



등록특허 10-2758180



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년01월23일
(11) 등록번호 10-2758180
(24) 등록일자 2025년01월17일

- (51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) *C07H 21/04* (2006.01)
- (52) CPC특허분류
C12N 15/113 (2013.01)
C07H 21/02 (2013.01)
- (21) 출원번호 10-2023-7043357(분할)
- (22) 출원일자(국제) 2017년12월29일
심사청구일자 2024년01월11일
- (85) 번역문제출일자 2023년12월14일
- (65) 공개번호 10-2023-0175330
- (43) 공개일자 2023년12월29일
- (62) 원출원 특허 10-2019-7022278
원출원일자(국제) 2017년12월29일
심사청구일자 2020년12월29일
- (86) 국제출원번호 PCT/US2017/069019
- (87) 국제공개번호 WO 2018/126176
국제공개일자 2018년07월05일

- (30) 우선권주장
62/441,046 2016년12월30일 미국(US)
62/492,001 2017년04월28일 미국(US)

- (56) 선행기술조사문현
WO2016186745 A1*
US20160215275 A1
WO2016164356 A1
WO2016182917 A1

*는 심사관에 의하여 인용된 문현

전체 청구항 수 : 총 19 항

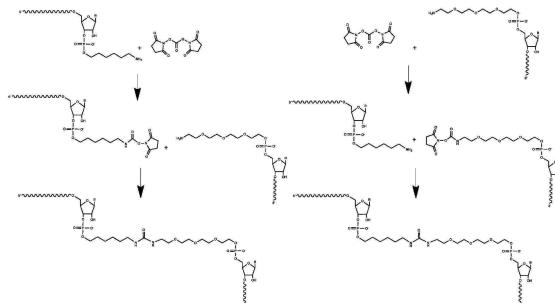
심사관 : 문동현

- (54) 발명의 명칭 합성 가이드 문자, 그와 관련된 조성물 및 방법

(57) 요 약

가이드 문자의 화학적 합성이 그와 관련된 조성물 및 방법과 함께 개시된다.

대 표 도



(52) CPC특허분류

C07H 21/04 (2013.01)
C12N 2310/20 (2017.05)
C12N 2310/315 (2013.01)
C12N 2310/318 (2013.01)
C12N 2310/319 (2013.01)
C12N 2320/53 (2013.01)
C12N 2330/30 (2013.01)

(72) 발명자

사코마노 셈

미국 콜로라도주 80301 볼더 스위트 208 노틸러스
코트 노스 4909 아이2 파마슈티컬스 내

케이프하트 스테이시

미국 콜로라도주 80301 볼더 스위트 208 노틸러스
코트 노스 4909 아이2 파마슈티컬스 내

페르난데스 세실리아

미국 매사추세츠주 02141 캠브리지 헬리 스트리트
11 에디타스 메디신, 인코포레이티드 내

자야람 하리하란

미국 캘리포니아주 94401 샌 마티오 사우스 프리몬
트 스트리트 #401 337

이튼 브루스

미국 콜로라도주 80301 볼더 스위트 208 노틸러스
코트 노스 4909 아이2 파마슈티컬스 내

베리 카린 쟁스키

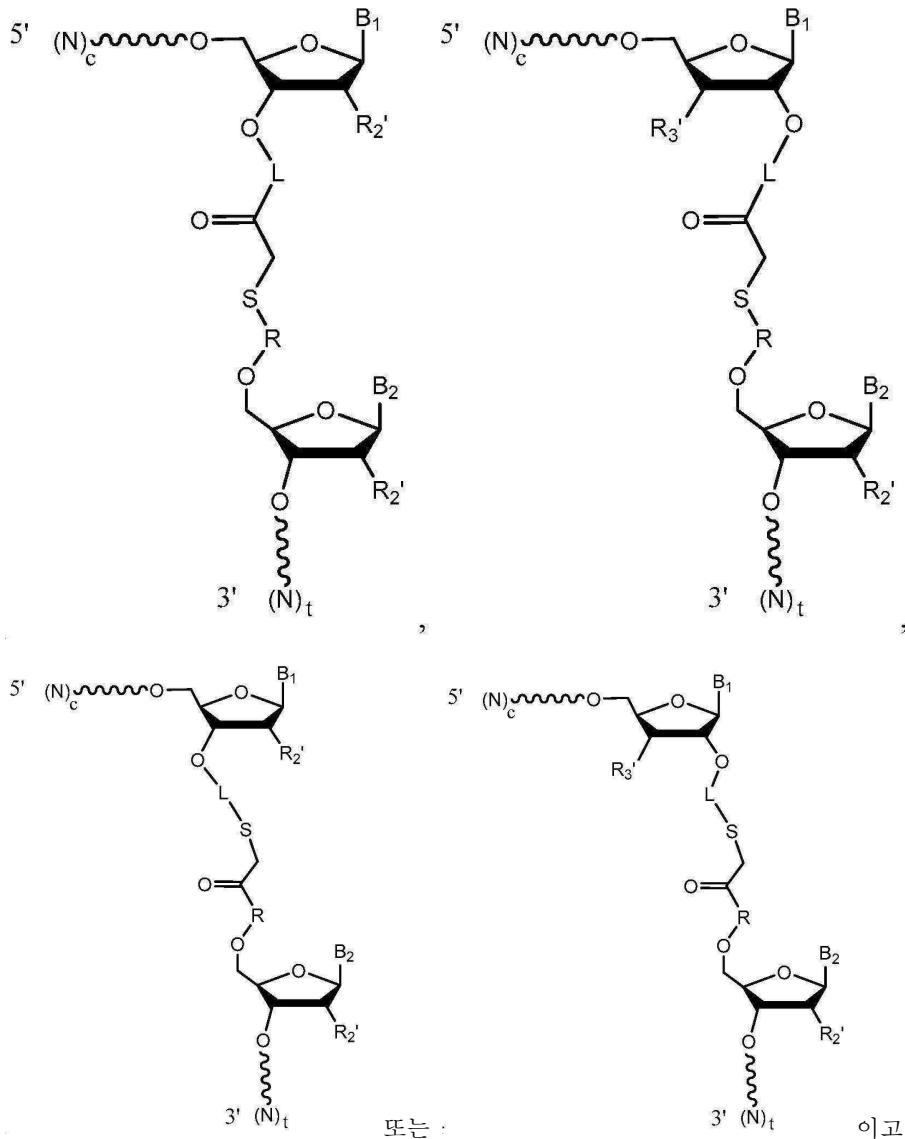
미국 콜로라도주 80301 볼더 스위트 208 노틸러스
코트 노스 4909 아이2 파마슈티컬스 내

명세서

청구범위

청구항 1

CRISPR 시스템용 단분자 가이드 분자로서, 상기 가이드 분자는 티오에테르를 포함하는 비 뉴클레오타이드 연결을 포함하며, 상기 가이드 분자는 화학식



상기 화학식에서,

$(N)_c$ 및 $(N)_t$ 의 각각의 N은 독립적으로, 각각 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 통해 그의 인접 뉴클레오타이드(들)에 연결된 뉴클레오타이드 잔기이고;

$(N)_c$ 는 $(N)_t$ 의 5' 영역에 상보성이거나 부분적으로 상보성이고, 그와 이중가닥을 형성하는 3' 영역을 포함하고;

c는 정수 20 이상이고;

t는 정수 20 이상이고;

각각의 ~~~~~는 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티

오포스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 나타내고;

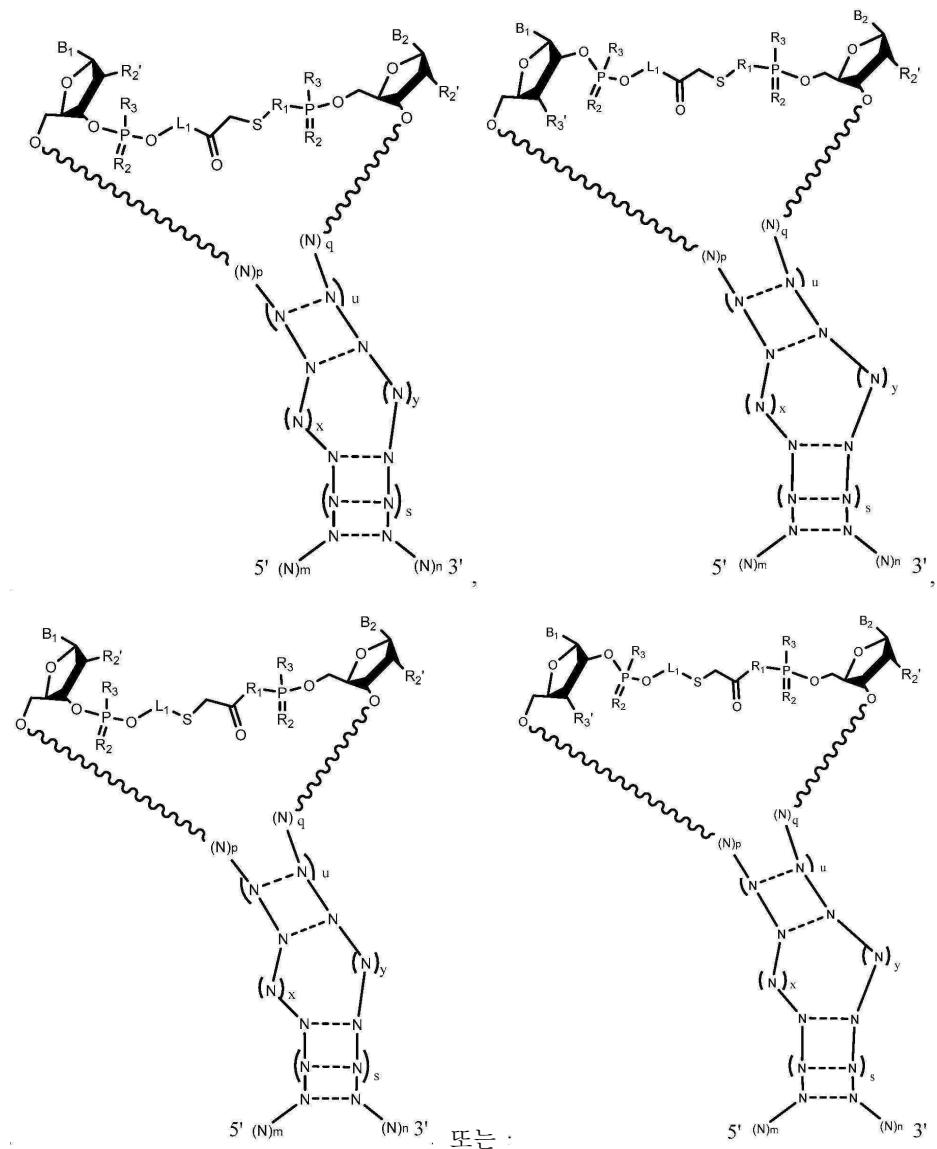
각각의 R_2' 및 R_3' 는 독립적으로 H , OH , 풀루오로, 클로로, 브로모, NH_2 , SH , $S-R'$, 또는 $O-R'$ 이고, 각각의 R' 는 독립적으로 보호기 또는 알킬기이고, 알킬기는 선택적으로 치환될 수 있고;

L 및 R 은 각각 독립적으로 비 뉴클레오타이드 링커이고;

B_1 및 B_2 는 각각 독립적으로 핵염기인, 단분자 가이드 분자.

청구항 2

제1항에 있어서, 상기 가이드 분자는 화학식



상기 화학식에서,

L_1 및 R_1 은 각각 독립적으로 비 뉴클레오타이드 링커이고;

각각의 R_2 는 독립적으로 O 또는 S 이고;

각각의 R_3 은 독립적으로 O^- 또는 COO^- 이고;

p 및 q 는 각각 독립적으로 정수 0 내지 6(상하한 포함)이고, $p+q$ 는 정수 0 내지 6(상하한 포함)이고;

u 는 정수 2 내지 22(상하한 포함)이고;

s 는 정수 1 내지 10(상하한 포함)이고;

x 는 정수 1 내지 3(상하한 포함)이고;

y는 x를 초과하고, 정수 3 내지 5(상하한 포함)이고;

m 은 정수 15 이상이고;

n은 정수 30 이상이고;

각각의 N은 독립적으로, 각각 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미테이트 연결을 통해 그의 인접 뉴클레오타이드(들)에 연결된 뉴클레오타이드 잔기, 선택적으로 변형된 뉴클레오타이드 잔기이고;

각각의 N---N은 독립적으로 2개의 상보성 뉴클레오타이드, 선택적으로, 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타내는, 가이드 문자.

청구항 3

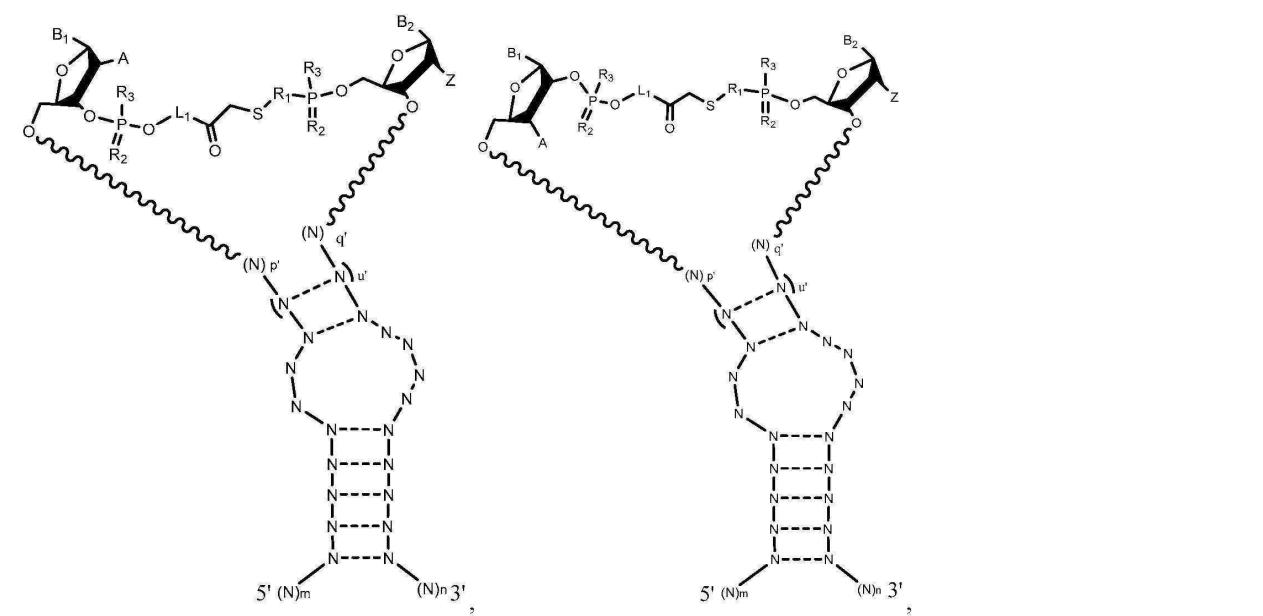
제2항에 있어서, p 및 q는 각각 0인, 가이드 문자.

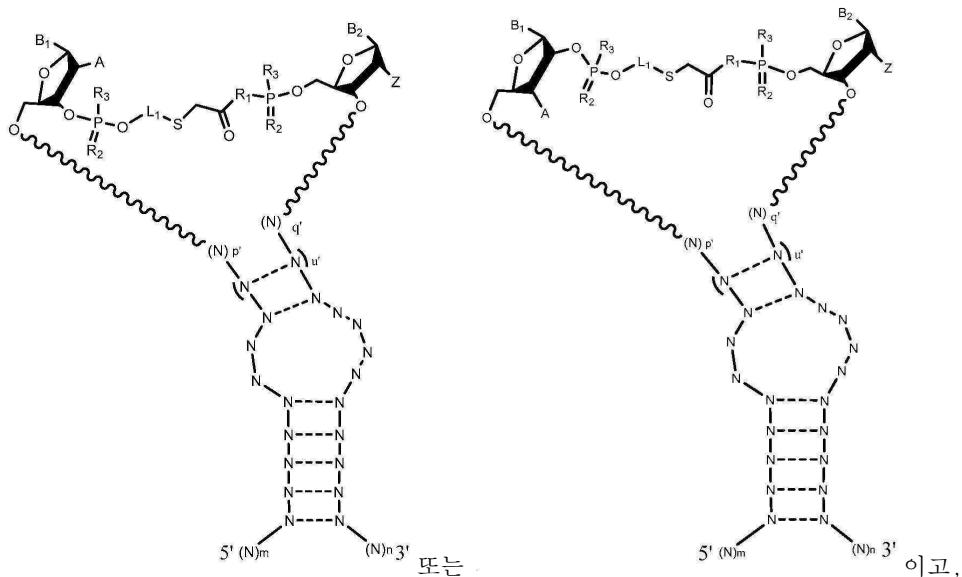
청구항 4

제2항에 있어서, u는 정수 4 내지 14(상하한 포함)인, 가이드 문자.

청구항 5

제2항에





상기 화학식에서,

u' 는 정수 2 내지 22(상하한 포함)이고;

p' 및 q' 는 각각 독립적으로 정수 0 내지 4(상하한 포함)이고, $p'+q'$ 는 정수 0 내지 4(상하한 포함)인, 가이드 문자.

청구항 6

제5항에 있어서, p' 및 q' 는 각각 0인, 가이드 문자.

청구항 7

제2항에

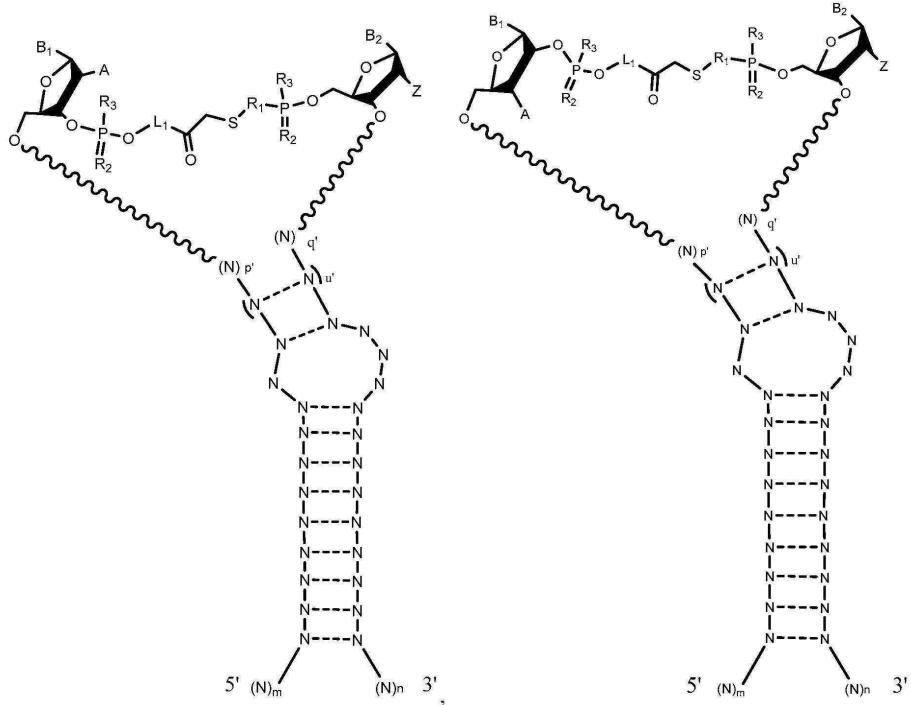
있어서,

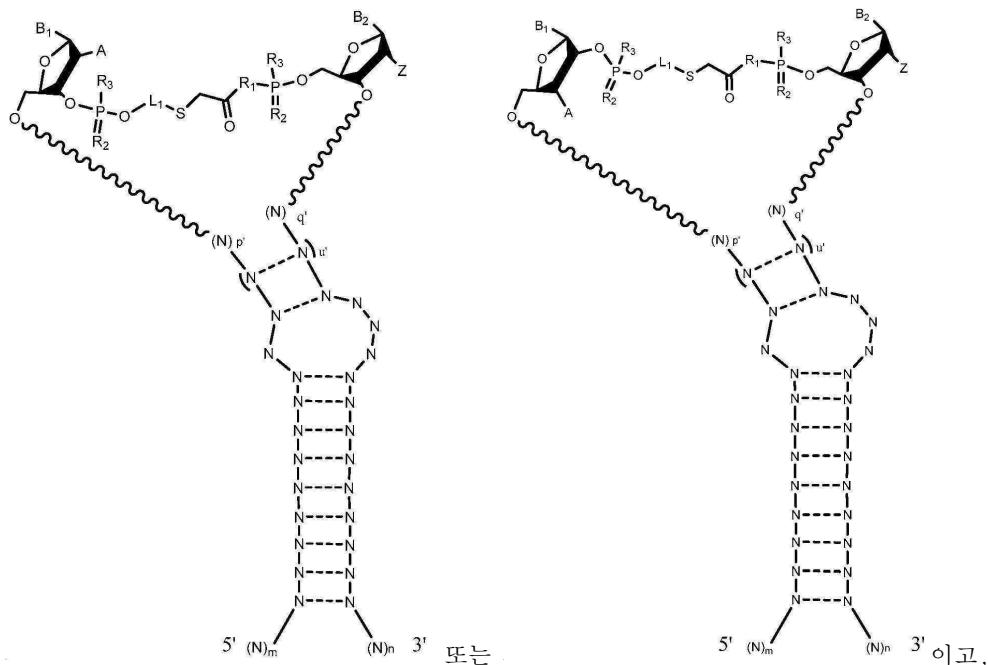
상기

가이드

분자는

화학식





상기 화학식에서,

u'는 정수 2 내지 22(상하한 포함)이고;

p' 및 q'는 각각 독립적으로 정수 0 내지 4(상하한 포함)이고, p'+q'는 정수 0 내지 4(상하한 포함)인, 가이드 문자.

청구항 8

제7항에 있어서, p' 및 q'는 각각 0인, 가이드 문자.

청구항 9

제2항에 있어서, L₁ 및 R₁은 -(CH₂)_w- , -(CH₂)_w-NH-C(O)-(CH₂)_w-NH- , -(OCH₂CH₂)_v-NH-C(O)-(CH₂)_w- 및 -(CH₂CH₂O)_v- 으로부터 각각 독립적으로 선택되고, 여기서, 각각의 w는 정수 1 내지 20(상하한 포함)이고, 각각의 v는 정수 1 내지 10(상하한 포함)인, 가이드 문자.

청구항 10

제9항에 있어서, L₁은 -(CH₂)₆-NH-C(O)-(CH₂)-NH-이고, R₁은 -(OCH₂CH₂)₄-NH-C(O)-(CH₂)₂-인, 가이드 문자.

청구항 11

제1항에 있어서, 상기 가이드 문자는

SEQ ID NO. 35 - [L] - SEQ ID NO. 57이고, 여기서, [L]은 티오에테르를 포함하는 비 뉴클레오타이드 연결인, 가이드 문자.

청구항 12

제1항 내지 제11항 중 어느 한 항의 가이드 문자 또는 이의 약제학적으로 허용 가능한 염을 포함하는, 세포 또는 대상체에서 표적 핵산 서열을 변형시키거나 표적 핵산 서열의 발현을 조절하기 위한, 조성물.

청구항 13

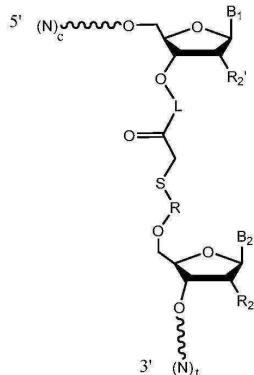
제12항에 있어서, 상기 가이드 문자가 용액 중에 또는 약제학적으로 허용되는 담체 중에 혼탁된, 조성물.

청구항 14

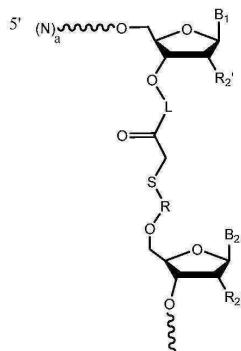
제12항에 있어서, Cas9 단백질을 추가로 포함하고, 상기 가이드 분자 및 Cas9 단백질은 (i) 상기 가이드 분자의 표적화 도메인 서열에 상보성인 서열 및 (ii) Cas9 단백질에 의해 인식되는 프로토스페이서 인접 모티프(PAM) 서열을 포함하는 표적 핵산과 상호작용할 수 있는 복합체를 형성하는, 조성물.

청구항 15

제12항에 있어서, 상기 가이드 분자는 화학식



또는 그의 약제학적으로 허용 가능한 염이고,



본 조성물은 화학식

의 분자 또는 그의 약제학적으로 허용 가능한 염을 실질적으로 무함유하고,

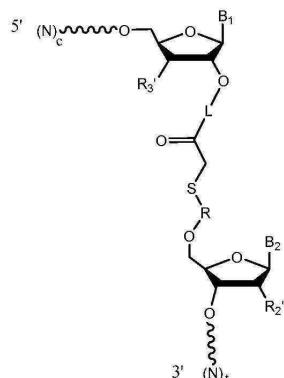
상기 화학식에서,

a는 c와 동일하지 않고/않거나;

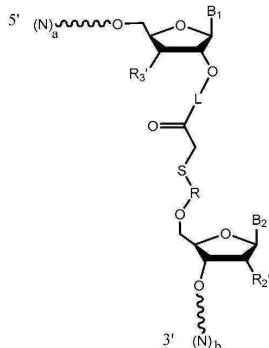
b는 t와 동일하지 않은, 조성물.

청구항 16

제12항에 있어서, 상기 가이드 분자는 화학식



또는 그의 약제학적으로 허용 가능한 염이고,



본 조성물은 화학식
함유하고,

상기 화학식에서,

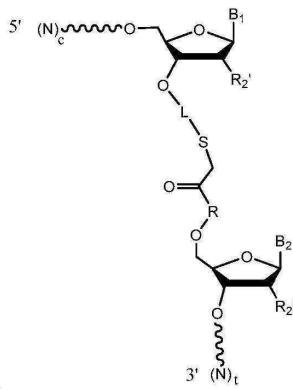
a는 c와 동일하지 않고/않거나;

b는 t와 동일하지 않은, 조성물.

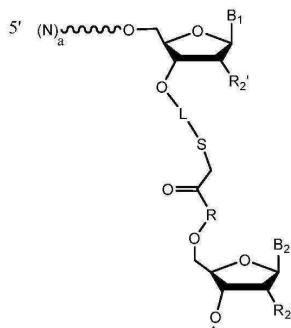
의 분자 또는 그의 약제학적으로 허용 가능한 염을 실질적으로 무

청구항 17

제12항에 있어서, 상기 가이드 분자는 화학식



또는 그의 약제학적으로 허용 가능한 염이고,



본 조성물은 화학식
무함유하고,

의 분자 또는 그의 약제학적으로 허용 가능한 염을 실질적으로

상기 화학식에서,

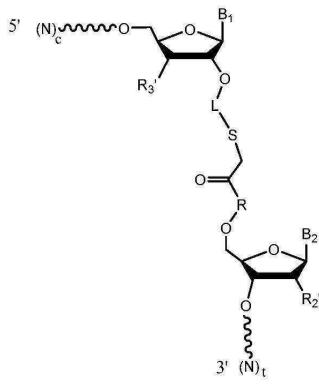
a는 c와 동일하지 않고/않거나;

b는 t와 동일하지 않은, 조성물.

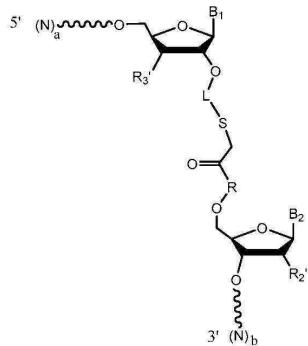
무

청구항 18

제12항에 있어서, 상기 가이드 분자는 화학식



또는 그의 약제학적으로 허용 가능한 염이고,



본 조성물은 화학식
무함유하고,

의 분자 또는 그의 약제학적으로 허용 가능한 염을 실질적으로

상기 화학식에서,

a는 c와 동일하지 않고/않거나;

b는 t와 동일하지 않은, 조성물.

청구항 19

제1항 내지 제11항 중의 어느 한 항의 가이드 분자를 합성하는 방법으로서, 본 방법은,

제1 올리고뉴클레오타이드 및 제2 올리고뉴클레오타이드를 어닐링함으로써, 상기 제1 올리고뉴클레오타이드의 3' 영역과 상기 제2 올리고뉴클레오타이드의 5' 영역 사이에 이중가닥을 형성하는 단계로서, 여기서, (i) 상기 제1 올리고뉴클레오타이드는 제1 반응성 기를 포함하고, 상기 제1 반응성 기는 브로모아세틸 모이어티를 포함하고 2' 반응성 기 또는 3' 반응성 기이고, 상기 제2 올리고뉴클레오타이드는 제2 반응성 기를 포함하고, 상기 제2 반응성 기는 술피드릴 모이어티를 포함하고 5' 반응성 기이거나, 또는 (ii) 상기 제1 올리고뉴클레오타이드는 제1 반응성 기를 포함하고, 상기 제1 반응성 기는 술피드릴 모이어티를 포함하고 2' 반응성 기 또는 3' 반응성 기이고, 상기 제2 올리고뉴클레오타이드는 제2 반응성 기를 포함하고, 상기 제2 반응성 기는 브로모아세틸 모이어티를 포함하고 5' 반응성 기인, 단계; 및

상기 어닐링된 제1 및 제2 올리고뉴클레오타이드를 상기 제1 및 제2 반응성 기를 통해 접합시킴으로써, 상기 제1 및 제2 올리고뉴클레오타이드를 연결하는 공유결합을 포함하는 상기 가이드 분자를 형성하는 단계를 포함하는, 방법.

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

청구항 111

삭제

청구항 112

삭제

청구항 113

삭제

청구항 114

삭제

청구항 115

삭제

청구항 116

삭제

청구항 117

삭제

청구항 118

삭제

청구항 119

삭제

청구항 120

삭제

청구항 121

삭제

청구항 122

삭제

청구항 123

삭제

청구항 124

삭제

청구항 125

삭제

청구항 126

삭제

청구항 127

삭제

청구항 128

삭제

청구항 129

삭제

청구항 130

삭제

청구항 131

삭제

청구항 132

삭제

청구항 133

삭제

청구항 134

삭제

청구항 135

삭제

청구항 136

삭제

청구항 137

삭제

청구항 138

삭제

청구항 139

삭제

청구항 140

삭제

청구항 141

삭제

청구항 142

삭제

청구항 143

삭제

청구항 144

삭제

청구항 145

삭제

청구항 146

삭제

청구항 147

삭제

청구항 148

삭제

청구항 149

삭제

청구항 150

삭제

청구항 151

삭제

청구항 152

삭제

청구항 153

삭제

청구항 154

삭제

청구항 155

삭제

청구항 156

삭제

청구항 157

삭제

청구항 158

삭제

청구항 159

삭제

청구항 160

삭제

청구항 161

삭제

청구항 162

삭제

청구항 163

삭제

청구항 164

삭제

청구항 165

삭제

청구항 166

삭제

청구항 167

삭제

청구항 168

삭제

청구항 169

삭제

청구항 170

삭제

청구항 171

삭제

청구항 172

삭제

청구항 173

삭제

청구항 174

삭제

청구항 175

삭제

청구항 176

삭제

청구항 177

삭제

청구항 178

삭제

청구항 179

삭제

청구항 180

삭제

청구항 181

삭제

청구항 182

삭제

청구항 183

삭제

청구항 184

삭제

청구항 185

삭제

청구항 186

삭제

청구항 187

삭제

청구항 188

삭제

청구항 189

삭제

청구항 190

삭제

청구항 191

삭제

청구항 192

삭제

청구항 193

삭제

청구항 194

삭제

청구항 195

삭제

청구항 196

삭제

청구항 197

삭제

청구항 198

삭제

청구항 199

삭제

청구항 200

삭제

청구항 201

삭제

청구항 202

삭제

청구항 203

삭제

청구항 204

삭제

청구항 205

삭제

청구항 206

삭제

청구항 207

삭제

청구항 208

삭제

청구항 209

삭제

청구항 210

삭제

청구항 211

삭제

청구항 212

삭제

청구항 213

삭제

청구항 214

삭제

청구항 215

삭제

청구항 216

삭제

청구항 217

삭제

청구항 218

삭제

청구항 219

삭제

청구항 220

삭제

청구항 221

삭제

청구항 222

삭제

청구항 223

삭제

청구항 224

삭제

청구항 225

삭제

청구항 226

삭제

청구항 227

삭제

청구항 228

삭제

청구항 229

삭제

청구항 230

삭제

청구항 231

삭제

청구항 232

삭제

청구항 233

삭제

청구항 234

삭제

청구항 235

삭제

청구항 236

삭제

청구항 237

삭제

청구항 238

삭제

청구항 239

삭제

청구항 240

삭제

청구항 241

삭제

청구항 242

삭제

청구항 243

삭제

청구항 244

삭제

청구항 245

삭제

청구항 246

삭제

청구항 247

삭제

청구항 248

삭제

청구항 249

삭제

발명의 설명

기술 분야

[0001]

관련 출원에 대한 상호 참조

[0002]

본 출원은 각 개시내용 전문이 본 명세서에 참조로 포함되는 2016년 12월 30일에 출원된 미국 출원 제 62/441,046호 및 2017년 4월 28일에 출원된 미국 출원 제 62/492,001호의 우선권 이익을 주장한다.

[0003]

본 개시내용은 표적 핵산 서열을 편집하거나 표적 핵산 서열의 발현을 조절하기 위한 CRISPR/Cas 관련 방법 및 성분에 관한 것이다. 더욱 구체적으로, 본 개시내용은 합성 가이드 분자 및 관련 시스템, 방법 및 조성물에 관한 것이다.

배경 기술

[0004]

CRISPR(주기적으로 간격을 띠고 분포하는 짧은 회문구조 반복서열)(Clustered Regularly Interspaced Short Palindromic Repeats)은 바이러스 공격으로부터 방어하기 위한 적응 면역 시스템으로서 박테리아 및 고세균에서 발달하였다. 바이러스에 노출되면, 바이러스 DNA의 짧은 절편이 CRISPR 유전자 좌위에 혼입된다. RNA는 바이러스 서열을 포함하는 CRISPR 유전자 좌위의 일부분으로부터 전사된다. 바이러스 게놈에 상보성인 서열을 포함하는 RNA는 바이러스 게놈의 표적 서열에 대해 Cas9 또는 Cpf1과 같은 RNA-가이드 뉴클레아제 단백질의 표적화를 매개한다. 결과적으로 RNA 가이드 뉴클레아제는 바이러스 표적을 절단하여, 침묵화시킨다.

[0005]

최근, CRISPR 시스템은 진핵 세포에서의 게놈 편집에 적용되었다. 이들 시스템은 일반적으로 단백질 성분(RNA-가이드 뉴클레아제) 및 핵산 성분(일반적으로 가이드 분자, 가이드 RNA 또는 "gRNA"로 지칭됨)을 포함한다. 이들 2개의 성분은 시스템의 2개의 성분에 의해 인식되거나 이에 상보성인 특이적 표적 DNA 서열과 상호 작용하며, 예를 들어 부위 특이적 DNA 절단에 의해 표적 서열을 선택적으로 편집하거나 변이시키는 복합체를 형성한다. 표적 서열의 편집 또는 변이는 또한 비 상동성 말단 결합(NHEJ) 또는 상동성 유도 수선(HDR)과 같은 세포 DNA 수선 메카니즘의 집합을 포함할 수 있다.

[0006]

유전적 질병을 치료하는 수단으로서의 CRISPR 시스템의 가치는 널리 인식되어 있지만, 광범위한 임상 적용을 달성하기 위해 이들 시스템을 기반으로 한 치료법에 대한 특정 기술적 문제점이 해결되어야 한다. 특히, 고품질의 CRISPR 시스템 성분의 비용 효율적이고, 복잡하지 않은 상업적 규모의 합성에 대한 필요성이 존재한다.

[0007]

예를 들어, 대부분의 가이드 분자는 현재, 시험관내 전사(IVT) 및 화학적 합성의 2가지 방법 중 하나에 의해 합성된다. IVT는 통상적으로 T7 중합효소와 같은 박테리아 RNA 중합효소에 의한 DNA 주형으로부터의 RNA의 전사를 포함한다. 현재 미국 및 해외 규제 당국이 요구하는 우량 제조 기준(GMP) 표준에 따른 가이드 분자의 IVT 제조는 비용이 많이 들고, 규모가 제한적일 수 있다. 추가로, IVT 합성은 모든 가이드 RNA 서열에 적합하지는 않을 수 있다: T7 중합효소는 다른 5' 염기로 개시된 것보다 더 효율적으로 5' 구아닌으로 개시하는 서열을 전사시키는 경향이 있고, 절두된 가이드 분자 전사체를 생성하는 것으로, 전사 종결 신호로서 특정 가이드 분자 내에 그 구조가 존재하는 폴리 우라실 판이 후행하는 줄기 루프 구조를 인식할 수 있다.

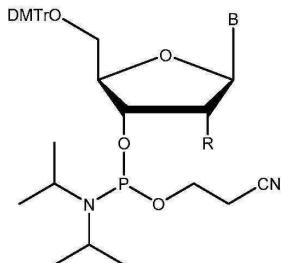
[0008]

반면에, 화학적 합성은 비용이 저렴하고, 더 짧은 올리고뉴클레오타이드(예를 들어, 100개 미만의 뉴클레오타이드 길이)의 GMP 생산이 용이하게 이용 가능하다. 화학적 합성 방법은, 예를 들어 그 전문이 본 명세서의 모든 목적을 위해 참조로 포함되는 문헌[Beaucage and Carruthers, Curr Protoc Nucleic Acid Chem. 2001 May; Chapter 3: Unit 3.3 (Beaucage & Carruthers)] 전반에 기재되어 있다. 이들 방법은 통상적으로 적절한 길이의 올리고뉴클레오타이드 서열에 도달할 때까지 반응성 뉴클레오타이드 단량체를 단계적으로 첨가하는 단계를 포함

한다. 가장 일반적으로 사용되는 합성 방식(예를 들어, 포스포르아미디트 방법)에서 단량체는 올리고뉴클레오타이드의 5' 말단에 첨가된다. 이들 단량체는 종종 3' 기능화되고(예를 들어, 포스포르아미디트), 예를 들어 하기 화학식 I에 따른 5' 보호기(예컨대, 4,4'-디메톡시트리틸)를 포함한다:

[0009]

[화학식 I]



[0010]

[0011]

화학식 I에서, DMTr은 4,4'-디메톡시트리틸이고, R은 H, F, O-알킬 또는 보호된 하이드록실기를 포함하는 비 제한적인 예인 올리고뉴클레오타이드 합성 조건에 적합한 기이고, B는 임의의 적합한 핵염기이다. (문헌[Beaucage & Carruthers]). 5' 보호된 단량체를 사용하는 경우, 각 라운드의 첨가 후 탈보호 단계가 필요하며, 5' 보호기가 제거되고, 하이드록실기가 유지된다.

[0012]

어떤 화학 반응이 이용되더라도, 5' 잔기의 단계적 첨가는 정량적으로 이루어지지 않으며; 일부 올리고뉴클레오타이드는 일부 잔기의 첨가가 "누락"될 것이다. 이는 적절한 올리고뉴클레오타이드를 포함하지만, 다양한 잔기가 누락된 더 짧은 올리고뉴클레오타이드(다른 절두 또는 결실 종뿐만 아니라, n-2, n-3 등을 포함할 수 있으나 "n-1 종"으로 지칭됨)로 오염된 합성 산물을 생성한다. n-1 종에 의한 오염을 최소화하기 위해, 다수의 화학적 합성 도식은 단계적 첨가 단계와 탈보호 단계 사이에 "캡핑(capping)" 반응을 포함한다. 캡핑 반응에서, 비 반응성 모이어티가 5' 보호기에 의해 종결되지 않는 올리고뉴클레오타이드의 5' 말단에 첨가되고; 이러한 비 반응성 모이어티는 올리고뉴클레오타이드에 단량체가 추가로 첨가되는 것을 방지하고, 약 60개 또는 70개의 염기 길이의 올리고뉴클레오타이드의 합성 동안 허용 가능한 낮은 수준으로 n-1 오염을 감소시키는 데 효과적이다. 그러나 캡핑 반응은 정량적이지 않으며, 단분자 가이드 RNA와 같은 더 긴 올리고뉴클레오타이드에서 n-1 오염을 방지하는데 효과적이지 않을 수 있다. 반면에 DMT 보호가 커플링 반응 중에 소실되어, 더 긴 올리고뉴클레오타이드(n+2, n+3, 등을 포함할 수 있지만, "n+1 종"으로 지칭됨)가 생성되는 경우가 있다. n-1 종 및/또는 n+1 종으로 오염된 단분자 가이드 RNA는 다른 방법으로 제조된 전장 가이드 RNA와 동일한 방식으로 작용하지 않을 수 있으므로, 이는 잠재적으로 치료법에서 합성 가이드 RNA의 사용을 복잡하게 한다.

발명의 내용

[0013]

본 개시내용은 특히, 2개 이상의 사전 어닐링된 가이드 단편을 교차 연결시키는 단계를 포함하는 단분자 가이드 분자의 합성 방법을 제공함으로써, n-1 및/또는 n+1 종, 절두 종 및 다른 오염물을 갖는 고 순도 단분자 가이드 분자의 비용 효율적이고, 복잡하지 않은 화학적 합성에 대한 필요성을 다룬다. 일부 구현예에서, 본 명세서에 제공되는 단분자 가이드 분자는 5' 말단에서 서열 정확도가 개선되어, 부적절한 표적외 편집을 감소시킨다. 또한 n-1 및/또는 n+1 오염을 실질적으로 무함유하는 전장 단분자 가이드 분자를 포함하거나 이로 필수적으로 이루어진 조성물이 또한 본 명세서에 제공된다.

[0014]

본 개시내용의 특정 양태는, 가이드 단편이 본 명세서에 기재된 우레아-기반 교차 연결 방법에 사용되는 아민-기능화된 단편과 같은 동종다중기능성(예를 들어, 동종이중기능성)인 경우, 가이드 단편의 사전 어닐링이 특히 유용할 수 있다는 인식을 포괄한다. 실제로 동종다중기능성 가이드 단편을 이종이량체로 사전 어닐링하면 부적절한 동종이량체의 형성을 감소시킬 수 있다. 따라서, 본 개시내용은 또한 부산물(예를 들어, 동종이량체)을 실질적으로 무함유하는 전장 단분자 가이드 분자를 포함하거나 이로 필수적으로 이루어진 조성물을 제공한다.

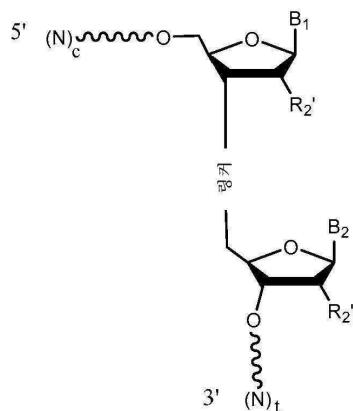
[0015]

일 양태에서, 본 개시내용은 CRISPR 시스템용 단분자 가이드 분자를 합성하는 방법에 관한 것이며, 본 방법은:

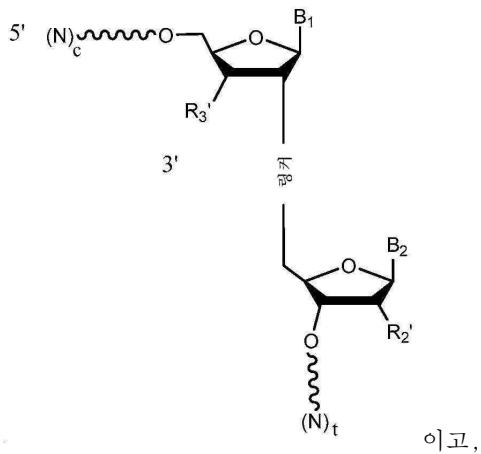
[0016]

제1 올리고뉴클레오타이드 및 제2 올리고뉴클레오타이드를 어닐링함으로써, 제1 올리고뉴클레오타이드의 3' 영역과 제2 올리고뉴클레오타이드의 5' 영역 사이에 이중가닥을 형성하는 단계로서, 제1 올리고뉴클레오타이드는 2' 반응성 기 및 3' 반응성 기 중 적어도 하나의 제1 반응성 기를 포함하고, 제2 올리고뉴클레오타이드는 5' 반응성 기인 제2 반응성 기를 포함하는 단계; 및

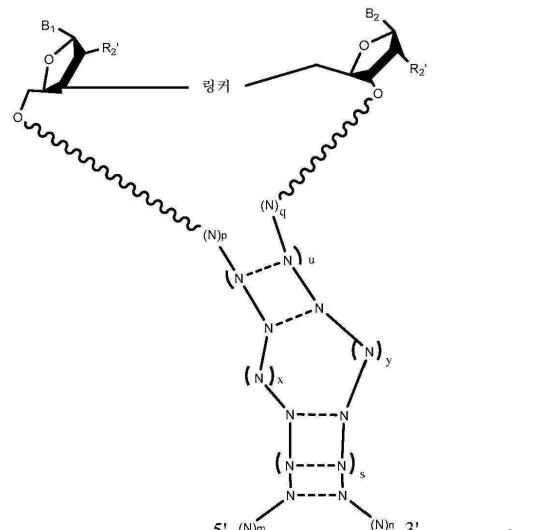
- [0017] 제1 및 제2 반응성 기를 통해 어닐링된 제1 및 제2 올리고뉴클레오타이드를 접합시켜, 제1 및 제2 올리고뉴클레오타이드를 연결하는 공유 결합을 포함하는 단분자 가이드 RNA 분자를 형성하는 단계를 포함한다.
- [0018] 일 양태에서, 본 개시내용은 CRISPR 시스템용 단분자 가이드 분자에 관한 것이다. 일부 구현예에서, 본 명세서에 제공되는 단분자 가이드 분자는 II형 CRISPR 시스템용이다.
- [0019] 일부 구현예에서, 제1 올리고뉴클레오타이드의 5' 영역은 표적 서열(예를 들어, 전핵생물 유전자 내의 표적 서열) 내의 표적 도메인에 완전히 또는 부분적으로 상보성인 표적화 도메인을 포함한다.
- [0020] 일부 구현예에서, 제2 올리고뉴클레오타이드의 3' 영역은 하나 이상의 줄기 루프 구조를 포함한다.
- [0021] 일부 구현예에서, 본 명세서에 제공되는 단분자 가이드 분자는 Cas9 분자와 상호작용하고, Cas9/가이드 분자 복합체의 형성을 매개할 수 있다.
- [0022] 일부 구현예에서, 본 명세서에 제공되는 단분자 가이드 분자는 Cas9 또는 RNA-가이드 뉴클레아제와의 복합체이다.
- [0023] 일부 구현예에서, 본 명세서에 제공되는 단분자 가이드 분자는 5'에서 3'으로:
- [0024] 표적화 도메인 서열;
- [0025] 제1 하부 줄기 서열;
- [0026] 제1 별지 서열(bulge sequence);
- [0027] 제1 상부 줄기 서열;
- [0028] 을 포함하는 제1 가이드 분자 단편
- [0029] 비 뉴클레오타이드 화학적 연결; 및
- [0030] 제2 상부 줄기 서열;
- [0031] 제2 별지 서열; 및
- [0032] 제2 하부 줄기 서열
- [0033] 을 포함하는 제2 가이드 분자 단편
- [0034] 을 포함하고,
- [0035] (a) 제1 하부 줄기 서열의 적어도 하나의 뉴클레오타이드는 제2 하부 줄기 서열의 뉴클레오타이드와 염기쌍을 형성하고, (b) 제1 상부 줄기 서열의 적어도 하나의 뉴클레오타이드는 제2 상부 줄기 서열의 뉴클레오타이드와 염기쌍을 형성한다.
- [0036] 일부 구현예에서, 단분자 가이드 분자는 제1 및 제2 상부 줄기 서열 사이에 테트라루프 서열을 포함하지 않는다. 일부 구현예에서, 제1 및/또는 제2 상부 줄기 서열은 4개 내지 22개(상하한 포함)의 수의 뉴클레오타이드를 포함한다.



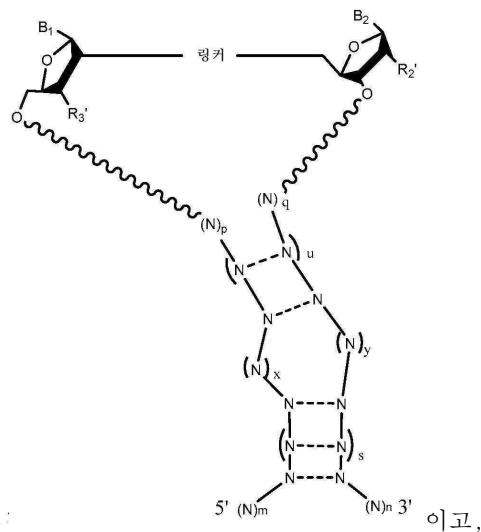
[0037] 일부 구현예에서, 단분자 가이드 분자는 화학식 또는



- [0038] $(N)_c$ 및 $(N)_t$ 의 각각의 N은 독립적으로, 각각 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 통해 그의 인접 뉴클레오타이드(들)에 연결된 뉴클레오타이드 잔기, 선택적으로 변형된 뉴클레오타이드 잔기이고;
- [0039] $(N)_c$ 는 $(N)_t$ 의 5' 영역에 상보성이거나 부분적으로 상보성이고, 그와 이중가닥을 형성하는 3' 영역을 포함하고;
- [0040] c는 정수 20 이상이고;
- [0041] t는 정수 20 이상이고;
- [0042] 링커는 비 뉴클레오타이드 화학적 연결이고;
- [0043] B_1 및 B_2 는 각각 독립적으로 핵염기이고;
- [0044] 각각의 R_2' 및 R_3' 는 독립적으로 H, OH, 플루오로, 클로로, 브로모, NH_2 , SH, S-R', 또는 O-R'이고, 각각의 R'는 독립적으로 보호기 또는 알킬기이고, 알킬기는 선택적으로 치환될 수 있고;
- [0045] 각각의 $\sim\sim\sim$ 는 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 나타낸다.
- [0046] 일부 구현예에서, $(N)_c$ 는 II형 CRISPR 시스템으로부터의 반복의 적어도 일부분을 포함하는 3' 영역을 포함한다. 일부 구현예에서, $(N)_c$ 는 표적 서열 내의 표적 도메인에 완전히 또는 부분적으로 상보성인 표적화 도메인을 포함하는 3' 영역을 포함한다. 일부 구현예에서, $(N)_t$ 는 하나 이상의 줄기 루프 구조를 포함하는 3' 영역을 포함한다.



- [0047] 일부 구현예에서, 단분자 가이드 분자는 화학식 또는



[0048] 각각의 N은 독립적으로, 각각 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오프스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 통해 그의 인접 뉴클레오타이드(들)에 연결된 뉴클레오타이드 잔기, 선택적으로 변형된 뉴클레오타이드 잔기이고;

[0049] 각각의 N---N은 독립적으로 2개의 상보성 뉴클레오타이드, 선택적으로, 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타내고;

[0050] p 및 q는 각각 0이고;

[0051] u는 정수 2 내지 22(상하한 포함)이고;

[0052] s는 정수 1 내지 10(상하한 포함)이고;

[0053] x는 정수 1 내지 3(상하한 포함)이고;

[0054] y는 x를 초과하고, 정수 3 내지 5(상하한 포함)이고;

[0055] m은 정수 15 이상이고;

[0056] n은 정수 30 이상이다.

[0057] 일부 구현예에서, u는 정수 2 내지 22(상하한 포함)이고;

[0058] s는 정수 1 내지 8(상하한 포함)이고;

[0059] x는 정수 1 내지 3(상하한 포함)이고;

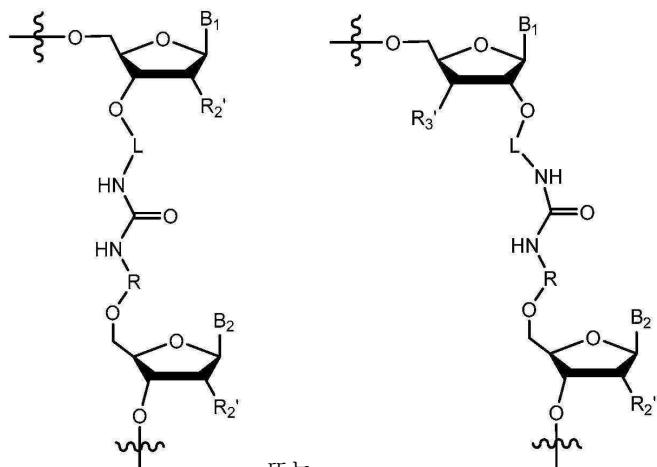
[0060] y는 x를 초과하고, 정수 3 내지 5(상하한 포함)이고;

[0061] m은 정수 15 내지 50(상하한 포함)이고;

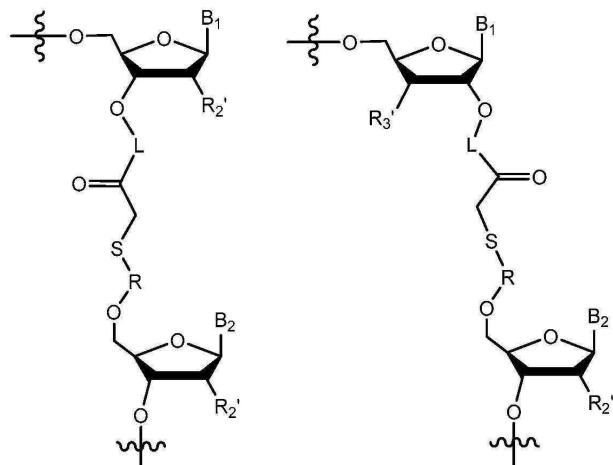
[0062] n은 정수 30 내지 70(상하한 포함)이다.

[0063] 일부 구현예에서, 가이드 분자는 테트라루프를 포함하지 않는다(p 및 q는 각각 0임). 일부 구현예에서, 하부 줄기 서열 및 상부 줄기 서열은 3개 초과의 뉴클레오타이드의 동일한 서열을 포함하지 않는다. 일부 구현예에서, u는 정수 3 내지 22(상하한 포함)이다.

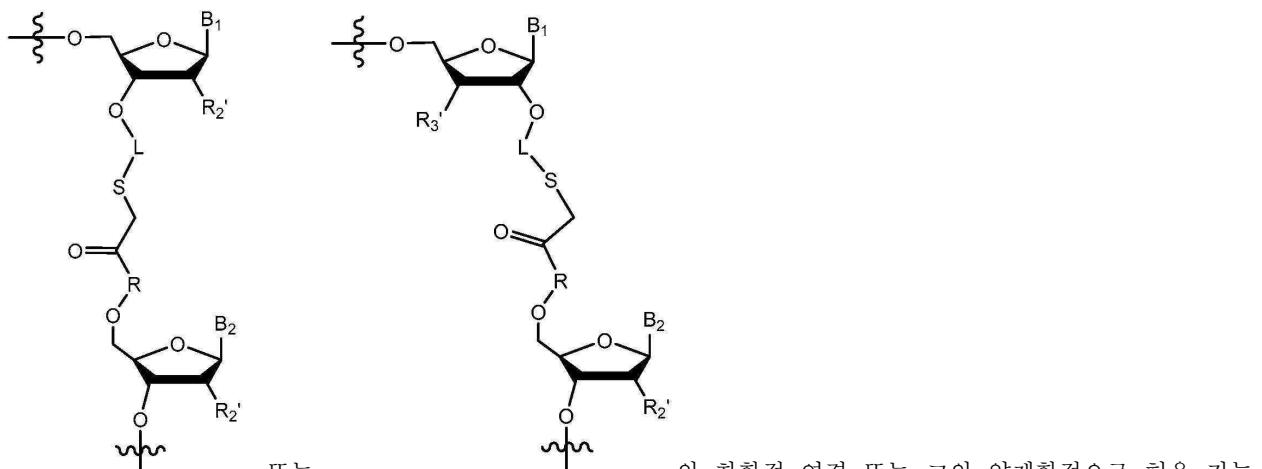
[0064] 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 각각 아미노기를 포함하고, 접합 단계는 이중기능성 교차 연결 시약을 포함하는 카보네이트에 의해 제1 및 제2 반응성 기의 아민 모이어티를 교차연결함으로써, 우레아 연결을 형성하는 단계를 포함한다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 브로모아세틸기 및 술 피드릴기이다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 포스페이트기 및 하이드록실기이다.



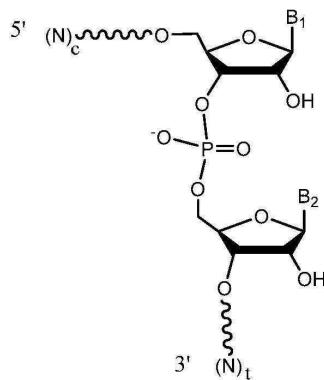
[0065] 일부 구현예에서, 단분자 가이드 분자는 화학식 또는 화학적 연결 또는 그의 약제학적으로 허용 가능한 염을 포함하고, L 및 R은 각각 독립적으로 비 뉴클레오파이드 화학적 링커이다.



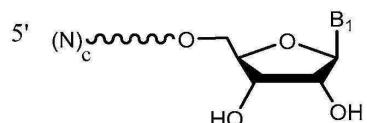
[0066] 일부 구현 예에서, 단문자 가이드 문자는 화학식



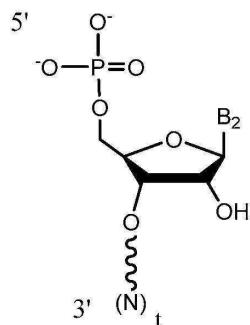
한 열을 포함하고, 1 및 R은 각각 돌령점으로 비-뉴클레오타이드 화합점 텁커이다.



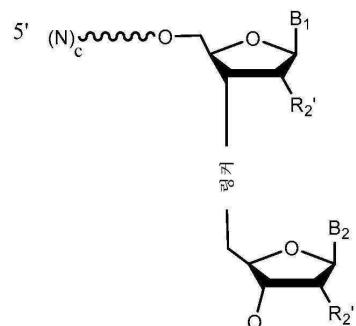
[0067] 일부 구현예에서, 단분자 가이드 분자는 화학식
능한 염이고,



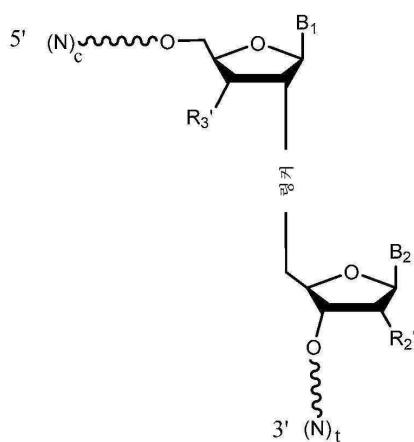
[0068] 포스포디에스테르 연결을 형성하기 위해, 활성화제의 존재하에 및



또는 그의 염 사이의 반응을 포함하는 방법에 의해 제조된다.

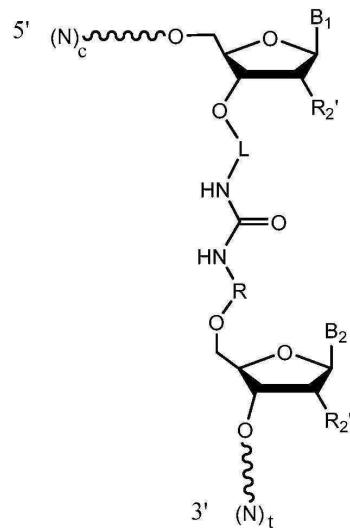


[0069] 일 양태에서, 본 개시내용은 화학식 또는



또는 그의 약제학적으로 허용 가능한 염인 단분자 가이드 분자를 포함하

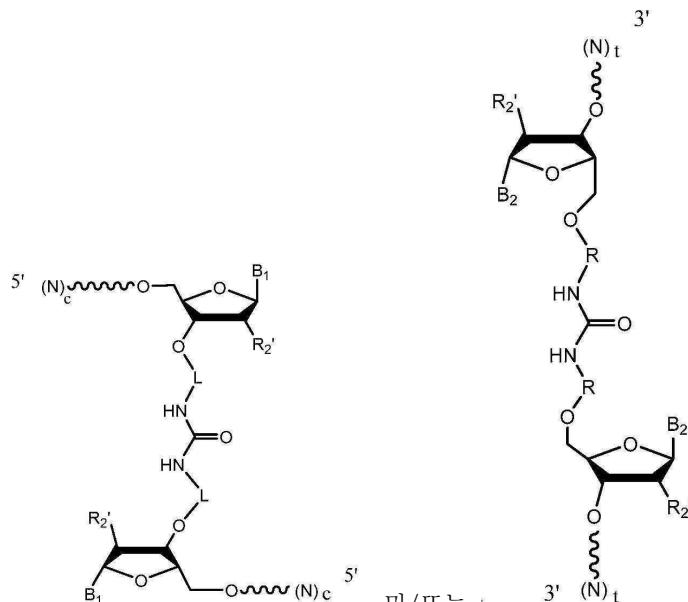
거나 이로 필수적으로 이루어진 CRISPR 시스템용 가이드 분자의 조성물에 관한 것이다. 일부 구현예에서, 약 10% 미만의 가이드 분자가 기준 가이드 분자 서열에 비교하여, 5' 말단에 절두를 포함한다. 일부 구현예에서, 적어도 약 99%의 가이드 분자는 기준 가이드 분자 서열의 해당 5' 서열과 100% 동일한 가이드 분자의 1개 내지 20개의 뉴클레오타이드를 포함하는 5' 서열을 포함한다.



[0070] 일부 구현예에서, 가이드 분자의 조성물은 화학식

또는 그의 약제학적으로 허용

가능한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고:

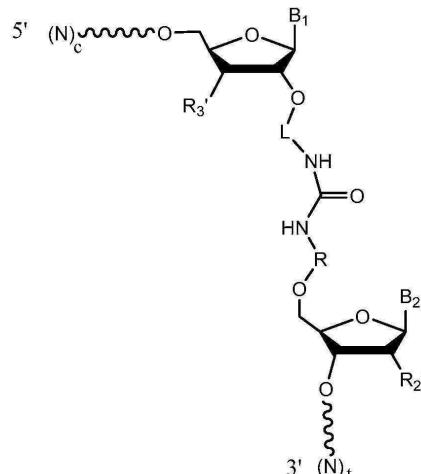


[0071] 본 조성물은 화학식

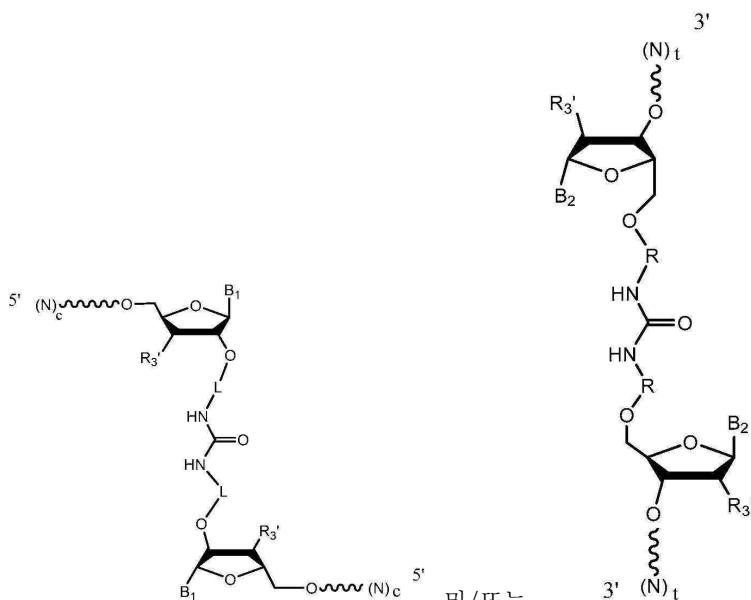
및/또는

또는 그의 약제학적으로 허

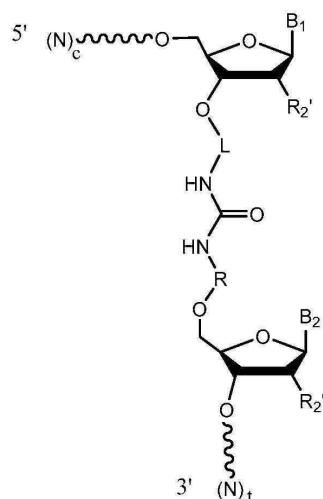
용 가능한 염의 분자를 실질적으로 무함유한다.



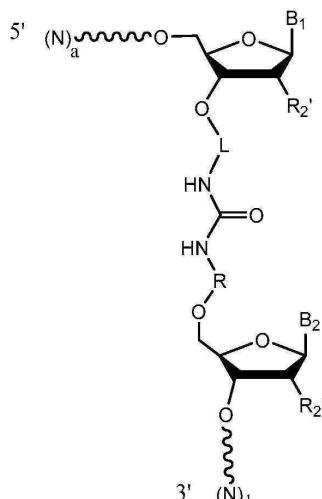
[0072] 일부 구현예에서, 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고:



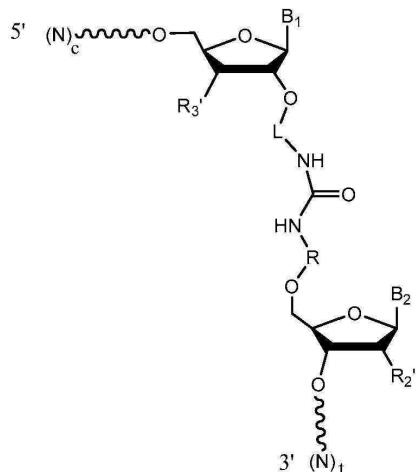
[0073] 본 조성물은 화학식 및/또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로 무함유한다.



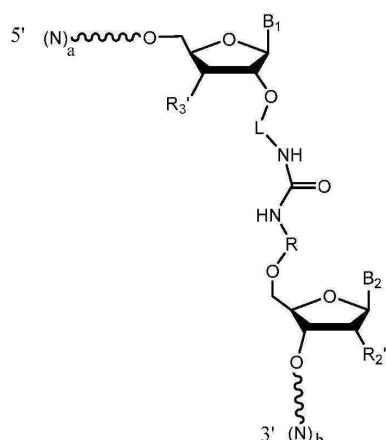
[0074] 일부 구현예에서, 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고,



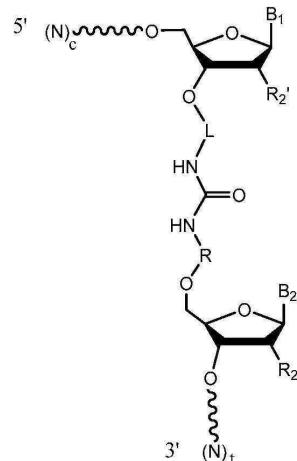
- [0075] 본 조성물은 화학식 .
또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로
무함유하고,
- [0076] a는 c와 동일하지 않고/않거나;
- [0077] b는 t와 동일하지 않다.



- [0078] 일부 구현예에서, 본 조성물은 화학식 .
또는 그의 약제학적으로 허용 가능
한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고,

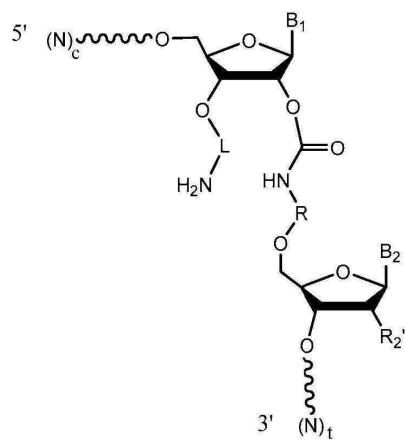


- [0079] 본 조성물은 화학식 .
또는 그의 약제학적으로 허용 가능한 염의 분자를 실질
적으로 무함유하고,
- [0080] a는 c와 동일하지 않고/않거나;
- [0081] b는 t와 동일하지 않다.



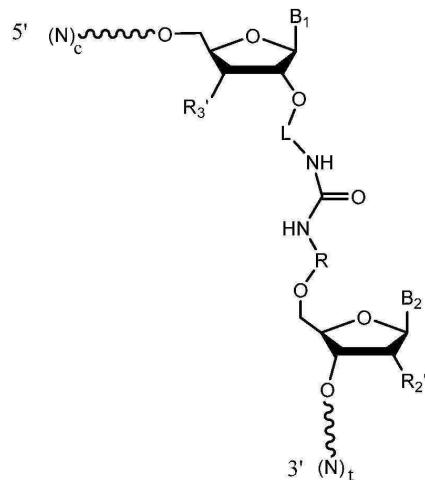
[0082]

일부 구현예에서, 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고:



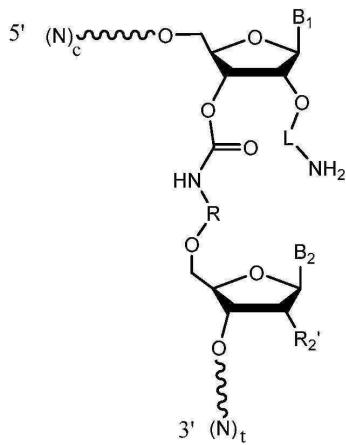
[0083]

본 조성물은 화학식 의 분자를 실질적으로 무함유한다.



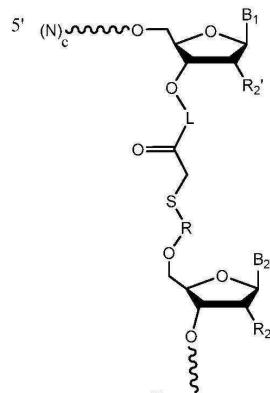
[0084]

일부 구현예에서, 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고:



[0085] 본 조성물은 화학식

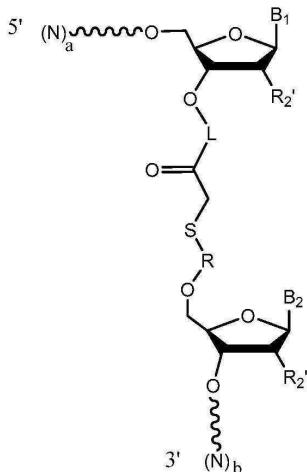
의 분자를 실질적으로 무함유한다.



[0086] 일부 구현예에서, 본 조성물은 화학식

또는 그의 약제학적으로 허용 가능한 염의 가이

드 분자를 포함하거나 이로 필수적으로 이루어지고,



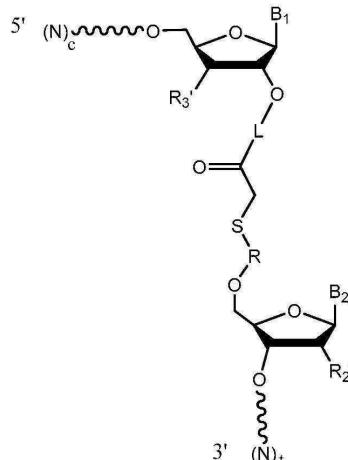
[0087] 본 조성물은 화학식

또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로

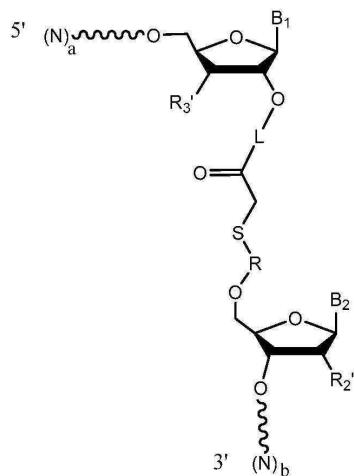
무함유하고,

[0088] a는 c와 동일하지 않고/않거나;

[0089] b는 t와 동일하지 않다.



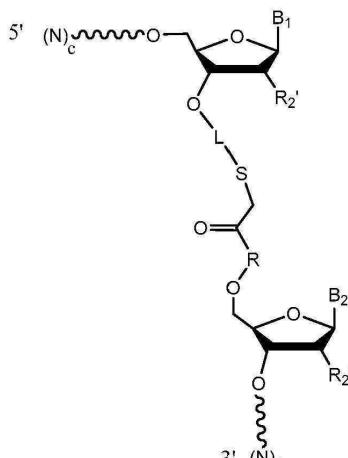
- [0090] 일부 구현예에서, 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 가이드 문자를 포함하거나 이로 필수적으로 이루어지고,



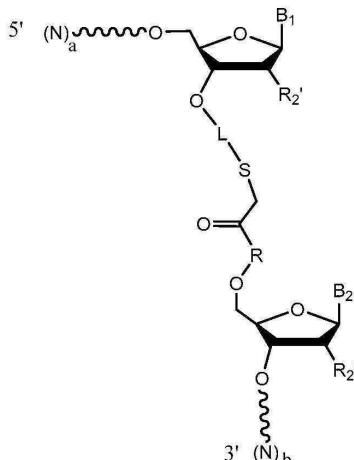
- [0091] 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로 무함유하고,

[0092] a는 c와 동일하지 않고/않거나;

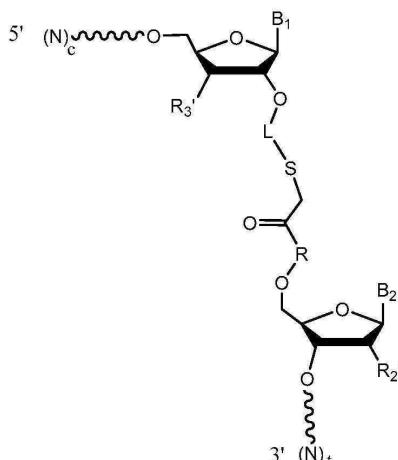
[0093] b는 t와 동일하지 않다.



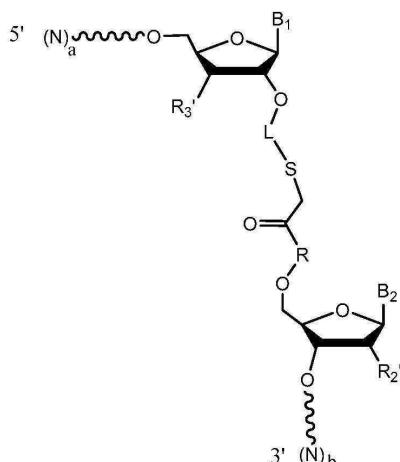
- [0094] 일부 구현예에서, 본 조성물은 화학식 또는 그의 약제학적으로 허용 가능한 염의 가이드 문자를 포함하거나 이로 필수적으로 이루어지고,



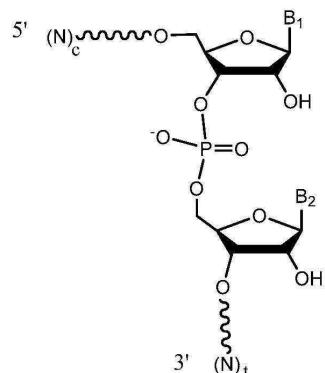
- [0095] 본 조성물은 화학식 .
으로 무함유하고,
[0096] a는 c와 동일하지 않고/않거나;
[0097] b는 t와 동일하지 않다.



- [0098] 일부 구현예에서, 본 조성물은 화학식 .
한 염의 가이드 분자를 포함하거나 이로 필수적으로 이루어지고,



- [0099] 본 조성물은 화학식 .
질적으로 무함유하고,
[0100] a는 c와 동일하지 않고/않거나;
[0101] b는 t와 동일하지 않다.
[0102] 일부 구현예에서, 본 조성물은

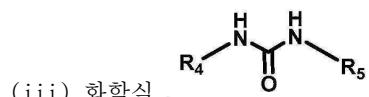


[0103] (a) CRISPR 시스템용 합성 단분자 가이드 분자로서, 가이드 분자는 화학식 그의 약제학적으로 허용 가능한 염인 CRISPR 시스템용 합성 단분자 가이드 분자; 및

[0104] (b)

[0105] (i) 카보디이미드, 또는 그의 염;

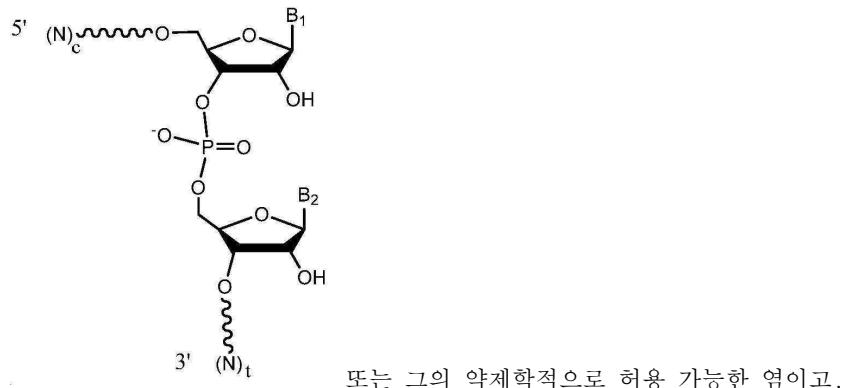
[0106] (ii) 이미다졸, 시아노이미다졸, 피리딘 및 디메틸아미노피리딘, 또는 그의 염; 및



[0107] (iii) 화학식 , 또는 그의 염의 화합물로서,

[0108] R₄ 및 R₅는 각각 독립적으로 치환되거나 치환되지 않은 알킬, 또는 치환되거나 치환되지 않은 카보사이클릭인 화합물 중 하나 이상을 포함한다.

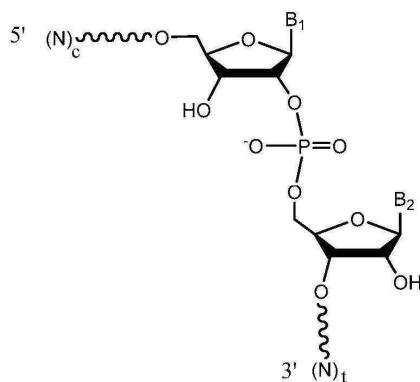
[0109] 일부 구현예에서, 본 조성물은 CRISPR 시스템용 합성 단분자 가이드 분자를 포함하고, 가이드 분자는 화학식



[0110] 본 조성물은 화학식 으로 무함유하고, 또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로

[0111] a+b는 c+t-k이고, k는 1 내지 10이다.

[0112] 일부 구현예에서, 본 조성물은 CRISPR 시스템용 합성 단분자 가이드 분자를 포함하거나 이로 필수적으로 이루어

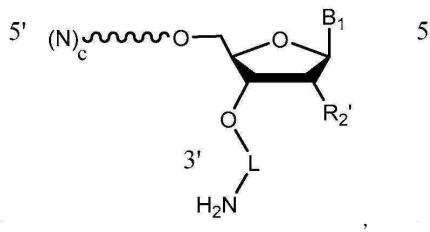


지고, 가이드 분자는 화학식
염이고:

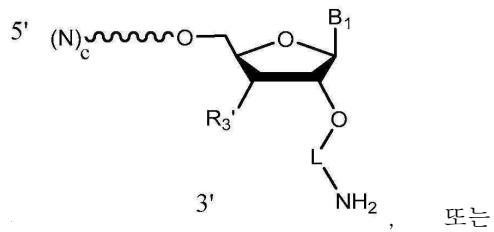
또는 그의 약제학적으로 허용 가능한

[0113] 화학식 내에 도시된 2'-5' 포스포디에스테르 연결은 (N)_c의 3' 영역과 (N)_t의 5' 영역 사이에 형성된 이중가닥
내의 2개의 뉴클레오타이드 사이에 존재한다.

