

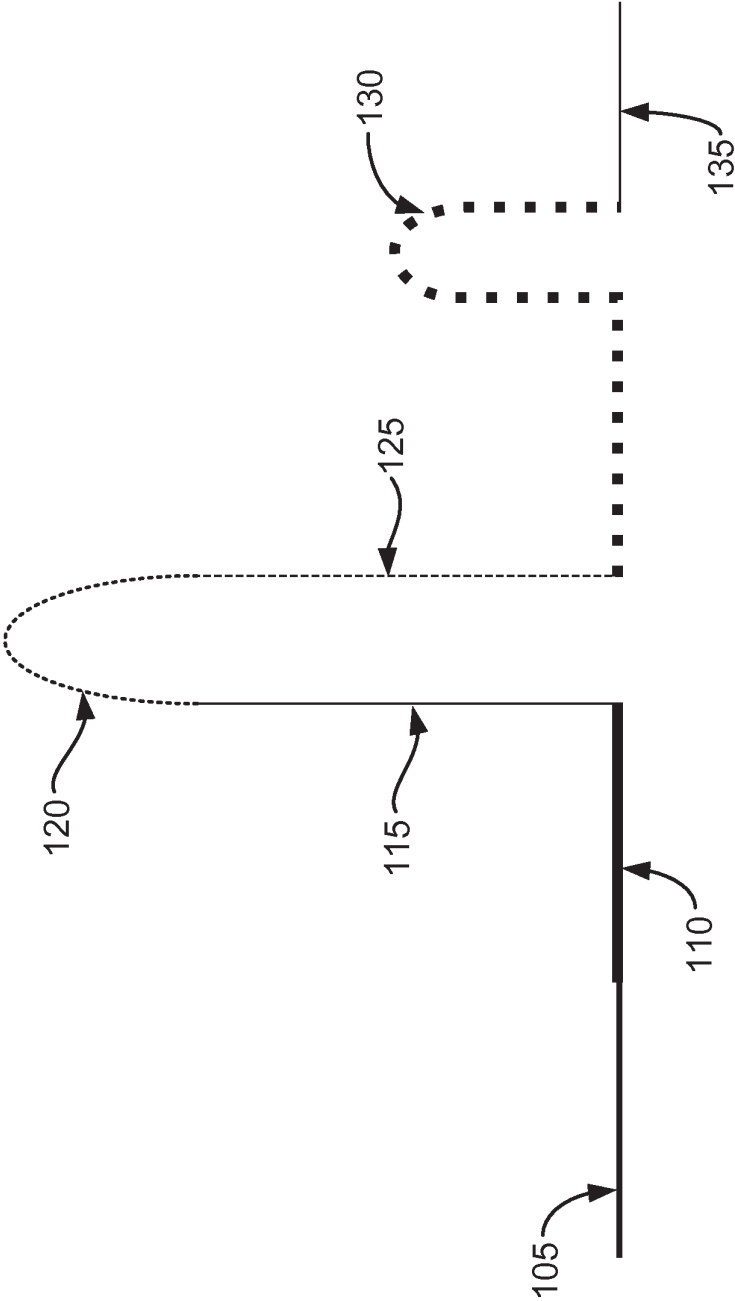


FIG. 1A

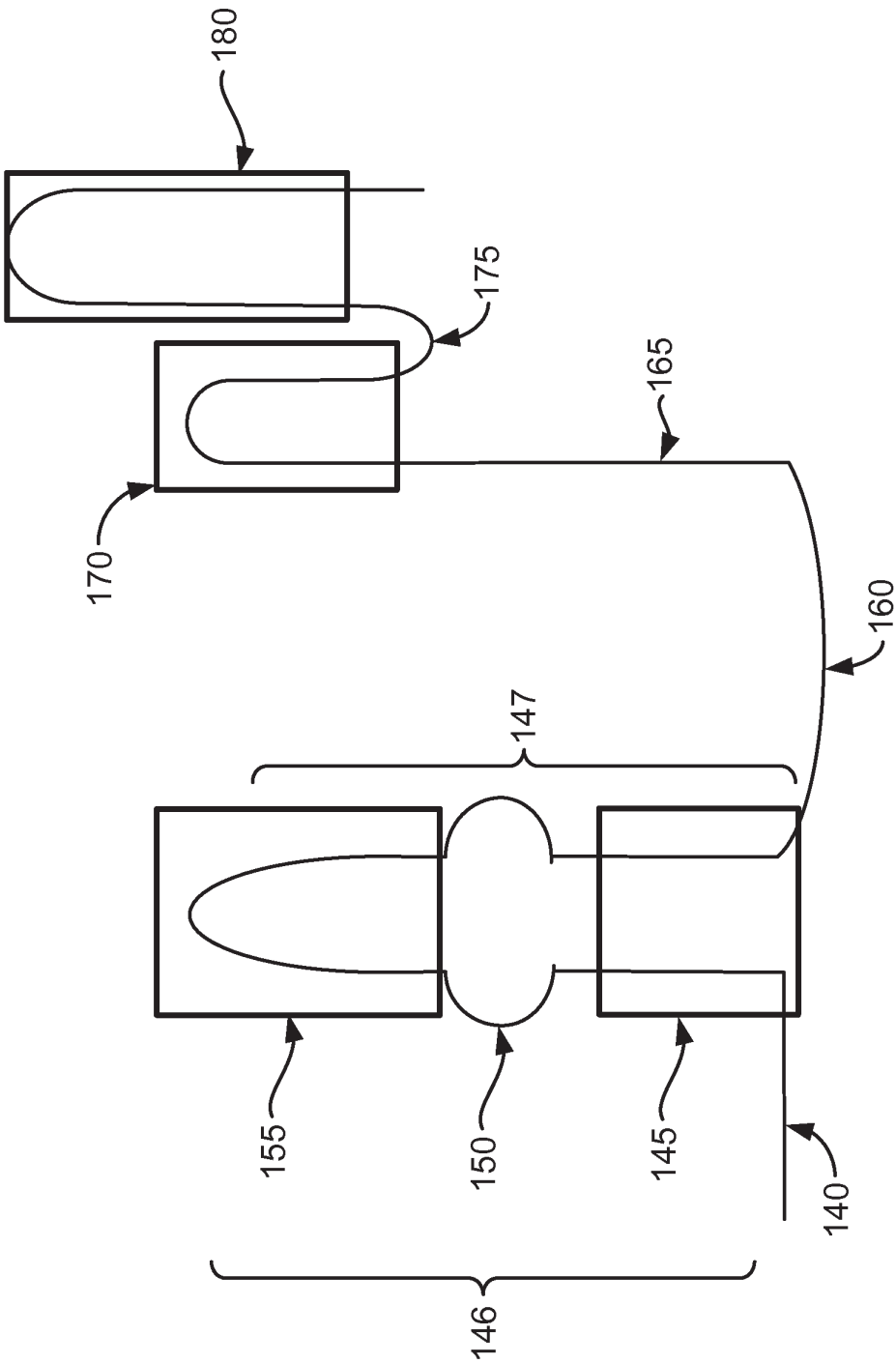


FIG. 1B

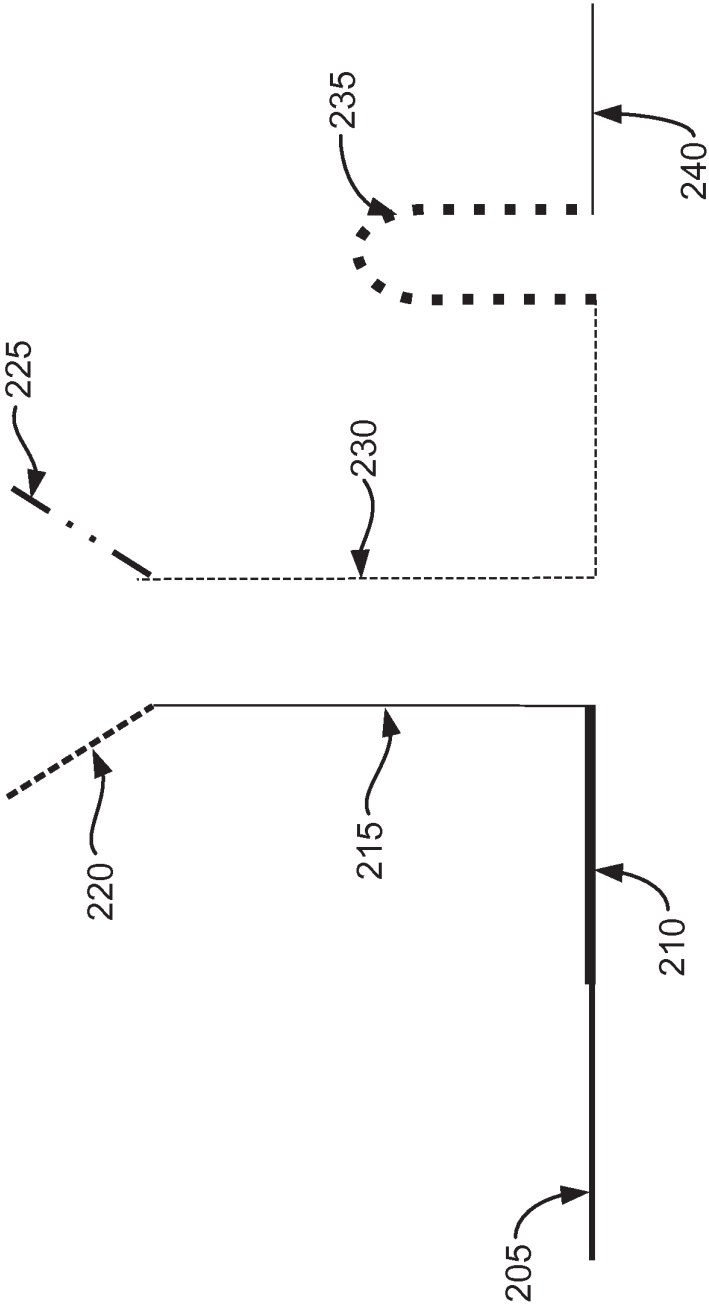


FIG. 2

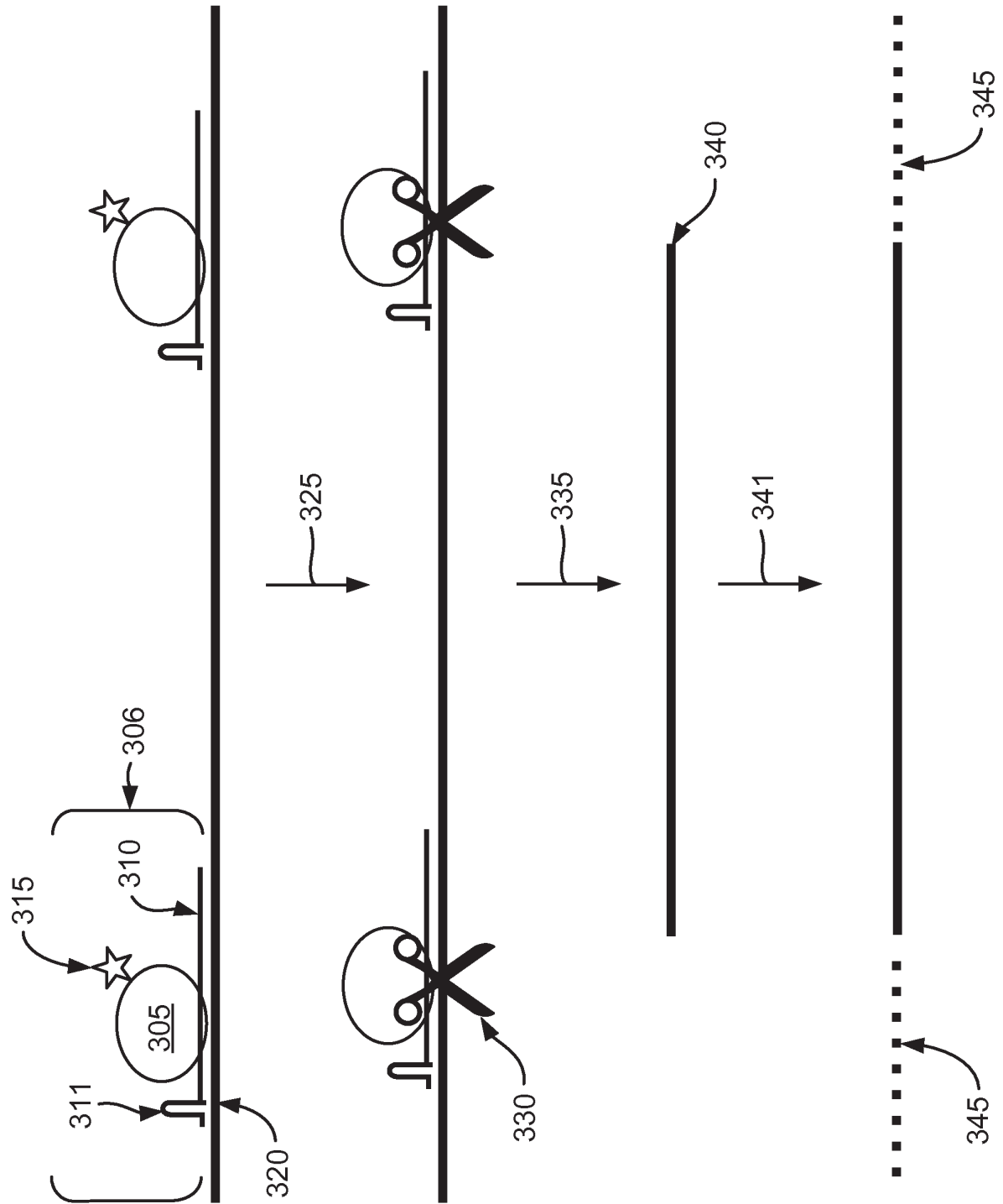


FIG. 3

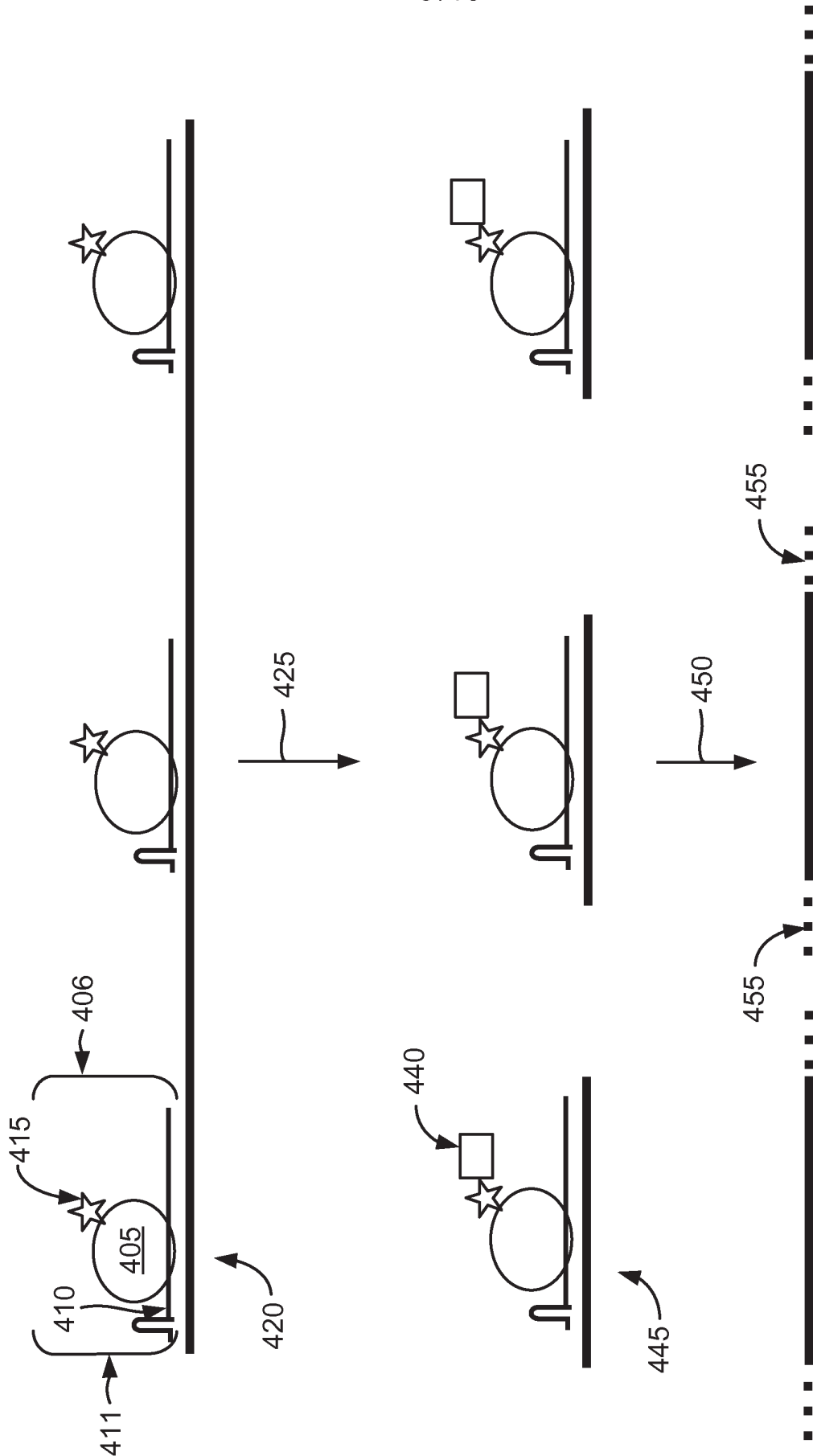


FIG. 4

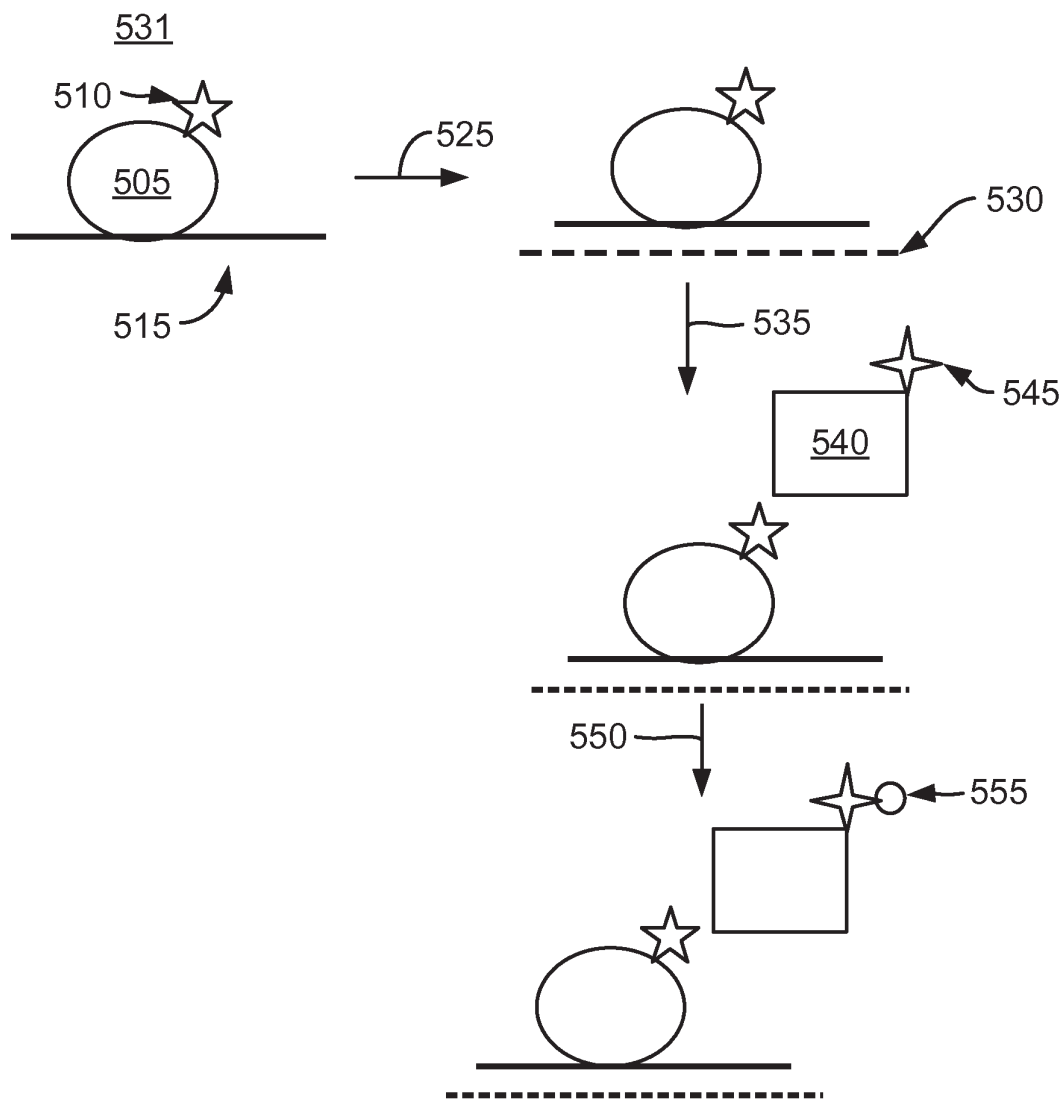


FIG. 5

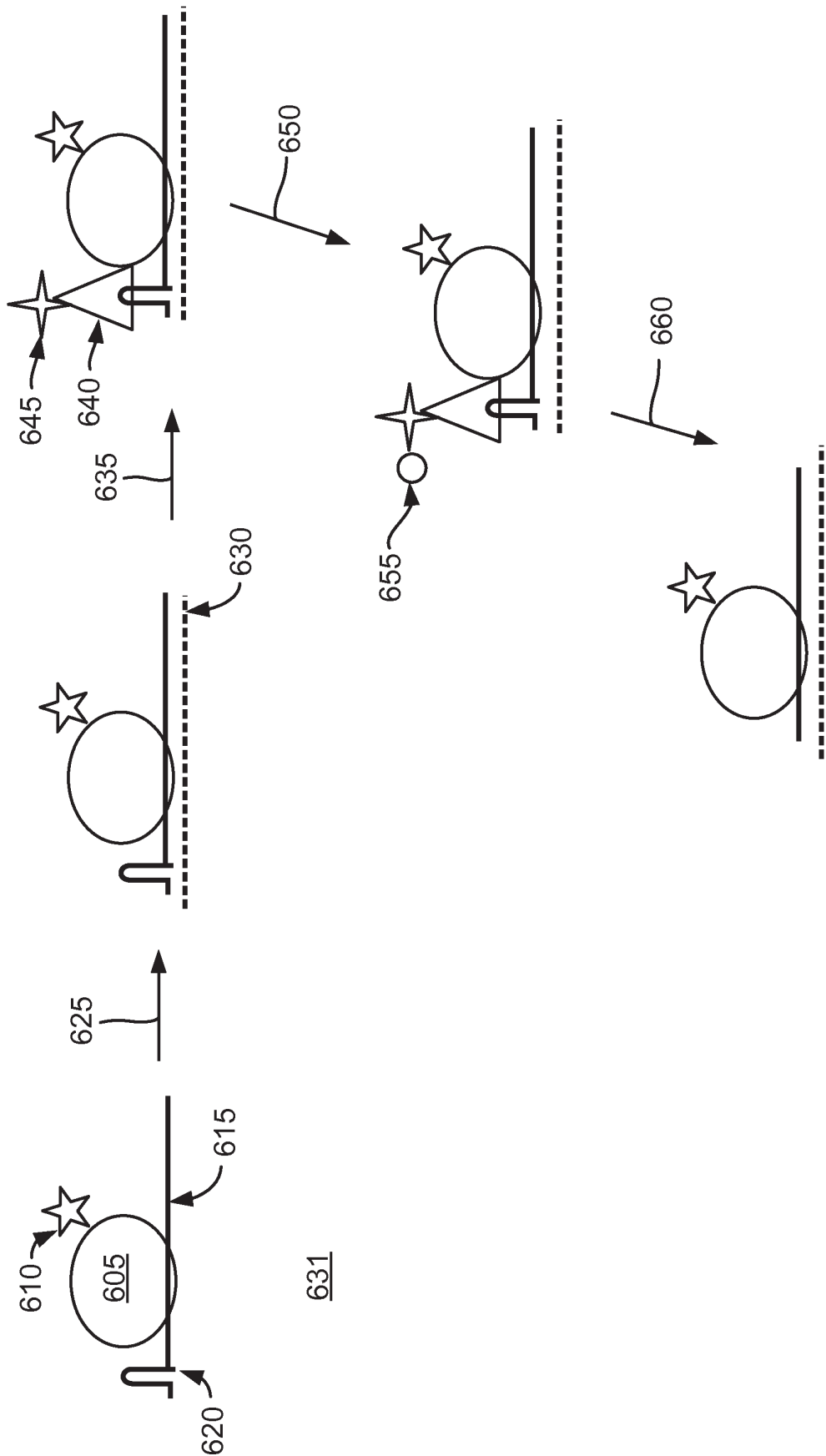


