(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)公表番号
特表2023-527764
(P2023-527764A)
(43)公表日　令和5年6月30日(2023.6.30)

(51)国際特許分類		FI			テーマコード（参考）
C07K	*14/195 (2006.01)*	C07K	14/195		2G045
C07K	*1/13 (2006.01)*	C07K	1/13		4B029
C07K	*19/00 (2006.01)*	C07K	19/00		4B050
C07K	*17/14 (2006.01)*	C07K	17/14		4B063
C12M	*1/34 (2006.01)*	C12M	1/34	Z	4H045

審査請求　未請求　予備審査請求　未請求　（全184頁）　最終頁に続く

(21)出願番号　特願2022-570635(P2022-570635)
(86)(22)出願日　令和3年5月20日(2021.5.20)
(85)翻訳文提出日　令和5年1月13日(2023.1.13)
(86)国際出願番号　PCT/US2021/033493
(87)国際公開番号　WO2021/236983
(87)国際公開日　令和3年11月25日(2021.11.25)
(31)優先権主張番号　63/059,919
(32)優先日　令和2年7月31日(2020.7.31)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　63/027,913
(32)優先日　令和2年5月20日(2020.5.20)
(33)優先権主張国・地域又は機関
米国(US)
(81)指定国・地域　AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA
最終頁に続く

(71)出願人　516144164
クアンタム－エスアイ インコーポレイテッド
ＱＵＡＮＴＵＭ－ＳＩ ＩＮＣＯＲＰＯＲＡＴＥＤ
アメリカ合衆国 ０６４３７ コネチカット州 ギルフォード オールド ウィット フィールド ストリート ５３０
(74)代理人　100105957
弁理士　恩田 誠
(74)代理人　100068755
弁理士　恩田 博宣
(74)代理人　100142907
弁理士　本田 淳
(74)代理人　100152489
最終頁に続く

(54)【発明の名称】　タンパク質シーケンスのための方法及び組成物

(57)【要約】
本願の態様は、タンパク質、ポリペプチド、及びアミノ酸を同定及びシーケンスする方法、並びにそれらに有用な組成物を提供する。いくつかの態様では、本願は、アミノ酸結合タンパク質及びそれらの融合ポリペプチドなどのアミノ酸認識分子を提供する。いくつかの態様では、本願は、ポリペプチドシーケンシング反応時の光安定性を増強するシールディングエレメントを含むアミノ酸認識分子を提供する。

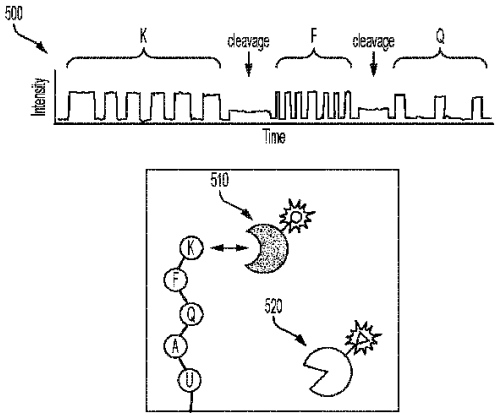


FIG. 5

【特許請求の範囲】
【請求項１】
表１又は表２から選択された配列に少なくとも８０％同一なアミノ酸配列を有するとともに１つ以上の標識を含む、組換えアミノ酸結合タンパク質。
【請求項２】
前記１つ以上の標識は、発光標識又は伝導率標識を含む、請求項１に記載の組換えアミノ酸結合タンパク質。
【請求項３】
前記発光標識は、少なくとも１つのフルオロフォア色素分子を含む、請求項２に記載の組換えアミノ酸結合タンパク質。
【請求項４】
前記発光標識は、２０以下のフルオロフォア色素分子を含む、請求項２又は３に記載の組換えアミノ酸結合タンパク質。
【請求項５】
前記発光標識は、ドナー標識とアクセプター標識とを含む少なくとも１つのＦＲＥＴペアを含む、請求項２～４のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項６】
前記ドナー標識と前記アクセプター標識との比は、１：１、２：１、３：１、４：１、又は５：１である、請求項５に記載の組換えアミノ酸結合タンパク質。
【請求項７】
前記アクセプター標識と前記ドナー標識との比は、１：１、２：１、３：１、４：１、又は５：１である、請求項５に記載の組換えアミノ酸結合タンパク質。
【請求項８】
前記伝導率標識は、荷電ポリマーを含む、請求項１～７のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項９】
前記１つ以上の標識は、タグ配列を含む、請求項１～８のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項１０】
前記タグ配列は、精製タグ、切断部位、及びビオチン化配列のうちの１つ以上を含む、請求項９に記載の組換えアミノ酸結合タンパク質。
【請求項１１】
前記ビオチン化配列は、少なくとも１つのビオチンリガーゼ認識配列を含む、請求項１０に記載の組換えアミノ酸結合タンパク質。
【請求項１２】
前記ビオチン化配列は、直列に配向された２つのビオチンリガーゼ認識配列を含む、請求項１０又は１１に記載の組換えアミノ酸結合タンパク質。
【請求項１３】
前記１つ以上の標識は、ビオチン部分を含む、請求項１～１２のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項１４】
前記ビオチン部分は、少なくとも１つのビオチン分子を含む、請求項１３に記載の組換えアミノ酸結合タンパク質。
【請求項１５】
前記ビオチン部分は、ビスビオチン部分である、請求項１３又は１４に記載の組換えアミノ酸結合タンパク質。
【請求項１６】
前記標識は、少なくとも１つのビオチン分子が装着された少なくとも１つのビオチンリガーゼ認識配列を含む、請求項１４又は１５に記載の組換えアミノ酸結合タンパク質。
【請求項１７】

前記１つ以上の標識は、１つ以上のポリオール部分を含む、請求項１～１６のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項１８】
前記１つ以上のポリオール部分は、デキストラン、ポリビニルピロリドン、ポリエチレングリコール、ポリプロピレングリコール、ポリオキシエチレングリコール、ポリビニルアルコール、又はこれらの組み合わせ若しくはバリエーションを含む、請求項１７に記載の組換えアミノ酸結合タンパク質。
【請求項１９】
前記組換えアミノ酸結合タンパク質は、前記１つ以上の標識が装着された１つ以上の非天然アミノ酸を含む、請求項１～１８のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項２０】
前記アミノ酸配列は、表１から選択された配列に８０～９０％、９０～９５％、又は少なくとも９５％同一である、請求項１～１９のいずれか一項に記載の組換えアミノ酸結合タンパク質。
【請求項２１】
請求項１～２０のいずれか一項に記載の組換えアミノ酸結合タンパク質を含む組成物。
【請求項２２】
２つ以上のアミノ酸認識分子を含むポリペプチドシーケンシング反応組成物であって、前記２つ以上のアミノ酸認識分子のうちの少なくとも１つは、請求項１～２０のいずれか一項に記載の組換えアミノ酸結合タンパク質である、ポリペプチドシーケンシング反応組成物。
【請求項２３】
前記組成物は、少なくとも１タイプの切断試薬を含む、請求項２２に記載のポリペプチドシーケンシング反応組成物。
【請求項２４】
ポリペプチドシーケンシング方法であって、
ポリペプチドと請求項２２又は２３に記載のポリペプチドシーケンシング反応組成物とを接触させることと、
前記ポリペプチドが分解されつつある間に前記ポリペプチドと少なくとも１つのアミノ酸認識分子との一連の相互作用を検出することによって前記ポリペプチドをシーケンシングすることと、を含む、方法。
【請求項２５】
アミノ酸結合タンパク質及びペプチダーゼを含むポリペプチドシーケンシング反応混合物であって、前記アミノ酸結合タンパク質は１つ以上の標識を含み、前記標識付アミノ酸結合タンパク質と前記ペプチダーゼとのモル比は、約１：１，０００～約１：１、又は約１：１～約１００：１である、ポリペプチドシーケンシング反応混合物。
【請求項２６】
前記モル比は、約１：１００～約１：１、又は約１：１～約１０：１である、請求項２５に記載のポリペプチドシーケンシング反応混合物。
【請求項２７】
前記モル比は、約１：１，０００、約１：５００、約１：２００、約１：１００、約１：１０、約１：５、約１：２、約１：１、約５：１、約１０：１、約５０：１、約１００：１である、請求項２５又は２６に記載のポリペプチドシーケンシング反応混合物。
【請求項２８】
前記アミノ酸結合タンパク質は、合成タンパク質又は組換えタンパク質である、請求項２５～２７のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項２９】
前記アミノ酸結合タンパク質は、分解経路タンパク質、不活性化ペプチダーゼ、抗体、アミノトランスフェラーゼ、ｔＲＮＡシンテターゼ、又はＳＨ２ドメイン含有タンパク質

若しくはその断片である、請求項２５～２８のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３０】
前記アミノ酸結合タンパク質は、ＣｌｐＳタンパク質、Ｇｉｄタンパク質、ＵＢＲボックスタンパク質若しくはそのＵＢＲボックスドメイン含有断片、又はｐ６２タンパク質若しくはそのＺＺドメイン含有断片である、請求項２５～２９のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３１】
前記アミノ酸結合タンパク質は、ＣｌｐＳタンパク質である、請求項２５～３０のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３２】
前記ＣｌｐＳタンパク質は、Ａ．ツミファシエンス（Ａ．ｔｕｍｉｆａｃｉｅｎｓ）、Ｃ．クレセンタス（Ｃ．ｃｒｅｓｃｅｎｔｕｓ）、Ｅ．コリ（Ｅ．ｃｏｌｉ）、Ｓ．エロンガタス（Ｓ．ｅｌｏｎｇａｔｕｓ）、Ｐ．ファルシパラム（Ｐ．ｆａｌｃｉｐａｒｕｍ）、Ｔ．エロンガタス（Ｔ．ｅｌｏｎｇａｔｕｓ）、又はこれらの相同種由来のＣｌｐＳ１又はＣｌｐＳ２である、請求項３１に記載のポリペプチドシーケンシング反応混合物。
【請求項３３】
前記アミノ酸結合タンパク質は、表１又は表２から選択された配列に少なくとも８０％、８０～９０％、９０～９５％、又は少なくとも９５％同一のアミノ酸配列を有するタンパク質である、請求項２５～３２のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３４】
前記アミノ酸結合タンパク質は、請求項Ａ１～Ａ１９のいずれか一項に記載の組換えアミノ酸結合タンパク質である、請求項２５～３３のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３５】
前記ペプチダーゼは、エキソペプチダーゼである、請求項２５～３４のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３６】
前記ペプチダーゼは、アミノペプチダーゼ又はカルボキシペプチダーゼである、請求項２５～３５のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３７】
前記ペプチダーゼは、１タイプより多くのアミノ酸をペプチダーゼの末端から切断する非特異的エキソペプチダーゼである、請求項２５～３６のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３８】
前記ペプチダーゼは、プロリンアミノペプチダーゼ、プロリンイミノペプチダーゼ、グルタミン酸／アスパラギン酸特異的アミノペプチダーゼ、メチオニン特異的アミノペプチダーゼ、又は亜鉛メタロプロテアーゼである、請求項２５～３７のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項３９】
前記ペプチダーゼは、表４又は表５から選択された配列に少なくとも８０％、８０～９０％、９０～９５％、又は少なくとも９５％同一のアミノ酸配列を有する酵素である、請求項２５～３８のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項４０】
前記反応混合物は、１つより多くのアミノ酸結合タンパク質を含む、請求項２５～３９のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項４１】
前記反応混合物は、１つより多くのペプチダーゼを含む、請求項２５～４０のいずれか一項に記載のポリペプチドシーケンシング反応混合物。

【請求項４２】
前記反応混合物は、表面に固定されたペプチド分子を含む、請求項２５～４１のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項４３】
単一ペプチド分子、少なくとも１つのペプチダーゼ分子、及び少なくとも３つのアミノ酸認識分子を含む、ポリペプチドシーケンシング反応混合物。
【請求項４４】
少なくとも１つ且つ１０までのペプチダーゼ分子を含む、請求項４３に記載のポリペプチドシーケンシング反応混合物。
【請求項４５】
少なくとも１つ且つ５つまでのペプチダーゼ分子を含む、請求項４３又は４４に記載のポリペプチドシーケンシング反応混合物。
【請求項４６】
少なくとも１つ且つ３つまでのペプチダーゼ分子を含む、請求項４３～４５のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項４７】
少なくとも３つ且つ３０までのアミノ酸認識分子を含む、請求項４３～４６のいずれか一項に記載のポリペプチドシーケンシング反応混合物。
【請求項４８】
２０まで、１０まで、又は５つまでのアミノ酸認識分子を含む、請求項４７に記載のポリペプチドシーケンシング反応混合物。
【請求項４９】
サンプルウェルのアレイを含む基体であって、前記アレイの少なくとも１つのサンプルウェルは、請求項４３～４８のいずれか一項に記載のポリペプチドシーケンシング反応混合物を含む、基体。
【請求項５０】
前記少なくとも１つのサンプルウェルはボトム表面を含み、前記単一ポリペプチド分子は前記ボトム表面に固定される、請求項４９に記載の基体。
【請求項５１】
サンプルウェルのアレイを含む基体であって、前記アレイの少なくとも１つのサンプルウェルは、単一ポリペプチド分子、切断手段、及び結合手段を含み、前記結合手段及び前記切断手段は、前記結合手段と前記ポリペプチド上の末端アミノ酸との間の少なくとも１０の会合イベントを、前記ポリペプチドから前記末端アミノ酸を前記切断手段によって除去する前に達成するように構成される、基体。
【請求項５２】
末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポリペプチドを含むアミノ酸認識分子であって、前記第１のアミノ酸結合タンパク質及び前記第２のアミノ酸結合タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている、アミノ酸認識分子。
【請求項５３】
前記第１のアミノ酸結合タンパク質と前記第２のアミノ酸結合タンパク質とは、同じである、請求項５２に記載のアミノ酸認識分子。
【請求項５４】
前記第１のアミノ酸結合タンパク質と前記第２のアミノ酸結合タンパク質とは、互いに異なる、請求項５２に記載のアミノ酸認識分子。
【請求項５５】
前記第１のアミノ酸結合タンパク質及び前記第２のアミノ酸結合タンパク質は、それぞれ独立して、表１又は表２から選択されたアミノ酸配列に対して少なくとも８０％のアミノ酸配列同一性を有する、請求項５２～５４のいずれか一項に記載のアミノ酸認識分子。
【請求項５６】

前記リンカーは、１００までのアミノ酸を含む、請求項５２～５５のいずれか一項に記載のアミノ酸認識分子。
【請求項５７】
前記リンカーは、約５から約５０のアミノ酸を含む、請求項５２～５６のいずれか一項に記載のアミノ酸認識分子。
【請求項５８】
式（Ｉ）：
$(Z^1-X^1)_n-Z^2$
（Ｉ）
（式中、Z^1及びZ^2は、独立したアミノ酸結合タンパク質であり、X^1は、少なくとも２つのアミノ酸を含むリンカーであり、前記アミノ酸結合タンパク質は前記リンカーによって末端同士で結合されており、且つｎは１～５（両端の値を含む）の整数である）
のポリペプチドを含むアミノ酸認識分子。
【請求項５９】
Z^1及びZ^2は、同じタイプのアミノ酸結合タンパク質を含む、請求項５８に記載のアミノ酸認識分子。
【請求項６０】
Z^1及びZ^2は、互いに異なるタイプのアミノ酸結合タンパク質を含む、請求項５８に記載のアミノ酸認識分子。
【請求項６１】
Z^1及びZ^2は、独立して任意選択で、少なくとも１つの検出標識を含む標識成分に関連する、請求項５８～６０のいずれか一項に記載のアミノ酸認識分子。
【請求項６２】
前記ポリペプチドは、さらにタグ配列を含む、請求項５８～６１のいずれか一項に記載のアミノ酸認識分子。
【請求項６３】
前記タグ配列は、前記ポリペプチドの末端か、又は末端の近くにある、請求項６２に記載のアミノ酸認識分子。
【請求項６４】
前記タグ配列は、少なくとも１つのビオチンリガーゼ認識配列を含む、請求項６２又は６３に記載のアミノ酸認識分子。
【請求項６５】
前記タグ配列は、直列に配向された２つのビオチンリガーゼ認識配列を含む、請求項６２～６４のいずれか一項に記載のアミノ酸認識分子。
【請求項６６】
Z^1及びZ^2は、それぞれ独立して、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有する、請求項５８～６５のいずれか一項に記載のアミノ酸認識分子。
【請求項６７】
X^1は、１００までのアミノ酸を含む、請求項５８～６６のいずれか一項に記載のアミノ酸認識分子。
【請求項６８】
X^1は、約５から約５０のアミノ酸を含む、請求項５８～６７のいずれか一項に記載のアミノ酸認識分子。
【請求項６９】
末端同士で連結したアミノ酸結合タンパク質と標識付タンパク質とを有するポリペプチドを含むアミノ酸認識分子であって、前記アミノ酸結合タンパク質及び前記標識付タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている、アミノ酸認識分子。
【請求項７０】

前記リンカーは、１００までのアミノ酸を含む、請求項６９に記載のアミノ酸認識分子。
【請求項７１】
前記リンカーは、約５から約５０のアミノ酸を含む、請求項６９又は７０に記載のアミノ酸認識分子。
【請求項７２】
前記標識付タンパク質は、少なくとも１０ｋＤａの分子量を有する、請求項６９～７１のいずれか一項に記載のアミノ酸認識分子。
【請求項７３】
前記標識付タンパク質は、約１０ｋＤａ～約１５０ｋＤａの分子量を有する、請求項６９～７２のいずれか一項に記載のアミノ酸認識分子。
【請求項７４】
前記標識付タンパク質は、約１５ｋＤａ～約１００ｋＤａの分子量を有する、請求項６９～７３のいずれか一項に記載のアミノ酸認識分子。
【請求項７５】
前記標識付タンパク質は、少なくとも５０アミノ酸を含む、請求項６９～７４のいずれか一項に記載のアミノ酸認識分子。
【請求項７６】
前記標識付タンパク質は、約５０～約１，０００のアミノ酸を含む、請求項６９～７５のいずれか一項に記載のアミノ酸認識分子。
【請求項７７】
前記標識付タンパク質は、約１００～約７５０のアミノ酸を含む、請求項６９～７６のいずれか一項に記載のアミノ酸認識分子。
【請求項７８】
前記標識付タンパク質は、ＤＮＡポリメラーゼ、マルトース結合タンパク質、グルタチオンＳトランスフェラーゼ、緑色蛍光タンパク質、及びＳＮＡＰタグからなる群から選択されたタンパク質を含む、請求項６９～７７のいずれか一項に記載のアミノ酸認識分子。
【請求項７９】
前記標識付タンパク質は、発光標識を含む、請求項６９～７８のいずれか一項に記載のアミノ酸認識分子。
【請求項８０】
前記発光標識は、少なくとも１つのフルオロフォア色素分子を含む、請求項７９に記載のアミノ酸認識分子。
【請求項８１】
前記アミノ酸結合タンパク質は、Ｇｉｄタンパク質、ＵＢＲボックスタンパク質若しくはそのＵＢＲボックスドメイン含有断片、ｐ６２タンパク質若しくはそのＺＺドメイン含有断片、又はＣｌｐＳタンパク質である、請求項６９～８０のいずれか一項に記載のアミノ酸認識分子。
【請求項８２】
前記アミノ酸結合タンパク質は、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有する、請求項６９～８１のいずれか一項に記載のアミノ酸認識分子。
【請求項８３】
ポリペプチドシーケンシング方法であって、
反応混合物中の単一ポリペプチド分子と結合手段及び切断手段を含む組成物とを接触させることを含み、
前記結合手段及び前記切断手段は、前記結合手段と前記ポリペプチド上の末端アミノ酸との間の少なくとも１０の会合イベントを、前記ポリペプチドから前記末端アミノ酸を前記切断手段によって除去する前に達成するように構成される、方法。
【請求項８４】

前記結合手段及び前記切断手段は、少なくとも１０且つ１，０００までの会合イベントを、前記末端アミノ酸を除去する前に達成するように構成される、請求項８３に記載の方法。

【請求項８５】

前記結合手段及び前記切断手段は、５００まで、２００まで、又は１００までの会合イベントを、前記末端アミノ酸を除去する前に達成するように構成される、請求項８３又は８４に記載の方法。

【請求項８６】

前記末端アミノ酸は、前記少なくとも１０の会合イベントの前の切断イベントにて前記ポリペプチドの末端に露出した、請求項８３～８５のいずれか一項に記載の方法。

【請求項８７】

前記少なくとも１０の会合イベントは、前記切断イベントの後に起こる、請求項８６に記載の方法。

【請求項８８】

前記結合手段及び前記切断手段は、切断イベント間に少なくとも１分の時間インターバルを達成するように構成される、請求項８３～８７のいずれか一項に記載の方法。

【請求項８９】

前記時間インターバルは、約１分～約２０分、約５分～約１５分、又は約１分～約１０分である、請求項８８に記載の方法。

【請求項９０】

前記結合手段は１つ以上のアミノ酸認識分子を含み、前記切断手段は１つ以上のペプチダーゼ分子を含む、請求項８３～８９のいずれか一項に記載の方法。

【請求項９１】

アミノ酸認識分子とペプチダーゼ分子とのモル比は、少なくとも１０の会合イベントを、前記末端アミノ酸を除去する前に達成するように構成される、請求項９０に記載の方法。

【請求項９２】

前記アミノ酸認識分子と前記ペプチダーゼ分子とのモル比は、約１：１，０００～約１：１又は約１：１～約１００：１である、請求項９１に記載の方法。

【請求項９３】

前記アミノ酸認識分子と前記ペプチダーゼ分子とのモル比は、約１：１００～約１：１又は約１：１～約１０：１である、請求項９１に記載の方法。

【請求項９４】

前記アミノ酸認識分子と前記ペプチダーゼ分子とのモル比は、約１：１，０００、約１：５００、約１：２００、約１：１００、約１：１０、約１：５、約１：２、約１：１、約５：１、約１０：１、約５０：１、約１００：１である、請求項９１に記載の方法。

【請求項９５】

ポリペプチドの分解プロセス時のデータを得ることと、

前記分解プロセス時に前記ポリペプチドの末端に逐次露出されるアミノ酸に対応する前記データの部分であって、独立した前記部分の各々は、少なくとも１つのパルス持続時間を含むパルスパターンを含み、前記パルスパターンは約１ミリ秒～約１０秒の平均パルス持続時間を含む、前記データの部分を決定するように前記データを解析することと、

前記ポリペプチドに相当するアミノ酸配列を出力することと、を含む、方法。

【請求項９６】

前記平均パルス持続時間は、約５０ミリ秒～約２秒である、請求項９５に記載の方法。

【請求項９７】

前記平均パルス持続時間は、約５０ミリ秒～約５００ミリ秒又は約５００ミリ秒～約２秒である、請求項９５又は９６に記載の方法。

【請求項９８】

第１のタイプのアミノ酸の前記パルスパターンは、第２のタイプのアミノ酸の前記パル

スパターンと少なくとも１０ミリ秒の平均パルス持続時間だけ異なる、請求項９５～９７のいずれか一項に記載の方法。
【請求項９９】
第１のタイプのアミノ酸の前記パルスパターンは、第２のタイプのアミノ酸の前記パルスパターンと約１０ミリ秒～約１００ミリ秒又は約１００ミリ秒～約１０秒の平均パルス持続時間だけ異なる、請求項９５～９７のいずれか一項に記載の方法。
【請求項１００】
ポリペプチドシーケンシング方法であって、
単一ポリペプチド分子と１つ以上のアミノ酸認識分子とを接触させることと、
前記１つ以上のアミノ酸認識分子と、前記単一ポリペプチドが分解されつつある間に前記単一ポリペプチド分子の末端に露出される逐次アミノ酸と、の会合の指標となる一連のシグナルパルスを検出することにより、前記単一ポリペプチド分子をシーケンスすることと、を含み、
前記１つ以上のアミノ酸認識分子と前記末端に露出される各タイプのアミノ酸との会合が、前記末端に露出される他のタイプのアミノ酸とは異なる前記一連のシグナルパルスの特性パターンを生成し、前記特性パターンのシグナルパルスは約１ミリ秒～約１０秒の平均パルス持続時間を含む、方法。
【請求項１０１】
ポリペプチドをシーケンシングする方法であって、
反応混合物中の単一ポリペプチド分子と１つ以上のアミノ酸認識分子及び切断試薬を含む組成物とを接触させることと、
前記１つ以上のアミノ酸認識分子と前記単一ポリペプチド分子の末端との前記切断試薬の存在下での会合の指標となる一連のシグナルパルスであって、前記一連のシグナルパルスは、前記切断試薬による末端アミノ酸切断の結果として、経時的に前記末端にて曝露された一連のアミノ酸の指標である、一連のシグナルパルスを検出することを含み、
前記１つ以上のアミノ酸認識分子と前記末端に露出される各タイプのアミノ酸との会合が、前記末端に露出される他のタイプのアミノ酸とは異なる前記一連のシグナルパルスの特性パターンを生成し、前記特性パターンのシグナルパルスは約１ミリ秒～約１０秒の平均パルス持続時間を含む、方法。
【請求項１０２】
ポリペプチドシーケンス方法であって、
ａ）単一ポリペプチド分子の末端の第１のアミノ酸を同定することと、
ｂ）前記第１のアミノ酸を除去して前記単一ポリペプチド分子の前記末端の第２のアミノ酸を露出させることと、
ｃ）前記単一ポリペプチド分子の前記末端の前記第２のアミノ酸を同定することと、を含み、
（ａ）～（ｃ）は、１つ以上のアミノ酸認識分子を含む単一反応混合物中で実施され、
前記第１のアミノ酸及び前記第２のアミノ酸は、前記１つ以上のアミノ酸認識分子と、前記単一ポリペプチド分子の末端との会合の指標となる一連のシグナルパルスを検出することによって同定され、
前記１つ以上のアミノ酸認識分子と前記第１のアミノ酸との会合により、前記第２のアミノ酸とは異なる前記一連のシグナルパルスの特性パターンが生成され、前記特性パターンのシグナルパルスは約１ミリ秒～約１０秒の平均パルス持続時間を含む、方法。
【請求項１０３】
ポリペプチドのアミノ酸を同定する方法であって、
単一ポリペプチド分子と前記単一ポリペプチド分子に結合する１つ以上のアミノ酸認識分子とを接触させることと、
前記１つ以上のアミノ酸認識分子と前記単一ポリペプチド分子とのポリペプチド分解条件下での会合の指標となる一連のシグナルパルスを検出することと、
前記一連のシグナルパルスにおける特性パターンであって、前記特性パターンのシグナ

ルパルスは、約１ミリ秒～約１０秒の平均パルス持続時間を含む、前記一連のシグナルパルスにおける特性パターンに基づいて、前記単一ポリペプチド分子の第１のタイプのアミノ酸を同定することと、を含む、方法。

【請求項１０４】

前記平均パルス持続時間は、約５０ミリ秒～約２秒である、請求項１００～１０３のいずれか一項に記載の方法。

【請求項１０５】

前記平均パルス持続時間は、約５０ミリ秒～約５００ミリ秒又は約５００ミリ秒～約２秒である、請求項１００～１０４のいずれか一項に記載の方法。

【請求項１０６】

前記特性パターンは、少なくとも１０のシグナルパルスを含む、請求項１００～１０５のいずれか一項に記載の方法。

【請求項１０７】

少なくとも１タイプのアミノ酸の前記特性パターンは、約５０～約２００のシグナルパルスを含む、請求項１００～１０６のいずれか一項に記載の方法。

【請求項１０８】

少なくとも１タイプのアミノ酸の前記特性パターンは、約２５～約１００のシグナルパルスを含む、請求項１００～１０７のいずれか一項に記載の方法。

【請求項１０９】

第１のタイプのアミノ酸の前記特性パターンは、第２のタイプのアミノ酸の前記特性パターンと少なくとも１０ミリ秒の平均パルス持続時間だけ異なる、請求項１００～１０８のいずれか一項に記載の方法。

【請求項１１０】

第１のタイプのアミノ酸の前記特性パターンは、第２のタイプのアミノ酸の前記特性パターンと約１０ミリ秒～約１００ミリ秒又は約１００ミリ秒～約１秒の平均パルス持続時間だけ異なる、請求項１００～１０９のいずれか一項に記載の方法。

【請求項１１１】

前記方法は、前記１つ以上のアミノ酸認識分子及び１つ以上の切断試薬を含む組成物を含む反応混合物中で実施される、請求項１００～１１０のいずれか一項に記載の方法。

【請求項１１２】

前記反応混合物中のアミノ酸認識分子と切断試薬とのモル比は、約１：１，０００～約１：１又は約１：１～約１００：１である、請求項１１１に記載の方法。

【請求項１１３】

前記反応混合物中のアミノ酸認識分子と切断試薬とのモル比は、約１：１００～約１：１又は約１：１～約１０：１である、請求項１１１に記載の方法。

【請求項１１４】

前記反応混合物中のアミノ酸認識分子と切断試薬とのモル比は、約１：１，０００、約１：５００、約１：２００、約１：１００、約１：１０、約１：５、約１：２、約１：１、約５：１、約１０：１、約５０：１、約１００：１である、請求項１１１に記載の方法。

【請求項１１５】

前記１つ以上のアミノ酸認識分子は、１つ以上の末端アミノ酸認識分子を含む、請求項１００～１１４のいずれか一項に記載の方法。

【請求項１１６】

前記１つ以上のアミノ酸認識分子の各々は、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有するアミノ酸結合タンパク質を含む、請求項１００～１１５のいずれか一項に記載の方法。

【請求項１１７】

少なくとも１つのハードウェアプロセッサーと、

前記少なくとも１つのハードウェアプロセッサーによって実行されるときに前記少なく

とも１つのハードウェアプロセッサーに請求項１００～１１６のいずれか一項に記載の方法を実施させるプロセッサー実行可能命令を記憶する少なくとも１つの非一時コンピュータ可読記憶媒体と、を含む、システム。

【請求項１１８】

少なくとも１つのハードウェアプロセッサーによって実行されるときに前記少なくとも１つのハードウェアプロセッサーに請求項１００～１１６のいずれか一項に記載の方法を実施させるプロセッサー実行可能命令を記憶する、少なくとも１つの非一時コンピュータ可読記憶媒体。

【請求項１１９】

式（ＩＩ）：

Ａ－（Ｙ）$_{n}$－Ｄ

（ＩＩ）

（式中、Ａは、表１又は表２から選択されたアミノ酸配列に対して少なくとも８０％のアミノ酸配列同一性を有する少なくとも１つのアミノ酸結合タンパク質を含むアミノ酸結合コンポーネントであり、

Ｙの各々は、共有結合性又は非共有結合性連結基を形成するポリマーであり、

ｎは、０～１０（両端の値を含む）の整数であり、

且つＤは、少なくとも１つの検出標識を含む標識コンポーネントであって、Ｄは２０ｎｍ（２００Å）未満の直径である）

の可溶性アミノ酸認識分子を含む組成物。

【請求項１２０】

前記可溶性アミノ酸認識分子は、末端同士で連結したＡ及びＹ^{1}を有するポリペプチドを含み、Ａ及びＹ^{1}は少なくとも２つのアミノ酸を含むリンカーによって分離されている、請求項１１９に記載の組成物。

【請求項１２１】

前記リンカーは、１００までのアミノ酸を含む、請求項１２０に記載の組成物。

【請求項１２２】

前記リンカーは、約５～約５０のアミノ酸を含む、請求項１２０又は１２１に記載の組成物。

【請求項１２３】

Ｙ^{1}は、少なくとも１０ｋＤａの分子量を有するタンパク質である、請求項１１９～１２２のいずれか一項に記載の組成物。

【請求項１２４】

Ｙ^{1}は、約１０ｋＤａ～約１５０ｋＤａの分子量を有するタンパク質である、請求項１１９～１２３のいずれか一項に記載の組成物。

【請求項１２５】

Ｙ^{1}は、約１５ｋＤａ～約１００ｋＤａの分子量を有するタンパク質である、請求項１１９～１２４のいずれか一項に記載の組成物。

【請求項１２６】

Ｙ^{1}は、少なくとも５０アミノ酸を含むタンパク質である、請求項１１９～１２５のいずれか一項に記載の組成物。

【請求項１２７】

Ｙ^{1}は、約５０～約１，０００のアミノ酸を含むタンパク質である、請求項１１９～１２６のいずれか一項に記載の組成物。

【請求項１２８】

Ｙ^{1}は、約１００～約７５０のアミノ酸を含むタンパク質である、請求項１１９～１２７のいずれか一項に記載の組成物。

【請求項１２９】

Ｙ^{1}は、ＤＮＡポリメラーゼ、マルトース結合タンパク質、グルタチオンＳトランスフェラーゼ、緑色蛍光タンパク質、及びＳＮＡＰタグからなる群から選択されたタンパク質

である、請求項１１９～１２８のいずれか一項に記載の組成物。
【請求項１３０】
Ａは、末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポリペプチドを含み、前記第１のアミノ酸結合タンパク質及び前記第２のアミノ酸結合タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている、請求項１１９～１２９のいずれか一項に記載の組成物。
【請求項１３１】
前記第１のアミノ酸結合タンパク質と前記第２のアミノ酸結合タンパク質とは、同じである、請求項１３０に記載の組成物。
【請求項１３２】
前記第１のアミノ酸結合タンパク質と前記第２のアミノ酸結合タンパク質とは、互いに異なる、請求項１３０又は１３１に記載の組成物。
【請求項１３３】
前記第１のアミノ酸結合タンパク質及び前記第２のアミノ酸結合タンパク質は、それぞれ独立して、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有する、請求項１３０～１３２のいずれか一項に記載の組成物。
【請求項１３４】
前記リンカーは、１００までのアミノ酸を含む、請求項１３０～１３３のいずれか一項に記載の組成物。
【請求項１３５】
前記リンカーは、約５～約５０のアミノ酸を含む、請求項１３０～１３４のいずれか一項に記載の組成物。
【請求項１３６】
式（ＩＩＩ）：
Ａ－Y^1－Ｄ
（ＩＩＩ）
（式中、Ａは、表１又は表２から選択されたアミノ酸配列に対して少なくとも８０％のアミノ酸配列同一性を有する少なくとも１つのアミノ酸結合タンパク質を含むアミノ酸結合コンポーネントであり、
Ｙ１は、核酸又はポリペプチドであり、
且つＤは、少なくとも１つの検出標識を含む標識コンポーネントであり、
Y^1が核酸であるとき、前記核酸は共有結合性又は非共有結合性連結基を形成し、且つ、
Y^1がポリペプチドであるとき、前記ポリペプチドは、５０×10^{-9}Ｍ未満の解離定数（K_D）により特徴付けられる非共有結合性連結基を形成する）
のアミノ酸認識分子。
【請求項１３７】
Ａは、末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポリペプチドを含み、前記第１のアミノ酸結合タンパク質及び前記第２のアミノ酸結合タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている、請求項１３６に記載のアミノ酸認識分子。
【請求項１３８】
前記第１のアミノ酸結合タンパク質と前記第２のアミノ酸結合タンパク質とは、同じである、請求項１３７に記載のアミノ酸認識分子。
【請求項１３９】
前記第１のアミノ酸結合タンパク質と前記第２のアミノ酸結合タンパク質とは、互いに異なる、請求項１３７又は１３８に記載のアミノ酸認識分子。
【請求項１４０】
前記第１のアミノ酸結合タンパク質及び前記第２のアミノ酸結合タンパク質は、それぞれ独立して、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ

酸配列を有する、請求項１３７～１３９のいずれか一項に記載のアミノ酸認識分子。
【請求項１４１】
前記リンカーは、１００までのアミノ酸を含む、請求項１３７～１４０のいずれか一項に記載のアミノ酸認識分子。
【請求項１４２】
前記リンカーは、約５～約５０のアミノ酸を含む、請求項１３７～１４１のいずれか一項に記載のアミノ酸認識分子。
【請求項１４３】
核酸と、
前記核酸上の第１の装着部位に装着された少なくとも１つのアミノ酸認識分子であって、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有するアミノ酸結合タンパク質を含む、少なくとも１つのアミノ酸認識分子と、
前記核酸上の第２の装着部位に装着された少なくとも１つの検出可能標識と、を含むアミノ酸認識分子であって、
前記核酸は、前記少なくとも１つのアミノ酸認識分子と前記少なくとも１つの検出可能標識との間に共有結合性又は非共有結合性連結基を形成する、アミノ酸認識分子。
【請求項１４４】
少なくとも２つのリガンド結合部位を含む多価タンパク質と、
前記タンパク質上の第１のリガンド結合部位に結合された第１のリガンド部分を介して前記タンパク質に装着された少なくとも１つのアミノ酸認識分子であって、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有するアミノ酸結合タンパク質を含む、少なくとも１つのアミノ酸認識分子と、
前記タンパク質上の第２のリガンド結合部位に結合された第２のリガンド部分を介して前記タンパク質に装着された少なくとも１つの検出可能標識と、を含む、アミノ酸認識分子。
【発明の詳細な説明】
【背景技術】
【０００１】
プロテオミクスは、生物学的系の研究においてゲノミクス及びトランスクリプトミクスに対する重要且つ必要な補完物として現われた。個別生物のプロテオミック解析は、改善された診断及び療法ストラテジーをもたらす細胞プロセス及び反応パターンについての洞察を提供可能である。タンパク質構造、組成、及び修飾を取り巻く複雑性は、生物学的サンプルに対する大規模タンパク質シーケンシング情報を決定するうえで課題を呈する。
【発明の概要】
【０００２】
いくつかの態様では、本願は、ポリペプチドからアミノ酸配列情報を決定するための（たとえば、１つ以上のポリペプチドをシーケンスするための）方法及び組成物を提供する。いくつかの実施形態では、アミノ酸配列情報は、単一ポリペプチド分子に対して決定可能である。いくつかの実施形態では、たとえば、単一ポリペプチド分子に対して、ポリペプチドの２つ以上のアミノ酸の相対位置が決定される。いくつかの実施形態では、ポリペプチドの１つ以上のアミノ酸が標識されて（たとえば、直接的又は間接的に）、ポリペプチドの標識付アミノ酸の相対位置が決定される。いくつかの実施形態では、アミノ酸配列情報は、ポリペプチドと１つ以上のアミノ酸認識分子（たとえば、１つ以上のアミノ酸結合タンパク質）との相互作用を検出することによって、決定されうる。
【０００３】
いくつかの態様では、本願は、ポリペプチドからアミノ酸配列情報を決定するための方法において使用されうるアミノ酸結合タンパク質を提供する。いくつかの態様では、本願は、表１又は表２から選択された配列に少なくとも８０％同一のアミノ酸配列を有するとともに１つ以上の標識を含む、組換えアミノ酸結合タンパク質を提供する。いくつかの態様では、１つ以上の標識は、発光標識又は伝導率標識を含む。いくつかの実施形態では、

１つ以上標識は、タグ配列を含む。いくつかの実施形態では、タグ配列は、精製タグ、切断部位、及びビオチン化配列のうちの１つ以上を含む（たとえば、少なくとも１つのビオチンリガーゼ認識配列）。いくつかの実施形態では、ビオチン化配列は、直列に配向された２つのビオチンリガーゼ認識配列を含む。いくつかの実施形態では、１つ以上標識は、少なくとも１つのビオチン分子を有するビオチン部分を含む（たとえば、ビスビオチン部分）。いくつかの実施形態では、標識は、少なくとも１つのビオチン分子が装着された少なくとも１つのビオチンリガーゼ認識配列を含む。いくつかの実施形態では、１つ以上の標識は、１つ以上のポリオール部分を含む（たとえば、ポリエチレングリコール）。いくつかの実施形態では、組換えアミノ酸結合タンパク質は、１つ以上の標識が装着された１つ以上の非天然アミノ酸を含む。いくつかの態様では、本願は、本明細書に記載の組換えアミノ酸結合タンパク質を含む組成物を提供する。

【０００４】

いくつかの態様では、本願は、２つ以上のアミノ酸認識分子を含むポリペプチドシーケンシング反応組成物であって、２つ以上のアミノ酸認識分子のうちの少なくとも１つは、本明細書に記載の組換えアミノ酸結合タンパク質である、ポリペプチドシーケンシング反応組成物を提供する。いくつかの実施形態では、２つ以上のアミノ酸認識分子は、互いに異なるタイプのアミノ酸認識分子を含む。たとえば、いくつかの実施形態では、１タイプのアミノ酸認識分子は、ポリペプチドシーケンシング反応組成物中の他のタイプのアミノ酸認識分子とは異なる（たとえば、検出できるほどに異なる）方法で対象ポリペプチドと相互作用する。いくつかの実施形態では、ポリペプチドシーケンシング反応組成物は、少なくとも１タイプの切断試薬を含む。いくつかの態様では、本願は、ポリペプチドと本明細書に記載のポリペプチドシーケンシング反応組成物とを接触させることを含む、ポリペプチドシーケンシングの方法を提供する。いくつかの実施形態では、方法は、ポリペプチドが分解されつつある間にポリペプチドと少なくとも１つのアミノ酸認識分子との一連の相互作用を検出することによってポリペプチドをシーケンシングすることをさらに含む。

【０００５】

いくつかの態様では、本願は、アミノ酸結合タンパク質及びペプチダーゼを含むポリペプチドシーケンシング反応混合物を提供する。いくつかの実施形態では、標識付アミノ酸結合タンパク質とペプチダーゼとのモル比は、約１：１，０００～約１：１又は約１：１～約１００：１である。いくつかの実施形態では、アミノ酸結合タンパク質は、１つ以上の標識を含む。いくつかの実施形態では、アミノ酸結合タンパク質は、ＣｌｐＳタンパク質である。いくつかの実施形態では、アミノ酸結合タンパク質は、表１又は表２から選択された配列に少なくとも８０％、８０～９０％、９０～９５％、又は少なくとも９５％同一のアミノ酸配列を有するタンパク質である。いくつかの実施形態では、ペプチダーゼは、エキソペプチダーゼである。いくつかの実施形態では、ペプチダーゼは、表４又は表５から選択された配列に少なくとも８０％、８０～９０％、９０～９５％、又は少なくとも９５％同一のアミノ酸配列を有する酵素である。いくつかの実施形態では、反応混合物は、１つ以上のアミノ酸結合タンパク質及び／又は１つ以上のペプチダーゼを含む。いくつかの実施形態では、反応混合物は、表面に固定されたペプチド分子を含む。

【０００６】

いくつかの態様では、本願は、単一ペプチド分子、少なくとも１つのペプチダーゼ分子、及び少なくとも３つのアミノ酸認識分子を含む、ポリペプチドシーケンシング反応混合物を提供する。いくつかの実施形態では、反応混合物は、少なくとも１且つ１０までのペプチダーゼ分子（たとえば、少なくとも１つ且つ５つまでのペプチダーゼ分子、少なくとも１つ且つ３つまでのペプチダーゼ分子）を含む。いくつかの実施形態では、反応組成物は、２つ以上のペプチダーゼ分子を含み、各ペプチダーゼ分子は、互いに異なるタイプである。たとえば、いくつかの実施形態では、１タイプのペプチダーゼ分子は、反応混合物中の他のタイプのペプチダーゼ分子のとは異なる切断優先度を有する。いくつかの実施形態では、反応混合物は、少なくとも３つ且つ３０までのアミノ酸認識分子を含む（たとえば、２０まで、１０まで、又は５つまでのアミノ酸認識分子）。いくつかの実施形態では

、少なくとも３つのアミノ酸認識分子は、互いに異なるタイプのアミノ酸認識分子を含む。たとえば、いくつかの実施形態では、１タイプのアミノ酸認識分子は、反応混合物中の他のタイプのアミノ酸認識分子とは異なる（たとえば、検出できるほどに異なる）方法で対象ポリペプチドと相互作用する。

【０００７】

いくつかの態様では、本願は、サンプルウェルのアレイを含む基体であって、アレイの少なくとも１つのサンプルウェルは、本明細書に記載のポリペプチドシーケンシング反応混合物を含む、基体を提供する。いくつかの実施形態では、少なくとも１つのサンプルウェルは、ボトム表面を含む。いくつかの実施形態では、単一ポリペプチド分子は、ボトム表面に固定される。

【０００８】

いくつかの態様では、本願は、末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポリペプチドを含むアミノ酸認識分子であって、第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている、アミノ酸認識分子を提供する。いくつかの実施形態では、第１のアミノ酸結合タンパク質と第２のアミノ酸結合タンパク質とは、同じである。いくつかの実施形態では、第１のアミノ酸結合タンパク質と第２のアミノ酸結合タンパク質とは、互いに異なる。

【０００９】

いくつかの態様では、本願は、式（Ｉ）：

$$（Z^{1}－X^{1}）_{n}－Z^{2}$$

（Ｉ）

（式中、Z^{1}及びZ^{2}は、独立したアミノ酸結合タンパク質であり、X^{1}は、少なくとも２つのアミノ酸を含むリンカーであり、アミノ酸結合タンパク質はリンカーによって末端同士で結合されており、且つｎは１～５（両端の値を含む）の整数である）

のポリペプチドを含むアミノ酸認識分子を提供する。いくつかの実施形態では、Z^{1}及びZ^{2}は、同じタイプのアミノ酸結合タンパク質を含む。いくつかの実施形態では、Z^{1}及びZ^{2}は、互いに異なるタイプのアミノ酸結合タンパク質を含む。いくつかの実施形態では、Z^{1}及びZ^{2}は、独立して任意選択で、少なくとも１つの検出可能標識を含む標識成分と関連する。いくつかの実施形態では、ポリペプチドは、タグ配列をさらに含む。

【００１０】

いくつかの態様では、本願は、ポリペプチドシーケンシングの方法を提供する。いくつかの実施形態では、ポリペプチドシーケンシングの方法は、反応混合物中の単一ポリペプチド分子と結合手段及び切断手段を含む組成物とを接触させることを含む。いくつかの実施形態では、結合手段及び切断手段は、結合手段とポリペプチド上の末端アミノ酸との間の少なくとも１０の会合イベントを、ペプチドから末端アミノ酸を切断手段によって除去する前に達成するように構成される。いくつかの実施形態では、結合手段及び切断手段は、少なくとも１０且つ１，０００までの会合イベントを、末端アミノ酸を除去する前に達成するように構成される。いくつかの実施形態では、末端アミノ酸は、少なくとも１０の会合イベントの前の切断イベントにてポリペプチドの末端に露出された。いくつかの実施形態では、少なくとも１０の会合イベントは、切断イベントの後に起こる。

【００１１】

いくつかの実施形態では、結合手段及び切断手段は、切断イベント間に少なくとも１分の時間インターバルを達成するように構成される（たとえば、約１分～約２０分、約５分～約１５分、又は約１分～約１０分）。いくつかの実施形態では、結合手段は１つ以上のアミノ酸認識分子を含み、切断手段は１つ以上のペプチダーゼ分子を含む。いくつかの実施形態では、アミノ酸認識分子とペプチダーゼ分子とのモル比は、少なくとも１０の会合イベントを、末端アミノ酸を除去する前に達成するように構成される。いくつかの実施形態では、アミノ酸認識分子とペプチダーゼ分子とのモル比は、約１：１，０００～約１：１又は約１：１～約１００：１である。いくつかの実施形態では、アミノ酸認識分子とペ

プチダーゼ分子とのモル比は、約１：１００～約１：１又は約１：１～約１０：１である。

【００１２】

いくつかの態様では、本願は、サンプルウェルのアレイを含む基体であって、アレイの少なくとも１つのサンプルウェルは、単一ポリペプチド分子、切断手段、及び結合手段を含む、基体を提供する。いくつかの実施形態では、結合手段及び切断手段は、結合手段とポリペプチド上の末端アミノ酸との間の少なくとも１０の会合イベントを、ペプチドから末端アミノ酸を切断手段によって除去する前に達成するように構成される。いくつかの実施形態では、結合手段及び切断手段は、少なくとも１０且つ１，０００までの会合イベントを、末端アミノ酸を除去する前に達成するように構成される。いくつかの実施形態では、末端アミノ酸は、少なくとも１０の会合イベントの前の切断イベントにてポリペプチドに末端に露出された。いくつかの実施形態では、少なくとも１０の会合イベントは、切断イベントの後に起こる。

【００１３】

いくつかの態様では、本願は、たとえば、ポリペプチドシーケンシング反応時の光安定性増強のために、シールディング（ｓｈｉｅｌｄｉｎｇ）エレメントを含むアミノ酸認識分子を提供する。いくつかの態様では、本願は、末端同士で連結したアミノ酸結合タンパク質及び標識付タンパク質を有するポリペプチドを含むアミノ酸認識分子を提供する。いくつかの実施形態では、アミノ酸結合タンパク質及び標識付タンパク質は、少なくとも２つのアミノ酸（たとえば、少なくとも２つ且つ１００までのアミノ酸、約５つ～約５０のアミノ酸）を含むリンカーによって分離されている。いくつかの実施形態では、標識付タンパク質は、少なくとも１０ｋＤａ（たとえば、約１０ｋＤａ～約１５０ｋＤａ、約１５ｋＤａから約１００ｋＤａ）の分子量を有する。いくつかの実施形態では、標識付タンパク質は、少なくとも５０のアミノ酸（たとえば、約５０～約１，０００のアミノ酸、約１００から約７５０のアミノ酸）を含む。いくつかの実施形態では、標識付タンパク質は、発光標識を含む。いくつかの実施形態では、発光標識は、少なくとも１つのフルオロフォア色素分子を含む。いくつかの実施形態では、アミノ酸結合タンパク質は、Ｇｉｄタンパク質、ＵＢＲボックスタンパク質若しくはそのＵＢＲボックスドメイン含有断片、ｐ６２タンパク質若しくはそのＺＺドメイン含有断片、又はＣｌｐＳタンパク質である。いくつかの実施形態では、アミノ酸結合タンパク質は、表１又は表２から選択されたアミノ酸配列に対して少なくとも８０％のアミノ酸配列同一性を有する。

【００１４】

いくつかの態様では、本願は、式（ＩＩ）：

Ａ－（Ｙ）$_{n}$－Ｄ

（ＩＩ）

（式中、Ａは、少なくとも１つのアミノ酸認識分子を含むアミノ酸結合成分であり；Ｙの各々は、共有結合性又は非共有結合性連結基を形成するポリマーであり；ｎは、１～１０（両端の値を含む）の整数であり；且つ、Ｄは、少なくとも１つの検出可能標識を含む標識成分である）

のアミノ酸認識分子を提供する。いくつかの実施形態では、Ａは、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有する少なくとも１つのアミノ酸結合タンパク質を含む。いくつかの実施形態では、アミノ酸認識分子は、末端同士で連結したＡ及びＹ^{1}を有するポリペプチドを含み、Ａ及びＹ^{1}は少なくとも２つのアミノ酸を含むリンカーによって分離されている。いくつかの実施形態では、Ｙ^{1}は、少なくとも１０ｋＤａ（たとえば、約１０ｋＤａ～約１５０ｋＤａ）の分子量を有するタンパク質である。いくつかの実施形態では、Ｙ^{1}は、少なくとも５０のアミノ酸（たとえば、約５０～約１，０００のアミノ酸）を含むタンパク質である。

【００１５】

いくつかの実施形態では、Ｄは２０ｎｍ（２００Å）未満の直径である。いくつかの実施形態では、－（Ｙ）$_{n}$－は、少なくとも２ｎｍの長さ（たとえば、少なくとも５ｎｍ、

少なくとも１０ｎｍ、少なくとも２０ｎｍ、少なくとも３０ｎｍ、少なくとも５０ｎｍ、又はそれ以上の長さ）である。いくつかの実施形態では、－（Ｙ）$_{n}$－は、約２ｎｍ～約２００ｎｍの長さ（たとえば、約２ｎｍ～約１００ｎｍ、約５ｎｍ～約５０ｎｍ、又は約１０ｎｍ～約１００ｎｍの長さ）である。いくつかの実施形態では、Ｙの各々は、独立して、バイオ分子又はデンドリティックポリマー（たとえば、ポリオール、デンドリマー）である。いくつかの実施形態では、Ａは、末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポチペプチドを含む（たとえば、融合ポリペプチド）。いくつかの実施形態では、本願は、式（ＩＩ）のアミノ酸認識分子を含む組成物を提供する。いくつかの実施形態では、アミノ酸認識分子は組成物に可溶である。

【００１６】

いくつかの態様では、本願は、式（ＩＩＩ）：

Ａ－Ｙ^{1}－Ｄ

（ＩＩＩ），

（式中、Ａは、少なくとも１つのアミノ酸認識分子を含むアミノ酸結合成分であり；Ｙ^{1}は、核酸又はポリペプチドであり；Ｄは、少なくとも１つの検出可能標識を含む標識成分である）

のアミノ酸認識分子を提供する。いくつかの実施形態では、Ａは、表１又は表２から選択されたアミノ酸配列に少なくとも８０％同一のアミノ酸配列を有する少なくとも１つのアミノ酸結合タンパク質を含む。いくつかの実施形態では、Ｙ^{1}が核酸であるとき、核酸は共有結合性又は非共有結合性連結基を形成する。いくつかの実施形態では、Ｙ^{1}がポリペプチドであるときならば、ポリペプチドは、５０×１０$^{-9}$Ｍ未満の解離定数（Ｋ$_{Ｄ}$）により特徴付けられる非共有結合性連結基を形成する。いくつかの実施形態では、Ｋ$_{Ｄ}$は、１×１０$^{-9}$Ｍ未満、１×１０$^{-10}$Ｍ未満、１×１０$^{-11}$Ｍ未満、又は１×１０$^{-12}$Ｍ未満である。

【００１７】

いくつかの態様では、本願は、核酸と、核酸上の第１の装着部位に装着された少なくとも１つのアミノ酸認識分子と、核酸上の第２の装着部位に装着された少なくとも１つの検出可能標識と、を含むアミノ酸認識分子であって、核酸が少なくとも１つのアミノ酸認識分子と少なくとも１つの検出可能標識との間に共有結合性又は非共有結合性連結基を形成する、アミノ酸認識分子を提供する。いくつかの実施形態では、核酸は、第１のオリゴヌクレオチド鎖を含む。いくつかの実施形態では、核酸は、第１のオリゴヌクレオチド鎖にハイブリダイズされた第２のオリゴヌクレオチド鎖をさらに含む。いくつかの実施形態では、少なくとも１つのアミノ酸認識分子は、末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポチペプチドを含む（たとえば、融合ポリペプチド）。いくつかの実施形態では、第１のアミノ酸結合タンパク質と第２のアミノ酸結合タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている。

【００１８】

いくつかの態様では、本願は、少なくとも２つのリガンド結合部位を含む多価タンパク質と、タンパク質上の第１のリガンド結合部位に結合された第１のリガンド部分を介してタンパク質に装着された少なくとも１つのアミノ酸認識分子と、タンパク質上の第２のリガンド結合部位に結合された第２のリガンド部分を介してタンパク質に装着された少なくとも１つの検出可能標識と、を含むアミノ酸認識分子を提供する。いくつかの実施形態では、多価タンパク質はアビジンタンパク質である。いくつかの実施形態では、少なくとも１つのアミノ酸認識分子は、末端同士で連結した少なくとも第１のアミノ酸結合タンパク質及び第２のアミノ酸結合タンパク質を有するポチペプチドを含む（たとえば、融合ポリペプチド）。いくつかの実施形態では、第１のアミノ酸結合タンパク質と第２のアミノ酸結合タンパク質は、少なくとも２つのアミノ酸を含むリンカーによって分離されている。

【００１９】

いくつかの実施形態では、シールドアミノ酸認識分子は、本願に係るポリペプチドシーケンシング方法又は当技術分野で公知のいずれかの方法で使用されうる。それゆえ、いくつかの態様では、本願は、ポリペプチド分子と本願の１つ以上のシールドアミノ酸認識分子とを接触させることを含むポリペプチドシーケンシング方法を提供する（たとえば、エドマン型分解反応で、動的シーケンシング反応で、又は当技術分野で公知の他の方法で）。たとえば、いくつかの実施形態では、本方法は、ポリペプチド分子と、本願に係るシールド又はシールディングエレメントを含む少なくとも１つのアミノ酸認識分子と、を接触させることと、少なくとも１つのアミノ酸認識分子とポリペプチド分子との会合を検出することと、を含む。

【００２０】

いくつかの態様では、本願は、ポリペプチドの分解プロセス時のデータを得ることを含む方法を提供する。いくつかの実施形態では、本方法は、分解プロセス時にポリペプチドの末端に逐次露出されるアミノ酸に対応するデータの部分を決定するようにデータを解析することをさらに含む。いくつかの実施形態では、本方法は、ポリペプチドに相当するアミノ酸配列を出力することをさらに含む。いくつかの実施形態では、データは、分解プロセス時のポリペプチドの末端のアミノ酸アイデンティティー（ｉｄｅｎｔｉｔｙ）の指標となる。いくつかの実施形態では、データは、分解プロセス時に末端の異なるタイプの末端アミノ酸に結合する１つ以上のアミノ酸認識分子により生成されるシグナルの指標となる。いくつかの実施形態では、データは、分解プロセス時に発生する発光シグナルの指標となる。いくつかの実施形態では、データは、分解プロセス時に発生する電気シグナルの指標となる。

【００２１】

いくつかの実施形態では、データを解析することは、一連の切断イベントを検出することと、逐次切断イベント間のデータの部分を決定することと、をさらに含む。いくつかの実施形態では、データを解析することは、個別部分の各々に対してアミノ酸のタイプを決定することをさらに含む。いくつかの実施形態では、個別部分の各々は、パルスパターン（たとえば特性パターン）を含み、且つデータを解析することは、そのそれぞれのパルスパターンに基づいて１つ以上の部分に対してアミノ酸のタイプを決定することをさらに含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、データが閾値超であるときの部分内の時間量を同定することと、部分に対して時間量と持続時間とを比較することと、をさらに含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、１つ以上の部分の各々に対して少なくとも１つのパルス持続時間を同定することをさらに含む。いくつかの実施形態では、パルスパターンは、約１ミリ秒～約１０秒の平均パルス持続時間を含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、１つ以上の部分の各々に対して少なくとも１つのパルス間持続時間を同定することをさらに含む。いくつかの実施形態では、アミノ酸配列は、部分に対応する一連のアミノ酸を含む。

【００２２】

いくつかの態様では、本願は、単一ポリペプチド分子と１つ以上のアミノ酸認識分子（たとえば、一つ以上の末端アミノ酸認識分子）とを接触させることを含むポリペプチドシーケンシング方法を提供する。いくつかの実施形態では、方法は、１つ以上のアミノ酸認識分子と、単一ポリペプチドが分解されつつある間に単一ポリペプチド分子の末端に露出される逐次アミノ酸との会合の指標となる一連のシグナルパルスを検出することにより、単一ポリペプチド分子についての配列情報を得ることをさらに含む。いくつかの実施形態では、ほとんど又は全ての単一ポリペプチド分子のアミノ酸配列が、決定される。いくつかの実施形態では、一連のシグナルパルスは、一連のリアルタイムシグナルパルスである。

【００２３】

いくつかの実施形態では、１つ以上のアミノ酸認識分子と末端に露出される各タイプのアミノ酸との会合は、末端に露出される他のタイプのアミノ酸とは異なる一連のシグナルパルスの特性パターンを生成する。いくつかの実施形態では、特性パルスのシグナルパル