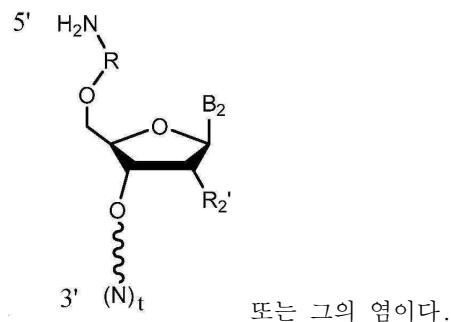
[0114] 일 양태에서, 본 개시내용은 본 명세서에 제공되는 단분자 가이드 분자의 합성 및/또는 본 명세서에 제공되는
방법에 의한 단분자 가이드 분자의 합성을 위한 올리고뉴클레오타이드에 관한 것이다. 일부 구현예에서, 올리고



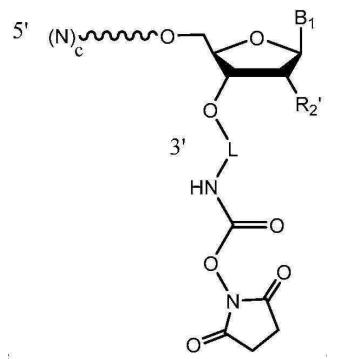
뉴클레오타이드는 화학식



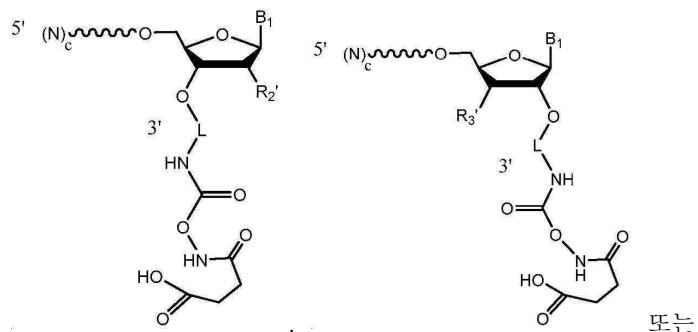
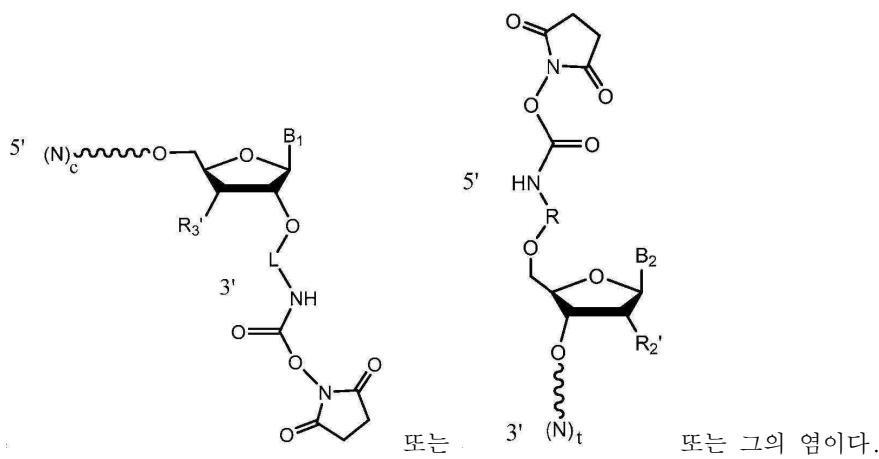
, 또는



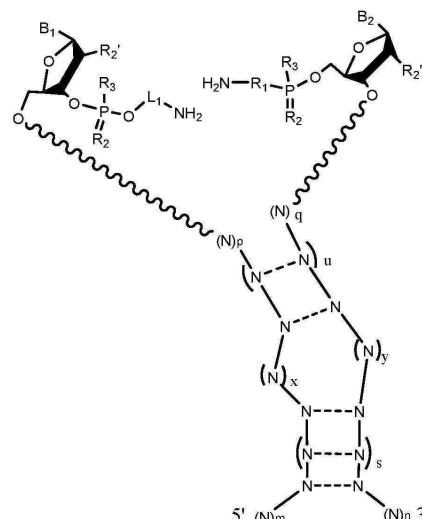
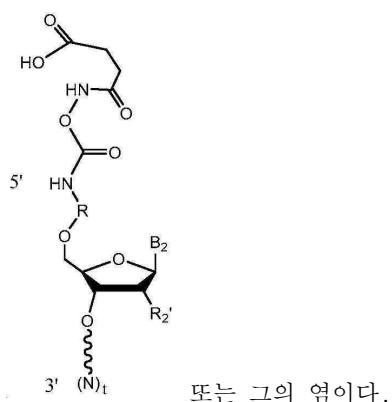
또는 그의 염이다.



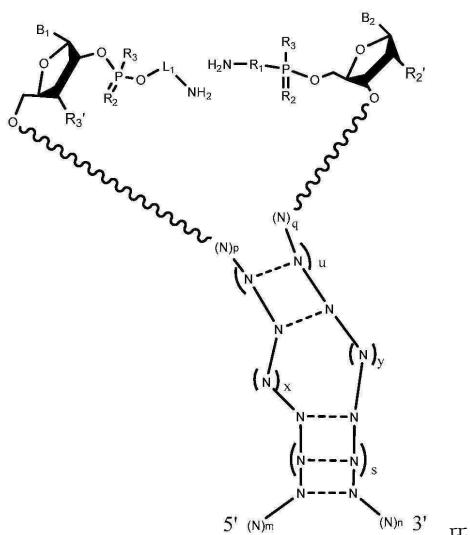
[0115] 일부 구현예에서, 올리고뉴클레오타이드는 화학식



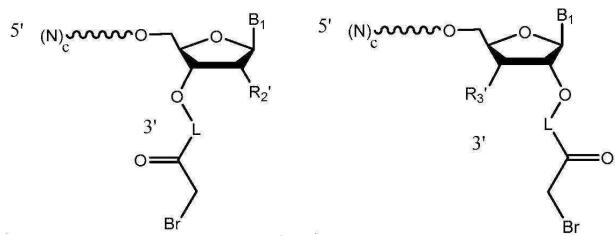
[0116] 일부 구현예에서, 올리고뉴클레오타이드는 화학식



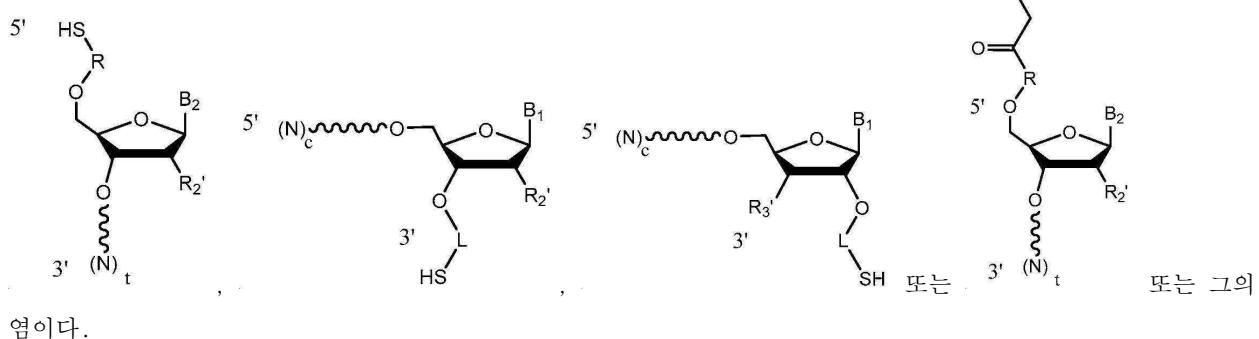
[0117] 일부 구현예에서, 본 조성물은 화학식 또는



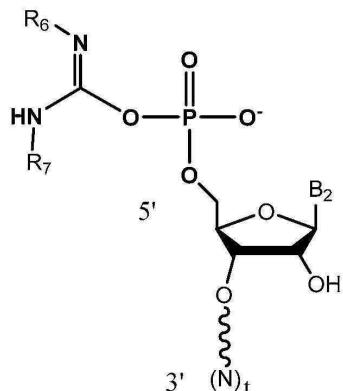
또는 그의 염의 어닐링된 이중가닥을 갖는 올리고뉴클레오타이드를 포함 한다.



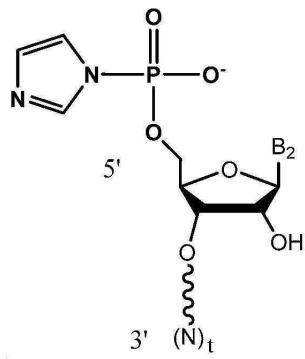
[0118] 일부 구현예에서, 올리고뉴클레오타이드는 화학식



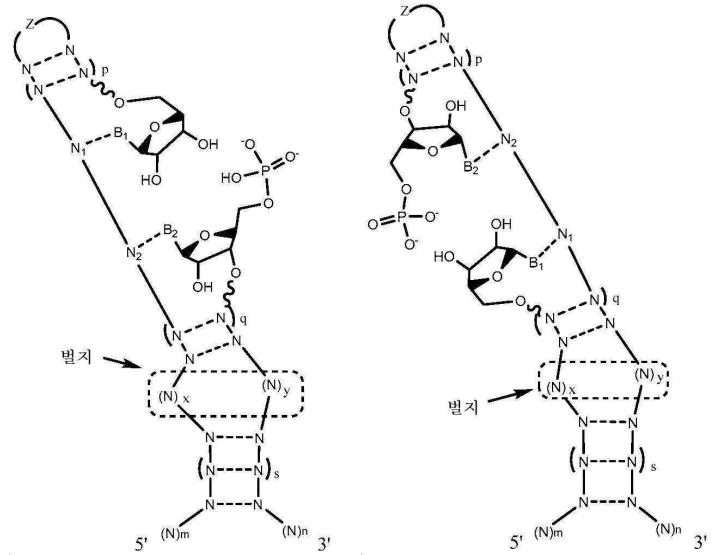
염이다.



[0119] 일부 구현예에서, 올리고뉴클레오타이드는 화학식

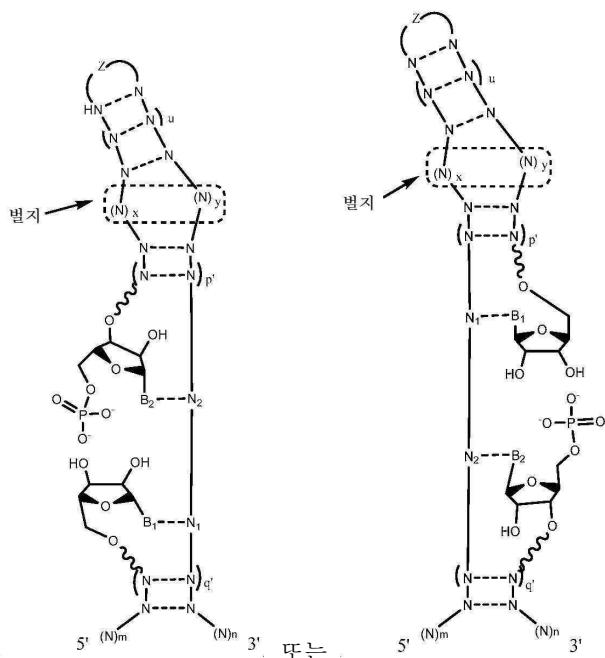


또는 그의 염이다.



[0120]

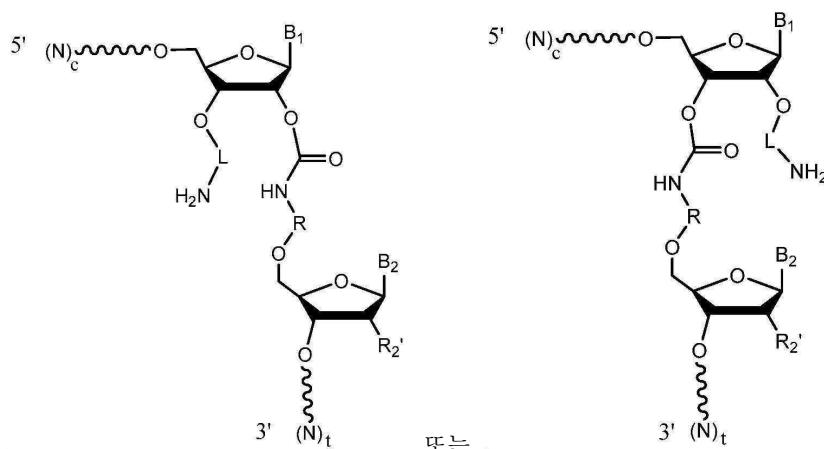
일부 구현예에서, 본 조성물은 화학식



[0121]

또는 그의 염의 어닐링된 이중가닥을 갖는 올리고뉴클레

오타이드를 포함한다.



[0122] 일 양태에서, 본 개시내용은 화학식
의 화합물에 관한 것이다.

[0123] 일 양태에서, 본 개시내용은 세포 또는 대상체에서 핵산을 변이시키는 방법에 관한 것이며, 이는 본 명세서에서 제공되는 가이드 문자 또는 조성물을 대상체에 투여하는 단계를 포함한다.

[0124] 일부 구현예에서, 본 명세서에서 제공되는 조성물은 임의의 정제 단계가 적용되지 않았다.

[0125] 일부 구현예에서, 본 명세서에서 제공되는 조성물은 용액 또는 약제학적으로 허용 가능한 담체 중에 혼탁된 단 문자 가이드 RNA 문자를 포함한다.

[0126] 일 양태에서, 본 개시내용은 본 명세서에 제공되는 가이드 문자를 포함하는 게놈 편집 시스템에 관한 것이다.
일부 구현예에서, 게놈 편집 시스템 및/또는 가이드 문자는 요법에 사용하기 위한 것이다. 일부 구현예에서, 게놈 편집 시스템 및/또는 가이드 문자는 약제의 생성에 사용하기 위한 것이다.

도면의 간단한 설명

[0127] 첨부 도면은 본 개시내용의 특정 양태 및 구현예의 포괄적인 예시가 아닌 예시적이고 개략적인 예를 제공하기 위한 것이다. 도면은 임의의 특정 이론이나 모델로 제한하거나 결부시키고자 하는 것은 아니며, 반드시 축척에 의하지 않는다. 전술한 것으로 제한하려는 것은 아니지만, 핵산 및 폴리펩타이드는 선형 서열 또는 개략적인 2차원 또는 3차원 구조로서 도시될 수 있으며; 이러한 도시는 그 구조에 관한 임의의 특정 모델 또는 이론으로 제한하거나 결부시키고자 하는 것이 아니며, 예시하기 위한 것이다.

도 1a는 본 개시내용의 특정 구현예에 따른 예시적인 교차 연결 반응 공정을 도시한다.

도 1b는 2차원의 개략적인 형태의 예시적인 *S. 피오게네스*(*S. pyogenes*) 가이드 문자를 도시하며, 본 개시내용의 다양한 구현예에 따라, 제1 및 제2 가이드 문자 단편이 함께 교차 연결된 위치(별표)가 강조 표시되어 있다.

도 1c는 2차원의 개략적인 형태의 예시적인 *S. 아우레우스*(*S. aureus*) 가이드 문자를 도시하며, 본 개시내용의 다양한 구현예에 따라, 제1 및 제2 가이드 문자 단편이 함께 교차 연결된 위치(별표)가 강조 표시되어 있다.

도 2a는 본 개시내용의 특정 구현예에 따른 예시적인 교차 연결 반응 공정의 단계를 도시한다.

도 2b는 본 개시내용의 특정 구현예에 따른 예시적인 교차 연결 반응 공정의 단계를 도시한다.

도 2c는 도 2a 및 도 2b로부터의 반응 산물을 사용하는 예시적인 교차 연결 반응 공정의 추가적 단계를 도시한다.

도 3a는 본 개시내용의 특정 구현예에 따른 예시적인 교차 연결 반응 공정의 단계를 도시한다.

도 3b는 본 개시내용의 특정 구현예에 따른 예시적인 교차 연결 반응 공정의 단계를 도시한다.

도 3c는 2차원의 개략적인 형태의 예시적인 *S. 피오게네스* 가이드 문자를 도시하며, 본 개시내용의 다양한 구현 예에 따라, 제1 및 제2 가이드 문자 단편이 함께 교차 연결된 위치가 강조 표시되어 있다.

도 3d는 2차원의 개략적인 형태의 예시적인 *S. 아우레우스* 가이드 문자를 도시하며, 본 개시내용의 다양한 구현 예에 따라, 제1 및 제2 가이드 문자 단편이 함께 교차 연결된 위치가 강조 표시되어 있다.

도 4는 시중 판매자로부터 입수한 IVT에 의해 제조된 결찰되지 않고 어닐링된 가이드 문자 단편 및 가이드 문자

와 비교하여, 본 개시내용의 특정 구현예에 따른 합성 단분자 가이드 분자에 대한 DNA 절단 투여량-반응 곡선을 도시한다. 본 명세서에 기재된 바와 같이, DNA 절단을 T7E1 분석에 의해 분석하였다. 그래프가 나타내는 바와 같이, 접합된 가이드 분자는 HEK293 세포에서 합성 단분자 가이드 분자 또는 IVT에 의해 생성된 단분자 가이드 분자에서 관찰된 것과 일치하는 투여량 의존적 방식으로 절단을 지원하였다. 접합되지 않고 어널링된 가이드 분자 단편은 유사한 투여량 의존적 방식이지만, 더 낮은 수준의 절단을 지원한다는 것을 유의해야 한다.

도 5a는 대표적인 이온 크로마토그래프를 도시하고, 도 5b는 실시예 1의 방법에 따라 우레아 링커와 접합된 이온 교환 정제 가이드 분자의 디콘볼루션(deconvolution)된 질량 스펙트럼을 도시한다. 도 5c는 대표적인 이온 크로마토그래프를 도시하고, 도 5d는 시판 제조된 합성 단분자 가이드 분자의 디콘볼루션된 질량 스펙트럼을 도시한다. 이온 크로마토그래프에서 강조 표시된 피크에 대해 질량 스펙트럼을 평가했다. 도 5e는 질량 스펙트럼의 확대된 버전을 도시한다. 좌측은 시판 제조된 합성 단분자 가이드 분자의 질량 스펙트럼이고(총 질량 순도 34%), 우측은 실시예 1의 방법에 따라 우레아 링커와 접합된 가이드 분자에 대한 질량 스펙트럼이다(총 질량 순도 72%).

도 6a는 우레아 연결을 포함하는 합성 단분자 가이드 분자로부터 생성된 상보성 DNA(cDNA)의 5' 말단으로부터 각 위치에서 개별 염기 및 길이 변화가 발생하는 빈도를 나타내는 그래프를 도시하고, 도 6b는 시판 제조된 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨)로부터 생성된 cDNA의 5' 말단으로부터 각 위치에서 개별 염기 및 길이 변화가 발생하는 빈도를 나타내는 그래프를 도시한다. 가이드 분자의 20 bp 표적화 도메인에 박스 표시하였다. 도 6c는 티오에테르 연결을 포함하는 합성 단분자 가이드 분자로부터 생성된 cDNA의 5' 말단으로부터 각 위치에서 개별 염기 및 길이 변화가 발생하는 빈도를 나타낸 그래프를 도시한다.

도 7a 및 도 7b는 우레아 연결을 포함하는 다양한 합성 단분자 가이드 분자(도 7a) 및 시판 제조된 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨)(도 7b)로부터 생성된 cDNA의 5' 말단으로부터의 처음 41개의 위치에서의 내부 서열 길이 변화(+5 내지 -5)를 도시하는 그래프이다.

도 8a 내지 도 8h는 2차원 개략적 형태로, 본 개시내용의 다양한 구현예에 따른 특정 예시적 가이드 분자의 구조를 도시한다. 염기쌍을 형성할 수 있는 상보성 염기는 염기 사이에 하나의 (A-U 또는 A-T 쌍) 또는 2개의 (G-C) 수평선으로 표시된다. 비 왓슨 크릭 쌍형성(non-Watson-Crick pairing)이 가능한 염기는 원이 포함된 하나의 수평선으로 표시된다.

도 9a 내지 도 9d는 2차원 개략적 형태로, 본 개시내용의 다양한 구현예에 따른 특정 예시적 가이드 분자의 구조를 도시한다. 염기쌍을 형성할 수 있는 상보성 염기는 염기 사이에 하나의 (A-U 또는 A-T 쌍) 또는 2개의 (G-C) 수평선으로 표시된다. 비 왓슨 크릭 쌍형성이 가능한 염기는 원이 포함된 하나의 수평선으로 표시된다.

도 10a 내지 도 10d는 2차원 개략적 형태로, 본 개시내용의 다양한 구현예에 따른 특정 예시적 가이드 분자의 구조를 도시한다. 염기쌍을 형성할 수 있는 상보성 염기는 염기 사이에 하나의 (A-U 또는 A-T 쌍) 또는 2개의 (G-C) 수평선으로 표시된다. 비 왓슨 크릭 쌍형성이 가능한 염기는 원이 포함된 하나의 수평선으로 표시된다.

도 11은 CD34+ 세포에서 표 10의 접합된 가이드 분자를 포함하는 일련의 리보핵단백질 복합체에 의한 DNA 절단 그래프를 도시한다. 차세대 시퀀싱 기법을 사용하여, 야생형 인간 기준 서열과 비교하여, 삽입 및 결실(삽입결실) %를 정량화하기 위해, 절단을 평가하였다. 실시예 1에 따라 생성된 결찰된 가이드 분자는 CD34+ 세포에서 DNA 절단을 지원한다. 줄기 루프 길이가 증가함에 따라, 삽입결실 %가 증가하는 것으로 나타났지만, 줄기 루프 서열에 인접한 U-A 교환의 혼입(gRNA 1E, 1F, 및 2D 참조)은 효과를 완화한다.

도 12a는 gRNA 1A의 T1 엔도뉴클레아제 분해 후 액체 크로마토그래피-질량 분광분석(LC-MS) 추적을 도시하고, 도 12b는 4.50분의 보유 시간에서 피크의 질량 스펙트럼을 도시한다(A34:G39). 특히, 우레아 연결, A-[UR]-AAUAG(A34:G39)을 포함하는 단편이 4.50분의 보유 시간에서 $m/z = 1190.7$ 로 검출되었다.

도 13a는 주된 산물(A-2, 보유 시간 3.25분) 및 부수 산물(A-1, 보유 시간 3.14분) 둘 모두가 존재하는 우레아 연결된 가이드 분자의 정제되지 않은 조성물에 대한 LC-MS 데이터를 도시한다. 본 발명자들은 설명하자면, 도 13a의 부수 산물(A-1)이 다량 존재하고, 통상적으로 실시예 1의 방법에 따른 가이드 분자의 합성시 최대 10%의 수율로 검출됨에 주목하였다. 도 13b는 피크 A-2의 디콘볼루션된 질량 스펙트럼(보유 시간 3.25분)을 도시하고, 도 13c는 피크 A-1의 디콘볼루션된 질량 스펙트럼(보유 시간 3.14분)을 도시한다. 질량 스펙트럼에 의한 각 피크의 분석은 두 산물 모두 동일한 분자량을 가짐을 시사한다.

도 14a는 실시예 10에 기재된 바와 같이, 화학적 변형 후의 가이드 분자 조성물에 대한 LC-MS 데이터를 도시한다. 주된 산물(B-1, 우레아)은 최초 분석과 동일한 보유 시간을 가지지만(3.26분), 부수 산물(B-2, 카바메이

트)의 보유 시간은 3.86분으로 변하였으며, 이는 유리 아민 모이어티의 화학적 기능화와 일치하였다. 도 14b는 피크 B-2의 질량 스펙트럼(보유 시간 3.86분)을 도시한다. 3.86분에서 피크($M + 134$)의 분석은 예측된 기능화가 발생했음을 나타낸다.

도 15a는 주된 산물(우레아) 및 화학적으로 변형된 부수 산물(카바메이트) 둘 모두를 포함하는 반응 혼합물을 T1 엔도뉴클레아제로 분해시킨 후의 단편 혼합물의 LC-MS 추적을 도시한다. 우레아 연결(G35-[UR]-C36) 및 화학적으로 변형된 카바메이트 연결(G35-[CA+PAA]-C36)이 각각 4.31분 및 5.77분의 보유 시간에서 검출되었다. 도 15b는 4.31분, $m/z = 532.13$ 이 $[M-2H]^{2-}$ 에 할당된 피크의 질량 스펙트럼을 도시하며, 도 15c는 5.77분, $m/z = 599.15$ 가 $[M-2H]^{2-}$ 에 할당된 피크의 질량 스펙트럼을 도시한다. 도 15d 및 도 15e는 도 15b의 $m/z = 532.1$ 및 도 15c의 $m/z = 599.1$ 의 LC-MS/MS 충돌 유도 해리(CID) 실험을 도시한다. 도 15d에서, 통상적인 a-d 및 x-z 이온이 관찰되었으며, 5' 말단($m/z = 487.1$ 및 461.1) 및 3' 말단($m/z = 603.1$ 및 577.1)으로부터 UR 연결의 어느 한쪽에 MS/MS 단편 이온이 관찰되었다. 도 15e에서, 카바메이트 연결의 5' 말단($m/z = 595.2$) 및 CA 연결의 3' 말단($m/z = 603.1$)으로부터의 MS/MS 단편 이온을 포함하는 단지 2개의 산물 이온이 관찰되었다.

도 16a는 동일한 5' 가이드 분자의 변형되지 않은 버전과의 반응의 미정제 반응 혼합물(하부 스펙트럼)과 비교한 2'-H 변형된 5' 가이드 분자 단편과의 반응의 미정제 반응 혼합물(상부 스펙트럼)의 LC-MS 데이터를 도시한다. 2'-H 변형된 5' 가이드 분자 단편의 경우 카바메이트 부산물 형성이 관찰되지 않는다(상부 스펙트럼). 이와 달리, 동일한 5' 가이드 분자 단편의 변형되지 않은 버전과의 반응의 미정제 반응 혼합물(하부 스펙트럼)은 우레아 연결된 주된 산물(A-2) 및 부수 카바메이트 부산물(A-1)의 혼합물을 포함하였다. 본 발명자들은 실시예 10과는 달리, 카바메이트 부산물은 다량 존재하지 않았으며, 따라서 실시예 10의 도 13a보다 훨씬 더 낮은 수준에서 검출되었음에 주목하였다. 도 16b는 피크 B의 디콘볼루션된 질량 스펙트럼(보유 시간 3.14분, 도 16a의 상부 스펙트럼)을 도시하고, 도 16c는 피크 A-2의 디콘볼루션된 질량 스펙트럼(보유 시간 3.45분, 도 16a의 하부 스펙트럼)을 도시한다. 2'-H 변형된 5' 가이드 분자 단편과의 반응의 산물(B)의 분석은 M - 16을 제공하였으며(변형되지 않고 우레아 연결된 주된 산물인 A-2와 비교), 이는 2'-OH가 2'-H로 대체된 분자에서 예측된 바와 같다(도 16b 및 도 16c 참조).

도 17a는 gRNA 1L의 T1 엔도뉴클레아제 분해 후의 LC-MS 추적을 도시하고, 도 17b는 4.65분의 보유 시간의 피크의 질량 스펙트럼(A34:G39)을 도시한다. 특히, 우레아 연결 A-[UR]-AAUAG(A34:G39)를 포함하는 단편이 4.65분의 보유 시간, $m/z = 1182.7$ 에서 검출되었다.

발명을 실시하기 위한 구체적인 내용

[0128]

정의 및 약자

[0129]

달리 명시되지 않는 한, 다음의 용어 각각은 본 섹션에서 그와 관련된 의미를 갖는다.

[0130]

부정 관사 단수형("a" 및 "an")은 관련된 명사 중 적어도 하나를 지칭하며, 용어 "적어도 하나" 및 "하나 이상"과 상호 혼용 가능하게 사용된다. 예를 들어, "하나의 모듈"은 적어도 하나의 모듈 또는 하나 이상의 모듈을 의미한다.

[0131]

접속사 "또는" 및 "및/또는"은 비배타적 논리합으로 상호 혼용 가능하게 사용된다.

[0132]

어구 "필수적으로 이루어진"은 열거된 종류가 주된 종류이지만, 다른 종류가 본 조성물의 구조, 기능 또는 양상에 영향을 미치지 않는 양 또는 미량으로 존재할 수 있음을 의미한다. 예를 들어, 특정 종류로 필수적으로 이루어진 조성물은 일반적으로 이러한 종류를 90%, 95%, 96% 이상(질량 또는 몰농도) 포함할 것이다.

[0133]

어구 "분자를 실질적으로 무함유하는"은 이러한 분자가 열거된 조성물 중의 주요 성분이 아니라는 것을 의미한다. 예를 들어, 분자를 실질적으로 무함유하는 조성물은 이러한 분자가 조성물 중 5%, 4%, 3%, 2%, 1%, 0.5%, 또는 0.1%(질량 또는 몰농도) 미만임을 의미한다. 분자의 양은, 예를 들어 실시예에 기재된 바와 같이 다양한 분석 기법에 의해 측정될 수 있다. 일부 구현예에서, 본 명세서에 제공된 조성물은 특정 분자를 실질적으로 무함유하고, 이러한 분자는 5%, 4%, 3%, 2%, 1%, 0.5%, 또는 0.1%(질량 또는 몰농도) 미만이며, 이는 젤 전기영동에 의해 측정되는 바와 같다. 일부 구현예에서, 본 명세서에 제공된 조성물은 특정 분자를 실질적으로 무함유하고, 이러한 분자는 5%, 4%, 3%, 2%, 1%, 0.5%, 또는 0.1%(질량 또는 몰농도) 미만이며, 이는 질량 분광분석에 의해 측정되는 바와 같다.

[0134]

"도메인"은 단백질 또는 핵산의 절편을 기재하기 위해 사용된다. 달리 명시하지 않는 한, 도메인은 임의의 특정

기능적 특성을 가질 필요가 없다.

[0135] 용어 "상보성"은 수소결합을 통해 안정한 염기쌍을 형성할 수 있는 뉴클레오타이드 쌍을 지칭한다. 예를 들어, U는 A에 상보성이고, G는 C에 상보성이다. 특정 쌍의 상보성 뉴클레오타이드가(예를 들어, 가이드 분자 이중가닥 내의) 수소 결합에 의한 염기쌍 형성을 통해 결합되었는지가 환경(예를 들어, 주위 뉴클레오타이드 및 화학적 연결) 및 외부 조건(예를 들어, 온도 및 pH)에 좌우될 수 있음이 당업자에게 이해될 것이다. 따라서, 상보성 뉴클레오타이드는 반드시 수소 결합에 의한 염기쌍 형성을 통해 결합되는 것은 아니라는 것을 이해해야 한다.

[0136] "공변이체" 서열은 기준 서열의 하나 이상의 뉴클레오타이드의 상보성 뉴클레오타이드로의 치환(예를 들어, 하나 이상의 U의 A로의 대체, 하나 이상의 G의 C로의 대체 등)에 의해 기준 서열과 상이하다. 이중가닥(예를 들어, 가이드 분자의 상부 줄기)을 형성하는 2개의 상보성 서열을 포함하는 영역과 관련하여 사용되는 경우, 용어 "공변이체"는 기준 이중가닥의 2개의 상보성 서열 사이에 하나 이상의 뉴클레오타이드가 대체된 이중가닥을 포괄하며(즉, 하나 이상의 A-U 교환 및/또는 하나 이상의 G-C 교환), 이는 아래 표 1에 예시된 바와 같다:

[0137]

[표 1]

3개의 뉴클레오타이드 서열의 공변이체 서열.

A----U	U----A
G----C	G----C
C----G	C----G
A----U	A----U
C----G	G----C
C----G	G----C
U----A	U----A
C----G	G----C
C----G	G----C
A----U	U----A
C----G	C----G
G----C	G----C

[0138]

[0139] 일부 구현예에서, 공변이체 서열은 기준 서열과 실질적으로 동일한 특정 어닐링 반응(예를 들어, 본 개시내용의 가이드 분자와 관련된 이중가닥 형성)의 에너지적 경향성을 실질적으로 나타낼 수 있다. 본 명세서의 다른 부분에 기재된 바와 같이, 특정 어닐링 반응의 에너지적 경향성은 경험을 기반으로 측정하거나 컴퓨터 모델을 사용하여 예측할 수 있다.

[0140]

"삽입결실"은 핵산 서열 내의 삽입 및/또는 결실이다. 삽입결실은 본 개시내용의 게놈 편집 시스템에 의해 형성된 이중 가닥 파손과 같은 DNA 이중 가닥 파손의 수선 산물일 수 있다. 삽입결실은 하기에 기재되는 NHEJ 경로와 같은 "오류가 발생하기 쉬운" 수선 경로에 의해 파손을 수선하는 경우, 가장 일반적으로 형성된다.

[0141]

"유전자 변환"은 내인성 상동성 서열(예를 들어, 유전자 배열 내의 상동성 서열)의 혼입에 의한 DNA 서열의 변이를 지칭한다. "유전자 보정"은 외인성 단일 또는 이중 가닥 공여체 주형 DNA와 같은 외인성 상동성 서열의 혼입에 의한 DNA 서열의 변이를 지칭한다. 유전자 변환 및 유전자 보정은 하기에 기재되는 것과 같은 HDR 경로에 의한 DNA 이중 가닥 파손의 수선 산물이다.

[0142]

삽입결실, 유전자 변환, 유전자 보정, 및 다른 게놈 편집 결과는 통상적으로 시퀀싱에 의해(Sanger 시퀀싱이 여전히 사용될 수 있지만, 가장 공통적으로는 "차세대" 또는 "합성에 의한 시퀀싱" 방법에 의해) 평가하며, 모든

시퀀싱 판독 중 대상 부위에서의 수적 변화(예를 들어, ±1개, ±2개 이상의 염기)의 상대적 빈도로 정량화한다. 시퀀싱을 위한 DNA 샘플은 당업계에 공지된 다양한 방법에 의해 제조할 수 있으며, 중합효소 연쇄 반응(PCR), 본 명세서에 참조로 포함되는 문헌[Tsai et al. (Nat. Biotechnol. 34(5): 483 (2016))]에 기재된 GUIDEseq 과정에서와 같은 이중 가닥 파손에 의해 생성된 DNA 말단의 포획 또는 당업계에 널리 공지된 다른 수단에 의한 대상 부위의 증폭을 포함할 수 있다. 게놈 편집 결과는 또한 Genomic Vision(Baguenoux, France)에 의해 시판되는 FiberComb™ 시스템과 같은 현장 혼성화 방법 및 당업계에 공지된 임의의 다른 적합한 방법에 의해 평가할 수 있다.

[0143] "Alt-HDR", "대안적 상동성 직접 수선", 또는 "대안적 HDR"은 상동성 핵산(예를 들어, 내인성 상동성 서열, 예를 들어 자매 염색분체, 또는 외인성 핵산, 예를 들어 주형 핵산)을 사용하여, DNA 손상을 수선하는 방법을 지칭하는 것으로 혼용하여 사용된다. Alt-HDR은, 상기 과정이 정규 HDR과는 상이한 경로를 사용하며, 정규 HDR 매개인자, RAD51 및 BRCA2에 의해 저해될 수 있다는 점에서, 정규 HDR과 구별된다. Alt-HDR은 또한 단일 가닥 또는 낙형성(nicked) 상동성 핵산 주형의 관여에 의해 구별되며, 정규 HDR은 일반적으로 이중 가닥의 상동성 주형을 포함한다.

[0144] "정규 HDR", "정규 상동성 직접 수선" 또는 "cHDR"은 상동성 핵산(예를 들어, 내인성 상동성 서열, 예를 들어 자매 염색분체, 또는 외인성 핵산, 예를 들어 주형 핵산)을 사용하여, DNA 손상을 수선하는 방법을 지칭한다. 정규 HDR은 통상적으로, 이중 가닥 파손에 상당한 절삭(resection)이 존재하는 경우, 작용함으로써, DNA의 적어도 하나의 단일 가닥 부분을 형성한다. 정상 세포에서, cHDR은 통상적으로 일련의 단계, 예를 들어 파손의 인식, 파손의 안정화, 절삭, 단일 가닥 DNA의 안정화, DNA 크로스오버 중간체의 형성, 크로스오버 중간체의 풀림 및 결찰을 포함한다. 이 방법에는 RAD51 및 BRCA2가 필요하며, 상동성 핵산은 통상적으로 이중 가닥이다.

[0145] 달리 명시되지 않는 한, 본 명세서에 사용되는 바와 같이, 용어 "HDR"은 정규 HDR 및 alt-HDR을 포함한다.

[0146] "비상동성 말단 결합" 또는 "NHEJ"는 결과적으로 마이크로상동성-매개 말단 결합(MMEJ), 단일 가닥 어닐링(SSA), 및 합성 의존적 마이크로상동성-매개 말단 결합(SD-MMEJ)을 포함하는 정규 NHEJ(cNHEJ), 및 대안적 NHEJ(altNHEJ)을 포함하는 비주형 매개 수선 및/또는 결찰 매개 수선을 지칭한다.

[0147] "대체" 또는 "대체된"은 문자(예를 들어, 핵산 또는 단백질)의 변형과 관련하여 사용되는 경우, 과정 제한이 요구되지는 않지만, 이는 단지 대체 물질이 존재함을 나타낸다.

[0148] "대상체"는 인간 또는 비 인간 동물을 의미한다. 인간 대상체는 임의의 연령(예를 들어, 유아, 소아, 청소년 또는 성인)일 수 있고, 질병을 앓고 있을 수 있거나 유전자 변형이 필요할 수 있다. 대안적으로, 대상체는 포유류, 조류, 어류, 파충류, 양서류, 및 더욱 특히 비 인간 영장류, 설치류(예컨대, 마우스, 랙트, 햄스터 등), 토끼, 기니피그, 개, 고양이 등을 포함하지만 이에 한정되지 않는 동물일 수 있다. 본 개시내용의 특정 구현예에서, 대상체는 가축, 예를 들어 소, 말, 양 또는 염소이다. 특정 구현예에서, 대상체는 가금류이다.

[0149] "치료하다", "치료하는" 및 "치료"는 대상체(예를 들어, 인간 대상체)에서, 질병을 억제하는 것, 즉 질병의 발생 또는 진행을 저지하거나 예방하는 것; 질병을 완화시키는 것, 즉, 질병 상태의 퇴행을 야기하는 것; 질병의 하나 이상의 증상을 완화하는 것; 및 질병을 치유하는 것 중 하나 이상을 포함하는 질병의 치료를 의미한다.

[0150] "예방하다", "예방하는" 및 "예방"은 포유류, 예를 들어 인간에서, (a) 질병을 배제하거나 방지하는 것; (2) 질병에 대한 소인에 영향을 주는 것; 및 (c) 질병의 적어도 하나의 증상을 예방하거나 발병을 지연시키는 것을 포함하는 질병의 예방을 지칭한다.

[0151] "키트"는 특정 목적을 위해 사용될 수 있는 기능적 단위를 함께 구성하는 2개 이상의 성분의 임의의 집합체를 지칭한다. 예시로서(및 제한적이지 않음), 본 개시내용에 따른 하나의 키트는 RNA-가이드 뉴클레아제와 복합체를 형성하거나 복합체를 형성할 수 있는 가이드 RNA를 포함할 수 있고, 약제학적으로 허용 가능한 담체에 수용될 수 있다(예를 들어, 담체 중에 혼탁되거나 혼탁 가능한 것일 수 있음). 키트는, 예를 들어 세포 또는 대상체에서 적절한 게놈 변이를 유발하기 위해, 이러한 세포 또는 대상체에 복합체를 도입하기 위해 사용될 수 있다. 키트의 성분은 함께 패키징되거나(packaged) 별도로 패키징될 수 있다. 본 개시내용에 따른 키트는 또한 선택적으로, 예를 들어 본 개시내용의 방법에 따른 키트의 용도를 기재하는 사용 지침(DFU)을 포함한다. DFU는 키트와 함께 물리적으로 패키징되거나, 예를 들어 전자 수단을 통해 키트 사용자가 이용 가능하도록 할 수 있다.

[0152] 용어 "폴리뉴클레오타이드", "뉴클레오타이드 서열", "핵산", "핵산 분자", "핵산 서열", 및 "올리고뉴클레오타이드"는 DNA 및 RNA의 일련의 뉴클레오타이드 염기(또한 "뉴클레오타이드"로 언급됨)를 지칭하며, 2개 이상의 뉴클레오타이드의 임의의 사슬을 의미한다. 폴리뉴클레오타이드, 뉴클레오타이드 서열, 핵산, 등은 단일 가닥

또는 이중 가닥의 이들의 키메라 혼합물 또는 유도체 또는 변형된 베전일 수 있다. 이들은, 예를 들어 분자의 안정성, 그의 혼성화 매개변수, 등을 개선하기 위해, 염기 모이어티, 당 모이어티, 또는 포스페이트 백본에서 변형될 수 있다. 뉴클레오타이드 서열은 통상적으로 단백질 및 효소 제조를 위한 세포 기전에 사용되는 정보를 포함하지만, 이에 제한되지 않는 유전 정보를 보유한다. 이를 용어는 이중 또는 단일 가닥 게놈 DNA, RNA, 임의의 합성 및 유전자 조작 폴리뉴클레오타이드, 및 센스 및 앤티센스 폴리뉴클레오타이드 둘 모두를 포함한다. 이를 용어는 또한 변형된 염기를 포함하는 핵산을 포함한다.

[0153] 종래의 IUPAC 표기법은 하기 표 2에 제시된 바와 같이 본 명세서에 제시된 뉴클레오타이드 서열에 사용된다(또한 본 명세서에 참조로 포함되는 문헌 [Cornish-Bowden A, Nucleic Acids Res. 1985 May 10; 13(9):3021-30] 참조). 그러나 "T"는 서열이 DNA 또는 RNA에 의해 인코딩될 수 있는 경우의 예에서, 예를 들어 가이드 분자 표적화 도메인에서 "티민 또는 우라실"을 나타낸다는 점에 유의해야 한다.

[0154] [표 2]

IUPAC 핵산 표기법

부호	염기
A	아데닌
T	티민 또는 우라실
G	구아닌
C	시로신
U	우라실
K	G 또는 T/U
M	A 또는 C
R	A 또는 G
Y	C 또는 T/U
S	C 또는 G
W	A 또는 T/U
B	C, G 또는 T/U
V	A, C 또는 G
H	A, C 또는 T/U
D	A, G 또는 T/U
N	A, C, G 또는 T/U

[0155]

[0156] 용어 "단백질," "펩타이드" 및 "폴리펩타이드"는 상호 혼용 가능하게 사용되어, 웨타이드 결합을 통해 함께 연결된 아미노산의 연속적인 사슬을 지칭한다. 이 용어는 개별 단백질, 함께 결합한 단백질의 그룹 또는 복합체뿐만 아니라, 이러한 단백질의 단편 또는 일부분, 변이체, 유도체 및 유사체를 포함한다. 웨타이드 서열은 좌측의 아미노 또는 N 말단에서 시작하여, 우측의 카복실 또는 C 말단으로 진행하는 통상적인 표기법을 사용하여 본 명세서에 제시된다. 한 글자 또는 세 글자의 표준 약어를 사용할 수 있다.

[0157] 용어 "변이체"는 기준 물질과 유의한 구조적 동일성을 나타내지만, 기준 물질과 비교하여, 하나 이상의 화학적 모이어티의 존재 또는 수준에서 기준 물질과 구조적으로 상이한 물질, 예컨대 폴리펩타이드, 폴리뉴클레오타이드 또는 소분자를 지칭한다. 다수의 구현예에서, 변이체는 또한 그의 기준 물질과 기능적으로 상이하다. 일반적으로 특정 물질이 기준 물질의 "변이체"로 적절하게 고려되는지는 기준 물질과의 구조적 동일성 정도를 기반으로 한다.

[0158] 개요

[0159] 본 개시내용의 특정 구현예는 일반적으로 2개 이상의 가이드 단편이 (a) 서로 어닐링된 다음, (b) 적절한 교차 연결 화학반응을 사용하여 교차 연결된 가이드 분자를 합성하는 방법에 관한 것이다. 본 발명자들은 가이드 단편의 교차 연결 전에 이를 사전 어닐링시키는 단계를 포함하는 방법이 교차 연결의 효율을 개선하고, 심지어 동종다중기능성 교차 링커가 사용되는 경우에도 적절한 이종이량체 산물의 형성 경향성이 있음을 발견하였다. 임의의 이론에 결부시키고자 하는 것은 아니지만, 교차 연결 효율의 개선, 결과적으로 적절한 반응 산물의 생성의

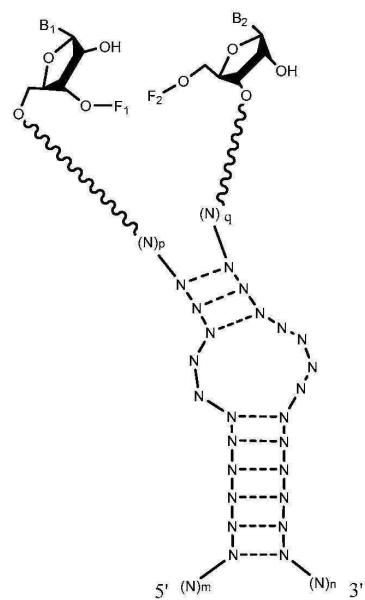
개선은 사전 어닐링에 의해 달성된 동종이량체 등의 형성에 적용 가능한 유리 RNA 단편의 감소 및/또는 어닐링되지 않은 동종이량체와 비교하여, 교차 연결 기질로서 어닐링된 이종이량체의 증가된 안정성으로 인한 것으로 고려된다.

[0160] 가이드 단편의 사전 어닐링 단계를 포함하는 본 개시내용의 방법은 제한 없이 다음을 포함하여, 다수의 이점을 갖는다: 이는 단편이 본 명세서에 기재된 우레아-기반 교차 연결 방법에 사용되는 아민 기능화 단편과 같은 동종다중기능성(예를 들어, 동종이중기능성)인 경우에도 높은 수율의 달성을 가능하게 하고; 부적절한 동종이량체 및 다른 반응 산물의 감소 또는 부재는 결과적으로 하류 정제 단계를 단순화할 수 있으며; 교차 연결에 사용되는 단편은 전장 가이드 문자보다 더 짧은 경향이 있기 때문에, 이는 전장 합성 가이드 문자에서 관찰되는 것보다 $n-1$ 종, 절두 종, $n+1$ 종 및 다른 오염물질에 의해 더 낮은 수준의 오염을 나타낼 수 있다.

[0161] 사전 어닐링과 관련하여, 당업자는, 어닐링된 염기의 더 긴 관이 더 짧은 관보다 더 안정하고, 유사한 길이의 2개의 관 사이에서 더 큰 어닐링 정도가 일반적으로 더 큰 안정성과 관련된다는 것을 인식할 것이다. 따라서, 본 개시내용의 특정 구현예에서, 단편 사이의 어닐링 정도가 최대화되고/최대화되거나 기능화된 3' 또는 5' 말단이 서로 및/또는 어닐링된 염기에 밀접하게 위치하도록 단편을 설계한다.

[0162] 이하에서 더욱 상세히 논의되는 바와 같이, 특정 단분자 가이드 문자, 특히 단분자 Cas9 가이드 문자는 비교적 큰 줄기 루프 구조를 특징으로 한다. 예를 들어, 도 1b 및 1c는 단분자 S. 피오게네스 및 S. 아우레우스 gRNA의 2차원 구조를 도시하며, 2개의 gRNA는 일반적으로 "별지"를 갖는 비교적 긴 줄기 루프 구조를 포함한다는 것이 도면으로부터 명백할 것이다. 특정 구현예에서, 합성 가이드 문자는 이러한 줄기 루프 구조 내의 단편들 사이에 교차 연결을 포함한다. 이는 일부 경우에는, 각각 3' 및 5' 말단에서 또는 그 주위에서 상보성 영역을 갖는 제1 및 제2 단편을 교차 연결시킴으로써 달성되고; 교차 연결 반응이 용이하도록 이를 단편의 3' 말단 및 5' 말단을 기능화시키며, 이는 예를 들어 하기 화학식 II 및 III에 나타낸 바와 같다:

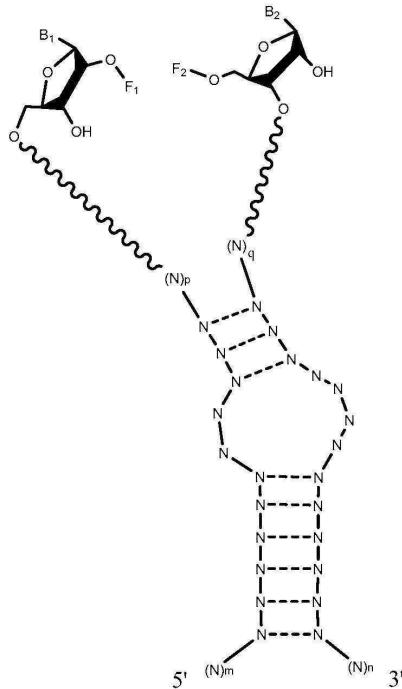
[화학식 II]



[0164]

[0165]

[화학식 III]



[0166]

[0167]

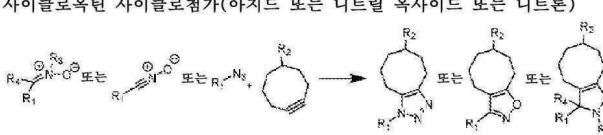
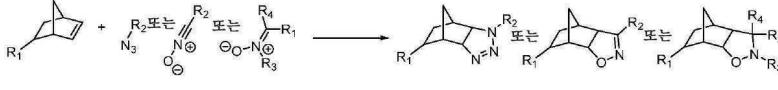
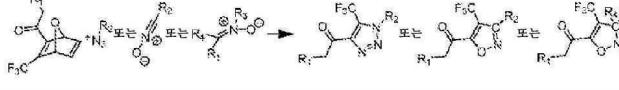
이들 화학식에서, p 및 q 는 각각 독립적으로 0 내지 6이고, $p+q$ 는 0 내지 6이고; m 은 20 내지 40이고; n 은 30 내지 70이고; 각각의 $---$ 는 독립적으로 해당 뉴클레오타이드 사이의 수소결합을 나타내고; 각각의 $\sim\sim$ 는 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 나타내고; $N---N$ 은 독립적으로 2개의 상보성 뉴클레오타이드, 선택적으로, 수소 결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타내고; F_1 및 F_2 는 각각 교차 연결 반응을 거쳐, 2개의 가이드 단편을 교차 연결할 수 있도록 하는 작용기를 포함한다. 예시적 교차 연결 화학반응이 하기 표 3에 제시되어 있다.

[0168]

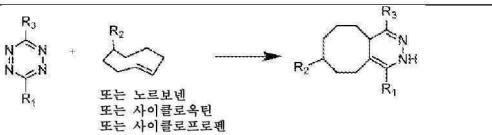
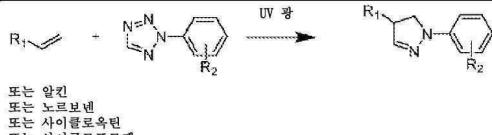
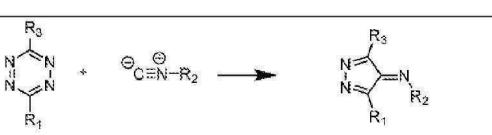
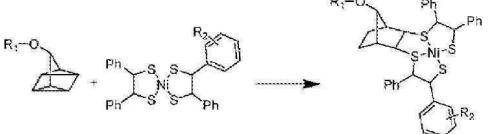
[표 3]

[0169]

예시적 교차 연결 화학반응

반응 유형	반응 요약
티올린	$R_1\equiv + HS-R_2 \longrightarrow R_1-S-R_2$
NHS 에스테르	$\begin{matrix} O \\ \\ R_1-C(=O)-O-N \\ \\ C(=O) \end{matrix} + H_2N-R_2 \longrightarrow R_1-C(=O)-NH-R_2$
티올린	$R_1\equiv + HS-R_2 \xrightarrow{h\nu, cat.} R_1-CH_2-S-R_2$
이소시아네이트	$R_1-NCO + HX-R_2 \longrightarrow R_1-N(CO)X-R_2 \quad X = S \text{ 또는 } NH$
에폭사이드 또는 아지리딘	$\begin{matrix} O \\ \\ R_1 \end{matrix} \text{ 또는 } \begin{matrix} H \\ \\ R_1 \end{matrix} + HS-R_2 \longrightarrow \begin{matrix} HO \\ \\ R_1 \end{matrix}-S-R_2 \text{ 또는 } \begin{matrix} H_2N \\ \\ R_1 \end{matrix}-S-R_2$
알데하이드-아민옥시	$R_1\text{-CHO} + H_2N-O-R_2 \longrightarrow R_1\text{-CH}_2-N(O-R_2)$
Cu-촉매-아지드-알킨 사이클로첨가	$R_1-N_3 + R_2\text{-C}\equiv \xrightarrow{Cu^{\oplus}} \text{cyclic product}$
변형-촉진 사이클로첨가	사이클로옥坦 사이클로첨가(아지드 또는 니트릴 옥사이드 또는 니트론)  노르보넨 사이클로첨가(아지드 또는 니트릴 옥사이드 또는 니트론)  옥사노르보나디엔 사이클로첨가 
Staudinger 결찰	$R_1-N_3 + \begin{matrix} O \\ \\ R_2-C(=O)-OMe \end{matrix} \longrightarrow \text{cyclic product}$

[0170]

테트라진 결찰	
광-유도 테트라졸-알켄 사이클로체가	
[4+1] 사이클로체가	
쿼드리시클란 결찰	

[0171]

화학식 II 및 III은 가이드 문자 반복-반복 방지 이중가닥의 "테트라루프" 구조(또는 "테트라루프" 구조를 대체하는 교차 링커) 내에 위치하는 교차 링커를 도시하며, 교차 링커는 문자, 예를 들어 자연 발생 줄기 루프 및 조작된 줄기 루프를 포함하는 가이드 문자 내에 발생하는 임의의 줄기 루프 구조 내의 임의의 부위에 위치할 수 있음이 이해될 것이다. 특히, 본 개시내용의 특정 구현예는 제1 및 제2 상보성 영역의 말단(예를 들어, 제1 상부 줄기 영역의 3' 말단 및 제2 상부 줄기 영역의 5' 말단)에 위치하는 교차 링커를 포함하고, 테트라루프 구조를 포함하지 않는 가이드 문자에 관한 것이다.

[0172]

화학식 II 및 III은 반복-반복 방지 이중가닥의 "테트라루프" 구조를 포함할 수 있거나($p > 0$ 및 $q > 0$), 포함하지 않을 수 있는($p = 0$ 및 $q = 0$) 가이드 문자를 나타낸다. 본 발명의 일 양태는 "테트라루프"를 포함하지 않는 가이드 문자가 밀접 부위에 적절한 배향의 기능화된 3' 및 5' 말단을 가진 결과로서, 증진된 결찰 효율을 나타낼 수 있다는 인식에 있다.

[0173]

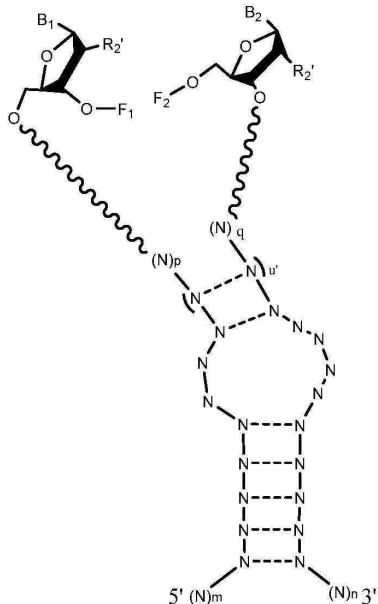
대안적으로, 또는 추가로, 본 개시내용에 따른 교차 연결 반응은 기능화된 3' 및 5' 말단 또는 그 주위의 서열에 혼성화하는 "부목" 또는 단일 가닥 올리고뉴클레오타이드를 포함할 수 있으며, 이는 기능화된 말단이 서로 인접한 부위에 안정하게 존재하도록 하기 위함이다.

[0174]

본 발명의 또 다른 양태는 더 긴 이중가닥의 가이드 문자(예를 들어, 연장된 상부 줄기 포함)는 더 짧은 이중가닥의 가이드 문자와 비교하여, 증진된 결찰 효율을 나타낼 수 있다는 인식에 있다. 더 긴 이들 이중가닥 구조는 본 개시내용에서 "연장된 이중가닥"으로서 상호 혼용 가능하게 지칭되며, 일반적으로(반드시 그러한 것은 아니나) 가이드 단편의 기능화된 뉴클레오타이드의 인접 부위에 위치한다. 따라서, 일부 구현예에서, 본 개시내용은 하기 화학식 VIII 및 IX의 가이드 문자를 제공한다:

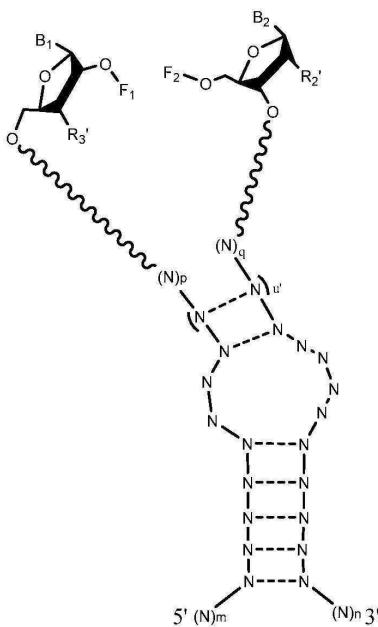
[0176]

[화학식 VIII]



[0177]

[화학식 IX]



[0179]

[0180]

화학식 VIII 및 IX에서, p' 및 q' 는 각각 독립적으로 정수 0 내지 4(상하한 포함)이고, $p'+q'$ 는 정수 0 내지 4(상하한 포함)이고, u' 는 정수 2 내지 22(상하한 포함)이고, 다른 변수는 화학식 II 및 III에 규정된 바와 같다. 화학식 VIII 및 IX는 선택적 테트라루프(즉, p 및 q 는 0인 경우)뿐만 아니라, 선택적으로 연장된 상부 줄기를 포함하는 이중가닥을 도시한다. 화학식 VIII 및 IX의 가이드 문자는 더 긴 상부 줄기로부터 야기되는 증가된 결찰 효율로 인해 유리할 수 있다. 추가로, 더 긴 상부 줄기 및 테트라루프의 부재의 조합은 결찰 반응에서 반응성 기 F_1 및 F_2 의 적절한 배향을 달성하는 데 유리할 수 있다.

[0181]

본 발명의 또 다른 양태는 가이드 단편이 단일 가이드 단편 내 및/또는 상이한 가이드 단편 간에 다수의 상보성 영역을 포함할 수 있다는 인식에 관한 것이다. 예를 들어, 본 개시내용의 특정 구현예에서, 완전히 어닐링되는 경우, (a) 제1 및 제2 작용기가 교차 연결 반응에 적절한 근위의 이중가닥 상부 줄기 영역의 말단에 위치하고/하거나, (b) 가이드 문자와 RNA-가이드 뉴클레아제 사이의 복합체의 형성을 지원할 수 있는 이중가닥 구조가 제1 및 제2 가이드 단편 사이에 형성된 이중이량체가 생성되는 상보성 상부 및 하부 줄기 영역을 갖는 제1 및 제2 가이드 단편을 설계한다. 그러나, 제1 및 제2 가이드 단편이 서로 불안전하게 어닐링되거나 내부 이중가닥 또는 동종이량체를 형성하는 것이 가능할 수 있으며, (a) 및/또는 (b)는 이루어지지 않는다. 일 예로서, 야생형

crRNA 및 tracrRNA 서열을 기반으로 하는 S. 피오게네스 가이드 분자에서 상부 줄기 영역이 서로 적절하게 어닐링되며 이중가닥을 형성하는 경우에, 이중가닥의 형성을 감소시킬 수 있는 하부 및 상부 줄기 사이에 어닐링된 것을 포함하는 부적절한 "엇갈린" 이중가닥을 야기할 수 있다. 유사하게, 가이드 단편의 표적화 도메인 서열과 동일한 가이드 단편 또는 상이한 단편의 또 다른 영역 사이에 부적절한 이중가닥이 형성될 수 있으며, 달리는, 제1 및 제2 가이드 단편의 상보성 영역 사이에 쌍 형성 오류가 발생함으로써, 잠재적으로, 불완전한 이중가닥, 별지 및/또는 비쌍 절편이 생성될 수 있다.

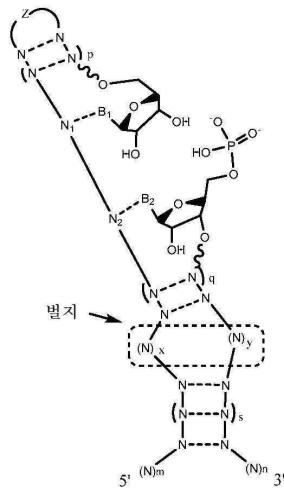
[0182] 가이드 단편 사이에 형성될 수 있는 모든 가능한 부적절한 내부 또는 문자간 이중가닥 구조를 예측하는 것은 실행 가능하지 않지만, 본 발명자들은 일부 경우, 특정 쌍 형성 오류 또는 부적절한 이중가닥의 형성을 감소시키거나 방지하기 위해, 유발시킨 변형이 교차 연결 반응에서 적절한 가이드 분자 산물의 수율에 유의한 영향을 미칠 수 있고/있거나 동일한 반응으로부터 하나 이상의 오염물질 종의 감소를 야기할 수 있음을 발견하였다. 따라서, 일부 구현예에서, 본 개시내용은 (예를 들어, 제1 및 제2 가이드 단편 사이에 2개의 상보성 뉴클레오타이드를 교환시킴으로써) 가이드 단편의 주 서열이 2개 또는 특정 쌍 형성 오류 또는 부적절한 이중가닥이 방지되도록 설계된 가이드 분자 및 방법을 제공한다. 예를 들어, 상기에 언급된 야생형 S. 피오게네스 가이드 단편의 상부 줄기에서의 A-U 교환은 동일하지 않은 UUUU 및 UAUU 서열을 포함하는 제1 가이드 단편 및 제1 단편의 변형된 서열에 상보성인 서열 즉, AAAA 및 AUAA 서열을 포함하는 제2 가이드 단편을 생성시킨다. 더욱 광범위하게, 가이드는 서열 변화, 예컨대 상부 또는 하부 줄기의 2개의 이중가닥 부분 사이의 뉴클레오타이드 교환, 상부 또는 하부 줄기에서의 서열의 삽입, 결실 또는 대체 또는 구조적 변화, 예컨대 2차 구조의 형성을 감소시키거나 제거하기 위해 선택된 위치에 잠긴 핵산(LNA)의 도입을 혼입시킬 수 있다.

[0183] 임의의 이론에 결부시키고자 하는 것은 아니지만, 본 명세서에 기재된 이중가닥 연장, 서열 변형 및 구조적 변형은 적절한 이중가닥의 형성을 촉진시키고, 쌍 형성 오류 또는 부적절한 이중가닥의 형성과 비교하여, 적절한 이중가닥의 형성의 에너지적 경향성을 증가시킴으로써 쌍 형성 오류 및 부적절한 이중가닥의 형성을 감소시키는 것으로 여겨진다. 특정 어닐링 반응의 에너지적 경향성은 갑스 자유 에너지(Gibbs free energy)(ΔG)로 나타낼 수 있고; 음의 ΔG 값은 자발적 반응과 관련되고, 제1 반응의 ΔG 가 제2 반응의 ΔG 보다 작은 경우(즉, 더 큰 음의 값을 갖는 경우), 제1 어닐링 반응은 제2 반응보다 더 큰 에너지적 경향성을 갖는다. 예를 들어, 모든 목적을 위해 본 명세서에 참조로 포함되는 문헌[You, Tatourov and Owczarzy, "Measuring Thermodynamic Details of DNA Hybridization Using Fluorescence" Biopolymers Vol. 95, No. 7, pp. 472-486 (2011)]에 기재된 바와 같이, NMR, 형광 퀸칭(fluorescence quenching), UV 흡광도, 열량계측 등을 사용하여, 특정 이중가닥의 열 안정성(용융 양상)을 기반으로 하여, ΔG 를 경험적으로 평가할 수 있다. (예를 들어, "도입" 472 내지 473 페이지 및 "재료 및 방법" 473 내지 475 페이지 참조). 그러나, 정확한 이중가닥화 및 선택된 쌍 형성 오류 또는 부적절한 이중가닥화 반응의 자유 에너지를 평가하기 위해, 컴퓨터 모델을 사용하여 가이드 단편 및 어닐링 반응을 설계하는 경우, 더 실행 가능할 수 있으며, Integrated DNA Technologies(Coralville, Iowa)가 개설한 biophysics.idtdna.com 를 포함하여, 다수의 룰이 이러한 모델링의 수행에 이용될 수 있다. 대안적으로 또는 추가로, 열역학 초인접 이웃 모델(TNN)을 사용한 다수의 알고리즘이 문헌에 기재되어 있다. 예를 들어, 문헌[Tulpan, Andronescu and Leger, "Free energy estimation of short DNA duplex hybridizations," BMC Bioinformatics, Vol. 11, No. 105 (2010)(TNN 모델 및 MultiRNAFold 패키지, Vienna 패키지 및 UNAFold 패키지가 기재된 1 내지 2 페이지의 "배경기술" 참조)]을 참조한다. 다른 알고리즘이 또한 각각의 참조 문헌의 전문이 모든 목적을 위해 참조로 포함되는 문헌, 예를 들어 문헌[Kim et al. "An evolutionary Monte Carlo algorithm for predicting DNA hybridization," J. Biosystems Vol. 7, No. 5 (2007)(이러한 모델이 기재된 71 내지 72 페이지의 섹션 2 참조)]에 기재되었다.

[0184] 작용기가 다수의 탄소를 포함하는 연결기에 위치하는 화학식 II 및 III에 도시된 정렬이 특히 유리할 수 있다. 규모가 덜 큰 교차 링커의 경우, 기능화된 3' 및 5' 말단 사이에 인접한 첨가를 달성하는 것이 적절할 수 있다. 도 3c 및 도 3d에 제한 없이 포스포디에스테르 결합을 포함하는 더 짧은 링커를 사용하는 경우에 적절한 S. 피오게네스 및 S. 아우레우스 gRNA의 이중가닥 부분이 표시되어 있다. 이들 위치는 일반적으로, 단편 사이에 어닐링이 가능하고, 기능화된 3' 및 5' 말단이 교차 연결 전에 서로 바로 인접하여 있도록 위치하는 것으로 선택된다. 어닐링된 잔기의 관 내에 위치한(그에 인접하기 보다는) 예시적 3' 및 5' 위치가 다음의 화학식 IV, V, VI 및 VII에 제시되어 있다:

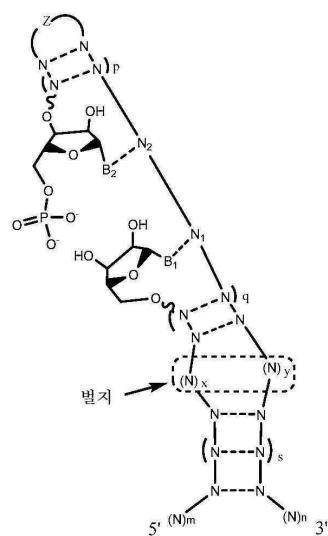
[0185]

[화학식 IV]



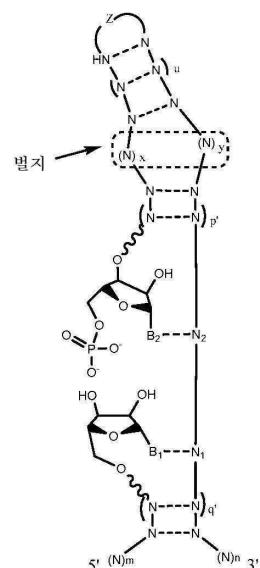
[0186]

[화학식 V]



[0188]

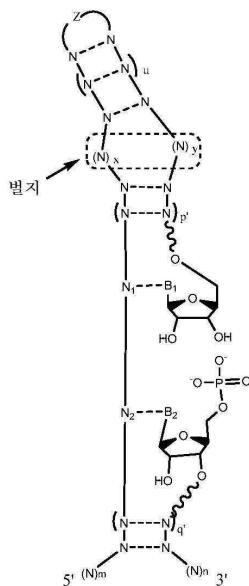
[화학식 VI]



[0190]

[0191]

[화학식 VII]



[0192]

[0193]

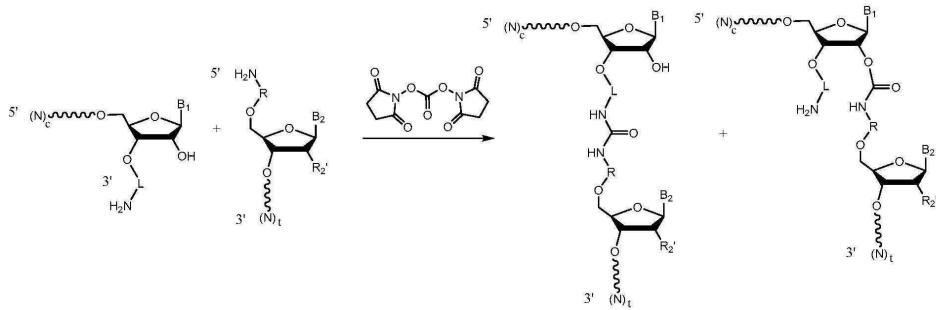
Z는 4개 내지 6개의 뉴클레오타이드 길이, 선택적으로 4개 또는 6개의 뉴클레오타이드 길이의 뉴클레오타이드 루프를 나타내고; p 및 q는 각각 독립적으로 정수 0 내지 2(상하한 포함), 선택적으로 0이고; p'는 정수 0 내지 4(상하한 포함), 선택적으로 0이고; q'는 정수 2 내지 4(상하한 포함), 선택적으로 2이고; x는 정수 0 내지 6(상하한 포함) 선택적으로 2이고; y는 정수 0 내지 6(상하한 포함), 선택적으로 4이고; u는 정수 0 내지 4(상하한 포함), 선택적으로 2이고; s는 정수 2 내지 6(상하한 포함), 선택적으로 4이고; m은 정수 20 내지 40(상하한 포함)이고; n은 정수 30 내지 70(상하한 포함)이고; B₁ 및 B₂는 각각 독립적으로 핵염기이고; (N)_m 및 (N)_n의 각각의 N은 독립적으로 뉴클레오타이드 잔기이고; N₁ 및 N₂는 각각 독립적으로 뉴클레오타이드 잔기이고; N---N은 독립적으로 2개의 상보성 뉴클레오타이드, 선택적으로, 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타내고; 각각의 ~~~~~는 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오프스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 나타낸다.

[0194]

본 발명의 또 다른 양태는 화학식 II, III, IV, V, VI 또는 VII 중 어느 하나에 도시된 정렬이 교차 연결 반응에서 부산물의 방지에 유리할뿐만 아니라, 동종이중기능성 반응이 동종이량체화 없이 이루어질 수 있다는 인식이다. 2개의 이종이량체 가닥의 사전 어닐링은 반응성 기가 적절하게 커플링되도록 유도하고, 가이드 분자 내의 다른 잠재적 반응성 기와의 반응이 일어나지 않도록 한다.

[0195]

본 발명의 또 다른 양태는 반응성 기에 인접한 하이드록실기(예를 들어, 제1 단편의 3' 말단의 2'-OH)를 바람직하게는 특정 부산물의 형성이 방지되도록 변형시킨다는 인식이다. 특히, 다음에 예시된 바와 같이, 본 발명자들은 본 명세서에 기재된 우레아-기반 교차 연결 방법에 아민-기능화된 단편이 사용되는 경우, 카바메이트 부산물이 형성될 수 있음을 발견하였다:



주된 산물(우레아)

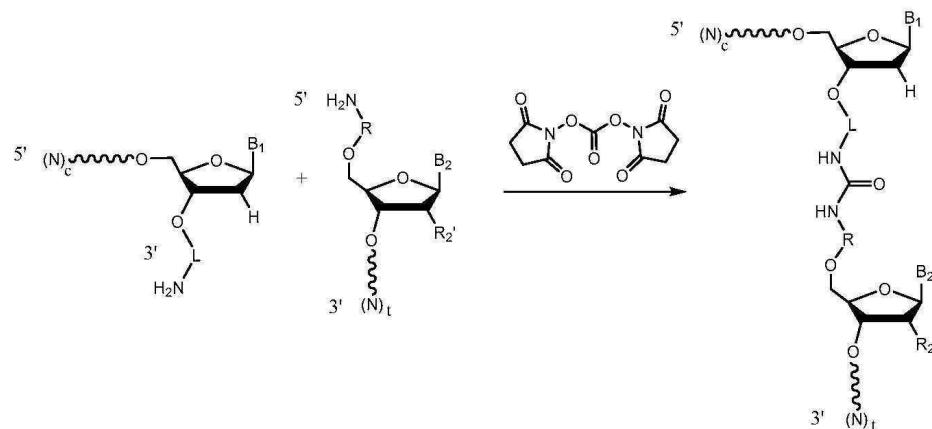
부수 산물(카바메이트)

[0196]

[0197]

따라서, 특정 구현예에서, 카바메이트 부산물의 형성을 방지하기 위해, 제1 단편의 3' 말단의 2'-OH를 (예를 들

어, H, 할로겐, O-Me, 등으로) 변형시킨다. 예를 들어, 2'-OH를 2'-H로 변형시킨다:



단일 산물(우레아)

[0198]

[0199] 다음에 교차 연결로 화제를 돌리자면, 교차 링커 연결 모이어티, 작용기 및 반응성 기의 선택과 관련하여 여러 고려 사항이 존재한다. 특히 링커의 크기, 수용액 중 용해도 및 생체적합성뿐만 아니라, 작용기 반응성, 교차 연결에 대한 최적 반응 조건 및 교차 연결에 필요한 임의의 시약, 촉매 등이 있다.

[0200]

일반적으로, 적절한 RNA 이차 구조가 유지되거나 달성되고, 가이드 분자 및 RNA-가이드 뉴클레아제 사이의 복합체의 파괴 또는 불안정화가 방지되는 링커 크기 및 용해도를 선택한다. 특정 길이 초파의 유기 링커가 수용액 중 용해도가 저조할 수 있고, 가이드 분자 내의 주변 뉴클레오타이드 및/또는 가이드 분자와 복합체화된 RNA-가이드 뉴클레아제의 아미노산과 입체적으로 간섭될 수 있는한, 이를 2개의 인자는 일정 정도 관련된다.

[0201]

[0201] 다수의 링커가 본 개시내용의 다양한 구현예에서 사용하기에 적합하다. 특정 구현예에서 제한 없이, 폴리비닐에 테르, 폴리에틸렌, 폴리프로필렌, 폴리에틸렌 글리콜(PEG), 폴리프로필렌 글리콜(PEG), 폴리비닐 알코올(PVA), 폴리글리콜라이드(PGA), 폴리락티드(PLA), 폴리카프로락톤(PCL), 및 이들의 공중합체를 포함하는 일반적인 연결 모이어티가 사용된다. 일부 구현예에서, 링커가 사용되지 않는다.

[0202]

[0202] 작용기에 대해, 가이드 단편의 5' 및 3' 말단을 연결시키기 위해, 이중기능성 교차 링커가 사용되는 구현예에서, 연결될 가이드 단편의 3' 또는 5' 말단을 교차 링커의 반응성 기와 반응하는 작용기로 변형시킨다. 일반적으로, 이들 변형은 아민, 슬피드릴, 카복실, 하이드록실, 알켄(예를 들어, 말단 알켄), 아지드 및/또는 또 다른 적절한 작용기 중 하나 이상을 포함한다. 다중기능성(예를 들어, 이중기능성) 교차 링커는 또한 일반적으로 당 업계에 공지되어 있고, 이중기능성 또는 동종기능성일 수 있으며, 제한 없이, 이소티오시아네이트, 이소시아네이트, 아실 아지드, NHS 에스테르, 염화술포닐, 토실 에스테르, 트레실 에스테르, 알데하이드, 아민, 에폭사이드, 카보네이트(예를 들어, 비스(p-니트로페닐) 카보네이트), 할로겐화아릴, 할로겐화알킬, 이미도 에스테르, 카복실레이트, 알킬 포스페이트, 무수물, 플루오로페닐 에스테르, HOEt 에스테르, 하이드록시메틸 포스핀, O-메틸이소우레아, DSC, NHS 카바메이트, 글루타르알데하이드, 활성화된 이중결합, 사이클릭 헤미아세탈, NHS 카보네이트, 이미다졸 카바메이트, 아실 이미다졸, 메틸피리디늄 에테르, 아줄락톤, 시아네이트 에스테르, 사이클릭 이미도 카보네이트, 클로로트리아진, 디하이드로아제핀, 6-술포-시토신 유도체, 말레이미드, 아지리딘, TNB 티올, 엘만 시약(Ellman's reagent), 퍼록사이드, 비닐술폰, 폐닐티오에스테르, 디아조알칸, 디아조아세틸, 에폭사이드, 디아조늄, 벤조페논, 안트라퀴논, 디아조 유도체, 디아지린 유도체, 소랄렌 유도체, 알켄, 폐닐 봉산 등을 포함하여, 임의의 적절한 작용기를 포함할 수 있다.

[0203]

[0203] 이들 및 다른 교차 연결 화학반응은 당업계에 공지되어 있으며, 그 전문이 모든 목적을 위해 본 명세서에 참조로 포함되는 문헌[Greg T. Hermanson, Bioconjugate Techniques, 3rd Ed. 2013, published by Academic Press]을 포함하여, 문헌에 정리되어 있다.

[0204]

본 개시내용에 의해 제공되는 방법에 의해 합성되는 가이드 분자를 포함하는 조성물은 특정 구현예에서, n-1 종, 절두, n+1 종, 가이드 단편 동종이량체, 반응하지 않은 기능화된 가이드 단편 등을 포함하여, 부적절한 종에 의한 오염 수준이 낮은 고 순도의 적절한 가이드 분자 반응 산물을 특징으로 한다. 본 개시내용의 특정 구현예에서, 합성 가이드 분자를 포함하는 정제 조성물은 조성물 내에 다수의 종류를 포함할 수 있다(즉, 가이드 분

자는 조성물 내에 가장 일반적인 종류임, 질량 또는 몰농도). 대안적으로, 또는 추가적으로, 본 개시내용의 구현예에 따른 조성물은 적절한 길이(예를 들어, 기준 가이드 분자 서열에 비해 5' 말단의 절두가 부재함) 및 적절한 서열(예를 들어, 기준 가이드 분자 서열의 5' 서열 포함)을 갖는 가이드 분자를 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 및/또는 99% 이상 포함할 수 있다.

[0205]

예를 들어, 일부 구현예에서, 본 개시내용에 따른 가이드 분자(예를 들어, 본 명세서에 기재된 적절한 교차 연결 화학반응을 사용하여 교차 연결된 단편을 포함하는 가이드 분자)를 포함하는 조성물은 기준 가이드 분자 서열과 비교하여, 5' 말단에 절두를 포함하는 가이드 분자를 약 20%, 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1% 이하를 포함한다. 추가적으로 또는 대안적으로, 본 개시내용에 따른 가이드 분자(예를 들어, 본 명세서에 기재된 적절한 교차 연결 화학반응을 사용하여 교차 연결된 단편을 포함하는 가이드 분자)를 포함하는 조성물은 기준 가이드 분자 서열의 상응하는 5' 서열과 100% 동일한 5' 서열(예를 들어, 가이드 분자의 뉴클레오타이드 1 내지 30, 1 내지 25, 또는 1 내지 20을 포함하거나 이로 이루어진 5' 서열)을 갖는 가이드 분자를 적어도 약 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 포함한다. 일부 구현예에서, 본 조성물이 기준 가이드 분자 서열의 상응하는 5' 서열에 100% 미만만큼 동일한 5' 서열을 갖는 가이드 분자를 포함하고, 이러한 가이드 분자가 0.1% 이상의 수준으로 존재하는 경우, 이러한 가이드 분자는 잠재적 표적의 부위에 대한 표적화 도메인을 포함하지 않는다. 일부 구현예에서, 본 개시내용에 따른 가이드 분자를 포함하는 조성물은(기준 가이드 분자 서열에 비해) 5' 말단의 절두를 포함하지 않는 가이드 분자를 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상 포함하고, 적어도 약 90%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 이러한 가이드 분자(즉, 5' 말단의 절두를 포함하지 않는 가이드 분자)는 기준 가이드 분자 서열의 상응하는 5' 서열에 100% 동일한 5' 서열(예를 들어, 가이드 분자의 뉴클레오타이드 1 내지 30, 1 내지 25, 또는 1 내지 20을 포함하거나 이로 이루어진 5' 서열)을 갖고, 본 조성물이 기준 가이드 분자 서열의 상응하는 5' 서열에 100% 미만만큼 동일한 5' 서열을 갖는 가이드 분자를 포함하고, 이러한 가이드 분자가 0.1% 이상의 수준으로 존재하는 경우, 이러한 가이드 분자는 잠재적 표적의 부위에 대한 표적화 도메인을 포함하지 않는다. 일부 구현예에서, 본 개시내용에 따른 가이드 분자를 포함하는 조성물은 (i) 기준 가이드 분자 서열의 상응하는 5' 서열과 100% 동일한 5' 서열(예를 들어, 가이드 분자의 뉴클레오타이드 1 내지 30, 1 내지 25, 또는 1 내지 20을 포함하거나 이로 이루어진 5' 서열)을 갖는 가이드 분자를 적어도 약 99% 포함하고, (ii) 조성물이 기준 가이드 분자 서열의 상응하는 5' 서열에 100% 미만만큼 동일한 5' 서열을 갖는 가이드 분자를 포함하고, 이러한 가이드 분자가 0.1% 이상의 수준으로 존재하는 경우, 이러한 가이드 분자는 잠재적 표적의 부위에 대한 표적화 도메인을 포함하지 않고, 본 조성물은 허용 가능한 수준의 특이성/안전성을 나타낸다. 조성물의 순도는 조성물 내의 총 가이드 분자의 분율(질량 또는 몰농도), 조성물 내의 모든 RNA 또는 모든 핵산의 분율(질량 또는 몰농도), 조성물 내의 모든 용질의 분율(질량) 및/또는 조성물의 총 질량의 분율로서 표시될 수 있다.

