

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-196214

(P2012-196214A)

(43) 公開日 平成24年10月18日(2012.10.18)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 N 9/16 (2006.01)	C 1 2 N 9/16 Z N A Z	4 B 0 2 4
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 A	4 B 0 5 0
C 1 2 N 1/15 (2006.01)	C 1 2 N 1/15	4 B 0 6 5
C 1 2 N 1/19 (2006.01)	C 1 2 N 1/19	
C 1 2 N 5/10 (2006.01)	C 1 2 N 5/00 1 0 1	

審査請求 有 請求項の数 13 O L 外国語出願 (全 190 頁)

(21) 出願番号	特願2012-110691 (P2012-110691)	(71) 出願人	512135735 デューク大学
(22) 出願日	平成24年5月14日 (2012.5.14)		
(62) 分割の表示	特願2008-536805 (P2008-536805) の分割	(74) 代理人	110000659 特許業務法人広江アソシエイツ特許事務所
原出願日	平成18年10月18日 (2006.10.18)	(72) 発明者	ヒーリング, オム, ダブリュ. アメリカ合衆国 ノースカロライナ州 2 7708, ダーハム, ピー. オー. ボック ス 90083
(31) 優先権主張番号	60/727, 512	(72) 発明者	スミス, ジェームス, ジェファーソン アメリカ合衆国 ノースカロライナ州 2 7703, ダーハム, 1808 カーネー ション ドライブ
(32) 優先日	平成17年10月18日 (2005.10.18)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 配列特異性およびDNA-結合親和度に変更された、合理設計メガヌクレアーゼ

(57) 【要約】 (修正有)

【課題】合理的に設計されたLAGLIDAGメガヌクレアーゼ及びそのようなメガヌクレアーゼの製造法の提供。

【解決手段】野生型I-CREIメガヌクレアーゼと比べ、2本鎖DNAに対する結合親和度に変更された組み換えメガヌクレアーゼであって、特定の配列のアミノ酸配列からなるI-CREIメガヌクレアーゼの残基2-153に対して少なくとも85%の配列類似性を有するポリペプチドを含み；DNA結合親和度が、(A)H、N、Q、S、T、K又はRによるE80、D137、I81、L112、P29、V64又はY66の置換；又は、(B)K又はRによるT46、T140又はT143の置換、から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇する、組み換えメガヌクレアーゼ。

【選択図】なし

【特許請求の範囲】

【請求項 1】

野生型 I - C R E I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 2、配列番号 3、配列番号 4 及び配列番号 5 から成る群より選ばれる I - C R E I メガヌクレアーゼ認識配列内の半分部位と、少なくとも 1 塩基対異なる認識配列半分部位に対して特異性を有し、

排外的修飾ではない表 5 の少なくとも一つの修飾を含む、組み換えメガヌクレアーゼ。

10

【請求項 2】

野生型 I - M S O I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 7 及び配列番号 8 から成る群より選ばれる I - M S O I メガヌクレアーゼ認識配列内の半分部位と、少なくとも 1 塩基対異なる認識配列半分部位に対して特異性を有し、

排外的修飾ではない表 7 の少なくとも一つの修飾を含む、組み換えメガヌクレアーゼ。

20

【請求項 3】

野生型 I - S C E I メガヌクレアーゼと比べ、認識配列に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 9 の I - S C E I メガヌクレアーゼの残基 3 - 1 8 6 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 10 及び配列番号 11 から成る群より選ばれる I - S C E I メガヌクレアーゼ認識配列と、少なくとも 1 塩基対異なる認識配列に対して特異性を有し、

排外的修飾ではない表 9 の少なくとも一つの修飾を含む、組み換えメガヌクレアーゼ。

【請求項 4】

野生型 I - C E U I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 12 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 13 及び配列番号 14 から成る群より選ばれる I - C E U I メガヌクレアーゼ認識配列内の半分部位と、少なくとも 1 塩基対異なる認識配列半分部位に対して特異性を有し、

排外的修飾ではない表 11 の少なくとも一つの修飾を含む、組み換えメガヌクレアーゼ。

30

【請求項 5】

野生型 I - C R E I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 2、配列番号 3、配列番号 4 及び配列番号 5 から成る群より選ばれる I - C R E I メガヌクレアーゼ認識配列内の半分部位と、少なくとも 1 塩基対異なる認識配列半分部位に対して特異性を有し、

(1) 位置 - 1 における特異性が、

(A) Q 7 0、C 7 0、L 7 0、Y 7 5、Q 7 5、H 7 5、H 1 3 9、Q 4 6 及び H 4 6 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) Y 7 5、L 7 5、C 7 5、Y 1 3 9、C 4 6 及び A 4 6 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

40

50

(C) K70、E70、E75、E46及びD46から成る群より選ばれる修飾によって、センス鎖においてGに変更され、

(D) H75、R75、H46、K46及びR46から成る群より選ばれる修飾によって、センス鎖においてCに変更され、又は、

(E) G70、A70、S70及びG46から成る群より選ばれる修飾によって、センス鎖において任意の塩基に変更され；

及び/又は

(2) 位置 - 2における特異性が、

(A) Q70、T44、A44、V44、I44、L44及びN44から成る群より選ばれる修飾によって、センス鎖においてAに変更され；

(B) E70、D70、K44及びR44から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(C) H70、D44及びE44から成る群より選ばれる修飾によって、センス鎖においてGに変更され；又は、

(D) C44を含む修飾によって、センス鎖においてA又はTに変更され；

及び/又は、

(3) 位置 - 3における特異性が、

(A) Q68及びC24から成る群より選ばれる修飾によって、センス鎖においてAに変更され；

(B) E68、F68、K24及びR24から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(C) M68、C68、L68及びF68から成る群より選ばれる修飾によって、センス鎖においてTに変更され；

(D) H68を含む修飾によってセンス鎖においてA又はCに変更され；

(E) Y68を含む修飾によってセンス鎖においてC又はTに変更され；

(F) K68を含む修飾によってセンス鎖においてG又はTに変更され；

及び/又は、

(4) 位置 - 4における特異性が、

(A) E77及びK26から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(B) E26及びR77から成る群より選ばれる修飾によって、センス鎖においてGに変更され；

(C) S77を含む修飾によってセンス鎖においてC又はTに変更され；又は、

(D) S26を含む修飾によってセンス鎖において任意の塩基に変更され；

及び/又は、

(5) 位置 - 5における特異性が、

(A) E42を含む修飾によってセンス鎖においてCに変更され；

(B) R42を含む修飾によってセンス鎖においてGに変更され；

(C) C28及びQ42から成る群より選ばれる修飾によって、センス鎖においてA又はCに変更され；又は、

(D) M66及びK66から成る群より選ばれる修飾によって、センス鎖において任意の塩基に変更され；

及び/又は、

(6) 位置 - 6における特異性が、

(A) C40、I40、V40、C79、I79、V79及びQ28から成る群より選ばれる修飾によって、センス鎖においてTに変更され；

(B) E40及びR28から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(C) R40を含む修飾によってセンス鎖においてGに変更され；

及び/又は、

10

20

30

40

50

(7) 位置 - 7 における特異性が、

(A) E 3 8、K 3 0 及び R 3 0 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(B) K 3 8、R 3 8 及び E 3 0 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) I 3 8 及び L 3 8 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(D) C 3 8 を含む修飾によってセンス鎖において A 又は G に変更され；

(E) H 3 8、N 3 8 及び Q 3 0 から成る群より選ばれる修飾によって、センス鎖において任意の塩基に変更され；

及び / 又は、

(8) 位置 - 8 における特異性が、

(A) L 3 3、V 3 3、I 3 3、F 3 3 及び C 3 3 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) E 3 3 及び D 3 3 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(C) K 3 3 から成る修飾によってセンス鎖において G に変更され；

(D) R 3 2 を含む修飾によってセンス鎖において A 又は C に変更され；

(E) R 3 3 を含む修飾によってセンス鎖において A 又は G に変更され；

及び / 又は、

(9) 位置 - 9 における特異性が、

(A) E 3 2 を含む修飾によってセンス鎖において C に変更され；

(B) R 3 2 及び K 3 2 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) L 3 2、V 3 2、A 3 2 及び C 3 2 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(D) D 3 2 及び I 3 2 から成る群より選ばれる修飾によって、センス鎖において C 又は T に変更され；又は、

(E) S 3 2、N 3 2、H 3 2、Q 3 2 及び T 3 2 から成る群より選ばれる修飾によって、センス鎖において任意の塩基に変更される、

ことを特徴とする、組み換えメガヌクレアーゼ。

【請求項 6】

野生型 I - M S O I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 7 及び配列番号 8 から成る群より選ばれる I - M S O I メガヌクレアーゼ認識配列内の半分部位と、少なくとも 1 塩基対異なる認識配列半分部位に対して特異性を有し、

(1) 位置 - 1 における特異性が、

(A) K 7 5、Q 7 7、A 4 9、C 4 9 及び K 7 9 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) C 7 7、L 7 7 及び Q 7 9 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(C) K 7 7、R 7 7、E 4 9 及び E 7 9 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び / 又は、

(2) 位置 - 2 における特異性が、

(A) Q 7 5、K 8 1、C 4 7、I 4 7 及び L 4 7 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

10

20

30

40

50

(B) E 7 5、D 7 5、R 4 7、K 4 7、K 8 1 及び R 8 1 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(C) K 7 5、E 4 7 及び E 8 1 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び / 又は、

(3) 位置 - 3 における特異性が、

(A) Q 7 2、C 2 6、L 2 6、V 2 6、A 2 6 及び I 2 6 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) E 7 2、Y 7 2、H 2 6、K 2 6 及び R 2 6 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(C) K 7 2、Y 7 2 及び H 2 6 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

及び / 又は、

(4) 位置 - 4 における特異性が、

(A) K 2 8、K 8 3 及び Q 2 8 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) R 8 3 及び K 8 3 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) K 2 8 及び Q 8 3 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

及び / 又は、

(5) 位置 - 5 における特異性が、

(A) R 4 5 及び E 2 8 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(B) Q 2 8 を含む修飾によってセンス鎖において T に変更され；

(C) R 2 8 を含む修飾によってセンス鎖において C に変更され；

及び / 又は、

(6) 位置 - 6 における特異性が、

(A) K 4 3、V 8 5、L 8 5 及び Q 3 0 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) E 4 3、E 8 5、K 3 0 及び R 3 0 から成る群より選ばれる修飾によって、センス鎖において C に変更され；または、

(C) R 4 3、K 4 3、K 8 5、R 8 5、E 3 0 及び D 3 0 より成る群から選ばれる修飾によって、センス鎖において G に変更され；

及び / 又は、

(7) 位置 - 7 における特異性が、

(A) E 3 2 及び E 4 1 から成る群から選ばれる修飾によって、センス鎖において C に変更され；

(B) R 3 2、R 4 1 及び K 4 1 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) K 3 2、M 4 1、L 4 1 及び I 4 1 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

及び / 又は、

(8) 位置 - 8 における特異性が、

(A) K 3 2 及び K 3 5 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) E 3 2 を含む修飾によってセンス鎖において C に変更され；

(C) K 3 2、K 3 5 及び R 3 5 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び / 又は、

10

20

30

40

50

(9) 位置 - 9 における特異性が、

(A) N 3 4 及び H 3 4 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) S 3 4、C 3 4、V 3 4、T 3 4 及び A 3 4 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(C) K 3 4、R 3 4 及び H 3 4 から成る群より選ばれる修飾によって、センス鎖において G に変更される、

ことを特徴とする、組み換えメガヌクレアーゼ。

【請求項 7】

野生型 I - S C E I メガヌクレアーゼと比べ、認識配列に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 9 の I - S C E I メガヌクレアーゼの残基 3 - 1 8 6 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 1 0 及び配列番号 1 1 の I - S C E I メガヌクレアーゼと、少なくとも 1 塩基対異なる認識配列に対して特異性を有し、

(1) 位置 4 における特異性が、

(A) K 5 0 を含む修飾によってセンス鎖において A に変更され；

(B) K 5 7、M 5 7 及び Q 5 0 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(C) E 5 0、R 5 7 及び K 5 7 から成る群より選ばれる修飾によって、センス鎖において G に変更され；及び / 又は、

(2) 位置 5 における特異性が、

(A) K 4 8、Q 1 0 2 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) E 4 8、K 1 0 2 及び R 1 0 2 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) Q 4 8、C 1 0 2、L 1 0 2 及び V 1 0 2 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

及び / 又は、

(3) 位置 6 における特異性が、

(A) K 5 9 を含む修飾によってセンス鎖において A に変更され；

(B) R 5 9 及び K 5 9 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(B) K 8 4 及び E 5 9 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び / 又は、

(4) 位置 7 における特異性が、

(A) R 4 6、K 4 6 及び E 8 6 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(B) K 8 6、R 8 6 及び E 4 6 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) C 4 6、L 4 6 及び V 4 6 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

及び / 又は、

(5) 位置 8 における特異性が、

(A) E 8 8、R 6 1 及び H 6 1 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(B) K 8 8、Q 6 1 及び H 6 1 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(C) K 6 1、S 6 1、V 6 1、A 6 1 及び L 6 1 から成る群より選ばれる修飾によ

10

20

30

40

50

って、センス鎖においてAに変更され；

及び/又は、

(6) 位置9における特異性が、

(A) C98、V98及びL98から成る群より選ばれる修飾によって、センス鎖においてAに変更され；

(B) R98及びK98から成る群より選ばれる修飾によって、センス鎖においてCに変更され；又は、

(C) E98及びD98から成る群より選ばれる修飾によって、センス鎖においてGに変更され；

及び/又は、

(7) 位置10における特異性が、

(A) K96及びR96から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(B) D96及びE96から成る群より選ばれる修飾によって、センス鎖においてGに変更され；

(C) C96及びA96から成る群より選ばれる修飾によって、センス鎖においてAに変更され；

及び/又は、

(8) 位置11における特異性が、

(A) Q90を含む修飾によってセンス鎖においてTに変更され；

(B) K90及びR90から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(C) E90を含む修飾によってセンス鎖においてGに変更され；

及び/又は、

(9) 位置12における特異性が、

(A) Q193を含む修飾によってセンス鎖においてAに変更され；

(B) E165、E193及びD193から成る群より選ばれる修飾によって、センス鎖においてCに変更され；

(C) K165及びR165から成る群より選ばれる修飾によって、センス鎖においてGに変更され；

及び/又は、

(10) 位置13における特異性が、

(A) Q193、C163及びL163から成る群より選ばれる修飾によって、センス鎖においてTに変更され；

(B) E193、D193、K163及びR192から成る群より選ばれる修飾によって、センス鎖においてGに変更され；

(C) C193およびL193から成る群より選ばれる修飾によって、センス鎖においてAに変更され；

及び/又は、

(11) 位置14における特異性が、

(A) K161及びQ192から成る群より選ばれる修飾によって、センス鎖においてTに変更され；

(B) L192及びC192から成る群より選ばれる修飾によって、センス鎖においてAに変更され；

(C) K147、K161、R161、R197、D192及びE192から成る群より選ばれる修飾によって、センス鎖においてGに変更され；または、

(D) K161及びQ192から成る群から選ばれる修飾によって、センス鎖においてTに変更され；

及び/又は

(12) 位置15における特異性が、

10

20

30

40

50

(A) C 1 5 1、L 1 5 1 及び K 1 5 1 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) K 1 5 1 を含む修飾によってセンス鎖において G に変更され；又は、

(C) E 1 5 1 を含む修飾によってセンス鎖において C に変更され；

及び / 又は、

(13) 位置 1 7 における特異性が、

(A) G 1 5 2 及び Q 1 5 0 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) K 1 5 2 及び K 1 5 0 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(C) N 1 5 2、S 1 5 2、D 1 5 2、D 1 5 0 及び E 1 5 0 から成る群より選ばれる修飾によってセンス鎖において G に変更され；

及び / 又は、

(14) 位置 1 8 における特異性が、

(A) H 1 5 5 及び Y 1 5 5 から成る群より選ばれる修飾によってセンス鎖において T に変更され；

(B) R 1 5 5 及び K 1 5 5 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(C) K 1 5 5 及び C 1 5 5 から成る群より選ばれる修飾によって、センス鎖において A に変更される

ことを特徴とする、組み換えメガヌクレアーゼ。

【請求項 8】

野生型 I - C E U I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対する特異性が変更された組み換えメガヌクレアーゼであって、

配列番号 1 2 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対し少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、かつ、

配列番号 1 3 及び配列番号 1 4 から成る群より選ばれる I - C E U I メガヌクレアーゼ認識配列内の半分部位と、少なくとも 1 塩基対異なる認識配列半分部位に対して特異性を有し、

(1) 位置 - 1 における特異性が、

(A) C 9 2、A 9 2 及び V 9 2 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) Q 1 1 6 及び Q 9 2 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(C) E 1 1 6 及び E 9 2 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び / 又は、

(2) 位置 - 2 における特異性が、

(A) Q 1 1 7、C 9 0、L 9 0 及び V 9 0 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) K 1 1 7、R 1 2 4、K 1 2 4、E 1 2 4、E 9 0 及び D 9 0 から成る群から選ばれる修飾によって、センス鎖において G に変更され；

(C) E 1 1 7、D 1 1 7、R 1 7 4、K 1 2 4、K 9 0 及び K 6 8 から成る群から選ばれる修飾によって、センス鎖において C に変更され；

及び / 又は、

(3) 位置 - 3 における特異性が、

(A) C 7 0、V 7 0、T 7 0、L 7 0 及び K 7 0 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) Q 7 0 を含む修飾によってセンス鎖において T に変更され；

(B) K 7 0 を含む修飾によってセンス鎖において C に変更され；

10

20

30

40

50

及び/又は、

(4) 位置 - 4 における特異性が、

(A) E 1 2 6、D 1 2 6、R 8 8、K 8 8 及び K 7 2 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(B) K 1 2 6、L 1 2 6 及び Q 8 8 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(C) Q 1 2 6、N 1 2 6、K 8 8、L 8 8、C 7 2、L 7 2 及び V 7 2 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

及び/又は、

(5) 位置 - 5 における特異性が、

(A) E 7 4、K 1 2 8、R 1 2 8 及び E 1 2 8 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(B) C 1 2 8、L 1 2 8、V 1 2 8 及び T 1 2 8 から成る群から選ばれる修飾によって、センス鎖において T に変更され；又は、

(C) C 7 4、L 7 4、V 7 4 及び T 7 4 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

及び/又は、

(6) 位置 - 6 における特異性が、

(A) K 8 6、C 8 6 及び L 8 6 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) D 8 6、E 8 6、R 8 4 及び K 8 4 から成る群より選ばれる修飾によって、センス鎖において C に変更され；又は、

(C) K 1 2 8、R 1 2 8、R 8 6、K 8 6 及び E 8 4 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び/又は、

(7) 位置 - 7 における特異性が、

(A) R 7 6、K 7 6 及び H 7 6 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(B) E 7 6 及び R 8 4 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

(C) H 7 6 及び Q 7 6 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

及び/又は、

(8) 位置 - 8 における特異性が、

(A) Y 7 9、R 7 9 及び Q 7 6 から成る群より選ばれる修飾によって、センス鎖において A に変更され；

(B) D 7 9、E 7 9、D 7 6 及び E 7 6 から成る群より選ばれる修飾によって、センス鎖において C に変更され；

(C) R 7 9、K 7 9、K 7 6 及び R 7 6 から成る群より選ばれる修飾によって、センス鎖において G に変更され；

及び/又は、

(9) 位置 - 9 における特異性が、

(A) K 7 8、V 7 8、L 7 8、C 7 8 及び T 7 8 から成る群より選ばれる修飾によって、センス鎖において T に変更され；

(B) D 7 8 及び E 7 8 から成る群より選ばれる修飾によって、センス鎖において C に変更され；又は、

(C) R 7 8、K 7 8 及び H 7 8 から成る群より選ばれる修飾によって、センス鎖において G に変更される、

ことを特徴とする、組み換えメガヌクレアーゼ。

【請求項 9】

10

20

30

40

50

野生型 I - C R E I メガヌクレアーゼと比べ、2本鎖 DNA に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、K 又は R による E 8 0、D 1 3 7、I 8 1、L 1 1 2、P 2 9、V 6 4 又は Y 6 6 の置換；

又は、

(B) K 又は R による T 4 6、T 1 4 0 又は T 1 4 3 の置換、

から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇することを特徴とする、組み換えメガヌクレアーゼ。

10

【請求項 1 0】

野生型 I - C R E I メガヌクレアーゼと比べ、2本鎖 DNA に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、D 又は E による K 3 4、K 4 8、R 5 1、K 8 2、K 1 1 6、または K 1 3 9 の置換；

又は、

(B) D 又は E による I 8 1、L 1 1 2、P 2 9、V 6 4、Y 6 6、T 4 6、T 1 4 0 又は T 1 4 3 の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって低下することを特徴とする、組み換えメガヌクレアーゼ。

20

【請求項 1 1】

野生型 I - M S O I メガヌクレアーゼと比べ、2本鎖 DNA に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、K 又は R による E 1 4 7、I 8 5、G 8 6 又は Y 1 1 8 の置換；

又は、

(B) K 又は R による Q 4 1、N 7 0、S 8 7、T 8 8、H 8 9、Q 1 2 2、Q 1 3 9、S 1 5 0 又は N 1 5 2 の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇することを特徴とする、組み換えメガヌクレアーゼ。

30

【請求項 1 2】

野生型 I - M S O I メガヌクレアーゼと比べ、2本鎖 DNA に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、D 又は E による K 3 6、R 5 1、K 1 2 3、K 1 4 3、または R 1 4 4 の置換；

又は

(B) D 又は E による I 8 5、G 8 6、Y 1 1 8、Q 4 1、N 7 0、S 8 7、T 8 8、H 8 9、Q 1 2 2、Q 1 3 9、S 1 5 0 又は N 1 5 2 の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって低下することを特徴とする、組み換えメガヌクレアーゼ。

40

【請求項 1 3】

野生型 I - S C E I メガヌクレアーゼと比べ、2本鎖 DNA に対する結合親和度が変更

50

された組み換えメガヌクレアーゼであって、

配列番号 9 の I - S C E I メガヌクレアーゼの残基 3 - 1 8 6 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、K 又は R による D 2 0 1、L 1 9、L 8 0、L 9 2、Y 1 5 1、Y 1 8 8、I 1 9 1、Y 1 9 9 又は Y 2 2 2 の置換；

又は

(B) K 又は R による N 1 5、N 1 7、S 8 1、H 8 4、N 9 4、N 1 2 0、T 1 5 6、N 1 5 7、S 1 5 9、N 1 6 3、Q 1 6 5、S 1 6 6、N 1 9 4 又は S 2 0 2 の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇することを特徴とする、組み換えメガヌクレアーゼ。

10

【請求項 1 4】

野生型 I - S C E I メガヌクレアーゼと比べ、2 本鎖 D N A に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 9 の I - S C E I メガヌクレアーゼの残基 3 - 1 8 6 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、D 又は E による K 2 0、K 2 3、K 6 3、K 1 2 2、K 1 4 8、K 1 5 3、K 1 9 0、K 1 9 3、K 1 9 5 又は K 2 2 3 の置換；

又は

(B) D 又は E による L 1 9、L 8 0、L 9 2、Y 1 5 1、Y 1 8 8、I 1 9 1、Y 1 9 9、Y 2 2 2、N 1 5、N 1 7、S 8 1、H 8 4、N 9 4、N 1 2 0、T 1 5 6、N 1 5 7、S 1 5 9、N 1 6 3、Q 1 6 5、S 1 6 6、N 1 9 4 又は S 2 0 2 の置換から成る群から選ばれる一つの置換に対応する少なくとも一つの修飾によって低下することを特徴とする、組み換えメガヌクレアーゼ。

20

【請求項 1 5】

野生型 I - C E U I メガヌクレアーゼと比べ、2 本鎖 D N A に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 1 2 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、K 又は R による D 2 5 又は D 1 2 8 の置換又は

(B) K 又は R による S 6 8、N 7 0、H 9 4、S 1 1 7、N 1 2 0、N 1 2 9、又は H 1 7 2 の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇することを特徴とする、組み換えメガヌクレアーゼ。

30

【請求項 1 6】

野生型 I - C E U I メガヌクレアーゼと比べ、2 本鎖 D N A に対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号 1 2 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

D N A 結合親和度が、

(A) H、N、Q、S、T、D 又は E による K 2 1、K 2 8、K 3 1、R 1 1 2、R 1 1 4、または R 1 3 0 の置換；

又は

(B) D 又は E による S 6 8、N 7 0、H 9 4、S 1 1 7、N 1 2 0、N 1 2 9 又は H 1 7 2 の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって低下することを特徴とする、組み換えメガヌクレアーゼ。

40

【請求項 1 7】

参照メガヌクレアーゼモノマーによるダイマー形成に対する親和度を変更させた組み換えメガヌクレアーゼモノマーであって、

50

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、

ダイマー形成に対する親和度が、

(A) D 又は E による K 7、K 5 7 又は K 9 6 の置換；
又は

(B) K 又は R による E 8 又は E 6 1 の置換から成る群より選ばれる置換に対応する少なくとも一つの修飾によって変えられることを特徴とする組み換えメガヌクレアーゼモノマー。

【請求項 1 8】

組み換えメガヌクレアーゼ・ヘテロダイマーであって、

10

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有する第 1 ポリペプチドであって、

ダイマー形成に対する親和度が、

(A) D 又は E による K 7、K 5 7 又は K 9 6 の置換から成る群から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第 1 ポリペプチド及び、

配列番号 1 の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有する第 2 ポリペプチドであって、

ダイマー形成に対する親和度が、

(B) K 又は R による E 8 又は E 6 1 の置換から成る群から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第 2 ポリペプチドを含む、

20

ことを特徴とする、組み換えメガヌクレアーゼ・ヘテロダイマー。

【請求項 1 9】

参照メガヌクレアーゼモノマーによるダイマー形成に対する親和度を変更させた組み換えメガヌクレアーゼモノマーであって、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、

ダイマー形成に対する親和度が、

(A) D 又は E による R 3 0 2 の置換；

又は

(B) K 又は R による D 2 0、E 1 1 又は Q 6 4 の置換から成る群から選ばれる置換に対応する少なくとも一つの修飾によって変えられることを特徴とする、組み換えメガヌクレアーゼモノマー。

30

【請求項 2 0】

組み換えメガヌクレアーゼ・ヘテロダイマーであって、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対して少なくとも 8 5 % の配列類似性を有する第 1 ポリペプチドであって、

ダイマー形成に対する親和度が、

(A) D 又は E による R 3 0 2 の置換から成る群より選ばれる置換に対応する少なくとも一つの修飾によって変えられる第 1 ポリペプチド及び、

配列番号 6 の I - M S O I メガヌクレアーゼの残基 6 - 1 6 0 に対して少なくとも 8 5 % の配列類似性を有する第 2 ポリペプチドであって、

40

ダイマー形成に対する親和度が、

(B) K 又は R による D 2 0、E 1 1 又は Q 6 4 の置換から成る群から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第 2 ポリペプチドを含む、

ことを特徴とする、組み換えメガヌクレアーゼ・ヘテロダイマー。

【請求項 2 1】

参照メガヌクレアーゼモノマーによるダイマー形成に対する親和度を変更させた組み換えメガヌクレアーゼモノマーであって、

配列番号 1 2 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み、

50

ダイマー形成に対する親和度が、

(A) D又はEによるR93の置換；

又は

(B) K又はRによるE152の置換から成る群から選ばれる置換に対応する少なく一つの修飾によって変えられることを特徴とする、前記組み換えメガヌクレアーゼモノマー。

【請求項22】

組み換えメガヌクレアーゼ・ヘテロダイマーであって、

配列番号12のI-C E U Iメガヌクレアーゼの残基5-211に対して少なくとも85%の配列類似性を有する第1ポリペプチドであって、

ダイマー形成に対する親和度が、

(A) D又はEによるR93の置換から成る群より選ばれる置換に対応する少なく一つの修飾によって変えられる第1ポリペプチド及び、

配列番号12のI-C E U Iメガヌクレアーゼの残基5-211に対して少なくとも85%の配列類似性を有する第2ポリペプチドであって、

ダイマー形成に対する親和度が、

(B) K又はRによるE152の置換から成る群より選ばれる置換に対応する少なく一つの修飾によって変えられる第2ポリペプチドを含む、

ことを特徴とする組み換えメガヌクレアーゼ・ヘテロダイマー。

【請求項23】

表1から選ばれる少なく一つの修飾で、排外的修飾ではない修飾をさらに含むことを特徴とする、

請求項17又は請求項18に記載の組み換えメガヌクレアーゼモノマー又はヘテロダイマー。

【請求項24】

表2から選ばれる少なく一つの修飾で、排外的修飾ではない修飾をさらに含むことを特徴とする、

請求項19又は請求項20に記載の組み換えメガヌクレアーゼモノマー又はヘテロダイマー。

【請求項25】

表4から選ばれる少なく一つの修飾で、排外的修飾ではない修飾をさらに含むことを特徴とする、

請求項21又は請求項22に記載の組み換えメガヌクレアーゼモノマー又はヘテロダイマー。

【請求項26】

真核細胞の染色体に挿入される興味の外因性配列を含む遺伝学的に修飾された真核細胞を生産するための方法であって、

真核細胞を、

(I)メガヌクレアーゼをコードする第1核酸配列及び、

(II)前記興味の配列を含む第2核酸配列

を含む一つ以上の核酸によってトランスフェクトすることを含み、

前記メガヌクレアーゼが前記染色体の中に切断部位を生産し、かつ、前記興味の配列が前記切断部位において前記染色体の中に挿入され、

前記メガヌクレアーゼが請求項1~25のいずれか1項による組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項27】

前記第2核酸が、前記切断部位に側接する配列に対し相同な配列をさらに含み、前記興味の配列が、相同組み換えによって前記切断部位において挿入されることを特徴とする、

請求項26に記載の方法。

【請求項28】

10

20

30

40

50

前記第 2 核酸が、前記切断部位に対し実質的相同性を欠き、前記興味の配列が、非相同的末端接合によって前記染色体に挿入されることを特徴とする、

請求項 26 に記載の方法。

【請求項 29】

真核細胞の染色体に挿入される興味の外因性配列を含む、遺伝学的に修飾された真核細胞を生産するための方法であって、

真核細胞にメガヌクレアーゼタンパクを導入すること；

及び、

前記興味の配列を含む核酸によって前記真核細胞をトランスフェクトすることを含み、前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記興味の配列が、前記切断部位において前記染色体の中に挿入され；

かつ、

前記メガヌクレアーゼが請求項 1 ~ 25 のいずれか 1 項による組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項 30】

前記核酸が、前記切断部位に側接する配列に対し相同な配列をさらに含み、前記興味の配列が、相同組み換えによって前記切断部位において挿入されることを特徴とする、

請求項 29 に記載の方法。

【請求項 31】

前記核酸が、前記切断部位に対し実質的相同性を欠き、前記興味の配列が、非相同的末端接合によって前記染色体に挿入されることを特徴とする、

請求項 29 に記載の方法。

【請求項 32】

真核細胞の染色体中の標的配列を破壊することによって遺伝学的に修飾された真核細胞を生産するための方法であって、

メガヌクレアーゼをコードする核酸によって真核細胞をトランスフェクトすることを含み、

前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記標的配列が、前記切断部位において非相同的末端接合によって破壊され；

かつ、

前記メガヌクレアーゼが、請求項 1 ~ 25 のいずれか 1 項による組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項 33】

遺伝学的に修飾された生物体を生産する方法であって、

請求項 26 ~ 32 のいずれか 1 項の方法に従って遺伝学的に修飾された真核細胞を生産するステップと、

及び、

前記遺伝学的に修飾された真核細胞を育成して遺伝学的に修飾された生物体を生産するステップと、

を含む方法。

【請求項 34】

前記真核細胞が、配偶子、接合子、胚盤胞細胞、胚幹細胞及びプロトプラスト細胞から成る群より選ばれることを特徴とする、

請求項 33 に記載の方法。

【請求項 35】

真核生物において遺伝子治療によって疾患を治療するための方法であって、

前記真核生物の少なくとも一つの細胞を、

(I) メガヌクレアーゼをコードする第 1 核酸配列及び、

(II) 興味の配列を含む第 2 核酸配列

によってトランスフェクトすることを含み；

10

20

30

40

50

前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記興味の配列が、前記切断部位において前記染色体の中に挿入され；

かつ、

前記メガヌクレアーゼが、請求項 1 - 25 のいずれか 1 項による組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項 36】

前記第 2 核酸配列が、前記切断部位に側接する配列に対し相同な配列をさらに含み、前記興味の配列が、相同組み換えによって前記切断部位において挿入されることを特徴とする、

請求項 35 に記載の方法。

10

【請求項 37】

前記第 2 核酸配列が、前記切断部位に対し実質的相同性を欠き、前記興味の配列が、非相同的末端接合によって前記染色体に挿入されることを特徴とする、

請求項 35 に記載の方法。

【請求項 38】

真核生物において遺伝子治療によって疾患を治療するための方法であって、前記真核生物の少なくとも一つの細胞の中にメガヌクレアーゼタンパクを導入することとステップと、

前記真核細胞を、興味の配列を含む核酸によってトランスフェクトするステップとを含み、

20

前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記興味の配列が、前記切断部位において前記染色体の中に挿入され；

前記メガヌクレアーゼが請求項 1 ~ 25 のいずれか 1 項による組み換えメガヌクレアーゼであり；

かつ、

前記興味の配列の挿入が、前記疾患に対する前記遺伝子治療を与えることを特徴とする方法。

【請求項 39】

前記核酸が、前記切断部位に側接する配列に対し相同な配列をさらに含み、前記興味の配列が、相同組み換えによって前記切断部位において挿入されることを特徴とする、

請求項 38 に記載の方法。

30

【請求項 40】

前記核酸が、前記切断部位に対し実質的相同性を欠き、前記興味の配列が、非相同的末端接合によって前記染色体に挿入されることを特徴とする、

請求項 38 に記載の方法。

【請求項 41】

真核生物において、前記真核生物の染色体中の標的配列の破壊による遺伝子治療によって疾患を治療するための方法であって、

前記真核生物の少なくとも一つの細胞を、メガヌクレアーゼをコードする核酸によってトランスフェクトすることを含み；

40

前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記標的配列が、前記切断部位において非相同的末端接合によって破壊され；

前記メガヌクレアーゼが、請求項 1 - 25 のいずれか 1 項による組み換えメガヌクレアーゼであり；

かつ、

前記標的配列の破壊が、前記疾患に対する前記遺伝子治療を与えることを特徴とする方法。

【請求項 42】

真核細胞宿主におけるウイルス性病原体感染を、前記ウイルス性病原体のゲノムにおける標的配列を破壊することによって治療する方法であって、

50

前記真核細胞宿主の少なくとも一つの感染細胞を、メガヌクレアーゼをコードする核酸によってトランスフェクトすることを含み；

前記メガヌクレアーゼが、前記ウィルスゲノムの中に切断部位を生産し、前記標的配列が、前記切断部位において非相同的末端接合によって破壊され、

前記メガヌクレアーゼが、請求項 1 - 25 のいずれか 1 項による組み換えメガヌクレアーゼであり；

かつ、

前記標的配列の破壊が、前記感染に対する治療を与えることを特徴とする方法。

【請求項 43】

真核細胞宿主におけるウィルス性病原体感染を、前記ウィルス病原体のゲノムにおける標的配列を破壊することによって治療する方法であって、

前記真核細胞宿主の少なくとも一つの感染細胞を、メガヌクレアーゼをコードする第 1 核酸、および第 2 核酸によってトランスフェクトすることを含み；

前記メガヌクレアーゼが、前記ウィルスゲノムの中に切断部位を生産し、前記標的配列が、前記切断部位において、前記ウィルスゲノムと前記第 2 核酸との相同組み換えによって破壊され；

前記メガヌクレアーゼが、請求項 1 - 25 のいずれか 1 項による組み換えメガヌクレアーゼであり；

前記第 2 核酸が、前記切断部位に側接する配列に対して相同な配列を含み；

かつ、

前記標的配列の破壊が、前記感染に対する治療を与えることを特徴とする方法。

【請求項 44】

真核細胞宿主における前核細胞病原体感染を、前記前核細胞病原体のゲノムにおける標的配列を破壊することによって治療する方法であって、

前記真核細胞宿主に感染する前記前核細胞病原体の少なくとも一つの細胞を、メガヌクレアーゼをコードする核酸によってトランスフェクトすることを含み；

前記メガヌクレアーゼが、前記前核細胞ゲノムの中に切断部位を生産し、前記標的配列が、前記切断部位において、非相同的末端接合によって破壊され；

前記メガヌクレアーゼが、請求項 1 ~ 25 のいずれか 1 項による組み換えメガヌクレアーゼであり；

かつ、

前記標的配列の破壊が、前記感染に対する治療を与えることを特徴とする方法。

【請求項 45】

真核細胞宿主における前核細胞病原体感染を、前記前核細胞病原体のゲノムにおける標的配列を破壊することによって治療する方法であって、

前記真核細胞宿主に感染する前記前核細胞病原体の少なくとも一つの細胞を、メガヌクレアーゼをコードする第 1 核酸、および第 2 核酸によってトランスフェクトすることを含み；

前記メガヌクレアーゼが、前記前核細胞ゲノムの中に切断部位を生産し、前記標的配列が、前記切断部位において、前記前核細胞ゲノムと前記第 2 核酸との相同組み換えによって破壊され；

前記メガヌクレアーゼが、請求項 1 ~ 25 のいずれか 1 項による組み換えメガヌクレアーゼであり；

前記第 2 核酸が、前記切断部位に側接する配列に対して相同な配列を含み；

かつ、

前記標的配列の破壊が、前記感染に対する治療を与えることを特徴とする方法。

【請求項 46】

少なくとも一つの所望の変化を含むように、認識配列の少なくとも一つの塩基位置に対する特異性を変更させた組み換えメガヌクレアーゼを合理的に設計する方法であって、

(1) 参照メガヌクレアーゼ - DNA 複合体の三次元構造の少なくとも一部を決定する

10

20

30

40

50

こと；

(2) 前記塩基位置において塩基接触面を形成するアミノ酸残基を特定すること；

(3) 前記接触面を含む少なくとも第1残基の -炭素と、前記塩基位置における少なくとも第1塩基との間の距離を決めること；

及び、

(4) 前記所望の変化を促進するアミノ酸置換を、表2に示される群の中から、

(A) 第1塩基から6オングストローム未満の距離にある第1残基については、G群、C群、T群又はA群のうち、前記所望の変化を促進するのに適切なもののメンバーである1群及び/又は2群の置換を選択すること；

又は、

(B) 前記第1塩基から6オングストロームより遠い距離にある第1残基については、G群、C群、T群又はA群のうち、前記所望の変化を促進するのに適切なもののメンバーである2群及び/又は3群の置換を選択することによって特定すること、

を含む方法。

【請求項47】

DNA-結合親和度を上昇させる組み換えメガヌクレアーゼを合理的に設計する方法であって、

(1) 参照メガヌクレアーゼ-DNA複合体の三次元構造の少なくとも一部を決定すること；

(2) バックボーン接触面を形成するアミノ酸接触残基を特定すること；

(3) 前記DNA-結合親和度を上昇させるアミノ酸置換を、

(A) 陰性荷電、または疎水性側鎖を持つ接触残基については、非荷電/極性、または陽性荷電側鎖を持つ置換を選択すること；

又は

(B) 非荷電/極性側鎖を持つ接触残基については、陽性荷電側鎖を持つ置換を選択することによって特定すること、

を含む方法。

【請求項48】

DNA-結合親和度を低下させる組み換えメガヌクレアーゼを合理的に設計する方法であって、

(1) 参照メガヌクレアーゼ-DNA複合体の三次元構造の少なくとも一部を決定すること；

(2) バックボーン接触面を形成するアミノ酸接触残基を特定すること；

(3) 前記DNA-結合親和度を低下させるアミノ酸置換を、

(A) 陽性荷電側鎖を持つ接触残基については、非荷電/極性、または陰性荷電側鎖を持つ置換を選択すること；

又は

(B) 疎水性、または非荷電/極性側鎖を持つ接触残基については、陰性荷電側鎖を持つ置換を選択することによって特定すること、

を含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、分子生物学および組み換え核酸工学の分野に関する。特に、本発明は、DNA認識配列特異性及び/又は親和度に変更された、合理的に設計された、非天然のメガヌクレアーゼに関する。さらに、本発明は、そのようなメガヌクレアーゼの生産法、および、そのようなメガヌクレアーゼによる、組み換え核酸および生物の生産法に関する。

【背景技術】

【0002】

(関連出願) 本出願は、2005年10月18日出願の米国特許仮出願第60/727

10

20

30

40

50

、512号の優先権の利益を主張する。なお、この引用出願の全開示を、参照することにより本出願に含める。

【0003】

(政府援助)本発明は、アメリカ合衆国国立衛生研究所、国立医科学研究所からの研究補助金2R01-GM-0498712、5F32-GM072322、および5DP1OD000122によって一部援助された。したがって、米国政府も、本発明において若干の権利を保有する場合がある。

【0004】

ゲノム工学は、あるゲノムの内部に、特定の遺伝子配列を、挿入し、欠失し、置換し、および、その他のやり方で操作することを可能とする能力を要求するが、多くの治療的および生物工学的応用を有する。ゲノム修飾のための効果的手段の開発は、これまでずっと、遺伝子治療、農工学、および合成生物学における主要目標となっている(PORTEUS ET AL.(2005), NAT. BIOTECHNOL. 23:967-73; TZFIRA ET AL.(2005), TRENDS BIOTECHNOL. 23:567-9; MCDANIEL ET AL.(2005), CURR. OPIN. BIOTECHNOL. 16:476-83)。DNA配列を挿入又は修飾するための一般的な方法は、ゲノム標的に対して相同な配列によって側接される、トランスジェニックDNA配列を導入すること、および、所望の相同組み換え事象を選択またはスクリーニングすることを含む。トランスジェニックDNAによる組み換えはめったに起こらないが、標的部位のゲノムDNAにおける二本鎖切断によって刺激することが可能である。従来から、DNA二本鎖切断部を作るために、例えば、放射線照射および化学的処理を含めた数多くの方法が用いられている。これらの方法は、効率的に組み換えを刺激することが可能ではあるけれども、二本鎖切断部は、ゲノムの中にランダムに挿入され、それらは、突然変異誘発性および毒性が高い可能性がある。現在、染色体バクグラウンドにある特定の部位に対し遺伝子修飾を照準することはできていないため、思い通りのゲノム加工に対する大きな障害となっている。

【0005】

この目標を達成するための一つの方法は、ゲノム内の単一部位にしか存在しないほど十分に大きい配列に対し特異性を持つヌクレアーゼを用いて、標的配座に二本鎖切断部を設けてその部位における相同組み換えを刺激することである(例えば、PORTEUS ET AL.(2005), NAT. BIOTECHNOL. 23:967-73を参照)。この戦略の有効性は、FOKI制限酵素の、加工されたジンクフィンガーDNA-結合ドメインと、非特異的ヌクレアーゼドメインとの間のキメラ融合を用いて、各種の生物体において明らかにされている(PORTEUS(2006), MOL. THER. 13:438-46; WRIGHT ET AL.(2005), PLANT J. 44:693-705; URNOV ET AL.(2005), NATURE 435:646-51)。これらの人工的ジンクフィンガーヌクレアーゼは、部位特異的組み換えを刺激するけれども、これらのヌクレアーゼは、ヌクレアーゼドメインの調整低調に由来する、非特異的な、残留切断活性を保持するので、不要の部位を切断することがしばしばある(SMITH ET AL.(2000), NUCLEIC ACIDS RES. 28:3361-9)。これらの不要な切断は、処置生物体において、突然変異および毒性を引き起こす可能性がある(PORTEUS ET AL.(2005), NAT. BIOTECHNOL. 23:967-73)。

【0006】

植物および菌類のゲノムの中に一般的に認められる、15-40塩基対切断部位を認識する、1群の天然ヌクレアーゼは、比較的毒性の低い、ゲノム加工代替品を提供する可能性がある。このような「メガヌクレアーゼ」または「ホーミング・エンドヌクレアーゼ」は、多くの場合、寄生的DNA要素、例えば、グループ1自己スプライシングイントロンおよびインテンとコンタクトされる。これらのヌクレアーゼは、自然状態で、宿主ゲノムの特異的位置における相同的組み換えまたは遺伝子挿入を促進する。これを、細胞のD

10

20

30

40

50

NA - 修復機構を招集する、染色体において二本鎖切断を生産することによって行う (STODDARD (2006), REV. BIOPHYS. 38: 49 - 95)。一般に、メガヌクレアーゼは、4つのファミリーに分けられる。すなわち、LAGLIDADGファミリー、GIY-YIGファミリー、HIS-CYSボックスファミリー、およびHNHファミリーである。これらのファミリーは、触媒活性および認識配列に影響を及ぼす、構造的モチーフによって特徴づけられる。例えば、LAGLIDADGファミリーのメンバーは、LAGLIDADG保存モチーフの、1または2コピーを持つことによって特徴づけられる (CHEVALIER ET AL. (2001), NUCLEIC ACIDS RES. 29 (18): 3757 - 3774を参照)。LAGLIDADGモチーフの単一コピーを持つLAGLIDADGメガヌクレアーゼは、ホモダイマーを形成するが、一方、2コピーのLAGLIDADGモチーフを持つメンバーは、モノマーであることが判明した。同様に、GIY-YIGファミリーメンバーは、70 - 100残基長のGIY-YIGモジュールを持ち、その内二つは活性のために必要な、四つの不変残基を持つ、4から5個の保存配列モチーフを含む (VAN ROEY ET AL. (2002), NATURE STRUCT. BIOL. 9: 806 - 811を参照)。HIS-CYSボックス・メガヌクレアーゼは、数百のアミノ酸残基を含む領域に分散する、高度に保存された、ヒスチジンおよびシステインの連続列によって特徴づけられる (CHEVALIER ET AL. (2001), NUCLEIC ACIDS RES. 29 (18): 3757 - 3774を参照)。NHNファミリーの場合は、そのメンバーは、2対の、アスパラギン残基によって囲まれる保存ヒスチジンを含むモチーフによって定義される (CHEVALIER ET AL. (2001), NUCLEIC ACIDS RES. 29 (18): 3757 - 3774を参照)。メガヌクレアーゼのこの4ファミリーは、構造保存部位に関してDNA認識配列の特異性及び触媒活性が互いに大きく異なっている。

【0007】

主にLAGLIDADGファミリーから得られる、天然のメガヌクレアーゼは、これまで、植物、酵母、ショウジョウバエ、哺乳類細胞及びマウスにおいて部位特異的なゲノム修飾に有効に使用されているが、この方法は、このメガヌクレアーゼ認識配列を保存する相同配列か (MONNAT ET AL. (1999), BIOCHEM. BIOPHYS. RES. COMMUN. 255: 88 - 93) 又は認識配列を導入したあらかじめ加工されたゲノム (ROUET ET AL. (1994), MOL. CELL BIOL. 14: 8096 - 106; CHILTON ET AL. (2003), PLANT PHYSIOL. 133: 956 - 65; PUCHTA ET AL. (1996), PROC. NATL. ACAD. SCI. USA 93: 5055 - 60; RONG ET AL. (2002), GENES DEV. 16: 1568 - 81; GOUBLE ET AL. (2006), J. GENE MED. 8 (5): 616 - 622) の修飾に限定される。

【0008】

ヌクレアーゼ刺激による遺伝子修飾の体系的実施は、ゲノムの既存部位に対する標的DNA中断部に向けて照準させた特異性を持つ加工酵素の使用が必要となる。したがって、医学的又は生物工学的関連部位において遺伝子修飾を促進するようにメガヌクレアーゼを適応させることについては大きな関心が寄せられている (PORTEUS ET AL. (2005), NAT. BIOTECHNOL. 23: 967 - 73; SUSSMAN ET AL. (2004), J. MOL. BIOL. 342: 31 - 41; EPINAT ET AL. (2003), NUCLEIC ACIDS RES. 31: 2952 - 62)。

【0009】

CHLAMYDOMONAS REINHARDTIIから得られたメガヌクレアーゼI-CREIは、LAGLIDADGファミリーの一員であり、葉緑体染色体の中の2塩基対認識配列を認識・切断し、かつ、メガヌクレアーゼ再設計の好個の標的となっている

。この野生型酵素は、各モノマーが、全体認識配列における9塩基対に対して直接接触する、ホモダイマーである。この認識配列の単一位置において(SUSSMAN ET AL. (2004), J. MOL. BIOL. 342: 31-41; CHAMES ET AL. (2005), NUCLEIC ACIDS RES. 33: E178; SELIGMAN ET AL. (2002), NUCLEIC ACIDS RES. 30: 3870-9)、あるいは、比較的最近では、認識配列の中の3位置において(ARRNOULD ET AL. (2006), J. MOL. BIOL. 355: 443-58)、塩基選択性を変えるI-CREIにおける突然変異を特定するために、遺伝子選択技術が用いられている。I-CREIタンパク-DNAインターフェイスは、DNA塩基に直接接触する9個のアミノ酸と、修飾されたインターフェイスにおいて接触を形成する可能性のある、少なくともさらに5個の残基を含む。このインターフェイスのサイズは、切断部位を大きく変更された酵素を選択するために構築された配列ライブラリーにおいても十分に抽出されることがほとんどないと思われるほど複雑な組み合わせを提示する。

10

20

30

40

50

【発明の開示】

【発明が解決しようとする課題】

【0010】

ゲノムの正確な修飾を促進するヌクレアーゼについては、依然として需要がある。さらに、特定の遺伝子座位における遺伝子配列の操作を可能とする、あらかじめ指定された、合理的に設計された認識配列を有するヌクレアーゼを生成する技術、および、正確な配列修飾を持つ生物体を遺伝学的に加工するために、そのようなヌクレアーゼを利用する技術に対しては、依然として需要がある。

【課題を解決するための手段】

【0011】

本発明は、メガヌクレアーゼLAGLIDAGファミリーの同定及び特徴解明に基づく。メガヌクレアーゼが2本鎖DNA認識配列とコンタクトする際、DNA塩基およびDNAバックボーンに結合し、そうすることによって、該酵素の特異性および活性に影響を及ぼす。この発見は、以下に詳細に説明するように、メガヌクレアーゼの認識配列特異性及び/又はDNA結合親和性を変えることが可能なアミノ酸置換を特定するため及び天然のメガヌクレアーゼが認識しない所望のDNA配列を認識することが可能なメガヌクレアーゼを合理的に設計し、開発するために用いられた。本発明はさらに、遺伝子治療、病原体感染の治療及び診断と研究におけるインビトロ応用のために、生物体ゲノム内の少数座位において所望の遺伝子配列の組み換えを実現するために、このようなメガヌクレアーゼを使用する方法を提供する。

【0012】

したがって、ある実施態様では、本発明は、野生型I-CREIメガヌクレアーゼに比べ、少なくとも一つの認識配列半分部位に対し変更特異性を持つ組み換えメガヌクレアーゼを提供する。この実施態様におけるメガヌクレアーゼは、配列番号1の野生型I-CREIメガヌクレアーゼの残基2-153に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、該組み換えメガヌクレアーゼは、配列番号2、配列番号3、配列番号4及び配列番号5から選ばれる、I-CREIメガヌクレアーゼ認識配列内の半分部位と、少なくとも1塩基対異なる認識配列半分部位に対して特異性を持ち、かつ、この組み換えメガヌクレアーゼは、従来技術に認められる排外的修飾ではない、表5に掲げられる少なくとも一つの修飾を含む。

