

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2009-509535

(P2009-509535A)

(43) 公表日 平成21年3月12日 (2009.3.12)

(51) Int. Cl.	F I	テーマコード (参考)
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 Z C C A	4 B O 2 4
C 1 2 P 21/02 (2006.01)	C 1 2 P 21/02 Z N A C	4 B O 6 4
C 4 O B 40/10 (2006.01)	C 4 O B 40/10	4 C O 8 4
C O 7 K 14/005 (2006.01)	C O 7 K 14/005	4 H O 4 5
A 6 1 K 38/00 (2006.01)	A 6 1 K 37/02	

審査請求 未請求 予備審査請求 未請求 (全 163 頁) 最終頁に続く

(21) 出願番号	特願2008-533574 (P2008-533574)	(71) 出願人	508091410
(86) (22) 出願日	平成18年9月27日 (2006.9.27)		アムニクス, インコーポレイテッド
(85) 翻訳文提出日	平成20年5月21日 (2008.5.21)		アメリカ合衆国 カリフォルニア 940
(86) 国際出願番号	PCT/US2006/037713		43, マウンテン ビュー, エリス
(87) 国際公開番号	W02007/038619		ストリート 500 - スイート ビー
(87) 国際公開日	平成19年4月5日 (2007.4.5)	(74) 代理人	100078282
(31) 優先権主張番号	60/721, 270		弁理士 山本 秀策
(32) 優先日	平成17年9月27日 (2005.9.27)	(74) 代理人	100062409
(33) 優先権主張国	米国 (US)		弁理士 安村 高明
(31) 優先権主張番号	60/721, 188	(74) 代理人	100113413
(32) 優先日	平成17年9月27日 (2005.9.27)		弁理士 森下 夏樹
(33) 優先権主張国	米国 (US)	(72) 発明者	ステマー, ウィレム ピー. シー.
(31) 優先権主張番号	60/743, 622		アメリカ合衆国 カリフォルニア 950
(32) 優先日	平成18年3月21日 (2006.3.21)		30, ロス ガトス, キャシー コー
(33) 優先権主張国	米国 (US)		ト 108

最終頁に続く

(54) 【発明の名称】 タンパク様薬剤およびその使用

(57) 【要約】

本発明は、システイン含有スカフォールドおよび／またはタンパク、および、そのようなシステイン含有産物を担持するか、および／または発現する発現ベクター、宿主細胞、および提示システムを提供する。本発明はさらに、そのような産物のライブラリーの設計法、そのようなライブラリーをスクリーニングして、標的分子に対し結合特異性を示す実体を生成する方法を提供する。さらに、本発明によって、本発明のシステイン含有産物を含む製薬組成物が提供される。

【特許請求の範囲】

【請求項 1】

35以下のアミノ酸を有するポリペプチドを含む、非天然システイン(C)－含有タンパクであって、

前記ポリペプチドのアミノ酸の少なくとも10%がシステインであり、

スカフォールド内のシステインを対合させることによって、少なくとも二つのジスルフィド結合が形成され、かつ前記対合は、3を超える複合指数を生成する、非天然システイン(C)－含有タンパク。

【請求項 2】

約60個以下のアミノ酸を有するポリペプチドを含む、非天然システイン(C)－含有タンパクであって、

前記ポリペプチドのアミノ酸の少なくとも10%はシステインであり、

ポリペプチドに含まれるシステインを対合させることによって、少なくとも四つのジスルフィド結合が形成され、かつ前記対合は、4を超える複合指数を生成する、非天然システイン(C)－含有タンパク。

【請求項 3】

複合指数が6よりも大きい、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 4】

複合指数が10よりも大きい、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 5】

標的分子に特異的に結合する、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 6】

約50℃よりも高い温度で加熱された後も標的結合能を維持する、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 7】

約80℃よりも高い温度で加熱された後も標的結合能を維持する、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 8】

約100℃よりも高い温度で、かつ0.1秒を超える間、加熱された後も標的結合能を維持する、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 9】

標識類、エフェクター類、抗体類、および半減期延長成分から成る群から選ばれる成分に接合される、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 10】

モノマーである、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 11】

マルチマーである、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 12】

前記タンパクが一タイプのスカフォールドを含む、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 13】

前記タンパクが一を超えるタイプのスカフォールドを含む、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 14】

前記タンパクが、標的結合部位および半減期延長成分を含む、請求項1または2に記載の非天然システイン(C)－含有タンパク。

【請求項 15】

前記タンパクが、標的に結合する反復単位を含む、請求項 1 または 2 に記載の非天然システイン (C) - 含有タンパク。

【請求項 16】

前記タンパクが、血清アルブミン、IgG、赤血球、および血清に接触可能なタンパクから成る群から選ばれる半減期延長成分を含む、請求項 1 または 2 に記載の非天然システイン (C) - 含有タンパク。

【請求項 17】

対応する天然システイン (C) - 含有タンパクまたはスカフォールドの天然標的とは異なる標的に対して特異的結合性を示す、非天然システイン (C) - 含有タンパク。

【請求項 18】

20-60 個のアミノ酸から成る単ドメインを含む非天然タンパクであって、3 個以上のジスルフィドを有し、ヒトの血清暴露タンパクに結合し、かつ 5 % 未満の脂肪族アミノ酸を有する、前記非天然タンパク。

【請求項 19】

20-60 個のアミノ酸から成る単ドメインを含む非天然タンパクであって、3 個以上のジスルフィドを有し、ヒトの血清暴露タンパクに結合し、かつ T-エピトーププログラムにおいてデータベースのタンパクの平均の 90 % よりも低いスコアを有する、前記非天然タンパク。

【請求項 20】

請求項 1、2、18、または 19 に記載の非天然タンパクから成るライブラリー。

【請求項 21】

請求項 20 のライブラリーを提示する遺伝子パッケージ。

【請求項 22】

標的と、遺伝子パッケージの上に提示される外来ポリペプチドとの間の特異的相互作用の存在を検出する方法であって、

(a) 請求項 20 の遺伝子パッケージの提示を準備する工程、

(b) 安定なポリペプチド-標的複合体を生成するのに好適な条件下で遺伝子パッケージを標的と接触させる工程、および

(c) 遺伝子パッケージにおいて安定なポリペプチド-標的複合体の形成を検出し、それによって特異的相互作用の存在を検出する工程、

を含む、前記方法。

【請求項 23】

所望の性質を有するポリペプチドを提示する遺伝子パッケージを単離する工程をさらに含む、請求項 22 に記載の方法。

【請求項 24】

請求項 1 または 2 に記載の非天然システイン (C) - 含有タンパク、および、製薬学的に受容可能な担体を含む製薬組成物。

【請求項 25】

標的分子に対して結合特異性を示す非天然システイン (C) - 含有スカフォールドであって、C¹⁻², 3-4、C¹⁻³, 2-4、および C¹⁻⁴, 2-3 から成る群から選ばれるパターンに従って、スカフォールド内システインを対合させることによって形成される二つのジスルフィド結合を有するポリペプチドを含み、前式において、ハイフンで結ばれる二つの数字は、ポリペプチドの N 末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す、前記非天然システイン (C) - 含有スカフォールド。

【請求項 26】

標的分子に対して結合特異性を示す非天然システイン (C) - 含有スカフォールドであって、

10

20

30

40

【化 1】

$C^{1-2, 3-4, 5-5}, C^{1-2, 3-5, 4-6}, C^{1-2, 3-6, 4-5}, C^{1-3, 2-4, 5-6}, C^{1-3, 2-5, 4-6}, C^{1-3, 2-6, 4-5}, C^{1-4, 2-3, 5-6}, C^{1-4, 2-6, 3-5}, C^{1-5, 2-3, 4-6}, C^{1-5, 2-4, 3-6}, C^{1-5, 2-6, 3-4}, C^{1-6, 2-3, 4-5}$, および $C^{1-6, 2-5, 3-4}$

から成る群から選ばれるパターンに従って、スカフォールド内システインを対合させることによって形成される三つのジスルフィド結合を有するポリペプチドを含み、前式において、ハイフンで結ばれる二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す、前記非天然システイン(C)－含有スカフォールド。

【請求項 27】

標的分子に対して結合特異性を示す非天然システイン(C)－含有スカフォールドであって、下記：

【化 2】

1-2 3-4 5-6 7-8 1-2 3-4 5-7 6-8 1-2 3-4 5-8 6-7 1-2 3-5 4-6 7-8 1-2 3-5 4-7 6-8 1-2 3-5 4-8 6-7
 1-2 3-6 4-5 7-8 1-2 3-6 4-7 5-8 1-2 3-6 4-8 5-7 1-2 3-7 4-5 6-8 1-2 3-7 4-6 5-8 1-2 3-7 4-8 5-6
 1-2 3-8 4-5 6-7 1-2 3-8 4-6 5-7 1-2 3-8 4-7 5-6 1-3 2-4 5-6 7-8 1-3 2-4 5-7 6-8 1-3 2-4 5-8 6-7
 1-3 2-5 4-6 7-8 1-3 2-5 4-7 6-8 1-3 2-5 4-8 6-7 1-3 2-6 4-5 7-8 1-3 2-6 4-7 5-8 1-3 2-6 4-8 5-7
 1-3 2-7 4-5 6-8 1-3 2-7 4-6 5-8 1-3 2-7 4-8 5-6 1-3 2-8 4-5 6-7 1-3 2-8 4-6 5-7 1-3 2-8 4-7 5-6
 1-4 2-3 5-6 7-8 1-4 2-3 5-7 6-8 1-4 2-3 5-8 6-7 1-4 2-5 3-6 7-8 1-4 2-5 3-7 6-8 1-4 2-5 3-8 6-7
 1-4 2-6 3-5 7-8 1-4 2-6 3-7 5-8 1-4 2-6 3-8 5-7 1-4 2-7 3-5 6-8 1-4 2-7 3-6 5-8 1-4 2-7 3-8 5-6
 1-4 2-8 3-5 6-7 1-4 2-8 3-6 5-8 1-4 2-8 3-7 5-6 1-5 2-3 4-6 7-8 1-5 2-3 4-7 6-8 1-5 2-3 4-8 6-7
 1-5 2-4 3-6 7-8 1-5 2-4 3-7 6-8 1-5 2-4 3-8 6-7 1-5 2-6 3-4 7-8 1-5 2-6 3-7 4-8 1-5 2-6 3-8 4-7
 1-5 2-7 3-4 6-8 1-5 2-7 3-6 4-8 1-5 2-7 3-8 4-6 1-5 2-8 3-4 4-7 1-5 2-8 3-6 4-7 1-5 2-8 3-7 4-6
 1-6 2-3 4-5 7-8 1-6 2-3 4-7 5-8 1-6 2-3 4-8 5-7 1-6 2-4 3-5 7-8 1-6 2-4 3-7 5-8 1-6 2-4 3-8 5-7
 1-6 2-5 3-4 7-8 1-6 2-5 3-7 4-8 1-6 2-5 3-8 4-7 1-6 2-7 3-4 5-8 1-6 2-7 3-5 4-8 1-6 2-7 3-8 4-5
 1-6 2-8 3-4 5-7 1-6 2-8 3-5 4-7 1-6 2-8 3-7 4-5 1-7 2-3 4-5 6-8 1-7 2-3 4-6 5-8 1-7 2-3 4-8 5-6
 1-7 2-4 3-5 6-8 1-7 2-4 3-6 5-8 1-7 2-4 3-8 5-6 1-7 2-5 3-4 6-8 1-7 2-5 3-6 4-8 1-7 2-5 3-8 4-6
 1-7 2-6 3-4 5-8 1-7 2-6 3-5 4-8 1-7 2-6 3-8 4-5 1-7 2-8 3-4 5-6 1-7 2-8 3-5 4-6 1-7 2-8 3-6 4-5
 1-8 2-3 4-5 6-7 1-8 2-3 4-6 5-7 1-8 2-3 4-7 5-6 1-8 2-4 3-5 6-7 1-8 2-4 3-6 5-7 1-8 2-4 3-7 5-6
 1-8 2-5 3-4 6-7 1-8 2-5 3-6 4-7 1-8 2-5 3-7 4-6 1-8 2-6 3-4 5-7 1-8 2-6 3-5 4-7 1-8 2-6 3-7 4-5
 1-8 2-7 3-4 5-6 1-8 2-7 3-5 4-6 1-8 2-7 3-6 4-5

から選ばれるパターンに従って、スカフォールド内システインを対合させることによって形成される少なくとも四つのジスルフィド結合を有するポリペプチドを含み、

前式において、ハイフンで結ばれる二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す、前記非天然システイン(C)－含有スカフォールド。

【請求項 28】

約 50℃ よりも高い温度で加熱された後も標的結合能を維持する、請求項 25、26、または 27 に記載の非天然システイン(C)－含有スカフォールド。

【請求項 29】

約 80℃ よりも高い温度で加熱された後も標的結合能を維持する、請求項 25、26、または 27 に記載の非天然システイン(C)－含有スカフォールド。

【請求項 30】

約 100℃ よりも高い温度で、かつ 0.1 秒を超える間、加熱された後も標的結合能を維持する、請求項 25、26、または 27 に記載の非天然システイン(C)－含有スカフォールド。

【請求項 31】

標識類、エフェクター類、および抗体類から成る群から選ばれる成分に接合される、請求項 25、26、または 27 に記載の非天然システイン (C) - 含有スカフォールド。

【請求項 32】

モノマーである、請求項 25、26、または 27 に記載の非天然システイン (C) - 含有スカフォールド。

【請求項 33】

半減期延長成分を含む、請求項 25、26、または 27 に記載の非天然システイン (C) - 含有スカフォールド。

【請求項 34】

前記半減期延長成分が、血清アルブミン、IgG、赤血球、および、血清に接触可能なタンパクから成る群から選ばれる、請求項 33 に記載の非天然システイン (C) - 含有スカフォールド。

10

【請求項 35】

対応する天然システイン (C) - 含有タンパクまたはスカフォールドの天然標的とは異なる標的に対して特異的結合性を示す、請求項 25、26、または 27 に記載の非天然システイン (C) - 含有スカフォールド。

【請求項 36】

請求項 25、26、または 27 に記載の非天然システイン (C) - 含有スカフォールドから成るライブラリー。

【請求項 37】

請求項 36 のライブラリーを提示する遺伝子パッケージ。

20

【請求項 38】

標的と、遺伝子パッケージの上に提示される外来ポリペプチドとの間の特異的相互作用の存在を検出する方法であって、

(d) 請求項 37 の遺伝子パッケージの提示を準備する工程、

(e) 安定なポリペプチド-標的複合体を生成するのに好適な条件下で遺伝子パッケージを標的と接触させる工程、および

(f) 遺伝子パッケージにおいて安定なポリペプチド-標的複合体の形成を検出し、それによって特異的相互作用の存在を検出する工程、

を含む、前記方法。

30

【請求項 39】

所望の性質を有するポリペプチドを提示する遺伝子パッケージを単離する工程をさらに含む、請求項 38 に記載の方法。

【請求項 40】

前記遺伝子パッケージがファージである、請求項 37 に記載の方法。

【請求項 41】

前記ファージが繊維状ファージである、請求項 36 に記載の方法。

【請求項 42】

非天然システイン (C) - 含有スカフォールドを生産する方法であって、

請求項 25-27 のいずれか一項に記載の非天然システイン (C) - 含有スカフォールドをコードする核酸を含む宿主細胞を準備する工程、

40

前記核酸から、前記スカフォールドの発現を実現する条件下で、前記宿主細胞を適切な培養媒体において培養する工程、

を含む、前記方法。

【請求項 43】

前記媒体から前記スカフォールドを回収する工程をさらに含む、請求項 38 に記載の方法。

【請求項 44】

請求項 25、26、または 27 に記載の前記非天然システイン (C) - 含有スカフォールド、および製薬学的に受容可能な担体を含む製薬組成物。

50

【発明の詳細な説明】

【技術分野】

【0001】

(相互参照)

本願は、双方とも2005年9月27日に出願された米国仮特許出願第60/721, 270号および米国仮特許出願第60/721, 188号、ならびに2006年3月21日に出願された米国仮特許出願第60/743, 622号の利益を請求し、それらはその全体が本明細書中に参考として援用される。

【背景技術】

【0002】

(発明の背景)

分子生物学の基本概念の一つは、各天然タンパク質が、単一の「天然」構造またはフォールドを採用していることである。天然のフォールド以外の任意のフォールドの採用は、「ミスフォールド形成」と見なされる。多数の天然の、機能的フォールドを採用する天然タンパクの例はほとんど、または全く存在しない。ミスフォールド形成は、プリオンの感染性によって例示されるように重大な問題である。プリオンでは、その「誤った」フォールドが、触媒的に他のプリオンタンパクがミスフォールド形成する原因となり、脳の疾患およびある場合には死へと導く。ほとんど全てのタンパク質が、変性されると、ミスフォールドを生成し、繊維状ポリマーを形成するが、これらは、いくつかの変性疾患に関与しているようである。一例として、アルツハイマー病に関与するベータアミロイド繊維がある。タンパクのミスフォールド形成は、不溶の凝集体の非可逆的形成をもたらすが、変性タンパクは、溶融球状物としても出現することがある。多種多様な不安定構造を探る溶融球状状態から、タンパクは、漏斗型経路を辿り、次第にフォールド形成中間体の多様性を減らしていき、最終的に、単一の、安定的に折りたたまれた天然の構造が実現されるまで進むと考えられている。天然タンパクは、アロステリック調整、他のドメインに対するドメインの相対的なリッド／フラップ型移動、リガンドに対する結合による誘発的適合、または、結晶化力によって構造的な変更が可能であるが、これらの変更は一般に、基礎フォールドにおける基本的変化ではなくヒンジ様構造の移動を含む。従来観察された例は全て、天然タンパクは、その生物機能を実行するために、単一の安定なフォールドを採用するように進化してきており、この天然構造からの逸脱は有害であるとする考えを支持する。

【0003】

同じタンパク配列でありながら（選択的スプライシング、糖化、またはタンパク分解酵素処理によって創出される変種を除く）天然に、一つを超える形として存在する例がこれまで二三見つかっているが、第2形は、通常単純にジスルフィド結合を失った不活性な副産物である（Schulz ら、, 2005; Peterson ら、, 2003; Lauber et al, 2003）。高いジスルフィド密度を有する小型タンパクを含むミクロタンパク・ファミリーにおいて（多くは、毒素および受容体ドメインである）、完全に形成されてはいるものの（単純に欠陥的ではない）、別形のジスルフィド結合パターンを採用するために異なる構造を取る、互いに緊密に関連する配列が見いだされている。実例は、ソマトメジン（Kamikubo et al, 2004）およびマウロトキシン（Fajloun et al, 2000）である。

【0004】

従来、タンパク提示ライブラリーは、非特許文献2に総覧されるように、単一の固定化されたタンパクフォールド、例えば、免疫グロブリンの各種ドメイン、インターフェロン、タンパクA、アンキリン、Aドメイン、T細胞受容体、フィブロネクチンIII、ガンマークリスタリン、ユビキチン、その他諸々のフォールドを用いていた。ある場合、例えば、ヒトの免疫レパートリーから得られた免疫グロブリンライブラリーでは、ある単一ライブラリーは、多様なV領域配列をスカフォールドとして用いるが、それらの配列は全て、基本的免疫グロブリンフォールドを共有する。別種ライブラリーとしてランダムペ

プチドまたは環状ペプチドライブラリーがあるが、これらは、定められたフォールドを持たず、かつ、単一の安定構造を取らないので、タンパクとは見なされない。

【0005】

例えば、一つ以上の所望の性質を示す製薬を創出するための指向性進化による合理的選択を受け容れることが可能な新規タンパク構造の設計に対しては相当な需要が依然として存在する。上記の所望の性質として、免疫原性の低下、安定性の強化または半減期の延長、多重特異性、多価性、および標的に対する高い結合アフィニティーが挙げられるが、ただしこれらに限定されない。

【非特許文献1】Binz., Aら、Nature Biotechnology (2005) 23: 1257

【発明の開示】

【課題を解決するための手段】

【0006】

(発明の要旨)

本発明の一局面は、高いジスルフィド密度を示す新規タンパク構造の設計である。このタンパク構造は、例えば、一つ以上の所望の性質を示す製薬を創出するための指向性進化による合理的設計および選択に対し特に受容的である。そのような所望の性質としては、標的に対する高い結合アフィニティーおよび／またはアビディティー、分子量の低下および組織浸透性の向上、熱およびプロテアーゼ安定性の強化、半減期の延長、疎水性の向上、処方強化（特に、高濃度）、および免疫原性の低下が挙げられるが、ただしこれらに限定されない。

【0007】

一実施態様において、本発明は、例えば、スカフォールドの形を取る各種タンパク構造、および、そのようなタンパク構造から成るライブラリーを提供する。一局面では、スカフォールドは、多様なフォールド、または、他の非一次的構造を示す。別の局面では、スカフォールドは、生物機能を実行するために定義された空間配置を有する。別の実施態様では、本発明は、このようなタンパク構造から成るライブラリーを構築する方法、そのようなライブラリーを、遺伝学的ベヒクルまたはパッケージ（例えば、ファージなどのウィルスパッケージ、および非ウィルスパッケージ（例えば、酵母ディスプレイ、大腸菌表面ディスプレイ、リボソームディスプレイ、またはCIS（DNA-連結）ディスプレイ）において提示する方法と共に、そのようなライブラリーをスクリーニングして製薬または製薬候補を生成する方法を提供する。本発明はさらに、主題のタンパク構造を発現するか、または利用するベクター、宿主細胞、および、その他のインビトロ系を提供する。

【0008】

別の実施態様では、本発明は、標的分子に対し結合特異性を示す、非天然のシステイン（C）-含有スカフォールドを提供する。この実施態様において、非天然システイン（C）-含有スカフォールドは、式：Error! Objects cannot be created from editing field codes. によって表される組み合わせの群から選ばれるパターンによるスカフォールド内システインを含む。式中、nは、システイン残基によって形成されるジスルフィド結合の予測数に等しく、かつ、Error! Objects cannot be created from editing field codes. は、 $(2i-1)$ の積を表し、この式においてiは、1からnまでの範囲の正の整数である。

【0009】

別の実施態様では、本発明は、35以下のアミノ酸を有するポリペプチドを含む、非天然システイン（C）-含有タンパクを提供する。前記ポリペプチドのアミノ酸の少なくとも10%はシステインであり、スカフォールド内のシステインを対合させることによって、少なくとも二つのジスルフィド結合が形成され、前記対合は、3を超える複合指数を生成する。

【0010】

10

20

30

40

50

一局面では、非天然システイン（C）－含有タンパクは、約60個以下のアミノ酸を有するポリペプチドを含んでもよく、前記ポリペプチドのアミノ酸の少なくとも10%はシステインであり、ポリペプチドに含まれるシステインを対合させることによって、少なくとも四つのジスルフィド結合が形成され、前記対合は、4、6、または10を超える複合指数を生成する。

【0011】

別の局面では、本発明の非天然システイン（C）－含有タンパクは、約50℃よりも高い温度、好ましくは約80℃よりも高い温度、さらに好ましくは100℃よりも高い温度で、0.001秒から10分の範囲における任意の期間加熱された後、標的結合能を発揮する。

【0012】

ある局面では、本明細書に記載される非天然システイン（C）－含有タンパクは、標識（すなわち、GFP、HAタグ、フラグ、Cy3、Cy5、FITC）、エフェクター（すなわち、酵素、細胞傷害剤、キレート剤）、抗体（すなわち、全体抗体、Fc領域、dAb、scFv、二重特異性抗体）、腫瘍など所望の組織または区画へ分子を集中させる標的狙撃分子（VEGFヘパリン結合エキソンなどのペプチドまたはドメイン）、組織の障壁（経皮、経口、腸内、頬内、腔内、直腸内、鼻腔内、肺内、血液脳関門、経強膜）を横断する輸送を強化する障壁輸送接合体、例えば、アルギニン富裕ペプチド、アルキルサッカリド、界面活性剤を模倣するか、またはタンパクを含むか、または表示するミセルを形成する、（イオン性、または非イオン性）両親媒性ペプチド、および、小型分子（例えば、アルブミンに結合するか、または細胞膜に挿入されるもの）を含む半減期延長成分、ポリエチレングリコール（PEG）などの化学的ポリマー、または各種ペプチドおよびタンパク配列（膜に挿入されるか、または非特異的に結合する疎水性ペプチドを含む）、（ヒト）血清アルブミン、トランスフェリン、ポリ（GGGS）リンカーなどのポリマー系グリシン富裕配列、から成る群から選ばれる成分に接合される。これらの接合体を形成する連結部は、遺伝学的に形成されてもよいし、あるいは化学的に形成されてもよい。このシステイン含有タンパクはまた、ホモまたはヘテロのマルチマーを生成して、2マー、3マー、4マー、5マー、6マー、7マー、8マー、9マー、10マー、11マー、12マー、14マー、16マー、18マー、20マー、またはそれよりもさらに高次のマルチマーを形成してもよく、これらのマルチマーは、タンパクの半減期を延長し、結合部位の濃度を増し、したがって見かけの解離定数を向上させ、標的に応じて、結合アビディティを増す場合もある。この高次マルチマーは、単一の大型遺伝子への融合、または、タンパクに対し遺伝学的にコードされるペプチド結合性ペプチド（「連結ペプチド」）を加え、それによって別々に発現されるタンパクが、N-および/またはC-末端において連結ペプチドを介して、あるいは、各種の化学的連結を通じて互いに結合してタンパクマルチマーを形成することによって創製することが可能である。好適な半減期延長成分としては、血清アルブミン、IgG、赤血球、および血清に接触可能なタンパクに結合する成分が挙げられるが、ただしこれらに限定されない。各標的および各製薬使用は、多数のこれらの要素の組み合わせの内それぞれ異なる組み合わせを好む。

【0013】

本発明はさらに、20-60個のアミノ酸から成る単一ドメインを含む非天然タンパクで、3個以上のジスルフィドを有し、ヒトの血清暴露タンパクに結合し、かつ、5%未満の脂肪族アミノ酸を有するタンパクを提供する。

【0014】

本発明はさらに、20-60個のアミノ酸から成る単一ドメインを含む非天然タンパクで、3個以上のジスルフィドを有し、ヒトの血清暴露タンパクに結合し、かつ、T-エピトーププログラムにおいてデータベースのタンパクの平均の90%よりも低い、好ましくはデータベースのタンパクの平均の99%よりも低いスコアを持つタンパクを提供する。さらに本発明に含まれるものは、主題の非天然タンパクのライブラリー、該タンパクをコ

10

20

30

40

50

ードする遺伝子パッケージを含む発現ベクター、および、該タンパクを発現または提示する他の宿主細胞である。

【0015】

さらに本発明に含まれるものは、本明細書に開示されるシステイン含有ミクロタンパクを生産する方法である。

【0016】

さらに本発明に含まれるものは、標的と、遺伝子パッケージの上に提示される外来ポリペプチドとの間の特異的相互作用の存在を検出する方法である。この方法は、(a) 本発明の遺伝子パッケージ提示を準備する工程；(b) 安定なポリペプチドー標的複合体を生成するのに好適な条件下で遺伝子パッケージを標的と接触させる工程；および(c) 遺伝子パッケージにおいて安定なポリペプチドー標的複合体の形成を検出し、それによって特異的相互作用の存在を検出する工程を含む。方法はさらに、所望の性質を有するポリペプチドを提示する遺伝子パッケージを単離するか、または、遺伝子パッケージによって担持される、所望のポリペプチドをコードする配列部分の配列を決定する工程を含んでもよい。例示の遺伝子パッケージとしては、ウィルス（例えば、ファージ）、細胞、および胞子が挙げられるが、ただしこれらに限定されない。

【0017】

（参照としての援用）

本明細書に言及される公刊物および特許出願は全て、各個別の公刊物または特許出願が、特異的および個別的に、参照により取り込まれると表示されるのと同程度に、全ての目的のために参照により本明細書に取り込まれる。

【発明を実施するための最良の形態】

【0018】

（発明の詳細な説明）

本明細書に言及される公刊物および特許出願は全て、各個別の公刊物または特許出願が、特異的および個別的に、参照により取り込まれると表示されるのと同程度に、全ての目的のために参照により本明細書に取り込まれる。

【0019】

本発明の好ましい実施態様が図示され、説明されるが、これらの実施態様は、ただ例として提供されるに過ぎないことは当業者には明白であろう。これから数多くの変種、変更、および置換が、本発明から逸脱することなく当業者には思い浮かぶことであろう。本発明を実行する際、本明細書に記載される、本発明の実施態様に対する様々な代替も採用が可能であることを理解しなければならない。

【0020】

（一般技術）

本発明の実施は、別様に指示しないかぎり、免疫学、生化学、化学、分子生物学、微生物学、細胞生物学、ゲノミクス、および組み換えDNAに関する、当業者の能力範囲内にある、通例技術を用いる。Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2版(1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, ら, eds., (1987)); シリーズ METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR2: A PRACTICAL APPROACH (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R. I. Freshney, ed. (1987))を参照されたい。

【0021】

（定義）

「タンパク」という用語は、任意の長さのアミノ酸のポリマーを指す。このポリマーは

直鎖でも分枝鎖でもよく、修飾されたアミノ酸を含んでもよく、非アミノ酸によって中断されてもよい。この用語はまた、修飾されたアミノ酸ポリマー、例えば、ジスルフィド結合形成、糖化、脂質化、アセチル化、リン酸化、または、他の任意の操作、例えば、標識化合物と接合されたアミノ酸ポリマーを含む。本明細書で用いる「アミノ酸」という用語は、天然および／または、非天然すなわち合成アミノ酸、例えば、グリシン、およびDまたはLの光学異性体、およびアミノ酸類縁体、およびペプチド様物質を含むアミノ酸を指す。タンパクは一つ以上のドメインを含んでもよい。

【0022】

「ドメイン」という用語は、サイズにはよらず、単一の、安定な三次元構造体を指す。典型的ドメインの三次元構造体は、溶液において安定で、そのメンバードメインが単離されるか、他のドメインに共有的に融合されるかとは無関係に同じ状態を持続する。本明細書に定義されるドメインは、ベータシート、アルファヘリックス、および未構造化ループなどの二次構造要素によって形成される、特定の三次元構造体を有する。ミクロタンパク・ファミリーのドメインでは、ジスルフィド架橋が、一般に、三次構造を決定する一次要素である。ある例では、ドメインは、アビディティ（同じ標的に対する複数の結合部位）、多重特異性（異なる標的に対する複数の結合部位）、ヒト血清アルブミン（HSA）、またはIgG（hIgG1、2、3、または4）、または赤血球に結合する（ドメイン、環状ペプチド、または直線状ペプチドによる）半減期などの、特定の機能的活性を付与することが可能なモジュールである。

【0023】

「ループ」は、標的との相互作用のアフィニティーおよび特異性に寄与するシステイン間配列であり、そのアミノ酸組成もまた、タンパクの可溶性に影響を及ぼす。可溶性は、高濃度処方、例えば、経口、腸内、経皮、鼻腔内、肺内、血液脳関門、在宅注射で使用される処方、およびその他の投与ルートおよび方式にとって重要である。

【0024】

「ミクロタンパク」という用語は、SCOPデータベースにおける、ある分類を指す。ミクロタンパクは通常、一定の構造を持つ最小タンパクであり、通常、2個のジスルフィド持つ15個ほどの小さなアミノ酸、または10個を超えるジスルフィドを持つ最大200個のアミノ酸を有するタンパクであるが、ただしそれらに限定されない。ミクロタンパクは、一つ以上のミクロタンパク・ドメインを含んでもよい。いくつかのミクロタンパク・ドメイン、またはドメイン・ファミリーは、様々なジスルフィド結合パターンによって与えられる、複数の幾分安定で、複数の幾分類似の構造を持つことが可能であり、そのため安定という用語は、ペプチド、および非ミクロタンパク・ドメインからミクロタンパクを区別するために相対的な意味で使用される。多くのミクロタンパク毒素は、単一ドメインから構成されるが、細胞表面の受容体ミクロタンパクは、多くの場合複数ドメインを有する。ミクロタンパクがそのように小型であることが可能なのは、それらのフォールド形成が、典型的な疎水性コアによって安定化されるのではなく、ジスルフィド結合および／またはイオン、例えば、カルシウム、マグネシウム、マンガン、銅、亜鉛、鉄、または様々な他の多価イオンによって安定化されるからである。

【0025】

「スカフォールド」という用語は、タンパクライブラリーの構築において、保存された共通配列として使用される、最小のポリペプチド「枠組み」または「配列モチーフ」を指す。スカフォールドの固定された、または保存された残基／位置の間には、可変性および超可変性位置が存在する。固定されたスカフォールド残基の間の可変領域には、きわめて多様なアミノ酸が供給され、標的分子に対する特異的結合を実現する。スカフォールドは、通常、1群の配列近縁タンパクを整列させたときに観察される保存残基によって定義される。固定された残基は、特に、その整列された複数のタンパクの機能がそれぞれ異なる場合、フォールド形成または構造にとって必要とされると考えられる。ミクロタンパク・スカフォールドの完全な記述には、システインの数、位置または間隔、および結合パターンの外、カルシウムなどのイオンに対する結合部位を含む、ループにおける任意の固定残基

10

20

30

40

50

の位置およびその名称を含んでもよい。

【0026】

「ミクロタンパク」のフォールドは、大抵ジスルフィド結合の連結パターン（すなわち、1-4、2-6、3-5）によって定義される。このパターンは、空間配置定数であり、一般に、例えば、還元・酸化（レドックス剤）によってジスルフィド連結の解除および回復することなしに別パターンに変換することはできない。一般に、近縁配列を有する天然タンパク同士は、同じジスルフィド結合パターンを取る。主要決定因子は、システイン距離パターン（CDP）、および、いくつかの固定された非cys残基の外、もしあれば金属結合部位である。少数例では、タンパクのフォールド形成はまた、周囲の配列（すなわち、ペプチド前駆体）、および、ある場合には、該タンパクが、そのフォールド形成を助ける2価の金属イオン（すなわち、Ca⁺⁺）に結合することを可能とする残基の誘導体の化学的形成（すなわち、ガンマ・カルボキシル化）によっても影響を受ける。大多数のミクロタンパクでは、そのようなフォールド形成支援は必要とされない。

【0027】

しかしながら、同じ結合パターンを持つタンパクにおいて、ループの長さおよび組成における差が、そのタンパクにむしろ異なる構造を与えるほど十分に大きいために、複数のフォールドを含む場合がやはりある。一例が、コノトキシン、シクロトキシン、およびアナトドメイン・ファミリーである。これらは、同じDBPを持つが、きわめて異なるCDPを有するために、別々のフォールドと見なされる。タンパクフォールドの決定因子は、異なるフォールドに対し構造を大きく変える任意の属性であり、例えば、システインの数および結合パターン、システインの間隔、システイン間ループ（特に、フォールド形成に、または、カルシウム（または他の金属、または補因子）結合部位の位置または組成に必要と思われる固定されたループ残基）の配列モチーフの差である。

【0028】

「ジスルフィド結合パターン」または「DBP」という用語は、システインの連結パターンを指し、タンパクのN末端からC末端に向けて1-nという数字で表される。ジスルフィド結合パターンは、空間配置的に一定であり、これは、このパターンが、1個以上のジスルフィドを、例えば、レドックス条件を用いて連結解除することによってのみ変えることが可能であることを意味する。可能な2-、3-、および4-ジスルフィド結合パターンを、下記の0048-0075パラグラフに掲げる。

【0029】

「システイン距離パターン」または「CDP」という用語は、タンパク直鎖においてシステインを隔てる、非システインアミノ酸の数を指す。いくつかの表示法が使用される：C5C0C3Cは、C5CC3Cに等しく、CxxxxxC CxxxxCに等しい。

【0030】

「位置n6」または「n7=4」という用語は、システイン間ループを指し、「n6」は、C6およびC7の間のループと定義され、「n7=4」は、C7およびC8の間のループが、システインを勘定に入れずに4アミノ酸長であることを意味する。

【0031】

「還元性フォールド解除」という用語は、還元剤（例えば、ジチオスレイトール）の存在下に、フォールドされていたタンパクのフォールドが解けることを含む。「酸化性フォールド再形成」という用語は、酸化剤の存在下に、完全にフォールドが解除された還元状態から、フォールド形成経路への再進入を含む。

【0032】

「複合体」という用語は、平均して、直線状α鎖バックボーンにおいて多数のアミノ酸によって隔てられる複数のシステインに対し、複数のシステインがジスルフィド結合される、システイン結合パターンを指す。「複合性」は、両ジスルフィド間の長さの、合計（累積）の、バックボーン直線距離として定量される。例えば、3-ジスルフィド空間配置における最大値は9（1-4 2-5 3-6 = 3 + 3 + 3）であり、最小値は3（すなわち、1-2 3-4 5-6）である。複合性パターンは、より低い複合性パターンと比

べると、長さの多様性のためにより多くの異なるフォールドを提供するが、出現頻度はより低くなるようである。例えば、天然配列ファミリーで、もっとも剛性の高い構造体の最大数は、パターン 1-4 2-5 3-6、1-6 2-4 3-5、1-5 2-4 3-6、および 1-4 2-6 3-5 において観察される。これらは全て、複合性最大のパターンであり（3 S S タンパクの 3-9 スケールにおいて複合性スコア 9）、これは、より複合性の高い空間配置の方が、より多くの異なるシステイン間隔を、すなわち、より多くのフォールドを形成することが可能であるようであることを示す。したがって、単純なジスルフィド結合パターン（例えば、1-2 3-4 5-6）の頻度を除去または低減することによって、各ジスルフィド結合パターンにおいて形成される平均フォールド数が増す（すなわち、コトキシシン対シクロチド対アナトのように、大きく異なる c y s 間隔となる）ことが期待される。単純結合パターンの大部分を除去する簡単な方法は、約 9 アミノ酸長未満のループ長を用いることである。なぜなら、天然タンパクでは、ジスルフィド連結される、c y s 残基間の最小距離（「スパン」と呼ばれる）は、一般に約 9 アミノ酸長だからである。2 S S タンパクの複合性は、2-4、4 S S タンパクでは 4-16、5 S S タンパクでは、5-25 の範囲を有する。

【0033】

ジスルフィド結合の「スパン」という用語は、システイン自身を排除した、連結システイン間のアミノ酸距離を指す。平均スパンは、下記の表 1 に示すように、10-14 AA であるが、約 12 であることが好ましい。11-14 aa の倍数が最大となるようにシステイン間隔を設けると、それは、近接ジスルフィド（隣接システイン間に形成される）を除去し、約 12 アミノ酸（その他 18、24 など）のスパンを有するシステイン残基の、多数の組み合わせを実現することになるので、構造的多様性を強化するために使用することが可能である。一例は、C X₆ C X₆ C X₆ C X₆ C X₆ C（'3 X 6'）、C X₆ C X₆ C X₆ C X₆ C X₆ C X₆ C（'4 X 6'）、C X₅ C X₅ C X₅ C X₅ C X₅ C（'3 X 5'）、C X₅ C X₅ C X₅ C X₅ C X₅ C X₅ C（'4 X 5'）、または 5-6、4-7、または 3-8 アミノ酸の範囲を持つループの組み合わせを有する類似のモチーフと考えられる。C X₆ C および C X₅ C は、一般に、この二つの隣接システイン同士を結合させるには短すぎ（最小スパンは通常約 9 アミノ酸である）、時に「サブドメイン」または「マイクロ・ドメイン」と呼ばれることもあるが、一般に完全なドメインとは見なされない、環状ペプチド構造の形成を阻止する。いくつかの例示のジスルフィドスパンを下表に示す。

【0034】

表 1. ジスルフィドスパン

【0035】

10

20

30

【化 3】

ファミリー	C1-C6 距離 (アミノ酸で表す)	ジスルフィドスパン (aa)		
		1	2	3
A	39	11	11	15
EGF	37	11	13	10
TNFR	42	12	12	17
クニツツ	52	50	23	20
ノッチ	34	23	12	15
DSL	43	24	15	28
トレフォイル	40	19	14	16
TSP1	45	33	36	10
アナト	37	25	31	19
サイログロブリン	81	32	9	20
デフェンシン1	29	27	14	19
シクロチド	24	16	14	14
SHKT	42	35	24	12
コトキシシ	29	15	13	10
トキシシ2	29	20	21	15

10

20

「システイン富裕反復タンパク（‘CRRP’）」という用語は、通常、ただしそれに限定されないが、単一ポリペプチド鎖を有し、約1%を超える、好ましくは約5%、あるいは場合によっては10%を超えるシステイン含量を持つ、ある特定の保存アミノ酸配列（「反復パターン」または「反復モチーフ」）から成る「反復単位」（また、「モジュール」、「反復列」、または「建設ブロック」とも呼ばれる）を含むタンパクを指す。このファミリーは、アンキリン・ファミリーを含むロイシン富裕反復タンパクとは配列上無関係である。CRRP単位は互いに相互作用を持ち、他のドメインとは独立して折りたたまる一つの大きなドメインを形成する。反復単位を加える、または欠失することによってCRRPのサイズを調整することが可能である。好ましい反復タンパクとしては、無関係配列によって隔てられる単一反復列とは一般に区別が可能な、同じモチーフからなる頭一尾反復列が挙げられるが、ただしこれに限定されない。

30

【0036】

本明細書で用いる「製薬学的に受容可能な担体」という用語は、標準的製薬学的担体の内の任意のもの、例えば、リン酸バッファー生食液、水、乳剤、例えば、油／水、または水／油乳剤、および各種の湿潤剤が挙げられる。組成物はさらに、安定剤および防腐剤を含むことも可能である。担体、安定剤、およびアジュバントの例については、Martin, REMINGTON'S PHARM. SCI., 15版 (Mack Publ. Co., Easton (1975)) を参照されたい。

40

【0037】

「製薬組成物」とは、活性剤と、不活性であるか活性であるかを問わない、担体との組み合わせを含み、それによって、インビトロ、インビボ、または体外での診断または治療用途に好適な組成物が形成されることが意図される。

【0038】

核酸またはタンパクに適用される場合の、「非天然性」という用語は、天然には認められない核酸またはタンパクを指す。非天然性核酸およびタンパクの例としては、組み換えによって修飾されるものが挙げられるが、ただしこれらに限定されない。

【0039】

50

(システイン含有タンパクおよびタンパクライブラリーの設計)

以下に詳述するように、本発明の一局面は、広範な構造的多様性を有するタンパクライブラリーであって、それから、治療、予防、獣医学、診断、試薬、または材料用途を含む多様な用途、ただしこれらに限定されないが、のために所望の性質を有する結合タンパクを選択し、進化させることが可能なタンパクライブラリーを創製することである。

【0040】

一実施態様では、本発明は、好ましくは空間配置的に異なる、少なくとも2、3、4、5、10、30、100、300、1000、3000、10000以上の異なる構造を有する、システイン含有タンパクライブラリーを提供する。ある実施態様では、システイン含有タンパクライブラリーは、高ジスルフィド密度(HDD)タンパクを含む。HDDファミリーのタンパクは、通常、5-50%(5、6、7、8、9、10、12、14、16、18、20、25、30、35、40、45、または50%)のシステイン残基を有し、各ドメインは、通常、少なくとも二つのジスルフィド、および、任意に、補因子、例えば、カルシウムまたは別のイオンを含む。

【0041】

HDDスカフォールドの存在によって、これらのタンパクは、小型ではあるが、比較的剛強な構造を取ることが可能となる。剛性は、高い結合アフィニティー、抗原処理に関与するプロテアーゼ(プロテアーゼの分類については下記参照)を含むプロテアーゼおよび熱に対する耐性の実現に重要であり、したがって、そのタンパクの低い、または非免疫原性に寄与する。ジスルフィド枠組みによって、タンパクは、疎水性コアと呼ばれる、多くのタンパク内部における多数の疎水性側鎖の相互作用を要することなく、折りたたまれる。非HDDスカフォールドは全て、特異性またはフォールド形成においてしばしば問題の原因となる疎水性コアを有する。HDDタンパクは、非HDDタンパクよりも親水性であり、このため結合特異性が改善される。小さいサイズも、速やかな組織浸透には有利であり、かつ、経口、鼻腔内、腸内、肺内、血液脳関門などの代替輸送においても有利である。さらに、小さなサイズは、免疫原性を下げるのにも役立つ。ジスルフィドの数を増すか、または、同数のジスルフィドを持つが、少ない数のアミノ酸を有するドメインを用いることによって、より高いジスルフィド密度を実現することが可能である。さらに、非システイン固定残基の数を減らし、より高いパーセンテージのアミノ酸を、標的結合に利用可能となるようにすることが望ましい。

【0042】

ジスルフィド枠組みによって、システイン間ループにおける各ファミリーの極端な配列多様性が可能になる。ファミリー間には、ループ長およびシステイン間隔において膨大なばらつきが存在する。ジスルフィド結合形成の組み合わせの性質により、ジスルフィド枠組みにより、多数の異なる結合パターンおよび異なる構造の形成が可能となり、かつ、フォールド形成は異種的であってもよい。ため、指向的進化によって構造および配列を最適化するための、ゆっくりと進行する進化経路が存在する。特に、HDDタンパクは、単一配列が、複数の異なる安定フォールドを取ることが可能とする独特の能力を有することが予想される。

【0043】

広範なジスルフィド結合パターンを生成するために、種々のジスルフィド結合パターン(DBP)を有する種々の異性形に有利に作用する、ある一定範囲の異なる条件下に、ライブラリーを置くことも可能である。例えば、種々のDBPの形成を実現するために、還元剤と酸化剤の相対的濃度および強度によって決定される、溶媒のレドックス電位を利用することも可能である。還元性溶媒を創製するために、各種の還元剤、例えば、2-メルカプトエタノール(ベーターメルカプトエタノール、BME)、2-メルカプトエチルアミン-HCl、TCEP(トリス(2-カルボキシエチル)フォスフィン)、水素化ホウ素ナトリウム、ジチオスレイトール(DTT、還元形)、グルタチオンの還元形(GSH)、およびシステインの還元形を含む還元剤を、ただしこれらに限定されないが、使用することが可能である。酸化性溶媒を創製するには、各種の酸化剤、例えば、ジチオスレイ

トール（D T T、酸化形）、過酸化水素、グルタチオン（酸化形、G S S G）、銅フェナントロリン（酸化形）、酸素（空気）、微量金属、およびシステインの酸化形（シスチン）を含む酸化剤を、ただしこれらに限定されないが、使用することが可能である。

【0044】

特に有用なのは、多様なジスルフィド結合パターンの探索を可能とし、時間と共に安定形の集積を可能とするほど十分急速に、タンパクがジスルフィドを繰り返し形成・破壊することを可能とするレドックス試薬の混合物および勾配である。安定性ではなく、D B Pの最大多様性を望む場合、混合物が平衡に達するのを阻止してもよい。多様な構造（完全に還元された、高温）に好都合な条件は、突然、高度に酸化性の、低温条件に変えて、構造体が、もっとも安定なD B Pを見出すことができないほどの時間で形成されるようにする。構造多様性を創出するもう一つの方法は、多様な条件、例えば、種々の薬品（すなわち、遠くに位置するシステインとの遅い／困難なジスルフィド結合を加速する、ポリエチレングリコールのような容積排除剤）、種々の溶媒（極性、非極性、アルコール）、種々の金属イオン（C a、Z n、C u、F e、M g、その他）、または様々のp H（p H 1、2、3、4、5、6、7、8、9、10、11、12）などの条件下でゆっくりとジスルフィドを形成することである。この各種条件は、同じタンパク配列に多様な別形フォールドを取らせるために、単独で、または任意の組み合わせで使用することが可能である。

10

【0045】

ジスルフィドの形成、および／または、補因子の存在は、還元剤または酸化剤の供給、または補因子の添加によって簡単に調節することが可能である。

20

【0046】

あるタンパクが、折りたたまれて複数の別形の安定構造に変わることができる能力は、通常、該タンパク内部の結合相互作用の数と強度の外に、利用可能なフォールド形成経路（単数または複数）の性質に依存する。ジスルフィドの不在下では、安定な折りたたみタンパクを得るためには、通常、多数の弱い側鎖接触（塩橋、ファンデルワールス接触、疎水性相互作用など）が要求される。したがって、異なる、別様の安定フォールドの形成、または、標的への結合を指令するためには、多くの残基の修飾が必要とされる。一方、タンパクに安定構造を与えるには、僅かに数個の（例えば、2または3個）ジスルフィド結合で十分であり、他のアミノ酸の位置（通常、約65－80％）は全て、所望の標的（コノトキシンでは80％を超え、このもっとも極端な例である）のための結合表面の創出のために利用することが可能である。したがって、ジスルフィドは、構造に対する低い情報含量対処法（すなわち、ランダム配列では、高い出現頻度を持つ）であり、アミノ酸の最大分画を、結合およびその他の各種機能のために利用可能とする。

30

【0047】

大型の、非ジスルフィド含有タンパクのフォールド形成経路および安定性は、大きな割合の残基部分が多少なりとも固定されるように、多数のアミノ酸側鎖相互作用を必要とする。そのため、タンパクが配列を適応させる能力が大きく低下する。この状況は、通常、比較的大型のスカフォールド・タンパク、例えば、免疫グロブリン、フィブロネクチン、およびリポカリンなどのタンパクに起こるが、これらのタンパクでは、通常、ミスフォールド形成を起こさないようにすると、少数のC D R様ループのみがランダム化されるだけにすぎない。しかも、疎水性コアを含む、これらのタンパクでは、ミスフォールド形成は、一般に非可逆的なタンパク凝集を意味する。数回の突然変異によって導入される単一ジスルフィド架橋は、多数のアミノ酸残基の構造的機能を引き受け、それらの配列を解放して、異なる目的、例えば、所望のタンパク標的に対する結合に向かって進化するように仕向ける。非H D Dタンパクの場合でも、ジスルフィドを徐々に付加することは、そのタンパクが、複合性の増大に向かって進化し続けるためのもっとも重要な鍵となる場合がある。システイン（C）は、20種の生物学的アミノ酸レパートリーに、後から付け加えられたように見えるが、システインの頻度は、タンパク進化中に徐々に上昇することが示された。

40

【0048】

50

さらに、ジスルフィド介在性フォールド形成によって、タンパクはより親水性になる（なぜなら、ジスルフィドは、疎水性コアを置換する）ので、このようなタンパクのミスフォールド形成は、一般に非可逆的凝集をもたらすことはなく、タンパクは可溶状態を持続し、最終的に再生が可能となる。

【0049】

ジスルフィド固有の特徴は、同じ一組のシステインが、原理的には、様々の別形のジスルフィド結合パターンとして連結することが可能であることである。なぜなら、ジスルフィドは組み合わせだからである。例えば、二つのジスルフィドタンパクは、3通りの異なるジスルフィド結合パターン（DBP）を有し、三つのジスルフィドタンパクは、15通りの異なるDBPを有し、かつ、四つのジスルフィドタンパクは、最大105通りの異なるDBPを有することが可能である。2SS DBPの全て、3SS DBPの大部分、および4SS DBPの半分未満において天然例が存在する。一局面において、ジスルフィド結合パターンの全数は、式：Error! Objects cannot be created from editing field codes. に従って計算することが可能である。式中、 n ＝システイン残基によって形成されるジスルフィド結合の予測数であり、かつ、式中、Error! Objects cannot be created from editing field codes. は、 $(2i-1)$ の積を表す。この式において i は、1から n までの範囲の正の整数である。

【0050】

したがって、一実施態様では、標的分子に対して結合特異性を示す、非天然システイン（C）－含有スカフォールドを提供する。この実施態様において、非天然システイン（C）－含有スカフォールドは、式：Error! Objects cannot be created from editing field codes. によって表される組み合わせの群から選ばれるパターンによるスカフォールド内システインを含む。式中、 n は、システイン残基によって形成されるジスルフィド結合の予測数に等しく、かつ、式中、Error! Objects cannot be created from editing field codes. は、 $(2i-1)$ の積を表し、この式において i は、1から n までの範囲の正の整数である。一局面において、非天然システイン（C）－含有タンパクは、 $C^{1-2, 3-4}$ 、 $C^{1-3, 2-4}$ 、および $C^{1-4, 2-3}$ から成る群から選ばれるパターンに従って、ポリペプチドに含まれるシステインを対合させることによって形成される二つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。別の局面では、非天然システイン（C）－含有タンパクは、 $C^{1-2, 3-4, 5-6}$ 、 $C^{1-2, 3-5, 4-6}$ 、 $C^{1-2, 3-6, 4-5}$ 、 $C^{1-3, 2-4, 5-6}$ 、 $C^{1-3, 2-5, 4-6}$ 、 $C^{1-3, 2-6, 4-5}$ 、 $C^{1-4, 2-3, 5-6}$ 、 $C^{1-4, 2-6, 3-5}$ 、 $C^{1-5, 2-3, 4-6}$ 、 $C^{1-5, 2-4, 3-6}$ 、 $C^{1-5, 2-6, 3-4}$ 、 $C^{1-6, 2-3, 4-5}$ 、および $C^{1-6, 2-5, 3-4}$ から成る群から選ばれるパターンに従って、ポリペプチドに含まれるシステインを対合させることによって形成される三つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。別の局面において、非天然システイン（C）－含有タンパクは、前述の式によって定義される組み合わせの群から選ばれるパターンに従って、ポリペプチドに含まれるシステインを対合させることによって形成される、少なくとも四つのジスルフィド結合を有するポリペプチドを含む、標的分子に対して結合特異性を示すポリペプチド、非天然システイン（C）－含有タンパクを含む。さらに別の局面では、非天然システイン（C）－含有タンパクは、 C^{1-9} 、 C^{1-10} 、 C^{2-9} 、 C^{2-10} 、 C^{3-9} 、 C^{3-10} 、 C^{4-9} 、 C^{4-10} 、 C^{5-9} 、 C^{5-10} 、 C^{6-9} 、 C^{6-10} 、 C^{7-9} 、 C^{7-10} 、 C^{8-9} 、 C^{8-10} 、および C^{9-10} から成る群から選ばれるパターンに従って、タンパク内のシステインを対合さ

せることによって形成される、少なくとも五つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。さらに別の局面では、標的分子に対し結合特異性を示す、非天然システイン(C)－含有タンパクは、 C^1-1^1 、 C^1-1^2 、 C^2-1^1 、 C^2-1^2 、 C^3-1^1 、 C^3-1^2 、 C^4-1^1 、 C^4-1^2 、 C^5-1^1 、 C^5-1^2 、 C^6-1^1 、 C^6-1^2 、 C^7-1^1 、 C^7-1^2 、 C^8-1^1 、 C^8-1^2 、および C^9-1^1 、 C^9-1^2 、 $C^{10}-1^1$ 、 $C^{10}-1^2$ 、および $C^{11}-1^2$ から成る群から選ばれるパターンに従って、タンパク内のシステインを対合させることによって形成される、少なくとも六つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。

10

【0051】

通常、全てのシステインが、同じドメイン中の他のシステインに対するジスルフィド結合に与る。二つのジスルフィド(2SS)を有するマイクロタンパクは、空間配置的に異なる(すなわち、単純な回転によっては相互変換されない)三つのジスルフィド結合パターン： $1-2 \quad 3-4$ 、 $1-3 \quad 2-4$ 、または $1-4 \quad 2-3$ 、これらはそれぞれ異なるアルファ鎖バックボーン構造を有する、を取ることが可能である。

【0052】

同様に、三つのジスルフィドを有するマイクロタンパクは、最大15通りの異なるジスルフィド結合パターンを持つことが可能であり、四つのジスルフィドを有するマイクロタンパクは、最大105通りのジスルフィド結合パターンを持つことが可能であり、五つのジスルフィドを有するマイクロタンパクは、最大945通りの結合パターンを持つことが可能であり、六つのジスルフィドを有するマイクロタンパクは、最大10,395通りのジスルフィド結合パターンを持つことが可能であり、七つのジスルフィドを有するタンパクは、最大135,135通りの異なる結合パターンを持つことが可能であり、さらに大きなジスルフィド数についても以下同様である(掛け数は、3、5、7、9、11、13倍)。下記は、二つ、三つ、または四つのジスルフィド結合を有するタンパクにおけるジスルフィド結合パターン(DBP)を掲げる。

20

【0053】

2SSタンパクにおける3通りのDBPパターンは：

$1-2 \quad 3-4$ 、 $1-3 \quad 2-4$ 、 $1-4 \quad 2-3$

3SSタンパクにおける15通りのDBPパターンは：

30

【0054】

【化4】

1-62-53-4, 1-42-53-6, 1-62-43-5, 1-52-63-4, 1-52-43-6, 1-42-63-5,

1-23-45-6, 1-23-54-6, 1-23-64-5, 1-62-34-5, 1-42-35-6, 1-52-34-6,

1-32-64-5, 1-32-45-6, 1-32-54-6.

4SSタンパクにおける105通りのDBPは：

40

【0055】

【化 5】

1-2 3-4 5-6 7-8 1-2 3-4 5-7 6-8 1-2 3-4 5-8 6-7 1-2 3-5 4-6 7-8 1-2 3-5 4-7 6-8 1-2 3-5 4-8

6-7

1-2 3-6 4-5 7-8 1-2 3-6 4-7 5-8 1-2 3-6 4-8 5-7 1-2 3-7 4-5 6-8 1-2 3-7 4-6 5-8 1-2 3-7 4-8
5-6 1-2 3-8 4-5 6-7 1-2 3-8 4-6 5-7 1-2 3-8 4-7 5-6 1-3 2-4 5-6 7-8 1-3 2-4 5-7 6-8 1-3 2-4
5-8 6-7

1-3 2-5 4-6 7-8 1-3 2-5 4-7 6-8 1-3 2-5 4-8 6-7 1-3 2-6 4-5 7-8 1-3 2-6 4-7 5-8 1-3 2-6 4-8
5-7

1-3 2-7 4-5 6-8 1-3 2-7 4-6 5-8 1-3 2-7 4-8 5-6 1-3 2-8 4-5 6-7 1-3 2-8 4-6 5-7 1-3 2-8 4-7
5-6

10

1-4 2-3 5-6 7-8 1-4 2-3 5-7 6-8 1-4 2-3 5-8 6-7 1-4 2-5 3-6 7-8 1-4 2-5 3-7 6-8 1-4 2-5 3-8
6-7

1-4 2-6 3-5 7-8 1-4 2-6 3-7 5-8 1-4 2-6 3-8 5-7 1-4 2-7 3-5 6-8 1-4 2-7 3-6 5-8 1-4 2-7 3-8
5-6

1-4 2-8 3-5 6-7 1-4 2-8 3-6 5-8 1-4 2-8 3-7 5-6 1-5 2-3 4-6 7-8 1-5 2-3 4-7 6-8 1-5 2-3 4-8
6-7

1-5 2-4 3-6 7-8 1-5 2-4 3-7 6-8 1-5 2-4 3-8 6-7 1-5 2-6 3-4 7-8 1-5 2-6 3-7 4-8 1-5 2-6 3-8
4-7

20

1-5 2-7 3-4 6-8 1-5 2-7 3-6 4-8 1-5 2-7 3-8 4-6 1-5 2-8 3-4 4-7 1-5 2-8 3-6 4-7 1-5 2-8 3-7
4-6

1-6 2-3 4-5 7-8 1-6 2-3 4-7 5-8 1-6 2-3 4-8 5-7 1-6 2-4 3-5 7-8 1-6 2-4 3-7 5-8 1-6 2-4 3-8
5-7

1-6 2-5 3-4 7-8 1-6 2-5 3-7 4-8 1-6 2-5 3-8 4-7 1-6 2-7 3-4 5-8 1-6 2-7 3-5 4-8 1-6 2-7 3-8
4-5

1-6 2-8 3-4 5-7 1-6 2-8 3-5 4-7 1-6 2-8 3-7 4-5 1-7 2-3 4-5 6-8 1-7 2-3 4-6 5-8 1-7 2-3 4-8
5-6

30

【 0 0 5 6 】

【化 6】

1-7 2-4 3-5 6-8 1-7 2-4 3-6 5-8 1-7 2-4 3-8 5-6 1-7 2-5 3-4 6-8 1-7 2-5 3-6 4-8 1-7 2-5 3-8
4-6

1-7 2-6 3-4 5-8 1-7 2-6 3-5 4-8 1-7 2-6 3-8 4-5 1-7 2-8 3-4 5-6 1-7 2-8 3-5 4-6 1-7 2-8 3-6

4-5

1-8 2-3 4-5 6-7 1-8 2-3 4-6 5-7 1-8 2-3 4-7 5-6 1-8 2-4 3-5 6-7 1-8 2-4 3-6 5-7 1-8 2-4 3-7

5-6

1-8 2-5 3-4 6-7 1-8 2-5 3-6 4-7 1-8 2-5 3-7 4-6 1-8 2-6 3-4 5-7 1-8 2-6 3-5 4-7 1-8 2-6 3-7

40

4-5

1-8 2-7 3-4 5-6 1-8 2-7 3-5 4-6 1-8 2-7 3-6 4-5.