[0206]

본 개시내용에 따른 가이드 분자를 포함하는 조성물의 순도는 당업계에 공지된 임의의 적합한 수단에 의해 평가된다. 예를 들어, 적절한 가이드 분자 종의 상대적 존재량은 젤 전기영동에 의해 정성적 또는 반정량적으로 평가될 수 있다. 대안적으로 또는 추가적으로, 적절한 가이드 분자 종의 순도는 크로마토그래피(예를 들어, 액체 크로마토그래피, HPLC, FPLC, 기체 크로마토그래피), 분광분석(예를 들어, 질량 분광분석, 전개 시간, 섹터 필드, 사중극자 질량, 이온 트랩, 오비트랩, 푸리에 변환 이온 사이클로트론 공명 또는 다른 기술 기반 여부불문), 핵 자기 공명(NMR) 분광분석(예를 들어, 가시광선, 적외선 또는 자외선), 열 안정성 방법(예를 들어, 시차 주사 열량계 등), 시퀀싱 방법(예를 들어, 주형 교환 올리고뉴클레오타이드 사용) 및 이들의 조합(예를 들어, 크로마토 그래피-분광분석 등)에 의해 평가된다.

[0207]

본 명세서에 제공되는 합성 가이드 분자는 임의의 다른 가이드 분자(예를 들어, gRNA)와 실질적으로 동일한 방식으로 작동하고, 일반적으로 (a) RNA-가이드 뉴클레아제, 예컨대 Cas9와 복합체를 형성하고, (b) 가이드 분자의 표적화 서열에 상보성이 있는 영역 및 RNA-가이드 뉴클레아제에 의해 인식되는 프로토스페이서(protospacer) 인접 모티프(PAM)를 포함하는 표적 서열과 상호작용하고, 선택적으로 (c) 예를 들어, 가이드 분자 및 RNA-가이드 뉴클레아제를 포함하는 세포 내에서 작용하는 DNA 수선 경로에 의해 수선될 수 있는 DNA 이중 가닥 파손, 단일 가닥 파손 등의 형성에 의해 표적 서열 내 또는 그에 인접한 DNA를 변형시킴으로써 작용한다.

[0208]

일부 구현예에서, 본 명세서에 기재된 가이드 분자, 예를 들어 본 명세서에 기재된 방법을 사용하여 생성된 가이드 분자는 RNA에 작용하는 효소(예를 들어, 역전사효소)의 기질로서 작용할 수 있다. 이론에 결부시키고자 하는 것은 아니지만, 본 명세서에 기재된 가이드 분자 내에 존재하는 교차 링커는 본 개시내용의 방법에 따른 사

전 어닐링에 의해 촉진되는 반응성 말단의 인접 첨가로 인해 이러한 처리 효소에 적합할 수 있다.

[0209]

상기에 기재된 예시적 구현예는 본 명세서에 기재된 합성 및 교차 연결 방법을 2개의 가이드 단편으로부터의 가이드 문자의 조립에 적용하는 것에 중점을 둔다. 그러나 본 명세서에 기재된 방법은 다양한 적용을 가지며, 이는 당업자에 자명할 것이다. 이를 적용은 본 개시내용의 범위 내에 있다. 일 예로서, 본 개시내용의 방법은 가이드 문자와 비상동성 서열의 연결에 사용될 수 있다. 비상동성 서열은 제한 없이, 모든 목적을 위해 본 명세서에 참조로 포함되는 Cotta-Ramusino 등에 의한 제 WO 2017/180711호에 기재된 바와 같은 DNA 공여체 주형을 포함할 수 있다. (예를 들어, 공유결합된 주형 핵산 및 가이드 문자의 3' 말단과 주형의 결찰을 용이하게 하는 고정 올리고의 사용이 기재된 23 페이지의 섹션 I, "gRNA 융합 문자" 참조). 비상동성 서열은 또한 상기의 제 WO 2017/180711호의 섹션 I에 기재된 MS2 루프와 같은 웹타이드 DNA 또는 RNA 결합 도메인에 의해 인식되는 핵산 서열을 또한 포함할 수 있다.

[0210]

이러한 개요는 가이드 문자의 합성에 관한 특정 원리 및 이러한 가이드 문자를 포함하는 조성물을 예시하는 소수의 예시적 구현예에 집중되어 있다. 그러나 명료성을 위해, 본 개시내용은 기재되어 있지는 않지만 당업자에게 자명할 수 있는 변형 및 변화를 포괄한다. 이를 염두에 두고, 하기 개시내용은 게놈 편집 시스템의 작동 원리를 더욱 일반적으로 설명하기 위한 것이다. 하기의 내용은 제한으로서 이해되어서는 안되며, 본 개시내용과 조합하여, 본 개시내용의 추가적 구현예 및 본 개시내용의 범위 내에 있는 변형에 대한 정보를 당업자에 제공할 것인 게놈 편집 시스템의 특정 원리를 설명하는 것이다.

[0211]

게놈 편집 시스템

[0212]

용어 "게놈 편집 시스템"은 RNA-가이드 DNA 편집 활성을 갖는 임의의 시스템을 지칭한다. 본 개시내용의 게놈 편집 시스템은 자연 발생 CRISPR 시스템: 가이드 문자(예를 들어 가이드 RNA 또는 gRNA) 및 RNA-가이드 뉴클레아제로부터 개조된 적어도 2개의 성분을 포함한다. 이들 2개의 성분은 단일 가닥 파손(SSB 또는 낙), 이중 가닥 파손(DSB) 및/또는 점 돌연변이 중 하나 이상의 발생에 의해, 핵산 서열 내 또는 그 주위에 특이적 핵산 서열 및 편집 DNA와 결합될 수 있는 복합체를 형성한다.

[0213]

자연 발생 CRISPR 시스템은 발전적으로 2개의 클래스 및 5개의 유형으로 체계화되고(본 명세서에 참조로 포함되는 문헌[Makarova et al. Nat Rev Microbiol. 2011 Jun; 9(6): 467-477 (Makarova)]), 본 개시내용의 게놈 편집 시스템은 자연 발생 CRISPR 시스템의 임의 유형 또는 클래스의 성분을 채택할 수 있지만, 본 명세서에 제시된 구현예는 일반적으로 클래스 2 및 유형 II 또는 V CRISPR 시스템으로부터 채택된다. 유형 II 및 V를 포괄하는 클래스 2 시스템은 상대적으로 큰 다수의 도메인 RNA-가이드 뉴클레아제 단백질(예를 들어, Cas9 또는 Cpf1) 및 하나 이상의 가이드 RNA(예를 들어, crRNA 및, 선택적으로, tracrRNA)를 특징으로 하고, 이는 crRNA의 표적화(또는 스페이서) 서열에 상보성이 특이적 유전자 좌우에 결합(즉, 표적화)하고, 이를 절단하는 리보핵단백질(RNP) 복합체를 형성한다. 본 개시내용에 따른 게놈 편집 시스템은 유사하게 세포 DNA 서열을 표적화하고 편집 하지만, 자연 발생하는 CRISPR 시스템과는 유의하게 상이하다. 예를 들어, 본 명세서에 기재된 단분자 가이드 분자는 자연 발생하지 않으며, 본 개시내용에 따른 가이드 문자 및 RNA-가이드 뉴클레아제는 둘 모두 임의의 수의 자연 발생적 변형을 포함할 수 있다.

[0214]

게놈 편집 시스템은 다양한 방식으로 구현(예를 들어, 세포 또는 대상체에 투여되거나 전달)될 수 있으며, 상이한 구현예가 별개의 적용에 적합할 수 있다. 예를 들어, 게놈 편집 시스템은 특정 구현예에서, 선택적으로 약제학적으로 허용 가능한 담체 및/또는 캡슐화제, 예컨대 지질 또는 중합체 마이크로 또는 나노입자, 미셀, 리포솜 등을 포함하는 약제학적 조성물 내에 포함될 수 있는 단백질/RNA 복합체(리보핵단백질, 또는 RNP)로서 구현된다. 특정 구현예에서, 게놈 편집 시스템은(선택적으로 하나 이상의 추가적 성분과 함께) 상기에 기재된 RNA-가이드 뉴클레아제 및 가이드 문자 성분을 인코딩하는 하나 이상의 핵산으로서 구현되고; 특정 구현예에서, 게놈 편집 시스템은 이러한 핵산을 포함하는 하나 이상의 벡터, 예를 들어 아데노-연관 바이러스와 같은 바이러스 벡터로서 구현되며; 특정 구현예에서, 게놈 편집 시스템은 전술한 것 중 어느 하나의 조합으로서 구현된다. 본 명세서에 제시된 원리들에 따라 작용하는 추가의 또는 변형된 구현예는 당업자에게 명백할 것이며, 본 개시내용의 범위 내에 있다.

[0215]

본 개시내용의 게놈 편집 시스템은 하나의 특이적 뉴클레오타이드 서열을 표적화할 수 있거나 2개 이상의 가이드 문자의 사용을 통해, 2개 이상의 특이적 뉴클레오타이드 서열을 표적화할 수 있고 - 동시에 편집할 수 있다. 다수의 가이드 문자의 사용은 본 개시내용 전체에 걸쳐 "다중복합화"로 지칭되며, 다수의 비관련 대상 표적 서열을 표적화하거나 단일 표적 도메인 내에 다수의 SSB 또는 DSB를 형성하고, 일부 경우, 이러한 표적 도메인 내에 특이적 편집을 수행하기 위해 사용될 수 있다. 예를 들어, 본 명세서에 참조로 포함되는 Maeder 등의 국제

특히 공개 제 WO 2015/138510호(Maeder)는 결과적으로 유전자의 기능을 감소시키거나 제거하는 불특정 스플라이싱 부위의 생성을 유발하는 인간 CEP290 유전자의 (C.2991+1655A에서 G로의) 점 돌연변이를 수정하기 위한 게놈 편집 시스템을 기재한다. Maeder의 게놈 편집 시스템은 점 돌연변이의 측면(즉, 측접하는) 서열에 표적화된 2개의 가이드 RNA를 사용하며, 돌연변이에 측접하는 DSB를 형성한다. 이는 결과적으로 돌연변이를 포함하는 개재 서열의 결실을 촉진함으로써 불특정 스플라이싱 부위를 제거하고, 정상적인 유전자 기능을 복원시킨다.

[0216] 또 다른 예로서, 본 명세서에 참조로 포함되는 Cotta-Ramusino 등에 의한 제 WO 2016/073990호("Cotta-Ramusino")는 일명 "이중 낙카제 시스템"의 정렬로서 Cas9 낙카제(S. 피오제네스 D10A와 같은 단일 가닥 낙을 형성하는 Cas9)와 조합하여, 2개의 gRNA를 사용하는 게놈 편집 시스템을 기재한다. 문헌[Cotta-Ramusino]의 이 중 낙카제 시스템은 하나 이상의 뉴클레오타이드에 의한 오프셋(offset)인 대상 서열의 맞은편 가닥에 2개의 낙을 형성하도록 구성되며, 낙은 조합하여 오버행(문헌[Cotta-Ramusino]의 경우 5', 또한 3' 오버행이 가능함)을 갖는 이중 가닥 파손을 생성한다. 결과적으로 오버행은 일부 상황에서 상동성 유도 수선 이벤트를 촉진할 수 있다. 그리고 또 다른 예로서, Palestrant 등에 의한 제 WO 2015/070083호("Palestrant", 본 명세서에 참조로 포함됨)는 Cas9를 인코딩하는 뉴클레오타이드 서열에 표적화된 gRNA("통제 RNA로 지칭됨")를 기재하며, 이는 달리는, 예를 들어 일부 바이러스 형질도입 세포에서 항시적으로 발현되는 Cas9의 일시적 발현을 가능하게 하는 하나 이상의 추가적 gRNA를 포함하는 게놈 편집 시스템에 포함될 수 있다. 이러한 다중복합 적용은 제한하는 것이 아닌 예시하기 위한 것이며, 다른 다중복합 적용은 일반적으로 본 명세서에 기재된 게놈 편집 시스템에 적합하다는 것을 당업자는 인식할 것이다.

[0217] 일부 경우, 게놈 편집 시스템은 NHEJ 또는 HDR과 같은 세포 DNA 이중가닥 파손 메커니즘에 의해 수선되는 이중 가닥 파손을 형성할 수 있다. 이를 메커니즘은 문헌, 예를 들어 문헌[Davis & Maizels, PNAS, 111(10):E924-932, March 11, 2014 (Davis) (describing Alt-HDR); Frit et al. DNA Repair 17(2014) 81-97 (Frit) (describing Alt-NHEJ); 및 Iyama and Wilson III, DNA Repair (Amst.) 2013-Aug; 12(8): 620-636 (Iyama)(정규 HDR 및 NHEJ 경로를 일반적으로 기재함)]에 전반적으로 기재되어 있다.

[0218] 게놈 편집 시스템은 DSB의 형성에 의해 작용하지만, 이러한 시스템은 선택적으로 이중가닥 파손 수선의 특정 방식 또는 특정 수선 결과를 촉진하거나 용이하게 하는 하나 이상의 성분을 포함한다. 예를 들어, 문헌[Cotta-Ramusino]은 또한 단일 가닥 올리고뉴클레오타이드 "공여체 주형"이 첨가되고; 게놈 편집 시스템에 의해 절단되고, 표적 서열 내에 변화를 야기할 수 있는 세포 DNA의 표적 영역 내로 공여체 주형이 혼입되는 게놈 편집 시스템을 기재한다.

[0219] 특정 구현예에서, 게놈 편집 시스템은 단일 또는 이중가닥 파손을 야기함이 없이, 표적 서열을 변형시키거나 표적 서열 또는 그 주위의 유전자 발현을 변형시킨다. 예를 들어, 게놈 편집 시스템은 DNA에 작용함으로써 표적 서열 또는 그 발현을 변형시키는 기능적 도메인에 융합된 RNA-가이드 뉴클레아제를 포함할 수 있다. 일 예로서, RNA-가이드 뉴클레아제는 시티딘 탈아미노 효소 기능적 도메인에 연결(예를 들어, 융합)될 수 있으며, 표적화된 C-대-A 치환의 유발에 의해 작용할 수 있다. 예시적 뉴클레아제/탈아미노 효소 융합은 참조로 포함되는 문헌 [Komor et al. Nature 533, 420-424 (19 May 2016) ("Komor")]에 기재되어 있다. 대안적으로, 게놈 편집 시스템은 사멸 Cas9(dCas9)와 같은 절단 불활성화(즉, "사멸")뉴클레아제를 사용할 수 있으며, 세포 DNA의 하나 이상의 표적화된 영역에 안정한 복합체를 형성함으로써, 제한 없이 mRNA 전사, 크로마틴 리모델링 등을 포함하여, 표적화된 영역(들)을 포함하는 기능에 간섭함으로써 작용할 수 있다.

가이드 분자

[0220] 용어 "가이드 분자"는 세포 내의 게놈 또는 에피솜 서열과 같은 표적 서열과 Cas9 또는 Cpf1과 같은 RNA-가이드 뉴클레아제의 특이적 결합(또는 "표적화")을 촉진하는 임의의 핵산을 지칭하는 것으로 본 명세서에서 사용된다. 가이드 분자는 RNA 분자 또는 혼성체 RNA/DNA 분자일 수 있다. 가이드 분자는 단분자(단일 분자를 포함하고, 대안적으로 키메라로 지칭됨) 또는 모듈식(예를 들어 일반적으로 이중가닥화에 의해 서로 결합되는 crRNA 및 tracrRNA와 같은 하나 초파의, 통상적으로 2개의 개별 분자 포함)일 수 있다. 가이드 분자 및 그의 성분 일부가 문헌, 예를 들어 참조로 포함되는 문헌[Briner et al. (Molecular Cell 56(2), 333-339, October 23, 2014 (Briner), 및 Cotta-Ramusino]에 전반적으로 기재되어 있다.

[0221] 박테리아 및 고세균에서, II형 CRISPR 시스템은 일반적으로 crRNA의 3' 영역에 상보성이며, 그와 이중가닥을 형성하는 5' 영역을 포함하는 트랜스-활성화 crRNA(tracrRNA) 및 외래 서열에 상보성인 5' 영역을 포함하는 Cas9, CRISPR RNA(crRNA)와 같은 RNA-가이드 뉴클레아제 단백질을 포함한다. 임의의 이론에 결부시키고자 하는 것은 아니지만, 이러한 이중가닥이 Cas9/가이드 분자 복합체의 형성을 촉진하고 - 그의 활성에 필요한 것으로 여겨진

다. II형 CRISPR 시스템이 유전자 편집에 사용 적용되므로, 하나의 비 제한적 예에서, crRNA(그의 3' 말단) 및 tracrRNA(그의 5' 말단)의 4개의 뉴클레오타이드(예를 들어, GAAA) "테트라루프" 또는 "링커" 서열 가교 상보성 영역에 의해 crRNA 및 tracrRNA가 단일 단분자 또는 키메라 가이드 RNA에 결합될 수 있는 것으로 밝혀졌다. (모두 본 명세서에 참조로 포함되는 문헌[Mali et al., Science. 2013 Feb 15; 339(6121): 823-826 ("Mali"); Jiang et al., Nat Biotechnol. 2013 Mar; 31(3): 233-239 ("Jiang")); 및 Jinek et al., 2012 Science Aug. 17; 337(6096): 816-821 ("Jinek")]).

[0223] 가이드 분자는, 단분자인지 또는 모듈식인지 상관없이, 편집이 요망되는 세포의 계놈 내의 DNA 서열과 같은 표적 서열 내의 표적 도메인에 완전히 또는 부분적으로 상보성인 "표적화 도메인"을 포함한다. 표적화 도메인은 제한 없이 "가이드 서열"(본 명세서에 참조로 포함되는 문헌[Hsu et al., Nat Biotechnol. 2013 Sep; 31(9): 827-832, ("Hsu")]), "상보성 영역"(문헌[Cotta-Ramusino]), "스페이서"(문헌[Briner]) 및 일반적으로 "crRNA"(문헌[Jiang])를 포함하여, 문헌에 다양한 명칭으로 지칭된다. 지정 명칭에 상관없이, 표적화 도메인은 통상적으로 10개 내지 30개의 뉴클레오타이드 길이이고, 특정 구현예에서, 16개 내지 24개의 뉴클레오타이드 길이(예를 들어, 16, 17, 18, 19, 20, 21, 22, 23 또는 24개의 뉴클레오타이드 길이)이고, Cas9 가이드 분자의 경우 5' 말단 또는 그 주위에, Cpf1 가이드 분자의 경우, 3' 말단 또는 그 주위에 존재한다.

[0224] 표적화 도메인 이외에, 가이드 분자는 통상적으로 (반드시 그러한 것은 아니며, 하기에 논의되는 바와 같음) 가이드 분자/Cas9 복합체의 형성 또는 활성에 영향을 줄 수 있는 다수의 도메인을 포함한다. 예를 들어, 상기에 언급된 바와 같이, 가이드 분자의 제1 및 제2 상보성 도메인에 의해 형성된 이중가닥 구조(또한 반복:반복 방지 이중가닥으로 지칭됨)는 Cas9의 인식(REC) 로브와 상호작용하며, Cas9/가이드 분자 복합체의 형성을 매개할 수 있다. (둘 모두 본 명세서에 참조로 포함되는 문헌[Nishimasu et al., Cell 156, 935-949, February 27, 2014 (Nishimasu 2014) 및 문헌[Nishimasu et al., Cell 162, 1113-1126, August 27, 2015 (Nishimasu 2015)])).

[0225] 제1 및 제2 상보성 도메인과 함께, Cas9 가이드 분자는 통상적으로 생체내에서 뉴클레아제 활성에 관여하는 2개 이상의 추가적 이중가닥 영역을 포함하지만, 시험관내에서는 반드시 그러하지 않은 것은 아니다. (문헌[Nishimasu 2015]). 제2 상보성 도메인의 3' 부분 주위의 제1 줄기 루프는 "근위 도메인"(문헌[Cotta-Ramusino]), "줄기 루프 1"(문헌[Nishimasu 2014 및 2015]) 및 "넥서스(nexus)"(문헌[Briner])로서 다양하게 지칭된다. 하나 이상의 추가적 줄기 루프 구조는 일반적으로 가이드 분자의 3' 말단 주위에 존재하고, 그 수는 종에 따라 다양하고: S. 피오계네스 gRNA는 통상적으로 2개의 3' 줄기 루프(반복:반복 방지 이중가닥을 포함하는 총 4개의 줄기 루프 구조)를 포함하며, S. 아우레우스 및 다른 종은 단지 하나(총 3개의 줄기 루프 구조)를 갖는다. 종에 따라 체계화된 보존된 줄기 루프 구조(및 더욱 일반적으로 가이드 분자 구조)의 기재가 문헌[Briner]에 제공되어 있다.

[0226] 전술한 기재는 Cas9와 함께 사용하기 위한 가이드 분자에 중점을 두었지만, 이 점에 대해 기재된 것과 일부 방식에서 상이한 가이드 분자를 사용하는 다른 RNA-가이드 뉴클레아제가 발견되었거나 발명되었다(또는 추후에 발견되거나 발명될 수 있음). 예를 들어, Cpf1("Prevotella 및 Franciscella 1로부터의 CRISPR")은 tracrRNA가 기능할 필요가 없는 RNA-가이드 뉴클레아제로서 최근 개발되었다. (본 명세서에서 참조로 포함되는 문헌[Zetsche et al., 2015, Cell 163, 759-771 October 22, 2015 (Zetsche I)]). Cpf1 계놈 편집 시스템에서 사용하기 위한 가이드 분자는 일반적으로 표적화 도메인 및 상보성 도메인(대안적으로 "핸들"로 지칭됨)을 포함한다. 또한, Cpf1과 함께 사용하기 위한 가이드 분자에서, 표적화 도메인은 Cas9 가이드 분자와 관련하여 상기 기재된 바와 같이, 5' 말단이 아닌 3' 말단 또는 그 주위에 통상적으로 존재한다는 점에 유의해야 한다(핸들은 Cpf1 가이드 분자의 5' 말단 또는 그 주위에 존재함).

[0227] 그러나 구조적 차이가 상이한 원핵생물 종으로부터의 가이드 분자 사이 또는 Cpf1 및 Cas9 가이드 분자 사이에 존재할 수 있지만, 가이드 분자가 작용하는 원리는 일반적으로 일치한다는 것을 당업자는 인식할 것이다. 이러한 작용의 일관성으로 인해, 광범위한 의미에서, 가이드 분자는 그의 표적화 도메인 서열에 의해 규정될 수 있으며, 일정한 표적화 도메인 서열이 하나 이상의 화학적 변형 및/또는 순차적 변형(치환, 추가적 뉴클레오타이드, 절두 등)을 포함하는 가이드 분자 또는 단분자 또는 키메라 가이드 분자를 포함하는 임의의 적절한 가이드 분자에 혼입될 수 있음을 당업자는 이해할 것이다. 따라서, 본 개시내용에 제시된 조작을 위해, 가이드 분자는 그의 표적화 도메인 서열에 관련하여서만 기재될 수 있다.

[0228] 더욱 일반적으로, 당업자는 본 개시내용의 일부 양태가 다수의 RNA-가이드 뉴클레아제를 사용하여 구현될 수 있는 시스템, 방법 및 조성물에 관한 것임을 이해할 것이다. 이러한 이유로, 달리 명시되지 않는한, 용어 가이드 분자는 Cas9 또는 Cpf1의 특정 종에 적합한 이를 가이드 분자뿐만 아니라, 임의의 RNA-가이드 뉴클레아제와 함

께 사용될 수 있는 임의의 적절한 가이드 문자(예를 들어, gRNA)를 포괄하는 것으로 이해되어야 한다. 설명을 위한 것으로, 용어 가이드 문자는 특정 구현예에서, 클래스 2 CRISPR 시스템, 예컨대 유형 II 또는 유형 V 또는 CRISPR 시스템에서 생성되는 임의의 RNA-가이드 뉴클레아제 또는 그로부터 유래하거나 적용되는 RNA-가이드 뉴클레아제와 함께 사용하기 위한 가이드 문자를 포함할 수 있다.

[0229] 교차 연결된 가이드 문자

본 개시내용의 특정 구현예는, 예를 들어 비 뉴클레오타이드 화학적 연결을 통해 교차 연결된 가이드 문자에 관한 것이다. 상기에 기재된 바와 같이, 연결 위치는 가이드 문자의 줄기 루프 구조 내에 존재할 수 있다. 일부 구현예에서, 가이드 문자는 포함한다.

[0231] 일부 구현예에서, 단분자 가이드 문자는 5'에서 3'으로:

[0232] 표적화 도메인 서열;

[0233] 제1 하부 줄기 서열;

[0234] 제1 벌지 서열;

[0235] 제1 상부 줄기 서열

[0236] 을 포함하는 제1 가이드 문자 단편;

[0237] 비 뉴클레오타이드 화학적 연결; 및

[0238] 제2 상부 줄기 서열;

[0239] 제2 벌지 서열; 및

[0240] 제2 하부 줄기 서열

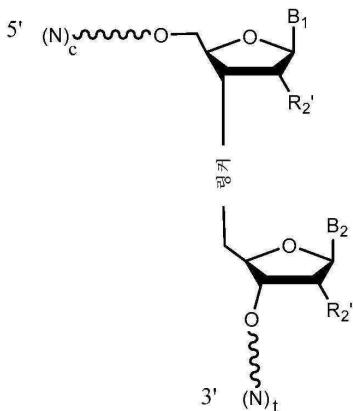
[0241] 을 포함하는 제2 가이드 문자 단편

[0242] 을 포함하며,

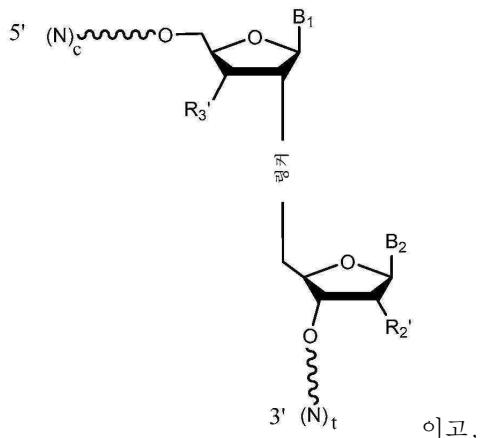
[0243] (a) 제1 하부 줄기 서열의 적어도 하나의 뉴클레오타이드는 제2 하부 줄기 서열의 뉴클레오타이드와 염기쌍을 형성하고, (b) 제1 상부 줄기 서열의 적어도 하나의 뉴클레오타이드는 제2 상부 줄기 서열의 뉴클레오타이드와 염기쌍을 형성한다.

[0244] 일부 구현예에서, 가이드 문자는 제1 및 제2 상부 줄기 서열 사이에 테트라루프 서열을 포함하지 않는다. 일부 구현예에서, 제1 및/또는 제2 상부 줄기 서열은 4개 내지 22개(상하한 포함)의 수의 뉴클레오타이드를 포함한다. 일부 구현예에서, 제1 및/또는 제2 상부 줄기 서열은 1개 내지 22개(상하한 포함)의 수의 뉴클레오타이드를 포함한다. 일부 구현예에서, 제1 및/또는 제2 상부 줄기 서열은 4개 내지 22개(상하한 포함)의 수의 뉴클레오타이드를 포함한다. 일부 구현예에서, 제1 및 제2 상부 줄기 서열은 8개 내지 22개(상하한 포함)의 수의 뉴클레오타이드를 포함한다. 일부 구현예에서, 제1 및 제2 상부 줄기 서열은 12개 내지 22개(상하한 포함)의 수의 뉴클레오타이드를 포함한다.

[0245] 일부 구현예에서, 가이드 문자는 제1 및 제2 가이드 문자 단편 사이의 이중가닥 형성의 깁스 자유 에너지(ΔG)는 2개의 제1 가이드 문자 단편 사이의 이중가닥 형성의 ΔG 미만인 것을 특징으로 한다. 일부 구현예에서, (i) 제1 및 제2 상부 줄기 서열 및 (ii) 제1 및 제2 하부 줄기 서열 각각 사이에 50%, 60%, 70%, 80%, 90% 또는 95% 초과의 염기쌍 형성을 특징으로 하는 제1 및 제2 가이드 문자 단편 사이의 이중가닥 형성의 ΔG 는 (i) 및 (ii) 사이에 50%, 60%, 70%, 80%, 90% 또는 95% 미만의 염기쌍 형성을 특징으로 하는 이중가닥 형성의 ΔG 보다 더 작다.



[0246] 일부 구현예에서, 합성 가이드 분자는 화학식 또는



[0247] (N_c) 및 (N_t)의 각각의 N은 독립적으로, 각각 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미테이트 연결을 통해 그의 인접 뉴클레오타이드(들)에 연결된 뉴클레오타이드 잔기, 선택적으로 변형된 뉴클레오타이드 잔기이고;

[0248] (N_c)는 (N_t)의 5' 영역에 상보성이거나 부분적으로 상보성이고, 그와 이중가닥을 형성하는 3' 영역을 포함하고;

[0249] c는 정수 20 이상이고;

[0250] t는 정수 20 이상이고;

[0251] 링커는 비 뉴클레오타이드 화학적 연결이고;

[0252] B_1 및 B_2 는 각각 독립적으로 해염기이고;

[0253] 각각의 R_2' 및 R_3' 는 독립적으로 H, OH, 플루오로, 클로로, 브로모, NH₂, SH, S-R', 또는 O-R'이고, 각각의 R'는 독립적으로 보호기 또는 알킬기이고, 알킬기는 선택적으로 치환될 수 있고;

[0254] 각각의 wave는 독립적으로 포스포디에스테르 연결, 포스포로티오에이트 연결, 포스포노아세테이트 연결, 티오포스포노아세테이트 연결, 또는 포스포로아미데이트 연결을 나타낸다.

[0255] 일부 구현예에서, (N_t) 및 (N_c)의 이중가닥 영역은 표 4에 열거된 서열을 포함한다.

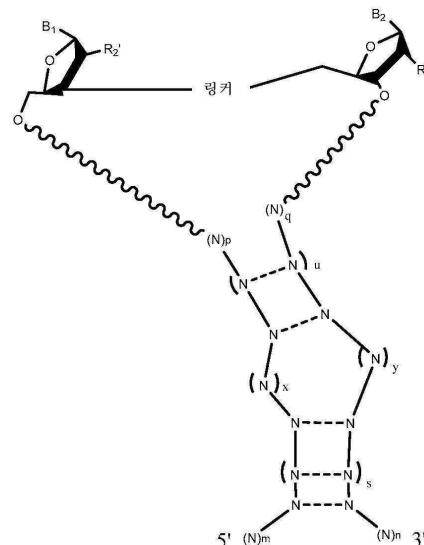
[0256]

[표 4]

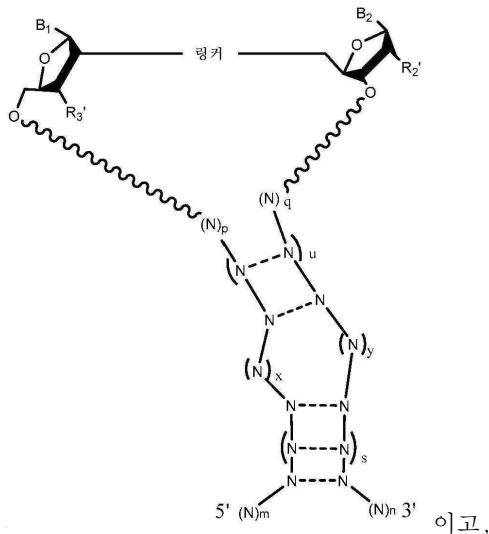
(N)t 및 (N)c의 예시적 서열

SEQ ID NO.	서열
1	GUUUUAGAGCUAG
2	AUAGCAAGUUAAAAAU
3	GUUUUAGAGCU
4	AGCAAGUUAAAAU
5	GUUUUAGAGCUAG
6	CUAGCAAGUUAAAAU
7	GUUUUAGAGCUAUG
8	CAUAGCAAGUUAAAAU
9	GUAUUAGAGCUAUGCUGUUUU
10	AAAACAGCAUAGCAAGUUAAAU
11	GUAUUAGAGCUAUGC
12	AGCAUAGCAAGUUAAAU
13	GUUUUAGAGCUAUGCUGUUU
14	AAAACAGCAUAGCAAGUUAAAAU
15	GUUUUAGAGCUAUGC
16	AGCAUAGCAAGUUAAAAA
17	GUUUUAGAGCUAAAG
18	AUUUAGCAAGUUAAAAU
19	GUUUUAGACCUAA
20	UUAGCAAGUUAAAAU
21	GUUUUAGACCUAAAGGG
22	ACCUUUAGCAAGUUAAAAU
23	GUUUUAGAGCUAG
24	GUUUUAGUACUCU
25	AGAAUCUACUAAAAC
26	GUUUUAGUACUCUGUA
27	UACAGAAUCUACUAAAAC
28	GUUUUAGUACUCUGUAUUUUAGG
29	CCUAAAUUACAGAACUACUAAAAC
30	GUUUUAGUACUCUGUAUUUUAGGUAGA
31	UCAUACCUAAAAUUACAGAACUACUACUAAAAC

[0257]



[0258] 일부 구현예에서, 가이드 분자는 화학식 또는



[0259] N, B₁, B₂, R_{2'}, R_{3'}, 링커, 및 는 상기에 정의된 바와 같고;

[0260] N---N은 독립적으로 2개의 상보성 뉴클레오파이드, 선택적으로, 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오파이드를 나타내고;

[0261] p 및 q는 각각 독립적으로 정수 0 내지 6(상하한 포함)이고, p+q는 정수 0 내지 6(상하한 포함)이고;

[0262] u는 정수 2 내지 22(상하한 포함)이고;

[0263] s는 정수 1 내지 10(상하한 포함)이고;

[0264] x는 정수 1 내지 3(상하한 포함)이고;

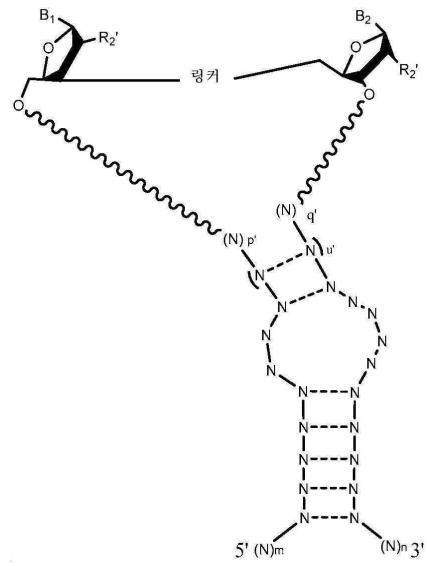
[0265] y는 x를 초과하고, 정수 3 내지 5(상하한 포함)이고;

[0266] m은 정수 15 이상이고;

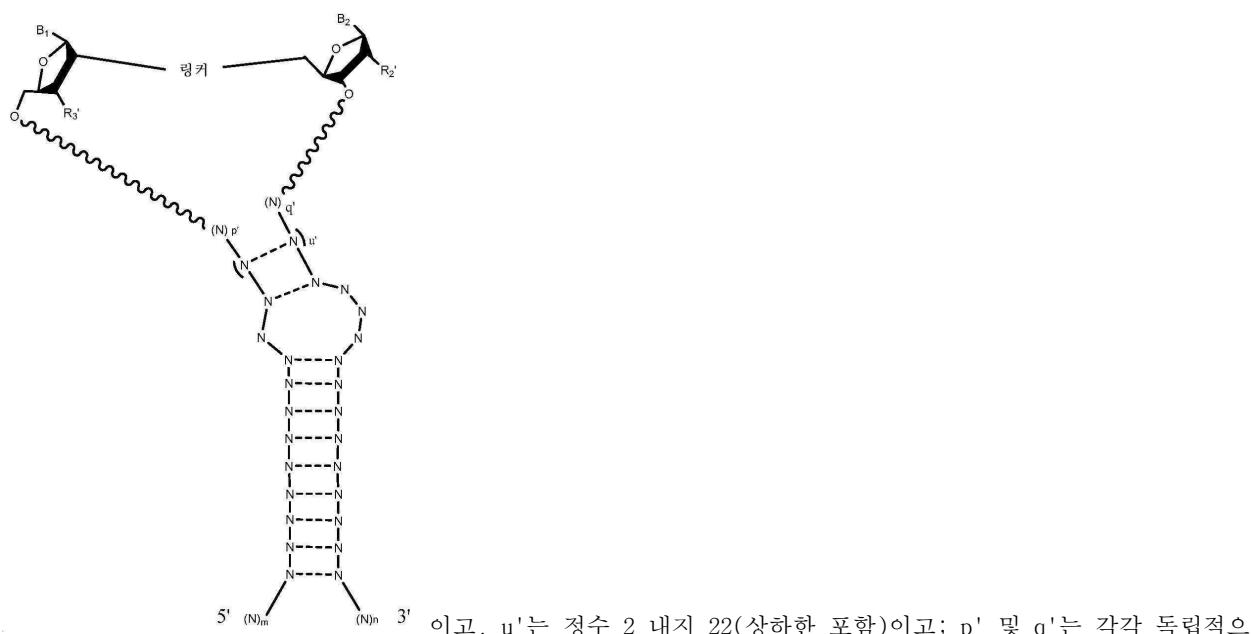
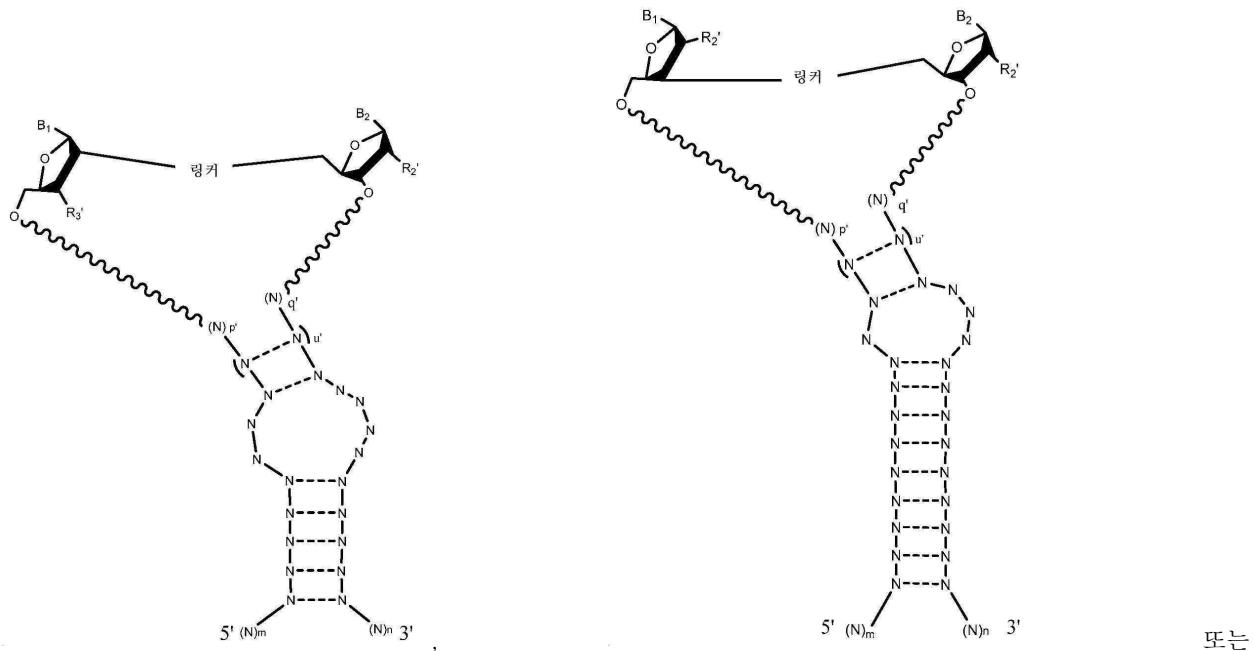
[0267] n은 정수 30 이상이다.

[0268] 일부 구현예에서, (N---N)_u 및 (N---N)_s는 3개 이상의 뉴클레오파이드의 동일한 서열을 포함하지 않는다. 일부 구현예에서, (N---N)_u 및 (N---N)_s는 4개 이상의 뉴클레오파이드의 동일한 서열을 포함하지 않는다. 일부 구현예에서, (N---N)_s는 N'UUU, UN'UU, UUN'U 또는 UUUN' 서열을 포함하고, (N---N)_u는 UUUU 서열을 포함하고, N'는 A, G 또는 C이다. 일부 구현예에서, (N---N)_s는 UUUU 서열을 포함하고, (N---N)_u는 N'UUU, UN'UU, UUN'U 또는 UUUN' 서열을 포함하고, N'는 A, G 또는 C이다. 일부 구현예에서, N'는 A이다. 일부 구현예에서, N'는 G이다. 일부 구현예에서, N'는 C이다.

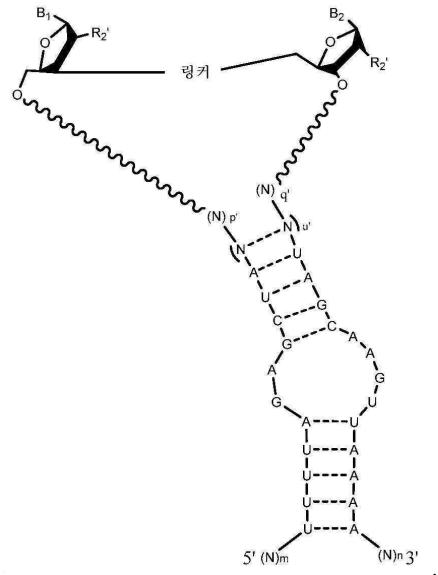
[0269] 일부 구현예에서, 가이드 분자는 S. 피오게네스 또는 S. 아우레우스 Cas9 시스템에 사용되는 gRNA를 기반으로



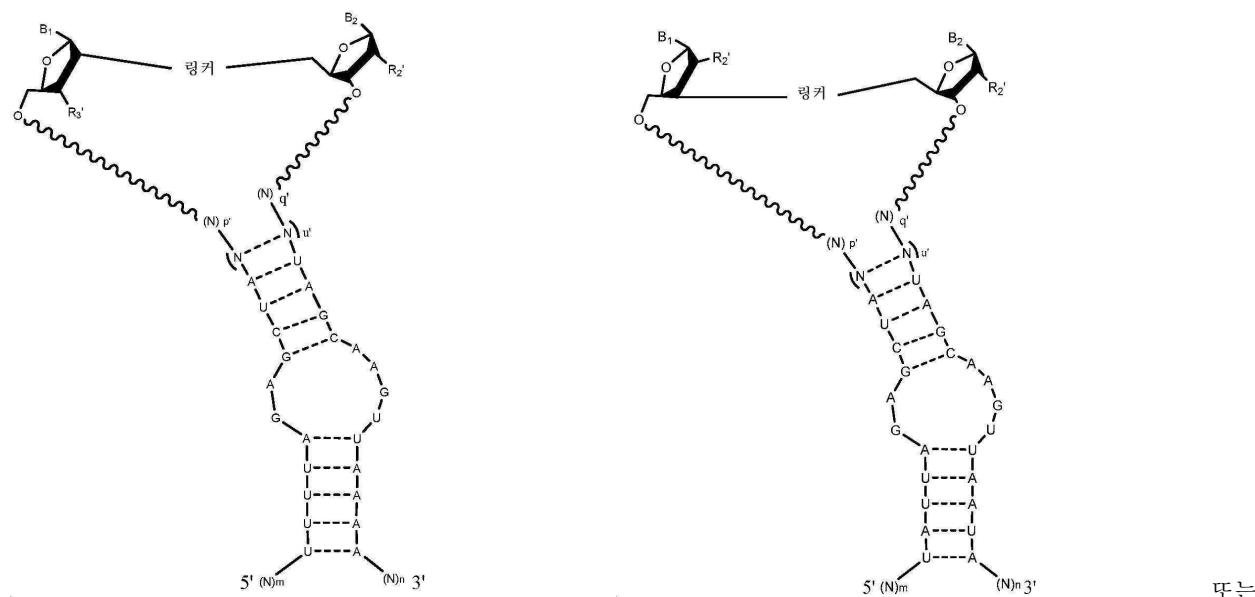
한다. 일부 구현예에서, 가이드 문자는



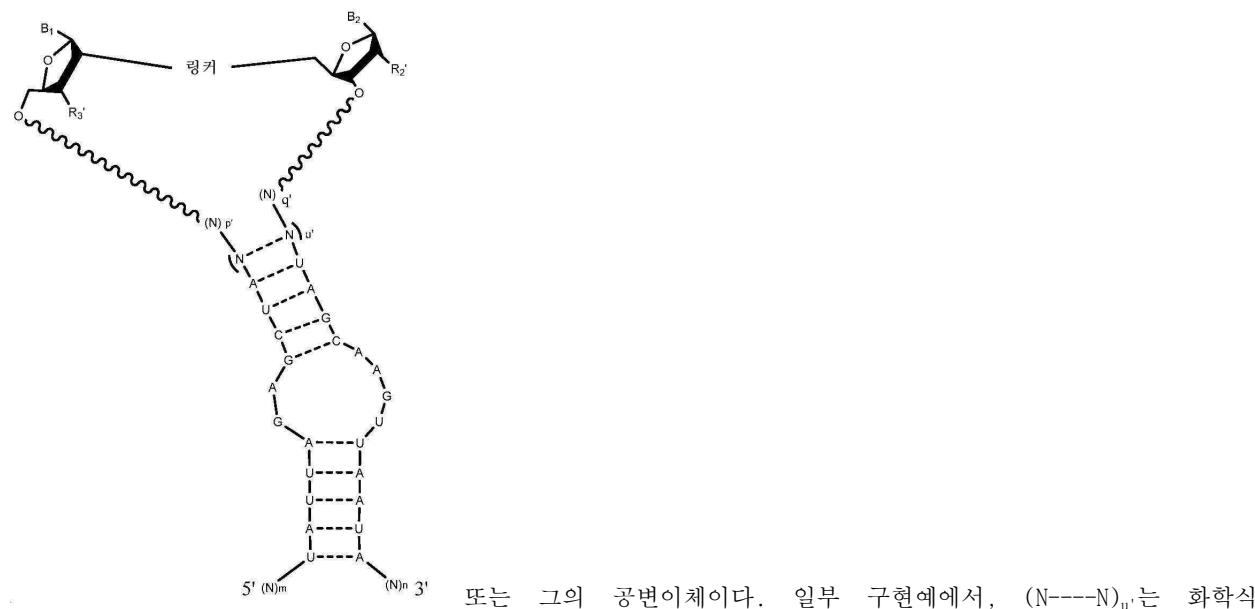
로 정수 0 내지 6(상하한 포함)이고, $p'+q'$ 는 정수 0 내지 6(상하한 포함)이다.



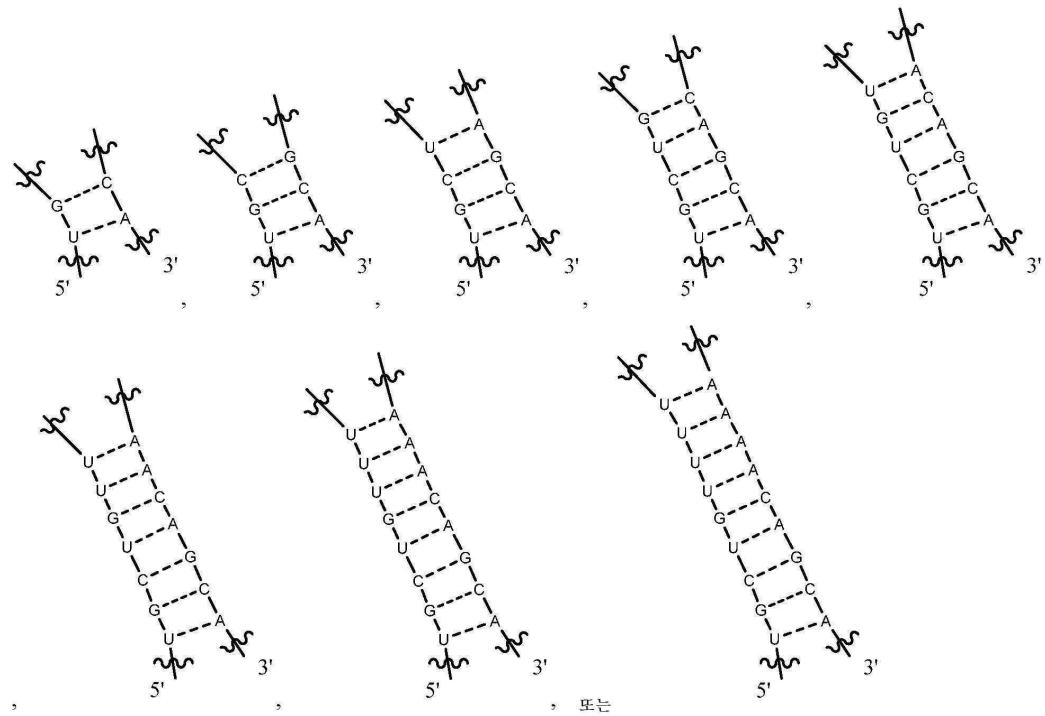
[0270] 일부 구현 예에서, 가이드 분자는 화학식



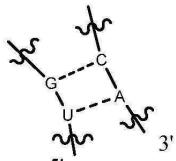
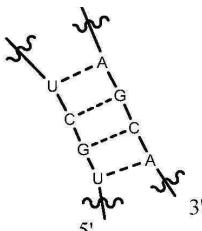
또는



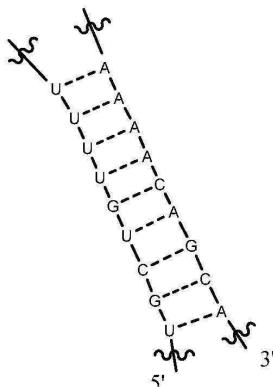
또는 그의 공변이체이다. 일부 구현 예에서, $(N----N)_{u'}$ 는 화학식

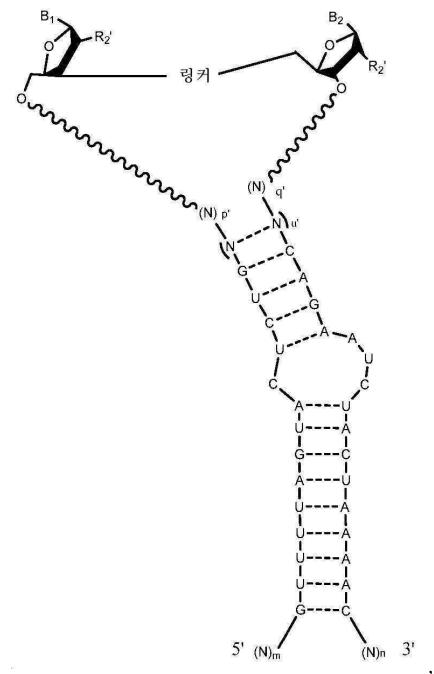


또는 그의 공변

이체이다. 일부 구현예에서, $(N---N)_u'$ 는 화학식또는 그의 공변이체이고, B_1 은 시토신 잔기

또는 그의 공변이체이

이고, B_2 는 구아닌 잔기이다. 일부 구현예에서, $(N---N)_u'$ 는 화학식고, B_1 은 구아닌 잔기이고, B_2 는 시토신 잔기이다. 일부 구현예에서, $(N---N)_u'$ 는 화학식또는 그의 공변이체이고, B_1 은 구아닌 잔기이고, B_2 는 시토신 잔기이다.



[0271]

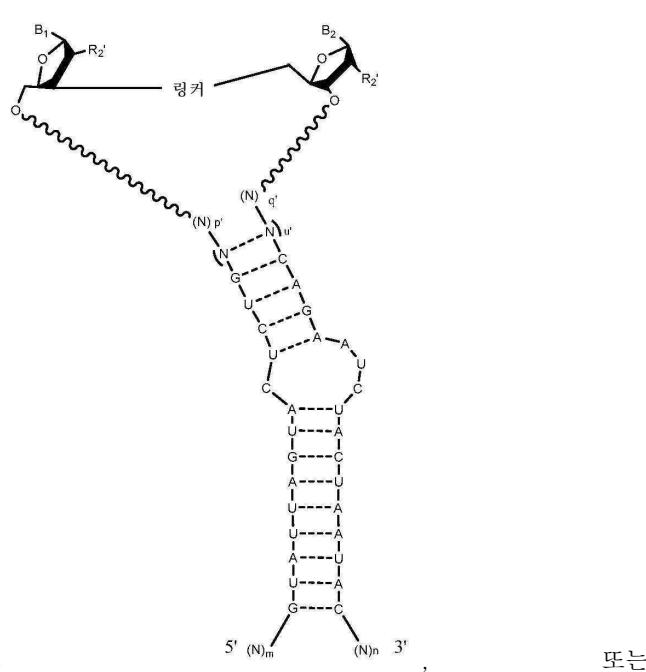
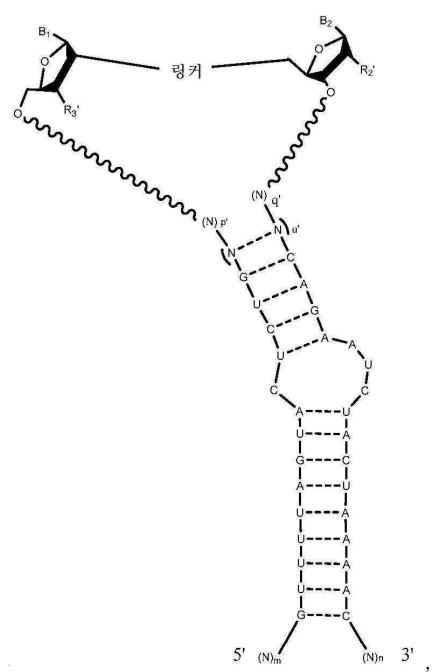
일부

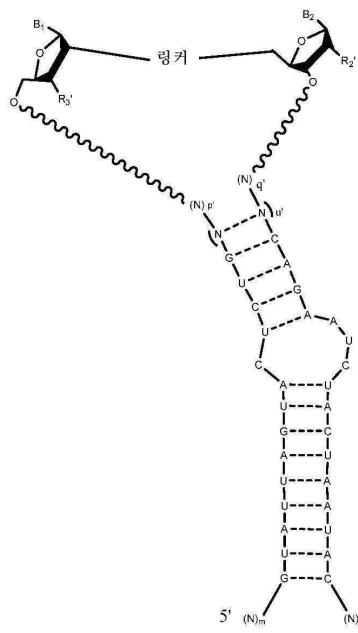
구현예에서,

가이드

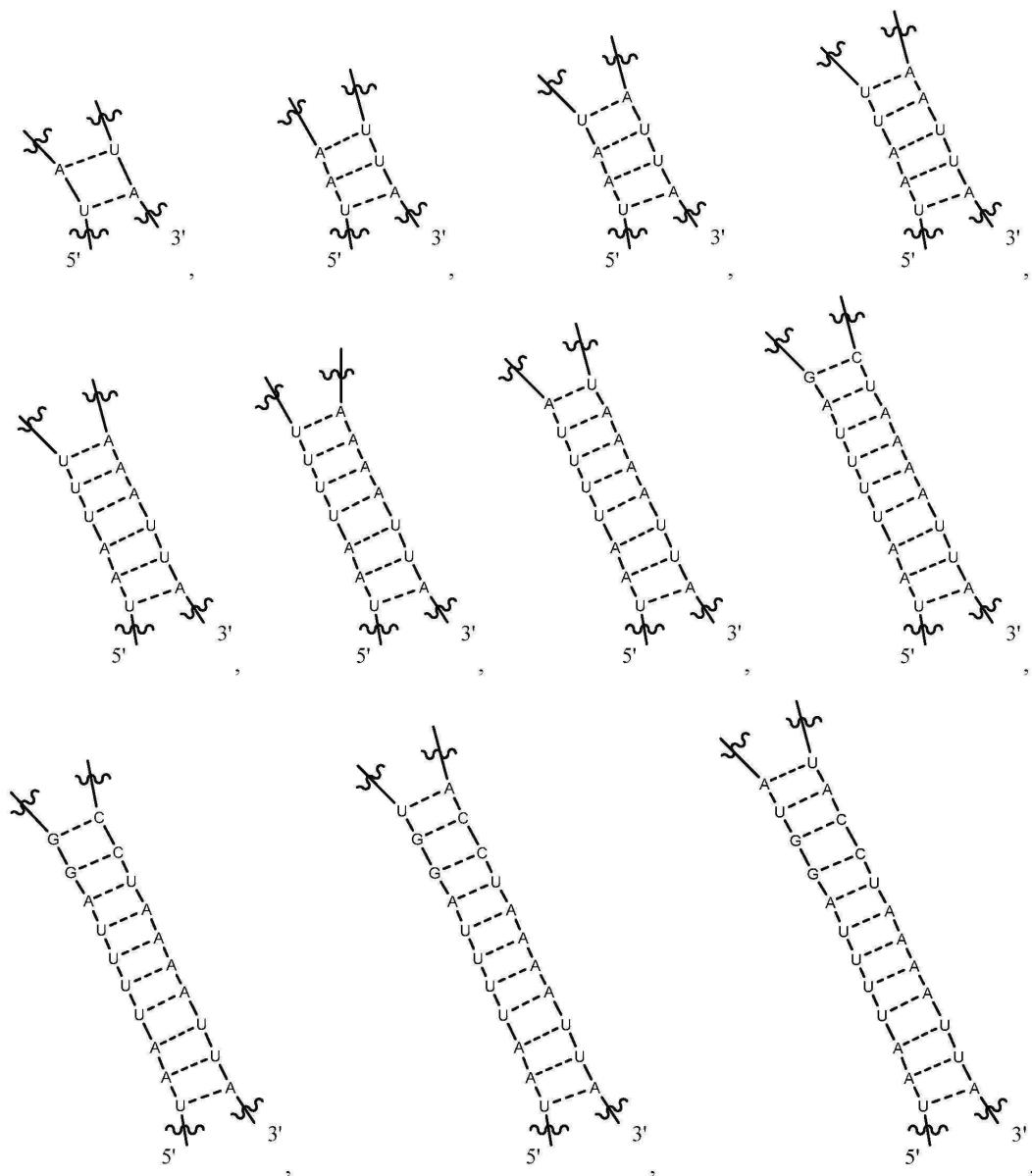
분자는

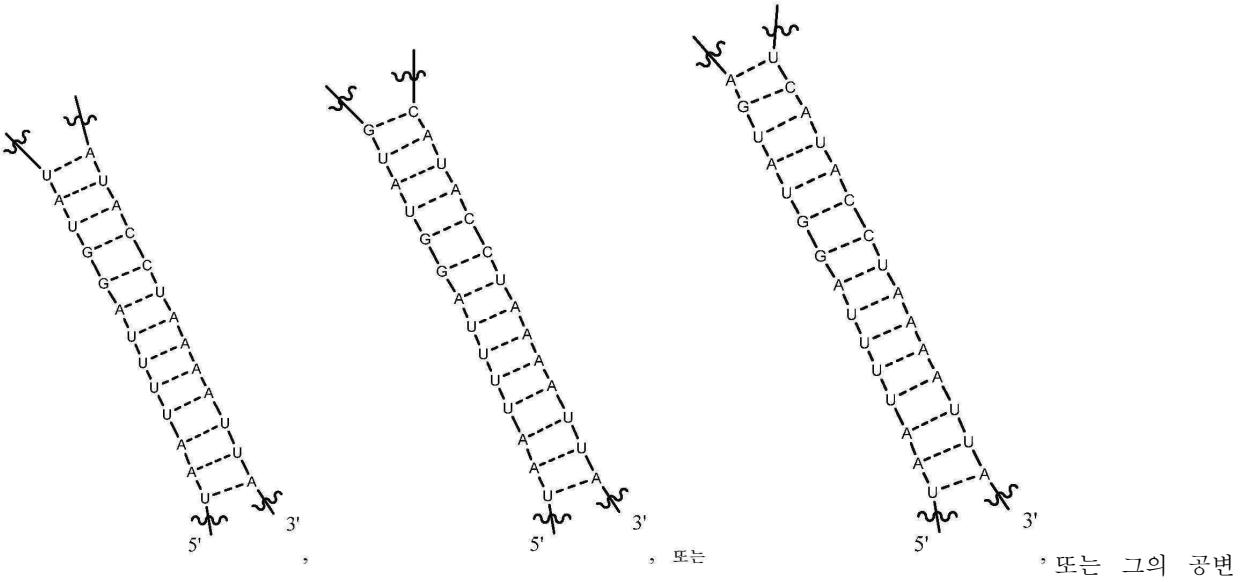
화학식



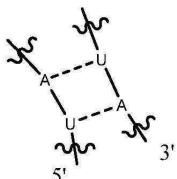


또는 그의 공변이체이다. 일부 구현예에서, $(N---N)_u$ '는 화학식

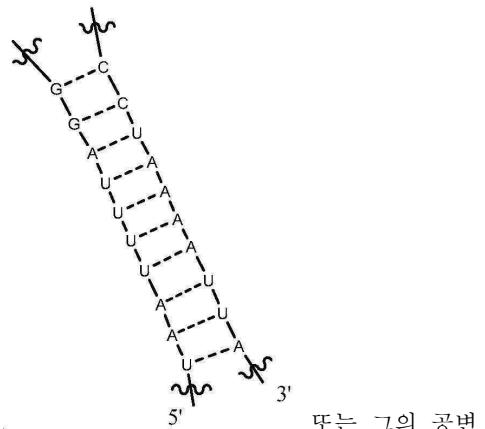




이체이다. 일부 구현예에서, $(N---N)_u'$ 는 화학식

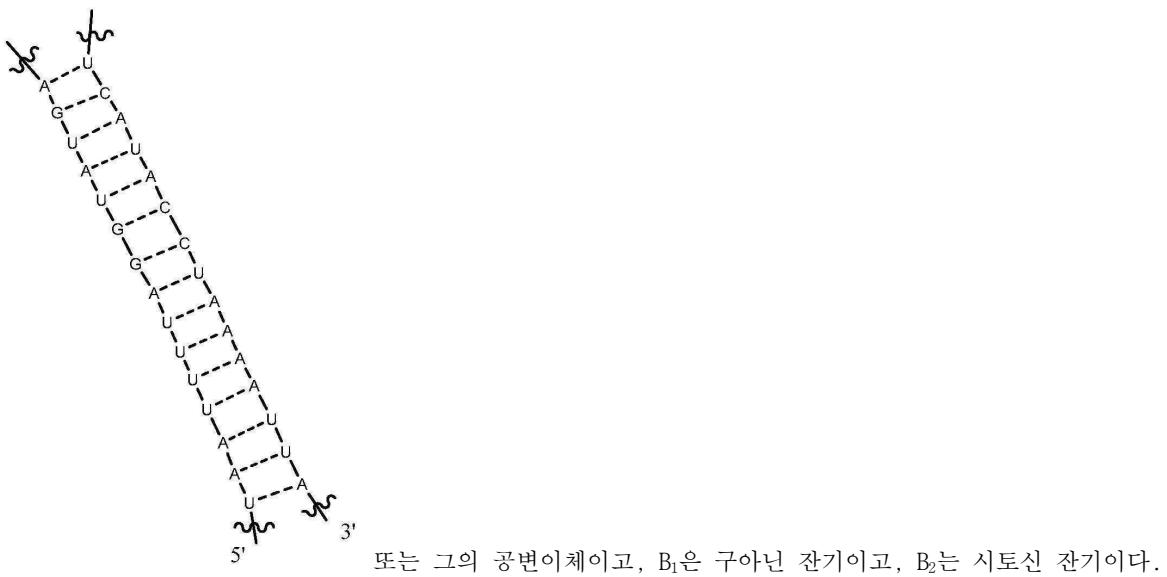


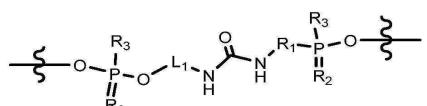
또는 그의 공변이체이고, B_1 은 아데닌 잔기



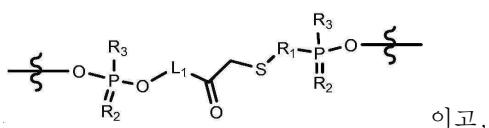
이고, B_2 는 우라실 잔기이다. 일부 구현예에서, $(N---N)_u'$ 는 화학식

이체이고, B_1 은 우라실 잔기이고, B_2 는 아데닌 잔기이다. 일부 구현예에서, $(N---N)_u'$ 는 화학식





[0272] 일부 구현예에서, 링커는 화학식 또는



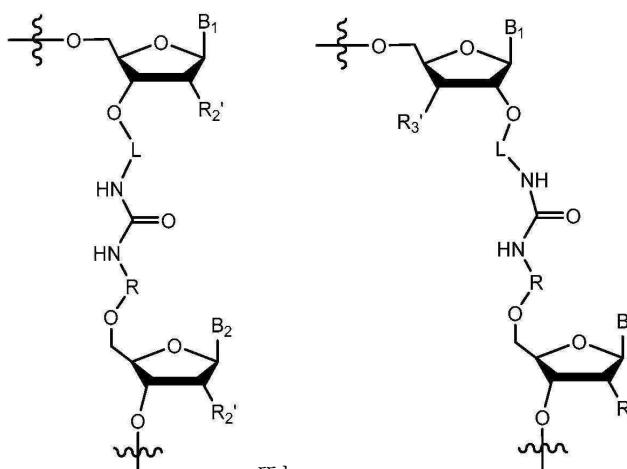
이고,

[0273] 각각의 R2는 독립적으로 O 또는 S이고;

[0274] 각각의 R3은 독립적으로 O⁻ 또는 COO⁻이고;

[0275] L₁ 및 R₁은 각각 비 뉴클레오타이드 화학적 링커이다.

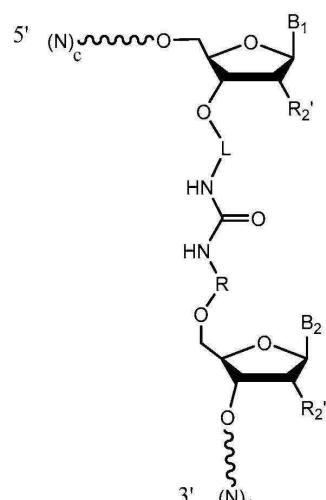
[0276] 일부 구현예에서, 교차 연결된 가이드 분자의 화학적 연결은 우레아를 포함한다. 일부 구현예에서, 우레아를 포함하는 가이드 분자는 화학식



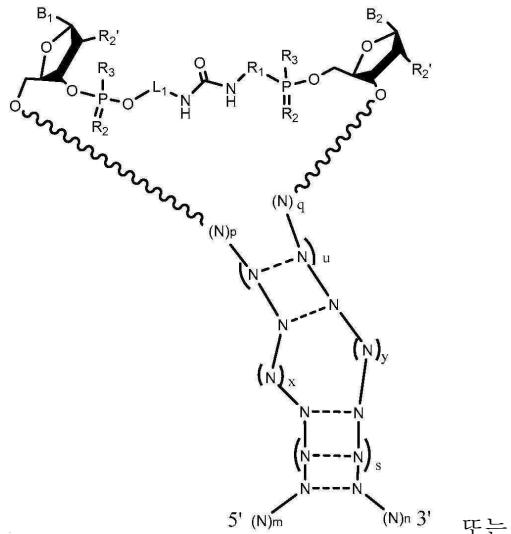
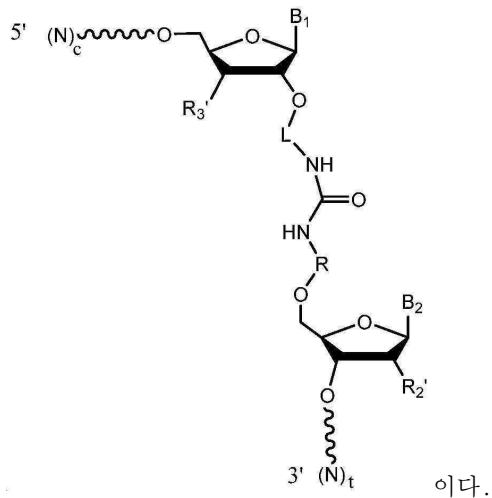
또는

이고,

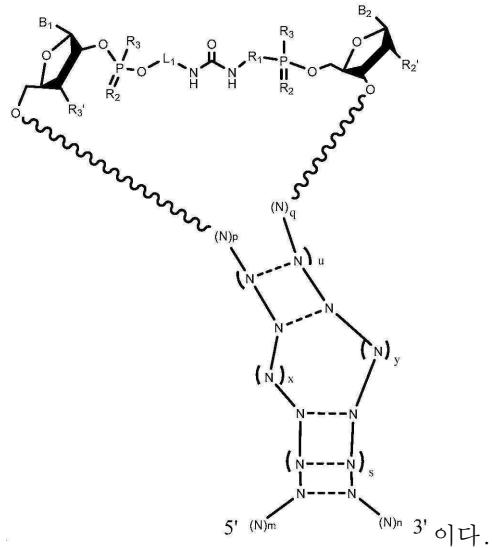
[0277] L 및 R은 각각 독립적으로 비 뉴클레오타이드 링커이다.

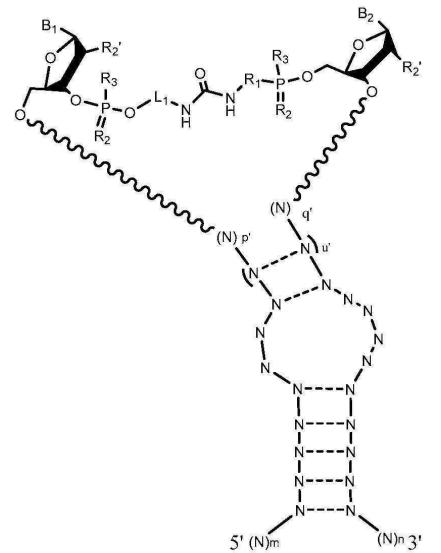


[0278] 일부 구현예에서, 우레아를 포함하는 가이드 분자는 화학식 3' (N)t 또는

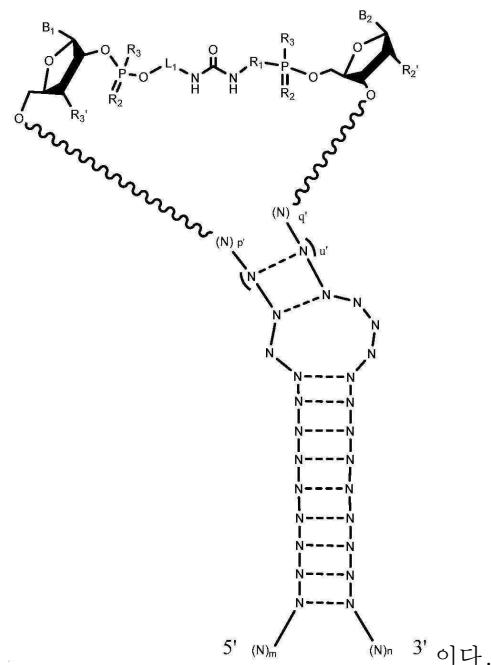
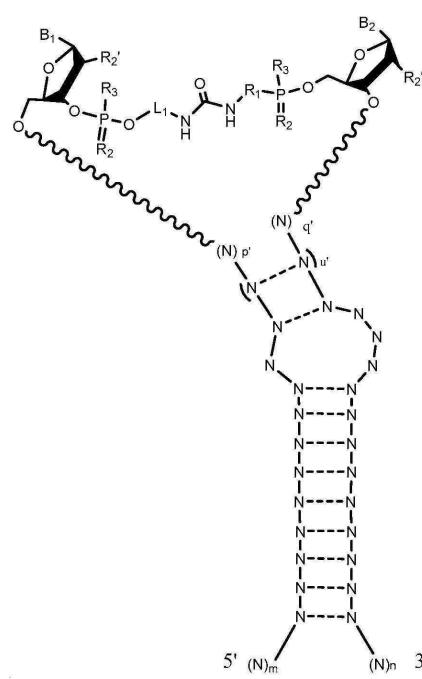
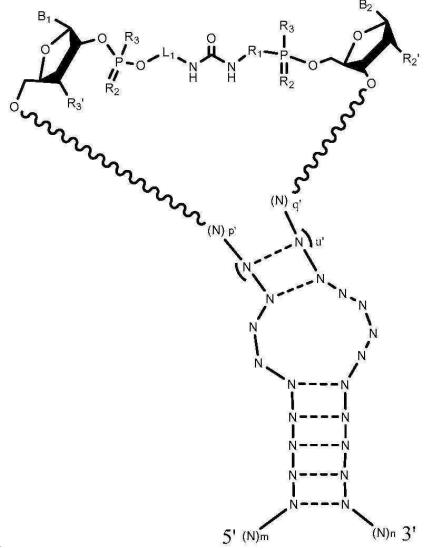


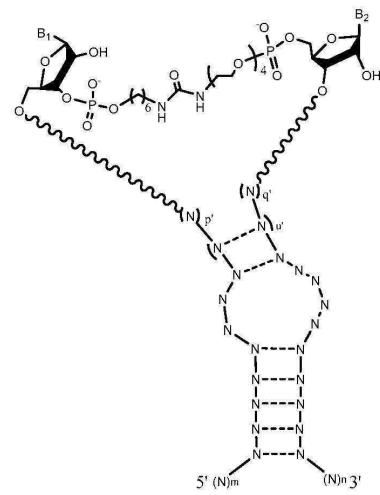
[0279] 일부 구현예에서, 우레아를 포함하는 가이드 분자는 화학식



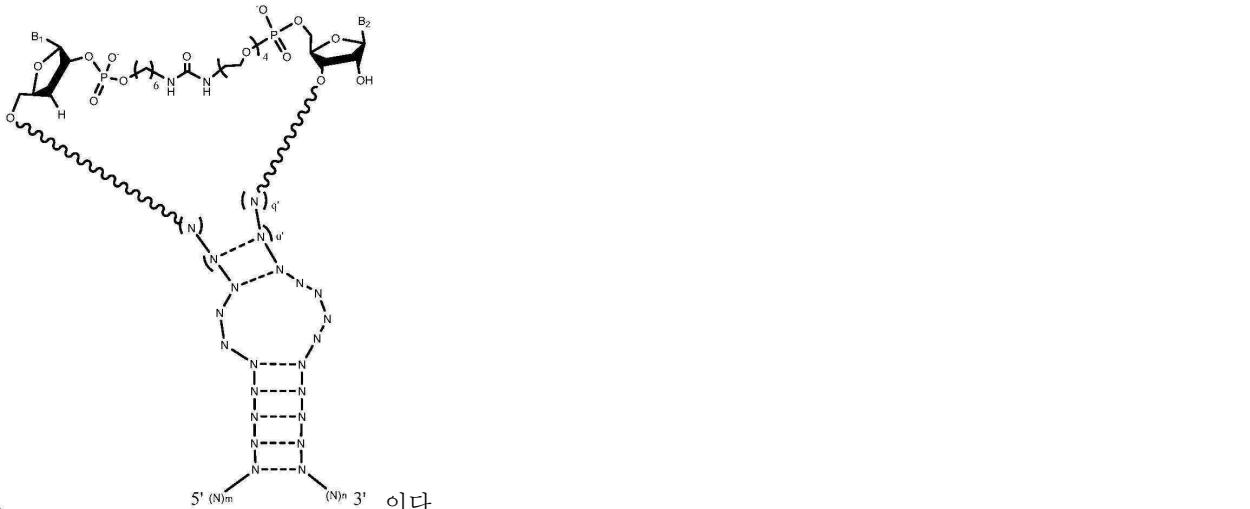
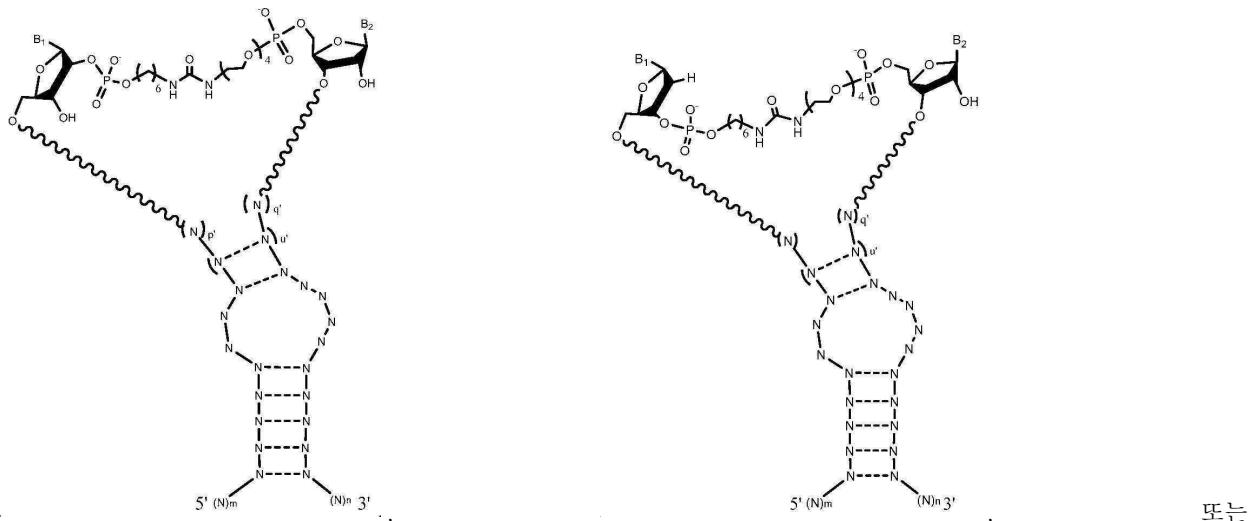


[0280] 일부 구현 예에서, 우레아를 포함하는 가이드 분자는 화학식

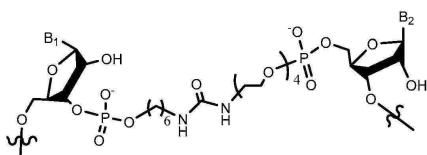




[0281] 일부 구현예에서, 우레아를 포함하는 가이드 분자는 화학식



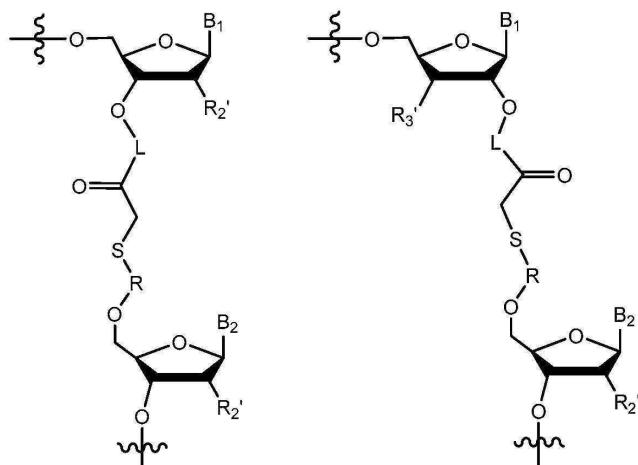
[0282] 일부 구현예에서, 우레아를 포함하는 가이드 분자는 실시예 섹션의 표 10에 열거되어 있는 서열을 가지며, [UR]은 우레아를 포함하는 비 뉴클레오타이드 연결이다. 일부 구현예에서, [UR]은 핵염기 B₁ 및 B₂를 갖는 2개의 뉴클레오타이드 사이의 다음 연결을 나타낸다:



[0283]

[0284] 일부 구현예에서, 교차 연결된 가이드 분자의 화학적 연결은 티오에테르를 포함한다. 일부 구현예에서, 티오에

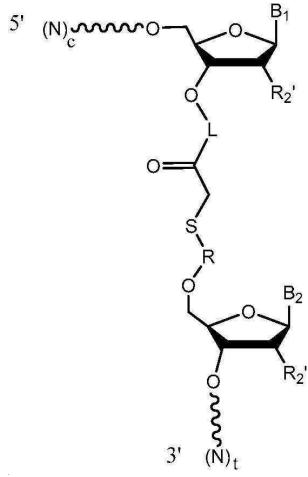
테르를 포함하는 가이드 분자는 화학식

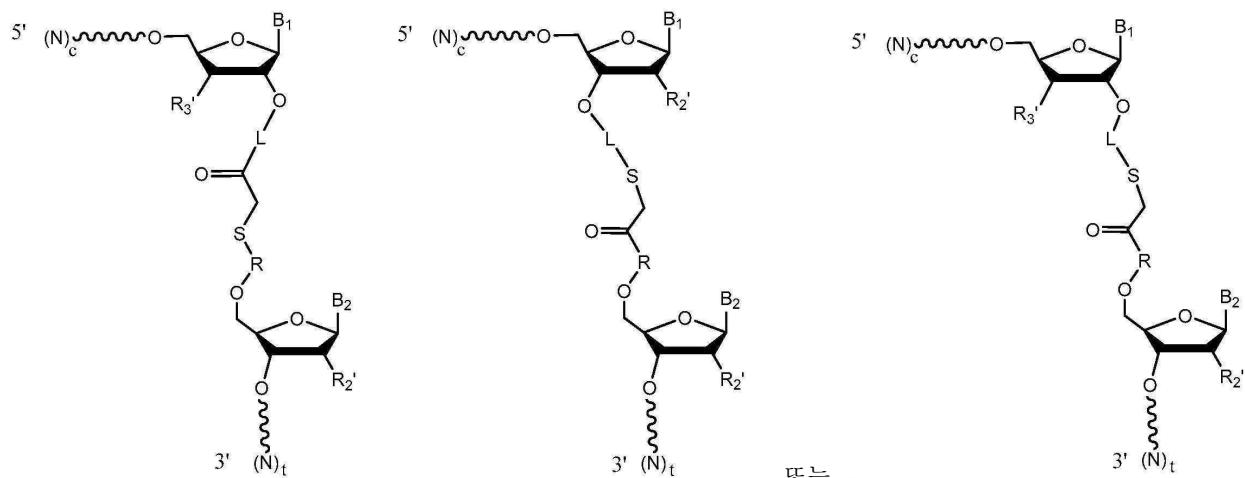


, 또는 :

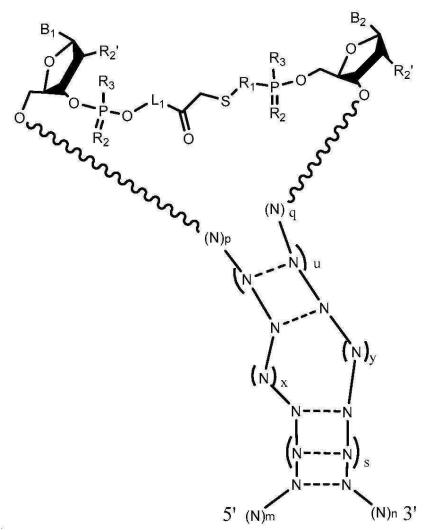
이고, L 및 R은 각각 독립적으로 비 뉴클레오타이드 링커이다.

