTGAGTTAAAT GCTACAGAAGCAATGGTTCCTATAGAAAACAAATTTGTTACATT GATCAACTATATTACTCAACAAGCTCGTAAGCCAATTCCAGAGG AGTTAGCAAACGTTTCTCAAGATGCAGCAGTAGTTCATTACTTT AATAACCAAAACCAAGACCGTATGGCTGTCACCTCGTTATGTTA TCAAGTTTATGATAATTCTTATCCTGAGATAAGAAAATACCATC AACATACGATTCTTAAGCCACATATTCGGCGTTCGGCAATCCAC GGTATTGTCCAAAAATATTTAGCTGAGTTGAGATTTGGAGATAT

		TACGTTAAAAGTTTCAACTATACCTGAACTAGTACCACAGGAAA TGTTTAACTTACCCGATTATTGTTTTGGTAATGATTATAAATTG TCAGTCAAGGGTAGTGAAGGAACTGCTCAAATAAGTTTGGATCA GGTTGGAAAACAACGACTAGAATTACTCAGCAAGGCAGAAGCAG GTATATATGTACAAGAGAAATTTGATAGACAATATATTCTACTA CCTCAAACAGTTGGAGATAGCTTTGGAAGTAGATTTATAGATGA TCTCAAAAAAACTGTTGATAAACTTTACCCAGCAGGTGGGGGAT ACGACCCTAAGATTATTTACTATCCGGACCGTGGTTTACGAACT TATATTGAACAAGGCAGAGCGATACTCAAAACTGTTGAGGAAAA CGAGCTTCAGCCTGGATACGGTATAGTGATGTTGCATGATTCAC CAGATAGACTACTACGTCAGCATGATAAGCTTGCAGCTTTAGTA ATTCGTGAATTGAAAGATTATGACTTATATGTTGCTGTAATTCA TTCTAAAACAGGTAGAGAATGCTATGAACTTCGGTATAACAATC AAGGAGAACCATTCTATGCAGTTATACATGAAAAACGTGGGAAG CTTTACGGGTACATGCGCGGTGTTGCTTTAAATAAGGTGTTGTT AACAAACGAACGTTGGCCTTTTGTCTTAAGTACACCTCTCAATG CGGATGTTGTTATTGGCATTGATGTTAAGCATCATACGGCAGGA TATATAGTTGTTAATAAAAATGGAAGTCGAATTTGGACTCTTCC AACAATTACTTCTAAACAAAAAGAGCGCTTGCCTAGTATCCAAA TAAAAGCTAGTTTAATAGAGATTATCACAAAGGAAGCAGAACAA ACAGTAGACCAGCTTCATAATATAGTTATTCATCGTGATGGTCG AATTCATGAATCAGAAATTGAAGGAGCAAAACAAGCAATGGCTG AATTAATATCAAGATGTACTTTGCCAGTAAATGCTACGCTTACT ATTCTAGAAGTTGCTAAATCTTCACCAGTATCATTTAGGCTATT TGACGTTTCAAATACTAATAGTAAAGACCCATTTGTTCAAAATC CTCAAGTAGGTTGCTATTATATAGCAAATTCTACGGATGCATAT CTCTGCTCTACAGGTCGGGCATTTCTCAAATTTGGAACTGTTAA TCCACTTCATATTAGATATGTGGAAGGTACACTTCCTTTAAAAT TATGCTTAGAAGATGTTTATTATCTCACGGCATTACCTTGGACA AAACCAGATGGCTGCATTCGTTATCCAATTACAGTAAAAATTAA CGATAGGCGCCTAGGGGAAGACGCAAGCGAGTACGATGAAGACG CACTTCGTTTTGAATTATTTGAAAGTTTAGAATCTGAAGATGAT TTTGACGAAATGACGGATAGTGATTTCAACCAAGAGGAAACAAT GGTATGA
19	04 *Deinococcus sp. YIM 77859*	TTGAAGCTCAACCATTTTCCCCTGAACCCTGATCTGCCGCTCTA CATCACGGAATATGCTCACCGCAACCCCCGCGCCCTGCTGGGCT TTGTGCGCGGGCAGGGCTTTTGGGCTCAGCAGGTGGGCGAACAG GTTCAGGTGTATCACGGCCGACCTCAGCCCACGTTTCGCGGCGT CCAAGTCATTTCGCACACGCGGCTTGACCCCGACCACCCGGCTT TCGACCAAGGGGTGTTGTCGCTGATTCGGCAGGCGCTTGTGCGA GCGGGATATGTGCTGACGTACCGGGAACGAATGGCCATCCATCC AAGGCTAGAACGGGTGGTCCTGCGCCCGCCCGACCGCCACCCGG CAGAACTCACTGTCCACGCCCATCTCCGTTGGGAATGGGAGCTG GAACGGCATTCCGGACAACGCTGGCTGGTGCTGCGGCCGGGGCG CCGTCATCTGTCTGCGCTGCCCTGGCCAGCCGAGGCGGTCCAGA TGTGGAGCGCAGCCTTGCCCGCCACCTGCCAAAAGCTTCATGCT CTGTGTCTCGACCGAGGCCAGCAGATGGCGCTTTTGCGCCAAGA GGACGGCTGGCACTTTGCCAACCCCGGAGCGGCGACCCAGGGCC GCTGGCATCTTAGCTTTTCTCCACAAGCGCTGCATGAGCTGGGC

		CTGGCCCAGGCCGCCCACCACGCCGCCGCGTTCCGCTGGGACGA GGTGCAGCGGCTCGTGCAGCTCACAGACCTCTGGAAACCCTTTG TGACCTCACTGGAACCGCTGGAGGTGGCTGCGCCCATCATTGCG GGGAAGAGGCTGCGCTTTGGGCGTGGCCTGGGGCGTGACGTGAC CGAGGTTCACAAGCGGGGAATTCTGGAACCGCCGCCCCTTCCCG TCCGACTGGCGGTGGTTTCACCCCACCTTCCCGATGAGCACGCC AACGCCCAACTGCGGCGCGAGCTGCTGGCCCACCTGCTGCCACG TCACCAGGTGCTGCGTTCTGCTGAATCGCGGCAGGGGCTCCACG AACACCTGCGGCGTCAGGACCAGGACGACACCCTGTATACCTTT TGGAGTGGCGGCGAGTACCGTAAACTGGGCCTCCCGCCTTTTGA TCTGGCGCGGGGCCTTCACACCTACGATCCCGCCAGCGGTCAGC TGCAGCAGCCCGCCGCACTGGCACCCGCACCCGCGCAGGCCACC CAAGCTGGCCGCCAACTGATCGCGCTGGTGGTCCTGCCCGACGA CCTCACCCGCAGCGTGCGCGACACCCTGTTTCAGCAGCTCCAGC AGCTTGGTCTCCGGTGCCTTTTTTCCGTAAGCCGCACACTCCTC CACCGGCCGCGCACCGAGTACATGGCCTGGGTCAATATGGCGGT CAAGCTGGCGCGCACCGCCGGCGCGGTGCCCTGGGATCTGGCCG ACCTCCCGGGCGTCACCGAGCAGACTTTCTTTGTGGGGGTGGAT TTGGGGCACGATCACACCCACCAACAGAGCCTGCCCGCCTTTAC CCTCCACGACCACCGGGGCCGACCCCTGCAGAGCTGGACTCCGC CTCGCCGCACCAACAACGAACGGCTCAGCCTGGCGGAGCTGAAA AAAGGGTTGCACCGCCTGTTGGCCCGCCGCTCAGTGGATCAGGT GATCGTGCACCGCGACGGCCGCTTTTTGGCGGGTGAGGTGGATG ATTTCACCCTTGCGCTGCACGACCTGGGCATTCCCCAGTTTTCG CTGCTGGCCATCAAGAAAAGCAACCACAGTGTGGCCGTGCAGGC AGAAGAGGGCAGCGTATTGTCTCTGGATGAGCGGCGCTGCCTGC TGGTCACCAACACCCAGGCGGCCCTGCCCCGACCCACGGAGCTT GAGCTTGTTCACAGTGACCGCCTCAGCCTAGCGACACTCACCGA GCAGGTGTTTTGGCTCACCCGCGTGTTTATGAATAACGCCCAGC ACGCCGGAAGTGACCCGGCCACCATCGAGTGGGCCAATGGAATC GCGCGCACAGGGCAGCGCGTGCCCCTCGCCGGTTGGAGGCTCTG A
20	37 *Chroococcidiopsis thermalis*	ATGAATAATGTTATGCAAGAATTTCCAGTTGCTTCATTTCCAAC TTTTTTAAGTGAAATTTCACTTCTAGATATTACTCCGAAAAATT TCATTTGTTTTCGATTAACTCCAGAAATCGAACGCAAAACTGGT AATAGCTTTAGTTGGCGATTCAGTCAAAAATTTCCTGATGCAGT TGTTATTTGGCACAATAAATTTTTCTGGGTTTTAGCCAAACCCA ATCGACCAATGCCAAGCCAAGAACAGTGGCGCGAGAAACTGCTA GAAATTTGCGAAGAATTAAAAAAAGATATTGGCGATCGCACTTA TGCAATACAATGGGTAAGCCAGCCACAGATTACACCTGAGATTC TTTCACAGTTAGCAGTGAGAGTATTAAAAATAAATTGCCGTTTT TCATCTCCATCTGTAATATCAGTAAATCAAGTAGAAGTCAAACG AGAGATTGATTTTTGGGCAGAAACGATTGAAATTCAAACTCAGA TTCAGCCAGCTTTGACAATTACCGTCCATAGTAGTTTCTTCTAT CAAAGGCATCTAGAAGAATTTTACAATAATCATCCCTATCGGCA GAATCCAGAGCAACTGTTAATTGGCTTAAAAGTACGAGATATCG AACGTAATAGCTTTGCAACAATAACTGATATTGTAGGTACTATT GCCGACCACAGACAAAAACTACTTGAGGATGCAACTGGCGCAAT TAGTAAACAAGCATTGATAGAAGCACCGGAAGAACAGCCAGTCG

		TTGCCGTACAGTTTGGTAAAAATCAACAACCTTTTTATTATGCA ATGGCAGCTTTACGTCCTTGCATAACAGCGGAAACTGCTAGAAA ATTCGATGTAGATTATGGAAAATTACTGTCTGCAACTAAAATTC CTTATTTAGAGCGAAAAGAACTTTTAGCATTGTACAAAAAAGAA GCTGGACAAAGTTTAGCTACTTATGGATTTCAACTGAAGATTAG TATAAATAGCCGCAGACATCCTGAATTATTCTTCTCTCCGTCAG TTAAATTATCAGAAACAAAACTGGTGTTTGGAAAAAATCAAATT GGCGTTCAAGGTCAAATTTTATCTGGTTTATCTAAAGGTGGCGT GTATCGCCGTCATGAAGATTTTAGCGATTTGTCAAGACCAATTC GGATTGCTGCATTGAAACTTTGCGATTATCCAGCAAATTCTTTT CTGCAAGAAACGCGACAGAGACTCAAACGCTATGGTTTTGAAAC TCTTCTTCCTGTTGAAAATAAAAAGACATTATTGGTAGATGATT TATCTGGAGTTGAAGCGAGAGCCAAAGCTGAAGAAGCTGTCGAT GAATTGATGGTAAATCATCCCGATATAGTTTTGACATTTTTACC AACTAGCGATCGCCACAGCGATAATACAGAAGGAGGCAGCTTAT ATTCTTGGATTTACTCACGCTTGCTCAGACGTGGAATTGCCAGC CAAGTTATTTACGAAGATACTTTGAAAAGCGTTGAAGCTAAATA TTTATTAAATCAGGTGATTCCAGGTATTCTAGCCAAGCTAGGCA ACTTGCCTTTTGTGTTAGCTGAACCATTAGGAATTGCCGACTAT TTCATCGGCTTAGATATTTCCAGAAGCGCCAAGAAGAGAGGTTC TGGAACTATGAATGCTTGCGCTAGCGTGCGTCTATACGGTCGCA AAGGAGAATTTATTCGCTATCGATTAGAAGATGCTTTAATTGAA GGGGAAGAAATTCCCCAAAGAATTTTAGAAAGCTTTCTTCCCGC AGCTCAACTGAAAGGTAAAGTCGTACTAATTTACCGCGATGGAC GCTTTTGTGGTGATGAAGTGCAGCATTTAAAAGAAAGAGCGAAA GCAATTGGTTCGGAGTTTATTTTAGTCGAGTGCTATAAATCTGG GATTCCTCGGCTCTACAACTGGGAAGAAGAAGTTATTAAAGCAC CAACGCTAGGATTAGCACTGCGCTTATCGGCGCGGGAAGTTATT TTAGTCACGACAGAACTAAATTCGGCGAAAATCGGTTTGCCTCT GCCTCTACGCTTGAGAATTCATGAAGCAGGACATCAGGTATCGT TAGAAAGTTTAGTAGAGGCGACTTTGAAACTTACCTTACTACAT CATGGTTCGCTGAACGAGCCGCGCTTACCAATACCGCTTTTTGG TTCTGACCGTATGGCTTATCGACGGTTGCAAGGCATTTATCCAG GTTTGCTAGAAGGCGATCGCCAATTTTGGTTATAA
21	27 *Thermosynechococcus elongates*	ATGCCCACCCAATTCCAAGAAGTTGAAGTTATACTCAATCGTTT TTTTGTAAAAAAATTGAGTCGACCTGATCTTACATTTCATGAAT ACCAATGCCAATTTACTCAAGTGCCAGAACAAGGTAGCGAGCAA AAAGCTATTTCCAGTGTTTGCTACAAACTAGGAGTCACTGCTGT TCGACTAGGGAGCTGCATTATTACAAGGGAGCCTATTGACCCTG AGAGAATGCGAACTAAGGATTGGCAGTTACAGCTAATAGGATGT AGAGAACTGAGCTGTCAAAATTATCGTGAAAGGCAGGCTCTGGA AACCTTTGAAAGAAAAATTCTAGAGGAAAAGTTAAAAGAAACAT TTAAGAAAACTATAATTGAAAAAGACTATGAATTAGGGTTGATT TGGTGGATTTCTGGCGAAGAAGGGTTAGAAAAAACAGGTCATGG CTGGGAAGTACATAGAGGCAGACAAATTGATCTCAAAATAGAAA CAGATGAAAAATTATACTTAGAAATTGATATTCATCATCGATTT TATACTCCATTCAAACTTGAGTGGTGGCTCAGTGAATATCCTAA TATCCAAATCAAGTACGTGAGGAATACCTACAAAGATAAGAAAA AGTGGATCCTTGAAAATTTTGCGGACAAAAGTCCAAATGAAATA

		CAAATAGAAGCCCTAGGAATAAGCCTAGCAGAGTATCACCGTCA AGAGGGAGCAACTCAACAAGAAATAGATGAGTCCCGAGTTGTAA TTGTTAAGAAAATCAGCGACTACAAGGCTAAGCCAGTATATCAT TTGTCTCAAAGACTATCACCAATTCTCACAATGGAAACGTTAGC GCAAATTGCTGAACAAGGAAGAGAAAAAAAGGAAATTCAAGGTG TTTTTGACTACATCAGGAAAAACATTGGTACACGTTTGCAAGAA TCACAAAAAATAGCTCAGGTCATTTTCAAAAACGTTTATAATCT CAGTAGTCAGCCAGAGATAATGAAAGTTAATGGTTTTGTCATGC CTCGTGCAAAACTATTAGCTCGAAACAATAAAGAAGTCAATCAA ACAGCTAGAATTAAATCCTTTGGCTGTGCCAAGATTGGCGAAAC AAAATTCGGCTGCCTAAATTTATTTGATAATAAACCCGAGTATC CAGAGGAAGTACACAAATGTTTACTGGCAATAGCAAGAAGTAGT GGGGTACAGATAAAAATAGACTCCTACTTTACAGGAAGTGACTA TCCAAAAGATGATTTAGCTCAACAACGATTTTGGCAACAATGGG CTGCTCAAGGAATTAAAACTGTTTTAGTAGTGATGCCTTGGTCC CCCCATGAGGAGAAAACAAGGCTACGAATTCAGGCATTAAAGGC GGGAATCGCTACTCAGTTCATGATACCCACACCCCAGGATAATC CCTACAAAGCTCTCAATGTTGCCTTGGGACTGTTGTGTAAAGCT AAGTGGCAGCCTGTCTATCTAAAACCATTAGATGATCCTCAGGC TGCTGACTTAATTATCGGTTTTGACACCAGTACAAACCGAAGGC TGTACTATGGTACGTCTGCTTTTGCAATTTTAGCCAATGGTCAA AGCTTGGGTTGGGAGTTGCCAGATATACAACGGGGCGAAACTTT CTCTGGTCAATCCATTTGGCAAGTTGTGTCTAAGCTAGTGCTGA AATTCCAAGACAACTATGATTCTTACCCTAAGAAGATCCTACTG ATGCGGGACGGGCTTGTTCAAGACGGGGAGTTTGAACAAACAAT CAGAGAACTAACTCACCAGGGTATCGATGTCGATATCTTAAGTG TCCGTAAGAGTGGCTCGGGGAGGATGGGACGTGAATTGACTTCA GGCAATACCGCAATAACCTATGATGATGCTGAAGTGGGCACTGT AATTTTTTATTCTGCAACTGACTCATTTATATTACAAACCACTG AGGTCATCAAGACGAAAACTGGCCCCCTTGGTAGTGCTAGACCT CTGCGGGTTGTGCGTCACTACGGCAATACACCTTTAGAGCTACT GGCTTTGCAGACCTATCACTTGACTCAATTGCATCCTGCCAGCG GATTTCGCTCCTGCCGGCTGCCTTGGGTACTGCACTTGGCTGAT CGGAGTAGTAAGGAGTTTCAGCGGATAGGTCAGATCAGCCTTTT GCAGAATGTTGATCGAGAAAAGTTAATTGCTGTTTGA

[0239] In some embodiments, the Ago is codon optimized for expression in particular cells, such as eukaryotic cells. In some embodiments, a polynucleotide encoding the Ago is codon optimized for expression in particular cells, such as eukaryotic cells. This type of optimization can entail the mutation of foreign-derived (*e.g.*, recombinant) nucleic acids to mimic the codon preferences of the intended host organism or cell while encoding the same protein.

[0240] The Ago may bind and/or modify (*e.g.*, cleave, methylate, demethylate, etc.) a target nucleic acid and/or a polypeptide associated with target nucleic acid. As described in further detail below, in some cases, a subject nuclease has enzymatic activity that modifies target nucleic

acid. Enzymatic activity may refer to nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity or glycosylase activity. In other cases, a subject Ago may have enzymatic activity that modifies a polypeptide associated with a target nucleic acid.

[0241] In some embodiments, in addition to or as a substitute for nucleic acid-cleaving activity, the compositions, fusion polypeptides, methods, and systems described herein have a "pasting" function. Accordingly, In some embodiments, the compositions, fusion polypeptides, methods, and systems can be used to insert a nucleic acid into a target sequence in addition to or instead of cleaving the target nucleic acid. Such exemplary nucleic acid-insertion activities include, but are not limited to, integrase, flippase, transposase, and recombinase activity. Thus, exemplary polypeptides having such function (nucleic acid-insertion polypeptides) include integrases, recombinases, and flippases. These nucleic acid-insertion polypeptides can, for example, insert a nucleic acid sequence at a site that has been cleaved by a polypeptide of the present disclosure.

[0242] In some cases, the Ago system comprises a nuclear localization sequence (NLS). In some embodiments, the nuclear localization sequence is from SV40. In some embodiments, the NLS is from at least one of: SV40, nucleoplasmin, importin alpha, C-myc, EGL-13, TUS, BORG, hnRNPA1, Mata2, or PY-NLS. In some embodiments, the NLS is on a C-terminus or an N-terminus of a nuclease polypeptide or nucleic acid. In some cases, the Ago system may contain from about 1 to about 10 NLS sequences. In some embodiments, the Ago system contains 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10 NLS sequences. The Ago system may contain a SV40 and Nucleoplasmin NLS sequence. In some cases, an NLS is from Simian Vacuolating Virus 40.

[0243] In some cases, the system comprises an Ago polypeptide or Ago polypeptide fragment, and, optionally, an Ago associated protein, that performs a genomic alternation with favorable thermodynamics. In some embodiments, the genomic alteration is exothermic. In some embodiments, the genomic alteration is endothermic. In some cases, a genomic alteration utilizing the disclosed system is energetically favorable over alternate gene editing systems. In some embodiments, the present disclosure provides an *ex vivo* system comprising an Ago polypeptide or fragment and a guide nucleic acid, wherein the guide nucleic acid binds to a predetermined gene or to a nucleic acid sequence adjacent to the predetermined gene, wherein the Ago polypeptide or fragment thereof is capable of introducing a double strand break in the

predetermined gene, wherein the Ago polypeptide or fragment comprises a nucleic acid unwinding sequence that lowers the energetic requirement for introducing the double strand break in comparison to introducing a double strand break with a comparable Ago polypeptide or fragment without the nucleic acid unwinding sequence, and the *ex vivo* system introduces the double strand break at a range of temperatures from 19 °C to 40 °C. Without wishing to be bound by theory, the nucleic acid unwinding sequence can overcome the energetic barrier that prevents Argonaute proteins without such sequences from inducing single- or double-stranded nucleic acid breaks because the nucleic acid unwinding polypeptide exposes a nucleic acid sequence such that the RHDC polypeptide can cleave in the exposed region. The Ago polypeptide or Ago polypeptide fragment system can be more thermodynamically favorable, as measured by a biochemical system, for example by providing a finite amount of ATP into the reaction and measuring an amount of gene editing before, during, and after the genomic alteration has occurred. In some cases, the disclosed editing system utilizing the Ago polypeptide or Ago polypeptide fragment can reduce an energetic requirement by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 40%, 50%, or up to about 60% as compared to a system that does not employ the Ago polypeptide or Ago polypeptide fragment. In some cases, the disclosed editing system utilizing the Ago polypeptide or Ago polypeptide fragment can reduce an immune response to the Ago polypeptide or Ago polypeptide fragment by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 40%, 50%, or up to about 60% as compared to a system that does not employ the disclosed Ago polypeptide or Ago polypeptide fragment. In some cases, the Ago polypeptide or Ago polypeptide fragment can be harvested from bacteria that are endogenously present in the human body to prevent eliciting an immune response.

[0244] In some cases, the Ago system comprises a nucleic acid unwinding polypeptide or a polynucleic acid encoding the same. For example, the system can comprise the Ago and the nucleic acid unwinding polypeptide individually or as a fused polypeptide.

(a) Clostridia Argonautes

[0245] In some cases, the Ago (or variant or functional fragment thereof) does not naturally occur in a bacterium (e.g., a bacterium of class Clostridia, or genus Clostridium); rather it is e altered or engineered based on a naturally-occurring polypeptide or protein of that bacterium (e.g., a bacterium of class Clostridia, or genus Clostridium).

[0246] In some cases, the Ago (or a functional fragment thereof) is derived from phylum *Firmicutes*.

[0247] In some embodiments, the Ago (or variant or functional fragment thereof) described herein, is derived from a bacterium of the class *Clostridia*. In some cases, the Ago does not naturally occur in a *Clostridia* bacterium; rather it is altered or engineered based on a naturally-occurring polypeptide or protein of that *Clostridia* bacterium.

[0248] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the class *Clostridia*.

[0249] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the order: *Candidatus Comantemales, Clostridiales, Halanaerobiales, Natranaerobiales,* or *Thermoanaerobacterales, or Negativicutes*.

[0250] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the family: *Caldicoprobacteraceae, Christensenellaceae, Clostridiaceae, Defluviitaleaceae, Eubacteriaceae, Graciibacteraceae, Heliobacteriaceae, Lachnospiraceae, Oscillospiraceae, Peptococcaceae, Peptostreptococcaceae, Ruminococcaceae*, or *Syntrophomonadaceae*.

[0251] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the family: *Halanaerobiaceae or Halobacteroidaceae.*

[0252] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the family: *Natranaerobiaceae.*

[0253] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the Family: *Thermoanaerobacteraceae or Thermodesulfobiaceae*.

[0254] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from the genus: *Clostridium, Acetanaerobacterium, Acetivibrio, Acidaminobacter, Alkaliphilus, Anaerobacter, Anaerostipes, Anaerotruncus, Anoxynatronum, Bryantella, Butyricicoccus, Caldanaerocella, Caldisalinibacter, Caloramator, Caloranaerobacter, Caminicella, Candidatus Arthromitus, Cellulosibacter, Coprobacillus, Crassaminicella, Dorea, Ethanologenbacterium, Faecalibacterium, Garciella, Guggenheimella, Hespellia, Linmingia, Natronincola, Oxobacter, Parasporobacterium, Sarcina, Soehngenia, Sporobacter, Subdoligranulum, Tepidibacter, Tepidimicrobium, Thermobrachium, Thermohalobacter, or Tindallia.*

[0255] In some cases, the Ago (or variant or functional fragment thereof) is derived from the genus *Clostridium*.

[0256] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a species of *Anaerococcus prevotii, Butyrivibrio proteoclasticus, Clostridiales genomosp., Clostridium acidurici, Clostridium cellulolyticum, Clostridium difficile, Clostridium lentocellum, Clostridium leptum, Clostridium phytofermentans, Clostridium sticklandii, Clostridium symbiosum, Clostridium thermocellum, Ethanoligenens harbinense, Eubacterium rectale, Filifactor alocis, Finegoldia magna, Peptostreptococcus anaerobius, Roseburia hominis, Ruminococcus albus, Candidatus Arthromitus, Clostridium acetobutylicum, Clostridium botulinum, Clostridium perfringens,* or *Clostridium tetani*.

[0257] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a species of *Clostridium absonum, Clostridium aceticum, Clostridium acetireducens, Clostridium acetobutylicum, Clostridium acidisoli, Clostridium aciditolerans, Clostridium acidurici, Clostridium aerotolerans, Clostridium aestuarii, Clostridium akagii, Clostridium aldenense, Clostridium aldrichii, Clostridium algidicarnis, Clostridium algidixylanolyticum, Clostridium algifaecis, Clostridium algoriphilum, Clostridium alkalicellulosi, Clostridium amazonense, Clostridium aminophilum, Clostridium aminovalericum, Clostridium amygdalinum, Clostridium amylolyticum, Clostridium arbusti, Clostridium arcticum, Clostridium argentinense, Clostridium asparagiforme, Clostridium aurantibutyricum, Clostridium baratii, Clostridium barkeri, Clostridium bartlettii, Clostridium beijerinckii, Clostridium bifermentans, Clostridium bolteae, Clostridium bornimense, Clostridium botulinum, Clostridium bowmanii, Clostridium bryantii, Clostridium budayi, Clostridium butyricum, Clostridium cadaveris, Clostridium caenicola, Clostridium caminithermale, Clostridium carboxidivorans, Clostridium carnis, Clostridium cavendishii, Clostridium celatum, Clostridium celerecrescens, Clostridium cellobioparum, Clostridium cellulofermentans, Clostridium cellulolyticum, Clostridium cellulosi, Clostridium cellulovorans, Clostridium chartatabidum, Clostridium chauvoei, Clostridium chromiireducens, Clostridium citroniae, Clostridium clariflavum, Clostridium clostridioforme, Clostridium coccoides, Clostridium cochlearium, Clostridium cocleatum, Clostridium colicanis, Clostridium colinum, Clostridium collagenovorans, Clostridium combesii, Clostridium cylindrosporum, Clostridium difficile, Clostridium diolis, Clostridium disporicum, Clostridium drakei, Clostridium durum, Clostridium estertheticum, Clostridium estertheticum* subsp. *Estertheticum, Clostridium estertheticum* subsp. *Laramiense, Clostridium fallax, Clostridium felsineum,*

Clostridium fervidum, Clostridium fimetarium, Clostridium formicaceticum,
Clostridium frigidicarnis, Clostridium frigoris, Clostridium ganghwense, Clostridium gasigenes,
Clostridium ghonii, Clostridium glycolicum, Clostridium glycyrrhizinilyticum,
Clostridium grantii, Clostridium guangxiense, Clostridium haemolyticum,
Clostridium halophilum, Clostridium hastiforme, Clostridium hathewayi,
Clostridium herbivorans, Clostridium hiranonis, Clostridium histolyticum,
Clostridium homopropionicum, Clostridium huakuii, Clostridium hungatei,
Clostridium hydrogeniformans, Clostridium hydroxybenzoicum, Clostridium hylemonae,
Clostridium indolis, Clostridium innocuum, Clostridium intestinale, Clostridium irregulare,
Clostridium isatidis, Clostridium jeddahense, Clostridium jejuense, Clostridium josui,
Clostridium kluyveri, Clostridium lactatifermentans, Clostridium lacusfryxellense,
Clostridium laramiense, Clostridium lavalense, Clostridium lentocellum,
Clostridium lentoputrescens, Clostridium leptum, Clostridium limosum, Clostridium liquoris,
Clostridium litorale, Clostridium lituseburense, Clostridium ljungdahlii, Clostridium lortetii,
Clostridium lundense, Clostridium luticellarii, Clostridium magnum,
Clostridium malenominatum, Clostridium mangenotii, Clostridium maximum,
Clostridium mayombei, Clostridium methoxybenzovorans, Clostridium methylpentosum,
Clostridium moniliforme, Clostridium neonatale, Clostridium neopropionicum,
Clostridium neuense, Clostridium nexile, Clostridium nitritogenes, Clostridium nitrophenolicum,
Clostridium novyi, Clostridium oceanicum, Clostridium orbiscindens, Clostridium oroticum,
Clostridium oryzae, Clostridium oxalicum, Clostridium pabulibutyricum,
Clostridium papyrosolvens, Clostridium paradoxum, Clostridium paraperfringens,
Clostridium paraputrificum, Clostridium pascui, Clostridium pasteurianum,
Clostridium peptidivorans, Clostridium perenne, Clostridium perfringens, Clostridium pfennigii,
Clostridium phytofermentans, Clostridium piliforme, Clostridium polyendosporum,
Clostridium polysaccharolyticum, Clostridium populeti, Clostridium propionicum,
Clostridium proteoclasticum, Clostridium proteolyticum, Clostridium psychrophilum,
Clostridium punense, Clostridium puniceum, Clostridium purinilyticum,
Clostridium putrefaciens, Clostridium putrificum, Clostridium quercicolum, Clostridium quinii,
Clostridium ramosum, Clostridium rectum, Clostridium roseum, Clostridium saccharobutylicum,
Clostridium saccharogumia, Clostridium saccharolyticum,
Clostridium saccharoperbutylacetonicum, Clostridium sardiniense, Clostridium sartagoforme,

Clostridium saudiense, Clostridium scatologenes, Clostridium schirmacherense, Clostridium scindens, *Clostridium senegalense, Clostridium septicum, Clostridium sordellii, Clostridium sphenoides, Clostridium spiroforme, Clostridium sporogenes, Clostridium sporosphaeroides*, *Clostridium stercorarium*, *Clostridium stercorarium* subsp. *Leptospartum, Clostridium stercorarium* subsp. *Stercorarium, Clostridium stercorarium* subsp. *Thermolacticum, Clostridium sticklandii, Clostridium straminisolvens, Clostridium subterminale, Clostridium sufflavum, Clostridium sulfidigenes, Clostridium swellfunianum, Clostridium symbiosum, Clostridium tarantellae, Clostridium tagluense, Clostridium tepidiprofundi, Clostridium tepidum, Clostridium termitidis, Clostridium tertium, Clostridium tetani, Clostridium tetanomorphum, Clostridium thermaceticum, Clostridium thermautotrophicum, Clostridium thermoalcaliphilum, Clostridium thermobutyricum, Clostridium thermocellum, Clostridium thermocopriae, Clostridium thermohydrosulfuricum, Clostridium thermolacticum, Clostridium thermopalmarium, Clostridium thermopapyrolyticum, Clostridium thermosaccharolyticum, Clostridium thermosuccinogenes, Clostridium thermosulfurigenes, Clostridium thiosulfatireducens, Clostridium tyrobutyricum, Clostridium uliginosum, Clostridium ultunense, Clostridium ventriculi*, *Clostridium villosum, Clostridium vincentii, Clostridium viride, Clostridium vulturis, and Clostridium xylanolyticum,* and *Clostridium xylanovorans*.

[0258] In some embodiments, the Ago or variant or functional fragment thereof is derived from a species of *Clostridium perfringens*, *Clostridium butyricum, or Clostridium sardiniense*.

[0259] In some embodiments, the Ago or variant or functional fragment thereof is derived from a species of *Clostridiales bacterium NK3B98, Geobacillus sp. FW23, [Clostridium] citroniae WAL-19142, Clostridium disporicum, Burkholderia vietnamiensis, Bacteroides fragilis str. 3397 T14, Leptolyngbya sp. 'hensonii', Acidobacterium capsulatum ATCC 51196, Clostridium perfringens WAL-14572, Geobacillus kaustophilus GBlys, Clostridium saudiense, Methylomicrobium buryatense 5G, Enterobacter kobei, or Deinococcus sp. RL.*

[0260] In some embodiments, the Ago or variant or functional fragment thereof is derived from a species *C. absonum, C. aerotolerans, C. aminobutyricum, C. caliptrosporurn, C. celatum, C. colinum, C. corinoforum, C. durum, C. favososporum, C. felsineum, C. jilarnentosum, C. formicoaceticum, C. glycolicum, C. halophilum, C. hastiforme, C. hornopropionicurn, C. intestinalis, C. kainantoi, C. lentocellum, C. litorale, C. longisporum, C. magnum, C.*

neopropionicum, C. oxalicum, C. pfennigii, C. polysaccharolyticum, C. propionicum, C. quinii, C. rectum, C. tetani, C. thermoamylolyticum, and *C. xylanolyticum.*

[0261] In some embodiments the clostridia Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NOs: 1-3. In some embodiments the clostridia Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NOs: 134-136. In some embodiments the clostridia Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NOs: 11-14. In some embodiments the clostridia Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NOs: 137-139

[0262] In some embodiments, the Ago comprises an amino acid sequence 100% identical to SEQ ID NO: 1. In some embodiments, the Ago comprises an amino acid sequence 100% identical to SEQ ID NO: 1, except there is a non-lysine amino acid residue at one or more of (e.g., 1, 2, 3, 4, or 5) positions 479, 522, 563, 581, 642 of SEQ ID NO: 1.

(b) Clostridia Argonaute 69 Homologues

[0263] In some embodiments, the Argonaute is a homologue of Ago69 (SEQ ID NO: 1). In some embodiments, the Ago69 homologue comprises an amino acid sequence of an Ago69 homologue described in Table 14. In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to an Ago69 homologue described in Table 14. In some embodiments, the Ago69 homologue comprises a nucleic acid sequence of an Ago69 homologue described in Table 15. In some embodiments, the Ago69 homologue comprises a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to an Ago69 homologue described in Table 15. In some embodiments, the Ago69 homologue is HG2, HG4, or HG5.

[0264] In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to an Ago69 homologue HG2. HG2 has 78.3% pairwise sequence identity with Ago69. In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to SEQ ID NO:

134. In some embodiments, the Ago69 homologue comprises a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to SEQ ID NO: 137.

[0265] In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to an Ago69 homologue HG4. HG4 has 39.9% pairwise sequence identity with Ago69. In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to SEQ ID NO: 135. In some embodiments, the Ago69 homologue comprises a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to SEQ ID NO: 138.

[0266] In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to an Ago69 homologue HG5. HG5 has 38.5% pairwise sequence identity with Ago69. In some embodiments, the Ago69 homologue comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to SEQ ID NO: 136. In some embodiments, the Ago69 homologue comprises a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 100% sequence identity to SEQ ID NO: 139.

Table 14. Amino Acid Sequence of Exemplary Ago69 Homologues

Ago69 Homologue	Amino Acid Sequence	SEQ ID NO
HG2 *Clostridium butyricum* Underlined: PIWI Domain	MNNLTFEAFEGIGQLNELNFYKYRLIGKGQIDNVHQA IWSVKYKLQANNFFKPVFVKGEILYSLDELKVIPEFE NVEVILDGNIILSISENTDIYKDVIVFYINNALKNIK DITNYRKYITKNTDEIICKSILTTNLKYQYMKSEKGF KLQRKFKISPVVFRNGKVILYLNCSSDFSTDKSIYEM LNDGLGVVGLQVKNKWTNANGNIFIEKVLDNTISDPG TSGKLGQSLIDYYINGNQKYRVEKFTDEDKNAKVIQA KIKNKTYNYIPQALTPVITREYLSHTDKKFSKQIENV IKMDMNYRYQTLKSFVEDIGVIKELNNLHFKNQYYTN FDFMGFESGVLEEPVLMGANGKIKDKKQIFINGFFKN PKENVKFGVLYPEGCMENAQSIARSILDFATAGKYNK QENKYISKNLMNIGFKPSECIFESYKLGDITEYKATA RKLKEHEKVGFVIAVIPDMNELEVENPYNPFKKVWAK LNIPSQMITLKTTEKFKNIVDKSGLYYLHNIALNILG KIGGIPWIIKDMPGNIDCFIGLDVGTREKGIHFPACS VLFDKYGKLINYYKPTIPQSGEKIAETILQEIFDNVL ISYKEENGEYPKNIVIHRDGFSRENIDWYKEYFDKKG	134

	IKFNIIEVKKNIPVKIAKVVGSNICNPIKGSYVLKND KAFIVTTDIKDGVASPNPLKIEKTYGDVEMKSILEQI YSLSQIHVGSTKSLRLPITTGYADKICKAIEYIPQGV VDNRLFFL	
HG4 *Clostridium saragoforme* Underlined: PIWI Domain	MKEFNVITEFKNGINSKSIEIYIYKMMVRDFEKRHNE NYDVVKELINLNNNSTIVFYEQYIASFKEIEKWGNEQ YINVEKRAINLESNEKKILERLLLKEIKNNIDNNKYK VVKDSIYINKPVYNEKGIKIDRYFNLDINVESNGDII IGFDISHNFEYINTLEYEIKNNNIKIGDRVKDYFYNL TYEYVGIAPFTISEENEYMGCSIVDYYENKNQSYIVN KLPKDMKAILVKNNKNSIFPYIPSRLKKVCRFENLPQ NVLRDFNTRVKQKTNEKMQFMVDEVINIVKNSEHIDV KKKNMMCDNIGYKIEDLQQPDLLFGNARAQRYPLYGL KNFGVYENKRIEIKYFIDPILAKSKMNLEKISKFCDE LEQFSSKLGVGLNRVKLNNIVNFKEIRMDNEDIFSYE IRKIVSNYNETTIVILSEENLNKYYNIIKKTFSGGNE VPTQCIGFNTLSYTEKNKDSIFLNILLGVYAKSGIQP WILNEKLNSDCFIGLDVSRENKVNKAGVIQVVGKDGR VLKTKVISSSQSGEKIKLETLREIVFEAINSYENTYR CKPKHITFHRDGINREELENLKNTMTNLGVEFDYIEI TKGINRRIATISEGEEWKTIMGRCYYKDNSAYVCTTK PYEGIGMAKPIRIRRVFGTLDIEKIVEDAYKLTFMHV GAINKIRLPITTYYADLSSTYGNRDLIPTNIDTNCLY FI	135
HG5 *Clostridium sp. 1-1-41A1FAA*	MVGLDREFNVITEFKNELKPEDIKIFLYSMPIKDINE RHSENYAIVQELKKINENPNIVFNEYIIASFNPIINW GKYKDIDVKPDNRNINLDNHTERKILERLLLCDIKNN INNNTTWEQQNKYEIRGNANPAVYLRKPIYLNDNLII RRKLNFDVNIDKKDIIIGFFLNHEFEYQKTLDEEIKC GNIQKGDKVKDFYNNITYEFLEMAPFSISQENKYMRS SIIEYYLNKGQSYIISGLDKNTKAVLVKNKEGSIFPY IPNRLKKICVFENLGNRQIIEGNKYIKMNPSQNMSES IKLAEDILKNSKYVKFNKANMIVEKIGYKKDIVKRPA LKFGKNESNFSAMYGLNKSGSYEQKNIKIDYFIDPKI LNNKRDYQIVYSFLNDIISKSKDLGVEINTDKSYINL TPINIKNENEFELNVMEIIKNYNNPVLVILEKENIDK YYETLKKIFGGRNSIATQFVDLDTIKRCDPKIDNKRG KESIFLNILLGIYCKSGIQPWVLANGLSADCYIGLDV CRENNMSTVGLIQVIGKDGRVLKSKTISSHQSGEKIQ INILKDIIFEAKQAYKNTYNKKLEHIVFHRDGINRED IDLLKEITNSLEIKFDYVEVTKNINRRMAMLEKSDEN YNHRDKENKKWITEIGMCLKKENEAYLITTNPSENMG MARPLRIKKVYGNQNMDDIVKDIYKLSFMHIGSIMKS RLPITTHYADLSSIYSHRELMPKSVDNNILHFI	136

Table 15. Nucleic Acid Sequence of Exemplary Ago69 Homologues

Ago69 Homologue	Amino Acid Sequence	SEQ ID NO

HG2	TGTACAAGCTTGCCACCATGGGTCCGAAGAAGAAACG CAAGGTCGAGGACCCAAAGAAGAAGCGTAAGGTTGGA TCGGGTTCTATGAATAATCTGACCTTCGAGGCCTTCG AGGGTATCGGACAATTGAACGAGTTAAACTTCTATAA GTACCGCCTCATTGGTAAGGGCCAAATCGACAATGTC CACCAGGCCATCTGGTCAGTCAAGTACAAACTTCAAG CGAATAATTTCTTCAAGCCGGTTTTCGTCAAGGGCGA AATTCTGTACTCACTTGACGAGCTGAAAGTCATCCCG GAATTCGAGAATGTCGAGGTTATTCTTGACGGGAACA TTATCCTGAGCATTAGCGAGAACACCGACATTTACAA GGATGTGATCGTGTTTTATATCAATAACGCGTTGAAG AACATCAAGGACATCACCAACTACCGTAAGTATATCA CTAAGAACACGGATGAAATCATTTGCAAGAGTATTTT AACGACGAATCTCAAGTATCAATATATGAAGTCAGAG AAAGGGTTCAAGTTACAGCGCAAGTTTAAGATCTCCC CGGTGGTATTCCGTAATGGGAAGGTCATCTTGTACCT TAATTGCAGTAGCGACTTCAGCACAGACAAATCCATC TACGAAATGTTAAATGATGGACTCGGTGTTGTGGGCC TGCAAGTGAAGAATAAGTGGACTAATGCGAATGGCAA TATCTTTATTGAAAAGGTGCTCGACAATACCATCTCC GATCCCGGCACGAGTGGAAAGCTGGGGCAGTCCCTGA TCGACTACTACATCAATGGGAATCAAAAGTACCGTGT AGAGAAATTTACCGACGAGGACAAGAATGCAAAGGTT ATCCAGGCCAAAATCAAGAATAAAACATACAACTACA TCCCGCAAGCTCTCACCCCCGTAATTACGCGCGAGTA TCTGAGTCATACCGATAAGAAGTTTAGCAAGCAAATC GAGAATGTGATTAAGATGGATATGAACTACCGCTACC AGACGTTGAAGTCTTTCGTTGAGGACATTGGCGTGAT CAAGGAGTTAAACAATCTGCACTTTAAGAACCAATAT TACACCAATTTTGACTTTATGGGGTTCGAGAGCGGGG TGCTGGAAGAACCTGTCCTGATGGGTGCGAACGGAAA GATCAAGGACAAGAAGCAGATTTTCATCAATGGGTTC TTTAAGAATCCCAAGGAGAACGTAAAATTCGGAGTAC TCTACCCAGAAGGCTGTATGGAGAATGCTCAGAGCAT TGCTCGTTCCATCCTCGACTTCGCTACGGCCGGTAAA TACAATAAGCAAGAGAACAAGTATATTTCGAAGAATT TAATGAACATCGGATTCAAACCTTCTGAGTGTATCTT TGAGTCGTATAAGTTGGGAGACATCACCGAGTATAAG GCGACGGCCCGTAAGCTCAAGGAGCATGAGAAAGTTG GGTTCGTTATCGCAGTGATCCCTGACATGAATGAGCT GGAAGTCGAGAACCCTTATAACCCCTTCAAGAAGGTC TGGGCGAAACTCAATATCCCATCCCAGATGATCACAT TGAAGACCACCGAAAAGTTCAAGAATATCGTCGACAA GTCAGGCTTGTACTACTTACACAATATCGCCCTTAAT ATTCTCGGCAAAATCGGCGGAATCCCGTGGATTATTA AAGACATGCCTGGCAACATCGACTGTTTCATCGGTTT AGACGTCGGCACGCGCGAGAAGGGCATCCACTTCCCG GCATGTTCTGTGTTGTTCGACAAGTACGGAAAGTTAA TCAATTATTACAAGCCGACTATTCCGCAGAGCGGAGA	137

	GAAGATTGCTGAGACAATTTTACAGGAGATCTTCGAC AACGTGTTAATCAGCTACAAAGAGGAAAACGGGGAGT ACCCCAAGAATATCGTTATCCATCGTGATGGCTTCAG CCGTGAGAACATCGATTGGTACAAAGAATACTTCGAT AAGAAGGGTATCAAGTTCAACATTATTGAGGTTAAGA AGAACATTCCCGTAAAGATCGCGAAGGTGGTTGGATC CAATATCTGCAACCCGATCAAGGGCTCTTATGTGCTT AAGAATGATAAGGCATTCATCGTAACCACCGATATCA AAGACGGTGTGGCTTCTCCAAATCCACTTAAAATCGA GAAAACCTATGGTGACGTTGAGATGAAGAGTATTCTG GAGCAGATCTACAGTCTGAGCCAAATTCATGTTGGCT CAACCAAGTCCCTGCGTCTTCCTATCACAACGGGATA TGCCGATAAGATCTGTAAGGCAATTGAATACATTCCG CAAGGAGTCGTAGACAATCGTTTGTTCTTTCTTTAAC GTCTCGAGGCGGCCGC	
HG4	TGTACAAGCTTGCCACCATGGGCCCTAAGAAGAAACG CAAGGTAGAGGATCCGAAGAAGAAGCGTAAGGTAGGT TCCGGTTCGATGAAGGAGTTTAACGTCATCACAGAGT TCAAGAACGGTATTAATTCGAAGAGCATCGAGATCTA TATTTACAAGATGATGGTTCGTGACTTTGAGAAGCGT CACAATGAAAATTATGACGTGGTAAAAGAGCTTATTA ACCTGAACAATAATAGTACGATTGTCTTTTATGAGCA ATATATCGCCTCATTCAAGGAAATCGAGAAGTGGGGT AACGAGCAATACATTAATGTTGAGAAACGCGCAATTA ACCTGGAAAGCAACGAGAAGAAGATTCTTGAACGCCT TCTGTTAAAGGAGATCAAGAACAACATCGATAACAAT AAGTACAAGGTAGTGAAGGATTCGATCTACATCAACA AGCCTGTGTATAACGAAAAGGGTATCAAAATCGACCG CTACTTCAACTTAGACATCAACGTAGAATCAAACGGA GACATCATTATTGGCTTCGATATTAGCCATAATTTCG AGTATATTAACACGTTAGAGTACGAAATCAAGAACAA CAATATCAAGATTGGAGACCGCGTAAAGGATTACTTT TACAACCTTACTTATGAATATGTTGGCATCGCGCCGT TCACTATTTCCGAAGAGAATGAATATATGGGATGTAG CATCGTGGACTACTATGAAAATAAGAACCAGAGCTAC ATCGTGAACAAGTTGCCAAAGGATATGAAGGCAATCT TAGTTAAGAACAATAAGAACAGCATTTTCCCGTACAT CCCTTCACGTCTTAAGAAGGTTTGTCGTTTCGAGAAT CTGCCCCAAAACGTACTCCGTGATTTTAACACGCGCG TCAAGCAGAAAACTAATGAGAAGATGCAATTTATGGT GGACGAGGTTATCAACATTGTAAAGAATAGCGAGCAT ATCGACGTAAAGAAGAAGAACATGATGTGTGACAATA TCGGGTACAAGATTGAGGACCTGCAACAACCTGACCT TTTGTTTGGAAACGCCCGCGCGCAGCGTTACCCACTG TATGGATTGAAGAACTTTGGCGTGTACGAAAACAAGC GCATTGAAATCAAGTACTTTATCGACCCGATTCTCGC CAAGAGCAAGATGAATCTGGAAAAGATCTCCAAGTTC TGTGATGAGCTGGAGCAGTTTAGCTCCAAGTTAGGAG TAGGATTAAATCGCGTAAAACTGAACAATATTGTTAA	138

	CTTCAAGGAGATTCGTATGGACAATGAGGACATCTTC TCCTACGAGATTCGCAAAATTGTGAGCAACTATAATG AGACAACGATCGTGATTCTGTCGGAAGAGAACCTTAA TAAGTATTACAACATCATTAAGAAAACCTTCAGCGGT GGCAACGAGGTTCCGACGCAATGCATTGGTTTCAACA CACTTTCCTACACGGAGAAGAACAAGGACTCAATTTT CTTAAATATTTTACTTGGTGTTTACGCCAAGTCAGGA ATCCAACCGTGGATCCTCAATGAGAAATTGAATTCCG ACTGTTTCATTGGTTTAGATGTCTCCCGTGAGAATAA GGTAAACAAGGCCGGCGTCATTCAAGTTGTCGGAAAA GATGGCCGCGTACTCAAGACCAAGGTCATCAGTTCGA GCCAAAGCGGGGAGAAGATCAAGCTGGAAACGTTACG CGAGATCGTGTTCGAGGCGATTAACTCGTATGAGAAT ACCTACCGCTGTAAACCAAAACACATTACATTCCACC GTGACGGTATTAATCGTGAGGAGCTGGAGAATCTTAA GAATACGATGACCAATCTTGGTGTTGAGTTTGACTAC ATCGAGATCACCAAGGGCATTAACCGCCGCATTGCCA CCATCAGTGAGGGCGAGGAGTGGAAGACTATCATGGG CCGCTGTTATTATAAGGACAATTCTGCCTACGTCTGC ACTACTAAGCCTTATGAGGGAATCGGAATGGCAAAGC CCATTCGCATCCGCCGCGTGTTTGGCACGCTTGATAT CGAGAAAATTGTTGAAGACGCGTATAAACTTACTTTT ATGCATGTAGGCGCGATCAATAAAATTCGTCTTCCAA TTACAACCTATTACGCAGATCTCAGCTCCACTTACGG AAATCGCGACTTAATTCCGACGAATATTGATACCAAT TGCCTCTACTTCATTTAACGTCTCGAGGCGGCCGC	
HG5	TGTACAAGCTTGCCACCATGGGACCGAAGAAGAAGCG TAAGGTCGAGGATCCCAAGAAGAAGCGTAAGGTGGGA TCCGGGTCGATGGTGGGCCTGGACCGCGAATTCAACG TGATCACCGAGTTCAAGAATGAGCTTAAGCCCGAGGA CATCAAGATCTTCTTATACTCGATGCCGATCAAGGAT ATTAATGAGCGCCATTCAGAGAATTATGCAATTGTCC AAGAGCTCAAGAAGATCAACGAGAACCCTAACATTGT ATTTAACGAGTACATCATCGCCAGCTTCAATCCTATT ATTAATTGGGGCAAGTACAAGGACATCGATGTTAAGC CGGACAATCGTAATATTAATCTGGATAACCACACTGA GCGCAAAATCCTGGAGCGTTTATTACTCTGTGACATT AAGAATAACATTAACAATAATACTACCTGGGAGCAAC AGAATAAATACGAGATTCGCGGTAATGCTAACCCGGC AGTATATCTTCGCAAGCCCATCTATCTGAACGATAAC TTGATTATCCGCCGTAAGCTGAATTTTGACGTTAATA TTGACAAGAAAGACATCATTATCGGCTTCTTCCTGAA TCATGAGTTTGAATACCAAAAGACGTTAGACGAGGAA ATCAAGTGTGGCAACATTCAGAAGGGCGACAAAGTGA AGGACTTCTATAATAATATTACATATGAATTCTTGGA GATGGCCCCATTTAGCATCTCACAAGAAAATAAATAC ATGCGCAGTAGCATTATCGAGTATTATTTGAACAAGG GCCAAAGCTACATCATTTCCGGCTTGGATAAGAACAC TAAGGCCGTACTTGTTAAGAACAAAGAGGGCAGTATC	139

	TTCCCCTATATCCCCAATCGCCTTAAGAAAATCTGCG TCTTTGAGAATCTCGGCAACCGCCAGATCATCGAAGG GAATAAGTACATCAAGATGAACCCTAGTCAAAATATG AGTGAAAGCATCAAGTTGGCGGAAGATATCCTTAAGA ATTCGAAGTATGTGAAGTTTAACAAGGCGAACATGAT CGTGGAGAAAATCGGTTACAAGAAGGACATCGTGAAG CGCCCTGCGTTAAAGTTTGGCAAGAATGAGAGCAATT TCAGCGCCATGTACGGCCTTAACAAGAGCGGTAGTTA CGAGCAGAAGAATATTAAGATCGACTATTTCATTGAC CCGAAGATTCTTAATAACAAGCGCGATTACCAGATCG TATACTCCTTCCTCAACGATATTATTAGTAAATCGAA GGACTTGGGAGTCGAGATCAACACGGACAAGAGCTAT ATCAATTTAACTCCAATCAACATTAAGAATGAAAATG AGTTTGAGCTGAACGTCATGGAAATCATTAAGAATTA CAATAACCCAGTACTTGTGATTCTTGAGAAGGAGAAT ATCGACAAGTATTATGAAACCCTTAAGAAGATCTTCG GCGGCCGTAACTCAATCGCAACCCAATTCGTGGATCT GGACACGATCAAGCGCTGCGACCCTAAGATCGATAAC AAGCGTGGAAAGGAATCGATCTTCTTAAACATCCTCT TGGGCATCTACTGTAAGTCGGGTATTCAACCTTGGGT TTTAGCGAATGGTCTGAGCGCTGACTGTTACATTGGC CTCGACGTTTGTCGCGAGAATAATATGTCCACTGTGG GGTTGATTCAAGTCATCGGGAAGGACGGTCGTGTACT CAAAAGTAAGACTATTAGCAGCCATCAAAGTGGGGAA AAGATTCAAATTAATATTTTGAAGGATATCATCTTCG AGGCCAAGCAAGCGTATAAGAATACGTATAACAAGAA GCTGGAACACATCGTTTTCCACCGCGACGGCATCAAT CGTGAAGACATTGACCTTTTGAAGGAGATTACGAACT CCCTGGAGATTAAGTTTGACTACGTCGAGGTAACAAA GAATATTAACCGCCGTATGGCGATGTTAGAGAAAAGC GATGAGAACTATAACCACCGTGACAAGGAGAATAAGA AGTGGATTACGGAAATTGGTATGTGCCTTAAGAAGGA AAATGAGGCCTATCTCATTACCACCAATCCTAGCGAG AATATGGGTATGGCCCGTCCTCTTCGCATTAAGAAGG TGTACGGTAACCAGAACATGGACGACATCGTTAAGGA CATCTACAAGCTGTCCTTCATGCACATTGGTAGCATT ATGAAGTCTCGTCTTCCAATCACAACCCATTACGCGG ATTTATCTTCTATCTACAGCCACCGTGAATTGATGCC TAAGTCCGTCGATAATAACATCCTGCACTTTATTTAA CGTCTCGAGGCGGCCGC	

Table 27: Amino Acid Sequences of Ago69 and Ago69 Homologue PIWI Domains

SEQ ID NO	Ago/Genus Species	Amino Acid Sequence
141	69 PIWI Domain	FVIAIVPNMSDEEIENSYNPFKKIWAELNLPSQMISVKTAEIFA NSRDNTALYYLHNIVLGILGKIGGIPWVVKDMKGDVDCFVGLDV GTREKGIHYPACSVVFDKYGKLINYYKPNIPQNGEKINTEILQE

		IFDKVLISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKNI KFNIIEVKKSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTD IKENLGSPKPLKIEKSYGDIDMLTALSQIYALTQIHVGATKSLR LPITTGYADKICKAIEF
142	HG2 PIWI Domain	FVIAVIPDMNELEVENPYNPFKKVWAKLNIPSQMITLKTTEKFK NIVDKSGLYYLHNIALNILGKIGGIPWIIKDMPGNIDCFIGLDV GTREKGIHFPACSVLFDKYGKLINYYKPTIPQSGEKIAETILQE IFDNVLISYKEENGEYPKNIVIHRDGFSRENIDWYKEYFDKKGI KFNIIEVKKNIPVKIAKVVGSNICNPIKGSYVLKNDKAFIVTTD IKDGVASPNPLKIEKTYGDVEMKSILEQIYSLSQIHVGSTKSLR LPITTGYADKICKAIEYI
143	HG4 PIWI Domain	TTIVILSEENLNKYYNIIKKTFSGGNEVPTQCIGFNTLSYTEKN KDSIFLNILLGVYAKSGIQPWILNEKLNSDCFIGLDVSRENKVN KAGVIQVVGKDGRVLKTKVISSSQSGEKIKLETLREIVFEAINS YENTYRCKPKHITFHRDGINREELENLKNTMTNLGVEFDYIEIT KGINRRIATISEGEEWKTIMGRCYYKDNSAYVCTTKPYEGIGMA KPIRIRRVFGTLDIEKIVEDAYKLTFMHVGAINKIRLPITTYYA DLSSTYGNRDLI

[0267] **FIG. 102** shows a sequence alignment and homology of Ago69, HG2, and HG4. **FIGS. 103A-103D** show a sequence alignment and homology of Ago69, HG2, and HG4 along with an indication of the PAZ, MID, and PIWI domains. **The percent sequence identity across Ago69, HG2, and HG4 is provided in Table 18.**

Table 18. Percent Sequence Identity between Ago69, HG2, and HG4

% Amino Acid Identity Between:	Ago69, HG2, HG4	Ago69, HG2	Ago69, HG4
PIWI Domains	27.7%	72.8%	34.1%
Whole Protein Sequence	19.1%	61.3%	25.2%

[0268] In some embodiments, the Ago polypeptide comprises a PIWI domain. In some embodiments, the Ago polypeptide comprises a PIWI domain that comprises a sequence that has at least 50%, 55%, 60%, 65%, or 70% sequence identity to one of SEQ ID NOS: 141-143. In some embodiments, the Ago polypeptide comprises a PIWI domain that comprises a sequence that has at least 50%, 55%, 60%, 65%, or 70% sequence identity to one of SEQ ID NO: 141. In some embodiments, the Ago polypeptide comprises a PIWI domain that comprises a sequence that has at least 50%, 55%, 60%, 65%, or 70% sequence identity to one of SEQ ID NO: 142. In

some embodiments, the Ago polypeptide comprises a PIWI domain that comprises a sequence that has at least 50%, 55%, 60%, 65%, or 70% sequence identity to one of SEQ ID NO: 143.

(c) Additional Mesophilic Argonautes

[0269] In some cases, the Ago (or variant or functional fragment thereof) does not naturally occur in a bacterium or archael organism; rather it is altered or engineered based on a naturally-occurring polypeptide or protein of that bacterium or archaeal organism.

[0270] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of Phylum planctomycetes, cyanobacteria, or firmicutes.

[0271] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of Phylum planctomycetes. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of class Planctomycetacia. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of order Planctomycetales. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of family Planctomycetaceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus *Rhodopirellula*. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Rhodopirellula bahusiensis, Rhodopirellula baltica, Rhodopirellula caenicola, Rhodopirellula europaea, Rhodopirellula lusitana, Rhodopirellula europaea, Rhodopirellula rosea, Rhodopirellula rubra, or Rhodopirellula sallentina*.

[0272] In some embodiments, an Ago polypeptide as described herein is a mesophilic Ago or a mesothermic Ago. In some embodiments the mesophilic Ago has an amino acid sequence at least 0%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 4. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NO: 15.

[0273] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of Phylum firmicutes. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of class bacilli. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of order bacillales. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of family paenibacillaceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus paenibacillus. In some embodiments, the Ago (or

variant or functional fragment thereof) is derived from a bacterium of species *P. agarexedens, P. agaridevorans, P. alginolyticus, P. alkaliterrae, P. alvei, P. amylolyticus, P. anaericanus, P. antarcticus, P. apiarius, P. assamensis, P. azoreducens, P. azotofixans, P. barcinonensis, P. borealis, P. brasilensis, P. brassicae, P. campinasensis, P. chinjuensis, P. chitinolyticus, P. chondroitinus, P. cineris, P. cookii, P. curdlanolyticus, P. daejeonensis, P. dendritiformis, P. durum, P. ehimensis, P. elgii, P. favisporus, P. glucanolyticus, P. glycanilyticus, P. gordonae, P. graminis, P. granivorans, P. hodogayensis, P. illinoisensis, P. jamilae, P. kobensis, P. koleovorans, P. koreensis, P. kribbensis, P. lactis, P. larvae, P. lautus, P. lentimorbus, P. macerans, P. macquariensis, P. massiliensis, P. mendelii, P. motobuensis, P. naphthalenovorans, P. nematophilus, P. odorifer, P. pabuli, P. peoriae, P. phoenicis, P. phyllosphaerae, P. polymyxa, P. popilliae, P. pulvifaciens, P. rhizosphaerae, P. sanguinis, P. stellifer, P. terrae, P. thiaminolyticus, P. timonensis, P. tylopili, P. turicensis, P. validus, P. vortex, P. vulneris, P. wynnii,* or *P. xylanilyticus*. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *P. odorifer*.

[0274] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 5. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 16.

[0275] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of Phylum proteobacteria. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of class alphaproteobacteria. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of order rhodobacterales. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of family hyphomonadaceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus hyphomonas. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Hyphomonas adhaerens, Hyphomonas hirschiana, Hyphomonas jannaschiana, Hyphomonas johnsonii, Hyphomonas neptunium, Hyphomonas oceanitis, Hyphomonas polymorpha, Hyphomonas rosenbergii, Hyphomonas sp., Hyphomonas sp. AP-32, Hyphomonas sp. BAL52, Hyphomonas sp. DG895, Hyphomonas sp. kbc20, Hyphomonas sp.*

MED623, Hyphomonas sp. MK02, Hyphomonas sp. MK06, Hyphomonas sp. MK08, or Hyphomonas taiwanensis.

[0276] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 6. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 17.

[0277] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of Phylum cyanobacteria. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of class cyanophyceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of order nostocales. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of family rivulariaceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus calothrix. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Calothrix sp. PCC 7103, Calothrix adscendens, Calothrix atricha, Calothrix braunii, Calothrix breviarticulata, Calothrix caespitora, Calothrix confervicola, Calothrix crustacea, Calothrix donnelli, Calothrix elenkinii, Calothrix epiphytica, Calothrix fusca, Calothrix juliana, Calothrix parasitica, Calothrix parietina, Calothrix pilosa, Calothrix pulvinata, Calothrix scopulorum, Calothrix scytonemicola, Calothrix simulans, Calothrix solitaria, Calothrix stagnalis, Calothrix stellaris, Calothrix thermalis, or Calothrix 336/3.* In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Calothrix sp. PCC 7103.*

[0278] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 7. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 18.

[0279] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus Thermosynechococcus. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Thermosynechococcus elongatus.*

[0280] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 10. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 21.

[0281] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of order chroococcidiopsidales. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of family chroococcidiopsidaceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus chroococcidiopsis. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Chroococcopsis gigantea* or *Chroococcidiopsis thermalis*. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Chroococcidiopsis thermalis*.

[0282] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 9. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 20.

[0283] In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of Phylum deinococcus-thermus. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of class deinocci. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of order deinoccoccales. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of family deinococcaceae. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of genus deinococcus. In some embodiments, the Ago (or variant or functional fragment thereof) is derived from a bacterium of species *Deinobacter Oyaizu* or *Deinococcus sp. YIM 77859.*

[0284] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of SEQ ID NO: 8. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 19.

[0285] In some embodiments the mesophilic Ago has an amino acid sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to an amino acid sequence of one of SEQ ID NOs: 4-10. In some embodiments the mesophilic Ago has an amino acid sequence encoded by a nucleic acid at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence of one of SEQ ID NOs: 15-21.

II. SSB Polypeptides

[0286] In some embodiments, described herein are fusion proteins that comprise an single strand DNA binding protein (SSB) polypeptide. In some embodiments, described herein are methods of engineering cells comprising introducing into a cell an Ago (e.g., described herein) and an SSB (e.g., as described herein). Such introduction can be made by separately introducing an Ago and SSB; or by introducing a fusion polypeptide (or nucleic acid encoding said polypeptide) that comprises both an Ago polypeptide and an SSB polypeptide (e.g., Ago-SSB fusions described herein).

[0287] In some embodiments, the SSB polypeptide component of an Ago-SSB fusion comprises an SSB polypeptide described herein (or a functional fragment or functional variant thereof). In some embodiments, the SSB polypeptide component of an Ago-SSB fusion comprises an SSB derived from a microorganism. In some embodiments, the microorganism is a bacterium. In some embodiments, the microorganism is a hyperthermophilic microorganism. In some embodiments, the SSB is from Saccharolobus solfataricus. In some embodiments, the SSB is active at a temperature between 32°C - 42°C. In some embodiments, the SSB is active at a temperature between 35°C - 40°C. In some embodiments, the SSB is active at about 37°C.

[0288] In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 22-35. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 36-49. In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 22, 24, 26, 58, 30, 32, or 34. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 36, 38, 40, 42, 44, 46, or 48. In some embodiments, the SSB polypeptide is one selected from Table 4 or Table 5.

[0289] In some embodiments, the SSB is ET-SSB (Sso-SSB), Neq SSB, TaqSSB, TmaSSB, or EcoSSB. In some embodiments, the SSB is an ET-SSB (also referred to herein as Sso-SSB). In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 22. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NOS: 36.

Table 4. Amino Acid Sequence of Exemplary SSB Proteins

SSB	Amino Acid Sequence	SEQ ID NO
ET-SSB (also referred to herein as Sso-SSB) *Saccharolobus solfataricus*	MEEKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVR TISEAIVGDETGRVKLTLWGKHAGSIKEGQVVKIENA WTTAFKGQVQLNAGSKTKIAEASEDGFPESSQIPENT PTAPQQMRGGGRGFRGGGRRYGRRGGRRQENEEGEEE	22
ET-SSB (also referred to herein as Sso-SSB) *Saccharolobus solfataricus* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Sso-SSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGA**MEEKVGNL KPNMESVNVTVRVLEASEARQIQTKNGVRTISEAIVG DETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQ VQLNAGSKTKIAEASEDGFPESSQIPENTPTAPQQMR GGGRGFRGGGRRYGRRGGRRQENEEGEEE**	23
Neq SSB *Nanoarchaeum equitans*	MDEEELIQLIIEKTGKSREEIEKMVEEKIKAFNNLIS RRGALLLVAKKLGVLYKNTPKEKKIGELESWEYVKVK GKILKSFGLISYSKGKFQPIILGDETGTIKAIIWNTD KELPENTVIEAIGKTKINKKTGNLELHIDSYKILESD LEIKPQKQEFVGICIVKYPKKQTQKGTIVSKAILTSL DRELPVVYFNDFDWEIGHIYKVYGKLKKNIKTGKIEF FADKVEEATLKDLKAFKGEAD	24
Neq SSB *Nanoarchaeum equitans* *Italicized: His tag* Underlined: GST **Bold: NeqSSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGALA**MDEEEL IQLIIEKTGKSREEIEKMVEEKIKAFNNLISRRGALL LVAKKLGVLYKNTPKEKKIGELESWEYVKVKGKILKS FGLISYSKGKFQPIILGDETGTIKAIIWNTDKELPEN TVIEAIGKTKINKKTGNLELHIDSYKILESDLEIKPQ KQEFVGICIVKYPKKQTQKGTIVSKAILTSLDRELPV	25

	VYFNDFDWEIGHIYKVYGKLKKNIKTGKIEFFADKVE EATLKDLKAFKGEAD	
TaqSSB *Thermus aquaticus*	MARGLNQVFLIGTLTARPDMRYTPGGLAILDLNLAGQ DAFTDESGQEREVPWYHRVRLLGRQAEMWGDLLEKGQ LIFVEGRLEYRQWEKDGEKKSEVQVRAEFIDPLEGRG RETLEDARGQPRLRRALNQVILMGNLTRDPDLRYTPQ GTAVVRLGLAVNERRRGQEEERTHFLEVQAWRELAEW ASELRKGDGLLVIGRLVNDSWTSSSGERRFQTRVEAL RLERPTRGPAQAGGSRPPTVQTGGVDIDEGLEDFPPE EDLPF	26
TaqSSB *Thermus aquaticus* *Italicized: His tag* Underlined: GST **Bold: TaqSSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGA**MARGLNQV FLIGTLTARPDMRYTPGGLAILDLNLAGQDAFTDESG QEREVPWYHRVRLLGRQAEMWGDLLEKGQLIFVEGRL EYRQWEKDGEKKSEVQVRAEFIDPLEGRGRETLEDAR GQPRLRRALNQVILMGNLTRDPDLRYTPQGTAVVRLG LAVNERRRGQEEERTHFLEVQAWRELAEWASELRKGD GLLVIGRLVNDSWTSSSGERRFQTRVEALRLERPTRG PAQAGGSRPPTVQTGGVDIDEGLEDFPPEEDLPF**	27
TmaSSB *Thermotoga maritima*	MGSFFNKIILIGRLVRDPEERYTLSGTPVTTFTIAVD RVPRKNAPDDAQTTDFFRIVTFGRLAEFARTYLTKGR LVLVEGEMRMRRWETPTGEKRVSPEVVANVVRFMDRK PAETVSETEEELEIPEEDFSSDTFSEDEPPF	28
TmaSSB *Thermotoga maritima* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: TmaSSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGA**MGSFFNKI ILIGRLVRDPEERYTLSGTPVTTFTIAVDRVPRKNAP DDAQTTDFFRIVTFGRLAEFARTYLTKGRLVLVEGEM RMRRWETPTGEKRVSPEVVANVVRFMDRKPAETVSET EEELEIPEEDFSSDTFSEDEPPF**	29
EcoSSB *Escherichia coli (strain K12)* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: EcoSSB**	MASRGVNKVILVGNLGQDPEVRYMPNGGAVANITLAT SESWRDKATGEMKEQTEWHRVVLFGKLAEVASEYLRK GSQVYIEGQLRTRKWTDQSGQDRYTTEVVVNVGGTMQ MLGGRQGGGAPAGGNIGGGQPQGGWGQPQQPQGGNQF SGGAQSRPQQSAPAAPSNEPPMDFDDDIPF	30

EcoSSB *Escherichia coli (strain K12)* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: EcoSSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGALA**MASRGV** **NKVILVGNLGQDPEVRYMPNGGAVANITLATSESWRD** **KATGEMKEQTEWHRVVLFGKLAEVASEYLRKGSQVY** **IEGQLRTRKWTDQSGQDRYTTEVVVNVGGTMQMLGGR** **QGGGAPAGGNIGGGQPQGGWGQPQQPQGGNQFSGGAQ** **SRPQQSAPAAPSNEPPMDFDDDIPF**	31
TthSSB *Thermus Thermophilus*	MARGLNRVFLIGALATRPDMRYTPAGLAILDLTL AGQDLLLSDNGGEREVSWYHRVRLLGRQAEMW GDLLDQGQLVFVEGRLEYRQWEREGEKRSELQIR ADFLDPLDDRGKERAEDSRGQPRLRAALNQVFLMGNL TRDPELRYTPQGTAVARLGLAVNERRQGAEERTHFVE VQAWRDLAEWAAELRKGDGLFVIGRLVNDSWTSSSGE RRFQTRVEALRLERPTRGPAQAGGSRSREAQTGGVDI DEGLEDFPPEEELPF	32
TthSSB *Thermus Thermophilus* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: TthSSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRL LLEYLEEKYEEHLYERDEGDKWRNKKFELGLEF PNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCP KERAEISMLEGAVLDIRYGVSRIAYSKDFETLKV DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDF MLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI PQIDKYLKSSKYIAWPLQGWQATFGGGDHPPTSG SGGGGGWMSENLYFQGA**MARGLNRVFLIGALA** **TRPDMRYTPAGLAILDLTLAGQDLLLSDNGGE** **REVSWYHRVRLLGRQAEMWGDLLDQGQLVF** **VEGRLEYRQWEREGEKRSELQIRADFLDPLDD** **RGKERAEDSRGQPRLRAALNQVFLMGNLTRDPELRYT** **PQGTAVARLGLAVNERRQGAEERTHFVEVQAWRDLAE** **WAAELRKGDGLFVIGRLVNDSWTSSSGERRFQTRVEA** **LRLERPTRGPAQAGGSRSREAQTGGVDIDEGLEDFPP** **EEELPF**	33
TneSSB *Thermotoga neapolitana*	MGSFFNRIILIGRLVRDPEERYTLSGTPVTTFTIAVD RVPRKNAPDDAQTTDFFRVVTFGRLAEFARTYLTKGR LILVEGEMRMRRWETQTGEKRVSPEVVANVVRFMDRK PVEMPSEDIEEKLEIPEEDFTDDTFSEDEPPF	34
TneSSB *Thermotoga neapolitana* Construct Sequence *Italicized: His tag*	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGA**MGSFFNRI** **ILIGRLVRDPEERYTLSGTPVTTFTIAVDRVPRKNAP**	35

Underlined: GST **Bold: TneSSB**	**DDAQTTDFFRVVTFGRLAEFARTYLTKGRLILVEGEM RMRRWETQTGEKRVSPEVVANVVRFMDRKPVEMPSED IEEKLEIPEEDFTDDTFSEDEPPF**	

Table 5. Nucleic Acid Sequence of Exemplary SSBs

SSB	Nucleic Acid Sequence	SEQ ID NO
ET-SSB (also referred to herein as Sso-SSB) *Saccharolobus solfataricus*	ATGGAAGAAAAAGTAGGCAACCTGAAGCCTAATATGG AATCCGTAAATGTAACCGTTCGCGTTTTAGAAGCCTC TGAAGCACGGCAGATCCAGACCAAAAATGGTGTTCGC ACCATTTCAGAGGCGATTGTAGGGGATGAAACCGGGC GCGTGAAACTGACTCTGTGGGGCAAACATGCGGGCAG CATCAAAGAAGGCCAGGTCGTTAAAATTGAGAACGCC TGGACAACCGCGTTCAAAGGCCAGGTACAGCTGAATG CCGGTAGCAAGACCAAAATTGCCGAGGCATCTGAAGA CGGTTTCCCTGAAAGCAGCCAGATCCCAGAAAATACT CCTACGGCACCGCAGCAGATGCGTGGCGGTGGGCGGG GCTTTCGTGGCGGAGGCCGCCGTTATGGCCGTCGCGG TGGGCGCCGGCAAGAAAACGAAGAAGGCGAAGAAGAA TAG	36
ET-SSB (also referred to herein as Sso-SSB) Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Sso-SSB**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATG**AGCGAGAATCTTTA TTTTCAGGGCGCCATGGAAGAAAAAGTAGGCAACCTG AAGCCTAATATGGAATCCGTAAATGTAACCGTTCGCG TTTTAGAAGCCTCTGAAGCACGGCAGATCCAGACCAA AAATGGTGTTCGCACCATTTCAGAGGCGATTGTAGGG GATGAAACCGGGCGCGTGAAACTGACTCTGTGGGGCA AACATGCGGGCAGCATCAAAGAAGGCCAGGTCGTTAA AATTGAGAACGCCTGGACAACCGCGTTCAAAGGCCAG GTACAGCTGAATGCCGGTAGCAAGACCAAAATTGCCG AGGCATCTGAAGACGGTTTCCCTGAAAGCAGCCAGAT**	37

	CCCAGAAAATACTCCTACGGCACCGCAGCAGATGCGT GGCGGTGGGCGGGGCTTTCGTGGCGGAGGCCGCCGTT ATGGCCGTCGCGGTGGGCGCCGGCAAGAAAACGAAGA AGGCGAAGAAGAATAG	
Neq SSB *Nanoarchaeum equitans*	ATGGATGAAGAAGAGCTTATTCAGCTTATTATTGAAA AAACTGGGAAGTCGCGCGAGGAAATTGAGAAGATGGT AGAAGAAAAAATCAAGGCTTTCAACAACCTGATCTCG CGCCGTGGCGCTTTGCTGTTAGTGGCGAAGAAACTTG GTGTACTTTATAAAAACACGCCAAAAGAAAAAAAGAT CGGGGAACTGGAGTCCTGGGAATACGTTAAGGTGAAA GGTAAAATCCTGAAGTCCTTCGGCCTGATTAGTTATT CAAAGGGCAAGTTTCAACCGATCATTCTTGGGGACGA AACTGGCACTATCAAAGCTATTATCTGGAATACTGAT AAGGAGTTACCTGAGAATACGGTCATTGAAGCTATTG GAAAGACTAAGATTAACAAAAAAACAGGGAATCTTGA GTTACATATTGATAGTTACAAAATTTTAGAGTCCGAC TTAGAGATTAAGCCTCAGAAACAGGAATTTGTCGGTA TTTGCATCGTTAAATACCCCAAGAAGCAGACCCAAAA GGGTACGATTGTAAGCAAAGCTATCCTGACATCATTA GACCGTGAGTTACCCGTCGTTTACTTTAATGATTTTG ACTGGGAGATCGGGCATATCTATAAGGTCTACGGGAA ACTGAAAAAAAATATCAAAACTGGCAAAATCGAGTTT TTTGCAGACAAGGTAGAGGAAGCCACCCTGAAAGATC TTAAGGCGTTCAAGGGGGAAGCAGACTAGTGA	38
Neq SSB *Nanoarchaeum equitans* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: NeqSSB**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCGCTAGCA**ATGGATGAAGAAGAGCTT ATTCAGCTTATTATTGAAAAAACTGGGAAGTCGCGCG AGGAAATTGAGAAGATGGTAGAAGAAAAAATCAAGGC TTTCAACAACCTGATCTCGCGCCGTGGCGCTTTGCTG TTAGTGGCGAAGAAACTTGGTGTACTTTATAAAAACA**	39

	CGCCAAAAGAAAAAAAGATCGGGGAACTGGAGTCCTG GGAATACGTTAAGGTGAAAGGTAAAATCCTGAAGTCC TTCGGCCTGATTAGTTATTCAAAGGGCAAGTTTCAAC CGATCATTCTTGGGGACGAAACTGGCACTATCAAAGC TATTATCTGGAATACTGATAAGGAGTTACCTGAGAAT ACGGTCATTGAAGCTATTGGAAAGACTAAGATTAACA AAAAAACAGGGAATCTTGAGTTACATATTGATAGTTA CAAAATTTTAGAGTCCGACTTAGAGATTAAGCCTCAG AAACAGGAATTTGTCGGTATTTGCATCGTTAAATACC CCAAGAAGCAGACCCAAAAGGGTACGATTGTAAGCAA AGCTATCCTGACATCATTAGACCGTGAGTTACCCGTC GTTTACTTTAATGATTTTGACTGGGAGATCGGGCATA TCTATAAGGTCTACGGGAAACTGAAAAAAAATATCAA AACTGGCAAAATCGAGTTTTTTGCAGACAAGGTAGAG GAAGCCACCCTGAAAGATCTTAAGGCGTTCAAGGGGG AAGCAGACTAGTGA**	
TaqSSB *Thermus aquaticus*	ATGGCGCGCGGTCTGAACCAGGTATTTCTGATCGGCA CCCTCACTGCCCGTCCAGATATGCGCTATACCCCGGG CGGGCTGGCAATTCTGGATCTCAATCTTGCTGGGCAG GATGCGTTTACCGATGAAAGTGGGCAAGAGCGTGAAG TCCCGTGGTATCATCGTGTGCGTCTGCTCGGCCGTCA AGCGGAAATGTGGGGTGACCTGCTGGAAAAAGGTCAG CTGATCTTTGTGGAAGGTCGCCTGGAATACCGCCAAT GGGAAAAAGACGGCGAAAAAAAGAGCGAAGTCCAAGT CCGTGCTGAGTTTATTGATCCGCTGGAAGGTCGCGGC CGTGAGACGCTCGAAGATGCTCGTGGTCAGCCCCGCT TACGTCGTGCACTGAACCAGGTTATTCTCATGGGTAA CCTCACCCGCGATCCCGATTTACGCTATACCCCCCAG GGTACGGCGGTGGTACGCCTGGGCCTTGCTGTGAACG AGCGGCGTCGTGGCCAAGAAGAAGAACGTACCCATTT TCTGGAAGTGCAGGCGTGGCGCGAGCTGGCCGAATGG GCTAGCGAATTACGCAAAGGCGACGGTCTTCTGGTCA TCGGTCGCTTGGTCAACGATTCCTGGACAAGCTCCTC GGGTGAACGTCGCTTCCAAACGCGTGTGGAGGCACTG CGGTTAGAACGTCCGACCCGCGGCCCGGCACAGGCGG GGGGATCCCGGCCGCCCACCGTGCAGACGGGGGGTGT GGATATCGATGAGGGGCTGGAAGACTTTCCGCCTGAA GAAGATCTGCCTTTCTAG	40
TaqSSB *Thermus aquaticus* *Italicized: His tag* Underlined: GST **Bold: TaqSSB**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT	41

	TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCC**ATGGCGCGCGGTCTGAACCAGGTA** **TTTCTGATCGGCACCCTCACTGCCCGTCCAGATATGC** **GCTATACCCCGGGCGGGCTGGCAATTCTGGATCTCAA** **TCTTGCTGGGCAGGATGCGTTTACCGATGAAAGTGGG** **CAAGAGCGTGAAGTCCCGTGGTATCATCGTGTGCGTC** **TGCTCGGCCGTCAAGCGGAAATGTGGGGTGACCTGCT** **GGAAAAAGGTCAGCTGATCTTTGTGGAAGGTCGCCTG** **GAATACCGCCAATGGGAAAAAGACGGCGAAAAAAAGA** **GCGAAGTCCAAGTCCGTGCTGAGTTTATTGATCCGCT** **GGAAGGTCGCGGCCGTGAGACGCTCGAAGATGCTCGT** **GGTCAGCCCCGCTTACGTCGTGCACTGAACCAGGTTA** **TTCTCATGGGTAACCTCACCCGCGATCCCGATTTACG** **CTATACCCCCCAGGGTACGGCGGTGGTACGCCTGGGC** **CTTGCTGTGAACGAGCGGCGTCGTGGCCAAGAAGAAG** **AACGTACCCATTTTCTGGAAGTGCAGGCGTGGCGCGA** **GCTGGCCGAATGGGCTAGCGAATTACGCAAAGGCGAC** **GGTCTTCTGGTCATCGGTCGCTTGGTCAACGATTCCT** **GGACAAGCTCCTCGGGTGAACGTCGCTTCCAAACGCG** **TGTGGAGGCACTGCGGTTAGAACGTCCGACCCGCGGC** **CCGGCACAGGCGGGGGGATCCCGGCCGCCCACCGTGC** **AGACGGGGGGGTGTGGATATCGATGAGGGGCTGGAAGA** **CTTTCCGCCTGAAGAAGATCTGCCTTTCTAG**	
TmaSSB *Thermotoga maritima*	ATGGGATCTTTCTTCAACAAAATTATCCTTATCGGCC GTCTGGTCCGCGACCCGGAAGAACGTTATACACTGTC TGGCACACCGGTCACCACCTTTACTATTGCCGTCGAT CGTGTTCCGCGCAAAAACGCACCGGATGATGCCCAGA CCACCGATTTTTTTCGCATTGTGACTTTCGGCCGCCT GGCGGAGTTTGCCCGTACTTATTTAACGAAAGGTCGT CTCGTGCTCGTAGAGGGCGAGATGCGCATGCGCCGTT GGGAAACACCAACGGGCGAAAAACGTGTGAGCCCGGA AGTGGTGGCCAATGTGGTTCGTTTTATGGACCGCAAA CCTGCCGAAACCGTCAGCGAAACGGAAGAGGAACTCG AAATCCCAGAGGAGGACTTCAGCTCAGACACCTTTTC GGAAGATGAACCCCCGTTTTAG	42
TmaSSB *Thermotoga maritima* Construct Sequence	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT	43

Italicized: His tag Underlined: GST **Bold: TmaSSB**	GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCC**ATGGGATCTTTCTTCAACAAAATT** **ATCCTTATCGGCCGTCTGGTCCGCGACCCGGAAGAAC** **GTTATACACTGTCTGGCACACCGGTCACCACCTTTAC** **TATTGCCGTCGATCGTGTTCCGCGCAAAAACGCACCG** **GATGATGCCCAGACCACCGATTTTTTTCGCATTGTGA** **CTTTCGGCCGCCTGGCGGAGTTTGCCCGTACTTATTT** **AACGAAAGGTCGTCTCGTGCTCGTAGAGGGCGAGATG** **CGCATGCGCCGTTGGGAAACACCAACGGGCGAAAAAC** **GTGTGAGCCCGGAAGTGGTGGCCAATGTGGTTCGTTT** **TATGGACCGCAAACCTGCCGAAACCGTCAGCGAAACG** **GAAGAGGAACTCGAAATCCCAGAGGAGGACTTCAGCT** **CAGACACCTTTTCGGAAGATGAACCCCCGTTTTAG**	
EcoSSB *Escherichia coli (strain K12)*	ATGGCATCACGTGGCGTCAACAAGGTCATTTTAGTCG GAAACCTTGGGCAGGATCCTGAAGTCCGCTACATGCC CAATGGAGGCGCTGTTGCGAATATCACATTGGCAACT AGTGAAAGCTGGCGCGATAAGGCTACGGGAGAGATGA AGGAGCAAACGGAGTGGCACCGTGTGGTATTGTTCGG CAAATTAGCTGAAGTGGCTAGTGAATATTTGCGTAAA GGTTCGCAAGTGTATATTGAGGGCCAGCTTCGTACCC GTAAGTGGACCGACCAAAGTGGACAGGACCGCTACAC TACGGAAGTAGTGGTCAATGTAGGCGGGACGATGCAA ATGCTTGGTGGACGTCAAGGTGGTGGAGCTCCAGCAG GAGGTAATATCGGTGGTGGACAGCCCCAAGGGGGTTG GGGCCAACCGCAACAGCCACAGGGGGGTAACCAATTT TCCGGTGGGGCTCAGAGCCGTCCACAGCAGTCGGCTC CCGCAGCACCAAGCAATGAACCCCCGATGGACTTTGA TGACGATATTCCTTTCTAGTGA	44
EcoSSB *Escherichia coli (strain K12)* Construct Sequence *Italicized: His tag*	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG	45

Underlined: GST **Bold: Eco-SSB**	TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCGCTAGCA**ATGGCATCACGTGGCGTC** **AACAAGGTCATTTTAGTCGGAAACCTTGGGCAGGATC** **CTGAAGTCCGCTACATGCCCAATGGAGGCGCTGTTGC** **GAATATCACATTGGCAACTAGTGAAAGCTGGCGCGAT** **AAGGCTACGGGAGAGATGAAGGAGCAAACGGAGTGGC** **ACCGTGTGGTATTGTTCGGCAAATTAGCTGAAGTGGC** **TAGTGAATATTTGCGTAAAGGTTCGCAAGTGTATATT** **GAGGGCCAGCTTCGTACCCGTAAGTGGACCGACCAAA** **GTGGACAGGACCGCTACACTACGGAAGTAGTGGTCAA** **TGTAGGCGGGACGATGCAAATGCTTGGTGGACGTCAA** **GGTGGTGGAGCTCCAGCAGGAGGTAATATCGGTGGTG** **GACAGCCCCAAGGGGGTTGGGGCCAACCGCAACAGCC** **ACAGGGGGGTAACCAATTTTCCGGTGGGGCTCAGAGC** **CGTCCACAGCAGTCGGCTCCCGCAGCACCAAGCAATG** **AACCCCCGATGGACTTTGATGACGATATTCCTTTCTA** **GTGA**	
TthSSB *Thermus Thermophilus*	ATGGCACGCGGCCTGAACCGCGTTTTTCTGATTGGTG CACTGGCCACCCGCCCGGATATGCGCTATACCCCGGC AGGCCTTGCAATTTTAGACCTGACCCTTGCGGGCCAA GATTTACTGCTTTCAGACAATGGCGGTGAACGTGAGG TGAGTTGGTACCATCGTGTACGCCTGTTAGGACGTCA GGCCGAGATGTGGGGCGATCTGCTTGACCAGGGCCAG CTGGTGTTTGTGGAGGGCCGCCTTGAGTATCGTCAAT GGGAACGTGAAGGTGAAAAACGCTCCGAACTGCAGAT TCGCGCTGATTTCCTCGATCCGTTGGATGATCGCGGT AAGGAACGCGCAGAAGATAGCCGGGGTCAGCCACGGC TCCGTGCCGCGCTGAACCAGGTATTTTTAATGGGCAA TCTGACCCGCGATCCCGAACTGCGCTACACTCCACAG GGCACCGCAGTCGCTCGTTTAGGCCTGGCTGTGAACG AACGCCGTCAGGGCGCGGAAGAACGTACCCACTTCGT TGAAGTCCAGGCCTGGCGCGACTTAGCAGAGTGGGCC GCAGAGCTGCGTAAGGGTGACGGCCTGTTCGTTATCG GGCGTCTCGTTAACGACTCTTGGACTAGCTCGTCAGG TGAGCGTCGCTTTCAAACCCGTGTCGAAGCCCTGCGG CTGGAACGCCCAACGCGGGGTCCGGCACAGGCCGGCG GGTCGCGCTCTCGCGAGGCACAGACAGGCGGGGTTGA	46

	TATTGACGAAGGGTTAGAGGATTTCCCGCCAGAGGAG GAGCTGCCTTTTTAG	
TthSSB *Thermus Thermophilus* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: TthSSB**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCC**ATGGCACGCGGCCTGAACCGCGTT** **TTTCTGATTGGTGCACTGGCCACCCGCCCGGATATGC** **GCTATACCCCGGCAGGCCTTGCAATTTTAGACCTGAC** **CCTTGCGGGCCAAGATTTACTGCTTTCAGACAATGGC** **GGTGAACGTGAGGTGAGTTGGTACCATCGTGTACGCC** **TGTTAGGACGTCAGGCCGAGATGTGGGGCGATCTGCT** **TGACCAGGGCCAGCTGGTGTTTGTGGAGGGCCGCCTT** **GAGTATCGTCAATGGGAACGTGAAGGTGAAAAACGCT** **CCGAACTGCAGATTCGCGCTGATTTCCTCGATCCGTT** **GGATGATCGCGGTAAGGAACGCGCAGAAGATAGCCGG** **GGTCAGCCACGGCTCCGTGCCGCGCTGAACCAGGTAT** **TTTTAATGGGCAATCTGACCCGCGATCCCGAACTGCG** **CTACACTCCACAGGGCACCGCAGTCGCTCGTTTAGGC** **CTGGCTGTGAACGAACGCCGTCAGGGCGCGGAAGAAC** **GTACCCACTTCGTTGAAGTCCAGGCCTGGCGCGACTT** **AGCAGAGTGGGCCGCAGAGCTGCGTAAGGGTGACGGC** **CTGTTCGTTATCGGGCGTCTCGTTAACGACTCTTGGA** **CTAGCTCGTCAGGTGAGCGTCGCTTTCAAACCCGTGT** **CGAAGCCCTGCGGCTGGAACGCCCAACGCGGGGTCCG** **GCACAGGCCGGCGGGTCGCGCTCTCGCGAGGCACAGA** **CAGGCGGGGTTGATATTGACGAAGGGTTAGAGGATTT** **CCCGCCAGAGGAGGAGCTGCCTTTTTAG**	47
TneSSB *Thermotoga neapolitana*	ATGGGATCCTTCTTTAACCGTATTATTTTAATTGGCC GCCTGGTTCGGGATCCTGAAGAACGCTATACCCTGTC AGGGACTCCGGTGACGACTTTTACTATCGCGGTCGAT CGCGTTCCTCGTAAGAATGCCCCTGATGATGCCCAGA CAACTGACTTTTTTCGTGTTGTAACCTTTGGTCGCTT	48

	GGCGGAATTCGCACGGACGTATCTGACCAAAGGCCGC CTTATCCTGGTCGAGGGTGAAATGCGCATGCGTCGCT GGGAAACCCAGACTGGCGAAAAACGCGTGAGCCCGGA AGTAGTTGCAAATGTCGTGCGTTTTATGGACCGCAAA CCCGTGGAAATGCCGAGCGAAGACATTGAAGAAAAAC TGGAAATTCCCGAAGAAGACTTTACGGACGATACGTT TTCGGAGGATGAACCCCCGTTTTAG	
TneSSB *Thermotoga neapolitana* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: TneSSB**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCC**ATGGGATCCTTCTTTAACCGTATT** **ATTTTAATTGGCCGCCTGGTTCGGGATCCTGAAGAAC** **GCTATACCCTGTCAGGGACTCCGGTGACGACTTTTAC** **TATCGCGGTCGATCGCGTTCCTCGTAAGAATGCCCCT** **GATGATGCCCAGACAACTGACTTTTTTCGTGTTGTAA** **CCTTTGGTCGCTTGGCGGAATTCGCACGGACGTATCT** **GACCAAAGGCCGCCTTATCCTGGTCGAGGGTGAAATG** **CGCATGCGTCGCTGGGAAACCCAGACTGGCGAAAAAC** **GCGTGAGCCCGGAAGTAGTTGCAAATGTCGTGCGTTT** **TATGGACCGCAAACCCGTGGAAATGCCGAGCGAAGAC** **ATTGAAGAAAAACTGGAAATTCCCGAAGAAGACTTTA** **CGGACGATACGTTTTCGGAGGATGAACCCCCGTTTTA** **G**	49

III. Helicase Polypeptides

[0290] In some embodiments, described herein are fusion proteins that comprise a helicase polypeptide. In some embodiments, described herein are methods of engineering cells comprising introducing into a cell an Ago (e.g., described herein) and a helicase (e.g., as described herein). Such introduction can be made by separately introducing an Ago and helicase;

or by introducing a fusion polypeptide (or nucleic acid encoding said polypeptide) that comprises both an Ago polypeptide and a helicase (e.g., Ago-helicase fusions described herein).

[0291] In some embodiments, the helicase polypeptide component of an Ago-helicase fusion comprises a helicase polypeptide described herein (or a functional fragment or functional variant thereof). In some embodiments, the helicase polypeptide component of an Ago-helicase fusion comprises a helicase derived from a microorganism. In some embodiments, the microorganism is a bacterium. In some embodiments, the microorganism is a hyperthermophilic microorganism. In some embodiments, the helicase is active at a temperature between 32°C - 42°C. In some embodiments, the helicase is active at a temperature between 35°C - 40°C. In some embodiments, the helicase is active at about 37°C.

[0292] In some embodiments, the helicase has an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 50-59. In some embodiments, the helicase is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 60-69. In some embodiments, the helicase has an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 50, 52, 54, 56, or 58. In some embodiments, the helicase is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 60, 62, 64, 66, or 68. In some embodiments, the helicase polypeptide is one selected from Table 6 or Table 7. In some embodiments, the helicase is Eco RecQ, Tth UvrD, Eco UvrD, HEL#100, HEL#75, or HEL#76.