【0013】

別の実施態様では、本発明は、野生型I-MSOIメガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対し変更特異性を持つ組み換えメガヌクレアーゼを提供する。この実施態様におけるメガヌクレアーゼは、配列番号6のI-MSOIメガヌクレアーゼの残基6-160に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、該組み換えメガヌクレアーゼは、配列番号7および配列番号8から選ばれる、I-MSOIメガヌクレアーゼ認識配列内の半分部位と、少なくとも1塩基対異なる、認識配列

半分部位に対して特異性を持ち、かつ、該組み換えメガヌクレアーゼは、従来技術に認められる排外的修飾ではない、表7に掲げられる少なくとも一つの修飾を含む。

【0014】

別の実施態様では、本発明は、野生型 I - S C E I メガヌクレアーゼと比べ、認識配列に対し変更特異性を持つ組み換えメガヌクレアーゼを提供する。この実施態様におけるメガヌクレアーゼは、配列番号9の I - S C E I メガヌクレアーゼの残基3 - 186に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、この組み換えメガヌクレアーゼは、配列番号10および配列番号11の I - S C E I メガヌクレアーゼ認識配列と少なくとも1塩基対異なる認識配列に対して特異性を持ち、かつ、この組み換えメガヌクレアーゼは、従来技術に認められる排外的修飾ではない、表9に掲げられる少なくとも一つの修飾を含む。

10

【0015】

別の実施態様では、本発明は、野生型 I - C E U I メガヌクレアーゼと比べ、少なくとも一つの認識配列半分部位に対し変更特異性を持つ組み換えメガヌクレアーゼを提供する。この実施態様におけるメガヌクレアーゼは、配列番号12の I - C E U I メガヌクレアーゼの残基5 - 211に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、該組み換えメガヌクレアーゼは、配列番号13および配列番号14から選ばれる I - C E U I メガヌクレアーゼ認識配列内の半分部位と少なくとも1塩基対異なる認識配列半分部位に対して特異性を持ち、かつ、この組み換えメガヌクレアーゼは、従来技術に認められる排外的修飾ではない、表11に掲げられる少なくとも一つの修飾を含む。

20

【0016】

本発明のメガヌクレアーゼは、認識配列内の1、2又は3以上の位置において組み換えメガヌクレアーゼの配列特異性に影響を及ぼすために、本出願に開示される修飾の内の一つ、二つ又は三つ以上を含むことが可能である。メガヌクレアーゼは、本出願に開示される新規修飾のみを含んでもよいし又は本出願に開示される新規修飾を従来技術で認められる修飾と組み合わせて含んでもよい。しかしながら、特異的に排除されるものは、従来技術の修飾のみを含む組み換えメガヌクレアーゼである。

【0017】

他の側面から考えると、本発明は、2本鎖DNAに対する配列特異的ではない結合親和度を変更させた組み換えメガヌクレアーゼを提供する。これは、2本鎖DNA認識配列のバックボーンに接触する、メガヌクレアーゼ残基を修飾することによって実現される。この修飾は、結合親和度を増大又は低減するため、酵素の全体活性を増大又は低減することが可能である。さらに、結合および活性の上昇/低下は、配列特異性における低下/上昇を引き起こすことが認められた。したがって、本発明は、DNA - 結合親和度を変えることによって配列特異性を全体的に変えるための手段を提供する。

30

【0018】

したがって、ある実施態様では、本発明は、野生型 I - C R E I メガヌクレアーゼと比べ、2本鎖DNAに対する結合親和度を変更させた組み換えメガヌクレアーゼを提供する。この発明におけるメガヌクレアーゼは、配列番号1の I - C R E I メガヌクレアーゼの残基2 - 153に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、DNA結合親和度は、(1)(A)H、N、Q、S、T、K若しくはRによるE80、D137、I81、L112、P29、V64若しくはY66の置換、又は(B)K若しくはRによるT46、T140若しくはT143の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇するか、又は、逆に、(2)(A)H、N、Q、S、T、D若しくはEによるK34、K48、R51、K82、K116若しくはK139の置換、又は(B)D若しくはEによるI81、L112、P29、V64、Y66、T46、T140若しくはT143の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって低下する。

40

【0019】

別の実施態様では、本発明は、野生型 I - M S O I メガヌクレアーゼに比べ、2本鎖D

50

NAに対する結合親和度を変更させた組み換えメガヌクレアーゼを提供する。この発明におけるメガヌクレアーゼは、配列番号6のI-MSOIメガヌクレアーゼの残基6-160に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、DNA結合親和度は、(1)(A)H、N、Q、S、T、K若しくはRによるE147、I85、G86若しくはY118の置換、又は(B)K若しくはRによるQ41、N70、S87、T88、H89、Q122、Q139、S150若しくはN152の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇するか、又は、逆に、(2)(A)H、N、Q、S、T、D若しくはEによるK36、R51、K123、K143若しくはR144の置換、又は(B)D若しくはEによるI85、G86、Y118、Q41、N70、S87、T88、H89、Q122、Q139、S150若しくはN152の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって低下する。

10

【0020】

別の実施態様では、本発明は、野生型I-SCEIメガヌクレアーゼに比べ、2本鎖DNAに対する結合親和度を変更させた組み換えメガヌクレアーゼを提供する。この発明におけるメガヌクレアーゼは、配列番号9のI-SCEIメガヌクレアーゼの残基3-186に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、DNA結合親和度は、(1)(A)H、N、Q、S、T、K若しくはRによるD201、L19、L80、L92、Y151、Y188、I191、Y199若しくはY222の置換、又は(B)K若しくはRによるN15、N17、S81、H84、N94、N120、T156、N157、S159、N163、Q165、S166、N194若しくはS202の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇するか、又は、逆に、(2)(A)H、N、Q、S、T、D若しくはEによるK20、K23、K63、K122、K148、K153、K190、K193、K195若しくはK223の置換、又は(B)D若しくはEによるL19、L80、L92、Y151、Y188、I191、Y199、Y222、N15、N17、S81、H84、N94、N120、T156、N157、S159、N163、Q165、S166、N194若しくはS202の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって低下する。

20

【0021】

別の実施態様では、本発明は、野生型I-CEUIメガヌクレアーゼに比べ、2本鎖DNAに対する結合親和度を変更させた組み換えメガヌクレアーゼを提供する。この発明におけるメガヌクレアーゼは、配列番号12のI-CEUIメガヌクレアーゼの残基5-211に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、DNA結合親和度は、(1)(A)H、N、Q、S、T、K若しくはRによるD25若しくはD128の置換、又は(B)K若しくはRによるS68、N70、H94、S117、N120、N129若しくはH172の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇するか、又は、逆に、(2)(A)H、N、Q、S、T、D若しくはEによるK21、K28、K31、R112、R114若しくはR130の置換、又は(B)D若しくはEによるS68、N70、H94、S117、N120、N129若しくはH172の置換から選ばれる一つの置換に対応する少なくとも一つの修飾によって低下する。

30

40

【0022】

本発明のメガヌクレアーゼは、DNA-結合親和度に影響を及ぼすために、本出願に開示される、バックボーン接触残基の修飾の内、一つ、二つ又は三つ以上を含むことが可能である。さらに、DNA-結合親和度に影響を及ぼすこれらの修飾は、認識配列内部の特定位置における組み換えメガヌクレアーゼの配列特異性を変える、前述の塩基接触残基の新規修飾の内一つ以上と組み合わせてもよいし、前述の従来技術の修飾と組み合わせてもよいし、新規修飾と従来技術の修飾との組み合わせと組み合わせてもよい。特に、バックボーン接触修飾および塩基接触修飾を組み合わせることによって、組み換えメガヌクレアーゼを、所望の特異性および活性を持つように合理的に設計することが可能である。例えば、塩基接触残基に対する設計変更由来する親和度の損失を逆転するDNA-結合親

50

和度の増加を設計することが可能であり、配列特異性を減少させ酵素に対する認識配列の組を広げる親和度減少を設計することも可能である。

【0023】

他の側面から考えると、本発明は、ホモ又はヘテロダイマー形成に対する親和度を変更させた合理的設計メガヌクレアーゼモノマーを提供する。ダイマー形成に対する親和度は、同じモノマー（すなわち、ホモダイマー形成）によって又は異なるモノマー（すなわち、ヘテロダイマー形成）、例えば、参照野生型メガヌクレアーゼによって測定することが可能である。これらの組み換えメガヌクレアーゼは、メガヌクレアーゼダイマーにおけるモノマー間のタンパク質とタンパク質との境界面に存在するアミノ酸残基に対する修飾を有する。これらの修飾は、ヘテロダイマー形成の促進及び非パリンドローム認識配列を持つメガヌクレアーゼの創製に使用することが可能である。

10

【0024】

したがって、他の実施態様では、本発明は、参照メガヌクレアーゼモノマーによるダイマー形成に対する親和度を変更させた組み換えメガヌクレアーゼを提供する。この実施態様では、組み換えモノマーは、配列番号1のI-CREIメガヌクレアーゼの残基2-153に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、ダイマー形成に対する親和度は、(A)D若しくはEによるK7、K57若しくはK96の置換、又は(B)K若しくはRによるE8若しくはE61の置換から選ばれる置換に対応する少なくとも一つの修飾によって変えられる。このような組み換えモノマーに基づき、本発明はさらに、(1)配列番号1のI-CREIメガヌクレアーゼの残基2-153に対して少なくとも85%の配列類似性を有する第1ポリペプチドであって、ダイマー形成に対する親和度が、(A)D若しくはEによるK7、K57若しくはK96の置換から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第1ポリペプチド、及び(2)配列番号1のI-CREIメガヌクレアーゼの残基2-153に対して少なくとも85%の配列類似性を有する第2ポリペプチドであって、ダイマー形成に対する親和度が、(B)K若しくはRによるE8若しくはE61の置換から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第2ポリペプチドを含む組み換えメガヌクレアーゼ・ヘテロダイマーを提供する。

20

【0025】

他の実施態様では、本発明は、参照メガヌクレアーゼモノマーによるダイマー形成に対する親和度を変更させた組み換えメガヌクレアーゼモノマーを提供する。この実施態様では、組み換えモノマーは、配列番号6のI-MSOIメガヌクレアーゼの残基6-160に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、ダイマー形成に対する親和度は、(A)D若しくはEによるR302の置換、又は(B)K若しくはRによるD20、E11若しくはQ64の置換から選ばれる置換に対応する少なくとも一つの修飾によって変えられる。このような組み換えモノマーに基づき、本発明はさらに、(1)配列番号6のI-MSOIメガヌクレアーゼの残基6-160に対して少なくとも85%の配列類似性を有する第1ポリペプチドであって、ダイマー形成に対する親和度が、(A)D若しくはEによるR302の置換から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第1ポリペプチド、及び(2)配列番号6のI-MSOIメガヌクレアーゼの残基6-160に対して少なくとも85%の配列類似性を有する第2ポリペプチドであって、ダイマー形成に対する親和度が、(B)K若しくはRによるD20、E11若しくはQ64の置換から選ばれる置換に対応する少なくとも一つの修飾によって変えられる第2ポリペプチドを含む組み換えメガヌクレアーゼ・ヘテロダイマーを提供する。

30

40

【0026】

他の実施態様では、本発明は、参照メガヌクレアーゼモノマーによるダイマー形成に対する親和度を変更させた組み換えメガヌクレアーゼモノマーを提供する。この実施態様では、組み換えモノマーは、配列番号12のI-CEUIメガヌクレアーゼの残基5-211に対して少なくとも85%の配列類似性を有するポリペプチドを含むが、ダイマー形成に対する親和度は、(A)DまたはEによるR93の置換、又は(B)KまたはRによる

50

E 1 5 2 の置換から選ばれる置換に対応する少なく共一つの修飾によって変えられる。このような組み換えモノマーに基づき、本発明はさらに、(1) 配列番号 1 2 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対して少なくとも 8 5 % の配列類似性を有する第 1 ポリペプチドであって、ダイマー形成に対する親和度が、(A) D または E による R 9 3 の置換から選ばれる置換に対応する少なく一つの修飾によって変えられる第 1 ポリペプチド、及び(2) 配列番号 1 2 の I - C E U I メガヌクレアーゼの残基 5 - 2 1 1 に対して少なくとも 8 5 % の配列類似性を有する第 2 ポリペプチドであって、ダイマー形成に対する親和度が、(B) K または R による E 1 5 2 の置換から選ばれる置換に対応する少なく一つの修飾によって変えられる第 2 ポリペプチドを含む組み換えメガヌクレアーゼ・ヘテロダイマーを提供する。

10

【 0 0 2 7 】

ダイマー形成に対する親和度を変更させた、組み換えメガヌクレアーゼモノマー又はヘテロダイマーはさらに、前述の塩基接触残基の修飾の内の一つ、二つまたは三つ以上；前述のバックボーン接触残基の修飾の内の一つ、二つまたは三つ以上、又は両者の組み合わせを含むことが可能である。したがって、例えば、モノマーの塩基接触部は、配列特異性を変えるように修飾することが可能であり、モノマーのバックボーン接触部は、DNA 結合親和度を変えるように修飾することが可能であり、タンパク質とタンパク質との境界面は、ダイマー形成に影響を及ぼすように修飾することが可能である。このような組み換えモノマーは、同様に修飾されたモノマーと組み合わせて、所望の配列特異性および活性を有する合理設計メガヌクレアーゼ・ヘテロダイマーを生産することが可能である。

20

【 0 0 2 8 】

他の側面から考えると、本発明は、本発明において記載され可能とされた合理設計メガヌクレアーゼのために使用される各種方法を提供する。これらの方法は、遺伝学的に修飾された細胞および生物体を生産すること、病気を遺伝子治療によって治療すること、病原体感染を治療すること及び診断学と研究用とのインビトロ応用のために組み換えメガヌクレアーゼを使用することを含む。

【 0 0 2 9 】

したがって、この側面では、本発明は、染色体に挿入された興味の外因性配列を含む遺伝学的に修飾された真核細胞を生産するための方法であって、これを、該細胞を、(I) 本発明のメガヌクレアーゼをコードする第 1 核酸配列、及び(I I) 前記興味のある配列を含む第 2 核酸配列によってトランスフェクトすることによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、興味のある配列は、相同組み換え又は非相同的末端接合によって、染色体中の切断部位に挿入される。

30

【 0 0 3 0 】

それとは別に、他の側面では、本発明は、染色体に挿入された興味のある外来配列を含む遺伝学的に修飾された真核細胞を生産するための方法であって、この細胞に本発明のメガヌクレアーゼタンパクを挿入すること及びこの細胞を興味のある配列を含む核酸によってトランスフェクトすることによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、興味のある配列は、相同組み換え又は非相同的末端接合によって、切断部位において染色体の中に挿入される。

40

【 0 0 3 1 】

他の側面では、本発明は、染色体中の標的配列を破壊することによって遺伝学的に修飾された真核細胞を生産するための方法であって、この細胞を本発明のメガヌクレアーゼタンパクをコードする核酸によってトランスフェクトすることによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、標的配列は切断部位における非相同的末端接合によって破壊される。

【 0 0 3 2 】

他の側面では、本発明は、遺伝学的に修飾された生物体を生産するための方法であって、前述の方法に従って遺伝学的に修飾された真核細胞を生産し、かつ、この遺伝学的に修飾された真核細胞を育成して遺伝学的に修飾された生物体を生産することによって実現す

50

る方法を提供する。これらの実施態様では、真核細胞は、配偶子、接合子、胚盤胞細胞、胚幹細胞及びプロトプラスト細胞から選ぶことが可能である。

【0033】

他の側面では、本発明は、真核生物において遺伝子治療によって疾患を治療するための方法であって、この真核生物の少なくとも一つの細胞を、(I)本発明のメガヌクレアーゼをコードする第1核酸配列、及び(II)興味の配列を含む第2核酸配列によってトランスフェクトすることによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、興味の配列は、相同組み換え又は非相同的末端接合によって染色体の中に挿入され、興味の配列の挿入は、疾患に対する遺伝子治療をもたらす。

【0034】

それとは別に、他の側面では、本発明は、真核生物において遺伝子治療によって疾患を治療するための方法であって、この真核生物の少なくとも一つの細胞の中に本発明のメガヌクレアーゼを導入すること及びこの細胞を、興味の配列を含む核酸によってトランスフェクトすることによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、興味の配列は、相同組み換え、または非相同的末端接合によって切断部位において染色体の中に挿入され、興味の配列の挿入は、該疾患に対する遺伝子治療をもたらす。

【0035】

他の側面では、本発明は、真核生物において、該真核生物の染色体における標的配列を破壊する遺伝子治療によって疾患を治療するための方法であって、この真核生物の少なくとも一つの細胞を、本発明のメガヌクレアーゼをコードする核酸によってトランスフェクトすることによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、標的配列は、相同組み換え又は非相同的末端接合によって破壊され、該標的配列の破壊が、該疾患に対する遺伝子治療をもたらす。

【0036】

別の局面では、本発明は、真核細胞宿主において、ウィルス性または前核細胞性病原体感染を治療するための方法であって、この病原体のゲノムにおける標的配列を破壊することによって実現する方法を提供する。その際、メガヌクレアーゼは、染色体の中に切断部位を生産し、標的配列は、(1)切断部位における非相同的末端接合によって、又は(2)第2核酸との相同組み換えによって破壊され、かつ、標的配列の破壊は、この感染に対する治療をもたらす。

【0037】

さらに一般的に、他の側面では、本発明は、認識配列の少なくとも一つの塩基位置に対する特異性を変更させた組み換えメガヌクレアーゼを合理的に設計する方法であって、(1)参照メガヌクレアーゼ-DNA複合体の三次元構造の少なくとも一部を決定すること、(2)該塩基位置において塩基接触面を形成するアミノ酸残基を特定すること、(3)接触面の少なくとも第1残基の炭素と、該塩基位置における少なくとも第1塩基との間の距離を決めること及び(4)(A)第1塩基から6オングストローム未満の距離にある第1残基については、G、C、T若しくはAの内の適切なもののメンバーである基1及び/又は基2の置換を選択すること；又は(B)前記第1塩基から6オングストロームより遠い位置にある第1残基については、G、C、T及びAの内の適切なもののメンバーである基2及び/又は基3の置換を選択することによって、所望の変化を促進するアミノ酸置換を特定することによって実現する方法を提供する。なお、上述したそれぞれの基は、後述するように定義する。この方法は、同じ塩基に対する別の接触残基についても、同じ位置における他の塩基に対する接触残基についても、さらに、別の位置についても繰り返してよい。

【0038】

さらに、他の側面では、本発明は、DNA-結合親和度を増大させた組み換えメガヌクレアーゼを合理的に設計する方法であって、(1)基準メガヌクレアーゼ-DNA複合体の三次元構造の少なくとも一部を決定すること、(2)バックボーン接触面を形成するア

10

20

30

40

50

ミノ酸接触残基を特定すること、(3)(A)陰性荷電若しくは疎水性側鎖を持つ接触残基について、非荷電/極性若しくは陽性荷電側鎖を持つ置換を選択すること;又は(B)非荷電/極性側鎖を持つ接触残基について、陽性荷電側鎖を持つ置換を選択することによって、DNA-結合親和度を増大させるアミノ酸置換を特定すること、によって実現する方法を提供する。逆に、本発明は、DNA-結合親和度を減少させた組み換えメガヌクレアーゼを合理的に設計する方法であって、これを、(1)基準メガヌクレアーゼ-DNA複合体の三次元構造の少なくとも一部を決定すること、(2)バックボーン接触面を形成するアミノ酸接触残基を特定すること、(3)(A)陽性荷電側鎖を持つ接触残基について、非荷電/極性若しくは陰性荷電側鎖を持つ置換を選択すること、又は(B)疎水性及び非荷電/極性側鎖を持つ接触残基について、陰性荷電側鎖を持つ置換を選択することによって、DNA-結合親和度を減少させるアミノ酸置換を特定することによって実現する方法を提供する。

10

【0039】

本発明の、上記、およびその他の側面および実施態様は、本発明の下記の詳細な説明に基づき当業者には明白であろう。

【発明を実施するための最良の形態】

【0040】

1.1 序論

本発明は、一部は、メガヌクレアーゼが2本鎖DNA認識配列とコンタクトする際、DNA塩基と特異的接触を形成し、かつ、DNAバックボーンと非特異的接触を形成し、そうすることによって、該酵素の認識配列特異性およびDNA-結合親和度に影響を及ぼす、メガヌクレアーゼLAGLIDAGファミリーにおける特異的アミノ酸残基の特定および特徴解明に基づく。この発見は、以下に詳細に説明するように、メガヌクレアーゼの特異性及び/又は親和性を変えることが可能な、メガヌクレアーゼにおけるアミノ酸置換を特定するため及び天然のメガヌクレアーゼが認識しない所望のDNA配列を認識することが可能であり、及び/又は、天然のメガヌクレアーゼに比べて特異性及び/又は親和度が増大または減少させたメガヌクレアーゼを合理的に設計し、開発するために用いられた。さらに、DNA-結合親和度は、酵素活性の外に、配列特異性にも影響を及ぼすので、本発明は、天然メガヌクレアーゼに比べ活性を変更させた、合理的設計メガヌクレアーゼを提供する。さらに、本発明は、ダイマーを形成するためにコンタクトされたモノマー間のインターフェイスにおける残基が、ヘテロダイマー形成を促進するように修飾された合理設計メガヌクレアーゼを提供する。最後に、本発明は、組み換え細胞および生物体の生産における使用ばかりでなく、本出願に開示されるように、遺伝子療法、抗病原体、抗癌及びインビトロ応用における合理設計メガヌクレアーゼの使用を提供する。

20

30

【0041】

一般的出来事として、本発明は、(1)2本鎖DNA認識配列中の個々の塩基に対する配列特異的結合、又は(2)2本鎖DNA分子のフォスフォジエステルバックボーンに対する非特異的結合による、メガヌクレアーゼ内の複数部位において変更アミノ酸残基を含む、合理設計LAGLIDAGメガヌクレアーゼを生成するための方法を提供する。しかしながら、酵素活性は、DNA-結合親和度と相関するのであるから、DNA認識配列に対する結合に与るアミノ酸を変えることは、2本鎖DNAに対する全体的結合親和度を増大又は減少することになり、特異的塩基対相互作用を介するメガヌクレアーゼの特異性ばかりでなく、メガヌクレアーゼの活性も変える可能性がある。同様に、DNAバックボーンに対する結合によるアミノ酸を変えることは、2本鎖DNAに対する全体的結合親和度を増大又は減少することによって、酵素の活性ばかりでなく、認識配列に対する結合の特異性又は縮重の程度を変える可能性がある。

40

【0042】

下記に詳述するように、メガヌクレアーゼの合理的設計法は、DNA認識/結合によるアミノ酸を特定すること及び適切なアミノ酸変化を選択するために一連の規則を適用することを含む。この規則は、メガヌクレアーゼの配列特異性に関して、メガヌクレアーゼの

50

アミノ酸側鎖とDNAのセンス鎖及びアンチセンス鎖における塩基の間とのメガヌクレアーゼ-DNA複合体距離に関する互いの立体位置、アミノ酸側鎖とDNA塩基との相対位置及び非共有的化学的相互作用を考慮に含んでいる。

【0043】

最後に、ホモダイマーとしてDNAに結合する天然メガヌクレアーゼの大部分は、擬似、又は完全パリンドロームの認識配列を認識する。長々としたパリンドロームはめったにないと予想されるので、ゲノムの興味の配列部位でパリンドロームに遭遇する確率はきわめて低い。したがって、これらの酵素を、ゲノムの興味の配列部位を認識するように再設計しなければならないとすると、異なる半分部位を認識する二つの酵素モノマーにおいて、ヘテロダイマー形成されると非パリンドローム的ハイブリッド認識配列を切断することが可能な二つの酵素モノマーを設計することが必要である。それゆえ、ある側面では、本発明は、少なくとも一つのアミノ酸位置で異なる二つのモノマーをダイマー化して、ヘテロダイマーを形成して得られる合理設計メガヌクレアーゼを提供する。ある場合、両モノマーは、非パリンドローム的認識配列を認識するヘテロダイマーを形成するように合理的設計される。二つの異なるモノマーから成る混合物は、最大三つの活性形メガヌクレアーゼダイマー、二つのホモダイマー及びヘテロダイマーをもたらす。それに加えて、またはそれとは別に、ある場合、ホモダイマーまたはヘテロダイマー形成の確率を増すか、または減らすために、モノマー同士が相互作用を持ってダイマーを形成する境界面においてアミノ酸残基が変えられる。

10

【0044】

したがって、一つの側面では、本発明は、LAGLIDADGメガヌクレアーゼにおいて、酵素の特異性及び/又は活性を変えるアミノ酸変化を含むLAGLIDADGメガヌクレアーゼの合理的設計法を提供する。他の側面では、本発明は、これらの方法によってもたらされる合理設計メガヌクレアーゼを提供する。他の側面では、本発明は、生物体ゲノム内部の所望のDNA配列または遺伝子座が、DNA配列の挿入、欠失、置換、またはその他の操作によって修飾される、組み換え核酸および生物体を生産するために、このような合理設計メガヌクレアーゼを使用する方法を提供する。さ他の側面では、本発明は、病原体または癌細胞において、病原体特異的又は癌特異的認識配列を有する合理設計メガヌクレアーゼを用いて、この病原体または癌細胞の生存率を下げるための方法を提供する。

20

30

【0045】

1.2 参照および定義

本出願において参照される特許および科学文献は、当業者にとって利用可能な知識を確定する。本出願において引用される、公刊された米国特許、承認済みの出願、出版された外国特許、および、GENBANKのデータベース配列を含む参照資料は、参照により、あたかも、それぞれが、特異的、個別的に、参照によって含まれるのと同程度に、本出願に含まれる。

【0046】

本出願で用いる「メガヌクレアーゼ」という用語は、12塩基対よりも大きい認識配列で2本鎖DNAに結合するエンドヌクレアーゼを指す。天然のメガヌクレアーゼはモノマー（例えば、I-SceI）又はダイマー（例えば、I-CreI）であることが可能である。本出願で用いるメガヌクレアーゼという用語は、モノマー・メガヌクレアーゼ、ダイマー・メガヌクレアーゼ、または、コンタクトしてダイマー形メガヌクレアーゼを形成するモノマーを指す。「ホーミング・エンドヌクレアーゼ」という用語は、「メガヌクレアーゼ」という用語と同義である。

40

【0047】

本出願で用いる「LAGLIDADGメガヌクレアーゼ」という用語は、天然ではダイマー形である単一のLAGLIDADGモチーフを含むメガヌクレアーゼ、又は天然ではモノマー形である二つのLAGLIDADGモチーフを含むメガヌクレアーゼを指す。本出願で用いる「モノLAGLIDADGメガヌクレアーゼ」という用語は、単一のLAG

50

L I D A D Gモチーフを含むメガヌクレアーゼを指し、本出願で用いる「ジ L A G L I D A D Gメガヌクレアーゼ」という用語は、二つを区別することが必要な場合には、二つの L A G L I D A D Gモチーフを含むメガヌクレアーゼを指す。L A G L I D A D Gモチーフを含むジ L A G L I D A D Gメガヌクレアーゼの二つの構造ドメインは、それぞれ、L A G L I D A D Gサブユニットと呼ぶことが可能である。

【0048】

本出願で用いる「合理的に設計された」という用語は、非天然及び/又は遺伝子工学的に加工されたことを意味する。本発明の合理的設計メガヌクレアーゼは、野生型又は天然のメガヌクレアーゼとアミノ酸配列または一次構造において異なり、さらに、二次、三次又は四次構造において異なる場合がある。加えて、本発明の合理設計メガヌクレアーゼは、認識配列の特異性及び/又は活性においても、野生型又は天然メガヌクレアーゼと異なる。

10

【0049】

タンパク質に関して本出願で用いる「組み換え」という用語は、タンパク質をコードする核酸に対し遺伝子工学技術を適用した結果アミノ酸配列が変化すること、及びこのタンパク質を発現する細胞又は生物体が見られることを意味する。核酸に関して「組み換え」という用語は、遺伝子工学技術を適用した結果、核酸配列が変化することを意味する。遺伝子工学技術としては、これらに限定されるものではないが、PCR、トランスフェクション、形質転換及びその他の遺伝子転移技術等のDNAクローニング技術、相同組み換え、部位特異的変異処理及び遺伝子融合が含まれる。この定義によれば、天然のタンパクと同じアミノ酸配列を持つが、異種宿主におけるクローニング及び発現によって生産されるタンパク質は、「組み換え」とは見なされない。

20

【0050】

組み換えタンパクに関して本出願で用いられる「修飾」という用語は、基準配列（例えば、野生型）と比べた場合の、組み換え配列におけるアミノ酸残基の、どのようなものであれ、挿入、欠失又は置換を含む意味である。

【0051】

本出願で用いる「遺伝学的に修飾された」という用語は、そのゲノムDNA配列が、組み換え技術によって意図的に修飾されたゲノムDNA配列を有する細胞、生物体又は祖先を指す。本出願で用いる「遺伝学的に修飾された」という用語は、「トランスジェニック」という用語を含む。

30

【0052】

本出願で用いる「野生型」という用語は、どのようなものであれ、メガヌクレアーゼの全ての天然形を指す。「野生型」という用語は、天然における酵素の最も一般的な変異を意味するものではなく、自然で見られるいかなる変異も意味する。野生型メガヌクレアーゼは、組み換え又は非天然メガヌクレアーゼと区別される。

【0053】

本出願で用いる「認識配列半分部位」又は単に「半分部位」は、モノ-L A G L I D A D Gメガヌクレアーゼのモノマー又はジ-L A G L I D A D Gメガヌクレアーゼの一つのL A G L I D A D Gサブユニットによって認識される2本鎖DNA分子間の核酸配列を意味する。

40

【0054】

本出願で用いる「認識配列」という用語は、モノ-L A G L I D A D Gメガヌクレアーゼダイマー又はジ-L A G L I D A D Gメガヌクレアーゼモノマーによって結合され、切断される一对の半分部位を指す。この二つの半分部位は、この酵素によって特異的に認識されない塩基対によって隔てられてもよいし、隔てられなくともよい。I - C R E I、I - M S O I及びI - C E U Iの場合、各モノマーの認識配列半分部位は9塩基対の幅を持ち、この二つの半分部位は、特異的には認識されないが実際の切断部位である塩基対（4塩基対のオーバーハングを有する）によって分割されている。したがって、I - C R E I、I - M S O I及びI - C E U Iメガヌクレアーゼダイマーの結合認識配列は、通常、4

50

塩基対切断部位に側接する二つの 9 塩基対を含む 2 2 塩基対の幅を持つ。各半分部位の塩基対は - 9 から - 1 と表示され、- 9 位は切断部位からもっとも遠位であり、- 1 位は N 1 - N 4 と表示される 4 個の中央塩基対に隣接する。- 9 から - 1 へと向かう方向（すなわち、切断部位に向かう方向）において 5 から 3 に方向づけられる各半分部位鎖は、「センス」鎖と表示され、反対鎖は、「アンチセンス鎖」と表示される。ただし、いずれの鎖もタンパクをコードしない。したがって、図 1 (A) に示すように、一方の半分部位の「センス」鎖は、他方の半分部位のアンチセンス鎖である。ジ - L A G L I D A D G メガヌクレアーゼモノマーである I - S C E I メガヌクレアーゼの場合、認識配列は約 1 8 B P の非パリンドローム配列であり、特異的には認識されない中央塩基対は無い。通例に従って、2 本鎖の一方は「センス」鎖と呼ばれ、他方は「アンチセンス」鎖と呼ばれるが、いずれの鎖もタンパクをコードしない。

10

【 0 0 5 5 】

本出願で用いる「特異性」という用語は、メガヌクレアーゼが認識配列と呼ばれる特定の塩基対配列又は特定の一組の認識配列においてのみ、2 本鎖 DNA 分子を認識し切断する能力を意味する。この一組の認識配列は、いくつかの保存位置または配列モチーフを共有するが、これらは 1 又は 2 以上の位置で縮重する場合がある。高い特異性を有するメガヌクレアーゼは、たった一つ又はきわめて少数の認識配列の切断が可能である。特異性は実施例 1 に記載される切断アッセイで決めることが可能である。本出願で用いる場合、メガヌクレアーゼは、基準メガヌクレアーゼ（例えば、野生型）によって結合及び切断されない認識配列に結合し、切断する場合、又は、認識配列の切断速度が、基準メガヌクレアーゼに比べて、統計的に有意な値 ($P < 0.05$) だけ上昇又は降下した場合に、特異性が「変更された」ことになる。

20

【 0 0 5 6 】

本出願で用いる「縮重」という用語は、「特異性」の反対を意味する。高縮重したメガヌクレアーゼは、非常に多岐にわたる認識配列を切断することができる。メガヌクレアーゼは単一の、複数の又は全ての箇所の半分部位内に配列縮重を有することができる。このような配列縮重は、(I) メガヌクレアーゼの DNA 結合ドメインにおける任意のアミノ酸が、認識配列内の 1 又は 2 以上の位置で任意の塩基と接触できないこと、(I I) メガヌクレアーゼの DNA 結合ドメインにおける 1 又は 2 以上のアミノ酸が、認識配列内の 1 又は 2 以上の位置で 1 以上の塩基と特異的にコンタクトできること、及び / 又は (I I I) 活性に十分なほど、非特異的 DNA 結合親和性を持つこと、からもたらされる可能性がある。「完全な」縮重位置は任意の 4 塩基で占められてもよく、半分部位内では " N " で表示することができる。「部分的な」縮重位置は、4 塩基の内の任意の 2 又は 3 塩基で占められてもよい。(例えば、プリン (P U) が、ピリミジン (P Y) で占められるが、G には占められない)。

30

【 0 0 5 7 】

メガヌクレアーゼに関して本出願で使用される場合、「DNA - 結合親和度」又は「結合親和度」は、メガヌクレアーゼが基準 DNA 分子（例えば、認識配列又は任意の配列）と非共有的に結合する傾向を意味する。結合親和度は、解離定数 K_D によって測定される（例えば、WT 認識配列に対する I - C R E I の K_D は約 0.1 nM である）。本出願で用いる場合、基準認識配列に対する組み換えメガヌクレアーゼの K_D が、基準メガヌクレアーゼに比べて、統計的に有意な値 ($P < 0.05$) だけ増加又は減少した場合に、結合親和度を「変更させた」ことになる。

40

【 0 0 5 8 】

メガヌクレアーゼモノマーに関して本出願で用いる場合、「ダイマー形成に対する親和度」という用語は、メガヌクレアーゼモノマーが、基準メガヌクレアーゼモノマーと非共有的に結合する傾向を意味する。ダイマー形成に対する親和度は、例えば、基準野生型メガヌクレアーゼのような同じモノマー（すなわち、ホモダイマー形成）又は異なるモノマー（すなわち、ヘテロダイマー形成）で測定することができる。結合親和度は、解離定数 K_D によって測定される。本出願で用いる場合、基準認識配列に対する組み換えメガヌク

50

レアーゼのKDが、基準メガヌクレアーゼに比べて、統計的に有意な値 ($P < 0.05$) だけ増加又は減少した場合に、結合親和度を「変更させた」ことになる。

【0059】

本出願で用いる「パリンδροーム的」という用語は、同じ半分部位の逆方向反復を含む認識配列を指す。しかしながら、この場合、パリンδροーム配列は酵素にコンタクトされない四つの中央塩基対に関してパリンδροーム的である必要はない。ダイマー形メガヌクレアーゼの場合、パリンδροームDNA配列は二つのモノマーが同じ半分部位にコンタクトするホモダイマーによって認識される。

【0060】

本出願で用いる「擬似パリンδροーム的」という用語は、非同一又は不完全なパリンδροーム的半分部位の逆方向反復を含む認識配列を指す。この場合、擬似パリンδροーム配列は、四つの中央塩基対に関してパリンδροーム的である必要はなく、二つの半分部位の間のパリンδροーム配列からもはずれることが可能である。擬似パリンδροームDNA配列は、二つの同一の酵素モノマーが、異なる半分部位にコンタクトする、野生型ホモダイマー形メガヌクレアーゼによって認識される天然DNA部位において典型的に見られる。

10

【0061】

本出願で用いる「非パリンδροーム的」という用語は、メガヌクレアーゼの、二つの、無関係の半分部位から構成される認識配列を指す。この場合、非パリンδροーム配列は、四つの中央塩基対に対しても、二つのモノマー半分部位に対しても、パリンδροーム的である必要はない。非パリンδροームDNA配列は、ジ-LAGLIDADGメガヌクレアーゼ、高縮重モノ-LAGLIDADGメガヌクレアーゼ(例えば、I-CeuI)、非同一半分配列を認識するモノ-LAGLIDADGメガヌクレアーゼのヘテロダイマーのいずれかによって認識される。

20

【0062】

本出願で用いる「活性」という用語は、本発明のメガヌクレアーゼが、特定の認識配列を切断する速度を指す。このような活性は、2本鎖DNAのフォスフォジエステル結合の加水分解に関する測定可能な酵素反応である。ある特定のDNA基質に作用するメガヌクレアーゼの活性は、その特定のDNA基質に対するメガヌクレアーゼの親和力又は結合力に影響され、DNAとの配列特異的相互作用及び非配列特異的相互作用によって影響される。

30

【0063】

本出願で用いる「相同組み換え」という用語は、2本鎖DNA切断が、修復鋳型として相同DNA配列を用いて修復される、天然の、細胞プロセスを指す(例えば、CAHILL ET AL. (2006), FRONT. BIOSCI. 11:1958-1976を参照)。相同のDNA配列は、内因性の染色体配列であってもよいし、細胞に輸送された外来性の核酸であってもよい。したがって、ある実施態様では、合理設計メガヌクレアーゼは、標的配列内の認識配列を切断するために使用され、標的配列に対して相同な又はほぼ同様の配列類似性を持つ外来性の核酸が細胞に輸送されて、相同性組み換えによる修復のための鋳型として使用される。したがって、標的配列とは著明に異なってもよい外来性のDNA配列が、染色体配列に組み込まれる。相同組み換えプロセスは、主に真核生物において起こる。本出願では、「相同性」という用語は「配列類似性」と等価的に用いられ、出自又は系統関係による同一性を要求することを意図するものではない。

40

【0064】

本出願で用いる「非相同的末端接合」という用語は、2本鎖DNA切断が、2本の非相同的DNAセグメントの直接的接合によって修復される、天然の細胞プロセスを指す(例えば、CAHILL ET AL. (2006), FRONT. BIOSCI. 11:1958-1976)。非相同的末端接合によるDNA修復は、間違いを起こしやすく、修復部位において鋳型と対応しないDNA配列の付加又は除去をもたらすことが多い。したがって、ある実施態様では、(例えば、塩基挿入、塩基欠失又はフレームシフト突然変異を行うことによって)非相同的末端接合によって遺伝子を破壊し、標的配列内のメガヌ

50

クレーゼ認識配列中に二本鎖切断点を作るために、合理設計メガヌクレーゼを用いることも可能である。他の実施態様では、標的配列に対する相同性を欠如するか又は標的配列に対する実質的配列類似性を欠如する外来性の核酸が、非相同的末端接合によって、メガヌクレーゼ刺激二本鎖DNA切断部位で取り込まれてもよい(例えば、SALOMON, ET AL. (1998), EMBO J. 17:6086-6095を参照)。非相同的末端接合プロセスは、真核生物でも、細菌のような原核生物でも起こる。

【0065】

本出願で用いる「興味の配列」という用語は、メガヌクレーゼタンパクを用いてゲノムへ挿入するか又はゲノムDNA配列を置換するのに使用することが可能な核酸配列であって、タンパク、RNA又は調整エレメント(例えば、エンハンサー、サイレンサー、またはプロモーター配列)のどれをコードするにはよらない任意の核酸配列を意味する。興味の配列は、興味の配列から発現されるタンパク質又はRNAの標識化を可能とする非相同的なDNA配列を有することも可能である。例えば、以下のものに限定されるものではないが、タンパク質は、エピトープ(例えば、C-MYC、FLAG)、または他のリガンド(例えば、ポリ-HIS)を含むタグによって標識することが可能である。更に、興味の配列は、従来技術で周知の技術に従って融合タンパクをコードしてもよい(例えば、AUSUBEL ET AL., 「分子生物学における最近のプロトコール(“CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, ”)」、WILEY 1999)。ある実施態様では、隣接配列は、組み換えメガヌクレーゼによって切断され、ゲノム組み替え配列の中に興味の配列の適切な挿入される。

10

20

【0066】

したがって、この側接配列は切断され、該組み換えメガヌクレーゼによって切断されるゲノム認識配列の中に、興味の配列を適正に挿入することを可能となる。ある実施態様では、興味の配列の全ては、ゲノム中の標的配列と相同であるか又は標的配列と実質的に配列類似性を有しており、相同組み換えによって、標的配列が興味の配列によって効果的に置換される。他の実施態様では、興味の配列は、標的配列と相同性を有するか又は実質的配列類似性を有するDNA配列によって側接され、相同組み換えによってゲノム内の標的配列座位において興味の配列が挿入される。ある実施態様では、興味の配列は、メガヌクレーゼ認識配列における突然変異又は他の修飾を除いては、標的配列とほぼ同一であり、そのため、メガヌクレーゼは、標的配列が興味の配列によって修飾された後に、標的配列を切断することができない。

30

【0067】

アミノ酸配列及び核酸配列に関して本出願で用いる場合、「類似性パーセント」及び「配列類似性」という用語は、アライメントされたアミノ酸残基又はヌクレオチド間の類似性を最大にするようにアライメントされた配列に基づく2つの配列の類似度の測定値を意味する。そしてこれは、アライメント配列中のギャップ及びギャップの数、核酸又は残基の総数、同一又は類似する核酸又は残基の数の関数による。一般的なパラメータを用いて配列類似性を決めるために、様々のアルゴリズムおよびコンピュータプログラムの利用が可能である。本出願では、配列類似度は、アミノ酸配列の場合はBLASTPプログラムを用い、核酸配列の場合はBLASTNプログラムを用いて算出される。両方とも、NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION(WWW.NCBI.NLM.NIH.GOV/)を通じて使用可能であり、例えば、ALTSCHUL ET AL. (1990), J. MOL. BIOL. 215:403-410; GISH AND STATES (1993), NATURE GENET. 3:266-272; MADDEN ET AL. (1996), METH. ENZYMOL. 266:131-141; ALTSCHUL ET AL. (1997), NUCLEIC ACIDS RES. 25:3389-3402; ZHANG ET AL. (2000) J. COMPUT. BIOL. 7(1-2):203-14に記載される。本出願では、二つのアミノ酸配列の類似パーセントは、BLASTPアルゴリズムにおける下記のパラメータに基づくスコアである。単語サイズ=3; ギャップ開放

40

50

ペナルティー = - 1 1 ; ギャップ延長ペナルティー = - 1 ; 及びスコアリングマトリックス = B L O S U M 6 2 である。本出願では、二つの核酸配列の類似パーセントは、B L A S T N アルゴリズムにおける下記のパラメータに基づくスコアである。単語サイズ = 1 1 ; ギャップ開放ペナルティー = - 5 ; ギャップ延長ペナルティー = - 2 ; マッチ評価 = 1 ; 及びミスマッチペナルティー = - 3 である。

【 0 0 6 8 】

二つのタンパク質又はアミノ酸配列の修飾に関して本出願で用いる場合、「対応する」という用語は、(例えば、B L A S T P プログラムを用いて)二つのタンパク質を一般的な方法で配列アライメントしたとき、第1のタンパク質の修飾が、第2のタンパク質の修飾と同じアミノ酸残基の置換であること、及び第1のタンパク質における修飾のアミノ酸位置が、第2のタンパク質の修飾のアミノ酸位置に対応する又は揃うことを意味する。したがって、もし残基 X と Y とが配列アライメントで互に対応するものであれば、X と Y の配列番号が異なるものであっても、第1のタンパク質における残基 “ A ” がアミノ酸 “ A ” に修飾されることは、残基 “ Y ” がアミノ酸 “ A ” に修飾されることと対応する。

10

【 0 0 6 9 】

本出願では、ある変数に対する数値範囲は、その範囲内であれば、任意の数値に等しい変数を用いて本発明が実行可能であることを意図する。したがって、本来的に不連続な変数である場合、この変数は、その数値範囲の両端を含む任意の整数に等しくなることが可能である。同様に、本来的に連続的な変数である場合、この変数は、その数値範囲の両端を含む任意の実数に等しくなることが可能である。例えば、ただしこれらに限定されるものではないが、0 と 2 の間の値を持つと記載される変数は、もしこの変数が本来的に不連続である場合は、0、1 又は 2 を取ることが可能であり、もしこの変数が本来的に連続的である場合は、数値 0 . 0、0 . 1、0 . 0 1、0 . 0 0 1 又は 0 及び 2 の実数の任意の実数を取ることができ。

20

【 0 0 7 0 】

本出願では、特に別段の指定が無い限り、「又は」という単語は、「及び / 又は」という包括的意味で使われ、「又は / どちらか」という排除的意味では使われない。

【 0 0 7 1 】

2 . 1 配列特異性を変更させた合理的設計メガヌクレアーゼ

本発明の様態では、組み換え L A G L I D A D G ファミリーのメガヌクレアーゼを合理的に設計する方法が提供される。この様態では、先ず、半分部位の各位置における塩基指向性を変えることが可能なアミノ酸置換を予測することによって、組み換えメガヌクレアーゼを合理的に設計する。これらの置換は、個別に又は組み合わせて、実験的にその有効性が検証され、所望の切断特異性を持つメガヌクレアーゼを生産する。

30

【 0 0 7 2 】

本発明によれば、塩基指向性において所望の変化をもたらすアミノ酸置換は、メガヌクレアーゼの D N A 認識配列及び D N A フォスフォジエステルバックボーンの核酸塩基との接触に参加が可能な基準メガヌクレアーゼ (例えば、野生型メガヌクレアーゼ又は非天然の基準メガヌクレアーゼ) のアミノ酸側鎖及びそれら接触部の空間的、化学的性質を決めることによって予測される。これらのアミノ酸としては、基準 D N A 半分部位との接触に与る側鎖が挙げられるが、ただしこれらに限定されない。一般に、この決定には、メガヌクレアーゼと、その 2 本鎖 D N A 認識配列との間の複合体の構造に関する知識、又はきわめて類似性の高い複合体の構造 (例えば、同じメガヌクレアーゼと、別の D N A 認識配列との間の複合体構造、又はメガヌクレアーゼの対立遺伝子変種または系統の変種と、その D N A 認識配列との間の複合体構造) に関する知識が必要である。

40

【 0 0 7 3 】

原子座標データとして記載される、ポリペプチド又は 2 個以上のポリペプチドから成る複合体の三次元構造は、いくつかの方法で獲得することが可能である。例えば、タンパク質の構造の決定は、以下のものに限定されるものではないが、X 線結晶学、N M R 及び質量分析を含む技術を用いて実行することが可能である。もう一つの方法は、興味のあるメガヌ

50

クレアーゼ又は関連ヌクレアーゼの、既存の構造座標のデータベースを分析することである。このような構造データは、多くの場合、三次元座標形式のデータベースから入手することが可能である。多くの場合、このデータは、オンラインデータベースを通じて（例えば、WWW.RCSB.ORG/PDBにおいてRCSB PROTEIN DATA BANKへ）アクセスすることが可能である。

【0074】

構造情報は、タンパク質またはタンパク質複合体の、規則的な、二次元又は三次元アレイ（例えば、結晶）によって得られる。例えば、X線又は電子の回折パターンを分析することによって実験的に得ることが可能である。回折データを空間における三次元原子座標に変換するためには、コンピュータ法が使用される。例えば、X線結晶学の分野は、メガヌクレアーゼを含めた、多くのタンパク質-DNA複合体の三次元構造情報の生成のために使用されている（例えば、CHEVALIER ET AL. (2001), NUCLEIC ACIDS RES. 29(18): 3757-3774を参照）。

10

【0075】

溶液中の分子の原子間距離の決定には、核磁気共鳴(NMR)も用いられている。コンピュータ法と組み合わせた多次元NMR法は、増大するサイズのポリペプチドの原子座標の決定に成功した（例えば、TZAKOS ET AL. (2006), ANNU. REV. BIOPHYS. BIOMOL. STRUCT. 35: 12-42を参照）。

【0076】

それとは別に、タンパク質/DNAの既知の一次構造及び利用可能であれば、二次、三次、及び/又は四次構造のほか、アミノ酸側鎖、核酸塩基及び結合性相互作用に関する既知の物理化学的性質に基づいたアルゴリズムを適用することによって、コンピュータモデルを使用することも可能である。このような方法は、任意に、繰り返し法又は実験的に誘導された制限を含むことも可能である。このようなコンピュータソフトウェアの一例として、ADAMS ET AL. (1999), ACTA CRYSTALLOGR. D. BIOL. CRYSTALLOGR. 55(P1): 181-90によって記載されるCNSプログラムがある。他にも、タンパク質の構造におけるアミノ酸の空間配置及びタンパクのアミノ酸側鎖と各種標的分子との相互作用を予測する、様々なコンピュータプログラムが開発されている（例えば、米国特許第6,988,041号を参照）。

20

【0077】

したがって、本発明のある実施態様では、DNA核酸塩基と特異的に相互作用を持ち、及び/又は、非特異的フォスホジエステルバックボーン相互作用を促進する、特異的アミノ酸残基を特定するために、コンピュータモデルが使用される。例えば、可能なメガヌクレアーゼ-DNA相互作用の全体に関するコンピュータモデルを、例えば、以下に限定されるものではないが、MOLSCRIPT(登録商標)2.0(AVATAR SOFTWARE AB, STOCKHOLM, SWEDEN)、グラフィックディスプレイ(JONES ET AL. (1991), ACTA CRYSTALLOGRAPHY, A47: 110)、グラフィックディスプレイ・プログラムGRASP(登録商標)(NICHOLLS ET AL. (1991), PROTEINS, STRUCTURE, FUNCTION AND GENETICS 11(4): 281FF)、グラフィックディスプレイ・プログラムINSIGHT(登録商標)(TSI, INC., SHOREVIEW、ミネソタ州)を含む、適切なソフトウェアプログラムを用いて生産することが可能である。タンパク質-DNA複合体の、三次元構造描画を生産し、視認し、かつ操作するために好適なコンピュータハードウェアは市販されており、従来技術で周知である（例えば、SILICON GRAPHICS WORKSTATION, SILICON GRAPHICS, INC., MOUNTAINVIEW、カリフォルニア州）。

30

40

【0078】

具体的に述べると、メガヌクレアーゼと、その2本鎖DNA認識配列との間の相互作用は、従来技術で既知の方法を用いて分解することが可能である。例えば、そのために結晶が生産される、複数成分の複合体構造の三次元構造の描画、またはモデルは、分子置換、

50

またはSIR/MIR(SINGLE/MULTIPLE ISOMORPHOUS REPLACEMENT)(例えば、BRUNGER(1997), METH. ENZYM. 276:558-580; NAVAZA AND SALUDIJIAN(1997), METH. ENZYM. 276:581-594; TONG AND ROSSMAN(1997), METH. ENZYM. 276:594-611; およびBENTLEY(1997), METH. ENZYM. 276:611-619を参照)を含む技術を用いて決めることが可能であり、例えば、AMORE/MOSFLM(NAVAZA(1994), ACTA CRYST. A50:157-163; CCP4(1994), ACTA CRYST. D50:760-763)、または、XPLORE(BRUENGER ET AL.(1992), X-PLOR VERSION 3.1.A. SYSTEM FOR X-RAY CRYSTALLOGRAPHY AND NMR, YALE UNIVERSITY PRESS, NEW HAVEN、コネチカット州)などのソフトウェアを用いて実行することが可能である。