スは、約１ミリ秒～約１０秒の平均パルス持続時間を含む。いくつかの実施形態では、特性パターンのシグナルパルスは、アミノ酸認識分子と末端に露出されるアミノ酸との間の個別会合イベントに対応する。いくつかの実施形態では、特性パターンは、単一ポリペプチドの末端に露出されるアミノ酸との一連の可逆アミノ酸認識分子結合相互作用に対応する。いくつかの実施形態では、特性パターンは、単一ポリペプチド分子の末端に露出されるアミノ酸及び連続位置のアミノ酸（たとえば、同じタイプ又は互いに異なるタイプのアミノ酸）の指標となる。

【００２４】

いくつかの実施形態では、単一ポリペプチド分子は、単一ポリペプチド分子の末端から少なくとも１つのアミノ酸を除去する切断試薬によって分解される。いくつかの実施形態では、方法は、切断試薬と末端との会合の指標となるシグナルを検出することをさらに含む。いくつかの実施形態では、切断試薬は、検出可能標識を含む（たとえば、発光標識、伝導率標識）。いくつかの実施形態では、単一ポリペプチド分子は、表面に固定される。いくつかの実施形態では、単一ポリペプチド分子は、１つ以上のアミノ酸認識分子が会合する末端から遠位の末端を介して表面に固定される。いくつかの実施形態では、単一ポリペプチド分子は、リンカー（たとえば、バイオ分子を含む可溶化性リンカー）を介して表面に固定される。

【００２５】

いくつかの態様では、本願は、反応混合物中の単一ポリペプチド分子と１つ以上のアミノ酸認識分子（たとえば、一つ以上の末端アミノ酸認識分子）及び切断試薬を含む組成物とを接触させることを含む、ポリペプチドをシーケンシングする方法を提供する。いくつかの実施形態では、方法は、切断試薬存在下で１つ以上のアミノ酸認識分子と単一ポリペプチド分子の末端との会合の指標となる一連のシグナルパルスを検出することをさらに含む。いくつかの実施形態では、一連のシグナルパルスは、切断試薬による末端アミノ酸切断の結果として、経時的に末端に露出されるの一連のアミノ酸の指標である。

【００２６】

いくつかの態様では、本願は、（ａ）単一ポリペプチド分子の末端の第１のアミノ酸を同定することと、（ｂ）第１のアミノ酸を除去して単一ポリペプチド分子の末端の第２のアミノ酸を露出させることと、（ｃ）単一ポリペプチド分子の末端の第２のアミノ酸を同定することと、を含む、ポリペプチドをシーケンシングする方法を提供する。いくつかの実施形態では、（ａ）～（ｃ）は、単一反応混合物中で実施される。いくつかの実施形態では、（ａ）～（ｃ）は、連続して起こる。いくつかの実施形態では、（ｃ）は、（ａ）及び（ｂ）の前に起こる。いくつかの実施形態では、単一反応混合物は、１つ以上のアミノ酸認識分子を含む（たとえば、１つ以上の末端アミノ酸認識分子）。いくつかの実施形態では、単一反応混合物は、切断試薬を含む。いくつかの実施形態では、第１のアミノ酸は、切断試薬によって除去される。いくつかの実施形態では、方法は、単一ポリペプチド分子の末端の１つ以上のアミノ酸を除去及び同定する工程を繰り返すことによって、単一ポリペプチド分子の配列（たとえば、部分配列又は完全配列）を同定することをさらに含む。

【００２７】

いくつかの態様では、本願は、単一ポリペプチド分子と単一ポリペプチド分子に結合する１つ以上のアミノ酸認識分子とを接触させることを含む、ポリペプチドのアミノ酸を同定する方法を提供する。いくつかの実施形態では、方法は、１つ以上のアミノ酸認識分子と単一ポリペプチド分子とのポリペプチド分解条件下での会合の指標となる一連のシグナルパルスを検出することをさらに含む。いくつかの実施形態では、方法は、一連のシグナルパルスの第１の特性パターンに基づいて単一ポリペプチド分子の第１のタイプのアミノ酸を同定することをさらに含む。いくつかの実施形態では、特性パルスのシグナルパルスは、約１ミリ秒～約１０秒の平均パルス持続時間を含む。

【００２８】

いくつかの態様では、本願は、ポリペプチドの末端アミノ酸（たとえば、Ｎ末端又はＣ

末端のアミノ酸）を同定する方法を提供する。いくつかの実施形態では、方法は、ポリペプチドと１タイプ以上の末端アミノ酸に選択的に結合する１つ以上の標識付認識分子とをポリペプチドの末端で接触させることを含む。いくつかの実施形態では、方法は、ポリペプチドと、１つ以上の標識付認識分子との相互作用を検出することによって、ポリペプチドの末端の末端アミノ酸を同定することをさらに含む。

【００２９】

さらに他の態様では、本願は、エドマン型分解反応によるポリペプチドシーケンスの方法を提供する。いくつかの実施形態では、エドマン型分解反応は、ポリペプチドと様々な反応混合物とを検出又は切断のいずれかの目的で接触させることによって実施されうる（たとえば、単一反応混合物を使用して検出及び切断を含みうる動的シーケンシング反応と比較して）。

【００３０】

それゆえ、いくつかの態様では、本願は、（ｉ）ポリペプチドと、ポリペプチドの末端の１タイプ以上の末端アミノ酸に選択的に結合する１つ以上の標識付認識分子とを接触させることを含む、ポリペプチドのアミノ酸配列を決定する方法を提供する。いくつかの実施形態では、方法は、（ｉｉ）ポリペプチドと１つ以上の標識付認識分子との相互作用を検出することによって、ポリペプチドの末端の末端アミノ酸（たとえば、Ｎ末端又はＣ末端のアミノ酸）を同定することをさらに含む。いくつかの実施形態では、方法は、（ｉｉｉ）末端アミノ酸を除去することをさらに含む。いくつかの実施形態では、方法は、（ｉｖ）ポリペプチドの末端で（ｉ）～（ｉｉｉ）を１回以上繰り返してポリペプチドのアミノ酸配列を決定することをさらに含む。

【００３１】

いくつかの実施形態では、方法は、（ｉ）の後且つ（ｉｉ）の前に、末端アミノ酸に選択的に結合しない１つ以上の標識付認識分子のいずれも除去することを含む。いくつかの実施形態では、方法は、（ｉｉ）の後且つ（ｉｉｉ）の前に、末端アミノ酸に選択的に結合する１つ以上の標識付認識分子のいずれも除去することを含む。

【００３２】

いくつかの実施形態では、末端アミノ酸を除去すること（たとえば、（ｉｉｉ））は、末端アミノ酸とイソチオシアネート（たとえば、フェニルイソチアシアネート）とを接触させることによって末端アミノ酸を修飾すること、及び修飾末端アミノ酸と修飾末端アミノ酸を特異的に結合及び除去するプロテアーゼとを接触させることを含む。いくつかの実施形態では、末端アミノ酸を切断すること（たとえば、（ｉｉｉ））は、末端アミノ酸とイソチオシアネートとを接触させることによって末端アミノ酸を修飾すること、及び修飾末端アミノ酸を、修飾末端アミノ酸を除去するのに十分な酸性条件又は塩基性条件にさらすことを含む。

【００３３】

いくつかの実施形態では、末端アミノ酸を同定することは、１つ以上の標識付認識分子が結合する１タイプ以上の末端アミノ酸のうちの１タイプとして末端アミノ酸を同定することを含む。いくつかの実施形態では、末端アミノ酸を同定することは、１つ以上の標識付認識分子が結合する１タイプ以上の末端アミノ酸以外のタイプとして末端アミノ酸を同定することを含む。

【００３４】

いくつかの態様では、本願は、混合サンプル中の対象タンパク質の同定方法を提供する。いくつかの実施形態では、本方法は、混合タンパク質サンプルを切断して複数のポリペプチド断片を生成することを含む。いくつかの実施形態では、本方法は、本願の方法に係る方法で複数のポリペプチド断片の少なくとも１つのアミノ酸配列を決定することをさらに含む。いくつかの実施形態では、本方法は、アミノ酸配列が対象タンパク質にユニークに同定可能である場合に混合サンプル中の対象タンパク質を同定することをさらに含む。

【００３５】

いくつかの実施形態では、混合サンプル中の対象タンパク質の同定方法は、混合タンパ

ク質サンプルを切断して複数のポリペプチド断片を生成することを含む。いくつかの実施形態では、本方法は、複数のポリペプチド断片中の１タイプ以上のアミノ酸を１つ以上の異なる発光標識で標識することをさらに含む。いくつかの実施形態では、本方法は、複数の標識付ポリペプチドの少なくとも１つについて経時的に発光を測定することをさらに含む。いくつかの実施形態では、本方法は、検出された発光に基づいて少なくとも１つの標識付ポリペプチドのアミノ酸配列を決定することをさらに含む。いくつかの実施形態では、本方法は、アミノ酸配列が対象タンパク質にユニークに同定可能である場合に混合サンプル中の対象タンパク質を同定することをさらに含む。

【００３６】

それゆえ、いくつかの実施形態では、本願に係る分析される対象ポリペプチド分子又はタンパク質は、混合又は精製サンプルでありうる。いくつかの実施形態では、対象ポリペプチド分子又はタンパク質は、生物学的サンプル（たとえば、血液、組織、唾液、尿、又は他の生物源）から得られる。いくつかの実施形態では、対象ポリペプチド分子又はタンパク質は、患者サンプル（たとえばヒトサンプル）から得られる。

【００３７】

いくつかの態様では、本願は、少なくとも１つのハードウェアプロセッサーと、少なくとも１つのハードウェアプロセッサーによって実行されるときに少なくとも１つのハードウェアプロセッサーに明細書に記載の方法を実施させるプロセッサー実行可能命令を記憶する少なくとも１つの非一時コンピュータ可読記憶媒体と、を含むシステムを提供する。いくつかの態様では、本願は、少なくとも１つのハードウェアプロセッサーによって実行されるときに少なくとも１つのハードウェアプロセッサーに明細書に記載の方法を実施させるプロセッサー実行可能命令を記憶する少なくとも１つの非一時コンピュータ可読記憶媒体を提供する。

【００３８】

本発明のある特定の実施形態の詳細は、以下に記載のある特定の実施形態の詳細な説明に示される。本発明の他の特徴、目的、及び利点は、実施例、図面、及び特許請求の範囲から明らかになるであろう。

【図面の簡単な説明】

【００３９】

本明細書の一部を構成する添付の図面は、本発明のいくつかの実施形態を示し、説明とともに、本発明の原理を説明するのに役立つ。

【図１Ａ】単一分子結合相互作用の検出（図１Ａ）及び解析（図１Ｂ）によるポリペプチドシーケンシングの例を示す。

【図１Ｂ】単一分子結合相互作用の検出（図１Ａ）及び解析（図１Ｂ）によるポリペプチドシーケンシングの例を示す。

【図２】１タイプ以上の末端アミノ酸に選択的に結合する標識付酵素及び標識付アプタマーをはじめとする標識付認識分子の構成例を描く。

【図３Ａ】シールディングエレメントを介して標識されたアミノ酸認識分子の非限定的例を示す。図３Ａは、従来の共有結合性連結を介して標識された認識分子による単一分子ペプチドシーケンシングを例示する。

【図３Ｂ】シールディングエレメントを介して標識されたアミノ酸認識分子の非限定的例を示す。図３Ｂは、シールディングエレメントを含む認識分子による単一分子ペプチドシーケンシングを例示する。

【図３Ｃ】シールディングエレメントを介して標識されたアミノ酸認識分子の非限定的例を示す。図３Ｃ～３Ｅは、本願に係るシールディングエレメントの各種例を例示する。

【図３Ｄ】シールディングエレメントを介して標識されたアミノ酸認識分子の非限定的例を示す。図３Ｃ～３Ｅは、本願に係るシールディングエレメントの各種例を例示する。

【図３Ｅ】シールディングエレメントを介して標識されたアミノ酸認識分子の非限定的例を示す。図３Ｃ～３Ｅは、本願に係るシールディングエレメントの各種例を例示する。

【図４】標識付認識分子を用いたポリペプチドシーケンシングの分解ベースプロセスを一

般的に描く。
【図５】末端及び／又は内部アミノ酸と標識付認識分子及び標識付切断試薬（たとえば標識付非特異的エキソペプチダーゼ）との結合相互作用を評価することによるリアルタイムのポリペプチドシーケンシングの例を示す。図５は、シグナル出力で一連のパルスを検出することによるリアルタイムシーケンシングの例を示す。
【図６】末端及び／又は内部アミノ酸と標識付認識分子及び標識付切断試薬（たとえば標識付非特異的エキソペプチダーゼ）との結合相互作用を評価することによるリアルタイムのポリペプチドシーケンシングの例を示す。図６は、温度依存シーケンシングプロセスを模式的に描く。
【図７】末端及び／又は内部アミノ酸と標識付認識分子及び標識付切断試薬（たとえば標識付非特異的エキソペプチダーゼ）との結合相互作用を評価することによるリアルタイムのポリペプチドシーケンシングの例を示す。図７は、末端及び内部アミノ酸と標識付認識分子及び標識付非特異的エキソペプチダーゼとの結合相互作用を評価することによるリアルタイムのポリペプチドシーケンシングの例を示す。
【図８】本願に係るポリペプチド及びタンパク質の分析のためのサンプル及びサンプルウェル表面の各種調製例を示す。図８は、タンパク質サンプルから末端修飾ポリペプチドを調製するプロセス例を一般的に描く。
【図９】本願に係るポリペプチド及びタンパク質の分析のためのサンプル及びサンプルウェル表面の各種調製例を示す。図９は、可溶化性リンカーをポリペプチドにコンジュゲートするプロセス例を一般的に描く。
【図１０】本願に係るポリペプチド及びタンパク質の分析のためのサンプル及びサンプルウェル表面の各種調製例を示す。図１０は、ボトム表面への単一分子固定を促進するために使用されうる修飾表面を有するサンプルウェルの模式図例を示す。
【図１１】本明細書に記載の技術のいくつかの実施形態に係るポリペプチド分解プロセス時に得られるデータを解析するための例示的配列データ処理パイプラインの図である。
【図１２】本明細書に記載の技術のいくつかの実施形態に係るポリペプチド分子のアミノ酸配列を決定するための例示的プロセスのフローチャートである。
【図１３】本明細書に記載の技術のいくつかの実施形態に係るポリペプチドに相当するアミノ酸配列を決定するための例示的プロセスのフローチャートである。
【図１４】本明細書に記載の技術のいくつかの実施形態の実現に使用されうる例示的コンピュータシステムのブロック図である。
【図１５Ａ】さまざまな可溶化性リンカーにより提供される溶解性増強のために調製及び評価された選ばれたペプチド－リンカーコンジュゲートの実験データを示す。図１５Ａは、合成及び評価されたペプチド－リンカーコンジュゲートの構造例を示す。
【図１５Ｂ】さまざまな可溶化性リンカーにより提供される溶解性増強のために調製及び評価された選ばれたペプチド－リンカーコンジュゲートの実験データを示す。図１５Ｂは、Ｎ末端のペプチド切断を実証するＬＣＭＳの結果を示す。
【図１５Ｃ－１】さまざまな可溶化性リンカーにより提供される溶解性増強のために調製及び評価された選ばれたペプチド－リンカーコンジュゲートの実験データを示す。図１５Ｃは、ローディング実験の結果を示す。
【図１５Ｃ－２】同上。
【図１６】実験結果に基づいて選ばれたエキソペプチダーゼのアミノ酸切断活性の概要を示す。
【図１７Ａ－１】末端アミノ酸の検出及び切断のための色素／ペプチドコンジュゲートアッセイの実験データを示す。図１７Ａは、色素／ペプチドコンジュゲートアッセイを実施するために使用されるスキーム例及び構造例を示す。
【図１７Ａ－２】同上。
【図１７Ｂ】末端アミノ酸の検出及び切断のための色素／ペプチドコンジュゲートアッセイの実験データを示す。図１７Ｂは、オンチップアッセイでのサンプルウェル中へのペプチド－リンカーコンジュゲートローディングのイメージング結果を示す。

【図１７Ｃ】末端アミノ酸の検出及び切断のための色素／ペプチドコンジュゲートアッセイの実験データを示す。図１７Ｃは、ペプチドコンジュゲートローディング及び末端アミノ酸切断を検出したシグナルトレース例を示す。
【図１８Ａ－１】末端アミノ酸の検出及び切断のためのＦＲＥＴ色素／ペプチドコンジュゲートアッセイの実験データを示す。図１８Ａは、ＦＲＥＴ色素／ペプチドコンジュゲートアッセイを実施するために使用されるスキーム例及び構造例を示す。
【図１８Ａ－２】同上。
【図１８Ｂ】末端アミノ酸の検出及び切断のためのＦＲＥＴ色素／ペプチドコンジュゲートアッセイの実験データを示す。図１８Ｂは、さまざまな時点のＦＲＥＴイメージング結果を示す。
【図１８Ｃ】末端アミノ酸の検出及び切断のためのＦＲＥＴ色素／ペプチドコンジュゲートアッセイの実験データを示す。図１８Ｃは、さまざまな時点のカッティング効率を示す。
【図１８Ｄ】末端アミノ酸の検出及び切断のためのＦＲＥＴ色素／ペプチドコンジュゲートアッセイの実験データを示す。図１８Ｄは、さまざまな時点の各々で表示されるカッティングを示す。
【図１８Ｅ】末端アミノ酸の検出及び切断のためのＦＲＥＴ色素／ペプチドコンジュゲートアッセイの実験データを示す。図１８Ｅは、エルシニア・ペスティス（Ｙｅｒｓｉｎｉａ　ｐｅｓｔｉｓ）由来のプロリンイミノペプチダーゼ（ｙＰＩＰ）によるさまざまな時点の追加のＦＲＥＴイメージング結果を示す。
【図１８Ｆ】末端アミノ酸の検出及び切断のためのＦＲＥＴ色素／ペプチドコンジュゲートアッセイの実験データを示す。図１８Ｆは、ビブリオ・プロテオリチカス（Ｖｉｂｒｉｏ　ｐｒｏｔｅｏｌｙｔｉｃｕｓ）由来のアミノペプチダーゼ（ＶＰｒ）によるさまざまな時点のＦＲＥＴイメージング結果を示す。
【図１９Ａ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ａは、これらの実験のために標識されたＣｌｐＳ２タンパク質の結晶構造を示す。
【図１９Ｂ－１】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｂは、標識付ＣｌｐＳ２タンパク質によるＮ末端アミノ酸識別を例示する単一分子強度トレースを示す。
【図１９Ｂ－２】同上。
【図１９Ｃ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｃは、さまざまな末端アミノ酸の平均パルス持続時間を示すプロットである。
【図１９Ｄ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｄは、さまざまな末端アミノ酸の平均パルス間持続時間を示すプロットである。
【図１９Ｅ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｅは、さまざまな末端アミノ酸間のディスクリミナントパルス持続時間をさらに例示するプロットを示す。
【図１９Ｆ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｆ、１９Ｇ、及び１９Ｈは、サーモシノコッカス・エロンガタス（Ｔｈｅｒｍｏｓｙｎｏｃｈｏｃｃｕｓ　ｅｌｏｎｇａｔｕｓ）由来のＣｌｐＳタンパク質（ｔｅＣｌｐＳ）によるロイシン認識を実証する滞留時間解析の結果例を示す。
【図１９Ｇ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｆ、１９Ｇ、及び１９Ｈは、サーモシノコッカス・エロンガタス（Ｔｈｅｒｍｏｓｙｎｏｃｈｏｃｃｕｓ　ｅｌｏｎｇａｔｕｓ）由来のＣｌｐＳタンパク質（ｔｅＣｌｐＳ）によるロイシン認識を実証する滞留時間解析の結果例を示す。
【図１９Ｈ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｆ、１９Ｇ、及び１９Ｈは、サーモシノコッカス・エロンガタス（Ｔｈｅｒｍｏｓｙｎｏｃｈｏｃｃｕｓ　ｅｌｏｎｇａｔｕｓ）由来のＣｌｐＳタンパク質（ｔｅＣｌｐＳ）によるロイシン認識を実証する滞留時間解析の結果例を示す。
【図１９Ｉ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｉは、

Ａ．ツメファシエンス（Ａ．ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ１によるフェニルアラニン、ロイシン、トリプトファン、及びチロシンの区別可能認識を実証する滞留時間解析の結果例を示す。
【図１９Ｊ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｊは、Ｓ．エロンガタス（ｅｌｏｎｇａｔｕｓ）ＣｌｐＳ２によるロイシン認識を実証する滞留時間解析の結果例を示す。
【図１９Ｋ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｋ～１９Ｌは、ＧＩＤ４によるプロリン認識を実証する滞留時間解析の結果例を示す。
【図１９Ｌ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｋ～１９Ｌは、ＧＩＤ４によるプロリン認識を実証する滞留時間解析の結果例を示す。
【図１９Ｍ】標識付認識分子による末端アミノ酸識別の実験データを示す。図１９Ｍは、さまざまなＮ末端アミノ酸を有するペプチドを有するａｔＣｌｐＳ２－Ｖ１の模範的な結合曲線を示す。
【図２０Ａ】同一反応混合物中の標識付ＣｌｐＳ２認識タンパク質及びアミノペプチダーゼ切断試薬を用いてリアルタイムで行われたポリペプチドシーケンシング反応の結果例を示す。図２０Ａは、第１のシーケンシング反応のシグナルトレースデータを示す。
【図２０Ｂ】同一反応混合物中の標識付ＣｌｐＳ２認識タンパク質及びアミノペプチダーゼ切断試薬を用いてリアルタイムで行われたポリペプチドシーケンシング反応の結果例を示す。図２０Ｂは、図２０Ａに示されるシグナルトレースデータのパルス持続時間統計を示す。
【図２０Ｃ】同一反応混合物中の標識付ＣｌｐＳ２認識タンパク質及びアミノペプチダーゼ切断試薬を用いてリアルタイムで行われたポリペプチドシーケンシング反応の結果例を示す。図２０Ｃは、第２のシーケンシング反応のシグナルトレースデータを示す。
【図２０Ｄ】同一反応混合物中の標識付ＣｌｐＳ２認識タンパク質及びアミノペプチダーゼ切断試薬を用いてリアルタイムで行われたポリペプチドシーケンシング反応の結果例を示す。図２０Ｄは、図２０Ｃに示されるシグナルトレースデータのパルス持続時間統計を示す。
【図２１Ａ】標識付エキソペプチダーゼによる末端アミノ酸同定及び切断の実験データを示す。図２１Ａは、これらの実験のために部位特異的に標識されたプロリンイミノペプチダーゼ（ｙＰＩＰ）の結晶構造を示す。
【図２１Ｂ】標識付エキソペプチダーゼによる末端アミノ酸同定及び切断の実験データを示す。図２１Ｂは、精製タンパク質産物の標識度を示す。
【図２１Ｃ】標識付エキソペプチダーゼによる末端アミノ酸同定及び切断の実験データを示す。図２１Ｃは、ｙＰＩＰの部位特異的標識化を確認するＳＤＳ　ｐａｇｅの画像である。
【図２１Ｄ】標識付エキソペプチダーゼによる末端アミノ酸同定及び切断の実験データを示す。図２１Ｄは、部位特異的標識化を確認するＳＤＳ　ｐａｇｅゲルのオーバー露出画像である。
【図２１Ｅ】標識付エキソペプチダーゼによる末端アミノ酸同定及び切断の実験データを示す。図２１Ｅは、標識付タンパク質産物の純度を確認するクーマシー染色ゲルの画像である。
【図２１Ｆ】標識付エキソペプチダーゼによる末端アミノ酸同定及び切断の実験データを示す。図２１Ｆは、標識付エキソペプチダーゼの切断活性を実証するＨＰＬＣトレースである。
【図２２Ａ】特異的翻訳後修飾を含有するアミノ酸の認識を評価する実験のデータを示す。図２２Ａは、ＳＨ２ドメイン含有タンパク質によるホスホチロシン認識を実証した代表的トレースを示す。
【図２２Ｂ】特異的翻訳後修飾を含有するアミノ酸の認識を評価する実験のデータを示す。図２２Ｂは、図２２Ａのトレースに対応するパルス持続時間データを示し、且つ図２２Ｃは、トレースに対して決定された統計を示す。

【図２２Ｃ】特異的翻訳後修飾を含有するアミノ酸の認識を評価する実験のデータを示す。図２２Ｂは、図２２Ａのトレースに対応するパルス持続時間データを示し、且つ図２２Ｃは、トレースに対して決定された統計を示す。
【図２２Ｄ】特異的翻訳後修飾を含有するアミノ酸の認識を評価する実験のデータを示す。図２２Ｄ～２２Ｆは、陰性対照実験の代表的トレースを示す。
【図２２Ｅ】特異的翻訳後修飾を含有するアミノ酸の認識を評価する実験のデータを示す。図２２Ｄ～２２Ｆは、陰性対照実験の代表的トレースを示す。
【図２２Ｆ】特異的翻訳後修飾を含有するアミノ酸の認識を評価する実験のデータを示す。図２２Ｄ～２２Ｆは、陰性対照実験の代表的トレースを示す。
【図２３】パルス持続時間に及ぼすペナルティメートアミノ酸の影響を評価する実験のパルス持続時間の中央値を示すプロットである。
【図２４Ａ】差次的標識付認識分子による同時アミノ酸認識を評価する実験のデータを示す。図２４Ａは、代表的トレースを示す。
【図２４Ｂ】差次的標識付認識分子による同時アミノ酸認識を評価する実験のデータを示す。図２４Ｂは、各認識分子に対してこれらの実験時に得られたパルス持続時間データを比較するプロットである。
【図２４Ｃ】差次的標識付認識分子による同時アミノ酸認識を評価する実験のデータを示す。図２４Ｃは、これらの実験のパルス持続時間統計を示す。
【図２５Ａ】単一分子認識時のペプチドの光安定性を評価する実験のデータを示す。図２５Ａは、アミノ酸結合部位から約２ｎｍに色素で標識されたａｔＣｌｐＳ２－Ｖ１を用いた認識の代表的トレースを示す。
【図２５Ｂ】単一分子認識時のペプチドの光安定性を評価する実験のデータを示す。図２５Ｂは、これらの実験に使用されるＣｌｐＳ２タンパク質の構造の可視化を示す。
【図２５Ｃ】単一分子認識時のペプチドの光安定性を評価する実験のデータを示す。図２５Ｃは、ＤＮＡ／タンパク質リンカーを介してアミノ酸結合部位から＞１０ｎｍに色素で標識されたＣｌｐＳ２を用いた認識の代表的トレースを示す。
【図２６Ａ－１】アミノペプチダーゼ切断試薬の存在下でＤＮＡ／ストレプトアビジンリンカーを介して標識されたＣｌｐＳ２認識タンパク質を用いて相補型金属－酸化物－半導体（ＣＭＯＳ）チップ上でリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースを示す。
【図２６Ａ－２】同上。
【図２６Ｂ－１】アミノペプチダーゼ切断試薬の存在下でＤＮＡ／ストレプトアビジンリンカーを介して標識されたＣｌｐＳ２認識タンパク質を用いて相補型金属－酸化物－半導体（ＣＭＯＳ）チップ上でリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースを示す。
【図２６Ｂ－２】同上。
【図２６Ｃ－１】アミノペプチダーゼ切断試薬の存在下でＤＮＡ／ストレプトアビジンリンカーを介して標識されたＣｌｐＳ２認識タンパク質を用いて相補型金属－酸化物－半導体（ＣＭＯＳ）チップ上でリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースを示す。
【図２６Ｃ－２】同上。
【図２６Ｄ－１】アミノペプチダーゼ切断試薬の存在下でＤＮＡ／ストレプトアビジンリンカーを介して標識されたＣｌｐＳ２認識タンパク質を用いて相補型金属－酸化物－半導体（ＣＭＯＳ）チップ上でリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースを示す。
【図２６Ｄ－２】同上。
【図２７】パイロコッカス・ホリコシイ（Ｐｙｒｏｃｏｃｃｕｓ　ｈｏｒｉｋｏｓｈｉｉ）ＴＥＴアミノペプチダーゼ切断試薬の存在下でＤＮＡ／ストレプトアビジンリンカーを介して標識されたａｔＣｌｐＳ２－Ｖ１認識タンパク質を用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースを示す。

【図２８Ａ－１】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ａは、図２８Ｂに示される拡大パルスパターン領域と共にｈＴＥＴエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２８Ａ－２】同上。
【図２８Ｂ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ａは、図２８Ｂに示される拡大パルスパターン領域と共にｈＴＥＴエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２８Ｃ－１】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｃは、図２８Ｄに示される拡大パルスパターン領域及び図２８Ｅに示される追加の代表的トレースと共にｈＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２８Ｃ－２】同上。
【図２８Ｄ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｃは、図２８Ｄに示される拡大パルスパターン領域及び図２８Ｅに示される追加の代表的トレースと共にｈＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２８Ｅ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｃは、図２８Ｄに示される拡大パルスパターン領域及び図２８Ｅに示される追加の代表的トレースと共にｈＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２８Ｆ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｆは、図２８Ｇに示される拡大パルスパターン領域及び図２８Ｈに示される追加の代表的トレースと共にｈＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施されたさらなる反応の代表的トレースを示す。
【図２８Ｇ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｆは、図２８Ｇに示される拡大パルスパターン領域及び図２８Ｈに示される追加の代表的トレースと共にｈＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施されたさらなる反応の代表的トレースを示す。
【図２８Ｈ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｆは、図２８Ｇに示される拡大パルスパターン領域及び図２８Ｈに示される追加の代表的トレースと共にｈＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施されたさらなる反応の代表的トレースを示す。
【図２８Ｉ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｉは、図２８Ｊに示される拡大パルスパターン領域と共にＰｆｕＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２８Ｊ】差次的切断特異性を有する複数のタイプのエキソペプチダーゼを用いてリアルタイムで行われたポリペプチドシーケンシング反応の代表的トレースデータを示す。図２８Ｉは、図２８Ｊに示される拡大パルスパターン領域と共にＰｆｕＴＥＴ及びｙＰＩＰの両方のエキソペプチダーゼで実施された反応の代表的トレースを示す。
【図２９Ａ】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ａは、精製ＣｌｐＳタンパク質のＳＤＳ－ＰＡＧＥゲルイメージングを示す。

【図２９Ｂ－１】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ｂは、ＣｌｐＳホモログのバイオレイヤー干渉スクリーニングの結果を示す。
【図２９Ｂ－２】同上。
【図２９Ｃ】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ｃは、スクリーニングの選択結果を示す。
【図２９Ｄ】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ｄは、ＣｌｐＳタンパク質（ＰＳ３７２）と、ＬＡ、ＩＡ、及びＶＡペプチドとの反応曲線を示す。
【図２９Ｅ】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ｅは、ＰＳ３７２及び４つの他のホモログ、並びに無タンパク質対照の偏光反応を示す。
【図２９Ｆ】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ｆは、ＩＡ、ＩＲ、ＩＱ、ＶＡ、及びＶＲペプチドのバイオレイヤー干渉反応曲線を示す。
【図２９Ｇ】新たに同定されたＣｌｐＳホモログを評価する実験のデータを示す。図２９Ｇは、ＩＲペプチド（トップパネル）及びＬＦペプチド（ボトムパネル）を有するＰＳ３７２のパルス幅ヒストグラム及び代表的トレースを示す。
【図３０Ａ】新たに同定されたＣｌｐＳホモログによる末端アミノ酸識別を評価する実験のデータを示す。図３０Ａは、ＳＤＳ－ＰＡＧＥによるＰＳ３７２のｉｎ　ｖｉｖｏビオチン化を示す。
【図３０Ｂ】新たに同定されたＣｌｐＳホモログによる末端アミノ酸識別を評価する実験のデータを示す。図３０Ｂは、ＳＶ－色素にコンジュゲートされたＰＳ３７２の精製プロファイルを示す。
【図３０Ｃ】新たに同定されたＣｌｐＳホモログによる末端アミノ酸識別を評価する実験のデータを示す。図３０Ｃは、ＳＶ－色素コンジュゲートされたＰＳ３７２の精製後のＳＤＳ－ＰＡＧＥを示す。
【図３０Ｄ】新たに同定されたＣｌｐＳホモログによる末端アミノ酸識別を評価する実験のデータを示す。図３０Ｄは、ＩからＬ（ａ）及びＬからＩ（ｂ）への遷移を示す代表的トレースを示す。
【図３０Ｅ】新たに同定されたＣｌｐＳホモログによる末端アミノ酸識別を評価する実験のデータを示す。図３０Ｅは、色素標識付ＰＳ３２７によるリアルタイム動的ペプチドシーケンシングアッセイの例示データを示す。
【図３１Ａ】メチオニン結合ＣｌｐＳタンパク質の工学のデータを示す。図３１Ａは、蛍光活性化セルソーティング（ＦＡＣＳ）を介して実施された選択の結果を示す。
【図３１Ｂ】メチオニン結合ＣｌｐＳタンパク質の工学のデータを示す。図３１Ｂは、Ｎ末端ＬＡを有するペプチドを有するメチオニン結合ＣｌｐＳタンパク質の例示的な反応曲線を示す。
【図３１Ｃ】メチオニン結合ＣｌｐＳタンパク質の工学のデータを示す。図３１Ｃは、Ｎ末端ＭＡを有するペプチドを有するメチオニン結合ＣｌｐＳタンパク質の例示的な反応曲線を示す。
【図３１Ｄ】メチオニン結合ＣｌｐＳタンパク質の工学のデータを示す。図３１Ｄは、Ｎ末端ＭＲを有するペプチドを有するメチオニン結合ＣｌｐＳタンパク質の例示的な反応曲線を示す。
【図３１Ｅ】メチオニン結合ＣｌｐＳタンパク質の工学のデータを示す。図３１Ｅは、Ｎ末端ＦＡを有するペプチドを有するメチオニン結合ＣｌｐＳタンパク質の例示的な反応曲線を示す。
【図３１Ｆ】メチオニン結合ＣｌｐＳタンパク質の工学のデータを示す。図３１Ｆは、Ｎ末端ＭＱを有するペプチドを有するメチオニン結合ＣｌｐＳタンパク質の例示的な反応曲線を示す。
【図３２Ａ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ａ～３２Ｂは、ペナルティメート位置のアミノ酸の前にＮ末端Ｒを含有する１４のペプチダ

ーゼと結合するＵＢＲボックスホモログＰＳ５３５（図３２Ａ）及びＰＳ５２２（図３２Ｂ）の例示的な結合曲線を示す。
【図３２Ｂ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ａ～３２Ｂは、ペナルティメート位置のアミノ酸の前にＮ末端Ｒを含有する１４のペプチダーゼと結合するＵＢＲボックスホモログＰＳ５３５（図３２Ａ）及びＰＳ５２２（図３２Ｂ）の例示的な結合曲線を示す。
【図３２Ｃ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｃは、Ｎ末端Ｒペプチドに結合する２４のＵＢＲボックスホモログを測定した結果を示すヒートマップである。
【図３２Ｄ－１】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｄは、Ｒ、Ｋ、又はＨをＮ末端位置に含有するポリペプチドと結合する拡大したセットのＵＢＲボックスホモログを測定した結果を示すヒートマップである。
【図３２Ｄ－２】同上。
【図３２Ｅ－１】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｅは、ペナルティメート位置の様々なアミノ酸の前にＮ末端Ｒを含有する１４のポリペプチドと結合する拡大したセットのＵＢＲボックスホモログを測定した結果を示すヒートマップである。
【図３２Ｅ－２】同上。
【図３２Ｆ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｆは、単一点蛍光偏光アッセイの結果を示す。
【図３２Ｇ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｇは、結合親和性決定の偏光結果の解析を示す。
【図３２Ｈ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｈは、認識アッセイにおけるＰＳ６２１の代表的トレースを示す。
【図３２Ｉ】ＵＢＲボックスドメインホモログを評価する実験のデータを示す。図３２Ｉは、３バインダー動的シーケンシング反応の例示的なシーケンシングトレースを示す。
【図３３Ａ】Ｐ３７２ホモログタンパク質を評価する実験のデータを示す。図３３Ａ～３３Ｅは、さまざまなＮ末端アミノ酸（Ｉ、Ｖ、Ｌ、Ｆ）を含有する４つのポリペプチドと結合するＰＳ３７２（図３３Ａ）及びホモログＰＳ５４５（図３３Ｂ）、ＰＳ５５１（図３３Ｃ）、ＰＳ５５７（図３３Ｄ）、並びにＰＳ５５８（図３３Ｅ）の例示的な結合曲線を示す。
【図３３Ｂ】Ｐ３７２ホモログタンパク質を評価する実験のデータを示す。図３３Ａ～３３Ｅは、さまざまなＮ末端アミノ酸（Ｉ、Ｖ、Ｌ、Ｆ）を含有する４つのポリペプチドと結合するＰＳ３７２（図３３Ａ）及びホモログＰＳ５４５（図３３Ｂ）、ＰＳ５５１（図３３Ｃ）、ＰＳ５５７（図３３Ｄ）、並びにＰＳ５５８（図３３Ｅ）の例示的な結合曲線を示す。
【図３３Ｃ】Ｐ３７２ホモログタンパク質を評価する実験のデータを示す。図３３Ａ～３３Ｅは、さまざまなＮ末端アミノ酸（Ｉ、Ｖ、Ｌ、Ｆ）を含有する４つのポリペプチドと結合するＰＳ３７２（図３３Ａ）及びホモログＰＳ５４５（図３３Ｂ）、ＰＳ５５１（図３３Ｃ）、ＰＳ５５７（図３３Ｄ）、並びにＰＳ５５８（図３３Ｅ）の例示的な結合曲線を示す。
【図３３Ｄ】Ｐ３７２ホモログタンパク質を評価する実験のデータを示す。図３３Ａ～３３Ｅは、さまざまなＮ末端アミノ酸（Ｉ、Ｖ、Ｌ、Ｆ）を含有する４つのポリペプチドと結合するＰＳ３７２（図３３Ａ）及びホモログＰＳ５４５（図３３Ｂ）、ＰＳ５５１（図３３Ｃ）、ＰＳ５５７（図３３Ｄ）、並びにＰＳ５５８（図３３Ｅ）の例示的な結合曲線を示す。
【図３３Ｅ】Ｐ３７２ホモログタンパク質を評価する実験のデータを示す。図３３Ａ～３３Ｅは、さまざまなＮ末端アミノ酸（Ｉ、Ｖ、Ｌ、Ｆ）を含有する４つのポリペプチドと結合するＰＳ３７２（図３３Ａ）及びホモログＰＳ５４５（図３３Ｂ）、ＰＳ５５１（図３３Ｃ）、ＰＳ５５７（図３３Ｄ）、並びにＰＳ５５８（図３３Ｅ）の例示的な結合曲線

を示す。
【図３３Ｆ】ＰＳ３７２ホモログタンパク質を評価する実験のデータを示す。図３３Ｆは、３４のＰＳ３７２ホモログを測定した結果を示すヒートマップである。
【図３４Ａ】ａｔＣｌｐＳ２－Ｖ１の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダー（ＰＳ６１０）を評価する実験のデータを示す。図３４Ａは、ペプチドオンチップ認識アッセイの代表的トレースデータを示す。
【図３４Ｂ】ａｔＣｌｐＳ２－Ｖ１の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダー（ＰＳ６１０）を評価する実験のデータを示す。図３４Ｂは、バインダー濃度の関数として平均パルス速度を示すプロットである。
【図３４Ｃ】ａｔＣｌｐＳ２－Ｖ１の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダー（ＰＳ６１０）を評価する実験のデータを示す。図３４Ｃは、色素標識付Ｐ６１０及びＰＳ３２７によるリアルタイム動的ペプチドシーケンシングアッセイの例示データを示す。
【図３４Ｄ】ａｔＣｌｐＳ２－Ｖ１の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダー（ＰＳ６１０）を評価する実験のデータを示す。図３４Ｄは、バインダーオンチップアッセイの代表的トレースデータを示す。
【図３５Ａ】さまざまなリンカーを含有する直列ＣｌｐＳ２構築物を評価する実験のデータを示す。図３５Ａは、一価バインダーａｔＣｌｐＳ２－Ｖ１の例示的な結合曲線を示す。
【図３５Ｂ】さまざまなリンカーを含有する直列ＣｌｐＳ２構築物を評価する実験のデータを示す。図３５Ｂ～３５Ｄは、リンカー１（図３５Ｂ）、リンカー２（図３５Ｃ）、又はリンカー３（図３５Ｄ）によって分離されるａｔＣｌｐＳ２－Ｖ１の２つのコピーを有する直列構築物の例示的な結合曲線を示す。
【図３５Ｃ】さまざまなリンカーを含有する直列ＣｌｐＳ２構築物を評価する実験のデータを示す。図３５Ｂ～３５Ｄは、リンカー１（図３５Ｂ）、リンカー２（図３５Ｃ）、又はリンカー３（図３５Ｄ）によって分離されるａｔＣｌｐＳ２－Ｖ１の２つのコピーを有する直列構築物の例示的な結合曲線を示す。
【図３５Ｄ】さまざまなリンカーを含有する直列ＣｌｐＳ２構築物を評価する実験のデータを示す。図３５Ｂ～３５Ｄは、リンカー１（図３５Ｂ）、リンカー２（図３５Ｃ）、又はリンカー３（図３５Ｄ）によって分離されるａｔＣｌｐＳ２－Ｖ１の２つのコピーを有する直列構築物の例示的な結合曲線を示す。
【図３６Ａ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ａは、一価バインダーａｔＣｌｐＳ２－Ｖ１（左側プロット）及び一価バインダーＰＳ３７２（右側プロット）の例示的な結合曲線を示す。
【図３６Ｂ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｂは、ａｔＣｌｐＳ２－Ｖ１及びＰＳ３７２の直列コピーを有する多価ポリペプチドの例示的な結合曲線を示す。
【図３６Ｃ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｃは、一価バインダーＰＳ３７２の例示的な結合曲線を示す。
【図３６Ｄ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｄは、Ｓ３７２の２つの直列コピーを有する多価ポリペプチドの例示的な結合曲線を示す。
【図３６Ｅ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｅは、一価バインダーＰＳ５５７の例示的な結合曲

線を示す。
【図３６Ｆ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｆ～３６Ｈは、リンカー１（図３６Ｆ）、リンカー２（図３６Ｇ）、又はリンカー３（図３６Ｈ）によって分離されるＰＳ５５７の２つのコピーを有する直列構築物の例示的な結合曲線を示す。
【図３６Ｇ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｆ～３６Ｈは、リンカー１（図３６Ｆ）、リンカー２（図３６Ｇ）、又はリンカー３（図３６Ｈ）によって分離されるＰＳ５５７の２つのコピーを有する直列構築物の例示的な結合曲線を示す。
【図３６Ｈ】同じＣｌｐＳタンパク質又は互いに異なるＣｌｐＳタンパク質の直列コピーを有する単一ポリペプチドとして生成された、工学操作された多価アミノ酸バインダーを評価する実験のデータを示す。図３６Ｆ～３６Ｈは、リンカー１（図３６Ｆ）、リンカー２（図３６Ｇ）、又はリンカー３（図３６Ｈ）によって分離されるＰＳ５５７の２つのコピーを有する直列構築物の例示的な結合曲線を示す。
【図３７Ａ】タンパク質シールドのＣ末端付加によって駆動されるバインダー及び融合タンパク質のｋｏｎ速度定数及びｋｏｆｆ速度決定のストップトフロー迅速速度論解析のデータを示す。図３７Ａは、アッセイ設計を例示する模式図（トップパネル）並びに会合速度定数（ｋｏｎ）を決定するための実験結果及び解析を示すプロット（ミドル及びボトムのパネル）を示す。
【図３７Ｂ】タンパク質シールドのＣ末端付加によって駆動されるバインダー及び融合タンパク質のｋｏｎ速度定数及びｋｏｆｆ速度決定のストップトフロー迅速速度論解析のデータを示す。図３７Ｂは、アッセイ設計を例示する模式図（トップパネル）並びに解離速度（ｋｏｆｆ）を測定するための実験結果及び解析を示すプロット（ボトムパネル）を示す。
【発明を実施するための形態】
【００４０】
本願の態様は、タンパク質シーケンシング及び同定の方法と、ポリペプチドシーケンシング及び同定の方法と、アミノ酸同定の方法と、かかる方法を実施するための組成物と、に関する。
【００４１】
いくつかの態様では、本願は、デバイス修正をほとんど又はまったく伴うことなく既存の分析機器を用いて実現されうるポリペプチドシーケンシング技術の発見に関する。たとえば、以前のポリペプチドシーケンシングストラテジーは、分析されるポリペプチドを含有する反応ベッセル（ｖｅｓｓｅｌ）を介するさまざまな試薬混合物の繰返しサイクリングを含んでいた。かかるストラテジーは、試薬サイクリングが可能なフローセルや類似の装置を備えていないこともある核酸シーケンシング機器などの既存の分析機器の修正を必要とすることもある。本願のある特定のポリペプチドシーケンシング技術は繰返し試薬サイクリングを必要としないので、機器サイズを増加させるおそれのある有意な修正を伴うことなく既存の機器の使用を可能にすることを、本発明者らは認識し高く評価した。それゆえ、いくつかの態様では、本願は、より小型のシーケンシング機器の使用を可能にするポリペプチドシーケンシング方法を提供する。いくつかの態様では、本願は、同一シーケンシング機器を用いて実施されるゲノミック解析及びプロテオミック解析の両方を可能にするポリペプチドシーケンシング技術の発見に関する。
【００４２】
差次的結合相互作用（ｄｉｆｆｅｒｅｎｔｉａｌ　ｂｉｎｄｉｎｇ　ｉｎｔｅｒａｃｔｉｏｎｓ）はポリペプチドシーケンシングの従来の標識化ストラテジーに対する追加又は代替のアプローチを提供可能であることを、本発明者らはさらに認識し高く評価した。従来のポリペプチドシーケンシングは、各タイプのアミノ酸をユニーク同定可能標識で標識

することを含みうる。数多くの翻訳後変異に加えて少なくとも２０の異なるタイプの天然に存在するアミノ酸があるので、こうしたプロセスは労力がかかるうえに誤りを起こしやすい可能性がある。いくつかの態様では、本願は、異なるタイプのアミノ酸に差次的に会合してポリペプチドのアミノ酸配列の指標となる検出可能特性シグネチャーを生成するアミノ酸認識分子の使用を含む技術の発見に関する。それゆえ、本願の態様は、ある特定の従来のポリペプチドシーケンシングに使用されるポリペプチド標識化及び／又は強烈な化学試薬を必要としない技術を提供するので、サンプルから得られる配列情報のスループット及び／又は確度を増加させる。

【００４３】

いくつかの態様では、本願は、単一反応混合物のみを用いて（たとえば、反応ベッセルを介する繰返し試薬サイクリングを必要とすることなく）ポリペプチドシーケンシング反応をリアルタイムでモニター可能であるという発見に関する。以上に詳述したように、従来のポリペプチドシーケンシング反応は、アミノ酸検出工程とアミノ酸切断工程との間でサイクルするようにポリペプチドをさまざまな試薬混合物に暴露することを含みうる。それゆえ、いくつかの態様では、本願は、進行中の分解反応全体にわたるリアルタイムのアミノ酸検出によりポリペプチド分析を可能にする次世代シーケンシングの進歩に関する。動的シーケンシングによるかかるポリペプチド分析のアプローチは以下に説明される。

【００４４】

本明細書に記載されるように、いくつかの態様では、本願は、ポリペプチド分解プロセス時のデータを得ることと、データを解析して分解プロセス時にポリペプチドの末端に逐次露出されるアミノ酸に対応するデータの部分を決定することと、によるポリペプチドのシーケンシング方法を提供する。いくつかの実施形態では、データの部分は、１つ以上のアミノ酸認識分子とポリペプチドの末端に露出される（たとえば分解時に）逐次アミノ酸との会合の指標となる一連のシグナルパルスを含む。いくつかの実施形態では、一連のシグナルパルスは、分解プロセス時のポリペプチドの末端での一連の可逆単一分子結合相互作用に対応する。

【００４５】

ポリペプチド分解プロセス時の単一分子結合相互作用を検出することによるポリペプチドシーケンシングの非限定的例は、図１Ａに模式的に例示される。シグナルトレース例（Ｉ）は、シグナルの変化に対応するときのさまざまな会合イベント（ａｓｓｏｃｉａｔｉｏｎ　ｅｖｅｎｔｓ）を描く一連のパネル（ＩＩ）で示される。示されるように、アミノ酸認識分子（斑点形状）とポリペプチド（ビーズオンストリング（ｂｅａｄｓ－ｏｎ－ａ－ｓｔｒｉｎｇ）として示される）の末端のアミノ酸との間の会合イベントは、ある持続時間にわたり持続するシグナルの大きさの変化を生成する。

【００４６】

パネル（Ａ）及び（Ｂ）は、アミノ酸認識分子とポリペプチドの末端に露出される第１のアミノ酸（たとえば第１の末端アミノ酸）との間のさまざまな会合イベントを描く。各会合イベントは、会合イベントの持続時間にわたり持続するシグナルの大きさの変化により特徴付けられるシグナルトレース（Ｉ）の変化を生成する。それゆえ、パネル（Ａ）及び（Ｂ）の会合イベント間の持続時間は、ポリペプチドがアミノ酸認識分子に検出可能に会合されない持続時間に対応する。

【００４７】

パネル（Ｃ）及び（Ｄ）は、アミノ酸認識分子とポリペプチドの末端に露出される第２のアミノ酸（たとえば第２の末端アミノ酸）との間のさまざまな会合イベントを描く。本明細書に記載されるように、ポリペプチドの末端に「露出」されるアミノ酸は、ポリペプチドに依然として装着されており且つ分解時に先行末端アミノ酸の除去により末端アミノ酸になる（たとえば、単独で又は１つ以上の追加のアミノ酸と共に、のいずれかで）アミノ酸である。それゆえ、一連のパネル（ＩＩ）の第１及び第２のアミノ酸は、ポリペプチドの末端に露出される逐次アミノ酸の例示的例を提供し、この場合、第２のアミノ酸は、第１のアミノ酸の除去により末端アミノ酸になったものである。

【００４８】
一般的に描かれるように、パネル（Ｃ）及び（Ｄ）の会合イベントは、パネル（Ａ）及び（Ｂ）のものよりも相対的に短い持続時間にわたり持続する大きさの変化により特徴付けられるシグナルトレース（Ｉ）の変化を生成し、且つパネル（Ｃ）及び（Ｄ）の会合イベント間の持続時間は、パネル（Ａ）及び（Ｂ）のものよりも相対的に短い。本明細書に記載されるように、いくつかの実施形態では、シグナルのこれらの識別可能な変化の一方又は両方のどちらかは、異なるタイプのアミノ酸間の識別が可能なシグナルトレース（Ｉ）の特性パターンを決定するために使用されうる。いくつかの実施形態では、一方の特性パターンから他方の特性パターンへの遷移は、アミノ酸切断の指標となる。本明細書で用いられる場合、いくつかの実施形態では、アミノ酸切断とは、ポリペプチドの末端からの少なくとも１つのアミノ酸の除去（たとえば、ポリペプチドからの少なくとも１つの末端アミノ酸の除去）を意味する。いくつかの実施形態では、アミノ酸切断は、特性パターン間の持続時間に基づく推測により決定される。いくつかの実施形態では、アミノ酸切断は、標識付切断試薬とポリペプチドの末端のアミノ酸との会合により生成されるシグナルの変化を検出することにより決定される。アミノ酸は分解時にポリペプチドの末端から逐次切断されるので、大きさの一連の変化又は一連のシグナルパルスが検出される。いくつかの実施形態では、シグナルパルスデータは、図１Ｂに例示されるように分析可能である。

【００４９】
いくつかの実施形態では、シグナルデータは、シグナルデータの１つ以上のパラメータに閾値レベルを適用することによりシグナルパルス情報を抽出するように解析可能である。たとえば、パネル（ＩＩＩ）は、シグナルトレース例（Ｉ）のシグナルデータに適用される閾値大きさレベル（「M_L」）を描く。いくつかの実施形態では、M_Lは、ある時点で検出されるシグナルと所与のデータセットに対して決定されるベースラインとの間の最小差である。いくつかの実施形態では、シグナルパルス（「ｓｐ」）は、M_Lを超える大きさの変化の指標となり且つある持続時間にわたり持続するデータの各部分に割り当てられる。いくつかの実施形態では、閾値持続時間は、M_Lを満たすデータの部分に適用されてシグナルパルスがその部分に割り当てられるかを決定しうる。たとえば、実験アーチファクトは、所望の信頼性でシグナルパルスを帰属するのに十分な持続時間にわたり持続しないM_Lを超える大きさの変化を引き起こしうる（たとえば、アミノ酸タイプを識別できない可能性のある一過性会合イベント、観測領域中への拡散や観測領域内に固着する試薬などの非特異的検出イベント）。それゆえ、いくつかの実施形態では、シグナルパルスは、閾値大きさレベル及び閾値持続時間に基づいてシグナルデータから抽出される。

【００５０】
抽出されたシグナルパルス情報は、例示を目的として重ね合わせられたシグナルトレース例（Ｉ）と共にパネル（ＩＩＩ）に示される。いくつかの実施形態では、シグナルパルスの大きさのピークは、M_L超を持続する持続時間にわたり検出された大きさを平均することにより決定される。いくつかの実施形態では、本明細書で用いられる「シグナルパルス」とは、ある持続時間にわたりベースライン超を持続するシグナルデータの変化（たとえば、シグナルトレース例（Ｉ）により例示される生のシグナルデータ）又はそれから抽出されたシグナルパルス情報（たとえば、パネル（ＩＶ）に例示される処理シグナルデータ）を意味しうることが認識されるべきである。

【００５１】
パネル（ＩＶ）は、シグナルトレース例（Ｉ）から抽出されたシグナルパルス情報を示す。いくつかの実施形態では、シグナルパルス情報は、一連のシグナルパルス中のさまざまな特性パターンに基づいて異なるタイプのアミノ酸を同定するように解析可能である。たとえば、パネル（ＩＶ）に示されるように、シグナルパルス情報は、第１の特性パターン（「CP_1」）に基づいて第１のタイプのアミノ酸及び第２の特性パターン（「CP_2」）に基づいて第２のタイプのアミノ酸の指標となる。例として、より早い時点で検出された２つのシグナルパルスは、CP_1に基づいてポリペプチドの末端の第１のアミノ酸の指標となる情報を提供し、且つより後の時点で検出された２つのシグナルパルスは、ＣＰ

２に基づいてポリペプチドの末端の第２のアミノ酸の指標となる情報を提供する。

【００５２】

また、パネル（ＩＶ）に示されるように、各シグナルパルスは、アミノ酸認識分子と特性パターンのアミノ酸との間の会合イベントに対応するパルス持続時間（「ｐｄ」）を含む。いくつかの実施形態では、パルス持続時間は、結合（ｂｉｎｄｉｎｇ）の解離速度に特有である。また、示されるように、特性パターンの各シグナルパルスは、パルス間持続時間（ｉｎｔｅｒｐｕｌｓｅ　ｄｕｒａｔｉｏｎ）（「ｉｐｄ」）により特性パターンの他のシグナルパルスから分離される。いくつかの実施形態では、パルス間持続時間は、結合の会合速度に特有である。いくつかの実施形態では、大きさの変化（「ΔＭ」）は、ベースラインとシグナルパルスのピークとの間の差に基づいてシグナルパルスに対して決定可能である。いくつかの実施形態では、特性パターンは、パルス持続時間に基づいて決定される。いくつかの実施形態では、特性パターンは、パルス持続時間及びパルス間持続時間に基づいて決定される。いくつかの実施形態では、特性パターンは、パルス持続時間、パルス間持続時間、及び大きさの変化のいずれか１つ以上に基づいて決定される。

【００５３】

それゆえ、図１Ａ～１Ｂにより例示されるように、いくつかの実施形態では、ポリペプチドシーケンシングは、１つ以上のアミノ酸認識分子と進行中の分解反応でポリペプチドの末端に露出される逐次アミノ酸との会合の指標となる一連のシグナルパルスを検出することにより実施される。一連のシグナルパルスは、一連のシグナルパルス中の特性パターンを決定するように解析可能であり、且つ特性パターンの経時変化は、ポリペプチドのアミノ酸配列を決定するために使用可能である。

【００５４】

いくつかの実施形態では、一連のシグナルパルスは、経時的に光学シグナルの大きさの一連の変化を含む。いくつかの実施形態では、光学シグナルの一連の変化は、会合イベント時に生成される発光の一連の変化を含む。いくつかの実施形態では、発光は、シーケンシング反応の１つ以上の試薬に関連付けられる検出可能標識により生成される。たとえば、いくつかの実施形態では、１つ以上のアミノ酸認識分子の各々は発光標識を含む。いくつかの実施形態では、切断試薬は発光標識を含む。本願に係る発光標識及びその使用の例は、本明細書の他の箇所に提供される。

【００５５】

いくつかの実施形態では、一連のシグナルパルスは、経時的に電気シグナルの大きさの一連の変化を含む。いくつかの実施形態では、電気シグナルの一連の変化は、会合イベント時に生成されるコンダクタンスの一連の変化を含む。いくつかの実施形態では、伝導率は、シーケンシング反応の１つ以上の試薬に会合される検出可能標識により生成される。たとえば、いくつかの実施形態では、１つ以上のアミノ酸認識分子の各々は伝導率標識を含む。本願に係る伝導率標識及びその使用の例は、本明細書の他の箇所に提供される。伝導率標識を用いた単一分子の同定方法は記載されている（たとえば、米国特許出願公開第２０１７／００３７４６２号明細書を参照されたい）。

【００５６】

いくつかの実施形態では、コンダクタンスの一連の変化は、ナノポアを介するコンダクタンスの一連の変化を含む。たとえば、ナノポアを用いたレセプター－リガンド相互作用の評価方法は記載されている（たとえば、タークル，Ａ．Ｋ．（Ｔｈａｋｕｒ，Ａ．Ｋ．）及びモビレアヌ，Ｌ．（Ｍｏｖｉｌｅａｎｕ，Ｌ．）著、２０１９年、ネイチャー・バイオテクノロジー（Ｎａｔｕｒｅ　Ｂｉｏｔｅｃｈｎｏｌｏｇｙ）、第３７巻、第１号を参照されたい）。かかるナノポアは本願に係るポリペプチドシーケンシング反応をモニターするために使用されうることを、本発明者らは認識し高く評価した。それゆえ、いくつかの実施形態では、本願は、単一ポリペプチド分子と１つ以上のアミノ酸認識分子とを接触させることを含むポリペプチドシーケンシング方法を提供し、この場合、単一ポリペプチド分子はナノポアに固定される。いくつかの実施形態では、本方法は、１つ以上の末端アミノ酸認識分子と、単一ポリペプチドが分解されつつある間に単一ポリペプチドの末端

に露出される逐次アミノ酸と、の会合の指標となるナノポアを介するコンダクタンスの一連の変化を検出することにより、単一ポリペプチド分子をシーケンスすることをさらに含む。