FIG. 6

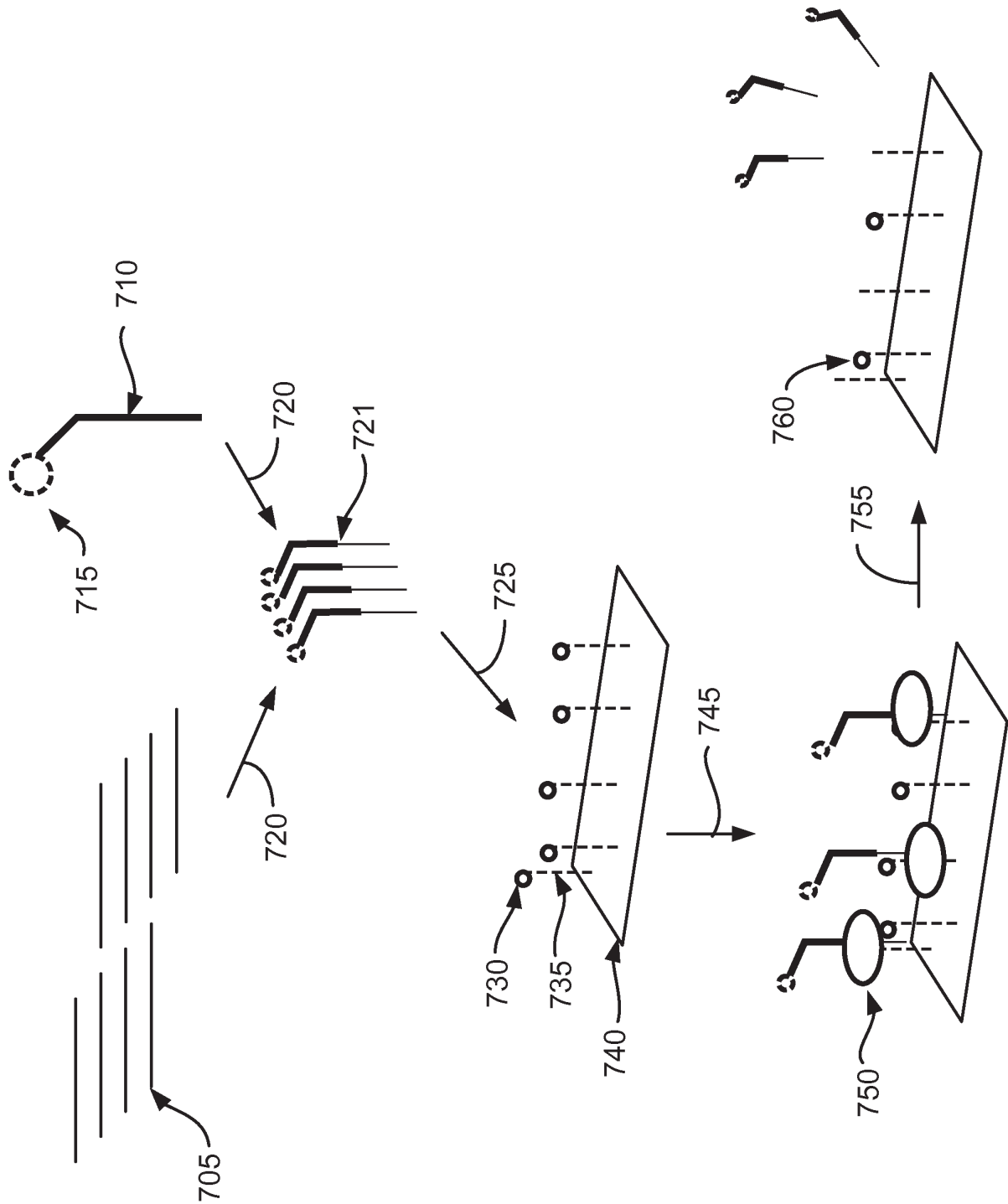


FIG. 7

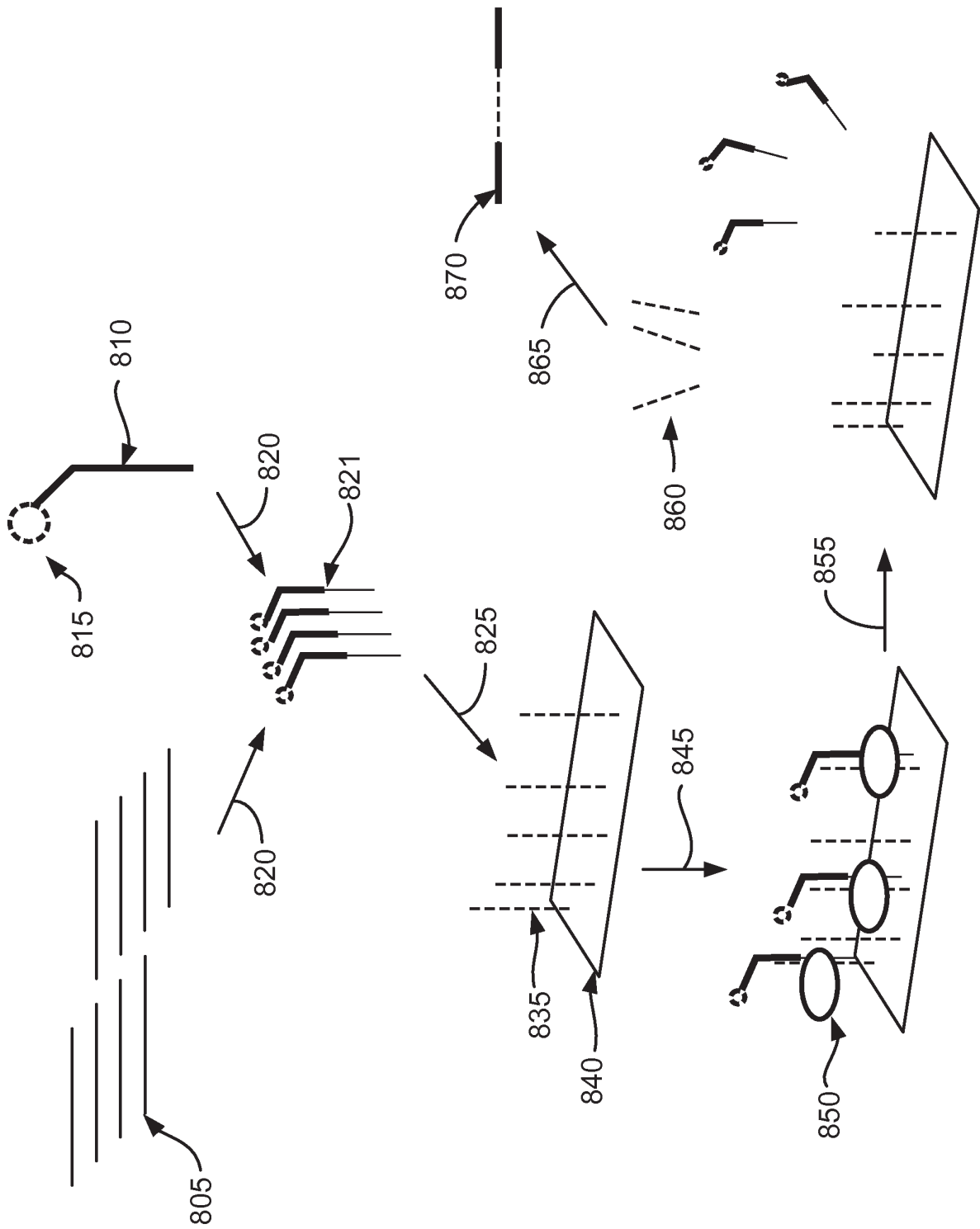


FIG. 8

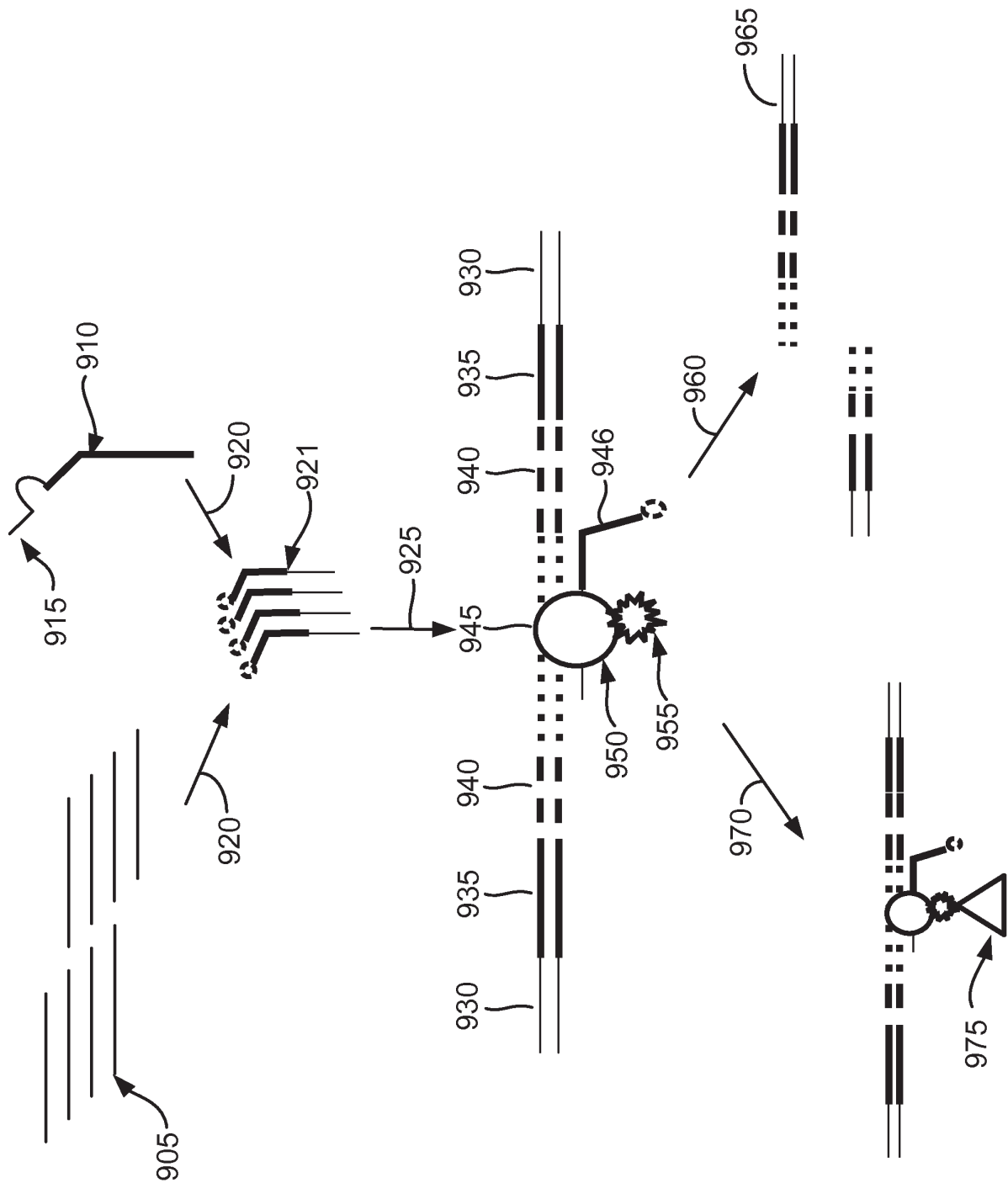


FIG. 9

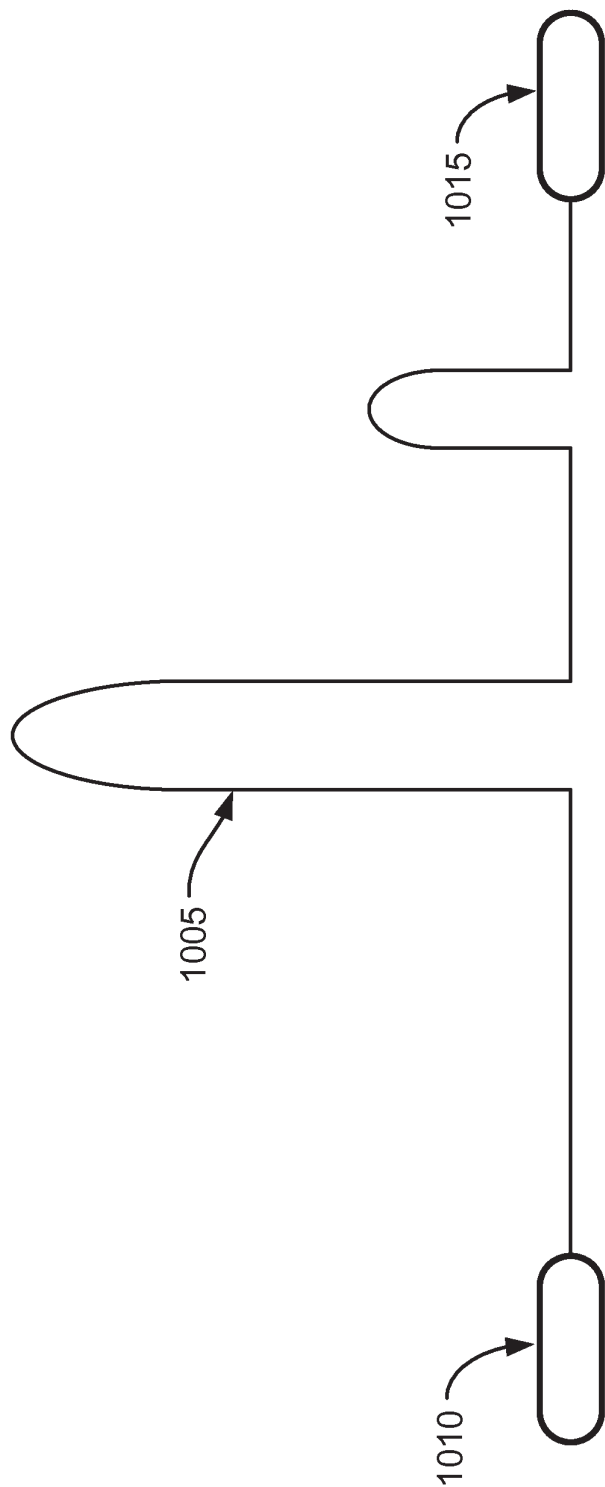


FIG. 10

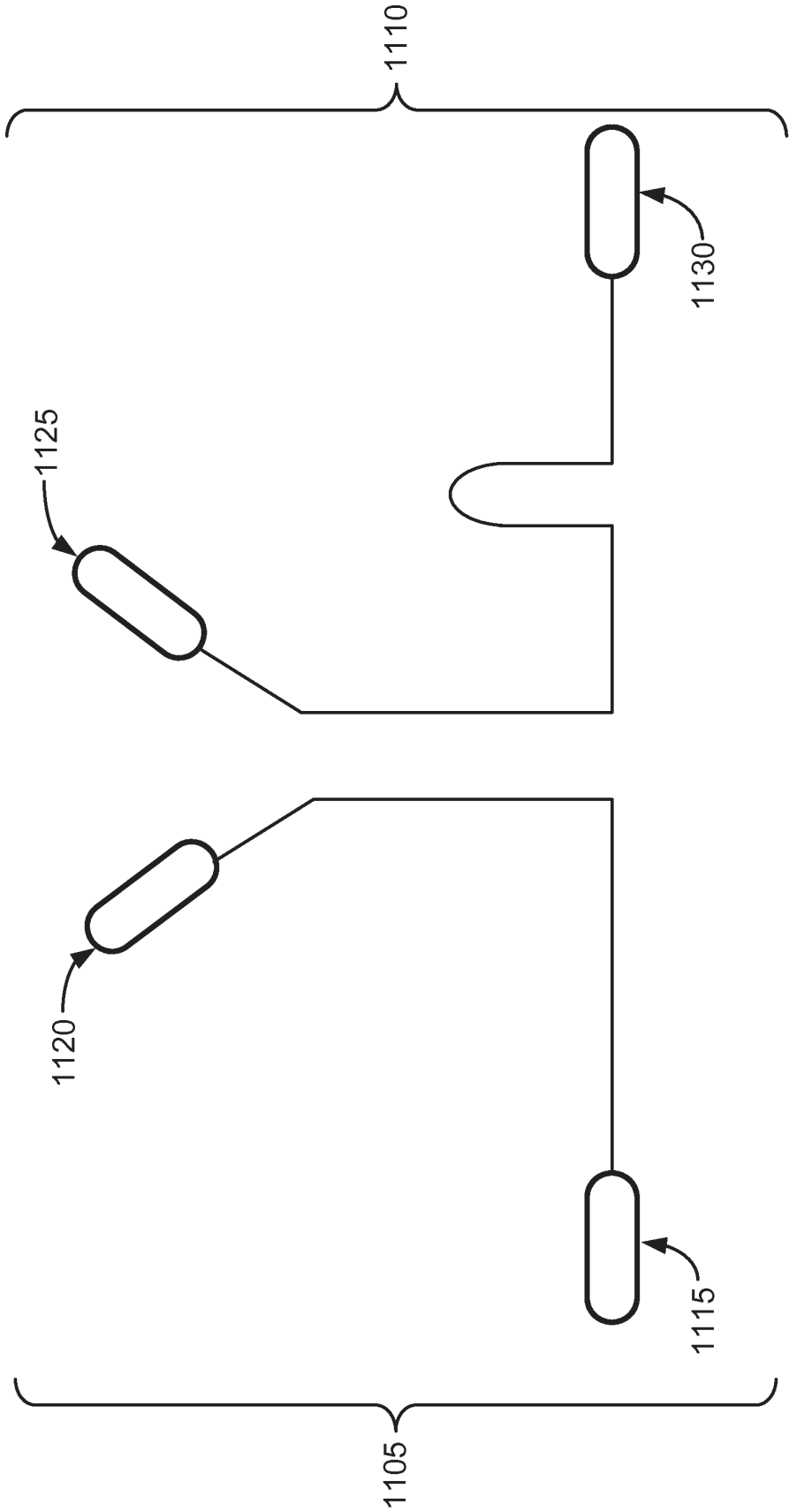


FIG. 11

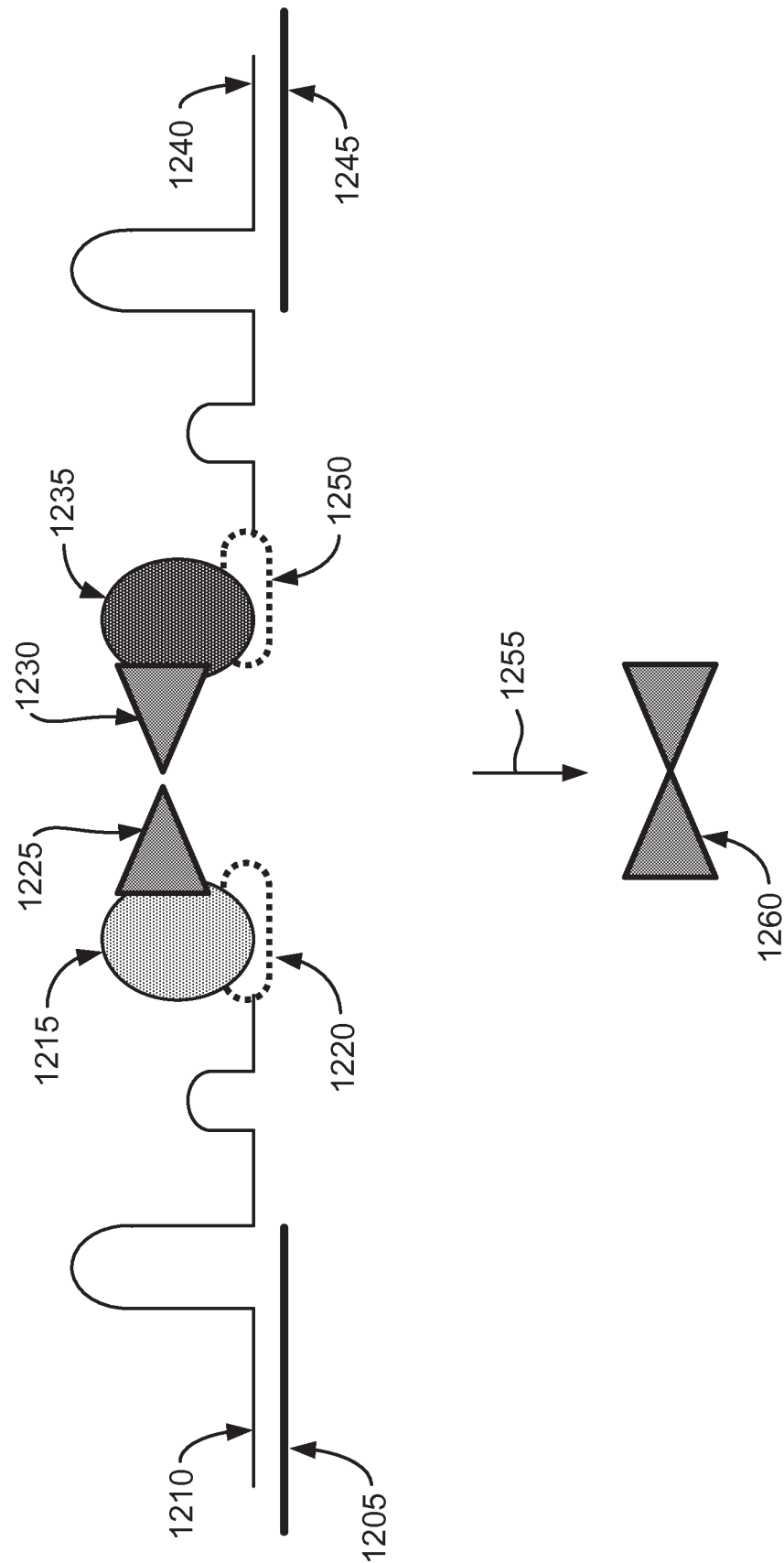
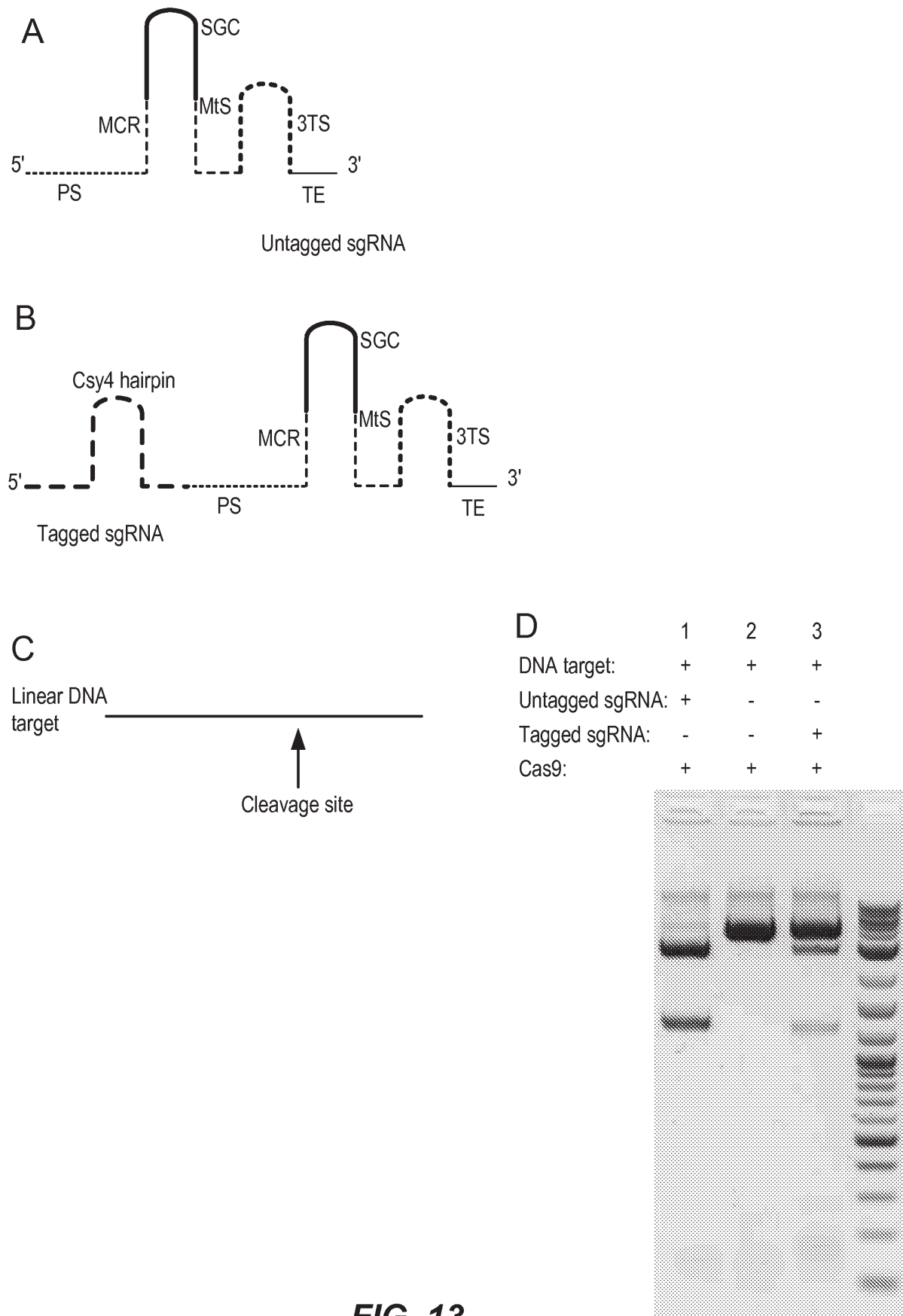