10

20

30

40

50

【0079】

タンパク質の構造及び可能なメガヌクレアーゼ-DNA相互作用の決定は、酵素活性および特異性に影響を及ぼすために変えることが可能なアミノ酸に関する合理的選択を可能にする。決定は、アミノ酸側鎖の、特定の塩基又はDNAフォスフォジエステルバックボーンとの相互作用に関するいくつかの因子に基づく。適切なアミノ酸置換を決めるために使用される化学的相互作用としては、以下に限定されるものではないが、ファンデルワールス力、立体障害、イオン結合、水素結合及び疎水性相互作用が挙げられる。メガヌクレアーゼと、可能な認識配列半分部位中のある特定の塩基との特定の相互作用を優先するか、又は忌避するアミノ酸置換が選択される。この選択は、その認識配列に対する特異性を、かつ、ある程度は、全体的結合親和度および活性を増すようにあるいは減らすように行われる。さらに、全体活性の増加又は減少のために、かつ、ある程度は、特異性を下げるか又は上げるかするために、2本鎖DNAのフォスフォジエステルバックボーンに対する結合性を増すか又は下げるアミノ酸置換を選択することが可能である。

【0080】

したがって、特定の実施態様では、メガヌクレアーゼ-DNA複合体の三次元構造が決定され、DNA認識配列半分部位における各塩基対について「接触面」が定義される。ある実施態様では、接触面は、その残基が、野生型メガヌクレアーゼ-DNA複合体において塩基接触を実行するか否かとは無関係に、ペアのいずれかの塩基における大溝水素結合ドナーまたはアクセプターから9.0オングストローム未満の炭素を有する、酵素中のアミノ酸を含む。別の実施態様では、残基が、野生型メガヌクレアーゼ-DNA複合体において接触を行わない場合、この残基を排除することが可能であり、又は考慮される残基の数または名称を変えようとする設計者の意図に基づいて残基を含めること、又は排除することが可能である。後述の一実施例では、野生型I-CREI半分部位の塩基位置-2、-7、-8及び-9において、接触面は、野生型酵素-DNA複合体において実際に相互作用を持つアミノ酸位置に限定された。しかしながら、位置-1、-3、-4、-5及び-6では、接触面は、野生型接触には与らないが、異なるアミノ酸によって置換された場合、塩基に接触する可能性のある、さらに新たなアミノ酸位置を含むように定義された。

【0081】

認識配列半分部位は、通常、DNAの1本鎖のみに関して表されるが、メガヌクレアーゼは、2本鎖DNAの大溝に結合し、両鎖の核酸塩基と接触することに注意しなければならない。さらに、「センス」および「アンチセンス」鎖という表示は、メガヌクレアーゼの結合と認識に関しては完全にその場まかせである。ある位置における配列特異性は、塩基対の一方のメンバーとの相互作用か又は塩基対の両方のメンバーとの相互作用の組み合わせのどちらかによって実現される。したがって、例えば、位置Xにおいて、該位置XではA塩基は「センス」鎖上にあり、T塩基が「アンチセンス」鎖上にある、A/T塩基対の存在を優先するためには、位置Xにおいてセンス鎖に接触できるほど十分に近接し、か

つ A の存在を好む残基が選択されるか、及び / 又は、位置 X においてアンチセンス鎖に接触できるほど十分に近接し、かつ T の存在を好む残基が選択される。本発明によれば、残基の - 炭素が、関連塩基の最近接原子の 9 オングストローム以内にある場合、その残基は十分に近接すると考慮される。

【 0 0 8 2 】

したがって、例えば、DNA センス鎖の 9 オングストローム以内ではあるが、アンチセンス鎖から 9 オングストロームより大きく離れる - 炭素を持つアミノ酸は、センス鎖に対してのみ相互作用を持つ可能性があると考えられる。同様に、DNA アンチセンス鎖の 9 オングストローム以内ではあるが、センス鎖から 9 オングストロームより大きく離れる - 炭素を持つアミノ酸は、アンチセンス鎖に対してのみ相互作用を持つ可能性があると考えられる。両方の DNA 鎖の 9 オングストローム以内に - 炭素を持つアミノ酸は、どちらの鎖とも相互作用を持つ可能性があると考えられる。

10

【 0 0 8 3 】

各接触面について、可能なアミノ酸置換について、4 種の DNA 塩基の内の一つ以上に対し好んで相互作用を持とうとする能力を予測し、その予測能力に基づいてアミノ酸置換が選択される。この選択プロセスは、二つの主要基準に基づく。(I) 異なる核酸塩基との立体相互作用に影響を及ぼすアミノ酸側鎖のサイズ、及び (I I) 異なる核酸塩基との静電的及び結合的相互作用に影響を及ぼすアミノ酸側鎖の化学的性質である。

【 0 0 8 4 】

側鎖のサイズに関しては、接触面におけるアミノ酸の - 炭素が、塩基から 6 オングストローム未満の距離にある場合には、比較的短い及び / 又は比較的小さい側鎖を持つアミノ酸を選択することが可能であり、接触面におけるアミノ酸の - 炭素が、塩基から 6 オングストロームより遠い距離にある場合には、比較的長い及び / 又は比較的大きい側鎖を持つアミノ酸を選択することが可能である。サイズが中間のアミノ酸側鎖は、接触面におけるアミノ酸 - 炭素が、塩基から 5 ~ 8 オングストロームである場合に選択することが可能である。

20

【 0 0 8 5 】

比較的短く、比較的小さい側鎖を持つアミノ酸は、グリシン (G)、アラニン (A)、セリン (S)、トレオニン (T)、システイン (C)、バリン (V)、ロイシン (L)、イソロイシン (I)、アスパラギン酸 (D)、アスパラギン (N) 及びプロリン (P) を含む 1 群に割り当てられる。しかしながら、プロリンは、比較的屈曲性が低いので、比較的頻繁な使用が期待されない。さらに、グリシンは、ペプチドバックボーンに不要な屈曲性を導入するため、かつ、きわめてサイズが小さいために、より大きな残基を置換した場合、効果的接触が得られる確率が低いと考えられるために、比較的頻繁な使用が期待されない。一方、グリシンは、ある場合、縮重位置を増進するために使用することが可能である。比較的中間的長さ及びサイズの側鎖を持つアミノ酸は、リシン (K)、メチオニン (M)、アルギニン (R)、グルタミン酸 (E)、およびグルタミン (Q) を含む 2 群に割り当てられる。比較的長い、及び / 又は比較的大きい側鎖を持つアミノ酸は、リシン (K)、メチオニン (M)、アルギニン (R)、ヒスチジン (H)、フェニルアラニン (F)、チロシン (Y)、およびトリプトファン (W) を含む 3 群に割り当てられる。しかしながら、トリプトファンは、比較的屈曲性が低いので、比較的頻繁な使用が期待されない。さらに、リシン、アルギニン、およびメチオニン側鎖の屈曲性によって、これらのアミノ酸は、長い距離または中間距離からの塩基接触が可能となるために、それらが、2 群と 3 群の両群に含められることを保証する。これらの群をさらに、下記に表 1 として示す。

30

40

【 0 0 8 6 】

【表 1】

1群	2群	3群
glycine (G)	glutamine (Q)	arginine (R)
alanine (A)	glutamate (E)	histidine (H)
serine (S)	lysine (K)	phenylalanine (F)
threonine (T)	methionine (M)	tyrosine (Y)
cysteine (C)	arginine (R)	tryptophan (W)
valine (V)		lysine (K)
leucine (L)		methionine (M)
isoleucine (I)		
aspartate (D)		
asparagine (N)		
proline (P)		

10

20

【0087】

側鎖の化学的性質に関しては、種々のアミノ酸が、種々の核酸塩基との相互作用の可能性について評価されており（例えば、ファンデルワールス力、イオン結合、水素結合及び疎水性相互作用）、2本鎖DNA認識配列半分部位中の特定位置における特定塩基と、メガヌクレアーゼとの特異的相互作用を好むか又は忌避する残基が選択される。ある場合には、一つ以上の、完全な又は部分的な縮重位置を持つ半分部位を創出することが望ましい場合がある。そのような場合、二つ以上の塩基の存在を好む残基又は一つ以上の塩基を忌避する残基を選んでもよい。例えば、部分的に縮重された塩基認識は、センス又はアンチセンス位置においてピリミジンを立体的に妨げることによって実現することが可能である。

30

【0088】

グアニン（G）塩基の認識は、該塩基のN7およびO6に対して水素結合を形成する塩基性側鎖を持つアミノ酸を用いて実現される。シトシン（C）の特異性は、C以外の全ての塩基の上に存在する、大溝の電氣的陰性基との相互作用を持ちたがらない、陰性荷電側鎖によって付与される。チミン（T）の認識は、疎水性側鎖と、塩基の上の大溝メチル基との間の疎水性およびファンデルワールス相互作用を用いて合理的に設計される。最後に、アデニン（A）塩基は、該塩基のN7およびN6に対する一对の水素結合を介して、ASNおよびGLNのカルボキサミド側鎖、またはTYRのヒドロキシル側鎖を用いて認識される。最終として、HISは、N7に対して水素結合を供与することによって、プリン塩基（AまたはG）に対して特異性を付与するために使用することが可能である。この、DNA認識に関する単純な規則を用いて、ここでは、ある特定の塩基対位置において、塩基の内の一つが、または両方が合理的設計接触を通じて認識される、そのような接触面を予測することが可能である。

40

【0089】

したがって、種々の核酸塩基との結合性相互作用及び接触を行う位置において好む塩基に基づいて、各アミノ酸残基を、それが好む種々の塩基（すなわち、G、C、T、またはA）に対応する一つ以上の異なる群に割り当てることが可能である。このようにして、G

50

群は、アルギニン（R）、リシン（K）、およびヒスチジン（H）を含み；C群は、アスパラギン酸（D）およびグルタミン酸（E）を含み；T群は、アラニン（A）、バリン（V）、ロイシン（L）、イソロイシン（I）、システイン（C）、トレオニン（T）、メチオニン（M）、およびフェニルアラニン（F）を含み；かつ、A群は、アスパラギン（N）、グルタミン（N）、チロシン（Y）、およびヒスチジン（H）を含む。なお、ヒスチジンは、G群とA群の両群に現れること；セリン（S）は、どの群にも含まれないが、縮重位置を好むものとして使用してもよいこと；および、プロリン、グリシン、およびトリプトファンは、主に立体的配慮のためにどの特定の群にも含まれないことに注意されたい。これらの群も下記に表2として示す。

【0090】

【表2】

G群	C群	T群	A群
arginine (R)	aspartate (D)	alanine (A)	asparagine (N)
lysine (K)	glutamate (E)	valine (V)	glutamine (Q)
histidine (H)		leucine (L)	tyrosine (Y)
		isoleucine (I)	histidine (H)
		cysteine (C)	
		threonine (T)	
		methionine (M)	
		phenylalanine (F)	

10

20

30

【0091】

したがって、本発明によれば、任意の位置Xにおいて、メガヌクレアーゼの認識配列半分部位に所望の変化を実行するために、(1)野生型、または参照メガヌクレアーゼ-DNA複合体の三次元構造の少なくとも関連部分及び位置Xにおける接触面を定めるアミノ酸残基側鎖を決定すること；(2)接触面を含む少なくとも一つの残基の炭素及び位置Xにおける塩基対の少なくとも一つの塩基との間の距離を決定すること；(3)所望の変化を促進するために、(A)塩基から6オングストローム未満の距離にある残基については、G群、C群、T群又はA群の内の適切な一つの群のメンバーである1群及び/又は2群から残基を選ぶこと、及び/又は、(B)塩基から6オングストロームより遠い距離にある残基については、G群、C群、T群又はA群の内の適切な一つの群のメンバーである2群及び/又は3群から残基を選ぶこと、である。接触面を含む、一つを超えるこのような残基を、分析および修飾のために選ぶことが可能であり、ある実施態様では、このような各残基が分析され複数の残基が修飾される。同様に、接触面に含まれる残基の炭素と、位置Xにおける塩基対の二つの塩基のそれぞれとの間の距離を測定することが可能であり、もしもその残基が両方の塩基の9オングストローム以内にある場合、該塩基対の二つの塩基に影響を及ぼすために異なる置換を実行することが可能である（例えば、一方の鎖の近位塩基に影響を及ぼすために1群の残基、又は他方の鎖の遠位塩基に影響を及ぼすために3群の残基）。それとは別に、ペアの両塩基と相互作用を持つことが可能な残基置換の組み合わせは、特異性に影響を及ぼすことが可能である（例えば、アンチセンス鎖

40

50

に接触する A 群の残基と組み合わせたセンス鎖に接触する T 群の残基は T / A を選択する)。最後に、残基の複数の候補修飾は、経験的に (例えば、組み換えメガヌクレアーゼを生産し、その配列認識を調べることによって)、又はコンピュータ的に (例えば、修飾された酵素のメガヌクレアーゼ - DNA 複合体のコンピュータモデル化によって) その有効性を検証して、複数候補の中から選択することが可能である。

【0092】

野生型又は基準ヌクレアーゼの、一つ以上の所望のアミノ酸修飾が選択されたならば、合理設計メガヌクレアーゼは、従来技術で周知の組み換え法および技術によって生産することが可能である。ある実施態様では、特異的配列修飾を創製するために、非ランダムな、部位指向性突然変異誘発技術が用いられる。非ランダム突然変異誘発技術の非限定的例としては、重複プライマー PCR (例えば、WANG ET AL. (2006), NUCLEIC ACIDS RES. 34 (2) : 517 - 527 を参照)、部位指向性突然変異誘発 (例えば、米国特許第 7,041,814 号を参照)、カセット突然変異誘発 (例えば、米国特許第 7,041,814 号参照)、および、PROMEGA BIOSCIENCES, INC. (SAN LUIS OBISPO、カリフォルニア州) から市販される ALTERED SITES (R) IMMUTAGENESIS SYSTEM KIT 用のメーカーのプロトコールが挙げられる。

10

【0093】

合理設計メガヌクレアーゼによる、特異的 DNA 配列の認識及び切断は、当業者の既知の任意の方法によって定量することが可能である (例えば、米国特許出願公開第 2006 / 0078552 号を参照)。ある実施態様では、メガヌクレアーゼ切断の定量は、インビトロ切断アッセイによって定量される。このようなアッセイは、定量されるメガヌクレアーゼの意図される認識配列を含むポリヌクレオチド基質のインビトロ切断を用いるが、ある実施態様では、片方又は両方の半分部位の一つ以上の塩基が、別の塩基に変えられた、意図された認識配列の変種を含むポリヌクレオチド基質のインビトロ切断を用いる。通常、ポリヌクレオチド基質は、合成され、ベクターにクローンされた標的部位を含む 2 本鎖 DNA 分子である。ポリヌクレオチド基質は、直線状であっても、円形であってもよいが、通常、唯一の認識配列を含む。メガヌクレアーゼは、適切な条件下で、ポリヌクレオチド基質とインキュベートされ、得られたポリヌクレオチドは、切断産物を特定するために、既知の方法で分析される (例えば、電気泳動又はクロマトグラフィー)。直線状 2 本鎖 DNA 基質において単一の認識配列がある場合には、メガヌクレアーゼ活性は、二つのバンドの出現及び最初の全長の基質バンドの消失によって検出される。一実施態様では、メガヌクレアーゼ活性は、例えば、WANG ET AL. (1997), NUCLEIC ACID RES., 25 : 3767 - 3776 に記載される通りに定量することが可能である。

20

30

【0094】

他の実施態様では、メガヌクレアーゼの切断パターンは、インビボ切断アッセイによって定量される (例えば、米国特許出願公開第 2006 / 0078552 号)。特定の実施態様では、このインビボ試験は、単一鎖アニーリング組み換え試験 (SSA) である。この種の試験は当業者には既知である (RUDIN ET AL. (1989), GENETICS 122 : 519 - 534 ; FISHMAN - LOBELLE ET AL. (1992), SCIENCE 258 : 480 - 4)。

40

【0095】

当業者には明らかなように、DNA 認識および結合に關与するドメイン以外の、メガヌクレアーゼ酵素のドメインに対し、活性を完全に消失させることなく、さらに新たなアミノ酸置換、挿入又は欠失を行うことが可能である。置換は、構造的または機能的に束縛された位置における、類似のアミノ酸残基の保存的置換であってもよく、又は構造的または機能的に比較的束縛されない位置における非保存的置換であってもよい。このような置換、挿入又は欠失は、当業者によって、不当な努力を要することなく、通例の実験操作によって特定することが可能である。したがって、ある実施態様では、本発明の組み換えメ

50

ガヌクレアーゼは、基準メガヌクレアーゼの配列に対し、85%から99%の間の任意のパーセント(例えば、85%、87.5%、90%、92.5%、95%、97.5%、99%)の配列類似性を持つタンパクを含む。各野生型I-CREI、I-MSOI、I-SCEI、およびI-CEUIタンパクに関して、多くのN-末端およびC-末端配列は、X線結晶実験でははっきりと見ることができない。これは、これらの位置が、構造的又は機能的に束縛されていないことを示唆する。それゆえ、これらの残基は、配列類似性の計算から排除することが可能であり、下記の基準メガヌクレアーゼ配列を使用することが可能である。I-CREIについては配列番号1の残基2-153; I-MSOIについては配列番号6の残基6-160; I-SCEIについては配列番号9の残基3-186; I-CEUIについては配列番号12の残基5-211である。

10

【0096】

2.2 LAGLIDADGファミリー・メガヌクレアーゼ

LAGLIDADGメガヌクレアーゼ・ファミリーは、多様な系統群の宿主生物体由来の200を超えるメンバーから構成される。このファミリーのメンバーは全て、特異的DNA配列の切断に与る他の構造的モチーフと共に、高度に保存されるLAGLIDADGモチーフの1個又は2個の複製を有する。LAGLIDADGモチーフの1つの複製を有する酵素(すなわち、モノ-LAGLIDADGメガヌクレアーゼ)はダイマーとして機能し、一方、このモチーフの2つの複製を有する酵素(すなわち、ジ-LAGLIDADGメガヌクレアーゼ)はモノマーとして機能する。

20

【0097】

LAGLIDADGファミリーメンバーは全て、比較的長い配列(>12BP)を認識、切断し、4個のヌクレオチドの、3オーバーハングを残す。これらの酵素はさらに、LAGLIDADGの外に、タンパク-DNAインターフェイスにおいて、逆平行鎖の類似配置を含むいくつかの構造的モチーフを共有する。これらの保存された構造モチーフ内のアミノ酸がDNA塩基と相互作用を持ち、認識配列特異性を付与する。ファミリーのいくつかのメンバー(例えば、I-CREI、I-MSOI、I-SCEI及びI-CEUI

30

)の間に全体構造の類似性が見られることは、X線結晶学によって明らかにされている。したがって、このファミリーのメンバーは、この構造モチーフ内部の特定のアミノ酸について修飾し、該酵素の全体活性又は配列特異性を変えることが可能であり、かつ、対応する修飾は、他のファミリーメンバーにおいても類似の結果をもたらすことが当然期待される。概論については、CHEVALIERET AL. (2001), NUCLEIC ACID RES. 29(18):3757-3774)を参照されたい。

【0098】

2.2.1 I-CREIから得られるメガヌクレアーゼ

1側面において、本発明は、CHLAMYDOMONAS REINHARDTIIのI-CREIメガヌクレアーゼに基づく、又はこのメガヌクレアーゼから得られた合理的設計メガヌクレアーゼに関する。I-CREIメガヌクレアーゼの野生型アミノ酸配列は、GENBANKアクセス#PO5725に対応する、配列番号1に示される。結晶構造PDB#1BP7における野生型I-CREIメガヌクレアーゼの、二つの、認識配列半分部位を下記の表3に示す。

40

【0099】

【表3】

Position	-9-8-7-6-5-4-3-2-1
	5'-G A A A C T G T C T C A C G A C G T T T T G-3' SEQ ID NO: 2
	3'-C T T T G A C A G A G T G C T G C A A A A C-5' SEQ ID NO: 3
Position	-1-2-3-4-5-6-7-8-9

【0100】

50

この天然の認識配列は、中央の4塩基の外側においても、完全にはパリンδροーム的ではないことに注意されたい。なお、二つの認識配列半分部位がそれぞれのセンス鎖に太字で示される。

【0101】

野生型 I - C R E I はさらに、下記の表4に示すように、(中央の N 1 - N 4 塩基を除き)完全にパリンδροームである配列も認識し、切断する。

【0102】

【表4】

Position	-9-8-7-6-5-4-3-2-1	
	5'- C A A A C T G T C G T G A G A C A G T T T G -3'	SEQ ID NO: 4
	3'- G T T T G A C A G C A C T C T G T C A A A C -5'	SEQ ID NO: 5
Position		-1-2-3-4-5-6-7-8-9

10

【0103】

配列番号4及び配列番号5のパリンδροーム配列は、野生型 I - C R E I にとってより優れた基質であると考えられる。なぜなら、この酵素は、天然の DNA 配列よりもより高い親和度の下にこの部位に結合し、より効率的に該部位を切断するからである。下記の開示のために、かつ、特に本出願において提示する実験結果との関連において、野生型 I - C R E I によって切断されるこのパリンδροーム配列を " W T " と呼ぶことにする(例えば、図2(A)参照)。二つの、認識配列半分部位が、それぞれのセンス鎖に太字で示される。

20

【0104】

図1(A)は、野生型 I - C R E I メガヌクレアーゼ・ホモダイマーの、2本鎖 DNA 認識配列との相互作用を描く。図1(B)は、野生型酵素の半分部位と野生型認識配列の - 4 位置におけるこの酵素のアミノ酸残基と塩基との間の特異的相互作用を示し、図1(C)~図1(E)は、半分部位の - 4 位置における特異性を変更させた、本発明の、三つの合理設計メガヌクレアーゼの、半分部位の - 4 位置における、酵素のアミノ酸残基と塩基との間の特異的相互作用を示す。

30

【0105】

したがって、半分部位の任意の、指定の塩基位置における塩基に対する好みは、本出願に開示される方法を用いて、他の三つの塩基のそれぞれに合理的に変えることが可能である。第一に、指定の塩基位置における野生型認識面が決められる(例えば、メガヌクレアーゼ - DNA 複合体共結晶構造を分析することによって、又はメガヌクレアーゼ - DNA 複合体のコンピュータモデル化によって)。第二に、指定の塩基位置における、周辺アミノ酸位置の - 炭素と、各 DNA 鎖における核酸塩基の間の距離に基づいて、既存の及び可能な接触残基が決められる。例えば、ただしこれらに限定されないが、図1(A)に示すように - 4 位置における、I - C R E I 野生型メガヌクレアーゼ - DNA 接触残基は、位置26のグルタミンを含み、これはアンチセンス DNA 鎖の A 塩基と水素結合する。さらに、残基77が、DNA アンチセンス鎖の - 4 塩基に接触し得る可能性を持つことが特定された。残基26の - 炭素は、アンチセンス DNA 鎖の A 塩基の N 7 から 5 . 9 離れており、残基77の - 炭素は、センス鎖の T の C 5 メチルから 7 . 1 5 オングストローム離れている。この距離及び本出願に記載される塩基化学規則によれば、センス鎖を C とすると、位置77のグルタミン酸と水素結合することが可能であり、アンチセンス鎖を G とすると、位置26のグルタミンと結合することが可能である(野生型 I - C R E I 結晶構造に観察されるように、水分子を介して)(図1(C)を参照)。センス鎖を G とすると、位置77のアルギニンと水素結合することが可能であり、アンチセンス鎖を C とすると、位置26のグルタミン酸と水素結合することが可能である(図1(D)を参照)。センス鎖を A とすると、位置77のグルタミンと水素結合することが可能であり、アンチセンス鎖を T とすると、位置26のアラニンと疎水性接触を形成することが可能である(

40

50

図 1 (E) を参照) 。 もしも塩基の特異的接触が位置 77 によって与えられるとすると、野生型接触 Q 26 を置換し (例えば、セリン残基によって) 、その特異性に対する影響を低下、または除去することが可能である。それとは別に、位置 26 および 77 における相補的突然変異を組み合わせ、特定の塩基対を指定することが可能である (例えば、A 26 は、アンチセンス鎖において T を指定し、Q 77 は、センス鎖において A を指定する (図 1 (E))) 。これらの予言された残基置換は、全て、実験的にその有効性が実証された。

【 0 1 0 6 】

したがって、本発明によれば、I - C R E I メガヌクレアーゼの DNA 認識ドメインに対する、ある確定数のアミノ酸置換が特定された。これらは、単独で又は組み合わせて用いられて、DNA 認識配列半分部位内の個別の塩基において特異性を変更された組み換えメガヌクレアーゼを、したがって、野生型酵素とは異なる半分部位を持つ合理設計メガヌクレアーゼをもたらすことが可能である。I - C R E I のアミノ酸修飾及びそれによってもたらされる、認識配列半分部位特異性における変化を、表 5 に示す。

【 0 1 0 7 】

【表 5】

Posn.	Favored Sense-Strand Base										
	A	C	G	T	A/T	A/C	A/G	C/T	G/T	A/G/T	A/C/G/T
-1	Y75 L75* C75* Y139* C46* A46*	R70* H75* R75* H46* K46* R46*	K70 E70* E75* E46* D46*	Q70* C70 L70 Y75* Q75* H75* H139 Q46* H46*				T46*			G70 A70 S70 G46*
-2	Q70 T44* A44* V44* I44* L44* N44*	E70 D70 K44* R44*	H70 D44* E44*	Q44*	C44*						
-3	Q68 C24* I24*	E68 F68 K24* R24*	R68	M68 C68 L68 F68		H68		Y68	K68		
-4	A26* Q77	E77 K26*	R77 E26*					S77 Q26*			S26*
-5		E42	R42			K28*	C28* Q42				M66 K66
-6	Q40 C28*	E40 R28*	R40	C40 I40 V40 C79 I79 V79 Q28*	A40 A79 A28* H28*						S40 S28*
-7	N30* Q38	E38 K30* R30*	K38 R38 E30*	I38 L38			C38				H38 N38 Q30*
-8	F33 Y33	E33 D33	F33 H33	L33 V33 I33 F33 C33		R32*	R33				
-9		E32	R32 K32	L32 V32 A32 C32				D32 I32			S32 N32 H32 Q32 T32

10

20

30

40

50

【0108】

太字入力は、野生型接触残基であり、本出願で使用される「修飾」を構成しない。
星印は、その残基が、アンチセンス鎖の塩基に接触することを示す。

【0109】

2.2.2. I - M S O L I から得られるメガヌクレアーゼ

他の側面では、本発明は、MONOMASTIX SP. の I - M S O I メガヌクレアーゼに基づく、又はこのヌクレアーゼから得られた合理設計メガヌクレアーゼに関する。
I - M S O I メガヌクレアーゼの野生型アミノ酸配列は、GENBANKアクセス# A A L 3 4 3 8 7 に対応し、配列番号 6 に示される。結晶構造 P D B # 1 M 5 X における野生型 I - M S O I メガヌクレアーゼの、二つの認識配列半分部位を下記の表 6 に示す。

10

【0110】

【表 6】

Position	-9-8-7-6-5-4-3-2-1	
	5'- C A G A A C G T C G T G A G A C A G T T C C -3'	SEQ ID NO: 7
	3'- G T C T T G C A G C A C T C T G T C A A G G -5'	SEQ ID NO: 8
Position		-1-2-3-4-5-6-7-8-9

【0111】

認識配列は、中央の 4 塩基の外側においても、完全にはパリンδροーム的ではないことに注意されたい。二つの認識配列半分部位が、それぞれのセンス鎖に太字で示される。

20

【0112】

本発明によれば、I - M S O I メガヌクレアーゼの DNA 認識ドメインに対する、ある確定数のアミノ酸置換が特定された。これらは、単独で又は組み合わせて用いられて、DNA 認識配列半分部位内の個別の塩基において特異性を変更された組み換えメガヌクレアーゼを、したがって、野生型酵素とは異なる半分部位を持つ合理設計メガヌクレアーゼをもたらすことが可能である。I - M S O I のアミノ酸修飾、および、予測される、認識配列半分部位特異性における変化を、表 7 に示す。

【0113】

【表 7】

Position	Favored Sense-Strand Base			
	A	C	G	T
-1	K75* Q77 A49* C49* K79*	D77 E77 K49* R75* K75* R79* K79*	K77 R77 E49* E79*	C77 L77 Q79*
-2	Q75 K81 C47* I47* L47*	E75 D75 R47* K47* K81* R81*	K75 E47* E81*	A75 C75 V75 I75 T75 Q47* Q81*
-3	Q72 C26* L26* V26* A26* I26*	E72 Y72 H26* K26* R26*	R72 K72 Y26* F26*	K72 Y72 H26*
-4	K28 Q83	K28* R28* E83	R83 K83	K28 K83 Q28*
-5	K28 C28* L28* I28*	K28* R28*	R45 E28*	Q28*
-6	I30* V30* S30* L30* Q43	E43 E85 K30* R30*	R43 K43 K85 R85 E30* D30*	K43 I85 V85 L85 Q30*
-7	Q41	E32 E41	R32 R41 K41	K32 M41 L41 I41
-8	Y35 K35	E32	R32 K32 K35 R35	K32 K35
-9	N34 H34	D34 E34 S34	K34 R34 H34	S34 C34 V34 T34 A34

10

20

30

40

【0114】

太字は、野生型接触残基を表し、本出願で使用される「修飾」を構成しない。
星印は、その残基がアンチセンス鎖の塩基に接触することを示す。

【0115】

2.2.3. I - S C E I から得られるメガヌクレアーゼ

他の側面では、本発明は、S A C C H A R O M Y C E S C E R E V I S I A E の I - S C E I メガヌクレアーゼに基づく、又はこのヌクレアーゼから得られた合理設計メガヌクレアーゼに関する。I - S C E I メガヌクレアーゼの野生型アミノ酸配列は、G E N B A N K アクセス # C A A 0 9 8 4 3 に対応し、配列番号 9 に示される。結晶構造 P D B # 1 R 7 M における野生型 I - S C E I メガヌクレアーゼの、認識配列を下記の表 8 に示す。

50

【 0 1 1 6 】

【 表 8 】

Sense	5'-T T A C C C T G T T A T C C C T A G-3'	SEQ ID NO:
	10	
Antisense	3'-A A T G G G A C A A T A G G G A T C-5'	SEQ ID NO:
	11	
Position	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18	

【 0 1 1 7 】

認識配列は、非パリンδροーム的であり、かつ、半分部位を分ける四つの塩基対がない
ことに注意されたい。 10

【 0 1 1 8 】

本発明によれば、I - S C E IメガヌクレアーゼのDNA認識ドメインに対する、ある
確定数のアミノ酸置換が特定された。これらは、単独で又は組み合わせて用いられて、D
N A 認識配列内の個別の塩基において特異性を変更された組み換えメガヌクレアーゼ、つ
まり、野生型酵素とは異なる認識配列を持つ合理設計メガヌクレアーゼをもたらすことが
可能である。I - S C E Iのアミノ酸修飾及び予測される認識配列特異性における変化を
、表9に示す。

【 0 1 1 9 】

【表 9】

Position	Favored Sense-Strand Base			
	A	C	G	T
4	K50	R50* K50* E57	E50* R57 K57	K57 M57 Q50*
5	K48 Q102	R48* K48* E102 E59	E48* K102 R102	Q48* C102 L102 V102
6	K59	R59* K59*	K84 E59*	Q59* Y46
7	C46* L46* V46*	R46* K46* E86	K86 R86 E46*	K68 C86 L86 Q46*
8	K61* S61* V61* A61* L61*	E88 R61* H61*	E61* R88 K88	K88 Q61* H61*
9	T98* C98* V98* L98*	R98* K98*	E98* D98*	Q98*
10	V96* C96* A96*	K96* R96*	D96* E96*	Q96*
11	C90* L90*	K90* R90*	E90*	Q90*
12	Q193	E165 E193 D193	K165 R165	C165 L165 C193 V193 A193 T193 S193
13	C193* L193*	K193* R193* D192	E193* D193* K163 R192	Q193* C163 L163
14	L192* C192*	E161 R192* K192*	K147 K161 R161 R197 D192* E192*	K161 Q192*
15		E151	K151	C151 L151 K151
17	N152* S152* C150* L150* V150* T150*	K152* K150*	N152* S152* D152* D150* E150*	Q152* Q150*
18	K155* C155*	R155* K155*	E155*	H155* Y155*

10

20

30

40

【 0 1 2 0 】

太字入力は、野生型接触残基を表し、本出願で使用される「修飾」を構成しない。星印は、その残基が、アンチセンス鎖の塩基に接触することを示す。

【 0 1 2 1 】

50

2.2.4 I-CEUI から得られるメガヌクレアーゼ

他の側面では、本発明は、CHLAMYDOMONASEUGAMETOSのI-CEUIメガヌクレアーゼに基づく、又はこのヌクレアーゼから得られた合理設計メガヌクレアーゼに関する。I-CEUIメガヌクレアーゼの野生型アミノ酸配列は、GENBANKアクセス# P32761に対応し、配列番号12に示される。結晶構造PDB# 2EX5における野生型I-CEUIメガヌクレアーゼの、二つの認識配列半分部位を下記に示す。

【0122】

【表10】

Position	-9-8-7-6-5-4-3-2-1	
	5'-A T A A C G G T C C T A A G G T A G C G A A-3'	SEQ ID NO:
13		
	3'-T A T T G C C A G G A T T C C A T C G C T T-5'	SEQ ID NO:
14		
Position	-1-2-3-4-5-6-7-8-9	

10

【0123】

I-CEUIはホモダイマーであるにも拘わらず、I-CEUIの認識インターフェイスにおける天然の縮重のために(SPIEGEL ET AL.(2006), STRUCTURE 14:869-80)、認識配列は、中央の4塩基の外側においても、非パンドローム的であることに注意されたい。二つの認識配列半分部位がそれぞれのセンス鎖に太字で示される。

20

【0124】

本発明によれば、I-CEUIメガヌクレアーゼのDNA認識ドメインに対する、ある確定数のアミノ酸置換が特定された。これらは、単独で又は組み合わせて用いられて、DNA認識配列半分部位内の個別の塩基において特異性を変更された組み換えメガヌクレアーゼを、したがって、野生型酵素とは異なる半分部位を持つ合理設計メガヌクレアーゼをもたらすことが可能である。I-CEUIのアミノ酸修飾及び予測される認識配列特異性における変化を、表11に示す。

【0125】

30

【表 1 1】

Position	Favored Sense-Strand Base			
	A	C	G	T
-1	C92* A92* V92*	K116* R116* D116* K92*	E116* E92*	Q116* Q92*
-2	Q117 C90* L90* V90*	E117 D117 R174* K124* K90* R90* K68*	K117 R124 K124 E124* E90* D90*	C117 V117 T117 Q90*
-3	C70* V70* T70* L70* K70*	K70*	E70* E88*	Q70*
-4	Q126 N126 K88* L88* C88* C72* L72* V72*	E126 D126 R88* K88* K72*	R126 K126 E88* D88*	K126 L126 Q88*
-5	C74* L74* V74* T74*	K74*	E74* K128 R128 E128	C128 L128 V128 T128
-6	Q86	D86 E86 R84* K84*	K128 R128 R86 K86 E84*	K86 C86 L86
-7	L76* C76* K76*	R76* K76* H76*	E76* R84	H76* Q76*
-8	Y79 R79 Q76	D79 E79 D76 E76	R79 K79 K76 R76	C79 L79 V79 L76
-9	Q78 N78 H78 K78	D78 E78	R78 K78 H78	K78 V78 L78 C78 T78

10

20

30

40

【 0 1 2 6 】

太字入力は、野生型接触残基を表し、本出願で使用される「修飾」を構成しない。星印は、その残基がアンチセンス鎖の塩基に接触することを示す。

【 0 1 2 7 】

2.2.5 特異的に排除される組み換えメガヌクレアーゼ

本発明は、従来技術において記載され、別法によって開発された、いくつかの組み換えメガヌクレアーゼを包含することを意図しない。これらの排除されたメガヌクレアーゼとしては、ARNOULD ET AL. (2006), J. MOL. BIOL. 355:

50

443 - 58 ; S U S S M A N E T A L . (2 0 0 4) , J . M O L . B I O L . 342 : 31 - 41 ; C H A M E S E T A L . (2 0 0 5) , N U C L E I C A C I D S R E S . 33 : E 178 ; S E L I G M A N E T A L . (2 0 0 2) , N U C L E I C A C I D S R E S . 30 : 3870 - 9 ; および A S H W O R T H E T A L . (2 0 0 6) N A T U R E 441 (7093) : 656 - 659 によって記載されるものが挙げられる。なお、上記の全開示を、C33、R33、A44、H33、K32、F33、R32、A28、A70、E33、V33、A26及びR66から選ばれる単一置換を持つ、I - C R E I に基づく組み換えメガヌクレアーゼを含め、引用により本出願に含める。さらに排除されるものは、A68 / N70 / N75及びD44 / D70 / N75から選ばれた3置換、K44 / T68 / G60 / N75及びR44 / A68 / T70 / N75から選ばれる4置換を持つI - C R E I に基づく組み換えメガヌクレアーゼである。最後に、特に指定して排除されるものは、置換L28及びR83のペアを持つI - M S O I に基づく組み換えメガヌクレアーゼである。これらの置換又は置換の組み合わせは、本出願では「排除修飾」と呼ばれる。

10

【0128】

2.2.6 認識配列半分部位において複数変化を有するメガヌクレアーゼ

他の側面では、本発明は、DNA認識配列半分部位の二つ以上の位置において半分部位指向を変えるために、上の節2.2.1 - 2.2.4に記載した、二つ以上のアミノ酸置換を組み合わせることによって得られる合理設計メガヌクレアーゼに関する。例えば、ただしこれらに限定されるものではないが、さらに詳しく後述されるように、酵素DJ1は、修飾R30 / E38（これは、位置 - 7においてCを好む）、R40（これは位置 - 6においてGを好む）、R42（これは、位置 - 5においてGを好む）及びN32（これは、位置 - 9において完全縮重を好む）を取り込むことによって、I - C R E I から得られる。この合理設計DJ1メガヌクレアーゼは、A₇A₆C₅に対する野生型の指向性に比べ、ほぼ変わらない程度でC₇G₆G₅を認識し、位置 - 9におけるAに対する耐性を増大させる。

20

【0129】

異なる塩基位置に影響を及ぼす残基置換を組み合わせることができる能力は、一部は、L A G L I D A D Gメガヌクレアーゼのモジュラー性による。L A G L I D A D G認識インターフェイスにおける塩基接触の大部分は、個々のアミノ酸側鎖によって形成され、インターフェイスは、隣接塩基と相互作用を持つ側鎖間の相互接続性、または水素結合ネットワークから比較的自由である。これによって、一般に、一塩基位置と相互作用を持つ残基を隣接塩基における側鎖同士の間相互作用に影響を及ぼさずに操作することが可能になる。さらに、上の2.2.1 - 2.2.4節に掲げた突然変異の加重性は、これらの突然変異を特定するために用いた本法の直接的結果である。この方法は、単一塩基と直接に相互作用を持つ側鎖置換を予測する。一般に、側鎖間の相互接続、または水素結合ネットワークは、認識インターフェイス内の置換の独立性を維持するために回避される。

30

【0130】

側鎖置換のある組み合わせは、完全に又は部分的に相互に不適合である。不適合ペア又は組のアミノ酸を合理設計メガヌクレアーゼに組み込んだ場合、得られた酵素の触媒活性は、低下するか又は排除される。通常、これらの不適合は、導入されたアミノ酸側鎖間の立体干渉によるもので、活性は、この干渉を特定し、除去することで回復が可能である。具体的には、大きな側鎖を持つ（例えば、2又は3群のアミノ酸）二つのアミノ酸が、メガヌクレアーゼ構造において、互いに隣接するアミノ酸位置に組み込まれた場合（例えば、I - C R E I由来のメガヌクレアーゼの場合、位置32及び33、28及び40、28及び42、42及び77又は68及び77）、これら二つのアミノ酸は互いに干渉しあって、酵素活性を下げる可能性が高い。この干渉は、一方の又は両方の不適合アミノ酸を、より小さな側鎖を持つアミノ酸（例えば、1群または2群）で置換することによって排除することが可能である。例えば、I - C R E I由来の合理設計メガヌクレアーゼでは、K28は、R40とR42の両方に干渉する。酵素活性を最大化するためには、R40及び

40

50

R 4 2 を、位置 2 8 においてセリン又はアスパラギン酸と組み合わせることが可能である。

【 0 1 3 1 】

本出願に記載のように特定される、アミノ酸置換の組み合わせを用いて、野生型メガヌクレアーゼ（又は以前に修飾されたメガヌクレアーゼ）の特異性を、元の認識配列から、興味のある核酸（例えば、ゲノム）中に存在する、所望の認識配列に合理的に変えることが可能である。例えば、図 2（A）は、I - C R E I メガヌクレアーゼ認識配列 W T（配列番号 4）の「センス」鎖の外に、それに対し合理的設計メガヌクレアーゼを創製したならば有用であろうと考えられる、他の、いくつかの配列を示す。W T 認識配列と所望の認識配列との間で保存される塩基には陰影を施す。本発明によれば、I - C R E I メガヌクレアーゼに基づく組み換えメガヌクレアーゼは、これらの所望の認識配列のそれぞれについてばかりでなく、他の任意の認識配列についても、本出願に記載される適切なアミノ酸置換によって、合理的に設計することが可能である。

10

【 0 1 3 2 】

3 . D N A 結合親和度に変更された合理設計メガヌクレアーゼ

前述のように、本発明の組み換えメガヌクレアーゼの D N A 結合親和度は、D N A のフォスフォジエステルバックボーンと接触面を形成するいくつかのアミノ酸を変えることによって修飾することが可能である。この接触面は、D N A バックボーンから 9 オングストローム未満の距離の炭素を持つ酵素のアミノ酸を、この残基が野生型メガヌクレアーゼ - D N A 複合体において D N A バックボーンと接触するかどうかとは無関係に含む。D N A 結合は、酵素活性に必要な前提条件であるから、D N A 結合親和度の上昇 / 低下は、それぞれ、酵素活性の上昇 / 低下を起こすことが示されている。しかしながら、D N A 結合親和度の上昇 / 低下はさらに、メガヌクレアーゼの配列特異性の低下 / 上昇を起こすことが示されている。それゆえ、活性および特異性は両方とも、フォスフォジエステルバックボーン接触を修飾することによって修飾することが可能である。

20

【 0 1 3 3 】

酵素活性の上昇 / 酵素の特異性の低下のためには、

【 0 1 3 4 】

（ I ）酵素と D N A バックボーンの間で静電的反発を取り除く。ある特定されたアミノ酸が、陰性荷電の D N A バックボーンに反発すると予想される、陰性に荷電した側鎖を持つ場合（例えば、アスパラギン酸、グルタミン酸）、この反発は、アミノ酸を無荷電の又は陽性荷電の側鎖で置換することによって除去することが可能である。ただし、この場合、立体障害の作用には影響される。実験的に確かめられた例は、I - C R E I におけるグルタミン酸 8 0 をグルタミンに突然変異させた場合である。

30

【 0 1 3 5 】

（ I I ）酵素と D N A バックボーンの間で静電的牽引相互作用を導入する。接触面の位置のいずれかにおける、陽性荷電側鎖を持つアミノ酸（例えば、リシン又はアルギニン）の導入は、立体干渉の作用には影響されるが、結合親和度を増すことが予想される。

【 0 1 3 6 】

（ I I I ）酵素と D N A バックボーンの間で水素結合を導入する。接触面のアミノ酸が、適切な水素結合官能基を欠如するため、又は D N A バックボーンと相互作用を持つには、側鎖が短すぎるか長すぎるか、及び / 又は、屈曲性が不足するという理由で水素結合を造ることができない場合、適切な長さで屈曲性を持つ、水素結合を供与することが可能な極性アミノ酸（例えば、セリン、トレオニン、チロシン、ヒスチジン、グルタミン、アスパラギン、リシン、システイン又はアルギニン）を導入することが可能である。ただし、立体障害の作用には影響される。

40

【 0 1 3 7 】

具体的に、酵素活性の低下 / 酵素の特異性の上昇のためには、

【 0 1 3 8 】

（ I ）酵素と D N A バックボーンの間で静電的反発を導入する。接触面の位置のいず

50

れかにおいて、陰性に荷電した側鎖を持つアミノ酸（例えば、グルタミン酸、アスパラギン酸）を導入することは、結合親和度を下げることが予想される。ただし、立体障害の作用には影響される。

【0139】

（II）酵素とDNAの間の静電的牽引を取り除く。接触面のいずれかのアミノ酸が、陰性に荷電したDNAバックボーンと相互作用を持つ、陽性に荷電した側鎖を持つ場合（例えば、リシン又はアルギニン）、この好ましい相互作用は、アミノ酸を無荷電の又は陰性荷電の側鎖で置換することによって除去することが可能である。ただし、立体干渉の作用には影響される。実験的に確かめられた例は、I-CREIにおけるリシン116をアスパラギン酸に突然変異させた場合である。

10

【0140】

（III）酵素とDNAバックボーンの水素結合を取り除く。接触面のいずれかのアミノ酸が、DNAバックボーンと水素結合を造る場合、そのアミノ酸をその側鎖が適切な官能基を持たないか又は必要な長さ/屈曲性を欠如するために、同様の水素結合を造るとは予想されないアミノ酸に置換することが可能である。

【0141】

例えば、I-CREIに基づくある組み換えメガヌクレアーゼにおいて、活性を上げるためにI-CREIメガヌクレアーゼの位置80におけるグルタミン酸が、リシン又はグルタミンに変更される。別の実施態様では、I-CREIの位置66のチロシンが、アルギニン又はリシンに変更される。これは、該メガヌクレアーゼの活性を上げる。さらに別の実施態様では、I-CREIの位置34におけるリシンをアスパラギン酸に、位置66におけるチロシンをアスパラギン酸に、及び/又は、位置116におけるリシンをアスパラギン酸に変えることによって、酵素活性を低下させる。

20

【0142】

組み換えメガヌクレアーゼの活性は、その組み換え酵素が、ある特定の認識配列に関して、無活性から極度に高い活性までの間の任意の活性レベルを持つように修飾することが可能である。例えば、DJ組み換えメガヌクレアーゼは、位置26においてグルタミン酸突然変異を抱えると活性を完全に失う。しかしながら、位置26におけるグルタミン酸置換と位置80におけるグルタミン置換の組み合わせは、認識配列半分部位内の-4におけるグアニンに対し高度の特異性および活性を持つ組み換えメガヌクレアーゼを創製する（図1(D)を参照）。

30

【0143】

本発明によれば、フォスフォジエステルDNAバックボーン近傍の各種位置におけるアミノ酸は、メガヌクレアーゼ活性および特異性の両方に同時に作用するように変えることが可能である。この、酵素の特異性と活性の「調整」は、フォスフォジエステルバックボーンに対し、アミノ酸によって造られる接触点の数を増すか又は減らすことによって達成される。フォスフォジエステルバックボーンに対する種々の接触は、アミノ酸側鎖によって促進することが可能である。ある実施態様では、イオン結合、塩橋、水素結合及び立体障害は、アミノ酸側鎖の、フォスフォジエステルバックボーンに対するコンタクトに影響を及ぼす。例えば、I-CREIメガヌクレアーゼの場合、位置116におけるリシンのアスパラギン酸への変更によって、位置-8および-9における核酸塩基間の塩橋が取り除かれ、これは酵素の切断速度は下げるが、特異性は上げる。

40

【0144】

野生型I-CREI（配列番号1）、I-MSOI（配列番号6）、I-SCEI（配列番号9）、およびI-CEUI（配列番号12）各メガヌクレアーゼの、バックボーン接触面を形成する残基は、下表12において特定される。

【0145】

【表 1 2】

I-CreI	I-MsoI	I-SceI	I-CeuI
P29, K34, T46, K48, R51, V64, Y66, E80, I81, K82, L112, K116, D137, K139, T140, T143	K36, Q41, R51, N70, I85, G86, S87, T88, H89, Y118, Q122, K123, Q139, K143, R144, E147, S150, N152	N15, N17, L19, K20, K23, K63, L80, S81, H84, L92, N94, N120, K122, K148, Y151, K153, T156, N157, S159, N163, Q165, S166, Y188, K190, I191, K193, N194, K195, Y199, D201, S202, Y222, K223	K21, D25, K28, K31, S68, N70, H94, R112, R114, S117, N120, D128, N129, R130, H172

10

20

【0146】

酵素の親和度を上げ、したがって、より活性を高め/特異性を下げるためには、

(1) 対応する酵素において、陰性荷電(D又はE)、疎水性(A、C、F、G、I、L、M、P、V、W、Y)、又は無荷電/極性(H、N、Q、S、T)のアミノ酸を、表12から選ぶ。

(2) アミノ酸が陰性荷電しているか又は疎水性である場合は、それを、無荷電/極性(作用が小さい)又は陽性荷電(KまたはR、作用が大きい)となるように突然変異させる。

(3) アミノ酸が無荷電/極性である場合、それを陽性荷電に突然変異させる。

【0147】

酵素の親和度を下げ、したがって、より活性を下げ/特異性を上げるためには、

(1) 対応する酵素において、陽性荷電(K又はR)、疎水性(A、C、F、G、I、L、M、P、V、W、Y)、又は無荷電/極性(H、N、Q、S、T)のアミノ酸を、表12から選ぶ。

(2) アミノ酸が陽性荷電している場合は、それを、無荷電/極性(作用が小さい)又は陰性荷電(作用が大きい)となるように突然変異させる。

(3) アミノ酸が疎水性又は無荷電である場合、それを陰性荷電に突然変異させる。

【0148】

4. ヘテロダイマー・メガヌクレアーゼ

他の側面では、本発明は、二つのモノマーで、その一方は野生型で、その一方または両方が、非天然形または組み換え形であるモノマーのコンタクトによって形成されるヘテロダイマーであるメガヌクレアーゼを提供する。例えば、野生型I-CREIメガヌクレアーゼは、通常、それぞれが、擬似パリンドローム認識配列において一つの半分部位に結合する、二つのモノマーから構成されるホモダイマーである。ヘテロダイマー組み換えメガヌクレアーゼは、異なる半分部位を認識する二つのメガヌクレアーゼを組み合わせることによって、例えば、細胞において二つのメガヌクレアーゼを共時発現することによって、又は溶液中で二つのメガヌクレアーゼを混合することによって生産することが可能である。二つのモノマーにおいて、それらのコンタクトによるダイマー形成に影響を及ぼす、それぞれのアミノ酸を変えることによって、ヘテロダイマーの形成を、ホモダイマーの形成よりも優先することが可能である。特定の実施態様では、二つのモノマーのインターフ

30

40

50

ェイスにおけるいくつかのアミノ酸が、第1モノマーでは、陰性荷電のアミノ酸（D又はE）から陽性荷電のアミノ酸（K又はR）に変えられ、第2モノマーでは、陽性荷電のアミノ酸から陰性荷電のアミノ酸に変えられる（表13参照）。例えば、I-CREI由来のメガヌクレアーゼの場合、位置7及び57におけるリシンが、第1モノマーではグルタミン酸に突然変異され、第2モノマーでは、位置8および61におけるグルタミン酸がリシンに突然変異される。このプロセスの結果は、第1モノマーが、ダイマーインターフェイスにおいて過剰な陽性荷電残基を持ち、第2モノマーが、ダイマーインターフェイスにおいて過剰な陰性荷電残基を持つ、一对のモノマーである。それゆえ、この第1及び第2モノマーは、インターフェイスにおける変更アミノ酸の間の静電的相互作用のために、その同じモノマーペアよりも優先的にコンタクトする。

10

【0149】

【表13】

I-CreI: 第1のモノマーの置換	I-CreI: 第2のモノマーの置換
K7 to E7 or D7 K57 to E57 or D57 K96 to E96 or D96	E8 to K8 or R8 E61 to K61 or R61
I-MsoI: 第1のモノマーの置換	I-MsoI: 第2のモノマーの置換
R302 to E302 or D302	D20 to K60 or R60 E11 to K11 or R11 Q64 to K64 or R64
I-CeuI: 第1のモノマーの置換	I-CeuI: 第1のモノマーの置換
R93 to E93 or D93	E152 to K152 or R152