Table 6. Amino Acid Sequence of Exemplary Helicases

Helicase	Amino Acid Sequence	SEQ ID NO
Eco RecQ *Escherichia coli*	MAQAEVLNLESGAKQVLQETFGYQQFRPGQEEIIDTV LSGRDCLVVMPTGGGKSLCYQIPALLLNGLTVVVSPL ISLMKDQVDQLQANGVAAACLNSTQTREQQLEVMTGC RTGQIRLLYIAPERLMLDNFLEHLAHWNPVLLAVDEA HCISQWGHDFRPEYAALGQLRQRFPTLPFMALTATAD DTTRQDIVRLLGLNDPLIQISSFDRPNIRYMLMEKFK PLDQLMRYVQEQRGKSGIIYCNSRAKVEDTAARLQSK GISAAAYHAGLENNVRADVQEKFQRDDLQIVVATVAF GMGINKPNVRFVVHFDIPRNIESYYQETGRAGRDGLP AEAMLFYDPADMAWLRRCLEEKPQGQLQDIERHKLNA MGAFAEAQTCRRLVLLNYFGEGRQEPCGNCDICLDPP KQYDGSTDAQIALSTIGRVNQRFGMGYVVEVIRGANN QRIRDYGHDKLKVYGMGRDKSHEHWVSVIRQLIHLGL VTQNIAQHSALQLTEAARPVLRGESSLQLAVPRIVAL	50

	KPKAMQKSFGGNYDRKLFAKLRKLRKSIADESNVPPY VVFNDATLIEMAEQMPITASEMLSVNGVGMRKLERFG KPFMALIRAHVDGDDEE	
Eco RecQ *Escherichia coli* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: EcoReq**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGALA**MAQAEV** **LNLESGAKQVLQETFGYQQFRPGQEEIIDTVLSGRDC** **LVVMPTGGGKSLCYQIPALLLNGLTVVVSPLISLMKD** **QVDQLQANGVAAACLNSTQTREQQLEVMTGCRTGQIR** **LLYIAPERLMLDNFLEHLAHWNPVLLAVDEAHCISQW** **GHDFRPEYAALGQLRQRFPTLPFMALTATADDTTRQD** **IVRLLGLNDPLIQISSFDRPNIRYMLMEKFKPLDQLM** **RYVQEQRGKSGIIYCNSRAKVEDTAARLQSKGISAAA** **YHAGLENNVRADVQEKFQRDDLQIVVATVAFGMGINK** **PNVRFVVHFDIPRNIESYYQETGRAGRDGLPAEAMLF** **YDPADMAWLRRCLEEKPQGQLQDIERHKLNAMGAFAE** **AQTCRRLVLLNYFGEGRQEPCGNCDICLDPPKQYDGS** **TDAQIALSTIGRVNQRFGMGYVVEVIRGANNQRIRDY** **GHDKLKVYGMGRDKSHEHWVSVIRQLIHLGLVTQNIA** **QHSALQLTEAARPVLRGESSLQLAVPRIVALKPKAMQ** **KSFGGNYDRKLFAKLRKLRKSIADESNVPPYVVFNDA** **TLIEMAEQMPITASEMLSVNGVGMRKLERFGKPFMAL** **IRAHVDGDDEE**	51
Tth UvrD *Thermus Thermophilus*	MSDALLAPLNEAQRQAVLHFEGPALVVAGAGSGKTRT VVHRVAYLVARRGVFPSEILAVTFTNKAAEEMRERLR GLVPGAGEVWVSTFHAAALRILRVYGERVGLRPGFVV YDEDDQTALLKEVLKELALSARPGPIKALLDRAKNRG VGLKALLGELPEYYAGLSRGRLGDVLVRYQEALKAQG ALDFGDILLYALEAFRGGRGGPQARAQRARFIHVDEY QDTSPVQYRFTRLLAGEEANLMAVGDPDQGIYSFRAA DIKNILDFTRDYPEARVYRLEENYRSTEAILRFANAV IVKNALRLEKALRPVKRGGEPVRLYRAEDAREEARFV AEEIARLGPPWDRYAVLYRTNAQSRLLEQALAGRGIP ARVVGGVGFFERAEVKDLLAYARLALNPLDAVSLKRV LNTPPRGIGPATWARVQLLAQEKGLPPWEALKEAART FSRPEPLRHFVALVEELQDLVFGPAEAFFRHLLEATD YPAYLREAYPEDAEDRLENVEELLRAAKEAEDLQDFL DRVALTAKAEEPAEAEGRVALMTLHNAKGLEFPVVFL VGVEEGLLPHRNSVSTLEGLEEERRLFYVGITRAQER LYLSHAEEREVYGRREPARPSRFLEEVEEGLYEVYDP YRRPPSPPPHRPRPGAFRGGERVVHPRFGPGTVVAAQ GDEVTVHFEGFGLKRLSLKYAELKPA	52
Tth UvrD *Thermus Thermophilus*	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG	53

Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Tht UvrD**	AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGALA**MSDALL** **APLNEAQRQAVLHFEGPALVVAGAGSGKTRTVVHRVA** **YLVARRGVFPSEILAVTFTNKAAEEMRERLRGLVPGA** **GEVWVSTFHAAALRILRVYGERVGLRPGFVVYDEDDQ** **TALLKEVLKELALSARPGPIKALLDRAKNRGVGLKAL** **LGELPEYYAGLSRGRLGDVLVRYQEALKAQGALDFGD** **ILLYALEAFRGGRGGPQARAQRARFIHVDEYQDTSPV** **QYRFTRLLAGEEANLMAVGDPDQGIYSFRAADIKNIL** **DFTRDYPEARVYRLEENYRSTEAILRFANAVIVKNAL** **RLEKALRPVKRGGEPVRLYRAEDAREEARFVAEEIAR** **LGPPWDRYAVLYRTNAQSRLLEQALAGRGIPARVVGG** **VGFFERAEVKDLLAYARLALNPLDAVSLKRVLNTPPR** **GIGPATWARVQLLAQEKGLPPWEALKEAARTFSRPEP** **LRHFVALVEELQDLVFGPAEAFFRHLLEATDYPAYLR** **EAYPEDAEDRLENVEELLRAAKEAEDLQDFLDRVALT** **AKAEEPAEAEGRVALMTLHNAKGLEFPVVFLVGVEEG** **LLPHRNSVSTLEGLEEERRLFYVGITRAQERLYLSHA** **EEREVYGRREPARPSRFLEEVEEGLYEVYDPYRRPPS** **PPPHRPRPGAFRGGERVVHPRFGPGTVVAAQGDEVTV** **HFEGFGLKRLSLKYAELKPA**	
Eco UvrD *Escherichia coli*	MDVSYLLDSLNDKQREAVAAPRSNLLVLAGAGSGKTR VLVHRIAWLMSVENCSPYSIMAVTFTNKAAAEMRHRI GQLMGTSQGGMWVGTFHGLAHRLLRAHHMDANLPQDF QILDSEDQLRLLKRLIKAMNLDEKQWPPRQAMWYINS QKDEGLRPHHIQSYGNPVEQTWQKVYQAYQEACDRAG LVDFAELLLRAHELWLNKPHILQHYRERFTNILVDEF QDTNNIQYAWIRLLAGDTGKVMIVGDDDQSIYGWRGA QVENIQRFLNDFPGAETIRLEQNYRSTSNILSAANAL IENNNGRLGKKLWTDGADGEPISLYCAFNELDEARFV VNRIKTWQDNGGALAECAILYRSNAQSRVLEEALLQA SMPYRIYGGMRFFERQEIKDALSYLRLIANRNDDAAF ERVVNTPTRGIGDRTLDVVRQTSRDRQLTLWQACREL LQEKALAGRAASALQRFMELIDALAQETADMPLHVQT DRVIKDSGLRTMYEQEKGEKGQTRIENLEELVTATRQ FSYNEEDEDLMPLQAFLSHAALEAGEGQADTWQDAVQ LMTLHSAKGLEFPQVFIVGMEEGMFPSQMSLDEGGRL EEERRLAYVGVTRAMQKLTLTYAETRRLYGKEVYHRP SRFIGELPEECVEEVRLRATVSRPVSHQRMGTPMVEN DSGYKLGQRVRHAKFGEGTIVNMEGSGEHSRLQVAFQ GQGIKWLVAAYARLESV	54
Eco UvrD *Escherichia coli* Construct Sequence	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA	55

Italicized: His tag Underlined: GST **Bold: Eco UvrD**	TFGGGDHPPTSGSGGGGGWMSENLYFQGALA**MDVSYL** **LDSLNDKQREAVAAPRSNLLVLAGAGSGKTRVLVHRI** **AWLMSVENCSPYSIMAVTFTNKAAAEMRHRIGQLMGT** **SQGGMWVGTFHGLAHRLLRAHHMDANLPQDFQILDSE** **DQLRLLKRLIKAMNLDEKQWPPRQAMWYINSQKDEGL** **RPHHIQSYGNPVEQTWQKVYQAYQEACDRAGLVDFAE** **LLLRAHELWLNKPHILQHYRERFTNILVDEFQDTNNI** **QYAWIRLLAGDTGKVMIVGDDDQSIYGWRGAQVENIQ** **RFLNDFPGAETIRLEQNYRSTSNILSAANALIENNNG** **RLGKKLWTDGADGEPISLYCAFNELDEARFVVNRIKT** **WQDNGGALAECAILYRSNAQSRVLEEALLQASMPYRI** **YGGMRFFERQEIKDALSYLRLIANRNDDAAFERVVNT** **PTRGIGDRTLDVVRQTSRDRQLTLWQACRELLQEKAL** **AGRAASALQRFMELIDALAQETADMPLHVQTDRVIKD** **SGLRTMYEQEKGEKGQTRIENLEELVTATRQFSYNEE** **DEDLMPLQAFLSHAALEAGEGQADTWQDAVQLMTLHS** **AKGLEFPQVFIVGMEEGMFPSQMSLDEGGRLEEERRL** **AYVGVTRAMQKLTLTYAETRRLYGKEVYHRPSRFIGE** **LPEECVEEVRLRATVSRPVSHQRMGTPMVENDSGYKL** **GQRVRHAKFGEGTIVNMEGSGEHSRLQVAFQGQGIKW** **LVAAYARLESV**	
HEL#100 *Clostridium perfringens*	MVLNPKYSIGVYYDELVEEDIEKVYSYLSRGIVVHLF LRGILKEELELNEYDLNTFKLPKDNNLLFVYEEETSL SSENIIIFVDNNILNKEAYKNITENRECEFNKDQYEI ITAPVDDNIIVTSGAGTGKTTTMINRLIYLRSVMSDF TFDQAVLITFTNKASIEMKERLLEVLDKYFRVTNDIK YLDYMEEAAKGSISTIHKFAKKILNKSGRHIGINKDI NVRSFKYKRQEAVNNALNKIYKEESELFSLIKYYPIY EVERVILKMWEILDNYSIDLLSNK VRVDFNFEEDKFTELISKTLKYAQEILDYDKENELEI SDLMKKLAYEDIFKGIDSTYKVIMIDEFQDSDNTQIE FISELEKKTGARILVVGDEKQSIYRFRGAEYTAFDKL KKLLSNSKREVKEYEMTRNYRTNYNILNEINRIFIEV DKKLECFNYKEKDYIYSNKDKDNPKEITCFNVSDNLK RKEFFDDLLENKKEDESIAVLFRSN SDIKEFKEFCDRNNILCMVDSTGGFYRHEAVRDFYIM IKSIIDERNSRTMYSFINTPYILEDIDKNIILNGNSK DKNEFLYYILEKNNWNYFRESSNFKNPIILIDEIIEK LKPVKNYYVKVLLEAKKNQHNYVNIAKMKALEYKLNL EHLVFILKKEFSENITSIEQIEQFLKVKISTDNLVDV RKPKDYENDYIQCSTVHKAKGLEYDYVVLDKLTNRFL SNSRKVNLILKPDGDKLLIGYKIRLGEDEFKNKIYSD NLKYEKKEIKGEEARLLYVAL TRCKKGIYLNMSGELAATESLNTWKSLIGGTINYV	56
HEL#100 *Clostridium perfringens*	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD	57

Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Hel#100**	AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA TFGGGDHPPTSGSGGGGGWMSENLYFQGALA**MVLNPK** **YSIGVYYDELVEEDIEKVYSYLSRGIVVHLFLRGILK** **EELELNEYDLNTFKLPKDNNLLFVYEEETSLSSENII** **IFVDNNILNKEAYKNITENRECEFNKDQYEIITAPVD** **DNIIVTSGAGTGKTTTMINRLIYLRSVMSDFTFDQAV** **LITFTNKASIEMKERLLEVLDKYFRVTNDIKYLDYME** **EAAKGSISTIHKFAKKILNKSGRHIGINKDINVRSFK** **YKRQEAVNNALNKIYKEESELFSLIKYYPIYEVERVI** **LKMWEILDNYSIDLLSNKVRVDFNFEEDKFTELISKT** **LKYAQEILDYDKENELEISDLMKKLAYEDIFKGIDST** **YKVIMIDEFQDSDNTQIEFISELEKKTGARILVVGDE** **KQSIYRFRGAEYTAFDKLKKLLSNSKREVKEYEMTRN** **YRTNYNILNEINRIFIEVDKKLECFNYKEKDYIYSNK** **DKDNPKEITCFNVSDNLKRKEFFDDLLENKKEDESIA** **VLFRSNSDIKEFKEFCDRNNILCMVDSTGGFYRHEAV** **RDFYIMIKSIIDERNSRTMYSFINTPYILEDIDKNII** **LNGNSKDKNEFLYYILEKNNWNYFRESSNFKNPIILI** **DEIIEKLKPVKNYYVKVLLEAKKNQHNYVNIAKMKAL** **EYKLNLEHLVFILKKEFSENITSIEQIEQFLKVKIST** **DNLVDVRKPKDYENDYIQCSTVHKAKGLEYDYVVLDK** **LTNRFLSNSRKVNLILKPDGDKLLIGYKIRLGEDEFK** **NKIYSDNLKYEKKEIKGEEARLLYVALTRCKKGIYLN** **MSGELAATESLNTWKSLIGGTINYV**	
HEL#75 *Clostridium perfringens*	MLGLNNESKEFFKGISRIWRNYKDYTYLDGIKLSQAQ IDIIEKEEDQLLIEGYAGTGKSLTLIYKFINVLVRED GKRVLYVTFNDTLIEDTKKRLSYCNEYNENKERHHVE ICTFHEIASNILKKKKIIDRGIEKLTAKKIEDYKGAA LRRIAGILARYIEGGKYYSELPKEERLYKTHDENFIR EEVAWIKAMGFIEKEKYFEKDRIGRSKSIRLTRSQRK TIFKIFEKYCEEQENKFFKSLDLEDYALKLIQNIDNF DDLKFDYIFVDEVQDLDPMQIKALCLLTNTSIVLSGD ANQRIYKKSPVKYEELGLRIKEKGKRKILNKNYRSTG EIVKLANSIKFFDESINKYNEKQFVKSGDRPIIRKVN DKKGAVKFLIGEIKKIHEEDPYKTIAIIHREKNELIG FQKSEFRKYLEGQLYMEKFSDIKSFESKFDLREKNQV FYTNGYDVKGLEFDVVFIINFNTANYPLSKELKKIKD ENDGKEMTLIKDDVLEFINREKRLLYVAMTRAKEKLY LVADCKNSNISSFIYDFNTKYYEAQNFKKKEIEENYN RYKINMEREYGIIIEDDDSNNVKNNDTKQENKFNTES KEKGKDDIDKIKVFFINKGIEVVDNRDKSGCLWIVAG KEAIPLMKKFGVLGYNFIFIANGGRASKNRPAWYLKN S	58
HEL#75 *Clostridium perfringens* Construct Sequence	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEY LEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEG AVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFE DRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLD AFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQA	59

Italicized: His tag Underlined: GST **Bold/Underlined: 2xSV40 NLS** **Bold: Hel#75**	TFGGGDHPPTSGSGGGGGWMSENLYFQGA**MPKKKRKV** **EDPKKKRKV**GSGS**LGLNNESKEFFKGISRIWRNYKDY** **TYLDGIKLSQAQIDIIEKEEDQLLIEGYAGTGKSLTL** **IYKFINVLVREDGKRVLYVTFNDTLIEDTKKRLSYCN** **EYNENKERHHVEICTFHEIASNILKKKKIIDRGIEKL** **TAKKIEDYKGAALRRIAGILARYIEGGKYYSELPKEE** **RLYKTHDENFIREEVAWIKAMGFIEKEKYFEKDRIGR** **SKSIRLTRSQRKTIFKIFEKYCEEQENKFFKSLDLED** **YALKLIQNIDNFDDLKFDYIFVDEVQDLDPMQIKALC** **LLTNTSIVLSGDANQRIYKKSPVKYEELGLRIKEKGK** **RKILNKNYRSTGEIVKLANSIKFFDESINKYNEKQFV** **KSGDRPIIRKVNDKKGAVKFLIGEIKKIHEEDPYKTI** **AIIHREKNELIGFQKSEFRKYLEGQLYMEKFSDIKSF** **ESKFDLREKNQVFYTNGYDVKGLEFDVVFIINFNTAN** **YPLSKELKKIKDENDGKEMTLIKDDVLEFINREKRLL** **YVAMTRAKEKLYLVADCKNSNISSFIYDFNTKYYEAQ** **NFKKKEIEENYNRYKINMEREYGIIIEDDDSNNVKNN** **DTKQENKFNTESKEKGKDDIDKIKVFFINKGIEVVDN** **RDKSGCLWIVAGKEAIPLMKKFGVLGYNFIFIANGGR** **ASKNRPAWYLKNS**	

Table 7. Nucleic Acid Sequence of Exemplary Helicases

Helicase	Nucleic Acid Sequence	SEQ ID NO
Eco RecQ *Escherichia coli*	ATGGCACAGGCAGAAGTTCTGAACCTGGAATCCGGTG CTAAACAAGTATTACAGGAGACCTTCGGTTATCAGCA GTTCCGTCCCGGACAAGAAGAAATTATTGATACCGTA CTGTCCGGTCGTGATTGTTTGGTAGTCATGCCAACTG GTGGAGGAAAGAGCCTGTGCTATCAAATCCCTGCCTT ATTATTGAATGGGTTAACGGTAGTCGTATCACCATTA ATTTCTTTGATGAAGGATCAAGTTGATCAGCTTCAGG CGAATGGTGTAGCAGCTGCATGCCTTAATAGTACCCA AACACGCGAGCAACAGTTAGAAGTGATGACAGGTTGT CGTACGGGCCAAATTCGCCTGTTGTACATCGCCCCCG AACGTCTGATGCTGGACAATTTTTTAGAGCACCTGGC TCACTGGAATCCAGTTTTGCTGGCGGTGGACGAGGCA CACTGTATCAGTCAGTGGGGGCACGACTTCCGCCCTG AGTATGCTGCCCTGGGTCAGTTGCGTCAGCGTTTTCC TACCCTGCCTTTTATGGCTCTGACGGCGACTGCTGAC GACACAACTCGTCAGGATATCGTACGCCTGTTAGGAT TGAATGACCCACTGATCCAGATCAGTTCGTTTGACCG CCCAAATATCCGCTATATGTTAATGGAAAAATTTAAA CCCTTGGATCAATTAATGCGCTACGTACAAGAGCAGC GTGGTAAGAGCGGCATTATTTACTGTAACAGTCGCGC GAAGGTTGAGGACACAGCGGCACGCCTGCAGAGCAAA GGCATTTCAGCGGCGGCATACCATGCAGGTTTGGAGA ACAATGTACGCGCAGACGTTCAGGAGAAGTTCCAGCG CGATGATTTGCAGATCGTTGTGGCCACTGTAGCGTTC GGTATGGGGATCAACAAACCTAATGTACGTTTCGTTG	60

	TCCACTTTGACATCCCACGCAATATTGAGAGCTACTA TCAAGAGACCGGACGCGCAGGGCGTGATGGTTTACCA GCCGAGGCCATGTTGTTCTACGATCCGGCTGATATGG CCTGGCTGCGTCGCTGTTTGGAGGAAAAACCTCAAGG TCAGTTGCAAGACATCGAACGCCACAAATTAAATGCT ATGGGTGCGTTTGCCGAAGCTCAAACATGCCGTCGCT TAGTTTTACTTAATTATTTTGGTGAGGGGCGTCAGGA GCCGTGTGGTAATTGCGATATTTGCTTGGACCCTCCT AAACAATATGACGGGTCAACAGACGCCCAGATTGCGT TATCGACTATTGGACGCGTCAATCAGCGTTTTGGTAT GGGGTACGTGGTCGAAGTAATTCGTGGAGCAAATAAC CAACGTATCCGTGATTATGGGCACGATAAACTGAAAG TATACGGTATGGGTCGCGATAAGAGTCATGAGCACTG GGTGTCAGTCATCCGCCAATTAATTCACCTTGGTCTG GTTACACAAAACATCGCGCAACACTCTGCACTGCAGC TTACTGAAGCCGCTCGTCCTGTATTGCGTGGTGAGAG CAGTCTGCAGTTGGCCGTGCCCCGCATTGTGGCCTTG AAACCAAAAGCCATGCAGAAAAGCTTTGGGGGAAATT ATGATCGCAAATTGTTTGCCAAGCTTCGCAAACTGCG CAAATCAATCGCGGATGAGTCAAACGTACCACCGTAT GTTGTCTTCAATGACGCAACTTTAATCGAGATGGCGG AGCAAATGCCAATCACAGCTTCAGAGATGCTGAGTGT AAATGGCGTTGGCATGCGCAAGCTTGAGCGCTTCGGA AAGCCGTTCATGGCATTAATTCGCGCCCACGTCGATG GGGATGACGAGGAGTAGTGA	
Eco RecQ *Escherichia coli* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: EcoReq**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCGCTAGCA**ATGGCACAGGCAGAAGTT** **CTGAACCTGGAATCCGGTGCTAAACAAGTATTACAGG** **AGACCTTCGGTTATCAGCAGTTCCGTCCCGGACAAGA** **AGAAATTATTGATACCGTACTGTCCGGTCGTGATTGT**	61

	TTGGTAGTCATGCCAACTGGTGGAGGAAAGAGCCTGT GCTATCAAATCCCTGCCTTATTATTGAATGGGTTAAC GGTAGTCGTATCACCATTAATTTCTTTGATGAAGGAT CAAGTTGATCAGCTTCAGGCGAATGGTGTAGCAGCTG CATGCCTTAATAGTACCCAAACACGCGAGCAACAGTT AGAAGTGATGACAGGTTGTCGTACGGGCCAAATTCGC CTGTTGTACATCGCCCCCGAACGTCTGATGCTGGACA ATTTTTTAGAGCACCTGGCTCACTGGAATCCAGTTTT GCTGGCGGTGGACGAGGCACACTGTATCAGTCAGTGG GGGCACGACTTCCGCCCTGAGTATGCTGCCCTGGGTC AGTTGCGTCAGCGTTTTCCTACCCTGCCTTTTATGGC TCTGACGGCGACTGCTGACGACACAACTCGTCAGGAT ATCGTACGCCTGTTAGGATTGAATGACCCACTGATCC AGATCAGTTCGTTTGACCGCCCAAATATCCGCTATAT GTTAATGGAAAAATTTAAACCCTTGGATCAATTAATG CGCTACGTACAAGAGCAGCGTGGTAAGAGCGGCATTA TTTACTGTAACAGTCGCGCGAAGGTTGAGGACACAGC GGCACGCCTGCAGAGCAAAGGCATTTCAGCGGCGGCA TACCATGCAGGTTTGGAGAACAATGTACGCGCAGACG TTCAGGAGAAGTTCCAGCGCGATGATTTGCAGATCGT TGTGGCCACTGTAGCGTTCGGTATGGGGATCAACAAA CCTAATGTACGTTTCGTTGTCCACTTTGACATCCCAC GCAATATTGAGAGCTACTATCAAGAGACCGGACGCGC AGGGCGTGATGGTTTACCAGCCGAGGCCATGTTGTTC TACGATCCGGCTGATATGGCCTGGCTGCGTCGCTGTT TGGAGGAAAAACCTCAAGGTCAGTTGCAAGACATCGA ACGCCACAAATTAAATGCTATGGGTGCGTTTGCCGAA GCTCAAACATGCCGTCGCTTAGTTTTACTTAATTATT TTGGTGAGGGGCGTCAGGAGCCGTGTGGTAATTGCGA TATTTGCTTGGACCCTCCTAAACAATATGACGGGTCA ACAGACGCCCAGATTGCGTTATCGACTATTGGACGCG TCAATCAGCGTTTTGGTATGGGGTACGTGGTCGAAGT AATTCGTGGAGCAAATAACCAACGTATCCGTGATTAT GGGCACGATAAACTGAAAGTATACGGTATGGGTCGCG ATAAGAGTCATGAGCACTGGGTGTCAGTCATCCGCCA ATTAATTCACCTTGGTCTGGTTACACAAAACATCGCG CAACACTCTGCACTGCAGCTTACTGAAGCCGCTCGTC CTGTATTGCGTGGTGAGAGCAGTCTGCAGTTGGCCGT GCCCCGCATTGTGGCCTTGAAACCAAAAGCCATGCAG AAAAGCTTTGGGGGAAATTATGATCGCAAATTGTTTG CCAAGCTTCGCAAACTGCGCAAATCAATCGCGGATGA GTCAAACGTACCACCGTATGTTGTCTTCAATGACGCA ACTTTAATCGAGATGGCGGAGCAAATGCCAATCACAG CTTCAGAGATGCTGAGTGTAAATGGCGTTGGCATGCG CAAGCTTGAGCGCTTCGGAAAGCCGTTCATGGCATTA ATTCGCGCCCACGTCGATGGGGATGACGAGGAGTAGT GA	
Tth UvrD	ATGAAACATCACCATCACCATCACAACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA	62

Thermus Thermophilus	GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCGCTAGCaATGTCTGACGCCTTGCTG GCACCATTAAACGAGGCACAACGCCAAGCCGTCCTGC ATTTTGAGGGTCCAGCATTAGTAGTGGCAGGGGCCGG ATCGGGGAAGACGCGTACCGTGGTTCACCGCGTCGCA TATCTGGTGGCCCGCCGTGGCGTGTTCCCATCCGAGA TTCTGGCGGTGACATTCACAAATAAGGCAGCCGAGGA GATGCGTGAACGCTTGCGTGGCTTAGTCCCTGGAGCC GGAGAAGTCTGGGTTTCGACTTTCCATGCTGCAGCGC TGCGTATCTTACGCGTATACGGAGAACGCGTGGGCCT GCGTCCCGGGTTCGTCGTATACGATGAGGATGACCAG ACAGCATTATTGAAGGAGGTGCTGAAAGAACTGGCTC TTTCGGCACGTCCCGGGCCGATTAAGGCATTGTTAGA CCGCGCCAAGAATCGTGGTGTTGGCCTGAAAGCCTTA CTGGGGGAACTTCCCGAGTACTACGCTGGGTTATCGC GCGGTCGTCTGGGAGACGTGCTGGTACGTTACCAGGA AGCCCTGAAGGCTCAAGGGGCTTTAGATTTCGGCGAC ATTTTGTTGTATGCTCTTGAAGCGTTCCGTGGAGGAC GCGGTGGTCCGCAGGCCCGCGCGCAACGTGCACGTTT CATCCATGTGGATGAGTACCAGGACACCTCGCCGGTT CAGTATCGTTTTACCCGTCTTTTGGCCGGTGAAGAAG CAAACCTTATGGCTGTAGGAGACCCCGATCAAGGGAT TTACTCTTTCCGCGCAGCGGATATTAAGAACATTTTA GACTTCACACGTGATTATCCTGAGGCACGTGTATATC GTCTTGAAGAGAACTATCGTTCGACCGAAGCCATTCT GCGTTTCGCCAACGCCGTAATCGTCAAAAACGCGCTT CGCTTGGAGAAAGCCTTACGCCCCGTCAAACGTGGGG GAGAGCCTGTCCGCTTATATCGCGCAGAGGACGCACG CGAAGAAGCACGCTTTGTCGCAGAAGAGATTGCTCGT TTGGGACCCCCGTGGGATCGCTATGCAGTCTTATACC GCACTAATGCTCAAAGCCGCCTTCTGGAACAGGCGTT AGCAGGTCGTGGGATCCCCGCACGCGTCGTTGGAGGT	

	GTGGGTTTTTTCGAGCGTGCAGAGGTGAAGGACTTGT TGGCGTACGCTCGTTTGGCCTTGAATCCCTTGGATGC CGTGTCCCTTAAGCGCGTCCTGAACACTCCCCCACGC GGTATCGGACCAGCCACGTGGGCCCGCGTGCAGTTAC TTGCCCAAGAGAAAGGATTACCCCCCTGGGAGGCTCT TAAAGAAGCGGCACGCACCTTTTCTCGCCCAGAACCA CTGCGCCATTTCGTAGCCCTTGTTGAAGAGTTGCAAG ATTTAGTATTCGGGCCTGCCGAGGCTTTCTTTCGCCA CTTGCTGGAGGCGACTGATTACCCCGCCTACCTGCGT GAAGCGTACCCAGAAGATGCGGAAGACCGCTTGGAAA ATGTAGAAGAACTGTTGCGCGCCGCGAAAGAAGCGGA GGATCTTCAGGACTTCCTTGATCGTGTCGCACTGACT GCCAAGGCCGAGGAGCCGGCCGAAGCAGAAGGACGCG TTGCATTGATGACATTGCATAACGCAAAGGGGTTGGA GTTTCCAGTCGTTTTCCTGGTTGGCGTAGAGGAAGGG TTACTGCCCCACCGTAACTCGGTGTCGACGTTAGAAG GACTTGAAGAGGAACGTCGTTTGTTCTATGTCGGTAT CACCCGTGCTCAGGAACGTTTGTACCTGTCACATGCG GAAGAGCGCGAGGTTTATGGCCGCCGCGAGCCCGCGC GTCCGTCCCGCTTTCTTGAAGAGGTTGAAGAGGGTTT ATACGAAGTATACGACCCATATCGTCGCCCACCGTCA CCCCCTCCACATCGCCCTCGCCCGGGGGCATTTCGTG GAGGTGAACGCGTCGTACATCCGCGCTTTGGACCTGG CACAGTCGTGGCCGCGCAGGGTGACGAGGTTACGGTC CATTTTGAGGGTTTTGGTCTGAAACGCCTTTCATTAA AATATGCAGAGCTGAAACCAGCTTAGTGA	
Tth UvrD *Thermus Thermophilus* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Tht UvrD**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCGCTAGCA**ATGAAACATCACCATCAC** **CATCACAACACTAGTAGCAATTCCATGTCCCCTATAC** **TAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCAC**	63

	TCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAA GAGCATTTGTATGAGCGCGATGAAGGTGATAAATGGC GAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAA TCTTCCTTATTATATTGATGGTGATGTTAAATTAACA CAGTCTATGGCCATCATACGTTATATAGCTGACAAGC ACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGA GATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGA TACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTG AAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGA AATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA ACATATTTAAATGGTGATCATGTAACCCATCCTGACT TCATGTTGTATGACGCTCTTGATGTTGTTTTATACAT GGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTT TGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTG ATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCC TTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGAC CATCCTCCAACTAGTGGATCTGGTGGTGGTGGCGGAT GGATGAGCGAGAATCTTTATTTTCAGGGCGCGCTAGC aATGTCTGACGCCTTGCTGGCACCATTAAACGAGGCA CAACGCCAAGCCGTCCTGCATTTTGAGGGTCCAGCAT TAGTAGTGGCAGGGGCCGGATCGGGGAAGACGCGTAC CGTGGTTCACCGCGTCGCATATCTGGTGGCCCGCCGT GGCGTGTTCCCATCCGAGATTCTGGCGGTGACATTCA CAAATAAGGCAGCCGAGGAGATGCGTGAACGCTTGCG TGGCTTAGTCCCTGGAGCCGGAGAAGTCTGGGTTTCG ACTTTCCATGCTGCAGCGCTGCGTATCTTACGCGTAT ACGGAGAACGCGTGGGCCTGCGTCCCGGGTTCGTCGT ATACGATGAGGATGACCAGACAGCATTATTGAAGGAG GTGCTGAAAGAACTGGCTCTTTCGGCACGTCCCGGGC CGATTAAGGCATTGTTAGACCGCGCCAAGAATCGTGG TGTTGGCCTGAAAGCCTTACTGGGGGAACTTCCCGAG TACTACGCTGGGTTATCGCGCGGTCGTCTGGGAGACG TGCTGGTACGTTACCAGGAAGCCCTGAAGGCTCAAGG GGCTTTAGATTTCGGCGACATTTTGTTGTATGCTCTT GAAGCGTTCCGTGGAGGACGCGGTGGTCCGCAGGCCC GCGCGCAACGTGCACGTTTCATCCATGTGGATGAGTA CCAGGACACCTCGCCGGTTCAGTATCGTTTTACCCGT CTTTTGGCCGGTGAAGAAGCAAACCTTATGGCTGTAG GAGACCCCGATCAAGGGATTTACTCTTTCCGCGCAGC GGATATTAAGAACATTTTAGACTTCACACGTGATTAT CCTGAGGCACGTGTATATCGTCTTGAAGAGAACTATC GTTCGACCGAAGCCATTCTGCGTTTCGCCAACGCCGT AATCGTCAAAAACGCGCTTCGCTTGGAGAAAGCCTTA CGCCCCGTCAAACGTGGGGGAGAGCCTGTCCGCTTAT ATCGCGCAGAGGACGCACGCGAAGAAGCACGCTTTGT CGCAGAAGAGATTGCTCGTTTGGGACCCCCGTGGGAT CGCTATGCAGTCTTATACCGCACTAATGCTCAAAGCC GCCTTCTGGAACAGGCGTTAGCAGGTCGTGGGATCCC CGCACGCGTCGTTGGAGGTGTGGGTTTTTTCGAGCGT	

	GCAGAGGTGAAGGACTTGTTGGCGTACGCTCGTTTGG **CCTTGAATCCCTTGGATGCCGTGTCCCTTAAGCGCGT** **CCTGAACACTCCCCCACGCGGTATCGGACCAGCCACG** **TGGGCCCGCGTGCAGTTACTTGCCCAAGAGAAAGGAT** **TACCCCCCTGGGAGGCTCTTAAAGAAGCGGCACGCAC** **CTTTTCTCGCCCAGAACCACTGCGCCATTTCGTAGCC** **CTTGTTGAAGAGTTGCAAGATTTAGTATTCGGGCCTG** **CCGAGGCTTTCTTTCGCCACTTGCTGGAGGCGACTGA** **TTACCCCGCCTACCTGCGTGAAGCGTACCCAGAAGAT** **GCGGAAGACCGCTTGGAAAATGTAGAAGAACTGTTGC** **GCGCCGCGAAAGAAGCGGAGGATCTTCAGGACTTCCT** **TGATCGTGTCGCACTGACTGCCAAGGCCGAGGAGCCG** **GCCGAAGCAGAAGGACGCGTTGCATTGATGACATTGC** **ATAACGCAAAGGGGTTGGAGTTTCCAGTCGTTTTCCT** **GGTTGGCGTAGAGGAAGGGTTACTGCCCCACCGTAAC** **TCGGTGTCGACGTTAGAAGGACTTGAAGAGGAACGTC** **GTTTGTTCTATGTCGGTATCACCCGTGCTCAGGAACG** **TTTGTACCTGTCACATGCGGAAGAGCGCGAGGTTTAT** **GGCCGCCGCGAGCCCGCGCGTCCGTCCCGCTTTCTTG** **AAGAGGTTGAAGAGGGTTTATACGAAGTATACGACCC** **ATATCGTCGCCCACCGTCACCCCCTCCACATCGCCCT** **CGCCCGGGGGCATTTCGTGGAGGTGAACGCGTCGTAC** **ATCCGCGCTTTGGACCTGGCACAGTCGTGGCCGCGCA** **GGGTGACGAGGTTACGGTCCATTTTGAGGGTTTTGGT** **CTGAAACGCCTTTCATTAAAATATGCAGAGCTGAAAC** **CAGCTTAGTGA**	
Eco UvrD *Escherichia coli*	ATGGACGTTTCCTACTTGCTGGACTCGTTGAACGATA AGCAACGTGAGGCCGTTGCCGCGCCTCGTTCCAACTT ATTGGTGCTTGCCGGCGCAGGTTCCGGCAAGACACGC GTCTTAGTTCATCGCATCGCGTGGTTAATGAGCGTGG AGAATTGCTCACCGTATAGCATCATGGCAGTTACGTT TACTAACAAGGCGGCCGCAGAAATGCGTCACCGCATT GGACAACTGATGGGAACAAGCCAGGGAGGTATGTGGG TAGGGACTTTCCACGGCCTTGCGCACCGTCTTCTTCG CGCACACCACATGGATGCCAATCTGCCGCAGGACTTT CAGATCCTTGATTCGGAGGATCAGTTGCGCTTGCTGA AGCGCTTAATCAAAGCGATGAATTTAGATGAGAAGCA GTGGCCACCCCGTCAGGCAATGTGGTACATCAATTCG CAAAAGGATGAGGGTTTGCGCCCTCACCATATCCAGT CGTATGGCAATCCAGTCGAGCAAACATGGCAGAAAGT TTACCAGGCATATCAGGAGGCCTGTGATCGCGCAGGA TTAGTAGACTTCGCAGAGCTTCTTCTTCGCGCCCACG AGTTATGGCTGAATAAACCTCACATTTTACAACATTA CCGTGAGCGTTTTACGAATATTTTAGTGGATGAGTTC CAGGATACTAACAACATTCAGTACGCTTGGATCCGCT TACTTGCCGGAGATACGGGGAAAGTTATGATCGTTGG TGATGACGACCAGTCGATCTACGGCTGGCGTGGGGCA CAGGTAGAGAACATCCAACGCTTCTTAAACGACTTCC CTGGTGCTGAGACGATCCGCCTTGAACAGAATTACCG	64

	TTCTACAAGCAATATCCTGTCCGCAGCGAATGCCCTT ATTGAGAACAACAACGGGCGCCTTGGCAAGAAGTTGT GGACTGACGGAGCTGATGGCGAACCGATCTCTCTGTA TTGCGCATTCAATGAACTGGACGAGGCACGCTTCGTT GTCAATCGCATTAAGACTTGGCAGGATAACGGCGGTG CCTTGGCTGAGTGCGCTATTCTGTACCGTTCAAACGC CCAGAGCCGTGTGCTGGAGGAAGCGTTACTGCAGGCT TCTATGCCGTATCGCATTTACGGTGGTATGCGCTTTT TTGAACGTCAAGAGATTAAGGACGCGCTGTCTTATCT GCGTCTGATCGCTAACCGCAATGACGACGCCGCATTT GAGCGTGTCGTCAATACCCCCACTCGCGGGATCGGGG ATCGCACACTGGACGTAGTCCGCCAAACAAGCCGCGA CCGTCAATTAACACTTTGGCAGGCGTGCCGTGAATTA CTTCAGGAAAAGGCATTGGCTGGTCGTGCCGCGAGCG CCCTTCAACGTTTTATGGAGCTTATCGACGCCCTGGC ACAAGAGACTGCAGACATGCCATTGCACGTACAGACT GACCGTGTGATTAAGGACAGCGGGCTGCGTACAATGT ATGAGCAAGAGAAAGGAGAGAAAGGGCAGACACGCAT TGAGAACTTAGAAGAATTGGTAACGGCGACTCGTCAA TTCTCCTACAACGAAGAAGATGAGGATTTAATGCCTC TTCAGGCGTTCTTAAGTCATGCTGCGTTGGAAGCAGG AGAAGGACAAGCTGATACCTGGCAAGACGCAGTCCAG CTTATGACTTTGCATTCAGCGAAGGGCTTGGAATTTC CGCAAGTTTTTATCGTCGGCATGGAAGAAGGGATGTT TCCCTCCCAGATGAGTCTTGACGAAGGGGGACGTTTG GAAGAGGAACGTCGTTTGGCTTATGTCGGGGTGACAC GCGCAATGCAAAAGCTTACTCTGACCTATGCAGAAAC CCGTCGCTTGTACGGGAAAGAAGTCTATCATCGTCCC AGCCGTTTCATTGGCGAGCTGCCCGAAGAATGTGTCG AAGAGGTACGCCTTCGTGCCACCGTATCTCGCCCGGT GTCTCACCAACGTATGGGGACGCCTATGGTAGAAAAT GACTCCGGTTACAAGTTGGGTCAACGTGTCCGCCATG CCAAGTTCGGCGAGGGGACCATTGTCAATATGGAAGG AAGCGGCGAACACTCGCGCTTGCAAGTGGCTTTCCAA GGACAGGGCATCAAATGGCTGGTTGCCGCCTATGCTC GCTTGGAGAGTGTGTAGTGA	
Eco UvrD *Escherichia coli* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Eco UvrD**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC	65

	ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA TTTTCAGGGCGCGCTAGCA**ATGGACGTTTCCTACTTG** **CTGGACTCGTTGAACGATAAGCAACGTGAGGCCGTTG** **CCGCGCCTCGTTCCAACTTATTGGTGCTTGCCGGCGC** **AGGTTCCGGCAAGACACGCGTCTTAGTTCATCGCATC** **GCGTGGTTAATGAGCGTGGAGAATTGCTCACCGTATA** **GCATCATGGCAGTTACGTTTACTAACAAGGCGGCCGC** **AGAAATGCGTCACCGCATTGGACAACTGATGGGAACA** **AGCCAGGGAGGTATGTGGGTAGGGACTTTCCACGGCC** **TTGCGCACCGTCTTCTTCGCGCACACCACATGGATGC** **CAATCTGCCGCAGGACTTTCAGATCCTTGATTCGGAG** **GATCAGTTGCGCTTGCTGAAGCGCTTAATCAAAGCGA** **TGAATTTAGATGAGAAGCAGTGGCCACCCCGTCAGGC** **AATGTGGTACATCAATTCGCAAAAGGATGAGGGTTTG** **CGCCCTCACCATATCCAGTCGTATGGCAATCCAGTCG** **AGCAAACATGGCAGAAAGTTTACCAGGCATATCAGGA** **GGCCTGTGATCGCGCAGGATTAGTAGACTTCGCAGAG** **CTTCTTCTTCGCGCCCACGAGTTATGGCTGAATAAAC** **CTCACATTTTACAACATTACCGTGAGCGTTTTACGAA** **TATTTTAGTGGATGAGTTCCAGGATACTAACAACATT** **CAGTACGCTTGGATCCGCTTACTTGCCGGAGATACGG** **GGAAAGTTATGATCGTTGGTGATGACGACCAGTCGAT** **CTACGGCTGGCGTGGGGCACAGGTAGAGAACATCCAA** **CGCTTCTTAAACGACTTCCCTGGTGCTGAGACGATCC** **GCCTTGAACAGAATTACCGTTCTACAAGCAATATCCT** **GTCCGCAGCGAATGCCCTTATTGAGAACAACAACGGG** **CGCCTTGGCAAGAAGTTGTGGACTGACGGAGCTGATG** **GCGAACCGATCTCTCTGTATTGCGCATTCAATGAACT** **GGACGAGGCACGCTTCGTTGTCAATCGCATTAAGACT** **TGGCAGGATAACGGCGGTGCCTTGGCTGAGTGCGCTA** **TTCTGTACCGTTCAAACGCCCAGAGCCGTGTGCTGGA** **GGAAGCGTTACTGCAGGCTTCTATGCCGTATCGCATT** **TACGGTGGTATGCGCTTTTTTGAACGTCAAGAGATTA** **AGGACGCGCTGTCTTATCTGCGTCTGATCGCTAACCG** **CAATGACGACGCCGCATTTGAGCGTGTCGTCAATACC** **CCCACTCGCGGGATCGGGGATCGCACACTGGACGTAG** **TCCGCCAAACAAGCCGCGACCGTCAATTAACACTTTG** **GCAGGCGTGCCGTGAATTACTTCAGGAAAAGGCATTG** **GCTGGTCGTGCCGCGAGCGCCCTTCAACGTTTTATGG** **AGCTTATCGACGCCCTGGCACAAGAGACTGCAGACAT** **GCCATTGCACGTACAGACTGACCGTGTGATTAAGGAC** **AGCGGGCTGCGTACAATGTATGAGCAAGAGAAAGGAG** **AGAAAGGGCAGACACGCATTGAGAACTTAGAAGAATT**	

	GGTAACGGCGACTCGTCAATTCTCCTACAACGAAGAA GATGAGGATTTAATGCCTCTTCAGGCGTTCTTAAGTC ATGCTGCGTTGGAAGCAGGAGAAGGACAAGCTGATAC CTGGCAAGACGCAGTCCAGCTTATGACTTTGCATTCA GCGAAGGGCTTGGAATTTCCGCAAGTTTTTATCGTCG GCATGGAAGAAGGGATGTTTCCCTCCCAGATGAGTCT TGACGAAGGGGGACGTTTGGAAGAGGAACGTCGTTTG GCTTATGTCGGGGTGACACGCGCAATGCAAAAGCTTA CTCTGACCTATGCAGAAACCCGTCGCTTGTACGGGAA AGAAGTCTATCATCGTCCCAGCCGTTTCATTGGCGAG CTGCCCGAAGAATGTGTCGAAGAGGTACGCCTTCGTG CCACCGTATCTCGCCCGGTGTCTCACCAACGTATGGG GACGCCTATGGTAGAAAATGACTCCGGTTACAAGTTG GGTCAACGTGTCCGCCATGCCAAGTTCGGCGAGGGGA CCATTGTCAATATGGAAGGAAGCGGCGAACACTCGCG CTTGCAAGTGGCTTTCCAAGGACAGGGCATCAAATGG CTGGTTGCCGCCTATGCTCGCTTGGAGAGTGTGTAGT GA**	
HEL#100 *Clostridium perfringens*	ATGGTGCTTAACCCTAAGTACTCAATCGGAGTGTATT ACGATGAATTAGTCGAAGAGGATATTGAGAAAGTCTA TTCGTACCTGAGCCGTGGAATCGTGGTACATTTATTT TTGCGTGGCATTTTAAAGGAAGAGCTGGAATTGAATG AGTATGATTTGAATACATTCAAGCTGCCGAAAGACAA TAACTTACTGTTTGTGTACGAGGAAGAGACCAGTTTG TCTTCCGAAAACATCATCATCTTTGTCGATAACAACA TTCTGAACAAGGAGGCGTATAAGAACATCACCGAAAA TCGCGAGTGCGAGTTCAACAAAGACCAATATGAGATT ATTACGGCGCCTGTAGATGATAACATCATTGTGACAA GCGGCGCAGGAACCGGAAAGACAACAACCATGATCAA CCGCCTTATTTATTTACGCTCCGTGATGTCAGACTTT ACGTTTGACCAAGCGGTGTTAATCACTTTCACTAACA AAGCATCGATTGAAATGAAAGAACGCCTTTTGGAAGT GCTGGATAAGTATTTCCGCGTCACAAACGACATTAAA TACTTGGACTATATGGAGGAAGCCGCAAAGGGGTCCA TCAGCACTATTCACAAATTTGCCAAGAAGATTCTTAA CAAGTCCGGACGTCATATTGGGATCAACAAAGACATT AACGTGCGCTCGTTCAAGTACAAGCGTCAGGAGGCCG TCAACAACGCCCTGAATAAAATCTATAAGGAAGAGTC TGAGCTGTTTTCCCTGATCAAATACTACCCAATCTAT GAAGTCGAACGTGTTATCTTAAAAATGTGGGAAATCT TAGACAATTACTCGATTGATCTTTTATCAAACAAAGT GCGTGTCGACTTCAATTTTGAGGAGGATAAGTTCACA GAGCTTATTAGCAAAACTTTAAAGTACGCACAGGAGA TTTTGGATTATGATAAAGAGAACGAGTTAGAGATCTC AGACTTGATGAAGAAATTAGCTTACGAAGATATTTTT AAGGGGATCGACAGTACGTACAAAGTGATTATGATCG ACGAATTTCAGGATAGCGACAACACCCAAATTGAGTT TATTTCTGAATTGGAAAAAAAAACAGGAGCCCGCATC TTGGTTGTGGGAGACGAAAAGCAATCAATTTACCGCT	66

	TCCGCGGGGCAGAATATACAGCATTCGACAAATTGAA GAAGCTTTTATCAAATTCTAAGCGTGAAGTCAAGGAA TATGAGATGACACGCAATTATCGCACAAACTACAACA TCTTGAATGAGATTAATCGTATTTTTATTGAGGTCGA TAAAAAGTTAGAGTGCTTTAATTATAAAGAGAAGGAC TACATCTATAGCAATAAGGACAAAGATAATCCTAAAG AAATCACGTGTTTCAACGTTTCTGACAATCTTAAACG TAAAGAGTTCTTTGACGACCTTCTGGAGAACAAAAAG GAAGACGAATCAATTGCTGTCTTATTTCGCTCTAATT CTGACATTAAAGAGTTCAAAGAGTTCTGCGATCGCAA TAATATTCTTTGTATGGTTGATTCGACAGGAGGTTTT TATCGCCACGAAGCTGTACGCGACTTCTATATTATGA TTAAATCGATTATTGATGAGCGCAACAGTCGCACGAT GTACTCTTTCATCAATACACCGTACATTTTAGAAGAC ATCGACAAAAACATTATTTTGAACGGTAACTCCAAAG ACAAAAATGAGTTCCTTTACTACATTTTAGAAAAAAA TAACTGGAACTATTTCCGCGAGTCCAGTAACTTTAAG AACCCCATTATCCTGATTGACGAGATTATCGAAAAGT TAAAGCCGGTCAAAAACTATTACGTTAAGGTGCTTCT GGAGGCAAAGAAAAACCAGCATAATTATGTTAACATT GCGAAAATGAAGGCGCTGGAATACAAGCTTAATCTGG AACACTTAGTATTTATTCTTAAGAAAGAGTTTAGTGA GAATATTACTTCAATCGAACAGATTGAACAGTTTCTG AAAGTGAAGATCAGCACTGATAATCTTGTAGACGTAC GCAAGCCAAAGGATTACGAGAATGACTACATCCAATG TTCAACAGTTCATAAGGCGAAAGGTTTGGAGTATGAT TACGTTGTGCTGGACAAGTTGACGAATCGCTTTTTGT CTAATTCGCGTAAAGTTAACTTGATCTTAAAGCCCGA CGGAGACAAGTTGTTAATTGGATACAAAATCCGTTTG GGAGAAGACGAGTTCAAGAACAAGATCTACAGCGACA ATCTGAAATACGAGAAGAAAGAGATTAAGGGGGAGGA GGCACGCTTGTTATATGTTGCGTTGACCCGTTGCAAA AAGGGGATCTATCTGAATATGTCTGGCGAACTGGCGG CGACCGAGTCGCTTAACACCTGGAAAAGCCTGATTGG AGGCACTATTAATTATGTTTAATAG	
HEL#100 *Clostridium perfringens* Construct Sequence *Italicized: His tag* Underlined: GST **Bold: Hel#100**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTA GCAATTCCATGTCCCCTATACTAGGTTATTGGAAAAT TAAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAA TATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGC GCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGA ATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATT GATGGTGATGTTAAATTAACACAGTCTATGGCCATCA TACGTTATATAGCTGACAAGCACAACATGTTGGGTGG TTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAA GGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAA TTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGA TTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTC GAAGATCGTTTATGTCATAAAACATATTTAAATGGTG ATCATGTAACCCATCCTGACTTCATGTTGTATGACGC	67

	TCTTGATGTTGTTTTATACATGGACCCAATGTGCCTG GATGCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTA TTGAAGCTATCCCACAAATTGATAAGTACTTGAAATC CAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAA GCCACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAATCT TTATTTTCAGGGCGCGCTAGCA**ATGGTGCTTAACCCT** **AAGTACTCAATCGGAGTGTATTACGATGAATTAGTCG** **AAGAGGATATTGAGAAAGTCTATTCGTACCTGAGCCG** **TGGAATCGTGGTACATTTATTTTTGCGTGGCATTTTA** **AAGGAAGAGCTGGAATTGAATGAGTATGATTTGAATA** **CATTCAAGCTGCCGAAAGACAATAACTTACTGTTTGT** **GTACGAGGAAGAGACCAGTTTGTCTTCCGAAAACATC** **ATCATCTTTGTCGATAACAACATTCTGAACAAGGAGG** **CGTATAAGAACATCACCGAAAATCGCGAGTGCGAGTT** **CAACAAAGACCAATATGAGATTATTACGGCGCCTGTA** **GATGATAACATCATTGTGACAAGCGGCGCAGGAACCG** **GAAAGACAACAACCATGATCAACCGCCTTATTTATTT** **ACGCTCCGTGATGTCAGACTTTACGTTTGACCAAGCG** **GTGTTAATCACTTTCACTAACAAAGCATCGATTGAAA** **TGAAAGAACGCCTTTTGGAAGTGCTGGATAAGTATTT** **CCGCGTCACAAACGACATTAAATACTTGGACTATATG** **GAGGAAGCCG** **CAAAGGGGTCCATCAGCACTATTCACAAATTTGCCAA** **GAAGATTCTTAACAAGTCCGGACGTCATATTGGGATC** **AACAAAGACATTAACGTGCGCTCGTTCAAGTACAAGC** **GTCAGGAGGCCGTCAACAACGCCCTGAATAAAATCTA** **TAAGGAAGAGTCTGAGCTGTTTTCCCTGATCAAATAC** **TACCCAATCTATGAAGTCGAACGTGTTATCTTAAAAA** **TGTGGGAAATCTTAGACAATTACTCGATTGATCTTTT** **ATCAAACAAAGTGCGTGTCGACTTCAATTTTGAGGAG** **GATAAGTTCACAGAGCTTATTAGCAAAACTTTAAAGT** **ACGCACAGGAGATTTTGGATTATGATAAAGAGAACGA** **GTTAGAGATCTCAGACTTGATGAAGAAATTAGCTTAC** **GAAGATATTTTTAAGGGGATCGACAGTACGTACAAAG** **TGATTATGATCGACGAATTTCAGGATAGCGACAACAC** **CCAAATTGAGTTTATTTCTGAATTGGAAAAAAAAACA** **GGAGCCCGCATCTTGGTTGTGGGAGACGAAAAGCAAT** **CAATTTACCGCTTCCGCGGGGCAGAATATACAGCATT** **CGACAAATTGAAGAAGCTTTTATCAAATTCTAAGCGT** **GAAGTCAAGGAATATGAGATGACACGCAATTATCGCA** **CAAACTACAACATCTTGAATGAGATTAATCGTATTTT** **TATTGAGGTCGATAAAAAGTTAGAGTGCTTTAATTAT** **AAAGAGAAGGACTACATCTATAGCAATAAGGACAAAG** **ATAATCCTAAAGAAATCACGTGTTTCAACGTTTCTGA** **CAATCTTAAACGTAAAGAGTTCTTTGACGACCTTCTG** **GAGAACAAAAAGGAAGACGAATCAATTGCTGTCTTAT** **TTCGCTCTAATTCTGACATTAAAGAGTTCAAAGAGTT** **CTGCGATCGCAATAATATTCTTTGTATGGTTGATTCG**	

	ACAGGAGGTTTTTATCGCCACGAAGCTGTACGCGACT **TCTATATTATGATTAAATCGATTATTGATGAGCGCAA** **CAGTCGCACGATGTACTCTTTCATCAATACACCGTAC** **ATTTTAGAAGACATCGACAAAAACATTATTTTGAACG** **GTAACTCCAAAGACAAAAATGAGTTCCTTTACTACAT** **TTTAGAAAAAAATAACTGGAACTATTTCCGCGAGTCC** **AGTAACTTTAAGAACCCCATTATCCTGATTGACGAGA** **TTATCGAAAAGTTAAAGCCGGTCAAAAACTATTACGT** **TAAGGTGCTTCTGGAGGCAAAGAAAAACCAGCATAAT** **TATGTTAACATTGCGAAAATGAAGGCGCTGGAATACA** **AGCTTAATCTGGAACACTTAGTATTTATTCTTAAGAA** **AGAGTTTAGTGAGAATATTACTTCAATCGAACAGATT** **GAACAGTTTCTGAAAGTGAAGATCAGCACTGATAATC** **TTGTAGACGTACGCAAGCCAAAGGATTACGAGAATGA** **CTACATCCAATGTTCAACAGTTCATAAGGCGAAAGGT** **TTGGAGTATGATTACGTTGTGCTGGACAAGTTGACGA** **ATCGCTTTTTGTCTAATTCGCGTAAAGTTAACTTGAT** **CTTAAAGCCCGACGGAGACAAGTTGTTAATTGGATAC** **AAAATCCGTTTGGGAGAAGACGAGTTCAAGAACAAGA** **TCTACAGCGACAATCTGAAATACGAGAAGAAAGAGAT** **TAAGGGGGAGGAGGCACGCTTGTTATATGTTGCGTTG** **ACCCGTTGCAAAAAGGGGATCTATCTGAATATGTCTG** **GCGAACTGGCGGCGACCGAGTCGCTTAACACCTGGAA** **AAGCCTGATTGGAGGCACTATTAATTATGTTTAATAG**	
HEL#75 *Clostridium perfringens*	ATGCTGGGGCTGAATAATGAGTCCAAAGAGTTCTTTA AGGGCATTAGCCGCATTTGGAGAAATTACAAGGACTA CACCTACCTTGACGGGATTAAGCTGAGCCAGGCGCAG ATCGATATCATCGAGAAGGAGGAGGACCAATTGCTTA TAGAGGGCTACGCCGGCACCGGTAAGTCCCTGACCCT TATATACAAGTTCATTAACGTGCTGGTTCGGGAAGAT GGGAAGAGGGTGCTGTATGTGACTTTTAACGATACGC TGATCGAGGATACGAAGAAACGCCTTAGTTATTGCAA CGAGTACAACGAGAATAAAGAGAGGCACCACGTAGAG ATTTGCACATTCCATGAGATCGCCAGTAATATCCTGA AGAAAAAGAAGATCATAGACAGGGGTATTGAGAAACT GACGGCTAAAAAGATAGAAGATTACAAAGGTGCCGCT CTCCGCAGAATTGCGGGAATCCTGGCTAGGTACATCG AGGGGGGAAAGTATTATAGCGAGTTGCCTAAAGAGGA ACGCCTCTACAAGACACATGACGAGAACTTTATCAGG GAGGAGGTGGCCTGGATCAAGGCCATGGGCTTTATAG AAAAGGAGAAGTATTTCGAGAAAGATCGCATTGGGAG GTCCAAGAGTATCAGGCTGACGCGCTCACAACGCAAA ACTATATTCAAGATATTTGAAAAGTACTGCGAGGAGC AAGAAAACAAATTCTTCAAAAGCCTCGACTTGGAGGA TTACGCCCTGAAGCTCATCCAGAACATAGATAATTTC GATGACCTTAAGTTCGACTACATTTTTGTGGACGAGG TACAGGATCTCGATCCCATGCAAATTAAGGCGCTGTG TCTGCTGACCAATACGAGCATCGTGCTGTCAGGCGAC GCGAATCAGCGGATTTACAAGAAATCTCCCGTGAAGT	68

	ACGAGGAGCTCGGCCTCAGAATCAAAGAGAAGGGGAA ACGGAAAATTCTGAACAAGAACTATCGGTCCACGGGT GAGATTGTCAAGCTCGCGAACTCAATCAAGTTCTTCG ACGAGTCCATCAATAAGTATAATGAAAAGCAGTTCGT AAAATCCGGTGATCGCCCGATCATCCGGAAGGTGAAC GACAAAAAGGGTGCGGTGAAGTTCCTGATCGGCGAGA TCAAGAAAATCCACGAAGAGGACCCCTACAAAACAAT CGCCATCATCCACCGAGAGAAAAACGAGCTTATCGGC TTCCAAAAGTCCGAGTTCCGAAAGTACCTGGAAGGCC AGCTGTACATGGAAAAATTCAGTGACATCAAGTCCTT TGAGTCAAAGTTTGATTTGAGGGAAAAGAACCAGGTG TTCTACACCAACGGCTACGATGTAAAGGGGCTGGAAT TTGATGTGGTGTTCATCATAAACTTCAACACGGCCAA CTACCCACTGAGTAAAGAGCTGAAGAAAATCAAGGAC GAAAACGACGGCAAGGAAATGACGCTCATTAAAGACG ATGTGCTCGAGTTTATCAATCGCGAGAAGAGGCTGCT GTACGTAGCTATGACCAGGGCCAAAGAAAAGCTGTAT CTCGTGGCCGACTGCAAAAACAGCAACATCAGCAGCT TCATCTACGACTTTAACACCAAGTACTATGAGGCACA AAATTTCAAGAAGAAAGAGATAGAGGAGAACTACAAC CGGTACAAGATTAACATGGAGCGCGAATACGGCATCA TCATTGAGGACGACGACTCCAACAACGTTAAGAACAA TGACACGAAACAAGAGAACAAGTTTAATACCGAATCT AAGGAAAAGGGCAAAGATGACATCGACAAGATAAAGG TGTTTTTCATCAACAAGGGAATCGAGGTGGTGGACAA CCGAGATAAGAGCGGGTGCTTGTGGATCGTCGCCGGG AAGGAAGCGATCCCTCTTATGAAGAAGTTCGGTGTCC TGGGCTATAACTTCATATTTATCGCAAACGGCGGTCG GGCATCTAAGAACCGGCCAGCCTGGTACCTCAAGAAT AGC	
HEL#75 *Clostridium perfringens* Construct Sequence *Italicized: His tag* Underlined: GST **Bold/Underlined: 2xSV40 NLS** **Bold: Hel#75**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAGCA ATTCCATGTCCCCTATACTAGGTTATTGGAAAATTAA GGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATAT CTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCG ATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATT GGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT GGTGATGTTAAATTAACACAGTCTATGGCCATCATAC GTTATATAGCTGACAAGCACAACATGTTGGGTGGTTG TCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGA GCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTG CATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAA GATCGTTTATGTCATAAAACATATTTAAATGGTGATC ATGTAACCCATCCTGACTTCATGTTGTATGACGCTCT TGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTG AAGCTATCCCACAAATTGATAAGTACTTGAAATCCAG CAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCC ACGTTTGGTGGTGGCGACCATCCTCCAACTAGTGGAT	69

CTGGTGGTGGTGGCGGATGGATGAGCGAGAATCTTTA
TTTTCAGGGCGCC**ATGCCTAAGAAAAAGCGGAAAGTT**
GAGGACCCCAAAAAGAAACGAAAAGTCGGAAGCGGCT
CA**CTGGGGCTGAATAATGAGTCCAAAGAGTTCTTTAA**
GGGCATTAGCCGCATTTGGAGAAATTACAAGGACTAC
ACCTACCTTGACGGGATTAAGCTGAGCCAGGCGCAGA
TCGATATCATCGAGAAGGAGGAGGACCAATTGCTTAT
AGAGGGCTACGCCGGCACCGGTAAGTCCCTGACCCTT
ATATACAAGTTCATTAACGTGCTGGTTCGGGAAGATG
GGAAGAGGGTGCTGTATGTGACTTTTAACGATACGCT
GATCGAGGATACGAAGAAACGCCTTAGTTATTGCAAC
GAGTACAACGAGAATAAAGAGAGGCACCACGTAGAGA
TTTGCACATTCCATGAGATCGCCAGTAATATCCTGAA
GAAAAAGAAGATCATAGACAGGGGTATTGAGAAACTG
ACGGCTAAAAAGATAGAAGATTACAAAGGTGCCGCTC
TCCGCAGAATTGCGGGAATCCTGGCTAGGTACATCGA
GGGGGGAAAGTATTATAGCGAGTTGCCTAAAGAGGAA
CGCCTCTACAAGACACATGACGAGAACTTTATCAGGG
AGGAGGTGGCCTGGATCAAGGCCATGGGCTTTATAGA
AAAGGAGAAGTATTTCGAGAAAGATCGCATTGGGAGG
TCCAAGAGTATCAGGCTGACGCGCTCACAACGCAAAA
CTATATTCAAGATATTTGAAAAGTACTGCGAGGAGCA
AGAAAACAAATTCTTCAAAAGCCTCGACTTGGAGGAT
TACGCCCTGAAGCTCATCCAGAACATAGATAATTTCG
ATGACCTTAAGTTCGACTACATTTTTGTGGACGAGGT
ACAGGATCTCGATCCCATGCAAATTAAGGCGCTGTGT
CTGCTGACCAATACGAGCATCGTGCTGTCAGGCGACG
CGAATCAGCGGATTTACAAGAAATCTCCCGTGAAGTA
CGAGGAGCTCGGCCTCAGAATCAAAGAGAAGGGGAAA
CGGAAAATTCTGAACAAGAACTATCGGTCCACGGGTG
AGATTGTCAAGCTCGCGAACTCAATCAAGTTCTTCGA
CGAGTCCATCAATAAGTATAATGAAAAGCAGTTCGTA
AAATCCGGTGATCGCCCGATCATCCGGAAGGTGAACG
ACAAAAAGGGTGCGGTGAAGTTCCTGATCGGCGAGAT
CAAGAAAATCCACGAAGAGGACCCCTACAAAACAATC
GCCATCATCCACCGAGAGAAAAACGAGCTTATCGGCT
TCCAAAAGTCCGAGTTCCGAAAGTACCTGGAAGGCCA
GCTGTACATGGAAAAATTCAGTGACATCAAGTCCTTT
GAGTCAAAGTTTGATTTGAGGGAAAAGAACCAGGTGT
TCTACACCAACGGCTACGATGTAAAGGGGCTGGAATT
TGATGTGGTGTTCATCATAAACTTCAACACGGCCAAC
TACCCACTGAGTAAAGAGCTGAAGAAAATCAAGGACG
AAAACGACGGCAAGGAAATGACGCTCATTAAAGACGA
TGTGCTCGAGTTTATCAATCGCGAGAAGAGGCTGCTG
TACGTAGCTATGACCAGGGCCAAAGAAAAGCTGTATC
TCGTGGCCGACTGCAAAAACAGCAACATCAGCAGCTT
CATCTACGACTTTAACACCAAGTACTATGAGGCACAA
AATTTCAAGAAGAAAGAGATAGAGGAGAACTACAACC
GGTACAAGATTAACATGGAGCGCGAATACGGCATCAT

	CATTGAGGACGACGACTCCAACAACGTTAAGAACAAT GACACGAAACAAGAGAACAAGTTTAATACCGAATCTA AGGAAAAGGGCAAAGATGACATCGACAAGATAAAGGT GTTTTTCATCAACAAGGGAATCGAGGTGGTGGACAAC CGAGATAAGAGCGGGTGCTTGTGGATCGTCGCCGGGA AGGAAGCGATCCCTCTTATGAAGAAGTTCGGTGTCCT GGGCTATAACTTCATATTTATCGCAAACGGCGGTCGG GCATCTAAGAACCGGCCAGCCTGGTACCTCAAGAATA GC	

IV. Linkers

[0293] In some embodiments, a linker is used herein to connect one component of a fusion polypeptide to another component of a fusion polypeptide. For example, a linker can be a polypeptide linker, such as a linker that is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more amino acids long. In some embodiments, the linker is a cleavable or non-cleavable linker. As described herein, two polypeptide sequences that are "fused" need not be directly adjacent to each other. Fused polypeptide sequences can be fused by a linker, or by an additional functional polypeptide sequence that is fused to the polypeptide sequences.

[0294] In some embodiments, a linker comprises glycine and serine amino acid residues. linker can comprise non-charged or charged amino acids. A linker can comprise alpha-helical domains. In some embodiments, a linker comprises a chemical cross linker. In some cases, a linker can be of different lengths to adjust the function of fused domains and their physical proximity. In some cases, a linker comprises peptides with ligand-inducible conformational changes.

[0295] Exemplary linkers are provided in Table 8. In some embodiments, the linker comprises a sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a linker in Table 8. In some embodiments, the linker comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOs: 70-72 or 140.

Table 8. Amino Acid Sequence of Exemplary Linkers

Linker	Amino Acid Sequence	SEQ ID NO
A	GGGGS	70
B	SGSGGGGS	71
C	SGSETPGTSESATPES	72

D	GSGSS	140

V. Nuclear Localization Signals (NLS)

[0296] In some embodiments, Ago fusion proteins described herein comprise at least 1, 2, 3, or 4 nuclear localization signal (NLS) polypeptides. In some embodiments, the Ago fusion protein comprises at least 1 NLS. In some embodiments, the Ago fusion protein comprises at least 2 NLS. In some embodiments, the Ago fusion protein comprises at least 3 NLS. In some embodiments, the Ago fusion protein comprises at least 4 NLS.

[0297] In some embodiments, the Ago fusion protein comprises at least 2 NLS, wherein each NLS is different. In some embodiments, the Ago fusion protein comprises at least 2 NLS, wherein each NSL is the same. In some embodiments, the Ago fusion protein comprises at least 3 NLS, wherein each NLS is different. In some embodiments, the Ago fusion protein comprises at least 3 NLS, wherein each NSL is the same. In some embodiments, the Ago fusion protein comprises at least 3 NLS, wherein two NLSs are the same and one is different. In some embodiments, at least one NLS is located between the Ago and another functional component (e.g., nucleic acid unwinding polypeptide) of the fusion polypeptide, optionally via one or more linkers.

[0298] In some embodiments, the NLS is derived from a microorganism. In some embodiments, the microorganism is a virus. In some embodiments, the NLS is an SV40 NLS.

[0299] Exemplary NLSs are provided in Table 9. In some embodiments, the NLS comprises a sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a linker in Table 9. In some embodiments, the linker comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOs: 73-78.

[0300] Exemplary NLS polypeptides are provided in Table 9.

Table 9. Amino Acid Sequence of Exemplary NLSs

NLS	Amino Acid Sequence	SEQ ID NO
SV40 Large T-antigen	PKKKRKV	73
2XSV40 Large T-antigen	PKKKRKVEDPKKKRKV	74
Nucleoplasmin (NPM)	KRPAATKKAGQAKKKK	75

c-Myc	PAAKRVKLD	76
EGL-13	MSRRRKANPTKLSENAKKLAKEVEN	77
TUS-protein	KLKIKRPVK	78

VI. Fusion Polypeptides

[0301] Described herein are fusion polypeptide constructs that comprise an Ago (e.g., an Ago described herein). Also described herein are nucleic acids encoding fusion polypeptide constructs comprising an Ago (e.g., an Ago described herein). In some embodiments, the fusion polypeptide comprises an Ago polypeptide that comprises an amino acid sequence having at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ IDs NO: 1-10 or 134-136. In some embodiments, the fusion polypeptide comprises a nucleic acid unwinding polypeptide. In some embodiments, the nucleic acid unwinding polypeptide is a helicase. In some embodiments, the nucleic acid unwinding polypeptide comprises a CRISPR associated (Cas) protein domain.

[0302] In some cases, the Ago polypeptide or Ago polypeptide fragment is fused to at least one additional element, for example a helicase. In some cases, the Ago polypeptide or Ago polypeptide fragment is fused to an ATPase. In some cases, the Ago polypeptide or Ago polypeptide fragment is fused to another Ago polypeptide or Ago polypeptide fragment. In some cases, the Ago polypeptide or Ago polypeptide fragment is fused with a guiding polynucleic acid or guiding protein. In some cases, the Ago polypeptide or Ago polypeptide fragment is a fusion construct of the Ago polypeptide or Ago polypeptide fragment and a nucleic acid unwinding polypeptide. In some cases, the Ago system comprises an Ago and a nucleic acid unwinding polypeptide fused together. In some cases, the Ago system comprises an Ago and a nucleic acid unwinding polypeptide, which are not fused together.

[0303] Fusion proteins can be synthesized using known technologies, for instance, recombination DNA technology where the coding sequences of various portions of the fusion proteins can be linked together at the nucleic acid level. Subsequently a fusion protein can be produced using a host cell. In some embodiments, a fusion protein comprises a cleavable or non-cleavable linker between the different sections or domains of the protein. For example, a linker can be a polypeptide linker, such as a linker that is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more amino acids long. As described herein, two polypeptide sequences that are "fused" need not be directly adjacent to each other. Fused polypeptide

sequences can be fused by a linker, or by an additional functional polypeptide sequence that is fused to the polypeptide sequences.

[0304] In some embodiments, a linker is a GSGSGS linker. In some cases, there are from 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10 linkers on a genome editing construct. For example, there can be from 1 to 10 GSGSGS linkers. linker can comprise non-charged or charged amino acids. A linker can comprise alpha-helical domains. In some embodiments, a linker comprises a chemical cross linker. In some cases, a linker can be of different lengths to adjust the function of fused domains and their physical proximity. In some cases, a linker comprises peptides with ligand-inducible conformational changes.

[0305] In some cases, a nucleic acid unwinding agent may be utilized with the Ago. A nucleic acid unwinding agent may be a polynucleic acid, protein, drug, or system that unwinds a nucleic acid. A nucleic acid unwinding agent can be energy. A nucleic acid unwinding agent can provide energy or heat. Unwinding can refer to the unwinding of a double helix (*e.g.*, of DNA) as well as to unwinding a double-stranded nucleic acid to convert it to a single-stranded nucleic acid or to unwinding DNA from histones. In some embodiments, an unwinding agent is a helicase. In some embodiments, helicases are enzymes that bind nucleic acid or nucleic acid protein complexes. In some embodiments, a helicase is a DNA helicase. In some embodiments, a helicase is an RNA helicase. In some embodiments, a helicase unwinds a polynucleic acid at any position. In some cases, a position that is unwound is found within an immune checkpoint gene. In some cases, a position of a nucleic acid that is unwound encodes a gene involved in disease. In some embodiments, an unwinding agent is an ATPase, helicase, synthetic associated helicase, or topoisomerase.

[0306] In some embodiments, a nucleic acid unwinding agent functions by breaking hydrogen bonds between nucleotide base pairs in double-stranded DNA or RNA. In some cases, unwinding a nucleic acid (*e.g.*, by breaking a hydrogen bond) requires energy. To break hydrogen bonds, nucleic acid unwinding agents can use energy stored in ATP. In some embodiments, a nucleic acid unwinding agent includes an ATPase. For example, in some embodiments, a polypeptide with nucleic acid unwinding activity comprises or be fused to an ATPase. In some embodiments, an ATPase is added to a cellular system.

[0307] In some embodiments, a nucleic acid unwinding agent is a polypeptide. For example, a nucleic acid unwinding peptide is of prokaryotic origin, archaeal origin, or eukaryotic origin. In some embodiments, a nucleic acid unwinding polypeptide comprises a helicase domain, a

topoisomerase domain, a Cas protein domain *e.g.*, a Cas protein domain selected from the group consisting of: Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx1S, Csf1, Csf2, CsO, Csf4, Cpf1, c2c1, c2c3, Cas9HiFi, xCas9, CasX, CasY, CasRX or a catalytically dead nucleic acid unwinding domain such as a dCas domain (*e.g.*, a dCas9 domain).

[0308] In some embodiments, a nucleic acid unwinding agent is a small molecule. For example, in some embodiments, a small molecule nucleic acid unwinding agent unwinds a nucleic acid through intercalation, groove binding or covalent binding to the nucleic acid, or a combination thereof. Exemplary small molecule nucleic acid unwinding agents include, but are not limited to, 9-aminoacridine, quinacrine, chloroquine, acriflavin, amsacrine, (Z)-3-(acridin-9-ylamino)-2-(5-chloro-1,3-benzoxazol-2-yl)prop-2-enal, small molecules that can stabilize quadruplex structures, quarfloxin, quindoline, quinoline-based triazine compounds, BRACO-19, acridines, pyridostatin, and derivatives thereof.

[0309] In some embodiments, the nucleic acid unwinding agent is a single strand DNA binding protein (SSB) polypeptide, e.g., as described herein. In some embodiments, In some embodiments, the SSB polypeptide comprises an SSB polypeptide described herein (or a functional fragment or functional variant thereof). In some embodiments, the SSB polypeptide comprises an SSB derived from a microorganism. In some embodiments, the microorganism is a bacterium. In some embodiments, the microorganism is a hyperthermophilic microorganism. In some embodiments, the SSB is from Saccharolobus solfataricus. In some embodiments, the SSB is active at a temperature between 32°C - 42°C. In some embodiments, the SSB is active at a temperature between 35°C - 40°C. In some embodiments, the SSB is active at about 37°C. In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 22-35. In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 22, 24, 26, 28, 30, 34, OR 34. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 36-49. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 36, 38, 40, 42, 44, 46, OR 48. In some

embodiments, the SSB polypeptide is one selected from Table 4. In some embodiments, the SSB is ET-SSB (Sso-SSB), Neq SSB, TaqSSB, TmaSSB, or EcoSSB. In some embodiments, the SSB is an ET-SSB (also referred to herein as Sso-SSB). In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 22. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NOS: 36.

[0310] In some embodiments, the nucleic acid unwinding agent is a helicase. In some embodiments, the helicase comprises a helicase polypeptide described herein (or a functional fragment or functional variant thereof). In some embodiments, the helicase polypeptide comprises a helicase derived from a microorganism. In some embodiments, the microorganism is a bacterium. In some embodiments, the microorganism is a hyperthermophilic microorganism. In some embodiments, the helicase is active at a temperature between 32°C - 42°C. In some embodiments, the helicase is active at a temperature between 35°C - 40°C. In some embodiments, the helicase is active at about 37°C. In some embodiments, the helicase comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 50-59. In some embodiments, the helicase comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 50, 52, 54, 56, or 58. In some embodiments, the helicase is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 60-69. In some embodiments, the helicase is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 60, 62, 64, or 68. In some embodiments, the helicase polypeptide is one selected from Table 6 or Table 7. In some embodiments, the helicase is Eco RecQ, Tth UvrD, Eco UvrD, HEL#100, HEL#75, or HEL#76.

[0311] In some embodiments, a polynucleic acid is unwound in a physical manner. A physical manner can include addition of heat or shearing for example. In some cases, a polynucleic acid such as DNA or RNA can be exposed to heat for nucleic acid unwinding. A DNA or RNA may denature at temperatures from about 50ºC to about 150ºC. DNA or RNA denatures from about 50 ºC to 60 ºC, from about 60 ºC to about 70 ºC, from about 70 ºC to about 80 ºC, from about 80 ºC to about 90 ºC, from about 90 ºC to about 100 ºC, from about 100 ºC to about 110 ºC, from

about 110 °C to about 120 °C, from about 120 °C to about 130 °C, from about 130 °C to about 140 °C, from about 140 °C to about 150 °C.

[0312] In some cases, a polynucleic acid can be denatured via changes in pH. For example, sodium hydroxide (NaOH) can be used to denature a polynucleic acid by increasing a pH to about 25 to about 29. In some cases, a polynucleic acid can be denatured via the addition of a salt.

[0313] In some cases, the disclosed editing system utilizing an unwinding agent can reduce a thermodynamic energetic requirement by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 40%, 50%, or up to about 60% as compared to a system that does not employ the disclosed unwinding agent. In some cases, the disclosed editing system utilizing an unwinding agent can reduce an immune response to the unwinding agent by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 40%, 50%, or up to about 60% as compared to a system that does not employ the disclosed unwinding agent. In some cases, an unwinding agent can be harvested from bacteria that are endogenously present in the human body to prevent eliciting an immune response.

VII. Ago-SSB Fusion Polypeptides

[0314] In one aspect, described herein are fusion polypeptides that comprises an Ago (or functional fragment or variant thereof) and a single strand DNA binding protein (SSB) described herein (or a functional fragment or variant thereof) (also referred to herein as an Ago-SSB fusion polypeptide).

[0315] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus Ago-SSB; SSB-Ago, Ago-linker-SSB; SSB-linker-Ago.

[0316] In some embodiments, the Ago-SSB fusion polypeptide comprises at least one nuclear localization signal polypeptide (NLS). In some embodiments, the Ago-SSB fusion polypeptide comprises at least two nuclear localization signal polypeptides. In some embodiments, the Ago-SSB fusion polypeptide comprises at least three nuclear localization signal polypeptides. In some embodiments, the Ago-SSB fusion polypeptide comprises at least four nuclear localization signal polypeptides. In some embodiments, the Ago-SSB fusion polypeptide comprises at least five nuclear localization signal polypeptides.

[0317] In some embodiments, wherein the Ago-SSB comprises two nuclear localization signal polypeptides, said nuclear localization signal polypeptides are the same. In some embodiments,

wherein the Ago-SSB comprises two nuclear localization signal polypeptides, said nuclear localization signal polypeptides are different. In some embodiments, wherein the Ago-SSB comprises three nuclear localization signal polypeptides, said nuclear localization signal polypeptides are the same. In some embodiments, wherein the Ago-SSB comprises three nuclear localization signal polypeptides, said nuclear localization signal polypeptides are different. In some embodiments, wherein the Ago-SSB comprises four nuclear localization signal polypeptides, said nuclear localization signal polypeptides are the same. In some embodiments, wherein the Ago-SSB comprises four nuclear localization signal polypeptides, said nuclear localization signal polypeptides are different.

[0318] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS-Ago-SSB; NLS-SSB-Ago; NLS-linker-Ago-SSB; NLS-linker-SSB-Ago, NLS-Ago-linker-SSB; NLS-SSB-linker-Ago; NLS-linker-Ago-linker-SSB; or NLS-linker-SSB-linker-Ago.

[0319] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-SSB, wherein NLS1 and NLS2 are different. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-SSB, wherein NLS1 and NLS2 are different.

[0320] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-Ago, wherein NLS1 and NLS2 are different.

[0321] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-Ago, wherein NLS1 and NLS2 are different.

[0322] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-SSB, wherein NLS1 and NLS2 are different.

[0323] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-linker-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-linker-SSB, wherein NLS1 and NLS2 are different.

[0324] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-Ago, wherein NLS1 and NLS2 are different.

[0325] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-linker-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-linker-Ago, wherein NLS1 and NLS2 are different.

[0326] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0327] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments,

the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0328] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0329] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0330] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-linker-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-linker-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-linker-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-

NLS3-Ago-linker-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-Ago-linker-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0331] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-linker-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-linker-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-linker-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-linker-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-Ago-linker-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0332] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-linker-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-linker-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-linker-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-linker-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-SSB-linker-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0333] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-linker-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-linker-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-linker-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In

some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-linker-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-linker-SSB-linker-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0334] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-Ago-SSB; wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-SSB-Ago, wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-Ago-linker-SSB, wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-SSB-linker-Ago, wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof.

[0335] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-linker-Ago-SSB; wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-linker-SSB-Ago, wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-linker-Ago-linker-SSB, wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-linker-SSB-linker-Ago, wherein each of NSL1, NSL2, NSL3, and NSL4 can be the same or different, or any combination thereof.

[0336] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-Ago-SSB, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-SSB-Ago, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N

to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-Ago-linker-SSB, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-SSB-linker-Ago, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof.

[0337] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-linker-Ago-SSB, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-linker-SSB-Ago, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-linker-Ago-linker-SSB, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-NLS3-NLS4-NLS5-linker-SSB-linker-Ago, wherein each of NSL1, NSL2, NSL3, NSL4, and NSL5 can be the same or different, or any combination thereof.

[0338] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-Ago-NLS2-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-Ago-NLS2-SSB, wherein NLS1 and NLS2 are different. In any of the embodiments described herein, any component may be linked to an adjacent component via a linker polypeptide. For example, NLS2 may be linker to Ago via a linker polypeptide. Multiple linkers may be used to connect different components of the polypeptide fusion. In embodiments, where fusion polypeptides contain multiple linkers, each linker may be the same or different, e.g., in a polypeptide fusion comprising three linkers two linkers may be the same and one different, all three may be the same, or all three may be different.

[0339] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-SSB-NLS2-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-SSB-NLS2-Ago, wherein NLS1 and NLS2 are different. In any of the embodiments described herein, any component may be linked to an adjacent component via a linker polypeptide. For example, NLS2 may be linker to Ago via a linker polypeptide. Multiple linkers may be used to connect different components of

the polypeptide fusion. In embodiments, where fusion polypeptides contain multiple linkers, each linker may be the same or different, e.g., in a polypeptide fusion comprising three linkers two linkers may be the same and one different, all three may be the same, or all three may be different.

[0340] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker-Ago-NLS2-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker-Ago-NLS2-SSB, wherein NLS1 and NLS2 are different.

[0341] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker-SSB-NLS2-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker-SSB-NLS2-Ago, wherein NLS1 and NLS2 are different.

[0342] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-Ago-linker-NLS2-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-Ago-linker-NLS2-SSB, wherein NLS1 and NLS2 are different.

[0343] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-Ago-linker1-NLS2-linker2-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-Ago-linker-NLS2-linker-SSB, wherein NLS1 and NLS2 are different. In any of the embodiments, described above, linker1 and linker 2 can be the same or different.

[0344] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-Ago-linker2-NLS2-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-Ago-linker2-NLS2-SSB, wherein NLS1 and NLS2 are different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0345] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-Ago-linker2-NLS2-linker3-SSB, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-Ago-linker2-NLS2-linker3-SSB, wherein NLS1 and NLS2 are different. In any of the embodiments described above, any of linker1, linker2, and linker3 can be the same or different.

[0346] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-SSB-linker-NLS2-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-SSB-linker-NLS2-Ago, wherein NLS1 and NLS2 are different.

[0347] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-SSB-linker1-NLS2-linker2-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-SSB-linker1-NLS2-linker2-Ago, wherein NLS1 and NLS2 are different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0348] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-SSB-linker2-NLS2-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-SSB-linker2-NLS2-Ago, wherein NLS1 and NLS2 are different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0349] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-SSB-linker2-NLS2-linker3-Ago, wherein NLS1 and NLS2 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-linker1-SSB-linker2-NLS2-linker3-Ago, wherein NLS1 and NLS2 are different. In any of the embodiments described above, any of linker1, linker2, and linker 3 can be the same or different.

[0350] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described herein, any component may be linked to an adjacent component via a linker polypeptide. For example, NLS2 may be linker to Ago via a linker polypeptide. Multiple linkers may be used to connect different components of the

polypeptide fusion. In embodiments, where fusion polypeptides contain multiple linkers, each linker may be the same or different, e.g., in a polypeptide fusion comprising three linkers two linkers may be the same and one different, all three may be the same, or all three may be different.

[0351] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-NLS3-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-NLS3-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-NLS3-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-Ago-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0352] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described herein, any component may be linked to an adjacent component via a linker polypeptide. For example, NLS2 may be linker to Ago via a linker polypeptide. Multiple linkers may be used to connect different components of the polypeptide fusion. In embodiments, where fusion polypeptides contain multiple linkers, each linker may be the same or different, e.g., in a polypeptide fusion comprising three linkers two linkers may be the same and one different, all three may be the same, or all three may be different.

[0353] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-NLS3-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-NLS3-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-NLS3-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker-SSB-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0354] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-NLS3-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-NLS3-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-NLS3-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0355] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-linker-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-linker-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-linker-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-linker-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments,

the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-NLS3-linker-SSB, wherein NLS1, NSL2, and NSL3 are each different.

[0356] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker1-NLS3-linker2-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker1-NLS3-linker2-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker1-NLS3-linker2-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker1-NLS3-linker2-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-Ago-linker1-NLS3-linker2-SSB, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0357] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-SSB, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0358] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-NLS3-linker2-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-NLS3-linker2-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-NLS3-

linker2-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-NLS3-linker2-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-NLS3-linker2-SSB, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0359] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-linker3-SSB, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-linker3-SSB, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-linker3-SSB, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-linker3-SSB, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-Ago-linker2-NLS3-linker3-SSB, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, any of linker1, linker2, and linker 3 can be the same or different.

[0360] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-NLS3-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-NLS3-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-NLS3-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0361] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-linker-Ago, wherein NLS1 and NLS2 are the same and NLS3 is

different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-linker-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-linker-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-linker-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-NLS3-linker-Ago, wherein NLS1, NSL2, and NSL3 are each different.

[0362] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker1-NLS3-linker2-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker1-NLS3-linker2-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker1-NLS3-linker2-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker1-NLS3-linker2-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-SSB-linker1-NLS3-linker2-Ago, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0363] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-

linker1-SSB-linker2-NLS3-Ago, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0364] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-NLS3-linker2-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-NLS3-linker2-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-NLS3-linker2-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-NLS3-linker2-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-NLS3-linker2-Ago, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, linker1 and linker 2 can be the same or different.

[0365] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-linker3-Ago, wherein NLS1 and NLS2 are the same and NLS3 is different from NSL1 and NLS2. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-linker3-Ago, wherein NLS1 and NLS3 are the same and NLS2 is different from NSL1 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-linker3-Ago, wherein NLS2 and NLS3 are the same and NLS1 is different from NSL2 and NLS3. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-linker3-Ago, wherein NLS1, NSL2, and NSL3 are the same. In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-linker3-Ago, wherein NLS1, NSL2, and NSL3 are each different. In any of the embodiments described above, any of linker1, linker2, and linker 3 can be the same or different. In some embodiments, linker2 and linker3 are the same and linker 1 is different. In some embodiments, linker1 and linker2 are the same and linker3 is different. In some embodiments, linker1 and linker3 are the same and linker2 is different.

[0366] In some embodiments, the Ago-SSB fusion polypeptide comprises from N to C terminus NLS1-NLS2-linker1-SSB-linker2-NLS3-linker3-Ago, wherein NLS1 and NLS2 are the same

and NLS3 is different from NSL1 and NLS2; and wherein linker2 and linker3 are the same and linker1 is different.

(a) SSB Polypeptides

[0367] In some embodiments, the SSB polypeptide component of an Ago-SSB fusion comprises an SSB polypeptide described herein (or a functional fragment or functional variant thereof). In some embodiments, the SSB polypeptide component of an Ago-SSB fusion comprises an SSB derived from a microorganism. In some embodiments, the microorganism is a bacterium. In some embodiments, the microorganism is a hyperthermophilic microorganism. In some embodiments, the SSB is from Saccharolobus solfataricus. In some embodiments, the SSB is active at a temperature between 32°C - 42°C. In some embodiments, the SSB is active at a temperature between 35°C - 40°C. In some embodiments, the SSB is active at about 37°C.

[0368] In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 22-35. In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 22, 24, 26, 28, 30, 32, or 34. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 36-49. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 36, 38, 40, 42, 44, or 48. In some embodiments, the SSB polypeptide is one selected from Table 4.

[0369] In some embodiments, the SSB is ET-SSB (Sso-SSB), Neq SSB, TaqSSB, TmaSSB, or EcoSSB. In some embodiments, the SSB is an ET-SSB (also referred to herein as Sso-SSB). In some embodiments, the SSB comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 22. In some embodiments, the SSB is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NOS: 36.

(b) Nuclear Localization Signals (NLS)

[0370] In some embodiments, Ago-SSB fusion proteins described herein comprise at least 1, 2, 3, or 4 nuclear localization signal (NLS) polypeptides. In some embodiments, the Ago-SSB fusion protein comprises at least 1 NLS. In some embodiments, the Ago-SSB fusion protein

comprises at least 2 NLS. In some embodiments, the Ago-SSB fusion protein comprises at least 3 NLS. In some embodiments, the Ago-SSB fusion protein comprises at least 4 NLS.

[0371] In some embodiments, the Ago-SSB fusion protein comprises at least 2 NLS, wherein each NLS is different. In some embodiments, the Ago-SSB fusion protein comprises at least 2 NLS, wherein each NSL is the same. In some embodiments, the Ago-SSB fusion protein comprises at least 3 NLS, wherein each NLS is different. In some embodiments, the Ago-SSB fusion protein comprises at least 3 NLS, wherein each NSL is the same. In some embodiments, the Ago-SSB fusion protein comprises at least 3 NLS, wherein two NLSs are the same and one is different. In some embodiments, at least one NLS is located between the Ago and SSB polypeptides of the fusion polypeptide, optionally via one or more linkers.

[0372] In some embodiments, the NLS is derived from a microorganism. In some embodiments, the microorganism is a virus. In some embodiments, the NLS is an SV40 NLS.

[0373] In some embodiments, the NLS comprises a sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a linker in Table 9. In some embodiments, the NLS comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOs: 73-78.

(c) Linkers

[0374] In some embodiments, Ago-SSB fusion proteins described herein comprise at least 1, 2, 3, 4, 5, or 6 linkers. In some embodiments, the linker is a linker described herein. In some embodiments in which a fusion construct has more than 1 linker, each linker may be the same or different from the other linkers, e.g., a in a fusion polypeptide construct have linker1, linker2, and linker3 – each of linker1, linker2, and linker3 can be the same (e.g., 100% sequence identity); each of linker1, linker2, and linker3 can be the same (e.g., less than 100% sequence identity); or two of linkers1-3 may be the same and the other different.

[0375] In some embodiments, a linker is used herein to connect one component of a fusion polypeptide to another component of a fusion polypeptide. For example, a linker can be a polypeptide linker, such as a linker that is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more amino acids long. In some embodiments, the linker is a cleavable or non-cleavable linker. As described herein, two polypeptide sequences that are "fused" need not be directly adjacent to each other. Fused polypeptide sequences can be fused

by a linker, or by an additional functional polypeptide sequence that is fused to the polypeptide sequences.

[0376] In some embodiments, a linker comprises glycine and serine amino acid residues. linker can comprise non-charged or charged amino acids. A linker can comprise alpha-helical domains. In some embodiments, a linker comprises a chemical cross linker. In some cases, a linker can be of different lengths to adjust the function of fused domains and their physical proximity. In some cases, a linker comprises peptides with ligand-inducible conformational changes.

[0377] Exemplary linkers include those described herein, e.g., Table 8, SEQ ID NOs: 70-72 or 140.

[0378] In some embodiments, the linker comprises a sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to a linker in Table 8. In some embodiments, the linker comprises an amino acid sequence with at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOs: 70-72 or 140.

(d) Exemplary Ago-SSB Fusion Polypeptides

[0379] In some embodiments, the Ago-SSB fusion protein comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 79-87. In some embodiments, the Ago-SSB fusion protein is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 88-96. The amino acid sequence of exemplary Ago-SSB fusion polypeptides are provided in Table 10. The nucleic acid sequence of exemplary Ago-SSB fusion polypeptides are provided in Table 11.

Table 10. Amino Acid Sequence of Exemplary Ago-SSB Fusion Polypeptides

Fusion Polypeptide	**Amino Acid Sequence**	**SEQ ID NO**
APO72 (also referred to herein is SSB-Ago69_v1) (See FIG. 68) N to C terminus: *Italicized: His Tag*	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGAM *PKKKRKVEDPKKKRKV*GSGSRLE**MEEKVGNLKPNMESVNVTV** **RVLEASEARQIQTKNGVRTISEAIVGDETGRVKLTLWGKHAG** **SIKEGQVVKIENAWTTAFKGQVQLNAGSKTKIAEASEDGFPE** **SSQIPENTPTAPQQMRGGGRGFRGGGRRYGRRGGRRQENEEG** **EEEGGGGS*MVGGYKVSNLTVEAFEGIGSVNPMLFYQYKVTGK*** ***GKYDNVYKIIKSARYKMHSKNRFKPVFIKDDKLYTLEKLPDI*** ***EDLDFANINFVKSEVLSIEDNMSIYGEVVEYYINLKLKKVKV***	79

Underlined: GST *Italicized/Underlined: 2XSV40NLS* **Bold: SsoSSB** **Underlined/Bold: G4S** ***Italicized/Bold: Ago69***	***LGKYPKYRINYSKEILSNTLLTRELKDEFKKSNKGFNLKRKF RISPVVNKMGKVILYLSCSADFSTNKNIYEMLKEGLEVEGLA VKSEWSNISGNLVIESVLETKISEPTSLGQSLIDYYKNNNQG YRVKDFTDEDLNANIVNVRGNKKIYMYIPHALKPIITREYLA KNDPEFSKEIEQLIKMNMNYRYETLKSFVNDIGVIEELNNLS FKNKYYEDVKLLGYSSGKIDEPVLMGAKGIIKNKMQIFSNGF YKLPEGKVRFGVLYPKEFDGVSRKAIRAIYDFSKEGKYHGES NKYIAEHLINVEFNPKECIFEGYELGDITEYKKAALKLNNYN NVDFVIAIVPNMSDEEIENSYNPFKKIWAELNLPSQMISVKT AEIFANSRDNTALYYLHNIVLGILGKIGGIPWVVKDMKGDVD CFVGLDVGTREKGIHYPACSVVFDKYGKLINYYKPNIPQNGE KINTEILQEIFDKVLISYEEENGAYPKNIVIHRDGFSREDLD WYENYFGKKNIKFNIIEVKKSTPLKIASINEGNITNPEKGSY ILRGNKAYMVTTDIKENLGSPKPLKIEKSYGDIDMLTALSQI YALTQIHVGATKSLRLPITTGYADKICKAIEFIPQGRVDNRL FFL***	
APO73 (See FIG. 68) N to C terminus: *Italicized: His Tag* Underlined: GST *Italicized/Underlined: 2XSV40NLS* **Bold: VP64** **Underlined/Bold: G4S** ***Italicized/Bold: Ago69***	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGAM *PKKKRKVEDPKKKRKV*GSGSRLE**MDALDDFDLDMLGSDALDD FDLDMLGSDALDDFDLDMLGSDALDDFDLDML**GGGGS***MVGGY KVSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKIIKSA RYKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVKS EVLSIEDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRINYSK EILSNTLLTRELKDEFKKSNKGFNLKRKFRISPVVNKMGKVI LYLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLV IESVLETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNA NIVNVRGNKKIYMYIPHALKPIITREYLAKNDPEFSKEIEQL IKMNMNYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVKLLG YSSGKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVL YPKEFDGVSRKAIRAIYDFSKEGKYHGESNKYIAEHLINVEF NPKECIFEGYELGDITEYKKAALKLNNYNNVDFVIAIVPNMS DEEIENSYNPFKKIWAELNLPSQMISVKTAEIFANSRDNTAL YYLHNIVLGILGKIGGIPWVVKDMKGDVDCFVGLDVGTREKG IHYPACSVVFDKYGKLINYYKPNIPQNGEKINTEILQEIFDK VLISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKNIKF NIIEVKKSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTD IKENLGSPKPLKIEKSYGDIDMLTALSQIYALTQIHVGATKS LRLPITTGYADKICKAIEFIPQGRVDNRLFFL***	80
APO46 (See FIG. 69B) N to C terminus:	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGAV SSPQGYPSL*MPKKKRKVEDPKKKRKV***GSGS*MVGGYKVSNLTV**	81

Italicized: His Tag Underlined: GST *Italicized/ Underlined: 2XSV40NLS* **Underlined/Bold: GSGS linker** ***Italicized/Bold: Ago69***	***EAFEGIGSVNPMLFYQYKVTGKGKYDNVYKIIKSARYKMHSK NRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIED NMSIYGEVVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTL LTRELKDEFKKSNKGFNLKRKFRISPVVNKMGKVILYLSCSA DFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVIESVLET KISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRG NKKIYMYIPHALKPIITREYLAKNDPEFSKEIEQLIKMNMNY RYETLKSFVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGKID EPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDG VSRKAIRAIYDFSKEGKYHGESNKYIAEHLINVEFNPKECIF EGYELGDITEYKKAALKLNNYNNVDFVIAIVPNMSDEEIENS YNPFKKIWAELNLPSQMISVKTAEIFANSRDNTALYYLHNIV LGILGKIGGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPACS VVFDKYGKLINYYKPNIPQNGEKINTEILQEIFDKVLISYEE ENGAYPKNIVIHRDGFSREDLDWYENYFGKKNIKFNIIEVKK STPLKIASINEGNITNPEKGSYILRGNKAYMVTTDIKENLGS PKPLKIEKSYGDIDMLTALSQIYALTQIHVGATKSLRLPITT GYADKICKAIEFIPQGRVDNRLFFL***	
APO25 (See FIG. 69B) N to C terminus: *Italicized: His Tag* Underlined: GST ***Bold/Italicized/Underlined: 2XSV40NLS*** **Underlined and Bold: GSGS linker** ***Italicized and Bold: Ago69*** *Italicized and underlined: GGGS linker* **Bold: V5 tag** *Italicized: NPM NLS*	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGAM ***PKKKRKVEDPKKKRKVGSGS***RLE***MVGGYKVSNLTVEAFEGIG SVNPMLFYQYKVTGKGKYDNVYKIIKSARYKMHSKNRFKPVF IKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMSIYGE VVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTRELKD EFKKSNKGFNLKRKFRISPVVNKMGKVILYLSCSADFSTNKN IYEMLKEGLEVEGLAVKSEWSNISGNLVIESVLETKISEPTS LGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIYMY IPHALKPIITREYLAKNDPEFSKEIEQLIKMNMNYRYETLKS FVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGKIDEPVLMGA KGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAIR AIYDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGD ITEYKKAALKLNNYNNVDFVIAIVPNMSDEEIENSYNPFKKI WAELNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGILGKI GGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDKYG KLINYYKPNIPQNGEKINTEILQEIFDKVLISYEEENGAYPK NIVIHRDGFSREDLDWYENYFGKKNIKFNIIEVKKSTPLKIA SINEGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLKIE KSYGDIDMLTALSQIYALTQIHVGATKSLRLPITTGYADKIC KAIEFIPQGRVDNRLFFL***TS*GGGGS*GKPIPNPLLGLDST*KRP AATKKAGQAKKKK*	82

APO71 (See FIG. 69B) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: 2XSV40NLS** **Underlined/Bold: GSGS linker** ***Italicized/Bold: Ago69***	MK*HHHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGA**M** ***PKKKRKVEDPKKKRKV*GSGS*RLE*MVGGYKVSNLTVEAFEGIG*** ***SVNPMLFYQYKVTGKGKYDNVYKIIKSARYKMHSKNRFKPVF*** ***IKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMSIYGE*** ***VVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTRELKD*** ***EFKKSNKGFNLKRKFRISPVVNKMGKVILYLSCSADFSTNKN*** ***IYEMLKEGLEVEGLAVKSEWSNISGNLVIESVLETKISEPTS*** ***LGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIYMY*** ***IPHALKPIITREYLAKNDPEFSKEIEQLIKMNMNYRYETLKS*** ***FVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGKIDEPVLMGA*** ***KGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAIR*** ***AIYDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGD*** ***ITEYKKAALKLNNYNNVDFVIAIVPNMSDEEIENSYNPFKKI*** ***WAELNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGILGKI*** ***GGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDKYG*** ***KLINYYKPNIPQNGEKINTEILQEIFDKVLISYEEENGAYPK*** ***NIVIHRDGFSREDLDWYENYFGKKNIKFNIIEVKKSTPLKIA*** ***SINEGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLKIE*** ***KSYGDIDMLTALSQIYALTQIHVGATKSLRLPITTGYADKIC*** ***KAIEFIPQGRVDNRLFFL***	83
SSB-AGO#69v3 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	MK*HHHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGA**M** **EEKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAI** **VGDETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQL** **NAGSKTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRG** **GGRRYGRRGGRRQENEEGEEE***SGSETPGTSESATPES* ***VGGYK*** ***VSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKIIKSAR*** ***YKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVKSE*** ***VLSIEDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRINYSKE*** ***ILSNTLLTRELKDEFKKSNKGFNLKRKFRISPVVNKMGKVIL*** ***YLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVI*** ***ESVLETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNAN*** ***IVNVRGNKKIYMYIPHALKPIITREYLAKNDPEFSKEIEQLI*** ***KMNMNYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVKLLGY*** ***SSGKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLY*** ***PKEFDGVSRKAIRAIYDFSKEGKYHGESNKYIAEHLINVEFN*** ***PKECIFEGYELGDITEYKKAALKLNNYNNVDFVIAIVPNMSD*** ***EEIENSYNPFKKIWAELNLPSQMISVKTAEIFANSRDNTALY*** ***YLHNIVLGILGKIGGIPWVVKDMKGDVDCFVGLDVGTREKGI*** ***HYPACSVVFDKYGKLINYYKPNIPQNGEKINTEILQEIFDKV***	84

	LISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKNIKFN IIEVKKSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTDI KENLGSPKPLKIEKSYGDIDMLTALSQIYALTQIHVGATKSL RLPITTGYADKICKAIEFIPQGRVDNRLFFL	
2SSB-AGO#69v1 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** **Underlined/Bold: G4S** *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGA**M EEKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAI VGDETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQL NAGSKTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRG GGRRYGRRGGRRQENEEGEEEGGGGSEEKVGNLKPNMESVNV TVRVLEASEARQIQTKNGVRTISEAIVGDETGRVKLTLWGKH AGSIKEGQVVKIENAWTTAFKGQVQLNAGSKTKIAEASEDGF PESSQIPENTPTAPQQMRGGGRGFRGGGRRYGRRGGRRQENE EGEEE*SGSETPGTSESATPES****VGGYKVSNLTVEAFEGIGSVN PMLFYQYKVTGKGKYDNVYKIIKSARYKMHSKNRFKPVFIKD DKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMSIYGEVVE YYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTRELKDEFK KSNKGFNLKRKFRISPVVNKMGKVILYLSCSADFSTNKNIYE MLKEGLEVEGLAVKSEWSNISGNLVIESVLETKISEPTSLGQ SLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIYMYIPH ALKPIITREYLAKNDPEFSKEIEQLIKMNMNYRYETLKSFVN DIGVIEELNNLSFKNKYYEDVKLLGYSSGKIDEPVLMGAKGI IKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAIRAIY DFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGDITE YKKAALKLNNYNNVDFVIAIVPNMSDEEIENSYNPFKKIWAE LNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGILGKIGGI PWVVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDKYGKLI NYYKPNIPQNGEKINTEILQEIFDKVLISYEEENGAYPKNIV IHRDGFSREDLDWYENYFGKKNIKFNIIEVKKSTPLKIASIN EGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLKIEKSY GDIDMLTALSQIYALTQIHVGATKSLRLPITTGYADKICKAI EFIPQGRVDNRLFFL***	85
2SSB-AGO#69v2 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB**	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGA**M EEKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAI VGDETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQL NAGSKTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRG GGRRYGRRGGRRQENEEGEEESGSGGGGSEEKVGNLKPNMES VNVTVRVLEASEARQIQTKNGVRTISEAIVGDETGRVKLTLW GKHAGSIKEGQVVKIENAWTTAFKGQVQLNAGSKTKIAEASE DGFPESSQIPENTPTAPQQMRGGGRGFRGGGRRYGRRGGRRQ	86

Underlined/Bold: SGSG4S *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	**ENEEGEEE***SGSETPGTSESATPES***VGGYKVSNLTVEAFEGIG** ***SVNPMLFYQYKVTGKGKYDNVYKIIKSARYKMHSKNRFKPVF*** ***IKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMSIYGE*** ***VVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTRELKD*** ***EFKKSNKGFNLKRKFRISPVVNKMGKVILYLSCSADFSTNKN*** ***IYEMLKEGLEVEGLAVKSEWSNISGNLVIESVLETKISEPTS*** ***LGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIYMY*** ***IPHALKPIITREYLAKNDPEFSKEIEQLIKMNMNYRYETLKS*** ***FVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGKIDEPVLMGA*** ***KGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAIR*** ***AIYDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGD*** ***ITEYKKAALKLNNYNNVDFVIAIVPNMSDEEIENSYNPFKKI*** ***WAELNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGILGKI*** ***GGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDKYG*** ***KLINYYKPNIPQNGEKINTEILQEIFDKVLISYEEENGAYPK*** ***NIVIHRDGFSREDLDWYENYFGKKNIKFNIIEVKKSTPLKIA*** ***SINEGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLKIE*** ***KSYGDIDMLTALSQIYALTQIHVGATKSLRLPITTGYADKIC*** ***KAIEFIPQGRVDNRLFFL***	
2SSB-AGO#69v3 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	MK*HHHHHH*NTSSNSMSPILGYWKIKGLVQPTRLLLEYLEEKY EEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMA IIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYS KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSS KYIAWPLQGWQATFGGGDHPPTSGSGGGGGWMSENLYFQGA**M** **EEKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAI** **VGDETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQL** **NAGSKTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRG** **GGRRYGRRGGRRQENEEGEEE***SGSETPGTSESATPES***MEEKV** **GNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAIVGDE** **TGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQLNAGS** **KTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRGGGRR** **YGRRGGRRQENEEGEEE***SGSETPGTSESATPES***VGGYKVSNL** ***TVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKIIKSARYKMH*** ***SKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVKSEVLSI*** ***EDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRINYSKEILSN*** ***TLLTRELKDEFKKSNKGFNLKRKFRISPVVNKMGKVILYLSC*** ***SADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVIESVL*** ***ETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNV*** ***RGNKKIYMYIPHALKPIITREYLAKNDPEFSKEIEQLIKMNM*** ***NYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGK*** ***IDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEF*** ***DGVSRKAIRAIYDFSKEGKYHGESNKYIAEHLINVEFNPKEC*** ***IFEGYELGDITEYKKAALKLNNYNNVDFVIAIVPNMSDEEIE*** ***NSYNPFKKIWAELNLPSQMISVKTAEIFANSRDNTALYYLHN*** ***IVLGILGKIGGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPA*** ***CSVVFDKYGKLINYYKPNIPQNGEKINTEILQEIFDKVLISY*** ***EEENGAYPKNIVIHRDGFSREDLDWYENYFGKKNIKFNIIEV*** ***KKSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTDIKENL***	87

	GSPKPLKIEKSYGDIDMLTALSQIYALTQIHVGATKSLRLPI TTGYADKICKAIEFIPQGRVDNRLFFL	

Table 11. Nucleic Acid Sequence of Exemplary Ago-SSB Fusion Polypeptides

Fusion Polypeptide	Nucleic Acid Sequence	SEQ ID NO
APO72 (also referred to herein is SSB-Ago69_v1) (See FIG. 68) N to C terminus: *Italicized: His Tag* Underlined: GST *Italicized/Underlined: 2XSV40NLS* **Bold: SsoSSB** **Underlined/Bold: G4S** ***Italicized/Bold: Ago69***	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCA*ATGCCCAAGAAAAAGCG* *AAAGGTAGAGGACCCCAAAAAGAAACGCAAAGTG*G GCTCCGGAAGCCGTCTCGAA**ATGGAAGAAAAAGTA** **GGCAACCTGAAGCCTAATATGGAATCCGTAAATGT** **AACCGTTCGCGTTTTAGAAGCCTCTGAAGCACGGC** **AGATCCAGACCAAAAATGGTGTTCGCACCATTTCA** **GAGGCGATTGTAGGGGATGAAACCGGGCGCGTGAA** **ACTGACTCTGTGGGGCAAACATGCGGGCAGCATCA** **AAGAAGGCCAGGTCGTTAAAATTGAGAACGCCTGG** **ACAACCGCGTTCAAAGGCCAGGTACAGCTGAATGC** **CGGTAGCAAGACCAAAATTGCCGAGGCATCTGAAG** **ACGGTTTCCCTGAAAGCAGCCAGATCCCAGAAAAT** **ACTCCTACGGCACCGCAGCAGATGCGTGGCGGTGG** **GCGGGGCTTTCGTGGCGGAGGCCGCCGTTATGGCC** **GTCGCGGTGGGCGCCGGCAAGAAAACGAAGAAGGC** **GAAGAAGAAGGCGGTGGTGGCTCA*ATGGTCGGCGG*** ***CTATAAAGTCAGCAATTTGACAGTGGAAGCGTTCG*** ***AAGGTATCGGGAGTGTCAACCCGATGCTGTTTTAC*** ***CAATACAAAGTCACCGGAAAGGGAAAGTACGATAA*** ***TGTGTATAAGATTATCAAAAGCGCACGGTACAAGA*** ***TGCATTCTAAGAACCGATTCAAGCCCGTGTTCATC***	88

	AAGGACGACAAACTGTACACCCTCGAGAAGCTCCC GGATATAGAGGACCTGGATTTCGCAAACATTAACT TCGTGAAAAGCGAGGTTCTCAGCATAGAGGATAAT ATGTCAATTTATGGCGAGGTGGTGGAATACTATAT CAATCTCAAGCTGAAAAAAGTGAAGGTGTTGGGAA AATACCCCAAGTACAGGATCAATTACAGCAAAGAG ATTCTCAGTAATACGCTGCTGACACGAGAGCTCAA AGACGAGTTTAAGAAATCAAATAAGGGTTTTAACC TGAAACGGAAGTTTAGAATTTCCCCCGTGGTGAAT AAGATGGGCAAAGTGATACTCTATTTGTCCTGCAG TGCTGATTTCAGCACCAACAAGAACATTTACGAAA TGTTGAAAGAGGGCTTGGAGGTTGAGGGGCTGGCC GTTAAGAGCGAGTGGAGCAATATCAGTGGCAACCT GGTGATCGAGAGCGTACTGGAAACCAAGATATCCG AGCCCACTAGCCTGGGCCAATCCCTGATAGACTAC TATAAGAATAACAACCAGGGCTATAGGGTGAAGGA TTTCACCGATGAGGATCTGAATGCCAACATTGTCA ACGTGAGAGGAAATAAGAAGATCTATATGTATATT CCGCACGCGTTGAAGCCGATAATCACCCGGGAGTA CCTGGCCAAGAACGATCCAGAGTTTTCTAAGGAGA TCGAGCAGCTTATCAAGATGAATATGAACTACCGA TATGAAACCCTCAAGTCATTTGTGAATGACATCGG GGTCATTGAGGAGCTGAACAACCTGAGCTTCAAAA ACAAATACTACGAAGATGTGAAACTGCTGGGTTAC TCCAGCGGCAAAATAGACGAACCCGTCCTGATGGG GGCAAAAGGGATCATAAAGAACAAAATGCAGATTT TTTCCAATGGATTCTACAAACTCCCCGAAGGCAAG GTACGATTTGGCGTTCTGTACCCAAAAGAATTTGA TGGCGTGTCAAGGAAAGCTATCCGCGCCATTTATG ACTTCAGTAAGGAGGGCAAATACCACGGCGAAAGC AACAAGTATATCGCGGAACACCTGATAAACGTGGA GTTCAATCCAAAGGAGTGCATATTTGAGGGATACG AACTGGGCGATATCACCGAATACAAGAAGGCGGCT CTGAAACTTAATAACTACAACAATGTCGACTTCGT AATCGCAATAGTCCCGAACATGTCCGACGAAGAGA TAGAGAACAGCTACAATCCGTTCAAGAAAATATGG GCCGAACTGAATCTGCCCAGCCAGATGATTAGCGT CAAGACGGCCGAAATCTTTGCCAATAGCAGGGATA ACACGGCGCTTTACTATCTGCATAACATCGTCCTC GGTATCCTGGGTAAGATAGGAGGGATTCCCTGGGT GGTTAAAGACATGAAGGGCGACGTGGATTGCTTCG TTGGACTCGATGTCGGCACCAGGGAGAAGGGCATA CATTACCCCGCCTGCAGCGTTGTGTTTGACAAGTA CGGCAAGCTTATTAACTATTACAAGCCTAACATCC CGCAGAACGGAGAGAAGATTAACACAGAAATACTT CAGGAAATTTTCGACAAGGTGCTCATAAGCTATGA GGAGGAGAATGGAGCCTACCCGAAGAATATCGTGA TCCACAGGGACGGCTTTAGCCGAGAGGACCTTGAC TGGTATGAGAACTACTTCGGTAAGAAAAACATAAA	

	GTTTAACATCATCGAAGTCAAAAAGTCAACTCCGT TGAAAATCGCCAGTATAAACGAGGGAAATATCACG AATCCTGAAAAGGGTTCCTACATCCTGCGCGGCAA CAAAGCCTACATGGTGACCACAGATATTAAGGAAA ACCTGGGAAGCCCAAAGCCCCTGAAGATAGAAAAG AGCTACGGCGACATAGACATGCTCACAGCTCTCAG CCAAATATACGCACTCACGCAAATCCATGTGGGGG CGACCAAAAGCCTGCGCCTCCCAATCACCACCGGC TACGCCGACAAGATTTGCAAGGCGATCGAGTTCAT CCCCCAAGGGCGCGTGGACAACCGCCTTTTCTTTC TG	
APO73 (See FIG. 68) N to C terminus: *Italicized: His Tag* Underlined: GST *Italicized/Underlined: 2XSV40NLS* **Bold: VP64** **Underlined/Bold: G4S** ***Italicized/Bold: Ago69***	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCA*ATGCCCAAGAAAAAGCG* *AAAGGTAGAGGACCCCAAAAAGAAACGCAAAGTG*G GCTCCGGAAGCCGTCTCGAA**ATGGACGCATTGGAC GATTTTGATCTGGATATGCTGGGAAGTGACGCCCT CGATGATTTTGACCTTGACATGCTTGGTTCGGATG CCCTTGATGACTTTGACCTCGACATGCTCGGCAGT GACGCCCTTGATGATTTCGACCTGGACATGCTGGG CGGTGGTGGCTCA*ATGGTCGGCGGCTATAAAGTCA *GCAATTTGACAGTGGAAGCGTTCGAAGGTATCGGG *AGTGTCAACCCGATGCTGTTTTACCAATACAAAGT *CACCGGAAAGGGAAAGTACGATAATGTGTATAAGA *TTATCAAAAGCGCACGGTACAAGATGCATTCTAAG *AACCGATTCAAGCCCGTGTTCATCAAGGACGACAA *ACTGTACACCCTCGAGAAGCTCCCGGATATAGAGG *ACCTGGATTTCGCAAACATTAACTTCGTGAAAAGC *GAGGTTCTCAGCATAGAGGATAATATGTCAATTTA	89

	TGGCGAGGTGGTGGAATACTATATCAATCTCAAGC TGAAAAAAGTGAAGGTGTTGGGAAAATACCCCAAG TACAGGATCAATTACAGCAAAGAGATTCTCAGTAA TACGCTGCTGACACGAGAGCTCAAAGACGAGTTTA AGAAATCAAATAAGGGTTTTAACCTGAAACGGAAG TTTAGAATTTCCCCCGTGGTGAATAAGATGGGCAA AGTGATACTCTATTTGTCCTGCAGTGCTGATTTCA GCACCAACAAGAACATTTACGAAATGTTGAAAGAG GGCTTGGAGGTTGAGGGGCTGGCCGTTAAGAGCGA GTGGAGCAATATCAGTGGCAACCTGGTGATCGAGA GCGTACTGGAAACCAAGATATCCGAGCCCACTAGC CTGGGCCAATCCCTGATAGACTACTATAAGAATAA CAACCAGGGCTATAGGGTGAAGGATTTCACCGATG AGGATCTGAATGCCAACATTGTCAACGTGAGAGGA AATAAGAAGATCTATATGTATATTCCGCACGCGTT GAAGCCGATAATCACCCGGGAGTACCTGGCCAAGA ACGATCCAGAGTTTTCTAAGGAGATCGAGCAGCTT ATCAAGATGAATATGAACTACCGATATGAAACCCT CAAGTCATTTGTGAATGACATCGGGGTCATTGAGG AGCTGAACAACCTGAGCTTCAAAAACAAATACTAC GAAGATGTGAAACTGCTGGGTTACTCCAGCGGCAA AATAGACGAACCCGTCCTGATGGGGGCAAAAGGGA TCATAAAGAACAAAATGCAGATTTTTTCCAATGGA TTCTACAAACTCCCCGAAGGCAAGGTACGATTTGG CGTTCTGTACCCAAAAGAATTTGATGGCGTGTCAA GGAAAGCTATCCGCGCCATTTATGACTTCAGTAAG GAGGGCAAATACCACGGCGAAAGCAACAAGTATAT CGCGGAACACCTGATAAACGTGGAGTTCAATCCAA AGGAGTGCATATTTGAGGGATACGAACTGGGCGAT ATCACCGAATACAAGAAGGCGGCTCTGAAACTTAA TAACTACAACAATGTCGACTTCGTAATCGCAATAG TCCCGAACATGTCCGACGAAGAGATAGAGAACAGC TACAATCCGTTCAAGAAAATATGGGCCGAACTGAA TCTGCCCAGCCAGATGATTAGCGTCAAGACGGCCG AAATCTTTGCCAATAGCAGGGATAACACGGCGCTT TACTATCTGCATAACATCGTCCTCGGTATCCTGGG TAAGATAGGAGGGATTCCCTGGGTGGTTAAAGACA TGAAGGGCGACGTGGATTGCTTCGTTGGACTCGAT GTCGGCACCAGGGAGAAGGGCATACATTACCCCGC CTGCAGCGTTGTGTTTGACAAGTACGGCAAGCTTA TTAACTATTACAAGCCTAACATCCCGCAGAACGGA GAGAAGATTAACACAGAAATACTTCAGGAAATTTT CGACAAGGTGCTCATAAGCTATGAGGAGGAGAATG GAGCCTACCCGAAGAATATCGTGATCCACAGGGAC GGCTTTAGCCGAGAGGACCTTGACTGGTATGAGAA CTACTTCGGTAAGAAAAACATAAAGTTTAACATCA TCGAAGTCAAAAAGTCAACTCCGTTGAAAATCGCC AGTATAAACGAGGGAAATATCACGAATCCTGAAAA GGGTTCCTACATCCTGCGCGGCAACAAAGCCTACA	

	TGGTGACCACAGATATTAAGGAAAACCTGGGAAGC CCAAAGCCCCTGAAGATAGAAAAGAGCTACGGCGA CATAGACATGCTCACAGCTCTCAGCCAAATATACG CACTCACGCAAATCCATGTGGGGGCGACCAAAAGC CTGCGCCTCCCAATCACCACCGGCTACGCCGACAA GATTTGCAAGGCGATCGAGTTCATCCCCCAAGGGC GCGTGGACAACCGCCTTTTCTTTCTG	
APO46 (See FIG. 69B) N to C terminus: *Italicized: His Tag* Underlined: GST *Italicized/ Underlined: 2XSV40NLS* **Underlined/Bold: GSGS linker** ***Italicized/Bold: Ago69***	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCGGTCTCATCTCCACAGGG GTACCCGTCTCTA*ATGCCCAAGAAGAAGAGAAAGG* *TCGAGGACCCGAAAAAGAAGCGAAAGGTA***GGTAGT** **GGTTCC*ATGGTCGGCGGCTATAAAGTCAGCAATTT*** ***GACAGTGGAAGCGTTCGAAGGTATCGGGAGTGTCA*** ***ACCCGATGCTGTTTTACCAATACAAAGTCACCGGA*** ***AAGGGAAAGTACGATAATGTGTATAAGATTATCAA*** ***AAGCGCACGGTACAAGATGCATTCTAAGAACCGAT*** ***TCAAGCCCGTGTTCATCAAGGACGACAAACTGTAC*** ***ACCCTCGAGAAGCTCCCGGATATAGAGGACCTGGA*** ***TTTCGCAAACATTAACTTCGTGAAAAGCGAGGTTC*** ***TCAGCATAGAGGATAATATGTCAATTTATGGCGAG*** ***GTGGTGGAATACTATATCAATCTCAAGCTGAAAAA*** ***AGTGAAGGTGTTGGGAAAATACCCCAAGTACAGGA*** ***TCAATTACAGCAAAGAGATTCTCAGTAATACGCTG*** ***CTGACACGAGAGCTCAAAGACGAGTTTAAGAAATC*** ***AAATAAGGGTTTTAACCTGAAACGGAAGTTTAGAA*** ***TTTCCCCCGTGGTGAATAAGATGGGCAAAGTGATA*** ***CTCTATTTGTCCTGCAGTGCTGATTTCAGCACCAA*** ***CAAGAACATTTACGAAATGTTGAAAGAGGGCTTGG*** ***AGGTTGAGGGGCTGGCCGTTAAGAGCGAGTGGAGC***	90

	AATATCAGTGGCAACCTGGTGATCGAGAGCGTACT GGAAACCAAGATATCCGAGCCCACTAGCCTGGGCC AATCCCTGATAGACTACTATAAGAATAACAACCAG GGCTATAGGGTGAAGGATTTCACCGATGAGGATCT GAATGCCAACATTGTCAACGTGAGAGGAAATAAGA AGATCTATATGTATATTCCGCACGCGTTGAAGCCG ATAATCACCCGGGAGTACCTGGCCAAGAACGATCC AGAGTTTTCTAAGGAGATCGAGCAGCTTATCAAGA TGAATATGAACTACCGATATGAAACCCTCAAGTCA TTTGTGAATGACATCGGGGTCATTGAGGAGCTGAA CAACCTGAGCTTCAAAAACAAATACTACGAAGATG TGAAACTGCTGGGTTACTCCAGCGGCAAAATAGAC GAACCCGTCCTGATGGGGGCAAAAGGGATCATAAA GAACAAAATGCAGATTTTTTCCAATGGATTCTACA AACTCCCCGAAGGCAAGGTACGATTTGGCGTTCTG TACCCAAAAGAATTTGATGGCGTGTCAAGGAAAGC TATCCGCGCCATTTATGACTTCAGTAAGGAGGGCA AATACCACGGCGAAAGCAACAAGTATATCGCGGAA CACCTGATAAACGTGGAGTTCAATCCAAAGGAGTG CATATTTGAGGGATACGAACTGGGCGATATCACCG AATACAAGAAGGCGGCTCTGAAACTTAATAACTAC AACAATGTCGACTTCGTAATCGCAATAGTCCCGAA CATGTCCGACGAAGAGATAGAGAACAGCTACAATC CGTTCAAGAAAATATGGGCCGAACTGAATCTGCCC AGCCAGATGATTAGCGTCAAGACGGCCGAAATCTT TGCCAATAGCAGGGATAACACGGCGCTTTACTATC TGCATAACATCGTCCTCGGTATCCTGGGTAAGATA GGAGGGATTCCCTGGGTGGTTAAAGACATGAAGGG CGACGTGGATTGCTTCGTTGGACTCGATGTCGGCA CCAGGGAGAAGGGCATACATTACCCCGCCTGCAGC GTTGTGTTTGACAAGTACGGCAAGCTTATTAACTA TTACAAGCCTAACATCCCGCAGAACGGAGAGAAGA TTAACACAGAAATACTTCAGGAAATTTTCGACAAG GTGCTCATAAGCTATGAGGAGGAGAATGGAGCCTA CCCGAAGAATATCGTGATCCACAGGGACGGCTTTA GCCGAGAGGACCTTGACTGGTATGAGAACTACTTC GGTAAGAAAAACATAAAGTTTAACATCATCGAAGT CAAAAAGTCAACTCCGTTGAAAATCGCCAGTATAA ACGAGGGAAATATCACGAATCCTGAAAAGGGTTCC TACATCCTGCGCGGCAACAAAGCCTACATGGTGAC CACAGATATTAAGGAAAACCTGGGAAGCCCAAAGC CCCTGAAGATAGAAAAGAGCTACGGCGACATAGAC ATGCTCACAGCTCTCAGCCAAATATACGCACTCAC GCAAATCCATGTGGGGGCGACCAAAAGCCTGCGCC TCCCAATCACCACCGGCTACGCCGACAAGATTTGC AAGGCGATCGAGTTCATCCCCCAAGGGCGCGTGGA CAACCGCCTTTTCTTTCTGTAGTGA	
APO25	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA	91