FIG. 12



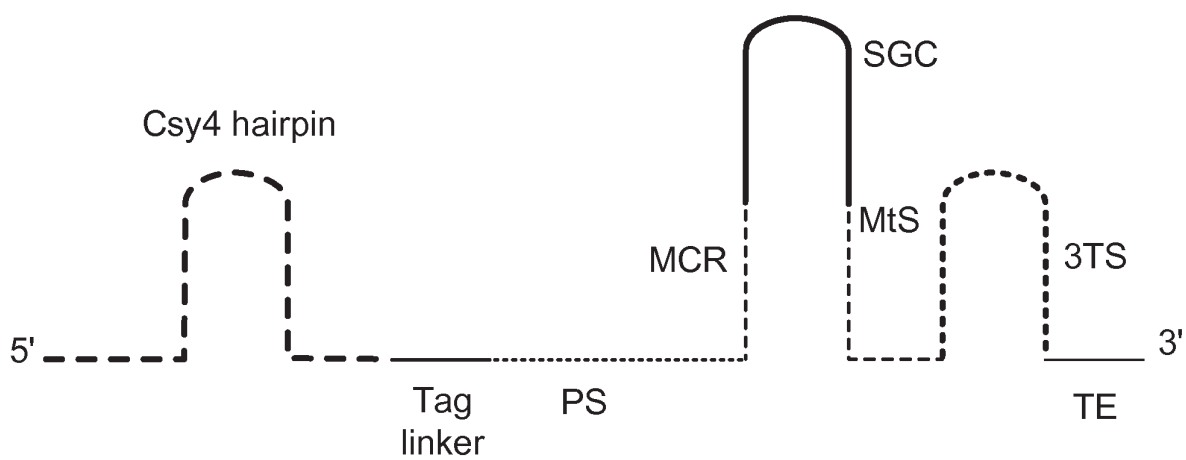


FIG. 14

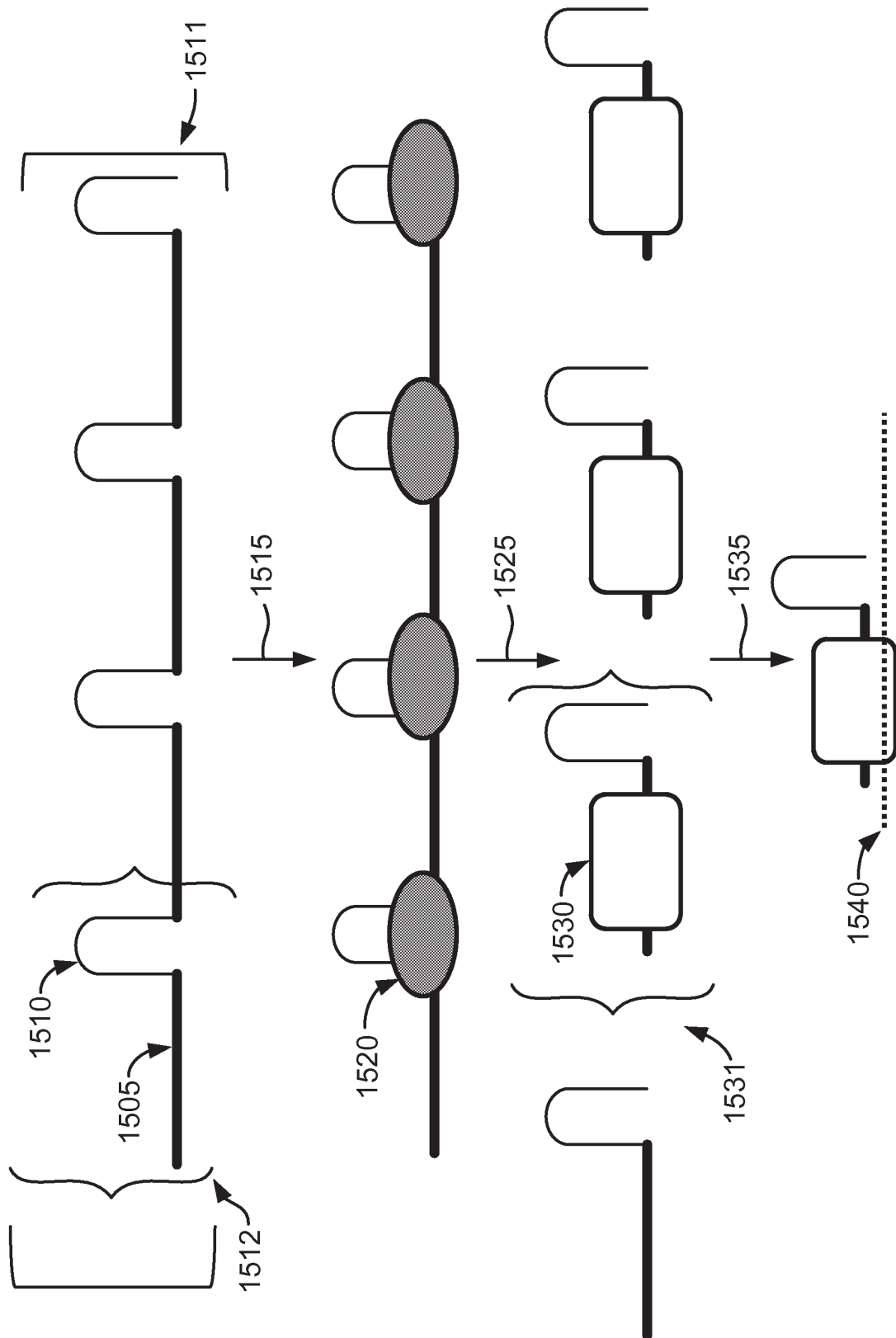


FIG. 15

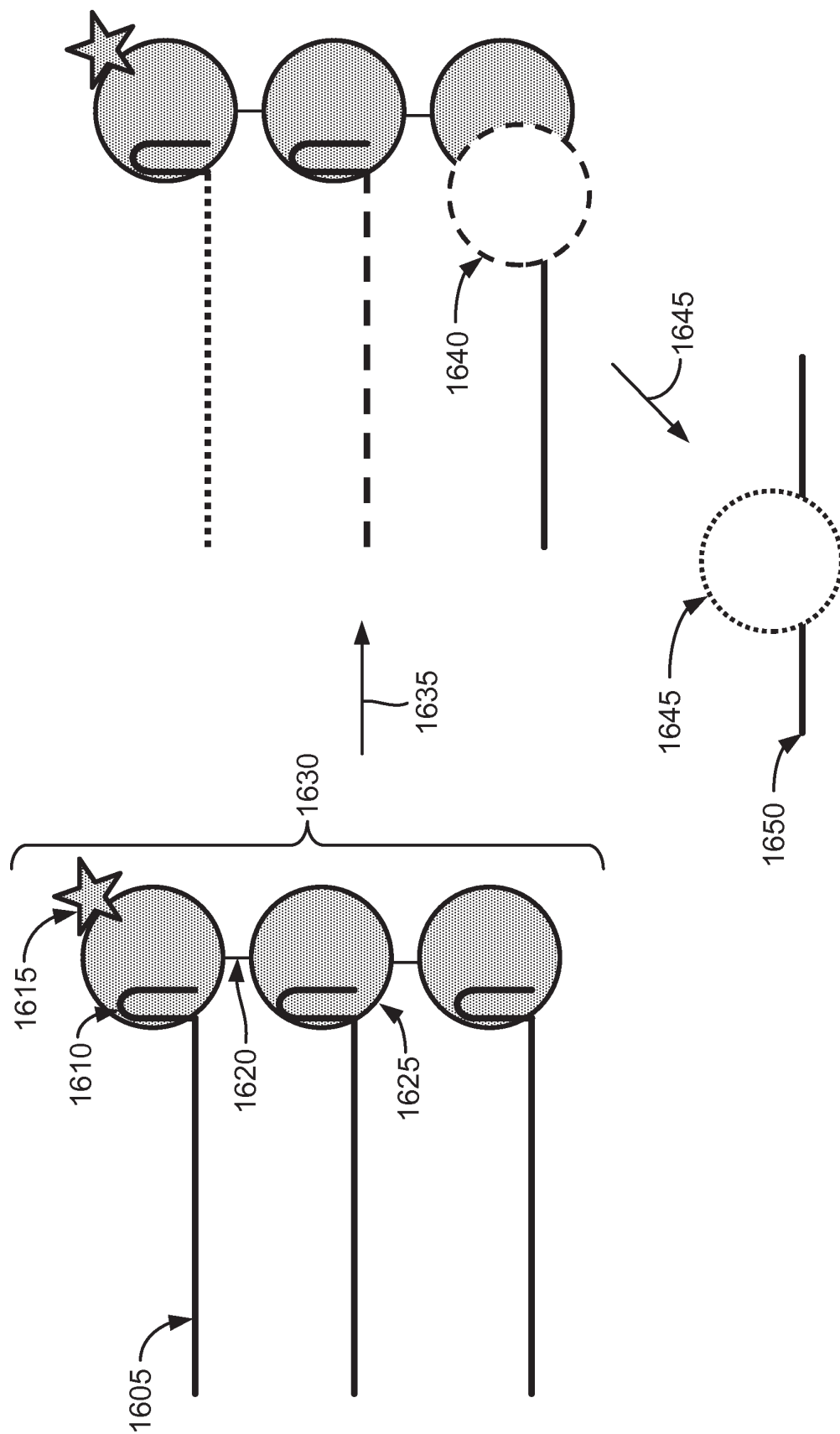
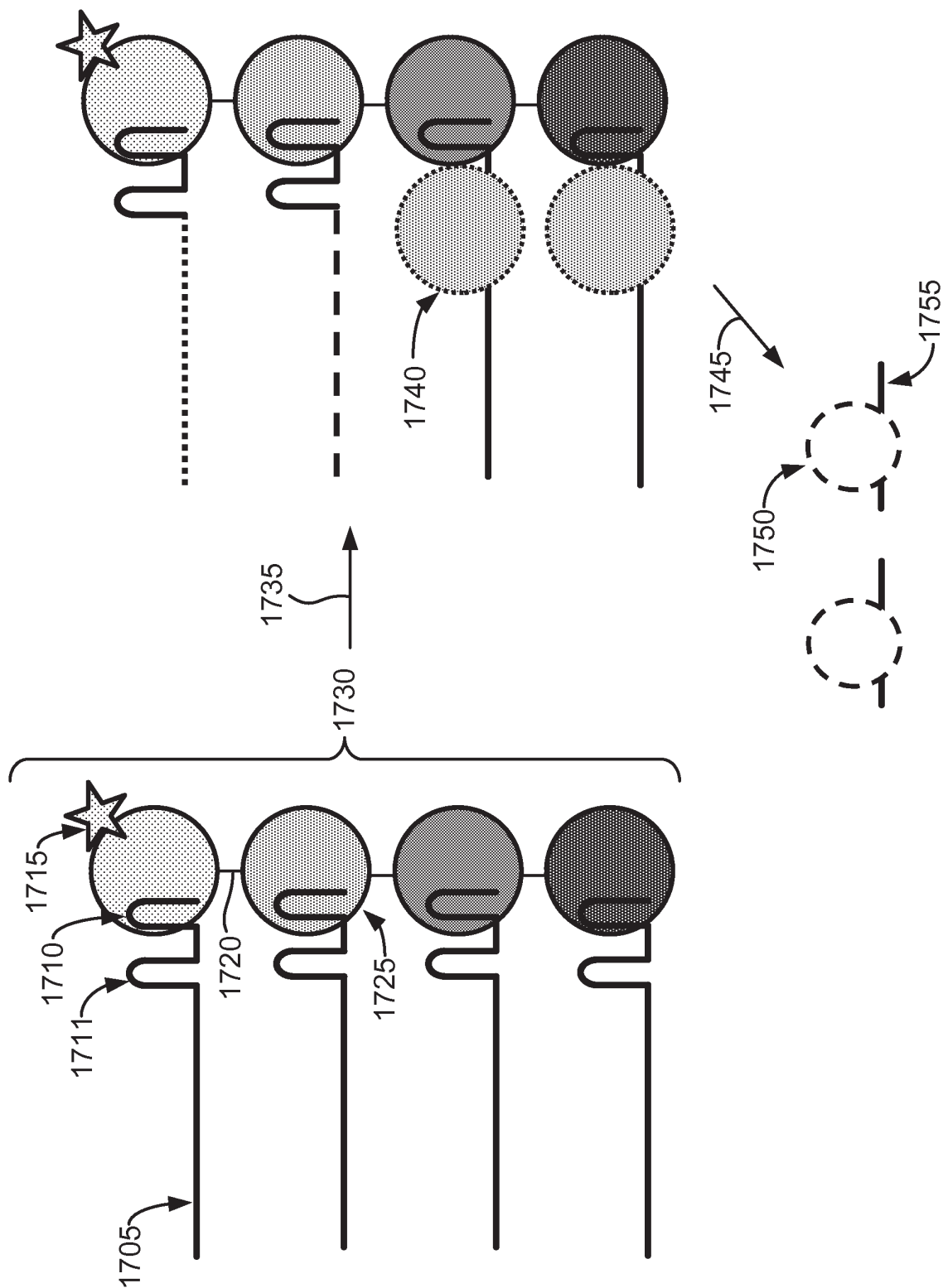


FIG. 16



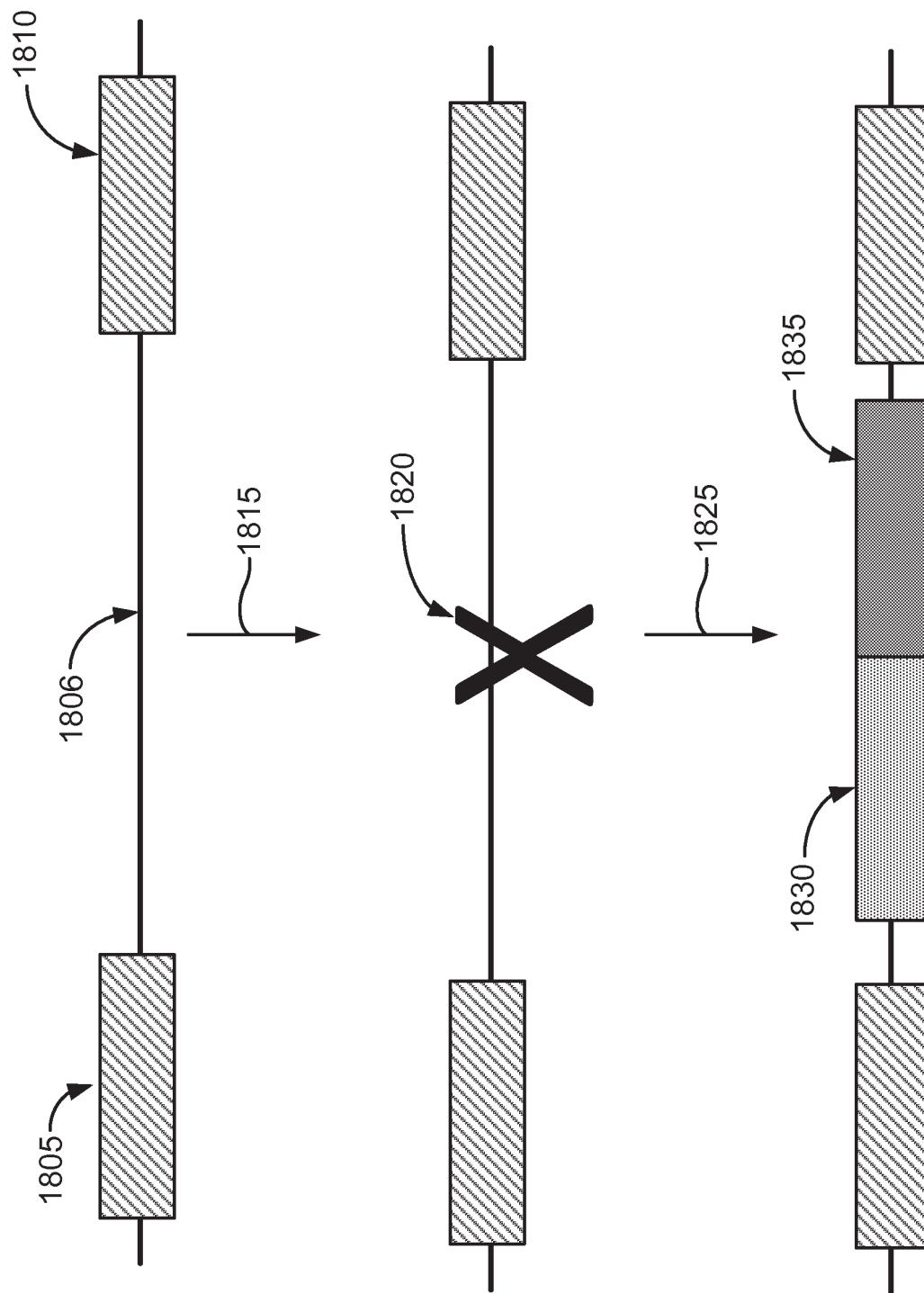


FIG. 18

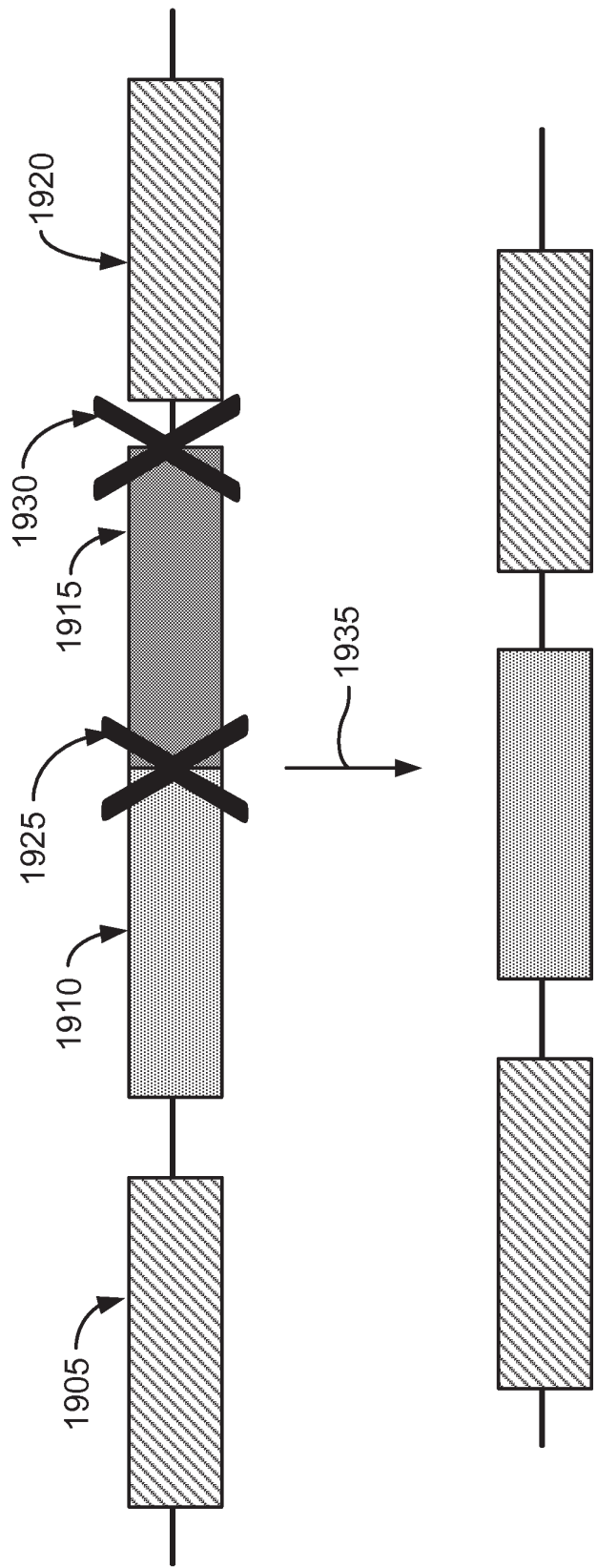


FIG. 19

5' -GUUUUAG-AGCUAUGCUGUUUUUGAAUGGUCGCCAAAC-3' - crRNA repeat sequence
 ||||| • ||||||| ||||||| ||||| |
3' -UAAAAUUGAACGAUACGACAAAACUUACCAAGGUUUU-5' - tracrRNA sequence

FIG. 20

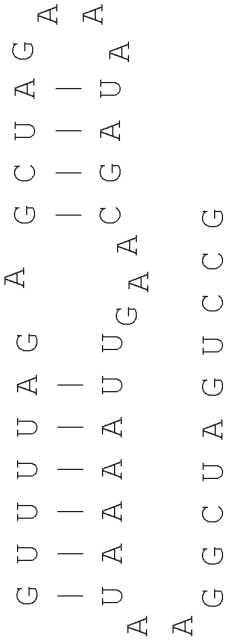
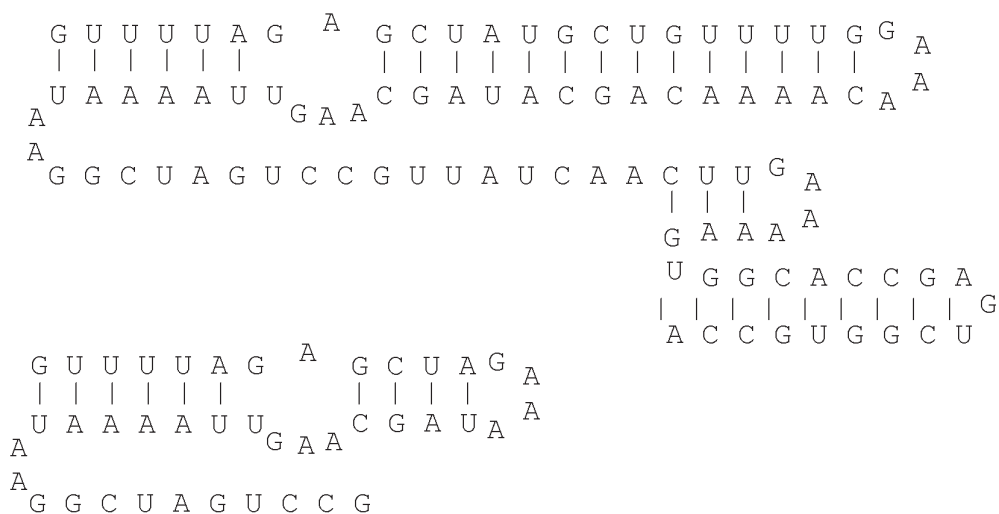


FIG. 21

22 / 73

FL- TRACR-CRRNA

Sequence



Duplex Variants Group 1

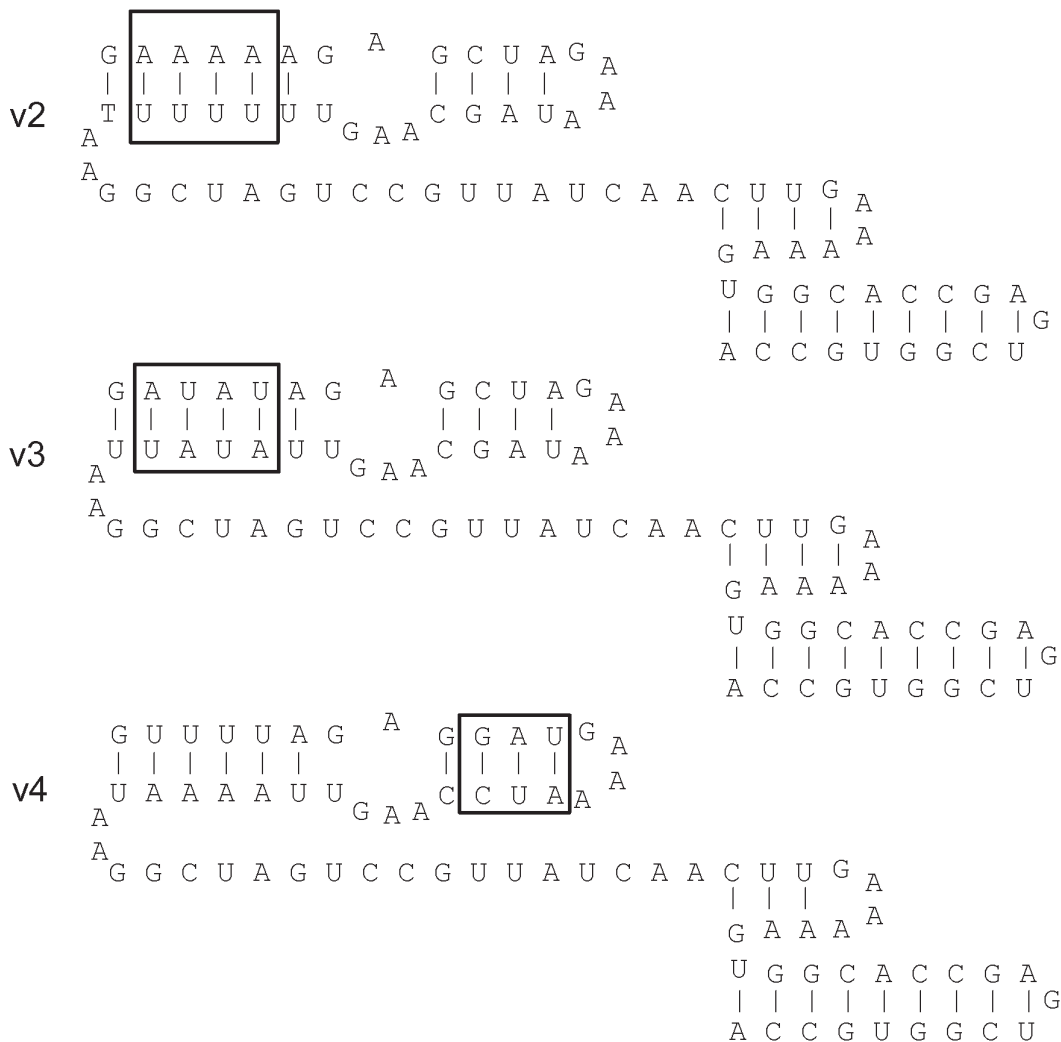


FIG. 22A

23 / 73

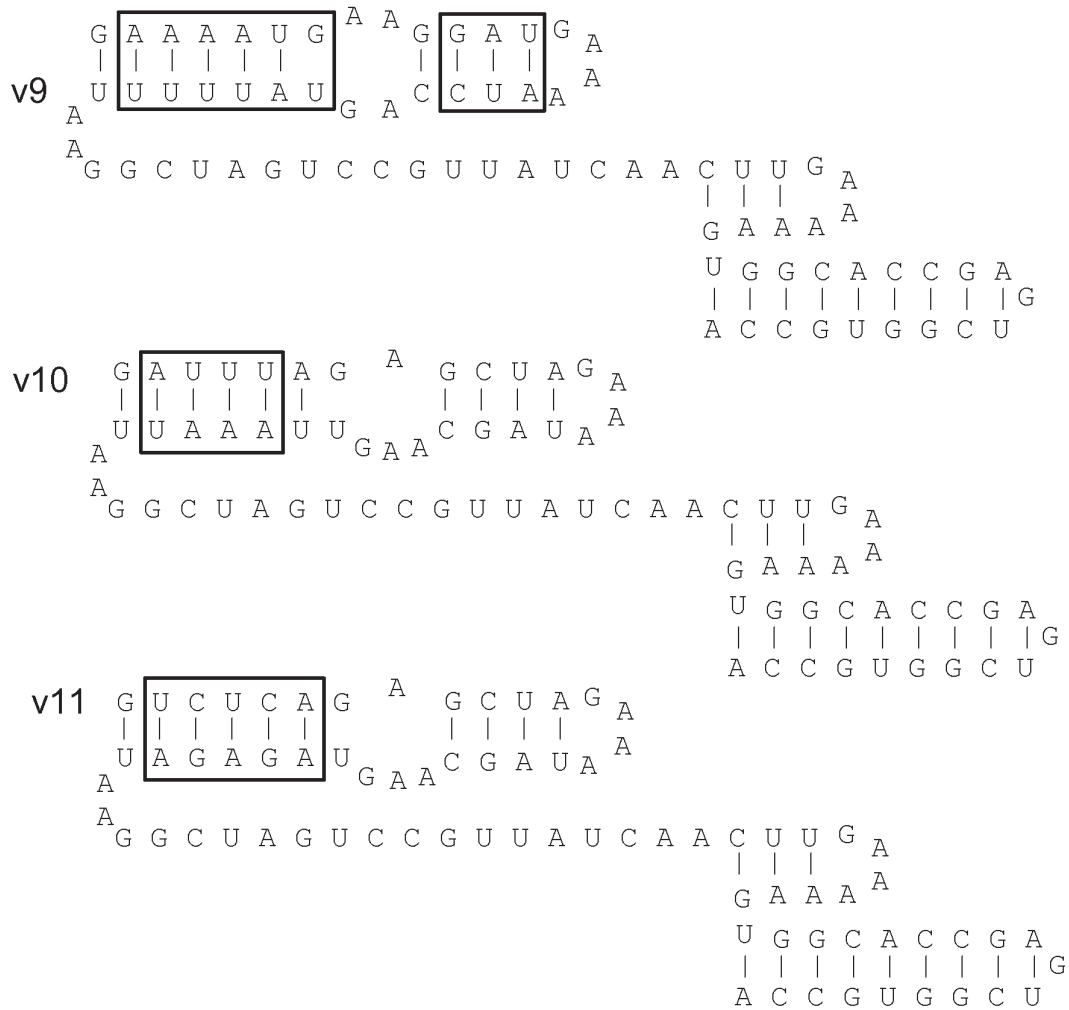
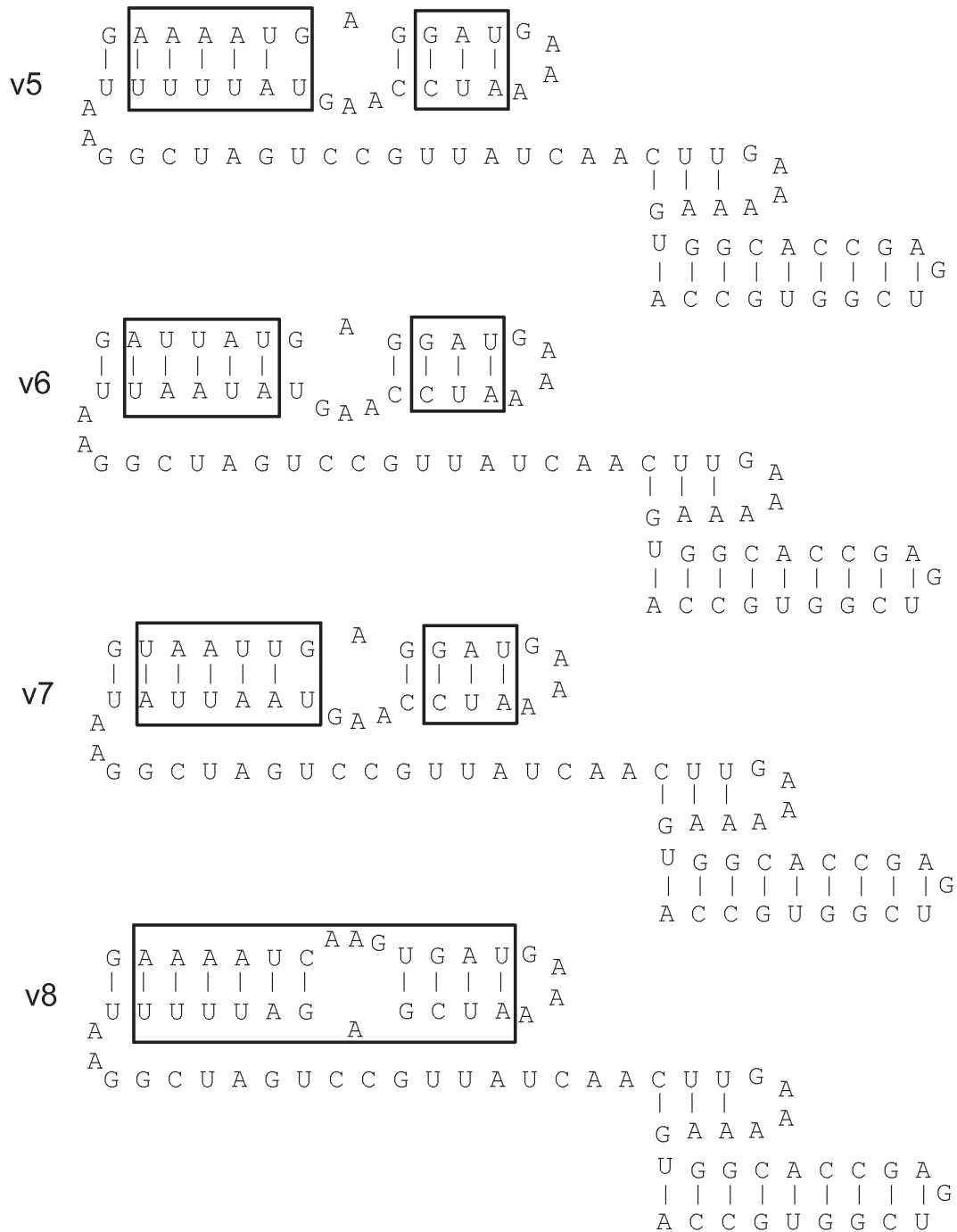


FIG. 22A
(Continued)

24 / 73

**FIG. 22B**

25 / 73

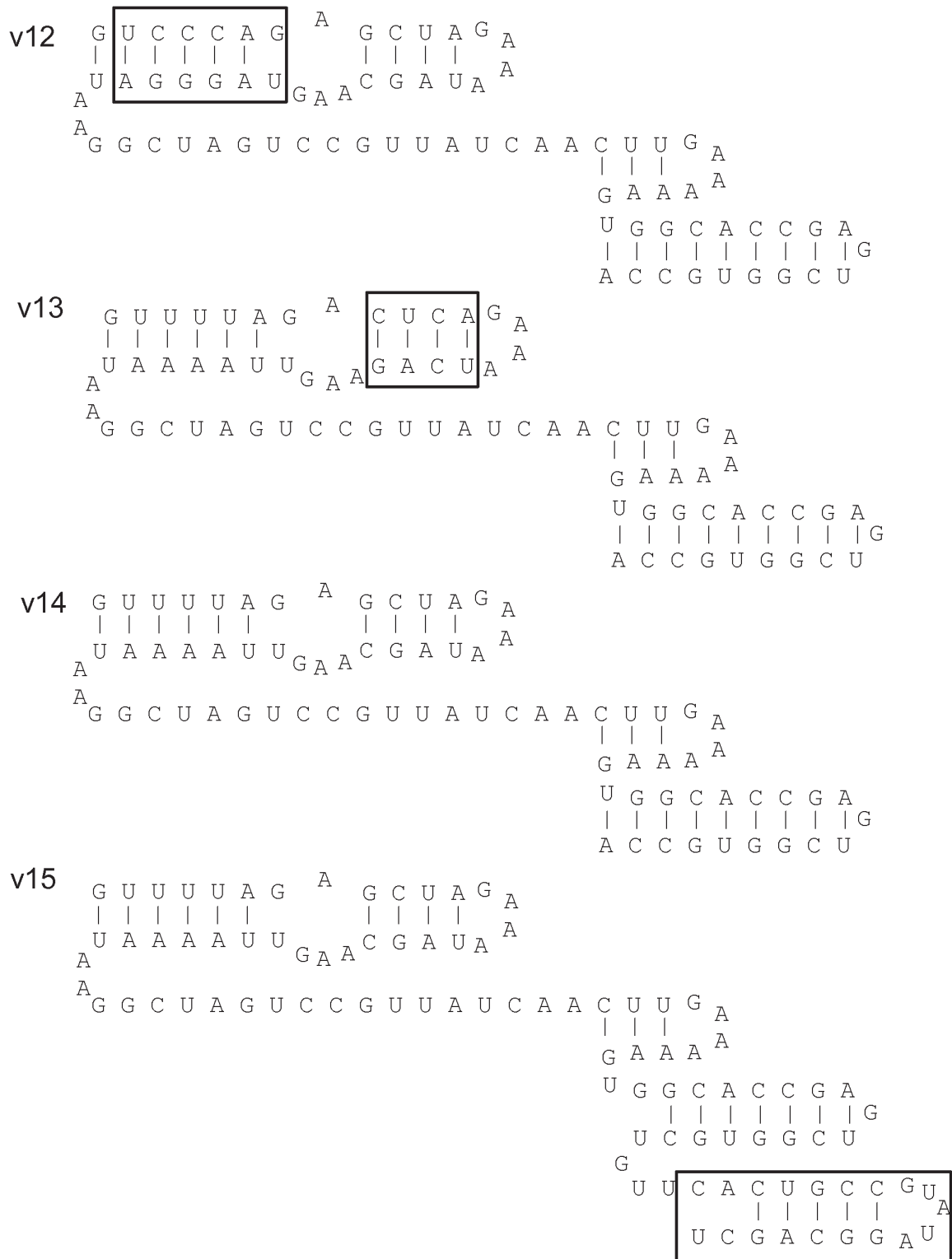


FIG. 22B
(Continued)