【００５７】

いくつかの態様では、本願は、混合物からポリペプチドの１タイプ以上のアミノ酸を同定することによるタンパク質の複合混合物中の個別タンパク質のシーケンシング及び／又は同定の方法を提供する。いくつかの実施形態では、ポリペプチドの１つ以上のアミノ酸（たとえば、末端アミノ酸及び／又は内部アミノ酸）が標識され（たとえば、直接的又は間接的に、たとえば、アミノ酸認識分子などの結合剤を用いて）、且つポリペプチド中の標識付アミノ酸の相対位置が決定される。いくつかの実施形態では、ポリペプチド中のアミノ酸の相対位置は、一連のアミノ酸標識及び切断工程を用いて決定される。しかしながら、いくつかの実施形態では、ポリペプチド中の標識付アミノ酸の相対位置は、ポリペプチドからアミノ酸を除去することなく、ポア（たとえばタンパク質チャネル）を介して標識付ポリペプチドをトランスロケートすることと、ポリペプチド分子中の標識付アミノ酸の相対位置を決定するためにポアを介するトランスロケーション時に標識付アミノ酸からのシグナル（たとえばＦＲＥＴシグナル）を検出することと、により決定可能である。

【００５８】

いくつかの実施形態では、末端アミノ酸（たとえば、Ｎ末端又はＣ末端アミノ酸）のアイデンティティーが評価され、その後で末端アミノ酸が除去され、そして末端のその次のアミノ酸のアイデンティティーが評価され、このプロセスがポリペプチド中の複数の逐次アミノ酸が評価されるまで繰り返される。いくつかの実施形態では、アミノ酸のアイデンティティーを評価することは、存在するアミノ酸のタイプを決定することを含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、たとえば、天然に存在する２０アミノ酸のどれであるかを決定することにより、実際のアミノ酸のアイデンティティーを決定することを含む（たとえば、個別末端アミノ酸に特異的な結合剤を用いて）。いくつかの実施形態では、アミノ酸のタイプは、アラニン、アルギニン、アスパラギン、アスパラギン酸、システイン、グルタミン、グルタミン酸、グリシン、ヒスチジン、イソロイシン、ロイシン、リシン、メチオニン、フェニルアラニン、プロリン、セレノシステイン、トレオニン、トリプトファン、チロシン、及びバリンから選択される。

【００５９】

しかしながら、いくつかの実施形態では、末端アミノ酸タイプのアイデンティティーを評価することは、ポリペプチドの末端に存在可能な潜在的アミノ酸のサブセットを決定することを含みうる。いくつかの実施形態では、これは、アミノ酸が１つ以上の特異的アミノ酸でないことを決定することにより達成可能である（したがって、他のアミノ酸のいずれかでありうる）。いくつかの実施形態では、これは、アミノ酸の特定サブセットのどれがポリペプチドの末端にありうるかを決定することにより（たとえば、サイズ、電荷、疎水性、翻訳後修飾、結合性に基づいて）達成可能である（たとえば、２つ以上の末端アミノ酸の特定サブセットに結合する結合剤を用いて）。

【００６０】

いくつかの実施形態では、末端アミノ酸タイプのアイデンティティーを評価することは、アミノ酸が翻訳後修飾を含むことを決定することを含む。翻訳後修飾の非限定的例としては、アセチル化、ＡＤＰリボシル化、カスパーゼ切断、シトルリン化、ホルミル化、Ｎ連結グリコシル化、Ｏ連結グリコシル化、ヒドロキシル化、メチル化、ミリストイル化、ＮＥＤＤ化（ｎｅｄｄｙｌａｔｉｏｎ）、ニトロ化、酸化、パルミトイル化、リン酸化、プレニル化、Ｓ－ニトロシル化、硫酸化、ＳＵＭＯ化（ｓｕｍｏｙｌａｔｉｏｎ）、及びユビキチン化が挙げられる。

【００６１】

いくつかの実施形態では、末端アミノ酸タイプのアイデンティティーを評価することは、アミノ酸が１つ以上の生化学的性質により特徴付けられる側鎖を含むことを決定することを含む。たとえば、アミノ酸は、非極性脂肪族側鎖、正荷電側鎖、負荷電側鎖、非極性

芳香族側鎖、又は極性非荷電側鎖を含みうる。非極性脂肪族側鎖を含むアミノ酸の非限定的例としては、アラニン、グリシン、バリン、ロイシン、メチオニン、及びイソロイシンが挙げられる。正荷電側鎖を含むアミノ酸の非限定的例としては、リシン、アルギニン、及びヒスチジンが挙げられる。負荷電側鎖を含むアミノ酸の非限定的例としては、アスパラギン酸及びグルタミン酸が挙げられる。非極性芳香族側鎖を含むアミノ酸の非限定的例としては、フェニルアラニン、チロシン、及びトリプトファンが挙げられる。極性非荷電側鎖を含むアミノ酸の非限定的例としては、セリン、トレオニン、システイン、プロリン、アスパラギン、及びグルタミンが挙げられる。

【００６２】

いくつかの実施形態では、タンパク質又はポリペプチドは、複数のより小さなポリペプチドに消化可能であり、且つ配列情報は、これらのより小さなポリペプチドの１つ以上から得ることが可能である（たとえば、ポリペプチドの末端アミノ酸を逐次評価することと、そのアミノ酸を除去してその次のアミノ酸を末端に露出することと、を含む方法を用いて）。

【００６３】

いくつかの実施形態では、ポリペプチドは、そのアミノ（Ｎ）末端からシーケンスされる。いくつかの実施形態では、ポリペプチドは、そのカルボキシ（Ｃ）末端からシーケンスされる。いくつかの実施形態では、ポリペプチドの第１の末端（たとえば、Ｎ又はＣ末端）は固定され、且つ他の末端（たとえば、Ｃ又はＮ末端）は本明細書に記載されるようにシーケンスされる。

【００６４】

本明細書で用いられる場合、ポリペプチドをシーケンスすることは、ポリペプチドの配列情報を決定することを意味する。いくつかの実施形態では、これは、ポリペプチドの一部分（又は全部）の各逐次アミノ酸のアイデンティティーを決定することを含みうる。しかしながら、いくつかの実施形態では、これは、ポリペプチド内のアミノ酸のサブセットのアイデンティティーを評価すること（たとえば、ポリペプチド中の各アミノ酸のアイデンティティーを決定することなく１つ以上のアミノ酸タイプの相対位置を決定すること）を含みうる。しかしながら、いくつかの実施形態では、ポリペプチド中のさまざまなタイプのアミノ酸の相対位置を直接決定することなく、ポリペプチドからアミノ酸含有率情報を得ることが可能である。アミノ酸含有率は単独で、存在するポリペプチドのアイデンティティーを推測するために使用されうる（たとえば、アミノ酸含有率とポリペプチド情報のデータベースとを比較することと、どのポリペプチドが同一アミノ酸含有率を有するかを決定することと、により）。

【００６５】

いくつかの実施形態では、（たとえば、酵素切断及び／又は化学切断を介して）より長いポリペプチド又はタンパク質から得られる複数のポリペプチド産物の配列情報は、より長いポリペプチド又はタンパク質の配列を再構築又は推測するように解析可能である。

【００６６】

それゆえ、いくつかの実施形態では、１タイプ以上のアミノ酸は、１タイプ以上のアミノ酸に選択的に結合する１つ以上の標識付認識分子の発光を検出することにより同定される。いくつかの実施形態では、１タイプ以上のアミノ酸は、標識付ポリペプチドの発光を検出することにより同定される。

【００６７】

本明細書に記載のポリペプチドシーケンシング技術は、とくに従来のポリペプチドシーケンシング技術とは対照的に、新規なポリペプチドシーケンシングデータを発生することを含みうることを、本発明者らはさらに認識し高く評価した。そのため、ポリペプチドシーケンシングデータを解析する従来の技術は、本明細書に記載のポリペプチドシーケンシング技術を用いて発生されるデータに適用されるとき十分でないおそれがある。

【００６８】

たとえば、繰返し試薬サイクリングを含む従来のポリペプチドシーケンシング技術は、

シーケンスされるポリペプチドの個別アミノ酸に関連付けられるデータを発生しうる。かかる場合には、検出されるデータが１つのアミノ酸のみに対応するので、発生されるデータを解析することは、特定時間でどのアミノ酸が検出されているかを決定することを単に含むにすぎないであろう。これとは対照的に、本明細書に記載のポリペプチドシーケンシング技術は、ポリペプチド分子の複数のアミノ酸が検出されている間にポリペプチド分解プロセス時のデータを発生しうるので、ポリペプチドのさまざまなアミノ酸に対応するデータのセクションを見分けることが困難なおそれのあるデータをもたらしうる。それゆえ、本発明者らは、個別アミノ酸に対応するデータのセクションを決定することを含む本明細書に記載のポリペプチドシーケンシング技術により、たとえば、それぞれのアミノ酸会合イベントに対応する部分にデータをセグメント化することにより、発生されるかかるデータを解析する新しい計算技術を開発した。次いで、それらのセクションは、それらの個別セクション時に検出されるアミノ酸を同定するようにさらに解析されうる。

【００６９】

他の一例として、各タイプのアミノ酸に対してユニーク同定可能標識を使用することを含む従来のシーケンシング技術は、どのように個別アミノ酸が他の分子と相互作用するかの動力学をなんら考慮することなくどの標識が特定時間に検出されているかを単に解析することを含みうるにすぎない。これとは対照的に、本明細書に記載のポリペプチドシーケンシング技術は、どのようにアミノ酸が認識分子と相互作用するかの指標となるデータを発生する。以上で考察されるように、データは、アミノ酸とそのそれぞれの認識分子との間の会合イベントに対応する一連の特性パターンを含みうる。それゆえ、本発明者らは、データのその部分に対応するアミノ酸のタイプを決定するように特性パターンを解析する新しい計算技術を開発し、一連のさまざまな特性パターンを解析することによりポリペプチドのアミノ酸配列の決定を可能にした。

【００７０】

いくつかの態様では、本明細書に記載のポリペプチドシーケンシング技術は、ポリペプチドが切断手段によって分解されつつある間に、どのようにポリペプチドが結合手段と相互作用するかの指標となるデータを発生する。以上で考察されるように、データは、末端の切断手段間におけるポリペプチドの末端の会合イベントに対応する一連の特性パターンを含みうる。いくつかの実施形態では、本明細書に記載のシーケンシングの方法は、単一ポリペプチド分子と、結合手段及び切断手段とを接触させることを含み、結合手段及び切断手段は、少なくとも１０の会合イベントを、切断イベントの前に達成するように構成される。いくつかの実施形態では、手段は、少なくとも１０の会合イベントを、２つの切断イベントの間に達成するように構成される。

【００７１】

本明細書に記載されるように、いくつかの実施形態では、複数の単一分子シーケンシング反応がサンプルウェルのアレイでパラレルに実施される。いくつかの実施形態では、アレイは、約１０，０００～約１，０００，０００のサンプルウェルを含む。いくつかの実現形態では、サンプルウェルの容積は、約10^{-21}リットル～約10^{-15}リットルでありうる。サンプルウェルは小さな容積を有するので、任意の所与の時間にサンプルウェル内に約１つのみのポリペプチドが存在しうるため、単一分子イベントの検出が可能でありうる。統計的に、いくつかのサンプルウェルは単一分子シーケンシング反応を含まないことがあり、いくつかは１つより多くの単一ポリペプチド分子を含むことがある。しかしながら、相当数のサンプルウェルは、各々に単一分子反応を含みうる（たとえば、いくつかの実施形態では少なくとも３０％）ため、単一分子解析は、多数のサンプルウェルでパラレルに行うことが可能である。いくつかの実施形態では、結合手段及び切断手段は、少なくとも１０の会合イベントを、単一分子反応が起こるサンプルウェルの少なくとも１０％（たとえば、１０～５０％、５０％より多く、２５～７５％、少なくとも８０％、又はそれ以上）における切断イベントの前に達成するように構成される。いくつかの実施形態では、結合手段及び切断手段は、少なくとも１０の会合イベントを、単一分子反応においてポリペプチドのアミノ酸の少なくとも５０％（たとえば、５０％より多く、少なくとも

８０％、又はそれ以上）への切断イベントの前に達成するように構成される。

【００７２】

アミノ酸認識分子

いくつかの実施形態では、本明細書に提供される方法は、ポリペプチドと、少なくとも１タイプの末端アミノ酸に選択的に結合する標識を含んでいても含んでいなくてもよいアミノ酸認識分子と、を接触させることを含む。本明細書で用いられる場合、いくつかの実施形態では、末端アミノ酸とは、ポリペプチドのアミノ末端アミノ酸又はポリペプチドのカルボキシ末端アミノ酸を意味しうる。いくつかの実施形態では、標識付認識分子は、他のタイプの末端アミノ酸よりも１タイプの末端アミノ酸に選択的に結合する。いくつかの実施形態では、標識付認識分子は、同一タイプの内部アミノ酸よりも１タイプの末端アミノ酸に選択的に結合する。さらに他の実施形態では、標識付認識分子は、ポリペプチドのいずれかの位置の１タイプのアミノ酸、たとえば、末端アミノ酸や内部アミノ酸と同一タイプのアミノ酸に選択的に結合する。

【００７３】

本明細書で用いられる場合、いくつかの実施形態では、アミノ酸のタイプとは、２０の天然に存在するアミノ酸の１つ又はそのタイプのサブセットを意味する。いくつかの実施形態では、アミノ酸のタイプとは、２０の天然に存在するアミノ酸の１つの修飾変異体又はその非修飾及び／若しくは修飾変異体のサブセットを意味する。修飾アミノ酸変異体の例としては、限定されるものではないが、翻訳後修飾変異体（たとえば、アセチル化、ＡＤＰリボシル化、カスパーゼ切断、シトルリン化、ホルミル化、Ｎ連結グリコシル化、Ｏ連結グリコシル化、ヒドロキシル化、メチル化、ミリストイル化、ＮＥＤＤ化、ニトロ化、酸化、パルミトイル化、リン酸化、プレニル化、Ｓ－ニトロシル化、硫酸化、ＳＵＭＯ化、及びユビキチン化）、化学修飾変異体、非天然アミノ酸、及びタンパク質原性アミノ酸、たとえば、セレノシステイン及びピロリシンが挙げられる。いくつかの実施形態では、アミノ酸のタイプのサブセットは、１つ以上の類似の生化学的性質を有する１超且つ２０未満のアミノ酸を含む。たとえば、いくつかの実施形態では、アミノ酸のタイプとは、荷電側鎖（たとえば、正及び／又は負荷電側鎖）を有するアミノ酸、極性側鎖（たとえば、極性非荷電側鎖）を有するアミノ酸、非極性側鎖（たとえば、非極性脂肪族及び／又は芳香族側鎖）を有するアミノ酸、及び疎水性側鎖を有するアミノ酸から選択される１タイプを意味する。

【００７４】

いくつかの実施形態では、本明細書に提供される方法は、ポリペプチドと、１タイプ以上の末端アミノ酸に選択的に結合する１つ以上の標識付認識分子と、を接触させることを含む。例示的且つ非限定的例として、４つの標識付認識分子が本願の方法で使用される場合、いずれか１つの認識分子は、他の３つのいずれかが選択的に結合する他の１タイプのアミノ酸とは異なる１タイプの末端アミノ酸に選択的に結合する（たとえば、第１の認識分子は第１のタイプに結合し、第２の認識分子は第２のタイプに結合し、第３の認識分子は第３のタイプに結合し、且つ第４の認識分子は第４のタイプの末端アミノ酸に結合する）。この考察の目的では、本明細書に記載の方法との関連での１つ以上の標識付認識分子は、標識付認識分子のセットと代替的にいいうる。

【００７５】

いくつかの実施形態では、標識付認識分子のセットは、少なくとも１つ且つ６つまでの標識付認識分子を含む。たとえば、いくつかの実施形態では、標識付認識分子のセットは、１つ、２つ、３つ、４つ、５つ、又は６つの標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、１０以下の標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、８つ以下の標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、６つ以下の標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、４つ以下の標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、３つ以下の標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、２つ以下の標識付認識分子を含む。いく

つかの実施形態では、標識付認識分子のセットは、４つの標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、少なくとも２つ且つ２０まで（たとえば、少なくとも２つ且つ１０まで、少なくとも２つ且つ８つまで、少なくとも４つ且つ２０まで、少なくとも４つ且つ１０まで）の標識付認識分子を含む。いくつかの実施形態では、標識付認識分子のセットは、２０超（たとえば、２０～２５、２０～３０）の認識分子を含む。しかしながら、いずれの数の認識分子も所望の使用に適応する本願の方法に従って使用されうることが認識されるべきである。

【００７６】

本願によれば、いくつかの実施形態では、１タイプ以上のアミノ酸は、標識付認識分子の発光を検出することにより同定される。いくつかの実施形態では、標識付認識分子は、１タイプのアミノ酸に選択的に結合する認識分子と、認識分子に関連付けられる発光を有する発光標識と、を含む。こうして、発光（たとえば、発光寿命、発光強度、及び本明細書の他の箇所に記載の他の発光性）は、ポリペプチドのアミノ酸を同定するように認識分子の選択的結合に関連付けられうる。いくつかの実施形態では、複数のタイプの標識付認識分子が本願に係る方法で使用されうるとともに、各タイプは、複数間でユニークに同定可能な発光を有する発光標識を含む。好適な発光標識は、フルオロフォア色素などの発光分子を含みうるとともに本明細書の他の箇所に記載されている。

【００７７】

いくつかの実施形態では、１タイプ以上のアミノ酸は、標識付認識分子の１つ以上の電気特性を検出することにより同定される。いくつかの実施形態では、標識付認識分子は、１タイプのアミノ酸に選択的に結合する認識分子と、認識分子に会合される伝導率標識と、を含む。こうして、１つ以上の電気特性（たとえば、電荷、電流振動カラー、及び他の電気特性）は、ポリペプチドのアミノ酸を同定するように認識分子の選択的結合に関連付けられうる。いくつかの実施形態では、複数のタイプの標識付認識分子が本願に係る方法で使用されうるとともに、各タイプは、複数間でユニークに同定可能な電気シグナルの変化（たとえば、コンダクタンスの変化、たとえば、特性パターンの伝導率の大きさの変化及び伝導率の遷移）を生成する伝導率標識を含む。いくつかの実施形態では、複数のタイプの標識付認識分子は各々、異なる数の荷電基（たとえば、異なる数の負及び／又は正荷電基）を有する伝導率標識を含む。それゆえ、いくつかの実施形態では、伝導率標識は電荷標識である。電荷標識の例としては、デンドリマー、ナノ粒子、核酸、及び複数の荷電基を有する他のポリマーが挙げられる。いくつかの実施形態では、伝導率標識は、その正味電荷（ｎｅｔ　ｃｈａｒｇｅ）（たとえば、正味正電荷又は正味負電荷）により、その電荷密度により、及び／又はその荷電基数によりユニークに同定可能である。

【００７８】

いくつかの実施形態では、アミノ酸認識分子は、従来から公知の技術を用いて当業者により工学操作されうる。いくつかの実施形態では、望ましい性質は、ポリペプチドの末端（たとえば、Ｎ末端又はＣ末端）に位置するときのみ１タイプのアミノ酸に選択的に且つ高親和性で結合する能力を含みうる。さらに他の実施形態では、望ましい性質は、ポリペプチドの末端（たとえば、Ｎ末端又はＣ末端）に位置するとき及びポリペプチドの内部位置に位置するとき１タイプのアミノ酸に選択的に且つ高親和性で結合する能力を含みうる。いくつかの実施形態では、望ましい性質は、１タイプ超のアミノ酸に選択的に且つ低親和性で（たとえば、約５０ｎＭ以上、たとえば、約５０ｎＭ～約５０μＭ、約１００ｎＭ～約１０μＭ、約５００ｎＭ～約５０μＭのＫＤで）結合する能力を含む。たとえば、いくつかの態様では、本願は、ポリペプチド分解プロセス時の可逆結合相互作用を検出することによるシーケンシング方法を提供する。有利には、かかる方法は、１タイプ超のアミノ酸（たとえば、アミノ酸タイプのサブセット）に低親和性で可逆結合する認識分子を用いて実施されうる。

【００７９】

本明細書で用いられる場合、いくつかの実施形態では、「選択的」及び「特異的」（及びそれらの変化形、たとえば、選択的に、特異的に、選択性、特異性）という用語は、優

先的結合相互作用を意味する。たとえば、いくつかの実施形態では、１タイプのアミノ酸に選択的に結合するアミノ酸認識分子は、１タイプのアミノ酸に他の１タイプのアミノ酸よりも優先的に結合する。選択的結合相互作用は、１タイプのアミノ酸（たとえば１タイプの末端アミノ酸）と他のタイプのアミノ酸（たとえば他のタイプの末端アミノ酸）とを典型的には約１０倍超～１００倍以上（たとえば、約１，０００倍超又は１０，０００倍超）で識別するであろう。それゆえ、選択的結合相互作用とは、１タイプのアミノ酸を他のタイプのアミノ酸よりもユニークに同定可能ないずれかの結合相互作用を意味しうることが認識されるべきである。たとえば、いくつかの態様では、本願は、１つ以上のアミノ酸認識分子とポリペプチド分子との会合の指標となるデータを得ることによるポリペプチドシーケンシング方法を提供する。いくつかの実施形態では、データは、ポリペプチド分子のアミノ酸との一連の可逆アミノ酸認識分子結合相互作用に対応する一連のシグナルパルスを含み、且つデータは、アミノ酸のアイデンティティーを決定するために使用されうる。したがって、いくつかの実施形態では、「選択的」又は「特異的」結合相互作用とは、１タイプのアミノ酸と他のタイプのアミノ酸とを識別する検出結合相互作用を意味する。

【００８０】

いくつかの実施形態では、アミノ酸認識分子は、他のタイプのアミノ酸に有意に結合することなく約10^{-6}Ｍ未満（たとえば、約10^{-7}Ｍ未満、約10^{-8}Ｍ未満、約10^{-9}Ｍ未満、約10^{-10}Ｍ未満、約10^{-11}Ｍ未満、約10^{-12}Ｍ未満、10^{-16}Ｍ程度まで）の解離定数（K_D）で１タイプのアミノ酸に結合する。いくつかの実施形態では、アミノ酸認識分子は、約１００ｎＭ未満、約５０ｎＭ未満、約２５ｎＭ未満、約１０ｎＭ未満、又は約１ｎＭ未満のK_Dで１タイプのアミノ酸（たとえば１タイプの末端アミノ酸）に結合する。いくつかの実施形態では、アミノ酸認識分子は、約５０ｎＭ～約５０μＭ（たとえば、約５０ｎＭ～約５００ｎＭ、約５０ｎＭ～約５μＭ、約５００ｎＭ～約５０μＭ、約５μＭ～約５０μＭ、又は約１０μＭ～約５０μＭ）のK_Dで１タイプのアミノ酸に結合する。いくつかの実施形態では、アミノ酸認識分子は、約５０ｎＭのK_Dで１タイプのアミノ酸に結合する。

【００８１】

いくつかの実施形態では、アミノ酸認識分子は、約10^{-6}Ｍ未満（たとえば、約10^{-7}Ｍ未満、約10^{-8}Ｍ未満、約10^{-9}Ｍ未満、約10^{-10}Ｍ未満、約10^{-11}Ｍ未満、約10^{-12}Ｍ未満、10^{-16}Ｍ程度まで）のK_Dで２タイプ以上のアミノ酸に結合する。いくつかの実施形態では、アミノ酸認識分子は、約１００ｎＭ未満、約５０ｎＭ未満、約２５ｎＭ未満、約１０ｎＭ未満、又は約１ｎＭ未満のK_Dで２タイプ以上のアミノ酸に結合する。いくつかの実施形態では、アミノ酸認識分子は、約５０ｎＭ～約５０μＭ（たとえば、約５０ｎＭ～約５００ｎＭ、約５０ｎＭ～約５μＭ、約５００ｎＭ～約５０μＭ、約５μＭ～約５０μＭ、又は約１０μＭ～約５０μＭ）のK_Dで２タイプ以上のアミノ酸に結合する。いくつかの実施形態では、アミノ酸認識分子は、約５０ｎＭのK_Dで２タイプ以上のアミノ酸に結合する。

【００８２】

いくつかの実施形態では、アミノ酸認識分子は、少なくとも０．１s^{-1}の解離速度（K_{off}）で少なくとも１タイプのアミノ酸に結合する。いくつかの実施形態では、解離速度は、約０．１s^{-1}～約１，０００s^{-1}（たとえば、約０．５s^{-1}～約５００s^{-1}、約０．１s^{-1}～約１００s^{-1}、約１s^{-1}～約１００s^{-1}、又は約０．５s^{-1}～約５０s^{-1}）である。いくつかの実施形態では、解離速度は、約０．５s^{-1}～約２０s^{-1}である。いくつかの実施形態では、解離速度は、約２s^{-1}～約２０s^{-1}である。いくつかの実施形態では、解離速度は、約０．５s^{-1}～約２s^{-1}である。

【００８３】

いくつかの実施形態では、K_D又はk_{off}の値は既知の文献の値でもよく、又は値は実験的に決定されてもよい。たとえば、K_D又はk_{off}の値は、単一分子アッセイ又はアンサンブルアッセイにおいて測定されうる（たとえば、実施例４及び図１９Ｍを参照さ

れたい）。いくつかの実施形態では、k_{off}の値は、本明細書の他の箇所に記載されるように単一分子アッセイで取得したシグナルパルス情報に基づいて実験的に決定されうる。たとえば、k_{off}の値は、平均パルス持続時間の逆数によって概算されてもよい。いくつかの実施形態では、アミノ酸認識分子は、２タイプ以上のアミノ酸に２タイプ以上の各々に対して異なるK_D又はk_{off}で結合する。いくつかの実施形態では、第１のタイプのアミノ酸に対する第１のK_D又はk_{off}は、第２のタイプのアミノ酸に対する第２のK_D又はk_{off}とは少なくとも１０％（たとえば、少なくとも２５％、少なくとも５０％、少なくとも１００％、又はそれ以上）だけ異なる。いくつかの実施形態では、K_D又はk_{off}の第１の値と第２の値とは、約１０～２５％、２５～５０％、５０～７５％、７５～１００％、又は１００％超、たとえば、約２倍、３倍、４倍、５倍、又はそれ以上に異なる。

【００８４】

本明細書に提供される方法及び組成物に従って、図２は、標識付認識分子の各種構成例及び使用例を示す。いくつかの実施形態では、標識付認識分子２００は、発光標識２１０（たとえば標識）と、ポリペプチド２２０の１タイプ以上の末端アミノ酸に選択的に結合する認識分子（斑点形状として示される）と、を含む。いくつかの実施形態では、認識分子は、末端位置又は末端位置及び内部位置の両方の１タイプのアミノ酸又はアミノ酸のタイプのサブセット（たとえば、２０未満の通常のタイプのアミノ酸）に対して選択的である。

【００８５】

本明細書に記載されるように、「アミノ酸認識分子」は、１分子に他の１分子よりも（たとえば、１タイプのアミノ酸に他の１タイプのアミノ酸よりも）選択的又は特異的に結合可能ないずれかのバイオ分子でありうる。いくつかの実施形態では、認識分子は、ペプチダーゼでないか又はペプチダーゼ活性を有していない。たとえば、いくつかの実施形態では、本願のポリペプチドシーケンシング方法は、ポリペプチド分子と１つ以上の認識分子及び切断試薬とを接触させることを含む。かかる実施形態では、１つ以上の認識分子は、ペプチダーゼ活性を有しておらず、且つポリペプチド分子からの１つ以上のアミノ酸の除去（たとえば、ポリペプチド分子の末端からのアミノ酸除去）は、切断試薬により実施される。

【００８６】

認識分子は、たとえば、合成又は組換えでありうるタンパク質及び核酸を含む。いくつかの実施形態では、認識分子は、抗体若しくは抗体の抗原結合部分、ＳＨ２ドメイン含有タンパク質若しくはその断片、又は酵素バイオ分子、たとえば、ペプチダーゼ、アミノトランスフェラーゼ、リボザイム、アプタザイム、若しくはｔＲＮＡシンテターゼ（「繰返しポリペプチド分析及び処理のための分子及び方法（ＭＯＬＥＣＵＬＥＳ ＡＮＤ ＭＥＴＨＯＤＳ ＦＯＲ ＩＴＥＲＡＴＩＶＥ ＰＯＬＹＰＥＰＴＩＤＥ ＡＮＡＬＹＳＩＳ ＡＮＤ ＰＲＯＣＥＳＳＩＮＧ）」という名称の２０１６年９月２日出願の米国特許出願第１５／２５５，４３３号明細書に記載のアミノアシルｔＲＮＡシンテターゼ及び関連分子を含む）でありうる。

【００８７】

いくつかの態様では、本願は、本明細書に記載の方法又は当分野で知られた方法に従って使用するためのアミノ酸認識分子の発見及び開発に関する。いくつかの実施形態では、本願は、タンパク質ファミリーの他のホモログメンバー間に存在することが以前は知られていなかった、結合性を有するアミノ酸結合タンパク質（たとえば、ＣｌｐＳタンパク質）を提供する。いくつかの実施形態では、本願は、工学操作されたアミノ酸結合タンパク質を提供する。たとえば、いくつかの実施形態では、本願は、２つ以上のアミノ酸結合タンパク質の直列コピーを有する単一ポリペプチドを含む融合構築物を提供する。

【００８８】

本発明者らは、本願の融合構築物が、標識バックグラウンドノイズ（たとえば、バックグラウンド蛍光）を増加させることなく、認識分子濃度を効果的に増加させることを可能

にすることを認識し、且つ評価している。発明者らは、本願の融合構築物が、シーケンシング反応の確度を増加させ、及び／又はシーケンシング反応を実施するために必要な時間量を減少させることを認識し、且つ評価している。そのほか、アミノ酸結合タンパク質の２タイプ以上の異なる直列コピーを有する融合構築物を提供することによって、反応により少ない試薬しか必要とされないので、これは、シーケンスのより効率的且つ安価なアプローチを提供する。

【００８９】

いくつかの実施形態では、本願の認識分子は分解経路タンパク質である。認識分子として使用するのに好適な分解経路タンパク質の例としては、限定されるものではないが、Ｎ末端則経路タンパク質、たとえば、Ａｒｇ／Ｎ末端則経路タンパク質、Ａｃ／Ｎ末端則経路タンパク質、及びＰｒｏ／Ｎ末端則経路タンパク質が挙げられる。いくつかの実施形態では、認識分子は、Ｇｉｄタンパク質（たとえば、Ｇｉｄ４又はＧｉｄ１０タンパク質）、ＵＢＲボックスタンパク質（たとえば、ＵＢＲ１、ＵＢＲ２）又はそのＵＢＲボックスドメイン含有タンパク質断片、ｐ６２タンパク質又はそのＺＺドメイン含有断片、及びＣｌｐＳタンパク質（たとえば、ＣｌｐＳ１、ＣｌｐＳ２）から選択されるＮ末端則経路タンパク質である。それゆえ、いくつかの実施形態では、標識付認識分子２００は、分解経路タンパク質を含む。いくつかの実施形態では、標識付認識分子２００は、ＣｌｐＳタンパク質である。

【００９０】

いくつかの実施形態では、本願の認識分子は、ＣｌｐＳタンパク質、たとえば、アグロバクテリウム・ツミファシエンス（Ａｇｒｏｂａｃｔｅｒｉｕｍ ｔｕｍｉｆａｃｉｅｎｓ）ＣｌｐＳ１、アグロバクテリウム・ツミファシエンス（Ａｇｒｏｂａｃｔｅｒｉｕｍ ｔｕｍｉｆａｃｉｅｎｓ）ＣｌｐＳ２、シネココッカス・エロンガタス（Ｓｙｎｅｃｈｏｃｏｃｃｕｓ ｅｌｏｎｇａｔｕｓ）ＣｌｐＳ１、シネココッカス・エロンガタス（Ｓｙｎｅｃｈｏｃｏｃｃｕｓ ｅｌｏｎｇａｔｕｓ）ＣｌｐＳ２、サーモシネココッカス・エロンガタス（Ｔｈｅｒｍｏｓｙｎｅｃｈｏｃｏｃｃｕｓ ｅｌｏｎｇａｔｕｓ）ＣｌｐＳ、エシェリキア・コリ（Ｅｓｃｈｅｒｉｃｈｉａ ｃｏｌｉ）ＣｌｐＳ、又はプラスモジウム・ファルシパラム（Ｐｌａｓｍｏｄｉｕｍ ｆａｌｃｉｐａｒｕｍ）ＣｌｐＳである。いくつかの実施形態では、認識分子は、Ｌ／Ｆトランスフェラーゼ、たとえば、エシェリキア・コリ（Ｅｓｃｈｅｒｉｃｈｉａ ｃｏｌｉ）ロイシル／フェニルアラニルｔＲＮＡタンパク質トランスフェラーゼである。いくつかの実施形態では、認識分子は、Ｄ／Ｅロイシルトランスフェラーゼ、たとえば、ビブリオ・バルニフィカス（Ｖｉｂｒｉｏ ｖｕｌｎｉｆｉｃｕｓ）アスパラギン酸／グルタミン酸ロイシルトランスフェラーゼＢｐｔである。いくつかの実施形態では、認識分子は、ＵＢＲタンパク質又はＵＢＲボックスドメイン、たとえば、ＵＢＲタンパク質又はヒトＵＢＲ１及びＵＢＲ２若しくはサッカロマイセス・セレビシエ（Ｓａｃｃｈａｒｏｍｙｃｅｓ ｃｅｒｅｖｉｓｉａｅ）ＵＢＲ１のＵＢＲボックスドメインである。いくつかの実施形態では、認識分子は、ｐ６２タンパク質、たとえば、Ｈ．サピエンス（Ｈ．ｓａｐｉｅｎｓ）ｐ６２タンパク質又はラタス・ノルベギカス（Ｒａｔｔｕｓ ｎｏｒｖｅｇｉｃｕｓ）ｐ６２タンパク質又は最低限でもＺＺドメインを含むそれらのトランケーション変異体である。いくつかの実施形態では、認識分子は、Ｇｉｄ４タンパク質、たとえば、Ｈ．サピエンス（Ｈ．ｓａｐｉｅｎｓ）ＧＩＤ４又はサッカロマイセス・セレビシエ（Ｓａｃｃｈａｒｏｍｙｃｅｓ ｃｅｒｅｖｉｓｉａｅ）ＧＩＤ４である。いくつかの実施形態では、認識分子は、Ｇｉｄ１０タンパク質、たとえば、サッカロマイセス・セレビシエ（Ｓａｃｃｈａｒｏｍｙｃｅｓ ｃｅｒｅｖｉｓｉａｅ）ＧＩＤ１０である。いくつかの実施形態では、認識分子は、Ｎ－メリストイルトランスフェラーゼ（Ｎ－ｍｅｒｉｓｔｏｙｌｔｒａｎｓｆｅｒａｓｅ）、たとえば、リーシュマニア・メジャー（Ｌｅｉｓｈｍａｎｉａ ｍａｊｏｒ）Ｎ－メリストイルトランスフェラーゼ（Ｎ－ｍｅｒｉｓｔｏｙｌｔｒａｎｓｆｅｒａｓｅ）又はＨ．サピエンス（Ｈ．ｓａｐｉｅｎｓ）Ｎ－メリストイルトランスフェラーゼ（Ｎ－ｍｅｒｉｓｔｏｙｌｔｒａｎｓｆｅｒａｓｅ）ＮＭＴ１である。いくつかの実施形態では、認識分子は、ＢＩ

Ｒ２タンパク質、たとえば、ドロソフィラ・メラノガステル（Ｄｒｏｓｏｐｈｉｌａ ｍｅｌａｎｏｇａｓｔｅｒ）ＢＩＲ２である。いくつかの実施形態では、認識分子は、チロシンキナーゼ又はチロシンキナーゼのＳＨ２ドメイン、たとえば、Ｈ．サピエンス（Ｈ．ｓａｐｉｅｎｓ）ＦｙｎＳＨ２ドメイン、Ｈ．サピエンス（Ｈ．ｓａｐｉｅｎｓ）Ｓｒｃチロシンキナーゼ ＳＨ２ドメイン、又はそれらの変異体、たとえば、Ｈ．サピエンス（Ｈ．ｓａｐｉｅｎｓ）ＦｙｎＳＨ２ドメイン三重突然変異体スーパーバインダーである。いくつかの実施形態では、認識分子は、抗体又は抗体フラグメント、たとえば、ホスホチロシンに対する一本鎖抗体可変フラグメント（ｓｃＦｖ）又は本明細書に記載の他の翻訳後修飾アミノ酸変異体である。

【００９１】

表１及び表２は、アミノ酸認識分子の配列例のリストを提供する。また、表１及び表２にとくに明記されていない限り、ポリペプチドの末端位置のアミノ酸アイデンティティーに対する各分子のアミノ酸結合優先度も示される。これらの配列及び本明細書に記載の他の例は、非限定的であることが意図され、且つ本願に係る認識分子は、最低限でもペプチド認識を担うドメイン又はサブドメインを含有するそのいずれかのホモログ、変異体、又はそれらの断片を含みうることが認識されるべきである。

【００９２】

【表１－１】

表１．ClpSアミノ酸認識タンパク質の非限定的例。

名称	結合 優先度*	配列 番号	配列
PS368	F, Y, W, L	1	MASAPSTTLDKSTQVVKKTYPNYKVIVLNDDLNTFDHVANCLIKYIPDMTTDRAWELTNQVHYQGQAIVWTGPQEQAELYHQQLRREGLTMAPLEAA
PS369	F, Y, W, L	2	MTSTLRARPARDTDLQHRPYPHYRIIVLDDDVNTFQHVVNCLVTFLPGMTRDQAWAMAQQVDGEGSAVVWTGPQEQAELYHVQLGNHGLTMAPLEPV
PS370	F, L	3	MFNSLGTVLDPKKSKAKYPEARVIVLDDNFNTFQHVANCLLAIIPRMCEQRAWDLTIKVDKAGSAEVWRGNLEQAELYHEQLFSKGLTMAPIEKT
PS371	F, Y, W, L	4	MATETIERPRTRDPGSGLGGHWLVIVLNDDHNTFDHVAKTLARVIPGVTVDDGYRFADQIHQRGQAIVWRGPKEPAEHYWEQLQDAGLSMAPLERH

【００９３】

【表１－２】

PS372	L, I, V	5	MAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIEM LQKIFGFPPEKGFQIAEEVDRTGRVILLTTSKEHAELKQ DQVHSYGPDPYLGRPCSGSMTCVIEPAV
PS373		6	MNRIKQEAVRTENLLICSESIRRTPGTMSNEESMFDEVV AVAVAEPETQHDERRGTKPKRQPPYHVILWDDTDHSFDY VIMMMKRLFRMPIEKGFQVAKEVDSSGRAICMTTTLELA ELKRDQIHAFGKDELLPRCKGSMSATIEPAEG
PS374	F, Y, W, L	7	MRWEDPLAAEPVTPGVAPVVEEETDAAVETPWRVILYDD DIHTFEEVILQLMKATGCTPEQGERHAWTVHTRGKDCVY QGDFFDCFRVQGVLREIQLVTEIEG
PS375	F, L	8	MEAEPETKVLASIPGVGTSEPFRVVLFNDEEHSFDEVIF QIIKAVRCSRAKAEALTMEVHNSGRSIVYTGPIEQCIRV SAVLEEIELRTEIQS
PS376	F, W, L	9	MPTNDLDLLEKQDVKIERPKMYQVVMYNDDFTPFDFVVA VLMQFFNKGMDEATAIMMQVHMQGKGICGVFPKDIAETK ATEVMKWAKVEQHPLRLQVEAQA
PS377	W	10	MADISKSRPEIGGPKGPQFGDSDRGGGVAVITKPVTKKK FKRKSQTEYEPYWHVLLHHDNVHTFEYATGAIVKVVRTV SRKTAHRITMQAHVSGVATVTTTWKAQAEEYCKGLQMHG LTSSIAPDSSFTH
PS378	F, Y, W, L	11	MXPQEVEEVSFLESKEHEIVLYNDDVNTFDHVIECLVKI CNHNYLQAEQCAYIVHHSGKCGVKTGSLEELIPKCNALL EEGLSAEVI
PS379		12	MSTQEEVLEEVKTTTQKENEIVLYNDDYNTFDHVIETLI YACEHTPVQAEQCAILVHYKGKCTVKTGSFDELKPRCSK LLEEGLSAEIV
PS380	F, W	13	MGDIYGESNPEEVSCIDSLSEEGNELILFNDNIHTFEYV IDCLVAICSLSYEQASNCAYIVDRKGLCTVKHGSYDELL IMYHALVEKDLKVEIR
PS381		14	MVAFSKKWKKDELDKSTGKQKMLILHNDSVNSFDYVIKT LCEVCDHDTIQAEQCAFLTHFKGQCEIAVGEVADLVPLK NKLLNKNLIVSIH
PS382	F, Y, W, L	15	MSDSPVIKEIKKDNIKEADEHEKKEREKETSAWKVILYN DDIHNFTYVTDVIVKVVGQISKAKAHTITVEAHSTGQAL ILSTWKSKAEKYCQELQQNGLTVSIIHESQLKDKQKK
PS388	F, Y, L	16	MVTTLSADVYGMATAPTVAPERSNQVVRKTYPNYKVIVL NDDFNTFQHVAECLMKYIPGMSSDRAWDLTNQVHYEGQA IVWVGPQEPAELYHQQLRRAGLTMAPLEAA
PS389	F, Y, L	17	MLNSAAFKAASASPVIAPERSGQVTQKPYPTYKVIVLND DFNTFQHVHDCLVKYIPGMTSDRAWQLTHQVHNDGQAIV WVGPQEQAELYHQQLSRAGLTMAPIEAA
PS390	F, Y, L	18	MLSIAAVTEAPSKGVQTADPKTVRKPYPNYKVIVLNDDF NTFQHVSSCLLKYIPGMSEARAWELTNQVHFEGLAVVWV GPQEQAELYYAQLKNAGLTMAPPEPA
PS391	F, Y, W, L	19	MGQTVEKPRVEGPGTGLGGSWRVIVRNDDHNTFDHVART LARFIPGVSLERGHEIAKVIHTTGRAVVYTGHKEAAEHY WQQLKGAGLTMAPLEQG
PS392	F, Y, W	20	MSVEIIEKRSTVRKLAPRYRVLLHNDDFNPMEYVVQTLM ATVPSLTQPQAVNVMMEAHTNGMGLVIVCALEHAEFYAE TLNNHGLGSSIEPDD
PS393	F, Y, W, L	21	MSDEDGEDGDENAVGIATRTRTRTKKPTPYRVLLLNDDY TPMEFVVLVLQRFFRMSIEDATRVMLQVHQKGVGVCGVF TYEVAETKVSQVIDFARQNQHPLQCTLEKA
PS394	F, Y, W, L	22	MAERRDTGDDEGTGLGIATKTRSKTKKPTPYRVLMLNDD YTPMEFVVLCLQRFFRMNMEEATRVMLHVHQKGVGVCGV FSYEVAETKVGQVIDFARANQHPLQCTLEKA

【００９４】

【表１－３】

PS395	F, Y, W, L	23	MTVSQSKTQGAPAAQSATELEYEGLWRVVVLNDPVNLMS YVVLVFKKVFGFDETTARKHMLEVHEQGRSVVWSGMREK AEAYAFTLQQWHLTTVLEQDEVR
PS396	F, W	24	MSDNDVALKPKIKSKPKLERPKLYKVILVNDDFTPREFV IAVLKMVFRMSEETGYRVMLTAHRLGTSVVVVCARDIAE TKAKEAVDFGKEAGFPLMFTTEPEE
PS397	F, Y, W	25	MSDNEVAPKRKTRVKPKLERPRLYKVILVNDDYTPRDFV VMVLKAIFRMSEEAGYRVMMTAHKLGTSVVVVCARDIAE TKAKEATDLGKEAGFPLMFTTEPEE
PS398	F, W	26	MPLKAQNRSIVGRRDEWPPPTTQSSSETKSESKRVSDTG ADTKRKTKTVPKVEKPRLYKVILVNDDYTPREFVLVVLK AVFRMSEDQGYKVMITAHQKGSCVVAVYTRDIAETKAKE AVDLAKEIGFPLMFRTEPEE
PS404	F, Y, W	27	MPVSVTAPQTKTKTKPKVERPKLYKVILVNDDFTPREFV VRVLKAEFRMSEDQAAKVMMTAHQRGVCVVAVFTRDVAE TKATRATDAGRAKGYPLLFTTEPEE
PS405	F, Y, W	28	MVSIGAATVACAEGRPIFSGYFDWLAAMPETVTVPRTRL RPKTERPKLHKVILVNDDYTPREFVVTVLKGEFHMSEDQ AQRVMITAHRRGVCVVAVFTKDVAETKATRASDAGRAKG YPLQFTTEPEE
PS406	F, Y, W	29	MPDATTTPRTKTLTRTARPPLHKVILVNDDFTPREFVVR LLKAEFRTTGDEAQRIMITAHMKGSCVVAVFTREIAESK ATRATETARAEGFPLLFTTEPEE
PS407	F, Y, W, L	30	MPSNKRQMCLSDIKNSFNESGIVDWHISPRLANEPSEEG DSDLAVQTVPPELKRPPLYAVVLLNDDYTPMEFVIEILQ QYFAMNLDQATQVMLTVHYEGKGVAGVYPRDIAETKANQ VNNYARSQGHPLLCQIEPKD
PS408	F, Y, W, L	31	MTDPPSKGREDVDLATRTKPKTQRPPLYKVLLLNDDFTP MEFVVHILERLFGMTHAQAIEIMLTVHRKGVAVVGVFSH EIAETKVAQVMELARRQQHPLQCTMEKE
PS409	F, Y, W, L	32	MPARLTDIEGEPNTDPVEDVLLADPELKKPQMYAVVMYN DDYTPMEFVVDVLQNHFKHTLDSAISIMLAIHQQGKGIA GIYPKDIAETKAQTVNRKARQAGYPLLSQIEPQG
PS410	F, W, L	33	MGDDDQSSREGEGDVAFQTADPELKRPSLYRVVLLNDDY TPMEFVVHILEQFFAMNREKATQVMLAVHTQGKGVCGVY TKDIAETKAALVNDYSRENQHPLLCEVEELDDESR
PS411	F, Y, W, L	34	MTRPDAPEYDDDLAVEPAEPELARPPLYKVVLHNDDFTP MEFVVEVLQEFFNMDSEQAVQVMLAVHTQGKATCGIFTR DIAETKSYQVNEYARECEHPLMCDIEAAD
PS412	F, Y, W, L	35	MATKREGSTLLEPTAAKVKPPPLYKVLLLNDDYTPMEFV VLVLKKFFGIDQERATQIMLKVHTEGVGVCGVYPRDIAH TKVEQVVDFARQHQHPLQCTMEES
PS413	F, Y, W, L	36	MMKQCGSYFLIKAVQDFKPLSKHRSDTDVITETKIQVKQ PKLYTVIMYNDNYTTMDFVVYVLVEIFQHSIDKAYQLMM QIHESGQAAVALLPYDLAEMKVDEVTALAEQESYPLLTT IEPA
PS414	F, Y, W, L	37	MQAAGNEPPDPQNPGDVGNGGDGGNQDGSNTGVVVKTRT RTRKPAMYKVLMLNDDYTPMEFVVHVLERFFQKNREEAT RIMLHVHRRGVGVCGVYTYEVAETKVTQVMDLARQNQHP LQCTIEKE
PS415	F, Y, W	38	MALPETRTKIKPDVNIKEPPNYRVIYLNDDKTSMEFVIG SLMQHFSYPQQQAVEKTEEVHEHGSSTVAVLPYEMAEQK GIEVTLDARAEGFPLQVKIEPAER
PS416	F, Y, W, L	39	MTSQTDTLVKPNIQPPSLFKVIYINDSVTTMEFVVESLM SVFNHSADEATRLTQLVHEEGAAVVAILPYELAEQKGME VTLLARNNGFPLAIRLEPAV

【００９５】

【表１－４】

PS417	F, Y, W	40	MSNLDTDVLIDEKVKVVTTEPEKYRVILLNDDVTPMDFV INILVSIFKHSTDTAKDLTLKIHKEGSAIVGVYTYEIAE QKGIEATNESRQHGFPLQVKIERENTL
PS418	F, Y, W, L	41	MSDHNIDHDTSVAVHLDVVVREPPMYRVVLLNDDFTPME FVVELLMHFFRKTAEQATQIMLNIHHEGVGVCGTYPREI AETKVAQVHQHARTNGHPLKCRMEPS
PS419	F, Y, W, L	42	MEKEQSLCKEKTHVELSEPKHYKVVFHNDDFTTMDFVVK VLQLVFFKSQLQAEDLTMKIHLEGSATAGIYSYDIAQSK AQKTTQMAREEGFPLRLTVEPEDN
PS420	F, Y, W, L	43	MSDYSNQISQAGSGVAEDASITLPPERKVVFYNDDFTTM EFVVDVLVSIFNKSHSEAEELMQTVHQEGSSVVGVYTYD IAVSRTNLTIQAARKNGFPLRVEVE
PS421	F, Y, W, L	44	MTTPNKRPEFEPEIGLEDEVGEPRKYKVLLHNDDYTTMD FVVQVLIEVFRKSETEATHIMLTIHEKGVGTCGIYPAEV AETKINEVHTRARREGFPLRASMEEV
PS422	F, Y, W, L	45	MTQIKPQTIPDTDVISQTQSDWQMPDLYAVIMHNDDYTT MDFVVFLLNAVFDKPIEQAYQLMMQIHQTGRAVVAILPY EIAEMKVDEATSLAEQEQFPLFISIEQA
PS423	F, Y, W	46	MAPTPAGAAVLDKQQQRRHKHASRYRVLLHNDPVNTMEY VVESLRQVVPQLSEQDAIAVMVEAHNTGVGLVIVCDIEP AEFYCEQLKTKGLTSSIEPED
PS424	F, Y, W	47	MSVETIEKRSTTRKLAPQYRVLLHNDDYNSMEYVVQVLM TSVPSITQPQAVNIMMEAHNSGLALVITCAQEHAEFYCE TLKGHGLSSTIEPD
PS425	F, Y, W, L	48	MTHYFSNILRDQESPKINPKELEQIDVLEEKEHQIILYN DDVNTFEHVIDCLVKICEHNYLQAEQCAYIVHHSGKCSV KTGSLDELVPKCNALLEEGLSAEVV
PS426	F, Y, W, L	49	MSIIEKTQENVAILEKVSINHEIILYNDDVNTFDHVIET LIRVCNHEELQAEQCAILVHYTGKCAVKTGSFDELQPLC LALLDAGLSAEIT
PS427	F, W	50	MSTKEKVKERVREKEAISFNNEIIVYNDDVNTFDHVIET LIRVCNHTPEQAEQCSLIVHYNGKCTVKTGSMDKLKPQC TQLLEAGLSAEIV
PS428	F	51	MSTKEKVKERVREKEAVGFNNEIIVYNDDVNTFDHVIDT LMRVCSHTPEQAEQCSLIVHYNGKCTVKTGPMDKLKPQC TQLLEAGLSAEIV
PS429		52	MSVQEEVLEEVKTKERVNKQNQIIVFNDDVNTFDHVIDM LIATCDHDPIQAEQCTMLIHYKGKCEVKTGDYDDLKPRC SKLLDAGISAEIQ
PS430	F, Y, W, L	53	MQPFEETYTDVLDEVVDTDVHNLVVFNDDVNTFDHVIET LIDVCKHTPEQAEQCTLLIHYKGKCSVKNGSWEELVPMR NEICRRGISAEVLK
PS431		54	MIISSVKSSPSTETLSRTELQLGGVWRVVVLNDPVNLMS YVMMIFKKIFGFNETVARRHMLEVHEKGRSVVWSGLREK AEAYVFTLQQWHLTAVLESDETH
PS432	F, W	55	MIGVEARTSSAPELAIETEIRLAGLWHVIVINDPVNLMS YVVMVLRKIFGFDDTKARKHMLEVHENGRSIVWSGEREP AEAYANTLHQWHLSAVLERDETD
PS433	F, Y, W, L	56	MMSSLKECSIQALPSLDEKTKTEEDLSVPWKVIVLNDPV NLMSYVVMVFRKVFGYNENKATKHMMEVHQLGKSVLWTG QREEAECYAYQLQRWRLQTILEKDD
PS434	F, Y, W, L	57	MSRLPWKQEAKFAATVIIDFPDATLEAPTIEKKEATEQQ IEMPWNVVVHNDPVNLMSYVTMVFQRVFGYPRERAEKHM LEVHHSGRSILWSGLRERAELYVQQLHGYLLLATIEKTV
PS435	F, Y, W, L	58	MTLSVALGPDTQESTQTGTAVSTDTLTAPDIPWNLVIWN DPVNLMSYVSYVFQSYFGYSETKANKLMMEVHKKGRSIV

【００９６】

【表１－５】

			AHGSKEQVEQHAVAMHGYGLWATVEKATGGNSGGGKSGG PGKGKGKRG
Planctomycetales bacterium (PS545)	I, L, V	59	MSEPMTLPAIPQPRLKERTQRQPPYNVILLNDDDHSYEY VIAMLQVLFGYPREKGYQMAKEVDSTGRVILLTTTREHA ELKQEQIHAFGPDPLMARCQGSMTAVIEPAV
Planctomycetia bacterium (PS546)	I, L, V	60	MSDTITLPGRPEVERDERTRRQPPYNVILHNDDDHTFEY VIVMLNQLFGYPPEKGYEMAKEVHLNGRVIVLTTSKEHA ELKRDQIHAFGPDPFSSKDCKGSMSASIEPAY
Gemmataceae bacterium (PS547)	I, L, V	61	MGFPTDFRQSIEISTPLGSQQPRFSNASSEPALADPVLV INPRIQPRYHVILLNDDDHTYRYVIEMMLIVFGHPPEKG FLIAKEVDKAGRAICLTTSLEHAEFKQEQVHAYGADPYF GPKCKGSMTAVLEPAE
Gemmataceae bacterium (PS548)	I, L, V	62	MSDTITLPEEKTDVRTKRQPPYHVILLNDDDHTYQYVIY MLQTLFGHPPETGFKMAQEVDKTGRVIVDTTSLERAELK RDQIHAFGPDPYIERCKGSMSAMIEPSE
Planctomycetes bacterium (PS549)	I, L, V	63	MSESITTLPKKSRRLKEEEEQKTKRQPPYNVILLNDDDH TFEYVIFMLQKLFGHPPERGMQMAKEVHTTGRVIVMTTA LELAELKRDQIHAFGPDPLIDRCKGSMSATIEPAPI
Planctomycetes bacterium (PS550)	I, L, V	64	MPTFTEPEVVNDTRILPPYHVILLNDDDHTYEYVIHMLQ TLFGHPQERGFQLAVEVDKKGKAIVFTTSKEHAEFKRDQ IHAFGADPLSSKNCKGSMSAVIEPSF
Rubrobacter indicoceani (PS551)	I, L, V	65	MPSAAPAKPKTKRQSRTQGMPPYNVVLLDDDDHTYGYVI EMLNKVFGHPPEKGFELATEVDKNGRVIVMTTNLEVAEL KRDEVHAFGPDPLMPRSKGSMSAVVERAG
Fimbriiglobus ruber (PS552)	I, L, V	66	MSKTSTLPEVESESAQKLKYQPPYHVILLNDDDHSYVYV ITMLKELFGHPEQKGYQLADAVDKQGRAIVFTTTREHAE LKQEQIHAYGPDPTIPRCKGAMTAVIEPAE
Planctomycetes bacterium (PS553)	I, L, V	67	MPASASAVTEPPVSLPEAAAPRPKDRPKRQPRYHVILWN DDDHTYQYVVAMLRQLFGHPPEKGFTLAKQVDKDGRVVV LTTTKEHAELKRDQIHAFGADRLLARSKGSMSASIEPEA STG
Planctomycetia bacterium (PS554)	I, L, V	68	MSDSASATVEVQADPPADATARSQPTPARSTGSKPKRQP RYHVVLWNDDDHTYEYVIAMLRRLFGIEPEKGFRIAEEV DQSGRAVVLTTTREHAELKRDQIHAFGADRLLARSKGSM SASIEPEA
Planctomycetes bacterium RBG_16_64_12 (PS555)	I, L, V	69	MADSAQTGVAEPIQETLRRRKLRDDRRPKRQPPYHVILW NDNDHTYAYVVVMLMQLFGYPAEKGYQLASEVDTQGRAV VLTTTKEHAELKRDQIHAYGKDGLIEKCKGSMWATIEPA PGE
Blastopirellula marina (PS556)	I, L, V	70	MGDSNTSVAEPGEVTVVTTKPAPKKAKPKRQPKYHVVLW NDDDHTYEYVILMMHELFGHPVEKGFQIAKTVDADGRAI CLTTTKEHAELKRDQIHAYGKDELIARCRGSMSSTIEPE C
Planctomycetia bacterium (PS557)	I, L, V	71	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVL WNDDDHTYQYVVVMLQSLFGHPPERGYRLAKEVDTQGRV IVLTTTREHAELKRDQIHAFGYDRLLARSKGSMKASIEA EE
Planctomycetia bacterium (PS558)	I, L, V	72	MTATTADPDRTTAEKTTKKARRSGQPKRQPRYHVILWND NDHTYQYVVAMLQQLFGHPATTGLKLATEVDRTGRAVIL TTTREHAELKRDQIHAFGADRLLARSKGAMSASIEPEAE
Planctomycetaceae bacterium (PS559)	I, L, V	73	MNQAAISPNPDIKPNPSTHKKRASQRQPRYHVILWNDND HTYHYVVTMLQKLFGHPPRTGIKMATEVDKKGKVIVLTT SREHAELKRDQIHAFGADKLIRRSKGAMAASIEPES
Planctomycetes bacterium (PS560)	I, L, V	74	MTETITTPAERTQTQAEPRSDRAWLWNVVLLDDDEHTYE YVIRMLHTLFGMPVERAFRLAEEVDARGRAVVLTTHKEH AELKRDQVHAFGKDALIASCAGSMSAVLEPAECGSDDED