(See FIG. 69B) N to C terminus: *Italicized: His Tag* Underlined: GST ***Bold/Italicized/Underlined: 2XSV40NLS*** **Underlined and Bold: GSGS linker** ***Italicized and Bold: Ago69*** *Italicized and underlined: GGGS linker* **Bold: V5 tag** *Italicized: NPM NLS*	TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCA***ATGCCCAAGAAAAAGCG*** ***AAAGGTAGAGGACCCCAAAAAGAAACGCAAAGTG***G GCTCCGGAAGCCGTCTCGAA***ATGGTCGGCGGCTAT*** ***AAAGTCAGCAATTTGACAGTGGAAGCGTTCGAAGG*** ***TATCGGGAGTGTCAACCCGATGCTGTTTTACCAAT*** ***ACAAAGTCACCGGAAAGGGAAAGTACGATAATGTG*** ***TATAAGATTATCAAAAGCGCACGGTACAAGATGCA*** ***TTCTAAGAACCGATTCAAGCCCGTGTTCATCAAGG*** ***ACGACAAACTGTACACCCTCGAGAAGCTCCCGGAT*** ***ATAGAGGACCTGGATTTCGCAAACATTAACTTCGT*** ***GAAAAGCGAGGTTCTCAGCATAGAGGATAATATGT*** ***CAATTTATGGCGAGGTGGTGGAATACTATATCAAT*** ***CTCAAGCTGAAAAAAGTGAAGGTGTTGGGAAAATA*** ***CCCCAAGTACAGGATCAATTACAGCAAAGAGATTC*** ***TCAGTAATACGCTGCTGACACGAGAGCTCAAAGAC*** ***GAGTTTAAGAAATCAAATAAGGGTTTTAACCTGAA*** ***ACGGAAGTTTAGAATTTCCCCCGTGGTGAATAAGA*** ***TGGGCAAAGTGATACTCTATTTGTCCTGCAGTGCT*** ***GATTTCAGCACCAACAAGAACATTTACGAAATGTT*** ***GAAAGAGGGCTTGGAGGTTGAGGGGCTGGCCGTTA*** ***AGAGCGAGTGGAGCAATATCAGTGGCAACCTGGTG*** ***ATCGAGAGCGTACTGGAAACCAAGATATCCGAGCC*** ***CACTAGCCTGGGCCAATCCCTGATAGACTACTATA*** ***AGAATAACAACCAGGGCTATAGGGTGAAGGATTTC*** ***ACCGATGAGGATCTGAATGCCAACATTGTCAACGT*** ***GAGAGGAAATAAGAAGATCTATATGTATATTCCGC*** ***ACGCGTTGAAGCCGATAATCACCCGGGAGTACCTG*** ***GCCAAGAACGATCCAGAGTTTTCTAAGGAGATCGA*** ***GCAGCTTATCAAGATGAATATGAACTACCGATATG*** ***AAACCCTCAAGTCATTTGTGAATGACATCGGGGTC***	

	ATTGAGGAGCTGAACAACCTGAGCTTCAAAAACAA ***ATACTACGAAGATGTGAAACTGCTGGGTTACTCCA*** ***GCGGCAAAATAGACGAACCCGTCCTGATGGGGGCA*** ***AAAGGGATCATAAAGAACAAAATGCAGATTTTTTC*** ***CAATGGATTCTACAAACTCCCCGAAGGCAAGGTAC*** ***GATTTGGCGTTCTGTACCCAAAAGAATTTGATGGC*** ***GTGTCAAGGAAAGCTATCCGCGCCATTTATGACTT*** ***CAGTAAGGAGGGCAAATACCACGGCGAAAGCAACA*** ***AGTATATCGCGGAACACCTGATAAACGTGGAGTTC*** ***AATCCAAAGGAGTGCATATTTGAGGGATACGAACT*** ***GGGCGATATCACCGAATACAAGAAGGCGGCTCTGA*** ***AACTTAATAACTACAACAATGTCGACTTCGTAATC*** ***GCAATAGTCCCGAACATGTCCGACGAAGAGATAGA*** ***GAACAGCTACAATCCGTTCAAGAAAATATGGGCCG*** ***AACTGAATCTGCCCAGCCAGATGATTAGCGTCAAG*** ***ACGGCCGAAATCTTTGCCAATAGCAGGGATAACAC*** ***GGCGCTTTACTATCTGCATAACATCGTCCTCGGTA*** ***TCCTGGGTAAGATAGGAGGGATTCCCTGGGTGGTT*** ***AAAGACATGAAGGGCGACGTGGATTGCTTCGTTGG*** ***ACTCGATGTCGGCACCAGGGAGAAGGGCATACATT*** ***ACCCCGCCTGCAGCGTTGTGTTTGACAAGTACGGC*** ***AAGCTTATTAACTATTACAAGCCTAACATCCCGCA*** ***GAACGGAGAGAAGATTAACACAGAAATACTTCAGG*** ***AAATTTTCGACAAGGTGCTCATAAGCTATGAGGAG*** ***GAGAATGGAGCCTACCCGAAGAATATCGTGATCCA*** ***CAGGGACGGCTTTAGCCGAGAGGACCTTGACTGGT*** ***ATGAGAACTACTTCGGTAAGAAAAACATAAAGTTT*** ***AACATCATCGAAGTCAAAAAGTCAACTCCGTTGAA*** ***AATCGCCAGTATAAACGAGGGAAATATCACGAATC*** ***CTGAAAAGGGTTCCTACATCCTGCGCGGCAACAAA*** ***GCCTACATGGTGACCACAGATATTAAGGAAAACCT*** ***GGGAAGCCCAAAGCCCCTGAAGATAGAAAAGAGCT*** ***ACGGCGACATAGACATGCTCACAGCTCTCAGCCAA*** ***ATATACGCACTCACGCAAATCCATGTGGGGGCGAC*** ***CAAAAGCCTGCGCCTCCCAATCACCACCGGCTACG*** ***CCGACAAGATTTGCAAGGCGATCGAGTTCATCCCC*** ***CAAGGGCGCGTGGACAACCGCCTTTTCTTTCTG***AC TAGTGGGGGAGGTGGATCT*GGGAAGCCCATCCCAA* *ACCCGCTGTTGGGCTTGGATTCCACGAAGCGACCC* *GCAGCGACTAAGAAAGCCGGCCAGGCCAAAAAGAA* *GAAA*	
APO71 (See FIG. 69B) N to C terminus: *Italicized: His Tag*	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA	92

Underlined: GST **Bold: 2XSV40NLS** **Underlined/Bold: GSGS linker** ***Italicized/Bold: Ago69***	ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGG**ATGAGCGAGAAT** **CTTTATTTTCAGGGCGCAATGCCCAAGAAAAAGCG** **AAAGGTAGAGGACCCCAAAAAGAAACGCAAAGTGG** **GCTCCGGAAGC**CGTCTCGAA***ATGGTCGGCGGCTAT*** ***AAAGTCAGCAATTTGACAGTGGAAGCGTTCGAAGG*** ***TATCGGGAGTGTCAACCCGATGCTGTTTTACCAAT*** ***ACAAAGTCACCGGAAAGGGAAAGTACGATAATGTG*** ***TATAAGATTATCAAAAGCGCACGGTACAAGATGCA*** ***TTCTAAGAACCGATTCAAGCCCGTGTTCATCAAGG*** ***ACGACAAACTGTACACCCTCGAGAAGCTCCCGGAT*** ***ATAGAGGACCTGGATTTCGCAAACATTAACTTCGT*** ***GAAAAGCGAGGTTCTCAGCATAGAGGATAATATGT*** ***CAATTTATGGCGAGGTGGTGGAATACTATATCAAT*** ***CTCAAGCTGAAAAAAGTGAAGGTGTTGGGAAAATA*** ***CCCCAAGTACAGGATCAATTACAGCAAAGAGATTC*** ***TCAGTAATACGCTGCTGACACGAGAGCTCAAAGAC*** ***GAGTTTAAGAAATCAAATAAGGGTTTTAACCTGAA*** ***ACGGAAGTTTAGAATTTCCCCCGTGGTGAATAAGA*** ***TGGGCAAAGTGATACTCTATTTGTCCTGCAGTGCT*** ***GATTTCAGCACCAACAAGAACATTTACGAAATGTT*** ***GAAAGAGGGCTTGGAGGTTGAGGGGCTGGCCGTTA*** ***AGAGCGAGTGGAGCAATATCAGTGGCAACCTGGTG*** ***ATCGAGAGCGTACTGGAAACCAAGATATCCGAGCC*** ***CACTAGCCTGGGCCAATCCCTGATAGACTACTATA*** ***AGAATAACAACCAGGGCTATAGGGTGAAGGATTTC*** ***ACCGATGAGGATCTGAATGCCAACATTGTCAACGT*** ***GAGAGGAAATAAGAAGATCTATATGTATATTCCGC*** ***ACGCGTTGAAGCCGATAATCACCCGGGAGTACCTG*** ***GCCAAGAACGATCCAGAGTTTTCTAAGGAGATCGA*** ***GCAGCTTATCAAGATGAATATGAACTACCGATATG*** ***AAACCCTCAAGTCATTTGTGAATGACATCGGGGTC*** ***ATTGAGGAGCTGAACAACCTGAGCTTCAAAAACAA*** ***ATACTACGAAGATGTGAAACTGCTGGGTTACTCCA*** ***GCGGCAAAATAGACGAACCCGTCCTGATGGGGGCA*** ***AAAGGGATCATAAAGAACAAAATGCAGATTTTTTC*** ***CAATGGATTCTACAAACTCCCCGAAGGCAAGGTAC*** ***GATTTGGCGTTCTGTACCCAAAAGAATTTGATGGC***	

	GTGTCAAGGAAAGCTATCCGCGCCATTTATGACTT CAGTAAGGAGGGCAAATACCACGGCGAAAGCAACA AGTATATCGCGGAACACCTGATAAACGTGGAGTTC AATCCAAAGGAGTGCATATTTGAGGGATACGAACT GGGCGATATCACCGAATACAAGAAGGCGGCTCTGA AACTTAATAACTACAACAATGTCGACTTCGTAATC GCAATAGTCCCGAACATGTCCGACGAAGAGATAGA GAACAGCTACAATCCGTTCAAGAAAATATGGGCCG AACTGAATCTGCCCAGCCAGATGATTAGCGTCAAG ACGGCCGAAATCTTTGCCAATAGCAGGGATAACAC GGCGCTTTACTATCTGCATAACATCGTCCTCGGTA TCCTGGGTAAGATAGGAGGGATTCCCTGGGTGGTT AAAGACATGAAGGGCGACGTGGATTGCTTCGTTGG ACTCGATGTCGGCACCAGGGAGAAGGGCATACATT ACCCCGCCTGCAGCGTTGTGTTTGACAAGTACGGC AAGCTTATTAACTATTACAAGCCTAACATCCCGCA GAACGGAGAGAAGATTAACACAGAAATACTTCAGG AAATTTTCGACAAGGTGCTCATAAGCTATGAGGAG GAGAATGGAGCCTACCCGAAGAATATCGTGATCCA CAGGGACGGCTTTAGCCGAGAGGACCTTGACTGGT ATGAGAACTACTTCGGTAAGAAAAACATAAAGTTT AACATCATCGAAGTCAAAAAGTCAACTCCGTTGAA AATCGCCAGTATAAACGAGGGAAATATCACGAATC CTGAAAAGGGTTCCTACATCCTGCGCGGCAACAAA GCCTACATGGTGACCACAGATATTAAGGAAAACCT GGGAAGCCCAAAGCCCCTGAAGATAGAAAAGAGCT ACGGCGACATAGACATGCTCACAGCTCTCAGCCAA ATATACGCACTCACGCAAATCCATGTGGGGGCGAC CAAAAGCCTGCGCCTCCCAATCACCACCGGCTACG CCGACAAGATTTGCAAGGCGATCGAGTTCATCCCC CAAGGGCGCGTGGACAACCGCCTTTTCTTTCTGTA G	
SSB-AGO#69v3 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG	93

CTATCCCACAAATTGATAAGTACTTGAAATCCAGC
AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC
CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG
GATCTGGTGGTGGTGGCGGATGG**ATGAGCGAGAAT**
CTTTATTTTCAGGGCGCAATGGAGGAGAAGGTCGG
TAACCTTAAGCCCAATATGGAATCTGTCAACGTTA
CTGTTAGAGTCCTGGAAGCCAGCGAGGCGCGCCAA
ATACAGACGAAGAACGGGGTAAGGACCATAAGCGA
GGCAATCGTCGGCGACGAGACTGGCAGAGTTAAAT
TGACCCTTTGGGGAAAACACGCTGGTTCTATAAAG
GAAGGTCAAGTCGTAAAGATCGAAAATGCTTGGAC
GACCGCATTCAAGGGCCAGGTTCAACTCAACGCCG
GATCTAAAACTAAGATAGCTGAGGCGTCAGAAGAC
GGCTTCCCAGAATCAAGCCAGATACCTGAGAACAC
TCCAACGGCTCCTCAACAAATGAGAGGAGGTGGAC
GAGGATTTCGCGGGGGGGGACGAAGATATGGCCGA
CGCGGAGGACGACGCCAGGAAAATGAAGAGGGTGA
AGAGGAA*AGCGGCTCTGAGACTCCCGGCACATCCG*
*AAAGCGCAACCCCTGAGTCT****GTCGGCGGCTATAAA***
GTCAGCAATTTGACAGTGGAAGCGTTCGAAGGTAT
CGGGAGTGTCAACCCGATGCTGTTTTACCAATACA
AAGTCACCGGAAAGGGAAAGTACGATAATGTGTAT
AAGATTATCAAAAGCGCACGGTACAAGATGCATTC
TAAGAACCGATTCAAGCCCGTGTTCATCAAGGACG
ACAAACTGTACACCCTCGAGAAGCTCCCGGATATA
GAGGACCTGGATTTCGCAAACATTAACTTCGTGAA
AAGCGAGGTTCTCAGCATAGAGGATAATATGTCAA
TTTATGGCGAGGTGGTGGAATACTATATCAATCTC
AAGCTGAAAAAAGTGAAGGTGTTGGGAAAATACCC
CAAGTACAGGATCAATTACAGCAAAGAGATTCTCA
GTAATACGCTGCTGACACGAGAGCTCAAAGACGAG
TTTAAGAAATCAAATAAGGGTTTTAACCTGAAACG
GAAGTTTAGAATTTCCCCCGTGGTGAATAAGATGG
GCAAAGTGATACTCTATTTGTCCTGCAGTGCTGAT
TTCAGCACCAACAAGAACATTTACGAAATGTTGAA
AGAGGGCTTGGAGGTTGAGGGGCTGGCCGTTAAGA
GCGAGTGGAGCAATATCAGTGGCAACCTGGTGATC
GAGAGCGTACTGGAAACCAAGATATCCGAGCCCAC
TAGCCTGGGCCAATCCCTGATAGACTACTATAAGA
ATAACAACCAGGGCTATAGGGTGAAGGATTTCACC
GATGAGGATCTGAATGCCAACATTGTCAACGTGAG
AGGAAATAAGAAGATCTATATGTATATTCCGCACG
CGTTGAAGCCGATAATCACCCGGGAGTACCTGGCC
AAGAACGATCCAGAGTTTTCTAAGGAGATCGAGCA
GCTTATCAAGATGAATATGAACTACCGATATGAAA
CCCTCAAGTCATTTGTGAATGACATCGGGGTCATT
GAGGAGCTGAACAACCTGAGCTTCAAAAACAAATA
CTACGAAGATGTGAAACTGCTGGGTTACTCCAGCG
GCAAAATAGACGAACCCGTCCTGATGGGGGCAAAA

	GGGATCATAAAGAACAAAATGCAGATTTTTTCCAA TGGATTCTACAAACTCCCCGAAGGCAAGGTACGAT TTGGCGTTCTGTACCCAAAAGAATTTGATGGCGTG TCAAGGAAAGCTATCCGCGCCATTTATGACTTCAG TAAGGAGGGCAAATACCACGGCGAAAGCAACAAGT ATATCGCGGAACACCTGATAAACGTGGAGTTCAAT CCAAAGGAGTGCATATTTGAGGGATACGAACTGGG CGATATCACCGAATACAAGAAGGCGGCTCTGAAAC TTAATAACTACAACAATGTCGACTTCGTAATCGCA ATAGTCCCGAACATGTCCGACGAAGAGATAGAGAA CAGCTACAATCCGTTCAAGAAAATATGGGCCGAAC TGAATCTGCCCAGCCAGATGATTAGCGTCAAGACG GCCGAAATCTTTGCCAATAGCAGGGATAACACGGC GCTTTACTATCTGCATAACATCGTCCTCGGTATCC TGGGTAAGATAGGAGGGATTCCCTGGGTGGTTAAA GACATGAAGGGCGACGTGGATTGCTTCGTTGGACT CGATGTCGGCACCAGGGAGAAGGGCATACATTACC CCGCCTGCAGCGTTGTGTTTGACAAGTACGGCAAG CTTATTAACTATTACAAGCCTAACATCCCGCAGAA CGGAGAGAAGATTAACACAGAAATACTTCAGGAAA TTTTCGACAAGGTGCTCATAAGCTATGAGGAGGAG AATGGAGCCTACCCGAAGAATATCGTGATCCACAG GGACGGCTTTAGCCGAGAGGACCTTGACTGGTATG AGAACTACTTCGGTAAGAAAAACATAAAGTTTAAC ATCATCGAAGTCAAAAAGTCAACTCCGTTGAAAAT CGCCAGTATAAACGAGGGAAATATCACGAATCCTG AAAAGGGTTCCTACATCCTGCGCGGCAACAAAGCC TACATGGTGACCACAGATATTAAGGAAAACCTGGG AAGCCCAAAGCCCCTGAAGATAGAAAAGAGCTACG GCGACATAGACATGCTCACAGCTCTCAGCCAAATA TACGCACTCACGCAAATCCATGTGGGGGCGACCAA AAGCCTGCGCCTCCCAATCACCACCGGCTACGCCG ACAAGATTTGCAAGGCGATCGAGTTCATCCCCCAA GGGCGCGTGGACAACCGCCTTTTCTTTCTGTAGTG A**	
2SSB-AGO#69v1 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** **Underlined/Bold: G4S**	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC	94

Italicized/Underlined: XTEN ***Italicized/Bold: Ago69***	ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCA**ATGGAGGAGAAGGTCGG** **TAACCTTAAGCCCAATATGGAATCTGTCAACGTTA** **CTGTTAGAGTCCTGGAAGCCAGCGAGGCGCGCCAA** **ATACAGACGAAGAACGGGGTAAGGACCATAAGCGA** **GGCAATCGTCGGCGACGAGACTGGCAGAGTTAAAT** **TGACCCTTTGGGGAAAACACGCTGGTTCTATAAAG** **GAAGGTCAAGTCGTAAAGATCGAAAATGCTTGGAC** **GACCGCATTCAAGGGCCAGGTTCAACTCAACGCCG** **GATCTAAAACTAAGATAGCTGAGGCGTCAGAAGAC** **GGCTTCCCAGAATCAAGCCAGATACCTGAGAACAC** **TCCAACGGCTCCTCAACAAATGAGAGGAGGTGGAC** **GAGGATTTCGCGGGGGGGGACGAAGATATGGCCGA** **CGCGGAGGACGACGCCAGGAAAATGAAGAGGGTGA** **AGAGGAAGGCGGTGGTGGCTCAGAAGAAAAAGTGG** **GTAACCTTAAACCCAACATGGAGAGCGTTAACGTC** **ACGGTCAGAGTACTCGAGGCGAGCGAGGCGCGGCA** **GATACAAACAAAAAATGGTGTGCGCACCATTTCCG** **AAGCTATAGTCGGTGACGAAACGGGCCGCGTTAAA** **TTGACGCTCTGGGGAAAACATGCAGGTTCTATTAA** **AGAGGGTCAGGTCGTGAAAATAGAGAACGCCTGGA** **CTACGGCGTTCAAGGGTCAGGTCCAACTGAATGCA** **GGGTCTAAAACTAAAATTGCGGAGGCTAGTGAAGA** **TGGTTTTCCCGAATCAAGCCAGATTCCAGAAAATA** **CACCTACGGCACCGCAACAGATGCGAGGAGGCGGG** **CGAGGATTTCGAGGTGGAGGTCGACGCTACGGTAG** **GAGGGGTGGGCGGCGCCAAGAGAACGAAGAAGGAG** **AGGAAGAA***AGCGGCTCTGAGACTCCCGGCACATCC* *GAAAGCGCAACCCCTGAGTCT****GTCGGCGGCTATAA*** ***AGTCAGCAATTTGACAGTGGAAGCGTTCGAAGGTA*** ***TCGGGAGTGTCAACCCGATGCTGTTTTACCAATAC*** ***AAAGTCACCGGAAAGGGAAAGTACGATAATGTGTA*** ***TAAGATTATCAAAAGCGCACGGTACAAGATGCATT*** ***CTAAGAACCGATTCAAGCCCGTGTTCATCAAGGAC*** ***GACAAACTGTACACCCTCGAGAAGCTCCCGGATAT*** ***AGAGGACCTGGATTTCGCAAACATTAACTTCGTGA*** ***AAAGCGAGGTTCTCAGCATAGAGGATAATATGTCA*** ***ATTTATGGCGAGGTGGTGGAATACTATATCAATCT*** ***CAAGCTGAAAAAAGTGAAGGTGTTGGGAAAATACC*** ***CCAAGTACAGGATCAATTACAGCAAAGAGATTCTC*** ***AGTAATACGCTGCTGACACGAGAGCTCAAAGACGA*** ***GTTTAAGAAATCAAATAAGGGTTTTAACCTGAAAC*** ***GGAAGTTTAGAATTTCCCCCGTGGTGAATAAGATG***	

	GGCAAAGTGATACTCTATTTGTCCTGCAGTGCTGA TTTCAGCACCAACAAGAACATTTACGAAATGTTGA AAGAGGGCTTGGAGGTTGAGGGGCTGGCCGTTAAG AGCGAGTGGAGCAATATCAGTGGCAACCTGGTGAT CGAGAGCGTACTGGAAACCAAGATATCCGAGCCCA CTAGCCTGGGCCAATCCCTGATAGACTACTATAAG AATAACAACCAGGGCTATAGGGTGAAGGATTTCAC CGATGAGGATCTGAATGCCAACATTGTCAACGTGA GAGGAAATAAGAAGATCTATATGTATATTCCGCAC GCGTTGAAGCCGATAATCACCCGGGAGTACCTGGC CAAGAACGATCCAGAGTTTTCTAAGGAGATCGAGC AGCTTATCAAGATGAATATGAACTACCGATATGAA ACCCTCAAGTCATTTGTGAATGACATCGGGGTCAT TGAGGAGCTGAACAACCTGAGCTTCAAAAACAAAT ACTACGAAGATGTGAAACTGCTGGGTTACTCCAGC GGCAAAATAGACGAACCCGTCCTGATGGGGGCAAA AGGGATCATAAAGAACAAAATGCAGATTTTTTCCA ATGGATTCTACAAACTCCCCGAAGGCAAGGTACGA TTTGGCGTTCTGTACCCAAAAGAATTTGATGGCGT GTCAAGGAAAGCTATCCGCGCCATTTATGACTTCA GTAAGGAGGGCAAATACCACGGCGAAAGCAACAAG TATATCGCGGAACACCTGATAAACGTGGAGTTCAA TCCAAAGGAGTGCATATTTGAGGGATACGAACTGG GCGATATCACCGAATACAAGAAGGCGGCTCTGAAA CTTAATAACTACAACAATGTCGACTTCGTAATCGC AATAGTCCCGAACATGTCCGACGAAGAGATAGAGA ACAGCTACAATCCGTTCAAGAAAATATGGGCCGAA CTGAATCTGCCCAGCCAGATGATTAGCGTCAAGAC GGCCGAAATCTTTGCCAATAGCAGGGATAACACGG CGCTTTACTATCTGCATAACATCGTCCTCGGTATC CTGGGTAAGATAGGAGGGATTCCCTGGGTGGTTAA AGACATGAAGGGCGACGTGGATTGCTTCGTTGGAC TCGATGTCGGCACCAGGGAGAAGGGCATACATTAC CCCGCCTGCAGCGTTGTGTTTGACAAGTACGGCAA GCTTATTAACTATTACAAGCCTAACATCCCGCAGA ACGGAGAGAAGATTAACACAGAAATACTTCAGGAA ATTTTCGACAAGGTGCTCATAAGCTATGAGGAGGA GAATGGAGCCTACCCGAAGAATATCGTGATCCACA GGGACGGCTTTAGCCGAGAGGACCTTGACTGGTAT GAGAACTACTTCGGTAAGAAAAACATAAAGTTTAA CATCATCGAAGTCAAAAAGTCAACTCCGTTGAAAA TCGCCAGTATAAACGAGGGAAATATCACGAATCCT GAAAAGGGTTCCTACATCCTGCGCGGCAACAAAGC CTACATGGTGACCACAGATATTAAGGAAAACCTGG GAAGCCCAAAGCCCCTGAAGATAGAAAAGAGCTAC GGCGACATAGACATGCTCACAGCTCTCAGCCAAAT ATACGCACTCACGCAAATCCATGTGGGGGCGACCA AAAGCCTGCGCCTCCCAATCACCACCGGCTACGCC GACAAGATTTGCAAGGCGATCGAGTTCATCCCCCA	

	AGGGCGCGTGGACAACCGCCTTTTCTTTCTGTAGTGA	
2SSB-AGO#69v2 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** **Underlined/Bold: SGSG4S** *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCA**ATGGAGGAGAAGGTCGG** **TAACCTTAAGCCCAATATGGAATCTGTCAACGTTA** **CTGTTAGAGTCCTGGAAGCCAGCGAGGCGCGCCAA** **ATACAGACGAAGAACGGGGTAAGGACCATAAGCGA** **GGCAATCGTCGGCGACGAGACTGGCAGAGTTAAAT** **TGACCCTTTGGGGAAAACACGCTGGTTCTATAAAG** **GAAGGTCAAGTCGTAAAGATCGAAAATGCTTGGAC** **GACCGCATTCAAGGGCCAGGTTCAACTCAACGCCG** **GATCTAAAACTAAGATAGCTGAGGCGTCAGAAGAC** **GGCTTCCCAGAATCAAGCCAGATACCTGAGAACAC** **TCCAACGGCTCCTCAACAAATGAGAGGAGGTGGAC** **GAGGATTTCGCGGGGGGGGACGAAGATATGGCCGA** **CGCGGAGGACGACGCCAGGAAAATGAAGAGGGTGA** **AGAGGAATCTGGTTCCGGTGGCGGTGGTAGCGAAG** **AAAAAGTGGGTAACCTTAAACCCAACATGGAGAGC** **GTTAACGTCACGGTCAGAGTACTCGAGGCGAGCGA** **GGCGCGGCAGATACAAACAAAAAATGGTGTGCGCA** **CCATTTCCGAAGCTATAGTCGGTGACGAAACGGGC** **CGCGTTAAATTGACGCTCTGGGGAAAACATGCAGG** **TTCTATTAAAGAGGGTCAGGTCGTGAAAATAGAGA** **ACGCCTGGACTACGGCGTTCAAGGGTCAGGTCCAA** **CTGAATGCAGGGTCTAAAACTAAAATTGCGGAGGC** **TAGTGAAGATGGTTTTCCCGAATCAAGCCAGATTC** **CAGAAAATACACCTACGGCACCGCAACAGATGCGA** **GGAGGCGGGCGAGGATTTCGAGGTGGAGGTCGACG** **CTACGGTAGGAGGGGTGGGCGGCGCCAAGAGAACG**	95

	AAGAAGGAGAGGAAGAA*AGCGGCTCTGAGACTCCC* *GGCACATCCGAAAGCGCAACCCCTGAGTCT***GTCGG** ***CGGCTATAAAGTCAGCAATTTGACAGTGGAAGCGT*** ***TCGAAGGTATCGGGAGTGTCAACCCGATGCTGTTT*** ***TACCAATACAAAGTCACCGGAAAGGGAAAGTACGA*** ***TAATGTGTATAAGATTATCAAAAGCGCACGGTACA*** ***AGATGCATTCTAAGAACCGATTCAAGCCCGTGTTC*** ***ATCAAGGACGACAAACTGTACACCCTCGAGAAGCT*** ***CCCGGATATAGAGGACCTGGATTTCGCAAACATTA*** ***ACTTCGTGAAAAGCGAGGTTCTCAGCATAGAGGAT*** ***AATATGTCAATTTATGGCGAGGTGGTGGAATACTA*** ***TATCAATCTCAAGCTGAAAAAAGTGAAGGTGTTGG*** ***GAAAATACCCCAAGTACAGGATCAATTACAGCAAA*** ***GAGATTCTCAGTAATACGCTGCTGACACGAGAGCT*** ***CAAAGACGAGTTTAAGAAATCAAATAAGGGTTTTA*** ***ACCTGAAACGGAAGTTTAGAATTTCCCCCGTGGTG*** ***AATAAGATGGGCAAAGTGATACTCTATTTGTCCTG*** ***CAGTGCTGATTTCAGCACCAACAAGAACATTTACG*** ***AAATGTTGAAAGAGGGCTTGGAGGTTGAGGGGCTG*** ***GCCGTTAAGAGCGAGTGGAGCAATATCAGTGGCAA*** ***CCTGGTGATCGAGAGCGTACTGGAAACCAAGATAT*** ***CCGAGCCCACTAGCCTGGGCCAATCCCTGATAGAC*** ***TACTATAAGAATAACAACCAGGGCTATAGGGTGAA*** ***GGATTTCACCGATGAGGATCTGAATGCCAACATTG*** ***TCAACGTGAGAGGAAATAAGAAGATCTATATGTAT*** ***ATTCCGCACGCGTTGAAGCCGATAATCACCCGGGA*** ***GTACCTGGCCAAGAACGATCCAGAGTTTTCTAAGG*** ***AGATCGAGCAGCTTATCAAGATGAATATGAACTAC*** ***CGATATGAAACCCTCAAGTCATTTGTGAATGACAT*** ***CGGGGTCATTGAGGAGCTGAACAACCTGAGCTTCA*** ***AAAACAAATACTACGAAGATGTGAAACTGCTGGGT*** ***TACTCCAGCGGCAAAATAGACGAACCCGTCCTGAT*** ***GGGGGCAAAAGGGATCATAAAGAACAAAATGCAGA*** ***TTTTTTCCAATGGATTCTACAAACTCCCCGAAGGC*** ***AAGGTACGATTTGGCGTTCTGTACCCAAAAGAATT*** ***TGATGGCGTGTCAAGGAAAGCTATCCGCGCCATTT*** ***ATGACTTCAGTAAGGAGGGCAAATACCACGGCGAA*** ***AGCAACAAGTATATCGCGGAACACCTGATAAACGT*** ***GGAGTTCAATCCAAAGGAGTGCATATTTGAGGGAT*** ***ACGAACTGGGCGATATCACCGAATACAAGAAGGCG*** ***GCTCTGAAACTTAATAACTACAACAATGTCGACTT*** ***CGTAATCGCAATAGTCCCGAACATGTCCGACGAAG*** ***AGATAGAGAACAGCTACAATCCGTTCAAGAAAATA*** ***TGGGCCGAACTGAATCTGCCCAGCCAGATGATTAG*** ***CGTCAAGACGGCCGAAATCTTTGCCAATAGCAGGG*** ***ATAACACGGCGCTTTACTATCTGCATAACATCGTC*** ***CTCGGTATCCTGGGTAAGATAGGAGGGATTCCCTG*** ***GGTGGTTAAAGACATGAAGGGCGACGTGGATTGCT*** ***TCGTTGGACTCGATGTCGGCACCAGGGAGAAGGGC***	

	ATACATTACCCCGCCTGCAGCGTTGTGTTTGACAA GTACGGCAAGCTTATTAACTATTACAAGCCTAACA TCCCGCAGAACGGAGAGAAGATTAACACAGAAATA CTTCAGGAAATTTTCGACAAGGTGCTCATAAGCTA TGAGGAGGAGAATGGAGCCTACCCGAAGAATATCG TGATCCACAGGGACGGCTTTAGCCGAGAGGACCTT GACTGGTATGAGAACTACTTCGGTAAGAAAAACAT AAAGTTTAACATCATCGAAGTCAAAAAGTCAACTC CGTTGAAAATCGCCAGTATAAACGAGGGAAATATC ACGAATCCTGAAAAGGGTTCCTACATCCTGCGCGG CAACAAAGCCTACATGGTGACCACAGATATTAAGG AAAACCTGGGAAGCCCAAAGCCCCTGAAGATAGAA AAGAGCTACGGCGACATAGACATGCTCACAGCTCT CAGCCAAATATACGCACTCACGCAAATCCATGTGG GGGCGACCAAAAGCCTGCGCCTCCCAATCACCACC GGCTACGCCGACAAGATTTGCAAGGCGATCGAGTT CATCCCCCAAGGGCGCGTGGACAACCGCCTTTTCT TTCTGTAGTGA**	
2SSB-AGO#69v3 (See FIG. 75) N to C terminus: *Italicized: His Tag* Underlined: GST **Bold: SsoSSB** *Italicized/Underlined: XTEN* ***Italicized/Bold: Ago69***	ATGAAA*CATCACCATCACCATCAC*AACACTAGTAG CAATTCCATGTCCCCTATACTAGGTTATTGGAAAA TTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTG GAATATCTTGAAGAAAAATATGAAGAGCATTTGTA TGAGCGCGATGAAGGTGATAAATGGCGAAACAAAA AGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCT TATTATATTGATGGTGATGTTAAATTAACACAGTC TATGGCCATCATACGTTATATAGCTGACAAGCACA ACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAG ATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAG ATACGGTGTTTCGAGAATTGCATATAGTAAAGACT TTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTA CCTGAAATGCTGAAAATGTTCGAAGATCGTTTATG TCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTT GTTTTATACATGGACCCAATGTGCCTGGATGCGTT CCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAG CTATCCCACAAATTGATAAGTACTTGAAATCCAGC AAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGC CACGTTTGGTGGTGGCGACCATCCTCCAACTAGTG GATCTGGTGGTGGTGGCGGATGGATGAGCGAGAAT CTTTATTTTCAGGGCGCA**ATGGAGGAGAAGGTCGG TAACCTTAAGCCCAATATGGAATCTGTCAACGTTA CTGTTAGAGTCCTGGAAGCCAGCGAGGCGCGCCAA ATACAGACGAAGAACGGGGTAAGGACCATAAGCGA GGCAATCGTCGGCGACGAGACTGGCAGAGTTAAAT TGACCCTTTGGGGAAAACACGCTGGTTCTATAAAG GAAGGTCAAGTCGTAAAGATCGAAAATGCTTGGAC GACCGCATTCAAGGGCCAGGTTCAACTCAACGCCG GATCTAAAACTAAGATAGCTGAGGCGTCAGAAGAC GGCTTCCCAGAATCAAGCCAGATACCTGAGAACAC**	96

	TCCAACGGCTCCTCAACAAATGAGAGGAGGTGGAC GAGGATTTCGCGGGGGGGGACGAAGATATGGCCGA CGCGGAGGACGACGCCAGGAAAATGAAGAGGGTGA AGAGGAAAGCGGCTCTGAGACTCCCGGCACATCCG AAAGCGCGAAGAAAAAGTGGGTAACCTTAAACCCA ACATGGAGAGCGTTAACGTCACGGTCAGAGTACTC GAGGCGAGCGAGGCGCGGCAGATACAAACAAAAAA TGGTGTGCGCACCATTTCCGAAGCTATAGTCGGTG ACGAAACGGGCCGCGTTAAATTGACGCTCTGGGGA AAACATGCAGGTTCTATTAAAGAGGGTCAGGTCGT GAAAATAGAGAACGCCTGGACTACGGCGTTCAAGG GTCAGGTCCAACTGAATGCAGGGTCTAAAACTAAA ATTGCGGAGGCTAGTGAAGATGGTTTTCCCGAATC AAGCCAGATTCCAGAAAATACACCTACGGCACCGC AACAGATGCGAGGAGGCGGGCGAGGATTTCGAGGT GGAGGTCGACGCTACGGTAGGAGGGGTGGGCGGCG CCAAGAGAACGAAGAAGGAGAGGAAGAAAGCGGCT CTGAGACTCCCGGCACATCCGAAAGCGCAACCCCT GAGTCTGTCGGCGGCTATAAAGTCAGCAATTTGAC AGTGGAAGCGTTCGAAGGTATCGGGAGTGTCAACC CGATGCTGTTTTACCAATACAAAGTCACCGGAAAG GGAAAGTACGATAATGTGTATAAGATTATCAAAAG CGCACGGTACAAGATGCATTCTAAGAACCGATTCA AGCCCGTGTTCATCAAGGACGACAAACTGTACACC CTCGAGAAGCTCCCGGATATAGAGGACCTGGATTT CGCAAACATTAACTTCGTGAAAAGCGAGGTTCTCA GCATAGAGGATAATATGTCAATTTATGGCGAGGTG GTGGAATACTATATCAATCTCAAGCTGAAAAAAGT GAAGGTGTTGGGAAAATACCCCAAGTACAGGATCA ATTACAGCAAAGAGATTCTCAGTAATACGCTGCTG ACACGAGAGCTCAAAGACGAGTTTAAGAAATCAAA TAAGGGTTTTAACCTGAAACGGAAGTTTAGAATTT CCCCCGTGGTGAATAAGATGGGCAAAGTGATACTC TATTTGTCCTGCAGTGCTGATTTCAGCACCAACAA GAACATTTACGAAATGTTGAAAGAGGGCTTGGAGG TTGAGGGGCTGGCCGTTAAGAGCGAGTGGAGCAAT ATCAGTGGCAACCTGGTGATCGAGAGCGTACTGGA AACCAAGATATCCGAGCCCACTAGCCTGGGCCAAT CCCTGATAGACTACTATAAGAATAACAACCAGGGC TATAGGGTGAAGGATTTCACCGATGAGGATCTGAA TGCCAACATTGTCAACGTGAGAGGAAATAAGAAGA TCTATATGTATATTCCGCACGCGTTGAAGCCGATA ATCACCCGGGAGTACCTGGCCAAGAACGATCCAGA GTTTTCTAAGGAGATCGAGCAGCTTATCAAGATGA ATATGAACTACCGATATGAAACCCTCAAGTCATTT GTGAATGACATCGGGGTCATTGAGGAGCTGAACAA CCTGAGCTTCAAAAACAAATACTACGAAGATGTGA AACTGCTGGGTTACTCCAGCGGCAAAATAGACGAA CCCGTCCTGATGGGGGCAAAAGGGATCATAAAGAA	

	CAAAATGCAGATTTTTTCCAATGGATTCTACAAAC TCCCCGAAGGCAAGGTACGATTTGGCGTTCTGTAC CCAAAAGAATTTGATGGCGTGTCAAGGAAAGCTAT CCGCGCCATTTATGACTTCAGTAAGGAGGGCAAAT ACCACGGCGAAAGCAACAAGTATATCGCGGAACAC CTGATAAACGTGGAGTTCAATCCAAAGGAGTGCAT ATTTGAGGGATACGAACTGGGCGATATCACCGAAT ACAAGAAGGCGGCTCTGAAACTTAATAACTACAAC AATGTCGACTTCGTAATCGCAATAGTCCCGAACAT GTCCGACGAAGAGATAGAGAACAGCTACAATCCGT TCAAGAAAATATGGGCCGAACTGAATCTGCCCAGC CAGATGATTAGCGTCAAGACGGCCGAAATCTTTGC CAATAGCAGGGATAACACGGCGCTTTACTATCTGC ATAACATCGTCCTCGGTATCCTGGGTAAGATAGGA GGGATTCCCTGGGTGGTTAAAGACATGAAGGGCGA CGTGGATTGCTTCGTTGGACTCGATGTCGGCACCA GGGAGAAGGGCATACATTACCCCGCCTGCAGCGTT GTGTTTGACAAGTACGGCAAGCTTATTAACTATTA CAAGCCTAACATCCCGCAGAACGGAGAGAAGATTA ACACAGAAATACTTCAGGAAATTTTCGACAAGGTG CTCATAAGCTATGAGGAGGAGAATGGAGCCTACCC GAAGAATATCGTGATCCACAGGGACGGCTTTAGCC GAGAGGACCTTGACTGGTATGAGAACTACTTCGGT AAGAAAAACATAAAGTTTAACATCATCGAAGTCAA AAAGTCAACTCCGTTGAAAATCGCCAGTATAAACG AGGGAAATATCACGAATCCTGAAAAGGGTTCCTAC ATCCTGCGCGGCAACAAAGCCTACATGGTGACCAC AGATATTAAGGAAAACCTGGGAAGCCCAAAGCCCC TGAAGATAGAAAAGAGCTACGGCGACATAGACATG CTCACAGCTCTCAGCCAAATATACGCACTCACGCA AATCCATGTGGGGGCGACCAAAAGCCTGCGCCTCC CAATCACCACCGGCTACGCCGACAAGATTTGCAAG GCGATCGAGTTCATCCCCCAAGGGCGCGTGGACAA CCGCCTTTTCTTTCTGTAGTGA	

[0380] In some embodiments, the Ago-SSB fusion protein comprises an amino acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 97-101. In some embodiments, the Ago-SSB fusion protein is encoded by a nucleic acid sequence with at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to one of SEQ ID NOS: 102-106. The amino acid sequence of exemplary Ago-SSB fusion polypeptides are provided in Table 12. The nucleic acid sequence of exemplary Ago-SSB fusion polypeptides are provided in Table 13.

Table 12. Amino Acid Sequence of Exemplary Ago-SSB 2XSV40NLS Fusion Polypeptides

Fusion Polypeptide	Amino Acid Sequence	SEQ ID NO
AP109 N to C terminus: *Italicized/Underlined: 2XSV40NLS* Underlined: GSGS linker *Italicized: V5 Tag* **Bold: Ago69**	*MPKKKRKVEDPKKKRKVGSGS**GKPIPNPLLGLDST**GSGSS***MVGGYKVSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKII KSARYKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINF VKSEVLSIEDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRIN YSKEILSNTLLTRELKDEFKKSNKGFNLKRKFRISPVVNKMG KVILYLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISG NLVIESVLETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDED LNANIVNVRGNKKIYMYIPHALKPIITREYLAKNDPEFSKEI EQLIKMNMNYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVK LLGYSSGKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRF GVLYPKEFDGVSRKAIRAIYDFSKEGKYHGESNKYIAEHLIN VEFNPKECIFEGYELGDITEYKKAALKLNNYNNVDFVIAIVP NMSDEEIENSYNPFKKIWAELNLPSQMISVKTAEIFANSRDN TALYYLHNIVLGILGKIGGIPWVVKDMKGDVDCFVGLDVGTR EKGIHYPACSVVFDKYGKLINYYKPNIPQNGEKINTEILQEI FDKVLISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKN IKFNIIEVKKSTPLKIASINEGNITNPEKGSYILRGNKAYMV TTDIKENLGSPKPLKIEKSYGDIDMLTALSQIYALTQIHVGA TKSLRLPITTGYADKICKAIEFIPQGRVDNRLFFL*	97
AP110 N to C terminus: ***Italicized/Underlined/Bold: 2XSV40NLS*** Underlined: GSGS linker *Italicized: V5 Tag* ***Italicized and Bold: Sso SSB*** *Italicized and underlined: GGGGS linker* **Bold: Ago69**	***MPKKKRKVEDPKKKRKV**GSGS**GKPIPNPLLGLDST**GSGSS**ME EKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAIV GDETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQLN AGSKTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRGG GRRYGRRGGRRQENEEGEEE**GGGGS**MVGGYKVSNLTVEAFEG IGSVNPMLFYQYKVTGKGKYDNVYKIIKSARYKMHSKNRFKP VFIKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMSIY GEVVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTREL KDEFKKSNKGFNLKRKFRISPVVNKMGKVILYLSCSADFSTN KNIYEMLKEGLEVEGLAVKSEWSNISGNLVIESVLETKISEP TSLGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIY MYIPHALKPIITREYLAKNDPEFSKEIEQLIKMNMNYRYETL KSFVNDIGVIEELNNLSFKNKYYEDVKLLGYSSGKIDEPVLM GAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKA IRAIYDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYEL GDITEYKKAALKLNNYNNVDFVIAIVPNMSDEEIENSYNPFK KIWAELNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGILG KIGGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDK YGKLINYYKPNIPQNGEKINTEILQEIFDKVLISYEEENGAY PKNIVIHRDGFSREDLDWYENYFGKKNIKFNIIEVKKSTPLK IASINEGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLK IEKSYGDIDMLTALSQIYALTQIHVGATKSLRLPITTGYADK ICKAIEFIPQGRVDNRLFFL*	98

SPL0389 N to C terminus: *Italicized/Underlined: 2XSV40NLS* Underlined: GSGS linker *Italicized: V5 Tag* **Bold: Homologue 2 (HG2)**	*MPKKKRKVEDPKKKRKVGSGS*GKPIPNPLLGLDSTGSGSS**MN** **NLTFEAFEGIGQLNELNFYKYRLIGKGQIDNVHQAIWSVKYK** **LQANNFFKPVFVKGEILYSLDELKVIPEFENVEVILDGNIIL** **SISENTDIYKDVIVFYINNALKNIKDITNYRKYITKNTDEII** **CKSILTTNLKYQYMKSEKGFKLQRKFKISPVVFRNGKVILYL** **NCSSDFSTDKSIYEMLNDGLGVVGLQVKNKWTNANGNIFIEK** **VLDNTISDPGTSGKLGQSLIDYYINGNQKYRVEKFTDEDKNA** **KVIQAKIKNKTYNYIPQALTPVITREYLSHTDKKFSKQIENV** **IKMDMNYRYQTLKSFVEDIGVIKELNNLHFKNQYYTNFDFMG** **FESGVLEEPVLMGANGKIKDKKQIFINGFFKNPKENVKFGVL** **YPEGCMENAQSIARSILDFATAGKYNKQENKYISKNLMNIGF** **KPSECIFESYKLGDITEYKATARKLKEHEKVGFVIAVIPDMN** **ELEVENPYNPFKKVWAKLNIPSQMITLKTTEKFKNIVDKSGL** **YYLHNIALNILGKIGGIPWIIKDMPGNIDCFIGLDVGTREKG** **IHFPACSVLFDKYGKLINYYKPTIPQSGEKIAETILQEIFDN** **VLISYKEENGEYPKNIVIHRDGFSRENIDWYKEYFDKKGIKF** **NIIEVKKNIPVKIAKVVGSNICNPIKGSYVLKNDKAFIVTTD** **IKDGVASPNPLKIEKTYGDVEMKSILEQIYSLSQIHVGSTKS** **LRLPITTGYADKICKAIEYIPQGVVDNRLFFL**	99
SPL0390 N to C terminus: *Italicized/Underlined: 2XSV40NLS* Underlined: GSGS linker *Italicized: V5 Tag* **Bold: Homologue (HG4)**	*MPKKKRKVEDPKKKRKVGSGS*GKPIPNPLLGLDSTGSGSS**MK** **EFNVITEFKNGINSKSIEIYIYKMMVRDFEKRHNENYDVVKE** **LINLNNNSTIVFYEQYIASFKEIEKWGNEQYINVEKRAINLE** **SNEKKILERLLLKEIKNNIDNNKYKVVKDSIYINKPVYNEKG** **IKIDRYFNLDINVESNGDIIIGFDISHNFEYINTLEYEIKNN** **NIKIGDRVKDYFYNLTYEYVGIAPFTISEENEYMGCSIVDYY** **ENKNQSYIVNKLPKDMKAILVKNNKNSIFPYIPSRLKKVCRF** **ENLPQNVLRDFNTRVKQKTNEKMQFMVDEVINIVKNSEHIDV** **KKKNMMCDNIGYKIEDLQQPDLLFGNARAQRYPLYGLKNFGV** **YENKRIEIKYFIDPILAKSKMNLEKISKFCDELEQFSSKLGV** **GLNRVKLNNIVNFKEIRMDNEDIFSYEIRKIVSNYNETTIVI** **LSEENLNKYYNIIKKTFSGGNEVPTQCIGFNTLSYTEKNKDS** **IFLNILLGVYAKSGIQPWILNEKLNSDCFIGLDVSRENKVNK** **AGVIQVVGKDGRVLKTKVISSSQSGEKIKLETLREIVFEAIN** **SYENTYRCKPKHITFHRDGINREELENLKNTMTNLGVEFDYI** **EITKGINRRIATISEGEEWKTIMGRCYYKDNSAYVCTTKPYE** **GIGMAKPIRIRRVFGTLDIEKIVEDAYKLTFMHVGAINKIRL** **PITTYYADLSSTYGNRDLIPTNIDTNCLYFI**	100
SPL0398 N to C terminus: ***Italicized/Bold/Underlined: 2XSV40NLS***	*MGKPIPNPLLGLDSTGSGS****MPKKKRKVEDPKKKRKV****GSGSS****M*** ***EEKVGNLKPNMESVNVTVRVLEASEARQIQTKNGVRTISEAI*** ***VGDETGRVKLTLWGKHAGSIKEGQVVKIENAWTTAFKGQVQL*** ***NAGSKTKIAEASEDGFPESSQIPENTPTAPQQMRGGGRGFRG*** ***GGRRYGRRGGRRQENEEGEEE****GGGGS****PAAKRVKLD****GGGGS***MV **GGYKVSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDNVYKII** **KSARYKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINF** **VKSEVLSIEDNMSIYGEVVEYYINLKLKKVKVLGKYPKYRIN**	101

Underlined: GSGS linker *Italicized: V5 Tag* ***Italicized and Bold: Sso SSB*** *Italicized and underlined: GGGGS linker* ***Italicized/Bold/Underlined: c-Myc NLS*** **Bold: Ago69**	**YSKEILSNTLLTRELKDEFKKSNKGFNLKRKFRISPVVNKMG KVILYLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISG NLVIESVLETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDED LNANIVNVRGNKKIYMYIPHALKPIITREYLAKNDPEFSKEI EQLIKMNMNYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVK LLGYSSGKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRF GVLYPKEFDGVSRKAIRAIYDFSKEGKYHGESNKYIAEHLIN VEFNPKECIFEGYELGDITEYKKAALKLNNYNNVDFVIAIVP NMSDEEIENSYNPFKKIWAELNLPSQMISVKTAEIFANSRDN TALYYLHNIVLGILGKIGGIPWVVKDMKGDVDCFVGLDVGTR EKGIHYPACSVVFDKYGKLINYYKPNIPQNGEKINTEILQEI FDKVLISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKN IKFNIIEVKKSTPLKIASINEGNITNPEKGSYILRGNKAYMV TTDIKENLGSPKPLKIEKSYGDIDMLTALSQIYALTQIHVGA TKSLRLPITTGYADKICKAIEFIPQGRVDNRLFFL**	

Table 13. Amino Acid Sequence of Exemplary Ago-SSB 2XSV40NLS Fusion Polypeptides

Fusion Polypeptide	Nucleic Acid Sequence	SEQ ID NO
AP109 N to C terminus: *Italicized/Underlined: 2XSV40NLS* Underlined: GSGS linker *Italicized: V5 Tag* **Bold: Ago69**	*ATGCCCAAGAAAAAGCGAAAGGTAGAGGACCCCAAAAAGAAA CGCAAAGTGGGCTCCGGAAGCGGGAAGCCCATCCCAAACCCG CTGTTGGGCTTGGATTCCACG*GGCAGCGGAAGCTCT**ATGGTC GGCGGCTATAAAGTCAGCAATTTGACAGTGGAAGCGTTCGAA GGTATCGGGAGTGTCAACCCGATGCTGTTTTACCAATACAAA GTCACCGGAAAGGGAAAGTACGATAATGTGTATAAGATTATC AAAAGCGCACGGTACAAGATGCATTCTAAGAACCGATTCAAG CCCGTGTTCATCAAGGACGACAAACTGTACACCCTCGAGAAG CTCCCGGATATAGAGGACCTGGATTTCGCAAACATTAACTTC GTGAAAAGCGAGGTTCTCAGCATAGAGGATAATATGTCAATT TATGGCGAGGTGGTGGAATACTATATCAATCTCAAGCTGAAA AAAGTGAAGGTGTTGGGAAAATACCCCAAGTACAGGATCAAT TACAGCAAAGAGATTCTCAGTAATACGCTGCTGACACGAGAG CTCAAAGACGAGTTTAAGAAATCAAATAAGGGTTTTAACCTG AAACGGAAGTTTAGAATTTCCCCCGTGGTGAATAAGATGGGC AAAGTGATACTCTATTTGTCCTGCAGTGCTGATTTCAGCACC AACAAGAACATTTACGAAATGTTGAAAGAGGGCTTGGAGGTT GAGGGGCTGGCCGTTAAGAGCGAGTGGAGCAATATCAGTGGC AACCTGGTGATCGAGAGCGTACTGGAAACCAAGATATCCGAG CCCACTAGCCTGGGCCAATCCCTGATAGACTACTATAAGAAT AACAACCAGGGCTATAGGGTGAAGGATTTCACCGATGAGGAT CTGAATGCCAACATTGTCAACGTGAGAGGAAATAAGAAGATC**	102

	TATATGTATATTCCGCACGCGTTGAAGCCGATAATCACCCGG **GAGTACCTGGCCAAGAACGATCCAGAGTTTTCTAAGGAGATC** **GAGCAGCTTATCAAGATGAATATGAACTACCGATATGAAACC** **CTCAAGTCATTTGTGAATGACATCGGGGTCATTGAGGAGCTG** **AACAACCTGAGCTTCAAAAACAAATACTACGAAGATGTGAAA** **CTGCTGGGTTACTCCAGCGGCAAAATAGACGAACCCGTCCTG** **ATGGGGGCAAAAGGGATCATAAAGAACAAAATGCAGATTTTT** **TCCAATGGATTCTACAAACTCCCCGAAGGCAAGGTACGATTT** **GGCGTTCTGTACCCAAAAGAATTTGATGGCGTGTCAAGGAAA** **GCTATCCGCGCCATTTATGACTTCAGTAAGGAGGGCAAATAC** **CACGGCGAAAGCAACAAGTATATCGCGGAACACCTGATAAAC** **GTGGAGTTCAATCCAAAGGAGTGCATATTTGAGGGATACGAA** **CTGGGCGATATCACCGAATACAAGAAGGCGGCTCTGAAACTT** **AATAACTACAACAATGTCGACTTCGTAATCGCAATAGTCCCG** **AACATGTCCGACGAAGAGATAGAGAACAGCTACAATCCGTTC** **AAGAAAATATGGGCCGAACTGAATCTGCCCAGCCAGATGATT** **AGCGTCAAGACGGCCGAAATCTTTGCCAATAGCAGGGATAAC** **ACGGCGCTTTACTATCTGCATAACATCGTCCTCGGTATCCTG** **GGTAAGATAGGAGGGATTCCCTGGGTGGTTAAAGACATGAAG** **GGCGACGTGGATTGCTTCGTTGGACTCGATGTCGGCACCAGG** **GAGAAGGGCATACATTACCCCGCCTGCAGCGTTGTGTTTGAC** **AAGTACGGCAAGCTTATTAACTATTACAAGCCTAACATCCCG** **CAGAACGGAGAGAAGATTAACACAGAAATACTTCAGGAAATT** **TTCGACAAGGTGCTCATAAGCTATGAGGAGGAGAATGGAGCC** **TACCCGAAGAATATCGTGATCCACAGGGACGGCTTTAGCCGA** **GAGGACCTTGACTGGTATGAGAACTACTTCGGTAAGAAAAAC** **ATAAAGTTTAACATCATCGAAGTCAAAAAGTCAACTCCGTTG** **AAAATCGCCAGTATAAACGAGGGAAATATCACGAATCCTGAA** **AAGGGTTCCTACATCCTGCGCGGCAACAAAGCCTACATGGTG** **ACCACAGATATTAAGGAAAACCTGGGAAGCCCAAAGCCCCTG** **AAGATAGAAAAGAGCTACGGCGACATAGACATGCTCACAGCT** **CTCAGCCAAATATACGCACTCACGCAAATCCATGTGGGGGCG** **ACCAAAAGCCTGCGCCTCCCAATCACCACCGGCTACGCCGAC** **AAGATTTGCAAGGCGATCGAGTTCATCCCCCAAGGGCGCGTG** **GACAACCGCCTTTTCTTTCTG**	
AP110 N to C terminus: ***Italicized/Underlined/Bold: 2XSV40NLS*** Underlined: GSGS linker	***ATGCCCAAGAAAAAGCGAAAGGTAGAGGACCCCAAAAAGAAA*** ***CGCAAAGTG***GGCTCCGGAAGC*GGGAAGCCCATCCCAAACCCG* *CTGTTGGGCTTGGATTCCACG****AGTATGGAAGAGAAGGTCGGA*** ***AATCTCAAACCCAACATGGAGAGCGTCAACGTGACTGTCAGA*** ***GTGCTTGAGGCTAGTGAGGCTCGTCAAATACAAACTAAGAAC*** ***GGCGTGCGAACGATCTCTGAAGCAATCGTGGGAGACGAGACA*** ***GGTCGGGTCAAGCTTACACTTTGGGGAAAGCACGCAGGGTCC*** ***ATTAAAGAGGGGCAAGTGGTCAAGATCGAAAATGCATGGACC*** ***ACGGCCTTTAAGGGTCAAGTCCAACTCAACGCTGGCTCTAAA*** ***ACAAAGATCGCGGAAGCCAGCGAGGATGGGTTTCCCGAGTCT***	103

Italicized: V5 Tag ***Italicized and Bold: Sso SSB*** *<u>Italicized and underlined: GGGGS linker</u>* **Bold: Ago69**	***TCCCAAATACCTGAGAATACCCCCACAGCCCCTCAACAAATG*** ***CGCGGCGGCGGAAGGGGTTTTAGGGGTGGCGGGAGACGGTAC*** ***GGGAGAAGGGGCGGAAGGAGACAGGAGAATGAGGAAGGAGAG*** ***GAAGAG****<u>GGTGGAGGAGGAAGT</u>***ATGGTCGGCGGCTATAAAGTC*** **AGCAATTTGACAGTGGAAGCGTTCGAAGGTATCGGGAGTGTC** **AACCCGATGCTGTTTTACCAATACAAAGTCACCGGAAAGGGA** **AAGTACGATAATGTGTATAAGATTATCAAAAGCGCACGGTAC** **AAGATGCATTCTAAGAACCGATTCAAGCCCGTGTTCATCAAG** **GACGACAAACTGTACACCCTCGAGAAGCTCCCGGATATAGAG** **GACCTGGATTTCGCAAACATTAACTTCGTGAAAAGCGAGGTT** **CTCAGCATAGAGGATAATATGTCAATTTATGGCGAGGTGGTG** **GAATACTATATCAATCTCAAGCTGAAAAAAGTGAAGGTGTTG** **GGAAAATACCCCAAGTACAGGATCAATTACAGCAAAGAGATT** **CTCAGTAATACGCTGCTGACACGAGAGCTCAAAGACGAGTTT** **AAGAAATCAAATAAGGGTTTTAACCTGAAACGGAAGTTTAGA** **ATTTCCCCCGTGGTGAATAAGATGGGCAAAGTGATACTCTAT** **TTGTCCTGCAGTGCTGATTTCAGCACCAACAAGAACATTTAC** **GAAATGTTGAAAGAGGGCTTGGAGGTTGAGGGGCTGGCCGTT** **AAGAGCGAGTGGAGCAATATCAGTGGCAACCTGGTGATCGAG** **AGCGTACTGGAAACCAAGATATCCGAGCCCACTAGCCTGGGC** **CAATCCCTGATAGACTACTATAAGAATAACAACCAGGGCTAT** **AGGGTGAAGGATTTCACCGATGAGGATCTGAATGCCAACATT** **GTCAACGTGAGAGGAAATAAGAAGATCTATATGTATATTCCG** **CACGCGTTGAAGCCGATAATCACCCGGGAGTACCTGGCCAAG** **AACGATCCAGAGTTTTCTAAGGAGATCGAGCAGCTTATCAAG** **ATGAATATGAACTACCGATATGAAACCCTCAAGTCATTTGTG** **AATGACATCGGGGTCATTGAGGAGCTGAACAACCTGAGCTTC** **AAAAACAAATACTACGAAGATGTGAAACTGCTGGGTTACTCC** **AGCGGCAAAATAGACGAACCCGTCCTGATGGGGGCAAAAGGG** **ATCATAAAGAACAAAATGCAGATTTTTTCCAATGGATTCTAC** **AAACTCCCCGAAGGCAAGGTACGATTTGGCGTTCTGTACCCA** **AAAGAATTTGATGGCGTGTCAAGGAAAGCTATCCGCGCCATT** **TATGACTTCAGTAAGGAGGGCAAATACCACGGCGAAAGCAAC** **AAGTATATCGCGGAACACCTGATAAACGTGGAGTTCAATCCA** **AAGGAGTGCATATTTGAGGGATACGAACTGGGCGATATCACC** **GAATACAAGAAGGCGGCTCTGAAACTTAATAACTACAACAAT** **GTCGACTTCGTAATCGCAATAGTCCCGAACATGTCCGACGAA** **GAGATAGAGAACAGCTACAATCCGTTCAAGAAAATATGGGCC** **GAACTGAATCTGCCCAGCCAGATGATTAGCGTCAAGACGGCC** **GAAATCTTTGCCAATAGCAGGGATAACACGGCGCTTTACTAT** **CTGCATAACATCGTCCTCGGTATCCTGGGTAAGATAGGAGGG** **ATTCCCTGGGTGGTTAAAGACATGAAGGGCGACGTGGATTGC** **TTCGTTGGACTCGATGTCGGCACCAGGGAGAAGGGCATACAT** **TACCCCGCCTGCAGCGTTGTGTTTGACAAGTACGGCAAGCTT** **ATTAACTATTACAAGCCTAACATCCCGCAGAACGGAGAGAAG** **ATTAACACAGAAATACTTCAGGAAATTTTCGACAAGGTGCTC** **ATAAGCTATGAGGAGGAGAATGGAGCCTACCCGAAGAATATC** **GTGATCCACAGGGACGGCTTTAGCCGAGAGGACCTTGACTGG** **TATGAGAACTACTTCGGTAAGAAAAACATAAAGTTTAACATC**	

	ATCGAAGTCAAAAAGTCAACTCCGTTGAAAATCGCCAGTATA **AACGAGGGAAATATCACGAATCCTGAAAAGGGTTCCTACATC** **CTGCGCGGCAACAAAGCCTACATGGTGACCACAGATATTAAG** **GAAAACCTGGGAAGCCCAAAGCCCCTGAAGATAGAAAAGAGC** **TACGGCGACATAGACATGCTCACAGCTCTCAGCCAAATATAC** **GCACTCACGCAAATCCATGTGGGGGCGACCAAAAGCCTGCGC** **CTCCCAATCACCACCGGCTACGCCGACAAGATTTGCAAGGCG** **ATCGAGTTCATCCCCCAAGGGCGCGTGGACAACCGCCTTTTC** **TTTCTG**	
SPL0389 N to C terminus: *Italicized/Underlined: 2XSV40NLS* Underlined: GSGS linker *Italicized: V5 Tag* **Bold: Homologue 2 (HG2)**	*ATGCCCAAGAAAAAGCGAAAGGTAGAGGACCCCAAAAAGAAA* *CGCAAAGTGGGCTCCGAAGCGGGAAGCCCATCCCAAACCCGC* *TGTTGGGCTTGGATTCCACG*GGCAGCGGAAGCTCT**ATGAATA** **ATCTGACCTTCGAGGCCTTCGAGGGTATCGGACAATTGAACG** **AGTTAAACTTCTATAAGTACCGCCTCATTGGTAAGGGCCAAA** **TCGACAATGTCCACCAGGCCATCTGGTCAGTCAAGTACAAAC** **TTCAAGCGAATAATTTCTTCAAGCCGGTTTTCGTCAAGGGCG** **AAATTCTGTACTCACTTGACGAGCTGAAAGTCATCCCGGAAT** **TCGAGAATGTCGAGGTTATTCTTGACGGGAACATTATCCTGA** **GCATTAGCGAGAACACCGACATTTACAAGGATGTGATCGTGT** **TTTATATCAATAACGCGTTGAAGAACATCAAGGACATCACCA** **ACTACCGTAAGTATATCACTAAGAACACGGATGAAATCATTT** **GCAAGAGTATTTTAACGACGAATCTCAAGTATCAATATATGA** **AGTCAGAGAAAGGGTTCAAGTTACAGCGCAAGTTTAAGATCT** **CCCCGGTGGTATTCCGTAATGGGAAGGTCATCTTGTACCTTA** **ATTGCAGTAGCGACTTCAGCACAGACAAATCCATCTACGAAA** **TGTTAAATGATGGACTCGGTGTTGTGGGCCTGCAAGTGAAGA** **ATAAGTGGACTAATGCGAATGGCAATATCTTTATTGAAAAGG** **TGCTCGACAATACCATCTCCGATCCCGGCACGAGTGGAAAGC** **TGGGGCAGTCCCTGATCGACTACTACATCAATGGGAATCAAA** **AGTACCGTGTAGAGAAATTTACCGACGAGGACAAGAATGCAA** **AGGTTATCCAGGCCAAAATCAAGAATAAAACATACAACTACA** **TCCCGCAAGCTCTCACCCCCGTAATTACGCGCGAGTATCTGA** **GTCATACCGATAAGAAGTTTAGCAAGCAAATCGAGAATGTGA** **TTAAGATGGATATGAACTACCGCTACCAGACGTTGAAGTCTT** **TCGTTGAGGACATTGGCGTGATCAAGGAGTTAAACAATCTGC** **ACTTTAAGAACCAATATTACACCAATTTTGACTTTATGGGGT** **TCGAGAGCGGGGTGCTGGAAGAACCTGTCCTGATGGGTGCGA** **ACGGAAAGATCAAGGACAAGAAGCAGATTTTCATCAATGGGT** **TCTTTAAGAATCCCAAGGAGAACGTAAAATTCGGAGTACTCT** **ACCCAGAAGGCTGTATGGAGAATGCTCAGAGCATTGCTCGTT** **CCATCCTCGACTTCGCTACGGCCGGTAAATACAATAAGCAAG** **AGAACAAGTATATTTCGAAGAATTTAATGAACATCGGATTCA** **AACCTTCTGAGTGTATCTTTGAGTCGTATAAGTTGGGAGACA** **TCACCGAGTATAAGGCGACGGCCCGTAAGCTCAAGGAGCATG** **AGAAAGTTGGGTTCGTTATCGCAGTGATCCCTGACATGAATG** **AGCTGGAAGTCGAGAACCCTTATAACCCCTTCAAGAAGGTCT**	104

	GGGCGAAACTCAATATCCCATCCCAGATGATCACATTGAAGA CCACCGAAAAGTTCAAGAATATCGTCGACAAGTCAGGCTTGT ACTACTTACACAATATCGCCCTTAATATTCTCGGCAAAATCG GCGGAATCCCGTGGATTATTAAAGACATGCCTGGCAACATCG ACTGTTTCATCGGTTTAGACGTCGGCACGCGCGAGAAGGGCA TCCACTTCCCGGCATGTTCTGTGTTGTTCGACAAGTACGGAA AGTTAATCAATTATTACAAGCCGACTATTCCGCAGAGCGGAG AGAAGATTGCTGAGACAATTTTACAGGAGATCTTCGACAACG TGTTAATCAGCTACAAAGAGGAAAACGGGGAGTACCCCAAGA ATATCGTTATCCATCGTGATGGCTTCAGCCGTGAGAACATCG ATTGGTACAAAGAATACTTCGATAAGAAGGGTATCAAGTTCA ACATTATTGAGGTTAAGAAGAACATTCCCGTAAAGATCGCGA AGGTGGTTGGATCCAATATCTGCAACCCGATCAAGGGCTCTT ATGTGCTTAAGAATGATAAGGCATTCATCGTAACCACCGATA TCAAAGACGGTGTGGCTTCTCCAAATCCACTTAAAATCGAGA AAACCTATGGTGACGTTGAGATGAAGAGTATTCTGGAGCAGA TCTACAGTCTGAGCCAAATTCATGTTGGCTCAACCAAGTCCC TGCGTCTTCCTATCACAACGGGATATGCCGATAAGATCTGTA AGGCAATTGAATACATTCCGCAAGGAGTCGTAGACAATCGTT TGTTCTTTCTT	
SPL0390 N to C terminus: *Italicized/Underlined: 2XSV40NLS* Underlined: GSGS linker *Italicized: V5 Tag* **Bold: Homologue 4 (HG4)**	*ATGCCCAAGAAAAAGCGAAAGGTAGAGGACCCCAAAAAGAAA CGCAAAGTG*GGCTCCGGAAGC*GGGAAGCCCATCCCAAACCCG CTGTTGGGCTTGGATTCCACG*GGCAGCGGAAGCTCT**ATGAAG GAGTTTAACGTCATCACAGAGTTCAAGAACGGTATTAATTCG AAGAGCATCGAGATCTATATTTACAAGATGATGGTTCGTGAC TTTGAGAAGCGTCACAATGAAAATTATGACGTGGTAAAAGAG CTTATTAACCTGAACAATAATAGTACGATTGTCTTTTATGAG CAATATATCGCCTCATTCAAGGAAATCGAGAAGTGGGGTAAC GAGCAATACATTAATGTTGAGAAACGCGCAATTAACCTGGAA AGCAACGAGAAGAAGATTCTTGAACGCCTTCTGTTAAAGGAG ATCAAGAACAACATCGATAACAATAAGTACAAGGTAGTGAAG GATTCGATCTACATCAACAAGCCTGTGTATAACGAAAAGGGT ATCAAAATCGACCGCTACTTCAACTTAGACATCAACGTAGAA TCAAACGGAGACATCATTATTGGCTTCGATATTAGCCATAAT TTCGAGTATATTAACACGTTAGAGTACGAAATCAAGAACAAC AATATCAAGATTGGAGACCGCGTAAAGGATTACTTTTACAAC CTTACTTATGAATATGTTGGCATCGCGCCGTTCACTATTTCC GAAGAGAATGAATATATGGGATGTAGCATCGTGGACTACTAT GAAAATAAGAACCAGAGCTACATCGTGAACAAGTTGCCAAAG GATATGAAGGCAATCTTAGTTAAGAACAATAAGAACAGCATT TTCCCGTACATCCCTTCACGTCTTAAGAAGGTTTGTCGTTTC GAGAATCTGCCCCAAAACGTACTCCGTGATTTTAACACGCGC GTCAAGCAGAAAACTAATGAGAAGATGCAATTTATGGTGGAC GAGGTTATCAACATTGTAAAGAATAGCGAGCATATCGACGTA**	105

	AAGAAGAAGAACATGATGTGTGACAATATCGGGTACAAGATT **GAGGACCTGCAACAACCTGACCTTTTGTTTGGAAACGCCCGC** **GCGCAGCGTTACCCACTGTATGGATTGAAGAACTTTGGCGTG** **TACGAAAACAAGCGCATTGAAATCAAGTACTTTATCGACCCG** **ATTCTCGCCAAGAGCAAGATGAATCTGGAAAAGATCTCCAAG** **TTCTGTGATGAGCTGGAGCAGTTTAGCTCCAAGTTAGGAGTA** **GGATTAAATCGCGTAAAACTGAACAATATTGTTAACTTCAAG** **GAGATTCGTATGGACAATGAGGACATCTTCTCCTACGAGATT** **CGCAAAATTGTGAGCAACTATAATGAGACAACGATCGTGATT** **CTGTCGGAAGAGAACCTTAATAAGTATTACAACATCATTAAG** **AAAACCTTCAGCGGTGGCAACGAGGTTCCGACGCAATGCATT** **GGTTTCAACACACTTTCCTACACGGAGAAGAACAAGGACTCA** **ATTTTCTTAAATATTTTACTTGGTGTTTACGCCAAGTCAGGA** **ATCCAACCGTGGATCCTCAATGAGAAATTGAATTCCGACTGT** **TTCATTGGTTTAGATGTCTCCCGTGAGAATAAGGTAAACAAG** **GCCGGCGTCATTCAAGTTGTCGGAAAAGATGGCCGCGTACTC** **AAGACCAAGGTCATCAGTTCGAGCCAAAGCGGGGAGAAGATC** **AAGCTGGAAACGTTACGCGAGATCGTGTTCGAGGCGATTAAC** **TCGTATGAGAATACCTACCGCTGTAAACCAAAACACATTACA** **TTCCACCGTGACGGTATTAATCGTGAGGAGCTGGAGAATCTT** **AAGAATACGATGACCAATCTTGGTGTTGAGTTTGACTACATC** **GAGATCACCAAGGGCATTAACCGCCGCATTGCCACCATCAGT** **GAGGGCGAGGAGTGGAAGACTATCATGGGCCGCTGTTATTAT** **AAGGACAATTCTGCCTACGTCTGCACTACTAAGCCTTATGAG** **GGAATCGGAATGGCAAAGCCCATTCGCATCCGCCGCGTGTTT** **GGCACGCTTGATATCGAGAAAATTGTTGAAGACGCGTATAAA** **CTTACTTTTATGCATGTAGGCGCGATCAATAAAATTCGTCTT** **CCAATTACAACCTATTACGCAGATCTCAGCTCCACTTACGGA** **AATCGCGACTTAATTCCGACGAATATTGATACCAATTGCCTC** **TACTTCATT**	
SPL0398 N to C terminus: ***Italicized/Bold/Underlined: 2XSV40NLS*** Underlined: GSGS linker *Italicized: V5 Tag* ***Italicized and Bold: Sso SSB***	*ATGGGCAAACCAATACCTAACCCACTCCTCGGACTGGACTCT* *ACC*GGGAGTGGCTCC***ATGCCAAAGAAGAAAAGGAAAGTGGAA*** ***GATCCTAAGAAGAAGCGAAAGGTC***GGTAGCGGTTCA***AGTATG*** ***GAAGAGAAGGTCGGAAATCTCAAACCCAACATGGAGAGCGTC*** ***AACGTGACTGTCAGAGTGCTTGAGGCTAGTGAGGCTCGTCAA*** ***ATACAAACTAAGAACGGCGTGCGAACGATCTCTGAAGCAATC*** ***GTGGGAGACGAGACAGGTCGGGTCAAGCTTACACTTTGGGGA*** ***AAGCACGCAGGGTCCATTAAAGAGGGGCAAGTGGTCAAGATC*** ***GAAAATGCATGGACCACGGCCTTTAAGGGTCAAGTCCAACTC*** ***AACGCTGGCTCTAAAACAAAGATCGCGGAAGCCAGCGAGGAT*** ***GGGTTTCCCGAGTCTTCCCAAATACCTGAGAATACCCCCACA*** ***GCCCCTCAACAAATGCGCGGCGGCGGAAGGGGTTTTAGGGGT*** ***GGCGGGAGACGGTACGGGAGAAGGGGCGGAAGGAGACAGGAG*** ***AATGAGGAAGGAGAGGAAGAG****GGTGGAGGAGGAAGT****CCAGCA*** ***GCCAAACGGGTCAAGCTTGAC****GGCGGCGGCGGGTCT***ATGGTG	106

Italicized and underlined: GGGGS linker ***Italicized/Bold/Underlined: c-Myc NLS*** **Bold: Ago69**	**GGCGGGTACAAGGTGTCAAACCTCACTGTCGAGGCCTTTGAG** **GGGATTGGCTCAGTAAATCCCATGCTCTTCTATCAGTATAAG** **GTGACAGGCAAAGGCAAATATGACAACGTCTACAAAATCATT** **AAGTCCGCCAGATATAAAATGCACTCCAAGAACCGGTTTAAA** **CCTGTATTCATTAAAGATGACAAGCTGTACACCCTCGAGAAG** **CTCCCGGATATAGAGGACCTGGATTTCGCAAACATTAACTTC** **GTGAAAAGCGAGGTTCTCAGCATAGAGGATAATATGTCAATT** **TATGGCGAGGTGGTGGAATACTATATCAATCTCAAGCTGAAA** **AAAGTGAAGGTGTTGGGAAAATACCCCAAGTACAGGATCAAT** **TACAGCAAAGAGATTCTCAGTAATACGCTGCTGACACGAGAG** **CTCAAAGACGAGTTTAAGAAATCAAATAAGGGTTTTAACCTG** **AAACGGAAGTTTAGAATTTCCCCCGTGGTGAATAAGATGGGC** **AAAGTGATACTCTATTTGTCCTGCAGTGCTGATTTCAGCACC** **AACAAGAACATTTACGAAATGTTGAAAGAGGGCTTGGAGGTT** **GAGGGGCTGGCCGTTAAGAGCGAGTGGAGCAATATCAGTGGC** **AACCTGGTGATCGAGAGCGTACTGGAAACCAAGATATCCGAG** **CCCACTAGCCTGGGCCAATCCCTGATAGACTACTATAAGAAT** **AACAACCAGGGCTATAGGGTGAAGGATTTCACCGATGAGGAT** **CTGAATGCCAACATTGTCAACGTGAGAGGAAATAAGAAGATC** **TATATGTATATTCCGCACGCGTTGAAGCCGATAATCACCCGG** **GAGTACCTGGCCAAGAACGATCCAGAGTTTTCTAAGGAGATC** **GAGCAGCTTATCAAGATGAATATGAACTACCGATATGAAACC** **CTCAAGTCATTTGTGAATGACATCGGGGTCATTGAGGAGCTG** **AACAACCTGAGCTTCAAAAACAAATACTACGAAGATGTGAAA** **CTGCTGGGTTACTCCAGCGGCAAAATAGACGAACCCGTCCTG** **ATGGGGGCAAAAGGGATCATAAAGAACAAAATGCAGATTTTT** **TCCAATGGATTCTACAAACTCCCCGAAGGCAAGGTACGATTT** **GGCGTTCTGTACCCAAAAGAATTTGATGGCGTGTCAAGGAAA** **GCTATCCGCGCCATTTATGACTTCAGTAAGGAGGGCAAATAC** **CACGGCGAAAGCAACAAGTATATCGCGGAACACCTGATAAAC** **GTGGAGTTCAATCCAAAGGAGTGCATATTTGAGGGATACGAA** **CTGGGCGATATCACCGAATACAAGAAGGCGGCTCTGAAACTT** **AATAACTACAACAATGTCGACTTCGTAATCGCAATAGTCCCG** **AACATGTCCGACGAAGAGATAGAGAACAGCTACAATCCGTTC** **AAGAAAATATGGGCCGAACTGAATCTGCCCAGCCAGATGATT** **AGCGTCAAGACGGCCGAAATCTTTGCCAATAGCAGGGATAAC** **ACGGCGCTTTACTATCTGCATAACATCGTCCTCGGTATCCTG** **GGTAAGATAGGAGGGATTCCCTGGGTGGTTAAAGACATGAAG** **GGCGACGTGGATTGCTTCGTTGGACTCGATGTCGGCACCAGG** **GAGAAGGGCATACATTACCCCGCCTGCAGCGTTGTGTTTGAC** **AAGTACGGCAAGCTTATTAACTATTACAAGCCTAACATCCCG** **CAGAACGGAGAGAAGATTAACACAGAAATACTTCAGGAAATT** **TTCGACAAGGTGCTCATAAGCTATGAGGAGGAGAATGGAGCC** **TACCCGAAGAATATCGTGATCCACAGGGACGGCTTTAGCCGA** **GAGGACCTTGACTGGTATGAGAACTACTTCGGTAAGAAAAAC**	

	ATAAAGTTTAACATCATCGAAGTCAAAAAGTCAACTCCGTTG AAAATCGCCAGTATAAACGAGGGAAATATCACGAATCCTGAA AAGGGTTCCTACATCCTGCGCGGCAACAAAGCCTACATGGTG ACCACAGATATTAAGGAAAACCTGGGAAGCCCAAAGCCCCTG AAGATAGAAAAGAGCTACGGCGACATAGACATGCTCACAGCT CTCAGCCAAATATACGCACTCACGCAAATCCATGTGGGGGCG ACCAAAAGCCTGCGCCTCCCAATCACCACCGGCTACGCCGAC AAGATTTGCAAGGCGATCGAGTTCATCCCCCAAGGGCGCGTG GACAACCGCCTTTTCTTTCTG	

VIII. Regulatory domain polypeptide (RDP)

[0381] In some cases, a regulatory domain polypeptide is part of a nucleic acid editing system. An RDP can regulate a level of an activity, such as editing, of a nucleic acid editing system. Non-limiting examples of RDPs include recombinases, epigenetic modulators, germ cell repair domains, or DNA repair proteins. In some cases, an RDP is mined by screening for co-localized DNA repair proteins in a region comprising an RNase-H like domain containing polypeptide. In some embodiments, the Agos described herein are an RNase-H like domain containing polypeptide.

[0382] Exemplary recombinases that can be used as RDPs include Cre, Hin, Tre, or FLP recombinases. In some cases, recombinases involved in homologous recombination are utilized. For example, in some embodiments, the RDP is RadA, Rad51, RecA, Dmc1, or UvsX.

[0383] In some embodiments, an epigenetic modulator is a protein that can modify an epigenome directly through DNA methylation, post-translational modification of chromatin, or by altering a structure of chromatin.

[0384] Exemplary germ cell repair domains include ATM, ATR, or DNA-PK to name a few. A germ cell repair domain can repair DNA damage though a variety of mechanisms such as nucleotide excision repair (NER), base excision repair (BER), mismatch repair (MMR), DNA double strand break repair (DSBR), and post replication repair (PRR).

[0385] An RDP can be a tunable component of a nucleic acid editing system. For example, an RDP can be swapped in the editing system to achieve a particular outcome. In some cases, an RDP can be selected based on a cell to be targeted, a level of editing efficiency that is sought, or in order to reduce off-target effects of a nucleic acid editing system. A dialing up or a tuning can enhance a parameter (efficiency, safety, speed, or accuracy) of a genomic break repair by about 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or up to about 100% as

compared to a comparable gene editing system. A dialing down or a tuning can be performed by interchanging a domain such as an RDP to achieve a different effect during a genomic modification. For example, a different effect may be a skewing towards a particular genomic break repair, a recombination, an epigenetic modulation, or a high fidelity repair. In some cases, an RDP may be used to enhance a transgene insertion into a genomic break. In some cases, interchanging a module of a gene editing system can allow for HDR of a double strand break as opposed to NHEJ or MMEJ. Use of a gene editing system disclosed herein can allow for preferential HDR of a double strand break over that of comparable or alternate gene editing systems. In some cases, an HDR repair can preferentially occur in a population of cells from about 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or up to about 100% over that which occurs in a comparable gene editing system without said RDP.

[0386] In some cases, the disclosed editing system utilizing an RDP can reduce a thermodynamic energetic requirement by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 40%, 50%, or up to about 60% as compared to a system that does not employ the disclosed RDP. In some cases, the disclosed editing system utilizing an RDP can reduce an immune response to the RDP by about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 40%, 50%, or up to about 60% as compared to a system that does not employ the disclosed RDP. In some cases, an RDP can be harvested from bacteria that are endogenously present in the human body to prevent eliciting an immune response.

IX. Guiding Polynucleic Acid and Target Polynucleic Acid

[0387] The guiding polynucleic acid can direct a gene editing system comprising the Ago polypeptide to a genomic location. The guiding polynucleic acid can direct a nucleic acid-cleaving activity of the described Ago polypeptides. The guiding polynucleic acid can also be capable of interacting with the Ago polypeptide. In some cases, the guiding polynucleic acid can be a DNA. In other cases, the guiding polynucleic acid can be RNA. The guiding polynucleic acid can be a combination of DNA and RNA. The guiding polynucleic acid can be single stranded, double stranded, or a combination thereof. The guiding polynucleic acid can be at least or at least about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides long. The guiding polynucleotide can be at most or at most about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides long. The guiding polynucleotide can be about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides long. In some cases, the guiding

polynucleic acid may be truncated. Truncated guiding polynucleic acids can be utilized to determine a minimum binding length.

[0388] The system described herein can comprise an exogenous guiding polynucleic acid. The system can comprise a non-naturally occurring guiding polynucleic acid. The system can also comprise a naturally occurring guiding polynucleic acid. In some cases, the system comprises one guiding polynucleic acid. In other cases, the system comprises two guiding polynucleic acids, each targeting an opposite strand of a double-stranded target polynucleic acid. In still other cases, the system comprises two or more guiding polynucleic acids targeting different sequences in the target polynucleic acid.

[0389] The guiding polynucleic acid can be a guide RNA (*i.e.*, "gRNA") that can associate with and direct an Ago polypeptide, or the Ago containing complex, to a specific target sequence within a target nucleic acid by virtue of hybridization to a target site of the target nucleic acid. Similarly the guiding polynucleic acid can be a guide RNA (*i.e.*, "gDNA") that can associate with and direct the Ago polypeptide or complex to a specific target sequence within a target nucleic acid by virtue of hybridization to a target site of the target nucleic acid. In some cases, the guiding polynucleic acid can hybridize with a mismatch between the guiding polynucleic acid and a target nucleic acid. The guiding polynucleic acid can comprise at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 25, 30, 35, or up to 40 mismatches when hybridized to a target nucleic acid. In some cases, the guiding polynucleic acid can tolerate mismatches in a recruiting domain, for example at g6, g7, and g8. In some cases, the guiding polynucleic acid can contain mismatches in a stabilization domain. A stabilization domain can be adjacent to a 3' end of the guiding molecule. For example, positions g6-g16, such as g6, g7, g8, g9, g10, g11, g12, g13, g14, g15, and g16 or any combination thereof, can be mismatched in 16 nucleotide long guide molecules. Mismatches in a recruiting domain can have mismatches preferably in positions g6, g7, and/or g8.

[0390] A method disclosed herein also can comprise introducing into a cell or embryo at least one guide RNA or polynucleic acid, *e.g.*, DNA encoding at least one guide RNA. A guide RNA can interact with a RNA-guided endonuclease to direct the endonuclease to a specific target site, at which site the 5' end of the guide RNA base pairs with a specific protospacer sequence in a chromosomal sequence. Similarly, the method can comprise introducing into a cell or embryo at least one guide DNA or polynucleic acid, e.g., RNA that is complementary to the guide DNA. A guide DNA can interact with a DNA-guided endonuclease to direct the endonuclease to a specific

target site. A guide DNA, or a DNA sequence that translates to a guide RNA, can be on the same polynucleic acid molecule that encodes for a chimeric polypeptide as described herein, or on a separate polynucleic acid molecule.

[0391] A guide RNA can comprise two RNAs, *e.g.*, CRISPR RNA (crRNA) and transactivating crRNA (tracrRNA). A guide RNA can sometimes comprise a single-guide RNA (sgRNA) formed by fusion of a portion (*e.g.*, a functional portion) of crRNA and tracrRNA.
A guide RNA can also be a dual RNA comprising a crRNA and a tracrRNA. A guide RNA can comprise a crRNA and lack a tracrRNA. Furthermore, a crRNA can hybridize with a target DNA or protospacer sequence. A guide DNA can be double-stranded or single-stranded DNA.

[0392] As discussed above, a guide RNA can be an expression product. For example, a DNA that encodes a guide RNA can be a vector comprising a sequence coding for the guide RNA. A guide RNA can be transferred into a cell or organism by transfecting the cell or organism with an isolated guide RNA or plasmid DNA comprising a sequence coding for the guide RNA and a promoter. Similarly, a guide DNA can be transferred into a cell or organism by transfecting the cell or organism with an isolated guide DNA or RNA that is complementary to the guide DNA. A guide RNA or DNA can also be transferred into a cell or organism in other way, such as using virus-mediated gene delivery.

[0393] The guiding polynucleic acid can be isolated. For example, a guide RNA or DNA can be transfected in the form of an isolated RNA or DNA into a cell or organism. A guide RNA or DNA can be prepared by *in vitro* transcription using any *in vitro* transcription system.
A guide RNA can be transferred to a cell in the form of isolated RNA rather than in the form of plasmid comprising encoding sequence for a guide RNA.

[0394] A guide RNA or DNA can comprise a DNA-targeting segment and a protein binding segment. A DNA-targeting segment (or DNA-targeting sequence, or spacer sequence) comprises a nucleotide sequence that can be complementary to a specific sequence within a target DNA (*e.g.*, a protospacer). A protein-binding segment (or protein-binding sequence) can interact with a site-directed modifying polypeptide, *e.g.* an RNA-guided endonuclease such as a Cas protein. By "segment" it is meant a segment/section/region of a molecule, *e.g.*, a contiguous stretch of nucleotides in RNA. A segment can also mean a region/section of a complex such that a segment can comprise regions of more than one molecule. For example, in some cases a protein-binding segment of a DNA-targeting RNA is one RNA molecule and the protein-binding segment therefore comprises a region of that RNA molecule. In other cases, the protein-binding segment

of a DNA-targeting RNA comprises two separate molecules that are hybridized along a region of complementarity.

[0395] The guiding polynucleic acid can comprise two separate polynucleic acid molecules or a single polynucleic acid molecule. An exemplary single molecule guiding polynucleic acid (*e.g.*, guide RNA) comprises both a DNA-targeting segment and a protein-binding segment.

[0396] In some cases, the Ago polypeptide or portion thereof can form a complex with the guiding polynucleic acid. In some cases, the system described herein comprises a complex comprising the Ago polypeptide and the guiding polynucleic acid. The guiding polynucleic acid can provide target specificity to a complex by comprising a nucleotide sequence that can be complementary to a sequence of a target nucleic acid. In some cases, a target nucleic acid can comprise at least a portion of a gene. In some cases, a target nucleic acid can be within an exon of a gene. In other cases, a target nucleic acid can be within an intron of a gene.

[0397] The guiding polynucleic acid can complex with the Ago polypeptide to provide the Ago polypeptide site-specific activity. In other words, the Ago polypeptide can be guided to a target site within a single stranded target nucleic acid sequence *e.g.* a single stranded region of a double stranded nucleic acid, a chromosomal sequence or an extrachromosomal sequence, *e.g.* an episomal sequence, a minicircle sequence, a mitochondrial sequence, a chloroplast sequence, an ssRNA, an ssDNA, etc. by virtue of its association with the guiding polynucleic acid.

[0398] In some cases, the guiding polynucleic acid can comprise one or more modifications (*e.g.*, a base modification, a backbone modification), to provide the nucleic acid with a new or enhanced feature (*e.g.*, improved stability). The guiding polynucleic acid can comprise a nucleic acid affinity tag. A nucleoside can be a base-sugar combination. A base portion of the nucleoside can be a heterocyclic base. The two most common classes of such heterocyclic bases can be purines and pyrimidines. Nucleotides can be nucleosides that further include a phosphate group covalently linked to a sugar portion of a nucleoside. For those nucleosides that include a pentofuranosyl sugar, a phosphate group can be linked to the 2', the 3', or the 5' hydroxyl moiety of a sugar. In forming guiding polynucleic acids, a phosphate group can covalently link adjacent nucleosides to one another to form a linear polymeric compound. In addition, linear compounds may have internal nucleotide base complementarity and may therefore fold in a manner as to produce a fully or partially double-stranded compound. Within guiding polynucleic acids, a phosphate groups can commonly be referred to as forming a internucleoside backbone of the guiding polynucleic acid. The linkage or backbone of the guiding polynucleic acid can be a 3' to

5' phosphodiester linkage. In some cases, the guiding polynucleic acid can comprise nucleoside analogs, which can be oxy- or deoxy-analogues of a naturally-occurring DNA and RNA nucleosides deoxycytidine, deoxyuridine, deoxyadenosine, deoxyguanosine and thymidine. The guiding polynucleic acid can also include a universal base, such as deoxyinosine, or 5-nitroindole.The guiding polynucleic acid can comprise a modified backbone and/or modified internucleoside linkages. Modified backbones can include those that can retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone. Suitable modified guiding polynucleic acid backbones containing a phosphorus atom therein can include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates such as 3'-alkylene phosphonates, 5'-alkylene phosphonates, chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, phosphorodiamidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', a 5' to 5' or a 2' to 2' linkage. Suitable guiding polynucleic acids having inverted polarity can comprise a single 3' to 3' linkage at the 3'-most internucleotide linkage (*i.e.* a single inverted nucleoside residue in which the nucleobase is missing or has a hydroxyl group in place thereof).

[0399] In some cases, the guiding polynucleic acid (*e.g.*, a guide RNA or DNA) can also comprise a tail region at a 5' or 3' end that can be essentially single-stranded. For example, a tail region is sometimes not complementarity to any chromosomal sequence in a cell of interest and can sometimes not be complementary to the rest of a guide polynucleic acid. Further, the length of a tail region can vary. A tail region can be more than or more than about 4 nucleotides in length. For example, the length of a tail region can range from or from about 5 to from or from about 60 nucleotides in length.

[0400] In some cases, the guiding polynucleic acid can bind to a region of a genome adjacent to a protospacer adjacent motif (PAM). A guide nucleic acid can comprise a nucleotide sequence (*e.g.*, a spacer), for example, at or near a 5' end or 3' end, that can hybridize to a sequence in a target nucleic acid (*e.g.*, a protospacer). A spacer of a guide nucleic acid can interact with a target nucleic acid in a sequence-specific manner via hybridization (*i.e.*, base pairing). A spacer sequence can hybridize to a target nucleic acid that is located 5' or 3' of a protospacer adjacent motif (PAM). The length of a spacer sequence can be at least or at least about 5, 10, 15, 16, 17,

18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides. The length of a spacer sequence can be at most or at most about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides. In some cases, the guiding polynucleic acid can bind to a region from about 1 to about 20 base pairs adjacent to a PAM. In other cases, the guiding polynucleic acid can bind from about 1, 2,3,4,5,6,7,8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or up to 85 base pairs away from a PAM. Generally, the guiding polynucleic acid binding region can be designed to complement or substantially complement the target nucleic acid sequence or sequences. In some cases, a binding region of the guiding polynucleic acid can incorporate wobble or degenerate bases to bind multiple sequences. In some cases, the binding region can be altered to increase stability. For example, non-natural nucleotides can be incorporated to increase RNA resistance to degradation. In some cases, the binding region can be altered or designed to avoid or reduce secondary structure formation in the binding region. In some cases, the binding region can be designed to optimize G-C content. In some cases, G-C content is preferably between about 40% and about 60% (*e.g.*, 40%, 45%, 50%, 55%, and 60%). In some cases, the binding region can contain modified nucleotides such as, without limitation, methylated or phosphorylated nucleotides.

[0401] In some cases, the guiding polynucleic acid can also comprise a double strand duplex region that can form a secondary structure. For example, a secondary structure formed by the guiding polynucleic acid can comprise a stem (or hairpin) and a loop. A length of a loop and a stem can vary. For example, a loop can range from about 3 to about 10 nucleotides in length, and a stem can range from about 6 to about 20 base pairs in length. A stem can comprise one or more bulges of 1 to about 10 nucleotides. The overall length of a second region can range from about 16 to about 60 nucleotides in length. For example, a loop can be or can be about 4 nucleotides in length and a stem can be or can be about 12 base pairs. In some cases, a 5 ' stem-loop region can be between about 15 and about 50 nucleotides in length (*e.g.*, about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or about 50 nucleotides in length). In some cases, a 5' stem-loop region is between about 30-45 nucleotides in length (e.g., about 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, or 45 nucleotides in length). In some cases, a 5 ' stem- loop region is at least about 31 nucleotides in length (e.g., at least about 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, or 45 nucleotides in length). In some cases, a 5' stem-loop structure contains one or more loops or bulges, each loop or bulge of about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides. In some cases, a 5' stem-loop

structure contains a stem of between about 10 and 30 complementary base pairs (e.g., 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 complementary base pairs). In some cases, a 5 ' stem-loop structure can contain protein-binding, or small molecule-binding structures. In some cases, a 5 ' stem- loop function (e.g., interacting or assembling with the guiding polynucleic acid-guided nuclease) can be conditionally activated by drugs, growth factors, small molecule ligands, or a protein that binds to the protein-binding structure of the 5' stem- loop. In some cases, a 5 ' stem-loop structure can contain non-natural nucleotides. For example, non-natural nucleotides can be incorporated to enhance protein-RNA interaction, protein DNA interaction, or to increase the thermal stability or resistance to degradation of the guiding polynucleic acid.

[0402] In some cases, the guiding polynucleic acid may have an intervening sequence between the 5' and 3' stem- loop structures that can be between about 10 and about 50 nucleotides in length (e.g., about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or about 50 nucleotides in length). In some cases, the intervening sequence is designed to be linear, unstructured, substantially linear, or substantially unstructured. In some embodiments, the intervening sequence can contain non-natural nucleotides. For example, non-natural nucleotides can be incorporated to enhance protein-RNA interaction or to increase the activity of the gRNA: nuclease complex. Similarly, non-natural nucleotides can be incorporated to enhance protein-DNA interaction or to increase the activity of the gDNA: nuclease complex. As another example, natural nucleotides can be incorporated to enhance the thermal stability or resistance to degradation of the gRNA or gDNA. In some cases, a 3 ' stem-loop structure can contain about 3, 4, 5, 6, 7, or 8 nucleotide loop and an about 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotide or longer stem. In some cases, the 3 ' stem-loop can contain a protein-binding, small molecule-binding, hormone-binding, or metabolite-binding structure that can conditionally stabilize the secondary and/or tertiary structure of the gRNA or gDNA. In some embodiments, the 3 ' stem- loop can contain non-natural nucleotides. For example, non-natural nucleotides can be incorporated to enhance protein- guiding nucleic acid interaction or to increase the activity of the guiding polynucleic acid: nuclease complex. As another example, natural nucleotides can be incorporated to enhance the thermal stability or resistance to degradation of the gRNA or gDNA.

[0403] In some cases, the guiding polynucleic acid can include a termination structure at its 3 ' end. In some cases, the guiding polynucleic acid can include an additional 3 ' hairpin structure, e.g., before the termination structure, that can interact with proteins, small-molecules, hormones, or the like, for stabilization or additional functionality, such as conditional stabilization or conditional regulation of a guiding polynucleic acid: nuclease assembly or activity. In some cases, the guiding polynucleic acid can be optimized to enhance stability, assembly, and/or expression. In some case, the guiding polynucleic acid can be optimized to enhance the activity of the guiding polynucleic acid: nuclease complex as compared to control or comparable guiding polynucleic acid: nuclease structures (gRNA, CRISPR RNP, unmodified gRNA, or unmodified guiding polynucleic acids). In some cases, the guiding polynucleic acid can be optimized for expression by substituting, deleting, or adding one or more nucleotides. In some cases, a nucleotide sequence that provides inefficient transcription from an encoding template nucleic acid can be deleted or substituted. For example, in some cases, the guiding polynucleic acid can be transcribed from a nucleic acid operably linked to an RNA polymerase III promoter. In some cases, the guiding polynucleic acid can be modified for increased stability. Stability can be enhanced by optimizing the stability of the guiding polynucleic acid: nuclease interaction, optimizing assembly of the guiding polynucleic acid: nuclease complex, removing or altering RNA or DNA destabilizing sequence elements, or adding RNA or DNA stabilizing sequence elements. In some embodiments, the guiding polynucleic acid can contain a 5' stem-loop structure proximal to, or adjacent to, the binding region that interacts with the guiding polynucleic acid-guided nuclease. Optimization of the 5 ' stem-loop structure can provide enhanced stability or assembly of the guiding polynucleic acid: nuclease complex. In some cases, the 5 ' stem-loop structure is optimized by increasing the length of the stem portion of the stem-loop structure. For example, a 5 ' stem-loop optimization can be combined with mutations for increased transcription to provide an optimized guiding polynucleic acid. For example, an A-U flip and an elongated stem loop can be combined to provide an optimized guiding polynucleic acid.

[0404] A double stranded-guiding polynucleic acid duplex region can comprise a protein-binding segment that can form a complex with an RNA or DNA-binding protein, such as an Argonaute protein, polypeptide, or functional portion thereof.

[0405] In some cases, the guiding polynucleic acid can comprise a modification. A modification can be a chemical modification. A modification can be selected from 5'adenylate, 5' guanosine-

triphosphate cap, 5'N7-Methylguanosine-triphosphate cap, 5'triphosphate cap, 3'phosphate, 3'thiophosphate, 5'phosphate, 5'thiophosphate, Cis-Syn thymidine dimer, trimers, C12 spacer, C3 spacer, C6 spacer, dSpacer, PC spacer, rSpacer, Spacer 18, Spacer 9,3'-3' modifications, 5'-5' modifications, abasic, acridine, azobenzene, biotin, biotin BB, biotin TEG, cholesteryl TEG, desthiobiotin TEG, DNP TEG, DNP-X, DOTA, dT-Biotin, dual biotin, PC biotin, psoralen C2, psoralen C6, TINA, 3'DABCYL, black hole quencher 1, black hole quencer 2, DABCYL SE, dT-DABCYL, IRDye QC-1, QSY-21, QSY-35, QSY-7, QSY-9, carboxyl linker, thiol linkers, 2'deoxyribonucleoside analog purine, 2'deoxyribonucleoside analog pyrimidine, ribonucleoside analog, 2'-0-methyl ribonucleoside analog, sugar modified analogs, wobble/universal bases, fluorescent dye label, 2'fluoro RNA, 2'O-methyl RNA, methylphosphonate, phosphodiester DNA, phosphodiester RNA, phosphothioate DNA, phosphorothioate RNA, UNA, pseudouridine-5'-triphosphate, 5-methylcytidine-5'-triphosphate, 2-O-methyl 3phosphorothioate or any combinations thereof. A modification can be a pseudouridine modification. In some cases, a modification cannot affect viability.

[0406] In some cases, a modification is a 2-O-methyl 3 phosphorothioate addition. A 2-O-methyl 3 phosphorothioate addition can be performed from 1 base to 150 bases. A 2-O-methyl 3 phosphorothioate addition can be performed from 1 base to 4 bases. A 2-O-methyl 3 phosphorothioate addition can be performed on 2 bases. A 2-O-methyl 3 phosphorothioate addition can be performed on 4 bases. A modification can also be a truncation. A truncation can be a 5 base truncation. Guiding polynucleic acids can be modified by methods known in the art. In some cases, the modifications can include, but are not limited to, the addition of one or more of the following sequence elements: a 5 ' cap (*e.g.*, a 7-methylguanylate cap); a 3 ' polyadenylated tail; a riboswitch sequence; a stability control sequence; a hairpin; a subcellular localization sequence; a detection sequence or label; or a binding site for one or more proteins. Modifications can also include the introduction of non-natural nucleotides including, but not limited to, one or more of the following: fluorescent nucleotides and methylated nucleotides. In some embodiments, the guiding polynucleic acid can contain from 5' to 3': (i) a binding region of between about 10 and about 50 nucleotides; (ii) a 5' hairpin region containing fewer than four consecutive uracil nucleotides, or a length of at least 31 nucleotides (e.g., from about 31 to about 41 nucleotides); (iii) a 3' hairpin region; and (iv) a transcription termination sequence, wherein the small guide RNA is configured to form a complex with the guiding polynucleic acid-guided nuclease, the complex having increased stability or activity relative to an unmodified complex.

[0407] A guide RNA or guide DNA can target a nucleic acid sequence of or of about 20 nucleotides. A target nucleic acid can be less than or less than about 20 nucleotides. A target nucleic acid can be at least or at least about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides. A target nucleic acid can be at most or at most about 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 or more nucleotides. A target nucleic acid sequence can be or can be about 20 bases immediately 5' of the first nucleotide of the PAM. A guide RNA or guide DNA can target a nucleic acid sequence comprising a gene or portion thereof.

[0408] A guide RNA or guide DNA can target a genomic sequence comprising a gene. A gene that can be targeted can be involved in a disease. A disease can be a cancer, a cardiovascular condition, a reproductive condition, a neurological disease, an immunological disease, an organ condition, degeneration, an ocular condition, diabetes, a vascular condition, or a gastrointestinal condition. A gene that can be targeted can be involved in a signaling biochemical pathway.

[0409] In some cases, the target polynucleic acid comprises a sequence of the gene to be targeted. The target polynucleic acid can be a sequence of the gene that is associated with a disease or disorder.

[0410] A gene that can be disrupted can be a member of a family of genes. For example, a gene that can be disrupted can improve therapeutic potential of cancer immunotherapy. A gene that can be disrupted can ameliorate one or more symptoms or complications associated with human genetic diseases. In some cases, a method of treating a disease or disorder comprises disruption of the gene.

[0411] A gene that can be disrupted can be involved in attenuating TCR signaling, functional avidity, or immunity to cancer. In some cases, a gene to be disrupted is upregulated when a TCR is stimulated. A gene can be involved in inhibiting cellular expansion, functional avidity, or cytokine polyfunctionality. A gene can be involved in negatively regulating cellular cytokine production. For example, a gene can be involved in inhibiting production of effector cytokines, IFN-gamma and/or TNF for example. A gene can also be involved in inhibiting expression of supportive cytokines such as IL-2 after TCR stimulation.