[0285] 일부 구현예에서, 티오에테르를 포함하는 가이드 분자는 화학식

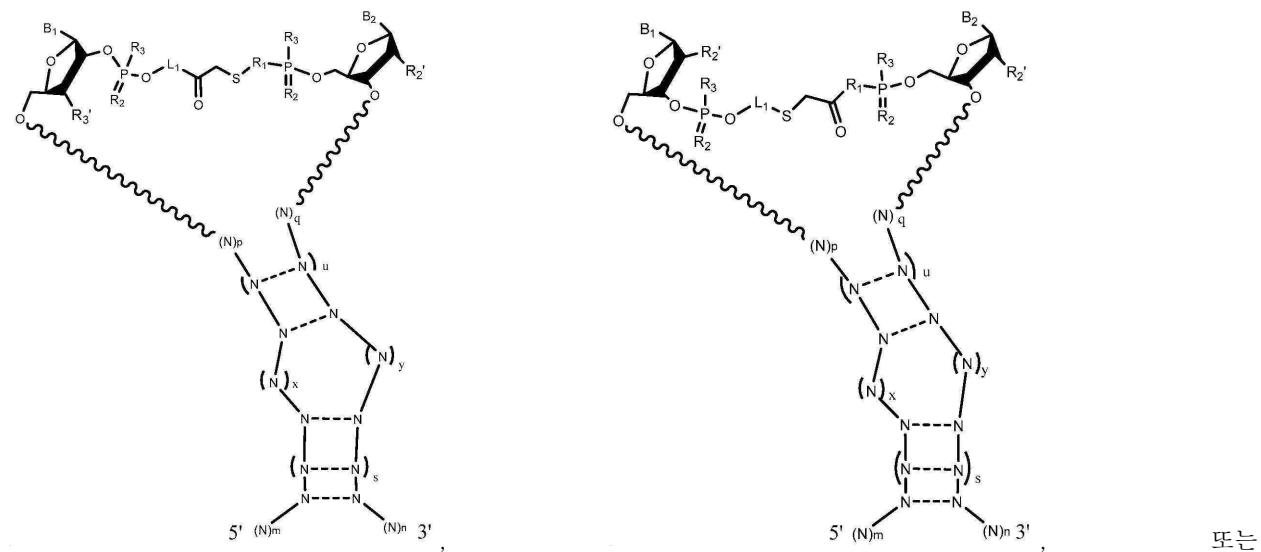


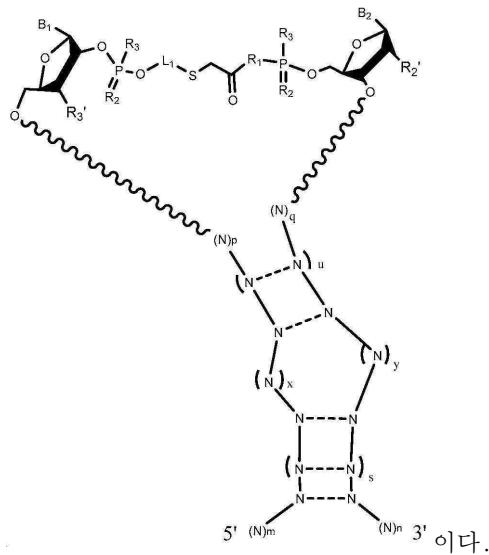


이다.

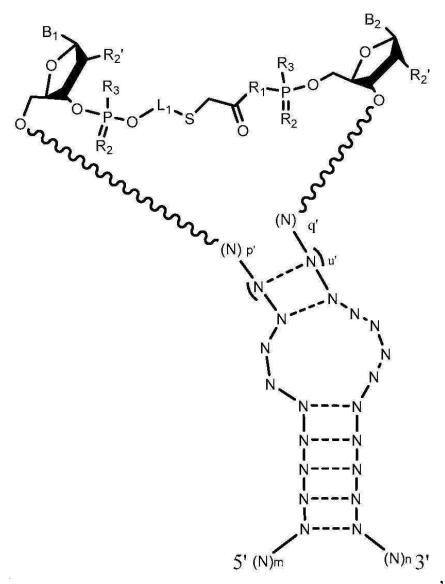
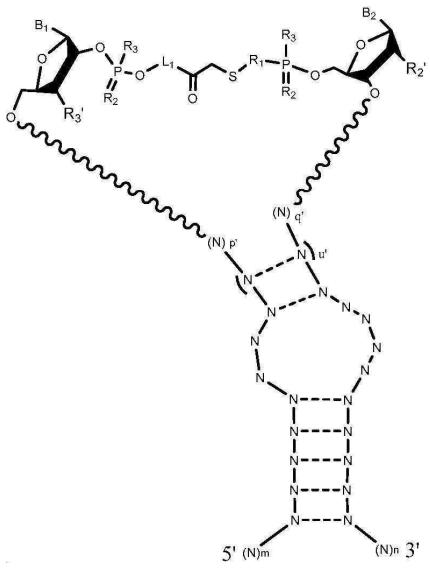


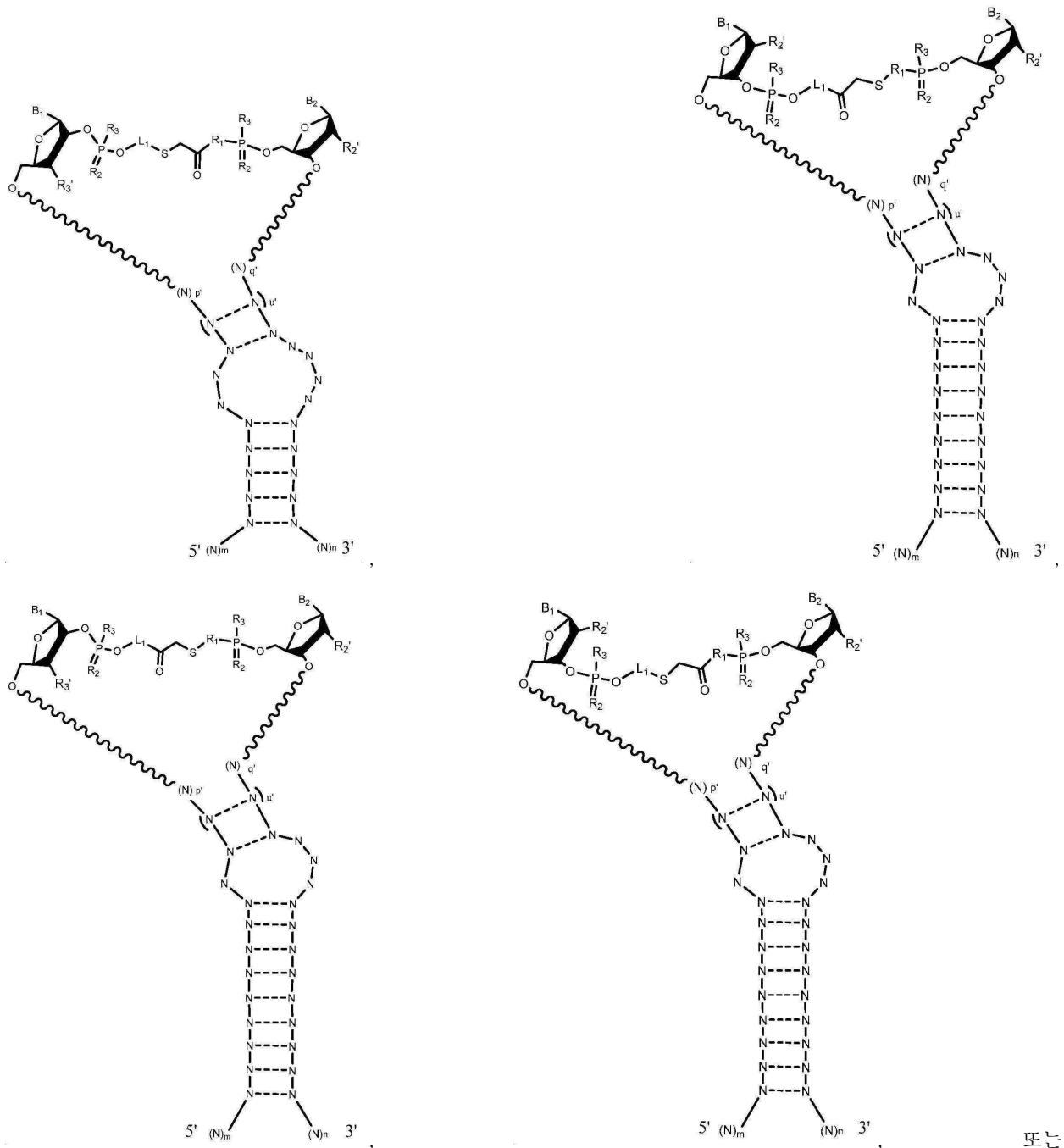
[0286] 일부 구현예에서, 티오에테르를 포함하는 가이드 분자는 화학식

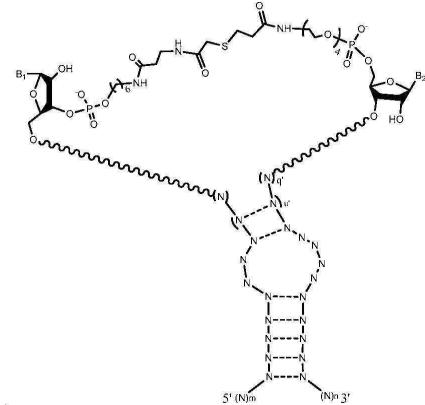
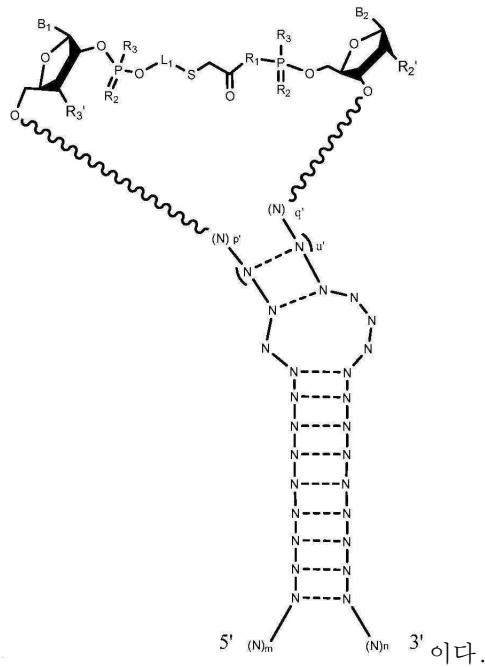




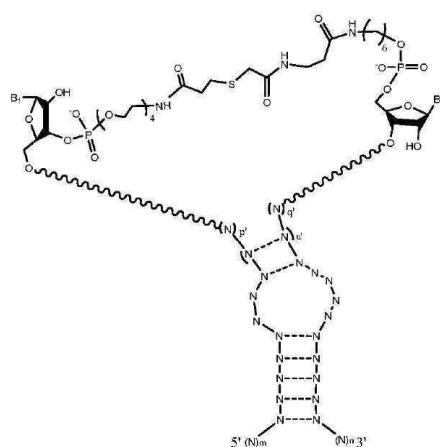
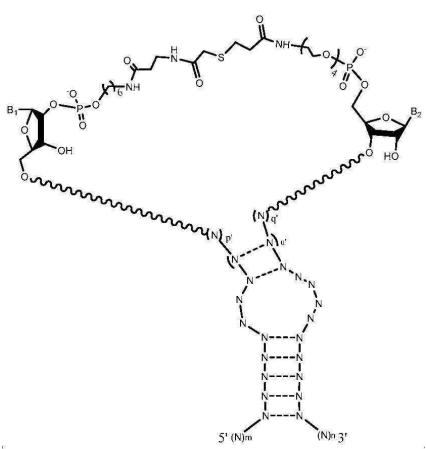
[0287] 일부 구현예에서, 티오에테르를 포함하는 가이드 분자는 화학식



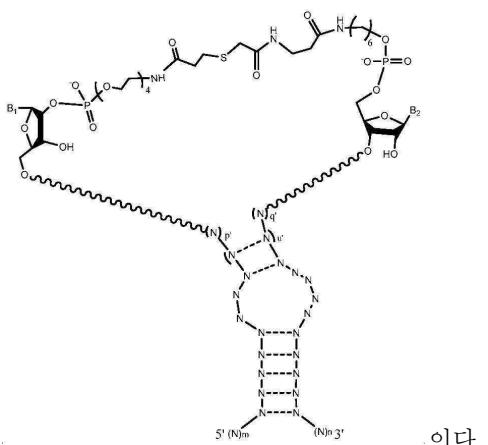




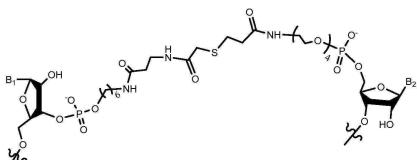
[0288] 일부 구현예에서, 티오에테르를 포함하는 가이드 분자는 화학식



또는



일부 구현 예에서, 우레아를 포함하는 가이드 문자는 실시 예 셙션의 표 9에 열거되어 있는 서열을 가지며, [L]은 티오에테르 연결이다. 일부 구현 예에서, [L]은 핵 염기 B₁ 및 B₂를 갖는 2개의 뉴클레오타이드 사이의 다음 연결을 나타낸다:



일부 구현예에서, 본 출원의 임의의 화학식에서, R_2' 및 R_3' 는 각각 독립적으로 H, OH, 플루오로, 클로로, 브로모, NH_2 , SH, S-R', 또는 O-R'이고, 각각의 R'는 독립적으로 보호기 또는 선택적으로 치환된 알킬기이다. 일부 구현예에서, R_2' 및 R_3' 는 각각 독립적으로 H, OH, 할로겐, NH_2 , 또는 O-R'이고, 각각의 R'는 독립적으로 보호기 또는 선택적으로 치환된 알킬기이다. 일부 구현예에서, R_2' 및 R_3' 는 각각 독립적으로 H, 플루오로, 및 O-R'이고, R'는 보호기 또는 선택적으로 치환된 알킬기이다. 일부 구현예에서, R_2' 는 H이다. 일부 구현예에서, R_3' 는 H이다. 일부 구현예에서, R_2' 는 할로겐이다. 일부 구현예에서, R_3' 는 할로겐이다. 일부 구현예에서, R_2' 는 플루오르이다. 일부 구현예에서, R_3' 는 플루오르이다. 일부 구현예에서, R_2' 는 O-R'이다. 일부 구현예에서, R_3' 는 O-R'이다. 일부 구현예에서, R_2' 는 O-Me이다. 일부 구현예에서, R_3' 는 O-Me이다.

일부 구현예에서, 본 출원의 임의의 화학식에서, p 및 q 는 각각 독립적으로 0, 1, 2, 3, 4, 5, 또는 6이다. 일부 구현예에서, p 및 q 는 각각 독립적으로 2이다. 일부 구현예에서, p 및 q 는 각각 독립적으로 0이다. 일부 구현예에서, p' 및 q' 는 각각 독립적으로 0, 1, 2, 3, 또는 4이다. 일부 구현예에서, p' 및 q' 는 각각 독립적으로 2이다. 일부 구현예에서, p' 및 q' 는 각각 독립적으로 0이다.

일부 구현예에서, 본 출원의 임의의 화학식에서, u 는 정수 2 내지 22(상하한 포함)이다. 일부 구현예에서, u 는 정수 3 내지 22(상하한 포함)이다. 일부 구현예에서, u 는 정수 4 내지 22(상하한 포함)이다. 일부 구현예에서, u 는 정수 8 내지 22(상하한 포함)이다. 일부 구현예에서, u 는 정수 12 내지 22(상하한 포함)이다. 일부 구현예에서, u 는 정수 0 내지 22(상하한 포함)이다. 일부 구현예에서, u 는 정수 2 내지 14(상하한 포함)이다. 일부 구현예에서, u 는 정수 4 내지 14(상하한 포함)이다. 일부 구현예에서, u 는 정수 8 내지 14(상하한 포함)이다. 일부 구현예에서, u 는 정수 0 내지 14(상하한 포함)이다. 일부 구현예에서, u 는 정수 0 내지 4(상하한 포함)이다. 일부 구현예에서, 본 출원의 임의의 화학식에서, u' 는 정수 2 내지 22(상하한 포함)이다. 일부 구현예에서, u' 는 정수 3 내지 22(상하한 포함)이다. 일부 구현예에서, u' 는 정수 4 내지 22(상하한 포함)이다. 일부 구현예에서, u' 는 정수 8 내지 22(상하한 포함)이다. 일부 구현예에서, u' 는 정수 12 내지 22(상하한 포함)이다. 일부 구현예에서, u' 는 정수 0 내지 22(상하한 포함)이다. 일부 구현예에서, u' 는 정수 2 내지 14(상하한 포함)이다. 일부 구현예에서, u' 는 정수 4 내지 14(상하한 포함)이다. 일부 구현예에서, u' 는 정수 8 내지 14(상하한 포함)이다. 일부 구현예에서, u' 는 정수 0 내지 14(상하한 포함)이다. 일부 구현예에서, u' 는 정수 0 내지 4(상하한 포함)이다.

일부 구현 예에서, 본 출원의 임의의 화학식에서, N을 독립적으로 리보뉴를 레오타이드, 데옥시리보뉴를 레오타이

드, 변형된 리보뉴클레오타이드, 또는 변형된 데옥시리보뉴클레오타이드이다. 뉴클레오타이드 변형은 하기에서 논의된다.

- [0295] 일부 구현예에서, 본 출원의 임의의 화학식에서, c는 정수 20 이상이다. 일부 구현예에서, c는 정수 20 내지 60(상하한 포함)이다. 일부 구현예에서, c는 정수 20 내지 40(상하한 포함)이다. 일부 구현예에서, c는 정수 40 내지 60(상하한 포함)이다. 일부 구현예에서, c는 정수 30 내지 60(상하한 포함)이다. 일부 구현예에서, c는 정수 20 내지 50(상하한 포함)이다.

[0296] 일부 구현예에서, 본 출원의 임의의 화학식에서, t는 정수 20 이상이다. 일부 구현예에서, t는 정수 20 내지 80(상하한 포함)이다. 일부 구현예에서, t는 정수 20 내지 50(상하한 포함)이다. 일부 구현예에서, t는 정수 50 내지 80(상하한 포함)이다. 일부 구현예에서, t는 정수 20 내지 70(상하한 포함)이다. 일부 구현예에서, t는 정수 30 내지 80(상하한 포함)이다.

[0297] 일부 구현예에서, 본 출원의 임의의 화학식에서, s는 정수 1 내지 10(상하한 포함)이다. 일부 구현예에서, s는 정수 3 내지 9(상하한 포함)이다. 일부 구현예에서, s는 정수 1 내지 8(상하한 포함)이다. 일부 구현예에서, s는 정수 0 내지 10(상하한 포함)이다. 일부 구현예에서, s는 정수 2 내지 6(상하한 포함)이다.

[0298] 일부 구현예에서, 본 출원의 임의의 화학식에서, x는 정수 1 내지 3(상하한 포함)이다. 일부 구현예에서, x는 1이다. 일부 구현예에서, x는 2이다. 일부 구현예에서, x는 3이다. 일부 구현예에서, 본 출원의 임의의 화학식에서, y는 x를 초과한다. 일부 구현예에서, y는 정수 3 내지 5(상하한 포함)이다. 일부 구현예에서, y는 3이다. 일부 구현예에서, y는 4이다. 일부 구현예에서, y는 5이다. 일부 구현예에서, x는 1이고, y는 3이다. 일부 구현예에서, x는 2이고, y는 4이다.

[0299] 일부 구현예에서, 본 출원의 임의의 화학식에서, m은 정수 15 이상이다. 일부 구현예에서, m은 정수 15 내지 50(상하한 포함)이다. 일부 구현예에서, m은 정수 16 이상이다. 일부 구현예에서, m은 정수 17 이상이다. 일부 구현예에서, m은 정수 18 이상이다. 일부 구현예에서, m은 정수 19 이상이다. 일부 구현예에서, m은 정수 20 이상이다. 일부 구현예에서, m은 정수 20 내지 40(상하한 포함)이다. 일부 구현예에서, m은 정수 30 내지 50(상하한 포함)이다. 일부 구현예에서, m은 정수 15 내지 30(상하한 포함)이다.

[0300] 일부 구현예에서, 본 출원의 임의의 화학식에서, n은 정수 30 이상이다. 일부 구현예에서, n은 정수 30 내지 70(상하한 포함)이다. 일부 구현예에서, n은 정수 30 내지 60(상하한 포함)이다. 일부 구현예에서, n은 정수 40 내지 70(상하한 포함)이다.

[0301] 일부 구현예에서, 본 출원의 임의의 화학식에서, L, R, L₁ 및 R₁은 각각 독립적으로 비 뉴클레오타이드 링커이다. 일부 구현예에서, L, R, L₁ 및 R₁은 각각 독립적으로 폴리에틸렌, 폴리프로필렌, 폴리에틸렌 글리콜, 및 폴리프로필렌 글리콜로 이루어진 군으로부터 선택되는 모이어티를 포함한다. 일부 구현예에서, L₁ 및 R₁은 각각 독립적으로 -(CH₂)_w- , -(CH₂)_w-NH-C(O)-(CH₂)_w-NH- , -(OCH₂CH₂)_v-NH-C(O)-(CH₂)_w- , 또는 -(CH₂CH₂O)_v-이고, 각각의 w는 정수 1 내지 20(상하한 포함)이고, 각각의 v는 정수 1 내지 10(상하한 포함)이다. 일부 구현예에서, L₁은 -(CH₂)_w-이다. 일부 구현예에서, L₁은 -(CH₂)_w-NH-C(O)-(CH₂)_w-NH-이다. 일부 구현예에서, L₁은 -(OCH₂CH₂)_v-NH-C(O)-(CH₂)_w-이다. 일부 구현예에서, L₁은 -(CH₂)₆-NH-C(O)-(CH₂)₁-NH-이다. 일부 구현예에서, L₁은 -(OCH₂CH₂)₄-NH-C(O)-(CH₂)₂-이다. 일부 구현예에서, R₁은 -(CH₂CH₂O)_v-이다. 일부 구현예에서, R₁은 -(CH₂)_w-NH-C(O)-(CH₂)_w-NH-이다. 일부 구현예에서, R₁은 -(OCH₂CH₂)_v-NH-C(O)-(CH₂)_w-이다. 일부 구현예에서, R₁은 -(CH₂CH₂O)₄-이다. 일부 구현예에서, L₁은 -(CH₂)₆-NH-C(O)-(CH₂)₁-NH-이다. 일부 구현예에서, R₁은 -(OCH₂CH₂)₄-NH-C(O)-(CH₂)₂-이다. 일부 구현예에서, L₁은 -(CH₂)₆-이고, R₁은 -(CH₂CH₂O)₄-이다. 일부 구현예에서, L₁은 -(CH₂)₆-NH-C(O)-(CH₂)₁-NH-이고, R₁은 -(OCH₂CH₂)₄-NH-C(O)-(CH₂)₂-이다.

[0302] 일부 구현예에서, 본 출원의 임의의 화학식에서, R₂는 0이고, 일부 구현예에서, R₂는 S이다. 일부 구현예에서, R₃은 0이고, 일부 구현예에서, R₃은 COO⁻이다. 일부 구현예에서, R₂는 0이고, R₃은 0⁻이다. 일부 구현예에서, R₂는 0이고, R₃은 COO⁻이다. 일부 구현예에서, R₂는 S이고, R₃은 0⁻이다. 일부 구현예에서, R₂는 S이고, R₃은 COO⁻이다. 닫힘자는 R₂이 또한 양성자화 혼태(OH 및 COOH)로 존재할 수 있음을 의식할 것이다. 본 출원 전반에 걸쳐, 본

발명자들은 R₃의 탈양성자화 형태 및 양성자화 형태 둘 모두를 포괄하고자 한다.

[0303] 일부 구현예에서, 본 출원의 임의의 화학식에서, 각각의 N----N은 독립적으로 2개의 상보성 뉴클레오타이드, 선택적으로 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타낸다. 일부 구현예에서, 모든 N----N은 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타낸다. 일부 구현예에서, 일부 N----N은 2개의 상보성 뉴클레오타이드를 나타내고, 일부 N----N은 수소결합에 의해 염기쌍을 형성하는 2개의 상보성 뉴클레오타이드를 나타낸다.

[0304] 일부 구현예에서, 본 출원의 임의의 화학식에서, B₁ 및 B₂는 각각 독립적으로 핵염기이다. 일부 구현예에서, B₁은 구아닌이고, B₂는 시토신이다. 일부 구현예에서, B₁은 시토신이고, B₂는 구아닌이다. 일부 구현예에서, B₁은 아데닌이고, B₂는 우라실이다. 일부 구현예에서, B₁은 우라실이고, B₂는 아데닌이다. 일부 구현예에서, B₁ 및 B₂는 상보성이다. 일부 구현예에서, B₁ 및 B₂는 상보성이고, 수소결합을 통해 염기쌍을 형성한다. 일부 구현예에서, B₁ 및 B₂는 상보성이고, 수소결합을 통해 염기쌍을 형성하지 않는다. 일부 구현예에서, B₁ 및 B₂는 상보성이 아니다.

가이드 문자의 합성

[0306] 본 발명의 또 다른 양태는 단문자 가이드 문자를 합성하는 방법이고, 본 방법은:

[0307] 제1 올리고뉴클레오타이드 및 제2 올리고뉴클레오타이드를 어닐링함으로써, 제1 올리고뉴클레오타이드의 3' 영역과 제2 올리고뉴클레오타이드의 5' 영역 사이에 이중가닥을 형성하는 단계로서, 제1 올리고뉴클레오타이드는 2' 반응성 기 및 3' 반응성 기 중 적어도 하나의 제1 반응성 기를 포함하고, 제2 올리고뉴클레오타이드는 5' 반응성 기인 제2 반응성 기를 포함하는 단계; 및

[0308] 어닐링된 제1 및 제2 올리고뉴클레오타이드를 제1 및 제2 반응성 기를 통해 접합시킴으로써, 제1 및 제2 올리고뉴클레오타이드를 연결하는 공유결합을 포함하는 단문자 가이드 문자를 형성하는 단계를 포함한다.

[0309] 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 "개요"하에 상기에 열거된 작용기로부터 선택된다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 각각 독립적으로 아민 모이어티, 술피드릴 모이어티, 브로모아세틸 모이어티, 하이드록실 모이어티, 또는 포스페이트 모이어티이다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 둘 모두 아민 모이어티이다. 일부 구현예에서, 제1 반응성 기는 술피드릴 모이어티이고, 제2 반응성 기는 브로모아세틸 모이어티이다. 일부 구현예에서, 제1 반응성 기는 브로모아세틸 모이어티이고, 제2 반응성 기는 술피드릴 모이어티이다. 일부 구현예에서, 제1 반응성 기는 하이드록실 모이어티이고, 제2 반응성 기는 포스페이트 모이어티이다. 일부 구현예에서, 제1 반응성 기는 포스페이트 모이어티이고, 제2 반응성 기는 하이드록실 모이어티이다.

[0310] 일부 구현예에서, 접합 단계는 10 μM 내지 1 mM의 범위의 제1 뉴클레오타이드의 농도를 포함한다. 일부 구현예에서, 접합 단계는 10 μM 내지 1 mM의 범위의 제2 뉴클레오타이드의 농도를 포함한다. 일부 구현예에서, 제1 또는 제2 뉴클레오타이드의 농도는 10 μM, 50 μM, 100 μM, 200 μM, 400 μM, 600 μM, 800 μM, 또는 1 mM이다.

[0311] 일부 구현예에서, 접합 단계는 5.0 내지 9.0의 범위의 pH를 포함한다. 일부 구현예에서, pH는 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 또는 9.0이다. 일부 구현예에서, pH는 6.0이다. 일부 구현예에서, pH는 8.0이다. 일부 구현예에서, pH는 8.5이다.

[0312] 일부 구현예에서, 접합 단계는 아르곤하에서 수행된다. 일부 구현예에서, 접합 단계는 주위 대기하에서 수행된다.

[0313] 일부 구현예에서, 접합 단계는 물 중에서 수행된다. 일부 구현예에서, 접합 단계는 조용매를 포함하는 물 중에서 수행된다. 일부 구현예에서, 조용매는 DMSO, DMF, NMP, DMA, 모르폴린, 피리딘, 또는 MeCN이다. 일부 구현예에서, 조용매는 DMSO이다. 일부 구현예에서, 조용매는 DMF이다.

[0314] 일부 구현예에서, 접합 단계는 0°C 내지 40°C의 범위의 온도에서 수행된다. 일부 구현예에서, 온도는 0°C, 4°C, 10°C, 20°C, 25°C, 30°C, 37°C, 또는 40°C이다. 일부 구현예에서, 온도는 25°C이다. 일부 구현예에서, 온도는 4°C이다.

[0315] 일부 구현예에서, 접합 단계는 2가 금속 양이온의 존재하에 수행된다. 일부 구현예에서, 2가 금속 양이온은

Mg^{2+} , Ca^{2+} , Sr^{2+} , Ba^{2+} , Cr^{2+} , Mn^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} , 또는 Zn^{2+} 이다. 일부 구현예에서, 2가 금속 양이온은 Mg^{2+} 이다.

[0316] 일부 구현예에서, 접합 단계는 교차 연결 시약 또는 교차 링커를 포함한다(상기 "개요" 참조). 일부 구현예에서, 교차 링커는 다중기능성 및 일부 구현예에서, 교차 링커는 이중기능성이다. 일부 구현예에서, 다중 기능성 교차 링커는 이중기능성 또는 동종기능성이다. 일부 구현예에서, 교차 링커는 카보네이트를 포함한다. 일부 구현예에서, 교차 링커를 포함하는 카보네이트는 디숙신이미딜 카보네이트, 디이미다졸 카보네이트 또는 비스-(p-니트로페닐) 카보네이트이다. 일부 구현예에서, 교차 링커를 포함하는 카보네이트는 디숙신이미딜 카보네이트이다.

[0317] 일부 구현예에서, 접합 단계는 1 mM 내지 100 mM의 범위의 이중기능성 교차 연결 시약의 농도를 포함한다. 일부 구현예에서, 이중기능성 교차 연결 시약의 농도는 1 mM, 10 mM, 20 mM, 40 mM, 60 mM, 80 mM, 또는 100 mM이다. 일부 구현예에서, 이중기능성 교차 연결 시약의 농도는 제1 및 제2 올리고뉴클레오타이드의 각각의 농도를 100배 내지 1000배 초과한다. 일부 구현예에서, 이중기능성 교차 연결 시약의 농도는 제1 올리고뉴클레오타이드의 농도를 100, 200, 400, 600, 800, 또는 1000배 초과한다. 일부 구현예에서, 이중기능성 교차 연결 시약의 농도는 제2 올리고뉴클레오타이드의 농도를 100, 200, 400, 600, 800, 또는 1000배 초과한다.

[0318] 일부 구현예에서, 접합 단계는 킬레이트 시약의 존재하에 수행된다. 일부 구현예에서, 킬레이트 시약은 에틸렌 디아민테트라아세트산(EDTA), 또는 그의 염이다.

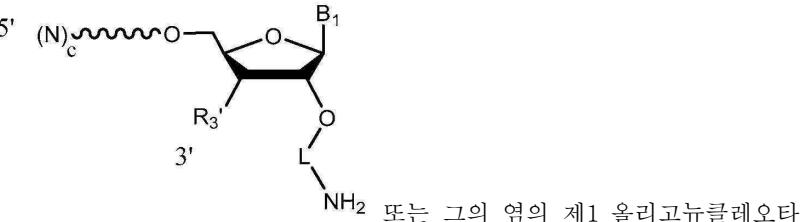
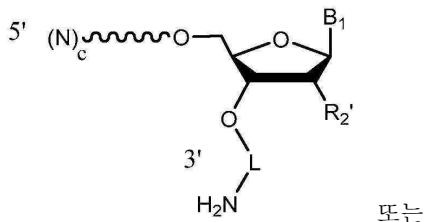
[0319] 일부 구현예에서, 접합 단계는 활성화제의 존재하에 수행된다. 일부 구현예에서, 활성화제는 카보디이미드, 또는 그의 염이다. 일부 구현예에서, 카보디이미드는 1-에틸-3-(3-디메틸아미노프로필)카보디이미드(EDC), N,N'-디사이클로헥실카보디이미드(DCC) 또는 N,N'-디이소프로필카보디이미드(DIC), 또는 그의 염이다. 일부 구현예에서, 카보디이미드는 1-에틸-3-(3-디메틸아미노프로필)카보디이미드(EDC), 또는 그의 염이다.

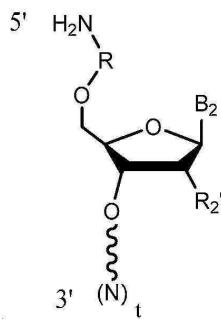
[0320] 일부 구현예에서, 접합 단계는 1 mM 내지 100 mM의 범위의 활성화제의 농도를 포함한다. 일부 구현예에서, 활성화제의 농도는 1 mM, 10 mM, 20 mM, 40 mM, 60 mM, 80 mM, 또는 100 mM이다. 일부 구현예에서, 활성화제의 농도는 제1 및 제2 올리고뉴클레오타이드의 각각의 농도를 100배 내지 1000배 초과한다. 일부 구현예에서, 활성화제의 농도는 제1 올리고뉴클레오타이드의 농도를 100, 200, 400, 600, 800, 또는 1000배 초과한다. 일부 구현예에서, 활성화제의 농도는 제2 올리고뉴클레오타이드의 농도를 100, 200, 400, 600, 800, 또는 1000배 초과한다.

[0321] 일부 구현예에서, 접합 단계는 안정화제의 존재하에 수행된다. 일부 구현예에서, 안정화제는 이미다졸, 시아노 이미다졸, 피리딘, 또는 디메틸아미노피리딘, 또는 그의 염이다. 일부 구현예에서, 안정화제는 이미다졸이다. 일부 구현예에서, 접합 단계는 활성화제 및 안정화제 둘 모두의 존재하에 수행된다. 일부 구현예에서, 접합 단계는 1-에틸-3-(3-디메틸아미노프로필)카보디이미드(EDC) 및 이미다졸, 또는 그의 염의 존재하에 수행된다.

[0322] 일부 구현예에서, 단분자 가이드 분자를 합성하는 방법은 상기에 개시된 임의의 화학식의 가이드 분자를 생성한다.

[0323] 일부 구현예에서, 단분자 가이드 분자를 합성하는 방법은 우레아 링커를 갖는 가이드 분자를 생성한다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 둘 모두 아민이고, 제1 및 제2 반응성 기는 이중기능성 교차연결 시약을 포함하는 카보네이트에 의해 교차 연결됨으로써 우레아 링커를 형성한다. 일부 구현예에서, 이중기능성 교차연결 시약을 포함하는 카보네이트는 디숙신이미딜 카보네이트이다. 일부 구현예에서, 본 방법은 화학식

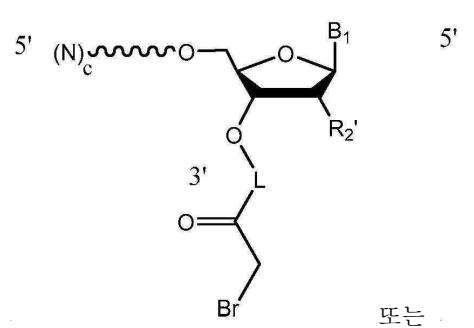




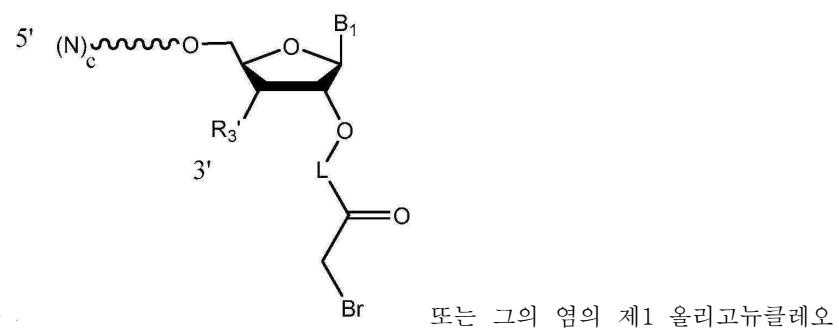
이드를 포함한다. 일부 구현예에서, 본 방법은 화학식 $5'-(N)_t-B_2-R'$, 또는 그의 염의 제2 올리고뉴클레오타이드를 포함한다.

[0324]

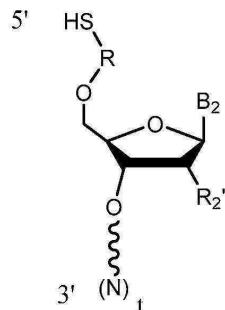
일부 구현예에서, 단분자 가이드 분자를 합성하는 방법은 티오에테르 링커를 갖는 가이드 분자를 생성한다. 일부 구현예에서, 제1 반응성 기는 술피드릴기이고, 제2 반응성 기는 브로모아세틸기이나, 제1 반응성 기는 브로모아세틸기이고, 제2 반응성 기는 술피드릴기이다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 킬레이트제의 존재하에 반응하여, 티오에테르 연결을 형성한다. 일부 구현예에서, 제1 반응성 기 및 제2 반응성 기는 치환 반응을 거쳐, 티오에테르 연결을 형성한다. 일부 구현예에서, 본 방법은 화학식



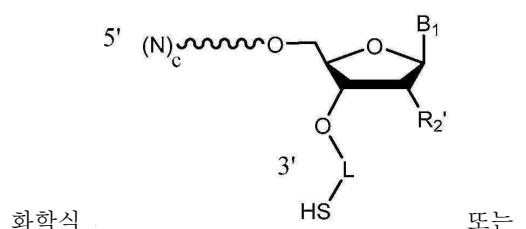
또는



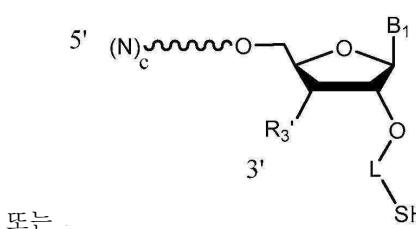
또는 그의 염의 제1 올리고뉴클레오타이드를 포함하고, 제2 올리고뉴클레오타이드는 화학식



타이드를 포함하고, 제2 올리고뉴클레오타이드는 화학식 $5'-(N)_t-B_2-R'$, 또는 그의 염이거나; 본 방법은

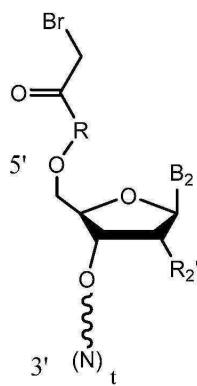


화학식



또는

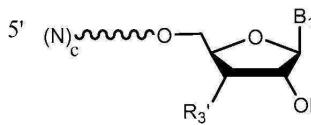
또는 그의 염의 제1 올리고뉴클레오타이드를 포함하고, 제2 올리고뉴클레오타이드는 화학식



레오타이드를 포함하고, 제2 올리고뉴클레오타이드는 화학식 $5'-(N)_t-B_2-R'$, 또는 그의 염이다.

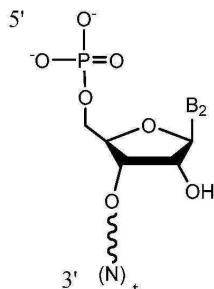
[0325]

일부 구현예에서, 단분자 가이드 분자를 합성하는 방법은 포스포디에스테르 링커를 갖는 가이드 분자를 생성한다. 일부 구현예에서, 제1 반응성 기는 2' 또는 3' 하이드록실기를 포함하고, 제2 반응성 기는 5' 포스페이트 모이어티를 포함한다. 일부 구현예에서, 제1 및 제2 반응성 기는 활성화제의 존재하에 접합되어, 포스포디에스테르 링커를 형성한다. 일부 구현예에서, 활성화제는 EDC이다. 일부 구현예에서, 본 방법은 화학식



3'

또는 그의 염의 제1 올리고뉴클레오타이드를 포함하고; 제2 올리고뉴클레오타이



3' (N)t

또는 그의 염이다.

드는 화학식

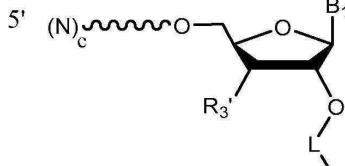
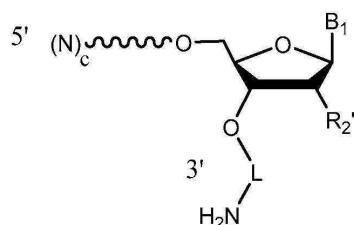
[0326] 일부 구현예에서, 단분자 가이드 분자를 합성하는 방법은 이중가닥 영역 내에 적어도 하나의 2'-5' 포스포디에스테르 연결을 갖는 단분자 가이드 분자를 생성한다.

올리고뉴클레오타이드 중간체

[0327] 본 개시내용의 특정 구현예는 교차 연결된 합성 가이드 분자의 합성에 유용한 올리고뉴클레오타이드 중간체에 관한 것이다. 일부 구현예에서, 올리고뉴클레오타이드 중간체는 우레아 연결, 티오에테르 연결 또는 포스포디에스테르 연결을 포함하는 가이드 분자의 합성에 유용하다. 일부 구현예에서, 올리고뉴클레오타이드 중간체는 어닐링된 이중가닥을 포함한다.

[0328] 특정 구현예에서, 올리고뉴클레오타이드 중간체는 우레아 연결을 포함하는 가이드 분자의 합성에 유용하다. 일

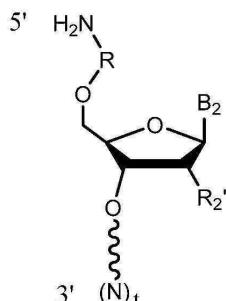
부 구현예에서, 올리고뉴클레오타이드 중간체는 화학식



3'

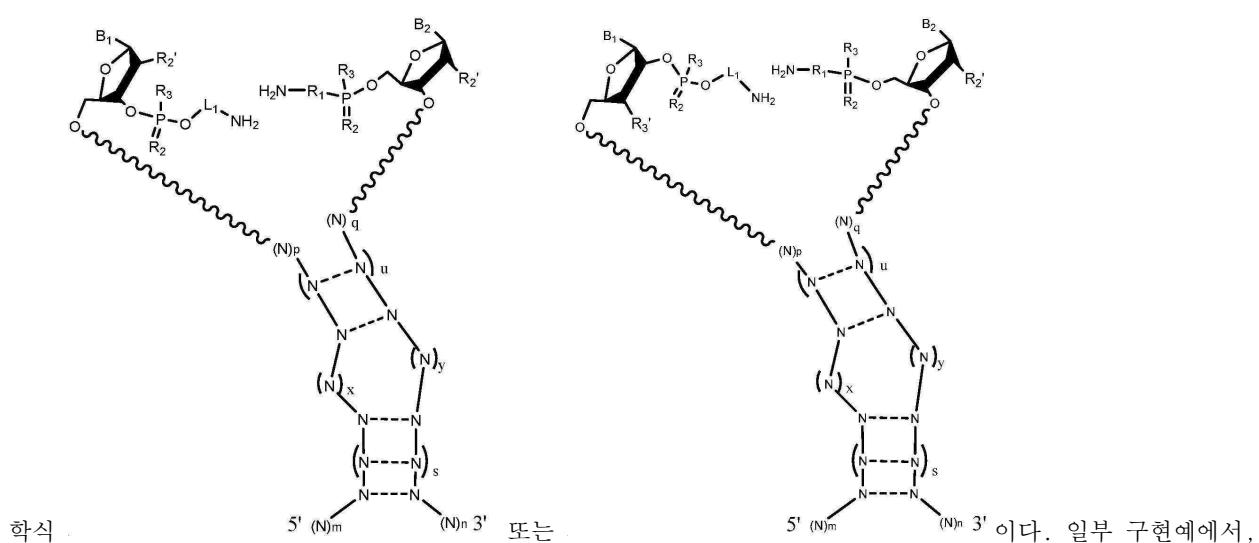
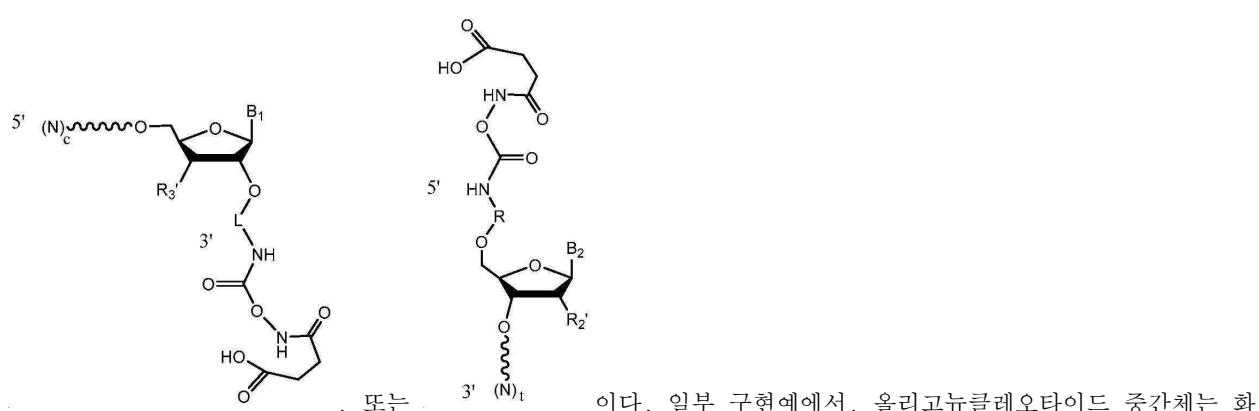
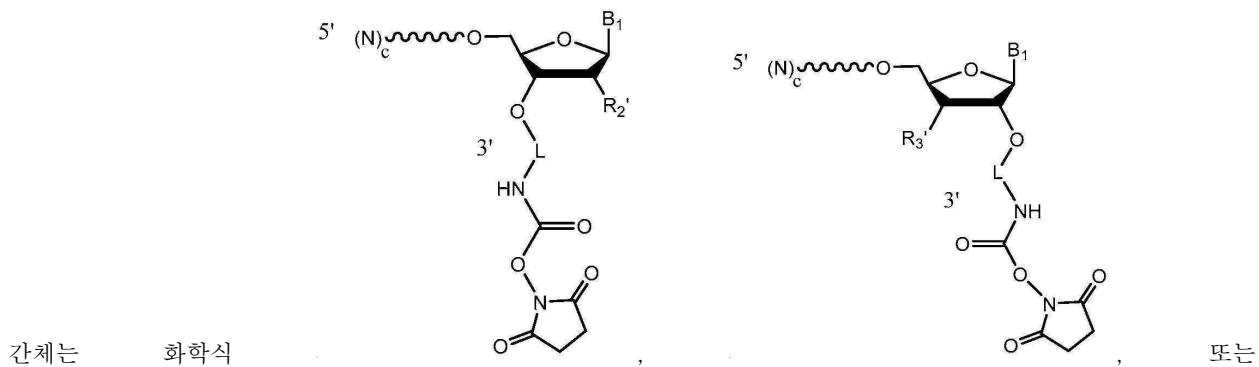
올리고뉴클레오타이드 중간체는

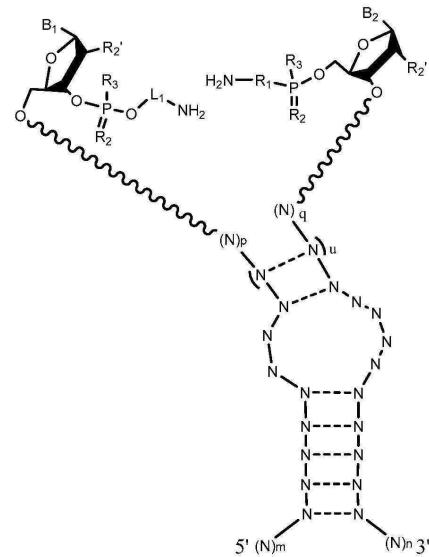
화학식



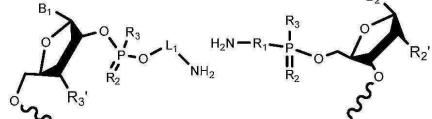
3' (N)t

이다. 일부 구현예에서, 올리고뉴클레오타이드 중





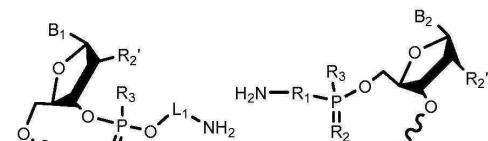
온리고뉴클레오타이드 중간체는 화학식 5' (N)_m (N)_n 3' 이다.



화학식

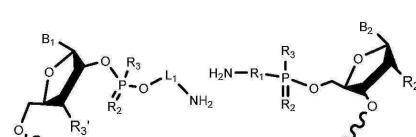
또는

구현예에서, 온리고뉴클레오타이드 중간체는 화학식



일부

온리고뉴클레오타이드 중간체는 화학식

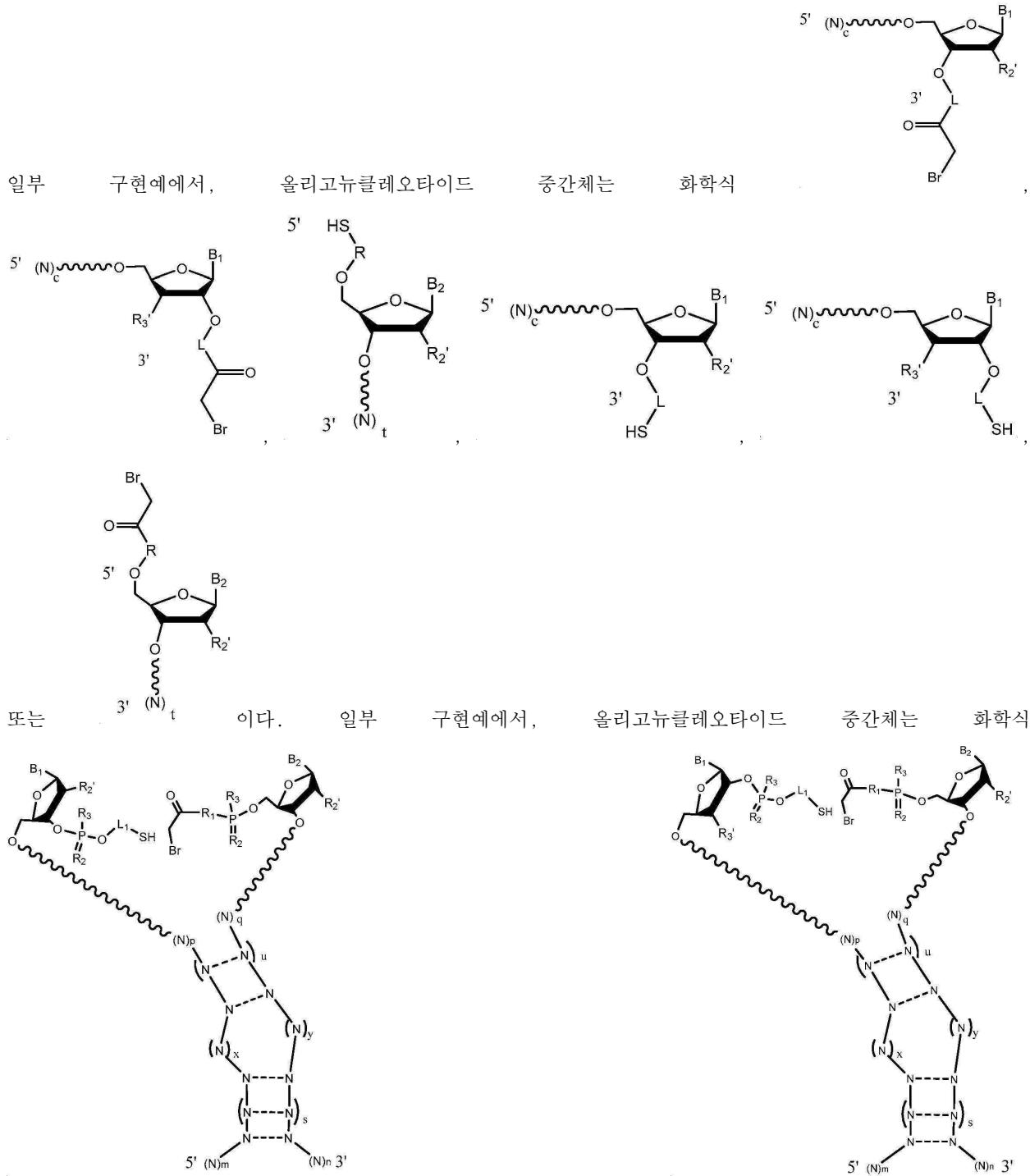


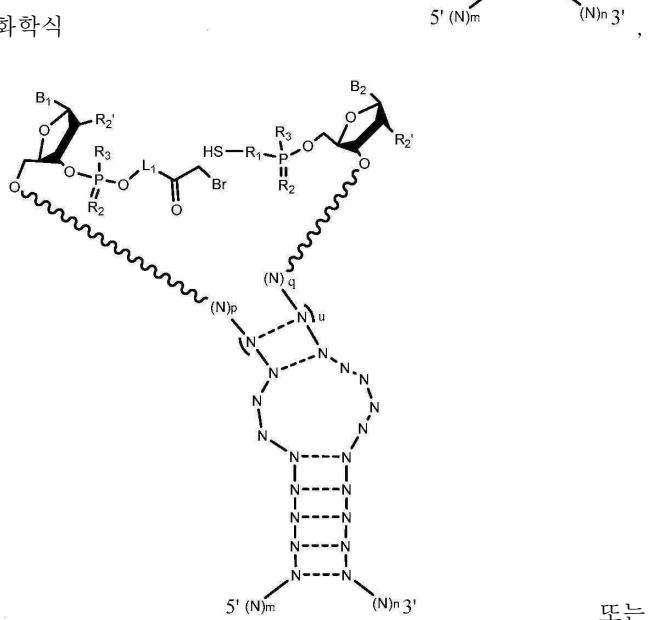
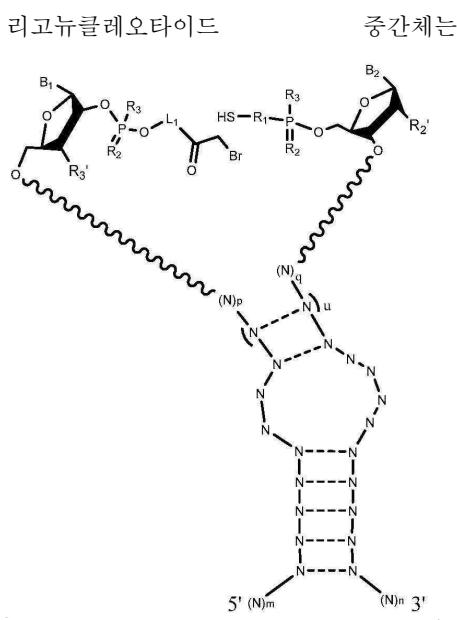
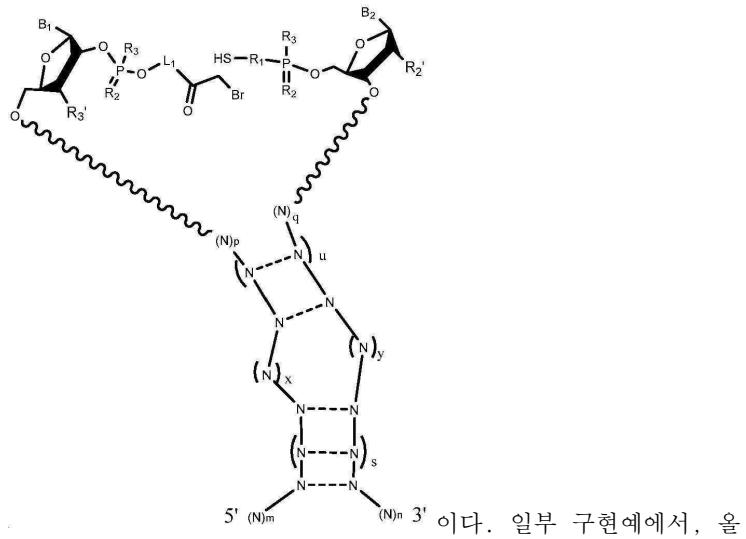
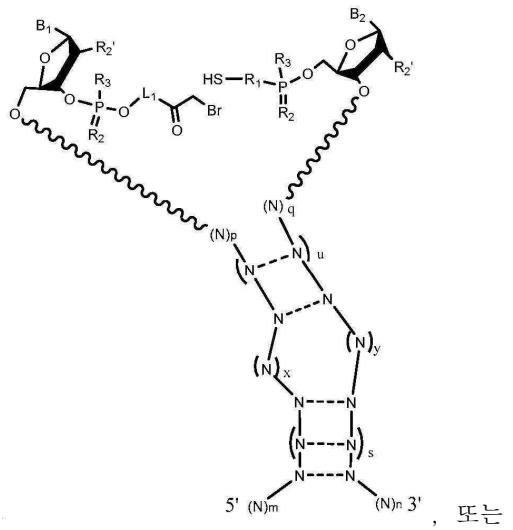
또는 5' (N)_m (N)_n 3'

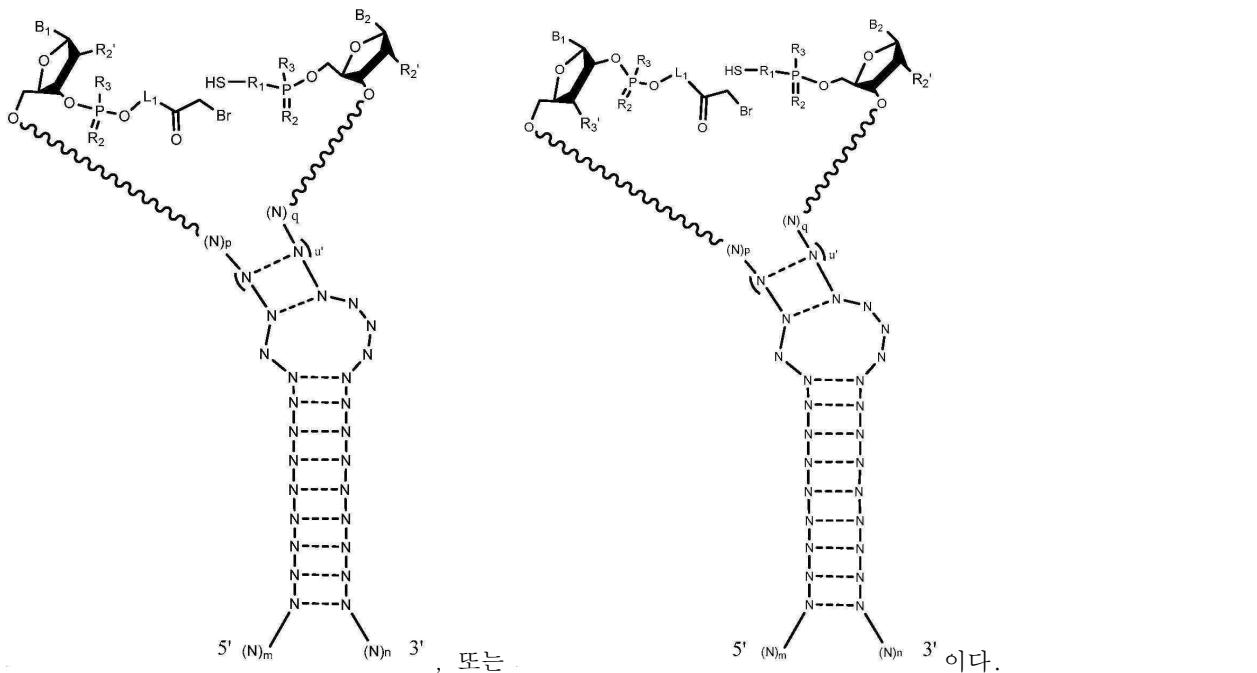
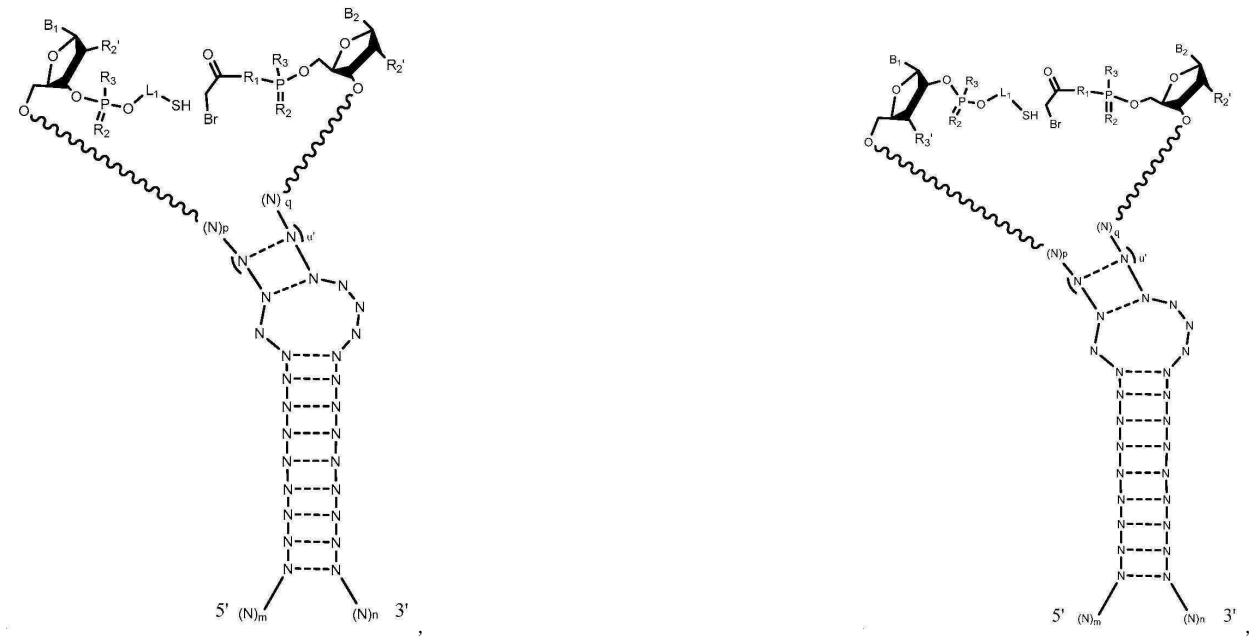
또는 5' (N)_m (N)_n 3']이다.

[0330]

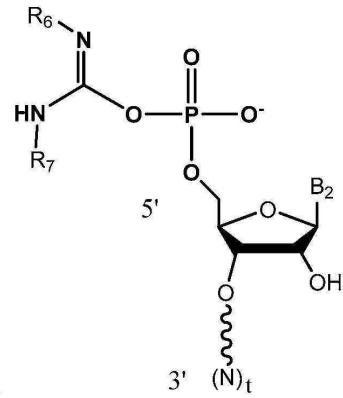
특정 구현예에서, 올리고뉴클레오타이드 중간체는 티오에테르 연결을 포함하는 가이드 분자의 합성에 유용하다.



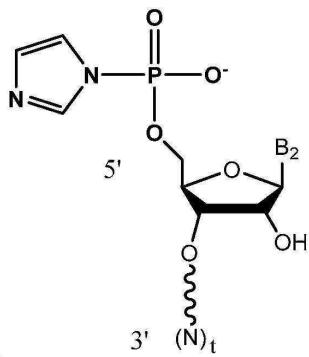




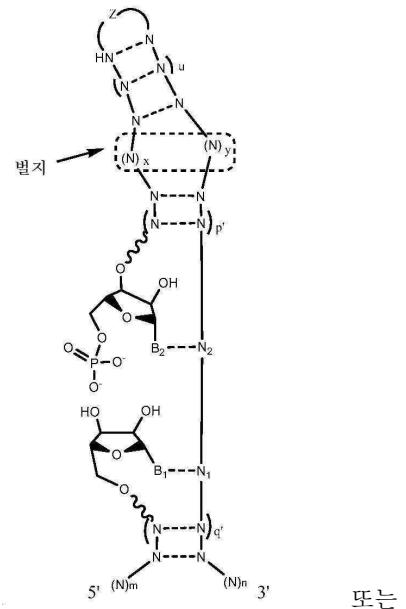
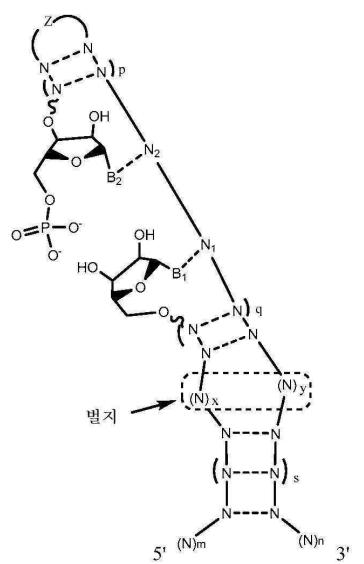
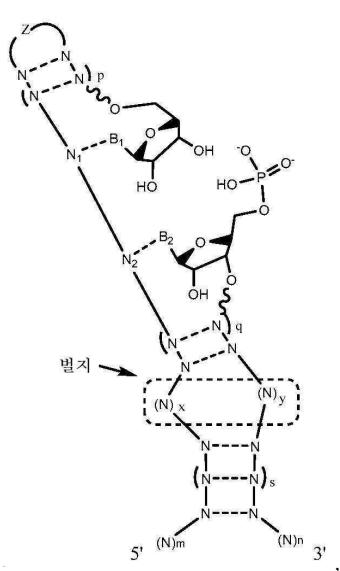
[0331] 특정 구현예에서, 올리고뉴클레오타이드 중간체는 포스포디에스테르 연결을 포함하는 가이드 분자의 합성에 유

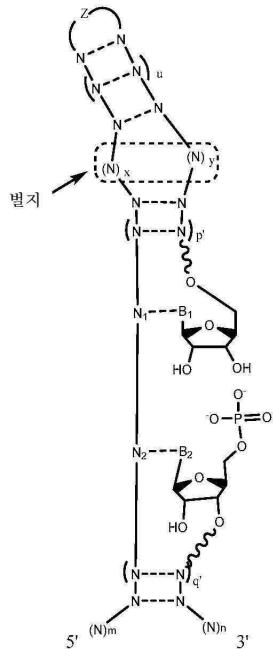


용하다. 일부 구현예에서, 올리고뉴클레오타이드 중간체는 화학식



이고, R6 및 R7은 각각 독립적으로 치환되거나 치환되지 않은 알킬, 또는 치환되거나 치환되지 않은 카보사이클릴이다. 일부 구현예에서, 올리고뉴클레오타이드 중간체는 화학식

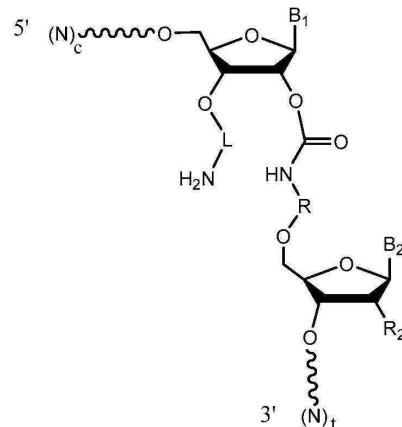




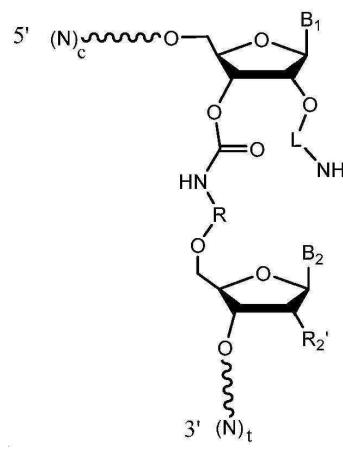
이고, Z는 4개 내지 6개의 뉴클레오타이드 길이, 선택적으로 4개 또는 6개의 뉴클레오타이드 길이의 뉴클레오타이드 루프를 나타낸다.

[0332]

본 개시내용의 특정 구현예는 교차 연결 반응에서 부산물로서 형성되는 올리고뉴클레오타이드 화합물에 관한 것이다. 이들 올리고뉴클레오타이드 화합물은 가이드 분자로서 유용할 수 있거나 유용하지 않을 수 있다. 일부 구



현 예에서, 올리고뉴클레오타이드 화합물은 화학식 또는



이다.

[0333]

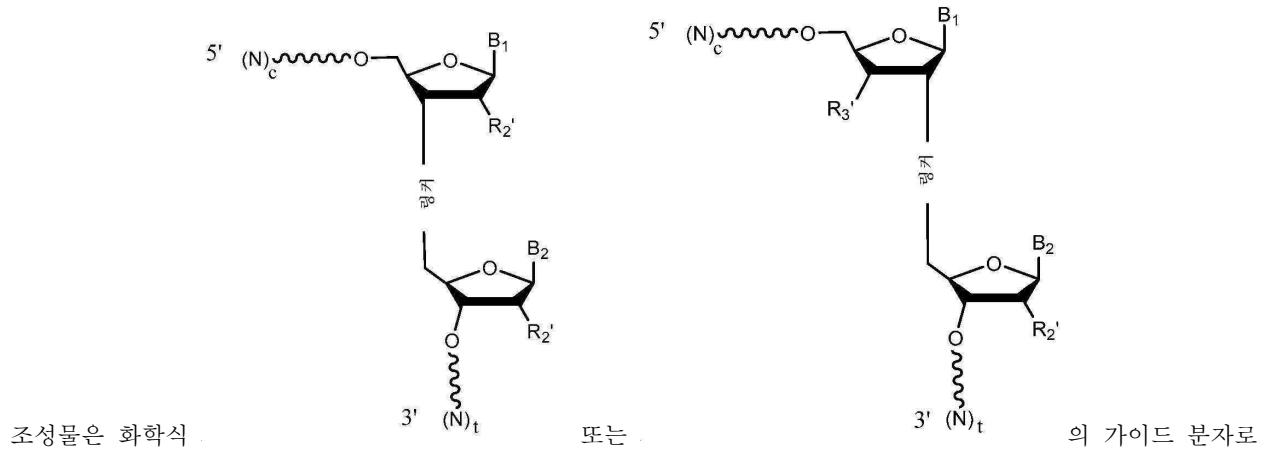
화학적으로 접합된 가이드 분자의 조성물

[0334]

본 개시내용의 특정 구현예는 상기에 기재된 합성 가이드 분자를 포함하는 조성물 및 상기에 기재된 방법에 의해 생성된 조성물에 관한 것이다. 일부 구현예에서, 본 조성물은 조성물 중 90% 초과의 가이드 분자가 전장 가이드 분자인 것을 특징으로 한다. 일부 구현예에서, 본 조성물은 조성물 중 85% 초과의 가이드 분자가 동일한

표적화 도메인 서열을 포함하는 것을 특징으로 한다.

일부 구현예에서, 본 조성물은 정제 단계가 적용되지 않았다. 일부 구현예에서, CRISPR 시스템용 가이드 분자의



조성물은 화학식

또는

의 가이드 문자로

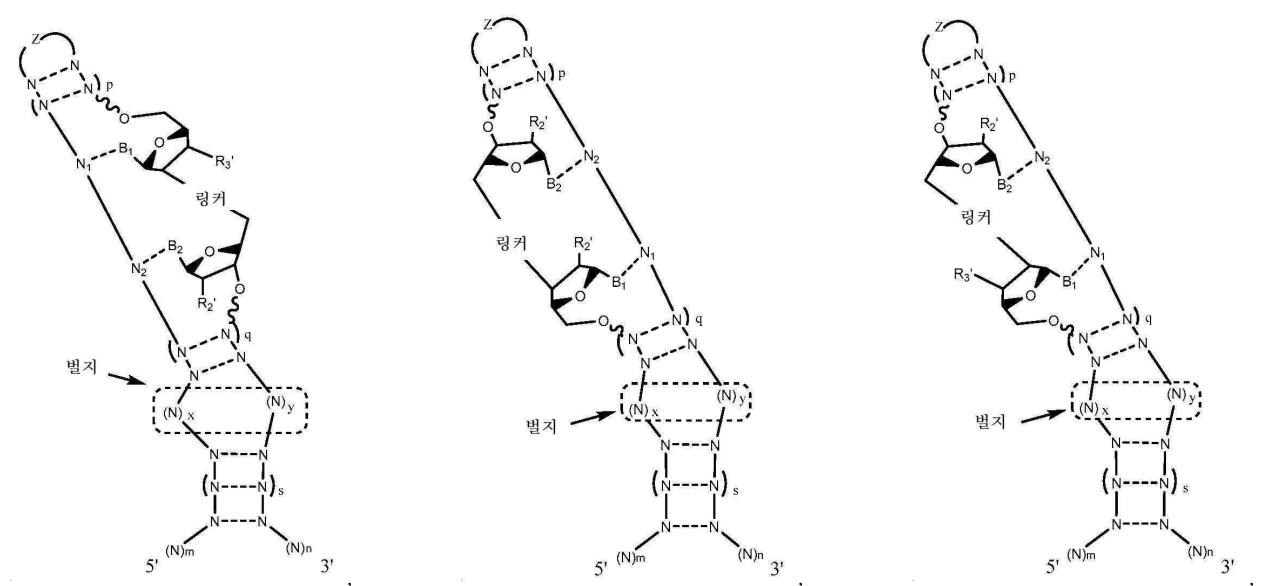
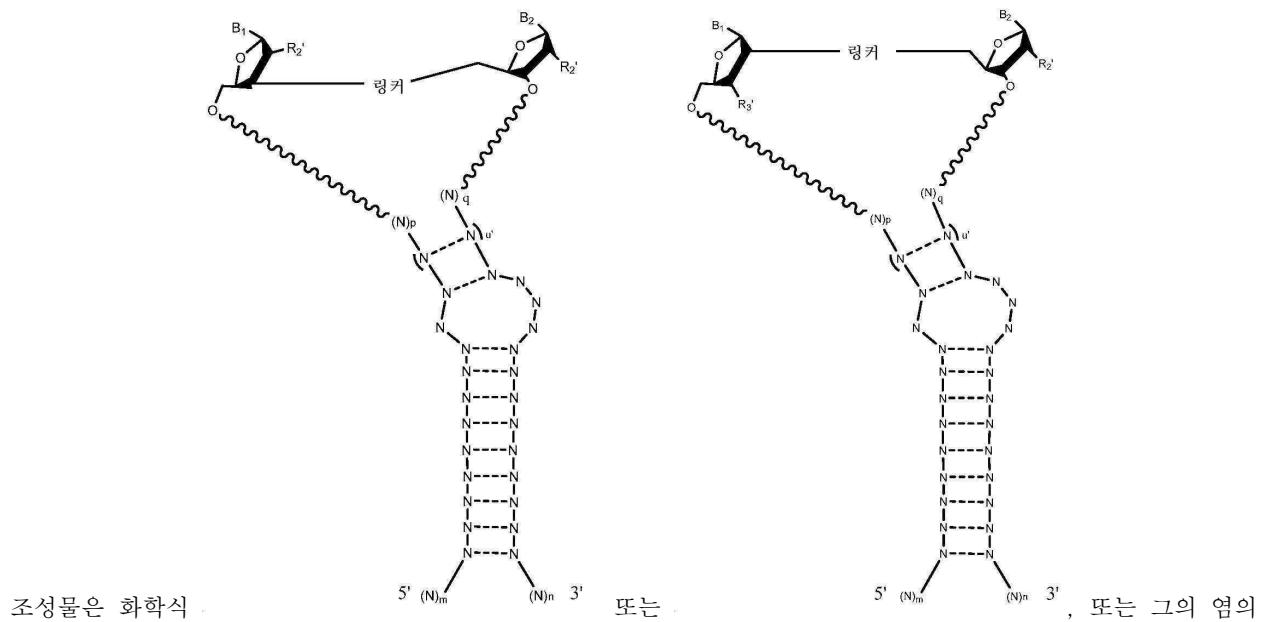


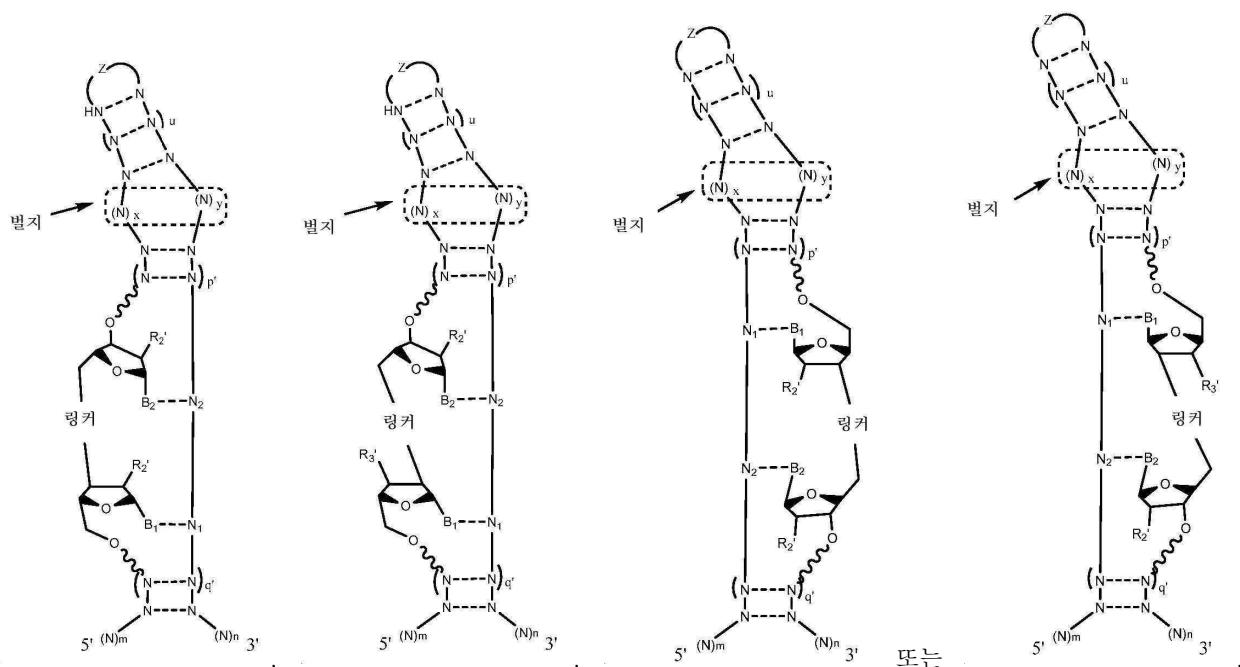
필수적으로 이루어진다. 일부 구현예에서, 본 조성물은 화학식

또는



³, 또는 그의 염의 가이드 문자로 필수적으로 이루어진다. 일부 구현예에서, 본

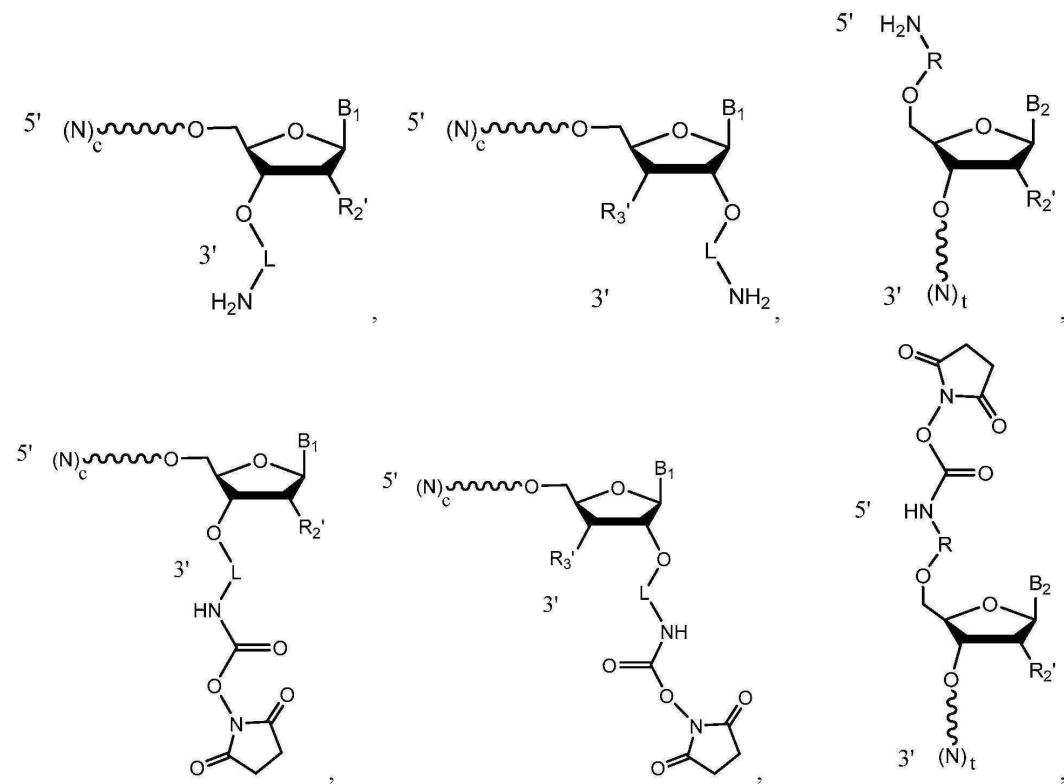


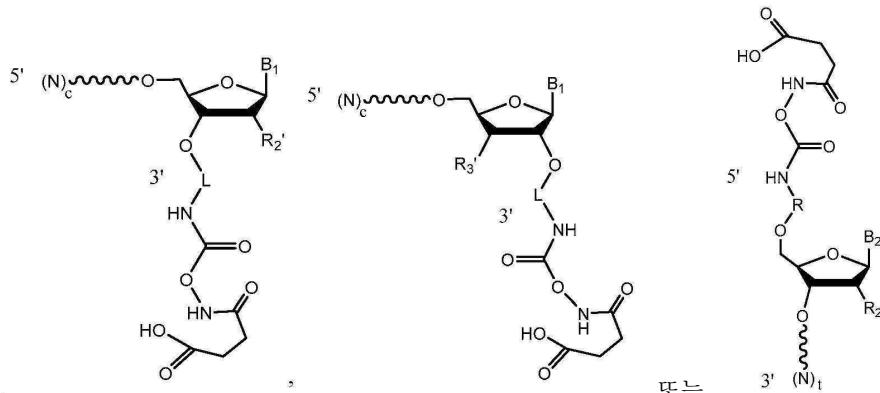


또는 그의 염의 가이드 분자로 필수적으로 이루어진다.