20

30

40

【0150】

それとは別に、またはそれに加えて、二つのモノマーのインターフェイスにおけるいくつかのアミノ酸は、立体的にホモダイマー形成を妨げるように変えることも可能である。具体的には、一方のモノマーのダイマーインターフェイスのアミノ酸を、ホモダイマーを立体的に阻止する比較的大きな又は嵩高な残基によって置換する。第2のモノマーのダイ

50

マーインターフェイスにおけるアミノ酸を、第1のモノマーの嵩高な残基を補正するために比較的小さな残基で置換し、ヘテロダイマーにおける衝突を取り除くことが可能であるし、または未修飾のままでもよい。

【0151】

さらに他の実施態様では、イオン架橋又は水素結合を、ヘテロダイマー・インターフェイスの疎水性コアの中に埋没させることが可能である。具体的には、インターフェイスのコアのモノマーにおける疎水性残基は、陽性荷電残基によって置換することが可能である。さらに、野生型ホモダイマーでは、第1モノマーにおいて置換された疎水性残基と相互作用を持つ、第2モノマーの疎水性残基は、陰性荷電残基によって置換することが可能である。したがって、この二つの置換残基は、イオン結語または水素結合を形成することが可能である。同時に、疎水性インターフェイスの中に埋没された満たされない電荷の静電的反発は、ホモダイマー形成を忌避する。

10

【0152】

最後に、前述のように、ヘテロダイマーの各モノマーは、DNA認識領域において、異なるアミノ酸を置換されて、そのため、それぞれが異なるDNA半分部位を持ち、かつ、組み合わせられたダイマー形DNA認識配列が非パリンドローム的となってもよい。

【0153】

5. 組み換え細胞および生物体を生産する方法

本発明の側面はさらに、合理設計メガヌクレアーゼを用いた、組み換え、トランスジェニック又は他のやり方で遺伝学的に修飾された細胞および生物体を生産する方法を提供する。したがって、ある実施態様では、相同組み換えによる興味配列の正確な挿入（単数又は複数）を可能とするために、細胞または生物体のゲノムDNAの単一部位又は比較的少数部位に、二本鎖切断点を特異的に引き起こすための、組み換えメガヌクレアーゼが開発された。別の実施態様では、（A）非相同的末端接合による興味配列の稀な挿入（単数又は複数）を可能とするために、又は（B）非相同的末端接合による標的配列の破壊を可能とするために、細胞または生物体のゲノムDNAの単一部位又は比較的少数部位に、二本鎖切断点を特異的に引き起こすための、組み換えメガヌクレアーゼが開発された。興味配列の、相同的組み換え又は非相同的末端接合に関して本出願で用いる場合、「挿入」という用語は、興味配列が染色体の中に組み込まれるように、興味配列が染色体にコンタクトすることを意味する。相同組み換えの場合、挿入された配列は内因性配列を置換し、そのため、元のDNAは等しい長さではあるがヌクレオチド配列が変更された外因性DNAによって置換される。それとは別に、挿入配列はそれが置換する配列よりもより多数の又はより少数の塩基を含むことが可能である。

20

30

【0154】

それゆえ、本発明のこの局面によれば、組み換え生物体としては、単子葉植物種、例えば、米、小麦、トウモロコシ（メイズ）並びにライ麦、及び双子葉植物種、例えば、豆類（例えば、ソラマメ、大豆、レンズ豆、ピーナッツ、エンドウ豆）、アルファルファ、クローバー、タバコ並びにARABIDOPSIS種が挙げられるが、これらに限定されるものではない。さらに、組み換え生物体としては、動物、例えば、ヒト並びに非ヒト霊長類、ウマ、ウシ、ヤギ、ブタ、ヒツジ、イヌ、ネコ、モルモット、ラット、マウス、トカゲ、魚類、及び昆虫類、例えば、ショウジョウバエ種が挙げられるが、これらに限定されるものではない。他の実施態様では、生物体は、真菌、たとえば、カンジダ、ニューロスポラ又はサッカロマイセスである。

40

【0155】

ある実施態様では、本発明の方法は、成熟組み換え生物体になることが可能な又は遺伝学的に修飾された生物体が、そのゲノムの中に興味配列を抱える子孫を出産することを可能とする。例えば、細胞、胚細胞又は幹細胞の中に興味配列を導入することを含む。

【0156】

メガヌクレアーゼタンパクは、ゲノムDNAを切断するために、細胞に輸送することが

50

可能であるが、これによって、従来技術で既知の様々な機序によって、その切断部位において、興味の特列による相同組み換え、または非相同末端接合が可能とされる。組み換えメガヌクレアーゼタンパクは、例えば、マイクロインジェクション又はリボソームトランスフェクション（例えば、L I P O F E C T A M I N E（登録商標）、I N V I T R O G E N C O R P . , C A R L S B A D、カリフォルニア州）を含む、ただしこれらに限定されないが、技術によって細胞中に導入することが可能である。このリボソーム処方は、標的細胞との脂質二重層融合を促進するのに用いることが可能であり、そのため、リボソームの内容物又はその表面にコンタクトするタンパクの細胞内部への取り込みを可能にする。それとは別に、酵素は、適切な取り込みタンパク、例えば、H I V T A T タンパクなどに融合させて細胞の取り込みを指示するようにすることも可能である（例えば、H U D E C Z E T A L . (2 0 0 5) , M E D . R E S . R E V . 2 5 : 6 7 9 - 7 3 6 を参照）。

10

【0157】

それとは別に、メガヌクレアーゼタンパクをコードする遺伝子配列は、従来技術で既知の技術（例えば、A U S U B E L E T A L . , 「分子生物学における最近のプロトコール（“ C U R R E N T P R O T O C O L S I N M O L E C U L A R B I O L O G Y , ”）」、W I L E Y 1 9 9 9 を参照されたい）を用いて、ベクターに挿入され、真核細胞にトランスフェクトされる。興味の特列は、同じベクターによって、異なるベクターによって又は従来技術で既知の他の手段によって導入することが可能である。

20

【0158】

D N A トランスフェクションの非限定的例としては、ウィルスベクター、プラスミド、コスミド及びY A C ベクターが挙げられる。D N A 配列のトランスフェクションは、当業者には既知の、様々な方法によって実現することが可能である。例えば、D N A 配列を細胞に輸送するために、リボソームおよびイムノリボソームが用いられる（例えば、L A S I C E T A L . (1 9 9 5) , S C I E N C E 2 6 7 : 1 2 7 5 - 7 6 を参照）。さらに、ベクターを細胞内に導入するためにウィルスを利用することが可能である（例えば、米国特許第7,037,492号を参照）。それとは別に、ベクターが裸のD N A として導入されるようにトランスフェクション戦略を利用することも可能である（例えば、R U I E T A L . (2 0 0 2) , L I F E S C I . 7 1 (1 5) : 1 7 7 1 - 8 を参照）。

30

【0159】

細胞内に核酸を輸送するための一般的方法としては、(1)化学的方法（G R A H A M E T A L . (1 9 7 3) , V I R O L O G Y 5 4 (2) : 5 3 6 - 5 3 9 ; Z A T L O U K A L E T A L . (1 9 9 2) , A N N . N . Y . A C A D . S C I . , 6 6 0 : 1 3 6 - 1 5 3) ; (2) 物理的方法、例えば、マイクロインジェクション（C A P E C C H I (1 9 8 0) , C E L L 2 2 (2) : 4 7 9 - 4 8 8)、電気穿孔（W O N G E T A L . (1 9 8 2) , B I O C H E M . B I O P H Y S . R E S . C O M M U N . 1 0 7 (2) : 5 8 4 - 5 8 7 ; F R O M M E T A L . (1 9 8 5) , P R O C . N A T ' L A C A D S C I . U S A 8 2 (1 7) : 5 8 2 4 - 5 8 2 8 ; 米国特許第5,384,253号）、および、弾丸注入法（J O H N S T O N E T A L . (1 9 9 4) , M E T H O D S C E L L . B I O L . 4 3 (A) : 3 5 3 - 3 6 5 ; F Y N A N E T A L . (1 9 9 3) , P R O C . N A T ' L A C A D . S C I . U S A 9 0 (2 4) : 1 1 4 7 8 - 1 1 4 8 2) など；(3)ウィルスベクター（C L A P P (1 9 9 3) , C L I N . P E R I N A T O L . 2 0 (1) : 1 5 5 - 1 6 8 ; L U E T A L . (1 9 9 3) , J . E X P . M E D . 1 7 8 (6) : 2 0 8 9 - 2 0 9 6 ; E G L I T I S E T A L . (1 9 8 8) , A V D . E X P . M E D . B I O L . 2 4 1 : 1 9 - 2 7 ; E G L I T I S E T A L . (1 9 8 8) , B I O T E C H N I Q U E S 6 (7) : 6 0 8 - 6 1 4) ; および、(4)受容体介在性機構（C U R I E L E T A L . (1 9 9 1) , P R O C . N A T ' L A C A D . S C I . U S A 8 8 (1 9) : 8 8 5 0 - 8 8 5 4 ; C U R I E L E T A L . (1 9 9 2) , H U M . G E N

40

50

. THER. 3 (2) : 147 - 154 ; WAGNER ET AL. (1992), PROC. NAT'L ACAD. SCI. USA 89 (13) : 6099 - 6103) が挙げられる。

【0160】

ある実施態様では、ゲノムに挿入された興味の配列を含む遺伝学的に修飾された植物が生産される。ある実施態様では、組み換えメガヌクレアーゼに対応するDNA配列及びこのメガヌクレアーゼ認識配列及び/又は標的配列とほぼ同一の配列によって側接されても側接されなくともよい興味の配列によって植物細胞をトランスフェクトし、遺伝学的に修飾された植物が生産される。別の実施態様では、遺伝学的に修飾された植物は、組み換えメガヌクレアーゼのみに対応するDNA配列によって植物細胞をトランスフェクトし、それによる切断が非相同性末端接合を促進し、認識配列を含む標的配列を破壊することによって生産される。このような実施態様では、メガヌクレアーゼ配列は、宿主植物細胞においてメガヌクレアーゼの発現を可能とする調整配列の調節下にある。これらの調整配列としては、構成的植物プロモーター、例えば、NOSプロモーター、化学的に誘発可能な遺伝子プロモーター、例えば、デキサメタゾン誘発性プロモーター（例えば、GREMILLON ET AL. (2004), PLANT J. 37 : 218 - 228を参照）、および、植物組織特異的プロモーター、例えば、LGC1プロモーター（例えば、SINGH ET AL. (2003), FEBS LETT. 542 : 47 - 52を参照）が挙げられるが、これらに限定されるものではない。

10

【0161】

植物細胞の中にDNAを導入するための適切な方法は、DNAを細胞内に導入することが可能であるならば、ほとんどどの方法であってもよく、例えば、AGROBACTERIUM感染、プロトプラストのPEG-介在性形質転換（OMIRULLEH ET AL. (1993), PLANT MOLECULAR BIOLOGY, 21 : 415 - 428）、DESICCATION/INHIBITION MEDIATED DNA UPTAKE、電気穿孔、シリコンカーバイド線維による揺動、弾丸注入、または微小弾丸発射などが挙げられるが、これらに限定されるものではない。

20

【0162】

他の実施態様では、組み換えメガヌクレアーゼを用いて、遺伝学的に修飾された動物が生産される。植物細胞の場合と同様、核酸配列を、胚細胞又は最終的にトランスジェニック生物体となる細胞の中に導入することが可能である。ある実施態様では、細胞は、受精卵であり、外因性DNA分子は、受精卵の前核の中に注入することが可能である。次に、このマイクロインジェクションされた卵は、擬似妊娠の代理母の卵管に移送され、育成される。組み換えメガヌクレアーゼは、（例えば、3-フォスフォグリセレートキナーゼなどの構成的プロモーターの調節下に）受精卵で発言され、ゲノムの一部又は不連続な少数部位における興味の配列の相同組み換えを促進する。それとは別に、GOSSLER ET AL. (1986), PROC. NAT'L ACAD. SCI. USA 83 : 9065 - 9069に記載されるように、トランスジェニック動物生成のための組み換え胚性幹（“ES”）細胞を利用することによって、遺伝学的に修飾された動物を入手することが可能である。

30

40

【0163】

ある実施態様では、組み換え哺乳類発現ベクターは、核酸の組織特異的発現を、ある特定の細胞タイプに優先的に向けて指示することが可能である。組織特異的調整要素は従来技術で既知である。適切な組織特異的プロモーターの非限定的例としては、アルブミンプロモーター（肝臓特異的、PINKERT ET AL. (1987), GENES DEV. 1 : 268 - 277）、リンパ系特異的プロモーター（CALAME AND EATON (1988), ADV. IMMUNOL. 43 : 235 - 275）、特に、T細胞受容体のプロモーター（WINOTO AND BALTIMORE (1989), EMBO J. 8 : 729 - 733）、および免疫グロブリンのプロモーター（BANERJI ET AL. (1983), CELL 33 : 729 - 740 ; QUEEN AND

50

BALTIMORE (1983), CELL 33:741-748)、ニューロン特異的プロモーター(例えば、ニューロフィラメント・プロモーター; BAYNE AND RUDDLE (1989), PROC. NATL. ACAD. SCI. USA 86:5473-5477)、膵臓特異的プロモーター(EDLUND ET AL. (1985), SCIENCE 230:912-916)、および、乳腺特異的プロモーター(例えば、ミルクフェイ・プロモーター、米国特許第4,873,316号、および欧州特許公開EP0264166)が挙げられる。発達時の調整に与るプロモーター、例えば、マウスのHOXプロモーター(KESSEL AND GRUSS (1990), SCIENCE 249:374-379)、および、 α -フェトプロテイン・プロモーター(CAMPES AND TILGHMAN (1989), GENES DEV. 3:537-546)なども含まれる。

【0164】

ある実施態様では、合理設計メガヌクレアーゼは、発現レベルまたは局在を監視するために、ペプチドエピトープ(例えば、HA、FLAG、またはMYCエピトープ)によって標識されてもよい。ある実施態様では、メガヌクレアーゼは、細胞内局在シグナル、例えば、核局在シグナル(例えば、SV40の核局在シグナル)又はクロロプラストもしくはミトコンドリア局在シグナルに融合されてもよい。別の実施態様では、メガヌクレアーゼは、それを、細胞質に局在させる核輸出シグナルに融合されてもよい。さらに、メガヌクレアーゼは、無関係のタンパク質又はタンパクドメイン、例えば、DNA-修復又は相同組み換えを刺激するタンパク質(例えば、RECA、RAD51、RAD52、RAD54、RAD57又はBRCA2)に融合されてもよい。

【0165】

6. 遺伝子治療法

本発明の局面は、遺伝子治療のための、組み換えメガヌクレアーゼの使用を可能とする。本出願で用いる「遺伝子治療」という用語は、その構造及び/又は機能において欠陥的である遺伝子又は遺伝子調整配列と置換するために、患者の中に、少なくとも一つの遺伝子の機能的コピー又は遺伝子調整配列、例えば、プロモーター、エンハンサー又はサイレンサーを導入することを含む治療処置を意味する。さらに、「遺伝子治療」という用語は、有害遺伝子または調整要素に対し、該遺伝子の発現を下げるか又は除去するために為される修飾を指すことが可能である。遺伝子治療は、先天性病態、患者の生涯に亘る特定の遺伝子座位における突然変異または損傷による病態又は感染生物体による病態を治療するために実行することが可能である。

【0166】

本発明のある局面では、機能不全遺伝子は、遺伝子発現に影響を及ぼすゲノム領域に、外因性核酸配列を挿入することによって置換されるか、不能にさせられる。ある実施態様では、組み換えメガヌクレアーゼは、病態を緩和するために、修飾すべきゲノム領域の特定配列を標的とする。この配列は、エキソン、イントロン、プロモーター内の領域又は遺伝子の機能不全発現を引き起こす、他の調整領域であってもよい。本出願で用いる「機能不全発現」という用語は、遺伝子産物の生産が少なすぎる細胞又は遺伝子産物の生産が多すぎる細胞又は異なる機能、例えば、必要機能の欠如又は必要以上の機能を持つ遺伝子産物を生産する細胞による、遺伝子産物の異常な発現を意味する。

【0167】

修飾領域に挿入される外因性核酸配列は、遺伝子を正常化する「修復」配列を提供するために使用することが可能である。遺伝子修復は、再生されるべき適正機能の実現を可能とする遺伝子中に適正遺伝子配列を導入することによって実現される。これらの実施態様では、挿入される核酸配列は、タンパクの全体コード配列であってもよいし又はある実施態様では、修復されるべき領域のみを含む遺伝子断片であってもよい。他の実施態様では、挿入される核酸配列は、異常発現または調整を引き起こす突然変異が修復されるように、プロモーター配列または、その他の調整要素を含む。他の実施態様では、挿入される核酸配列は、突然変異遺伝子には欠如する適切な翻訳終止コドンを含む。核酸配列はさらに

、適切な翻訳終止シグナルを欠如する組み換え遺伝子において翻訳を停止するための配列を持つことが可能である。

【0168】

それとは別に、核酸配列は、遺伝子の調整配列を破壊することによって又は遺伝子機能を除去するサイレンサーを提供することによって、遺伝子機能を全く除去することが可能である。ある実施態様では、外因性核酸配列は、遺伝子産物の発現を阻止するために翻訳終止コドンを提供する。別の実施態様では、外因性核酸配列は、全長のRNA分子の発現を阻止するために翻訳終止コドンを提供する。さらに別の実施態様では、遺伝子機能は、非相同的末端接合を介して、塩基挿入、塩基欠失及び/又はフレームシフト突然変異を導入することによって、メガヌクレアーゼによって直接破壊される。

10

【0169】

多くの例において、病態の原因である標的細胞又は細胞集団に対し適切な遺伝子配列を方向づけることが望ましい。治療薬の、このような標的指向は、健康な細胞が、治療薬によって標的とされることを防ぐ。これは、治療の効力を増し、一方では、治療が、健康な細胞に及ぼす可能性のある、有害となる可能性のある作用を下げる。

【0170】

興味のある細胞に対する、組み換えメガヌクレアーゼ遺伝子及び挿入すべき興味のある配列の輸送は、様々の機構によって実現することが可能である。ある実施態様では、核酸は、ウィルスの再生を阻止するように不活性化された特定のウィルス遺伝子を持つウィルスを介して細胞に輸送される。したがって、ウィルスは、標的細胞内に輸送され、維持されることは可能であるが、標的細胞または組織の中で複製する能力は保持しないように変えられる。ベクターのように活動するウィルスゲノムを生産ように、この変更ウィルスゲノムに、一つ以上のDNA配列を挿入することが可能であるが、これらの配列は、宿主のゲノムの中に導入されてその後発現されてもよいし、導入し発現されなくともよい。より具体的には、ある実施態様は、レトロウィルス、例えば、ただしこれらに限定されるものではないが、MFG又はPLJベクターなどの採用を含む。MFGベクターは、単純化されたモロニー Maus 白血病ウィルスベクター(MOMLV)であり、このベクターでは、複製欠損とするために、POL及びENVタンパク質をコードするDNA配列は欠失される。PLJレトロウィルスベクターも、MOMLVの一形である(例えば、KORMAN ET AL. (1987), PROC. NAT'L ACAD. SCI., 84: 2150 - 2154)。別の実施態様では、組み換えアデノウィルス又はアデノ関連ウィルスを、輸送ベクターとして使用することが可能である。

20

30

【0171】

別の実施態様では、標的細胞に対する、組み換えメガヌクレアーゼタンパク質及び/又は組み換えメガヌクレアーゼ遺伝子配列の輸送は、リボソームの使用によって実現される。核酸及び/又はタンパク積載物を含むリボソームの生産は、従来技術で既知である(例えば、LASIC ET AL. (1995), SCIENCE 267: 1275 - 76)。イムノリボソームは、リボソームの中に、細胞関連抗原に対する抗体を含み、特定の細胞タイプに対しメガヌクレアーゼのDNA配列、またはメガヌクレアーゼそのものを輸送することが可能である(例えば、LASIC ET AL. (1995), SCIENCE 267: 1275 - 76; YOUNG ET AL. (2005), J. CELL PHYSIOL. 187: 967 - 71; PFEIFFER ET AL. (2006), J. VASC. SURG. 43(5): 1021 - 7を参照)。リボソーム処方生産および使用のための方法は、従来技術で周知である(例えば、米国特許第6,316,024号、米国特許第6,379,699号、米国特許第6,387,397号、米国特許第6,511,676号、および米国特許第6,593,308号、およびそれらの中に引用される参考文献を参照)。ある実施態様では、リボソームは、興味のある配列の外、組み換えメガヌクレアーゼタンパク、または組み換えメガヌクレアーゼ遺伝子配列を輸送するために使用される。

40

【0172】

50

7. 病原体感染の治療法

本発明の局面はさらに、病原体による感染の治療法を提供する。病原性生物体としては、ウイルス、例えば、ただしこれらに限定されるものではないが、単純ヘルペスウイルス1、単純ヘルペスウイルス2、ヒト免疫不全ウイルス1、ヒト免疫不全ウイルス2、痘瘡ウイルス、ポリオウイルス、エプスタインバーウイルスもしくはヒトパピローマウイルスなど、及び細菌生物体、例えば、これらに限定されるものではないが、BACILLUS ANTHRACIS、HAEMOPHILUS種、PNEUMOCOCCUS種、STAPHYLOCOCCUS AUREUS、STREPTOCOCCUS種、メシシリン耐性STAPHYLOCOCCUS AUREUS及びMYCOPLASMA TUBERCULOSISなどが挙げられる。病原生物体としてはさらに、真菌生物体、例えば、

10

【0173】

ある実施態様では、合理設計メガヌクレアーゼは、病原体ゲノム内の認識配列、例えば、該病原体の増殖、生殖又は毒性にとって必須である、遺伝子または調整要素を標的とすることが可能である。ある実施態様では、認識配列は、細菌プラスミド内であってもよい。病原体ゲノムの認識配列のメガヌクレアーゼ介在性切断は、非相同的末端接合を刺激することによって、挿入、欠失又はフレームシフトの形で、標的必須遺伝子内の突然変異を刺激することが可能である。それとは別に、細菌プラスミドの切断は、その上にコードされる任意の遺伝子、例えば、トキシン遺伝子（例えば、B. ANTHRACIS致死因子遺伝子）又は抗生物質耐性遺伝子などと共にプラスミドの消失を招く可能性がある。前述のように、メガヌクレアーゼは、従来技術で一般的な技術を用いて、タンパク質又は核酸の形で、感染患者、動物又は植物に輸送してよい。ある実施態様では、メガヌクレアーゼ遺伝子は、病原体細菌類に輸送するために、バクテリオファージゲノムの中に組み込んで

20

【0174】

本発明の局面はさらに、ある形の癌の治療のための治療薬を提供する。ヒトのウイルスは、多くの場合、腫瘍形成と関連するので（例えば、エプスタインバーウイルスと鼻咽頭癌；ヒトパピローマウイルスと子宮頸部癌）、これらのウイルス性病原体の不活性化は、癌の発達または進行を阻止することがある。それとは別に、合理設計メガヌクレアーゼによる、これらの腫瘍関連ウイルスのゲノムを標的とする、二本鎖切断点を用いて、DNA損傷反応経路を介してアポトーシスを誘発してもよい。このようにして、ウイルスゲノムを抱える腫瘍細胞において選択的にアポトーシスを誘発することが可能となる場合がある。

30

【0175】

8. 遺伝子型分析法および病原体の特定

本発明の局面はさらに、インビトロ分子生物学研究および開発のためのツールを提供する。核酸、例えば、プラスミド、PCR産物、BAC配列、YAC配列、ウイルス及び真核および前核生物体由来のゲノム配列などの核酸の単離、クローニング及び操作のために、部位特異的エンドヌクレアーゼ（例えば、制限酵素）を用いることは、従来技術で普通のことである（例えば、AUSUBEL ET AL., 「分子生物学における最近のプロトコル（「CURRENT PROTOCOLS IN MOLECULAR BIOLOGY,」）」、WILEY 1999を参照されたい）。したがって、ある実施態様では、インビトロにおける核酸の操作に合理設計メガヌクレアーゼを使用してもよい。例えば、同じDNA分子の中に一对の認識配列を認識する、合理設計メガヌクレアーゼは、その後の操作、例えば、細菌プラスミド、BAC又はYACへのコンタクトなどのための介在DNAセグメントを単離するのに使用することが可能である。

40

【0176】

別の局面では、本発明は、病原遺伝子及び生物体特定用のツールを提供する。一実施態様では、合理設計メガヌクレアーゼは、健康な対立遺伝子から病的対立遺伝子を区別する

50

ために、病気と相関する多型遺伝子領域において対応する認識部位を切断するために使用することが可能である（例えば、ヒトのCFTR遺伝子のF-508対立遺伝子を認識する合理設計メガヌクレアーゼ、実施例4を参照）。この実施態様では、ヒトの患者、又は他の生物体から単離されたDNA配列が、合理設計メガヌクレアーゼによって、場合によっては、追加の部位特異的ヌクレアーゼと組み合わせて消化され、得られたDNA断片のパターンが、ゲル電気泳動、毛細管電気泳動、質量分析又はその他の従来技術で既知の方法によって分析される。この断片化パターン及び具体的には合理設計メガヌクレアーゼによる切断の有無は、ゲノム中に認識配列が存在するか否かを明らかにすることによって該生物体の遺伝子型を示す。別の実施態様では、合理設計メガヌクレアーゼは、病原性ウイルス、真菌、または細菌のゲノムにおける多型領域を標的とし、該生物体を特定するのに使用される。この実施態様では、合理設計メガヌクレアーゼは、病原体に固有の認識配列を切断するが（例えば、細菌における16Sおよび23S rRNA遺伝子の間のスペーサー領域；例えば、VAN DER GIESSEN ET AL. (1994), MICROBIOLOGY 140:1103-1108を参照）、これを用いて、ゲノムのエンドヌクレアーゼ消化及びそれに続く、電気泳動、質量分析、または、従来技術で既知の、その他の方法による断片化パターンの分析に基づいて、他の、近縁生物体と区別することが可能である。

10

【0177】

9. 特注DNA-結合ドメインの生産法

他の側面で、本発明は、エンドヌクレアーゼ切断活性を除去した合理設計DNA-結合タンパク質を提供する。合理設計メガヌクレアーゼの触媒活性の除去は、触媒に与るアミノ酸を突然変異させることによって実行することが可能である（例えば、I-CREIにおけるQ47のEへの突然変異、CHEVALIER ET AL. (2001), BIOCHEMISTRY, 43:14015-14026を参照）；I-SCEIにおけるD44またはHD145のNへの突然変異；I-CREIにおけるE66のQへの突然変異；I-MSOIにおけるD22のNへの突然変異）。次に、この不活性化メガヌクレアーゼを別のタンパク質、例えば、ただしこれらに限定されるものではないが、転写アクチベーター（例えば、GAL4トランスアクチベーションドメイン、またはVP16トランスアクチベーションドメイン）、転写リプレッサー（例えば、KRUPPELタンパクのKRABドメイン）、DNAメチラーゼドメイン（例えば、M.CVIP1、またはM.S35I）、またはヒストンアセチルトランスフェラーゼ・ドメイン（例えば、HDAC1、またはHDAC2）などのエフェクタードメインに融合させることが可能である。もっとも著明なものとして加工されたジンクフィンガードメインがある、加工されたDNA結合ドメイン及び、エフェクタードメインから成るキメラタンパクは、従来技術で既知である（例えば、PAPWORTH ET AL. (2006), GENE 366:27-38参照）。

20

30

【0178】

本発明は、下記の実施例によってさらに具体的に説明される。ただし、これらの実施例を限定的なものと考えてはならない。当業者であれば、たかだか通例の実験操作を用いるだけで、本出願に記載される特定物質および手順に対し数多くの等価物を認識し、確認することが可能であろう。そのような等価物は、下記の実施例に先行する特許請求項の範囲に含まれることが意図される。下記の実施例1-4は、I-CREIに基づく合理設計メガヌクレアーゼを特異的に言及するが、その外、I-SCEI、I-MSOI、I-CEUIに基づく合理設計メガヌクレアーゼ及びその他のLAGLIDADGメガヌクレアーゼも、本出願に記載するように、同様に生産し、使用することが可能である。

40

【実施例1】

【0179】

HIV-1 TAT遺伝子を認識するメガヌクレアーゼの合理設計

1. メガヌクレアーゼの設計

一対のメガヌクレアーゼが、HIV-1 TAT遺伝子に認められるDNA部位5' -

50

G A A G A G C T C A T C A G A A C A G T C A - 3' (配列番号15)を認識し、切断するように設計された。表5に従って、二つのメガヌクレアーゼ、T A T 1及びT A T 2が、下記の塩基接触部(非WT接触部は太字で表す)を用い、それぞれ、半分部位5' - G A A G A G C T C - 3' (配列番号16)、および5' - T G A C T G T T C - 3' (配列番号17)に結合するように設計された。

【0180】

T A T 1 :

【表14】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	G	A	A	G	A	G	C	T	C
Contact Residues	S32	Y33	N30/ Q38	R40	K28	S26/ R77	K24/ Y68	Q44	R70

10

【0181】

T A T 2 :

【表15】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	G	A	C	T	G	T	T	C
Contact Residues	C32	R33	N30/ Q38	R28/ E40	M66	S26/ R77	Y68	Q44	R70

20

【0182】

二つの酵素はクローンされ、大腸菌において発現され、下記に記載するように、対応するDNA認識配列に対する酵素活性について定量された。いずれの場合も、合理設計メガヌクレアーゼは、不活性であることが認められた。次に、DNAバックボーンとの接触を向上させるためE80をQに突然変異させた、それぞれの第2世代を生産した。この第2世代T A T 2酵素は、その意図される認識配列に対し活性を持つことが認められたが、一方、T A T 1酵素は不活性のままであった。野生型I - C R E Iとの共晶構造の肉眼検査から、T A T 1は、R40とK28の間の立体障害のために不活性であることが示唆された。この衝突を緩和するため、K28をより小さい側鎖を持つアミノ酸(A、S、T又はC)に突然変異させ、一方、Q80突然変異はそのまま維持させた、T A T 1変種を生産した。これらの酵素をE . C O L Iにおいて生産し、定量したところ、S28及びT28を有するT A T 1変種は、共に、所望の認識配列に対して活性を持つが、一方、位置-7において所望の塩基指向性を維持することが認められた。

30

【0183】

2. 組み換えメガヌクレアーゼの構築

再設計I - C R E I酵素のための突然変異を、重複PCR戦略において突然変異誘発性プライマーを用いて導入した。一次PCRで生成されたI - C R E Iの組み換えDNA断片を、二次PCRにおいて接合し、全長の組み換え核酸を生産した。組み換えI - C R E I構築体は全て、精製用遺伝子の3'末端に、6個のヒスチジンタグを融合させたPET21Aベクター(NOVAGEN CORP., SAN DIEGO、カリフォルニア州)にクローンした。核酸配列は全て、サンガーのジデオキシヌクレオチド配列決定法(SANGER ET AL. (1977), PROC. NATL. ACADEM. SCI. U.S.A. 74(12): 5463-7を参照)を用いて確認した。

40

【0184】

野生型I - C R E I及び全ての加工メガヌクレアーゼは、下記の方法を用いて発現し、精製した。PET21Aベクターにクローンした構築体は、化学的にコンピテントとした

50

B L 2 1 (D E 3) P L Y S S において形質転換され、200 μ G / M L のカルバニシリンを含む、2 X Y T プレートに撒いた。一晚育成の後、形質転換細菌コロニーを、プレートから掻き落とし、それを用いて50 M L の2 X Y T プロスに接種した。細胞を、波長600 N M において光学的濁度が0.9に達するまで、振とうしながら37 で育成した。次に、育成温度を37 から22 に下げた。タンパク質の発現は、1 M M I P T G を添加することによって誘発し、細胞は、振とうしながら2時間半インキュベートした。次に、10分間6000 X G の遠心によって細胞をペレットとした。ペレットを、1 M L の結合バッファー(20 M M T R I S - H C L , P H 8 . 0 , 500 M M N A C L , 10 M M イミダゾール)において渦巻き攪拌によって再懸濁させた。次に、細胞を、50%電力における12パルスの超音波処理によって破壊し、細胞碎片を、15分間14,000 X G の遠心によってペレットとした。細胞上清を、4 M L の結合バッファーで希釈し、200 μ L のニッケル充填金属キレート性セファローズカラム(P H A R M A C I A) に負荷した。

10

20

30

40

50

【0185】

次いで、カラムを、4 M L の洗浄バッファー(20 M M T R I S H C L , P H 8 . 0 , 500 M M N A C L , 60 M M イミダゾール)及び0.2 M L の溶出バッファー(20 M M T R I S - H C L , P H 8 . 0 , 500 M M N A C L , 400 M M イミダゾール)で洗浄した。メガヌクレアーゼ酵素は、さらに0.6 M L の溶出バッファーを添加することによって溶出し、V I V O S P I N ディスポーザブル濃縮器(I S C , I N C . , K A Y S V I L L E , ユタ州)を用いて50 - 130 μ L となるまで濃縮した。酵素は、Z E B A スピン脱塩カラム(P I E R C E B I O T E C H N O L O G Y , I N C . , R O C K F O R D , イリノイ州)を用い、定量および保存のためにS A バッファー(25 M M T R I S - H C L , P H 8 . 0 , 100 M M N A C L , 5 M M M G C L 2 , 5 M M E D T A) と交換した。酵素濃度は、23,590 M - 1 C M - 1 の吸光係数を用い280 N M における吸収によって決めた。次に、酵素の純度および分子量は、M A L D I - T O F 質量分析によって確認した。

【0186】

ヘテロダイマー酵素は、二つのタンパクを独立に精製し、それらをインビトロで混合することによって又は大腸菌における二つのタンパクの縦列発現用人工オペロンを構築することによって生産した。前者の場合、精製メガヌクレアーゼを、溶液中で1:1で混合し、DNA 基質の添加20分前に42 で予備インキュベーションした。後者の場合、二つの遺伝子を、N D E I / E C O R I 及びE C O R I / H I N D I I I を用いて順次P E T - 2 1 A 発現ベクターにクローンした。12塩基対の核酸スパーサー及びP E T 2 1 ベクター由来のシャイン - デルガルノ配列が、人工オペロン中の第1及び第2遺伝子を隔てた。

【0187】

3. 切断アッセイ

前述のように精製した酵素は全て、メガヌクレアーゼの認識配列を含む、直線状の2本鎖DNA 基質とインキュベーションすることによって活性について定量した。認識配列のセンス及びアンチセンスの両鎖に対応する合成オリゴヌクレオチドをアニールし、平滑末端コンタクトによって、P U C 1 9 プラスミドのS M A I 部位にクローンした。このクローンされた結合部位の配列は、サンガーのジデオキシヌクレオチド配列決定法によって確認した。プラスミド基質は、全て、メガヌクレアーゼ消化と同時に、X M N I , S C A I , またはB P M I によって直線化された。この酵素消化物は、5 μ L の0.05 μ M DNA 基質、2.5 μ L の5 μ M 組み換えI - C R E I メガヌクレアーゼ、9.5 μ L のS A バッファー及び0.5 μ L のX M N I , S C A I 又はB P M I を含んでいた。消化物は、37 で又はあるメガヌクレアーゼ酵素用として42 で、4時間インキュベートした。消化は、0.3 M G / M L のプロテイナーゼK及び0.5% S D S を加えることによって停止させ、37 で1時間インキュベートした。消化物を、1.5% アガロースにおいて分析し、臭化エチジウム染色によって視像化した。

【 0 1 8 8 】

メガヌクレアーゼ半分部位の指向性を評価するために、合理設計メガヌクレアーゼを、意図する半分部位の完全パリンδροームに対応する、一組のDNA基質のほか、この半分部位における27通りの可能な単一塩基対置換の内の一つとインキュベートした。

【 0 1 8 9 】

4. 認識配列 - 特異性

精製組み換えTAT1およびTAT2メガヌクレアーゼは、野生型メガヌクレアーゼ配列とは異なるDNAを認識した(図2(B))。野生型I-CREIメガヌクレアーゼは、WT認識配列を切断するが、TAT1のために意図された配列、TAT2のために意図された配列のいずれも切断しない。同様に、TAT1及びTAT2は、その意図された認識配列は切断するが、野生型配列は切断しない。次に、これらのメガヌクレアーゼを、半分部位指向性および全体特異性について評価した(図3)。野生型I-CREIは、その天然の半分部位における単一塩基対置換についてきわめて耐性が高いことが認められた。一方、TAT1及びTAT2は、きわめて特異的で、TAT1の場合は、位置-1、-2、-3、-6及び-8において、TAT2の場合は、位置-1、-2及び-6における塩基置換に対して完全に耐性を持たなかった。

【 実施例 2 】

【 0 1 9 0 】

変更DNA - 結合親和度を持つメガヌクレアーゼの合理的設計

1. 親和度を上昇および活性を上昇させたメガヌクレアーゼ

メガヌクレアーゼCCR1およびBRP2を、それぞれ、半分部位5'-AACCCCTCTC-3'(配列番号18)および5'-CTCCGGGTC-3'(配列番号19)を切断するように設計した。これらの酵素を、実施例1の場合と同様、表1に従って生産した。

【 0 1 9 1 】

CCR1:

【 表 1 6 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	A	C	C	C	T	C	T	C
Contact Residues	N32	Y33	R30/ E38	R28/ E40	E42	Q26	K24/ Y68	Q44	R70

【 0 1 9 2 】

BRP2:

【 表 1 7 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	C	T	C	C	G	G	G	T	C
Contact Residues	S32	C33	R30/ E38	R28/ E40	R42	S26/ R77	R68	Q44	R70

【 0 1 9 3 】

両酵素を大腸菌において発現させ、精製し、実施例1と同様にして定量した。両第1世代酵素は、天然の認識配列を持つ野生型I-CREIの速度よりも相当に低い速度で、その意図された認識配列を切断することが認められた。活性におけるこの低下を緩和するために、CCR1およびBRP2のDNA-結合親和度を、両酵素においてE80をQに突然変異させることによって上昇させた。これらの、CCR1及びBRP2の第2世代パー

ジョンは、触媒速度をはっきりと上昇させて、その意図された認識配列を切断することが認められた。

【0194】

2. DNA - 結合親和度を低下させ、活性を低下させたが、特異性を上昇させたメガヌクレアーゼ

野生型 I - C R E I は、その半分部位に対する置換に対して高い耐性を持つことが認められている (図 3 (A))。この酵素の特異性をさらに高めるための試みとして、通常、DNA バックボーンのリン酸塩と塩橋を造る、酵素の位置 116 のリシンを、アスパラギン酸に突然変異させて DNA - 結合親和度を下げた。この合理設計酵素の、野生型認識配列に対する切断は、はっきりと低下していたが、この組み換え酵素の特異性は、野生型よりも相当に高かった。K 116 D 変種の半分部位の指向性を、実施例 1 と同様に評価したところ、この酵素は、位置 - 1、- 2 及び - 3 における天然半分部位からの逸脱に対しては全く耐性を持たないことが認められ、半分部位の残りの 6 位置では、少なくとも部分的な塩基指向性を示した (図 3 (B))。

【実施例 3】

【0195】

合理設計メガヌクレアーゼ・ヘテロダイマー

1. 溶液の中で形成されたメガヌクレアーゼ・ヘテロダイマーによる非パリンδροーム DNA 部位の切断

二つのメガヌクレアーゼ L A M 1 及び L A M 2 が、それぞれ、半分部位 5' - T G C G G T G T C - 3' (配列番号 20) 及び 5' - C A G G C T G T C - 3' (配列番号 21) を切断するように設計された。これら二つの酵素から成るヘテロダイマーは、バクテリオファージ P 05 遺伝子に見られる DNA 配列 5' - T G C G G T G T C C G G C G A C A G C C T G - 3' (配列番号 22) を認識することが予想された。

【0196】

L A M 1 :

【表 18】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	G	C	G	G	T	G	T	C
Contact Residues	C32	R33	R30/ E38	D28/ R40	R42	Q26	R68	Q44	R70

【0197】

L A M 2 :

【表 19】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	C	A	G	G	C	T	G	T	C
Contact Residues	S32	Y33	E30/ R38	R40	K28/ E42	Q26	R68	Q44	R70

【0198】

L A M 1 及び L A M 2 は、実施例 1 に記載されるように、クローンし、大腸菌において発現し、個別に精製した。次に、二つの酵素を 1 : 1 で混合し、42 °C で 20 分インキュベートし、二つの酵素が、サブユニットを交換し、再度平衡に達するのを可能とした。L A M 1 ホモダイマー、L A M 2 ホモダイマー及び L A M 1 / L A M 2 ヘテロダイマーの混合物と予想される、得られた酵素液を、L A M 1 半分部位の完全パリンδροーム、L A M

2 半分部位の完全パリンδροーム及びバクテリオファージ ゲノムに認められる非パリンδροームハイブリッド部位に対応する、3種の異なる認識配列と共にインキュベートした。精製LAM1酵素単独は、LAM1パリンδροーム部位を切断するが、LAM2パリンδροーム部位も、LAM1/LAM2ハイブリッド部位も切断しない。同様に、精製LAM2酵素単独は、LAM2パリンδροーム部位を切断するが、LAM1部位も、LAM1/LAM2ハイブリッド部位も切断しない。しかしながら、LAM1及びLAM2の1:1混合物は、3種全てのDNA部位を切断する。LAM1/LAM2ハイブリッド部位の切断は、二つの、異なる、再設計メガヌクレアーゼが溶液の中で混合されて、非パリンδροーム的DNA部位を切断することが可能なヘテロダイマー酵素を形成することが可能であることを示す。

10

【0199】

2. 共時発現によって形成されるメガヌクレアーゼ・ヘテロダイマーによる非パリンδροームDNA部位の切断

前述のLAM1及びLAM2酵素をコードする遺伝子は、実施例1において記載したように、大腸菌における同時発現のためにオペロンの中に配置された。この共時発現酵素は、実施例1と同様に精製され、かつ、酵素混合物は、前述の、三つの可能性のある認識配列とインキュベートされた。この共時発現酵素は、LAM1/LAM2ハイブリッド部位を含む三つ全ての部位を切断することが認められたが、これは、二つの、異なる合理設計メガヌクレアーゼが共時発現されて、非パリンδροームDNA部位を切断することが可能なヘテロダイマー酵素を形成することを示す。

20

【0200】

3. タンパク-タンパクインターフェイスを修飾されたメガヌクレアーゼ・ヘテロダイマーによる、非パリンδροームDNA部位の優先的切断

非パリンδροームDNA部位の切断を必要とする応用のために、異なる(パリンδροーム的)DNA部位を認識し、切断するホモダイマーの形成を最小に留めながら、酵素ヘテロダイマーの形成を増進することが望ましい。このために、位置7、57及び96におけるリシンがグルタミン酸に変更されたLAM1酵素の変種が生産された。次に、この酵素を、位置8及び61のグルタミン酸をリシンに変えたLAM2の変種と、前述と同様に共時発現し、精製した。この場合、LAM1ホモダイマーの形成は、一方のモノマーにおけるE7、E57及びE96と、他方のモノマーにおけるE8及びE61の間における静電的反発によって低下すると予想された。同様に、LAM2ホモダイマーの形成は、一方のモノマーにおけるK7、K57及びK96と、他方のモノマーにおけるK8及びK61の間における静電的反発によって低下すると予想された。逆に、LAM1/LAM2ヘテロダイマーは、LAM1におけるE7、E57、およびE96、およびLAM2におけるK8及びK61の間の静電的牽引によって好まれることが予想された。インターフェイスが修飾されたこの二つのメガヌクレアーゼが、前述のように、共時発現され、定量されると、LAM1/LAM2ハイブリッド部位が、二つのパリンδροーム部位に対し優先的に切断されることが認められた。これは、メガヌクレアーゼのタンパク-タンパクインターフェイスにおける置換が、ヘテロダイマーの優先的形成の駆動が可能であることを示す。

30

40

【実施例4】

【0201】

生理的DNA配列を切断する、追加のメガヌクレアーゼ・ヘテロダイマー

1. 遺伝子治療に関連するDNA配列を切断するメガヌクレアーゼ・ヘテロダイマー

軟骨無形成症を引き起こす突然変異である、ヒトのFGFR3遺伝子における配列5'-CTGGGAGTCTCAGGACAGCCCTG-3'(配列番号23)を切断する、合理設計メガヌクレアーゼ・ヘテロダイマー(ACH1/ACH2)を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【0202】

50

A C H 1 :

【表 2 0】

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>C33</u>	<u>E30/</u> <u>R38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>A26/</u> <u>Q77</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

【 0 2 0 3】

10

A C H 2 :

【表 2 1】

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>Y33</u>	<u>E30/</u> <u>R38</u>	<u>R40</u>	<u>K28/</u> <u>E42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

【 0 2 0 4】

20

ヒトの成長ホルモン遺伝子のプロモーターにおいて配列 5' - C C A G G T G T C T C T G G A C T C C T C C - 3' (配列番号 24) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (HGH1 / HGH2) を生産することが可能である。例えば、前述したように、I - C R E I メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【 0 2 0 5】

H G H 1 :

【表 2 2】

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>C33</u>	<u>N30/</u> <u>Q38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

30

【 0 2 0 6】

H G H 2 :

【表 2 3】

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>K32</u>	<u>R33</u>	<u>N30/</u> <u>Q38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>A26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

40

【 0 2 0 7】

ヒトの C F T R 遺伝子の F 5 0 8 対立遺伝子における 5' - G A A A A T A T C A T T G G T G T T T C C T - 3' (配列番号 25) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (CF1 / CF2) を生産することが可能である。例えば、前述したように、I - C R E I メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位

50

を持つメガヌクレアーゼが設計された。

【0208】

CF1:

【表24】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	G	A	A	A	A	T	A	T	C
Contact Residues	S32	Y33	N30/ Q38	Q40	K28	Q26	H68/ C24	Q44	R70

10

【0209】

CF2:

【表25】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	G	G	A	A	A	C	A	C
Contact Residues	N32	R33	E30/ R38	Q40	K28	A26	Y68/ C24	T44	R70

20

【0210】

ヒトのCCR5遺伝子(HIVコレセプター)の配列5'-AACCCCTCTCCAGTGAGATGCCT-3'(配列番号26)を切断する合理設計メガヌクレアーゼ・ヘテロダイマー(CCR1/CCR2)を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【0211】

CCR1:

【表26】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	A	C	C	C	T	C	T	C
Contact Residues	N32	Y33	R30/ E38	E40/ R28	E42	Q26	Y68/ K24	Q44	R70

30

【0212】

CCR2:

【表27】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	G	G	C	A	T	C	T	C
Contact Residues	N32	R33	E30/ R38	E40	K28	Q26	Y68/ K24	Q44	R70

40

【0213】

ヒトのDMキナーゼ遺伝子の3'非翻訳領域における5'-GACCTCGTCCCTCCGACTCGCTG-3'(配列番号27)を切断する合理設計メガヌクレアーゼ・ヘテロダイマー(MYD1/MYD2)を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

50

【 0 2 1 4 】

M Y D 1 :

【 表 2 8 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	G	A	C	C	T	C	G	T	C
Contact Residues	S32	Y33	R30/ E38	E40/ R28	K66	Q26/ E77	R68	Q44	R70

【 0 2 1 5 】

M Y D 1 :

【 表 2 9 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	C	A	G	C	G	A	G	T	C
Contact Residues	S32	Y33	E30/ R38	E40/ R28	R42	A26/ Q77	R68	Q44	R70

【 0 2 1 6 】

2. 病原体ゲノムのDNA配列を切断するメガヌクレアーゼ・ヘテロダイマー

単純ヘルペスウイルス-1および単純ヘルペスウイルス-2のUL36遺伝子の配列5'-CTCGATGTCGGACGACACGGCA-3' (配列番号28)を切断する合理設計メガヌクレアーゼ・ヘテロダイマー(HSV1/HSV2)を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【 0 2 1 7 】

H S V 1 :

【 表 3 0 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	C	T	C	G	A	T	G	T	C
Contact Residues	S32	C33	R30/ E38	R40/ R28	Q42/ K28	Q26	R68	Q44	R70

【 0 2 1 8 】

H S V 2 :

【 表 3 1 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	G	C	C	G	T	G	T	C
Contact Residues	C32	R33	R30/ E38	E40/ R28	R42	Q26	R68	Q44	R70

【 0 2 1 9 】

BACILLUS ANTHRACISゲノムの配列5'-ACAAGTGTCTATGGACAGTTTA-3' (配列番号29)を切断する合理設計メガヌクレアーゼ・ヘテロダイマー(ANT1/ANT2)を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部

10

20

30

40

50

位を持つメガヌクレアーゼが設計された。

【 0 2 2 0 】

A N T 1 :

【 表 3 2 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	C	A	A	G	T	G	T	C
Contact Residues	<u>N32</u>	<u>C33</u>	<u>N30/</u> <u>Q38</u>	<u>Q40/</u> <u>A28</u>	<u>R42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

10

【 0 2 2 1 】

A N T 2 :

【 表 3 3 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	A	A	A	C	T	G	T	C
Contact Residues	<u>C32</u>	<u>Y33</u>	<u>N30/</u> <u>Q38</u>	<u>Q40</u>	<u>E42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

20

【 0 2 2 2 】

痘瘡（天然痘）ウィルス G P 0 0 9 遺伝子の配列 5' - A A A A C T G T C A A A T G A C A T C G C A - 3'（配列番号 30）を切断する合理設計メガヌクレアーゼ・ヘテロダイマー（P O X 1 / P O X 2）を生産することが可能である。例えば、前述したように、I - C R E I メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【 0 2 2 3 】

P O X 1 :

【 表 3 4 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	A	A	A	C	T	G	T	C
Contact Residues	<u>N32</u>	<u>C33</u>	<u>N30/</u> <u>Q38</u>	<u>Q40</u>	<u>K28</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

30

【 0 2 2 4 】

P O X 2 :

【 表 3 5 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	G	C	G	A	T	G	T	C
Contact Residues	<u>C32</u>	<u>R33</u>	<u>R30/</u> <u>E38</u>	<u>R40</u>	<u>C28/</u> <u>Q42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

40

【 0 2 2 5 】

エプスタインバーウィルス B A L F 2 遺伝子の配列 5' - C G G G G T C T C G T G C G A G G C C T C C - 3'（配列番号 31）を切断する合理設計メガヌクレアーゼ・ヘテ

50

ロダイマー (E B B 1 / E B B 2) を生産することが可能である。例えば、前述したように、 I - C R E I メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【 0 2 2 6 】

E B B 1 :

【 表 3 6 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
Contact Residues	<u>S32</u>	<u>R33</u>	<u>D30/</u> <u>Q38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

10

【 0 2 2 7 】

E B B 1 :

【 表 3 7 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>
Contact Residues	<u>S32</u>	<u>R33</u>	<u>D30/</u> <u>Q38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

20

【 0 2 2 8 】

3 . 植物ゲノムにおける DNA 配列を切断するメガヌクレアーゼ・ヘテロダイマー

A R A B I D O P S I S T H A L I A N N A G L 2 遺伝子の配列 5' - C A C T A A C T C G T A T G A G T C G G T G - 3' (配列番号 3 2) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (G L A 1 / G L A 2) を生産することが可能である。例えば、前述したように、 I - C R E I メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

30

【 0 2 2 9 】

G L A 1 :

【 表 3 8 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	<u>C</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>C</u>
Contact Residues	<u>S32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>S40/</u> <u>C79</u>	<u>K28</u>	<u>A26/</u> <u>Q77</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