【００９７】

【表１－６】

Roseimaritima sp. JC651 (PS561)	I, L, V	75	MAELQTAVVEPTTRPEQDEKQSQSRPKRQPRYNVILWDD PDHSYDYVIMMLKELFGHPRQRGHQMAEEVDTTGRVICL TTTMEHAELKRDQIHAYGSDEGITRCKGSMSASIEPVPE
Rubripirellula amarantea (PS562)	I, L, V	76	MSDQQSMVAEPEVVVHTQDEKKLEKQNKRKKQPRYNVVL WDDTDHSYDYVVLMMKQLFHHPIETGFQIAKQVDKGGKA ICLTTTMEHAELKRDQIHAFGKDDLIARCTGSMSATIEP VPE
Acidobacteria bacterium (PS563)	I, L, V	77	MSSRSATAYPEVEDDTSDQLQPLYHVILLNDEDHTYDYV IEMLQKIFGFPESKAFSHAVEVDTKGTTILLTCDLEQAE RKRDLIHSYGPDWRLPRSLGSMAAVVEPAAG
Planctomycetes bacterium Poly21 (PS564)	I, L, V	78	MFEEVVSVAVAEPKTKKQSRTKPKRQPPYHVILWDDTDH TFDYVIKMMGELFRMPREKGYQLAKEVDTSGRAICMTTT LELAELKRDQIHAFGRDDASAHCKGSMSATIEPAEG
Aquisphaera sp. JC650 (PS565)	I, L, V	79	MSEFDHEHSGDTSVADPIVTTKTAPKPQKHAENETETRR QPPYNVIILNDEEHTFDYVIELLCKVFRHSLATAQELTW RIHLTGRAVVLTTHKELAELKRDQVLAYGPDPRMSVSKG PLDCFIEPAPGG
Planctomycetaceae bacterium (PS566)	I, L, V	80	MSSPSSLDDVQVSTSRAKPANETRTRKQPPYAVIVENDD HHTFLYVIEALMKVCGHAPEKGFVLAQQIHTQGKAMVWS GTLELAELKRDQLRGFGPDNYAPRPVTFPLGVTIEPLP
Planctomycetaceae bacterium (PS567)	I, L, V	81	MADYEDAGEDALEDDFDHGTVTVAPQKPEPKKQSENKRQ ANRQPRYNVLLWDSEDHTFEYVEKMLRELFGHIKKQCQI IAEQVDQEGRAVVLTTTLEHAELKRDQIHAYGKDQLEGS KGSMWSTIEAVD
Dehalococcoidia bacterium (PS568)	I, L, V	82	MTTPSLPTRETEVEERTEVEPERLYHLVLLDDDQHSYQY VIEMLASIFGYGSEKAWTLARIVDTEGRAILETASHAQC ERHQSQIHAYGADSRIPTSVGSMSAVIEEAGTPPQT
Planctomycetes bacterium (PS569)	I, L, V	83	MYSKNQIKIYCSEDDKGQTATPLLEKKPKFAPLYHVILW DDNTHTYEYVIKLLMSLFRMTFEKAYQHTLEVDKKGRTI CITTHLEKAELKQEQISNFGPDILMQNSKGPMSATIEPA N
Leptospira congkakensis (PS570)	I, L, V	84	MTGAGASQPSILEETEVRPRLSDGPWKVVLWDDDFHTYE YVIEMLMDVCQMPWEKAFQHAVEVDTRKKTIVFSGELEH AEFVHERILNYGPDPRMGSSKGSMTATLEQ
Leptospira meyeri (PS571)	I, L, V	85	MTSSGASQPSILEETERKPRLSDGPWKVVLWDDDFHTYE YVIEMLMDVCQMPWEKAFQHAVEVDTRKKTIVFFGELEH AEFVHERILNYGPDPRMGTSKGSMTATLEK
Blastopirellula marina (PS572)	I, L, V	86	MSSEELSLQTRPKRQPPFGVILHNDDLNSFDYVIDSIRK VFHYELEKCFQLTLEAHETGRSLLWTGTLEGAELKQELL LSCGPDPIMLDKGGLPLKVTLEELPQ
Leptospira fluminis (PS573)	I, L, V	87	MSQTPVIEETTVKDPVKTGGPWKVVLWDDDEHTYDYVIE MLMEVCVMTMEQAFHHAVEVDTQKKTVVYSGEFEHAEHI QELILEYGPDPRMAVSKGSMSATLEKS
Gemmata obscuriglobus (PS574)	I, L, V	88	MANATPTPDVVPEEETETRTRRQPPYAVVLHNDDTNTMD FVVTVLRKVFGYTVEKCVELMLEAHTQGKVAVWIGALEV AELKADQIKSFGPDPHVTKNGHPLGVTVEPAA
Leptospira kmetyi (PS575)	I, L, V	89	MASTQTPDLNEITEESTKSTGGPWRVVLWDDNEHTYEYV IEMLMEICTMTVEKAFLHAVQVDQEKRTVVFSGEFEHAE HVQERILTYGADPRMSNSKGSMSATLEK
Leptospira interrogans (PS576)	I, L, V	90	MASTQTPDLNEITEESTKSTGGPWRVVLWDDNEHTYEYV IEMLVEICMMTVEKAFLHAVQVDKEKRTVDFSGELEHAE HVQERILNYGADPRMSNSKGSMSATLER
Tuwongella immobilis (PS577)	I, L, V	91	MSASSSQPGTTTKPDLDIQPRLLPPFHVILENDEFHSME FVIDTLRKVLGVSIERAYQLMMTAHESGQAIIWTGPKEV AELKYEQVIGFHEKRSDGRDLGPLGCRIEPAV

【００９８】

【表１－７】

Planctomycetes bacterium (PS578)	I, L, V	92	MSGTVVESKPRNSTQLAPRWKVIVHDDPVTTFDFVLGVL RRVFAKPPGEARRITREAHDTGSALVDVLALEQAEFRRD QAHSLARAEGFPLTLTLEPAD
Agrobacterium tumifaciens ClpS1 (atClpS1)	F, W, Y, L	93	MIAEPICMQGEGDGEDGGTNRGTSVITRVKPKTKRPNLY RVLLLNDDYTPMEFVIHILERFFQKDREAATRIMLHVHQ HGVGECGVFTYEVAETKVSQVMDFARQHQHPLQCVMEKK
Agrobacterium tumifaciens ClpS2 (atClpS2)	F, W, Y	94	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPREFV TVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAE TKAKEATDLGKEAGFPLMFTTEPEE
atClpS2 熱安定性変異体	F, W, Y	95	MSDSPVDLKPKPKVKPKLERPKLYKVILLNDDYTPMEFV VEVLKRVFNMSEEQARRVMMTAHKKGKAVVGVCPRDIAE TKAKQATDLAREAGFPLMFTTEPEE
PS489	M	96	MSDSPVDLKPKPKVKPKLERLKLYKVILLNDDYTTAFFV VKVLKRVFNMSEEQARRVMMTAHKKGKAVVGVCPRDIAE TKAKQATDLAREAGFPLMFTTEPEE
PS490	M	97	MSDSPVDLKPKPKVKPKLERLKLYKVILLNDDYTTMRFV VLVLKRVFNMSEEQARRVMMTAHKKGKAVVGVCPRDIAE TKAKQATDLAREAGFPLMFTTEPEE
PS218	F, W, Y, L	98	MIAEPICMQGEGDGEDGGTNRGTSVITRVKPKTKRPNLY RVLLLNDDYTPFQFVIHILERFFQKDREAAWRITLHVHQ HGVGECGVFTYEVAETKVSQVMDFARQHQHPLQCVMEKK
atClpS2-V1	F, W, Y	99	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFV TVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAE TKAKEATDLGKEAGFPLMFTTEPEE
atClpS2 C72S	F, W, Y	100	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPREFV TVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVSERDIAE TKAKEATDLGKEAGFPLMFTTEPEE
atClpS2-V1 + C72S	F, W, Y	101	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFV TVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVSERDIAE TKAKEATDLGKEAGFPLMFTTEPEE
atClpS2 熱安定性変異体 + C72S	F, W, Y	102	MSDSPVDLKPKPKVKPKLERPKLYKVILLNDDYTPMEFV VEVLKRVFNMSEEQARRVMMTAHKKGKAVVGVSPRDIAE TKAKQATDLAREAGFPLMFTTEPEE
atClpS1 C7S	F, W, Y, L	103	MIAEPISMQGEGDGEDGGTNRGTSVITRVKPKTKRPNLY RVLLLNDDYTPMEFVIHILERFFQKDREAATRIMLHVHQ HGVGECGVFTYEVAETKVSQVMDFARQHQHPLQCVMEKK
atClpS1 C7S, C84S, C112S	F, W, Y, L	104	MIAEPISMQGEGDGEDGGTNRGTSVITRVKPKTKRPNLY RVLLLNDDYTPMEFVIHILERFFQKDREAATRIMLHVHQ HGVGESGVFTYEVAETKVSQVMDFARQHQHPLQSVMEKK
Synechococcus elongatus ClpS1	F, W, Y	105	MAVETIQKPETTTKRKIAPRYRVLLHNDDFNPMEYVVMV LMQTVPSLTQPQAVDIMMEAHTNGTGLVITCDIEPAEFY CEQLKSHGLSSSIEPDD
Synechococcus elongatus ClpS2	F, W, Y, L	106	MSPQPDESVLSILGVPRPCVKKRSRNDAFVLTVLTCSLQ AIAAPATAPGTTTTRVRQPYPHFRVIVLDDDVNTFQHVA ECLLKYIPGMTGDRAWDLTNQVHYEGAATVWSGPQEQAE LYHEQLRREGLTMAPLEAA
Thermosynechococcus elongatus ClpS	F, W, Y, L	107	MPQERQQVTRKHYPNYKVIVLNDDFNTFQHVAACLMKYI PNMTSDRAWELTNQVHYEGQAIVWVGPQEQAELYHEQLL RAGLTMAPLEPE

【００９９】

【表１－８】

Escherichia coli ClpS	F, W, Y, L	108	MGKTNDWLDFDQLAEEKVRDALKPPSMYKVILVNDDYTP MEFVIDVLQKFFSYDVERATQLMLAVHYQGKAICGVFTA EVAETKVAMVNKYARENEHPLLCTLEKA
Escherichia coli ClpS M40A	F, W, Y, L	109	MGKTNDWLDFDQLAEEKVRDALKPPSMYKVILVNDDYTP AEFVIDVLQKFFSYDVERATQLMLAVHYQGKAICGVFTA EVAETKVAMVNKYARENEHPLLCTLEKA
Plasmodium falciparum ClpS	F, W, Y, L	110	MFKDLKPFFLCIILLLLLIYKCTHSYNIKNKNCPLNFMN SCVRINNVNKNTNISFPKELQKRPSLVYSQKNFNLEKIK KLRNVIKEIKKDNIKEADEHEKKEREKETSAWKVILYND DIHNFTYVTDVIVKVVGQISKAKAHTITVEAHSTGQALI LSTWKSKAEKYCQELQQNGLTVSIIHESQLKDKQKK

*結合優先度は、既刊科学文献から推測され、且つ／又は本明細書に記載されるように単一分子及び／又はアンサンブル実験で本発明者によりさらに実証される。

**ホスホチロシンへの結合は、ペプチド末端又は内部位置で行われうる。

【０１００】

【表２－１】

表２．アミノ酸認識タンパク質の非限定的例。

名称	結合 優先度*	配列 番号	配列
Escherichia coli ロイシル／フェニルアラニル-tRNA-タンパク質トランスフェラーゼ	K, R	111	MRLVQLSRHSIAFPSPEGALREPNGLLALGGDLSPARLL MAYQRGIFPWFSPGDPILWWSPDPRAVLWPESLHISRSM KRFHKRSPYRVTMNYAFGQVIEGCASDREEGTWITRGVV EAYHRLHELGHAHSIEVWREDELVGGMYGVAQGTLFCGE SMFSRMENASKTALLVFCEEFIGHGGKLIDCQVLNDHTA SLGACEIPRRDYLNYLNQMRLGRLPNNFWVPRCLFSPQE LE
Vibrio vulnificus アスパラギン酸／グルタミン酸ロイシルトランスフェラーゼ Bpt	D, E	112	MSSDIHQIKIGLTDNHPCSYLPERKERVAVALEADMHTA DNYEVLLANGFRRSGNTIYKPHCDSCHSCQPIRISVPDI ELSRSQKRLLAKARSLSWSMKRNMDENWFDLYSRYIVAR HRNGTMYPPKKDDFAHFSRNQWLTTQFLHIYEGQRLIAV AVTDIMDHCASAFYTFFEPEHELSLGTLAVLFQLEFCQE EKKQWLYLGYQIDECPAMNYKVRFHRHQKLVNQRWQ
H. sapiens GID4	P	113	MSGSKFRGHQKSKGNSYDVEVVLQHVDTGNSYLCGYLKI KGLTEEYPTLTTFFEGEIISKKHPFLTRKWDADEDVDRK HWGKFLAFYQYAKSFNSDDFDYEELKNGDYVFMRWKEQF LVPDHTIKDISGASFAGFYYICFQKSAASIEGYYYHRSS EWYQSLNLTHV
Saccharomyces cerevisiae GID4	P	114	MINNPKVDSVAEKPKAVTSKQSEQAASPEPTPAPPVSRN QYPITFNLTSTAPFHLHDRHRYLQEQDLYKCASRDSLSS LQQLAHTPNGSTRKKYIVEDQSPYSSENPVIVTSSYNHT VCTNYLRPRMQFTGYQISGYKRYQVTVNLKTVDLPKKDC TSLSPHLSGFLSIRGLTNQHPEISTYFEAYAVNHKELGF LSSSWKDEPVLNEFKATDQTDLEHWINFPSFRQLFLMSQ KNGLNSTDDNGTTNAAKKLPPQQLPTTPSADAGNISRIF SQEKQFDNYLNERFIFMKWKEKFLVPDALLMEGVDGASY DGFYYIVHDQVTGNIQGFYYHQDAEKFQQLELVPSLKNK VESSDCSFEFA
ホスホチロシンに対する	phospho-Y	115	MMEVQLQQSGPELVKPGASVMISCRTSAYTFTENTVHWV KQSHGESLEWIGGINPYYGGSIFSPKFKGKATLTVDKSS

【０１０１】

【表２－２】

一本鎖抗体可変 フラグメント(scFV)**			STAYMELRSLTSEDSAVYYCARRAGAYYFDYWGQGTTLT VSSGGGSGGGSGGGSENVLTQSPAIMSASPGEKVTMTCR ASSSVSSSYLHWYRQKSGASPKLWIYSTSNLASGVPARF SGSGSGTSYSLTISSVEAEDAATYYCQQYSGYRTFGGGT KLEIKR
H. sapiens Fyn SH2 ドメイン**	リン酸化Y	116	MGAMDSIQAEEWYFGKLGRKDAERQLLSFGNPRGTFLIR ESETTKGAYSLSIRDWDDMKGDHVKHYKIRKLDNGGYYI TTRAQFETLQQLVQHYSERAAGLSSRLVVPSHK
H. sapiens Fyn SH2 ドメイン 三重突然変異体 スーパーバインダー**	リン酸化Y	117	MGAMDSIQAEEWYFGKLGRKDAERQLLSFGNPRGTFLIR ESETVKGAYALSIRDWDDMKGDHVKHYLIRKLDNGGYYI TTRAQFETLQQLVQHYSERAAGLSSRLVVPSHK
H. sapiens Src チロシンキナーゼ SH2 ドメイン**	リン酸化Y	118	MGAMDSIQAEEWYFGKITRRESERLLLNAENPRGTFLVR ESETTKGAYSLSVSDFDNAKGLNVKHYKIRKLDSGGFYI TSRTQFNSLQQLVAYYSKHADGLCHRLTTVCPTSK
H. sapiens Src チロシンキナーゼ SH2 ドメイン 三重突然変異体**	リン酸化Y	119	MGAMDSIQAEEWYFGKITRRESERLLLNAENPRGTFLVR ESEVTKGAYALSVSDFDNAKGLNVKHYLIRKLDSGGFYI TSRTQFNSLQQLVAYYSKHADGLCHRLTTVCPTSK
H. sapiens p62 フラグメント1-310	K, R, H, W, F, Y	120	MASLTVKAYLLGKEDAAREIRRFSFCCSPEPEAEAEAAA GPGPCERLLSRVAALFPALRPGGFQAHYRDEDGDLVAFS SDEELTMAMSYVKDDIFRIYIKEKKECRRDHRPPCAQEA PRNMVHPNVICDGCNGPVVGTRYKCSVCPDYDLCSVCEG KGLHRGHTKLAFPSPFGHLSEGFSHSRWLRKVKHGHFGW PGWEMGPPGNWSPRPPRAGEARPGPTAESASGPSEDPSV NFLKNVGESVAAALSPLGIEVDIDVEHGGKRSRLTPVSP ESSSTEEKSSSQPSSCCSDPSKPGGNVEGATQSLAEQ
H. sapiens p62 フラグメント1-180	K, R, H, W, F, Y	121	MASLTVKAYLLGKEDAAREIRRFSFCCSPEPEAEAEAAA GPGPCERLLSRVAALFPALRPGGFQAHYRDEDGDLVAFS SDEELTMAMSYVKDDIFRIYIKEKKECRRDHRPPCAQEA PRNMVHPNVICDGCNGPVVGTRYKCSVCPDYDLCSVCEG KGLHRGHTKLAFPSPFGHLSEGFSHSRWLRKVKHGHFGW PGWEMGPPGNWSPRPPRAGEARPGPTAESASGPSEDPSV NFLKNVGESVAAALSPLGIEVDIDVEHGGKRSRLTPVSP ESSSTEEKSSSQPSSCCSDPSKPGGNVEGATQSLAEQ
H. sapiens p62 フラグメント126-180	K, R, H, W, F, Y	122	MASLTVKAYLLGKEDAAREIRRFSFCCSPEPEAEAEAAA GPGPCERLLSRVAALFPALRPGGFQAHYRDEDGDLVAFS SDEELTMAMSYVKDDIFRIYIKEKKECRRDHRPPCAQEA PRNMVHPNVICDGCNGPVVGTRYKCSVCPDYDLCSVCEG KGLHRGHTKLAFPSPFGHLSEGFSHSRWLRKVKHGHFGW PGWEMGPPGNWSPRPPRAGEARPGPTAESASGPSEDPSV NFLKNVGESVAAALSPLGIEVDIDVEHGGKRSRLTPVSP ESSSTEEKSSSQPSSCCSDPSKPGGNVEGATQSLAEQ
H. sapiens p62 タンパク質	K, R, H, W, F, Y	123	MASLTVKAYLLGKEDAAREIRRFSFCCSPEPEAEAEAAA GPGPCERLLSRVAALFPALRPGGFQAHYRDEDGDLVAFS SDEELTMAMSYVKDDIFRIYIKEKKECRRDHRPPCAQEA PRNMVHPNVICDGCNGPVVGTRYKCSVCPDYDLCSVCEG KGLHRGHTKLAFPSPFGHLSEGFSHSRWLRKVKHGHFGW PGWEMGPPGNWSPRPPRAGEARPGPTAESASGPSEDPSV NFLKNVGESVAAALSPLGIEVDIDVEHGGKRSRLTPVSP

【０１０２】

【表２－３】

			ESSSTEEKSSSQPSSCCSDPSKPGGNVEGATQSLAEQMR KIALESEGRPEEQMESDNCSGGDDDWTHLSSKEVDPSTG ELQSLQMPESEGPSSLDPSQEGPTGLKEAALYPHLPPEA DPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALD TIQYSKHPPPL
Rattus norvegicus p62 タンパク質	K, R, H, W, F, Y	124	MASLTVKAYLLGKEEAAREIRRFSFCFSPEPEAEAAAGP GPCERLLSRVAVLFPALRPGGFQAHYRDEDGDLVAFSSD EELTMAMSYVKDDIFRIYIKEKKECRREHRPPCAQEARS MVHPNVICDGCNGPVVGTRYKCSVCPDYDLCSVCEGKGL HREHSKLIFPNPFGHLSDSFSHSRWLRKLKHGHFGWPGW EMGPPGNWSPRPPRAGDGRPCPTAESASAPSEDPNVNFL KNVGESVAAALSPLGIEVDIDVEHGGKRSRLTPTSAESS STGTEDKSGTQPSSCSSEVSKPDGAGEGPAQSLTEQMKK IALESVGQPEELMESDNCSGGDDDWTHLSSKEVDPSTGE LQSLQMPESEGPSSLDPSQEGPTGLKEAALYPHLPPEAD PRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDT IQYSKHPPPL
Saccharomyces cerevisiae GID10	P, M, V	125	MTSLNIMGRKFILERAKRNDNIEEIYTSAYVSLPSSTDT RLPHFKAKEEDCDVYEEGTNLVGKNAKYTYRSLGRHLDF LRPGLRFGGSQSSKYTYYTVEVKIDTVNLPLYKDSRSLD PHVTGTFTIKNLTPVLDKVVTLFEGYVINYNQFPLCSLH WPAEETLDPYMAQRESDCSHWKRFGHFGSDNWSLTERNF GQYNHESAEFMNQRYIYLKWKERFLLDDEEQENQMLDDN HHLEGASFEGFYYVCLDQLTGSVEGYYYHPACELFQKLE LVPTNCDALNTYSSGFEIA
Leishmania major N-ミリストリルトランスフェラーゼ	G	126	MSRNPSNSDAAHAFWSTQPVPQTEDETEKIVFAGPMDEP KTVADIPEEPYPIASTFEWWTPNMEAADDIHAIYELLRD NYVEDDDSMFRFNYSEEFLQWALCPPNYIPDWHVAVRRK ADKKLLAFIAGVPVTLRMGTPKYMKVKAQEKGEGEEAAK YDEPRHICEINFLCVHKQLREKRLAPILIKEATRRVNRT NVWQAVYTAGVLLPTPYASGQYFHRSLNPEKLVEIRFSG IPAQYQKFQNPMAMLKRNYQLPSAPKNSGLREMKPSDVP QVRRILMNYLDSFDVGPVFSDAEISHYLLPRDGVVFTYV VENDKKVTDFFSFYRIPSTVIGNSNYNLLNAAYVHYYAA TSIPLHQLILDLLIVAHSRGFDVCNMVEILDNRSFVEQL KFGAGDGHLRYYFYNWAYPKIKPSQVALVML
H. sapiens N-ミリストリルトランスフェラーゼNMT1	G	127	MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDN SYNRGGLSPANDTGAKKKKKKQKKKKEKGSETDSAQDQP VKMNSLPAERIQEIQKAIELFSVGQGPAKTMEEASKRSY QFWDTQPVPKLGEVVNTHGPVEPDKDNIRQEPYTLPQGF TWDALDLGDRGVLKELYTLLNENYVEDDDNMFRFDYSPE FLLWALRPPGWLPQWHCGVRVVSSRKLVGFISAIPANIH IYDTEKKMVEINFLCVHKKLRSKRVAPVLIREITRRVHL EGIFQAVYTAGVVLPKPVGTCRYWHRSLNPRKLIEVKFS HLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVVHQ LLTRYLKQFHLTPVMSQEEVEHWFYPQENIIDTFVVENA NGEVTDFLSFYTLPSTIMNHPTHKSLKAAYSFYNVHTQT PLLDLMSDALVLAKMKGFDVFNALDLMENKTFLEKLKFG IGDGNLQYYLYNWKCPSMGAEKVGLVLQ

【０１０３】

【表２－４】

Drosophila melanogaster BIR2	A	128	MGDVQPETCRPSAASGNYFPQYPEYAIETARLRTFEAWP RNLKQKPHQLAEAGFFYTGVGDRVRCFSCGGGLMDWNDN DEPWEQHALWLSQCRFVKLMKGQLYIDTVAAKPVLAEEK EESTSIGGDT
Amanita thiersii Skay4041 UBRボックスドメイン (PS501)	K, R, H	129	MICGQIIGKGESCFRCRDCGLDESCVMCSQCFHATDHIN HNVSFFVSQQPGGCCDCGDEEAWKKPMNCPYHPP
Helobdella robusta UBRボックスドメイン (PS502)	K, R, H	130	MVCLKVFKLGEPTYSCRSVTCGMDPTCVLCVDCFQNSSH KLHKYKMSTSGGGGYCDCGDLEAWKADPLCDLHKL
Hydra vulgaris UBRボックスドメイン (PS503)	K, R, H	131	MFCGRLFKVGDPTYTCKDCAADPTCVFCHDCFHQSVHTK HKYKLFASQGRGGYCDCGDKEAWTNDPACNKHKE
Galleria mellonella UBRボックスドメイン (PS504)	K, R, H	132	MLCGKVFKQGEPAYSCRECGMDNTCVLCVECFKVSPHRH HKYKMGQSGGGGCCDCGDTEAWKRDPFCERHAK
Brachionus plicatilis UBRボックスドメイン (PS505)	K, R, H	133	MVCGRVFKSGEPSYFCRECGTDPTCVLCSICFRHSKHRY HKYVMMTSGGGGYCDCGDPEAWKSDPCCELHMP
Capitella teleta UBRボックスドメイン (PS506)	K, R, H	134	MLCGKVFKMGELTYSCRDCGTDPTCVLCMDCFQHSAHKK HRYKMAASGGGGYCDCGDREAWKAEPFCDVHKR
Sparassis crispa UBRボックスドメイン (PS507)	K, R, H	135	MPCGHIFKKGESCFRCKDCALDDSCVLCSKCFEATDHAN HNVSFFIAQQSGGCCDCGDIEAWLVPIDCPFHPV
Anabarilius graham UBRボックスドメイン (PS508)	K, R, H	136	MLCGRVFKEGETVYSCRDCAIDPTCVLCIECFQKSVHKS HRYKMHASAGGGFCDCGDLEAWKTGPCCSQHDP
Lottia gigantean UBRボックスドメイン (PS509)	K, R, H	137	MICGHGFKTGEPTYSCRDCATDPTCVLCISCFQKSPHRE HRYKMSASGGGGYCDCGDPEAWKIEPFCEQHKP
Camponotus floridanus UBRボックスドメイン (PS510)	K, R, H	138	MICGRMFKMGEPTYSCRQCGMDSTCVLCVDCFKQSAHRN HKYKMGTSSGGGCCDCGDTEAWKNEPFCKIHLA
Habropoda laboriosa UBRボックスドメイン (PS511)	K, R, H	139	MICGKVFKMGEATYSCKECGVDPTCVLCADCFKQSAHRH HKYRMGTSSGGGFCDCGDIEAWKKEPFCNTHLA
Mastacembelus armatus UBRボックスドメイン (PS512)	K, R, H	140	MLCGRVFKEGETVYSCRDCAIDPTCVLCMDCFQDSVHKS HRYKMHASAGGGFCDCGDVEAWKIGPYCSKHDP
Pyrenophora seminiperda CCB06 UBRボックスドメイン (PS513)	K, R, H	141	MPCGHIFKNGEATYRCKTCTADDTCVLCARCFDASDHEG HQVFVSVSPGNSGCCDCGDDEAWVRPVHCNIHSA
Tribolium castaneum UBRボックスドメイン (PS514)	K, R, H	142	MVCGRVFKLGEPTYSCRECGMDNTCVLCVNCFKNSEHRF HKYKMGTSQGGGCCDCGDVEAWKKAPFCDVHIA

【０１０４】

【表２－５】

Wasmannia auropunctata UBRボックスドメイン (PS515)	K, R, H	143	MICGKMFKIGEPTYSCRECGMDSTCVLCVDCFKQSAHRN HKYKMGTSSGGGCCDCGDTEAWKKEPFCKTHVV
Crassostrea gigas UBRボックスドメイン (PS516)	K, R, H	144	MLCGKVFKTGEPTYSCRDCANDPTCVLCIDCFQNGAHKN HRYKMNTSGGGGYCDCGDQEAWTSHPFCNLHSP
Harpegnathos saltator UBRボックスドメイン (PS517)	K, R, H	145	MMCGRVFKMGEPTYSCRECGVDSTCVLCVGCFQQSAHRD HKYKMGTSGGGGCCDCGDTEAWKRDPFCEIHMV
Nilaparvata lugens UBRボックスドメイン (PS518)	K, R, H	146	MVCGRVFKMGEPSYHCRECGMDATCVLCVDCFKKSSHRN HKYKMGTSIGGGCCDCGDVEAWKTEPYCEVHIA
Manduca sexta UBRボックスドメイン (PS519)	K, R, H	147	MLCGRVFKQGEPAYSCRECGMDNTCVLCVECFKVSAHRH HKYKMGQSGGGGCCDCGDTEAWKRDPFCELHAA
Monopterus albus UBRボックスドメイン (PS520)	K, R, H	148	MLCGRVFKEGETVYSCRDCAIDPTCVLCMDCFQDSVHKS HRYKMHASSGGGFCDCGDVEAWKIGPCCSKHDP
Lingula anatine UBRボックスドメイン (PS521)	K, R, H	149	MLCGRVFRSGEPTYSCRDCAVDPTCVLCIDCFNNGAHRK HKYRMSTSSGGGYCDCGDKEAWKTDPLCEIHRK
Vombatus ursinus UBRボックスドメイン (PS522)	K, R, H	150	MLCGKVFKSGETTYSCRDCAIDPTCVLCMNCFQSSVHKN HRYKMHTSTGGGFCDCGDTEAWKTGPFCTIHEP
Saccharomycetaceae sp. Ashbya aceri UBRボックスドメイン (PS523)	K, R, H	151	MAKSHRHTGRNCGRAFQPGEPLYRCQECAYDDTCVLCIS CFNPDDHVNHHVSTHICNELHDGICDCGDAEAWNVPLHC KAEED
Drosophila ficusphila UBRボックスドメイン (PS524)	K, R, H	152	MVCGKVFKNGEPTYSCRECGVDPTCVLCVNCFKRSAHRF HKYKMSTSGGGGCCDCGDDEAWKKDHYCQLHLA
Mus musculus UBRボックスドメイン (PS525)	K, R, H	153	MLCGKVFKSGETTYSCRDCAIDPTCVLCMDCFQSSVHKN HRYKMHTSTGGGFCDCGDTEAWKTGPFCVDHEP
Maylandia zebra UBRボックスドメイン (PS526)	K, R, H	154	MLCGRVFKEGETVYSCRDCAIDPTCVLCMDCFQDSVHKS HRYKMHASSGGGFCDCGDVEAWKIGPYCSKHDP
Mizuhopecten yessoensis UBRボックスドメイン (PS527)	K, R, H	155	MLCGKVFKYGEPTYSCRDCANDPTCVLCIDCFQKSAHKK HRYKMSTSGGGGYCDCGDSEAWKTAPFCSNHKA
Kluyveromyces lactis UBRボックスドメイン (PS528)	K, R, H	156	MHSKFNHAGRICGAKFRVGEPIYRCKECSFDDTCVLCVN CFNPKDHVGHHVYTSICTEFNNGICDCGDKEAWNHELNC KGAED
Chelonia mydas UBRボックスドメイン (PS529)	K, R, H	157	MLCGKVFKGGETTYSCRDCAIDPTCVLCMDCFQNSIHKN HRYKMHTSTGGGFCDCGDTEAWKTGPLCANHEP
Acropora millepora UBRボックスドメイン (PS530)	K, R, H	158	MLCGKVFKVGEPTYSCRDCGYDNTCVLCINCFQKSIHKN HHYKMNTSGGGGVCDCGDVEAWKEGEACEIHQQ
Musca domestica UBRボックスドメイン	K, R, H	159	MVCGKVFKIGEPTYSCRECGMDQTCVLCVNCFKQSAHRY HKYKMSTSGGGGCCDCGDEEAWKKDHYCEEHLR

【０１０５】

【表２－６】

(PS531)			
Schizosaccharomyces cryophilus OY26 UBRボックスドメイン (PS532)	K, R, H	160	MSCGRIFKKGEVFYRCKTCSVDSNSALCVKCFRATDHHG HETSFTISAGSGGCCDCGNSAAWIRDMPCKIHNR
Contarinia nasturtii UBRボックスドメイン (PS533)	K, R, H	161	MVCGRVFKMNEPFYSCRECGMDPTCVLCVNCFKQSAHRH HKYKMGTSAGGGCCDCGDNEAWKQDHYCDEHTK
Schizosaccharomyces pombe UBRボックスドメイン (PS534)	K, R, H	162	MKCGHIFRKGEVFYRCKTCSVDSNSALCVKCFRATSHKD HETSFTVSAGSGGCCDCGNAAAWIGDVSCKIHSH
Mus musculus UBRボックスドメイン (PS535)	K, R, H	163	MLCGRVFKVGEPTYSCRDCAVDPTCVLCMECFLGSIHRD HRYRMTTSGGGGFCDCGDTEAWKEGPYCQKHKL
Aphis gossypii UBRボックスドメイン (PS536)	K, R, H	164	MVCGRVFKMGEPTYNCRECGMDSTCVLCVDCFKRSPHKN HKYKMGTSYGGGCCDCGDVEAWKHDPYCQTHKL
Aedes aegypti UBRボックスドメイン (PS537)	K, R, H	165	MVCGRVFKIGEPTYSCRECSMDPTCVLCSSCFKKSSHRL HKYKMSTSGGGGCCDCGDHEAWKRDPSCEEHAV
Saccharomyces cerevisiae UBRボックスドメイン (PS538)	K, R, H	166	MGDVHKHTGRNCGRKFKIGEPLYRCHECGCDDTCVLCIH CFNPKDHVNHHVCTDICTEFTSGICDCGDEEAWNSPLHC KAEEQ
Saccharomyces cerevisiae UBR1 D3S 変異体 (PS25)	K, R, H	167	MGSVHKHTGRNCGRKFKIGEPLYRCHECGCDDTCVLCIH CFNPKDHVNHHVCTDICTEFTSGICDCGDEEAWNSPLHC KAEEQ
Kazachstania africana CBS 2517 UBRボックスドメイン (PS539)	K, R, H	168	MQTSFTHKGRNCGRKFKVGEPLYRCHECGFDDTCVLCIH CFNPADHENHHIYTDICNDFTSGICDCGDTEAWNGDLHC KAEEI
Clathrospora elynae UBRボックスドメイン (PS540)	K, R, H	169	MPCGHIFKNGEATYRCKTCTADDTCVLCARCFDASDHEG HQVFVSVSPGNSGCCDCGDDEAWVRPVHCNMHSA
Aspergillus neoniger CBS 115656 UBRボックスドメイン (PS541)	K, R, H	170	MRCGHIFRAGEATYRCITCAADDTCVLCSRCFDASDHTG HQYQISLSSGNCGCCDCGDEEAWRLPLFCAIHTD
Trichuris suis UBRボックスドメイン (PS542)	K, R, H	171	MRCNHVFANGEATYSCRGCAADPTCVLCASCFELSAHKE HKYMITTSSGTGYCDCGDPEAWKADPFCQQHQP
Trichinella spiralis UBRボックスドメイン (PS543)	K, R, H	172	MKCNRQLICGEPTYCCLDCACDQTCIFCHACFQSSEHKN HRYSMSTSEGSGTCDCGDKEAWKSNYYCLNHKP
Homo sapiens UBR1 (PS544)	K, R, H	173	MGPLGSLCGRVFKSGETTYSCRDCAIDPTCVLCMDCFQD SVHKNHRYKMHTSTGGGFCDCGDTEAWKTGPFCVNHEP
Homo sapiens UBR2	K, R, H	174	MGPLGSLCGRVFKVGEPTYSCRDCAVDPTCVLCMECFLG SIHRDHRYRMTTSGGGGFCDCGDTEAWKEGPYCQKHE

【０１０６】

【表２－７】

Kluyveromyces marxianus UBR2 (PS615)	K, R, H	175	MVNEHRGSQCSKQCHGTETVYYCFDCTKNPLYEICEECF DETQHMGHRYTSRVVTRPEGKVCHCGDISGYNNPEKAFQ CKI
Kluyveromyces lactis UBR2 (PS616)	K, R, H	176	MHNDHRGSQCSKQCHGTETVYYCFDCTKNPLYEICEDCF DESQHIGHRYTSRVVTRPEGKVCHCGDISSYNDPKKAFQ CRI
Eremothecium sinecaudum UBR2 (PS617)	K, R, H	177	MPKEHRGTSCNKHCQPTETVYYCFDCTKNPLYEICEECF DADKHLGHRWTSKVVSRPEGKICHCGDPSGLTDPENGYE CKN
Zygosaccharomyces bailii UBR2 (PS618)	K, R, H	178	MNASHKGAMCSKQCYPTETVFYCFTCTTNPLYEICESCF DEEKHRGHLYTAKVVVRPEGRVCHCGDPFVFKEPRFAFL CKN
Vanderwaltozyma polyspora UBR2 (PS619)	K, R, H	179	MENLHIGSCCNRQCYPTQTVYYCLTCTINPLYEICELCF DEDKHVGHTYISKSVIRPEGKVCHCGNPNVFKKPEFAFN CKN
Saccharomyces cerevisiae UBR2 (PS620)	K, R, H	180	MGNMHIGTACTRLCFPSETIYYCFTCSTNPLYEICELCF DKEKHVNHSYVAKVVMRPEGRICHCGDPFAFNDPSDAFK CKN
Kluyveromyces marxianus UBR1 (PS621)	K, R, H	181	MHSKFSHAGRICGAKFKVGEPIYRCKECSFDDTCVLCVN CFNPKDHTGHHVYTTICTEFNNGICDCGDKEAWNHTLFC KAEEG
Kluyveromyces dobzhanskii UBR1 (PS622)	K, R, H	182	MHSRFNHAGRICASKFKVGEPIYRCKECSFDDTCVICVN CFNPKDHVGHHVYTSICSEFNNGICDCGDTEAWNHDMHC KADEN
Kazachstania naganishii UBR1 (PS623)	K, R, H	183	MSKQFRHKGRNCGRKFRLGEPLYRCQECGYDDTCVLCIN CFNPKDHEGHHIYTDICNDFTSGICDCGDEEAWLSPLHC KAEED
Eremothecium sinecaudum UBR1 (PS624)	K, R, H	184	MPKNHNHKGRNCGRSFQPGEPLYRCQECAYDDTCVLCIR CFNPLDHVNHHVSTHICSEFNDGICDCGDVEAWNVELNC KAEED
Saccharomyces eubayanus UBR1 (PS625)	K, R, H	185	MGDVHKHTGRNCGRKFKIGEPLYRCHECGCDDTCVLCIH CFNPKDHINHHVCTDICSEFTSGICDCGDEEAWNSSLHC KAEEQ
Zygosaccharomyces parabailii UBR1 (PS626)	K, R, H	186	MYHVYKHSGRNCGRKFKVGEPIYRCHECGYDETCVLCIH CFNPKDHDSHHVYIDICSEFSTGICDCGDTEAFVNPLHC KAEED
Zygosaccharomyces mellis UBR1 (PS627)	K, R, H	187	MPKYHQHSGRYCGRKFKVGEPIYRCHECGFDETCVICIH CFNAKDHETHHVSVSICSEYSTGICDCGDTEAFVNPLHC RAEEV
Candida albicans UBR1 (PS628)	K, R, H	188	MSHRAYHKNSPCGRIFRKGEPIHRCLTCGFDDTCALCSH CFQPEYHEGHKVHIGICQRENGGVCDCGDPEAWTQELFC PYAVD
Pichia pastoris UBR1 (PS629)	K, R, H	189	MCPNYKHHGRPCARQFKQGEPIYRCYECGFDETCVMCMH CFNREQHRDHEVSISIASSSNDGICDCGDPQAWNIELHC QSELD

*結合優先度は、既刊科学文献から推測され、且つ／又は本明細書に記載されるように単一分子及び／又はアンサンブル実験で本発明者によりさらに実証される。
**ホスホチロシンへの結合は、ペプチド末端又は内部位置で行われうる。

【０１０７】

それゆえ、いくつかの実施形態では、本願は、表１又は表２から選択されるアミノ酸配列を有する（又は表１又は表２から選択されるアミノ酸配列に対して少なくとも５０％、少なくとも６０％、少なくとも７０％、少なくとも８０％、８０～９０％、９０～９５％、９５～９９％、若しくはそれ以上のアミノ酸配列同一性を有するアミノ酸配列を有する）アミノ酸認識分子を提供する。いくつかの実施形態では、アミノ酸認識分子は、表１又は表２に列挙されるアミノ酸認識分子に対して２５～５０％、５０～６０％、６０～７０％、７０～８０％、８０～９０％、９０～９５％、若しくは９５～９９％、又はそれ以上のアミノ酸配列同一性を有する。いくつかの実施形態では、アミノ酸認識分子は、修飾ア

ミノ酸認識分子であり、且つ表１又は表２に示される配列に対して１つ以上のアミノ酸の欠失、付加、又は突然変異を含む。いくつかの実施形態では、修飾アミノ酸認識分子は、表１又は表２に示される配列に対して、１、２、３、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、３６、３７、３８、３９、４０、４１、４２、４３、４４、４５、４６、４７、４８、４９、５０、又はそれ以上のアミノ酸（連続したアミノ酸でもそうでなくてもよい）の欠失、付加、又は突然変異を含む。

【０１０８】

いくつかの実施形態では、アミノ酸認識分子は、２つ以上のアミノ酸結合タンパク質（たとえば、２つ以上のバインダー）の直列コピーを有する単一ポリペプチドを含む。本明細書で用いられる場合、いくつかの実施形態では、分子内のエレメントの直列配置又は配向は、エレメントが連続して融合されるように、各エレメントが線形融合内のその次のエレメントに末端同士で連結していることを意味する。たとえば、いくつかの実施形態では、２つのバインダーの直列コピーを有するポリペプチドは、一方のバインダーのＣ末端が他方のバインダーのＮ末端に融合された融合ポリペプチドを意味する。同様に、２つ以上のバインダーの直列コピーを有するポリペプチドは、第１のバインダーのＣ末端が第２のバインダーのＮ末端に融合され、第２のバインダーのＣ末端が第３のバインダーのＮ末端に融合された融合ポリペプチドなどを意味する。かかる融合ポリペプチドは、同じバインダーの複数コピー又は互いに異なるバインダーの複数コピーを含みうる。いくつかの実施形態では、本願の融合ポリペプチドは、少なくとも２つ且つ１０までのバインダー（たとえば、少なくとも２つのバインダー且つ８つ、６つ、５つ、４つ、又は３つまでのバインダー）を有する。いくつかの実施形態では、本願の融合ポリペプチドは、５つ以下のバインダー（たとえば、２つ、３つ、４つ、又は５つのバインダー）を有する。それゆえ、いくつかの実施形態では、標識付認識分子２００は、本願の融合ポリペプチドを含む。

【０１０９】

いくつかの実施形態では、単一コーディング配列の発現により、２つ以上の独立結合部位を有する単一全長ポチペプチドが生成される場合、融合ポリペプチドは、フレキシブルリンカーをコードするセグメントによって分離されるモノマー形バインダーサブユニットをコードするセグメントを含む単一コーディング配列の発現によって提供される。いくつかの実施形態では、１つ以上のモノマー形サブユニット（たとえば、バインダー）は、ＣｌｐＳタンパク質である。いくつかの実施形態では、ＣｌｐＳサブユニットは、同一でも非同一でもよい。非同一の場合、ＣｌｐＳサブユニットは、同親ＣｌｐＳタンパク質の識別可能な変異体であってもよく、又はそれらは、互いに異なる親ＣｌｐＳに由来してもよい。いくつかの実施形態では、融合ポリペプチドは、１つ以上のＣｌｐＳモノマー及び１つ以上の非ＣｌｐＳモノマーを含む。いくつかの実施形態では、モノマー形サブユニットは、非ＣｌｐＳモノマーを含む。いくつかの実施形態では、モノマー形サブユニットは、１つ以上の分解経路タンパク質を含む。たとえば、いくつかの実施形態では、モノマー形サブユニットは、Ｇｉｄタンパク質、ＵＢＲボックスタンパク質若しくはそのＵＢＲボックスドメイン含有タンパク質断片、ｐ６２タンパク質若しくはそのＺＺドメイン含有断片、及びＣｌｐＳタンパク質（たとえば、ＣｌｐＳ１、ＣｌｐＳ２）のうちの１つ以上を含む。

【０１１０】

いくつかの実施形態では、融合ポリペプチドの少なくとも１つのバインダーは、表１又は表２から選択されるアミノ酸配列を有する（又は表１又は表２から選択されるアミノ酸配列に対して少なくとも５０％、少なくとも６０％、少なくとも７０％、少なくとも８０％、８０～９０％、９０～９５％、９５～９９％、若しくはそれ以上のアミノ酸配列同一性を有するアミノ酸配列を有する）。いくつかの実施形態では、融合ポリペプチドの各バインダーは、表１又は表２から選択されるアミノ酸配列に少なくとも８０％（たとえば、８０～９０％、９０～９５％、９５～９９％、又はそれ以上）のアミノ酸配列同一性を有

する（又は表１又は表２から選択されるアミノ酸配列に対して少なくとも５０％、少なくとも６０％、少なくとも７０％、少なくとも８０％、８０～９０％、９０～９５％、９５～９９％、若しくはそれ以上のアミノ酸配列同一性を有するアミノ酸配列を有する）。いくつかの実施形態では、融合ポリペプチドのバインダーは、修飾され、表１又は表２に示される配列に対して１つ以上のアミノ酸欠失、付加、又は突然変異を含む。いくつかの実施形態では、融合ポリペプチドのバインダーは、表１又は表２に示される配列に対して、１、２、３、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、３６、３７、３８、３９、４０、４１、４２、４３、４４、４５、４６、４７、４８、４９、５０、又はそれ以上のアミノ酸（連続したアミノ酸でもそうでなくてもよい）の欠失、付加、又は突然変異を含む。

【０１１１】

いくつかの実施形態では、融合ポリペプチドのバインダーは、１つ以上のアミノ酸の同一セットを認識する。いくつかの実施形態では、融合ポリペプチドのバインダーは、１つ以上のアミノ酸の識別可能なセットを認識する。いくつかの実施形態では、融合ポリペプチドのバインダーは、アミノ酸の重複するセットを認識する。いくつかの実施形態では、融合ポリペプチドのバインダーが同じアミノ酸を認識する場合、それらは、同じ特性パルスパターンを有するアミノ酸又は互いに異なる特性パルスパターンを有するアミノ酸を認識しうる。

【０１１２】

いくつかの実施形態では、融合ポリペプチドのバインダーは、共有結合、又は一方のバインダーのＣ末端ともう一方のバインダーのＮ末端とを共有結合的に連結するリンカーのいずれかによって、末端同士で連結される。本願の融合ポリペプチドとの関連で、リンカーは、２つのバインダーを連結するが２つのバインダーのいずれに対応するポリペプチド配列の部分も形成しない、融合ポリペプチド内の１つ以上のアミノ酸を意味する。いくつかの実施形態では、リンカーは、少なくとも２つのアミノ酸（たとえば、少なくとも２つ、３つ、４つ、５つ、６つ、８つ、１０、１５、２５、５０、１００、又はそれ以上のアミノ酸）を含む。いくつかの実施形態では、リンカーは、５つまで、１０まで、１５まで、２５まで、５０まで、又は１００までのアミノ酸を含む。いくつかの実施形態では、リンカーは、約２つ～約２００のアミノ酸（たとえば、約２つ～約１００、約５つ～約５０、約２つ～約２０、約５つ～約２０、又は約２つ～約３０のアミノ酸）を含む。

【０１１３】

いくつかの実施形態では、本願は、２つ以上のアミノ酸結合タンパク質の直列コピーを有する単一ポリペプチドをコードする核酸を提供する。いくつかの実施形態では、核酸は、本願の融合ポリペプチドをコードする発現構築物である。いくつかの実施形態では、発現構築物は、少なくとも２つ且つ１０までのバインダー（たとえば、少なくとも２つのバインダー且つ８つ、６つ、５つ、４つ、又は３つまでのバインダー）を有する融合ポリペプチドをコードする。いくつかの実施形態では、発現構築物は、５つ以下のバインダー（たとえば、２つ、３つ、４つ、又は５つのバインダー）を有する融合ポリペプチドをコードする。

【０１１４】

いくつかの実施形態では、アミノ酸認識分子は、１つ以上の標識を含む。いくつかの態様では、１つ以上の標識は、明細書の他の箇所に記載されるように発光標識又は伝導率標識を含む。いくつかの実施形態では、１つ以上の標識は、１つ以上のポリオール部分（たとえば、デキストラン、ポリビニルピロリドン、ポリエチレングリコール、ポリプロピレングリコール、ポリオキシエチレングリコール、及びポリビニルアルコールから選択された１つ以上の部分）を含む。たとえば、いくつかの実施形態では、アミノ酸認識分子は、ＰＥＧ化される（ＰＥＧｙｌａｔｅｄ）。いくつかの実施形態では、ポリオール修飾（たとえば、ＰＥＧ化）は、基体（たとえば、シーケンシングチップ）表面への非特異的な固着を制限しうる。いくつかの実施形態では、ポリオール修飾は、アミノ酸認識分子とその

他の認識分子との間、アミノ酸認識分子と切断試薬との間、又はアミノ酸認識分子とシーケンシング反応混合物中に存在するその他の種との間のアグリゲーション又は相互作用を制限しうる。ＰＥＧ化は、認識分子（たとえば、ＣｌｐＳタンパク質などのアミノ酸結合タンパク質）をｍＰＥＧ４－ＮＨＳエステルとインキュベートすることによって実施可能であり、これによって表面露出リシン側鎖などの一級アミンが標識される。他のタイプのＰＥＧ、及びポリオール修飾の他の方法は、当分野で知られている。

【０１１５】

いくつかの実施形態では、１つ以上標識は、タグ配列を含む。たとえば、いくつかの実施形態では、アミノ酸認識分子は、アミノ酸結合以外の１つ以上の機能を提供するタグ配列を含む。いくつかの実施形態では、タグ配列は、認識分子のビオチン化（たとえば、ビオチン部分及びビスビオチン部分をはじめとする１つ以上のビオチン分子の取込み）を可能にする少なくとも１つのビオチンリガーゼ認識配列を含む。いくつかの実施形態では、タグ配列は、直列に配向された２つのビオチンリガーゼ認識配列を含む。いくつかの実施形態では、ビオチンリガーゼ認識配列は、配列とビオチン分子との間の共有結合を触媒するビオチンリガーゼによって認識されるアミノ酸配列を意味する。タグ配列の各ビオチンリガーゼ認識配列は、ビオチン部分に共有結合されうるため、複数のビオチンリガーゼ認識配列を有するタグ配列は、複数のビオチン分子に共有結合されうる。１つ以上のビオチンリガーゼ認識配列を有するタグ配列の領域は、ビオチン化タグ又はビオチン化配列と一般的にいいうる。いくつかの実施形態では、ビスビオチン又はビスビオチン部分は、直列に配向された２つのビオチンリガーゼ認識配列に結合する２つのビオチンを意味しうる。

【０１１６】

タグ配列中の機能配列の追加の例としては、認識分子の精製及び／又は修飾に有用な精製タグ、切断部位、及び他の部分が挙げられる。表３は、タグ配列の非限定的配列のリストを提供し、そのうちのいずれか１つ以上は、本願のアミノ酸認識分子のいずれか１つとの組合せで（たとえば、表１に示される配列との組合せで）使用されうる。表３に示されるタグ配列は、非限定的であることが意図され、且つ本願に係る認識分子は、認識分子ポリペプチド或いは内部位置のＮ末端又はＣ末端にタグ配列（たとえば、Ｈｉｓタグ及び／又はビオチン化タグ）のいずれか１つ以上を含みうるか、Ｎ末端とＣ末端との間がスプリットされうるか、さもなければ当技術分野で実用されるように他の形で再構成されうることが認識されるべきである。

【０１１７】

【表３】

表３．タグ配列の非限定的例。

名称	配列番号	配列
ビオチン化タグ	190	GGGSGGGSGGGSGLNDFFEAQKIEWHE
ビスビオチン化タグ	191	GGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
ビスビオチン化タグ	192	GSGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
His／ビオチン化タグ	193	GHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHE
His／ビスビオチン化タグ	194	GHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
His／ビスビオチン化タグ	195	GGSHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
His／ビスビオチン化タグ	196	GSHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
ビスビオチン化／Hisタグ	197	GGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHEGHHHHHH

【０１１８】
いくつかの実施形態では、本願の認識分子は、シーケンシング方法のペプチダーゼ及び／又は核酸アプトマーなどの他のタイプのアミノ酸結合分子と共に使用されうるアミノ酸結合タンパク質である。プロテアーゼ又はプロテイナーゼともいわれるペプチダーゼは、ペプチド結合の加水分解を触媒する酵素である。ペプチダーゼは、ポリペプチドをより短い断片に消化するとともに、内部又は末端でそれぞれポリペプチド鎖を切断するエンドペプチダーゼ及びエキソペプチダーゼに一般に分類されうる。いくつかの実施形態では、標識付認識分子２００は、エキソペプチダーゼ活性又はエンドペプチダーゼ活性を不活性化するように修飾されたペプチダーゼを含む。こうして、標識付認識分子２００は、ポリペプチドからアミノ酸を切断することもなく選択的に結合する。さらに他の実施形態では、本願のアミノ酸結合タンパク質とともに、エキソペプチダーゼ又はエンドペプチダーゼ活性を不活性化するように修飾されていないペプチダーゼが使用されうる。たとえば、いくつかの実施形態では、標識付認識分子は標識付エキソペプチダーゼ２０２を含む。

【０１１９】
本願のある特定の実施形態によれば、タンパク質シーケンシング方法は、ポリペプチドの末端の繰返し検出及び切断を含みうる。いくつかの実施形態では、アミノ酸の検出及び切断の両工程を実施する単一試薬として標識付エキソペプチダーゼ２０２が使用されうる。一般的に描かれるように、いくつかの実施形態では、標識付エキソペプチダーゼ２０２は、ポリペプチドのＮ末端又はＣ末端のアミノ酸にそれぞれ選択的に結合してそれを切断するようにアミノペプチダーゼ活性又はカルボキシペプチダーゼ活性を有する。ある特定の実施形態では、標識付エキソペプチダーゼ２０２は、本明細書に記載されるように標識付エキソペプチダーゼ２０２を非切断標識付認識分子２００として使用すべく選択的結合性を保持するように当業者により触媒的に不活性化されうることが認識されるべきである。

【０１２０】
エキソペプチダーゼは、一般に、そのアミノ末端の遊離アミノ基又はそのカルボキシ末端の遊離カルボキシル基の少なくとも１つを含むポリペプチド基質を必要とする。いくつかの実施形態では、本願に係るエキソペプチダーゼは、ポリペプチドの末端又はその近くの結合を加水分解する。いくつかの実施形態では、エキソペプチダーゼは、ポリペプチド末端から３残基以下の結合を加水分解する。たとえば、いくつかの実施形態では、エキソペプチダーゼにより触媒される単一加水分解反応は、ポリペプチド末端から単一アミノ酸、ジペプチド、又はトリペプチドを切断する。

【０１２１】
いくつかの実施形態では、本願に係るエキソペプチダーゼは、アミノ末端又はカルボキシ末端からそれぞれ単一アミノ酸を切断するアミノペプチダーゼ又はカルボキシペプチダーゼである。いくつかの実施形態では、本願に係るエキソペプチダーゼは、アミノ末端又はカルボキシ末端からそれぞれジペプチドを切断するジペプチジルペプチダーゼ又はペプチジルジペプチダーゼである。さらに他の実施形態では、本願に係るエキソペプチダーゼは、アミノ末端からトリペプチドを切断するトリペプチジルペプチダーゼである。各クラス又はそのサブクラスのペプチダーゼの分類及び活性は、周知であり且つ文献に記載されている（たとえば、グルプリヤ，Ｖ．Ｓ．（Ｇｕｒｕｐｒｉｙａ，Ｖ．Ｓ．）及びロイ，Ｓ．Ｃ．（Ｒｏｙ，Ｓ．Ｃ．）著、「男性生殖におけるプロテアーゼ及びプロテアーゼ阻害剤（Ｐｒｏｔｅａｓｅｓ ａｎｄ Ｐｒｏｔｅａｓｅ Ｉｎｈｉｂｉｔｏｒｓ ｉｎ Ｍａｌｅ Ｒｅｐｒｏｄｕｃｔｉｏｎ）」、プロテアーゼ・イン・フィジオロジー・アンド・パソロジー（Ｐｒｏｔｅａｓｅｓ ｉｎ Ｐｈｙｓｉｏｌｏｇｙ ａｎｄ Ｐａｔｈｏｌｏｇｙ）、ｐ．１９５～２１６、２０１７年、並びにブリックス，Ｋ．（Ｂｒｉｘ，Ｋ．）及びストッカー，Ｗ．（Ｓｔｏｃｋｅｒ，Ｗ．）著、「プロテアーゼ：構造及び機能（Ｐｒｏｔｅａｓｅｓ：Ｓｔｒｕｃｔｕｒｅ ａｎｄ Ｆｕｎｃｔｉｏｎ）」、第１章）を参照されたい）。いくつかの実施形態では、本願に係るペプチダーゼは、ポリペプチド末端から３アミノ酸超を除去する。それゆえ、いくつかの実施形態では、ペプチダーゼ

は、たとえば、特定位置（たとえば、特定アミノ酸の前又は後）で優先的に切断するエンドペプチダーゼである。いくつかの実施形態では、エンドペプチダーゼ活性のポリペプチド切断産物のサイズは、分析されるポリペプチド内の切断部位（たとえばアミノ酸）の分布に依存するであろう。

【０１２２】

本願に係るエキソペプチダーゼは、シーケンシング反応の方向性に基づいて選択又は工学操作されうる。たとえば、ポリペプチドのアミノ末端からカルボキシ末端へのシーケンシングの実施形態では、エキソペプチダーゼはアミノペプチダーゼ活性を含む。逆に、ポリペプチドのカルボキシ末端からアミノ末端へのシーケンシングの実施形態では、エキソペプチダーゼはカルボキシペプチダーゼ活性を含む。特異的カルボキシ末端アミノ酸を認識するとともに標識付エキソペプチダーゼとして使用されうるか又は本明細書に記載の非切断標識付認識分子として使用するように不活性化されうるカルボキシペプチダーゼの例は、文献に記載されている（たとえば、ガルシア－ゲレロ，Ｍ．Ｃ．（Ｇａｒｃｉａ－Ｇｕｅｒｒｅｒｏ，Ｍ．Ｃ．）ら著、２０１８年、米国科学アカデミー紀要（ＰＮＡＳ）、第１１５巻、第１７号を参照されたい）。

【０１２３】

切断試薬及び／又は認識分子として使用するのに好適なペプチダーゼとしては、１タイプ以上のアミノ酸に選択的に結合するアミノペプチダーゼが挙げられる。いくつかの実施形態では、アミノペプチダーゼ認識分子は、アミノペプチダーゼ活性を不活性化するように修飾される。いくつかの実施形態では、アミノペプチダーゼ切断試薬は、非特異的であるため、ポリペプチドの末端からほとんど又はすべてのタイプのアミノ酸を切断する。いくつかの実施形態では、アミノペプチダーゼ切断試薬は、ポリペプチドの末端の他のタイプのアミノ酸と比較してより効率良くポリペプチドの末端から１タイプ以上のアミノ酸を切断する。たとえば、本願に係るアミノペプチダーゼは、アラニン、アルギニン、アスパラギン、アスパラギン酸、システイン、グルタミン、グルタミン酸、グリシン、ヒスチジン、イソロイシン、ロイシン、リシン、メチオニン、フェニルアラニン、プロリン、セレノシステイン、セリン、トレオニン、トリプトファン、チロシン、及び／又はバリンを特異的に切断する。いくつかの実施形態では、アミノペプチダーゼはプロリンアミノペプチダーゼである。いくつかの実施形態では、アミノペプチダーゼはプロリンイミノペプチダーゼである。いくつかの実施形態では、アミノペプチダーゼはグルタミン酸／アスパラギン酸特異的アミノペプチダーゼである。いくつかの実施形態では、アミノペプチダーゼはメチオニン特異的アミノペプチダーゼである。いくつかの実施形態では、アミノペプチダーゼは表４に示されるアミノペプチダーゼである。いくつかの実施形態では、アミノペプチダーゼ切断試薬は表４に示されるペプチド基質を切断する。