[0412] A disease can be a neoplasia. Genes associated with neoplasia can be: PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4; Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF; HIF1a; HIF3a; Met; HRG; Bcl2; PPAR alpha; PPAR gamma; WT1 (Wilms Tumor); FGF Receptor Family members (5 members: 1, 2, 3, 4, 5); CDKN2a; APC; RB (retinoblastoma); MEN1; VHL; BRCA1; BRCA2; AR (Androgen Receptor); TSG101; IGF; IGF Receptor; Igf1 (4

variants); Igf2 (3 variants); Igf 1 Receptor; Igf 2 Receptor; Bax; Bcl2; caspases family (9 members: 1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc. A disease can be age-related macular degeneration. Genes associated with macular degeneration can be: Abcr; Ccl2; Cc2; cp (ceruloplasmin); Timp3; cathepsinD; Vldlr; Ccr2. A disease can be schizophrenia. Genes associated with schizophrenia can be: Neuregulin1 (Nrg1); Erb4 (receptor for Neuregulin); Complexin1 (Cplx1); Tph1 Tryptophan hydroxylase; Tph2 Tryptophan hydroxylase 2; Neurexin 1; GSK3; GSK3a; GSK3b. A disorder can be associated with a gene such as: 5-HTT (Slc6a4); COMT; DRD (Drd1a); SLC6A3; DAOA; DTNBP1; Dao (Dao1). A disease can be a trinucleotide repeat disorder. A trinucleotide repeat disorder can be associated with genes such as: HTT (Huntington's Dx); SBMA/SMAX1/AR (Kennedy's Dx); FXN/X25 (Friedrich's Ataxia); ATX3 (Machado- Joseph's Dx); ATXN1 and ATXN2 (spinocerebellar ataxias); DMPK (myotonic dystrophy); Atrophin-1 and Atn1 (DRPLA Dx); CBP (Creb-BP - global instability); VLDLR (Alzheimer's); Atxn7; Atxn10. A disease can be fragile X syndrome. Genes associated with fragile X syndrome can be: FMR2; FXR1; FXR2; mGLUR5. A disease can be secretase related with associated genes selected from: APH-1 (alpha and beta); Presenilin (Psen1); nicastrin, (Ncstn); PEN-2; Nos1; Parp1; Nat1; Nat2. A disease can be a prion related disorder with relevant genes being selected from: Prp. A disease can be ALS with relevant genes being: SOD1; ALS2; STEX; FUS; TARDBP; VEGF (VEGF-a; VEGF-b; VEGF-c). A disease can be drug addiction with relevant genes being; Prkce (alcohol); Drd2; Drd4; ABAT (alcohol); GRIA2; Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1 (alcohol). A disease can be autism with relevant genes being selected from: Mecp2; BZRAP1; MDGA2; Sema5A; Neurexin 1; Fragile X (FMR2 (AFF2); FXR1; FXR2; Mglur5). A disease can be Alzheimer's disease with relevant genes being selected from: E1; CHIP; UCH; UBB; Tau; LRP; PICALM; Clusterin; PS1; SORL1; CR1; Vldlr; Uba1; Uba3; CHIP28 (Aqp1, Aquaporin 1); Uchl1; Uchl3; APP. A disorder can be inflammation with relevant genes being selected from: IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL- 17b; IL-17c; IL-17d; IL-17f); II-23; Cx3cr1; ptpn22; TNFa; NOD2/CARD15 for IBD; IL-6; IL-12 (IL-12a; IL-12b); CTLA4; Cx3cl1. A disease can be Parkinson's disease with relevant genes being selected from: x-Synuclein; DJ-1; LRRK2; Parkin; PINK1. A disease can be a blood and coagulation disorders: Anemia (CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT); Bare lymphocyte syndrome (TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5), Bleeding disorders

(TBXA2R, P2RX1, P2X1); Factor H and factor H-like 1 (HF1, CFH, HUS); Factor V and factor VIII (MCFD2); Factor VII deficiency (F7); Factor X deficiency (F10); Factor XI deficiency (F11); Factor XII deficiency (F12, HAF); Factor XIIIA deficiency (F13A1, F13A); Factor XIIIB deficiency (F13B); Fanconi anemia (FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCB, FANCC, FACC, BRCA2, FANCD1, FANCD2, FANCD, FACD, FAD, FANCE, FACE, FANCF, XRCC9, FANCG, BRIP1, BACH1, FANCJ, PHF9, FANCL, FANCM, KIAA1596); Hemophagocytic lymphohistiocytosis disorders (PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3); Hemophilia A (F8, F8C, HEMA); Hemophilia B (F9, HEMB), Hemorrhagic disorders (PI, ATT, F5); Leukocyte deficiencies and disorders (ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH, CLE, EIF2B4); Sickle cell anemia (HBB); Thalassemia (HBA2, HBB, HBD, LCRB, HBA1).Cell dysregulation and oncology diseases and disorders: B-cell non-Hodgkin lymphoma (BCL7A, BCL7); Leukemia (TAL1 TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B, AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN). A disease can be an inflammation and/or an immune related diseases and disorders: AIDS (KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1); Autoimmune lymphoproliferative syndrome (TNFRSF6, APT1, FAS, CD95, ALPS1A); Combined immunodeficiency, (IL2RG, SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228), HIV susceptibility or infection (IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCKR5 (CCR5)); Immunodeficiencies (CD3E, CD3G, AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TACI); Inflammation (IL-10, IL-1 (IL-1a, IL-1b), IL-13, IL-17 (IL-17a (CTLA8), IL-17b, IL-17c, IL-17d, IL-17f), II-23, Cx3cr1, ptpn22, TNFa, NOD2/CARD15 for IBD, IL-6, IL-12 (IL-12a, IL-12b), CTLA4, Cx3cl1); Severe combined immunodeficiencies (SCIDs)(JAK3, JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4). A disease can be metabolic, liver, kidney and protein diseases

and disorders: Amyloid neuropathy (TTR, PALB); Amyloidosis (APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); Cirrhosis (KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988); Cystic fibrosis (CFTR, ABCC7, CF, MRP7); Glycogen storage diseases (SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); Hepatic adenoma, 142330 (TCF1, HNF1A, MODY3), Hepatic failure, early onset, and neurologic disorder (SCOD1, SCO1), Hepatic lipase deficiency (LIPC), Hepatoblastoma, cancer and carcinomas (CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; Medullary cystic kidney disease (UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); Phenylketonuria (PAH, PKU1, QDPR, DHPR, PTS); Polycystic kidney and hepatic disease (FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63). A disease can be muscular/skeletal diseases and disorders: Becker muscular dystrophy (DMD, BMD, MYF6), Duchenne Muscular Dystrophy (DMD, BMD); Emery-Dreifuss muscular dystrophy (LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A); Facioscapulohumeral muscular dystrophy (FSHMD1A, FSHD1A); Muscular dystrophy (FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPN1, SELN, RSMD1, PLEC1, PLTN, EBS1); Osteopetrosis (LRP5, BMND1, LRP7, LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1); Muscular atrophy (VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1). A disease can be neurological and neuronal diseases and disorders: ALS (SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b, VEGF-c); Alzheimer disease (APP, AAA, CVAP, AD1, APOE, AD2, PSEN2, AD4, STM2, APBB2, FE65L1, NOS3, PLAU, URK, ACE, DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3); Autism (Mecp2, BZRAP1, MDGA2, Sema5A, Neurexin 1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4, KIAA1260, AUTSX2); Fragile X Syndrome (FMR2, FXR1, FXR2, mGLUR5); Huntington's disease and disease like disorders (HD, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); Parkinson disease (NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA, NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1, PARK6,

UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN, PARK2, PDJ, DBH, NDUFV2); Rett syndrome (MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-Synuclein, DJ-1); Schizophrenia (Neuregulin1 (Nrg1), Erb4 (receptor for Neuregulin), Complexin1 (Cplx1), Tph1 Tryptophan hydroxylase, Tph2, Tryptophan hydroxylase 2, Neurexin 1, GSK3, GSK3a, GSK3b, 5-HTT (Slc6a4), COMT, DRD (Drd1a), SLC6A3, DAOA, DTNBP1, Dao (Dao1)); Secretase Related Disorders (APH-1 (alpha and beta), Presenilin (Psen1), nicastrin, (Ncstn), PEN-2, Nos1, Parp1, Nat1, Nat2); Trinucleotide Repeat Disorders (HTT (Huntington's Dx), SBMA/SMAX1/AR (Kennedy's Dx), FXN/X25 (Friedrich's Ataxia), ATX3 (Machado- Joseph's Dx), ATXN1 and ATXN2 (spinocerebellar ataxias), DMPK (myotonic dystrophy), Atrophin-1 and Atn1 (DRPLA Dx), CBP (Creb-BP - global instability), VLDLR (Alzheimer's), Atxn7, Atxn10). A disease can be an Ocular disease and/or disorder: Age-related macular degeneration (Abcr, Ccl2, Cc2, cp (ceruloplasmin), Timp3, cathepsinD, Vldlr, Ccr2); Cataract (CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1); Corneal clouding and dystrophy (APOA1, TGFBI, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD); Cornea plana congenital (KERA, CNA2); Glaucoma (MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A); Leber congenital amaurosis (CRB1, RP12, CRX, CORD2, CRD, RPGRIP1, LCA6, CORD9, RPE65, RP20, AIPL1, LCA4, GUCY2D, GUC2D, LCA1, CORD6, RDH12, LCA3); Macular dystrophy (ELOVL4, ADMD, STGD2, STGD3, RDS, RP7, PRPH2, PRPH, AVMD, AOFMD, VMD2).

[0413] In some cases a disease that can be treated with the disclosed editing system can be associated with a cellular condition. For example, genes associated with cellular performance may be disrupted with the disclosed editing system: PI3K/AKT Signaling: PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2; PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1; AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2; PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2; ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3; PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7; YWHAZ; ILK; TP53; RAF1; IKBKG; RELB;

DYRK1A; CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1; PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2; TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK; HSP90AA1; RPS6KB1. For example, ERK/MAPK Signaling: PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1; PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A; PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN; EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ; PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1; PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1; CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK. Glucocorticoid Receptor Signaling: RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1; MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I; PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2; MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1; MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13; RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1; PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3; MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP; CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1; ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1; STAT1; IL6; HSP90AA1. Axonal Guidance Signaling: PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12; IGF1; RAC1; RAP1A; E1F4E; PRKCZ; NRP1; NTRK2; ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2; CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11; PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1; GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B; AKT3; PRKCA. Ephrin Receptor Signaling: PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2; DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14; CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4, AKT1; JAK2; STAT3; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4; AKT3; SGK. Actin Cytoskeleton Signaling: ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1; PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6;

ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8; F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7; PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1; MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3; ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL; BRAF; VAV3; SGK. Huntington's Disease Signaling: PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2; MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2; PIK3CA; HDAC5; CREB1; PRKC1; HSPA5; REST; GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1; GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2; HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A; HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1; PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX; ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3. Apoptosis Signaling: PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1; BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB; CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8; BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF; RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2; CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2; BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK; CASP3; BIRC3; PARP1. B Cell Receptor Signaling: RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11; AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A; MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1; MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9; EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1; NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN; GSK3B; ATF4; AKT3; VAV3; RPS6KB1. Leukocyte Extravasation Signaling: ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA; RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11; MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12; PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB; MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK; MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2; CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK; CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9. Integrin Signaling: ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A; TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2; CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8; CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA; SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1; TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2; CRKL; BRAF; GSK3B; AKT3. Acute Phase Response

Signaling: IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11; AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14; PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1; TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN; AKT3; IL1R1; IL6. PTEN Signaling: ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11; MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA; CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1; MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR; RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2; AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1; NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2; GSK3B; AKT3; FOXO1; CASP3; RPS6KB1. p53 Signaling: PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A; BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2; PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1; PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9; CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A; HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1; SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN; SNAI2; GSK3B; BAX; AKT3. Aryl Hydrocarbon Receptor Signaling: HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1; NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1; SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1; MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1; SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF; CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1; CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1; HSP90AA1. Xenobiotic Metabolism Signaling: PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1; NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A; PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1; ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD; GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL; NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1; CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1; NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1; HSP90AA1. SAPK/JNK Signaling: PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1; GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA; FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1; GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2; PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1; CRKL; BRAF; SGK. PPAr/RXR Signaling: PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN; RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2; ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8; IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A; NCOA3; MAPK14; INSR;

RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1; TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1; ADIPOQ. NF-KB Signaling: IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ: TRAF6; TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2; MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2; KRAS; RELA; PIK3C2A; TRAF2; TLR4: PDGFRB; TNF; INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1; PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10; GSK3B; AKT3; TNFAIP3; IL1R1. Neuregulin Signaling: ERBB4; PRKCE; ITGAM; ITGA5: PTEN; PRKCZ; ELK1; MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI; CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS; PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2; ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3; EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL; AKT3; PRKCA; HSP90AA1; RPS6KB1. Wnt & Beta catenin Signaling: CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO; AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A; WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2: ILK; LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1; PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1; GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B; AKT3; SOX2. Insulin Receptor Signaling: PTEN; INS; EIF4E; PTPN1; PRKCZ; MAPK1; TSC1; PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3; MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1; SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1; GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK; RPS6KB1. IL-6 Signaling: HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11; IKBKB; FOS; NFKB2: MAP3K14; MAPK8; MAPK3; MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3; MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6. Hepatic Cholestasis: PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA; RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8; PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1; TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8; CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4; JUN; IL1R1; PRKCA; IL6. IGF-1 Signaling: IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2; PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8; IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A; YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3; FOXO1; SRF; CTGF; RPS6KB1. NRF2-mediated Oxidative Stress Response: PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1; NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8; PRKD1;

MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL; NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP; MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1; GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1. Hepatic Fibrosis/Hepatic Stellate Cell Activation: EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF; SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9; IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8; PDGFRA; NFKB1; TGFBR1; SMAD4; VEGFA; BAX; IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9. PPAR Signaling: EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB; NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3; NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2; PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA; MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1. Fc Epsilon RI Signaling: PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD; MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3; VAV3; PRKCA. G-Protein Coupled Receptor Signaling: PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB; PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB; PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1; IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK; PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3; PRKCA, Inositol Phosphate Metabolism: PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6; MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3; MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1; MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK. PDGF Signaling: EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC; PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2; PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC; JUN; CRKL; PRKCA; SRF; STAT1; SPHK2. VEGF Signaling: ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF; AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3; BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN; RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN; VEGFA; AKT3; FOXO1; PRKCA. Natural Killer Cell Signaling: PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11; KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6; PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1; PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA. Cell Cycle: G1/S Checkpoint Regulation: HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC; ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11; HDAC9;

CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1; E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1; GSK3B; RBL1; HDAC6. T Cell Receptor Signaling: RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS; NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; RELA, PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB, FYN; MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10; JUN; VAV3. Death Receptor Signaling: CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD; FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8; DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB; CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3; BIRC3. FGF Signaling: RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11; AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8; MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1; AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4; AKT3; PRKCA; HGF. GM-CSF Signaling: LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A; STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3; ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3; STAT1. Amyotrophic Lateral Sclerosis Signaling: BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2; PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1; PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1; APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3. JAK/Stat Signaling: PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B; PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A; PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3; STAT1. Nicotinate and Nicotinamide Metabolism: PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1; PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1; PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2; MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK. Chemokine Signaling: CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ; CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13; RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1; MAP2K2; MAP2K1; JUN; CCL2; PRKCA. IL-2 Signaling: ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS; STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2; JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3. Synaptic Long Term Depression: PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS; PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3; KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA; YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA. Estrogen Receptor Signaling: TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2; SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1; HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP; MAP2K2;

NCOA2; MAP2K1; PRKDC; ESR1; ESR2. Protein Ubiquitination Pathway: TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4; CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7; USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8; USP1; VHL; HSP90AA1; BIRC3. IL-10 Signaling: TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF; IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1; JUN; IL1R1; IL6. VDR/RXR Activation: PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1; NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD; RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1; LRP5; CEBPB; FOXO1; PRKCA. TGF-beta Signaling: EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1; FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2; SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2; MAP2K1; TGFBR1; SMAD4; JUN; SMAD5. Toll-like Receptor Signaling: IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK; NFKB1; TLR2; JUN. p38 MAPK Signaling: HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS; CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2; MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1; SRF; STAT1. Neurotrophin/TRK Signaling: NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A; RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; CDC42; JUN; ATF4. FXR/RXR Activation: INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8; APOB; MAPK10; PPARG; MTTP; MAPK9; PPARGC1A; TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1. Synaptic Long Term Potentiation: PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1; PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS; PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1; ATF4; PRKCA. Calcium Signaling: RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1; CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11; HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4; HDAC6. EGF Signaling: ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1; STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1. Hypoxia Signaling in the Cardiovascular System: EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT; HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM; VEGFA; JUN; ATF4; VHL; HSP90AA1. LPS/IL-1 Mediated Inhibition of RXR Function LXR/RXR Activation: IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1, MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2; TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1 FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA; NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1; SREBF1; IL1R1;

CCL2; IL6; MMP9. Amyloid Processing: PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2; CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1; PSEN1; CSNK1A1; GSK3B; AKT3; APP. IL-4 Signaling: AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1; PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1; FRAP1; AKT3; RPS6KB1. Cell Cycle: G2/M DNA Damage Checkpoint Regulation: EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC; CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A; PRKDC; ATM; SFN; CDKN2A. Nitric Oxide Signaling in the Cardiovascular System: KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3; CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1; VEGFA; AKT3; HSP90AA1. Purine Metabolism: NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4; PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C; NT5E; POLD1; NME1. cAMP-mediated Signaling: RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3; SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4. Mitochondrial Dysfunction Notch Signaling: SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9; PARK7; PSEN1; PARK2; APP; CASP3 HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2; PSEN1; NOTCH3; NOTCH1; DLL4. Endoplasmic Reticulum Stress Pathway: HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4; EIF2AK3; CASP3. Pyrimidine Metabolism: NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B; NT5E; POLD1; NME1. Parkinson's Signaling: UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7; PARK2; CASP3. Cardiac & Beta Adrenergic Signaling: GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC; PPP2R5C. Glycolysis/ Gluconeogenesis: HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1. Interferon Signaling: IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3. Sonic Hedgehog Signaling: ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRKIB. Glycerophospholipid Metabolism: PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2. Phospholipid Degradation: PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2. Tryptophan Metabolism: SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1. Lysine Degradation: SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C. Nucleotide Excision Repair Pathway: ERCC5; ERCC4; XPA; XPC; ERCC1. Starch and Sucrose Metabolism: UCHL1; HK2; GCK; GPI; HK1. Aminosugars Metabolism: NQO1; HK2; GCK; HK1. Arachidonic Acid Metabolism: PRDX6; GRN; YWHAZ; CYP1B1. Circadian Rhythm Signaling: CSNK1E; CREB1; ATF4; NR1D1. Coagulation System: BDKRB1; F2R; SERPINE1; F3. Dopamine Receptor Signaling: PPP2R1A; PPP2CA; PPP1CC; PPP2R5C. Glutathione Metabolism: IDH2; GSTP1; ANPEP; IDH1. Glycerolipid Metabolism: ALDH1A1; GPAM; SPHK1; SPHK2. Linoleic Acid Metabolism: PRDX6; GRN; YWHAZ; CYP1B1.

Methionine Metabolism: DNMT1; DNMT3B; AHCY; DNMT3A. Pyruvate Metabolism: GLO1; ALDH1A1; PKM2; LDHA. Arginine and Proline Metabolism: ALDH1A1; NOS3; NOS2A. Eicosanoid Signaling: PRDX6; GRN; YWHAZ. Fructose and Mannose Metabolism: HK2; GCK; HK1. Galactose Metabolism: HK2; GCK; HK1. Stilbene, Coumarine and Lignin Biosynthesis: PRDX6; PRDX1; TYR. Antigen Presentation Pathway: CALR; B2M. Biosynthesis of Steroids: NQO1; DHCR7. Butanoate Metabolism: ALDH1A1; NLGN1. Citrate Cycle: IDH2; IDH1. Fatty Acid Metabolism: ALDH1A1; CYP1B1. Glycerophospholipid Metabolism: PRDX6; CHKA. Histidine Metabolism: PRMT5; ALDH1A1. Inositol Metabolism: ERO1L; APEX1. Metabolism of Xenobiotics by Cytochrome p450: GSTP1; CYP1B1. Methane Metabolism: PRDX6; PRDX1. Phenylalanine Metabolism: PRDX6; PRDX1. Propanoate Metabolism: ALDH1A1; LDHA. Selenoamino Acid Metabolism: PRMT5; AHCY. Sphingolipid Metabolism: SPHK1; SPHK2. Aminophosphonate Metabolism: PRMT5. Androgen and Estrogen Metabolism: PRMT5. Ascorbate and Aldarate Metabolism: ALDH1A1. Bile Acid Biosynthesis: ALDH1A1. Cysteine Metabolism: LDHA. Fatty Acid Biosynthesis: FASN. Glutamate Receptor Signaling: GNB2L1. NRF2-mediated Oxidative Stress Response: PRDX1. Pentose Phosphate Pathway: GPI. Pentose and Glucuronate Interconversions: UCHL1. Retinol Metabolism: ALDH1A1. Riboflavin Metabolism: TYR. Tyrosine Metabolism: PRMT5, TYR. Ubiquinone Biosynthesis: PRMT5. Valine, Leucine and Isoleucine Degradation: ALDH1A1. Glycine, Serine and Threonine Metabolism: CHKA. Lysine Degradation: ALDH1A1. Pain/Taste: TRPM5; TRPA1. Pain: TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2; Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca; Prkacb; Prkar1a; Prkar2a. Mitochondrial Function: AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2. Developmental Neurology: BMP-4; Chordin (Chrd); Noggin (Nog); WNT (Wnt2; Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b; Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); beta-catenin; Dkk-1; Frizzled related proteins; Otx-2; Gbx2; FGF-8; Reelin; Dab1; unc-86 (Pou4fl or Brn3a); Numb; Reln.

[0414] In some cases, an editing system can be used to improve an immune cell performance. Examples of genes involved in cancer or tumor suppression may include ATM (ataxia telangiectasia mutated), ATR (ataxia telangiectasia and Rad3 related), EGFR (epidermal growth factor receptor), ERBB2 (v-erb-b2 erythroblastic leukemia viral oncogene homolog 2), ERBB3 (v-erb-b2 erythroblastic leukemia viral oncogene homolog 3), ERBB4 (v-erb-b2 erythroblastic leukemia viral oncogene homolog 4), Notch 1, Notch2, Notch 3, or Notch 4, for example. A gene

and protein associated with a secretase disorder may also be disrupted or introduced and can include PSENEN (presenilin enhancer 2 homolog (*C. elegans*)), CTSB (cathepsin B), PSEN1 (presenilin 1), APP (amyloid beta (A4) precursor protein). APH1B (anterior pharynx defective 1 homolog B (*C. elegans*)), PSEN2 (presenilin 2 (Alzheimer disease 4)), or BACE1 (beta-site APP-cleaving enzyme 1). It is contemplated that genetic homologues (*e.g.*, any mammalian version of the gene) of the genes within this applications are covered. For example, genes that can be targeted can further include CD27, CD40, CD122, OX40, GITR, CD137, CD28, ICOS, A2AR, B7-H3, B7-H4, BTLA, CTLA-4, IDO, KIR, LAG3, PD-1, TIM-3, VISTA, HPRT, CCR5, AAVS SITE (*e.g.* AAVS1, AAVS2, ETC.), PPP1R12C, TRAC, TCRB, or CISH. Therefore, it is contemplated that any one of the aforementioned gene that exhibits or exhibits about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity (at the nucleic acid or protein level) can be disrupted. It is also contemplated that any of the aforementioned genes that exhibits or exhibits about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity (at the nucleic acid or protein level) can be disrupted. Some genetic homologues are known in the art, however, in some cases, homologues are unknown. However, homologous genes between mammals can be found by comparing nucleic acid (DNA or RNA) sequences or protein sequences using publicly available databases such as NCBI BLAST. Also disclosed herein can be non-human gene equivalents of any one of the aforementioned genes. A non-human equivalent of any of the aforementioned genes can be disrupted with the gene editing system disclosed herein.

[0415] In some cases, a genome that can be disrupted or modified can be from an organism or subject that can be a eukaryote (including mammals including human) or a non-human eukaryote or a non-human animal or a non-human mammal. In some cases, an organism or subject can be a non-human animal, and may be an arthropod, for example, an insect, or may be a nematode. In some cases, an organism or subject can be a plant. In some cases, an organism or subject can be a mammal or a non-human mammal. A non-human mammal may be for example a rodent (preferably a mouse or a rat), an ungulate, or a primate. In some methods of the invention the organism or subject is algae, including microalgae, or is a fungus. In some cases, a subject can be a human. A human subject can be an adult or a pediatric subject. A pediatric subject can be under the age of 18. An adult subject can be about 18 or over 18 years of age. In some cases, a subject

can be a fetus or an embryo. In some cases, a genome that can be disrupted can be from a cell, tissue, or organ of an organism or subject. In some cases, a genome that can be disrupted may be from a stem cell. In some cases, a genome that can be disrupted may be from a germ cell.

[0416] A guide RNA can be introduced into a cell or embryo as an RNA molecule. For example, a RNA molecule can be transcribed *in vitro* and/or can be chemically synthesized. A guide RNA can then be introduced into a cell or embryo as an RNA molecule. A guide RNA can also be introduced into a cell or embryo in the form of a non-RNA nucleic acid molecule, *e.g.*, DNA molecule. For example, a DNA encoding a guide RNA can be operably linked to promoter control sequence for expression of the guide RNA in a cell or embryo of interest. A RNA coding sequence can be operably linked to a promoter sequence that is recognized by RNA polymerase III (Pol III).

[0417] A nucleic acid encoding a guide RNA or guide DNA can be linear. A nucleic acid encoding a guide RNA or guide DNA can also be circular. A nucleic acid encoding the guiding polynucleic acid can also be part of a vector. Some examples of vectors can include plasmid vectors, phagemids, cosmids, artificial/mini-chromosomes, transposons, and viral vectors. For example, a DNA encoding a RNA-guided endonuclease is present in a plasmid vector. Other non-limiting examples of suitable plasmid vectors include pUC, pBR322, pET, pBluescript, and variants thereof. Further, a vector can comprise additional expression control sequences (*e.g.*, enhancer sequences, Kozak sequences, polyadenylation sequences, transcriptional termination sequences, etc.), selectable marker sequences (*e.g.*, antibiotic resistance genes), origins of replication, and the like.

[0418] Suitable methods for introduction of the guiding polynucleic acid, protein, or guiding polynucleic acid: nuclease complex are known in the art and include, for example, electroporation; calcium phosphate precipitation; or PEI, PEG, DEAE, nanoparticle, or liposome mediated transformation. Other suitable transfection methods include direct micro-injection. In some cases, the guiding polynucleic acid and nuclease are introduced separately and the guiding polynucleic acid: nuclease complexes are formed in a cell. In other cases, the guiding polynucleic acid: nuclease complex can be formed and then introduced into a cell. In some cases, multiple, differentially labeled, guiding polynucleic acid: nuclease complexes, each directed to a different genomic targets are formed and then introduced into a cell. When both a nucleic acid guided nuclease and a guide polynucleic acid are introduced into a cell, each can be part of a separate molecule (*e.g.*, one vector containing fusion protein coding sequence and a second vector

containing guide polynucleic acid coding sequence) or both can be part of a same molecule (*e.g.*, one vector containing coding (and regulatory) sequence for both a fusion protein and the guiding polynucleic acid). In some cases, a nuclease can be pre-complexed with the guiding polynucleic acid. A complex can be a ribonucleoprotein (RNP) complex.

[0419] In some cases, a GUIDE-Seq analysis can be performed to determine the specificity of engineered guiding polynucleic acids. The general mechanism and protocol of GUIDE-Seq profiling of off-target cleavage by CRISPR system nucleases is discussed in Tsai, S. *et al.*, "GUIDE-Seq enables genome-wide profiling of off-target cleavage by CRISPR system nucleases," Nature, 33: 187-197 (2015).

[0420] The guiding polynucleic acid can be introduced at any functional concentration. For example, the guiding polynucleic acid can be introduced to a cell at 10micrograms. In other cases, the guiding polynucleic acid can be introduced from 0.5 micrograms to 100 micrograms. A gRNA can be introduced from 0.5, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 micrograms.

[0421] A sequence of a guiding polynucleic acid need not be 100% complementary to that of its target polynucleic acid to be specifically hybridizable or hybridizable. Moreover, a guiding polynucleic acid may hybridize over one or more segments such that intervening or adjacent segments are not involved in the hybridization event (*e.g.*, a loop structure or hairpin structure). For example, a polynucleotide can comprise 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more, 98% or more, 99% or more, 99.5%, or 100% sequence complementarity to a target region within the target nucleic acid sequence to which it will hybridize. For example, an antisense nucleic acid in which 18 of 20 nucleotides of the antisense compound are complementary to a target region, and would therefore specifically hybridize, would represent 90 percent complementarity. In this example, the remaining non-complementary nucleotides may be clustered or interspersed with complementary nucleotides and need not be contiguous to each other or to complementary nucleotides. Percent complementarity between particular stretches of nucleic acid sequences within nucleic acids can be determined using any convenient method. Exemplary methods include BLAST programs (basic local alignment search tools) and PowerBLAST programs (Altschul et al., J. Mol. Biol., 1990, 215, 403-410; Zhang and Madden, Genome Res., 1997, 7, 649-656) or by using the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group,

University Research Park, Madison Wis.), using default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489).

[0422] The guiding polynucleic acid can target a gene or portion thereof. In some cases, a cell that is modified can comprise one or more suppressed, disrupted, or knocked out genes and one or more transgenes, such as a receptor.

[0423] The target nucleic acid molecule can be DNA or RNA. The target nucleic acid can be double stranded or single stranded. The target nucleic acid can be double stranded DNA (dsDNA), single stranded DNA (ssDNA), double stranded RNA (dsRNA), or single stranded RNA (ssRNA). The Ago may be capable of cleaving 1, 2, 3, or 4 of dsDNA, ssDNA, dsRNA, or ssRNA.

[0424] Methods and compositions described herein can be used to target a gene from a mammal. A gene that can be targeted can be from any organ or tissue. A gene that can be targeted can be from skin, eyes, heart, liver, lung, kidney, reproductive tract, brain, to name a few. A gene that can be targeted can also be from a number of conditions and diseases

[0425] In some cases, a disruption can result in a reduction of copy number of genomic transcript of a disrupted gene or portion thereof. For example, a target gene that can be disrupted can have reduced transcript quantities compared to the same target gene in an undisrupted cell. A disruption can result in disruption results in less than 145 copies/μL, 140 copies/μL, 135 copies/μL, 130 copies/μL, 125 copies/μL, 120 copies/μL, 115 copies/μL, 110 copies/μL, 105 copies/μL, 100 copies/μL, 95 copies/μL, 190 copies/μL, 185 copies/μL, 80 copies/μL, 75 copies/μL, 70 copies/μL, 65 copies/μL, 60 copies/μL, 55 copies/μL, 50 copies/μL, 45 copies/μL, 40 copies/μL, 35 copies/μL, 30 copies/μL, 25 copies/μL, 20 copies/μL, 15 copies/μL, 10 copies/μL, 5 copies/μL, 1 copies/μL, or 0.05 copies/μL. In some cases, a disruption can result in less than 100 copies/μL.

[0426] One or more genes in a cell can be knocked out or disrupted using any method. For example, knocking out one or more genes can comprise deleting one or more genes from a genome of a cell. Knocking out can also comprise removing all or a part of a gene sequence from a cell. It is also contemplated that knocking out can comprise replacing all or a part of a gene in a genome of a cell with one or more nucleotides. Knocking out one or more genes can also comprise inserting a sequence in one or more genes thereby disrupting expression of the one or more genes. For example, inserting a sequence can generate a stop codon in the middle of one or more genes. Inserting a sequence can also shift the open reading frame of one or more genes.

[0427] An animal or cell may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more disrupted genomic sequences encoding a protein associated with a disease and zero, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more genomically integrated sequences encoding a protein associated with a disease.

X. Delivery and Cells

[0428] The Ago system, the fusion polypeptide, the polynucleotides encoding the same, and/or any transgene polynucleotides and compositions comprising the polypeptides and/or polynucleotides described herein can be delivered to one or more target cells by any suitable means. Accordingly, described herein are one or more cells that comprise the disclosed system and one or more cells that comprise the disclosed fusion polypeptide. The one or more cells can be ex vivo, in vivo, or in vitro cells. In some cases, the one or more cells are ex vivo cells. Similarly, the one or more cells can comprise an exogenous nucleic acid molecule that encodes the disclosed fusion polypeptide or the Ago. For example, described herein is a cell that comprises an exogenous nucleic acid molecule that encodes a disclosed Ago polypeptide. The Ago polypeptide can comprise an amino acid sequence having 70 % or more sequence identity with one of SEQ ID NOs: 1-10 or 134-136.

[0429] The cells can include but are not limited to eukaryotic and prokaryotic cells and/or cell lines. The cells can be engineered cells. The one or more cells can comprise or can be a mammalian cell. The cells can be from an animal selected from a group consisting of mice, rats, rabbits, sheep, cattle, horses, dogs, cats, and humans. The one or more cells can comprise or can be a human primary cell.

[0430] The primary cell can be taken directly from living tissue (i.e. biopsy material) and established for growth *in vitro*, that have undergone very few population doublings and are therefore more representative of the main functional components and characteristics of tissues from which they are derived from, in comparison to continuous tumorigenic or artificially immortalized cell lines.

[0431] The primary cell can be acquired from a variety of sources such as an organ, vasculature, buffy coat, whole blood, apheresis, plasma, bone marrow, tumor, cell-bank, cryopreservation bank, or a blood sample. The cell can be a stem cell. The cell can be a germ cell. The cells that can be edited with a genomic editing system comprising the Ago can be epithelial cells, fibroblast cells, neural cells, keratinocytes, hematopoietic cells, melanocytes, chondrocytes,

lymphocytes (B, NK, and T), macrophages, monocytes, mononuclear cells, cardiac muscle cells, other muscle cells, granulosa cells, cumulus cells, epidermal cells, endothelial cells, pancreatic islet cells, blood cells, blood precursor cells, bone cells, bone precursor cells, neuronal stem cells, primordial stem cells, hepatocytes, keratinocytes, umbilical vein endothelial cells, aortic endothelial cells, microvascular endothelial cells, fibroblasts, liver stellate cells, aortic smooth muscle cells, cardiac myocytes, neurons, Kupffer cells, smooth muscle cells, Schwann cells, and epithelial cells, erythrocytes, platelets, neutrophils, lymphocytes, monocytes, eosinophils, basophils, adipocytes, chondrocytes, pancreatic islet cells, thyroid cells, parathyroid cells, parotid cells, tumor cells, glial cells, astrocytes, red blood cells, white blood cells, macrophages, epithelial cells, somatic cells, pituitary cells, adrenal cells, hair cells, bladder cells, kidney cells, retinal cells, rod cells, cone cells, heart cells, pacemaker cells, spleen cells, antigen presenting cells, memory cells, T cells, B cells, plasma cells, muscle cells, ovarian cells, uterine cells, prostate cells, vaginal epithelial cells, sperm cells, testicular cells, germ cells, egg cells, leydig cells, peritubular cells, sertoli cells, lutein cells, cervical cells, endometrial cells, mammary cells, follicle cells, mucous cells, ciliated cells, nonkeratinized epithelial cells, keratinized epithelial cells, lung cells, goblet cells, columnar epithelial cells, dopamiergic cells, squamous epithelial cells, osteocytes, osteoblasts, osteoclasts, dopaminergic cells, embryonic stem cells, fibroblasts and fetal fibroblasts. Further, the one or more cells can be pancreatic islet cells and/or cell clusters or the like, including, but not limited to pancreatic α cells, pancreatic β cells, pancreatic δ cells, pancreatic F cells (*e.g.*, PP cells), or pancreatic ε cells. In one instance, the one or more cells can be pancreatic α cells. In another instance, the one or more cells can be pancreatic β cells.

[0432] A human primary cell can be an immune cell. An immune cell can be a T cell, B cell, NK cell, and/or TIL. Non-limiting examples of such cells or cell lines generated from such cells include COS, CHO (*e.g.*, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NSO, SP2/0-Ag14, HeLa, HEK293 (*e.g.*, HEK293-F, HEK293-H, HEK293-T), and perC6 cells as well as insect cells such as Spodopterafugiperda (Sf), or fungal cells such as Saccharomyces, Pichia and Schizosaccharomyces. In some cases, a cell line can be a CHO-K1, MDCK or HEK293 cell line. In some cases, suitable primary cells include peripheral blood mononuclear cells (PBMC), peripheral blood lymphocytes (PBL), and other blood cell subsets such as, but not limited to, T cell, a natural killer cell, a monocyte, a natural killer T cell, a monocyte-precursor cell, a

hematopoietic stem cell or a non-pluripotent stem cell. In some cases, the cell can be any immune cells including any T-cell such as tumor infiltrating cells (TILs), such as CD3+ T-cells, CD4+ T-cells, CD8+ T-cells, or any other type of T-cell. The T cell can also include memory T cells, memory stem T cells, or effector T cells. The T cells can also be selected from a bulk population, for example, selecting T cells from whole blood. The T cells can also be expanded from a bulk population. The T cells can also be skewed towards particular populations and phenotypes. For example, the T cells can be skewed to phenotypically comprise, CD45RO(-), CCR7(+), CD45RA(+), CD62L(+), CD27(+), CD28(+) and/or IL-7Rα(+). Suitable cells can be selected that comprise one of more markers selected from a list comprising: CD45RO(-), CCR7(+), CD45RA(+), CD62L(+), CD27(+), CD28(+) and/or IL-7Rα(+). Suitable cells also include stem cells such as, by way of example, embryonic stem cells, induced pluripotent stem cells, hematopoietic stem cells, neuronal stem cells and mesenchymal stem cells. Suitable cells can comprise any number of primary cells, such as human cells, non-human cells, and/or mouse cells. Suitable cells can be progenitor cells. Suitable cells can be derived from the subject to be treated (*e.g.*, subject). Suitable cells can be derived from a human donor. Suitable cells can be stem memory T_{SCM} cells comprised of CD45RO (-), CCR7(+), CD45RA (+), CD62L+ (L-selectin), CD27+, CD28+ and IL-7Rα+, stem memory cells can also express CD95, IL-2Rβ, CXCR3, and LFA-1, and show numerous functional attributes distinctive of stem memory cells. Suitable cells can be central memory T_{CM} cells comprising L-selectin and CCR7, central memory cells can secrete, for example, IL-2, but not IFNγ or IL-4. Suitable cells can also be effector memory T_{EM} cells comprising L-selectin or CCR7 and produce, for example, effector cytokines such as IFNγ and IL-4.

[0433] In some cases, modified cells can be a stem memory T_{SCM} cell comprised of CD45RO (-), CCR7(+), CD45RA (+), CD62L+ (L-selectin), CD27+, CD28+ and IL-7Rα+, stem memory cells can also express CD95, IL-2Rβ, CXCR3, and LFA-1, and show numerous functional attributes distinctive of stem memory cells. Engineered cells, such as Argonaute polypeptide modified cells can also be central memory T_{CM} cells comprising L-selectin and CCR7, where the central memory cells can secrete, for example, IL-2, but not IFNγ or IL-4. Engineered cells can also be effector memory T_{EM} cells comprising L-selectin or CCR7 and produce, for example, effector cytokines such as IFNγ and IL-4. In some cases a population of cells can be introduced to a subject. For example, a population of cells can be a combination of T cells and NK cells. In other cases, a population can be a combination of naïve cells and effector cells.

[0434] A method of attaining suitable cells, such as human primary cells, can comprise selecting cells. In some cases, a cell can comprise a marker that can be selected for the cell. For example, such marker can comprise GFP, a resistance gene, a cell surface marker, an endogenous tag. Cells can be selected using any endogenous marker. Suitable cells can be selected using any technology. Such technology can comprise flow cytometry and/or magnetic columns. The selected cells can then be infused into a subject. The selected cells can also be expanded to large numbers. The selected cells can be expanded prior to infusion.

[0435] In some cases, a suitable cell can be a recombinant cell. A recombinant cell can be an immortalized cell line. A cell line can be: CHO- K1 cells; HEK293 cells; Caco2 cells; U2-OS cells; NIH 3T3 cells; NSO cells; SP2 cells; CHO- S cells; DG44 cells; K-562 cells, U-937 cells; MRC5 cells; IMR90 cells; Jurkat cells; HepG2 cells; HeLa cells; HT-1080 cells; HCT-1 16 cells; Hu-h7 cells; Huvec cells; Molt 4 cells. All these cell lines can be modified by the method described herein to provide cell line models to produce, express, quantify, detect, study a gene or a protein of interest; these models can also be used to screen biologically active molecules of interest in research and production and various fields such as chemical, biofuels, therapeutics and agronomy as non-limiting examples.

[0436] The system as described herein can be delivered using vectors, for example containing sequences encoding one or more of the proteins or polypeptides. Accordingly, the system can comprise one or more vectors such as recombinant expression vectors. In some cases, the system as described herein can be delivered absent a viral vector. In some cases, the system as described herein can be delivered absent a viral vector, for example, when the system is greater than one kilobase, without affecting cellular viability. Transgenes encoding polynucleotides can be similarly delivered. Any vector systems can be used including, but not limited to, plasmid vectors, retroviral vectors, lentiviral vectors, adenovirus vectors, poxvirus vectors, herpesvirus vectors, split-intron retroviral vectors, adeno-associated virus vectors, any combination thereof, etc. Furthermore, any of these vectors can comprise one or more Ago or fragments thereof, Ago associated genes, transcription factors, nucleases, and/or transgenes. Thus, when one or more Ago or Ago associated molecules and/or transgenes are introduced into the cell, they can be carried on the same vector or on different vectors.

[0437] Split-intron based vectors, such as split-intron retroviral vectors, can be used for delivery. The methods and compositions of split-intron vectors are described in, e.g., Ismail et al, Journal of Virology, Mar. 2000, p. 2365–2371, and US20060281180, which are hereby incorporated by

reference in their entirety. Further, intron vectors like the ones described in Ding et al., Molecular Plant, vol 11 (4), p542, 2018, can be used for delivery. Ding et al., 2018, is hereby incorporated by reference in its entirety.

[0438] Conventional viral and non-viral based gene transfer methods can be used to introduce nucleic acids encoding engineered Ago, and Ago associated genes and/or transgenes in cells (*e.g.*, mammalian cells) and target tissues. In some examples, nucleic acids encoding Ago, and Ago associated genes, can be administered for *in vivo* or *ex vivo* immunotherapy uses. Non-viral vector delivery systems can include DNA plasmids, naked nucleic acid, lipid nanoparticles, and nucleic acid complexed with a delivery vehicle such as a liposome or poloxamer. Viral vector delivery systems can include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell.

[0439] Methods of non-viral delivery of nucleic acids include electroporation, lipofection, nucleofection, gold nanoparticle delivery, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycationic lipid: nucleic acid conjugates, naked DNA, mRNA, artificial virions, and agent-enhanced uptake of DNA. Sonoporation using, *e.g.*, the Sonitron 2000 system (Rich-Mar) can also be used for delivery of nucleic acids. Additional exemplary nucleic acid delivery systems include those provided by AMAXA® Biosystems (Cologne, Germany), Life Technologies (Frederick, Md.), MAXCYTE, Inc. (Rockville, Md.), BTX Molecular Delivery Systems (Holliston, Mass.) and Copernicus Therapeutics Inc. (see for example U.S. Pat. No. 6,008,336). Lipofection reagents are sold commercially (*e.g.*, TRANSFECTAM® and LIPOFECTIN®). Delivery can be to cells (ex vivo administration) or target tissues (*in vivo* administration). Additional methods of delivery include the use of packaging the nucleic acids to be delivered into EnGeneIC delivery vehicles (EDVs). These EDVs are specifically delivered to target tissues using bispecific antibodies where one arm of the antibody has specificity for the target tissue and the other has specificity for the EDV. The antibody brings the EDVs to the target cell surface and then the EDV is brought into the cell by endocytosis.

[0440] In some cases, the system, fusion polypeptide, and/or polynucleic acid is delivered using lipid nanoparticles. The methods and compositions of suitable lipid nanoparticles are described in, e.g., US20160375134, US20180147298, US20180200186, US20180263907, US20180092848, US20070087045, US9758795, US9687448, US9415109, US7858117, US7780983, US9504651, US 6586410, US 8969543, US9061063, and US9365610, which are hereby incorporated by reference in their entirety. The lipid nanoparticles can comprise a cationic

lipid, or a pharmaceutically acceptable salt thereof. The lipid nanoparticles can comprise a steroid, a neutral lipid, a polyethyleneglycol-containing lipid (PEGylated lipid), a phospholipid, or any combination thereof. The amount of the cationic lipid component can be from about 10 mol% to about 90 mol% of the overall lipid content of the formulation. In some cases, the cationic lipid component is from about 50 mol% to about 85 mol% of the overall lipids in the lipid nanoparticle. The amount of the steroid can be from about 10 mol% to about 50 mol% of the overall lipid in the lipid particle formulation. In some cases, the steroid is present in the lipid particles in an amount of from about 20 mol% to about 45 mol% of the total lipid. In some cases, the steroid is cholesterol or a derivative thereof. The amount of the phospholipid can be from about 1 mol% to about 20 mol% of the overall lipids in the lipid particle formulation. In some cases, from about 2 mol% to about 15 mol% of the total lipids are phospholipids.

[0441] Vectors including viral and non-viral vectors containing nucleic acids encoding engineered Ago, and Ago associated genes can also be administered directly to an organism for transduction of cells *in vivo*. Alternatively, naked DNA or mRNA can be administered. Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells including, but not limited to, injection, infusion, topical application and electroporation. More than one route can be used to administer a particular composition. Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition.

[0442] Vectors can be delivered *in vivo* by administration to an individual subject, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion) or topical application, as described below. Alternatively, vectors can be delivered to cells *ex vivo*, such as cells explanted from an individual subject (*e.g.*, lymphocytes, T cells, bone marrow aspirates, tissue biopsy), followed by reimplantation of the cells into a subject, usually after selection for cells which have incorporated the vector. Prior to or after selection, the cells can be expanded.

[0443] A cell can be transfected with a mutant or chimeric adeno-associated viral vector encoding one or more components of the editing system comprising the Ago, Ago fragment, the fusion polypeptide, the polynucleic acid, and/or the Ago associated genes. An AAV vector concentration can be from 0.5 nanograms to 50 micrograms. In some cases, the amount of nucleic acid (*e.g.*, ssDNA, dsDNA, RNA) that can be introduced into the cell by electroporation

can be varied to optimize transfection efficiency and/or cell viability. In some cases, less than about 100 picograms of nucleic acid can be added to each cell sample (*e.g.,* one or more cells being electroporated). In some cases, at least about 100 picograms, at least about 200 picograms, at least about 300 picograms, at least about 400 picograms, at least about 500 picograms, at least about 600 picograms, at least about 700 picograms, at least about 800 picograms, at least about 900 picograms, at least about 1 microgram, at least about 1.5 micrograms, at least about 2 micrograms, at least about 2.5 micrograms, at least about 3 micrograms, at least about 3.5 micrograms, at least about 4 micrograms, at least about 4.5 micrograms, at least about 5 micrograms, at least about 5.5 micrograms, at least about 6 micrograms, at least about 6.5 micrograms, at least about 7 micrograms, at least about 7.5 micrograms, at least about 8 micrograms, at least about 8.5 micrograms, at least about 9 micrograms, at least about 9.5 micrograms, at least about 10 micrograms, at least about 11 micrograms, at least about 12 micrograms, at least about 13 micrograms, at least about 14 micrograms, at least about 15 micrograms, at least about 20 micrograms, at least about 25 micrograms, at least about 30 micrograms, at least about 35 micrograms, at least about 40 micrograms, at least about 45 micrograms, or at least about 50 micrograms, of nucleic acid can be added to each cell sample (*e.g.*, one or more cells being electroporated). For example, 1 microgram of dsDNA can be added to each cell sample for electroporation. In some cases, the amount of nucleic acid (*e.g.*, dsDNA) required for optimal transfection efficiency and/or cell viability can be specific to the cell type. In some cases, the amount of nucleic acid (*e.g.*, dsDNA) used for each sample can directly correspond to the transfection efficiency and/or cell viability.

[0444] The transfection efficiency of cells with any of the nucleic acid delivery platforms described herein, for example, nucleofection or electroporation, can be or can be about 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more than 99.9%.

[0445] Vectors, plasmids, and genomic editing systems described herein can be delivered by any suitable method, including transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion. The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, *e.g.*, Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY. Electroporation using, for example, the Neon®

Transfection System (ThermoFisher Scientific) or the AMAXA® Nucleofector (AMAXA® Biosystems) can also be used for delivery of nucleic acids into a cell. Electroporation parameters can be adjusted to optimize transfection efficiency and/or cell viability. Electroporation devices can have multiple electrical wave form pulse settings such as exponential decay, time constant and square wave. Every cell type has a unique optimal Field Strength (E) that is dependent on the pulse parameters applied (*e.g.*, voltage, capacitance and resistance). Application of optimal field strength causes electropermeabilization through induction of transmembrane voltage, which allows nucleic acids to pass through the cell membrane. In some cases, the electroporation pulse voltage, the electroporation pulse width, number of pulses, cell density, and tip type can be adjusted to optimize transfection efficiency and/or cell viability.

[0446] In some cases, electroporation pulse voltage can be varied to optimize transfection efficiency and/or cell viability. In some cases, the electroporation voltage can be less than about 500 volts. In some cases, the electroporation voltage can be at least about 500 volts, at least about 600 volts, at least about 700 volts, at least about 800 volts, at least about 900 volts, at least about 1000 volts, at least about 1100 volts, at least about 1200 volts, at least about 1300 volts, at least about 1400 volts, at least about 1500 volts, at least about 1600 volts, at least about 1700 volts, at least about 1800 volts, at least about 1900 volts, at least about 2000 volts, at least about 2100 volts, at least about 2200 volts, at least about 2300 volts, at least about 2400 volts, at least about 2500 volts, at least about 2600 volts, at least about 2700 volts, at least about 2800 volts, at least about 2900 volts, or at least about 3000 volts. In some cases, the electroporation pulse voltage required for optimal transfection efficiency and/or cell viability can be specific to the cell type. For example, an electroporation voltage of 1900 volts can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for macrophage cells. In another example, an electroporation voltage of about 1350 volts can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for Jurkat cells or primary human cells such as T cells. In some cases, a range of electroporation voltages can be optimal for a given cell type. For example, an electroporation voltage between about 1000 volts and about 1300 volts can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for human 578T cells.

[0447] In some cases, electroporation pulse width can be varied to optimize transfection efficiency and/or cell viability. In some cases, the electroporation pulse width can be less than about 5 milliseconds. In some cases, the electroporation width can be at least about 5 milliseconds, at least about 6 milliseconds, at least about 7 milliseconds, at least about 8

milliseconds, at least about 9 milliseconds, at least about 10 milliseconds, at least about 11 milliseconds, at least about 12 milliseconds, at least about 13 milliseconds, at least about 14 milliseconds, at least about 15 milliseconds, at least about 16 milliseconds, at least about 17 milliseconds, at least about 18 milliseconds, at least about 19 milliseconds, at least about 20 milliseconds, at least about 21 milliseconds, at least about 22 milliseconds, at least about 23 milliseconds, at least about 24 milliseconds, at least about 25 milliseconds, at least about 26 milliseconds, at least about 27 milliseconds, at least about 28 milliseconds, at least about 29 milliseconds, at least about 30 milliseconds, at least about 31 milliseconds, at least about 32 milliseconds, at least about 33 milliseconds, at least about 34 milliseconds, at least about 35 milliseconds, at least about 36 milliseconds, at least about 37 milliseconds, at least about 38 milliseconds, at least about 39 milliseconds, at least about 40 milliseconds, at least about 41 milliseconds, at least about 42 milliseconds, at least about 43 milliseconds, at least about 44 milliseconds, at least about 45 milliseconds, at least about 46 milliseconds, at least about 47 milliseconds, at least about 48 milliseconds, at least about 49 milliseconds, or at least about 50 milliseconds. In some cases, the electroporation pulse width required for optimal transfection efficiency and/or cell viability can be specific to the cell type. For example, an electroporation pulse width of 30 milliseconds can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for macrophage cells. In another example, an electroporation width of about 10 milliseconds can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for Jurkat cells. In some cases, a range of electroporation widths can be optimal for a given cell type. For example, an electroporation width between about 20 milliseconds and about 30 milliseconds can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for human 578T cells.

[0448] In some cases, the number of electroporation pulses can be varied to optimize transfection efficiency and/or cell viability. In some cases, electroporation can comprise a single pulse. In some cases, electroporation can comprise more than one pulse. In some cases, electroporation can comprise 2 pulses, 3 pulses, 4 pulses, 5 pulses 6 pulses, 7 pulses, 8 pulses, 9 pulses, or 10 or more pulses. In some cases, the number of electroporation pulses required for optimal transfection efficiency and/or cell viability can be specific to the cell type. For example, electroporation with a single pulse can be optimal (*e.g.,* provide the highest viability and/or transfection efficiency) for macrophage cells. In another example, electroporation with a 3 pulses can be optimal (*e.g.,* provide the highest viability and/or transfection efficiency) for primary

cells. In some cases, a range of electroporation widths can be optimal for a given cell type. For example, electroporation with between about 1 to about 3 pulses can be optimal (*e.g.,* provide the highest viability and/or transfection efficiency) for human cells.

[0449] In some cases, the starting cell density for electroporation can be varied to optimize transfection efficiency and/or cell viability. In some cases, the starting cell density for electroporation can be less than about $1x10^5$ cells. In some cases, the starting cell density for electroporation can be at least about $1x10^5$ cells, at least about $2x10^5$ cells, at least about $3x10^5$ cells, at least about $4x10^5$ cells, at least about $5x10^5$ cells, at least about $6x10^5$ cells, at least about $7x10^5$ cells, at least about $8x10^5$ cells, at least about $9x10^5$ cells, at least about $1x10^6$ cells, at least about $1.5x10^6$ cells, at least about $2x10^6$ cells, at least about $2.5x10^6$ cells, at least about $3x10^6$ cells, at least about $3.5x10^6$ cells, at least about $4x10^6$ cells, at least about $4.5x10^6$ cells, at least about $5x10^6$ cells, at least about $5.5x10^6$ cells, at least about $6x10^6$ cells, at least about $6.5x10^6$ cells, at least about $7x10^6$ cells, at least about $7.5x10^6$ cells, at least about $8x10^6$ cells, at least about $8.5x10^6$ cells, at least about $9x10^6$ cells, at least about $9.5x10^6$ cells, at least about $1x10^7$ cells, at least about $1.2x10^7$ cells, at least about $1.4x10^7$ cells, at least about $1.6x10^7$ cells, at least about $1.8x10^7$ cells, at least about $2x10^7$ cells, at least about $2.2x10^7$ cells, at least about $2.4x10^7$ cells, at least about $2.6x10^7$ cells, at least about $2.8x10^7$ cells, at least about $3x10^7$ cells, at least about $3.2x10^7$ cells, at least about $3.4x10^7$ cells, at least about $3.6x10^7$ cells, at least about $3.8x10^7$ cells, at least about $4x10^7$ cells, at least about $4.2x10^7$ cells, at least about $4.4x10^7$ cells, at least about $4.6x10^7$ cells, at least about $4.8x10^7$ cells, or at least about $5x10^7$ cells. In some cases, the starting cell density for electroporation required for optimal transfection efficiency and/or cell viability can be specific to the cell type. For example, a starting cell density for electroporation of $1.5x10^6$ cells can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for macrophage cells. In another example, a starting cell density for electroporation of $5x10^6$ cells can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for human cells. In some cases, a range of starting cell densities for electroporation can be optimal for a given cell type. For example, a starting cell density for electroporation between of $5.6x10^6$ and $5 x10^7$ cells can optimal (*e.g.*, provide the highest viability and/or transfection efficiency) for human cells such as T cells.

[0450] In some cases, the guiding polynucleic acid and the Ago can be introduced into cells as a complex. The complex can comprise a DNA and the Ago or it can comprise an RNA and the Ago. The complex can be a ribonuclear protein complex (RNP). Introduction of an RNP

complex can be timed. In some cases, a cell can be synchronized with other cells at G1, S, and/or M phases of the cell cycle prior to introduction of a guiding polynucleic acid and the Ago. In some cases, an RNP complex can be delivered at a cell phase such that HDR, MMEJ, or NHEJ can be enhanced. In some cases an RNP complex can facilitate homology directed repair.

[0451] Non-homologous end joining (NHEJ) and Homology-directed repair (HDR) can be quantified using a variety of methods. In some cases, a percent of NHEJ, HDR, or a combination of both can be determined by co-delivering the gene editing molecules, for example a guiding polynucleic acid and an RNase H like domain containing polypeptide, with a donor DNA template that encodes a promoterless GFP into cells. After about 72 hrs, flow cytometry can be performed to quantify the total cell number (N_{Total}), GFP-positive cell number (N_{GFP+}), and GFP-negative cell number (N_{GFP-}). Among the GFP negative cells, next-generation sequencing can be performed to identify cells without mutations (N_{GFP-}^{0}), and with mutations (N_{GFP-}^{1}). HDR efficiency can be calculated as N_{GFP+}/N_{Total} X 100%, and NHEJ efficiency will be calculated as N_{GFP-}^{1}/N_{Total} X 100%.

[0452] In some cases, activity of a DNA editing system may be assayed using a cell expressing a reporter protein or containing a reporter gene. For example, a reporter protein may be engineered to contain an obstruction, such as a stop codon, a frameshift mutation, a spacer, a linker, or a transcriptional terminator; the DNA editing system may then be used to remove the obstruction and the resultant functional reporter protein may be detected. In some cases, the obstruction may be designed such that a specific sequence modification is required to restore functionality of the reporter protein. In other cases, the obstruction may be designed such that any insertion or deletion which results in a frame shift of one or two bases may be sufficient to restore functionality of the reporter protein. Examples of reporter proteins include colorimetric enzymes, metabolic enzymes, fluorescent proteins, enzymes and transporters associated with antibiotic resistance, and luminescent enzymes. Examples of such reporter proteins include β-galactosidase, Chloramphenicol acetyltransferase, Green fluorescent protein, Red fluorescent protein, luciferase, and renilla. Different detection methods may be used for different reporter proteins. For example, the reporter protein may affect cell viability, cell growth, fluorescence, luminescence, or expression of a detectable product. In some cases, the reporter protein may be detected using a colorimetric assay. In some cases, the reporter protein may be a fluorescent protein, and DNA editing may be assayed by measuring the degree of fluorescence in treated cells, or the number of treated cells with at least a threshold level of fluorescence. In some cases,

transcript levels of a reporter gene may be assessed. In other cases, a reporter gene may be assessed by sequencing. In some cases, an assay for measuring DNA editing may use a split fluorescence protein system, such as the self-complementing split $GFP_{1-10/11}$ systems, in which two fragments (G_{1-10} and G_{11}) of the GFP protein which can associate by themselves to form a functional GFP signal are linked using a frameshifting linker. Insertions or deletions within the frameshifting linker can restore the frame of the G_{11} fragment allowing the two fragments to form a functional GFP signal. In some cases, the Ago polypeptides as described herein may result in at least about 1%, 1.1%, 1.2%, 1.3%, 1.4%, 1.5%, 1.6%, 1.7%, 1.8%, 1.9%, 2%, 2.5%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, 98%, or 99% of cells exhibiting restored activity of a reporter protein. In some cases, the Ago polypeptides as described herein may result in at least about 1% to 99%, 1% to 10%, 1% to 5%, 1% to 2%, 5% to 50%, 10% to 80%, 10% to 50%, 30% to 70%, or 50% to 80% of cells exhibiting restored activity of a reporter protein. In some cases, Ago polypeptides as described herein may result in at least about a 1.5 fold, 2 fold, 3 fold, 4 fold, 5 fold, 6 fold, 7 fold, 8 fold, 9 fold, 10 fold, 15 fold, 20 fold, 25 fold, 30 fold, 40 fold, 50 fold, 60 fold, 70 fold, 80 fold, 90 fold, or 100 fold increase in the percentage of cells with restored activity of a reporter as compared to baseline. In some cases, the Ago polypeptides as described herein may result in at least about a 1.2 fold to 10 fold, 1.5 fold to 10 fold, 2 fold to 10 fold, 2 fold to 5 fold, 2 fold to 20 fold, 3 fold to 5 fold, 4 fold to 10 fold, 5 fold to 20 fold, 10 fold to 100 fold, 10 fold to 50 fold or 1.2 fold to 100 fold increase in the percentage of cells with restored activity of a reporter as compared to baseline.

[0453] The percent occurrence of a genomic break repair utilizing HDR over NHEJ or MMEJ can be or can be about 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more than 99.9% of cells that are contacted with a genomic editing system comprising the Ago or Ago fragment. The percent occurrence of a genomic break repair utilizing NHEJ over HDR or MMEJ can be or can be about 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more than 99.9% of cells that are contacted with a genomic editing system comprising the Ago or Ago fragment. The percent occurrence of a genomic break repair utilizing MMEJ over HDR or NHEJ can be or can be about 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, or more

than 99.9% of cells that are contacted with a genomic editing system comprising the Ago or Ago fragment.

[0454] Integration of an exogenous polynucleic acid can be measured using any technique. For example, integration can be measured by flow cytometry, surveyor nuclease assay, tracking of indels by decomposition (TIDE), junction PCR, or any combination thereof. In other cases, transgene integration can be measured by PCR. A TIDE analysis can also be performed on engineered cells. *Ex vivo* cell transfection can also be used for diagnostics, research, or for gene therapy (*e.g.*, via re-infusion of the transfected cells into the host organism). In some cases, cells are isolated from the subject organism, transfected with a nucleic acid (*e.g.*, gene or cDNA), and re-infused back into the subject organism (*e.g.*, subject).

[0455] The amount of the Ago or Ago fragment polypeptide-containing modified cells that can be necessary to be therapeutically effective in a subject can vary depending on the viability of the cells, and the efficiency with which the cells have been genetically modified (*e.g.*, the efficiency with which a transgene has been integrated into one or more cells). In some cases, the product (*e.g.*, multiplication) of the viability of cells post genetic modification and the efficiency of integration of a transgene can correspond to the therapeutic aliquot of cells available for administration to a subject. In some cases, an increase in the viability of cells post genetic modification can correspond to a decrease in the amount of cells that are necessary for administration to be therapeutically effective in a subject. In some cases, an increase in the efficiency with which a transgene has been integrated into one or more cells can correspond to a decrease in the amount of cells that are necessary for administration to be therapeutically effective in a subject. In some cases, determining an amount of cells that are necessary to be therapeutically effective can comprise determining a function corresponding to a change in the viability of cells over time. In some cases, determining an amount of cells that are necessary to be therapeutically effective can comprise determining a function corresponding to a change in the efficiency with which a transgene can be integrated into one or more cells with respect to time dependent variables (*e.g.*, cell culture time, electroporation time, cell stimulation time).

[0456] As described herein, viral particles, such as AAV, can be used to deliver a viral vector comprising a gene of interest or a transgene, such as the polynucleic acid described herein, into a cell *ex vivo* or *in vivo*. In some embodiments, a mutated or chimeric adeno-associated viral vector as disclosed herein can be measured as pfu (plaque forming units). In some cases, the pfu of recombinant virus or mutated or chimeric adeno-associated viral vector of the compositions and

methods of the disclosure can be about 10^8 to about 5×10^{10} pfu. In some cases, recombinant viruses of this disclosure are at least about 1×10^8, 2×10^8, 3×10^8, 4×10^8, 5×10^8, 6×10^8, 7×10^8, 8×10^8, 9×10^8, 1×10^9, 2×10^9, 3×10^9, 4×10^9, 5×10^9, 6×10^9, 7×10^9, 8×10^9, 9×10^9, 1×10^{10}, 2×10^{10}, 3×10^{10}, 4×10^{10}, and 5×10^{10} pfu. In some cases, recombinant viruses of this disclosure are at most about 1×10^8, 2×10^8, 3×10^8, 4×10^8, 5×10^8, 6×10^8, 7×10^8, 8×10^8, 9×10^8, 1×10^9, 2×10^9, 3×10^9, 4×10^9, 5×10^9, 6×10^9, 7×10^9, 8×10^9, 9×10^9, 1×10^{10}, 2×10^{10}, 3×10^{10}, 4×10^{10}, and 5×10^{10} pfu. In some aspects, a mutated or chimeric adeno-associated viral vector of the disclosure can be measured as vector genomes. In some cases, recombinant viruses of this disclosure are 1×10^{10} to 3×10^{12} vector genomes, or 1×10^9 to 3×10^{13} vector genomes, or 1×10^8 to 3×10^{14} vector genomes, or at least about 1×10^1, 1×10^2, 1×10^3, 1×10^4, 1×10^5, 1×10^6, 1×10^7, 1×10^8, 1×10^9, 1×10^{10}, 1×10^{11}, 1×10^{12}, 1×10^{13}, 1×10^{14}, 1×10^{15}, 1×10^{16}, 1×10^{17}, and 1×10^{18} vector genomes, or are 1×10^8 to 3×10^{14} vector genomes, or are at most about 1×10^1, 1×10^2, 1×10^3, 1×10^4, 1×10^5, 1×10^6, 1×10^7, 1×10^8, 1×10^9, 1×10^{10}, 1×10^{11}, 1×10^{12}, 1×10^{13}, 1×10^{14}, 1×10^{15}, 1×10^{16}, 1×10^{17}, and 1×10^{18} vector genomes.

[0457] In some cases, a mutated or chimeric adeno-associated viral vector of the disclosure can be measured using multiplicity of infection (MOI). In some cases, MOI can refer to the ratio, or multiple of vector or viral genomes to the cells to which the nucleic can be delivered. In some cases, the MOI can be 1×10^6 GC/mL. In some cases, the MOI can be 1×10^5 GC/mL to 1×10^7 GC/mL. In some cases, the MOI can be 1×10^4 GC/mL to 1×10^8 GC/mL. In some cases, recombinant viruses of the disclosure are at least about 1×10^1 GC/mL, 1×10^2 GC/mL, 1×10^3 GC/mL, 1×10^4 GC/mL, 1×10^5 GC/mL, 1×10^6 GC/mL, 1×10^7 GC/mL, 1×10^8 GC/mL, 1×10^9 GC/mL, 1×10^{10} GC/mL, 1×10^{11} GC/mL, 1×10^{12} GC/mL, 1×10^{13} GC/mL, 1×10^{14} GC/mL, 1×10^{15} GC/mL, 1×10^{16} GC/mL, 1×10^{17} GC/mL, and 1×10^{18} GC/mL MOI. In some cases, a mutated or chimeric adeno-associated viruses of this disclosure are from about 1×10^8 GC/mL to about 3×10^{14} GC/mL MOI, or are at most about 1×10^1 GC/mL, 1×10^2 GC/mL, 1×10^3 GC/mL, 1×10^4 GC/mL, 1×10^5 GC/mL, 1×10^6 GC/mL, 1×10^7 GC/mL, 1×10^8 GC/mL, 1×10^9 GC/mL, 1×10^{10} GC/mL, 1×10^{11} GC/mL, 1×10^{12} GC/mL, 1×10^{13} GC/mL, 1×10^{14} GC/mL, 1×10^{15} GC/mL, 1×10^{16} GC/mL, 1×10^{17} GC/mL, and 1×10^{18} GC/mL MOI.

[0458] In some aspects, a non-viral vector or nucleic acid can be delivered without the use of a mutated or chimeric adeno-associated viral vector and can be measured according to the quantity of nucleic acid. Generally, any suitable amount of nucleic acid can be used with the compositions and methods of this disclosure. In some cases, nucleic acid can be at least about 1 pg, 10 pg, 100

pg, 1 pg, 10 pg, 100 pg, 200 pg, 300 pg, 400 pg, 500 pg, 600 pg, 700 pg, 800 pg, 900 pg, 1 μg, 10 μg, 100 μg, 200 μg, 300 μg, 400 μg, 500 μg, 600 μg, 700 μg, 800 μg, 900 μg, 1 ng, 10 ng, 100 ng, 200 ng, 300 ng, 400 ng, 500 ng, 600 ng, 700 ng, 800 ng, 900 ng, 1 mg, 10 mg, 100 mg, 200 mg, 300 mg, 400 mg, 500 mg, 600 mg, 700 mg, 800 mg, 900 mg, 1 g, 2 g, 3 g, 4 g, or 5 g. In some cases, nucleic acid can be at most about 1 pg, 10 pg, 100 pg, 1 pg, 10 pg, 100 pg, 200 pg, 300 pg, 400 pg, 500 pg, 600 pg, 700 pg, 800 pg, 900 pg, 1 μg, 10 μg, 100 μg, 200 μg, 300 μg, 400 μg, 500 μg, 600 μg, 700 μg, 800 μg, 900 μg, 1 ng, 10 ng, 100 ng, 200 ng, 300 ng, 400 ng, 500 ng, 600 ng, 700 ng, 800 ng, 900 ng, 1 mg, 10 mg, 100 mg, 200 mg, 300 mg, 400 mg, 500 mg, 600 mg, 700 mg, 800 mg, 900 mg, 1 g, 2 g, 3 g, 4 g, or 5 g.

[0459] Cells (*e.g.*, engineered cells or engineered primary cells) before, after, and/or during transplantation can be functional. For example, transplanted cells can be functional for at least or at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 6, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, or 100 days after transplantation. Transplanted cells can be functional for at least or at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months after transplantation. Transplanted cells can be functional for at least or at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, or 30 years after transplantation. In some cases, transplanted cells can be functional for up to the lifetime of a recipient.

[0460] Further, transplanted cells can function at 100% of its normal intended operation. Transplanted cells can also function 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% of its normal intended operation.

[0461] Transplanted cells can also function over 100% of its normal intended operation. For example, transplanted cells can function 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000 or more % of its normal intended operation.

[0462] One or more cytokines can be introduced with cells of the invention. Cytokines can be utilized to boost cytotoxic T lymphocytes (including adoptively transferred tumor-specific cytotoxic T lymphocytes) to expand within a tumor microenvironment. In some cases, IL-2 can be used to facilitate expansion of the cells described herein. Cytokines such as IL-15 can also be employed. Other relevant cytokines in the field of immunotherapy can also be utilized, such as IL-2, IL-7, IL-12, IL-15, IL-21, or any combination thereof.

[0463] In some cases, IL-2 can be administered beginning within 24 hours of cell infusion and continuing for up to about 4 days (maximum 12 doses). In some cases, IL-2 can be administered for up to about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 days after an initial administration. Doses of IL-2 can be administered every eight hours. In some cases, IL-2 can be administered from about every 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 hours after an initial administration. In some cases, IL-2 dosing can be stopped if toxicities are detected. In some cases, doses can be delayed or stopped if subjects reach Grade 3 or 4 toxicity due to aldesleukin except for the reversible Grade 3 toxicities common to Aldesleukin such as diarrhea, nausea, vomiting, hypotension, skin changes, anorexia, mucositis, dysphagia, or constitutional symptoms and laboratory changes. In some cases, if these toxicities can be easily reversed within 24 hours by supportive measures, then additional doses can be given. In addition, dosing can be held or stopped at the discretion of a treating physician.

XI. Pharmaceutical Compositions

[0464] The Ago systems, polypeptides, and polynucleic acid described throughout can be formulated into a pharmaceutical composition. The pharmaceutical composition can comprise the Ago polypeptide, the Ago system, the fusion polypeptide, the polynucleic acid encoding the same, or any combination thereof. The pharmaceutical composition can further comprise a pharmaceutically acceptable excipient, diluent, carrier, or a combination thereof. A pharmaceutically acceptable excipient, carrier, or diluent can refer to an excipient, carrier or diluent that can be administered to a subject, together with an agent, and which does not destroy the pharmacological activity thereof and is nontoxic when administered in doses sufficient to deliver a therapeutic amount of the agent.

[0465] The pharmaceutical composition can be in a unit dosage form. The pharmaceutical composition can be administered in both single and multiple dosages. In some cases, for example, in the compositions, formulations and methods of treatment, the unit dosage of the composition or formulation administered can be 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 mg. In some cases, the total amount of the composition or formulation administered can be 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90, or 100 g.

[0466] The pharmaceutical composition can be in the form of tablets, capsules, lozenges, troches, hand candies, powders, sprays, aqueous suspensions, injectable solutions, elixirs, syrups, and the like. In some cases, the pharmaceutical composition is in a form of parenteral administration formulation. For example, the pharmaceutical composition can be in a form of intravenous, subcutaneous, or intramuscular administration formulation.

[0467] The pharmaceutical composition can include solid diluents or fillers, sterile aqueous media and various non-toxic organic solvents, etc. In some cases, the carrier can be water, saline, ethanol, glycerol, lactose, sucrose, calcium phosphate, gelatin, dextran, agar, pectin, peanut oil, sesame oil, etc. For parenteral formulations, the carrier usually comprises sterile water or aqueous sodium chloride solution, though other ingredients including those which aid dispersion may be included. Injectable suspensions may also be prepared, in which case appropriate liquid carriers, suspending agents and the like may be employed. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain antioxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. If administered intravenously, carriers can include, for example, physiological saline or phosphate buffered saline (PBS). The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example, water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets of the kind previously described. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's pharmaceutical sciences (Mack Pub. Co., N.J. 1991) which is incorporated by reference herein.

[0468] The compositions can be co-administered with one or more T cells (*e.g.*, engineered T cells) and/or one or more chemotherapeutic agents or chemotherapeutic compounds to a human or mammal. A chemotherapeutic agent can be a chemical compound useful in the treatment of cancer. The chemotherapeutic cancer agents that can be used in combination with the disclosed T cell include, but are not limited to, mitotic inhibitors (vinca alkaloids). These include vincristine, vinblastine, vindesine and Navelbine™ (vinorelbine, 5'-noranhydroblastine). In yet other cases, chemotherapeutic cancer agents include topoisomerase I inhibitors, such as camptothecin compounds. As used herein, "camptothecin compounds" include Camptosar™

(irinotecan HCL), Hycamtin™ (topotecan HCL) and other compounds derived from camptothecin and its analogues. Another category of chemotherapeutic cancer agents that can be used in the methods and compositions disclosed herein can be podophyllotoxin derivatives, such as etoposide, teniposide and mitopodozide. The present disclosure further encompasses other chemotherapeutic cancer agents known as alkylating agents, which alkylate the genetic material in tumor cells. These include without limitation cisplatin, cyclophosphamide, nitrogen mustard, trimethylene thiophosphoramide, carmustine, busulfan, chlorambucil, belustine, uracil mustard, chlomaphazin, and dacarbazine. The disclosure encompasses antimetabolites as chemotherapeutic agents. Examples of these types of agents include cytosine arabinoside, fluorouracil, methotrexate, mercaptopurine, azathioprime, and procarbazine. An additional category of chemotherapeutic cancer agents that can be used in the methods and compositions disclosed herein includes antibiotics. Examples include without limitation doxorubicin, bleomycin, dactinomycin, daunorubicin, mithramycin, mitomycin, mytomycin C, and daunomycin. There are numerous liposomal formulations commercially available for these compounds. The present disclosure further encompasses other chemotherapeutic cancer agents including without limitation anti-tumor antibodies, dacarbazine, azacytidine, amsacrine, melphalan, ifosfamide and mitoxantrone.

[0469] The pharmaceutical composition can comprise one or more herein described cells. Cells can be extracted from a human as described herein. Cells can be genetically altered *ex vivo* and used accordingly. These cells can be used for cell-based therapies. These cells can be used to treat disease in a recipient (*e.g.*, a human). For example, these cells can be used to treat cancer.

[0470] In some cases, a subject may receive a percentage of described engineered cells in a total population of cells that can be introduced. A patient may be infused with as many cells that can be generated for them. In some cases, cells that are infused into a patient are not all engineered. For example, at least 90% of cells that can be introduced into a patient can be engineered. In other instances, at least 40% of cells that are introduced into a patient can be engineered. For example, a patient may receive any number of engineered cells, 10%, 15%, 20%, 25%, 30%,35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the total introduced population.

[0471] The disclosed cells herein can be administered in combination with other anti-tumor agents, including cytotoxic/antineoplastic agents and anti-angiogenic agents. Cytotoxic/anti-neoplastic agents can be defined as agents who attack and kill cancer cells. Anti-angiogenic

agents can also be used. Suitable anti-angiogenic agents for use in the disclosed methods and compositions include anti-VEGF antibodies, including humanized and chimeric antibodies, anti-VEGF aptamers and antisense oligonucleotides. Other inhibitors of angiogenesis include angiostatin, endostatin, interferons, interleukin 1 (including α and β) interleukin 12, retinoic acid, and tissue inhibitors of metalloproteinase-1 and -2. (TIMP-1 and -2). Small molecules, including topoisomerases such as razoxane, a topoisomerase II inhibitor with anti-angiogenic activity, can also be used.

XII. Methods of Use

[0472] Described herein are methods of treating a disease (*e.g.*, cancer) or disorder. The methods can comprise administering to a subject in need thereof the Ago system, the Ago fusion polypeptide, the polynucleic acid, the cell, the pharmaceutical composition, or any combination thereof. In some cases, the method comprises parenteral injection such as intravenous, intramuscular, or subcutaneous injection.

[0473] Described herein is a method of treating a disease (*e.g.*, cancer) in a recipient comprising transplanting to the recipient one or more Argonaute modified cells (including organs and/or tissues). Generally, modified cells described herein can be expanded by contact with a surface having attached thereto an agent that can stimulate a CD3 TCR complex associated signal and a ligand that can stimulate a co-stimulatory molecule on the surface of the T cells. In particular, cell populations can be stimulated *in vitro* such as by contact with an anti-CD3 antibody or antigen-binding fragment thereof, or an anti-CD2 antibody immobilized on a surface, or by contact with a protein kinase C activator (*e.g.*, bryostatin) sometimes in conjunction with a calcium ionophore. For co-stimulation of an accessory molecule on the surface of modified cells, a ligand that binds the accessory molecule can be used. For example, a population of cells can be contacted with an anti-CD3 antibody and an anti-CD28 antibody, under conditions that can stimulate proliferation of the T cells. In some cases, 4-1BB can be used to stimulate cells. For example, cells can be stimulated with 4-1BB and IL-21 or another cytokine. In some cases $5x10^{10}$ cells will be administered to a subject. In other cases, $5x10^{11}$ cells will be administered to a subject.

[0474] In some embodiments, about $5x10^{10}$ cells are administered to a subject. In some embodiments, about $5x10^{10}$ cells represent the median amount of cells administered to a subject. In some embodiments, about $5x10^{10}$ cells are necessary to affect a therapeutic response in a

subject. In some embodiments, at least about at least about $1x10^7$ cells, at least about $2x10^7$ cells, at least about $3x10^7$ cells, at least about $4x10^7$ cells, at least about $5x10^7$ cells, at least about $6x10^7$ cells, at least about $6x10^7$ cells, at least about $8x10^7$ cells, at least about $9x10^7$ cells, at least about $1x10^8$ cells, at least about $2x10^8$ cells, at least about $3x10^8$ cells, at least about $4x10^8$ cells, at least about $5x10^8$ cells, at least about $6x10^8$ cells, at least about $6x10^8$ cells, at least about $8x10^8$ cells, at least about $9x10^8$ cells, at least about $1x10^9$ cells, at least about $2x10^9$ cells, at least about $3x10^9$ cells, at least about $4x10^9$ cells, at least about $5x10^9$ cells, at least about $6x10^9$ cells, at least about $6x10^9$ cells, at least about $8x10^9$ cells, at least about $9x10^9$ cells, at least about $1x10^{10}$ cells, at least about $2x10^{10}$ cells, at least about $3x10^{10}$ cells, at least about $4x10^{10}$ cells, at least about $5x10^{10}$ cells, at least about $6x10^{10}$ cells, at least about $6x10^{10}$ cells, at least about $8x10^{10}$ cells, at least about $9x10^{10}$ cells, at least about $1x10^{11}$ cells, at least about $2x10^{11}$ cells, at least about $3x10^{11}$ cells, at least about $4x10^{11}$ cells, at least about $5x10^{11}$ cells, at least about $6x10^{11}$ cells, at least about $6x10^{11}$ cells, at least about $8x10^{11}$ cells, at least about $9x10^{11}$ cells, or at least about $1x10^{12}$ cells. For example, about $5x10^{10}$ cells can be administered to a subject. In another example, starting with $3x10^6$ cells, the cells can be expanded to about $5x10^{10}$ cells and administered to a subject. In some cases, cells are expanded to sufficient numbers for therapy. For example, 5 $x10^7$ cells can undergo rapid expansion to generate sufficient numbers for therapeutic use. In some cases, sufficient numbers for therapeutic use can be $5x10^{10}$. Any number of cells can be infused for therapeutic use. For example, a subject can be infused with a number of cells between $1x10^6$ to $5x10^{12}$ inclusive. A subject can be infused with as many cells that can be generated for them. In some cases, cells that are infused into a subject are not all engineered. For example, at least 90% of cells that are infused into a subject can be engineered. In other instances, at least 40% of cells that are infused into a subject can be engineered.

[0475] In some embodiments, a method of the present disclosure comprises calculating and/or administering to a subject an amount of modified cells necessary to affect a therapeutic response in the subject. In some embodiments, calculating the amount of engineered cells necessary to affect a therapeutic response comprises the viability of the cells and/or the efficiency with which a transgene has been integrated into the genome of a cell. In some embodiments, in order to affect a therapeutic response in a subject, modified cells that can be administered to a subject can be viable. In some embodiments, in order to effect a therapeutic response in a subject, at least about 95%, at least about 90%, at least about 85%, at least about 80%, at least about 75%, at least about 70%, at least about 65%, at least about 60%, at least about 55%, at least about 50%, at least

about 45%, at least about 40%, at least about 35%, at least about 30%, at least about 25%, at least about 20%, at least about 15%, at least about 10% of the cells are viable cells. In some embodiments, in order to affect a therapeutic response in a subject, the Argonaute polypeptide modified cells administered to a subject can be cells that have had one or more transgenes successfully integrated into the genome of the cell. In some embodiments, in order to effect a therapeutic response in a subject, at least about 95%, at least about 90%, at least about 85%, at least about 80%, at least about 75%, at least about 70%, at least about 65%, at least about 60%, at least about 55%, at least about 50%, at least about 45%, at least about 40%, at least about 35%, at least about 30%, at least about 25%, at least about 20%, at least about 15%, at least about 10% of the cells have had one or more transgenes successfully integrated into the genome of the cell.

[0476] The methods disclosed herein can be used for treating or preventing disease including, but not limited to, cancer, cardiovascular diseases, lung diseases, liver diseases, skin diseases, or neurological diseases by administering to a subject in need thereof Ago modified cells.

[0477] In some embodiments, described herein are methods of treating cancer by administering to a subject in need thereof Ago modified cells. In some embodiments, the cancer is a solid tumor. In some embodiments, the cancer is a hematological malignancy. In some embodiments, the cancer is acute lymphocytic cancer, acute myeloid leukemia, alveolar rhabdomyosarcoma, bladder cancer, bone cancer, brain cancer, breast cancer, cancer of the anus, anal canal, rectum, cancer of the eye, cancer of the intrahepatic bile duct, cancer of the joints, cancer of the neck, gallbladder, or pleura, cancer of the nose, nasal cavity, or middle ear, cancer of the oral cavity, cancer of the vulva, chronic lymphocytic leukemia, chronic myeloid cancer, colon cancer, esophageal cancer, cervical cancer, fibrosarcoma, gastrointestinal carcinoid tumor, Hodgkin lymphoma, hypopharynx cancer, kidney cancer, larynx cancer, leukemia, liquid tumors, liver cancer, lung cancer, lymphoma, malignant mesothelioma, mastocytoma, melanoma, multiple myeloma, nasopharynx cancer, non-Hodgkin lymphoma, ovarian cancer, pancreatic cancer, peritoneum, omentum, and mesentery cancer, pharynx cancer, prostate cancer, rectal cancer, renal cancer, skin cancer, small intestine cancer, soft tissue cancer, solid tumors, stomach cancer, testicular cancer, thyroid cancer, ureter cancer, and/or urinary bladder cancer.

[0478] Transplanting can be by any type of transplanting. Sites can include, but not limited to, liver subcapsular space, splenic subcapsular space, renal subcapsular space, omentum, gastric or intestinal submucosa, vascular segment of small intestine, venous sac, testis, brain, spleen, or

cornea. For example, transplanting can be subcapsular transplanting. Transplanting can also be intramuscular transplanting. Transplanting can be intraportal transplanting.

[0479] Transplanting can be of one or more cells from a human. For example, the one or more cells can be from an organ, which can be a brain, heart, lungs, eye, stomach, pancreas, kidneys, liver, intestines, uterus, bladder, skin, hair, nails, ears, glands, nose, mouth, lips, spleen, gums, teeth, tongue, salivary glands, tonsils, pharynx, esophagus, large intestine, small intestine, rectum, anus, thyroid gland, thymus gland, bones, cartilage, tendons, ligaments, suprarenal capsule, skeletal muscles, smooth muscles, blood vessels, blood, spinal cord, trachea, ureters, urethra, hypothalamus, pituitary, pylorus, adrenal glands, ovaries, oviducts, uterus, vagina, mammary glands, testes, seminal vesicles, penis, lymph, lymph nodes or lymph vessels. The one or more cells can also be from a brain, heart, liver, skin, intestine, lung, kidney, eye, small bowel, or pancreas. The one or more cells can be from a pancreas, kidney, eye, liver, small bowel, lung, or heart. The one or more cells can be from a pancreas. The one or more cells can be pancreatic islet cells, for example, pancreatic β cells. The one or more cells can be any blood cells, such as peripheral blood mononuclear cell (PBMC), lymphocytes, monocytes or macrophages. The one or more cells can be any immune cells such as lymphocytes, B cells, or T cells.

[0480] The method disclosed herein can also comprise transplanting one or more cells (*e.g.*, autologous cells or allogeneic cells), wherein the one or more cells can be any types of cells. For example, the one or more cells can be epithelial cells, fibroblast cells, neural cells, keratinocytes, hematopoietic cells, melanocytes, chondrocytes, lymphocytes (B and T), macrophages, monocytes, mononuclear cells, cardiac muscle cells, other muscle cells, granulosa cells, cumulus cells, epidermal cells, endothelial cells, pancreatic islet cells, blood cells, blood precursor cells, bone cells, bone precursor cells, neuronal stem cells, primordial stem cells, hepatocytes, keratinocytes, umbilical vein endothelial cells, aortic endothelial cells, microvascular endothelial cells, fibroblasts, liver stellate cells, aortic smooth muscle cells, cardiac myocytes, neurons, Kupffer cells, smooth muscle cells, Schwann cells, and epithelial cells, erythrocytes, platelets, neutrophils, lymphocytes, monocytes, eosinophils, basophils, adipocytes, chondrocytes, pancreatic islet cells, thyroid cells, parathyroid cells, parotid cells, tumor cells, glial cells, astrocytes, red blood cells, white blood cells, macrophages, epithelial cells, somatic cells, pituitary cells, adrenal cells, hair cells, bladder cells, kidney cells, retinal cells, rod cells, cone cells, heart cells, pacemaker cells, spleen cells, antigen presenting cells, memory cells, T cells, B cells, plasma cells, muscle cells, ovarian cells, uterine cells, prostate cells, vaginal epithelial cells,

sperm cells, testicular cells, germ cells, egg cells, leydig cells, peritubular cells, sertoli cells, lutein cells, cervical cells, endometrial cells, mammary cells, follicle cells, mucous cells, ciliated cells, nonkeratinized epithelial cells, keratinized epithelial cells, lung cells, goblet cells, columnar epithelial cells, dopamiergic cells, squamous epithelial cells, osteocytes, osteoblasts, osteoclasts, dopaminergic cells, embryonic stem cells, fibroblasts and fetal fibroblasts. Further, the one or more cells can be pancreatic islet cells and/or cell clusters or the like, including, but not limited to pancreatic α cells, pancreatic β cells, pancreatic δ cells, pancreatic F cells (*e.g.*, PP cells), or pancreatic ε cells. In one instance, the one or more cells can be pancreatic α cells. In another instance, the one or more cells can be pancreatic β cells.

[0481] A donor can be at any stage of development including, but not limited to, fetal, neonatal, young and adult. For example, donor T cells can be isolated from an adult human. Donor human T cells can be under the age of 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 year(s). For example, T cells can be isolated from a human under the age of 6 years. T cells can also be isolated from a human under the age of 3 years. A donor can be older than 10 years.

[0482] The instant disclosure also provides materials and methods comprising modified polynucleotides and methods of using such polynucleotides for ameliorating one or more symptoms or complications associated with human genetic diseases. For example, the method can comprise genome editing using the polynucleotides.

[0483] Disclosed herein is also a method of genomically editing a target polynucleic acid utilizing the system, the polypeptide, or the polynucleic acid described herein. The method of modifying a target polynucleic acid can comprise (a) contacting the target polynucleic acid with an Ago polypeptide and a guiding polynucleic acid and (b) modifying the target polynucleic acid. For example, the method can comprise introducing the Ago system or the fusion polypeptide into a cell that contains the target polynucleic acid. For another example, the method can comprise introducing into a cell the system that comprises an Ago and a nucleic acid unwinding polypeptide. The Ago and the polynucleic acid unwinding polypeptide can be introduced into the cell individually or as a fused polypeptide. The method can also comprise introducing into the cell the described polynucleic acid. As described herein, the Ago system, the fusion polypeptide, and/or the polynucleotides encoding the same can be delivered, i.e., introduced, into a cell by any suitable means such as vectors and lipid nanoparticles.

[0484] In some cases, the method also comprises contacting the target polynucleic acid with a protein expressed by a gene of the microbiome prokaryotic organism located in an adjacent

operon to a gene encoding the Ago polypeptide. The gene located in an adjacent operon can be one that is involved in defense, stress response, a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), DNA replication, DNA recombination, DNA repair, and transcription.

[0485] The instant disclosure describes a method of identifying an Ago polypeptide. The method of identifying an Ago polypeptide can comprise comparing the genome sequences with a nucleic acid sequence of a known Ago polypeptide. The known Ago polypeptide can be a Clostridium Argonaute. The method of identifying an Ago polypeptide can comprise identifying a sequence that has 20% or more sequence identity to the nucleic acid sequence of a known Ago polypeptide, as measured by Needleman-Wunsch algorithm. In some cases, the identified sequence encodes an Ago polypeptide having at least 900 amino acid residues.

XIII. Kits

[0486] Disclosed herein can be kits comprising the compositions, the Ago, the fusion polypeptides, the polynucleic acid, or any combination thereof. Disclosed herein can also be kits for the treatment or prevention of a cancer, pathogen infection, immune disorder or allogeneic transplant. The kit can comprise a disclosed Ago system. The kit can comprise a fusion polypeptide comprising the Ago. The kit can comprise a herein described polynucleic acid, such as one that encodes the Ago or the fusion polypeptide. The kit can comprise one or more of the cells. The kit can also comprise the pharmaceutical composition. The kit can further comprise instructions for using the component therein.

[0487] In one embodiment, the kit can include a therapeutic or prophylactic composition containing an effective amount of a composition of nuclease modified cells in unit dosage form. In some cases, the kit comprises one or more sterile containers, which can be boxes, ampules, bottles, vials, tubes, bags, pouches, blister-packs, or other suitable container forms known in the art. Such containers can be made of plastic, glass, laminated paper, metal foil, or other materials suitable for holding medicaments. In some cases, Ago modified cells can be provided together with instructions for administering the cells to a subject having or at risk of developing a cancer, pathogen infection, immune disorder or allogeneic transplant. Instructions can generally include information about the use of the composition for the treatment or prevention of cancer, pathogen infection, immune disorder or allogeneic transplant. In some cases, a kit can include from about 1 x 10^4 cells to about 1 x 10^{12} cells. In some cases a kit can include at least about 1x10^5 cells, at

least about $1x10^6$ cells, at least about $1x10^7$ cells, at least about $4x10^7$ cells, at least about $5x10^7$ cells, at least about $6x10^7$ cells, at least about $6x10^7$ cells, at least about $8x10^7$ cells, at least about $9x10^7$ cells, at least about $1x10^8$ cells, at least about $2x10^8$ cells, at least about $3x10^8$ cells, at least about $4x10^8$ cells, at least about $5x10^8$ cells, at least about $6x10^8$ cells, at least about $6x10^8$ cells, at least about $8x10^8$ cells, at least about $9x10^8$ cells, at least about $1x10^9$ cells, at least about $2x10^9$ cells, at least about $3x10^9$ cells, at least about $4x10^9$ cells, at least about $5x10^9$ cells, at least about $6x10^9$ cells, at least about $6x10^9$ cells, at least about $8x10^9$ cells, at least about $9x10^9$ cells, at least about $1x10^{10}$ cells, at least about $2x10^{10}$ cells, at least about $3x10^{10}$ cells, at least about $4x10^{10}$ cells, at least about $5x10^{10}$ cells, at least about $6x10^{10}$ cells, at least about $6x10^{10}$ cells, at least about $8x10^{10}$ cells, at least about $9x10^{10}$ cells, at least about $1x10^{11}$ cells, at least about $2x10^{11}$ cells, at least about $3x10^{11}$ cells, at least about $4x10^{11}$ cells, at least about $5x10^{11}$ cells, at least about $6x10^{11}$ cells, at least about $6x10^{11}$ cells, at least about $8x10^{11}$ cells, at least about $9x10^{11}$ cells, or at least about $1x10^{12}$ cells. For example, about $5x10^{10}$ cells can be included in a kit. In another example, a kit can include $3x10^6$ cells; the cells can be expanded to about $5x10^{10}$ cells and administered to a subject.

[0488] In some cases, a kit can include allogenic cells. In some cases, a kit can include cells that can comprise a genomic modification. In some cases, a kit can comprise "off-the-shelf" cells. In some cases, a kit can include cells that can be expanded for clinical use. In some cases, a kit can contain contents for a research purpose.

[0489] In some cases, the instructions include at least one of the following: description of the therapeutic agent; dosage schedule and administration for treatment or prevention of a neoplasia, pathogen infection, immune disorder or allogeneic transplant or symptoms thereof; precautions; warnings; indications; counter-indications; overdosage information; adverse reactions; animal pharmacology; clinical studies; and/or references. The instructions can be printed directly on the container (when present), or as a label applied to the container, or as a separate sheet, pamphlet, card, or folder supplied in or with the container. In some cases, instructions provide procedures for administering nuclease modified cells at least about 1, 2, 3, 4, 5, 6, 7, 8, 9,10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or up to 2 days, 3 days, 4 days, 5 days, 6 days, or 7 days after administering a chemotherapeutic agent. In some cases, instructions provide procedures for administering engineered cells at least 24 hours after administering a chemotherapeutic agent. Nuclease modified cells can be formulated for intravenous injection.

Nuclease modified cells can be formulated for infusion. In some cases a kit can contain products at a pediatric dosage.

[0490] Further uses of the methods, compositions, or kits described herein can include one or more of the following: genome editing, transcriptional or epigenetic regulation, genome imaging, copy number analysis, analysis of living cells, detection of highly repetitive genome sequence or structure, detection of complex genome sequences or structures, detection of gene duplication or rearrangement, enhanced FISH labeling, unwinding of target nucleic acid, large scale diagnostics of diseases and genetic disorders related to genome deletion, duplication, and rearrangement, use of an RNA oligo chip with multiple unique gRNAs or gDNAs for high-throughput imaging and/or diagnostics, multicolor differential detection of target sequences, identification or diagnosis of diseases of unknown cause or origin, and 4-dimensional (*e.g.*, time-lapse) or 5-dimensional (*e.g.*, multicolor time-lapse) imaging of cells (*e.g.,* live cells), tissues, or organisms.

EXAMPLES

Example 1. Identification of Clostridia Argonautes

[0491] Argonautes of class Clostridia were identified as phylogenetic branch Ago41/69/70 (**FIG. 2**), including taxonomy (**FIG. 3**) and host and environmental information gathered from JGI database (**FIG. 4**). The exemplary taxonomy-specificity of the Ago41 branch is presented in **FIG. 5** and **FIG. 6**. The sequence specificity for the Ago41/69/70 branch was determined and a pairwise sequence comparison using the Needleman-Wunsch algorithm for global sequence pairwise comparison was conducted (**FIG. 7**). The amino acid sequence and nucleic acid sequence of Clostridia Agos, Ago69, Ago41, and Ago70, were determined and are disclosed in Table 1 (amino acid sequences) and Table 2 (nucleic acid sequences).

Example 2. Cleavage of ssDNA by Clostridia Argonautes

[0492] The cleavage of single stranded DNA (ssDNA) by Ago41 with a guide DNA (gDNA) was tested. The reaction buffer contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, and 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250nM : 250 nM). The time course included two replicates of 5 minutes, 15 minutes, 30 minutes, 60 minutes, 120 minutes, and 240 minutes. The gDNA was preloaded with Ago41 by incubation of the gDNA and Ago41 protein at 37°C for 15 min. As shown in **FIG. 8**, Ago41 is able to cleave ssDNA at each time point tested.

[0493] The cleavage of single stranded DNA (ssDNA) by Ago69 with a guide DNA (gDNA) was also tested. The reaction buffer contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, and 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The time course included two replicates of 5 minutes, 15 minutes, 30 minutes, 60 minutes, 120 minutes, and 240 minutes. The gDNA was preloaded with Ago41 by incubation of Ago69 and gDNA at 37°C for 15 min. As shown in **FIG. 9**, Ago69 is able to cleave ssDNA at each time point tested.

[0494] The cleavage of single stranded DNA (ssDNA) by Ago69 with a guide DNA (gDNA) was tested as above, but with varying cleavage times. The time course included two replicates of 0 minutes, 0.5 minutes, 1 minute, 60 minutes, 2 minutes, 3 minutes, 4 minutes, 5 minutes, and 10 minutes. The reaction buffer contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The gDNA was preloaded with Ago69 by incubation of the gDNA and Ago69 protein at 37°C for 15 min. As shown in **FIG. 10**, the Ago69 is able to cleave ssDNA at each time point tested. Cleavage at the 0 minute time point reflects that it take several seconds stop the reaction.

Example 3. The effect of mutating DEDX domain in Clostridia Argonautes

[0495] The effect of mutating the DEDX domain of Ago41 on Ago41 mediated cleavage of ssDNA with guide DNA (gDNA) was evaluated. The cleavage assay was allowed to proceed for 1 hour with ssDNA template, gDNA, and either wild type (WT) Ago41 or mutant Ago41. The ssDNA template is 90 nucleotides in length with expected cleavage products of 64 and 24 nucleotides each. The mutated Ago41 (MUT) contained the following amino acid substitutions in the DEDX domain: D559A, E595A, and D629A. The template DNA was 90 nucleotides in length. The results show that inclusion of the MUT Ago41 inhibited Ago41 mediated cleavage of the template ssDNA (**FIG. 48**). This suggests that the catalytic activity of Ago41 (e.g., as shown in Example 2) is dependent on the known intact catalytic domain of the Ago.

[0496] The corresponding mutation sites used in Ago41 (DEDX domain) were mapped for Ago69 and presented in **FIG. 49**. These include D544A, E580A, and D730A. Potential additional mutations sites we also mapped on Ago69, including conserved lysine residues putatively involved in DNA binding specificity (**FIG. 50**).

Example 4. The effect of temperature on secondary structure of ssDNA template and gDNA

[0497] The effect of temperature changes on the structure of single stranded DNA (ssDNA) template was analyzed by NUPAK. As shown in **FIG. 11**, increasing temperature to each of 37°C, 55°C, 65°C, and 75°C changes (e.g., decreases the number of) secondary structures in the ssDNA template sequence, with no secondary structures present at 75°C. The effect of temperature changes on the structure of gDNA was also analyzed by NUPAK. As shown in **FIG. 12**, increasing the temperature to each of 37°C, 55°C, 65°C, and 75°C changes (e.g., decreases the number of) secondary structures in the ssDNA template sequence, with no secondary structures present at 65°C or 75°C.

Example 5. The effect of temperature on Clostridia Ago cleavage

[0498] The effect of temperature on single strand DNA (ssDNA) cleavage by Ago69 with a ssDNA guide (gDNA) was analyzed. The gDNA was preloaded with Ago69 by incubation of the gDNA and Ago69 protein at 37°C for 15 min. The target ssDNA was added to the reaction for 15 minutes at 25°C, 37°C, 42.1°C, 46.5°C, 55°C, 65°C, and 75°C, with a subsequent denaturation step utilizing TBE/Urea sample buffer. The nucleic acids were then resolved by gel electrophoresis. The reaction buffer used contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM MnCl2, 1.6 mM b-MeOH, and 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The results are shown in **FIG. 13**, with Ago69 cleaving at each temperature, including at physiological temperature of 37°C.

[0499] Cleavage of single strand DNA (ssDNA) by Ago69 at different temperatures with target (D) and non-target (NT) ssDNA guides was also analyzed. The target ssDNA was added to the reaction for 15 minutes at 37°C, 65°C, and 75°C, with a subsequent denaturation step utilizing TBE/Urea sample buffer. The nucleic acids were then resolved by gel electrophoresis. The reaction buffer used contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM MnCl2, 1.6 mM b-MeOH, 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The results are shown in **FIG. 14**.

Example 6. The effect of unique ssDNA guides on Clostridia Ago cleavage of ssDNA

[0500] The effect using different ssDNA guides on Ago69 cleavage of ssDNA was evaluated. The targeting guide DNAs as labeled D1, D2, D3, D4, D40, D41, D42, in **FIG. 15A** are shown in

the corresponding map of the secondary structure of the nucleic acid in **FIG. 15B**. The non-targeting guide DNAs are labeled D30, D31, D32, D33, NT1, and NT2 in **FIG. 15A**. The results are presented in **FIG. 15A**.

Example 7. The effect of Clostridia Ago denaturation on cleavage

[0501] The effect of denaturing Ago69 before gDNA binding on cleavage of ssDNA cleavage was evaluated. The Ago69 alone was incubated for 15 minutes at 37°C. The Ago69 protein was then denatured by incubation for 60 minutes at 25°C, 37°C, 42.1°C, 46.5°C, 55°C, 65°C, and 75°C. The gDNA was loaded with Ago69 by incubation at 37°C for 15 min. Then the target ssDNA was added and incubated at 37°C for 15 min to allow for cleavage, with a subsequent denaturation step utilizing TBE/Urea sample buffer. The nucleic acids were then resolved by gel electrophoresis. The reaction buffer used contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM MnCl2, 1.6 mM b-MeOH, and 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The results are shown in **FIG. 16**.

[0502] The effect of denaturing Ago69 after gDNA binding on cleavage of ssDNA cleavage was evaluated. The gDNA was loaded with Ago69 by incubation at 37°C for 15 min. The protein was then denatured by incubation for 60 minutes at 25°C, 37°C, 42.1°C, 46.5°C, 55°C, 65°C, and 75°C. Then the target ssDNA was added and incubated at 37°C for 15 min to allow for cleavage, with a subsequent denaturation step utilizing TBE/Urea sample buffer. The nucleic acids were then resolved by gel electrophoresis. The reaction buffer used contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM MnCl2, 1.6 mM b-MeOH, 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The results are shown in **FIG. 17**.

Example 8. Cleavage activity of Ago41, 69, and 70

[0503] The ssDNA cleavage by Ago41, 69, and 70 with ssDNA guide (gDNA) (D1) or ssRNA guide (gRNA) (R1) was assessed according to methods described herein. The results are presented in **FIG. 19**, showing Ago41, 69, and 70 each catalyze ssDNA cleavage with gDNA (D1) or gRNA (R1).

Example 9. Cleavage of ssDNA by Clostridia Ago 69 with guide RNA

[0504] The ability of Ago69 to cleave ssDNA with a guide RNA (gRNA) was evaluated. The reaction buffer used contained 20 mM Tris/HCl at pH7.5, 125 mM NaCl, 5 mM MnCl2, 1.6 mM b-MeOH, and 0.3% BSA. The template (T1) was a 90 nucleotide ssDNA with expected cleavage products of 66 nucleotides and 24 nucleotides. The guide RNA has phosphorothioate bonds on the 5' and 3' ends. The Ago41:gDNA:Template were added at a ratio 1:1:1 (equaling 250 nM : 250 nM : 250 nM). The gDNA was loaded with Ago69 by incubation at 37°C for 15 min. The cleavage reactions were allowed to proceed for 5 minus, 15 minutes, 30 minutes, 60 minutes, 120 minutes, or 240 minutes. The results are presented in **FIG. 20**, showing Ago69 mediated cleavage of the ssDNA at each time point measured.

Example 10. The effect of Clostridia Ago level and guide DNA length on cleavage

[0505] The effect of the level of Ago70 in a cleavage reaction was evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used with a gDNA. The amount of Ago70 added to each cleavage reaction included 150ng, 300ng, 600ng, 900ng, 1200ng, and 1500ng. The results show a clear dose response with an increase in cleavage as the level of Ago70 increases, with saturation between 900ng and 1200ng of Ago70 (**FIG. 25A**).

[0506] The effect of the length of the guide DNA (gDNA) on Ago70 cleavage was also evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used. The length of the gDNA used in each reaction included 30 nucleotides, 25 nucleotides, 21 nucleotides, 20 nucleotides, 19 nucleotides, 18 nucleotides, 17 nucleotides, 16 nucleotides, 15 nucleotides, 15 nucleotides, 14 nucleotides, and 13 nucleotides. The results show that Ago70 cleaves ssDNA with a gDNA of 14-21 nucleotides in length (**FIG. 25B**).

Example 11. Effect of ionic strength and divalent cation concentration on Clostridia Ago cleavage

[0507] The effect of Mg^{2+} and Mn^{2+} concentration on Ago70 was evaluated. Cleavage was allowed to proceed for 1 hour with a template ssDNA of 90 nucleotides in all reactions and a gDNA. $MgCl_2$ concentrations were varied from 1mM $MgCl_2$, 5mM $MgCl_2$, 10 mM $MgCl_2$, 20mM $MgCl_2$. In a separate experiment, $MnCl_2$ concentrations were varied from 1mM $MnCl_2$, 5mM $MnCl_2$, 10 mM $ngCl_2$, 20mM $MnCl_2$. The results show no obvious sensitivity of Ago70 to the Mg^{2+} (**FIG. 26A**) or Mn^{2+} (**FIG. 26B**) at concentrations tested.

[0508] The effect of NaCl concentration was evaluated for Ago70 mediated cleavage of ssDNA with a DNA guide. The NaCl concentrations tested included 50mM, 125mM, 250mM, and 500mM. Template ssDNA 90 nucleotides in length was used in all reactions. Cleavage was allowed to proceed for 1 hour. The results show no obvious sensitive of Ag70 to the NaCl2 concentrations tested (**FIG. 27**).

[0509] The effect of NaCl concentration on Ago 41 and Ago69 mediated cleavage of ssDNA with a gDNA was also evaluated. The NaCl concentrations tested included 666mM, 333mM, 166mM, 66mM, or 33mM. The results are presented in **FIG. 21**.

Example 12. Analysis of mesophilic Ago using guide DNA - Ago02

[0510] The effect of the level of Ago02 on cleavage of ssDNA was evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used. The amount of Ago02 added to each cleavage reaction included 150ng, 300ng, 600ng, 900ng, 1200ng, and 1500ng. The results show a clear dose response with an increase in cleavage as the level of Ago70 increases, with no saturation at the concentrations of Ago20 tested (**FIG. 22A**).

[0511] The effect of the length of the guide DNA (gDNA) on Ago02 cleavage was also evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used. The length of the gDNA used in each reaction included 30 nucleotides, 25 nucleotides, 21 nucleotides, 20 nucleotides, 19 nucleotides, 18 nucleotides, 17 nucleotides, 16 nucleotides, 15 nucleotides, 15 nucleotides, 14 nucleotides, and 13 nucleotides. The results show that Ago20 cleaves ssDNA with a gDNA of 13-21 nucleotides in length (**FIG. 22B**).

[0512] The effect of Mg^{2+} and Mn^{2+} concentration on Ago02 mediated cleavage was evaluated. Cleavage was allowed to proceed for 1 hour with a template ssDNA of 90 nucleotides in all reactions, and guide DNA. $MgCl_2$ concentrations were varied from 1mM $MgCl_2$, 5mM $MgCl_2$, 10 mM $MgCl_2$, 20mM $MgCl_2$. In a separate experiment, $MnCl_2$ concentrations were varied from 1mM $MnCl_2$, 5mM $MnCl_2$, 10 mM $ngCl_2$, 20mM $MnCl_2$. The results show no obvious sensitivity of Ago02 to the Mg^{2+} (**FIG. 23A**). The results indicate that Ago02 may cleave less efficiently with the high Mn^{2+} concentrations tested (**FIG. 23B**).

[0513] The effect of NaCl concentration was evaluated for Ago02 mediated cleavage of ssDNA with a DNA guide. The NaCl concentrations tested included 50mM, 125mM, 250mM, and 500mM. Template ssDNA 90 nucleotides in length was used in all reactions. Cleavage was

allowed to proceed for 1 hour. The results show no obvious sensitivity of Ag02 to the $NaCl_2$ concentrations tested (**FIG. 24**).

Example 13. Effect of guide RNA stability on mesophilic Ago cleavage of ssDNA

[0514] The effect of the stability of the guide RNA (gRNA) was evaluated for Ago23, Ago29, and Ago51. RNase was inhibited with the addition of RNasin to the cleavage reactions (40U/reaction). The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used in each reaction. For the Ago29 experiments, 125ng of protein was used. The results show no obvious increase in cutting efficiency with RNasin (**FIG. 28**).

Example 14. Analysis of mesophilic Ago using guide RNA – Ago23

[0515] The effect of the level of Ago23 on cleavage of ssDNA was evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used with a guide RNA (gRNA) (R1p). The amount of Ago23 added to each cleavage reaction included 150ng, 300ng, 600ng, 900ng, 1200ng, and 1500ng. The results show a clear dose response with an increase in cleavage as the level of Ago23 increases (**FIG. 29A**).

[0516] The effect of the length of the guide RNA (gRNA) on Ago23 cleavage was also evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used. The length of the gRNA used in each reaction included 30 nucleotides, 25 nucleotides, 21 nucleotides, 20 nucleotides, 19 nucleotides, 18 nucleotides, 17 nucleotides, 16 nucleotides, 15 nucleotides, 15 nucleotides, 14 nucleotides, and 13 nucleotides. The results show that Ago23 cleaves ssDNA with a gRNA of 13-21 nucleotides in length (**FIG. 29B**).

[0517] The effect of Mg^{2+} and Mn^{2+} concentration on Ago23 mediated cleavage was evaluated. Cleavage was allowed to proceed for 1 hour with a template ssDNA of 90 nucleotides and a gRNA (R1p) in all reactions. $MgCl_2$ concentrations were varied from 1mM $MgCl_2$, 5mM $MgCl_2$, 10 mM $MgCl_2$, 20mM $MgCl_2$. In a separate experiment, $MnCl_2$ concentrations were varied from 1mM $MnCl_2$, 5mM $MnCl_2$, 10 mM $ngCl_2$, 20mM $MnCl_2$. The results indicate that Ago23 may cleave less efficiently with the low Mg^{2+} concentrations tested (**FIG. 30A**). The results also indicate that Ago23 may cleave less efficiently with the low Mn^{2+} concentrations tested (**FIG. 30B**). The results further indicate that Ago23 may cleave with better efficiency with Mg^{2+} versus Mn^{2+} (**FIG. 30A-30B**).

[0518] The effect of NaCl concentration was evaluated for Ago23 mediated cleavage of ssDNA with a RNA guide (gRNA). The NaCl concentrations tested included 50mM, 125mM, 250mM,

and 500mM. Template ssDNA 90 nucleotides in length was used in all reactions. Cleavage was allowed to proceed for 1 hour. The results show that Ago23 cleaves ssDNA only at $NaCl_2$ concentrations above 250mM and has better cleavage efficiency at 500mM (**FIG. 31**).

Example 15. Analysis of mesophilic Ago – Ago29

[0519] The effect of the level of Ago29 on cleavage of ssDNA was evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length and a guide RNA (R1p) was used. The amount of Ago29 added to each cleavage reaction included 150ng, 300ng, 600ng, 900ng, 1200ng, and 1500ng. The protein titration shows a strong non-specific DNA degradation, which is stronger without targeting gRNA (**FIG. 32A**).

[0520] The effect of the length of the guide RNA (gRNA) on Ago29 cleavage was also evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used. Ago29 was added at a concentration of 125ng/reaction. The length of the gRNA used in each reaction included 30 nucleotides, 25 nucleotides, 21 nucleotides, 20 nucleotides, 19 nucleotides, 18 nucleotides, 17 nucleotides, 16 nucleotides, 15 nucleotides, 15 nucleotides, 14 nucleotides, and 13 nucleotides. The results show that Ago29 cleaves ssDNA with a gDNA of 13-21 nucleotides in length (**FIG. 32B**). The results further show strong non-specific DNA degradation, which is stronger without targeting gRNA (**FIG. 32B**).

[0521] The effect of Mg^{2+} and Mn^{2+} concentration on Ago29 mediated cleavage was also evaluated. Cleavage was allowed to proceed for 1 hour with a template ssDNA of 90 nucleotides and a gRNA (R1p) in all reactions. $MgCl_2$ concentrations were varied from 1mM $MgCl_2$, 5mM $MgCl_2$, 10 mM $MgCl_2$, 20mM $MgCl_2$. In a separate experiment, $MnCl_2$ concentrations were varied from 1mM $MnCl_2$, 5mM $MnCl_2$, 10 mM $ngCl_2$, 20mM $MnCl_2$. The results indicate that the non-specific activity of Ago29 is weaker with the low Mn^{2+} concentrations tested (**FIG. 33B**). The results further indicate that Ago29 may cleave with better efficiency with Mn^{2+} versus Mg^{2+} (**FIG. 33A-33B**).