40

【 0 2 3 0 】

G L A 2 :

【 表 3 9 】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	<u>C</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>C</u>
Contact Residues	<u>S32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>E40/</u> <u>R28</u>	<u>R42</u>	<u>A26</u> <u>Q77</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

50

【 0 2 3 1 】

ARABIDOPSIS THALI ANNA BP1 遺伝子の配列 5' - T G C C T C C T C T A G A G A C C C G G A G - 3' (配列番号 33) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (BRP1 / BRP2) を生産することが可能である。例えば、前述したように、I - CREI メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

【 0 2 3 2 】

BRP1 :

【表 4 0】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	G	C	C	T	C	C	T	C
Contact Residues	C32	R33	R30/ E38	R28/ E40	K66	Q26/ E77	Y68/ K24	Q44	R70

10

【 0 2 3 3 】

BRP2 :

【表 4 1】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	C	T	C	C	G	G	G	T	C
Contact Residues	S32	C33	R30/ E38	E40/ R28	R42	S26 R77	R68	Q44	R70

20

【 0 2 3 4 】

NICOTIANA TABACUM マグネシウムケラターゼ遺伝子の配列 5' - T A A A T C T C T A A G G T C T G T G C A - 3' (配列番号 34) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (MGC1 / MGC2) を生産することが可能である。例えば、前述したように、I - CREI メガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

30

【 0 2 3 5 】

MGC1 :

【表 4 2】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	A	A	A	A	T	C	T	C
Contact Residues	C32	Y33	N30/ Q38	Q40/	K28	Q26	Y68/ K24	Q44	R70

40

【 0 2 3 6 】

MGC2 :

【表 4 3】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	G	C	A	C	A	G	A	C
Contact Residues	S32	R33	R30/ E38	Q40	K28	A26 Q77	R68	T44	R70

【0237】

NICOTIANA TABACUM CYP82E4 遺伝子の配列 5' - CAAGA
ATTCAAGCGAGCATTA - 3' (配列番号 35) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (CYP/HGH2) を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

10

【0238】

CYP:

【表 4 4】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	C	A	A	G	A	A	T	T	C
Contact Residues	D32	Y33	N30/ Q38	R40/	K28	Q77/ A26	Y68	Q44	R70

20

【0239】

HGH2:

【表 4 5】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	T	A	A	T	G	C	T	C
Contact Residues	S32	C33	N30/ Q38	Q40	K66	R77/ S26	Y68 K24	Q44	R70

30

【0240】

4. 酵母ゲノムにおける DNA 配列を切断するメガヌクレアーゼ・ヘテロダイマー

SACCHAROMYCES CEREVISIAE URA3 遺伝子の配列 5' - T
TAGATGACAAGGGAGACGCAT - 3' (配列番号 36) を切断する合理設計メガヌクレアーゼ・ヘテロダイマー (URA1/URA2) を生産することが可能である。例えば、前述したように、I-CREIメガヌクレアーゼに基づいて、下記の接触残基および認識配列半分部位を持つメガヌクレアーゼが設計された。

40

【0241】

URA1:

【表 4 6】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	T	T	A	G	A	T	G	A	C
Contact Residues	S32	C33	N30/ Q38	R40	K28	Q26	R68	T44	R70

【0242】

50

U R A 2 :
【表 4 7】

Position	-9	-8	-7	-6	-5	-4	-3	-2	-1
Base	A	T	G	C	G	T	C	T	C
Contact Residues	N32	C33	E30/ R38	E40/ R28	R42	Q26	Y68/ K24	Q44	R70

【 0 2 4 3】

5 . 認識配列特異性

本実施例において上に概説した合理設計メガヌクレアーゼを、実施例 1 と同様にして、クローンし、大腸菌において発現し、精製した。次いで、各精製メガヌクレアーゼを、そのヘテロダイマー形成パートナー（例えば、A C H 1 を A C H 2 と、H G H 1 を H G H 2 と、など）と 1 : 1 で混合し、各メガヌクレアーゼ・ヘテロダイマーのために意図される非パリンドローム D N A 認識配列を含む、直線化 D N A 基質とインキュベートした。図 3 に示すように、各合理設計メガヌクレアーゼ・ヘテロダイマーは、その意図された D N A 部位を切断する。

【図面の簡単な説明】

【 0 2 4 4】

【図 1】図 1 (A) は、結晶学データに基づいて、I - C R E I ホモダイマー及びその天然の 2 本鎖認識配列の相互作用を示す。本模式図は、二つの卵形として示すホモダイマーによって結合される、図示目的のためだけに巻き戻された状態で示される認識配列（配列番号 2 及び配列番号 3）を描いたものである。各 D N A 半分部位の塩基には、- 1 から - 9 までの番号がふられ、認識表面を形成する、I - C R E I のアミノ酸残基は、1 文字アミノ酸表示および残基位置を示す数字によって表される。実線：D N A 塩基に対する水素結合。鎖線：酵素設計では新たな接触点を形成するが、野生型複合体では D N A と接触しないアミノ酸位置。矢印：D N A バックボーンと相互作用を持ち、切断活性に影響を及ぼす残基。また、図 1 (B) は、図 1 (A) の右側の切断半分部位の - 4 位置における A - T 塩基対間の野生型接触を示す。具体的に言うと、残基 Q 2 6 が、A 塩基と相互作用を持つところが示される。残基 I 7 7 は、該塩基対に近接するが、特異的相互作用は持たない。更に、図 1 (C) は、残基 I 7 7 が修飾されて E 7 7 に変更された、I - C R E I メガヌクレアーゼの、合理設計変種との間の相互作用を示す。この変更の結果、- 4 位置では G - C 塩基対が好まれる。Q 2 6 と G 塩基との間の相互作用は、図 1 (A) の左側の切断半分部位において結晶学的に観察されるように、水分子によって仲介される。更に又、図 1 (D) は、残基 Q 2 6 が修飾されて E 2 6 に変更され、残基 I 7 7 が修飾されて R 7 7 に変更された、I - C R E I メガヌクレアーゼの合理設計変種との間の相互作用を示す。この変更の結果、- 4 位置では G - C 塩基対が好まれる。そして更に、図 1 (E) は、残基 Q 2 6 が修飾されて A 2 6 に変更され、残基 I 7 7 が修飾されて Q 7 7 に変更された、I - C R E I メガヌクレアーゼの合理設計変種との間の相互作用を示す。この変更の結果、- 4 位置では T - A 塩基対が好まれる。

【図 2】図 2 (A) は、野生型 I - C R E I メガヌクレアーゼ (W T) 及び本発明の 1 1 種の合理設計メガヌクレアーゼ・ヘテロダイマーそれぞれの一認識配列の比較を示す。W T 認識配列に対して保存される塩基には陰影を施す。9 B P の半分部位は太字で示す。W T : 野生型 (配列番号 4) ; C F : 多くの嚢胞性線維症例の原因となる、ヒト C F T R 遺伝子の F 5 0 8 対立遺伝子 (配列番号 2 5) ; M Y D : 筋緊張性ジストロフィーと関連するヒト D M キナーゼ遺伝子 ; C C R : ヒト C C R 5 遺伝子 (大型 H I V コレセプター) (配列番号 2 6) ; A C H ; 軟骨無形成症と関連するヒトの F G F R 3 遺伝子 (配列番号 2 3) ; T A T : H I V - 1 T A T / R E V 遺伝子 (配列番号 1 5) ; H S V : H S V - 1 U L 3 6 遺伝子 (配列番号 2 8) ; L A M : パクテリオファージ P 0 5 遺伝子 (配列番号 2 2) ; P O X : 痘瘡 (天然痘) ウィルス G P 0 0 9 遺伝子 (配列番号 3 0) ;

10

20

30

40

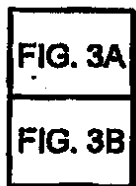
50

URA : SACCHAROMYCES CERVISIAE URA3 遺伝子 (配列番号 36) ; GLA : ARABIDOPSIS THALIANA GL2 遺伝子 (配列番号 32) ; BRP : ARABIDOPSIS THALIANA BP-1 遺伝子 (配列番号 33) 。 図 2 (B) は、野生型 I - CREI (WT) 及び 11 種の合理設計メガヌクレアーゼ・ヘテロダイマーのそれぞれと、12 種全ての酵素の認識部位を担持するプラスミドと、37 で 6 時間インキュベートした場合の結果を示す。各ボックスに切断パーセントが示される。

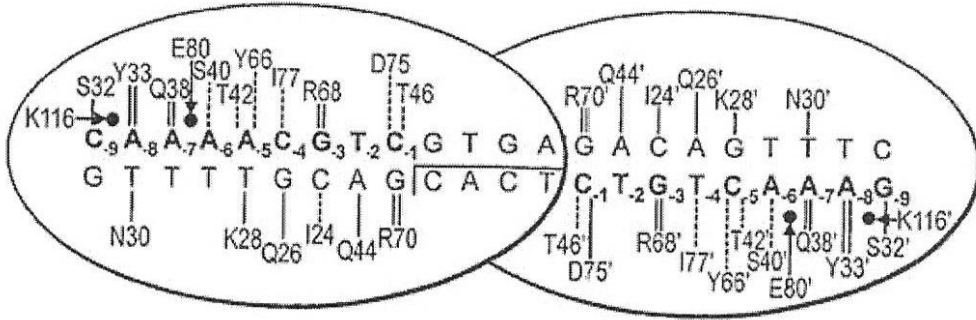
【図 3】 図 3 は、野生型および合理設計 I - CREI ホモダイマーの切断パターンを示す。(A) 野生型 I - CREI 。 (B) I - CREI K116D 。 (C - L) 本発明の合理設計メガヌクレアーゼ。酵素は、27 通りの対応する単一塩基対変動を持つ、意図される切断半分部位のパリンドロームを担持する、一組のプラスミドと共にインキュベートした。黒色バー：表 1 に基づいて予想される切断パターン。灰色バー：予想される切断パターンから変位する DNA 部位。円は、意図される認識部位における塩基を示す。さらに、2 時間における切断の時間経過も示される。C および L における白抜き円の時間経過プロットは、E80Q 突然変異を欠く CCR1 および BRP2 酵素による切断に対応する。切断部位は、図 2 (A) に記載されるヘテロダイマー酵素の、5' (左コラム) 、および 3' (右コラム) 半分部位に対応する。

10

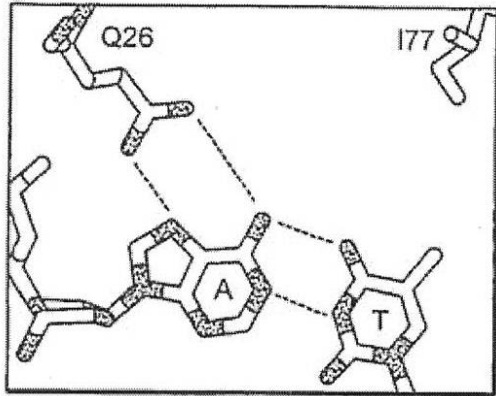
【図 3】



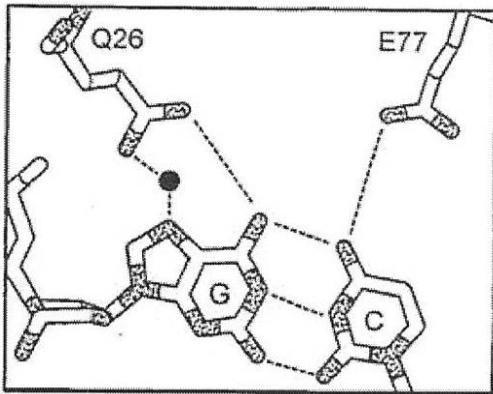
【 図 1 A 】



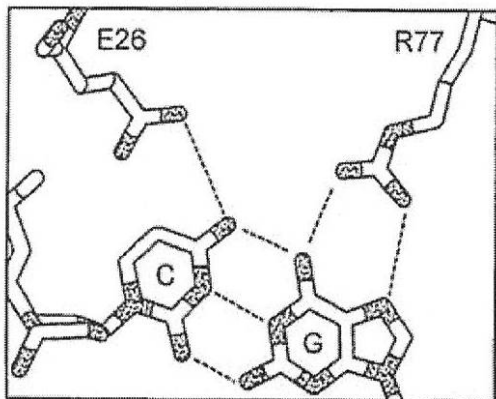
【 図 1 B 】



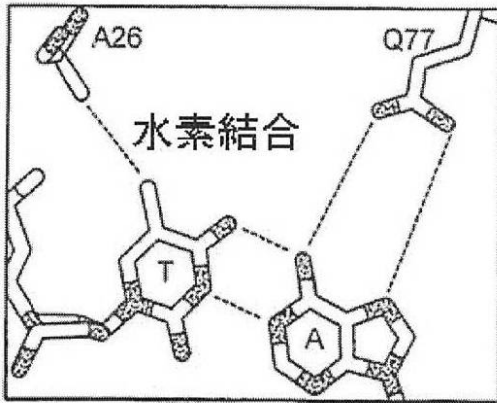
【 図 1 C 】



【 図 1 D 】



【図1E】



【図2A】

WT **C A A A C T G T C G T G A G A C A G T T T G**

CF1/CF2 **G A A A A T A T C A T T G G T G T T T C C T**

MYD1/MYD2 **G A C C T C G T C C T C C G A C T C G C T G**

CCR1/CCR2 **A A C C C T C T C C A G T G A G A T G C C T**

ACH1/ACH2 **C T G G G A G T C T C A G G A C A G C C T G**

TAT1/TAT2 **G A A G A G C T C A T C A G A A C A G T C A**

HSV1/HSV2 **C T C G A T G T C G G A C G A C A C G G C A**

LAM1/LAM2 **T G C G G T G T C C G G C G A C A G C C T G**

POX1/POX2 **A A A A C T G T C A A A T G A C A T C G C A**

URA1/URA2 **T T A G A T G A C A A G G G A G A C G C A T**

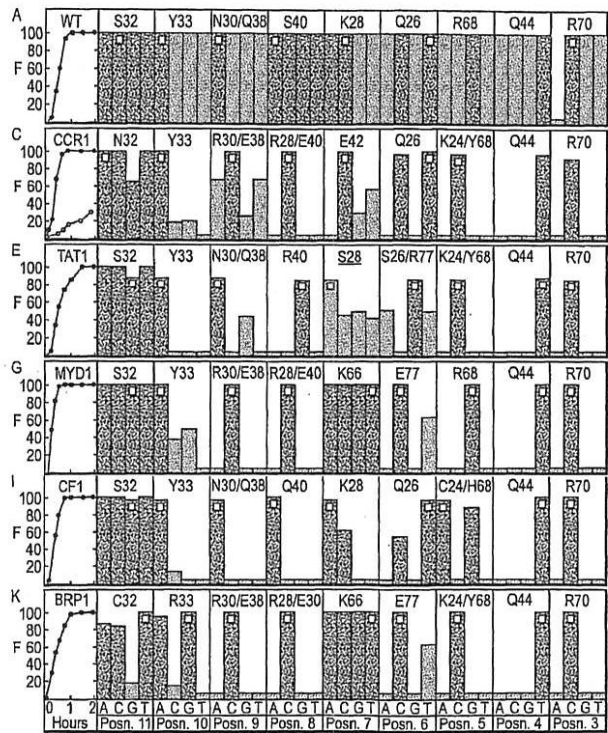
GLA1/GLA2 **C A C T A A C T C G T A T G A G T C G G T G**

BRP1/BRP2 **T G C C T C C T C T A G A G A C C C G G A G**

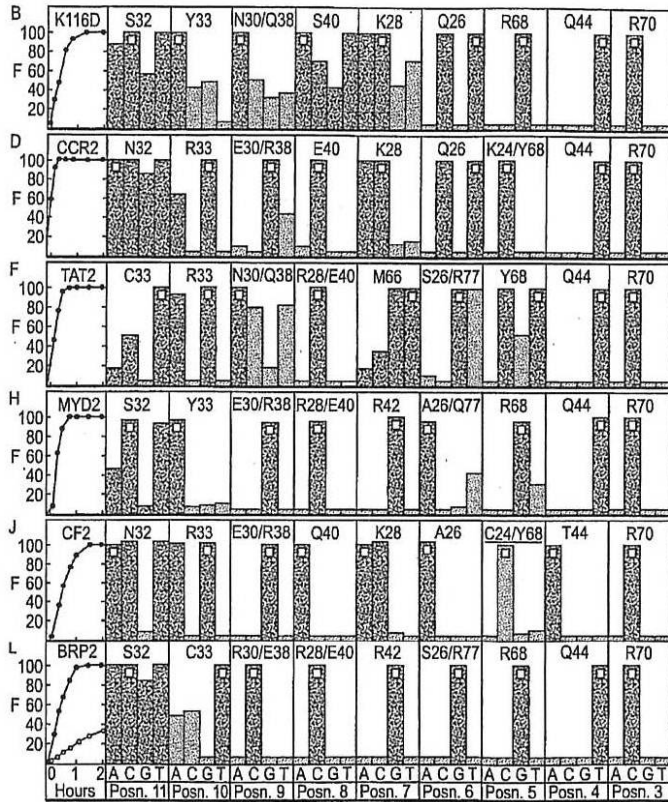
【 図 2 B 】

		酵素											
サイト	WT	CF	CCR	TAT	LAM	URA	BRP						
		MYD	ACH	HSV	POX	GLA	WT	CF	CCR	TAT	LAM	URA	BRP
WT	100	0	0	0	0	0	47	0	100	100	0	0	Wild Type
CF1/CF2	0	63	0	0	0	0	0	0	0	0	0	0	Cystic Fibrosis
MYD1/MYD2	0	0	94	0	0	0	16	0	0	0	0	0	Myotonic Dystrophy
CCR1/CCR2	0	0	0	100	0	0	0	0	0	0	0	0	CCR5
ACH1/ACH2	0	0	0	0	87	0	0	0	0	0	0	0	Achondroplasia
TAT1/TAT2	0	0	0	0	0	81	0	0	0	0	0	0	HIV-1
HSV1/HSV2	0	0	0	0	0	0	67	0	0	0	0	0	Herpes Simplex 1
LAM1/LAM2	0	0	0	0	0	0	0	87	0	0	0	0	Lambda Phage
POX1/POX2	100	0	0	0	0	0	0	0	100	0	0	0	Variola Virus
URA1/URA2	0	0	0	0	0	0	0	0	0	70	0	0	<i>S. cerevisiae</i> URA3
GLA1/GLA2	0	0	0	0	0	0	0	0	0	0	88	0	<i>A. thaliana</i> GL2
BRP1/BRP2	0	0	0	0	0	0	0	0	0	0	0	100	<i>A. thaliana</i> BP1

【 図 3 A 】



【図 3 B】



【配列表】

2012196214000001.app

【手続補正書】

【提出日】平成24年6月1日(2012.6.1)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

野生型 I - C R E I メガヌクレアーゼと比べ、2本鎖DNAに対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号1の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

DNA 結合親和度が、

(A) H、N、Q、S、T、K 又は R による E 8 0、D 1 3 7、I 8 1、L 1 1 2、P 2 9、V 6 4 又は Y 6 6 の置換；

又は、

(B) K 又は R による T 4 6、T 1 4 0 又は T 1 4 3 の置換、

から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって上昇することを特徴とする、組み換えメガヌクレアーゼ。

【請求項 2】

野生型 I - C R E I メガヌクレアーゼと比べ、2本鎖DNAに対する結合親和度が変更された組み換えメガヌクレアーゼであって、

配列番号1の I - C R E I メガヌクレアーゼの残基 2 - 1 5 3 に対して少なくとも 8 5 % の配列類似性を有するポリペプチドを含み；

DNA結合親和度が、

(A) H、N、Q、S、T、D又はEによるK34、K48、R51、K82、K116、またはK139の置換；

又は、

(B) D又はEによるI81、L112、P29、V64、Y66、T46、T140又はT143の置換から成る群より選ばれる一つの置換に対応する少なくとも一つの修飾によって低下することを特徴とする、組み換えメガヌクレアーゼ。

【請求項3】

真核細胞の染色体に挿入される興味の外因性配列を含む遺伝学的に修飾された真核細胞を生産するための方法であって、

真核細胞を、

(I) メガヌクレアーゼをコードする第1核酸配列及び、

(II) 前記興味の配列を含む第2核酸配列

を含む一つ以上の核酸によってトランスフェクトすることを含み、

前記メガヌクレアーゼが前記染色体の中に切断部位を生産し、かつ、前記興味の配列が前記切断部位において前記染色体の中に挿入され、

前記メガヌクレアーゼが請求項1又は2記載の組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項4】

27

前記第2核酸が、前記切断部位に側接する配列に対し相同な配列をさらに含み、前記興味の配列が、相同組み換えによって前記切断部位において挿入される、請求項3記載の方法。

【請求項5】

前記第2核酸が、前記切断部位に対し実質的相同性を欠き、前記興味の配列が、非相同的末端接合によって前記染色体に挿入される、請求項3記載の方法。

【請求項6】

真核細胞の染色体に挿入される興味の外因性配列を含む、遺伝学的に修飾された真核細胞を生産するための方法であって、

真核細胞にメガヌクレアーゼタンパクを導入すること；

及び、

前記興味の配列を含む核酸によって前記真核細胞をトランスフェクトすることを含み、

前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記興味の配列が、前記切断部位において前記染色体の中に挿入され；

かつ、

前記メガヌクレアーゼが請求項1又は2記載の組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項7】

前記核酸が、前記切断部位に側接する配列に対し相同な配列をさらに含み、前記興味の配列が、相同組み換えによって前記切断部位において挿入される、請求項6記載の方法。

【請求項8】

前記核酸が、前記切断部位に対し実質的相同性を欠き、前記興味の配列が、非相同的末端接合によって前記染色体に挿入される、請求項6記載の方法。

【請求項9】

真核細胞の染色体中の標的配列を破壊することによって遺伝学的に修飾された真核細胞を生産するための方法であって、

メガヌクレアーゼをコードする核酸によって真核細胞をトランスフェクトすることを含み、

前記メガヌクレアーゼが、前記染色体の中に切断部位を生産し、前記標的配列が、前記切断部位において非相同的末端接合によって破壊され；

かつ、

前記メガヌクレアーゼが、請求項 1 又は 2 記載の組み換えメガヌクレアーゼであることを特徴とする方法。

【請求項 10】

遺伝学的に修飾された生物体を生産する方法であって、

請求項 3 ~ 9 のいずれか 1 項の方法に従って遺伝学的に修飾された真核細胞を生産するステップと、

及び、

前記遺伝学的に修飾された真核細胞を育成して遺伝学的に修飾された生物体を生産するステップと、

を含む方法。

【請求項 11】

前記真核細胞が、配偶子、接合子、胚盤胞細胞、胚幹細胞及びプロトプラスト細胞から成る群より選ばれる、請求項 10 記載の方法。

【請求項 12】

DNA - 結合親和度を上昇させる組み換えメガヌクレアーゼを合理的に設計する方法であって、

(1) 参照メガヌクレアーゼ - DNA 複合体の三次元構造の少なくとも一部を決定すること；

(2) バックボーン接触面を形成するアミノ酸接触残基を特定すること；

(3) 前記 DNA - 結合親和度を上昇させるアミノ酸置換を、

(A) 陰性荷電、または疎水性側鎖を持つ接触残基については、非荷電 / 極性、または陽性荷電側鎖を持つ置換を選択すること；

又は

(B) 非荷電 / 極性側鎖を持つ接触残基については、陽性荷電側鎖を持つ置換を選択することによって特定すること、

を含む方法。

【請求項 13】

DNA - 結合親和度を低下させる組み換えメガヌクレアーゼを合理的に設計する方法であって、

(1) 参照メガヌクレアーゼ - DNA 複合体の三次元構造の少なくとも一部を決定すること；

(2) バックボーン接触面を形成するアミノ酸接触残基を特定すること；

(3) 前記 DNA - 結合親和度を低下させるアミノ酸置換を、

(A) 陽性荷電側鎖を持つ接触残基については、非荷電 / 極性、または陰性荷電側鎖を持つ置換を選択すること；

又は

(B) 疎水性、または非荷電 / 極性側鎖を持つ接触残基については、陰性荷電側鎖を持つ置換を選択することによって特定すること、

を含む方法。

フロントページの続き

(72)発明者 ジャンツ, デイルク

アメリカ合衆国 ノースカロライナ州 27705, ダーハム, 914 ナインス ストリート

Fターム(参考) 4B024 AA01 AA20 CA01 CA03 CA09 CA11 CA20 DA01 DA02 DA11

EA04 GA11 HA01 HA11 HA17

4B050 CC03 CC04 DD11 LL01

4B065 AA57X AA83X AA87X AA90X AA90Y AB01 AC14 BA01 CA44

【外国語明細書】

**RATIONALLY-DESIGNED MEGANUCLEASES
WITH ALTERED SEQUENCE SPECIFICITY AND DNA-BINDING AFFINITY**

RELATED APPLICATION

[0001] This application claims benefit of priority to U.S. Provisional Patent Application No. 60/727,512, filed October 18, 2005, the entire disclosure of which is hereby incorporated by reference.

GOVERNMENT SUPPORT

[0002] The invention was supported in part by grants 2R01-GM-0498712, 5F32-GM072322 and 5 DP1 OD000122 from the National Institute of General Medical Sciences of National Institutes of Health of the United States of America. Therefore, the U.S. government may have certain rights in the invention.

FIELD OF THE INVENTION

[0003] The invention relates to the field of molecular biology and recombinant nucleic acid technology. In particular, the invention relates to rationally-designed, non-naturally-occurring meganucleases with altered DNA recognition sequence specificity and/or altered affinity. The invention also relates to methods of producing such meganucleases, and methods of producing recombinant nucleic acids and organisms using such meganucleases.

BACKGROUND OF THE INVENTION

[0004] Genome engineering requires the ability to insert, delete, substitute and otherwise manipulate specific genetic sequences within a genome, and has numerous therapeutic and biotechnological applications. The development of effective means for genome modification remains a major goal in gene therapy, agrotechnology, and synthetic biology (Porteus *et al.* (2005), *Nat. Biotechnol.* 23: 967-73; Tzfira *et al.* (2005), *Trends Biotechnol.* 23: 567-9; McDaniel *et al.* (2005), *Curr. Opin. Biotechnol.* 16: 476-83). A common method for inserting or modifying a DNA sequence involves introducing a transgenic DNA sequence flanked by sequences homologous to the genomic target and

selecting or screening for a successful homologous recombination event. Recombination with the transgenic DNA occurs rarely but can be stimulated by a double-stranded break in the genomic DNA at the target site. Numerous methods have been employed to create DNA double-stranded breaks, including irradiation and chemical treatments. Although these methods efficiently stimulate recombination, the double-stranded breaks are randomly dispersed in the genome, which can be highly mutagenic and toxic. At present, the inability to target gene modifications to unique sites within a chromosomal background is a major impediment to successful genome engineering.

[0005] One approach to achieving this goal is stimulating homologous recombination at a double-stranded break in a target locus using a nuclease with specificity for a sequence that is sufficiently large to be present at only a single site within the genome (see, e.g., Porteus *et al.* (2005), *Nat. Biotechnol.* 23: 967-73). The effectiveness of this strategy has been demonstrated in a variety of organisms using chimeric fusions between an engineered zinc finger DNA-binding domain and the non-specific nuclease domain of the FokI restriction enzyme (Porteus (2006), *Mol Ther* 13: 438-46; Wright *et al.* (2005), *Plant J.* 44: 693-705; Urnov *et al.* (2005), *Nature* 435: 646-51). Although these artificial zinc finger nucleases stimulate site-specific recombination, they retain residual non-specific cleavage activity resulting from under-regulation of the nuclease domain and frequently cleave at unintended sites (Smith *et al.* (2000), *Nucleic Acids Res.* 28: 3361-9). Such unintended cleavage can cause mutations and toxicity in the treated organism (Porteus *et al.* (2005), *Nat. Biotechnol.* 23: 967-73).

[0006] A group of naturally-occurring nucleases which recognize 15-40 base-pair cleavage sites commonly found in the genomes of plants and fungi may provide a less toxic genome engineering alternative. Such "meganucleases" or "homing endonucleases" are frequently associated with parasitic DNA elements, such as group 1 self-splicing introns and inteins. They naturally promote homologous recombination or gene insertion at specific locations in the host genome by producing a double-stranded break in the chromosome, which recruits the cellular DNA-repair machinery (Stoddard (2006), *Q. Rev. Biophys.* 38: 49-95). Meganucleases are commonly grouped into four families: the LAGLIDADG family, the GIY-YIG family, the His-Cys box family and the HNH family. These families are characterized by structural motifs, which affect catalytic activity and

recognition sequence. For instance, members of the LAGLIDADG family are characterized by having either one or two copies of the conserved LAGLIDADG motif (see Chevalier *et al.* (2001), *Nucleic Acids Res.* 29(18): 3757-3774). The LAGLIDADG meganucleases with a single copy of the LAGLIDADG motif form homodimers, whereas members with two copies of the LAGLIDADG motif are found as monomers. Similarly, the GIY-YIG family members have a GIY-YIG module, which is 70–100 residues long and includes four or five conserved sequence motifs with four invariant residues, two of which are required for activity (see Van Roey *et al.* (2002), *Nature Struct. Biol.* 9: 806-811). The His-Cys box meganucleases are characterized by a highly conserved series of histidines and cysteines over a region encompassing several hundred amino acid residues (see Chevalier *et al.* (2001), *Nucleic Acids Res.* 29(18): 3757-3774). In the case of the NHN family, the members are defined by motifs containing two pairs of conserved histidines surrounded by asparagine residues (see Chevalier *et al.* (2001), *Nucleic Acids Res.* 29(18): 3757-3774). The four families of meganucleases are widely separated from one another with respect to conserved structural elements and, consequently, DNA recognition sequence specificity and catalytic activity.

[0007] Natural meganucleases, primarily from the LAGLIDADG family, have been used to effectively promote site-specific genome modification in plants, yeast, *Drosophila*, mammalian cells and mice, but this approach has been limited to the modification of either homologous genes that conserve the meganuclease recognition sequence (Monnat *et al.* (1999), *Biochem. Biophys. Res. Commun.* 255: 88-93) or to pre-engineered genomes into which a recognition sequence has been introduced (Rouet *et al.* (1994), *Mol. Cell. Biol.* 14: 8096-106; Chilton *et al.* (2003), *Plant Physiol.* 133: 956-65; Puchta *et al.* (1996), *Proc. Natl. Acad. Sci. USA* 93: 5055-60; Rong *et al.* (2002), *Genes Dev.* 16: 1568-81; Gouble *et al.* (2006), *J. Gene Med.* 8(5):616-622).

[0008] Systematic implementation of nuclease-stimulated gene modification requires the use of engineered enzymes with customized specificities to target DNA breaks to existing sites in a genome and, therefore, there has been great interest in adapting meganucleases to promote gene modifications at medically or biotechnologically relevant sites (Porteus *et al.* (2005), *Nat. Biotechnol.* 23: 967-73;

Sussman *et al.* (2004), *J. Mol. Biol.* 342: 31-41; Epinat *et al.* (2003), *Nucleic Acids Res.* 31: 2952-62).

[0009] The meganuclease I-CreI from *Chlamydomonas reinhardtii* is a member of the LAGLIDADG family which recognizes and cuts a 22 base-pair recognition sequence in the chloroplast chromosome, and which presents an attractive target for meganuclease redesign. The wild-type enzyme is a homodimer in which each monomer makes direct contacts with 9 base pairs in the full-length recognition sequence. Genetic selection techniques have been used to identify mutations in I-CreI that alter base preference at a single position in this recognition sequence (Sussman *et al.* (2004), *J. Mol. Biol.* 342: 31-41; Chames *et al.* (2005), *Nucleic Acids Res.* 33: e178; Seligman *et al.* (2002), *Nucleic Acids Res.* 30: 3870-9) or, more recently, at three positions in the recognition sequence (Arnould *et al.* (2006), *J. Mol. Biol.* 355: 443-58). The I-CreI protein-DNA interface contains nine amino acids that contact the DNA bases directly and at least an additional five positions that can form potential contacts in modified interfaces. The size of this interface imposes a combinatorial complexity that is unlikely to be sampled adequately in sequence libraries constructed to select for enzymes with drastically altered cleavage sites.

[0010] There remains a need for nucleases that will facilitate precise modification of a genome. In addition, there remains a need for techniques for generating nucleases with pre-determined, rationally-designed recognition sequences that will allow manipulation of genetic sequences at specific genetic loci and for techniques utilizing such nucleases to genetically engineer organisms with precise sequence modifications.

SUMMARY OF THE INVENTION

[0011] The present invention is based, in part, upon the identification and characterization of specific amino acid residues in the LAGLIDADG family of meganucleases that make contacts with DNA bases and the DNA backbone when the meganucleases associate with a double-stranded DNA recognition sequence, and thereby affect the specificity and activity of the enzymes. This discovery has been used, as described in detail below, to identify amino acid substitutions which can alter the recognition sequence specificity and/or DNA-binding affinity of the meganucleases, and

to rationally design and develop meganucleases that can recognize a desired DNA sequence that naturally-occurring meganucleases do not recognize. The invention also provides methods that use such meganucleases to produce recombinant nucleic acids and organisms by utilizing the meganucleases to cause recombination of a desired genetic sequence at a limited number of loci within the genome of the organism, for gene therapy, for treatment of pathogenic infections, and for *in vitro* applications in diagnostics and research.

[0012] Thus, in some embodiments, the invention provides recombinant meganucleases having altered specificity for at least one recognition sequence half-site relative to a wild-type I-CreI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 2-153 of the wild-type I-CreI meganuclease of SEQ ID NO: 1, but in which the recombinant meganuclease has specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-CreI meganuclease recognition sequence selected from SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5, and in which the recombinant meganuclease includes at least one modification listed in Table 1 which is not an excluded modification found in the prior art.

[0013] In other embodiments, the invention provides recombinant meganucleases having altered specificity for at least one recognition sequence half-site relative to a wild-type I-MsoI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6, but in which the recombinant meganuclease has specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-MsoI meganuclease recognition sequence selected from SEQ ID NO: 7 and SEQ ID NO: 8, and in which the recombinant meganuclease includes at least one modification listed in Table 2 which is not an excluded modification found in the prior art.

[0014] In other embodiments, the invention provides recombinant meganucleases having altered specificity for a recognition sequence relative to a wild-type I-SceI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 3-186 of the I-SceI meganuclease of SEQ ID NO: 9, but in which the recombinant meganuclease has specificity for a recognition sequence which

differs by at least one base pair from an I-SceI meganuclease recognition sequence of SEQ ID NO: 10 and SEQ ID NO: 11, and in which the recombinant meganuclease includes at least one modification listed in Table 3 which is not an excluded modification found in the prior art.

[0015] In other embodiments, the invention provides recombinant meganucleases having altered specificity for at least one recognition sequence half-site relative to a wild-type I-CeuI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12, but in which the recombinant meganuclease has specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-CeuI meganuclease recognition sequence selected from SEQ ID NO: 13 and SEQ ID NO: 14, and in which the recombinant meganuclease includes at least one modification listed in Table 4 which is not an excluded modification found in the prior art.

[0016] The meganucleases of the invention can include one, two, three or more of the modifications which have been disclosed herein in order to affect the sequence specificity of the recombinant meganucleases at one, two, three or more positions within the recognition sequence. The meganucleases can include only the novel modifications disclosed herein, or can include the novel modifications disclosed herein in combination with modifications found in the prior art. Specifically excluded, however, are recombinant meganucleases comprising only the modifications of the prior art.

[0017] In another aspect, the invention provides for recombinant meganucleases with altered binding affinity for double-stranded DNA which is not sequence-specific. This is accomplished by modifications of the meganuclease residues which make contacts with the backbone of the double-stranded DNA recognition sequence. The modifications can increase or decrease the binding affinity and, consequently, can increase or decrease the overall activity of the enzyme. Moreover, increases/decreases in binding and activity have been found to cause decreases/increases in sequence specificity. Thus, the invention provides a means for altering sequence specificity generally by altering DNA-binding affinity.

[0018] Thus, in some embodiments, the invention provides for recombinant meganucleases having altered binding affinity for double-stranded DNA relative to a

wild-type I-CreI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1, and in which the DNA-binding affinity has been either (1) increased by at least one modification corresponding to a substitution selected from (a) substitution of E80, D137, I81, L112, P29, V64 or Y66 with H, N, Q, S, T, K or R, or (b) substitution of T46, T140 or T143 with K or R; or, conversely, (2) decreased by at least one modification corresponding to a substitution selected from (a) substitution of K34, K48, R51, K82, K116 or K139 with H, N, Q, S, T, D or E, or (b) substitution of I81, L112, P29, V64, Y66, T46, T140 or T143 with D or E.

[0019] In other embodiments, the invention provides for recombinant meganucleases having altered binding affinity for double-stranded DNA relative to a wild-type I-MsoI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6, and in which the DNA-binding affinity has been either (1) increased by at least one modification corresponding to a substitution selected from (a) substitution of E147, I85, G86 or Y118 with H, N, Q, S, T, K or R, or (b) substitution of Q41, N70, S87, T88, H89, Q122, Q139, S150 or N152 with K or R; or, conversely, (2) decreased by at least one modification corresponding to a substitution selected from (a) substitution of K36, R51, K123, K143 or R144 with H, N, Q, S, T, D or E, or (b) substitution of I85, G86, Y118, Q41, N70, S87, T88, H89, Q122, Q139, S150 or N152 with D or E.

[0020] In other embodiments, the invention provides for recombinant meganucleases having altered binding affinity for double-stranded DNA relative to a wild-type I-SceI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 3-186 of the I-SceI meganuclease of SEQ ID NO: 9, and in which the DNA-binding affinity has been either (1) increased by at least one modification corresponding to a substitution selected from (a) substitution of D201, L19, L80, L92, Y151, Y188, I191, Y199 or Y222 with H, N, Q, S, T, K or R, or (b) substitution of N15, N17, S81, H84, N94, N120, T156, N157, S159, N163, Q165, S166, N194 or S202 with K or R; or, conversely, (2) decreased by at least one modification corresponding to a substitution selected from (a) substitution of K20, K23, K63, K122, K148, K153, K190, K193, K195 or K223 with H, N, Q, S, T, D or E, or (b) substitution

of L19, L80, L92, Y151, Y188, I191, Y199, Y222, N15, N17, S81, H84, N94, N120, T156, N157, S159, N163, Q165, S166, N194 or S202 with D or E.

[0021] In other embodiments, the invention provides for recombinant meganucleases having altered binding affinity for double-stranded DNA relative to a wild-type I-CeuI meganuclease, in which the meganuclease includes a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12, and in which the DNA-binding affinity has been either (1) increased by at least one modification corresponding to a substitution selected from (a) substitution of D25 or D128 with H, N, Q, S, T, K or R, or (b) substitution of S68, N70, H94, S117, N120, N129 or H172 with K or R; or, conversely, (2) decreased by at least one modification corresponding to a substitution selected from (a) substitution of K21, K28, K31, R112, R114 or R130 with H, N, Q, S, T, D or E, or (b) substitution of S68, N70, H94, S117, N120, N129 or H172 with D or E.

[0022] The meganucleases of the invention can include one, two, three or more of the modifications of backbone contact residues which have been disclosed herein in order to affect DNA-binding affinity. In addition, these modifications affecting DNA-binding affinity can be combined with one or more of the novel modifications of the base contact residues described above which alter the sequence specificity of the recombinant meganucleases at specific positions within the recognition sequence, or with the prior art modifications described above, or with a combination of the novel modifications and prior art modifications. In particular, by combining backbone contact modifications and base contact modifications, recombinant meganucleases can be rationally-designed with desired specificity and activity. For example, increases in DNA-binding affinity can be designed which may offset losses in affinity resulting from designed changes to base contact residues, or decreases in affinity can be designed which may also decrease sequence specificity and broaden the set of recognition sequences for an enzyme.

[0023] In another aspect, the invention provides for rationally-designed meganuclease monomers with altered affinity for homo- or heterodimer formation. The affinity for dimer formation can be measured with the same monomer (*i.e.*, homodimer formation) or with a different monomer (*i.e.*, heterodimer formation) such as a reference wild-type meganuclease. These recombinant meganucleases have modifications to the

amino acid residues which are present at the protein-protein interface between monomers in a meganuclease dimer. The modifications can be used to promote heterodimer formation and create meganucleases with non-palindromic recognition sequences.

[0024] Thus, in some embodiments, the invention provides recombinant meganuclease monomers having altered affinity for dimer formation with a reference meganuclease monomer, in which the recombinant monomer includes a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (a) substitution of K7, K57 or K96 with D or E, or (b) substitution of E8 or E61 with K or R. Based upon such recombinant monomers, the invention also provides recombinant meganuclease heterodimers including (1) a first polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (a) substitution of K7, K57 or K96 with D or E, and (2) a second polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (b) substitution of E8 or E61 with K or R.

[0025] In other embodiments, the invention provides recombinant meganuclease monomers having altered affinity for dimer formation with a reference meganuclease monomer, in which the recombinant monomer includes a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (a) substitution of R302 with D or E, or (b) substitution of D20, E11 or Q64 with K or R. Based upon such recombinant monomers, the invention also provides recombinant meganuclease heterodimers including (1) a first polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (a) substitution of R302 with D or E, and (2) a second polypeptide having at least 85%

sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (b) substitution of D20, E11 or Q64 with K or R.

[0026] In other embodiments, the invention provides recombinant meganuclease monomers having altered affinity for dimer formation with a reference meganuclease monomer, in which the recombinant monomer includes a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (a) substitution of R93 with D or E, or (b) substitution of E152 with K or R. Based upon such recombinant monomers, the invention also provides recombinant meganuclease heterodimers including (1) a first polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (a) substitution of R93 with D or E, and (2) a second polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12, but in which affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from (b) substitution of E152 with K or R.

[0027] The recombinant meganuclease monomers or heterodimers with altered affinity for dimer formation can also include one, two, three or more of the modifications of base contact residues described above; one, two, three or more of the modifications of backbone contact residues described above; or combinations of both. Thus, for example, the base contacts of a monomer can be modified to alter sequence specificity, the backbone contacts of a monomer can be modified to alter DNA-binding affinity, and the protein-protein interface can be modified to affect dimer formation. Such a recombinant monomer can be combined with a similarly modified monomer to produce a rationally-designed meganuclease heterodimer with desired sequence specificity and activity.

[0028] In another aspect, the invention provides for various methods of use for the rationally-designed meganucleases described and enabled herein. These methods include producing genetically-modified cells and organisms, treating diseases by gene therapy,

treating pathogen infections, and using the recombinant meganucleases for *in vitro* applications for diagnostics and research.

[0029] Thus, in one aspect, the invention provides methods for producing a genetically-modified eukaryotic cell including an exogenous sequence of interest inserted in a chromosome, by transfecting the cell with (i) a first nucleic acid sequence encoding a meganuclease of the invention, and (ii) a second nucleic acid sequence including said sequence of interest, wherein the meganuclease produces a cleavage site in the chromosome and the sequence of interest is inserted into the chromosome at the cleavage site either by homologous recombination or non-homologous end-joining.

[0030] Alternatively, in another aspect, the invention provides methods for producing a genetically-modified eukaryotic cell including an exogenous sequence of interest inserted in a chromosome, by introducing a meganuclease protein of the invention into the cell, and transfecting the cell with a nucleic acid including the sequence of interest, wherein the meganuclease produces a cleavage site in the chromosome and the sequence of interest is inserted into the chromosome at the cleavage site either by homologous recombination or non-homologous end-joining.

[0031] In another aspect, the invention provides methods for producing a genetically-modified eukaryotic cell by disrupting a target sequence in a chromosome, by transfecting the cell with a nucleic acid encoding a meganuclease of the invention, wherein the meganuclease produces a cleavage site in the chromosome and the target sequence is disrupted by non-homologous end-joining at the cleavage site.

[0032] In another aspect, the invention provides methods of producing a genetically-modified organism by producing a genetically-modified eukaryotic cell according to the methods described above, and growing the genetically-modified eukaryotic cell to produce the genetically-modified organism. In these embodiments, the eukaryotic cell can be selected from a gamete, a zygote, a blastocyst cell, an embryonic stem cell, and a protoplast cell.

[0033] In another aspect, the invention provides methods for treating a disease by gene therapy in a eukaryote, by transfecting at least one cell of the eukaryote with one or more nucleic acids including (i) a first nucleic acid sequence encoding a meganuclease of the invention, and (ii) a second nucleic acid sequence including a sequence of interest,

wherein the meganuclease produces a cleavage site in the chromosome and the sequence of interest is inserted into the chromosome by homologous recombination or non-homologous end-joining, and insertion of the sequence of interest provides gene therapy for the disease.

[0034] Alternatively, in another aspect, the invention provides methods for treating a disease by gene therapy in a eukaryote, by introducing a meganuclease protein of the invention into at least one cell of the eukaryote, and transfecting the cell with a nucleic acid including a sequence of interest, wherein the meganuclease produces a cleavage site in the chromosome and the sequence of interest is inserted into the chromosome at the cleavage site by homologous recombination or non-homologous end-joining, and insertion of the sequence of interest provides gene therapy for the disease.

[0035] In another aspect, the invention provides methods for treating a disease by gene therapy in a eukaryote by disrupting a target sequence in a chromosome of the eukaryotic, by transfecting at least one cell of the eukaryote with a nucleic acid encoding a meganuclease of the invention, wherein the meganuclease produces a cleavage site in the chromosome and the target sequence is disrupted by non-homologous end-joining at the cleavage site, wherein disruption of the target sequence provides the gene therapy for the disease.

[0036] In another aspect, the invention provides methods for treating a viral or prokaryotic pathogen infection in a eukaryotic host by disrupting a target sequence in a genome of the pathogen, by transfecting at least one infected cell of the host with a nucleic acid encoding a meganuclease of the invention, wherein the meganuclease produces a cleavage site in the genome and the target sequence is disrupted by either (1) non-homologous end-joining at the cleavage site or (2) by homologous recombination with a second nucleic acid, and wherein disruption of the target sequence provides treatment for the infection.

[0037] More generally, in another aspect, the invention provides methods for rationally-designing recombinant meganucleases having altered specificity for at least one base position of a recognition sequence, by (1) determining at least a portion of a three-dimensional structure of a reference meganuclease-DNA complex; (2) identifying amino acid residues forming a base contact surface at the base position; (3) determining a

distance between a β -carbon of at least a first residue of the contact surface and at least a first base at the base position; and (4) identifying an amino acid substitution to promote the desired change by either (a) for a first residue which is $< 6 \text{ \AA}$ from the first base, selecting a substitution from Group 1 and/or Group 2 which is a member of an appropriate one of Group G, Group C, Group T or Group A; or (b) for a first residue which is $> 6 \text{ \AA}$ from said first base, selecting a substitution from Group 2 and/or Group 3 which is a member of an appropriate one of Group G, Group C, Group T or Group A, where each of the Groups is defined herein. This method may be repeated for additional contact residues for the same base, and for contact residues for the other base at the same position, as well as for additional positions.

[0038] In addition, in another general aspect, the invention provides methods for rationally-designing a recombinant meganuclease having increased DNA-binding affinity, by (1) determining at least a portion of a three-dimensional structure of a reference meganuclease-DNA complex; (2) identifying amino acid contact residues forming a backbone contact surface; and (3) identifying an amino acid substitution to increase the DNA-binding affinity by (a) for a contact residue having a negatively-charged or hydrophobic side chain, selecting a substitution having an uncharged/polar or positively-charged side chain; or (b) for a contact residue having an uncharged/polar side chain, selecting a substitution having a positively-charged side chain. Conversely, the invention also provides methods for rationally-designing a recombinant meganuclease having decreased DNA-binding affinity, by (1) determining at least a portion of a three-dimensional structure of a reference meganuclease-DNA complex; (2) identifying amino acid contact residues forming a backbone contact surface; (3) identifying an amino acid substitution to decrease the DNA-binding affinity by (a) for a contact residue having a positively-charged side chain, selecting a substitution having an uncharged/polar or negatively-charged side chain; or (b) for a contact residue having an hydrophobic or uncharged/polar side chain, selecting a substitution having a negatively-charged side chain.

[0039] These and other aspects and embodiments of the invention will be apparent to one of ordinary skill in the art based upon the following detailed description of the invention.

BRIEF DESCRIPTION OF THE FIGURES

[0040] Figure 1(A) illustrates the interactions between the I-CreI homodimer and its naturally-occurring double-stranded recognition sequence, based upon crystallographic data. This schematic representation depicts the recognition sequence (SEQ ID NO: 2 and SEQ ID NO: 3), shown as unwound for illustration purposes only, bound by the homodimer, shown as two ovals. The bases of each DNA half-site are numbered -1 through -9, and the amino acid residues of I-CreI which form the recognition surface are indicated by one-letter amino acid designations and numbers indicating residue position. Solid black lines: hydrogen bonds to DNA bases. Dashed lines: amino acid positions that form additional contacts in enzyme designs but do not contact the DNA in the wild-type complex. Arrows: residues that interact with the DNA backbone and influence cleavage activity.

[0041] Figure 1(B) illustrates the wild-type contacts between the A-T base pair at position -4 of the cleavage half-site on the right side of Figure 1(A). Specifically, the residue Q26 is shown to interact with the A base. Residue I77 is in proximity to the base pair but not specifically interacting.

[0042] Figure 1(C) illustrates the interactions between a rationally-designed variant of the I-CreI meganuclease in which residue I77 has been modified to E77. As a result of this change, a G-C base pair is preferred at position -4. The interaction between Q26 and the G base is mediated by a water molecule, as has been observed crystallographically for the cleavage half-site on the left side of Figure 1(A).

[0043] Figure 1(D) illustrates the interactions between a rationally-designed variant of the I-CreI meganuclease in which residue Q26 has been modified to E26 and residue I77 has been modified to R77. As a result of this change, a C-G base pair is preferred at position -4.

[0044] Figure 1(E) illustrates the interactions between a rationally-designed variant of the I-CreI meganuclease in which residue Q26 has been modified to A26 and residue I77 has been modified to Q77. As a result of this change, a T-A base pair is preferred at position -4.

[0045] Figure 2(A) shows a comparison of one recognition sequence for each of the wild type I-CreI meganuclease (WT) and 11 rationally-designed meganuclease heterodimers of the invention. Bases that are conserved relative to the WT recognition sequence are shaded. The 9 bp half-sites are bolded. WT: wild-type (SEQ ID NO: 4); CF: Δ F508 allele of the human CFTR gene responsible for most cases of cystic fibrosis (SEQ ID NO: 25); MYD: the human DM kinase gene associated with myotonic dystrophy (SEQ ID NO: 27); CCR: the human CCR5 gene (a major HIV co-receptor) (SEQ ID NO: 26); ACH: the human FGFR3 gene correlated with achondroplasia (SEQ ID NO: 23); TAT: the HIV-1 TAT/REV gene (SEQ ID NO: 15); HSV: the HSV-1 UL36 gene (SEQ ID NO: 28); LAM: the bacteriophage λ p05 gene (SEQ ID NO: 22); POX: the Variola (smallpox) virus gp009 gene (SEQ ID NO: 30); URA: the *Saccharomyces cerevisiae* URA3 gene (SEQ ID NO: 36); GLA: the *Arabidopsis thaliana* GL2 gene (SEQ ID NO: 32); BRP: the *Arabidopsis thaliana* BP-1 gene (SEQ ID NO: 33).

[0046] Figure 2(B) illustrates the results of incubation of each of wild-type I-CreI (WT) and 11 rationally-designed meganuclease heterodimers with plasmids harboring the recognition sites for all 12 enzymes for 6 hours at 37°C. Percent cleavage is indicated in each box.

