

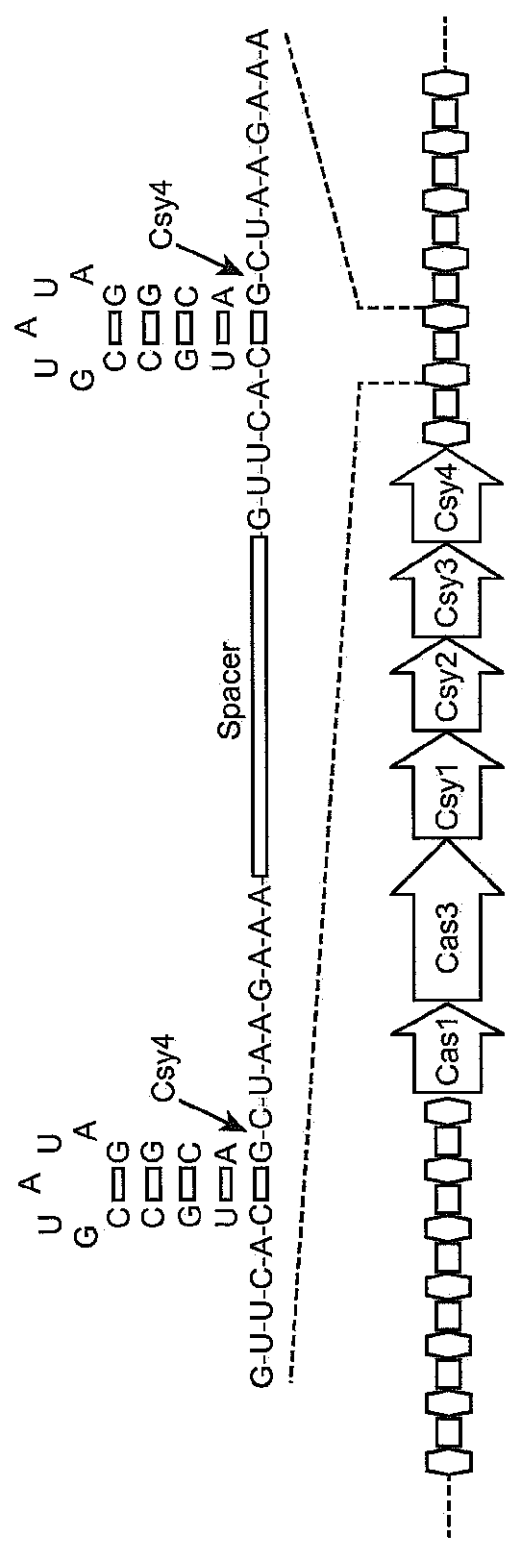


FIG. 1A

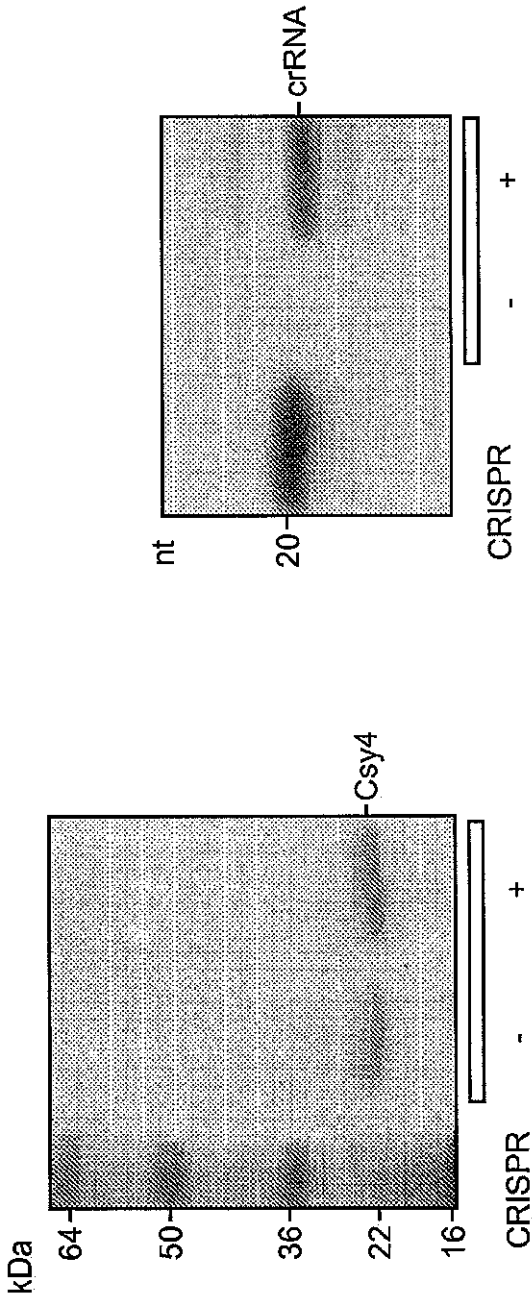


FIG. 1B

FIG. 1C

2 / 29

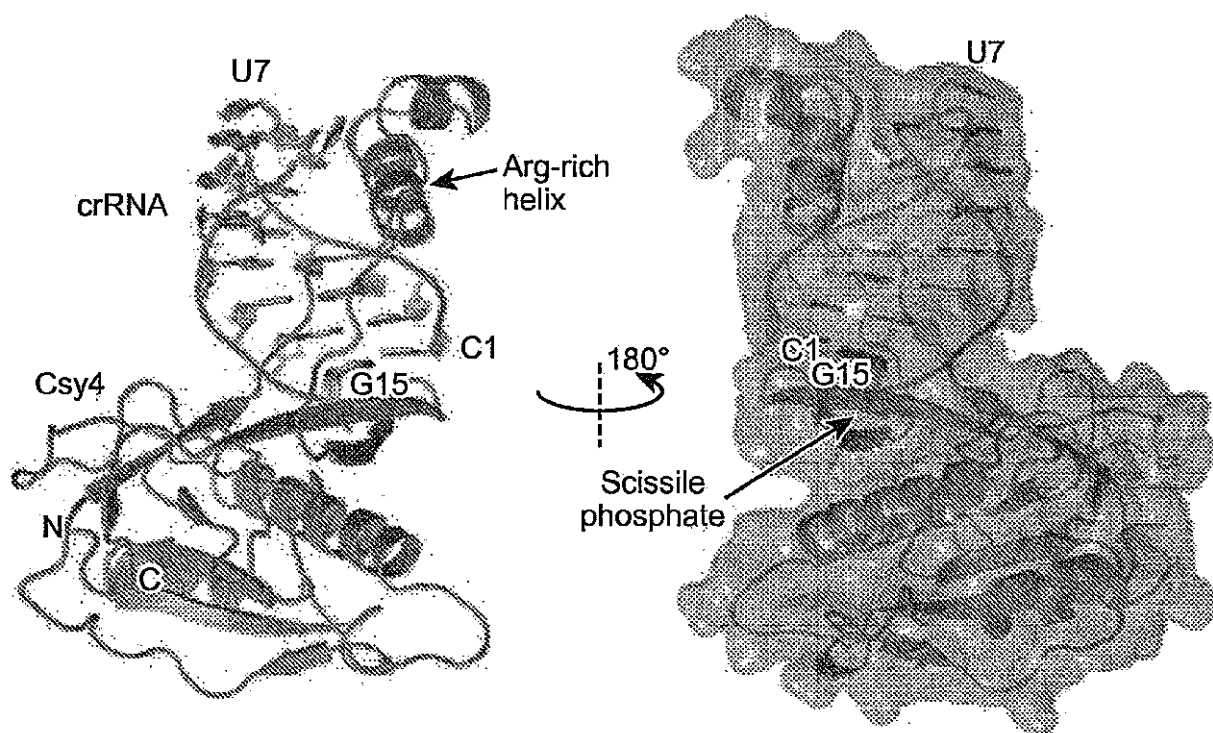


FIG. 2A

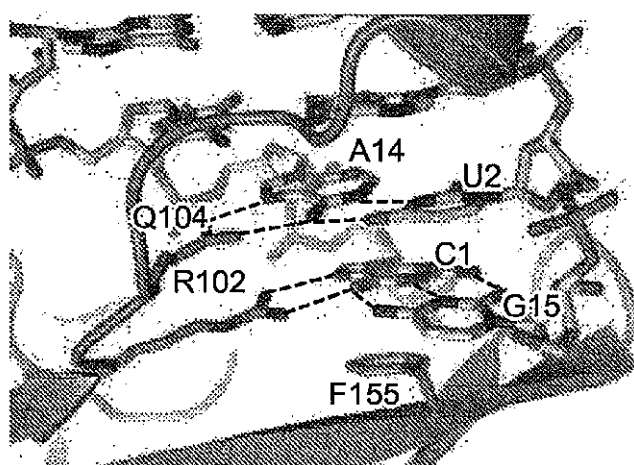


FIG. 2B