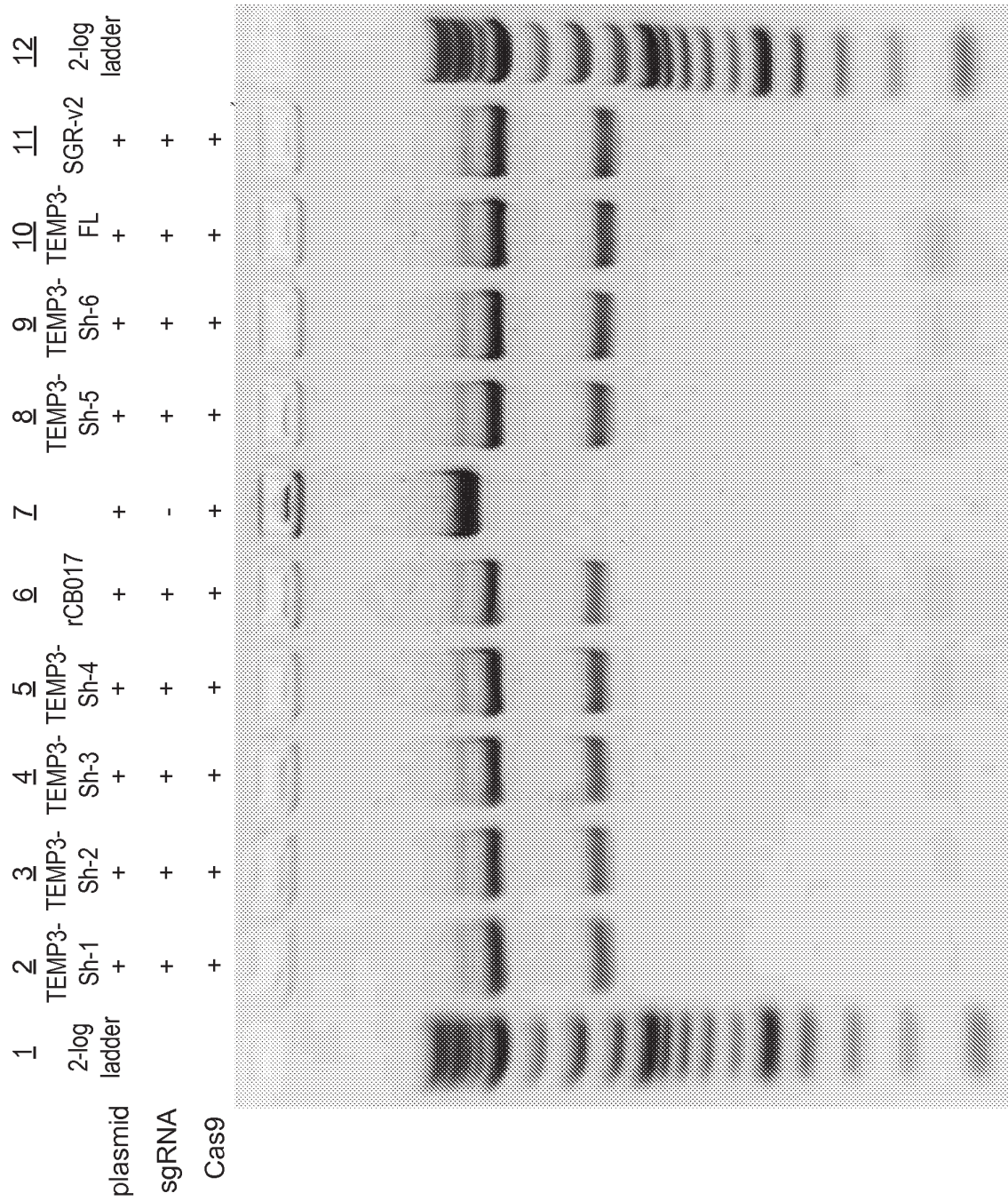


FIG. 23A

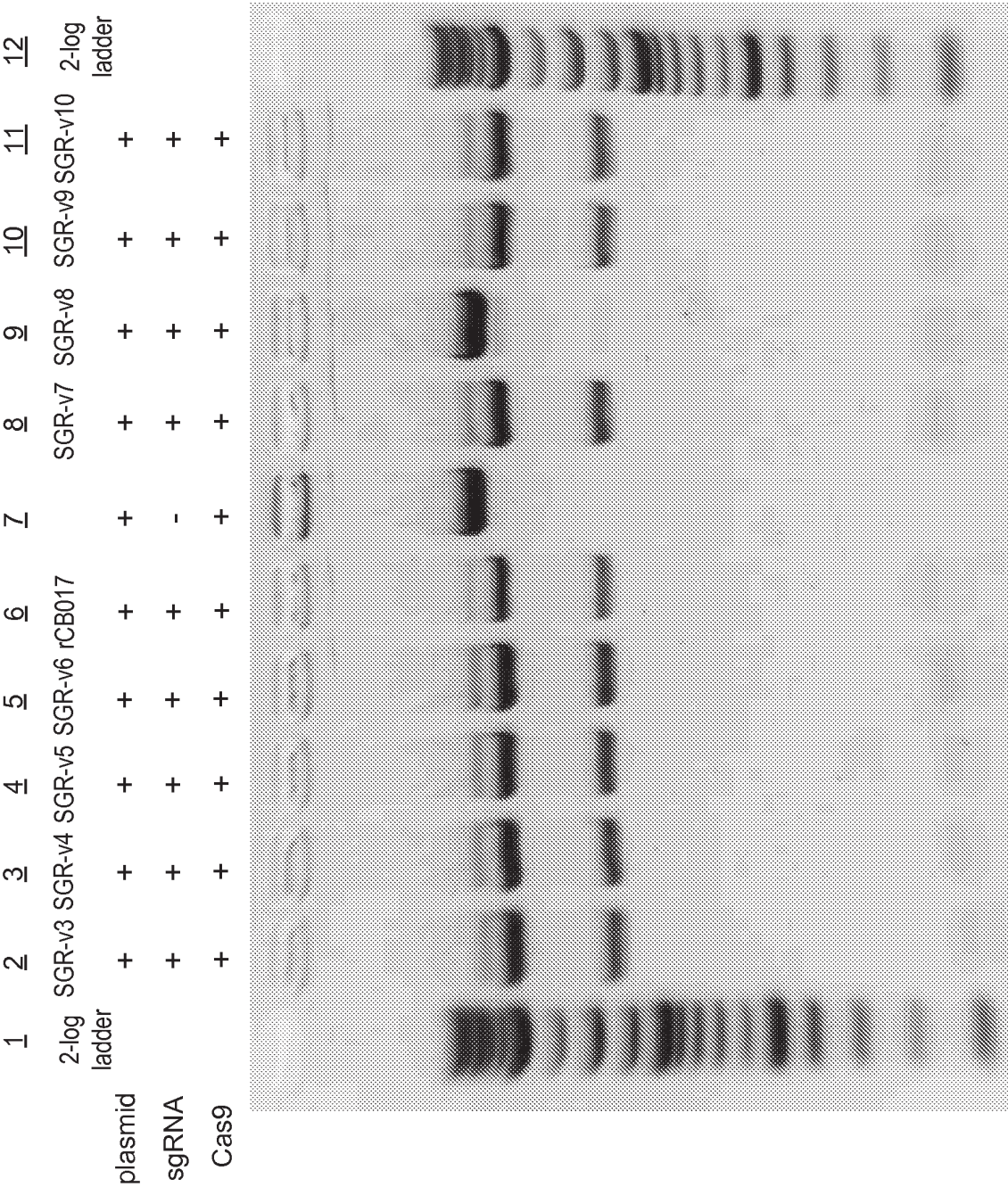


FIG. 23B

	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>
	2-log ladder	SGR- V11	SGR- V12	SGR- V13	SGR- V14	SGR- V15	
plasmid		+	+	+	+	+	+
sgRNA		+	+	+	+	+	-
Cas9		+	+	+	+	+	+

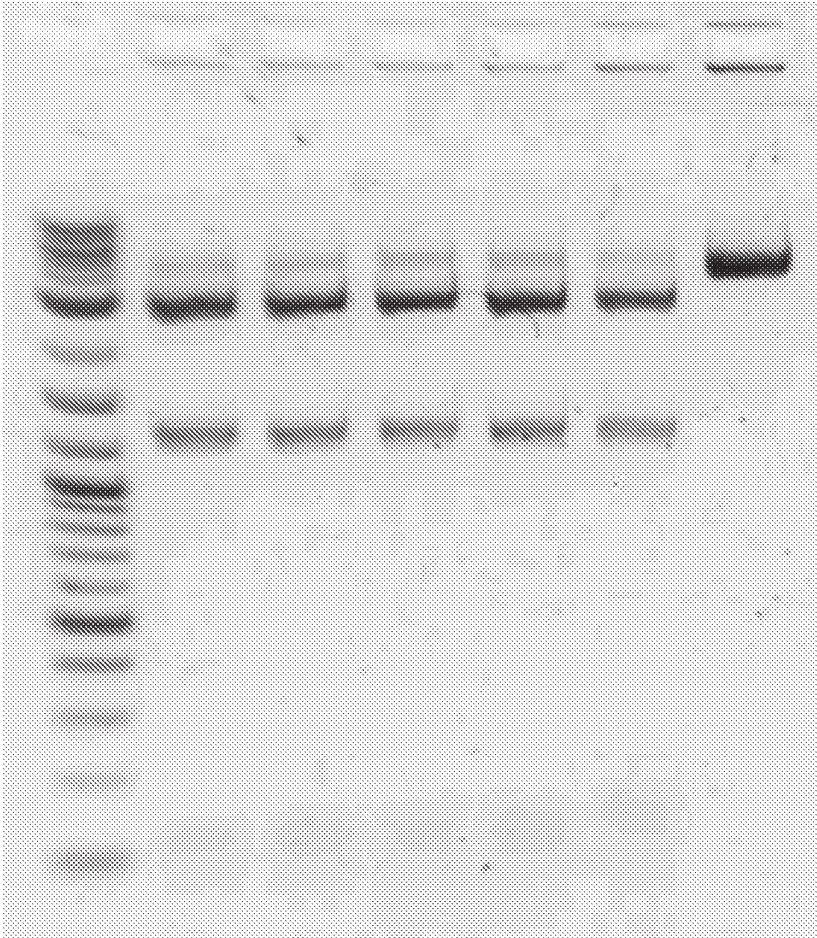
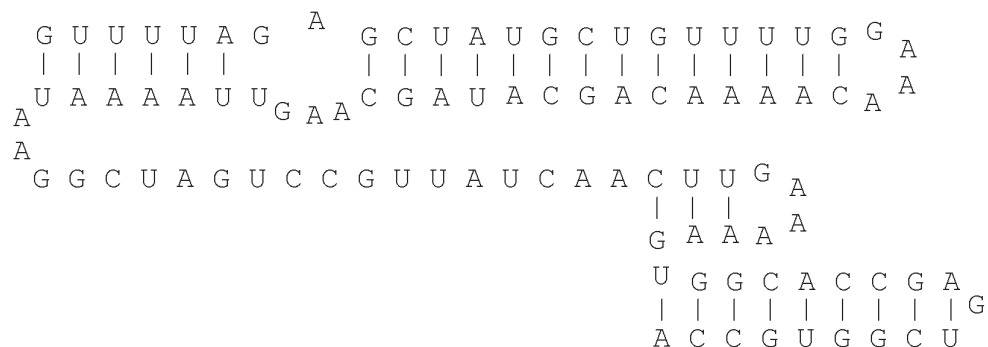


FIG. 23C

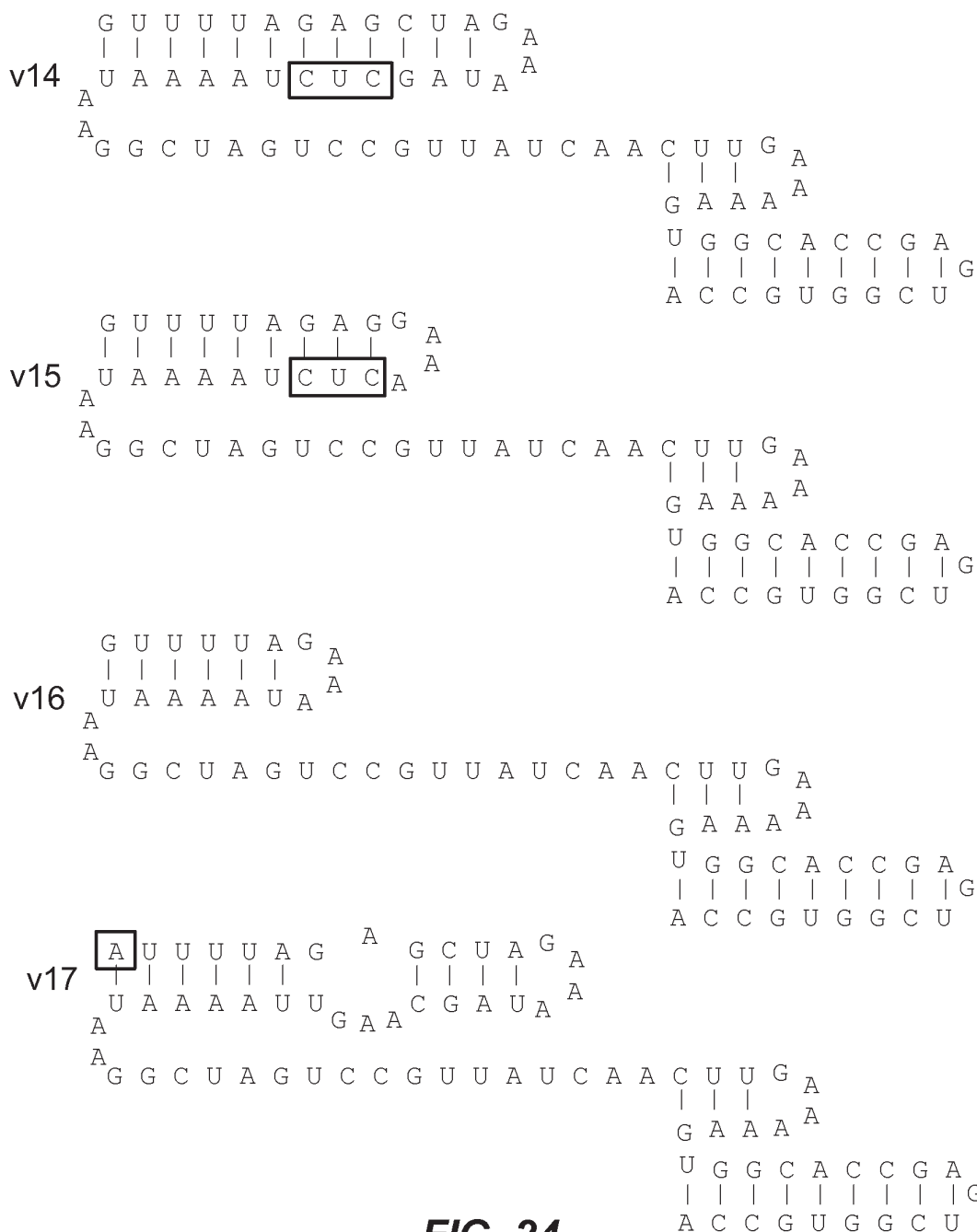
29 / 73

FL- TRACR-CRRNA

Sequence



Duplex Variants Group 2

**FIG. 24**

30 / 73

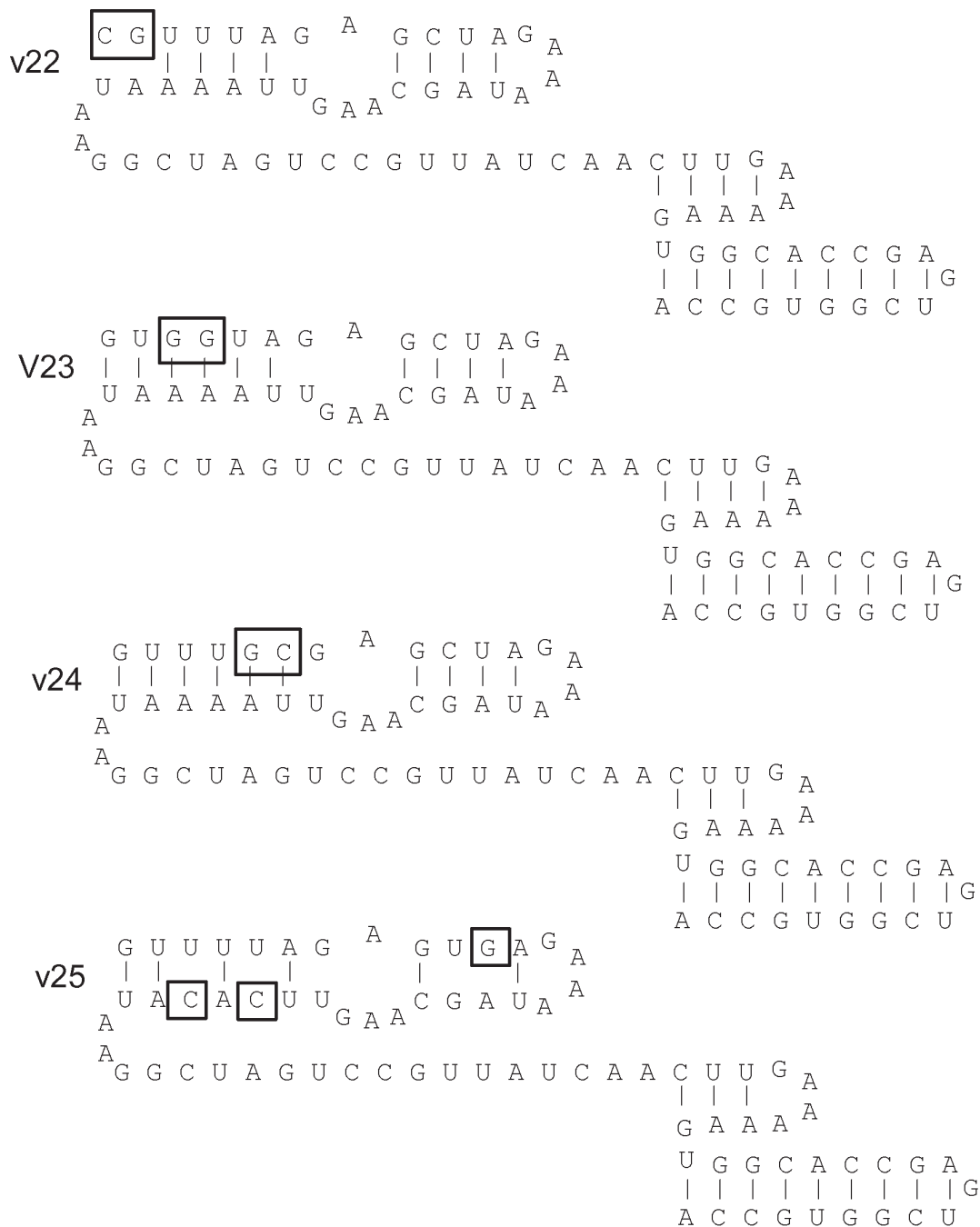
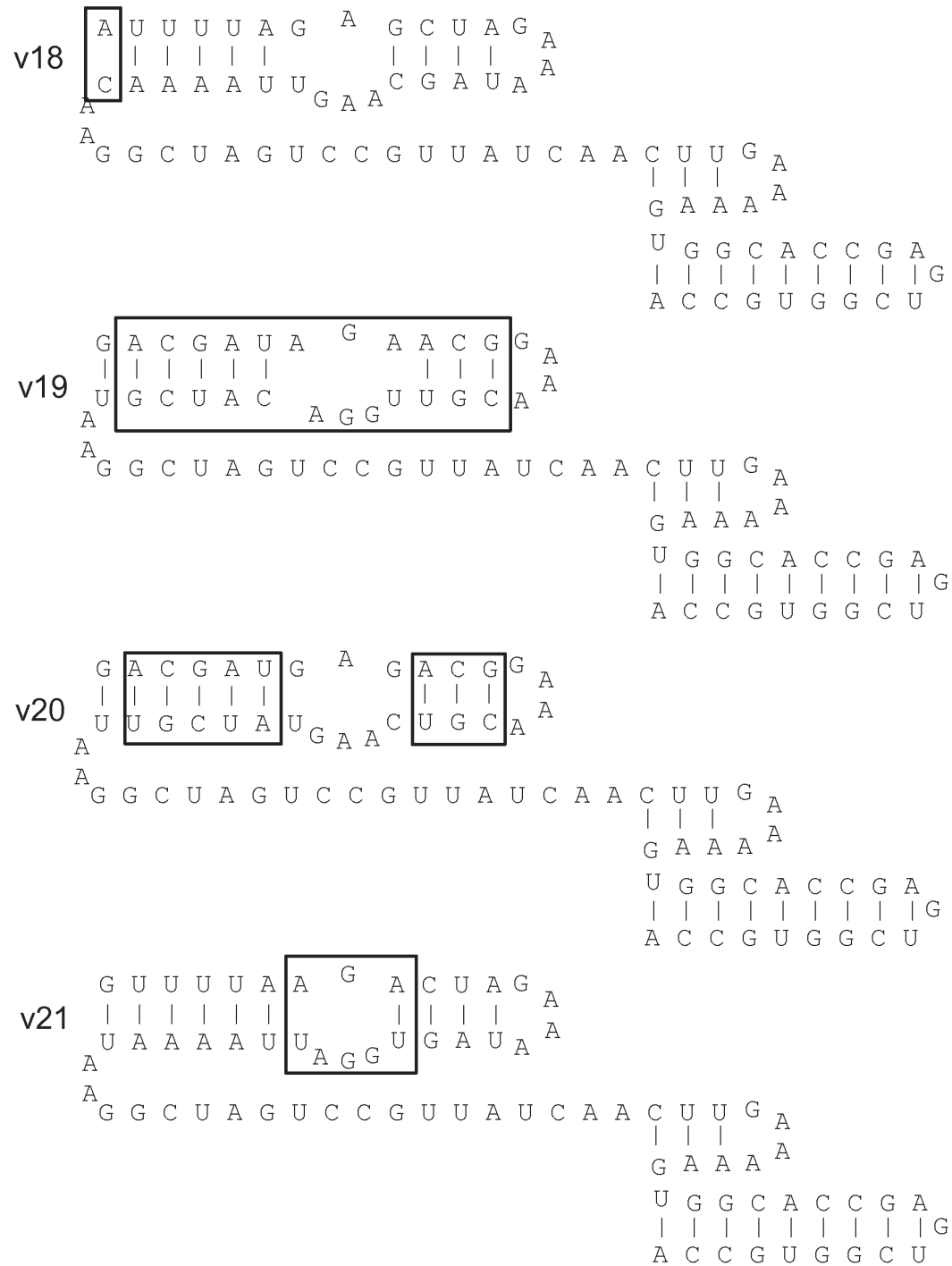


FIG. 24
(Continued)

31 / 73

**FIG. 25**

32 / 73

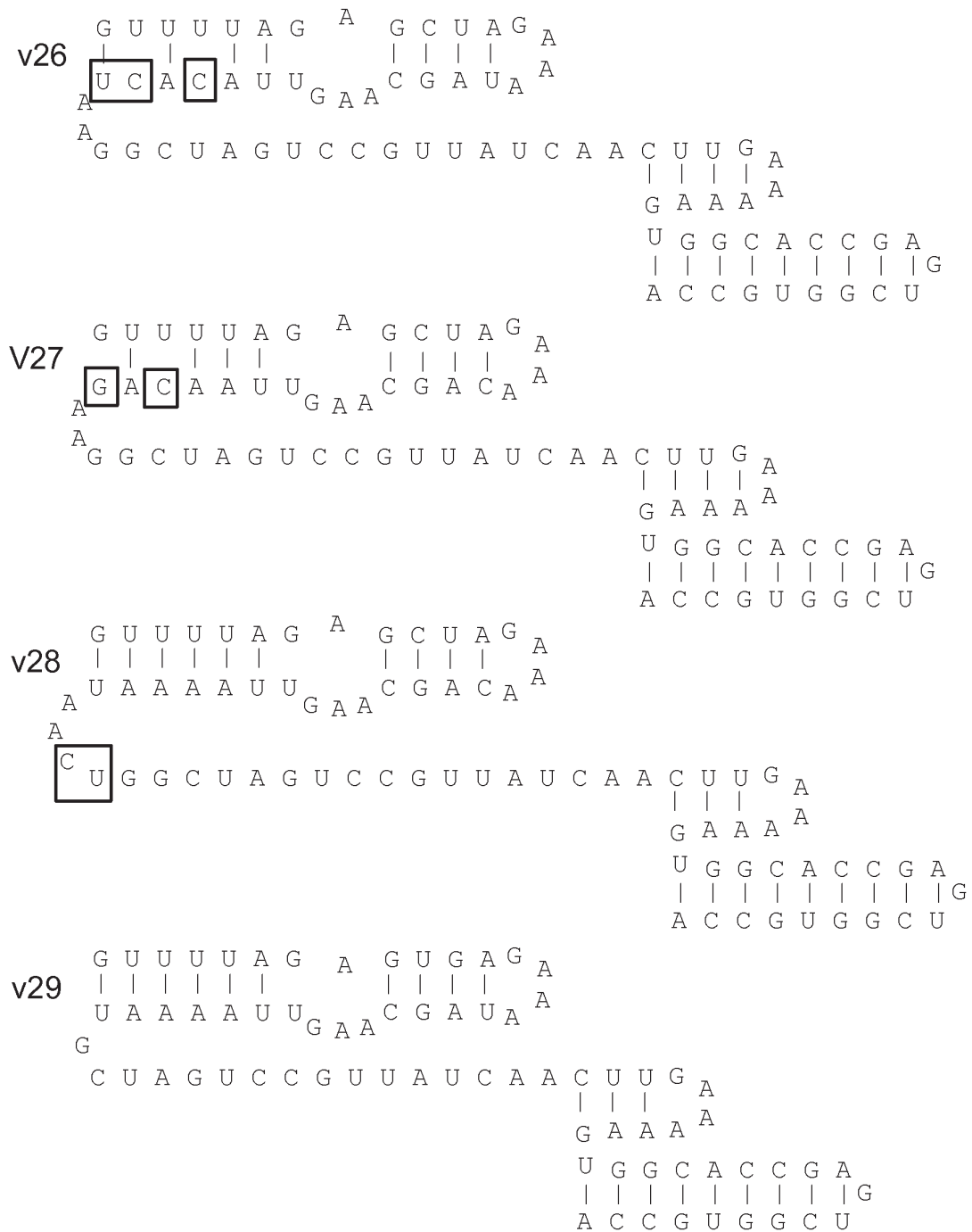
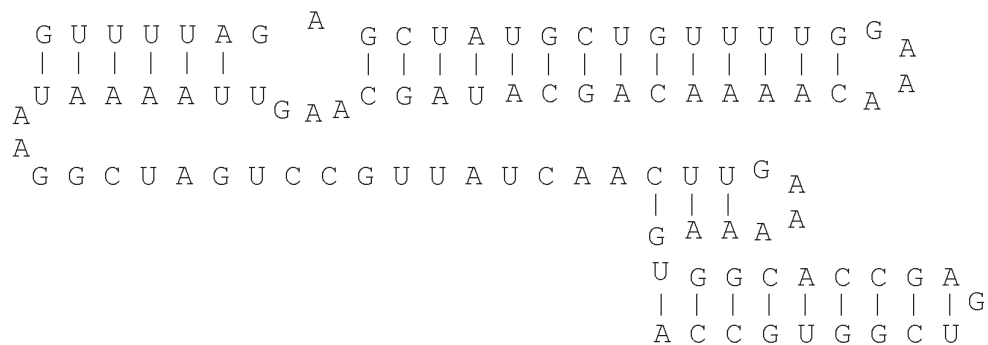


FIG. 25B
(Continued)