大型で、システイン少数タンパクは、広範な二次、三次構造、または場合によっては四次構造を必要とし、別形ジスルフィド結合パターンによって仲介される別形フォールドの形成を阻止する。ミクロタンパクでは、ジスルフィド誘発構造以外の二次または三次構造は、ほとんど、または全く無く、システイン間のループ配列（一次構造）は、アミノ酸組成において例外的に変動性が高い。したがって、他のタンパクに比べ、ミクロタンパクは、様々の異なる結合パターンを取ることが可能な配列弾力性を持つ可能性はるかに高い。

50

【0057】

少数のシステインは、多様な、完全に異なる空間配置構造を提供することが可能であるが、これは、ジスルフィドを破ることなしには相互変換が不可能であることを意味する。これらの構造は、通常、ループに対しては全く、またはごく僅かしか要求を課すことなく実現されるので、結合特異性、および特異的標的に対するアフィニティーの創出について、ループ配列を思いのままに利用することが可能である。ある特異的タンパク配列は、あるフォールドに対しては、別のフォールドよりもはっきりした好みを示すようであり、あるフォールドを全く取ることができない場合がある。天然のミクロタンパクのファミリーの配列モチーフから、システイン間隔はDBPに寄与するが、非-cysのループ残基からの寄与はごく僅かであるようである。高ジスルフィド密度タンパクにおけるシステイン間ループの平均長は、多くの好ましいスカフォールドでは約0から約10の範囲であり、大部分のスカフォールドでは約3から約15アミノ酸である。これは、あるスカフォールドでは約50%、25%–20%（もっとも好ましい）、15%–10%（少し好ましくない）、または場合によって5%の範囲の、高密度のシステインを提供する。これらはいずれも、僅か0.8%という平均的タンパクにおけるシステインの密度よりもはるかに高い。要すれば、システイン同士をごく緊密に近づくように加工して、ジスルフィドを効率的、適正に形成することも可能である。効率的結合形成によって、もっとも弱い結合の破壊および新規結合の再生を多数サイクル繰り返すことが可能となり、それによって、もっとも安定に結合されるタンパクが次第に得られる。大型タンパクではシステインが低密度であることが、ジスルフィドの非効率的で、したがって不適正となり易い形成に寄与するようである。

10

20

【0058】

種々のジスルフィド結合パターンは、温度およびプロテアーゼに対するその安定性において異なることが予想される。したがって、本発明は、(a)標的分子に結合することが可能であり、(b)スカフォールド内システインを対合させることによって形成される少なくとも二つのジスルフィド結合を有し、(c)約50℃よりも高い温度、好ましくは約80℃よりも高い温度、場合によっては約100℃よりも高い温度で、0.01秒から10秒までの範囲の任意の期間加熱された後も標的結合性を発揮する、非天然システイン(C)–含有スカフォールドを提供する。要すれば、この非天然システイン(C)–含有スカフォールドは、スカフォールド内システインを対合させることによって形成される、少なくとも、三つ、四つ、五つ、六つ、七つ、八つ、九つ、十、十一、十二、またはそれ以上のジスルフィド結合を含むように設計されてもよい。

30

【0059】

より高度に架橋される（例えば、高い複合数を有する）タンパクの方が、「サブドメイン」を形成し、一つ以上のジスルフィドを含むが、タンパクの他の部分に対して自由に回転することが可能なタンパクよりも安定であることが予想される。より高い安定性は、直線状ペプチドの上に引かれたジスルフィドの（累積）長さ、および、直線状ペプチド配列を用いたDBPダイアグラムにおいてジスルフィドが互いに交わる回数と相関する。しかしながら、種々のジスルフィド結合パターンは、様々な収率で形成され、架橋数の最大ものの抽出が最小となることが予想される。システイン同士の近傍性がジスルフィド形成の駆動実体である限り、近接システイン同士の間に形成されるジスルフィドは、出現する確率が高いが、同時に安定性の視点からはもっとも好ましくない、なぜなら、そのようなジスルフィドは、ミクロ、またはサブドメインを形成するからである。

40

【0060】

したがって、ある実施態様では、本発明は、非天然システイン(C)–含有タンパクを有するライブラリーを提供する。これらのタンパクにおいては、それぞれが、35未満のアミノ酸を含み、そのポリペプチド中のアミノ酸の少なくとも10%がシステインであり、少なくとも二つのジスルフィド結合が、スカフォールド内システインを対合させることによって形成され、かつ、この対合が、3を超える複合指数を生成する。他の、いくつかの実施態様では、本発明は、非天然システイン(C)–含有タンパクを有するライブラリ

50

ーを提供する。これらのタンパクにおいては、それぞれが、約60未満のアミノ酸を含み、そのポリペプチド中のアミノ酸の少なくとも10%がシステインであり、少なくとも四つのジスルフィド結合が、ポリペプチドに含まれるシステインを対合させることによって形成され、かつ、前記対合が、4、6、または10を超える複合指数を生成する。

【0061】

ある局面では、主題のマイクロタンパクは、任意の標的に対しピコモル活性を発揮し、加熱（場合によっては沸騰）およびプロテアーゼに対し高い耐性を持っていてもよい。別の局面では、主題のマイクロタンパクは、高度の親水性傾向を持ち、ドメイン当たり二つの異なる結合面（二面性）を持つ傾向を有していてもよい。

【0062】

各ジスルフィド結合パターンは、理論上は、システインの、広範な各種間隔と適合するが、あるシステイン間隔パターンは、ある特異的結合パターンに対し、別のシステイン間隔パターンよりも適合する。天然配列では、各ジスルフィド結合パターンに関連して、複数の優勢システイン間隔パターンがある。例えば、コノトキシン、シクロチド、およびアナト・ファミリー（異なるフォールドと考えられる）は、大きく異なるシステイン間隔を有するが、ジスルフィド結合パターンは同じである。したがって、ジスルフィド結合パターンの頻度分布を主に決定するものはシステインの間隔であり、CDPの設計が、DBPおよび構造を調節し、その進化を方向づける実際的方法である。システインの間隔は、システイン間ループの長さを決定し、かつ、そのタンパクの「フォールド」を大きく決定する。同じ配列ファミリーに属するタンパク同士は、高度に保存されるアミノ酸位置、およびそれらの優勢間隔の全てから構成される、同じスカフォールド配列、またはスカフォールドモチーフを共有するので、これらは、通常、同じ「フォールド」を持つと考えられる。

【0063】

主題マイクロタンパクは、モノマー、ダイマー、トリマー、またはより高次のマルチマーであってもよい。複数ドメイン・マイクロタンパクは、ホモマルチマーであってもよいし、あるいは、ヘテロマルチマーであってもよい。すなわち、ドメイン同士が、ジスルフィド数、ジスルフィド結合パターン、構造、フォールド、配列、またはスカフォールドにおいて異なるマルチマーであってもよい。主題マイクロタンパクは、種々の異なる構造体、例えば、種々の異なる長さ、アミノ酸組成および機能を持つペプチド（直線状、または環状）を含む構造体に融合させることも可能である。各ドメインは、天然の毒素と同じ、または異なる、種々の標的に対し、一つ以上の結合面（二面性）を持つことが可能である。

【0064】

本発明はさらに、一つ以上のドメイン、および任意に一つ以上の（環状または直線状）ペプチドを含む単一タンパク鎖を有する非天然マイクロタンパクを提供する。一般に、各ドメインは、別々に折たたまり、機能する。マイクロタンパク・ドメインは、該ドメインのサイズ、その温度に対する安定性、および、大腸菌におけるその発現レベル（したがって、商品のコスト）の大部分を決定する高ジスルフィド密度「スカフォールド」を有する。このスカフォールドはさらに、該タンパクの免疫原性の決定に重要な役割を果たすことが予想される。スカフォールドは、4、6、8、10、12、14、16、18、またはそれ以上のシステインを含み、これらのシステインは、同じドメイン内に2、3、4、5、6、7、8、またはそれ以上のジスルフィド結合を形成する。

【0065】

複数の性質において改善を示す、好ましい特異的3-ジスルフィドスカフォールドの内のいくつかは、コノトキシン（29aa合計、7aa固定化、Ca-部位無し、1-4 2-5 3-6ジスルフィド結合パターンによる剛性構造）、シクロチド（24aa合計、10aa固定、Ca-部位無し、剛強な1-4 2-5 3-6構造）、アナト・スカフォールド（37aa合計、10aa固定、Ca-部位無し、剛強な1-4 2-5 3-6ジスルフィド結合パターン）、デフェンシンIスカフォールド（29aa、10aa固定、Ca-部位無し、剛強な1-6 2-4 3-5結合パターン）、トキシン2スカ

フォールド（29aa合計、10aa固定、Ca-部位無し、剛強な1-4 2-6 3-5ジスルフィド結合スカフォールド）であるが、他にも、既存、および新規の、広範なスカフォールドが特異的利点を提供する。他の好ましいスカフォールドとして、173メンバーを擁するPfamファミリーPF00734で、1-3 2-4連結の4システイン、C10C5C9CのCDPを有する26AA長（第1から最終Cysまで）である、セルロース結合ドメイン（CB, CEB）；25メンバーを擁するPF00451ファミリーで、15AA長、1-3 2-4連結の4システイン、C0C4CCのCDPを有する、アルファ-コノトキシン（AC）；68メンバーを擁するPF00451ファミリーで、1-3 2-5 3-6連結の6システイン、C5C3C10C4C1CのCDPを有する28AA長、オメガートキシン様（OT）；39メンバーを擁するPF05375ファミリーで、29AA長、1-4 2-6 3-5連結の6システイン、C9C2C1C8C4CのCDPを有する、パシファスチン（PC）；35メンバーを擁するPF00299ファミリーで、26AA長、1-4 2-5 3-6連結の6システイン、C6C5C3C1C6CのCDPを有する、セリンプロテアーゼ阻害剤（SP）；175メンバーを擁するPF00066ファミリーで、33AA長、1-5 2-4 3-6連結の6システイン、C7C8C3C4C6CのCDPを有する、ノッチ（NO）；126メンバーを擁するPF00088ファミリーで、39AA長、1-5 2-4 3-6連結の6システイン、C10C10C4C0C10CのCDPを有する、トレフォイル（TR）；123メンバーを擁するPF01821ファミリーで、42AA長、1-2 3-5 4-6連結の6システイン、C14C2C2C11C7CのCDPを有する、TNF-受容体様（TN）；123メンバーを擁するPF01821ファミリーで、42AA長、1-2 3-5 4-6連結の6システイン、C14C2C2C11C7CのCDPを有する、TNF-受容体様（TN）；123メンバーを擁するPF01821ファミリーで、37AA長、1-4 2-5 3-6連結の6システイン、C5C2C8C2C5C1CのCDPを有する、アナフィロトキシン様（AT）；410メンバーを擁するPF01437ファミリーで、61AA長、1-4 2-8 3-6 4-7連結の8システイン、C5C2C8C2C5C12C19CのCDPを有する、プレキシシン（PL）がある。他の好ましいスカフォールドとして、58AA長（第1から最終Cysまで）で、1-3 2-4 5-6 7-8連結の8システイン、C13C6C16C1C10C0C4CのCDPを有する、三指トキシン（PL）；35AA長で、1-2 3-4 5-6 7-8連結（交互連結DBPは既知であることに注意）の8システイン、C3C9C1C3C5C0C6CのCDPを有する、ソマトメジン；47AA長で、8システイン、C3C8C11C2C0C5C10CのCDPを有する、じゃがいもプロテアーゼ阻害剤（PI）；37AA長で、1-4 2-5 3-6 7-8連結の8システイン、C5C2C8C2C5C12C19CのCDPを有する、キチンビンジンドメイン（CHB）；34AA長で、6システイン、C6C6C0C4C6CのCDPを有する、蜘蛛トキシン（ST）；34AA長で、6システイン、C6C5C0C3C8CのCDPを有する、トキシンB（TB）；26AA長で、1-3 2-4連結の4システイン、C10C5C9CのCDPを有する、セルロース結合ドメイン（CEB）；15AA長で、1-3 2-4連結の4システイン、C0C4C8CのCDPを有する、アルファ-コノトキシン（AC）がある。

【0066】

主題の非天然ミクロタンパクは、天然のタンパク配列に基づいて設計されてもよい。例えば、本明細書に含まれる多くの天然タンパクまたはドメインは、スカフォールドタンパクとして使用するのに魅力的な特徴を有する。

【0067】

【化 7】

表 2

タンパクファミリー	ファミリーにおける他の例示のメンバー
インスリン様	
有毒ヘアピン	熱安定性エンテロトキシン、ニューロトキシンB-IV
ノッチン	植物レクチン、抗菌ペプチド(ヘベイン様 アグルチニン(レクチン)ドメイン)、 抗菌ペプチド2、AC-AMP2)
プロテイナーゼおよびアミラーゼの植物阻害剤	トリプシン阻害剤、カルボキシペプチダーゼA 阻害剤、 α -アミラーゼ阻害剤
シクロチド	カラタBI、シクロピオラクチンO1、サークリンA、 パリクレイン
グルマリン様	
アグーチ関連タンパク	
オメガトキシン様	コオトキシン、蜘蛛トキシン、昆虫トキシン、 アルブミンI
さそりトキシン様	長鎖さそりトキシン(さそりトキシン、 アルファトキシン、Tx10アルファ様トキシン、 LQH IIIアルファ様トキシン) 短鎖さそりトキシン、デフェンシンMGD-1、 昆虫デフェンシン、植物デフェンシン
セルロース結合ドメイン	セロビオヒドロラーゼI

【0068】

【化 8】

タンパクファミリー	ファミリーにおける他の例示のメンバー
増殖因子受容体ドメイン	インスリン様増殖因子結合タンパク5 IGFB-5、1型インスリン様増殖因子受容体 Cys-富裕ドメイン、受容体タンパクチロシン キナーゼErbB-3、Cys-富裕ドメイン、EGF受 容体Cys-富裕ドメイン、癌原タンパクHer2 細胞外ドメイン
コリパーゼ様 EGF/ラミニン	(プロ)コリパーゼ腸内トキシ1 EDF-型モジュール(因子IX、凝固因子 VIIa、E-セレクチン、因子X、N-末端モジ ュール、活性化タンパクC(オートプロト ロンピンIIa)、プロスタグランジンH2シン ターゼ-1、EGF-様モジュール、P-セレク チン、上皮増殖因子(EGF)、トランスフォ ーミング増殖因子アルファ、エピレグリン 、EGFドメイン、ペーテセルリン-2、ヘパ リン結合性上皮増殖因子HBEGF、プラ スミノゲン活性化因子(ウロキナーゼ 型)、ヘレグリンアルファ、EGFDメイン、 トロンボモジュリン、フィブリリン-1、マン ノース結合タンパク関連セリンプロテア ーゼ2、補体CIS、補体プロテアーゼCIR 、プラスミノゲン活性化因子(組織型) (tPA)、低密度脂質タンパク(LDL)受容 体)、インテグリンベータEGF様ドメイン、 ニドゲン-1のEGF様ドメイン、ラミニン 型モジュール、ラミニン γ 1鎖、フォリス タチンモジュールN-末端ドメインFS-N、 BM-40/SPARC/オステオネクチンのド メイン、フォリスタチンのドメイン、メロゾ イト表面タンパク1(MSP-1)
プロメライン阻害剤VI (システインプロテイナーゼ阻害剤)	
ボウマン-バーク阻害剤	
エラフィン様	エラフィン、エラスターゼ特異的阻害剤、 ナワブリン
蛭抗止血タンパク	プリスタチン様、ヒルジン様
グラニューリン反復列	グラニューリン-1のN-末端ドメイン、 オリザインベータ鎖
満腹因子CART (コカインおよびアンフェタミン調節転写物)	
DPYモジュール	Dumpy
バブルタンパク	
PMP阻害剤	
TSP-1型1反復列	トロンボスポンジン-1
AmbV	
蛇毒様	蛇毒トキシ(エラプトキシB、 γ -カルジ オトキシ、ファシクリン、ムスカリニントキ シン、エラプトキシA、ニューロトキシI、 カルジオトキシV4II(トキシIII)、カルジ オトキシV、 α -コブラトキシ、長ニュー ロトキシ1、FS2トキシ、ブンガロトキシ 、プカンジン、カルジオトキシCTXI、カル ジオトキシCTX IIB、カルジオトキシII、 カルジオトキシIII、カルジオトキシIV、コ ブラトキシ2、 α -トキシ、ニューロキシ II(コブラトキシB)、トキシB(長ニュー ロトキシ)、カントトキシ、プカイン)、 デンドロアスピン
BPTI-様	
(ヒト)細胞表面受容体の細胞外ドメイン	CD59、II型アクトチン受容体、BMP受容体 Iaエクドメイン、TGF-ベータII型受容体 細胞外ドメイン
デフェンシン様	デフェンシン、デフェンシン2、ミオトキシ
ヘアピンループ含有ドメイン様	APPLEドメイン
ニューロトキシIII(ATXIII)	
LDL-受容体様モジュール	
クランピン様	
クリングル様	クリングルモジュール、フィブロネクチンII型
Kazal型セリンプロテアーゼ阻害剤	
植物プロテイナーゼ阻害剤	

【化 9】

タンパクファミリー	ファミリーにおける他の例示のメンバー
トレフォイル/プレキシンドメイン様	トレフォイル、プレキシシ
壊死誘発タンパク1、NIP1	
シスチン-ノットサイトカイン	PDGF-様、TGF- β -様、ノギン、ニューロトロフィン、ゴナドトロピン/フォリトロピン、インターロイキン17F、コアグロゲン
補体調節モジュール、SCRドメイン	CD46、 β 2-糖タンパク、補体受容体1.2 (cr1, cr2)、補体C1RおよびC1Sプロテアーゼドメイン、MASP-2
いそざんちゃくトキシシ	
血液凝固阻害剤 (ジスインテグリン)	エキスタチン、フラボリジン、キストリン、オブツスタチン、サルモシン、シスタチン
メチルアミンジヒドロゲナーゼ、L鎖	
セリンプロテアーゼ阻害剤	ATI-様、BSTI-様
TB-モジュール/8-cysドメイン	フィブリリン、TGF β -結合タンパク-1
TNF受容体様	TGF-R、NGF-R、BAFF-受容体
血管内皮増殖因子のヘパリン結合ドメイン	
抗真菌タンパク (AGA-FP)	
フィブロネクチンI型モジュール	フィブロネクチン、組織プラスミノゲン活性化因子、t-PA
サイログロブリンI型ドメイン	
X型セルロース結合ドメイン、CBDX	
セルロースドッキングドメイン、docking	
カルボキシペプチダーゼ阻害剤	
無脊椎動物キチン結合タンパク	
フェロモンER-23	
軟体動物フェロモン	
尖頂膜抗原	
ソマトメジンBドメイン	
ノッチドメイン	
ミニコラーゲンI、C-末端ドメイン	
ホルモン受容体ドメイン (HRM)	
レジスチン	
YAPレドックスドメイン	
GLAFドメイン	
コレシストキニンA受容体N-ドメイン	
HIV-1 VPU細胞原形質ドメイン	
HIPIP (高電位鉄タンパク)	
フェレドキシンチオレドキシ還元酵素 (FTR)	
触媒性ペーザ鎖	
C2H2およびC2HCジンクフィンガー	
Zn2/Cys6 DNA結合ドメイン	
グルココルチコイド受容体様	
SBTDドメイン	
レトロウィルス・ジンクフィンガー様ドメイン	
ルブレドキシン様	
リボソームタンパクL36	
翻訳起動因子2 β の亜鉛結合ドメイン	
B-ボックス亜鉛結合ドメイン	
RING/U-ボックス	
Pyk2-関連タンパク β ARF-GAPドメイン	
メタロチオネイン	
酵母銅調節転写因子に保存される亜鉛ドメイン	
Ada DNA修復ドメイン	
システイン富裕ドメイン	
FYVE/PHDジンクフィンガー	
ADDBPのZn-結合ドメイン	
アポトーシス抑制 (IAP) 反復列	
CCCHジンクフィンガー	
DNAポリメラーゼ α のジンクフィンガー・ドメイン	
TAZドメイン	
システイン富裕DNA結合ドメイン (DM)	
DnaJ/Hsp40システイン富裕ドメイン	
CCCHドメイン	
SecCモチーフ	

【化 10】

タンパクファミリー	ファミリーにおける他の例示のメンバー
TSP3型反復列	

プロテアーゼ耐性マイクロタンパクの設計は、免疫原性を最小化するという観点から重要である。天然の、多くのマイクロタンパクは、プロテアーゼ阻害剤である。Rao, M. B. ら、(1998) *Molecular and Biotechnological Aspects of Microbial Proteases*. *Microbiol Mol Biol Rev.* 62 (3) : 597-635を参照されたい。生化学および分子生物学国際単位命名委員会によれば、プロテアーゼは、グループ3のサブグループ4（ヒドロラーゼ）に分類される。しかしながら、プロテアーゼは、その作用および構造の巨大な多様性のために、酵素命名の一般的システムには簡単に馴染まない。現在、プロテアーゼは、下記の三つの主要基準に基づいて分類される。(i) 触媒される反応のタイプ、(ii) 触媒部位の化学的性質、および(iii) 構造を基準とする進化的関係である。

10

【0071】

プロテアーゼは、その作用部位に応じて、大きく二つの主要グループ、エキソペプチダーゼおよびエンドペプチダーゼに分割される。エキソペプチダーゼは、基質のアミノまたはカルボキシ末端に近いペプチド結合を切断するが、一方、エンドペプチダーゼは、基質の末端から遠いペプチド結合を切断する。活性部位に存在する官能基に基づいて、プロテアーゼはさらに、四つの主要グループに分類される。すなわち、セリンプロテアーゼ、アスパラギン酸プロテアーゼ、システインプロテアーゼ、およびメタロプロテアーゼである。この標準的分類に正確に合致しない雑多なプロテアーゼが少数ある。例えば、活性にATPを必要とするATP依存性プロテアーゼがそれである。プロテアーゼは、そのアミノ酸配列に基づいて、種々のファミリーに分類され、さらに、共通の祖先から別れた、数組のペプチドを入れるために「クラン」に分類される。ペプチダーゼの各ファミリーには、触媒タイプを示すコード文字が割り当てられる。すなわち、セリン、システイン、アスパラギン酸、メタロ、または未知のタイプには、それぞれ、S、C、A、M、またはUが割り当てられる。

20

30

【0072】

エキソペプチダーゼ：エキソペプチダーゼは、ポリペプチド鎖の末端近くにおいてのみ作用する。それらの、NまたはC末端における作用部位に基づいて、それぞれ、アミノ、およびカルボキシペプチダーゼに分類される。

【0073】

アミノペプチダーゼ：アミノペプチダーゼは、ポリペプチドの遊離N末端に作用し、単一アミノ酸残基、ジペプチド、またはトリペプチドを開放する。

【0074】

カルボキシペプチダーゼ：カルボキシペプチダーゼは、ポリペプチドのC末端に作用し、単一アミノ酸、またはジペプチドを開放する。カルボキシペプチダーゼは、酵素の活性部位におけるアミノ酸残基の性質に基づいて、三つの主要グループ、セリン・カルボキシペプチダーゼ、メタロ・カルボキシペプチダーゼ、およびシステイン・カルボキシペプチダーゼに分けられる。

40

【0075】

エンドペプチダーゼ：エンドペプチダーゼは、NおよびC末端から離れた、ポリペプチド鎖の内部領域のペプチド結合に対し好んで作用を及ぼすという特徴を持つ。アミノ基またはカルボキシル基の存在は、酵素活性に対し否定的影響を及ぼす。エンドペプチダーゼは、その触媒機序に基づいて四つのサブグループに分けられる。すなわち、(i) セリンプロテアーゼ、(ii) アスパラギン酸プロテアーゼ、(iii) システインプロテアーゼ、(iv) メタロプロテアーゼである。

50

【0076】

ヒトのプロテアーゼ：カテプシンB、C、H、L、S、V、X/Z/P、および1は、パイン・ファミリーのシステインプロテアーゼである。カテプシンLおよびカテプシンSは、抗原提示細胞において抗原処理に関与することが知られている。カテプシンCは、別にDPP I（ジペプチジルペプチダーゼI）とも呼ばれる。カテプシンAは、セリンカルボキシペプチダーゼであり、カテプシンDおよびEは、アスパラギン酸プロテアーゼである。カテプシンは、リソソーム・プロテアーゼとして、タンパク分解において重要な役割を演じる。カテプシンは、ヒトおよび動物の腫瘍において再分布が行われるか、あるいはレベルの上昇があるので、侵入および転移においても役割を演じている可能性がある。カテプシンは、不活性な酵素前駆体として合成され、処理されて、成熟した、活性を有する酵素となる。シスタチン、およびいくつかのセルピンなどの内因性タンパク阻害剤は、活性酵素を抑制する。他のカテプシンは、カテプシンG、D、およびEである。

10

【0077】

それに対して耐性を持つようにタンパク剤を加工することが可能な、他のヒト・プロテアーゼとしては、トリプターゼ、キマーゼ、トリプシン、カルボキシペプチダーゼA、カルボキシペプチダーゼB、アジプシン／因子D、カリクライン、ヒトタンパク分解酵素3（Sigma）、トロンビンがある。

【0078】

さらに、主題のマイクロタンパクの設計に天然のHDDタンパクを用いてもよい。天然のHDDタンパクとしては、動物細胞表面受容体タンパクの数多くのファミリーの外に、防衛的（すなわち、嚥下吸収された）および攻撃的（注入可能な）動物トキシン、例えば、蛇、蜘蛛、さそり、腹足類、およびアネモネの有毒タンパクが挙げられる。これらのタンパククラスの共通項は、それらが、宿主－環境／病原体境界面に存在することである。上記、および、本明細書に記載される、他の、任意のタンパクは、本発明の、非天然システインスカフォールドを生成するのに使用することが可能な例示のスカフォールドとして役立つ。

20

【0079】

特に興味あるのは、この境界部（宿主および病原体の両方）には、その配列を急速に適応させることを可能とする特殊な分子支援システムを有する傾向のあるタンパクが存在することである。例としては、ナイセリアおよび他の細菌のピリン、脊椎動物の抗体システム、トリパノソームの可変表面糖タンパク、プラスモジウムの表面タンパク（実際はマイクロタンパク）、および他の多くの例がある。AA配列の急速な適応は、マイクロタンパクで明瞭に観察される。すなわち、マイクロタンパクの配列は、ゲノム配列の近似性から予想するよりもはるかにまちまちとなる傾向を持つ。プロテアーゼによる攻撃を防ぐ剛強構造（しかしながら、必ずしも同じ構造というわけではない）を保持しながら、急速に配列を適応させる能力が多分、このクラスのタンパクが、トキシンの開始物質として使用されるために、動物の進化とは独立に複数（7）回招集された理由であろう。このように繰り返し招集されたことは、このクラスのタンパクが、トキシン構築に特に有用な特性を提供することを示している。その他の、定常な特性としては、小さなサイズ（これらは最小のフォールド形成タンパクである）、および、プロテアーゼおよび温度に対する極端な安定性がある。

30

40

【0080】

受容体タンパクおよびトキシンは、急速な配列変動率を示し、このため、ごく近縁の腹足類トキシン同士が完全に無関係に見えるほどである。急速な進化は、トキシンの必須の特性であると考えられる。なぜなら、毒物は、多様で、常に変化する犠牲動物種における広範な受容体タンパクの変化（トキシンに対する耐性を実現するための進化速度の上昇を示す）に追従しなければならないからである。このグループの一つのきわめて有用な特性は、複数の公刊物に記載されるように、高いジスルフィド密度のスカフォールドのプロテアーゼ安定性によって付与される低い免疫原性である。これは、噛まれたが、逃げおおせた犠牲動物においてトキシンに対する耐性の創出を免れるのに重要と考えられる。受容体

50

とトキシンの両方が、急速に配列を適応させる必要があるのであるから、ある場合には、両方が、HDDミクロタンパク・ドメインから構成されるということになっても驚くべきことではない。例えば、蛇毒様タンパクの構造準拠クラス（タンパク構造分類（SCOP）データベースによって定義される）は、蛇毒トキシンの外、そのいくつかは同じ構造のリガンドと相互作用を持つ（すなわち、TGFベーター-TGFベーター受容体）、ヒトの細胞表面受容体の細胞外ドメインを含む。例示のタンパクとしては、蛇毒様タンパク、例えば、蛇毒トキシン、およびヒトの細胞表面受容体の細胞外ドメインが挙げられる。蛇毒トキシンの非限定的例としては、エラプトキシンB、ガンマーカルジオトキシン、ファシクリン、ムスカリニントキシン、エラプトキシンA、ニューロトキシンI、カルジオトキシンV4II（トキシンIII）、カルジオトキシンV、アルファコブラトキシン、長ニューロトキシン1、FS2トキシン、ブンガロトキシン、ブカンジン、カルジオトキシンCTXI、カルジオトキシンCTXIIB、カルジオトキシンII、カルジオトキシンIII、カルジオトキシンIV、コブラトキシン2、アルファトキシン、ニューロトキシンII（コブラトキシンB）、トキシンB（長ニューロトキシン）、カンドトキシン、ブカインがある。（ヒト）細胞表面受容体の細胞外ドメインの、非限定的例としては、CD59、II型アクチビン受容体、BMP受容体Iaエクトドメイン、TGF-ベータI型受容体細胞外ドメインが挙げられる。

10

【0081】

大抵の天然HDDタンパクファミリーでは、ジスルフィド・スカフォールド単独で高度の剛性を与えることが可能である。これは、誘発適合性、およびそれと関連するエントロピー・ペナルティを回避することによって高いアフィニティの獲得にとって有利である。多くのミクロタンパク・ファミリーにおいて、単に4、6、8、または10個のシステイン残基のみで、そのタンパクの構造、耐熱性、およびプロテアーゼ耐性などの主要な性質を完全に決定することを可能とし、一方、ループの他の残基の全て（例えば、コトキシンの場合のように）、またはほとんど全ては、結合特異性のために望ましい任意の配列を自由に採用可能とするようにしているようである。システインは、最小の配列定義（「低情報含量」）によって最重要機能を実現するが、これは、より多くの固定アミノ酸を含み、したがってより高い情報含量を有する別形スカフォールドよりも、このスカフォールドの方が統計的に優先的に独立して招集される理由となる。例えば、2個の余分な固定アミノ酸は、情報含量を増し、複数の配列から成るランダムプールから招集される、または該プールにおいて出現する予想頻度を $20 \times 20 = 400$ 倍下げる。非-cysアミノ酸に基づくタンパク安定性の同様レベルは、はるかに多くの残基を必要とし、その結果、より大型で、および／または、進化的にはより適応性の低いタンパクしか得られない。

20

30

【0082】

天然トキシンの構造的多様性の一つの起源は、HDD（高ジスルフィド密度）タンパクが、進化的時間スケールにおいて発揮することが示されている長さ変動によってもたらされる。このことは、蛇のジスインテグリンについて詳細に記載されている（Clavete, J. J., Moreno-Murcano, M. P., Theakston, R. D. G., Kisiel, D. G. and Marcinkiewicz, C. (2003) Snake venom disintegrins: Novel dimeric disintegrins and structural diversification by disulfide bond engineering. *Biochem J.* 372: 725-734. Clavete, J. J., Marcinkiewicz, C., Monleon, D., Esteve, V., Celda, B., Juarez, P. and Sanz, L. (2005) Snake venom disintegrins: Evolution of structure and function. *Toxicon* 45: 1063-1074)。

40

【0083】

大型HDDタンパクをコードする遺伝子の一部の欠失（または挿入／添加）は、元の配列と相同ではあるが、異なる構造と見なしてもよい、多数の、より小さい（またはより大

50

きい) 変動をもたらすことが可能である。公刊された例では、ジスルフィドの多くは保存されるが、少数のシステインは、新しい結合パターンを形成する。これに関する天然の機序は、DNAレベルにおける修飾、mRNA選択的スプライシング、分解、タンパクの(トランス-)スプライシング、または、いずれかの末端における、別形の短縮または付加、選択的翻訳の外、分解、または別形の短縮を含むことが考えられる。その天然の機序が何であれ、この原理は、最適能力および安定性、および最小サイズを有するタンパクを進化させるために、分子生物学および(ファージ)ディスプレイライブラリを用いて実地に実行することが可能である。

【0084】

さらに、下記の好ましいファミリー：Aドメイン、EGF、Ca-EGF、TNF-R、ノッチ、DSL、トレフォイル、PD、TSP1、TSP2、TSP3、アナト、インテグリンベータ、サイログロブリン、デフェンシン1に加え、本明細書に開示されるさらに別のファミリー、を含む天然のタンパク配列から、新規の、修飾スカフォールドを生成することが可能である。動物トキシンとして機能する、二つ以上のジスルフィドを有する、既存のタンパクドメイン・ファミリーとしては、下記の好ましいファミリー：トキシン1、2、3、4、5、6、7、9、11、12、デフェンシン1、デフェンシン2、シクロチド、SHKT、ジスイнтеグリン、ミオトキシン、ガンマーチオネイン、コノトキシン、Muコノトキシン、オメガアトラコトキシン、デルターアトラコトキシンの外、本明細書に掲載されるさらに別のファミリーが挙げられる。修飾スカフォールドは、システイン数、ジスルフィド結合パターン、間隔、第1から最終システインまでのサイズ/長さ、ループ構造(種々の固定残基またはサイズを持つ)、イオン結合部位(異なる位置、アミノ酸組成、およびイオン特異性)、性能関連特性(安全性、非免疫原性、ヒトのものにより似ていること、ヒトのものにより似ていないこと、温度安定性、プロテアーゼ安定性、疎水性指数、親水性アミノ酸のパーセンテージ、共融点、高濃度、特異的残基の不在などの処方性質、剛性、ジスルフィド密度、ライブラリー残基のパーセンテージ、ジスルフィド結合パターンの複合性などを含む)において、天然のものと異なってもよい。

【0085】

ある場合には、天然の多様性において出現するサブファミリーを反映することが有用であるが、これは、それぞれが異なるサブファミリーに属し、サブファミリー間の長さおよび配列差を反映する、特異的ループ設計の複数の長さ変動を、同じスカフォールドライブラリーに含めることによって(通常、別々のオリゴヌクレオチドを用いて)実行することが可能である。

【0086】

ある応用では、既存のスカフォールドの改良変種を生成することが有用である場合がある。例えば、LDL受容体A型ドメイン(「Aドメイン」)、またはEGFドメインの新規変種は、元のものに比べて改良されたスカフォールドをもたらすと考えられる、様々の比較的保存性の高い方法によって生成することが可能である。変種を修飾するには種々の方法が、例えば、N-末端をC-末端に切り替えることによって、Aドメインのシステインモチーフ(間隔を含む)だけを、または、Aドメインの保存残基(非-cysを含む)のモチーフを反転させることを含む方法が存在する。いくつかの小型ペプチドでは反転の実行が可能であることが示されているが、この場合、少数のアミノ酸だけが反転される。他の修飾としては、タンパクの長さを変えて(より短く、またはより長く)、公刊されたライブラリーの、または天然の配列の長さ範囲から逸脱させること、カルシウム結合部位を別組のループに移動すること、および、ループの固定非-cys残基の一つ以上を変えることを含めてもよい。固定残基がDである場合、目的は、この位置に非D残基を置くことになる。これを実行し、特異的アミノ酸位置について新規の、多数の組成物を試験する良い方法は、天然のアミノ酸の、または、公刊ライブラリーで使用されるミックスの反対(すなわち、相補的)となるアミノ酸ミックスを与えるコドンを使用することである。公刊ライブラリーが、ある位置にI、L、Vを含む場合、新規モチーフは、その位置にI

、L、V以外の20AA全てを与えることによって実現することが可能と考えられる。各位置は、構造に関し、機能に関してはさらに、アミノ酸要求において差を示す。

【0087】

スカフォールドのライブラリーはさらに、既存のスカフォールド配列モチーフについてより良い変種を探すために使用することも可能である。下記の局面：異なるジスルフィド結合パターン、および／またはジスルフィドの異なる間隔、および／またはループの異なる配列モチーフ、および／または固定ループ残基における違い、および／または、カルシウム結合部位の異なる位置、不在、またはAA組成、またはイオン特異性の内の一つ以上において、既知のスカフォールドよりも優れるスカフォールドを探し求めることが可能である。

10

【0088】

Aードメイン以外のスカフォールド、例えば、ドメインファミリーEGF、Ca-EGF、TNF-R、クニッツ、ノッチ/LNR/DSL、トレフォイル/PD/P-型、TSP1、TSP2、TSP3、アナト、インテグリンベータ、サイログロブリン、トキシシン1、2、3、4、5、6、7、9、11、12、デフェンシン1、デフェンシン2、シクロチド、SHKT、ジスイнтеグリン、ミオトキシシン、ガンマーチオネイン、コノトキシシン、Mu-コノトキシシン、オメガアトラコトキシシン、デルターアトラコトキシシンの外、表に掲載されるさらに別のファミリーを含むスカフォールドに対し、これらの原理をどのようにして適用したらよいかを、当業者であれば知っている。

【0089】

Aドメインから得られた、例示の、修飾および新規スカフォールドとしては、下記の配列：

20

C₁ (xx) E D s x D x C₂ D x x G D C₃ x W x x [p s] x C₄ (xx) x x x C₅ x F x x x (xx) C₆ プラスさらに一つのジスルフィド

を含む非天然配列(50aa未満)を有するタンパクドメインが挙げられる。例えば、3-ジスルフィドAードメインと類似するが、比較的屈曲性の大きいAードメイン構造を安定化する位置に余分のシステインを持つために、剛性がより大きい、いくつかの4ジスルフィドドメインがある。一例は、Aードメインの3SSフォールド(1-3 2-5 4-6)を含むが、そのAードメインのどちらかの側面にある1ジスルフィドによって該フォールドを安定化する1-8 2-4 3-6 5-7結合パターンである。Aードメインの、別の、高品質4-ジスルフィド版(「A+ドメイン」と呼ばれる)としては、1-5 2-4 3-7 6-8、1-3 2-6 4-8 5-7、1-4 2-7 3-6 5-8、1-4 2-7 3-6 5-8を始め、外にも多くのものがある。サイズは、ほんの数AAだけより長い(2-12、好ましくは8AA未満)だけで、Aードメインとほぼ同じでなければならない。この同じ分析法および解決法は、他の、全ての3-ジスルフィドファミリーに対し、さらに、下記の一般構造を持つ2-および4-ジスルフィドファミリーに対しても使用することが可能である。

30

【0090】

(非天然配列および50aa未満の)タンパクドメインで、C₁およびC₆の間に、C₁ x (xxx) x F x C₂ x x x (xxx) C₃ x x (xx) x x x C₄ D G x x D C₅ x D x S D E (xxxx) x C₆、および36を超えるaaを含むドメイン。

40

【0091】

(非天然配列および50aa未満の)タンパクドメインで、C₁およびC₆の間に、C₁ x (xxx) x F x C₂ x x x (xxx) C₃ x x (xx) x x x C₄ D G x x D C₅ x D x S D E (xxxx) x C₆、および32未満のaaを含むドメイン。

【0092】

非天然配列および50未満のaaを有するタンパクドメインで、1-3 2-5 4-6連結の3ジスルフィド、およびC₁とC₆の間に36を超えるaaを有するドメイン。

50

【0093】

(非天然配列および50aa未満の)タンパクドメインで、 C_1 および C_6 の間に、 C_1 x (x x x) x F x C_2 x x x (x x x) C_3 x x (x x) x x x C_4 D G x x D C_5 x D x S D E (x x x x) x C_6 、および32未満のaaを有するドメイン。

【0094】

非天然配列(50aa未満)のタンパクドメインで、 C_1 (x x) x x x x x x x x C_2 x x x x x C_3 x x x x x x C_4 (x x) x x x C_5 x x x x x (x x) C_6 (反転Aードメイン)を含むドメイン。

10

【0095】

下線アミノ酸の内の一つが存在しないタンパクドメイン(非天然配列および50aa未満)：

C_1 x [a p s] (x) [e k q] F x C_2 x x x x (x) C_3 [i l v] [p s] x x [l w] [l r v] C_4 D G [d e v] [p n d] D C_5 x D [d g n s] S D E (a p s) (l p s) x x C_6

同じ方法による異なる表示法は(三つの異なるモチーフレベルが示され；所望の変化には下線を施す)である。

【0096】

C_1 x (x x) x x x 非 F x C_2 x x x x (x x) C_3 x x x x x C_4 x x x x 非 D C_5 x (x) x x x 非 D 非 F (x) x x x C_6 、または

20

C_1 x (x x) x x x 非 F x C_2 x x x x (x x) C_3 [非 I L V] [非 P S] x x x x C_4 非 D 非 G x x 非 D C_5 x (x) 非 D x 非 S 非 D 非 E (x) x x x C_6

(非天然配列、および)、Aードメインフォールドと同じ結合パターンを有するが、システインについてはきわめて異なる間隔、および完全に無関係のタンパク配列を持つHuwentキシンIIフォールドを有するタンパクドメイン。

【0097】

二重配列を含まないドメイン・ファミリー：このクラスは、大抵の動物トキシンスカフォールド、および細胞表面受容体由来のスカフォールドを含む。蛇、蜘蛛、さそり、腹足類、およびいそぎんちゃく毒のタンパクトキシンは、天然の、注入可能な生体製剤と考えることが可能である。これらの毒は、通常、一定範囲の受容体および生物種特異性を有する、関連および無関連の、100を超える異なるトキシンを含む。これらのトキシンの大部分は、高密度のジスルフィドを有する小型タンパクである。典型的サイズは、2ジスルフィドでは15-22aa、3ジスルフィドでは25-45aa、4ジスルフィドでは35-50aaであるが、5、6、7、8、またはそれ以上のジスルフィドについても多くの例がある。例は、デルターアトラコトキシ(1-4 2-6 3-7 5-8)、さそりトキシ(1-8 2-5 3-6 4-7)、オメガアガトキシ(1-4 2-5 3-4 7-8)、マウロトキシ(1-5 2-6 3-4 7-8)、およびJアトラコトキシ(1-4 2-7 3-4 5-8)である。

30

【0098】

系統発生的分析から、これらのタンパクは、無関係の動物群が、無関係の開始点から独立に類似のトキシ構造を生成する、集約的進化の例であることが示されている。同じ設計原理が、少なくとも七つの独立の場合(それぞれ、無関係の分類学的群に属する)において支配的であることを考えると、この設計は、他のタイプのトキシ(すなわち、微生物タンパクトキシ)を構築するために用いられる他のスカフォールドよりも重要な利点を有することが期待される。

40

【0099】

これらのタンパクによって共有される唯一の特性は、高密度のジスルフィド結合である。これらのタンパクの(cys以外の)アミノ酸配列は、きわめて変動性が高く(コノトキシンの整列を参照)、広範囲の、種々の構造体(タンパクフォールド)が創製されてい

50

る。

【0100】

これらのタンパクの望ましい性質の一つは、それらの並外れて小さいサイズである。ミクロタンパクは、もっとも小さい剛強タンパクで、急速な組織浸透に必要とされる。第2の共通性質は、同じサイズの他のタンパクよりも高い、その剛性であり、これによって、これらのタンパクは、標的に結合したときに誘発適合を回避することができ、より高度な結合アフィニティーが可能となる。第3の特性は、これらのタンパクの並外れた安定性で、熱安定性（多くのミクロプロテインは、変性されることなく煮沸することが可能である）、および広範囲のプロテアーゼに対する耐性の両方を備える。安定性は、静脈内（IV）、または皮下（SC）に注入される生物製剤にとって重要であり、経皮的、鼻腔内、経口、腸内、血液脳関門を介して輸送されるタンパクにとってはさらに重要である。安定性はまた、長い半減期、好適な搬出・保存にとっても重要である。大いに興味をそそるもう一つの性質は、これらのタンパクの非免疫原性である。これは、抗原提示細胞（APC）におけるタンパク分解に対し、これらのタンパクの示す抵抗性によって仲介されることが報告されている。この抵抗性は、高ジスルフィド密度構造によって付与されることが公表されている。免疫原性を低く抑える他の因子としては、タンパクの小さなサイズおよび親水性がある。

10

【0101】

二重配列を含むドメインのファミリーはまた、主題のミクロタンパクおよびそのライブラリーの生成にも用いられる。多数の例が、下記の実施例に記載される。

20

【0102】

反復列を含むドメインのファミリー：システイン富裕反復列タンパク。システイン富裕反復列タンパクの高いシステイン含量によって、その反復列内において、および／または二つの反復単位間において、複数のジスルフィド結合の形成を可能とする。このために、ジスルフィド結合の反復パターンが得られる。このパターンは固定された空間配置を与える。ただし、稀な場合であるが、同じ配列が、別形のジスルフィド結合パターンを取ることもある（または、取るように進化することも可能である）。反復タンパクにおけるジスルフィド結合は、 $CRRP$ モチーフ（ X_{A1} , X_{A2} ）／（ X_{B1} , X_{B2} ）／（ X_C ）によって特徴づけられる。前式において X_A は、連結システイン間のシステイン距離であり、これは、同じジスルフィド結合における第1システインと第2システインの間のシステイン数である。このシステイン距離は、1、2、3、4、5、6、7、8、9、または10であってもよい。 $CRRP$ モチーフにおける2（またはそれ以上）の数は、二つ（またはそれ以上）の異なるタイプの結合を示し、 X_{A1} はそのような第1結合を表し、 X_{A2} はそのような第2ジスルフィド結合を表す。例えば、1-4 2-3 空間配置を有する $CXCXCXCXCXCXCXCXCXC$ は、第1ジスルフィド結合タイプについては+3のシステイン距離、第2ジスルフィド結合タイプについては+1のシステイン距離を有する（「3、1」）。

30

【0103】

X_B は、一ジスルフィド結合の第1システインから、次のジスルフィド結合の第1システインまでのシステイン距離（システイン数）を表す（例えば、1-4 2-3の空間配置を有する $CXCXCXCXCXCXCXCXCXC$ の場合、 X_B は+1である。二つの異なるタイプのジスルフィド結合 X_{B1} は、一タイプのジスルフィド結合の第1システインから、隣接ジスルフィド結合の第1システインまでのシステイン距離を表し、一方、 X_{B2} は、この場合では、次の反復列に配される第2タイプのジスルフィド結合の第1システインから次のジスルフィド結合の第1システインまでのシステイン距離を表す。この例では、 X_{B2} は+3である（C2からC5まで）が、 X_{B2} は1、2、3、4、5、6、7、8、9、10となることも可能である。 X_C は、螺旋反復タンパクにおいて螺旋転回当たりのジスルフィド結合数を表し、1の分数か、または、例えば、1、2、3、4、5、6、7、8、9、10のような整数となることも可能である。

40

【0104】

50

各ドメインは通常（必ずしもそうとは限らないが）、N-および／またはC-末端に末端キャップを有する。末端キャップは、通常、正規の反復列よりも一つまたは二つ少ないシステインを持つ。なぜなら、末端キャップは、二つの反復列ではなく、一反復列のみに接続しなければならないからである。

【0105】

反復列に関するさらに詳細な記載は、タンパク中の、各タイプのジスルフィド結合の「スパン」（二つの連結システイン間の非-cysアミノ酸数）を含むと考えられる。反復タンパクを表すもう一つの方法は、反復単位の配列を、例えば、 $(C_x x x C_x C_x x x C_x x C C x x)_n$ と表すことである。どのシステイン同士が連結されるかを表示するのにC_aおよびC_b表記を、例えば、 $(C_a x x x C_a x C_b x x x x C_c x x C_b C_c x x)_n$ に見られるように使用することが可能である。

10

【0106】

システイン富裕反復タンパクの重要な特性は、どちらの末端でも、N-末端またはC-末端のいずれにおいても延長が可能であるということである。ライブラリー設計のための二つの方法として、1)天然の反復タンパクのランダム化、および2)通常、天然の反復タンパクからの抽出によって得られるが、天然の反復列とは幾分異なる間隔（より理想化された）を有する合成反復列がある。天然のCRRPとしては、グラニュリン（PF00396）、昆虫抗凍結タンパク（PF02420）、フリン様ドメイン（PF00757）、CxCxCx反復列（PF03128）、パラメシウム表面抗原（PF01508）、および機能が未知のショウジョウバエドメイン（PF05444）が挙げられる。

20

【0107】

要すれば、主題のシステイン含有タンパクおよび／またはスカフォールドは、生体反応修飾因子と融合されてもよい。生体反応修飾因子の例としては、蛍光タンパク、例えば、緑色蛍光タンパク（GFP）、サイトカインまたはリンフォカイン、例えば、インターロイキン-2（IL-2）、インターロイキン4（IL-4）、GM-CSF、およびγ-インターフェロンが挙げられるが、ただしこれらに限定されない。その他の有用な融合配列として、精製をやり易くするものがある。そのような配列の例は、従来技術で既知であるが、例えば、Myc、HA（インフルエンザウィルス血球凝集素由来）、His-₆、またはFLAGなどのエピトープをコードするものが挙げられる。精製をやり易くする他の融合配列が、例えば、グルタチオンS-トランスフェラーゼ（GST）、マルトース結合タンパク（MBP）、または、免疫グロブリンのFc部分などのタンパクから得られる。

30

【0108】

ライブラリーの構築：本発明は、主題のシステイン含有スカフォールドのライブラリーを提供する。天然の選択を受けるタンパクは、均一に折りたたまれる必要があるが、新規の、進化していない配列を有するタンパクは、折りたたまれて複数の安定構造を形成することが可能であるか、または、条件を変えることによって少なくともそのように誘起することが可能である。同じタンパク配列の異なるコピーを折りたたんで異なる安定構造を形成させることは、ライブラリーの構造的多様性を、ライブラリー中の独立クローンを超えてさらに拡大する。ライブラリー中の独立クローンの数は、一般に、異なる配列の数に等しく、「ライブラリーサイズ」と呼ばれ、ファージディスプレイ・ライブラリーでは約 10^{10} である。しかしながら、ファージライブラリーのパンニング選択時に使用されるファージ粒子の実際数は、通常、ライブラリーサイズよりも $10-10,000$ 倍大きい。過剰倍数は、「ライブラリー当量数」と呼ばれるが、より大きなライブラリー性能を獲得するためにこの差を利用する方法がある。もしもクローンの $10-10,000$ コピーのそれぞれが（すなわち、同じアミノ酸配列を有する全てが）、異なる、安定なDBPおよび構造を取ると仮定すると、一過性に異なる構造を取る不安定構造を用いて、構造多様性をさらに増すことが可能である。一方、もしも各ファージ粒子が、その利点および欠点も含めて、ランダムペプチドと同様に、広く多様な構造を取ることが可能な不安定タンパクも表示するとするならば、多様性はさらに増加する可能性がある。多数の不安定構造を

40

50

取ることが可能なタンパクは、ファージ粒子の数 ($10^{12} - 10^{15}$) を超えて多様性を拡大する。低アフィニティークローンの回収は、大きな数のライブラリー当量を必要とするが（すなわち、1%回収効率においてクローンを回収するためには約100ライブラリー当量が必要である）、高アフィニティークローンの回収は、100%効率に近く起こるので（アフィニティークロマトグラフィーによって実証されるように）、構造的多様性の増加は、高アフィニティークローンの割合を大きく増すことが期待される。不安定構造を伴う構造的多様性の増加には妥協点がある。なぜなら、標的結合の際、表示タンパクにおける構造誘発の必要（標的ではなく、結合するタンパクの方の誘発適合）は、これらのクローンの結合アフィニティーを下げることで予想されるからである。

【0109】

一つの方法は、4システイン（最大2ジスルフィド、および最大3通りの結合パターン）、6システイン（最大3ジスルフィド、および最大15通りの異なるジスルフィド結合パターン）、8システイン（最大4ジスルフィド、および最大105通りの結合パターン）、または10システイン（最大5ジスルフィド、および最大945通りの結合パターン）、あるいは12、14、16、18、20、またはそれ以上のシステインを有するライブラリーを構築することである。

【0110】

一局面では、ジスルフィド結合パターンの全数は、下記の式によって一般化することが可能である。

【0111】

式: `Error! Objects cannot be created from editing field codes.` 式中、 n = システイン残基によって形成されるジスルフィド結合の予測数であり、かつ、`Error! Objects cannot be created from editing field codes.` は、 $(2i-1)$ の積を表し、この式において i は、1 から n までの正の整数である。

【0112】

要すれば、大きい、変動する数（すなわち、10-30）のシステインをコードする、はるかに大きな構築体を生成することも可能である。得られたシステイン含有産物は、広く多様なやり方でフォールド形成し、それぞれが2、3、4、または5ジスルフィドを含む構造要素が、それぞれの間に交差性架橋の可能性のある、異なる組み合わせを創出する。これら比較的大型の構築体の指向性進化過程の際、以前に選択された構築体を、例えば、ランダム断片化、PCR（例えば、ランダムプライマーによる）、または（例えば、4bp）制限消化によってより小型の断片に破断することも可能である。長いタンパクのライブラリー多様性が低下した場合、各大型構築体から種々の断片を創製し、その後、組み換え、または他の指向性進化法によって再度多様性を増すことが可能である。

【0113】

HDDタンパクから成るこのようなライブラリーに関する、一つの予想される不安は、ジスルフィドの多くが形成された後の、対合されない複数のシステインの存在である。遊離チオールは、互いに相互作用を持ち、凝集体を創成するが、この凝集体は、標的に対する多価性結合のために、ブロッキングアッセイにおいてははっきりと高いスコアを呈する傾向を持つ。しかしながら、これらの遊離チオールは、例えば、イオドアセタミド、すなわち、スルフヒドリルに対する周知の遮断剤によって遮断され、凝集体を形成することがないよう、または適正に形成されたジスルフィドを攻撃することがないよう阻止することが可能である。

【0114】

同数のジスルフィドを有する（すなわち、三つのジスルフィドは、15通りの可能な連結パターンを与える）、マイクロタンパクの複数のファミリーの共通配列を整列することによって、システイン間の間隔は、0 から約12アミノ酸のほぼ均等な分布を形成することが明らかにされる。簡単のためと、平均ループ長を小さく保持するために、我々は、システイン間ループ当たり0-10アミノ酸を有するファミリーを好む。

【0115】

合成オリゴヌクレオチドを用いて、DNAが、6個のシステインと、システイン間ループの0-10NNK（または類似のあいまいコドン）残基をコードするようにライブラリーを構築することが可能である。NNKコドンは、20aa全てをコードするが、1/64コドンだけは終止コドンとなる（NNNコドンを用いた場合よりも3倍少ない）。これは、早すぎる終止コドンを含むタンパクの割合を下げる。5個のシステイン間ループを仮定すると、これらのタンパクは、平均25NNKコドンを含むことになり（ループ当たり0から10aa、平均5を仮定する）、これは、早すぎる終止コドンを含むクロンの割合を下げる。終止コドンを除く、10よりも少ない数のコドン、またはあいまい（混合塩基組成）コドンを用いることによって、完全タンパクの割合を増すことが可能である。図に示すように、各オリゴヌクレオチドは、システイン・コドンから開始し、かつシステイン・コドンで停止し（一端ではセンス方向、他端ではアンチセンス方向）、これらのシステイン・コドン間には0-10NNK（または逆方向）コドンがある。この合成ライブラリー作製法では、ループ配列の全てが、任意のループ位置において使用することが可能であり、したがって、システインは全て、通常、同じコドンによってコードされる。これらのオリゴは全て混ぜ合わされ、合成遺伝子プールが、前述のように（Stemmerら、1995. Gene）重複PCRによって創製される。

10

【0116】

ファージライブラリーを創成するための、別の、強力な方法は、Kunkel突然変異法のScholleの変法である（Scholle, M. ら、(2005) Comb. Chem. & HTP Screening 8:545-551）。この方法では、ライブラリーコード・オリゴヌクレオチドによって、プラスミド中の終止コドンが、非終止コドンに変換される。この方法の新規バージョンとして、任意の二つの終止コドン（通常、アンバーコドンとオークルコドン）の間の往復サイクルが挙げられる。これによって、突然変異発生各サイクル毎に終止コドンを再挿入する手間を要することなく、進化するクローンプールに対しScholle法を反復帰納的に適用することが可能となる。

20

【0117】

3SS（3-ジスルフィド；15通りの可能な構造）、および4SS（105通りの可能な構造）混合のスcaf-foldライブラリーは特に有用である。ジスルフィド結合パターンに関して我々が所有する主な調節手段は、システインの間隔である。タンパクがどの構造（ジスルフィド結合パターン、‘DBP’）を取るかは、例えば、f-fold再形成のために、ある一定範囲の環境を提供することによってある程度調節することが可能である。DBPは、トリプシン消化および／またはMS/MS分析によって分析することが可能である。

30

【0118】

構造的多様性の問題点は、複数scaf-foldライブラリーでも、単一scaf-foldライブラリーでも同じであり、大きさの差は連続的に調整することが可能である。既存の天然ファミリーから多少変動し（ループ当たり平均0から15アミノ酸）、かつ、多少類似するが、システイン間隔に基づくライブラリー設計には連続性がある。通常、単一scaf-foldライブラリーも、相当の長さ変動（天然の変動を模倣する）を含む。配列類似性によって複数のファミリーが創製されること、および、通常、僅かな少数のメンバーについてのみ、構造（結合パターン）が実験的に決定されることに注意されたい。したがって、かなりの数の天然配列が、配列から予想されるものとは別の構造を持つ可能性がある。天然の、高度に進化し、微細に調整された（すなわち、高い情報含量の）配列は、一般に、高い信頼度で一定のやり方でf-fold形成するが、しかし、低い情報含量の、調整度のより低いタンパク（例えば、初期のファージディスプレイ・ライブラリーのもの、および／または、1サイクルのpanning後で、指向性進化前の構造的に多様なライブラリーから得られたもの）は、しばしば、いくつかの異なるf-fold形成をすることが予想される。

40

【0119】

50

タンパクの、ある特異的な、天然ファミリーの保存スカフォールドに基づくライブラリーは、IgドメインまたはフィブロネクチンIIIのように、通常、各種の問題点（すなわち、ヘテロのフォールド形成、フォールド未形成、凝集、または低レベルの発現）を抱える約5-10%のクローンを含む。長さ多様性を増すこと、より大きな配列および構造多様性を許容することは、より振る舞いの低劣なクローンを生成する可能性がある。突然変異発生の新たなサイクルを適用する前に、不要なモノマーを選択排除し、ダイマーおよびより高次のマルチマーを作製するのが一般的である。しかしながら、指向性進化は、非適正クローンの振る舞いを向上させる点できわめて有効である傾向を持つので、指向性進化によって、クローン排除によって、および／または、配列変更および／または構造変更によって、クローンプールの平均品質を徐々に改善することが可能である。指向性進化は、改良活性を選択するようにスクリーニングするが、改良フォールド形成は、活性を改善するための簡便な方法であるので、活性の指向性進化は、タンパクフォールド形成効率の向上（Leong, S. R., ら、(2003) Proc. Natl. Acad. Sci. USA 100:1163-1168; Cramer i, A. ら、(1996) Nature Biotechnology 14:315-319）、および温度安定性の向上（多くの報告例）を実現するための、実証済みの、効率的な方法である。理由は、活性構造をより効率的に採用するクローンは、より活性が高く、したがって、選択過程においてより好まれるからである。我々の目指す過程は、パンニングの初期ラウンドは、種々のフォールドを有するたくさんのクローンを生成し、一方、高レベルの様々な問題点（不完全なフォールド形成、ヘテロのフォールド形成、低レベル発現、凝集など）が見られる可能性の高いものであるが、（ファージ）パンニングによる強力な機能的選択と組み合わせて、指向性進化（誤作動性PCR、相同組み換え、カセットによる組み換え、または場合によっては単純に複数回のスクリーニングを含む、多くの可能な方式）を用いることによって、均一なフォールド形成を持つクローンを積極的に選び出すことが予想される過程である。均一なフォールド形成をするクローンの頻度を高めるために、同じライブラリーを、複数回（ファージ増幅付きで、または無しで）、改めて還元、フォールド形成、およびパンニングを行うことも可能である。不完全なフォールド形成のタンパクを除去するために、各サイクル毎に、遊離チオール・アフィニティーカラムを用いることが可能であり、あるいは、遊離チオールを、各種キャップ剤（FITC-マレイミド、イオドアセタミド、ヨード酢酸、DTNBなど）と反応させることが可能である。さらに、遊離チオールの頻度を下げるために、全体ライブラリーを再度フォールド形成することも可能であるし、あるいは、部分的に還元、および再度酸化することも可能である。ファージディスプレイおよび可溶性タンパク結合アッセイは、しばしば、多価の溶液を積極的に選ぶ。タンパク間ジスルフィドを有するタンパクは、多価性の一般的供給源であるが、これらは製造不能なので除去する必要がある。ファージディスプレイを複数回行う（その間に可溶性タンパクのアッセイを行うことなく）ことは、ファージ上においてのみ作動する溶液を進化させる傾向がある。したがって、これらのクローンが支配的になることを阻止するために、一般に、可溶性タンパクのスクリーニングは望ましい。タンパク構造の多様性は、初期には望ましいが、タンパク間ジスルフィド結合を形成するクローンは、その除去を次第に増やすことが望ましい。構造の多様性は、不定のフォールド形成、およびタンパク間ジスルフィドと相関し、かつ、構造進化は、不均一フォールド形成と切り離すことができないので、ある程度の不均一性を許容する方法を開発する必要がある。

【0120】

構造の多様性とフォールドの均一性に関する所望のバランスについて種々のライブラリー設計を評価するために、ライブラリー設計の多様性を手早く評価する目的で、小型のライブラリーを作製し、少数のクローン（30-1000）をスクリーニングすることが可能である。

【0121】

同じタンパクの中の種々のジスルフィドは、別様に反応するので、ある程度の調節を可能とする。タンパク間ジスルフィドを有するクローンをファージライブラリーから除去す

る方法の内の一つは、そのファージライブラリーを、低濃度の還元剤に暴露することである。この還元剤は、もっとも弱いジスルフィド、例えば、きわめて弱いので、クローンから除去したいと思う、タンパク間ジスルフィドおよびタンパク内ジスルフィドのみを還元するので、次に、この部分的に還元されたライブラリーを遊離チオールカラムの上に通過させると、これらのクローンは除去される。

【0122】

(HDDタンパクの構造進化)

上述したように、HDDタンパクは、各レベル、例えば、一次（配列）、二次（アルファヘリックス、ベータ・シートなど）、三次（フォールド、ジスルフィド結合パターン）、および四次（他のタンパクとの連結）を含むレベルにおけるタンパクの構造進化を受容することが可能である。三次構造を完全に変えることができれば、それは、HDDタンパクを、治療薬または製薬組成物に関する原理的設計にもっとも適応するものとするようになる。既存の指向性進化法でも、限られたものではあるが、二次的構造進化（アルファヘリックス、ベータ・シート）が起こる場合があるが、三次構造において高品質の修飾を創出することは、指向性および原理設計では実行が困難であった。

【0123】

2SSから、ジスルフィド付加によって3SSからさらに4SSへと進化させること、および、欠失による逆転は、頻繁に起こるようであり、蛇ジスインテグリンに関しては報告されてもいる（Calvete, J. J. ら、(2003) Biochem. J. 372: 725-734）。天然のファミリーのDBP同士の近縁性は、天然でもDBPの再構成が起こる可能性のあることを示唆しており、特異的ファミリー、例えば、ソマトメジンについては公刊物によって支持されている。

【0124】

15種の異なる3SS構造、105種の4SS、または945種の4SS構造は、空間配置的に異なるが、これは、これらの構造は、相互変換が可能となるためには、破壊して、ジスルフィド結合を再形成しなければならないことを意味する。各3SSタンパクは、「最近接」変種である、6種の（完全）ジスルフィド結合異性形（変更結合パターンを有する2ジスルフィド、保存された結合パターンを有する1ジスルフィド）を有し、各4SSタンパクは、それぞれ2個の保存ジスルフィド、および2個の変更ジスルフィドを有する12種の最近接異性形を有する。このようにして、徐々に、構造進化のための経路が創出される。

【0125】

構造の指向性進化過程は、先ず、多様な構造（必ずしも全てが可能とは限らず、かつ、頻度も異なる）を奨励し、次いで、徐々に構造を締めつけるばかりでなく、構造を部分的に修飾し（緩慢なDBP変化によって）、一方では、だんだんより優れた結合因子を選択する。構造の、最初の大きな多様性は、異なるAA配列数を超えて、ライブラリーの実効サイズを拡大するのに役立つ。しかしながら、構造が多様になればなるほど、それらのフォールド形成は不均一となり、そのため、これらのタンパクは、有用となるほどの均一なフォールドを形成するためには相当の進化を必要とする。最適ループ長を持つ構造は、より均一なフォールドを形成し、プロテアーゼ耐性がより高く、免疫原性がより低い。場合によって見られる特定位置を除き、ループの配列は、三次構造に影響を及ぼすようには見えず、かつ、ループは概して二次構造を持たない。

【0126】

ループ長を最適化する好ましい方法は、比較的長いループ（すなわち、6、7、8アミノ酸）から始め、それから徐々にその長さを短縮し、各ループを、ある範囲の、異なるサイズ（低い平均サイズ）の他のループによって置換することである。この過程は、結び目の締めつけに似ている。ループの位置は、通常一定とされるが（すなわち、C2-C3）、それらの位置は、特に、タンパクにおける、複数の小さい結合部位が有用な解決策である場合には、変動させることも可能である。

【0127】

一つの好ましい方法は、選ばれたクローンのプールにおけるループ（すなわち、ループ C 1 - C 2、C 2 - C 3、C 3 - C 4、C 4 - C 5、C 5 - C 6、C 6 - C 7、または C 7 - C 8、C 8 - C 9、C 9 - C 10）を、多くはランダム配列からなる、これまで一度も選択されたことのない新しい 1 組のループで置換することである。異なるシステインに対し、また、要すれば、それらのシステインに側接する二三の固定塩基に対し別のコドンを用いることによって、PCR 重複反応（好ましい）においてループ交換を実行するための PCR 部位を創出することが可能であり、あるいは、制限部位法を用いることも可能である。

【0128】

あるタンパク標的に結合するように選ばれた、プール中の異なるクローンは、そのタンパクの異なる部位に結合する可能性が高い。同じ部位に結合するために近似配列を用いる場合でも、それらのクローンは関与部分が異なる可能性が高い。例えば、あるクローンは、ループ 1 に活性配列を持ち、別のクローンはループ 5 に持つことがある。より多くの固定アミノ酸を持つことは、同じ関与部分を有するクローンがより多いことになり、これは、相同組み換えによる指向性進化のためには有利と考えられる。

【0129】

選ばれたクローンプールに対し組み換えを実行するにはたくさんの方法がある。大抵の方式では、ループはそのまま手をつけず、相互に組み合わされるだけであるが、相同組み換えを駆動するのに、ループ間の相同性が用いられる方式もある。一般に、各ループは、同じ位置に留まるが（すなわち、C 4 - C 5）、これも変動させてよい。ある方式では、選ばれたクローンプールのループが全て連結されておらず、次に改めて連結されるが、もっと保存的な方法は、一つの特定のループ（すなわち、C 4 - C 5）のみを未連結とし、他のループは連結状態に保持し、たくさんの交差ではなく、僅かに 1 - 2 個の交差を有するクローンライブラリーを創製することである。この目的は、たくさんの保存的変化の組み合わせを必要とする、たくさんの、異なる、緩慢な経路を創製することである。

【0130】

たくさんのフォールドを有するライブラリー、または、ただ一つのフォールドを有するライブラリーを作製するのではなく、比較的少数の構造体（すなわち、最低限界が 2、5、10、30、100、300、最高限界が 10、30、100、300、1000、3000）で、その結合パターンは剛強構造を生ずるか、または天然ファミリーに出現するもので、最良のシステイン間隔に関する詳細な情報を提供する構造体を可能とするように設計された、間隔に僅かな変動性を有するライブラリーを作製できることが望ましい。一例は、c x x x (x) c x x c x x x x (x x) c x x x c x x x (x) x x c x x x x (x) c x x x c である。

【0131】

ライブラリーの実効的多様性および品質は、共にきわめて重要であるが、正反対の設計要求を持つ傾向がある。品質は、主に、適正にフォールド形成するクローンの割合によって決定される。ライブラリーの理論的多様性（よりランダム化された A A 位置）の展開は、非フォールド形成クローンの割合を増やす傾向を持つ。フォールド形成を増す工程は、各 A A 位置において天然 A A を使用すること、および天然の保存残基を保存することを含む。これは、単一スカフォールドライブラリーでは簡単に実行されるが、複数スカフォールドライブラリーではそうではない。したがって、複数スカフォールドライブラリーでは、非フォールド形成クローンの割合がより高くなる。フォールド形成のために固定化される必要のある 2 A A だけをランダム化すると、フォールド形成クローンの割合は、400 倍低下し、実効ライブラリーサイズを下げる。

【0132】

種々のライブラリーを創製し、FITC-マレイミドを用いて残存する遊離チオールを割合を測定（反応、洗浄、結合 FITC の測定）することによって、フォールド形成クローンの割合を測定することは有用である。さらに、遊離チオールを含む固相支持体を用いてフォールド未形成クローンを除去すること、および／または、全体ライブラリーまたは

フォールド未形成クローンのフォールドを再形成することも有用である場合がある。一つの方法は、部分的にフォールド形成した、または劣悪なフォールド形成したタンパクは還元するが、安定にフォールド形成したタンパクは還元しないと期待される濃度の還元剤にライブラリーを暴露することである。

【0133】

しかしながら、劣悪なライブラリー設計でも、はるかに低い、あるレベルのフォールド形成クローンを有することがある。一つの方法は、たくさんの、単一スカフォールドライブラリーを別々に構築し、パンニングの前にそれらを混ぜ合わせることである。これによって、高品質の、多様なライブラリーが得られるはずである。

【0134】

ヘテロ性フォールド形成は、適切に処理すれば利点となるべきものである。通例のライブラリーはサイズが $10 - e8 - 10 e9$ であり、また、約 $10 e13$ 個のファージ粒子を創出するから、各配列は、 $10 e4 - 10 e5$ 個の粒子によって提示される。もしもパンニングが 100% の効率で実行される（すなわち、各 $1 nM$ 以上のクローンが全て捕捉される）とすると、各配列を、 $10 e3$ 個の異なる構造として所有することは、有効な多様性、的中率、および品質にとって莫大な利点となるはずである。効率的パンニングは、高濃度のファージ、高濃度の標的、高温（より速やかな平衡）、 $10 - 15\%$ のポリエチレングリコール（PEG）などの容積排除剤、可溶性標的・対・固定標的などを必要とする。

【0135】

タンパクの適正なフォールド形成を促進するための、一つの方法は、PEGなどの容積排除剤の存在下にフォールド形成（最初に）をすることであってもよい。容積排除剤は、オリゴヌクレオチドのハイブリダイゼーション速度、さらにはシャッフリング反応（複合断片重複PCR）の効率を急激に増す。PEGは単にチオールの実効濃度を増し、これによってより多くの鎖間、および鎖内ジスルフィドが増す。

【0136】

一般に、フォールド未形成クローンは不要であるが、ヘテロ性フォールド形成は望ましい。未フォールド形成およびヘテロ・フォールド形成は明らかに相伴っている。他の場合は未フォールド形成のクローンが標的誘発性にフォールド形成するのは特に有用であるが、めったに起こらない出来事である。混合スカフォールドライブラリーの実効ライブラリーサイズに予想通りの減少が見られるため、効果的な突然変異発生策が一般に好まれる。組み換えを選んでもよいし、あるいは、長さ変動と点突然変異の両方を選んでもよい。ランダムライブラリーから得られた配列の組み換えは困難である場合がある。このような短いクローンに対しては、誤作動性PCRの誤作動率はやや低く（ 0.7% ）、再クローニングを必要とする。再合成は、選ばれたクローンの配列決定、ライブラリーの再合成、および再クローニングを必要とする。別法として、大腸菌の、主題の突然変異株に対し、適正にフォールド形成されたクローンを選択的に採取するために、多数サイクルのパンニングおよび増幅を行うことも可能である。さらに、EvoGenixの方法を応用することも可能である。

【0137】

2-3-4法の魅力は、PCRによる各ステップでランダム配列を付加するが、他の形の突然変異発生を必要としないことである。マイクロタンパクは、新規または既存のペプチドリガンドまたはタンパク断片から建設することが可能である。この方法は、あらかじめ結合性を有する、または有しない、短いアミノ酸配列を利用する。結合性アミノ酸配列は、その一端または両端において、単一システインをコードするランダムまたは固定アミノ酸によって側接されてもよい。オリゴヌクレオチドが、この結合配列および側接システインコードDNAをコードするように設計される。新たに導入されたシステインは、任意に、ランダムまたは非ランダム配列によって側接されてもよい。システイン含有側接配列の変動は全て、混合され、集められ、2本鎖DNAに変換される。これらの集合配列は、任意に、制限酵素認識部位をコードするか、または、あらかじめ存在するDNA配列にアニ

10

20

30

40

50

ールするDNAによって側接されてもよい。この方法は、新規または既存のシステイン距離パターンを生成することが可能である。

【0138】

(システイン富裕反復タンパク (CRRP))

甲虫 *Tenebrio molitor* 由来のシステイン富裕反復抗凍結タンパクは、C-末端において延長が可能であることが示された (C. B. Marshall, ら、(2004) *Biochemistry*, 43:11637-46)。この延長は、CRRPモチーフ1/2/1を含む。螺旋状であるが、ベータ・シートを含む(「ベーターヘリックス」)抗凍結タンパクの極端な規則性は、抗凍結活性と、氷結合部位の面積との関係を調べるために系統的に探索された。ベータ・ヘリックスの、ジスルフィド結合の中心コイルである、12-アミノ酸のそれぞれが、Thr-Xaa-Thr氷結合モチーフを含んでいる。7コイルの親抗凍結タンパクにコイルを付加し、かつ、該タンパクからコイルを取り去ることによって、6-11コイルを持つ、一連の構築体が作製される。これらの抗凍結体のミスフォールド形は、氷アフィニティー精製によって除去し、各構築体の比活性を正確に比較した。比較される濃度に応じて、6から9コイルに進むと、抗凍結活性に10-100倍の増加が見られた。