[0522] The effect of NaCl concentration was evaluated for Ago29 mediated cleavage of ssDNA with a RNA guide. The NaCl concentrations tested included 50mM, 125mM, 250mM, and 500mM. Template ssDNA 90 nucleotides in length was used in all reactions. Cleavage was allowed to proceed for 1 hour. The results are presented in **FIG. 34**.

Example 16. Analysis of mesophilic Ago – Ago51

[0523] The effect of the level of Ago51 on cleavage of ssDNA was evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length and a guide RNA (R1p) was used. The amount of Ago51 added to each cleavage reaction included 150ng, 300ng, 600ng, 900ng, 1200ng, and 1500ng. The protein titration shows good cleavage activity from 300ng Ago51 (**FIG. 35A**).

[0524] The effect of the length of the guide RNA (gRNA) on Ago51 cleavage was evaluated. The cleavage reactions were allowed to proceed for 1 hour. Template ssDNA of 90 nucleotides in length was used. The length of the gRNA used in each reaction included 30 nucleotides, 25 nucleotides, 21 nucleotides, 20 nucleotides, 19 nucleotides, 18 nucleotides, 17 nucleotides, 16 nucleotides, 15 nucleotides, 15 nucleotides, 14 nucleotides, and 13 nucleotides. The results show that Ago51 cleaves ssDNA with a gDNA of 13-25 nucleotides in length (**FIG. 35B**).

[0525] The effect of Mg^{2+} and Mn^{2+} concentration on Ago51 mediated cleavage was also evaluated. Cleavage was allowed to proceed for 1 hour with a template ssDNA of 90 nucleotides and a gRNA (R1p) in all reactions. $MgCl_2$ concentrations were varied from 1mM $MgCl_2$, 5mM $MgCl_2$, 10 mM $MgCl_2$, 20mM $MgCl_2$. In a separate experiment, $MnCl_2$ concentrations were varied from 1mM $MnCl_2$, 5mM $MnCl_2$, 10 mM $ngCl_2$, 20mM $MnCl_2$. The results indicate that Ago51 cleavage efficiency may be weaker with the lower Mg^{2+} concentrations (**FIG. 36A**) tested and the lower Mn^{2+} concentrations (**FIG. 36B**) tested. The results further indicate that Ago29 may cleave with better efficiency with Mg^{2+} versus Mn^{2+} (**FIG. 36A-36B**).

[0526] The effect of NaCl concentration was evaluated for Ago51 mediated cleavage of ssDNA with a RNA guide. The NaCl concentrations tested included 50mM, 125mM, 250mM, and 500mM. Template ssDNA 90 nucleotides in length was used in all reactions. Cleavage was allowed to proceed for 1 hour. The results show that Ago51 only cuts at NaCl concentrations greater than 125mM, and has much better cleavage efficiency at over 250mM NaCl (**FIG. 37**).

Example 17. The effect of dsDNA nicking on Clostridia Ago cleavage

[0527] The effect of dsDNA nicking on Ago69 was evaluated. The experimental protocol utilized for the dsDNA "bubble" nicking assay is outlined in **FIG. 38**. The bubble template used was a ssDNA oligo with complementary regions to ensure that no ssDNA is present. The bubble template is 84 nucleotides in length with expected cleavage products of 58 nucleotides and 26 nucleotides. The ssDNA template was 43 nucleotides in length with expected cleavage products of 26 nucleotides and 17 nucleotides. The RecQ helicase unwinds substrates with 3' overhangs. The Nt.AlwI site is included as a positive control. The reaction includes, ssDNA

template:gDNA/cleavage control. The reaction buffer includes 20 mM Tris/HCl pH7.5, 5 mM $MnCl_2$, 125 mM NaCl. The results show Ago69 ssDNA guide dependent nicking of dsDNA bubble template (**FIG. 39**).

Example 18. The effect of GC content on Clostridia Ago cleavage

[0528] The effect of GC content of guide DNA (gDNA) on Ago69 mediated cleavage was evaluated. The sequences of each of the different guide DNAs (D1, D40, D41, D42, D43, and D44), the GC content, and the expected cleavage products are presented in Table 3.

Table 3. gDNA sequence, GC content, and cleavage products

gDNA#	Sequence	GC content	Expected cleavage products
$D1^P$	5'-GCTGCCATCCAGATCGTTATC-3'	52%	66 + 24
$D40^P$	5'-CGTTATCGCCCATGGGGTGCA-3'	62%	79 + 11
$D41^P$	5'-GATCGTTATCGCCCATGGGGT-3'	57%	76 + 14
$D42^P$	5'-GGTGCGGGTGAAGCTGCCATC-3'	67%	53 + 37
$D43^P$	5'-ACTTAGACTGAAGGTGCGGGT-3'	52%	49 + 41
$D44^P$	5'-AGTAATCGTCATCACTTAGAC-3'	38%	62 + 28

[0529] The positioning of each gDNA within the larger nucleic acid sequence is presented in **FIG. 40**. The results of the cleavage assay are presented in **FIG. 40**.

[0530] The effect of GC content of guide DNA (gDNA) on Ago41 mediated cleavage was also evaluated. The sequences of each of the different guide DNAs (D1, D40, D41, D42, D43, and D44), the GC content, and the expected cleavage products are presented in Table 3. The positioning of each gDNA within the larger nucleic acid sequence is presented in **FIG. 42**. The results of the cleavage assay are presented in **FIG. 42**.

[0531] The effect of GC content of guide DNA (gDNA) on Ago70 mediated cleavage was evaluated. The sequences of each of the different guide DNAs (D1, D40, D41, D42, D43, and D44), the GC content, and the expected cleavage products are presented in Table 3. The positioning of each gDNA within the larger nucleic acid sequence is presented in **FIG. 43**. The results of the cleavage assay are presented in **FIG. 43**.

Example 19. The effect of GC content on mesophilic Ago cleavage

[0532] The effect of GC content of guide DNA (gDNA) on Ago02 mediated cleavage was evaluated. The sequences of each of the different guide DNAs (D1, D40, D41, D42, D43, and D44), the GC content, and the expected cleavage products are presented in Table 3. The

positioning of each gDNA within the larger nucleic acid sequence is presented in **FIG. 41**. The results of the cleavage assay are presented in **FIG. 41**.

Example 20. Double stranded DNA cleavage by Clostridia Ago

[0533] The ability of Ago69 to cleave double stranded DNA (dsDNA) was evaluated. The cleavage assay was performed in 1X CutSmart buffer (NEB). The guide DNA (gDNA) was preloaded with Ago69 by incubation of the gDNA with Ago69 at 37°C for 15 minutes. Each of the 8 gDNAs tested were preloaded separately. The gDNAs including the location, CG content, and Tm are presented in **FIG. 51**. Double stranded target plasmid DNA was preincubated for 15 minutes at 37°C. Half reactions were combined before the plasmid template was added. Reactions were incubated for 15 or 60 minutes at 75°C. Linearization was completed with XhoI incubation for 30 minutes at 37°C. The DNA was resolved on a 1% agarose gel stained with SYBR gold. The expected cleavage products are approximately 1.5kb and 3.5kb (**FIG. 52A**). The results show that Ago69 can cleave dsDNA (**FIG. 52B**).

Example 21. The impact of ET-SSB and Eco-SSB proteins on DNA unwinding

[0534] The impact of single strand DNA binding (SSB) proteins on the processivity of DNA unwinding by RecQ helicase was evaluated. The experimental design is outlined in **FIG. 44**. A helicase substrate was used which contained the guide DNA1 (gDNA1) sequence. The initial experiment was conducted with RecQ and ET-SSB on a 3' overhang long substrate. The results show the ET-SSB has a beneficial effect on DNA unwinding (**FIG. 45**). The experiment was repeated with shorter substrate, which produced a better noise/signal ratio (**FIG. 46**). The experiment confirmed that ET-SSB has a beneficial effect on DNA unwinding with the short substrate, with no strong dose dependency effect observed (**FIG. 46**). A third experiment was conducted with RecQ and Eco-SSB on a 3' overhang short substrate. The initial experiment with Eco-SSB showed saturation within 5 minutes, this for the Eco-SSB experiment 10x less RecQ was used. The results show the Eco-SSB has the same beneficial effect on DNA unwinding as ET-SSB, with no strong dose dependency effect (**FIG. 47**). As used herein the terms "ET-SSB" and "Sso-SSB" refer to the same SSB protein from *Saccharolobus solfataricus* of SEQ ID NO: 22, or a functional fragment or variant thereof.

Example 22. The impact of ET-SSB and Eco-SSB proteins on Ago mediated DNA cleavage

[0535] A separate experiment was conducted to evaluate dsDNA cleavage with and without preincubation of plasmid DNA at 75°C and with and without ET-SSB protein. The cleavage assay was performed in 1X CutSmart buffer (NEB). The guide DNA (gDNA) was preloaded

with Ago69 by incubation of the gDNA with Ago69 at 37°C for 15 minutes. Guides were preloaded separately. Plasmid DNA was incubated at 75°C for 15 minutes/no preincubation. Half reactions were combined with ET-SSB protein (0.5µg/reaction) (or control) and plasmid template; incubated for 30 minutes at either 37°C or 39°C (**FIG. 54**), 41.5°C or 44.9°C (**FIG. 55**), 49.1°C or 67°C (**FIG. 56**); and a MluI plasmid restriction digest was carried out for 30 minutes at 37°C. Proteinase K was added to stop the reaction. The DNA was run on a 1% agarose gel and stained with SYBR gold. The expected cleavage products of the M1uI plasmid digest are 4487 and 1827 bp (**FIG. 53A**). The expected cleavage products of the M1uI plasmid digest and Ago69 cleavage are 3816, 1827, and 671 bp (**FIG. 53B**). As shown in **FIG. 54 – FIG. 56**, Ago69 cleavage was dependent on the inclusion of guide DNA (gDNA54 and gDNA 55) and was increased with inclusion of ET-SSB across the temperatures measured, including at 37°C.

[0536] A separate experiment, to confirm the effect of ET-SSB and Eco-SSB on Ago69 mediated dsDNA cleavage was conducted. The cleavage assay was performed in 1X CutSmart buffer (NEB). The guide DNA (gDNA) was preloaded with Ago69 by incubation of the gDNA with Ago69 at 37°C for 15 minutes. Guides were preloaded separately. Half reactions were combined with ET-SSB protein (0.5µg/reaction), Eco-SSB protein, or control, and plasmid template; incubated for 30 minutes at either 37°C, and a MluI-HF or BsmI plasmid restriction digest was carried out for 30 minutes at 37°C. Proteinase K was added to stop the reaction. The DNA was run on a 1% agarose gel and stained with SYBR gold. The expected cleavage products of the MluI-HF plasmid digest are 4487 and 1827 bp (**FIG. 53A**). The expected cleavage products of the Mlul-HF plasmid digest and Ago69 cleavage are 3816, 1827, and 671 bp (**FIG. 53B**). The expected cleavage products of the BsmI plasmid digest are 4596, 1641, and 77 bp (**FIG. 57A**). The expected cleavage products of the BsmI plasmid digest and Ago69 cleavage are 4596, 1089, 552, and 77 bp (**FIG. 57B**). **FIG. 58** shows the cleavage of the plasmid DNA from both the MluI-HF (left) and BsmI (right) digests (high exposure agarose gel) with ET-SSB at 37°C. **FIG. 59** shows the cleavage of the plasmid DNA from both the MluI-HF (left) and BsmI (right) digests (high exposure agarose gel) with Eco-SSB at 37°C. **FIG. 60** (low exposure gel) and **FIG. 61** (high exposure gel) shows a dose response cleavage of plasmid DNA with ET SSB at the indicate ng/reaction (i.e. 1000ng/reaction, 500ng/reaction, 250 ng/reaction, 125 ng/reaction, 75 ng/reaction, or 0 ng/reaction) in a BsmI digestion. The expected cleavage products are the same as previously shown in **FIG. 57A** and **FIG. 57B**.

Example 23. The impact of ET-SSB and TteUvrD helicase on Ago mediated DNA cleavage

[0537] As set out in **FIG. 62**, Ago69 was preloaded with guide DNA 54 (guide 54) (3.75μM) in 1X CutSmart buffer (NEB) supplemented with 5mM ATP and the preloading reaction allowed to proceed for 15 minutes at 37°C to produce a guide 1 reaction. Similarly, Ago69 was preloaded with guide DNA 55 (guide 55) (3.75μM) in 1X CutSmart buffer (NEB) supplemented with 5mM ATP and the preloading reaction allowed to proceed for 15 minutes at 37°C to produce a guide 2 reaction. A reaction mixture of 7μl of guide 1 reaction, 7μl of guide 2 reaction, 1μl ET SSB (500 ng), 1μl Tte UvrD helicase (20 ng), and 1μl plasmid #56 DNA (250-300 ng) was incubated for 30 minutes at 37°C. 1μl of restriction enzyme (Bsal-HF) was added and incubated for 30 minutes at 37°C to mediate digestion. 1μl of proteinase K was added to stop the digestion. The protocol as described is further set out in **FIG. 62**. The expected cleavage products with Bsal-HF digest alone are 6314bp (linearized plasmid). The expected cleavage products with Ago mediated cleavage with D54 and D55 guides and Bsal-HF digest are 4937 and 1341 bp. The results of the cleavage analysis are shown in **FIG. 63**.

Example 24. Expression and purification of helicases and SSBs

[0538] All SSB proteins were expressed and purified as shown in **FIG. 64A-64B**, including TnsSSB, TthSSB, and NeqSSB. Repeat helicase expression and purification were also carried out as shown in **FIG. 64A-64B**, including Eco RecQ, Tth UvrD, Eco UvrD, HEL#100, HEL#75, HEL#76. The helicases and SSBs were then tested in combination with Ago69 as shown in Example 25.

Example 25. Effect of helicases and SSBs on Ago69 mediated cleavage

[0539] Ago69 was preloaded with guides 54 and 55 at a ratio of Ago69:gDNA of 1:1. Guides 54 and 55 were preloaded separately for 15 min at 37°C. Half reactions were combined with 1000 ng of SSB (TneSSB, Tth SSB, Neq SSB, Taq SSB, Tma SSB, Sso SSB, Eco SSB, ET SSB, or control with no SSB protein) and 40 ng of helicase (tTE uVRd, hel#65, HEL#71, HEL#78, HEL#92, or no helicase control). Plasmid #56 DNA was added last (~250 ng/reaction) and incubated for 30 minutes at 37°C. Mlul-HF restriction enzyme was added for 30 minutes at 37°C. Proteinase K was added and incubated for 30 minutes at room temperature to the to stop the digestion. The expected cleavage products of Mlul-HF digest alone are 4487 and 1827 bp. The cleavage products of Ago69 cleavage and Mlul-HF digestion are 3816, 1827, and 671. The results are shown in **FIG. 65**, **FIG. 66**, and **FIG. 67**.

Example 26. First round expression and purification of Ago69 fusion proteins

[0540] Three Ago69 fusion constructs were expressed and purified. Each of the constructs is shown graphically in **FIG. 68**. One construct comprises Ago69 fused to SV40 nuclear localization signal via a linker (Ago69 construct (APO71)); a second construct comprises Ago69 further fused to SsoSSB via a linker (SsoSSB-AGO#69 construct (AP072)); a third construct comprises Ago69 further fused to VP64 (transcriptional activator) (VP64-AGO#69 construct (AP073)). The purified constructs were detected via western blot as shown in **FIG. 69A** and **FIG. 69B**.

Example 27. Cleavage of plasmid DNA mediated by Ago69 fusion proteins

[0541] Plasmid DNA cleavage mediated by the Ago69 containing fusion proteins described in Example 26 was carried out as previously described. The expected cleavage products of Plasmid # 56 using XbaI restriction enzyme and Ago mediated cleavage were 4604, 1388, and 35 bp. The results are shown in **FIG. 70** and **FIG. 71**.

Example 28. Second round expression and purification of Ago69 fusion proteins

[0542] The three Ago fusion constructs described in Example 26 were expressed and purified again in a second round experiment. Each of the constructs is shown graphically in **FIG. 68**. One construct comprises Ago69 fused to SV40 nuclear localization signal via a linker (Ago69 construct (APO71)); a second construct comprises Ago69 further fused to SsoSSB via a linker (SsoSSB-AGO#69 construct (AP072)); a third construct comprises Ago69 further fused to VP64 (transcriptional activator) (VP64-AGO#69 construct (AP073)). The purified constructs were detected via western blot as shown in **FIG. 72A** and **FIG. 72B**.

Example 29. Cleavage of plasmid DNA mediated by Ago69 fusion proteins

[0543] Plasmid DNA cleavage mediated by the Ago69 containing fusion proteins described in Example 28 was carried out as previously described. The expected cleavage products of Plasmid # 56 using XbaI restriction enzyme and Ago mediated cleavage were 4604, 1388, and 35 bp. The results of one cleavage experiment are presented in **FIG. 73**; and the results from a second cleavage experiment are presented in **FIG. 74**.

Example 30. Expression and purification of SsoSSB-Ago69 fusion proteins.

[0544] Six SsoSSB-Ago69 fusion constructs were expressed and purified. The constructs are shown in **FIG. 75**. The constructs included an N-terminal His tag. The purified fusion preparations are shown in **FIG. 76A** and **FIG. 76B**.

Example 31. Cleavage of plasmid DNA mediated by SsoSSB-Ago69 fusion proteins.

[0545] Plasmid DNA cleavage mediated by the SsoSSB-Ago69 containing fusion proteins described in Example 30 was carried out per the protocol below. SsoSSB-Ago69 fusion constructs were separately preloaded with guide DNA 55 (guide 55) or guide DNA 54 (guide 54) in 1X CutSmart buffer (NEB) and the preloading reaction allowed to proceed for 15 minutes at 37°C to produce a guide 2 reaction. Half reaction mixtures were combined with plasmid DNA template (plasmid 56) and incubated for 30 minutes at 37°C (in some samples ET-SSB was also added to the reaction mixture where indicated). Restriction enzyme (Kpn1-HF) was added and incubated for 30 minutes at 37°C to mediate digestion. Proteinase K was added to stop the digestion through incubation at 50°C for 30 minutes. The expected cleavage products using Kpn1-HF restriction enzyme and Ago mediated cleavage were 4723 and 1591 bp. The results of one cleavage experiment are presented in **FIG. 77**; and the results from a second cleavage experiment are presented in **FIG. 78**.

[0546] The ability of the SsoSSB fusion constructs to mediate cleavage at 75°C was also tested. SsoSSB-Ago69 fusion constructs were separately preloaded with guide DNA 55 (guide 55) or guide DNA 54 (guide 54) in 1X CutSmart buffer (NEB) and the preloading reaction allowed to proceed for 15 minutes at 37°C to produce a guide 2 reaction. Half reaction mixtures were combined with plasmid DNA template (plasmid 56) and incubated for 30 minutes at 75°C. Restriction enzyme (Kpn1-HF) was added and incubated for 30 minutes at 37°C to mediate digestion. Proteinase K was added to stop the digestion through incubation at 50°C for 30 minutes. The expected cleavage products using Kpn1-HF restriction enzyme and Ago mediated cleavage were 4723 and 1591 bp. The results of the cleavage experiment are presented in **FIG. 79**.

Example 32. Localization of Ago69 fusion constructs to the nucleus

[0547] Gene editing happens in the nucleus of a mammalian cell. Consequently, delivery of an Argonaute to the nucleus becomes a key prerequisite for its ability to function as a gene editing machine. We determined that Ago69 that it may only tolerate N-terminal fusions as its C-terminus is quite hydrophobic and folds back into the hydrophobic core of the molecule. Consequently, a construct was created in which two SV40-derived nuclear localization signals were fused to the N-terminus of Ago#69 as outlined in the schematic presented in **FIG.80** (SEQ ID NO: 97; AP109). To assess the subcellular localization, this construct was expressed in Hela cells and localization was assessed by immunofluorescence microscopy, staining for Ago#69 using the V5-specific antibody R960-25 (Invitrogen). The data shown in **FIG. 81**, **FIG. 82**, **FIG.**

85, and **FIG. 86** suggest that Ago#69, Ago homologs 2 (SEQ ID NO: 99; SPL0389) and 4 (SEQID NO: 100; SPL0390) upon fusion with two SV40-derived NLSs localize in the nucleus.

[0548] Fusion constructs were created that include the SsoSSB in the fusion construct as outlined in FIG.80 (SEQ ID NO: 98; AP110), essentially adding SsoSSB between the V5-tag and the N-terminus of Ago69. This construct was found to localize almost exclusively in the cytosol (**FIG. 83**). This suggested that the presence of SsoSSB at this position hampered the two SV40-derived NLSs to function.

[0549] To solve this issue, a new series of constructs were made in which we changed the identity and positioning of the nuclear localization signals. In some constructs, the SV40 NLS was exchanged for NPM NLS and the spacing was adjusted as outlined in **FIG.80**. Flexible linkers (GSGS or GSGSS) were also added to ensure accessibility (see **FIG.80**). The constructs were tested for their subcellular localization in Hela cells as described above. The V5 – 2xSV40 NLS – SsoSSB – MYC NLS – AGO69 (SEQ ID NO: 101; SPL0398) was found localized almost exclusively in the nucleus (**FIG. 84**). This is in stark contrast to the SsoSSB fusion tested above (AP110) which showed compromised nuclear localization.

[0550] In summary, these data suggest an effective configuration/construct design that allows the nuclear localization of Ago69 fusions with SsoSSB. This represents a key prerequisite for Ago69 functioning as gene editing tool.

Example 33. Importance of sequence context of guide DNA recognition site

[0551] Argonautes are guided by short DNA or RNA sequences (so called guide DNAs or RNAs) to their target sequence which is complementary. As is the case with Cas9 endonuclease, Argonaute guide DNA sequences may differ in terms of their ability to induce target cleavage by the Argonaute. To test this experimentally, a set of guide DNA pairs were designed (*see* Table 16) targeting two plasmids (plasmid #56 and #70). Linearization of the plasmid by the Argonaute-induced DNA double-strand break was followed by the digestion with a cognate restriction enzyme, leading to a defined cleavage pattern.

Table 16. Guide DNA Sequences

Guide ID	Guide sequence	SEQ ID NO
D1	5Phos/GCTGCCATCCAGATCGTTATC	107
D2	5Phos/GGAGCTGTAGTAGCCGCCGTC	108
D3	5Phos/TAGCCGCCGTCGCGCAGGCTG	109
D30	5Phos/GATAACGATCTGGATGGCAGC	110
D31	5Phos/GGAGCTGTAGTAGCCGCCGTC	111
D32	5Phos/CAGCCTGCGCGACGGCGGCTA	112

D54	5Phos/AATTCGCGTTAAATTTTTGTT	113
D55	5Phos/AACAAAAATTTAACGCGAATT	114
D82	5Phos/GCCCTGAAAATAAAGATTCTC	115
D83	5Phos/GAGAATCTTTATTTTCAGGGC	116
D86	5Phos/GCCCCCGATTTAGAGCTTGAC	117
D87	5Phos/GTCAAGCTCTAAATCGGGGGC	118
D88	5Phos/CCACACCCGCCGCGCTTAATG	119
D89	5Phos/CATTAAGCGCGGCGGGTGTGG	120
D90	5Phos/GGGGAAAGCCGGCGAACGTGG	121
D91	5Phos/CCACGTTCGCCGGCTTTCCCC	122
D92	5Phos/GGCCCCAAGGGGTTATGCTAG	123
D93	5Phos/CTAGCATAACCCCTTGGGGCC	124
D94	5Phos/TATTATTTTCTCCCATGAAGA	125
D95	5Phos/TCTTCATGGGAGAAAATAATA	126
D102	5Phos/AGAACGTGGACTCCAACGTCA	127
D103	5Phos/TGACGTTGGAGTCCACGTTCT	128
D104	5Phos/TAACCAATAGGCCGAAATCGG	129
D105	5Phos/CCGATTTCGGCCTATTGGTTA	130
D106	5Phos/TATTTAGAAAAATAAACAAAT	131
D107	5Phos/ATTTGTTTATTTTTCTAAATA	132

[0552] When testing a set of ~10 guide DNA pairs on these two plasmids, the guide DNAs clearly differed in their ability to induce target cleavage (**FIG. 87**). It was initially hypothesized that the GC content of the guide DNA sequences was responsible for this difference, but upon closer inspection, we found guide DNA pairs which were effective despite a high GC content (e.g. guide pair 92/93 targeting a sequence with 62% GC content) (**FIG. 87**).

[0553] Consequently, it was hypothesized that it may not be guide DNA sequence itself, but rather the sequence context in which the guide DNA recognition lies. If that hypothesis were true, it should be possible to take an inactive guide DNA recognition site and "transplant" it into a region in which guide DNA cutting is permitted. Likewise, transplanting an active guide DNA recognition site into an inactive region should prevent cutting from occurring.

[0554] To test this hypothesis, a series of "guide swapping constructs" were created (**FIG. 88A**, **FIG. 88B, FIG. 89**). We started with plasmid #p56 which bears two guide RNA recognition sites: AE1 is recognized by the guide pair D54/55 and recognition leads to the effective cleavage of the target sequence. AE2 is recognized by the guide pair D82/83, but its recognition does not trigger effective cleavage. We then created two derivates of plasmid #56: In plasmid #114, the

AE1 site was replaced by AE2 (**FIG. 89**). In plasmid #115 the position of AE1 and AE2 were swapped, i.e. AE1 now lies within the sequence context of AE2 and vice versa (**FIG. 89**).

[0555] When testing these plasmids, we made the following key observations: 1) In plasmid #56, the AE1 is effectively cleaved, whereas the AE2 remains uncleaved. This was the starting assumption of the experiment (**FIG. 90** and **FIG. 91**); 2) in plasmid #114, the AE2 site which was previously inactive, is now accessible for cleavage (**FIG. 90** and **FIG. 91**); and 3) in plasmid #115, the AE2 site which was previously inactive, can now induce the target site cleavage and the AE1 which was previously active is now inactive (**FIG. 90** and **FIG. 91**).

[0556] These data suggest that, at least in these assays, the guide DNA recognition site is not the only determinant for enzymatic AGO activity. Instead, the sequence context has a significant impact on the ability of guide DNAs to trigger target cleavage. This implies that certain regions in the plasmid may be more accessible and that accessibility may be a limiting factor for the Argonaute to exert its action.

Example 34. Generation of HAT plasmid

[0557] HAT plasmids were generated in order to test the cleavage efficiency of Agos on regions of DNA with low GC content. HAT versions of plasmid #70 and plasmid #56 were generated. To generate plasmid #70-HAT (high AT region) plasmid #70 was digested with BamH1 and BsrG1. HAT_high AT region was subcloned by using NEBuilder HiFi DNA Assembly. HAT sequence is 144 bp with a 20.14% GC content. The HAT region comprises the following sequence: ATTAGACATAATTTATAGTAGAAATATAGAAATTCTATCTAACTATATTTAAGTTCA ATTGATATCTTTAAAGATTATAGTCACAGTAATAAGAATTGTTAACTATACTTTGATA TCTTTGACTTATTAGTTAAGTCTTAGAAA) (SEQ ID NO: 133).

Example 35. Ago69 cleavage of HAT plasmid

[0558] The ability of Ago69 to cleavage HAT plasmid DNA was assessed. The HAT plasmid was generated as described in Example 34. The cleavage assay was performed in 1X CutSmart buffer (NEB). The guide DNAs (gDNAs) was separately preloaded with Ago69 by incubation of the gDNA with Ago69 at 37°C for 15 minutes. The guides used in the analysis included H1P (D166 (H1F), D167 (H1R)); H2P (D168 (H2F), D169 (H2R)); H3P (D170 (H3F), D171 (H3R)); AE1, H1F, H1R, H2F, and H2R. (see **FIG. 92**). Half reactions were combined with ET-SSB protein (0.5μg/reaction) (or control) and plasmid #70-HAT (see below) template DNA; incubated for 30 minutes at either 37°C; and a SacI-HF plasmid restriction digest was carried out for 30 minutes at 37°C. Proteinase K was added to stop the reaction. The DNA was run on a 1% agarose

gel and stained. The data shows that single guides show detectable cleavage only in presence of ET SSB (**FIG. 93** and **FIG. 94**).

Example 36. Single strand DNA cleavage mediated by Ago69 Homologues

[0559] Nine Ago69 homologue proteins were expressed and purified, denoted HG1, HG2, HG3, HG4, HG5, HG6, HG7, HG8 and HG9 (**FIG. 97**, **FIG. 98A**, **FIG. 98B**, **FIG. 99**). HG1, HG3, and HG7 appeared to be insoluble. The ability of Ago69 homologues to cleave single strand plasmid DNA was assessed. In one experiment the homologues tested included HG2 (SEQ ID NO: 134), HG4 (SEQ ID NO: 135), and HG5 (SEQ ID NO: 136). The cleavage assay was performed in 1X CutSmart buffer (NEB). The guide DNAs (gDNAs) was separately preloaded with Ago69 or Ago69 homologue by incubation of the gDNA with Ago69 or Ago69 homologue at 37°C for 15 minutes. The guides used in the analysis included H1P (D166 (H1F), D167 (H1R)) and AE1 (see **FIG. 92**). Half reactions were combined with ET-SSB protein (0.5µg/reaction) (or control) and plasmid #70-HAT (as described in Example 34) template DNA; incubated for 30 minutes at either 37°C; and a SacI-HF plasmid restriction digest was carried out for 30 minutes at 37°C. Proteinase K was added to stop the reaction. The DNA was run on a 1% agarose gel and stained. The expected cleavage products of the SacI-HF plasmid digest are 1402 and 1118 bp. The data shows that Ago69 homologues HG2, HG4, and HG5 all show cleavage activity on single strand DNA, and the cleavage efficiency is increased with the inclusion of Sso-SSB (**FIG. 95**).

[0560] In another experiment the homologues tested included HG2 (SEQ ID NO: 134), HG4 (SEQ ID NO: 135), and HG6. The cleavage assay was performed in 1X CutSmart buffer (NEB). The guide DNAs (gDNAs) was separately preloaded with Ago69 or Ago69 homologue by incubation of the gDNA with Ago69 or Ago69 homologue at 37°C for 15 minutes. The guides used in the analysis included AE1. The restriction digests were run without column purification (**FIG. 100**) and with column purification (**FIG. 101**) of the Ago69 homologue. As shown in **FIG. 100** and **FIG. 101**, HG2 and HG4 showed the highest cleavage efficiency of the homologues tested, while HG6 did not appear to show cleavage activity.

Example 37. HPRT Assay

[0561] The HPRT1 gene encodes for an enzyme called hypoxanthine phosphoribosyltransferase 1 which is involved in purine metabolism. Addition of 6-thioguanine (6-TG) to cells harbouring the wild-type HPRT1 will lead to cell death, mediated by the product of 6-TG conversion by

HPRT1. Cells harbouring inactive HPRT1 can no longer convert 6-TG to its toxic metabolite. Conversely, these cells will be resistant to 6-TG (A4882; Sigma Aldrich).

[0562] Hela cells were transfected with the following series of constructs using Turbofectin 8.0 (TF81001; BioCat GmbH) according to manufacturer's instructions: AP109, SPL0390, SPL0398. Importantly, cells were co-transfected with the set of DNA guides shown in Table 19.

Table 19. DNA Guides

Guide ID	Guide Sequence	SEQ ID NO:
D176	5Phos/T*C*A*TGGACTAATTATGGA*C*A*G	145
D177	5Phos/C*T*G*TCCATAATTAGTCCA*T*G*A	146
D178	5Phos/T*A*T*GAAACTTTCTATTAA*A*T*T	147
D179	5Phos/A*A*T*TTAATAGAAAGTTTC*A*T*A	148
D180	5Phos/T*T*T*TTTACTTTTTCTTGT*G*T*T	149
D181	5Phos/A*A*C*ACAAGAAAAAGTAAA*A*A*A	150
D182	5Phos/A*A*T*TCGCGTTAAATTTTT*G*T*T	151
D183	5Phos/A*A*C*AAAAATTTAACGCGA*A*T*T	152

* Phosphorothioate bonds

[0563] Some of these guides (D176-D181) target the HPRT1 gene, whereas other guides (D182-183) were non-targeting controls which served as negative controls. Guides were either included as single guides or as pairs of guides. Importantly, when used as pairs, guides were designed to targeting opposing DNA strands.

[0564] Following transfection, the cells were allowed to rest for 2-3 days. Then, they were seeded in 96 well plates and treated with 5μM 6-TG. Cells were analysed by microscopy at day 4 post drug treatment. Controls in the experiment included Hela cells that were not treated with 6-TG (alive) and Hela cells treated with 6-TG (rounded shape; compromised viability). In addition, HPRT1 was also targeted using Cas9 and the following sgRNA targeting the human HPRT1 gene: CATGGACTAATTATGGACAG (SEQ ID NO: 144). Inactivation of cells using Cas9 and the HPRT1-specific sgRNA lead to 6-TG resistance as expected. The results for the SPL0390 construct are presented in **FIG. 104**. The results for the AP109 construct are presented in **FIG. 105**. The results for the SPL0398 construct are presented in **FIG. 106**.

[0565] The only condition in the experiment conferring resistance to 6-TG is the condition in which Ago69 homolog 4 (HG4) (SPL0390 construct) was combined with the guide pair D178/179 (**FIG. 104**). Ago69 homolog 4 did not induce this survival phenotype when combined

with the single guide (D178), suggesting that the pair of guides is necessary to establish the 6-TG resistance phenotype. Overall, these data suggest that Ago69 homolog 4 can edit the HPRT1 gene in human cells, leading to the establishment of 6-TG-resistant colonies that can be visualized in the microscope.

CLAIMS

What is claimed is:

1. A system comprising:

a. an Argonaute (Ago) polypeptide, or a polynucleic acid encoding the same, wherein said Ago polypeptide is a Clostridia Ago polypeptide, or a functional fragment or functional variant thereof; and

b. a non-naturally occurring guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence.

2. The system of claim 1, wherein the Ago polypeptide is a mesophilic Clostridia Ago polypeptide.

3. The system of any one of claims 1-2, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid.

4. The system of any one of claims 1-3, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 40 °C, 19 °C to about 50 °C, 19 °C to about 60 °C, 19 °C to about 70 °C, 19 °C to about 80 °C, 20 °C to about 40 °C, 20 °C to about 30 °C, 20 °C to about 50 °C, 20 °C to about 60 °C, 20 °C to about 70 °C, 20 °C to about 80 °C, 25 °C to about 40 °C, 25 °C to about 30 °C, or 25 °C to about 50 °C.

5. The system of any one of claims 1-4, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 19 °C, 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, or 40°C.

6. The system of any one of claims 1-5, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 37 °C.

7. The system of any one of claims 1-6, wherein the Ago polypeptide demonstrates a maximal nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 45 °C, 19 °C to about 40 °C, 20 °C to about 45 °C, 25 °C to about 45 °C, 30 °C to about 45 °C, or 30 °C to about 40 °C, as compared to nucleic acid-cleaving activity at a different temperature.

8. The system of any one of claims 3-7, wherein the nucleic acid-cleaving activity of the target polynucleic acid is directed by the guiding polynucleic acid.

9. The system of any one of claims 3-8, wherein the Ago polypeptide demonstrates one, two, three, or four of: single stranded DNA (ssDNA) cleaving activity, double stranded DNA (dsDNA) cleaving activity, single stranded RNA (ssRNA) cleaving activity, or double stranded RNA (dsRNA) cleaving activity.

10. The system of claim 9, wherein the Ago polypeptide demonstrates single stranded DNA (ssDNA) cleaving activity

11. The system of any one of claims 1-10, wherein the target polynucleic acid is a single stranded DNA (ssDNA) sequence, a double stranded DNA (dsDNA) sequence, a single stranded RNA (ssRNA) sequence, or a double stranded RNA (dsRNA) sequence.

12. The system of claim 11, wherein the target polynucleic acid is a single stranded DNA (ssDNA) sequence.

13. The system of any one of claims 1-12, wherein the target polynucleic acid is DNA.

14. The system of claim 13, wherein a region of the target DNA sequence that the Ago polypeptide cleaves is about at least 50%, 60%, 70%, 80%, or 90% deoxyadenosine and deoxythymidine.

15. The system of any one of claims 1-14, wherein said target polynucleic acid comprises a gene sequence.

16. The system of any one of claims 1-15, wherein said Ago polypeptide produces a disruption in said gene sequence when introduced into a cell.

17. The system of claim 16, wherein said disruption comprises a double strand break or a single strand break.

18. The system of any one of claims 1-17, wherein said guiding polynucleic acid is capable of interacting with said Ago polypeptide and directing said Ago polypeptide to said target polynucleic acid.

19. The system of any one of claim 1-18, wherein the guiding polynucleic acid is a guide DNA or a guide RNA.

20. The system of any one of claims 1-19, wherein said guiding polynucleic acid is from about 1 nucleotide to about 30 nucleotides in length.

21. The system of any one of claims 1 to 20, wherein said system comprises a complex, and wherein said complex comprises said Ago polypeptide and said guiding polynucleic acid.

22. The system of any one of claims 1-21, wherein the Ago polypeptide comprises a PIWI-like domain.

23. The system of any one of claims 1-22, wherein the Ago polypeptide comprises a PIWI domain.

24. The system of any one of claims 1-23, wherein the Ago polypeptide comprises a PAZ domain.

25. The system of any one of claims 1-24, wherein the Ago polypeptide comprises a PAZ-like domain.

26. The system of any one of claim 1-25, wherein the Ago polypeptide is an Ago polypeptide, or a functional fragment or a functional variant thereof, from: *Candidatus Comantemales, Clostridiales, Halanaerobiales, Natranaerobiales, Thermoanaerobacterales, or Negativicutes.*

27. The system of any one of claims 1-25, wherein the Ago polypeptide is an Ago polypeptide, or a functional fragment or a functional variant thereof, from: *Caldicoprobacteraceae, Christensenellaceae, Clostridiaceae, Defluviitaleaceae, Eubacteriaceae, Graciibacteraceae, Heliobacteriaceae, Lachnospiraceae, Oscillospiraceae, Peptococcaceae, Peptostreptococcaceae, Ruminococcaceae, Syntrophomonadaceae, Halanaerobiaceae, Halobacteroidaceae, Natranaerobiaceae, Thermoanaerobacteraceae, or Thermodesulfobiaceae.*

28. The system of any one of claims 1-25, wherein the Ago polypeptide is a *Clostridiaceae* Ago polypeptide, or a functional fragment or a functional variant thereof.

29. The system of claim 28, wherein the Ago polypeptide is a *Clostridium, Acetanaerobacterium, Acetivibrio, Acidaminobacter, Alkaliphilus, Anaerobacter, Anaerostipes, Anaerotruncus, Anoxynatronum, Bryantella, Butyricicoccus, Caldanaerocella, Caldisalinibacter, Caloramator, Caloranaerobacter, Caminicella, Candidatus Arthromitus, Cellulosibacter, Coprobacillus, Crassaminicella, Dorea, Ethanologenbacterium, Faecalibacterium, Garciella, Guggenheimella, Hespellia, Linmingia, Natronincola, Oxobacter, Parasporobacterium, Sarcina, Soehngenia, Sporobacter, Subdoligranulum, Tepidibacter, Tepidimicrobium, Thermobrachium, Thermohalobacter, or Tindallia* Ago polypeptide, or a functional fragment or a functional variant thereof.

30. The system of claim 28, wherein the Ago polypeptide is a *Clostridium* Ago polypeptide, or a functional fragment or a functional variant thereof.

31. The system of claim 28, wherein the Ago polypeptide is a *Clostridium absonum, Clostridium aceticum, Clostridium acetireducens, Clostridium acetobutylicum, Clostridium acidisoli, Clostridium aciditolerans, Clostridium acidurici, Clostridium aerotolerans,*

Clostridium aestuarii, *Clostridium akagii*, *Clostridium aldenense*, *Clostridium aldrichii*, *Clostridium algidicarnis*, *Clostridium algidixylanolyticum*, *Clostridium algifaecis*, *Clostridium algoriphilum*, *Clostridium alkalicellulosi*, *Clostridium amazonense*, *Clostridium aminophilum*, *Clostridium aminovalericum*, *Clostridium amygdalinum*, *Clostridium amylolyticum*, *Clostridium arbusti*, *Clostridium arcticum*, *Clostridium argentinense*, *Clostridium asparagiforme*, *Clostridium aurantibutyricum*, *Clostridium baratii*, *Clostridium barkeri*, *Clostridium bartlettii*, *Clostridium beijerinckii*, *Clostridium bifermentans*, *Clostridium bolteae*, *Clostridium bornimense*, *Clostridium botulinum*, *Clostridium bowmanii*, *Clostridium bryantii*, *Clostridium budayi*, *Clostridium butyricum*, *Clostridium cadaveris*, *Clostridium caenicola*, *Clostridium caminithermale*, *Clostridium carboxidivorans*, *Clostridium carnis*, *Clostridium cavendishii*, *Clostridium celatum*, *Clostridium celerecrescens*, *Clostridium cellobioparum*, *Clostridium cellulofermentans*, *Clostridium cellulolyticum*, *Clostridium cellulosi*, *Clostridium cellulovorans*, *Clostridium chartatabidum*, *Clostridium chauvoei*, *Clostridium chromiireducens*, *Clostridium citroniae*, *Clostridium clariflavum*, *Clostridium clostridioforme*, *Clostridium coccoides*, *Clostridium cochlearium*, *Clostridium cocleatum*, *Clostridium colicanis*, *Clostridium colinum*, *Clostridium collagenovorans*, *Clostridium combesii*, *Clostridium cylindrosporum*, *Clostridium difficile*, *Clostridium diolis*, *Clostridium disporicum*, *Clostridium drakei*, *Clostridium durum*, *Clostridium estertheticum*, *Clostridium estertheticum* subsp. *Estertheticum*, *Clostridium estertheticum* subsp. *Laramiense*, *Clostridium fallax*, *Clostridium felsineum*, *Clostridium fervidum*, *Clostridium fimetarium*, *Clostridium formicaceticum*, *Clostridium frigidicarnis*, *Clostridium frigoris*, *Clostridium ganghwense*, *Clostridium gasigenes*, *Clostridium ghonii*, *Clostridium glycolicum*, *Clostridium glycyrrhizinilyticum*, *Clostridium grantii*, *Clostridium guangxiense*, *Clostridium haemolyticum*, *Clostridium halophilum*, *Clostridium hastiforme*, *Clostridium hathewayi*, *Clostridium herbivorans*, *Clostridium hiranonis*, *Clostridium histolyticum*, *Clostridium homopropionicum*, *Clostridium huakuii*, *Clostridium hungatei*, *Clostridium hydrogeniformans*, *Clostridium hydroxybenzoicum*, *Clostridium hylemonae*, *Clostridium indolis*, *Clostridium innocuum*, *Clostridium intestinale*, *Clostridium irregulare*, *Clostridium isatidis*, *Clostridium jeddahense*, *Clostridium jejuense*, *Clostridium josui*, *Clostridium kluyveri*, *Clostridium lactatifermentans*, *Clostridium lacusfryxellense*, *Clostridium laramiense*, *Clostridium lavalense*, *Clostridium lentocellum*,

Clostridium lentoputrescens, Clostridium leptum, Clostridium limosum, Clostridium liquoris, Clostridium litorale, Clostridium lituseburense, Clostridium ljungdahlii, Clostridium lortetii, Clostridium lundense, Clostridium luticellarii, Clostridium magnum, Clostridium malenominatum, Clostridium mangenotii, Clostridium maximum, Clostridium mayombei, Clostridium methoxybenzovorans, Clostridium methylpentosum, Clostridium moniliforme, Clostridium neonatale, Clostridium neopropionicum, Clostridium neuense, Clostridium nexile, Clostridium nitritogenes, Clostridium nitrophenolicum, Clostridium novyi, Clostridium oceanicum, Clostridium orbiscindens, Clostridium oroticum, Clostridium oryzae, Clostridium oxalicum, Clostridium pabulibutyricum, Clostridium papyrosolvens, Clostridium paradoxum, Clostridium paraperfringens, Clostridium paraputrificum, Clostridium pascui, Clostridium pasteurianum, Clostridium peptidivorans, Clostridium perenne, Clostridium perfringens, Clostridium pfennigii, Clostridium phytofermentans, Clostridium piliforme, Clostridium polyendosporum, Clostridium polysaccharolyticum, Clostridium populeti, Clostridium propionicum, Clostridium proteoclasticum, Clostridium proteolyticum, Clostridium psychrophilum, Clostridium punense, Clostridium puniceum, Clostridium purinilyticum, Clostridium putrefaciens, Clostridium putrificum, Clostridium quercicolum, Clostridium quinii, Clostridium ramosum, Clostridium rectum, Clostridium roseum, Clostridium saccharobutylicum, Clostridium saccharogumia, Clostridium saccharolyticum, Clostridium saccharoperbutylacetonicum, Clostridium sardiniense, Clostridium sartagoforme, Clostridium saudiense, Clostridium scatologenes, Clostridium schirmacherense, Clostridium scindens, Clostridium senegalense, Clostridium septicum, Clostridium sordellii, Clostridium sphenoides, Clostridium spiroforme, Clostridium sporogenes, Clostridium sporosphaeroides, Clostridium stercorarium, Clostridium stercorarium subsp. *Leptospartum, Clostridium stercorarium* subsp. *Stercorarium, Clostridium stercorarium* subsp. *Thermolacticum, Clostridium sticklandii, Clostridium straminisolvens, Clostridium subterminale, Clostridium sufflavum, Clostridium sulfidigenes, Clostridium swellfunianum, Clostridium symbiosum, Clostridium tarantellae, Clostridium tagluense, Clostridium tepidiprofundi, Clostridium tepidum, Clostridium termitidis, Clostridium tertium, Clostridium tetani, Clostridium tetanomorphum, Clostridium thermaceticum, Clostridium thermautotrophicum, Clostridium thermoalcaliphilum, Clostridium thermobutyricum, Clostridium thermocellum, Clostridium thermocopriae,*

Clostridium thermohydrosulfuricum, Clostridium thermolacticum, Clostridium thermopalmarium, Clostridium thermopapyrolyticum, Clostridium thermosaccharolyticum, Clostridium thermosuccinogenes, Clostridium thermosulfurigenes, Clostridium thiosulfatireducens, Clostridium tyrobutyricum, Clostridium uliginosum, Clostridium ultunense, Clostridium ventriculi, Clostridium villosum, Clostridium vincentii, Clostridium viride, Clostridium vulturis, and Clostridium xylanolyticum, or *Clostridium xylanovorans* Ago polypeptide, or a functional fragment or a functional variant thereof.

32. The system of claim 31, wherein the Ago polypeptide is a *Clostridium perfringens*, *Clostridium butyricum*, *Clostridium saudiense*, or *Clostridium disporicum* Ago polypeptide, or a functional fragment or a functional variant thereof.

33. The system of any one of claims 1-32, wherein said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 1-3 or 134-136.

34. The system of any one of claims 1-33, wherein said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity with one of SEQ ID NOs: 11-14 or 137-139.

35. The system of any one of claims 1-34, wherein said system comprises a nucleic acid unwinding polypeptide or a polynucleic acid encoding the same.

36. The system of claim 35, wherein said nucleic acid unwinding polypeptide is a helicase, a single strand DNA binding (SSB) protein, or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

37. The system of claim 36, wherein said nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide.

38. The system of claim 37, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35.

39. The system of claim 37 or 38, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49.

40. The system of claim 38, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22.

41. The system of claim 39, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

42. The system of claim 36, wherein said nucleic acid unwinding polypeptide is a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

43. The system of claim 42, wherein said Cas protein domain is a catalytically dead Cas polypeptide.

44. The system of any one of claims 1-43, wherein said Ago polypeptide is fused either directly or indirectly to a nuclear localization signal (NLS).

45. The system of any one of claims 35-44, wherein said nucleic acid unwinding polypeptide is fused either directly or indirectly to a NLS.

46. The system of any one of claims 35-45, wherein said Ago polypeptide and said nucleic acid unwinding polypeptide are fused either directly or indirectly.

47. The system of claim 46, wherein said Argonaute polypeptide and said nucleic acid unwinding polypeptide are fused and a NLS is in between said Ago polypeptide and said nucleic acid unwinding polypeptide.

48. The system of any one of claims 1-47, wherein said Ago polypeptide is encoded by a gene located in an adjacent operon to at least one of a gene involved in defense, stress response, gene editing, CRISPR, DNA replication, DNA recombination, DNA repair, and transcription.

49. The system of any one of claims 1-48, wherein said system comprises one or more recombinant expression vectors.

50. The system of claim 49, wherein said one or more recombinant expression vectors comprise an adeno-associated virus vector, a plasmid vector, a retroviral vector, a lentiviral vector, an adenovirus vectors, a poxvirus vectors, a herpesvirus vector, or a split-intron vector.

51. The system of any one of claims 1-50, wherein said Ago polypeptide, or functional fragment or variant thereof, comprises a DEDX motif sequence.

52. The system of claim 51, wherein said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

53. An ex vivo cell comprising the system of any one of claims 1-52.

54. The ex vivo cell of claim 53, wherein the cell is a human cell.

55. The ex vivo cell of claim 54, wherein the cell is an immune cell, a stem cell, or a germ cell.

56. A recombinant expression vector encoding said system of any one of claims 1-52.

57. A pharmaceutical composition comprising the system of any one of claims 1-52, and at least one of: an excipient, a diluent, or a carrier.

58. The pharmaceutical composition of claim 57, wherein said pharmaceutical composition is in a form of intravenous, subcutaneous, or intramuscular administration formulation.

59. A kit comprising: (a) the system of any one of claims 1-52; and (b) instructions for use thereof, and optionally (c) a container.

60. A polypeptide construct, said construct comprising a mesophilic Clostridia Ago (C-Ago) polypeptide sequence, or a functional fragment or a functional variant thereof, wherein said C-Ago polypeptide sequence cleaves a nucleic acid in a target polynucleic acid sequence at a mesophilic temperature, wherein said target polynucleic acid sequence is bound by a non-naturally occurring guide polynucleic acid sequence.

61. The polypeptide construct of claim 60, wherein said C-Ago polypeptide sequence or functional fragment or variant thereof comprises a DEDX motif sequence.

62. The polypeptide construct of claim 61, wherein said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said C-Ago polypeptide as compared to a corresponding C-Ago polypeptide without said mutation in said DEDX motif sequence.

63. A nucleic acid molecule encoding said polypeptide construct of any one of claims 60-62.

64. A recombinant fusion polypeptide, said fusion polypeptide comprising: (a) an Argonaute (Ago) polypeptide, wherein said Ago polypeptide is a Clostridia Ago (C-Ago) polypeptide; and (b) a nucleic acid unwinding polypeptide.

65. The recombinant fusion polypeptide of claim 64, wherein the nucleic acid unwinding polypeptide comprises a helicase, a single strand DNA binding protein (SSB) polypeptide, or a Cas protein domain.

66. The recombinant fusion polypeptide of claim 65, wherein the nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide.

67. The recombinant fusion polypeptide of claim 66, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35.

68. The recombinant fusion polypeptide of claim 66 or 67, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49.

69. The recombinant fusion polypeptide of claim 67, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22.

70. The recombinant fusion polypeptide of claim 68, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

71. The recombinant fusion polypeptide of claim 65, wherein said nucleic acid unwinding polypeptide is a Cas protein domain.

72. The recombinant fusion polypeptide of claim 71, wherein said Cas protein domain is a catalytically dead Cas polypeptide.

73. The recombinant fusion polypeptide of any one of claims 64-72, wherein said fusion polypeptide comprises at least one nuclear localization signal (NLS) polypeptide.

74. The recombinant fusion polypeptide of any one of claims 64-73, wherein said fusion polypeptide comprises at least two, three, or four NLSs polypeptides.

75. The recombinant fusion polypeptide of any one of claims 64-74, wherein said fusion polypeptide comprises a nuclear localization signal between said nucleic acid unwinding polypeptide and said C-Ago.

76. The recombinant fusion polypeptide of any one of claims 64-75, wherein said C-Ago polypeptide comprises a DEDX motif sequence.

77. The recombinant fusion polypeptide of claim 76, wherein said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said C-Ago polypeptide

as compared to a corresponding C-Ago polypeptide without said mutation in said DEDX motif sequence.

78. A nucleic acid encoding said recombinant fusion polypeptide of any one of claims 64-77.

79. A method of modifying a target polynucleic acid, said method comprising:

a. introducing into a cell the system of any one of claims 1-52; or a polypeptide construct of any one of claims 60-62; or a recombinant fusion polypeptide of any one of claims 64-77 and a non-naturally occurring guiding polynucleic acid that is complementary to said target polynucleic acid; and
b. modifying said target polynucleic acid.

80. A method of treating a disease or disorder in a subject in need thereof, said method comprising administering to the subject:

a. the system of any one of claims 1-52,
b. the polypeptide construct of any one of claims 60-62,
c. the recombinant fusion polypeptide of any one of claims 64-77,
d. the cell of any one of claims 53-55,
e. the vector of claim 56, or
f. the pharmaceutical composition of any one of claims 57-58.

81. The method of claim 80, wherein said disease is cancer, an autoimmune disease, a genetic disease, or an infection.

82. The method of claim 81, wherein said disease is cancer.

83. A system comprising:

a. a mesophilic Argonaute (Ago) polypeptide, or a polynucleic acid encoding the same, or a functional fragment or variant thereof; and
b. an exogenous non-naturally occurring guiding polynucleic acid comprising a sequence that is complementary to a target polynucleic acid sequence.

84. The system of claim 83, wherein said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%sequence identity with one of SEQ ID NOs: 4-10 or 134-136.

85. The system of claim 83, wherein said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 15-21.

86. The system of any one of claims 83-85, wherein said Ago polypeptide comprises a DEDX motif sequence.

87. The system of claim 86, wherein said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

88. The system of any one of claims 83-87, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid.

89. The system of any one of claims 83-88, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 40 °C, 19 °C to about 50 °C, 19 °C to about 60 °C, 19 °C to about 70 °C, 19 °C to about 80 °C, 20 °C to about 40 °C, 20 °C to about 30 °C, 20 °C to about 50 °C, 20 °C to about 60 °C, 20 °C to about 70 °C, 20 °C to about 80 °C, 25 °C to about 40 °C, 25 °C to about 30 °C, or 25 °C to about 50 °C.

90. The system of any one of claims 83-89, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 19 °C, 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, or 40°C.

91. The system of any one of claims 83-90, wherein the Ago polypeptide demonstrates nucleic acid-cleaving activity of the target polynucleic acid at about 37 °C.

92. The system of any one of claims 83-91, wherein the Ago polypeptide demonstrates a maximal nucleic acid-cleaving activity of the target polynucleic acid in a range of temperature of from about 19 °C to about 45 °C, 19 °C to about 40 °C, 20 °C to about 45 °C, 25 °C to about 45 °C, 30 °C to about 45 °C, or 30 °C to about 40 °C, as compared to nucleic acid-cleaving activity at a different temperature.

93. The system of any one of claims 88-92, wherein the nucleic acid-cleaving activity of the target polynucleic acid is directed by the guiding polynucleic acid.

94. The system of any one of claims 83-93, wherein the Ago polypeptide demonstrates one, two, three, or four of: single stranded DNA (ssDNA) cleaving activity, double stranded DNA (dsDNA) cleaving activity, single stranded RNA (ssRNA) cleaving activity, or double stranded RNA (dsRNA) cleaving activity.

95. The system of claim 94, wherein the Ago polypeptide demonstrates single stranded DNA (ssDNA) cleaving activity.

96. The system of any one of claims 83-95, wherein the target polynucleic acid is a single stranded DNA (ssDNA) sequence, a double stranded DNA (dsDNA) sequence, a single stranded RNA (ssRNA) sequence, or a double stranded RNA (dsRNA) sequence.

97. The system of claim 96, wherein the target polynucleic acid is a single stranded DNA (ssDNA) sequence.

98. The system of any one of claims 83-97, wherein the target polynucleic acid is DNA.

99. The system of claim 98, wherein a region of the target DNA sequence that the C-Ago polypeptide cleaves is about at least 50%, 60%, 70%, 80%, or 90% deoxyadenosine and deoxythymidine.

100. The system of any one of claims 83-99, wherein said target polynucleic acid comprises a gene sequence.

101. The system of claim 100, wherein said Ago polypeptide sequence produces a disruption in said gene sequence when introduced into a cell.

102. The system of claim 101, wherein said disruption comprises a double strand break or a single strand break.

103. The system of any one of claims 83-102, wherein said guiding polynucleic acid is capable of interacting with said Ago polypeptide and directing said Ago polypeptide to said target polynucleic acid.

104. The system of any one of claim 83-103, wherein the guiding polynucleic acid is a guide DNA or a guide RNA.

105. The system of any one of claims 83-104, wherein said guiding polynucleic acid is from about 1 nucleotide to about 30 nucleotides in length.

106. The system of any one of claims 83-105, wherein said system comprises a complex, and wherein said complex comprises said Ago polypeptide and said guiding polynucleic acid.

107. The system of any one of claims 83-106, wherein the Ago polypeptide comprises a PIWI-like domain.

108. The system of any one of claims 83-107, wherein the Ago polypeptide comprises a PIWI domain.

109. The system of any one of claims 83-108, wherein the Ago polypeptide comprises a PAZ domain.

110. The system of any one of claims 83-109, wherein the Ago polypeptide comprises a PAZ-like domain.

111. The system of any one of claims 83-110, wherein said system comprises a nucleic acid unwinding polypeptide or a polynucleic acid encoding the same.

112. The system of claim 111, wherein said nucleic acid unwinding polypeptide is a helicase, a single strand DNA binding (SSB) protein, or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

113. The system of claim 112, wherein said nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide.

114. The system of claim 113, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35.

115. The system of any one of claims 113-114, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49.

116. The system of claim 114, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22.

117. The system of claim 115, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

118. The system of claim 112, wherein said nucleic acid unwinding polypeptide is a Cas protein domain.

119. The system of claim 118, wherein said Cas protein domain is a catalytically dead Cas polypeptide.

120. The system of any one of claims 83-119, wherein said Ago polypeptide is fused either directly or indirectly to a NLS.

121. The system of any one of claims 83-120, wherein said nucleic acid unwinding polypeptide is fused either directly or indirectly to a NLS.

122. The system of any one of claims 83-121, wherein said Ago polypeptide and said nucleic acid unwinding polypeptide are fused either directly or indirectly.

123. The system of claim 122, wherein said Ago polypeptide and said nucleic acid unwinding polypeptide are fused and a NLS is in between said Ago polypeptide and said nucleic acid unwinding polypeptide.

124. The system of any one of claims 83-123, wherein said Ago polypeptide is encoded by a gene located in an adjacent operon to at least one of a gene involved in defense, stress response, gene editing, CRISPR, DNA replication, DNA recombination, DNA repair, and transcription.

125. The system of any one of claims 83-124, wherein said system comprises one or more recombinant expression vectors.

126. The system of claim 125, wherein said one or more recombinant expression vectors comprise an adeno-associated virus vector, a plasmid vector, a retroviral vector, a lentiviral vector, an adenovirus vectors, a poxvirus vectors, a herpesvirus vector, or a split-intron vector.

127. An ex vivo cell comprising the system of any one of claims 83-126.

128. The ex vivo cell of claim 127, wherein the cell is a human cell.

129. The ex vivo cell of claim 128, wherein the cell is an immune cell, a stem cell, or a germ cell.

130. A recombinant expression vector encoding said system of any one of claims 83-126.

131. A pharmaceutical composition comprising the system of any one of claims 83-126, and at least one of: an excipient, a diluent, or a carrier.

132. The pharmaceutical composition of claim 131, wherein said pharmaceutical composition is in a form of intravenous, subcutaneous, or intramuscular administration formulation.

133. A kit comprising: (a) the system of any one of claims 83-126; and (b) instructions for use thereof, and optionally (c) a container.

134. A polypeptide construct, said construct comprising a mesophilic Ago polypeptide sequence, or a functional fragment or a functional variant thereof, wherein said Ago polypeptide sequence cleaves a nucleic acid in a target polynucleic acid sequence at a mesophilic temperature, wherein said target polynucleic acid sequence is bound by a non-naturally occurring guide polynucleic acid sequence.

135. The polypeptide construct of claim 134, wherein said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 4-10.

136. The polypeptide construct of claim 134, wherein said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 15-21.

137. The polypeptide construct of any one of claims 134-136, wherein said Ago polypeptide comprises a DEDX motif sequence.

138. The polypeptide construct of claim 137, wherein said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

139. A nucleic acid sequence encoding the polypeptide construct of any one of claims 134-138.

140. A recombinant fusion polypeptide, said fusion polypeptide comprising:

a. a mesophilic Argonaute (Ago) polypeptide; and

b. a nucleic acid unwinding polypeptide.

141. The recombinant fusion polypeptide of claim 140, wherein the nucleic acid unwinding polypeptide comprises a helicase, a single strand DNA binding protein (SSB), or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated (Cas) protein domain.

142. The recombinant fusion polypeptide of claim 140 or 141, wherein said Ago polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 4-10.

143. The recombinant fusion polypeptide of claim 140 or 141, wherein said Ago polypeptide is encoded by a polynucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more sequence identity with one of SEQ ID NOs: 15-21.

144. The recombinant fusion polypeptide of any one of claims 140-143, wherein the nucleic acid unwinding polypeptide is a single strand DNA binding protein (SSB) polypeptide.

145. The recombinant fusion polypeptide of claim 144, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 22-35.

146. The recombinant fusion polypeptide of claim 144 or 145, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with one of SEQ ID NOs: 36-49.

147. The recombinant fusion polypeptide of claim 145, wherein said SSB polypeptide comprises an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 22.

148. The recombinant fusion polypeptide of claim 146, wherein said SSB polypeptide is encoded by a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 36.

149. The recombinant fusion polypeptide of claim 141, wherein said nucleic acid unwinding polypeptide is a Cas protein domain.

150. The recombinant fusion polypeptide of claim 149, wherein said Cas protein domain is a catalytically dead Cas polypeptide.

151. The recombinant fusion polypeptide of any one of claims 140-150, wherein said fusion polypeptide comprises at least one nuclear localization signal (NLS) polypeptide.

152. The recombinant fusion polypeptide of any one of claims 140-151, wherein said fusion polypeptide comprises at least two, three, or four NLS polypeptides.

153. The recombinant fusion polypeptide of any one of claims 140-152, wherein said fusion polypeptide comprises a NLS between said nucleic acid unwinding polypeptide and said Ago polypeptide.

154. The recombinant fusion polypeptide of any one of claims 140-153, wherein said Ago polypeptide comprises a DEDX motif sequence.

155. The recombinant fusion polypeptide of claim 154, wherein said DEDX motif sequence comprises a mutation, wherein said mutation reduces catalytic activity of said Ago polypeptide as compared to a corresponding Ago polypeptide without said mutation in said DEDX motif sequence.

156. A nucleic acid encoding said recombinant fusion polypeptide of any one of claims 140-155.

157. A method of modifying a target polynucleic acid, said method comprising:

a. introducing into a cell the system of any one of claims 83-126; or
b. the polypeptide construct of any one of claims 134-138; or
c. the recombinant fusion polypeptide of any one of claims 140-155,and

a non-naturally occurring guiding polynucleic acid that is complementary to said target polynucleic acid; and

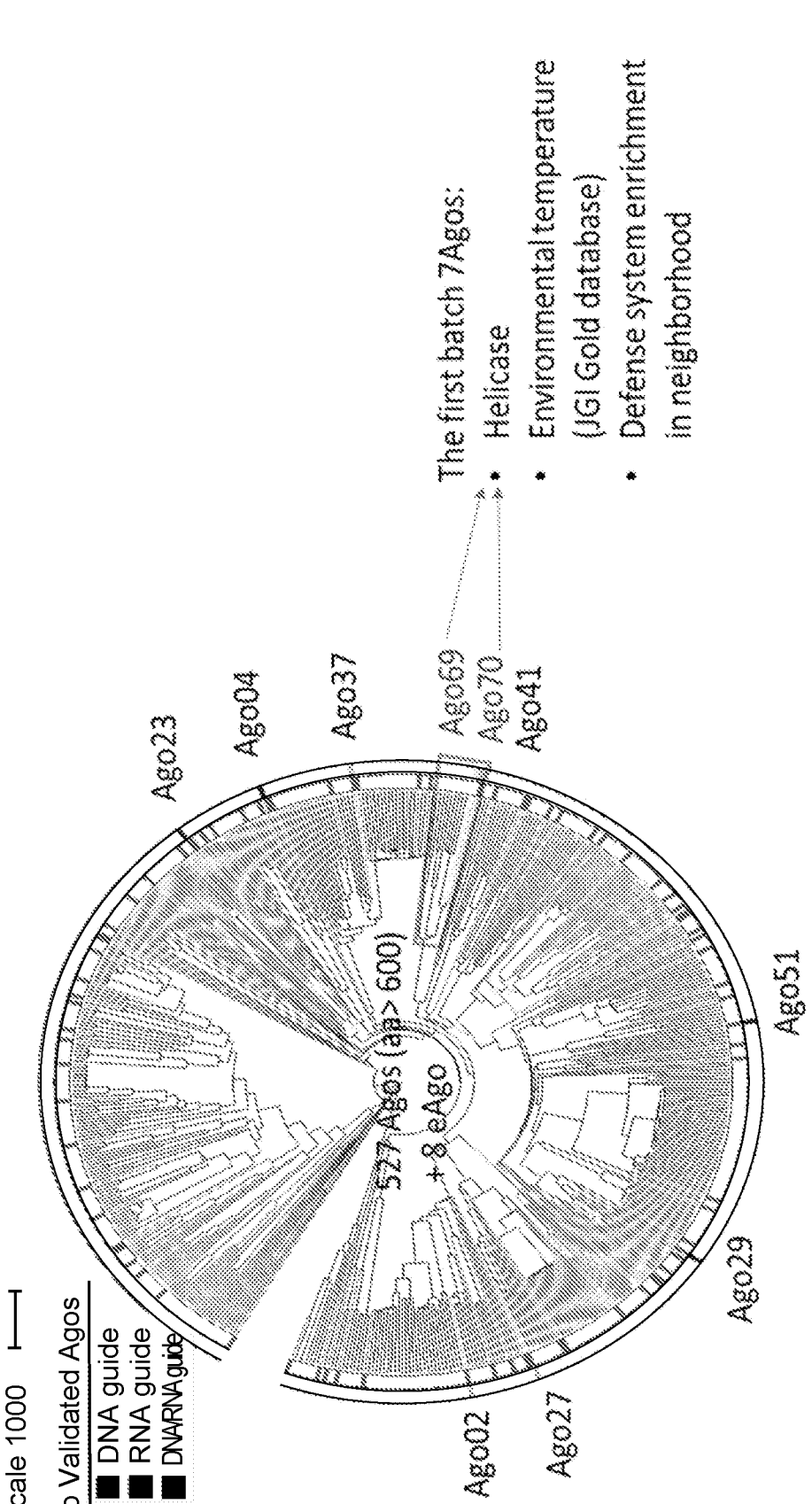
modifying said target polynucleic acid.

158. A method of treating a disease or disorder in a subject in need thereof, said method comprising administering to the subject:

a. the system of any one of claims 83-126,
b. the polypeptide construct of any one of claims 134-138,
c. the recombinant fusion polypeptide of any one of claims 140-155,
d. the cell of any one of claims 127-129,
e. the vector of claim 130, or
f. the pharmaceutical composition of any one of claims 131-132.

159. The method of claim 158, wherein said disease is cancer, an autoimmune disease, a genetic disease, or an infection.

160. The method of claim 159, wherein said disease is cancer.



	Selected Agos	Positive In-Vitro
Branch representatives	80	8/80 (10%)
refined selection	7	2/7 (28.5%)

FIG. 1

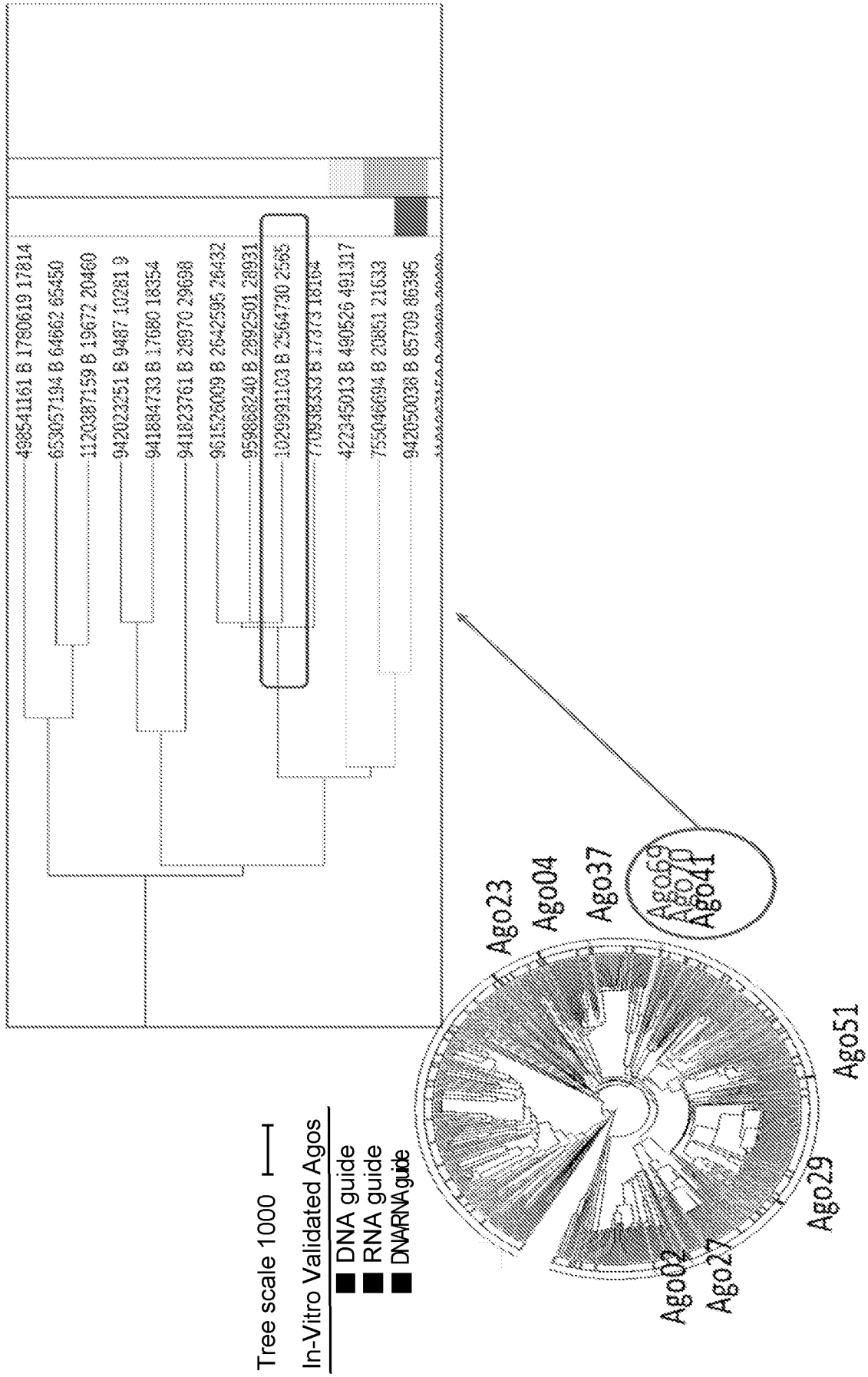

498541161 B 1780619 17814
653057194 B 64962 65450
1120387159 B 19672 20460
942022251 B 9487 10281 9
941884733 B 17680 18354
941823761 B 26970 29698
961526409 B 2642595 26432
959663240 B 2892501 28931
1029991103 B 25647730 25685
770938333 B 17373 18154
422345013 B 480526 491317
755046694 B 26851 21633
942050038 B 85709 86395
Ago23
Ago04
Ago37
Ago69
Ago70
Ago41
Ago51
Ago29
Ago27
Ago02
Tree scale 1000
In-Vitro Validated Agos
DNA guide
RNA guide
DNA/RNA guide

FIG. 2

Id(NCBI)	Organism	Taxonomy
498541161	Thermobrachium celere DSM 8682	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Thermobrachium
653057194	Caloramator sp. ALD01	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Caloramator
1120387159	Caloramator proteoclasticus DSM 10124	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Caloramator
942023251	Fusicatenibacter saccharivorans	Bacteria Firmicutes Clostridia Clostridiales Lachnospiraceae Fusicatenibacter
941884733	Fusicatenibacter saccharivorans	Bacteria Firmicutes Clostridia Clostridiales Lachnospiraceae Fusicatenibacter
941823761	Dorea longicatena	Bacteria Firmicutes Clostridia Clostridiales Lachnospiraceae Dorea
961526009	Clostridium butyricum	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium
959868240	Clostridium butyricum	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium
1029991103	Clostridium butyricum	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium
770938333	Clostridium butyricum CWB1009	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium
422345013(Ago69)	Clostridium perfringens WAL 14572	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium
755046694(Ago70)	Clostridium saudiense	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium
942050038(Ago41)	Clostridium disporicum	Bacteria Firmicutes Clostridia Clostridiales Clostridiaceae Clostridium

FIG. 3

Id(NCBI)	Temperature_Range	Sample_Collection_Site	Sample_Isolation_Comments	Ecosystem_Subtype	Specific_Ecosystem
498541161					
653057194			Etolika lagoon in Western Greece	Unclassified	Unclassified
1120387159	55 C	Thermophile	Mesophilic granular sludge from a lab-scale UASB reactor that was treating whey	Unclassified	Unclassified
942023251					
941884733					
941823761					
961536009			the pit mud of a Chinese liquor factory		
959968240			fecal sample of pig	Digestive system	Fecal
1029991103			probiotics		
770938333					
422345013(Ago69)		Mesophile	Human feces	Digestive system	Large intestine
755046694(Ago70)		25-37 C	Human fecal sample	Digestive system	Large intestine
942050038(Ago41)			rat caecum	Digestive system	Large intestine

FIG. 4

Ago41 (942050038) Taxonomy

NCBI TAXONOMY	
NCBI Organism Name	Clostridium disporicum
NCBI Tax ID	84024
NCBI Superkingdom	Bacteria
NCBI Kingdom	
NCBI Phylum	Firmicutes
NCBI Class	Clostridia
NCBI Order	Clostridiales
NCBI Family	Clostridiaceae
NCBI Genus	Clostridium
NCBI Species	Clostridium disporicum

FIG. 5

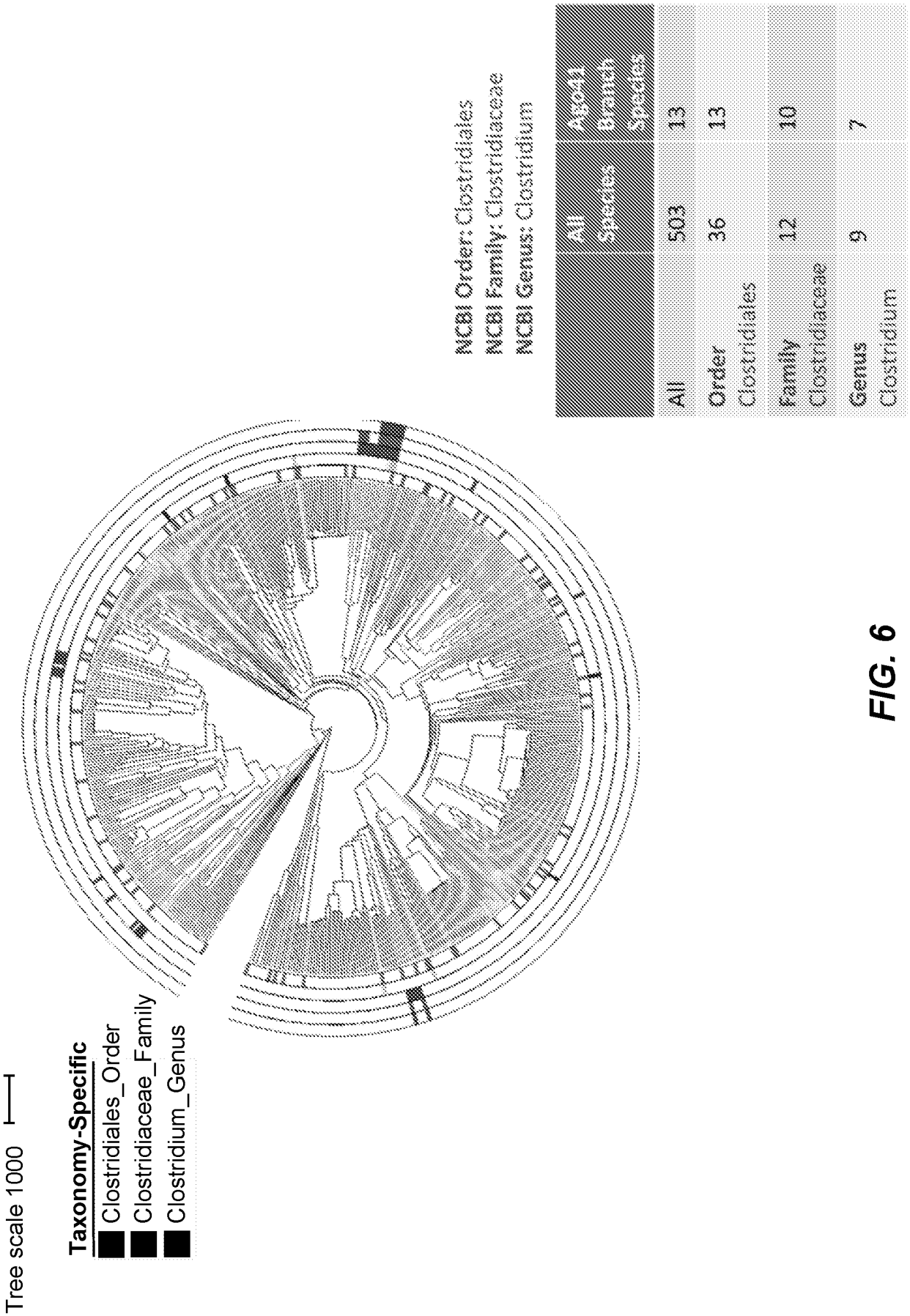

Tree scale 1000
Taxonomy-Specific
Clostridiales_Order
Clostridiaceae_Family
Clostridium_Genus
NCBI Order: Clostridiales
NCBI Family: Clostridiaceae
NCBI Genus: Clostridium
All Species
Ab41 Branch Species
All 503 13
Order Clostridiales 36 13
Family Clostridiaceae 12 10
Genus Clostridium 9 7

FIG. 6

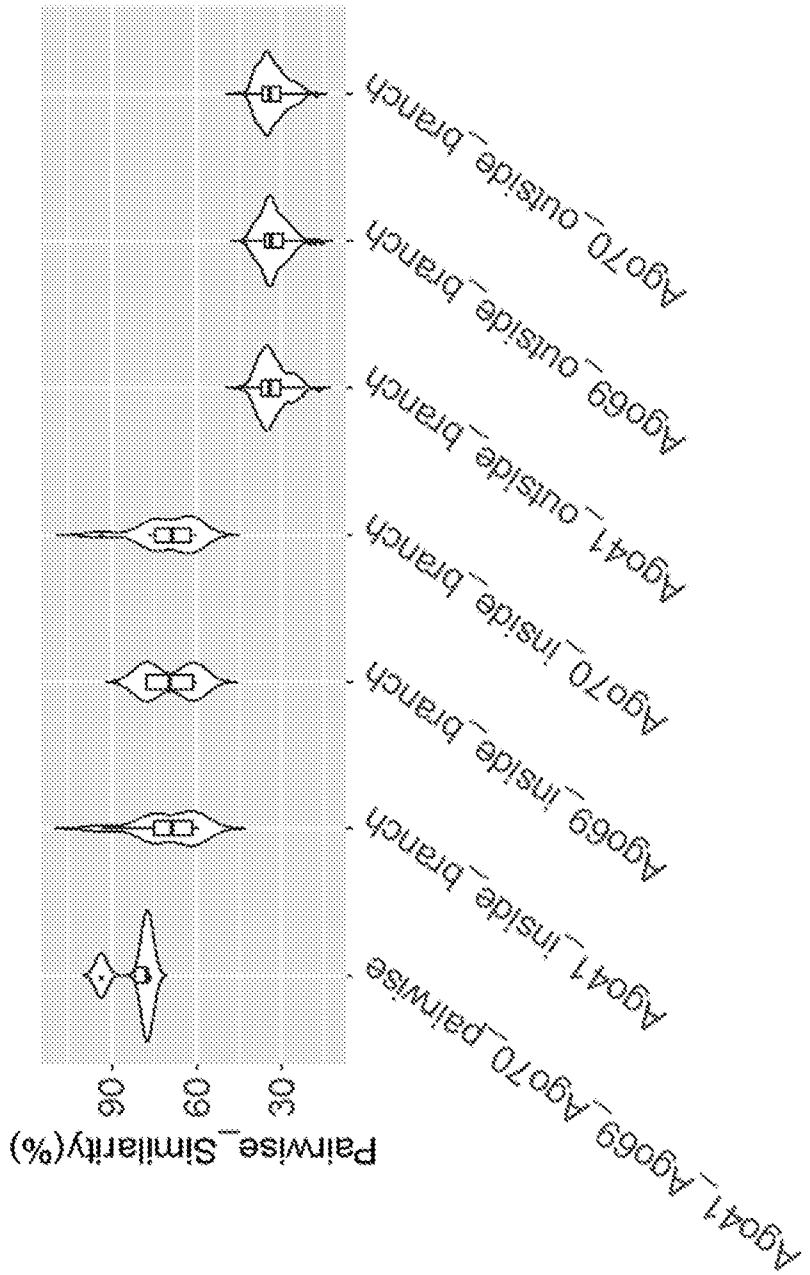
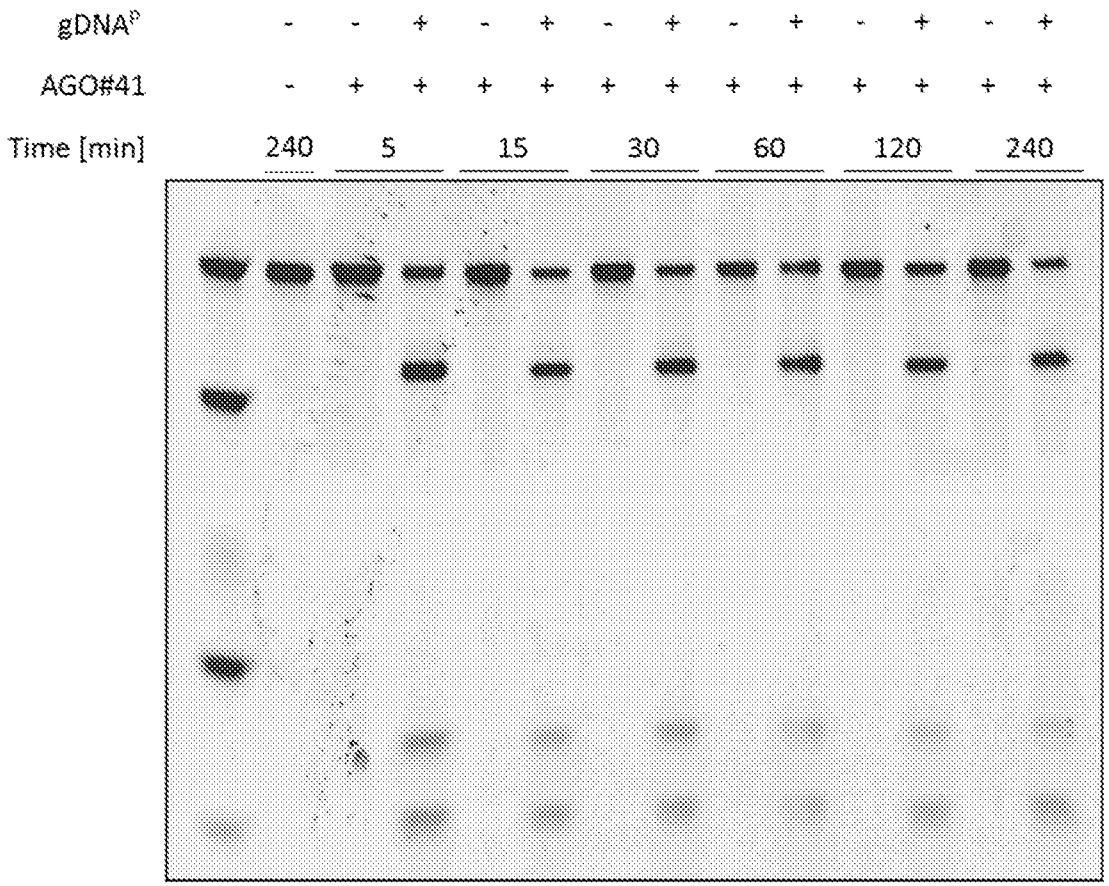

Pairwise_Similarity(%)
90
95
100
Ago41_Ago69_Ago70_pairwise
Ago41_inside_branch
Ago69_inside_branch
Ago70_inside_branch
Ago41_outside_branch
Ago69_outside_branch
Ago70_outside_branch

FIG. 7



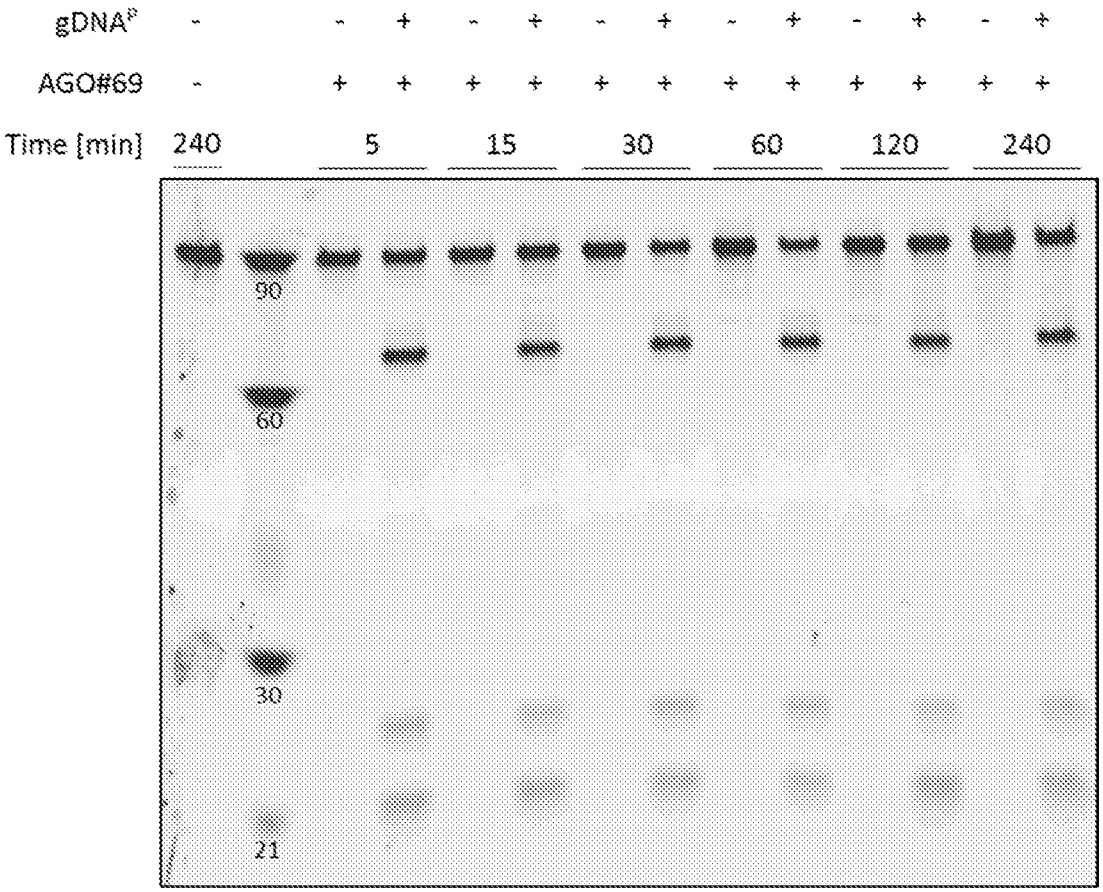
Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA

Template: T1 - 90 nt ssDNA; expected cleavage product: 66 + 24nt

AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

Preloading: gDNA + AGO at 37°C for 15 min

FIG. 8



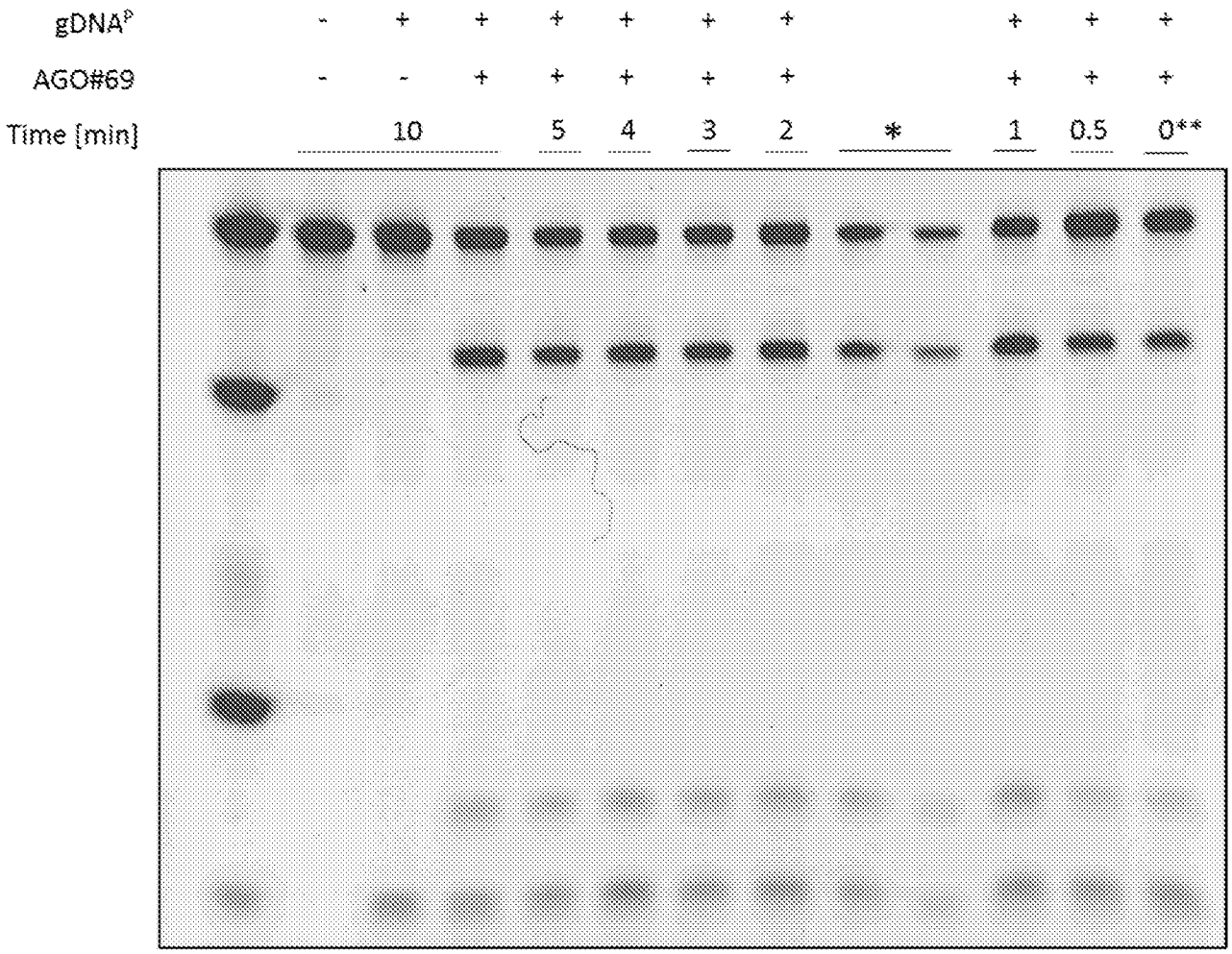
Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA

Template: T1 - 90 nt ssDNA; expected cleavage product: 66 + 24nt

AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

Preloading: gDNA + AGO at 37°C for 15 min

FIG. 9



*...pipetting mistake
**...it takes a couple of seconds to stop the reaction

FIG. 10

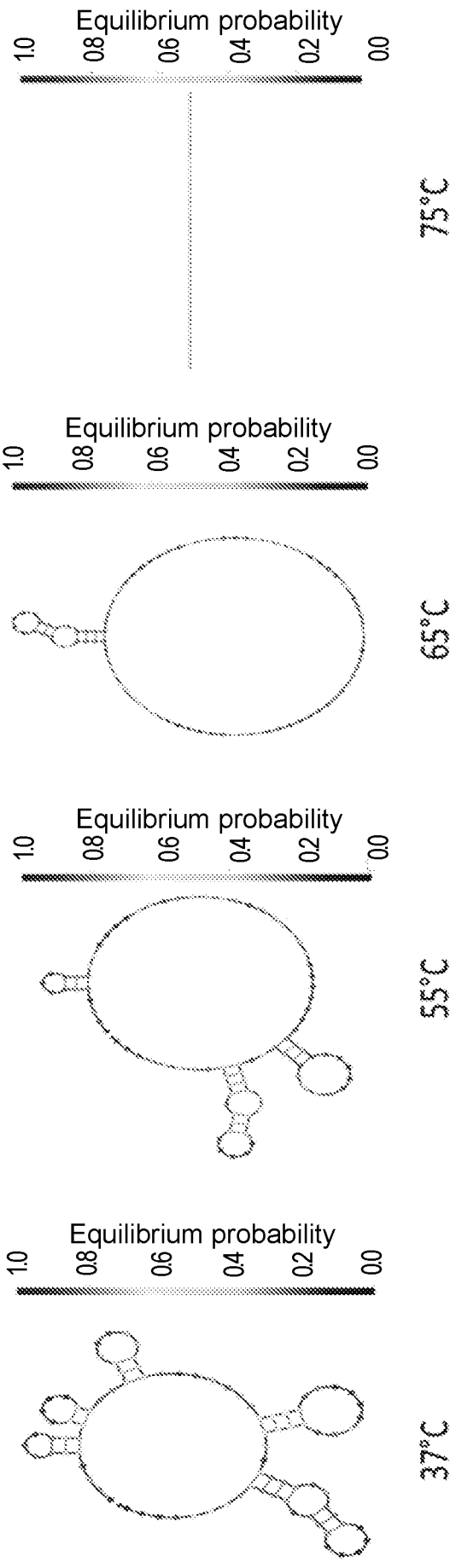

Equilibrium probability
1.0 0.8 0.6 0.4 0.2 0.0
75°C
Equilibrium probability
1.0 0.8 0.6 0.4 0.2 0.0
65°C
Equilibrium probability
1.0 0.8 0.6 0.4 0.2 0.0
55°C
Equilibrium probability
1.0 0.8 0.6 0.4 0.2 0.0
37°C

FIG. 11

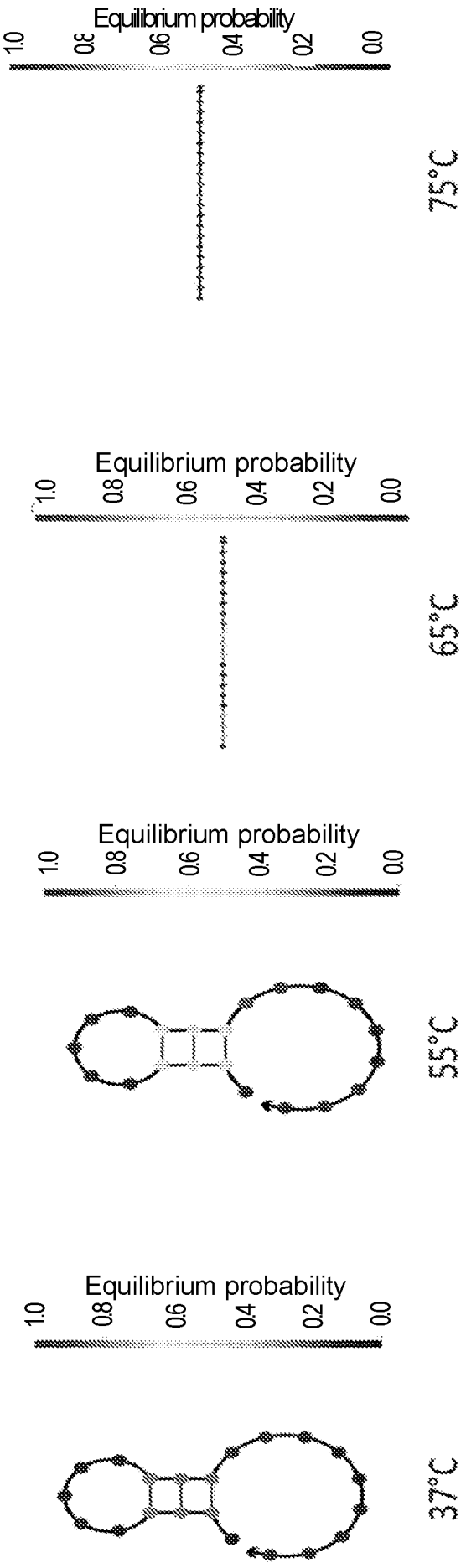
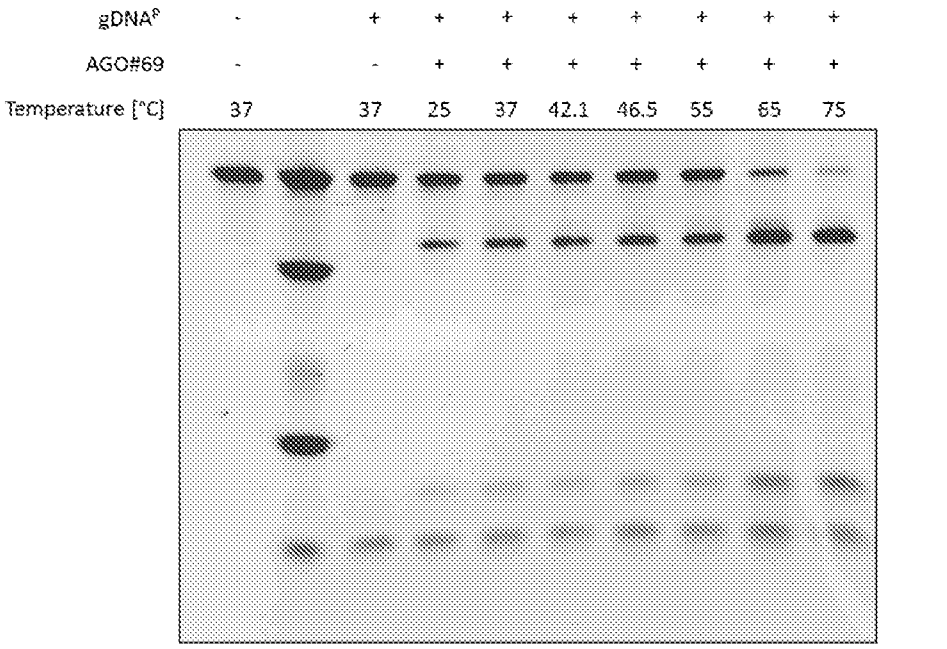

Equilibrium probability
1.0
0.8
0.6
0.4
0.2
0.0
75°C
Equilibrium probability
1.0
0.8
0.6
0.4
0.2
0.0
65°C
Equilibrium probability
1.0
0.8
0.6
0.4
0.2
0.0
55°C
Equilibrium probability
1.0
0.8
0.6
0.4
0.2
0.0
37°C

FIG. 12


gDNA$^{\delta}$ - + + + + + + + +
AGO#69 - - + + + + + + +
Temperature [°C] 37 37 25 37 42.1 46.5 55 65 75
Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA
Template: T1 - 90 nt ssDNA; expected cleavage product: 66 + 24nt
AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

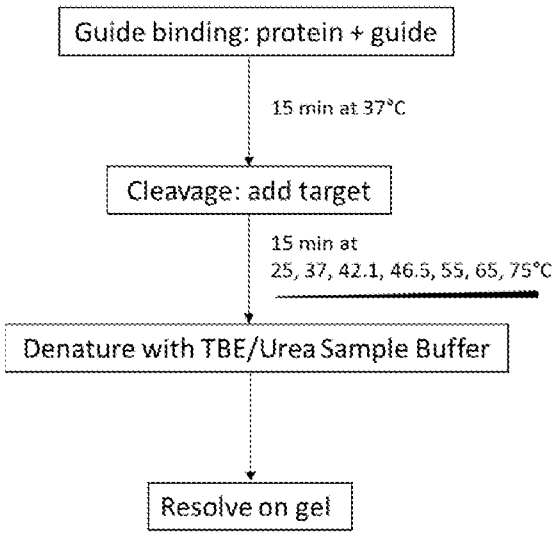
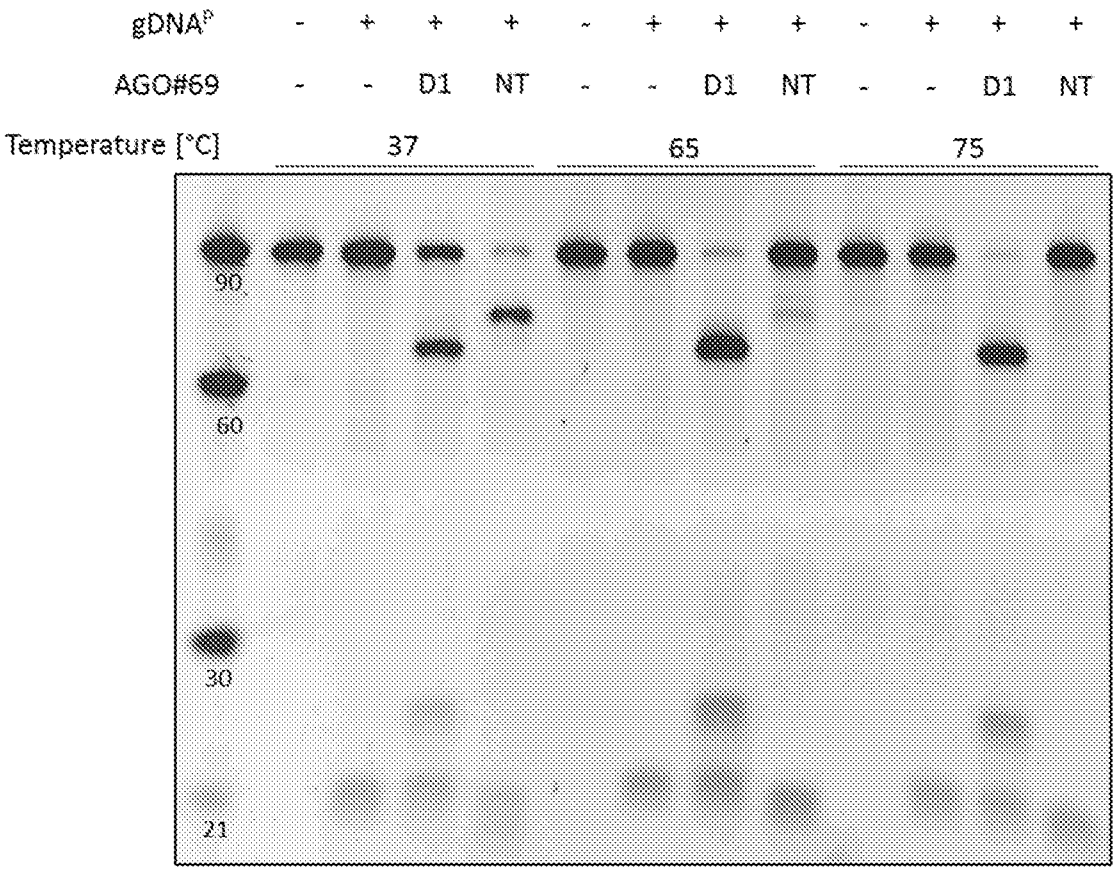

Guide binding: protein + guide
15 min at 37°C
Cleavage: add target
15 min at
25, 37, 42.1, 46.5, 55, 65, 75°C
Denature with TBE/Urea Sample Buffer
Resolve on gel

FIG. 13



Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA
Template: T1 - 90 nt ssDNA; expected cleavage product: 66 + 24nt
AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

FIG. 14

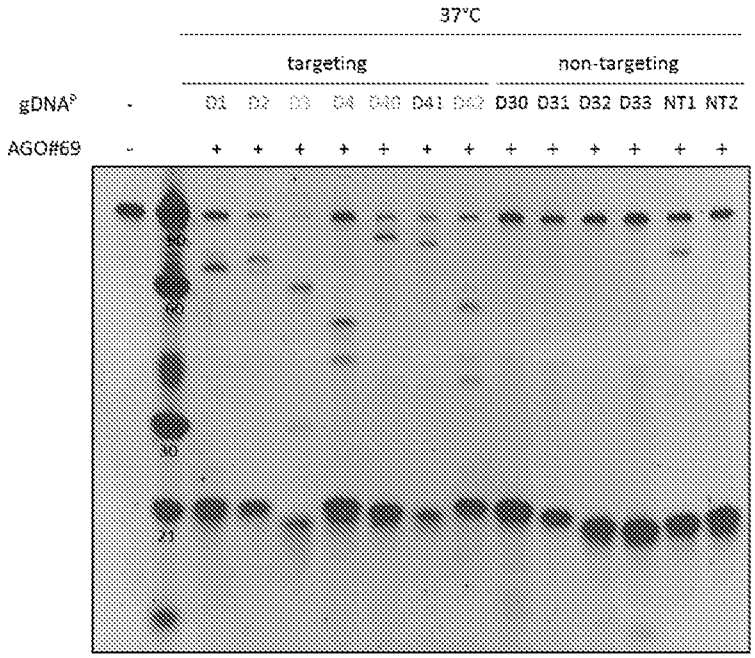
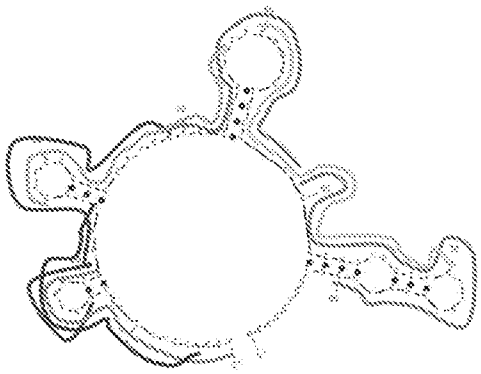
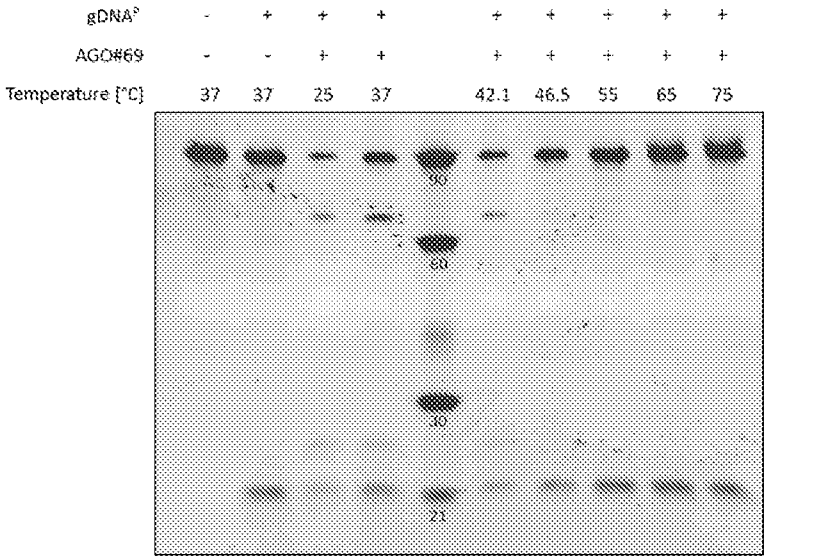


FIG. 15A



Location of guides relative to ssDNA target sequence and secondary structure

FIG. 15B


gDNA°
AGO#69
Temperature [°C]
37 37 25 37 42.1 46.5 55 65 75
90
60
30
21
Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM MnCl2, 1.6 mM b-MeOH, 0.3% BSA
Template: T1 - 90 nt ssDNA; expected cleavage product: 66 + 24nt
AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

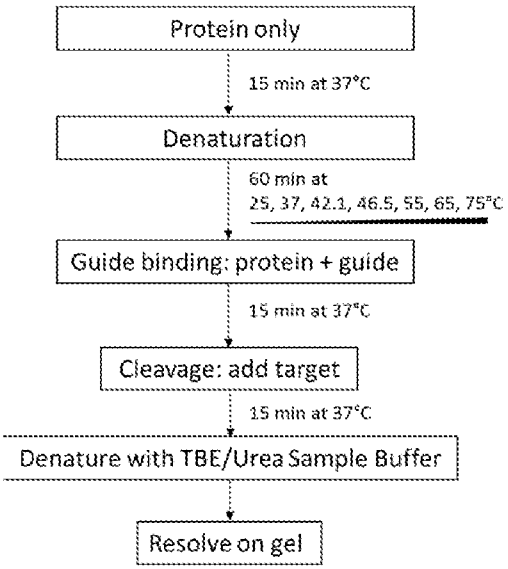
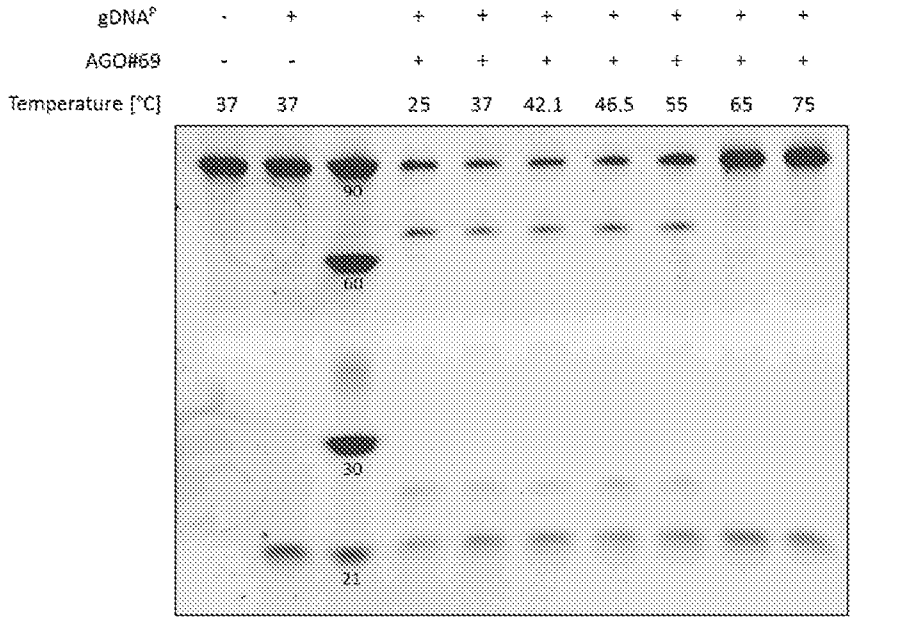