【０１２４】

いくつかの実施形態では、アミノペプチダーゼは非特異的アミノペプチダーゼである。いくつかの実施形態では、非特異的アミノペプチダーゼは亜鉛メタロプロテアーゼである。いくつかの実施形態では、非特異的アミノペプチダーゼは表５に示されるアミノペプチダーゼである。いくつかの実施形態では、非特異的アミノペプチダーゼは表５に示されるペプチド基質を切断する。

【０１２５】

それゆえ、いくつかの実施形態では、本願は、表４若しくは表５から選択されるアミノ酸配列を有する（又は表４若しくは表５から選択されるアミノ酸配列に対して少なくとも５０％、少なくとも６０％、少なくとも７０％、少なくとも８０％、８０～９０％、９０～９５％、９５～９９％、若しくはそれ以上のアミノ酸配列同一性を有するアミノ酸配列を有する）アミノペプチダーゼ（たとえば、アミノペプチダーゼ認識分子、アミノペプチダーゼ切断試薬）を提供する。いくつかの実施形態では、アミノペプチダーゼは、表４又は表５に列挙されるアミノペプチダーゼに対して２５～５０％、５０～６０％、６０～７０％、７０～８０％、８０～９０％、９０～９５％、若しくは９５～９９％、又はそれ以上のアミノ酸配列同一性を有する。いくつかの実施形態では、アミノペプチダーゼは、修

飾アミノペプチダーゼであり、且つ表４又は表５に示される配列に対して１つ以上のアミノ酸突然変異を含む。

【０１２６】

【表４－１】

表４．アミノペプチダーゼの非限定的例

名称	配列番号	配列
L. pneumophila M1 アミノペプチダーゼ (Glu/Asp 特異的)	198	MMVKQGVFMKTDQSKVKKLSDYKSLDYFVIHVDLQIDLSKKPVESK ARLTVVPNLNVDSHSNDLVLDGENMTLVSLQMNDNLLKENEYELTK DSLIIKNIPQNTPFTIEMTSLLGENTDLFGLYETEGVALVKAESEG LRRVFYLPDRPDNLATYKTTIIANQEDYPVLLSNGVLIEKKELPLG LHSVTWLDDVPKPSYLFALVAGNLQRSVTYYQTKSGRELPIEFYVP PSATSKCDFAKEVLKEAMAWDERTFNLECALRQHMVAGVDKYASGA SEPTGLNLFNTENLFASPETKTDLGILRVLEVVAHEFFHYWSGDRV TIRDWFNLPLKEGLTTFRAAMFREELFGTDLIRLLDGKNLDERAPR QSAYTAVRSLYTAAAYEKSADIFRMMMLFIGKEPFIEAVAKFFKDN DGGAVTLEDFIESISNSSGKDLRSFLSWFTESGIPELIVTDELNPD TKQYFLKIKTVNGRNRPIPILMGLLDSSGAEIVADKLLIVDQEEIE FQFENIQTRPIPSLLRSFSAPVHMKYEYSYQDLLLLMQFDTNLYNR CEAAKQLISALINDFCIGKKIELSPQFFAVYKALLSDNSLNEWMLA ELITLPSLEELIENQDKPDFEKLNEGRQLIQNALANELKTDFYNLL FRIQISGDDDKQKLKGFDLKQAGLRRLKSVCFSYLLNVDFEKTKEK LILQFEDALGKNMTETALALSMLCEINCEEADVALEDYYHYWKNDP GAVNNWFSIQALAHSPDVIERVKKLMRHGDFDLSNPNKVYALLGSF IKNPFGFHSVTGEGYQLVADAIFDLDKINPTLAANLTEKFTYWDKY DVNRQAMMISTLKIIYSNATSSDVRTMAKKGLDKVKEDLPLPIHLT FHGGSTMQDRTAQLIADGNKENAYQLH
E. coli メチオニン アミノペプチダーゼ (Met 特異的)	199	MGTAISIKTPEDIEKMRVAGRLAAEVLEMIEPYVKPGVSTGELDRI CNDYIVNEQHAVSACLGYHGYPKSVCISINEVVCHGIPDDAKLLKD GDIVNIDVTVIKDGFHGDTSKMFIVGKPTIMGERLCRITQESLYLA LRMVKPGINLREIGAAIQKFVEAEGFSVVREYCGHGIGRGFHEEPQ VLHYDSRETNVVLKPGMTFTIEPMVNAGKKEIRTMKDGWTVKTKDR SLSAQYEHTIVVTDNGCEILTLRKDDTIPAIISHD
M. smegmatis プロリン イミノペプチダーゼ (Pro 特異的)	200	MGTLEANTNGPGSMLSRMPVSSRTVPFGDHETWVQVTTPENAQPHA LPLIVLHGGPGMAHNYVANIAALADETGRTVIHYDQVGCGNSTHLP DAPADFWTPQLFVDEFHAVCTALGIERYHVLGQSWGGMLGAEIAVR QPSGLVSLAICNSPASMRLWSEAAGDLRAQLPAETRAALDRHEAAG TITHPDYLQAAAEFYRRHVCRVVPTPQDFADSVAQMEAEPTVYHTM NGPNEFHVVGTLGDWSVIDRLPDVTAPVLVIAGEHDEATPKTWQPF VDHIPDVRSHVFPGTSHCTHLEKPEEFRAVVAQFLHQHDLAADARV

【０１２７】

【表４－２】

Y. pestis プロリン イミノペプチダーゼ (Pro 特異的)	201	MTQQEYQNRRQALLAKMAPGSAAIIFAAPEATRSADSEYPYRQNSD FSYLTGFNEPEAVLILVKSDETHNHSVLFNRIRDLTAEIWFGRRLG QEAAPTKLAVDRALPFDEINEQLYLLLNRLDVIYHAQGQYAYADNI VFAALEKLRHGFRKNLRAPATLTDWRPWLHEMRLFKSAEEIAVLRR AGEISALAHTRAMEKCRPGMFEYQLEGEILHEFTRHGARYPAYNTI VGGGENGCILHYTENECELRDGDLVLIDAGCEYRGYAGDITRTFPV NGKFTPAQRAVYDIVLAAINKSLTLFRPGTSIREVTEEVVRIMVVG LVELGILKGDIEQLIAEQAHRPFFMHGLSHWLGMDVHDVGDYGSSD RGRILEPGMVLTVEPGLYIAPDADVPPQYRGIGIRIEDDIVITATG NENLTASVVKDPDDIEALMALNHAGENLYFQE
P. furiosus メチオニン アミノペプチダーゼ	202	MDTEKLMKAGEIAKKVREKAIKLARPGMLLLELAESIEKMIMELGG KPAFPVNLSINEIAAHYTPYKGDTTVLKEGDYLKIDVGVHIDGFIA DTAVTVRVGMEEDELMEAAKEALNAAISVARAGVEIKELGKAIENE IRKRGFKPIVNLSGHKIERYKLHAGISIPNIYRPHDNYVLKEGDVF AIEPFATIGAGQVIEVPPTLIYMYVRDVPVRVAQARFLLAKIKREY GTLPFAYRWLQNDMPEGQLKLALKTLEKAGAIYGYPVLKEIRNGIV AQFEHTIIVEKDSVIVTQDMINKSTLE
Aeromonas sobria プロリン アミノペプチダーゼ	203	HMSSPLHYVLDGIHCEPHFFTVPLDHQQPDDEETITLFGRTLCRKD RLDDELPWLLYLQGGPGFGAPRPSANGGWIKRALQEFRVLLLDQRG TGHSTPIHAELLAHLNPRQQADYLSHFRADSIVRDAELIREQLSPD HPWSLLGQSFGGFCSLTYLSLFPDSLHEVYLTGGVAPIGRSADEVY RATYQRVADKNRAFFARFPHAQAIANRLATHLQRHDVRLPNGQRLT VEQLQQQGLDLGASGAFEELYYLLEDAFIGEKLNPAFLYQVQAMQP FNTNPVFAILHELIYCEGAASHWAAERVRGEFPALAWAQGKDFAFT GEMIFPWMFEQFRELIPLKEAAHLLAEKADWGPLYDPVQLARNKVP VACAVYAEDMYVEFDYSRETLKGLSNSRAWITNEYEHNGLRVDGEQ ILDRLIRLNRDCLE
Pyrococcus furiosus プロリン アミノペプチダーゼ (X-/-Pro)	204	MKERLEKLVKFMDENSIDRVFIAKPVNVYYFSGTSPLGGGYIIVDG DEATLYVPELEYEMAKEESKLPVVKFKKFDEIYEILKNTETLGIEG TLSYSMVENFKEKSNVKEFKKIDDVIKDLRIIKTKEEIEIIEKACE IADKAVMAAIEEITEGKRREREVAAKVEYLMKMNGAEKPAFDTIIAS GHRSALPHGVASDKRIERGDLVVIDLGALYNHYNSDITRTIVVGSP NEKQREIYEIVLEAQKRAVEAAKPGMTAKELDSIAREIIKEYGYGD YFIHSLGHGVGLEIHEWPRISQYDETVLKEGMVITIEPGIYIPKLG GVRIEDTVLITENGAKRLTKTERELL
Elizabethkingia meningoseptica プロリン アミノペプチダーゼ	205	MIPITTPVGNFKVWTKRFGTNPKIKVLLLHGGPAMTHEYMECFETF FQREGFEFYEYDQLGSYYSDQPTDEKLWNIDRFVDEVEQVRKAIHA DKENFYVLGNSWGGILAMEYALKYQQNLKGLIVANMMASAPEYVKY AEVLSKQMKPEVLAEVRAIEAKKDYANPRYTELLFPNYYAQHICRL KEWPDALNRSLKHVNSTVYTLMQGPSELGMSSDARLAKWDIKNRLH EIATPTLMIGARYDTMDPKAMEEQSKLVQKGRYLYCPNGSHLAMWD DQKVFMDGVIKFIKDVDTKSFN
N. gonorrhoeae プロリン イミノペプチダーゼ	206	MYEIKQPFHSGYLQVSEIHQIYWEESGNPDGVPVIFLHGGPGAGAS PECRGFFNPDVFRIVIIDQRGCGRSHPYACAEDNTTWDLVADIEKV REMLGIGKWLVFGGSWGSTLSLAYAQTHPERVKGLVLRGIFLCRPS ETAWLNEAGGVSRIYPEQWQKFVAPIAENRRNRLIEAYHGLLFHQD EEVCLSAAKAWADWESYLIRFEPEGVDEDAYASLAIARLENHYFVN GGWLQGDKAILNNIGKIRHIPTVIVQGRYDLCTPMQSAWELSKAFP EAELRVVQAGHCAFDPPLADALVQAVEDILPRLL

【０１２８】

【表５－１】

表５．非特異的アミノペプチダーゼの非限定的例

名称	配列番号	配列
E. coli アミノペプチダーゼ N*	207	MTQQPQAKYRHDYRAPDYQITDIDLTFDLDAQKTVVTAVSQAVRHG ASDAPLRLNGEDLKLVSVHINDEPWTAWKEEEGALVISNLPERFTL KIINEISPAANTALEGLYQSGDALCTQCEAEGFRHITYYLDRPDVL

【０１２９】
【表５－２】

(亜鉛メタロプロテアーゼ)		ARFTTKIIADKIKYPFLLSNGNRVAQGELENGRHWVQWQDPFPKPC YLFALVAGDFDVLRDTFTTRSGREVALELYVDRGNLDRAPWAMTSL KNSMKWDEERFGLEYDLDIYMIVAVDFFNMGAMENKGLNIFNSKYV LARTDTATDKDYLDIERVIGHEYFHNWTGNRVTCRDWFQLSLKEGL TVFRDQEFSSDLGSRAVNRINNVRTMRGLQFAEDASPMAHPIRPDM VIEMNNFYTLTVYEKGAEVIRMIHTLLGEENFQKGMQLYFERHDGS AATCDDFVQAMEDASNVDLSHFRRWYSQSGTPIVTVKDDYNPETEQ YTLTISQRTPATPDQAEKQPLHIPFAIELYDNEGKVIPLQKGGHPV NSVLNVTQAEQTFVFDNVYFQPVPALLCEFSAPVKLEYKWSDQQLT FLMRHARNDFSRWDAAQSLLATYIKLNVARHQQGQPLSLPVHVADA FRAVLLDEKIDPALAAEILTLPSVNEMAELFDIIDPIAIAEVREAL TRTLATELADELLAIYNANYQSEYRVEHEDIAKRTLRNACLRFLAF GETHLADVLVSKQFHEANNMTDALAALSAAVAAQLPCRDALMQEYD DKWHQNGLVMDKWFILQATSPAANVLETVRGLLQHRSFTMSNPNRI RSLIGAFAGSNPAAFHAEDGSGYLFLVEMLTDLNSRNPQVASRLIE PLIRLKRYDAKRQEKMRAALEQLKGLENLSGDLYEKITKALA
P. falciparum M1 アミノペプチダーゼ**	208	PKIHYRKDYKPSGFIINQVTLNINIHDQETIVRSVLDMDISKHNVG EDLVFDGVGLKINEISINNKKLVEGEEYTYDNEFLTIFSKFVPKSK FAFSSEVIIHPETNYALTGLYKSKNIIVSQCEATGFRRITFFIDRP DMMAKYDVTVTADKEKYPVLLSNGDKVNEFEIPGGRHGARFNDPPL KPCYLFAVVAGDLKHLSATYITKYTKKKVELYVFSEEKYVSKLQWA LECLKKSMAFDEDYFGLEYDLSRLNLVAVSDFNVGAMENKGLNIFN ANSLLASKKNSIDFSYARILTVVGHEYFHQYTGNRVTLRDWFQLTL KEGLTVHRENLFSEEMTKTVTTRLSHVDLLRSVQFLEDSSPLSHPI RPESYVSMENFYTTTVYDKGSEVMRMYLTILGEEYYKKGFDIYIKK NDGNTATCEDFNYAMEQAYKMKKADNSANLNQYLLWFSQSGTPHVS FKYNYDAEKKQYSIHVNQYTKPDENQKEKKPLFIPISVGLINPENG KEMISQTTLELTKESDTFVFNNIAVKPIPSLFRGFSAPVYIEDQLT DEERILLLKYDSDAFVRYNSCTNIYMKQILMNYNEFLKAKNEKLES FQLTPVNAQFIDAIKYLLEDPHADAGFKSYIVSLPQDRYIINFVSN LDTDVLADTKEYIYKQIGDKLNDVYYKMFKSLEAKADDLTYFNDES HVDFDQMNMRTLRNTLLSLLSKAQYPNILNEIIEHSKSPYPSNWLT SLSVSAYFDKYFELYDKTYKLSKDDELLLQEWLKTVSRSDRKDIYE ILKKLENEVLKDSKNPNDIRAVYLPFTNNLRRFHDISGKGYKLIAE VITKTDKFNPMVATQLCEPFKLWNKLDTKRQELMLNEMNTMLQEPQ ISNNLKEYLLRLTNK
ピューロマイシン感受性 アミノペプチダーゼ (“NPEPPS”)	209	MWLAAAAPSLARRLLFLGPPPPPLLLLVFSRSSRRRLHSLGLAAMP EKRPFERLPADVSPINYSLCLKPDLLDFTFEGKLEAAAQVRQATNQ IVMNCADIDIITASYAPEGDEEIHATGFNYQNEDEKVTLSFPSTLQ TGTGTLKIDFVGELNDKMKGFYRSKYTTPSGEVRYAAVTQFEATDA RRAFPCWDEPAIKATFDISLVVPKDRVALSNMNVIDRKPYPDDENL VEVKFARTPVMSTYLVAFVVGEYDFVETRSKDGVCVRVYTPVGKAE QGKFALEVAAKTLPFYKDYFNVPYPLPKIDLIAIADFAAGAMENWG LVTYRETALLIDPKNSCSSSRQWVALVVGHELAHQWFGNLVTMEWW THLWLNEGFASWIEYLCVDHCFPEYDIWTQFVSADYTRAQELDALD NSHPIEVSVGHPSEVDEIFDAISYSKGASVIRMLHDYIGDKDFKKG MNMYLTKFQQKNAATEDLWESLENASGKPIAAVMNTWTKQMGFPLI YVEAEQVEDDRLLRLSQKKFCAGGSYVGEDCPQWMVPITISTSEDP NQAKLKILMDKPEMNVVLKNVKPDQWVKLNLGTVGFYRTQYSSAML ESLLPGIRDLSLPPVDRLGLQNDLFSLARAGIISTVEVLKVMEAFV NEPNYTVWSDLSCNLGILSTLLSHTDFYEEIQEFVKDVFSPIGERL GWDPKPGEGHLDALLRGLVLGKLGKAGHKATLEEARRRFKDHVEGK QILSADLRSPVYLTVLKHGDGTTLDIMLKLHKQADMQEEKNRIERV LGATLLPDLIQKVLTFALSEEVRPQDTVSVIGGVAGGSKHGRKAAW KFIKDNWEELYNRYQGGFLISRLIKLSVEGFAVDKMAGEVKAFFES HPAPSAERTIQQCCENILLNAAWLKRDAESIHQYLLQRKASPPTV
NPEPPS E366V	210	MWLAAAAPSLARRLLFLGPPPPPLLLLVFSRSSRRRLHSLGLAAMP EKRPFERLPADVSPINYSLCLKPDLLDFTFEGKLEAAAQVRQATNQ

【０１３０】

【表５－３】

		IVMNCADIDIITASYAPEGDEEIHATGFNYQNEDEKVTLSFPSTLQ TGTGTLKIDFVGELNDKMKGFYRSKYTTPSGEVRYAAVTQFEATDA RRAFPCWDEPAIKATFDISLVVPKDRVALSNMNVIDRKPYPDDENL VEVKFARTPVMSTYLVAFVVGEYDFVETRSKDGVCVRVYTPVGKAE QGKFALEVAAKTLPFYKDYFNVPYPLPKIDLIAIADFAAGAMENWG LVTYRETALLIDPKNSCSSSRQWVALVVGHVLAHQWFGNLVTMEWW THLWLNEGFASWIEYLCVDHCFPEYDIWTQFVSADYTRAQELDALD NSHPIEVSVGHPSEVDEIFDAISYSKGASVIRMLHDYIGDKDFKKG MNMYLTKFQQKNAATEDLWESLENASGKPIAAVMNTWTKQMGFPLI YVEAEQVEDDRLLRLSQKKFCAGGSYVGEDCPQWMVPITISTSEDP NQAKLKILMDKPEMNVVLKNVKPDQWVKLNLGTVGFYRTQYSSAML ESLLPGIRDLSLPPVDRLGLQNDLFSLARAGIISTVEVLKVMEAFV NEPNYTVWSDLSCNLGILSTLLSHTDFYEEIQEFVKDVFSPIGERL GWDPKPGEGHLDALLRGLVLGKLGKAGHKATLEEARRRFKDHVEGK QILSADLRSPVYLTVLKHGDGTTLDIMLKLHKQADMQEEKNRIERV LGATLLPDLIQKVLTFALSEEVRPQDTVSVIGGVAGGSKHGRKAAW KFIKDNWEELYNRYQGGFLISRLIKLSVEGFAVDKMAGEVKAFFES HPAPSAERTIQQCCENILLNAAWLKRDAESIHQYLLQRKASPPTV
Francisella tularensis アミノペプチダーゼ N	211	MIYEFVMTDPKIKYLKDYKPSNYLIDETHLIFELDESKTRVTANLY IVANRENRENNTLVLDGVELKLLSIKLNNKHLSPAEFAVNENQLII NNVPEKFVLQTVVEINPSANTSLEGLYKSGDVFSTQCEATGFRKIT YYLDRPDVMAAFTVKIIADKKKYPIILSNGDKIDSGDISDNQHFAV WKDPFKKPCYLFALVAGDLASIKDTYITKSQRKVSLEIYAFKQDID KCHYAMQAVKDSMKWDEDRFGLEYDLDTFMIVAVPDFNAGAMENKG LNIFNTKYIMASNKTATDKDFELVQSVVGHEYFHNWTGDRVTCRDW FQLSLKEGLTVFRDQEFTSDLNSRDVKRIDDVRIIRSAQFAEDASP MSHPIRPESYIEMNNFYTVTVYNKGAEIIRMIHTLLGEEGFQKGMK LYFERHDGQAVTCDDFVNAMADANNRDFSLFKRWYAQSGTPNIKVS ENYDASSQTYSLTLEQTTLPTADQKEKQALHIPVKMGLINPEGKNI AEQVIELKEQKQTYTFENIAAKPVASLFRDFSAPVKVEHKRSEKDL LHIVKYDNNAFNRWDSLQQIATNIILNNADLNDEFLNAFKSILHDK DLDKALISNALLIPIESTIAEAMRVIMVDDIVLSRKNVVNQLADKL KDDWLAVYQQCNDNKPYSLSAEQIAKRKLKGVCLSYLMNASDQKVG TDLAQQLFDNADNMTDQQTAFTELLKSNDKQVRDNAINEFYNRWRH EDLVVNKWLLSQAQISHESALDIVKGLVNHPAYNPKNPNKVYSLIG GFGANFLQYHCKDGLGYAFMADTVLALDKFNHQVAARMARNLMSWK RYDSDRQAMMKNALEKIKASNPSKNVFEIVSKSLES
Pyrococcus horikoshii TET アミノペプチダーゼ	212	MEVRNMVDYELLKKVVEAPGVSGYEFLGIRDVVIEEIKDYVDEVKV DKLGNVIAHKKGEGPKVMIAAHMDQIGLMVTHIEKNGFLRVAPIGG VDPKTLIAQRFKVWIDKGKFIYGVGASVPPHIQKPEDRKKAPDWDQ IFIDIGAESKEEAEDMGVKIGTVITWDGRLERLGKHRFVSIAFDDR IAVYTILEVAKQLKDAKADVYFVATVQEEVGLRGARTSAFGIEPDY GFAIDVTIAADIPGTPEHKQVTHLGKGTAIKIMDRSVICHPTIVRW LEELAKKHEIPYQLEILLGGGTDAGAIHLTKAGVPTGALSVPARYI HSNTEVVDERDVDATVELMTKALENIHELKI
T. aquaticus アミノペプチダーゼ T	213	MDAFTENLNKLAELAIRVGLNLEEGQEIVATAPIEAVDFVRLLAEK AYENGASLFTVLYGDNLIARKRLALVPEAHLDRAPAWLYEGMAKAF HEGAARLAVSGNDPKALEGLPPERVGRAQQAQSRAYRPTLSAITEF VTNWTIVPFAHPGWAKAVFPGLPEEEAVQRLWQAIFQATRVDQEDP VAAWEAHNRVLHAKVAFLNEKRFHALHFQGPGTDLTVGLAEGHLWQ GGATPTKKGRLCNPNLPTEEVFTAPHRERVEGVVRASRPLALSGQL VEGLWARFEGGVAVEVGAEKGEEVLKKLLDTDEGARRLGEVALVPA DNPIAKTGLVFFDTLFDENAASHIAFGQAYAENLEGRPSGEEFRRR GGNESMVHVDWMIGSEEVDVDGLLEDGTRVPLMRRGRWVI
Bacillus stearothermophilus ペプチダーゼ M28	214	MAKLDETLTMLKALTDAKGVPGNEREARDVMKTYIAPYADEVTTDG LGSLIAKKEGKSGGPKVMIAGHLDEVGFMVTQIDDKGFIRFQTLGG WWSQVMLAQRVTIVTKKGDITGVIGSKPPHILPSEARKKPVEIKDM FIDIGATSREEAMEWGVRPGDMIVPYFEFTVLNNEKMLLAKAWDNR

【０１３１】

【表５－４】

		IGCAVAIDVLKQLKGVDHPNTVYGVGTVQEEVGLRGARTAAQFIQP DIAFAVDVGIAGDTPGVSEKEAMGKLGAGPHIVLYDATMVSHRGLR EFVIEVAEELNIPHHFDAMPGVGTDAGAIHLTGIGVPSLTIAIPTR YIHSHAAILHRDDYENTVKLLVEVIKRLDADKVKQLTFDE
Vibrio cholera アミノペプチダーゼ	215	MEDKVWISMGADAVGSLNPALSESLLPHSFASGSQVWIGEVAIDEL AELSHTMHEQHNRCGGYMVHTSAQGAMAALMMPESIANFTIPAPSQ QDLVNAWLPQVSADQITNTIRALSSFNNRFYTTTSGAQASDWLANE WRSLISSLPGSRIEQIKHSGYNQKSVVLTIQGSEKPDEWVIVGGHL DSTLGSHTNEQSIAPGADDDASGIASLSEIIRVLRDNNFRPKRSVA LMAYAAEEVGLRGSQDLANQYKAQGKKVVSVLQLDMTNYRGSAEDI VFITDYTDSNLTQFLTTLIDEYLPELTYGYDRCGYACSDHASWHKA GFSAAMPFESKFKDYNPKIHTSQDTLANSDPTGNHAVKFTKLGLAY VIEMANAGSSQVPDDSVLQDGTAKINLSGARGTQKRFTFELSQSKP LTIQTYGGSGDVDLYVKYGSAPSKSNWDCRPYQNGNRETCSFNNAQ PGIYHVMLDGYTNYNDVALKASTQ
Photobacterium halotolerans アミノペプチダーゼ	216	MEDKVWISIGSDASQTVKSVMQSNARSLLPESLASNGPVWVGQVDY SQLAELSHHMHEDHQRCGGYMVHSSPESAIAASNMPQSLVAFSIPE ISQQDTVNAWLPQVNSQAITGTITSLTSFINRFYTTTSGAQASDWL ANEWRSLSASLPNASVRQVSHFGYNQKSVVLTITGSEKPDEWIVLG GHLDSTIGSHTNEQSVAPGADDDASGIASVTEIIRVLSENNFQPKR SIAFMAYAAEEVGLRGSQDLANQYKAEGKQVISALQLDMTNYKGSV EDIVFITDYTDSNLTTFLSQLVDEYLPSLTYGFDTCGYACSDHASW HKAGFSAAMPFEAKFNDYNPMIHTPNDTLQNSDPTASHAVKFTKLG LAYAIEMASTTGGTPPPTGNVLKDGVPVNGLSGATGSQVHYSFELP AQKNLQISTAGGSGDVDLYVSFGSEATKQNWDCRPYRNGNNEVCTF AGATPGTYSIMLDGYRQFSGVTLKASTQ
Yersinia pestis アミノペプチダーゼN	217	MTQQPQAKYRHDYRAPDYTITDIDLDFALDAQKTTVTAVSKVKRQG TDVTPLILNGEDLTLISVSVDGQAWPHYRQQDNTLVIEQLPADFTL TIVNDIHPATNSALEGLYLSGEALCTQCEAEGFRHITYYLDRPDVL ARFTTRIVADKSRYPYLLSNGNRVGQGELDDGRHWVKWEDPFPKPS YLFALVAGDFDVLQDKFITRSGREVALEIFVDRGNLDRADWAMTSL KNSMKWDETRFGLEYDLDIYMIVAVDFFNMGAMENKGLNVFNSKYV LAKAETATDKDYLNIEAVIGHEYFHNWTGNRVTCRDWFQLSLKEGL TVFRDQEFSSDLGSRSVNRIENVRVMRAAQFAEDASPMAHAIRPDK VIEMNNFYTLTVYEKGSEVIRMMHTLLGEQQFQAGMRLYFERHDGS AATCDDFVQAMEDVSNVDLSLFRRWYSQSGTPLLTVHDDYDVEKQQ YHLFVSQKTLPTADQPEKLPLHIPLDIELYDSKGNVIPLQHNGLPV HHVLNVTEAEQTFTFDNVAQKPIPSLLREFSAPVKLDYPYSDQQLT FLMQHARNEFSRWDAAQSLLATYIKLNVAKYQQQQPLSLPAHVADA FRAILLDEHLDPALAAQILTLPSENEMAELFTTIDPQAISTVHEAI TRCLAQELSDELLAVYVANMTPVYRIEHGDIAKRALRNTCLNYLAF GDEEFANKLVSLQYHQADNMTDSLAALAAAVAAQLPCRDELLAAFD VRWNHDGLVMDKWFALQATSPAANVLVQVRTLLKHPAFSLSNPNRT RSLIGSFASGNPAAFHAADGSGYQFLVEILSDLNTRNPQVAARLIE PLIRLKRYDAGRQALMRKALEQLKTLDNLSGDLYEKITKALAA
Vibrio anguillarum アミノペプチダーゼ	218	MEEKVWISIGGDATQTALRSGAQSLLPENLINQTSVWVGQVPVSEL ATLSHEMHENHQRCGGYMVHPSAQSAMSVSAMPLNLNAFSAPEITQ QTTVNAWLPSVSAQQITSTITTLTQFKNRFYTTSTGAQASNWIADH WRSLSASLPASKVEQITHSGYNQKSVMLTITGSEKPDEWVVIGGHL DSTLGSRTNESSIAPGADDDASGIAGVTEIIRLLSEQNFRPKRSIA FMAYAAEEVGLRGSQDLANRFKAEGKKVMSVMQLDMTNYQGSREDI VFITDYTDSNFTQYLTQLLDEYLPSLTYGFDTCGYACSDHASWHAV GYPAAMPFESKFNDYNPNIHSPQDTLQNSDPTGFHAVKFTKLGLAY VVEMGNASTPPTPSNQLKNGVPVNGLSASRNSKTWYQFELQEAGNL SIVLSGGSGDADLYVKYQTDADLQQYDCRPYRSGNNETCQFSNAQP GRYSILLHGYNNYSNASLVANAQ

【０１３２】

【表５－５】

Salinivibrio spYCSC6 アミノペプチダーゼ	219	MEDKKVWISIGADAQQTALSSGAQPLLAQSVAHNGQAWIGEVSESE LAALSHEMHENHHRCGGYIVHSSAQSAMAASNMPLSRASFIAPAIS QQALVTPWISQIDSALIVNTIDRLTDFPNRFYTTTSGAQASDWIKQ RWQSLSAGLAGASVTQISHSGYNQASVMLTIEGSESPDEWVVVGGH LDSTIGSRTNEQSIAPGADDDASGIAAVTEVIRVLAQNNFQPKRSI AFVAYAAEEVGLRGSQDVANQFKQAGKDVRGVLQLDMTNYQGSAED IVFITDYTDNQLTQYLTQLLDEYLPTLNYGFDTCGYACSDHASWHQ VGYPAAMPFEAKFNDYNPNIHTPQDTLANSDSEGAHAAKFTKLGLA YTVELANADSSPNPGNELKLGEPINGLSGARGNEKYFNYRLDQSGE LVIRTYGGSGDVDLYVKANGDVSTGNWDCRPYRSGNDEVCRFDNAT PGNYAVMLRGYRTYDNVSLIVE
Vibrio proteolyticus アミノペプチダーゼI	220	MPPITQQATVTAWLPQVDASQITGTISSLESFTNRFYTTTSGAQAS DWIASEWQALSASLPNASVKQVSHSGYNQKSVVMTITGSEAPDEWI VIGGHLDSTIGSHTNEQSVAPGADDDASGIAAVTEVIRVLSENNFQ PKRSIAFMAYAAEEVGLRGSQDLANQYKSEGKNVVSALQLDMTNYK GSAQDVVFITDYTDSNFTQYLTQLMDEYLPSLTYGFDTCGYACSDH ASWHNAGYPAAMPFESKFNDYNPRIHTTQDTLANSDPTGSHAKKFT QLGLAYAIEMGSATGDTPTPGNQLE
Vibrio proteolyticus アミノペプチダーゼI (A55F)	221	MPPITQQATVTAWLPQVDASQITGTISSLESFTNRFYTTTSGAQAS DWIASEWQFLSASLPNASVKQVSHSGYNQKSVVMTITGSEAPDEWI VIGGHLDSTIGSHTNEQSVAPGADDDASGIAAVTEVIRVLSENNFQ PKRSIAFMAYAAEEVGLRGSQDLANQYKSEGKNVVSALQLDMTNYK GSAQDVVFITDYTDSNFTQYLTQLMDEYLPSLTYGFDTCGYACSDH ASWHNAGYPAAMPFESKFNDYNPRIHTTQDTLANSDPTGSHAKKFT QLGLAYAIEMGSATGDTPTPGNQLE
P. furiosus アミノペプチダーゼI	222	MVDWELMKKIIESPGVSGYEHLGIRDLVVDILKDVADEVKIDKLGN VIAHFKGSAPKVMVAAHMDKIGLMVNHIDKDGYLRVVPIGGVLPET LIAQKIRFFTEKGERYGVVGVLPPHLRREAKDQGGKIDWDSIIVDV GASSREEAEEMGFRIGTIGEFAPNFTRLSEHRFATPYLDDRICLYA MIEAARQLGEHEADIYIVASVQEEIGLRGARVASFAIDPEVGIAMD VTFAKQPNDKGKIVPELGKGPVMDVGPNINPKLRQFADEVAKKYEI PLQVEPSPRPTGTDANVMQINREGVATAVLSIPIRYMHSQVELADA RDVDNTIKLAKALLEELKPMDFTPLE

*切断効率（最大から最小へ）：アルギニン＞リシン＞疎水性残基（アラニン、ロイシン、メチオニン、及びフェニルアラニンを含む）＞プロリン（例えば、マシューズ（Matthews）著、バイオケミストリー（Biochemistry）、第47巻、2008年、p.5303~5311を参照されたい）。

**切断効率（最大から最小へ）：ロイシン＞アラニン＞アルギニン＞フェニルアラニン＞プロリン、ただし、グルタミン酸及びアスパラギン酸の後では切断しない。

【０１３３】

２つ以上のアミノ酸配列を比較する目的では、第１のアミノ酸配列と第２のアミノ酸配列との間の「配列同一性（ｓｅｑｕｅｎｃｅ　ｉｄｅｎｔｉｔｙ）」（本明細書では「アミノ酸同一性」ともいう）のパーセントは、［第２のアミノ酸配列中の対応する位置のアミノ酸残基と同一の第１のアミノ酸配列中のアミノ酸残基の数］を［第１のアミノ酸配列中のアミノ酸残基の合計数］で除算して［１００］を乗算することにより計算されうる。この場合、第１のアミノ酸配列と比較して第２のアミノ酸配列中のアミノ酸残基の各欠失、挿入、置換、又は付加は、単一アミノ酸残渣（位置）の差とみなされる。代替的に、２つのアミノ酸配列間の配列同一性の程度は、既知のコンピュータアルゴリズムを用いて計算されうる（たとえば、スミス（Ｓｍｉｔｈ）及びウォーターマン（Ｗａｔｅｒｍａｎ）著、１９７０年、アドバンシズ・イン・アプライド・マセマティックス（Ａｄｖ．Ａｐｐｌ．Ｍａｔｈ．）、第２巻、ｐ．４８２ｃの局所相同性アルゴリズムにより、ニードルマン（Ｎｅｅｄｌｅｍａｎ）及びブンシュ（Ｗｕｎｓｃｈ）著、ジャーナル・オブ・モレキ

ュラー・バイオロジー（Ｊ．Ｍｏｌ．Ｂｉｏｌ．）、１９７０年、第４８巻、ｐ．４４３の相同性アライメントアルゴリズムにより、ピアソン（Ｐｅａｒｓｏｎ）及びリップマン（Ｌｉｐｍａｎ）著、米国科学アカデミー紀要（Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ）、１９９８年、第８５巻、ｐ．２４４４の類似性検索法により、又はブラスト（Ｂｌａｓｔ）、クラスタル（Ｃｌｕｓｔａｌ）、オメガ（Ｏｍｅｇａ）として利用可能なアルゴリズム若しくは他の配列アライメントアルゴリズムをたとえば標準的設定値を用いてコンピュータで実行することにより）。通常、以上で概説した計算方法に従って２つのアミノ酸配列間の「配列同一性」のパーセントを決定する目的では、最大数のアミノ酸残基を有するアミノ酸配列は「第１」のアミノ酸配列とみなされ、他のアミノ酸配列は「第２」のアミノ酸配列とみなされるであろう。

【０１３４】

追加的又は代替的に、配列間の同一性に関して２つ以上の配列が評価されうる。２つ以上の核酸配列又はアミノ酸配列との関連での「同一」又はパーセント「同一性」という用語は、同一の２つ以上の配列又はサブ配列を意味する。以上の配列比較アルゴリズムの１つを用いて又は手動アライメント及び目視検査により測定される比較ウィンドウ又は指定領域にわたり比較して最大の対応が得られるようにアライメントしたときに、２つの配列が指定領域にわたり又は全配列にわたり同一（たとえば、少なくとも８０％、８５％、９０％、９５％、９６％、９７％、９８％、９９％、９９．５％、９９．６％、９９．７％、９９．８％、又は９９．９％同一）の特定パーセントのアミノ酸残基又はヌクレオチドを有する場合、２つの配列は「実質的に同一」である。任意選択的に、同一性は、少なくとも約２５、５０、７５、若しくは１００アミノ酸長の領域にわたり、又は１００～１５０、１５０～２００、１００～２００、若しくは２００アミノ酸長以上の領域にわたり存在する。

【０１３５】

追加的又は代替的に、配列間のアライメントに関して２つ以上の配列が評価されうる。２つ以上の核酸配列又はアミノ酸配列との関連での「アライメント」又はパーセント「アライメント」という用語は、同一の２つ以上の配列又はサブ配列を意味する。以上の配列比較アルゴリズムの１つを用いて又は手動アライメント及び目視検査により測定される比較ウィンドウ又は指定領域にわたり比較して最大の対応が得られるようにアライメントしたときに、２つの配列が指定領域にわたり又は全配列にわたり同一（たとえば、少なくとも８０％、８５％、９０％、９５％、９６％、９７％、９８％、９９％、９９．５％、９９．６％、９９．７％、９９．８％、又は９９．９％同一）の特定パーセントのアミノ酸残基又はヌクレオチドを有する場合、２つの配列は「実質的にアラインメント」させる。任意選択的に、アライメントは、少なくとも約２５、５０、７５、若しくは１００アミノ酸長の領域にわたり、又は１００～１５０、１５０～２００、１００～２００、若しくは２００アミノ酸長以上の領域にわたり存在する。

【０１３６】

タンパク質分子のほか、核酸分子は、本願に係るアミノ酸認識分子として使用するのに有利なさまざまな性質を有する。

核酸アプタマーは、高親和性且つ選択性で所望の標的に結合するように工学操作された核酸分子である。それゆえ、核酸アプタマーは、当技術分野で公知の選択及び／又は富化技術を用いて所望のタイプのアミノ酸に選択的に結合するように工学操作されうる。そのため、いくつかの実施形態では、認識分子は、核酸アプタマー（たとえば、ＤＮＡアプタマー、ＲＮＡアプタマー）を含む。図２に示されるように、いくつかの実施形態では、標識付認識分子は、１タイプの末端アミノ酸に選択的に結合する標識付アプタマー２０４である。たとえば、いくつかの実施形態では、標識付アプタマー２０４は、本明細書に記載されるように、ポリペプチドの末端の１タイプのアミノ酸（たとえば、単一タイプのアミノ酸又はアミノ酸のタイプのサブセット）に選択的に結合する。示されていないが、標識付アプタマー２０４は、本願の方法に従ってポリペプチドのいずれかの位置（たとえば、ポリペプチドの末端位置又は末端位置及び内部位置）の１タイプのアミノ酸に選択的に結

合するように工学操作されうることが認識されるべきである。

【０１３７】

いくつかの実施形態では、標識付認識分子は、結合誘起発光を有する標識を含む。たとえば、いくつかの実施形態では、標識付アプタマー２０６は、ドナー標識２１２とアクセプター標識２１４とを含み、図２のパネル（Ｉ）及び（ＩＩ）に例示されるように機能する。パネル（Ｉ）に描かれるように、遊離分子としての標識付アプタマー２０６は、ドナー標識２１２及びアクセプター標識２１４が標識間の検出可能ＦＲＥＴを制限する距離（たとえば約１０ｎｍ以上）だけ分離されるコンフォメーションをとる。パネル（ＩＩ）に描かれるように、選択的結合分子としての標識付アプタマー２０６は、ドナー標識２１２及びアクセプター標識２１４が標識間の検出可能ＦＲＥＴを促進する距離内（たとえば約１０ｎｍ以下）にあるコンフォメーションをとる。さらに他の実施形態では、標識付アプタマー２０６は、クエンチング部分を含むとともに分子ビーコンと同様に機能し、標識付アプタマー２０６の発光は、遊離分子として内部クエンチされ、且つ選択的結合分子として回復させる（たとえば、ハマグチ（Ｈａｍａｇｕｃｈｉ）ら著、２００１年、アナリティカル・バイオケミストリー（Ａｎａｌｙｔｉｃａｌ　Ｂｉｏｃｈｅｍｉｓｔｒｙ）、第２９４巻、ｐ．１２６～１３１を参照されたい）。理論により拘束されることを望むものではないが、結合誘起発光のこれらの及び他のタイプの機構は、バックグラウンド発光を有利に低減又は排除して本明細書に記載の方法の全体的感度及び確度を増加させると考えられる。

【０１３８】

シールド認識分子（Ｓｈｉｅｌｄｅｄ　Ｒｅｃｏｇｎｉｔｉｏｎ　Ｍｏｌｅｃｕｌｅｓ）

本明細書に記載の実施形態によれば、単一分子ポリペプチドシーケンシング方法は、励起光で表面固定ポリペプチドに照明することと、アミノ酸認識分子に装着された標識により生成される発光を検出することと、により行うことが可能である。いくつかの場合には、標識により生成される放射及び／又は非放射減衰は、ポリペプチドに光損傷をもたらしうる。たとえば、図３Ａは、表面に固定されたポリペプチドに認識分子が会合して示されるシーケンシング反応例を例示する。

【０１３９】

励起照明の存在下で、標識は、検出可能会合イベントをもたらす放射減衰を介して蛍光を生成可能である。しかしながら、いくつかの場合には、標識は、反応性酸素種３００の形成をもたらす可能性のある非放射減衰を生成する。反応性酸素種３００は、最終的に固定ペプチドを損傷する可能性があり、その場合、ポリペプチドの完全配列情報を得る前に反応が終了することになる。この光損傷は、たとえば、露出ポリペプチド末端（トップの中抜き矢印）、内部位置（中間の中抜き矢印）、表面にポリペプチドを装着する表面リンカー（ボトムの中抜き矢印）で発生する可能性がある。

【０１４０】

本発明者らは、光損傷を軽減可能であり且つアミノ酸認識分子へのシールディングエレメントの取込みにより認識時間を延長可能であることを見いだした。図３Ｂは、シールド３０２を含むシールド認識分子を用いたシーケンシング反応例を例示する。シールド３０２は、標識とポリペプチドとの間の距離の増加を提供する共有結合性又は非共有結合性連結基を形成するので、標識－ポリペプチド分離距離にわたるフリーラジカル減衰に起因して反応性酸素種３００からの損傷作用を低減可能である。シールド３０２はまた、反応性酸素種３００並びに放射及び／又は非放射減衰からの損傷を吸収することにより標識からポリペプチドをシールドする立体バリアを提供可能である。

【０１４１】

理論により拘束されることを望むものではないが、認識成分と標識成分との間に位置決めされるシールドは、標識成分により放出される放射及び／又は非放射減衰を吸収、偏向、さもなければブロックできると考えられる。いくつかの実施形態では、シールドは、１つ以上の標識（たとえば発光標識）と１つ以上のアミノ酸認識分子との相互作用を防止又は制限する。いくつかの実施形態では、シールドは、１つ以上の標識とアミノ酸認識分子

に会合する１つ以上の分子（たとえば、認識分子に会合するポリペプチド）との相互作用を防止又は制限する。それゆえ、いくつかの実施形態では、シールドという用語は、認識成分と標識成分との間に形成される連結基のいずれかの部分により提供される保護効果又はシールディング効果を一般に意味しうる。

【０１４２】

いくつかの実施形態では、シールドは、１つ以上のアミノ酸認識分子（たとえば認識成分）と１つ以上の標識（たとえば標識成分）とに装着される。いくつかの実施形態では、認識成分及び標識成分は、シールド上の非近接部位に装着される。たとえば、１つ以上のアミノ酸認識分子は、シールドの第１の側に装着可能であり、且つ１つ以上の標識は、シールドの第２の側に装着可能であり、この場合、シールドの第１及び第２の側は互いに離れている。いくつかの実施形態では、装着部位は、シールドのおおよそ両側である。

【０１４３】

シールドが認識分子に装着される部位とシールドが標識に装着される部位との間の距離は、スペースを介する線形測定又はシールドの表面全体にわたる非線形測定でありうる。シールド上の認識分子結合部位と標識結合部位との間の距離は、シールドの３次元構造のモデリングにより測定可能である。いくつかの実施形態では、この距離は、少なくとも２ｎｍ、少なくとも４ｎｍ、少なくとも６ｎｍ、少なくとも８ｎｍ、少なくとも１０ｎｍ、少なくとも１２ｎｍ、少なくとも１５ｎｍ、少なくとも２０ｎｍ、少なくとも３０ｎｍ、少なくとも４０ｎｍ、又はそれ以上でありうる。代替的に、シールド上の認識分子と標識との相対位置は、シールドの構造を二次曲面（たとえば、楕円体、楕円柱）として扱うことにより記述可能である。いくつかの実施形態では、認識分子結合部位及び標識結合部位は、シールドを表す楕円体形状の周りの距離の少なくとも１／８の距離だけ離れている。いくつかの実施形態では、認識分子及び標識は、シールドを表す楕円体形状の周りの距離の少なくとも１／４の距離だけ離れている。いくつかの実施形態では、認識分子及び標識は、シールドを表す楕円体形状の周りの距離の少なくとも１／３の距離だけ離れている。いくつかの実施形態では、認識分子及び標識は、シールドを表す楕円体形状の周りの距離の１／２の距離だけ離れている。

【０１４４】

シールドのサイズは、アミノ酸認識分子がポリペプチドに会合するときに標識がポリペプチドに直接接触できないか又はその可能性が低くなるようにすべきである。シールドのサイズはまた、アミノ酸認識分子がポリペプチドに会合するときに装着標識が検出可能になるようにすべきである。たとえば、サイズは、装着発光標識が励起される照明体積内にあるようにすべきである。

【０１４５】

実施者がシールディング効果を評価可能なさまざまなパラメータが存在することが認識されるべきである。一般に、シールディングエレメントの効果は、シールディングエレメントを有する組成物とシールディングエレメントの欠如した組成物との間の比較アセスメントを行うことにより評価可能である。たとえば、シールディングエレメントは、アミノ酸認識分子の認識時間を増加させることが可能である。いくつかの実施形態では、認識時間とは、本明細書に記載のポリペプチドシーケンシング反応時に認識分子とポリペプチドとの間の会合イベントが観測可能な時間の長さを意味する。いくつかの実施形態では、認識時間は、アミノ酸認識分子がシールディングエレメントを欠如しているがそれ以外は類似又は同一である点を除けば同一条件下で実施されるポリペプチドシーケンシング反応と比べて、約１０～２５％、２５～５０％、５０～７５％、７５～１００％、又は１００％超、たとえば、約２倍、３倍、４倍、５倍、又はそれ以上に増加する。いくつかの実施形態では、シールディングエレメントは、シーケンシング確度（ｓｅｑｕｅｎｃｉｎｇ　ａｃｃｕｒａｃｙ）及び／又は配列リード長を増加させることが可能である（たとえば、以上に記載の比較条件下で実施されるシーケンシング反応と比べて、少なくとも５％、少なくとも１０％、少なくとも１５％、少なくとも２５％、又はそれ以上に）。

【０１４６】

それゆえ、いくつかの態様では、本願は、少なくとも１つのアミノ酸認識分子と、少なくとも１つの検出可能標識と、認識分子と標識との間に共有結合性又は非共有結合性連結基を形成するシールディングエレメント（たとえば「シールド」）と、を含むシールド認識分子を提供する。いくつかの実施形態では、シールディングエレメントは、少なくとも２ｎｍ、少なくとも５ｎｍ、少なくとも１０ｎｍ、少なくとも１２ｎｍ、少なくとも１５ｎｍ、少なくとも２０ｎｍ、又はそれ以上の長さである（たとえば、水性溶液中で）。いくつかの実施形態では、シールディングエレメントは、約２ｎｍ～約１００ｎｍの長さ（たとえば、約２ｎｍ～約５０ｎｍ、約１０ｎｍ～約５０ｎｍ、約２０ｎｍ～約１００ｎｍ）である。

【０１４７】

いくつかの実施形態では、シールド（たとえばシールディングエレメント）は、１つ以上のアミノ酸認識分子（たとえば認識成分）と１つ以上の標識（たとえば標識成分）との間に共有結合性又は非共有結合性連結基を形成する。本明細書で用いられる場合、いくつかの実施形態では、共有結合性及び非共有結合性の連結又は連結基とは、シールドへの認識成分及び標識成分の装着の性状を意味する。

【０１４８】

いくつかの実施形態では、共有結合性連結又は共有結合性連結基とは、共有結合又は一連のコンティグ共有結合を介して認識成分及び標識成分の各々に装着されるシールドを意味する。一方又は両方の成分の共有結合性装着は、当技術分野で公知の共有結合性コンジュゲーション方法により達成可能である。たとえば、いくつかの実施形態では、一方又は両方の成分をシールドに装着するために、クリックケミストリー技術（たとえば、銅触媒、歪み促進、銅フリーのクリックケミストリーなど）を使用可能である。かかる方法は、一般に、反応性部分間に１つ以上の共有結合を形成するように一方の反応性部分を他方の反応性部分にコンジュゲートすることを含む。それゆえ、いくつかの実施形態では、シールドの第１の反応性部分は、共有結合性装着を形成するように認識成分又は標識成分の第２の反応性部分に接触される。反応性部分の例としては、限定されるものではないが、反応性のアミン、アジド、アルキン、ニトロン、アルケン（たとえばシクロアルケン）、テトラジン、テトラゾール、並びにクリック反応及び類似のカップリング技術に好適な他の反応性部分が挙げられる。

【０１４９】

いくつかの実施形態では、非共有結合性連結又は非共有結合性連結基とは、１つ以上の非共有結合性カップリング手段、たとえば、限定されるものではないがレセプター－リガンド相互作用及びオリゴヌクレオチド鎖ハイブリダイゼーションを介して認識成分及び標識成分の一方又は両方に装着されるシールドを意味する。レセプター－リガンド相互作用の例は、本明細書に提供されており、限定されるものではないが、タンパク質－タンパク質複合体、タンパク質－リガンド複合体、タンパク質－アプタマー複合体、及びアプタマー－核酸複合体が挙げられる。オリゴヌクレオチド鎖ハイブリダイゼーションに対する種々の構成及びストラテジーは、本明細書に記載されており、当技術分野で公知である（たとえば、米国特許出願公開第２０１９／００２４１６８号明細書を参照されたい）。

【０１５０】

いくつかの実施形態では、シールド３０２は、バイオ分子やデンドリティックポリマーなどのポリマーを含む。図３Ｃは、ポリマーシールドの例及び本願のシールド認識分子の構成を示す。第１のシールド構築物３０４は、タンパク質シールド３３０の例を示す。いくつかの実施形態では、タンパク質シールド３３０は、認識分子と標識との間に共有結合性連結基を形成する。たとえば、いくつかの実施形態では、タンパク質シールド３３０は、１つ以上の共有結合を介して、たとえば、タンパク質シールド３３０の天然又は非天然アミノ酸の側鎖を介する共有結合性装着により、認識分子及び標識の各々に装着される。いくつかの実施形態では、アミノ酸認識分子は、末端同士で連結された少なくとも１つのアミノ酸結合タンパク質及びタンパク質シールド３３０を有する単一ポリペプチドを含む。

【０１５１】
それゆえ、いくつかの態様では、本願は、末端同士で連結（たとえば、Ｃ末端とＮ末端との融合）したアミノ酸結合タンパク質及びタンパク質シールドを有する融合ポリペプチドを含むシールド認識分子を提供する。いくつかの実施形態では、バインダー及びタンパク質シールドは、共有結合、又は一方のタンパク質のＣ末端と他方のタンパク質のＮ末端とを共有結合的に連結するリンカーのいずれかによって、末端同士で連結される。いくつかの実施形態では、融合ポリペプチドとの関連でのリンカーは、バインダーとタンパク質シールドとを連結するがバインダー又はタンパク質シールドのいずれに対応するポリペプチド配列の部分も形成しない、融合ポリペプチド内の１つ以上のアミノ酸を意味する。いくつかの実施形態では、リンカーは、少なくとも２つのアミノ酸（たとえば、少なくとも２つ、３つ、４つ、５つ、６つ、８つ、１０、１５、２５、５０、１００、又はそれ以上のアミノ酸）を含む。いくつかの実施形態では、リンカーは、５つまで、１０まで、１５まで、２５まで、５０まで、又は１００までのアミノ酸を含む。いくつかの実施形態では、リンカーは、約２つ～約２００のアミノ酸（たとえば、約２つ～約１００、約５つ～約５０、約２つ～約２０、約５つ～約２０、又は約２つ～約３０のアミノ酸）を含む。
【０１５２】
いくつかの実施形態では、融合ポリペプチドのタンパク質シールドは、少なくとも１０ｋＤａの分子量を有するタンパク質である。たとえば、いくつかの実施形態では、タンパク質シールドは、少なくとも１０ｋＤａ且つ５００ｋＤａまで（たとえば、約１０ｋＤａ～約２５０ｋＤａ、約１０ｋＤａ～約１５０ｋＤａ、約１０ｋＤａ～約１００ｋＤａ、約２０ｋＤａ～約８０ｋＤａ、約１５ｋＤａ～約１００ｋＤａ、又は約１５ｋＤａ～約５０ｋＤａ）の分子量を有するタンパク質である。いくつかの実施形態では、融合ポリペプチドのタンパク質シールドは、少なくとも２５のアミノ酸を含むタンパク質である。たとえば、いくつかの実施形態では、タンパク質シールドは、少なくとも２５且つ１，０００までのアミノ酸（たとえば、約１００～約１，０００のアミノ酸、約１００～約７５０のアミノ酸、約５００～約１，０００のアミノ酸、約２５０～約７５０のアミノ酸、約５０～約５００のアミノ酸、約１００～約４００のアミノ酸、又は約５０～約２５０までのアミノ酸）を含むタンパク質である。
【０１５３】
いくつかの実施形態では、タンパク質シールドは、１つ以上のタグタンパク質を含むポチペプチドである。いくつかの実施形態では、タンパク質シールドは、少なくとも２つのタグタンパク質を含むポチペプチドである。いくつかの実施形態では、少なくとも２つのタグタンパク質は同一である（たとえば、ポリペプチドは、タグタンパク質配列の少なくとも２つのコピーを含む）。いくつかの実施形態では、少なくとも２つのタグタンパク質は互いに異なる（たとえば、ポリペプチドは、少なくとも２つの互いに異なるタグタンパク質配列を含む）。タグタンパク質の例は、限定されるものではないが、Ｆａｓｃｉｏｌａ ｈｅｐａｔｉｃａ ８ｋＤａ抗原（Ｆｈ８）、マルトース結合タンパク質（ＭＢＰ）、Ｎ利用物質（Ｎ－ｕｔｉｌｉｚａｔｉｏｎ ｓｕｂｓｔａｎｃｅ）（ＮｕｓＡ）、チオレドキシン（Ｔｒｘ）、ユビキチン類似因子（ＳＵＭＯ）、グルタチオンＳトランスフェラーゼ（ＧＳＴ）、溶解性エンハンサーペプチド配列（ＳＥＴ）、プロテインＧのＩｇＧドメインＢ１（ＧＢ１）、プロテインＡのＩｇＧ繰り返しドメインＺＺ（ＺＺ）、突然変異デハロゲナーゼ（ＨａｌｏＴａｇ）、溶解性エンハンシングユビキタスタグ（ＳＮＵＴ）、１７キロダルトンタンパク質（Ｓｋｐ）、Ｔ７ファージタンパク質キナーゼ（Ｔ７ＰＫ）、Ｅ．コリ（Ｅ．ｃｏｌｉ）選択プロテインＡ（ＥｓｐＡ）、モノマー形バクテリオファージＴ７ ０．３タンパク質（Ｏｒｃタンパク質；Ｍｏｃｒ）、Ｅ．コリ（Ｅ．ｃｏｌｉ）トリプシン阻害剤（Ｅｃｏｔｉｎ）、カルシウム結合タンパク質（ＣａＢＰ）、ストレス応答性ヒ酸レダクターゼ（ＡｒｓＣ）、翻訳開始因子ＩＦ２のＮ末端（ＩＦ２ドメインＩ）、ストレス応答性タンパク質（たとえば、ＲｐｏＡ、ＳｌｙＤ、Ｔｓｆ、ＲｐｏＳ、ＰｏｔＤ、Ｃｒｒ）、及びＥ．コリ（Ｅ．ｃｏｌｉ）酸性タンパク質（たとえば、ｍｓｙＢ、ｍｓｙＤ、ｒｐｏＤ）が挙げられる。たとえば、コスタ，Ｓ．（Ｃｏｓｔａ，Ｓ

．）ら著、「エシェリキア・コリ（Ｅｓｃｈｅｒｉｃｈｉａ　ｃｏｌｉ）におけるタンパク質溶解度、精製、及び免疫原性のための融合タグ：新規Ｆｈ８システム（Ｆｕｓｉｏｎ　ｔａｇｓ　ｆｏｒ　ｐｒｏｔｅｉｎ　ｓｏｌｕｂｉｌｉｔｙ，　ｐｕｒｉｆｉｃａｔｉｏｎ　ａｎｄ　ｉｍｍｕｎｏｇｅｎｉｃｉｔｙ　ｉｎ　Ｅｓｃｈｅｒｉｃｈｉａ　ｃｏｌｉ：　ｔｈｅ　ｎｏｖｅｌ　Ｆｈ８　ｓｙｓｔｅｍ）」、フロンティア・イン・マイクロバイオロジー（Ｆｒｏｎｔ　Ｍｉｃｒｏｂｉｏｌ．）、２０１４年２月１９日、第５巻、ｐ．６３（関連の内容は参照により本明細書に組み込まれる）を参照されたい。

【０１５４】

本明細書に記載されるように、本願のシールディングエレメントは、アミノ酸認識分子の標識成分により放出される放射及び／又は非放射減衰を有利に吸収、偏向、さもなければブロックしうる。そのため、融合ポリペプチドの好適なタンパク質シールドは、当業者によって容易に選択可能であることが認識されるべきである。たとえば、発明者らは、酵素（たとえば、ＤＮＡポリメラーゼ、グルタチオンＳトランスフェラーゼ）、輸送タンパク質（たとえば、マルトース結合タンパク質）、蛍光タンパク質（たとえば、ＧＦＰ）、及び商業的に入手可能なタグタンパク質（ＳＮＡＰ－ｔａｇ（登録商標））に融合したアミノ酸結合タンパク質を有するポリペプチドを含む融合ポリペプチドとの関連でのさまざまなタイプのタンパク質シールドの使用を実証してきた発明者らは、直列に配向されたタンパク質シールドの複数コピーを有する融合ポリペプチドの使用をさらに実証してきた。

【０１５５】

それゆえ、いくつかの実施形態では、本願は、１つ以上の直列配向タンパク質シールドに融合された１つ以上の直列配向アミノ酸結合タンパク質を有する融合ポリペプチドを提供する。いくつかの実施形態では、融合ポリペプチドが、２つ以上の直列配向バインダー及び／又は２つ以上の直列配向シールドを含む場合、２つ以上のバインダーのうちの１つの末端は、２つ以上のシールドのうちの１つの末端と末端同士で連結される。２つ以上のバインダーの直列コピーを有する融合ポリペプチドは、本明細書の他の箇所に記載され、いくつかの実施形態では、かかる融合物は、２つ以上のバインダーのうちの１つと末端同士で連結されたタンパク質シールドをさらに含む。