10

【0139】

我々の関心は、保存性のもっとも低いアミノ酸位置においてランダム化され、選ばれたヒトの治療標的に対する結合剤(作用剤または拮抗剤)の選択に使用される、複数の反復列を有する抗凍結タンパクを作製することである。

20

【0140】

グラニュリン(図102および103)は、3/2/2なるCRRPモチーフを持つ天然のCRRPである(ヘリックス、図130-132参照)。各反復単位は、高いモジュール性を持ち、したがって、C-末端に複数の反復列を付加することによってコア単位を延長するのに有用であることを示す証拠が提出されている(D. Tolkatchev ら、(2000) *Biochemistry*, 39:2878-86; W. F. Vranken, ら、(1999) *J Rept Res*, 53:590-7)。空気酸化されると、鯉グラニュリン1の、30残基のN-末端サブドメインに対応するペプチドは、天然タンパクに観察されるジスルフィド対合を自発的に形成した。NMRによる構造解明から、このペプチド中には、定められた二次構造の存在が示された。このペプチドの構造計算から、ペプチド断片が、天然ペプチド内に形成されるものと同じ立体配置を取ることが示された。鯉グラニュリン-1の、30残基の、N-末端ペプチドは、二つのヘアピン間ジスルフィド結合によって補強された、二つのベータ・ヘアピンが独立にフォールド形成した積層体の最初の例である。

30

【0141】

我々の関心は、保存性のもっとも低いアミノ酸位置においてランダム化され、選ばれたヒトの治療標的に対する結合剤(作用剤または拮抗剤)の選択に使用される、複数の反復列を有するグラニュリン由来タンパクを作製することである(図102)。

【0142】

反復性タンパク構造およびアフィニティー熟成：CRRPの利点は、それらが、他の多くのドメインと違って、特定の用途に必要なだけ長くも、短くも作製することが可能であることである。したがって、これらには、同じ、または異なる標的のために、1、2、3、4、5、6、7、8、9、10、またはそれ以上の結合部位を付与することが可能である。

40

【0143】

ロイシン富裕タンパクおよび、他の、非システイン含有反復タンパクに対し、CRRPの有利な点は、ライブラリーにおいて、より多くのアミノ酸のランダム化が可能であることである。これは、CRRPのフォールド形成が、たくさんの、より多くの固定残基を必要とする疎水性コアの存在ではなく、むしろジスルフィド結合の存在に依存するためである。したがって、CRRPのライブラリーは、より多くの変動性位置を持つ(>50、6

50

0、70、または80%) クローンを含み、これは、可能な表面接触面積を増し、かつ、標的に対する高いアフィニティーの可能性を増す。ロイシン富裕反復性タンパク、例えば、アンキリンは、通常、末端キャップがランダム化されないために、各33AA反復列の内の僅かに6AA、あるいは、6-反復列ドメイン当たり24AAにおいて変動するだけである。

【0144】

各種アフィニティー熟成法が、図140、141、142、および160に示される。これらのアフィニティー熟成原理は、反復性タンパクについて説明するのがもっともよいが、本出願に記載される他の、全てのスカフォールドにも同様に適用が可能である。

【0145】

CRRPのアフィニティー熟成は、異なる二つの方策：モジュール付加およびモジュール置換によって実行することが可能である。

【0146】

「モジュール付加法」は、比較的少数（例えば、1-3）の反復単位から始め、アフィニティー熟成の各工程でランダム化反復単位を加え、次いで、結合因子に関にして選択する。各進化サイクルにおいて、ランダム化モジュールを加え、次いでもっとも活性の高いクローンを選択する。この過程は、各サイクルにおいてタンパクのサイズを増し、一方、各延長ラウンド後、所望の結合活性について選択する。この方法は、ランダム化配列を、選択配列に変換する。

【0147】

「モジュール置換法」は、比較的多数（例えば、4-10；「最終数」）の反復列から始め、各ライブラリー生成ラウンド毎に、新しい反復列（通常1-3）群をランダム化し、次いで標的結合性について選択する。この方法では、タンパクのサイズは一定である。選択されない配列（通常、固定）は、徐々にランダム化配列に変換され、これが次に選択配列に変換される。

【0148】

両方法とも、1、2、3、4、5、6、またはそれ以上の標的に対する結合アフィニティーの向上について選ばれた、大きな単一結合部位、または、複数の別々の結合部位を持つ反復性タンパクを生成する。反復列の付加は、結合部位（単複）の延長を可能とし、これは、その標的に対し単一部位で結合するドメインと比べ、より高い結合アフィニティーをもたらす。反復性タンパクドメインは、反復列を含まない短いリンカー配列を介して、他の反復性タンパクドメインに連結させることが可能である。これは、天然の反復性タンパクに観察されるものと同様の反復性タンパク組織である。天然の反復性タンパクは、短いアミノ酸配列によって縦列に連結され、かつ、その間に非反復性タンパクが点在することがしばしば起こる（H. K. Binz ら、(2005) Nature Biotechnology）。

【0149】

しかしながら、反復性タンパクはまた、二つの結合部位の間に剛強な接続部を形成するのに用い、これらの部位が同時に標的に結合できるようにしてもよい。別々のドメインの間に通常存在する、屈曲性のペプチドリinkerと違って、反復性タンパクによる剛強なコネクターは、より高い結合アフィニティーを生成することが期待される。結合部位の間に剛強なコネクターを創製するためのもう一つの方法は、それ自身コイル形成するプロリン富裕配列、またはコラーゲン様配列を用いることである。

【0150】

アフィニティー熟成は、単一の連続配列、または複数の不連続配列を標的として、DNAレベルにおける（部分的）ランダム化によって実行される。DNAランダム化の順序工程も、DNAレベルにおいて不連続的か、連続的（すなわち、順序的）であってもよい。タンパクレベルでは、突然変異発生も、用途に応じて、不連続的、または連続的であってもよい。例えば、螺旋状反復タンパクでは、タンパクの同じ側に連続的結合表面を実現するために、DNAおよびタンパク鎖レベルにおいて不連続的熟成を使用するのが典型的と

10

20

30

40

50

考えられる。これは、タンパクの表面ではランダム化区域は連続的であるにも拘わらず、ランダム化アミノ酸が、アルファ鎖バックボーンおよびDNAレベルにおいて不連続的であるために不連続的と呼ばれる。他方、順序的熟成は、DNAレベルおよびタンパク・バックボーンレベルにおいて連続的な、一組のアミノ酸について、螺旋の全ての側面がランダム化され、標的に対する結合部位となり、この反復性タンパクと標的タンパクの間に、より複雑な三次元相互作用を可能とするように、ランダム化することを含む。不連続（DNAレベル）アフィニティー熟成の場合は、ランダム化配列の間に挟まる、一般的固定配列が、ライブラリー内において、または、複数のライブラリー間において、制限酵素または重複PCRによる組み換えを実行するのに利用され、これによって、結合アフィニティーの向上についてスクリーニングされるクローンの数を増す、さらに新たな工程が提供される。

10

【0151】

アフィニティー熟成にとって好ましい方法は、配列のランダム化であり、これは、スカフォールドタンパクの一区域を第1（部分的）ランダム化すること、最良クローンのプールを選択すること、この選択されたプールのクローンにおける第2区域をランダム化すること、最良クローンの（第2）プールを選択すること、この第2プールのクローンの第3区域をランダム化すること、および、向上クローンの（第3）プールを選択することを含む。これは、例えば、図136に示される。好ましい方法は、三つの突然変異発生区域（ n -項、中間、および c -項）を非重複性とすることである。任意の順序の突然変異発生を使用することが可能であるが、 n -項／中間／ c -項、および n -項／ c -項／中間が、好ましい選択である。突然変異発生区域の間の、15-20bpの、突然変異非発生スカフォールド配列はそのまま放置しておくのが有用である。その配列が、Kunkel型突然変異発生においてオリゴヌクレオチドのためのアニール区域として役立つからである。この方法は、以前に突然変異発生した配列の、合成的な、突然変異の再発生、これは、通常、クローンの配列決定、配列同士の整列、ファミリーモチーフの抽出、およびこれらのモチーフをコードするオリゴの再合成、および新規合成ライブラリーの創製を含む手間のかかるプロセスである、を回避する。好ましい方式は、ランダム化が、各位置において、多くは天然に見られるアミノ酸を生成するようにコドン選択を用いることである。

20

合成CRRP

合成CRRPは、モチーフ $C_a X_{0-n} C_b X_{0-n} C_c X_{0-n} C_d X_{0-n} C_e X_{0-n} C_f X_{0-n} C_g X_{0-n} C_i X_{0-n} C_j X_{0-j}$ から成る。式中、 C は、定められた位置におけるシステイン残基であり、 x は、各個別のシステイン間における0と12の間の、任意の数のアミノ酸である。これらの設計は、CRRPモチーフによって、例えば、個別のジスルフィド結合間のシステイン距離、ジスルフィド結合の第1システインから、次のジスルフィド結合の第1システインまでのシステイン距離によって定義される。ライブラリー設計には下記のモチーフが有用である。 $3/4/1$ 、 $C_a X_{0-n} C_b X_{0-n} C_c X_{0-n} C_d X_{0-n} C_e X_{0-n} C_f X_{0-n} C_g X_{0-n}$ で、 C_a が C_d とジスルフィド結合を形成する； $(3, 4) (1, 4) / 2$ 、 $C_a X_{0-n} C_b X_{0-n} C_c X_{0-n} C_d X_{0-n} C_e X_{0-n} C_f X_{0-n} C_g X_{0-n}$ で、 C_a が C_d とジスルフィド結合を形成し、 C_c が C_g とジスルフィド結合を形成する； $(4/2)$ 、 $(3/1)$ 、 $C_a X_{0-n} C_b X_{0-n} C_c X_{0-n} C_d X_{0-n} C_e X_{0-n} C_f X_{0-n} C_g X_{0-n}$ で、 C_a が C_e とジスルフィド結合を形成する； $(3, 5) / (1, 2) / 2$ 、 $C_a X_{0-n} C_b X_{0-n} C_c X_{0-n} C_d X_{0-n} C_e X_{0-n} C_f X_{0-n} C_g X_{0-n}$ で、 C_a が C_f とジスルフィド結合を形成し、 C_b が C_e とジスルフィド結合を形成し、 C_d が C_i とジスルフィド結合を形成する； $(3, 5, 7) / (1, 2, 3) / 3$ で、 C_a が C_f とジスルフィド結合を形成し、 C_b が C_e とジスルフィド結合を形成し、 C_c が C_j とジスルフィド結合を形成する； $(4, 5) / (1, 4) / 2$ で、 C_d が C_i とジスルフィド結合を形成し、 C_f が C_j とジスルフィド結合を形成する（図125-133を参照）。

30

40

【0152】

50

10

20

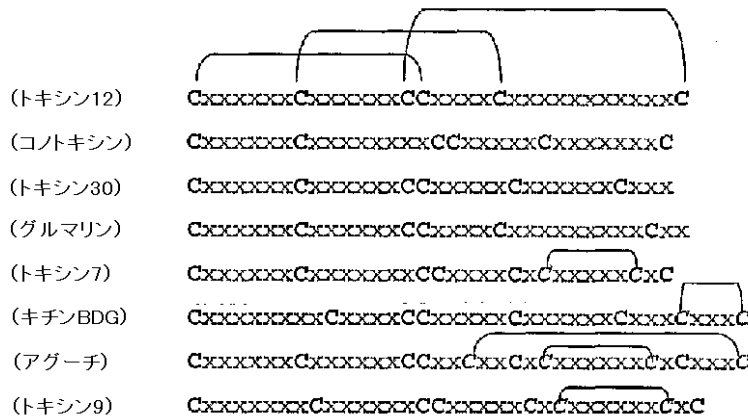
30

40

C X C X C C X C X Cモチーフによる合成スカフォールドの設計：多くのマイクロタンパク・ファミリーが、ジスルフィド空間配置1-4 2-5 3-6を持つ、下式：
C x x x x x x (x x x x x x x) C x x x x x x (x x x x x x x) C C x x x x x x (x x x x x x x) C x x x x x x (x x x x x x x) Cから成るモチーフを含む。この一般的共通配列がライブラリー設計に使用される。間隔は、さらに別のシステインおよびジスルフィド結合を含んでもよい。各ジスルフィド結合間の間隔は、平均13-15である。基本的モチーフの外に加わる余分なシステイン・ペアは、青または緑のイタリック体で示され、連結されるシステインは同じ色を共有する。

【 0 1 5 6 】

【化 1 2】



10

20

30

40

50

	1-4	2-5	3-6	追加SS
トキシシン12	13	12	17	
コノトキシシン	15	15	14	
トキシシン30	14	13	13	
グルマリリン	14	12	15	
トキシシン7	15	13	15	6-7
キチンBDG	14	11	13	7-8
アグーチ	14	13	16	5-10, 7-8
トキシシン9	15	15	15	
平均	14	13	15	

Swissprotは、間隔6、5、0、3を持つ44メンバー、および間隔6、5、0、4を持つ57メンバー、および間隔6、6、03を持つ34メンバー、および間隔6、6、0、4を持つ27メンバーを含む。最後の間隔(Cys5とCys6の間の)は、4から6アミノ酸まで変動してもよい。

【0157】

システイン距離パターン(CDP)：天然のタンパクをファミリーに分類するためにもっとも一般的に使用される方法は、タンパク配列の相同性に基づく。これらのアルゴリズムの目的は、多くの場合、進化距離を反映する近縁性に基づいてタンパク配列をグループ分けすることである。これらのアルゴリズムは、各位置について、同一マッチするか、または化学的に近縁なアミノ酸の数を最大とするように配列を整列させる。しばしば、整列を向上させるためにギャップが導入される。有意な配列変動を可能とし、したがって、新規結合タンパクのための基礎となるタンパクスカフォールドを特定するために、一般に、このような相同性に基づく配列ファミリーが用いられる。しかしながら、相同性に基づくファミリーは、近縁マイクロタンパク間には配列保存性の程度が低いために、マイクロタンパク系ライブラリーの設計には限られた有用性しか持たない。ごく近縁なマイクロタンパクの配列は、そのシステイン残基の保存を除いては、ほとんど配列相同性を持たない場合が多い。相同性に基づく探索アルゴリズムによるギャップの導入は、突然変異残基、および、タンパク構造および／または安定性に重要な残基の特定において決定的な、マイクロタンパク配列の整列を複雑にする。マイクロタンパクは、極端に高密度のシステイン残基を有することで、他の多くのタンパクとは異なるので、この群は、システイン間隔をキー・パラメータと位置づけ、同じシステイン距離パターン(CDP)を共有するマイクロタンパクをまとめてクラスターとすることを可能とする整列法を必要とする。したがって、システイン距離クラスターとは、同じ数のアミノ酸によって隔てられるいくつかのシステイン残基を有するタンパク配列群である。あるシステイン距離クラスターの全メンバーの配列は、全

てのクラスターメンバーは同じ合計長を持つので、整列される。さらに、配列内の各位置の平均アミノ酸組成を簡単に計算することができる。これは、マイクロタンパク・ライブラリーを構築する際、変動する可能性のある残基の特定ばかりでなく、変動の程度の特定制までごく簡単なものとする。同じCDPを持つマイクロタンパクの大型クラスターは、マイクロタンパク・ライブラリーを設計するために特に有用である。なぜなら、該クラスターは、各位置における天然の変動について詳細な情報を提供するからである。

【0158】

CDPクラスターは、通常、近縁マイクロタンパク配列のサブセットである。多くの場合、CDPクラスターのメンバーは全て、相同タンパクから成る同じファミリーの出自である。しかしながら、複数のタンパクファミリー由来のメンバーを含むCDPクラスターもある。一例は、51メンバーを含む、CDPクラスター3__5__4__1__8（場合によってC3C5C4C1C8、またはCxxx Cxxxx Cxxxx Cxx Cxxxx Cxxxx Cとも表示される）である。このクラスターでは、あるものはファミリーPF00008から、別のあるものはPF07974から得られる。このCDPを持つ配列は、（原理的には）、両方の構造を取ることが可能である。これらの構造的に異なるCDPは、構造進化を実現するために好ましい。

【0159】

DBPは直接調節することが難しいが、CDPは遺伝子合成によって簡単に調節されるので、CDPは、DBPおよび構造を調節するのにもっとも好ましい方法となっている。

【0160】

有用CDPの特定：有用CDPは、Swiss-Protまたは翻訳EMBL（TrEMBL）などのタンパク配列データベースを分析することによって見つけ出すことが可能である。Swiss-ProtおよびPfamからの情報を組み合わせ、システイン結合パターンを注記するデータベースが、Guptaによって記載されている（Gupta, A., ら、(2004) Protein Sci, 13: 2045-58）。このデータベースは、マイクロタンパクに典型的な、高いパーセンテージのシステイン残基を含むタンパク配列を求めて探索することが可能である。連続または隣接システイン残基間の距離を計算し、CDPを得、次いで多数回出現するCDPを探索することが可能である。多くの天然配列が同じCDPを共有する場合、CDPは特に興味深い。なぜなら、これは、CDPが多様な配列を許容することを示すからである。有用なCDPは、隣接システイン残基間の距離が長い（「長いループ」）のを避ける。なぜなら、長距離ループは、プロテアーゼによる攻撃を受け易く、かつ、MHC分子の裂け目に結合するほど長いペプチドを生成する可能性がより高くなるからである。特に興味深いのは、距離がどれも15、14、13、12、または11アミノ酸を超えないCDPである。より好ましいのは、隣接システイン間の距離がどれも10、9、または8残基を超えないCDPである。特に興味深いのは、トリプトファン、フェニルアラニン、チロシン、ロイシン、バリン、メチオニン、イソロイシンなどの疎水性アミノ酸を少量しか含まないファミリー由来のCDPである。これらの疎水性残基は、典型的タンパクでは約34%の頻度で起こり、非特異的疎水性結合と関連する。特に興味深いCDPは、疎水性残基を30、28、26、24、または22%未満しか持たないメンバーをたくさん含む。好ましいCDPおよび個別メンバーは、疎水性残基を、20、18、16、14、12、10%未満、場合によっては8または6%しか含まない。特に興味深いのは、個別メンバーが大きな配列多様性を示すCDPである。表2は、マイクロタンパク・ライブラリーのきわめて有用なスカフォールドとして役立つことが可能なCDPの例である。[表3]は、もっとも好ましいCDPを示す。

【0161】

【化 1 3】

表2. 例示のCDPのリスト

メンバー番号	ジスルフィド数	ドメイン 長さ	C-C, AA表示	ループ長n1	n2	n3	n4	n5	n6	n7
124	3		37	6	4	8	1	12		
107	3		43	3	10	11	9	4		
103	3		51	8	15	7	12	3		
93	3		58	12	12	3	13	12		
92	3		49	7	7	10	2	17		
90	3		36	6	3	8	1	12		
77	4		46	1	9	6	1	8	2	11
74	4		37	8	4	0	5	6	3	3
70	7		65	1	5	3	0	4	7	4
69	4		57	10	6	16	3	10	0	4
65	3		46	15	2	12	3	8		
60	2		22	4	13	1				
59	2		40	3	29	4				
54	3		38	6	5	6	5	10		
54	6		61	1	6	0	4	7	4	0
49	3		31	6	4	9	6	0		
49	4		61	1	6	17	2	8	2	17
47	3		56	11	28	0	3	8		
45	2		21	4	12	1				
45	4		38	8	4	0	5	6	4	3
44	3		45	3	7	10	6	13		
44	4		48	1	6	6	2	8	2	15
42	4		58	13	6	16	1	10	0	4
41	4		47	3	8	11	2	0	5	10
40	4		52	3	5	3	9	9	1	14
40	5		59	8	3	3	6	10	3	1
39	2		15	1	7	3				
39	3		35	5	3	8	1	12		
38	4		31	1	4	0	5	6	3	4
37	3		30	12	0	0	10	2		
36	4		38	8	4	0	5	6	3	4

10

20

30

40

【0 1 6 2】

【化 1 4】

36	7	65	1	5	4	0	3	7	4
35	3	36	0	12	12	6	0		
34	3	38	9	9	4	0	10		
33	3	29	12	0	0	9	2		
31	3	45	2	5	16	2	14		
31	7	76	2	7	5	0	5	9	9
30	3	36	7	4	10	1	8		
29	3	34	6	5	8	1	8		
29	2	40	13	9	14				
29	3	47	16	2	12	3	8		
28	2	9	0	3	2				
28	3	26	6	5	3	1	5		
28	3	46	3	10	12	11	4		
27	3	39	9	7	12	3	2		
26	2	23	5	11	3				
26	4	48	1	9	6	1	8	2	13
25	3	26	8	2	1	8	1		
25	3	36	6	5	8	1	10		
24	2	25	3	7	11				
24	3	47	3	9	10	6	13		
23	3	41	12	6	12	3	2		
23	3	42	10	8	13	3	2		
23	4	45	1	9	5	1	8	2	11
23	3	46	3	8	10	6	13		
23	5	61	2	4	5	6	17	3	10
22	2	14	3	1	6				
22	3	24	0	4	7	1	6		
22	3	29	4	5	5	1	8		
22	3	29	5	3	10	4	1		
22	3	31	12	0	0	9	4		
22	3	38	0	11	9	5	7		
22	4	51	1	11	6	1	8	2	14
22	7	77	2	7	5	0	5	9	9
21	3	37	7	5	6	5	8		
21	3	48	6	7	10	2	17		
20	3	30	13	0	0	9	2		
20	2	33	9	10	10				
20	4	50	1	11	6	2	8	2	12

「メンバー」と表示されるコラムは、Gupta (Gupta, A., ら、(2004) Protein Sci, 13:2045-58) によって記載されたデータベースにおいて特定された、特定CDPを持つ天然配列の番号を示す。'n'は、そのクラスターにおけるジスルフィド数である。「ドメイン長」は、CDPのアミノ酸残基数(第1cysから最終cysまで)である。コラムn1からn7は、クラスターのシステイン残基を隔てる、非システイン残基の数を掲げる。n2=6は、C2およびC3の間のループが、

10

20

30

40

50

両システムを除いて、6 A A長であることを意味する。

【0 1 6 3】

【化 1 5】

表3. 例示のCDPのリスト

メンバー番号	ジスルフィド数	ドメイン 長さ	AA	ループ長								
					n1	n2	n3	n4	n5	n6	n7	
575	3		35		6	4	6	5	8			10
518	3		32		4	5	8	1	8			
190	3		37		6	4	6	5	10			
155	3		36		6	5	6	5	8			
93	3		36		6	4	6	5	9			
72	3		38		7	4	6	5	10			
71	3		23		2	1	7	1	6			
67	3		37		6	6	6	5	8			
64	3		36		5	4	8	1	12			20
62	3		36		7	4	6	5	8			
59	3		34		4	5	10	1	8			
57	3		28		3	5	5	1	8			
57	3		33		4	5	9	1	8			
56	3		35		6	6	12	3	2			
54	4		44		1	9	6	1	8	2	9	
51	3		27		3	5	4	1	8			
49	3		29		1	4	9	9	0			
45	3		37		6	5	6	5	9			30
43	3		31		4	4	8	1	8			
43	4		45		10	5	3	9	6	1	3	
38	4		45		1	9	6	1	8	2	10	
34	5		54		8	3	3	8	3	3	1	
33	3		41		3	10	9	9	4			
29	2		23		6	5	8					
27	3		37		6	3	9	1	12			
26	4		35		3	9	1	3	5	0	6	
25	3		26		4	3	10	2	1			40
25	3		35		4	5	11	1	8			
24	3		34		5	4	6	5	8			
24	3		37		7	3	8	1	12			
24	3		44		3	10	10	11	4			
23	3		35		6	8	10	3	2			
22	3		33		5	5	8	1	8			
22	3		37		3	10	5	9	4			

【0 1 6 4】

【化 16】

21	3	33	9	9	4	0	5	
21	3	36	3	10	4	9	4	
20	2	18	9	0	5			
20	3	34	5	5	9	1	8	
20	3	42	3	10	10	9	4	
20	4	43	1	9	5	1	8	2 9

「メンバー」は、Gupta (Gupta, A., ら、(2004) Protein Sci, 13: 2045-58) によって記載されたデータベースにおいて特定された、特定CDPを持つ天然配列の番号を示す。‘n’は、そのクラスターにおけるジスルフィド数を示す。「ドメイン長」は、CDPのアミノ酸残基数（第1cysから最終cysまで）を示す。コラムn1からn7は、クラスターのシステイン残基を隔てる、非システイン残基の数（「ループ長」）を掲げる。

【0165】

システイン間ループの内のあるものは、サイズを固定させる必要があるが、一方、別のループは、ある程度の長さの変動性を受容することが可能である。天然配列のファミリーに見られる長さの多様性は、特定のループにおいてどの程度の長さ変動が受容可能であるのかを評価する一つの目安となる。このように許容される長さ変動は、マイナス10、9、8、7、6、5、4、3、2、1アミノ酸からプラス1、2、3、4、5、6、7、8、9、または10アミノ酸の範囲に亘る。

【0166】

DBPの指向性進化、およびクローンプールのタンパクフォールド：多数のジスルフィド結合パターン（DBP）は、非HDDタンパクには、たとえそれが多くのジスルフィドを持つものであっても利用不可能であるが、HDD（「高ジスルフィド密度」）タンパクの最適化のためには使用することが可能な、さらにもう一つの自由度である。一つの要因は、大型のタンパクでは、ジスルフィドは互いに遠くに隔てられるので、他の固定配列がタンパクを折りたたみ、それによって、システイン同士が互いに局所的に高濃度となるように、かつ適切な方向性をもって近づけられる場合を除いては、反応する可能性がまず無いということである。したがって、システインは、大型のタンパクのフォールド形成では比較的小さな役割しか果たしていない。疎水性コアを持つ大型タンパクは、3D構造の創出に与る、たくさんの側鎖接触点を持つ傾向がある。この、Hubert Yockey (1974) によって定義される、いわゆる高情報含量解決策では、DBPは、統計学的に場所的に固定され、DBPの進化的変化はほとんど起こりえない。構造的進化は、低情報含量のタンパク、すなわち、構造および機能のために必要とされる残基をほとんど持たないタンパクにのみ利用可能であると考えられる。ランダム突然変異発生に対する感受性と定義される、タンパクの情報含量は、該タンパクの進化年齢の関数として単純に時間と共に増加するものではない。例えば、遺伝子が二重になると、二つのコピーの内の一方は、もしもその遺伝子について一コピーしかなかったならその情報含量は高くなると予想されるのに、自由に進化し、その情報含量は効果的にきわめて低い。情報含量が低い状況では、多数のアミノ酸の突然変異、および大きな構造変化が起こり得、それらの変化が単一コピー遺伝子に起こったならば、致命的と考えられる。あるタンパクの情報含量はまた、考慮の対象とされる特異的機能的側面に依存する。ある機能（すなわち、触媒作用）は、別の機能（すなわち、9AAT細胞エピトープに基づくワクチン）よりもはるかに高い情報含量を持つ。冗長性は、有毒動物では一般的であり、通常、有毒動物はそれぞれ、その毒の中に、同じ、または異なる遺伝子から得られた100種を超えるトキシンを持つ。冗長性は、同じ遺伝子の複数コピーとして、および／または、多様なトキシンをコードする種々の遺伝子の単一コピーとして、HDDタンパクの急速な進化を助けるようである。

【0167】

10

20

30

40

50

標的に対する結合に関して選択されたクロンのプールは、その結合機能を与えるドメインのほんの一部（サブ、またはマイクロ・ドメイン、または1本以上のループ）を持つだけでよい。典型的な10e10ライブラリーの最良クロンでも、平均して、完全に最適化されたアミノ酸はほんの約7個にすぎない。これは、1サイクルのパンニングで付加される最大（平均）情報含量は、ライブラリーのサイズ（すなわち、10e10）だからである。それ以上に情報含量を集積するためには、一般に、複数サイクルのライブラリー生成およびスクリーニングが必要とされる。10e10の3サイクルは、理論上、最大10e30の情報含量をもたらすことになるが、通常、この数字は、加算性に対する実地の限界によってそれよりもはるかに少ないと予想される。通常、ドメインのアミノ酸の多くは、標的に直接接触することはせず、全部ではないにしろ、種々のアミノ酸によって置換されることが考えられる。構造的進化の一つの目的は、非結合部のDBPを進化させて、標的に結合する部分のアミノ酸配列をまったく変えることなく、より高いアフィニティーの標的結合を実現する修飾構造をもたらすことである。

10

【0168】

好ましい方法は、各単一配列から、複数の構造の形成を奨励することである。これを、第1サイクルで行うか、または、1サイクル以上のパンニングで多様性が低下した後に行う。それによって、それぞれが異なるDBPおよび構造を取ることが可能な、各ファージクロンの多数コピー（ $> 10e4$ ）が得られる。パンニング前にライブラリーの構造の多様性を増すための一つの方法は、ライブラリーを10-30秒加熱した後、部分的にフォールド形成した構造の全てを除去するために、ライブラリーに高濃度の酸化剤を急激に添加することである。タンパクがアニールし、そのフォールド形成経路を探索する機会を持つ前に、ジスルフィドを急激に形成させることによって多様性の増加が得られるはずである。もっとも得られたフォールドの平均的品質は、この方法では低下するおそれがあるが。均一なフォールド形成を実現するためには、反対の方法が使用され、これは、通常、透析によって還元剤をゆっくりと除去することを含む。これによって、ゆっくりしたフォールド形成およびゆっくりしたスルフヒドリル酸化が得られる。この方法はまた、オリゴヌクレオチドのアニリングと同様の、温度のゆっくりした低下を含む。DBP-多様性を、第1ラウンドのパンニングでライブラリーに適用する場合、大きなライブラリーの余剰分、例えば、異なるクロン数（通常、10e9-10e10）よりも10e5倍たくさんの粒子を創製することが重要である。これは、各配列から創製される可能性のある、大きな数の、異なる構造をカバーするためである。

20

30

【0169】

DBPの多様化：DBPのスペクトラムおよび分布は、同じライブラリーの分液を、多様な、種々の条件に暴露することによって多様化することが可能である。これらの条件として、ある範囲のpH、温度、酸化剤、還元剤、例えば、DTT（ジチオスレイトール）、BME（ベータメルカプトエタノール）、グルタチオン、ポリエチレングリコール（分子密集化し、散発的DBPをより高頻度にする）などが挙げられる。

【0170】

複数スカフォールドライブラリー：高いアフィニティーで標的に結合するマイクロタンパク・ドメインを特定するために、下記の3工程プロセスに従って、複数スカフォールドライブラリーが用いられる。

40

【0171】

1. 複数スカフォールド、またはシステイン距離パターン（CDP）、および各種ランダム化スキームに基づいてサブライブラリーを構築する。

【0172】

2. 対象標的に対しいくつかのサブライブラリーをパンニングすることによって初回ヒットを特定する。これは、各ライブラリーを別々にパンニングすることによって行ってもよいし、あるいは、サブライブラリーの混合物をパンニングして行ってもよい。

【0173】

3. 突然変異発生および選択またはスクリーニングを含む反復過程であるアフィニテ

50

ィー熟成によって初回ヒットを最適化する。

【0174】

複数スカフォールドライブラリーの使用は、個々のスカフォールドに注目する従来の方法とは大きく異なる。単一スカフォールドライブラリーでは、大抵のライブラリーメンバーは、類似の全体構造、すなわちフォールドを共有するが、主に、そのアミノ酸側枝において異なる。単一スカフォールドライブラリーの例は、フィブロネクチンによるもの（Koide, A., ら、(1998) J Mol Biol, 284:1141-51）、リポカリン（Beste, G., ら、(1992) Proc Natl Acad Sci USA 96:1898-903）、またはタンパクAードメイン（Nord, K., ら、(1997) Nat Biotechnol, 15:772-）によるものであった。さらに別の、多くのスカフォールドが Binz, H. K., ら、(2005) Nat Biotechnol, 23:1257-68 に記載されている。ある場合では、単一スカフォールドライブラリーは、個別のループ長において僅かな差を示すメンバー、例えば、抗体ライブラリーに CDR を含んでいた。単一スカフォールドライブラリーは、限られた量の形状空間しかカバーしない傾向がある。そのため、低アフィニティー結合子しか得られないことがよくある。これらの分子は、その標的の形状にぴったりと合致しない。しかしながら、接触区域を形成するアミノ酸は、形状の相補性の不足を部分的に補うように最適化されてきている。スカフォールドと標的の間の接触区域におけるアミノ酸の多様性を改善するため、多くの公刊物が、ライブラリーサイズを増大させる試み（すなわち、リボソームディスプレイ、組み換えファージディスプレイ）を記載する。しかしながら、このプロセスは、通常、結合タンパクにおける外部の、CDR 様ループにおける小さな変化を強調するが、ドメインの全体構造には影響を及ぼさない。固定スカフォールドのアフィニティー熟成が、全体フォールドおよび結合タンパクの構造に大きな変化をもたらした例はない。大きな変化が起こった稀な場合では、そのようなクローンは一般には除去される。それらの免疫原性および製造性質が予測不能と考えられるからである。

【0175】

複数スカフォールドライブラリーは、全体構造が大きく異なる、多様なスカフォールド（多くの場合、無関係の）を有するクローンを含む。一般に、各 CDP は、異なる形状を表し、各サブライブラリーは、ある特定の CDP 周囲の配列空間をほとんど抽出しない突然変異集団を含む。（それぞれが異なる CDP を持つたくさんのサブライブラリーから得られる）多くの、異なる形状を有する分子を調べることによって、その構造が、標的表面に緊密に当てはまる結合タンパクを特定する機会が増す。各サブライブラリーは、CDP の周辺の配列空間に関する比較的小さいサンプルを表すにすぎないので、この過程から最適結合配列を獲得する可能性は低い。複数スカフォールドライブラリーの初回ヒットは、その標的の形状を模倣するが、ヒットと標的の間の接触面の構造は、最適以下の可能性がある。そのため、構造を大きく変えることなく特定のタンパク配列の最適化に重点を注ぐ、その後のアフィニティー熟成において、結合アフィニティーの改善が達成される可能性が高い。単純化して言うと、目的は、標的に適合する最良構造を見出し、次いで、この構造に適合する最良配列を見出し、標的に対する最適相補性を実現することである。

【0176】

新規スカフォールドを見つけ出すための実験法：ライブラリー設計改良のためのもう一つの方法は、多様な設計を競合させることによって、タンパク自身に、その最良解決案を計算させることである。完全にフォールド形成し、十分に発現されたタンパクを選択し、配列決定する。フォールド形成タンパクをもっとも高い割合（入力メンバーについて補正した）で持つ設計が好ましい。好ましい CDP および配列モチーフを見出すには、いくつかの異なる方法がある。

【0177】

方法 1：ランダム CDP、ランダム配列

ランダム空間および配列法は、天然の多様性に存在する空間または配列にはよらない、したがって、ランダム配列を受け容れる、その方法の能力に比例して、新規および既存の

c y s -空間パターンを見出すことが可能である。

【0178】

この方法は、設計 C X (0 - 8) C X (0 - 8) C X 0 - 8) C X (0 - 8) C X (0 - 8) C によって、例えば、1 0 e 1 0 ディスプレイライブラリーなどの広範な、開放ライブラリー作製し、次いで、アガロースゲル、大腸菌における発現を用いて、2 5 - 3 5 A A 全長に関して選択し、このディスプレイライブラリーから、遊離チオールカラムを用いてフォールド未形成タンパクを全て除去し（または、発現レベルについて個々のクローンをスクリーニングし）、かつ、十分に発現され、完全にフォールド形成されたタンパクをコードする 2 0 0 - 1 0 0 0 クローンについて配列決定を行うことを含む。

【0179】

このライブラリーでは、距離パターンは全て近似頻度で起こる。天然タンパクに見られる間隔／距離パターンにおいて強いバイアスを見出すこと、多くの間隔パターンが新規であることが期待される。例えば、距離パターン A が、0 . 0 1 % しかフォールド形成タンパクを許容せず、パターン B が、1 0 % のフォールド形成タンパクを生成するとしたならば、パターン B を有するクローンは、パターン A を有するクローンよりも 1 0 0 0 倍より高頻度に出現するはずである。1 0 0 0 クローンの配列決定は、ループ配列とは無関係に、フォールド形成能力のもっとも高い 1 0 - 3 0 間隔を特定するのに十分であるはずである。この方法によって見出される多くの間隔パターンは、新規である可能性が高く、これらの間隔に基づいて別々のライブラリーを作製するために使用することが考えられる。

【0180】

この方法によって見出される新規間隔は、通常、次の方法において、天然のファミリーにおける間隔と組み合わせられる。

【0181】

方法 2：天然 C D P、ランダム配列

1 0 - 1 0 0 の特異的天然ファミリーの C D P が、ランダム A A 組成物（すなわち、N N N、N N K、N N S、または類似のコドン）を用いて合成され、次にライブラリーに変換され単一プールとされ、前述のようにフォールド形成および発現について選択またはスクリーニングされ、次いで、最良フォールド形成および発現クローンの配列が決定される。この方法は、ランダム配列を受容する能力について、天然ファミリーのスカフォールドのランク付けをもたらす。この方法は、より高い平均品質レベルを生成する傾向がある。これは、フォールド形成クローンの割合が、ランダム C D P 法よりもはるかに高いからであるが、同じぐらいたくさんのスカフォールドという点では評価されない。

【0182】

好ましい間隔パターンを選択した後、フォールド形成を改善するためには、ある特異的間隔パターンにおいてどの非 - c y s 残基が必要なのかを決定する。

【0183】

方法 3：天然 C D P、天然 A A 配列混合物

1 0 - 1 0 0 特異的天然ファミリーの間隔パターンを、各位置に現れる（整列で決定した場合に）A A 組成物の天然の混合物を用いて合成し、次にライブラリーに変換、単一プールとし、前述のようにフォールド形成および発現について選択またはスクリーニングし、次いで、最良フォールド形成および発現クローンの配列を決定する。この方法は、最高の平均品質レベルを生成する傾向があり、フォールド形成クローンの割合が、従来法よりもはるかに高い。しかし、この方法は、自然がもう既に探索した配列空間に関する高密度探索に、多少限定される。

【0184】

最高品質ライブラリー（すなわち、商品標的に対しすぐに有用な）は、固定された非 - c y s 残基の全てを有するが、各位置にある程度の変動を持つ天然のファミリー（天然 C D P）を合成することによって得られると考えられる。次に、十分にフォールド形成したクローンの配列を分析することによって、固定残基の内のどれが真に必要なのか、どの残基において変動が許容されるのかが分かると考えられる。

【0185】

構造進化：ジスルフィド含有タンパクがフォールド形成して、はっきりと定められた 3-D 構造へ変換されるのは、主に、インビボでもインビトロでも、現在の還元性環境の性質に依存する。例えば、ジスルフィド結合の還元は、タンパク構造の完全な喪失をもたらす、構造の維持におけるジスルフィド結合の重要性を強調する。その反対極では、完全に還元され、フォールドの解けたタンパクのフォールド形成時には、フォールド形成時近接するジスルフィド同士の酸化によって理論的に多数のジスルフィド異性形が可能である。4 個のシステインを含むタンパクでは、理論的に 3 通りのジスルフィド異性形、6 個のシステインでは 15 種の異性形、8 個のシステインでは 105 種の異性形が存在する。このような多様ではあるが、多くの場合非生産的異性形も、タンパクのフォールド形成過程で観察されるが、天然の立体配置では、通常、システイン対合の内ただ一つだけが代表とされる。これが、インビトロのフォールド再形成実験において、ジスルフィド異性体形成が、多くの研究者によって重大問題と見なされる理由である。一方、ジスルフィド富裕マイクロタンパクの構造多様性の進化のために、ジスルフィド異性体形成を利用することが可能である。小型であり、かつ、高いジスルフィド含量を持つために、これらのタンパクは、折りたたまれた立体配置を維持するために、システインの共有結合のみに依存することが多い。多くのマイクロタンパクは、大型タンパクのフォールド形成において一般的基礎力と見なされる疎水性コアを完全に欠如する。それぞれ異なるジスルフィド異性形が、マイクロタンパクファミリー、例えば、ソマトメジン B および蛇のコノトキシシン (Y. Kamikubo, ら、(2004) *Biochemistry*, 43: 6519-34; J. L. Dutton, ら、(2002) *J Biol Chem*, 277: 48849-57) の単一メンバーにおいて実験的に観察されている。しかしながら、これらの公刊物は、複数の異性形の存在を、克服すべき問題点として記載しており、タンパク設計に利用すべき機会としては見ていない。したがって、マイクロタンパクの構造進化の駆動力としてジスルフィド異性形形成を使用するための、一般的に適用が可能な概念および実験手順の開発が可能とされる。

10

20

【0186】

ジスルフィドシャッフリングによる構造進化：下記の文節は、構造進化のためにジスルフィド異性形を利用する、特定の実験法を示す。特定のマイクロタンパクに融合されたファージ粒子の分泌後、混合物を、ミリモル濃度の還元グルタチオンとインキュベートすることによって、これらの粒子を高度の還元条件に暴露して、レドックス活性、ジスルフィド含有トリペプチドを生成する。次に、遊離チオールを空気酸化を防ぐために、ミリモル濃度の EDTA を含むバッファーにおいて、ファージ粒子を還元剤から精製する。このライブラリーは、多数の、還元され、構造的に多様なポリペプチド鎖を含む。これらの、異性形の還元された混合物を接触させた後、ライブラリーを、標的結合の際、酸化条件、例えば、ミリモル濃度の酸化されたグルタチオンに暴露し、そのチオールの酸化によって好適なマイクロタンパク立体配置に固定する。この方法は、最初に、その還元状態において標的と相互作用を持ち、次いで、急速な酸化によって結合立体配置に固定されたマイクロタンパク結合因子を選択的に選び出す。選ばれたマイクロタンパクプールは、標的タンパクに対し形状として相補的であり、このプロセスは、ジスルフィド依存性標的誘発フォールド形成と呼ばれる。最良結合因子が選択され、指向性進化（突然変異発生およびパンニング）の、さらに追加のサイクルに暴露され、最終的に、標的独立的なやり方で、すなわち、所望の立体配置を誘発するのに標的はもはや必要でなくなるので、タンパクの製造がよりやり易くなるやり方で、活性の高い、十分酸化された立体配置が実現される。

30

40

【0187】

別法として、ファージライブラリーは、ジスルフィドシャッフリングを可能とするために、中間レドックス電位を持つバッファーに暴露される。これは、様々な比率の酸化および還元グルタチオンを有するバッファー組成物を選択することによって簡単に実行される。これによって、システイン残基のサブセットの部分的酸化、およびそれに続くジスルフィドシャッフリング、例えば、既存のボンドの破壊と、多くのジスルフィド結合の集積に

50

好都合となる再形成が可能となる。したがって、この条件下では、多くの異なる構造的組み合わせ（ある任意のミクロタンパクのシステイン残基数に依存する）のプールが存在する。次に、もっとも強力なクローンが選択され、さらにもう1ラウンドのジスルフィドシャッフリングに暴露される（アミノ酸配列の最適化付きで、または無しで）。

【0188】

ジスルフィド結合を介する標的との共有結合：長く保持されてきた考えとは違って、最近の研究では、ジスルフィド結合の特異的還元は、細胞外環境で起こることが示されている（P. J. Hogg (2003) Trends Biochem Sci, 28:210-4）。内皮細胞は、トロンボスポンジン1と特定することが可能な還元性活性物を上清に分泌することが示された。このトロンボスポンジン1は、カルシウム結合ドメインにおいてレドックス活性を持つチオールを有する糖タンパクである（J. E. Pimanda, ら、(2002) Blood, 100:2823-8）。注目すべきことに、トロンボスポンジン1の遊離チオールは、分子間ジスルフィド結合を還元することによって、von Willebrand 因子の接着タンパクの長さを調節する。この所見は、ジスルフィド含有標的タンパクに対する、新規ミクロタンパクを共有的に連結するの利用することが可能である。方法は、標的タンパクのジスルフィド結合の近傍に結合する、部分的に還元された、レドックス活性を持つミクロタンパクを選択することになると考えられる。例えば、標的タンパクに対する結合後、ミクロタンパク変種のファージディスプレイ・ライブラリーについて、酸化条件下での洗浄には耐性を示すが、還元条件下での洗浄には特異的に溶出されるように選択が行われる。したがって、タンパク進化の際、ミクロタンパク構造を安定化する、あるジスルフィド結合は形成されるが、一方、別のあるものは、レドックス活性遊離チオールの選択のために、捨てられる。

【0189】

構造多様性の進化とは、ある特異的クローンによって経験される構造の変化を指す。構造変化は、通常、配列変化に依存するが、二つの同一配列であっても、異なる構造を取ることが可能である。構造の差は、ジスルフィド結合パターン、またはフォールドレベルでも起こり得るが、これは、一般に、ループ長における構造的に著明な変化による。構造進化は、構造の多様性（例えば、多くの複数スカフォールドライブラリーによって使用される）とは異なる。後者の場合、複数のスカフォールド構造が使用されるが、各クローンは常に、その親配列の構造を取る。構造進化では、各クローンは、親配列とは異なる構造を有することが可能である。

【0190】

図155は、優勢3SS結合パターン（18の異なる天然ファミリー）、および、一工程でそれから創製されるジスルフィド変種を示す。天然ファミリーの多くは、優勢3SSパターン（14 25 36）の1工程以内にある。図155はさらに、既存のジスルフィドのいずれも変えることなく、優勢3SSパターン（14 25 36）に1ジスルフィドを加えることによって創製される4SS変種を示す。天然の4SS結合パターンの内の11/15は、3SSジスルフィド結合のいずれも破壊することなく、優勢3SSパターンに1ジスルフィドを加えることによって獲得することが可能である。合計が105通りであるので、データは、既存の3SSタンパクに対するジスルフィドの添加に向けて強力な優先的指向性のあることを示唆する。この分析によって、この逆、すなわち、4SSタンパクからジスルフィドを取り去って3SSタンパクを創製することが好ましい経路であるかどうかについても解答が得られると思われる。データベースの不完全性がこの結果に影響を及ぼさない限り（可能である）、14 25 36、および、1ジスルフィドの付加によって得られる、その4SS誘導体は好ましいスタート点となるようである。

【0191】

ミクロタンパク構築法：この構築法の目的は、標的に対する結合タンパクの段階的アフィニティー熟成を実現することである。各サイクルにおいて、先行選択サイクルで得られた産物に、一対のシステイン、プラス、ランダム配列（通常、新規ループ）を加えてライブラリーが創製され、次いで、ライブラリーパンニングを行って、標的に対し最高のアフ

ィニティー、または活性を持つクロンを選抜する。スタート点は、単一配列であってもよいし、あるいは、配列プールであってもよく、開始点のランダム区域の配列は、既知、または未知でもよい。

【0192】

スタート点としての1-ジスルフィド（'1SS'）の創製：二つ以上のジスルフィドを有する新規マイクロタンパクを、構築法を用いて、単一ジスルフィド含有タンパクから創製することが可能である。一つの構築法は、二つの固定システイン残基（1-ジスルフィド、すなわち'1SS'タンパクのための）を含むタンパクから始める。要すれば任意に、このタンパクは、好ましい（通常、天然の）ジスルフィド結合パターンの一ループの中に認められるものと同じシステイン間隔または長さ（「スパン」と呼ばれる、システインは除く）を持ってもよい。このような相似性は、既存の2SS、3SS、4SS、またはさらに高次のスカフォールドへの、1SSペプチドの移植を容易にする。1SSライブラリーのスパンは、通常、0から20アミノ酸長であるが、好ましくは、5、6、7、8、9、10、11、12、13、14、15であり、さらに好ましくは、7、8、9、10、11、12であり、理想的には9、10、11アミノ酸長である。前記システインペアの外側（すなわち、ループまたは「スパン」の外側）の残基に対しさらにランダム化が行われてもよい。初期の1SSタンパクは、通常、システイン間において完全に、または部分的にランダム化されるが、時に、フォールド形成を行う、または標的分子（単複）に対しアフィニティーを持つ固定アミノ酸（システイン以外の）を含むことがある。

【0193】

1SSから2SSまたはさらに高次のスカフォールドの構築：以前に選択された1SSタンパクを熟成させる一つの方法は、固定タンパクに、あるいは、ライブラリーとして種々の好ましい位置に、二つの新規cys残基を設けることである。通常、この二つの新規システイン、および新規ループに側接する残基はランダム化される。

【0194】

偶数のシステインを有するタンパクは、毒性を持ち、および／または、発現が低い傾向があり、発現宿主によって効率的に排除される。したがって、たとえランダムな数のシステインをコードしても、偶数のシステインをコードするDNA配列のみが、機能的ファージ粒子として発現される。したがって、以前に選択された1SSペプチド（単複）（プール）を、2SSペプチド（単複）（プール）に拡大する方法の一つの方法は、単一の、第3固定システイン、および、それより大きな（かつ変動する）数のランダム残基、その内のいくつかは、Cys残基をコードすることが統計的に期待される、を有するライブラリーを創製することである。これらのランダム位置のある既知の割合が、システイン残基をコードし、ファージ増殖によって奇数システインを有する配列が除去された後、第2ペアシステインを有する2SSタンパクは、ファージライブラリーの>50%、好ましくは>60-80%、場合によっては>90-95%を構成する。この新規システイン（単複）および／または新規ランダム化区域は、開始タンパクのN-末端か両末端、または、該タンパクのC-末端か両末端、あるいは、比較的一般的ではないが、開始タンパク配列の内部に存在してもよい。ジスルフィド結合パターンは、構築過程において変化することも可能である。元のジスルフィド結合（単複）は、異なるシステインを連結するジスルフィド結合によって置換されてもよい（新規DBP）。

【0195】

延長法：標的に結合するタンパク（任意の長さ、またはジスルフィド数を持つ）は、それらを、ランダム化ライブラリー配列に融合させることによって延長することが可能である。配列は、通常、いくつかのランダム位置によって隔てられた、一（またはそれ以上）ペア（単複）の、任意に変動する間隔を有するシステインを含む。このようなタンパクのライブラリーを、標的分子に対する結合アフィニティーの強化について選択する。この方法は、最初の結合部位とは別にフォールド形成する、異なる配列の第2結合部位をもたらす可能性が高い。

【0196】

二量化法：特に、ホモ・マルチマーであるか、細胞表面に局在する標的に対しては、以前に選択された結合部位を二重化し、それによって、それぞれが標的の同じ部位に結合することが可能な、同じジスルフィド含有配列の二量体、三量体、四量体、五量体、または六量体を創製することは興味そそられる。もしも標的の複数部位に同時に結合することが可能であるなら、結合アビディティは増す。最適アビディティの実現は、通常、種々の長さ、および、場合によってはさらに種々の組成を持つ各種スペーサを調べることによって、結合部位間の間隔を最適化することを必要とする。ヒトの V E G F に結合する、ホモの二量体の例が本明細書に記載される。結合部位の間に、G l y - S e r から成るスペーサを使用し、長さは、二量体 V E G F 標的に対し最適アビディティを実現するように調節することが可能である。

10

【0197】

既存の C D P シリーズ：各 1 S S、2 S S、または 3 S S 構築体の間隔（「システイン距離パターン」、C D P）が、既存のタンパクファミリーの C D P と同じになるように、例えば、構築の各段階が天然の C D P を使用するように、ジスルフィドを加えることが可能である。さらに、既存の 3 S S、4 S S、または 5 S S スカフォールドに、選ばれた 1 S S または 2 S S タンパクを、同様のループ長となる場所に移植することも可能である。既存のジスルフィド結合パターンを変える目的でジスルフィドを加え、構造変種または D B P 変種のライブラリーを創製することも可能であるし、あるいは、既存の結合パターンを維持することも可能である。D B P に対する調節は、主に、新規システインペアおよび新規ランダム配列が、既存タンパクの一端のみに付加されるのか（既存 D B P を保存する傾向がある）、あるいは、それらの配列は、既存タンパクの両端（すなわち、各端に一システイン）に付加されるのか（D B P に変更をもたらす傾向がある）に依存する。既存のジスルフィド結合（単複）を保存したいのならば、古いシステインペアと、新たに付加されたシステインペア（単複）の間に、いくつかの余分のスペーサ残基を残すのが有効である。このようなスペーサは、任意の配列を持っていてよいが、グリシン富裕スペーサが好ましい（すなわち、G G S、または G G G G S のマルチマー）。標的分子が二量体（可溶性）か、細胞結合性である場合、二つのマイクロタンパクモチーフが、その標的に結合するのを可能とするほど長いスペーサは、両部位における同時結合をもたらし、アビディティの上昇、あるいは見かけのアフィニティーの増加を招く。

20

【0198】

メガプライマー法による構築：メガプライマー法は、配列ライブラリーの存在から生じる複雑性を回避しながら、先行ライブラリーからの新規ライブラリーの創製を可能とする。以前に選択された 1 S S タンパクのプールを含む P C R 断片が生成され、この断片を、一つまたは二つの新規 C y s 残基を有する新規ライブラリーをコードする新規 D N A 断片（オリゴ、または P C R 産物）と重ねる。この重複断片から創製される s s D N A ラン・オフ P C R 産物（「メガプライマー」）は、ベクターに対して相同な末端を含むので、ベクターにアニールされ、このメガプライマーによって置換される区域に終止コドンを含む鋳型による、K u n k e l 様ポリメラーゼ伸長反応を駆動するために使用される。別法として、一对の独自の制限部位を用いて、以前に選択されたベクターのライブラリーの内部に新規ライブラリーを創製することも可能である。ファージタンパク p I I I または p V I I I に対する遺伝学的融合によって、ファージカプシドにおけるタンパクの提示が可能となる。偶数システインを有するタンパクは、i) ファージ増殖、i i) アフィニティー選択、i i i) 遊離チオール精製、および／または、i v) D N A 配列のスクリーニングによって選択することが可能である。この方法を 1 サイクル、または複数サイクル用いることによって、1 S S、2 S S、3 S S、4 S S、5 S S、6 S S、またはより大きなジスルフィド数までジスルフィド含量を積み上げることが可能である。スタート点として任意のジスルフィド数を使用することが可能である。

30

40

【0199】

いくつかの特定の、例示の構築過程を下記に述べる。

【0200】

50

234 設計プロセス：図138を参照されたい。好ましい一方法は‘234’と呼ばれる。その理由は、先ず、三つの結合パターン全てから成る混合物を含む2ジスルフィドライブラリーを創製し、パンニングし、次いで、最良クローンを選び、これを用いて、さらに別の（部分的）ランダム化アミノ酸位置、およびさらに別の1ペアのシステインを有する新規ライブラリーを創製し、このようにして、最大15種の異なる構造を取ることが可能な3ジスルフィドライブラリーを形成し、その内のいくつかは、元の四つのシステインに異なる結合パターンを取らせ、それによって、元の2SS配列の構造進化を可能とするからである。各「ライブラリー伸長セグメント」は、通常、複数のアミノ酸（すなわち、NNK、NNS、または類似の混合コドンによってコードされる）の混合物をコードするいくつかのコドン、プラス、一つ以上のシステイン（外側に配される）を含み、以前に選ばれた配列プールの5'またはN-末端に付加されるか、あるいは、以前に選ばれた配列プールの3'またはC-末端に付加されるか、あるいは両端に付加される。遊離チオールを回避するために、偶数のシステイン（2、4、6）を、各クローンに付加することが望ましい。これは、両端にライブラリー延長セグメントを付加することによって（各端に1システイン、および4-5ランダムコドン）、あるいは、C-末端のみ、またはN-末端のみに付加される二つ（または4または6）のシステイン、および6-8のあいまいコドン（アミノ酸の所望の混合物をコードする）として付加することによって実行することが可能である。このプロセスを複数回繰り返すことが可能である。

10

【0201】

したがって、234指向性進化プロセスは、下記の諸工程から構成される。初期のライブラリー構築（2SS）、標的パンニング（任意、個々のクロンのスクリーニングおよび最良クロンのプール蓄積）、伸長ライブラリーの構築（3SS）、標的パンニング（任意、個々のクロンのスクリーニングおよび最良クロンのプール蓄積）、伸長ライブラリーの構築（4SS）、標的パンニング、および個々のクロンの最終的スクリーニングで、最良の4SS結合因子を特定すること。

20

【0202】

このプロセスについては多くの変法を工夫することが可能である。4、5、6、7、またはそれ以上のジスルフィドを用いること、または、例えば、1-ジスルフィドジャンプの代わりに2ジスルフィドジャンプを用いること、あるいは、一つの標的に対して一つのライブラリーをパンニングし、第2の標的に対し下記のライブラリーの構築を用いることが可能である。前記において、二つの標的は近縁であっても、無関係であってもよい。

30

【0203】

好ましい方法は、天然の3SSにも認められる（好ましくは普通に）CDPを有する2SSライブラリーを作製すること、および、天然の4SSタンパクにも認められるCDPを有する3SSライブラリーを作製することである。このようにして、2SSが熟成して3SSになること、3SSが熟成して4SSになることを正当に確信することが可能になる。

【0204】

3x0-8および4x0-8設計プロセス：図139を参照されたい。この‘3x0-8’および‘4x0-8’の好ましい設計プロセスは、標的パンニングのために最大の構造多様性および配列多様性を提示するために、15通りの3ジスルフィド構造の全て、105通りの4ジスルフィド構造の全てを創製することを目指す。同じ方法を、5-、6-、または場合によっては7-ジスルフィドマイクロタンパク（5x0-8、6x0-8、7x0-8）にまで拡張することが可能である。

40

【0205】

天然の3-ジスルフィドマイクロタンパクの全てのループ長の分析から、ループは、0-10アミノ酸のサイズ範囲を持つ傾向のあることが示された。五つのループ（C1-C2、C2-C3、C3-C4、およびC5-C6）の平均は、きわめて近似している（最長ループのいくつかは望ましくないのを除くと、0-8から3-12の範囲）。ただし、異なるスカフォールドファミリーの間には、ループサイズに大きな差がある。例えば、コノ

50

トキシンのループ C 1 - C 2 は、アナトドメインの 0 A A 長に対し 6 A A 長であり、それ
 でいて両者は、同じジスルフィド結合パターンを有する。

【0206】

配列モチーフ C 1 x 0 - 8 C 2 x 3 - 1 0 C 3 x 0 - 1 0 C 4 x 0 - 8 C 5 x 0 - 9 C
 6 は、天然の 3 S S タンパクの 9 0 % 以上、および、有用な性質を持つ、未知の全ての 3
 S S ミクロタンパクの圧倒的多数をカバーすることが予想される。ライブラリーの構築プ
 ロセスは、等しい長さ、例えば、0 - 8 を持つループでは比較的容易であり、C 1 x 0 -
 8 C 2 x 0 - 8 C 3 x 0 - 8 C 4 x 0 - 8 C 5 x 0 - 8 C 6 のライブラリー配列モチーフ
 、または、この設計の 4 S S バージョン、C 1 x 0 - 8 C 2 x 0 - 8 C 3 x 0 - 8 C 4 x
 0 - 8 C 5 x 0 - 8 C 6 x 0 - 8 C 7 x 0 - 8 C 8 が得られる。使用が可能な他のループ
 長としては、0 - 1 0、0 - 9、0 - 8、0 - 7、0 - 6、0 - 5、0 - 4、1 - 5、1
 - 6、1 - 7、1 - 8、1 - 9、または 1 - 1 0 がある。ただし大抵のループ長が有効で
 あることが期待される。

10

【0207】

このタイプのライブラリーは、ヘテロ性にフォールド形成する多数の配列を含むことが
 予想される。これは、これらの配列は、たくさんの異なる構造を取ることは可能であるが
 、均一な形としては簡単に製造することはできないことを意味する。この異質性は、タン
 パクの生産には不利であるが、多様性の増大は、パンニングと早期のリガンド発見のため
 には有利である。

【0208】

合成タンパク多様性の従来のディスプレイライブラリーでは、クローンは全て、同じ固
 定タンパクスカフォールドを共有する。膨大な多様性を持つ配列が創製されるが、それら
 は皆同じ構造を共有し、著明な構造的多様性は存在しない。逆に、3 x 0 - 8 および 4 x
 0 - 8 ライブラリーは、1 5 通り、場合によっては 1 0 5 通りもの異なる構造体から成る
 等量混合物を含む。

20

【0209】

典型的ファージディスプレイ・ライブラリーは、1 0 e 9 から 1 0 e 1 0 の異なるクロ
 ーンを含み、これらは、通常、それぞれ異なる配列を持つ。しかしながら、パンニングの
 対象になるのは、各配列またはクローンの、平均して、約 1 0 0 0 - 1 0、0 0 0 コピー
 を含む約 1 0 e 1 3 のファージ粒子である。このコピー数は、「ライブラリー当量数」と
 呼ばれる。同じ配列の 1 0 0 0 - 1 0、0 0 0 コピーのそれぞれが、ジスルフィド結合形
 成によって仲介されるフォールド形成の異質性によって異なる構造を取ることが可能で
 ある。したがって、3 x 0 - 8、4 x 0 - 8、または 5 x 0 - 8 ライブラリーの実効ライブ
 ラリーサイズは、単一スカフォールドライブラリーよりも 1 0、1 0 0、または 1 0 0 0
 倍大きい。したがって、この設計のライブラリーは、理論的に可能な構造、ジスルフィド
 結合パターンおよびフォールドの全て、またはほとんどを含むことが予想される。

30

【0210】

平均タンパクを小さいままに維持し、それによって不要の構造の形成を阻止し、かつ、
 所望の構造の頻度を増すために、ループの長さ範囲を狭めることも可能である。中間のル
 ープ長、例えば、2 - 6、2 - 7、2 - 8、2 - 9、または 2 - 1 0 アミノ酸、または 3
 - 4、3 - 5、3 - 6、3 - 7、3 - 8、3 - 9、または 3 - 1 0 アミノ酸、または 4 -
 5、4 - 6、4 - 7、4 - 8、4 - 9、または 4 - 1 0 アミノ酸、または 5 - 6、5 - 7
 、5 - 8、5 - 9、または 5 - 1 0 アミノ酸を使用することも可能である。

40

【0211】

さらに、ライブラリーのために、単一の固定ループ長、通常、1、2、3、4、5、6
 、7、8、9、または 1 0 アミノ酸長を取ることも可能である。

【0212】

平均タンパクサイズを小さいままに維持するための相補性法とは、上限として：2 0、
 2 1、2 2、2 3、2 4、2 5、2 6、2 7、2 8、2 9、3 0、3 1、3 2、3 3、3
 4、3 5、3 6、3 7、3 8、3 9、4 0、4 1、4 2、4 3、4 4、4 5、5 0、5 5

50

、60アミノ酸、および下限として：13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、または35アミノ酸をコードするDNA断片を選択するために、DNA断片サイズ分画ゲルを用いることである。

【0213】

4X6設計プロセス：図140を参照されたい。好ましい方法は、‘3x6’または‘4x6’プロセスであり、これは、3または4個のジスルフィド、および、変動配列を持つことが可能な6個のアミノ酸から成る固定ループサイズを有するライブラリーから始める。4X6ライブラリーのタンパク配列モチーフは、C1x₆ C2x₆ C3x₆ C4x₆ C5x₆ C6x₆ C7x₆ C8である（下付文字は、塩基の混合物（多くの場合、NNK、NNS、または類似のあいまいコドンによってコードされるもの；Cの後の数字は、タンパクにおいてN-末端からC-末端に向かうシステインの順番を指す））。天然のミクロタンパクファミリーでは、結合されるシステイン同士は、タンパク鎖バックボーンにおいて、平均10-14アミノ酸（平均12）隔てられる。我々は、この距離を「ジスルフィドスパン」と呼ぶ。このスパンは、約8-9アミノ酸未満になることは滅多にない。隣接システイン同士がジスルフィド結合を形成する場合、それらは、多くの用途において望ましくないサブドメインを形成する。なぜなら、そのサブドメインは、自身の熱およびプロテアーゼ不安定性プロファイルを持つからである。この不要なサブドメインは、隣接システイン同士の結合を不可能とするほど短いループ長、すなわち、9アミノ酸未満を選ぶことによって根絶することが可能である。6AAの固定間隔が、特に好都合のようである。それは、この間隔が、サブドメインを阻止し、かつ、（非隣接の）システインが、天然タンパクの平均と同じという点で理想的な12アミノ酸隔てられる複数位置を創出するからである。サブドメインの根絶は、69通りの最悪の4SSジスルフィド結合パターンを取り除くので、36通りの最良の4SSジスルフィド結合パターンが得られる。4、5、7、または8アミノ酸、またはそれらの組み合わせから成る固定間隔も実行可能である。

【0214】

既知の、天然の3SSトキシンの大多数は、下記の組成：C1-(x₀₋₁₀)-C2-(x₂₋₁₂)-C3-(x₀₋₁₀)-C4-(x₀₋₁₀)-C5-(x₀₋₁₂)-C6を有する、単一「全スカフォールド」ライブラリーに含まれると考えられる。このようなライブラリーはさらに、未知の天然トキシンの大部分、さらに、より多数の、非天然トキシンを含むと考えられる。このようなライブラリーによってコードされるタンパクの平均長は、1+5+1+7+1+5+1+5+1+5+1=33アミノ酸長となると考えられる。

【0215】

より短いタンパクを創製するには、長い配列をコードするオリゴに対し、より高いモル比の、短い配列をコードするオリゴを使用すること、あるいは、最大ループ長を、10-12aaではなく、僅かな8aaに限定することが可能と考えられる。

【0216】

同様に、下記の組成を有する全スカフォールドライブラリーは、105通りの異なるジスルフィド結合パターン、および可能な1000を超えるフォールドを有する、4-ジスルフィドHDDトキシンの大多数を含むと考えられる。

【0217】

C1-(x₀₋₁₀)-C2-(x₀₋₁₀)-C3-(x₀₋₁₀)-C4-(x₀₋₁₀)-C5-(x₀₋₁₀)-C6-(x₀₋₁₀)-C7-(x₀₋₁₀)-C8
および、5-ジスルフィドの「全スカフォールド」ライブラリーは、

C1-(x₀₋₁₀)-C2-(x₀₋₁₀)-C3-(x₀₋₁₀)-C4-(x₀₋₁₀)-C5-(x₀₋₁₀)-C6-(x₀₋₁₀)-C7-(x₀₋₁₀)-C8
(x₀₋₁₀)-C9-(x₀₋₁₀)-C10によって指定されることが考えられる。

【0218】

xは通常、アミノ酸の望ましい混合物を指す。アミノ酸の混合物をコードするのにNN

10

20

30

40

50

Nコドンを使用することが可能であるが、他のコドンも利点を持つ。各コドンは、異なるアミノ酸混合物を提供する。

【0219】

例えば、NNKは、終止コドンの頻度を3倍下げる。異なる用途には異なるコドンが有用である。親水性アミノ酸を優先する混合物が望ましく、終止コドン、トリプトファン、その他の疎水性アミノ酸の回避、および、ループにおけるシステインの回避も望ましい。分子生物学者であれば、所望の混合物を生産するコドンの選択法を知っている。コドンは、通常、コドンの第1塩基として、A、C、G、T、または、混合塩基文字N、M、K、S、W、Y、R、B、D、V、またはHを含み、コドンの第2塩基として、A、C、G、T、または、混合塩基文字N、M、K、S、W、Y、R、B、D、V、またはHを含み、かつ、コドンの第3塩基として、A、C、G、T、または、混合塩基文字N、M、K、S、W、Y、R、B、D、V、またはHを含むように選ばれ、それぞれが、異なるアミノ酸から成る混合物をコードする、多数の可能なコドンをもたらす。

10

【0220】

天然のHDDタンパクのループ配列は、タンパクのフォールド形成に役割を果たすと思われる、少数の固定残基を含む。従来の方法は、単純に、ランダムコドンを用い、それらが真にフォールド形成に重要である場合は、多様性にそれらの残基の支給をまかせるというものである。このランダムコドン法は、各位置においてアミノ酸の天然組成を用いるライブラリーに比べてライブラリーの質が低下するが、新規フォールドの可能性を探る点では最良である。