Protein only
15 min at 37°C
Denaturation
60 min at
25, 37, 42.1, 46.5, 55, 65, 75°C
Guide binding: protein + guide
15 min at 37°C
Cleavage: add target
15 min at 37°C
Denature with TBE/Urea Sample Buffer
Resolve on gel

FIG. 16



Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA
Template: 71 - 90 nt ssDNA; expected cleavage product: 66 + 24nt
AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

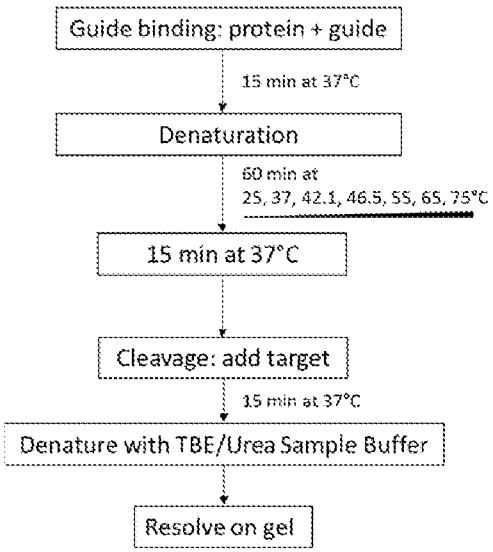


FIG. 17

FIG. 18

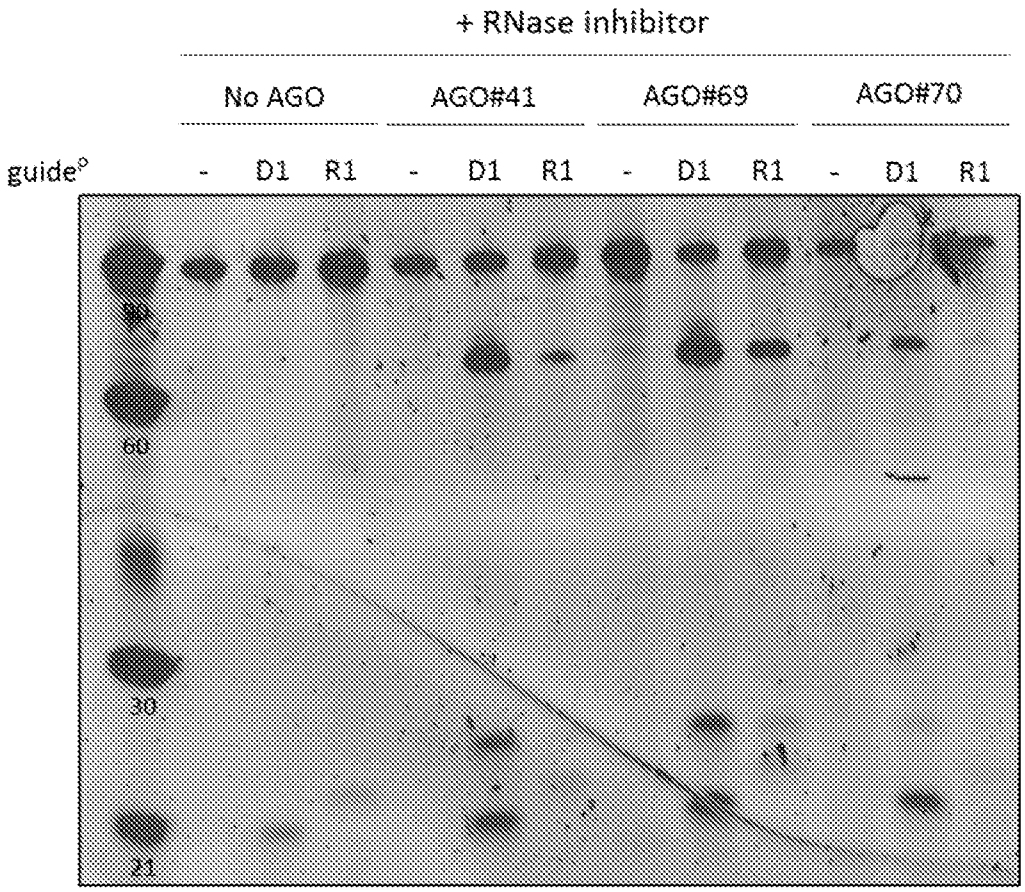
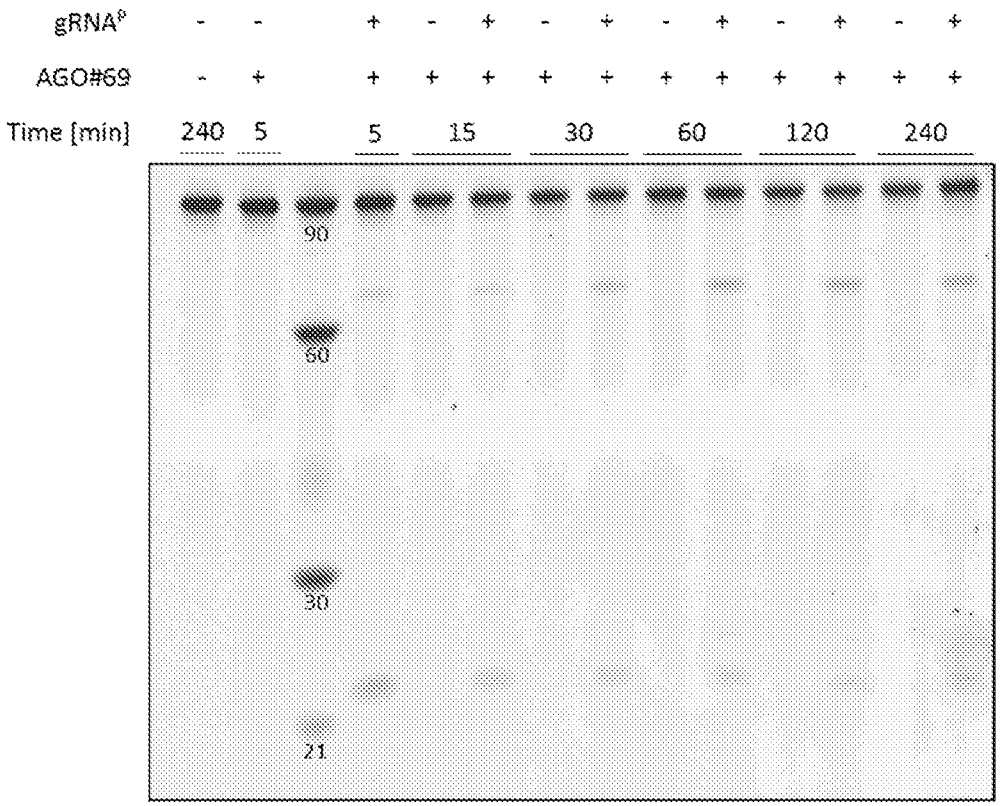

+ RNase inhibitor
No AGO
AGO#41
AGO#69
AGO#70
guide
-
D1
R1
-
D1
R1
-
D1
R1
-
D1
R1
60
30
21

FIG. 19



Reaction Buffer: 20 mM Tris/HCl pH7.5, 125 mM NaCl, 5 mM $MnCl_2$, 1.6 mM b-MeOH, 0.3% BSA, **RNase inhibitor**

Template: T1 - 90 nt ssDNA; expected cleavage product: 66 + 24nt

gRNA: Phosphorothioate bonds on 5' and 3' end

AGO : gDNA : Template = 1 : 1 : 1 (250 nM : 250 nM : 250 nM)

Preloading: gDNA + AGO at 37°C for 15 min

FIG. 20

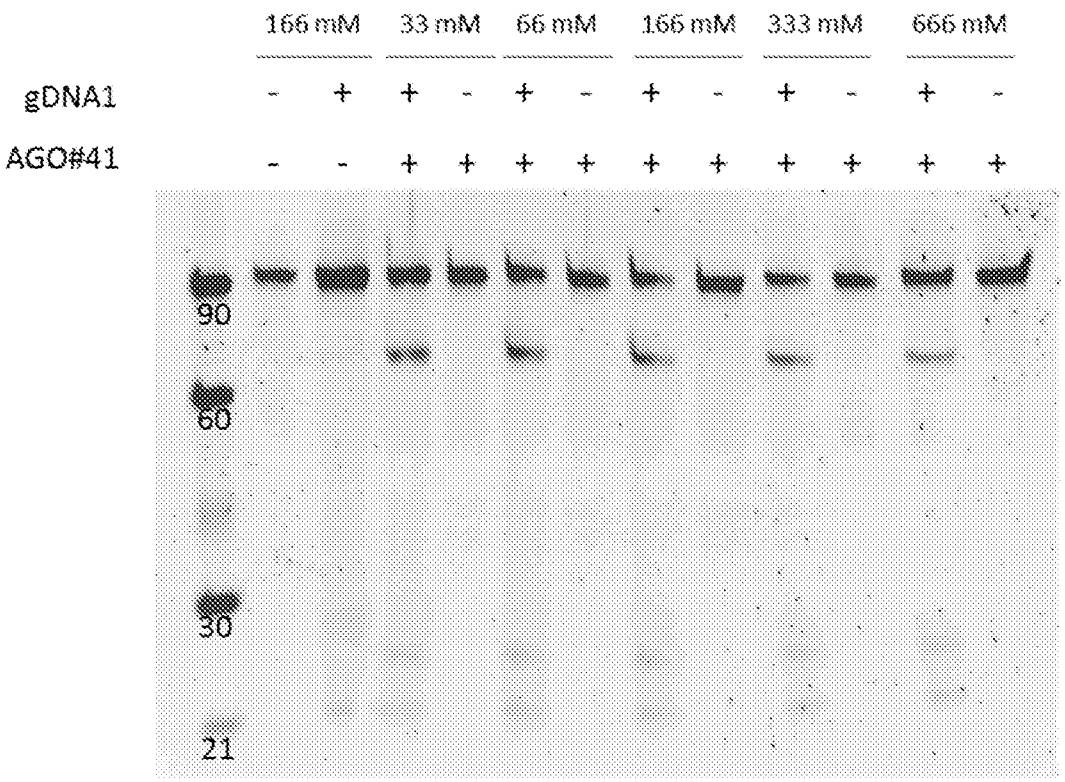


FIG. 21A

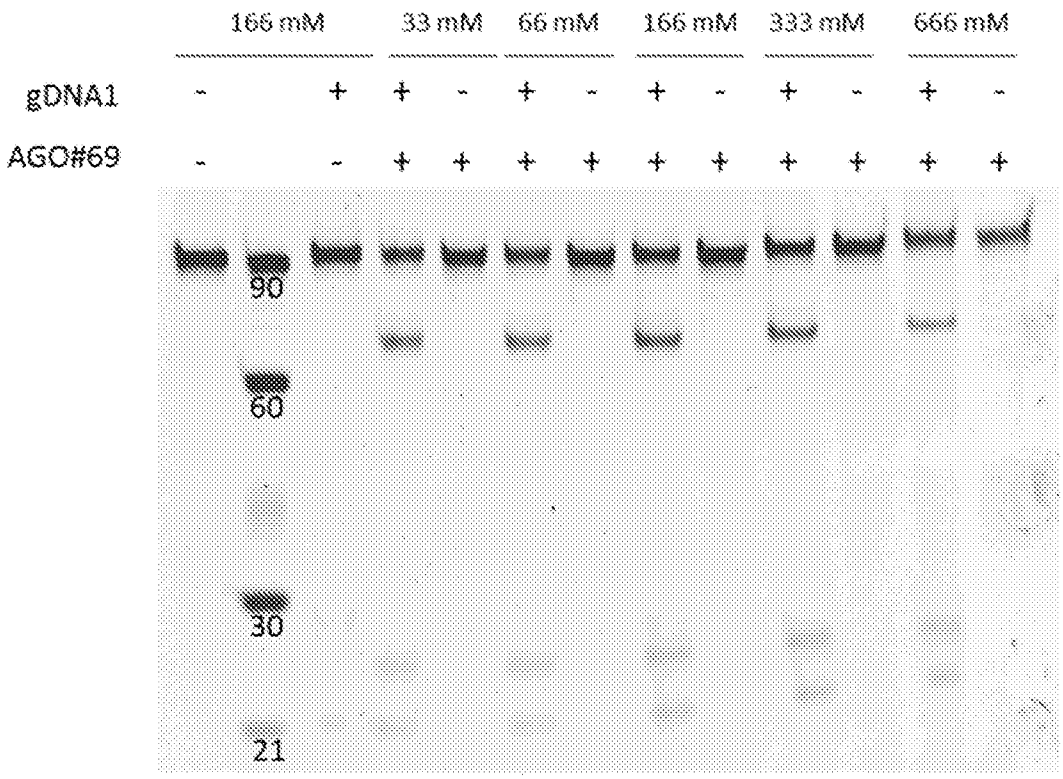


FIG. 21B

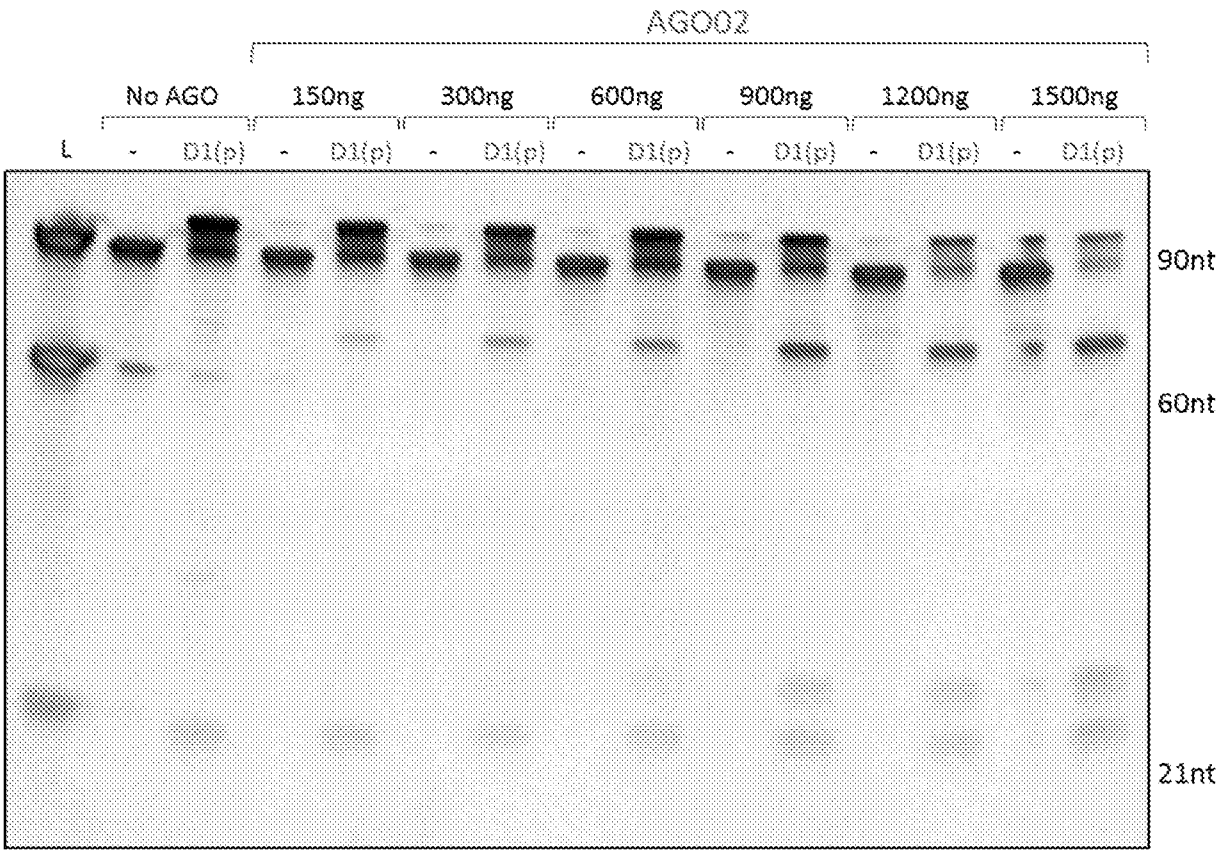

AGO2
No AGO
150ng
300ng
600ng
900ng
1200ng
1500ng
L
-
D1(p)
90nt
60nt
21nt

FIG. 22A

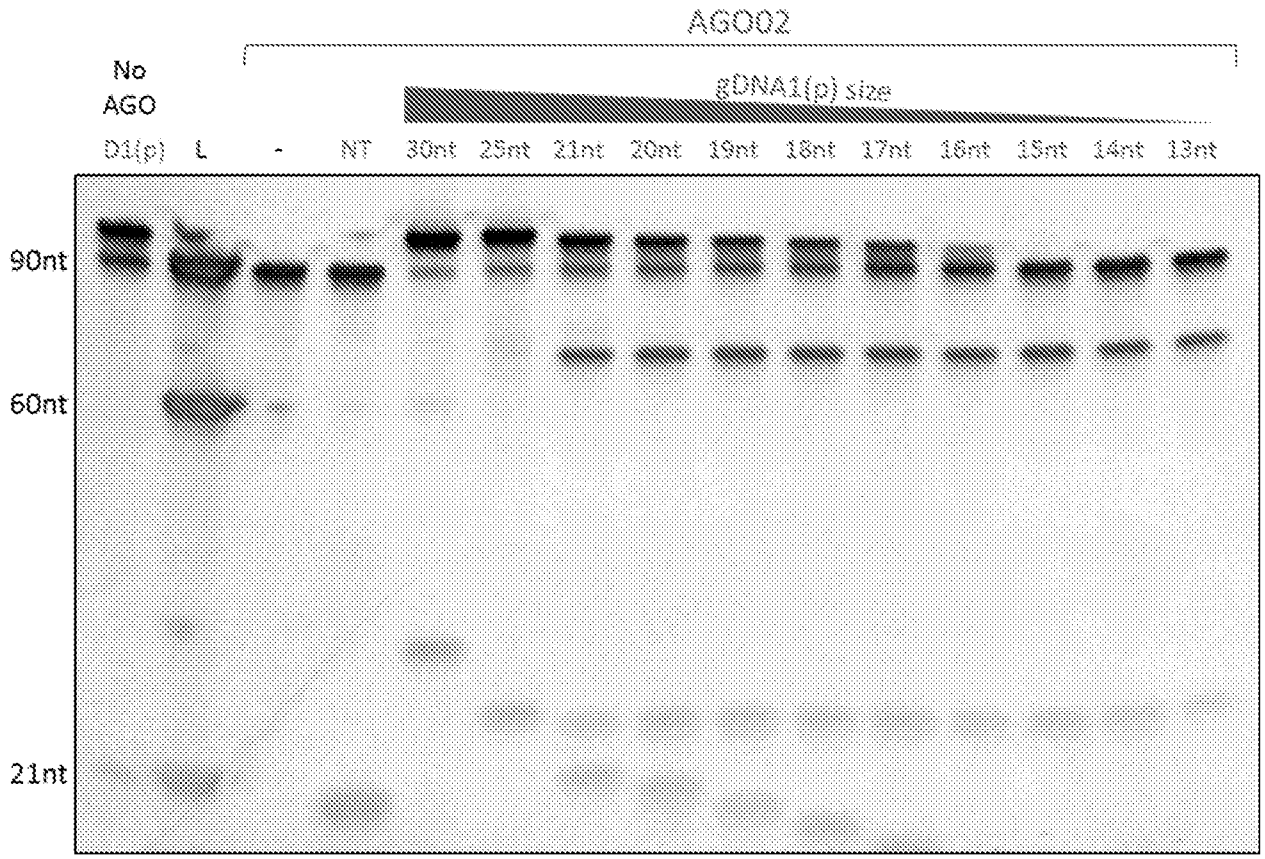

AGO2
No AGO
gDNA1(p) size
D1(p)
L
-
NT
30nt
25nt
21nt
20nt
19nt
18nt
17nt
16nt
15nt
14nt
13nt
90nt
60nt
21nt

FIG. 22B

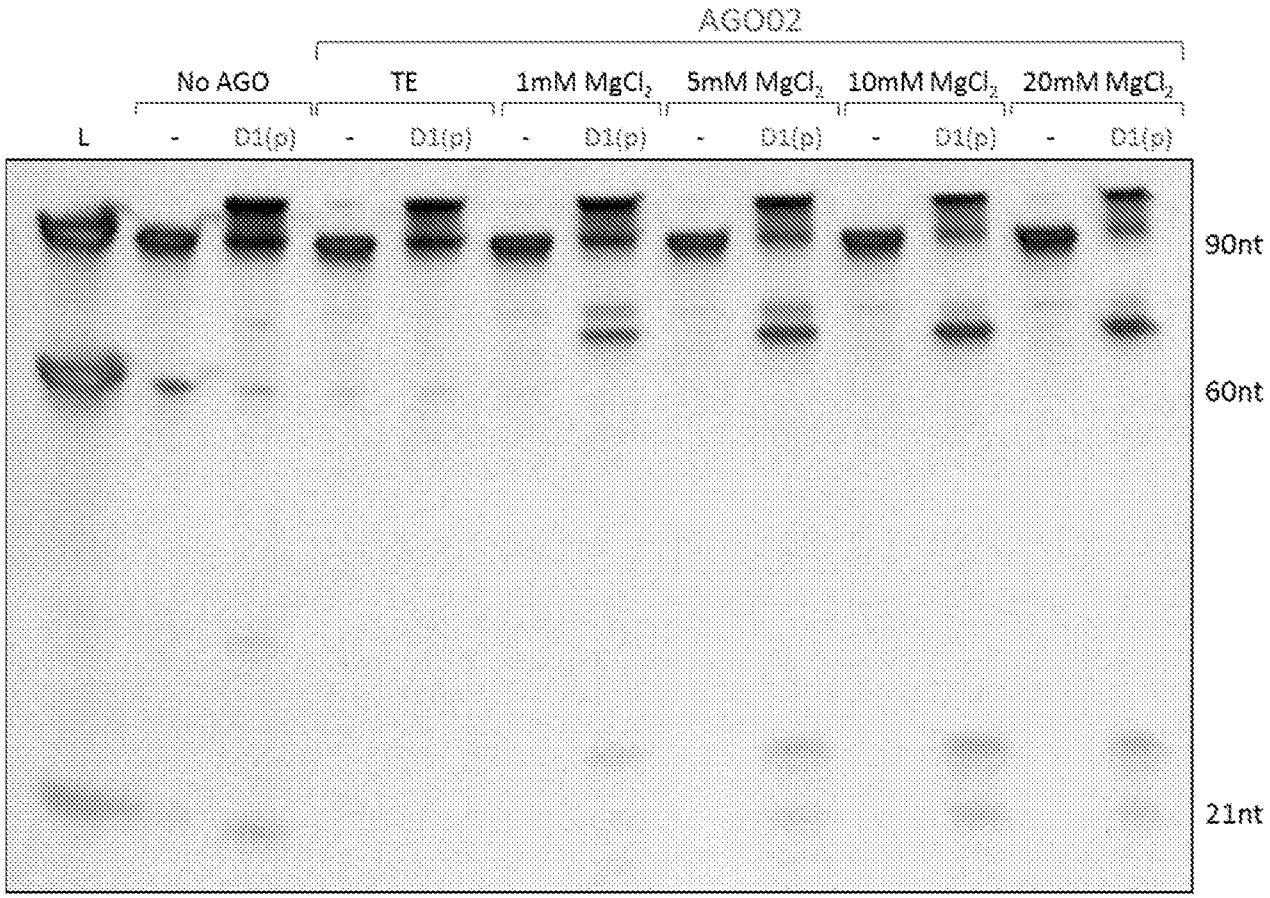

AGO2
No AGO
TE
1mM MgCl2
5mM MgCl2
10mM MgCl2
20mM MgCl2
L
-
D1(p)
90nt
60nt
21nt

FIG. 23A

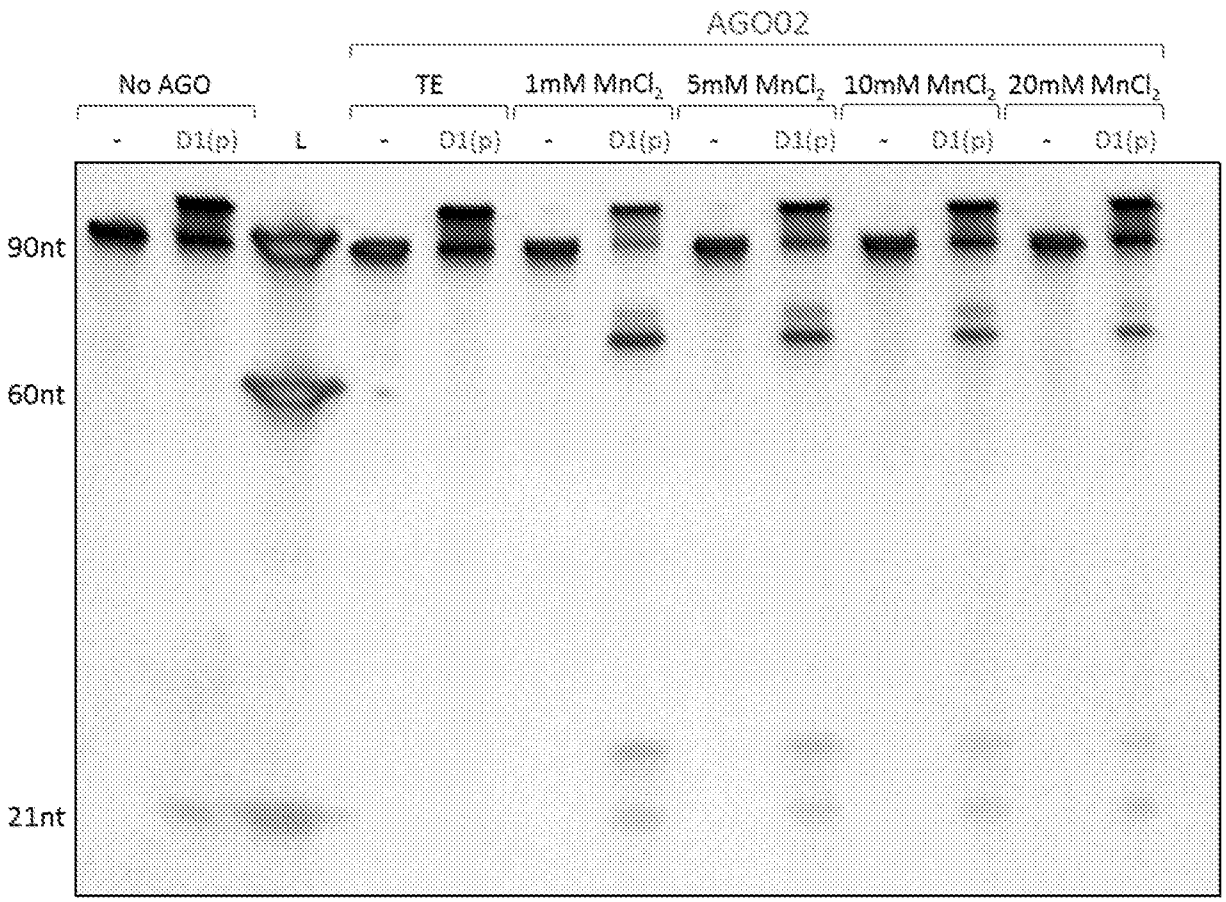

AGO02
No AGO
TE
1mM MnCl2
5mM MnCl2
10mM MnCl2
20mM MnCl2
-
D1(p)
L
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)
90nt
60nt
21nt

FIG. 23B

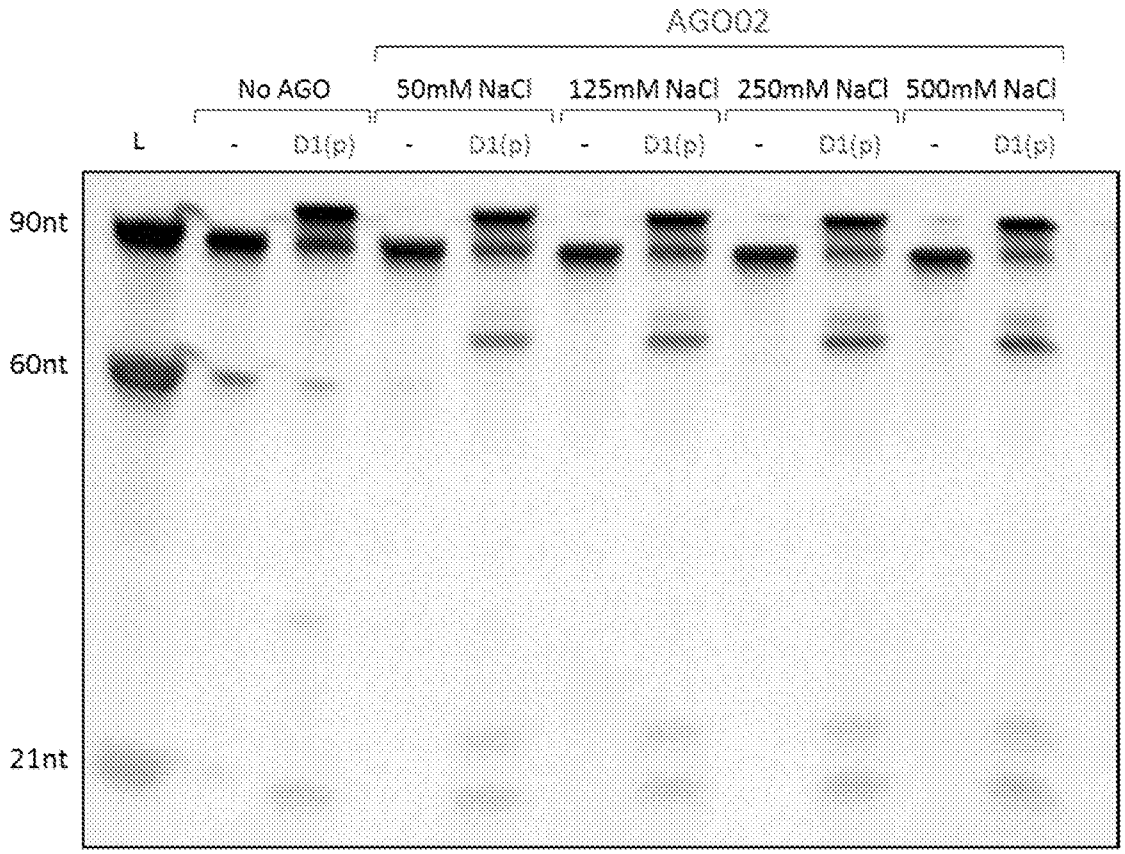

AGO02
No AGO
50mM NaCl
125mM NaCl
250mM NaCl
500mM NaCl
L
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)
90nt
60nt
21nt

FIG. 24

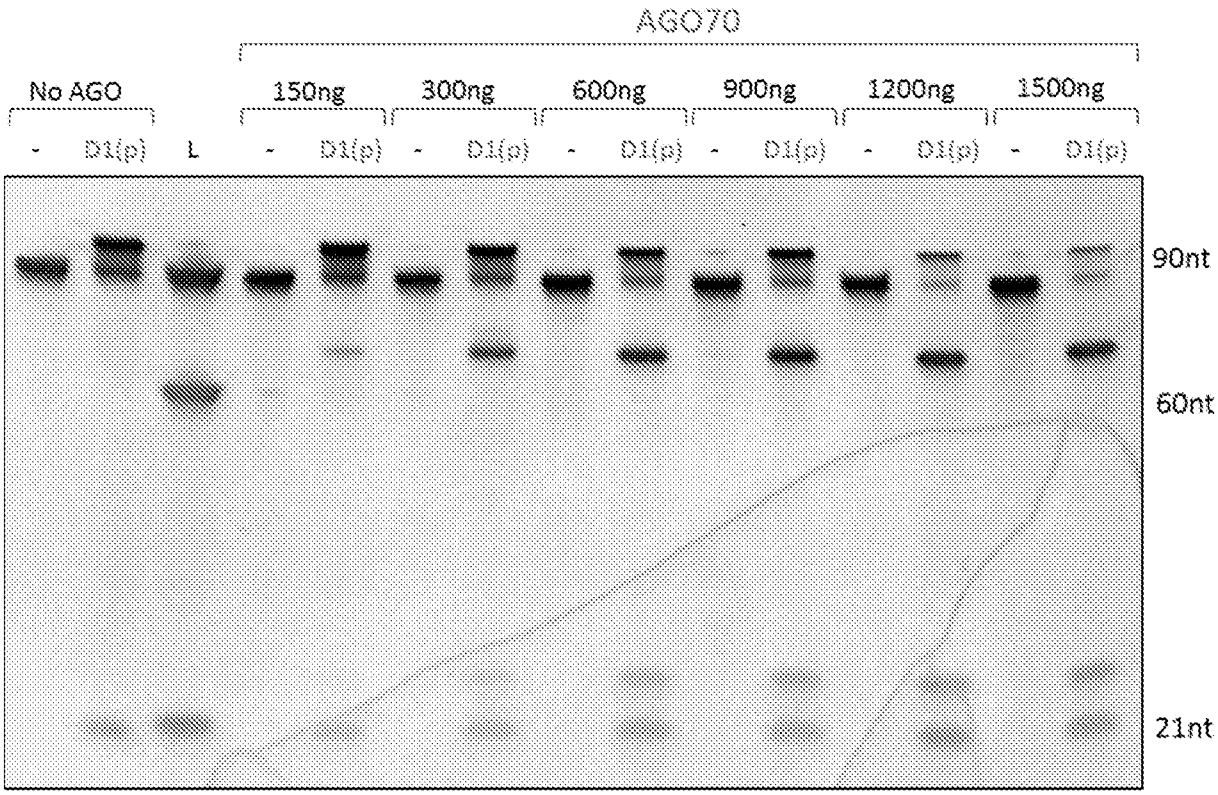

AGO70
No AGO
150ng
300ng
600ng
900ng
1200ng
1500ng
-
D1(p)
L
90nt
60nt
21nt

FIG. 25A

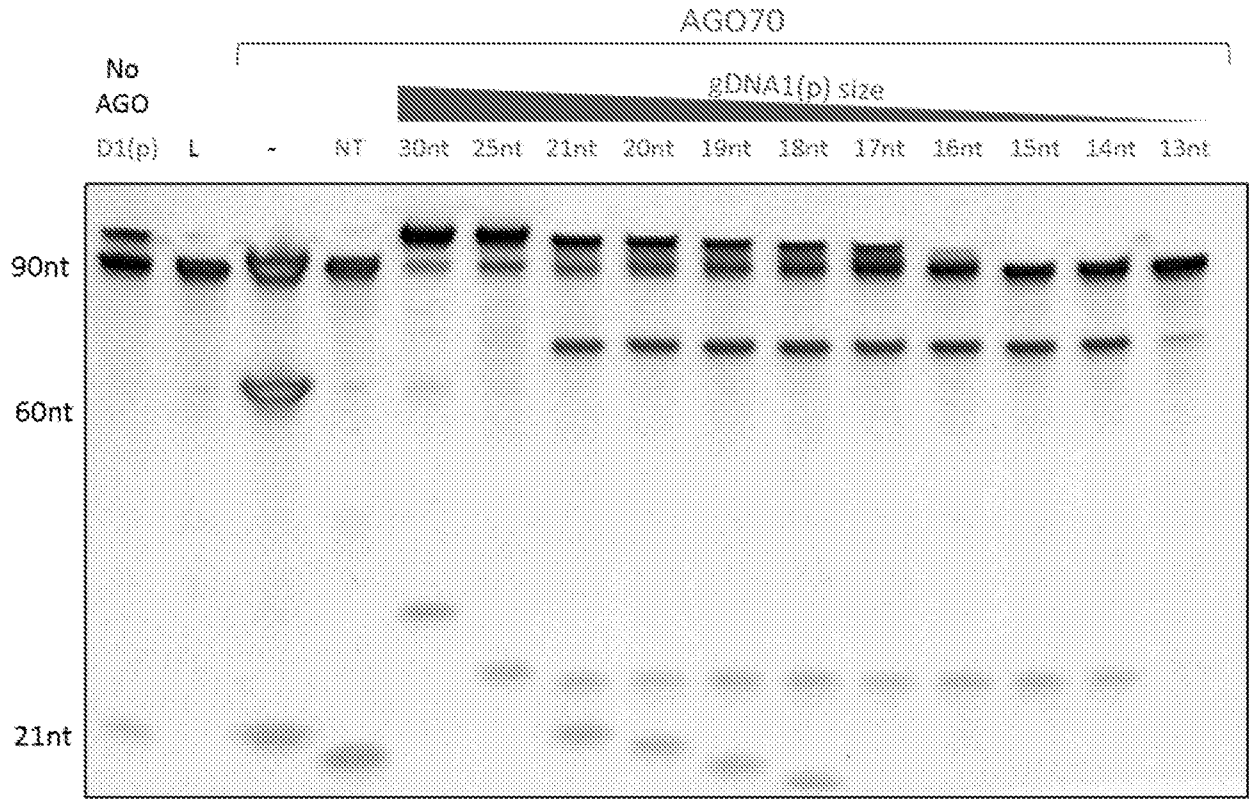

AGO70
No AGO
gDNA1(p) size
D1(p)
L
-
NT
30nt
25nt
21nt
20nt
19nt
18nt
17nt
16nt
15nt
14nt
13nt
90nt
60nt
21nt

FIG. 25B

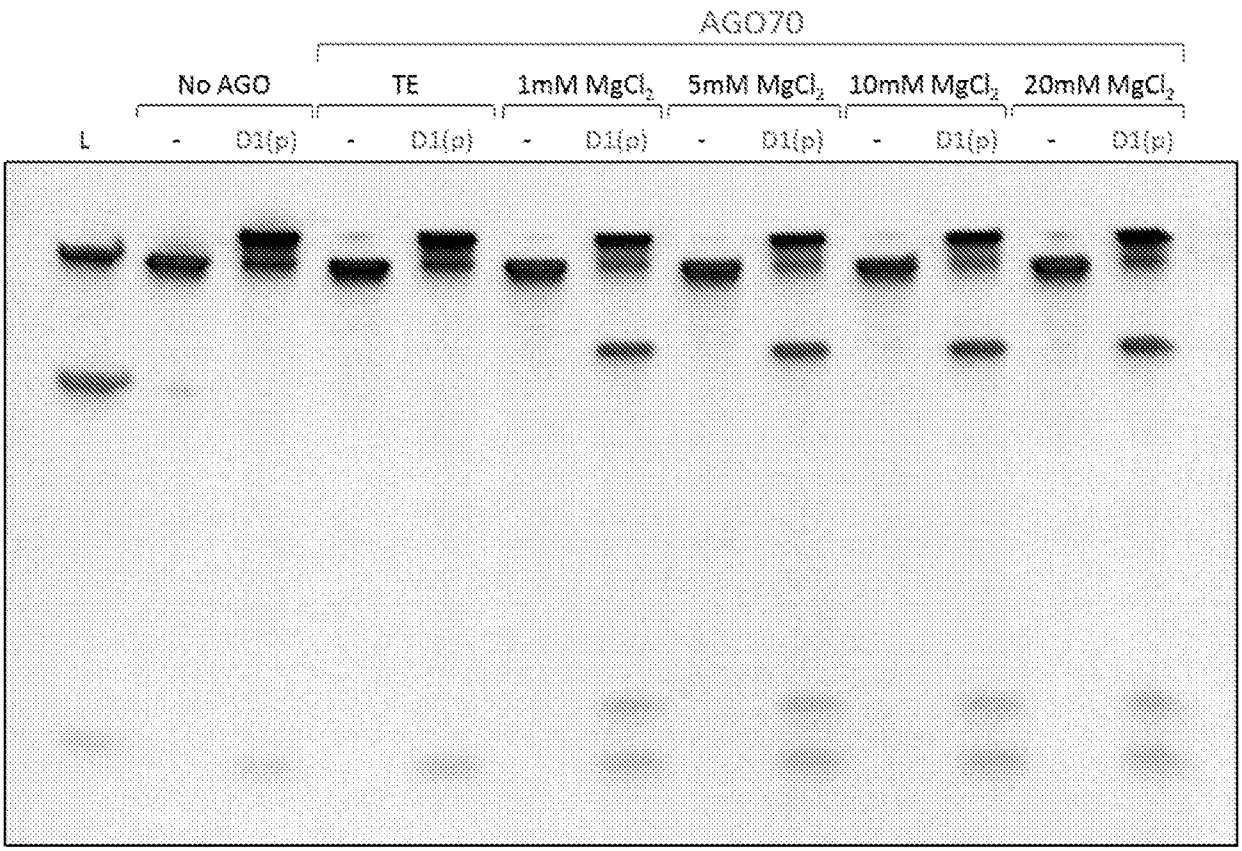
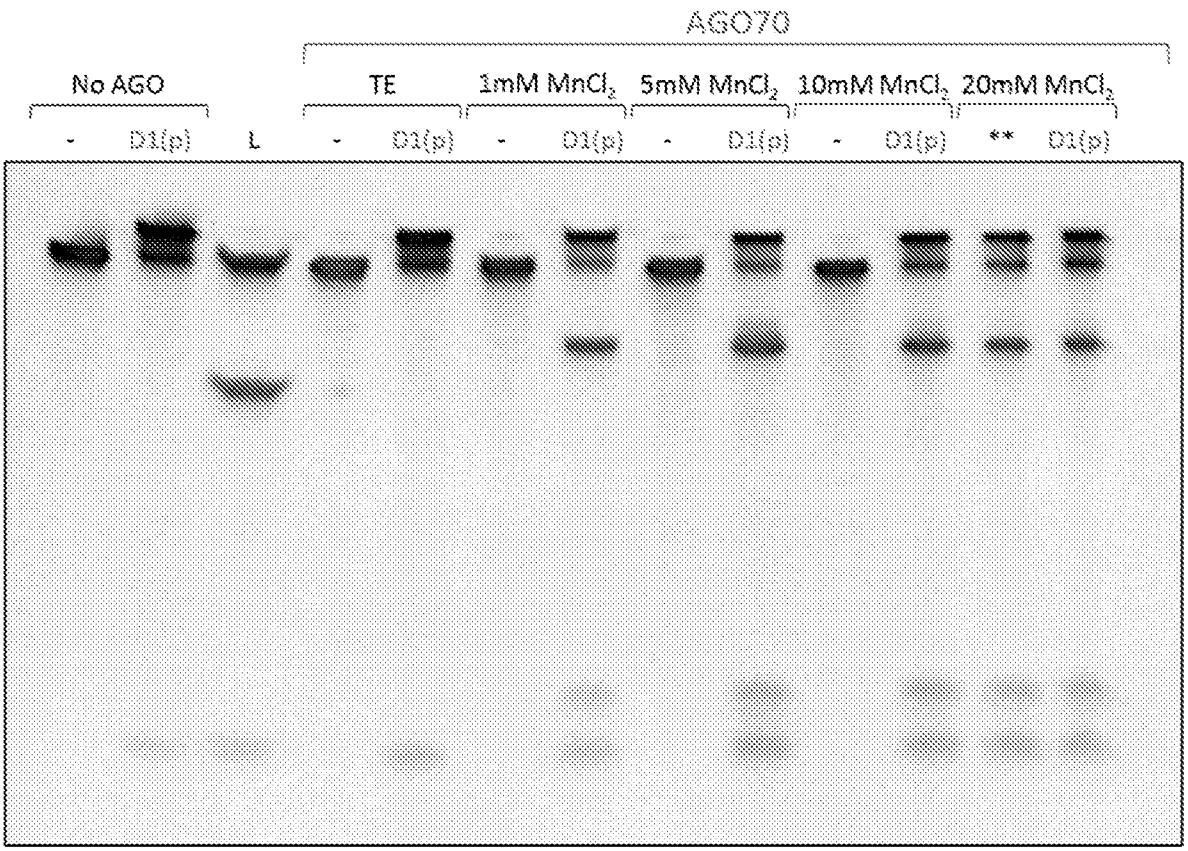

AGO70
No AGO
TE
1mM MgCl2
5mM MgCl2
10mM MgCl2
20mM MgCl2
L
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)
-
D1(p)

FIG. 26A



*** experimental error: D1(p) instead of no gDNA*

FIG. 26B

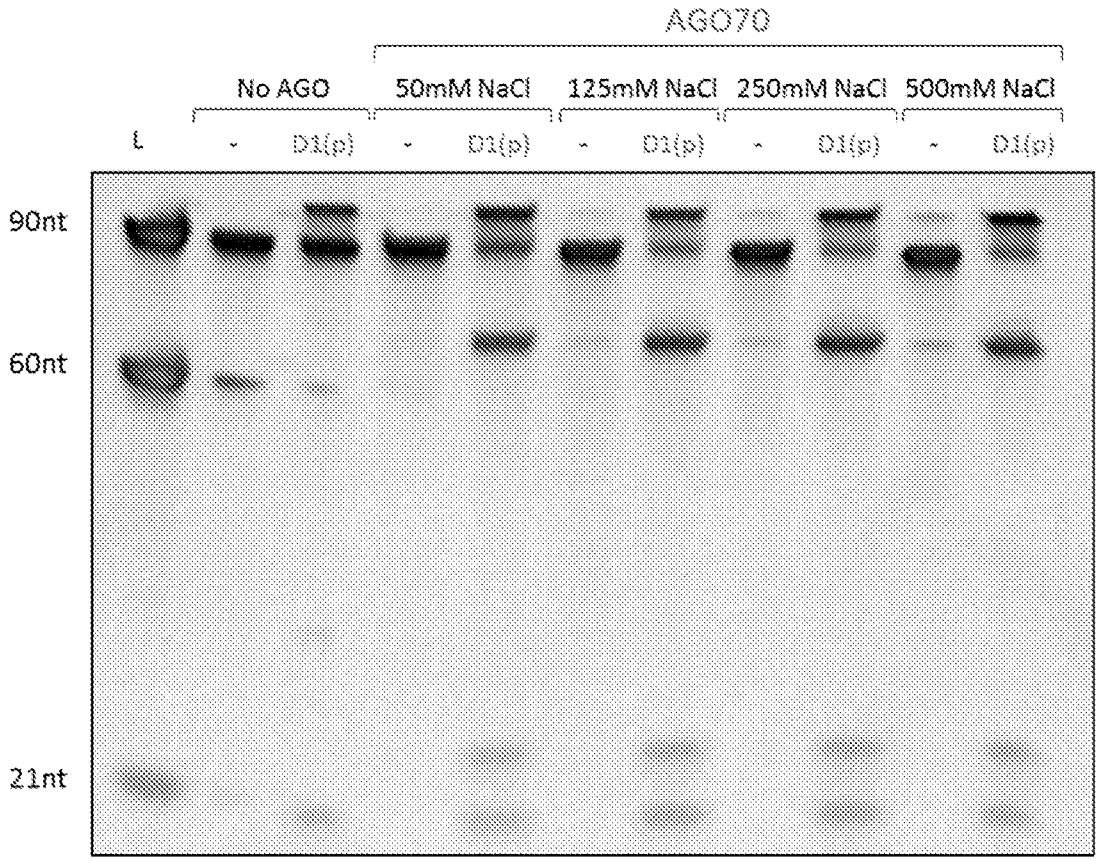
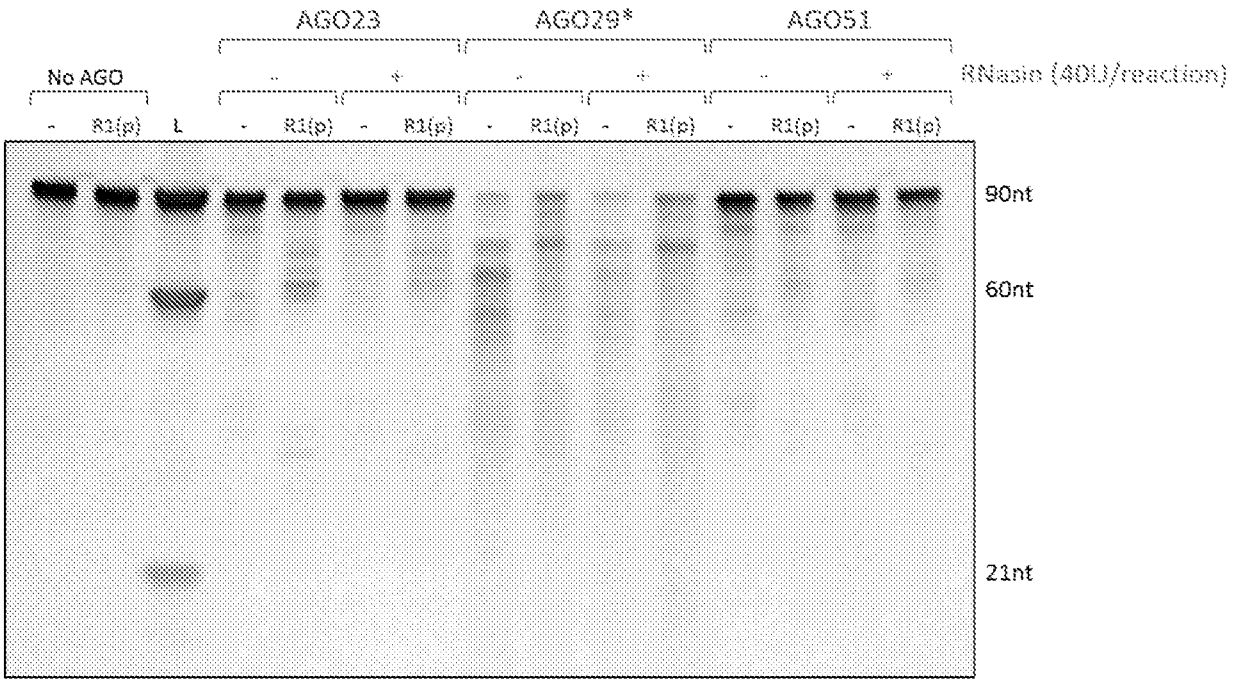

AGO70
No AGO
50mM NaCl
125mM NaCl
250mM NaCl
500mM NaCl
L
-
D1(p)
90nt
60nt
21nt

FIG. 27



AGO29*: protein amounts used = ½ volume of protein (125ng)

FIG. 28

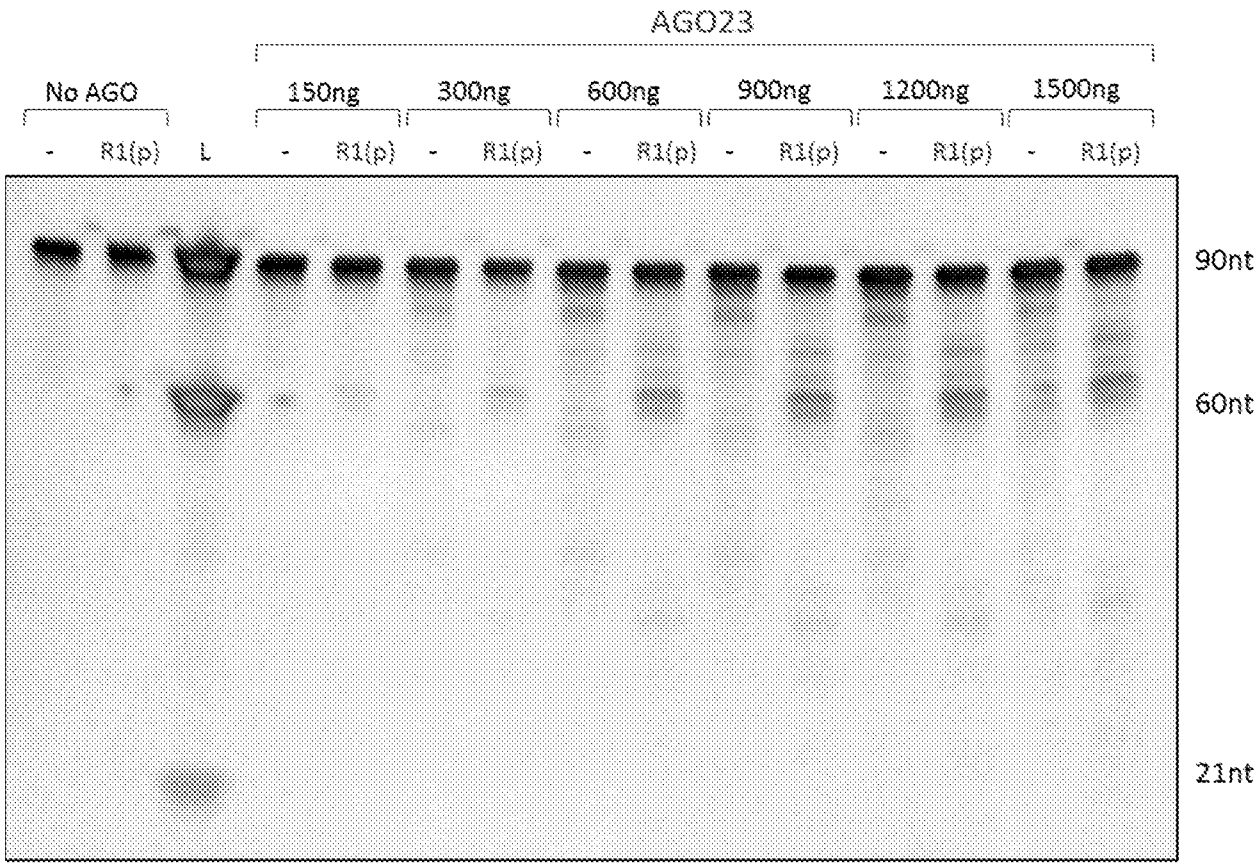

AGO23
No AGO
150ng
300ng
600ng
900ng
1200ng
1500ng
- R1(p) L - R1(p) - R1(p) - R1(p) - R1(p) - R1(p) - R1(p)
90nt
60nt
21nt

FIG. 29A

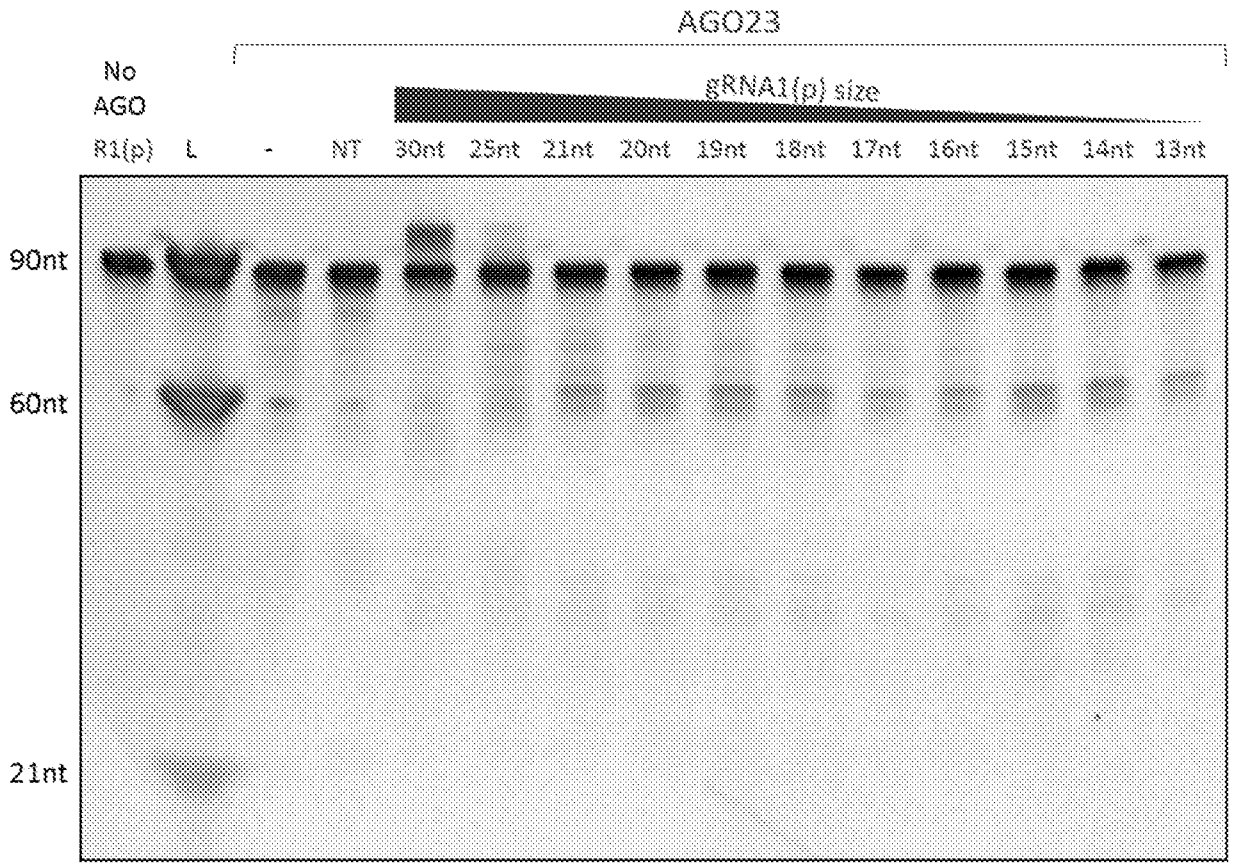

AGO23
No AGO
gRNA1(p) size
R1(p)
L
-
NT
30nt
25nt
21nt
20nt
19nt
18nt
17nt
16nt
15nt
14nt
13nt
90nt
60nt
21nt

FIG. 29B

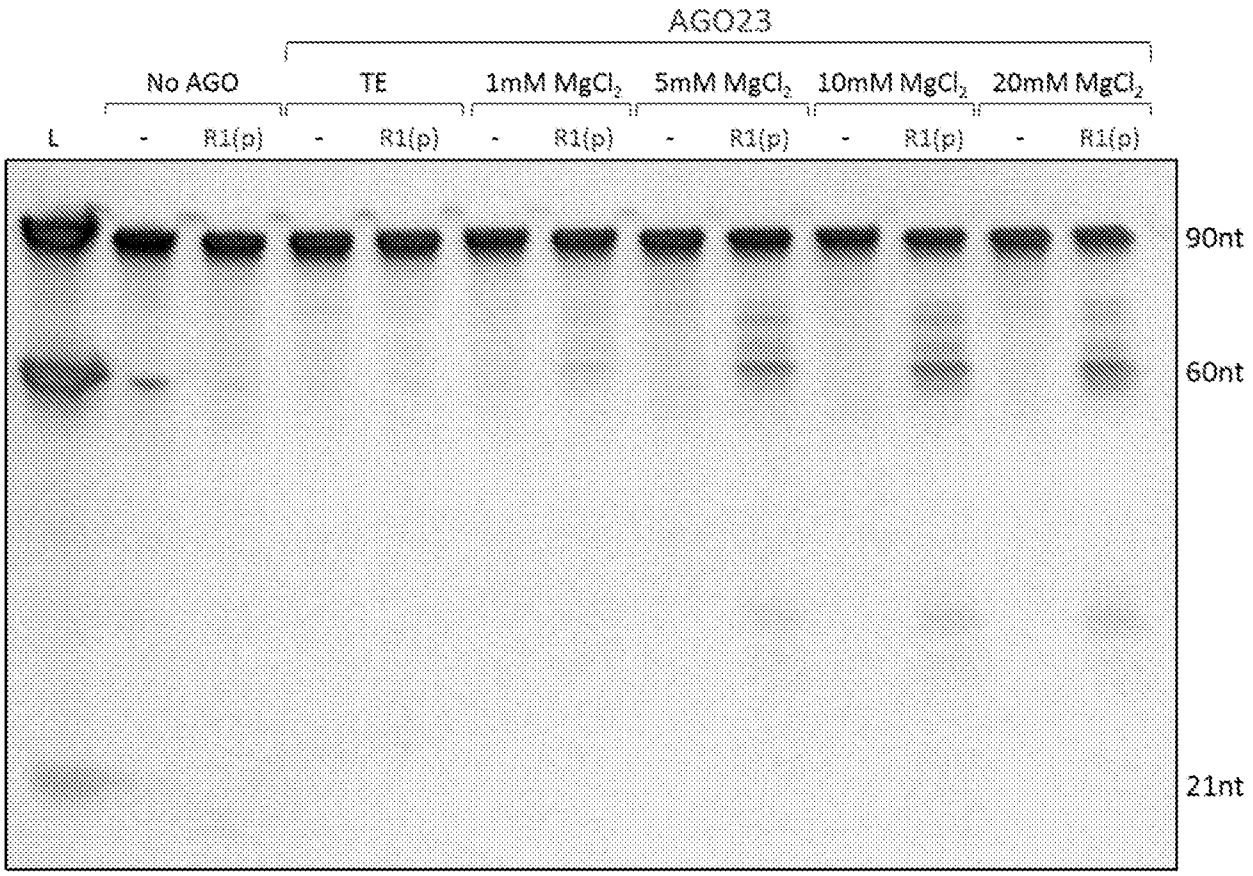

AGO23
No AGO
TE
1mM MgCl2
5mM MgCl2
10mM MgCl2
20mM MgCl2
L
-
R1(p)
90nt
60nt
21nt

FIG. 30A

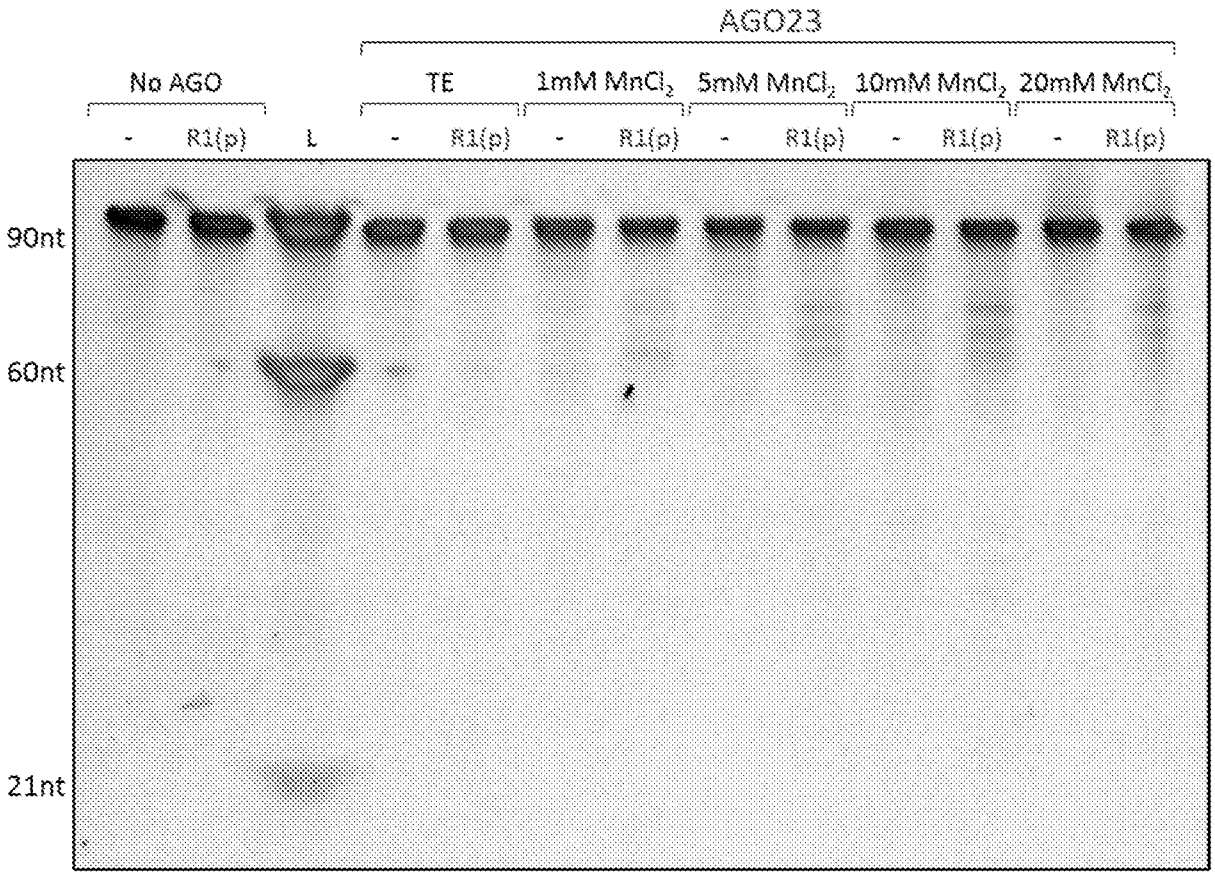

AGO23
No AGO
TE
1mM MnCl2
5mM MnCl2
10mM MnCl2
20mM MnCl2
-
R1(p)
L
90nt
60nt
21nt

FIG. 30B

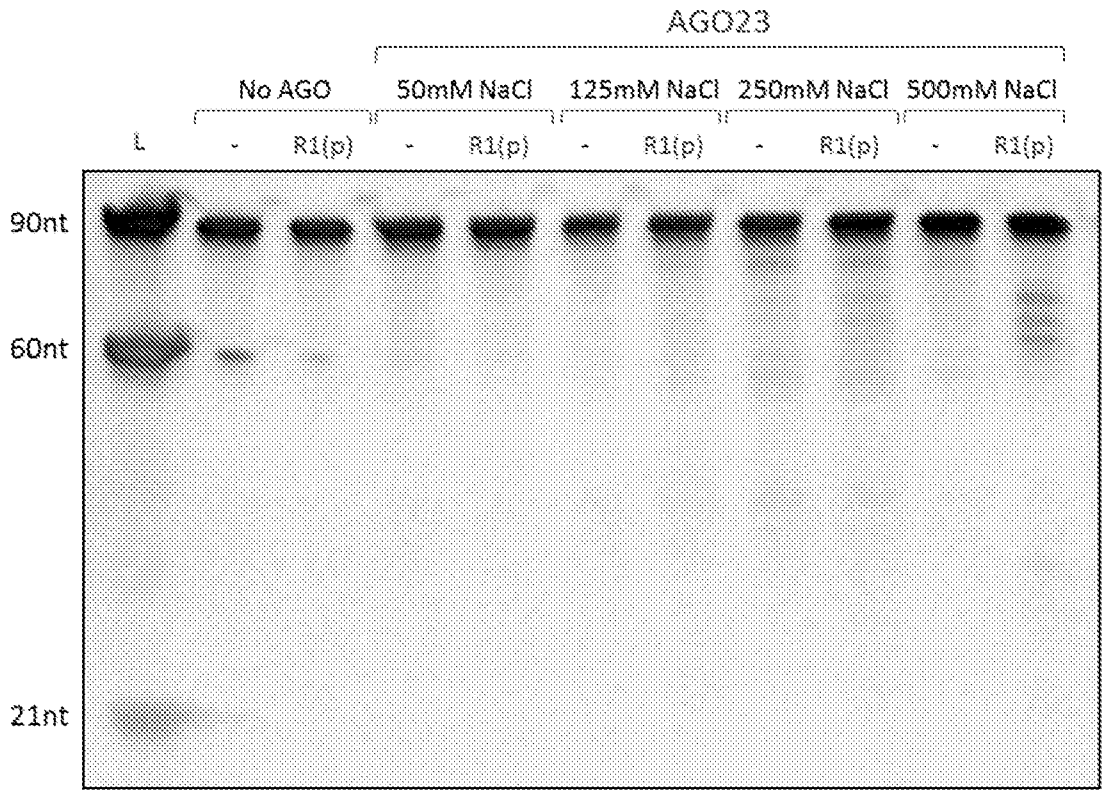

AGO23
No AGO
50mM NaCl
125mM NaCl
250mM NaCl
500mM NaCl
L
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)
90nt
60nt
21nt

FIG. 31

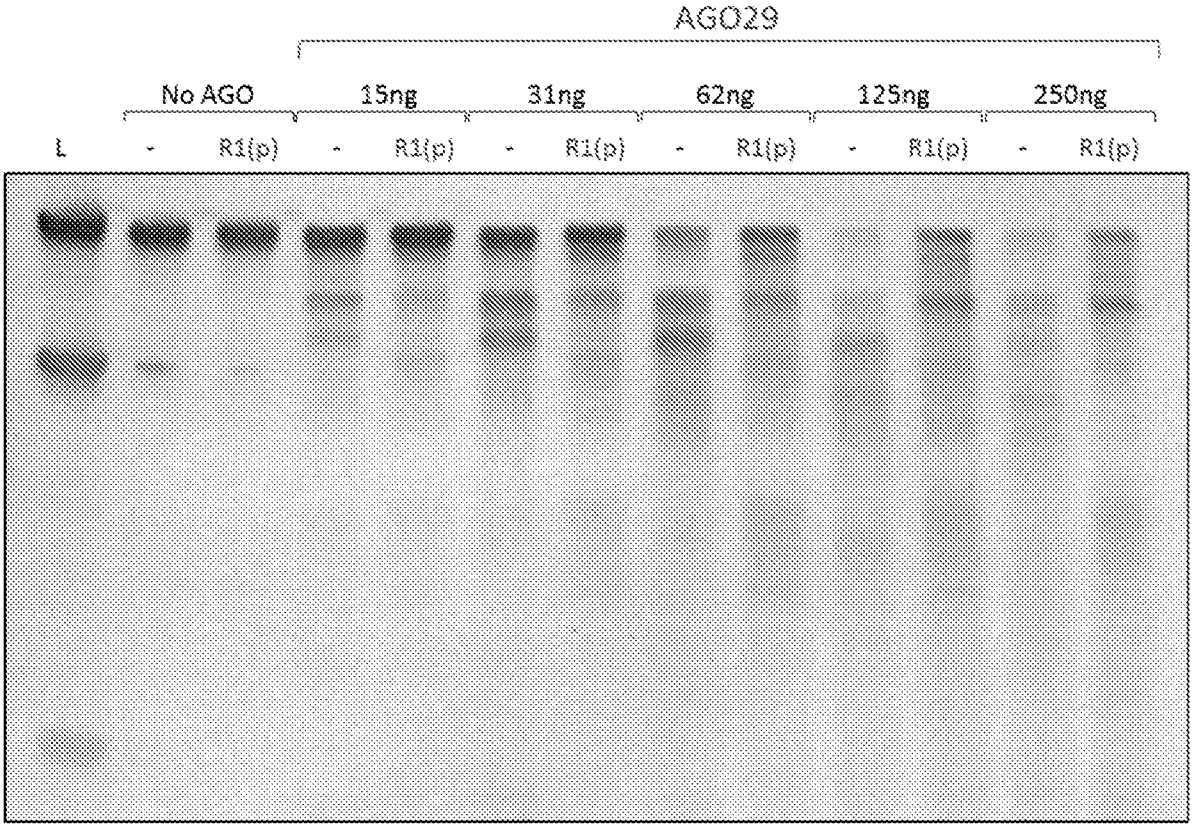
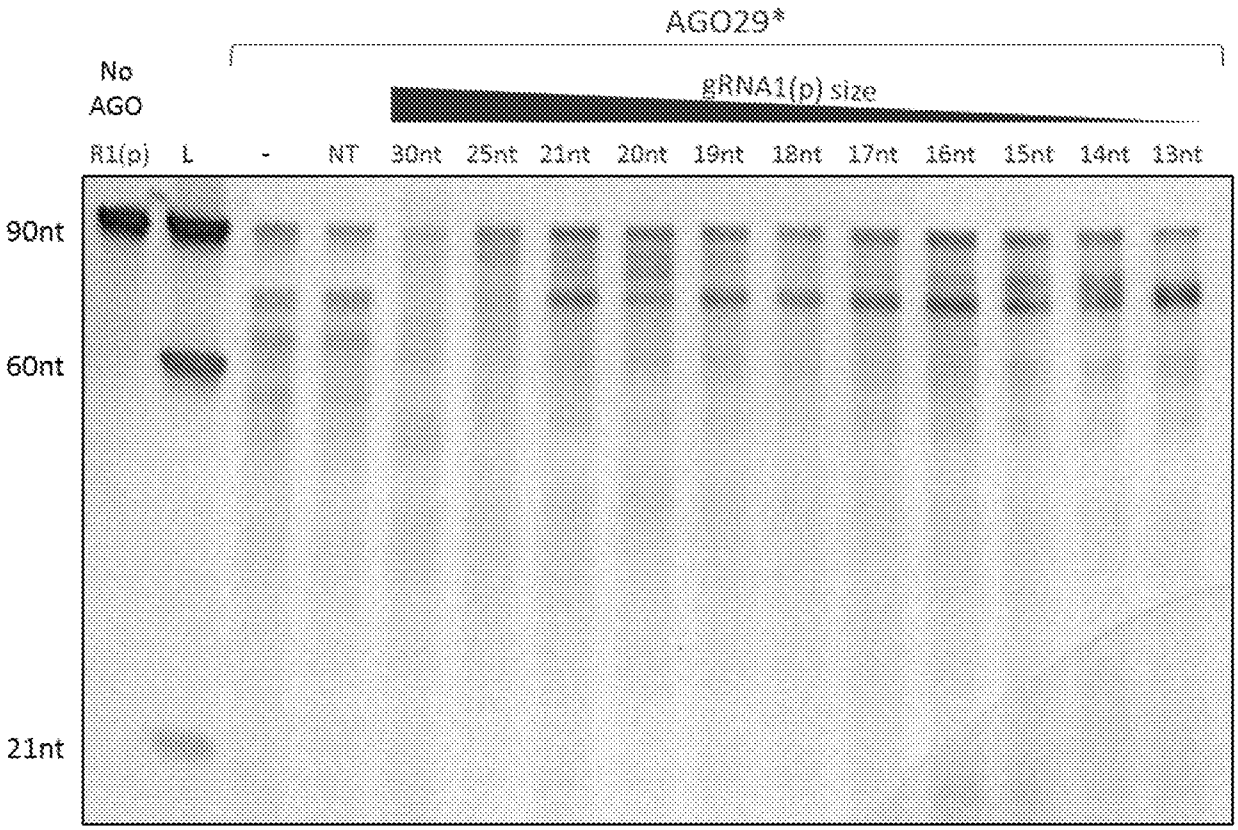

AGO29
No AGO
15ng
31ng
62ng
125ng
250ng
L
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)

FIG. 32A



AGO29*: 125ng/reaction

FIG. 32B

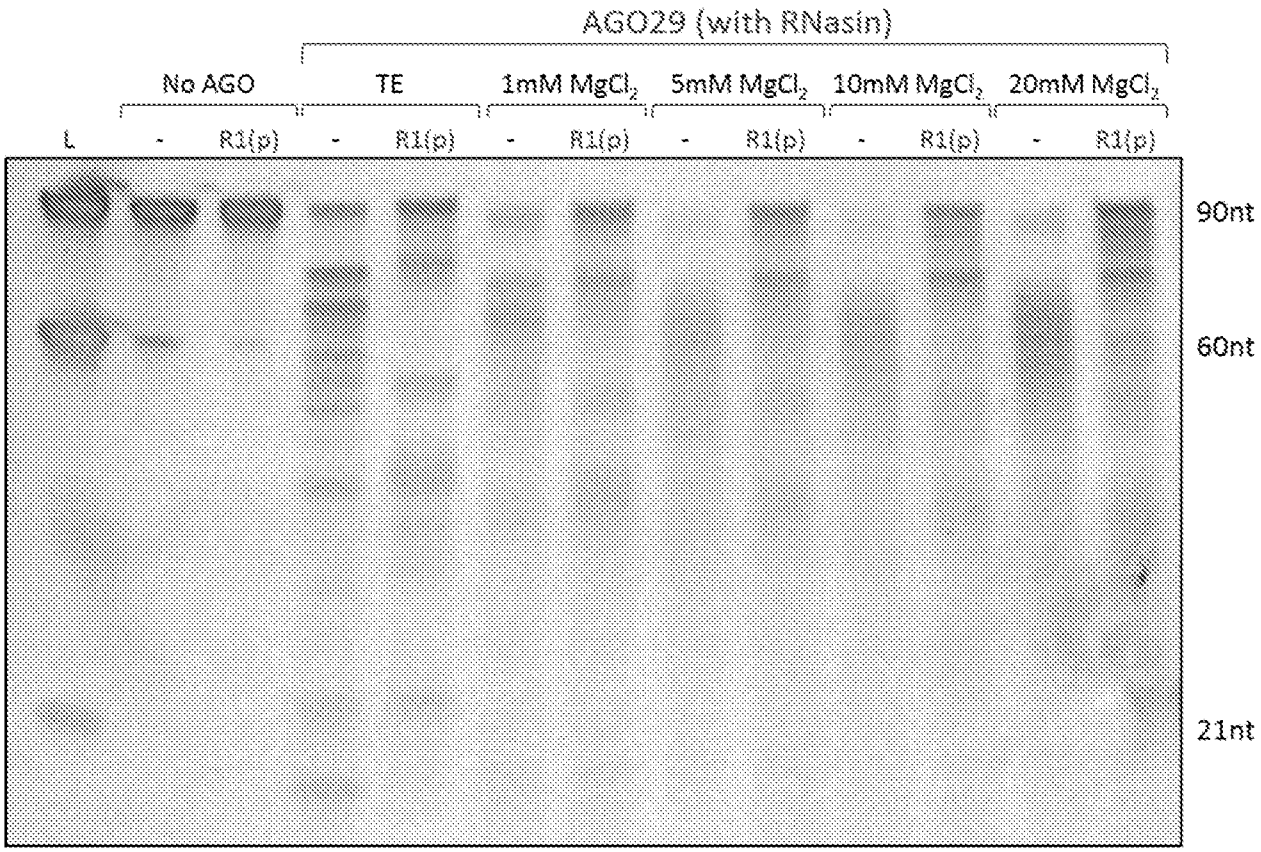

AGO29 (with RNasin)
No AGO
TE
1mM MgCl2
5mM MgCl2
10mM MgCl2
20mM MgCl2
L
-
R1(p)
90nt
60nt
21nt

FIG. 33A

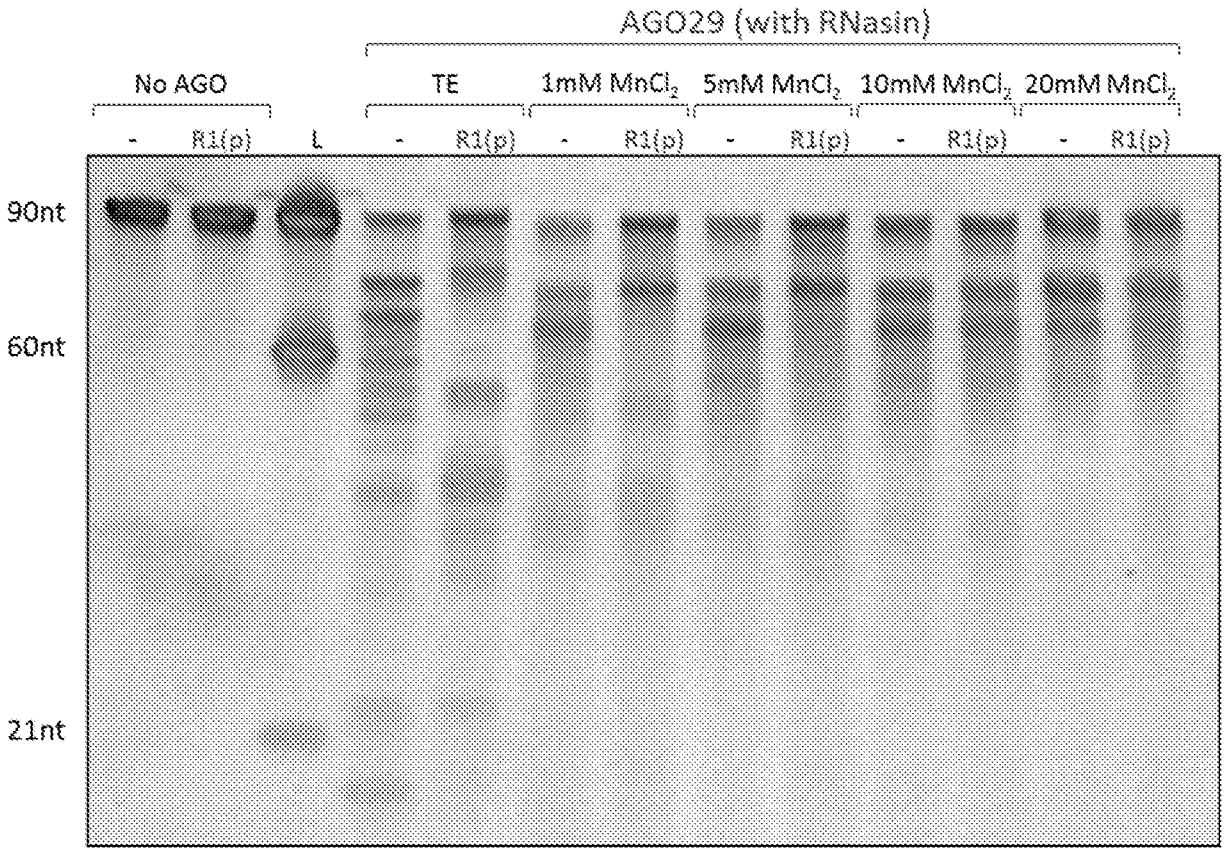
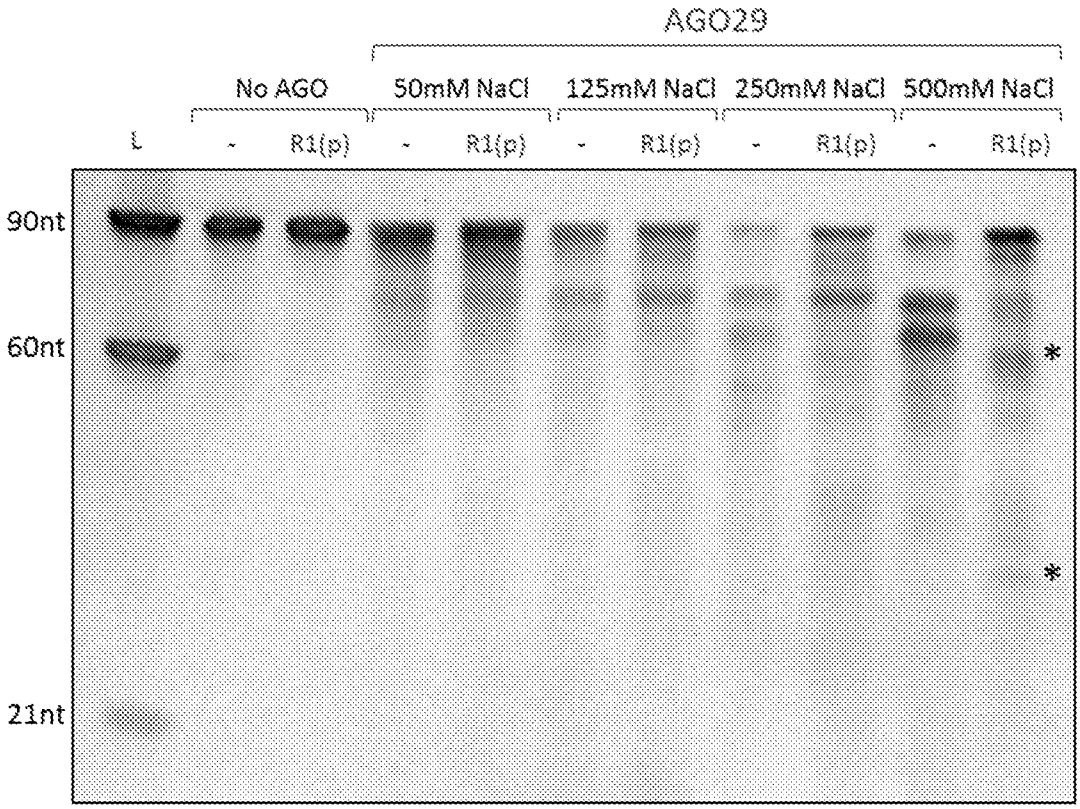

AGO29 (with RNasin)
No AGO
TE
1mM MnCl2
5mM MnCl2
10mM MnCl2
20mM MnCl2
-
R1(p)
L
90nt
60nt
21nt

FIG. 33B



>> * Expected cutting pattern at 500mM NaCl

FIG. 34

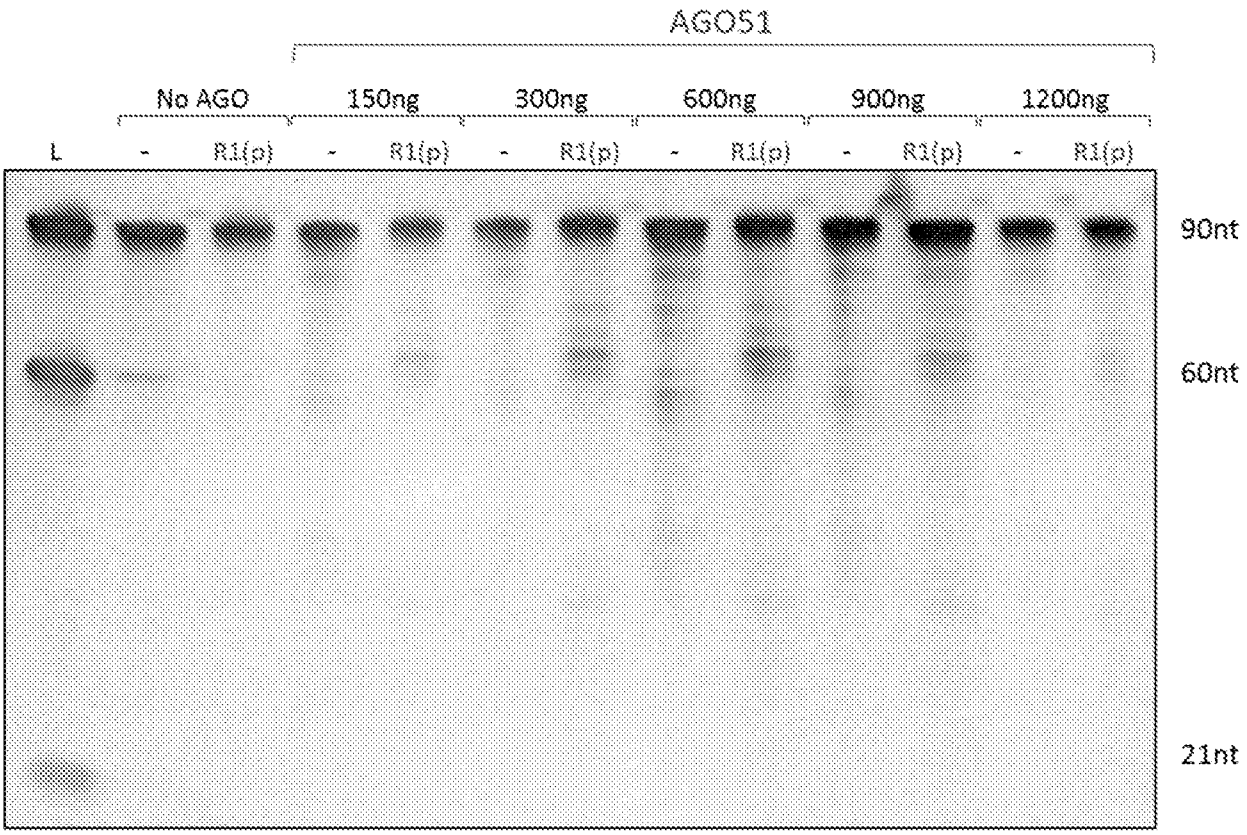

AGO51
No AGO
150ng
300ng
600ng
900ng
1200ng
L
-
R1(p)
90nt
60nt
21nt

FIG. 35A

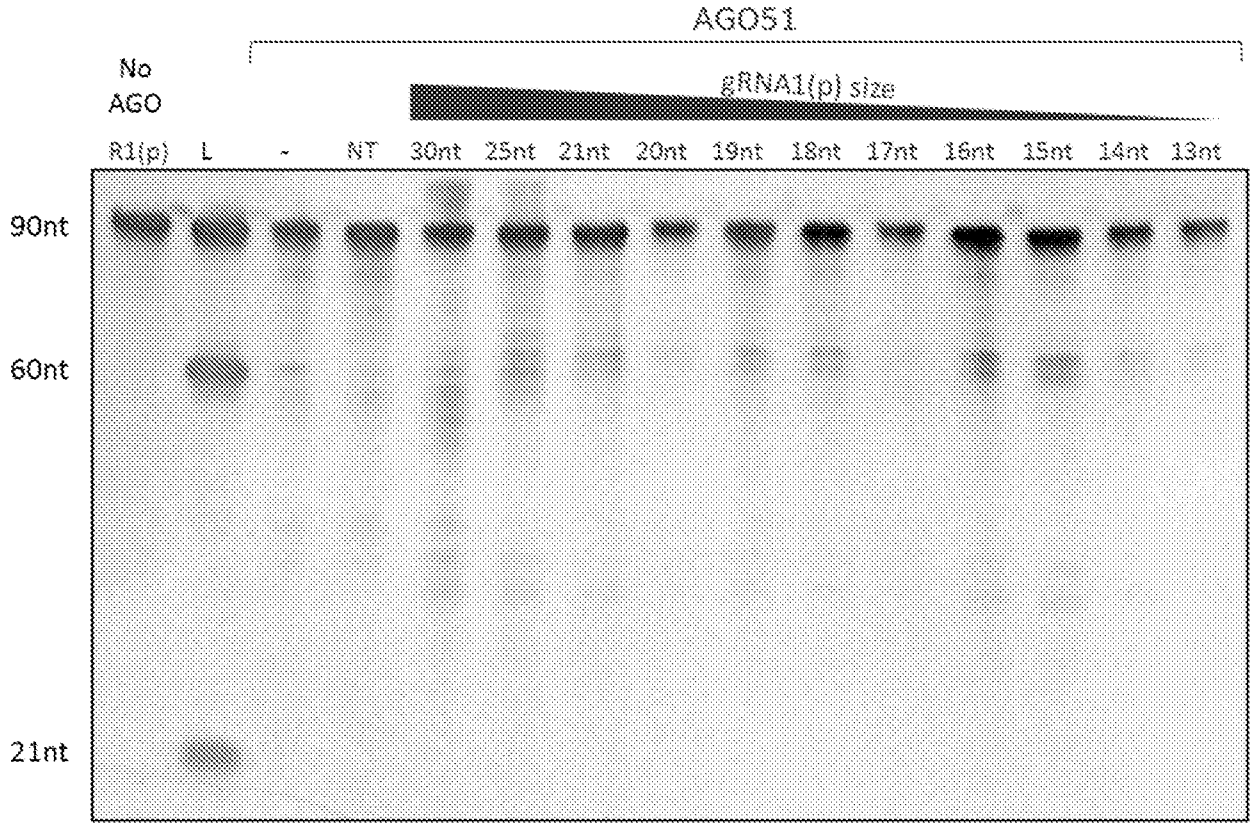

AGO51
No AGO
gRNA1(p) size
R1(p)
L
-
NT
30nt
25nt
21nt
20nt
19nt
18nt
17nt
16nt
15nt
14nt
13nt
90nt
60nt
21nt

FIG. 35B

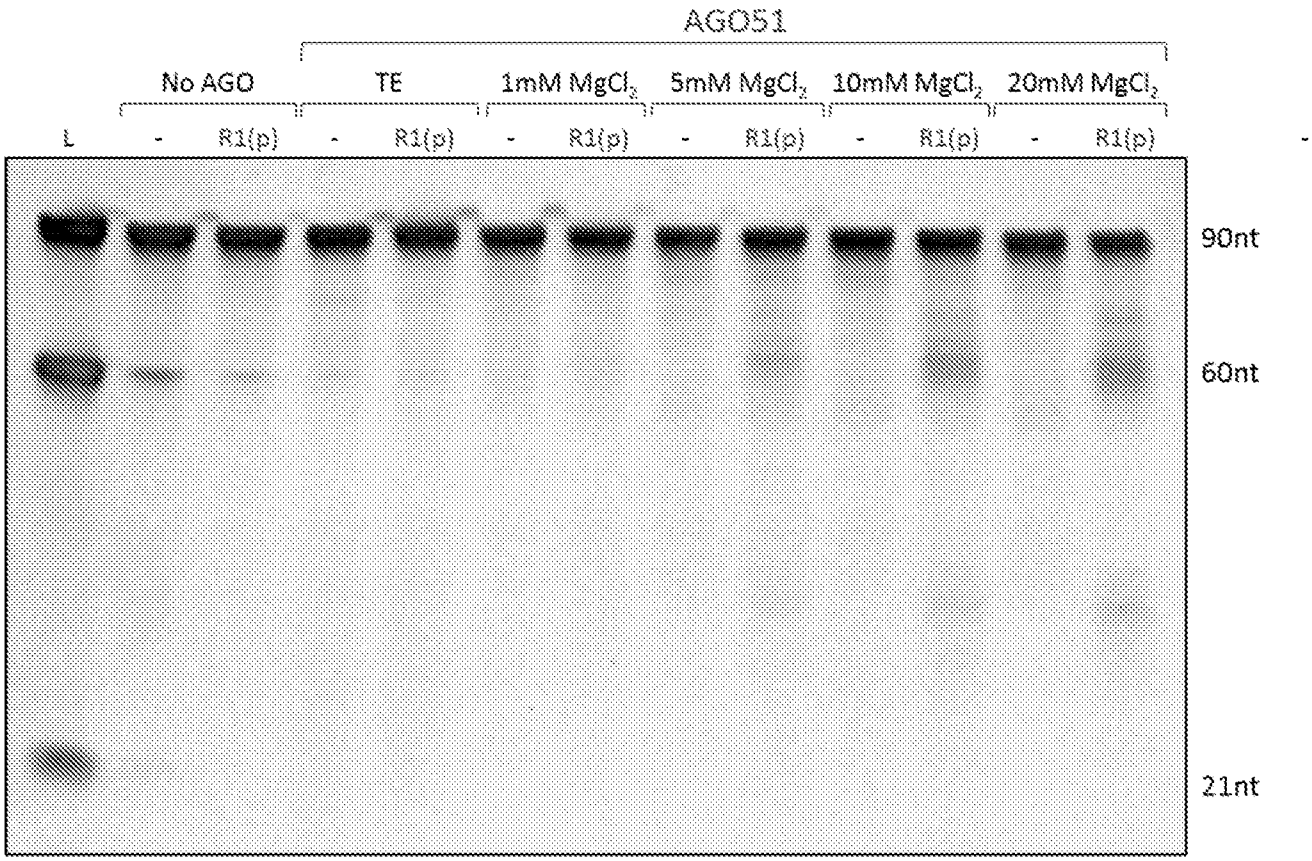

AGO51
No AGO
TE
1mM MgCl2
5mM MgCl2
10mM MgCl2
20mM MgCl2
L
-
R1(p)
90nt
60nt
21nt

FIG. 36A

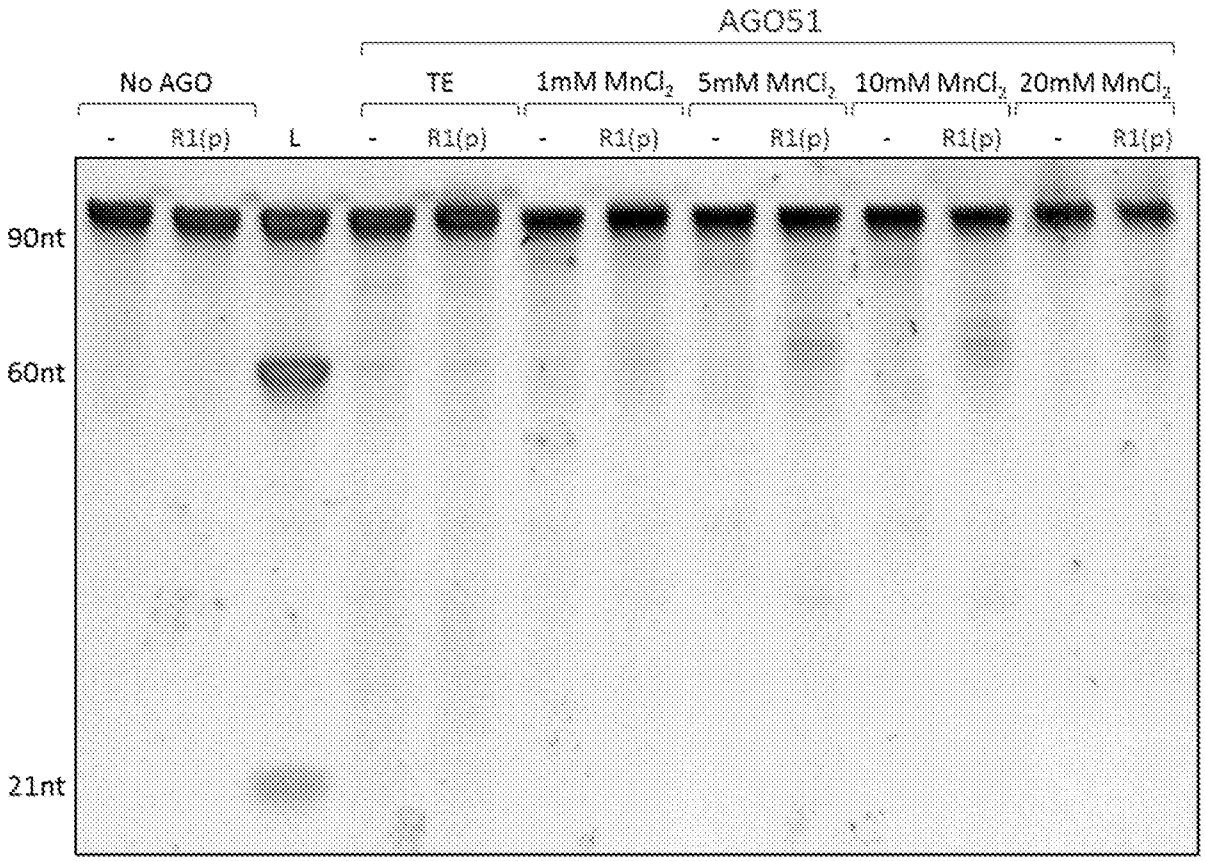

AGO51
No AGO
TE
1mM MnCl2
5mM MnCl2
10mM MnCl2
20mM MnCl2
- R1(p) L - R1(p) - R1(p) - R1(p) - R1(p) - R1(p)
90nt
60nt
21nt

FIG. 36B

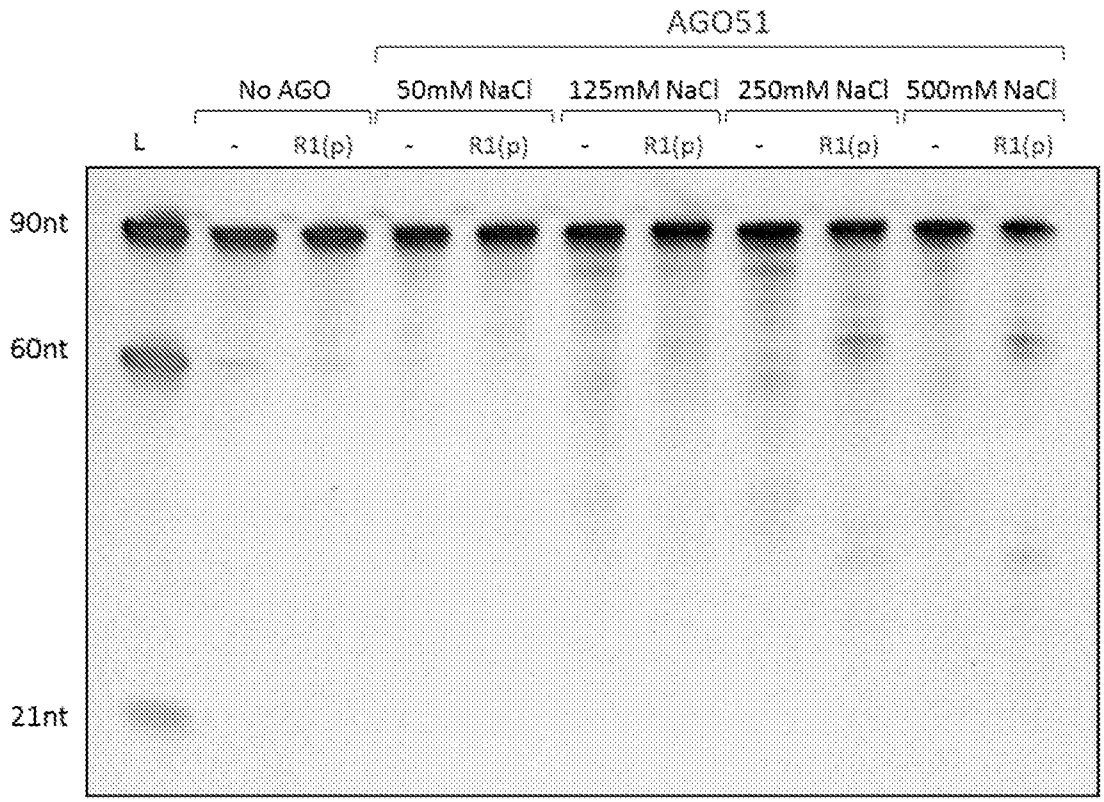

AGO51
No AGO
50mM NaCl
125mM NaCl
250mM NaCl
500mM NaCl
L
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)
-
R1(p)
90nt
60nt
21nt

FIG. 37

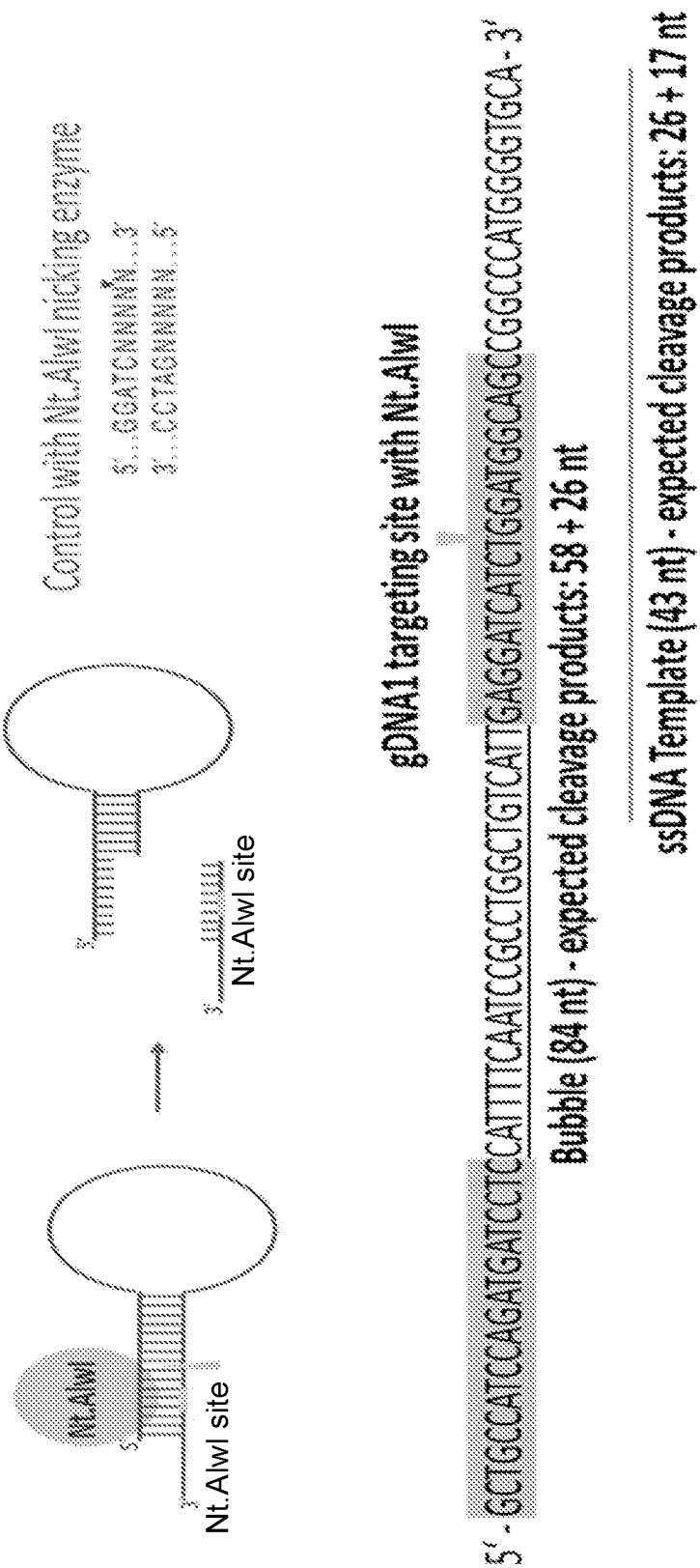

Control with Nt.AlwI nicking enzyme
5'...GGATCNNNNN...3'
3'...CCTAGNNNNN...5'
Nt.AlwI
Nt.AlwI site
Nt.AlwI site
gDNA1 targeting site with Nt.AlwI
5' - GCTGCCATCCAGATGATCCTCCATTTTCAATCCGCCTGGCTGTCATTGAGGATCATCTGGATGGCAGCCGGCCCATGGGGTGCA - 3'
Bubble (84 nt) - expected cleavage products: 58 + 26 nt
ssDNA Template (43 nt) - expected cleavage products: 26 + 17 nt

FIG. 38

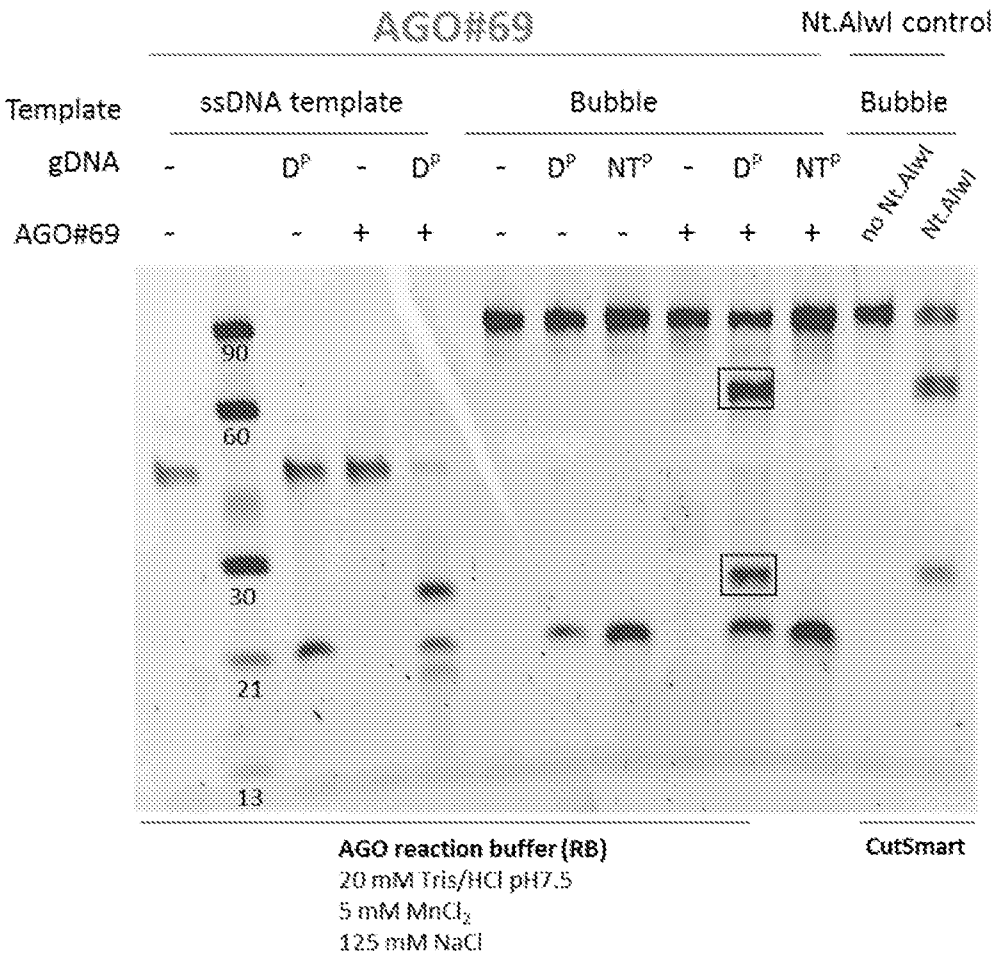
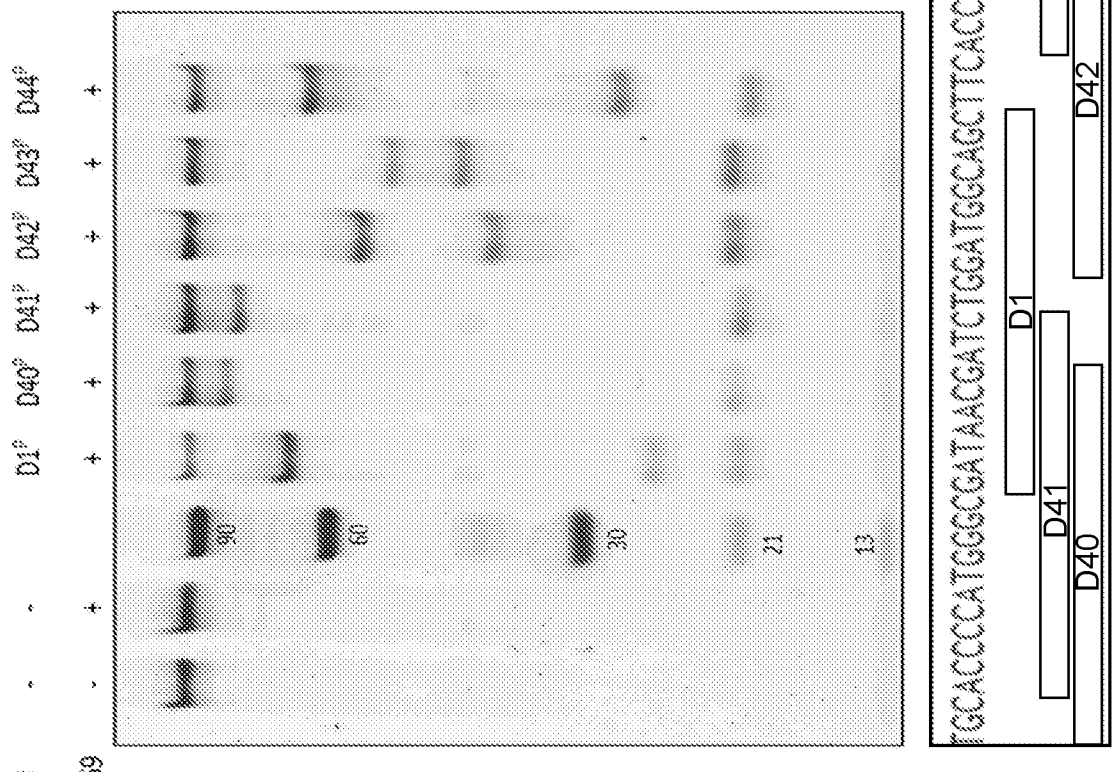