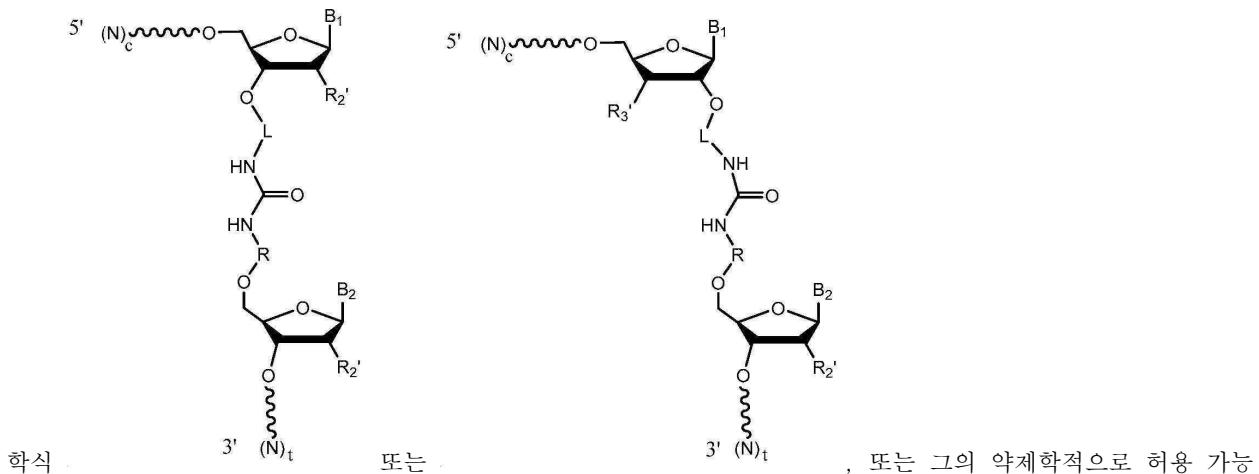
[0336]

일부 구현예에서, 본 조성물은 합성 가이드 분자의 존재 또는 부재하의 올리고뉴클레오타이드 중간체(상기에 기재됨)를 포함한다. 일부 구현예에서, 조성물 중 올리고뉴클레오타이드 중간체는 화학식



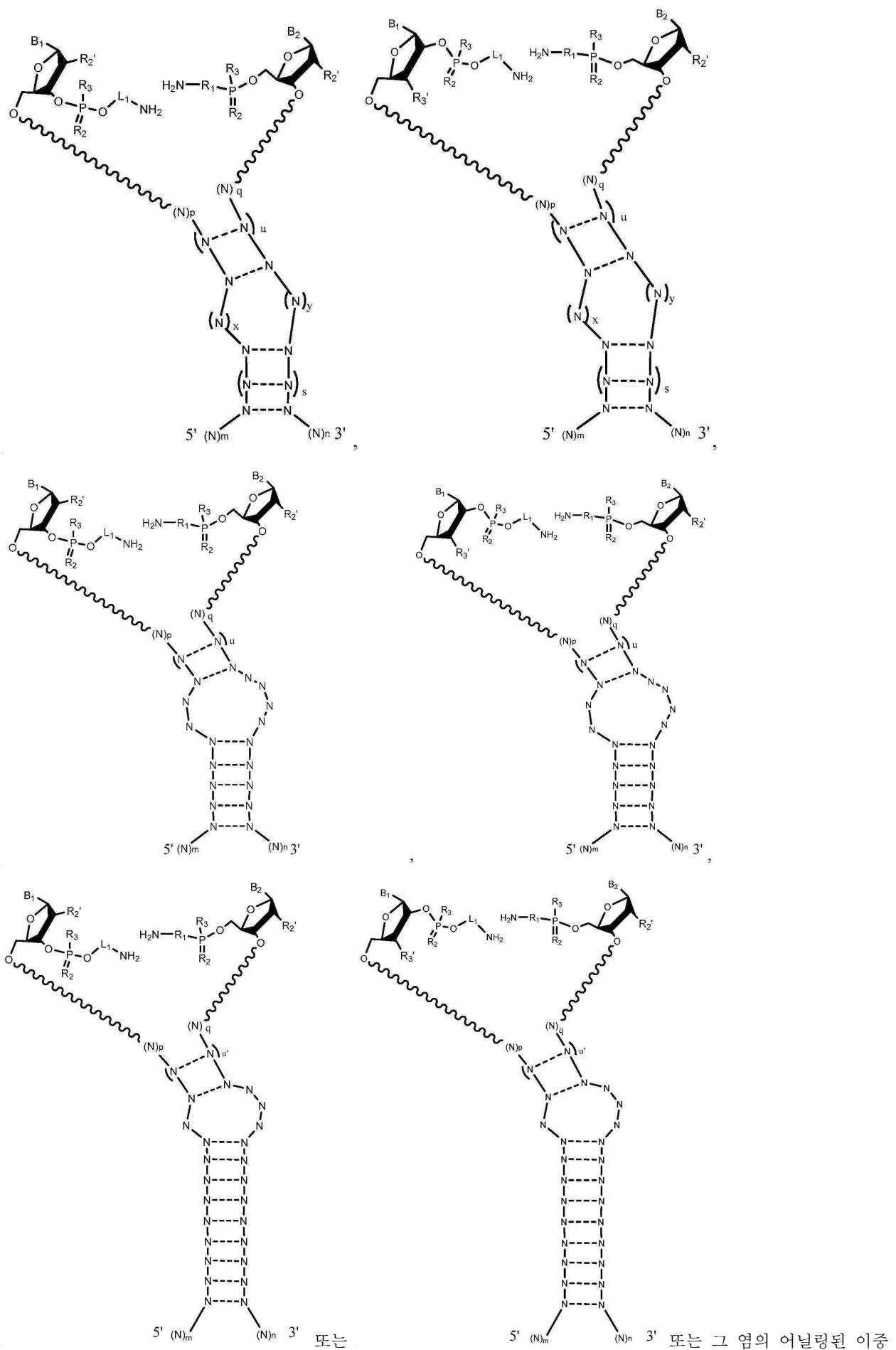


[0337]

, 또는 $3' (N)_t$ 이고, 합성 가이드 분자는 화

한 염이다. 일부 구현예에서, 본 조성물은 화학식





가닥을 갖는 올리고뉴클레오타이드 중간체를 포함한다.

[0338]

일부

구현예에서,

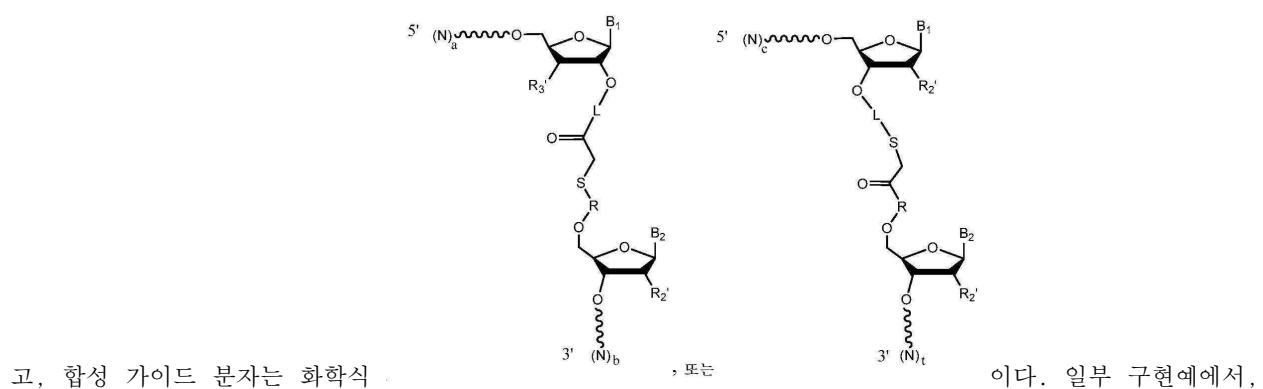
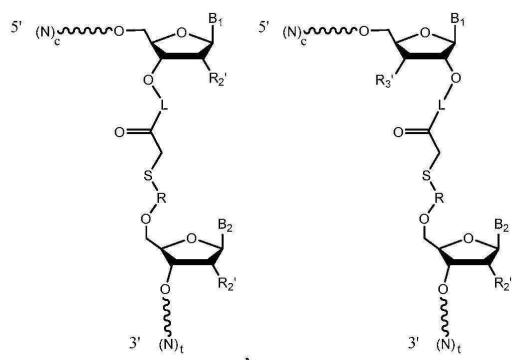
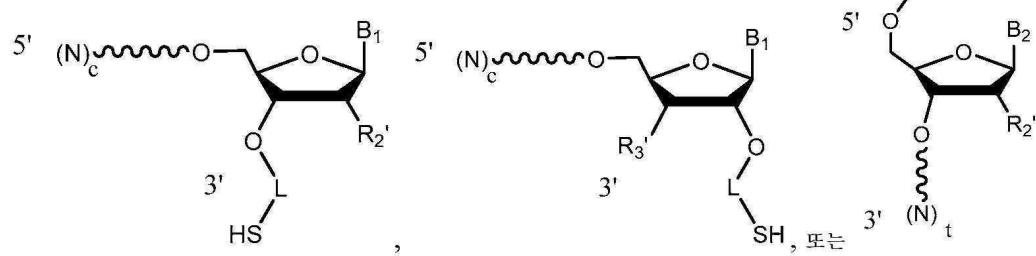
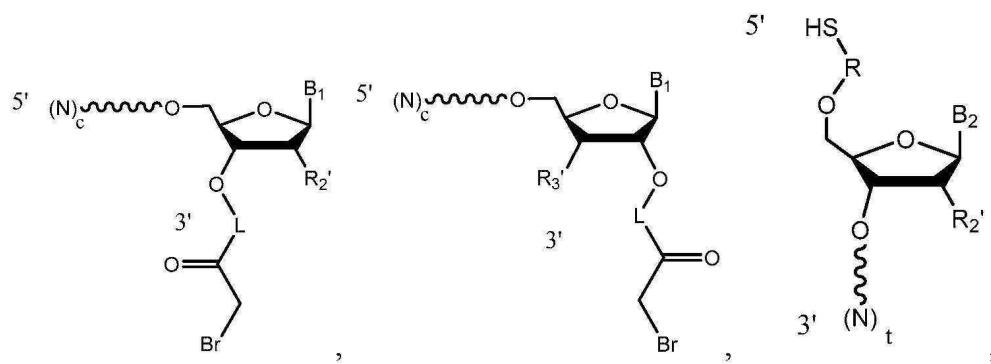
조성물

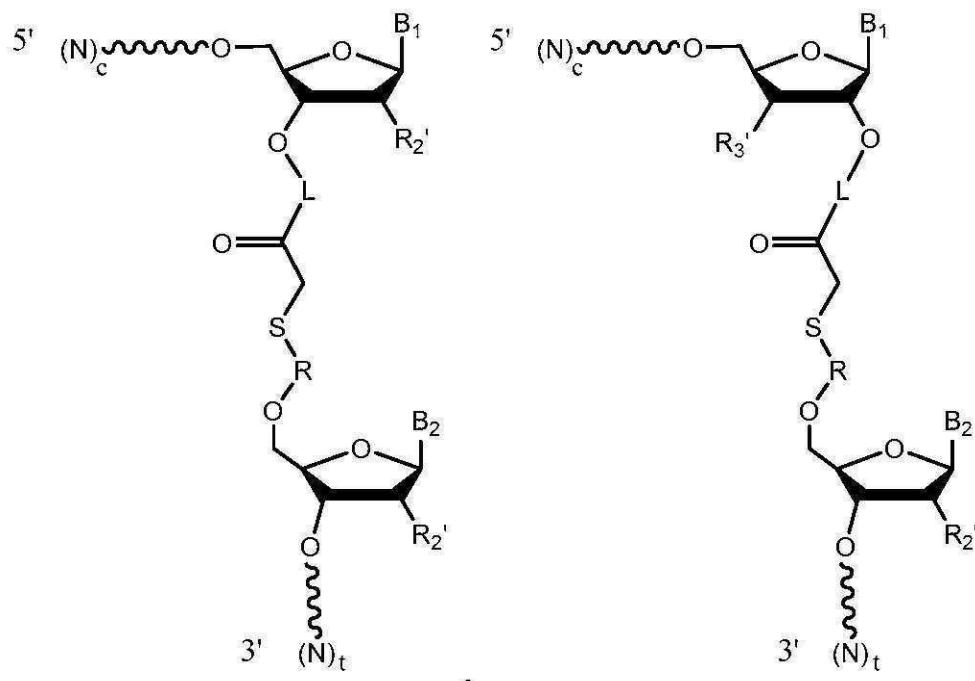
종

올리고뉴클레오타이드

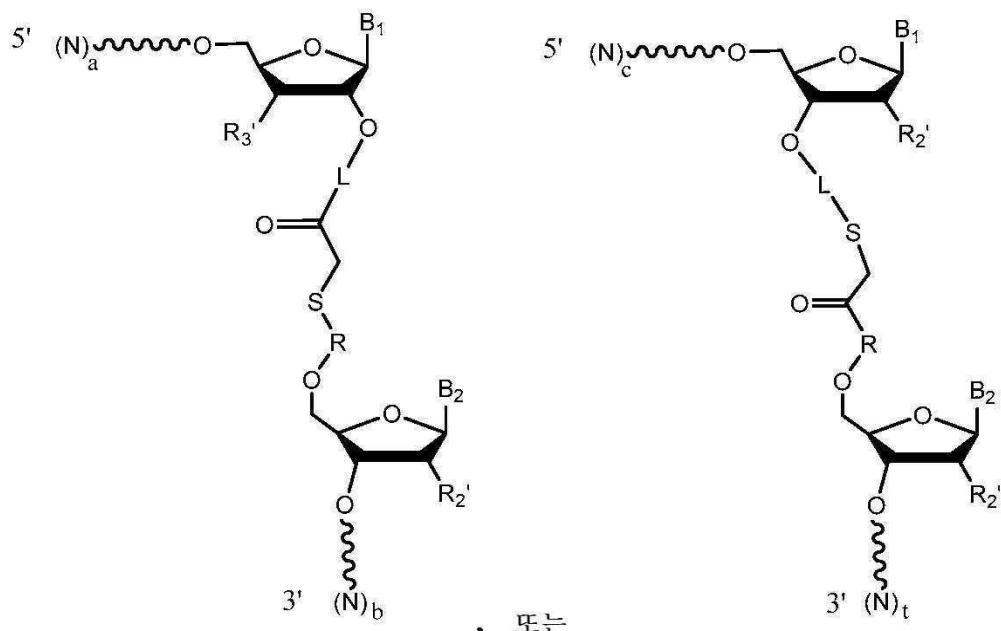
중간체는

화학식





본 조성물은

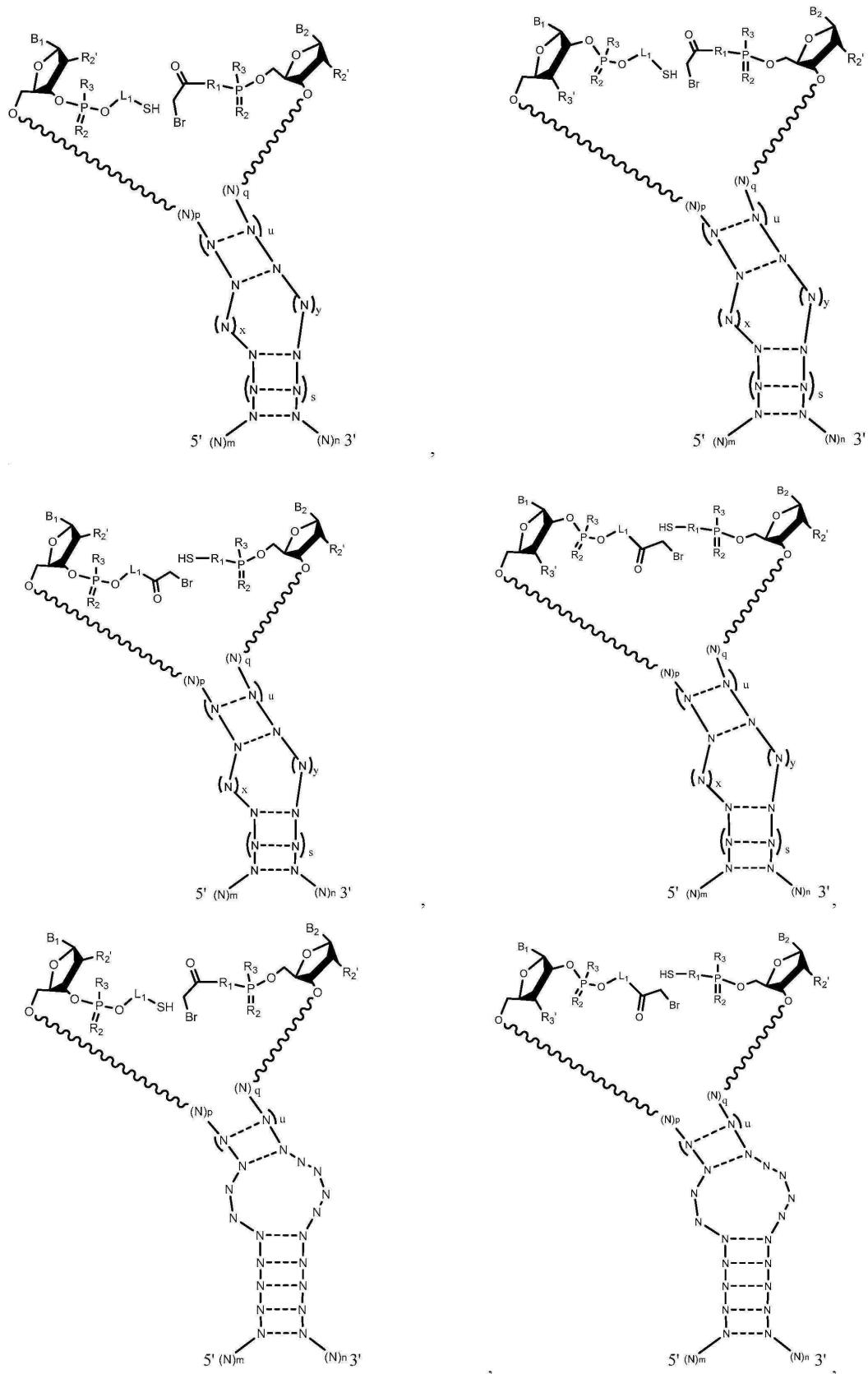


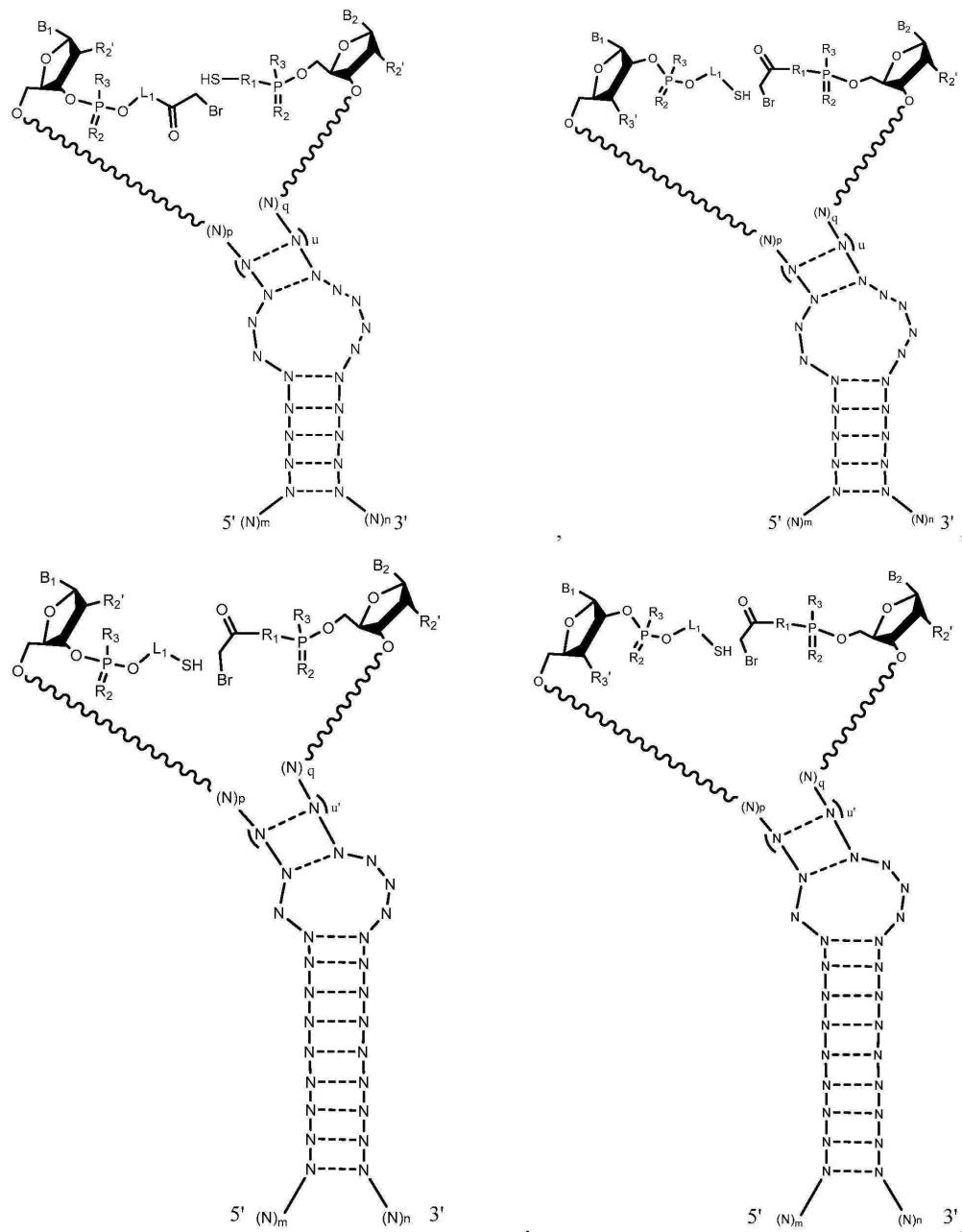
의

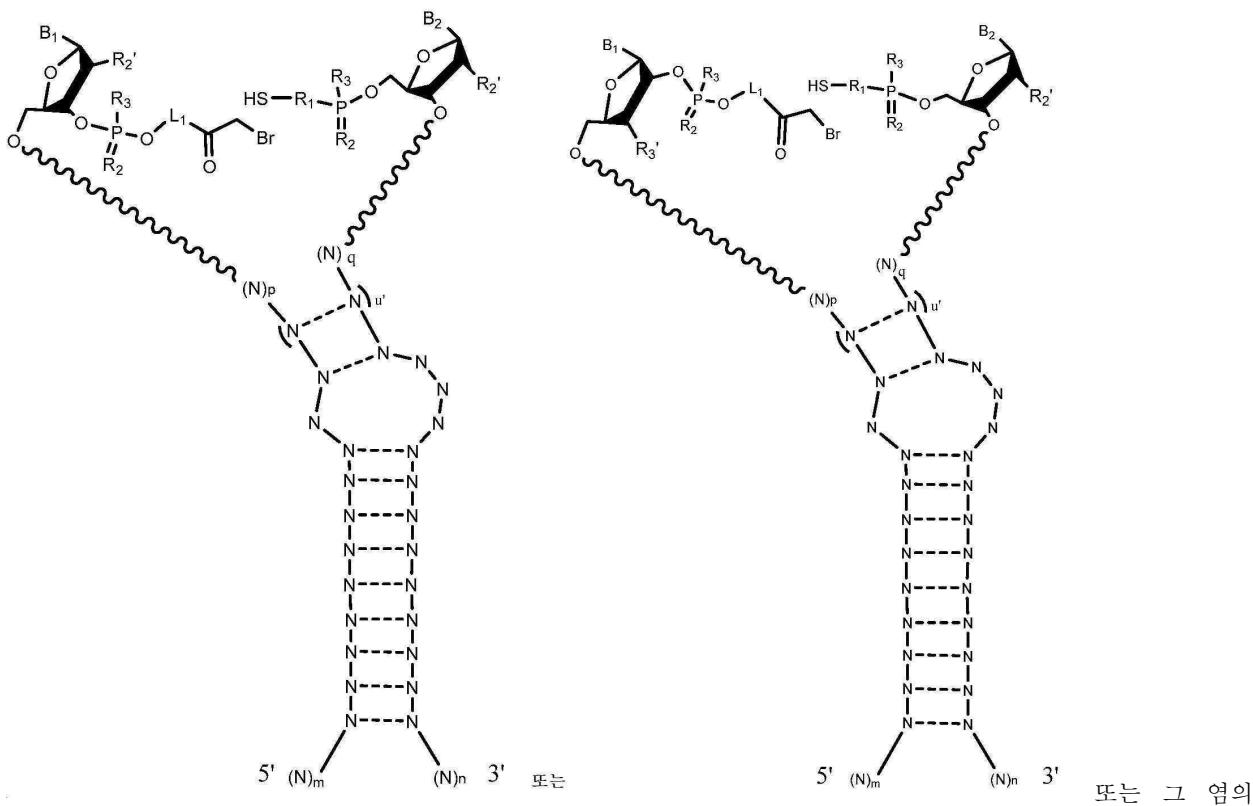
존재

또는

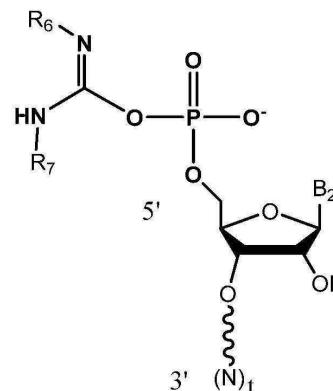
의 합성 가이드 분자
부재하에





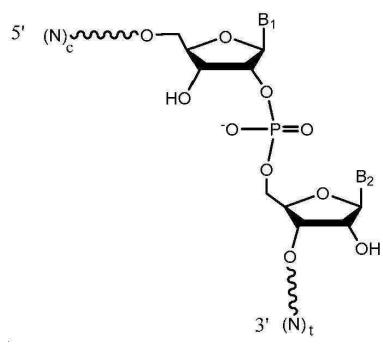


어닐링된 이중가닥을 갖는 올리고뉴클레오타이드 중간체를 포함한다.



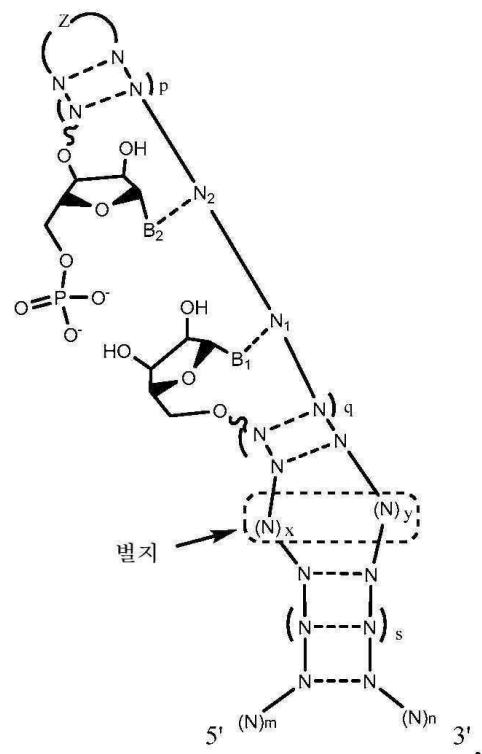
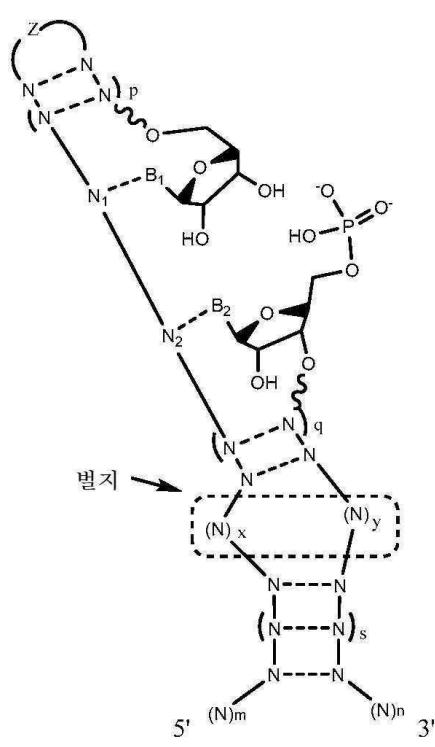
일부 구현 예에서, 조성을 중 을 리고 뉴클레오타이드 중간체는 화학식 3'-(N)t , 또는

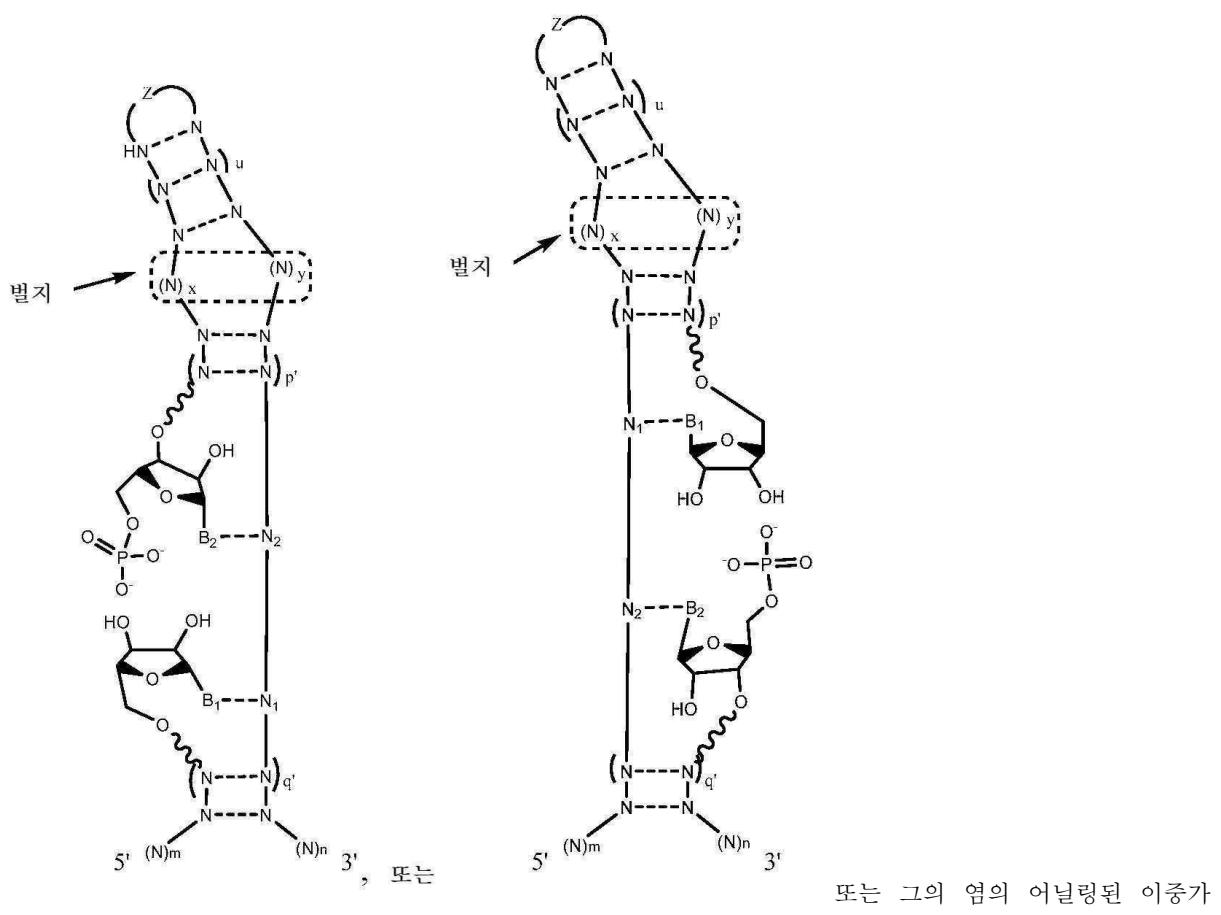




이다.

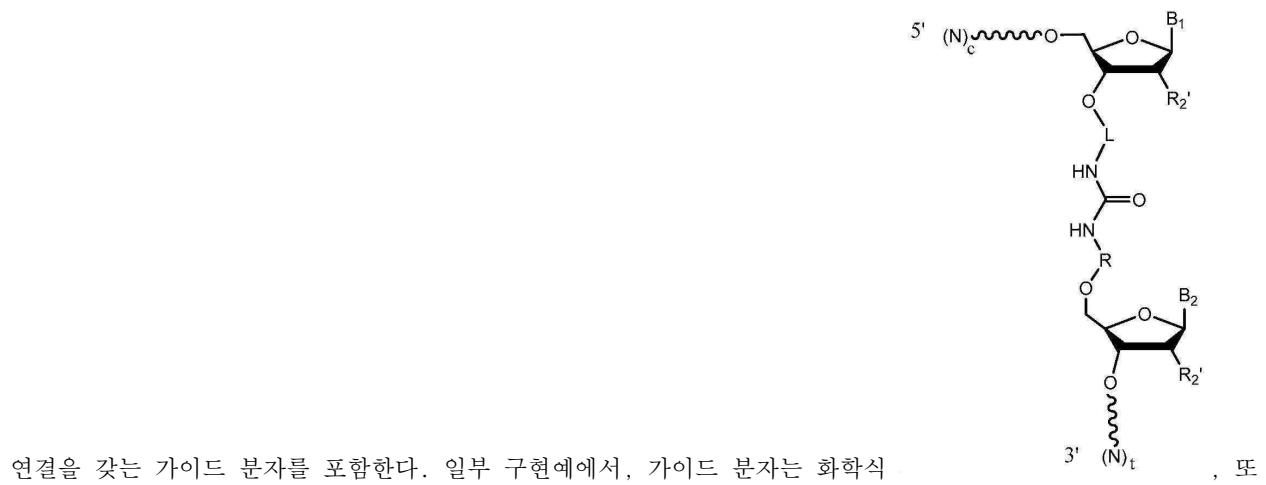
일부 구현예에서, 본 조성물은 화학식

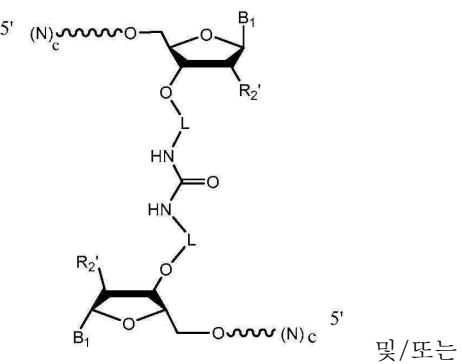




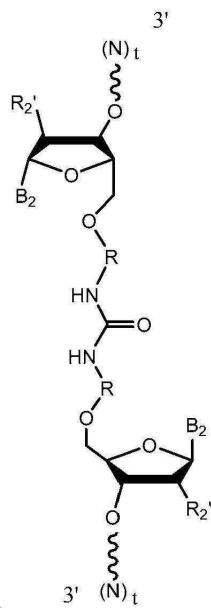
닥을 갖는 올리고뉴클레오타이드 중간체를 포함한다.

일부 구현예에서, 본 조성물은 동종이량체를 실질적으로 무함유한다. 일부 구현예에서, 동종이량체 및/또는 부산물을 실질적으로 무함유하는 조성물은 동종이중기능성 교차 연결 시약을 포함하는 방법을 사용하여 합성된 가이드 분자를 포함한다. 일부 구현예에서, 동종이량체 및/또는 부산물을 실질적으로 무함유하는 조성물은 우레이아

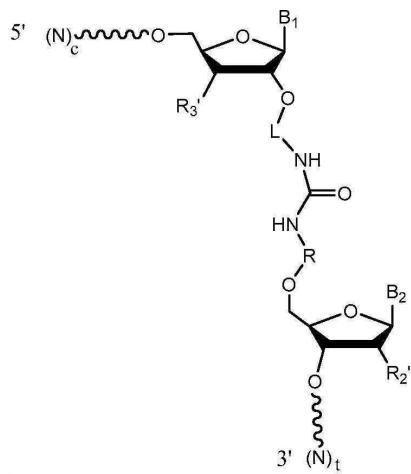




는 그의 약제학적으로 허용 가능한 염이고, 본 조성물은 화학식

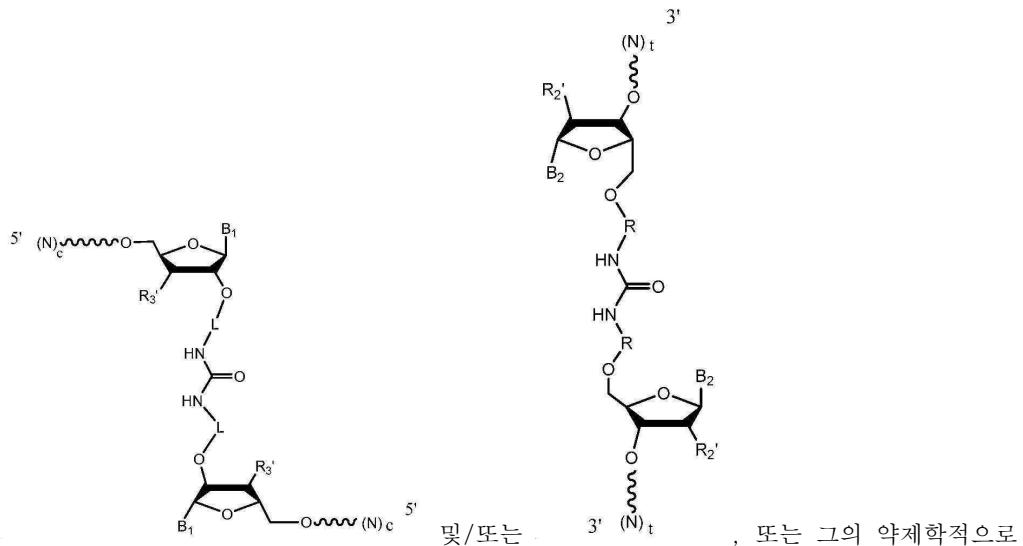


, 또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로 무함유한다. 일부 구현예에

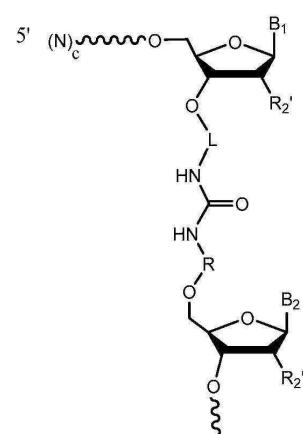


서, 가이드 분자는 화학식

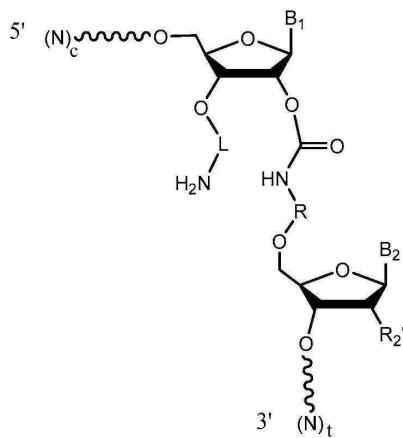
, 또는 그의 약제학적으로 허용 가능한 염이고,



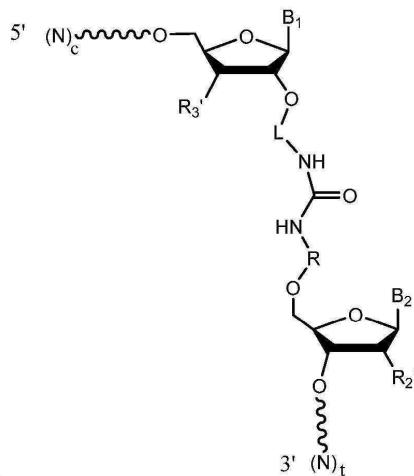
[0341] 일부 구현예에서, 본 조성물은 부산물을 실질적으로 무함유한다. 일부 구현예에서, 본 조성물은 우레아 연결을



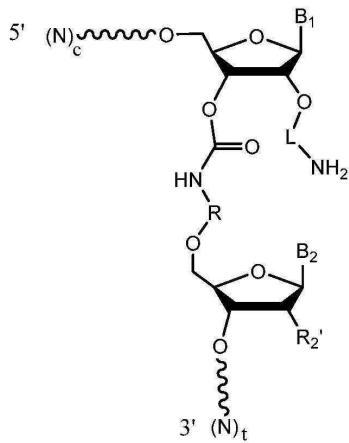
포함하는 가이드 분자를 포함한다. 일부 구현예에서, 본 조성물은 화학식 $3'-(N)_t$, 또는 그 의 약제학적으로 허용 가능한 염의 가이드 분자를 포함하고, 본 조성물은 화학식



의 분자를 실질적으로 무함유한다. 일부 구현예에서, 본 조성물은 화학식

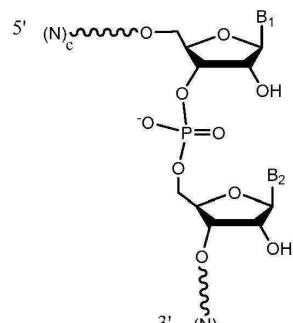


, 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 포함하고, 본

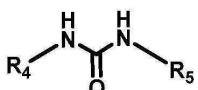


조성물은 화학식 $3' \text{ (N)}_t$ 의 분자를 실질적으로 무함유한다.

[0342] 일부 구현예에서, 본 조성물은 부산물을 실질적으로 무함유하지 않는다. 일부 구현예에서, 본 조성물은 (a)

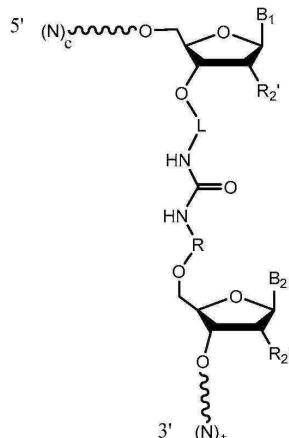


CRISPR 시스템용 합성 단분자 가이드 분자로서, 가이드 분자는 화학식 $3' \text{ (N)}_t$, 또는 그의 약제학적으로 허용 가능한 염인 CRISPR 시스템용 합성 단분자 가이드 분자; 및 (b) (i) 카보디이미드, 또는 그의 염; (ii) 이미다졸, 시아노이미다졸, 피리딘 및 디메틸아미노피리딘, 또는 그의 염; 및 (iii) 화학식



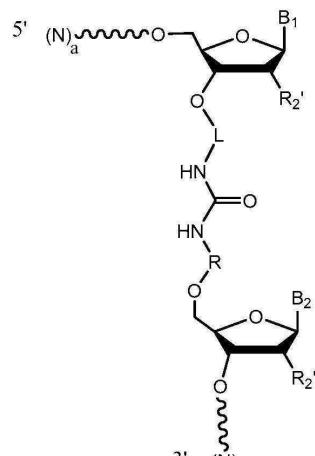
또는 그의 염의 화합물로서, R_4 및 R_5 는 각각 독립적으로 치환되거나 치환되지 않은 알킬, 또는 치환되거나 치환되지 않은 카보사이클릴인 화합물 중 하나 이상을 포함한다. 일부 구현예에서, 카보디이미드는 EDC, DCC, 또는 DIC이다. 일부 구현예에서, 본 조성물은 EDC를 포함한다. 일부 구현예에서, 본 조성물은 이미다졸을 포함한다.

[0343] 일부 구현예에서, 본 조성물은 $n+1$ 및/또는 $n-1$ 종을 실질적으로 무함유한다. 일부 구현예에서, 본 조성물은 기준 가이드 분자 서열과 비교하여 약 10%, 5%, 2%, 1%, 또는 0.1% 미만의 절두를 포함하는 가이드 분자를 포함한다. 일부 구현예에서, 적어도 약 85%, 90%, 95%, 98%, 또는 99%의 가이드 분자는 기준 가이드 분자 서열의 해당 5' 서열과 100% 동일한 가이드 분자의 1개 내지 20개의 뉴클레오타이드를 포함하는 5' 서열을 포함한다.



[0344] 일부 구현예에서, 본 조성물은 화학식

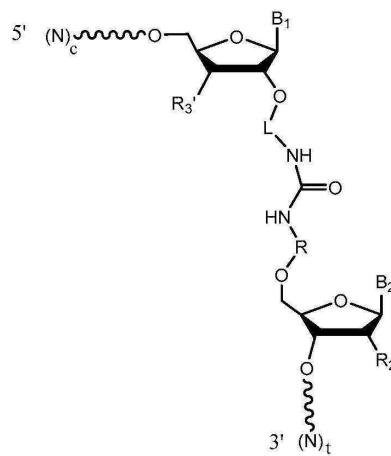
, 또는 그의 약제학적으로 허용 가능한 염의



가이드 분자를 필수적으로 포함하고, 본 조성물은 화학식

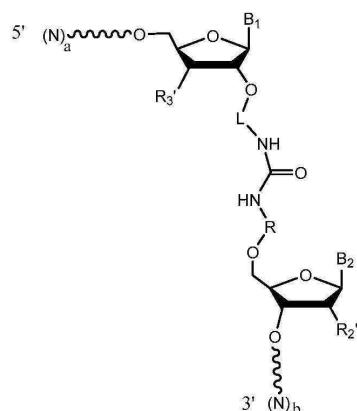
, 또는 그의 약제학적으

로 허용 가능한 염의 분자를 실질적으로 무함유하고, a는 c와 동일하지 않고/않거나; b는 t와 동일하지 않다.



일부 구현예에서, 본 조성물은 화학식

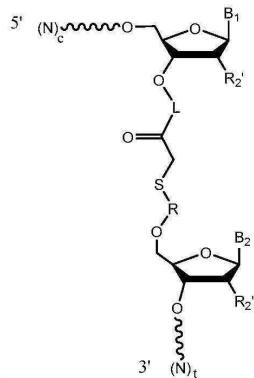
, 또는 그의 약제학적으로 허용 가능



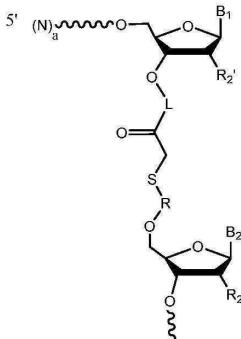
한 염을 필수적으로 포함하고, 본 조성물은 화학식

, 또는 그의 약제학적으로

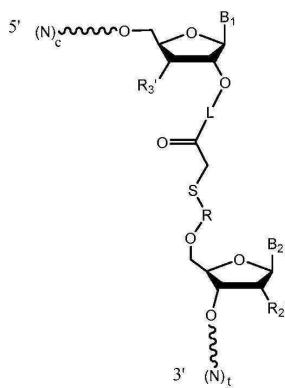
허용 가능한 염의 분자를 실질적으로 무함유하고, a는 c와 동일하지 않고/않거나; b는 t와 동일하지 않다.



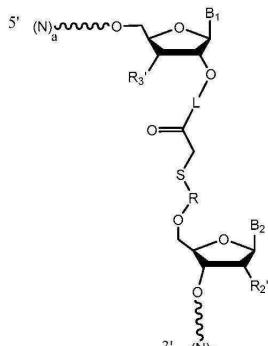
일부 구현예에서, 본 조성물은 화학식 3' (N_t), 또는 그의 약제학적으로 허용 가능한 염의 가이



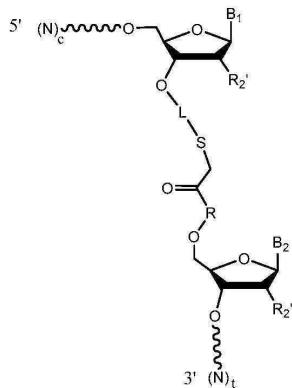
드 분자를 필수적으로 포함하고, 본 조성물은 화학식 $3' \text{--} (\text{N})_b \text{--} 5'$, 또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로 무함유하고, a는 c와 동일하지 않고/않거나; b는 t와 동일하지 않다. 일부 구현 예



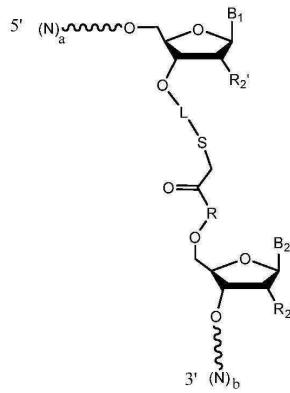
에서, 본 조성물은 화학식 $^3\text{N}_t$, 또는 그의 약제학적으로 허용 가능한 염의 가이드 문자를



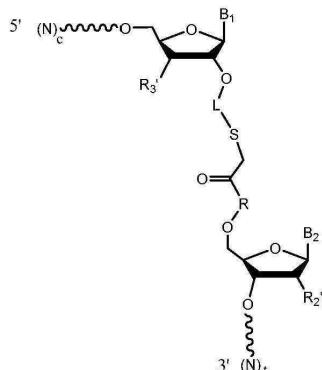
필수적으로 포함하고, 본 조성물은 화학식 $\text{c}^{\text{t}}(\text{N})_{\text{b}}$, 또는 그의 약제학적으로 허용 가능한 염의 문자를 실질적으로 무함유하고, a는 c와 동일하지 않고/않거나; b는 t와 동일하지 않다. 일부 구현예에서, 본



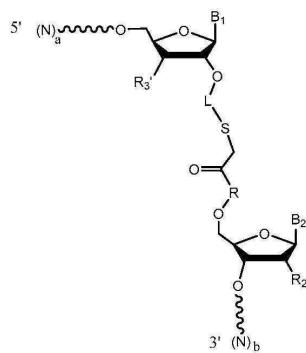
조성물은 화학식 . , 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 필수적으



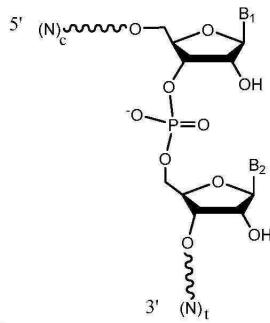
로 포함하고, 본 조성물은 화학식 . , 또는 그의 약제학적으로 허용 가능한 염의 분자
를 실질적으로 무함유하고, a는 c와 동일하지 않고/않거나; b는 t와 동일하지 않다. 일부 구현예에서, 본 조성



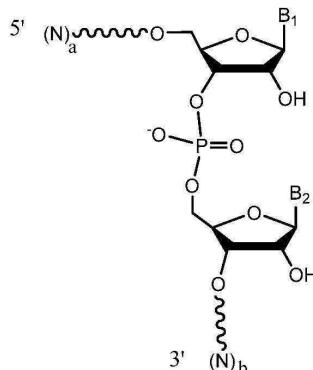
물은 화학식 . , 또는 그의 약제학적으로 허용 가능한 염의 가이드 분자를 필수적으



로 포함하고, 본 조성물은 화학식 . , 또는 그의 약제학적으로 허용 가능한 염의 분자
를 실질적으로 무함유하고, a는 c와 동일하지 않고/않거나; b는 t와 동일하지 않다. 일부 구현예에서, a는 c 미
만이고/이거나, b는 t 미만이다.

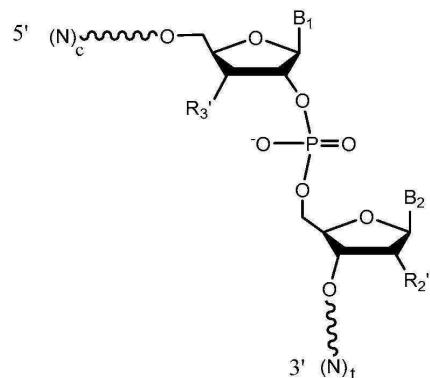


[0346] 일부 구현예에서, 본 조성물은 화학식 , 또는 그의 약제학적으로 허용 가능한 염의 가

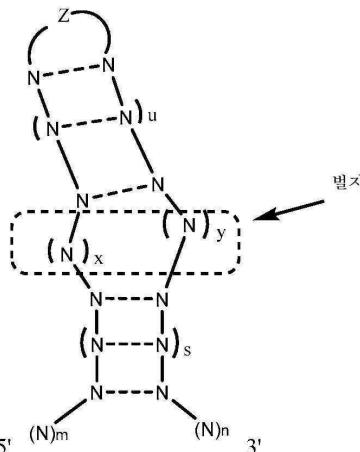


이드 분자를 포함하고, 본 조성물은 화학식 , 또는 그의 약제학적으로 허용 가능한 염의 분자를 실질적으로 무함유하고, $a+b$ 는 $c+t-k$ 이고, k 는 정수 1 내지 10(상하한 포함)이다.

[0347] 일부 구현예에서, 본 조성물은 CRISPR 시스템용 합성 단분자 가이드 분자를 포함하고, 가이드 분자는 화학식



, 또는 그의 약제학적으로 허용 가능한 염이고, 화학식 내에 도시된 2'-5' 포스포디에스테르 연결은 이중가닥 내의 2개의 뉴클레오타이드 사이에 존재한다. 일부 구현예에서, 가이드



분자는 화학식 , 또는 그의 약제학적으로 허용 가능한 염이고, 화학식 내에 도시된 이중가닥 영역의 2개의 뉴클레오타이드 사이의 적어도 하나의 포스포디에스테르 연결은 2'-5' 포스포디에스테르 연결이다. 일부 구현예에서, 2'-5' 포스포디에스테르 연결은 별지의 5'에 위치한 2개의 뉴클레오타이드 사이에 존재한다. 일부 구현예에서, 2'-5' 포스포디에스테르 연결은 뉴클레오타이드 루프 Z의 5' 및 별지의 3'에 위치한 2개의 뉴클레오타이드 사이에 존재한다. 일부 구현예에서, 2'-5' 포스포디에스테르 연결은 뉴클레

오타이드 루프 Z의 3' 및 벌지의 5'에 위치한 2개의 뉴클레오타이드 사이에 존재한다. 일부 구현예에서, 2'-5' 포스포디에스테르 연결은 벌지의 3'에 위치한 2개의 뉴클레오타이드 사이에 존재한다.

[0348] 가이드 분자 설계

표적외 분석뿐만 아니라, 표적 서열의 선택 및 검증을 위한 방법은 종래에, 예를 들어 문헌[Mali; Hsu; Fu et al., 2014 Nat biotechnol 32(3): 279-84, Heigwer et al., 2014 Nat methods 11(2):122-3; Bae et al. (2014) Bioinformatics 30(10): 1473-5; 및 Xiao A et al. (2014) Bioinformatics 30(8): 1180-1182]에 기재되었다. 각각의 이들 문헌은 본 명세서에 참조로 포함된다. 비 제한적인 예로서, 가이드 분자 설계는, 예를 들어 게놈을 통한 전체 표적외 활성을 최소화하기 위해, 사용자의 표적 서열에 상응하는 잠재적 표적 서열의 선택을 최적화하기 위한 소프트웨어 툴의 사용을 포함할 수 있다. 표적외 활성은 절단에 제한되는 것은 아니며, 각각의 표적외 서열에서의 절단 효율은, 예를 들어 실험적으로 유도된 각종 방식을 사용하여 예측될 수 있다. 이들 및 다른 가이드 선택 방법이 문헌[Meder 및 Cotta-Ramusino]에 상세하게 기재되어 있다.

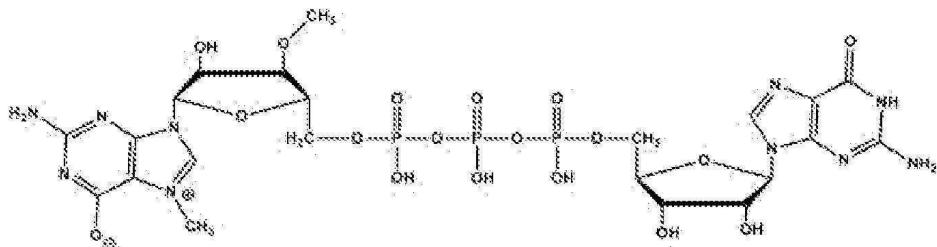
[0350] 합성 단분자 가이드 분자에서의 줄기 루프 구조 및 화학적 연결의 위치를 또한 설계 할 수 있다. 본 발명자들은 길스 자유 에너지 차이(ΔG)를 사용하여, 화학적 접합 반응의 결찰 효율을 예측하는 것의 가치를 확인하였다. ΔG 의 계산은 OligoAnalyzer(www.idtdna.com/calc/analyzer에서 입수 가능) 또는 유사한 툴을 사용하여 수행한다. 적절한 어닐링된 이중가닥 형성하기 위한 이중이랑체화의 ΔG 및 2개의 동일한 올리고뉴클레오타이드의 동종이랑체화의 ΔG 를 비교하면, 화학적 접합의 실험 결과를 예측할 수 있다. 이중이랑체화의 ΔG 가 동종이랑체화의 ΔG 보다 작으면, 결합 효율이 높을 것으로 예측된다. 이 예측 방법은 실시예 XX에서 추가로 설명된다.

[0351] 가이드 분자 변형

[0352] 가이드 분자의 활성, 안정성 또는 다른 특성은 특정 변형의 혼입을 통해 변이될 수 있다. 하나의 예로서, 일시적으로 발현되거나 전달된 핵산은, 예를 들어 세포 뉴클레아제에 의해 쉽게 분해될 수 있다. 따라서, 본 명세서에 기재된 가이드 분자는 뉴클레아제에 대한 안정성을 도입하는 하나 이상의 변형된 뉴클레오사이드 또는 뉴클레오타이드를 포함할 수 있다. 이론에 결부시키고자 하는 것은 아니지만, 본 명세서에 기재된 변형된 특정 가이드 분자는 또한, 세포 내로 도입되는 경우, 감소된 선천성 면역 반응을 나타낼 수 있는 것으로 여겨진다. 당업자는 외인성 핵산, 특히 바이러스 또는 박테리아 유래의 외인성 핵산에 반응하여, 세포, 예를 들어 포유동물 세포에서 일반적으로 관찰되는 특정 세포 반응을 인식할 것이다. 사이토카인 발현 및 방출 및 세포 사멸 유도를 포함할 수 있는 이러한 반응은 본 명세서에 제시된 변형에 의해 감소되거나 완전히 제거될 수 있다.

[0353] 이 섹션에서 논의된 특정 예시적 변형은 제한 없이 5' 말단 또는 그 주위(예를 들어, 5' 말단의 1개 내지 10개, 1개 내지 5개, 1개 내지 3개 또는 1개 내지 2개의 뉴클레오타이드 내) 및/또는 3' 말단 또는 그 주위(3' 말단의 1개 내지 10개, 1개 내지 5개, 1개 내지 3개 또는 1개 내지 2개의 뉴클레오타이드 내)를 포함하여, 가이드 분자 서열 내의 임의의 위치에 포함될 수 있다. 일부 경우, 변형은 기능성 모티프, 예컨대 Cas9 가이드 분자의 반복-반복 방지 이중가닥, Cas9 또는 Cpf1 가이드 분자의 줄기 루프 구조, 및/또는 가이드 분자의 표적화 도메인 내에 위치한다.

[0354] 일 예로서, 가이드 분자의 5' 말단은 다음에 제시된 바와 같은 진핵세포 mRNA 캡 구조 또는 캡 유사체(예를 들어, $G(5')ppp(5')G$ 캡 유사체, $m7G(5')ppp(5')G$ 캡 유사체, 또는 $3'-O-Me-m7G(5')ppp(5')G$ 역회전 방지 캡 유사체(ARCA))를 포함할 수 있다:



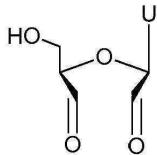
[0355]

[0356] 캡 또는 캡 유사체는 가이드 분자의 화학적 또는 효소적 합성 중에 포함될 수 있다.

[0357] 유사한 맥락에서, 가이드 분자의 5' 말단은 5' 트리포스페이트기가 부재할 수 있다. 예를 들어, 시험관내 전사된 가이드 분자를 포스파타제로 처리하여, 5' 트리포스페이트기를 제거할 수 있다(예를 들면, 송아지 장내 알칼리성 포스파타제 사용).

[0358] 또 다른 통상적인 변형은 가이드 분자의 3' 말단에 폴리A 관으로 지칭되는 다수의(예를 들어, 1개 내지 10개, 10개 내지 20개, 또는 25개 내지 200개의) 아데닌(A) 잔기의 첨가를 포함한다. 폴리아데노신 중합효소(예를 들어, E. coli) 폴리(A)중합효소를 사용한 화학적 또는 효소적 합성 동안 폴리A 관을 가이드 분자에 첨가할 수 있다.

[0359] 가이드 RNA를 3' 말단 U 리보스에서 변형시킬 수 있다. 예를 들어, U 리보스의 2개의 말단 하이드록실기를 알데하이드기로 산화시키고, 동시에 리보스 고리를 개방시켜, 다음에 제시된 바와 같은 변형된 뉴클레오사이드를 제공할 수 있고:

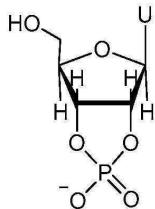


[0360]

"U"는 변형되지 않았거나 변형된 우리딘일 수 있다.

[0362]

3' 말단 U 리보스는 다음에 제시된 바와 같이 2'3' 사이클릭 포스페이트에 의해 변형될 수 있고:



[0363]

"U"는 변형되지 않았거나 변형된 우리딘일 수 있다.

[0365]

가이드 RNA는, 예를 들어 본 명세서에 기재된 하나 이상의 변형된 뉴클레오타이드를 혼입시킴으로써, 분해에 대해 안정화될 수 있는 3' 뉴클레오타이드를 포함할 수 있다. 특정 구현예에서, 우리딘은 변형된 우리딘, 예를 들어 5-(2-아미노)프로필 우리딘, 및 5-브로모 우리딘, 또는 본 명세서에 기재된 변형된 우리딘 중 어느 하나에 의해 대체될 수 있으며; 아데노신 및 구아노신은 변형된 아데노신 및 구아노신, 예를 들어 8 위치에서 변형된 것, 예를 들어 8-브로모 구아노신, 또는 본 명세서에 기재된 변형된 아데노신 또는 구아노신 중 어느 하나에 의해 대체될 수 있다.

[0366]

특정 구현예에서, 당-변형된 리보뉴클레오타이드가 가이드 분자에 혼입될 수 있으며, 예를 들어 2' OH기는 H, -OR, -R(R은, 예를 들어 알킬, 사이클로알킬, 아릴, 아르알킬, 헤테로아릴 또는 당일 수 있음), 할로, -SH, -SR(R은, 예를 들어 알킬, 사이클로알킬, 아릴, 아르알킬, 헤테로아릴 또는 당일 수 있음), 아미노(아미노는, 예를 들어 NH₂; 알킬아미노, 디알킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 디헤테로아릴아미노, 또는 아미노산일 수 있음); 또는 시아노(-CN)로부터 선택되는 군에 의해 대체된다. 특정 구현예에서, 포스페이트 백본은 본 명세서에 기재된 바와 같이, 예를 들어 포스포로티오에이트(PhTx) 기에 의해 변형될 수 있다. 특정 구현예에서, 가이드 분자의 하나 이상의 뉴클레오타이드는 각각 독립적으로, 예를 들어 2'-F 또는 2'-O-메틸, 아데노신(A), 2'-F 또는 2'-O-메틸, 시티딘(C), 2'-F 또는 2'-O-메틸, 우리딘(U), 2'-F 또는 2'-O-메틸, 티미딘(T), 2'-F 또는 2'-O-메틸, 구아노신(G), 2'-O-메톡시에틸-5-메틸우리딘(Teo), 2'-O-메톡시에틸아데노신(Aeo), 2'-O-메톡시에틸-5-메틸시티딘(m5Ceo), 및 이들의 임의의 조합을 포함하여, 2'-당 변형, 예컨대, 2'-O-메틸, 2'-O-메톡시에틸, 또는 2'-플루오로 변형을 포함하지만, 이에 제한되지 않는 변형되거나 변형되지 않은 뉴클레오타이드일 수 있다.

[0367]

가이드 RNA는 또한 2' OH-기가, 예를 들어 C1-6 알킬렌 또는 C1-6 헤테로알킬렌 가교에 의해, 동일한 리보스 당의 4' 탄소에 연결될 수 있는 "잠진" 핵산(LNA)을 포함할 수 있다. 임의의 적절한 모이어티를 사용하여, 이러한 가교를 제공할 수 있으며, 제한 없이 메틸렌, 프로필렌, 에테르, 또는 아미노 가교; O-아미노(아미노는, 예를 들어, NH₂; 알킬아미노, 디알킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 또는 디헤테로아릴아미노, 에틸렌디아민, 또는 폴리아미노일 수 있음) 및 아미노알콕시 또는 O(CH₂)_n-아미노(아미노는, 예를 들어 NH₂; 알킬아미노, 디알킬아미노, 헤테로사이클릴, 아릴아미노, 디아릴아미노, 헤테로아릴아미노, 또는

디헵테로아릴아미노, 에틸렌디아민, 또는 폴리아미노일 수 있음)를 포함한다.

[0368] 특정 구현예에서, 가이드 분자는 다중사이클릭인 변형된 뉴클레오타이드(예를 들어, 트리사이클로; 및 "잡기지 않은" 형태, 예컨대 글리콜 핵산(GNA)(예를 들어, R-GNA 또는 S-GNA, 리보스는 포스포디에스테르 결합에 부착된 글리콜 단위에 의해 대체됨), 또는 트레오스 핵산(TNA, 리보스는 α -L-트레오푸라노실-(3'→2')에 의해 대체됨)을 포함할 수 있다.

[0369] 일반적으로, 가이드 분자는 산소를 포함하는 5원 고리인 당 기 리보스를 포함한다. 예시적 변형된 가이드 분자는 제한 없이, (예를 들어, 황(S), 셀레늄(Se), 또는 알킬렌, 예컨대, 예를 들어, 메틸렌 또는 에틸렌에 의한) 리보스 내의 산소의 대체; 이중 결합의 첨가(예를 들어, 사이클로펜테닐 또는 사이클로헥세닐에 의한 리보스의 대체); 리보스의 고리 축소(예를 들어, 사이클로부탄 또는 옥세탄의 4원 고리의 형성); 리보스의 고리 확장(예를 들어, 또한 포스포르아미네이트 백본을 갖는 안하이드로헥시톨, 알트리톨, 만니톨, 사이클로헥사닐, 사이클로헥세닐 및 모르폴리노와 같은 추가적 탄소 또는 헤테로원자를 갖는, 예를 들어 6원 또는 7원 고리의 형성)을 포함할 수 있다. 당 유사체 변이의 대부분이 2' 위치에 국한되지만, 4' 위치를 포함하여 다른 부위가 변형될 수 있다. 특정 구현예에서, 가이드 분자는 4'-S, 4'-Se 또는 4'-C-아미노메틸-2'-O-Me 변형을 포함한다.

[0370] 특정 구현예에서, 데아자 뉴클레오타이드, 예를 들어 7-데아자-아데노신을 가이드 분자에 혼입시킬 수 있다. 특정 구현예에서, O- 및 N-알킬화된 뉴클레오타이드, 예를 들어 N6-메틸 아데노신을 가이드 분자에 혼입시킬 수 있다. 특정 구현예에서, 가이드 분자의 뉴클레오타이드 중 하나 이상 또는 전부가 데옥시뉴클레오타이드이다.

[0371] 가이드 분자의 뉴클레오타이드는 또한 포스포디에스테르 연결에서 변형될 수 있다. 이러한 변형은 포스포노아세테이트, 포스포로티오에이트, 티오포스포노아세테이트, 또는 포스포로아미네이트 연결을 포함할 수 있다. 일부 구현예에서, 뉴클레오타이드는 포스포로티오에이트 연결을 통해 그의 인접 뉴클레오타이드에 연결될 수 있다. 추가로, 포스포디에스테르 연결에 대한 변형은 뉴클레오타이드에 대한 단일 변형일 수 있거나 상기에 기재된 다른 뉴클레오타이드 변형과 조합될 수 있다. 예를 들어, 변형된 포스포디에스테르 연결은 뉴클레오타이드의 당기에 대한 변형과 조합될 수 있다. 일부 구현예에서, 5' 또는 3' 뉴클레오타이드는 포스포로티오에이트 연결을 통해 그의 인접 뉴클레오타이드(들)에 연결된 2'-OMe 변형된 리보뉴클레오타이드 잔기를 포함한다.

RNA-가이드 뉴클레아제

[0373] 본 개시내용에 따른 RNA-가이드 뉴클레아제는 자연 발생 클래스 2 CRISPR 뉴클레아제, 예컨대, Cas9, 및 Cpf1뿐만 아니라 그로부터 유래되거나 얻어진 다른 뉴클레아제를 포함하지만, 이에 제한되지 않는다. 기능적 의미에서, RNA-가이드 뉴클레아제는 (a) 가이드 분자(예를 들어, gRNA)와 상호작용(예를 들어, 복합체화)하고; (b) 가이드 분자(예를 들어, gRNA)와 함께, (i) 가이드 분자(예를 들어, gRNA)의 표적화 도메인에 상보성인 서열 및 선택적으로 (ii) 하기에서 더욱 상세하게 기재되는 "프로토스페이서 인접 모티프" 또는 "PAM"으로 지칭되는 추가적 서열을 포함하는 DNA의 표적 영역과 결합하고, 선택적으로 절단하거나 변형하는 뉴클레아제로서 정의된다. 다음의 예에 예시되는 바와 같이, RNA-가이드 뉴클레아제는 광범위한 의미에서, 동일한 PAM 특이성 또는 절단 활성을 공유하는 개개의 RNA-가이드 뉴클레아제 간에 변화가 존재할 수는 있지만, 그의 PAM 특이성 및 절단 활성을 의해 정의될 수 있다. 당업자는 특정 PAM 특이성 및/또는 절단 활성을 갖는 임의의 적절한 RNA-가이드 뉴클레아제를 사용하여, 구현될 수 있는 시스템, 방법 및 조성물에 관한 것임을 이해할 것이다. 이러한 이유로, 달리 명시되지 않는 한, 용어 RNA-가이드 뉴클레아제는 일반적인 용어로서 이해되어야 하며, RNA-가이드 뉴클레아제의 임의의 특정 유형(예를 들어, Cas9 대 Cpf1), 종(예를 들어, S. 피오카네스 대 S. 아우레우스) 또는 변화(예를 들어, 전장의 것 대 절두된 것 또는 분할된 것; 자연 발생 PAM 특이성 대 조작된 PAM 특이성 등)에 제한되지 않아야 한다.

[0374] PAM 서열은 그 명칭이 가이드 분자 표적화 도메인(또는 "스페이서")에 상보성인 "프로토스페이서" 서열에 대한 그의 순차적 관계로부터 도출된다. 프로토스페이서 서열과 함께, PAM 서열은 특이적 RNA-가이드 뉴클레아제/가이드 분자 조합에 대한 표적 영역 또는 서열을 정의한다.

[0375] 다양한 RNA-가이드 뉴클레아제는 PAM과 프로토스페이서 사이에 상이한 순차적 관계를 요구할 수 있다. 일반적으로, Cas9는 가이드 분자에 대비하여, 예상된 바와 같이 프로토스페이서의 3'인 PAM 서열을 인식한다.

[0376] Cpf1은, 다른 한편으로, 가이드 분자와 비교하여 관찰되는 바와 같이 일반적으로 프로토스페이서의 5'인 PAM 서열을 인식한다.

[0377] PAM 및 프로토스페이서의 특이적인 순차적 배향을 인식하는 것 이외에, RNA-가이드 뉴클레아제는 또한 특이적 PAM 서열을 인식할 수 있다. S. 아우레우스 Cas9는, 예를 들어, NNGRRT 또는 NNGRRV의 PAM 서열을 인식하며, N

잔기는 가이드 분자 표적화 도메인에 의해 인식되는 영역의 3'의 바로 옆에 존재한다. *S. 피오게네스 Cas9*는 NGG PAM 서열을 인식한다. 그리고, *F. 노비시다(F. novicida) Cpf1*은 TTN PAM 서열을 인식한다. 다양한 RNA-가이드 뉴클레아제에 대한 PAM 서열이 확인되었으며, 신규한 PAM 서열을 확인하기 위한 전략이 문헌[Shmakov et al., 2015, Molecular Cell 60, 385-397, November 5, 2015]에 기재되어 있다. 조작된 RNA-가이드 뉴클레아제는 기준 분자(예를 들어, 조작된 RNA-가이드 뉴클레아제의 경우, 기준 분자는 RNA-가이드 뉴클레아제가 유래한 자연 발생 변이체, 또는 조작된 RNA-가이드 뉴클레아제에 대해 최대 아미노산 서열 상동성을 갖는 자연 발생 변이체일 수 있음)의 PAM 특이성과 상이한 PAM 특이성을 가질 수 있음에 유의해야 한다.

[0378] 그의 PAM 특이성 이외에, RNA-가이드 뉴클레아제는 그의 DNA 절단 활성에 의해 특성화될 수 있다: 자연 발생 RNA-가이드 뉴클레아제는 통상적으로 표적 핵산 내에 DSB를 형성하지만, SSB(상기에서 논의됨)만을 발생시키거나(본 명세서에 참조로 포함되는 문헌[Ran & Hsu, et al., Cell 154(6), 1380-1389, September 12, 2013, (Ran)]), 전혀 커팅하지 않는 조작된 변이체를 생성하였다.

Cas9

[0380] 단분자 가이드 RNA 및 표적 DNA와의 복합체로서 *S. 피오게네스 Cas9* (Jinek 2014), 및 *S. 아우레우스 Cas9*의 결정 구조가 결정되었다(문헌[Nishimasu 2014; Anders 2014; and Nishimasu 2015]).

[0381] 자연 발생 Cas9 단백질은 각각 특정 구조 및/또는 기능성 도메인을 포함하는 2개의 로브, 즉 인식(REC) 로브 및 뉴클레아제(NUC) 로브를 포함한다. REC 로브는 아르기닌이 다량 존재하는 가교 나선(BH) 도메인 및 적어도 하나의 REC 도메인(예를 들어, REC1 도메인 및, 선택적으로, REC2 도메인)을 포함한다. REC 로브는 공지된 다른 단백질과 구조적 유사성을 공유하지 않으며, 이는 고유한 기능적 도메인임을 시사한다. 임의의 이론에 결부시키고자 하는 것은 아니지만, 돌연변이 분석은 BH 및 REC 도메인에 특이적인 기능적 역할을 제안한다: BH 도메인은 가이드 분자:DNA 인식에 관여하는 것으로 나타나고, REC 도메인은 가이드 분자의 반복:반복 방지 이중가닥과 상호작용하고, Cas9/가이드 분자 복합체의 형성을 매개하는 것으로 여겨진다.

[0382] NUC 로브는 RuvC 도메인, HNH 도메인, 및 PAM-상호작용(PI) 도메인을 포함한다. RuvC 도메인은 레트로바이러스 인테그라제 슈퍼페밀리원(retroviral integrase superfamily member)과 구조적 유사성을 공유하며, 표적 핵산의 비-상보성(즉, 하부) 가닥을 절단한다. 이는 2개 이상의 분할 RuvC 모티프(예컨대, *S. 피오게네스* 및 *S. 아우레우스* 내의 RuvC I, RuvCII, 및 RuvCIII)로부터 형성될 수 있다. 한편, HNH 도메인은 HNN 엔도뉴클레아제 모티프와 구조적으로 유사하며, 표적 핵산의 상보성(즉, 상부) 가닥을 절단한다. 그 명칭이 시사하는 바와 같이, PI 도메인은 PAM 특이성에 관여한다.

[0383] Cas9의 특정 기능이 상기에 제시된 특정 도메인에 연결되지만(반드시 필연적으로 완전히 결정되는 것은 아님), 이를 및 다른 기능은 다른 Cas9 도메인 또는 로브 상의 다수의 도메인에 의해 매개되거나 그의 영향을 받을 수 있다. 예를 들어, 문헌[Nishimasu 2014]에 기재된 바와 같이, *S. 피오게네스 Cas9*에서 가이드 분자의 반복:반복 방지 이중가닥은 REC와 NUC 로브 사이의 주름에 존재하고, 이중가닥의 뉴클레오타이드는 BH, PI, 및 REC 도메인의 아미노산과 상호작용한다. 제1 줄기 루프 구조 내의 일부 뉴클레오타이드는 또한 다수의 도메인(PI, BH 및 REC1) 내의 아미노산과 상호작용하며, 일부 뉴클레오타이드는 제2 및 제3 줄기 루프(RuvC 및 PI 도메인) 내의 아미노산과 상호작용한다.

Cpf1

[0385] TTTN PAM 서열을 포함하는 이중 가닥(ds) DNA 표적 및 crRNA와의 복합체의 아시다미노코커스 종 (*Acidaminococcus sp.*) Cpf1의 결정 구조가 본 명세서에 참조로 포함되는 문헌[Yamano et al. (Cell. 2016, May 5; 165(4): 949-962 (Yamano))에 의해, 밝혀졌다. Cas9와 같이, Cpf1은 REC(인식) 로브, 및 NUC(뉴클레아제) 로브의 2개의 로브를 갖는다. REC 로브는 임의의 공지된 단백질 구조와 유사성이 없는 REC1 및 REC2 도메인을 포함한다. 한편, NUC 로브는 3개의 RuvC 도메인(RuvC-I, -II 및 -III) 및 BH 도메인을 포함한다. 그러나, Cas9와 달리, Cpf1 REC 로브는 HNH 도메인을 갖지 않으며, 또한 공지된 단백질 구조와 유사성이 없는 다른 도메인을 포함한다: 구조적으로 고유한 PI 도메인, 3개의 Wedge(WED) 도메인(WED-I, -II 및 -III), 및 뉴클레아제(Nuc) 도메인.

[0386] Cas9 및 Cpf1이 구조 및 기능에서 유사성을 공유하지만, 특정 Cpf1 활성이 임의의 Cas9 도메인과 유사하지 않은 구조적 도메인에 의해 매개된다는 것이 이해되어야 한다. 예를 들어, 표적 DNA의 상보성 가닥의 절단은 Cas9의 HNH 도메인과 공간적으로 순차적으로 상이한 Nuc 도메인에 의해 매개되는 것으로 나타난다. 추가적으로, Cpf1 가이드 분자의 비표적화 부분(핸들(handle))은 반복에 의해 형성되는 줄기 루프 구조가 아닌 유사매듭 구조를

채용한다: Cas9 가이드 분자에서의 반복 방지 이중가닥.

[0387] RNA-가이드 뉴클레아제의 변형

[0388] 전술된 RNA-가이드 뉴클레아제는 다양한 적용 예에 유용할 수 있는 활성 및 특성을 가지지만, 당업자는 특정 예에서 절단 활성, PAM 특이성, 또는 다른 구조적 또는 기능적 특징을 변이시키기 위해, RNA-가이드 뉴클레아제가 또한 변형될 수 있음을 이해할 것이다.