20

【0221】

しかしながら、もしも例えば、フォールド形成または機能のためにWが必要とされるのに、その位置にNNKコドンが使用される場合、ライブラリーの内の僅かに1/64クロンがこの要求を満たすだけであるので、このライブラリーの実効サイズは64倍下がる。これは、有用な結合因子の獲得を阻止するのに十分である。したがって、天然配列において固定されるように見える任意の残基は、ライブラリーにおいても固定するのが重要と考えられる。

【0222】

ランダムコドン(NNK、または上述の他の、多くのものの内の一つ)の使用に代わる代替法は、ある特定のタンパクファミリーのループの正確な共通配列によってオリゴヌクレオチドを合成することである。この方法は、ループ2設計は、ライブラリーのループ2の位置にのみ組み込むこと、ループ3配列は、ループ3の位置にのみ組み込むことを要求する。これは、複数のシステインが、重複反応が起こる場合、それぞれ、三つのシステインコドンの内の異なる一つによってコードされる場合に実現される。より効率的な重複PCR反応を実現するために、cysコドンの前後の1から3塩基も同様に固定してよい。重複反応効率は、ライブラリーの多様性を制限する可能性があるので、これは、検出されない、または簡単に調節されない重要な危険因子である。一般に、数塩基の付加が、低ライブラリー多様性の重大な危機を下げるための効果的方法である。

30

【0223】

種々のファミリーのループ配列の全てを混ぜ合わせ、それらを重複PCRで組み込んだ後は、合成ループ配列は全て、その天然位置においてのみ出現するはずである。このライブラリー法は、種々のファミリー由来のループの相互関係にシャッフリングをもたらす。

40

【0224】

ライブラリーの多様性の増大：天然および指向性進化の力は、選択圧に曝される多様性に関連する。より多くの、より多様なクロンから得られた選抜体は、一般に、より良い結果を生む。生物は、遺伝子の数を超えてタンパク構造の多様性を増すために複数の方法を用いる。この拡大された天然の多様性によって、選択に対するより多くの解法が得られ、かつ、天然進化力の増大が得られる。

【0225】

指向性進化の力を増すことを目的として、同じ数のクロン、または同数の配列から得

50

ることが可能な構造の多様性を増すには、様々な、たくさんの方がある。

【0226】

この原理は、単一遺伝子、複数遺伝子経路、全体ゲノム（前核細胞、古細胞、真核細胞）、および生物の全共同体（すなわち、細菌共同体）の最適化にも応用することが可能である。

【0227】

一般に、単一遺伝子の発現は、様々な、異なるmRNA配列を生み出す。これは、複数のプロモーター、選択的スプライシング、トランススプライシング、または分解のためである。各mRNA配列は、違ったやり方でフォールド形成し、様々な、異なった構造を取り、かつ、その結果も、他のRNA（マイクロ、tRNA、またはmRNA）の存在によって修飾されるばかりでなく、RNAと相互作用を持つタンパクによっても修飾される。これらのmRNA構造はそれぞれ、複数の翻訳開始および停止シグナルの存在によって幾分違ったやり方で翻訳されたり、リボソーム上の異なる休止による変種、あるいは、低度ではあるが、様々な度合いの、「非天然」アミノ酸を含む、アミノ酸の組み込みミスがある。さらに、各タンパク翻訳産物は異なるやり方でフォールド形成する：あるものは凝集し、あるものはミスフォールド形成し、あるものはプロテアーゼによって分解され、あるものはユビキチン化され、あるものは折りたたまれて複数の安定構造を形成する。重要で、実際の差別機序は、タンパクの誘導体形成、アミノ酸側鎖の化学的変化、および、小型分子、例えば、糖類、PEGのようなポリマーの、タンパク鎖に対する化学的連結である。これらの化学法は、全体ライブラリー（大抵の）、または精製単一タンパクに対して適用することが可能である。

【0228】

ライブラリーに適用した場合、小型分子は、特に低量で与えた場合、異種集団が得られるので多様性を劇的に増す。例えば、5リシン残基を含むタンパクライブラリーにおけるリシン残基に対し、PEGまたは炭水化物分子を、非網羅的に接合した場合、5一階乗＋1種類の分子（122変種）が得られる。最良変種をパンニングによって選び、次に、数種の標識レシピを、どのレシピが最良結果を与えるかを見出すために、ライブラリー等価物、クローンプール、または単一クローンに付与する。さらに、このタンパク配列を進化させ、所望の活性の保持および改善について選択する。例えば、最良突然変異は、活性に寄与しない4個のリシンを失うかもしれないが、誘導体形成された場合、活性レベルの増加をもたらすリシンは保持する。タンパクの誘導体形成に用いられる試薬（すなわち、Pierce Chemicalのオンライン・カタログ）は全て、原理的には、この方法に使用することが可能である。細胞機能と多様性のための一意で、安定な構造と、細胞の進化を加速する、ある程度の不安定性の間には微妙な平衡がある。

【0229】

これらの機序はそれぞれ、実験的介入のためのポテンシャル点となる。これらの調節はそれぞれ、自然進化によって現在の変動レベルに設定されるが、その多様性は、指向性進化の目標に応じて増加または減少することが可能である。

【0230】

特に商業的関心の集まる領域は、ディスプレイライブラリー（ファージ、酵母、細菌表面、ポリソーム、リボソーム、プロフュージョン、または遺伝子融合ライブラリー）による結合タンパクの指向性進化である。最良精選クロンの頻度および品質は、直接、ライブラリーのサイズに相関することが十分に確かめられている。ライブラリーが大きければ大きいほど、結合因子の数は大きくなり、かつ、最良クローンはより良くなる。このため、どんどん大きいライブラリーを創製するために様々な方法が開発されており、例えば、10e6のクローンから成る二つの免疫グロブリンライブラリーを組み合わせ、10e12クローンから成る単一ライブラリーを形成する組み換え法がそうである。しかしながら、この例では、ライブラリータンパクは全て、同じ免疫グロブリンフォールドを持つので、単一構造における多様性に重点を置いており、ある用途、すなわち、全体抗体産物にとっては有利であるが、種々の構造の多様性を創出するためには好適ではない。ライブラ

リーにおけるクローン数の増加ではなく、単一配列から創製される構造の数を増すことによって、ライブラリーの実効サイズを増すことも可能である。

【0231】

ライブラリーにおけるクローン数の増加ではなく、ライブラリーの多様性を増す別の方法は、各クローンによって採用される構造の多様性を増すことである。これは、脱安定化タンパクを用いることによって実現が可能である。この脱安定化タンパクは、それぞれが、短時間ではあるが、多様な構造として存在する点で、溶融小球により似る。この方法は、高度に構造化されたタンパクから成るライブラリーにおいてはアクセスされない、新規バックボーン構造を含む、はるかに大きい空間の探索を可能とする。この比較的大規模な探索によって、比較的大規模な最適フォールドの特定が可能となり、さらに、指向性進化を用いて、この新規フォールドの安定なフォールド形成変種および均一に製造することが不可能な変種を創製することが可能となる。

10

【0232】

標的は通常タンパクであるが、核酸（DNA、RNA、PNA）、炭水化物、脂質、代謝産物、または任意の生物的、または非生物的材料であってもよい。ライブラリータンパクは、（部分的に）構造未形成であるので、該タンパクは、それぞれ短時間であるが、たくさんの異なる構造を取る。これは、ライブラリーの分子の多様性を増すので、多数のライブラリー等価物の使用には好都合である。標準的ファージライブラリーのパンニングのためには、通常、100個のライブラリー等価物、または、ライブラリーが10e10の多様性を持つ場合、10e12ファージを用いる。ライブラリーからある特定の（構造化された）クローンを高い信頼度で回収するためにはこの100倍の余剰が必要であることが実験的に見出されている。高アフィニティークローンのためには、使用する余剰はもっと低くてもよいが、低アフィニティークローンのためには、使用する余剰はもっと高くなければならない。

20

【0233】

多様性を創出する他の方法と区別して、我々は、これを、「一過性多様性」と呼ぶ。なぜなら、この多様性は、それぞれが短時間起こる多数の構造体によって実現されるからである。同じ単一遺伝子からの多様な構造体の創製は、生物進化にとって重要な原理であり、生物組織の多数レベルに存在する。

【0234】

ディスプレイライブラリーの多様性の拡大：ファージライブラリーは、通常、10e10の異なる配列から成る多様性を持つ、約10e14ファージを含む。アフィニティークロマトグラフィーは、このようなライブラリー（10e10濃縮体）から、結合タンパクを発現する単一配列を選択することが可能であることが十分確かめられている。高いアフィニティで結合することが可能なファージのほぼ100%がアフィニティークラムによって結合されるのであるから、この方法によって（10e14濃縮体）ファージの単一コピーも、高い信頼度で選択することが可能である。

30

【0235】

ファージに提示されるペプチドは、通常、10e3-10e6の異なる、不安定な立体配置において存在し、その内の一つだけがカラムに結合する。カラム結合は、そのペプチドの活性的立体配置を安定化するので、そのようなペプチドは効率的に濃縮され、濃縮体（10e17-10e20）を生成する。したがって、バックボーン立体配置における屈曲性は、ライブラリーの実効サイズを10e20に増す。パンニングの第1ラウンド後、多様性は既に1000倍低下し、そのため、その後のライブラリーでは、各クローンは、1000以上のコピーで表されることになる。これは、タンパクが取る、異なる一過性構造の全てが、統計的に十分抽出されることを意味する。さらにその後の指向性進化の過程において、目標は、標的に対して高いアフィニティを示す構造体として存在する時間割合の大きいクローンを選ぶことである。目的は、選択と組み合わせて各種の突然変異法を用いることによって、アフィニティを徐々に改善するばかりでなく、タンパクの安定性を向上させることである。

40

50

【0236】

標的誘発性フォールド形成：ミクロタンパクの構造は、標的結合によって誘発することが可能である（標的結合後ジスルフィドを形成することによって）、あるいは、ミクロタンパクの構造は、その標的に結合している際に最適化することが可能である。

【0237】

標的への結合は、必ず、ある程度の誘発性適合を含み、したがって、ジスルフィドの内のいくつか（結合された部分のもの）を安定化し、他のジスルフィドを脱安定化し、還元剤に対する差別的感受性をもたらすことが予想される。還元剤および酸化剤において（各種濃度および間隔で）滴定することによって、もっとも安定性の低いジスルフィドの急速な還元および再酸化が可能とされ、これは、結合パターンに変化があると、構造的適応をもたらす、結合標的に対するより優れた適合を生ずる。この方法は、最高の結合アフィニティを有するクローンの生存率を高める。

10

【0238】

生産のためには、タンパクのフォールド形成が、標的独立性となるまで進化することが望ましい場合がある。

【0239】

ミクロタンパクのアミノ酸組成の最適化：大抵のタンパクまたはタンパクドメインは、タンパクの安定性および立体配置にとって決定的に重要な疎水性コアを含む。これらのタンパクの疎水性コアは、高い割合の疎水性アミノ酸を含む。アミノ酸の特性は、その疎水性によって決めることが可能である。いくつかのスケールが開発されている。一般的に使用されるスケールは、(Levit, M (1976) J Mol Biol 104, 59, #3233) によって開発されたもので、これは、(Hopp, TP, ら、(1981) Proc Natl Acad Sci USA 78, 3824, #3232) に掲載される。疎水性残基はさらに、脂肪族残基のロイシン、イソロイシン、バリン、およびメチオニン、および芳香族残基のトリプトファン、フェニルアラニン、およびチロシンに分けられる。図1は、Brooks, DJ, ら、(2002) Mol Biol Evol 19, 1645, #3234 に発表された全てのタンパクにおけるアミノ酸量を比較し、さらに、Gupta, A., ら、(2004) Protein Sci, 13: 2045-58 に発表されたデータベースに含まれる8550のミクロタンパクドメインについて計算されたアミノ酸の平均量を付け加える。

20

30

【0240】

図13を参照されたい。タンパクにおけるアミノ酸の出現率である。図は、他のタンパクに比べ、ミクロタンパクでは、脂肪族疎水性アミノ酸の量が有意に低い傾向のあることが明らかに示す。このことは、従来技術でこれまで認識されていない。逆に、芳香族疎水性アミノ酸(W, F, Y)は、平均的タンパクと同様である。脂肪族アミノ酸の量がこのように低いのは、ミクロタンパク構造は、疎水性コアを不要とする、いくつかのジスルフィド結合によって安定化されるという事実を反映する。脂肪族炭素原子を含む、他の、いくつかのアミノ酸残基（グルタミン酸塩、リシン、アラニン）も、他のタンパクに比べ、ミクロタンパクにおける量は低い。

【0241】

40

低疎水性を有するスカフォールドの有用性：タンパクにおける脂肪族アミノ酸量の低下は、製薬、およびその他の用途におけるそれらのタンパクの有用性を著明に増す。多くのタンパクは、フォールド形成時凝集体を形成する傾向を有する。これは、タンパクが、異種宿主において高濃度で生産される場合、および、インビトロで再度天然化される場合、さらに悪化する可能性がある。凝集およびミスフォールド形成は、商業的生産時、タンパクの収率を著明に下げる可能性がある。タンパク配列における脂肪族アミノ酸の割合を下げることによって、凝集体を形成する傾向を下げ、したがって、適正にフォールド形成したタンパクの収率を上げることが可能となる。

【0242】

脂肪族アミノ酸量の低いタンパクは、他のタンパクに比べて、免疫原性がより低い。脂

50

脂肪族アミノ酸は、免疫反応の形成において決定的ステップとなる、MHCに対するペプチドの結合の度合いを増す傾向がある。その結果、低い割合の脂肪族アミノ酸しか含まないタンパクは、大抵の、他のタンパクに比べ、含むT細胞エпитープの数はより少ない。

【0243】

脂肪族残基は、親油性相互作用を形成する傾向を有する。その結果、大きい割合の脂肪族アミノ酸を有するタンパクは、他のタンパク、膜、および、他の表面に、非特異的に結合する可能性がより高い。タンパクの表面に暴露される脂肪族残基は、他のタンパクと、非特異的相互作用を持つ傾向が特に高い。マイクロタンパク中のアミノ酸の多くは、マイクロタンパクの小型サイズのため、ある程度表面を露出させる。

【0244】

したがって、本発明は、3個以上のジスルフィドを有し、20-60アミノ酸から成る単ドメインを含み、かつ、ヒトの血清暴露タンパクに結合し、5%未満の脂肪族アミノ酸しか含まない非天然タンパクを提供する。要すれば、この非天然タンパクは、4%、3%、2%未満、場合によっては1%未満の脂肪族アミノ酸しか含まない。さらに、本発明は、このような性質を有する非天然タンパクのライブラリーを提供する。

【0245】

疎水性の低いスカフォールドの特定：大抵のマイクロタンパクは、大抵の通常タンパクに比べ、その含む脂肪族アミノ酸の数が少ないが、種々のマイクロタンパクファミリーにおいて、脂肪族アミノ酸の含量には著明なばらつきがある。表4は、脂肪族含量の低い製薬タンパクの加工においてスタート点として特に有用な、いくつかのマイクロタンパク・ファミリーを列挙する。

【0246】

免疫原性の低いタンパクの設計：免疫原性の低いタンパクは、ヒトに投与した場合、有害な免疫反応を引き起こす確率が低いので、治療薬としてより望ましい。いくつかの点で、望ましい標的結合特性を有する主題のマイクロタンパクは、一般に、同じ標的に結合するが、所望のシステイン結合パターンまたはフォールドを持たないタンパクよりも免疫原性が低い。一実施態様では、主題のマイクロタンパクは、免疫原性が、1倍低い、好ましくは2倍低い、好ましくは3倍低い、好ましくは5倍低い、好ましくは10倍低い、好ましくは100倍低い、好ましくは500倍低い、さらに好ましくは1000倍低い。ある実施態様では、免疫原性の低いマイクロタンパクは、本明細書に記載されるHDDタンパクである。

【0247】

タンパクの免疫原性は、TEPITOPEのようなプログラムを用いて予測することが可能である。このプログラムは、大きな1組のアフィニティー測定値を用いて、主要なヒトのHMCクラスII対立遺伝子全てにたいする、免疫原由来の、重複する9個のアミノ酸ペプチド全てについて、その結合アフィニティーを計算する（Sturnioloら、1999；www.biovation.com；www.epivax.com；www.algonomics.com）。このようなプログラムは、ヒトのT細胞エпитープの予測および除去に広く使用され、その使用は、FDAによって奨励されている。

【0248】

これらのアルゴリズムを用いて、我々は、25-90残基、および10%システインを有するマイクロタンパクが、平均的タンパクよりも、通常、MHCI Iに対する予測結合アフィニティーが316倍低いことを見出した。図166の赤色曲線は、中央長が372アミノ酸である、全て26,000のヒト・タンパクの予測免疫原性を示す。青色曲線は、中央長が38アミノ酸である、全て10,500のマイクロタンパクの予測免疫原性を示す。緑色曲線は、マイクロタンパクと同じ長さ分布を持つ、タンパク断片から成る非天然グループの予測免疫原性を示す。各グループの平均スコアの比較から、マイクロタンパク単独の1-logのサイズ低下は、免疫原性の67倍の低下をもたらし、マイクロタンパクのアミノ酸組成は、さらに4.7倍の低下を生ずることが示された。図167の上段パネルは、脂肪族疎水性アミノ酸（I、V、M、L）が、TEPITOPEアルゴリズム（Stur

10

20

30

40

50

niolo et al 1999) においてもっとも強力な接触因子としてランクされ、予測免疫原性にもっとも多く寄与することを示す。図167下段パネルは、上記脂肪族残基が、ヒトのタンパクと比べると、マイクロタンパクではもっとも出現率が低く、予測免疫原性における、組成物由来1-log低下の大部分を説明することを示す。

【0249】

マイクロタンパクにおける脂肪族疎水性残基の低レベルは、他のタンパクでは通常である疎水性コアの欠如によって可能とされる。その代わりに、マイクロタンパクは、架橋結合して、鎖内ジスルフィドを形成する、少数のシステインを含む。この、少数のジスルフィドによる、多数の疎水性アミノ酸の置換は、タンパクが安定でいられる最小サイズを下げ、それによって、マイクロタンパクがより小さくなり、脂肪族アミノ酸の頻度を減らし、予測免疫原性において3-logの低下をもたらすことを可能にする。

【0250】

免疫原性低下は、様々の表示、例えば、1) 樹状細胞(DC)などの抗原提示細胞(APC)が、免疫タンパクからペプチドを放出する能力(抗原処理); 2) HLA II分子に対する結合を決める、これらのペプチドにおけるT細胞エпитオプの存在; 3) APC表面におけるペプチド-HLA II複合体を認識する、血中の未洗練T細胞の数; および、4) 血清における抗体レベルを含む表示によって測定することが可能である。

【0251】

タンパクの免疫原性を下げるためには多くの方法があり、その全てが、HDDおよび非HDDタンパクに適用することが可能である。一つの方法は、コンピュータのモデル化および合理的設計を通じてジスルフィドを加えることである。もう一つの方法は、指向性進化または合理的設計を用いてタンパクを微調整することによって既存のジスルフィドを改善することである。ジスルフィドをタンパクの内部に納めるか、または、そのシステインに、保護作用を持つアミノ酸側鎖を側接させることによって、ジスルフィドを化学的攻撃から保護することが可能である。タンパクの免疫原性はまた、TEPITOPEまたはPropredなどのプログラムを用いて予測することが可能である。このプログラムは、大きな1組のアフィニティー測定値を用いて、主要なヒトのHMCクラスII対立遺伝子全てにたいする(MHCクラスIIに対して使用される別のプログラムもある)、免疫原由来の、重複する9個のアミノ酸ペプチド全てについて、その結合アフィニティーを計算する。Sturniolo, T., ら、(1999) Generation of tissue-specific and promiscuous HLA ligand database using DNA microarrays and virtual HLA class II matrices. Nature Biotechnol, 17:555. さらに、www.algonomics.com; www.biovation.com; およびwww.genencor.comも参照されたい。このようなプログラムは、ヒトT細胞エпитオプの予測および除去のために広く使用されており、その使用はFDAによって奨励されている。

【0252】

免疫原性の低いマイクロタンパクを生成するためのもう一つの方法は、化学的架橋剤による、タンパク内架橋を介するものである。多様な架橋剤が、Pierceなどの販売業者から入手が可能である。適用可能な架橋剤としては、アルギニン反応性架橋剤、ホモ双官能性架橋剤、例えば、アミン反応性ホモ双官能性架橋剤、スルフヒドリル反応性ホモ双官能性架橋剤、ヘテロ双官能性架橋剤、例えば、アミン-カルボキシル反応性ヘテロ双官能性架橋剤、およびアミノ基反応性ヘテロ双官能性架橋剤が挙げられる。

【0253】

さらにもう一つの方法は、複数結合部位を持つ小型分子を作製し、各ドメインを、2から3の結合部位に分離することである。例えば、ドメインの一方の面は、一つの標的に結合し、他の半分は、もう一つの標的に結合する。この二つの面は平行して設計し(すなわち、同時に別々のライブラリーで)、融合して一つのドメインにまとめることが可能である。それに代わるものとして、二つの面を継時的に設計し、面1の残基において一ライブ

ライブラリーを創製し、このライブラリーを、標的 1 に対する結合についてパンニングし、一つ以上の最良クローンを選択し、残余のアミノ酸、ライブラリー 1 で使用されなかったアミノ酸において新規ライブラリー 2 を創製し、次いで、標的 2 に対してパンニングを行い、標的 2 に対する結合因子、および標的 1 に対する結合の保持についてスクリーニングを行う方法がある。面 1 に対するアミノ酸は、面 2 に対するアミノ酸と相互にかみ合う傾向を持つので、これらのライブラリーを、異なる配列を持つクローンプールとして構築することは、いくつかのアミノ酸を固定し、これらの固定アミノ酸が、PCR による重複伸長に必要な接触点となるようにすれば、簡単に実行が可能である。システインは固定化される傾向を持つので、これらは、種々のオリゴヌクレオチドのための重複点として論理的選択となる。しかしながら、重複点は、それが 4 個以上の塩基を持つと、よりよく活動するので、システインの一侧にさらに別のアミノ酸を固定するのが有用である。したがって、2 面ライブラリーのスカフォールドは、3 組のアミノ酸および塩基を有する：面 1 / ライブラリー 1 用に 1 組、面 2 / ライブラリー 2 用に 1 組、および、重複伸長によってこの二つのライブラリーを結合するための固定された 1 組である。制限部位を用いることは原理的には可能であるが、重複法の方が一般にうまく作動する。

10

【0254】

さらにもう一つの方法は、システイン間ループの長さを最小化することによってタンパクのサイズを下げることである。典型的方法は、ライブラリーにおいて、あるものは天然に見られ、あるものは天然に見られるものよりも短い、ある範囲のループ長を用いることである。

20

【0255】

さらにもう一つの方法は、親水性を増すことである。HDD タンパクの多くは、きわめて親水性が高く、これは、機能（特異性、非免疫原性）のためばかりでなく、タンパクのフォールド形成のためにも重要である場合がある。親水性は、タンパクライブラリーの各位置に使用されるアミノ酸混合物において、オリゴヌクレオチドの合成のために所望のコドン（の混合物）を選ぶことによって調整することが可能である。好ましい一般的方法は、各アミノ酸位置について天然組成を模倣することであり、これを、いくつかの所望の残基に都合のよいように変形する。クローンを、サイズ、および DNA 配列決定によって親水性についてスクリーニングする。前述の各種の方法を、単独で、あるいは組み合わせて使用することが可能である。

30

【0256】

さらに続く修飾のために、主題のミクロタンパクのいずれのものも使用が可能である。非限定的例としては、HDD タンパク、例えば、修飾された A ドメイン、LNR / DSL / PD、TNFR、アナト、ベータインテグリン、Kunitz、および動物のトキシンファミリー、トキシン 1、2、3、4、5、6、7、8、9、10、11、12、ミオトキシン、コノトキシン、および、デルター、およびオメガアトラコトキシンが挙げられる。本明細書に記載される脱免疫化法は、多様なヒトまたは霊長類タンパク、例えば、サイトカイン、増殖因子、受容体細胞外ドメイン、ケモカインなどに適用することが可能である。この方法はまた、他の、非 HDD スカフォールドタンパク、例えば、フィブロネクチン III を含む免疫グロブリン、および、アンキリン、タンパク A、ユビキチン、クリスタリン、リポカリンに適用することも可能である。免疫原性が最小化される限り、非ヒトスカフォールドの方が、天然のヒトタンパクとの、免疫性交差反応の可能性が低下するために、（近一）天然ヒトタンパクおよびヒト由来スカフォールドよりも好ましい。

40

【0257】

HDD タンパクの免疫原性低下を定量するにはいくつかの方法がある。例えば、ヒトまたは動物 APC によるタンパク分解を定量することが可能である。このアッセイは、ヒトまたは動物の抗原提示細胞、APC 由来リソソーム、または APC プロテアーゼに、対象タンパクを添加すること、および、該タンパクの分解を、例えば、SDS-PAGE によって求めることを含む。APC は、血中単球由来の、または、他の標準法によって得られた樹状細胞であってもよい。ヒトの APC ではなく、動物のものを使用してもよく、あ

50

るいは、丸ごとの細胞ではなく、細胞分解物を使用してもよく、あるいは、一つ以上の酵素分画、または、リソソームのような細胞分画を使用してもよい。細胞の分解は、分解性 SDS-PAGE ゲル分析によってもっとも簡単に定量される。分解されたタンパクは、ゲルにおいて見かけ上、より低い分子量としてより速やかに移動する。一つの方法は、各クローンを蛍光的に、または放射能的に（放射能： ^3H 、 ^{14}C 、 ^{35}S ；染料および蛍光標識、例えば、FITC、ローダミン、Cy5、Cy3 など）、あるいは、他の、任意の、適切な化学的標識で、対象タンパク、および、その分解産物のみが、UV 暴露またはオートラジオグラフィーによってゲル上に可視化されるように標識することである。ウェスタンブロットにおいて抗体を用いて検出することが可能な、ペプチドタグ・タンパクを用いることも可能である。

10

【0258】

免疫原性を定量するためのもう一つの方法は、タンパクの凝集傾向を定量することである。タンパク凝集は、光散乱によって簡単に定量され、ダイナミック光散乱装置（DLS）、または分光光度計（すなわち、OD300-600 対 OD280）によって実行することが可能である。

【0259】

また、T 細胞刺激およびサイトカイン活性化レベルを定量することも可能である。サイトカイン活性化は、ヒトの PBMC において、FACS によって、免疫系が刺激されたことを示す、樹状細胞（CD83 など）における活性化抗原の有無、T 細胞活性化（CD69、IL-2r など）の有無の外、多くの共刺激因子（CD28、CD80、CD86）の有無に関して測定される。さらに細胞は、標準的 ELISA アッセイを用いて、サイトカイン、例えば、IL-2、4、5、6、8、10、TNF アルファ、ベータ、IFN ガンマ、IL-1 ベータなどの生産について調べることが可能である。標準的マイトゲンおよび LPS が好ましいコントロールとして使用が可能である。

20

【0260】

さらに、T 細胞受容体に対する結合を定量することも可能である。Toll 様受容体 1-9（TLR1-TLR9）に対する、治療タンパクの結合は、自然免疫の有用なインディケータである。いくつかの販売業者、例えば、Invivogen が、細胞構築体において、トランスジェニック Toll-受容体接続リポーター遺伝子を全て供給する。

【0261】

さらに、タンパクを、宿主動物、例えば、ウサギおよびマウスに直接注入することによって、タンパクの免疫原性を評価するための動物実験を行うことも可能である。

30

【0262】

下記は、HLA II に対する結合アフィニティーの低いマイクロタンパクの加工法の一例を与える。図 161 を参照されたい。ヘルパー T 細胞の活性化は、最重要ステップであり、外来タンパクに対する免疫反応の起動にとって必須である。T 細胞活性化は、抗原提示細胞（APC）による抗原摂取、抗原の分解によるペプチド変換、および、APC 表面におけるペプチドの提示、およびヒト白血球抗原 DR 群（HLA-DR）のタンパクとの複合体形成を含む。HLA-DR 分子は、提示されたペプチドと相互作用を持つ、複数の結合ポケットを含む。これら HLA-DR ポケットの特異性は、インビトロで測定することが可能であり、得られた特異性プロフィールを用いて、各種 HLA-DR タイプに対するペプチドの結合アフィニティーを予測することが可能である（Hammer, J. (1995) Curr Opin Immunol, 7: 263-9)。それを用いると HLA-DR 結合配列を特定することが可能なコンピュータプログラムが記載されている（Sturniolo, T., ら、(1999) Nat Biotechnol, 17: 555-61）。本発明は、親マイクロタンパクの所望の製薬学的、およびその他の性質は維持しながら、HLA-DR に対する結合を下げるやり方でマイクロタンパク配列を修飾する目的で、これらのアルゴリズムを利用する。第 1 ステップとして、親マイクロタンパクの配列を、HLA-DR 予測アルゴリズムを用いて分析する。親配列における非システイン残基の、全ての、単一アミノ酸突然変異を、親配列と比較し、HLA-DR タイプに対する結

40

50

合を予測する。目的は、親ミクロタンパク、またはその誘導体によって治療される予定の患者集団において高頻度に出現する、HLA-D Rに対する結合を下げると予測される、一組の突然変異を特定することである。その後、ライブラリー中の変異種が、HLA-D R結合を下げると予測される、一つ以上の突然変異を含む、組み合わせライブラリーを構築する。計画された突然変異のサブセットを含む、いくつかのサブライブラリーを構築することは有利な場合がある。次に、得られたライブラリー、またはサブライブラリーについて、適切な標的に結合する変異種を特定するためにスクリーニングが行われてもよい。さらに、ライブラリーのメンバーについて、安定性、可溶性、発現レベル、および、最終的性質にとって重要な、他の性質に関してスクリーニングを行ってもよい。スクリーニングの前に、所望の標的結合アフィニティーおよび特異性を保持する、組み合わせ変異種を単離するために、ファージパンニングまたは同様の濃縮法を、この組み合わせライブラリーに対して実行してもよい。このプロセスによって、親タンパクの所望の性質は全て保持するが、HLA-D Rに対する結合が低下し、したがって、免疫原性が低下すると予測された、親ミクロタンパクの変異種が特定される。要すれば任意に、得られた改良変異種に対し、HLA-D R結合性配列除去の、さらにもう1ラウンドを実行してもよい。この後続の1ラウンドは、単純に前述の手順の繰り返しであってもよい。それに代わるものとして、第2の組み合わせライブラリーを、所望のミクロタンパク機能と適合し、HLA-D R結合をさらに下げると予測されるプロセスの第1ラウンドにおいて特定された突然変異のみに限定してもよい。このプロセスの第2ラウンドを、これらのあらかじめ選択された突然変異に限定することによって、比較的小さいライブラリーを構築しながら、改良変異種単離の頻度を増すことが可能である。

【0263】

【化17】

表4. 脂肪酸アミノ酸が低量なミクロタンパクファミリー

PFAM	ファミリーサイズ	長さ	脂肪酸アミノ酸 (%)	摘要	供給源
PF02977	3	27.0	0.00	カルボキシペプチダーゼA	植物
PF05374	4	19.0	2.63	Mu-コノトキシン	イモ貝
PF00734	42	18.1	4.07	菌類セルロース結合ドメイン	菌類
PF00187	228	36.2	4.93	キチン認識タンパク	植物
PF06357	7	33.0	6.06	オメガ-アトラトキシン	蜘蛛
PF05294	11	32.6	7.24	さそり短小トキシン	さそり
PF05453	6	24.0	7.64	BmTXKS1トキシンファミリー	さそり
PF05353	5	42.2	8.06	デルタアトラトキシン	
PF05375	24	29.5	8.63	パシファスチン阻害剤	いなご
PF00200	285	64.1	8.68	ジスインテグリン	蛇
PF01033	68	35.6	9.00	ソマトメジン	哺乳類
PF00304	105	44.8	9.08	ガンマーチオニン	植物

平均的タンパクは、26.1%の脂肪酸アミノ酸を含む。

治療タンパクにおける疎水性アミノ酸の割合を下げる方法

前述のように、脂肪酸アミノ酸が低量のミクロタンパクを創製する方法の一つは、僅かな脂肪酸アミノ酸しか含まないスカフォールドおよびライブラリーから開始することによる。さらに、各種のタンパク工学技術を用いることによって、タンパク中の脂肪酸アミノ

酸の量を下げることが可能である。例えば、1個または数個の脂肪族アミノ酸が、たくさんの親水性アミノ酸の出現を可能とするランダムコドンによって置換される、タンパクライブラリーを構築することが可能である。特に興味深いのは、親水性アミノ酸の割合を大きくするが、脂肪族または疎水性アミノ酸の割合を下げるあいまいコドンである。例えば、コドン V V K は、12 種のアミノ酸（アラニン、アスパラギン酸塩、グルタミン酸塩、グリシン、ヒスチジン、リシン、アルパラギン、プロリン、グルタミン、アルギニン、セリン、トレオニン）の出現を可能にするが、脂肪族および芳香族アミノ酸は全て回避する。このようなライブラリーから所望の性質を持つタンパクを単離すること、したがって、芳香族疎水性および脂肪族疎水性アミノ酸の量を下げることが可能である。さらに、脂肪族アミノ酸を含む、複数のアミノ酸位置をランダム化する、組み合わせタンパクライブラリーを構築することが可能である。このようなライブラリーから得られる複数の変異種についてその配列および性能を決定することによって、前記タンパクにおいて、親水性アミノ酸による置換を可能とする位置を特定することが可能となる。

10

スカフォールドの有用性を評価する方法

天然配列の特定ファミリーに基づいて設計を創出せよ。各アミノ酸位置において、その位置におけるアミノ酸の天然の多様性を反映するアミノ酸の混合物が使用される。タンパクの N-末端に H A タグが付加され、C-末端には H i s 6 タグが付加される。

【0264】

このタンパク設計をコードするオリゴヌクレオチドが合成される。同時に、単一的に、または、異なる設計の混合物として、1-30 の異なる設計が構築される。

20

【0265】

（主題組成物の発現）

細胞内・対・細胞外環境

ジスルフィド結合は、主に、分泌（細胞外ゾル）タンパクである。その形成は、多細胞生物の小胞体（E R）の中に存在するいくつかの酵素によって触媒される。一方、ジスルフィド結合は、一般に、非ストレス条件下では、細胞質ゾルには認められない。これは、遊離システインを酸化から保護する、グルタチオン還元酵素およびチオレドキシン還元酵素などの還元系の存在による。例えば、リボヌクレオチド還元酵素は、その反応サイクルにおいてジスルフィド結合を形成するが、このジスルフィド結合の還元は、反応の進行にとって必須である（P r i n z , J B i o l C h e m . 2 7 2 (2 5) : 1 5 6 6 1) 。

30

【0266】

天然のミクロタンパクは、細菌、動物（いそぎんちゃく、腹足類、昆虫、さそり、蛇）、および植物によって発現される。しかしながら、組み換えミクロタンパクの異種発現は、一般に、大腸菌において行われている。ただし、枯草菌（B a c i l l u s s u b t i l i s）、酵母（S a c c h a r o m y c e s , K l u y v e r o m y c e s , P i c c h i a）、および糸状真菌、例えば、A s p e r g i l l u s および F u s a r i u m の外、哺乳類細胞系統、例えば、C H O、C O S、または P e r C 6 も、ミクロタンパクの発現に使用することが可能である。文献の例では、異種的に発現されるミクロタンパクは、通常、大腸菌の細胞原形質において生産される。

40

【0267】

組み換え発現に対する代替法は化学的合成である。ミクロタンパクは、化学的合成を可能とするほど十分に小さく、経済的に十分見合うコストで合成によって生産することが可能と考えられる。

【0268】

ジスルフィドを含む無関係産物（多くは、A b 断片、および丸ごとの A b を含む I g -ドメイン含有産物）は、一般に、周辺ゾルまたは培養液に対する分泌によって、哺乳類組織培養、または大腸菌において生産される。分泌産物は、タンパク分解的に取り除かれるシグナルペプチドを持ち、フォルミル化されない N-末端を残す。一方、大腸菌の細胞原形質において生産されるタンパクは、f M e t に続くアミノ酸（単複）に依存して、N-

50

末端フォルミルメチオニンを保持することが多い。文献は、f M e t に続くどのアミノ酸が、f M e t 除去を実行するかを記載する。

【0269】

ミクロタンパクは、細菌および古細菌ではほぼ完全に見られないが（若干の例外）、親水性ミクロタンパクは全て大腸菌内で簡単に作製することが可能である。

【0270】

少数の細菌性ミクロタンパクとして、例えば、大腸菌および関連腸内細菌由来の熱安定性エンテロトキシン（S T - I a および S T - I b と呼ばれる）がある。熱安定性エンテロトキシン、例えば、S T a （P F A M 0 2 0 4 8）および S T b は、配列レベルでは無関係である。S t - I a の配列整列は、7 2 a a の前駆体を示す。このタンパクは、二つの独立したタンパク分解事象によって処理されて、熟成トキシンを生成する。このトキシンは、1 4 2 5 3 6 の空間配置を持つ三つのジスルフィド結合を含む。S T - I a のモチーフは、C x x x x x x x x x x x x x x x x x x x x x x C C x x C C x x x C x x C である。

【0271】

ミクロタンパクを発現し、ミクロタンパクを培養液に分泌するための有望な方法は、S T - I a プロモーター、およびリーダーペプチドおよび前駆体で、ただし異なるミクロタンパクに接続されるものを用い、現在の 3 S S 1 4 2 5 3 6 モジュールを、異なるミクロタンパクで置換することである。S T - I a は、培養液に（周辺ゾルではなく）分泌されるが、これは、大腸菌ではきわめて稀であり、ジスルフィドがどのように形成されるかを説明する。S T - I a は、特別のリーダーペプチドを持っていて、そのため、S T - I a は、3 または 4 通りの異なる特別分泌系の内の一つを介して、大腸菌から分泌されることが可能となるらしい。ミクロタンパクと接続されるこのリーダーペプチドは、効率的分泌、および他のミクロタンパクのジスルフィド結合形成を可能とするばかりでなく、培養上清の速やかなスクリーニングにも有用であるようである。

【0272】

ミクロタンパクは、前核細胞および真核細胞を含む、種々の発現システムにおいて生産することが可能である。好適な発現宿主としては、例えば、酵母、真菌、哺乳類細胞培養体、昆虫細胞がある。特に興味深いのは、大腸菌、桿菌、または他の宿主生物を用いる細菌発現システムである。ミクロタンパクの異種発現は、通常、大腸菌の細胞原形質において行われる。ジスルフィド結合は、一般に、細胞原形質では、還元的環境であるために形成されないが、細胞が分解された後には形成される。ミクロタンパクの特徴解明および精製は、タンパク発現後に細胞を加熱することによって促進される。このプロセスによって、細胞は分解され、多くの大腸菌タンパクの沈殿が起こる。（S i l v e r m a n , J . , ら、（2005）N a t B i o t e c h n o l ）。大腸菌における種々のミクロタンパクの発現レベルは、該ミクロタンパクを、G F P のようなリポーター、または H R P 、ベータ・ラクタマーゼ、またはアルカリフォスファターゼのような酵素と融合させた場合、コロニースクリーニングによって互いに比較することが可能である。特に興味深いのは、熱およびプロテアーゼ安定酵素である。なぜなら、これらは、熱またはプロテアーゼストレスの条件下で、ミクロタンパクの安定性の定量を可能とするからである。例としては、子牛腸内アルカリフォスファターゼ、またはベータ・ラクタマーゼの熱安定性変異種がある（A m i n , N . , ら、（2004）P r o t e i n E n g D e s S e l , 1 7 : 7 8 7 - 9 3 ）。酵素またはリポーターに対するミクロタンパクの融合も、ミクロタンパクの結合性の分析をやり易くする。リポーター酵素の存在によって標的結合ミクロタンパクを検出することが可能となるからである。ミクロタンパクは、1 種以上のエピトープタグとの融合物として発現させることも可能である。例として、H A - タグ、H i s - タグ、m y c - タグ、s t r e p - タグ、E - タグ、T 7 - タグがある。このようなタグは、サンプルの精製をやり易くするが、サンドイッチ E L I S A 、またはその他の方法によって結合性を測定するのに使用することも可能である。タンパク、またはペプチドリガンドの結合性を検出するために、他にも多くのアッセイが記載されているが、これらの方

法もミクロタンパクに適用することが可能である。例としては、表面プラズモン共鳴、シンチレーション近接アッセイ、ELISA、アフラスクリーン（Perkin Elmer）、ベータガラクトシダーゼ酵素断片相補性アッセイ（CEDIA）がある。

【0273】

ミクロタンパクの異種発現は、通常、大腸菌の細胞原形質において行われる。ジスルフィド結合は、一般に、細胞原形質では、還元的環境であるために形成されないが、細胞が分解された後には形成される。大腸菌における種々のミクロタンパクの発現レベルは、該ミクロタンパクを、GFPのようなりポーター、またはHRP、またはアルカリフォスファターゼ（好ましくは、熱安定型、例えば、子牛腸内アルカリフォスファターゼ）のような酵素と融合させた場合、コロニースクリーニングによって互いに比較することが可能である。

10

【0274】

本発明はさらに、本明細書に開示されるシステイン含有スカフォールド、およびその断片を含む融合タンパクを含む。この融合は、本発明の二つ以上のスカフォールドと、関連または無関係スカフォールドとの間で行われてもよい。有用な融合パートナーとしては、ポリペプチドの細胞内局在を促進する配列、血清半減期反応性を延長する配列、または、免疫アッセイ支持体、またはワクチン担体に対するペプチドの結合を促進する配列である。

【0275】

（ジスルフィド結合の安定性における変動）

20

一般に、タンパクのジスルフィド結合の安定性にはある程度の変動がある。例えば、分泌タンパクにおけるジスルフィド結合は、細胞ゾルタンパクの「不要の」ジスルフィド結合よりも安定である傾向を持つ。一般に、ジスルフィド結合は、埋め込まれると還元に対して抵抗性を持ち、Wedemeyerらによれば、ジスルフィド結合は、一般に、埋め込まれる。したがって、分泌タンパクにおけるジスルフィド結合は、完全にフォールド形成すると、還元に対してむしろ抵抗性を持ち、ジスルフィド結合に接触可能とするように、局所的巻き戻しを誘発するためには、低濃度の変性剤を加えなければならない。

【0276】

複数のジスルフィド結合を持つタンパクが、フォールド形成状態で細胞ゾルに向けて輸送され、摂取の間タンパクがフォールド形成状態を続けている場合、そのジスルフィド結合は、還元に対し耐性を持つ。このための前提条件は、どのジスルフィド結合も還元剤に接触されてはならないということである。細胞ゾル内では、チオレドキシンおよびグルタチオンが、ジスルフィド結合の直接の酸化剤として働く。DTTと比べてそれらの分子量は比較的大きいため、フォールド状態のタンパクに埋め込まれたジスルフィド結合に対するアクセスはごく制限される。

30

【0277】

タンパクのジスルフィド結合の接触可能度は、結晶構造を用いてインシリコで定量するか、または、実験的にNMRによって定量することが可能であり、変性感度の定量と比較することが可能である（すなわち、D50は、野生型のジスルフィドの50%が存在するが、50%が存在しなくなる、還元剤の濃度である）。

40

標的に対する共有結合

あるタンパクは、ジスルフィド結合を交換することによって別のタンパクに共有的に結合し、例外的結合アフィニティーを実現することが可能である。一つの有用な例はミニコラーゲンであり、この場合、c-末端尾部配列が、N-末端頭部配列に共有的に結合し、二つのタンパクの間に6ジスルフィドの形成を実現する。図113を参照されたい。

【0278】

（スクリーニングおよび特徴解明ツール）

前述の234、3x0-8、4x0-8、および4x6法の早期サイクルにおいて得られるタンパクライブラリーおよび個別のタンパククローンは、異種的にフォールド形成する傾向を持つ。

50

【0279】

ある程度まで、この異種性を無視することは可能であるが、所望の性質、特に、高いアフィニティー（通常ピコモル）および高い特異性の外、均一フォールド形成および高い発現レベルが得られるまで、指向性進化によって、タンパクを進化し続ける。

ファージライブラリーを構築し、選択する方法

（ディスプレイのタイプ）

変異種から成る大きなライブラリーにおいて結合分子を特定することを可能とする様々な方法がこれまでに記載されている。一つの方法は化学的合成である。ビーズにおいてライブラリーメンバーを、各ビーズが異なるペプチド配列を担持するように合成する。所望の特異性を持つリガンドを担持するビーズを、標的された結合パートナーを用いることによって特定することが可能である。もう一つの方法は、繰り返し手順において特異的結合配列の特定を可能とする、ペプチドのサブライブラリーの作製である（Pinilla, C., ら、(1992) *BioTechnique*, 13:901-905）。さらに一般的に使用されるのは、変異種のライブラリーが、ファージ、タンパク、または細胞表面において発現されるディスプレイ法である。この方法では、共通の特性として、ライブラリーの各変異種をコードするDNAまたはRNAが、物理的にリガンドに連結される。これによって、対象とするリガンドを検出または回収することが可能となり、かつ、付属のDNAまたはRNAの配列を決定することによってそのペプチド配列を決めることが可能となる。ディスプレイ法によって、当業者は、ランダムな配列から成る大きなライブラリーから、所望の結合性を有するライブラリーメンバーを濃縮することができる。多くの場合、所望の結合性を有する変異種は、濃縮ライブラリーの個別の分離体について所望の性質の有無に関してスクリーニングすることによって特定することが可能である。ディスプレイ法の例として、lacレプレッサーに対する融合（Cull, M., ら、(1992) *Proc. Natl. Acad. Sci. USA*, 89:1865-1869）、細胞表面ディスプレイ（Witttrup, K. D. (2001) *Curr Opin Biotech*, 12:395-9）がある。特に興味深いのは、ランダムペプチドまたはタンパクが、ファージ粒子に連結される方法である。一般的に使用されるのは、M13ファージ（Smith, G. P., ら、(1997) *Chem. Rev*, 97:391-410）およびT7ファージ（Danner, S., ら、(2001) *Proc Natl Acad Sci USA*, 98:12954-9）である。M13ファージにおいてペプチドまたはタンパクを提示するためには複数の方法の利用が可能である。多くの場合、ライブラリー配列は、M13ファージのペプチドpIIIのN-末端に融合される。ファージは、通常、このタンパクの3-5コピーを担持するので、このライブラリーのファージは、多くの場合、ライブラリーメンバーの3-5コピーを担持する。この方法は、多価性ディスプレイと呼ばれる。別法は、ファージミドディスプレイである。この方法では、ライブラリーはファージミドにコードされる。ファージ粒子は、ファージミドを担持する細胞を、ヘルパーファージによって感染させることによって形成することが可能である。（Lowman, H. B., ら、(1991) *Biochemistry*, 30:10832-10838）。このプロセスは、通常、一価性ディスプレイをもたらす。ある場合には、一価性ディスプレイは、高アフィニティー結合因子をもたらす点で好まれる。別の場合では、多価性ディスプレイが好まれる（O'Connell, D., ら、(2002) *J Mol Biol*, 321:49-56）。

【0280】

所望の特徴を持つ配列を、ファージディスプレイによって濃縮するための種々の方法が記載されている。対象標的は、免疫試験管、マイクロタイタープレート、磁気ビーズ、または他の表面に結合させることによって固定することが可能である。次いで、ファージライブラリーを、固定標的に接触させ、結合リガンドを欠如するファージを洗い流し、標的的特異的リガンドを担持するファージを、種々の条件によって溶出することが可能である。溶出は、低pH、高pH、尿素、または、タンパク-タンパク接触を破断する傾向を持つ他の条件によって実行することが可能である。さらに、結合ファージに、大腸菌細胞を、

溶出ファージが直接その添加大腸菌宿主に感染できるようなやり方で添加することによって溶出させることが可能である。興味深いプロトコルは、ファージ結合リガンド、または固定標的を分解することが可能なプロテアーゼによる溶出である。プロテアーゼはまた、プロテアーゼ耐性ファージ結合リガンドを濃縮するためのツールとして利用することが可能である。例えば、対象標的によるパンニングの前に、1種以上の（ヒトまたはマウス）プロテアーゼと共に、ファージ結合リガンドのライブラリーをインキュベートしてもよい。このプロセスは、前記ライブラリーのプロテアーゼ感受性リガンドを分解し、除去する（Kristensen, P., ら、(1998) Fold Des, 3:321-8）。リガンドのファージディスプレイ・ライブラリーは、複雑な生物サンプルに対する結合のために濃縮することも可能である。例は、固定された細胞膜分画（Tur, M. K., ら、(2003) Int J Mol Med, 11:523-7）、または丸ごとの細胞（Rasmussen, U. B., ら、(2002) Cancer Gene Ther, 9:606-12; Kelly, K. A., ら、(2003) Neoplasia, 5:437-44）によるパンニングである。ある場合、ファージライブラリーの細胞特異的結合因子の濃縮度を上げるためにパンニング条件を最適化しなければならない（Watters, J. M., ら、(1997) Immunotechnology, 3:21-9）。ファージパンニングはまた、生きている患者または動物において実行することも可能である。この方法は、血管標的に結合するリガンドの特定のために特に興味深い（Arap, W., ら、(2002) Nat Med, 8:121-7）。

【0281】

ライブラリーを構築するためのクローニング法

文献は、当業者に対し、ペプチドリガンドのライブラリーをコードするDNA配列のライブラリーを作製することを可能とする、広く様々の方法を記載する。ヌクレオチドのランダム混合物を用い、一つ、または複数のランダム位置を含むオリゴヌクレオチドを合成してもよい。このプロセスによって、ランダム位置の数ばかりでなく、ランダム化の程度を調節することが可能である。さらに、生物サンプル由来のDNAを部分的に消化することによってランダムな、または半ランダムなDNA配列を得ることが可能である。ランダムオリゴヌクレオチドは、あらかじめ定められた位置においてランダム化されたプラスミドまたはファージのライブラリーの構築に使用することが可能である。これは、（de Kruijf, J., ら、(1995) J Mol Biol, 248:97-105）に記載されるように、PCR融合によって実行される。他のプロトコルは、DNA連結に基づく（Felicci, F., ら、(1991) J Mol Biol, 222:301-10; Kay, B. K., ら、(1993) Gene, 128:59-65）。もう一つの一般的に使用される方法は、Kunkelの突然変異発生である。この方法では、プラスミドまたはファージミドの突然変異1本鎖が、鋳型として、1本鎖サイクリックDNAによって合成される。Sidhu, S. S., ら、(2000) Methods Enzymol, 328:333-63; Kunkel, T. A., ら、(1987) Methods Enzymol, 154:367-82を参照されたい。

【0282】

Kunkel突然変異発生法は、CJ236のような大腸菌株から得られた、ランダムに組み込まれたウラシル塩基を含む鋳型を使用する。このウラシル含有鋳型鎖は、大腸菌において形質転換されると分解されることが好ましい。一方、インビトロで合成された突然変異発生鎖は保持される。その結果、多くの形質転換細胞は、ファージミドまたはファージの突然変異発生バージョンを担持する。ライブラリーの多様性を増すための貴重な方法は、複数のサブライブラリーを合わせることである。これらのサブライブラリーは、前述の方法の内の任意の方法によって作製されてよく、同じ、または異なるスカフォールドに基づいていてもよい。

【0283】

短いペプチドから成る大きなファージライブラリーを作製するための有用な方法が最近記載された（Schollee, M. D., ら、(2005) Comb Chem Hig

h Throughput Screen, 8:545-51)。この方法は、Kunkel法に関連するが、ランダムウラシル塩基を含む、単一鎖のDNA鋳型の作製を要しない。その代わり、この方法は、突然変異を誘発すべき区域の近くに一つ以上の突然変異を担持する鋳型ファージから始め、かつ、前記突然変異は、ファージを非感染性とする。方法は、いくつかの位置にランダムコドンに担持し、かつ、鋳型におけるファージ不活性化突然変異を補正する突然変異発生オリゴヌクレオチドを用いる。その結果、形質転換後、突然変異発生ファージ粒子のみが感染性を持ち、そのライブラリーには親ファージはほとんど含まれなくなる。この方法は、いくつかのやり方でさらに修飾することが可能である。例えば、複数の突然変異発生オリゴヌクレオチドを用い、ファージの複数の、不連続領域において同時に突然変異を発生させてもよい。我々は、これを、<10、15、または30アミノ酸ではなく、さらに新たな目標として、>25、30、35、40、45、50、55、および60アミノ酸から成るミクロタンパク全体に適用することによって、この方法をさらに1ステップ推し進めた。この方法は、現在では、1回の形質転換で、10e10を超える（最大10e11）のライブラリーを作製し、このため、10回の形質転換では、10e12の多様性を持つ単一ライブラリーが期待される。

【0284】

（反復突然変異発生法）

Scholle法の新規変法は、突然変異発生オリゴヌクレオチドを、鋳型のアンバー終止コドンが、オークル終止コドンに変換され、突然変異発生の次回サイクルにおいてオークルがアンバーに変換されるように、設計することである。この場合、鋳型ファージおよび突然変異発生ライブラリーメンバーは、大腸菌の異なるサプレッサー株において、オークルサプレッサーとアンバーサプレッサーと交互に、培養されなければならない。このように、2種類の終止コドンおよび二つのサプレッサー株の間を交互に行き来することによって、ファージの突然変異発生の連続ラウンドを実行することが可能である。

【0285】

Schoelle法の、もう一つの新規変法は、単一鎖のファージDNA鋳型と共にメガプライマーを使用することを含む。メガプライマーは、パンニングの前回ラウンドから選ばれたファージプールのライブラリー挿入体から作製された長いssDNAである。目的は、一つ以上の区域で突然変異を発生させた、前回プールのライブラリー挿入体の完全な多様性を捕捉し、これを、新規ライブラリーに、さらに別の区域において突然変異発生が可能となるように転送することである。このメガプライマープロセスは、対象遺伝子の中に終止コドンを含む、同じ鋳型を用いて複数サイクル反復することが可能である。メガプライマーは、下記を含むssDNA（任意にPCRによって作製される）である：1）前記ssDNA鋳型に対し相補性を持つ少なくとも15塩基の5'および3'重複区域、および、2）以前に選ばれたクローンのプールからコピーされた（任意にPCRによって）、一つ以上の、以前に選ばれたライブラリー区域（1、2、3、4、またはそれ以上）、および、3）パンニングの次回ラウンドにおいて選ばれるはずの新規突然変異発生区域。メガプライマーは、1）新たに合成されるライブラリー区域をコードする一つ以上のオリゴヌクレオチドを合成すること；および、2）これを、任意に重複PCRを用いて、以前に最適化された、他の、任意のライブラリー区域を含むDNA断片（任意にPCRによって得られた）に融合すること、によって任意に調製される。組み合わせた（重複）PCR産物のランオフまたは1本鎖PCRを用い、以前に最適化された区域全てを含むばかりでなく、次のパンニング実験において最適化される予定の別区域の新規ライブラリーを含む1本鎖メガプライマーを作製する。図28を参照されたい。この方法は、サイクル当たり10e11から10e12の多様性を生成する、パンニングを伴う、速やかな、複数のライブラリー創製サイクルによるタンパクのアフィニティー熟成を可能とすることが期待される。

【0286】

ミクロタンパクの結合性または、製造、安定性、または免疫原性のようなその他の性質を強化する目的で、（前回選択された、または未洗練の）ミクロタンパク・ライブラリー

に配列多様性を導入するため、または、個々のマイクロタンパク・クローンを突然変異させるために、様々の方法の適用が可能である。原理的に、ライブラリーの作製に使用することが可能な方法は全て、マイクロタンパクの濃縮（以前に選択された）ライブラリーに多様性を導入するために使用することが可能である。特に、所望の結合またはその他の性質を有する変異種を合成し、これらの配列に基づいて部分的にランダム化されたオリゴヌクレオチドを設計することが可能である。このプロセスによって、ランダム化の位置、およびランダム化の程度を調節することが可能である。各種コンピュータアルゴリズムを用いて、複数の変異種の配列データから、タンパク中の個別の突然変異の有用性を導き出すことが可能である（Jonsson, J., ら、(1993) *Nucleic Acids Res*, 21:733-9; Amin, N., ら、(2004) *Protein Eng Des Sel*, 17:787-93）。濃縮ライブラリーの再度の突然変異発生のために特に興味深いのは、DNAシャッフリングである（Stemmer, W. P. C. (1994) *Nature*, 370:389-391）。これは、濃縮ライブラリーにおいて個々の配列の組み換えを生成する。シャッフリングは、様々の改変PCR条件を用いて実行が可能であるが、組み換えを強調するために鋳型は部分的に分解してもよい。それに代わるものとして、制限酵素クローニング用い、あらかじめ定められた位置に組み換えを実行する方法がある。特に興味深いのは、DNAを、その配列認識部位の外側で切断する、IIS型制限酵素を利用する方法である（Collins, J., ら、(2001) *J Biotechnol*, 74:317-38）。非パリンδροーム・オーバーハングを生成する制限酵素を利用して、複数位置において変異種混合物をコードするプラスミド、またはその他のDNAを切断し、連結によって完全プラスミドを再集合することが可能である（Berger, S. L., ら、(1993) *Anal Biochem*, 214:571-9）。多様性を導入するもう一つの方法は、PCR突然変異発生法である。この方法では、ライブラリーメンバーをコードするDNA配列に対し、突然変異発生条件下でPCRを行う。比較的高い突然変異頻度で突然変異を起こすPCR条件は記載されている（Leung, D., ら、(1989) *Technique*, 1:11-15）。さらに、忠実性の低下したポリメラーゼを採用することも可能である（Vanhercke, T., ら、(2005) *Anal Biochem*, 339:9-14）。特に興味深い方法は、ミューテーター株による（Irving, R. A., ら、(1996) *Immunotechnology*, 2:127-43; Coia, G., ら、(1997) *Gene*, 201:203-9）。これらは、一つ以上のDNA修復遺伝子に欠陥を持つ株である。これらの株におけるプラスミド、またはファージ、または他のDNAは、正常の複製の間に突然変異を蓄積する。ミューテーター株中の、個々のクローン、または濃縮集団を伝播して、遺伝子の多様性を導入することが可能である。前述の方法の多くは、反復プロセスに利用することが可能である。複数ラウンドの突然変異発生および、スクリーニングまたはパンニングを、遺伝子全体、または遺伝子の一部に適用することが可能であり、あるいは、各後続ラウンド毎に、タンパクの様々な部分に突然変異を起こさせることが可能である（Yang, W. P., ら、(1995) *J Mol Biol*, 254:392-403）。

【0287】

（ライブラリー処理）

ファージパンニングの既知のアーチファクトとして、1) 疎水性に基づく非特異的結合、および2) 標的に対する多価性結合で、a) pIIIファージタンパクの5価性によるもの、または、b) 異なるマイクロタンパク間のジスルフィド形成によるマルチマー形成によるもの、または、c) 固相支持体における標的の高密度コーティングによるもの、および3) 背景依存性標的結合で、標的の背景、またはマイクロタンパクの背景が、結合または阻害活性を決定しているもの、が挙げられる。これら問題点の大きさを最小化するために、種々の処理工程の採用が可能である。理想的には、この処理は、ライブラリー全体に適用されるが（ライブラリー処理）、不正クローンを除去するいくつかの有用な処理は、可溶性タンパクプール、または個々の可溶性タンパクにのみ適用される。

【0288】

ミクロタンパク・ライブラリーは、他のタンパクに架橋結合して指向性進化を複雑化する可能性のある遊離チオールを含む確率が高い。一つの方法は、ライブラリーを遊離チオールカラムを通過させ、一つ以上の遊離スルフヒドリルを有する全てのクローンを取り除くことによって、該ライブラリーから、最悪クローンを除去することである。遊離SH基を有するクローンはさらに、ビオチン-SH試薬と反応することが可能であり、ストレプトアビジン・カラムによる、反応性SH基を有するクローンの効率的除去を可能とする。もう一つの方法は、遊離チオールを除去せず、それらを、スルフヒドリル反応性薬品、例えば、ヨード酢酸によってキャップ封印して不活性化することである。特に興味深いのは、非特異的標的結合を抑える、嵩張った、または親水性のスルフヒドリル試薬、またはその修飾変種である。

10

【0289】

背景依存性の例は、全ての定常配列、例えば、pI I Iタンパク、リンカー、ペプチドタグ、ビオチン-ストレプトアビジン、Fc、および、相互作用に寄与する、その他の融合タンパクを含む配列である。背景依存を回避するための典型的な方法は、蓄積を避けるために、できるだけ頻繁に背景を切り替えることを含む。これは、異なるディスプレイシステム（すなわち、M13・対・T7、またはM13・対・酵母）の間、使用されるタグ同士およびリンカー同士の間、固定のために使用される（固相）支持体（すなわち、固定化学）、および、標的タンパクそのもの（異なる販売業者、異なる融合バージョン）を交互に行き来することを含む。

20

【0290】

さらにライブラリー処理を用いて、好ましい性質を持つタンパクを選択することが可能である。一つの選択肢は、ライブラリーから不安定変異種を除くために、ライブラリーをプロテアーゼで処理することである。使用されるプロテアーゼは、通常、用途の中で遭遇することが予想されるものである。肺輸送のためには、例えば、肺洗浄によって得られる肺プロテアーゼを用いることが考えられる。同様に、血清、唾液、胃、小腸、皮膚、鼻などから、プロテアーゼの混合物を得られよう。しかしながら、単一の精製プロテアーゼの混合物を使用することも可能である。プロテアーゼの広範なリストが付録Eに示される。

【0291】

例えば、ライブラリーを、低濃度の還元剤（すなわち、DTT、またはベータメルカプトエタノール）に暴露して、先ず、安定性のもっとも低い構造を除去し、それから次第に濃度を増すことによって、もっとも安定な構造、すなわち、最強のジスルフィド結合を有する構造に関してライブラリーを選択することが可能である。通常、還元剤（すなわち、DTT、BMEなど）の濃度は、所望の安定性に応じて、2.5 mMから5 mM、10 mM、20 mM、30 mM、40 mM、50 mM、60 mM、70 mM、80 mM、90 mM、または場合によっては100 mMと考えられる。

30

【0292】

さらに、全体ディスプレイライブラリーを、高濃度の還元剤で還元し、次いで、このタンパクライブラリーを徐々に再び酸化して再度ジスルフィドを形成し、次いで、前述のように遊離SH基を持つクローンを除去することによって、インビトロにおいて効率的に再度フォールド形成することが可能なクローンを選択することが可能である。このプロセスは、インビトロにおいて再度フォールド形成する効率が低いクローンを除去するために、1回または複数回適用することが可能である。

40

【0293】

一つの方法は、A. C. Fisher ら、(2006) Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Science (オンライン)によって記載されるように、タンパクの発現レベル、フォールド形成、および可溶性に関して遺伝子選択を適用することである。ディ

50

スプレイライブラリーのパンニングの後では（任意）、標的結合、発現レベル、およびフォールド形成に関して、タンパクレベルで、数千のクローンをスクリーニングするのは避けたいであろう。別法として、選ばれた挿入体の全プールを、ベータラクタマーゼ融合ベクターにおいてクローンすることがある。このクローンを、ベータラクタムにプレートしたところ、発現性の高い、ジスルフィド結合の完全な、かつ可溶なタンパクについて選択的であることを作者らは証明した。

【0294】

M13ファージによるタンパクライブラリーのディスプレイ、および、1サイクル以上の標的に対するパンニングの後、その後の進行には様々の道がある。

【0295】

ファージELISAによる個々のファージクローンのスクリーニング。これは、固定標的に結合する、ファージ粒子の数（抗M13抗体を用いて）を測定する。

【0296】

M13からT7ファージディスプレイ・ライブラリーへの転送。単一ライブラリー方式はどれも、標的に対し高アビディティ接触を形成することが可能なクローンを優先的に取る傾向がある。これが、可溶性タンパクのスクリーニングが重要である理由である。ただし、これは手間ひまのかかる解決策ではある。T7ファージディスプレイにおいて実現される多価性は、M13ディスプレイにおいて実現されるものとはきわめて異なるようであり、T7およびM13の間で往復サイクルすることは、原子価に基づく擬似陽性の出現を抑えるための優れた方法と考えられる。

【0297】

フィルターリフト。大きな寒天プレートにおいて高密度（ $10^2 - 10^5$ ）で育成した細菌コロニーについてはフィルターリフトを実行することが可能である。少量の、いくつかのタンパクは、培養液に分泌され、フィルター膜（ニトロセルロース、またはナイロン）に結合して終わる。次に、フィルターは、非脂肪ミルク、1%カゼイン加水分解物、または1%BSA液においてブロックされ、蛍光染料、またはインディケーター酵素（直接、または、間接的に、抗体またはビオチンーストレプトアビジンを介して）によって標識された標的タンパクとインキュベートされる。コロニーの位置は、フィルターを、プレートの背面に重ねることによって決定され、全ての陽性コロニーが選択され、さらなる特徴解明のために使用される。フィルターリフトの利点は、様々の期間洗浄した後シグナルを読み取ることによって、アフィニティー選択的とすることが可能であることである。高アフィニティークローンのシグナルは、ゆっくりと褪色するが、低アフィニティークローンのシグナルは、速やかに褪色する。このアフィニティー特徴の解明は、通常、ウェル使用アッセイでは3点アッセイを必要とするが、ウェル使用アッセイよりも、クローンからクローンの近似性はより優れる。コロニーを格子で区切ってアレイとするやり方は、コロニー・サイズまたは位置による差を最小化するので有用である。

【0298】

（製薬組成物）

本発明はさらに、主題のシステイン含有タンパクを含む製薬組成物を提供する。この組成物は、経口的に、鼻腔内に、非経口的に、または吸引治療によって投与されてもよく、錠剤、ローゼンジ、顆粒、カプセル、丸剤、アンプル、坐剤として、またはエロゾル形として摂取されてもよい。組成物はまた、水性、または非水性希釈剤に溶解した活性成分の懸濁液、溶液、または乳液、シロップ、顆粒、または粉末の形で摂取されてもよい。さらに、この製薬組成物は、さらに、他の製薬学的に活性な化合物、または、複数の、本発明の化合物を含んでもよい。

【0299】

本発明のシステイン含有タンパクはまた、各種液相担体、例えば、滅菌、または水溶液、製薬学的に受容可能な担体、懸濁液、および乳液と混ぜ合わせられてもよい。非水性溶媒の例としては、プロピルエチレングリコール、ポリエチレングリコール、および植物油が挙げられる。

【0300】

さらに詳細には、本発明の製薬組成物は、任意の適切なルート、例えば、経口、直腸内、鼻腔内、局所（経皮、エロゾル、頬内、および舌下）、腔内、非経口（皮下、筋肉内、静脈内、および皮内）、および肺内を含むルートを通じて投与してよい。さらに、好ましいルートは、レシピエントの状態および年齢、および治療される病気に応じて変動することが理解される。

【0301】

（製品形式）

多様な用途、例えば、試薬、診断薬、予防薬、体外治療剤、および、静脈内、皮下、硬膜下腔内、眼内、経強膜、腹腔内、経皮、経口、頬内、腸内、腔内、鼻腔内、肺内、および他の剤形などのインビボ治療薬のための、種々の薬剤輸送法用の特殊形式を含む用途に使用されるために、広範な製品形式が考えられている（図159を参照）。