AGO#69
Nt.AlwI control
Template
ssDNA template
Bubble
Bubble
gDNA
AGO#69
no Nt.AlwI
Nt.AlwI
90
60
30
21
13
AGO reaction buffer (RB)
20 mM Tris/HCl pH7.5
5 mM MnCl2
125 mM NaCl
CutSmart

FIG. 39

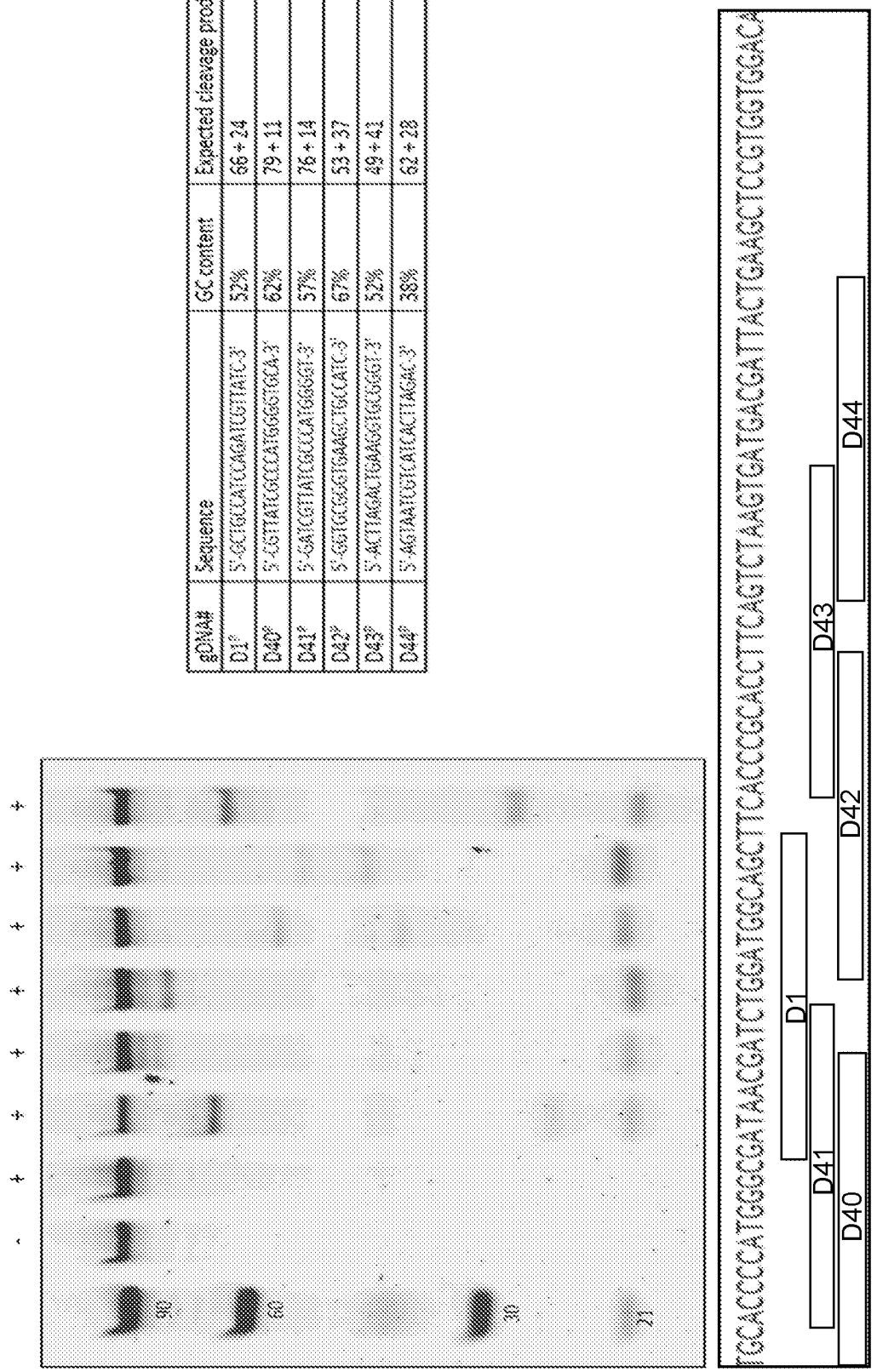


gDNA#	Sequence	GC content	Expected cleavage products
D1*	5'-GCTGCCATCCAAATCGTTATC-3'	52%	66 + 24
D40*	5'-CGTTATCGCCATGGGGTGCA-3'	62%	79 + 11
D41*	5'-GATCGTTATCGCCATGGGGT-3'	57%	76 + 14
D42*	5'-GGTGCGGTTGAAGCTGCCATC-3'	67%	53 + 37
D43*	5'-ACTTAGATCGAAAGGTGCGGGT-3'	52%	49 + 41
D44*	5'-AGTAATCGTCATCACTTAGAC-3'	38%	62 + 28

Data for AGO#02, AGO#41 and AGO#70 - Slides 13, 14 & 15

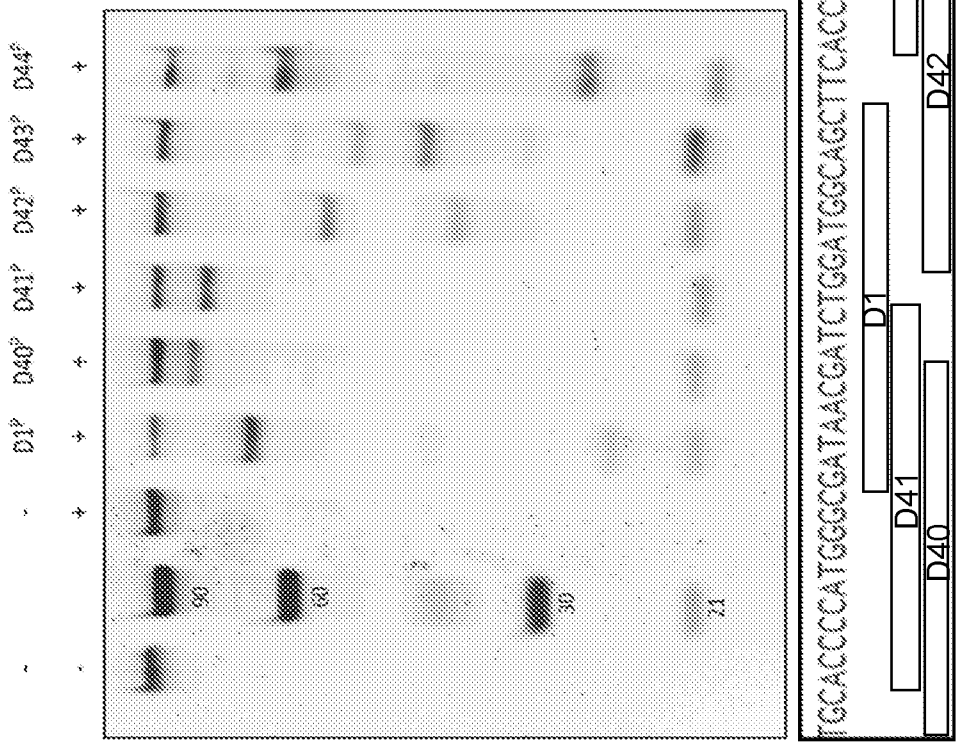
TGCACCCCATGGCGATAAGCGATCTCGATGGCAGCTTCAGCCGCAGCTTCAGTCTAAGTGATGACGATTACTGAAGCTCCGTGGTGCAC

FIG. 40



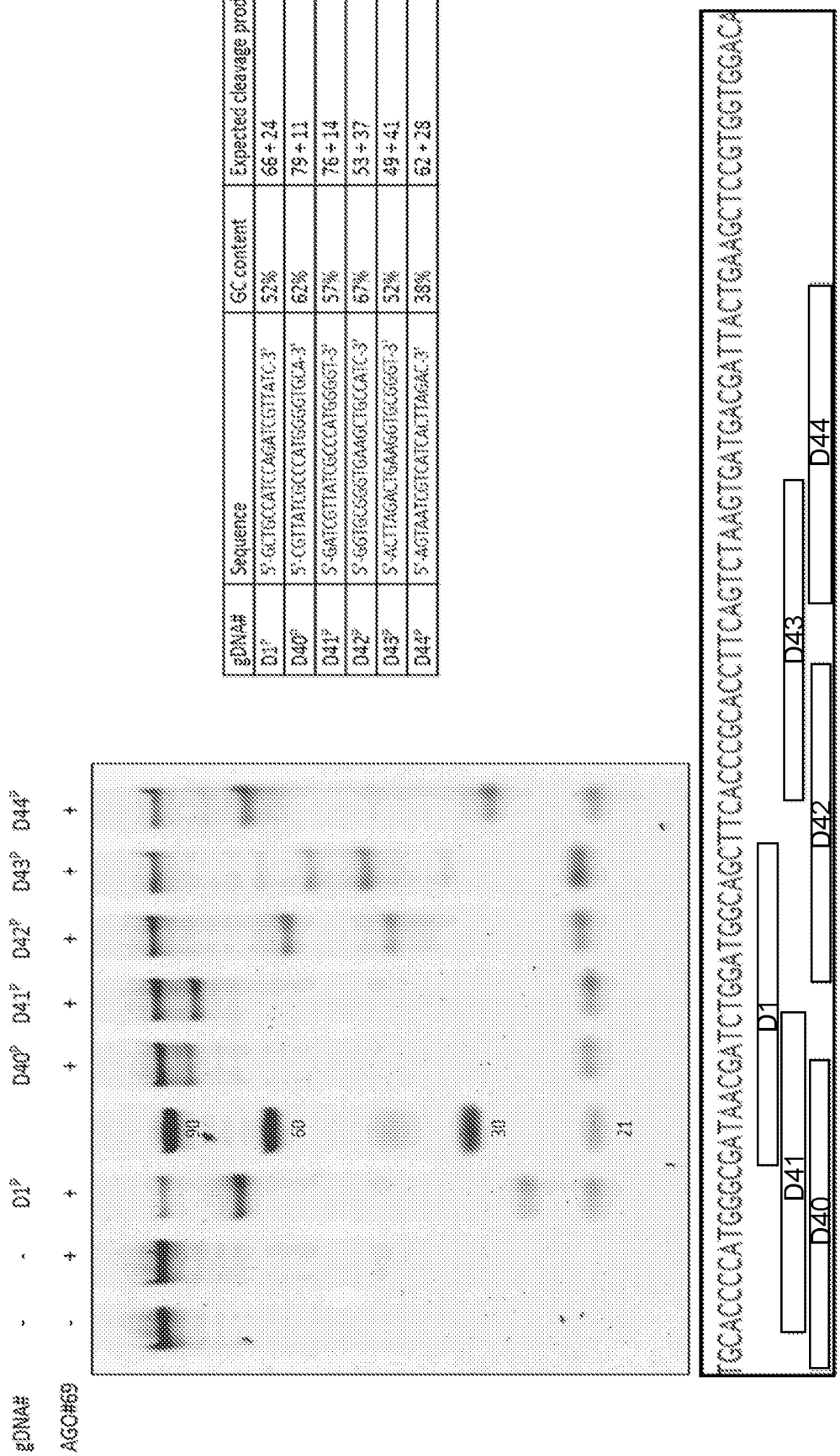
gDNA#	Sequence	GC content	Expected cleavage products
D1[P]	5'-GCTGCCATCCAGATCGTTATC-3'	52%	66 + 24
D40[P]	5'-CGTTATCGCCCATGGGCTGCA-3'	62%	79 + 11
D41[P]	5'-GATCGTTATCGCCCATGGGCT-3'	57%	76 + 14
D42[P]	5'-GGTGCGCAGTGAAGCTTGCCATC-3'	67%	53 + 37
D43[P]	5'-ACTTAGACTGAAGCTGCGGGT-3'	52%	49 + 41
D44[P]	5'-AGTAATCGTCATCACTTAGAC-3'	38%	62 + 28

FIG. 41



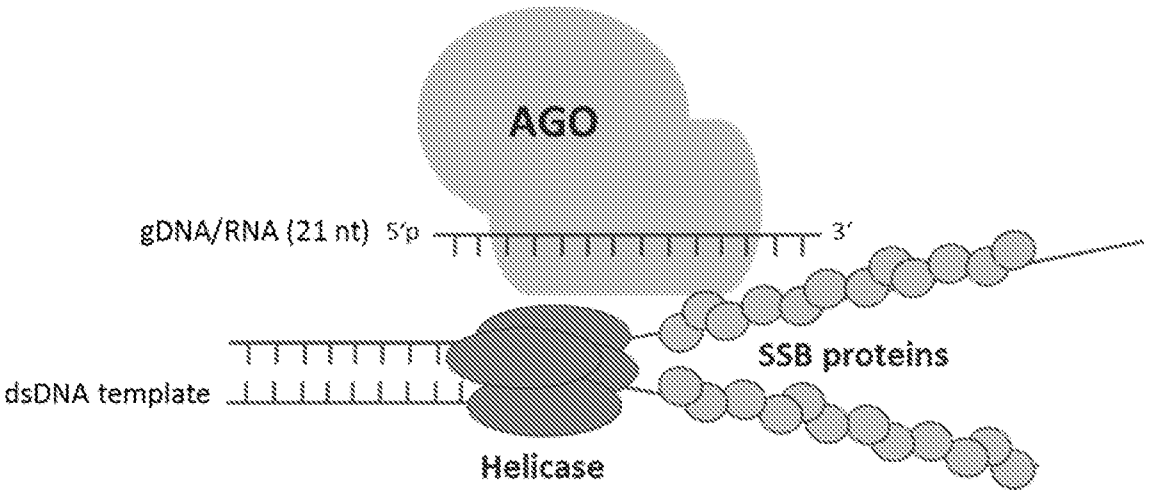
gDNA#	Sequence	GC content	Expected cleavage products
D1[P]	5'-GCTGCCATCCAGATCGTTATC-3'	52%	66 + 24
D40[P]	5'-CGTTATCGCCATGGGGTGCA-3'	62%	79 + 11
D41[P]	5'-GATCGTTATCGCCATGGGGT-3'	57%	76 + 14
D42[P]	5'-CGTTGCGGGTGAAGCTGCCATC-3'	67%	53 + 37
D43[P]	5'-ATTAGACTTGAAGCTGCGGGT-3'	52%	49 + 41
D44[P]	5'-AGTAATCGTCATCACTTAGAC-3'	38%	62 + 28

FIG. 42



gDNA#	Sequence	GC content	Expected cleavage products
D1P	5'-GUTGCCAUCCAGAUCGUUAUC-3'	52%	66 + 24
D40P	5'-CGUUAUGCCCAUGGGGUGCA-3'	62%	79 + 11
D41P	5'-GAUCGUUAUGCCCAUGGGGT-3'	57%	76 + 14
D42P	5'-GGUGCGGGUGAAGCUGCAUC-3'	67%	53 + 37
D43P	5'-ACUUAGACUGAAGGUGCGGT-3'	52%	49 + 41
D44P	5'-AGUAAUCGUCAUCAUUAGAC-3'	38%	62 + 28

FIG. 43



dsDNA nicking/cleavage

FIG. 44

Time (min)	without ET-SSB					
	+ATP			-ATP		
0	502	510	486	376	379	380
5	1105	1015	1056	376	371	373
10	1261	1190	1252	374	370	377
15	1375	1283	1366	382	381	386

Time (min)	with ET-SSB (250 ng/reaction)					
	+ATP			-ATP		
0	708	677	604	404	399	398
5	1741	2184	2096	406	399	396
10	1961	2587	2536	411	406	412
15	2054	2761	2734	426	422	426

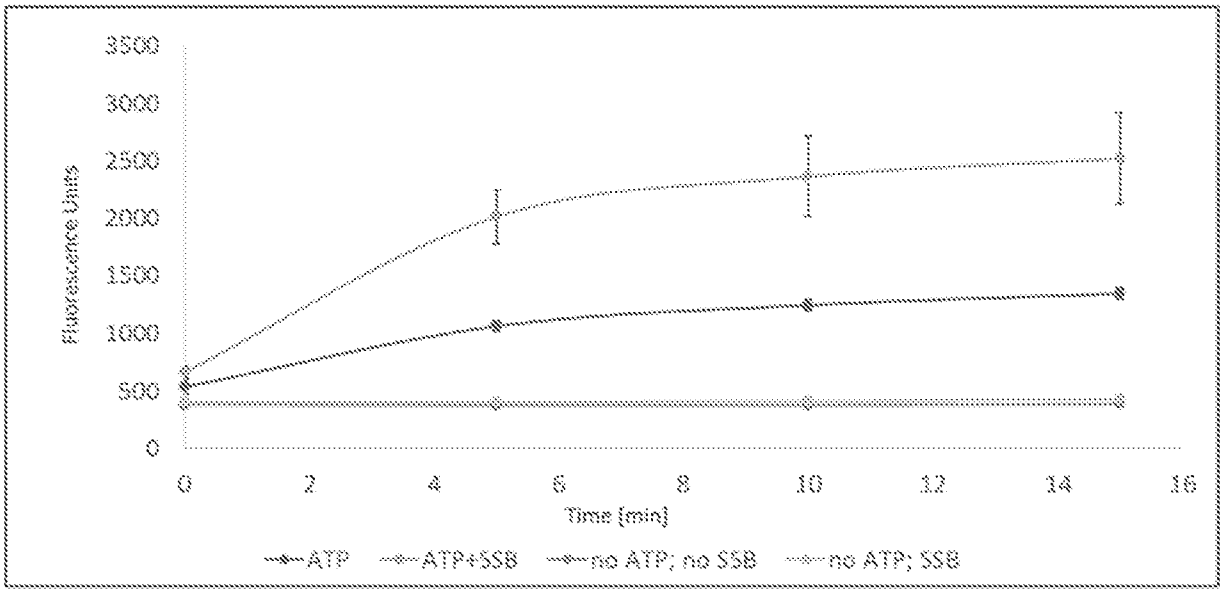
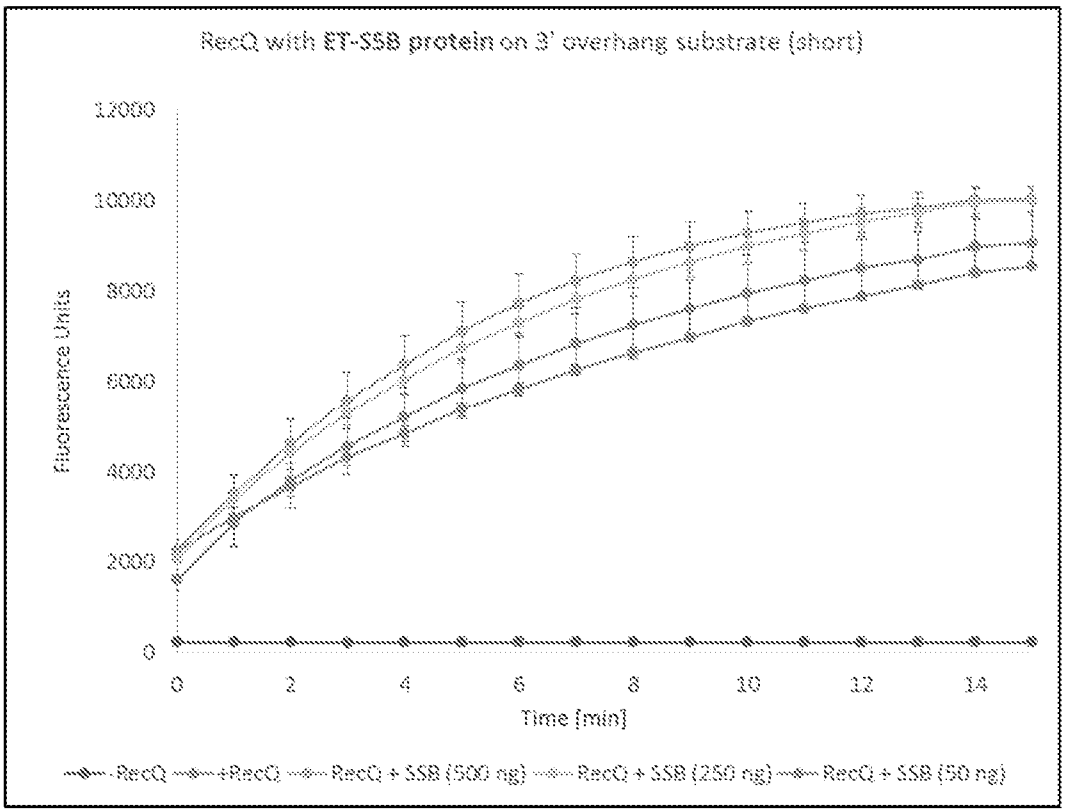
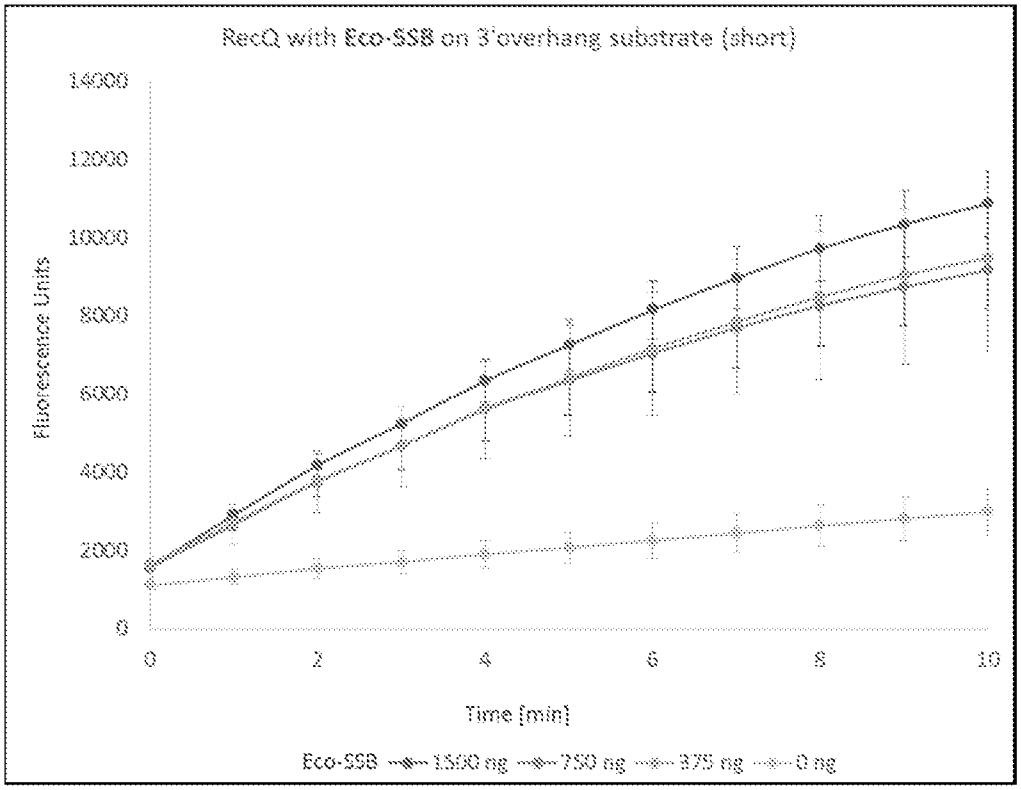


FIG. 45



		ET-SSB + RecQ				ET-SSB + RecQ			
Time [min]	no Helicase	0	500 ng	250 ng	50 ng	no Helicase	0	500 ng	250 ng	50 ng
0	241	2418	2059	2141	1432	224	2058	2414	1983	1784
1	231	3130	3174	3465	2491	217	2839	3794	3255	3205
2	230	3793	4203	4563	3389	217	3515	5000	4238	4198
3	228	4489	5136	5481	4122	211	4186	6026	5112	5003
4	228	4955	5916	6269	4743	214	4713	6822	5813	5662
5	228	5479	6655	6986	5370	217	5270	7568	6500	6302
6	227	5926	7263	7527	5879	215	5714	8187	7056	6823
7	228	6342	7805	8041	6362	217	6170	8654	7582	7305
8	227	6725	8243	8518	6784	215	6544	9037	7987	7700
9	228	7043	8625	8870	7155	216	6905	9378	8403	8067
10	230	7397	8934	9234	7491	217	7259	9630	8743	8400
11	229	7677	9206	9513	7756	220	7560	9835	9022	8688
12	233	7927	9418	9747	8060	220	7827	10080	9300	8960
13	229	8205	9597	9973	8264	220	8067	10280	9506	9139
14	230	8437	9798	10180	8525	223	8253	10222	9736	9414
15	230	8573	9824	10214	8636	220	8517	10240	9770	9498

FIG. 46



Eco-SSB	1500 ng	750 ng	375 ng	0 ng	1500 ng	750 ng	375 ng	0 ng	1500 ng	750 ng	375 ng	0 ng
Time (min)	+RecQ				+RecQ				+RecQ			
0	1613	1646	1639	1129	1576	1780	1668	1219	1586	1390	1627	979
1	3053	2852	2802	1372	2866	3044	2860	1472	2788	2078	2491	1118
2	4489	4085	3969	1647	4098	4309	4041	1716	3954	2842	3307	1258
3	5712	5129	4998	1850	5116	5361	5059	1907	4873	3490	3965	1367
4	6944	6139	6047	2079	6215	6468	6149	2137	5879	4166	4673	1505
5	7983	6960	6893	2256	7078	7365	7054	2330	6724	4717	5295	1624
6	8984	7693	7717	2478	7958	8213	7891	2539	7535	5234	5907	1750
7	9850	8338	8423	2689	8784	9037	8658	2746	8290	5724	6489	1887
8	10647	8960	9083	2901	9536	9715	9400	2962	9003	6137	7061	2027
9	11280	9362	9543	3106	10143	10253	10014	3161	9633	6525	7586	2157
10	11795	9782	9979	3300	10690	10879	10490	3358	10126	6875	8020	2296
11	#Sat	10108	10336	3500	11215	11324	10977	3568	10617	7193	8444	2439
12	#Sat	10365	10652	3691	11556	11610	11282	3731	10927	7398	8722	2543
13	#Sat	10594	10917	3866	11805	11825	11591	3916	11265	7677	9081	2678
14	#Sat	10856	11178	4076	12171	12237	11900	4124	11568	7906	9365	2832
15	#Sat	11043	11423	4275	#Sat	#Sat	#Sat	4324	11926	8193	9706	2981

FIG. 47

Lane	1	2	3	4	5	6	7	8	9	10	11	12
AGO41	WT	WT	WT	WT	MUT	MUT	MUT	MUT	-	-	-	-
Guide	-	D1	D1(p)	DNT(p)	-	D1	D1(p)	DNT(p)	-	D1	D1(p)	DNT(p)
Temp 1	+	+	+	+	+	+	+	+	+	+	+	+

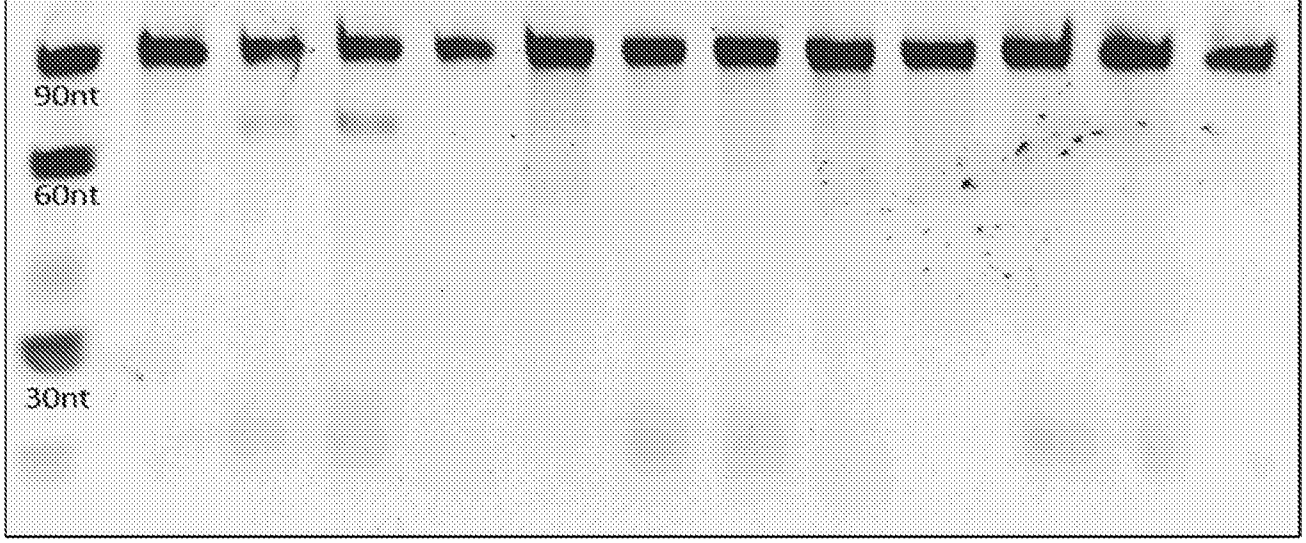


FIG. 48

AGO#69

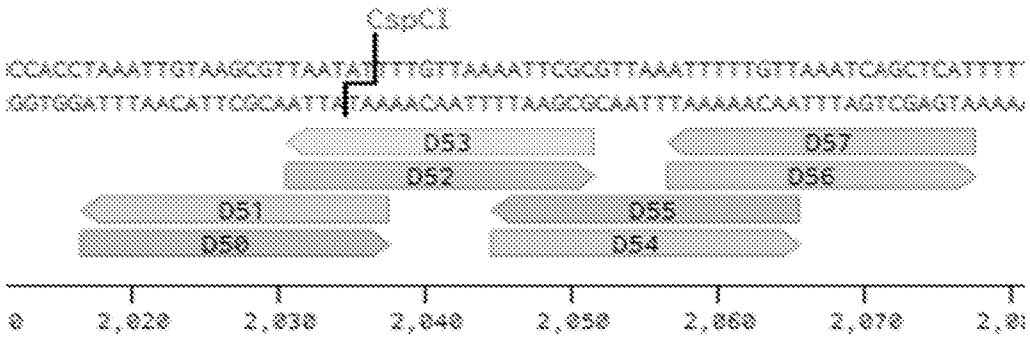
MPKKKRKVEDPKKKRKVGSGSMVGGYKVSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDN
VYKIIKSARYKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMS
IYGEVVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTRELKDEFKKSNKGFNLKRK
FRISPVVNKMGKVILYLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVIESV
LETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIYMYIPHALKPI
ITREYLAKNDPEFSKEIEQLIKMNMNYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVKL
LGYSSGKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAIRAI
YDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGDITEYKKAALKLNNYNNVDFV
IAIVPNMSDEEIENSYNPFKKIWAELNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGI
LGKIGGIPWVVKDMKGDVDCFVGL▓VGTREKGIHYPACSVVFDKYGKLINYYKPNIPQNG
▓KINTEILQEIFDKVLISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKNIKFNIIE
VKKSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLKIEKSYGDIDM
LTALSQIYALTQIHVGATKSLRLPITTGYA▓KICKAIEFIPQGRVDNRLFFL

FIG. 49

AGO#69

MPKKKRKVEDPKKKRKVGSGSMVGGYKVSNLTVEAFEGIGSVNPMLFYQYKVTGKGKYDN
VYKIIKSARYKMHSKNRFKPVFIKDDKLYTLEKLPDIEDLDFANINFVKSEVLSIEDNMS
IYGEVVEYYINLKLKKVKVLGKYPKYRINYSKEILSNTLLTRELKDEFKKSNKGFNLKRK
FRISPVVNKMGKVILYLSCSADFSTNKNIYEMLKEGLEVEGLAVKSEWSNISGNLVIESV
LETKISEPTSLGQSLIDYYKNNNQGYRVKDFTDEDLNANIVNVRGNKKIYMYIPHALKPI
ITREYLAKNDPEFSKEIEQLIKMNMNYRYETLKSFVNDIGVIEELNNLSFKNKYYEDVKL
LGYSSGKIDEPVLMGAKGIIKNKMQIFSNGFYKLPEGKVRFGVLYPKEFDGVSRKAIRAI
YDFSKEGKYHGESNKYIAEHLINVEFNPKECIFEGYELGDITEYKKAALKLNNYNNVDFV
IAIVPNMSDEEIENSYNPFKKIWAELNLPSQMISVKTAEIFANSRDNTALYYLHNIVLGI
LGKIGGIPWVVKDMKGDVDCFVGLDVGTREKGIHYPACSVVFDKYGKLINYYKPNIPQNG
EKINTEILQEIFDKVLISYEEENGAYPKNIVIHRDGFSREDLDWYENYFGKKNIKFNIIE
VKKSTPLKIASINEGNITNPEKGSYILRGNKAYMVTTDIKENLGSPKPLKIEKSYGDIDM
LTALSQIYALTQIHVGATKSLRLPITTGYADKICKAIEFIPQGRVDNRLFFL

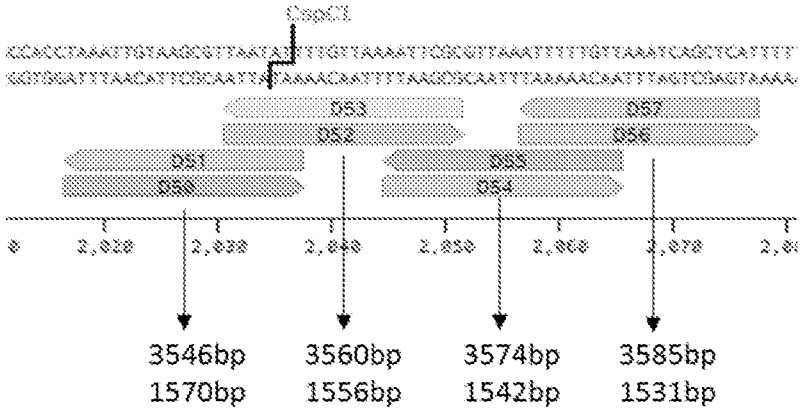
FIG. 50



D51 – GC 19.05% - Tm 43.2°C
D53 – GC 19.05% - Tm 43.2°C
D55 – GC 23.81% - Tm 48.2°C
D57 – GC 23.81% - Tm 46.2°C

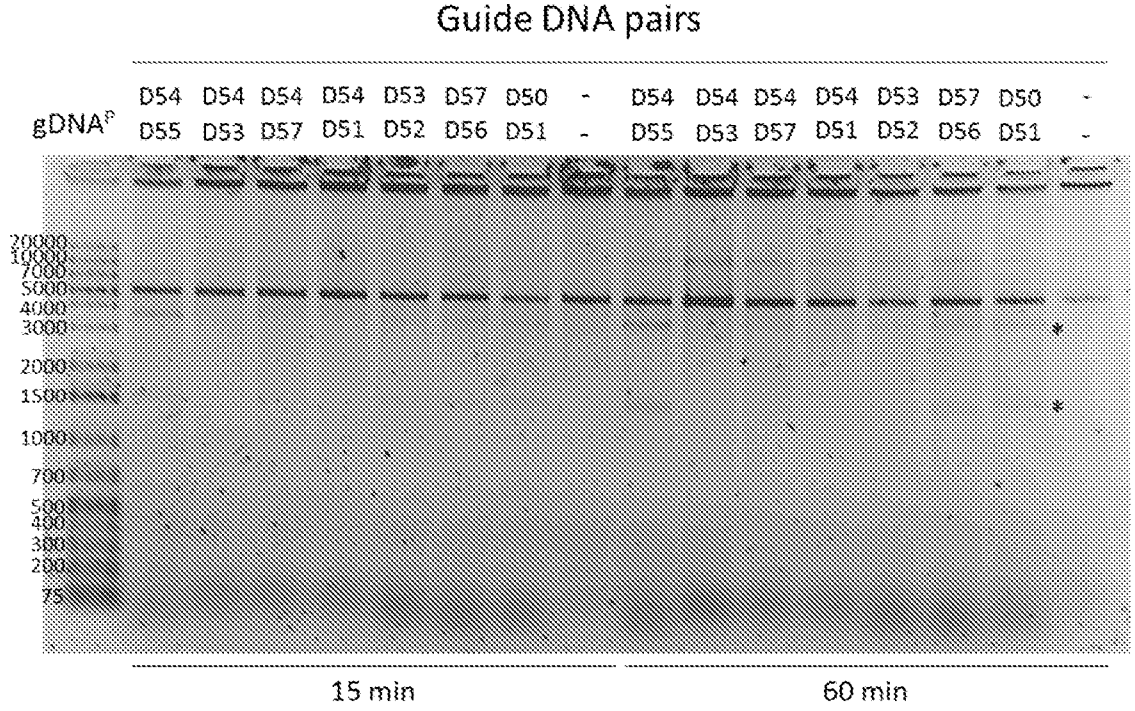
Guide DNA location and Tm

FIG. 51



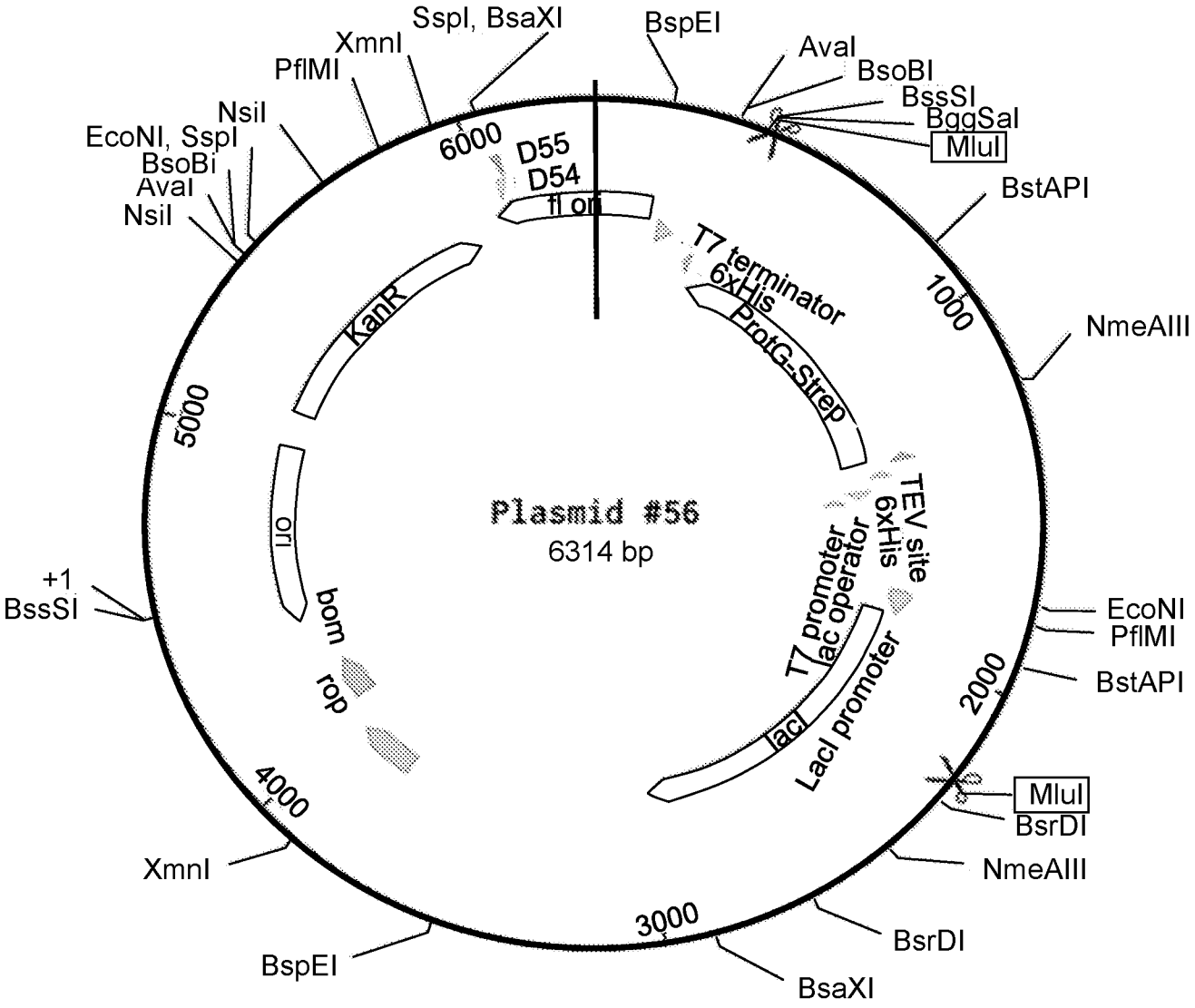
Expected cleavage product size

FIG. 52A



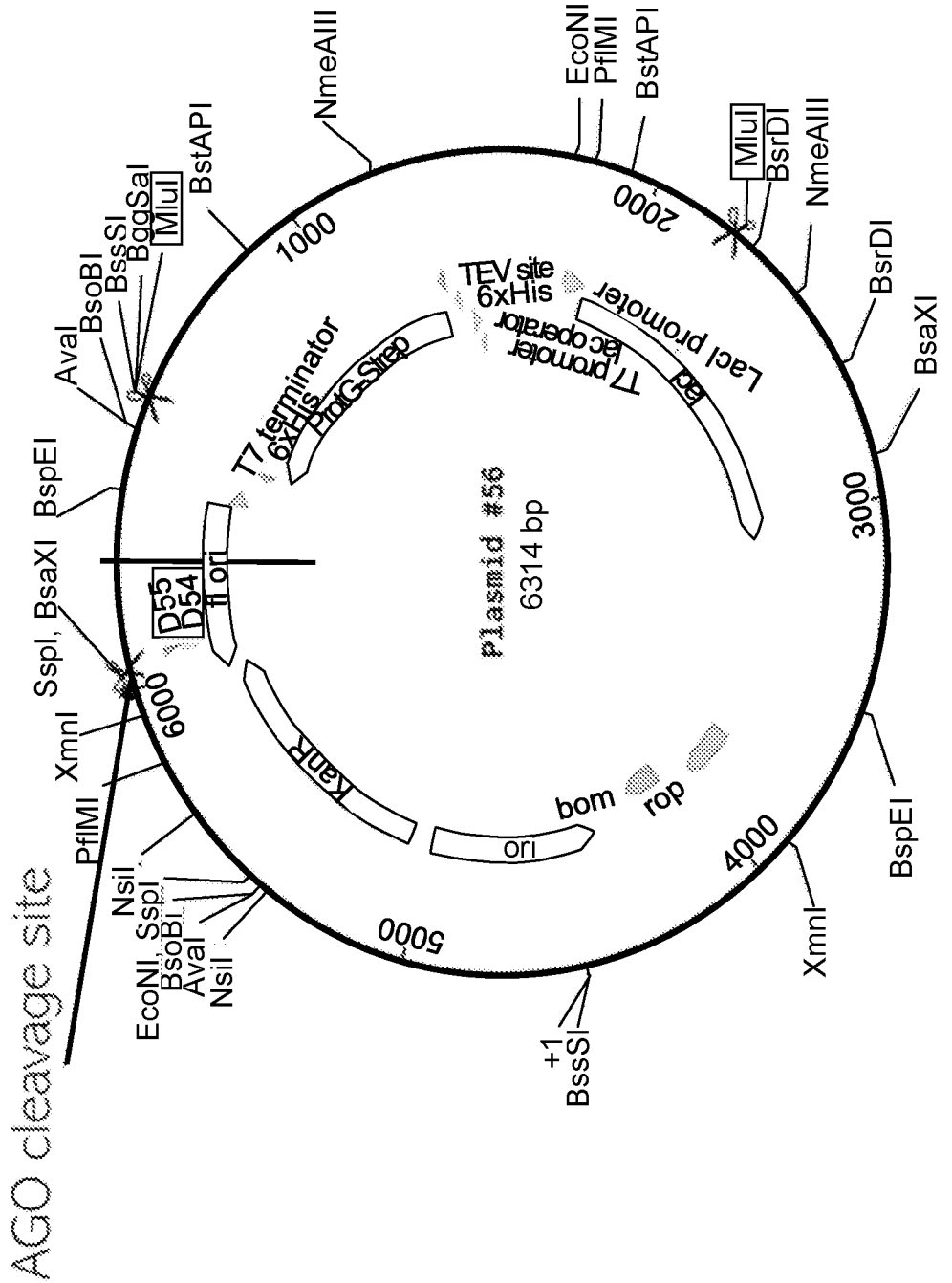
Incubation time

FIG. 52B



Expected cleavage product: 4487 + 1827

FIG. 53A



Expected cleavage product: 3816 + 1827 + 671

FIG. 53B

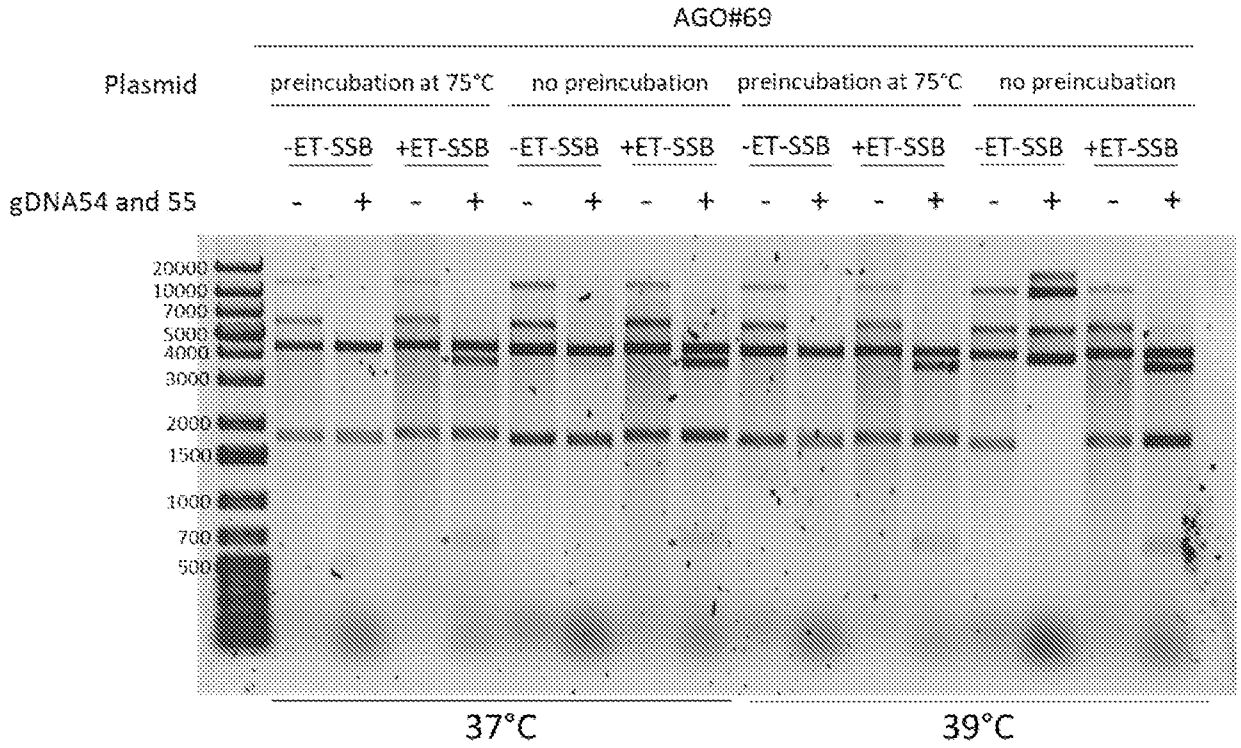

AGO#69
Plasmid
preincubation at 75°C
no preincubation
preincubation at 75°C
no preincubation
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
gDNA54 and 55
- + - + - + - + - + - + - + - +
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
37°C
39°C

FIG. 54

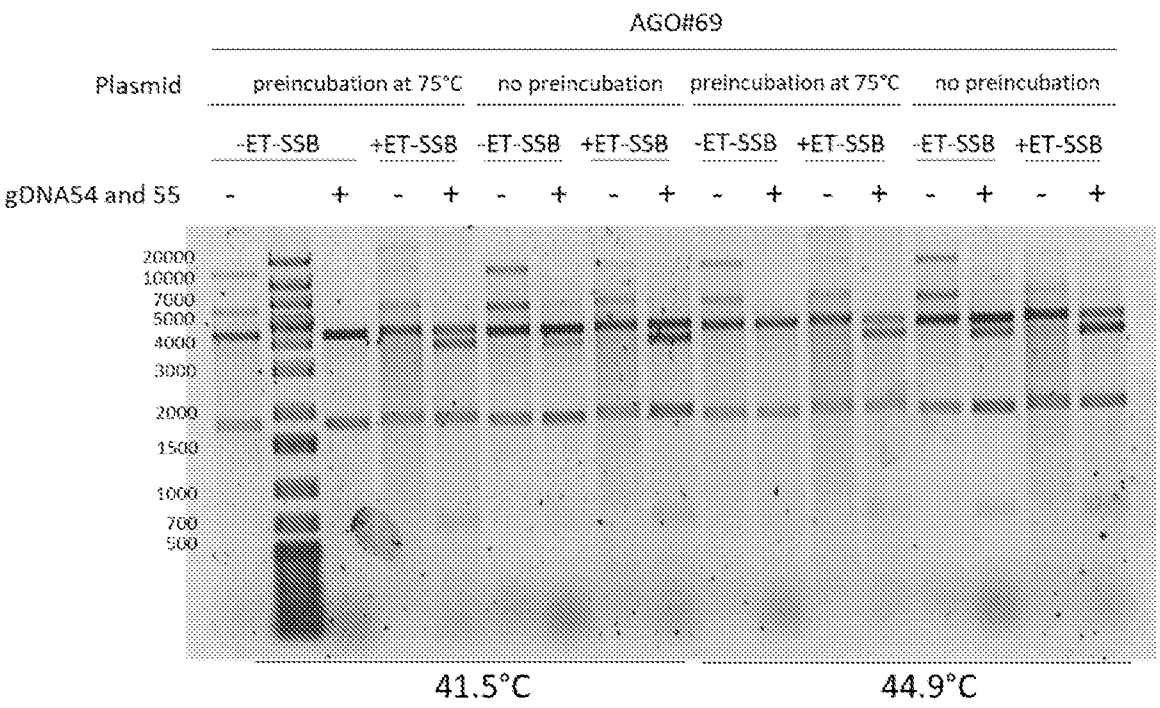

AGO#69
Plasmid
preincubation at 75°C
no preincubation
preincubation at 75°C
no preincubation
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
gDNA54 and 55
- + - + - + - + - + - + - + - +
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
41.5°C
44.9°C

FIG. 55

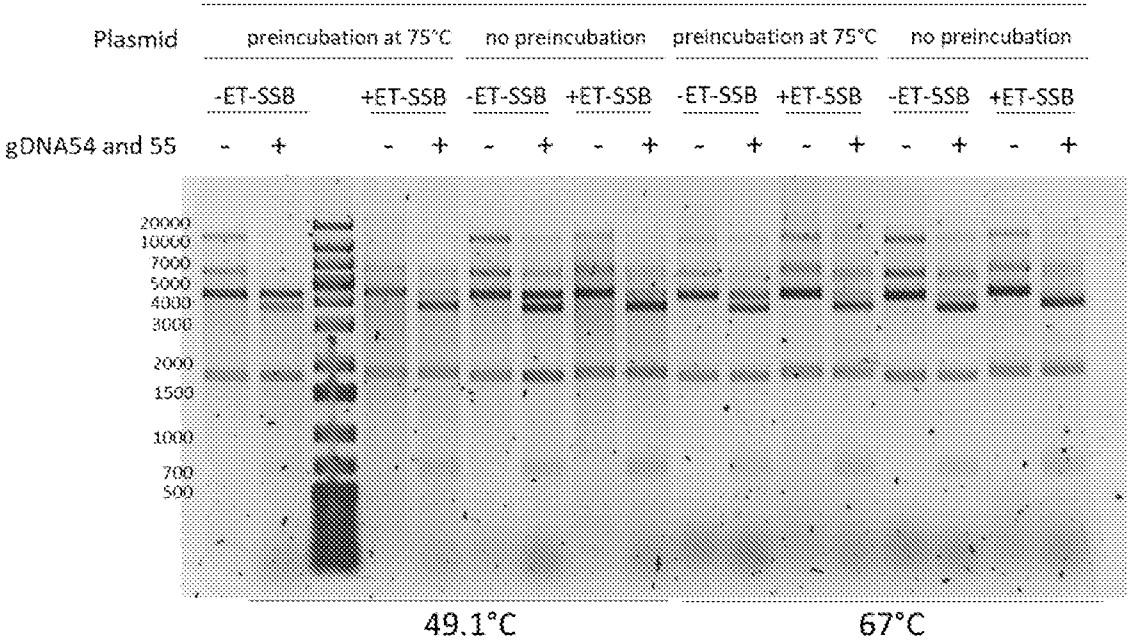
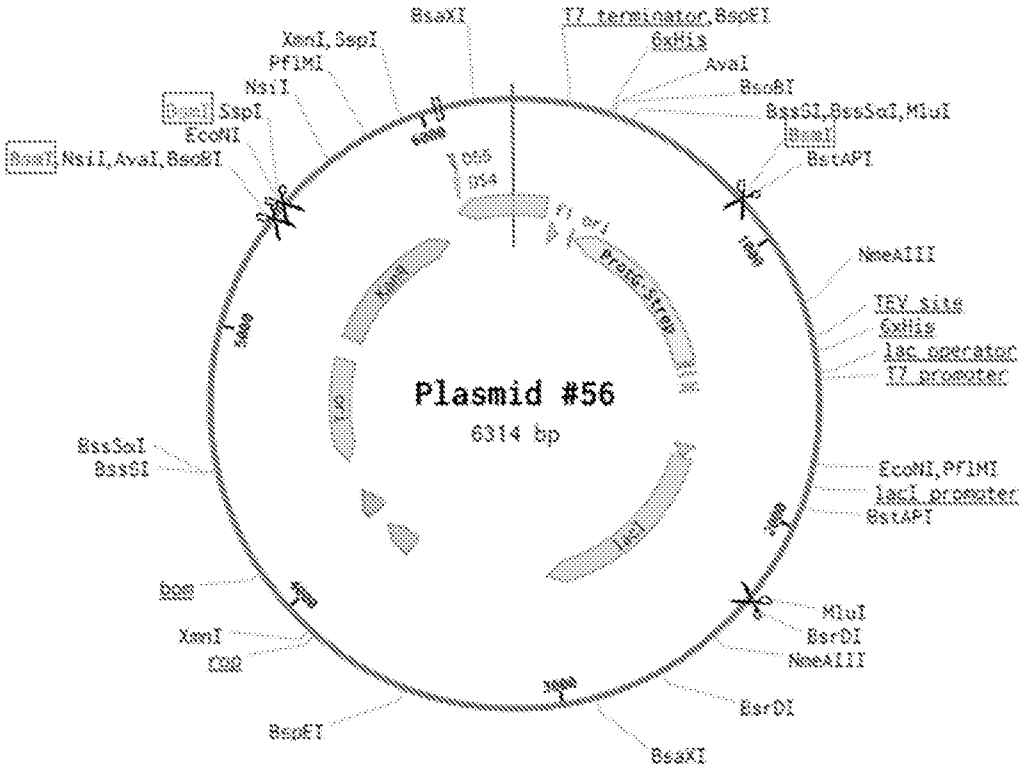
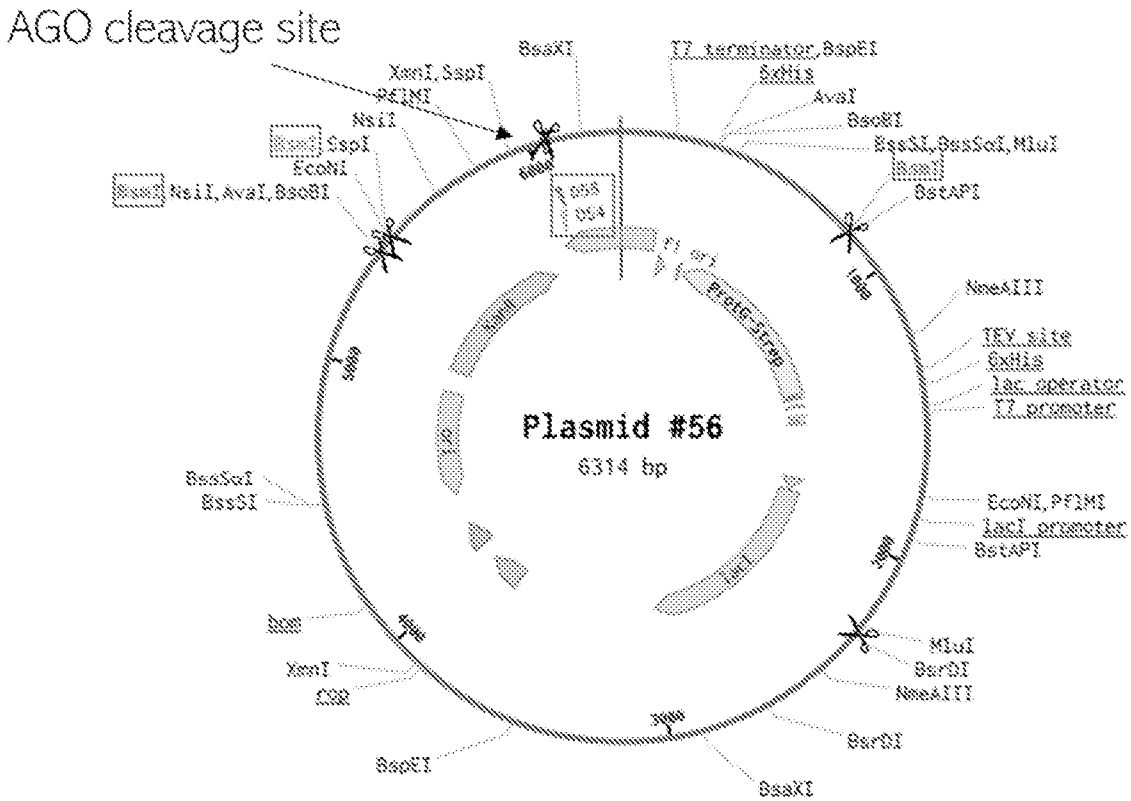

Plasmid
preincubation at 75°C
no preincubation
preincubation at 75°C
no preincubation
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
-ET-SSB
+ET-SSB
gDNA54 and 55
- + - + - + - + - + - + - + - +
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
49.1°C
67°C

FIG. 56



Expected cleavage product: 4596 + 1641 + 77

FIG. 57A



Expected cleavage product: 4596 + 1089 + 552 + 77

FIG. 57B

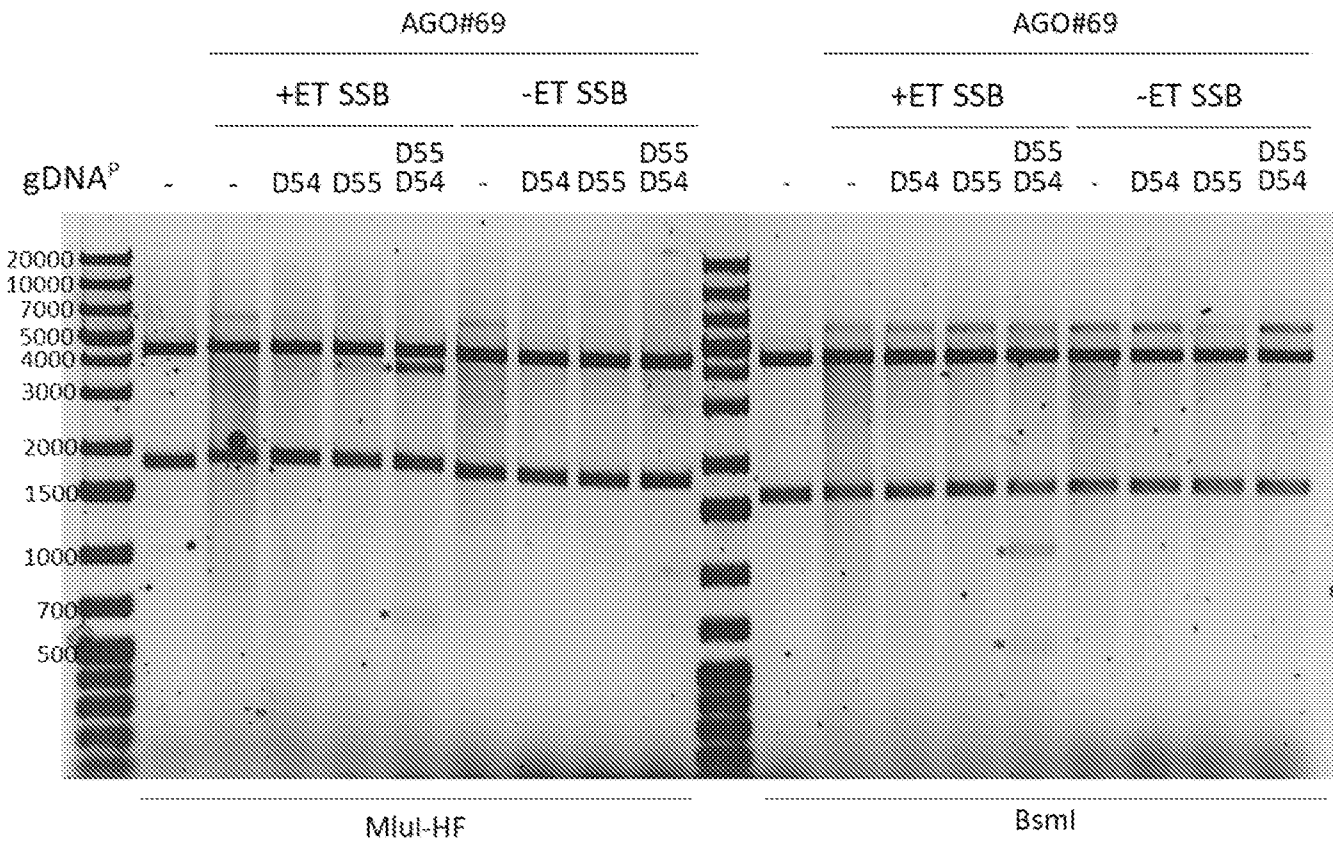

AGO#69
AGO#69
+ET SSB
-ET SSB
+ET SSB
-ET SSB
gDNA^P
- - D54 D55 D55 D54 - D54 D55 D55 D54 - - D54 D55 D55 D54 - D54 D55 D55 D54
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
MluI-HF
BsmI

FIG. 58

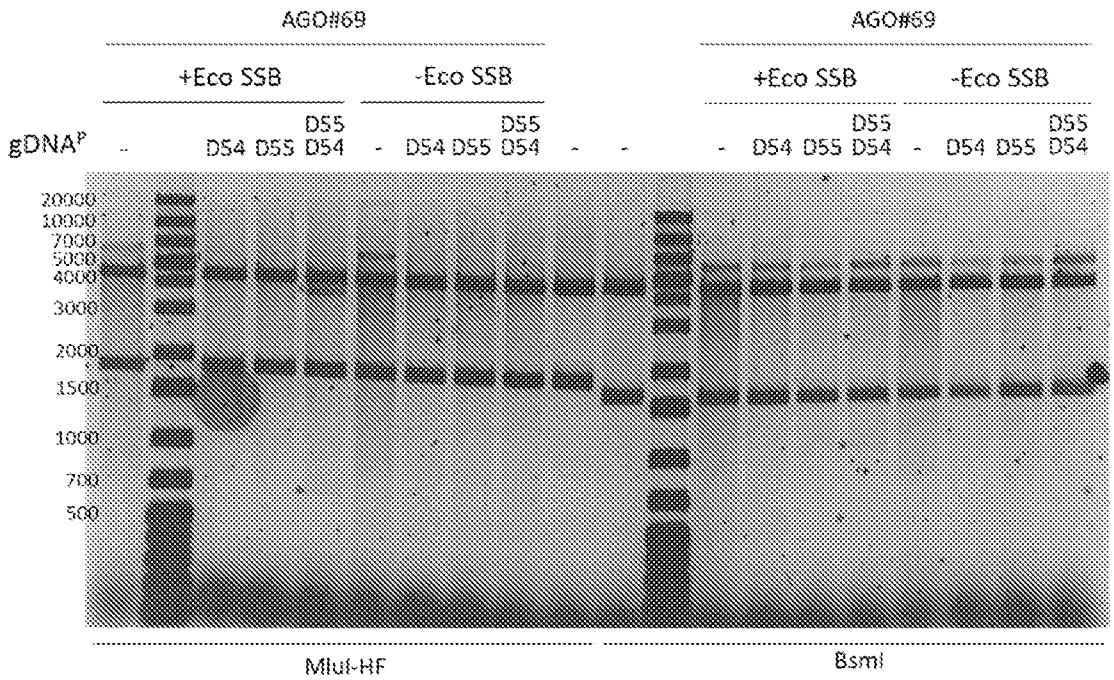
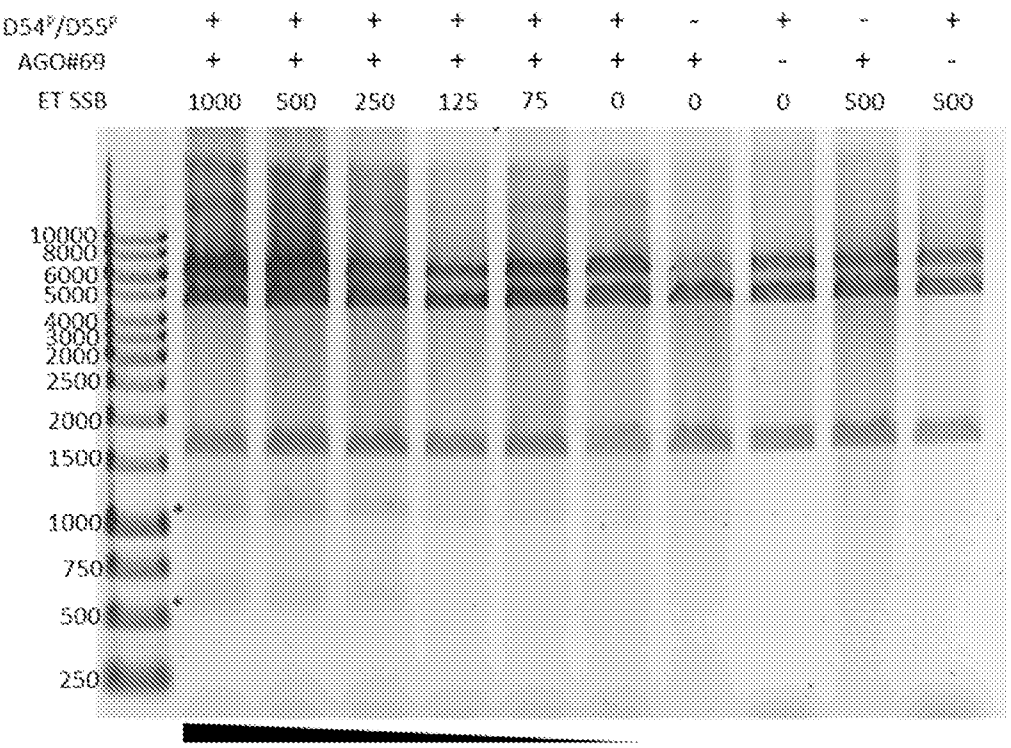
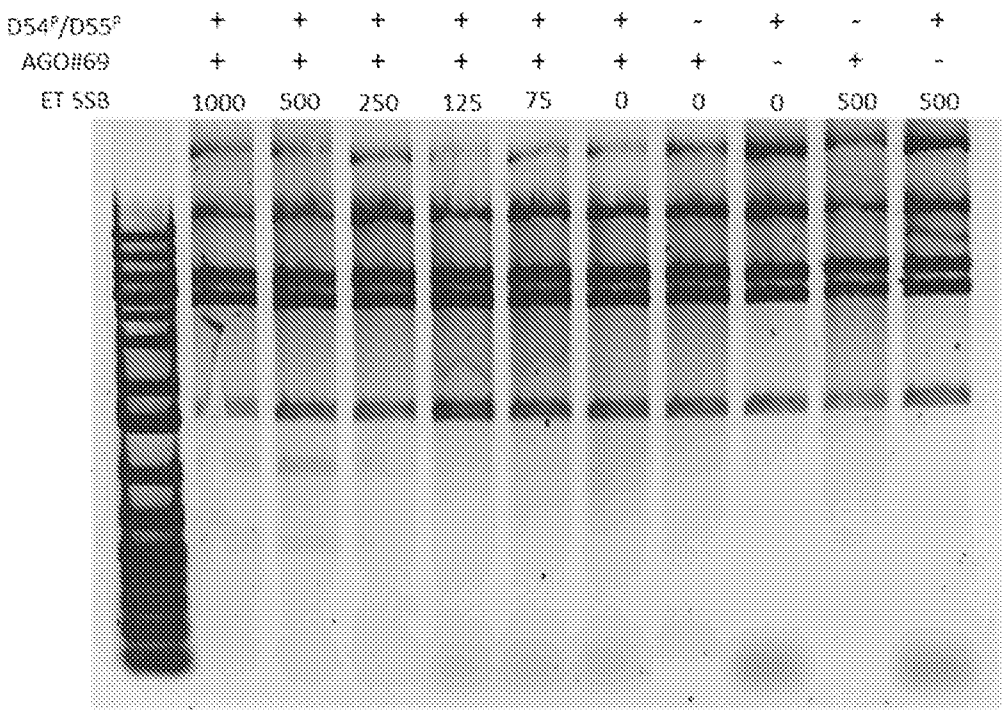

AGO#69
AGO#69
+Eco SSB
-Eco SSB
+Eco SSB
-Eco SSB
gDNA
D55
D54 D55 D54
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
MluI-HF
BsmI

FIG. 59



Expected cleavage product: 4596 + 1089 + 552 + 77

FIG. 60



Expected cleavage product: 4596 + 1089 + 552 + 77

FIG. 61

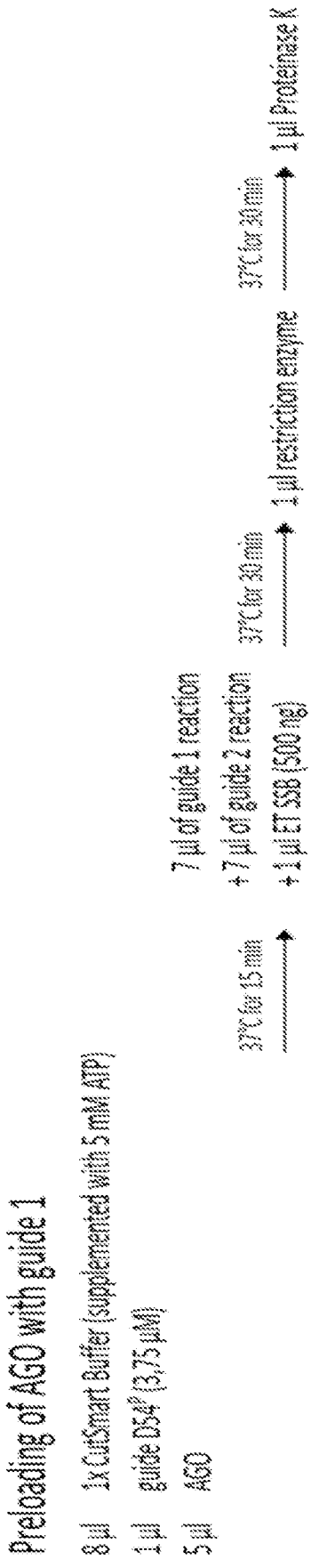

Preloading of AGO with guide 1
8 µl 1x CutSmart Buffer (supplemented with 5 mM ATP)
1 µl guide DS4P (3,75 µM)
5 µl AGO
37°C for 15 min
7 µl of guide 1 reaction
+ 7 µl of guide 2 reaction
+ 1 µl ET SSB (500 ng)
+ 1 µl Tte UvrD (20 ng)
+ 1 µl Plasmid (250-300 ng)
37°C for 30 min
1 µl restriction enzyme
37°C for 30 min
1 µl Proteinase K
Preloading of AGO with guide 2
8 µl 1x CutSmart Buffer (supplemented with 5 mM ATP)
1 µl guide DS3P (3,75 µM)
5 µl AGO

FIG. 62

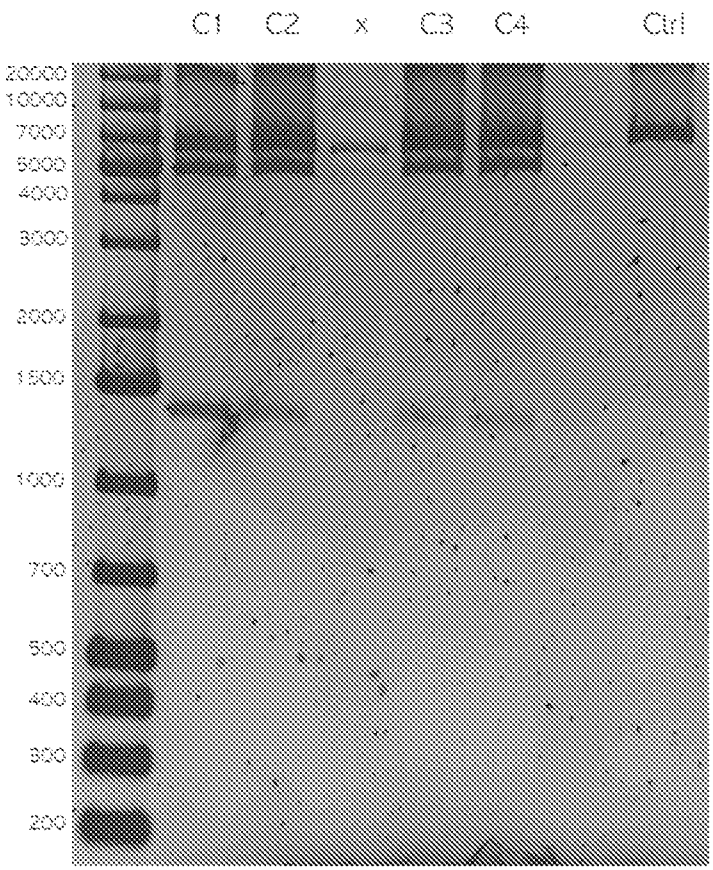

C1
C2
x
C3
C4
Ctrl
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
400
300
200

FIG. 63

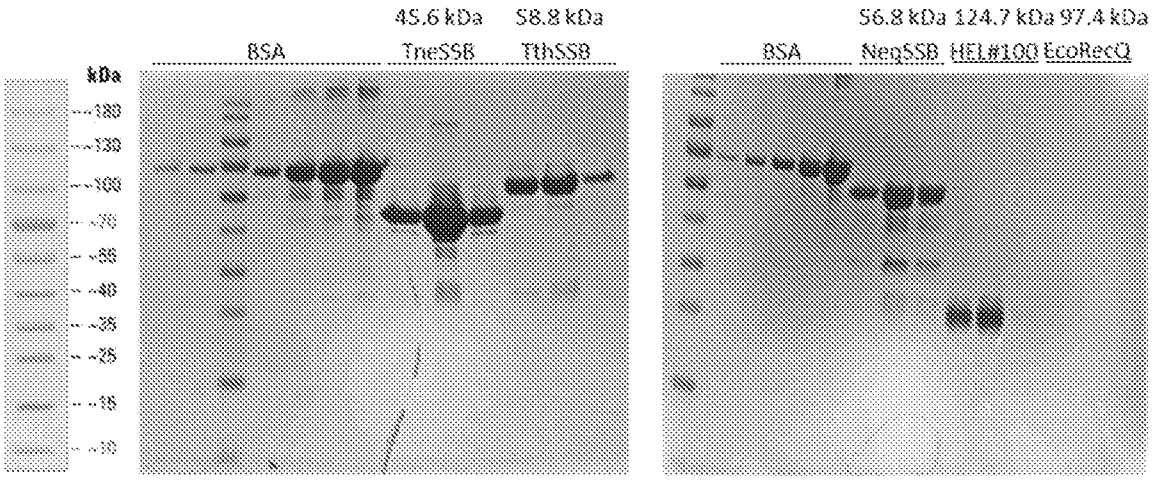


FIG. 64A

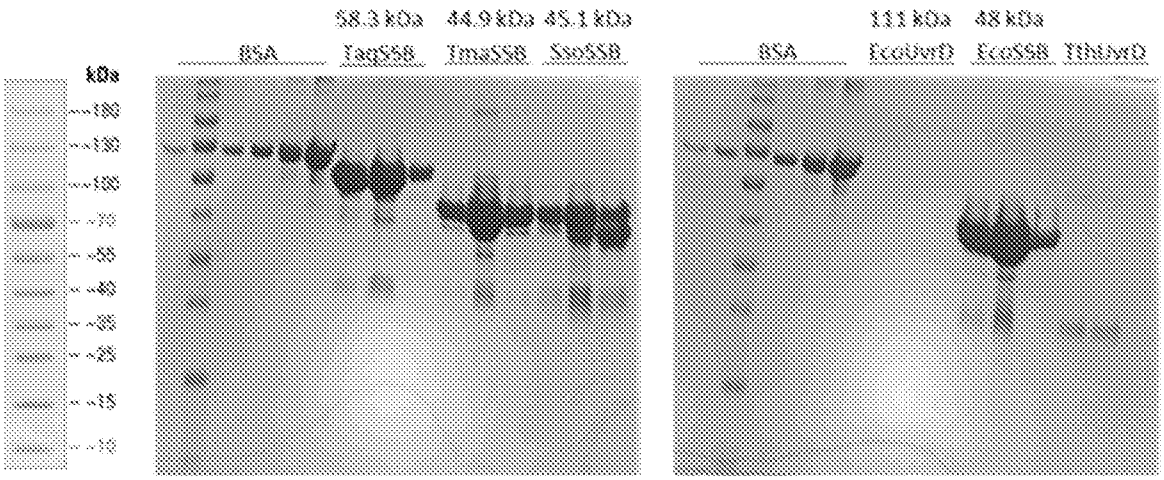


FIG. 64B

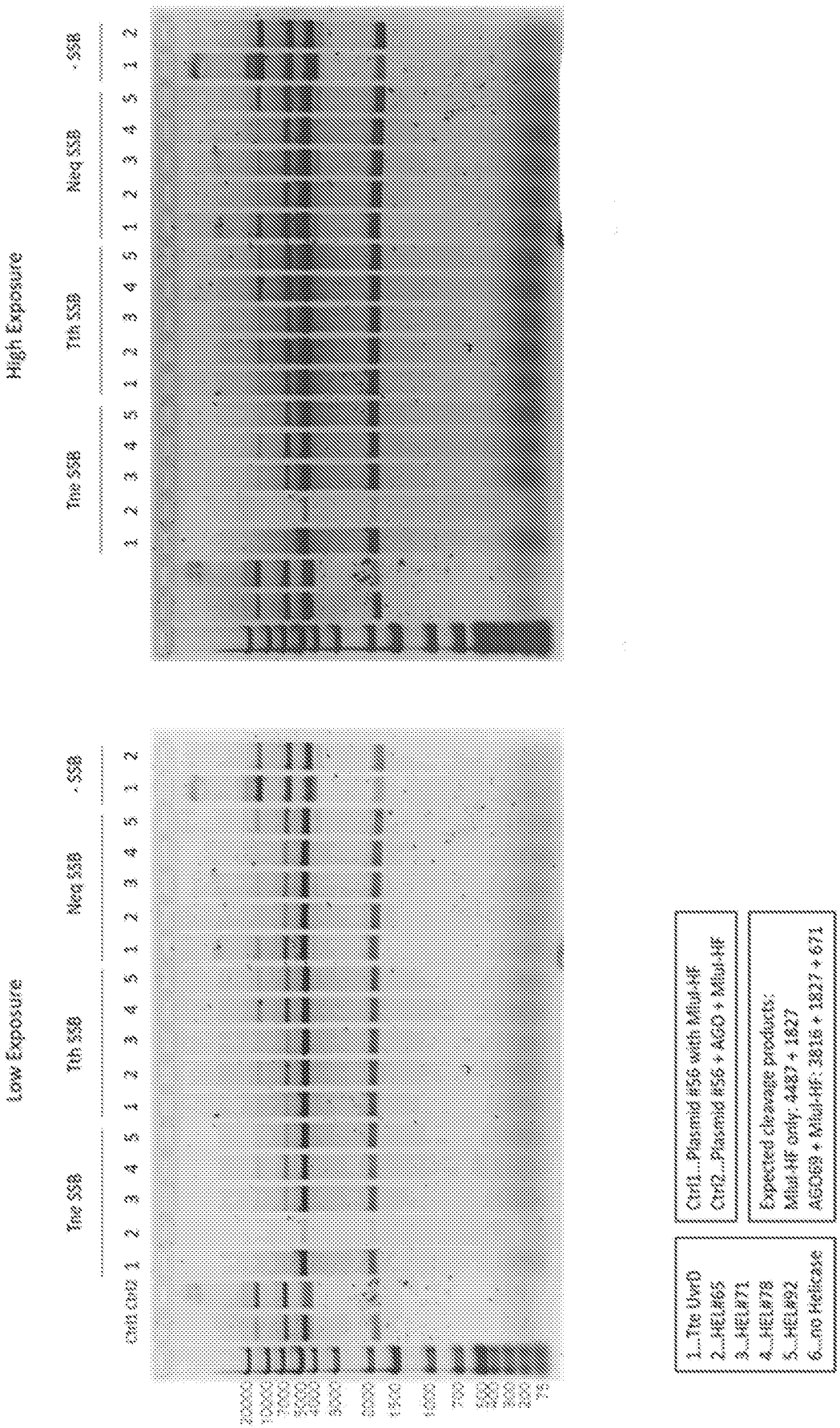

High Exposure
Low Exposure
Tne SSB
Tth SSB
Neq SSB
-SSB
Ctrl1 Ctrl2
1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2
1...Tne UvrD
2...HEL#65
3...HEL#71
4...HEL#78
5...HEL#92
6...no Helicase
Ctrl1...Plasmid #56 with MluI-HF
Ctrl2...Plasmid #56 + AGO + MluI-HF
Expected cleavage products:
MluI-HF only: 4487 + 1827
AGO69 + MluI-HF: 3818 + 1827 + 671

FIG. 65

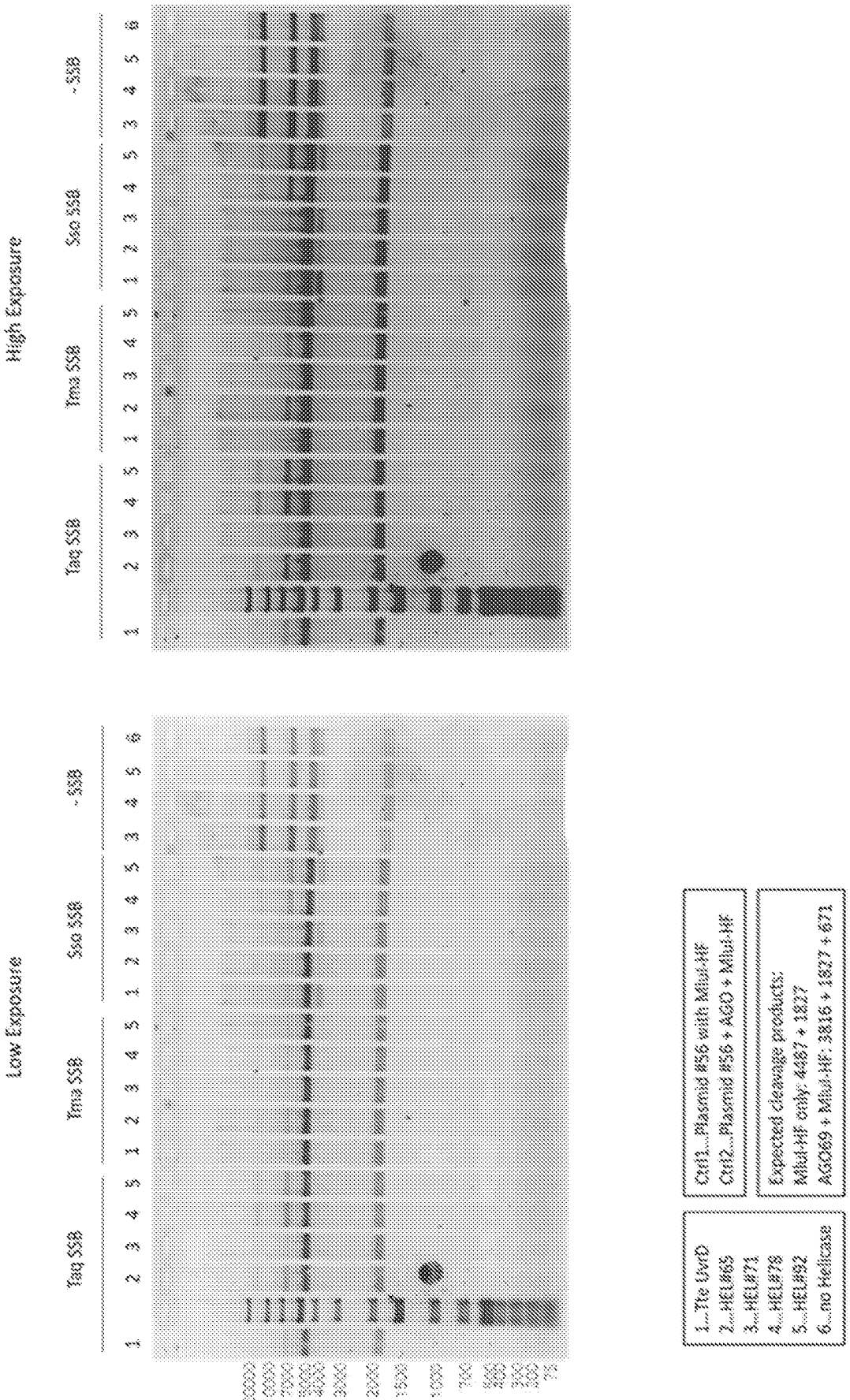

High Exposure
Low Exposure
Taq SSB
Tma SSB
Sso SSB
-SSB
1...No Unwinding
2...HEL#65
3...HEL#71
4...HEL#78
5...HEL#92
6...no Helicase
Ctrl1...Plasmid #56 with Mlu1-HF
Ctrl2...Plasmid #56 + AGO + Mlu1-HF
Expected cleavage products:
Mlu1-HF only: 4487 + 1827
AGO69 + Mlu1-HF: 3816 + 1827 + 671

FIG. 66

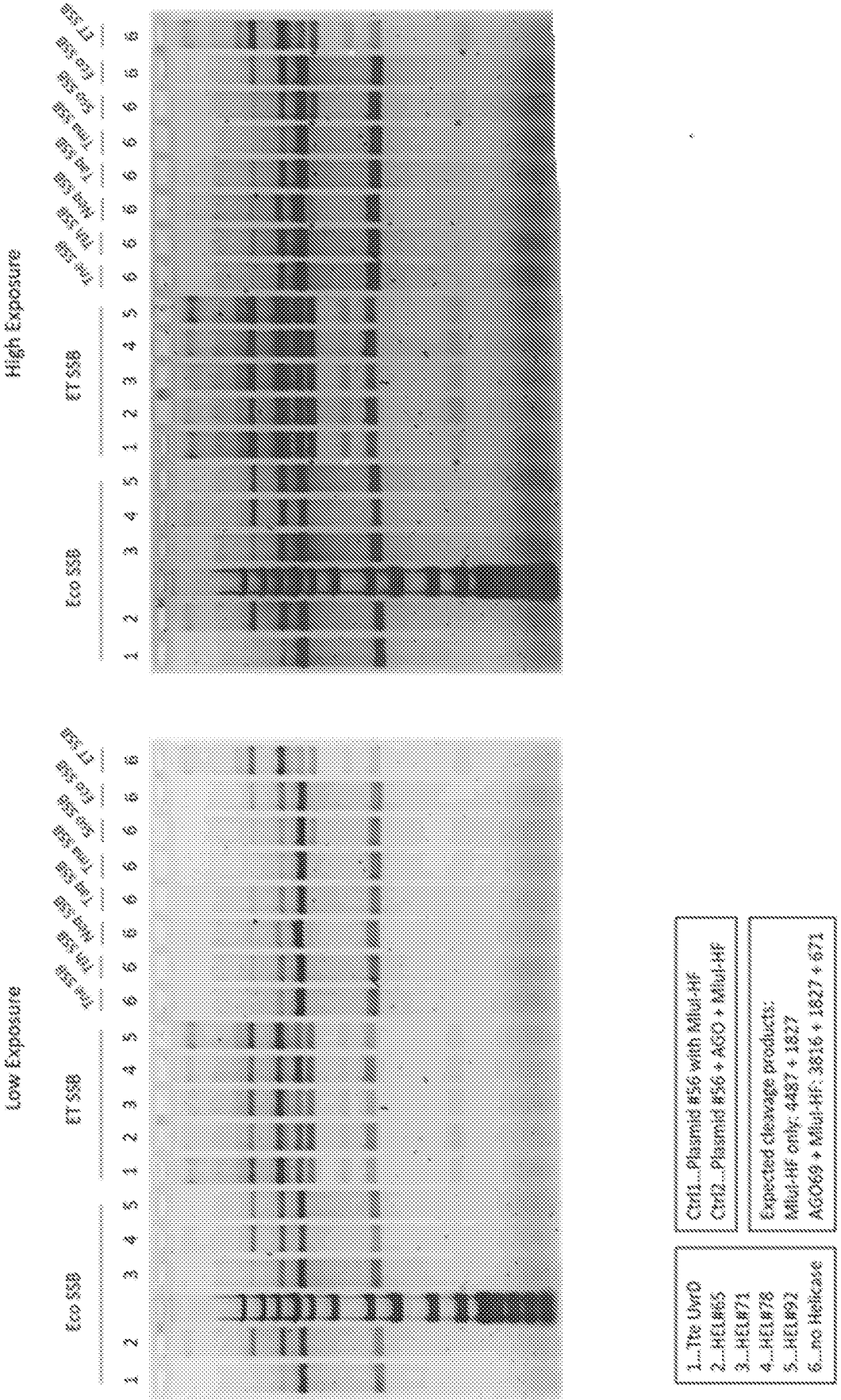

Low Exposure
High Exposure
Eco SSB
ET SSB
1...Tte UvrD
2...HEL885
3...HEL871
4...HEL878
5...HEL892
6...no Helicase
Ctrl1...Plasmid #56 with MluI-HF
Ctrl2...Plasmid #56 + AGO + MluI-HF
Expected cleavage products:
MluI-HF only: 4487 + 1827
AGO69 + MluI-HF: 3816 + 1827 + 671

FIG. 67

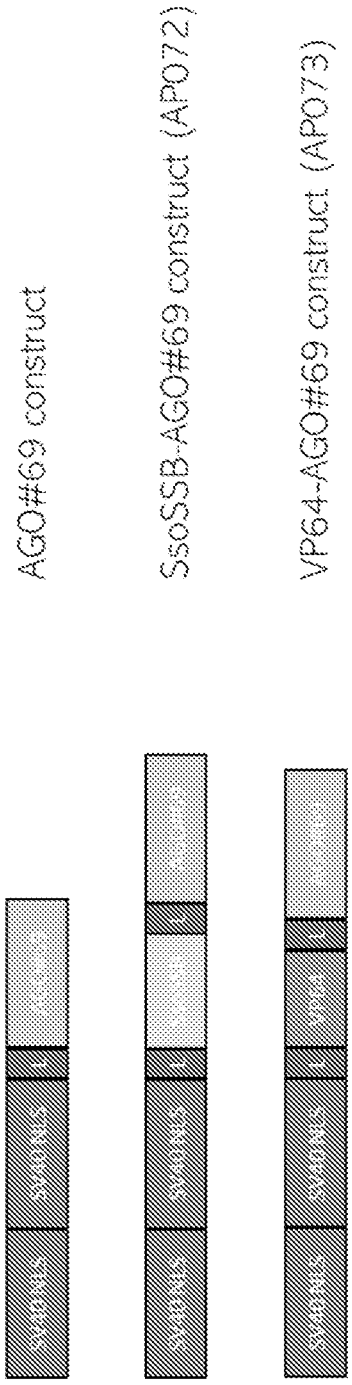

AGO#69 construct
SsoSSB-AGO#69 construct (AP072)
VP64-AGO#69 construct (AP073)

FIG. 68

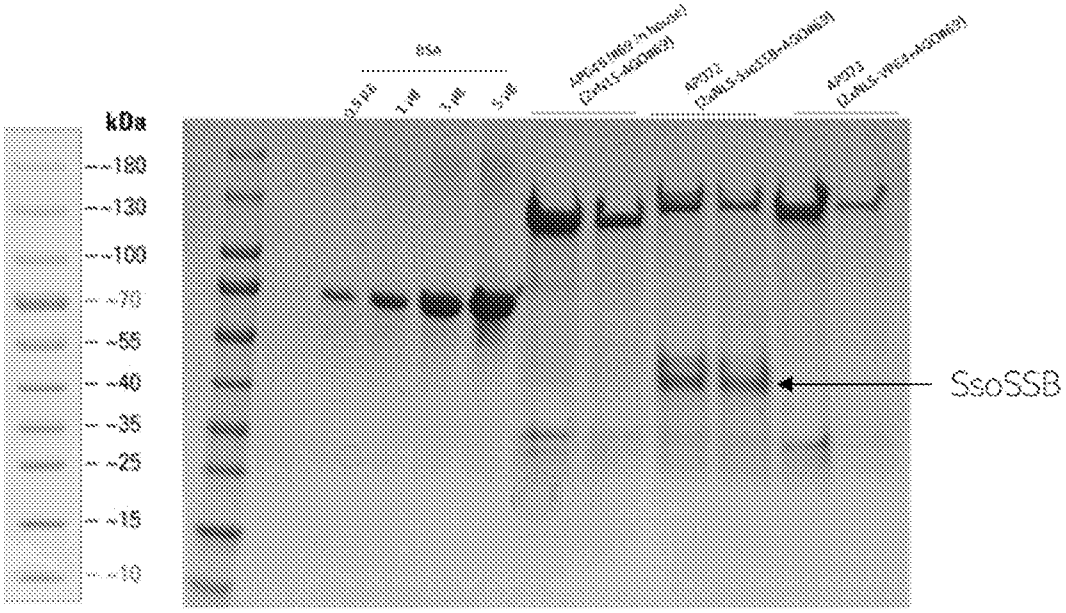


FIG. 69A

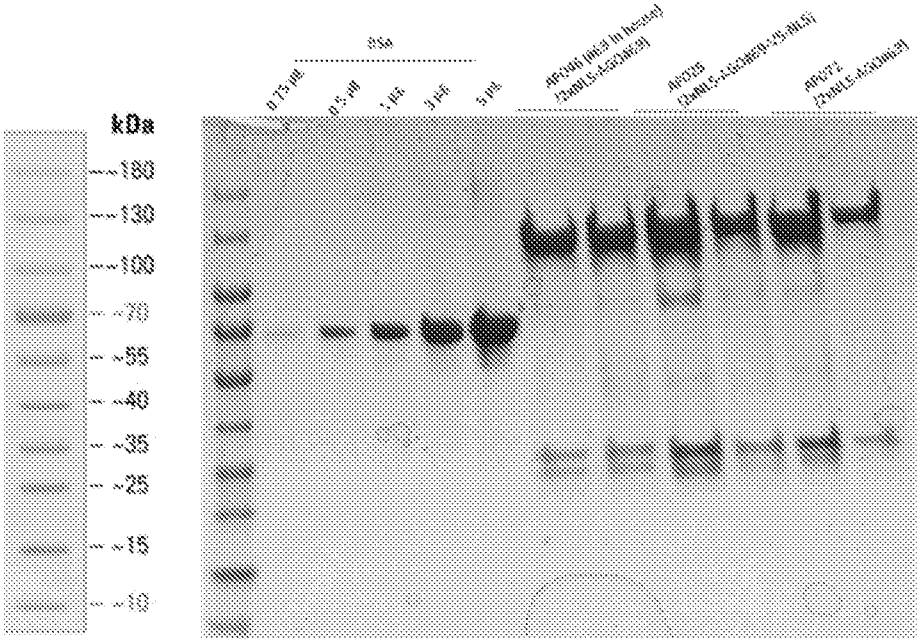
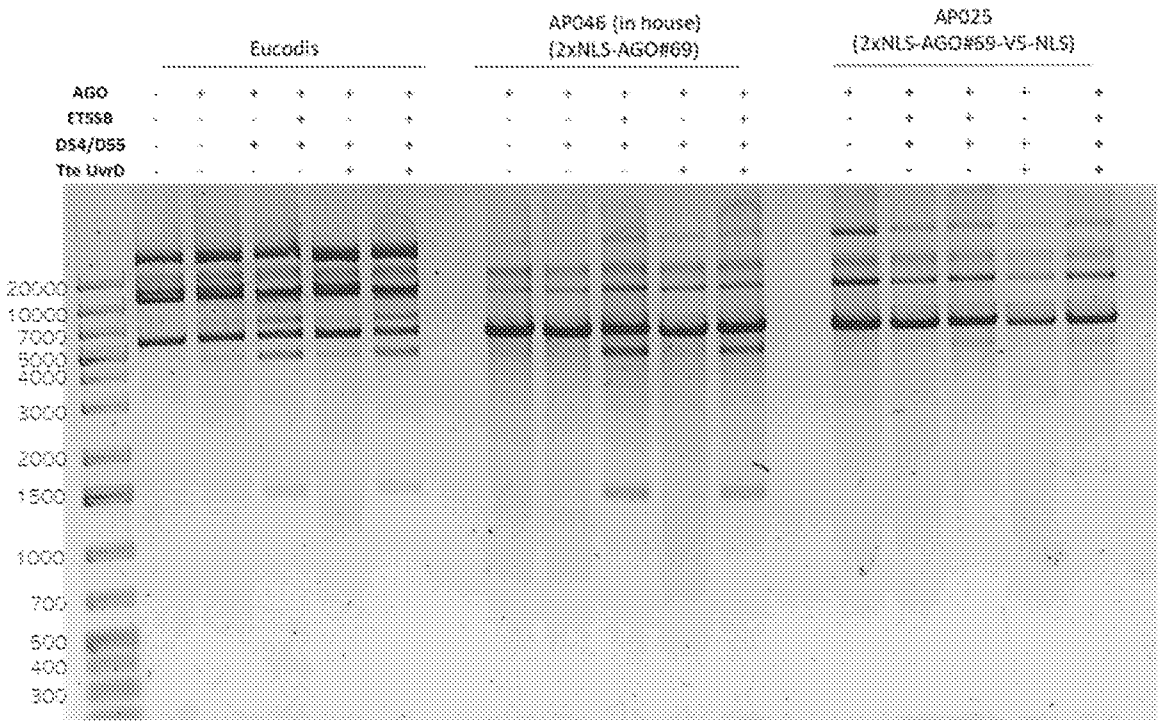
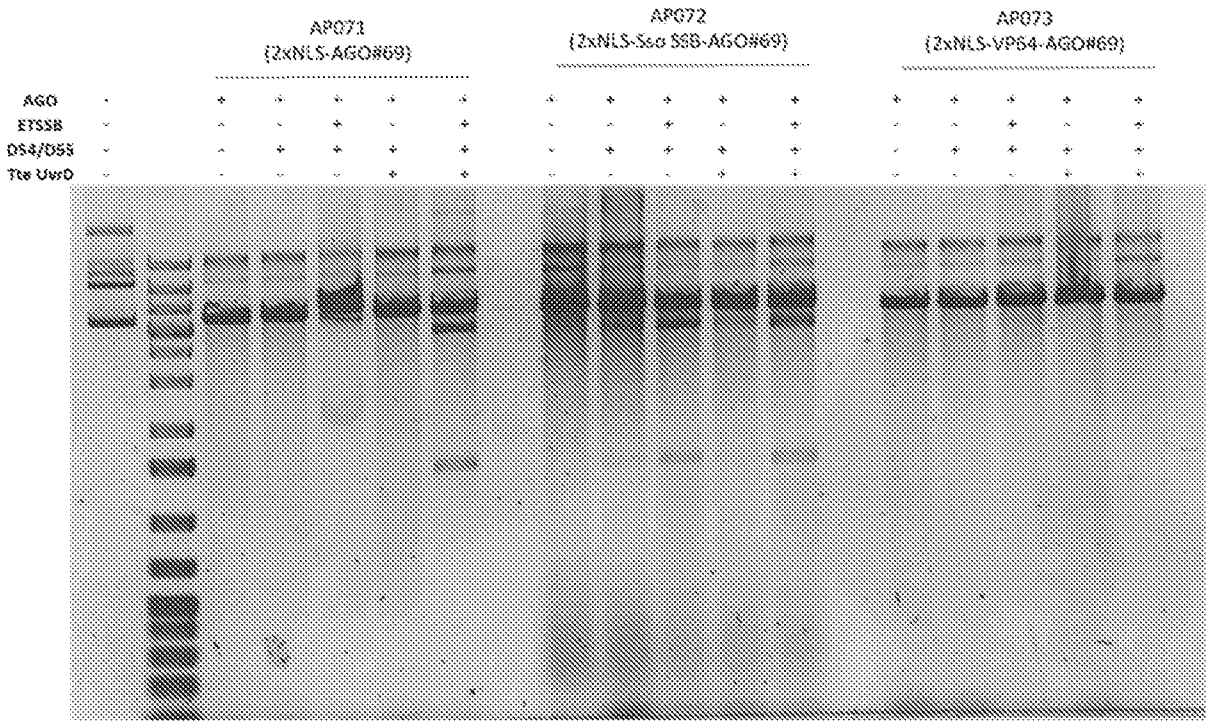


FIG. 69B



Expected Cleavage product:4604 + 1388 + 35 bp

FIG. 70



Expected Cleavage product:4604 + 1388 + 35 bp

FIG. 71

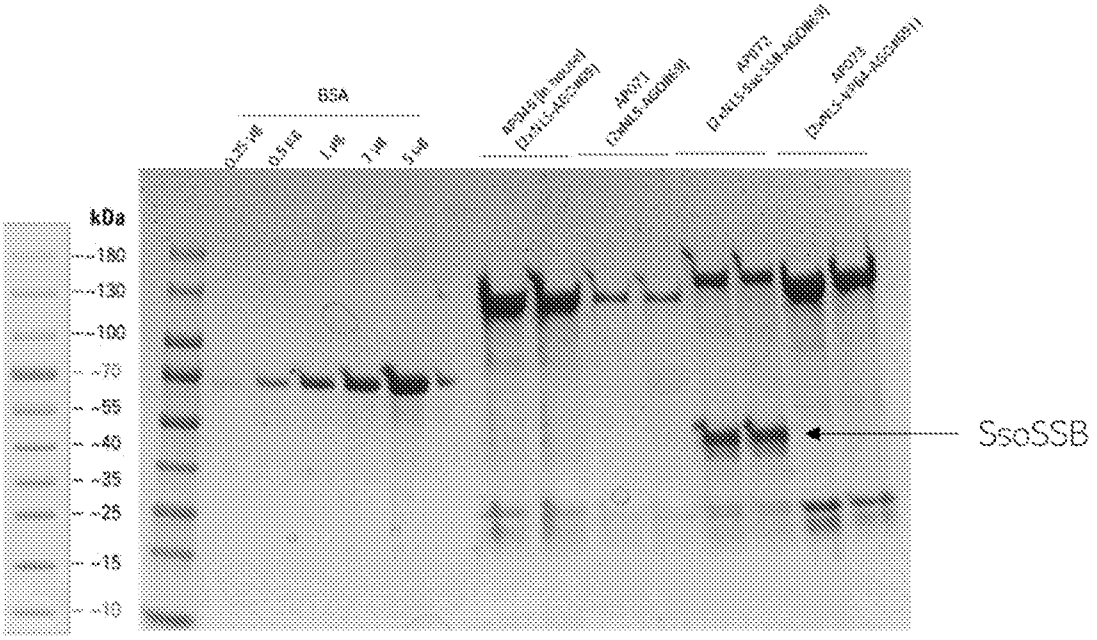


FIG. 72A

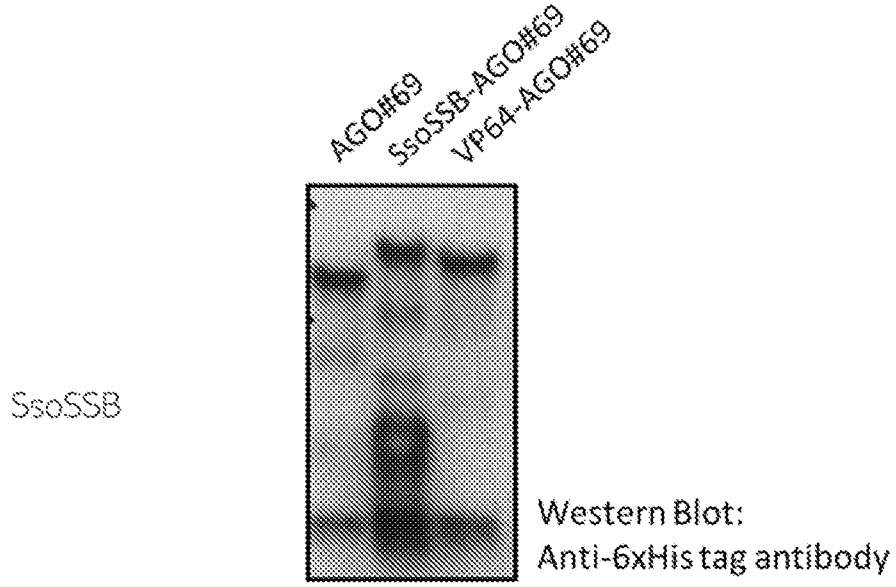
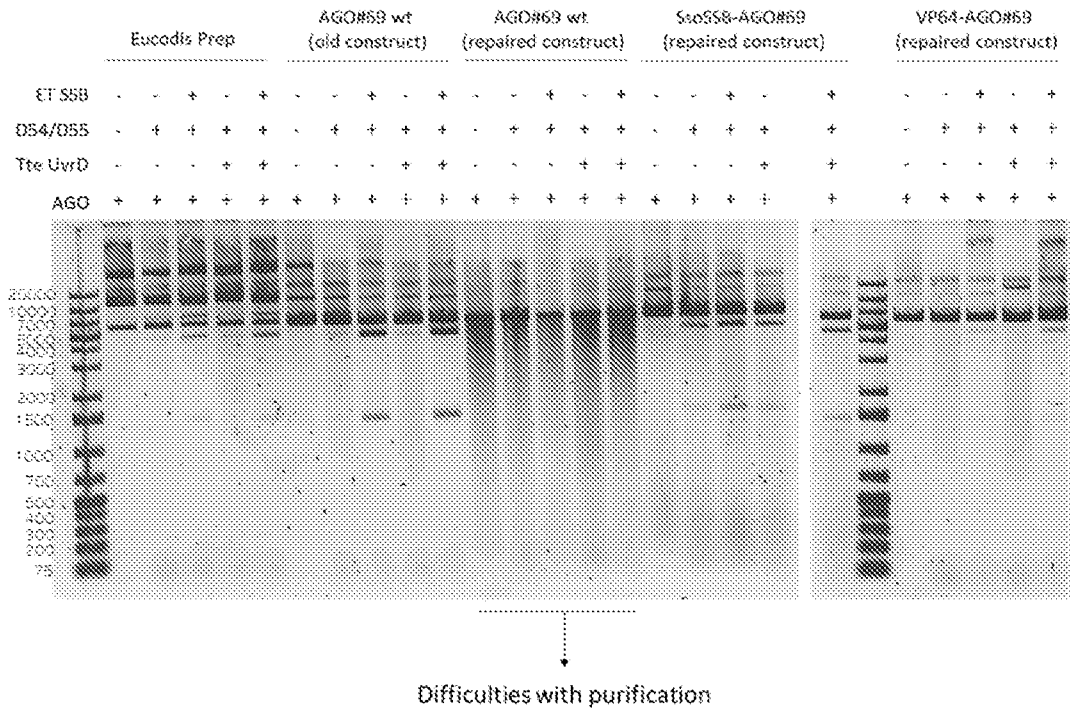
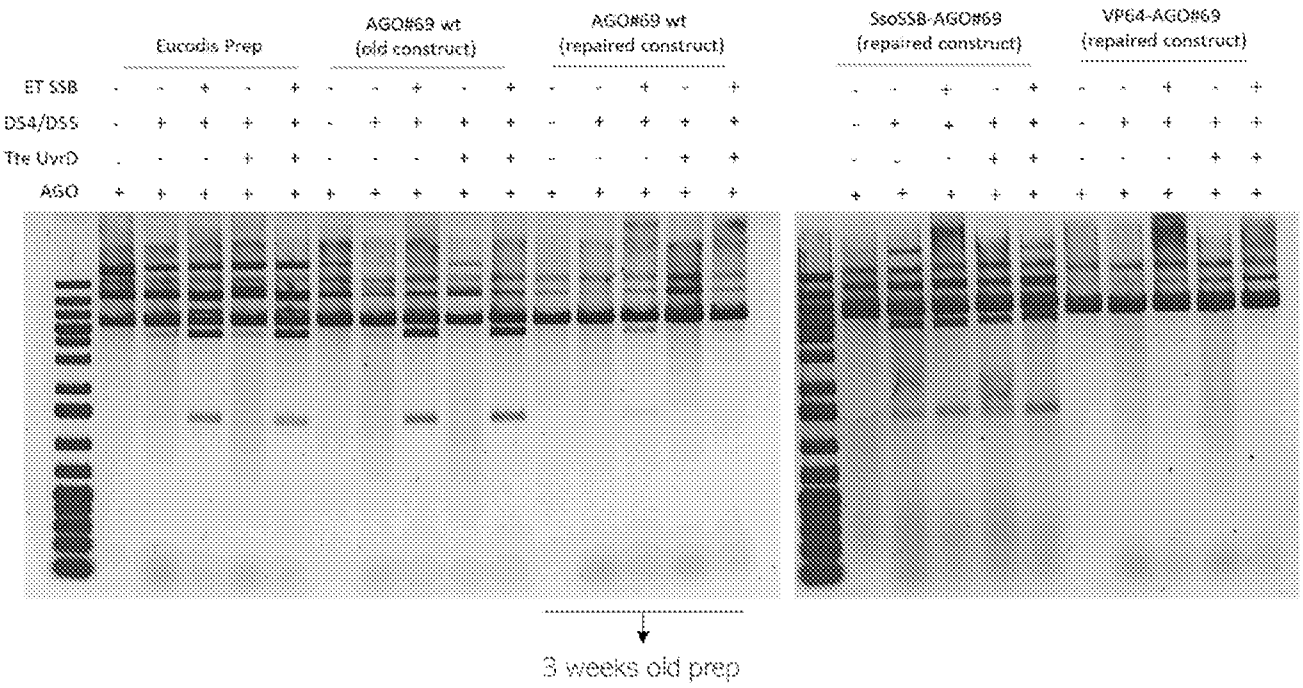