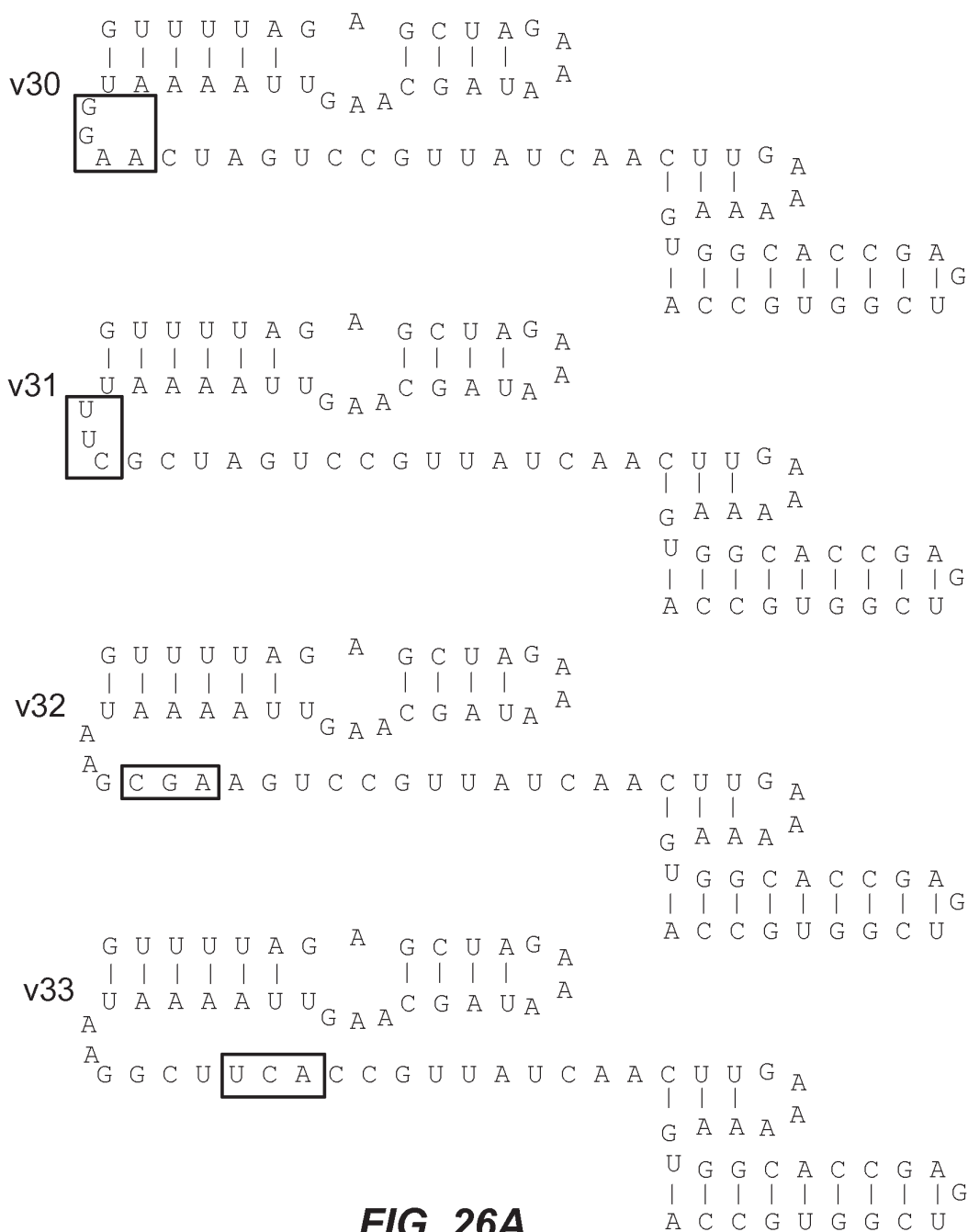
33 / 73

FL- TRACR-CRRNA

Sequence



Tracr Variants – Group 1

**FIG. 26A**

34 / 73

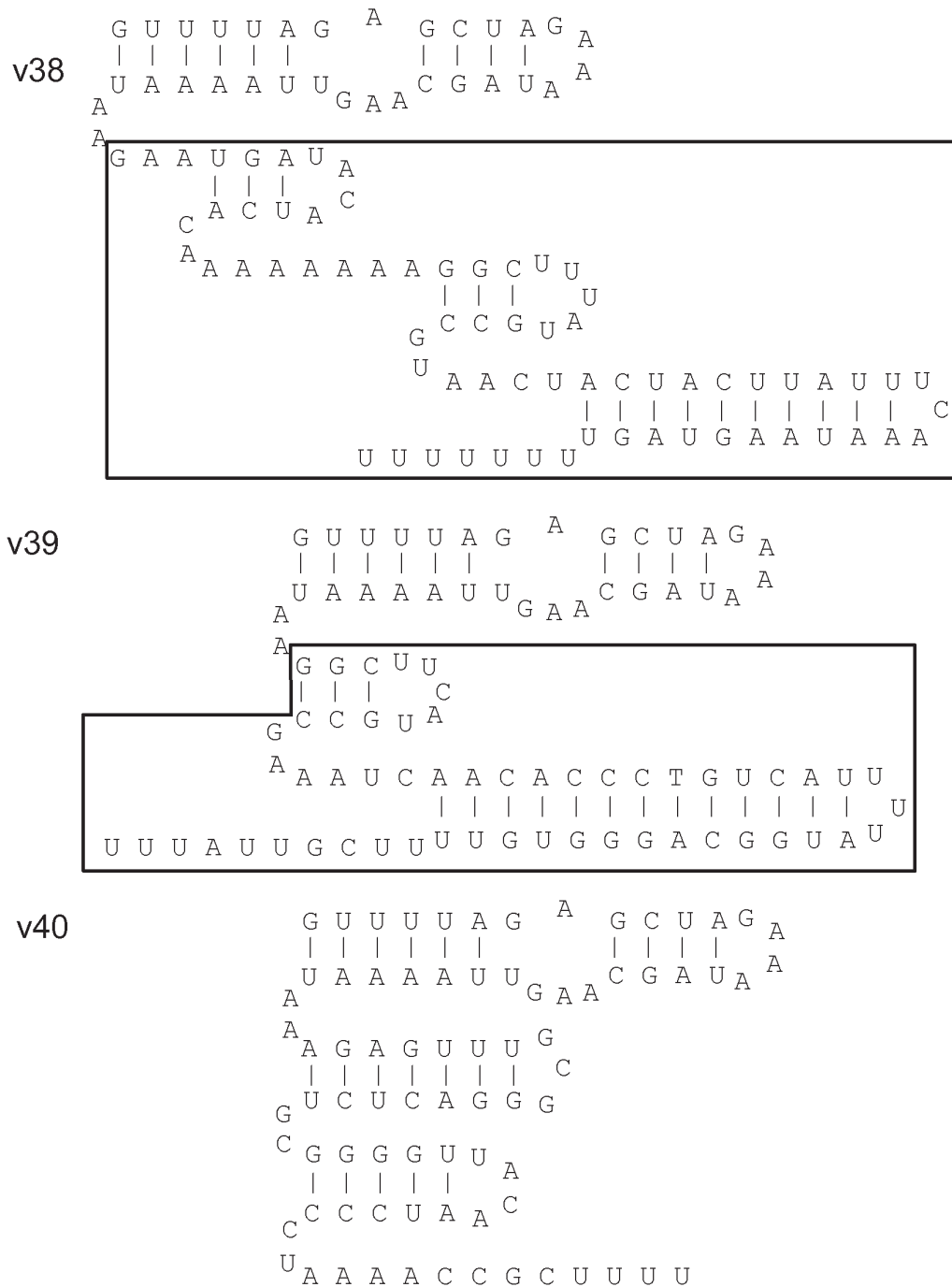
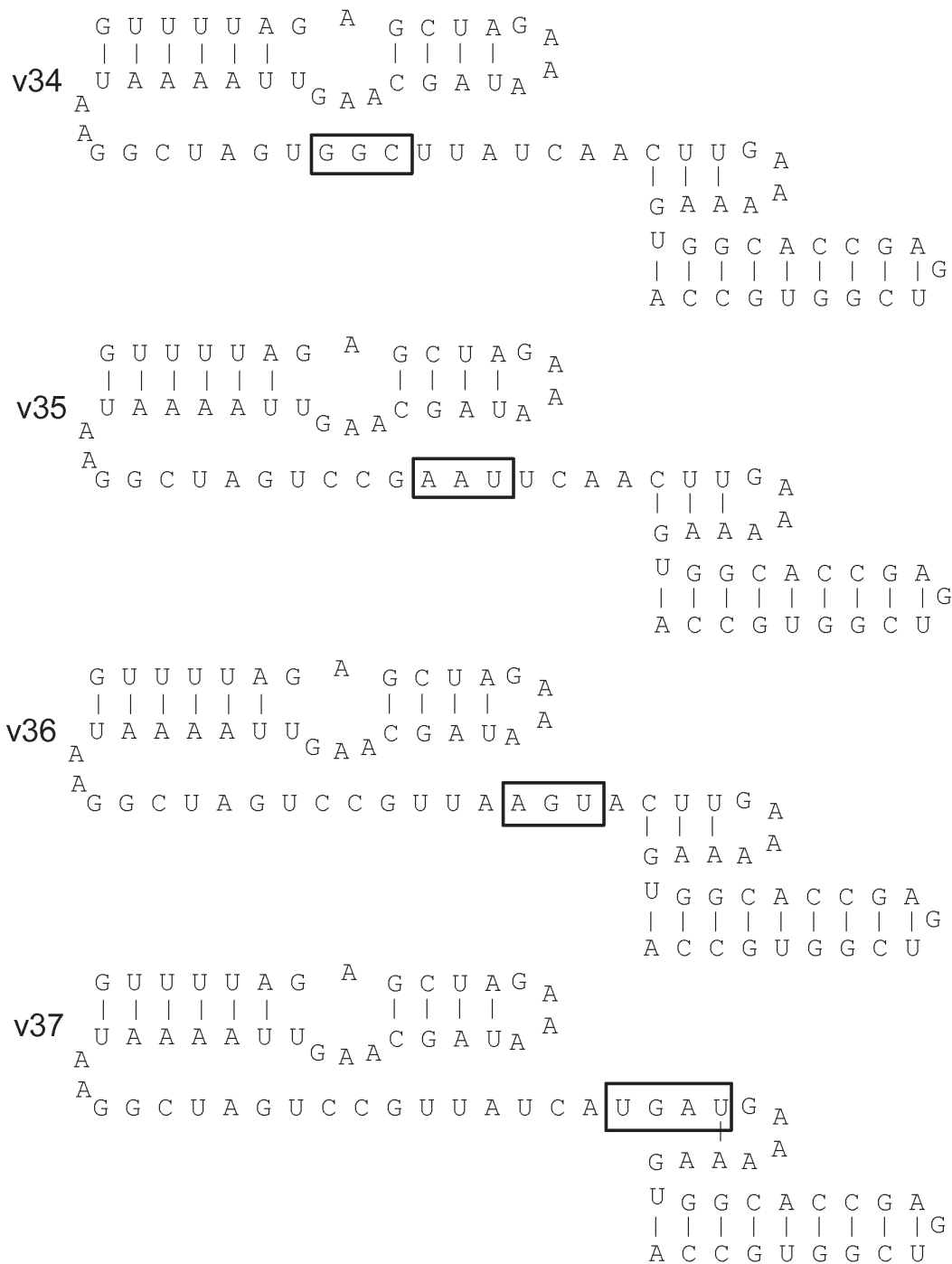


FIG. 26A
(Continued)

35 / 73

**FIG. 26B**

36 / 73

v41

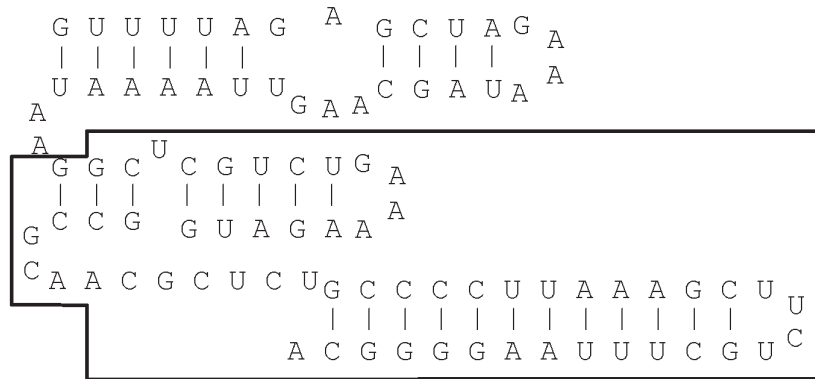


FIG. 26B
(Continued)

Csy4-tag Group 1

v43
 G U U U A G A G C U A G A
 | | | | | | | | | |
 U A A A U U G A A
 A
 A G C U A G U C C G U U A U C A A C U U
 A U A C U G C C G U A U A C G U A
 A G A G A C G G A U
 A
 A U G G A C U C G G A A
 A

FIG. 27A

FIG. 27A

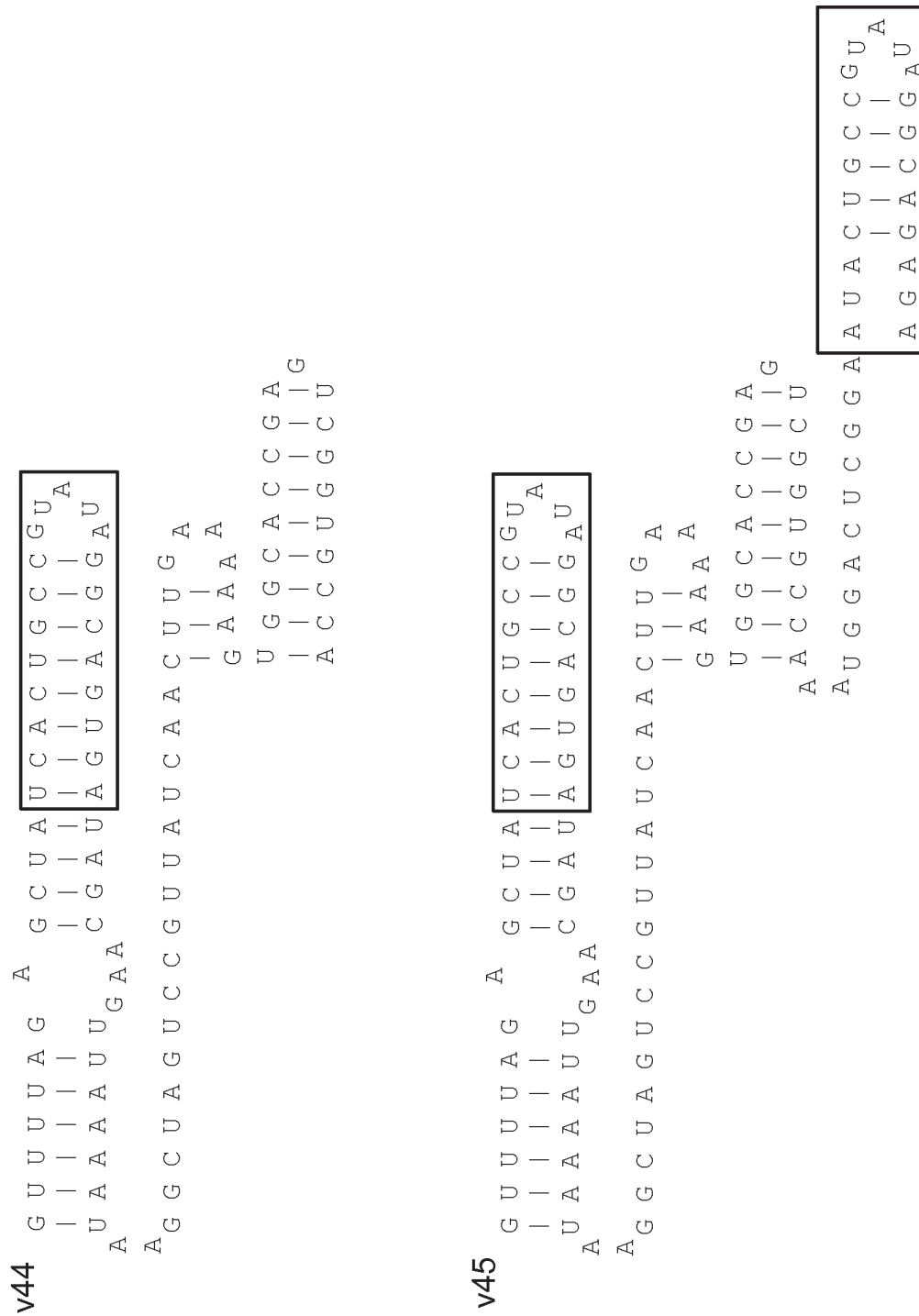


FIG. 27B

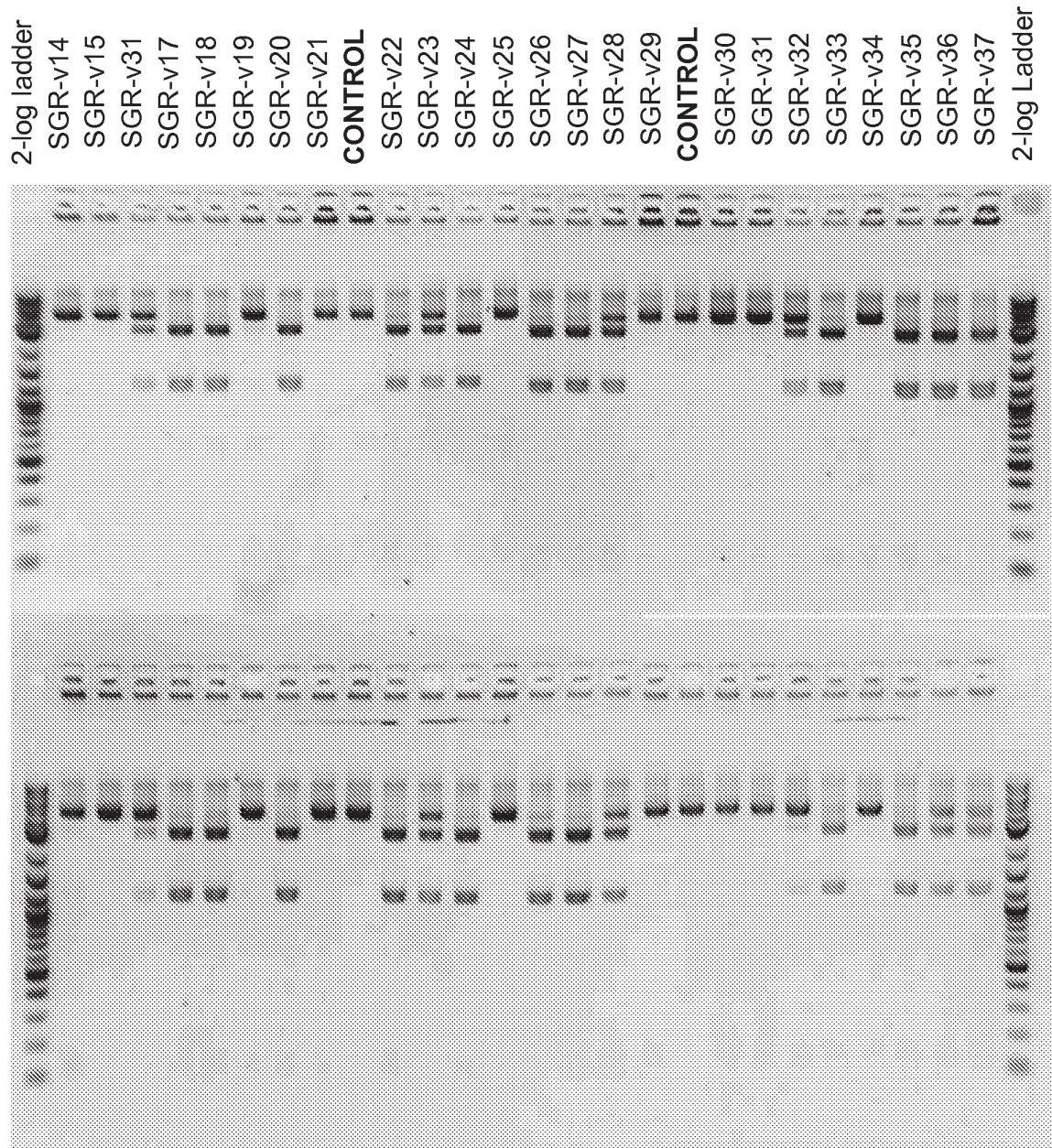
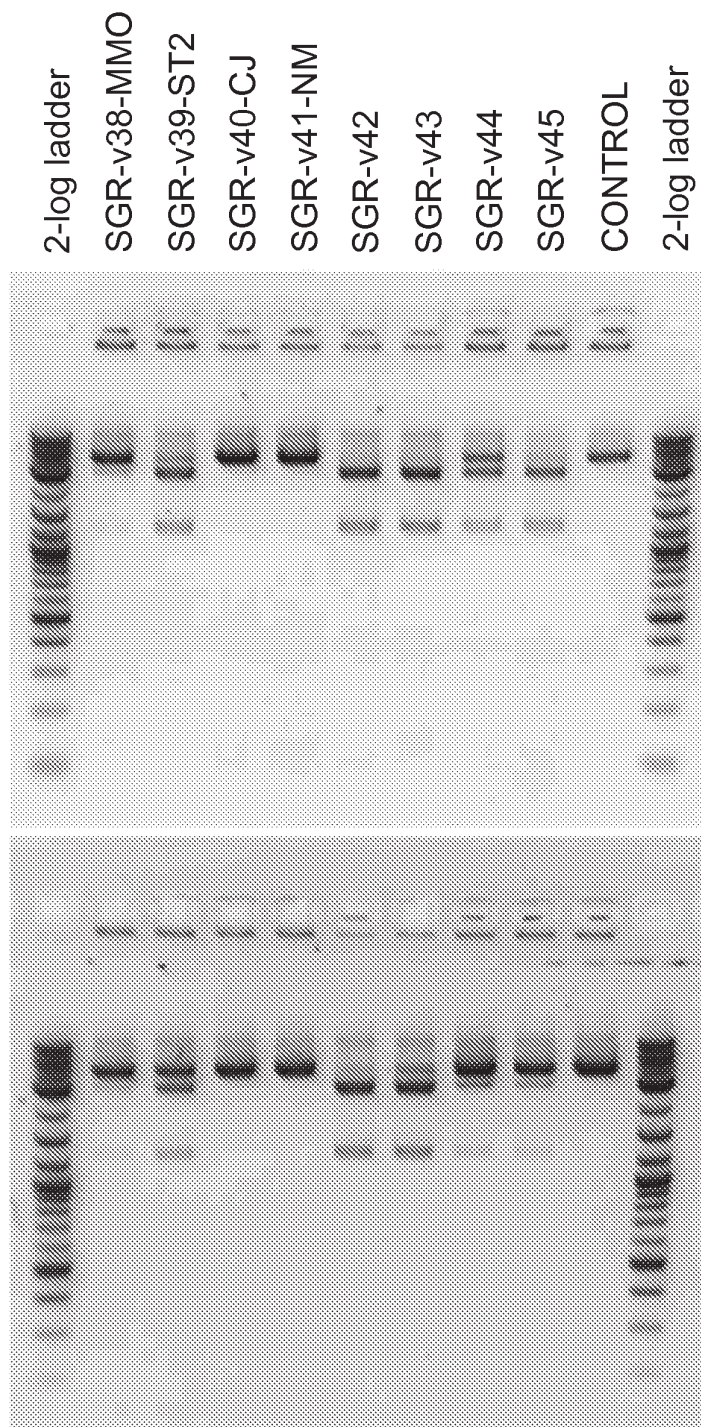
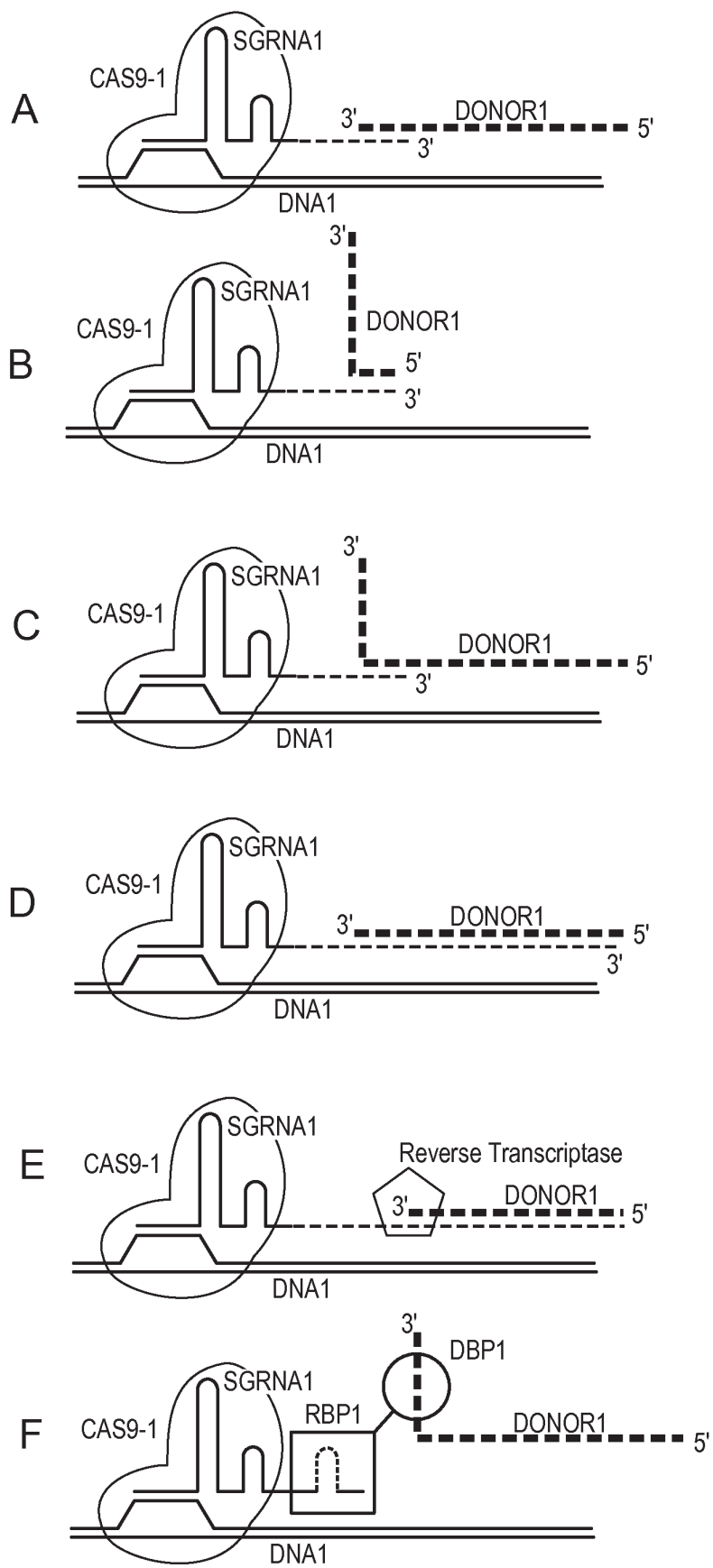