[0047] Figure 3 illustrates cleavage patterns of wild-type and rationally-designed I-CreI homodimers. (A) wild type I-CreI. (B) I-CreI K116D. (C-L) rationally-designed meganucleases of the invention. Enzymes were incubated with a set of plasmids harboring palindromes of the intended cleavage half-site the 27 corresponding single-base pair variations. Bar graphs show fractional cleavage (F) in 4 hours at 37°C. Black bars: expected cleavage patterns based on Table 1. Gray bars: DNA sites that deviate from expected cleavage patterns. White circles indicate bases in the intended recognition site. Also shown are cleavage time-courses over two hours. The open circle time-course plots in C and L correspond to cleavage by the CCR1 and BRP2 enzymes lacking the E80Q mutation. The cleavage sites correspond to the 5' (left column) and 3' (right column) half-sites for the heterodimeric enzymes described in Fig. 2(A).

DETAILED DESCRIPTION OF THE INVENTION

1.1 Introduction

[0048] The present invention is based, in part, upon the identification and characterization of specific amino acids in the LAGLIDADG family of meganucleases that make specific contacts with DNA bases and non-specific contacts with the DNA backbone when the meganucleases associate with a double-stranded DNA recognition sequence, and which thereby affect the recognition sequence specificity and DNA-binding affinity of the enzymes. This discovery has been used, as described in detail below, to identify amino acid substitutions in the meganucleases that can alter the specificity and/or affinity of the enzymes, and to rationally design and develop meganucleases that can recognize a desired DNA sequence that naturally-occurring meganucleases do not recognize, and/or that have increased or decreased specificity and/or affinity relative to the naturally-occurring meganucleases. Furthermore, because DNA-binding affinity affects enzyme activity as well as sequence-specificity, the invention provides rationally-designed meganucleases with altered activity relative to naturally-occurring meganucleases. In addition, the invention provides rationally-designed meganucleases in which residues at the interface between the monomers associated to form a dimer have been modified in order to promote heterodimer formation. Finally, the invention provides uses for the rationally-designed meganucleases in the production of recombinant cells and organisms, as well as in gene therapy, anti-pathogen, anti-cancer, and *in vitro* applications, as disclosed herein.

[0049] As a general matter, the invention provides methods for generating rationally-designed LAGLIDADG meganucleases containing altered amino acid residues at sites within the meganuclease that are responsible for (1) sequence-specific binding to individual bases in the double-stranded DNA recognition sequence, or (2) non-sequence-specific binding to the phosphodiester backbone of a double-stranded DNA molecule. Because enzyme activity is correlated to DNA-binding affinity, however, altering the amino acids involved in binding to the DNA recognition sequence can alter not only the specificity of the meganuclease through specific base pair interactions, but also the activity of the meganuclease by increasing or decreasing overall binding affinity for the

double-stranded DNA. Similarly, altering the amino acids involved in binding to the DNA backbone can alter not only the activity of the enzyme, but also the degree of specificity or degeneracy of binding to the recognition sequence by increasing or decreasing overall binding affinity for the double-stranded DNA.

[0050] As described in detail below, the methods of rationally-designing meganucleases include the identification of the amino acids responsible for DNA recognition/binding, and the application of a series of rules for selecting appropriate amino acid changes. With respect to meganuclease sequence specificity, the rules include both steric considerations relating to the distances in a meganuclease-DNA complex between the amino acid side chains of the meganuclease and the bases in the sense and anti-sense strands of the DNA, and considerations relating to the non-covalent chemical interactions between functional groups of the amino acid side chains and the desired DNA base at the relevant position.

[0051] Finally, a majority of natural meganucleases that bind DNA as homodimers recognize pseudo- or completely palindromic recognition sequences. Because lengthy palindromes are expected to be rare, the likelihood of encountering a palindromic sequence at a genomic site of interest is exceedingly low. Consequently, if these enzymes are to be redesigned to recognize genomic sites of interest, it is necessary to design two enzyme monomers recognizing different half-sites that can heterodimerize to cleave the non-palindromic hybrid recognition sequence. Therefore, in some aspects, the invention provides rationally-designed meganucleases in which monomers differing by at least one amino acid position are dimerized to form heterodimers. In some cases, both monomers are rationally-designed to form a heterodimer which recognizes a non-palindromic recognition sequence. A mixture of two different monomers can result in up to three active forms of meganuclease dimer: the two homodimers and the heterodimer. In addition or alternatively, in some cases, amino acid residues are altered at the interfaces at which monomers can interact to form dimers, in order to increase or decrease the likelihood of formation of homodimers or heterodimers.

[0052] Thus, in one aspect, the invention provide methods for rationally designing LAGLIDADG meganucleases containing amino acid changes that alter the specificity and/or activity of the enzymes. In another aspect, the invention provides the rationally-

designed meganucleases resulting from these methods. In another aspect, the invention provides methods that use such rationally-designed meganucleases to produce recombinant nucleic acids and organisms in which a desired DNA sequence or genetic locus within the genome of an organism is modified by the insertion, deletion, substitution or other manipulation of DNA sequences. In another aspect, the invention provides methods for reducing the survival of pathogens or cancer cells using rationally-designed meganucleases which have pathogen-specific or cancer-specific recognition sequences.

1.2 References and Definitions

[0053] The patent and scientific literature referred to herein establishes knowledge that is available to those of skill in the art. The issued U.S. patents, allowed applications, published foreign applications, and references, including GenBank database sequences, that are cited herein are hereby incorporated by reference to the same extent as if each was specifically and individually indicated to be incorporated by reference.

[0054] As used herein, the term "meganuclease" refers to an endonuclease that binds double-stranded DNA at a recognition sequence that is greater than 12 base pairs. Naturally-occurring meganucleases can be monomeric (*e.g.*, I-SceI) or dimeric (*e.g.*, I-CreI). The term meganuclease, as used herein, can be used to refer to monomeric meganucleases, dimeric meganucleases, or to the monomers which associate to form a dimeric meganuclease. The term "homing endonuclease" is synonymous with the term "meganuclease."

[0055] As used herein, the term "LAGLIDADG meganuclease" refers either to meganucleases including a single LAGLIDADG motif, which are naturally dimeric, or to meganucleases including two LAGLIDADG motifs, which are naturally monomeric. The term "mono-LAGLIDADG meganuclease" is used herein to refer to meganucleases including a single LAGLIDADG motif, and the term "di-LAGLIDADG meganuclease" is used herein to refer to meganucleases including two LAGLIDADG motifs, when it is necessary to distinguish between the two. Each of the two structural domains of a di-LAGLIDADG meganuclease which includes a LAGLIDADG motif can be referred to as a LAGLIDADG subunit.

[0056] As used herein, the term "rationally-designed" means non-naturally occurring and/or genetically engineered. The rationally-designed meganucleases of the invention differ from wild-type or naturally-occurring meganucleases in their amino acid sequence or primary structure, and may also differ in their secondary, tertiary or quaternary structure. In addition, the rationally-designed meganucleases of the invention also differ from wild-type or naturally-occurring meganucleases in recognition sequence-specificity and/or activity.

[0057] As used herein, with respect to a protein, the term "recombinant" means having an altered amino acid sequence as a result of the application of genetic engineering techniques to nucleic acids which encode the protein, and cells or organisms which express the protein. With respect to a nucleic acid, the term "recombinant" means having an altered nucleic acid sequence as a result of the application of genetic engineering techniques. Genetic engineering techniques include, but are not limited to, PCR and DNA cloning technologies; transfection, transformation and other gene transfer technologies; homologous recombination; site-directed mutagenesis; and gene fusion. In accordance with this definition, a protein having an amino acid sequence identical to a naturally-occurring protein, but produced by cloning and expression in a heterologous host, is not considered recombinant.

[0058] As used herein with respect to recombinant proteins, the term "modification" means any insertion, deletion or substitution of an amino acid residue in the recombinant sequence relative to a reference sequence (*e.g.*, a wild-type).

[0059] As used herein, the term "genetically-modified" refers to a cell or organism in which, or in an ancestor of which, a genomic DNA sequence has been deliberately modified by recombinant technology. As used herein, the term "genetically-modified" encompasses the term "transgenic."

[0060] As used herein, the term "wild-type" refers to any naturally-occurring form of a meganuclease. The term "wild-type" is not intended to mean the most common allelic variant of the enzyme in nature but, rather, any allelic variant found in nature. Wild-type meganucleases are distinguished from recombinant or non-naturally-occurring meganucleases.

[0061] As used herein, the term "recognition sequence half-site" or simply "half site" means a nucleic acid sequence in a double-stranded DNA molecule which is recognized by a monomer of a mono-LAGLIDADG meganuclease or by one LAGLIDADG subunit of a di-LAGLIDADG meganuclease.

[0062] As used herein, the term "recognition sequence" refers to a pair of half-sites which is bound and cleaved by either a mono-LAGLIDADG meganuclease dimer or a di-LAGLIDADG meganuclease monomer. The two half-sites may or may not be separated by base pairs that are not specifically recognized by the enzyme. In the cases of I-CreI, I-MsoI and I-CeuI, the recognition sequence half-site of each monomer spans 9 base pairs, and the two half-sites are separated by four base pairs which are not recognized specifically but which constitute the actual cleavage site (which has a 4 base pair overhang). Thus, the combined recognition sequences of the I-CreI, I-MsoI and I-CeuI meganuclease dimers normally span 22 base pairs, including two 9 base pair half-sites flanking a 4 base pair cleavage site. The base pairs of each half-site are designated -9 through -1, with the -9 position being most distal from the cleavage site and the -1 position being adjacent to the 4 central base pairs, which are designated N₁-N₄. The strand of each half-site which is oriented 5' to 3' in the direction from -9 to -1 (*i.e.*, towards the cleavage site), is designated the "sense" strand and the opposite strand is designated the "antisense strand", although neither strand may encode protein. Thus, the "sense" strand of one half-site is the antisense strand of the other half-site. See, for example, Figure 1(A). In the case of the I-SceI meganuclease, which is a di-LAGLIDADG meganuclease monomer, the recognition sequence is an approximately 18 bp non-palindromic sequence, and there are no central base pairs which are not specifically recognized. By convention, one of the two strands is referred to as the "sense" strand and the other the "antisense" strand, although neither strand may encode protein.

[0063] As used herein, the term "specificity" means the ability of a meganuclease to recognize and cleave double-stranded DNA molecules only at a particular sequence of base pairs referred to as the recognition sequence, or only at a particular set of recognition sequences. The set of recognition sequences will share certain conserved positions or sequence motifs, but may be degenerate at one or more positions. A highly-specific

meganuclease is capable of cleaving only one or a very few recognition sequences. Specificity can be determined in a cleavage assay as described in Example 1. As used herein, a meganuclease has "altered" specificity if it binds to and cleaves a recognition sequence which is not bound to and cleaved by a reference meganuclease (*e.g.*, a wild-type) or if the rate of cleavage of a recognition sequence is increased or decreased by a statistically significant ($p < 0.05$) amount relative to a reference meganuclease.

[0064] As used herein, the term "degeneracy" means the opposite of "specificity." A highly-degenerate meganuclease is capable of cleaving a large number of divergent recognition sequences. A meganuclease can have sequence degeneracy at a single position within a half-site or at multiple, even all, positions within a half-site. Such sequence degeneracy can result from (i) the inability of any amino acid in the DNA-binding domain of a meganuclease to make a specific contact with any base at one or more positions in the recognition sequence, (ii) the ability of one or more amino acids in the DNA-binding domain of a meganuclease to make specific contacts with more than one base at one or more positions in the recognition sequence, and/or (iii) sufficient non-specific DNA binding affinity for activity. A "completely" degenerate position can be occupied by any of the four bases and can be designated with an "N" in a half-site. A "partially" degenerate position can be occupied by two or three of the four bases (*e.g.*, either purine (Pu), either pyrimidine (Py), or not G).

[0065] As used herein with respect to meganucleases, the term "DNA-binding affinity" or "binding affinity" means the tendency of a meganuclease to non-covalently associate with a reference DNA molecule (*e.g.*, a recognition sequence or an arbitrary sequence). Binding affinity is measured by a dissociation constant, K_D (*e.g.*, the K_D of I-CreI for the WT recognition sequence is approximately 0.1 nM). As used herein, a meganuclease has "altered" binding affinity if the K_D of the recombinant meganuclease for a reference recognition sequence is increased or decreased by a statistically significant ($p < 0.05$) amount relative to a reference meganuclease.

[0066] As used herein with respect to meganuclease monomers, the term "affinity for dimer formation" means the tendency of a meganuclease monomer to non-covalently associate with a reference meganuclease monomer. The affinity for dimer formation can be measured with the same monomer (*i.e.*, homodimer formation) or with a different

monomer (*i.e.*, heterodimer formation) such as a reference wild-type meganuclease. Binding affinity is measured by a dissociation constant, K_D . As used herein, a meganuclease has "altered" affinity for dimer formation if the K_D of the recombinant meganuclease monomer for a reference meganuclease monomer is increased or decreased by a statistically significant ($p < 0.05$) amount relative to a reference meganuclease monomer.

[0067] As used herein, the term "palindromic" refers to a recognition sequence consisting of inverted repeats of identical half-sites. In this case, however, the palindromic sequence need not be palindromic with respect to the four central base pairs, which are not contacted by the enzyme. In the case of dimeric meganucleases, palindromic DNA sequences are recognized by homodimers in which the two monomers make contacts with identical half-sites.

[0068] As used herein, the term "pseudo-palindromic" refers to a recognition sequence consisting of inverted repeats of non-identical or imperfectly palindromic half-sites. In this case, the pseudo-palindromic sequence not only need not be palindromic with respect to the four central base pairs, but also can deviate from a palindromic sequence between the two half-sites. Pseudo-palindromic DNA sequences are typical of the natural DNA sites recognized by wild-type homodimeric meganucleases in which two identical enzyme monomers make contacts with different half-sites.

[0069] As used herein, the term "non-palindromic" refers to a recognition sequence composed of two unrelated half-sites of a meganuclease. In this case, the non-palindromic sequence need not be palindromic with respect to either the four central base pairs or the two monomer half-sites. Non-palindromic DNA sequences are recognized by either di-LAGLIDADG meganucleases, highly degenerate mono-LAGLIDADG meganucleases (*e.g.*, I-CeuI) or by heterodimers of mono-LAGLIDADG meganuclease monomers that recognize non-identical half-sites.

[0070] As used herein, the term "activity" refers to the rate at which a meganuclease of the invention cleaves a particular recognition sequence. Such activity is a measurable enzymatic reaction, involving the hydrolysis of phosphodiester bonds of double-stranded DNA. The activity of a meganuclease acting on a particular DNA substrate is affected by the affinity or avidity of the meganuclease for that particular DNA

substrate which is, in turn, affected by both sequence-specific and non-sequence-specific interactions with the DNA.

[0071] As used herein, the term “homologous recombination” refers to the natural, cellular process in which a double-stranded DNA-break is repaired using a homologous DNA sequence as the repair template (see, *e.g.* Cahill *et al.* (2006), *Front. Biosci.* 11:1958-1976). The homologous DNA sequence may be an endogenous chromosomal sequence or an exogenous nucleic acid that was delivered to the cell. Thus, in some embodiments, a rationally-designed meganuclease is used to cleave a recognition sequence within a target sequence and an exogenous nucleic acid with homology to or substantial sequence similarity with the target sequence is delivered into the cell and used as a template for repair by homologous recombination. The DNA sequence of the exogenous nucleic acid, which may differ significantly from the target sequence, is thereby incorporated into the chromosomal sequence. The process of homologous recombination occurs primarily in eukaryotic organisms. The term "homology" is used herein as equivalent to "sequence similarity" and is not intended to require identity by descent or phylogenetic relatedness.

[0072] As used herein, the term “non-homologous end-joining” refers to the natural, cellular process in which a double-stranded DNA-break is repaired by the direct joining of two non-homologous DNA segments (see, *e.g.* Cahill *et al.* (2006), *Front. Biosci.* 11:1958-1976). DNA repair by non-homologous end-joining is error-prone and frequently results in the untemplated addition or deletion of DNA sequences at the site of repair. Thus, in certain embodiments, a rationally-designed meganuclease can be used to produce a double-stranded break at a meganuclease recognition sequence within a target sequence to disrupt a gene (*e.g.*, by introducing base insertions, base deletions, or frameshift mutations) by non-homologous end-joining. In other embodiments, an exogenous nucleic acid lacking homology to or substantial sequence similarity with the target sequence may be captured at the site of a meganuclease-stimulated double-stranded DNA break by non-homologous end-joining (see, *e.g.* Salomon, *et al.* (1998), *EMBO J.* 17:6086-6095). The process of non-homologous end-joining occurs in both eukaryotes and prokaryotes such as bacteria.

[0073] As used herein, the term "sequence of interest" means any nucleic acid sequence, whether it codes for a protein, RNA, or regulatory element (*e.g.*, an enhancer, silencer, or promoter sequence), that can be inserted into a genome or used to replace a genomic DNA sequence using a meganuclease protein. Sequences of interest can have heterologous DNA sequences that allow for tagging a protein or RNA that is expressed from the sequence of interest. For instance, a protein can be tagged with tags including, but not limited to, an epitope (*e.g.*, c-myc, FLAG) or other ligand (*e.g.*, poly-His). Furthermore, a sequence of interest can encode a fusion protein, according to techniques known in the art (see, *e.g.*, Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley 1999). In some embodiments, the sequence of interest is flanked by a DNA sequence that is recognized by the recombinant meganuclease for cleavage. Thus, the flanking sequences are cleaved allowing for proper insertion of the sequence of interest into genomic recognition sequences cleaved by the recombinant meganuclease. In some embodiments, the entire sequence of interest is homologous to or has substantial sequence similarity with the a target sequence in the genome such that homologous recombination effectively replaces the target sequence with the sequence of interest. In other embodiments, the sequence of interest is flanked by DNA sequences with homology to or substantial sequence similarity with the target sequence such that homologous recombination inserts the sequence of interest within the genome at the locus of the target sequence. In some embodiments, the sequence of interest is substantially identical to the target sequence except for mutations or other modifications in the meganuclease recognition sequence such that the meganuclease can not cleave the target sequence after it has been modified by the sequence of interest.

[0074] As used herein with respect to both amino acid sequences and nucleic acid sequences, the terms "percentage similarity" and "sequence similarity" refer to a measure of the degree of similarity of two sequences based upon an alignment of the sequences which maximizes similarity between aligned amino acid residues or nucleotides, and which is a function of the number of identical or similar residues or nucleotides, the number of total residues or nucleotides, and the presence and length of gaps in the sequence alignment. A variety of algorithms and computer programs are available for determining sequence similarity using standard parameters. As used herein, sequence

similarity is measured using the BLASTp program for amino acid sequences and the BLASTn program for nucleic acid sequences, both of which are available through the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/), and are described in, for example, Altschul *et al.* (1990), *J. Mol. Biol.* 215:403-410; Gish and States (1993), *Nature Genet.* 3:266-272; Madden *et al.* (1996), *Meth. Enzymol.* 266:131-141; Altschul *et al.* (1997), *Nucleic Acids Res.* 25:3389-3402; Zhang *et al.* (2000), *J. Comput. Biol.* 7(1-2):203-14.. As used herein, percent similarity of two amino acid sequences is the score based upon the following parameters for the BLASTp algorithm: word size = 3; gap opening penalty = -11; gap extension penalty = -1; and scoring matrix = BLOSUM62. As used herein, percent similarity of two nucleic acid sequences is the score based upon the following parameters for the BLASTn algorithm: word size = 11; gap opening penalty = -5; gap extension penalty = -2; match reward = 1; and mismatch penalty = -3.

[0075] As used herein with respect to modifications of two proteins or amino acid sequences, the term "corresponding to" is used to indicate that a specified modification in the first protein is a substitution of the same amino acid residue as in the modification in the second protein, and that the amino acid position of the modification in the first proteins corresponds to or aligns with the amino acid position of the modification in the second protein when the two proteins are subjected to standard sequence alignments (*e.g.*, using the BLASTp program). Thus, the modification of residue "X" to amino acid "A" in the first protein will correspond to the modification of residue "Y" to amino acid "A" in the second protein if residues X and Y correspond to each other in a sequence alignment, and despite the fact that X and Y may be different numbers.

[0076] As used herein, the recitation of a numerical range for a variable is intended to convey that the invention may be practiced with the variable equal to any of the values within that range. Thus, for a variable which is inherently discrete, the variable can be equal to any integer value within the numerical range, including the end-points of the range. Similarly, for a variable which is inherently continuous, the variable can be equal to any real value within the numerical range, including the end-points of the range. As an example, and without limitation, a variable which is described as having values between 0 and 2 can take the values 0, 1 or 2 if the variable is inherently discrete, and can

take the values 0.0, 0.1, 0.01, 0.001, or any other real values ≥ 0 and ≤ 2 if the variable is inherently continuous.

[0077] As used herein, unless specifically indicated otherwise, the word "or" is used in the inclusive sense of "and/or" and not the exclusive sense of "either/or."

2.1 Rationally-Designed Meganucleases with Altered Sequence-Specificity

[0078] In one aspect of the invention, methods for rationally designing recombinant LAGLIDADG family meganucleases are provided. In this aspect, recombinant meganucleases are rationally-designed by first predicting amino acid substitutions that can alter base preference at each position in the half-site. These substitutions can be experimentally validated individually or in combinations to produce meganucleases with the desired cleavage specificity.

[0079] In accordance with the invention, amino acid substitutions that can cause a desired change in base preference are predicted by determining the amino acid side chains of a reference meganuclease (*e.g.*, a wild-type meganuclease, or a non-naturally-occurring reference meganuclease) that are able to participate in making contacts with the nucleic acid bases of the meganuclease's DNA recognition sequence and the DNA phosphodiester backbone, and the spatial and chemical nature of those contacts. These amino acids include but are not limited to side chains involved in contacting the reference DNA half-site. Generally, this determination requires having knowledge of the structure of the complex between the meganuclease and its double-stranded DNA recognition sequence, or knowledge of the structure of a highly similar complex (*e.g.*, between the same meganuclease and an alternative DNA recognition sequence, or between an allelic or phylogenetic variant of the meganuclease and its DNA recognition sequence).

[0080] Three-dimensional structures, as described by atomic coordinates data, of a polypeptide or complex of two or more polypeptides can be obtained in several ways. For example, protein structure determinations can be made using techniques including, but not limited to, X-ray crystallography, NMR, and mass spectrometry. Another approach is to analyze databases of existing structural co-ordinates for the meganuclease of interest or a related meganuclease. Such structural data is often available from

databases in the form of three-dimensional coordinates. Often this data is accessible through online databases (e.g., the RCSB Protein Data Bank at www.rcsb.org/pdb).

[0081] Structural information can be obtained experimentally by analyzing the diffraction patterns of, for example, X-rays or electrons, created by regular two- or three-dimensional arrays (e.g., crystals) of proteins or protein complexes. Computational methods are used to transform the diffraction data into three-dimensional atomic coordinates in space. For example, the field of X-ray crystallography has been used to generate three-dimensional structural information on many protein-DNA complexes, including meganucleases (see, e.g., Chevalier *et al.* (2001), *Nucleic Acids Res.* 29(18): 3757-3774).

[0082] Nuclear Magnetic Resonance (NMR) also has been used to determine inter-atomic distances of molecules in solution. Multi-dimensional NMR methods combined with computational methods have succeeded in determining the atomic coordinates of polypeptides of increasing size (see, e.g., Tzakos *et al.* (2006), *Annu. Rev. Biophys. Biomol. Struct.* 35:19-42.).

[0083] Alternatively, computational modeling can be used by applying algorithms based on the known primary structures and, when available, secondary, tertiary and/or quaternary structures of the protein/DNA, as well as the known physiochemical nature of the amino acid side chains, nucleic acid bases, and bond interactions. Such methods can optionally include iterative approaches, or experimentally-derived constraints. An example of such computational software is the CNS program described in Adams *et al.* (1999), *Acta Crystallogr. D. Biol. Crystallogr.* 55 (Pt 1): 181-90. A variety of other computational programs have been developed that predict the spatial arrangement of amino acids in a protein structure and predict the interaction of the amino acid side chains of the protein with various target molecules (see, e.g., U.S. Pat. No. 6,988,041).

[0084] Thus, in some embodiments of the invention, computational models are used to identify specific amino acid residues that specifically interact with DNA nucleic acid bases and/or facilitate non-specific phosphodiester backbone interactions. For instance, computer models of the totality of the potential meganuclease-DNA interaction can be produced using a suitable software program, including, but not limited to, MOLSCRIPT™ 2.0 (Avatar Software AB, Stockholm, Sweden), the graphical display

program O (Jones *et al.* (1991), *Acta Crystallography*, A47: 110), the graphical display program GRASP™ (Nicholls *et al.* (1991), *PROTEINS, Structure, Function and Genetics* 11(4): 281ff), or the graphical display program INSIGHT™ (TSI, Inc., Shoreview, MN). Computer hardware suitable for producing, viewing and manipulating three-dimensional structural representations of protein-DNA complexes are commercially available and well known in the art (*e.g.*, Silicon Graphics Workstation, Silicon Graphics, Inc., Mountainview, CA).

[0085] Specifically, interactions between a meganuclease and its double-stranded DNA recognition sequences can be resolved using methods known in the art. For example, a representation, or model, of the three dimensional structure of a multi-component complex structure, for which a crystal has been produced, can be determined using techniques which include molecular replacement or SIR/MIR (single/multiple isomorphous replacement) (see, *e.g.*, Brunger (1997), *Meth. Enzym.* 276: 558-580; Navaza and Saludjian (1997), *Meth. Enzym.* 276: 581-594; Tong and Rossmann (1997), *Meth. Enzym.* 276: 594-611; and Bentley (1997), *Meth. Enzym.* 276: 611-619) and can be performed using a software program, such as AMoRe/Mosfilm (Navaza (1994), *Acta Cryst.* A50: 157-163; CCP4 (1994), *Acta Cryst.* D50: 760-763) or XPLOR (see, Brünger *et al.* (1992), X-PLOR Version 3.1. A System for X-ray Crystallography and NMR, Yale University Press, New Haven, CT).

[0086] The determination of protein structure and potential meganuclease-DNA interaction allows for rational choices concerning the amino acids that can be changed to affect enzyme activity and specificity. Decisions are based on several factors regarding amino acid side chain interactions with a particular base or DNA phosphodiester backbone. Chemical interactions used to determine appropriate amino acid substitutions include, but are not limited to, van der Waals forces, steric hindrance, ionic bonding, hydrogen bonding, and hydrophobic interactions. Amino acid substitutions can be selected which either favor or disfavor specific interactions of the meganuclease with a particular base in a potential recognition sequence half-site in order to increase or decrease specificity for that sequence and, to some degree, overall binding affinity and activity. In addition, amino acid substitutions can be selected which either increase or decrease binding affinity for the phosphodiester backbone of double-stranded DNA in

order to increase or decrease overall activity and, to some degree, to decrease or increase specificity.

[0087] Thus, in specific embodiments, a three-dimensional structure of a meganuclease-DNA complex is determined and a "contact surface" is defined for each base-pair in a DNA recognition sequence half-site. In some embodiments, the contact surface comprises those amino acids in the enzyme with β -carbons less than 9.0 Å from a major groove hydrogen-bond donor or acceptor on either base in the pair, and with side chains oriented toward the DNA, irrespective of whether the residues make base contacts in the wild-type meganuclease-DNA complex. In other embodiments, residues can be excluded if the residues do not make contact in the wild-type meganuclease-DNA complex, or residues can be included or excluded at the discretion of the designer to alter the number or identity of the residues considered. In one example, as described below, for base positions -2, -7, -8, and -9 of the wild-type I-CreI half-site, the contact surfaces were limited to the amino acid positions that actually interact in the wild-type enzyme-DNA complex. For positions -1, -3, -4, -5, and -6, however, the contact surfaces were defined to contain additional amino acid positions that are not involved in wild-type contacts but which could potentially contact a base if substituted with a different amino acid.

[0088] It should be noted that, although a recognition sequence half-site is typically represented with respect to only one strand of DNA, meganucleases bind in the major groove of double-stranded DNA, and make contact with nucleic acid bases on both strands. In addition, the designations of "sense" and "antisense" strands are completely arbitrary with respect to meganuclease binding and recognition. Sequence specificity at a position can be achieved either through interactions with one member of a base pair, or by a combination of interactions with both members of a base-pair. Thus, for example, in order to favor the presence of an A/T base pair at position X, where the A base is on the "sense" strand and the T base is on the "antisense" strand, residues are selected which are sufficiently close to contact the sense strand at position X and which favor the presence of an A, and/or residues are selected which are sufficiently close to contact the antisense strand at position X and which favor the presence of a T. In accordance with the

invention, a residue is considered sufficiently close if the β -carbon of the residue is within 9Å of the closest atom of the relevant base.

[0089] Thus, for example, an amino acid with a β -carbon within 9Å of the DNA sense strand but greater than 9Å from the antisense strand is considered for potential interactions with only the sense strand. Similarly, an amino acid with a β -carbon within 9Å of the DNA antisense strand but greater than 9Å from the sense strand is considered for potential interactions with only the antisense strand. Amino acids with β -carbons that are within 9Å of both DNA strands are considered for potential interactions with either strand.

[0090] For each contact surface, potential amino acid substitutions are selected based on their predicted ability to interact favorably with one or more of the four DNA bases. The selection process is based upon two primary criteria: (i) the size of the amino acid side chains, which will affect their steric interactions with different nucleic acid bases, and (ii) the chemical nature of the amino acid side chains, which will affect their electrostatic and bonding interactions with the different nucleic acid bases.

[0091] With respect to the size of side chains, amino acids with shorter and/or smaller side chains can be selected if an amino acid β -carbon in a contact surface is <6 Å from a base, and amino acids with longer and/or larger side chains can be selected if an amino acid β -carbon in a contact surface is >6 Å from a base. Amino acids with side chains that are intermediate in size can be selected if an amino acid β -carbon in a contact surface is 5-8 Å from a base.

[0092] The amino acids with relatively shorter and smaller side chains can be assigned to Group 1, including glycine (G), alanine (A), serine (S), threonine (T), cysteine (C), valine (V), leucine (L), isoleucine (I), aspartate (D), asparagine (N) and proline (P). Proline, however, is expected to be used less frequently because of its relative inflexibility. In addition, glycine is expected to be used less frequently because it introduces unwanted flexibility in the peptide backbone and its very small size reduces the likelihood of effective contacts when it replaces a larger residue. On the other hand, glycine can be used in some instances for promoting a degenerate position. The amino acids with side chains of relatively intermediate length and size can be assigned to Group 2, including lysine (K), methionine (M), arginine (R), glutamate (E) and glutamine (Q).

The amino acids with relatively longer and/or larger side chains can be assigned to Group 3, including lysine (K), methionine (M), arginine (R), histidine (H), phenylalanine (F), tyrosine (Y), and tryptophan (W). Tryptophan, however, is expected to be used less frequently because of its relative inflexibility. In addition, the side chain flexibility of lysine, arginine, and methionine allow these amino acids to make base contacts from long or intermediate distances, warranting their inclusion in both Groups 2 and 3. These groups are also shown in tabular form below:

Group 1	Group 2	Group 3
glycine (G)	glutamine (Q)	arginine (R)
alanine (A)	glutamate (E)	histidine (H)
serine (S)	lysine (K)	phenylalanine (F)
threonine (T)	methionine (M)	tyrosine (Y)
cysteine (C)	arginine (R)	tryptophan (W)
valine (V)		lysine (K)
leucine (L)		methionine (M)
isoleucine (I)		
aspartate (D)		
asparagine (N)		
proline (P)		

[0093] With respect to the chemical nature of the side chains, the different amino acids are evaluated for their potential interactions with the different nucleic acid bases (e.g., van der Waals forces, ionic bonding, hydrogen bonding, and hydrophobic interactions) and residues are selected which either favor or disfavor specific interactions of the meganuclease with a particular base at a particular position in the double-stranded DNA recognition sequence half-site. In some instances, it may be desired to create a half-site with one or more complete or partial degenerate positions. In such cases, one

may choose residues which favor the presence of two or more bases, or residues which disfavor one or more bases. For example, partial degenerate base recognition can be achieved by sterically hindering a pyrimidine at a sense or antisense position.

[0094] Recognition of guanine (G) bases is achieved using amino acids with basic side chains that form hydrogen bonds to N7 and O6 of the base. Cytosine (C) specificity is conferred by negatively-charged side chains which interact unfavorably with the major groove electronegative groups present on all bases except C. Thymine (T) recognition is rationally-designed using hydrophobic and van der Waals interactions between hydrophobic side chains and the major groove methyl group on the base. Finally, adenine (A) bases are recognized using the carboxamide side chains Asn and Gln or the hydroxyl side chain of Tyr through a pair of hydrogen bonds to N7 and N6 of the base. Lastly, His can be used to confer specificity for a purine base (A or G) by donating a hydrogen bond to N7. These straightforward rules for DNA recognition can be applied to predict contact surfaces in which one or both of the bases at a particular base-pair position are recognized through a rationally-designed contact.

[0095] Thus, based on their binding interactions with the different nucleic acid bases, and the bases which they favor at a position with which they make contact, each amino acid residue can be assigned to one or more different groups corresponding to the different bases they favor (i.e., G, C, T or A). Thus, Group G includes arginine (R), lysine (K) and histidine (H); Group C includes aspartate (D) and glutamate (E); Group T includes alanine (A), valine (V), leucine (L), isoleucine (I), cysteine (C), threonine (T), methionine (M) and phenylalanine (F); and Group A includes asparagine (N), glutamine (N), tyrosine (Y) and histidine (H). Note that histidine appears in both Group G and Group A; that serine (S) is not included in any group but may be used to favor a degenerate position; and that proline, glycine, and tryptophan are not included in any particular group because of predominant steric considerations. These groups are also shown in tabular form below:

Group G	Group C	Group T	Group A
arginine (R)	aspartate (D)	alanine (A)	asparagine (N)
lysine (K)	glutamate (E)	valine (V)	glutamine (Q)
histidine (H)		leucine (L)	tyrosine (Y)
		isoleucine (I)	histidine (H)
		cysteine (C)	
		threonine (T)	
		methionine (M)	
		phenylalanine (F)	

[0096] Thus, in accordance with the invention, in order to effect a desired change in the recognition sequence half-site of a meganuclease at a given position X, (1) determine at least the relevant portion of the three-dimensional structure of the wild-type or reference meganuclease-DNA complex and the amino acid residue side chains which define the contact surface at position X; (2) determine the distance between the β -carbon of at least one residue comprising the contact surface and at least one base of the base pair at position X; and (3)(a) for a residue which is $< 6 \text{ \AA}$ from the base, select a residue from Group 1 and/or Group 2 which is a member of the appropriate one of Group G, Group C, Group T or Group A to promote the desired change, and/or (b) for a residue which is $> 6 \text{ \AA}$ from the base, select a residue from Group 2 and/or Group 3 which is a member of the appropriate one of Group G, Group C, Group T or Group A to promote the desired change. More than one such residue comprising the contact surface can be selected for analysis and modification and, in some embodiments, each such residue is analyzed and multiple residues are modified. Similarly, the distance between the β -carbon of a residue included in the contact surface and each of the two bases of the base pair at position X

can be determined and, if the residue is within 9Å of both bases, then different substitutions can be made to affect the two bases of the pair (*e.g.*, a residue from Group 1 to affect a proximal base on one strand, or a residue from Group 3 to affect a distal base on the other strand). Alternatively, a combination of residue substitutions capable of interacting with both bases in a pair can affect the specificity (*e.g.*, a residue from the T Group contacting the sense strand combined with a residue from the A Group contacting the antisense strand to select for T/A). Finally, multiple alternative modifications of the residues can be validated either empirically (*e.g.*, by producing the recombinant meganuclease and testing its sequence recognition) or computationally (*e.g.*, by computer modeling of the meganuclease-DNA complex of the modified enzyme) to choose amongst alternatives.

[0097] Once one or more desired amino acid modifications of the wild-type or reference meganuclease are selected, the rationally-designed meganuclease can be produced by recombinant methods and techniques well known in the art. In some embodiments, non-random or site-directed mutagenesis techniques are used to create specific sequence modifications. Non-limiting examples of non-random mutagenesis techniques include overlapping primer PCR (see, *e.g.*, Wang *et al.* (2006), *Nucleic Acids Res.* 34(2): 517–527), site-directed mutagenesis (see, *e.g.*, U.S. Pat. No. 7,041,814), cassette mutagenesis (see, *e.g.*, U.S. Pat. No. 7,041,814), and the manufacturer's protocol for the Altered Sites® II Mutagenesis Systems kit commercially available from Promega Biosciences, Inc. (San Luis Obispo, CA).

[0098] The recognition and cleavage of a specific DNA sequence by a rationally-designed meganuclease can be assayed by any method known by one skilled in the art (see, *e.g.*, U.S. Pat. Pub. No. 2006/0078552). In certain embodiments, the determination of meganuclease cleavage is determined by *in vitro* cleavage assays. Such assays use *in vitro* cleavage of a polynucleotide substrate comprising the intended recognition sequence of the assayed meganuclease and, in certain embodiments, variations of the intended recognition sequence in which one or more bases in one or both half-sites have been changed to a different base. Typically, the polynucleotide substrate is a double-stranded DNA molecule comprising a target site which has been synthesized and cloned into a vector. The polynucleotide substrate can be linear or circular, and typically comprises

only one recognition sequence. The meganuclease is incubated with the polynucleotide substrate under appropriate conditions, and the resulting polynucleotides are analyzed by known methods for identifying cleavage products (*e.g.*, electrophoresis or chromatography). If there is a single recognition sequence in a linear, double-strand DNA substrate, the meganuclease activity is detected by the appearance of two bands (products) and the disappearance of the initial full-length substrate band. In one embodiment, meganuclease activity can be assayed as described in, for example, Wang *et al.* (1997), *Nucleic Acid Res.*, 25: 3767-3776.

[0099] In other embodiments, the cleavage pattern of the meganuclease is determined using *in vivo* cleavage assays (see, *e.g.*, U.S. Pat. Pub. No. 2006/0078552). In particular embodiments, the *in vivo* test is a single-strand annealing recombination test (SSA). This kind of test is known to those of skill in the art (Rudin *et al.* (1989), *Genetics* 122: 519-534; Fishman-Lobell *et al.* (1992), *Science* 258: 480-4).

[0100] As will be apparent to one of skill in the art, additional amino acid substitutions, insertions or deletions can be made to domains of the meganuclease enzymes other than those involved in DNA recognition and binding without complete loss of activity. Substitutions can be conservative substitutions of similar amino acid residues at structurally or functionally constrained positions, or can be non-conservative substitutions at positions which are less structurally or functionally constrained. Such substitutions, insertions and deletions can be identified by one of ordinary skill in the art by routine experimentation without undue effort. Thus, in some embodiments, the recombinant meganucleases of the invention include proteins having anywhere from 85% to 99% sequence similarity (*e.g.*, 85%, 87.5%, 90%, 92.5%, 95%, 97.5%, 99%) to a reference meganuclease sequence. With respect to each of the wild-type I-CreI, I-MsoI, I-SceI and I-CeuI proteins, the most N-terminal and C-terminal sequences are not clearly visible in X-ray crystallography studies, suggesting that these positions are not structurally or functionally constrained. Therefore, these residues can be excluded from calculation of sequence similarity, and the following reference meganuclease sequences can be used: residues 2-153 of SEQ ID NO: 1 for I-CreI, residues 6-160 of SEQ ID NO: 6 for I-MsoI, residues 3-186 of SEQ ID NO: 9 for I-SceI, and residues 5-211 of SEQ ID NO: 12 for I-CeuI.

2.2 LAGLIDADG Family Meganucleases

[0101] The LAGLIDADG meganuclease family is composed of more than 200 members from a diverse phylogenetic group of host organisms. All members of this family have one or two copies of a highly conserved LAGLIDADG motif along with other structural motifs involved in cleavage of specific DNA sequences. Enzymes that have a single copy of the LAGLIDADG motif (*i.e.*, mono-LAGLIDADG meganucleases) function as dimers, whereas the enzymes that have two copies of this motif (*i.e.*, di-LAGLIDADG meganucleases) function as monomers.

[0102] All LAGLIDADG family members recognize and cleave relatively long sequences (> 12bp), leaving four nucleotide 3' overhangs. These enzymes also share a number of structural motifs in addition to the LAGLIDADG motif, including a similar arrangement of anti-parallel β -strands at the protein-DNA interface. Amino acids within these conserved structural motifs are responsible for interacting with the DNA bases to confer recognition sequence specificity. . The overall structural similarity between some members of the family (*e.g.*, I-CreI, I-MsoI, I-SceI and I-CeuI) has been elucidated by X-ray crystallography. Accordingly, the members of this family can be modified at particular amino acids within such structural motifs to change the over-all activity or sequence-specificity of the enzymes, and corresponding modifications can reasonable be expected to have similar results in other family members. See, generally, Chevalier *et al.* (2001), *Nucleic Acid Res.* 29(18): 3757-3774).

2.2.1 Meganucleases Derived from I-CreI

[0103] In one aspect, the present invention relates to rationally-designed meganucleases which are based upon or derived from the I-CreI meganuclease of *Chlamydomonas reinhardtii*. The wild-type amino acid sequence of the I-CreI meganuclease is shown in SEQ ID NO: 1, which corresponds to Genbank Accession # PO5725. Two recognition sequence half sites of the wild-type I-CreI meganuclease from crystal structure PDB # 1BP7 are shown below:

computer modeling of the meganuclease-DNA complexes). Second, existing and potential contact residues are determined based on the distances between the β -carbons of the surrounding amino acid positions and the nucleic acid bases on each DNA strand at the specified base position. For example, and without limitation, as shown in Figure 1(A), the I-CreI wild type meganuclease-DNA contact residues at position -4 involve a glutamine at position 26 which hydrogen bonds to an A base on the antisense DNA strand. Residue 77 was also identified as potentially being able to contact the -4 base on the DNA sense strand. The β -carbon of residue 26 is 5.9 Å away from N7 of the A base on the antisense DNA strand, and the β -carbon of residue 77 is 7.15 Å away from the C5-methyl of the T on the sense strand. According to the distance and base chemistry rules described herein, a C on the sense strand could hydrogen bond with a glutamic acid at position 77 and a G on the antisense strand could bond with glutamine at position 26 (mediated by a water molecule, as observed in the wild-type I-CreI crystal structure) (see Fig 1(C)); a G on the sense strand could hydrogen bond with an arginine at position 77 and a C on the antisense strand could hydrogen bond with a glutamic acid at position 26 (see Fig 1(D)); an A on the sense strand could hydrogen bond with a glutamine at position 77 and a T on the antisense strand could form hydrophobic contacts with an alanine at position 26 (see Fig. 1(E)). If the base specific contact is provided by position 77, then the wild-type contact, Q26, can be substituted (*e.g.*, with a serine residue) to reduce or remove its influence on specificity. Alternatively, complementary mutations at positions 26 and 77 can be combined to specify a particular base pair (*e.g.*, A26 specifies a T on the antisense strand and Q77 specifies an A on the sense strand (Fig. 1(E)). These predicted residue substitutions have all been validated experimentally.

[0108] Thus, in accordance with the invention, a substantial number of amino acid modifications to the DNA recognition domain of the I-CreI meganuclease have been identified which, singly or in combination, result in recombinant meganucleases with specificities altered at individual bases within the DNA recognition sequence half-site, such that these rationally-designed meganucleases have half-sites different from the wild-type enzyme. The amino acid modifications of I-CreI and the resulting change in recognition sequence half-site specificity are shown in Table 1:

TABLE 1

Posn.	Favored Sense-Strand Base										
	A	C	G	T	A/T	A/C	A/G	C/T	G/T	A/G/T	A/C/G/T
-1	Y75 L75* C75* Y139* C46* A46*	R70* H75* R75* H46* K46* R46*	K70 E70* E75* E46* D46*	Q70* C70 L70 Y75* Q75* H75* H139 Q46* H46*				T46*			G70 A70 S70 G46*
-2	Q70 T44* A44* V44* I44* L44* N44*	E70 D70 K44* R44*	H70 D44* E44*	Q44*	C44*						
-3	Q68 C24* I24*	E68 F68 K24* R24*	R68	M68 C68 L68 F68		H68		Y68	K68		
-4	A26* Q77	E77 K26*	R77 E26*					S77 Q26*			S26*
-5		E42	R42			K28*	C28* Q42				M66 K66

-6	Q40 C28*	E40 R28*	R40	C40 I40 V40 C79 I79 V79 Q28*	A40 A79 A28* H28*					S40 S28*
-7	N30* Q38	E38 K30* R30*	K38 R38 E30*	I38 L38		C38				H38 N38 Q30*
-8	F33 Y33	E33 D33	F33 H33	L33 V33 I33 F33 C33		R32*				
-9		E32	R32 K32	L32 V32 A32 C32				D32 I32		S32 N32 H32 Q32 T32

Bold entries are wild-type contact residues and do not constitute "modifications" as used herein.

An asterisk indicates that the residue contacts the base on the antisense strand.

2.2.2 Meganucleases Derived from I-MsoI

[0109] In another aspect, the present invention relates to rationally-designed meganucleases which are based upon or derived from the I-MsoI meganuclease of *Monomastix sp.* The wild-type amino acid sequence of the I-MsoI meganuclease is shown in SEQ ID NO: 6, which corresponds to Genbank Accession # AAL34387. Two recognition sequence half-sites of the wild-type I-MsoI meganuclease from crystal structure PDB # 1M5X are shown below:

```

Position      -9-8-7-6-5-4-3-2-1
5'-C A G A A C G T C G T G A G A C A G T T C C-3' SEQ ID NO: 7
3'-G T C T T G C A G C A C T C T G T C A A G G-5' SEQ ID NO: 8
Position      -1-2-3-4-5-6-7-8-9

```

Note that the recognition sequence is not perfectly palindromic, even outside the central four base pairs. The two recognition sequence half-sites are shown in bold on their respective sense strands.

[0110] In accordance with the invention, a substantial number of amino acid modifications to the DNA recognition domain of the I-MsoI meganuclease have been identified which, singly or in combination, can result in recombinant meganucleases with specificities altered at individual bases within the DNA recognition sequence half-sites, such that these rationally-designed meganucleases have recognition sequences different from the wild-type enzyme. Amino acid modifications of I-MsoI and the predicted change in recognition sequence half-site specificity are shown in Table 2:

TABLE 2

Position	Favored Sense-Strand Base			
	A	C	G	T
-1	K75* Q77 A49* C49* K79*	D77 E77 K49* R75* K75* R79* K79*	K77 R77 E49* E79*	C77 L77 Q79*
-2	Q75 K81 C47* I47* L47*	E75 D75 R47* K47* K81* R81*	K75 E47* E81*	A75 C75 V75 I75 T75 Q47* Q81*
-3	Q72 C26* L26* V26* A26* I26*	E72 Y72 H26* K26* R26*	R72 K72 Y26* F26*	K72 Y72 H26*
-4	K28 Q83	K28* R28* E83	R83 K83	K28 K83 Q28*
-5	K28 C28* L28* I28*	K28* R28*	R45 E28*	Q28*
-6	I30* V30* S30* L30* Q43	E43 E85 K30* R30*	R43 K43 K85 R85 E30* D30*	K43 I85 V85 L85 Q30*
-7	Q41	E32 E41	R32 R41 K41	K32 M41 L41 I41
-8	Y35 K35	E32	R32 K32 K35 R35	K32 K35
-9	N34 H34	D34 E34 S34	K34 R34 H34	S34 C34 V34 T34 A34

Bold entries are represent wild-type contact residues and do not constitute "modifications" as used herein. An asterisk indicates that the residue contacts the base on the antisense strand.

2.2.3 Meganucleases Derived from I-SceI

[0111] In another aspect, the present invention relates to rationally-designed meganucleases which are based upon or derived from the I-SceI meganuclease of *Saccharomyces cerevisiae*. The wild-type amino acid sequence of the I-SceI meganuclease is shown in SEQ ID NO: 9, which corresponds to Genbank Accession # CAA09843. The recognition sequence of the wild-type I-SceI meganuclease from crystal structure PDB # 1R7M is shown below:

Sense	5'	T	T	A	C	C	C	T	G	T	T	A	T	C	C	C	T	A	G	-3'	SEQ ID NO:	10
Antisense	3'	A	A	T	G	G	G	A	C	A	A	T	A	G	G	G	A	T	C	-5'	SEQ ID NO:	11
Position		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18			

Note that the recognition sequence is non-palindromic and there are not four base pairs separating half-sites.

[0112] In accordance with the invention, a substantial number of amino acid modifications to the DNA recognition domain of the I-SceI meganuclease have been identified which, singly or in combination, can result in recombinant meganucleases with specificities altered at individual bases within the DNA recognition sequence, such that these rationally-designed meganucleases have recognition sequences different from the wild-type enzyme. The amino acid modifications of I-SceI and the predicted change in recognition sequence specificity are shown in Table 3:

TABLE 3

Position	Favored Sense-Strand Base			
	A	C	G	T
4	K50	R50* K50* E57	E50* R57 K57	K57 M57 Q50*
5	K48 Q102	R48* K48* E102 E59	E48* K102 R102	Q48* C102 L102 V102
6	K59	R59* K59*	K84 E59*	Q59* Y46
7	C46* L46* V46*	R46* K46* E86	K86 R86 E46*	K68 C86 L86 Q46*
8	K61* S61* V61* A61* L61*	E88 R61* H61*	E61* R88 K88	K88 Q61* H61*
9	T98* C98* V98* L98*	R98* K98*	E98* D98*	Q98*
10	V96* C96* A96*	K96* R96*	D96* E96*	Q96*
11	C90* L90*	K90* R90*	E90*	Q90*
12	Q193	E165 E193 D193	K165 R165	C165 L165 C193 V193 A193 T193 S193
13	C193* L193*	K193* R193* D192	E193* D193* K163 R192	Q193* C163 L163
14	L192* C192*	E161 R192* K192*	K147 K161 R161 R197 D192* E192*	K161 Q192*

15		E151	K151	C151 L151 K151
17	N152* S152* C150* L150* V150* T150*	K152* K150*	N152* S152* D152* D150* E150*	Q152* Q150*
18	K155* C155*	R155* K155*	E155*	H155* Y155*

Bold entries are wild-type contact residues and do not constitute "modifications" as used herein.

An asterisk indicates that the residue contacts the base on the antisense strand.

2.2.4 Meganucleases Derived from I-CeuI

[0113] In another aspect, the present invention relates to rationally-designed meganucleases which are based upon or derived from the I-CeuI meganuclease of *Chlamydomonas eugametos*. The wild-type amino acid sequence of the I-CeuI meganuclease is shown in SEQ ID NO: 12, which corresponds to Genbank Accession # P32761. Two recognition sequence half sites of the wild-type I-CeuI meganuclease from crystal structure PDB # 2EX5 are shown below:

```

Position      -9-8-7-6-5-4-3-2-1
5'-A T A A C G G T C C T A A G G T A G C G A A-3' SEQ ID NO:
13
3'-T A T T G C C A G G A T T C C A T C G C T T-5' SEQ ID NO:
14
Position                                -1-2-3-4-5-6-7-8-9

```

Note that the recognition sequence is non-palindromic, even outside the central four base pairs, despite the fact that I-CeuI is a homodimer, due to the natural degeneracy in the I-CeuI recognition interface (Spiegel *et al.* (2006), *Structure* 14:869-80). The two recognition sequence half-sites are shown in bold on their respective sense strands.