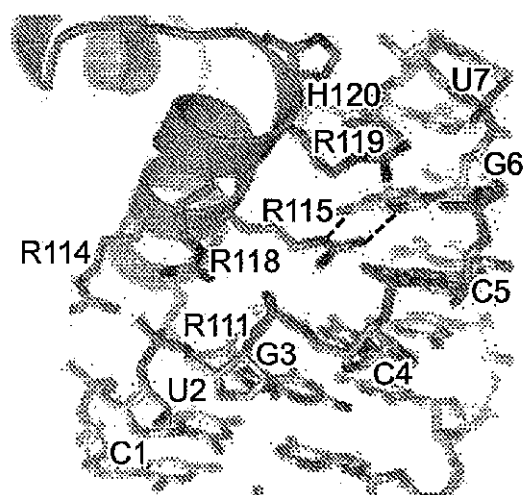


FIG. 2C

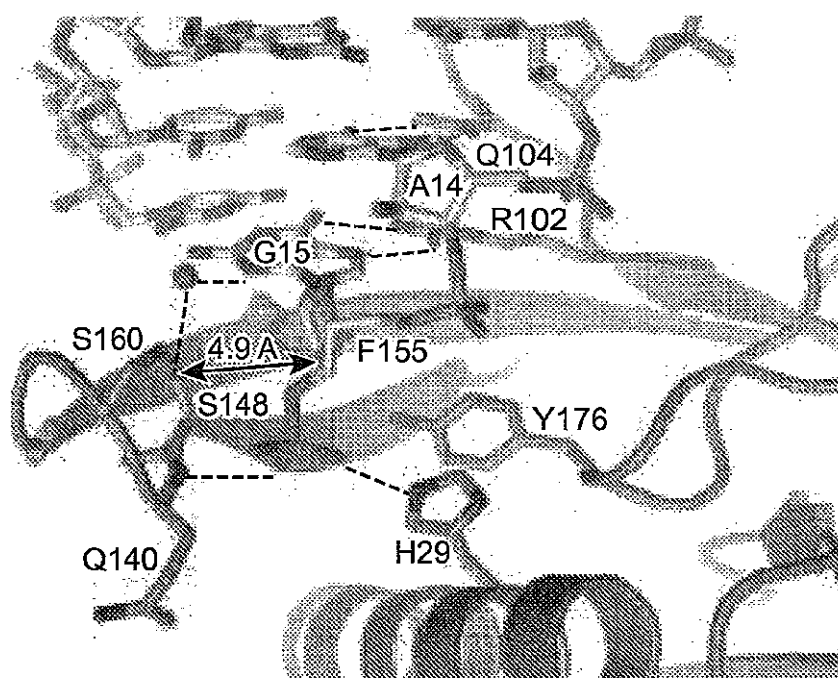


FIG. 3A

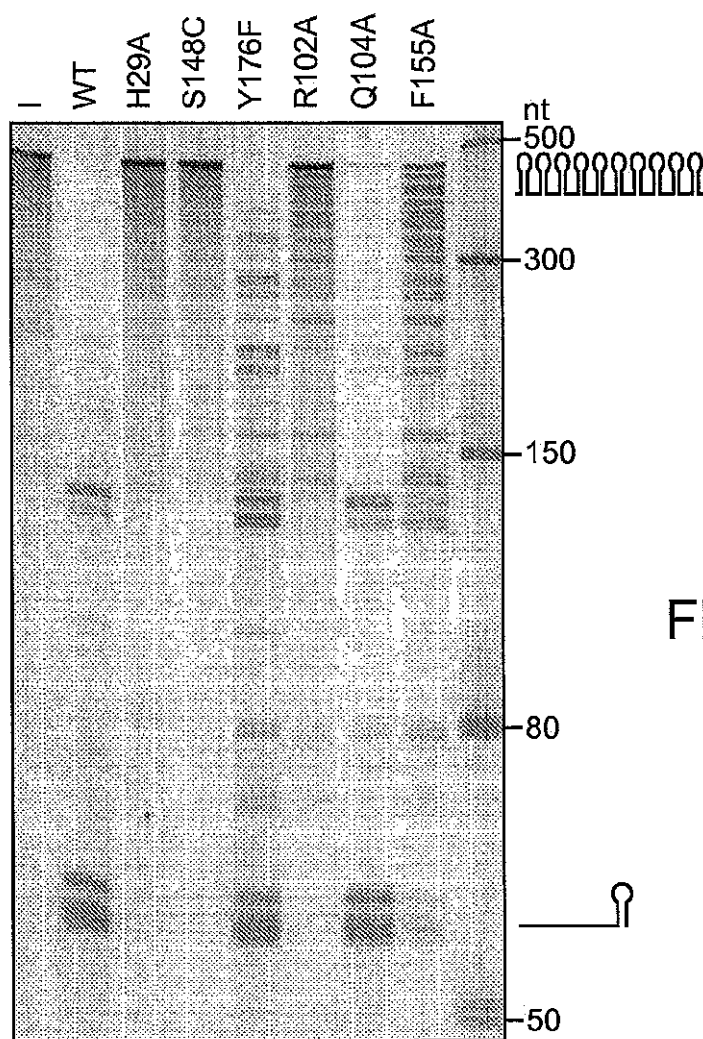


FIG. 3B

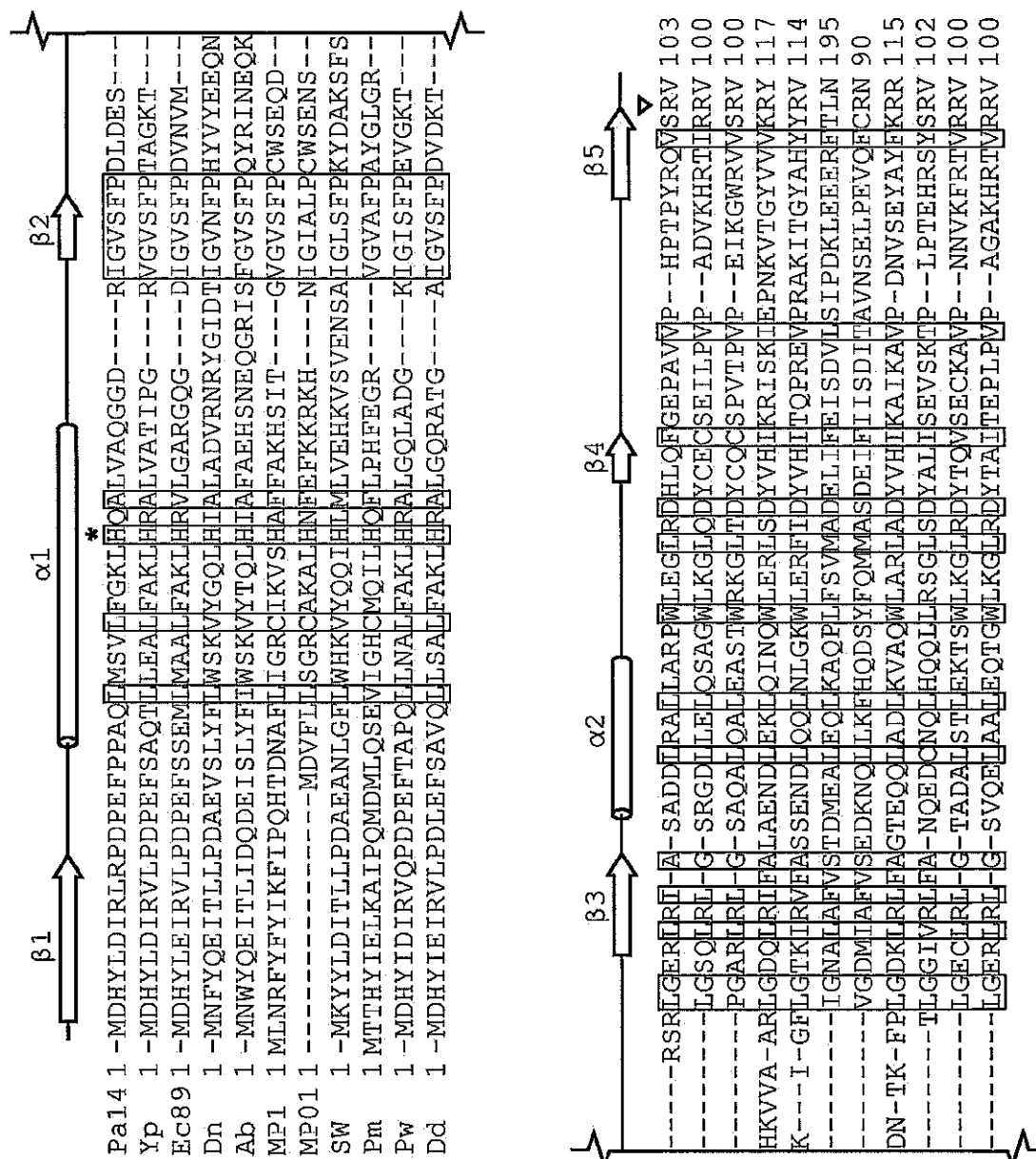


FIG. 4

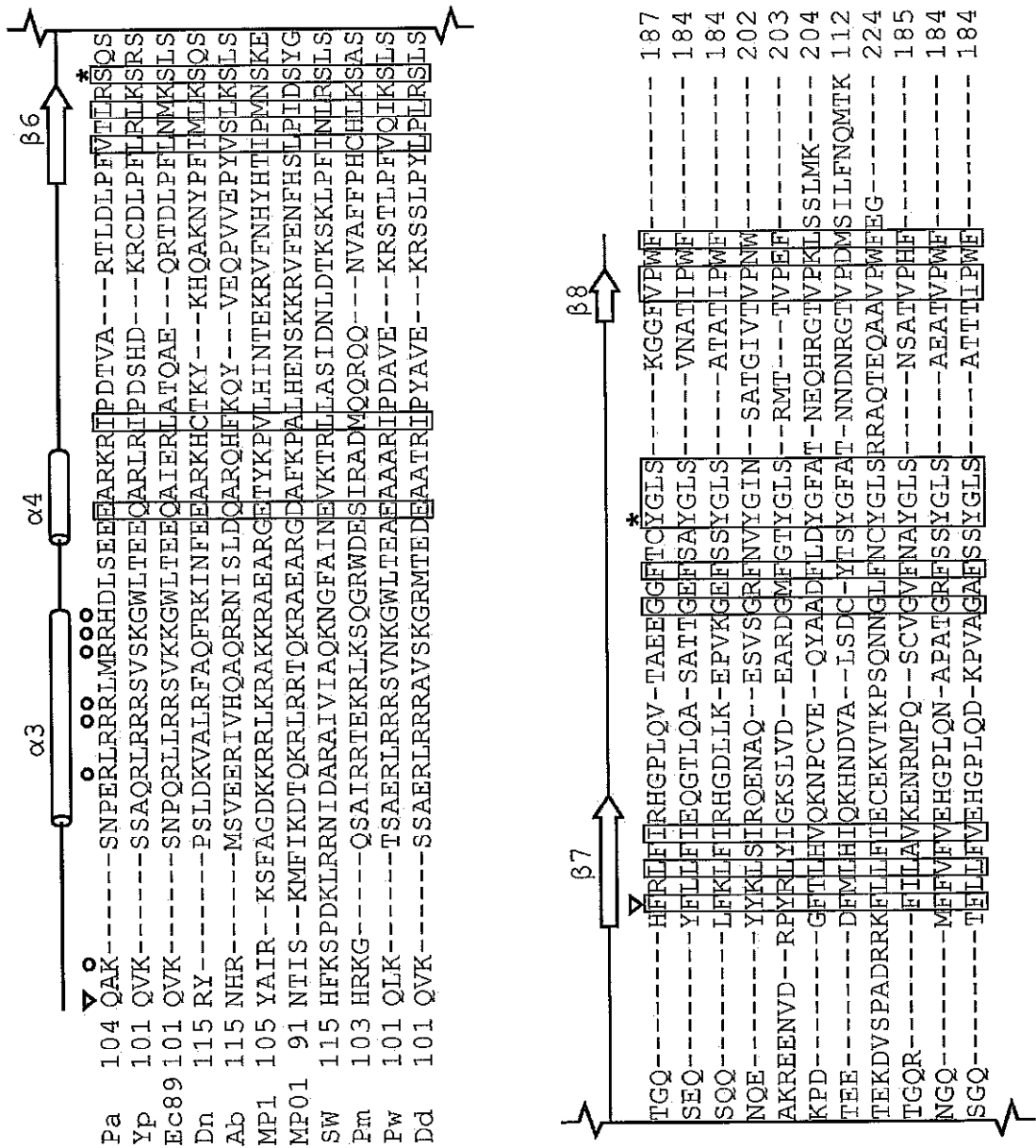


FIG. 4 (Cont.)