【０１５６】

いくつかの実施形態では、タンパク質シールド３３０は、認識分子と標識との間に非共有結合性連結基を形成する。たとえば、いくつかの実施形態では、タンパク質シールド３３０は、１つ以上のリガンド結合部位を含むモノマー形又はマルチマー形のタンパク質である。いくつかの実施形態では、非共有結合基は、１つ以上のリガンド結合部位に結合する１つ以上のリガンド部分を介して形成される。タンパク質シールドによって形成される非共有結合のさらなる例は、本明細書の他の箇所に記載される。

【０１５７】

第２のシールド構築物３０６は、第２のオリゴヌクレオチド鎖３３４にハイブリダイズされた第１のオリゴヌクレオチド鎖３３２を含む２本鎖核酸シールドの例を示す。示されるように、いくつかの実施形態では、２本鎖核酸シールドは、第１のオリゴヌクレオチド鎖３３２に装着された認識分子と、第２のオリゴヌクレオチド鎖３３４に装着された標識と、を含みうる。こうして、２本鎖核酸シールドは、オリゴヌクレオチド鎖ハイブリダイゼーションを介して認識分子と標識との間に非共有結合性連結基を形成する。いくつかの実施形態では、認識分子及び標識は、同一オリゴヌクレオチド鎖に装着可能であり、これは一本鎖核酸シールド又は他のオリゴヌクレオチド鎖とのハイブリダイゼーションを介して２本鎖核酸シールドを提供可能である。いくつかの実施形態では、鎖ハイブリダイゼーションは、認識分子と標識との間の分離をさらに増強するために連結基内の剛性の増加を提供可能である。

【０１５８】

シールディングエレメント３０２が核酸を含む場合、標識と認識分子との間の分離距離は、核酸上の装着部位間の距離により測定可能である（たとえば、直接装着又は１つ以上の追加のシールドポリマーを介するなどによる間接装着）。いくつかの実施形態では、核

酸上の装着部位間の距離は、標識と認識分子との間に存在する核酸内のヌクレオチドの数により測定可能である。ヌクレオチドの数とは、一本鎖核酸中のヌクレオチド塩基の数又は２本鎖核酸中のヌクレオチド塩基対の数のどちらかを意味しうることを理解すべきである。

【０１５９】

それゆえ、いくつかの実施形態では、認識分子の装着部位及び標識の装着部位は、５～２００ヌクレオチド（たとえば、５～１５０ヌクレオチド、５～１００ヌクレオチド、５～５０ヌクレオチド、１０～１００ヌクレオチド）だけ分離されうる。核酸中のいずれの位置も、認識分子、標識、又は１つ以上の追加のポリマーシールドに対する装着部位として機能しうることが認識されるべきである。いくつかの実施形態では、装着部位は、５’末端若しくは３’末端又はおおよそその位置或いは核酸の鎖に沿った内部位置でありうる。

【０１６０】

第２のシールド構築物３０６の非限定的構成は、鎖ハイブリダイゼーションを介して非共有結合性連結を形成するシールドの例を例示する。非共有結合性連結のさらなる例は、オリゴヌクレオチドシールド３３６を含む第３のシールド構築物３０８により例示される。いくつかの実施形態では、オリゴヌクレオチドシールド３３６は、認識分子に結合して非共有結合性連結を形成する核酸アプタマーである。いくつかの実施形態では、認識分子は、核酸アプタマーであり、且つオリゴヌクレオチドシールド３３６は、アプタマーにハイブリダイズして非共有結合性連結を形成するオリゴヌクレオチド鎖を含む。

【０１６１】

第４のシールド構築物３１０は、デンドリティックポリマーシールド３３８の例を示す。本明細書で用いられる場合、いくつかの実施形態では、デンドリティックポリマーとは、一般にポリオール又はデンドリマーを意味する。ポリオール及びデンドリマーは、当技術分野に記載されており、特定構成に最適化された分岐状デンドリティック構造を含みうる。いくつかの実施形態では、デンドリティックポリマーシールド３３８は、ポリエチレングリコール、テトラエチレングリコール、ポリ（アミドアミン）、ポリ（プロピレンイミン）、ポリ（プロピレンアミン）、カルボシラン、ポリ（Ｌ－リシン）、又はそれらの１つ以上の組合せを含む。

【０１６２】

デンドリマー又はデンドロンは、コアの周りが典型的には対称であり且つ球状３次元形態をとりうる繰返し分岐状分子である。たとえば、アストリュク（Ａｓｔｒｕｃ）ら著、２０１０年、ケミカル・レビューズ（Ｃｈｅｍ．Ｒｅｖ．）、第１１０巻、ｐ．１８５７を参照されたい。本願のシールドへのかかる構造の取込みは、標識とそれに会合する１つ以上のバイオ分子（たとえば、認識分子及び／又は認識分子に会合するポリペプチド）との間の接触の立体阻害を介して保護効果を提供可能である。分子の１次構造の変動を介するデンドリマーの化学的性質及び物理的性質のリファインメントは、デンドリマー表面の潜在的機能化を含めて要望に応じたシールディング効果の調整を可能にする。デンドリマーは、当技術分野で公知の広範にわたる材料及び分岐反応を用いてさまざまな技術により合成されうる。かかる合成変動は、必要に応じたデンドリマーの性質のカスタマイズを可能にする。本願のシールドに従って使用可能なポリオール及びデンドリマー化合物の例としては、限定されるものではないが米国特許出願公開第２０１８０３４６５０７号明細書に記載の化合物が挙げられる。

【０１６３】

図３Ｄは、本願のシールド認識分子のさらなる構成例を示す。タンパク質－核酸構築物３１２は、タンパク質の形態の１つ超のポリマーと２本鎖核酸とを含むシールドの例を示す。いくつかの実施形態では、シールドのタンパク質部分は、共有結合性連結を介してシールドの核酸部分に装着される。いくつかの実施形態では、装着は、非共有結合性連結を介する。たとえば、いくつかの実施形態では、シールドのタンパク質部分は、１価又は多価タンパク質のリガンド結合部位に装着されたリガンド部分を介して少なくとも１つの非共有結合性連結を形成する１価又は多価タンパク質である。いくつかの実施形態では、シ

ールドのタンパク質部分はアビジンタンパク質を含む。

【０１６４】

いくつかの実施形態では、本願のシールド認識分子はアビジン核酸構築物３１４である。いくつかの実施形態では、アビジン核酸構築物３１４は、アビジンタンパク質３４０と２本鎖核酸とを含むシールドを含む。本明細書に記載されるように、直接的又は本明細書に記載の１つ以上の追加のシールドポリマーを介するなどにより間接的のどちらかで、１つ以上のアミノ酸認識分子と１つ以上の標識との間に非共有結合性連結を形成するために、アビジンタンパク質３４０が使用されうる。

【０１６５】

アビジンタンパク質はビオチン結合タンパク質であり、一般にアビジンタンパク質の４つのサブユニットの各々にビオチン結合部位を有する。アビジンタンパク質としては、たとえば、アビジン、ストレプトアビジン、トラプタビジン、タマビジン、ブラダビジン、キセナビジン（ｘｅｎａｖｉｄｉｎ）、並びにそれらのホモログ及び変異体が挙げられる。いくつかの場合には、モノマー形、ダイマー形、又はテトラマー形のアビジンタンパク質を使用可能である。いくつかの実施形態では、アビジンタンパク質複合体のアビジンタンパク質は、テトラマー形（たとえばホモテトラマー）のストレプトアビジンである。いくつかの実施形態では、アビジンタンパク質のビオチン結合部位は、１つ以上のアミノ酸認識分子、１つ以上の標識、及び／又は本明細書に記載の１つ以上の追加のシールドポリマーに対する装着部位を提供する。

【０１６６】

アビジンタンパク質複合体の例示的な図は、図３Ｄのインセットパネルに示される。インセットパネルに示されるように、アビジンタンパク質３４０は、ビオチン部分（白丸として示される）に結合可能なタンパク質の４つのサブユニットの各々に結合部位３４２を含みうる。アビジンタンパク質３４０の多価性は、例示目的で一般的に示される各種連結構成を可能にしうる。たとえば、いくつかの実施形態では、アビジンタンパク質３４０への単一装着点を提供するためにビオチン連結部分３４４を使用可能である。いくつかの実施形態では、アビジンタンパク質３４０への２つの装着点を提供するためにビスビオチン連結部分３４６を使用可能である。アビジン核酸構築物３１４により例示されるように、アビジンタンパク質複合体は、認識分子と標識との間の分離距離の増加を提供するためにトランス構成を形成する２つのビスビオチン連結部分により形成されうる。

【０１６７】

アビジンタンパク質シールド構成の各種さらなる例が示される。第１のアビジン構築物３１６は、ビスビオチン連結部分を介して認識分子に及び個別ビオチン連結部分を介して２つの標識に装着されたアビジンシールドの例を示す。第２のアビジン構築物３１８は、個別ビオチン連結部分を介して２つ認識分子に及びビスビオチン連結部分を介して標識に装着されたアビジンシールドの例を示す。第３のアビジン構築物３２０は、個別ビオチン連結部分を介して２つ認識分子に及び核酸の各鎖のビオチン連結部分を介して標識付核酸に装着されたアビジンシールドの例を示す。第４のアビジン構築物３２２は、個別ビスビオチン連結部分を介して認識分子及び標識付核酸に装着されたアビジンシールドの例を示す。示されるように、標識は、標識と核酸との間のデンドリティックポリマーにより認識分子からさらにシールドされる。第５のアビジン構築物３２４は、２つのアビジンシールド認識分子に装着された内部標識３２５の例を示す。示されるように、各認識分子は、ビスビオチン連結部分を介して異なるアビジンタンパク質に装着され、内部標識３２６は、個別ビスビオチン連結部分を介して両方のアビジンタンパク質に装着される。

【０１６８】

図３Ａ～３Ｄに示されるシールド認識分子の構成例は、例示する目的で提供されることが認識されるべきである。本発明者らは、シールド認識分子の認識成分と標識成分との間に共有結合性又は非共有結合性の連結を形成する１つ以上の異なるポリマーを用いた各種他のシールド構成を構想した。例として、図３Ｅは、本願に係るシールド構成のモジュール性を例示する。

【０１６９】
図３Ｅのトップに示されるように、シールド認識分子は、一般に認識成分３５０とシールディングエレメント３５２と標識成分３５４とを含む。例示を容易にするために、認識成分３５０は、１つのアミノ酸認識分子として描かれ、且つ標識成分３５４は、１つの標識として描かれる。
【０１７０】
本願のシールド認識分子は、１つ以上のアミノ酸認識分子と１つ以上の標識とに装着されたシールディングエレメント３５２を含みうることが認識されるべきである。認識成分３５０が１つ超の認識分子を含む場合、各認識分子は、シールディングエレメント３５２上の１つ以上の装着部位でシールディングエレメント３５２に装着可能である。いくつかの実施形態では、認識成分３５０は、本明細書の他の箇所に記載されるように、２つ以上のアミノ酸結合タンパク質の直列コピーを有する単一ポリペプチド融合構築物を含む。標識成分３５４が１つ超の標識を含む場合、各標識は、シールディングエレメント３５２上の１つ以上の装着部位でシールディングエレメント３５２に装着可能である。標識成分３５４は、単一の装着点を有するように一般的に示されるが、この態様に限定されない。たとえば、いくつかの実施形態では、アビジン構築物３２４によって例示されるように、１つ超の装着点を有する内部標識が使用されて１つ超の認識成分３５０及び／又はシールディングエレメント３５２に連結してもよい。
【０１７１】
いくつかの実施形態では、シールディングエレメント３５２はタンパク質３６０を含む。いくつかの実施形態では、タンパク質３６０は１価又は多価タンパク質である。いくつかの実施形態では、タンパク質３６０は、モノマー又はマルチマータンパク質、たとえば、タンパク質ホモダイマー、タンパク質ヘテロダイマー、タンパク質オリゴマー、又は他のタンパク質分子である。いくつかの実施形態では、シールディングエレメント３５２は、少なくとも１つの他の分子に非共有結合されたタンパク質により形成されるタンパク質複合体を含む。たとえば、いくつかの実施形態では、シールディングエレメント３５２は、タンパク質－タンパク質複合体３６２を含む。いくつかの実施形態では、タンパク質－タンパク質複合体３６２は、他のタンパク質分子に特異的に結合された１つのタンパク質分子を含む。いくつかの実施形態では、タンパク質－タンパク質複合体３６２は、抗原に結合された抗体又は抗体フラグメント（たとえばｓｃＦｖ）を含む。いくつかの実施形態では、タンパク質－タンパク質複合体３６２は、タンパク質リガンドに結合されたレセプターを含む。タンパク質－タンパク質複合体の追加の例としては、限定されるものではないが、トリプシン－アプロチニン、バルナーゼ－バルスター、及びコリシンＥ９－Ｉｍ９免疫タンパク質が挙げられる。
【０１７２】
いくつかの実施形態では、シールディングエレメント３５２は、タンパク質－リガンド複合体３６４を含む。いくつかの実施形態では、タンパク質－リガンド複合体３６４は、１価タンパク質と非タンパク質リガンド部分とを含む。たとえば、いくつかの実施形態では、タンパク質－リガンド複合体３６４は、低分子阻害剤部分に結合された酵素を含む。いくつかの実施形態では、タンパク質－リガンド複合体３６４は、非タンパク質リガンド部分に結合されたレセプターを含む。
【０１７３】
いくつかの実施形態では、シールディングエレメント３５２は、１つ以上のリガンド部分に非共有結合された多価タンパク質により形成される多価タンパク質複合体を含む。いくつかの実施形態では、シールディングエレメント３５２は、１つ以上のビオチン連結部分に非共有結合されたアビジンタンパク質により形成されるアビジンタンパク質複合体を含む。構築物３６６、３６８、３７０、及び３７２は、アビジンタンパク質複合体の例示的例を提供するとともに、そのうちのいずれか１つ以上は、シールディングエレメント３５２に取り込まれうる。
【０１７４】

いくつかの実施形態では、シールディングエレメント３５２は、２つのビスビオチン連結部分に結合されたアビジンタンパク質を含むツーウェイアビジン複合体３６６を含む。いくつかの実施形態では、シールディングエレメント３５２は、２つのビオチン連結部分とビスビオチン連結部分とに結合されたアビジンタンパク質を含むスリーウェイアビジン複合体３６８を含む。いくつかの実施形態では、シールディングエレメント３５２は、４つのビオチン連結部分に結合されたアビジンタンパク質を含むフォーウェイアビジン複合体３７０を含む。

【０１７５】

いくつかの実施形態では、シールディングエレメント３５２は、アビジンタンパク質に工学操作により導入された１つ又は２つの非機能性結合部位を含むアビジンタンパク質を含む。たとえば、いくつかの実施形態では、シールディングエレメント３５２は、２つのサブユニットの各々でビオチン連結部分に結合されたアビジンタンパク質を含む２価アビジン複合体３７２を含み、この場合、アビジンタンパク質は、２つの他のサブユニットの各々で非機能性リガンド結合部位３４８を含む。示されるように、いくつかの実施形態では、２価アビジン複合体３７２はトランス２価アビジンタンパク質を含むが、所望の実現形態に依存してシス２価アビジンタンパク質が使用されうる。いくつかの実施形態では、アビジンタンパク質は３価アビジンタンパク質である。いくつかの実施形態では、３価アビジンタンパク質は、１つのサブユニットに非機能性リガンド結合部位３４８を含むとともに、３つのビオチン連結部分に結合されるか又は１つのビオチン連結部分及び他のサブユニットでビスビオチン連結部分に結合される。

【０１７６】

いくつかの実施形態では、シールディングエレメント３５２は、デンドリティックポリマー３７４を含む。いくつかの実施形態では、デンドリティックポリマー３７４は、本明細書の他の箇所に記載されるようにポリオール又はデンドリマーである。いくつかの実施形態では、デンドリティックポリマー３７４は、分岐状ポリオール又は分岐状デンドリマーである。いくつかの実施形態では、デンドリティックポリマー３７４は、単糖－ＴＥＧ、二糖、Ｎ－アセチル単糖、ＴＥＭＰＯ－ＴＥＧ、トロロックス－ＴＥＧ又はグリセロールデンドリマーを含む。本願のシールド認識分子に係る有用なポリオールの例としては、ポリエーテルポリオール及びポリエステルポリオール、たとえば、ポリエチレングリコール、ポリプロピレングリコール、及び当技術分野で周知の類似のかかるポリマーが挙げられる。いくつかの実施形態では、デンドリティックポリマー３７４は、次式：－（$CH_2CH_2O)_n$－（式中、ｎは１～５００（両端の値を含む）の整数である）の化合物を含む。いくつかの実施形態では、デンドリティックポリマー３７４は、次式：－（$CH_2CH_2O)_n$－（式中、ｎは１～１００（両端の値を含む）の整数である）の化合物を含む。

【０１７７】

いくつかの実施形態では、シールディングエレメント３５２は核酸を含む。いくつかの実施形態では、核酸は一本鎖である。いくつかの実施形態では、標識成分３５４は、一本鎖核酸の一方の末端（たとえば、５’末端又は３’末端）に直接的又は間接的に装着され、且つ認識成分３５０は、一本鎖核酸の他方の末端（たとえば、３’末端又は５’末端）に直接的又は間接的に装着される。たとえば、一本鎖核酸は、核酸の５’末端に装着された標識と、核酸の３’末端に装着されたアミノ酸認識分子と、を含みうる。

【０１７８】

いくつかの実施形態では、シールディングエレメント３５２は二本鎖核酸３７６を含む。示されるように、いくつかの実施形態では、二本鎖核酸３７６は、鎖ハイブリダイゼーションを介して認識成分３５０と標識成分３５４との間に非共有結合性連結を形成可能である。しかしながら、いくつかの実施形態では、二本鎖核酸３７６は、同一オリゴヌクレオチド鎖への装着を介して認識成分３５０と標識成分３５４との間に共有結合性連結を形成可能である。いくつかの実施形態では、標識成分３５４は、二本鎖核酸の一方の末端に直接的又は間接的に装着され、且つ認識成分３５０は、二本鎖核酸の他方の末端に直接的

又は間接的に装着される。たとえば、二本鎖核酸は、一方の鎖の５’末端に装着された標識と、他方の鎖の５’末端に装着されたアミノ酸認識分子と、を含みうる。

【０１７９】

いくつかの実施形態では、シールディングエレメント３５２は、シールドの立体バルクを増加させるのに有用でありうる１つ以上の構造モチーフを形成する核酸を含む。核酸構造モチーフの例としては、限定されるものではないが、ステム－ループ、スリーウェイジャンクション（たとえば、２つ以上のステム－ループモチーフにより形成される）、フォーウェイジャンクション（たとえばホリデイジャンクション）、及びバルジループが挙げられる。

【０１８０】

いくつかの実施形態では、シールディングエレメント３５２は、ステム－ループ３７８を形成する核酸を含む。ステム－ループ又はヘアピンループは、オリゴヌクレオチド鎖がフォールドして同一鎖の他のセクションと塩基対を形成するときに形成されるオリゴヌクレオチド鎖上のヌクレオチドの不対ループである。いくつかの実施形態では、ステム－ループ３７８の不対ループは、３～１０ヌクレオチドを含む。それゆえ、ステム－ループ３７８は、ハイブリダイズしてステムを形成する反転相補配列を有するオリゴヌクレオチド鎖の２つの領域による形成可能であり、この場合、２つの領域は、不対ループを形成する３～１０ヌクレオチドにより分離される。いくつかの実施形態では、ステム－ループ３７８のステムは、Ａ／Ｔヌクレオチドと比較して生じる追加の水素結合相互作用により安定性の増加を提供可能な１つ以上のＧ／Ｃヌクレオチドを有するように設計可能である。いくつかの実施形態では、ステム－ループ３７８のステムは、不対ループ配列に直接近接するＧ／Ｃヌクレオチドを含む。いくつかの実施形態では、ステム－ループ３７８のステムは、不対ループ配列に近接する最初の２、３、４、又は５ヌクレオチド内にＧ／Ｃヌクレオチドを含む。いくつかの実施形態では、ステム－ループ３７８の不対ループは、１つ以上の装着部位を含む。いくつかの実施形態では、装着部位は、不対ループ中の脱塩基部位に発生する。いくつかの実施形態では、装着部位は、不対ループの塩基に発生する。

【０１８１】

いくつかの実施形態では、ステム－ループ３７８は、二本鎖核酸により形成される。本明細書に記載されるように、いくつかの実施形態では、二本鎖核酸は、第１及び第２のオリゴヌクレオチド鎖の鎖ハイブリダイゼーションを介して非共有結合性連結基を形成可能である。しかしながら、いくつかの実施形態では、シールディングエレメント３５２は、ステム－ループモチーフを形成する一本鎖核酸を含んで共有結合性連結基を提供する。いくつかの実施形態では、シールディングエレメント３５２は、２つ以上のステム－ループモチーフを形成する核酸を含む。たとえば、いくつかの実施形態では、核酸は２つのステム－ループモチーフを含む。いくつかの実施形態では、モチーフが一緒になってスリーウェイジャンクションを形成するように、１つのステム－ループモチーフのステムは他のステムに近接する。いくつかの実施形態では、シールディングエレメント３５２は、フォーウェイジャンクション３７８を形成する核酸を含む。いくつかの実施形態では、フォーウェイジャンクション３７８は、２つ以上のオリゴヌクレオチド鎖（たとえば、２、３、又は４つのオリゴヌクレオチド鎖）のハイブリダイゼーションを介して形成される。

【０１８２】

いくつかの実施形態では、シールディングエレメント３５２は、図３Ｅの３６０、３６２、３６４、３６６、３６８、３７０、３７２、３７４、３７６、３７８、及び３８０から選択される１つ以上のポリマーを含む。３６０、３６２、３６４、３６６、３６８、３７０、３７２、３７４、３７６、３７８、及び３８０の各々に示される連結部分及び装着部位は、例示目的で示されており、好ましい連結部位又は装着部位の構成を描くことを意図するものではないことが認識されるべきである。

【０１８３】

いくつかの態様では、本願は、式（ＩＩ）

$$A-(Y)_n-D$$

（ＩＩ）
（式中、Ａは、少なくとも１つのアミノ酸認識分子を含むアミノ酸結合成分であり；Ｙの各々は、共有結合性又は非共有結合性連結基を形成するポリマーであり；ｎは、１～１０（両端の値を含む）の整数であり；且つ、Ｄは、少なくとも１つの検出可能標識を含む標識成分である）
のアミノ酸認識分子を提供する。いくつかの実施形態では、本願は、式（ＩＩ）の可溶性アミノ酸認識分子を含む組成物を提供する。

【０１８４】
いくつかの実施形態では、Ａは、複数のアミノ酸認識分子を含む。いくつかの実施形態では、複数の各アミノ酸認識分子は、Ｙ上の異なる装着部位に装着される。いくつかの実施形態では、複数のアミノ酸認識分子の少なくとも２つは、Ｙ上の単一装着部位に装着される。いくつかの実施形態では、アミノ酸認識分子は、認識タンパク質又は核酸アプタマー、たとえば、本明細書の他の箇所に記載のものである。

【０１８５】
いくつかの実施形態では、検出可能標識は、発光標識又は伝導率標識である。いくつかの実施形態では、発光標識は、少なくとも１つのフルオロフォア色素分子を含む。いくつかの実施形態では、Ｄは、２０以下のフルオロフォア色素分子を含む。いくつかの実施形態では、フルオロフォア色素分子の数とアミノ酸認識分子の数との比は、１：１～２０：１である。いくつかの実施形態では、発光標識は、ドナー標識とアクセプター標識とを含む少なくとも１つのＦＲＥＴペアを含む。いくつかの実施形態では、ドナー標識とアクセプター標識との比は、１：１、２：１、３：１、４：１、又は５：１である。いくつかの実施形態では、アクセプター標識とドナー標識との比は、１：１、２：１、３：１、４：１、又は５：１である。

【０１８６】
いくつかの実施形態では、Ｄは２０ｎｍ（２００Å）未満の直径である。いくつかの実施形態では、－（Ｙ）$_{n}$－は少なくとも２ｎｍの長さである。いくつかの実施形態では、－（Ｙ）$_{n}$－は少なくとも５ｎｍの長さである。いくつかの実施形態では、－（Ｙ）$_{n}$－は少なくとも１０ｎｍの長さである。いくつかの実施形態では、Ｙの各々は、独立して、バイオ分子、ポリオール、又はデンドリマーである。いくつかの実施形態では、バイオ分子は、核酸、ポリペプチド、又は多糖である。

【０１８７】
いくつかの実施形態では、アミノ酸認識分子は、以下の式：
Ａ－Ｙ^{1}－（Ｙ）$_{m}$－Ｄ又はＡ－（Ｙ）$_{m}$－Ｙ^{1}－Ｄ
（式中、Ｙ^{1}は、核酸又はポリペプチドであり、且つｍは、０～１０（両端の値を含む）の整数である）
のうちの１つである。

【０１８８】
いくつかの実施形態では、核酸は、第１のオリゴヌクレオチド鎖を含む。いくつかの実施形態では、核酸は、第１のオリゴヌクレオチド鎖にハイブリダイズされた第２のオリゴヌクレオチド鎖を含む。いくつかの実施形態では、核酸は、第１のオリゴヌクレオチド鎖を介して共有結合性連結を形成する。いくつかの実施形態では、核酸は、ハイブリダイズされた第１及び第２のオリゴヌクレオチド鎖を介して非共有結合性連結を形成する。

【０１８９】
いくつかの実施形態では、ポリペプチドは、１価又は多価タンパク質である。いくつかの実施形態では、１価又は多価タンパク質は、１価又は多価タンパク質のリガンド結合部位に装着されたリガンド部分を介して少なくとも１つの非共有結合性連結を形成する。いくつかの実施形態では、Ａ、Ｙ、又はＤは、リガンド部分を含む。

【０１９０】
いくつかの実施形態では、アミノ酸認識分子は、以下の式：
Ａ－（Ｙ）$_{m}$－Ｙ^{2}－Ｄ又はＡ－Ｙ^{2}－（Ｙ）$_{m}$－Ｄ

（式中、Y^2は、ポリオール又はデンドリマーであり、且つｍは、０～１０（両端の値を含む）の整数である）
のうちの１つである。いくつかの実施形態では、ポリオール又はデンドリマーは、ポリエチレングリコール、テトラエチレングリコール、ポリ（アミドアミン）、ポリ（プロピレンイミン）、ポリ（プロピレンアミン）、カルボシラン、ポリ（Ｌ－リシン）、又はそれらの１つ以上の組合せを含む。

【０１９１】

いくつかの態様では、本願は、式（ＩＩＩ）：

Ａ－Y^1－Ｄ

（ＩＩＩ）

（式中、Ａは、少なくとも１つのアミノ酸認識分子を含むアミノ酸結合成分であり；Y^1は、核酸又はポリペプチドであり；Ｄは、少なくとも１つの検出可能標識を含む標識成分である）
のアミノ酸認識分子を提供する。いくつかの実施形態では、Y^1が核酸であるとき、核酸は共有結合性又は非共有結合性連結基を形成する。いくつかの実施形態では、Y^1がポリペプチドであるとき、ポリペプチドは、５０×１０^{-9}Ｍ未満の解離定数（K_D）により特徴付けられる非共有結合性連結基を形成する。

【０１９２】

いくつかの実施形態では、Y^1は、第１のオリゴヌクレオチド鎖を含む核酸である。いくつかの実施形態では、核酸は、第１のオリゴヌクレオチド鎖にハイブリダイズされた第２のオリゴヌクレオチド鎖を含む。いくつかの実施形態では、Ａは、第１のオリゴヌクレオチド鎖に装着され、且つＤは、第２のオリゴヌクレオチド鎖に装着される。いくつかの実施形態では、Ａは、第１のオリゴヌクレオチド鎖上の第１の装着部位に装着され、且つＤは、第１のオリゴヌクレオチド鎖上の第２の装着部位に装着される。いくつかの実施形態では、核酸の各オリゴヌクレオチド鎖は、１５０未満、１００未満、又は５０未満のヌクレオチドを含む。

【０１９３】

いくつかの実施形態では、Y^1は、１価又は多価タンパク質である。いくつかの実施形態では、１価又は多価タンパク質は、１価又は多価タンパク質のリガンド結合部位に装着されたリガンド部分を介して少なくとも１つの非共有結合性連結を形成する。いくつかの実施形態では、Ａ及びＤの少なくとも１つは、リガンド部分を含む。いくつかの実施形態では、ポリペプチドは、アビジンタンパク質（たとえば、アビジン、ストレプトアビジン、トラプタビジン、タマビジン、ブラダビジン、キセナビジン（ｘｅｎａｖｉｄｉｎ）、又はそれらのホモログ若しくは変異体）である。いくつかの実施形態では、リガンド部分はビオチン部分である。

【０１９４】

いくつかの実施形態では、アミノ酸認識分子は、以下の式：

Ａ－Y^1－（Ｙ）$_n$－Ｄ又はＡ－（Ｙ）$_n$－Y^1－Ｄ

（式中、Ｙの各々は、共有結合性又は非共有結合性連結基を形成するポリマーであり、且つｎは、１～１０（両端の値を含む）の整数である）
のうちの１つである。いくつかの実施形態では、Ｙの各々は、独立して、バイオ分子、ポリオール、又はデンドリマーである。

【０１９５】

他の態様では、本願は、核酸と、核酸上の第１の装着部位に装着された少なくとも１つのアミノ酸認識分子と、核酸上の第２の装着部位に装着された少なくとも１つの検出可能標識と、を含むアミノ酸認識分子を提供する。いくつかの実施形態では、核酸は、少なくとも１つのアミノ酸認識分子と少なくとも１つの検出可能標識との間に共有結合性又は非共有結合性連結基を形成する。

【０１９６】

いくつかの実施形態では、核酸は、第２のオリゴヌクレオチド鎖にハイブリダイズされ

た第１のオリゴヌクレオチド鎖を含む二本鎖核酸である。いくつかの実施形態では、第１の装着部位は、第１のオリゴヌクレオチド鎖上にあり、且つ第２の装着部位は、第２のオリゴヌクレオチド鎖上にある。いくつかの実施形態では、少なくとも１つのアミノ酸認識分子は、少なくとも１つのアミノ酸認識分子と核酸との間に共有結合性又は非共有結合性連結基を形成するタンパク質を介して第１の装着部位に装着される。いくつかの実施形態では、少なくとも１つの検出可能標識は、少なくとも１つの検出可能標識と核酸との間に共有結合性又は非共有結合性連結基を形成するタンパク質を介して第２の装着部位に装着される。いくつかの実施形態では、第１及び第２の装着部位は、核酸上の５～１００ヌクレオチド塩基又はヌクレオチド塩基対だけ分離される。

【０１９７】

さらに他の態様では、本願は、少なくとも２つのリガンド結合部位を含む多価タンパク質と、タンパク質上の第１のリガンド結合部位に結合された第１のリガンド部分を介してタンパク質に装着された少なくとも１つのアミノ酸認識分子と、タンパク質上の第２のリガンド結合部位に結合された第２のリガンド部分を介してタンパク質に装着された少なくとも１つの検出可能標識と、を含むアミノ酸認識分子を提供する。

【０１９８】

いくつかの実施形態では、多価タンパク質は、４つのリガンド結合部位を含むアビジンタンパク質である。いくつかの実施形態では、リガンド結合部位は、ビオチン結合部位であり、且つリガンド部分は、ビオチン部分である。いくつかの実施形態では、ビオチン部分の少なくとも１つは、ビスビオチン部分であり、且つビスビオチン部分は、アビジンタンパク質上の２つビオチン結合部位に結合される。いくつかの実施形態では、少なくとも１つのアミノ酸認識分子は、第１のリガンド部分を含む核酸を介してタンパク質に装着される。いくつかの実施形態では、少なくとも１つの検出可能標識は、第２のリガンド部分を含む核酸を介してタンパク質に装着される。

【０１９９】

本明細書の他の箇所に記載されるように、本願のシールド認識分子は、本願に係るポリペプチドシーケンシング方法又は当技術分野で公知のいずれかの方法で使用されうる。たとえば、いくつかの実施形態では、本明細書に提供されるシールド認識分子は、本明細書に提供されるか又はポリペプチドシーケンシング反応で複数の反応混合物の繰返しサイクリングを含みうる当技術分野で従来から知られるエドマン型分解反応で使用されうる。いくつかの実施形態では、本明細書に提供されるシールド認識分子は、単一反応混合物でアミノ酸認識及び分解を含む本願の動的シーケンシング反応で使用されうる。

【０２００】

ポリペプチドシーケンシング

ポリペプチドの末端アミノ酸を同定する方法に加えて、本願は、標識付認識分子を使用してポリペプチドをシーケンシングする方法を提供する。いくつかの実施形態では、シーケンシングの方法は、ポリペプチド末端を、末端アミノ酸検出及び末端アミノ酸切断の繰返しサイクルにさらすことを含みうる。たとえば、いくつかの実施形態では、本願は、ポリペプチドと本明細書に記載の１つ以上の標識付認識分子とを接触させること、及びポリペプチドをエドマン分解にさらすことを含む、ポリペプチドのアミノ酸配列を決定する方法を提供する。

【０２０１】

従来のエドマン分解は、ポリペプチドの末端アミノ酸を修飾及び切断する繰返しサイクルを含み、各連続して切断されるアミノ酸が同定されて、ポリペプチドのアミノ酸配列を決定する。従来のエドマン分解の例示的例のように、ポリペプチドのＮ末端アミノ酸が、フェニルイソチオシアネート（ＰＩＴＣ）を使用して修飾され、ＰＩＴＣ誘導Ｎ末端アミノ酸を形成する。次いで、ＰＩＴＣ誘導Ｎ末端アミノ酸は、酸性条件、塩基性条件、及び／又は高温を使用して切断される。また、ＰＩＴＣ誘導Ｎ末端アミノ酸を切断する工程は、中性のｐＨ又は中性に近いｐＨの比較的穏やかな切断条件を含む、原虫トリパノソーマ・クルーズ（Ｔｒｙｐａｎｏｓｏｍａ　ｃｒｕｚｉ）由来の修飾システインプロテアーゼ

を使用して酵素的に達成されてもよいことが示されている。有用な酵素の非限定的例は、「繰返しポリペプチド分析及び処理のための分子及び方法（ＭＯＬＥＣＵＬＥＳ ＡＮＤ ＭＥＴＨＯＤＳ ＦＯＲ ＩＴＥＲＡＴＩＶＥ ＰＯＬＹＰＥＰＴＩＤＥ ＡＮＡＬＹＳＩＳ ＡＮＤ ＰＲＯＣＥＳＳＩＮＧ）」という名称の２０１６年９月２日出願の米国特許出願第１５／２５５，４３３号明細書に記載されている。

【０２０２】

本願にかかる標識付認識分子を使用したエドマン分解によるシーケンシングの例は、図４に描かれる。いくつかの実施形態では、エドマン分解によるシーケンシングは、リンカー４２４を介して固形担体の表面４３０に固定される（たとえば、サンプルウェルのボトム表面又は側壁表面）ポリペプチド４２０を提供することを含む。いくつかの実施形態では、本明細書に記載されるように、ポリペプチド４２０は一方の末端（たとえば、アミノ末端アミノ酸又はカルボキシ末端アミノ酸）で固定されるため、他方の末端は、末端アミノ酸の検出及び切断のために遊離している。それゆえ、いくつかの実施形態では、本明細書に記載のエドマン分解法で使用される試薬は、ポリペプチド４２０の非固定（たとえば、遊離）末端で末端アミノ酸と選択的に相互作用する。こうして、ポリペプチド４２０は、検出及び切断の繰返しサイクルにわたって固定されたままである。この目的で、いくつかの実施形態では、リンカー４２４は、たとえば、化学的切断条件下での表面４３０からのポリペプチド４２０の離脱を制限するなどのために、検出及び切断に使用される条件の所望のセットに従って設計されうる。ポリペプチドを表面に固定するために好適なリンカー組成及び技術は、本明細書の他の箇所に記載されている。

【０２０３】

本願によれば、いくつかの実施形態では、エドマン分解によるシーケンシングの方法は、ペプチド４２０と、１タイプ以上の末端アミノ酸に選択的に結合する１つ以上の標識付認識分子とを接触させる、工程（Ｉ）を含む。示されるように、いくつかの実施形態では、標識付認識分子４００は、末端アミノ酸に選択的に結合することによってポリペプチド４２０と相互作用する。いくつかの実施形態では、工程（Ｉ）は、ポリペプチド４２０の末端アミノ酸（たとえば、遊離アミノ酸）に選択的に結合しない１つ以上の標識付認識分子のいずれかを除去することをさらに含む。

【０２０４】

いくつかの実施形態では、方法は、標識付認識分子４００を検出することによってポリペプチド４２０の末端アミノ酸を同定することをさらに含む。いくつかの実施形態では、検出は、標識付認識分子４００からの発光を検出することを含む。本明細書に記載されるように、いくつかの実施形態では、発光は、標識付認識分子４００とユニークに関連付けられるため、発光は、標識付認識分子４００が選択的に結合するアミノ酸のタイプと関連付けられる。このように、いくつかの実施形態では、このアミノ酸のタイプは、標識付認識分子４００の１つ以上の発光性を決定することによって同定される。

【０２０５】

いくつかの実施形態では、エドマン分解によるシーケンシングの方法は、ポリペプチド４２０の末端アミノ酸を除去する、工程（２）を含む。いくつかの実施形態では、工程（２）は、ポリペプチド４２０から標識付認識分子４００（たとえば、末端アミノ酸に選択的に結合する１つ以上の標識付認識分子のいずれか）を除去することを含む。いくつかの実施形態では、工程（２）は、末端アミノ酸と、イソチオシアネート（たとえば、ＰＩＴＣ）とを接触させてイソチオシアネート修飾末端アミノ酸を形成することによって、ポリペプチド４２０の末端アミノ酸（たとえば、有利アミノ酸）を修飾することを含む。いくつかの実施形態では、イソチオシアネート修飾末端アミノ酸は、非修飾末端アミノ酸よりも切断試薬（たとえば、化学的又は酵素的切断試薬）による除去の影響を受けやすい。

【０２０６】

いくつかの実施形態では、工程（２）は、ポリペプチド４２０と、イソチオシアネート修飾末端アミノ酸に特異的に結合及び切断するプロテアーゼ４４０とを接触させることによって末端アミノ酸を除去することを含む。いくつかの実施形態では、プロテアーゼ４４

０は、修飾システインプロテアーゼを含む。いくつかの実施形態では、プロテアーゼ４４０は、トリパノソーマ・クルーズ（Ｔｒｙｐａｎｏｓｏｍａ　ｃｒｕｚｉ）由来のシステインプロテアーゼなどの修飾システインプロテアーゼを含む（たとえば、ボーゴ（Ｂｏｒｇｏ）ら著、２０１５年、プロテイン・サイエンス（Ｐｒｏｔｅｉｎ　Ｓｃｉｅｎｃｅ）、２４巻、ｐ．５７１～５７９を参照されたい）。さらに他の実施形態では、工程（２）は、ポリペプチド４２０を、イソチオシアネート修飾末端アミノ酸を切断するのに十分な化学的（たとえば、酸性、塩基性）条件にさらすことによって末端アミノ酸を除去することを含む。

【０２０７】

いくつかの実施形態では、エドマン分解によるシーケンシングの方法は、末端アミノ酸切断の後に、ポリペプチド４２０を洗浄する、工程（３）を含む。いくつかの実施形態では、洗浄は、プロテアーゼ４４０を除去することを含む。いくつかの実施形態では、洗浄は、中性のｐＨ条件にポリペプチド４２０を回復させることを含む（たとえば、酸性条件又は塩基性条件による化学的切断の後に）。いくつかの実施形態では、エドマン分解によるシーケンシングに方法は、複数のサイクルに工程（１）から（３）を繰返すことを含む。

【０２０８】

いくつかの態様では、本願は、末端アミノ酸と、標識付アミノ酸認識分子及び標識切断分子（たとえば、標識非特異的エキソペプチダーゼ）との結合相互作用を評価することによる、リアルタイムのポリペプチドシーケンシングの方法を提供する。図５は、個別の会合イベントがシグナル出力５００のシグナルパルスを生じさせるシーケンシングの方法の例を示す。図５のインセットパネルは、このアプローチによるリアルタイムシーケンシングの一般的なスキームを例示する。示されるように、標識付認識分子５１０は、選択的に、末端アミノ酸（明細書ではリシンとして示される）と会合（たとえば、結合）し、末端アミノ酸から解離する。これにより、末端アミノ酸を同定するために使用されうるシグナルパルス５００における一連のパルスが生じる。いくつかの実施形態では、一連のパルスは、対応する末端アミノ酸のアイデンティティーの診断となりうるパルスパターン（たとえば、特性パターン）を提供する。

【０２０９】

理論により拘束されることを望むものではないが、標識付認識分子５１０は、会合速度すなわち結合の「ｏｎ」速度（k_{on}）及び解離速度すなわち結合の「ｏｆｆ」速度（k_{off}）によって定義される結合親和性（K_D）に従って選択的に結合する。速度定数k_{off}及びk_{on}は、パルス持続時間（たとえば、検出可能な会合イベントに対応する時間）及びパルス間持続時間（たとえば、検出可能会合イベント同士の間の時間）それぞれの重要な決定因子である。いくつかの実施形態では、これらの速度は、最高のシーケンシング確度を与えるパルス持続時間及びパルス速度（たとえば、シグナルパルスの頻度）を達成するために工学操作されうる。

【０２１０】

インセットパネルに示すように、シーケンシング反応混合物は、標識付認識分子５１０のものとは異なる発光標識を含む標識付非特異的エキソペプチダーゼ５２０をさらに含む。いくつかの実施形態では、標識付非特異的エキソペプチダーゼ５２０は、標識付認識分子５１０の濃度よりも低い濃度で混合物中に存在する。いくつかの実施形態では、標識付特異的エキソペプチダーゼ５２０は、ほとんど又はすべてのタイプのアミノ酸を切断するように広範な特異性を示す。それゆえ、動的シーケンシングアプローチは、エキソペプチダーゼ切断活性によって触媒される分解反応の過程全体にわたり、ポリペプチドの末端に結合する認識分子をモニタリングすることを含みうる。

【０２１１】

シグナル出力５００の進行によって例示されるように、いくつかの実施形態では、標識付非特異的エキソペプチダーゼ５２０による末端アミノ酸切断はシグナルパルスを生じさせ、これらのイベントは、標識付認識分子５１０の結合パルスよりも低頻度で発生する。

こうして、ポリペプチドのアミノ酸は、リアルタイムシーケンシングプロセスにてカウント及び／又は同定しうる。シグナル出力５００にさらに例示されるように、いくつかの実施形態では、複数の標識付認識分子が使用されてもよく、各々は対応する末端アミノ酸を同定するために使用されうる診断パルスパターン（たとえば、特性パターン）を有する。たとえば、いくつかの実施形態では、互いに異なる特性パターン（シグナル出力５００においてリシン、フェニルアラニン、及びグルタミンの各々によって例示されるような）は、異なるタイプの末端アミノ酸を有する１つ以上の標識付認識分子の会合に対応する。本明細書に記載されるように、１タイプより多くのアミノ酸と会合する単一認識分子は、本願に従って使用されうることが認識されるべきである。それゆえ、いくつかの実施形態では、互いに異なる特性パターンは、１つの標識付認識分子と、互いに異なるタイプの末端アミノ酸との会合に対応する。

【０２１２】

本明細書に記載されるように、シグナルパルス情報は、一連のシグナルパルスの特性パターンに基づいてアミノ酸を同定するために使用されうる。いくつかの実施形態では、特性パターンは、各シグナルパルスがパルス持続時間を含む、複数のシグナルパルスを含む。いくつかの実施形態では、複数のシグナルパルスは、特性パターンにおけるパルス持続時間の分布の要的統計（たとえば、平均、中央値、時間減衰定数）によって特徴付けられうる。いくつかの実施形態では、特性パターンの平均パルス持続時間は、約１ミリ秒～約１０秒（たとえば、約１ｍｓ～１ｓ、約１ｍｓ～約１００ｍｓ、約１ｍｓ～約１０ｍｓ、約１０ｍｓ～約１０ｓ、約１００ｍｓ～約１０ｓ、約１ｓ～約１０ｓ、約１０ｍｓ～約１００ｍｓ、又は約１００ｍｓ～約５００ｍｓ）である。いくつかの実施形態では、平均パルス持続時間は、約５０ミリ秒～約２秒、約５０ミリ秒～約５００ミリ秒、又は約５００ミリ秒～約２秒である。

【０２１３】

いくつかの実施形態では、単一ポリペプチドの互いに異なるタイプのアミノ酸に対応する互いに異なる特性パルスパターンは、要的統計における統計学的有意差に基づいて互いに識別可能である。たとえば、いくつかの実施形態では、一方の特性パターンは、少なくとも１０ミリ秒（たとえば、約１０ｍｓ～約１０ｓ、約１０ｍｓ～約１ｓ、約１０ｍｓ～約１００ｍｓ、約１００ｍｓ～約１０ｓ、約１ｓ～約１０ｓ、又は約１００ｍｓ～約１ｓ）の平均パルス持続時間の差に基づいて他方の特性パターンと識別可能である。いくつかの実施形態では、平均パルス持続時間の差は、少なくとも５０ｍｓ、少なくとも１００ｍｓ、少なくとも２５０ｍｓ、少なくとも５００ｍｓ、又はそれ以上である。いくつかの実施形態では、平均パルス持続時間の差は、約５０ｍｓ～約１ｓ、約５０ｍｓ～約５００ｍｓ、約５０ｍｓ～約２５０ｍｓ、約１００ｍｓ～約５００ｍｓ、約２５０ｍｓ～約５００ｍｓ、又は約５００ｍｓ～約１ｓである。いくつかの実施形態では、一方の特性パターンの平均パルス持続時間は、約１０～２５％、２５～５０％、５０～７５％、７５～１００％、又は１００％超、たとえば、約２倍、３倍、４倍、５倍、又はそれ以上に他方の特性パターンの平均パルス持続時間と異なる。いくつかの実施形態では、互いに異なる特性パターン間の平均パルス持続時間の差がより小さいとき、統計学的信頼性をもって一方を他方と識別するために、各特性パターン内でより多くの数のパルス持続時間が必要とされうる。

【０２１４】

いくつかの実施形態では、特性パターンは、ポリペプチドのアミノ酸とアミノ酸に結合するための手段（たとえば、アミノ酸認識分子）との間の複数の会合イベントを一般的に意味する。いくつかの実施形態では、特性パターンは、少なくとも１０の会合イベント（たとえば、少なくとも２５、少なくとも５０、少なくとも７５、少なくとも１００、少なくとも２５０、少なくとも５００、少なくとも１，０００、又はそれ以上の会合イベント）を含む。いくつかの実施形態では、特性パターンは、約１０～約１００の会合イベント（たとえば、ｙ約１０～約５００の会合イベント、約１０～約２５０の会合イベント、約１０～約１００の会合イベント、又は約５０～約５００の会合イベント）を含む。いくつ

かの実施形態では、複数の会合イベントは、複数のシグナルパルスとして検出される。

【０２１５】

いくつかの実施形態では、特性パターンは、本明細書に記載されるように要的統計によって特徴付けられうる複数のシグナルパルスを意味する。いくつかの実施形態では、特性パターンは、少なくとも１０のシグナルパルス（たとえば、少なくとも２５、少なくとも５０、少なくとも７５、少なくとも１００、少なくとも２５０、少なくとも５００、少なくとも１，０００、又はそれ以上のシグナルパルス）を含む。いくつかの実施形態では、特性パターンは、約１０～約１００のシグナルパルス（たとえば、約１０～約５００のシグナルパルス、約１０～約２５０のシグナルパルス、約１０～約１００のシグナルパルス、又は約５０～約５００のシグナルパルス）を含む。

【０２１６】

いくつかの実施形態では、特性パターンは、アミノ酸認識分子とポリペプチドのアミノ酸との間で、アミノ酸の除去の前の時間インターバルに発生する複数の会合イベントを意味する。いくつかの実施形態では、特性パターンは、２つの切断イベント（たとえば、アミノ酸の除去の前と以前に末端に露出していたアミノ酸の除去の後と）の間の時間インターバルに発生する複数の会合イベントを意味する。いくつかの実施形態では、特性パターンの時間インターバルは、約１分～約３０分（たとえば、約１分～約２０分、約１分～約１０分、約５分～約２０分、約５分～約１５分、又は約５分～約１０分）である。

【０２１７】

いくつかの実施形態では、ポリペプチドシーケンシング反応条件は、特性パターンを有する所望の信頼水準を提供する十分な会合イベントを可能にする時間インターバルを達成するように構成されうる。これは、たとえば、試薬濃度、１つの試薬と別の試薬とのモル比（たとえば、アミノ酸認識分子と切断試薬とのモル比、１つの認識分子と別の認識分子との比、１つの切断試薬と別の切断試薬との比）、さまざまな試薬のタイプの数（たとえば、互いに異なるタイプの認識分子及び／又は切断試薬の数、切断試薬のタイプの数に対する認識分子のタイプの数）、切断活性（たとえば、ペプチダーゼ活性）、結合性（たとえば、認識分子結合の速度論的及び／又は熱力学的結合パラメータ）、試薬修飾（たとえば、相互作用ダイナミックスを変化させうるポリオール及び他のタンパク質修飾）、反応混合物組成（たとえば、ｐＨ、緩衝剤、塩、２価カチオン、界面活性剤、及び本明細書に記載の他の反応混合物組成などの１つ以上の組成物）、反応の温度、及び当業者に明らかのさまざまな他のパラメータ、並びにこれらの組み合わせを含むさまざまな性質に基づいて反応条件を構成することによって達成可能である。反応条件は、たとえば、シグナルパルス情報（たとえば、パルス持続時間、パルス間持続時間、大きさの変化）、標識化ストラトジー（たとえば、フルオロフォアの数及び／又はタイプ、リンカーがシールディングエレメントを有するか有しないか）、表面修飾（たとえば、ポリペプチド固定を含むサンプルウェル表面の修飾）、サンプル調製（たとえば、ポリペプチド断片サイズ、固定のためのポリペプチド修飾）、及び本明細書に記載のその他の態様を含む、本明細書に記載の１つ以上の態様に基づいて構成されうる。

【０２１８】

いくつかの実施形態では、本願に係るポリペプチドシーケンシング反応は、アミノ酸の認識及び切断が単一反応混合物中で同時に発生しうる条件下で実施される。たとえば、いくつかの実施形態では、ポリペプチドシーケンシング反応は、会合イベント及び切断イベントが発生可能なｐＨを有する反応混合物において実施される。いくつかの実施形態では、ポリペプチドシーケンシング反応は、約６．５～約９．０のｐＨの反応混合物において実施される。いくつかの実施形態では、ポリペプチドシーケンシング反応は、約７．５～約８．０（たとえば、約７．０～約８．０、約７．５～約８．５、約７．５～約８．０、又は８．０～約８．５）のｐＨの反応混合物において実施される。

【０２１９】

いくつかの実施形態では、ポリペプチドシーケンシング反応は、１つ以上の緩衝剤を含む反応混合物中で実施される。いくつかの実施形態では、反応混合物は、少なくとも１０

ｍＭ（たとえば、少なくとも２０ｍＭ且つ２５０ｍＭまで、少なくとも５０ｍＭ、１０～２５０ｍＭ、１０～１００ｍＭ、２０～１００ｍＭ、５０～１００ｍＭ、又は１００～２００ｍＭ）の濃度で緩衝剤を含む。いくつかの実施形態では、反応混合物は、約１０ｍＭ～約５０ｍＭ（たとえば、約１０ｍＭ～約２５ｍＭ、約２５ｍＭ～約５０ｍＭ、又は約２０ｍＭ～約４０ｍＭ）の濃度の緩衝剤を含む。緩衝剤の例としては、限定されるものではないが、ＨＥＰＥＳ（４－（２－ヒドロキシエチル）ー１ーピペラジンエタンスルホン酸）、Ｔｒｉｓ（トリス（ヒドロキシメチル）アミノメタン）、及びＭＯＰＳ（３－（Ｎ－モノホリノ）プロパンスルホン酸）が挙げられる。

【０２２０】

いくつかの実施形態では、ポリペプチドシーケンシング反応は、少なくとも１０ｍＭの濃度の塩を含む反応混合物中で実施される。いくつかの実施形態では、反応混合物は、少なくとも１０ｍＭ（たとえば、少なくとも２０ｍＭ、少なくとも５０ｍＭ、少なくとも１００ｍＭ、又はそれ以上）の塩を含む。いくつかの実施形態では、反応混合物は、約１０ｍＭ～約２５０ｍＭ（たとえば、約２０ｍＭ～約２００ｍＭ、約５０ｍＭ～約１５０ｍＭ、約１０ｍＭ～約５０ｍＭ、又は約１０ｍＭ～約１００ｍＭ）の濃度の塩を含む。塩の例としては、限定されるものではないが、塩化ナトリウム（ＮａＣｌ）、酢酸ナトリウム（ＮａＯＡｃ）、及び酢酸カリウム（ＫＯＡｃ）などのナトリウム塩、カリウム塩、及び酢酸塩が挙げられる。

【０２２１】

反応混合物中に使用するための組成物の追加の例としては、２価カチオン（たとえば、Mg^{2+}、Co^{2+}）及び界面活性剤（ポリソルベート２０）が挙げられる。いくつかの実施形態では、反応混合物は、約０．１ｍＭ～約５０ｍＭ（たとえば、約１０ｍＭ～約５０ｍＭ、約０．１ｍＭ～約１０ｍＭ、又は約１ｍＭ～約２０ｍＭ）の濃度の２価カチオンを含む。いくつかの実施形態では、反応混合物は、少なくとも０．０１％（たとえば、約０．０１％～約０．１０％）の濃度の界面活性剤を含む。いくつかの実施形態では、反応混合物は、酸素排除系（ＰＣＡ／ＰＣＤ系又はピラノース／カタラーゼ／グルコース系）及び／又は１つ以上のトリプレット状態クエンチャー（たとえば、トロロックス、ＣＯＴ、及びＮＢＡ）などの単一分子解析に有用な１つ以上の成分を含む。

【０２２２】

いくつかの実施形態では、ポリペプチドシーケンシング反応は、会合イベント及び切断イベントが発生しうる温度で実施される。いくつかの実施形態では、ポリペプチドシーケンシング反応は、少なくとも１０℃の温度で実施される。いくつかの実施形態では、ポリペプチドシーケンシング反応は、約１０℃～約５０℃（たとえば、１５～４５℃、２５℃前後、３０℃前後、３５℃前後、３７℃前後）の温度で実施される。いくつかの実施形態では、ポリペプチドシーケンシング反応は、室温前後で実施される。

【０２２３】

以前に詳述したように、図５に例示されるリアルタイムシーケンシングプロセスは、末端アミノ酸認識及び末端アミノ酸切断のサイクルを通常含み、認識及び切断の相対的出現は、標識付認識分子５１０と標識付非特異的エキソペプチダーゼ５２０との間の濃度差によって制御可能である。いくつかの実施形態では、濃度差は、個別アミノ酸の認識の間に検出されるシグナルパルスの数が同定に望ましい信頼区間を提供するように最適化されうる。たとえば、初期のシーケンシング反応が、望ましい信頼区間を有する特性パターンの決定を可能にするには少なすぎる切断イベント間のシグナルパルスしか有しないシグナルデータを提供する場合、シーケンシング反応は、認識分子に対して減少した濃度の非特異的エキソペプチダーゼを使用して繰り返されうる。

【０２２４】

いくつかの実施形態では、本願に係るポリペプチドシーケンシングは、ペプチドと、１つ以上のアミノ酸認識分子及び／又は１つ以上の切断試薬（たとえば、ペプチダーゼ）を含むシーケンシング反応混合物とを接触させることによって、行うことが可能である。いくつかの実施形態では、シーケンシング反応混合物は、約１０ｎＭ～約１０μＭの濃度で

アミノ酸認識分子を含む。いくつかの実施形態では、シーケンシング反応混合物は、約５００ｎＭ～約５００μＭの濃度で切断試薬を含む。

【０２２５】

いくつかの実施形態では、シーケンシング反応混合物は、約１００ｎＭ～約１０μＭ、約２５０ｎＭ～約１０μＭ、約１００ｎＭ～約１μＭ、約２５０ｎＭ～約１μＭ、約２５０ｎＭ～約７５０ｎＭ、又は約５００ｎＭ～約１μＭの濃度でアミノ酸認識分子を含む。いくつかの実施形態では、シーケンシング反応混合物は、約１００ｎＭ、約２５０ｎＭ、約５００ｎＭ、約７５０ｎＭ、又は約１μＭの濃度でアミノ酸認識分子を含む。

【０２２６】

いくつかの実施形態では、シーケンシング反応混合物は、約５００ｎＭ～約２５０ｎＭ、約５００ｎＭ～約１００ｎＭ、約１μＭ～約１００μＭ、約５００ｎＭ～約５０μＭ、約１μＭ～約１００μＭ、約１０μＭ～約２００μＭ、又は約１０μＭ～約１００μＭの濃度で切断試薬を含む。いくつかの実施形態では、シーケンシング反応混合物は、約１μＭ、約５μＭ、約１０μＭ、約３０μＭ、約５０μＭ、約７０μＭ、又は約１００μＭの濃度で切断試薬を含む。

【０２２７】

いくつかの実施形態では、シーケンシング反応混合物は、約１０ｎＭ～約１０μＭの濃度のアミノ酸認識分子、及び約５００ｎＭ～約５００μＭの濃度の切断試薬を含む。いくつかの実施形態では、シーケンシング反応混合物は、約１００ｎＭ～約１μＭの濃度のアミノ酸認識分子、及び約１μＭ～約１００μＭの濃度の切断試薬を含む。いくつかの実施形態では、シーケンシング反応混合物は、約２５０ｎＭ～約１μＭの濃度のアミノ酸認識分子、及び約１０μＭ～約１００μＭの濃度の切断試薬を含む。いくつかの実施形態では、シーケンシング反応混合物は、約５００ｎＭの濃度のアミノ酸認識分子、及び約２５μＭ～約７５μＭの濃度の切断試薬を含む。いくつかの実施形態では、反応混合物中のアミノ酸認識分子の濃度及び／又は切断試薬の濃度は、本明細書の他の箇所に記載されるようである。