[0114] In accordance with the invention, a substantial number of amino acid modifications to the DNA recognition domain of the I-CeuI meganuclease have been identified which, singly or in combination, result in recombinant meganucleases with specificities altered at individual bases within the DNA recognition sequence half-site,

such that these rationally-designed meganucleases can have recognition sequences different from the wild-type enzyme. The amino acid modifications of I-CeuI and the predicted change in recognition sequence specificity are shown in Table 4:

TABLE 4

Position	Favored Sense-Strand Base			
	A	C	G	T
-1	C92* A92* V92*	K116* R116* D116* K92*	E116* E92*	Q116* Q92*
-2	Q117 C90* L90* V90*	E117 D117 R174* K124* K90* R90* K68*	K117 R124 K124 E124* E90* D90*	C117 V117 T117 Q90*
-3	C70* V70* T70* L70* K70*	K70*	E70* E88*	Q70*
-4	Q126 N126 K88* L88* C88* C72* L72* V72*	E126 D126 R88* K88* K72*	R126 K126 E88* D88*	K126 L126 Q88*
-5	C74* L74* V74* T74*	K74*	E74* K128 R128 E128	C128 L128 V128 T128
-6	Q86	D86 E86 R84* K84*	K128 R128 R86 K86 E84*	K86 C86 L86
-7	L76* C76* K76*	R76* K76* H76*	E76* R84	H76* Q76*
-8	Y79 R79 Q76	D79 E79 D76	R79 K79 K76	C79 L79 V79

		E76	R76	L76
-9	Q78	D78	R78	K78
	N78	E78	K78	V78
	H78		H78	L78
	K78			C78
				T78

Bold entries are wild-type contact residues and do not constitute "modifications" as used herein.

An asterisk indicates that the residue contacts the base on the antisense strand.

2.2.5 Specifically-Excluded Recombinant Meganucleases

[0115] The present invention is not intended to embrace certain recombinant meganucleases which have been described in the prior art, and which have been developed by alternative methods. These excluded meganucleases include those described by Arnould *et al.* (2006), *J. Mol. Biol.* 355: 443-58; Sussman *et al.* (2004), *J. Mol. Biol.* 342: 31-41; Chames *et al.* (2005), *Nucleic Acids Res.* 33: e178; Seligman *et al.* (2002), *Nucleic Acids Res.* 30: 3870-9; and Ashworth *et al.* (2006), *Nature* 441(7093):656-659; the entire disclosures of which are hereby incorporated by reference, including recombinant meganucleases based on I-CreI with single substitutions selected from C33, R33, A44, H33, K32, F33, R32, A28, A70, E33, V33, A26, and R66. Also excluded are recombinant meganucleases based on I-CreI with three substitutions selected from A68/N70/N75 and D44/D70/N75, or with four substitutions selected from K44/T68/G60/N75 and R44/A68/T70/N75. Lastly, specifically excluded is the recombinant meganuclease based on I-MsoI with the pair of substitutions L28 and R83. These substitutions or combinations of substitutions are referred to herein as the "excluded modifications."

2.2.6 Meganucleases with Multiple Changes in the Recognition Sequence Half-Site

[0116] In another aspect, the present invention relates to rationally-designed meganucleases which are produced by combining two or more amino acid modifications as described in sections 2.2.1-2.2.4 above, in order to alter half-site preference at two or more positions in a DNA recognition sequence half-site. For example, without limitation,

and as more fully described below, the enzyme DJ1 was derived from I-CreI by incorporating the modifications R30/E38 (which favor C at position -7), R40 (which favors G at position -6), R42 (which favors G at position -5), and N32 (which favors complete degeneracy at position -9). The rationally-designed DJ1 meganuclease invariantly recognizes C₋₇G₋₆G₋₅ compared to the wild-type preference for A₋₇A₋₆C₋₅, and has increased tolerance for A at position -9.

[0117] The ability to combine residue substitutions that affect different base positions is due in part to the modular nature of the LAGLIDADG meganucleases. A majority of the base contacts in the LAGLIDADG recognition interfaces are made by individual amino acid side chains, and the interface is relatively free of interconnectivity or hydrogen bonding networks between side chains that interact with adjacent bases. This generally allows manipulation of residues that interact with one base position without affecting side chain interactions at adjacent bases. The additive nature of the mutations listed in sections 2.2.1-2.2.4 above is also a direct result of the method used to identify these mutations. The method predicts side chain substitutions that interact directly with a single base. Interconnectivity or hydrogen bonding networks between side chains is generally avoided to maintain the independence of the substitutions within the recognition interface.

[0118] Certain combinations of side chain substitutions are completely or partially incompatible with one another. When an incompatible pair or set of amino acids are incorporated into a rationally-designed meganuclease, the resulting enzyme will have reduced or eliminated catalytic activity. Typically, these incompatibilities are due to steric interference between the side chains of the introduced amino acids and activity can be restored by identifying and removing this interference. Specifically, when two amino acids with large side chains (*e.g.*, amino acids from group 2 or 3) are incorporated at amino acid positions that are adjacent to one another in the meganuclease structure (*e.g.*, positions 32 and 33, 28 and 40, 28 and 42, 42 and 77, or 68 and 77 in the case of meganucleases derived from I-CreI), it is likely that these two amino acids will interfere with one another and reduce enzyme activity. This interference can be eliminated by substituting one or both incompatible amino acids to an amino acid with a smaller side chain (*e.g.*, group 1 or group 2). For example, in rationally-designed meganucleases

derived from I-CreI, K28 interferes with both R40 and R42. To maximize enzyme activity, R40 and R42 can be combined with a serine or aspartic acid at position 28.

[0119] Combinations of amino substitutions, identified as described herein, can be used to rationally alter the specificity of a wild-type meganuclease (or a previously modified meganuclease) from an original recognition sequence to a desired recognition sequence which may be present in a nucleic acid of interest (*e.g.*, a genome). Figure 2A, for example, shows the "sense" strand of the I-CreI meganuclease recognition sequence WT (SEQ ID NO: 4) as well as a number of other sequences for which a rationally-designed meganuclease would be useful. Conserved bases between the WT recognition sequence and the desired recognition sequence are shaded. In accordance with the invention, recombinant meganucleases based on the I-CreI meganuclease can be rationally-designed for each of these desired recognition sequences, as well as any others, by suitable amino acid substitutions as described herein.

3. Rationally-Designed Meganucleases with Altered DNA-Binding Affinity

[0120] As described above, the DNA-binding affinity of the recombinant meganucleases of the invention can be modulated by altering certain amino acids that form the contact surface with the phosphodiester backbone of DNA. The contact surface comprises those amino acids in the enzyme with β -carbons less than 9 Å from the DNA backbone, and with side chains oriented toward the DNA, irrespective of whether the residues make contacts with the DNA backbone in the wild-type meganuclease-DNA complex. Because DNA-binding is a necessary precursor to enzyme activity, increases/decreases in DNA-binding affinity have been shown to cause increases/decreases, respectively, in enzyme activity. However, increases/decreases in DNA-binding affinity also have been shown to cause decreases/increases in the meganuclease sequence-specificity. Therefore, both activity and specificity can be modulated by modifying the phosphodiester backbone contacts.

[0121] Specifically, to increase enzyme activity/decrease enzyme specificity:

[0122] (i) Remove electrostatic repulsion between the enzyme and DNA backbone. If an identified amino acid has a negatively-charged side chain (*e.g.*, aspartic acid, glutamic acid) which would be expected to repulse the negatively-charged DNA

backbone, the repulsion can be eliminated by substituting an amino acid with an uncharged or positively-charged side chain, subject to effects of steric interference. An experimentally verified example is the mutation of glutamic acid 80 in I-CreI to glutamine.

[0123] (ii) Introduce electrostatic attraction interaction between the enzyme and the DNA backbone. At any of the positions of the contact surface, the introduction of an amino acid with a positively-charged side chain (*e.g.*, lysine or arginine) is expected to increase binding affinity, subject to effects of steric interference.

[0124] (iii) Introduce a hydrogen-bond between the enzyme and the DNA backbone. If an amino acid of the contact surface does not make a hydrogen bond with the DNA backbone because it lacks an appropriate hydrogen-bonding functionality or has a side chain that is too short, too long, and/or too inflexible to interact with the DNA backbone, a polar amino acid capable of donating a hydrogen bond (*e.g.*, serine, threonine, tyrosine, histidine, glutamine, asparagine, lysine, cysteine, or arginine) with the appropriate length and flexibility can be introduced, subject to effects of steric interference.

[0125] Specifically, to decrease enzyme activity/increase enzyme specificity:

[0126] (i) Introduce electrostatic repulsion between the enzyme and the DNA backbone. At any of the positions of the contact surface, the introduction of an amino acid with a negatively-charged side chain (*e.g.*, glutamic acid, aspartic acid) is expected to decrease binding affinity, subject to effects of steric interference.

[0127] (ii) Remove electrostatic attraction between the enzyme and DNA. If any amino acid of the contact surface has a positively-charged side chain (*e.g.*, lysine or arginine) that interacts with the negatively-charged DNA backbone, this favorable interaction can be eliminated by substituting an amino acid with an uncharged or negatively-charged side chain, subject to effects of steric interference. An experimentally verified example is the mutation of lysine 116 in I-CreI to aspartic acid.

[0128] (iii) Remove a hydrogen-bond between the enzyme and the DNA backbone. If any amino acid of the contact surface makes a hydrogen bond with the DNA backbone, it can be substituted to an amino acid that would not be expected to make

a similar hydrogen bond because its side chain is not appropriately functionalized or it lacks the necessary length/flexibility characteristics.

[0129] For example, in some recombinant meganucleases based on I-CreI, the glutamic acid at position 80 in the I-CreI meganuclease is altered to either a lysine or a glutamine to increase activity. In another embodiment, the tyrosine at position 66 of I-CreI is changed to arginine or lysine, which increases the activity of the meganuclease. In yet another embodiment, enzyme activity is decreased by changing the lysine at position 34 of I-CreI to aspartic acid, changing the tyrosine at position 66 to aspartic acid, and/or changing the lysine at position 116 to aspartic acid.

[0130] The activities of the recombinant meganucleases can be modulated such that the recombinant enzyme has anywhere from no activity to very high activity with respect to a particular recognition sequence. For example, the DJ1 recombinant meganuclease when carrying glutamic acid mutation at position 26 loses activity completely. However, the combination of the glutamic acid substitution at position 26 and a glutamine substitution at position 80 creates a recombinant meganuclease with high specificity and activity toward a guanine at -4 within the recognition sequence half-site (see Figure 1(D)).

[0131] In accordance with the invention, amino acids at various positions in proximity to the phosphodiester DNA backbone can be changed to simultaneously affect both meganuclease activity and specificity. This "tuning" of the enzyme specificity and activity is accomplished by increasing or decreasing the number of contacts made by amino acids with the phosphodiester backbone. A variety of contacts with the phosphodiester backbone can be facilitated by amino acid side chains. In some embodiments, ionic bonds, salt bridges, hydrogen bonds, and steric hindrance affect the association of amino acid side chains with the phosphodiester backbone. For example, for the I-CreI meganuclease, alteration of the lysine at position 116 to an aspartic acid removes a salt bridge between nucleic acid base pairs at positions -8 and -9, reducing the rate of enzyme cleavage but increasing the specificity.

[0132] The residues forming the backbone contact surface of each of the wild-type I-CreI (SEQ ID NO: 1), I-MsoI (SEQ ID NO: 6), I-SceI (SEQ ID NO: 9) and I-CeuI (SEQ ID NO: 12) meganucleases are identified in Table 5 below:

TABLE 5

I-CreI	I-MsoI	I-SceI	I-CeuI
P29, K34, T46, K48, R51, V64, Y66, E80, I81, K82, L112, K116, D137, K139, T140, T143	K36, Q41, R51, N70, I85, G86, S87, T88, H89, Y118, Q122, K123, Q139, K143, R144, E147, S150, N152	N15, N17, L19, K20, K23, K63, L80, S81, H84, L92, N94, N120, K122, K148, Y151, K153, T156, N157, S159, N163, Q165, S166, Y188, K190, I191, K193, N194, K195, Y199, D201, S202, Y222, K223	K21, D25, K28, K31, S68, N70, H94, R112, R114, S117, N120, D128, N129, R130, H172

[0133] To increase the affinity of an enzyme and thereby make it more active/less specific:

- (1) Select an amino acid from Table 5 for the corresponding enzyme that is either negatively-charged (D or E), hydrophobic (A, C, F, G, I, L, M, P, V, W, Y), or uncharged/polar (H, N, Q, S, T).
- (2) If the amino acid is negatively-charged or hydrophobic, mutate it to uncharged/polar (less effect) or positively-charged (K or R, more effect).
- (3) If the amino acid is uncharged/polar, mutate it to positively-charged.

[0134] To decrease the affinity of an enzyme and thereby make it less active/more specific:

- (1) Select an amino acid from Table 5 for the corresponding enzyme that is either positively-charged (K or R), hydrophobic (A, C, F, G, I, L, M, P, V, W, Y), or uncharged/polar (H, N, Q, S, T).

- (2) If the amino acid is positively-charged, mutate it to uncharged/polar (less effect) or negatively-charged (more effect).
- (3) If the amino acid is hydrophobic or uncharged/polar, mutate it to negatively-charged.

4. Heterodimeric Meganucleases

[0135] In another aspect, the invention provides meganucleases which are heterodimers formed by the association of two monomers, one of which may be a wild-type and one or both of which may be a non-naturally-occurring or recombinant form. For example, wild-type I-CreI meganuclease is normally a homodimer composed of two monomers that each bind to one half-site in the pseudo-palindromic recognition sequence. A heterodimeric recombinant meganuclease can be produced by combining two meganucleases that recognize different half-sites, for example by co-expressing the two meganucleases in a cell or by mixing two meganucleases in solution. The formation of heterodimers can be favored over the formation of homodimers by altering amino acids on each of the two monomers that affect their association into dimers. In particular embodiments, certain amino acids at the interface of the two monomers are altered from negatively-charged amino acids (D or E) to positively-charged amino acids (K or R) on a first monomer and from positively-charged amino acids to negatively-charged amino acids on a second monomer (Table 6). For example, in the case of meganucleases derived from I-CreI, lysines at positions 7 and 57 are mutated to glutamic acids in the first monomer and glutamic acids at positions 8 and 61 are mutated to lysines in the second monomer. The result of this process is a pair of monomers in which the first monomer has an excess of positively-charged residues at the dimer interface and the second monomer has an excess of negatively-charged residues at the dimer interface. The first and second monomer will, therefore, associate preferentially over their identical monomer pairs due to the electrostatic interactions between the altered amino acids at the interface.

TABLE 6

I-CreI: First Monomer Substitutions	I-CreI: Second Monomer Substitutions
--	---

K7 to E7 or D7 K57 to E57 or D57 K96 to E96 or D96	E8 to K8 or R8 E61 to K61 or R61
I-MsoI: First Monomer Substitutions	I-MsoI: Second Monomer Substitutions
R302 to E302 or D302	D20 to K60 or R60 E11 to K11 or R11 Q64 to K64 or R64
I-CeuI: First Monomer Substitutions	I-CeuI: Second Monomer Substitutions
R93 to E93 or D93	E152 to K152 or R152

[0136] Alternatively, or in addition, certain amino acids at the interface of the two monomers can be altered to sterically hinder homodimer formation. Specifically, amino acids in the dimer interface of one monomer are substituted with larger or bulkier residues that will sterically prevent the homodimer. Amino acids in the dimer interface of the second monomer optionally can be substituted with smaller residues to compensate for the bulkier residues in the first monomer and remove any clashes in the heterodimer, or can be unmodified.

[0137] In another alternative or additional embodiment, an ionic bridge or hydrogen bond can be buried in the hydrophobic core of a heterodimeric interface. Specifically, a hydrophobic residue on one monomer at the core of the interface can be substituted with a positively charged residue. In addition, a hydrophobic residue on the second monomer, that interacts in the wild type homodimer with the hydrophobic residue substituted in the first monomer, can be substituted with a negatively charged residue. Thus, the two substituted residues can form an ionic bridge or hydrogen bond. At the same time, the electrostatic repulsion of an unsatisfied charge buried in a hydrophobic interface should disfavor homodimer formation.

[0138] Finally, as noted above, each monomer of the heterodimer can have different amino acids substituted in the DNA recognition region such that each has a different DNA half-site and the combined dimeric DNA recognition sequence is non-palindromic.

5. Methods of Producing Recombinant Cells and Organisms

[0139] Aspects of the present invention further provide methods for producing recombinant, transgenic or otherwise genetically-modified cells and organisms using rationally-designed meganucleases. Thus, in certain embodiments, recombinant meganucleases are developed to specifically cause a double-stranded break at a single site or at relatively few sites in the genomic DNA of a cell or an organism to allow for precise insertion(s) of a sequence of interest by homologous recombination. In other embodiments, recombinant meganucleases are developed to specifically cause a double-stranded break at a single site or at relatively few sites in the genomic DNA of a cell or an organism to either (a) allow for rare insertion(s) of a sequence of interest by non-homologous end-joining or (b) allow for the disruption of the target sequence by non-homologous end-joining. As used herein with respect to homologous recombination or non-homologous end-joining of sequences of interest, the term "insertion" means the ligation of a sequence of interest into a chromosome such that the sequence of interest is integrated into the chromosome. In the case of homologous recombination, an inserted sequence can replace an endogenous sequence, such that the original DNA is replaced by exogenous DNA of equal length, but with an altered nucleotide sequence. Alternatively, an inserted sequence can include more or fewer bases than the sequence it replaces.

[0140] Therefore, in accordance with this aspect of the invention, the recombinant organisms include, but are not limited to, monocot plant species such as rice, wheat, corn (maize) and rye, and dicot species such as legumes (*e.g.*, kidney beans, soybeans, lentils, peanuts, peas), alfalfa, clover, tobacco and *Arabidopsis* species. In addition, the recombinant organisms can include, but are not limited to, animals such as humans and non-human primates, horses, cows, goats, pigs, sheep, dogs, cats, guinea pigs, rats, mice, lizards, fish and insects such as *Drosophila* species. In other embodiments, the organism is a fungus such as a *Candida*, *Neurospora* or *Saccharomyces* species.

[0141] In some embodiments, the methods of the invention involve the introduction of a sequence of interest into a cell such as a germ cell or stem cell that can become a mature recombinant organism or allow the resultant genetically-modified organism to give rise to progeny carrying the inserted sequence of interest in its genome.

[0142] Meganuclease proteins can be delivered into cells to cleave genomic DNA, which allows for homologous recombination or non-homologous end-joining at the cleavage site with a sequence of interest, by a variety of different mechanisms known in the art. For example, the recombinant meganuclease protein can be introduced into a cell by techniques including, but not limited to, microinjection or liposome transfections (see, e.g., Lipofectamine™, Invitrogen Corp., Carlsbad, CA). The liposome formulation can be used to facilitate lipid bilayer fusion with a target cell, thereby allowing the contents of the liposome or proteins associated with its surface to be brought into the cell. Alternatively, the enzyme can be fused to an appropriate uptake peptide such as that from the HIV TAT protein to direct cellular uptake (see, e.g., Hudecz *et al.* (2005), *Med. Res. Rev.* 25: 679-736).

[0143] Alternatively, gene sequences encoding the meganuclease protein are inserted into a vector and transfected into a eukaryotic cell using techniques known in the art (see, e.g., Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley 1999). The sequence of interest can be introduced in the same vector, a different vector, or by other means known in the art.

[0144] Non-limiting examples of vectors for DNA transfection include virus vectors, plasmids, cosmids, and YAC vectors. Transfection of DNA sequences can be accomplished by a variety of methods known to those of skill in the art. For instance, liposomes and immunoliposomes are used to deliver DNA sequences to cells (see, e.g., Lasic *et al.* (1995), *Science* 267: 1275-76). In addition, viruses can be utilized to introduce vectors into cells (see, e.g., U.S. Pat. No. 7,037,492). Alternatively, transfection strategies can be utilized such that the vectors are introduced as naked DNA (see, e.g., Rui *et al.* (2002), *Life Sci.* 71(15): 1771-8).

[0145] General methods for delivering nucleic acids into cells include: (1) chemical methods (Graham *et al.* (1973), *Virology* 54(2):536-539; Zatloukal *et al.* (1992), *Ann. N.Y. Acad. Sci.*, 660:136-153; (2) physical methods such as microinjection

(Capecchi (1980), *Cell* 22(2):479-488, electroporation (Wong *et al.* (1982), *Biochim. Biophys. Res. Commun.* 107(2):584-587; Fromm *et al.* (1985), *Proc. Nat'l Acad. Sci. USA* 82(17):5824-5828; U.S. Pat. No. 5,384,253) and ballistic injection (Johnston *et al.* (1994), *Methods Cell. Biol.* 43(A): 353-365; Fynan *et al.* (1993), *Proc. Nat'l Acad. Sci. USA* 90(24): 11478-11482); (3) viral vectors (Clapp (1993), *Clin. Perinatol.* 20(1): 155-168; Lu *et al.* (1993), *J. Exp. Med.* 178(6):2089-2096; Eglitis *et al.* (1988), *Adv. Exp. Med. Biol.* 241:19-27; Eglitis *et al.* (1988), *Biotechniques* 6(7):608-614); and (4) receptor-mediated mechanisms (Curiel *et al.* (1991), *Proc. Nat'l Acad. Sci. USA* 88(19):8850-8854; Curiel *et al.* (1992), *Hum. Gen. Ther.* 3(2):147-154; Wagner *et al.* (1992), *Proc. Nat'l Acad. Sci. USA* 89 (13):6099-6103).

[0146] In certain embodiments, a genetically-modified plant is produced, which contains the sequence of interest inserted into the genome. In certain embodiments, the genetically-modified plant is produced by transfecting the plant cell with DNA sequences corresponding to the recombinant meganuclease and the sequence of interest, which may or may not be flanked by the meganuclease recognition sequences and/or sequences substantially identical to the target sequence. In other embodiments, the genetically-modified plant is produced by transfecting the plant cell with DNA sequences corresponding to the recombinant meganuclease only, such that cleavage promotes non-homologous end-joining and disrupts the target sequence containing the recognition sequence. In such embodiments, the meganuclease sequences are under the control of regulatory sequences that allow for expression of the meganuclease in the host plant cells. These regulatory sequences include, but are not limited to, constitutive plant promoters such as the NOS promoter, chemically-inducible gene promoters such as the dexamethasone-inducible promoter (see, *e.g.*, Gremillon *et al.* (2004), *Plant J.* 37:218-228), and plant tissue specific promoters such as the LGC1 promoter (see, *e.g.*, Singh *et al.* (2003), *FEBS Lett.* 542:47-52).

[0147] Suitable methods for introducing DNA into plant cells include virtually any method by which DNA can be introduced into a cell, including but not limited to *Agrobacterium* infection, PEG-mediated transformation of protoplasts (Omirulleh *et al.* (1993), *Plant Molecular Biology*, 21:415-428), desiccation/inhibition-mediated DNA

uptake, electroporation, agitation with silicon carbide fibers, ballistic injection or microprojectile bombardment, and the like.

[0148] In other embodiments, a genetically-modified animal is produced using a recombinant meganuclease. As with plant cells, the nucleic acid sequences can be introduced into a germ cell or a cell that will eventually become a transgenic organism. In some embodiments, the cell is a fertilized egg, and exogenous DNA molecules can be injected into the pro-nucleus of the fertilized egg. The micro-injected eggs are then transferred into the oviducts of pseudopregnant foster mothers and allowed to develop. The recombinant meganuclease is expressed in the fertilized egg (*e.g.*, under the control of a constitutive promoter, such as 3-phosphoglycerate kinase), and facilitates homologous recombination of the sequence of interest into one or a few discrete sites in the genome. Alternatively, the genetically-modified animals can be obtained by utilizing recombinant embryonic stem ("ES") cells for the generation of the transgenics, as described by Gossler *et al.* (1986), *Proc. Natl. Acad. Sci. USA* 83:9065-9069.

[0149] In certain embodiments, a recombinant mammalian expression vector is capable of directing tissue-specific expression of the nucleic acid preferentially in a particular cell type. Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987), *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton (1988), *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989), *EMBO J.* 8: 729-733) and immunoglobulins (Banerji *et al.* (1983), *Cell* 33: 729-740; Queen and Baltimore (1983), *Cell* 33: 741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989), *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund *et al.* (1985), *Science* 230: 912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Pat. Pub. EP 0 264 166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990), *Science* 249: 374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989), *Genes Dev.* 3: 537-546).

[0150] In certain embodiments, a rationally-designed meganuclease may be tagged with a peptide epitope (*e.g.*, an HA, FLAG, or Myc epitope) to monitor expression

levels or localization. In some embodiments, the meganuclease may be fused to a sub-cellular localization signal such as a nuclear-localization signal (*e.g.*, the nuclear localization signal from SV40) or chloroplast or mitochondrial localization signals. In other embodiments, the meganuclease may be fused to a nuclear export signal to localize it to the cytoplasm. The meganuclease may also be fused to an unrelated protein or protein domain such as a protein that stimulates DNA-repair or homologous recombination (*e.g.*, recA, RAD51, RAD52, RAD54, RAD57 or BRCA2).

6. Methods for Gene Therapy

[0151] Aspects of the invention allow for the use of recombinant meganuclease for gene therapy. As used herein, “gene therapy” means therapeutic treatments that comprise introducing into a patient a functional copy of at least one gene, or gene regulatory sequence such as a promoter, enhancer, or silencer to replace a gene or gene regulatory region that is defective in its structure and/or function. The term “gene therapy” can also refer to modifications made to a deleterious gene or regulatory element (*e.g.*, oncogenes) that reduce or eliminate expression of the gene. Gene therapy can be performed to treat congenital conditions, conditions resulting from mutations or damage to specific genetic loci over the life of the patient, or conditions resulting from infectious organisms.

[0152] In some aspects of the invention, dysfunctional genes are replaced or disabled by the insertion of exogenous nucleic acid sequences into a region of the genome affecting gene expression. In certain embodiments, the recombinant meganuclease is targeted to a particular sequence in the region of the genome to be modified so as to alleviate the condition. The sequence can be a region within an exon, intron, promoter, or other regulatory region that is causing dysfunctional expression of the gene. As used herein, the term “dysfunctional expression” means aberrant expression of a gene product either by the cell producing too little of the gene product, too much of the gene product, or producing a gene product that has a different function such as lacking the necessary function or having more than the necessary function.

[0153] Exogenous nucleic acid sequences inserted into the modified region can be used to provide “repaired” sequences that normalize the gene. Gene repair can be

accomplished by the introduction of proper gene sequences into the gene allowing for proper function to be reestablished. In these embodiments, the nucleic acid sequence to be inserted can be the entire coding sequence for a protein or, in certain embodiments, a fragment of the gene comprising only the region to be repaired. In other embodiments the nucleic acid sequence to be inserted comprises a promoter sequence or other regulatory elements such that mutations causing abnormal expression or regulation are repaired. In other embodiments, the nucleic acid sequence to be inserted contains the appropriate translation stop codon lacking in a mutated gene. The nucleic acid sequence can also have sequences for stopping transcription in a recombinant gene lacking appropriate transcriptional stop signals.

[0154] Alternatively, the nucleic acid sequences can eliminate gene function altogether by disrupting the regulatory sequence of the gene or providing a silencer to eliminate gene function. In some embodiments, the exogenous nucleic acid sequence provides a translation stop codon to prevent expression of the gene product. In other embodiments, the exogenous nucleic acid sequences provide transcription stop element to prevent expression of a full length RNA molecule. In still other embodiments, gene function is disrupted directly by the meganuclease by introducing base insertions, base deletions, and/or frameshift mutations through non-homologous end-joining.

[0155] In many instances, it is desirable to direct the proper genetic sequences to a target cell or population of cells that is the cause of the disease condition. Such targeting of therapeutics prevents healthy cells from being targeted by the therapeutics. This increases the efficacy of the treatment, while decreasing the potentially adverse effects that the treatment could have on healthy cells.

[0156] Delivery of recombinant meganuclease genes and the sequence of interest to be inserted into the genome to the cells of interest can be accomplished by a variety of mechanisms. In some embodiments, the nucleic acids are delivered to the cells by way of viruses with particular viral genes inactivated to prevent reproduction of the virus. Thus, a virus can be altered so that it is capable only of delivery and maintenance within a target cell, but does not retain the ability to replicate within the target cell or tissue. One or more DNA sequences can be introduced to the altered viral genome, so as to produce a viral genome that acts like a vector, and may or may not be inserted into a host genome

and subsequently expressed. More specifically, certain embodiments include employing a retroviral vector such as, but not limited to, the MFG or pLJ vectors. An MFG vector is a simplified Moloney murine leukemia virus vector (MoMLV) in which the DNA sequences encoding the pol and env proteins have been deleted to render it replication defective. A pLJ retroviral vector is also a form of the MoMLV (see, e.g., Korman *et al.* (1987), *Proc. Nat'l Acad. Sci.*, 84:2150-2154). In other embodiments, a recombinant adenovirus or adeno-associated virus can be used as a delivery vector.

[0157] In other embodiments, the delivery of recombinant meganuclease protein and/or recombinant meganuclease gene sequences to a target cell is accomplished by the use of liposomes. The production of liposomes containing nucleic acid and/or protein cargo is known in the art (see, e.g., Lasic *et al.* (1995), *Science* 267: 1275-76). Immunoliposomes incorporate antibodies against cell-associated antigens into liposomes, and can deliver DNA sequences for the meganuclease or the meganuclease itself to specific cell types (see, e.g., Lasic *et al.* (1995), *Science* 267: 1275-76; Young *et al.* (2005), *J. Calif. Dent. Assoc.* 33(12): 967-71; Pfeiffer *et al.* (2006), *J. Vasc. Surg.* 43(5):1021-7). Methods for producing and using liposome formulations are well known in the art, (see, e.g., U.S. Pat. No. 6,316,024, U.S. Pat. No. 6,379,699, U.S. Pat. No. 6,387,397, U.S. Pat. No. 6,511,676 and U.S. Pat. No. 6,593,308, and references cited therein). In some embodiments, liposomes are used to deliver the sequences of interest as well as the recombinant meganuclease protein or recombinant meganuclease gene sequences.

7. Methods for Treating Pathogen Infection.

[0158] Aspects of the invention also provide methods of treating infection by a pathogen. Pathogenic organisms include viruses such as, but not limited to, herpes simplex virus 1, herpes simplex virus 2, human immunodeficiency virus 1, human immunodeficiency virus 2, variola virus, polio virus, Epstein-Barr virus, and human papilloma virus and bacterial organisms such as, but not limited to, *Bacillus anthracis*, *Haemophilus* species, *Pneumococcus* species, *Staphylococcus aureus*, *Streptococcus* species, methicillin-resistant *Staphylococcus aureus*, and *Mycoplasma tuberculosis*.

Pathogenic organisms also include fungal organisms such as, but not limited to, *Candida*, *Blastomyces*, *Cryptococcus*, and *Histoplasma* species.

[0159] In some embodiments, a rationally-designed meganuclease can be targeted to a recognition sequence within the pathogen genome, *e.g.*, to a gene or regulatory element that is essential for growth, reproduction, or toxicity of the pathogen. In certain embodiments, the recognition sequence may be in a bacterial plasmid. Meganuclease-mediated cleavage of a recognition sequence in a pathogen genome can stimulate mutation within a targeted, essential gene in the form of an insertion, deletion or frameshift, by stimulating non-homologous end-joining. Alternatively, cleavage of a bacterial plasmid can result in loss of the plasmid along with any genes encoded on it, such as toxin genes (*e.g.*, *B. anthracis* Lethal Factor gene) or antibiotic resistance genes. As noted above, the meganuclease may be delivered to the infected patient, animal, or plant in either protein or nucleic acid form using techniques that are common in the art. In certain embodiments, the meganuclease gene may be incorporated into a bacteriophage genome for delivery to pathogenic bacteria.

[0160] Aspects of the invention also provide therapeutics for the treatment of certain forms of cancer. Because human viruses are often associated with tumor formation (*e.g.*, Epstein-Barr Virus and nasopharyngeal carcinomas; Human Papilloma Virus and cervical cancer) inactivation of these viral pathogens may prevent cancer development or progression. Alternatively, double-stranded breaks targeted to the genomes of these tumor-associated viruses using rationally-designed meganucleases may be used to trigger apoptosis through the DNA damage response pathway. In this manner, it may be possible to selectively induce apoptosis in tumor cells harboring the viral genome.

8. Methods for Genotyping and Pathogen Identification

[0161] Aspects of the invention also provide tools for *in vitro* molecular biology research and development. It is common in the art to use site-specific endonucleases (*e.g.*, restriction enzymes) for the isolation, cloning, and manipulation of nucleic acids such as plasmids, PCR products, BAC sequences, YAC sequences, viruses, and genomic sequences from eukaryotic and prokaryotic organisms (see, *e.g.*, Ausubel *et al.*, Current

Protocols in Molecular Biology, Wiley 1999). Thus, in some embodiments, a rationally-designed meganuclease may be used to manipulate nucleic acid sequences *in vitro*. For example, rationally-designed meganucleases recognizing a pair of recognition sequences within the same DNA molecule can be used to isolate the intervening DNA segment for subsequent manipulation such as ligation into a bacterial plasmid, BAC, or YAC.

[0162] In another aspect, this invention provides tools for the identification of pathogenic genes and organisms. In one embodiment, rationally-designed meganucleases can be used to cleave recognition sites corresponding to polymorphic genetic regions correlated to disease to distinguish disease-causing alleles from healthy alleles (*e.g.*, a rationally-designed meganuclease which recognizes the ΔF -508 allele of the human CFTR gene, see example 4). In this embodiment, DNA sequences isolated from a human patient or other organism are digested with a rationally-designed meganuclease, possibly in conjunction with additional site-specific nucleases, and the resulting DNA fragment pattern is analyzed by gel electrophoresis, capillary electrophoresis, mass spectrometry, or other methods known in the art. This fragmentation pattern and, specifically, the presence or absence of cleavage by the rationally-designed meganuclease, indicates the genotype of the organism by revealing whether or not the recognition sequence is present in the genome. In another embodiment, a rationally-designed meganuclease is targeted to a polymorphic region in the genome of a pathogenic virus, fungus, or bacterium and used to identify the organism. In this embodiment, the rationally-designed meganuclease cleaves a recognition sequence that is unique to the pathogen (*e.g.*, the spacer region between the 16S and 23S rRNA genes in a bacterium; see, *e.g.*, van der Giessen *et al.* (1994), *Microbiology* 140:1103-1108) and can be used to distinguish the pathogen from other closely-related organisms following endonuclease digest of the genome and subsequent analysis of the fragmentation pattern by electrophoresis, mass spectrometry, or other methods known in the art.

9. Methods for the Production of Custom DNA-binding Domains.

[0163] In another aspect, the invention provides rationally-designed DNA-binding proteins that lack endonuclease cleavage activity. The catalytic activity of a rationally-designed meganuclease can be eliminated by mutating amino acids involved in catalysis

(e.g., the mutation of Q47 to E in I-CreI, see Chevalier *et al.* (2001), *Biochemistry*. 43:14015-14026); the mutation of D44 or D145 to N in I-SceI; the mutation of E66 to Q in I-CeuI; the mutation of D22 to N in I-MsoI). The inactivated meganuclease can then be fused to an effector domain from another protein including, but not limited to, a transcription activator (e.g., the GAL4 transactivation domain or the VP16 transactivation domain), a transcription repressor (e.g., the KRAB domain from the Kruppel protein), a DNA methylase domain (e.g., M.CviPI or M.SssI), or a histone acetyltransferase domain (e.g., HDAC1 or HDAC2). Chimeric proteins consisting of an engineered DNA-binding domain, most notably an engineered zinc finger domain, and an effector domain are known in the art (see, e.g., Papworth *et al.* (2006), *Gene* 366:27-38).

EXAMPLES

[0164] This invention is further illustrated by the following examples, which should not be construed as limiting. Those skilled in the art will recognize, or be able to ascertain, using no more than routine experimentation, numerous equivalents to the specific substances and procedures described herein. Such equivalents are intended to be encompassed in the scope of the claims that follow the examples below. Examples 1-4 below refer specifically to rationally-designed meganucleases based on I-CreI, but rationally-designed meganucleases based on I-SceI, I-MsoI, I-CeuI, and other LAGLIDADG meganucleases can be similarly produced and used, as described herein.

EXAMPLE 1

Rational Design of Meganucleases Recognizing the HIV-1 TAT Gene

1. Meganuclease Design.

[0165] A pair of meganucleases were designed to recognize and cleave the DNA site 5'-GAAGAGCTCATCAGAACAGTCA-3' (SEQ ID NO: 15) found in the HIV-1 TAT Gene. In accordance with Table 1, two meganucleases, TAT1 and TAT2, were designed to bind the half-sites 5'-GAAGAGCTC-3' (SEQ ID NO: 16) and 5'-TGACTGTTC-3' (SEQ ID NO: 17), respectively, using the following base contacts (non-WT contacts are in bold):

TAT1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>G</u>	<u>A</u>	<u>A</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>N30/ Q38</u>	<u>R40</u>	<u>K28</u>	<u>S26/ R77</u>	<u>K24/ Y68</u>	<u>Q44</u>	<u>R70</u>

TAT2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>G</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>R33</u>	<u>N30/ Q38</u>	<u>R28/ E40</u>	<u>M66</u>	<u>S26/ R77</u>	<u>Y68</u>	<u>Q44</u>	<u>R70</u>

[0166] The two enzymes were cloned, expressed in *E. coli*, and assayed for enzyme activity against the corresponding DNA recognition sequence as described below. In both cases, the rationally-designed meganucleases were found to be inactive. A second generation of each was then produced in which E80 was mutated to Q to improve contacts with the DNA backbone. The second generation TAT2 enzyme was found to be active against its intended recognition sequence while the second generation TAT1 enzyme remained inactive. Visual inspection of the wild-type I-CreI co-crystal structure suggested that TAT1 was inactive due to a steric clash between R40 and K28. To alleviate this clash, TAT1 variants were produced in which K28 was mutated to an amino acid with a smaller side chain (A, S, T, or C) while maintaining the Q80 mutation. When these enzymes were produced in *E. coli* and assayed, the TAT1 variants with S28 and T28 were both found to be active against the intended recognition sequence while maintaining the desired base preference at position -7.

2. Construction of Recombinant Meganucleases.

[0167] Mutations for the redesigned I-CreI enzymes were introduced using mutagenic primers in an overlapping PCR strategy. Recombinant DNA fragments of I-CreI generated in a primary PCR were joined in a secondary PCR to produce full-length recombinant nucleic acids. All recombinant I-CreI constructs were cloned into pET21a vectors with a six histidine tag fused at the 3' end of the gene for purification (Novagen

Corp., San Diego, CA). All nucleic acid sequences were confirmed using Sanger Dideoxynucleotide sequencing (see Sanger *et al.* (1977), *Proc. Natl. Acad. Sci. USA.* 74(12): 5463-7).

[0168] Wild-type I-CreI and all engineered meganucleases were expressed and purified using the following method. The constructs cloned into a pET21a vector were transformed into chemically competent BL21 (DE3) pLysS, and plated on standard 2xYT plates containing 200 µg/ml carbanicillin. Following overnight growth, transformed bacterial colonies were scraped from the plates and used to inoculate 50 ml of 2XYT broth. Cells were grown at 37°C with shaking until they reached an optical density of 0.9 at a wavelength of 600 nm. The growth temperature was then reduced from 37°C to 22°C. Protein expression was induced by the addition of 1mM IPTG, and the cells were incubated with agitation for two and a half hours. Cells were then pelleted by centrifugation for 10 min. at 6000 xg. Pellets were resuspended in 1ml binding buffer (20 mM Tris-HCL, pH 8.0, 500 mM NaCl, 10 mM imidazole) by vortexing. The cells were then disrupted with 12 pulses of sonication at 50% power and the cell debris was pelleted by centrifugation for 15 min. at 14,000 xg. Cell supernatants were diluted in 4 ml binding buffer and loaded onto a 200 µl nickel-charged metal-chelating Sepharose column (Pharmacia).

[0169] The column was subsequently washed with 4 ml wash buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 60 mM imidazole) and with 0.2 ml elution buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 400 mM imidazole). Meganuclease enzymes were eluted with an additional 0.6 ml of elution buffer and concentrated to 50-130 µl using Vivospin disposable concentrators (ISC, Inc., Kaysville, UT). The enzymes were exchanged into SA buffer (25 mM Tris-HCL, pH 8.0, 100 mM NaCl, 5 mM MgCl₂, 5mM EDTA) for assays and storage using Zeba spin desalting columns (Pierce Biotechnology, Inc., Rockford, IL). The enzyme concentration was determined by absorbance at 280 nm using an extinction coefficient of 23,590 M⁻¹cm⁻¹. Purity and molecular weight of the enzymes was then confirmed by MALDI-TOF mass spectrometry.

[0170] Heterodimeric enzymes were produced either by purifying the two proteins independently, and mixing them *in vitro* or by constructing an artificial operon for tandem expression of the two proteins in *E. coli*. In the former case, the purified

meganucleases were mixed 1:1 in solution and pre-incubated at 42°C for 20 minutes prior to the addition of DNA substrate. In the latter case, the two genes were cloned sequentially into the pET-21a expression vector using *NdeI/EcoRI* and *EcoRI/HindIII*. The first gene in the operon ends with two stop codons to prevent read-through errors during transcription. A 12-base pair nucleic acid spacer and a Shine-Dalgarno sequence from the pET21 vector separated the first and second genes in the artificial operon.

3. Cleavage Assays.

[0171] All enzymes purified as described above were assayed for activity by incubation with linear, double-stranded DNA substrates containing the meganuclease recognition sequence. Synthetic oligonucleotides corresponding to both sense and antisense strands of the recognition sequence were annealed and were cloned into the *SmaI* site of the pUC19 plasmid by blunt-end ligation. The sequences of the cloned binding sites were confirmed by Sanger dideoxynucleotide sequencing. All plasmid substrates were linearized with *XmnI*, *ScaI* or *BpmI* concurrently with the meganuclease digest. The enzyme digests contained 5 µl 0.05 µM DNA substrate, 2.5 µl 5 µM recombinant I-CreI meganuclease, 9.5 µl SA buffer, and 0.5 µl *XmnI*, *ScaI*, or *BpmI*. Digests were incubated at either 37°C, or 42°C for certain meganuclease enzymes, for four hours. Digests were stopped by adding 0.3 mg/ml Proteinase K and 0.5% SDS, and incubated for one hour at 37°C. Digests were analyzed on 1.5% agarose and visualized by ethidium bromide staining.

[0172] To evaluate meganuclease half-site preference, rationally-designed meganucleases were incubated with a set of DNA substrates corresponding to a perfect palindrome of the intended half-site as well as each of the 27 possible single-base-pair substitutions in the half-site. In this manner, it was possible to determine how tolerant each enzyme is to deviations from its intended half-site.

4. Recognition Sequence-Specificity.

[0173] Purified recombinant TAT1 and TAT2 meganucleases recognized DNA sequences that were distinct from the wild-type meganuclease recognition sequence (Fig. 2(B)). The wild-type I-CreI meganuclease cleaves the WT recognition sequence, but cuts

neither the intended sequence for TAT1 nor the intended sequence for TAT2. TAT1 and TAT2, likewise, cut their intended recognition sequences but not the wild-type sequence. The meganucleases were then evaluated for half-site preference and overall specificity (Fig 3). Wild-type I-CreI was found to be highly tolerant of single-base-pair substitutions in its natural half-site. In contrast, TAT1 and TAT2 were found to be highly-specific and completely intolerant of base substitutions at positions -1, -2, -3, -6, and -8 in the case of TAT1, and positions -1, -2, and -6 in the case of TAT2.

EXAMPLE 2

Rational Design of Meganucleases with Altered DNA-Binding Affinity

1. Meganucleases with increased affinity and increased activity.

[0174] The meganucleases CCR1 and BRP2 were designed to cleave the half-sites 5'-AACCCTCTC-3' (SEQ ID NO: 18) and 5'-CTCCGGGTC-3' (SEQ ID NO: 19), respectively. These enzymes were produced in accordance with Table 1 as in Example 1:

CCR1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>R28/</u> <u>E40</u>	<u>E42</u>	<u>Q26</u>	<u>K24/</u> <u>Y68</u>	<u>Q44</u>	<u>R70</u>

BRP2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>C33</u>	<u>R30/</u> <u>E38</u>	<u>R28/</u> <u>E40</u>	<u>R42</u>	<u>S26/</u> <u>R77</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0175] Both enzymes were expressed in *E. coli*, purified, and assayed as in Example 1. Both first generation enzymes were found to cleave their intended recognition sequences with rates that were considerably below that of wild-type I-CreI with its natural recognition sequence. To alleviate this loss in activity, the DNA-binding affinity of CCR1 and BRP2 was increased by mutating E80 to Q in both enzymes. These

second-generation versions of CCR1 and BRP2 were found to cleave their intended recognition sequences with substantially increased catalytic rates.

2. Meganucleases with decreased DNA-binding affinity and decreased activity but increased specificity.

[0176] Wild-type I-CreI was found to be highly-tolerant of substitutions to its half-site (Fig. 3(A)). In an effort to make the enzyme more specific, the lysine at position 116 of the enzyme, which normally makes a salt-bridge with a phosphate in the DNA backbone, was mutated to aspartic acid to reduce DNA-binding affinity. This rationally-designed enzyme was found to cleave the wild-type recognition sequence with substantially reduced activity but the recombinant enzyme was considerably more specific than wild-type. The half-site preference of the K116D variant was evaluated as in Example 1 and the enzyme was found to be entirely intolerant of deviation from its natural half-site at positions -1, -2, and -3, and displayed at least partial base preference at the remaining 6 positions in the half-site (Fig. 3(B)).

EXAMPLE 3

Rationally-Designed Meganuclease Heterodimers

1. Cleavage of non-palindromic DNA sites by meganuclease heterodimers formed in solution.

[0177] Two meganucleases, LAM1 and LAM2, were designed to cleave the half-sites 5'-TGCGGTGTC-3' (SEQ ID NO: 20) and 5'-CAGGCTGTC-3' (SEQ ID NO: 21), respectively. The heterodimer of these two enzymes was expected to recognize the DNA sequence 5'-TGCGGTGTCCGGCGACAGCCTG-3' (SEQ ID NO: 22) found in the bacteriophage λ p05 gene.

LAM1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>R33</u>	<u>R30/ E38</u>	<u>D28/ R40</u>	<u>R42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

LAM2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>E30/ R38</u>	<u>R40</u>	<u>K28/ E42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0178] LAM1 and LAM 2 were cloned, expressed in *E. coli*, and purified individually as described in Example 1. The two enzymes were then mixed 1:1 and incubated at 42°C for 20 minutes to allow them to exchange subunits and re-equilibrate. The resulting enzyme solution, expected to be a mixture of LAM1 homodimer, LAM2 homodimer, and LAM1/LAM2 heterodimer, was incubated with three different recognition sequences corresponding to the perfect palindrome of the LAM1 half-site, the perfect palindrome of the LAM2 half-site, and the non-palindromic hybrid site found in the bacteriophage λ genome. The purified LAM1 enzyme alone cuts the LAM1 palindromic site, but neither the LAM2 palindromic site, nor the LAM1/LAM2 hybrid site. Likewise, the purified LAM2 enzyme alone cuts the LAM2 palindromic site but neither the LAM1 palindromic site nor the LAM1/LAM2 hybrid site. The 1:1 mixture of LAM1 and LAM2, however, cleaves all three DNA sites. Cleavage of the LAM1/LAM2 hybrid site indicates that two distinct redesigned meganucleases can be mixed in solution to form a heterodimeric enzyme capable of cleaving a non-palindromic DNA site.

2. Cleavage of non-palindromic DNA sites by meganuclease heterodimers formed by co-expression.

[0179] Genes encoding the LAM1 and LAM2 enzymes described above were arranged into an operon for simultaneous expression in *E. coli* as described in Example 1. The co-expressed enzymes were purified as in Example 1 and the enzyme mixture

incubated with the three potential recognition sequences described above. The co-expressed enzyme mixture was found to cleave all three sites, including the LAM1/LAM2 hybrid site, indicating that two distinct rationally-designed meganucleases can be co-expressed to form a heterodimeric enzyme capable of cleaving a non-palindromic DNA site.

3. Preferential cleavage of non-palindromic DNA sites by meganuclease heterodimers with modified protein-protein interfaces.

[0180] For applications requiring the cleavage of non-palindromic DNA sites, it is desirable to promote the formation of enzyme heterodimers while minimizing the formation of homodimers that recognize and cleave different (palindromic) DNA sites. To this end, variants of the LAM1 enzyme were produced in which lysines at positions 7, 57, and 96 were changed to glutamic acids. This enzyme was then co-expressed and purified as in above with a variant of LAM2 in which glutamic acids at positions 8 and 61 were changed to lysine. In this case, formation of the LAM1 homodimer was expected to be reduced due to electrostatic repulsion between E7, E57, and E96 in one monomer and E8 and E61 in the other monomer. Likewise, formation of the LAM2 homodimer was expected to be reduced due to electrostatic repulsion between K7, K57, and K96 on one monomer and K8 and K61 on the other monomer. Conversely, the LAM1/LAM2 heterodimer was expected to be favored due to electrostatic attraction between E7, E57, and E96 in LAM1 and K8 and K61 in LAM2. When the two meganucleases with modified interfaces were co-expressed and assayed as described above, the LAM1/LAM2 hybrid site was found to be cleaved preferentially over the two palindromic sites, indicating that substitutions in the meganuclease protein-protein interface can drive the preferential formation of heterodimers.