[0389] 우선, 절단 활성을 변이시키는 변형으로 화제를 전환하자면, NUC 로브 내의 도메인의 활성을 감소시키거나 제거하는 돌연변이는 전술되어 있다. RuvC 도메인, Cas9 HNH 도메인 또는 Cpf1 Nuc 도메인에 이루어질 수 있는 예시적인 돌연변이가 문헌[Ran 및 Yamano]뿐만 아니라, 문헌[Cotta-Ramusino]에 기재되어 있다. 일반적으로 2개의 뉴클레아제 도메인 중 하나에서의 활성을 감소시키거나 제거하는 돌연변이는 낙카제 활성을 갖는 RNA-가이드 뉴클레아제를 야기하지만, 낙카제 활성의 유형은 어느 도메인이 불활성화되는지에 따라 달라진다는 점에 유의해야 한다. 일 예로서, Cas9의 RuvC 도메인의 불활성화는 상보성 또는 상부 가닥을 절단하는 낙카제를 야기할 것이다.

[0390] 다른 한편으로, Cas9 HNH 도메인의 불활성화는 하부 또는 비-상보성 가닥을 절단하는 낙카제를 야기한다.

[0391] 자연 발생 Cas9 기준 분자에 대한 PAM 특이성의 변형은 S. 피오케네스(문헌[Kleininstiver et al., Nature. 2015 Jul 23;523(7561):481-5 (Kleininstiver I)]) 및 S. 아우레우스(문헌[Kleininstiver et al., Nat Biotechnol. 2015 Dec; 33(12): 1293-1298 (Kleininstiver II)]) 돌 모두에 대해 Kleininstiver 등에 의해 기재된 바 있다. Kleininstiver 등은 또한 Cas9의 표적화 정확도를 개선하는 변형을 기재하였다(문헌[Nature, 2016 January 28; 529, 490-495 (Kleininstiver III)]). 이를 참조문헌 각각은 본 명세서에 참조로 포함된다.

[0392] 참조로 포함되는 문헌[Zetsche et al. (Nat Biotechnol. 2015 Feb;33(2):139-42 (Zetsche II))], 및 참조로 포함되는 문헌[Fine et al. (Sci Rep. 2015 Jul 1;5:10777 (Fine))]에 기재된 바와 같이, RNA-가이드 뉴클레아제는 2개 이상의 부분으로 분할되었다.

[0393] RNA-가이드 뉴클레아제는 특정 구현예에서, 예를 들어 뉴클레아제의 크기를 감소시키는 하나 이상의 결실을 통해 크기-최적화되거나 절두되지만, 가이드 분자 결합, 표적 및 PAM 인식, 및 절단 활성을 여전히 유지한다. 특정 구현예에서, RNA 가이드 뉴클레아제는 선택적으로 링커에 의해, 또 다른 폴리펩타이드, 뉴클레오타이드, 또는 다른 구조에 공유적으로 또는 비공유적으로 결합된다. 예시적 결합 뉴클레아제 및 링커가 본 명세서에서 모든 목적을 위해 참조로 포함되는 문헌[Guilinger et al., Nature Biotechnology 32, 577-582 (2014)]에 기재되어 있다.

[0394] RNA-가이드 뉴클레아제는 또한 선택적으로 핵내로의 RNA-가이드 뉴클레아제 단백질의 이동을 촉진하는 핵 국소화 신호와 같은 태그를 포함하지만, 이에 제한되지 않는다. 특정 구현예에서, RNA-가이드 뉴클레아제는 C- 및/ 또는 N-말단 핵 국소화 신호를 혼입시킬 수 있다. 핵 국소화 서열은 당업계에 공지되어 있으며, 문헌[Maeder] 등에 기재되어 있다.

[0395] 전술한 변형 목록은 본질적으로 예시적인 것으로 의도되며, 당업자는 본 개시내용을 고려하여 특정 적용 예에서 다른 변형이 가능하거나 요망될 수 있음을 이해할 것이다. 간략히 말하면, 따라서 본 개시내용의 예시적인 시스템, 방법 및 조성물은 특정 RNA-가이드 뉴클레아제와 관련하여 제시되지만 사용된 RNA-가이드 뉴클레아제가 그의 작동 원리를 변이시키지 않는 방식으로 변형될 수 있음을 이해해야 한다. 이러한 변형은 본 개시내용의 범위내에 있다.

[0396] RNA-가이드 뉴클레아제를 인코딩하는 핵산

[0397] RNA-가이드 뉴클레아제, 예를 들어 Cas9, Cpf1 또는 이들의 기능적 단편을 인코딩하는 핵산이 본 명세서에서 제공된다. RNA-가이드 뉴클레아제를 인코딩하는 예시적 핵산은 종래에 기재된 바 있다(예를 들어, 문헌[Cong 2013; Wang 2013; Mali 2013; Jinek 2012] 참조).

[0398] 일부 경우, RNA-가이드 뉴클레아제를 인코딩하는 핵산은 합성 핵산 서열일 수 있다. 예를 들어, 합성 핵산 분자는 화학적으로 변형될 수 있다. 특정 구현예에서, RNA-가이드 뉴클레아제를 인코딩하는 mRNA는 다음 특성 중 하나 이상(예를 들어, 전부)을 가질 것이다: 이는 캡핑되고, 폴라아데닐화되고, 5-메틸시티딘 및/또는 슈도우리딘에 의해 치환될 수 있다.

[0399] 합성 핵산 서열은 또한 코돈 최적화될 수 있으며, 예를 들어 적어도 하나의 보편적이지 않은 코돈 또는 덜 보편

적인 코돈이 보편적인 코돈에 의해 대체되었다. 예를 들어, 합성 핵산은 최적화된 메신저 mRNA, 예를 들어 본 명세서에 기재된, 예를 들어 포유류 발현 시스템에서의 발현을 위해 최적화된 메신저 mRNA의 합성을 유도할 수 있다. 코돈 최적화된 Cas9 코딩 서열의 예가 Cotta-Ramusino에 제시되어 있다.

[0400] 추가적으로 또는 대안적으로, RNA-가이드 뉴클레아제를 인코딩하는 핵산은 핵 국소화 서열(NLS)을 포함할 수 있다. 핵 국소화 서열은 당업계에 공지되어 있다.

후보 분자의 기능적 분석

[0402] 후보 RNA-가이드 뉴클레아제, 가이드 분자, 및 그의 복합체는 당업계에 공지된 표준 방법에 의해 평가될 수 있다. 예를 들어, Cotta-Ramusino를 참조한다. RNP 복합체의 안정성을 하기에 기재되는 바와 같이, 시차 주사 형광 측정기에 의해 평가할 수 있다.

시차 주사 형광 측정기(DSF)

[0404] 가이드 분자 및 RNA-가이드 뉴클레아제를 포함하는 리보핵단백질(RNP) 복합체의 열 안정성을 DSF를 통해 측정할 수 있다. DSF 기법에 의해 단백질의 열 안정성을 측정하는데, 이는 결합 RNA 분자, 예를 들어 가이드 분자의 첨가와 같은 유리한 조건하에서 증가할 수 있다.

[0405] 제한 없이 (a) 상이한 조건(예를 들어, 가이드 분자: RNA-가이드 뉴클레아제 단백질의 상이한 화학량론적 비, 상이한 완충액 용액 등)을 시험하여, RNP 형성을 위한 최적 조건을 확인하는 단계; 및 (b) RNA-가이드 뉴클레아제 및/또는 가이드 분자의 변형(예를 들어, 서열의 화학적 변형, 변이 등)을 시험하여, RNP 형성 또는 안정성을 개선하는 변형을 확인하는 단계를 포함하여, DSF 분석을 임의의 적합한 프로토콜에 따라 수행할 수 있으며, 임의의 적절한 설정하에 수행할 수 있다. DSF 분석의 한 가지 판독 결과는 RNP 복합체의 용융 온도의 변화이고; 상대적으로 높은 변화는 RNP 복합체가 더 낮은 변화를 특징으로 하는 기준 RNP 복합체에 비해 더 안정함을 나타낸다(따라서 더 큰 활성 또는 더 유리한 생성 동력학, 분해 동력학 또는 또 다른 기능적 특성을 가질 수 있음). DSF 분석을 스크리닝 도구로 사용하는 경우, 생성물이 임계값 이상의 용융 온도 변화를 갖는 하나 이상의 RNP가 되도록, 임계값 용융 온도 변화를 지정할 수 있다. 예를 들어, 임계값은 5°C 내지 10°C(예를 들어 5°, 6°, 7°, 8°, 9°, 10°) 이상일 수 있으며, 생성물은 임계값 이상의 용융 온도 변화를 특징으로 하는 하나 이상의 RNP일 수 있다.

[0406] DSF 분석 조건의 2개의 비 제한적인 예는 다음에 제시된다:

[0407] RNP 복합체를 형성하기 위한 최적의 용액을 결정하기 위해, 물+10x SYPRO Orange® (Life Technologies cat#S-6650) 중 고정 농도(예를 들어 2 μM)의 Cas9를 384 웰 플레이트에 분주한다. 이어서, 다양한 pH의 용액 중에 희석시킨 등몰량의 가이드 분자 및 염을 첨가한다. 실온에서 10' 동안 인큐베이션하고, 짧게 원심분리함으로써, 거품을 제거한 후에, Bio-Rad CFX Manager 소프트웨어가 구비된 Bio-Rad CFX384™ 실시간 시스템 C1000 Touch™ 열 순환기를 사용하여, 매 10초마다 온도를 1°C씩 증가시키면서, 20°C에서 90°C의 구배로 구동시킨다.

[0408] 제2 분석은 분석 1로부터의 최적의 완충액 중에 다양한 농도의 가이드 분자와 고정 농도(예를 들어, 2 μM)의 Cas9를 혼합하는 단계 및 이를 384 웰 플레이트에서 인큐베이션(예를 들어, 실온에서 10' 동안)하는 단계로 이루어진다. 동일한 용적의 최적의 완충액 + 10x SYPRO Orange®(Life Technologies cat#S-6650)를 첨가하고, Microseal® B 접착제(MSB-1001)를 사용하여 플레이트를 밀봉한다. 짧게 원심분리함으로써, 거품을 제거한 후에, Bio-Rad CFX Manager 소프트웨어가 구비된 Bio-Rad CFX384™ 실시간 시스템 C1000 Touch™ 열 순환기를 사용하여, 매 10초마다 온도를 1°C씩 증가시키면서, 20°C에서 90°C의 구배로 구동시킨다.

게놈 편집 전략

[0410] 상기에 기재된 게놈 편집 시스템을 본 개시내용의 다양한 구현예에서 사용하여, 세포 내 또는 그로부터 획득된 DNA의 표적화된 영역의 편집을 수행한다(즉, 변이시킴). 특정 편집을 수행하기 위한 다양한 전략이 본 명세서에 기재되어 있으며, 이들 전략은 일반적으로 적절한 수선 결과, 개별 편집(예를 들어, SSB 또는 DSB)의 수 및 위치, 이러한 편집의 표적 부위의 관점에서 기재된다.

[0411] SSB 또는 DSB의 형성을 포함하는 게놈 편집 전략은 (a) 표적화된 영역의 전부 또는 일부의 결실; (b) 표적화된 영역의 전부 또는 일부의 삽입 또는 대체; 또는 (c) 표적화된 영역의 전부 또는 일부의 단절을 포함하는 수선 결과를 특징으로 한다. 이러한 분류는 임의의 특정 이론 또는 모델로 제한하거나 결부시키려는 의도는 아니며, 단지 표현의 간결성을 위해 제공된다. 당업자는 열거된 결과가 상호 배타적이지 않으며, 일부 수선으로 인해 다른 결과가 발생할 수 있음을 이해할 것이다. 특정 편집 전략 또는 방법에 대한 기재는 달리 명시하지 않는 한,

특정 수선 결과가 요구되는 것으로 이해되어서는 안된다.

[0412] 표적화된 영역의 대체는 일반적으로, 예를 들어 HDR 경로에 의해 매개되는 2가지 수선 결과인 유전자 보정 또는 유전자 변환을 통해 표적 영역 내의 기존 서열의 전부 또는 일부를 상동성 서열로 대체하는 단계를 포함한다. HDR은 아래에서 더욱 상세히 기재되는 바와 같이, 단일 가닥 또는 이중 가닥일 수 있는 공여체 주형의 사용에 의해 용이해진다. 단일 또는 이중 가닥 주형이 외인성일 수 있는 경우, 이는 유전자 보정을 촉진할 것이지만, 이들이 내인성(예를 들어, 세포 계놈 내의 상동성 서열)일 수 있는 경우, 유전자 전환을 촉진할 것이다. 예를 들어, 참조로 포함되는 문헌[Richardson et al. (Nature Biotechnology 34, 339-344 (2016), (Richardson)]에 기재된 바와 같이, 외인성 주형은 비대칭 오버행을 가질 수 있다(즉, DSB의 부위에 상보성인 주형 부분은 공여체 주형 내의 중간 부분이 아닌 3' 또는 5' 방향으로 오프셋될 수 있음). 주형이 단일 가닥인 경우, 이는 표적화된 영역의 상보성(상) 또는 비 상보성(하) 가닥에 상응할 수 있다.

[0413] 유전자 변환 및 유전자 보정은 Ran 및 Cotta-Ramusino에 기재된 바와 같이, 일부 경우, 표적화된 영역 또는 그 주위에 하나 이상의 닉의 형성에 의해 촉진된다. 일부 경우, 이중 닉카제 전략을 사용하여, 2개의 오프셋 SSB를 형성하고, 결과적으로 오버행(예를 들어, 5' 오버행)을 갖는 단일 DSB를 형성한다.

[0414] 표적화된 서열의 전부 또는 일부의 단절 및/또는 결실은 다양한 수선 결과에 의해 달성될 수 있다. 일 예로서, 후에 DSB가 수선될 때 절제되는 표적화된 영역에 측접하는 2개 이상의 DSB를 동시에 생성함으로써 서열을 결실 시킬 수 있으며, 이는 LCA10 돌연변이에 대한 Maeder에 기재된 바와 같다. 또 다른 예로서, 서열은 단일 가닥 오버행을 갖는 이중 가닥 파손을 형성시킨 후, 수선 전에 오버행을 엑소핵산분해 처리(exonucleolytic processing)함으로써, 생성된 결실에 의해 서열을 단절시킬 수 있다.

[0415] 표적 서열 단절의 하나의 특정 하위세트는 표적화된 서열 내의 삽입결실의 형성에 의해 매개되며, 수선 결과는 통상적으로 NHEJ 경로(AIt-NHEJ 포함)에 의해 매개된다. NHEJ는 삽입결실 돌연변이와의 연관 때문에 "오류가 발생하기 쉬운" 수선 경로로 지칭된다. 그러나, 일부 경우 DSB는 NHEJ에 의해 그 주위의 서열이 변이되지 않고 수선되고(이른바 "완벽한" 또는 "흔적없는" 수선); 이는 일반적으로 DSB의 두 말단이 완벽하게 결찰될 것을 요구 한다. 한편, 삽입결실은 유리된 DNA 말단을 효소 처리한 후, 이를 결찰시켜 유리된 말단 중 어느 하나 또는 둘 모두의 가닥 중 어느 하나 또는 둘 모두에 뉴클레오파이드를 첨가하고/하거나 제거함으로써 유발되는 것으로 고려된다.

[0416] 유리된 DSB 말단의 효소 처리는 본질적으로 확률적이기 때문에, 삽입결실 돌연변이는 분포에 따라 가변적으로 발생하는 경향이 있고, 특정 표적 부위, 사용된 세포 유형, 사용된 계놈 편집 전략 등을 포함하는 다양한 인자에 의해 영향을 받을 수 있다. 그러나, 삽입결실 형성에 대한 제한된 일반화의 도출이 가능하며: 단일 DSB의 수선에 의해 형성된 결실은 가장 일반적으로 1 bp 내지 50 bp 범위이지만, 100 bp 내지 200 bp를 초과할 수 있다. 단일 DSB의 수선에 의해 형성된 삽입은 더 짧아지는 경향이 있고, 종종 파손 부위 바로 옆 서열의 짧은 복제를 포함한다. 그러나 큰 삽입의 획득이 가능하며, 이러한 경우에, 삽입된 서열은 종종 세포 내에 존재하는 플라스미드 DNA 또는 계놈의 다른 영역으로 추적된다.

[0417] 삽입결실 돌연변이 및 삽입결실을 생성하도록 구성된 계놈 편집 시스템은, 예를 들어 특정 최종 서열의 생성이 요구되지 않는 경우 및/또는 구조이동 돌연변이가 허용되는 경우, 표적 서열을 단절시키는 데 유용하다. 이는 또한, 요망되는 특정 서열이 소정의 부위에서 SSB 또는 DSB의 수선으로부터 우선적으로 발생하는 경향이 있는 한, 특정 서열이 요망되는 상황에서 유용할 수 있다. 삽입결실 돌연변이는 또한 특정 계놈 편집 시스템 및 그 성분의 작용을 평가하거나 스크리닝하는 데 유용한 툴이다. 이러한 상황 및 다른 상황에서, 삽입결실은 (a) 계놈 편집 시스템과 접촉한 세포의 계놈에서의 상대 및 절대 빈도 및 (b) 편집되지 않은 서열에 대비한 수치 차이, 예를 들어 ± 1 , ± 2 , ± 3 등의 분포를 특징으로 할 수 있다. 하나의 예로서, 판독 결과 상황에서 다수의 가이드 문자를 스크리닝하여, 제어된 조건하에서 삽입결실 판독 값을 기반으로 하여 표적 부위에서 가장 효과적으로 절단을 유도하는 가이드 문자를 확인할 수 있다. 임계 빈도 이상으로 삽입결실을 생성하거나 삽입결실의 특정 분포를 생성하는 가이드를 추가 연구 및 개발을 위해 선택할 수 있다. 삽입결실 빈도 및 분포는 또한, 예를 들어 가이드 문자를 일정하게 유지하고 다른 특정 반응 조건 또는 전달 방법을 변경함으로써 상이한 계놈 편집 시스템 구현 또는 제형 및 전달 방법을 평가하기 위한 판독 값으로 유용할 수 있다.

다중복합 전략

[0419] 상기 논의된 예시적 전략은 단일 DSB에 의해 매개되는 수선 결과에 역점을 두었지만, 본 개시내용에 따른 계놈 편집 시스템은 또한 동일한 유전자 좌위 또는 상이한 유전자 좌위에서 2개 이상의 DSB를 생성하는 데 사용될 수

있다. 다수의 DSB 또는 SSB의 형성을 수반하는 편집 전략은, 예를 들어 Cotta-Ramusino에 기재되어 있다.

[0420] 공여체 주형 설계

[0421] 공여체 주형 설계는 문헌, 예를 들어 문헌[Cotta-Ramusino]에 상세히 기재되어 있다. 단일 가닥(ssODN) 또는 이 중 가닥(dsODN)일 수 있는 DNA 올리고머 공여체 주형(올리고데옥시뉴클레오타이드 또는 ODN)은 DSB의 HDR-기반 수선을 용이하게 하는 데 사용될 수 있으며, 표적 DNA 서열에 변이를 도입하거나 새로운 서열을 표적 서열에 삽입하거나 표적 서열을 모두 교체하는 데 특히 유용하다.

[0422] 단일 가닥 또는 이중 가닥의 공여체 주형은 일반적으로 절단될 표적 서열 내 또는 그 주위(예를 들어, 측접 또는 인접)의 DNA 영역에 상동성인 영역을 포함한다. 이를 상동성 영역은 본 명세서에서 "상동성 아암(homology arm)"으로 지칭되며, 아래에 개략적으로 예시되어 있다:

[0423] [5' 상동성 아암] – [대체 서열] – [3' 상동성 아암].

[0424] 상동성 아암은 임의의 적절한 길이(단지 하나의 상동성 아암이 사용되는 경우, 0개의 뉴클레오타이드 포함)를 가질 수 있고, 3' 및 5' 상동성 아암은 동일한 길이를 가질 수 있거나 길이가 상이할 수 있다. 적절한 상동성 아암 길이의 선택은 Alu 반복 또는 다른 고도의 공통 요소와 같은 특정 서열과의 상동성 또는 마이크로상동성(microhomology)을 방지하고자 하는 목적과 같은 다양한 인자에 의해 영향을 받을 수 있다. 예를 들어, 서열 반복 요소를 방지하기 위해 5' 상동성 아암을 단축시킬 수 있다. 다른 구현예에서, 3' 상동성 아암을 단축시켜, 서열 반복 요소를 방지할 수 있다. 일부 구현예에서, 특정 서열 반복 요소를 포함하는 것을 방지하기 위해, 5' 및 3' 상동성 아암 둘 모두를 단축시킬 수 있다. 추가로, 일부 상동성 아암 설계는 적절한 수선 결과의 빈도를 증가시키거나 편집 효율을 개선시킬 수 있다. 예를 들어 참조로 포함되는 문헌[Richardson et al. Nature Biotechnology 34, 339-344 (2016) (Richardson)]은 단일 가닥 공여체 주형의 3' 및 5' 상동성 아암의 상대적 비대칭이 수선율 및/또는 결과에 영향을 미친다는 것을 발견했다.

[0425] 공여체 주형의 대체 서열은 문헌[Cotta-Ramusino et al.]에 포함된 다른 부분에 기재되어 있다. 대체 서열은 임의의 적절한 길이(적절한 수선 결과가 결실인 경우, 0개의 뉴클레오타이드 포함)일 수 있고, 통상적으로 편집이 요구되는 세포 내의 자연 발생 서열과 비교하여, 1, 2, 3개 이상의 서열 변형을 포함한다. 하나의 공통 서열 변형은 치료가 요구되는 질병 또는 질환과 관련된 돌연변이를 수선하기 위한 자연 발생 서열의 변이를 포함한다. 또 다른 공통 서열 변형은, 대체 서열이 표적 부위에 혼입된 후, 표적 부위의 반복적인 절단을 감소시키거나 제거하기 위해, SSB 또는 DSB를 생성하는 데 사용되는 가이드 분자(들)의 표적화 도메인 또는 RNA 가이드 뉴클레아제의 PAM 서열에 상보성이거나 이를 코딩하는 하나 이상의 서열의 변이를 포함한다.

[0426] 선형 ssODN이 사용되는 경우, 이는 (i) 표적 핵산의 낙형성 가닥에 어닐링되고, (ii) 표적 핵산의 손상되지 않은 가닥에 어닐링되고, (iii) 표적 핵산의 플러스 가닥에 어닐링되고/되거나, (iv) 표적 핵산의 마이너스 가닥에 어닐링되도록 구성될 수 있다. ssODN은 임의의 적절한 길이, 예를 들어 약, 적어도, 또는 150개 내지 200개 이하의 뉴클레오타이드(예를 들어, 150, 160, 170, 180, 190, 또는 200개의 뉴클레오타이드)를 가질 수 있다.

[0427] 주형 핵산은 또한 바이러스 게놈 또는 원형 이중 가닥 DNA, 예를 들어 플라스미드와 같은 핵산 벡터일 수 있음에 유의해야 한다. 공여체 주형을 포함하는 핵산 벡터는 다른 코딩 또는 비 코딩 요소를 포함할 수 있다. 예를 들어, 주형 핵산은 특정 게놈 백본 요소(예를 들어, AAV 게놈의 경우에는 역위 말단 반복)를 포함하는 바이러스 게놈(예를 들어, AAV 또는 렌티바이러스 게놈)의 일부로 전달될 수 있으며, 선택적으로 가이드 분자 및/또는 RNA 가이드 뉴클레아제를 코딩하는 추가 서열을 포함한다. 특정 구현예에서, 공여체 주형은 하나 이상의 가이드 분자에 의해 인식되는 표적 부위에 인접하거나 측접할 수 있어, 동일한 가이드 분자를 사용하여 세포 DNA에 형성된 상응하는 SSB 또는 DSB의 수선에 관여할 수 있는 공여체 주형의 말단 중 하나 또는 둘 모두에 유리 DSB의 형성을 촉진한다. 공여체 주형으로서 사용하기에 적합한 예시적인 핵산 벡터는 문헌[Cotta-Ramusino]에 기재되어 있다.

[0428] 어떠한 형식이 사용되든지 간에, 부적절한 서열을 방지하도록 주형 핵산을 설계할 수 있다. 특정 구현예에서, 상동성 아암 중 하나 또는 둘 모두는 특정 서열 반복 요소, 예를 들어 Alu 반복, LINE 요소 등과의 중첩을 방지하기 위해 단축될 수 있다.

[0429] 표적 세포

[0430] 본 개시내용에 따른 게놈 편집 시스템은, 예를 들어 표적 핵산을 편집하거나 변이시키기 위해 세포를 조작하거나 변이시키는 데 사용될 수 있다. 조작은 다양한 구현예에서 생체내 또는 생체외에서 발생할 수 있다.

- [0431] 다양한 세포 유형을 본 개시내용의 구현예에 따라 조작하거나 변이시킬 수 있으며, 생체내 적용과 같은 일부 경우에는, 예를 들어 다수의 세포 유형에 본 개시내용에 따른 계놈 편집 시스템을 전달함으로써 다수의 세포 유형을 변이시키거나 조작한다. 그러나 다른 경우에는 조작이나 변이를 특정 세포 유형 또는 유형들로 제한하는 것이 바람직할 수 있다. 예를 들어, 일부 경우, 제한된 분화 가능성을 갖는 세포 또는 문현[Maeder]의 경우의 광수용체 세포와 같은 최종 분화 세포를 편집하는 것이 바람직할 수 있으며, 유전자형의 변형은 세포 표현형의 변화를 유발할 것으로 예상된다. 그러나 다른 경우에는 덜 분화된 다능성 또는 전 능성 줄기 또는 전구세포를 편집하는 것이 바람직할 수 있다. 예로서, 세포는 배아 줄기세포, 유도된 전능성 줄기세포(iPSC), 조혈모/전구세포(HSPC), 또는 일정한 응용 또는 적용과 관련된 세포 유형으로 분화하는 다른 줄기 또는 전구 세포 유형일 수 있다.
- [0432] 결과로서, 변이되거나 조작될 세포는 표적화될 세포 유형(들) 및/또는 적절한 편집 결과에 따라, 다양한 분열 세포 또는 비분열 세포이다.
- [0433] 세포가 생체외에서 조작되거나 변경되는 경우, 세포는 즉시 사용(예를 들어, 대상체에 투여)될 수 있거나 이는 추후 사용하기 위해 유지되거나 저장될 수 있다. 당업자는 당업계에 공지된 임의의 적합한 방법을 사용하여 세포를 배양 유지하거나 저장(예를 들어, 액체 질소 중 동결)할 수 있음을 이해할 것이다.
- [0434] 계놈 편집 시스템 구현: 전달, 제형화 및 투여 경로
- [0435] 상기에 논의한 바와 같이, 본 개시내용의 계놈 편집 시스템은 임의의 적합한 방식으로 구현될 수 있으며, 이는 제한 없이 RNA-가이드 뉴클레아제, 가이드 문자, 및 선택적 공여체 주형 핵산을 포함하는 이러한 시스템의 성분이 계놈 편집 시스템의 형질도입, 발현 또는 도입을 유발하고/하거나, 세포, 조직 또는 대상체에서 적절한 수선 결과를 야기하는 임의의 적합한 형태 또는 형태의 조합으로 전달되거나 제형화되거나 투여될 수 있다. 표 5 및 표 6은 계놈 편집 시스템 구현의 수 개의 비 제한적인 예를 제시한다. 그러나, 당업자는 이를 목록이 포괄적이지 않으며, 다른 구현예가 가능하다는 것을 이해할 것이다. 특히 표 5와 관련하여, 표는 단일 가이드 문자 및 선택적 공여체 주형을 포함하는 계놈 편집 시스템의 수 개의 예시적인 구현예를 열거한다. 그러나, 본 개시내용에 따른 계놈 편집 시스템은 다수의 가이드 문자, 다수의 RNA-가이드 뉴클레아제, 및 다른 성분, 예컨대 단백질을 흡입시킬 수 있으며, 다양한 구현예가 표에 예시된 원리를 기반으로 하여 당업자에게 명백할 것이다. 표에서 [N/A]는 계놈 편집 시스템에 표시된 성분이 포함되어 있지 않음을 나타낸다.

[0436]

[표 5]

개념 편집 시스템 성분			
RNA-가이드 뉴클레아제	가이드 분자	공여체 주형	설명
단백질	RNA	[N/A]	gRNA 분자와 복합체화된 RNA-가이드 뉴클레아제 단백질(RNP 복합체)
단백질	RNA	DNA	상기에 기재된 바와 같은 RNP 복합체 + 단일 가탁 또는 이중 가탁 공여체 주형.
단백질	DNA	[N/A]	RNA-가이드 뉴클레아제 단백질 + DNA로부터 전사된 gRNA.
단백질	DNA	DNA	RNA-가이드 뉴클레아제 단백질 + gRNA-인코딩 DNA 및 개별 DNA 공여체 주형.
단백질	DNA		RNA-가이드 뉴클레아제 단백질 및 gRNA 및 공여체 주형 둘 모두를 인코딩하는 단일 DNA.
DNA			RNA-가이드 뉴클레아제, gRNA 및 공여체 주형을 인코딩하는 DNA 또는 DNA 벡터.
DNA	DNA	[N/A]	각각 RNA-가이드 뉴클레아제 및 gRNA를 인코딩하는 2개의 개별 DNA, 또는 2개의 개별 DNA 벡터.
DNA	DNA	DNA	각각 RNA-가이드 뉴클레아제, gRNA 및 공여체 주형을 인코딩하는 3개의 개별 DNA, 또는 3개의 개별 DNA 벡터
DNA		[N/A]	RNA-가이드 뉴클레아제 및 gRNA를 인코딩하는 DNA 또는 DNA 벡터
DNA		DNA	RNA-가이드 뉴클레아제 및 gRNA를 인코딩하는 제1 DNA 또는 DNA 벡터, 및 공여체 주형을 인코딩하는 제2 DNA 또는 DNA 벡터.
DNA	DNA		RNA-가이드 뉴클레아제 및 제2 DNA를 인코딩하는 제1 DNA 또는 DNA 벡터 또는 gRNA 및 공여체 주형을 인코딩하는 DNA 벡터.

[0437]

DNA	DNA	RNA-가이드 뉴클레아제 및 공여체 주형을 인코딩하는 제1 DNA 또는 DNA 벡터 및 gRNA를 인코딩하는 제2 DNA 또는 DNA 벡터
DNA	RNA	RNA-가이드 뉴클레아제 및 공여체 주형, 및 gRNA를 인코딩하는 DNA 또는 DNA 벡터
RNA	[N/A]	RNA-가이드 뉴클레아제를 인코딩하고, gRNA를 포함하는 RNA 또는 RNA 벡터
RNA	DNA	RNA-가이드 뉴클레아제를 인코딩하고, gRNA를 포함하는 RNA 또는 RNA 벡터 및 공여체 주형을 인코딩하는 DNA 또는 DNA 벡터.

[0438]

[0439] 표 6에는, 예를 들어 본 명세서에 기재된 바와 같은 게놈 편집 시스템의 성분에 대한 다양한 전달 방법이 정리되어 있다. 다시 말하면, 이 목록은 제한하려는 것이 아니며, 예시하고자 하는 것이다.

[0440] [표 6]

전달 벡터/방식		비-분열 세포로의 전달	발현 지속	게놈 혼입	전달된 분자의 유형
물성(예를들어, 전기천공, 입자 총, 인산칼륨 형질 감염, 세포 압착 또는 스퀴징)		있음	일시적	없음	핵산 및 단백질
바이러스	레트로바이러스	없음	안정	있음	RNA
	렌티바이러스	있음	안정	있음/없음 변형 예	RNA
	아데노바이러스	있음	일시적	없음	DNA
	아데노 관련 바이러스(AAV)	있음	안정	없음	DNA
	백시니아 바이러스	있음	매우 일시적	없음	DNA

[0441]

	단순 포진 바이러스	있음	안정	없음	DNA
비-바이러스	양이온성 리포솜	있음	일시적	전달된 것에 좌우됨	핵산 및 단백질
	중합체성 나노입자	있음	일시적	전달된 것에 좌우됨	핵산 및 단백질
생물학적 비-바이러스 전달 비히클	약독화된 박테리아	있음	일시적	없음	핵산
	조작된 박테리오파지	있음	일시적	없음	핵산
	포유류 바이러스-유사 입자	있음	일시적	없음	핵산
	생물학적 리포솜: 적혈구 고스트 및 엑소좀	있음	일시적	없음	핵산

[0442]

게놈 편집 시스템의 핵산 기반 전달

[0443]

본 개시내용에 따른 게놈 편집 시스템의 다양한 요소를 인코딩하는 핵산은 당업계에 공지된 방법에 의해 또는 예를 들어 본 명세서에 기재된 바와 같이 대상체에 투여되거나 세포로 전달될 수 있다. 예를 들어, RNA-가이드 뉴클레아제-인코딩 및/또는 가이드 분자-인코딩 DNA뿐만 아니라, 공여체 주형 핵산은, 예를 들어 벡터(예를 들어, 바이러스 또는 비 바이러스 벡터), 비 벡터 기반 방법(예를 들어, 네이키드 DNA(naked DNA) 또는 DNA 복합체 사용), 또는 이들의 조합에 의해 전달될 수 있다.

[0444]

게놈 편집 시스템 또는 그의 성분을 인코딩하는 핵산은, 예를 들어 형질감염 또는 전기천공에 의해, 네이키드 DNA 또는 RNA로서 세포에 직접 전달될 수 있거나 표적 세포(예를 들어, 적혈구, HSC)에 의한 흡수를 촉진하는 분자(예를 들어, N-아세틸갈락토사민)에 접합될 수 있다. 표 6에 정리되어 있는 벡터와 같은 핵산 벡터를 또한 사용할 수 있다.

[0445]

핵산 벡터는 게놈 편집 시스템 성분, 예컨대 RNA-가이드 뉴클레아제, 가이드 분자 및/또는 공여체 주형을 인코딩하는 하나 이상의 서열을 포함할 수 있다. 벡터는 또한 단백질을 코딩하는 서열과 결합된(예를 들어, 그 내부에 삽입되거나 그에 융합된) 신호 웨بت아이드(예를 들어, 핵 국소화, 핵소체 국소화, 또는 미토콘드리아 국소화)를 인코딩하는 서열을 포함할 수 있다. 일 예로서, 핵산 벡터는 하나 이상의 핵 국소화 서열(예를 들어, SV40으로부터의 핵 국소화 서열)을 포함하는 Cas9 코딩 서열을 포함할 수 있다.

[0446]

핵산 벡터는 임의의 적합한 수의 조절/제어 요소, 예를 들어 프로모터, 인핸서, 인트론, 폴리아데닐화 신호, Kozak 공통 서열 또는 내부 리보솜 진입 부위(IRES)를 포함할 수 있다. 이러한 요소는 당업계에 널리 공지되어 있으며, 문헌[Cotta-Ramusino]에 기재되어 있다.

[0447]

본 개시내용에 따른 핵산 벡터는 재조합 바이러스 벡터를 포함한다. 예시적 바이러스 벡터가 표 6에 제시되어 있으며, 추가의 적절한 바이러스 벡터 및 그 용도 및 생성이 문헌[Cotta-Ramusino]에 기재되어 있다. 당업계에 공지되어 있는 다른 바이러스 벡터를 또한 사용할 수 있다. 추가로, 바이러스 입자를 사용하여, 핵산 및/또는 웨بت아이드 형태로 게놈 편집 시스템 성분을 전달할 수 있다. 예를 들어, "빈" 바이러스 입자를 임의의 적절한 내

용물을 포함하도록 조립할 수 있다. 또한 바이러스 벡터 및 바이러스 입자를 조작하여, 표적화 리간드를 혼입시킴으로써 표적 조직 특이성을 변이시킬 수 있다.

[0449] 바이러스 벡터 이외에, 비 바이러스 벡터를 사용하여 본 개시내용에 따른 계놈 편집 시스템을 인코딩하는 핵산을 전달할 수 있다. 하나의 주요 범주의 비 바이러스 핵산 벡터는 유기 또는 무기물일 수 있는 나노입자이다. 나노입자는 당업계에 널리 공지되어 있으며, 문헌[Cotta-Ramusino]에 정리되어 있다. 임의의 적절한 나노입자 설계를 사용하여, 계놈 편집 시스템 성분 또는 이러한 성분을 인코딩하는 핵산을 전달할 수 있다. 본 개시내용의 특정 구현예에서, 예를 들어 유기(예를 들어 지질 및/또는 중합체) 입자가 전달 비히클로서 사용하기에 적합할 수 있다. 나노입자 제형 및/또는 유전자 전달에 사용하기 위한 예시적 지질이 표 7에 제시되어 있으며, 표 8에는 유전자 전달 및/또는 나노입자 제형에 사용하기 위한 예시적 중합체가 열거되어 있다.

[0450]

[표 7]

유전자 전달에 사용되는 지질

지질	약자	특징
1,2-디올레오일-sn-글리세로-3-포스파티딜콜린	DOPC	헬퍼
1,2-디올레오일-sn-글리세로-3-포스파티딜에탄올아민	DOPE	헬퍼
콜레스테롤		헬퍼
N -(1-(2,3-디올레오일옥시)프로필] N,N,N -트리메틸암모늄 클로라이드	DOTMA	양이온성
1,2-디올레일옥시-3-트리메틸암모늄-프로판	DOTAP	양이온성
디옥타데실아미도글리실스페르민	DOGS	양이온성
N -(3-아미노프로필)- N,N -디메틸-2,3-비스(도데실옥시)-1- 프로판아미늄 브로마이드	GAP-DLRIE	양이온성
세틸트리메틸암모늄 브로마이드	CTAB	양이온성
6-라우록시헥실 오르니티네이트	LHON	양이온성
1-(2,3-디올레오일옥시프로필)-2,4,6-트리메틸피리디늄	20c	양이온성
2,3-디올레일옥시- N -(2(스페르민카르복사미도-에틸)- N,N - 디메틸-1-프로판아미늄 트리플루오로아세테이트	DOSPA	양이온성
1,2-디올레일-3-트리메틸암모늄-프로판	DOPA	양이온성
N -(2-하이드록시에틸)- N,N -디메틸-2,3-비스(테트라데실옥시)- 1-프로판아미늄 브로마이드	MDRIE	양이온성
디미리스토옥시프로필 디메틸 하이드록시에틸 암모늄 브로마이드	DMRI	양이온성
3 β -[N -(N , N' -디메틸아미노에탄)-카르바모일]콜레스테롤	DC-Chol	양이온성
비스-구아니듐-트렌-콜레스테롤	BGTC	양이온성

[0451]

1,3-디오데옥시-2-(6-카르복시-스페르밀)-프로필아미드	DOSPER	양이온성
디메틸옥타데실암모늄 브로마이드	DDAB	양이온성
디옥타데실아미도글리실스페르미딘	DSL	양이온성
rac-[(2,3-디옥타데실옥시프로필)(2-하이드록시에틸)]- 디메틸암모늄 클로라이드	CLIP-1	양이온성
rac-[2(2,3-디헥사데실옥시프로필)- 옥시메틸옥시)에틸]트리메틸암모늄 브로마이드	CLIP-6	양이온성
에틸디미리스토일포스파티딜콜린	EDMPC	양이온성
1,2-디스테아릴옥시- <i>N,N</i> -디메틸-3-아미노프로판	DSDMA	양이온성
1,2-디미리스토일-트리메틸암모늄 프로판	DMTAP	양이온성
0,0' -디미리스틸- <i>N</i> -리실 아스파테이트	DMKE	양이온성
1,2-디스테아로일-sn-글리세로-3-에틸포스포콜린	DSEPC	양이온성
<i>N</i> -팔미토일 D-에리트로-스팽고실 카르바모일-스페르민	CCS	양이온성
<i>N</i> -t-부틸- <i>N</i> -테트라데실-3-테트라데실아미노프로파온아미딘	diC14-amidine	양이온성
옥타데케노일옥시[에틸-2-헵타데세닐-3-하이드록시에틸] 이미다졸리늄 클로라이드	DOTIM	양이온성
<i>M</i> -콜레스테릴옥시카보닐-3,7-디아자노난-1,9-디아민	CDAN	양이온성
2-(3-[비스(3-아미노-프로필)-아미노]프로필아미노)- <i>N</i> - 디테트라데실카르바모일메틸-아세트아미드	RPR209120	양이온성
1,2-딜리놀레닐옥시-3- 디메틸아미노프로판	DLinDMA	양이온성
2,2-딜리놀레닐-4-디메틸아미노에틸-[1,3]- 디옥솔란	DLin-KC2-DMA	양이온성
딜리놀레닐- 메틸-4-디메틸아미노부티레이트	DLin-MC3-DMA	양이온성

[0452]

[0453]

[표 8]

유전자 전달에 사용되는 중합체

중합체	약자
폴리(에틸렌)글리콜	PEG
폴리에틸렌이민	PEI
디티오비스(숙신이미딜프로파오네이트)	DSP
디메틸-3,3' -디티오비스프로파온이미데이트	DTBP
폴리(에틸렌 이민) 비스카바메이트	PEIC
폴리(L-리신)	PLL
히스티딘 변형 PLL	
폴리(N -비닐파롤리돈)	PVP
폴리(프로필렌이민)	PPI
폴리(아미도아민)	PAMAM
폴리(아미도 에틸렌이민)	SS-PAEI
트리에틸렌테트라민	TETA
폴리(β -아미노에스테르)	
폴리(4-하이드록시-L-프롤린 에스테르)	PHP
폴리(알릴아민)	
폴리(α -[4-아미노부틸]-L-글리콜산)	PAGA
폴리(D,L-락틱-co-글리콜산)	PLGA
폴리(N -에틸-4-비닐파리디늄 브로마이드)	
폴리(포스파젠)들	PPZ
폴리(포스포에스테르)들	PPE

[0454]

폴리(포스포라미데이트)들	PPA
폴리(N-2-하이드록시프로필메타크릴아미드)	pHPMA
폴리(2-(디메틸아미노)에틸 메타크릴레이트)	pDMAEMA
폴리(2-아미노에틸 프로필렌 포스페이트)	PPE-EA
키토산	
갈락토실화 키토산	
N-도다실화 키토산	
히스톤	
콜라겐	
렉스트란-스페르민	D-SPM

[0455]

[0456] 비 바이러스 벡터는 선택적으로 특정 세포 유형을 선택적으로 표적화하고/하거나 흡수를 개선하기 위한 표적화 변형을 포함한다. 이를 표적화 변형은, 예를 들어 세포 특이적 항원, 단일클론 항체, 단쇄 항체, 앱타마, 중합체, 당(예를 들어, N-아세틸갈락토사민(GalNAc)), 및 세포 침투 웹타이드를 포함할 수 있다. 이러한 벡터는 또한 선택적으로 융합생성 및 엔도솜 불안정화 웹타이드/중합체를 사용하고, (예컨대, 내용물의 엔도솜 이탈의 가속화를 위한) 산 촉발 입체형태 변화를 적용하고/하거나, 예를 들어 세포 구획에서의 방출을 위해 자극 절단성 중합체를 혼입시킨다. 예를 들어, 환원성 세포 환경에서 절단되는 이황화물 기반 양이온성 중합체가 사용될 수 있다.

[0457]

특정 구현예에서, 본 명세서에 기재된 게놈 편집 시스템의 성분, 예를 들어 RNA-가이드 뉴클레아제 성분 및/또는 가이드 분자 성분 이외에, 하나 이상의 핵산 분자(예를 들어, DNA 분자)가 전달된다. 특정 구현예에서, 핵산 분자는 게놈 편집 시스템의 성분 중 하나 이상과 동일한 시간에 전달된다. 특정 구현예에서, 핵산 분자는 게놈 편집 시스템의 성분 중 하나 이상의 전달 전 또는 후(예를 들어, 약 30분, 1시간, 2시간, 3시간, 6시간, 9시간, 12시간, 1일, 2일, 3일, 1주, 2주, 또는 4주 미만)에 전달된다. 특정 구현예에서, 핵산 분자는 게놈 편집 시스템의 성분, 예를 들어 RNA-가이드 뉴클레아제 성분 및/또는 가이드 분자 성분 중 하나 이상의 전달과 상이한 수 단에 의해 전달된다. 핵산 분자는, 예를 들어 본 명세서에 기재된 전달 방법 중 어느 하나에 의해 전달될 수 있다. 예를 들어, 핵산 분자는 바이러스 벡터, 예를 들어 혼입 부재 렌티바이러스에 의해 전달될 수 있으며, RNA-가이드 뉴클레아제 분자 성분 및/또는 가이드 분자 성분은, 예를 들어 핵산(예를 들어, DNA)에 의해 야기되는 독성이 감소할 수 있도록 하기 위해 전기천공에 의해 전달될 수 있다. 특정 구현예에서, 핵산 분자는 치료 단백질, 예를 들어, 본 명세서에 기재된 단백질을 인코딩한다. 특정 구현예에서, 핵산 분자는, 예를 들어 본 명세서에 기재된 RNA 분자와 같은 RNA 분자를 인코딩한다.

게놈 편집 시스템 성분을 인코딩하는 RNP 및/또는 RNA의 전달

[0459]

RNP(가이드 분자 및 RNA-가이드 뉴클레아제의 복합체) 및/또는 RNA-가이드 뉴클레아제 및/또는 가이드 분자를 인코딩하는 RNA를 당업계에 공지된 방법에 의해 세포에 전달하거나 대상체에 투여할 수 있으며, 상기 방법 중 일부는 문헌[Cotta-Ramusino]에 기재되어 있다. 시험관내에서, RNA-가이드 뉴클레아제-인코딩 및/또는 가이드 분자-인코딩 RNA는, 예를 들어 미세주입, 전기천공, 일시적인 세포 압착 또는 스queezeing(squeezing)에 의해 전달될 수 있다(예를 들어, 문헌[Lee 2012] 참조). 지질-매개 형질감염, 웹타이드-매개 전달, GalNAc- 또는 다른 접합체-매개 전달, 및 이들의 조합이 또한 전달 시험관내 및 생체내에서의 전달에 사용될 수 있다.

[0460]

시험관내에서, 전기천공을 통한 전달은 카트리지, 챔버 또는 큐벳에서 공여체 주형 핵산 분자의 존재 또는 부재 하에 RNA-가이드 뉴클레아제 및/또는 가이드 분자를 인코딩하는 RNA와 세포를 혼합하고, 규정된 지속기간 및 진폭의 전기 자극을 1회 이상 적용하는 단계를 포함한다. 전기천공을 위한 시스템 및 프로토콜은 당업계에 공지되어 있으며, 임의의 적합한 전기천공 툴 및/또는 프로토콜을 본 개시내용의 다양한 구현예와 연계하여 사용할 수 있다.

[0461] 투여 경로

이러한 시스템을 사용하여 변이되거나 조작된 계놈 편집 시스템은 국소 또는 전신 투여인지에 상관없이, 임의의 적절한 모드 또는 경로에 의해 대상체에 투여될 수 있다. 전신 투여 방식은 경구적 경로 및 비경구적 경로를 포함한다. 비경구적 경로는, 예로서 정맥내, 골수내, 동맥내, 근육내, 피내, 피하, 비강내 및 복강내 경로를 포함한다. 전신 투여된 성분은, 예를 들어 HSC, 조혈모/전구 세포, 또는 적혈구 전구 또는 전구체 세포를 표적화하기 위해 변형되거나 제형화될 수 있다.

[0463] 국소 투여 방식은, 예로서 소주골 내로의 골수내 주입 또는 골수 공간 내로의 대퇴부내 주입 및 간문맥 내로의 주입을 포함한다. 특정 구현예에서, 훨씬 더 소량의 성분(전신 접근법과 비교하여)은, 전신(예를 들어, 정맥내) 투여되는 경우와 비교하여, 국소(예를 들어, 골수 내로 직접) 투여되는 경우, 효과를 나타낼 수 있다. 국소 투여 방식은 치료 유효량의 성분이 전신 투여되는 경우 발생할 수 있는 잠재적 독성 유해 효과의 발생률을 감소시키거나 제거할 수 있다.

[0464] 투여는 주기적 볼루스(예를 들어, 정맥내)로서 또는 내부 저장기 또는 외부 저장기(예를 들어, 정맥내 백 또는 이식 가능한 펌프)로부터의 연속 주사로서 제공될 수 있다. 성분은, 예를 들어 지속 방출 약물 전달 장치로부터의 연속 방출에 의해 국소 투여될 수 있다.

[0465] 이에 더하여, 장기간에 걸친 방출이 가능하도록 성분을 제형화할 수 있다. 방출 시스템은 생분해성 재료 또는 혼입된 성분을 확산에 의해 방출하는 재료의 매트릭스를 포함할 수 있다. 성분은 방출 시스템 내에서 균질하게 또는 이질적으로 분포될 수 있다. 다양한 방출 시스템이 유용할 수 있지만, 적절한 시스템의 선택은 특정 용도에 요구되는 방출 속도에 좌우될 것이다. 비분해성 및 분해성 방출 시스템 둘 모두가 사용될 수 있다. 적합한 방출 시스템은 중합체 및 중합체성 매트릭스, 비중합체성 매트릭스 또는 무기 및 유기 부형제 및 희석제, 예를 들어 탄산칼슘 및 당(예를 들어, 트레할로스)을 포함하지만, 이에 제한되지 않는다. 방출 시스템은 천연 또는 합성적일 수 있다. 그러나 합성적 방출 시스템은, 일반적으로 더 신뢰성 있고 더 재생 가능하며, 더욱 규정된 방출 프로파일을 생성하므로 바람직하다. 방출 시스템 재료는 상이한 분자량을 갖는 성분이 재료를 통한 확산 또는 분해에 의해 방출되도록 선택될 수 있다.

[0466] 합성 생분해성 중합체의 대표 예는, 예를 들어: 폴리아미드 예를 들어 폴리(아미노산) 및 폴리(펩타이드); 폴리 에스테르, 예를 들어 폴리(락트산), 폴리(글리콜산), 폴리(락티-코-글리콜산), 및 폴리(카프로락톤); 폴리(무수 물); 폴리오르소에스테르; 폴리카보네이트; 및 이들의 화학적 유도체(치환, 화학적 기, 예를 들어 알킬, 알킬렌의 추가, 하이드록실화, 산화, 및 당업자에 의해 통상적으로 이루어지는 다른 변형), 공중합체 및 이들의 혼합물을 포함한다. 합성 비분해성 중합체의 대표 예는, 예를 들어: 폴리에테르, 예를 들어 폴리(에틸렌 옥사이드), 폴리(에틸렌 글리콜), 및 폴리(테트라메틸렌 옥사이드); 비닐 중합체-폴리아크릴레이트 및 폴리메타크릴레이트 예를 들어 메틸, 에틸, 다른 알킬, 하이드록시에틸 메타크릴레이트, 아크릴산 및 메타크릴산, 및 기타의 것, 예를 들어 폴리(비닐 알코올), 폴리(비닐 피롤리돈), 및 폴리(비닐 아세테이트); 폴리(우레탄); 셀룰로스 및 이들의 유도체, 예를 들어 알킬, 하이드록시알킬, 에테르, 에스테르, 니트로셀룰로스, 및 다양한 셀룰로스 아세테이트; 폴리실록산; 및 이들의 임의의 화학적 유도체(치환, 화학적 기, 예를 들어, 알킬, 알킬렌의 추가, 하이드록실화, 산화, 및 당업자에 의해 통상적으로 이루어지는 다른 변형), 공중합체 및 이들의 혼합물을 포함한다.

[0467] 폴리(락티드-코-글리콜리드) 마이크로스피어(microsphere)가 또한 사용될 수 있다. 통상적으로, 마이크로스피어는 중공 구체를 형성하는 구조의 락트산 및 글리콜산의 중합체로 구성된다. 구체는 직경이 대략 15 마이크론 내지 30 마이크론일 수 있으며, 본 명세서에 기재된 성분에 의해 로딩될 수 있다.

[0468] 성분의 다중 방식 또는 차등 전달

[0469] 본 개시내용의 관점에서, 본 명세서에 개시된 계놈 편집 시스템의 상이한 성분이 함께 또는 개별적으로 동시에 또는 순차적으로 전달될 수 있음을 당업자는 이해할 것이다. 계놈 편집 시스템 성분의 개별적 및/또는 순차적 전달은 계놈 편집 시스템의 기능에 대한 시간적 또는 공간적 제어를 제공하고, 그 작용으로 인해 야기되는 특정 효과를 제한하기 위한 경우에 특히 바람직할 수 있다.

[0470] 본 명세서에서 사용되는 바와 같이, 차등 또는 차등적 방식은 대상 성분 분자, 예를 들어 RNA-가이드 뉴클레아제 분자, 가이드 분자, 주형 핵산 또는 내포물에 차등적 약역학적 또는 약동학적 특성을 부여하는 전달 방식을 지칭한다. 예를 들어, 상기 전달 방식은 차등적 조직 분포, 차등적 반감기, 또는 예를 들어 선택된 구획, 조직 또는 기관에서의 차등적 시간 분포를 야기할 수 있다.

- [0471] 일부 전달 방식, 예를 들어 자율 복제 또는 세포 핵산 내로의 삽입에 의해, 세포 또는 세포의 자손에서 지속되는 핵산 벡터에 의한 전달은 성분의 지속적인 발현 및 존재를 야기한다. 예는 바이러스, 예를 들어 AAV 또는 렌티바이러스 전달을 포함한다.
- [0472] 예로서, 계놈 편집 시스템의 성분, 예를 들어 RNA-가이드 뉴클레아제 및 가이드 분자는 결과적인 반감기 또는 신체 또는 특정 구획, 조직 또는 기관에서의 전달된 성분의 지속성의 관점에서 상이한 방식으로 전달될 수 있다. 특정 구현예에서, 가이드 분자는 이러한 방식으로 전달될 수 있다. RNA-가이드 뉴클레아제 분자 성분은 신체 또는 특정 구획 또는 조직 또는 기관에 대한 더 낮은 지속성 또는 더 낮은 노출을 야기하는 방식으로 전달될 수 있다.
- [0473] 더욱 일반적으로, 특정 구현예에서, 제1 전달 방식은 제1 성분을 전달하기 위해 사용되고, 제2 전달 방식은 제2 성분을 전달하기 위해 사용된다. 제1 전달 방식은 제1의 약역학적 또는 약동학적 특성을 부여한다. 제1 약역학적 특성은, 예를 들어 신체, 구획, 조직 또는 기관에서 성분 또는 성분을 인코딩하는 핵산의 분포, 지속성 또는 노출일 수 있다. 제2 전달 방식은 제2의 약역학적 또는 약동학적 특성을 부여한다. 제2 약역학적 특성은, 예를 들어 신체, 구획, 조직 또는 기관에서 성분 또는 성분을 인코딩하는 핵산의 분포, 지속성 또는 노출일 수 있다.
- [0474] 특정 구현예에서, 제1 약역학적 또는 약동학적 특성, 예를 들어 분포, 지속성 또는 노출은 제2 약역학적 또는 약동학적 특성보다 더 제한된다.
- [0475] 특정 구현예에서, 제1 전달 방식은 약역학적 또는 약동학적 특성, 예를 들어 분포, 지속성 또는 노출이 최적화, 예를 들어 최소화되도록 선택된다.
- [0476] 특정 구현예에서, 제2 전달 방식은 약역학적 또는 약동학적 특성, 예를 들어 분포, 지속성 또는 노출이 최적화, 예를 들어 최대화되도록 선택된다.
- [0477] 특정 구현예에서, 제1 전달 방식은 상대적 지속 요소, 예를 들어 핵산, 예를 들어 플라스미드 또는 바이러스 벡터, 예를 들어, AAV 또는 렌티바이러스의 사용을 포함한다. 이러한 벡터는 상대적으로 지속적이므로, 이들로부터 전사된 산물은 상대적으로 지속적이다.
- [0478] 특정 구현예에서, 제2 전달 방식은 상대적으로 일시적인 요소, 예를 들어 RNA 또는 단백질을 포함한다.
- [0479] 특정 구현예에서, 제1 성분은 가이드 분자를 포함하고, 전달 방식은 상대적으로 지속적이며, 예를 들어 가이드 분자는 플라스미드 또는 바이러스 벡터, 예를 들어 AAV 또는 렌티바이러스로부터 전사된다. 이들 유전자의 전사는, 상기 유전자가 단백질 산물을 인코딩하지 않고, 가이드 분자가 단리시에 작용할 수 없으므로 생리적 결과가 거의 없다. 제2 성분인 RNA-가이드 뉴클레아제 분자는, 예를 들어 mRNA 또는 단백질로서 일시적 방식으로 전달되며, 이는 전체 RNA-가이드 뉴클레아제 분자/가이드 분자 복합체가 단기간 동안만 존재하고 활성인 것을 보장한다.
- [0480] 추가로, 성분은 서로 보완적인 상이한 분자 형태 또는 상이한 전달 벡터로 전달되어, 안전성 및 조직 특이성이 향상될 수 있다.
- [0481] 차등적 전달 방식의 사용은 성능, 안전성 및/또는 효능을 향상시킬 수 있으며, 예를 들어 결과적인 표적외 변형의 가능성이 감소될 수 있다. 지속성이 더 낮은 방식에 의한 면역원성 성분, 예를 들어 Cas9 분자의 전달은, 박테리아-유래 Cas 효소로부터의 웨타이드가 MHC 분자에 의해 세포의 표면상에 발현되므로, 면역원성을 감소시킬 수 있다. 2-부분 전달 시스템은 이러한 단점을 완화할 수 있다.
- [0482] 차등적 전달 방식은 성분을 상이하지만, 중첩되는 표적 영역에 전달하기 위해 사용될 수 있다. 활성 복합체의 형성은 중첩 표적 영역 외에서 최소화된다. 따라서, 특정 구현예에서, 제1 성분, 예를 들어 가이드 분자는 제1 전달 방식에 의해 전달되어, 제1 공간, 예를 들어 조직, 분포가 야기된다. 제2 성분, 예를 들어 RNA-가이드 뉴클레아제 분자는 제2 전달에 의해 전달되어, 제2 공간, 예를 들어 조직, 분포가 야기된다. 특정 구현예에서, 제1 방식은 리포솜, 나노입자, 예를 들어 중합체 나노입자, 및 핵산, 예를 들어 바이러스 벡터로부터 선택되는 제1 요소를 포함한다. 제2 방식은 상기 군으로부터 선택되는 제2 요소를 포함한다. 특정 구현예에서, 제1 전달 방식은 제1 표적화 요소, 예를 들어 세포 특이적 수용체 또는 항체를 포함하며, 제2 전달 방식은 상기 요소를 포함하지 않는다. 특정 구현예에서, 제2 전달 방식은 제2 표적화 요소, 예를 들어 제2 세포 특이적 수용체 또는 제2 항체를 포함한다.
- [0483] RNA-가이드 뉴클레아제 분자가 바이러스 전달 벡터, 리포솜 또는 중합체 나노입자로 전달될 때, 단일 조직만을 표적화하는 것이 바람직할 수 있는 경우, 다수 조직에서 전달 및 치료 활성의 가능성이 있다. 2-부분 전달 시스

템은 이러한 문제를 해결하고, 조직 특이성을 향상시킬 수 있다. 가이드 분자 및 RNA-가이드 뉴클레아제 분자가 별개이지만 중첩되는 조직 특성을 갖는 개별 전달 비히클에 패키징되는 경우, 완전히 기능적인 복합체는 두 벡터 모두에 의해 표적화된 조직에서만 형성된다.

[0484] 실시예

본 개시내용의 특정 원리가 다음의 비 제한적 실시예에 의해 예시된다.

실시예 1: 아민-기능화된 가이드 분자 단편과 디숙신이미딜 카보네이트의 예시적 접합 방법

도 1a에 도시된 바와 같이, 제1 5' 가이드 분자 단편(예를 들어, 34량체)을 3' 말단에서 (C₆)-NH₂ 링커에 의해 합성하고, 제2 3' 가이드 분자 단편(예를 들어, 66량체)을 5' 말단에서 TEG-NH₂ 링커에 의해 합성한다. 2개의 가이드 분자 단편을 10 mM 봉산나트륨, 150 mM NaCl, 및 5 mM MgCl₂를 포함하는 pH 8.5 완충액 중에서 1:1의 몰비로 혼합한다. 생성된 가이드 분자 농도는 약 50 μM 내지 100 μM이다. 2개의 가이드 분자 단편을 어닐링한 후, DMF(2.5 mM 최종 농도) 중 디숙신이미딜 카보네이트(DSC)를 첨가한다. 반응 혼합물을 간단하게 와류시킨 후, 실온에서 1시간 동안 혼합한 다음, 과량의 디숙신이미딜 카보네이트를 제거하고, 음이온 교환 HPLC 정제를 수행한다.

실시예 2: 티올-기능화된 가이드 분자 단편과 브로모아세틸-기능화된 가이드 분자 단편의 예시적 접합 방법

도 2a에 도시된 바와 같이, 제1 5' 가이드 분자 단편(예를 들어, 34량체)을 3' 말단에서 (C₆)-NH₂ 링커에 의해 합성한다. 이를 pH 8.5의 100 mM 봉산염 완충액 중에 혼탁시킨다. 가이드 분자 농도는 약 100 μM 내지 1 mM이다. DMSO(50 당량) 중 숙신이미딜-3-(브로모아세트아미도)프로파이오네이트(SBAP) 0.2배 용적을 가이드 분자 용액에 첨가한다. 30분 동안 실온에서 혼합한 후, pH 7.0의 100 mM 포스페이트 완충액 10배 용적을 첨가한다. 혼합물을 10,000 MW Amicon에서 10X 이상 농축시킨다. (a) 10배 용적의 물을 첨가하고, (b) 10,000 MW Amicon에서 10X 이상으로 농축시킴으로써, 혼합물을 추가로 처리한다. 단계 (a) 및 (b)를 3회 반복함으로써, 3' 말단에 브로모아세틸 모이어티를 갖는 제1 5' 가이드 분자 단편(예를 들어, 34량체)을 제공한다.

도 2b에 도시된 바와 같이, 제2 3' 가이드 분자 단편(예를 들어, 66량체)을 5' 말단에서 TEG-NH₂ 링커에 의해 합성한다. 이를 1 mM EDTA를 포함하는 pH 8.5의 100 mM 봉산염 완충액 중에 혼탁시킨다. 가이드 분자 농도는 약 100 μM 내지 1 mM이다. DMSO(50 당량) 중 숙식이미딜-3-(2-피리딜디티오)프로파이오네이트(SPDP) 0.2배 용적을 가이드 분자 용액에 첨가한다. 1시간 동안 실온에서 혼합한 후, 1 M 디티오트레이탈(DTT)을 1x PBS 중에 첨가한다. 혼합물 중 DTT의 최종 농도는 20 mM이다. 30분 동안 실온에서 혼합한 후, 5 M NaCl을 첨가하여, 혼합물 중 0.3 M NaCl의 최종 농도를 생성한 다음, 3배 용적의 에탄올을 첨가한다. (a) -20°C에서 15분 동안 냉각시키고; (b) 17,000 g에서 (바람직하게는 4°C) 5분 동안 원심분리하고; (c) 상청액을 제거하고; (d) 잔여물을 0.3 M NaCl 중에 혼탁시키고(아르곤을 분산시킴); (e) 3배 용적의 에탄올을 첨가함으로써, 혼합물을 추가로 처리한다. 단계 (a) 내지 (e)를 3회 반복한다. 생성된 펠릿(즉, 5' 말단에 티올을 갖는 제2 3' 가이드 분자 단편)을 진공건조시킨다.

도 2c에 도시된 바와 같이, 5' 말단에 티올을 갖는 제2 3' 가이드 분자 단편(예를 들어, 66량체)을 2 mM EDTA를 포함하는 pH 8의 100 mM 포스페이트 완충액 중에 혼탁시킨다(아르곤을 분산시킴). 가이드 분자 농도는 약 100 μM 내지 1 mM이다. 3' 말단에 브로모아세틸 모이어티를 갖는 제1 5' 가이드 분자 단편(예를 들어, 34량체)을 물(제2 3' 가이드 분자 단편 혼합물의 용적 대비 약 0.1배 용적) 중에 혼탁시킨다. 가이드 분자 농도는 약 100 μM 내지 1 mM이다. 제1 5' 가이드 분자 단편 혼합물을 제2 3' 가이드 분자 단편 혼합물에 첨가한다(아르곤을 분산시킴). 반응 혼합물을 밤새 실온에서 혼합한 후, 음이온 교환 HPLC 정제를 수행한다.

실시예 3: 포스페이트 가이드 분자 단편과 카보디이미드를 갖는 3' 하이드록실 가이드 분자 단편의 예시적 접합 방법

도 3a 및 도 3b에 도시된 바와 같이, 제1 5' 가이드 분자 단편(예를 들어, 34량체)을 표준 포스포르아미디트 화학반응을 사용하여 합성한다. 5'-포스페이트를 포함하는 제2 3' 가이드 분자 단편(예를 들어, 66량체)를 합성한다. 제1 및 제2 가이드 분자 단편을 커플링 완충액(100 mM 2-(N-모르폴리노)에탄슬픈산(MES), pH 6, 150 mM NaCl, 5 mM MgCl₂, 및 10 mM ZnCl₂) 중에서 1:1 몰비로 혼합한다. 2개의 가이드 분자 단편을 어닐링한 후, 100 mM 1-에틸-3-(3-디메틸아미노프로필)카보디이미드(EDC) 및 90 mM 이미다졸을 첨가한다. 반응 혼합물을 4°C에서 1일 내지 5일 동안 혼합한 후, 탈염 및 음이온 교환 HPLC 정제를 수행한다.

[0494] 실시예 4: HEK293T 세포에서 가이드 분자의 평가

[0495] 실시예 2의 방법에 따라 접합된 가이드 분자의 활성을 HEK293T 세포에서 T7E1 절단 분석을 통해 평가하였다. 명확히 하기 위해, 이 실시예에 사용된 모든 가이드 분자는 다음의 표 9에 제시된 바와 같은 동일한 표적화 도메인 서열, 및 실질적으로 유사한 RNA 백본 서열을 포함하였다. 표에서, 표적 도메인 서열은 "N"에 의해 축퇴 서열로 표시되고, 2개의 가이드 분자 단편 사이의 교차 연결 위치는 [L]로 표시된다.

[0496] [표 9]

가이드 분자 또는 가이드 분자 단편	SEQ ID NO.	서열
100량체 gRNA	32	NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUA GCAAGUUAAAUAAGGCUAGUCCGUUAUCACUUGAAA AAGUGGCACCGAGUCGGUGCUUUU
34량체 5' gRNA 단편	33	NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGA
66량체 3' gRNA 단편	34	AAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCACUU GAAAAAGUGGCACCGAGUCGGUGCUUUU
100량체 접합 gRNA	35	NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGA [L] A AUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCACUUG AAAAAGUGGCACCGAGUCGGUGCUUUU

[0497] [0498] IVT에 의해 생성된 단분자 가이드 분자, 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨), 또는 실시예 2의 브로모아세틸-티올 방법에 의해 접합된 합성 단분자 가이드 분자를 다양하게 포함하는 다양한 농도의 리보핵 단백질 복합체를 리포렉션(CRISPR-Max, Thermo Fisher Scientific, Waltham, MA)에 의해 HEK293T 세포에 도입시키고, 게놈 DNA를 추후 수집하였다. 시판 키트(Integrated DNA 시스템, Coralville, Iowa로부터 상업적으로 입수 가능한 Surveyor™)를 사용하고, 표준 T7E1 절단 분석을 사용하여 절단을 평가하였다. 결과는 도 4에 제시되어 있다.

[0499] 결과가 나타내는 바와 같이, 접합된 가이드 분자는 HEK293 세포에서 IVT 또는 합성 단분자 가이드 분자에 의해 생성된 단분자 가이드 분자에 의해 관찰된 것과 일치하는 투여량 의존적 방식으로 절단을 지원하였다. 접합되지 않은 어닐링된 가이드 분자 단편은 유사한 투여량 의존적 방식이지만, 더 낮은 수준의 절단을 지원한다는 것을 유의해야 한다. 이들 결과는 본 개시내용의 방법에 따라 접합된 가이드 분자가 IVT 또는 합성 단분자 가이드 분자에 의해 생성된 단분자 가이드 분자와 실질적으로 동일한 방식으로 높은 수준의 DNA 절단을 지원한다는 것을 시사한다.

[0500] 실시예 5: 겔 전기영동 및 질량 분광분석에 의한 가이드 분자 순도의 평가

[0501] [0500] 실시예 1의 방법에 따라 우레아 링커와 접합된 가이드 분자의 조성물의 순도를 총 이온 전류 크로마토그래피 및 질량 분광분석에 의해 시판 제조된 합성 단분자 가이드 분자(접합 과정 없이 제조됨)의 조성물의 순도와 비교하였다. 100 pmol의 분석물을 질량 분석을 위해 주입하였다. Waters ACQUITY UPLC 시스템이 장착된 Bruker microTOF-QII 질량 분광분석기에서 LC-MS에 의해 분석을 획득하였다. 분리를 위해 ThermoDNAPac C18 컬럼을 사용하였다. 결과를 도 5에 도시하였다.

[0502] 도 5a는 대표적인 이온 크로마토그래프를 도시하고, 도 5b는 실시예 1의 방법에 따라 우레아 링커와 접합된 이온 교환 정제 가이드 분자의 디콘볼루션된 질량 스펙트럼을 도시한다. 도 5c는 대표적인 이온 크로마토그래피를 도시하고, 도 5d는 시판 제조된 합성 단분자 가이드 분자의 디콘볼루션된 질량 스펙트럼을 도시한다. 이온 크로마토그래프에서 강조 표시된 피크에 대해 질량 스펙트럼을 평가했다. 도 5e는 질량 스펙트럼의 확대 버전을 도시한다. 시판 제조된 합성 단분자 가이드 분자의 질량 스펙트럼은 좌측에 도시되어 있으며(총 질량 순도 34%), 실시예 1의 방법에 따라 우레아 링커와 접합된 가이드 분자에 대한 질량 스펙트럼이 우측에 도시되어 있다(총 질량 순도 72%).

[0503] 실시예 6: 서열 분석에 의한 가이드 분자 순도의 평가

[0504] [0503] 실시예 1에서 기재된 바와 같이, 우레아 연결에 의해 접합된 가이드 분자의 조성물의 순도를 시판 제조된 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨)의 조성물 및 실시예 2에 기재된 바와 같이, 티오에테르 연결에

의해 접합된 가이드 분자의 조성물의 순도와 비교하였다. 모든 가이드 분자의 조성물은 소정의 동일한 가이드 분자 서열을 기반으로 하였다.

[0505] 도 6a는 우레아 연결을 포함하는 합성 단분자 가이드 분자로부터 생성된 상보성 DNA(cDNA)의 5' 말단으로부터 각 위치에서 개별 염기 및 길이 변화가 발생하는 빈도를 나타내는 그래프를 도시하고, 도 6b는 시판 제조된 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨)로부터 생성된 cDNA의 5' 말단으로부터 각 위치에서 개별 염기 및 길이 변화가 발생하는 빈도를 나타내는 그래프를 도시한다. 가이드 분자의 20 bp 표적화 도메인에 박스 표시하였다. 본 실시예에서, 시판 제조된 합성 단분자 가이드 분자로부터의 가이드 분자와 비교하여(즉, 10% 미만의 가이드 분자가 임의의 소정의 위치에서 결실을 포함하고, 5% 미만이 임의의 소정의 위치에서 치환을 포함함), 우레아 연결을 포함하는 가이드 분자는 표적화 도메인에서 더 큰 서열 정확도를 야기하였다(즉, 1% 미만의 가이드 분자가 임의의 소정의 위치에서 결실을 포함하고, 1% 미만의 가이드 분자가 임의의 소정의 위치에서 치환을 포함함).

[0506] 도 6c는 티오에테르 연결을 포함하는 합성 단분자 가이드 분자로부터 생성된 cDNA의 5' 말단으로부터 각 위치에서 개별 염기 및 길이 변화가 발생하는 빈도를 나타낸 그래프를 도시한다. 도 6c에 도시된 바와 같이, 높은 수준의 5' 서열 정확도가 관찰되었으며, 이는 서열 정확도 및 순도가 높은 가이드 분자 조성물의 생성을 나타내었다. 도 6a(우레아 연결) 및 도 6c(티오에테르 연결)의 정렬은 또한 연결 부위(위치 34)에서 미스매치/삽입결실의 빈도가 상대적으로 높은 영역을 나타내었다. 이들 데이터는 본 개시내용의 방법에 의해 합성된 가이드 분자가 시판용 가이드 분자와 비교하여, 저하된 결실 및 치환 빈도를 나타냄을 시사한다.

[0507] 또한, 도 7a 및 도 7b는 우레아 연결을 포함하는 합성 단분자 가이드 분자(도 7a) 및 시판 제조된 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨)(도 7b)로부터 생성된 cDNA의 5' 말단으로부터의 처음 41개의 위치에서의 내부 서열 길이 변화(+5 내지 -5)를 도시하는 그래프이다. 도시된 바와 같이, 우레아 연결을 포함하는 가이드 분자는 시판 제조된 합성 단분자 가이드 분자(즉, 접합 단계 없이 제조됨)와 비교하여, 삽입/결실의 빈도 및 길이가 감소하였다.

실시예 7: CD34+ 세포에서 가이드 분자 활성의 평가.

[0509] 실시예 1의 방법에 따라 우레아 연결 접합된 가이드 분자의 활성을 차세대 시퀀싱 기법을 통해 CD34+ 세포에서 평가하였다. 이 실시예에서 논의된 가이드 분자는 하기 표 10 및 도 8a 내지 도 81, 도 9a 내지 도 9e, 및 도 10a 내지 도 10d에 나타낸 바와 같이, 3개의 표적화 도메인 서열 및 다양한 가이드 분자 백본 서열 중 하나를 포함하였다. 2개의 가이드 분자 단편 사이의 우레아 연결의 위치는 표 10에서 [UR]로, 도 8a 내지 도 81, 도 9a 내지 도 9e, 및 도 10a 내지 도 10d에서 ⑧로 표시되어 있다. 처음 2개의 표적화 도메인 서열의 가이드 분자(gRNA 1 다음 문자 또는 gRNA 2 다음 문자로 표기됨)는 S. 피오게네스 gRNA 백본을 기반으로 하며, 세 번째 표적화 도메인 서열의 가이드 분자(gRNA 3 다음 문자로 표기됨)는 S. 아우레우스 gRNA 백본을 기반으로 한다.

[0510] 접합된 가이드 분자를 pH 7.5 완충액 중에 재현탁시키고, 용융시키고, 재어닐링시킨 후, S. 피오게네스 Cas9의 혼탁액에 첨가하여, 55 μM 완전 복합체화된 리보핵단백질을 포함하는 용액을 생성하였다.

[0511] 인간 CD34+ 세포를 계수하고, 펠렛에 원심분리하여, P3 혁전달 완충액에 재현탁시킨 다음, 인간 HSC 배지 (StemSpan™ Serum-Free Expansion Medium, StemCell Technologies, 밴쿠버, 브리시티 콜롬비아, 캐나다)를 미리 충전시킨 96 웰 Nucleocuvette 플레이트의 각 웰에 분주하여, 50,000 세포/웰을 생성하였다. 상기 기재된 바와 같이 완전 복합체화된 리보핵단백질 용액을 Nucleocuvette 플레이트의 각 웰에 첨가한 후, 천천히 혼합하였다. Amaxa Nucleofector 시스템(Lonza, 바젤, 스위스)에서 혁전달을 수행하였다. 혁전달된 세포를 37°C 및 5% CO₂에서 72시간 동안 인큐베이션하여, 안정한 상태에서 유지되도록 하였다. 제조 업체의 지침에 따라 DNAAdvance DNA 단리 키트를 사용하여 혁전달된 세포로부터 게놈 DNA를 추출하였다. 야생형 인간 참조 서열과 비교하여, 삽입 및 결실(삽입결실)의 %를 정량화하기 위해 차세대 시퀀싱 기법을 사용하여 절단을 평가하였다. CD34+ 세포에서 시험된 표 10의 gRNA에 대한 결과가 도 11에 제시되어 있다.

[0512] 도 11의 결과가 나타내는 바와 같이, 실시예 1에 따라 생성된 결찰된 가이드 분자는 CD34+ 세포에서 DNA 절단을 지원한다. 줄기 루프 길이가 증가함에 따라, 삽입결실 %가 증가하는 것으로 나타났지만, 줄기 루프 서열에 인접한 U-A 교환의 혼입(gRNA 1E, gRNA 1F, 및 gRNA 2D 참조)은 효과를 완화한다. 이들 데이터는 더 긴 줄기 루프 특성을 갖는 화학적으로 접합된 합성 단분자 가이드 분자가 세포에서 더 높은 수준의 DNA 절단을 유발함을 시사한다. 추가로, DNA 절단 활성은 결찰 효율과 무관하며, 경험적으로 결정되어야 한다.

[0513]

[표 10]

가이드 분자	SEQ ID NO.	서열
gRNA 1A	36	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAGA[UR]AAUAGCAAGUAAAAAUAGGCUA GUCCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1B	37	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUA[UR]UAGCAAGUAAAAAUAGGCUAGUCC GUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1C	38	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAGG[UR]CCUAGCAAGUAAAAAUAGGCUA GUCCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1D	39	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAUGC[UR]GCAUAGCAAGUAAAAAUAGGC UAGUCGGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1E	40	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAUGCUGUUUUG[UR]CAAACAGCAUAGCA AGUAAAUAUAAAGGCCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1F	41	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAUGCUG[UR]CAGCAUAGCAAGUAAAUA AGGCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1G	42	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAUGCUGUUUUG[UR]CAAACAGCAUAGCA AGUAAAUAUAAAGGCCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1H	43	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAUGCUG[UR]CAGCAUAGCAAGUAAAAAA AGGCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1I	44	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAAGA[UR]AAUUAAGCAAGUAAAAUA GCCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1J	45	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAAA[UR]UUAGCAAGUAAAAAUAGGCU AGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1K	46	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAAGGA[UR]AACUUUAGCAAGUAAAA UAAGGCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 1L	47	GUAACGGCAGACUUUCUCCUCGUUUUAGAGCUAG[UR]AAUAGCAAGUAAAAAUAGGCU AGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 2A	48	CUAACAGUUGGUUUUAUCAGGUUUUAGAGCUAUGC[UR]GCAUAGCAAGUAAAAAUAGG CUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 2B	49	CUAACAGUUGGUUUUAUCAGGUUUUAGAGCUAUGCUG[UR]CAGCAUAGCAAGUAAA AAGGCUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 2C	50	CUAACAGUUGGUUUUAUCAGGUUUUAGAGCUAUGCUGUUUUG[UR]CAAACAGCAUAGC AAGUAAAUAUAGGUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 2D	51	CUAACAGUUGGUUUUAUCAGGUUUUAGAGCUAUGCUGUUUUG[UR]CAAACAGCAUAGC AAGUAAAUAUAGGUAGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 2E	52	CUAACAGUUGGUUUUAUCAGGUUUUAGAGCUAAGA[UR]AAUAGCAAGUAAAAAUAGGCU AGUCGUUAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 3A	53	GUAACGGCAGACUUUCUCCUCGUUUUAGACUCUG[UR]CAGAAUCUACAAAACAAGGCAA AAUGCCGUGUUUAUCUCGUUAACUUGUUGCGAGAUUUU
gRNA 3B	54	GUAACGGCAGACUUUCUCCUCGUUUUAGACUCUGUAA[UR]UUACAGAAUCUACAAAACA AGGCAAAAUUCGGUGUUUAUCUCGUUAACUUGUUGCGAGAUUUU
gRNA 3C	55	GUAACGGCAGACUUUCUCCUCGUUUUAGACUCUGUAAUUUAGGU[UR]ACCUAAAUA CAGAAUCUACAAAACAAGGCAAUAUCUCGUUAACUUGUUGCGAGAU UU
gRNA 3D	56	GUAACGGCAGACUUUCUCCUCGUUUUAGACUCUGUAAUUUAGGUAGAG[UR]CUCAUAC CUAAAUAUACAGAAUCUACAAAACAAGGCAAUAUCUCGUUAACUUGU GGCGAGAUUUU

[0514]

실시예 8: 결찰 효율의 컴퓨터 모델 평가

[0515]

실시예 1에서 기재된 반응의 결찰 효율은 특정 가이드 분자 구조의 적합성을 측정하는 하나의 척도이다. 실시예 1의 제1 및 제2 가이드 분자 단편의 반응성 작용기가 동일하므로(아민), 경쟁적인 동종 커플링은 잠재적인 부산물이다. 본 실시예는 <http://www.idtdna.com/calc/analyzer>에서 이용 가능한 OligoAnalyzer 3.1 툴 사용하여 이종 커플링 반응의 자유 에너지 차이(ΔG_2)와 비교하여, 동종 커플링 반응의 자유 에너지 차이(ΔG_1)의 컴퓨터 모델링을 통해 결찰 효율(즉, 반응 산물 내의 이종 커플링된 산물의 %)을 예측할 수 있는지를 평가하였다. 이 분석의 결과가 표 11에 제시되어 있다.