FIG. 28

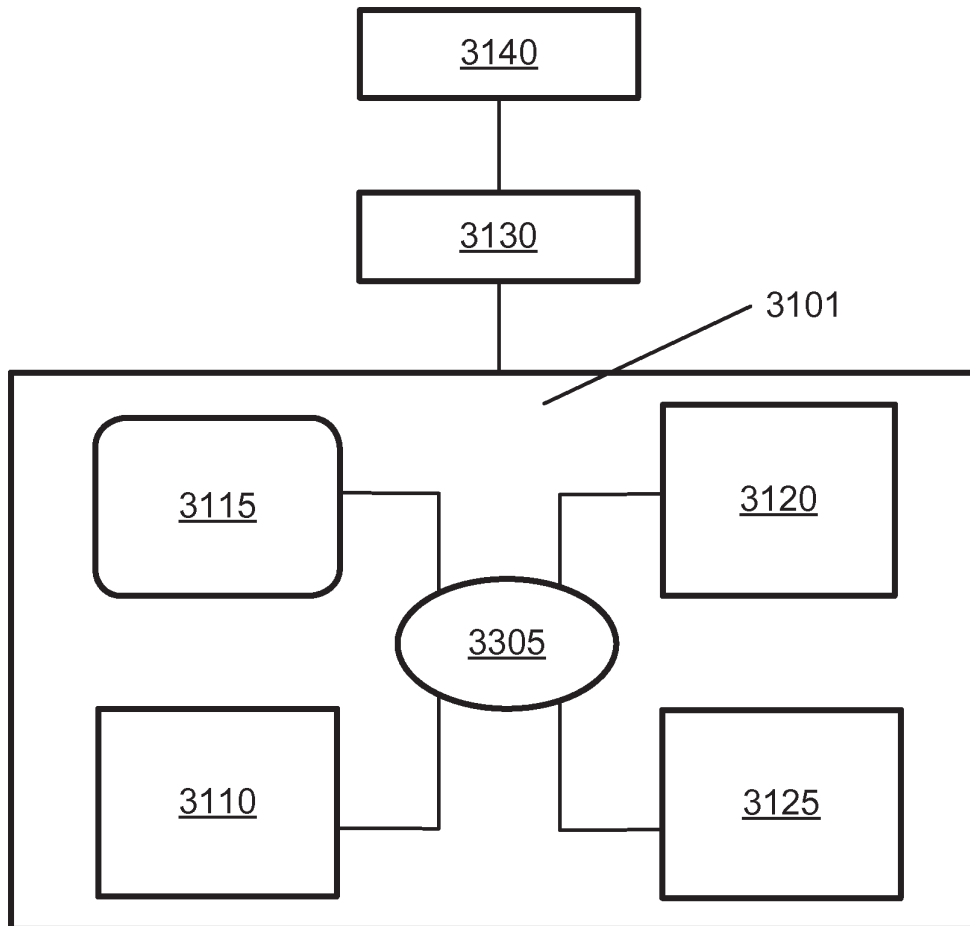
40 / 73

**FIG. 29**

41 / 73

**FIG. 30**

42 / 73

**FIG. 31**

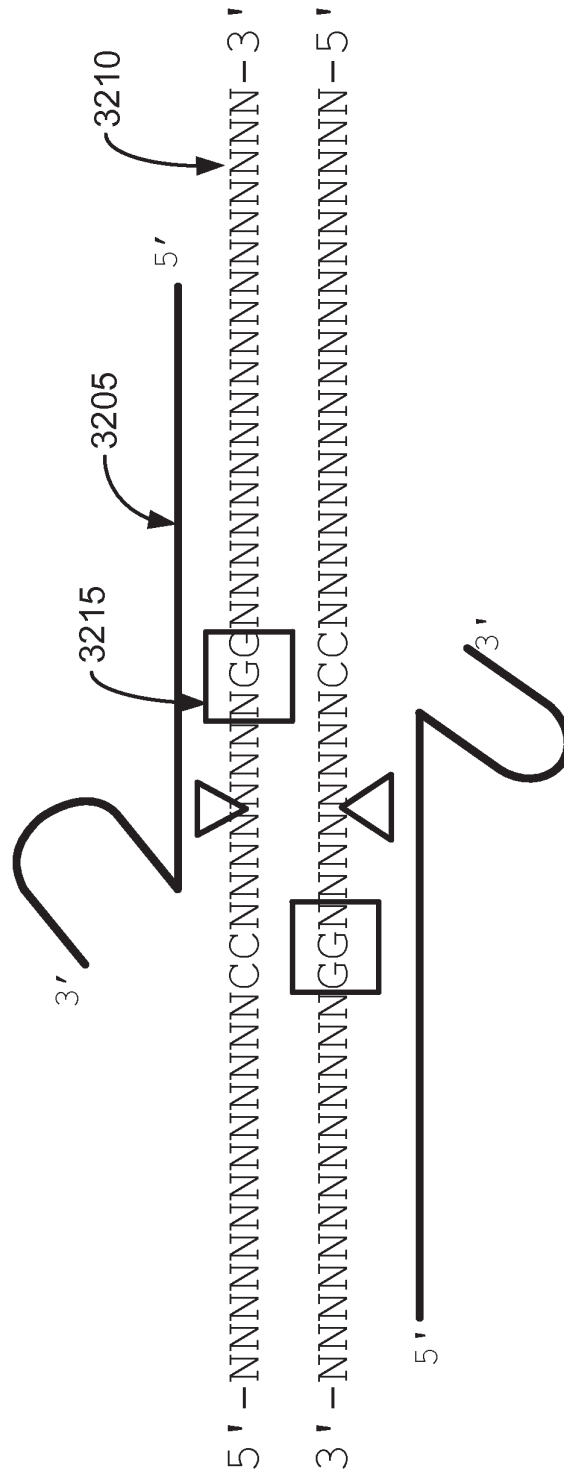


FIG. 32

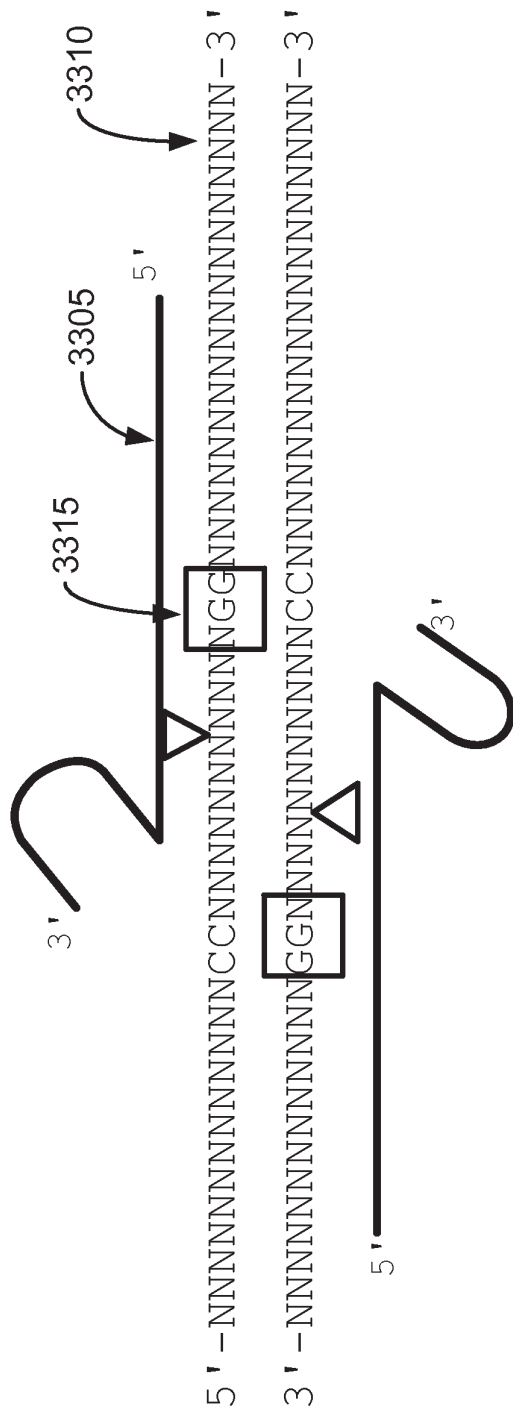


FIG. 33

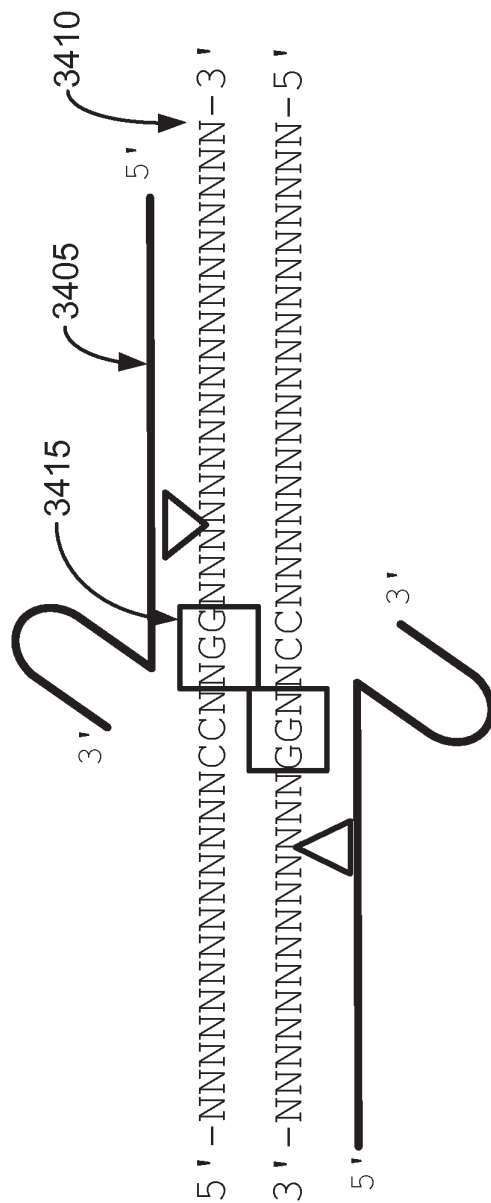


FIG. 34

TIGR01865: cas_Csn1 ?

CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI

CRISPR loci appear to be mobile elements with a wide host range. This model represents a protein found only in CRISPR-containing species, near other CRISPR-associated proteins(cas), as part of the NMEMI subtype of CRISPR/Cas locus. The species range so far for this protein is animal pathogens and commensals only [Mobile andextrachromosomal element functions, Other].

Links ?

Statistics ?

Structure ?

PubMed References ?

A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* 2002 Jan 15;30(2):482-496

TIGR01865 is classified as a model that may span more than one domain.

TIGR01865 is not assigned to any domain superfamily.

Sequence Alignment

Reformat Format: Hypertext Row Display: All 12 rows Color Bits: 3.0 bit Type Selection: top listed sequences

	10	20	30	40	50	60	70	80
consensus	* *	*	*	*	*	*	*	*
gi 22533915	1 EYILGLDIGIASVGWAI	VEDDYKVPAAKRLIDGG	-----VRNFTGAELPKTGETAA	LDRLARGARRIRRRKHRLRL	74			
gi 34483507	4 PYSIGLDIGTNSVGWS	IITDDYKVPAAKMRVLG	NTdkshieKKNLIGALLFDG	GNTAA-DRLLKRTARRRY	TRRRNRILYL	82		
gi 12721472	3 ERILGVDLGISSLGWA	IVEYDKDDEAANRIIDCG	-----VRLFTAAETPKKKE	SPNKARREARIGIRRV	LNRRVRMNM	76		
gi 24377777	7 SYILGLDLGIASVGWA	VEINEN-EDPIGLIDVG	-----VRIFERAEVPKTGES	LALSRRLARSTRRLIR	RRRAHRLLLA	79		
gi 13622193	4 PYSIGLDIGTNSVGWA	VVTDYKVPAAKMKVLG	NTdkshieKKNLIGALLFD	SGNTAE-DRLLKRTARR	RYTRRRNRILYL	82		
gi 41815893	7 DYFLGLDVGTSGVGA	VTDTDYKLLKANR-----	KDLWGMRCFETAETAE	-VRLHARGARRIERR	KKRIKLL	73		
gi 157150687	3 GLVLGLDIGIASVGW	IVDNETN-----GK-----	IHASRFLFPAA	TADNNVERRSNRQ	RRINPRKKHRSVRL	66		
gi 294660600	10 EVTIGLDLGVSGVGA	IVDNETN-----IHHLSR	LFSQAKTAE-DRRS	FRGVRRLI	IRRRKYKLR	71		
gi 218563121	2 ARILAFDIGISSIGWA	FAFSENDE-----LKDCG	-----VRIFTKVEN	PKTGESLALPRRL	ARSARKRLARRKAR	LNHL	68	
gi 370792169	7 PYTIGLDIGTNSVGWA	VLTDQYDLVKRMKIAGD	sekkqiKKNFWGVRL	FDEGQTAA-DRRMAR	TARRRIERRNR	ISYL	85	
	XXXXXXXXXXXXXXXXXXXX		XXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXX				

FIG. 35

46 / 73

		90	100	110	120	130	140	150	160
		*		*		*		*	
consensus	75	QELFSREGSLTDFDFrLENSFLVEEDKRN	---	---	---	---	---	---	---
gi	22533915	83	QEIFAEEMSKVDDSSFFhrLEDSFLVEEDKRGskypifatlq	---	---	---	---	---	---
gi	34483507	77	KKLFLRAGLIQDVLDL	---	---	---	---	---	---
gi	12721472	80	KRFLKREGILSTIDLEK	---	---	---	---	---	---
gi	24377777	83	QEIFSEEMGKVDDSSFFhrLEDSFLVTEDEKRGherpifgnle	---	---	---	---	---	---
gi	13622193	83	QEIFSNEMAKVDDSSFFhrLEESFLVEEDKKHerhpfifgniv	---	---	---	---	---	---
gi	41815893	74	QELFSQEIAKTDEGFFqrmKESPFYAEDKTIlgentlfndkdfadktyhkaypt	---	---	---	---	---	---
gi	218767588	83	RRLKREGVLQAADFdenGLIKSLP	---	---	---	---	---	---
gi	157150687	67	QDLFEGYGLLTDFSKVsmNLN	---	---	---	---	---	---
gi	294660600	72	VNLIWKYNSYFGFKNKedILNNYQEQQKLHN	---	---	---	---	---	---
gi	218563121	69	KHLIANEFKLNIEDYQsfDESakAYKGSli	---	---	---	---	---	---
gi	370792169	86	QGIFAEEMSKTDANFFcrLSDSFYVDNEKRNsrhpfatie	---	---	---	---	---	---
			XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX

		170	180	190	200	210	220	230	240
		*		*		*		*	
consensus	131	IIKHRGHFLIEGNDfdtAN	---	---	---	---	---	---	---
gi	22533915	161	IIKFRGHFLIEDSfdvRNtdiskqyqdfleifnttfenndllsqnvdveailtdkisksakkdrilaqypnqkstgifa	---	---	---	---	---	---
gi	34483507	131	IAKHRGYKFIGDDE	---	---	---	---	---	---
gi	12721472	126	LIKHRGYLSKRKNesqtNN	---	---	---	---	---	---
gi	24377777	161	IIKFRGHFLIEG-KfdtRNndvqrlfgeflavydntfensslgeqnqvveeilttdkisksakkdrvlklfpneksngrfa	---	---	---	---	---	---
gi	13622193	161	MIKFRGHFLIEG-DlnpDNsdvdklfiqlvtynqlfeenpinasgvdakailsarlsksrrlenliaqlpgekknglfg	---	---	---	---	---	---
gi	41815893	154	IIKFRGHFLIEG-DfdesENqfdtsiqalfeylredmevdidadsqkvkeilkdsslnkseksqsrlnkilglkpsdkqkka	---	---	---	---	---	---
gi	218767588	134	LIKHRGYLSQRKNEgetAD	---	---	---	---	---	---
gi	157150687	112	IVKRRGISYLDdASedgGT	---	---	---	---	---	---
gi	294660600	128	YLKNRGYFYEDNRDfnvYP	---	---	---	---	---	---
gi	218563121	125	IAKRRGYDDIKNSD	---	---	---	---	---	---
gi	370792169	164	IIKYRGNFLIEG-AldtQNTsvdgiykqfiqtynqvfasgiedgelkklednkdvakilvekvtrkeklkilypgek	---	---	---	---	---	---
			XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX	XXXXXX

FIG. 35
(Continued)

47 / 73

FIG. 35
(Continued)

consensus 193
 gi 22533915 376 ylsklltkqedsenflekiknedflrkqrtfdngsiphqvhlTELKAIIRRRQSeypflken-----iklltTFR 235
 gi 34483507 186 -----RNKHGDYEISIHrdLLVEEVEAIFVAQQ-----qdriekilTFR 448
 gi 12721472 185 -----PNQrGAYThRFNRLDLLAEINLLFAQQH-----EMR 216
 gi 24377777 375 ylgllnkiegsgyfldkieredflrkqrtfdngsiphQIHLQEMRAIIRrQAefypfladn-----qdriekilTFR 215
 gi 13622193 375 fikpilekmdgteellvklhnredllrkqrtfdngsiphQIHLGELHAILRRQEdfypflkdn-----rekiekilTFR 447
 gi 41815893 385 tilsakseikevndilteietgtflPKQISKSNAEIPYQLRKMELEKILSNAEkHfslkqkdekglshsekiimllTfK 464
 gi 218767588 193 -----RNQrGDYSHtFSRKDLQAEILLLFEKQK-----EFG 223
 gi 157150687 172 -----eeNGEKHRLINVFSTsAYRKEAERILRKQqefnskitde-----fiedyliilTgK 222
 gi 294660600 188 -----wleEVKkVLSNQTGLPEKfKEEYESLFSYVRnyseg-----pgsiNSV 230
 gi 218563121 181 -----ftnvRNKKESYERCIaQsFLKDELKLIffKKQR-----EFG 215
 gi 370792169 384 ymktleniegadyfiakienfllrkqrtfdngAIPHQLHLEELEAILHQQAkyypflken-----ydkikslvTFR 456
 YYYXXXXXXXXXXXXXXXXXXXXXXXXX

48 / 73

consensus 236 IPYYIGPLAEGKSEFA-----FVDKPSAenfiENMTGKCTYLPEEKRAPKHSLLAEKFTV 291
 gi 22533915 449 IPYYIGPLAREKSDFAwmtrktdds-----irpwnfedLVdKEKSaEakIHRMTNNDfYLPEEKVLPKHSLLIYEKFTV 521
 gi 34483507 217 STIATDALKAAyREIA-----FFVRPMQR---IEKMVGhCTYfPEERRAPKSAptAEKFIA 269
 gi 12721472 216 NPHCKEHIQqYMTELL-----MWQKPALSgeaILKMLGKCTHEKNEfKAAKHTYsAERfVW 271
 gi 24377777 448 IPYYVGPLARGKSDRAWlsrksadk-----itpwnfdeIVdKESSAeakINRMtNYDLYLPNQKVLPHSLLYEKFTV 520
 gi 13622193 448 IPYYVGPLARGNSRFawmtrkseat-----itpwnfeeVVDKGASAgSkIERMTNfDKNLpNEKVLPHSLLYEYFTV 520
 gi 41815893 465 IPYYIGPINDNHKKfFpdrCwvVkkekspsgkttPwnffdhIDKEKTAeafITSRtNfCTYLVGESVLPKSSLLYSEYTV 544
 gi 218767588 224 NPHVSGGLKEGIEtLL-----MTQRPALSgdaVQKMLGHCTfEPAEPKAAKNTYtAERfIW 279
 gi 157150687 223 RKYYHGPGNEKSRtDYg-----rFRtDGTtldniFGILIGKCTfYtEEYrASKASyTAQEFNL 280
 gi 294660600 231 SPYGIYHLDEKEGKV-----QYnniWDKTIGKCSIFPDEYRAPKNSPIAMIFNE 281
 gi 218563121 216 FSFS-KKFEeeVLSVA-----FYKRALKD---FshLVGNCsFFtDEKRAPKNSPLAFMFVA 267
 gi 370792169 457 IPYFVGPLANGQSEFAwltRkadge-----irpwnieeKVDFGKSAdvfIEKMTNkDtyLPkENVLPHSLLCYQKYLv 529
 XXXX XXXXXXXXXXXX
 XXX XXXXXXXXXXXXXXXXXXXX

FIG. 35
 (Continued)

49 / 73

	650	660	670	680	690	700	710	720									
	*		*		*		*										
consensus	354	---	LDNAEKAFNISL	KTYHLKRKALg	DKDILLDn	---	PKNPKDLD	DEIVKILTL	YKDREM	---	iK	407					
gi 22533915	583	---	LDKENKAFNASL	GTYHDLEKIL	-DKDFLDn	---	PDNESILED	IVQTLTLF	EDREM	---	iK	635					
gi 34483507	347	nepte	lefdkVEAEKKAWISL	RGAAKLREALg	NEFYGRf	---	VALGKHADEAT	KILTYYK	DEG	---	Q	407					
gi 12721472	331	---	KENAEASATFMEL	KAWHAIRKALe	NQGLKDt	wqdl	AKKPDLLD	DEIGTAFSL	YKTDE	---	D	386					
gi 24377777	582	---	LDKENKVFNASY	GTYHDLCIL	-DKDFLDn	---	SKNEKILED	IVLTLTLF	EDREM	---	iR	634					
gi 13622193	583	---	VEDR	---	FNASLGYHDLLK	IIk	DKDFLDn	---	EENEDILED	IVLTLTLF	EDREM	---	iE	633			
gi 41815893	604	---	LGID	-KECTSSLKS	YIELKNIF	---	GKQVDe	---	ISTKNMLEE	IIRWATIY	DEGEgkt	---	ilkt	658			
gi 218767588	339	---	KDNAEASTLMEM	KAYHAISRALe	KEGLKDKk	spL	NLSPELQ	DEIGTAFSL	FKTDE	---	D	394					
gi 157150687	339	---	VDVNNKPEMH	TFEYVRKMQS	LEtIK	VEELp	---	RKVDEL	LAHILTLN	TEREgi	---	eE	390				
gi 294660600	361	hkwlk	gkLNTNGKIQYND	LSSLAKFV	HKLkQHLK	Ldfl	ledQY	TPLDKIN	FLQSLVYV	LgKHLr	ysnr	vd	sanl	keys	440		
gi 218563121	324	---	---	EKGTYFIE	FKKYKEFI	KALgEHNl	---	---	SQDDLNEI	AKDITLIK	DEi	---	---	K	367		
gi 370792169	590	---	---	LEDS	---	FNSSYS	TYHDLLK	VGiKQ	EILDn	---	PVNTEML	ENIVKIL	TVFED	KRM	---	iK	640
		X	XXXXXXXXXX	XXXXX	XXX				XXXXXXXXXX	XXXXX							X

FIG. 35
(Continued)

50 / 73

FIG. 35
(Continued)

51 / 73

FIG. 35
(Continued)

52 / 73

[illegible]

FIG. 35
(Continued)

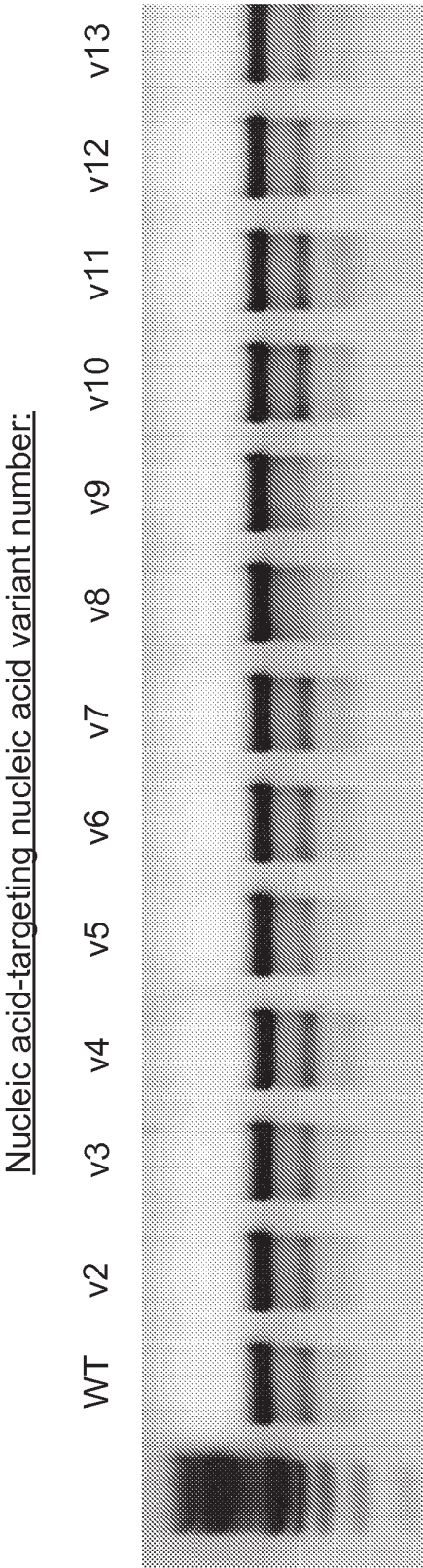
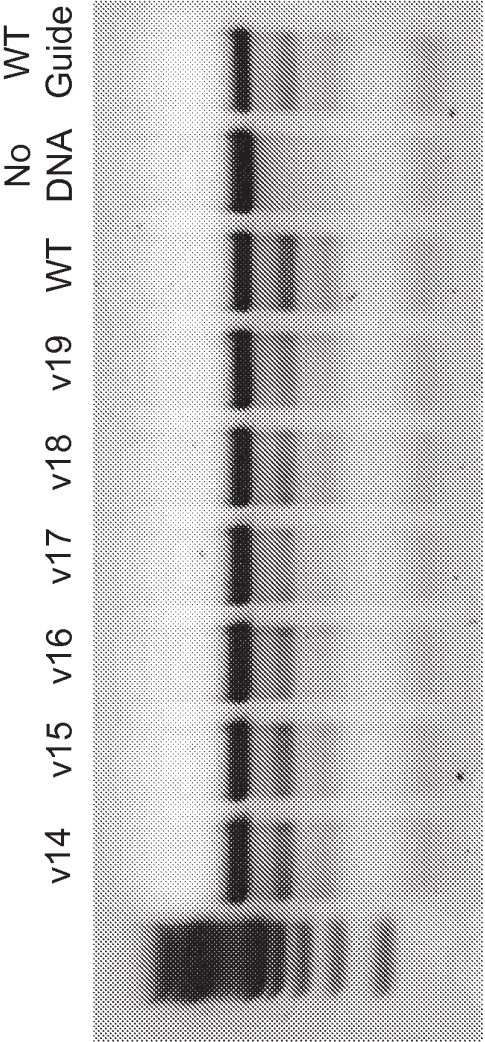
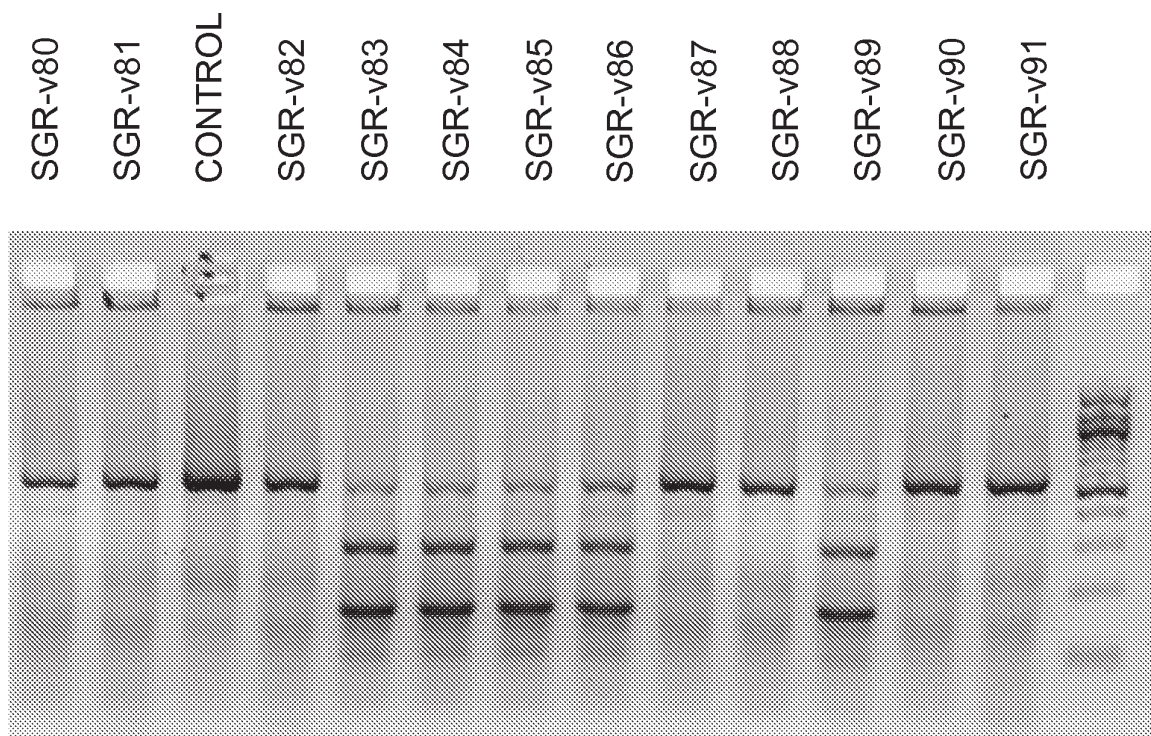