【０２２８】

いくつかの実施形態では、シーケンシング反応混合物は、約５００：１、約４００：１、約３００：１、約２００：１、約１００：１、約７５：１、約５０：１、約２５：１、約１０：１、約５：１、約２：１、又は約１：１のモル比で、アミノ酸認識分子と切断試薬とを含む。いくつかの実施形態では、シーケンシング反応混合物は、約１０：１～約２００：１のモル比で、アミノ酸認識分子と切断試薬とを含む。いくつかの実施形態では、シーケンシング反応混合物は、約５０：１～約１５０：１のモル比で、アミノ酸認識分子と切断試薬とを含む。いくつかの実施形態では、反応混合物中のアミノ酸認識分子と切断試薬とのモル比は、約１：１，０００～約１：１又は約１：１～約１００：１（たとえば、約１：１，０００、約１：５００、約１：２００、約１：１００、約１：１０、約１：５、約１：２、約１：１、約５：１、約１０：１、約５０：１、約１００：１）である。いくつかの実施形態では、反応混合物中のアミノ酸認識分子と切断試薬とのモル比は、約１：１００～約１：１又は約１：１～約１０：１である。いくつかの実施形態では、反応混合物中のアミノ酸認識分子と切断試薬とのモル比は、本明細書の他の箇所に記載されるようである。

【０２２９】

いくつかの実施形態では、シーケンシング反応混合物は、１つ以上のアミノ酸認識分子及び１つ以上の切断試薬を含む。いくつかの実施形態では、シーケンシング反応混合物は、少なくとも３つのアミノ酸認識分子と、少なくとも１つの切断試薬とを含む。いくつかの実施形態では、シーケンシング反応混合物は、２つ以上の切断試薬を含む。いくつかの実施形態では、シーケンシング反応混合物は、少なくとも１つ且つ１０までの切断試薬（たとえば、１つ～３つの切断試薬、２つ～１０の切断試薬、１つ～５つの切断試薬、３つ～１０の切断試薬）を含む。いくつかの実施形態では、シーケンシング反応混合物は、少なくとも３つ且つ３０までのアミノ酸認識分子（たとえば、３～２５、３～２０、３～１

０、３～５、５～３０、５～２０、５～１０、又は１０～２０のアミノ酸認識分子）を含む。いくつかの実施形態では、１つ以上のアミノ酸認識分子は、表１又は表２から選択された少なくとも１つの認識分子を含む。いくつかの実施形態では、１つ以上の切断試薬は、表４から選択された少なくとも１つのペプチダーゼ分子を含む。

【０２３０】

いくつかの実施形態では、シーケンシング反応混合物は、１つより多くのアミノ酸認識分子及び／又は１つより多くの切断試薬を含む。いくつかの実施形態では、１つより多くのアミノ酸認識分子（又は切断試薬）を含むとして記載されるシーケンシング反応混合物は、１タイプより多くのアミノ酸認識分子（又は切断試薬）を有するような混合物を意味する。たとえば、いくつかの実施形態では、シーケンシング反応混合物は、２つ以上のアミノ酸結合タンパク質を含む。いくつかの実施形態では、２つ以上のアミノ酸結合タンパク質は、２タイプ以上のアミノ酸結合タンパク質を意味する。いくつかの実施形態では、１タイプのアミノ酸結合タンパク質は、反応混合物中の別のタイプのアミノ酸結合タンパク質とは異なるアミノ酸配列を有する。いくつかの実施形態では、１タイプのアミノ酸結合タンパク質は、反応混合物中の別のタイプのアミノ酸結合タンパク質の標識とは異なる標識を有する。いくつかの実施形態では、１タイプのアミノ酸結合タンパク質は、反応混合物中の別のタイプのアミノ酸結合タンパク質が会合するアミノ酸とは異なるアミノ酸と会合（たとえば、結合）する。いくつかの実施形態では、１タイプのアミノ酸結合タンパク質は、反応混合物中の別のタイプのアミノ酸結合タンパク質が会合するアミノ酸のサブセットとは異なるアミノ酸のサブセットと会合（たとえば、結合）する。

【０２３１】

図５に例示される例は、標識付切断試薬を使用したシーケンシングプロセスに関するが、シーケンシングプロセスは、この態様に限定されるものではない。本明細書の他の箇所に記載されるように、本発明者らは、非標識切断分子を使用した単一分子シーケンシングを実証した。いくつかの実施形態では、切断試薬が逐次末端アミノ酸を除去するおおよその頻度が知られており、たとえば、使用される酵素の周知の活性及び／又は濃度に基づく。いくつかの実施形態では、試薬による末端アミノ酸切断は、たとえばアミノ酸認識のために検出されるシグナル又は検出されるシグナルがないことに基づいて推測される。

【０２３２】

本発明者らは、さらに、記載されるような濃度差アプローチと組み合わせて又はこれに代わって使用されうる、リアルタイムシーケンシング反応を制御するための技術を認識している。温度依存性リアルタイムシーケンシングプロセスの例は、図６に示される。パネル（Ｉ）～（ＩＩＩ）は、温度依存性末端アミノ酸認識及びアミノ酸切断のサイクルを含むシーケンシング反応を例示する。シーケンシング反応の各サイクルを、２つの温度範囲で行った。第１の温度範囲（「T_1」）は、エキソペプチダーゼ活性よりも認識分子活性に最適であり（たとえば、末端アミノ酸認識を促進するため）、第２の温度範囲（「T_2」）は、認識分子活性よりもエキソペプチダーゼ活性に最適である（たとえば、末端アミノ酸切断を促進するため）。シーケンシング反応は、反応混合物温度を第１の温度範囲T_1（アミノ酸認識を開始するため）と第２の温度範囲T_2（アミノ酸切断を開始するため）との間で変化させることによって、進行する。それゆえ、温度依存性シーケンシングプロセスの進行は、温度によって制御され、互いに異なる温度範囲の間（たとえば、T_1とT_2との間）の変化は、手動プロセス又は自動プロセスによって行われうる。いくつかの実施形態では、第１の温度範囲T_1の認識分子活性（たとえば、アミノ酸に対する結合親和性（K_D））は、第２の温度範囲T_2と比較した場合、少なくとも１０倍、少なくとも１００倍、少なくとも１，０００倍、少なくとも１０，０００倍、少なくとも１００，０００倍、又はそれ以上に上昇する。いくつかの実施形態では、第２の温度範囲T_2のエキソペプチダーゼ活性（たとえば、切断産物への基質変換の速度）は、第１の温度範囲T_1と比較した場合、少なくとも２倍、１０倍、少なくとも２５倍、少なくとも５０倍、少なくとも１００倍、少なくとも１，０００倍、又はそれ以上に上昇する。

【０２３３】

いくつかの実施形態では、第１の温度範囲T_1は、第２の温度範囲T_2よりも低い。いくつかの実施形態では、第１の温度範囲T_1は、約１５℃～約４０℃（たとえば、約２５℃～約３５℃、約１５℃～約３０℃、約２０℃～約３０℃）である。いくつかの実施形態では、第２の温度範囲T_2は、約４０℃～約１００℃（たとえば、約５０℃～約９０℃、約６０℃～約９０℃、約７０℃～約９０℃）である。いくつかの実施形態では、第１の温度範囲T_1は、約２０℃～約４０℃（たとえば、おおよそ３０℃）であり、第２の温度範囲T_2は、約６０℃～約１００℃（たとえば、おおよそ８０℃）である。

【０２３４】

いくつかの実施形態では、第１の温度範囲T_1は、第２の温度範囲T_2よりも高い。いくつかの実施形態では、第１の温度範囲T_1は、約４０℃～約１００℃（たとえば、約５０℃～約９０℃、約６０℃～約９０℃、約７０℃～約９０℃）である。いくつかの実施形態では、第２の温度範囲T_2は、約１５℃～約４０℃（たとえば、約２５℃～約３５℃、約１５℃～約３０℃、約２０℃～約３０℃）である。いくつかの実施形態では、第１の温度範囲T_1は、約６０℃～約１００℃（たとえば、おおよそ８０℃）であり、第２の温度範囲T_2は、約２０℃～約４０℃（たとえば、おおよそ３０℃）である。

【０２３５】

パネル（Ｉ）は、エキソペプチダーゼ活性よりも認識分子活性に最適な第１の温度範囲T_1以内の温度のシーケンシング反応混合物を描く。例示のために、アミノ酸配列「ＫＦＶＡＧ．．．」のポリペプチドが示される。反応混合物温度が第１の温度範囲T_1以内のとき、混合物内の標識付認識分子は、活性化（たとえば、復元）されて、ポリペプチド末端との会合によってアミノ酸認識を開始する。また、第１の温度範囲T_1では、混合物中の標識付エキソペプチダーゼは、不活性化（たとえば、変性）されて、認識中のアミノ酸切断が阻害される。パネル（Ｉ）では、第１の認識分子は、ポリペプチド末端のリシンと可逆的に会合するように示されると同時に、標識付エキソペプチダーゼ（たとえば、ＰｆｕアミノペプチダーゼＩ（Ｐｆｕ　ＡＰＩ））は変性して示される。いくつかの実施形態では、アミノ酸認識は、アミノ酸の切断開始前の時に、既定の持続時間だけ発生する。いくつかの実施形態では、アミノ酸認識は、アミノ酸の切断開始前の時に、認識に所望の信頼区間に到達するために必要な持続時間だけ発生する。アミノ酸認識の後に、第２の温度範囲T_2以内に混合物の温度を変化させることによって、反応が進む。

【０２３６】

パネル（ＩＩ）は、認識分子活性よりもエキソペプチダーゼ活性に最適な第２の温度範囲T_2以内の温度のシーケンシング反応混合物を描く。この例の例示目的のために、第２の温度範囲T_2は、第１の温度範囲T_1よりも高いが、試薬活性は、任意の所望の温度範囲に最適化されうることが認識されるべきである。それゆえ、パネル（Ｉ）からパネル（ＩＩ）への進行は、適切な熱源を使用して反応混合物温度を上昇させることによって行われる。反応混合物が第２の温度範囲T_2以内の温度に到達したとき、混合物内の標識付エキソペプチダーゼは、活性化（たとえば、復元）されて、エキソペプチダーゼ活性による末端アミノ酸切断を開始する。また、第２の温度範囲T_2では、混合物中の標識付認識分子は、不活性化（たとえば、変性）されて、切断中のアミノ酸認識が阻害される。パネル（ＩＩ）では、標識付エキソペプチダーゼは、末端リシン残基を切断しているように示されると同時に、標識付認識分子は変性されている。いくつかの実施形態では、アミノ酸切断は、ポリペプチド末端の逐次アミノ酸の認識開始前の時に、既定の持続時間だけ発生する。いくつかの実施形態では、アミノ酸切断は、逐次アミノ酸の認識開始前の時に、切断を検出するために必要な持続時間だけ発生する。アミノ酸切断の後に、第１の温度範囲T_1以内に混合物の温度を変化させることによって、反応が進む。

【０２３７】

パネル（ＩＩＩ）は、シーケンシング反応における次のサイクルの開始を描き、反応混合物温度は、第１の温度範囲T_1以内まで低下されている。それゆえ、この例では、パネル（ＩＩ）からパネル（ＩＩＩ）への進行は、熱源から反応混合物を除去するか、又は別の方法で反応混合物を第１の温度範囲T_1以内に冷却（たとえば、能動的又は受動的に）

することによって行われうる。示されるように、ポリペプチド末端のフェニルアラニンと可逆的に会合する第２の認識分子を含む標識付認識分子は、復元されているが、標識付エキソペプチダーゼは変性して示される。この例に例示されるように温度依存性の方法でアミノ酸認識及びアミノ酸切断をさらに循環させることによって、シーケンシング反応が継続される。

【０２３８】

それゆえ、動的シーケンシングアプローチは、反応混合物中の１つ以上のタンパク質のタンパク質活性又は機能のレベルで制御される反応サイクルを含みうる。図６に描かれるとともに上述される温度依存性ポリペプチドシーケンシングプロセスは、条件依存的な認識及び切断の制御可能なサイクルによるポリペプチドシーケンシングの全般的なアプローチの例示でありうる。たとえば、いくつかの実施形態では、アプローチは、発光活性化試薬を使用した発光依存的なシーケンシングプロセスを提供する。いくつかの実施形態では、発光依存的なシーケンシングプロセスは、発光依存的なアミノ酸認識及び切断のサイクルを含む。シーケンシング反応の各サイクルは、シーケンシング反応混合物を２つの異なる発光条件にさらすことによって行われうる。第１の発光条件は、エキソペプチダーゼ活性よりも認識分子活性に最適であり（たとえば、アミノ酸認識を促進するため）、第２の発光条件は、認識分子活性よりもエキソペプチダーゼ活性に最適である（たとえば、アミノ酸切断を促進するため）。シーケンシング反応は、第１の発光条件に反応混合物をさらすこと（アミノ酸認識を開始するため）と、第２の発光条件に反応混合物をさらすこと（アミノ酸切断を開始するため）との間で変化させることによって進行する。限定されるものではないが例として、いくつかの実施形態では、２つの異なる発光条件は、第１の波長及び第２の波長を含む。

【０２３９】

いくつかの態様では、本願は、１つ以上の標識付認識分子と末端アミノ酸及び内部アミノ酸との結合相互作用、並びに標識付非特異的エキソペプチダーゼと末端アミノ酸との結合相互作用を評価することによる、リアルタイムのポリペプチドシーケンシングの方法を提供する。図７は、図５及び６のアプローチを説明及び例示する方法が、末端位置及び内部位置の両方で１タイプのアミノ酸に選択的に結合し、解離する標識付認識アミノ酸を使用することによって改変された、シーケンシングの方法の例を示す。以前のアプローチで記載されるように、選択的結合は、シグナル出力７００における一連のパルスを生じさせる。しかしながら、このアプローチでは、一連のパルスは、ポリペプチド全体のアミノ酸のタイプの数によって決定される速度で発生する。それゆえ、いくつかの実施形態では、会合イベントに対応するパルスの速度は、ポリペプチド中に現在存在するコグネート（ｃｏｇｎａｔｅ）アミノ酸の数の診断でありうる。

【０２４０】

以前のアプローチでのように、標識付非特異的ペプチダーゼ７２０は、標識付認識分子７１０よりも比較的低い濃度で存在して、たとえば、切断イベント間に最適な時間ウィンドウを与える（図７、インセットパネル）。そのほか、特定の実施形態では、標識付非特異的ペプチダーゼ７２０のユニークに同定可能な発光標識は、切断イベントが発生した時を示す。ポリペプチドが繰返し切断されるにつれて、標識付認識分子７１０による結合に対応するパルスの速度は、末端アミノ酸が標識付非特異的ペプチダーゼ７２０によって切断されるたびに段階的に低下しうる。この概念は、矢印によって示される時間に切断イベントがある、時間の関数としてパルス速度を一般的に描くプロット７０２によって例示される。そのため、いくつかの実施形態では、切断イベント間に検出されるパターン内で生じるパルスパターン及び／又はパルスの速度に基づいて、アミノ酸を同定し、それによってポリペプチドをこのアプローチで配列決定することができる。

【０２４１】

いくつかの実施形態では、末端ポリペプチド配列情報（たとえば、本明細書に記載されるように決定される）は、１つ以上の他の情報源から得られたポリペプチド配列情報と組み合わされてもよい。たとえば、末端ポリペプチド配列情報は、内部ポリペプチド配列情

報と組み合わされる。いくつかの実施形態では、内部ポリペプチド配列情報は、本明細書に記載されるように、内部アミノ酸と会合する１つ以上のアミノ酸認識分子を使用して得ることができる。内部又は他のポリペプチド配列情報は、ポリペプチド分解プロセスの前又はその間に得ることができる。いくつかの実施形態では、これらの方法から得られた配列情報は、たとえば、１つ以上の内部アミノ酸認識分子を用いて得られた配列情報などの他の技術を用いたポリペプチド配列情報と組み合わされてもよい。

【０２４２】

シーケンシング用サンプルの調製

ポリペプチドサンプルは、シーケンシング前に修飾可能である。いくつかの実施形態では、ポリペプチドのＮ末端アミノ酸又はＣ末端アミノ酸を修飾可能である。図８は、タンパク質サンプルから末端修飾ポリペプチドを調製するための末端修飾の非限定的例を例示する。工程（１）では、タンパク質サンプル８００は、断片化されてポリペプチド断片８０２を生成する。ポリペプチドは、対象ポリペプチドを切断（たとえば化学的に）及び／又は消化（たとえば、トリプシンなどのペプチダーゼを用いて酵素的に）することにより断片化可能である。断片化は、標識化の前又は後に実施可能である。いくつかの実施形態では、断片化は、全タンパク質の標識化後に実施される。１つ以上のアミノ酸は、切断の前又は後に標識されて標識付ポリペプチドを生成可能である。いくつかの実施形態では、ポリペプチドは、化学的又は酵素的断片化の後にサイズ選択される。いくつかの実施形態では、より小さなポリペプチド（たとえば＜２ｋＤａ）は除去され、より大きなポリペプチドは配列解析のために保持される。サイズ選択は、ゲル濾過、ＳＥＣ、透析、ＰＡＧＥゲル抽出、マイクロ流体張力フローなどの技術又はいずれかの他の好適な技術を用いて達成可能である。工程（２）では、ポリペプチド断片８０２のＮ末端又はＣ末端は、修飾されて末端修飾ポリペプチド８０４を生成する。いくつかの実施形態では、修飾は、固定部分を付加することを含む。いくつかの実施形態では、修飾は、カップリング部分を付加することを含む。

【０２４３】

それゆえ、本明細書に提供されるのは、表面（たとえば、タンパク質分析に使用されるチップ上のサンプルウェルの表面）への固定を可能にする部分でタンパク質及びポリペプチドの末端を修飾する方法である。いくつかの実施形態では、かかる方法は、本願に従って分析される標識付ポリペプチドの末端を修飾することを含む。さらに他の実施形態では、かかる方法は、タンパク質又は本願に従ってポリペプチド基質を分解若しくはトランスロケートする酵素の末端を修飾することを含む。

【０２４４】

いくつかの実施形態では、タンパク質又はポリペプチドのカルボキシ末端は、（ｉ）タンパク質又はポリペプチドの遊離カルボン酸基をブロックすることと、（ｉｉ）タンパク質又はポリペプチドを変性することと（たとえば、熱及び／又は化学的手段により）、（ｉｉｉ）タンパク質又はポリペプチドの遊離チオール基をブロックすることと、（ｉｖ）タンパク質又はポリペプチドを消化して遊離Ｃ末端カルボン酸基を含む少なくとも１つのポリペプチド断片を生成することと、（ｖ）遊離Ｃ末端カルボン酸基に機能部分をコンジュゲートすることと（たとえば化学的に）、を含む方法で修飾される。いくつかの実施形態では、本方法は、（ｉ）の後且つ（ｉｉ）の前にタンパク質又はポリペプチドを含むサンプルを透析することをさらに含む。

【０２４５】

いくつかの実施形態では、タンパク質又はポリペプチドのカルボキシ末端は、（ｉ）タンパク質又はポリペプチドを変性することと（たとえば、熱及び／又は化学的手段により）、（ｉｉ）タンパク質又はポリペプチドの遊離チオール基をブロックすることと、（ｉｉｉ）タンパク質又はポリペプチドを消化して遊離Ｃ末端カルボン酸基を含む少なくとも１つのポリペプチド断片を生成することと、（ｉｖ）遊離Ｃ末端カルボン酸基をブロックしてブロックＣ末端カルボン酸基を含む少なくとも１つのポリペプチド断片を生成することと、（ｖ）ブロックＣ末端カルボン酸基に機能部分をコンジュゲートすることと（たと

えば酵素的に）、を含む方法で修飾される。いくつかの実施形態では、本方法は、（ｉｖ）の後且つ（ｖ）の前にタンパク質又はポリペプチドを含むサンプルを透析することをさらに含む。

【０２４６】

いくつかの実施形態では、遊離カルボン酸基をブロックすることは、非修飾カルボン酸と比べて化学反応性を変化させるこの基の化学修飾を意味する。好適なカルボン酸ブロッキング方法は、当技術分野で公知であり、官能基化されるポリペプチドのカルボキシ末端カルボン酸基とは化学的に異なるように側鎖カルボン酸基を修飾するべきである。いくつかの実施形態では、遊離カルボン酸基をブロックすることは、ポリペプチドの遊離カルボン酸基をエステル化又はアミド化することを含む。いくつかの実施形態では、遊離カルボン酸基をブロックすることは、たとえば、ポリペプチドとメタノール性ＨＣｌとを反応させることにより、ポリペプチドの遊離カルボン酸基をメチルエステル化することを含む。遊離カルボン酸基をブロックするのに有用な試薬及び技術の追加の例としては、限定されるものではないが、４－スルホ－２，３，５，６－テトラフルオロフェノール（ＳＴＰ）及び／又はカルボジイミド、たとえば、Ｎ－（３－ジメチルアミノプロピル）－Ｎ’－エチルカルボジイミド塩酸塩（ＥＤＡＣ）、ウロニウム試薬、ジアゾメタン、フィッシャーエステル化用のアルコール及び酸、ＮＨＳエステルを形成するためのＮ－ヒドロキシルスクシンイミド（ＮＨＳ）の使用（潜在的に後続のエステル形成又はアミン形成への中間体として）、又はカルボニルジイミダゾール（ＣＤＩ）との反応、又は混合アンヒドリドの形成、又は潜在的にエステル若しくはアミドのどちらかの形成を介して、カルボン酸を修飾若しくはブロックするいずれかの他の方法が挙げられる。

【０２４７】

いくつかの実施形態では、遊離チオール基をブロックすることは、非修飾チオールと比べて化学反応性を変化させるこの基の化学修飾を意味する。いくつかの実施形態では、遊離チオール基をブロックすることは、タンパク質又はポリペプチドの遊離チオール基を還元及びアルキル化することを含む。いくつかの実施形態では、還元及びアルキル化は、ポリペプチドと、ジチオトレイトール（ＤＴＴ）並びにヨードアセトアミド及びヨード酢酸の一方又は両方と、を接触させることにより行われる。使用されうる追加及び代替のシステイン還元試薬の例は周知であり、限定されるものではないが、２－メルカプトエタノール、トリス（２－カルボキシエチル）ホスフィン塩酸塩（ＴＣＥＰ）、トリブチルホスフィン、ジチオブチルアミン（ＤＴＢＡ）、又はチオール基を還元可能ないずれかの試薬が挙げられる。使用されうる追加及び代替のシステインブロッキング（たとえば、システインアルキル化）試薬の例は周知であり、限定されるものではないが、アクリルアミド、４－ビニルピリジン、Ｎ－エチルマレイミド（ＮＥＭ）、Ｎ－ε－マレイミドカプロン酸（ＥＭＣＡ）、又はジスルフィド結合形成を防止するようにシステインを修飾するいずれかの試薬が挙げられる。

【０２４８】

いくつかの実施形態では、消化は酵素消化を含む。いくつかの実施形態では、消化は、消化条件下でタンパク質又はポリペプチドとエンドペプチダーゼ（たとえばトリプシン）とを接触させることにより行われる。いくつかの実施形態では、消化は化学消化を含む。化学消化及び酵素消化に好適な試薬の例は、当技術分野で公知であり、限定されるものではないが、トリプシン、ケモトリプシン（ｃｈｅｍｏｔｒｙｐｓｉｎ）、Ｌｙｓ－Ｃ、Ａｒｇ－Ｃ、Ａｓｐ－Ｎ、Ｌｙｓ－Ｎ、ＢＮＰＳ－スカトール、ＣＮＢｒ、カスパーゼ、ギ酸、グルタミルエンドペプチダーゼ、ヒドロキシルアミン、ヨードソ安息香酸、好中球エラスターゼ、ペプシン、プロリン－エンドペプチダーゼ、プロテイナーゼＫ、ブドウ球菌ペプチダーゼＩ、サーモリシン、及びトロンビンが挙げられる。

【０２４９】

いくつかの実施形態では、機能部分はビオチン分子を含む。いくつかの実施形態では、機能部分は、アルキニルなどの反応性化学部分を含む。いくつかの実施形態では、機能部分をコンジュゲートすることは、当技術分野で公知のように、カルボキシペプチダーゼＹ

によるカルボキシ末端カルボキシメチルエステル基のビオチン化を含む。

【０２５０】

いくつかの実施形態では、可溶化性部分がポリペプチドに付加される。図９は、たとえば、ポリペプチドに可溶化性リンカーをコンジュゲートするプロセスを用いて、ポリペプチドの末端アミノ酸に付加される可溶化性部分の非限定的例を例示する。

【０２５１】

いくつかの実施形態では、リンカーコンジュゲーティング部分９１２を含む末端修飾ポリペプチド９１０は、ポリペプチドコンジュゲーティング部分９２２を含む可溶化性リンカー９２０にコンジュゲートされる。いくつかの実施形態では、可溶化性リンカーは、バイオ分子（たとえば、斑点形状として示される）などの可溶化性ポリマーを含む。いくつかの実施形態では、９１２と９２２との間に形成された連結９３２を含む得られるリンカーコンジュゲートポリペプチド９３０は、表面コンジュゲーティング部分９３４をさらに含む。それゆえ、いくつかの実施形態では、本明細書に提供される方法及び組成物は、その溶解性を増加させる部分でポリペプチドの末端を修飾するのに有用である。いくつかの実施形態では、可溶化性部分は、断片化（たとえば、トリプシンを用いるなどの酵素的断片化）から生じる相対的に不溶性の小さなポリペプチドに有用である。たとえば、いくつかの実施形態では、ポリペプチドプール中の短いポリペプチドは、ポリマー（たとえば、短いオリゴ、糖、又は他の荷電ポリマー）をポリペプチドにコンジュゲートすることにより可溶化可能である。

【０２５２】

いくつかの実施形態では、サンプルウェルの１つ以上の表面（たとえば、サンプルウェルの側壁）を修飾可能である。サンプルウェル側壁のパッシベーション及び／又はアンチファウリングの非限定的例は、ボトム表面への単一分子固定を促進するために使用されうる修飾表面を用いてサンプルウェルの模式図例が例示される図１０に示される。いくつかの実施形態では、１０４０はSiO_2である。いくつかの実施形態では、１０４２は、ポリペプチドコンジュゲーティング部分（たとえば、ＴＣＯ、テトラジン、Ｎ３、アルキン、アルデヒド、ＮＣＯ、ＮＨＳ、チオール、アルケン、ＤＢＣＯ、ＢＣＮ、ＴＰＰ、ビオチン、又は他の好適なコンジュゲーティング部分）である。いくつかの実施形態では、１０５０はTiO_2又はAl_2O_3である。いくつかの実施形態では、１０５２は、疎水性$C_{4\sim18}$分子、ポリテトラフルオロエチレン化合物（たとえば$(CF_2)_{4\sim12}$）、ポリオール、たとえば、ポリエチレングリコール（たとえばＰＥＧ$_{3\sim100}$）、ポリプロピレングリコール、ポリオキシエチレングリコール、又はそれらの組合せ若しくは変化形、或いは双性イオンたとえばスルホベタインである。いくつかの実施形態では、１０６０はＳｉである。いくつかの実施形態では、１０７０はＡｌである。いくつかの実施形態では、１０８０はＴｉＮである。

【０２５３】

発光標識

本明細書で用いられる場合、発光標識とは、１つ以上の光子を吸収し且つ１つ以上の持続時間後に１つ以上の光子を後続的に放出しうる分子のことである。いくつかの実施形態では、この用語は、状況に依存して「標識」又は「発光分子」と同義的に用いられる。本明細書に記載されるある特定の実施形態に係る発光標識は、標識付認識分子の発光標識、標識付ペプチダーゼ（たとえば、標識付エキソペプチダーゼ、標識付非特異的エキソペプチダーゼ）の発光標識、標識付ペプチドの発光標識、標識付補因子の発光標識、又は本明細書に記載の他の標識付組成物を意味しうる。いくつかの実施形態では、本願に係る発光標識は、１つ以上の標識付アミノ酸を含む標識付ポリペプチドの標識付アミノ酸を意味する。

【０２５４】

いくつかの実施形態では、発光標識は、第１及び第２のクロモフォアを含みうる。いくつかの実施形態では、第１のクロモフォアの励起状態は、第２のクロモフォアへのエネルギー移動を介して緩和可能である。いくつかの実施形態では、エネルギー移動は、フェル

スター共鳴エネルギー移動（ＦＲＥＴ）である。かかるＦＲＥＴペアは、たとえば、図２の標識付アプタマー２０６に対して本明細書に例示及び記載されるように、混合物中の複数の発光標識間の標識ごとの区別をより容易にする性質を有する発光標識を提供するのに有用でありうる。さらに他の実施形態では、ＦＲＥＴペアは、第１の発光標識の第１のクロモフォアと第２の発光標識の第２のクロモフォアとを含む。ある特定の実施形態では、ＦＲＥＴペアは、第１のスペクトル領域で励起エネルギーを吸収し、且つ第２のスペクトル領域で発光を放出しうる。

【０２５５】

いくつかの実施形態では、発光標識とは、フルオロフォア又は色素を意味する。典型的には、発光標識は、芳香族又はヘテロ芳香族化合物を含むとともに、ピレン、アントラセン、ナフタレン、ナフチルアミン、アクリジン、スチルベン、インドール、ベンゾインドール、オキサゾール、カルバゾール、チアゾール、ベンゾチアゾール、ベンゾオキサゾール、フェナントリジン、フェノキサジン、ポルフィリン、キノリン、エチジウム、ベンズアミド、シアニン、カルボシアニン、サリチレート、アントラニレート、クマリン、フルオロセイン、ローダミン、キサンテン、又は他の類似化合物でありうる。

【０２５６】

いくつかの実施形態では、発光標識は、下記：５／６－カルボキシローダミン６Ｇ、５－カルボキシローダミン６Ｇ、６－カルボキシローダミン６Ｇ、６－ＴＡＭＲＡ、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）４４０ＳＸＰ、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）４７０ＳＸＰ、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）４８８、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）５１２、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）５２０ＳＸＰ、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）５８０、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）６００、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）６３５、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スター（ＳＴＡＲ）６３５Ｐ、アベリア（Ａｂｂｅｒｉｏｒ）（登録商標）スターレッド（ＳＴＡＲ　ＲＥＤ）、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）３５０、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）４０５、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）４３０、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）４８０、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）４８８、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）５１４、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）５３２、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）５４６、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）５５５、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）５６８、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）５９４、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）６１０×、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）６３３、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）６４７、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）６６０、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）６８０、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）７００、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）７５０、アレキサフルオル（Ａｌｅｘａ　Ｆｌｕｏｒ）（登録商標）７９０、ＡＭＣＡ、アトー（ＡＴＴＯ）３９０、アトー（ＡＴＴＯ）４２５、アトー（ＡＴＴＯ）４６５、アトー（ＡＴＴＯ）４８８、アトー（ＡＴＴＯ）４９５、アトー（ＡＴＴＯ）５１４、アトー（ＡＴＴＯ）５２０、アトー（ＡＴＴＯ）５３２、アトー（ＡＴＴＯ）５４２、アトー（ＡＴＴＯ）５５０、アトー（ＡＴＴＯ）５６５、アトー（ＡＴＴＯ）５９０、アトー（ＡＴＴＯ）６１０、アトー（ＡＴＴＯ）６２０、アトー（ＡＴＴＯ）６３３、アトー（ＡＴＴＯ）６４７、アトー（ＡＴＴＯ）６４７Ｎ、アトー（ＡＴＴＯ）６５５、アトー（ＡＴＴＯ）６６５、アトー（ＡＴＴＯ）６８０、アトー（ＡＴＴＯ）７００、アトー（ＡＴＴＯ）７２５、アトー（ＡＴＴＯ）７４０、アトー（ＡＴＴＯ）Ｏｘａ１２、アトー（ＡＴＴＯ）Ｒｈｏ１０１、アトー（ＡＴＴＯ）Ｒｈｏ１

１、アトー（ＡＴＴＯ）Ｒｈｏ１２、アトー（ＡＴＴＯ）Ｒｈｏ１３、アトー（ＡＴＴＯ）Ｒｈｏ１４、アトー（ＡＴＴＯ）Ｒｈｏ３Ｂ、アトー（ＡＴＴＯ）Ｒｈｏ６Ｇ、アトー（ＡＴＴＯ）Ｔｈｉｏ１２、ＢＤホライゾン（Ｈｏｒｉｚｏｎ）（商標）Ｖ４５０、ボディピー（ＢＯＤＩＰＹ）（登録商標）４９３／５０１、ボディピー（ＢＯＤＩＰＹ）（登録商標）５３０／５５０、ボディピー（ＢＯＤＩＰＹ）（登録商標）５５８／５６８、ボディピー（ＢＯＤＩＰＹ）（登録商標）５６４／５７０、ボディピー（ＢＯＤＩＰＹ）（登録商標）５７６／５８９、ボディピー（ＢＯＤＩＰＹ）（登録商標）５８１／５９１、ボディピー（ＢＯＤＩＰＹ）（登録商標）６３０／６５０、ボディピー（ＢＯＤＩＰＹ）（登録商標）６５０／６６５、ボディピー（ＢＯＤＩＰＹ）（登録商標）ＦＬ、ボディピー（ＢＯＤＩＰＹ）（登録商標）ＦＬ－Ｘ、ボディピー（ＢＯＤＩＰＹ）（登録商標）Ｒ６Ｇ、ボディピー（ＢＯＤＩＰＹ）（登録商標）ＴＭＲ、ボディピー（ＢＯＤＩＰＹ）（登録商標）ＴＲ、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）ゴールド５４０、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）グリーン５１０、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）オレンジ５６０、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）レッド５９０、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）レッド６１０、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）レッド６１５、ＣＡＬフルオル（Ｆｌｕｏｒ）（登録商標）レッド６３５、カスケード（Ｃａｓｃａｄｅ）（登録商標）ブルー、ＣＦ（商標）３５０、ＣＦ（商標）４０５Ｍ、ＣＦ（商標）４０５Ｓ、ＣＦ（商標）４８８Ａ、ＣＦ（商標）５１４、ＣＦ（商標）５３２、ＣＦ（商標）５４３、ＣＦ（商標）５４６、ＣＦ（商標）５５５、ＣＦ（商標）５６８、ＣＦ（商標）５９４、ＣＦ（商標）６２０Ｒ、ＣＦ（商標）６３３、ＣＦ（商標）６３３－Ｖ１、ＣＦ（商標）６４０Ｒ、ＣＦ（商標）６４０Ｒ－Ｖ１、ＣＦ（商標）６４０Ｒ－Ｖ２、ＣＦ（商標）６６０Ｃ、ＣＦ（商標）６６０Ｒ、ＣＦ（商標）６８０、ＣＦ（商標）６８０Ｒ、ＣＦ（商標）６８０Ｒ－Ｖ１、ＣＦ（商標）７５０、ＣＦ（商標）７７０、ＣＦ（商標）７９０、クロメオ（Ｃｈｒｏｍｅｏ）（商標）６４２、クロミス（Ｃｈｒｏｍｉｓ）４２５Ｎ、クロミス（Ｃｈｒｏｍｉｓ）５００Ｎ、クロミス（Ｃｈｒｏｍｉｓ）５１５Ｎ、クロミス（Ｃｈｒｏｍｉｓ）５３０Ｎ、クロミス（Ｃｈｒｏｍｉｓ）５５０Ａ、クロミス（Ｃｈｒｏｍｉｓ）５５０Ｃ、クロミス（Ｃｈｒｏｍｉｓ）５５０Ｚ、クロミス（Ｃｈｒｏｍｉｓ）５６０Ｎ、クロミス（Ｃｈｒｏｍｉｓ）５７０Ｎ、クロミス（Ｃｈｒｏｍｉｓ）５７７Ｎ、クロミス（Ｃｈｒｏｍｉｓ）６００Ｎ、クロミス（Ｃｈｒｏｍｉｓ）６３０Ｎ、クロミス（Ｃｈｒｏｍｉｓ）６４５Ａ、クロミス（Ｃｈｒｏｍｉｓ）６４５Ｃ、クロミス（Ｃｈｒｏｍｉｓ）６４５Ｚ、クロミス（Ｃｈｒｏｍｉｓ）６７８Ａ、クロミス（Ｃｈｒｏｍｉｓ）６７８Ｃ、クロミス（Ｃｈｒｏｍｉｓ）６７８Ｚ、クロミス（Ｃｈｒｏｍｉｓ）７７０Ａ、クロミス（Ｃｈｒｏｍｉｓ）７７０Ｃ、クロミス（Ｃｈｒｏｍｉｓ）８００Ａ、クロミス（Ｃｈｒｏｍｉｓ）８００Ｃ、クロミス（Ｃｈｒｏｍｉｓ）８３０Ａ、クロミス（Ｃｈｒｏｍｉｓ）８３０Ｃ、Ｃｙ（登録商標）３、Ｃｙ（登録商標）３．５、Ｃｙ（登録商標）３Ｂ、Ｃｙ（登録商標）５、Ｃｙ（登録商標）５．５、Ｃｙ（登録商標）７、ディライト（ＤｙＬｉｇｈｔ）（登録商標）３５０、ディライト（ＤｙＬｉｇｈｔ）（登録商標）４０５、ディライト（ＤｙＬｉｇｈｔ）（登録商標）４１５－Ｃｏ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）４２５Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）４８５－ＬＳ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）４８８、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５０４Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５１０－ＬＳ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５１５－ＬＳ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５２１－ＬＳ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５３０－Ｒ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５４３Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５５０、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５５４－Ｒ０、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５５４－Ｒ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５９０－Ｒ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）５９４、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６１０－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６１５－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６３３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６３３－Ｂ１、デ

ィライト（ＤｙＬｉｇｈｔ）（登録商標）６３３－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６５０、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６５５－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６５５－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６５５－Ｂ３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６５５－Ｂ４、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６６２Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６７５－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６７５－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６７５－Ｂ３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６７５－Ｂ４、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６７９－Ｃ５、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６８０、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６８３Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６９０－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６９０－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）６９６Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７００－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７００－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７３０－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７３０－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７３０－Ｂ３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７３０－Ｂ４、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７４７、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７４７－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７４７－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７４７－Ｂ３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７４７－Ｂ４、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７５５、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７６６Ｑ、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７７５－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７７５－Ｂ３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７７５－Ｂ４、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７８０－Ｂ１、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７８０－Ｂ２、ディライト（ＤｙＬｉｇｈｔ）（登録商標）７８０－Ｂ３、ディライト（ＤｙＬｉｇｈｔ）（登録商標）８００、ディライト（ＤｙＬｉｇｈｔ）（登録商標）８３０－Ｂ２、ディオミクス（Ｄｙｏｍｉｃｓ）－３５０、ディオミクス（Ｄｙｏｍｉｃｓ）－３５０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－３６０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－３７０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－３７５ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－３８０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－３９０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－４０５、ディオミクス（Ｄｙｏｍｉｃｓ）－４１５、ディオミクス（Ｄｙｏｍｉｃｓ）－４３０、ディオミクス（Ｄｙｏｍｉｃｓ）－４３１、ディオミクス（Ｄｙｏｍｉｃｓ）－４７８、ディオミクス（Ｄｙｏｍｉｃｓ）－４８０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－４８１ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－４８５ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－４９０、ディオミクス（Ｄｙｏｍｉｃｓ）－４９５、ディオミクス（Ｄｙｏｍｉｃｓ）－５０５、ディオミクス（Ｄｙｏｍｉｃｓ）－５１０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－５１１ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－５２０ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－５２１ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－５３０、ディオミクス（Ｄｙｏｍｉｃｓ）－５４７、ディオミクス（Ｄｙｏｍｉｃｓ）－５４７Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－５４８、ディオミクス（Ｄｙｏｍｉｃｓ）－５４９、ディオミクス（Ｄｙｏｍｉｃｓ）－５４９Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－５５０、ディオミクス（Ｄｙｏｍｉｃｓ）－５５４、ディオミクス（Ｄｙｏｍｉｃｓ）－５５５、ディオミクス（Ｄｙｏｍｉｃｓ）－５５６、ディオミクス（Ｄｙｏｍｉｃｓ）－５６０、ディオミクス（Ｄｙｏｍｉｃｓ）－５９０、ディオミクス（Ｄｙｏｍｉｃｓ）－５９１、ディオミクス（Ｄｙｏｍｉｃｓ）－５９４、ディオミクス（Ｄｙ
ｏｍｉｃｓ）－６０１ＸＬ、ディオミクス（Ｄｙｏｍｉｃｓ）－６０５、ディオミクス（Ｄｙｏｍｉｃｓ）－６１０、ディオミクス（Ｄｙｏｍｉｃｓ）－６１５、ディオミクス（Ｄｙｏｍｉｃｓ）－６３０、ディオミクス（Ｄｙｏｍｉｃｓ）－６３１、ディオミクス（Ｄｙｏｍｉｃｓ）－６３２、ディオミクス（Ｄｙｏｍｉｃｓ）－６３３、ディオミクス（Ｄｙｏｍｉｃｓ）－６３４、ディオミクス（Ｄｙｏｍｉｃｓ）－６３５、ディオミクス（

Ｄｙｏｍｉｃｓ）－６３６、ディオミクス（Ｄｙｏｍｉｃｓ）－６４７、ディオミクス（Ｄｙｏｍｉｃｓ）－６４７Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－６４８、ディオミクス（Ｄｙｏｍｉｃｓ）－６４８Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－６４９、ディオミクス（Ｄｙｏｍｉｃｓ）－６４９Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－６５０、ディオミクス（Ｄｙｏｍｉｃｓ）－６５１、ディオミクス（Ｄｙｏｍｉｃｓ）－６５２、ディオミクス（Ｄｙｏｍｉｃｓ）－６５４、ディオミクス（Ｄｙｏｍｉｃｓ）－６７５、ディオミクス（Ｄｙｏｍｉｃｓ）－６７６、ディオミクス（Ｄｙｏｍｉｃｓ）－６７７、ディオミクス（Ｄｙｏｍｉｃｓ）－６７８、ディオミクス（Ｄｙｏｍｉｃｓ）－６７９Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－６８０、ディオミクス（Ｄｙｏｍｉｃｓ）－６８１、ディオミクス（Ｄｙｏｍｉｃｓ）－６８２、ディオミクス（Ｄｙｏｍｉｃｓ）－７００、ディオミクス（Ｄｙｏｍｉｃｓ）－７０１、ディオミクス（Ｄｙｏｍｉｃｓ）－７０３、ディオミクス（Ｄｙｏｍｉｃｓ）－７０４、ディオミクス（Ｄｙｏｍｉｃｓ）－７３０、ディオミクス（Ｄｙｏｍｉｃｓ）－７３１、ディオミクス（Ｄｙｏｍｉｃｓ）－７３２、ディオミクス（Ｄｙｏｍｉｃｓ）－７３４、ディオミクス（Ｄｙｏｍｉｃｓ）－７４９、ディオミクス（Ｄｙｏｍｉｃｓ）－７４９Ｐ１、ディオミクス（Ｄｙｏｍｉｃｓ）－７５０、ディオミクス（Ｄｙｏｍｉｃｓ）－７５１、ディオミクス（Ｄｙｏｍｉｃｓ）－７５２、ディオミクス（Ｄｙｏｍｉｃｓ）－７５４、ディオミクス（Ｄｙｏｍｉｃｓ）－７７６、ディオミクス（Ｄｙｏｍｉｃｓ）－７７７、ディオミクス（Ｄｙｏｍｉｃｓ）－７７８、ディオミクス（Ｄｙｏｍｉｃｓ）－７８０、ディオミクス（Ｄｙｏｍｉｃｓ）－７８１、ディオミクス（Ｄｙｏｍｉｃｓ）－７８２、ディオミクス（Ｄｙｏｍｉｃｓ）－８００、ディオミクス（Ｄｙｏｍｉｃｓ）－８３１、ｅフルオル（ｅＦｌｕｏｒ）（登録商標）４５０、エオシン、ＦＩＴＣ、フルオレセイン、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）４０５、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）４８８、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）５３２、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）５５５、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）５９４、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）６４７、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）６８０、ハイライト（ＨｉＬｙｔｅ）（商標）フルオル（Ｆｌｕｏｒ）７５０、ＩＲダイ（ＩＲＤｙｅ）（登録商標）６８０ＬＴ、ＩＲダイ（ＩＲＤｙｅ）（登録商標）７５０、ＩＲダイ（ＩＲＤｙｅ）（登録商標）８００ＣＷ、ＪＯＥ、ライトサイクラー（ＬｉｇｈｔＣｙｃｌｅｒ）（登録商標）６４０Ｒ、ライトサイクラー（ＬｉｇｈｔＣｙｃｌｅｒ）（登録商標）レッド６１０、ライトサイクラー（ＬｉｇｈｔＣｙｃｌｅｒ）（登録商標）レッド６４０、ライトサイクラー（ＬｉｇｈｔＣｙｃｌｅｒ）（登録商標）レッド６７０、ライトサイクラー（ＬｉｇｈｔＣｙｃｌｅｒ）（登録商標）レッド７０５、リサミンローダミンＢ、ナフトフルオレセイン、オレゴングリーン（Ｏｒｅｇｏｎ　Ｇｒｅｅｎ）（登録商標）４８８、オレゴングリーン（Ｏｒｅｇｏｎ　Ｇｒｅｅｎ）（登録商標）５１４、パシフィックブルー（Ｐａｃｉｆｉｃ　Ｂｌｕｅ）（商標）、パシフィックグリーン（Ｐａｃｉｆｉｃ　Ｇｒｅｅｎ）（商標）、パシフィックオレンジ（Ｐａｃｉｆｉｃ　Ｏｒａｎｇｅ）（商標）、ＰＥＴ、ＰＦ３５０、ＰＦ４０５、ＰＦ４１５、ＰＦ４８８、ＰＦ５０５、ＰＦ５３２、ＰＦ５４６、ＰＦ５５５Ｐ、ＰＦ５６８、ＰＦ５９４、ＰＦ６１０、ＰＦ６３３Ｐ、ＰＦ６４７Ｐ、クエーサー（Ｑｕａｓａｒ）（登録商標）５７０、クエーサー（Ｑｕａｓａｒ）（登録商標）６７０、クエーサー（Ｑｕａｓａｒ）（登録商標）７０５、ローダミン１２３、ローダミン６Ｇ、ローダミンＢ、ローダミングリーン、ローダミングリーン－Ｘ、ローダミンレッド、ＲＯＸ、セタ（Ｓｅｔａ）（商標）３７５、セタ（Ｓｅｔａ）（商標）４７０、セタ（Ｓｅｔａ）（商標）５５５、セタ（Ｓｅｔａ）（商標）６３２、セタ（Ｓｅｔａ）（商標）６３３、セタ（Ｓｅｔａ）（商標）６５０、セタ（Ｓｅｔａ）（商標）６６０、セタ（Ｓｅｔａ）（商標）６７０、セタ（Ｓｅｔａ）（商標）６８０、セタ（Ｓｅｔａ）（商標）７００、セタ（Ｓｅｔａ）（商標）７５０、セタ（Ｓｅｔａ）（商標）７８０、セタ（Ｓｅｔａ）（商標）ＡＰＣ－７８０、セタ（Ｓｅｔａ）（商標）ＰｅｒＣＰ－６８０、セタ（Ｓｅｔａ）（商標）Ｒ－ＰＥ－６７

０、セタ（Ｓｅｔａ）（商標）６４６、セタウ（ＳｅＴａｕ）３８０、セタウ（ＳｅＴａｕ）４２５、セタウ（ＳｅＴａｕ）６４７、セタウ（ＳｅＴａｕ）４０５、スクエア（Ｓｑｕａｒｅ）６３５、スクエア（Ｓｑｕａｒｅ）６５０、スクエア（Ｓｑｕａｒｅ）６６０、スクエア（Ｓｑｕａｒｅ）６７２、スクエア（Ｓｑｕａｒｅ）６８０、スルホローダミン１０１、ＴＡＭＲＡ、ＴＥＴ、テキサスレッド（Ｔｅｘａｓ　Ｒｅｄ）（登録商標）、ＴＭＲ、ＴＲＩＴＣ、ヤキマイエロー（Ｙａｋｉｍａ　Ｙｅｌｌｏｗ）（商標）、ゼノン（Ｚｅｎｏｎ）（登録商標）、Ｚｙ３、Ｚｙ５、Ｚｙ５．５、及びＺｙ７の１つ以上から選択される色素を含む。

【０２５７】

発光

いくつかの態様では、本願は、発光標識の１つ以上の発光性に基づくポリペプチドシーケンシング及び／又は同定に関する。いくつかの実施形態では、発光標識は、発光寿命、発光強度、輝度、吸収スペクトル、発光スペクトル、発光量子収率、又はそれらの２つ以上の組合せに基づいて同定される。いくつかの実施形態では、複数のタイプの発光標識は、異なる発光寿命、発光強度、輝度、吸収スペクトル、発光スペクトル、発光量子収率、又はそれらの２つ以上の組合せに基づいて互いに識別可能である。同定することは、発光標識に会合する１タイプのアミノ酸（たとえば、単一タイプ又はタイプのサブセット）の厳密なアイデンティティー及び／又は量を割り当てることを意味しうるとともに、他のタイプのアミノ酸と比べてポリペプチド中のアミノ酸位置を割り当てることも意味しうる。

【０２５８】

いくつかの実施形態では、発光は、発光標識を一連の個別光パルスに暴露することと、標識から放出される各光子のタイミング又は他の性質を評価することと、により検出される。いくつかの実施形態では、標識から逐次放出される複数の光子の情報は、アグリゲートされ、且つ標識を同定することにより関連アミノ酸のタイプを同定するように評価される。いくつかの実施形態では、標識の発光寿命は、標識から逐次放出される複数の光子から決定され、且つ発光寿命は、標識を同定するために使用可能である。いくつかの実施形態では、標識の発光強度は、標識から逐次放出される複数の光子から決定され、且つ発光強度は、標識を同定するために使用可能である。いくつかの実施形態では、標識の発光寿命及び発光強度は、標識から逐次放出される複数の光子から決定され、且つ発光寿命及び発光強度は、標識を同定するために使用可能である。

【０２５９】

本願のいくつかの態様では、単一ポリペプチド分子は、複数の個別光パルスに露出され、且つ一連の放出光子は、検出及び解析される。いくつかの実施形態では、一連の放出光子は、実験の時間にわたり反応サンプル中で変化しない存在する単一ポリペプチド分子に関する情報を提供する。しかしながら、いくつかの実施形態では、一連の放出光子は、反応サンプル中に異なる時間で存在する一連の異なる分子に関する情報を提供する（たとえば、反応又はプロセスが進行する際）。限定されるものではないが例として、かかる情報は、本願に従って化学分解又は酵素分解に付されたポリペプチドをシーケンス及び／又は同定するために使用されうる。

【０２６０】

ある特定の実施形態では、発光標識は、一光子を吸収し且つある持続時間後に一光子を放出する。いくつかの実施形態では、標識の発光寿命は、持続時間を測定することにより決定又は推定可能である。いくつかの実施形態では、標識の発光寿命は、複数のパルスイベント及び放出イベントに対して複数の持続時間を測定することにより決定又は推定可能である。いくつかの実施形態では、標識の発光寿命は、持続時間を測定することにより複数のタイプの標識の発光寿命間で区別可能である。いくつかの実施形態では、標識の発光寿命は、複数のパルスイベント及び放出イベントに対して複数の持続時間を測定することにより複数のタイプの標識の発光寿命間で区別可能である。ある特定の実施形態では、標識は、標識の発光寿命を決定又は推定することにより複数のタイプの標識間で同定又は区別される。ある特定の実施形態では、標識は、複数のタイプの標識の複数の発光寿命間で

標識の発光寿命を区別することにより複数のタイプの標識間で同定又は区別される。

【０２６１】

発光標識の発光寿命の決定は、いずれかの好適な方法を用いて実施可能である（たとえば、好適な技術を用いて寿命を測定することにより又は放出の時間依存特性を決定することにより）。いくつかの実施形態では、１つの標識の発光寿命を決定することは、他の標識と比べて寿命を決定することを含む。いくつかの実施形態では、標識の発光寿命を決定することは、参照と比べて寿命を決定することを含む。いくつかの実施形態では、標識の発光寿命を決定することは、寿命（たとえば蛍光寿命）を測定することを含む。いくつかの実施形態では、標識の発光寿命を決定することは、寿命の指標となる１つ以上の時間特性を決定することを含む。いくつかの実施形態では、標識の発光寿命は、励起パルスに対する１つ以上の時間ゲートウィンドウ全体にわたり発生する複数の放出イベント（たとえば、１、２、３、４、５、６、７、８、９、１０、１１、１２、１３、１４、１５、１６、１７、１８、１９、２０、３０、４０、５０、６０、７０、８０、９０、１００、又はそれ以上の放出イベント）の分布に基づいて決定可能である。たとえば、標識の発光寿命は、励起パルスに対して測定された光子到着時間の分布に基づいて異なる発光寿命を有する複数の標識から識別可能である。

【０２６２】

発光標識の発光寿命は、標識が励起状態に達した後に放出される光子のタイミングの指標となり、標識は、光子のタイミングの指標となる情報により区別可能であることが認識されるべきである。いくつかの実施形態は、標識により放出される光子に関連付けられる時間を測定することにより標識の発光寿命に基づいて標識を複数の標識から区別することを含みうる。時間の分布は、分布から決定されうる発光寿命の指標を提供しうる。いくつかの実施形態では、標識は、たとえば、時間の分布と既知の標識に対応する参照分布とを比較することにより、時間の分布に基づいて複数の標識から区別可能である。いくつかの実施形態では、発光寿命の値は、時間の分布から決定された。

【０２６３】

本明細書で用いられる場合、いくつかの実施形態では、発光強度とは、パルス励起エネルギーの送達により励起されている発光標識により放出される単位時間当たりの放出光子数を意味する。いくつかの実施形態では、発光強度とは、パルス励起エネルギーの送達により励起されている標識により放出され且つ特定センサー又はセンサーのセットにより検出される単位時間当たりの検出放出光子数を意味する。

【０２６４】

本明細書で用いられる場合、いくつかの実施形態では、輝度とは、発光標識当たりの平均発光強度について報告するパラメータを意味する。そのため、いくつかの実施形態では、「発光強度」は、１つ以上の標識を含む組成物の輝度を一般に意味するように使用されうる。いくつかの実施形態では、標識の輝度は、その量子収率と吸光係数との積に等しい。

【０２６５】

本明細書で用いられる場合、いくつかの実施形態では、発光量子収率とは、放出イベントをもたらす所与の波長又は所与のスペクトル領域内での励起イベントの割合を意味し、典型的には１未満である。いくつかの実施形態では、本明細書に記載の発光標識の発光量子収率は、０～約０．００１、約０．００１～約０．０１、約０．０１～約０．１、約０．１～約０．５、約０．５～０．９、又は約０．９～１である。いくつかの実施形態では、標識は、発光量子収率を同定又は推定することにより同定される。

【０２６６】

本明細書で用いられる場合、いくつかの実施形態では、励起エネルギーは光源からの光パルスである。いくつかの実施形態では、励起エネルギーは可視スペクトル域にある。いくつかの実施形態では、励起エネルギーは紫外スペクトル域にある。いくつかの実施形態では、励起エネルギーは赤外スペクトル域にある。いくつかの実施形態では、励起エネルギーは、複数の放出光子が検出される発光標識の吸収極大又はその近くにある。ある特定

の実施形態では、励起エネルギーは、約５００ｎｍ～約７００ｎｍ（たとえば、約５００ｎｍ～約６００ｎｍ、約６００ｎｍ～約７００ｎｍ、約５００ｎｍ～約５５０ｎｍ、約５５０ｎｍ～約６００ｎｍ、約６００ｎｍ～約６５０ｎｍ、又は約６５０ｎｍ～約７００ｎｍ）である。ある特定の実施形態では、励起エネルギーは、単色でありうるか又はあるスペクトル領域に閉じ込められうる。いくつかの実施形態では、スペクトル領域は、約０．１ｎｍ～約１ｎｍ、約１ｎｍ～約２ｎｍ、又は約２ｎｍ～約５ｎｍの領域を有する。いくつかの実施形態では、スペクトル領域は、約５ｎｍ～約１０ｎｍ、約１０ｎｍ～約５０ｎｍ、又は約５０ｎｍ～約１００ｎｍの領域を有する。

【０２６７】

シーケンシング

本願の態様は、ポリペプチドやタンパク質などの生体ポリマーのシーケンシングに関する。本明細書で用いられる場合、「シーケンシング」、「配列決定」、「配列を決定する」などの用語は、ポリペプチド又はタンパク質に関して、ポリペプチド又はタンパク質の部分配列情報さらには全配列情報の決定を含む。すなわち、この用語は、配列比較、フィンガープリンティング、確率的フィンガープリンティング、及び標的分子に関する同レベルの情報、さらには対象領域内の標的分子の各アミノ酸の明確な同定及び順序付けを含む。いくつかの実施形態では、この用語は、ポリペプチドの単一アミノ酸を同定することを含む。さらに他の実施形態では、ポリペプチドの１アミノ酸超が同定される。本明細書で用いられる場合、いくつかの実施形態では、「同定すること」、「アイデンティティーを決定すること」などの用語は、アミノ酸に関して、アミノ酸の明確なアイデンティティーの決定、さらにはアミノ酸の明確なアイデンティティーの確率の決定を含む。たとえば、いくつかの実施形態では、アミノ酸は、アミノ酸が特定のタイプである確率（たとえば０％～１００％）を決定することにより、又は複数の特定のタイプの各々の確率を決定することにより、同定される。それゆえ、いくつかの実施形態では、本明細書で用いられる「アミノ酸配列」、「ポリペプチド配列」、及び「タンパク質配列」という用語は、ポリペプチド又はタンパク質の物質それ自体を意味しうるとともに、特定のポリペプチド又はタンパク質を生化学的に特徴付ける特定の配列情報（たとえば、一方の末端から他方の末端へのアミノ酸の順序を表す一連の文字）に制限されるものではない。

【０２６８】

いくつかの実施形態では、ポリペプチド分子のシーケンシングは、ポリペプチド分子中の少なくとも２つ（たとえば、少なくとも３つ、少なくとも４つ、少なくとも５つ、少なくとも６つ、少なくとも７つ、少なくとも８つ、少なくとも９つ、少なくとも１０、少なくとも１１、少なくとも１２、少なくとも１３、少なくとも１４、少なくとも１５、少なくとも１６、少なくとも１７、少なくとも１８、少なくとも１９、少なくとも２０、少なくとも２５、少なくとも３０、少なくとも３５、少なくとも４０、少なくとも４５、少なくとも５０、少なくとも６０、少なくとも７０、少なくとも８０、少なくとも９０、少なくとも１００、又はそれ以上）のアミノ酸を同定することを含む。いくつかの実施形態では、少なくとも２つのアミノ酸は連続アミノ酸である。いくつかの実施形態では、少なくとも２つのアミノ酸は非連続アミノ酸である。

【０２６９】

いくつかの実施形態では、ポリペプチド分子のシーケンシングは、ポリペプチド分子中のすべてのアミノ酸の１００％未満（たとえば、９９％未満、９５％未満、９０％未満、８５％未満、８０％未満、７５％未満、７０％未満、６５％未満、６０％未満、５５％未満、５０％未満、４５％未満、４０％未満、３５％未満、３０％未満、２５％未満、２０％未満、１５％未満、１０％未満、５％未満、１％未満、又はそれ以下）の同定を含む。たとえば、いくつかの実施形態では、ポリペプチド分子のシーケンシングは、ポリペプチド分子中の１タイプのアミノ酸の１００％未満の同定（たとえば、ポリペプチド分子中の１タイプのすべてのアミノ酸の一部分の同定）を含む。いくつかの実施形態では、ポリペプチド分子のシーケンシングは、ポリペプチド分子中の各タイプのアミノ酸の１００％未満の同定を含む。