FIG. 5A**Csy4 sequence**

>gi|116050369|ref|YP_790814.1| hypothetical protein PA14_33300 [Pseudomonas aeruginosa UCBPP-PA14]
 MSVLFGKLHQALVAQGGDRIGVSFPDLDESRRLGERLRIHASADLRAILARPWLEGLRDHLQFGEPVAVPHPTPYRQVSRVQAKSNPERLRRRLMRRH
 DLSEEEARKRIIPDTVARALDLPFVTLRSQSTGQHFLFIRHGPLQVTAEEGGTCYGLSKGGFVPWF (SEQ ID NO:5)

RNA recognition sequence

GUUCACUGCCGUAGGACGUAAGAAA (SEQ ID NO:1)

FIG. 5B**Csy4 sequence**

>gi|107101871|ref|ZP_01365789.1| hypothetical protein PaerPA_01002916 [Pseudomonas aeruginosa PACS2]
 MDHYLDIRLPDPEFPPAQLMVLFGLHQALVAQGGDRIGVSFPDLDESRRLGERLRIHASADLRAILARPWLEGLRDHLQFGEPVAVPHPTPYRQVSRVQAKSN
 PERLRRRLMRRHDLSEEEARKRIIPDTVARALDLPFVTLRSQSTGQHFLFIRHGPLQVTAEEGGTCYGLSKGGFVPWF (SEQ ID NO:6)

RNA recognition sequences

GUUCACUGCCGUGUAGGACGUAAGAAA (SEQ ID NO:7)
 GUUCACUGCCGUAGGACGUAAGAAA (SEQ ID NO:1)

FIG. 5C

>gi|254235433|ref|ZP_04928756.1| hypothetical protein PACG_01340 [Pseudomonas aeruginosa C3719]
 MDHYLDIRLPDPEFPPAQLMVLFGLHQALVAQGGDRIGVSFPDLDESRRLGERLRIHASADLRAIL
 LARPWLEGLRDHLQFGEPVAVPHPTPYRQVSRVQAKSNPERLRRRLMRRHDLSEEEARKRIIPDTVARTLD
 LPFVTLRSQSTGQHFLFIRHGPLQVTAEEGGTCYGLSKGGFVPWF
 (SEQ ID NO:8)

RNA recognition sequences

GUUCACUGCCGUGUAGGACGUAAGAAA (SEQ ID NO:7)
 GUUCACUGCCGUAGGACGUAAGAAA (SEQ ID NO:1)

FIG. 5D

>gi|254240857|ref|ZP_04934179.1| hypothetical protein PA2G_01531 [Pseudomonas aeruginosa 2192]
 MDHYLDIRLPDPEFPPAQLMVLFGLHQALVAQGGDRIGVSFPDLDESRRLGERLRIHASADLRAIL
 LARPWLEGLRDHLQFGEEAAVPHPTPYRQVSRVQAKSNPERLRRRLMRRHDLSEEEARKRIIPDTVARTLD
 LPFVTLRSQSTGQHFLFIRHGPLQVTAEEGGTCYGLSKGGFVPWF (SEQ ID NO:9)

RNA recognition sequences
 GUUCACUGCCGUGAGGCAGCUAAGAAA (SEQ ID NO:7)
 GUUCACUGCCGUAUAGGCAGCUAAGAAA (SEQ ID NO:1)

FIG. 5E

>gi|242238181|ref|YP_002986362.1| CRISPR-associated protein, Csy4 family [Dickeya dadantii Ech703]
 MDHYIEIRVLPDLEFSAVQLISALFAKLHRAIGQATGAIGVSFPDVKTLGERIRLHGSVOELAALEQT
 GWLKLGRDYTAITEPLPVPAGAKHRTVRRVQVKSSAEIRLRRAVSKGRMTEDEAATRIPIYAVEKRSSLIPY
 LPLRSLSSGQTFLFVEHGPLQDKPVAGAFSSYGLSATTTIPWF (SEQ ID NO:10)

RNA recognition sequences
 GUUCACUGCCGCGUAGGCAGCUUAGAAA (SEQ ID NO:11)
 GUUCACUGCCGAGUAGGCAGCUUAGAAA (SEQ ID NO:12)

FIG. 5F

>gi|261822890|ref|YP_003260996.1| CRISPR-associated protein, Csy4 family [Pectobacterium wasabiae WPP163]
 MDHYIDIRVQDPDEFTAPQLINALFAKLHRAIGQADKIGISFEVGTGECIRLHGTADALSTLEKT
 SWLKLGRDYTQVSECKAVPNVKFTVRRVQVKSSAEIRLRRAVSKGRMTEDEAATRIPIYAVEKRSSLIPY
 VQIKLSLNGQMFVFEHGPLQNAATGTFSSYGLSAEATVPWF (SEQ ID NO:13)

RNA recognition sequence
 GUUCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:14)

FIG. 5G

>gi|253990195|ref|YP_003041551.1| hypothetical protein PAU_02718 [Phototribadus asymbiotica subsp. asymbiotica ATCC 43949]
 MDYFFEILVLPDEFFSKQSIMAEALFAKLHRAIGQVGNRIGVSPFCARKTLGDKLRIHGASEALNDLQAL
 PWLKLGRDYTEIMDIQVPQDTQYRRVSRVQVKSSAEIRLRRAVSKGRMTEDEAATRIPIYAVEKRSSLIPY
 LLVKSLSRQTFFLFI EQGPIEDKPTPGVSSYGLSASATVPWF (SEQ ID NO:15)

RNA recognition sequences
 GUUCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:16)
 GUUCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:17)
 ACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:18)
 GUUCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:19)
 GUUCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:20)
 GUUCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5H

>gi|307132482|ref|YP_003884498.1| hypothetical protein Dda3937_03453 [Dickeya dadantii 3937]
 MDHYIEIRVLDPDFSGVQLLSALEFAKLHRLALGORATGAIGVSFPDAGKTLGERLRLHGSVQELAALEQT
 GWLRLRDYTAITEPLVPAGVKHRTVRRVQVKSSAERLRRRAVNKGRMTVDADARIPYTVKEKRTSLPY
 LPLRLSNGQTFLLFVEHGPLQDKPVAGAFSSYGLSAVATIPWF (SEQ ID NO:22)

RNA recognition sequences

GUUCACUGCCGUGUAGGCAGCUUAGAAA (SEQ ID NO:23)
 GUUCACUGCCGAGUAGGCAGCUUAGAAA (SEQ ID NO:12)

FIG. 5I

>gi|285019813|ref|YP_003377524.1| CRISPR-associated protein, csy4 family [Xanthomonas albilineans GPE PC73]
 MQHYLDLHLRPDPELAPYQLIGALYARLHRSVLTINTTRIGVSFPGHDNRVPTLGTHLRLHGDDSTLHHL
 MATWNLHGVRDHTITSIGAVPSEAVHROVTRVQAKSSPERLRRRAMRRHGISEDLAQVRIIPDSAAEQLR
 LPFVVLGSRSTGQTAFPFVVRHGPVQQEFPVPGDFSSYGLSRGATVPWF (SEQ ID NO:24)

RNA recognition sequences

GUUCACUGCCGUGUAGGCAGCUCAGAAA (SEQ ID NO:25)
 UUCACUGCCGUGUAGGCAGCUCAGAAA (SEQ ID NO:26)
 GUUCACUGCCGUGUAGGCAGCUCAGAAA (SEQ ID NO:27)

FIG. 5J

>gi|50122605|ref|YP_051772.1| hypothetical protein ECA3684 [Pectobacterium atrosepticumSCRI1043]
 MDHYIDIRVQPDPEFTASQLLNALFAKLHRLVGLQANGKIGISFPEVGKTLGECRLRHGTEDALSTLEKT
 SWLKGLRDYTQVSECKVVPNGVKFTRVRRVQLKSSAERLRRRSVSKGWLTAEEAARIPDAVEKRSALPF
 VQIKLSLNGQMFVFEVHGPLQNAPTAGRFSSYGLSTEATVPWF (SEQ ID NO:28)

RNA recognition sequence

GUUCACUGCCGACAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5K

>gi|37525730|ref|NP_929074.1| hypothetical protein plul796 [Photorhabdus luminescens subsp. laumondii TT01]
 MDHYLEIRVLDPDLFSQSQSIFEALFAKLHRLALGOLSNQGVGFPCARKTLGDTLRIHGSSEALNDLQAL
 PWLKGLRDYTEVIDIQPIPOETKYRCVSRVQVKSSAERLRRRAIKKGWLTGEQARQIPISKEQRTHLPP
 LFLKLSLSSGQSFLLFVKQGPIDKPTSGIFSSYGLSSSATIPWF (SEQ ID NO:29)

RNA recognition sequences
 GUGCACUGCCGACAGGCAGCUUAGAAA (SEQ ID NO:17)
 GUUCACUGCCGACAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5L