FIG. 72B



Expected Cleavage Product:4604 + 1388 + 35 bp

FIG. 73



Expected Cleavage Product:4604 + 1388 + 35 bp

FIG. 74

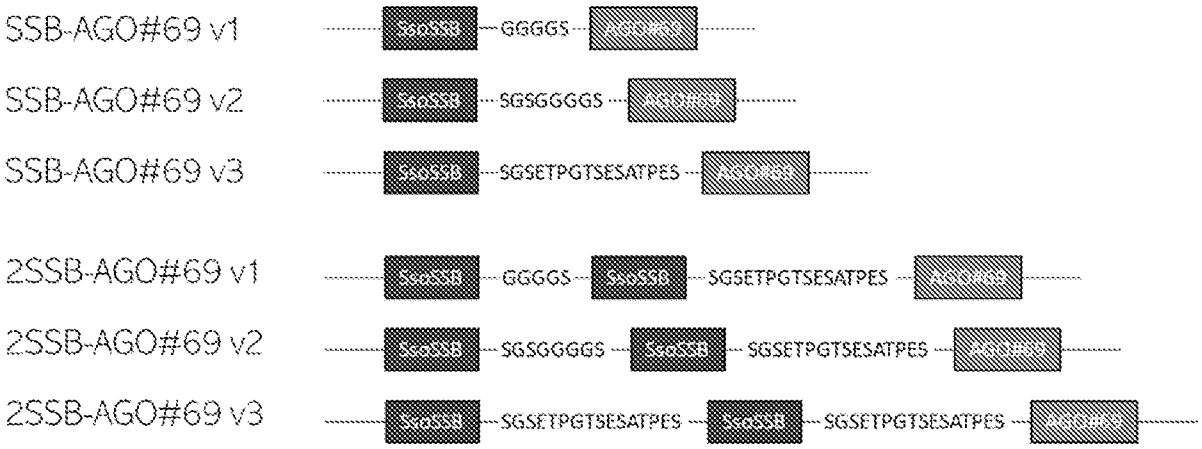

SSB-AGO#69 v1
SsoSSB
GGGGS
SSB-AGO#69 v2
SsoSSB
SGSGGGGS
SSB-AGO#69 v3
SsoSSB
SGSETPGTSESATPES
2SSB-AGO#69 v1
SsoSSB
GGGGS
SsoSSB
SGSETPGTSESATPES
2SSB-AGO#69 v2
SsoSSB
SGSGGGGS
SsoSSB
SGSETPGTSESATPES
2SSB-AGO#69 v3
SsoSSB
SGSETPGTSESATPES
SsoSSB
SGSETPGTSESATPES

FIG. 75

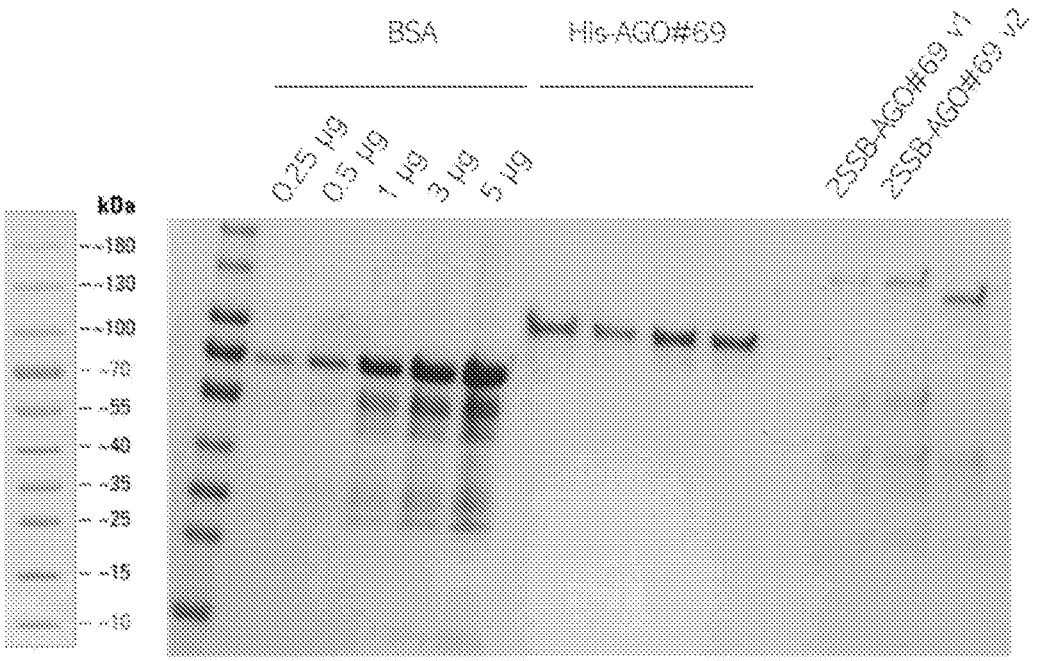


FIG. 76A

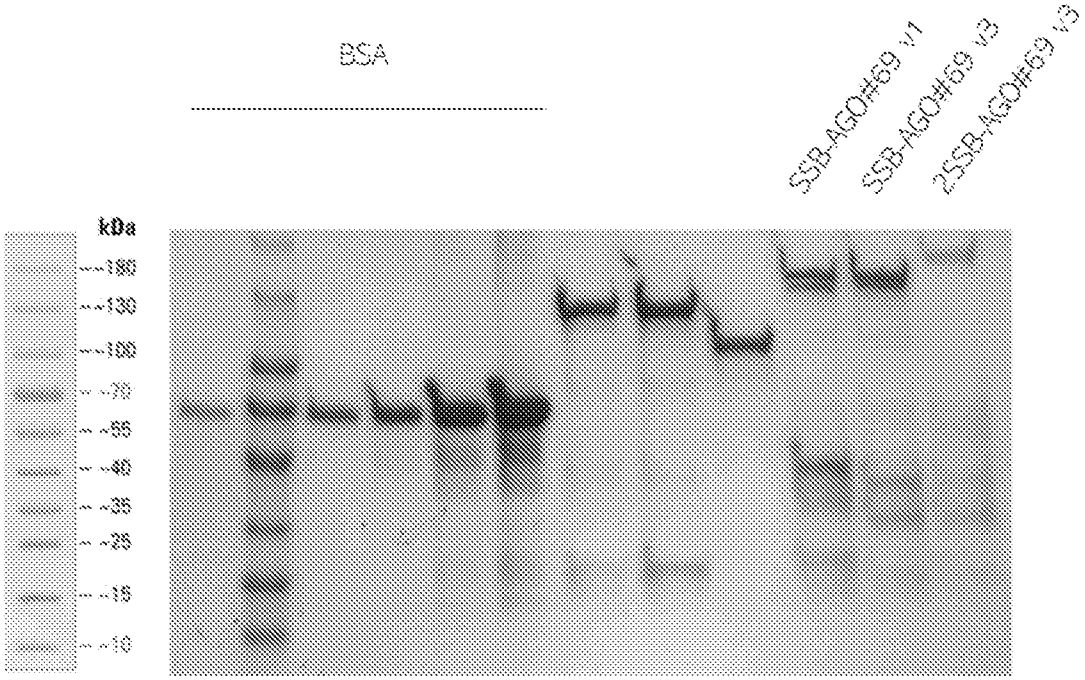


FIG. 76B

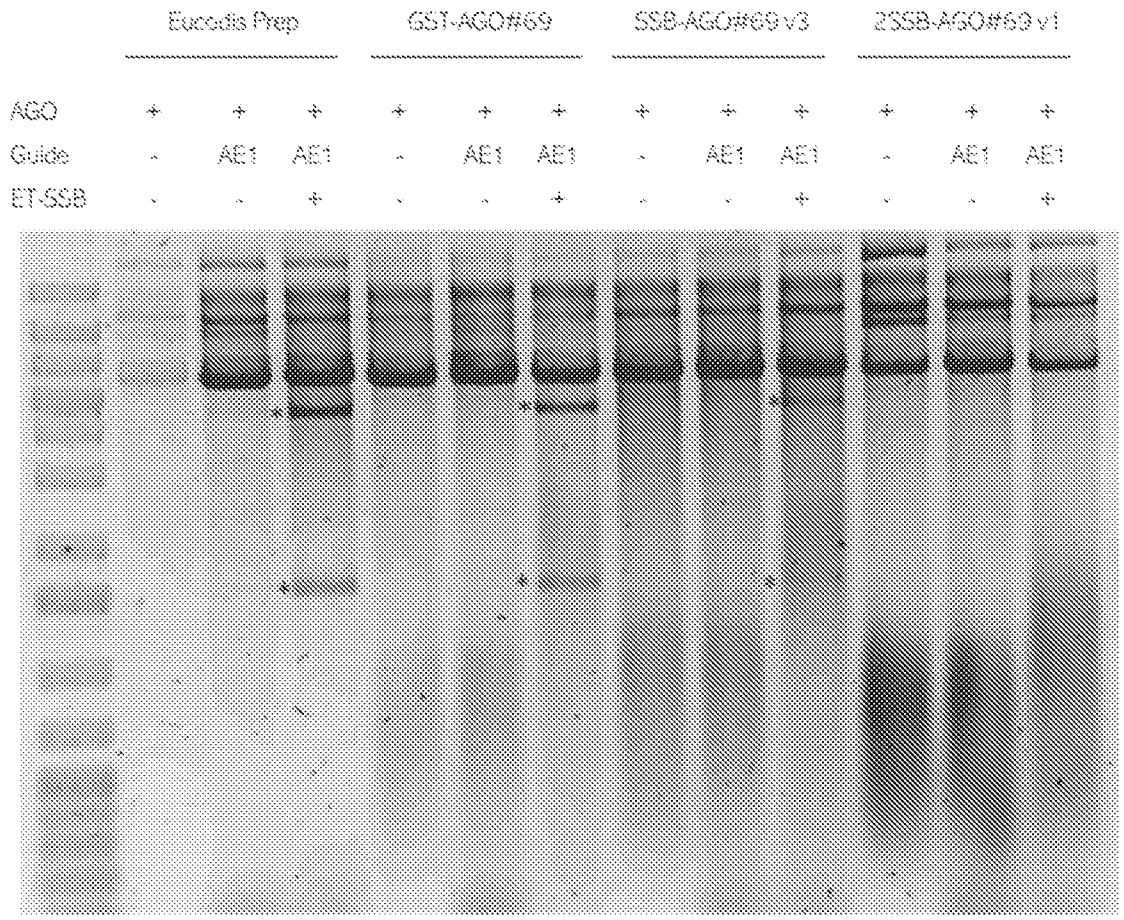

Eucodis Prep
GST-AGO#69
SSB-AGO#69 v3
2SSB-AGO#69 v1
AGO
Guide
ET-SSB
AE1
AE1 (D54/55)- AP056- RE KpnI-HF: 4723 + 1591

FIG. 77

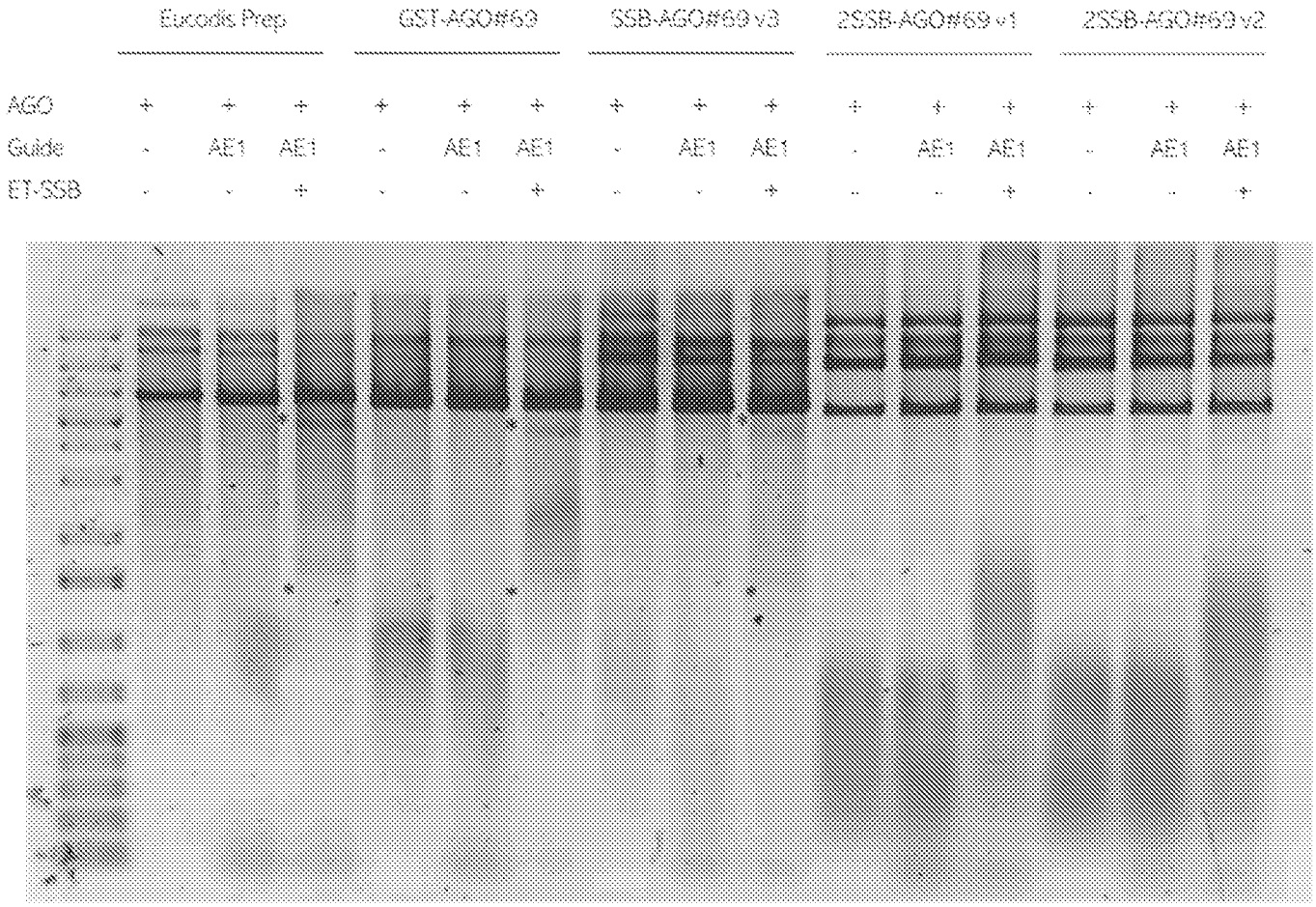
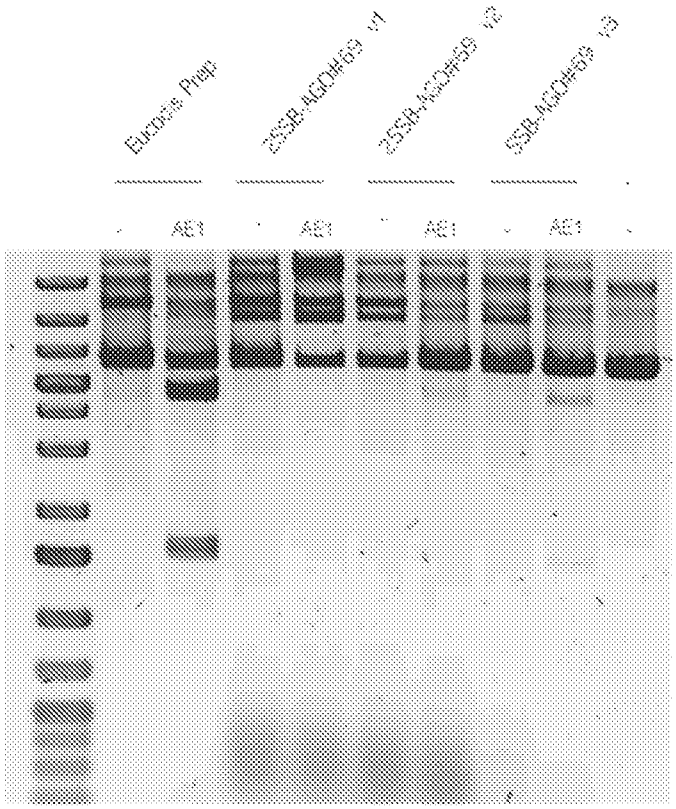

Eucodis Prep
GST-AGO#69
SSB-AGO#69 v3
2SSB-AGO#69 v1
2SSB-AGO#69 v2
AGO
Guide
ET-SSB
AE1
AE1 (D54/55)- AP056- RE KpnI-HF: 4723 + 1591

FIG. 78



AE1 (D54/55)- AP056- RE KpnI-HF: 4723 + 1591

FIG. 79

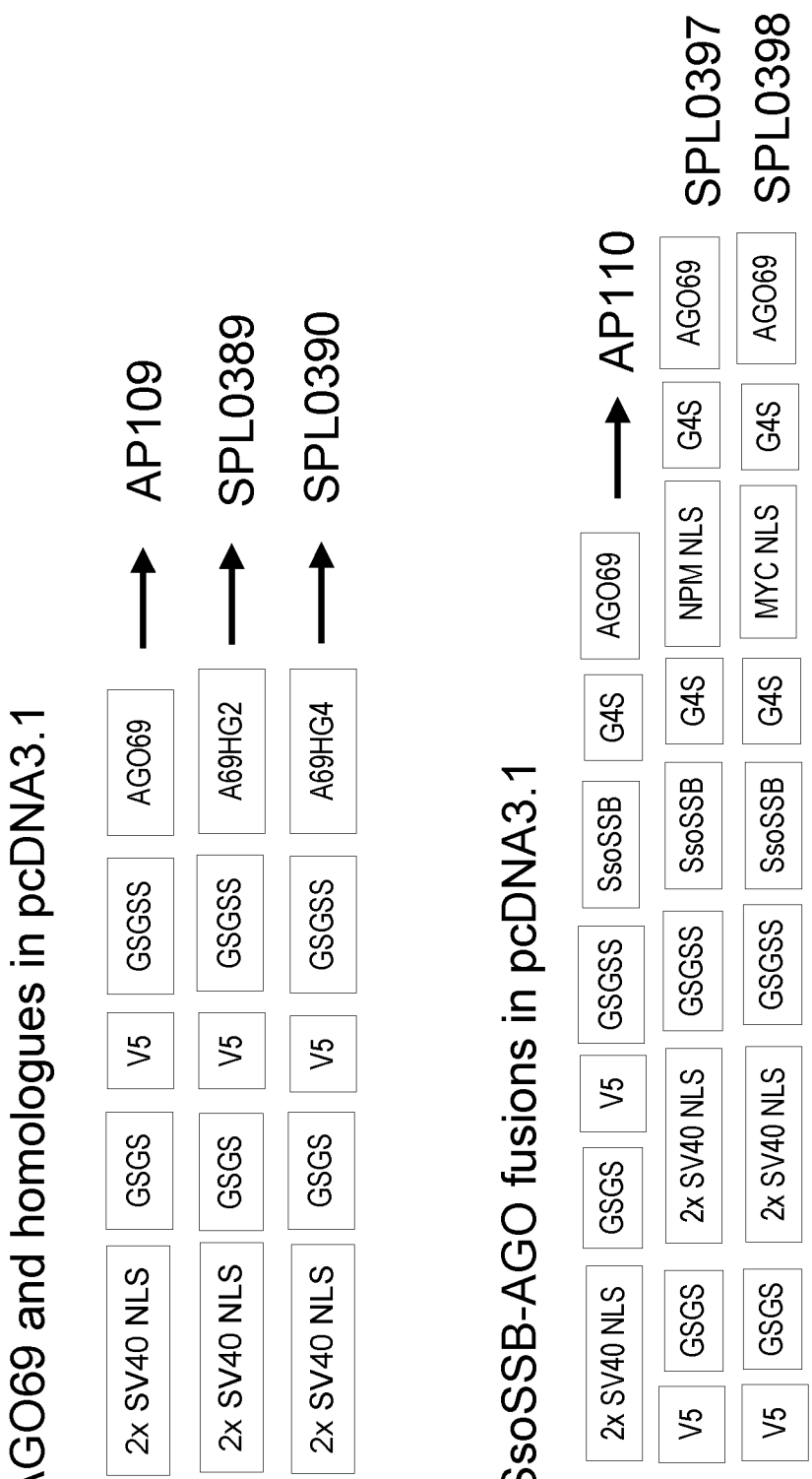

AGO69 and homologues in pcDNA3.1
2x SV40 NLS
GSGS
V5
GSGSS
AGO69
AP109
2x SV40 NLS
GSGS
V5
GSGSS
A69HG2
SPL0389
2x SV40 NLS
GSGS
V5
GSGSS
A69HG4
SPL0390
SsoSSB-AGO fusions in pcDNA3.1
2x SV40 NLS
GSGS
V5
GSGSS
SsoSSB
G4S
AGO69
AP110
V5
GSGS
2x SV40 NLS
GSGSS
SsoSSB
G4S
NPM NLS
G4S
AGO69
SPL0397
V5
GSGS
2x SV40 NLS
GSGSS
SsoSSB
G4S
MYC NLS
G4S
AGO69
SPL0398

FIG. 80

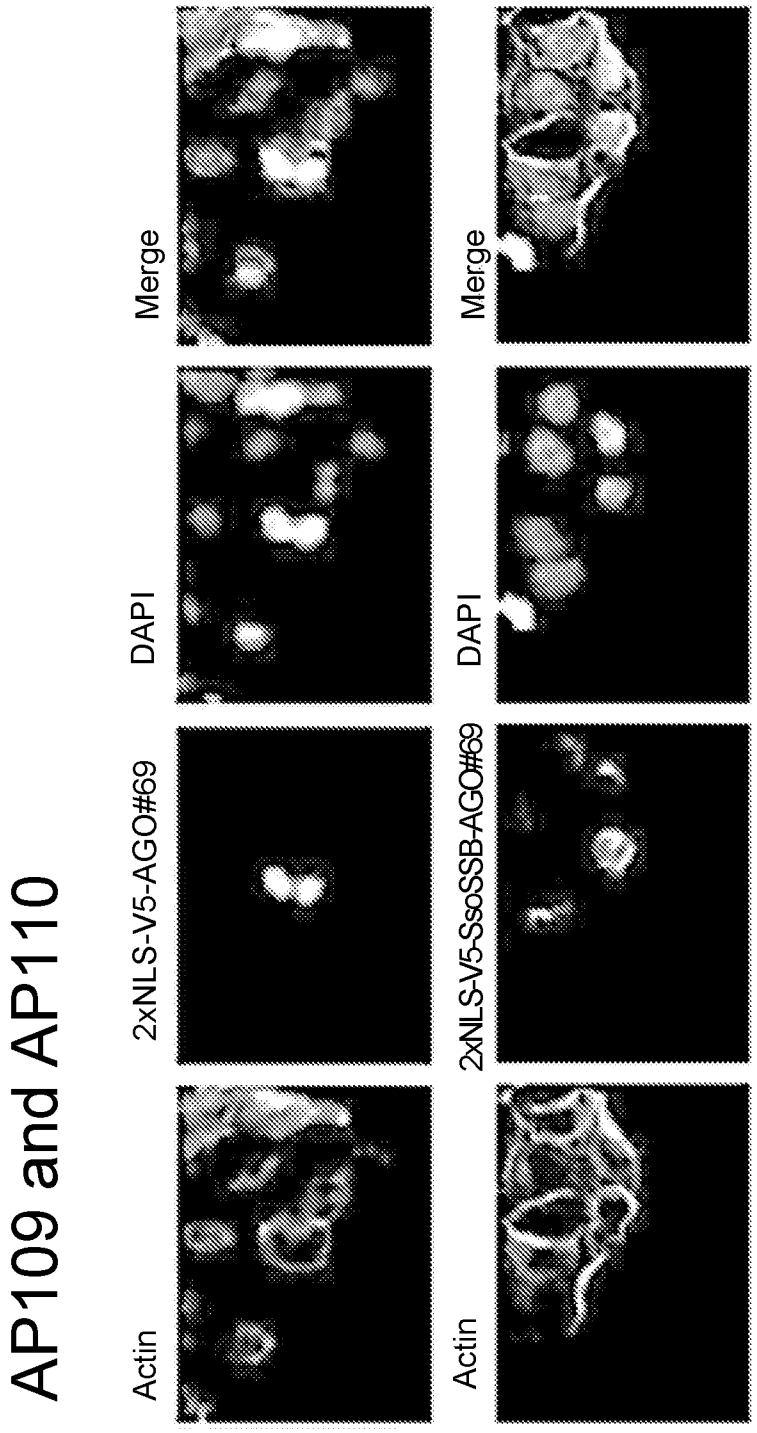

AP109 and AP110
Actin
2xNLS-V5-AGO#69
DAPI
Merge
Actin
2xNLS-V5-SsoSSB-AGO#69
DAPI
Merge

FIG. 81

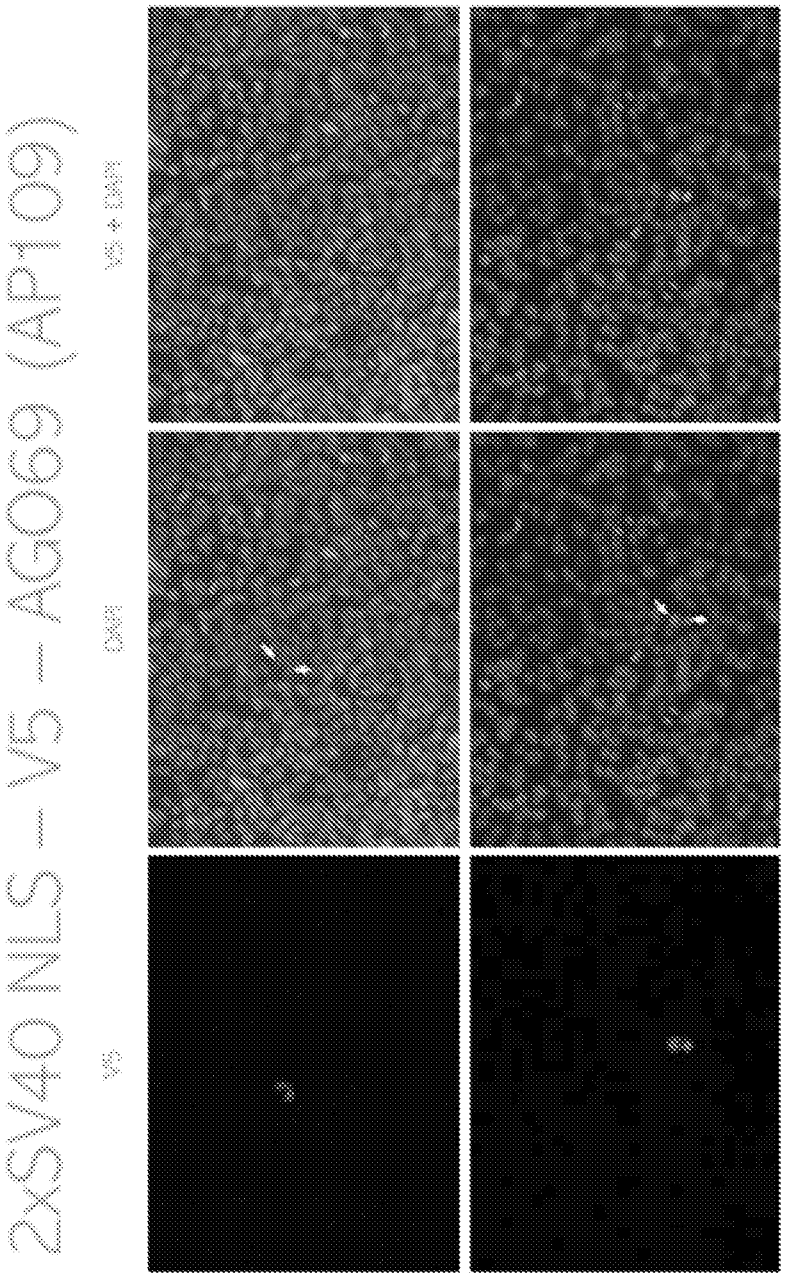


FIG. 82

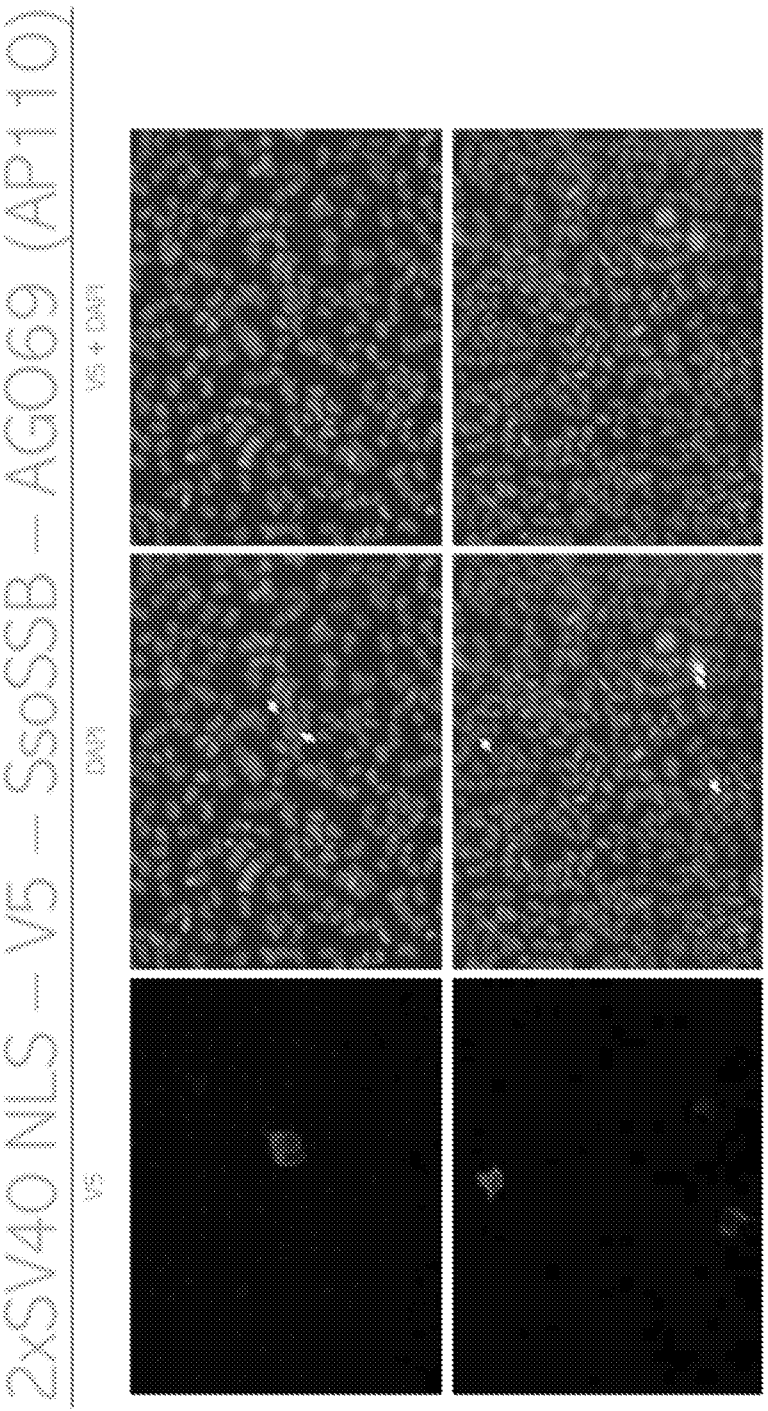


FIG. 83

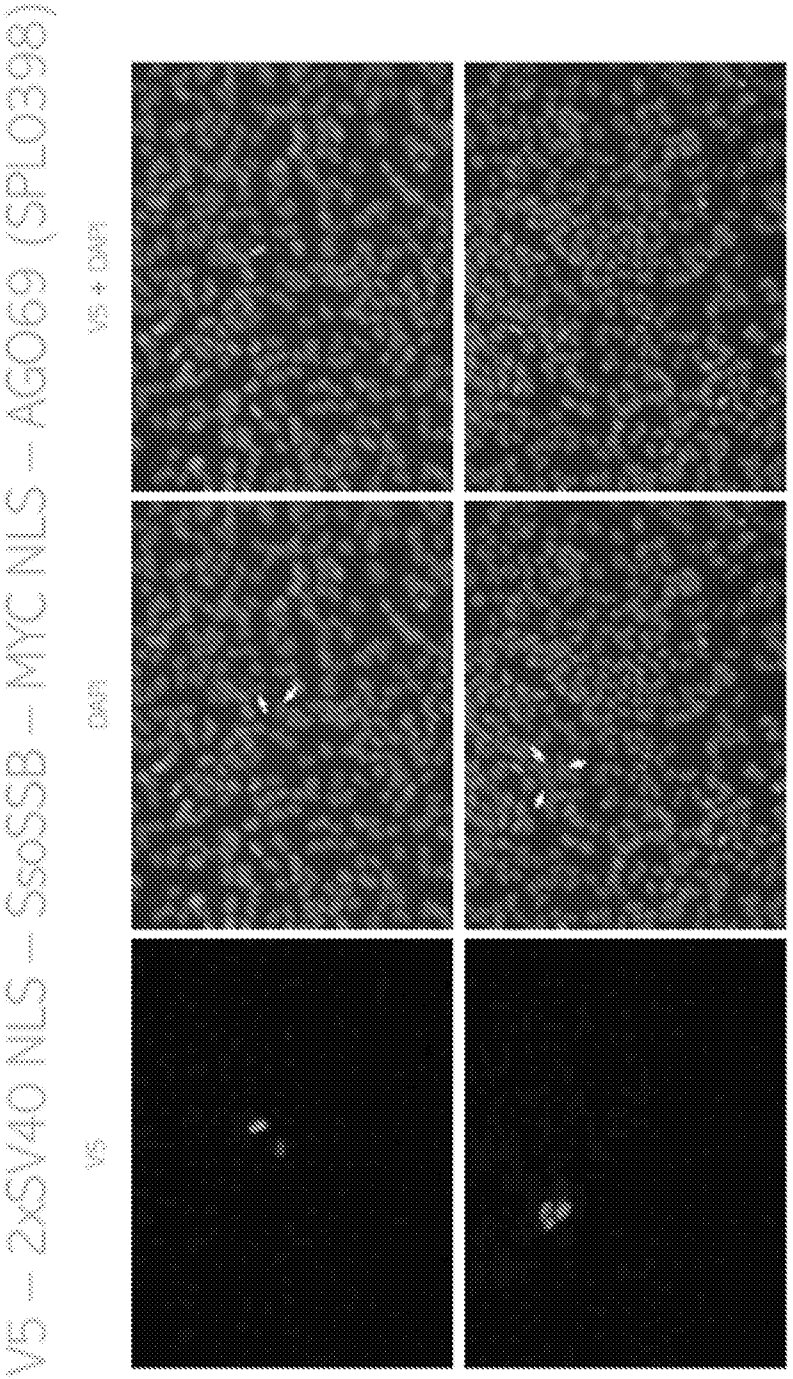


FIG. 84

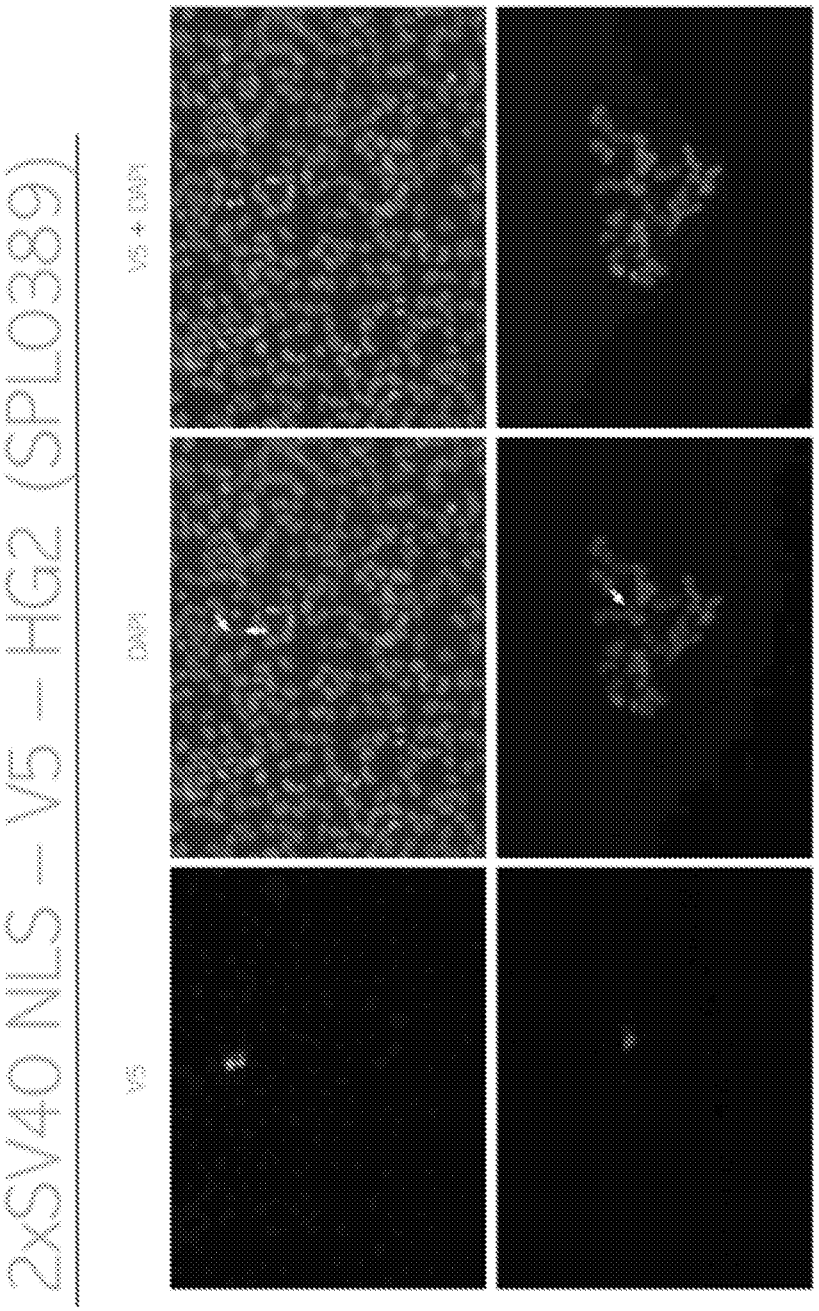


FIG. 85

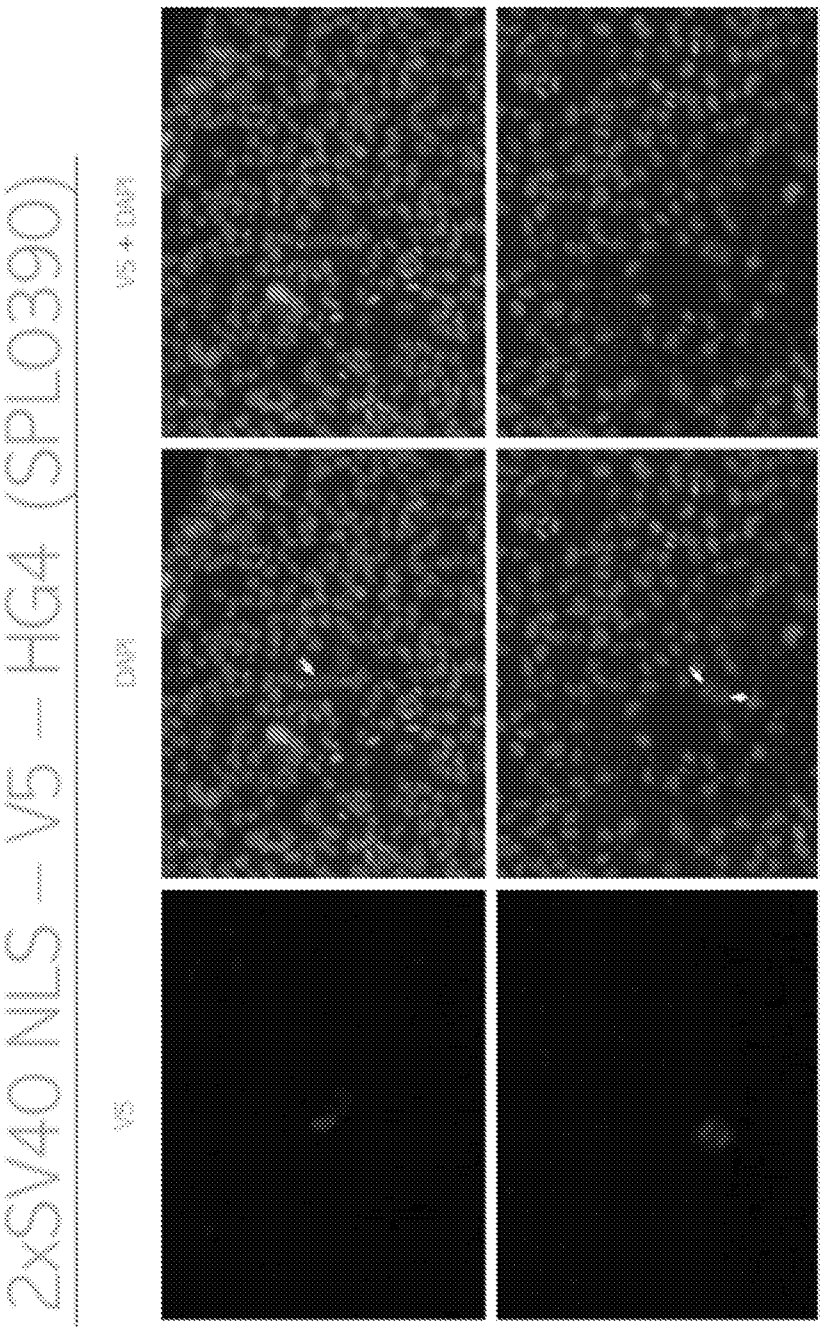
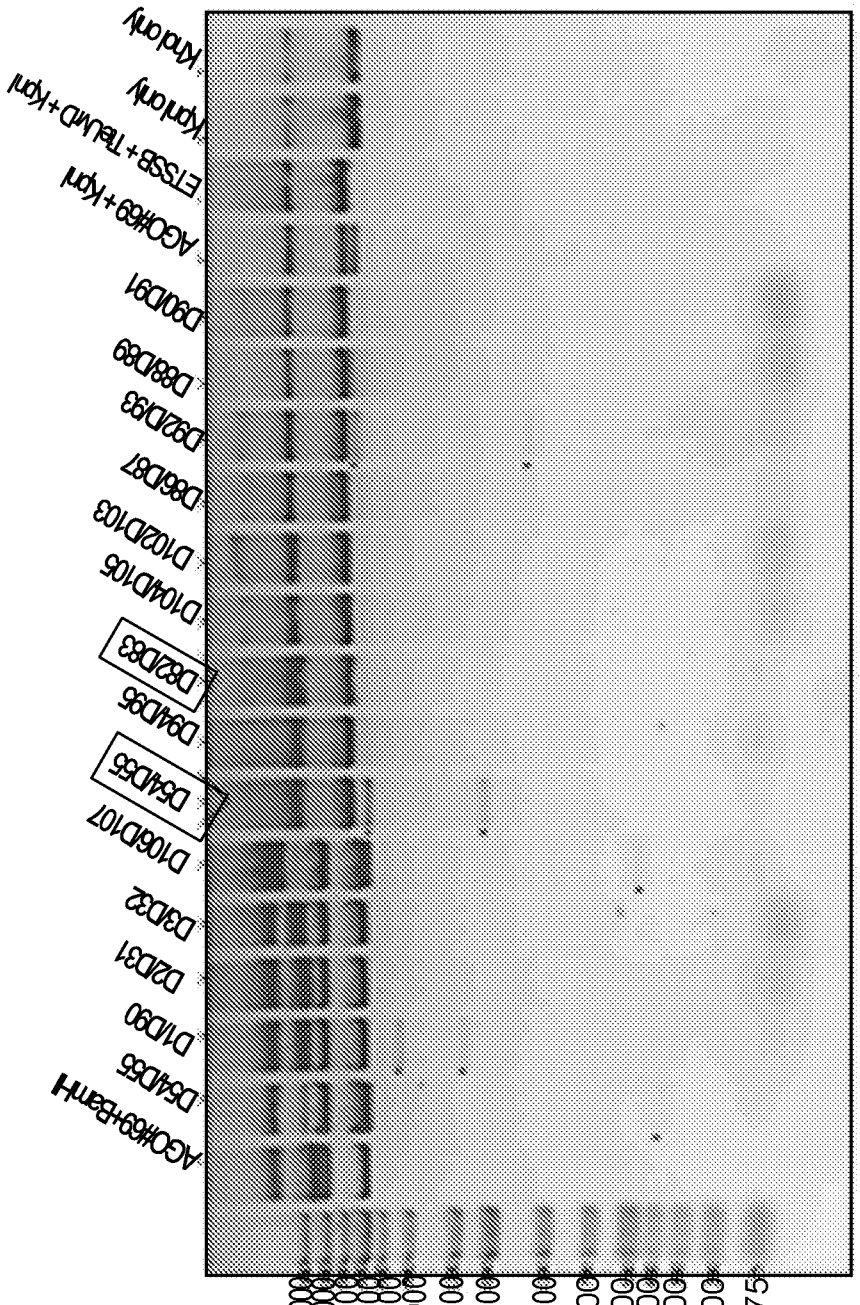


FIG. 86



Plasmid # 70

Guides	RE	% GC	Expected cleavage product
D54/D55	BamHI	23.8	375 + 4741
D1/D30	BamHI	52.38	1857 + 3259
D2/D31	BamHI	66.67	1902 + 3214
D3/D32	BamHI	76.19	1892 + 3224
D106/D107	BamHI	9.52	460 + 4656

Plasmid # 56

Guides	RE	% GC	Expected cleavage product
D54/D55	KpnI	23.8	1591 + 4723
D94/D95	KpnI	28.57	1040 + 5274
D82/D83	KpnI	38.10	1022 + 5292
D104/D105	KpnI	47.62	1553 + 4761
D102/D103	KpnI	52.38	1452 + 4862
D86/D87	KpnI	57.14	1320 + 4994
D92/D/93	KpnI	61.90	1098 + 5216
D88/D89	KpnI	66.67	1200 + 5114
D90/D91	KpnI	71.43	1299 + 5015

FIG. 87

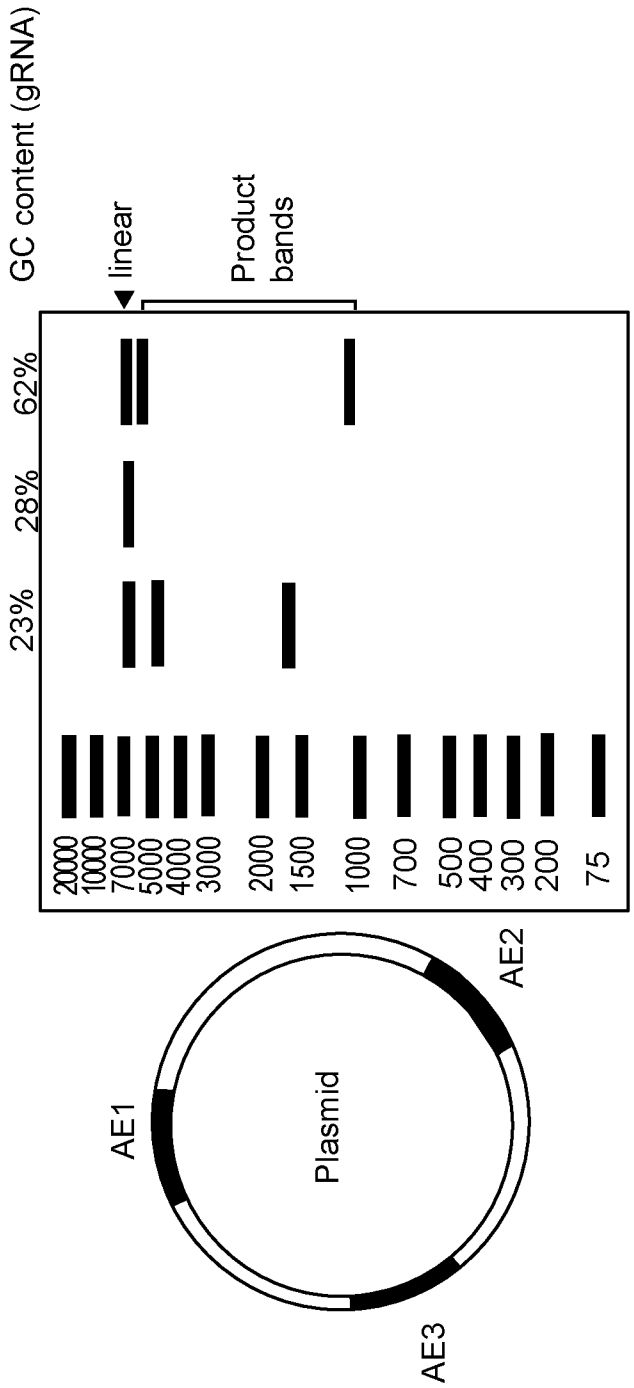

GC content (gRNA)
linear
Product bands
62%
28%
23%
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
400
300
200
75
AE1
AE2
AE3
Plasmid

FIG. 88A

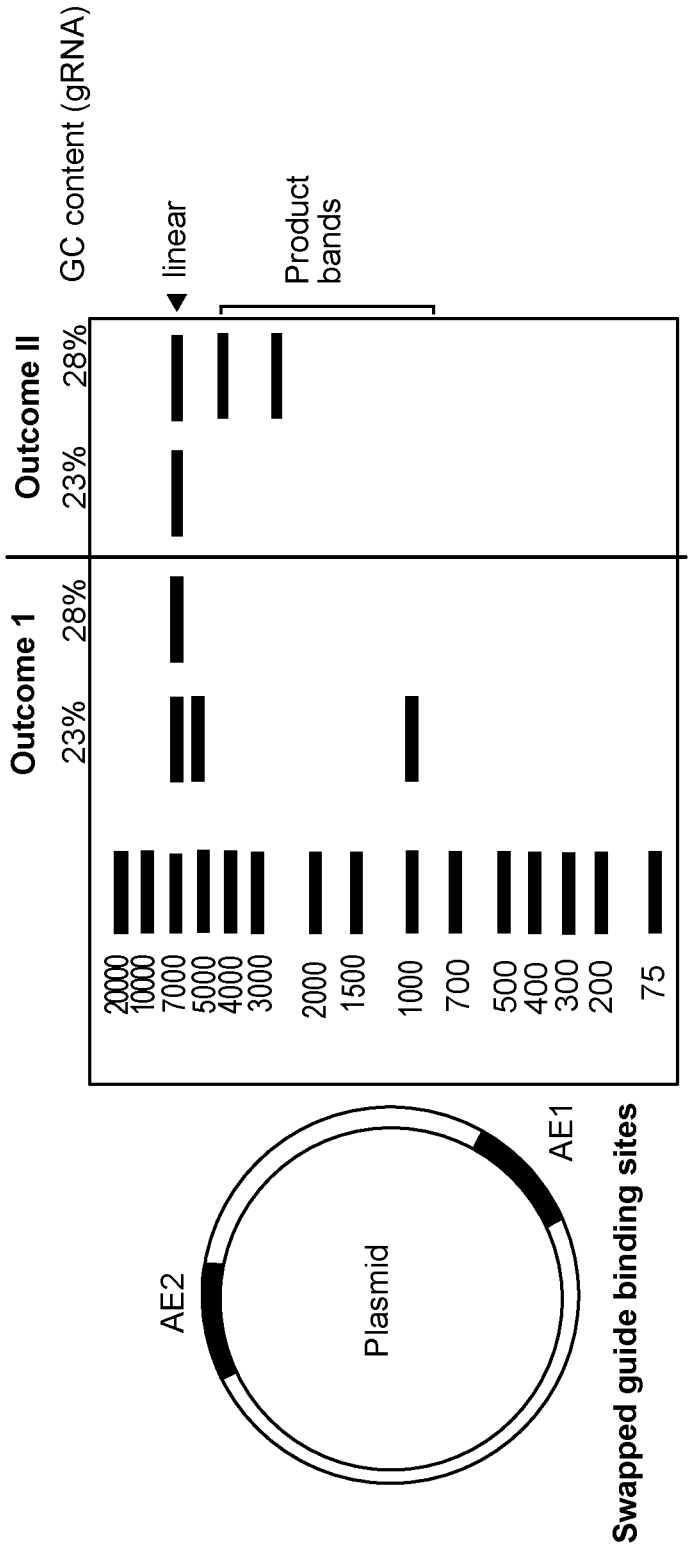
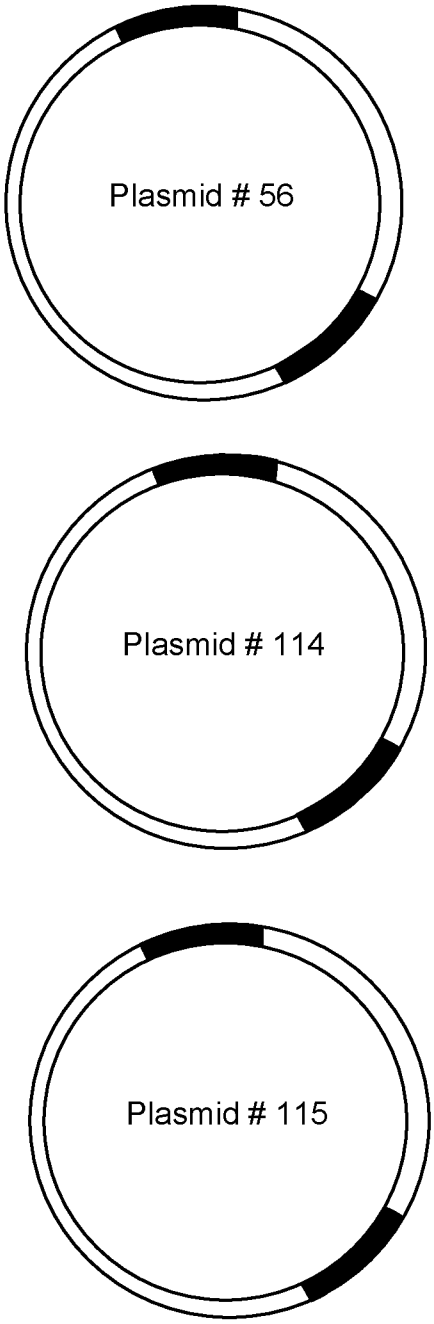

GC content (gRNA)
linear
Product bands
Outcome II
28%
23%
Outcome 1
28%
23%
20000
10000
7000
5000
4000
3000
2000
1500
1000
700
500
400
300
200
75
AE2
Plasmid
AE1
Swapped guide binding sites

FIG. 88B



Swapped guide binding sites

FIG. 89

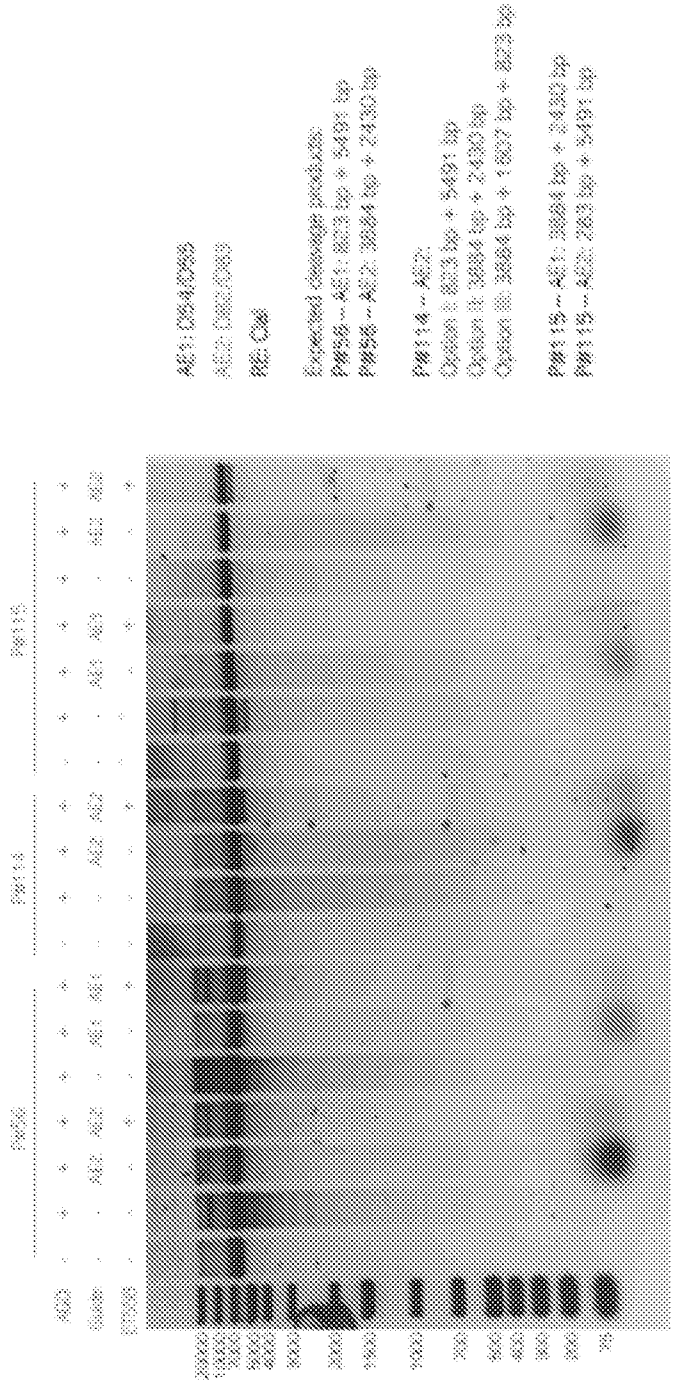


FIG. 90

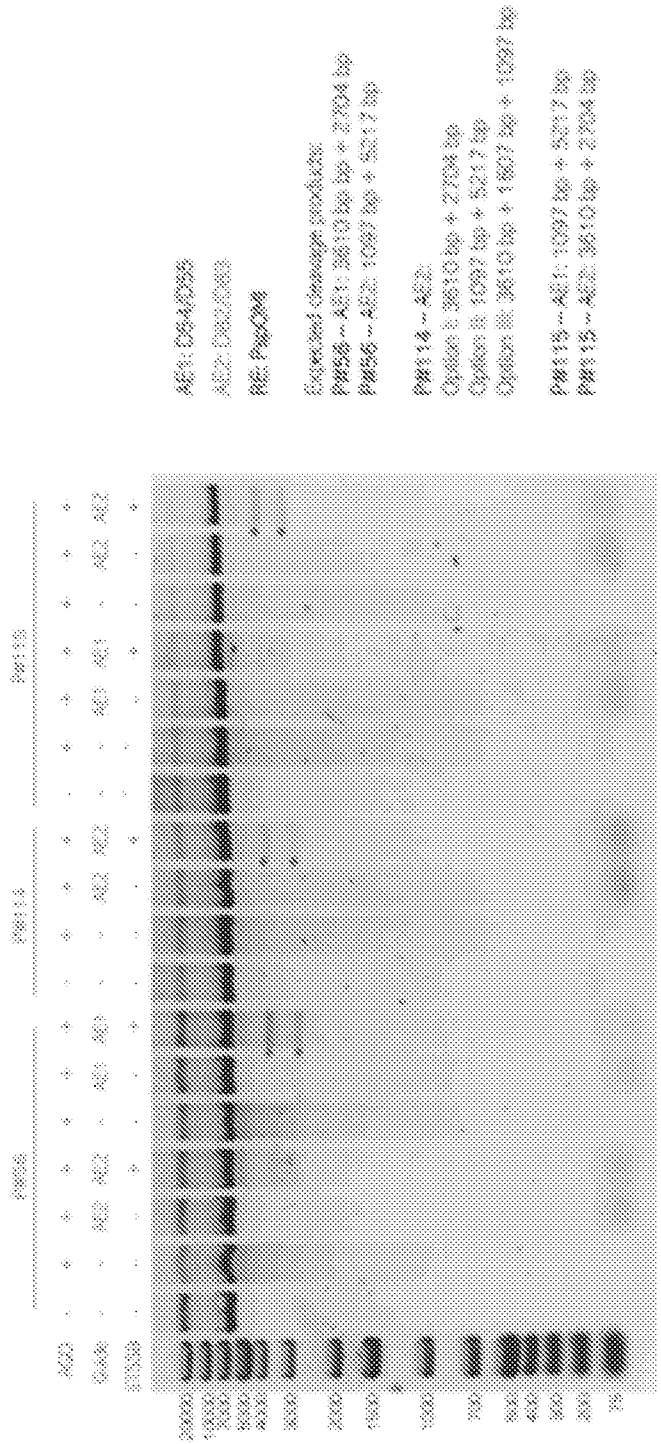

AE1: DE4/DE5
AE2: DE2/DE3
RE: PspOMI
Expected cleavage products:
PM56 -- AE1: 3610 bp + 2704 bp
PM56 -- AE2: 1097 bp + 5217 bp
PM116 -- AE2:
Option I: 3610 bp + 2704 bp
Option II: 1097 bp + 5217 bp
Option III: 3610 bp + 1607 bp + 1097 bp
PM115 -- AE1: 1097 bp + 5217 bp
PM115 -- AE2: 3610 bp + 2704 bp
PM56
PM116
PM115
AE2
DE4/DE5
RE
2000
1800
1600
1400
1200
1000
800
600
400
200
100
50
25

FIG. 91

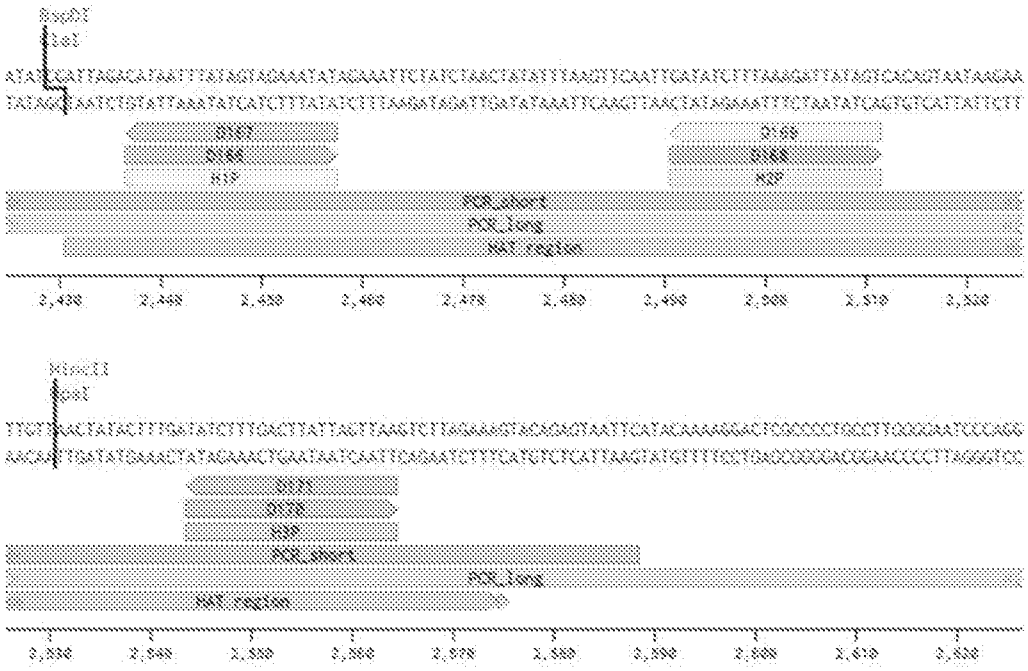
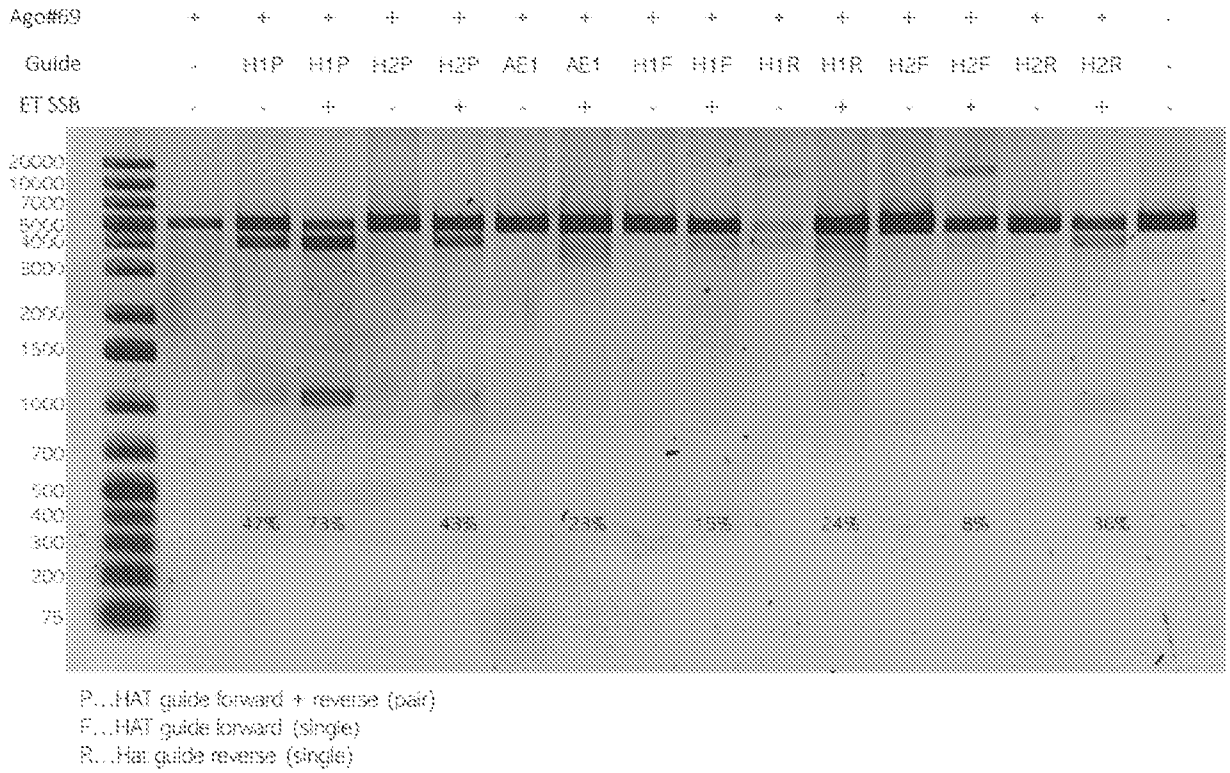

PCR_short
PCR_long
MAT region
2,430
2,440
2,450
2,460
2,470
2,480
2,490
2,500
2,510
2,520
PCR_short
PCR_long
MAT region
2,530
2,540
2,550
2,560
2,570
2,580
2,590
2,600
2,610
2,620

FIG. 92



P...HAT guide forward + reverse (pair)
F...HAT guide forward (single)
R...Hat guide reverse (single)

FIG. 93

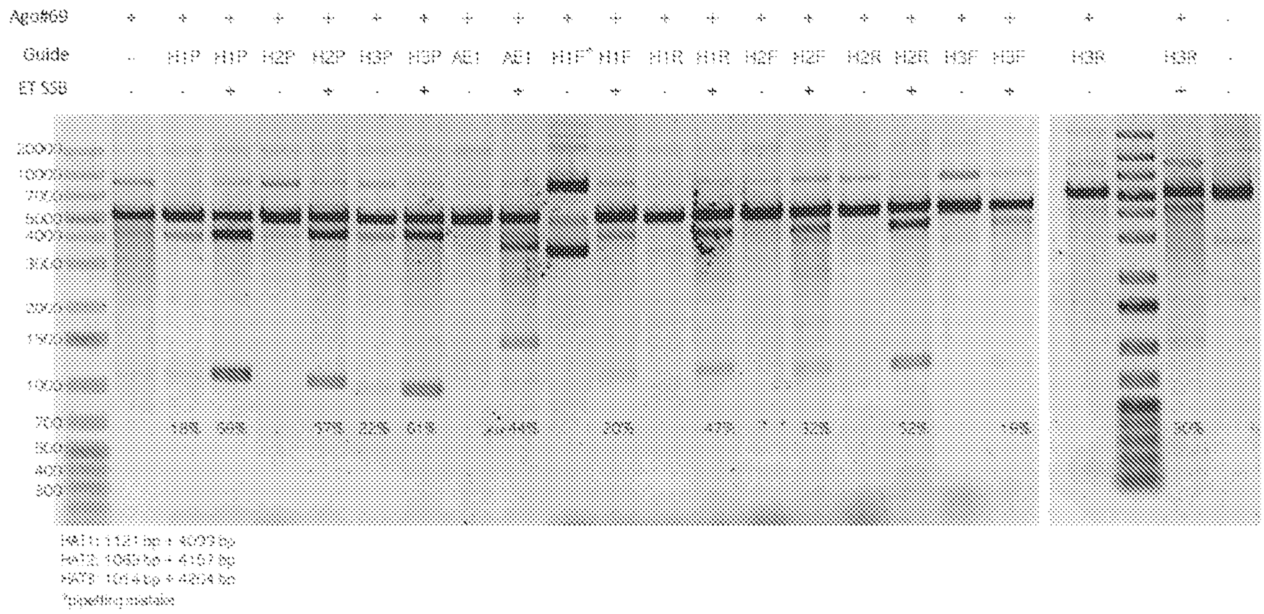
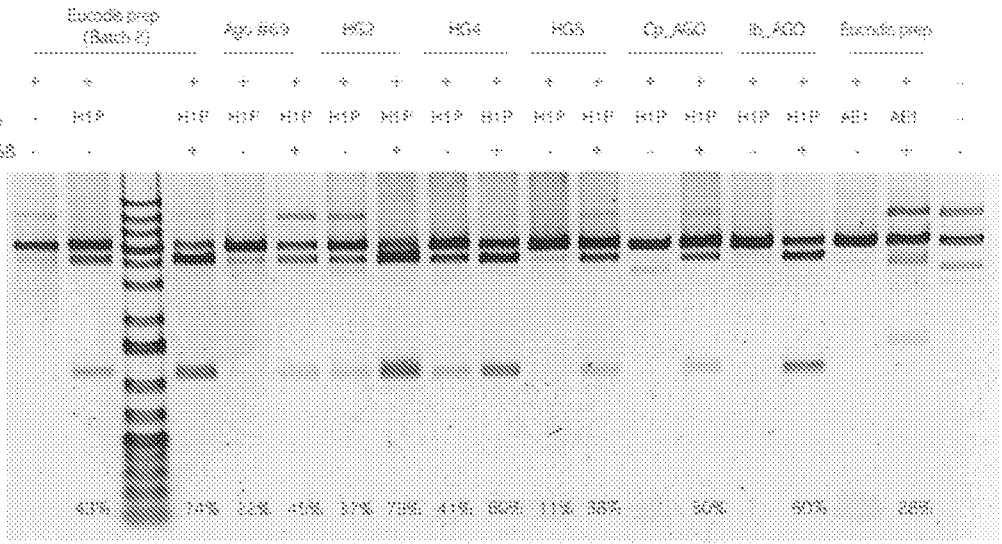

Ago#69
Guide
ET SSB
H1P
H2P
H3P
AE1
H1F
H1R
H2F
H2R
H3F
H3R
*pipetting mistake

FIG. 94

Plasmid#70-HAT



H1P...HAT1 guide (fwd + rev)

FIG. 95

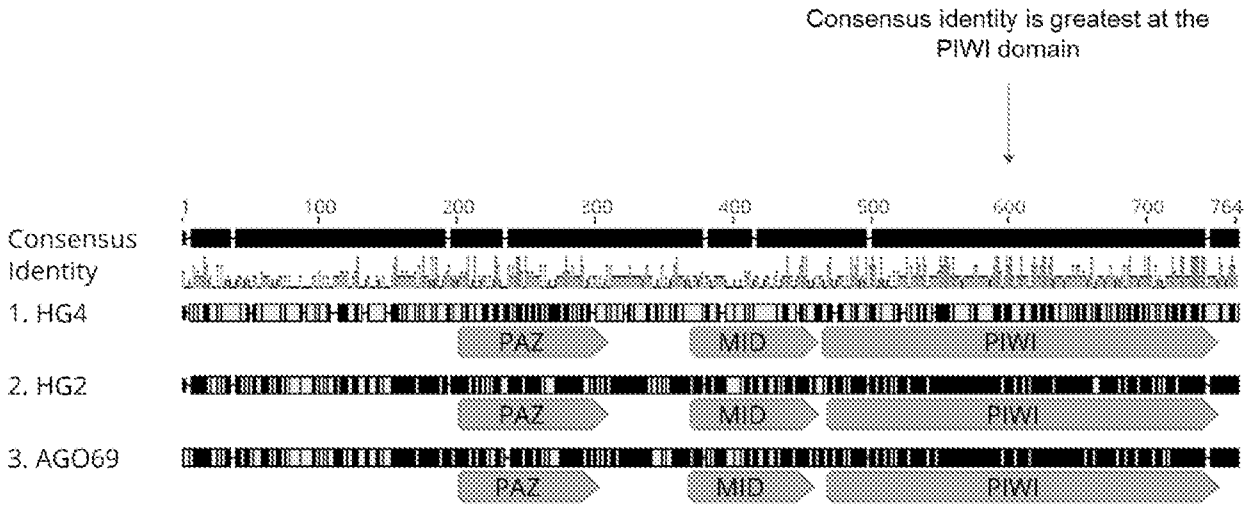


FIG. 96A

% Amino Acid Identity Between:	Ago69, HG2, HG4	Ago69, HG2	Ago69, HG4
PIWI Domains	27.7%	72.8%	34.1%
Whole Protein Sequence	19.1%	61.3%	25.2%

FIG. 96B

Accession #	Internal ID	Species	MW with Tags	
WP_002566035.1	HG1	Clostridium bolteae	152.2 kDa	
WP_045143632.1	HG2	Clostridium butyricum	117.3 kDa	
WP_061307001.1	HG3	Clostridium botulinum	85.2 kDa	
WP_016205751.1	HG4	Clostridium sartagoforme	113.98 kDa	
WP_048925809.1	HG5	Clostridium sp. 1_1_41A1FAA	117.5 kDa	
WP_119866036.1	HG6		117.8 kDa	
WP_118721317.1	HG7		120.3 kDa	
CpAgo	HG8	Clostridium perfringens	117.0 kDa	
IbAgo	HG9	Intestinibacter bartlettii	117.4 kDa	

FIG. 97

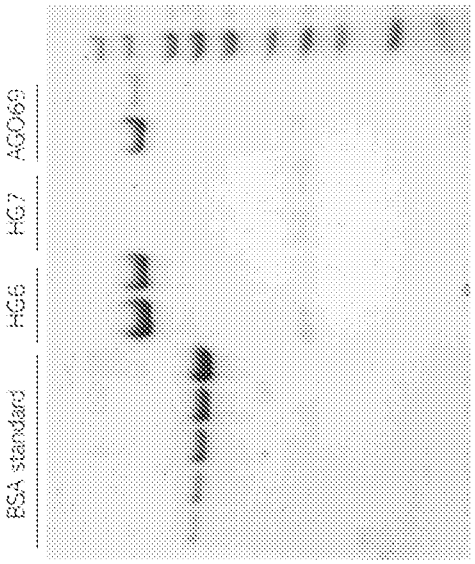


FIG. 98B

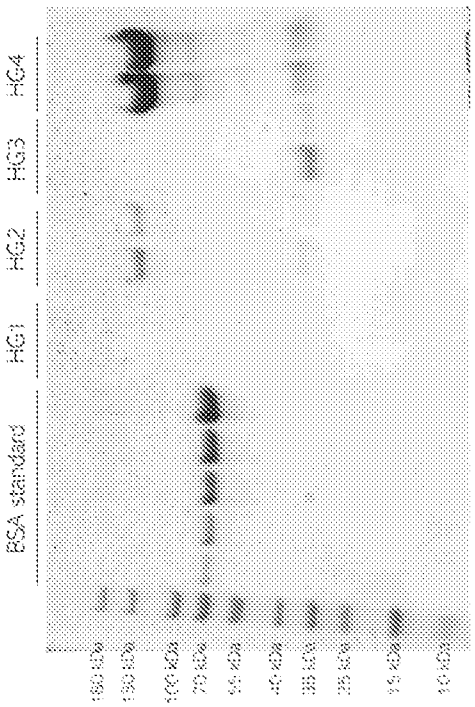


FIG. 98A

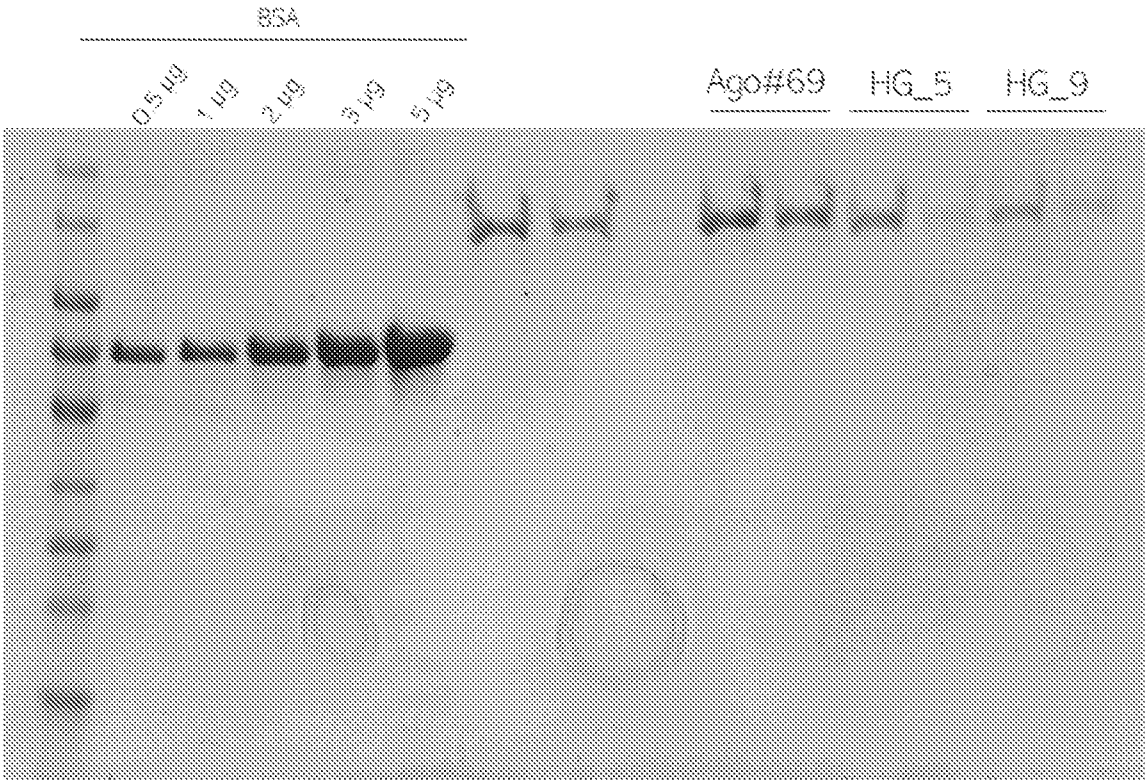


FIG. 99

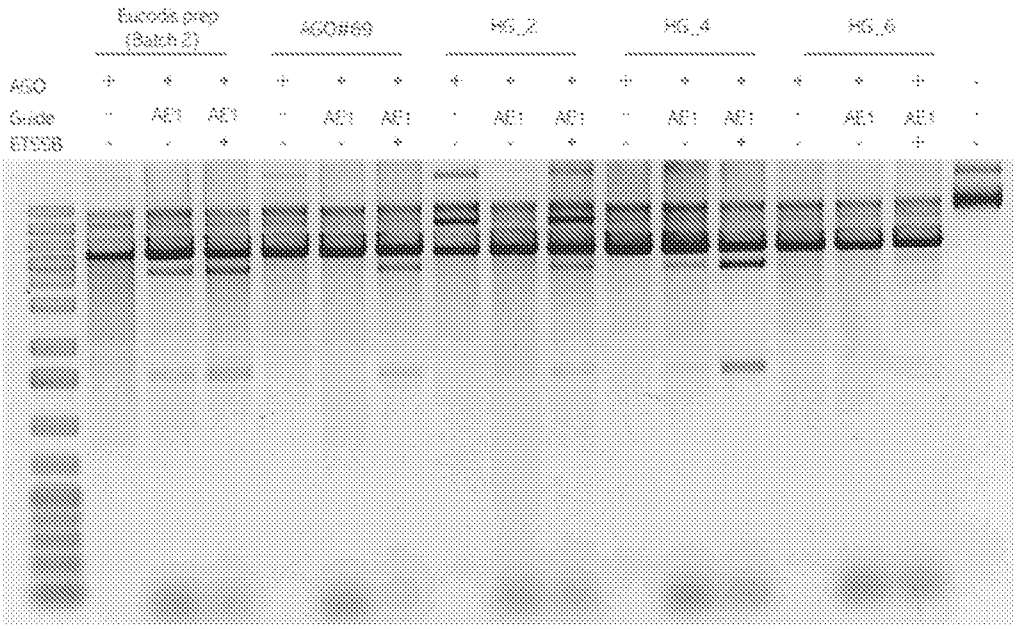

Exicode prep (Batch 2)
ASO#69
HG_2
HG_4
HG_6
ASO
Guide
ETS2B
AE1
AE1
AE1
AE1
AE1
AE1

FIG. 100

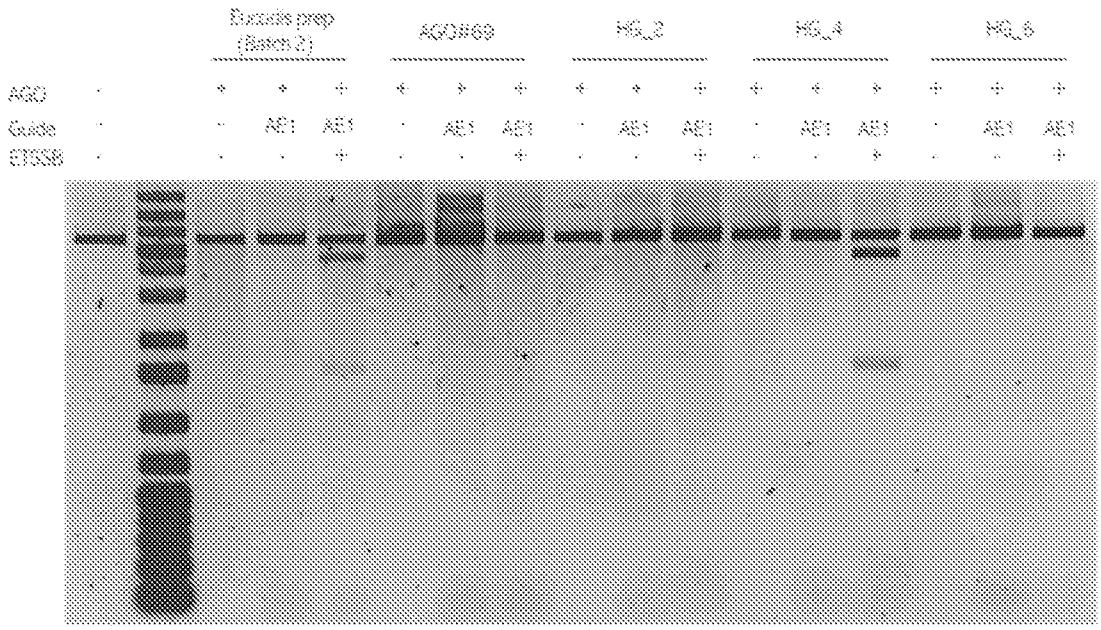

Exzact prep (Exon 2)
AGO#69
MG_2
MG_4
MG_6
AGO
Guide
ETSSB
AE1
AE1
AE1
AE1
AE1
AE1
AE1
AE1
AE1
AE1
AE1
AE1

FIG. 101

FIG. 102

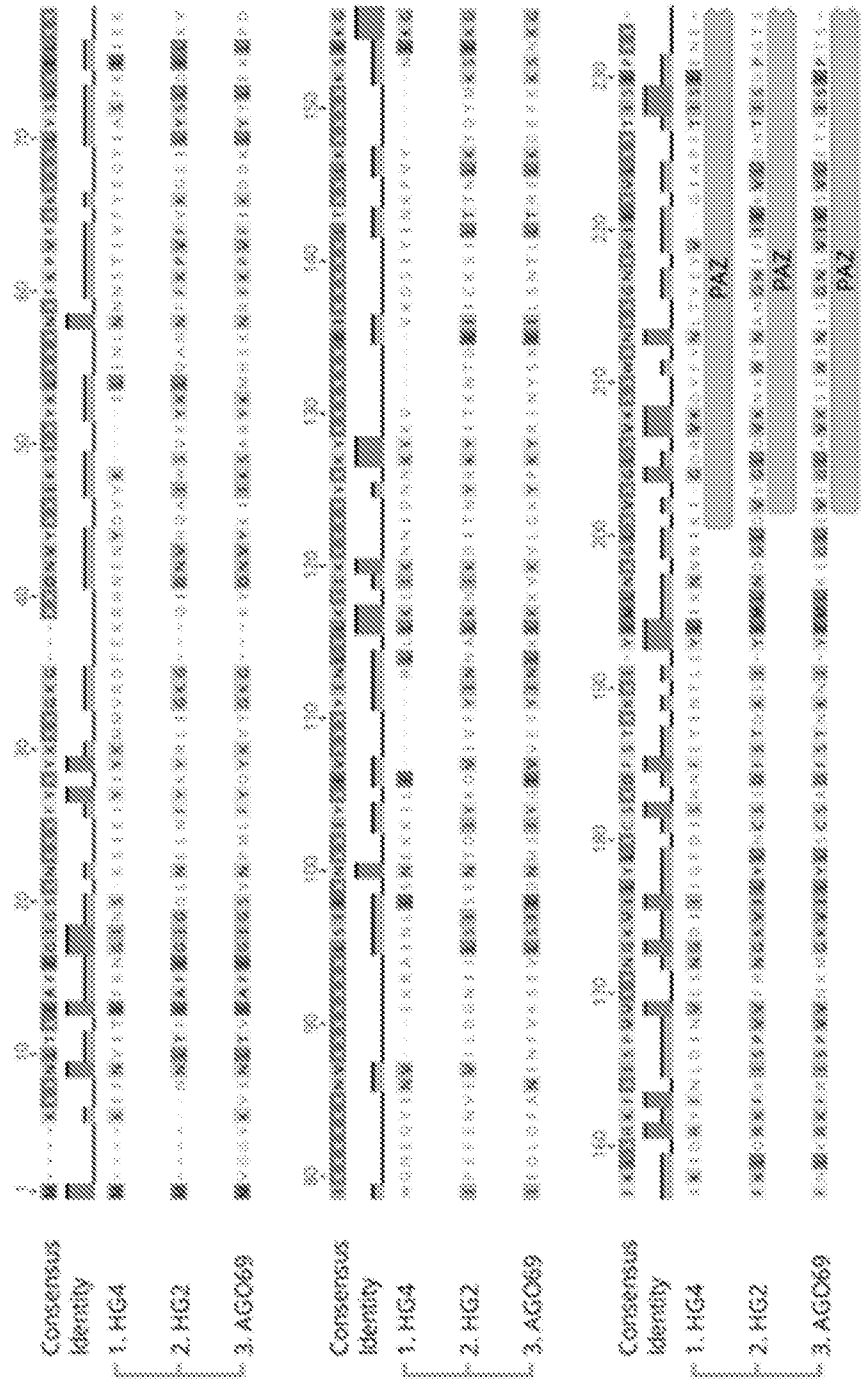


FIG. 103A

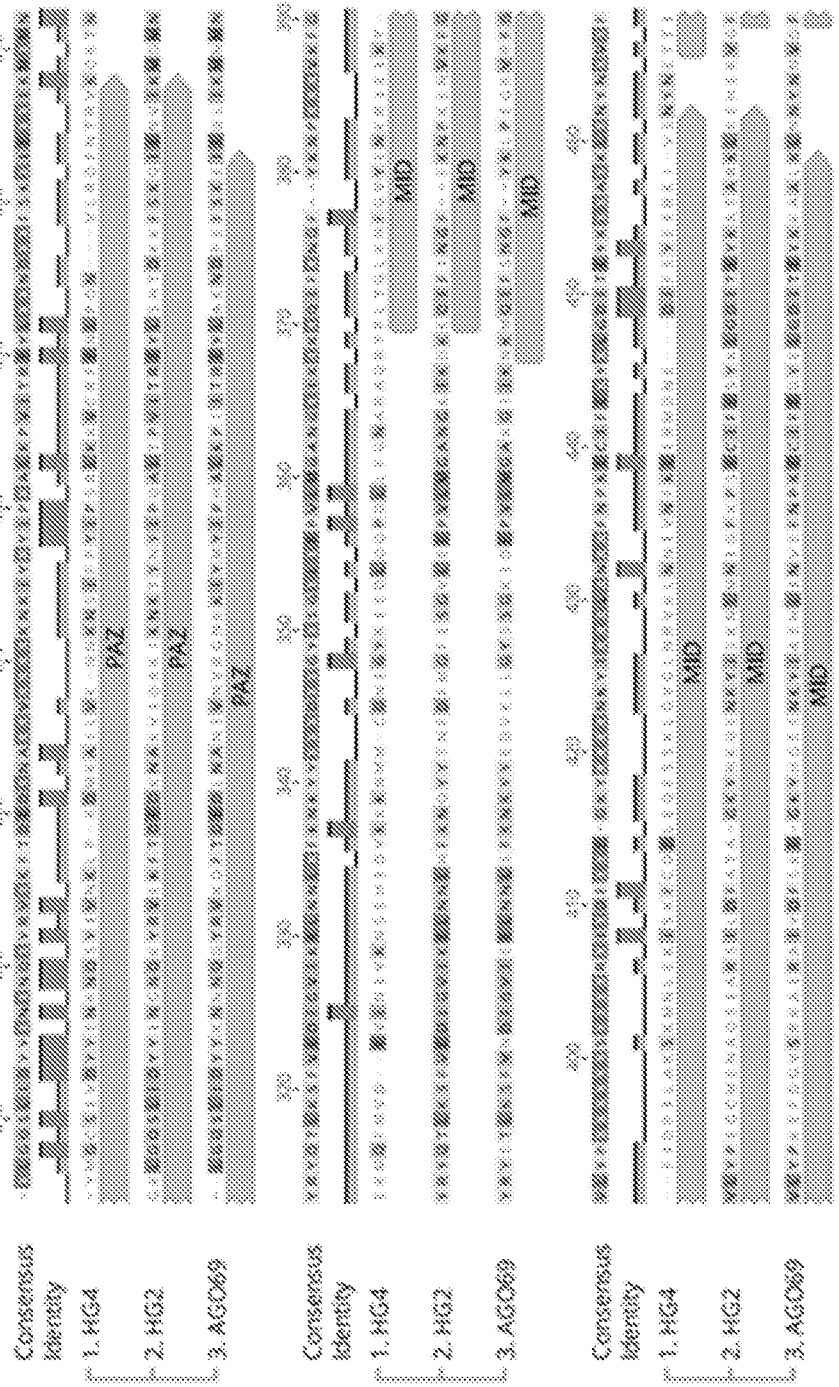


FIG. 103B

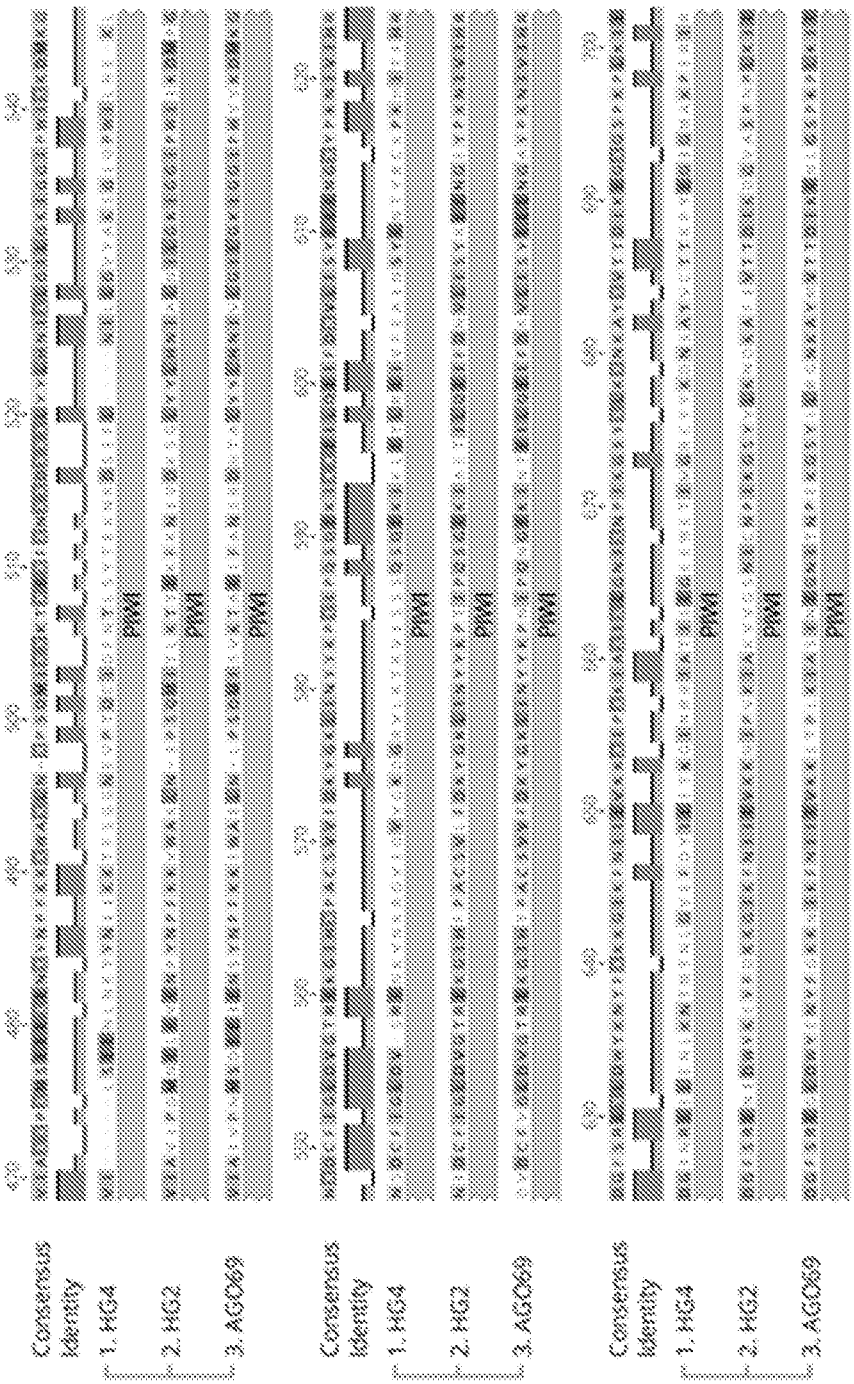


FIG. 103C

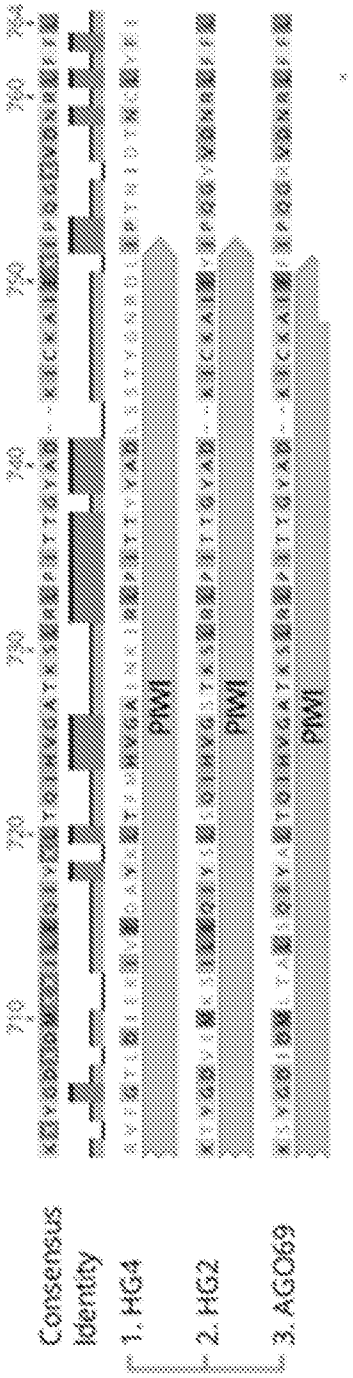


FIG. 103D

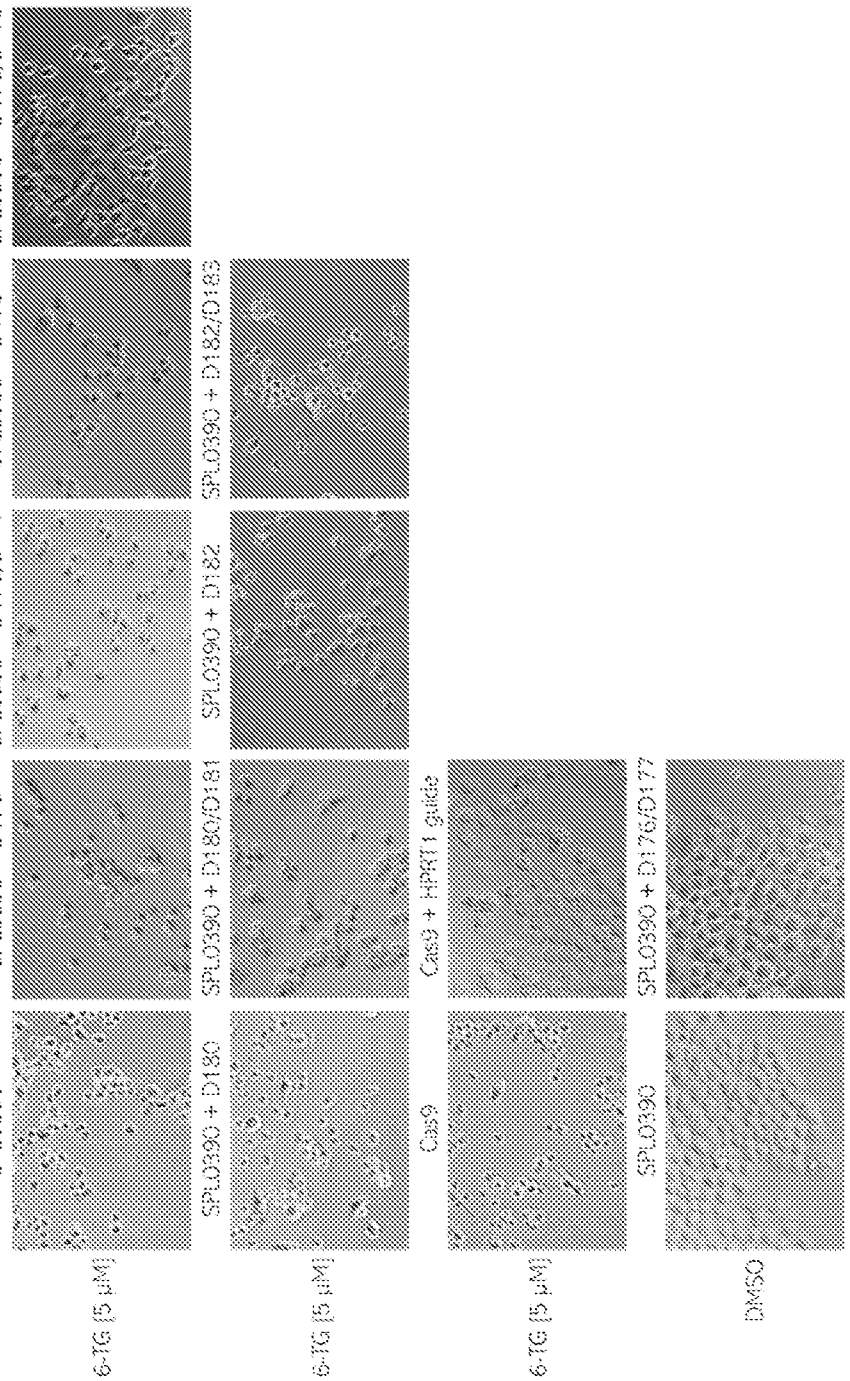

SPL0390
SPL0390 + D176
SPL0390 + D176/D177
SPL0390 + D178
SPL0390 + D178/D179
SPL0390 + D180
SPL0390 + D180/D181
SPL0390 + D182
SPL0390 + D182/D183
Cas9
Cas9 + HPRT1 guide
SPL0390
SPL0390 + D176/D177
6-TG (5 µM)
6-TG (5 µM)
6-TG (5 µM)
DMSO

FIG. 104

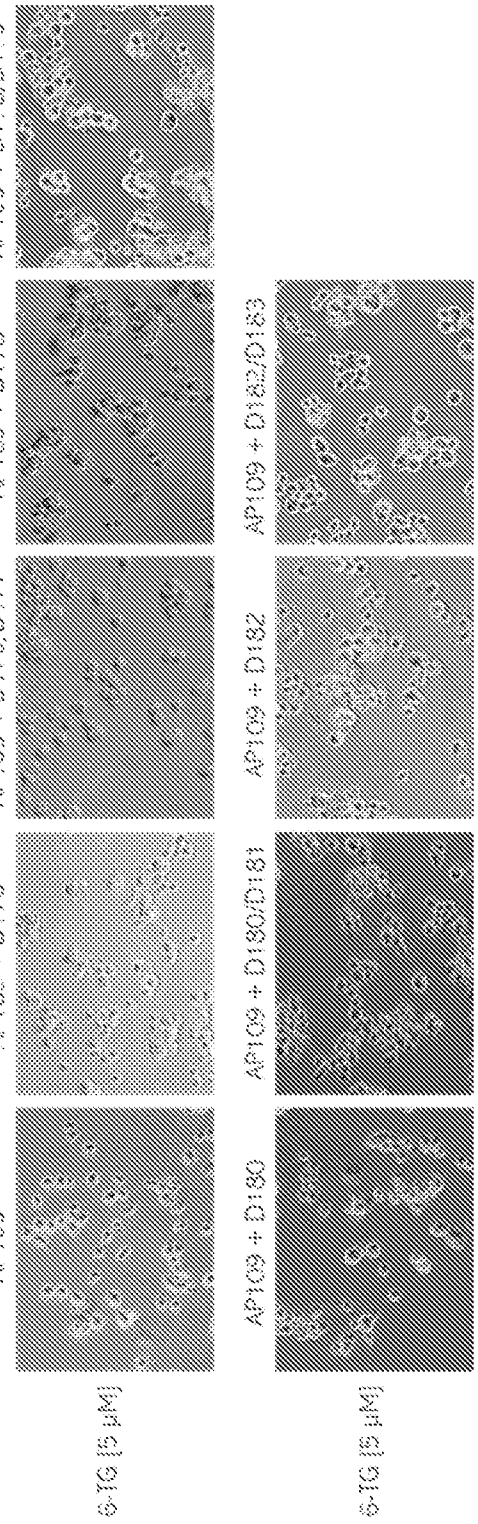

AP109
AP109 + D176
AP109 + D176/D177
AP109 + D178
AP109 + D178/D179
6-TG [5 µM]
AP109 + D180
AP109 + D180/D181
AP109 + D182
AP109 + D182/D183
6-TG [5 µM]

FIG. 105

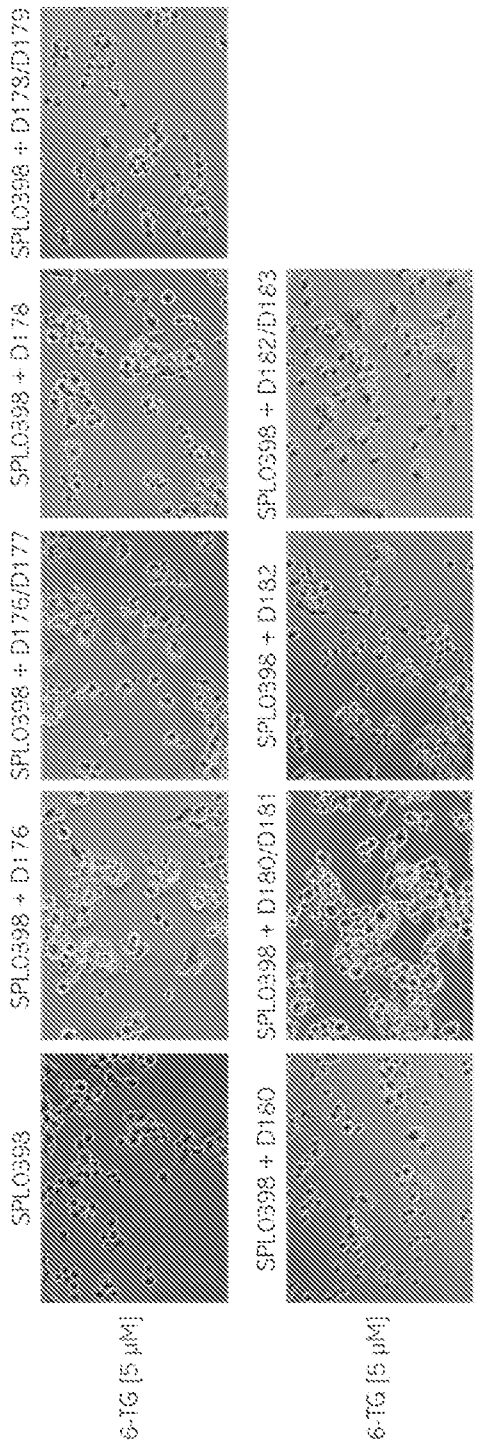
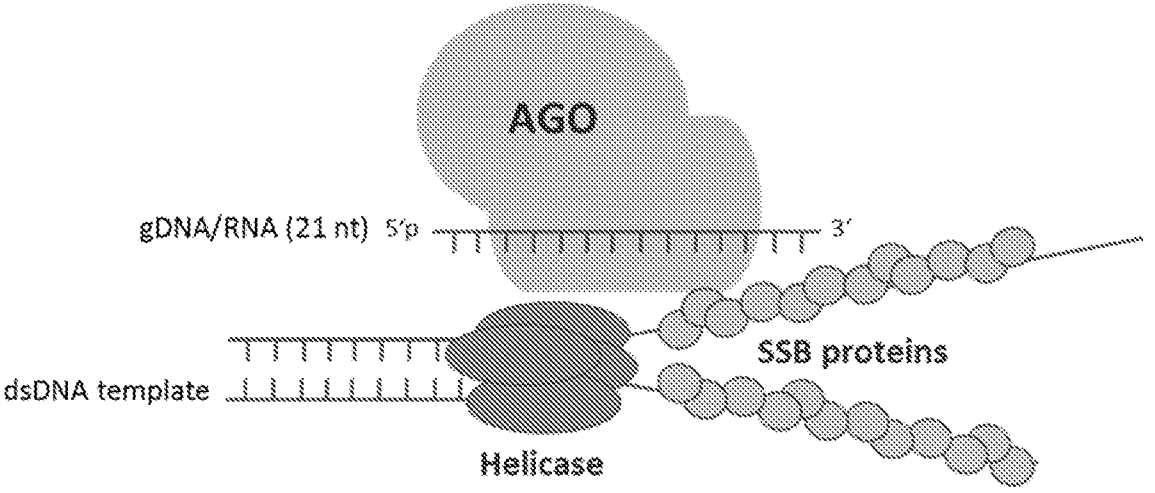

SPL0398
SPL0398 + D176
SPL0398 + D176/D177
SPL0398 + D178
SPL0398 + D178/D179
SPL0398 + D180
SPL0398 + D180/D181
SPL0398 + D182
SPL0398 + D182/D183
6-TG (5 μM)
6-TG (6 μM)

FIG. 106



dsDNA nicking/cleavage

FIG. 44