FIG. 36

54 / 73

**FIG. 37A**

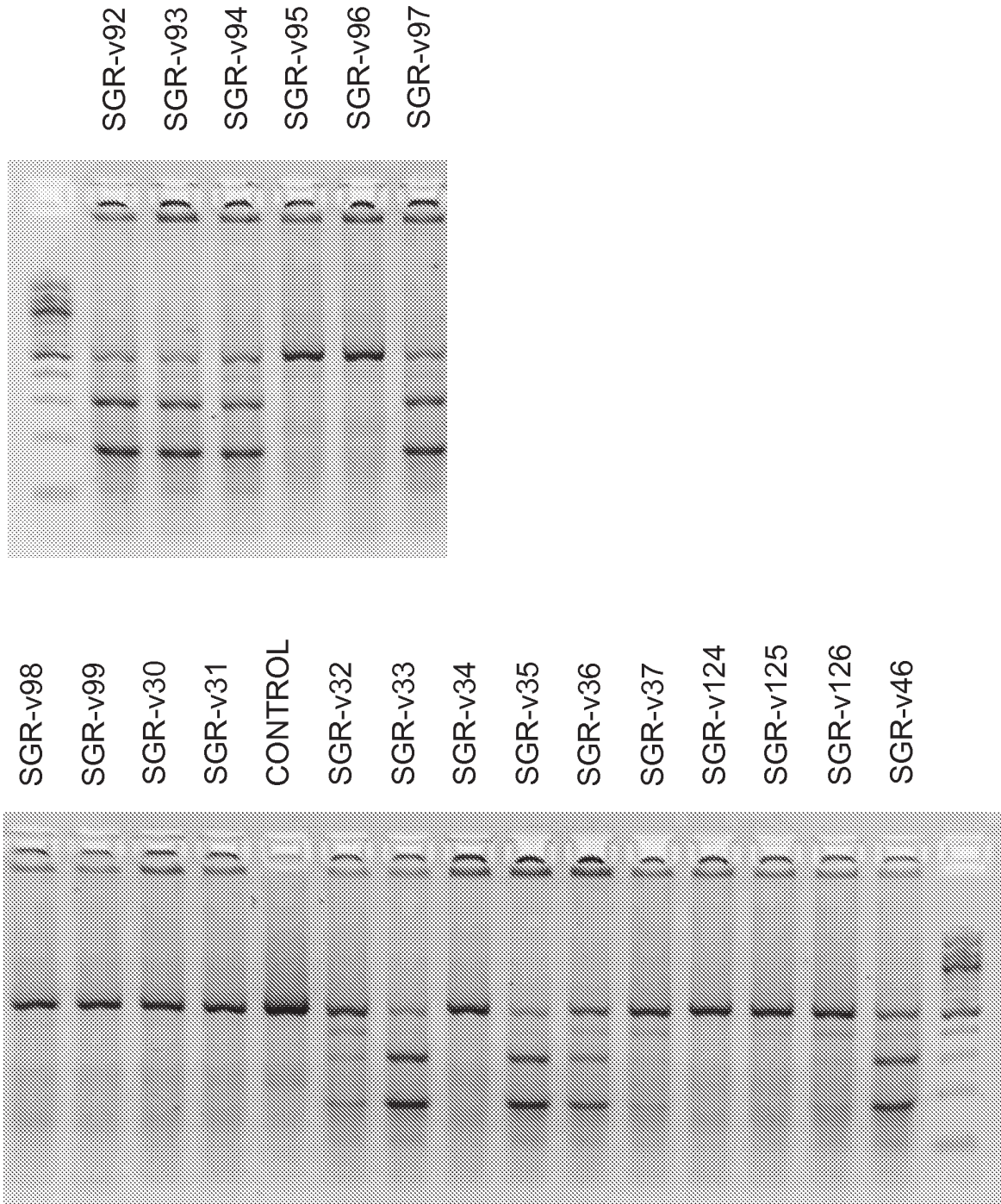


FIG. 37B

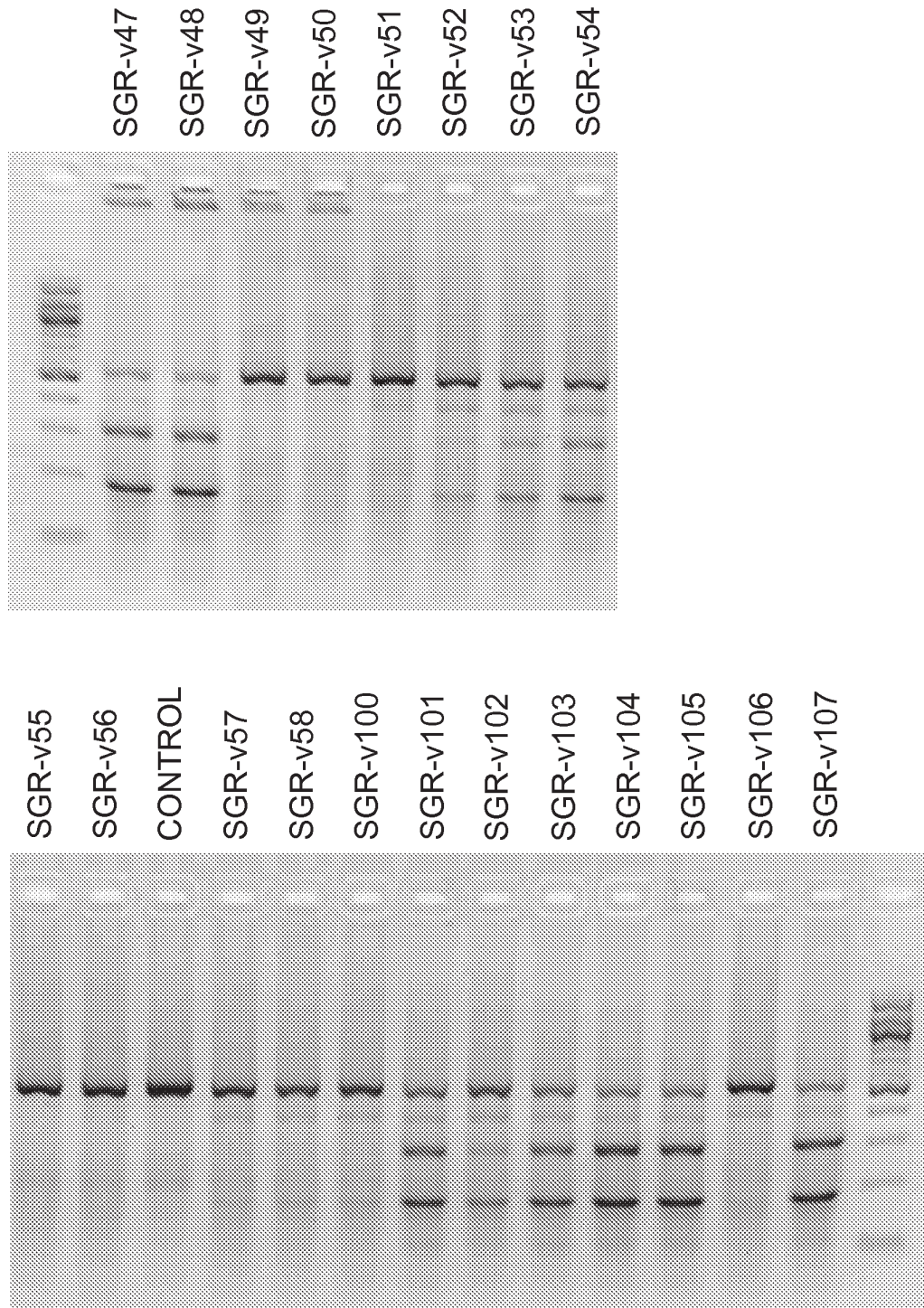
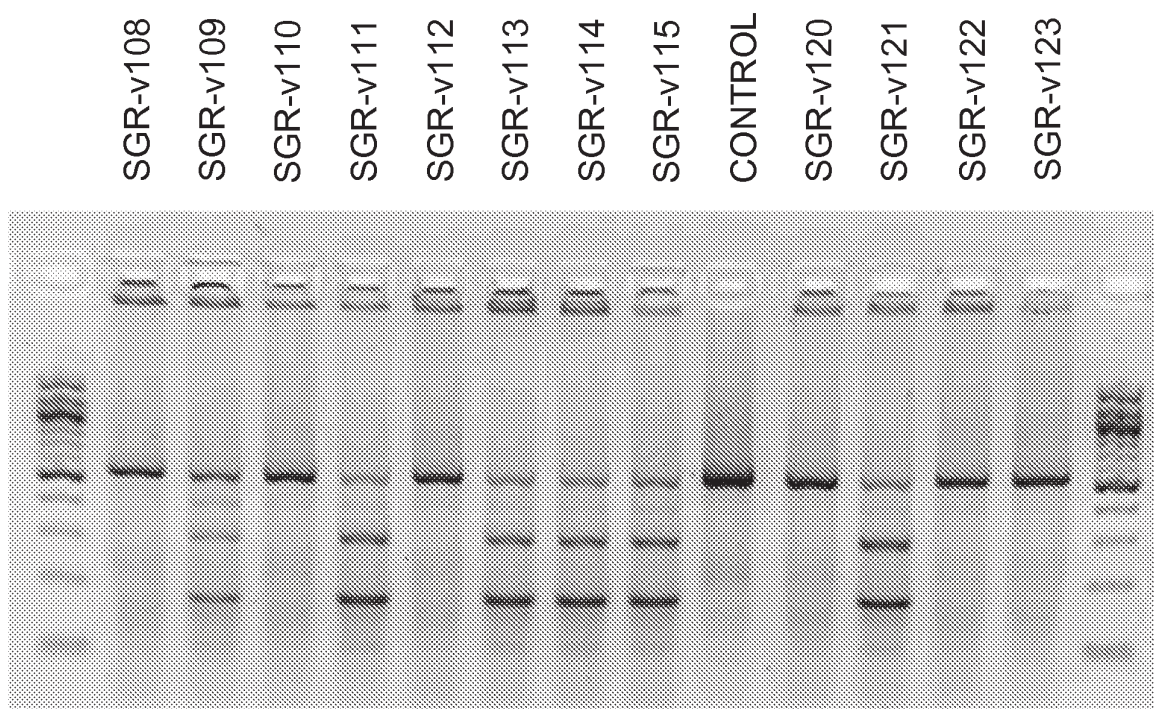


FIG. 37C

57 / 73

**FIG. 37D**

Csy4 sequence

>gi|107101871|ref|ZP 01365789.1| hypothetical protein PaerPA 01002916 [Pseudomonas aeruginosa PACS2]
 MDHYLDIRLRPDPEFPFPAQLMSVLFQKLHQALVAQGGDRIGVSFPDLDESRSLGERLRIHASADDLRALLARPWLEGLRDHLQFGEPAVVPHPPTPYRQVS
 RVQVKSNNPERLRRRLMRRHELSEEEARKRIPDTPVARALDLPFVTLRSQSTGQHFRLLFIRHGPLQVTAEEGGFTCYGLSKGGFVPWF

>gi|254235433|ref|ZP 04928756.1| hypothetical protein PACG 01340 [Pseudomonas aeruginosa C3719]
 MDHYLDIRLRPDPEFPFPAQLMSVLFQKLHQALVAQGGDRIGVSFPDLDESRSLGERLRIHASADDLRAL
 LARPWLEGLRDHLQFGEPAVVPHPPTPYRQVSRVQAQKSNPERLRRRLMRRHDLSEEEARKRIPDTPVARTLD
 LPFVTLRSQSTGQHFRLLFIRHGPLQATAEEGGFTCYGLSKGGFVPWF

58 / 73

>gi|254240857|ref|ZP 04934179.1| hypothetical protein PA2G 01531 [Pseudomonas aeruginosa 2192]
 MDHYLDIRLRPDPEFPFPAQLMSVLFQKLHQALVAQGGDRIGVSFPDLDESRSLGERLRIHASADDLHAL
 LARPWLEGLRDHLQFGEAAVVPHPPTPYRQVSRVQAQKSNPERLRRRLMRRHDLSEEEARKRIPDTPVARTLD
 LPFVTLRSQSTGQHFRLLFIRHGPLQATAEEGGFTCYGLSKGGFVPWF

FIG. 38

59 / 73

His29Ala
1 mdhyldirlr pdpefppaql msvlfgkl**a**q alvaqggdri gvsfpdlides rsrlgerlri
61 hasaddlral larpwleglr dhlqfgepav vphptpyrqv srvqvksnpe rlrmlmrrh
121 dlseeeeark ipdtvarald lpfvtlrsqs tgqhfrlfir hgplqataee ggftcyglsk
181 ggfvpwf

His29Ala/Ser50Cys
1 mdhyldirlr pdpefppaql msvlfgkl**a**q alvaqggdri gvsfpdlide**c** rsrlgerlri
61 hasaddlral larpwleglr dhlqfgepav vphptpyrqv srvqvksnpe rlrmlmrrh
121 dlseeeeark ipdtvarald lpfvtlrsqs tgqhfrlfir hgplqataee ggftcyglsk
181 ggfvpwf

FIG. 39

60 / 73

Csy4 sequences

<213> Pseudomonas aeruginosa

Met Asp His Tyr Leu Asp Ile Arg Leu Arg Pro Asp Pro Glu Phe Pro
 Pro Ala Gln Leu Met Ser Val Leu Phe Gly Lys Leu His Gln Ala Leu
 Val Ala Gln Gly Gly Asp Arg Ile Gly val Ser Phe Pro Asp Leu Asp
 Glu Ser Arg Ser Arg Leu Gly Glu Arg Leu Arg Ile His Ala Ser Ala
 Asp Asp Leu Arg Ala Leu Leu Ala Arg Pro Trp Leu Glu Gly Leu Arg
 Asp His Leu Gln Phe Gly Glu Pro Ala Val Val Pro His Pro Thr Pro
 Tyr Arg Gln Val Ser Arg Val Gln Val Lys Ser Asn Pro Glu Arg Leu
 Arg Arg Arg Leu Met Arg Arg His Asp Leu Ser Glu Glu Glu Ala Arg
 Lys Arg Ile Pro Asp Thr val Ala Arg Ala Leu Asp Leu Pro Phe Val
 Thr Leu Arg Ser Gln Ser Thr Gly Gln His Phe Arg Leu Phe Ile Arg
 His Gly Pro Leu Gln val Thr Ala Glu Glu Gly Gly Phe Thr Cys Tyr
 Gly Leu Ser Lys Gly Gly Phe Val Pro Trp Phe

<213> Pseudomonas aeruginosa

Met Asp His Tyr Leu Asp Ile Arg Leu Arg Pro Asp Pro Glu Phe Pro
 Pro Ala Gln Leu Met Ser Val Leu Phe Gly Lys Leu His Gln Ala Leu
 Val Ala Gln Gly Gly Asp Arg Ile Gly val Ser Phe Pro Asp Leu Asp
 Glu Ser Arg Ser Arg Leu Gly Glu Arg Leu Arg Ile His Ala Ser Ala
 Asp Asp Leu Arg Ala Leu Leu Ala Arg Pro Trp Leu Glu Gly Leu Arg
 Asp His Leu Gln Phe Gly Glu Pro Ala Val Val Pro His Pro Thr Pro
 Tyr Arg Gln Val Ser Arg Val Gln Ala Lys Ser Asn Pro Glu Arg Leu
 Arg Arg Arg Leu Met Arg Arg His Asp Leu Ser Glu Glu Glu Ala Arg
 Lys Arg Ile Pro Asp Thr Val Ala Arg Thr Leu Asp Leu Pro Phe Val
 Thr Leu Arg Ser Gln Ser Thr Gly Gln His Phe Arg Leu Phe Ile Arg
 His Gly Pro Leu Gln Ala Thr Ala Glu Glu Gly Gly Phe Thr Cys Tyr
 Gly Leu Ser Lys Gly Gly Phe Val Pro Trp Phe

<213> Pseudomonas aeruginosa

Met Asp His Tyr Leu Asp Ile Arg Leu Arg Pro Asp Pro Glu Phe Pro
 Pro Ala Gln Leu Met Ser Val Leu Phe Gly Lys Leu His Gln Ala Leu
 Val Ala Gln Gly Gly Asp Arg Ile Gly Val Ser Phe Pro Asp Leu Asp
 Glu Ser Arg Ser Arg Leu Gly Glu Arg Leu Arg Ile His Ala Ser Ala
 Asp Asp Leu His Ala Leu Leu Ala Arg Pro Trp Leu Glu Gly Leu Arg
 Asp His Leu Gln Phe Gly Glu Ala Ala Val Val Pro His Pro Thr Pro
 Tyr Arg Gln Val Ser Arg Val Gln Ala Lys Ser Asn Pro Glu Arg Leu
 Arg Arg Arg Leu Met Arg Arg His Asp Leu Ser Glu Glu Glu Ala Arg
 Lys Arg Ile Pro Asp Thr Val Ala Arg Thr Leu Asp Leu Pro Phe Val
 Thr Leu Arg Ser Gln Ser Thr Gly Gln His Phe Arg Leu Phe Ile Arg
 His Gly Pro Leu Gln Ala Thr Ala Glu Glu Gly Gly Phe Thr Cys Tyr
 Gly Leu Ser Lys Gly Gly Phe val Pro Trp Phe

FIG. 40

61 / 73

Cas6 sequences

Thermus thermophilus HB8

Met Val Leu Ala Ala Leu Val Leu Val Leu Glu Gly Glu Gly Leu Pro
 Glu Pro Leu Gly Leu Arg Gly Phe Phe Tyr Gly Leu Leu Arg Glu Val
 Ala Pro Glu Val His Asp Gln Gly Glu Asn Pro Phe Ala Leu Gly Phe
 Gly Gly Arg Glu Gly Ala Ala Trp Ala Arg Val Ser Leu Leu Val Glu
 Gly Leu Tyr Ala Arg Leu Ala Pro Arg Leu Tyr Ala Leu Glu Gly Glu
 Glu Val Arg Leu Gly Pro Pro Phe Arg Val Arg Ala Val Leu Gln Glu
 Gly His Pro Trp Ala Gly Val Ser Thr Tyr Pro Arg Leu Phe Gln Gly
 Pro Pro Ser Arg Asp Leu Ala Leu Arg Phe Ala Ser Pro Thr Phe Phe
 Arg Arg Lys Gly Val His Tyr Pro Val Pro Glu Pro Arg Leu Val Leu
 Glu Ser Leu Leu Arg Arg Leu Glu Ala Phe Gly Pro Leu Lys Ala Pro
 Glu Gly Val Arg Glu Ala Leu Leu Glu Arg Thr Thr Val Arg Ser Leu
 Glu Gly Arg Thr Leu Pro Ala Arg Thr Glu Val Asp Thr Ala Gly Phe
 Val Gly Arg Val Val Tyr His Leu Pro Arg Ala Thr Glu Glu Glu Ala
 Leu Trp Leu Ser Ala Leu Gly Arg Phe Ala Phe Tyr Ser Gly Val Gly
 Ala Lys Thr Ser Leu Gly Tyr Gly Arg Ala Arg Ala Glu Ser Ala

Thermus thermophilus HB8

Met Pro Gln Ala Val Val Leu Glu Leu Val Gly Glu Lys Pro Pro Leu
 Tyr Pro Ala Arg Tyr Ala His Gly Leu Phe Phe Ala Leu Leu Ser Arg
 Val Ser Pro Glu Leu Ala Gln Lys Leu His Glu Ala Pro Arg Lys Pro
 Phe Thr Leu Ala Pro Leu Pro Arg Ala Gly Pro Glu Gly Ala Thr Leu
 Lys Gly Thr Leu Arg Leu Arg Leu Thr Thr Leu Asp Asp Gly Leu Phe
 Ala Pro Phe Leu Arg Ala Leu Leu Glu Ala Ala Pro Asp Gly Leu Pro
 Leu Gly Asp Ser Ser Tyr Arg Leu Ala Arg Val Leu Ala Thr Arg Glu
 Gly His Pro Leu Ala Gly Ala Thr Ser Trp Glu Glu Leu Lys Glu Ala
 Pro Lys Arg Glu Lys Ala Thr Phe Arg Phe Leu Thr Pro Thr Val Phe
 Ala Thr Ser Lys Pro Gly Gly Arg Thr Arg Tyr Thr Pro Leu Pro Asp
 Pro Arg Leu Ile Ala Gly Ser Leu Leu Asp Lys Trp Gln Ala His Ser
 Pro Phe Pro Tyr Asn Pro Lys Glu Glu Ala Ala Leu Arg Glu Leu Phe
 Glu Leu Asp Leu Glu Val Ala Gly Phe Arg Asn Leu Arg Phe His Arg
 Val Gln Ala Gly Lys Gly Phe Phe Pro Gly Phe Thr Gly Glu Ala Thr
 Leu Arg Leu Trp Ser Gln Ser Leu Glu Ala Gln Glu Ala Leu Gly Arg
 Leu His Ala Leu Ala Phe Phe Ser Gly Val Gly Ala Lys Thr Pro Tyr
 Gly Met Gly Leu Ala Val Pro Leu

FIG. 41A

62 / 73

Cas6 sequences

Staphylococcus epidermidis RP62A

Met Ile Asn Lys Ile Thr Val Glu Leu Asp Leu Pro Glu Ser Ile Arg
Phe Gln Tyr Leu Gly Ser Val Leu His Gly Val Leu Met Asp Tyr Leu
Ser Asp Asp Ile Ala Asp Gln Leu His His Glu Phe Ala Tyr Ser Pro
Leu Lys Gln Arg Ile Tyr His Lys Asn Lys Lys Ile Ile Trp Glu Ile
Val Cys Met Ser Asp Asn Leu Phe Lys Glu Val Val Lys Leu Phe Ser
Ser Lys Asn Ser Leu Leu Leu Lys Tyr Tyr Gln Thr Asn Ile Asp Ile
Gln Ser Phe Gln Ile Glu Lys Ile Asn Val Gln Asn Met Met Asn Gln
Leu Leu Gln Val Glu Asp Leu Ser Arg Tyr Val Arg Leu Asn Ile Gln
Thr Pro Met Ser Phe Lys Tyr Gln Asn Ser Tyr Met Ile Phe Pro Asp
Val Lys Arg Phe Phe Arg Ser Ile Met Ile Gln Phe Asp Ala Phe Phe
Glu Glu Tyr Arg Met Tyr Asp Lys Glu Thr Leu Asn Phe Leu Glu Lys
Asn Val Asn Ile Val Asp Tyr Lys Leu Lys Ser Thr Arg Phe Asn Leu
Glu Lys Val Lys Ile Pro Ser Phe Thr Gly Glu Ile Val Phe Lys Ile
Lys Gly Pro Leu Pro Phe Leu Gln Leu Thr His Phe Leu Leu Lys Phe
Gly Glu Phe Ser Gly Ser Gly Ile Lys Thr Ser Leu Gly Met Gly Lys
Tyr Ser Ile Ile

Mycobacterium tuberculosis H37Rv

Met Ala Ala Arg Arg Gly Gly Ile Arg Arg Thr Asp Leu Leu Arg Arg
Ser Gly Gln Pro Arg Gly Arg His Arg Ala Ser Ala Ala Glu Ser Gly
Leu Thr Trp Ile Ser Pro Thr Leu Ile Leu Val Gly Phe Ser His Arg
Gly Asp Arg Arg Met Thr Glu His Leu Ser Arg Leu Thr Leu Thr Leu
Glu Val Asp Ala Pro Leu Glu Arg Ala Arg Val Ala Thr Leu Gly Pro
His Leu His Gly Val Leu Met Glu Ser Ile Pro Ala Asp Tyr Val Gln
Thr Leu His Thr Val Pro Val Asn Pro Tyr Ser Gln Tyr Ala Leu Ala
Arg Ser Thr Thr Ser Leu Glu Trp Lys Ile Ser Thr Leu Thr Asn Glu
Ala Arg Gln Gln Ile Val Gly Pro Ile Asn Asp Ala Ala Phe Ala Gly
Phe Arg Leu Arg Ala Ser Gly Ile Ala Thr Gln Val Thr Ser Arg Ser
Leu Glu Gln Asn Pro Leu Ser Gln Phe Ala Arg Ile Phe Tyr Ala Arg
Pro Glu Thr Arg Lys Phe Arg Val Glu Phe Leu Thr Pro Thr Ala Phe
Lys Gln Ser Gly Glu Tyr Val Phe Trp Pro Asp Pro Arg Leu Val Phe
Gln Ser Leu Ala Gln Lys Tyr Gly Ala Ile Val Asp Gly Glu Glu Pro
Asp Pro Gly Leu Ile Ala Glu Phe Gly Gln Ser Val Arg Leu Ser Ala
Phe Arg Val Ala Ser Ala Pro Phe Ala Val Gly Ala Ala Arg Val Pro
Gly Phe Thr Gly Ser Ala Thr Phe Thr Val Arg Gly Val Asp Thr Phe
Ala Ser Tyr Ile Ala Ala Leu Leu Trp Phe Gly Glu Phe Ser Gly Cys
Gly Ile Lys Ala Ser Met Gly Met Gly Ala Ile Arg Val Gln Pro Leu
Ala Pro Arg Glu Lys Cys Val Pro Lys Pro

FIG. 41B

63 / 73

Cas6 sequences

Streptococcus thermophilus LMG 18311

Met Lys Lys Leu Val Phe Thr Phe Lys Arg Ile Asp His Pro Ala Gln
 Asp Leu Ala Val Lys Phe His Gly Phe Leu Met Glu Gln Leu Asp Ser
 Asp Tyr Val Asp Tyr Leu His Gln Gln Gln Thr Asn Pro Tyr Ala Thr
 Lys Val Ile Gln Gly Lys Glu Asn Thr Gln Trp Val Val His Leu Leu
 Thr Asp Asp Ile Glu Asp Lys Val Phe Met Thr Leu Leu Gln Ile Lys
 Glu Val Ser Leu Asn Asp Leu Pro Lys Leu Ser Val Glu Lys Val Glu
 Ile Gln Glu Leu Gly Ala Asp Lys Leu Leu Glu Ile Phe Asn Ser Glu
 Glu Asn Gln Thr Tyr Phe Ser Ile Ile Phe Glu Thr Pro Thr Gly Phe
 Lys Ser Gln Gly Ser Tyr Val Ile Phe Pro Ser Met Arg Leu Ile Phe
 Gln Ser Leu Met Gln Lys Tyr Gly Arg Leu Val Glu Asn Gln Pro Glu
 Ile Glu Glu Asp Thr Leu Asp Tyr Leu Ser Glu His Ser Thr Ile Thr
 Asn Tyr Arg Leu Glu Thr Ser Tyr Phe Arg Val His Arg Gln Arg Ile
 Pro Ala Phe Arg Gly Lys Leu Thr Phe Lys Val Gln Gly Ala Gln Thr
 Leu Lys Ala Tyr Val Lys Met Leu Leu Thr Phe Gly Glu Tyr Ser Gly
 Leu Gly Met Lys Thr Ser Leu Gly Met Gly Gly Ile Lys Leu Glu Glu
 Arg Lys Asp

Streptococcus sanguinis SK36

Met Lys Lys Ile Arg Leu His Leu Ser Lys Val Ser Leu Lys Asp Asp
 Asp Leu Val Cys Lys Leu Gln Gly Phe Leu Met Glu Lys Leu Ser Asp
 Asp Phe Ala Ser Phe Leu His Gln Gln Glu Thr Asn Pro Tyr Ser Met
 Asn Leu Arg Ser Glu Arg Glu Glu Ser Ile Trp Thr Val Asn Leu Leu
 Ser Glu Glu Ala Glu Gln Gln Ile Leu Pro Gln Leu Leu Ser Leu Glu
 Met Ile Lys Leu Glu Thr Tyr Ser Glu Glu Ile Leu Val Lys Asn Ile
 Glu Ile Gln Ser Leu Ser Ser Gln Ser Leu Leu Glu Val Phe Gln Gly
 Asp Glu Ala Ser His Leu Ile Ser Leu Asn Phe Tyr Thr Pro Thr Thr
 Phe Lys Arg Gln Gly Gln Phe Val Leu Phe Pro Asp Thr Arg Leu Ile
 Phe Gln Ser Leu Met Gln Lys Tyr Ser Arg Leu Val Glu Gly Lys Ala
 Glu Ile Glu Glu Glu Thr Leu Glu Phe Leu Ala Glu His Ser Gln Ile
 Ser Ser Tyr Arg Leu Lys Ser His Tyr Phe Pro Ile His Gly Arg Lys
 Tyr Pro Ala Phe Glu Gly Arg Val Thr Ile Arg Ile Gln Gly Ala Ser
 Thr Leu Lys Ala Tyr Ala Gln Met Leu Leu Arg Phe Gly Glu Tyr Ser
 Gly Val Gly Ala Lys Cys Ser Leu Gly Met Gly Gly Met Arg Ile Glu
 Glu Arg Lys Thr

FIG. 41C

64 / 73

Cas6 sequences

Microcystis aeruginosa NIES-843

Met Pro Tyr Ser Leu Val Leu Asn Leu Thr Pro Arg Ser Pro Ile Tyr
 Pro Asn Phe Leu Thr Gly Arg His Leu His Ala Leu Phe Leu Thr Leu
 Val Ser Ser Val Asp Gln Glu Leu Gly Lys Ile Leu His Thr Ala Glu
 Ala Asp Lys Ala Phe Thr Leu Ser Pro Leu Gln Met Gln Ser Gly Gly
 Lys Thr Ile Asn Ser Pro Gln Trp Arg Tyr Glu Arg Pro Ile Ala Pro
 Glu Thr Pro Cys Trp Trp Arg Ile Ser Leu Leu Asp Asp Arg Leu Phe
 Gly Lys Leu Thr Pro Leu Trp Leu Asn Leu Asn Pro Lys His Pro Trp
 His Leu Gly Ser Ala Asp Leu Val Ile Thr Ser Val Leu Ala Thr Pro
 Gln Ser Val Gln Pro Trp Ala Asn Ser Cys Thr Tyr Gln Tyr Leu Tyr
 Glu Asn Ala Ser Glu Thr Asn Arg Glu Phe Asp Phe Leu Phe Ala Thr
 Pro Val Thr Phe Arg Gln Gly Lys Phe Asp Ser Ala Leu Pro Thr Arg
 Glu Leu Val Phe Asn Ser Leu Leu Asn Arg Trp Asn Arg Tyr Ser Ala
 Ile Pro Phe Asp Ser Ile Val Leu Glu Ser Ile Phe Pro Ser Phe Phe
 Asp Ile Gln Thr Lys Leu Ala Asp Glu Ala Tyr Lys Asn Gln Ser Phe
 Gly Cys Val Gly Glu Ile His Tyr Arg Leu Leu Gly Glu Val Glu Pro
 Ala Lys Ile Lys Ala Ile Asn Val Leu Ala Asp Phe Ala Leu Tyr Ala
 Gly Val Gly Arg Lys Thr Thr Met Gly Met Gly Met Thr Arg Arg Ile
 Ala Lys Glu Lys Arg

Meiothermus silvanus DSM 9946

Met Met Leu Ala Ala Leu Val Leu Pro Leu Glu Gly Gln Ala Arg Pro
 Asp Pro Asp Gly Trp Arg Gly Leu Val Tyr Gly Leu Leu Lys Glu Ile
 Asp Pro Glu Leu His Thr Ala Gln His Asn Pro Phe Ser Leu Gly Leu
 Gly Gly Ala Glu Gly Gln Trp Trp Val Arg Ile Ala Leu Leu Glu Glu
 Gly Leu Tyr Ala Arg Leu Ser Pro His Leu Phe Gly Leu Val Gly Gln
 Ser Val Lys Leu Lys Glu Pro Phe Arg Val Arg Ala Val Leu Gln Glu
 Glu His Pro Trp Ala Ser Leu Ser Thr Tyr Pro Arg Leu Phe Gln Gly
 Gln Ala Ser Pro Ser Leu Gly Leu Gln Phe Ala Ser Pro Thr Phe Phe
 Arg Arg Lys Gly Asn Ser Tyr Pro Leu Pro Glu Pro Lys Leu Val Phe
 Asp Ser Leu Thr Gln Arg Trp Asn Ala Phe Ala Pro Val Lys Val Pro
 Pro Glu Met Ala Glu Thr Trp Glu Arg Val Thr Ile Thr Arg Leu Gln
 Gly His Thr Gln Ala Ile Arg Pro Asn Pro Asp Glu Arg Gly Val Gly
 Phe Val Gly Arg Val Val Tyr His Leu Pro Ala Ala Lys Pro Thr Glu
 Ala Gln Trp Met Gln Ala Leu Gly Arg Phe Ala Phe Tyr Ala Gly Val
 Gly Ala Lys Thr Ser Leu Gly Phe Gly Arg Val Arg Gly Phe Asp Pro
 Ile Leu Lys Glu Glu Ser Ala Asn Gly Arg Leu Asp Ala Glu Asp Ser
 Ser Ser Leu Ala Thr Pro Gln Asp Pro Gly Ala

FIG. 41D

65 / 73

Cas6 sequences

Meiothermus silvanus DSM 9946

Met Met Leu Ala Ala Leu Val Leu Pro Leu Glu Gly Gln Ala Arg Pro
 Asp Pro Asp Gly Trp Arg Gly Leu Val Tyr Gly Leu Leu Lys Glu Ile
 Asp Pro Glu Leu His Thr Ala Gln His Asn Pro Phe Ser Leu Gly Leu
 Gly Gly Ala Glu Gly Gln Trp Trp Val Arg Ile Ala Leu Leu Glu Glu
 Gly Leu Tyr Ala Arg Leu Ser Pro His Leu Phe Gly Leu Val Gly Gln
 Ser Val Lys Leu Lys Glu Pro Phe Arg Val Arg Ala Val Leu Gln Glu
 Glu His Pro Trp Ala Ser Leu Ser Thr Tyr Pro Arg Leu Phe Gln Gly
 Gln Ala Ser Pro Ser Leu Gly Leu Gln Phe Ala Ser Pro Thr Phe Phe
 Arg Arg Lys Gly Asn Ser Tyr Pro Leu Pro Glu Pro Lys Leu Val Phe
 Asp Ser Leu Thr Gln Arg Trp Asn Ala Phe Ala Pro Val Lys Val Pro
 Pro Glu Met Ala Glu Thr Trp Glu Arg Val Thr Ile Thr Arg Leu Gln
 Gly His Thr Gln Ala Ile Arg Pro Asn Pro Asp Glu Arg Gly Val Gly
 Phe Val Gly Arg Val Val Tyr His Leu Pro Ala Ala Lys Pro Thr Glu
 Ala Gln Trp Met Gln Ala Leu Gly Arg Phe Ala Phe Tyr Ala Gly Val
 Gly Ala Lys Thr Ser Leu Gly Phe Gly Arg Val Arg Gly Phe Asp Pro
 Ile Leu Lys Glu Glu Ser Ala Asn Gly Arg Leu Asp Ala Glu Asp Ser
 Ser Ser Leu Ala Thr Pro Gln Asp Pro Gly Ala