>gi|297569494|ref|YP_003690838.1| CRISPR-associated protein, Csy4 family [Desulfurivibrio alkaliphilus AHT2]
 MVMAMDCYVEISLLPDPPEFPDSILMNALFAKLHRAAENGKQEI GVSFPEFGKLNKSLRIHGSEESLKR
 LMDLNWIQGMKDYTRVSGIAKVPDSCQYRTVKRVQAKSSVDRLYRRSVKKGWLSEENAEQQKERAREGRL
 KLPFVQLKSQTGTGQQRLFIQHGLQEKPVTFGRFSSYGLSNEATVPWF (SEQ ID NO:30)

RNA recognition sequence
 GUUCACUGCCGACAGGCAGCUUAGAAA (SEQ ID NO:31)

FIG. 5M

>gi|25178340|ref|YP_003003061.1| Csy4 family CRISPR-associated protein [Dickeya zeae Ech1591]
 MDHYIEIRVLPDLEFSAVQLISALFAKLHRLALGQQAIGAVSFDPDVGKTLGERLRLHGSQALTALEQT
 GWRTGLRDYSTITDVLTVPTGAQYRTVRRVQVKSSAERLRRRAVSKGWLTADEAAARIPYAVEKRTSLPY
 LPLRSLSSGQPFLLFVEHGPIQDKPVAGTFSSYGLSATATIPWF (SEQ ID NO:32)

RNA recognition sequences

GUUCACUGCCGUGUAGGCAGCUUAGAAA (SEQ ID NO:23)
 GUGCACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:33)

FIG. 5N

>gi|22125621|ref|NP_669044.1| hypothetical protein y1727 [Yersinia pestis KIM 10]
 MDHYLDIRVLPDPEFSAOTLLEALFAKLHRLALVATIPGRVGVSEFTAGKTLGSQLRLHGSRGDLLEIQSA
 GWLKGLDYCECSEILPVPADVKHRTIRRVQVKSSAQRLLRRRSVSKGWLTEEQARLRIPDSHDKRCDLPF
 LRLKRSSEQYFLLFIEQGTIQASATTGEFSAYGLSVNATIPWF (SEQ ID NO:34)

RNA recognition sequences

GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)
 UGUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:35)

FIG. 5O

>gi|271501952|ref|YP_0033334978.1| Csy4 family CRISPR-associated protein [Dickeya dadantii Ech586]
 MDHYIEIRVLPDPEFSAVQLISALFAKLHRLALGQATGDIGVSFPDAGKTLGERLRLHGSVQALAALEQT
 GWLKGRLDYSTITDVLTVPTGAQYRTVRRVQVKSSAERLRRRAVSKGRMTADEAAARIPYAAEKRTSLPY
 LPLRSLSSGQTFLLFVEHGPIQEKPVAGVFSSYGLSAIATIPWF (SEQ ID NO:36)

RNA recognition sequences

GUGAACUGCCGCAUAGGCAGCUUAGAAA (SEQ ID NO:37)
 GUUCACUGCCGAGUAGGCAGCUUAGAAA (SEQ ID NO:12)

FIG. 5P

>gi|117623067|ref|YP_851980.1| hypothetical protein APEC01_1206 [Escherichia coli APEC 01]
 MAVSLVRNRNKLPMHYLEIRVLPDPEFSSEMMAALFAKLHRLVLGARGQDIGVSFPDVMVMPGARLR
 LHGSAQALQALEASTWRKGLTDYCCQSPVTPVEIKGWRVSRVQVKSNPQRLRRSVKKGWLTEEQAIE
 RLATQAEQRTDLPFLNMKSLSSQQLFKLFIRHGDLLKEPVKGEFSSYGLSATATIPWF (SEQ ID NO:38)

RNA recognition sequence
 GUUCACUGCCGUACAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5Q

>gi|91209927|ref|YP_539913.1| hypothetical protein UTI89 C0896 [Escherichia coli UTI89]
 MDHYLEIRVLPDPESSEMMAALFAKLHRVIGARGQGDIGVSFPDVMMPGARLRHLHGSAQALQALEAS
 TWRKGLTDYCCQSPVTPVPEIKGWRVSRVQVKSNPORLIRSVKRWLTTEEQAIERLATQAEQRTDLPF
 LNKSLSSQQLFKLFIRHGDLLKEPVKGEFFSSYGLSATIPWF (SEQ ID NO:39)

RNA recognition sequence
 GUUCACUGCCGUACAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5R

>gi|237808124|ref|YP_002892564.1| CRISPR-associated protein, Csy4 family [Tolomonas auensis DSM 9187]
 MDHYLDIRLLPEEPEVSESFILNALFAKLHVRLGQQAQGRVGVSPDHHKRLGDLRLHGRQRTDLQALMA
 DDWLQGLKGYTQCSEVLPIPATVSYRAVKRVQAKSAHNKRQRSIAKGWLTSESAQIRIPDTQQKELHLLPF
 VQLKRSRNGQMMRVYVEHGPVLAVPVSGYFNAYGLSSATIPWF (SEQ ID NO:40)

RNA recognition sequence
 CUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:41)

FIG. 5S

>gi|259907505|ref|YP_002647861.1| CRISPR-associated protein Csy4 [Erwinia pyrifoliae Ep1/96]
 MDHYQDIRVRVDEENGAEVLLAQVFMHLHQVLMFAANGRIGISFENVKRTLGDRIHLHGTLLDLSALQQS
 GWNKCLRDIACSDIAPVPKGAARVTVRVQVKSSAERLRRSVNKGWLSQEAAERISVLNEQRSNLPP
 LQIKSGSNGQAWRLFIEHGSILVSAPSDGSFSSYGLSAAATIPWF (SEQ ID NO:42)

RNA recognition sequences

GUUCACUGCCGACAGGCAGCUUAGAAA (SEQ ID NO:19)
 UUCACUGCCGUACAGGCAGCUUAGAAA (SEQ ID NO:43)

FIG. 5T

>gi|218688670|ref|YP_002396882.1| hypothetical protein ECED1_0855 [Escherichia coli ED1a]
 MAVSLVRNRNKELPMDHYLETRVLPDPEFSSEMMAALFAKLHRVLGARGQGDIGVSFPDVMVMPGTHLR
 LHGSAQALQELEASTWRKGLTDYCQCSPTVPPEIKGWRVSVRVQVKSNPORLLRRSVKKGWLTTEEQAIE
 RLATQAEQRTDLPFLNMKSLSSQQFKLFIHGDLLKEPVKGEFSSYGLSATATIPWF (SEQ ID NO:44)

RNA recognition sequence

GUUCACUGCCGUAACAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5U

>gi|121608426|ref|YP_996233.1| CRISPR-associated Csy4 family protein [Verminephrobacter eiseniae EF01-2]
 MSTHYIDITLRPDPEFSPAHLINAHQHLALVQLGTGDVGVSFPGFILRGEHSHLGTTLRLHGATSAL
 QRLQALSWLRGMRDHVKTSEVAPVPTHQHRVVRVQAKSSPERSRRRLMRRLIDEAQALQRIIPDQEGR
 RLALPYLRLQASAKGQVFRLFIEHGPLLDTSPGSGTYGLSTQATIPWF (SEQ ID NO:45)

RNA recognition sequences

GUUCACUGCCGGAUAGGCAGCUCAGAAA (SEQ ID NO:46)

FIG. 5V

>gi|34497206|ref|NP_901421.1| hypothetical protein CV 1751 [Chromobacterium violaceum ATCC 12472]
 MDHYLDIRLLPDADFPPVIMNALYAKLHRAIAAQQRQDVGVSFPGYDPAPSSHDGKPLPPTLGLTLRLH
 GSAAALDGLMARRWLSGFADHAIIVGDIRPVAGASAVSVRRQAKSSPARARDRLMRROGISAEERRI
 PDETAQRINLFLYLTVDASTGTGQCFRLFVEQQAAPSIAGSFNAYGLSAAALPAW (SEQ ID NO:47)

RNA recognition sequence

GUUCACUGCCGGUAAGGCAGCUUAGAAA (SEQ ID NO:48)

FIG. 5W

>gi|188532992|ref|YP_001906789.1| hypothetical protein ETA 08450 [Erwinia tasmaniensis Et1/99]
 MDRYQDIRVRVDAEMTAPVLLAQVFMRLHQVLMRAANGRIGISFPDVKLTLGDRIRLHGTLLDLSLQQS
 GWDKGLTDYIACSAIDVPVPGAAWRTVRRVQVKSSAFRLRRSVNKGWLNAAERINVLSEQRSDLPY
 LQIKSGSNCHAWRLFIEHGPLVSVPVNGGFSYGLSATATVPWF (SEQ ID NO:49)

RNA recognition sequences

GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)
 GUUCACUGCCGGUACAGGCAGCUUAGAAAG (SEQ ID NO:50)

FIG. 5X

>gi|160896663|ref|YP_001562245.1| CRISPR-associated Csy4 family protein [Delftia acidovorans SPH-1]
 MAMTSHYIDTTLLPDPEFSAHLILGALVAKLHRAIVQLGSTDIGISFPGYSLRPRTLGTILRHGSEAL
 RGLMEQPWLQMRDHVHCTPPALVPEGAVPCLVQRQFKTSPDLRRRRMRKGETAEQAAAIPDSVER
 TPDLPYVQLRSASTGQPFCLFVEQKAVQGTAGQEGNTYGLSLGTAVPWF (SEQ ID NO:51)

RNA recognition sequence

GUUCGUGCCGGUAGGCAGCUUAGAAA (SEQ ID NO:52)

FIG. 5Y

>gi|146311064|ref|YP_001176138.1| CRISPR-associated Csy4 family protein [Enterobacter sp. 638]
 MDHYLEIRVLSDPEFSEETIMAALEAKLHRAIGARGQGDIGVSFPYSLKPGDTILRLHGSAQSLDELEKM
 AWKGLSDYCLCKGVLPAPDVNAWRCVSRVQVKSSPQRLMRRSVKKGWLTETEEAQQRLLNLQEARLDLPW
 LNLQSLSTGQSFRFLFIRHGDIVDMPMCGEFSSYGLSATATIPWF (SEQ ID NO:53)