[0517]

[표 11]

가이드 분자	SEQ ID NO.	결찰 효율	ΔG_1 (kcal/mol)	ΔG_2 (kcal/mol)	$\Delta G_2 - \Delta G_1$ (kcal/mol)
gRNA 1A	36	약 55%	-6.90	-10.93	-4.03
gRNA 1C	38	18%	-6.90	-10.93	-4.03
gRNA 1D	39	50%	-6.90	-12.27	-5.37
gRNA 1E	40	50%	-6.34	-24.95	-18.61
gRNA 1F	41	31%	-6.34	-15.82	-9.48
gRNA 1G	42	12%	-6.90	-24.95	-18.05
gRNA 1H	43	60%	-6.90	-15.82	-8.92
gRNA 1I	44	약 50%	-6.90	-10.93	-4.03
gRNA 1J	45	약 50%	-6.90	-10.93	-4.03
gRNA 1K	46	약 55%	-6.90	-10.93	-4.03
gRNA 2A	48	18%	-6.34	-12.27	-5.93
gRNA 2C	50	48%	-6.84	-24.95	-18.11
gRNA 2D	51	45%	-6.84	-24.95	-18.11
gRNA 2E	52	5%	-6.34	-8.64	-2.30

[0518]

[0519] 표 11에 나타낸 바와 같이, 결찰 효율(겔 분석 후 농도계에 의해 측정된 바와 같음)은 대부분의 서열에 대해 $\Delta G_2 - \Delta G_1$ 값이 더 큰 음의 값인 경우, 더 유리한 결찰 효율에 상응하는 것으로 적절히 예측되었다(예를 들어, gRNA 2A 및 2C 비교). 그러나 특정 가이드 분자를 형성하기 위한 결찰 효율은 항상 $\Delta G_2 - \Delta G_1$ 값과 상관이 있는 것은 아니었으며(예를 들어, 더 큰 음의 값의 $\Delta G_2 - \Delta G_1$ 값이 더 높은 결찰 효율로 이어지지 않은 gRNA 1G 참조), 이는 변형 및 실험이 특정 가이드 분자 단편을 접합시키는 데 필요할 수 있음을 시사한다. 예를 들어 gRNA 1G의 결찰 효율은 하부 줄기의 서열에 U-A 교환을 구현함으로써 개선되었으며(gRNA 1G와 gRNA 1E의 결찰 효율 비교), U-A 교환은 결찰 전에 2개의 가이드 분자 단편의 엇갈린 어닐링을 방지하기 위해 설계되었다.

[0520]

실시예 9: 질량 분광분석에 의한 우레아 연결의 특성화

[0521]

우레아 연결을 포함하고, 실시예 1에 기재된 바와 같이 합성된 화학적으로 접합된 가이드 분자를 질량 분광분석에 의해 특성화하였다. 합성, 화학적 결찰 및 정제 후, gRNA 1A(표 10 참조)를 T1 엔도뉴클레아제를 사용하여 일차 서열의 각 G 뉴클레오타이드의 3' 말단에서 단편으로 절단하였다. 이들 단편을 LC-MS를 사용하여 분석하였다. 특히, 우레아 연결, A-[UR]-AAUAG(A34:G39)을 포함하는 단편이 4.50분의 보유 시간에서 m/z = 1190.7로 검출되었다(도 12a 및 도 12b). 이 전구체 이온의 LC/MS-MS 분석에 의해 gRNA 1A에서 우레아 연결과 일치하는 충돌 유도 해리 단편 이온이 나타났다.

[0522]

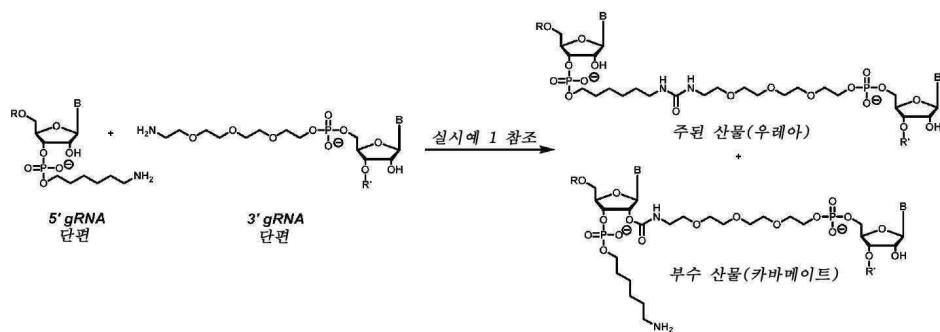
실시예 10: 카바메이트 부산물의 특성화

[0523]

도 13a는 주된 산물(A-2, 보유 시간 3.25분) 및 부수 산물(A-1, 보유 시간 3.14분) 둘 모두가 존재하는 우레아 연결된 가이드 분자의 정제되지 않은 조성물에 대한 LC-MS 데이터를 도시한다. 본 발명자들은 설명하자면, 도 13a의 부수 산물(A-1)이 더 높은 비율의 카바메이트 부수 산물을 포함하는 음이온 교환 정제로부터의 분획을 조합함으로써 다양화되었음을 주목한다. 부산물은 통상적으로 실시예 1의 방법에 따른 가이드 분자의 합성시 최대 10%의 수율로 검출된다. 질량 분광분석에 의한 각 피크의 분석에 의해 두 산물 모두가 동일한 분자량을 갖는 것으로 나타났다(도 13b 및 도 13c 참조).

[0524]

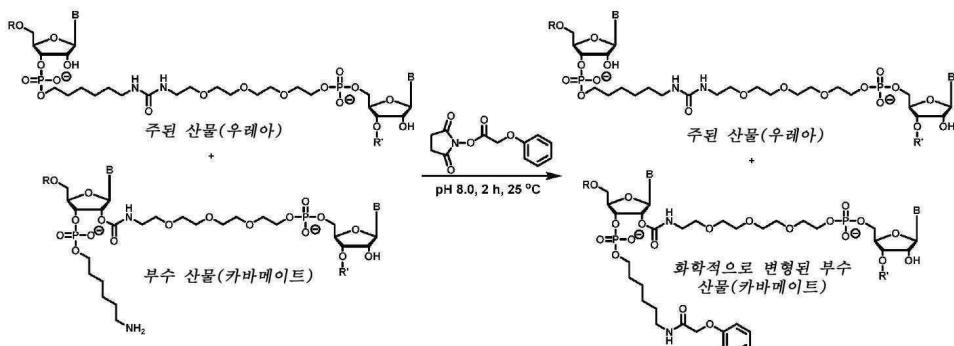
이에 비추어, 본 발명자들은 부수 산물이 다음과 같은 3' 가이드 분자 단편의 5' 말단 상의 5'-NH₂와 5' 가이드 분자 단편의 3' 말단 상의 2'-OH 사이의 반응으로부터 생성된 카바메이트 부산물이라고 가정하였다:



[0525]

[0526]

카바메이트 부산물의 지정을 추가로 확인하기 위해, 페녹시 아세트산 N-하이드록시 숙신이미드 에스테르로 화학적 변형을 수행하였다. 기본적인 화학적 원리에 의해 단지 부수 산물(카바메이트)만이 반응성 친핵성 중심을 가지며(유리 아민), 따라서 단지 부수 산물만이 화학적으로 기능화될 것으로 예측된다. 우레아 연결 가이드 분자의 미정제 조성물에 페녹시아세트산 N-하이드록시숙신이미드 에스테르를 첨가하면, 주된 산물(우레아) 및 화학적으로 변형된 부수 산물(카바메이트)이 생성되어야 한다:



[0527]

[0528]

도 14a는 화학적 변형 후의 가이드 분자 조성물에 대한 LC-MS 데이터를 도시한다. 주된 산물(B-1, 우레아)은 최초 분석과 동일한 보유 시간을 가지지만(3.26분, 도 13a), 부수 산물(B-1, 카바메이트)의 보유 시간은 3.86분으로 변하였으며, 이는 유리 아민 모이어티의 화학적 기능화와 일치하였다. 또한, 3.86분에서 피크($M + 134$)의 질량 분광 분석은 예측된 기능화가 발생했음을 나타낸다(도 14b 참조). 이를 결과는 부수 산물이 실제로 카바메이트 부산물임을 시사한다.

[0529]

카바메이트 부산물의 동일성을 추가로 확인하기 위해, 주된 산물(우레아) 및 화학적으로 변형된 부수 산물(카바메이트)의 혼합물을 일차 서열의 각 G 뉴클레오타이드의 3' 말단에서 가이드 분자를 절단하는 리보뉴클레아제 A로 분해시켰다(실시예 9 참조). 그 다음 단편을 LC-MS로 분석하였고, 우레아 연결(G35-[UR]-C36) 및 화학적으로 변형된 연결(G35-[CA+PAA]-C36)이 둘 모두 검출되었다. 도 15a는 단편 혼합물의 LC-MS 추적을 도시하고, 우레아 연결은 4.31분의 보유 시간을 화학적으로 변형된 카바메이트 연결은 5.77분의 보유 시간을 나타내었다. 도 15b는 4.31분, $m/z = 532.1$ 이 $[M-2H]^{2-}$ 에 할당된 피크의 질량 스펙트럼을 도시하며, 도 15c는 5.77분, $m/z = 599.1$ 이 $[M-2H]^{2-}$ 에 할당된 피크의 질량 스펙트럼을 도시한다. 질량 스펙트럼을 LC-MS/MS 기법을 사용하여 추가로 분석하였다. $m/z 532.1$, $[M-2H]^{2-}$ 에서의 우레아 연결 산물의 LC-MS/MS 스펙트럼(도 15d)은 올리고뉴클레오타이드 충돌 유도 해리(CID) 실험에서 관찰되는 통상적인 a-d 및 x-z 이온을 포함한다. 추가로, 5' 말단($m/z = 487.1$ 및 461.1) 및 3' 말단($m/z = 603.1$ 및 577.1)으로부터 UR 연결의 어느 한쪽에 MS/MS 단편 이온이 관찰되었다. 이와 달리, 카바메이트 연결의 5' 말단($m/z = 595.2$) 및 카바메이트 연결의 3' 말단($m/z = 603.1$)으로부터의 MS/MS 단편 이온을 포함하는 $m/z 599.1$, $[M-2H]^{2-}$ 에서의 화학적으로 변형된 카바메이트 연결 산물의 LC-MS/MS 스펙트럼(도 15e)에서 단지 2개의 산물 이온이 관찰되었다.

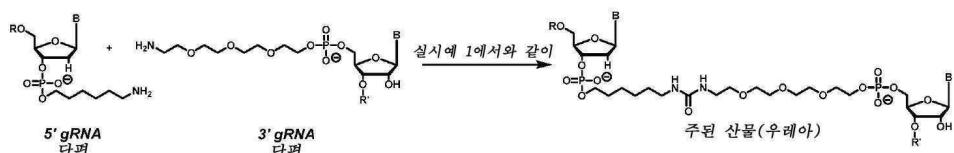
[0530]

실시예 11: 하나의 산물 형성의 뉴클레오타이드 변형

[0531]

본 발명자들은 실시예 10에 기재된 바와 같은 카바메이트 부산물의 형성이 5' 가이드 분자 단편의 3' 말단에서 뉴클레오타이드에서의 전략적 2' 변형을 통해 방지될 수 있다고 가정했다. 예를 들어, 5' 가이드 분자 단편의 3' 말단의 뉴클레오타이드에서 2'-OH를 2'-H로 대체하면, 실시예 1의 방법에 따른 우레아 연결 가이드 분자의

합성은 카바메이트 부산물이 부재하는 단일 우레아 연결 산물이 생성된다고 가정하였다:



[0532]

[0533] 도 16a는 동일한 5' 가이드 분자의 변형되지 않은 버전과의 반응의 미정제 반응 혼합물(하부 스펙트럼)과 비교한 2'-H 변형된 5' 가이드 분자 단편과의 반응의 미정제 반응 혼합물(상부 스펙트럼)의 LC-MS 데이터를 도시한다. 2'-H 변형된 5' 가이드 분자 단편의 경우 카바메이트 부산물 형성이 관찰되지 않는다(상부 스펙트럼). 이와 달리, 동일한 5' 가이드 분자 단편의 변형되지 않은 버전과의 반응의 미정제 반응 혼합물(하부 스펙트럼)은 우레아 연결된 주된 산물(A-2) 및 부수 카바메이트 부산물(A-1)의 혼합물을 포함하였다. 본 발명자들은 실시예 10과는 달리, 카바메이트 부산물은 다량 존재하지 않았으며, 따라서 실시예 10의 도 13a보다 훨씬 더 낮은 수준에서 검출되었음에 주목하였다. 추가로, 2'-H 변형된 5' 가이드 분자 단편과의 반응의 산물(B)의 질량 분광분석은 M - 16을 제공하였으며(변형되지 않고 우레아 연결된 주된 산물인 A-2와 비교), 이는 2'-OH가 2'-H로 대체된 분자에서 예측된 바와 같다(도 16b 및 도 16c 참조).

[0534]

동일한 2'-H 변형을 포함하는 표 10의 gRNA 1L을 사용하여 유사한 실험을 수행하였다. 2'-H 변형되고 우레아-연결된 가이드 분자의 형성을 T1 엔도뉴클레아제 분해 후, 질량 분광분석에 의해 확인하였다(실시예 9 참조). 우레아 연결(2'-H-A)-[UR]-AAUAG(A34:G39)를 포함하는 단편이 4.65분의 보유 시간(도 17a), m/z = 1182.7(도 17b)에서 검출되었다. 이러한 전구체 이온의 LC-MS/MS 분석에 의해 2'-H 변형된 뉴클레오타이드와의 반응의 우레아 연결과 일치하는 단편 이온이 밝혀졌다.

[0535]

[0535] 이러한 결과는 5' 가이드 분자 단편의 3' 말단에서의 뉴클레오타이드의 2'-OH 변형을 통해, 카바메이트 부산물의 형성을 방지할 수 있음을 시사한다. 결과적으로, 우레아 결찰된 가이드 분자가 고 순도로 합성되었으며, 이는 접합 가이드 분자의 전체 생성 방법을 간소화한다.

[0536]

참조문헌의 포함

[0537]

본 명세서에 언급된 모든 간행물, 특히 및 특허 출원은 각각의 개별 간행물, 특히 또는 특허 출원이 참조로서 포함되는 것으로 구체적이고, 개별적으로 지적된 것과 같이, 전체로서 본 명세서에 참조로 포함된다. 모순된 경우, 본 명세서의 임의의 정의를 포함하여, 본 출원이 우선할 것이다.

[0538]

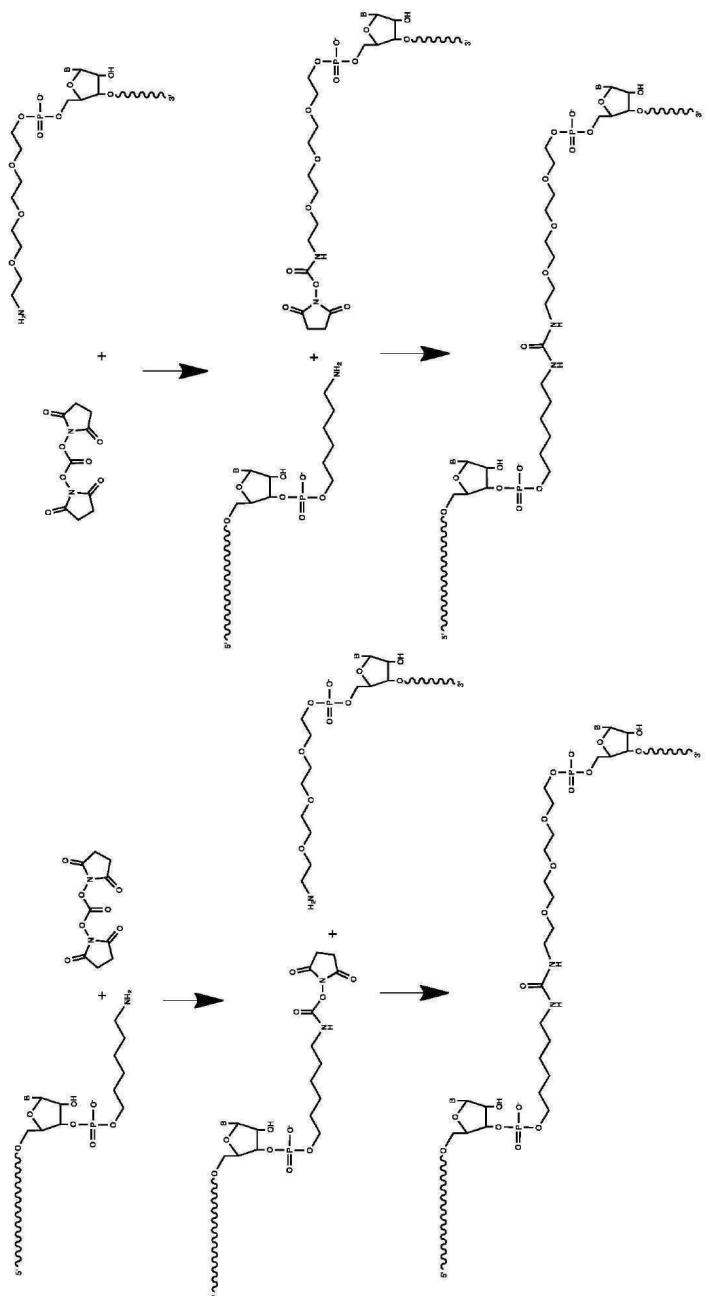
균등률

[0539]

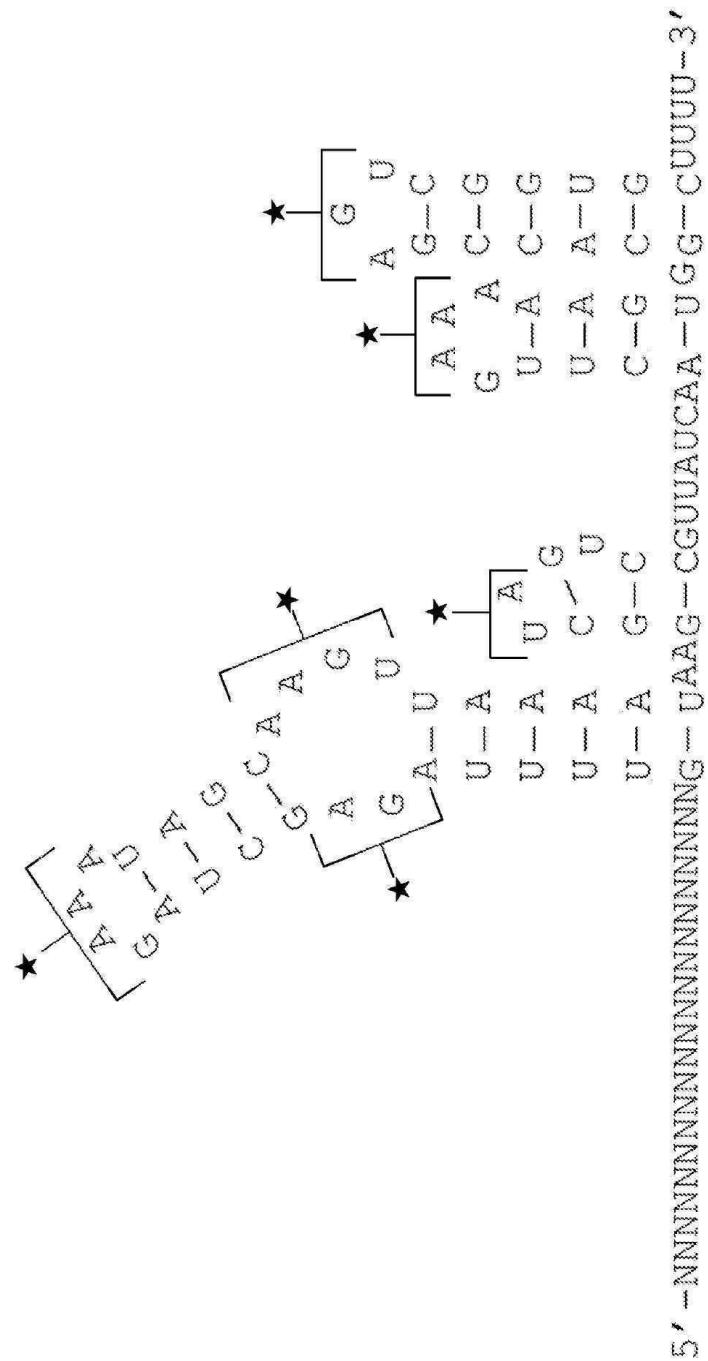
당업자는 일상적인 실험만을 사용하여, 본 명세서에 기재된 특정 구현예에 대한 다수의 균등물을 인식할 수 있거나 확인할 수 있을 것이다. 이러한 균등률은 하기 청구범위에 포함되는 것으로 의도된다.