Meiothermus ruber DSM 1279

Met Ile Leu Ala Ala Leu Ile Leu Pro Leu Glu Gly Pro Thr Arg Pro
 Asp Pro Asp Gly Trp Arg Gly Leu Val Tyr Gly Leu Leu Lys Glu Ile
 Asp Pro Glu Leu His Ala Ala Gln His Asn Pro Phe Ser Leu Gly Leu
 Gly Gly Ala Leu Gly Gln Trp Trp Val Arg Ile Ala Phe Leu Glu Glu
 Gly Leu Tyr Ala Arg Leu Ser Pro His Leu Phe Gly Leu Ala Gly Gln
 Thr Val Arg Leu Lys Glu Ala Phe Gln Val Arg Ala Val Leu Gln Glu
 Ala His Pro Trp Ala Gly Val Ser Thr Tyr Pro Lys Leu Phe Gln Gly
 Gln Ala Thr Ala Ser Leu Gly Leu Gln Phe Ala Ser Pro Thr Phe Phe
 Arg Arg Lys Gly His Ser Tyr Pro Leu Pro Glu Pro Arg Leu Val Phe
 Glu Ser Leu Thr Gln Arg Trp Asn Ala Phe Ala Pro Val Lys Val Pro
 Gln Glu Val Gln Glu Ala Trp Glu Arg Leu Leu Val Gly Gln Phe Gln
 Gly Arg Thr His His Ile Ala Pro Asn Gln Asp Glu Arg Gly Val Gly
 Phe Val Gly Arg Val Val Tyr Tyr Leu Pro Lys Ala Ser Pro Thr Glu
 Ala Gln Trp Leu Gln Ala Leu Gly Arg Phe Ala Phe Tyr Ala Gly Val
 Gly Ala Lys Thr Ser Leu Gly Phe Gly Arg Val Arg Met Phe Asp Pro
 Leu Gln Gln Glu Arg Arg Pro Asp Glu Ser Glu Gln Gly Ala Leu Thr
 Gly Thr Val Gly Gly Val

FIG. 41E

Cas6 sequences

Thermus thermophilus HB8

Met Val Leu Ala Ala Leu Val Leu Val Leu Glu Gly Glu Gly Leu Pro
 Glu Pro Leu Gly Leu Arg Gly Phe Phe Tyr Gly Leu Leu Arg Glu Val
 Ala Pro Glu Val His Asp Gln Gly Glu Asn Pro Phe Ala Leu Gly Phe
 Gly Gly Arg Glu Gly Ala Ala Trp Ala Arg Val Ser Leu Leu Val Glu
 Gly Leu Tyr Ala Arg Leu Ala Pro Arg Leu Tyr Ala Leu Glu Gly Glu
 Glu Val Arg Leu Gly Pro Pro Phe Arg Val Arg Ala Val Leu Gln Glu
 Gly His Pro Trp Ala Gly Val Ser Thr Tyr Pro Arg Leu Phe Gln Gly
 Pro Pro Ser Arg Asp Leu Ala Leu Arg Phe Ala Ser Pro Thr Phe Phe
 Arg Arg Lys Gly Val His Tyr Pro Val Pro Glu Pro Arg Leu Val Leu
 Glu Ser Leu Leu Arg Arg Leu Glu Ala Phe Gly Pro Leu Lys Ala Pro
 Glu Gly Val Arg Glu Ala Leu Leu Glu Arg Thr Thr Val Arg Ser Leu
 Glu Gly Arg Thr Leu Pro Ala Arg Thr Glu Val Asp Thr Ala Gly Phe
 Val Gly Arg Val Val Tyr His Leu Pro Arg Ala Thr Glu Glu Glu Ala
 Leu Trp Leu Ser Ala Leu Gly Arg Phe Ala Phe Tyr Ser Gly Val Gly
 Ala Lys Thr Ser Leu Gly Tyr Gly Arg Ala Arg Ala Glu Ser Ala

Thermus thermophilus JL-18

Met Val Leu Ala Ala Leu Val Leu Val Leu Glu Gly Glu Gly Leu Pro
 Glu Pro Leu Gly Leu Arg Gly Phe Phe Tyr Gly Leu Leu Arg Glu Val
 Ala Pro Glu Val His Asp Gln Gly Glu Asn Pro Phe Ala Leu Gly Phe
 Gly Gly Arg Glu Gly Ala Ser Trp Ala Arg Val Ser Leu Leu Arg Glu
 Glu Leu Tyr Ala Arg Leu Ala Pro Arg Leu Tyr Ala Leu Glu Gly Glu
 Glu Val Arg Leu Gly Pro Pro Phe Arg Val Arg Ala Val Leu Gln Glu
 Gly His Pro Trp Ala Gly Val Ser Thr Tyr Pro Arg Leu Phe Gln Gly
 Pro Pro Ser Arg Asp Leu Ala Leu Arg Phe Ala Ser Pro Thr Phe Phe
 Arg Arg Lys Gly Val His Tyr Pro Val Pro Glu Pro Arg Leu Val Leu
 Glu Ser Leu Leu Arg Arg Leu Glu Ala Phe Gly Pro Leu Lys Ala Pro
 Glu Gly Val Arg Glu Ala Leu Leu Glu Arg Thr Thr Val Arg Ser Leu
 Glu Gly Arg Thr Leu Pro Ala Arg Thr Glu Val Asp Thr Ala Gly Phe
 Val Gly Arg Val Val Tyr His Leu Pro Arg Ala Thr Glu Glu Glu Ala
 Leu Trp Leu Ser Ala Leu Gly Arg Phe Ala Phe Tyr Ser Gly Val Gly
 Ala Lys Thr Ser Leu Gly Tyr Gly Arg Ala Arg Ala Glu Ser Pro

FIG. 41F

67 / 73

Cas6 sequences

Thermus sp. CCB US3 UF1

Met Leu Ala Ala Leu Val Leu Thr Leu Glu Gly Glu Ala Pro Pro Glu
 Pro Arg Gly Leu Arg Gly Phe Phe Tyr Gly Leu Leu Gln Glu Val Ala
 Pro Glu Val His Asp Gln Gly Glu Asn Pro Phe Ala Leu Gly Phe Gly
 Gly Lys Glu Gly Ala Tyr Trp Ala Arg Phe Ser Leu Leu Gln Glu Gly
 Leu Tyr Ala Arg Leu Ala Pro Arg Leu Phe Ala Leu Glu Gly Lys Glu
 Val Arg Leu Gly Lys Pro Phe Arg Val Arg Gly Val Leu Gln Glu Gly
 His Pro Trp Ala Gly Val Ser Thr Tyr Ala Arg Leu Phe Gln Gly Glu
 Ala Leu Pro Asp Leu Pro Leu Arg Phe Ala Ser Pro Thr Phe Phe Arg
 Arg Lys Gly Val His Tyr Pro Leu Pro Glu Pro Arg Leu Val Val Glu
 Ser Leu Leu Arg Arg Leu Glu Ala Phe Gly Pro Leu Lys Ala Pro Glu
 Gly Val Arg Glu Ala Leu Leu Glu Arg Thr Thr Val Arg Trp Phe Glu
 Gly Lys Thr Leu Lys Ala Glu Thr Glu Val Glu Ala Val Gly Phe Val
 Gly Lys Val Val Tyr His Leu Pro Arg Ala Thr Glu Glu Glu Ala Arg
 Trp Leu Gln Ala Leu Gly Arg Phe Ala Phe Tyr Ser Gly Val Gly Ala
 Lys Thr Gly Leu Gly Tyr Gly Arg Ala Arg Val Gly

Thermus aquaticus Y51MC23

Met Val Leu Val Ala Leu Val Leu Val Leu Glu Gly Glu Gly Pro Pro
 Glu Pro Leu Gly Leu Arg Gly Phe Phe Tyr Thr Leu Leu Lys Glu Ala
 Phe Pro Glu Leu His Asp Gln Gly Glu Asn Pro Phe Ala Leu Gly Phe
 Gly Leu Arg Gly Gly Glu Pro Trp Ala Arg Val Ser Leu Leu Arg Glu
 Asp Leu Tyr Gly Arg Leu Ser Pro Ala Leu Phe Gly Leu Glu Gly Arg
 Glu Val Arg Leu Gly Arg Leu Phe Arg Val Arg Ala Val Leu Gln Glu
 Gly His Pro Trp Ala Gly Leu Thr Thr Tyr Ala Arg Leu Phe Gln Gly
 Pro His Ser Pro Asn Leu Pro Leu Arg Phe Tyr Ser Pro Thr Phe Phe
 Arg Arg Lys Gly Val Gln Tyr Pro Leu Pro Glu Pro Arg Leu Val Leu
 Glu Ser Leu Leu Arg Arg Leu Glu Ala Phe Gly Pro Leu Lys Ala Pro
 Gln Glu Val Arg Glu Ala Leu Leu Glu Arg Thr Thr Val Arg Phe Leu
 Glu Gly Arg Thr Gln Met Ala Arg Thr Glu Val Asp Thr Val Gly Phe
 Val Gly Lys Val Val Tyr His Leu Pro Lys Ala Thr Glu Glu Glu Ala
 Leu Trp Leu Ser Ala Leu Gly Arg Tyr Ala Phe Phe Ser Gly Val Gly
 Ala Lys Thr Ser Leu Gly Tyr Gly Leu Ala Arg Ala Phe Thr Gln Val
 Gly Pro Gln Asp Ala Glu Thr

FIG. 41G

68 / 73

Cas6 sequences

Marinithermus hydrothermalis DSM 14884

Met Leu Leu Ala Ala Leu Val Leu Pro Leu Glu Gly Pro Asp Arg Pro
 Gln Pro Leu His Ala Arg Gly Trp Val Tyr Arg Leu Leu Arg Glu Ala
 Ala Pro Glu Ile His Asp Ala Glu Gly Pro Lys Pro Phe Thr Val Gly
 Val Gly Gly Arg Pro Asn Ala Val Trp Val Arg Leu Thr Cys Leu Ala
 Glu Glu Val Tyr Ala Ala Leu Ser Pro Arg Leu Trp Ser Gln Val Gly
 Leu Glu Val Arg Leu Gly Glu Asp Thr Tyr Arg Ile Lys Ala Val Leu
 Glu Ala Glu His Pro Trp Ala Gly Leu Ala Thr Trp Pro Arg Leu Phe
 Gln Gly Glu Ala Gly Pro Asp Leu Gly Leu Glu Phe Ala Ser Pro Thr
 Phe Phe Arg Arg Gln Gly Ala Asn Tyr Pro Leu Pro Glu Pro Arg Leu
 Val Leu Gly Ser Leu Ile Glu Arg Trp Asn Ala His Ala Pro Thr Pro
 Val Pro Pro Glu Val Ala Glu Arg Leu Val Glu Ala Thr Thr Leu Arg
 Tyr Leu Lys Gly His Thr Val Ser Ala Val Gly His Asp Arg Thr Val
 Gly Phe Arg Gly Arg Val Thr Tyr His Leu Pro Arg Ala Ser Thr Glu
 Glu Ala Arg Trp Leu Ala Ala Leu Gly Arg Phe Ala Phe Phe Ser Gly
 Val Gly Ala Lys Thr Thr Leu Gly Phe Gly Gln Val Arg Pro Tyr Pro
 Leu Leu Ala Pro Ser Ala Ala Pro Pro Gly Pro

Thermus thermophilus HB8

Met Pro Gln Ala Val Val Leu Glu Leu Val Gly Glu Lys Pro Pro Leu
 Tyr Pro Ala Arg Tyr Ala His Gly Leu Phe Phe Ala Leu Leu Ser Arg
 Val Ser Pro Glu Leu Ala Gln Lys Leu His Glu Ala Pro Arg Lys Pro
 Phe Thr Leu Ala Pro Leu Pro Arg Ala Gly Pro Glu Gly Ala Thr Leu
 Lys Gly Thr Leu Arg Leu Arg Leu Thr Thr Leu Asp Asp Gly Leu Phe
 Ala Pro Phe Leu Arg Ala Leu Leu Glu Ala Ala Pro Asp Gly Leu Pro
 Leu Gly Asp Ser Ser Tyr Arg Leu Ala Arg Val Leu Ala Thr Arg Glu
 Gly His Pro Leu Ala Gly Ala Thr Ser Trp Glu Glu Leu Lys Glu Ala
 Pro Lys Arg Glu Lys Ala Thr Phe Arg Phe Leu Thr Pro Thr Val Phe
 Ala Thr Ser Lys Pro Gly Gly Arg Thr Arg Tyr Thr Pro Leu Pro Asp
 Pro Arg Leu Ile Ala Gly Ser Leu Leu Asp Lys Trp Gln Ala His Ser
 Pro Phe Pro Tyr Asn Pro Lys Glu Glu Ala Ala Leu Arg Glu Leu Phe
 Glu Leu Asp Leu Glu Val Ala Gly Phe Arg Asn Leu Arg Phe His Arg
 Val Gln Ala Gly Lys Gly Phe Phe Pro Gly Phe Thr Gly Glu Ala Thr
 Leu Arg Leu Trp Ser Gln Ser Leu Glu Ala Gln Glu Ala Leu Gly Arg
 Leu His Ala Leu Ala Phe Phe Ser Gly Val Gly Ala Lys Thr Pro Tyr
 Gly Met Gly Leu Ala Val Pro Leu

FIG. 41H

69 / 73

Cas6 sequences

Thermus thermophilus HB27

Met Pro Gln Ala Val Val Leu Glu Leu Val Gly Glu Lys Pro Pro Leu
 Tyr Pro Ala Arg Tyr Ala His Gly Leu Phe Phe Ala Leu Leu Ser Arg
 Val Ser Pro Glu Leu Ala Gln Lys Leu His Glu Ala Pro Arg Lys Pro
 Phe Thr Leu Ala Pro Leu Pro Arg Ala Gly Pro Glu Gly Ala Thr Leu
 Lys Gly Thr Leu Arg Leu Arg Leu Thr Thr Leu Asp Asp Gly Leu Phe
 Ala Pro Phe Leu Arg Ala Leu Leu Glu Ala Ala Pro Asp Gly Leu Pro
 Leu Gly Asp Ser Ser Tyr Arg Leu Ala Arg Val Leu Ala Thr Arg Glu
 Gly His Pro Leu Ala Gly Ala Thr Ser Trp Glu Glu Leu Lys Glu Ala
 Pro Lys Arg Glu Lys Val Thr Phe Arg Phe Leu Thr Pro Thr Val Phe
 Ala Thr Ser Lys Pro Gly Gly Arg Thr Arg Tyr Thr Pro Leu Pro Asp
 Pro Arg Leu Ile Ala Gly Ser Leu Leu Asp Lys Trp Gln Ala His Ser
 Pro Phe Pro Tyr Asn Pro Lys Glu Glu Ala Ala Leu Arg Gly Leu Phe
 Glu Leu Asp Leu Glu Val Ala Gly Phe Arg Asn Leu Arg Phe His Arg
 Val Gln Ala Gly Lys Gly Phe Phe Pro Gly Phe Thr Gly Glu Met Thr
 Leu Arg Leu Trp Ser Gln Ser Leu Glu Ala Arg Glu Ala Leu Gly Arg
 Leu His Ala Leu Ala Phe Phe Ser Gly Val Gly Ala Lys Thr Pro Tyr
 Gly Met Gly Leu Ala Val Pro Leu

Thermus thermophilus SG0.5JP17-16

Met Pro Gln Ala Val Val Leu Glu Leu Val Gly Glu Lys Pro Pro Leu
 Tyr Pro Gly Arg Tyr Ala His Gly Leu Phe Phe Ala Leu Leu Ser Arg
 Val Ser Pro Glu Leu Ala Gln Lys Leu His Glu Ala Pro Arg Lys Pro
 Phe Thr Leu Ala Pro Leu Pro Arg Val Gly Pro Glu Gly Ala Thr Leu
 Lys Gly Ile Leu Arg Leu Arg Leu Thr Ala Leu Asp Asp Gly Leu Phe
 Ala Pro Phe Leu Arg Ala Leu Leu Glu Ala Ala Pro Asp Gly Leu Pro
 Leu Gly Asp Ser Ser Tyr Arg Leu Ala Arg Val Leu Ala Thr Arg Glu
 Gly His Pro Leu Ala Gly Ala Thr Ser Trp Glu Glu Leu Lys Glu Ala
 Pro Lys Arg Glu Lys Ala Thr Phe Arg Phe Leu Thr Pro Thr Val Phe
 Ala Thr Ser Lys Pro Gly Gly Arg Thr Arg Tyr Thr Pro Leu Pro Asp
 Pro Arg Leu Ile Ala Gly Ser Leu Leu Asp Lys Trp Gln Ala His Ser
 Pro Phe Pro Tyr Asn Pro Lys Glu Glu Ala Ala Leu Arg Glu Leu Phe
 Glu Leu Asp Leu Glu Val Ala Gly Phe Arg Asn Leu Arg Phe His Arg
 Val Gln Ala Gly Lys Gly Phe Phe Pro Gly Phe Thr Gly Glu Ala Thr
 Leu Arg Leu Trp Ser Gln Ser Leu Glu Ala Gln Glu Ala Leu Gly Arg
 Leu His Ala Leu Ala Phe Phe Ser Gly Val Gly Ala Lys Thr Pro Tyr
 Gly Met Gly Leu Ala Val Pro Leu

FIG. 41I

70 / 73

Cas6 sequences

Thermus thermophilus JL-18

Met Pro Gln Ala Val Val Leu Glu Leu Val Gly Glu Glu Ser Pro Leu
Tyr Pro Ala Arg Tyr Ala His Gly Leu Phe Phe Ala Leu Leu Ser Arg
Val Ser Pro Glu Leu Ala Gln Lys Leu His Glu Ala Pro Arg Lys Pro
Phe Thr Leu Ala Pro Leu Pro Arg Val Gly Ser Glu Gly Ala Thr Leu
Lys Gly Ile Leu Arg Leu Arg Leu Thr Ile Leu Asp Asp Gly Leu Phe
Ala Pro Phe Leu Arg Ala Leu Leu Glu Ala Ala Pro Asp Gly Leu Pro
Leu Gly Asp Ser Ser Tyr Arg Leu Ala Arg Val Leu Ala Thr Arg Glu
Gly His Pro Leu Ala Gly Ala Thr Ser Trp Glu Glu Leu Lys Glu Ala
Pro Lys Arg Glu Lys Ala Thr Phe Arg Phe Leu Thr Pro Thr Val Phe
Ala Thr Ser Lys Pro Gly Gly Arg Thr Arg Tyr Thr Pro Leu Pro Asp
Pro Arg Leu Ile Ala Gly Ser Leu Leu Asp Lys Trp Gln Ala His Ser
Pro Phe Pro Tyr Asn Pro Lys Glu Glu Ala Ala Leu Arg Glu Leu Phe
Glu Leu Asp Leu Glu Val Ala Gly Phe Arg Asn Leu Arg Phe His Arg
Val Gln Ala Gly Lys Ser Phe Phe Pro Gly Phe Thr Gly Glu Met Thr
Leu Arg Leu Trp Ser Gln Ser Leu Glu Ala Gln Gly Ala Leu Gly Arg
Leu His Ala Leu Ala Phe Phe Ser Gly Val Gly Ala Lys Thr Pro Tyr
Gly Met Gly Leu Ala Val Pro Leu

FIG. 41J

71 / 73

Cas5 sequences

Desulfovibrio vulgaris str. Hildenborough

Met Thr His Gly Ala Val Lys Thr Tyr Gly Ile Arg Leu Arg Val Trp
 Gly Asp Tyr Ala Cys Phe Thr Arg Pro Glu Met Lys Val Glu Arg Val
 Ser Tyr Asp Val Met Pro Pro Ser Ala Ala Arg Gly Ile Leu Glu Ala
 Ile His Trp Lys Pro Ala Ile Arg Trp Ile Val Asp Arg Ile His Val
 Leu Arg Pro Ile Val Phe Asp Asn Val Arg Arg Asn Glu Val Ser Ser
 Lys Ile Pro Lys Pro Asn Pro Ala Thr Ala Met Arg Asp Arg Lys Pro
 Leu Tyr Phe Leu Val Asp Asp Gly Ser Asn Arg Gln Gln Arg Ala Ala
 Thr Leu Leu Arg Asn Val Asp Tyr Val Ile Glu Ala His Phe Glu Leu
 Thr Asp Lys Ala Gly Ala Glu Asp Asn Ala Gly Lys His Leu Asp Ile
 Phe Arg Arg Arg Ala Arg Ala Gly Gln Ser Phe Gln Gln Pro Cys Leu
 Gly Cys Arg Glu Phe Pro Ala Ser Phe Glu Leu Leu Glu Gly Asp Val
 Pro Leu Ser Cys Tyr Ala Gly Glu Lys Arg Asp Leu Gly Tyr Met Leu
 Leu Asp Ile Asp Phe Glu Arg Asp Met Thr Pro Leu Phe Phe Lys Ala
 Val Met Glu Asp Gly Val Ile Thr Pro Pro Ser Arg Thr Ser Pro Glu
 Val Arg Ala

Neisseria mucosa ATCC 25996

Met Asn Gln Ile Arg Leu His Val Trp Gly Asp Tyr Ala Cys Phe Thr
 Arg Pro Glu Met Lys Val Glu Arg Val Ser Tyr Asp Val Ile Thr Pro
 Ser Ala Ala Arg Gly Ile Leu Ala Ala Val His Trp Lys Pro Ala Ile
 Arg Trp Val Ile Asp Arg Ile Tyr Val Leu Lys Pro Ile Arg Phe Glu
 Ser Val Arg Arg Asn Glu Leu Gly Gly Lys Ile Ser Ala Gly Lys Val
 Ser Gly Ala Met Lys Arg Lys Ser Val Ala Asp Leu Tyr Thr Leu Ile
 Glu Asp Asp Arg Gln Gln Arg Ala Ala Thr Val Leu Lys Asp Val Ala
 Tyr Val Ile Glu Ala His Ala Val Leu Thr Ala Lys Ala Gly Ala Asp
 Glu Thr Val Thr Lys His Ile Glu Met Phe Lys Arg Arg Ala Lys Lys
 Gly Gln Cys Phe Gln Gln Pro Cys Leu Gly Val Arg Glu Phe Pro Ala
 Asp Phe Ala Leu Ile Asp Glu Gly Glu Pro Leu Pro Pro Ser Ala Leu
 Ser Glu Ser Glu Ala Asn Arg Asp Leu Gly Trp Met Leu His Asp Ile
 Asp Phe Asp His Gly Asn Thr Pro His Phe Phe Arg Ala Gln Met Lys
 Asp Gly Val Ile Asp Val Pro Pro Phe Tyr Ala Glu Glu Val Lys Ala

FIG. 42A

72 / 73

Cas5 sequences

Bacillus halodurans C-125

Met Arg Asn Glu Val Gln Phe Glu Leu Phe Gly Asp Tyr Ala Leu Phe
 Thr Asp Pro Leu Thr Lys Ile Gly Gly Glu Lys Leu Ser Tyr Ser Val
 Pro Thr Tyr Gln Ala Leu Lys Gly Ile Ala Glu Ser Ile Tyr Trp Lys
 Pro Thr Ile Val Phe Val Ile Asp Glu Leu Arg Val Met Lys Pro Ile
 Gln Met Glu Ser Lys Gly Val Arg Pro Ile Glu Tyr Gly Gly Gly Asn
 Thr Leu Ala His Tyr Thr Tyr Leu Lys Asp Val His Tyr Gln Val Lys
 Ala His Phe Glu Phe Asn Leu His Arg Pro Asp Leu Ala Phe Asp Arg
 Asn Glu Gly Lys His Tyr Ser Ile Leu Gln Arg Ser Leu Lys Ala Gly
 Gly Arg Arg Asp Ile Phe Leu Gly Ala Arg Glu Cys Gln Gly Tyr Val
 Ala Pro Cys Glu Phe Gly Ser Gly Asp Gly Phe Tyr Asp Gly Gln Gly
 Lys Tyr His Leu Gly Thr Met Val His Gly Phe Asn Tyr Pro Asp Glu
 Thr Gly Gln His Gln Leu Asp Val Arg Leu Trp Ser Ala Val Met Glu
 Asn Gly Tyr Ile Gln Phe Pro Arg Pro Glu Asp Cys Pro Ile Val Arg
 Pro Val Lys Glu Met Glu Pro Lys Ile Phe Asn Pro Asp Asn Val Gln
 Ser Ala Glu Gln Leu Leu His Asp Leu Gly Gly Glu

Bacillus cereus F65185

Met Lys Lys Glu Glu Glu Ser Leu Arg Asn Ser Ile Glu Phe Glu Val
 Phe Gly Asp Tyr Ala Leu Phe Thr Asp Pro Leu Met Lys Met Gly Gly
 Glu Lys Leu Thr Tyr Gln Val Pro Thr Tyr Gln Ala Ile Lys Gly Ile
 Val Glu Ser Ile Tyr Trp Lys Pro Thr Leu Leu Met Ile Val Asp Lys
 Ile Arg Ile Met Asn Ala Ile Lys Met Glu Ser Lys Gly Ile Arg Pro
 Ile Glu Tyr Gly Gly Gly Asn Thr Leu Ala Asn Tyr Thr Tyr Leu Lys
 Asn Val Arg Tyr Gln Val Gln Ala His Phe Ile Phe Asn Pro His Arg
 Pro Asp Leu Ala Phe Asp Arg Asn Glu Tyr Lys His His Asn Ile Leu
 Lys Arg Ser Leu Lys Val Gly Gly Arg Arg Asp Ile Phe Leu Gly Thr
 Arg Glu Cys Gln Gly Tyr Val Glu Pro Cys Val Phe Gly Glu Gly Glu
 Gly Phe Tyr Asp Asn Tyr Gly Gly Asp Ile His Leu Gly Thr Met Val
 His Gly Leu Asn Tyr Pro Asp Glu Thr Gly Arg Asn Glu Leu Glu Val
 Arg Leu Trp Asn Pro Val Met Arg Asp Gly Ile Ile Gln Phe Ile Arg
 Pro Glu Glu Cys Thr Lys Ile Arg Lys Ile Ser Lys Met Glu Pro Lys
 Ile Phe Asp Ser Ser Asn Val Glu Ser Val Asp Glu Leu Ile Lys Gln
 Leu Glu Glu Gly Gly Glu

FIG. 42B

73 / 73

Cas5 sequences

Selenomonas noxia ATCC 43541

Met	Arg	Asn	Ser	Ile	Glu	Phe	Gln	Val	Tyr	Gly	Arg	Met	Ala	Leu	Phe
Thr	Asp	Pro	Ile	Thr	Lys	Ile	Gly	Gly	Glu	Lys	Ala	Ser	Tyr	Ser	Val
Pro	Thr	Tyr	Gln	Ala	Leu	Lys	Gly	Ile	Thr	Glu	Ser	Ile	Tyr	Trp	Lys
Pro	Thr	Ile	Ile	Trp	Phe	Ile	Asp	Glu	Val	Arg	Val	Met	Lys	Arg	Ile
Thr	Thr	Gln	Val	Arg	Gly	Val	Lys	Pro	Leu	Lys	Tyr	Gly	Asp	Ser	Gly
Asn	Asp	Leu	Ser	Tyr	Tyr	Lys	Tyr	Leu	Ser	Asp	Val	Cys	Tyr	Gln	Val
Arg	Ala	His	Phe	Glu	Phe	Asn	Met	His	Arg	Glu	Glu	Leu	Lys	Glu	Asp
Arg	Asp	Glu	His	Lys	His	His	Asn	Ile	Ala	Lys	Arg	Met	Val	Glu	Arg
Gly	Gly	Arg	Arg	Asp	Ile	Phe	Leu	Gly	Thr	Arg	Glu	Cys	Gln	Gly	Tyr
Val	Glu	Pro	Val	Lys	Tyr	Gly	Val	Gly	Lys	Gly	Tyr	Tyr	Asp	Asn	Val
Asp	Glu	Leu	Pro	Leu	Gly	Ile	Met	Leu	His	Gly	Phe	Asn	Tyr	Pro	Asp
Glu	Thr	Gly	Glu	Asp	Lys	Leu	Gln	Val	Arg	Phe	Trp	Lys	Pro	Thr	Met
Lys	Lys	Gly	Ile	Ile	His	Phe	Arg	Arg	Pro	Glu	Lys	Cys	Glu	Met	Val
Arg	Asp	Ile	Arg	Glu	Val	Arg	Thr	Lys	Gln	Phe	Asp	Ala	Asp	Asn	Val
Phe	Phe	Ala	Glu	Glu	Glu	Glu	Lys	Gln	Leu	Glu	Gly	Gly	His	Leu	

FIG. 42C