RNA recognition sequence
 GUUCACUGCCGUACAGGCAGCUUAGAAA (SEQ ID NO:21)

FIG. 5Z

>gi|289209612|ref|YP_003461678.1| CRISPR-associated protein, Csy4 family [Thioalkalivibrio sp. K90mix]
 MDHYLDLRVMDPEEFKETTLLGALVSKLHRRRLVMSADDIGISLPDHEQEPEPLGRRRLRVHGTQGRLLNLLM
 QDEWLGMQSLVDATPVQVPDQVTPYRPVRRRQYKTNAERLRRRRMRHGESYEEARQHIDIVERRVNT
 PFLSVQSASTGQRFSLFTEHGPQQHASPGRFNTYGLSQDATVPWF (SEQ ID NO:54)

RNA recognition sequence
 GUUAGCUGCCGCACAGGCAGCUCAGAAA (SEQ ID NO:55)

FIG. 5AA

>gi|283856310|ref|YP_162420.2| Csy4 family CRISPR-associated protein [Zymomonas mobilis subsp. mobilis ZM4]
 MLANPVDSYQDIYILPNOETAPHIIMEKLFSLHLLELVRLGSOHIGISFPEHDNNKPCLGSRRLRHGTGA
 DLHELALSGWITRLDDYLYCEDIKSVPEIRQYCVVSRVQAKSSPARLRRRAIRRHGFHDEEAKKVIPTDA
 FERLELPFIMTGSCKTKQPRFPVFISHKLIQNKLMNGNFNSYGLSLGASVPWF (SEQ ID NO:56)

RNA recognition sequence
 GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)

FIG. 5AB

>gi|260752821|ref|YP_003225714.1| CRISPR-associated protein, Csy4 family [Zymomonas mobilis subsp. mobilis NCIMB 11163] MLANPVDSYQDIYILPNQEIAPHIMEKLFSLHLEIVRIGSQHIGISFPEDHNNKPCLSRLRLHGAGA DLHELALSGWITRLDDYLYCEDIKSVPEIRYOYCVSRVQAKSSPARLRRRAIRRHGFHDEAKKVIPTDA FERLELPFIMTGSCSTKQPRFPVFISHKLIQDKLMNGNFNSYGLSGASVPWF (SEQ ID NO:57)

RNA recognition sequence
GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)

FIG. 5AC

>gi|121592915|ref|YP_984811.1| CRISPR-associated Csy4 family protein [Acidovorax sp. JS42] MTTHYINITLLPDPEFSAHLILGALVAKLHRAIVQGHITTDIGVYPQHVSQPLTKRTIGAVIRLHGTPFA IQRMEEDWLKGRDHTQVGEILLPVANAQHRVRRRQKTNADRLRRRMQRKGETAEQAAAAIPDTVE RRPDLFPVQLRSSSTGQSFCLCVHGHGPLQLPVPAGAFNAYGLGHDAIVPWF (SEQ ID NO:58)

RNA recognition sequences
GUUCACUGCCGCAUAGGCAGCUCAGAAA (SEQ ID NO:59)

FIG. 5AD

>gi|317051103|ref|YP_004112219.1| Csy4 family CRISPR-associated protein [Desulfurispirillum indicum S5] MDSYIEIRILPDQEFEATTIMSTVFAKLHRAIVESGSDIGVSFPEAGKTPGALLRLHGSIAALESIMTL SWLFGLODYTOTSGILOVPAQAAAYQVARVQSKMTASRIIRALKRGSLSSEERALELLQSRDQINQPFRL LSASTAQKFFPLFIEQRNAEKAGKQSVYSAYGLSVGGSTVPWF (SEQ ID NO:60)

RNA recognition sequence
GUUCACUGCCGCAUAGGCAGCUCAGAAA (SEQ ID NO:59)

FIG. 5AE

>gi|54303643|ref|YP_133636.1| hypothetical protein PBPRB1991 [Photobacterium profundum SS9]
 MNDSYVDIQLKPDAREAEISSKVFTHKALATINTNKIGISFPQMNKLGRLFRIHGNASLLKDLQG
 IKWLGALAGYCQVGEITVVPDQVQYRVISVKRSNLSKAKLRLLIARGSIDKDGKRYKVRMLSQGFDPFY
 LDLFSSSTGQVYRKFFEFQDQIATSVSDEFDSYGLSNTATIPWF (SEQ ID NO:61)

RNA recognition sequence

GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)

FIG. 5AF

>gi|54292953|ref|YP_122340.1| hypothetical protein plp10047 [Legionella pneumophila str. Lens]
 MDHYLDISILPDSEFTTPVIMNAIYTNLHKALHTLASTNIGVSFPKYSSTLGNLLRIHGKKEALQELQNL
 NWIGGMIGYCEASLIKTVPADTKFRTVSRKQPTMSQSKLRLLKRNSLTETDEIRQYKAKMFSKGLDNPYI
 ELVSVSNGQRHRRYIEFGELFNEPIPGLFQDFGLSNSATVPWF (SEQ ID NO:62)

RNA recognition sequences

GUUCACUGCCGUACAGGCAGCUUAGAAA (SEQ ID NO:21)
 GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)

FIG. 5AG

>gi|54295752|ref|YP_128167.1| hypothetical protein lpl2842 [Legionella pneumophila str. Lens]
 MDHYLEISILPDSEFTTPVIMNAIYTNLHKALHTLASTSIGVSFPKYSSTLGNILRIHGKKEVLDLQNL
 NWIGGMIGYCEASLIKTVPADTKFRTVSRKQPTMSQSKLRLLKRNSLTETDEIRQYKAKMFSKGLDNPYI
 ELVSVSNGQRHRRYIEFGELFNEIPSPGLFDQFGLSNSATVPWF (SEQ ID NO:63)

RNA recognition sequences

GUUCACUGCCGUACAGGCAGCUUAGAAA (SEQ ID NO:21)
 GUUCACUGCCGCACAGGCAGCUUAGAAA (SEQ ID NO:19)

FIG. 5AH

>gi|146292921|ref|YP_001183345.1| CRISPR-associated Csy4 family protein [Shewanella putrefaciens CN-32]
 MNSYIDIRIKPDAMREAEISSKVFTKFKALVTLNSHKIGISFPQMKLSIGQLFRIHGDA SLLHDLQGL
 DWLGPLAGYCQVTAVSAVPDHVQYRIVSVKRSNLSKAKIKRLIARGSIDKDGEKRYKVKMLGQGFDNPYL
 DLFSSSTGQYRKFFEFSDIQAHPLDGEFDSYGLSKTATVPWF (SEQ ID NO:64)

RNA recognition sequence

GUUACCGCCGCACAGCGGCUUAGAAA (SEQ ID NO:65)

FIG. 5AI

>gi|296106567|ref|YP_003618267.1| hypothetical protein lpa 01476 [Legionella pneumophila 2300/99 Alcoy]
 MDYVVDILIKPDSEKSNFLSTLYTKLHKVLHDMASTNIGVSFPHYNTILGNILRIHKKVVLDELLGM
 NFLSGINNYEVSPIKSVPADSKFRIISRKQITMSQSKMRLLFKRGSMTVGDIRQYKAKMFAKSIDNPYL
 ELVSGSNGYRYYRIEFGELLDPVYGEFDRFGLSKTATVPWF (SEQ ID NO:66)

RNA recognition sequence

GUUACUGCCGCACAGCGGCUUAGAA (SEQ ID NO:67)

FIG. 5AJ

>gi|260772736|ref|ZP_05881652.1| hypothetical protein VIB 001192 [Vibrio metschnikovii CIP 69.14]
 MDSYIEIRLOPDAEMPEAEISSKVFTKFKALVLLHSNQIGISFPEVNVKLGRFLRHGEASFLHDLQGL
 NWLGPLSGYCOVSEILAIPEQVQYRIVSVKRSNLSQAKLRRLIARGSIDKEGERRYKVKMLSQGFNDNPYL
 DLFSSSTKQVHRKFFEFGEIQPLPVSGKFD SYGLSHTTVPWF (SEQ ID NO:68)

RNA recognition sequences

GUUACUGCCGCACAGCGGCUUAGAAA (SEQ ID NO:69)
 GUUACUGCCGCAUAGCGGCUUAGAAA (SEQ ID NO:69)

FIG. 5AK

>gi|157146437|ref|YP_001453756.1| hypothetical protein CKO_02197 [Citrobacter koseri ATCC BAA-895]
 MATTVPVAVKGRVTSRVQVKSSPQRLRRSVRKGLTFFEQALRLVESTEQHSDDLPLYLNVKSLSNQQQF
 RVFIRHSELSEPVSGTFTSYGLSSTATIPWF (SEQ ID NO:70)

RNA recognition sequence
 UGUUCACUGCCGUACAGGAGCUUAGAAA (SEQ ID NO:21)

FIG. 5AL

>gi|262402803|ref|ZP_06079364.1| hypothetical protein VOA_000785 [Vibrio sp. RC586]
 MDAYIDIRLMPDAEMREAEISSKVFIKFKALVKLRNSKIGISFPEANIKLGRFLRHGEMSAHDLQGL
 NWLGLAGYCKITTTVTHVPDQVYRIISVKRSNLSKAKLRLIARGSIDKDGEKRYKVKMLSQGFDPY
 DLSSTSGQVYRKFFEFSDIQADPVDGEFDSYGLSKTATVPWF (SEQ ID NO:71)

RNA recognition sequences
 GUUCACUGCCGACAGGAGCUUAGAAA (SEQ ID NO:19)
 AGUGUUCUGCCGAUAGGACGCUAAGAA (SEQ ID NO:72)

FIG. 5AM

>gi|229523353|ref|ZP_04412760.1| hypothetical protein VIF_000211 [Vibrio cholerae TM 11079-80]
 MMDAYIDIRLMPDAEMREAEISSKVFIKFKALVKLRNSKIGISFPEANIKLGRFLRHGEMSAHDLQGL
 LNWLGPLAGYCKITTTVTHVPDQVEYRIISVKRSNLSKAKLRLIARGSIDKDGEKRYKVKMLRQGFDPY
 LDLSSTSGQVYRKFFEFSDIQAEPPVDGEFDSYGLSKTATVPWF (SEQ ID NO:73)