【0302】

このような製品形式としては、ドメイン・モノマー、およびドメインマルチマー（単一または複数タンパク鎖の中に、2、3、4、5、6、7、8、9、10、15、20、30、40、50、場合によっては100ドメインを有する製品）が挙げられる。ドメインは、固有配列または構造モチーフしか含んでなくともよいし、あるいは、二重配列または構造モチーフ、またはより高次の反復配列または構造モチーフ（反復タンパク列）を含んでもよい。各ドメインは、1、2、3、4、5、6、7、8、9、または10種の異なる標的に対する、単一連続的、または複数不連続的（空間的、または配列定義による）結合部位を持ってもよい。標的は、治療、診断（インビボ、インビトロ）、試薬、または材料標的であってもよく、かつ、タンパク、炭水化物、脂質、金属、または、他の、任意の、生物的、または非生物的材料（その組み合わせ）であってもよい。ドメインモノマーおよびマルチマーは、同じ標的に対し複数の結合部位を持ち、任意にアビディティをもたらしてもよい。ドメインマルチマーはまた、異なる標的に対し、1、2、3、4、5、6、7、8、またはそれ以上の結合部位を持ち、多重特異性をもたらしてもよい。ドメインマルチマーは、任意に、長さが、1、2、3、4、5、6、7、8、9、10、12、14、16、18、20、25、30 AAの範囲のペプチドリンカーを含む。このドメインには、様々な要素、例えば、タグ（例えば、抗体またはN i -NTAによる検出または精製のための）を含む、直鎖または環状ペプチドが融合されてもよい。

【0303】

半減期延長形式：製品の血清排泄半減期を所望の分泌半減期、1、2、4、8、または16時間から、1、2、3、4、5、または6日から、1週、2週、3週、または1、2、3ヶ月の範囲に延長するために好ましい方法は、融合ペプチド（直鎖、単環、または二環、それぞれ、0、1、または2ジスルフィドを意味する）またはタンパクドメインで、血清アルブミン、免疫グロブリン（すなわち、IgG）、赤血球、または他の血液分子または、血清接触分子への結合を実現するものを使用することが好ましい。それに代わる別法は、部分的に重複する、または重複しない種々の結合部位を用いて、ドメインが、製薬学的標的に対してだけでなく、半減期延長標的、例えば、血清アルブミンにも結合するようにドメインを設計することである。望ましい方法は、一区域においてランダム化され、かつ、半減期標的（すなわちHSA）に結合するように選ばれるスカフォールドを創出し、次に、これらの構築体を用いて、一つ以上の製薬学的標的に対して結合するように設計される、さらに別の区域をランダム化し、このようにして、半減期標的にも、製薬学的標的にも結合するドメインを創製することである。血清タンパク、または血清接触タンパクに対して結合することによって半減期延長を実現するドメインは、非ミクロタンパク、例えば、ヒトのサイトカイン、増殖因子、およびケモカインに融合させてもよい。任意の応用として、このようなヒトタンパクの半減期の延長、または、ヒトタンパクの特定組織に対する標的照準がある。このような相互作用において好ましいアフィニティーは、10 uM、1 uM、100 nM、10 nM、1 nM、0.1 nM未満（または上回るレベル）である。もう一つの選択肢は、ドメインのStokes流体力学半径を拡張し、それによっ

て、その血清分泌半減期を延長するために、ドメイン（単複）に、長い、非構造的で、屈曲性に富むグリシン富裕配列を融合させることである。もう一つの選択肢は、ペプチド結合ではなく、ジスルフィド結合、または他の化学的連結を通じて、ドメインを、他のドメインに共有的に連結することである。もう一つの選択肢は、小型分子（製薬学的に活性な薬剤担体を含む）、放射性標識（すなわち、キレート剤）、およびPEGまたはPEG様分子または炭水化物を、タンパクに化学的に接合することである。

【0304】

薬剤輸送の別形式：平均的マイクロタンパクの性質は、大抵の代替（非注入）輸送形式に好適であるが（サイズ、プロテアーゼ安定性、可溶性、親水性）、特定の好ましい輸送形式に対する可能性をさらに改善するために加工法の使用が考えられる。Werle, M. 10
ら、(2006) J. Drug Targeting 14:137-146は、3種の異なるマイクロタンパクが、プロテアーゼに対し、例えば、エラスターゼ、ペプシン、キモトリプシンに対してばかりでなく、血漿プロテアーゼ（血清）および腸膜プロテアーゼ（2/3）に対しても高い耐性を持つことを示す。彼らはまた、二つのマイクロタンパクの見かけの移動係数（Papp）が、各種ペプチドおよび小型タンパクについて創製された標準曲線から予想されるものよりも3倍高いことを示す。組織障壁を横断する輸送、例えば、鼻腔内、経皮、経口、頬内、腸内、または経強膜輸送では、効率およびバイオアベイラビリティは、主に、タンパクのサイズによって決められる。各種の賦形剤、例えば、アルキルサッカリドは、タンパク製剤の輸送を最大約10倍改善することが報告されている（Maggio, E. (2006) Drug Delivery Reports; Magg 20
gio, E. (2006) Expert Opinion in Drug Delivery 3:1-11）。これらの輸送強化剤の内のいくつかは、GRASであるか、または、食品添加剤として使用されており、そのため、製剤におけるその使用は、長々しいFDAの承認手続きを必要としない。これらの強化剤のいくつかは、両親媒性であり、親水性部分（すなわち、炭水化物）および疎水性部分（すなわち、アルキル鎖）を持つために、ミセルを形成することが可能である。マイクロタンパク、および非マイクロタンパクペプチドまたはタンパクに遺伝学的に融合された、親水性および疎水性タンパク配列を用いてこれを模倣することが可能である。例えば、親水性タンパクは、グリシン（非イオン性）、グルタミン酸塩、およびアスパラギン酸塩（負に荷電）、またはリシンおよびアルギニン（正に荷電）に富んでもよく、疎水性配列は、トリプトファンに富んでもよい。疎水性 30
突出尾部（すなわち、5-20トリプトファン残基）を持つタンパクを、膜挿入によって長い半減期を実現する疎水性薬剤と同様、そのポリ・トリプトファンを細胞膜に挿入させて半減期の延長を実現するように使用してもよい。タンパクそのものは不変であるので、その結合特異性は、低下しないことが予想され、ただ、その（マイクロの）生体分布が変化するだけである。それに代わる別法は、マイクロタンパクに、薬剤輸送体、例えば、PepT1、PepT2、HPT1、ABC輸送体に結合する、またはそのような輸送体によって取り込まれることが知られる、ペプチドまたは小型分子を接合することである。参考文献は、Lee, VHL (2001) Mucosal drug delivery. J. Natl Cancer Inst Monogr 29:41-44; および、Kun 40
ta JR and Sinko, PJ (2004) Intestinal drug transporters: in vivo function and clinical importance. Current Drug Metabolism 5:109-124; Nielsen, CU and Brodin, B (2003) Di-
/Tri-peptide transporters as drug delivery targets: Regulation of transport under physiological and pathological conditions. Current Drug Targets 4:373-388; Blanchette, J. 50
ら、(2004) Principles of transmucosal delivery of therapeutic agents, Biomedicine & Pharmacotherapy 58:142-152. Dietri

ch, CG ら、(2005); ABC of oral bioavailability: transporters as gatekeepers in the gut. Gut 52:1788-1795; Yang CY ら、(1999) Intestinal Peptide transport systems and oral availability. Pharmaceutical Research 16:1331-1343である。

【0305】

ミクロタンパクは、理想的には、半減期延長が必要とされないので、局所輸送に好適である。ミクロタンパクは、単回投与によって持続輸送が実現されるため、沈着処方を通じて輸送されてもよい。

【0306】

沈着処方（例えば、インプラント、ナノ小球、ミクロ小球、およびゲルなどの注入可能溶液）は、ある程度の半減期延長は有利ではあるけれども、薬剤（溶液形）の半減期の延長を要求しない。

【0307】

ミクロタンパク・ドメイン、および、様々のアミノ酸組成を持つポリペプチドスペーサを重合して、粘ちょうな、長いポリマーに変換することは、可溶性薬剤がそれからゆっくりと放出される沈着物を生成することが期待される。これらのポリマーは、ミクロタンパクに融合されてもよいし、あるいは、それらは別々のタンパクであってもよい。粘ちょう液体は、皮下、または筋肉下に注入されることも考えられる。タンパクポリマーを用いる代わりに、タンパクを、他の、様々の生物分解性基質、例えば、ポリ無水物、またはポリエステル、またはPLG（ポリ（D、L-ラクチド-ε-グリコリド））、またはSAILB（スクロースアセテートイソブチレート）、またはポリエチレングリコール（PEG）、および、その他のヒドロゲル、脂質発泡体、コラーゲン、およびヒアルロン酸などの基質と混合してもよい。小型のサイズ、プロテアーゼ、機械的、および熱に対する耐性、および高度の親水性のため、ミクロタンパクは、他のタンパクが実現できない過酷な処方にも好適である。小型のサイズのために、ミクロタンパクは、イオン導入、粉末弾丸輸送、可聴輸送、および電気穿孔輸送にも好適である（Clelland, JL ら、(2001) Emerging protein delivery methods. Current Opinion in Biotechnology 12:212-219）。

【0308】

融合タンパクの経口輸送：経口輸送に対する別の方法は、既存の細菌トキシン、例えば、Pseudomonasエキソトキシン（PE38、PE40）に対するミクロタンパク薬剤の融合を含む。このトキシンは、細胞膜を横断して、薬剤を細胞の原形質の中に輸送することが可能である。この方法は、タンパク薬剤を、細胞（すなわち、腫瘍細胞）内へ輸送するばかりでなく、効率的経口輸送、すなわち、小腸腔から血流への輸送にも有効であることが証明されている（Mrsny, RJ ら、(2002) Bacterial toxins as tools for mucosal vaccination. Drug Discovery Today 4:247-258）。

【0309】

経口（および肺内）輸送に対するもう一つの方法は、ミクロタンパクをFc受容体に融合し、小腸からの、新生児Fc受容体仲介摂取を用い、トランスサイトシスを通じて血流に輸送することである（Low, SC ら、(2005) Oral and pulmonary delivery of FSH-Fc fusion protein via neonatal Fc receptor-mediated transcytosis. Human Reproduction (印刷中)）。

【0310】

ミクロタンパクの細胞内輸送：Rothbardらは、天然のアルギニン富裕ペプチド、例えば、HIV-tatが、細胞膜を横切って輸送されること、および合成arg-富

10

20

30

40

50

裕ペプチドも同様であることを実証した。これを模倣する一つの方法は、ミクロタンパクのN-またはC-末端にarg-富裕ペプチドを付着させることであり、第2の方法は、ライブラリーの設計時、ミクロタンパクのアルギニン含量を増し、かつ、スクリーニング時、高いarg含量のクローンを優先的に取ることである。アルギニン含量は、最大約3%、好ましくは5%、多くの場合7.5%、時に10%、理想的には15、20、25、30、または35%に増してもよい。

【0311】

マルチマー形式：ミクロタンパクは、アビディティの上昇および半減期の延長を含む、多くの理由でマルチマーとしてもよい。我々は、複数のドメインが、グリシン富裕の、長い親水性スペーサで隔てられる形式に注目したが、ただし、複数のドメインを、スペーサ無しで、または天然のスペーサを用いて重合することも可能である。

10

【0312】

長いグリシン富裕配列は、大きな流体力学半径を持ち、そのため、PEG化による半減期延長を模倣する。各グリシン富裕配列スペーサは、20、25、30、35、40、50、60、70、80、100、120、140、160、180、200、240、280、320アミノ酸長、またはそれ以上であってもよい。ホモのマルチマー標的、および細胞表面標的では、さらにモノマー標的でも、ミクロタンパク結合部位をマルチマー化し、結合部位間に、さらに（任意に）、N-末端およびC-末端に、グリシン富裕スペーサを配することが有用である。このタンパクでは、タンパクにおけるグリシンポリマーの全長は、100、150、200、250、300、350、または場合によっては400アミノ酸に達してもよい。このタンパクは、それぞれが、同じ標的（同じコピー、または異なるコピー）の異なる部位に結合する、複数の、異なる結合部位を含んでもよい。このようにして、例えば、一部はその長さおよび半径のために、一部は、血清アルブミン、または免疫グロブリン、またはその他の、血清接触タンパクに対する（ミクロタンパクの）結合部位の存在のために、きわめて長い半減期を有するタンパクを創製することが可能である。

20

【0313】

抗体も、その長い半減期を獲得するために、サイズおよび受容体結合の両方を利用するが、最長の半減期のためには両方の機序が必要とされるようである。結合性および非結合性要素から成るこのようなポリマーを実現するためには、様々な方法および組成物がある。1) 単一タンパク鎖において結合された（遺伝学的融合）結合モチーフの複数コピー；コピーは同じであっても、異なっているもよい、2) 結合部位の単一（または複数）コピーが、別々のタンパクとして発現され、化学的結合によって、NからC末端にかけてマルチマー化される。各種化学的結合法の使用が可能である（結合剤のリストについてはwww.pierce.com参照）；コピーは同じであっても、異なっているもよい、3) 単一タンパク鎖における結合部位の複数コピーで、非結合リンカーによって分離される、4) 結合部位および非結合部位が、それぞれ、別々のタンパクとして発現され、化学的結合によってマルチマー化される。各種の化学的結合の使用が可能である（Pierceの結合剤リストを加えよ）；コピーは同じであっても、異なっているもよい、5) 各タンパクは、一つの結合部位、および一つの非結合部位を含み、これらのタンパクが、化学的結合によってマルチマー化される。各種化学的結合法の使用が可能である（www.pierce.com参照）；コピーは同じであっても、異なっているもよい、6) 各タンパクは、結合部位、および、任意に非結合リンカーを含み、各タンパクは、N-およびC-の両末端において、互いに結合してタンパクの、方向性を持つ直線的マルチマーを創製する「連結ペプチド」を有する。各種のペプチド配列、例えば、SKVILF(E)、またはRARADADARARADADA、およびその誘導体などの使用が可能であり、コピーは同じであっても、異なっているもよい。SKVILF(E)は、反対平行的に、ホモの二量体を形成し（Bodenmuller et al (1986) EMBO J.）、かつ、DADADA（または[DA]n）に結合するRARARA（または[RA]n）は、RARADADARARADADAから得られることが、Narmoneva, DA

30

40

50

ら、(2005) Self-assembling sort oligonucleotides and the promotion of angiogenesis. Biomaterials 26:4837-4846に報告されている。ドメイン、またはドメインのマルチマーの一端に[RA]n、他端(C-またはN-末端)に[DA]nを配置することは、一タンパクのN-末端を、同じタンパクの別コピーのC-末端に連結することによって、直線的で、方向性を持つポリマーの創製を可能とする。ポリマーをきわめて長く作製するか、または架橋させて、皮下の注入部位を効率的に離脱することがないようにした場合、沈着、または徐放処方を実現される。一つの方法は、血清プロテアーゼのためのプロテアーゼ切断部位を、ゆっくりと崩壊するポリマーを生成するように設計することである。

10

【0314】

製薬標的：主題のミクロタンパクは、一般に、ある任意の標的に対して特異的な結合特異性を示す。ある実施態様では、主題のミクロタンパクは、下記の、非限定的リストから選ばれる一つの標的に結合することが可能である。すなわち、VEGF、VEGF-R1、VEGF-R2、VEGF-R3、Her-1、Her-2、Her-3、EGF-1、EGF-2、EGF-3、アルファ3、cMet、ICOS、CD40L、LFA-1、c-Met、ICOS、LFA-1、IL-6、B7.1、B7.2、OX40、IL-1b、TACI、IgE、BAFFまたはBLys、TPO-R、CD19、CD20、CD22、CD33、CD28、IL-1-R1、TNF α 、TRAIL-R1、補体受容体1、FGFa、オステオポンチン、ビトロネクチン、エフリンA1-A5、エフリンB1-B3、アルファ-2-マクログロブリン、CCL1、CCL2、CCL3、CCL4、CCL5、CCL6、CCL7、CXCL8、CXCL9、CXCL10、CXCL11、CXCL12、CCL13、CCL14、CCL15、CXCL16、CCL16、CCL17、CCL18、CCL19、CCL20、CCL21、CCL22、PDGF、TGFb、GMCSF、SCF、p40(IL12/IL23)、IL1b、IL1a、IL1ra、IL2、IL3、IL4、IL5、IL6、IL8、IL10、IL12、IL15、Fas、FasL、Flt3リガンド、41BB、ACE、ACE-2、KGF、FGF-7、SCF、ネトリン1、2、IFNa、b、g、カスパーゼ2、3、7、8、10、ADAM S1、S5、8、9、15、TS1、TS5；アジポネクチン、ALCAM、ALK-1、APRIL、アネキシンV、アンギオゲニン、アンフィレグリン、アンギオポエチン1、2、4、Bcl-2、BAK、BCAM、BDNF、bNGF、bECGF、BMP2、3、4、5、6、7、8；CRP、カドヘリン6、8、11；カテプシンA、B、C、D、E、L、S、V、X；CD11a/LFA-1、LFA-3、GP2b3a、GH受容体、RSV Fタンパク、IL-23(p40、p19)、IL-12、CD80、CD86、CD28、CTLA-4、 $\alpha 4\beta 1$ 、 $\alpha 4\beta 7$ 、TNF/リンフォトキシン、VEGF、IgE、CD3、CD20、IL-6、IL-6R、BLYS/BAFF、IL-2R、HER2、EGFR、CD33、CD52、ジゴキシン、Rho(D)、水痘、肝炎、CMV、破傷風、ワクシニア、抗毒、ボツリヌス、Trail-R1、Trail-R2、cMet、TNF-Rファミリー、例えば、LANKG-R、CD27、CD30、CD40、CD95、リンフォトキシンa/b受容体、Ws1-1、TL1A/TNFSF15、BAFF-R/TNFRSF13C、TRAIL R2/TNFRSF10B、TRAIL R2/TNFRSF10B、Fas/TNFRSF6、CD27/TNFRSF7、DR3/TNFRSF25、HVEM/TNFRSF14、TROY/TNFRSF19、CD40リガンド/TNFSF5、BCMA/TNFRSF17、CD30/TNFRSF8、LIGHT/TNFSF14、4-1BB/TNFRSF9、CD40/TNFRSF5、GITR/TNFRSF18、オステオプロテグリン/TNFRSF11B、RANK/TNFRSF11A、TRAIL R3/TNFRSF10C、TRAIL/TNFSF10、TRANCE/RANKL/TNFSF11、4-1BBリガンド/TNFSF9、TWEAK/TNFSF12、CD40リガンド/TNFSF5、Fasリガンド/TNFSF6、RELT/TNF

20

30

40

50

RSF19L、APRIL/TNFSF13、DcR3/TNFRSF6B、TNFRI/TNFRSF1A、TRAILR1/TNFRSF10A、TRAILR4/TNFRSF10D、CD30リガンド/TNFSF8、GITRリガンド/TNFSF18である。

【0315】

GITRリガンド/TNFSF18、TACI/TNFRSF13B、NGFR/TNFRSF16、OX40リガンド/TNFSF4、TRAILR2/TNFRSF10B、TRAILR3/TNFRSF10C、TWEAKR/TNFRSF12、BAFF/BLyS/TNFSF13、DR6/TNFRSF21、TNF-アルファ/TNFSF1A、プロ-TNF-アルファ/TNFSF1A、リンフォトキシン・ベータR/TNFRSF3、リンフォトキシン・ベータR(LTbR)/Fcキメラ、TNFRI/TNFRSF1A、TNF-ベータ/TNFSF1B、PGRP-S、TNFRII/TNFRSF1B、EDA-A2、TNF-アルファ/TNFSF1A、EDAR、TNFRI/TNFRSF1A。

10

【0316】

下記の実施例は、本発明の方法に有用な物質の製造法、および、本発明の方法の操作的实施態様を提供することによって本発明を具体的に説明することを意図するものであって、本発明を限定することを意図するものではない。

【実施例】

【0317】

20

(実施例1：CDP6__6__12__3__2のランダム化)

下記の実施例は、CDP6__6__12__3__2に基づくライブラリーの設計を記載する。タンパク配列のTrEMBLデータベースについて、CDP6__6__12__3__2に合致する部分的配列を探索した。合計71個の配列がこのCDPに合致した。表5に示す通り、各位置について、アミノ酸の出現率を計算した。各非システイン位置について、我々は、下記の基準に基づいてランダム化スキームを選んだ。a) 終止コドンの導入を避ける、b) 余分のシステイン残基の導入を避ける、c) 特定位置に>3%で観察されるアミノ酸の多数を許容する、d) CDPに合致する71個の天然配列のいずれにも観察されないアミノ酸の導入を最小とする。

【0318】

30

【化 1 8】

表5:CDP6_6_12_3_2のアミノ酸組成、およびライブラリー設計

位	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	9	109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	6	4	0	1	10	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	45	0	0	6	6	0	5	1	3	1	7	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	31	0	0	0	1	0	0	0	11	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	4	0	6	1	0	0	0	0	11	13	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	0	59	4	1	7	0	0	1	1	0	15	0	1	1	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
8	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	46	0	6	6	0	13	0	0	0	0	0	0	3	4	7	6	6	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	4	3	0	0	1	1	1	4	0	0	0	0	0	8	3	7	3	1	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	52	0	11	0	1	3	0	0	1	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	10	0	0	0	0	0	0	0	23	8	17	5	0	3	1	15	0	10	10	0	0	0	0	0	0	0	0	0	0	0	0	0
13	3	0	6	1	0	1	1	0	3	0	0	4	6	3	1	65	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	1	0	0	0	4	0	0	0	24	0	20	1	0	0	0	4	3	3	4	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	1	1	7	6	0	0	0	0	3	6	1	30	0	21	1	4	17	0	0	0	0	0	0	0	0	0	0	0	0	0
17	17	0	10	5	0	4	8	0	1	0	3	18	0	0	0	6	11	11	7	0	0	0	0	0	0	0	0	0	0	0	0	0
18	5	0	0	4	1	0	0	0	14	6	0	0	1	17	7	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	11	0	3	1	4	20	0	0	0	4	0	1	1	7	0	3	3	10	1	3	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	8	0	0	1	0	10	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	1	0	0	7	3	0	0	0	0	10	0	0	0	0	0	62	0	11	1	3	0	0	0	0	0	0	0	0	0	0	0	0
22	3	0	32	11	1	0	0	0	0	1	0	6	1	14	3	10	6	15	1	0	0	0	0	0	0	0	0	0	0	0	0	0
23	6	0	0	0	51	0	0	4	0	7	6	0	0	0	0	1	5	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	3	0	0	0	6	0	11	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	8	0	0	3	0	0	1	3	8	1	3	51	0	7	10	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	1	0	0	6	0	6	0	0	6	14	0	4	0	23	4	17	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	3	0	1	0	3	3	4	3	0	21	0	4	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	14	0	0	1	0	0	0	0	4	0	0	0	14	40	13	3	1	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	1	0	0	1	1	0	0	0	0	0	42	11	0	0	0	1	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	5	0	7	3	0	41	0	0	16	0	0	20	0	4	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	20	0	4	7	4	6	0	0	0	0	54	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
位	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ
2	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ
1	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ	タカオ

表の最後の三つのコラムは、「ライブラリー1」と表示されるコラムに列挙されるアミノ酸をもたらずコドン混合物を示す。

(実施例 2：大腸菌におけるタンパクの発現およびフォールド形成)

オリゴヌクレオチドを、発現バクテリアにクローンし、大腸菌の細胞原形質におけるタンパクの発現を駆動する。好ましいプロモーターは、大腸菌株BL21 DE3のT7である (NovagenのpETベクターシリーズ; Kanマーカー)。このオリゴを挿入するのに好ましいプロセスは、Kunkelの変法である (Scholle, D., Kehoe, J W and Kay, B. K. (2005) Efficient construction of a large collection of phage-displayed combinatorial peptide libraries, Comb. Chem. & HTP Screening 8: 545-551)。別の方法は、ベクター (全体または部分的) の2-オリゴPCR、次いで、断片のオリゴ由来末

端の、特有の制限部位の消化、次いで、適合性の、非パリンδροーム・オーバーハングの連結（効率的断片内連結）である。第3の方法は、重複PCRによる、2または4個のオリゴからの挿入体の集合、集合挿入体の末端における制限酵素部位の消化、次いで、連結による消化ベクターへの変換である。連結DNAは、コンピテント大腸菌細胞において形質転換され、LB-Kanプレートに撒かれ、一晚増殖された後、個別のコロニーが採取され、2xYT培養液を含む96ウェルプレートに接種され、培養体は、シェーカーにおいて37℃で一晩育成される。

【0319】

プレートは80℃で25分加熱され、6000gで遠心され、凝集大腸菌タンパクのペレットを形成する。

【0320】

（実施例3：抗凍結タンパクの設計工程）

目的：抗凍結反復性タンパクのライブラリーを設計する

戦略：ライブラリー設計の開始配列は、Tenebrio molitor由来の抗凍結タンパク（Genbankアクセス番号AF160494）から得られる。このタンパクは、大腸菌でよく発現することが知られる。結晶構造およびNMR構造の両方が知られる。このタンパクは、円筒形を形成する反復単位から構築される。構造のコアは、疎水性アミノ酸を欠くが、反復単位当たり一つのジスルフィド結合、および一つの不変のセリンおよびアラニン残基を含む。最初の2回の回転は、3個のジスルフィド結合を有するキャップモチーフを形成する。このキャップモチーフが、フォールド形成の核を構成することが仮定される。したがって、インビトロ進化の際、最初の二つの反復単位は、通常、不変のまま維持される。図127を参照されたい。

【0321】

交差点を選び、Scholleの突然変異発生のためのグルタミン残基の位置を見出すために、抗凍結タンパクの構造特徴を分析した。

【0322】

交差点は、赤で示されるが、構造に認められるベータ・シート積層体を保存するように選ばれた。末端キャップのループについては、抗凍結オープン読み枠の外に位置する、一般的上流成熟部位を用いて後期段階で突然変異を発生させることが可能である。突然変異発生のためのコドンを選ぶために、抗凍結タンパクファミリーを記載するPfamのウェブページから、215反復単位から成る整列（PfamデータベースにおいてPF02420）をダウンロードした。このテキストファイルを、プログラムProfile analyzer v1.0において、システイン位置“2.8”、および反復列の全長“12”に設定して分析した。この設定によって、12アミノ酸反復当たり3個のシステインを含むN-末端反復単位が排除された。したがって、プログラムは、89個の配列を排除し、抗凍結反復列における各アミノ酸の保存および出現を示す、残余の126個の配列を分析する。この出力を、Excelのスプレッドシートに貼り付け、ライブラリー設計の開始点として用いた。

【0323】

（実施例4：3本指トキシシン（エラプトキシシン）の設計工程）

目的：3本指トキシシンのフカフォールドを用いてライブラリーを設計する。

【0324】

背景：3本指トキシシンは、四つのジスルフィドコア、およびこのコアから突出する3本の長いループを有する独特の構造を呈する。これらのループは、様々のタンパク・タンパク相互作用に関与することが知られ、かつ、指向性進化の標的対象となり得る。

【0325】

方法：もっとも一般的なシステインの間隔パターンは、10-6-16-3-10-0-4、13-6-16-1-10-0-4、および13-5-16-1-10-0-4である。エラプトキシシンの配列：

TRICFNHQSSQPQTTKTCS PGESSCYNKQWSDFRGI I ERG

C G C P T V K P G I K L S C C E S E V C N N A 開始配列として選ばれ、これは、1 3 - 6 - 1 6 - 1 - 1 0 - 0 - 4 に一致する。この配列が選ばれたのは、これは大腸菌で発現可能であるからである。

【0326】

ループ領域において最大多数の突然変異を可能とするために二つの交差点を選んだ。

【0327】

(実施例5：プレキシンの設計工程)

目的：プレキシソおよびP S I スカフォールドを利用するライブラリーを設計する。

【0328】

このスカフォールドの利点：このスカフォールドは、個々のシステイン残基間に長さ変動を導入するのに独特の利点を提供する。天然では、P S I フォールドのシステイン間には著明な長さ変動が認められるが、これは、この設計原理を支持する。ループ長の多様性は、ミクロタンパク・ファミリーにおいて最高にランクされる。図135は、A A 残基の付加による、ゆっくりとした長さ延長によって創製される「マルチ・プレキシソ」を示す。

【0329】

戦略：P f a m データベースは、468ファミリーメンバーを列挙する。C y s 5 / C y s 6、C y s 6 / C y s 7、およびC y s 7 / C y s 8間のシステイン間隔は、きわめて変動性が高い。したがって、開始の共通配列を選ぶことが困難である。M e t 受容体のP S I ドメインのNMR構造は解明されており、5, 2, 8, 2, 3, 5, 9のパターンを示す。このタンパクは、やや低レベルではあるが(細胞9リットル当たり1mg)、大腸菌において発現されている。データベースについて、5, 2, 8, 2間隔を呈するメンバーを探索したところ、99個の配列が見つかった。しかしながら、これらの内の僅かに11%だけがモチーフ5, 2, 8, 2, 3を持つにすぎず、5, 2, 8, 2, 3, 5, 9を持つのは3メンバーにしかない。したがって、この空間パターンは無視し、このファミリーのもっとも一般的間隔パターンを求めた。5, 2, 7, 2, 5による探索から、54個の配列が得られた。これらの配列を、E x c e l のスプレッドシートに整列させ、各位置におけるもっとも一般的コドンを得た。最終間隔がもっとも変動性が高く、全体タンパクドメインの偶数挿入が認められる。5, 2, 7, 2, 5を有する54メンバーの最終位置におけるもっとも一般的間隔は、“15”である。まとめると、P S I フォールドの共通配列は、パターン5, 2, 7, 2, 5, 15を有するファミリーメンバーから得られた。

【0330】

構造“1 s s l”は、M e t 受容体のP S I ドメインを示す。交差点は、もっともよく保存されるファミリーモチーフC G W Cが不変となるように設計した。これによって、スカフォールドの最初の半分のランダム化が可能となる。第2の交差点は、C y s 7に挿入された。これによって、天然では大きな長さ変動を示す、システイン間隔5、6、および7のランダム化を最大とすることが可能になる。図119を参照されたい。

【0331】

図120：ライブラリー共通配列と、共通配列5, 2, 8, 2, 3, 5(11メンバーのみ)との整列は、25%の同一性を示す。最大の多様性は、最後のc y s 間隔にあり、これは、式、および他のメンバーとの比較と一致する。

【0332】

(実施例6：ソマトメジンの設計工程)

目的：ソマトメジンスカフォールドを利用してライブラリーを設計する。

【0333】

戦略：共通配列E E S C K G R C G E G F N R G K E C Q C D E L C K Y Y Q S C C P D Y E S V C K P Kは、同じシステイン間隔パターンを持つ44配列から得られた。

【0334】

交差点は、配列の両半分における突然変異発生を可能とするために、タンパクのほぼ中点に選んだ。図121を参照されたい。

【0335】

(実施例7：マイクロタンパクのスカフォールド発現の評価)

抗凍結タンパク (AF)、3本指トキシシン (TF)、ソマトメジン (SM)、およびプレキシシン (PL) の、マイクロタンパク・オープンリーディングフレームを、pET30由来ベクターでクローンし、大腸菌株BL21 (DE3) において発現させた。一晚の培養体を、1:200で希釈し20 mL LBとし、3時間育成し、2 mM IPTGで誘発し、さらに4時間育成した。培養体を5000 x gで10分回転させ、PBSに再懸濁した。250 μ lのサンプルを、30°Cで30分加熱し、RTで10分遠心した。加熱工程 (50 μ lサンプル) の上清を、5% BMEを含む25 μ lのサンプルと混ぜ合わせ、再懸濁細胞 (50 μ l) を直接、5% BMEを含む25 μ lのサンプルバッファーと混ぜ合わせた。サンプルを10分煮沸し、次いで16% SDS-PAGEに負荷した。

10

【0336】

結果：図122を参照されたい。左から右へ (16% SDS-PAGE)：部分的に精製されたタンパク：陽性コントロール、新規AFスカフォールド、新規TFスカフォールド、新規SMスカフォールド、PL (短小バージョン)、コントロール、NEB広範囲、次に、同じタンパクの全体細胞標本について同じ順序。

【0337】

結論：タンパクTF、SM、PLは、上清に高濃度に存在し、極めて耐熱性が高い。

【0338】

(実施例8：ファージミドベクターpMP0003の構築)

20

我々は、マイクロタンパク・ライブラリーの効率的構築のためにベクターを構築した。ベクターの背景は、pBluescriptファージミドベクターに基づく。lacZプロモーターによって駆動される発現カセットを挿入した。コード配列は下記の要素を含む。ompAシグナルペプチド、SfiIおよびBstXI部位によって側接される、短い詰め物配列、リンカー要素、ヘキサヒスチジン・タグ、ヘムアグルチニン (HA) タグ、アンバー終止コドン、M13ファージのpIIIタンパクのC-末端断片、終止コドンである。詰め物配列は僅かに40 bp長である。これは、TAAおよびTGAの二重終止コドン、および固有のBssHII部位を含む。大型のファージミドライブラリーの構築は、消化された精製ベクター断片の十分な量が得られないことによって制限されることが多い。pMP0003の設計は、予備的なアガロースゲル電気泳動によるベクター断片の精製の必要を回避するために、準備工程を大いに簡便化する。プラスミドpMP0003を、SfiI、BstXI、およびBssHIIによって三重に消化することによって、二つの、きわめて短い、19および21 bp長の、詰め物断片が放出される。これらは、YM-100カラム (Microcon) による限外ろ過によって除去することが可能である。詰め物におけるBssHII部位の存在によって、pMP0003に基づくライブラリーにおける非組み換えクロンの頻度の著明な減少が得られる。

30

【0339】

(実施例9：ライブラリーLMB0020の設計と構築)

ランダムクローンライブラリーは、多くのマイクロタンパク配列に基づいて構築される。プロセスは、いくつかの工程：1) 適切なマイクロタンパク・スカフォールドを特定すること、2) ランダム化のための残基を特定すること、3) 各ランダム化位置に対しランダム化スキームを選択すること、4) マイクロタンパク・スカフォールドをコードし、かつ、ランダム化スキームに従って特定位置にヌクレオチド混合物を組み込む、部分的にランダムなオリゴヌクレオチドを設計すること、5) マイクロタンパク断片を集合すること、6) 制限消化および精製、7) 断片を、消化されたベクター断片に連結すること、7) コンピテント細胞における形質転換、を含む。

40

【0340】

ライブラリーLMB0020は、squashファミリープロテアーゼ阻害剤の1メンバーである、トリプシン阻害剤EETI-Iの配列に基づく (Christmann, A., ら、(1999) Protein Eng, 12:797-806)。EETI-I

50

IIの結晶構造が調べられ、10位置がランダム化のために選ばれた。9位置は、ランダムコドンNNKを用いてランダム化された。これは、16種のアミノ酸（A、D、E、F、H、I、K、L、M、N、P、Q、S、T、V、Y）の導入を可能にする。1位置では、ランダムコドンVNKが使用され、これは16種のアミノ酸（A、D、E、F、H、I、K、L、M、N、P、Q、S、T、V、Y）を可能にする。得られたランダム配列は：GCPXXXXXXCKQDSDCXGCVCPXGXCGSPであり、式中、XはコドンNHKを表し、ZはコドンVNKを表す。このランダム化スキームによって、10¹²を上回る異なるアミノ酸配列から成る理論的多様性が可能となる。ランダム化トリプシン阻害剤をコードする遺伝子断片は、下記の配列を有する、二つのオリゴヌクレオチドの重複伸長によって集合した。

【0341】

LMB0020F=CAGGCAGCGGGCCCGTCTTGGCCCCGGGTGTCTCTNHKKNHKNHKNHKNHKNHKTGTAAACAAGACTCTTGACTG

【0342】

【化19】

LMB0020R=TGTAACAAGACTCTGACTGTNHKNIHGGTTGCGTTTGCVNKCCGNHKGGTNHKTGTGGCTCTCCGGGCCAGTCTGGTGGTTCCGGTCACGTGACCGGAACCACCAGACTGGCCCCGGAGAGCCACAMDNACCMDNCGGMNBGCAAACGCAACCMDNMDNACAGTCAGAGTCTTGTITACA.

オリゴヌクレオチドLMB00020FおよびLMB0020Rは、20ヌクレオチドの相補性領域を共有する。二つの相補的プライマーをアニールした後、充填反応を行って、2段階PCR増幅を実行した。次に、産物を、制限部位を含むスカフォールド・プライマーLIBPTFおよびLIBPTRを用いて増幅した。

【0343】

得られた産物は、YM-30フィルター（Microcon）を用いて濃縮し、1.2%アガロースによる予備的アガロースゲル電気泳動によって精製した。

【0344】

10μgの産物を、SfiI/BstXIによって50℃で5時間消化し、PCRカラム（Qiagen）で急速精製したところ、4μgの精製断片が得られた。ベクターpMP0003は、QIAGEN HiSpeed Maxiキットを用いて調製した。150μgのベクターDNAを、3本の、別々のエッペンドルフチューブにおいて、SfiI/BstXI/BssHIによって50℃で4時間消化し、YM-100カラム（Microcon）において精製した。全体収率は、消化ベクターの112.5μg（75%）であった。ライブラリーにおける形質転換株の数を最大化するために、様々な挿入体・対・ベクター比を、小規模実験で調べた。7本の連結チューブにおいて大規模連結を行った。各チューブは、総容量400μlに溶解した、3μgの消化ベクター、0.5μgの消化挿入体（1:2.5比）、40μlのリガーゼバッファー、20μlのT4 DNAリガーゼを含む。連結は、16℃で一晩行った。得られた産物は、各ライブラリーについて、8本のチューブにおいて、-20℃で一晩エタノール沈殿によって精製した。各チューブにおける連結DNAは、30mlの蒸留水に溶解し、2x15μlに分割し、ライブラリー当たり形質転換のために16本のチューブを生成した。

【0345】

電氣的受容能を備えた大腸菌ER2738を下記のプロセスによって調製した。1) LB寒天に新たに線条接種したグリセロールから得た、大腸菌単一コロニーを、50-mlポリプロピレンチューブにおいて、15mlのあらかじめ温めたSuperbroth培養液（SB）に接種する。テトラサイクリンを30μg/ml（90μlの、5mg/mlテトラサイクリン）まで加え、37℃においてシェーカーで250rpmで一晩育成する。2) 2.5mlの培養体を、4本の2リットルフラスコのそれぞれにおいて、500mlのSB培養液にて希釈し、10mlの20%グルコース、5mlの1M MgCl₂、および500μlの5mg/mlのテトラサイクリンを加える。250rpm、37℃

10

20

30

40

50

で、600 nmにおける吸光度が約0.9になるまで（2時間45分）振とうする。3）培養体、および4本の500-ml瓶を氷上で15分冷却する。4）培養体を、4本の500-ml瓶に移し、4℃で20分4000 rpmで回転する。5）上清を傾斜除去し、各ペレットを、25-mlのあらかじめ冷却したピペットを用い、あらかじめ冷却した、25 mlの10%グリセロールに再懸濁する。1本の250-ml瓶に2ペレットを合わせ、10%グリセロールを加え、250 mlを生成する。前と同様に遠心する。6）上清を傾斜除去し、工程5を繰り返す。7）上清を捨て、残余の容量（3.5 ml）に各ペレットを再懸濁する。全ての懸濁液を合わせる。ライブラリーの電気穿孔には300 μ lの分液を使用する。任意に行ってよいもの：保存するために、320 μ lの分液をエッペンドルフチューブに移し、エタノールおよびドライアイスを用いて瞬間凍結する。チューブをキャップで閉じ、それらを-80℃で保存する。8）LB寒天（100 mg/lのカルベニシリン）の上に50 μ lの細胞懸濁液をプレートし、ベクターファージの汚染について調べる。LB寒天（50 mg/lのカナマイシン）の上に50 μ lの細胞懸濁液をプレートし、ヘルパーファージの汚染について調べる。

【0346】

ライブラリーの電気穿孔は、下記の工程を用いて実行した。1）連結されたDNA（通常16）および対応する数のキュベットを、氷上に10分置く。2）新しく調製したER2738細胞を、各連結ライブラリーサンプルに加え、1回パイペッティングすることによって混ぜ合わせ、キュベットに移す。氷上に1分保存する。2.5 kV、25 μ F、および200 オームで電気穿孔する。ただちに、2 ml、次いで1 mlのSOC培養液でキュベットを室温で噴射する。10 mlの培養チューブにおいて3 mlの培養体を合わせる。37℃で1時間300 rpmで振とうする。3）二つの3 mlサンプルを合わせ、50-mlのポリエチレンチューブに移す。9 mlのあらかじめ温めておいた（37℃）SB培養液、3 μ lの100 mg/mlのカルベニシリン、および15 μ lの5 mg/mlのテトラサイクリンを加える。形質転換細菌の力価定量のために、培養体を、200 μ lのSB培養液に希釈し、この1:100希釈液の10 μ lおよび1 μ lをLB寒天（100 mg/lのカルベニシリン）にプレートする。プレートを37℃で一晩インキュベートする。コロニーの数を数え、培養体容量を掛け、プレート容量で割ることによって形質転換細胞の総数を計算する。15-mlの培養体を37℃で1時間300 rpmで振とうし、4.5 μ lの100 mg/mlのカルベニシリンを加え、さらに1時間37℃において300 rpmで振とうする。4）二つの15 mlサンプルを合わせ、3 mlのVC SM13ヘルパーファージを加える。500-mlのポリプロピレン遠心瓶に移す。167 mlの、あらかじめ温めた（37℃）SB培養液、92.5 μ lの100 mg/mlのカルベニシリン、および185 μ lの5 mg/mlのテトラサイクリンを加える。この200-ml培養体を300 rpmで37℃において1.5-2時間振とうする。5）280 μ lの50 mg/mlのカナマイシンを加え、37℃で一晩300 rpmで振とうを続ける。6）4℃で15分4000 rpmで遠心する。上清を、清潔な500-mlの遠心瓶に移し、50 mlの20%PEG-8000/NaCl 2.5 Mを加える。氷上に30分保存する。7）4℃で15分9000 rpmで遠心する。上清を捨て、遠心瓶をペーパータオルの上に少なくとも10分倒立して液体を排除し、遠心瓶の上部から、残余の液体をペーパータオルで拭い去る。8）ファージペレットを、2 mlの、トリスバッファー生食液（TBS）に溶解した1%（w/v）ウシ血清アルブミン（BSA）において、遠心瓶の側面におけるパイペッティングによって再懸濁し、2-mlのマイクロ遠心チューブに移す。1-mlのピペットの先端を用いてさらにパイペッティングすることによって再懸濁し、マイクロ遠心チューブにおいて4℃で5分間フルスピードで遠心し、上清を、0.2 μ mフィルターを通過させて、滅菌した2-mlマイクロ遠心チューブに移す。このファージ標本を4℃で保存する。長期の保存のために、アジ化ナトリウムを0.02%（w/v）まで加えてもよい。得られた、LMB0020のライブラリーサイズは、 2.4×10^9 形質転換体であった。

【0347】

(実施例 10 : ライブラリー LMB0020 のパンニング)

1) Costar の 96-ウェル ELISA プレートのウェルに、 $25\mu\text{l}$ の PBS に溶解した $0.25\mu\text{g}$ の CD22 抗原をコートする。プレートを、プレートシーラーにて覆う。コーティングは、 4°C で一晚、または 37°C で 1 時間行ってもよい。第 1 ラウンドのパンニングでは、スクリーニングされるライブラリー当たり 2 ウェルをコートする。1 ウェルで、後続ラウンドのそれぞれに十分である。標的濃度は、パンニングのラウンド 3 から 6 においてウェル当たり $0.1\mu\text{g}$ に下げる。

【0348】

2) コーティング液を十分に振とうした後、 $150\mu\text{l}$ の TBS / BSA 3 % (3 % ウシ血清アルブミンを含むトリスバッファー生食液) を加えることによってウェルをブロックする。シールし、 37°C で 1 時間インキュベートする。

10

【0349】

3) ブロッキング液を十分に振とうした後、 $50\mu\text{l}$ の新たに調製したファージライブラリーをウェルに加える (インプットサンプル)。プレートをシールし、 37°C で 2 時間インキュベートする。その間、 2ml の SB 培養液、プラス $2\mu\text{l}$ の $5\text{mg}/\text{ml}$ テトラサイクリンに対し、 $2\mu\text{l}$ の ER2738 細胞標本を接種し、 37°C で 2.5 時間 250rpm で育成させる。各ライブラリー当たり、スクリーニングされる 1 培養体、インプット力価測定のためにさらに 1 培養体を育成する。

【0350】

4) ファージ溶液を十分に振とうした後、 $150\mu\text{l}$ の、TBS / Tween-20 0.05 % をウェルに加え、活発に 5 回パイペッティングを行う。5 分待ち、十分振とうし、この洗浄工程を繰り返す。第 1 ラウンドのパンニングではこのようにして 4 回洗浄し、第 2 ラウンドでは 6 回、第 3 ラウンドでは 8 回、以下同じ。

20

【0351】

5) 最終洗浄液を十分に振とうした後、 $50\mu\text{l}$ の、TBS に溶解した新たに調製した $10\text{mg}/\text{ml}$ トリプシンを加え、シールし、 37°C で 30 分インキュベートする。活発に 10 回パイペッティングを行い、溶出液 (第 1 ラウンドでは $2 \times 50\mu\text{l}$ 、その後のラウンドでは $1 \times 50\mu\text{l}$) を、調製された 2- ml の大腸菌培養体に移し、室温で 15 分インキュベートする。

【0352】

30

6) 6ml の、あらかじめ温めておいた SB 培養液、および、 $1.6\mu\text{l}$ の $100\text{mg}/\text{ml}$ のカルベニシリン、および $6\mu\text{l}$ の $5\text{mg}/\text{ml}$ のテトラサイクリンを加える。アウトプット力価測定のために、 $2\mu\text{l}$ のサンプルを $200\mu\text{l}$ の SB 培養液で希釈し、 $100\mu\text{l}$ および $10\mu\text{l}$ のこのサンプルを、LB 寒天 ($100\text{mg}/\text{ml}$ のカルベニシリン) にプレートする (アウトプットサンプル)。平行して、 $50\mu\text{l}$ の、調製された 2- ml の大腸菌培養体に対し、 $1\mu\text{l}$ の、ファージ標本の 10^{-8} 希釈液を感染させることによってインプット力価測定を進め、室温で 15 分インキュベートし、LB 寒天 ($100\text{mg}/\text{l}$ のカルベニシリン) にプレートする。

【0353】

7) 8- ml の培養体を、 37°C で 1 時間 250rpm で振とうし、 $2.4\mu\text{l}$ の $100\text{mg}/\text{ml}$ のカルベニシリンを加え、さらに 1 時間 37°C において 250rpm で振とうする。

40

【0354】

8) 1ml の VCSM13 ヘルパーファージを加え、500- ml のポリプロピレン遠心瓶に移す。91 ml の、あらかじめ温めておいた (37°C) SB 培養液、および $46\mu\text{l}$ の、 $100\text{mg}/\text{ml}$ のカルベニシリン、および $92\mu\text{l}$ の、 $5\text{mg}/\text{ml}$ のテトラサイクリンを加える。この 100- ml 培養体を、 37°C で 1 1/2 から 2 時間 300rpm で振とうする。

【0355】

9) $140\mu\text{l}$ の、 $50\text{mg}/\text{ml}$ のカナマイシンを加え、 37°C で一晚 300rpm

50

で振とうを続ける。

【0356】

10) 4℃で15分4000rpmで遠心する。上清を、清潔な500-mlの遠心瓶に移し、25mlの20%PEG-8000/NaCl 2.5Mを加える。氷上に30分保存する。

【0357】

11) 4℃で15分9000rpmで遠心する。上清を捨て、遠心瓶をペーパータオルの上に少なくとも10分倒立して液体を排除し、遠心瓶の上部から、残余の液体をペーパータオルで拭い去る。

【0358】

12) ファージペレットを、2mlの、TBS/BSA 1%バッファーにおいて、遠心瓶の側面におけるパイペッティングによって再懸濁し、2-mlのマイクロ遠心チューブに移す。1-mlのピペットの先端を用いてさらにパイペッティングすることによって再懸濁し、マイクロ遠心チューブにおいて4℃で5分間フルスピードで遠心し、上清を、0.2μmフィルターを通過させて、滅菌した2-mlマイクロ遠心チューブに移す。

【0359】

13) 次のラウンドのために工程3)から続けるか、または、ファージ標本を4℃で保存する。長期の保存のために、アジ化ナトリウムを0.02%(w/v)まで加えてもよい。各ラウンドには新鮮な調製ファージのみを使用しなければならない。

【0360】

【化20】

表6は、6ラウンドのライブラリーパンニングにおけるインプットおよびアウトプット液のファージカ価を示す

ラウンド	インプット (10^{11})	アウトプット(10^6)	回収 (% $\times 10^3$)	濃縮度
1	12	1.9	0.16	-
2	0.45	0.032	0.007	neg
3	4.7	2.14	0.46	2.87
4	2.5	0.064	0.032	neg
5	0.52	1.2	2.3	14.37
6	0.6	2.0	3.33	20.8

(実施例11：標的結合に関する個別分離体のスクリーニング)

ER2738に、アウトプットファージを感染させ、LB寒天(100mg/lカルベニシリン)にプレートした。プレートを37℃で一晩インキュベートした。次いで、個々のコロニーについて、下記のように、標的タンパクに対する結合についてスクリーニングすることが可能である。

【0361】

1) 50μg/mlのカルベニシリンを含む、0.75mlのSB培養液を、ディーブ・ディーブウェルを有する96ウェルに加える。滅菌楊枝を用いて個々のコロニーを各ウェルに移す。2) 細菌培養体を含むプレートを37℃で数時間300rpmで振とうする。

【0362】

2) 接種後6時間において、1μlの各培養体を、LB寒天(100mg/mlのカルベニシリン)に滴下する。プレートを37℃で一晩インキュベートする。プレートをパラフィンでシールし、次いで4℃で保存する。これらのプレートは、後に、ELISA陽性シグナルを示す分離体を回収し、配列決定するのに用いられる。

【0363】

3) IPTGを、1 mM (7.5 μ l の、水で1:10に希釈した1 M IPTGストック) となるまで加えて培養体を誘起し、37℃で一晩培養する。

【0364】

4) 誘発された大腸菌培養体を遠心する (4000 rpm、20分)。

【0365】

5) Bugbuster液 (Novagen) を調製する (1.5 ml の試薬、プラス13.5 ml のTBSおよび15 μ l のベンゾナーゼ)。

【0366】

6) ペレットを150 μ l のbugbusterに再懸濁する。プレート室温で30分インキュベートし、プレートを4000 rpmで20分遠心する。

10

【0367】

7) ウェル当たり50 μ l の上清を、PBSに溶解し、3% BSAを含む、ウェル当たり150 μ l のTBSで1時間ブロックした、ウェル当たり100 ngの標的タンパクで、4℃で一晩コートしたマイクロタイタープレートに移した。

【0368】

8) プレートを37℃で2時間インキュベートする。

【0369】

9) 水道水で10回洗浄する。

【0370】

10) ラットの、ビオチニル化抗HA抗体 (3F10、Roche Biosciences) を、TBS/BSA 1%で希釈する (1:500希釈)。50 μ l の希釈抗体をウェルに加え、37℃で1時間インキュベートする。

20

【0371】

11) 水道水で10回洗浄する。

【0372】

12) ストレプトアビジン/HRPをTBS/BSA 1%で希釈し (1:2500希釈)、37℃で30分インキュベートする。

【0373】

13) ABTS液 (2.94 ml のクエン酸バッファー+60 μ l のABTS+1 μ l のH₂O₂) を調製する。

30

【0374】

14) 水道水で10回洗浄する。

【0375】

15) 各ウェルに50 μ l の基質液を加える。

【0376】

16) RTでインキュベートし、室温で20分インキュベーション後、ELISAプレートリーダーを用いて405 nmにおけるO.D.を読み取る。

【0377】

ライブラリーLMB0020の外、他の、二つのマイクロタンパクライブラリーの、ラウンド5から得られたアウトプットに対し、前述のようにスクリーニングを行った。下表は、IgGでコートしたプレート、および、BSAでコートしたプレートから得られた結合データを示す。いくつかの分離体は、BSAでコートしたプレートに比べて、IgGでコートしたプレートにおいて有意に高い結合シグナルを示す。

40

【0378】

【化 2 1】

IgG		1	2	3	4	5	6	7	8	9	10	11	12	
A		0.14	0.11	0.10	0.10	0.10	0.11	0.10	0.12	0.14	0.11	0.13	0.13	SMP3S5
B		0.11	0.29	0.11	0.10	0.10	0.11	0.10	0.12	0.12	0.17	0.59	0.33	SMP3S5
C		0.24	0.27	0.16	0.23	0.11	0.19	0.12	0.10	0.10	0.10	0.11	0.16	SMP3S5
D		0.12	0.10	0.10	0.14	0.12	0.11	0.09	0.15	0.09	0.09	0.10	0.10	SMP3S5
E		0.10	0.11	0.10	0.17	0.09	0.09	0.10	0.15	0.15	0.11	0.10	0.10	SMP3S5
F		0.10	0.10	0.10	0.11	0.11	0.09	0.11	0.10	0.10	0.10	0.10	0.14	SMP3S5
G		0.46	0.12	0.33	0.20	0.40	0.11	0.09	0.33	0.09	0.09	0.10	0.30	SMP4S5
H		0.12	0.12	0.11	0.10	0.13	0.07	0.09	0.41	0.09	0.12	0.48	0.15	SMP5S5

10

BSA	A	1	2	3	4	5	6	7	8	9	10	11	12	
B		0.10	0.10	0.10	0.10	0.09	0.10	0.10	0.10	0.12	0.10	0.10	0.10	SMP3S5
C		0.10	0.14	0.09	0.09	0.09	0.09	0.09	0.10	0.10	0.11	0.15	0.12	SMP3S5
D		0.12	0.12	0.10	0.13	0.09	0.12	0.10	0.11	0.10	0.09	0.10	0.10	SMP3S5
E		0.10	0.09	0.09	0.10	0.10	0.10	0.10	0.11	0.09	0.09	0.13	0.09	SMP3S5
F		0.09	0.10	0.09	0.12	0.09	0.09	0.09	0.10	0.12	0.09	0.09	0.10	SMP3S5
G		0.09	0.09	0.09	0.09	0.10	0.09	0.09	0.09	0.09	0.09	0.09	0.10	SMP3S5
H		0.14	0.09	0.11	0.09	0.11	0.09	0.09	0.12	0.09	0.09	0.09	0.11	SMP4S5
		0.10	0.09	0.10	0.09	0.10	0.09	0.09	0.15	0.09	0.11	0.18	0.11	SMP5S5

20

三つの I g G-結合分離体の配列を決定した。分離体は全て、トリプシン抑制因子スカフォールドの 6 個のシステイン残基間の間隔を維持していた。三つの分離体は全て、そのアミノ酸配列において異なっていた。これは、この方法が、それぞれが、その後の最適化の開始点となる、複数の結合ドメインを生成することが可能であることを証明する。

LMB0020 / SMP003S5. B2

G P S G P G C P I L Y A H C K Q D S D C V T G C V C R P L G M C G S P G Q S G G S G H H H H H H

LMB0020 / SMP003S5. B12

G P S G P G C P S L P T P C K Q D S D C D E G C V C K P N G T C G S P G Q S G G S G H H H H H H

LMB0020 / SMP003S5. C2

G P S G P G C P L Y S P V C K Q D S D C D N G C V C R P A G P C G S P G Q S G G S G H H H H H H

30

(実施例 12 : ミクロタンパク設計に対する構築法)

VEGF に結合する 1-ジスルフィドタンパク (1SS) を、プロテアーゼに対してより安定で、より免疫原性の低い 2SS ミクロタンパクへと段階的に進化させた。図 1 は、1SS フェージ由来ペプチド (「VEGF ペプチド」) から得られた、二つの別々の 2SS タンパク (「クローン 2」および「クローン 7」) の ELISA 結果を示す。三つとも全て VEGF に対して特異的であるが、BSA などの他のタンパクに対しては結合を示さない。ミクロタンパクを持たない M13 も、VEGF または BSA に結合しない。この 2SS タンパクは、VEGF 結合を決めた 1SS 配列を、2SS スカフォールド (アルファ-コノトキシン) に移動させることによって創製した。得られたタンパクは、VEGF に対して特異的であるが、無関係なタンパク、例えば、ウシ血清アルブミン (BSA) に対しては結合しない。野生型フェージ粒子 (M13) は、VEGF、BSA のいずれにも結合を示さない。図 168 を参照されたい。

40

【0379】

(実施例 13 : メガプライマー突然変異発生によるライブラリーの構築)

メガプライマー法とは、Kunkel 型ポリメラーゼ伸長反応 (終止コドン置換を用

50

いて、組み込みをきわめて高効率としたことを除いて)において、二つ(またはそれ以上)の異なるプライマーを組み合わせて、その両末端における相同性を通じて、プラスミドに組み込まれる、単一の大型プライマーを形成する方法である。メガプライマー法は、DNAおよびコードされたタンパク配列の複合プールの導入または転送のために、60、70、80、90、100、110、または、好ましくは、120を超えるヌクレオチドまたは塩基対から成る2本鎖、または1本鎖DNAを使用する。我々の例では、これらのプールは、マイクロタンパク・ライブラリーをコードするが、同じ方法は、任意のDNAまたはタンパクライブラリーをコードすることが可能である。メガプライマーは、通常、以前に選ばれた配列から成るプール(「先行ライブラリー」)の外、新たにランダム化された配列から成るプール(「新規ライブラリー」)を含む。したがって、メガプライマー法は、先行ライブラリーの配列決定を要することなく、先行ライブラリーから、新規ライブラリーの盲目的創製を可能とする。

10

【0380】

通常、PCR断片は、以前に選ばれた配列プールのライブラリー区域(「ランダム化区域」)から創製され、この断片は、新たにランダム化されたライブラリー・セグメント(未選択)をコードする合成オリゴに対して連結され(PCR重複によって)、新規(未選択)および先行(選択)ランダム化区域の両方を含むdsDNA断片を創製する。同じ最終結果は、プライマーの一方が新規ライブラリーを導入する場合、「先行ライブラリー」区域の両側のプライマーを用いて、単一PCRにおいて実現することが可能である。このdsDNA断片は、非対称またはランオフPCRによってssDNAメガプライマーに変換される。このssDNAメガプライマーの末端は、ベクターに対し、約10-25塩基の配列相同性を持つように設計され、これが、適正な位置への挿入を確実なものとする。

20

【0381】

2本鎖メガプライマーは、重複PCRを用い、2本以上のPCR断片および/または合成オリゴヌクレオチドから生成され、1本鎖DNAは、変性2本鎖PCR産物および/または1本鎖DNA「非対称PCR」(「ランオフPCR」)を用いて生成することが可能である。この非対称PCRは、前記1本鎖DNA鑄型に対し相補的な1本鎖配列を増幅する。メガプライマー配列は、単一配列を含むことも可能であるが、より一般的には、配列(例えば、マイクロタンパクの)ライブラリーを含む(図143に記載されるように)。1本鎖鑄型DNA(ベクターまたはファージ)は、ユリジン含有であってもよく、あるいは、終止コドンを持たないメガプライマー配列と交換される、抑制性終止コドン(TAG、TAA、TGA)をコードしていてもよい。アニールされたメガプライマーは、ポリメラーゼによって、DNAの第2鎖の合成のプライマーとして作動し、かつ、この合成鎖の連結は、バッファー、DNAポリメラーゼ、DNAリガーゼ、およびデオキシヌクレオチド3リン酸(dNTP)の存在下で、共有的に閉鎖された環状DNA(ccc-DNA)を生成するのに使用される。得られたccc-DNAは、マイクロタンパクを、不溶性タンパク、可溶性タンパク、またはタンパク融合物として発現させるために、細菌の細胞系統中に形質転換される。

30

【0382】

メガプライマー結果の例を下表に示す。図は、最初の15位置において突然変異発生したマイクロタンパクのアミノ酸配列を示す。最初のマイクロタンパク鑄型に合致する保存残基には灰色に影をつけた。図2の配列を含むマイクロタンパク・ライブラリーを、メガプライマー合成の開始点として用いた。「先行ライブラリー」区域、および新規ライブラリーを含むPCR断片を創製するために二つのDNAプライマーを用いた。i)マイクロタンパクの上流でアニールするプライマー、およびii)マイクロタンパク特異的アニーリング領域、およびDNA鑄型アニーリング領域によって側接される、新たにランダム化されたマイクロタンパク配列(「新規ライブラリー」)を含むプライマー、である。マイクロタンパク・ライブラリーを、前記二つのプライマーを用い、PCRによって増幅し、非対称PCRによって増幅し、1本鎖DNA鑄型にクローンし、二次的マイクロタンパク・ライブラリーを

40

50

生成した。得られたクローン（図2下段）から、元の配列の第1半分および第2半分の両方においてランダム化されたマイクロタンパクが明らかになった。

【0383】

【化22】

メガプライマー突然変異発生またはクローニングのためのインプット配列

マイクロタンパク	E	E	S	G	K	G	R	C	G	E	G	F	N	R	G	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン1	D	V	S	C	D	S	R	C	K	K	A	H	Q	L	H	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン2	V	G	S	C	K	G	R	C	K	P	T	I	V	E	G	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン3	L	L	S	C	P	G	R	C	P	T	R	F	V	L	V	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン4	I	S	S	C	P	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン5	I	V	S	C	S	G	R	C	A	H	D	S	A	S	Q	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン6	I	T	S	C	P	G	R	C	N	N	S	H	P	A	I	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン7	L	S	S	C	P	G	R	C	R	G	Q	P	L	P	P	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン8	T	Q	S	C	N	G	R	C	G	T	G	D	A	P	R	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン9	D	V	S	C	P	G	R	C	T	R	T	F	E	A	D	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン10	I	S	S	C	P	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン11	I	V	S	C	S	G	R	C	A	H	D	S	A	S	Q	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン12	A	V	S	C	K	G	R	C	T	R	T	T	H	L	T	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン13	T	S	F	C	L	G	R	C	G	R	K	T	T	M	H	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン14	T	A	S	C	T	G	R	C	P	H	P	V	R	G	P	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン15	I	V	S	C	S	G	R	C	A	H	D	S	A	S	Q	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン16	N	K	S	C	L	G	R	C	A	P	G	S	I	S	A	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン17	V	A	S	C	V	G	R	C	T	P	A	I	N	S	P	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン18	T	L	S	C	L	G	R	C	R	P	G	N	M	V	I	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン19	T	L	S	C	L	G	R	C	R	P	G	N	M	V	I	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K
クローン20	M	S	S	C	T	G	R	C	A	P	A	T	R	P	L	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K

ライブラリー区域1

メガプライマー突然変異発生またはクローニング後

マイクロタンパク	E	E	S	C	K	G	R	C	G	E	G	F	N	R	G	K	E	C	Q	C	D	E	L	C	K	Y	Y	Q	S	C	C	P	D	Y	E	S	V	C	K	P	K								
クローン21	L	T	S	S	C	P	G	R	C	R	G	Q	P	L	P	P	K	E	C	Q	C	D	E	L	C	R	P	S	T	S	P	C	C	L	D	F	E	E	I	C	E	P	R	L					
クローン22	T	S	S	F	C	P	L	G	R	C	R	G	K	T	L	M	H	K	E	C	Q	C	D	E	L	C	K	A	A	S	S	C	C	L	D	F	E	H	L	C	E	P	R	L					
クローン23	L	T	S	S	C	P	P	G	R	C	G	R	G	A	P	H	T	K	E	C	Q	C	D	E	L	C	S	P	S	L	S	C	C	L	D	F	E	A	N	L	C	E	P	R	L				
クローン24	I	T	S	S	C	P	T	G	R	C	G	A	P	H	T	K	E	C	Q	C	D	E	L	C	R	G	C	P	P	H	T	G	C	C	L	D	F	E	A	A	C	E	P	R	L				
クローン25	T	A	S	S	C	T	L	G	R	C	R	P	T	G	N	V	I	K	E	C	Q	C	D	E	L	C	E	F	H	H	Q	C	C	L	D	F	E	A	P	H	C	E	P	R	L				
クローン26	T	T	A	S	C	L	G	R	C	R	P	T	G	N	V	I	K	E	C	Q	C	D	E	L	C	N	P	C	H	Y	P	R	T	C	C	L	D	F	E	A	A	C	E	P	R	L			
クローン27	A	L	V	S	C	R	L	G	R	C	T	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	P	A	C	Q	L	C	C	S	D	F	E	A	A	C	E	P	R	L				
クローン28	T	I	S	F	C	L	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	T	A	C	S	H	H	A	T	C	C	S	D	F	E	A	A	C	E	P	R	L			
クローン29	I	I	S	S	C	P	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	E	T	C	G	S	T	R	Q	C	C	L	D	F	E	A	A	C	E	P	R	L			
クローン30	I	S	S	S	C	P	P	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	E	T	C	G	S	T	R	Q	C	C	L	D	F	E	A	A	C	E	P	R	L		
クローン31	A	N	K	S	C	L	G	R	C	T	A	R	T	N	S	I	K	E	C	Q	C	D	E	L	C	H	I	C	K	L	P	H	P	C	C	V	D	F	E	A	A	C	E	P	R	L			
クローン32	I	A	V	S	C	P	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	R	T	C	L	V	H	N	A	C	C	R	D	F	E	A	A	C	E	P	R	L			
クローン33	I	A	S	S	C	R	G	R	C	T	R	G	Q	P	L	P	K	E	C	Q	C	D	E	L	C	P	R	C	P	H	T	Q	R	C	C	P	D	F	E	A	A	C	E	P	R	L			
クローン34	A	L	V	S	C	R	P	G	R	C	T	R	G	Q	P	L	P	K	E	C	Q	C	D	E	L	C	K	P	C	V	I	S	S	P	C	C	N	D	F	E	A	A	C	E	P	R	L		
クローン35	L	L	G	P	C	P	G	R	C	R	K	P	T	I	V	E	G	K	E	C	Q	C	D	E	L	C	H	T	C	N	T	L	P	H	C	C	I	D	F	E	A	A	C	E	P	R	L		
クローン36	L	V	A	S	C	V	G	R	C	T	P	A	I	N	S	P	K	E	C	Q	C	D	E	L	C	G	Q	C	E	N	D	G	N	C	C	T	D	F	E	A	A	C	E	P	R	L			
クローン37	V	I	T	S	C	L	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	A	L	C	L	P	L	Q	S	C	C	E	D	F	E	A	A	C	E	P	R	L			
クローン38	V	I	T	S	F	C	L	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	S	N	C	K	L	I	P	C	C	H	D	F	E	A	A	C	E	P	R	L			
クローン39	I	L	L	S	C	P	G	R	C	P	T	R	F	V	L	V	K	E	C	Q	C	D	E	L	C	A	M	C	R	A	A	D	P	C	C	P	D	F	E	A	A	C	E	P	R	L			
クローン40	L	V	A	S	C	P	V	G	R	C	T	R	G	Q	P	L	P	K	E	C	Q	C	D	E	L	C	R	T	C	L	P	A	H	G	C	C	A	D	F	E	A	A	C	E	P	R	L		
クローン41	T	I	S	S	C	T	L	G	R	C	G	A	T	N	P	H	T	K	E	C	Q	C	D	E	L	C	H	H	C	K	T	F	H	A	C	C	H	D	F	E	A	A	C	E	P	R	L		
クローン42	I	L	L	S	C	P	G	R	C	P	T	R	F	V	L	V	K	E	C	Q	C	D	E	L	C	A	M	C	R	A	A	D	P	C	C	P	D	F	E	A	A	C	E	P	R	L			
クローン43	L	L	S	S	C	P	G	R	C	P	T	R	F	V	L	V	K	E	C	Q	C	D	E	L	C	A	M	C	R	A	A	D	P	C	C	P	D	F	E	A	A	C	E	P	R	L			
クローン44	L	V	A	S	C	P	V	G	R	C	T	R	G	Q	P	L	P	K	E	C	Q	C	D	E	L	C	P	P	C	R	S	N	L	R	C	K	V	E	E	N	P	R	L	L	K	V			
クローン45	I	M	S	S	C	T	G	R	C	A	P	A	T	R	P	L	K	E	C	Q	C	D	E	L	C	H	A	C	R	A	L	G	P	C	C	Q	D	F	E	A	A	C	E	P	R	L			
クローン46	L	L	S	S	C	P	G	R	C	P	T	R	F	V	L	V	K	E	C	Q	C	D	E	L	C	K	I	C	V	A	D	L	T	C	C	L	D	F	E	A	A	C	E	P	R	L			
クローン47	I	L	A	S	C	N																																											

った。

【0384】

(実施例15：DBPの複合性の変性)

複合性は、累積ジスルフィドスパンであり、これは、タンパク鎖におけるアミノ酸数で測定した、連結システイン間の累積距離に等しい。

【0385】

複合性は、架橋結合の程度、したがって、スカフォールドの剛性を表す測定値であり、複合性が高ければ高いほど、剛性も高い。剛性は、プロテアーゼ耐性の予測因子であるから、剛性はまた、免疫原性の有用な予測因子でもある。複合性がより高いことは、プロテアーゼ分解の減少、および免疫原性の低下を予測させる。

10

【0386】

複合性 = (C a - C b) + (C c - C d) + (C e - C f)

【0387】

【化23】

Ca-Cb Cc-Cd Ce-Cf Cg-Ch 複合性

1	2	3	4	2
1	3	2	4	4
1	4	2	3	4

20

1	6	2	5	3	4	9
1	4	2	5	3	6	9
1	6	2	4	3	5	9
1	5	2	6	3	4	9
1	5	2	4	3	6	9
1	4	2	6	3	5	9
1	2	3	4	5	6	3
1	2	3	5	4	6	5
1	2	3	6	4	5	5
1	6	2	3	4	5	7
1	4	2	3	5	6	5
1	5	2	3	4	6	7
1	3	2	6	4	5	7
1	3	2	4	5	6	5
1	3	2	5	4	6	7

30

1	2	3	4	5	6	7	8	4
---	---	---	---	---	---	---	---	---

40

(実施例16：反復モチーフを持たないスカフォールド)

トキシンファミリーのスーパーファミリー

1) uPAR / Ly6 / CD59 / 蛇トキシン受容体スーパーファミリー。ファミリー：Activin_recp; BAMBI; PLA2_inh; Toxin_1; UPAR_LY6を含む。

【0388】

2) さそりトキシン様ノッチンスーパーファミリーは、ファミリー、Toxin_2; Toxin_17; Gamma_thionin; Defensin_2; Toxin_

50

3 ; T o x i n _ 5 を含む。

【0389】

3) デフェンシン／ミオトキシン様スーパーファミリーは、ファミリー、B D S _ I _ I I ; D e f e n s i n _ 1 ; D e f e n s i n _ b e t a ; T o x i n _ 4 を含む。

【0390】

4) オメガトキシン様スーパーファミリーは、ファミリー、T o x i n _ 7 ; T o x i n _ 3 0 ; T o x i n _ 2 7 ; T o x i n _ 2 4 ; T o x i n _ 2 1 ; T o x i n _ 1 6 ; T o x i n _ 1 2 ; T o x i n _ 1 1 ; O m e g a _ t o x i n ; A l b u m i n _ I ; T o x i n _ 9 を含む。

【0391】

5) コノトキシンOースーパーファミリーは、同じ構造群に属するC o n u s ペプチドの3群から成る。これらの3群は、その薬理学的性質において異なる。カルシウムチャンネルを抑制するwーコノトキシン、電圧感受性を持つナトリウムチャンネルの不活性化速度を遅延するデルターコノトキシン、および、電圧感受性を持つナトリウム電流を阻止するm u Oーコノトキシンである。

【0392】

6) コノトキシンIースーパーファミリーは、トキシン19ファミリーのみを含む。

【0393】

7) コノトキシンTースーパーファミリーは、トキシン26ファミリーのみを含む。

【0394】

(個別のトキシンファミリー)

P F 0 0 0 8 7 : トキシン1

蛇トキシン。有毒なニューロトキシンおよびサイトトキシンから成るファミリー。構造は小型で、ジスルフィドに富み、ほぼ全てベータシートである。図61を参照されたい。

【0395】

【化24】

- 1) Cxxxx(xxxx)xxxCxxxxxCxxx(xxx)C(xx)xxxxxxxxCxxxC
- 2) Cxxxx(xxxx)xxxCxxxxxCYxkx(wf)(xx)C(xx)xxxxxxxxGCxxxC

P F 0 0 4 5 1 : トキシン2

「蛇トキシン短小」。蛇毒は、哺乳類、昆虫、および甲殻類に有毒な、様々のペプチドを含む。これらのペプチドの中に、カルシウム活性化カリウムチャンネルを阻害する短いトキシン(30から40残基)のファミリーがある。図55を参照されたい。空間配置は1-4 2-6 3-5.