도면

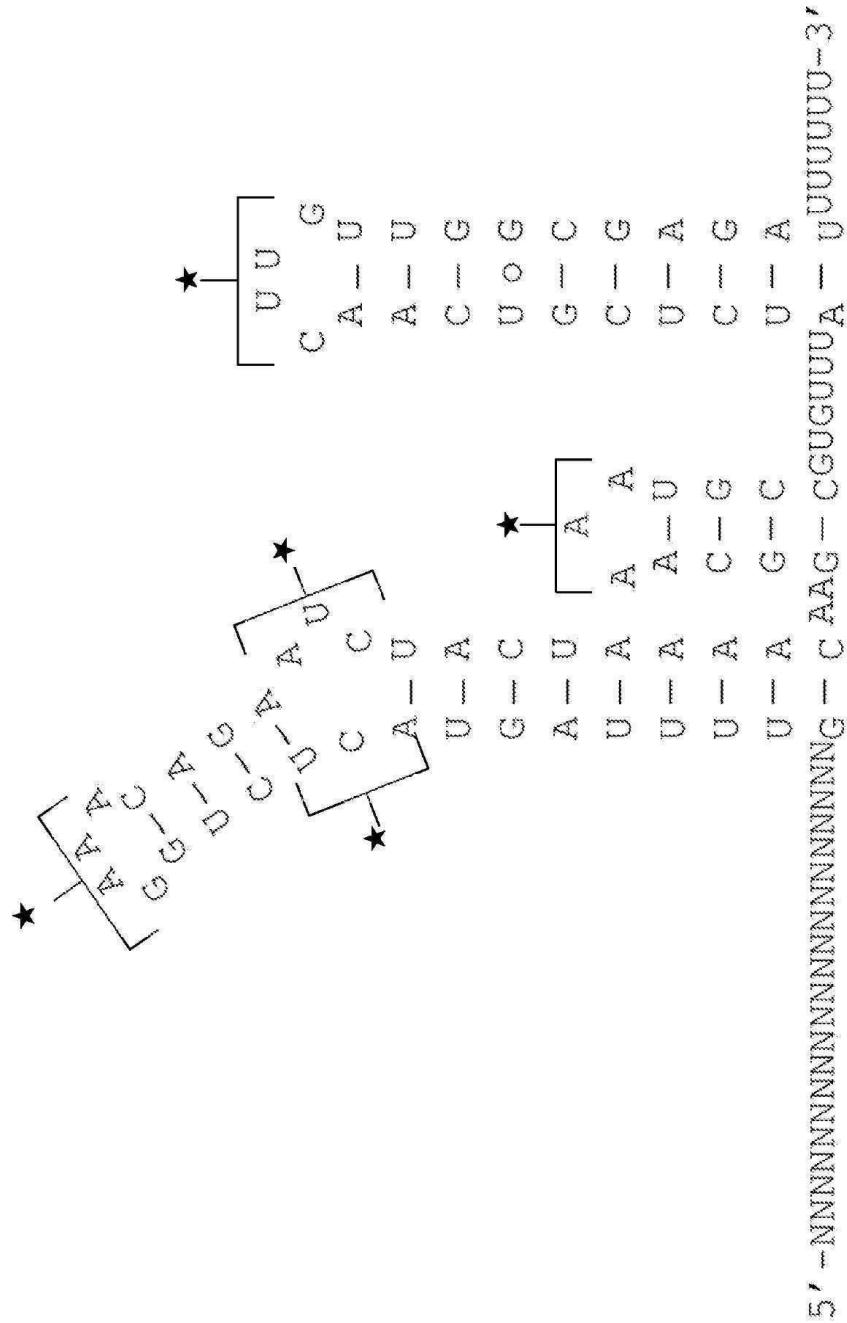
도면 1a



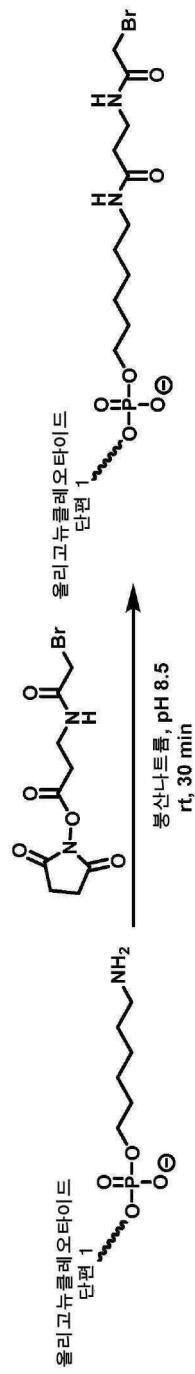
도면 1b



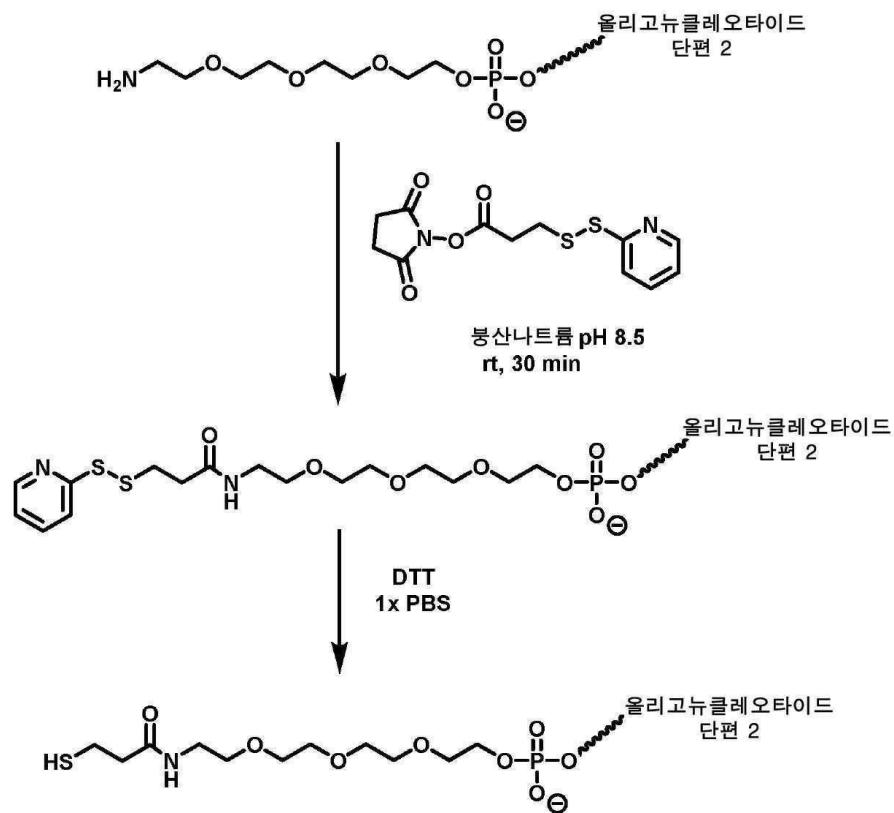
도면 1c



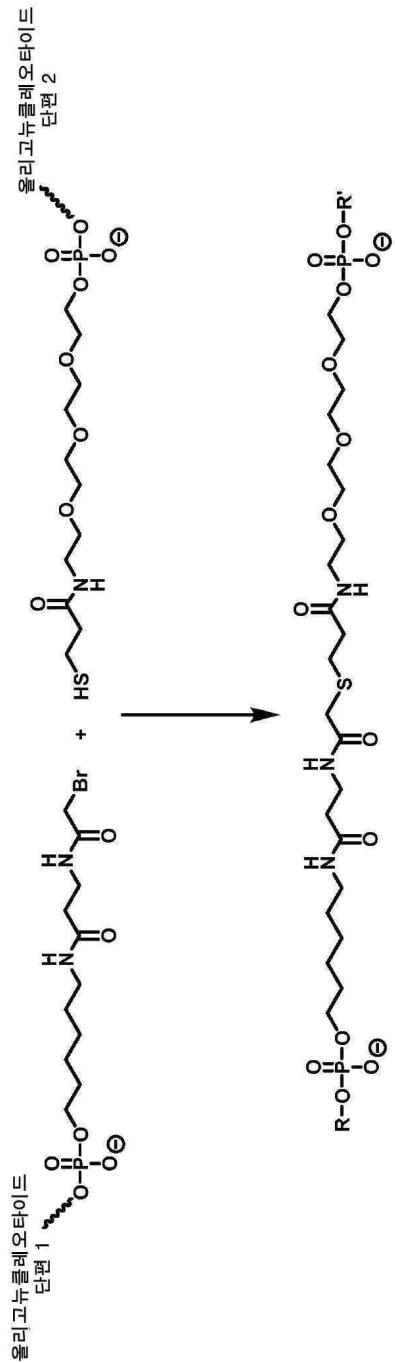
도면 2a



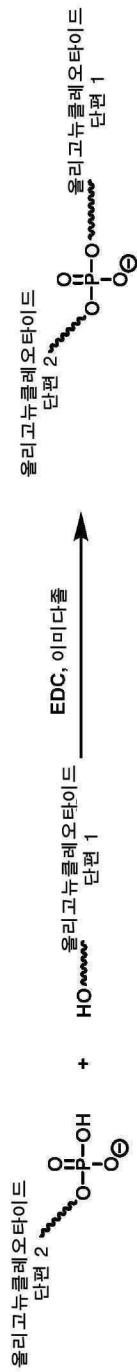
도면2b



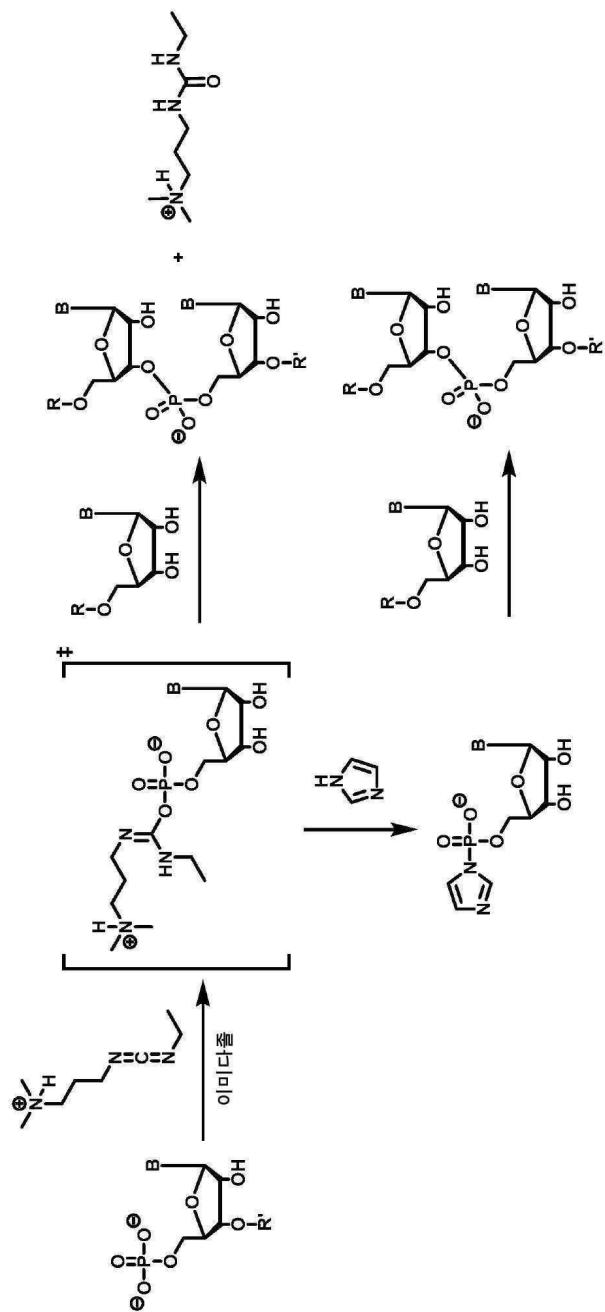
도면 2c



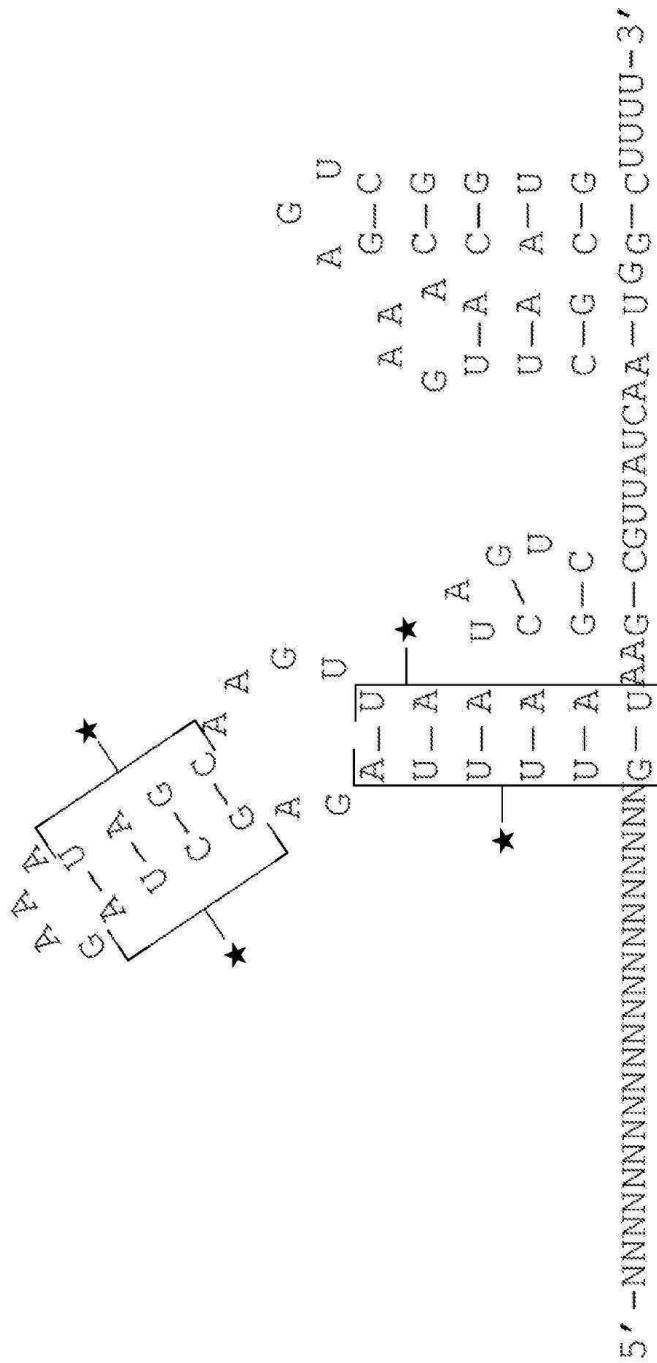
도면 3a



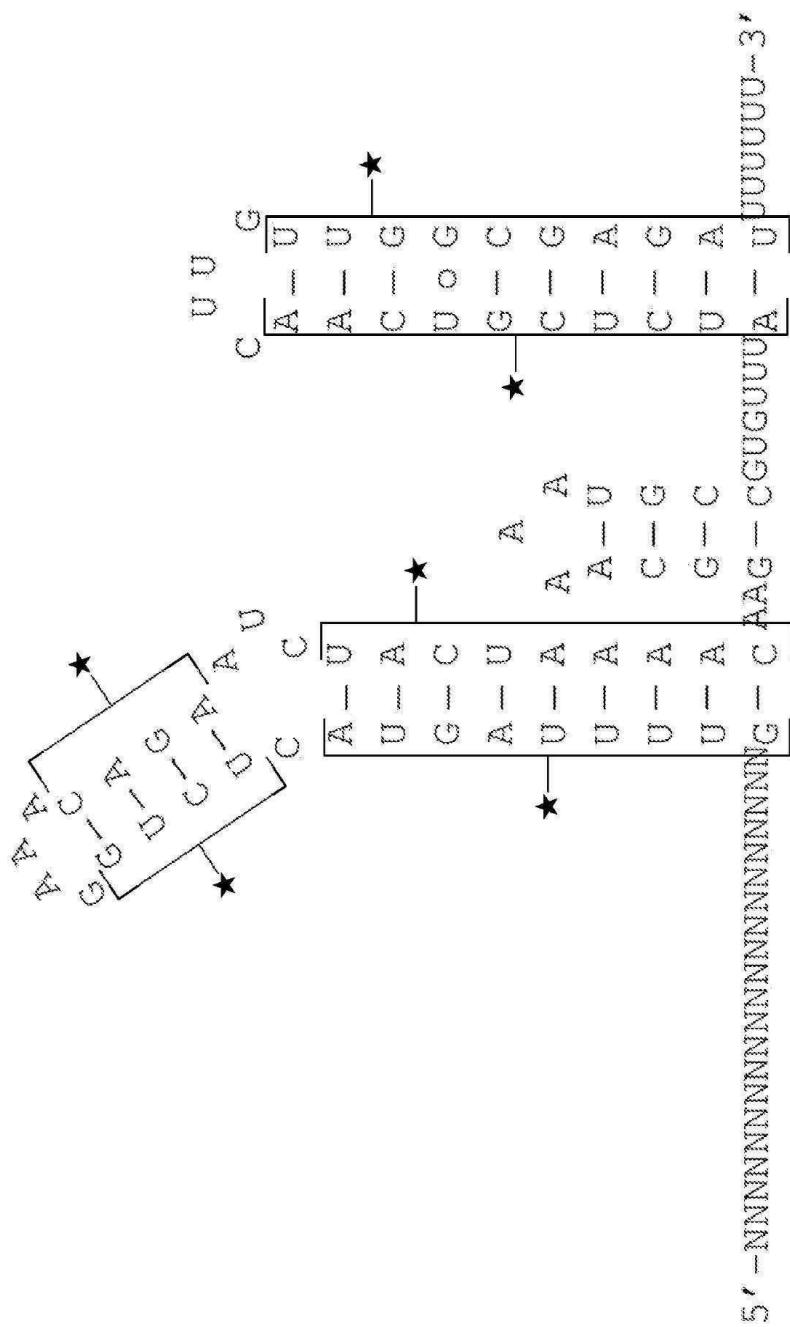
도면3b



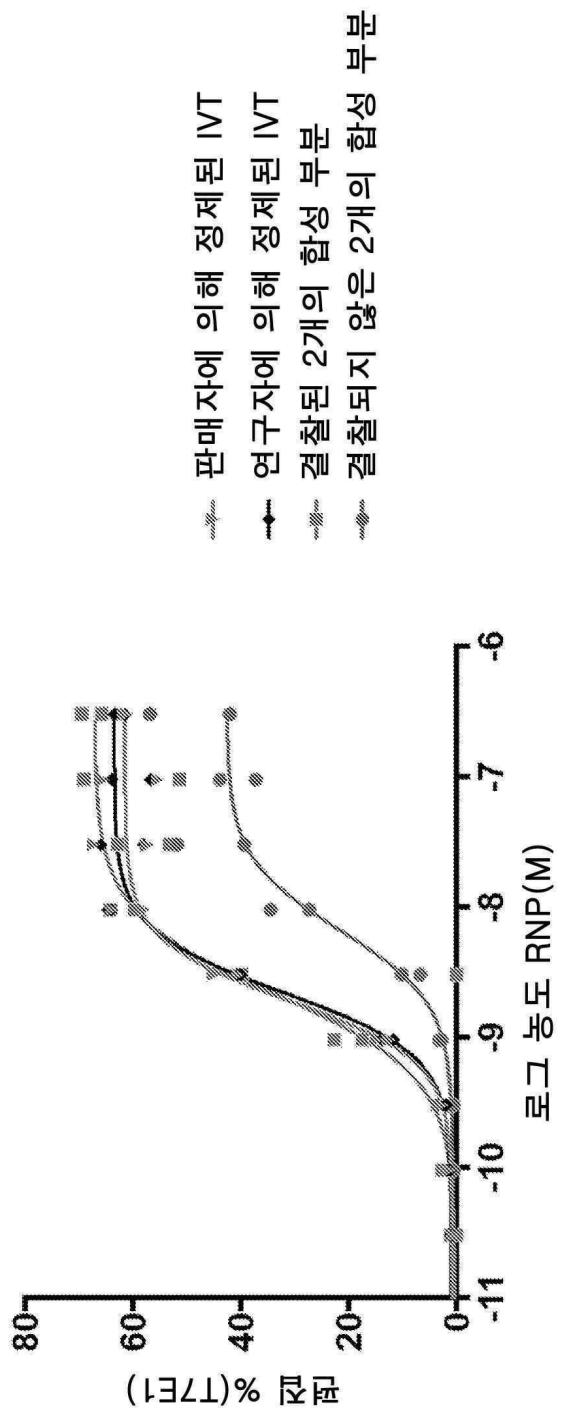
도면3c



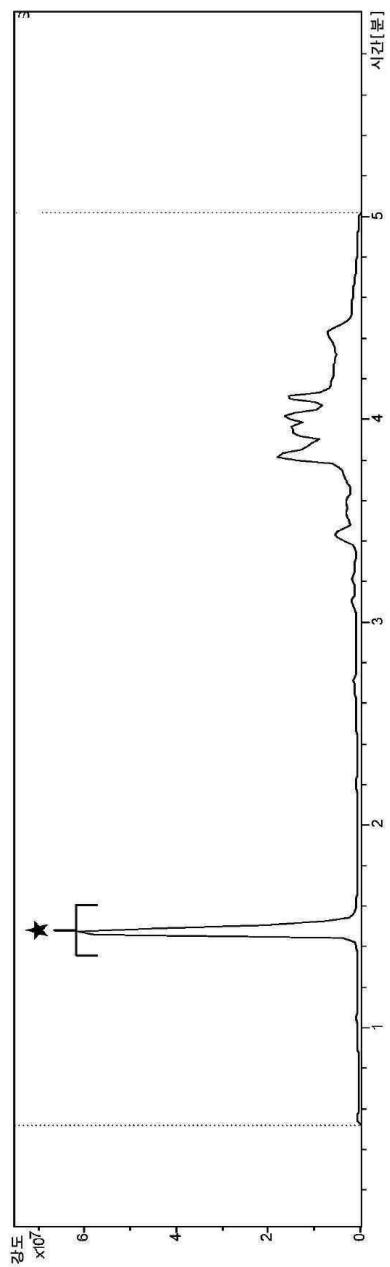
도면3d



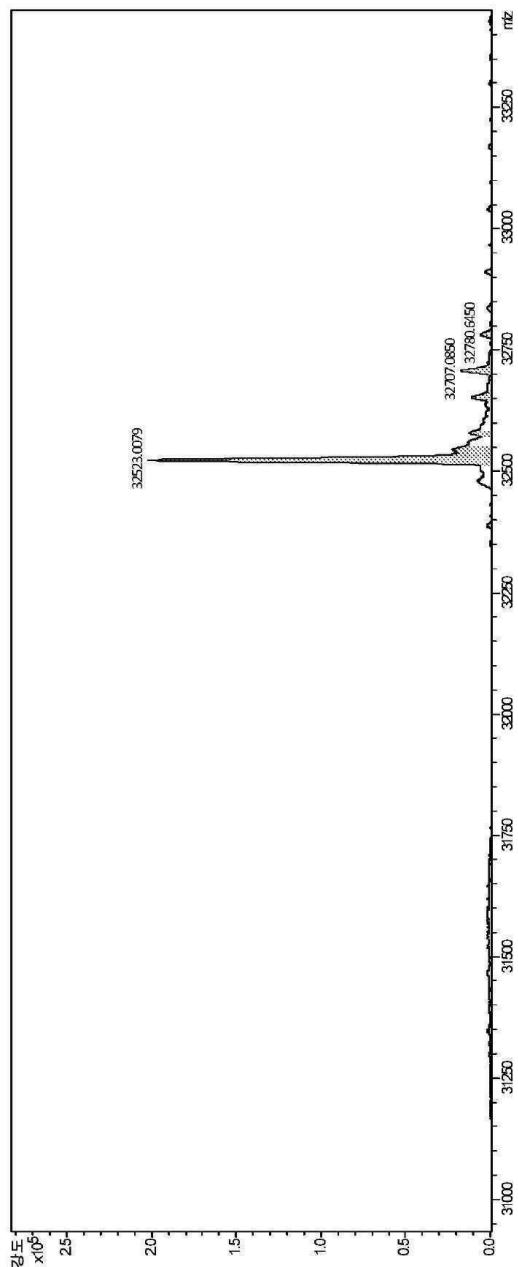
도면4



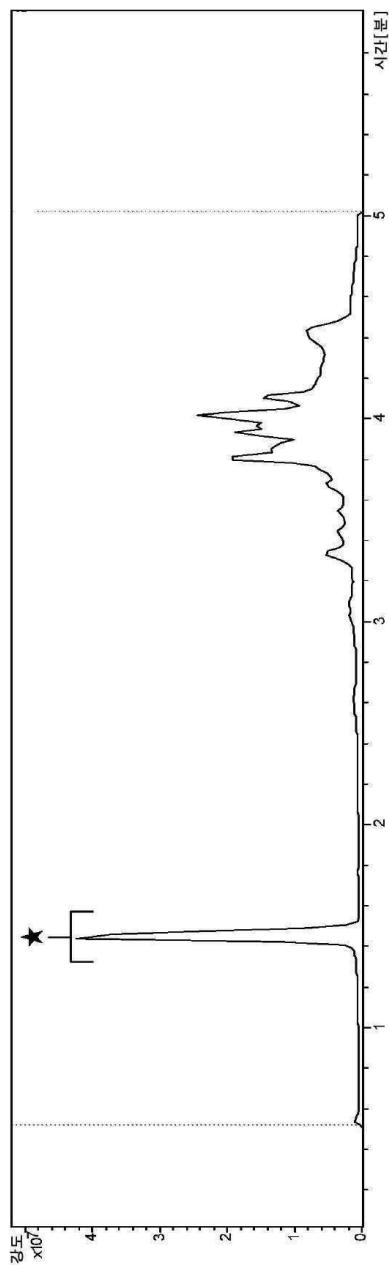
도면5a



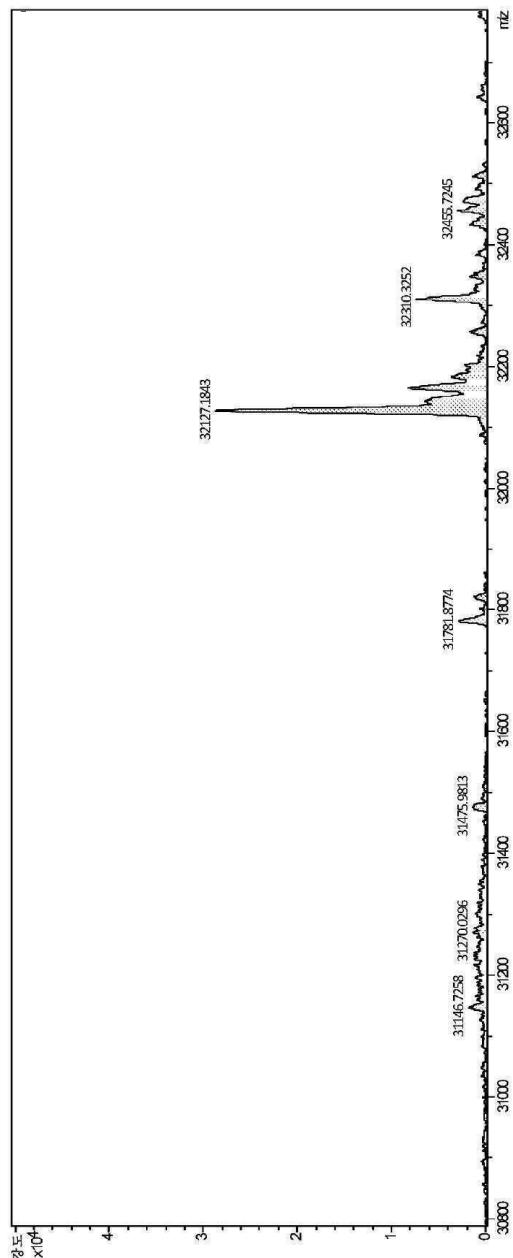
도면5b



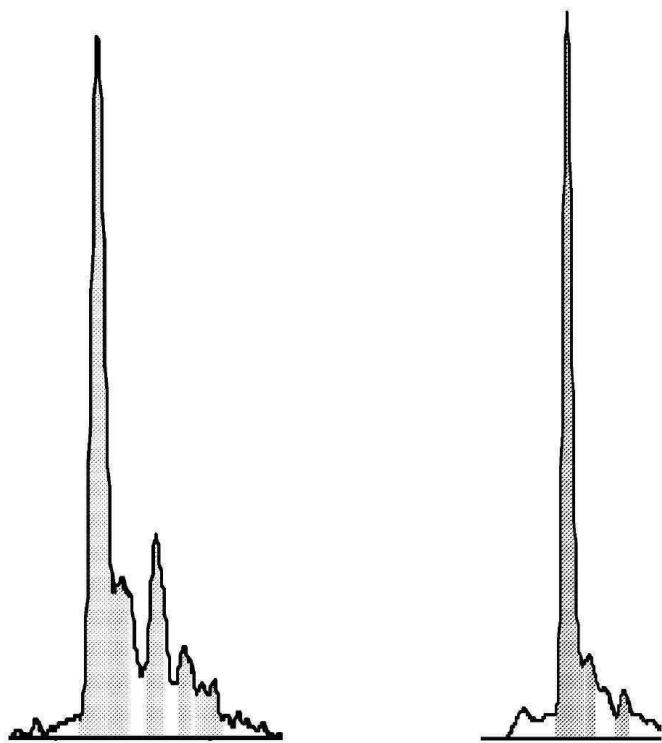
도면5c



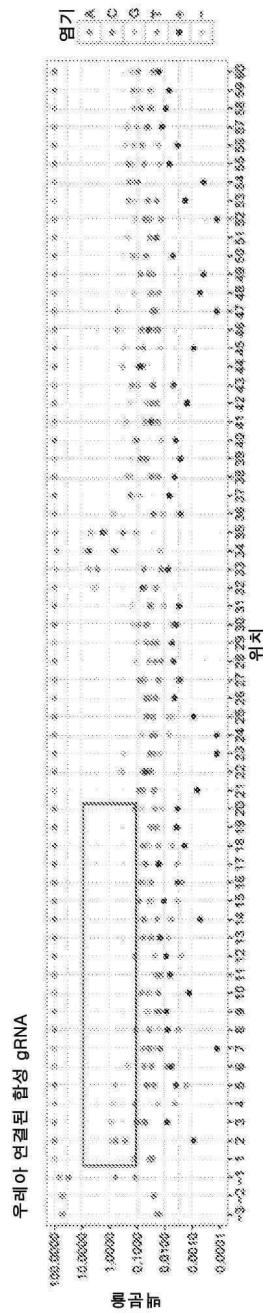
도면5d



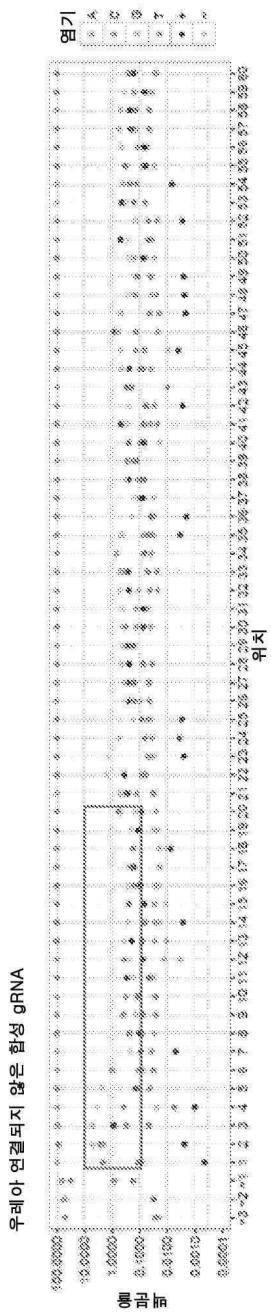
도면5e



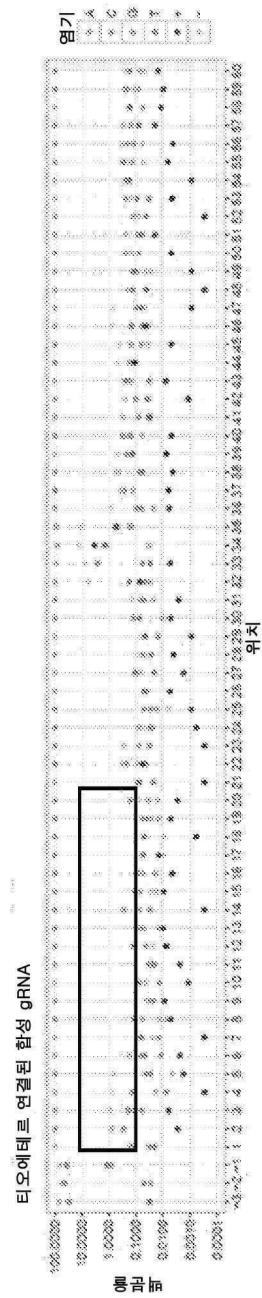
도면 6a



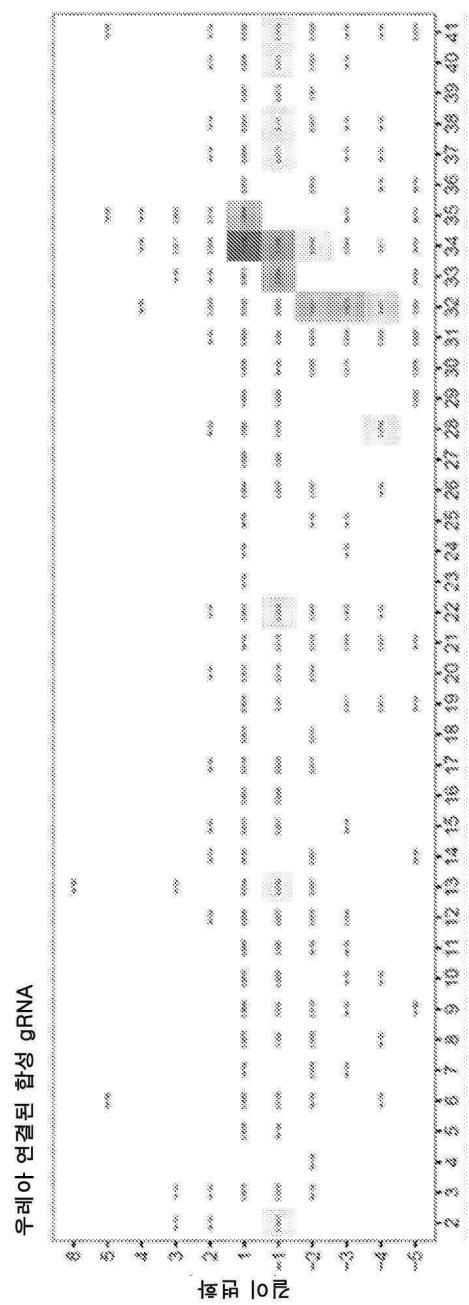
도면 6b



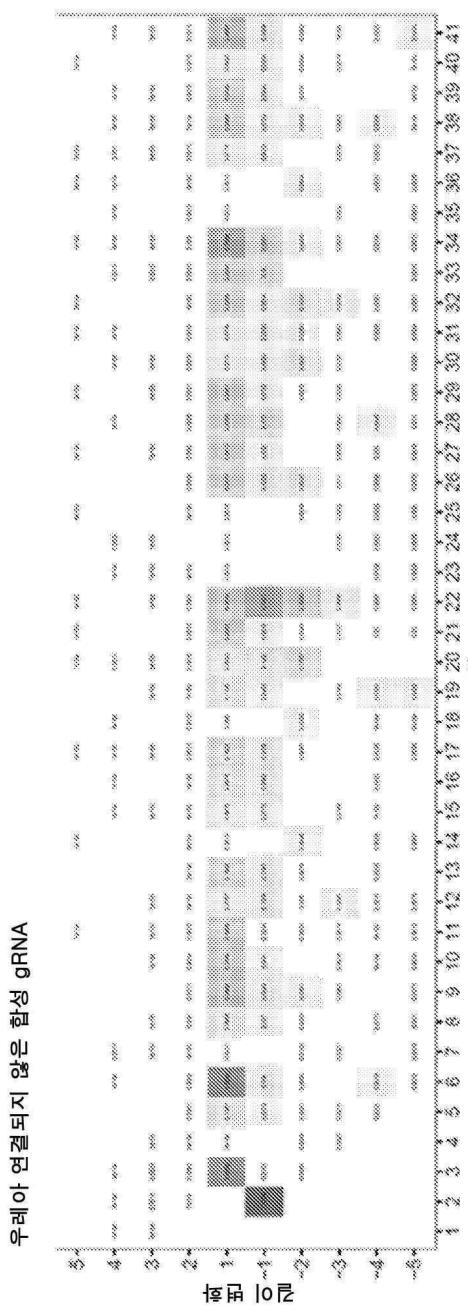
도면6c



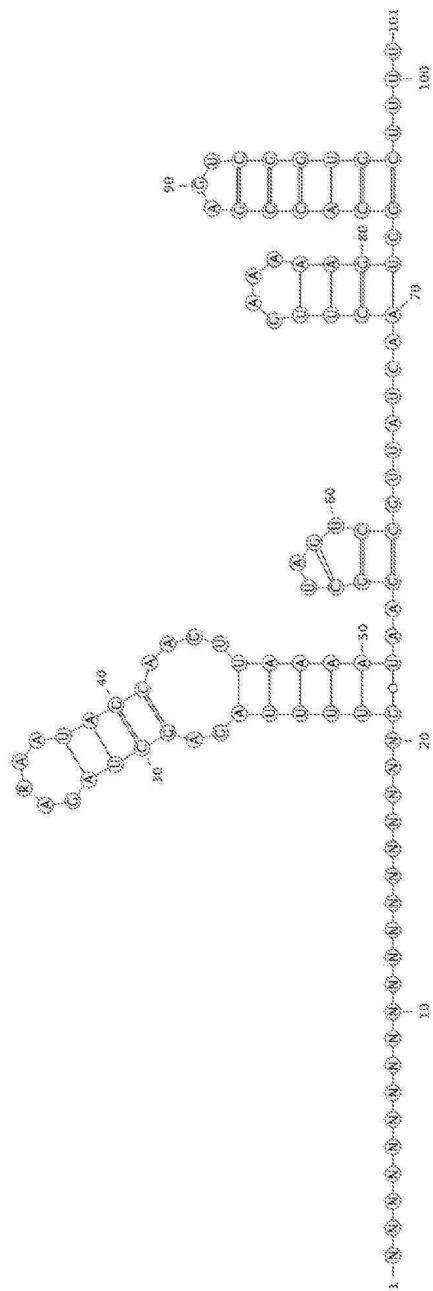
도면7a



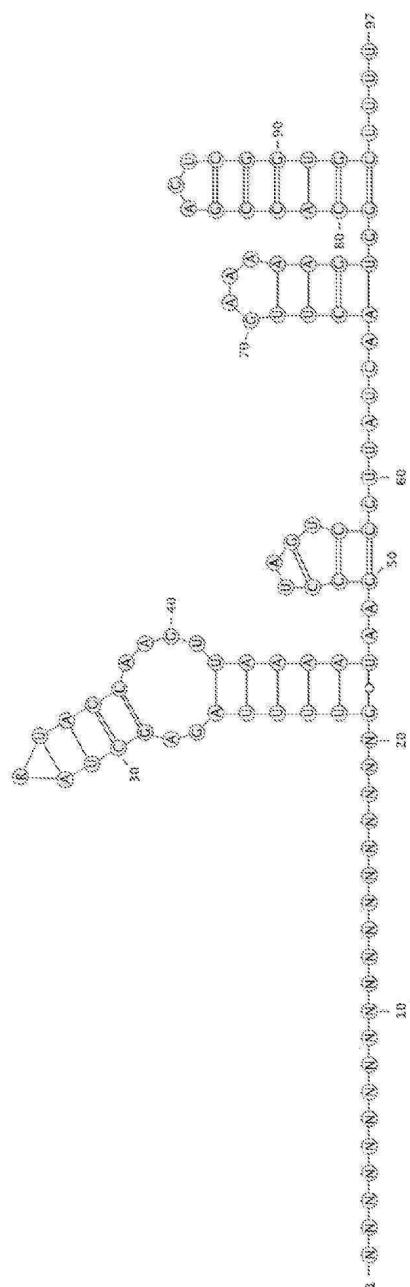
도면 7b



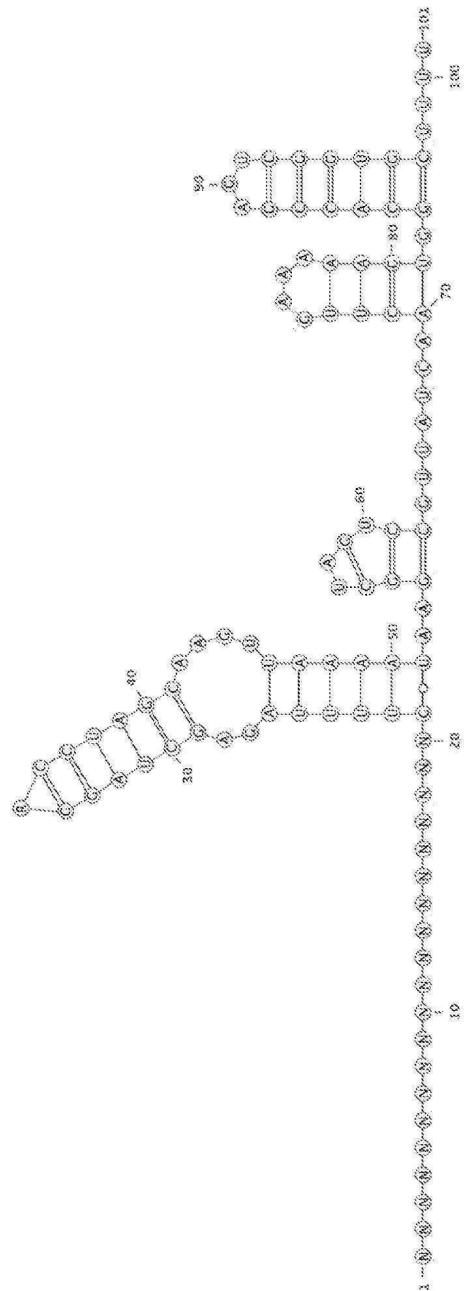
도면8a



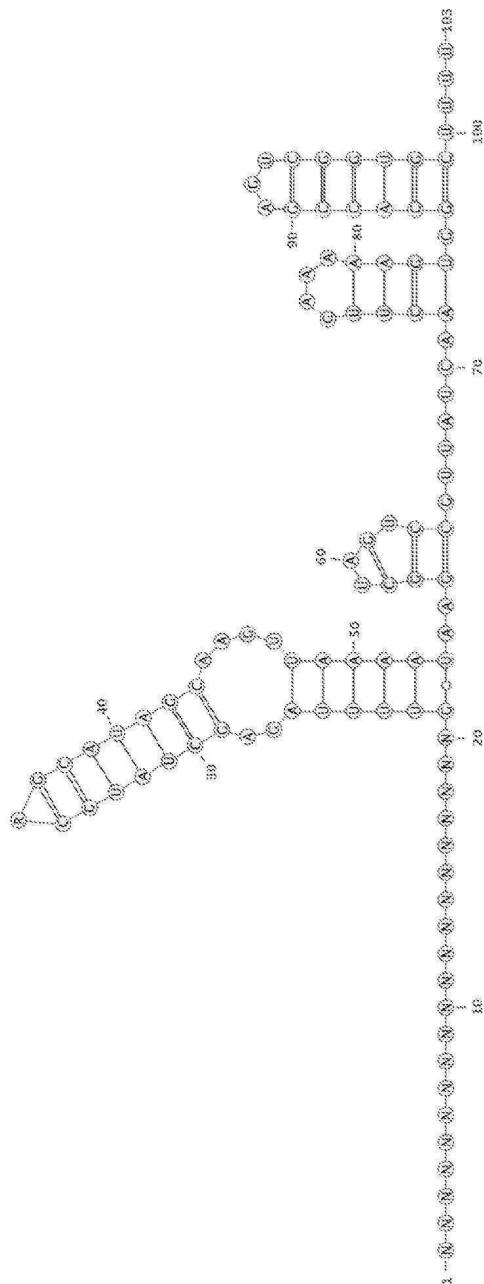
도면8b



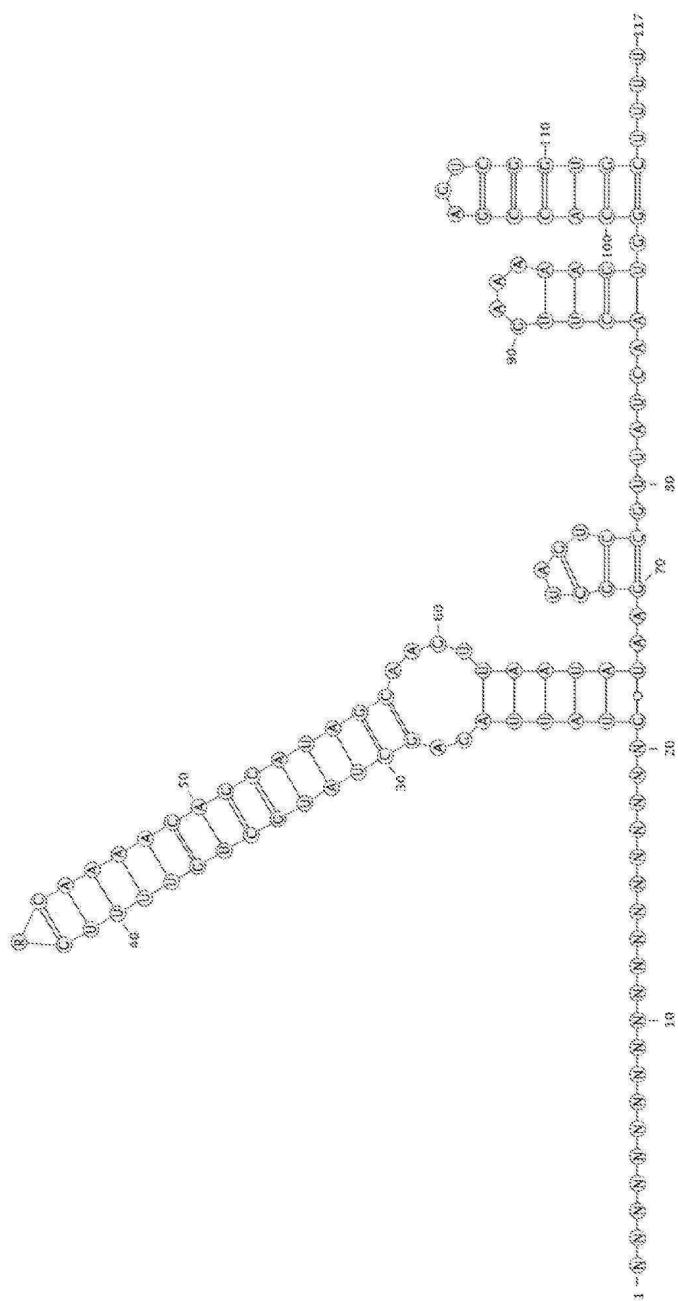
도면8c



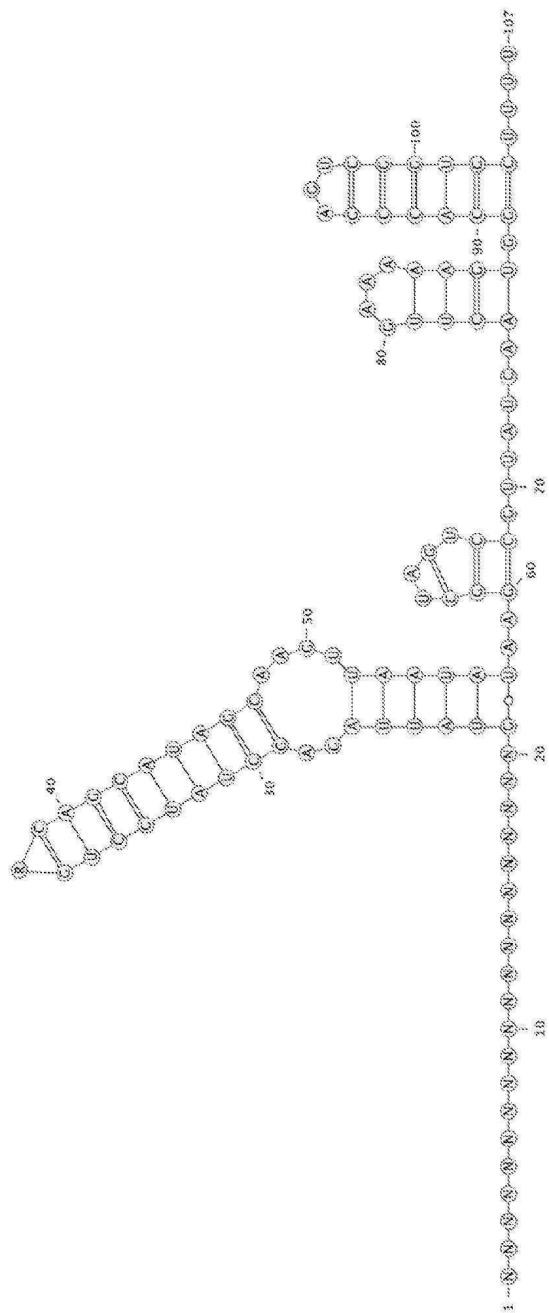
도면 8d



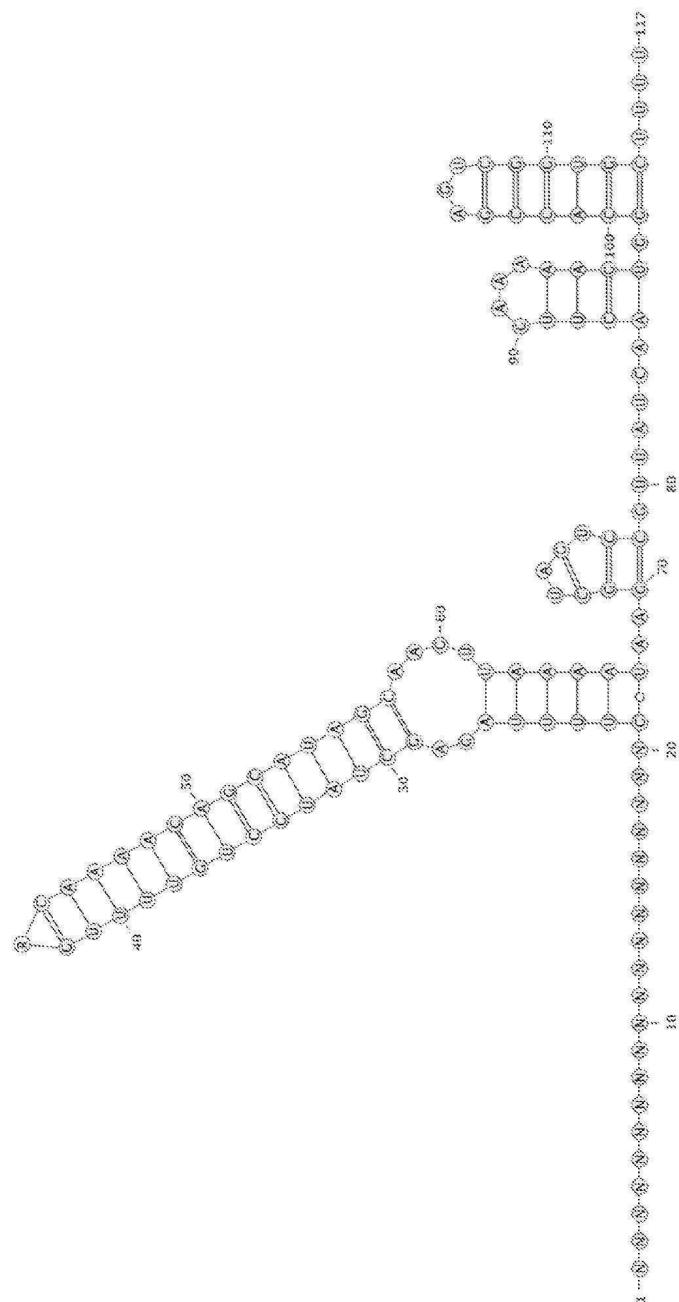
도면8e



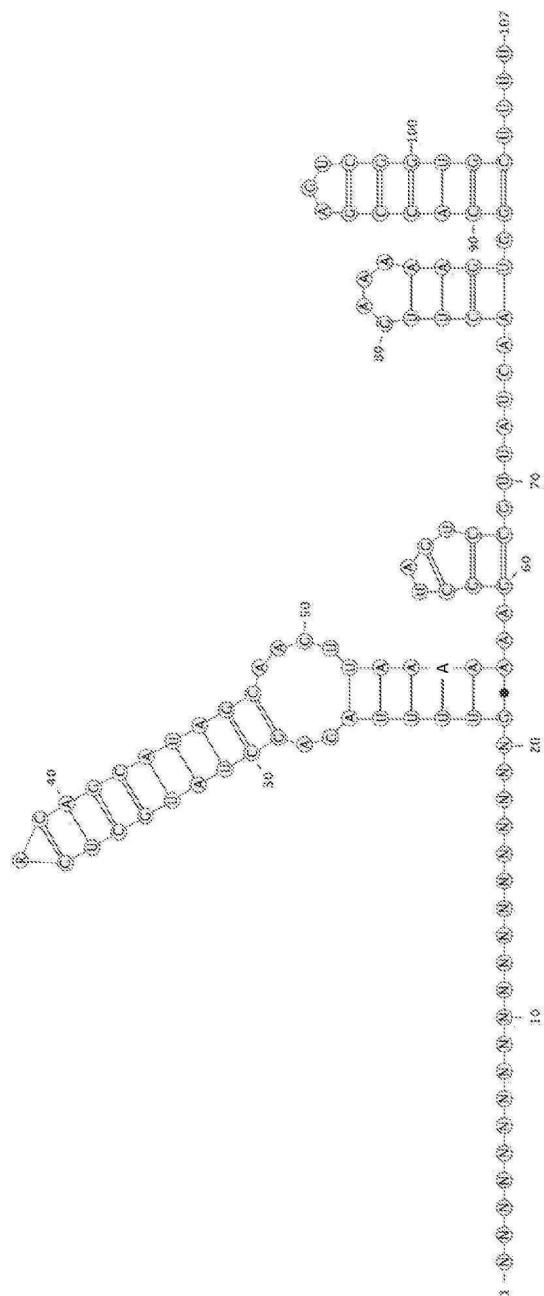
도면8f



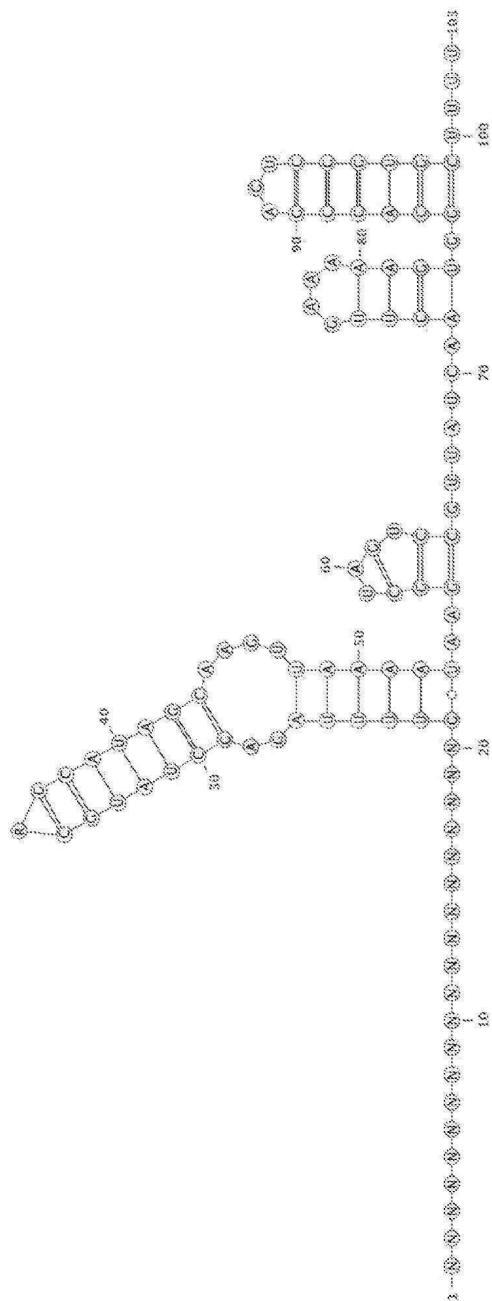
도면8g



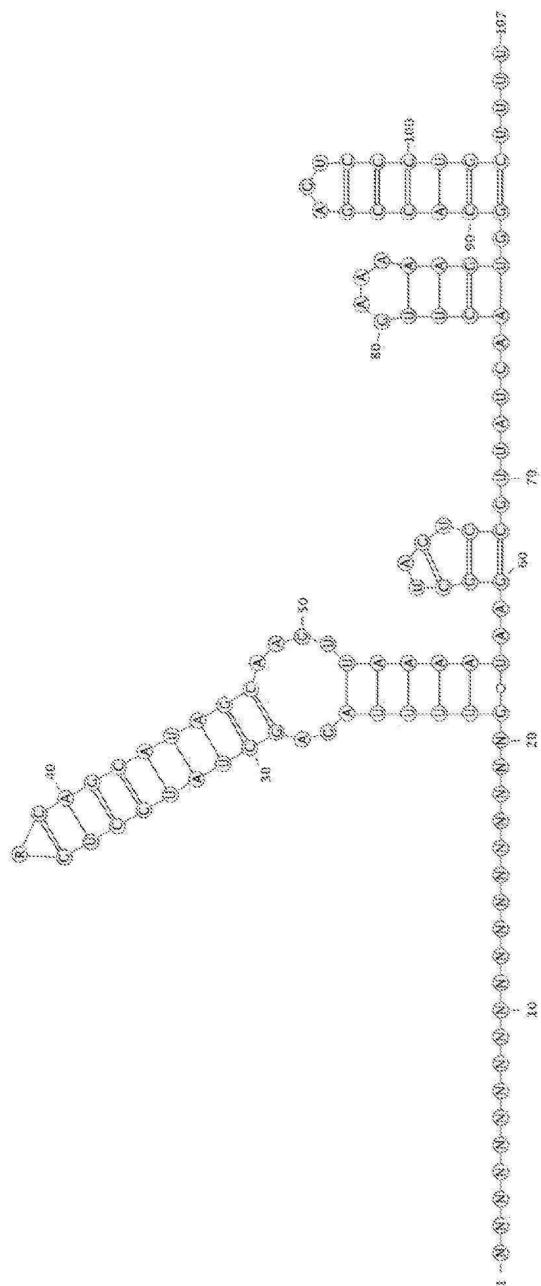
도면 8h



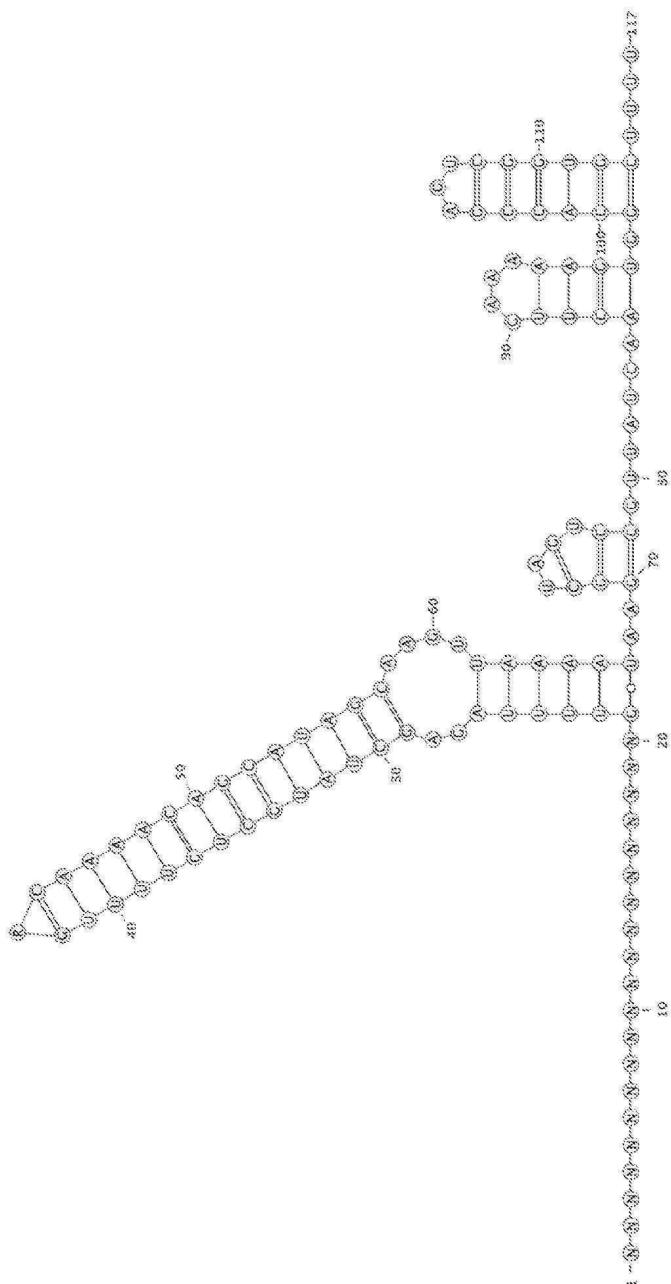
도면 9a



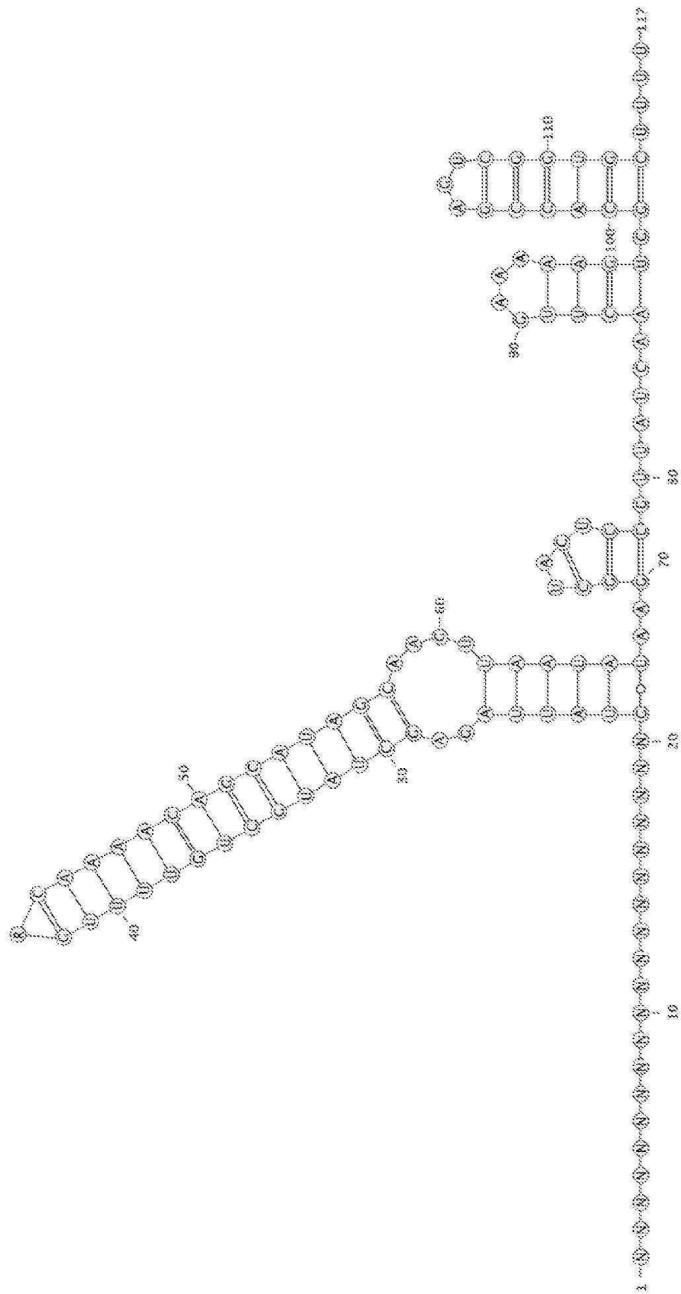
도면9b



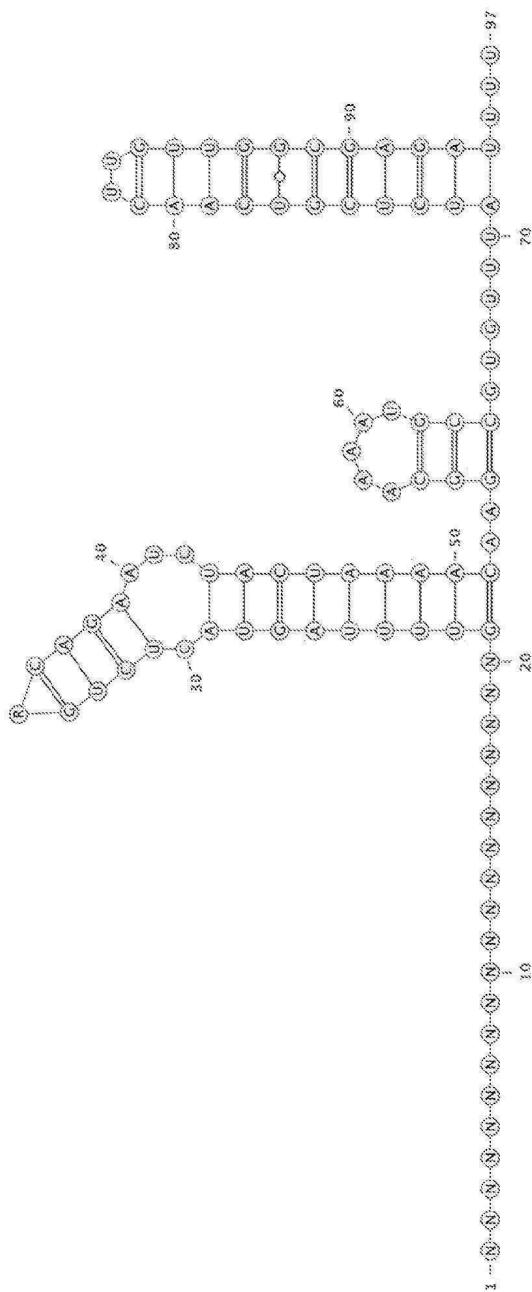
도면9c



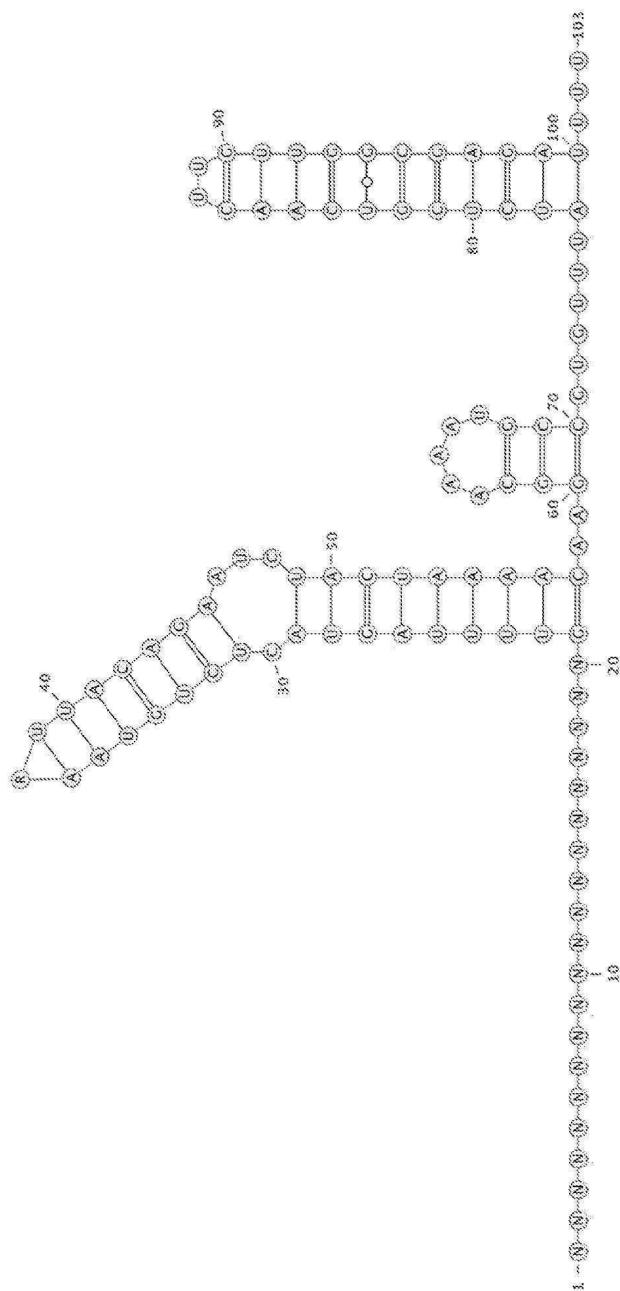
도면 9d



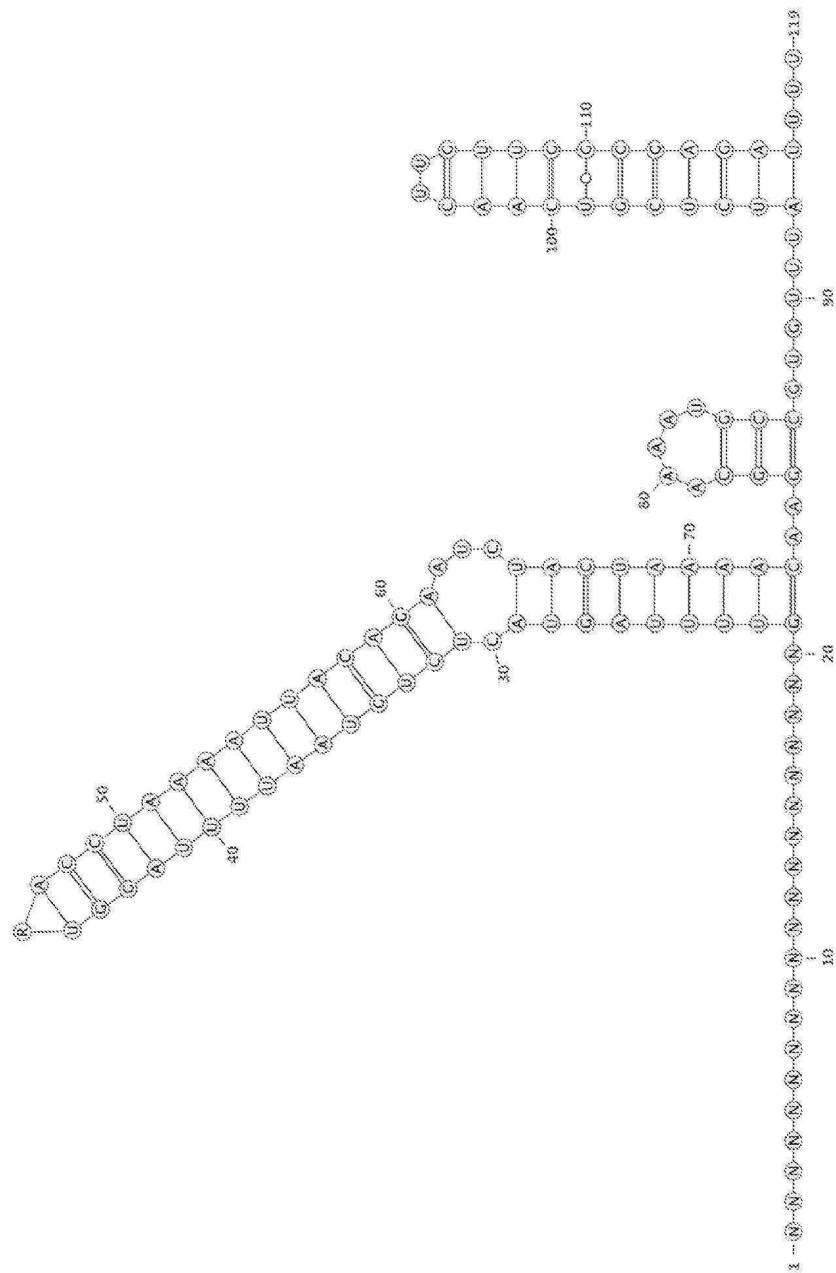
도면 10a



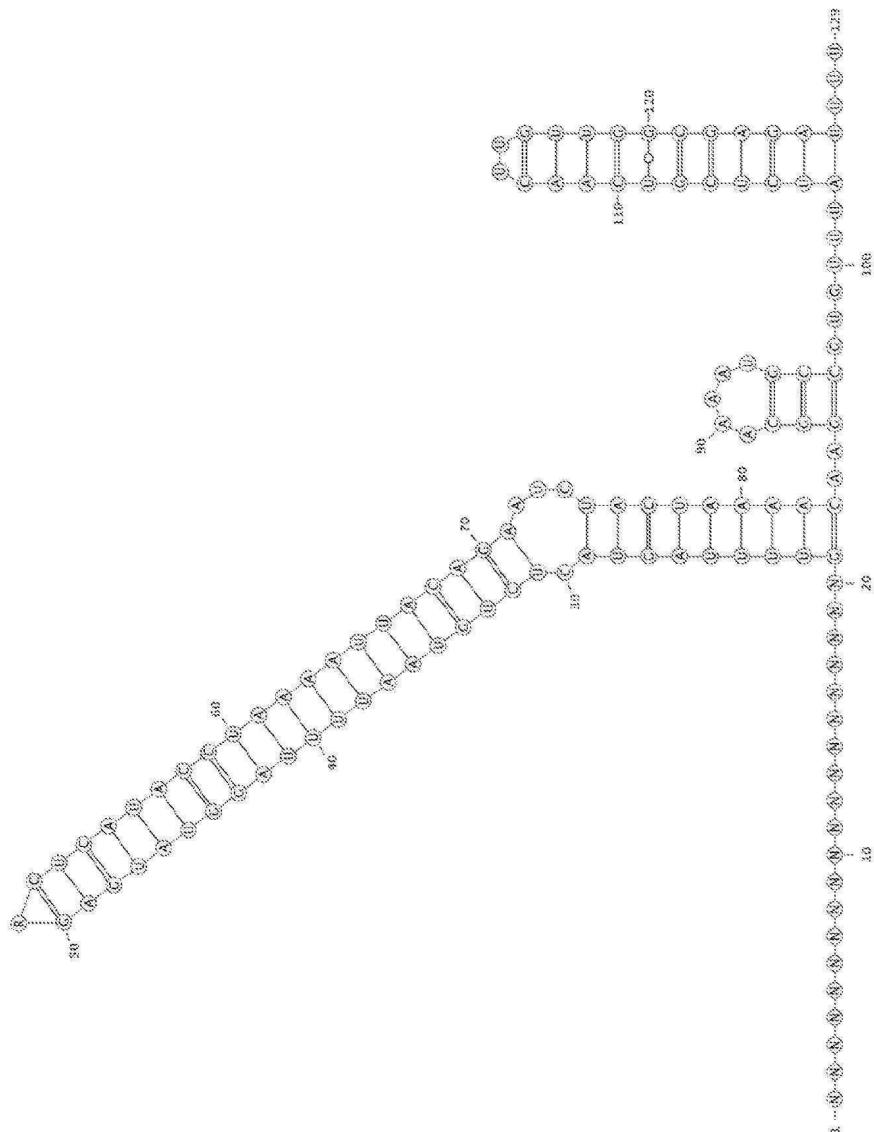
도면 10b



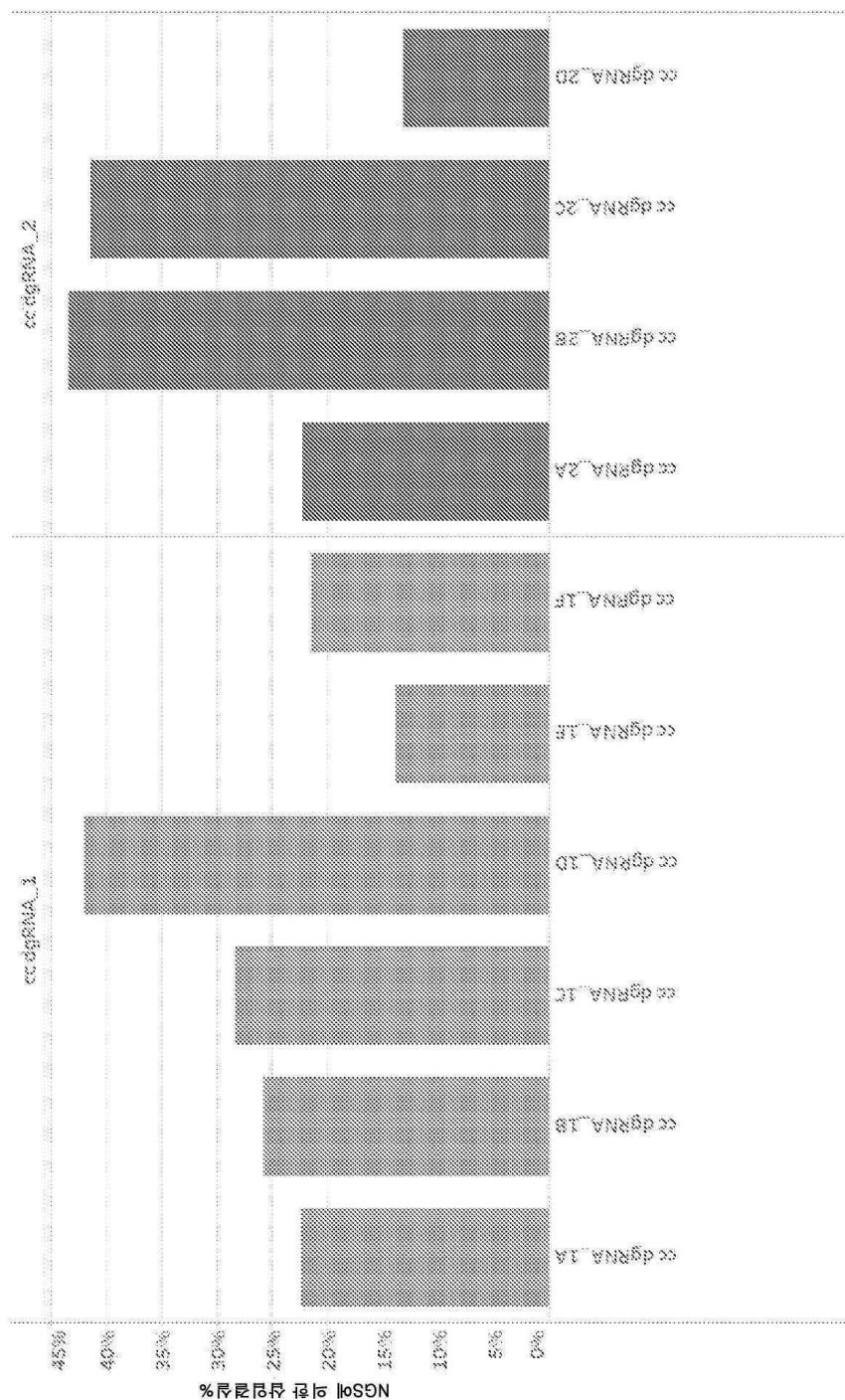
도면 10c



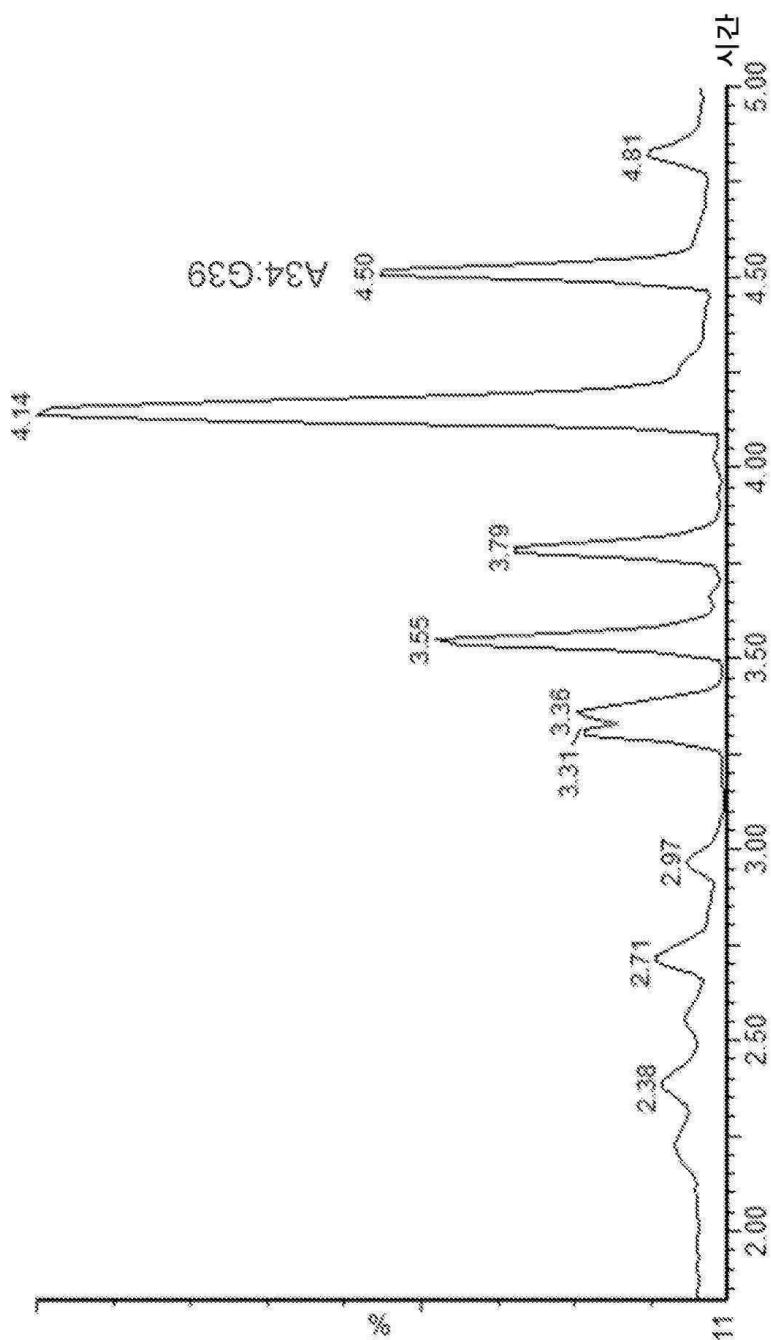
도면 10d



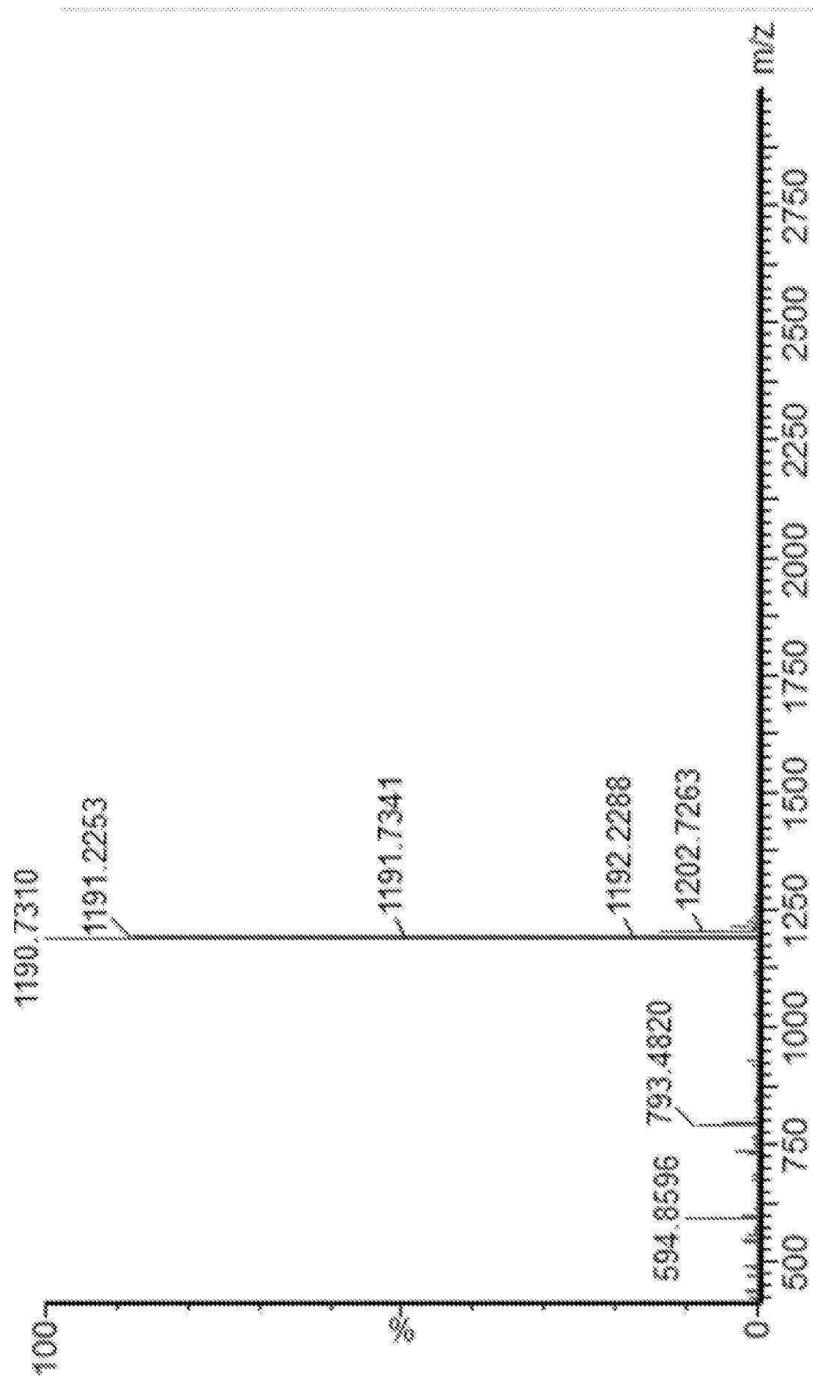
도면11



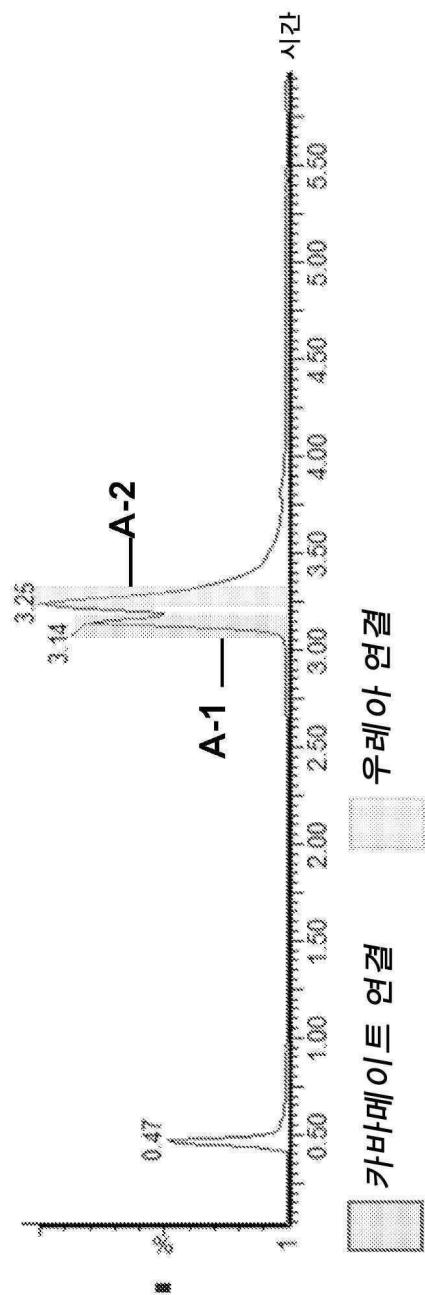
도면 12a



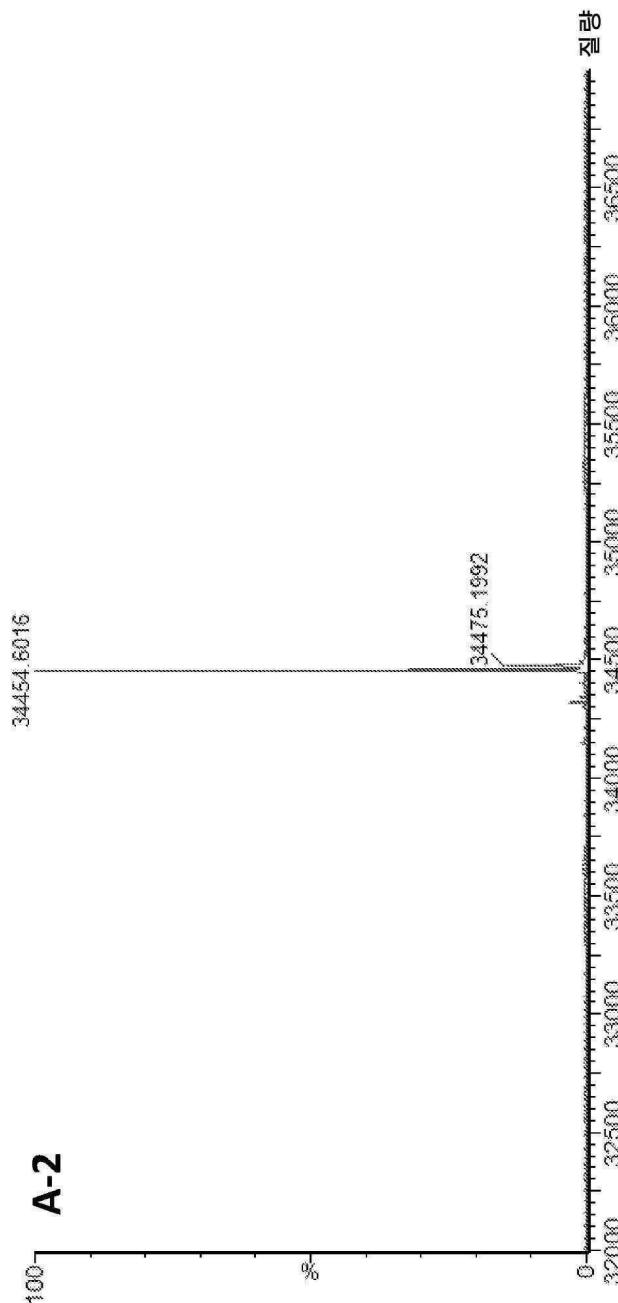
도면 12b



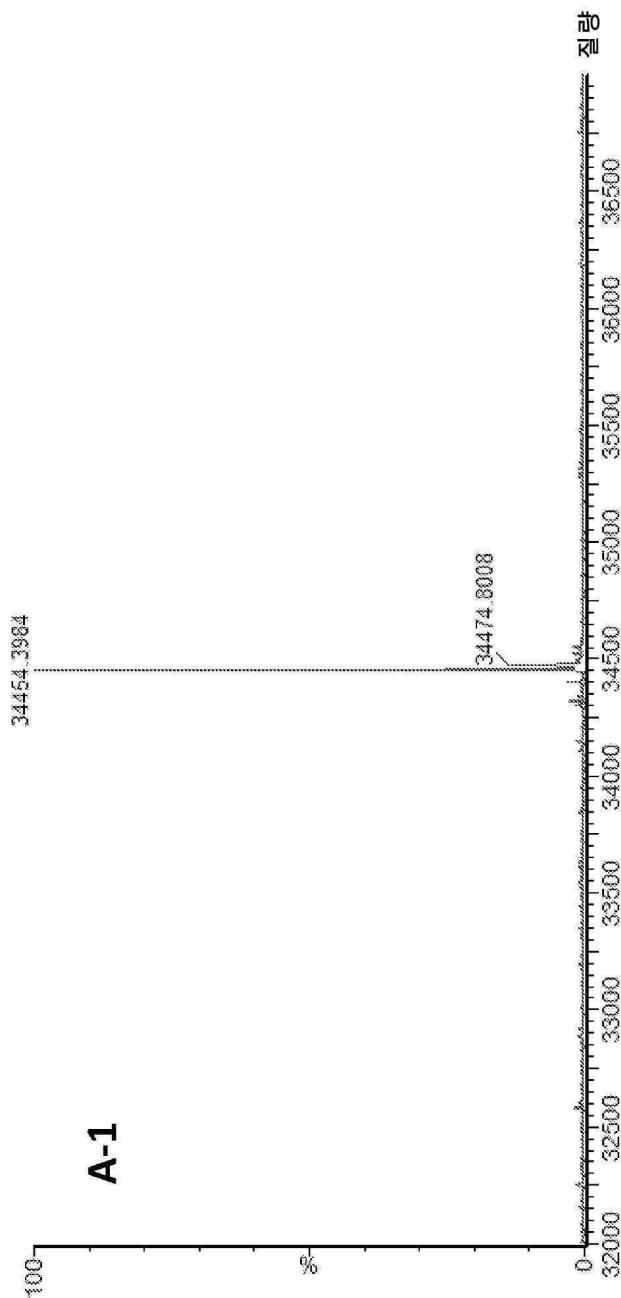
도면 13a



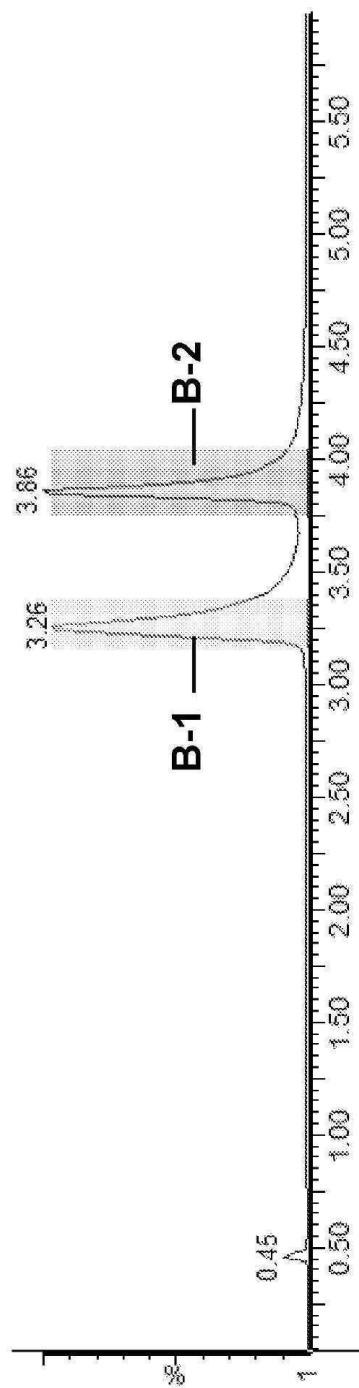
도면 13b



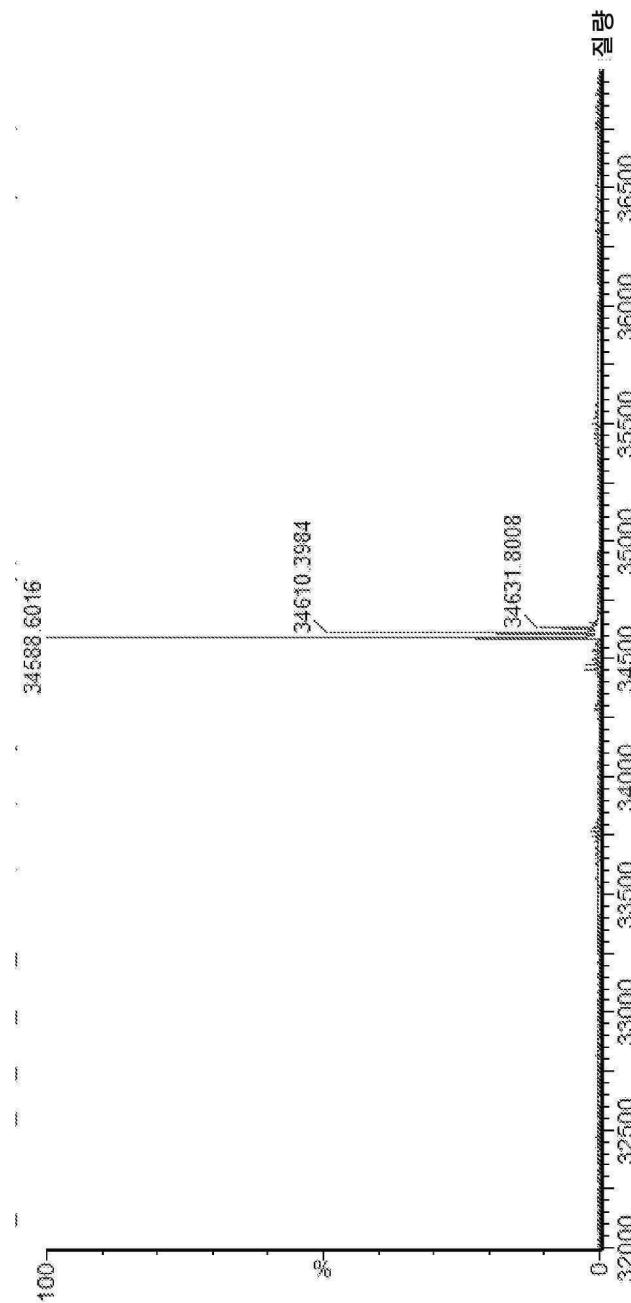
도면 13c



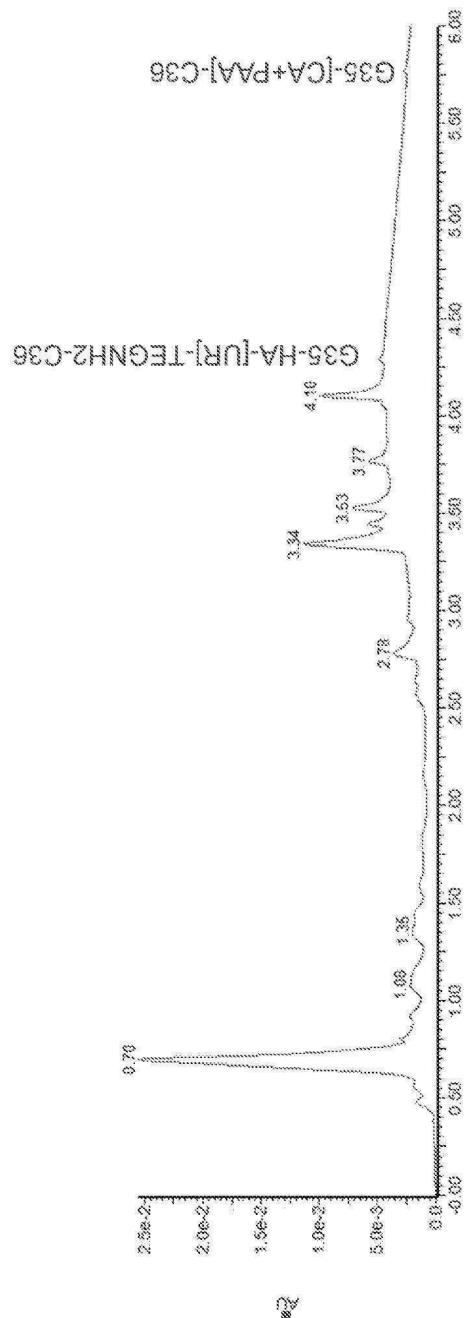
도면 14a



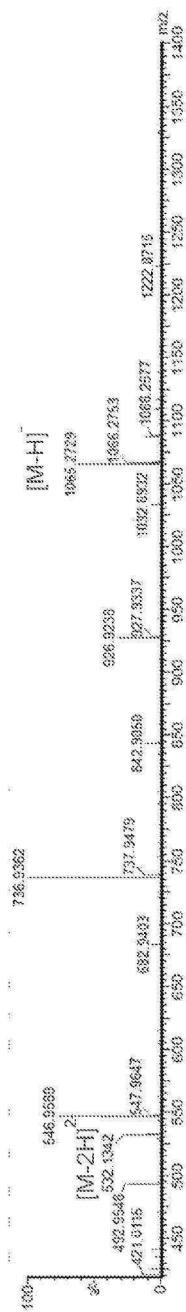
도면 14b



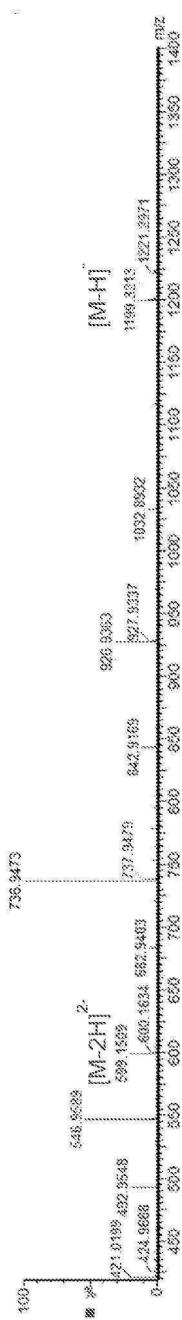
도면 15a



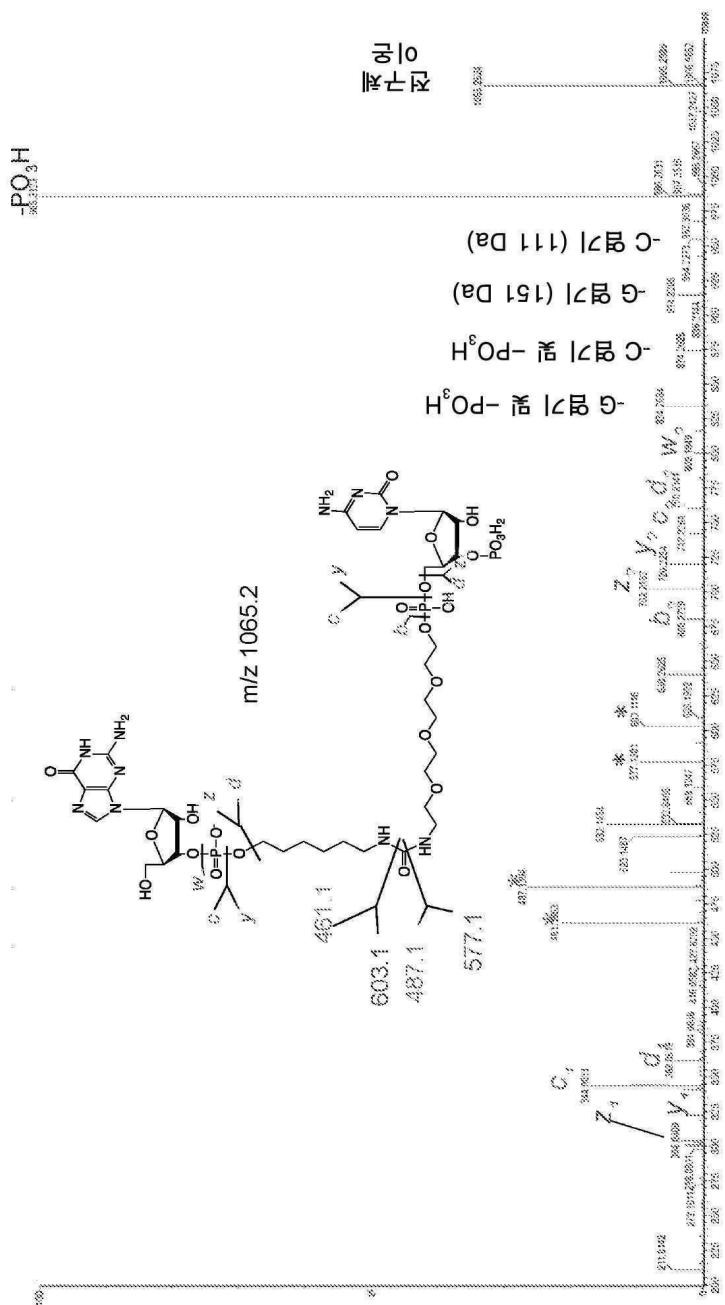
도면 15b



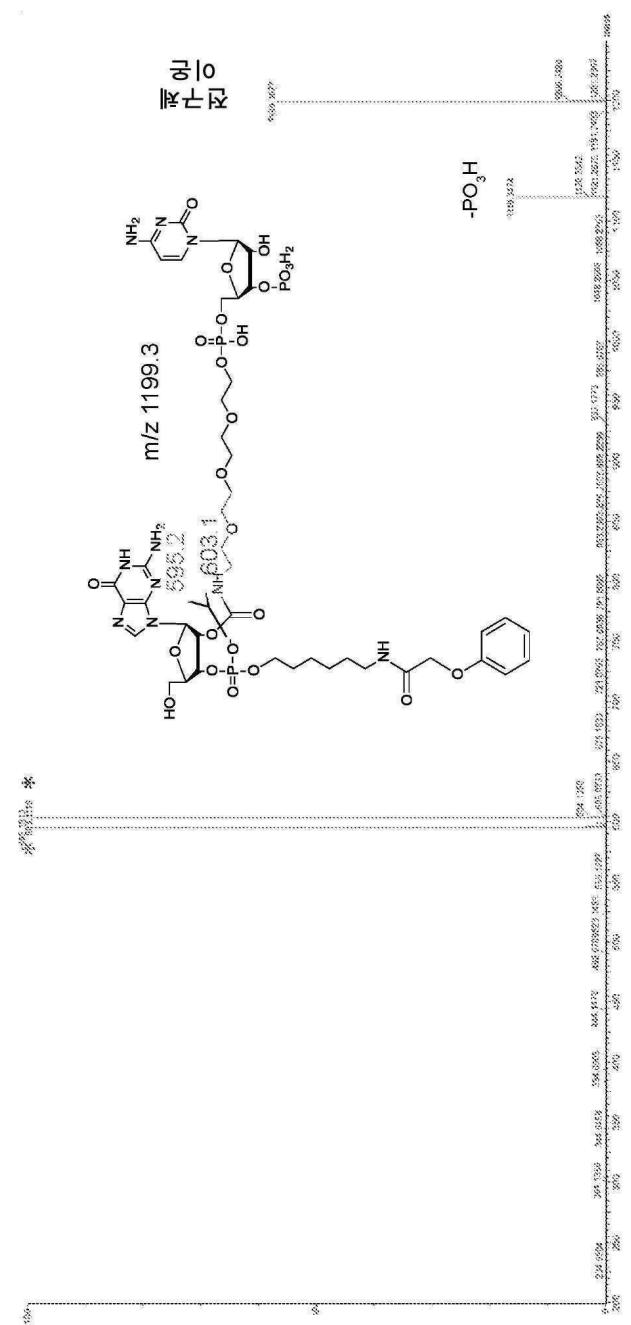
도면 15c



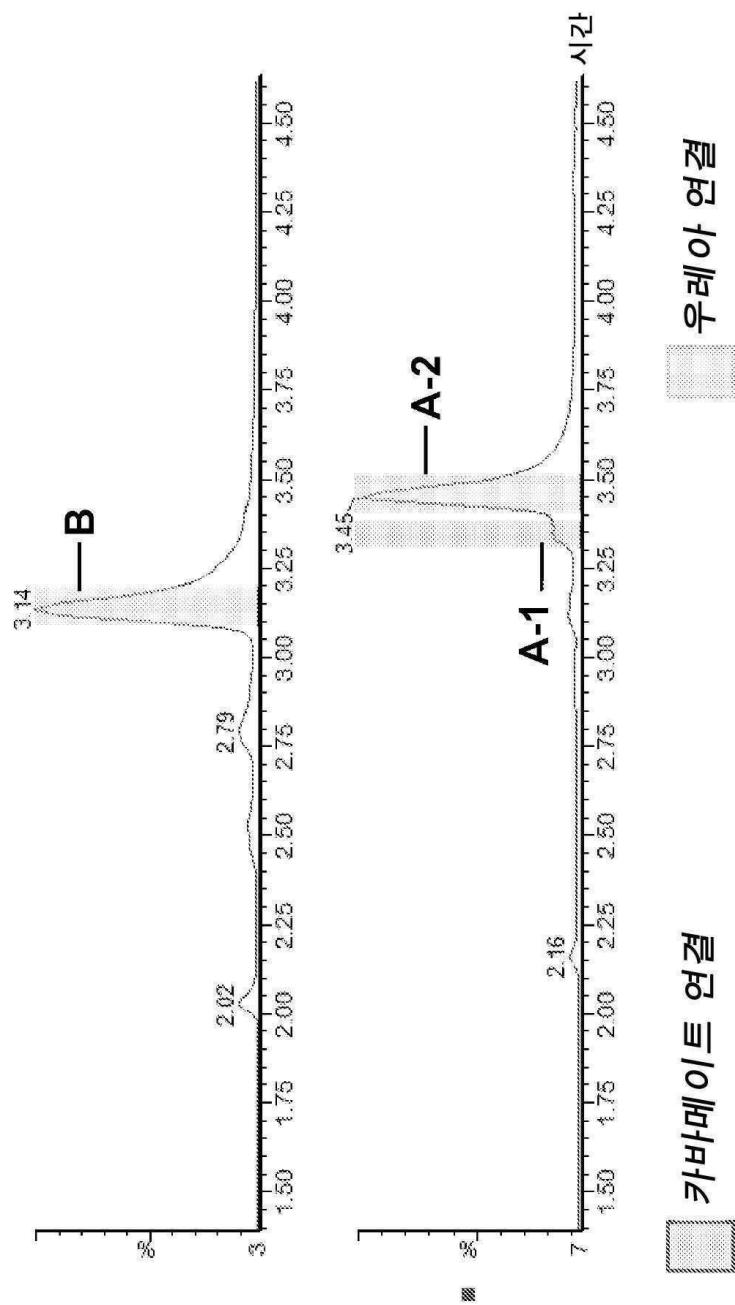
도면 15d



도면 15e



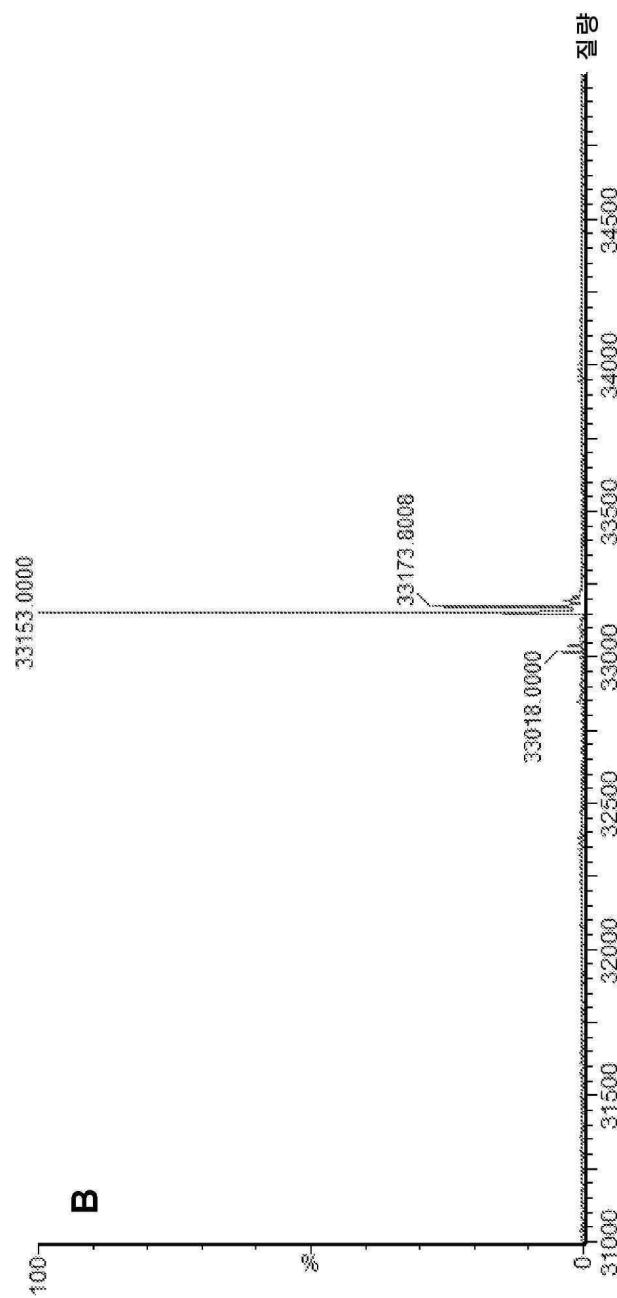
도면 16a



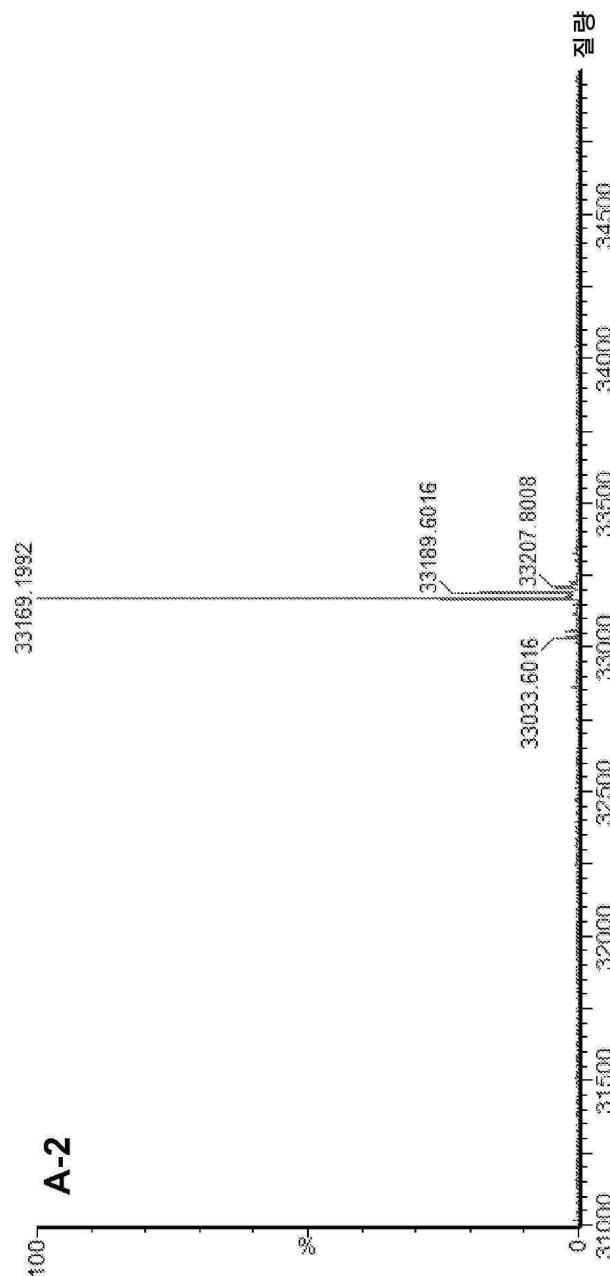
우레아 연결

카비메이트 연결

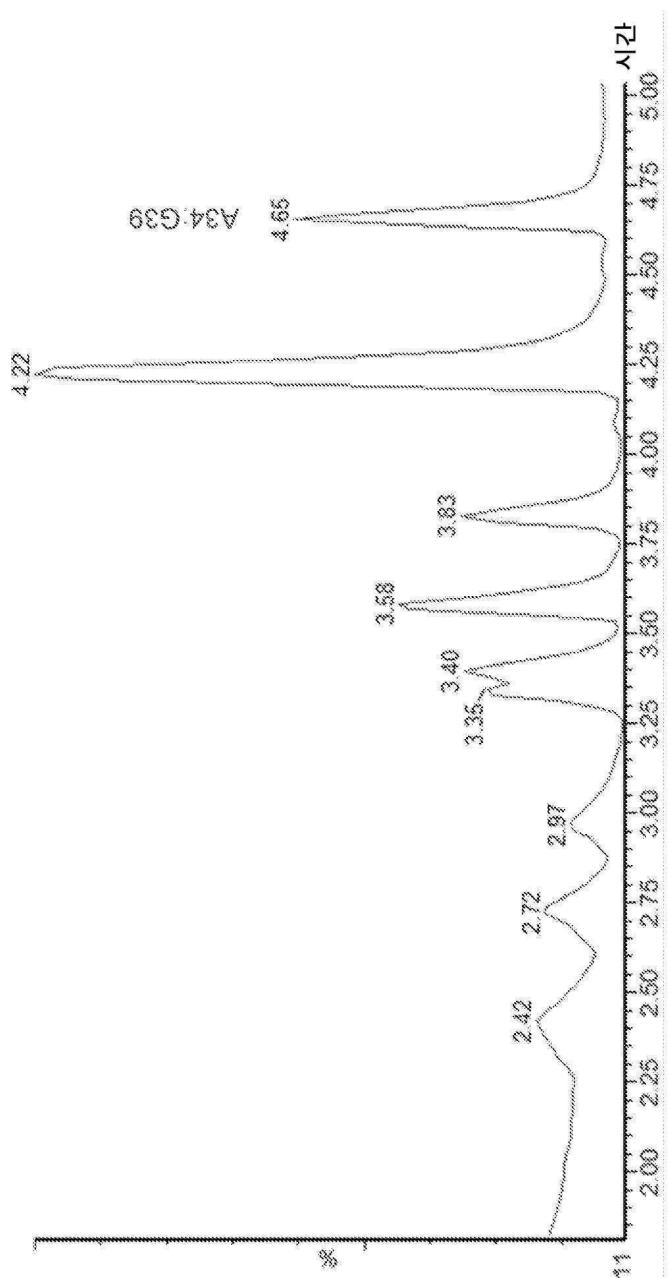
도면 16b



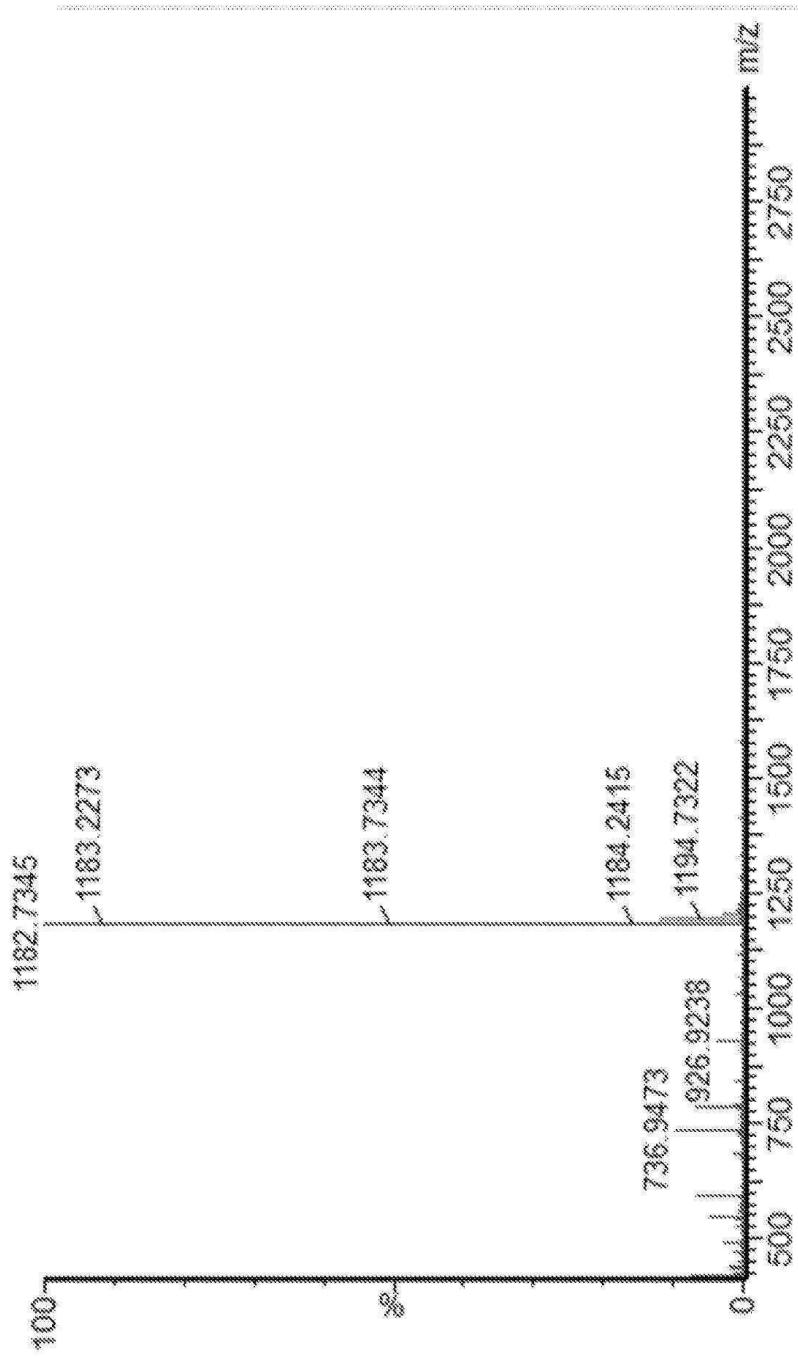
도면 16c



도면 17a



도면17b



서 열 목 록

<110> EDITAS MEDICINE, INC.

<120> SYNTHETIC GUIDE MOLECULES, COMPOSITIONS AND METHODS RELATING

THERETO

<130> 2011271-0066

<140> PCT/US2017/069019

<141> 2017-12-29

<150> 62/492,001
<151> 2017-04-28
<150> 62/441,046
<151> 2016-12-30
<160> 123
<170> KoPatentIn 3.0
<210> 1
<211> 13
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"
<400> 1

guuuuagagc uag 13

<210> 2
<211> 15
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 2
auagcaaguu aaaaau 15

<210> 3
<211> 11
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 3
guuuuagagc u 11

<210>

4

<211> 13

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 4

agcaaguuaa aau

13

<210> 5

<211> 13

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 5

guuuuagagc uag

13

<210> 6

<211> 15

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 6

cuagcaaguu aaaaau

15

<210> 7

<211> 14

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 7

guuuuagac uaug

14

<210> 8

<211>

16

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 8

cauagcaagu uaaaaau

16

<210> 9

<211> 21

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 9

guauuagagc uaugcuguuu u

21

<210> 10

<211> 23

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 10

aaaacagcau agcaaguuaa uau

23

<210> 11

<211> 16

<212> RNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 11

guauuagagc uaugcu 16

<210> 12

<211> 18

<

212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 12

agcauagcaa guuaauau 18

<210> 13

<211> 21

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 13

guuuuagagc uaugcuguuu u 21

<210> 14

<211> 23

<212

> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 14

aaaacagcau agcaaguuaa aa 23

<210> 15

<211> 16

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 15

guuuuagagc uaugcu

16

<210> 16

<211> 18

<212>

RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 16

agcauagcaa guaaaaaa

18

<210> 17

<211> 15

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 17

guuuuagagc uaaag

15

<210> 18

<211> 17

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 18

auuuagcaag uuaaaau

17

<210> 19

<211> 13

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 19

guuuuagagc uaa

13

<210> 20

<211> 15

<212> RNA

<

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 20

uuagcaaguu aaaaau

15

<210> 21

<211> 17

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 21

guuuuagagc uaaaggg

17

<210> 22

<211> 19

<212> RNA

<213

> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 22

accuuuagca aguuaaaaau	19
-----------------------	----

<210> 23

<211> 13

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 23

guuuuagagc uag	13
----------------	----

<210> 24

<211> 13

<212> RNA

<213>

Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 24

guuuuaguac ucu	13
----------------	----

<210> 25

<211> 15

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 25

agaaucuacu aaaaac	15
-------------------	----

<210> 26

<211> 16

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 26

guuuuaguac ucugua

16

<210> 27

<211> 18

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 27

uacagaaucu acuaaaac

18

<210> 28

<211> 24

<212> RNA

<213> Artificial Sequence

<

220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 28

guuuuaguac ucuguaauuu uagg

24

<210> 29

<211> 26

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 29

ccuaaaauua cagaaucuac uaaaac	26
<210> 30	
<211> 29	
<212> RNA	
<213> Artificial Sequence	
<220>	
><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 30	
guuuuaguac ucuguaauuu uagguauga	29
<210> 31	
<211> 31	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 31	
ucauaccuaa aaauacagaa ucuacuaaaa c	31
<210> 32	
<211> 100	
<212> RNA	
<213> Artificial Sequence	
<220>	
><221> source	
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"	
<220><221> modified_base	
<222> (1)..(20)	
<223> a, c, u, g, unknown or other	
<400> 32	
nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 33	

<211> 34

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 33

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaga 34

<210> 34

<211> 66

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 34

aauagcaagu uaaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucggu 60

gcuuuu 66

<210> 35

<211> 34

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 35

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaga 34

<210> 36
<211> 34
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 36
guaacggcag acuuucuccuc guuuuagagc uaga

34

<210> 37
<211> 32
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 37
guaacggcag acuuucuccuc guuuuagagc ua

32

<210> 38
<211> 34
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 38
guaacggcag acuuucuccuc guuuuagagc uagg

34

<
210> 39
<211> 35
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 39

guaacggcag acuuucuccuc guuuuagagc uaugc 35

<210> 40

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 40

guaacggcag acuuucuccuc guauuagagc uaugcuguuu ug 42

<210

> 41

<211> 37

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 41

guaacggcag acuuucuccuc guauuagagc uaugcug 37

<210> 42

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 42

guaacggcag acuuucuccuc guuuuagagc uaugcuguuu ug 42

<210>

43

<211> 37

<212> RNA

<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 43

guaacggcag acuuucuccuc guuuuagagc uaugcug 37

<210> 44

<211> 36

<212> RNA

<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 44

guaacggcag acuuucuccuc guuuuagagc uaaaga 36

<210> 45

<211> 34

<212> RNA

<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 45

guaacggcag acuuucuccuc guuuuagagc uaaa 34

<210> 46

<211> 38

<212> RNA

<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 46

guaacggcag acuuucuccuc guuuuagagc uaaaggga 38

<210> 47

<

<211> 35

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 47

guaacggcag acuuucuccuc guuuuagagc uagda 35

<210> 48

<211> 35

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 48

cuaacaguug cuuuuaucac guuuuagagc uaugc 35

<210> 49

<211

> 37

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 49

cuaacaguug cuuuuaucac guuuuagagc uaugcug 37

<210> 50

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 50
cuaacaguug cuuuuaucac guuuuagagc uaugcuguuu ug 42

<210> 51
<211>
42
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"
<400> 51
cuaacaguug cuuuuaucac guauuagagc uaugcuguuu ug 42

<210> 52
<211> 34
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"
<400> 52
cuaacaguug cuuuuaucac guuuuagagc uaga 34
<210> 53
<211> 34
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"
<400> 53
guaacggcag acuucuccuc guuuuaguac ucug 34
<210> 54
<211> 37
<212> RNA
<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 54

guaacggcag acuuucuccuc guuuuaguac ucuguaa 37

<210> 55

<211> 45

<

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 55

guaacggcag acuuucuccuc guuuuaguac ucuguaauuu uaggu 45

<210> 56

<211> 50

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 56

guaacggcag acuuucuccuc guuuuaguac ucuguaauuu uagguaugag 50

<210> 57

<211> 66

<212

> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 57

aauagcaagu uaaaaauaagg cuaguccguu auacaacuuga aaaaguggca ccgagucggu 60

gcuuuu 66

<210> 58
<211> 66
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 58
aauagcaagu uaaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucggu 60

gcuuuu 66
<210> 59
<211> 64
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 59
uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc gagucggugc 60
uuuu 64
<210> 60
<211> 66
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 60
ccuagcaagu uaaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucggu 60
gcuuuu 66
<210> 61
<211> 67
<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 61

gcauagcaag uuaaaauaag gcuaguccgu uaucaacuug aaaaaguggc accgagucgg	60
ugcuuuu	67

<210> 62

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 62

caaaacagca uagcaaguua auauaaggcu aguccguuau caacuugaaa aaguggcacc	60
gagucggugc uuuu	74

<210> 63

<211> 69

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 63

cagcauagca aguuaauaaua aggcuagucc guuaucacu ugaaaaagug gcaccgaguc	60
---	----

ggugcuuuu

69

<210> 64

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 64

caaaacagca uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc 60

gagucggugc uuuu 74

<210> 65

<211> 69

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 65

cagcauagca aguuaaaaaa aggcuagucc guuaucaacu ugaaaaagug gcaccgaguc 60

ggugcuuuu 69

<210> 66

<211> 68

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 66

aauuuagcaa guuaaaauaa ggcuaguccg uuaucaacuu gaaaaagugg caccgagucg 60

gugcuuuu 68

<210> 67

<211> 66

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 67

uuuagcaagu uaaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucgg 60

gcuuuu	66
<210> 68	
<211> 70	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 68	
accuuuagc aaguuaaaau aaggcuaguc cguuaucAAC uugaaaaAGU ggcaccgagu	60
cggugcuuuu	70
<210> 69	
<211> 66	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 69	
aauagcaagu uaaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucgg	60
gcuuuu	66
<210> 70	
<211> 67	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 70	
gcauagcaag uuaaaaauaag gcuaguccgu uaucacuug aaaaaguggc accgagucgg	60
ugcuuuu	67
<210> 71	
<211> 69	

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 71

cagcauagca aguuaaaaaua aggcuagucc guuaucacu ugaaaaagug gcaccgaguc	60
ggugcuuuu	69

<210> 72

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 72

caaaacagca uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc	60
gagucggugc uuuu	74

<210> 73

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 73

caaaacagca uagcaaguua auauaaggcu aguccguuau caacuugaaa aaguggcacc	60
---	----

gagucggugc uuuu

74

<210> 74

<211> 66

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 74

aauagcaagu	aaaaauaagg	cuaguccguu	auacaacuuga	aaaaguggca	ccgagucggu	60
gcuuuu						66

<210> 75

<211> 62

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 75

cagaaucuac	aaaaacaagg	caaaaugccg	uguuuuaucuc	gucaacuugu	uggcgagauu	60
uu						62

<210> 76

<211> 65

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 76

uuacagaauc	uacuaaaaca	aggcaaaaug	ccguguuuau	cucgucaacu	uguuggcag	60
auuuu						65

<210> 77

<211> 73

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 77

accuaaaaauu acagaaucua cuaaaacaag gcaaaugcc guguuaucu cgucaacuug	60
uuggcgagau uuu	73
<210> 78	
<211> 78	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 78	
cucauaccua aaauuacaga aucuacuaaa acaaggcaa augccgugu uaucucguca	60
acuuguuggc gagauuuu	78
<210> 79	
<211> 54	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> modified_base	
<222> (1)..(15)	
<223> a, c, u, g, unknown or other	
<220><221> modified_base	
<222> (27)..(48)	
<223> a, c, u, g, unknown or other	
<220><221> misc_feature	
<222> (27)..(48)	
<223> /note="This region may encompass 2-22 nucleotides"	
<220><221> modified_base	
<222> (49)..(54)	
<223> a, c, u, g, unknown or other	
<220><221> misc_feature	
<222> (49)..(54)	

<223> /note="This region may encompass 0-6 nucleotides"

<220>

<221> source

<223> /note="See specification as filed for detailed description of substitutions and preferred embodiments"

<400> 79

nnnnnnnnnn nnnnnuuua gagcuannn nnnnnnnnnn nnnnnnnnn nnnn	54
--	----

<210> 80

<211> 71

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(6)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (1)..(6)

<223> /note="This region may encompass 0-6 nucleotides"

<220><221> modified_base

<222> (7)..(28)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (7)..(28)

<223> /note="This region may encompass 2-22 nucleotides"

<220><221> modified_base

<222> (42)..(71)

<223> a, c, u, g, unknown or other

<220>

<221> source

<223> /note="See specification as filed for detailed description of substitutions and preferred embodiments"

<400> 80	
nnnnnnnnnn nnnnnnnnnn nnnnnnnnuu gcaaguuaaa annnnnnnnnn nnnnnnnnnn	60
nnnnnnnnnn n	71
<210> 81	
<211> 54	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> modified_base	
<222> (1)..(15)	
<223> a, c, u, g, unknown or other	
<220><221> modified_base	
<222> (27)..(48)	
<223	
> a, c, u, g, unknown or other	
<220><221> misc_feature	
<222> (27)..(48)	
<223> /note="This region may encompass 2-22 nucleotides"	
<220><221> modified_base	
<222> (49)..(54)	
<223> a, c, u, g, unknown or other	
<220><221> misc_feature	
<222> (49)..(54)	
<223> /note="This region may encompass 0-6 nucleotides"	
<220>	
<221> source	
<223> /note="See specification as filed for detailed description of substitutions and preferred embodiments"	
<400> 81	
nnnnnnnnnn nnnnnuauua gagcuannnn nnnnnnnnnn nnnnnnnnnn nnnn	54
<210> 82	

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(6)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (1)..(6)

<223> /note="This region may encompass 0-6 nucleotides"

<220><221> modified_base

<222> (7)..(28)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (7)..(28)

<223> /note="This region may encompass 2-22 nucleotides"

<220><221> modified_base

<222> (42)..(71)

<223> a, c, u, g, unknown or other

<220>

<221> source

<223> /note="See specification as filed for detailed description of substitutions and preferred embodiments"

<400> 82

nnnnnnnnnn nnnnnnnnnn nnnnnnnnua gcaaguuaau annnnnnnnnn nnnnnnnnnn	60
nnnnnnnnnnn n	71

<210> 83

<211> 57

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(15)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (30)..(51)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (30)..(51)

<223> /note="This region may encompass 2-22 nucleotides"

<220><221> modified_base

<222> (52)..(57)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (52)..(57)

<223> /note="This region may encompass 0-6 nucleotides"

<220>

<221> source

<223> /note="See specification as filed for detailed description of

substitutions and preferred embodiments"

<400> 83

nnnnnnnnnn nnnnnguuu aguacucugn nnnnnnnnnn nnnnnnnnn nnnnnnn	57
--	----

<210> 84

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(6)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (1)..(6)

<223> /note="This region may encompass 0-6 nucleotides"

<220><221> modified_base

<222> (7)..(28)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (7)..(28)

<223> /note="This region may encompass 2-22 nucleotides"

<220><221> modified_base

<222> (45)..(74)

<223> a, c, u, g, unknown or other

<220>

<221> source

<223> /note="See specification as filed for detailed description of

substitutions and preferred embodiments"

<400> 84

nnnnnnnnnn nnnnnnnnnn nnnnnnnnca gaaucuacua aaacnnnnn nnnnnnnnnn	60
--	----

nnnnnnnnnn nnnn	74
-----------------	----

<210> 85

<211> 57

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(15)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (30)..(51)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (30)..(51)

<223> /note="This region may encompass 2-22 nucleotides"

<220><221> modified_base
<222> (52)..(57)
<223> a, c, u, g, unknown or other

<220><221> misc_feature
<222> (52)..(57)
<223> /note="This region may encompass 0-6 nucleotides"
<220>
<221> source
<223> /note="See specification as filed for detailed description of substitutions and preferred embodiments"

<400> 85

nnnnnnnnnn nnnnnnguauu aguacucugn nnnnnnnnnn nnnnnnnnnn nnnnnnnn 57

<210> 86

<211> 74

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(6)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (1)..(6)

<223> /note="This region may encompass 0-6 nucleotides"

<220><221> modified_base

<222> (7)..(28)

<223> a, c, u, g, unknown or other

<220><221> misc_feature

<222> (7)..(28)

<223> /note="This region may encompass 2-22 nucleotides"

<220><221> modified_base

<222> (45)..(74)

<223> a, c, u, g, unknown or other

<220>

<221> source

<223> /note="See specification as filed for detailed description of

substitutions and preferred embodiments"

<400> 86

nnnnnnnnnn nnnnnnnnnn nnnnnnnca gaaucuacua auacnnnnnn nnnnnnnnnn	60
--	----

nnnnnnnnnn nnnn	74
-----------------	----

<210> 87

<211> 10

<212> RNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 87

uaauuuuagg	10
------------	----

<210> 88

<211> 11

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 88

uaauuuuagg u	11
--------------	----

<210> 89

<211> 12

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 89

uaauuuuagg ua 12

<210> 90

<211> 13

<212> RNA

<213> Artificial Sequence

<

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 90

uaauuuuagg uau 13

<210> 91

<211> 14

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 91

uaauuuuagg uaug 14

<210> 92

<211> 15

<212> RNA

<213> Artificial Sequence

<220

><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 92

uaauuuuagg uauga 15

<210> 93

<211> 101

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (1)..(21)

<223> a, c, u, g, unknown or other

<400> 93

nnnnnnnnnn nnnnnnnnnn nguuuuagag cuagaaauag caaguuaaaa uaaggcuagu	60
---	----

ccguaucaa cuugaaaaag ugacaccgag ucggugcuh u	101
---	-----

<210> 94

<211> 104

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (1)..(21)

<223> a, c, u, g, unknown or other

<400> 94

nnnnnnnnnn nnnnnnnnnn nguuuuagua cucuggaaac agaaucuacu aaaacaaggc	60
---	----

aaaaugccgu guuaaucucg ucaacuuguu ggcgagauuu uuuu	104
--	-----

<210> 95

<211> 34

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 95

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaga	34
<210> 96	
<211> 66	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 96	
aauagcaagu uaaaaauaagg cuaguccguu auacaacuuga aaaaguggca ccgagucggu	60
gcuuuu	66
<210> 97	
<211> 32	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> modified_base	
<222> (1)..(20)	
<223> a, c, u, g, unknown or other	
<400> 97	
nnnnnnnnnn nnnnnnnnnn guuuuagagc ua	32
<210> 98	
<211> 64	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 98	
uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc gagucggugc	60
uuuu	64

<210> 99
<211> 34
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 99

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagg

34

<210> 100
<211> 66
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 100

ccuagcaagu	aaaaauaagg	cuaguccguu	aucaacuuga	aaaaguggca	ccgagucggu	60
gcuuuu						66

<210> 101

<211> 35

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221>

> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 101

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugc	35
<210> 102	
<211> 67	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 102	
gcauagcaag uuaaaaauaag gcuaguccgu uaucaacuug aaaaaguggc accgagucgg	60
ugcuuuu	67
<210> 103	
<211	
> 42	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> modified_base	
<222> (1)..(20)	
<223> a, c, u, g, unknown or other	
<400> 103	
nnnnnnnnnn nnnnnnnnnn guauuagagc uaugcuguuu ug	42
<210> 104	
<211> 74	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 104	
caaaacagca uagcaaguua auauaaggcu aguccguuau caacuugaaa aaguggcacc	60
gagucggugc uuuu	74

<210> 105

<211> 37

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 105

nnnnnnnnnn nnnnnnnnnn guauuagagc uaugcug

37

<210> 106

<211> 69

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 106

cagcauagca aguuauaua aggcuaguucc guuaucacu ugaaaaagug gcaccgaguc

60

ggugcuuuu

69

<210> 107

<211> 42

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 107

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu ug	42
<210> 108	
<211> 74	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 108	
caaaacagca uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc	60
gagucggugc uuuu	74
<210> 109	
<211> 37	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221>	
> modified_base	
<222> (1)..(20)	
<223> a, c, u, g, unknown or other	
<400> 109	
nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcug	37
<210> 110	
<211> 69	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 110	
cagcauagca aguuaaaaaa aggcuagucc guuaucacu ugaaaaagug gcaccgaguc	60
ggugcuuuu	69

<210> 111
<211>
> 37
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 111

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcug

37

<210> 112

<211> 69

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 112

cagcauagca aguuaaaaaua aggcuaguucc guuaucacu ugaaaaagug gcaccgaguc

60

ggugcuuuu

69

<210> 113

<211> 34

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 113

nnnnnnnnnn nnnnnnnnnn guuuuaguac ucug	34
<210> 114	
<211> 62	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 114	
cagaaucuac uaaaacaagg caaaaugccg uguuuauucuc gucaacuugu uggcgagauu	60
uu	62
<210> 115	
<211> 37	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> modified_base	
<222> (1)..(20)	
<223> a, c, u, g, unknown or other	
<400> 115	
nnnnnnnnnn nnnnnnnnnn guuuuaguac ucuguaa	37
<210> 116	
<211> 65	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 116	
uuacagaauc uacaaaaaca aggcaaaaug ccguguuau cucgucaacu uguuggcag	60
auuuu	65

<210> 117

<211> 45

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221>

> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 117

nnnnnnnnnn nnnnnnnnnn	guuuuaguac ucuguaauuu uaggu	45
-----------------------	-----------------------------	----

<210> 118

<211> 73

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 118

accuaaaauu acagaaucua cuaaaacaag gcaaaaugcc guguuuaucu cgucaacuug	60
---	----

uuggcgagau uuu	73
----------------	----

<210> 119

<211>

> 50

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 119

nnnnnnnnnn nnnnnnnnnn guuuuaguac ucuguaauu uagguauagag	50
<210> 120	
<211> 78	
<212> RNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<400> 120	
cucauaccua aaauuacaga aucsacuaaa acaaggcaaa augccgugu uaucucguca	60
acuuguuggc gagauuuu	78
<210> 121	
<211> 10	
<212> DNA	
<213> Artificial Sequence	
<220><221> source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> misc_feature	
<222> (1)..(10)	
<223> /note="This sequence may encompass 1-10 nucleotides"	
<400> 121	
aaaaaaaaaaaa	10
<210> 122	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><221>	
source	
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"	
<220><221> misc_feature	
<222> (1)..(20)	
<223> /note="This sequence may encompass 10-20 nucleotides"	

<400> 122

aaaaaaaaaaa aaaaaaaaaaa 20

<210> 123

<211> 200

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220><221> misc_feature

<222> (1)..(200)

<223> /note="This sequence may encompass 25-200 nucleotides"

<400> 123

aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa 60

aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa 120

aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa 180

aaaaaaaaaaa aaaaaaaaaaa 200