RNA recognition sequence
 GUUCACUGCCGACAGGAGCUUAGAAA (SEQ ID NO:74)

FIG. 5AN

>gi|237748015|ref|ZP_04578495.1| crispr-associated protein [Oxalobacter formigenes OXCC13]
 MKHYIEITLTGSPDFLYHLWSKLYTQLHIALVENRDASQVNIGVSFPEYFNEKGMGLGTKLRLEA
 EDETSLOKIDIQKWFVRINDCIHTTPVCRVPLNEITGYATFSRHHKIKSNAERLARQMKRHKDLSFHETV
 QRYQKNIAKSPLPFIQLESLTNSHPFKLFIEKKPAINASLKVFTTYGLSAESTIPEF (SEQ ID NO:75)

RNA recognition sequence

GUUCACUGCCGUUAUAGGCAGCUUAGAAG (SEQ ID NO:76)

FIG. 5AO

>gi|119945137|ref|YP_942817.1| CRISPR-associated Csy4 family protein [Psychromonas ingrahamii 37]
 MKYYLDITLLPDIEIPLGFIWQVFOOVHIALADNKGVENESDIALSLPNYGDKAFFPLGNKLRLEFSVSEQ
 ALERLAITKWKIRFTDHTITSVKAVPESANEVACFTRKQFNTNISRLARRRAKRHMETFEKALQYYDNF
 AEEQTKLPFMNIKSLTNNAQFRIFERSITKIPKQGTFCYGLSQAIAIVPWF (SEQ ID NO:77)

RNA recognition sequence

GUGUUCGCCGUGCCCCACGGGAUAGACCG (SEQ ID NO:78)

FIG. 5AP

>gi|146328647|ref|YP_001209099.1| hypothetical protein DNO_0170 [Dichelobacter nodosus VCS1703A]
 MNFYQEITLLPDAEVSLYFLWSKVYQQLHIALADVNRNRYGIDTIGVNFPHYVVEEQNHKVVAARLGDQLR
 IFALAENDLEKLQINOWLERLSDYVHIKRIKIEPNKVTGYVVKRYRYPSLDKVALRFAQFRKINFEEA
 RKHCITKYKHQAKNYPFIMLKSSQSNQEYYKLSIRQENAQESVSGRFNVYGINSATGIVTPNW (SEQ ID NO:79)

RNA recognition sequence

GUUCACGCCGCACAGCGGCUUAGAAA (SEQ ID NO:65)

FIG. 5AQ

>gi|160876478|ref|YP_001555794.1| CRISPR-associated Csy4 family protein [Shewanella baltica OS195]
 MNHYLDITLLPNEEVGHYFLWEKLYHQVHLALVEHKNRVGFQFEIAAFAFPQFNEMDNLSGSKIRLLATQPO
 HLEDLKVSNNWLRHFTDYLIHSSIRPVPEKIEVVAYSRRPAIRANKAREIARRMKRHNETHLEQATAHFEFG
 KPKKTKAPFVVMQSYTKDSRFPFLFIQQTHSAVVKEGVSFDSYGLSSRGYLPKF (SEQ ID NO:80)

RNA recognition sequence
 GUUCACGCGGCACAGGCGGCUUAGAAA (SEQ ID NO:65)

FIG. 5AR

>gi|153001745|ref|YP_001367426.1| CRISPR-associated Csy4 family protein [Shewanella baltica OS185]
 MNHYLDITLLPNEEVGHYFLWEKLYHQVHLALVEHKNRVGFQFEIAAFAFPQFNEMDNLSGSKIRLLATQPO
 HLEDLKVSNNWLRHFTDYLIHSSIRPVDPKIEVVAYSRRPAIRANKAREIARRMKRHNETHLEQATAHFEFG
 KPKKTKAPFVVMQSYTKDSRFPFLFIQQTHSAVVKEGVSFDSYGLSSRGYLPKF (SEQ ID NO:81)

RNA recognition sequence
 GUUCACGCGGCACAGGCGGCUUAGAAA (SEQ ID NO:65)

FIG. 5AS

>gi|169795154|ref|YP_001712947.1| hypothetical protein ABAYE1000 [Acinetobacter baumannii AYE]
 MNWYQETITLIDDEISLFTWSKVYTLHIAFAEHSNEQGRISFGVSFPQYRINEQKKIGFLGKIRVF
 ASSENDLQQLNLGKWLERFDYVHITOPREVPRAKITGYAHYRVNHRMSVEERIVHQQRNLSLDQAR
 QHFQKVEQPVVEPVVSLKLSAKREENVDRPYRLYIGKSLVDEARDGMFGTYGLSRMTTVPEF (SEQ ID NO:82)

RNA recognition sequence
 GUUCAUGGCGGCAUACGCCAUUAGAAA (SEQ ID NO:83)

FIG. 5AT

>gi|213158184|ref|YP_002320235.1| crispr-associated protein, Csy4 family [Acinetobacter baumannii AB0057]
 MNWQETLIDQDEISLYFIWISKVYTLHIAFAEHSNEQGRISFGVSFPQYRINEQKKGIFLGTKIRVFA
 SSENDLQQLNIGKWLRFIDYVHITQPREVPRAKITGYAHYRVNHRMSVEERIVHQAQRNISLDQARQ
 HFQKVVEQPVVEFYVSLKSLSAKREENVDRPYRLYIGKSLVDEARDGMFGYGLSRMTTVPEF (SEQ ID NO:84)

RNA recognition sequence
 GUUC AUGGCGCAUACGCCAUUUAGAAA (SEQ ID NO:83)

FIG. 5AU

>gi|15602173|ref|NP_245245.1| hypothetical protein PM0308 [Pasteurella multocida subsp. multocida str. Pm70]
 MTTHYTELKAIPQMDMLQSEVIGHCMQILHQFLPHFEGRVGVAFPAYGLGRTLGIVRLFANQEDCNQLH
 QOILRSGLSDYALISEVSKTPLPTEHRSYSRVHRKQSSAIRTEKRLKSQGRWDESIRADMQQRQQNVAF
 FPHCHLSASTGTGQRFILAVKENRMPQSCVGVFNAYGLSNSATVPHF (SEQ ID NO:85)

RNA recognition sequences
 GUUCACCAUCGUGUAGAUAGCCUUAGAAA (SEQ ID NO:86)
 GUUACUGCCGUAUAGGCAGCUUAGAAA (SEQ ID NO:87)

FIG. 5AV

>gi|293390434|ref|ZP_06634768.1| Csy4 family CRISPR-associated protein [Aggregatibacter actinomycetemcomitans D7S-1]
 MTVQTHYIEIKAIPOVDMLQTEVIGFCLOKLHQILPHFEGRIGLAFFPAYGNDKTLGGIIRLFGTENDCGF
 IHFKLOSLRDYALISEVMPPEKVRSYRIYQRIQPKGOSSIRRAEKRLTAQGWNEEVLQNMLOKQATQR
 IYPAHLKSSSTKQQLAILAIKSVHQTKAVEGVFSAYGLSQTTVPHF (SEQ ID NO:88)

RNA recognition sequence
 CUUCACUGCCGAUAGGCAGCUUAGAAA (SEQ ID NO:89)

FIG. 5AW

>gi|152996699|ref|YP_001341534.1| CRISPR-associated Csy4 family protein [Marinomonas sp. MWYL1]
 MKHYIDITLLPSSDDIGVHFLWSKLMQVHLALVEIQNEOKQVPVAVSFPKYQPRENEKLGFGVGNKLRIFA
 NDKTDLFRINFGKWLHRLDYVHIKSIADVNDVISYSEFNRRSKSGSPDKHIKRRMQRHNETWEQAAAF
 FKGYSMEKADKDLPFIRMKSLHSDNEFCMSIIRKEAAPSNKHIMFNTYGLSAEGVLPKF (SEQ ID NO:90)

RNA recognition sequence
 GUUCGCCGCCGAGCAGCGGCUUAGAAA (SEQ ID NO:91.)

FIG. 5AY

>gi|258645690|ref|ZP_05733159.1| CRISPR-associated protein, Csy4 family [Dialister invisus DSM 15470]
 MEYYQEIITLLPCAENVSLAFLTKVFTQLHIAFADEKNKSGHNLVAVSFPEYRETGLGEKIRVFAEAQEELE
 RLNLKVLGRLLDVHCTSIKRVPERKLRGYAVYSRYQEGSIWVKARRYAKRHPGVTFEEAARLLQGKR
 KSVRLPYIQMKSLSRGGTFSLFIKRVEKESALTECGTYGLSNRNVPEF (SEQ ID NO:92)

RNA recognition sequences
 GUUACUGCCGCAUAGGUAGUUUAGAAA (SEQ ID NO:93)
 GUUAUCUGCCGUUAGGCAGCUUAGAAA (SEQ ID NO:94)

FIG. 5AZ

>gi|165975671|ref|YP_001651264.1| hypothetical protein APJL_0216 [Actinobacillus pleuropneumoniae serovar 3 str. J103]
 MSELTHYIELKAIPQVDILQTDVIAHGLQILHKFLPLYQGEIGLSFPAYGLGRTLGGIIRVFGNEQHCTQ
 IKTLIGELQDYVLITSVTPVEEIVEYHRYQVRHKGOSAIRRTEQFLVQQGKWTTEEIRQEMLIHQON
 QKVFPYVKLKSGSTKQHFVLAIKRLAEPASGLFNAYGLSQAATVPHF (SEQ ID NO:95)

RNA recognition sequence
 CUUCACUGCCGUUAGGCAGCUUAGAAA (SEQ ID NO:96)

FIG. 5BA

>gi|190149486|ref|YP_001968011.1| hypothetical protein APP7_0217 [Actinobacillus pleuropneumoniae serovar 7 str. AP76]
 MSELTHYIELKAIPQVDILQTDVIAHGLQILHKFLPLYQGEIGLSFPAYGLGRTLGGLIRVFGNEQHCTQ
 IKTQLIGEGLDYVLITSVTPVEEIVYHRYQVRHKGQSARIRTEQFLVQQGKWTEEIRQEMLIHQON
 QKVFPVVKLKGSTKQHFVLAIRQLRLAEPVSGLFNAYGLSKIATVPHF (SEQ ID NO:97)