【0396】

【化25】

- 1) CxxxxCxxxCxxxxxxxxxCxxxCx
- 2) CxxxxCxxxCkxxxxxxxxgKCxxxKCxC

P F 0 0 5 3 7 : トキシン3

このファミリーは、ニューロトキシンと植物デフェンシンの両方を含む(F. M. A s s a d i - P o r t e r , ら、(2000) Arch. Biochem Biophys , 376: 259-65)。マスタードトリプシン阻害剤、MTI-2は植物デフェンシンである。これは、トリプシンの強力な阻害剤である。MTI-2は、鱗翅類昆虫にとっては有毒である。さそりトキシン(ニューロトキシン)は、ナトリウムチャンネルに結合して、該チャンネルの活性化機構を抑制し、神経伝達を阻止する。図22を参照されたい。空間配置は1-8 2-5 3-6 4-7である。

【0397】

10

20

30

40

【化 2 6】

- 1) C(xxx)x(xx)xxxxCxxCxx(xx)xxCxxCxx(x)xxxxCxxxx(xx)xxCx
 2) C(xxx)Y(xx)xxxxCxxCxx(xx)xxCxxCxx(x)xxGxCxxxx(xx)xxC(W,Y)C

P F 0 0 7 0 6 : トキシン 4

いそぎんちゃくニューロトキシン。いそぎんちゃくは、関連構造および機能を持つ、多くの、異なるニューロトキシンを生産する。このファミリーに属するタンパクはニューロトキシンを含み、その内には、カリトキシンおよびアントプレウリンを含むいくつかがある。ニューロトキシンは、ナトリウムチャンネルに特異的に結合し、シグナル伝達中その不活性化を遅らせ、そのため哺乳類心筋収縮の強力な刺激をもたらす。カリトキシン 1 は、甲殻類の神経筋標本に見出された。該標本において、カリトキシン 1 は、伝達物質の放出を増し、軸索の発火を引き起こす。このタンパクには、三つのジスルフィド結合が存在する。このファミリーは、デフェンシン／ミオトキシン様スーパーファミリー族のメンバーである。この族は、下記の P f a m メンバーを含む：B D S _ _ I _ _ I I ; D e f e n s i n _ _ 1 ; D e f e n s i n _ _ b e t a ; T o x i n _ _ 4 . いそぎんちゃくは、関連構造および機能を持つ、多くの、異なるニューロトキシンを生産する。このファミリーに属するタンパクはニューロトキシンを含み、その内には、カリトキシンおよびアントプレウリンを含むいくつかがある。ニューロトキシンは、ナトリウムチャンネルに特異的に結合し、シグナル伝達中その不活性化を遅らせ、そのため哺乳類心筋収縮の強力な刺激をもたらす。カリトキシン 1 は、甲殻類の神経筋標本に見出された。該標本において、カリトキシン 1 は、伝達物質の放出を増し、軸索の発火を引き起こす。このタンパクには、三つのジスルフィド結合が存在する。25 の既知のファミリーメンバーがある。空間配置は 1 - 5 2 - 4 3 - 6 である。図 8 7。

【0 3 9 8】

【化 2 7】

- 1) CxCxxxxxxxxxxxxxxxxxx(xx)xxxxC(xx)xxxxxxxxCxxxxxxxxCC
 2) CxCxxxxPxxxxxxxxGxx(xx)xxxxC(xx)xxxWxxCxxxxxxxxCC

P F 0 5 2 9 4 : トキシン 5

さそり短小トキシン。図 4 6 .

P F 0 5 4 5 3 : トキシン 6

図 9 0 . このファミリーは、B u t h u s m a r t e n s i i K a r s c h さそりから単離されたトキシン様ペプチドから成る。前駆体は、シグナルペプチド候補 2 8 残基および 1 余分残基、成熟ペプチド 3 1 残基、およびアミド化 C - 末端を有する、6 0 アミノ酸残基から成る。ペプチドは、他の、さそり K + チャンネルトキシンと緊密な相同性を共有し、共通の、三次元フォールド、システイン - 安定化アルファベータ (C S アルファベータ) モチーフを提示する。このファミリーは、その被害者の、小さい、コンダクタン型、カルシウム活性化カリウムイオンチャンネルを阻止することによって活動する。空間配置は 1 - 4 2 - 5 3 - 6 である。モチーフは：

【0 3 9 9】

【化 2 8】

CxxCxxxCxxxxxx(xx)C(xx)xxxxCx

P F 0 5 9 8 0 : トキシン 7

このファミリーは、蜘蛛の、いくつかの短小ニューロトキシン・タンパクから成り、ジョウゴグモのトキシンの多くのものを含んでいる (W. S. S k i n n e r , ら、(1989) J B i o l C h e m , 2 6 4 : 2 1 5 0 - 5 5) 。図 6 4 を参照。

【0 4 0 0】

空間配置は 1 - 4 2 - 5 3 - 8 6 - 7 である。

【0 4 0 1】

【化 2 9】

- 1) CxxxxxCxxxxxxxxCCxxxxxCxxxxxCxC
 2) CxxxxxCxxWxxxxCCgxxYCxCxxxpxCxC

P F 0 7 3 6 5 : トキシン 8

アルファコノトキシンおよび前駆体。このファミリーは、イモ貝種のメンバーから得られる、いくつかのアルファコノトキシン前駆体タンパクから成る。アルファコノトキシンは、魚を狩るイモ貝の毒から得られる、ニコチン性アセチルコリン受容体 (n A C h R) を阻止する、小型ペプチドニューロトキシンである。図 7 2。

【0 4 0 2】

P F 0 0 0 9 5 : トキシン 9

蜘蛛ニューロトキシンの、このファミリーは、カルシウムイオンチャンネル阻害剤と考えられている。

【0 4 0 3】

図 6 3 を参照。空間配置は 1 - 4 2 - 5 3 - 8 6 - 7 である。

【0 4 0 4】

【化 3 0】

- 1) Cxx(x)xxxxCxxxxCCxxx(x)xCxCxxxxxCxC
 2) Cxx(x)yxxxCxxgxxCCxx(x)xCxCxxxxxCxC

P F 0 7 4 7 3 : トキシン 1 1

このファミリーは、いくつかの痙攣性ペプチド g m 9 a 配列から成る (M. B. L i r a z a n, ら、(2000) B i o c h e m i s t r y, 39:1583-8)。図 2 7 参照。D B P : 1 - 5 2 - 4 3 - 6

【0 4 0 5】

【化 3 1】

モチーフ: CxxxxCxxxxCxxxCxC

P F 0 7 7 4 0 : トキシン 1 2

H a T x 1 は、チリのタランチュラ毒から単離された 3 5 アミノ酸ペプチドトキシンである。このものは、d r k 1 電圧依存性 K (+) チャンネルを、孔を塞ぐのではなく、ゲートのエナージェティクスを変えることによって阻害する (H. T a k a h a s h i, ら、(2000) J M o l B i o l, 297:771-80)。図 5 0 参照。

【0 4 0 6】

空間配置は 1 - 4 2 - 5 3 - 6 である。モチーフは、C x x x x x x C x x x x x (x) C C x x x x C x x x (x x x) x (x x) x x C である。

【0 4 0 7】

P F 0 7 8 2 2 : トキシン 1 3

このファミリーのメンバーは、海産ワーム *C e r e b r a t u l u s l a c t e u s* によって生産される甲殻類選択的ニューロトキシンである、ニューロトキシン B - I V によく似る。この高度に陽イオン性のペプチドは、約 5 5 残基であり、ヘアピン構造の明瞭な形状を持つループによって接続される、2 本の逆平行螺旋を形成するように配される。ヘアピンの側枝は、四つのジスルフィド結合によって連結される。活性に重要であると特定された三つの残基、すなわち、A r g - 1 7、- 2 5、および - 3 4 は、分子の同じ面に見出されるが、一方、活性に重要なもう一つの残基、T r p 3 0 は、反対側にある。このタンパクの活動方式は、完全には理解されていないが、電圧依存性ナトリウムチャンネルに対し、おそらく、これらのタンパクの、まだ解明されていない部位に結合することによって作用すると考えられる。その相互作用部位も比較的特異性が低く、例えば、負に帯電する膜脂質と相互作用を持つと考えられる。図 6 5 参照。

【0 4 0 8】

P F 0 7 8 2 9 : トキシン 1 4

10

20

30

40

50

アルファ-AコノトキシンPIVAは、魚食性イモ貝*Conus purpurascens*によって生産される毒に認められる、主要な麻痺性トキシンである。このペプチドは、ニコチン性アセチルコリン受容体のアセチルコリン結合部位をブロックすることによって作動する(K. J. Nielsen, ら、(2002) *J Biol Chem*, 277: 27247-55)。図66参照。

【0409】

【化32】

モチーフ1: CCxxxxxxCxxCxCx(x)xxxxxC, Motif 2: CCgxxpxxxClpCxCx(x)xxpxxC

PF07945: トキシン16

10

ヤヌスアトラコトキシン・ファミリー。このファミリーは、蜘蛛*Hadronyche versuta*によって分泌される3種のペプチドを含む。これらは、筋肉のアセチルコリン受容体、または、他の、無脊椎動物に存在するアセチルコリン受容体サブタイプに拮抗することによって働く、昆虫選択的、興奮性ニューロトキシンである。ヤヌスアトラコトキシン-Hv1cは組織されて、ジスルフィド富裕小球状コア(残基3-19)、およびベーターヘアピン(残基20-34)となる。4ジスルフィド架橋があり、その内の一つは、隣接ジスルフィド架橋である。これは、構造の維持には重要ではないが、殺昆虫活性には重要であることが知られる。ファミリーには三つのメンバーが知られる。空間配置は1-6 2-7 3-4 5-8. 図91参照。

【0410】

20

【化33】

1) CxxxxxxCxxCCxCxCxxxxCxxxxxxxxxC

2) CxgxxpCxxCCpCCpgxxCxxxxxgxyC

PF08086: トキシン17

このファミリーは、さそりによって分泌されるトキシンであるエルゴトキシンペプチドから成る。エルゴトキシンは、K⁺チャンネルの機能を阻止することが可能である。さそり毒から、100を超えるエルゴトキシンが見出されており、それらは、その一次構造に基づいて、三つのサブファミリーに分類される(K. Frenal, ら、(2004) *Proteins*, 56: 367-75)。

30

【0411】

ファミリーには25の既知のメンバーがある。空間配置は1-4 2-6 3-7 5-8. 図60参照。

【0412】

【化34】

1) CxxxxxCxxxxxxxxCxxCCxxxxxxxxCxxxxCxC

2) dldxCxDxxxCxygxyxCxxCxxxgxxgxCxxxxCxC

PF08087: トキシン18

コノトキシンO-スーパーファミリー。このファミリーは、コノトキシンO-スーパーファミリーのメンバーから成る。コノトキシンO-スーパーファミリーは、同じ構造群に属する*Conus* ペプチドの3群から成る。これらの3群は、その薬理的性質において異なる。カルシウムチャンネルを抑制するw-コノトキシン、電圧感受性を持つナトリウムチャンネルの不活性化速度を遅延するデルターコノトキシン、および、電圧感受性を持つナトリウム電流を阻止するmuO-コノトキシンである。図31参照。

40

【0413】

【化35】

モチーフ1: CxxxxxCxxxxCCx(xx)xxCxxxxxC,

モチーフ2: CxxxgxxCxxxxCCx(xx)gxCxxxfxC

50

P F 0 8 0 8 8 : トキシシン 1 9

コノトキシシン I ースーパーファミリー。図 6 参照。このファミリーは、コノトキシシンの I ースーパーファミリーから成る。これは、いくつかの *Conus* 種の毒における新規クラスのペプチドである。これらのトキシシンは、四つのジスルフィドによって特徴づけられ、神経細胞のイオンチャンネルを抑制または修飾する。I ースーパーファミリーは、イモ貝の 5 または 6 の主要クラスにおいて認められており、さらに多くの生物種に見出される可能性がある。

【 0 4 1 4 】

P F 0 8 0 8 9 : トキシシン 9

Huwentoxinファミリー。このファミリーは、蜘蛛によって分泌されるトキシシンの、huwentoxin-I I (HWTX-I I) ファミリーから成る。これらのトキシシンは、鳥蜘蛛 *Selenocosmia huwena* Wang から分泌される毒の中に見出される。HWTX-I I は、他の huwentoxin に見られる ICKモチーフとは異なる新規スカフォールドを採用する。HWTX-I I は、三つのジスルフィド架橋に関与する六つのシステインを含む、37 アミノ酸残基から成る。図 5 参照。

10

【 0 4 1 5 】

P F 0 8 0 9 1 : トキシシン 2 1

このファミリーは、オメガトキシシン様族のメンバーである。このファミリーは、蜘蛛毒から単離された、殺昆虫ペプチドから成る。図 5 8 参照。ファミリーには 4 の既知のメンバーがある。空間配置は未知である。構造データは無い。

20

【 0 4 1 6 】

【 化 3 6 】

1) CxxxxxCxxxxCCxxxCxxxxxCxxxxxCxxxC

2) CxxxxPCxxxxCCxxxWxCxxxxxCskxC

P F 0 8 0 9 2 : トキシシン 2 2

図 4 参照。このファミリーは、Hexathelidae 蜘蛛の毒から単離された Magi ペプチドトキシシンから成る (Magi 1、2、および 5)。これらの、殺昆虫ペプチドトキシシンは、ナトリウムチャンネルに結合し、鱗翅目幼虫に注入されると弛緩麻痺を誘発する。しかしながら、これらのペプチドは、マウスに対しては、20 pmol / g で脳内に注入しても、毒性を持たない。

30

【 0 4 1 7 】

P F 0 8 0 9 3 : トキシシン 2 3

図 3 参照。このファミリーは、Hexathelidae 蜘蛛の毒に見出される有毒ペプチド (Magi 5) から成る。Magi 5 は、哺乳類のナトリウムチャンネルの部位 4 に対し結合アフィニティを持つ、最初の蜘蛛トキシシンであり、このトキシシンは、幼虫に対し殺昆虫作用を持ち、幼虫に注入されると麻痺を起こす。

【 0 4 1 8 】

P F 0 8 0 9 4 : トキシシン 2 4

コノトキシシン TVIIA / GS ファミリー。このファミリーは、イモ貝 *Conus tulipa* および *Conus geographus* の毒から単離されたコノトキシシンから成る。*Conus tulipa* から単離されたコノトキシシン TVIIA は、他の、よく解明された、薬理的ペプチドクラスとはほとんど相同性を示さないが、*Conus geographus* から得られたペプチド、コノトキシシン GS に対しては類似性を示す。これらのペプチドは両方とも、骨格筋のナトリウムチャンネルをブロックし、いくつかの生化学的特性を共有し、4 ループコノトキシシンから成る独立サブグループを代表する (J. M. Hill, ら、(2000) Eur J Biochem, 267: 4642-8)。図 2 8 参照。

40

【 0 4 1 9 】

【化 3 7】

- 1) CxxxxxxCxxxCCxxxxCxxxxxxxC
 2) CxGxxxxCPPxCCxGxxCxxGxxxxC

P F 0 8 0 9 5 : トキシシン 2 5

H e f u トキシシンファミリー。このファミリーは、さそりH e t e r o m e t r u s f u l v i p e s の毒に見出されるh u f e トキシシンから成る。これらのトキシシン、カップーh e f u トキシシン1およびカップーh e f u トキシシン2は、既知のいずれのトキシシンにたいしても相同性を示さない。h e f u トキシシンは、カリウムチャンネルトキシシンであり、1-4 2-3 空間配置を示す。図 1 7 3 参照。

10

【0 4 2 0】

P F 0 8 0 9 7 : トキシシン 2 6

コノトキシシンTスーパーファミリー。図 2 参照。このファミリーは、コノトキシシンのTスーパーファミリーから成る。五つのC o n u s 種から得られた、異なる八つのTスーパーファミリーが特定されている。これらのペプチドは、共通シグナル配列、およびシステイン残基の保存配置を共有する。Tスーパーファミリーは、全ての主要な採食型C o n u s の毒管において発現することが見出されており、これは、Tスーパーファミリーが、500の異なるC o n u s 種において広く分布する、ペプチドの大きな、多様な群であることを示唆する。

【0 4 2 1】

20

P F 0 8 0 9 9 : トキシシン 2 7

さそりカルシンファミリー。図 1 参照。このファミリーは、さそりトキシシンのカルシンファミリーから成る。カルシンファミリーは、M a u r o c a l c i n e および I m p e r o t o x i n から成る。これらのトキシシンは、骨格筋の、リアノジン感受性カルシウムチャンネルの強力な作用因子であることが示されている。したがって、これらのトキシシンは、ジヒドロピリジン受容体／リアノジン受容体との相互作用実験に有用である。

【0 4 2 2】

P F 0 8 1 1 6 : トキシシン 2 9

このファミリーは、ブラジルのP h o n e u t r i a n i g r i v e n t e r の毒に見出されたP h T x 殺昆虫ニューロトキシシンから成る。P h o n e u t r i a n i g r i v e n t e の毒は、ある範囲の生物学的作用を持つ、30-140アミノ酸から成る、数多くの神経毒ポリペプチドを含む。これらのニューロトキシシンのあるものは、マウスに対し、脳室内注入後、致命的であるが、他のものは、双翅目およびD i c t y o p t e r a 目の昆虫に対してはきわめて有毒であるが、マウスに対する毒作用ははるかに低い。図 7 参照。

30

【0 4 2 3】

P F 0 8 1 1 7 : トキシシン 3 0

P t u ファミリーとも呼ばれる。このファミリーは、サシガメの唾液から単離された毒性ペプチドから成る。この唾液は、餌食を不動化するか、または、それを消化するために、その虫によって用いられる、タンパクの複雑な混合物を含む。このタンパクの一つ（P t u 1）は、精製されて、N-型カルシウムチャンネルを可逆的にブロックするが、B H K 細胞において発現されるL-およびP/Q型カルシウムチャンネルに対する特異性はより低いことが示されている。

40

【0 4 2 4】

空間配置 1-4 2-5 3-6. 図 7 9 参照。

【0 4 2 5】

【化 3 8】

- 1) CxxxxxxCxxxxxxCCxxxxCxxxxxxxC
 2) CxxgxxxCxgxxxCCxxxxCxxyanxC

50

P F 0 8 1 1 9 : トキシシン 3 1

このファミリーは、酸性の、アルファ K T x 短鎖さそりトキシシンから成る。パラプトキシシンと名づけられるこれらのトキシシンは、電圧依存性 K チャンネルをブロックし、きわめて低い p I 値を有する。さらに、これらは、きわめて重要な、孔閉鎖性リシンを欠く。さらに、二連子の、第 2 の重要な残基、疎水性残基 (P h e または T y r) も見られない。図 8 参照。

【 0 4 2 6 】

P F 0 8 1 2 0 : トキシシン 3 2

図 9 参照。このファミリーは、インド赤さそり (M e s o b u t h u s t a m u l u s) の毒に見出される t a m u l u s トキシシンから成る。t a m a l u s トキシシンは、他のさそり毒トキシシンとはまったく類似点を共有しないが、その、六つのシステイン残基の位置は、それが、同じ構造的スカフォールドを共有することを示唆する。t a m u l u s トキシシンは、カリウムチャンネル遮断剤として活動する。

【 0 4 2 7 】

【 化 3 9 】

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Abstract&list_uids=11361010

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Abstract&list_uids=11361010

P F 0 8 3 9 6 : トキシシン 3 4

蜘蛛トキシシンオメガアゴトキシシン / T x 1 ファミリー。T x 1 ファミリー、致命的蜘蛛ニューロトキシシンは、マウスにおいて興奮性症状を誘発する。図 1 0 参照。

【 0 4 2 8 】

P F 0 1 0 3 3 : ソマトメジン

図 1 4 参照。ソマトメジン B、機能未知の血清因子は、細胞基質の接着タンパク、ビトロネクチンの N-末端からタンパク分解によって得られる、小型の、システイン富裕ペプチドである。S M B ドメインは、四つのジスルフィド結合に配される、八つの C y s 残基を含む (Y . K a m i k u b o , ら、(2 0 0 4) B i o c h e m i s t r y , 4 3 : 6 5 1 9 - 3 4) 。活性な S M B ドメインは、C y s 2 5 - C y s 3 1 ジスルフィド結合が保存される限り、相当なジスルフィド結合の異種性または変動性を許容することが示唆されている。S M B ドメインの三次元構造はきわめてコンパクトであり、かつ、ジスルフィド結合は、共有結合コアを形成するドメインの中心にバックされる。このタンパクは、チオレドキシシンの C-末端ドメインと共に可溶性融合タンパクとして発現される。

【 0 4 2 9 】

【 化 4 0 】

1) Cxx(x)xCxxxxxxxxxCxCxxxCxxxxCCxxxxxxxxC

2) Cxx(x)rCxxxxxxxxxCxCxxxCxxxxCCxDxxxxC

3) Cxx(x)RCxxxxxxxxxCxCxxxCxxxxCCxd[yf]xxxC

1 - 2 3 - 4 5 - 6 7 - 8 の空間配置が記載されているが、他の異性形も可能であり、N M R 構造計算とも一致する。

【 0 4 3 0 】

P F 0 0 0 8 7、P F 0 0 0 2 1 : 3 本指トキシシンファミリー

図 1 4 - 1 8 参照。有毒ニューロトキシシンおよびサイトトキシシンから成るファミリー。構造は、小型で、ジスルフィド富裕、ほとんど全てがベータシート。このファミリーは、u P A R / L y 6 / C D 5 9 / 蛇トキシシン-受容体スーパーファミリー族である。この族は、下記の P f a m メンバー : A c t i v i t i n _ r e c p ; B A M B I ; P L A 2 _ i n h ; T o x i n _ 1 ; U P A R _ L Y 6 を含む。

【 0 4 3 1 】

好ましいライブラリー戦略は、C y s 1 - C y s 2、C y s 3 - C y s 4、および C y s 5 - C y s 6 の間にある、三つの最長ループをランダム化することである。二つの異な

10

20

30

40

50

る設計戦略が用いられる。1) ジスルフィドコアはそのまま、三つのループだけを突然変異発生させる、2) ジスルフィドコアにおける突然変異発生を可能とし、それによって、より高い多様性のループ配置を生成する。もっとも保存されるシステイン間隔は、位置 $n = 0$ および $n = 7$ である（‘ $n 6$ ’は、C 6とC 7の間と定義され、‘ $n 7$ ’は、C 7とC 8の間と定義される）。残りのCDPを評価するのにこの情報が使用される。もっとも一般的CDPは、69メンバーについて、10、6、16、3、10、0、4である。

【0432】

【化41】

1) Cxxxxxxxxxxx(xxx)Cxxxx(xx)Cxxxxxxxxxxxxxx(x)xxxxCx(xx)CxxxxxxxxxxCCxxxxC

2) Cyxxxxxxxxxxx(xxx)Cpxgx(xx)Cyxkx(wf)xxxxxx(x)xxxxGCx(xt)CPxxxxxxxxxxCCx(ts)DxC

10

PF01607, PF00187: キチン結合タンパク

二つの、異なるシステイン富裕キチン結合ファミリーがある (Z. Shen, ら、(1998) J Biol Chem, 273:17665-70); T. Suetake, ら、(2000) J Biol Chem, 275:17929-32; T. Suetake, ら、(2002) Protein Eng, 15:763-9)。PF00187は、菌類および植物に認められ、小麦胚アグルチニンを含む。Heveinは、四つのジスルフィド結合を含む、前駆体メンバーである。このファミリーは、きわめて保存されたシステイン位置、および空間配置1-4 2-5 3-6 7-8を有する、382の、既知のファミリーメンバーを含む。ライブラリー設計においてスカフォールドとして使用した場合の、このファミリーの利点として、第1システインのN-末端位置、および最終システインのC-末端位置におけるアミノ酸の少数であること (<3) が挙げられる。個々のシステイン間の距離は、10未満であり、ドメインは、ジスルフィド結合に富む (四つのジスルフィド結合について約50アミノ酸)。もっとも一般的DBPは、1-4 2-5 3-6の空間配置である。ドメインは、天然では反復列の中に見出される。

20

【0433】

PF01607は、Peritrophinドメインとも呼ばれ、細胞外基質タンパクの一部として動物および昆虫に認められる。このドメインはまた、小型ペプチドタキサイチンにも出現する。タキサイチンおよびhevein (PF00187)の構造比較から、構造的類似性が明らかになった (整列を参照)。タキサイチンは、五つのジスルフィド結合をふくむが、このファミリーのメンバーは、通常、3SSを含む (式参照)。タキサイチンの、3署名SSは、1-3 2-6 4-5の空間配置を示す。1075の既知のファミリーメンバーがある。システイン位置は高度に保存される。第1システインのN-末端、および最終システインのC-末端におけるアミノ酸数は多くない (<3)。

30

【0434】

図19-21参照。

【0435】

PF00187 キチン結合タンパク

【0436】

【化42】

CxxxxxxxxCxxxxCCxxxxCxxxxxxxxCxxxCxxxC

CgxqxxxxxxxxCxxxxCCsxxGxCgxxxxxCxxxCxxxC

40

PF01607 キチン結合ドメイン:

【0437】

【化43】

1) Cxxx(x)xxxxxxxx(x)xxxC(x)xxxxxCxxxxxxxxxCxxxxxxxxxCxxxxxxxx

2) Cxxx(x)xxgxxx(x)xxxC(x)xx[yf]xxCxxxxxxxxCxxgxxxxxxxxxxxCxxxxxxxx

PF01826: トリプシン阻害剤

50

このファミリーは、トリプシン阻害剤の外、多くの細胞外ドメインに見られるドメインを含む (N. D. Rawlings, ら、(2004) Biochem J, 378: 705-16)。このドメインは通常、五つのジスルフィド結合を形成する、10個のシステイン残基を含む。DBPは、1-7 2-6 3-5 4-10 8-9である。414のファミリーメンバーが既知である。システイン位置は高度に保存される。図23参照。

【0438】

【化44】

CxxxxxxxxCxxxGxxxGxxxx(xxxxx)xxxCx(xxxxxxx)xxGxxx(x)CxGxxxxxxxxxx(xx)xGxxxxxC

10

PF02428: ジャがいもタンパク阻害剤

このファミリーは、遺伝子レベルの反復列に認められる。このタンパクは、大型の前駆体タンパクとして合成される。タンパク分解酵素による切断は、反復列の間ではなく、反復列の中で起こり、成熟ミクロタンパクを生成する (E. Barta, ら、(2002) Trends Genet, 18: 600-3) (N. Antcheva, ら、(2001) Protein Sci, 10: 2280-90)。

【0439】

大型前駆体タンパクは合成されるが、前駆体のジスルフィドの空間配置は不明である。

【0440】

反復単位は発現され、そのNMR構造は解決されている。フォールドは、成熟ミクロタンパクと同様である。これは、円形組み合わせが起こっていること、この単位は祖形であることを示唆する。これは、この反復単位に一致する、円形組み合わせタンパクの発見によって支持される。リンカーまたはプロテアーゼ部位 (EEKKN) が、この祖形の構造の中に無秩序ループとして存在する。図24参照。

20

【0441】

【化45】

1) CxxxGxxxxxxxxCxxxxxxxx(x)xxxxxCxxCCxxxxxCxxxxxxxxxxxC

2) CxxxGxxxxxxxxCPxxxx(x)xxxxxCxxCCxxxxGCxxxxxxxxGxxxC

タンパク分解酵素処理によって、成熟ミクロタンパクの配列は、上に示した式とは異なる。

30

【0442】

2C2CC5C10C11C3C8C2 (成熟式-タンパクレベル)

3C3C8C12C2CC5C10C2 (反復式-遺伝子レベル)

PF00304: ガンマチオニン

この熟成形では、これらの小型の植物タンパクは、一般に、約45から50アミノ酸残基から成る。ガンマープロチオニンのフォールド構造は、境界の明瞭な、3本の、逆平行シートおよび短い螺旋で特徴づけられる。三つのジスルフィド架橋が、螺旋とシートの間の疎水性コアに配され、システイン安定化螺旋モチーフを形成する (P. B. Pellegrini, ら、(2005) Int J Biochem Cell Biol, 37: 2239-53)。この構造は、さそりのトキシンおよび昆虫のデフェンシンに類似する (C. Bloch, Jr., ら、(1998) Proteins, 32: 334-49)。

40

【0443】

このドメインは、約50アミノ酸当たり4ジスルフィド結合という高いジスルフィド密度を示し、1-8 2-5 3-6 4-7の空間配置を持つ。個々のシステイン間のシステイン間隔は、10よりは小さく、したがってライブラリー設計には好ましい。このファミリーの異なるメンバーの間で、システイン位置は高度に保存される。図25参照。

【0444】

PF00304-ガンマーチオニン

50

【0445】

【化46】

モチーフ1: CxxxxxxxxCxxxxCxxxCxxxxxxxx(x)xxxCxx(x)xxxxCxCCCC

モチーフ2: CxxxSxxFxxGxCxxxxxCxxxCCCCxx(x)xGxCxx(x)xxxxCxCCCC

PF02950: オメガーコノトキシン

コノトキシンは、イオンチャンネルをブロックする、小型のニューロトキシンである。オメガーコノトキシンは、シナプス前膜に作用し、カルシウムチャンネルをブロックする (W. R. Gray, ら、(1988) *Annu Rev Biochem*, 57: 665-700)。ドメインは、約24アミノ酸当たり三つのジスルフィド結合という高いジスルフィド密度を示す。380を超える既知のファミリーメンバーがある。個々のシステイン間のシステイン間隔は10よりも小さく、したがってライブラリー設計には好ましい。このファミリーの異なるメンバーの間で、システイン位置は高度に保存される。ファミリーは、1-4 2-5 3-6の空間配置を有数る。

10

【0446】

図26参照。

【0447】

【化47】

モチーフ: C(xx)xxxxxCxx(xx)xCx(xxx)xxCC

20

ジコノチドは、FDA承認の25AAコノトキシン「プリアルト」である。ジコノチドは、>7000患者に試験され、非免疫原性であった(<1%出現率)。これにより、これは、ヒトに使用される新規結合タンパクとして有望なスカフォールドと期待される。配列、および1-4 2-5 3-6DBPを図12に示す。

【0448】

PF05374: Muーコノトキシン

Muーコノトキシンは、電圧依存性ナトリウムチャンネルのペプチド阻害剤である (K. J. Nielsen, ら、(2002) *J Biol Chem*, 277: 27247-55)。図29参照。DBP: 1-4 2-5 3-6

【0449】

30

【化48】

モチーフ1: CCxxxxxCxxxxCxxxxCC モチーフ2: CCxxpxxCxxxxCxPxxCC

PF02822: アンチスタシン

ペプチドプロテイナーゼ阻害剤は、タンパクの中の単ドメインタンパク、または、単一または複数ドメインとして認められる。これらは、それぞれ、単純または複合阻害剤と呼ばれる (R. Lapatto, ら、(1997) *Embo J*, 16: 5151-61)。多くの場合、これらは、より大型の前駆体タンパクの一部として、プレペプチドとして、または、不活性ペプチダーゼまたはチモーゲンと関連するN-末端ドメインとして合成される。Pfam定義は、1-4 2-5 3-6のDBPを有する、僅かに六つのシステインしか含まない。しかしながら、このファミリーの多くのメンバー (1bx7、1hia) は、さらに二つのN-末端ジスルフィドを含む。したがって、このファミリーは、N-末端において延長することが可能である。

40

【0450】

このドメインは、39-54アミノ酸当たり3-5ジスルフィドという高いジスルフィド密度、および、1-3 2-4 5-8 6-9 7-10の空間配置を示す。個々のシステイン間のシステイン間隔は10よりも小さく、したがってライブラリー設計には好ましい。このファミリーの異なるメンバーの間で、システイン位置は高度に保存される。図32参照。

【0451】

50

このファミリーのメンバーは、きわめて親水性が高く、これは、ライブラリー設計には好ましい（非特異的結合が低く、T-細胞エпитープは少数である）。例えば、ヒルスタシンは、合計して、僅かに6個の疎水性残基しか含まない。結晶構造は、ほとんど二次構造要素を欠如することを示す。これを、5SSで可能なジスルフィド異性形の多数であることと組み合わせると、ライブラリー設計のためのスカフォールドとしてきわめて有用となる。

【0452】

システイン位置は、5ジスルフィドについては高度に保存される：C4C5C6C1C4C4C10C5C1C

PF02822-アンチスタチン

10

【0453】

【化49】

- 1) CxxxxCxxxxCxxxxC(x)xxxCxxxxxxC(xxx)xCx
- 2) CxxxxCxxxxCxxxxC(x)xxxGxxdxxgC(xxx)xCx
- 3) CxxxxCxxxxCxxxxC(x)xxxGpyGxxdxxgC(xxx)xCx

N-末端の四つのシステイン残基を欠いた短小バージョン。

【0454】

【化50】

- 1) CxxxxC(x)xxxCxxxxxxC(xxx)xCx
- 2) CxxxxC(x)xxxGxxdxxgC(xxx)xCx
- 3) CxxxxC(x)xxxGpyGxxdxxgC(xxx)xCx

20

PF05039：アグーチ関連

図33参照。アグーチタンパクは、先端下に黄色バンドを持つ黒色毛を発生するマウス毛嚢における色素形成を調整する。きわめて相同性の高いタンパク、アグーチシグナルタンパク（ASIP）がヒトにおいて存在し、脂肪組織において高レベルで発現されるが、ここで該タンパクは、エネルギーホメオスターシスおよび、おそらくヒトの色素形成においても関与すると考えられる（J. C. McNulty, ら、(2001) *Biochemistry*, 40:15520-7; J. Voisey, ら、(2002) *Pigment Cell Res.* 15:10-8）。

30

【0455】

Cys5とCys10の間のジスルフィド結合は、構造と機能には不要である。除去すると、DBPは1-4 2-5 3-8 6-7となる。最初の三つのジスルフィド結合は、署名のシステインノットモチーフを形成する。受容体結合部位は、Cys7およびCys8の間のRFFモチーフ、および最初の16アミノ酸によって形成されるループを含む。C末端は、無秩序であり、除去してもよい（Cys1とCys10は、Pfam式には存在しないことに注意）。

【0456】

下記の式は、ライブラリー設計に好ましい：PF05039-アグーチ：

40

【0457】

【化51】

- 1) CxxxxCxxxxCCxxCxxCxxxxCxxxxxxC
- 2) CxxxxSCxxxxCCDPCxxCxxCRFFxxxCRxxxxxxC
- 3) CxxxxSCxxGxxPCCDPCAxxCxxCRFFxxxCRxLxxxxxC

比較的短いC-末端、およびシステイン5およびシステイン10を欠くようにタンパクを加工して、フォールド形成させたところ、天然タンパクと類似の構造を取る。この加工バージョンを、ライブラリー設計のためのスカフォールドとして用いた。これは、下記の式を有する：

50

【0458】

【化52】

ロゴス: CxxxxxCxxxxxCCxxxxxCxCCCCxxCx, CxxxxxCxxxxxCCDPxxxCxCRFFxxxCxCRxx,
CxGxxxCCCCxxCCDPxxxCYCRFFxxxCxCRxx

完全長アグーチタンパクは、可溶性タンパクとして大腸菌において発現させることが可能である (R. D. Rosenfeld, ら、(1998) *Biochemistry*, 37:16041-52)。

【0459】

PF05375: PMP阻害剤／パシファスチン

10

このファミリーのメンバーの構造から、これらは、三つのジスルフィド架橋によって接続される、3重鎖逆平行ベータシートから構成されることが示される。これは、このファミリーが、セリンプロテアーゼ阻害剤の新規ファミリーであることを規定する (G. Simonet, ら、(2002) *Comp Biochem Physiol B Biochem Mol Biol*, 132:247-55; A. Roussel, ら、(2001) *J Biol Chem*, 276:38893-8)。図34参照。

【0460】

39のファミリーメンバーがある。システイン位置は、ジスルフィド空間配置1-4 2-6 3-5を有し、高度に保存される。個々のシステイン間の距離は<10である。C-末端は、構造の中で見えない。これは、ライブラリー設計において省略可能であることを示唆する。二つの、強力に保存されるアミノ酸としてN15およびT29がある。これらは、プロテアーゼ結合ループの形成と安定化に関与する。これらは、結合の多様性を増すために、ライブラリー設計から省略することも可能である。

20

【0461】

【化53】

- 1) CxxxxxxxxxCxxCxCCCC(x)xxxCCCCC
- 2) CxpGxxxKxxCNxCxCCCC(x)xxxCTxxxC

PF01549: ShTKファミリーおよびStecrisp

Stecrispは、ShTKファミリーと、酷似する3D構造を示すが、ShTKファミリー (PF01549) の一部ではない (M. Guo, ら、(2005) *J Biol Chem*, 280:12405-12)。Stecrisp配列に基づいてBlast探索したところ、30-100%の同一性を持つ48の合致例が得られたが、ShTKファミリーのメンバーは全く得られなかった。図35-36参照。

30

【0462】

Pfam01549は、機能不明のドメインであるが、いくつかのC. elegansタンパクの中に認められる。このドメインは、30アミノ酸長で、6個の保存されたシステイン位置を持ち、これらは3個のジスルフィド架橋を形成する。このドメインは、ShKトキシンにちなんで名づけられた (SMARTによって) (M. Dauplais, ら、(1997) *J Biol Chem*, 272:4302-9)。

40

【0463】

このドメインは、39アミノ酸当たり3ジスルフィドという高いジスルフィド密度、および、1-6 2-4 3-5の空間配置を示す。個々のシステイン間のシステイン間隔は10よりも小さく、したがってライブラリー設計には有用である。このファミリーの異なるメンバーの間で、システイン位置は高度に保存される。

【0464】

PF01549-ShTK。図35参照。

【0465】

【化 5 4】

- 1) Cx(xxx)xxx(x)xxCxxxxx(xx)Cxxx(x)xxxxxxxxCxxCxxC
- 2) Cx(dx)x(dx)x(x)xxCxxxxx(xx)Cxxx(x)xxxxxxxxCxtCxxC

S T E C R I S P の C-末端ドメインおよび関連配列：図 3 6 を参照されたい。

【0 4 6 6】

P F 0 7 9 7 4 : E G F 2 ドメイン

このファミリーのメンバーは、E G F スーパーファミリーに属する。このスーパーファミリーは、E G F フォールドの安定性に必須の、1-3 2-4 5-6 の順序で 3-4 個のジスルフィド結合を形成する、6-8 個のシステインを有することで特徴づけられる。これらのジスルフィド結合は、梯子状に配されて積層される。ラミニン E G F ファミリーは、さらにもう一つのジスルフィド結合を持つことによって区別される。このファミリーにおけるこれらのドメインの機能は不明であるが、主に構造的役割を果たしていると考えられている。通常、これらのドメインは、細胞外タンパクとして縦列反復列として配置されている。

10

【0 4 6 7】

P F 0 7 9 7 4 - E G F 2 : 図 3 7 参照。

【0 4 6 8】

【化 5 5】

- 1) Cx(xxxxx)Cxx(x)xxxCxxx(xxxxxxx)CxCxxx(xxxx)xxxxxC
- 2) Cx(xxxxx)Cxx(x)xGxCxxx(xxxxxxx)CxCxxx(xxxx)xxGxxC

20

他の E G F-様ドメイン

P F 0 0 0 0 8 - E G F : 図 3 8 参照。

【0 4 6 9】

【化 5 6】

- 1) CxxxxxCxxxxCxxxx(xx)xxxCxCxxx(xxxx)xxxxxC
- 2) CxxxxxCxxxgxCxxxx(xx)xxxCxCxxg(xxxx)xxgxxC

P F 0 0 0 5 3 - L a m - E G F : 図 3 9 参照。D B P : 1-3 2-4 5-6 7 30
- 8

【0 4 7 0】

【化 5 7】

- 1) CxCCCCCCCC(xx)CxxCCCCC(xxxx)CxxCCCCCCCCCxxCxxxxxxxxxx(xxxx)C
- 2) CxCCCCCCCC(xx)CCCCCCCC(xxGx)CxCCCCGxxC(DF)xCCCCCCCCC(xxxx)C

P F 0 7 6 4 5 : C a - E G F : 図 4 0 参照。

【0 4 7 1】

【化 5 8】

- 1) CxxxxxCxxxx(xx)CCCCCxCx(xxxx)CCCCCCCCC(xxxxx)C
- 2) CxxxxxCxxxx(xx)CxNxxGx(F,Y)xCx(xxxx)Cxx(G,Y)xxxxxxxx(xxxxx)C

40

P F 0 4 8 6 3 : アリナーゼ E G F-様 : 図 4 1 参照。

【0 4 7 2】

【化 5 9】

- 1) Cxxxxxxxxxxxxxxxx(xxxx)CxCxxCxxxxCxxxxxC
- 2) Cxxxxxxxxxxxxxxxx(xxxx)CxCxxCxxxxCxxxxxC

P F 0 0 3 2 3 : 哺乳類 デフェンシン ; デフェンシン 1
図 4 5 参照。D B P : 1-6 2-4 3-5

50

【0473】

【化60】

- 1) CxCxxxxCxxxxxxxxCxxxxxxxxCC
- 2) CxCRxxxCxxxLrxxGxCxxxgxxxxCC

P F 0 1 0 9 7 : 節足動物デフェンシン : デフェンシン 2

図 4 4 参照。D B P : 1 - 4 2 - 5 3 - 6

【0474】

【化61】

- 1) CxxxCxxxxxxxxC(xxx)xxxCx
- 2) CxxHCxxxgxxGGxCxx(xx)xxxCx

10

P F 0 0 7 1 1 : デフェンシン B、ベーターデフェンシン

図 4 3 参照。D B P : 1 - 4 2 - 5 3 - 6、または 1 - 5 _ 2 - 4 _ 3 - 6

【0475】

【化62】

- 1) CxxxxxxxxCxxxCxxxxxxxxCxxxxxCC
- 2) CxxxgxCxxxCxxxxxxxxgxCxxxxxCC

P F 0 8 1 3 1 : デフェンシン様 ; デフェンシン 3、図 4 2 .

20

【0476】

【化63】

- 1) CxxxGxCrxlxxxnCxxxxxxxxCnxxqkCC
- 2) CxxxGxCrxlxxxnCxxxxxxxxCnxxqkCC

デフェンシン (様) 3 ファミリーは、カモノハシ毒から単離されたデフェンシン様ペプチド (D L P) から成る (A. M. T o r r e s, ら、(1999) B i o c h e m J . 3 4 1 (P t 3) : 7 8 5 - 9 4)。これらの D L P は、ベーターデフェンシン-12、およびナトリウムチャンネル・ニューロトキシン S h 1 のものと、近似の三次元フォールドを示す。しかしながら、ベーターデフェンシン-12および S h 1 にとって機能的に重要なことが知られる側鎖が、D L P では保存されない。このことは、異なる生物学的機能を示唆する。この意見と一致して、D L P は、抗菌性を持たないことが示され、かつ、ラットの後根神経節のナトリウムチャンネル電流に対し観察可能な活性を持たない。3 メンバーしか知られていないが、ベータデフェンシンに近似することは、このスカフォールドを魅力あるものにする。

30

【0477】

このドメインは、36 アミノ酸当たり 3 ジスルフィドという高いジスルフィド密度、および、1 - 5 _ 2 - 4 _ 3 - 6 の空間配置を示す。個々のシステイン間のシステイン間隔は 10 よりも小さく、したがってライブラリー設計には有用である。このファミリーの異なるメンバーの間で、システイン位置は高度に保存される。

40

【0478】

P F 0 0 3 2 1 : クランビン

クランビンは、小型の、基本的植物タンパクであり、45 から 50 のアミノ酸長を持ち、三つから四つの保存されたジスルフィド結合を含む。このタンパクは、動物細胞に有毒であるが、恐らく細胞膜を攻撃し、それを透過的にするものと考えられる。これは、糖分接種の抑制を招き、カリウムイオンおよびリン酸イオン、タンパク、およびヌクレオチドを細胞から漏出させる。このファミリーは、ガンマーチオニン P F 0 0 3 0 4 とは異なる (P. B. P e l e g r i n i, ら、, (2005) I n t J B i o c h e m C e l l B i o l, 3 7 : 2 2 3 9 - 5 3)。

【0479】

50

このドメインは、約46アミノ酸当たり4ジスルフィドという高いジスルフィド密度を示す。個々のシステイン間のシステイン間隔は10よりも小さく、したがってライブラリー設計には有用である。このファミリーの異なるメンバーの間で、システイン位置は高度に保存される。図46参照。

【0480】

システイン位置は高度に保存され、個々のシステイン間の距離は10以下であり、空間配置は1-6 2-5 3-4であり、ドメインは小さく、6システインを有する。

【0481】

三つのジスルフィド結合を含むメンバーのモチーフは、

P F 0 0 3 2 1-克蘭ビン

10

【0482】

【化64】

- 1) xxCCxxxxxxxxCxxxxxxxxCxxxxCxxxxxxxxCxxxx
- 2) xxCCxxxxRxxYxxCxxxGxxxxCxxxxCxJxxxxCxxxx
- 3) xxCCxxxxRxxYxxCRxxGxxxxCxxxxCxIISGxxCPxx(Y,F)xx

四つのジスルフィド結合、および空間配置1-8 2-7 3-6 4-5を有するメンバーのモチーフは、下式によって特徴づけられる：

【0483】

【化65】

20

xxCCxxxxxxxxCxxxCxxxxxxxxCxxxCxCxxxxxxxxC

P F 0 6 3 6 0 : R a i k o v i

僅かに6個のファミリーメンバーしか持たないが、システイン間アミノ酸は高密度である、拡散性ペプチドフェロモン (M. S. Weiss, ら、(1995) Proc Natl Acad Sci USA, 92:10172-6)。システイン位置は高度に保存され、空間配置は1-4 2-6 3-5である。個々のシステイン間の距離は<10。図47参照。

【0484】

【化66】

30

- 1) CxxxxxxxxCxxxCxxxCxxxxxxxxCxxxxxxxxC
- 2) CxxxxxCxxxCxxxCxxxxxxxxCxxxxxxxxC

P F 0 0 6 8 3 : T B ドメイン

トランスフォーミング増殖因子 (TGF)-結合タンパク様 (TB) ドメインは、ヒトのフィブリンから得られた。このドメインは、細胞外基質の繊維状構造に局在する、フィブリン、および潜在的TGF-結合タンパク (LTBP) に認められる (X. Yuan, ら、(1997) Embo J, 16:6659-66)。反復列は、このドメインが、フィブリンおよびLTBPにおいて複数コピーとして見出されることを意味する。図49参照。

【0485】

40

式は、僅か6個の保存システインのみを示す。3種類の構造を分析した (luzq, lapj, lksq)。消失システインは、Cys1と、Cys三重体の間に挿入される (位置8/12、4/12、9/12) が、式では最後のシステインが消失したままである。空間配置は1-3 2-6 4-7 5-8である。

【0486】

【化67】

- 1) CxxxxxxxxxxxxCCCxxx(xx)xxxxCxxCPxxxxxxxxC
- 2) Cxxxxxxxx(x)xxkxxCCCxxx(xx)xxgxxCexCPxxxxxxxxC

P F 0 0 0 9 3 : v o n W i l l e b r a n d 因子型Cドメイン

50

vWFドメインは、各種血漿タンパク、補体因子、インテグリン、コラーゲンVI、VII、XII、およびXIV型、および、その他の細胞外タンパクに認められる（P. Bork (1993) FEBS Lett, 327:125-30）。高度に保存されたシステイン残基を有する、488の既知のファミリーメンバーがある。構造および配列比較から、CRモジュールのN-末端サブドメインと、フィブロネクチン1型ドメインとの間には進化関係のあることが明らかにされており、これらのドメインが共通の祖先を共有することを示唆する（J. M. O'Leary, ら、(2004) J Biol Chem, 279:53857-66）。図50参照。

【0487】

ミニ・コラーゲンシステイン富裕ドメイン

ミニ・コラーゲンは、ヒドラの細胞壁に認められる。ミニ・コラーゲンは、分子内ジスルフィド結合前駆体として合成される、C-末端システイン富裕ドメインを含む。このC-末端ドメインは、独特のフォールドを有するマイクロタンパクである（S. Meier, ら、(2004) FEBS Lett, 569:112-6; E. Pokidysheva, ら、(2004) J Biol Chem, 279:30395-401）。16種のファミリーメンバーの間で、システイン残基だけが高度に保存される。ジスルフィド結合は、分子内ジスルフィド結合に対してシャッフルされて、細胞壁安定基質を形成するものと考えられている。ジスルフィド空間配置は、1-5 2-4 3-6である。C-末端ドメインが、互いに分子内ジスルフィド結合を形成するという所見は、分子間ジスルフィド結合によって連結される二量体分子から成る組み合わせライブラリーを創製するのに利用することが可能である。図136参照。

モチーフ：ミニコラーゲンにおけるC3C3C3C3C3C3C、およびヒドラHOWAタンパクにおけるC5C3C3C3C3C3C3C、これらにおいてこのドメインは反復体として出現する。

【0488】

PF03784：シクロチド

このファミリーは、各種活性を有する、一組の環状ペプチドを含む。構造は、歪んだ、3重鎖ベータシート、および、ジスルフィド結合のシステインノット配置から成る（D. J. Craik, ら、(1999) J Mol Biol, 294:1327-36）。図51参照。

【0489】

空間配置は1-4 2-5 3-6である。

【0490】

【化68】

1) CxxxCxxxxCxxxxxxxxCxCCCC

2) CxExCxxxxCxxxxxxxxGCxCxxxxC

PF06446：ヘプシジン

ヘプシジンは、肝臓において発現される抗菌および抗真菌タンパクであり、さらに、鉄代謝におけるシグナル伝達分子である。ヘプシジンタンパクは、システイン富裕であり、ヘアピンの曲がり角に見られる、異常なジスルフィド結合を有する、歪んだベータシートを形成する。

【0491】

図52参照。空間配置は1-8 2-7 3-6 4-5である。

【0492】

【化69】

モチーフ 1: xxxCxxCCxCxxxxCxxCC

モチーフ 2: FPxCxFCCxCxxxxCGxCC

PF05353：デルターアトラクトキシン

アトラコトキシンの構造は、ベータ領域から突出する、3重鎖親指状延長体を含むコアベータ領域、およびC-末端螺旋を含む。ベータ領域は、他のニューロトキシンポリペプチドにも見られる特徴である、システインノットモチーフを含む。図53参照。

【0493】

空間配置は1-4 2-6 3-7 5-8である。

【0494】

【化70】

モチーフ 1: CxxxxxCxxxxCCCxxCxxxxxxxxCxxxxxxxxxC

モチーフ 2: CxxxxWCxxxxCCCPxxCxxWxxxxCxxxxxxxxxC

10

P F 0 0 2 9 9 : セリンプロテアーゼ阻害剤

このsquash阻害剤は、いくつかのセリンプロテアーゼ阻害剤ファミリーの一つである。これらは、約30残基長で、3ジスルフィド結合を形成する、6Cys残基を含む。空間配置は1-4 2-5 3-6である。図56参照。

【0495】

【化71】

1) CxxxxxCxxxxCxxxCxCxxxx(x)xC

2) CPxxxxxCxxpCxxxCxCxxxx(x)xC

P F 0 1 8 2 1 : アナフィロトキシン様ドメイン

20

C3a、C4a、およびC5aアナフィロトキシンは、補体分子C3、C4、およびC5の活性化の際血清中に酵素的に生成されるタンパク断片である。これらは、平滑筋収縮を誘発する。これらの断片は、フィブリンの3重反復体に対して相同である。空間配置は1-4 2-5 3-6である。このファミリーには123の既知のメンバーがある。図57参照。

【0496】

【化72】

1) CCxxxxxx(xxxx)xxCxxxxxxxx(xx)xxCxxxxxxCC

2) CCxxGxxx(xxxx)xxCxxxxxxxx(xx)xxCxxxFxxCC

30

P F 0 5 1 9 6 : ミドカイン/P T N

成長および分化の調節に参与する、いくつかの細胞外ヘパリン結合タンパクは、新しい、増殖因子ファミリーに属する(W. Iwasaki, ら、(1997) Embo J, 16: 6936-46)。33のファミリーメンバーがある。システイン位置は高度に保存され、ジスルフィドの空間配置は1-4 2-5 3-6である。個々のシステイン間の距離は<10である。ミドカインのNMR構造は、きわめて無秩序なN-およびC-末端を示す。これは、これらは、ライブラリー設計から省略することが可能であることを示唆する。正に荷電する残基は、ヘアピン結合に参与するので、ライブラリー設計から省略することが可能である。図59参照。

【0497】

【化73】

1) CxxxxxxxxCxxxxxxxxCxxxxxxxxCxxxxxxxxCxxxC

2) CxxWxxxxCxxxxDCGxGRExxCxxxxxxxxCxxPCxW

40

P F 0 2 8 1 9 : W A P 「4-ジスルフィドコア」

保存されたシステインのパターンは、これらの配列が、相似のフォールドを取ることを示唆するけれども、配列類似性の総合度は低い(L. G. Henningshausen, ら、(1982) Nucleic Acids Res, 10: 2677-84)。25の既知のファミリーメンバーがある。図62参照。

【0498】

50

空間配置は 1-6 2-7 3-5 4-8 である。

【0499】

【化74】

- 1) Cxxxx(xx)xxxxCxxx(xxx)CxxxxCxxxxCCxxxC
- 2) CPxxx(xx)xxxxCxxx(xxx)CxxDxxCxxxxKCCxxxC

P F 0 2 0 4 8, P F 0 7 8 2 2 : 毒性ヘアピン

トキシシン 1 3 (P F 0 7 8 2 2) は折りたたまれて、4 S S ジスルフィドアルファ螺旋ヘアピンを形成する。S C O P データベースはさらに、熱安定性エンテロトキシシン (P F 0 2 0 4 8) を、1-4 2-5 3-6 の D B P を持つ毒性ヘアピンと掲載する。

10

【0500】

このファミリーのメンバーは、海産ワーム *Cerebratulus lacteus* によって生産される甲殻類選択的ニューロトキシシンである、ニューロトキシシン B-I V によく似る。この高度に陽イオン性のペプチドは、約 55 残基であり、ヘアピン構造の明瞭な形状を持つループによって接続される、2 本の逆平行螺旋を形成するように配される。ヘアピンの側枝は、四つのジスルフィド結合によって連結される。活性に重要であると特定された三つの残基は、分子の同じ面に見出されるが、一方、活性に重要なもう一つの残基、T r p 3 0 は、反対側にある。このタンパクの活動方式は、完全には理解されていないが、電圧依存性ナトリウムチャンネルに対し、おそらく、これらのタンパクの、まだ解明されていない部位に結合することによって作用すると考えられる。図 6 5 参照。トキシシン 1 3 の空間配置は 1-8 2-5 3-6 4-5 である。

20

【0501】

【化75】

- 1) CxxxCxxxxxxxxCxxCxxxxxxxxCxxxCxxxxxxxxCxxxC
- 2) CxxxCxxxxyxxCxxCxxgWxxgxxCxxhCxxxxxxxxCxxxC

P F 0 6 3 5 7 : オメガアトラコトキシシン

オメガアトラコトキシシン-H v 1 a は、昆虫特異的ニューロトキシシンであり、その系統発生的特異性は、脊椎動物ではできないが、昆虫の電圧依存性カルシウムチャンネルに拮抗する能力から得られる (X. Wang, ら、(1999) Eur J Biochem 30 264:488-94)。空間配置は 1-6 2-7 3-4 5-8。

30

【0502】

図 6 6 参照。空間配置は 1-4 2-5 3-6。

【0503】

【化76】

CxPxxxPCPYxxxxCCxxxCxxxxxxGxxxxxxC

P F 0 6 9 5 4 : レジスチン

このファミリーは、いくつかの哺乳類のレジスチンタンパクから成る。レジスチンの循環レベルの上昇は、インスリンの一定の生理的レベルの存在下にグルコース生産を刺激し、他方では、インスリンはレジスチンの発現を抑制することが実証されている。

40

【0504】

レジスチンは、C 末端のジスルフィド富裕部分のマルチマー化に関与する N-末端アルファ螺旋を含む。図 6 7 参照。空間配置は 1-10 2-9 3-6 4-7 5-8。

【0505】

ジスルフィド-富裕ミクロタンパクのみが示される。N-末端アルファ螺旋モチーフは、ミクロタンパクのマルチマー化に使用することが可能である。

【0506】

【化 7 7】

- 1) CxxxxxxxxxxxxCxxxxxxxxCxCxxxCxxxxxxxxCxCxxxxxxxxCC
 2) CxxxxxxxxxxxxCPxGxxxxxCxCGxxCGxWxxxxxCxCxxxDWxxRCC

P F 0 0 0 6 6 : ノッチ / D S L

動物の発達過程に与る膜貫通タンパクの細胞外ドメイン (J. C. Aster, ら、(1999) Biochemistry, 38:4736-42; D. Vardar, ら、(2003) Biochemistry, 42:7061-7)。DSL 反復列は縦列で起こる (3x)。三つの保存された Asp または Asn 残基。NMR 構造では、D12、N12、D30、D33 が Ca²⁺ 結合部位を形成する。一異性形だけが、ミリモルの Ca²⁺ の存在下に形成され、複数の異性形は、Mg²⁺ または EDTA の存在下に観察される。システイン位置は高度に保存され、1-5 2-4 3-6 空間配置を持つ。第1システインのN-末端および最終システインのC-末端のアミノ酸は多くはない (<3)。個々のシステイン間の距離は<10。図68参照。

【0507】

【化 7 8】

- 1) Cx(xx)xxxCxxxxxxxxCxxxCxxxCxxxxxC
 2) Cx(xx)xxxCxxxxxxgxCxxxCnxxxCxxDGxDC

P F 0 0 0 2 0 : T N F R

いくつかのタンパクは、その内のあるものは増殖因子の受容体であることが知られるが、N-末端領域にシステイン富裕ドメインを含むことが判明しており、このドメインは、さらに、その全てが鎖内ジスルフィド結合に関与する、六つの保存システインを含む、4個の (ある場合には、3個の) 反復列に分割される (M. D. Jones, ら、(1997) Biochemistry, 36:14914-23)。このドメインは、空間配置 1-2 3-5 4-6 を有する、高度に保存されるシステイン残基を含む。

【0508】

図69参照。

【0509】

【化 7 9】

- 1) Cxxx(x)xxxxxxxx(x)xxCx(x)CxxCxx(xx)xxxxxxxxCxxxxxxxxC
 2) Cxxx(x)x[yf]xxxxx(x)xxCx(x)CxxCxx(xx)gxxxxxxxxCxxxxxtxC

P F 0 0 0 3 9 : フィブロネクチン I I 型ドメイン

フィブロネクチンは、複数ドメイン糖タンパクであり、血漿中では可溶形として見られ、細胞表面、および、各種化合物、例えば、コラーゲン、フィブリン、ヘパリン、DNA、およびアクチンなどに結合する。

【0510】

図70参照。1-3 2-4 空間配置。

【0511】

【化 8 0】

モチーフ: CxfpxxxxxxxxCxxxxxxxxwCxxxxxxxxDxxxxxC

P F 0 2 0 1 3 : セルロースまたはタンパク結合ドメイン

好気性細菌に見られるものは、セルロース (または、他の炭水化物) に結合するが、嫌気性真菌類では、それらは、タンパク結合ドメインであり、ドッケリンドメイン、またはドッキングドメインと呼ばれる。

【0512】

1-2 3-4 の空間配置。図71参照。

【0513】

10

20

30

40

【化 8 1】

モチーフ: Cxx(xxx)xxxwCxxxxxxxxxxwxxxxxxxxDxxxxCxxx(xxx)xxxxxxxxwxxxxxxxxC

P F 0 0 7 3 4 : 菌類セルロース結合ドメイン

構造的に、セルラーゼおよびキシラーゼは、一般的に、プロリンおよび／またはヒドロキシアミノ酸富裕な短いリンカー配列によって、セルロース結合ドメイン (C O D) に接続される触媒ドメインから成る (N. R. G i l k e s, ら、(1991) M i c r o b i o l R e v, 55: 303-15)。いくつかの菌類セルラーゼの C B D は、36 アミノ酸残基から成ることが示されており、C B D は、該酵素の N-末端または C-末端のどちらかに見出される。このファミリーのメンバーは、二つのジスルフィド結合を持ち、空間配置は 1-3 2-4 である。図 7 2 参照。

10

【0 5 1 4】

【化 8 2】

モチーフ: qCGGxxxxGxxxCxxgxxCxxxxxy

P F 0 0 2 1 9 : インスリン様増殖因子結合タンパク

インスリン様増殖因子 (I G F - I および I G F - I I) は、細胞外液において、特異的結合タンパクに高いアフィニティーをもって結合する。このファミリーのメンバーは、二つのジスルフィド結合を持ち、空間配置は 1-3 2-4 である。図 7 4, 7 5 参照。

【0 5 1 5】

20

P F 0 0 3 2 2 : エンドセリンファミリー

エンドセリン (E T) は、知られるもっとも強力な血管収縮因子である。21 残基長の、これらのペプチドは、二つの分子内ジスルフィド結合を含み、1-4 2-3 の空間配置を持つ。図 7 6 参照。

【0 5 1 6】

P F 0 2 0 5 8 : グラニュリン前駆体

15 アミノ酸ペプチドであるグラニュリンは、S t a R という名で知られる小腸受容体グアニレートシクラーゼ C の内因性リガンドである。これらのペプチドは、二つの分子内ジスルフィド結合を含み、1-3 2-4 の空間配置を持つ。図 7 7 参照。

【0 5 1 7】

30

P F 0 2 9 7 7 : カルボキシペプチダーゼ阻害剤

ペプチドプロテイナーゼ阻害剤は、タンパクの中の単一ドメインタンパク、または、単一または複数ドメインとして認められる。これらは、それぞれ、単純または複合阻害剤と呼ばれる。多くの場合、これらは、より大型の前駆体タンパクの一部として、プレペプチドとして、または、不活性ペプチダーゼまたはチモーゲンと関連する N-末端ドメインとして合成される。第 2 ペプチダーゼとの相互作用、または自己触媒的切断による、N-末端阻害剤ドメインの除去は、チモーゲンを活性化する。