EXAMPLE 4

Additional Meganuclease Heterodimers Which Cleave Physiologic DNA Sequences

1. Meganuclease heterodimers which cleave DNA sequences relevant to gene therapy.

[0181] A rationally-designed meganuclease heterodimer (ACH1/ACH2) can be produced that cleaves the sequence 5'-CTGGGAGTCTCAGGACAGCCTG-3' (SEQ ID NO: 23) in the human FGFR3 gene, mutations in which cause achondroplasia. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

ACH1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>C33</u>	<u>E30/</u> <u>R38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>A26/</u> <u>Q77</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

ACH2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>Y33</u>	<u>E30/</u> <u>R38</u>	<u>R40</u>	<u>K28/</u> <u>E42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0182] A rationally-designed meganuclease heterodimer (HGH1/HGH2) can be produced that cleaves the sequence 5'-CCAGGTGTCTCTGGACTCCTCC-3' (SEQ ID NO: 24) in the promoter of the Human Growth Hormone gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

HGH1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>C33</u>	<u>N30/</u> <u>Q38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

HGH2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>K32</u>	<u>R33</u>	<u>N30/ Q38</u>	<u>R40/ D28</u>	<u>R42</u>	<u>A26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0183] A rationally-designed meganuclease heterodimer (CF1/CF2) can be produced that cleaves the sequence 5'-GAAAATATCATTGGTGTTCCT-3' (SEQ ID NO: 25) in the Δ F508 allele of the human CFTR gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

CF1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>G</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>T</u>	<u>A</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>N30/ Q38</u>	<u>Q40</u>	<u>K28</u>	<u>Q26</u>	<u>H68/ C24</u>	<u>Q44</u>	<u>R70</u>

CF2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>A</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>R33</u>	<u>E30/ R38</u>	<u>Q40</u>	<u>K28</u>	<u>A26</u>	<u>Y68/ C24</u>	<u>T44</u>	<u>R70</u>

[0184] A rationally-designed meganuclease heterodimer (CCR1/CCR2) can be produced that cleaves the sequence 5'-AACCCCTCTCCAGTGAGATGCCT-3' (SEQ ID NO: 26) in the human CCR5 gene (an HIV co-receptor). For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

CCR1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>E40/</u> <u>R28</u>	<u>E42</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

CCR2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	<u>A</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>R33</u>	<u>E30/</u> <u>R38</u>	<u>E40</u>	<u>K28</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

[0185] A rationally-designed meganuclease heterodimer (MYD1/MYD2) can be produced that cleaves the sequence 5'-GACCTCGTCCTCCGACTCGCTG-3' (SEQ ID NO: 27) in the 3' untranslated region of the human DM kinase gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

MYD1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>G</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>E40/</u> <u>R28</u>	<u>K66</u>	<u>Q26/</u> <u>E77</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

MYD1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>E30/</u> <u>R38</u>	<u>E40/</u> <u>R28</u>	<u>R42</u>	<u>A26</u> <u>Q77</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

2. Meganuclease heterodimers which cleave DNA sequences in pathogen genomes.

[0186] A rationally-designed meganuclease heterodimer (HSV1/HSV2) can be produced that cleaves the sequence 5'-CTCGATGTCGGACGACACGGCA-3' (SEQ ID NO: 28) in the UL36 gene of Herpes Simplex Virus-1 and Herpes Simplex Virus-2. For

example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

HSV1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>C33</u>	<u>R30/</u> <u>E38</u>	<u>R40/</u>	<u>Q42/</u> <u>K28</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

HSV2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>R33</u>	<u>R30/</u> <u>E38</u>	<u>E40/</u> <u>R28</u>	<u>R42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0187] A rationally-designed meganuclease heterodimer (ANT1/ANT2) can be produced that cleaves the sequence 5'-ACAAGTGTCTATGGACAGTTTA-3' (SEQ ID NO: 29) in the *Bacillus anthracis* genome. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

ANT1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>C</u>	<u>A</u>	<u>A</u>	<u>G</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>C33</u>	<u>N30/</u> <u>Q38</u>	<u>Q40/</u> <u>A28</u>	<u>R42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

ANT2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>Y33</u>	<u>N30/</u> <u>Q38</u>	<u>Q40</u>	<u>E42</u>	<u>Q26</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0188] A rationally-designed meganuclease heterodimer (POX1/POX2) can be produced that cleaves the sequence 5'-AAACTGTCAAATGACATCGCA-3' (SEQ ID NO: 30) in the Variola (smallpox) virus gp009 gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

POX1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>C33</u>	<u>N30/ Q38</u>	<u>Q40</u>	<u>K28</u>	<u>Q26</u>	<u>R68</u>	<u>O44</u>	<u>R70</u>

POX2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>T</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>R33</u>	<u>R30/ E38</u>	<u>R40</u>	<u>C28/ Q42</u>	<u>Q26</u>	<u>R68</u>	<u>O44</u>	<u>R70</u>

[0189] A rationally-designed meganuclease homodimer (EBB1/EBB1) can be produced that cleaves the pseudo-palindromic sequence 5'-CGGGTCTCGTGCGAGGCCTCC-3' (SEQ ID NO: 31) in the Epstein-Barr Virus *BALF2* gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

EBB1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>R33</u>	<u>D30/ Q38</u>	<u>R40/ D28</u>	<u>R42</u>	<u>Q26</u>	<u>Y68/ K24</u>	<u>O44</u>	<u>R70</u>

EBB1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>R33</u>	<u>D30/</u> <u>Q38</u>	<u>R40/</u> <u>D28</u>	<u>R42</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

3. Meganuclease heterodimers which cleave DNA sequences in plant genomes.

[0190] A rationally-designed meganuclease heterodimer (GLA1/GLA2) can be produced that cleaves the sequence 5'-CACTAACTCGTATGAGTCGGTG-3' (SEQ ID NO: 32) in the *Arabidopsis thaliana* GL2 gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

GLA1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>A</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>S40/</u> <u>C79</u>	<u>K28</u>	<u>A26/</u> <u>Q77</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

GLA2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>Y33</u>	<u>R30/</u> <u>E38</u>	<u>E40/</u> <u>R28</u>	<u>R42</u>	<u>A26</u> <u>Q77</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

[0191] A rationally-designed meganuclease heterodimer (BRP1/BRP2) can be produced that cleaves the sequence 5'-TGCCTCCTCTAGAGACCCGGAG-3' (SEQ ID NO: 33) in the *Arabidopsis thaliana* BP1 gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

BRP1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>R33</u>	<u>R30/</u> <u>E38</u>	<u>R28/</u> <u>E40</u>	<u>K66</u>	<u>Q26/</u> <u>E77</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

BRP2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>T</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>C33</u>	<u>R30/</u> <u>E38</u>	<u>E40/</u> <u>R28</u>	<u>R42</u>	<u>S26</u> <u>R77</u>	<u>R68</u>	<u>Q44</u>	<u>R70</u>

[0192] A rationally-designed meganuclease heterodimer (MGC1/MGC2) can be produced that cleaves the sequence 5'-TAAATCTCTAAGGTCTGTGCA-3' (SEQ ID NO: 34) in the *Nicotiana tabacum* Magnesium Chelatase gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

MGC1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>A</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>C32</u>	<u>Y33</u>	<u>N30/</u> <u>Q38</u>	<u>Q40/</u>	<u>K28</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

MGC2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>A</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>A</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>R33</u>	<u>R30/</u> <u>E38</u>	<u>Q40</u>	<u>K28</u>	<u>A26</u> <u>Q77</u>	<u>R68</u>	<u>T44</u>	<u>R70</u>

[0193] A rationally-designed meganuclease heterodimer (CYP/HGH2) can be produced that cleaves the sequence 5'-CAAGAATTCAAGCGAGCATTAA-3' (SEQ ID NO: 35) in the *Nicotiana tabacum* CYP82E4 gene. For example, a meganuclease was

designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

CYP:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>C</u>	<u>A</u>	<u>A</u>	<u>G</u>	<u>A</u>	<u>A</u>	<u>T</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>D32</u>	<u>Y33</u>	<u>N30/ Q38</u>	<u>R40/</u>	<u>K28</u>	<u>Q77/ A26</u>	<u>Y68</u>	<u>Q44</u>	<u>R70</u>

HGH2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>T</u>	<u>A</u>	<u>A</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>C33</u>	<u>N30/ Q38</u>	<u>Q40</u>	<u>K66</u>	<u>R77/ S26</u>	<u>Y68</u>	<u>Q44</u>	<u>R70</u>

4. Meganuclease heterodimers which cleave DNA sequences in yeast genomes.

[0194] A rationally-designed meganuclease heterodimer (URA1/URA2) can be produced that cleaves the sequence 5'-TTAGATGACAAGGGAGACGCAT-3' (SEQ ID NO: 36) in the *Saccharomyces cerevisiae* URA3 gene. For example, a meganuclease was designed based on the I-CreI meganuclease, as described above, with the following contact residues and recognition sequence half-sites:

URA1:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>T</u>	<u>T</u>	<u>A</u>	<u>G</u>	<u>A</u>	<u>T</u>	<u>G</u>	<u>A</u>	<u>C</u>
<u>Contact Residues</u>	<u>S32</u>	<u>C33</u>	<u>N30/ Q38</u>	<u>R40</u>	<u>K28</u>	<u>Q26</u>	<u>R68</u>	<u>T44</u>	<u>R70</u>

URA2:

<u>Position</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<u>Base</u>	<u>A</u>	<u>T</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>
<u>Contact Residues</u>	<u>N32</u>	<u>C33</u>	<u>E30/</u> <u>R38</u>	<u>E40/</u> <u>R28</u>	<u>R42</u>	<u>Q26</u>	<u>Y68/</u> <u>K24</u>	<u>Q44</u>	<u>R70</u>

5. Recognition Sequence Specificity.

[0195] The rationally-designed meganucleases outlined above in this Example were cloned, expressed in *E. coli*, and purified as in Example 1. Each purified meganuclease was then mixed 1:1 with its corresponding heterodimerization partner (*e.g.*, ACH1 with ACH2, HGH1 with HGH2, etc.) and incubated with a linearized DNA substrate containing the intended non-palindromic DNA recognition sequence for each meganuclease heterodimer. As shown in Figure 3, each rationally-designed meganuclease heterodimer cleaves its intended DNA site.

SEQUENCE LISTING

SEQ ID NO: 1 (wild-type I-CreI, Genbank Accession # PO5725)

1 MNTKYNKEFL LYLAGEFVDGD GSIIAQIKPN QSYKFKHQLS LAFQVTQKTQ RRWFLDKLVLD
 61 EIGVGYVRDR GSVSDYILSE IKPLHNFLTQ LQPFLKLLKQK QANLVLKIIW RLPSAKESPD
 121 KFLEVCTWVD QIAALNDSKT RKTTSSETVRA VLDSLSEKKK SSP

SEQ ID NO: 2 (wild-type I-CreI recognition sequence)

1 GAAACTGTCT CACGACGTTT TG

SEQ ID NO: 3 (wild-type I-CreI recognition sequence)

1 GAAAACGTCG TGAGACAGTT TC

SEQ ID NO: 4 (wild-type I-CreI recognition sequence)

1 CAAACTGTCTG TGAGACAGTT TG

SEQ ID NO: 5 (wild-type I-CreI recognition sequence)

1 CAAACTGTCT CACGACAGTT TG

SEQ ID NO: 6 (wild-type I-MsoI, Genbank Accession # AAL34387)

1 MTTKNTLQPT EAAYIAGFLD GDGSIYAKLI PRPDYKDIKY QVSLAISFIQ RKDKFPYLQD
 61 IYDQLGKRGV LRKDRGDGIA DYTIIGSTHL SILLPDLVPY LRIKKKQANR ILHIINLYPQ
 121 AQKNPSKFLD LVKIVDDVQN LNKRADELKS TNYDRLLLEEF LKAGKIESSP

SEQ ID NO: 7 (wild-type I-MsoI, recognition sequence)

1 CAGAACGTCG TGAGACAGTT CC

SEQ ID NO: 8 (wild-type I-MsoI, recognition sequence)

1 GGAACTGTCT CACGACGTTT TG

SEQ ID NO: 9 (wild-type I-SceI, Genbank Accession # CAA09843)

1 MKNIKKNQVM NLGPNKLLK EYKSQLELNL IEQFEAGIGL ILGDAYIRSR DEGKTYCMQF
 61 EWKNKAYMDH VCLLYDQWVL SPPHKKERVN HLGNLVITWG AQTFKHQAFN KLANLFIVNN
 121 KKTIPNNLVE NYLTPMSLAY WFMDGGKWD YNKNSTNKSI VLNTQSFTFE EVEYLVKGLR
 181 NKFQLNCYVK INKNKPIIYI DSMSYLIFYN LIKPYLIPQM MYKLPNTISS ETFLK

SEQ ID NO: 10 (wild-type I-SceI, recognition sequence)

1 TTACCCTGTT ATCCCTAG

SEQ ID NO: 11 (wild-type I-SceI, recognition sequence)

1 CTAGGGATAA CAGGGTAA

SEQ ID NO: 12 (wild-type I-CeuI, Genbank Accession # P32761)

1 MSNFILKPGE KLPQDKLEEL KKINDAVKKT KNFSKYLIDL RKLFDIDEVQ VTSESKLFLA
61 GFLEGEASLN ISTKKLATS K FGLVVDPEFN VTQHVNGVKV LYLALFVFKT GRIRHKSGSN
121 ATLVLTIDNR QSLEEKVLPF YEQYVAFSS PEKVKRVANF KALLELFNND AHQDLEQLVN
181 KILPIWDQMR KQQGQSNEGF PNLEAAQDFA RNYKKGIK

SEQ ID NO: 13 (wild-type I-CeuI, recognition sequence)

1 ATAACGGTCC TAAGGTAGCG AA

SEQ ID NO: 14 (wild-type I-CeuI, recognition sequence)

1 TTCGCTACCT TAGGACCGTT AT

SEQ ID NO: 15 (HIV-1 TAT gene, partial sequence)

1 GAAGAGCTCA TCAGAACAGT CA

SEQ ID NO: 16 (rationally-designed TAT1 recognition sequence half-site)

1 GAAGAGCTC

SEQ ID NO: 17 (rationally-designed TAT2 recognition sequence half-site)

1 TGACTGTTC

SEQ ID NO: 18 (rationally-designed CCR1 recognition sequence half-site)

1 AACCCCTCTC

SEQ ID NO: 19 (rationally-designed BRP2 recognition sequence half-site)

1 CTCCGGGTC

SEQ ID NO: 20 (rationally-designed LAM1 recognition sequence half-site)

1 TGCGGTGTC

SEQ ID NO: 21 (rationally-designed LAM2 recognition sequence half-site)

1 CAGGCTGTC

SEQ ID NO: 22 (LAM1/LAM2 recognition sequence in bacteriophage λ p05 gene)

1 TCCGGTGTCC GGCGACAGCC TG

SEQ ID NO: 23 (potential recognition sequence in human FGFR3 gene)

1 CTGGGAGTCT CAGGACAGCC TG

SEQ ID NO: 24 (potential recognition sequence in human growth hormone promoter)

1 CCAGGTGTCT CTGGACTCCT CC

SEQ ID NO: 25 (potential recognition sequence in human CFTR gene Δ F508 allele)

1 GAAAATATCA TTGGTGTTTC CT

SEQ ID NO: 26 (potential recognition sequence in human CCR5 gene)

1 AACCTCTCC AGTGAGATGC CT

SEQ ID NO: 27 (potential recognition sequence in human DM kinase gene 3' UTR)

1 GACCTCGTCC TCCGACTCGC TG

SEQ ID NO: 28 (potential recognition sequence in Herpes Simplex Virus-1 and Herpes Simplex Virus-2 UL36 gene)

1 CTCGATGTCC GACGACACGG CA

SEQ ID NO: 29 (potential recognition sequence in *Bacillus anthracis* genome)

1 ACAAGTGTCT ATGGACAGTT TA

SEQ ID NO: 30 (potential recognition sequence in the Variola (smallpox) virus gp009 gene)

1 AAAACTGTCA AATGACATCG CA

SEQ ID NO: 31 (potential recognition sequence in the Epstein-Barr Virus *BALF2* gene)

1 CGGGGTCTCG TGCGAGGCCT CC

SEQ ID NO: 32 (potential recognition sequence in the *Arabidopsis thaliana* GL2 gene)

1 CACTAACTCG TATGAGTCGG TG

SEQ ID NO: 33 (potential recognition sequence in the *Arabidopsis thaliana* BP1 gene)

1 TGCCTCCTCT AGAGACCCGG AG

SEQ ID NO: 34 (potential recognition sequence in the *Nicotiana tabacum* Magnesium Chelatase gene)

1 TAAAATCTCT AAGGTCTGTG CA

SEQ ID NO: 35 (potential recognition sequence in the *Nicotiana tabacum* CYP82E4 gene)

1 CAAGAATTCA AGCGAGCATT AA

SEQ ID NO: 36 (potential recognition sequence in the *Saccharomyces cerevisiae* URA3 gene)

1 TTAGATGACA AGGGAGACGC AT

CLAIMS

1. A recombinant meganuclease having altered specificity for at least one recognition sequence half-site relative to a wild-type I-CreI meganuclease, comprising:
 - a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1; and
 - having specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-CreI meganuclease recognition sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5;
 - wherein said recombinant meganuclease comprises at least one modification of Table 1 which is not an excluded modification.

2. A recombinant meganuclease having altered specificity for at least one recognition sequence half-site relative to a wild-type I-MsoI meganuclease, comprising:
 - a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6; and
 - having specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-MsoI meganuclease recognition sequence selected from the group consisting of SEQ ID NO: 7 and SEQ ID NO: 8;
 - wherein said recombinant meganuclease comprises at least one modification of Table 2 which is not an excluded modification.

3. A recombinant meganuclease having altered specificity for a recognition sequence relative to a wild-type I-SceI meganuclease, comprising:
 - a polypeptide having at least 85% sequence similarity to residues 3-186 of the I-SceI meganuclease of SEQ ID NO: 9; and
 - having specificity for a recognition sequence which differs by at least one base pair from an I-SceI meganuclease recognition sequence of SEQ ID NO: 10 and SEQ ID NO: 11;

wherein said recombinant meganuclease comprises at least one modification of Table 3 which is not an excluded modification.

4. A recombinant meganuclease having altered specificity for at least one recognition sequence half-site relative to a wild-type I-CeuI meganuclease, comprising:
a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12; and

having specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-CeuI meganuclease recognition sequence selected from the group consisting of SEQ ID NO: 13 and SEQ ID NO: 14;

wherein said recombinant meganuclease comprises at least one modification of Table 4 which is not an excluded modification.

5. A recombinant meganuclease having altered specificity for at least one recognition sequence half-site relative to a wild-type I-CreI meganuclease, comprising:
a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1; and

having specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-CreI meganuclease recognition sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 and SEQ ID NO: 5;

wherein:

(1) specificity at position -1 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of Q70, C70, L70, Y75, Q75, H75, H139, Q46 and H46;

(b) to an A on a sense strand by a modification selected from the group consisting of Y75, L75, C75, Y139, C46 and A46;

(c) to a G on a sense strand by a modification selected from the group consisting of K70, E70, E75, E46 and D46;

(d) to a C on a sense strand by a modification selected from the group consisting of H75, R75, H46, K46 and R46; or

(e) to any base on a sense strand by a modification selected from the group consisting of G70, A70, S70 and G46; and/or

(2) specificity at position -2 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of Q70, T44, A44, V44, I44, L44, and N44;

(b) to a C on a sense strand by a modification selected from the group consisting of E70, D70, K44 and R44;

(c) to a G on a sense strand by a modification selected from the group consisting of H70, D44 and E44; or

(d) to an A or T on a sense strand by a modification comprising C44; and/or

(3) specificity at position -3 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of Q68 and C24;

(b) to a C on a sense strand by a modification selected from the group consisting of E68, F68, K24 and R24;

(c) to a T on a sense strand by a modification selected from the group consisting of M68, C68, L68 and F68;

(d) to an A or C on a sense strand by a modification comprising H68;

(e) to a C or T on a sense strand by a modification comprising Y68; or

(f) to a G or T on a sense strand by a modification comprising K68; and/or

(4) specificity at position -4 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of E77 and K26;

(b) to a G on a sense strand by a modification selected from the group consisting of E26 and R77;

(c) to a C or T on a sense strand by a modification comprising S77; or

(d) to a any base on a sense strand by a modification comprising S26;

and/or

(5) specificity at position -5 has been altered:

(a) to a C on a sense strand by a modification comprising E42;

- (b) to a G on a sense strand by a modification comprising R42;
 - (c) to an A or G on a sense strand by a modification selected from the group consisting of C28 and Q42; or
 - (d) to any base on a sense strand by a modification of selected from the group consisting of M66 and K66; and/or
- (6) specificity at position -6 has been altered:
- (a) to a T on a sense strand by a modification selected from the group consisting of C40, I40, V40, C79, I79, V79, and Q28;
 - (b) to a C on a sense strand by a modification selected from the group consisting of E40 and R28; or
 - (c) to a G on a sense strand by a modification comprising R40; and/or
- (7) specificity at position -7 has been altered:
- (a) to a C on a sense strand by a modification selected from the group consisting of E38, K30 and R30;
 - (b) to a G on a sense strand by a modification selected from the group consisting of K38, R38 and E30;
 - (c) to a T on a sense strand by a modification selected from the group consisting of I38 and L38; or
 - (d) to an A or G on a sense strand by a modification comprising C38; or
 - (e) to any base on a sense strand by a modification selected from the group consisting of H38, N38 and Q30; and/or
- (8) specificity at position -8 has been altered:
- (a) to a T on a sense strand by a modification selected from the group consisting of L33, V33, I33, F33 and C33;
 - (b) to a C on a sense strand by a modification selected from the group consisting of E33 and D33;
 - (c) to a G on a sense strand by a modification consisting of K33;
 - (d) to an A or C on a sense strand by a modification comprising R32; or
 - (e) to an A or G on a sense strand by a modification comprising R33;
- and/or
- (9) specificity at position -9 has been altered:

- (a) to a C on a sense strand by a modification comprising E32;
- (b) to a G on a sense strand by a modification selected from the group consisting of R32 and K32;
- (c) to a T on a sense strand by a modification selected from the group consisting of L32, V32, A32 and C32;
- (d) to a C or T on a sense strand by a modification selected from the group consisting of D32 and I32; or
- (e) to any base on a sense strand by a modification selected from the group consisting of S32, N32, H32, Q32 and T32.

6. A recombinant meganuclease having altered specificity for at least one recognition sequence half-site relative to a wild-type I-MsoI meganuclease, comprising:
a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6; and

having specificity for a recognition sequence half-site which differs by at least one base pair from a half-site within an I-MsoI meganuclease recognition sequence selected from the group consisting of SEQ ID NO: 7 and SEQ ID NO: 8;

wherein:

- (1) specificity at position -1 has been altered:
 - (a) to an A on a sense strand by a modification selected from the group consisting of K75, Q77, A49, C49 and K79;
 - (b) to a T on a sense strand by a modification selected from the group consisting of C77, L77 and Q79; or
 - (c) to a G on a sense strand by a modification selected from the group consisting of K77, R77, E49 and E79; and/or
- (2) specificity at position -2 has been altered:
 - (a) to an A on a sense strand by a modification selected from the group consisting of Q75, K81, C47, I 47 and L47;
 - (b) to a C on a sense strand by a modification selected from the group consisting of E75, D75, R47, K47, K81 and R81; or
 - (c) to a G on a sense strand by a modification selected from the group

consisting of K75, E47 and E81; and/or

(3) specificity at position -3 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of Q72, C26, L26, V26, A26 and I26;

(b) to a C on a sense strand by a modification selected from the group consisting of E72, Y72, H26, K26 and R26; or

(c) to a T on a sense strand by a modification selected from the group consisting of K72, Y72 and H26; and/or

(4) specificity at position -4 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of K28, K83 and Q28;

(b) to a G on a sense strand by a modification selected from the group consisting of R83 and K83; or

(c) to an A on a sense strand by a modification selected from the group consisting of K28 and Q83; and/or

(5) specificity at position -5 has been altered:

(a) to a G on a sense strand by a modification selected from the group consisting of R45 and E28;

(b) to a T on a sense strand by a modification comprising Q28; or

(c) to a C on a sense strand by a modification comprising R28; and/or

(6) specificity at position -6 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of K43, V85, L85 and Q30;

(b) to a C on a sense strand by a modification selected from the group consisting of E43, E85, K30 and R30; or

(c) to a G on a sense strand by a modification selected from the group consisting of R43, K43, K85, R85, E30 and D30; and/or

(7) specificity at position -7 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of E32 and E41;

(b) to a G on a sense strand by a modification selected from the group

consisting of R32, R41 and K41;

(c) to a T on a sense strand by a modification selected from the group consisting of K32, M41, L41 and I41; and/or

(8) specificity at position -8 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of K32 and K35;

(b) to a C on a sense strand by a modification comprising E32; or

(c) to a G on a sense strand by a modification consisting of K32, K35 and R35; and/or

(9) specificity at position -9 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of N34 and H34;

(b) to a T on a sense strand by a modification selected from the group consisting of S34, C34, V34, T34 and A34; or

(c) to a G on a sense strand by a modification selected from the group consisting of K34, R34 and H34.

7. A recombinant meganuclease having altered specificity for a recognition sequence relative to a wild-type I-SceI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 3-186 of the I-SceI meganuclease of SEQ ID NO: 9; and

having specificity for a recognition sequence which differs by at least one base pair from an I-SceI meganuclease recognition sequence of SEQ ID NO: 10 and SEQ ID NO: 11;

wherein:

(1) specificity at position 4 has been altered:

(a) to an A on a sense strand by a modification comprising K50;

(b) to a T on a sense strand by a modification selected from the group consisting of K57, M57 and Q50; or

(c) to a G on a sense strand by a modification selected from the group consisting of E50, R57 and K57; and/or

(2) specificity at position 5 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of K48, Q102;

(b) to a G on a sense strand by a modification selected from the group consisting of E48, K102 and R102; or

(c) to a T on a sense strand by a modification selected from the group consisting of Q48, C102, L102 and V102; and/or

(3) specificity at position 6 has been altered:

(a) to an A on a sense strand by a modification comprising K59;

(b) to a C on a sense strand by a modification selected from the group consisting of R59 and K59; or

(b) to a G on a sense strand by a modification selected from the group consisting of K84 and E59; and/or

(4) specificity at position 7 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of R46, K46 and E86;

(b) to a G on a sense strand by a modification selected from the group consisting of K86, R86 and E46; or

(c) to an A on a sense strand by a modification selected from the group consisting of C46, L46 and V46; and/or

(5) specificity at position 8 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of E88, R61 and H61;

(b) to a T on a sense strand by a modification selected from the group consisting of K88, Q61 and H61; or

(c) to an A on a sense strand by a modification selected from the group consisting of K61, S61, V61, A61 and L61; and/or

(6) specificity at position 9 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of C98, V98 and L98;

(b) to a C on a sense strand by a modification selected from the group

consisting of R98 and K98; or

(c) to a G on a sense strand by a modification selected from the group consisting of E98 and D98; and/or

(7) specificity at position 10 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of K96 and R96;

(b) to a G on a sense strand by a modification selected from the group consisting of D96 and E96; or

(c) to an A on a sense strand by a modification selected from the group consisting of C96 and A96; and/or

(8) specificity at position 11 has been altered:

(a) to a T on a sense strand by a modification comprising Q90;

(b) to a C on a sense strand by a modification selected from the group consisting of K90 and R90; or

(c) to a G on a sense strand by a modification comprising E90; and/or

(9) specificity at position 12 has been altered:

(a) to an A on a sense strand by a modification comprising Q193;

(b) to a C on a sense strand by a modification selected from the group consisting of E165, E193 and D193; or

(c) to a G on a sense strand by a modification selected from the group consisting of K165 and R165; and/or

(10) specificity at position 13 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of Q193, C163 and L163;

(b) to a G on a sense strand by a modification selected from the group consisting of E193, D193, K163 and R192; or

(c) to an A on a sense strand by a modification selected from the group consisting of C193 and L193; and/or

(11) specificity at position 14 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of K161 and Q192;

(b) to an A on a sense strand by a modification selected from the group consisting of L192 and C192;

(c) to a G on a sense strand by a modification selected from the group consisting of K147, K161, R161, R197, D192 and E192; or

(d) to a T on a sense strand by a modification selected from the group consisting of K161 and Q192; and/or

(12) specificity at position 15 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of C151, L151 and K151;

(b) to a G on a sense strand by a modification comprising K151; or

(c) to a C on a sense strand by a modification comprising E151; and/or

(13) specificity at position 17 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of G152 and Q150;

(b) to a C on a sense strand by a modification selected from the group consisting of K152 and K150; or

(c) to a G on a sense strand by a modification selected from the group consisting of N152, S152, D152, D150 and E150; and/or

(14) specificity at position 18 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of H155 and Y155;

(b) to a C on a sense strand by a modification selected from the group consisting of R155 and K155; or

(c) to an A on a sense strand by a modification selected from the group consisting of K155 and C155.

8. A recombinant meganuclease having altered specificity for at least one recognition sequence half-site relative to a wild-type I-CeuI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12; and

having specificity for a recognition sequence half-site which differs by at least one

base pair from a half-site within an I-CeuI meganuclease recognition sequence selected from the group consisting of SEQ ID NO: 13 and SEQ ID NO: 14;

wherein:

(1) specificity at position -1 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of C92, A92 and V92;

(b) to a T on a sense strand by a modification selected from the group consisting of Q116 and Q92; or

(c) to a G on a sense strand by a modification selected from the group consisting of E116 and E92; and/or

(2) specificity at position -2 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of Q117, C90, L90 and V90;

(b) to a G on a sense strand by a modification selected from the group consisting of K117, R124, K124, E124, E90 and D90; or

(c) to a C on a sense strand by a modification selected from the group consisting of E117, D117, R174, K124, K90, R90 and K68; and/or

(3) specificity at position -3 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of C70, V70, T70, L70 and K70;

(b) to a T on a sense strand by a modification comprising Q70;

(b) to a C on a sense strand by a modification consisting of K70; and/or

(4) specificity at position -4 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of E126, D126, R88, K88 and K72;

(b) to a T on a sense strand by a modification selected from the group consisting of K126, L126 and Q88; or

(c) to an A on a sense strand by a modification selected from the group consisting of Q126, N126, K88, L88, C88, C72, L72 and V72; and/or

(5) specificity at position -5 has been altered:

(a) to a G on a sense strand by a modification selected from the group

consisting of E74, K128, R128 and E128;

(b) to a T on a sense strand by a modification selected from the group consisting of C128, L128, V128 and T128; or

(c) to an A on a sense strand by a modification selected from the group consisting of C74, L74, V74 and T74; and/or

(6) specificity at position -6 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of K86, C86 and L86;

(b) to a C on a sense strand by a modification selected from the group consisting of D86, E86, R84 and K84; or

(c) to a G on a sense strand by a modification selected from the group consisting of K128, R128, R86, K86 and E84; and/or

(7) specificity at position -7 has been altered:

(a) to a C on a sense strand by a modification selected from the group consisting of R76, K76 and H76;

(b) to a G on a sense strand by a modification selected from the group consisting of E76 and R84; or

(c) to a T on a sense strand by a modification consisting of H76 and Q76;
and/or

(8) specificity at position -8 has been altered:

(a) to an A on a sense strand by a modification selected from the group consisting of Y79, R79 and Q76;

(b) to a C on a sense strand by a modification selected from the group consisting of D79, E79, D76 and E76; or

(c) to a G on a sense strand by a modification selected from the group consisting of R79, K79, K76 and R76; and/or

(9) specificity at position -9 has been altered:

(a) to a T on a sense strand by a modification selected from the group consisting of K78, V78, L78, C78 and T78;

(b) to a C on a sense strand by a modification selected from the group consisting of D78 and E78; or

(c) to a G on a sense strand by a modification selected from the group consisting of R78, K78 and H78.

9. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-CreI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1;

wherein DNA-binding affinity has been increased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of E80, D137, I81, L112, P29, V64 or Y66 with H, N, Q, S, T, K or R; or

(b) substitution of T46, T140 or T143 with K or R.

10. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-CreI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1;

wherein DNA-binding affinity has been decreased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of K34, K48, R51, K82, K116 or K139 with H, N, Q, S, T, D or E; or

(b) substitution of I81, L112, P29, V64, Y66, T46, T140 or T143 with D or E.

11. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-MsoI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6;

wherein DNA-binding affinity has been increased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of E147, I85, G86 or Y118 with H, N, Q, S, T, K or R; or

(b) substitution of Q41, N70, S87, T88, H89, Q122, Q139, S150 or N152 with K

or R.

12. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-MsoI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6;

wherein DNA-binding affinity has been decreased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of K36, R51, K123, K143 or R144 with H, N, Q, S, T, D or E; or

(b) substitution of I85, G86, Y118, Q41, N70, S87, T88, H89, Q122, Q139, S150 or N152 with D or E.

13. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-SceI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 3-186 of the I-SceI meganuclease of SEQ ID NO: 9;

wherein DNA-binding affinity has been increased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of D201, L19, L80, L92, Y151, Y188, I191, Y199 or Y222 with H, N, Q, S, T, K or R; or

(b) substitution of N15, N17, S81, H84, N94, N120, T156, N157, S159, N163, Q165, S166, N194 or S202 with K or R.

14. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-SceI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 3-186 of the I-SceI meganuclease of SEQ ID NO: 9;

wherein DNA-binding affinity has been decreased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of K20, K23, K63, K122, K148, K153, K190, K193, K195 or K223 with H, N, Q, S, T, D or E; or

(b) substitution of L19, L80, L92, Y151, Y188, I191, Y199, Y222, N15, N17, S81, H84, N94, N120, T156, N157, S159, N163, Q165, S166, N194 or S202 with D or E.

15. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-CeuI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12;

wherein DNA-binding affinity has been increased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of D25 or D128 with H, N, Q, S, T, K or R; or

(b) substitution of S68, N70, H94, S117, N120, N129 or H172 with K or R.

16. A recombinant meganuclease having altered binding affinity for double-stranded DNA relative to a wild-type I-CeuI meganuclease, comprising:

a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12;

wherein DNA-binding affinity has been decreased by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of K21, K28, K31, R112, R114 or R130 with H, N, Q, S, T, D or E; or

(b) substitution of S68, N70, H94, S117, N120, N129 or H172 with D or E.

17. A recombinant meganuclease monomer having altered affinity for dimer formation with a reference meganuclease monomer, comprising:

a polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1;

wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of K7, K57 or K96 with D or E; or

(b) substitution of E8 or E61 with K or R.

18. A recombinant meganuclease heterodimer comprising:
a first polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1;
wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:
(a) substitution of K7, K57 or K96 with D or E; and
a second polypeptide having at least 85% sequence similarity to residues 2-153 of the I-CreI meganuclease of SEQ ID NO: 1;
wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:
(b) substitution of E8 or E61 with K or R.
19. A recombinant meganuclease monomer having altered affinity for dimer formation with a reference meganuclease monomer, comprising:
a polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6;
wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:
(a) substitution of R302 with D or E; or
(b) substitution of D20, E11 or Q64 with K or R.
20. A recombinant meganuclease heterodimer comprising:
a first polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6;
wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:
(a) substitution of R302 with D or E; and
a second polypeptide having at least 85% sequence similarity to residues 6-160 of the I-MsoI meganuclease of SEQ ID NO: 6;
wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:

(b) substitution of D20, E11 or Q64 with K or R.

21. A recombinant meganuclease monomer having altered affinity for dimer formation with a reference meganuclease monomer, comprising:

a polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12;

wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of R93 with D or E; or

(b) substitution of E152 with K or R.

22. A recombinant meganuclease heterodimer comprising:

a first polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12;

wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:

(a) substitution of R93 with D or E; and

a second polypeptide having at least 85% sequence similarity to residues 5-211 of the I-CeuI meganuclease of SEQ ID NO: 12;

wherein affinity for dimer formation has been altered by at least one modification corresponding to a substitution selected from the group consisting of:

(b) substitution of E152 with K or R.

23. A recombinant meganuclease monomer or heterodimer of claim 17 or claim 18, further comprising:

at least one modification selected from Table 1;

wherein said modification is not an excluded modification.

24. A recombinant meganuclease monomer or heterodimer of claim 19 or claim 20, further comprising:

at least one modification selected from Table 2;

wherein said modification is not an excluded modification.

25. A recombinant meganuclease monomer or heterodimer of claim 21 or claim 22, further comprising:

at least one modification selected from Table 4;

wherein said modification is not an excluded modification.

26. A method for producing a genetically-modified eukaryotic cell including an exogenous sequence of interest inserted in a chromosome of said eukaryotic cell, comprising:

transfecting a eukaryotic cell with one or more nucleic acids including

(i) a first nucleic acid sequence encoding a meganuclease, and

(ii) a second nucleic acid sequence including said sequence of interest;

wherein said meganuclease produces a cleavage site in said chromosome and said sequence of interest is inserted into said chromosome at said cleavage site; and

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25.

27. A method as in claim 26 wherein:

said second nucleic acid further comprises sequences homologous to sequences flanking said cleavage site and said sequence of interest is inserted at said cleavage site by homologous recombination.

28. A method as in claim 26 wherein:

said second nucleic acid lacks substantial homology to said cleavage site and said sequence of interest is inserted into said chromosome by non-homologous end-joining.

29. A method for producing a genetically-modified eukaryotic cell including an exogenous sequence of interest inserted in a chromosome of said eukaryotic cell, comprising:

introducing a meganuclease protein into a eukaryotic cell; and

transfecting said eukaryotic cell with a nucleic acid including said sequence of interest;

wherein said meganuclease produces a cleavage site in said chromosome and said sequence of interest is inserted into said chromosome at said cleavage site; and

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25.

30. A method as in claim 29 wherein:

said nucleic acid further comprises sequences homologous to sequences flanking said cleavage site and said sequence of interest is inserted at said cleavage site by homologous recombination.

31. A method as in claim 29 wherein:

said nucleic acid lacks substantial homology to said cleavage site and said sequence of interest is inserted into said chromosome by non-homologous end-joining.

32. A method for producing a genetically-modified eukaryotic cell by disrupting a target sequence in a chromosome of said eukaryotic cell, comprising:

transfecting a eukaryotic cell with a nucleic acid encoding a meganuclease;
wherein said meganuclease produces a cleavage site in said chromosome and said target sequence is disrupted by non-homologous end-joining at said cleavage site; and
wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25.

33. A method of producing a genetically-modified organism comprising:

producing a genetically-modified eukaryotic cell according to the method of any one of claims 26-32; and
growing said genetically-modified eukaryotic cell to produce said genetically-modified organism.

34. A method as in claim 33 wherein:
said eukaryotic cell is selected from the group consisting of a gamete, a zygote, a blastocyst cell, an embryonic stem cell, and a protoplast cell.
35. A method for treating a disease by gene therapy in a eukaryote, comprising:
transfecting at least one cell of said eukaryote with one or more nucleic acids including
- (i) a first nucleic acid sequence encoding a meganuclease, and
 - (ii) a second nucleic acid sequence including a sequence of interest;
- wherein said meganuclease produces a cleavage site in said chromosome and said sequence of interest is inserted into said chromosome at said cleavage site;
wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25; and
wherein insertion of said sequence of interest provides said gene therapy for said disease.
36. A method as in claim 35 wherein:
said second nucleic acid sequence further comprises sequences homologous to sequences flanking said cleavage site and said sequence of interest is inserted at said cleavage site by homologous recombination.
37. A method as in claim 35 wherein:
said second nucleic acid sequence lacks substantial homology to said cleavage site and said sequence of interest is inserted into said chromosome by non-homologous end-joining.
38. A method for treating a disease by gene therapy in a eukaryote, comprising:
introducing a meganuclease protein into at least one cell of said eukaryote; and
transfecting said eukaryotic cell with a nucleic acid including a sequence of interest;

wherein said meganuclease produces a cleavage site in said chromosome and said sequence of interest is inserted into said chromosome at said cleavage site;

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25; and

wherein insertion of said sequence of interest provides said gene therapy for said disease.

39. A method as in claim 38 wherein:

said nucleic acid further comprises sequences homologous to sequences flanking said cleavage site and said sequence of interest is inserted at said cleavage site by homologous recombination.

40. A method as in claim 38 wherein:

said nucleic acid lacks substantial homology to said cleavage site and said sequence of interest is inserted into said chromosome by non-homologous end-joining.

41. A method for treating a disease by gene therapy in a eukaryote by disrupting a target sequence in a chromosome of said eukaryotic cell, comprising:

transfecting at least one cell of said eukaryote with a nucleic acid encoding a meganuclease;

wherein said meganuclease produces a cleavage site in said chromosome and said target sequence is disrupted by non-homologous end-joining at said cleavage site;

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25; and

wherein disruption of said target sequence provides said gene therapy for said disease.

42. A method for treating a viral pathogen infection in a eukaryotic host by disrupting a target sequence in a genome of said viral pathogen, comprising:

transfecting at least one infected cell of said eukaryotic host with a nucleic acid encoding a meganuclease;

wherein said meganuclease produces a cleavage site in said viral genome and said target sequence is disrupted by non-homologous end-joining at said cleavage site;

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25; and

wherein disruption of said target sequence provides treatment for said infection.

43. A method for treating a viral pathogen infection in a eukaryotic host by disrupting a target sequence in a genome of said viral pathogen, comprising:

transfecting at least one infected cell of said eukaryotic host with a first nucleic acid encoding a meganuclease and a second nucleic acid;

wherein said meganuclease produces a cleavage site in said viral genome and said target sequence is disrupted by homologous recombination of said viral genome and said second nucleic acid at said cleavage site;

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25;

wherein said second nucleic acid comprises sequences homologous to sequences flanking said cleavage site; and

wherein disruption of said target sequence provides treatment for said infection.

44. A method for treating a prokaryotic pathogen infection in a eukaryotic host by disrupting a target sequence in a genome of said prokaryotic pathogen, comprising:

transfecting at least cell of said prokaryotic pathogen infecting said eukaryotic host with a nucleic acid encoding a meganuclease;

wherein said meganuclease produces a cleavage site in said prokaryotic genome and said target sequence is disrupted by non-homologous end-joining at said cleavage site;

wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25; and

wherein disruption of said target sequence provides treatment for said infection.

45. A method for treating a prokaryotic pathogen infection in a eukaryotic host by disrupting a target sequence in a genome of said prokaryotic pathogen, comprising:
transfecting at least cell of said prokaryotic pathogen infecting said eukaryotic host with a first nucleic acid encoding a meganuclease and a second nucleic acid;
wherein said meganuclease produces a cleavage site in said prokaryotic genome and said target sequence is disrupted by homologous recombination of said prokaryotic genome and said second nucleic acid at said cleavage site;
wherein said meganuclease is a recombinant meganuclease of any one of claims 1-25;
wherein said second nucleic acid comprises sequences homologous to sequences flanking said cleavage site; and
wherein disruption of said target sequence provides treatment for said infection.

46. A method for rationally-designing a recombinant meganuclease having altered specificity for at least one base position of a recognition sequence, said altered specificity including at least one desired change, the method comprising:

(1) determining at least a portion of a three-dimensional structure of a reference meganuclease-DNA complex;

(2) identifying amino acid residues forming a base contact surface at said base position;

(3) determining a distance between a β -carbon of at least a first residue comprising said contact surface and at least a first base at said base position;

(4) identifying an amino acid substitution to promote said desired change by

(a) for a first residue which is $< 6 \text{ \AA}$ from said first base, selecting a substitution from Group 1 and/or Group 2 which is a member of an appropriate one of Group G, Group C, Group T or Group A to promote said desired change; or

(b) for a first residue which is $> 6 \text{ \AA}$ from said first base, selecting a substitution from Group 2 and/or Group 3 which is a member of an appropriate one of Group G, Group C, Group T or Group A to promote said desired change.

47. A method for rationally-designing a recombinant meganuclease having increased DNA-binding affinity, comprising:

- (1) determining at least a portion of a three-dimensional structure of a reference meganuclease-DNA complex;
- (2) identifying amino acid contact residues forming a backbone contact surface;
- (3) identifying an amino acid substitution to increase said DNA-binding affinity by
 - (a) for a contact residue having a negatively-charged or hydrophobic side chain, selecting a substitution having an uncharged/polar or positively-charged side chain;or
 - (b) for a contact residue having an uncharged/polar side chain, selecting a substitution having a positively-charged side chain.

48. A method for rationally-designing a recombinant meganuclease having decreased DNA-binding affinity, comprising:

- (1) determining at least a portion of a three-dimensional structure of a reference meganuclease-DNA complex;
- (2) identifying amino acid contact residues forming a backbone contact surface;
- (3) identifying an amino acid substitution to decrease said DNA-binding affinity by
 - (a) for a contact residue having a positively-charged side chain, selecting a substitution having an uncharged/polar or negatively-charged side chain; or
 - (b) for a contact residue having an hydrophobic or uncharged/polar side chain, selecting a substitution having a negatively-charged side chain.

Abstract: Rationally-designed LAGLIDADG meganucleases and methods of making such meganucleases are provided. In addition, methods are provided for using the meganucleases to generate recombinant cells and organisms having a desired DNA sequence inserted into a limited number of loci within the genome, as well as methods of gene therapy, for treatment of pathogenic infections, and for in vitro applications in diagnostics and research.

【 図 1 A 】

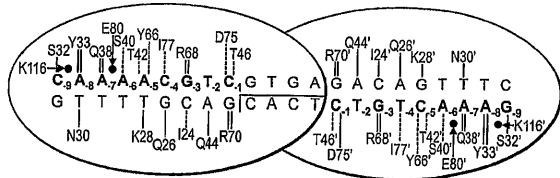


FIG. 1A

【 図 1 B 】

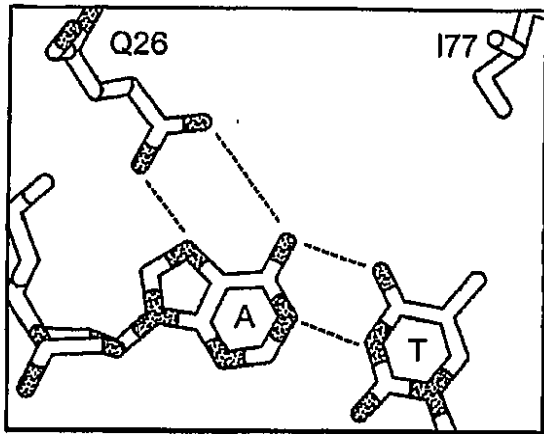


FIG. 1B

【 図 1 C 】

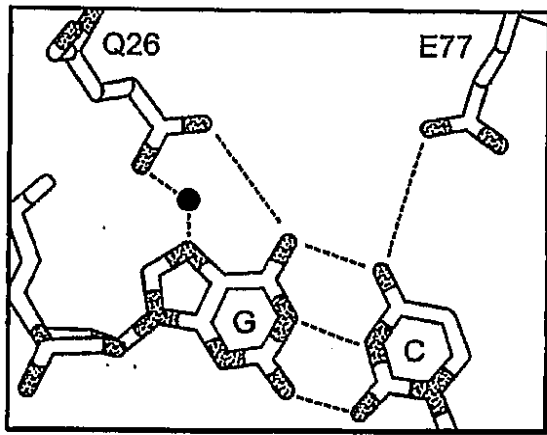


FIG. 1C

【 図 1 D 】

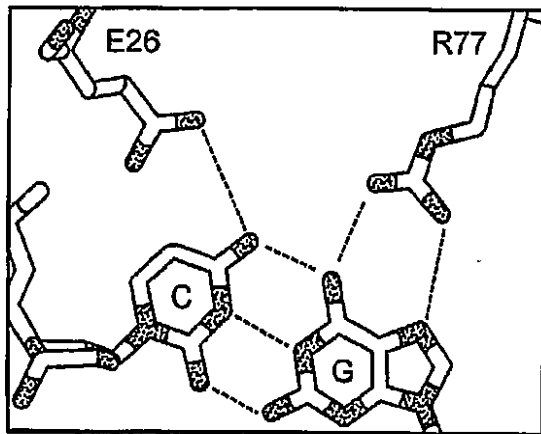


FIG. 1D

【 図 1 E 】

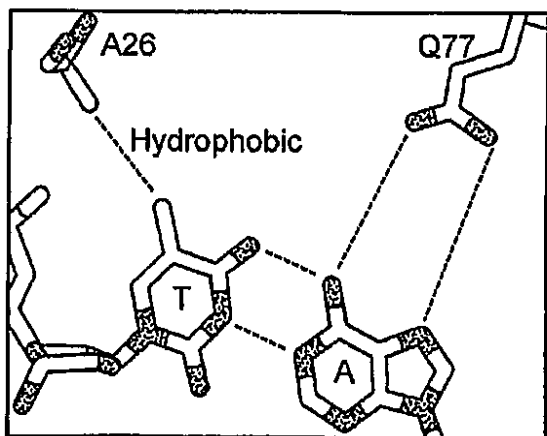


FIG. 1E

【 2 A 】

WT CAAACTGTCGTGAGACAGTTTG
 CF1/CF2 GAAAATATTCATTGTTTTCCT
 MYD1/MYD2 GACCTCGTCCCTCCGACTCGCTG
 CCR1/CCR2 AACCCCTCTCCAGTGAAGATGCCT
 ACH1/ACH2 CTGGGAGTCTCAGGACAGCCTG
 TAT1/TAT2 GAGAGCTCAATCAGAACAGTCA
 HSV1/HSV2 CTCGATGTCGACGACACGGCA
 LAM1/LAM2 TGGGGTGTCCGCGGACAGCCTG
 POX1/POX2 AAACTGTCAAAATGACATCGCA
 URA1/URA2 TTAGATGACAAAGGAGACGCAT
 GLA1/GLA2 CACTAACTCGTATGAGTCCGTTG
 BRP1/BRP2 TGCCTCCCTATAGACCCGGAG

FIG. 2A

【 2 B 】

Enzyme Site	CF		CCR		TAT		LAM		URA		BRP		
	WT	0	MYD	0	ACH	0	HSV	0	POX	0			GLA
WT	100	0	0	0	0	0	47	0	100	100	0	0	Wild Type
CF1/CF2	0	63	0	0	0	0	0	0	0	0	0	0	Cystic Fibrosis
MYD1/MYD2	0	0	94	0	0	0	16	0	0	0	0	0	Myotonic Dystrophy
CCR1/CCR2	0	0	0	100	0	0	0	0	0	0	0	0	CCR5
ACH1/ACH2	0	0	0	0	87	0	0	0	0	0	0	0	Achondroplasia
TAT1/TAT2	0	0	0	0	0	81	0	0	0	0	0	0	HIV-1
HSV1/HSV2	0	0	0	0	0	0	67	0	0	0	0	0	Herpes Simplex 1
LAM1/LAM2	0	0	0	0	0	0	0	87	0	0	0	0	Lambda Phage
POX1/POX2	100	0	0	0	0	0	0	0	100	0	0	0	Variola Virus
URA1/URA2	0	0	0	0	0	0	0	0	0	70	0	0	<i>S. cerevisiae</i> URA3
GLA1/GLA2	0	0	0	0	0	0	0	0	0	0	88	0	<i>A. thaliana</i> GL2
BRP1/BRP2	0	0	0	0	0	0	0	0	0	0	0	100	<i>A. thaliana</i> BP1

FIG. 2B

【 3 】

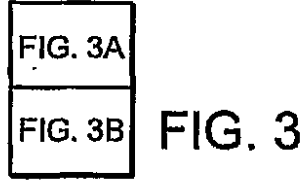


FIG. 3

【 3 A 】

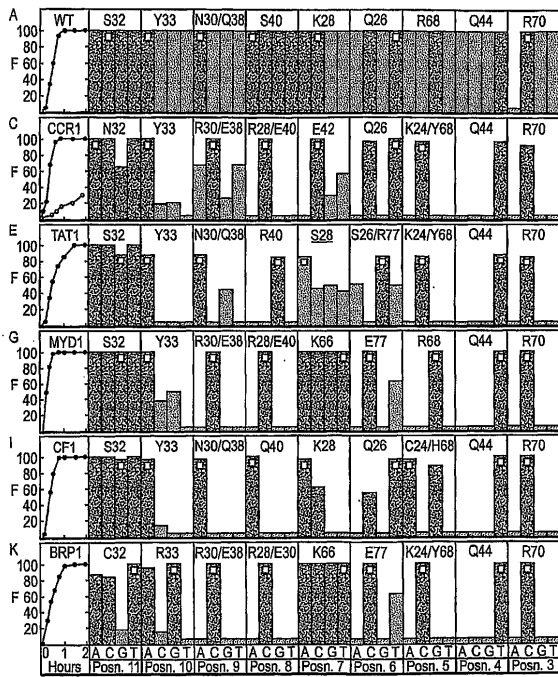


FIG. 3A

【 3 B 】

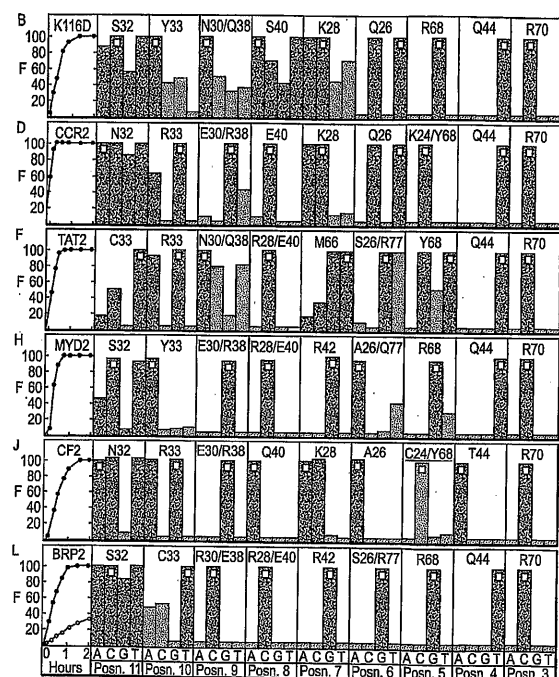


FIG. 3B