RNA recognition sequence
 CUUCACUGCCGUAGGAGCUUAGAAA (SEQ ID NO:96)

FIG. 5BB

>gi|303253029|ref|ZP_07339182.1| hypothetical protein APP2_1978 [Actinobacillus pleuropneumoniae serovar 2 str. 4226]
 MSELTHYIELKAIPQVDILQTDVIAHGLQILHKFLPLYQGEIGLSFPAYGLGRTLGGLIRVFGNEQHCTQ
 IKTQLIGEGLDYVLITSVTPVEEIVYHRYQVRHKGQSARIRTEQFLVQQGKWTEEIRQEMLIHQON
 QKVFPVVKLKGSTKQHFVLAIRQLRLAEPVSGLFNTYGLSKIATVPHF (SEQ ID NO:98)

FIG. 5BC

>gi|303251662|ref|ZP_07337835.1| hypothetical protein APP6_0866 [Actinobacillus pleuropneumoniae serovar 6 str. Femo]
 MSELTHYIELKAIPQVDILQTDVIAHGLQILHKFLPLYQGEIGLSFPAYGLGRTLGGLIRVFGNEQHCTQ
 IKTQLIGEGLDYVLITSVTPVEEIVYHRYQVRHKGQSARIRTEQFLVQQGKWTEEIRQEMLIHQON
 QKVFPVVKLKGSTKQHFVLAIRQLRLAEPVSGLFNTYGLSKIATVPHF (SEQ ID NO:99)

FIG. 5BD

>gi|116627507|ref|YP_820126.1| dephospho-CoA kinase [Streptococcus thermophilus LMD-9]
 MSKTMIGLTGGASGKSTVVEIKDAGYKVIDADQLVHDMQVKGRLYQALLDWLGDGILLPNEGELNRP
 KLGLIFSEEMRYQSAEIQGKIIFREELAAKRDCLAKEEDVFFMDIPLLFENDYQDFDQIWLVAVSPQV
 QGQRLMKRNHLSAEFAGMRIAQQMPLAEKLPYASLVIDNNGNIDDLKKVKGAIKDLANIV (SEQ ID NO:100)

24 / 29

FIG. 6

His29Ala

1 mdhyldirlr pdpefppaql msvlfgk**la**q alvaqggdri gvsfpdl**des** rslgerlri
61 hasaddlral larpwleglr dhlqfgepav vphptpyrqv srqvksnpe rlrmlmrh
121 dlseeearkr ipdtvarald lpfvtlrsqs tgqhfrl**fir** hgplqataee ggftcyglsk
181 ggfvpwf (SEQ ID NO:101)

His29Ala/Ser50Cys

1 mdhyldirlr pdpefppaql msvlfgk**la**q alvaqggdri gvsfpdl**dec** rslgerlri
61 hasaddlral larpwleglr dhlqfgepav vphptpyrqv srqvksnpe rlrmlmrh
121 dlseeearkr ipdtvarald lpfvtlrsqs tgqhfrl**fir** hgplqataee ggftcyglsk
181 ggfvpwf (SEQ ID NO:102)