【0 5 1 8】

35 の既知のファミリーメンバーがある。空間配置は 1-4 2-5 3-6 である。

【0 5 1 9】

40

【化 8 3】

1) CxxxxxCxxxxCxxxCxCxxxxxC

2) CPxxxxCxxdxCxxxCxCxxxxxCg

P F 0 6 3 7 3 : C A R T

C A R T は、主に、システインノットに共通な逆平行ベータ・シートから成る、数枚の僅かな延長体によって構成されるコンパクトな枠組みによって囲まれる、巻き回転およびループから成る。13 の既知のファミリーメンバーがある。

【0 5 2 0】

空間配置は 1-3 2-5 4-6。図 8 1 参照。

50

【0521】

他の全てのファミリーとは違って、非cys-残基は、むしろ保存されるので、このファミリーは、ランダム化には好ましい選択であるようには見えない。

【0522】

フォリスタチン

ヒトのフォリスタチンは、FDA承認製品であり、非免疫原性で、したがって、70-72 AAのフォリスタチンドメインは、魅力的なスカフォールドである。これは、合計36システイン残基を含むが、これらは、四つの自律的フォールド形成単位に対応する、ジスルフィド架橋の、非重複組の中に配置されると考えられている。これらの単位の内の第1は、我々はこれをFs0と呼ぶが、成熟ポリペプチドの、63N-末端残基を含むが、既知の構造を持つ他の、どのタンパクとも配列類似性を持たない。それと対照的に、フォリスタチン鎖の残余部分は、フォールド形成して、一連の、三つの連続する、70-74残基長のフォリスタチンドメインを形成する。これらは、Fs1、Fs2、およびFs3と呼ばれる構造的反復列であるが、細胞外基質タンパクBM-40のフォリスタチン様ドメインに対し相同性を示し、さらに、他の細胞外基質タンパク、例えば、アグリン、トモレグリン、および補体タンパクC6およびC7にも認められる。各69-72 AAフォリスタチンドメインは、1-3 2-4 5-9 6-8 7-10のDBPを有する。

10

【0523】

PF00713：ヒルジン

ヒルジンファミリーは、MEROPS阻害剤ファミリー、IM族に属するプロテイナーゼ阻害剤の1群である。これらは、S1ファミリーのセリンペプチダーゼを抑制する。

20

【0524】

ヒルジンは、*Hirudinaria manillensis* (バッファロー・ヒル) および *Hirudo medicinalis* (医用ヒル) の唾液腺によって分泌される強力なトロンビン阻害剤である。これは、アルファートロンビンと安定な非共有複合体を形成し、トロンビンのフィブリノーゲン切断能力を破壊する。ヒルジンの構造は、NMRによって解決されており、組み換えヒルジントロンビン複合体の構造は、X線結晶学によって2.3 Åまで決定されている。ヒルジンは、N-末端の小球ドメインと、延長したC-ドメインとから成る。残基1-3は、トロンビンの残基214-217と平行なベータ鎖を形成し、残基1の窒素原子は、触媒部位のSer195 Oガンマ原子と水素結合を形成する。C-末端は、トロンビンの陰イオン結合性細胞外部位と、数多くの静電的相互作用を形成し、一方、最後の5個の残基は、多くの疎水性接触を形成する螺旋ループの中に納まる。図123参照。

30

【0525】

PF06410：グルマリン

グルマリンは、トウワタ蔓植物 *Gymnema sylvestre* から得られた35-残基ポリペプチドである。これは、ラットにおいて甘味に対する神経反応を選択的に抑制する能力を持つので、従来、甘味転換の研究において薬学的ツールとして利用されている。

【0526】

二つの既知のファミリーメンバーがある。空間配置は1-4 2-5 3-6. 図82参照。

40

【0527】

【化84】

- 1) CxxxxxCxxxxxCxxxxCxxxxxxxxxC
- 2) CxxxxxCxxxxxCxxxxCxxxxwxxxxC

PF08027：アルブミン-1

このアルブミン-1タンパクは、ホルモン様ペプチドであるが、膜結合43 kDa受容体に結合するとキナーゼ活性を刺激する。このドメインの構造は、フォールド様ノッチン

50

の存在を明らかにし、3本のベータ鎖から構成される。34の既知のファミリーメンバーがある。空間配置は1-4 2-5 3-6. 図83-84参照。

【0528】

P F 0 8 0 2 8 : ニューロトキシン (A T X I I I)

このファミリーは、*Anemonia sulcata* トキシン I I I (A T X I I I) ニューロトキシンファミリーから成る。A T X I I I は、いそぎんちゃくによって生産されるニューロトキシンである。これは、四つの反転巻き回転および、他の、二つの鎖翻転を含むコンパクトな構造を採用するが、正規のアルファ螺旋またはベータシートを持たない。このペプチドの表面に認められる疎水性パッチは、ナトリウムチャンネル結合表面の一部を構成するようである。二つの既知のファミリーメンバーがある。空間配置は1-4 2-5 3-6。

10

【0529】

図85参照。

【0530】

【化85】

モチーフ: C C x C x x x x x x x C x x x x x x x x C

P F 0 1 1 4 7 : C H H / M I H / G I H 神経ホルモン

節足動物は、ニューロペプチドのファミリーを発現するが、その中には、甲殻類から分泌される高血糖ホルモン (C H H)、脱皮抑制ホルモン (M I H)、性腺抑制ホルモン (G I H)、および下顎器官抑制ホルモン (M O I H)、および、いなごから分泌されるイオン輸送ペプチド (I T P) が含まれる。

20

【0531】

131の既知のファミリーメンバーがある。空間配置は1-5 2-4 3-6. 図86参照。

【0532】

P F 0 4 7 3 6 : エクロシオン

エクロシオンファミリーは、脱皮行動の実行を起動する昆虫ニューロペプチドであり、脱皮の最後に古いクチクラの脱ぎ捨てを起こさせる。5の既知のファミリーメンバーがある。空間配置は1-5 2-4 3-6. 構造データは無い。図88参照。

【0533】

30

【化86】

1) C x x C x x C x x x x x x x x x x C x x C x x x x x x x x C

2) C x x C x q C k x x x g x x f g x x C x x C x x x g x x p x C

P F 0 1 1 6 0 : 内因性オピオイドニューロペプチド

脊椎動物の、内因性オピオイドニューロペプチドは、前駆体タンパクの翻訳後タンパク分解酵素切断によって放出される。この前駆体は、下記の成分から成る。約50残基の保存領域に先行するシグナル配列、可変長領域、およびニューロペプチドそのものの配列。配列分析から、前駆体の保存されたN-末端領域が、ジスルフィド結合の形成に与えられる6個のシステインを含むことが明らかになった。この領域は、ニューロペプチドの処理に重要と想定される。50の既知のファミリーメンバーがある。空間配置は1-5 2-4 3-6. 構造データは無い。図89参照。

40

【0534】

【化87】

1) C x x C x x C x x x x x x x x x x C x x C x x x x x x x x C

2) C x x C x x C x x x x x x x x x x C x l x C x x x x x x x W x x C

P F 0 8 0 3 7 : 軟体動物フェロモン

このファミリーは、水性フェロモンのアトラクチンファミリーから成る。A p l y s i a の交尾誘引は、産卵時に放出される、アトラクチンペプチドの形を取る長距離の水性シ

50

グナルを含む。これらのペプチドは、6個の保存システインを含み、折りたたまれて二つの逆平行螺旋を形成する。第2螺旋は、A p l y s i a アトラクチンに保存される I E E C K T S 配列を含む。5の既知のファミリーメンバーがある。空間配置は 1-6 2-5 3-4。図90参照。

【0535】

【化88】

1) CxxxxxxxxCxxxxxxxxCxxxxxxxxCxxxxxxxxCxxxxxxxxC

2) CxxxxxxxxCqxxxxxxxxCxxxxxxxxCxxxxxxxxCxxxxxxxxC

P F 0 3 9 1 3 : A M B V タンパク

10

A m b V は、A m b r o s i a s p (ブタクサ) タンパクである。A m b V は、主要な T 細胞エピトープとして C-末端螺旋を含むことが示されている。遊離スルフィド基も、これらの関連アレルゲンにおける交差反応性 T 細胞エピトープに対する T 細胞の認識において大きな役割を果たしている。

【0536】

3の既知のファミリーメンバーがある。空間配置は 1-7 2-5 3-6 4-8。図92参照。

【0537】

【化89】

1) CxxxxxxxxCxxxxxxxxC(x)xxxxCxxxxxxxxCxxxC

20

2) CgxxxxxxxxCxxxgxyC(x)xxxxCyxxxxxxxxCxxxC

付録 B : 二重モチーフを含む H D D ドメイン

P F 0 1 4 3 7 : プレキシン P S I

システイン富裕反復列が、いくつかの異なる細胞外受容体に認められている (J. S t a m o s, ら、(2004) E m b o J, 23:2325-35; J. P. X i o n g, ら、(2004) J B i o l C h e m, 279:40252-4)。反復列の機能は未知である。プレキシンには反復列の三つのコピーが認められる。反復列の2コピーは、マホガニータンパクの中に認められる。関連の C. e l e g a n s タンパクは、反復列の4コピーを含む。M e t 受容体は、反復列の単一コピーを含む。P f a m 整列によって、6個の高度に保存されるシステイン残基が示された。これらは、三つの保存されたジスルフィド架橋を形成するようであるが、一方、さらに二つのシステインが、位置5および7にも観察され、これらもジスルフィド結合に関与しているようである。空間配置は 1-4 2-8 3-6 5-7 (構造 1 s h y)。セマフォリン (構造 1 o l z) は、空間配置 1-4 2-6 3-5 を持つ、僅かに三つのジスルフィド結合しか含まない。図93参照。

30

【0538】

【化90】

1) CxxxxxCxxCxxxxxxxx(x)xCxxCxxxxxxxxCxxxx(xxxxxx)xCxxxx(xxxxxxxxxx)xxxxxxxxC

2) CxxxxxCxxCxxxxxxxx(x)xCxWCxxxxxxxxCxxxx(xxxxxx)xCxxxx(xxxxxxxxxx)xxxxxxxxC

40

C y s 7 および C y s 8 の間のループは、挿入に対し極めて寛容である。例えば、インテグリンのベータサブユニット構造では、ハイブリッドドメインが、これらのシステインの間に挿入される (J. P. X i o n g, ら、(2004) J B i o l C h e m, 279:40252-4) が、C y s 8 は、依然として C y s 2 とジスルフィド結合を形成する。これを利用して、C y s 7 の後に任意の配列を挿入することが可能である。

【0539】

設計: C x x x x x C x x C x x x x x x (x) x C x x C x x x x x C x x x x (x x x x x) x C x x x x x x x x x (x x x x x) (" a n y s e q u e n c e ") C

これは、複数プレキシンを創製するのに使用することが可能である。

50

【0540】

第1挿入：C x x x x x C x x C x x x x x x (x) x C x x C x x x x x C x x x x (x x x x x x) x C x x x x x x x x (x x x x x) (“ P L E X ”) C。式中、P L E Xは、C x x x x x C x x C x x x x x x (x) x C x x C x x x x x C x x x x (x x x x x) x C x x x x x x x x (x x x x x) x x x x x x Cに対応する。

【0541】

第2挿入：

C x x x x x C x x C x x x x x x (x) x C x x C x x x x x C x x x x (x x x x x x) x C x x x x x x x x (x x x x x) (“ P L E X I N ” (“ P L E X I N ”)) C。式中、 (“ P L E X I N ” (“ P L E X I N ”)) は、C y s 7の後、“ P L E X ”、および、挿入プレキシン配列の中に、複数の、その後の挿入のC y s 7の後、C x x x x x C x x C x x x x x x (x) x C x x C x x x x x C x x x x (x x x x x x) x C x x x x x x x x (x x x x x) (“ P L E X ”) Cへの、C x x x x x C x x C x x x x x x (x) x C x x C x x x x x C x x x x (x x x x x x) x C x x x x (x x x x x x x x x x x) x x x x x x Cの挿入に対応する。

【0542】

P F 0 0 0 8 8：トレフォイルおよび大型トレフォイル

いくつかの真核細胞細胞外タンパクにおいて、約45アミノ酸残基から成るシステイン富裕モジュールが認められている(M. D. C a r r ら、(1994) P r o c N a t l A c a d S c i U S A, 91:2206-10; T. Y a m a z a k i, ら、(2003) E u r J B i o c h e m, 270:1269-76)。ヒトのT F F 3は、大腸菌の周辺ゾルにおいて高濃度に発現される(15mg/l培養体)。このモジュールは、45アミノ酸当たり3ジスルフィド結合を有する、高いジスルフィド密度を示し、空間配置は1-5 2-4 3-6である。大型トレフォイルは、もう一つのジスルフィド結合によって連結される二つの隣接モジュールから成り、接続性は、1-14 2-6 3-5 4-7 8-12 9-11 10-13である。個々のシステイン間のシステイン間隔は10未満であり、したがってライブラリー設計には有用である。システイン位置は、このファミリーの異なるメンバーの間で高度に保存される。図94-95参照。

【0543】

【化91】

- 1) C(x)xxxxxxxxCxx(x)xxxxxxxxCxxxxCCxxxx(x)xxxxCx
- 2) C(x)xxxxxxxxRxxCxx(x)xxxxxxxxCxxxxCCfxxxx(x)xxxxwCf
- 3) C(x)xxxxxxxxRxxCgx(x)xxxxxxxxCxxxgCC[fwy]dxxx(x)xxxxwC[fy]

二つの隣接モジュール、および余分の1-14ジスルフィド連結を有する大型トレフォイル変異種の式は：

【0544】

【化92】

CxC(x)xxxxxxxxCxx(x)xxxxxxxxCxxxxCCxxxx(x)xxxxCxxxxxxxxxxxxC(x)xxxxxxxxCxx(x)xxxxxxxxCxxxxCCxxxx(x)xxxxCxxxxxxxxC

および誘導体。

【0545】

図134は、トレフォイルモチーフから創製することが可能な、反復性「ポリ・トレフォイル」構造を示す。

【0546】

P F 0 0 0 9 0：トロンボスポンジン1

このモジュールは、トロンボスポンジタンパクの中に存在し、該タンパクの、補体経路に関与するいくつかのタンパク、および細胞外基質タンパクにおいて3回反復される。

これは、細胞－細胞相互作用、血管形成およびアポトーシスの抑制に関与することが示されている（P. Bork (1993) FEBS Lett, 327:125-30）。図96参照。

【0547】

このドメインは、約50アミノ酸当たり3ジスルフィド結合を有する、高いジスルフィド密度を示し、空間配置は1-5-2-6-3-4である（T. M. Misenheimer, ら、(2005) J Biol Chem）。個々のシステイン間のシステイン間隔は10未満であり、したがってライブラリー設計には有用である。システイン位置は、このファミリーの異なるメンバーの間で保存される。

【0548】

【化93】

CxxxGxxxxxxGxxxxxx(xxx)xxxxxCxxxxxx(xxx)xxxC(x)xxxxC
 CxxxGxxGxxxxRxxxxxx(Pxxx)xxxxxCxxxxxx(xxx)xxxC(x)xxxxC
 CxvtGxxGxxxxRxxGxxxx(Pxxx)xxxxxCxxxxxx(xxx)xxxC(x)xxxxC

PF00228：ボウマン・バーク阻害剤

ボウマン・バーク阻害剤ファミリーは、数多くの、セリンプロテイナーゼ阻害剤ファミリーの一つである。これらは、二重構造を持ち、一般に、二つの、異なる抑制部位を有する。これら阻害剤は、主に、植物、特に、豆類の種子および穀粒に認められる（R. F. Qi, ら、(2005) Acta Biochim Biophys Sin (Shanghai), 37:283-92）。

【0549】

二つの異なるクラスがある。1) 14システインおよび空間配置1-14 2-6 3-13 4-5 7-9 8-12 10-11を有するドメイン、または、10システインおよび空間配置1-10 2-5 3-4 6-8 7-9を有するドメイン。これらのサブファミリーのために、式におけるCys位置は十分に保存されているようには見えないが、各サブファミリーについては保存されている。

【0550】

このドメインは、約50アミノ酸当たり5から7ジスルフィド結合を有する、高いジスルフィド密度を示す。個々のシステイン間のシステイン間隔は10未満であり、したがってライブラリー設計には有用である。システイン位置は、このファミリーの異なるメンバーの間で保存される。図97-98参照。

【0551】

PF00184：神経下垂体ホルモン、C-末端ドメイン

ノナペプチドホルモンであるバソプレッシン、およびオキシトシンは、ニューロフィジンとも呼ばれるジスルフィド富裕タンパククラスと1:1比で複合体を形成する神経分泌性顆粒に高濃度で見出される。二つの緊密に関連するNPクラスが特定されており、一つはバソプレッシンと、他方はオキシトシンと複合体を形成する（L. Q. Chen, ら、(1991) Proc Natl Acad Sci USA, 88:4240-4）。このファミリーには75メンバーがあり、システイン位置は高度に保存される。式では、システイン富裕モジュールが二重化される。図99参照。

【0552】

両モジュールとも相同のジスルフィド空間配置を有する。一ジスルフィドは、Cys1およびCys8によって二つのモジュールを接続する。このジスルフィド結合を無視すると、各モジュールのジスルフィド空間配置は、1-3 2-6 4-5である。図100参照。

【0553】

ニューロフィジンの結晶構造から、一モノマーは、それぞれ四つの逆平行ベータ鎖を有する、二つの相同の層から成ることが明らかにされた。この二つの領域は、螺旋とそれに続く長いループによって接続される。モノマー－モノマー接触は、逆平行ベータ・シート

10

20

30

40

50

相互作用を含み、これは、8枚のベータ・シートから成る2層を有する二量体を形成する。

【0554】

P F 0 0 2 0 0 : 拡張可能二量体ジスインテグリン

ジスインテグリンは、全てジスルフィド結合に関与するたくさんのシステインを含む、約50-80のアミノ酸残基から成るペプチドである。ジスインテグリンは、多くの接着タンパクの認識部位である、A r g - G l y - A s p (R G D) 配列を含む。ジスインテグリンのR G D配列は、糖タンパクI I b - I I I a複合体と相互作用を持つと考えられている。

【0555】

ジスインテグリンは、長さとしステイン含量にしたがって分類される(J. J. C a l v e t e , ら、(2005) T o x i c o n , 45:1063-74)。

【0556】

短小: C x x x x C C x x C x x x x x x x x C x x x x x x x x (x x) C x x x x C x C、4 S Sおよびジスルフィド空間配置1-4 2-6 3-7 5-8。

【0557】

中間: x C x x x x x x C C x x x x C x x x x (x) x x x C x (x x x) x x x C C x x C x x x x x x x x C x x x x x x x x x x x C x x x x x x x C

6 S Sおよびジスルフィド空間配置1-5、2-4、3-8、6-8、7-11、10-12。

【0558】

長大:

x x x x x x x x x x C x C x x x x C x x x C C x x x x C x x x (x) x x x C x (x x x) x x x C C x x C x x x x x x x x x C x x x x x x x x x x C x x x x x x x x C、7 S Sおよびジスルフィド空間配置1-4、2-7、3-6、5-11、8-10、9-13、12-14。

【0559】

二量体: C C x x x x C x x x x (x) x x x C x (x x x) x x x C C x x C x x x x x x x x C x x x x x x x x x x C x x x x x x C、4 S S、およびジスルフィド空間配置1-7、4-6、5-10、8-10、および、C y s 2およびC y s 3を含む、二つの分子間S Sが、二量体インテグリンを生成する。図101および157参照。ジスルフィド結合の欠失/付加によって特徴づけられる、これらの異なる群の間の進化的関係が認められた。したがって、このモチーフは、インビトロ進化の際、延長することが可能である。

【0560】

付録C: 高度に反復されるモチーフを有するスカフォールド

システイン富裕反復性タンパク(C R R P)

P F 0 0 3 9 6 : グラニュリン

グラニュリンは、複数の生物学的活性を持つと考えられる、約6 K dのシステイン富裕ペプチドである(A. B a t e m a n , ら、(1998) J E n d o c r i n o l , 158:145-51)。前駆体タンパク(アクログラニンと呼ばれる、配列については後記参照)は、七つの異なるグラニュリン形(g r n Aからg r n C)をコードすると予想され、これらは、翻訳後のタンパク分解処理によって放出されると考えられる。グラニュリンは、移動イナゴの脳間部から抽出されるペプチドである、P M P - D 1と進化的に関連する。図103参照。グラニュリン間隔:

【0561】

【化94】

CxxxxxCxxxxCCxxxxxxxxCCxxxxCCxxxxCxxxxCxxxxCxx

DBP: 1-3_2-5_4-7_6-9_8-11_10-12

10

20

30

40

50

サイズを拡張するための設計（キャッピングモチーフには下線を施し、1 反復列はイタリック体、1 反復列はゴシック体）

【0 5 6 2】

【化 9 5】

3C6C5CC8CC6CC5CC5CC8CC6CC5CC5C6C2

振れを導入するための設計

【0 5 6 3】

【化 9 6】

3C6C5CC₄4G3CC₄P5CC₂2G2CC₄P4C6C2

天然の 8-6-5-5 パターン、またはより規則的な 5-5-5-5 パターンを使用することも可能である。構造はベータシートであるので、一つの方法は、良好なベータシート形成材料であるアミノ酸を積極的に選び、ベータシート形成材料でないアミノ酸を回避することである。下記のアミノ酸は好ましく、混合コドンによって得ることが可能である。バリン、イソロイシン、フェニルアラニン、チロシン、トリプトファン、およびトレオニン。

【0 5 6 4】

5 A A ランダム配列を取ることを設計する。

【0 5 6 5】

【化 9 7】

3C6C5CC5CC5CC5CC5CC5CC5CC5C6C2

最小スタータンパクは二つのエンドキャップしか持っていない。

C 6 C 5 C 5 C (1 7 ランダム A A)

最小単位増加を加える。

【0 5 6 6】

【化 9 8】

C6C5CC5C6C

加工工程：ライブラリー作製、水平延長、ランダム化 5 C C 5 単位の付加、水平延長、5 C C 5 単位の付加など。

【0 5 6 7】

P F 0 2 4 2 0 : 抗凍結タンパク

抗凍結タンパクは、ベータ螺旋構造を形成する、8 k D a タンパクである (M. E. D a l e y, ら、(2002) B i o c h e m i s t r y, 41: 5515-25)。N-末端キャッピングモチーフは、ミクロタンパクドメインおよび 1-3 2-5 4-6 空間配置によって形成される。ジスルフィド結合 1-2 を有する 2 C 5 C 3 の反復単位が、このモチーフに付加される。トレオニンは、氷結合に関与するので保存されるが、設計のためには省略してよい。セリンおよびアラニンは、小さな側鎖のみが螺旋の内部に適合するので保存される。疎水性コアの完全な欠失は注目すべきことである。図 1 0 4 は、いくつかの抗凍結由来反復性タンパクである。図 1 0 4 は、いくつかのモチーフを示す。図 1 2 7 参照。

【0 5 6 8】

天然配列：

【0 5 6 9】

【化 9 9】

QCTGGADCTSGTACTGCGNCPNA(VTCTNSQHCVKA)(NTCTGSTDCNTA)(QTCTNSKDCFEA)(NTCTDSTNCKA)(TACTNSSGCPGH)

反復列は、このように示すと多少明確になる。

【0 5 7 0】

【化 1 0 0】

QCTGGADCTSCTGACTGCGNCPNA

(VTCTNSQHCVKA)

(NTCTGSTDCNTA)

(QCTCTNSKDCFEA)

(NTCTDSTNCKYA)

(TACTNSSGCPGH)

様々な設計（キャッピングドメインは下線；反復はイタリック体）

【 0 5 7 1】

【化 1 0 1】

1) 1C5C2C3C2C2C3(*2C5C3*)_n

2) 1C5C2C3C2C2C3(*xtCbxxxxCxxa*)_n

3) QCTGGA(*DCTSCTGACTGCG*)(*DCTSCTGACTGCG*)_n

4) QCTGGA(*DCTSCTGACTGCGA*)(*DCTSCTGACTGCGA*)_n

P F 0 0 7 5 7 : フリン様ドメイン

フリン様システイン富裕領域は、真核細胞由来の各種タンパク、受容体凝集に与る、受容体チロシンキナーゼによるシグナル変換の機構に参与するタンパクに認められている。

図 1 0 5 参照。

【 0 5 7 2】

式のサブセットは折りたたまれてラセン状の反復列を形成し、ライブラリ設計：

【 0 5 7 3】

【化 1 0 2】

CxxxCxxxCxxxxxxCCxxxCxxxCxxxxxxxC.

のためのスカフォールドとして使用される。このモチーフの空間配置は、1－3 __ 2－4 __ 5－7 __ 6－8である。このファミリーのメンバーは、そのシステイン位置および間隔に高度の保守性を示す。この反復列は、

【 0 5 7 4】

【化 1 0 3】

(CxxxCxxxCxxxxxxxC)_n

を、前記モチーフの C－末端に付加することによって延長させることが可能である。

【 0 5 7 5】

P F 0 3 1 2 8 : C x C x C x

この反復列は、Xが任意のアミノ酸である、保存パターン C X C X Cを含む。この反復列は、血管内皮細胞増殖因子 C において最大 5 コピー認められる。双翅類 *Chironomus tentatus* の唾液腺では、特定のメッセンジャーリボ核タンパク (mRNP) 粒子、*Balbani* リング (BR) 顆粒が、遺伝子にたいするその集合時、および、核・原形質輸送の際に、視認される。この反復列は、その 70 を超えるコピーが、*balbani* リングのタンパク 3 (下記参照) の中に認められる。反復列はまた、いくつかの絹タンパクにも認められる。

【 0 5 7 6】

C X C X C 反復列は、内部ではジスルフィド結合を形成せず、したがって、ループは、三つのアミノ酸にまたがるだけであり、データベースのどのマイクロタンパクも、3 のシステインスパンを持たない。図 1 0 9 に示すように、C x C x C x モチーフのシステインは、ジスルフィドが反復列の異なるコピーを連結する、真の反復列の形成に与る。通常、C x C x C x 反復列の間にはただ一つのシステインしか認められない (式では保存されるが、位置は変動する可能性がある)。図 1 0 6、1 0 7、1 0 8 参照。

10

20

30

40

50

【0577】

実際：

【0578】

【化104】

C10C1C1C8C10C1C1C8C10C1C1C3C10C1C1C6C11C

抽出、開始と終末：

【0579】

【化105】

C1C8C10C1C1C8C10C1C1C8C10C1

10

ジスルフィド結合構造のモデルを図109に示す。

【0580】

P F 0 5 4 4 4 : D U F 7 5 3

ショウジョウバエにおいて機能未知のいくつかのドメインで反復される配列。

【0581】

図110.

P F 0 1 5 0 8 : パラメシウム

前述の反復列の37コピーを含む表面抗原。構造的役割が示唆される。二次構造予測から、アルファヘリックスの欠如、およびベータ・シート構造の存在が示唆された（これを行うのに行ったのかは分からない、ジスルフィドの存在は、予測を妨げる可能性がある）。図111-112参照。

20

【0582】

P F 0 0 5 2 6 : ジクティール

いくつかの*Dictyostelium*（タマホコリカビ）種は、保存反復列を含むタンパクを有する。これらのタンパクは、細胞外基質タンパクB、環状ヌクレオチドフォスフォジエステラーゼ阻害剤前駆体、前柄タンパク前駆体、カルモジュリン結合タンパク候補CamBP64、システイン富裕、酸性統合膜タンパク前駆体、および、仮定的タンパクなど様々に記載されている。図113参照。

【0583】

P F 0 3 8 6 0 : D U F 3 2 6

このファミリーは、小型のシステイン富裕反復列である。システインは、多くの場合、C x x C x x x C x x C x x x C x x Cパターンに従う。もっともシステインは、しばしば、反復列の他の位置にも現れる。図114参照。

30

【0584】

P F 0 2 3 6 3 : システイン富裕反復列

このシステイン反復列C x x x C x x x C x x x Cは、このファミリーの配列の中で、O17970__CAEELでは34回繰り返される。この反復列の機能は、それらの出現するタンパクの機能と同様、不明である。このファミリーの配列の多くは、*C. elegans*（線虫）由来である。

40

【0585】

図115-116参照。

【0586】

【化 1 0 6】

名称	スカフォールド	Cys	無作為化	差異	サイズ	質 %
LMP0020	CB	8	29 AA	1027	2.6x107	78
LMP0021	CB	8	29 AA	1027	6.3x109	65
LMS0040	CB	8	16 AA	1019	2.9x108	77
LMS0041	CB	8	16 AA	1014	na	設計された
LMP0040	TF	8	4x7 AA	109	na	設計された
LMB0030	PL	8	13 AA	1012	na	設計された
LMP0030	PL	8	8 AA	109	na	設計された
LMP0010	TB	6	23 AA	1027	7.6x108	87
LMS0043	TB	6	14 AA	1018	5.1x109	92
LMS0044	TB	6	14 AA	1013	1.0x109	96
LMB0020	TI	6	10 AA	1012	2.4x109	92
LMB0010	BC	4	12 AA	1014	na	設計された
LMP0050	BC	4	8 AA	109	7.9x108	100

10

【 0 5 8 7】

20

【化 1 0 7】

参考文献:

- [00876] Artavanis-Tsokanas, S et al. (1995) *Science* 268:225-232.
- [00877] Aster, JC et al. (1999) *Biochemistry* 38:4736.
- [00878] Bensch KW et al. (1995) *FEBS Lett* 368:331-335.
- [00879] Bork, P (1993) *FEBS Lett* 327:125-30
- [00880] Carr, MD et al. (1994) *PNAS* 91:2206-2210.
- [00881] Chirino AJ, Ary ML, Marshall SA. (2004) Minimizing the immunogenicity of protein therapeutics. *Drug Discovery Today* 9:82-90
- [00882] Chong JM et al. (2001) *J. Biol. Chem.* 277:5134-5144.
- [00883] Chong, JM and Speicher, DW (2001) *J. Biol. Chem.* 276:5804-5813.
- [00884] Conticello SG, Gilad Y, Avidan N, Ben-Asher E, Levy Z, Fainzilber M. (2001) Mechanisms for evolving hypervariability: the case of conopeptides. *Mol Biol Evol.* 18:120-31.
- [00885] Cornet B et al (1995) *Structure* 3:435-448.
- [00886] DeA, et al. (1994) *PNAS* 91:1084-1088
- [00887] Dufton MJ (1984) *J. Mol. Evol.* 20:128-134.
- [00888] Fajloun, Z et al (2000) *J. Biol. Chem.* 275:39394-402.
- [00889] Fitzgerald, K et al. (1995) *Development* 121:4275-82.
- [00890] Gray WR et al (1988) *Annu Rev Biochem* 57:665-700.
- [00891] Guncar G et al (1999) *EMBO J* 18:793-803.
- [00892] Hermeling S, Crommelin DJ, Schellekens H, Jiskoot W. (2004) Structure-immunogenicity relationships of therapeutic proteins. *Pharm Res.* 21, 897-903
- [00893] Higgins, JM et al. (1995) *J. Immunol.* 155:5777-85
- [00894] Hoffman, W et al. (1993) *Trends Biochem Sci* 18:239-243.
- [00895] Hugli, TB (1990) *Curr Topics Microbiol Immunol.* 153:181-208.
- [00896] Jonassen I et al (1995) *Protein Sci* 4:1587-1595.
- [00897] Kamikubo, Y et al (2004)
- [00898] Kim, JI et al (1995) *J. Mol. Biol.* 250:659-671.
- [00899] Kimble, J et al.(1997) *Annu Rev Cell Dev Biol* 13:333-361.
- [00900] Koduri, V & Blacklow, SC (2001) 40:12801
- [00901] Lauber, T. et al (2003) *J. Mol. Biol.* 328:205-219.
- [00902] Léonetti et al. (1998) *J. Immunol.* 160; 3820-3827 (1998)

10

20

30

【 0 5 8 8】

【化 1 0 8】

- [00903] Léonetti M, Thai R, Cotton J, Leroy S, Drevet P, Ducancel F, Boulain JC, Ménez A. (1998) Increasing immunogenicity of antigens fused to Ig-binding proteins by cell surface targeting. *J. Immunol.*, **160**; 3820-3827.
- [00904] Leung-Hagesteijn, C et al. (1992) *Cell* 71:289-99
- [00905] Liu L et al (1997) *Genomics* 43:316-320.
- [00906] Maillère B, Mourier G, Hervé M, Cotton J, Leroy S, Ménez A. (1995) Immunogenicity of a disulphide-containing neurotoxin: presentation to T-cells requires a reduction step. *Toxicon*, **4**, 475-482; Maillère B. et al., unpublished data.
- [00907] Maillère, B., Cotton, J., Mourier, G., Léonetti, M., Leroy, S. and Ménez, A. (1993). Role of thiols in the presentation of a snake toxin to murine T cells. *J. Immunol.* **150**:5270-5280. 10
- [00908] Martin L, Stricher F, Misse D, Sironi F, Pugniere M, Barthe P, Prado-Gotor R, Freulon I, Magne X, Roumestand C, Ménez A, Lusso P, Veas F, Vita C (2003) Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes. *Nat Biotechnol.* **21**:71-6.
- [00909] Ménez, A. (1991) Immunology of snake toxins, p. 35-90. In: Snake Toxins. AL Harvey (Ed), Pergamon Press, Inc., New York.
- [00910] Miljanich, G.P. (2004), Ziconotide: neuronal calcium channel blocker for treating severe chronic pain. *Curr. Med. Chem.* **23**, 3029. 20
- [00911] Misenheimer, TM et al. (2001) *J. Biol. Chem.* 276:45882
- [00912] Molina F et al (1996) *Eur. J. Biochem.* 240:125-133.
- [00913] Mourier et al. (1995) *Toxicon* 4:475-482.
- [00914] Nielsen, KJ et al (2002) *J. Biol. Chem.* 277:27247-27255.
- [00915] Pallaghy PK et al (1993) *J. Mol Biol* 234:405-420.
- [00916] Pallaghy, P et al. *Protein Sci* 3:1833 (1994)
- [00917] Pan, TC et al. (1993) *J. Cell. Biol.* 123: 1269-1277
- [00918] Patten, P.A. and Schellekens, H. (2003) The immunogenicity of Biopharmaceuticals. In: Immunogenicity of Therapeutic Biological Products. Brown, F. and Mire-Sluis, A.R. (eds). Dev. Biol. Basel, Karger, **112**:81-97. 30
- [00919] Pereira, C.M., Guth, B.E.C., Sbrogio-Almeida, M.E. and Castilho, B.A. (2001) *Microbiology* 147:861-867.
- [00920] Petersen, SV et al (2003) *Proc. Natl. Acad. Sci. USA* 100:13875-80.
- [00921] Rebayl, et al. (1991) *Cell* 67:687-699
- [00922] Roszmusz, E. et al. (2002) *BBRC* 296:156
- [00923] Sands, BE & Podolsky, DK (1996) *Annu. Rev. Physiol.* 58:253-273.
- [00924] Schultz-Cherry, S et al. (1995) *J. Biol. Chem.* 270:7304-7310
- [00925] Schultz-Cherry, S et al. J. (1994) *J. Biol. Chem.* 269:26783-8
- [00926] Schulz A. et al (2005) *Biopolymers* 80:34-49.
- Singh H, Raghava GP (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics* **17**: 1236-7. 40
- [00927] Skinner WS et al, *J. Biol. Chem.* (1989) 264:2150-2155.
- [00928] So, T., Ito, H., Hirata, M., Ueda, T. and Imoto, T. (2001) Contribution of conformational stability of hen lysozyme to induction of type 2 T-helper immune responses. *Immunology* **104**:259-268.
- [00929] Sturniolo, T., et al. (1999) Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nature Biotechnol.*, **17**: 555
- [00930] Tam, JP and Lu, YA. *Protein Sci.* 7:1583 (1998)
- [00931] Tax, FE et al. (1994) *Nature* 368:150-154.

【 0 5 8 9】

【化 1 0 9】

- [00932] Thai R, Moine G, Desmadril M, Servent D, Tarride JL, Ménez A, Léonetti M. (2004) Antigen stability controls antigen presentation. *J. Biol. Chem.* 279, 50257-50266.
- [00933] Van den Hooven, HW et al. (2001) *Biochemistry* 40:3458-3466.
- [00934] van Vlijmen HW, Gupta A, Narasimhan S, Singh J (2004). A novel database of disulfide patterns and its application to the discovery of distantly related homologs. *J Mol Biol* 335: 1083-92.
- [00935] Vardar, D et al. (2003) *Biochemistry* 42:7061
- [00936] White, CE et al. (1996) *PNAS* 93:10177.
- [00937] Xu Y et al (2000) *Biochemistry* 39:13669-13675.
- [00938] Zaffarella GC et al (1988) *Biochemistry* 27:7102-7105.
- [00939] Zhu S et al (1999) *FEBS Lett* 457:509-514.
- [00940] Zuiderweg, ER et al. (1989) *Biochemistry* 28:172-85.

10

【図面の簡単な説明】

【 0 5 9 0 】

【図 1】図 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 のモチーフ：

【化 1 1 0】

20

- 1) CxPhxxxCxxxxdCCxxxCxxxGxxxxxC
 - 2) CxPxxxxCxxxxCCxxxCxxxxGxxxxxC
 - 3) CxxxxxCxxxxCCxxxCxxxxxxxxxC
- CDP: C6C5C0C3C10C

【図 2】図 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 のモチーフ：

【化 1 1 1】

- 1) fCCPxxxrCCw
 - 2) CCPxxxxCCW
 - 3) CCxxxxCC
- CDP: C0C5C0C

30

【図 3】図 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 のモチーフ：

【化 1 1 2】

- 1) CxxxWxCxxxxCCgWxxCxxgxC
 - 2) CxxxWxCxxxxCCxWxxCxxxxC
 - 3) CxxxxxCxxxxCCxxxCxxxxC
- CDP: C6C5C0C4C4

40

【図 4】図 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 のモチーフ：

【化 1 1 3】

- 1) CxgydxxCxxxpCxxxxxxxxCxxxgyWWyxxxY
 - 2) CxxxxxCxxxxCCxxxxxxxxCxxxxWWxxxxC
 - 3) CxxxxxCxxxxCCxxxxxxxxCxxxxxxxxxC
- CDP: C6C5C0C7C13C

【図 5】図 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを

50

示す。図 5 のモチーフ：

【化 1 1 4】

- 1) CxfxCxxxxgxxpCxxxxxxxxxxxxxxxxCxxgWxCxxxxC
 - 2) CxxxCxxxxxxxxCxxxxxxxxxxxxxxxxCxxxWxCxxxxC
 - 3) CxxxCxxxxxxxxCxxxxxxxxxxxxxxxxCxxxxCxxxxC
- CDP: C3C9C17C5C4C

【図 6】図 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 のモチーフ：

【化 1 1 5】

- 1) CxxxxxCxxHxxCCxxxCxxgxCxxxxwxxxgC
 - 2) CxxxxxCxxHxxCCxxxCxxxxCxxxxxxxxxC
 - 3) CxxxxxCxxxxxCxxxCxxxxCxxxxxxxxxC
- CDP: C6C5C0C3C4C10C

【図 7】図 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 のモチーフ：

【化 1 1 6】

- 1) CxxxgxxCxxdgcCxxgxCxxfxgxxC
 - 2) CxxxxxCxxxxCCxxxCxxxxxxxxC
- CDP: C6C5C0C3C8C

【図 8】図 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 のモチーフ：

【化 1 1 7】

- 1) CxdxxCxyCgxyxxgxCdgpxxCx
 - 2) CxxxxCxxxCxxxxxxxxxCxxxxCx
- CDP: C4C3C9C5C1C

【図 9】図 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 のモチーフ：

【化 1 1 8】

- 1) ChfxxCxxdCfxxPGxyGxCxxxxGxxCx
 - 2) CxxxxCxxxCxxxxPGxxGxCxxxxGxxCx
 - 3) CxxxxCxxxCxxxxxxxxCxxxxxxxxCx
- CDP: C4C3C10C8C1C

【図 10】図 10 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 10 のモチーフ：

【化 1 1 9】

- 1) CixxgxxCxG(xx)xxxxCxCxxxxyCxCxxx(xx)FG(x)xxxxCxC(x)xxxxCxxxxxx(x)xxxxC
- 2) CxxxxxCxG(xx)xxxxCxCxxxxCxCxxx(xx)FG(x)xxxxCxC(x)xxxxCxxxxxx(x)xxxxC
- 3) CxxxxxCxx(xx)xxxxCxCxxxxCxCxxx(xx)xx(x)xxxxCxC(x)xxxxCxxxxxx(x)xxxxC

【図 11】図 11 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。

10

20

30

40

【化 1 2 0】

- 1) CxPCftbxxxxxxCxxCCxxx(x)gxgCxxxxqCx C
- 2) CxPCxxxxxxxxxxCxxCCxxx(x)xxx CxxxxCx C
- 3) CxxCxxxxxxxxxxCxxCCxxx(x)xxx CxxxxCx C

CDP: C2C10C2C0C6(7)C4C1C

【図 1 2】図 1 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 1 のモチーフ：

【化 1 2 1】

CxxxxxxxxCxxxxxxxxCCxxx CxxxxC

CDP: C6C6C0C3C4C

10

【図 1 3】図 1 3 は、タンパクにおけるアミノ酸の出現プロフィールを示す。

【図 1 4】図 1 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 のモチーフ：

【化 1 2 2】

- 1) Cxx(x)xCxxxxxxxxxCxCxxx CxxxxCCxxxxxxxxC
- 2) Cxx(x)RCxExxxxxxxxxCxCxxx CxxxxCCxD[yf]xxx C

CDP: C3-4C10C1C3C5C6C

20

【図 1 5】図 1 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 のモチーフ：

【化 1 2 3】

- 1) Cxxxx(x)x(x)xxxxxCpxgxxx C[yf]xlxxxx(xx)CxxxxxxxxG CxttCPxxxx(x)xxxxxC CxtdxCN
- 2) Cxxxx(x)x(x)xxxxxCxxxxxCxxxxxxxx(xx)CxxxxxxxxG CxxxCPxxxx(x)xxxxxC CxxxxCN
- 3) Cxxxx(x)x(x)xxxxxCxxxxxCxxxxxxxx(xx)CxxxxxxxxCxxx Cxxxx(x)xxxxxC CxxxxC

CDP: C6-8C6C7-9C10C3C10-11C0C4C

【図 1 6】図 1 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 のモチーフ：

【化 1 2 4】

1) CxxCxxxxxxxxC(xxx)xxxxxxxxCxxxxxxxxxxxxxxxxxxxxCxxx(xx)x C(p)xx(x)xxxxxxxxxx(x)xxxxxC Cxx
xxC

【図 1 7】図 1 7 は、例示配列の一次および二次構造を示す。

【図 1 8】図 1 8 は、例示配列の一次および二次構造を示す。

【図 1 9】図 1 9 は、は、各種無脊椎動物および植物タンパク間における配列整列を示す。

【図 2 0】図 2 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 0 のモチーフ：

【化 1 2 5】

- 1) CgxqxxxxxCxxxxCCsxxGxCGxxxxxCxx(x)x Cx(x)xxC
- 2) CxxxxxxxxCxxxxCCxxxxxCxxxxxCxx(x)x Cx(x)xxC

CDP: C8C4C0C5C6C3-4C3-4C

40

【図 2 1】図 2 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 1 のモチーフ：

【化 1 2 6】

- 1) Cxxx(x)xxxxxxxx(xx)xxxC(x)xxxxCxxx(x)xxxCxxxxxxxxxxxxCxxx(xx)xC
 - 2) Cxxx(x)xxxxxxxx(xx)xxxC(xx)[yf]xxCxxxx(x)xxxCxxxx[yf]xxxxxxxxCxxxx(xx)xC
- CDP: C13-16C5-6C9-10C12C7-9C

【図 2 2】図 2 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 2 のモチーフ：

【化 1 2 7】

- 1) C(xx)xY(gg)xxxxxCxxxCxx(x)xxxCxxxCxx(x)xxgaxgxCxxx(x)xxxxxC[wylf]C
 - 2) C(xx)xx(xx)xxxxxxxxCxxxCxx(x)xxxCxxxCxx(x)xxxxxxxxCxxx(x)xxxxxCxC
- CDP: C8-12C3C5-6C3C9-10C9-10C1C

10

【図 2 3】図 2 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 3 のモチーフ：

【化 1 2 8】

- 1) CxxxxxxxxCxxxCxxxCxxxx(xxxx)xxxCxxxx(xxxx)xxCxxxxCxCxxxxxxxxxxx(x)xCxxxxC
 - 2) CpxxxxxxxxxCxxxCxxxCxxxx(xxxx)xxxCxxxx(xxxx)xxCxxxxCxCxxgxxxxxxxx(x)xCvxxxxC
- CDP: C8C3C3C8-12C6-10C4C1C

【図 2 4】図 2 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 4 のモチーフ：

20

【化 1 2 9】

- 1) CxxxCxxxxxxxxCPxxxx(x)xxxxCxxCCxxxxCxxxxxxxxxxxC
 - 2) CtxxCdxxxxxxxxCPxxxx(xx)xxxxCxxCCxxgxGCx[yfl][yfl]xxxxGxx[ivl]C
- CDP: C3C8C11-12C2C0C5C10C

【図 2 5】図 2 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 5 のモチーフ：

【化 1 3 0】

- 1) CxxxxSxx[Fwy]xGxCxxxxCxxxCxxxxx(xx)xGxCxx(xx)xx[tk]CxCxxxC
 - 2) CxxxxSxxFxGxCxxxxCxxxCxxxxxx(xx)xGxCxx(xx)xxxxCxCxxxC
 - 3) CxxxxxxxxCxxxCxxxCxxxxxx(xx)xxxCxx(xx)xxxxCxCxxxC
- CDP: C11C5C3C9-11C6-8C1C3C

30

【図 2 6】図 2 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 6 のモチーフ：

【化 1 3 1】

- C(xxx)xxxxxxxxCCxx(x)xCxx(xx)xxxC
- CDP: C6-9C0C4-5C5-7C

40

【図 2 7】図 2 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 7 のモチーフ：

【化 1 3 2】

- 1) CxxxCxshxxCxxxCxCxxx[xc]x[xc]

【図 2 8】図 2 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 8 のモチーフ：

【化 1 3 3】

1) CxgrxxrCprrCCxgxxCrgxxxxC
 2) CxxxxxxCxxxCCxxxCxxxxxxC
 CDP: C6C3C0C4C7C

【図 2 9】図 2 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 2 9 のモチーフ：

【化 1 3 4】

1) CCxxrrxCxxixCxrxxCC
 2) CCxxxxxCxxxCxxxxCC
 CDP: C0C5C4C4C0C

10

【図 3 0】図 3 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 0 のモチーフ：

【化 1 3 5】

1) CCgyprxxCpCxCxxirxyC
 2) CCxxxxxxxxCxCxxxxxxxC
 CDP: C0C7C2C1C7C

【図 3 1】図 3 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 1 のモチーフ：

20

【化 1 3 6】

1) CxxtGxxCxxxx[ex]Csx(x)Ga[ex]sxxFxxC
 2) CxxxxxCxxxx[ex]Cxx(x)xx[ex]xxxxxC

【図 3 2】図 3 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 2 のモチーフ：

【化 1 3 7】

1) CxxxxC(x)xxxCxxGxxxDxxgCxx(xx)xCxC
 2) CxxxxC(x)xxxCxxxxxxxxxxCxx(xx)xCxC
 CDP: C4C3-4C10C2-4C1C

30

【図 3 3】図 3 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 3 のモチーフ：

【化 1 3 8】

1) CxxxxxCDDPCaxCxCRFFxxxCxCR
 2) CxxxxxCxxCxxCxCCCCxxCx
 CDP: C6C0C2C2C1C6C1C

【図 3 4】図 3 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 4 のモチーフ：

40

【化 1 3 9】

1) CxpgxxxxCNxCxxxx(x)xxxTxxxC
 2) CxxxxxxxxCNxCxxxx(x)xxxTxxxC
 3) CxxxxxxxxCxxCxxxxx(x)xxxxxxxxC
 CDP: C9C2C1C11-12C

【図 3 5】図 3 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 5 のモチーフ：

50

【化 1 4 0】

- 1) Cxx(xx)xxxxxCxxxxxx(x)CxxxxxxxxxxxxCxxxC
 - 2) Cxx(xx)DxxxxCxxxxxx(x)CxxxxxxxxxxxxCxxxC
 - 3) Cxx(xx)DxxxxCxx[wylfim]xxxx(x)CxxxxxxxxxxxxCxxxC
- CDP: C7-9C7-8C12C3C2C

【図 3 6】図 3 6 は、各種無脊椎動物および植物タンパク間における配列整列を示す。

【図 3 7】図 3 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 7 のモチーフ：

【化 1 4 1】

10

- 1) C(xxxx)CxxxxCxx(xxxxxx)xxxCxCxxx(xx)xxxxxC
 - 2) C(xxxx)CxxxGxCxx(xxxxxx)xxxCxCxxx(xx)xxGxxC
 - 3) C(xxxx)CxxxGxCxx(xxxxxx)xxxCxCxxx(xx)[ywflh]xGxxC
- CDP: C0-4C5C6-13C1C9-11C

【図 3 8】図 3 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 8 のモチーフ：

【化 1 4 2】

20

- 1) Cxxxx(x)xCxxxxCxxxx(xx)xxxCxCxxx(xx)xxxxxC
 - 2) Cxxxx(x)xCxxxgxCxxxx(xx)xxxCxCxxg(xxx)xxxgxxC
- CDP: C5-6C5C8-10C1C9-12C

【図 3 9】図 3 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 3 9 のモチーフ：

【化 1 4 3】

- 1) CxCxxxxxxxx(xx)xxCxxx(xxxxxxx)xxxxxCxCxxxxxxxxCxCxxxxxxxx(xx)xxxxxC
 - 2) CxCxxxxxxxx(xx)xxCxxx(xxxxxxx)xxxxGxCxCxxxxGxxCxCxxxxxxxx(xx)xxxxxC
- CDP: C1C9-11C9-17C1C8C2C14-16C

【図 4 0】図 4 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 0 のモチーフ：

30

【化 1 4 4】

- 1) DxxECxxxxxCx(xx)xxxxxCNxxGx[fy]xCx(xxx)xCxxg[yf]x(xxxx)xxxxxC
 - 2) DxxECxxxxxCx(xx)xxxxxCNxxGxxxCx(xxx)xCxxxxx(xxxx)xxxxxC
 - 3) CxxxxxCx(xx)xxxxxCxxxxxxxxC(xxx)xCxxxxx(xxx)xxxxxC
- CDP: C6C6-8C8C2-5C12-16C

【図 4 1】図 4 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 1 のモチーフ：

【化 1 4 5】

40

- 1) CsxHGxxxxDGxx(x)xxGxxPxCeCxxCyxGxxCsxxxxxC
 - 2) CxxHGxxxxDGxx(x)xxGxxPxCxCxxCxxGxxCxxxxxC
 - 3) Cxxxxxxxxxx(x)xxxxxCxCxxCxxxxCxxxxxC
- CDP: C19-20C1C2C5C6C

【図 4 2】図 4 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 2 のモチーフ：

【化 1 4 6】

- 1) CxxxxGxCRxkxxxxCxxxxxxxxCxnxxqkCC
 - 2) CxxxxGxCRxkxxxxCxxxxxxxxCxxxxxxxxCC
 - 3) CxxxxxxxxCxxxxxxxxCxxxxxxxxCxxxxxxxxCC
- CDP: C6C7C7C6C0C

【図 4 3】図 4 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 3 のモチーフ：

【化 1 4 7】

- 1) CxxxxxxxxCxxxxCxxxxxxxxCxxxxxxxxCC
 - 2) Cxxxxgx CxxxxCxxxxxxxxgx CxxxxxxxxCC
- CDP: C6C4C9C6C0C

10

【図 4 4】図 4 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 4 のモチーフ：

【化 1 4 8】

- 1) CxxHCxxxgxxxgxxCxx(xxx)xxx CxC
 - 2) CxxHCxxxxxxxxCxx(xxx)xxx CxC
 - 3) Cxxx Cxxxxxxxx Cxx(xxx)xxx CxC
- CDP: C3C8C5-8C1C

20

【図 4 5】図 4 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 5 のモチーフ：

【化 1 4 9】

- 1) CxCRCxxx CxxxExxxGxCxxxxxxxx[yfh]x[yfl]CC
 - 2) CxCRCxxx CxxxExxxGxCxxxxxxxxCC
 - 3) CxCxxx Cxxxxxxxx CxxxxxxxxCC
- CDP: C1C3C9C9C0C

30

【図 4 6】図 4 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 6 のモチーフ：

【化 1 5 0】

- 1) CCxxxxRxx[yf]nxCrxxGxxxxCaxxxx CxiisgxxC
 - 2) CCxxxxRxxxx CxxxGxxxx Cxxxx CxxxxxxxxC
 - 3) CCxxxxxxxx Cxxxxxxxx Cxxxx CxxxxxxxxC
- CDP: C0C11C9C5C7C

【図 4 7】図 4 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 7 のモチーフ：

【化 1 5 1】

- 1) Cxxaxxx Cxxxx Cxxx Cxx(x)xxxx Cxxx[vi]xx(x)xxC
- 2) Cxxxx Cxxxx Cxxx Cxx(x)xxxx Cxxxxxx(x)xxC

40

【図 4 8】図 4 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 8 のモチーフ：

【化 1 5 2】

1) Cxxxxxxxx(x)xxxxxCCCxxxx(x)xxxxxxxxCxxC
 2) Cxxxxxxxx(x)xxkxxCCCxxxx(x)xx[wfiv]gxxCexC
 CDP: C12-13C0C0C10-11C2C

【図 4 9】図 4 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 4 9 のモチーフ：

【化 1 5 3】

1) Cxxxxxxxx[yfh]xxxxxWxxxx(xxxx)xxxCx(x)xCxCxx(xxxxxxxxx)xxxxCxxxxCxx(yyyy)xxCxxx(xxx)xxxxxxxxgcC
 Cx(xx)xC
 2) CxxxxxxxxxxxxxWxxxx(xxxx)xxxCx(x)xCxCxx(xxxxxxxxx)xxxxCxxxxCxx(xxxx)xxCxxx(xxx)xxxxxxxxxCx(x
 x)xC
 3) Cxxxxxxxxxxxxxxxxxxxx(xxxx)xxxCx(x)xCxCxx(xxxxxxxxx)xxxxCxxxxCxx(xxxx)xxCxxx(xxx)xxxxxxxxxCx(xx
)xC

10

【図 5 0】図 5 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 0 のモチーフ：

【化 1 5 4】

1) CxxxxxxCxxxxxCxxxxCxxx(xxx)x(xx)x[wylfi]C
 2) CxxxxxxCxxxxxCxxxxCxxx(xxx)x(xx)xxC
 CDP: C6C5C0C4C6-11C

20

【図 5 1】図 5 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 1 のモチーフ：

【化 1 5 5】

1) CxxxCvxxxxCxxxxxxxxGCxCxxxvC
 2) CxxxCxxxxCxxxxxxxxCxCxxxxC
 CDP: C3C4C7C1C4C

【図 5 2】図 5 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 2 のモチーフ：

【化 1 5 6】

1) CxfCCxCCxxxxCgxCC
 2) CxxCCxCCxxxxCxxCC
 CDP: C2C0C1C4C2C0C

30

【図 5 3】図 5 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 3 のモチーフ：

【化 1 5 7】

1) CxxxxxWCgxxedCCcpmxCxxxWyxqgxCqxxxxxxxxkxxC
 2) CxxxxxWCxxxxxCCCxxxCxxxWxxxxxxxxCxxxxxxxxxxxC
 3) CxxxxxxCxxxxxCxxxCxxxxxxxxxxCxxxxxxxxxxxC
 CDP: C6C5C0C0C3C10C12C

40

【図 5 4】図 5 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 4 のモチーフ：

【化 1 5 8】

1) CxxCxxxCxxxxxxxxCxxx(xx)xCxC

【図 5 5】図 5 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 5 のモチーフ：

50

【化 1 5 9】

1) CxxxxxCxxxCxxxx(x)xxxxCxxxxCx

2) CxxxxxCxxxCxxxx(x)xxxgkCxxxkCx

CDP: C5C3C10-11C4C1C

【図 5 6】図 5 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 6 のモチーフ：

【化 1 6 0】

1) CPxxxxxCxxdxdCxxxCxCxxx(x)x

2) CPxxxxxCxxxxCxxxCxCxxx(x)x

2) CxxxxxCxxxxCxxxCxCxxx(x)x

CDP: C6C5C3C1C5-6C

10

【図 5 7】図 5 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 7 のモチーフ：

【化 1 6 1】

1) CCxdgxxxx(x)xxxxCxxxxxxxxCxxfxxCC

2) Cxxxxxxxx(x)xxxxCxxxxxxxxCxxxixCC

CDP: C0C12-13C12C6C0C

20

【図 5 8】図 5 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 8 のモチーフ：

【化 1 6 2】

1) CxxxxPCxxxxCCxgxCxxxWxCxxxxxCskxC

2) CxxxxPCxxxxCCxxxCxxxWxCxxxxxCxxx

3) CxxxxxCxxxxCCxxxCxxxxxCxxxxxCxxx

CDP: C6C5C0C3C6C6C3C

【図 5 9】図 5 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 5 9 のモチーフ：

【化 1 6 3】

1) CxxWx[wylf]xxCxxxxdCgxgxrexx(xx)CxxxxxxxxCxxPC

2) CxxWxxxCxxxxxCxxxxxx(xx)CxxxxxxxxCxxPC

3) CxxxxxxxxCxxxxxCxxxxxx(xx)CxxxxxxxxCxxx

CDP: C7C6C8-10C8C3C

30

【図 6 0】図 6 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 0 のモチーフ：

【化 1 6 4】

1) CdxxxCxygyxxCxxCCxxxgxxxgxCxxxCx

2) CxxxxCxxxxxCxxCCxxxxxxxxCxxxCx

CDP: C5C8C2C0C9C4C1C

40

【図 6 1】図 6 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 1 のモチーフ：

【化 1 6 5】

- 1) Cxxxxx(x)x(x)xxxxxCpxgxxxC[yf]x|xxxx(x)x)CxxxxxxxxGxxtCPxxxx(x)xxxxxCxXdxC
 - 2) Cxxxxx(x)x(x)xxxxxCxxxxxCxxxxxxx(x)x)CxxxxxxxxGxxtCPxxxx(x)xxxxxCxxxxC
 - 3) Cxxxxx(x)x(x)xxxxxCxxxxxCxxxxxxx(x)x)CxxxxxxxxCxxxGxxxx(x)xxxxxCxxxxC
- CDP: C11-13C6C7-9C10C3C10-11C0C4C

【図 6 2】図 6 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 2 のモチーフ：

【化 1 6 6】

- 1) CPxxx(xx)xxxxxCxxx(xxx)CxxDxxxkCCxxxCxxxC
 - 2) CPxxx(xx)xxxxxCxxx(xxx)CxxDxxxkCCxxxCxxxC
 - 3) Cxxxx(xx)xxxxxCxxx(xxx)CxxxxxCxxxxCCxxxCxxxC
- CDP: C9-11C3-6C5C5C0C3C3C

10

【図 6 3】図 6 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 3 のモチーフ：

【化 1 6 7】

- 1) Cxx(x)xyxxCxxgxxxCCxxl(x)x)Cx CxxxxNCxC
 - 2) Cxx(x)xxxxCxxxxxCxxx(x)x)Cx CxxxxNCxC
 - 3) Cxx(x)xxxxCxxxxxCxxx(x)x)Cx CxxxxCx C
- CDP: C6-7C6C0C4-5C1C6C1C

20

【図 6 4】図 6 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 4 のモチーフ：

【化 1 6 8】

- 1) CxxxxxCxdWxxxxCCxgxyCx CxxpxCx C
 - 2) CxxxxxCxxWxxxxCCxxxxCx CxxxxCx C
 - 3) CxxxxxCxxxxxCxxxxCx CxxxxCx C
- CDP: C6C7C0C4C1C5C1C

30

【図 6 5】図 6 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 5 のモチーフ：

【化 1 6 9】

- 1) CxxxCxxxydx CxxCxgxWxgxxgx Cxh CxxxxxCxxx C
 - 2) Cxxx CxxxxxCxx CxxxWxxxxxCxxx CxxxxxCxxx C
 - 3) Cxxx CxxxxxCxx Cxxxxxxxxxx Cxxx CxxxxxCxxx C
- CDP: C3C6C2C10C3C6C3C

【図 6 6】図 6 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 6 のモチーフ：

【化 1 7 0】

- 1) CxPxGxPCPyxxx Cxxx CxxxxxxgxxxxxC
 - 2) CxxxxxCxxxxCCxxx CxxxxxxxxxxxxxC
 - 3) CxPxGxPCPxxx Cxxx CxxxxxxxxxxxxxC
- CDP: C6C5C0C3C13C

40

【図 6 7】図 6 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 7 のモチーフ：

【化 1 7 1】

- 1) CxxxxxxxxxxxxCPxxxxxxxxCxGxxCgsWxxxxxxxxCxCxCxxdWxxxrCC
 - 2) CxxxxxxxxxxxxCPxxxxxxxxCxCxxCxWxxxxxxxxCxCxCxxxWxxxrCC
 - 3) CxxxxxxxxxxxxCxxxxxxxxCxCxxCxxxxxxxxCxCxCxxxxxxxxCC
- CDP: C11C8C1C3C10C1C1C9C0C

【図 6 8】図 6 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 8 のモチーフ：

【化 1 7 2】

- 1) Cx(xx)xxxCxxxx[nd]gxCx[wylf]DGxDC
 - 2) Cx(xx)xxxCxxxxxxxxCxDCGxDC
 - 3) Cx(xx)xxxCxxxxxxxxCxxxxxC
- CDP: C4-6C8C6C

10

【図 6 9】図 6 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 6 9 のモチーフ：

【化 1 7 3】

- 1) Cxxxx[yf]xx(xx)xxCxxCxxCxx(xx)gxxxxCxxxxxC
- 2) Cxxxxxx(xx)xxx(x)xxCxxCxxCxx(xx)xxxxxCxxxxxC

20

【図 7 0】図 7 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 0 のモチーフ：

【化 1 7 4】

- 1) CxPFx[yf]xxxxxxxxCxxxgxxxxxWCxtxxxxDxxxx[fy]C
 - 2) CxxPFxxxxxxxxCxxxxxxxxWCxxxxxxxxDxxxxxC
 - 3) CxxxxxxxxxxxxCxxxxxxxxCxxxxxxxxxxxxxC
- CDP: C13C11C14C

【図 7 1】図 7 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 1 のモチーフ：

【化 1 7 5】

- 1) Cxx(xx)xxxxyxCxxx(xx)xxxxdxxxxWgxxxxxC
 - 2) Cxx(xx)xxxxxCxxx(xx)xxxxxxxWxxxxxxxxC
 - 3) Cxx(xx)xxxxxCxxx(xx)xxxxxxxxxxxxxxxxxC
- CDP: C8-10C0C22-24C

30

【図 7 2】図 7 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 2 のモチーフ：

【化 1 7 6】

- 1) CCxxx(x)CxxxpxxxCG
 - 2) CCxxx(x)CxxxxxxxC
- CDP: C0C4-5C8C

40

【図 7 3】図 7 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 3 のモチーフ：

【化 1 7 7】

- 1) CGGxxxxGxxxCxxgxxC
 - 2) CGGxxxxGxxxCxxxxxC
- CDP: C10C5C

【図 7 4】図 7 4 は、例示配列の一次および二次構造を示す。

【図 7 5】図 7 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 5 のモチーフ：

【化 1 7 8】

- 1) Cx(xxc)xxxCxxxxxxxxCpxxx(xxxx)xxxx(c)xxxxxxxxGCgCCxxCxxxxgxxCxxxxxxxx(dx)xxglxCxxg(xx)xxxxlxC
- 2) Cx(xxc)xxxCxxxxxxxxCxxxx(xxxx)xxxx(c)xxxxxxxxGCxCCxxCxxxxxxxxCxxxxxxxx(xx)xxxxCxxx(xx)xxxxxxxxC
- 3) Cx(xxc)xxxCxxxxxxxxCxxxx(xxxx)xxxx(c)xxxxxxxxCxCxCCxxCxxxxxxxxCxxxxxxxx(xx)xxxxCxxx(xx)xxxxxxxxC

10

【図 7 6】図 7 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 6 のモチーフ：

【化 1 7 9】

- 1) CxCxxxxdkeCx[yfli]xChxd[ivl][ivl]W
 - 2) CxCxxxxdkeCx[yfli]xC
 - 3) CxCxxxxxxxxCxxxC
- CDP: C1C7C3C

20

【図 7 7】図 7 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 7 のモチーフ：

【化 1 8 0】

- 1) CExCxxxxaCtGC
 - 2) CExCxxxxxCxGC
 - 3) CxxCxxxxxCxxC
- CDP: C2C5C2C

【図 7 8】図 7 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 8 のモチーフ：

30

【化 1 8 1】

- 1) CxrxCWregxdeetCkerC
 - 2) CxxxCWxxxxxxxxCxxxC
- CDP: C3C9C3C

【図 7 9】図 7 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 7 9 のモチーフ：

【化 1 8 2】

- 1) DCxxxGxxCxGxxkxCCpxxxCxxYanxC
 - 2) CxxxGxxCxGxxxxCxxxxxCxxYxxxC
 - 3) CxxxxxCxxxxCCxxxxxCxxxxxC
- CDP: C6C5C0C5C6C

40

【図 8 0】図 8 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 0 のモチーフ：

【化 1 8 3】

- 1) CPx[ivlf]xxxCxxdxdCxxxCxCxxxxxCg
 - 2) CPxxxxxCxxxxxCxxxCxCxxxxxC
 - 3) CxxxxxCxxxxxCxxxCxCxxxxxC
- CDP: C6C5C3C1C6C

【図 8 1】図 8 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 1 のモチーフ：

【化 1 8 4】

- 1) CdxgeqCaxrkgxrgkxCdCPrgxxCnxllkC
 - 2) CxxxxxCxxxxxxxxxxxCxCxxxxxCxxxxxC
- CDP: C5C11C1C5C6C

10

【図 8 2】図 8 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 2 のモチーフ：

【化 1 8 5】

- 1) CvkkdelCxyyxdCCpxxCxxxxWWdhkC
 - 2) CxxxxxCxxxxxCxxxxxCxxxxWWxxxC
 - 3) CxxxxxCxxxxxCxxxxCxxxxxxxxxC
- CDP: C6C6C0C4C9C

20

【図 8 3】図 8 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 3 のモチーフ：

【化 1 8 6】

- 1) CxGxCsPFExPPCxxxCxCxPxllxxGxcxxPxxxxxxxxxHxnlCxssxCxkkxsGcFCxxYPNxxixxGWC
- 2) CxGxCsPFExPPCxxxxxCxCxPxllxxGxcxxPxxxxxxxxxHxxxCxxxxxCxxxxGxFCxxYPNxxxxxGWC
- 3) CxxxCxxxxxxxxCxxxCxCxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxCxxxxCxxxxxxxxCxxxxxxxxGxC

30

【図 8 4】図 8 4 は、例示配列の一次および二次構造を示す。

【図 8 5】図 8 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 5 のモチーフ：

【化 1 8 7】

- 1) CCPCxxCxYxxGCPWgqxxxxxC
 - 2) CCPCxxCxYxxGCPWgxxxxxC
 - 3) CCxCxxCxxxxCxxxxxxxxxC
- CDP: C0C1C2C5C10C