25 / 29

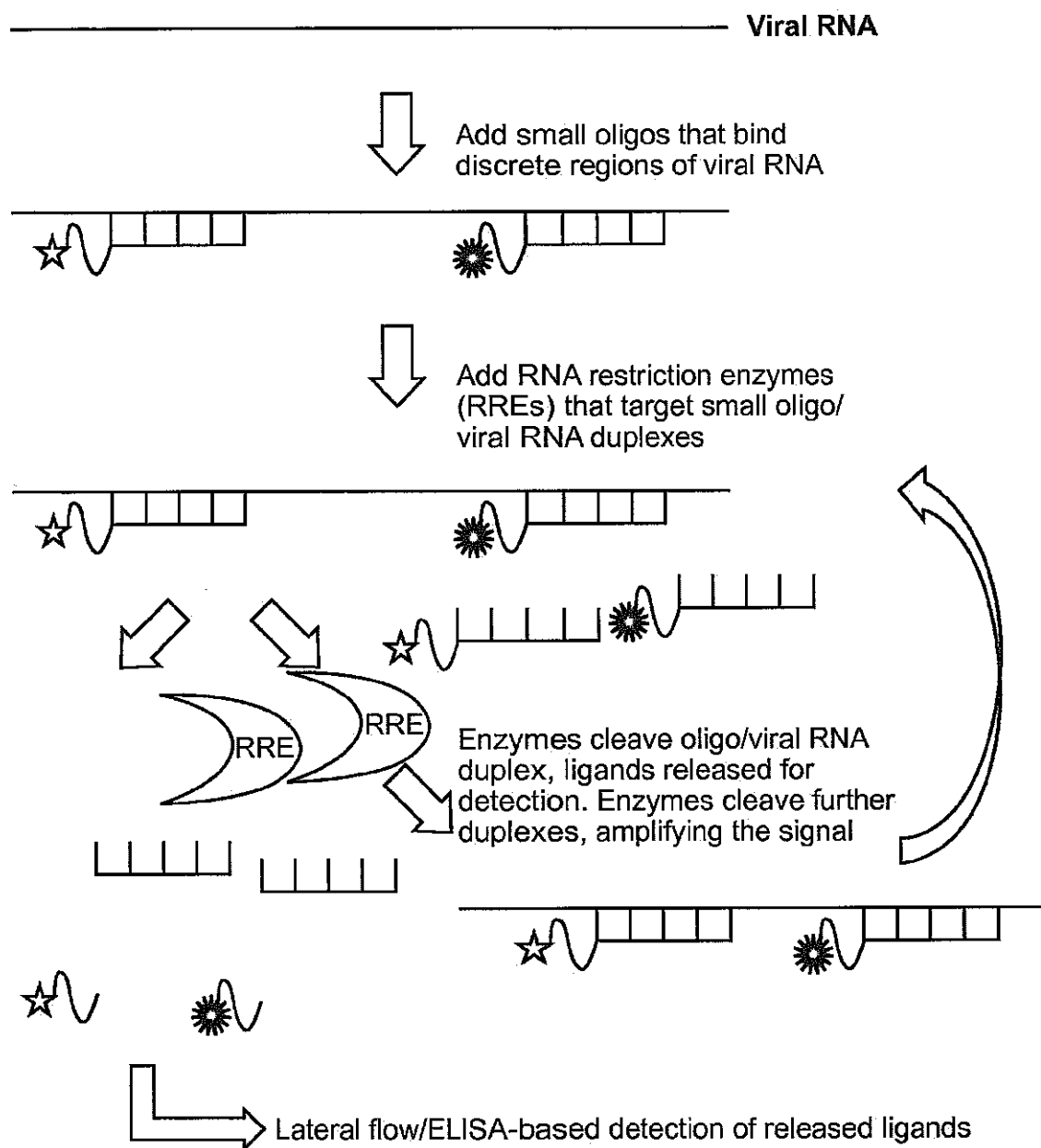


FIG. 7

26 / 29

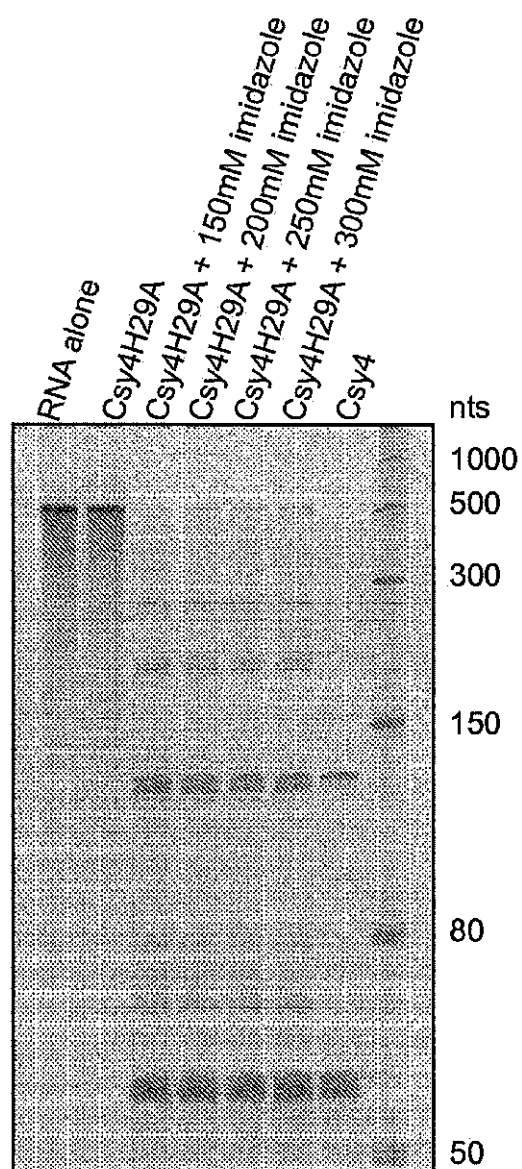


FIG. 8

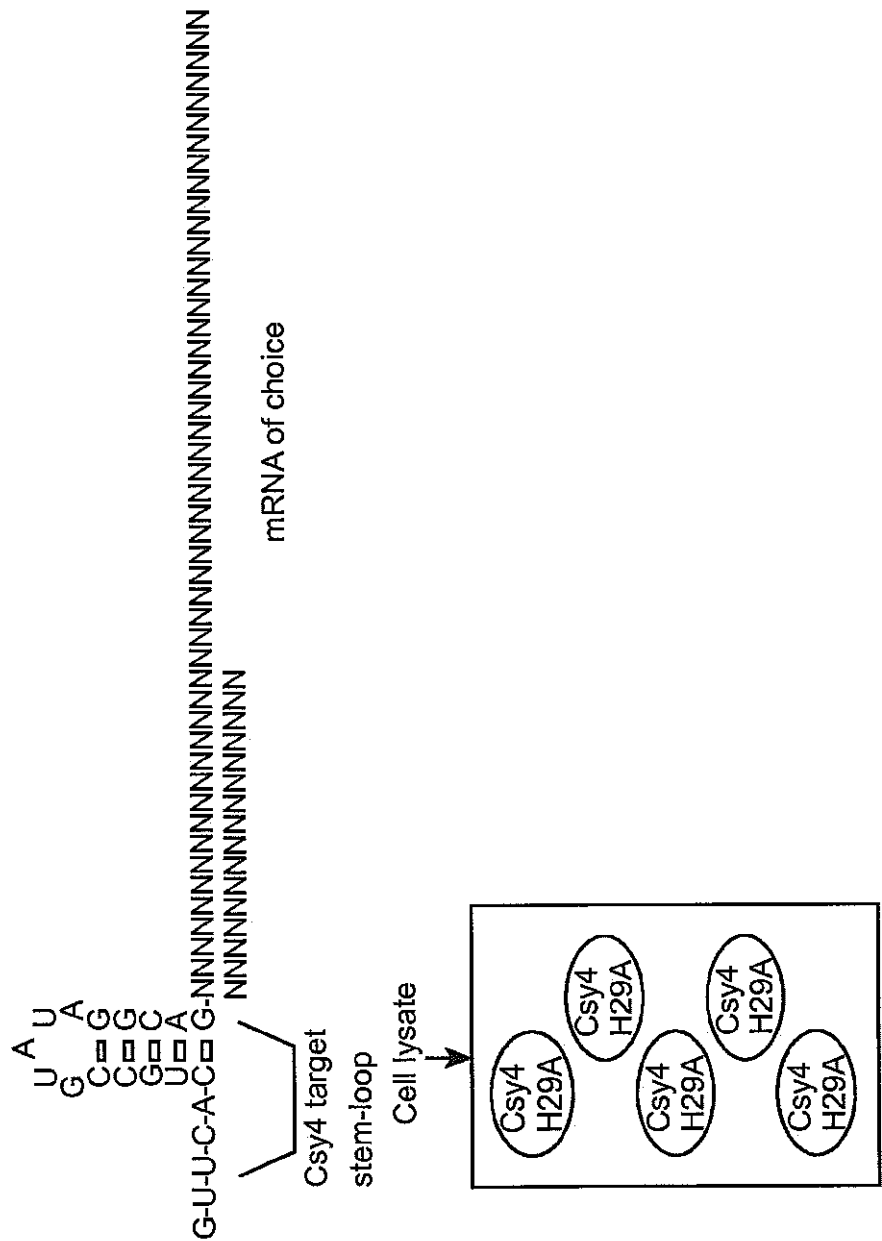
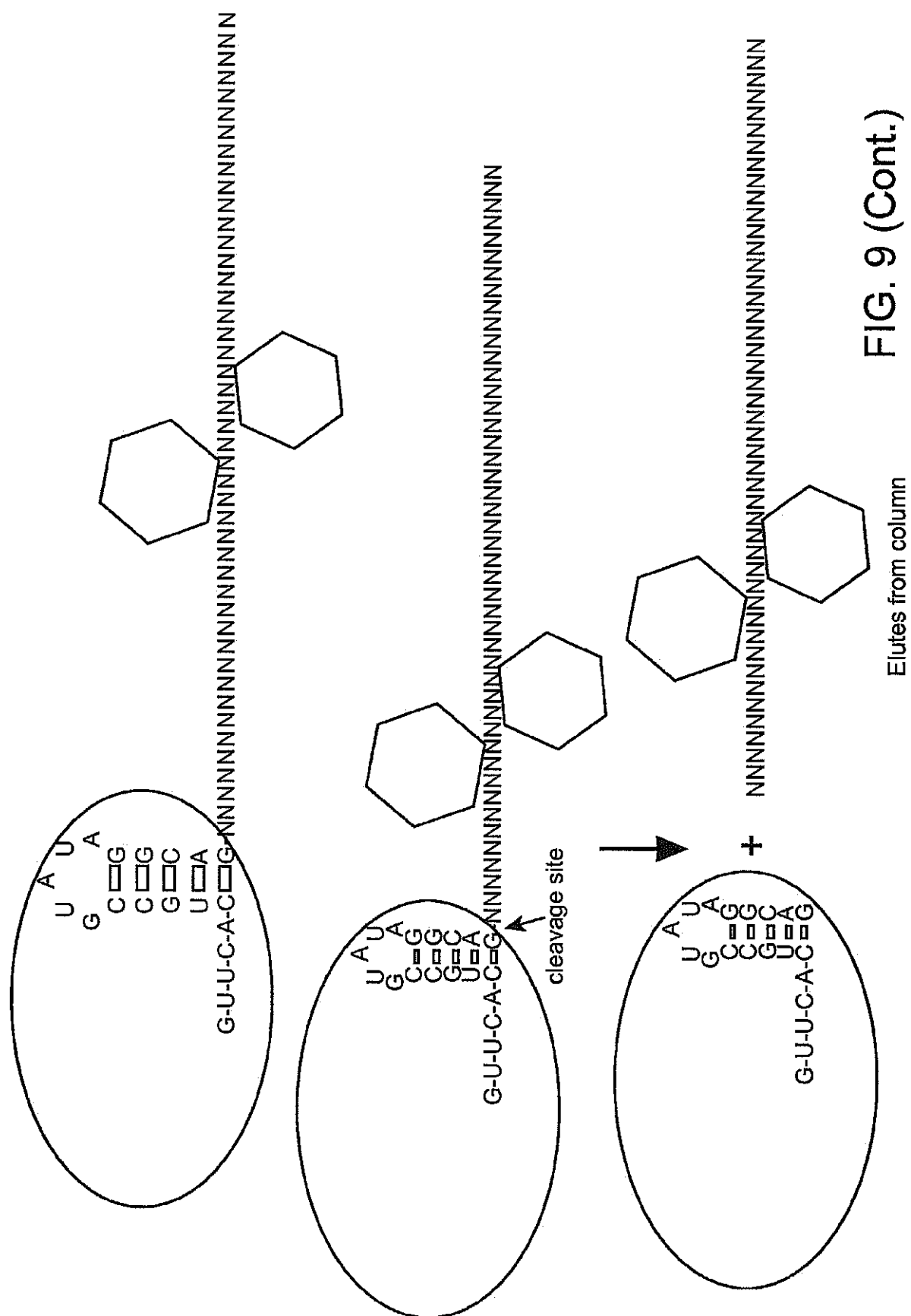


FIG. 9



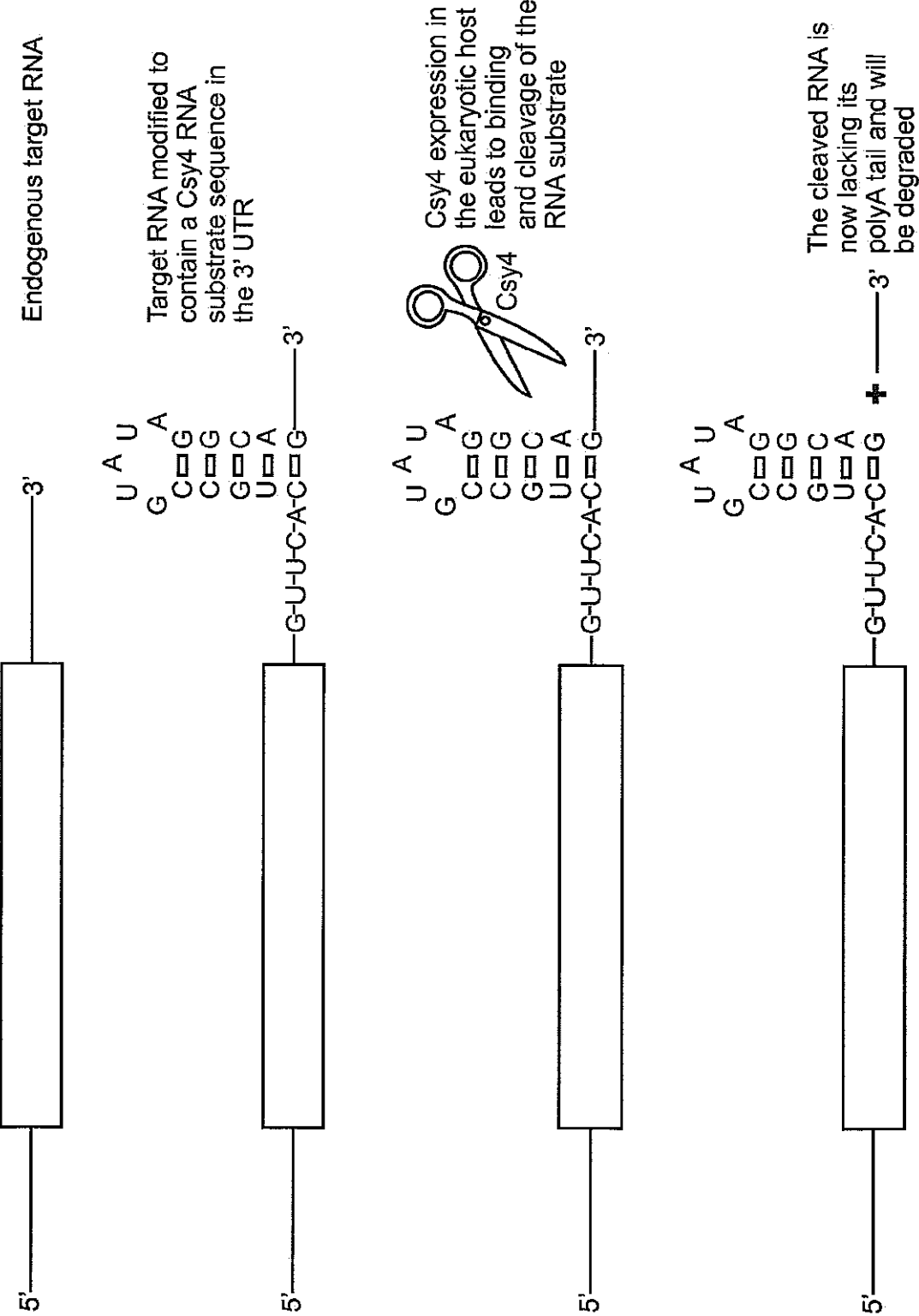


FIG. 10