【図 8 6】図 8 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 6 のモチーフ：

40

【化 1 8 8】

- 1) CgxxgxRxxxxxxxxCxDCxNxxRxxxxxxxxCrxxCxxxxFxxC
 - 2) CxxxxxxxxRxxxxxxxxCxDCxNxxRxxxxxxxxCxxxCxxxxFxxC
 - 3) CxxxxxxxxxxxxxxxxCxCCCCxxxxxxxxCxxxCxxxxxxxxC
- CDP: C16C2C12C3C8C

【図 8 7】図 8 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 7 のモチーフ：

【化 1 8 9】

- 1) CxCxxxxPxxxxxxGxx(x)xxxxxC(x)xxxxWxxCxxxxxxxxCC
- 2) CxCxxxxPxxxxxxGxx(x)xxxxxC(x)xxxxWxxCxxxxxxxxCC
- 3) CxCxxxxxxxxxxxxxx(x)xxxxxC(x)xxxxxxxxCxxxxxxxxCC

CDP: C1C21-22C8-9C9C0C

【図 8 8】図 8 8 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 8 のモチーフ：

【化 1 9 0】

- 1) CxxxCxqCxxxxgxxfxgxxCaxsCxltxxGkxxPx C
- 2) CxxxCxxCxxxxxxxxxxxxCxxx CxxxxGxxxPx C
- 3) CxxxCxxCxxxxxxxxxxxxCxxx Cxxxxxxxx C

CDP: C3C2C12C3C10C

10

【図 8 9】図 8 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 8 9 のモチーフ：

【化 1 9 1】

- 1) CxxxCxxCxxxxxxxxxxxxCxlCxxxxxxxxWxxC
- 2) CxxxCxxCxxxxxxxxxxxxCxxx CxxxxxxxxWxxC
- 3) CxxxCxxCxxxxxxxxxxxxCxxx Cxxxxxxxx C

CDP: C3C2C15C3C12C

20

【図 9 0】図 9 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 0 のモチーフ：

【化 1 9 2】

- 1) CdxxxxxsxCqmxcccCxxaxxCxxxieeCktsxxexC
- 2) CxxxxxxxxCxxxxxCxxxxCxxxxxCxxxxxC

CDP: C8C6C5C6C7

【図 9 1】図 9 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 1 のモチーフ：

【化 1 9 3】

- 1) CxGxdrPCxxCCPCCPGxxCxxxexgxyC
- 2) CxGxxxPCxxCCPCCPGxxCxxxxxxxxxC
- 3) CxxxxxCxxCCxCxxxxCxxxxxxxxxC

CDP: C6C2C0C1C4C10C

30

【図 9 2】図 9 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 2 のモチーフ：

【化 1 9 4】

- 1) CxxxxxxxxCxxxxxxxxCxxxxxCxxxxxC
- 2) CgxxxxxCsxxgxyCwxxvCyxxxxCxlxC
- 3) CxxxxxCxxxxxxxxCxxxxxCxxxxxCxxx C

CDP: C6C0C6C5C6C3C

40

【図 9 3】図 9 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 3 のモチーフ：

【化 1 9 5】

1) CxxxxxCxxCxxxxxx(x)xCxWCxx(x)xxxCxxxx(xxxxxx)xCxxxx(xxxxxxxx)xxxxxC
 2) CxxxxxCxxCxxxxxx(x)xCxxCxx(x)xxxCxxxx(xxxxxx)xCxxxx(xxxxxxxx)xxxxxC
 CDP: C5C2C7-8C2C5-6C5-11C10-19C

【図 9 4】図 9 4 は、例示配列の一次および二次構造を示す。

【図 9 5】図 9 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 5 のモチーフ：

【化 1 9 6】

1) CxxxxxxxRxxCgxxxixxxCxxxgCCfdxxxxxxxxwC
 2) CxxxxxxxRxxCxxxxxxxixCxxxxCCxxxxxxxxxxC
 3) CxxxxxxxxxxCxxxxxxxixCxxxxCCxxxxxxxxxxC
 CDP: C10C9C4C0C10C

10

【図 9 6】図 9 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 6 のモチーフ：

【化 1 9 7】

1) CxvtCgxGxxxRxxxCxxxx(pxx)xxxxxCxxxxxx(xxx)xxxC(x)xxxxC
 2) CxxxCxxGxxxRxxxCxxxx(xxx)xxxxxCxxxxxx(xxx)xxxC(x)xxxxC
 3) CxxxCxxxxxxxixCxxxx(xxx)xxxxxCxxxxxx(xxx)xxxC(x)xxxxC
 CDP: C3C10C9-12C9-12C4-5C

20

【図 9 7】図 9 7 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 7 のモチーフ：

【化 1 9 8】

1) CxxCxCxx(x)sxppxCxCxDxxxx(x)C
 2) CxxCxCxx(x)xxxxxCxCxDxxxx(x)C
 3) CxxCxCxx(x)xxxxxCxCxxxxxx(x)C
 CDP: C2C1C7-8C1C6-7C

30

【図 9 8】図 9 8 は、例示配列の一次および二次構造を示す。

【図 9 9】図 9 9 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 9 9 のモチーフ：

【化 1 9 9】

1) CxxCGPxxxGxCxGPxiCCGxxxGxxxGxxxxxxCxxexxxxxPCxxxxxCxxxxGxCxxxGxCxxxxCxxdxxC
 2) CxxCGPxxxGxCxGPxxCCGxxxGxxxGxxxxxxCxxxxxxxPCxxxxxCxxxxGxCxxxGxCxxxxCxxxxxC
 3) CxxCxxxxxxxCxxxxxCxxxxxCxxxxxxxixCxxxxxxxixCxxxxxxCxxxxxCxxxxxCxxxxxCxxxxxC
 CDP: C2C7C5C0C5C9C9C6C6C5C0C4C5C

40

【図 1 0 0】図 1 0 0 は、例示配列の一次および二次構造を示す。

【図 1 0 1】図 1 0 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 0 1 のモチーフ：

【化 2 0 0】

1) CDCGxxxxC(xx)xxxCC(x)xxxxCxlxxxxCxx(xx)xxCCxx(x)xCxxxxxxxxCtxxxx(x)xCxxxxxCxGxxxxC
 2) CDCGxxxxC(xx)xxxCC(x)xxxxCxxxxxxCxx(xx)xxxCCxx(x)xCxxxxxxxxCxxxx(x)xCxxxxxCxGxxxxC
 3) CxCxxxxxC(xx)xxxCC(x)xxxxCxxxxxxCxx(xx)xxxCCxx(x)xCxxxxxxxxCxxxx(x)xCxxxxxCxxxxxC
 CDP: C1C5C3-5C0C4-5C7C4-6C0C1-3C8C6-7C5C6C

50

【図 1 0 2】図 1 0 2 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 0 2 のモチーフ：

【化 2 0 1】

1) CCxxxxgxxxCCPxxxxCCxDxxHCCPxxgxxCxxxxxxC
 2) CCxxxxxxxxCCPxxxxCCxDxxHCCPxxxxCxxxxxxC
 3) CCxxxxxxxxCCxxxxxxCCxxxxxxCCxxxxxxCxxxxxxC
 CDP: C0C8C0C6C0C5C0C5C6C

【図 1 0 3】図 1 0 3 は、グラニユリンの配列および三次構造を示す。

【図 1 0 4】図 1 0 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 0 4 のモチーフ：

【化 2 0 2】

1) Cap(tCxxxxCxxax)_n 2) Cap(xCxxxxxCxxxx)_n

10

【図 1 0 5】図 1 0 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 0 5 のモチーフ：

【化 2 0 3】

1) Cxx(x)Cxx(xxxx)xxxxCxxxx(xxxx)xxxxRCWxxxxxCQxxxxxCxxxCxx(x)xxCxxxxxxxCChxxCxxggCx(xx)xPxx
 (x)xxCxxCxxfxxxgxCxxxCP
 2) Cxx(x)Cxx(xxxx)xxxxCxxxx(xxxx)xxxxRCWxxxxxCQxxxxxCxxxCxx(x)xxCxxxxxxxCChxxCxxgxCx(xx)xPxx
 (x)xxCxxCxxxxxxxCxxxCP
 3) Cxx(x)Cxx(xxxx)xxxxCxxxx(xxxx)xxxxCxxxxxxxCxxxxxxCxxxCxx(x)xxCxxxxxxxCChxxCxxCx(xx)xxxx(x)
)xxCxxCxxxxxxxCxxxC

20

【図 1 0 6】図 1 0 6 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 0 6 のモチーフ：

【化 2 0 4】

1) xxx[wylf]xxxxCxCxCx
 2) xxxxxxxxCxCxCx

【図 1 0 7】図 1 0 7 は、C X C モチーフ反復列を示す。

30

【図 1 0 8】図 1 0 8 は、V E G F の C 末端ドメインおよびバルビアニ環分泌タンパクの配列を示す。

【図 1 0 9】図 1 0 9 は、システイン含有反復列の予想構造を示す。

【図 1 1 0】図 1 1 0 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 1 0 のモチーフ：

【化 2 0 5】

1) CxxxxxxxCxxxxxx(x)xxxxxCxx(x)xxxxxCxxxxxx(x)xxxxxCxxxxxxxCxx(x)xxxxxCxxCxxx(x)xCNxxxxxp
 xxxxCxqCxxgxxxx[ex]xxxxxxCxxxx(x)xxxxxCyxxxx(xxx)xxxxRGxxxxxx[ex]xxxxCxxC
 2) CxxxxxxxCxxxxxx(x)xxxxxCxx(x)xxxxxCxxxxxx(x)xxxxxCxxxxxxxCxx(x)xxxxxCxxCxxx(x)xCNxxxxxx
 xxxxCxxCxxxxxx[ex]xxxxxxCxxxxCxxxx(x)xxxxxCxxxxxx(xxx)xxxxRGxxxxxx[ex]xxxxCxxC
 3) CxxxxxxxCxxxxxx(x)xxxxxCxx(x)xxxxxCxxxxxx(x)xxxxxCxxxxxxxCxx(x)xxxxxCxxCxxx(x)xCxxxxxx
 xxxxCxxCxxxxxx[ex]xxxxxxCxxxxCxxxx(x)xxxxxCxxxxxx(xxx)xxxxxCxxxxxx[ex]xxxxCxxC

40

【図 1 1 1】図 1 1 1 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 1 1 のモチーフ：

【化 2 0 6】

xxxxxCxxxxxC)Chxx(xx)xxg(x)xxCxxxxxCxyxxxxxCxxxx(xx)xxxxCxWxxxx(x)xxCxxxx(xx)Cx
 xxxxCxxxxxxC)Cxxxx(xx)xx(x)xxCxxxxxCxxxxxxCxxxx(xx)xxxxCxWxxxx(x)xxCxxxx(xx)Cx
 xxxxCxxxxxxC)Cxxxx(xx)xx(x)xxCxxxxxCxxxxxxCxxxx(xx)xxxxCxWxxxx(x)xxCxxxx(xx)Cx

50

【図 1 1 2】図 1 1 2 は、例示のシステイン含有反復タンパクの配列を示す。

【図 1 1 3】図 1 1 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 1 3 のモチーフ：

【化 2 0 7】

1) nxCtxdxCxxxxgCxxxxxCxxx

2) CxxxxCxxxxxCxxxxxCxxx

CDP: C4C5C6C3

【図 1 1 4】図 1 1 4 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 1 4 のモチーフ：

【化 2 0 8】

xxxx[xx]xxCxxx[xx]xxCxxxCxxxx

【図 1 1 5】図 1 1 5 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。

【化 2 0 9】

xxCxxxCxxxCxx(x)xCxx CDP: 2C3C3C3-4C2

【図 1 1 6】図 1 1 6 は、例示のシステイン含有反復タンパクの配列を示す。

【図 1 1 7】図 1 1 7 は、例示の抗凍結タンパクの構造を示す。

【図 1 1 8】図 1 1 8 は、エラプトキシンの構造を示す。

【図 1 1 9】図 1 1 9 は、プレキシンの構造を示す。

【図 1 2 0】図 1 2 0 は、プレキシンの配列を示す。

【図 1 2 1】図 1 2 1 は、ソマトメチンの構造を示す。

【図 1 2 2】図 1 2 2 は、分子量に従って発現ミクロタンパクを分離する SDS-PAGE ゲルを示す。

【図 1 2 3】図 1 2 3 は、遺伝子パッケージの中に含まれる各種スカフォールドおよびモチーフを示す。図 1 2 3 のモチーフ：

【化 2 1 0】

1) CtxxGxxxC(vilm)CxGxxxCGxGxxCxxxxxCxxxnC

2) CxxxGxxxCxGxxxCGxGxxCxxxxxCxxxnC

3) CxxxxxCxCxxxxxCxxxxxCxxxxxCxxxxxC

CDP: C7C1C5C5C10C

【図 1 2 4】図 1 2 4 は、システイン富裕反復タンパクのアフィニティー熟成過程スキームを示す。

【図 1 2 5】図 1 2 5 は、グラニュリン反復性タンパクの構造を示す。

【図 1 2 6】図 1 2 6 は、ランダム化のためのスキームを示す。

【図 1 2 7】図 1 2 7 は、抗凍結タンパク由来反復タンパクの構造および配列を示す。

【図 1 2 8】図 1 2 8 は、螺旋状反復性タンパクスカフォールドを示す。

【図 1 2 9】図 1 2 9 は、反復性タンパクのアフィニティー熟成過程のスキームを示す。

【図 1 3 0】図 1 3 0 は、システイン含有反復性タンパクの命名法を示す。

【図 1 3 1】図 1 3 1 は、システイン含有反復性タンパクの命名法を示す。

【図 1 3 2】図 1 3 2 は、システイン含有反復性タンパクの命名法を示す。

【図 1 3 3】図 1 3 3 は、A-ドメイン由来の反復性タンパクを示す。

【図 1 3 4】図 1 3 4 は、ポリ・トレフォイルスカフォールドを示す。

【図 1 3 5】図 1 3 5 は、マルチ・プレキシンスカフォールドを示す。

【図 1 3 6】図 1 3 6 は、ミニコラーゲンスカフォールドを示す。

【図 1 3 7】図 1 3 7 は、アフィニティー熟成過程の各種スキームを示す。

【図 1 3 8】図 1 3 8 は、アフィニティー熟成過程の各種スキームを示す。

- 【図 1 3 9】図 1 3 9 は、アフィニティー熟成過程の各種スキームを示す。
- 【図 1 4 0】図 1 4 0 は、アフィニティー熟成過程の各種スキームを示す。
- 【図 1 4 1】図 1 4 1 は、アフィニティー熟成過程の各種スキームを示す。
- 【図 1 4 2】図 1 4 2 は、アフィニティー熟成過程の各種スキームを示す。
- 【図 1 4 3】図 1 4 3 は、プラスミドサイクリングおよびメガプライマーを示す。
- 【図 1 4 4】図 1 4 4 は、疎水性プロットである。
- 【図 1 4 5】図 1 4 5 は、低分子量システイン含有ドメインを拡大するための各種方法を示す。
- 【図 1 4 6】図 1 4 6 は、抗凍結タンパクによる異なる構造体を接続する各種方法を示す。
- 【図 1 4 7】図 1 4 7 は、抗凍結タンパクによる異なる構造体を接続する各種方法を示す。
- 【図 1 4 8】図 1 4 8 は、ライブラリーを設計するための戦略を示す。
- 【図 1 4 9】図 1 4 9 は、A ドメイン構造を示す。
- 【図 1 5 0】図 1 5 0 は、ミクロタンパクの、標的誘発性フォールド形成の模式図である。
- 【図 1 5 1】図 1 5 1 は、フォリスタチンドメインの構造組織および配列を示す。
- 【図 1 5 2】図 1 5 2 は、システイン含有タンパクの構造多様性を示す。
- 【図 1 5 3】図 1 5 3 は、システイン含有タンパクの構造多様性を示す。
- 【図 1 5 4】図 1 5 4 は、ジスルフィドシャフリングによる構造進化、および、天然のシステイン含有タンパクの進化を示す。
- 【図 1 5 5】図 1 5 5 は、ジスルフィドシャフリングによる構造進化、および、天然のシステイン含有タンパクの進化を示す。
- 【図 1 5 6】図 1 5 6 は、508 個のジスルフィド含有タンパクから成るファミリーを示す。
- 【図 1 5 7】図 1 5 7 は、異なるインテグリン間の配列関係を示す。
- 【図 1 5 8】図 1 5 8 は、各種産物フォーマットの比較を示す。
- 【図 1 5 9】図 1 5 9 は、各種ミクロタンパク産物フォーマットを示す。
- 【図 1 6 0】図 1 6 0 は、アフィニティー熟成過程の各種スキームを示す。
- 【図 1 6 1】図 1 6 1 は、免疫原性を下げるための機序を示す。
- 【図 1 6 2】図 1 6 2 は、大腸菌由来の各種スカフォールドの発現を示すゲルである。図 1 6 2 のモチーフ：

【化 2 1 1】

1) CxxxxCxxxxCxxx(x)xxxxxCx(x)CxxxCxxxxxx(x)xxxCxxdxttyxxxCxxxxaxCxxxxxxxxxxgxC

2) CxxxxCxxxxCxxx(x)xxxxxCx(x)CxxxCxxxxxx(x)xxxCxxxxxxxxxxCxxxxxCxxxxxxxxxxxC

CDP: C4C5C9-10C1-2C3C9-10C10C6C13C

- 【図 1 6 3】図 1 6 3 は、HLA 結合の組み合わせ低下を示す。
- 【図 1 6 4】図 1 6 4 は、各種 TNFR ファミリー・ミクロタンパクの配列および構造を示す。
- 【図 1 6 5】図 1 6 5 は、2-3-4 積み上げ法を示す。
- 【図 1 6 6】図 1 6 6 は、ヒトおよびミクロタンパクの、予測 MHC II 結合アフィニティーを示す。グラフは、五つの主要 HLA 対立遺伝子について計算された各タンパクのスコアの分布を示す。赤色曲線：長さ中央値 372 AA を持つ、全長 26,000 のヒトタンパク。青色曲線：ジスルフィドパターンのデータベース (22) から採取した、少なくとも 10% のシステインで、偶数のシステインを有する、25-90 AA (中央値 38 AA) の 10,525 のミクロタンパク。緑色曲線：ミクロタンパク・データベースのサイズ分布に適合する 26,000 個のヒトタンパク断片。各ヒトタンパク配列について、我々のミクロタンパク・データベースからランダムに選ばれたタンパクの長さに一致する断片をランダムに生成した。コーカソイド集団において高頻度に出現する 5 種類の HLA 対

立遺伝子、HLA*101、HLA*301、HLA*401、HLA*701、HLA*1501について、MHCI結合を分析した。TEPITOPEに基づくMHCI結合基質を用いた。結合基質は、プログラムPropredからダウンロードした。TEPITOPE基質は、システイン残基に関するスコアを含まないので、代わりにアラニンスコアを用いた。各タンパクおよび各HLA対立遺伝子について、TEPITOPEの最高スコアを特定した。各対立遺伝子のデータは、全てのヒトタンパクにおける最高スコアの平均を差し引くことによって正規化した。

【図167】図167の上段パネルは、MHCI結合に対するアミノ酸のアフィニティー寄与を示す。TEPITOPE基質における全ての非疎水性残基のP1スコアは、そのP1スコアが平均スコアを支配できないように-999から-2に変えられた。アミノ酸は、各エピトープに対するその平均スコアに従ってランクづけされた。図は、5種のもっとも優勢なHLA対立遺伝子(*101、*301、*401、*701、*1501)における平均ランクを示す。下段パネルは、マイクロタンパク対ヒトタンパクにおけるアミノ酸の相対量を示す。アミノ酸量は、図166に示す配列を用いて、ヒトタンパクおよびマイクロタンパクについて計算した。データから、脂肪族疎水性残基I、V、M、Lが、免疫原性に対しもっとも強力な寄与をすること、および、平均的ヒトタンパクと比べ、マイクロタンパクではもっとも出現頻度が低いことが示される。したがって、タンパクの免疫原性の低下は、高い方から低い方に向かって下記の順序で高得点アミノ酸の含量を下げることによって実現することが可能である：IVMLFYSNRAHQ TGWKPED。

【図168】図168は、2-3-4積み上げ法のデモンストレーションとして、ファージクローンから発現されたVEGFマイクロタンパクについてELISAした結果を示す。

【図169】図169は、還元条件下におけるマイクロタンパクのSDS-PAGEゲルを示す。レーン1：ソマトメジン、レーン2：プレキシシ、レーン3：トキシシB、レーン4：トマト・プロテアーゼ阻害剤、レーン5：蜘蛛トキシシ、レーン6：アルカリフォスファターゼ・コントロール、レーン9：分子量マーカー

【図170】図170は、レドックス処理ライブラリー、および未処理ライブラリーの比較を示す。

10

20

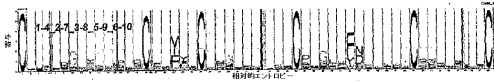
【図 20】

FIG. 20



【図 21】

FIG. 21



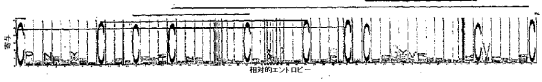
【図 22】

FIG. 22



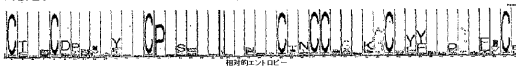
【図 23】

FIG. 23



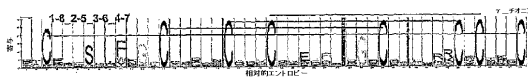
【図 24】

FIG. 24



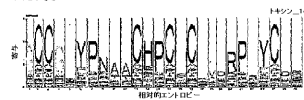
【図 25】

FIG. 25



【図 30】

FIG. 30



【図 31】

FIG. 31



【図 32】

FIG. 32



【図 33】

FIG. 33



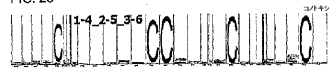
【図 34】

FIG. 34



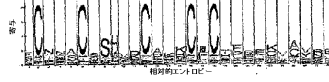
【図 26】

FIG. 26



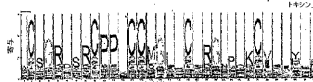
【図 27】

FIG. 27



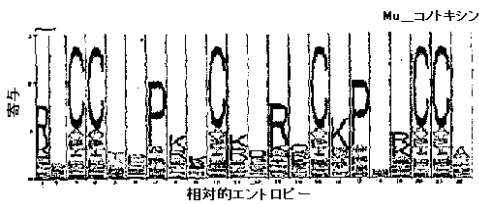
【図 28】

FIG. 28



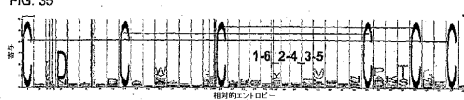
【図 29】

FIG. 29



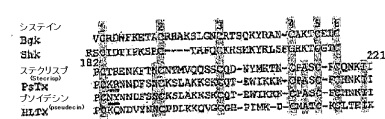
【図 35】

FIG. 35



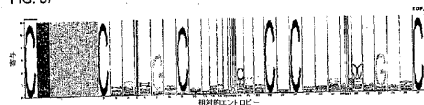
【図 36】

FIG. 36



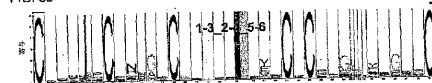
【図 37】

FIG. 37



【図 38】

FIG. 38



【図 39】

FIG. 39



【図 40】

FIG. 40



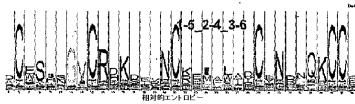
【図 41】

FIG. 41



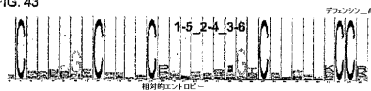
【図 42】

FIG. 42



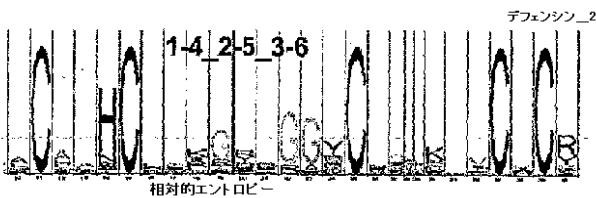
【図 43】

FIG. 43



【図 44】

FIG. 44



【図 50】

FIG. 50



【図 51】

FIG. 51



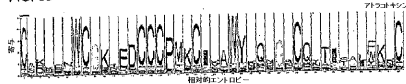
【図 52】

FIG. 52



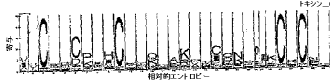
【図 53】

FIG. 53



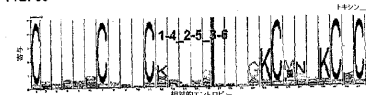
【図 54】

FIG. 54



【図 55】

FIG. 55



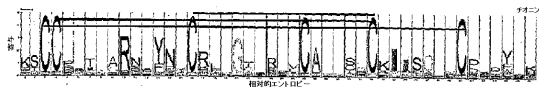
【図 45】

FIG. 45



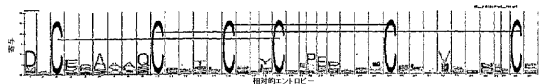
【図 46】

FIG. 46



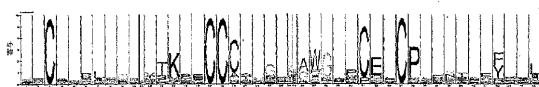
【図 47】

FIG. 47



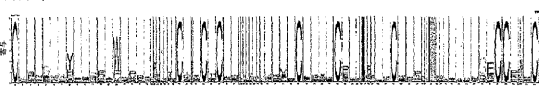
【図 48】

FIG. 48



【図 49】

FIG. 49



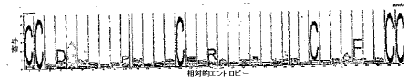
【図 56】

FIG. 56



【図 57】

FIG. 57



【図 58】

FIG. 58



【図 59】

FIG. 59



【図 60】

FIG. 60



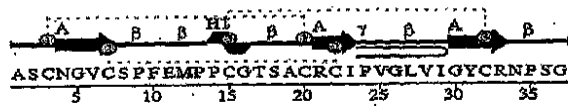
【図 61】

FIG. 61



【図 84】

FIG. 84



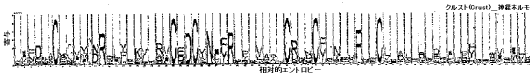
【図 85】

FIG. 85



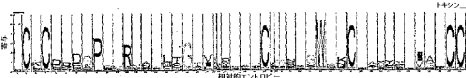
【図 86】

FIG. 86



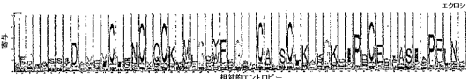
【図 87】

FIG. 87



【図 88】

FIG. 88



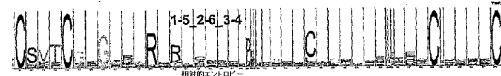
【図 95】

FIG. 95



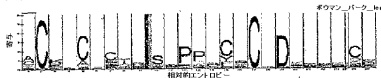
【図 96】

FIG. 96



【図 97】

FIG. 97



【図 98】

FIG. 98



【図 99】

FIG. 99



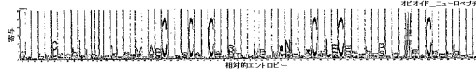
【図 100】

FIG. 100



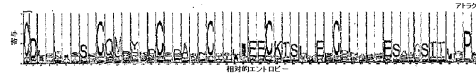
【図 89】

FIG. 89



【図 90】

FIG. 90



【図 91】

FIG. 91



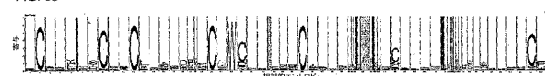
【図 92】

FIG. 92



【図 93】

FIG. 93



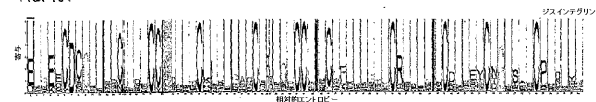
【図 94】

FIG. 94



【図 101】

FIG. 101



【図 102】

FIG. 102



【図 103】

FIG. 103

> グラニリン
 1 mwdsvval taglvgrc pdgqcpvac ddpqgasy occpildkwp ttismlggp
 61 qvdahcsag hscitvsgt escopfeav aogdghoop rghcsadgr scfqrsgn
 121 ygakqpdsg fecpdrfsc kmvtpgscoppppdpd nntdphgaf cdvlt
 181 ptgthpalk lpaqrtnav abssvmcpd anrcpdgst cclpsaky lcpthpao
 241 sghpccpd bvdllgskc lskentatl ltkphtvg dvkcdmevsc pdgytce
 301 sghpccpd bvdllgskc lskentatl ltkphtvg dvkcdmevsc pdgytce
 361 krdivpcnvs scpsctctt tseuvgcd pssvvsst qvdpqytc vaegqo
 421 eivaglekmp arrgslshpr dgcqntsc pvggtcpsa gssvact dltavvz
 481 pnygylcnw karsckevv saqpafar sphvgykive cpgshcdn ocdz
 541 wacpysag ctdmncap agfncargt kdmreapnw daplndplr qll



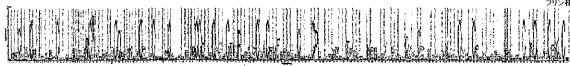
【図 104】

FIG. 104



【図 105】

FIG. 105



【図 106】

FIG. 106



【図 107】

FIG. 107



【図 108】

FIG. 108

>VEGF-C, C末端ドメイン

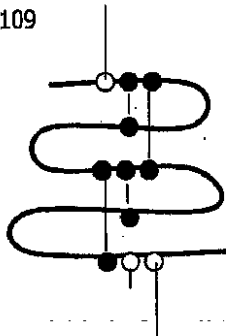
1 mltlyfiva clallallp gneapaae afegklid sepdagata yaskidneq
61 musvdelm blyparyem ykqlykgyw qinnearln srtetklid aahyrlklk
121 slinewrlkz cmrepcvdy gkefyvrlt flkpcovay rogacnag locmlstzy
181 lklidulv plagplqvlt lsfahlsr omklvlyq vhlslrp altpccaan
241 ktpnyman nltordage clmfedag cltdghdk gmkleat clldaglr
301 pascpkhl dmsagkz klfpagca nnefentkl clldcpm qplpklz
361 klldagkz klldghbp clldcpm nklacpklz ysewrcip syklpklz

>パルビニア環分泌タンパク

MKTLSSLLV LAVNVLQA SPDFONRCPA RXYWNERKQA CYKREHRCY PIGALDPES 60
CTFSMGQSG CSAKQFNKD KEGGKNDQ PKDGGAGRY WQVQSSGCTPTMSAGGS 120
GSQWCBGSG KEGGKNDQ TAPQVWIKDT GEGGKNDQ EPADGCTKPL IWDKVDGSG 180
GHLKDCGRN RDWDSGSG KEGGKNDQ SKWQVWIKDT GEGGKNDQ EPADGCTKPL IWDKVDGSG 240
DKSGGKAK MESGKGVYE SKWVWIKDT GEGGKNDQ EPADGCTKPL IWDKVDGSG 300
NGGQPKKOS GEGGKNDQ DKTCTAPQVY DGVAEGGKNDQ VIMQVADGCTKPL IWDKVDGSG 360
GEGGKNDQ GNGKVDWET IWDKVDGCTA PVCTAGKRC GEGGKNDQ REKESGAMP 420
LVWVNTGKADGADQSP GEGGKNDQ KLYGEGGKNDQ SASGGLRW NADTGGG 480
PGMPGEGGK QTWISGEGGK EPTITQQA PQDLNTEGKRWMLAQ KECKSPRQW 540
TDSKVGSG TPTATGSGQ TWGEGGKNDQ GEGGKNDQ KFFKPSGEGGKNDQ 600
SPQVWADGCTKPL IWDKVDGCTA PVCTAGKRC GEGGKNDQ REKESGAMP 660
GEGGKNDQ KQVWIKDT GEGGKNDQ PVGSGAKTW LDOEGGKNDQ PGKSGGCTG 720
AQKWDKCTKPL IWDKVDGCTA PVCTAGKRC GEGGKNDQ REKESGAMP 780
PPKNSQVLT KOTEGGKNDQ YSKCAQKPL IWDKVDGCTA PVCTAGKRC 840
PNSASMTKL SPKRWIKVTEGKNDQ CPPTGQVMD DKGGKNDQ GEGGKNDQ 900
NDPTGEGGKNDQ GEGGKNDQ KWSAETGEGGKNDQ GEGGKNDQ LONFNDGCTA PVCTAGKRC 960
NOKSPRTWY DTEGKNDQ GEGGKNDQ DSDGQVQ IWDGEGGKNDQ GEGGKNDQ ANKRFEGGKNDQ 1020
GEGGKNDQ PPGGKNDQ KQVWIKDT GEGGKNDQ PVGSGAKTW LDOEGGKNDQ PGKSGGCTG 1080
FIESKVGSG ETQTGQKGF RWSNLEGGKNDQ GEGGKNDQ VDKNTGEGGKNDQ GEGGKNDQ 1140
GEGGKNDQ PPLD GEGGKNDQ ANGTGQVQW NEKGGKNDQ KQVWIKDT GEGGKNDQ 1200
GEGGKNDQ PPTG GEGGKNDQ KQVWIKDT GEGGKNDQ PVGSGAKTW LDOEGGKNDQ PGKSGGCTG 1260
KTKWVNDGCTKPL IWDKVDGCTA PVCTAGKRC GEGGKNDQ REKESGAMP 1320
GEGGKNDQ GEGGKNDQ KOTEGGKNDQ GEGGKNDQ GEGGKNDQ GEGGKNDQ GEGGKNDQ 1380
CSAKQVCP AQKDCSPKT WDENSQSG PKNRPPTGG CHAGRTWDA TCEKQVATP 1440
KDSKVPFPT TGGGKNDQ PKLADAGRTW DEQKQVCTD LPAQCHYE GQVYDNTCK 1500
GEGGKNDQ QGFSFSGSK OKLLEGNK DPGGKNDQ QGEGGKNDQ SEPRMGCGF 1560
NRYWKEGCA CQLKTKVCT DSDPSFSD TDCESNAK DFLNQLKL NKVSGQPCI 1620
QRPPEGCAE KTEGFSWCP KADECKGEG KTYKGRVD KATONKCTN EPKTYKCAN 1680
TAESVDEGQK AKWSTORH 1700

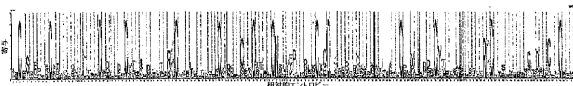
【図 109】

FIG. 109



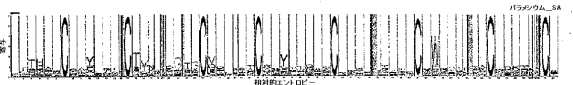
【図 110】

FIG. 110



【図 111】

FIG. 111



【図 112】

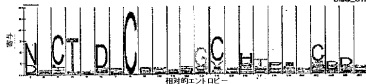
FIG. 112

1 mnnklfisi llaivasqy sltscaql lsegdclkn slgklitk ktogvstpv
61 tptlyayc dfaetdcpk alpctdogny aagavveskc tftgctpfa kltdsecqai
121 snroldgh cvevdasty kklpcaakna agstgclt nntcvdanc klipattad
181 klcdvistic tktgggcvd sgncsdqtl elqklk tsgldgaa clklcdnap
241 tslltdack tftdgtct kangcvrt tcaatqas clnssggd klgtacvdk
301 acantptia trisacagvt gclksgggc vngacsvan vqacvkms nfdagdtc

361 kektcanst tmthldcls ylstcvksq ggcqnrta aptmtnda ceaytgnnc
421 itksgggcv nttcaalte aacvkssgs clldtass clklctvnap atnthldcl
481 pflntcvns tsagvektc ensklald kdtssracw kgkykkqcv lassatttha
541

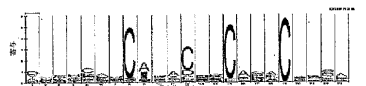
【図 113】

FIG. 113



【図 114】

FIG. 114



【図 1 1 5】

FIG. 115



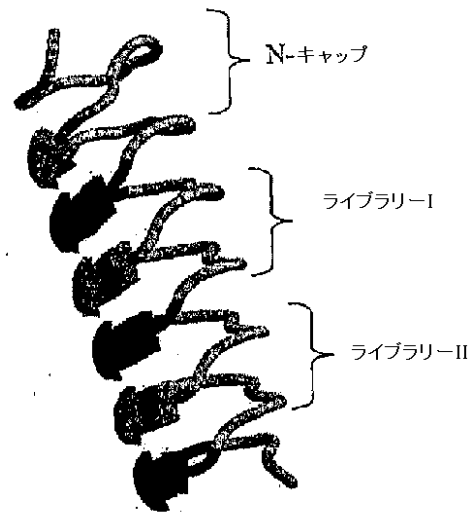
【図 1 1 6】

FIG. 116

MGRSLSRLIL FLFCLLALS TAASRLKRQ CGCSNFCNCQ QQPVFFAQIS 50
 LPACSCQAP ~~LCRCQFRAEINSDPSATCQ~~ RACIPSCSKS TGNTFACSTNIP
~~SESTGRTTA~~ SAAQQAAMSHI QVSPNPQPL VPAAAPTVD ~~DSGQNVQNVV~~
~~VOGFAVSQNS~~ PPAVE ~~QQLKRIUSQCHGQ~~ ATN EQLPTTSSTS TNAPTIKTL 200
 NINDAYFDS ~~QVQKPIGSC~~ H ~~SC~~ SQGNPA ASCSNSCNTE CSDKCSTRPV 250
 QAQVQIQPQ QVQITIPQC QSRCENRCLS TCTASQPSVC APSCSTACQL 300
 SCDSIGAATA PQSTNPTPVE VRALCTPTCM PQCLPSCAT TTTTITQTPY 350

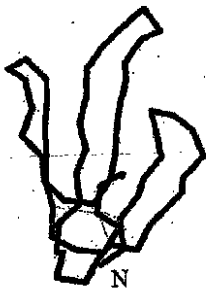
【図 1 1 7】

FIG. 117



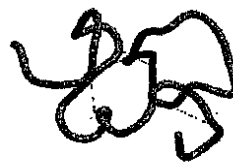
【図 1 1 8】

FIG. 118



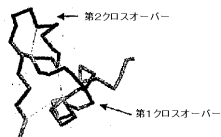
【図 1 2 1】

FIG. 121



【図 1 1 9】

FIG. 119



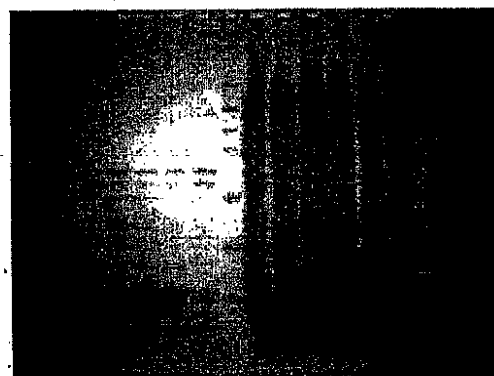
【図 1 2 0】

FIG. 120

ライブラリー-VESCEQYTSQGECLGSGDPH-CGWCVLENRCRRSDCQRAEEPNRWASSISQCVKL
 1esl GPGRHFLTQSMCLRAPRFMGCGWC-GGVCSRQHEC-DGGWWQDS-CPPP
 10 20 30 40 50

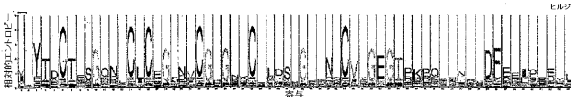
【図 1 2 2】

FIG. 122



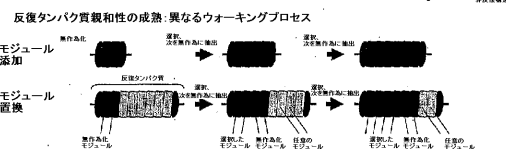
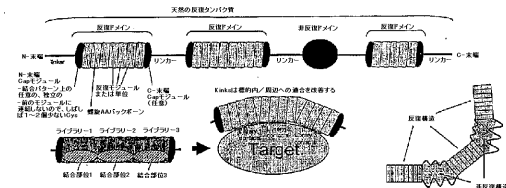
【 図 1 2 3 】

FIG. 123



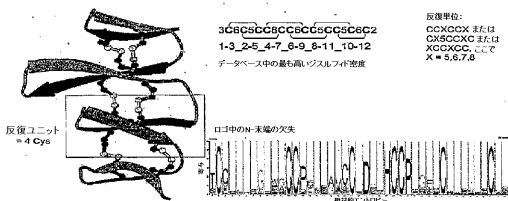
【図 1 2 4】

FIG. 124 システインリッチの反復タンパク質



【図 1 2 5】

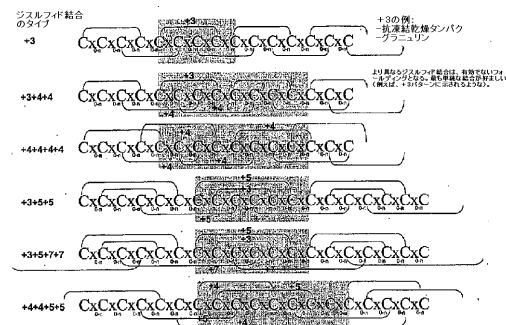
FIG. 125 グラニュリン反復性タンパク



Nat Struct Biol. 1996 Sep;3(9):747-52.
The hairpin stack fold, a novel protein architecture for a new family of protein growth factors
Jabal R, Chen Z, Jones K, Bonfield W, NUB

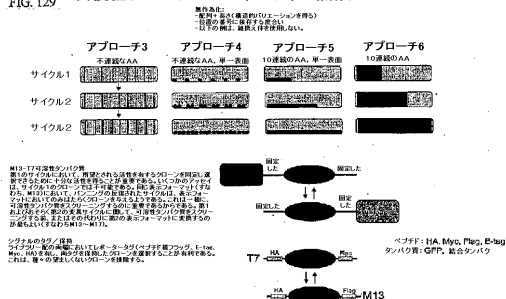
【图 1 2 8】

| FIG. 128 螺旋状反復性タンパクスカフォールドのための単純設計



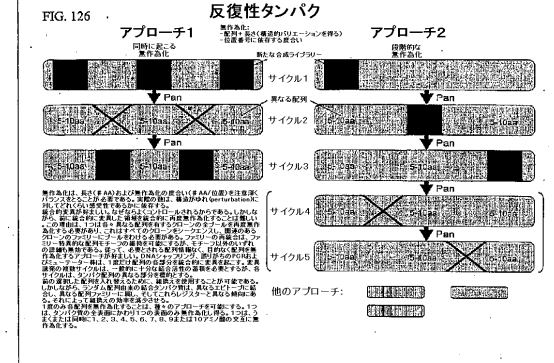
【图 129】

FIG. 129 反復性タンパクのアフィニティー熟成



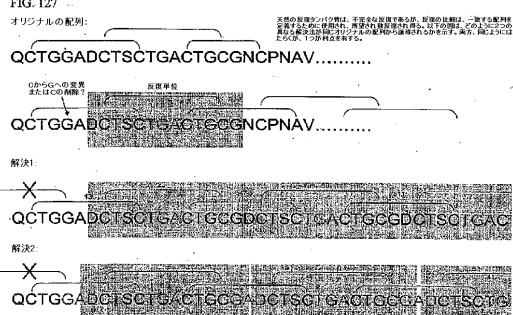
【図 1 2 6】

FIG. 126



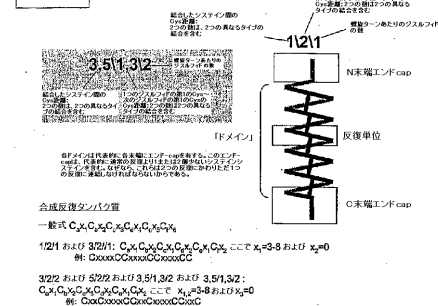
【例 127】

FIG. 127 抗凍結タンパク由来反復性タンパク



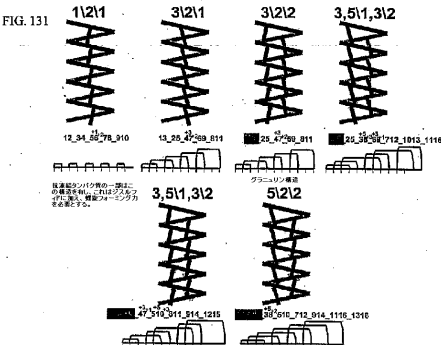
【☒ 1 3 0】

FIG. 130 反復タンパク質の命名法



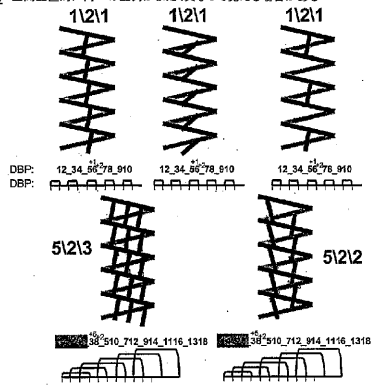
【 図 1 3 1 】

FIG. 131



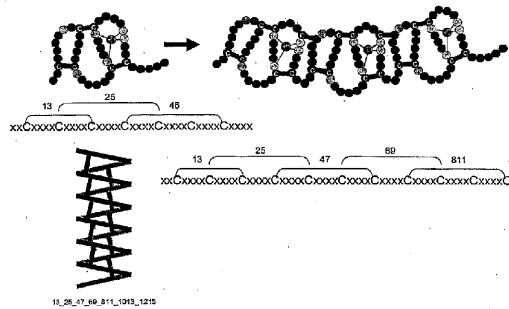
【図 132】

FIG. 132 空間配位的に同一の配列がまったく異なって見える場合がある



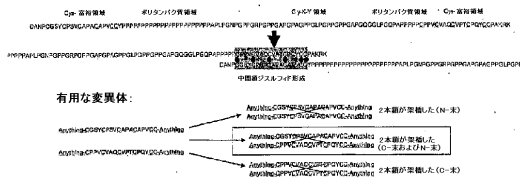
【図 133】

FIG. 133 A-ドメイン由来の反復性タンパク



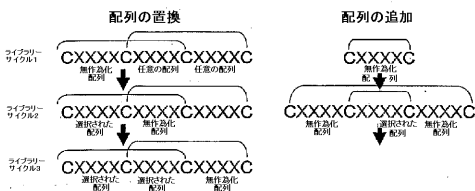
【図 136】

FIG. 136 ミニコラーゲン 2本鎖ドメイン



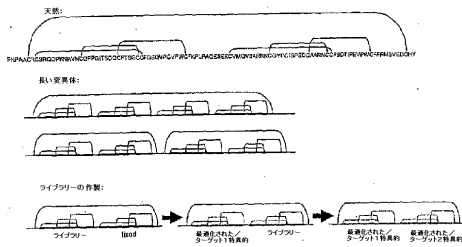
【図 137】

FIG. 137 アフィニティ熟成:ランダム化戦略



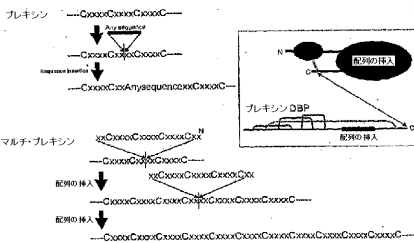
【図 134】

FIG. 134 ポリ・トレフォイル



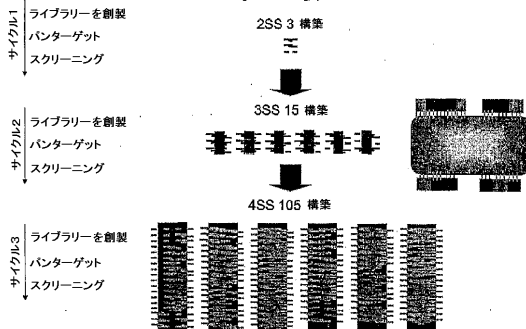
【図 135】

FIG. 135 マルチ・プレキシシ



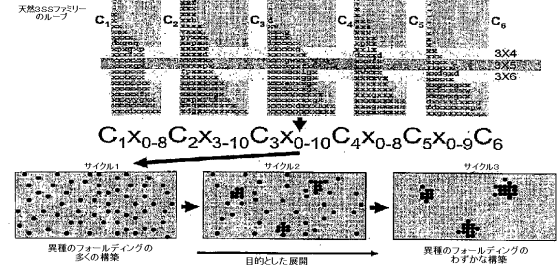
【図 138】

FIG. 138 アフィニティ熟成 - '234' スカフォールド独立



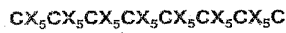
【図 139】

FIG. 139 設計プロセス - '3x0-8', '4x0-8' スカフォールド独立



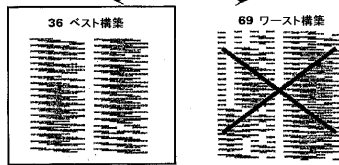
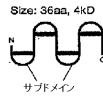
【 図 1 4 0 】

FIG. 140 設計プロセス - '4x5'
スカフォールド独立



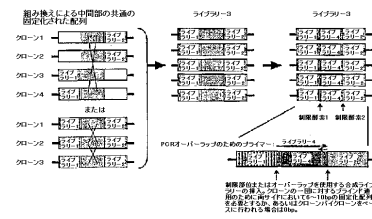
平均ジスルフィドスパンは14AAである。
決して<9 AAではない。

X₅ スペーシングはサブドメインを排除し、そして多くの異なる構築を可能にする。



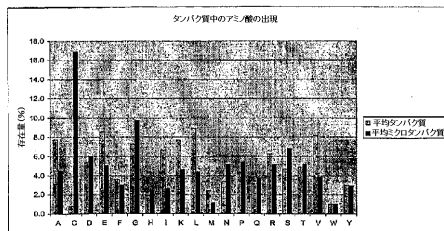
【 図 1 4 1 】

FIG. 141 アフィニティー熟成



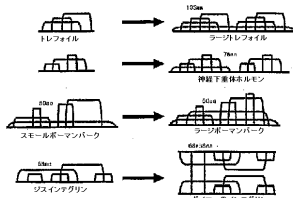
【图 1 4 4】

FIG. 144 疎水性



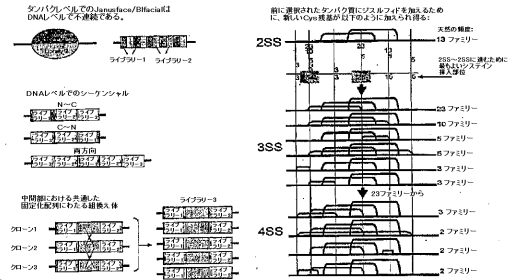
【図 1 4 5】

FIG. 145. 小型ドメインを拡大する



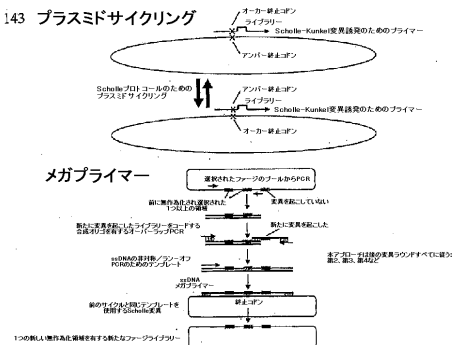
【図 1 4 2】

FIG. 142 アフィニティー熟成



【図 1 4 3】

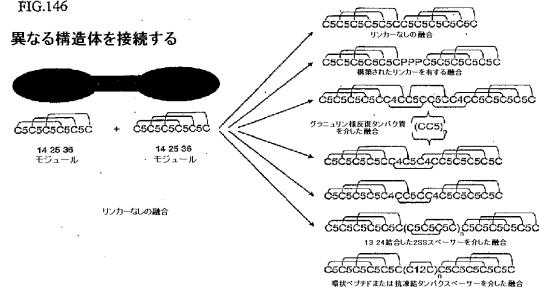
FIG. 143 プラスミドサイクリング



【图 1 4 6】

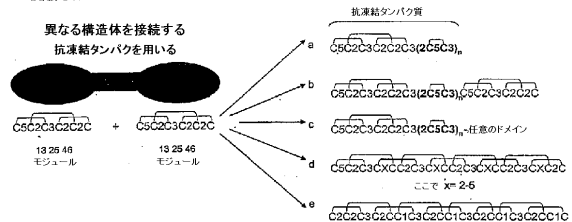
FIG.146

異なる構造体を接続する



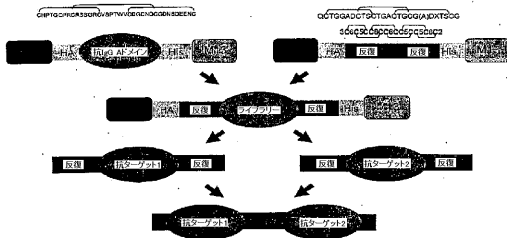
【图 1 4 7】

FIG. 147



【図 1 4 8】

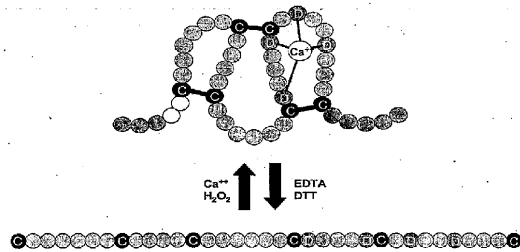
FIG. 148 リード設計戦略



【 図 1 4 9 】

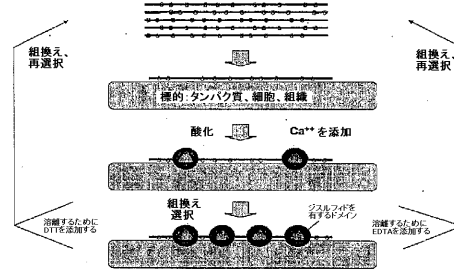
FIG. 149

Aードメイン構造



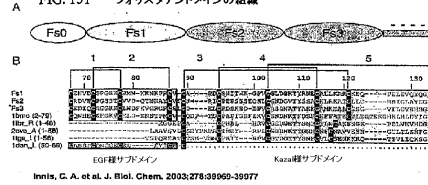
【図 1 5 0】

FIG. 150 ミクロタンパクの標的誘発性フォールド形成



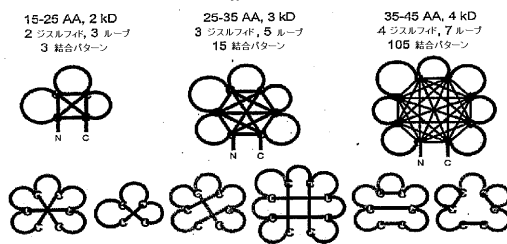
【図 1 5 1】

FIG. 151 フォリスタチンドメインの組織



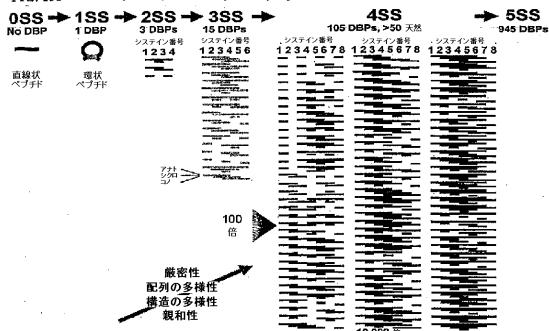
【図 1 5 2】

FIG. 152 ミクロタンパク構造



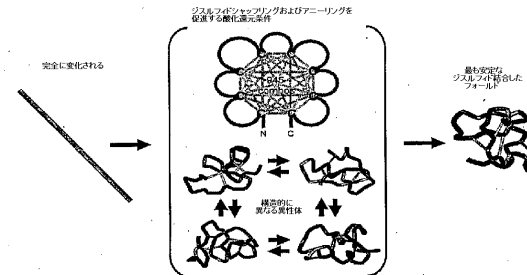
【図 1 5 3】

FIG. 153 ミクロタンパクのクラス



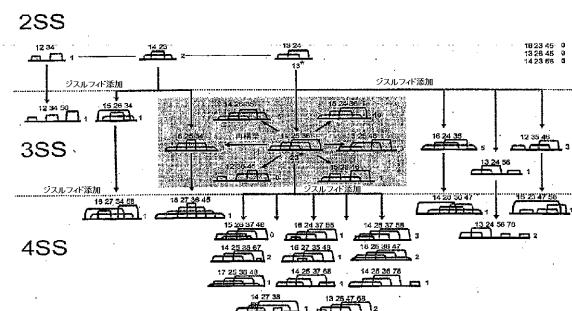
【図 1 5 4】

FIG. 154 ジスルフィドシャッフリングによる構造進化

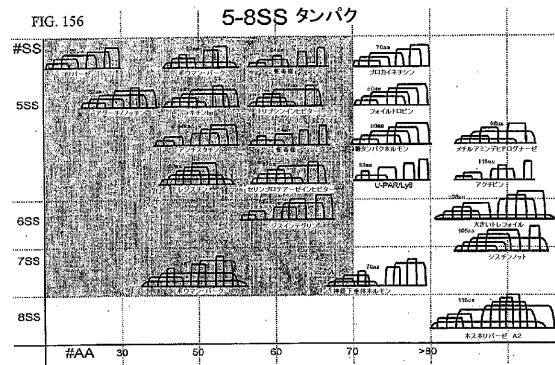


【図 1 5 5】

FIG. 155 天然ミクロタンパクは、構造、長さ、および配列を適応させる

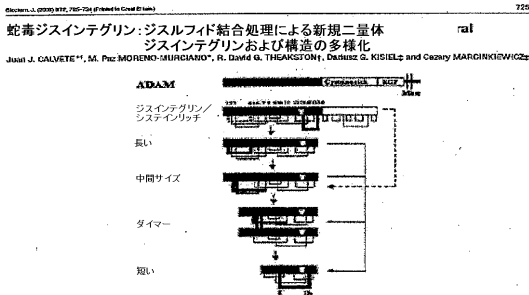


【図 156】

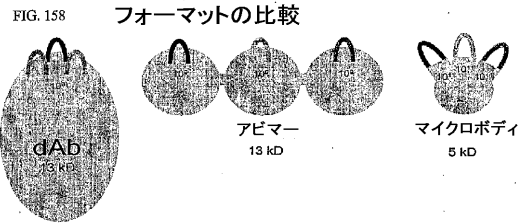


【図 157】

FIG. 157

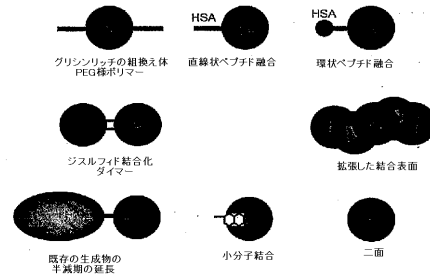


【図 158】



【図 159】

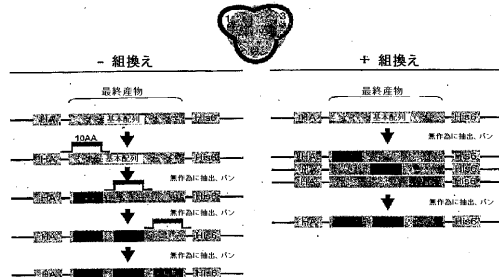
FIG. 159 ミクロタンパク製品フォーマット



【図 160】

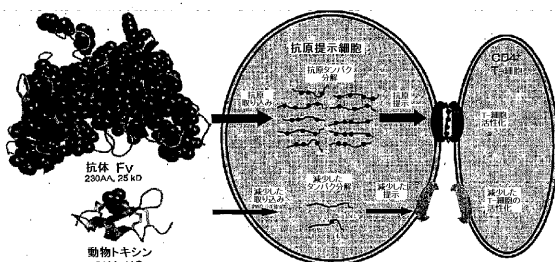
FIG. 160

アフィニティー熟成



【図 161】

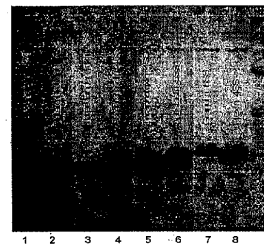
FIG. 161



【図 162】

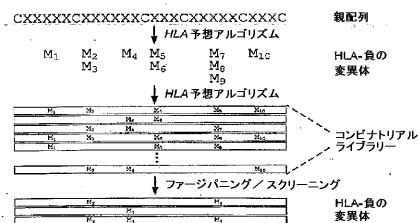
FIG. 162

大腸菌におけるスカフォールド発現レベルの例



【図 163】

FIG. 163 HLA結合の組み合わせによる低下



【手続補正書】

【提出日】平成20年12月8日(2008.12.8)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0008

【補正方法】変更

【補正の内容】

【0008】

別の実施態様では、本発明は、標的分子に対し結合特異性を示す、非天然のシステイン(C)－含有スカフォールドを提供する。この実施態様において、非天然システイン(C)－含有スカフォールドは、式：

【数1】

$$\prod_{i=1}^n 2i-1$$

によって表される組み合わせの群から選ばれるパターンによるスカフォールド内システインを含む。式中、nは、システイン残基によって形成されるジスルフィド結合の予測数に等しく、かつ、

【数2】

$$\prod$$

は、(2i-1)の積を表し、この式においてiは、1からnまでの範囲の正の整数である。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0049

【補正方法】変更

【補正の内容】

【0049】

ジスルフィド固有の特徴は、同じ一組のシステインが、原理的には、様々の別形のジスルフィド結合パターンとして連結することが可能であることである。なぜなら、ジスルフィドは組み合わせだからである。例えば、二つのジスルフィドタンパクは、3通りの異なるジスルフィド結合パターン(DBP)を有し、三つのジスルフィドタンパクは、15通りの異なるDBPを有し、かつ、四つのジスルフィドタンパクは、最大105通りの異なるDBPを有することが可能である。2SS DBPの全て、3SS DBPの大部分、および4SS DBPの半分未満において天然例が存在する。一局面において、ジスルフィド結合パターンの全数は、式：

【数3】

$$\prod_{i=1}^n 2i-1$$

に従って計算することが可能である。式中、n＝システイン残基によって形成されるジスルフィド結合の予測数であり、かつ、式中、

【数4】

$$\prod$$

は、(2i-1)の積を表す。この式においてiは、1からnまでの範囲の正の整数である。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】 0 0 5 0

【補正方法】 変更

【補正の内容】

【 0 0 5 0 】

したがって、一実施態様では、標的分子に対して結合特異性を示す、非天然システイン（C）－含有スカフォールドを提供する。この実施態様において、非天然システイン（C）－含有スカフォールドは、式：

【数 5】

$$\prod_{i=1}^n 2i-1$$

によって表される組み合わせの群から選ばれるパターンによるスカフォールド内システインを含む。式中、nは、システイン残基によって形成されるジスルフィド結合の予測数に等しく、かつ、式中、

【数 6】

$\prod$

は、 $(2i-1)$ の積を表し、この式においてiは、1からnまでの範囲の正の整数である。一局面において、非天然システイン（C）－含有タンパクは、 $C^{1-2, 3-4}$ 、 $C^{1-3, 2-4}$ 、および $C^{1-4, 2-3}$ から成る群から選ばれるパターンに従って、ポリペプチドに含まれるシステインを対合させることによって形成される二つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。別の局面では、非天然システイン（C）－含有タンパクは、 $C^{1-2, 3-4, 5-6}$ 、 $C^{1-2, 3-5, 4-6}$ 、 $C^{1-2, 3-6, 4-5}$ 、 $C^{1-3, 2-4, 5-6}$ 、 $C^{1-3, 2-5, 4-6}$ 、 $C^{1-3, 2-6, 4-5}$ 、 $C^{1-4, 2-3, 5-6}$ 、 $C^{1-4, 2-6, 3-5}$ 、 $C^{1-5, 2-3, 4-6}$ 、 $C^{1-5, 2-4, 3-6}$ 、 $C^{1-5, 2-6, 3-4}$ 、 $C^{1-6, 2-3, 4-5}$ 、および $C^{1-6, 2-5, 3-4}$ から成る群から選ばれるパターンに従って、ポリペプチドに含まれるシステインを対合させることによって形成される三つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。別の局面において、非天然システイン（C）－含有タンパクは、前述の式によって定義される組み合わせの群から選ばれるパターンに従って、ポリペプチドに含まれるシステインを対合させることによって形成される、少なくとも四つのジスルフィド結合を有するポリペプチドを含む、標的分子に対して結合特異性を示すポリペプチド、非天然システイン（C）－含有タンパクを含む。さらに別の局面では、非天然システイン（C）－含有タンパクは、 C^{1-9} 、 C^{1-10} 、 C^{2-9} 、 C^{2-10} 、 C^{3-9} 、 C^{3-10} 、 C^{4-9} 、 C^{4-10} 、 C^{5-9} 、 C^{5-10} 、 C^{6-9} 、 C^{6-10} 、 C^{7-9} 、 C^{7-10} 、 C^{8-9} 、 C^{8-10} 、および C^{9-10} から成る群から選ばれるパターンに従って、タンパク内のシステインを対合させることによって形成される、少なくとも五つのジスルフィド結合を有するポリペプチドを含む。前式において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。さらに別の局面では、標的分子に対し結合特異性を示す、非天然システイン（C）－含有タンパクは、 C^{1-11} 、 C^{1-12} 、 C^{2-11} 、 C^{2-12} 、 C^{3-11} 、 C^{3-12} 、 C^{4-11} 、 C^{4-12} 、 C^{5-11} 、 C^{5-12} 、 C^{6-11} 、 C^{6-12} 、 C^{7-11} 、 C^{7-12} 、 C^{8-11} 、 C^{8-12} 、および C^{9-11} 、 C^{9-12} 、 C^{10-11} 、 C^{10-12} 、および C^{11-12} から成る群から選ばれるパターンに従って、タンパク内のシステインを対合させることによって形成される、少なくとも六つのジスルフィド結合を有するポリペプチドを含む。前式

において、ハイフンで連結される二つの数字は、ポリペプチドのN末端から数えて、どの二つのシステインがペア形成をしてジスルフィド結合を形成するかを示す。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0 1 1 1

【補正方法】変更

【補正の内容】

【0 1 1 1】

【数7】

$$\prod_{i=1}^n 2i-1$$

式中、 n = システイン残基によって形成されるジスルフィド結合の予測数であり、かつ、

【数8】

$\prod$

は、 $(2i-1)$ の積を表し、この式において i は、1 から n までの正の整数である。

フロントページの続き

(51)Int.Cl. F I テーマコード (参考)
A 6 1 P 35/00 (2006.01) A 6 1 P 35/00
G 0 1 N 33/53 (2006.01) G 0 1 N 33/53 D

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(72)発明者 シェレンバーガー, フォルカー
 アメリカ合衆国 カリフォルニア 9 4 3 0 3, パロ アルト, モレノ アベニュー 9 1 4

(72)発明者 バーダー, マーティン
 アメリカ合衆国 カリフォルニア 9 4 0 4 1, マウンテン ビュー, レインボー ドライブ
 6 0 0, アpartment 2 1 4

(72)発明者 ショール, マイケル
 アメリカ合衆国 カリフォルニア 9 4 0 4 3, マウンテン ビュー, ノース ホイットマン
 ロード 1 0 0, ナンバー 3 6 1 3

Fターム(参考) 4B024 AA01 AA11 BA80 CA02 DA06 EA03
 4B064 AG01 CA02 CA19 CC24 DA01 DA13
 4C084 AA02 AA07 BA01 BA09 BA19 BA20 BA23 CA59 DC50 NA14
 ZB261 ZB262
 4H045 AA10 AA20 AA30 BA01 BA18 BA20 EA28 EA50 FA74