【０２７０】
いくつかの実施形態では、ポリペプチド分子のシーケンシングは、ポリペプチド中の少なくとも１つ、少なくとも５つ、少なくとも１０、少なくとも１５、少なくとも２０、少なくとも２５、少なくとも３０、少なくとも３５、少なくとも４０、少なくとも４５、少なくとも５０、少なくとも５５、少なくとも６０、少なくとも６５、少なくとも７０、少なくとも７５、少なくとも８０、少なくとも８５、少なくとも９０、少なくとも９５、少なくとも１００、又はそれ以上のタイプのアミノ酸の同定を含む。
【０２７１】
いくつかの実施形態では、本願は、経時的にポリペプチドの末端に存在する一連のアミノ酸を同定することにより（たとえば、末端のアミノ酸の繰返し検出及び切断により）ポリペプチドをシーケンスするための組成物及び方法を提供する。さらに他の実施形態では、本願は、ポリペプチドの標識付アミノ含有率を同定することと、参照配列データベースと比較することと、によりポリペプチドをシーケンスするための組成物及び方法を提供する。
【０２７２】
いくつかの実施形態では、本願は、ポリペプチドの複数の断片をシーケンスすることによりポリペプチドをシーケンスするための組成物及び方法を提供する。いくつかの実施形態では、ポリペプチドをシーケンスすることは、複数のポリペプチド断片の配列情報を組み合わせてポリペプチドの配列を同定及び／又は決定する工程を含む。いくつかの実施形態では、配列情報を組み合わせることは、コンピュータハードウェア及びソフトウェアにより実施されうる。本明細書に記載の方法は、関連ポリペプチドのセット、たとえば、生物の全プロテオームのシーケンシングを可能にし得る。いくつかの実施形態では、本出願の態様に従って複数の単一分子シーケンシング反応がパラレルに実施される（たとえば、単一チップ上で）。たとえば、いくつかの実施形態では、複数の単一分子シーケンシング反応は、各々、単一チップ上の個別サンプルウェル中で実施される。
【０２７３】
いくつかの実施形態では、本明細書に提供される方法は、タンパク質の複合混合物を含むサンプル中の個別タンパク質のシーケンシング及び同定に使用されうる。いくつかの実施形態では、本願は、タンパク質の複合混合物中の個別タンパク質をユニークに同定する方法を提供する。いくつかの実施形態では、個別タンパク質は、タンパク質の部分アミノ酸配列を決定することにより混合サンプルで検出される。いくつかの実施形態では、タンパク質の部分アミノ酸配列は、おおよそ５～５０アミノ酸の連続ストレッチ内にある。
【０２７４】
いかなる特定理論にも拘束されることを望むものではないが、ほとんどのヒトタンパク質は、プロテオミックデータベースを参照して不完全配列情報を用いて同定可能であると考えられる。たとえば、ヒトプロテオームの単純モデリングは、６～４０アミノ酸のストレッチ内のちょうど４つのタイプのアミノ酸を検出することによりタンパク質のおおよそ９８％をユニークに同定可能であることを示した（たとえば、スワミナサン（Ｓｗａｍｉｎａｔｈａｎ）ら著、プロス・コンピーテーショナル・バイオロジー（ＰＬｏＳ　Ｃｏｍｐｕｔ　Ｂｉｏｌ．）、２０１５年、第１１巻、第２号、ｐ．ｅ１００４０８０、及びヤオ（Ｙａｏ）ら著、フィジカル・バイオロジー（Ｐｈｙｓ．Ｂｉｏｌ．）、２０１５年、第１２巻、第５号、ｐ．０５５００３を参照されたい）。したがって、タンパク質の複合混合物は、おおよそ６～４０アミノ酸の短いポリペプチド断片に分解（たとえば、化学分解、酵素分解）可能であり、且つこのポリペプチドライブラリーのシーケンシングは、元の複合混合物中に存在するタンパク質の各々のアイデンティティー及び存在量を明らかにするであろう。選択的アミノ酸標識化と部分配列情報を決定することによるポリペプチドの同定とを行うための組成物及び方法は、「単一分子ペプチドシーケンシング（ＳＩＮＧＬＥ　ＭＯＬＥＣＵＬＥ　ＰＥＰＴＩＤＥ　ＳＥＱＵＥＮＣＩＮＧ）」という名称の２０１５年９月１５日出願の米国特許出願第１５／５１０，９６２号明細書（その全体が参照により組み込まれる）に詳細に記載されている。

【０２７５】
実施形態は、少なくとも約５０％、６０％、７０％、７５％、８０％、８５％、９０％、９５％、９６％、９７％、９８％、９９％、９９．９％、９９．９９％、９９．９９９％、又は９９．９９９９％の確度などの高確度で単一ポリペプチド分子をシーケンス可能である。いくつかの実施形態では、単一分子シーケンシングに使用される標的分子は、サンプルウェルのボトム表面や側壁表面などの固形担体の表面に固定されるポリペプチドである。サンプルウェルはまた、本願に係るシーケンシング反応に必要とされるいずれかの他の試薬、たとえば、１つ以上の好適な緩衝液、補因子、標識付認識分子、及び酵素（たとえば、発光標識又は非標識でありうる触媒活性又は不活性のエキソペプチダーゼ酵素）を含有可能である。

【０２７６】
以上に記載されるように、いくつかの実施形態では、本願に係るシーケンシングは、アミノ酸が特定のタイプである確率を決定することによりアミノ酸を同定することを含む。従来のタンパク質同定システムは、ポリペプチドを同定するためにポリペプチド中の各アミノ酸の同定を必要とする。しかしながら、ポリペプチド中の各アミノ酸を正確に同定するのは困難である。たとえば、第１の認識分子が第１のアミノ酸に会合する相互作用から収集されるデータは、第２の認識分子が第２のアミノ酸に会合する相互作用から収集されるデータと２つのアミノ酸を区別するのに十分な程度に異ならないおそれがある。いくつかの実施形態では、本願に係るシーケンシングは、従来のタンパク質同定システムとは異なりタンパク質中の各アミノ酸の同定を必要としない（ただし、除外するものではない）タンパク質同定システムを用いることによりこの問題を回避する。

【０２７７】
それゆえ、いくつかの実施形態では、本願に係るシーケンシングは、タンパク質を同定するために機械学習技術を使用するタンパク質同定システムを用いて行われうる。いくつかの実施形態では、システムは、（１）リアルタイムタンパク質シーケンシングデバイスを用いてタンパク質のポリペプチドに関するデータを収集することと、（２）機械学習モデル及び収集データを用いてある特定のアミノ酸がそれぞれの位置でポリペプチドの一部である確率を同定することと、（３）「確率的フィンガープリント」として同定確率を用いてタンパク質を同定することと、により動作する。いくつかの実施形態では、タンパク質のポリペプチドに関するデータは、アミノ酸に選択的に結合する試薬を用いて得られうる。例として、試薬及び／又はアミノ酸は、励起エネルギーの適用に反応して光を放出する発光標識で標識されうる。この例では、タンパク質シーケンシングデバイスは、試薬とサンプル中のアミノ酸との結合相互作用時にタンパク質（たとえばポリペプチド）のサンプルに励起エネルギーを適用しうる。いくつかの実施形態では、配列決定デバイスの１つ以上のセンサー（たとえば、光検出器、電気センサー、及び／又はいずれかの他の好適なタイプのセンサー）は、結合相互作用時に検出しうる。一方、検出光放出から収集及び／又は取得されるデータは、機械学習モデルに提供されうる。機械学習モデル並びに関連システム及び方法は、「機械学習で実行可能なタンパク質同定（ＭＡＣＨＩＮＥ　ＬＥＡＲＮＩＮＧ　ＥＮＡＢＬＥＤ　ＰＲＯＴＥＩＮ　ＩＤＥＮＴＩＦＩＣＡＴＩＯＮ）」という名称の２０１９年６月１２日出願の米国仮特許出願第６２／８６０，７５０号明細書（その全体が参照により組み込まれる）に詳細に記載されている。

【０２７８】
本願に係るシーケンシングは、いくつかの態様では、ポリペプチドを基体（たとえば、固形担体、たとえば、チップ、たとえば、本明細書に記載のインテグレートデバイス）の表面上に固定することを含みうる。いくつかの実施形態では、ポリペプチドは、基体上のサンプルウェルの表面上（たとえば、サンプルウェルのボトム表面上）に固定されうる。いくつかの実施形態では、ポリペプチドのＮ末端アミノ酸が固定される（たとえば、表面に装着される）。いくつかの実施形態では、ポリペプチドのＣ末端アミノ酸が固定される（たとえば、表面に装着される）。いくつかの実施形態では、１つ以上の非末端アミノ酸が固定される（たとえば、表面に装着される）。固定アミノ酸は、たとえば、本願に記載

されるように、いずれかの好適な共有結合性又は非共有結合性連結を用いて装着可能である。いくつかの実施形態では、複数のポリペプチドが複数のサンプルウェルに、たとえば、基体上のサンプルウェルのアレイ中に装着される（たとえば、各サンプルウェルの表面たとえばボトム表面に１つのポリペプチドが装着される）。

【０２７９】

本願に係るシーケンシングは、いくつかの態様では、単一分子分析を可能にするシステムを用いて実施されうる。システムは、インテグレートデバイスと、インテグレートデバイスにインターフェースするように構成された機器と、を含みうる。インテグレートデバイスは、ピクセルのアレイを含みうるとともに、個別ピクセルは、サンプルウェルと少なくとも１つの光検出器とを含む。インテグレートデバイスのサンプルウェルは、インテグレートデバイスの表面上又は表面全体にわたり形成されうるとともに、インテグレートデバイスの表面上に配置されたサンプルを受け取るように構成されうる。全体として、サンプルウェルは、サンプルウェルのアレイとみなされうる。複数のサンプルウェルは、サンプルウェルの少なくとも一部分が単一サンプル（たとえば、ポリペプチドなどの単一分子）を受け取るように好適なサイズ及び形状を有しうる。いくつかの実施形態では、サンプルウェル内のサンプルの数は、いくつかのサンプルウェルが１つのサンプルを含有し、一方、他のものがゼロ、２、又はそれ以上のサンプルを含有するように、インテグレートデバイスのサンプルウェル内に分配されうる。

【０２８０】

励起光は、インテグレートデバイスの外部の１つ以上の光源からインテグレートデバイスに提供される。インテグレートデバイスの光学コンポーネントは、光源から励起光を受け取り、光をインテグレートデバイスのサンプルウェルのアレイにダイレクトし、そしてサンプルウェル内の照射領域を照明しうる。いくつかの実施形態では、サンプルウェルは、サンプルウェルの表面に近接してサンプルを保持できるようにするとともにサンプルへの励起光の送達及びサンプルからの放出光の検出を容易にしうる構成を有しうる。照明領域内に位置決めされるサンプルは、励起光により照明されるとそれに反応して放出光を放出しうる。たとえば、サンプルは、励起光の照明を介して励起状態を達成するとそれに反応して光を放出する蛍光マーカーで標識されうる。次いで、サンプルにより放出される放出光は、分析されるサンプルを有するサンプルウェルに対応するピクセル内の１つ以上の光検出器により検出されうる。いくつかの実施形態によれば数がおおよそ１０，０００ピクセル～１，０００，０００ピクセルの範囲内でありうるサンプルウェルのアレイ全体にわたり実施されるとき、複数のサンプルはパラレルに分析可能である。

【０２８１】

インテグレートデバイスは、励起光を受け取って励起光をサンプルウェルアレイ間にダイレクトするための光学システムを含みうる。光学システムは、インテグレートデバイスに励起光を結合して他の光学コンポーネントに励起光をダイレクトするように構成された１つ以上のグレーティングカプラーを含みうる。光学システムは、グレーティングカプラーからサンプルウェルアレイに向けて励起光をダイレクトする光学コンポーネントを含みうる。かかる光学コンポーネントは、光学スプリッターと光学コンバイナーとウェーブガイドとを含みうる。いくつかの実施形態では、１つ以上の光学スプリッターは、グレーティングカプラーからの励起光を結合して励起光をウェーブガイドの少なくとも１つに送達しうる。いくつかの実施形態によれば、光学スプリッターは、ウェーブガイドの各々が実質的に類似した量の励起光を受け取るようにすべてのウェーブガイドにわたり実質的に均一な励起光の送達を可能にする構成を有しうる。かかる実施形態は、インテグレートデバイスのサンプルウェルが受け取る励起光の均一性を改善することによりインテグレートデバイスの性能を改善しうる。たとえば、サンプルウェルに励起光を結合するために及び／又は光検出器に放出光をダイレクトするために、インテグレートデバイスに含めるのに好適なコンポーネントの例は、「分子をプローブ、検出、及び分析するためのインテグレートデバイス（ＩＮＴＥＧＲＡＴＥＤ　ＤＥＶＩＣＥ　ＦＯＲ　ＰＲＯＢＩＮＧ，ＤＥＴＥＣＴＩＮＧ　ＡＮＤ　ＡＮＡＬＹＺＩＮＧ　ＭＯＬＥＣＵＬＥＳ）」という名称の２０１

５年８月７日出願の米国特許出願第１４／８２１，６８８号明細書、及び「分子をプローブ、検出、及び分析するための外部光源付きインテグレートデバイス（ＩＮＴＥＧＲＡＴＥＤ ＤＥＶＩＣＥ ＷＩＴＨ ＥＸＴＥＲＮＡＬ ＬＩＧＨＴ ＳＯＵＲＣＥ ＦＯＲ ＰＲＯＢＩＮＧ，ＤＥＴＥＣＴＩＮＧ，ＡＮＤ ＡＮＡＬＹＺＩＮＧ ＭＯＬＥＣＵＬＥＳ）」という名称の２０１４年１１月１７日出願の米国特許出願第１４／５４３，８６５号明細書（両方ともその全体が参照により組み込まれる）に記載されている。インテグレートデバイスに実装されうる好適なグレーティングカプラー及びウェーブガイドの例は、「光カプラー及びウェーブガイドシステム（ＯＰＴＩＣＡＬ ＣＯＵＰＬＥＲ ＡＮＤ ＷＡＶＥＧＵＩＤＥ ＳＹＳＴＥＭ）」という名称の２０１７年１２月１５日出願の米国特許出願第１５／８４４，４０３号明細書（その全体が参照により組み込まれる）に記載されている。

【０２８２】

追加のフォトニック構造は、サンプルウェルと光検出器との間に位置決めされうるとともに、放出光の検出時にシグナルノイズに寄与するおそれのある光検出器に達する励起光を低減又は予防するように構成されうる。いくつかの実施形態では、インテグレートデバイスの回路として作用しうる金属層はまた、空間フィルターとしても作用しうる。好適なフォトニック構造の例は、分光フィルターと偏光フィルターと空間フィルターとを含みうるとともに、「光拒絶フォトニック構造（ＯＰＴＩＣＡＬ ＲＥＪＥＣＴＩＯＮ ＰＨＯＴＯＮＩＣ ＳＴＲＵＣＴＵＲＥＳ）」という名称の２０１８年７月２３日出願の米国特許出願第１６／０４２，９６８号明細書（その全体が参照により組み込まれる）に記載されている。

【０２８３】

インテグレートデバイスから離れて位置決めされるコンポーネントは、インテグレートデバイスに励起源を位置決め及びアライメントするために使用されうる。かかるコンポーネントは、レンズ、ミラー、プリズム、ウィンドウ、アパーチャ、アテニュエータ、及び／又は光ファイバーをはじめとする光学コンポーネントを含みうる。追加の機械的コンポーネントは、１つ以上のアライメントコンポーネントの制御を可能にするために機器に含まれうる。かかる機械的コンポーネントは、アクチュエーター、ステッパーモータ、及び／又はノブを含みうる。好適な励起源及びアライメント機構の例は、「パルスレーザー及びシステム（ＰＵＬＳＥＤ ＬＡＳＥＲ ＡＮＤ ＳＹＳＴＥＭ）」という名称の２０１６年５月２０日出願の米国特許出願第１５／１６１，０８８号明細書（その全体が参照により組み込まれる）に記載されている。ビームステアリングモジュールの他の一例は、「コンパクトビーム整形及びステアリングアセンブリー（ＣＯＭＰＡＣＴ ＢＥＡＭ ＳＨＡＰＩＮＧ ＡＮＤ ＳＴＥＥＲＩＮＧ ＡＳＳＥＭＢＬＹ）」という名称の２０１７年１２月１４日出願の米国特許出願第１５／８４２，７２０号明細書（参照により本明細書に組み込まれる）に記載されている。好適な励起源の追加の例は、「分子をプローブ、検出、及び分析するためのインテグレートデバイス（ＩＮＴＥＧＲＡＴＥＤ ＤＥＶＩＣＥ ＦＯＲ ＰＲＯＢＩＮＧ，ＤＥＴＥＣＴＩＮＧ ＡＮＤ ＡＮＡＬＹＺＩＮＧ ＭＯＬＥＣＵＬＥＳ）」という名称の２０１５年８月７日出願の米国特許出願第１４／８２１，６８８号明細書（その全体が参照により組み込まれる）に記載されている。

【０２８４】

インテグレートデバイスの個別ピクセルと共に位置決めされる光検出器は、ピクセルの対応するサンプルウェルからの放出光を検出するように構成及び位置決めされうる。好適な光検出器の例は、「受け取った光子の時間ビニングのためのインテグレートデバイス（ＩＮＴＥＧＲＡＴＥＤ ＤＥＶＩＣＥ ＦＯＲ ＴＥＭＰＯＲＡＬ ＢＩＮＮＩＮＧ ＯＦ ＲＥＣＥＩＶＥＤ ＰＨＯＴＯＮＳ）」という名称の２０１５年８月７日出願の米国特許出願第１４／８２１，６５６号明細書（その全体が参照により組み込まれる）に記載されている。いくつかの実施形態では、サンプルウェル及びそのそれぞれの光検出器は、共通軸に沿ってアライメントされうる。こうすると、光検出器は、ピクセル内のサンプルウェルにオーバーラップしうる。

【０２８５】
検出放出光の特性は、放出光に関連付けられるマーカーを同定するための指標を提供しうる。かかる特性は、光検出器により検出される光子の到着時間、光検出器により経時的に蓄積される光子の量、及び／又は２つ以上の光検出器にわたる光子の分布をはじめとするいずれかの好適なタイプの特性を含みうる。いくつかの実施形態では、光検出器は、サンプルの放出光に関連付けられる１つ以上のタイミング特性（たとえば発光寿命）の検出を可能にする構成を有しうる。光検出器は、励起光のパルスがインテグレートデバイスを介して伝搬した後の光子到着時間の分布を検出しうるとともに、到着時間の分布は、サンプルの放出光のタイミング特性の指標（たとえば、発光寿命のプロキシ）を提供しうる。いくつかの実施形態では、１つ以上の光検出器は、マーカーにより放出される放出光の確率の指標（たとえば発光強度）を提供する。いくつかの実施形態では、複数の光検出器は、放出光の空間分布をキャプチャーするようにサイズ決め及び配置されうる。次いで、１つ以上の光検出器からの出力シグナルは、複数のマーカー間でマーカーを区別するために使用されうるとともに、複数のマーカーは、サンプル中のサンプルを同定するために使用されうる。いくつかの実施形態では、サンプルは、複数の励起エネルギーにより励起されうるとともに、複数の励起エネルギーに反応してサンプルにより放出される放出光及び／又は放出光のタイミング特性は、マーカーを複数のマーカーから区別しうる。

【０２８６】
動作時、励起光を用いてウェル内のサンプルの一部又は全部を励起することと、サンプル放出からのシグナルを光検出器で検出することと、により、サンプルウェル内のサンプルのパラレル分析が行われる。サンプルからの放出光は、対応する光検出器により検出されて少なくとも１つの電気シグナルに変換されうる。電気シグナルは、インテグレートデバイスの回路内の伝導ラインに沿って伝達されうるとともに、インテグレートデバイスにインターフェースされた機器に接続されうる。電気シグナルは、後続処理及び／又は解析されうる。電気シグナルの処理又は解析は、機器上又は機器外のどちらかに位置する好適なコンピューティングデバイスで行われうる。

【０２８７】
機器は、機器及び／又はインテグレートデバイスの動作を制御するためのユーザインターフェースを含みうる。ユーザインターフェースは、ユーザによる機器への情報の入力、たとえば、機器の機能を制御するために使用されるコマンド及び／又は設定値などの情報の入力を可能にするように構成されうる。いくつかの実施形態では、ユーザインターフェースは、ボタン、スイッチ、ダイヤル、及び音声コマンド用マイクロフォンを含みうる。ユーザインターフェースは、機器及び／又はインテグレートデバイスの性能に関するフィードバック、たとえば、適正アライメント及び／又はインテグレートデバイス上の光検出器からのリードアウトシグナルにより得られる情報、のユーザによる受取りを可能にしうる。いくつかの実施形態では、ユーザインターフェースは、オーディブルフィードバックを提供するためにスピーカーを用いたフィードバックを提供しうる。いくつかの実施形態では、ユーザインターフェースは、ユーザにビジュアルフィードバックを提供するためにインジケータライト及び／又はディスプレイスクリーンを含みうる。

【０２８８】
いくつかの実施形態では、機器は、コンピューティングデバイスに接続するように構成されたコンピュータインターフェースを含みうる。コンピュータインターフェースは、ＵＳＢインターフェース、ファイヤワイヤ（ＦｉｒｅＷｉｒｅ）（商標）インターフェース、又はいずれかの他の好適なコンピュータインターフェースでありうる。コンピューティングデバイスは、ラップトップコンピュータやデスクトップコンピュータなどのいずれかの汎用コンピュータでありうる。いくつかの実施形態では、コンピューティングデバイスは、好適なコンピュータインターフェースを経由してワイヤレスネットワークを介してアクセス可能なサーバ（たとえば、クラウドベースサーバ）でありうる。コンピュータインターフェースは、機器とコンピューティングデバイスとの間の情報の通信を促進しうる。コンピューティングデバイスへの機器を制御及び／又は構成するための入力情報は、提供

され得、及びコンピュータインターフェースを介して機器に伝達され得る。機器により発生する出力情報は、コンピュータインターフェースを経由してコンピューティングデバイスにより受け取られうる。出力情報は、機器の性能に関するフィードバック、インテグレートデバイスの性能、及び／又は光検出器のリードアウトシグナルから発生するデータを含みうる。

【０２８９】

いくつかの実施形態では、機器は、インテグレートデバイスの１つ以上の光検出器から受け取ったデータを解析するように及び／又は励起源に制御シグナルを伝達するように構成された処理デバイスを含みうる。いくつかの実施形態では、処理デバイスは、汎用プロセッサー、特化（ｓｐｅｃｉａｌｌｙ－ａｄａｐｔｅｄ）プロセッサー（たとえば、中央処理ユニット（ＣＰＵ）、たとえば、１つ以上のマイクロプロセッサーコア若しくはマイクロコントローラコア、フィールドプログラマブルゲートアレイ（ＦＰＧＡ）、特定用途向け集積回路（ＡＳＩＣ）、カスタム集積回路、ディジタルシグナルプロセッサー（ＤＳＰ）、又はそれらの組合せ）を含みうる。いくつかの実施形態では、１つ以上の光検出器からのデータの処理は、機器の処理デバイス及び外部コンピューティングデバイスの両方により実施されうる。他の実施形態では、外部コンピューティングデバイスは省略されうるとともに、１つ以上の光検出器からのデータの処理は、インテグレートデバイスの処理デバイスのみにより実施されうる。

【０２９０】

いくつかの実施形態によれば、発光放出特性に基づいてサンプルを解析するように構成された機器は、異なる発光分子間の発光寿命及び／若しくは強度の差並びに／又は異なる環境の同一発光分子の寿命及び／若しくは強度の差を検出しうる。発光放出寿命の差は、異なる発光分子の存在若しくは不在を見分けるために及び／又は発光分子が晒される異なる環境若しくは条件を見分けるために使用可能であることを、本発明者らは認識し高く評価した。いくつかの場合には、寿命（たとえば、発光波長ではない）に基づいて発光分子を見分けることは、システムの態様を単純化しうる。例として、波長識別光学素子（たとえば、波長フィルター、各波長の専用検出器、異なる波長の専用パルス光源、及び／又は回折光学素子）は、寿命に基づいて発光分子を見分けるときに数が低減されうるか又は排除されうる。いくつかの場合には、単一特性波長で動作する単一パルス光源は、光学スペクトルの同一波長領域内で放出するが測定可能に異なる寿命を有する異なる発光分子を励起するために使用されうる。同一波長領域に放出する異なる発光分子を励起して見分けるために、異なる波長で動作する複数の光源ではなく単一のパルス光源を使用する解析システムは、動作及び維持の複雑さが軽減され、よりコンパクトになりうるとともに、より低コストで製造されうる。

【０２９１】

発光寿命解析に基づく解析システムはある特定の利益を有しうるが、解析システムにより得られる情報量及び／又は検出確度は、追加の検出技術を許容すれば増加させうる。たとえば、システムのいくつかの実施形態は、発光波長及び／又は発光強度に基づいてサンプルの１つ以上の性質を見分けるように追加的に構成されうる。いくつかの実現形態では、発光強度は、異なる発光標識を区別するために追加的又は代替的に使用されうる。たとえば、いくつかの発光標識は、それらの減衰率が類似しているおそれがあったとしても、有意に異なる強度で放出しうるか又はそれらの励起確率に有意差（たとえば、少なくとも約３５％の差）を有しうる。測定励起光に対してビンシグナルを参照することにより、強度レベルに基づいて異なる発光標識を区別可能でありうる。

【０２９２】

いくつかの実施形態によれば、異なる発光寿命は、発光標識の励起に続く発光放出イベントを時間ビンするように構成された光検出器を用いて区別されうる。時間ビニングは、光検出器に対する単一電荷蓄積サイクル時に発生しうる。電荷蓄積サイクルは、光発生キャリヤーが時間ビニング光検出器のビンに蓄積されるリードアウトイベント間のインターバルである。時間ビニング光検出器の例は、「受け取った光子の時間ビニングのためのイ

ンテグレートデバイス（ＩＮＴＥＧＲＡＴＥＤ　ＤＥＶＩＣＥ　ＦＯＲ　ＴＥＭＰＯＲＡＬ　ＢＩＮＮＩＮＧ　ＯＦ　ＲＥＣＥＩＶＥＤ　ＰＨＯＴＯＮＳ）」という名称の２０１５年８月７日出願の米国特許出願第１４／８２１，６５６号明細書（参照により本明細書に組み込まれる）に記載されている。いくつかの実施形態では、時間ビニング光検出器は、光子吸収／キャリヤー発生領域で電荷キャリヤーを発生して電荷キャリヤー貯蔵領域の電荷キャリヤー貯蔵ビンに電荷キャリヤーを直接移送しうる。かかる実施形態では、時間ビニング光検出器は、キャリヤー移動／キャプチャー領域を含まないこともありうる。かかる時間ビニング光検出器は、「直接ビニングピクセル」といいうる。直接ビニングピクセルを含む時間ビニング光検出器の例は、「直接ビニングピクセルを備えたインテグレート光検出器（ＩＮＴＥＧＲＡＴＥＤ　ＰＨＯＴＯＤＥＴＥＣＴＯＲ　ＷＩＴＨ　ＤＩＲＥＣＴ　ＢＩＮＮＩＮＧ　ＰＩＸＥＬ）」という名称の２０１７年１２月２２日出願の米国特許出願第１５／８５２，５７１号明細書（参照により本明細書に組み込まれる）に記載されている。

【０２９３】

いくつかの実施形態では、各試薬が発光強度に基づいて同定されうるように、同一タイプの異数のフルオロフォアがサンプル中の異なる試薬に連結されうる。たとえば、２つのフルオロフォアは、第１の標識付認識分子に連結されうるとともに、４つ以上のフルオロフォアは、第２の標識付認識分子に連結されうる。異数のフルオロフォアであるため、異なる認識分子に関連付けられる異なる励起確率及びフルオロフォア放出確率が存在しうる。たとえば、第２の標識付認識分子では、シグナル蓄積インターバル時、より多くの放出イベントが存在しうるので、ビンの見掛けの強度は、第１の標識付認識分子よりも有意に高くなる。

【０２９４】

フルオロフォア減衰率及び／又はフルオロフォア強度に基づいてヌクレオチド又はいずれかの他の生物学的若しくは化学的サンプルを区別することは、光学励起及び検出システムの単純化を可能にしうることを、本発明者らは認識し高く評価した。たとえば、光学励起は、単一波長光源（たとえば、複数の光源ではなく１つの特性波長を生成する光源又は複数の異なる特性波長で動作する光源）で実施されうる。そのほか、波長識別光学素子及びフィルターは、検出システムに必要とされないこともありうる。また、異なるフルオロフォアからの放出を検出するために、各サンプルウェルに対して単一光検出器が使用されうる。「特性波長」又は「波長」という語句は、放射線の限定バンド幅内の中心波長又は主波長（たとえば、パルス光源による２０ｎｍバンド幅出力内の中心波長又はピーク波長）を意味するものとして用いられる。いくつかの場合には、「特性波長」又は「波長」は、光源による放射線出力の合計バンド幅内のピーク波長を意味するものとして用いられうる。

【０２９５】

計算技術

本出願の態様は、本明細書に記載のポリペプチドシーケンシング技術により発生したデータを解析するための計算技術に関する。以上で考察されるように、たとえば、図１Ａ及び１Ｂとの関連では、これらのシーケンシング技術を用いて発生したデータは、シーケンスされるポリペプチドの末端に露出されるアミノ酸にアミノ酸認識分子が会合するインスタンスの指標となる一連のシグナルパルスを含みうる。一連のシグナルパルスは、分解プロセスが逐次アミノ酸を除去して進行するにつれて経時的にその時点の末端のアミノ酸のタイプに依存して、さまざまな１つ以上のフィーチャー（たとえば、パルス持続時間、パルス間持続時間、大きさの変化）を有しうる。得られるシグナルトレースは、それぞれのアミノ酸に関連付けられるさまざまな１つ以上のフィーチャーから生じる特性パターンを含みうる。本明細書に記載の計算技術は、アミノ酸配列を同定するためにこれらのシーケンシング技術を用いて得られるかかるデータを解析する部分として実現されうる。

【０２９６】

いくつかの実施形態は、ポリペプチドの分解プロセス時のデータを得ることと、分解プ

ロセス時にポリペプチドの末端に逐次露出されるアミノ酸に対応するデータの部分を決定するようにデータを解析することと、ポリペプチドに相当するアミノ酸配列を出力すること、を含みうる。図１１は、本明細書に記載のポリペプチドシーケンシング技術を用いて得られるデータを解析することによりアミノ酸配列を同定するための例示的処理パイプライン１１００の図である。図１１に示されるように、シーケンシングデータ１１０２を解析することは、会合イベント同定技術１１０４とアミノ酸同定技術１１０６とを用いてアミノ酸配列１１０８を出力することを含みうる。

【０２９７】

本明細書で考察されるように、シーケンシングデータ１１０２は、ポリペプチドの分解プロセス時に得られうる。いくつかの実施形態では、シーケンシングデータ１１０２は、分解プロセス時のポリペプチドの末端のアミノ酸アイデンティティーの指標となる。いくつかの実施形態では、シーケンシングデータ１１０２は、分解プロセス時に末端の異なるタイプの末端アミノ酸に結合する１つ以上のアミノ酸認識分子により生成されるシグナルの指標となる。模範的シーケンシングデータは、以上で考察される図１Ａ及び１Ｂに示される。

【０２９８】

分解プロセス時にシグナルがどのように発生するかに依存して、シーケンシングデータ１１０２は、１つ以上の異なるタイプのシグナルの指標となりうる。いくつかの実施形態では、シーケンシングデータ１１０２は、分解プロセス時に発生する発光シグナルの指標となる。たとえば、発光標識は、アミノ酸認識分子を標識するために使用されうるとともに、発光標識により放出される発光は、アミノ酸認識分子が特定アミノ酸に会合して発光シグナルをもたらすときに検出されうる。いくつかの実施形態では、シーケンシングデータ１１０２は、分解プロセス時に発生する電気シグナルの指標となる。たとえば、シーケンスされるポリペプチド分子は、ナノポアに固定されうるとともに、電気シグナル（たとえば、コンダクタンスの変化）は、アミノ酸認識分子が特定アミノ酸に会合するときに検出されうる。

【０２９９】

いくつかの実施形態は、分解プロセス時にポリペプチドの末端に逐次露出されるアミノ酸に対応するシーケンシングデータ１１０２の部分を決定するようにシーケンシングデータ１１０２を解析することを含む。図１１に示されるように、会合イベント同定技術１１０４は、シーケンシングデータ１１０２にアクセスして会合イベントに対応するシーケンシングデータ１１０２の部分を同定するようにシーケンシングデータを解析しうる。会合イベントは、データ中では図１Ｂに示されるＣＰ$_1$やＣＰ$_2$などの特性パターンに対応しうる。いくつかの実施形態では、会合イベント同定技術１１０４は、一連の切断イベントを検出することと、逐次切断イベント間のシーケンシングデータ１１０２の部分を決定することと、を含みうる。例として、図１Ｂに示されるＣＰ$_1$とＣＰ$_2$との間の切断イベントは、ＣＰ$_1$に対応するデータの第１の部分が第１の会合イベントとして同定されうるとともに、ＣＰ$_2$に対応するデータの第２の部分が第２の会合イベントとして同定されうるように、検出されうる。

【０３００】

いくつかの実施形態は、シーケンシングデータ１１０２の決定部分の１つ以上に対してアミノ酸のタイプを同定することを含む。図１１に示されるように、アミノ酸同定技術１１０６は、会合イベント同定技術１１０４により同定された会合イベントの１つ以上に対してアミノ酸のタイプを決定するために使用されうる。いくつかの実施形態では、会合イベント同定技術１１０４より同定されるデータの個別部分は、パルスパターンを含みうるとともに、アミノ酸同定技術１１０６は、そのそれぞれのパルスパターンに基づいて１つ以上の部分に対してアミノ酸のタイプを決定しうる。図１Ｂを参照して、アミノ酸同定技術１１０６は、ＣＰ$_1$に対して第１のタイプのアミノ酸及びＣＰ$_2$に対して第２のタイプのアミノ酸を同定しうる。いくつかの実施形態では、アミノ酸のタイプを決定することは、データが閾値超であるときの会合イベント同定技術１１０４を用いて同定される部分な

どのデータの部分内の時間量を同定することと、データの部分に対して時間量と持続時間とを比較することと、を含みうる。たとえば、CP_1に対してアミノ酸のタイプを同定することは、シグナルが閾値超である場合のCP_1内の時間量、たとえば、シグナルがM_L超である場合の時間帯ｐｄを決定することと、CP_1に対してそれを合計持続時間と比較することと、を含みうる。いくつかの実施形態では、アミノ酸のタイプを決定することは、会合イベント同定技術１１０２により同定されたデータの１つ以上の部分に対して１つ以上のパルス持続時間を同定することを含みうる。たとえば、CP_1に対してアミノ酸のタイプを同定することは、CP_1に対して時間帯ｐｄなどのパルス持続時間を決定することを含みうる。いくつかの実施形態では、アミノ酸のタイプを決定することは、会合イベント同定技術１１０４を用いて同定されたデータの１つ以上の部分に対して１つ以上のパルス間持続時間を同定することを含みうる。たとえば、CP_1に対してアミノ酸のタイプを同定することは、ｉｐｄなどのパルス間持続時間を同定することを含みうる。

【０３０１】

シーケンシングデータ１１０２の逐次部分に対してアミノ酸のタイプを同定することにより、アミノ酸同定技術１１０６は、ポリペプチドに相当するアミノ酸配列１１０８を出力しうる。いくつかの実施形態では、アミノ酸配列は、会合イベント同定技術１１０４を用いて同定されたデータの部分に対応する一連のアミノ酸を含む。

【０３０２】

図１２は、本明細書に記載の技術のいくつかの実施形態に係るポリペプチド分子のアミノ酸配列を決定するための例示的プロセス１２００のフローチャートである。プロセス１２００は、本明細書に記載の技術の態様がこの態様に限定されないので、いずれかの好適なコンピューティングデバイス（たとえば、単一コンピューティングデバイス、単一物理的位置に共配置された又は互いに離れた複数の物理的位置に配置された複数のコンピューティングデバイス、クラウドコンピューティングシステムの１つ以上のコンピューティングデバイス部分など）で実施されうる。いくつかの実施形態では、会合イベント同定技術１１０４及びアミノ酸同定技術１１０６は、アミノ酸配列を決定するためにプロセス１２００の一部又は全部を実施しうる。

プロセス１２００は、単一ポリペプチド分子と１つ以上の末端アミノ酸認識分子とを接触させることを含む動作１２０２で始まる。次に、プロセス１２００は、１つ以上の末端アミノ酸認識分子と、単一ポリペプチドが分解されつつある間に単一ポリペプチドの末端に露出される逐次アミノ酸と、の会合の指標となる一連のシグナルパルスを検出することを含む動作１１０４に進む。一連のパルスは、たとえば、会合イベント同定技術１１０４及びアミノ酸同定技術１１０６を用いることにより、単一ポリペプチド分子のシーケンシングを可能にしうる。

【０３０３】

いくつかの実施形態では、プロセス１２００は、たとえば、アミノ酸同定技術１１０６を用いることにより、一連のシグナルパルスの第１の特性パターンに基づいて単一ポリペプチド分子の第１のタイプのアミノ酸を同定することを含む動作１２０６を含みうる。

【０３０４】

図１３は、本明細書に記載の技術のいくつかの実施形態に係るポリペプチドに相当するアミノ酸配列を決定するための例示的プロセス１３００のフローチャートである。プロセス１３００は、本明細書に記載の技術の態様がこの態様に限定されないので、いずれかの好適なコンピューティングデバイス（たとえば、単一コンピューティングデバイス、単一物理的位置に共配置された又は互いに離れた複数の物理的位置に配置された複数のコンピューティングデバイス、クラウドコンピューティングシステムの１つ以上のコンピューティングデバイス部分など）で実施されうる。いくつかの実施形態では、会合イベント同定技術１１０４及びアミノ酸同定技術１１０６は、アミノ酸配列を決定するためにプロセス１３００の一部又は全部を実施しうる。

【０３０５】

プロセス１３００は、ポリペプチドの分解プロセス時のデータが得られる動作１３０２

で始まる。いくつかの実施形態では、データは、分解プロセス時のポリペプチドの末端のアミノ酸アイデンティティー（ｉｄｅｎｔｉｔｙ）の指標となる。いくつかの実施形態では、データは、分解プロセス時に末端の異なるタイプの末端アミノ酸に結合する１つ以上のアミノ酸認識分子により生成されるシグナルの指標となる。いくつかの実施形態では、データは、分解プロセス時に発生する発光シグナルの指標となる。いくつかの実施形態では、データは、分解プロセス時に発生する電気シグナルの指標となる。

【０３０６】

次に、プロセス１３００は動作１３０４に進み、そこでデータが解析されて、たとえば、会合イベント同定技術１１０４及びアミノ酸同定技術１１０６を用いることにより、分解プロセス時にポリペプチドの末端に逐次露出されるアミノ酸に対応するデータの部分が決定される。いくつかの実施形態では、データを解析することは、一連の切断イベントを検出することと、たとえば、会合イベント同定技術１１０４を用いることにより、逐次切断イベント間のデータの部分を決定することと、をさらに含む。

【０３０７】

いくつかの実施形態では、データを解析することは、たとえば、アミノ酸同定技術１１０６を用いることにより、個別部分の各々に対してアミノ酸のタイプを決定することをさらに含む。いくつかの実施形態では、個別部分の各々は、パルスパターンを含み、且つデータを解析することは、そのそれぞれのパルスパターンに基づいて１つ以上の部分に対してアミノ酸のタイプを決定することをさらに含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、データが閾値超であるときの部分内の時間量を同定することと、部分に対して時間量と持続時間とを比較することと、をさらに含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、１つ以上の部分の各々に対して少なくとも１つのパルス持続時間を同定することをさらに含む。いくつかの実施形態では、アミノ酸のタイプを決定することは、１つ以上の部分の各々に対して少なくとも１つのパルス間持続時間を同定することをさらに含む。

【０３０８】

次に、プロセス１３００は動作１３０６に進み、そこでたとえば、ユーザインターフェースを介して、ポリペプチドに相当するアミノ酸配列が出力される。いくつかの実施形態では、アミノ酸配列は、部分に対応する一連のアミノ酸を含む。

【０３０９】

本明細書に記載の技術の実施形態のいずれかと組み合わせて使用されうるコンピュータシステム１４００の例示的実現形態が図１４に示される。コンピュータシステム１４００は、１つ以上のプロセッサー１４１０と、非一時コンピュータ可読記憶媒体（たとえば、メモリー１４２０及び１つ以上の不揮発性記憶媒体１４３０）を含む１つ以上の製造品と、を含む。プロセッサー１４１０は、メモリー１４２０及び不揮発性記憶デバイス１４３０に対するデータ書込み及びデータ読出しをいずれかの好適な方式で制御しうるが、本明細書に記載の技術の態様は、この点に限定されるものではない。本明細書に記載の機能性のいずれかを実施するために、プロセッサー１４１０は、プロセッサー１４１０により実行されるプロセッサー実行可能命令を記憶する非一時コンピュータ可読記憶媒体として機能しうる１つ以上の非一時コンピュータ可読記憶媒体（たとえば、メモリー１４２０）に記憶された１つ以上のプロセッサー実行可能命令を実行しうる。

【０３１０】

コンピューティングデバイス１４００はまた、コンピューティングデバイスと他のコンピューティングデバイスとの通信（たとえば、ネットワークを介して）を媒介しうるネットワーク入出力（Ｉ／Ｏ）インターフェース１４４０を含みうるとともに、コンピューティングデバイスによるユーザへの出力の提供及びユーザからの入力の受取りを媒介しうる１つ以上のユーザＩ／Ｏインターフェース１４５０も含みうる。ユーザＩ／Ｏインターフェースとしては、キーボード、マウス、マイクロフォン、ディスプレイデバイス（たとえば、モニター又はタッチスクリーン）、スピーカー、カメラ、及び／又は各種他のタイプのＩ／Ｏデバイスなどのデバイスが挙げられうる。

【０３１１】
以上に記載の実施形態は、多くの手段のいずれかで実現可能である。たとえば、実施形態は、ハードウェア、ソフトウェア、又はそれらの組合せを用いて実現しうる。ソフトウェアで実現する場合、ソフトウェアコードは、単一コンピューティングデバイスに提供されるか又は複数のコンピューティングデバイスに分散されるかにかかわらず、いずれかの好適なプロセッサー（たとえば、マイクロプロセッサー）又は一群のプロセッサーで実行可能である。以上に記載の機能を実施する任意のコンポーネント又は一群のコンポーネントは、以上で考察された機能を制御する１つ以上のコントローラであると一般にみなしうることを認識すべきである。１つ以上のコントローラは、多くの方法で、たとえば、以上に述べた機能を実施するマイクロコード又はソフトウェアを用いてプログラムされた専用ハードウェアを用いて又は汎用ハードウェア（たとえば、１つ以上のプロセッサー）を用いて、実現可能である。

【０３１２】
これに関連して、本明細書に記載の実施形態の一実現形態（ｉｍｐｌｅｍｅｎｔａｔｉｏｎ）は、１つ以上のプロセッサーで実行するときに１つ以上の実施形態の以上で考察される機能を実施するコンピュータプログラム（すなわち、複数の実行可能命令）によりコードされる少なくとも１つのコンピュータ可読記憶媒体（たとえば、ＲＡＭ、ＲＯＭ、ＥＥＰＲＯＭ、フラッシュメモリー、又は他のメモリー技術、ＣＤ－ＲＯＭ、ディジタル多用途ディスク（ＤＶＤ）、又は他の光ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置、又は他の磁気記憶デバイス、又は他のタンジブル非一時コンピュータ可読記憶媒体）を含むことが認識されるべきである。コンピュータ可読媒体は、本明細書で考察される技術の態様を実現するために、記憶されたプログラムをいずれかのコンピューティングデバイスにロードできるようにトランスポート可能でありうる。そのほか、実行時に以上で考察された機能のいずれかを実施するコンピュータプログラムへの参照は、ホストコンピュータ上で実行されるアプリケーションプログラムに限定されるものではないことが認識されるべきである。むしろ、コンピュータプログラム及びソフトウェアという用語は、本明細書で考察される技術の態様を実現するように１つ以上のプロセッサーをプログラムするために利用可能ないずれかのタイプのコンピュータコード（たとえば、アプリケーションソフトウェア、ファームウェア、マイクロコード、又はいずれかの他の形態のコンピュータ命令）を参照するために一般的な意味で本明細書で用いられる。

【実施例】

【０３１３】
実施例１．ペプチド表面固定用の可溶化性リンカー
水性緩衝液へのオリゴペプチドの溶解性の改善を試みて、ペプチド断片をオリゴヌクレオチドリンカーにコンジュゲートすることにより、水への溶解性を改善しうるとともに単一分子レベルでペプチドを表面固定するための機能部分も提供しうると判断された。さまざまなペプチド－リンカーコンジュゲートが合成され、ペプチド－ＤＮＡコンジュゲート及びペプチド－ＰＥＧコンジュゲートに関して構造例が図１５Ａに描かれる。リンカーコンジュゲーションは、評価されたさまざまなペプチド－リンカーコンジュゲートの各々で水性溶液へのペプチドの溶解性を大きく増強することが観測された。

【０３１４】
Ｎ末端アミノペプチダーゼによるペプチドＮ末端のアミノ酸切断に関して、ペプチド－リンカーコンジュゲートを評価した（以下の表６）。

【０３１５】

【表６－１】

表６．ペプチドリンカーコンジュゲートの末端アミノ酸切断。

エントリー	ペプチド	配列番号	クラス	リンカー	ラットAPNによる切断	PIPによる切断
1	KF	223	正	オリゴ	無	
2	KKMKKM{LYS(N3)}	224	正	オリゴ	無	
3	KKMKKM{LYS(N3)}	225	正	オリゴ-PEG	無	
4	KKMKKM{LYS(N3)}	226	正	PEG4	有	
5	DDMDDM{LYS(N3)}	227	負	オリゴ	有	
6	FFMFFM{LYS(N3)}	228	芳香族	オリゴ	有	

【０３１６】
【表６－２】

7	AAMAAM{LYS(N3)}	229	疎水性	オリゴ	有	
8	FPFPFP{LYS(N3)}	230	芳香族	オリゴ		有
9	DPDPDP{LYS(N3)}	231	負	オリゴ		有
10	KPKPKP{LYS(N3)}	232	正	オリゴ		無
11	KPKPKP{LYS(N3)}	233	正	PEG4		有

【０３１７】
表６に示されるペプチド－リンカーコンジュゲートは、プロリンイミノペプチダーゼ（「ＰＩＰ」）又はラットアミノペプチダーゼＮ（「ラットＡＰＮ」）のどちらかと共にインキュベートされ、且つペプチド切断は、ＬＣＭＳによりモニターされた。表６のエントリー５の切断を実証するＬＣＭＳの例は、図１５Ｂに示される。すべての他の切断反応は、同様に測定された。表６に示されるように、正荷電ペプチド－ＤＮＡコンジュゲート（「オリゴ」及び「オリゴ－ＰＥＧ」リンカー）は、試験アミノペプチダーゼにより切断されず、ＤＮＡオリゴヌクレオチドリンカーを有するすべての他のコンジュゲートクラス（負荷電、芳香族、疎水性）は切断された。比較により、正荷電ペプチド－ＰＥＧコンジュゲートは、アミノペプチダーゼの少なくとも１つにより切断されることが示された。
【０３１８】
標識付ペプチド－リンカーコンジュゲートを用いて、異なる組成のペプチドを単一分子分析のために個別サンプルウェル表面に固定可能であることが示された。これらの実験では、ＤＮＡリンカーは、色素で標識され（たとえば、ペプチド－ＤＮＡコンジュゲートの場合、図１５Ａに描かれる通りである）、且つ個別サンプルウェルへの異なるペプチド－ＤＮＡコンジュゲートのローディングは、色素蛍光により測定された。ローディング実験例は、図１５Ｃに示される。標識付ペプチド－ＤＮＡコンジュゲート（５０ｐＭ）の蛍光放出を測定することにより、チップ上のサンプルウェルの少なくとも１８％は、表面固定コンジュゲートを有してサンプルウェル当たり単一占有でロードされたと判断された。これらの実験は、ペプチド－リンカーコンジュゲートが非コンジュゲートペプチドカウンターパートと比較して増強された水への溶解性を呈すること、コンジュゲートリンカーが異なるアミノペプチダーゼによるペプチドの末端アミノ酸切断を防止しないこと、且つ異なる組成のペプチド－リンカーコンジュゲートが単一分子レベルでチップ表面に固定されうることを実証した。
【０３１９】
実施例２．ポリペプチド基質のエキソペプチダーゼ切断
さまざまなアミノペプチダーゼの切断能力を試験した。ペプチド基質の切断は、高速液

体クロマトグラフィー（ＨＰＬＣ）を使用してアッセイされた。選ばれたエキソペプチダーゼのアミノ酸切断活性の概要は、図１６に示される。特異的切断活性は、次の酵素：「ｃＶＰｒ」（Ｖ．プロテオリティカス（Ｖ．ｐｒｏｔｅｏｌｙｔｉｃｕｓ）アミノペプチダーゼ）、「ｙＰＩＰ」（Ｙ．ペスティス（Ｙ．ｐｅｓｔｉｓ）プロリンイミノペプチダーゼ）、「Ｄ／Ｅ　ＡＰＮ」（Ｌ．ニューモフィラ（Ｌ．ｐｎｅｕｍｏｐｈｉｌａ）Ｍ１アミノペプチダーゼ）、ｈＴＥＴ（パイロコッカス・ホリコシイ（Ｐｙｒｏｃｏｃｃｕｓ　ｈｏｒｉｋｏｓｈｉｉ）ＴＥＴアミノペプチダーゼ）、及びＰｆｕ　ＡＰＩ（「ＰｆｕＴＥＴ」）に対して示される。末端アミノ酸に対する比活性は、アミノ酸に一文字略号を用いて、示されるように分類される（「ＸＰ－」は、隣接又はペナルティメートのプロリン（Ｐ）残基を有するいずれかの末端アミノ酸（Ｘ）を表す）。

【０３２０】

実施例３．単一分子レベルの固定ペプチドの末端アミノ酸切断

固定ペプチドのオンチップアミノ酸切断のアッセイは、標識付ペプチドコンジュゲートを用いて開発された。アッセイは、固定ペプチドに対するエキソペプチダーゼの酵素認識及び切断活性を決定する方法を提供するように設計された。これは速度論的結合パラメータ及び一般的結合親和性の測定を可能にしうる。

【０３２１】

ペプチドのＮ末端アミノ酸切断を評価するために、ＰＥＧスペーサーを介して色素に装着されたＮ末端アスパラギン酸を含有する色素標識付ペプチドを設計及び合成した。このペプチドはまた、酵素プロリンイミノペプチダーゼ（エルシニア・ペスティス（Ｙｅｒｓｉｎｉａ　ｐｅｓｔｉｓ）由来であり、他でも知られ、本明細書では「ｙＰＩＰ」といわれる）により特異的に認識される修飾アスパラギン酸に隣接するプロリン残基も含有していた。酵素ｙＰＩＰは、プロリン残基の上流のＮ末端アミノ酸のみを切断するはずである。

【０３２２】

この及び他の標識付ペプチドがバルクでｙＰＩＰにより効率的にカットされることを示した後（たとえば、実施例１に記載の通り）、単一分子レベルでＮ末端アミノ酸切断を観測するために、オンチップ色素／ペプチドコンジュゲートアッセイを開発した。図１７Ａは、色素／ペプチドコンジュゲートアッセイの一般的スキームを示す（インセットパネル）。示されるように、スペーサーを介してＮ末端アミノ酸に装着された標識を有するペプチドは、リンカーを介して表面に固定される。ペプチダーゼに暴露された後、Ｎ末端アミノ酸切断は、検出可能観測容積からの標識付残基の除去をもたらすとともに、標識からのシグナルの付随的損失により測定される。インセットパネルの右側の酵素－ペプチド複合体は、Ｎ末端切断部位を一般的に描く。

【０３２３】

図１７Ａは、色素／ペプチドコンジュゲートアッセイに使用するように設計及び合成された標識付ペプチド構築物を示す（ボトム）。これらの実験では、Ｎ末端にペナルティメート（ｐｅｎｕｌｔｉｍａｔｅ）プロリン残基を有するペプチドのＮ末端アスパラギン酸残基にローダミン色素（アトー（ＡＴＴＯ）Ｒｈｏ６Ｇ）を装着した。示されるように、表面固定のためにビオチン部分を有する可溶化性ＤＮＡリンカーにペプチドをさらにコンジュゲートした。

【０３２４】

標識付ペプチドコンジュゲートは、サンプルウェルのアレイを有するガラスチップ上にロードされた。ローダミン蛍光により単一占有のサンプルウェルのパーセントローディングを決定するために、ローディングの前及び後でチップの画像を取得した。次いで、ロードされたチップ上に酵素ｙＰＩＰを導入し、３７℃で２時間インキュベートした。ｙＰＩＰの導入に続いてチップの画像を取得して失われた緑色色素のパーセントを計算することにより、Ｎ末端アミノ酸切断を評価した。図１７Ｂは、ローディングステージへの６～７％ローディングと、ｙＰＩＰと共にインキュベートした後、事前にロードされたウェルで９１％シグナル損失（これはＮ末端アミノ酸切断の指標となる）と、を呈した実験のイメ

ージング結果を示す。図１７Ｃは、これらの実験の代表的シグナルトレースを示しており、標識付ペプチドのローディング時に色素シグナルの増加が検出されるとともに、ｙＰＩＰへの暴露に続いて色素シグナルの損失が検出されることが実証される。

【０３２５】

単一分子レベルでのＮ末端アミノ酸切断のさらなる確認として、エキソペプチダーゼ認識及び切断活性を評価するために、オンチップＦＲＥＴアッセイを開発した。図１８Ａは、ＦＲＥＴペプチドコンジュゲートアッセイ（パネルＡ）及びＦＲＥＴ酵素コンジュゲートアッセイ（パネルＢ）を一般的に描く。ＦＲＥＴペプチドコンジュゲートアッセイ（パネルＡ）では、固定ペプチド構築物は、リンカーに装着されたＦＲＥＴドナー標識と、Ｎ末端に装着されたＦＲＥＴアクセプター標識と、を含む。Ｎ末端アミノ酸切断は、ペプチダーゼに暴露したときのＦＲＥＴアクセプター標識からのシグナルの損失により検出される。そのほか、この設計は、ＦＲＥＴドナー標識からの放出を追跡することにより、実験全体を通してペプチドコンジュゲートのローディングのモニタリングを可能にする。

【０３２６】

ＦＲＥＴ酵素コンジュゲートアッセイ（パネルＢ）では、固定ペプチド構築物は、リンカーに装着されたＦＲＥＴペアの第１の標識を含み、且つペプチダーゼは、ＦＲＥＴペアの第２の標識で標識される。Ｎ末端アミノ酸切断は、ペプチドへのペプチダーゼの近接が十分であり且つＮ末端での滞留時間が十分であるときに発生するＦＲＥＴ相互作用に帰属可能な蛍光の増強により検出される。そのほか、このアッセイは、プロセッシブ切断によるＦＲＥＴシグナルの経時的増加を検出することにより、プロセッシブエキソペプチダーゼによるプロセッシブアミノ酸切断の評価を可能にする。

【０３２７】

図１８Ａはまた、パネルＡのＦＲＥＴペプチドコンジュゲートアッセイに使用するために設計及び合成されたＦＲＥＴペプチド構築物をパネルＡに示す。示されるように、ＦＲＥＴペプチド構築物は、Ｎ末端ペナルティメートプロリン残基を有するペプチドのＮ末端アスパラギン酸残基に装着されたローダミン色素（アトー（ＡＴＴＯ）６４７Ｎ）を含んでいた。ペプチドは、ＦＲＥＴのためにシアニン色素（Ｃｙ３Ｂ）と表面固定のためにビオチン部分とに装着された可溶化性ＤＮＡリンカーにさらにコンジュゲートされた。

【０３２８】

この実験では、ＦＲＥＴペプチド構築物は、サンプルウェルのアレイを有するガラスチップ上にロードされ、且つ捕集光は、最初に緑色フィルター及び次いで赤色フィルターによりフィルターされた。ＦＲＥＴペプチド構築物のローディングは、緑色及び赤色フィルターの両方を通り抜けるシグナルを測定することにより検出された。末端アミノ酸切断は、シグナルが緑色フィルターのみで測定可能であるときに検出されたことから、ＦＲＥＴペプチド構築物の赤色色素コンジュゲートＮ末端アミノ酸がｙＰＩＰにより切断されたことが示唆される。この検出パターンはパネルＣに例示される。示されるように、ｙＰＩＰの添加前に両方の色素が検出可能であり且つｙＰＩＰと共にインキュベートした後に緑色色素のみが視認可能である場合、こうした検出パターンの変化はペプチドの切断に起因するものであり全体的な光漂白にもペプチドの損失にも起因しないと、合理的に結論付け可能である。そのほか、単独の緑色色素からの蛍光の増加が予想されるであろう。なぜなら、その放出がもはや赤色色素により吸収されないからである。

【０３２９】

ホスホン酸とシランとを用いて表面パッシベーションにより修飾されたチップへのＦＲＥＴペプチド構築物のローディングに続いて、ｙＰＩＰを導入していくつかの時間点で画像を得た。全体的切断トレンドを評価するために、各実験に対して（緑色）／（緑色＋赤色）の比を計算した。この比は、発生する切断の程度と共に増加する。図１８Ｂは、ｙＰＩＰと共にインキュベートした異なる時間点のすべてのアパーチャにわたるＦＲＥＴ放出比のプロットである。示されるように、蛍光放出比への緑色色素の寄与が経時的に増加することから、ｙＰＩＰと共にインキュベートしたとき、より多くのＮ末端アスパラギン酸残基がカットされて、緑色色素のみのトランケートペプチドが残ることが示唆される。

【０３３０】
次いで、どの時間点で色素蛍光が観測されたかを決定することにより、異なる時間点でカッティング効率を評価した。これは単純な閾値化により行われた。たとえば、平均色素発光シグナルが励起時に＞２．５であった場合、色素は存在すると決定された（各対応するフィルターを適用したとき）。次いで、カッティングを呈するアパーチャは、実験のローディングフェーズ時に緑色色素及び赤色色素の両方を呈するであろうが、ｙＰＩＰに暴露された時間点で緑色色素のみを呈するであろう。図１８Ｃに示されるように、チップがｙＰＩＰと共により長いインキュベーション時間に露出されると、徐々により多くのカッティングが観測された。３つのｙＰＩＰ処理時間点の各々で呈するカッティングを示すシグナルトレース例を図１８Ｄに示す。

【０３３１】
デキストランを用いて表面パッシベーションにより修飾されたチップを用いてｙＰＩＰ及び他のペプチダーゼにより追加の実験を実施したところ、チップ上へのペプチダーゼの導入に続いて経時的な末端アミノ酸切断の増加を示す類似の結果を生じた。図１８Ｅは、ｙＰＩＰと共にインキュベートした異なる時間点のロードされたアパーチャにわたるＦＲＥＴ放出比のプロットである。図１８Ｆは、アミノペプチダーゼと共にインキュベートした異なる時間点のロードされたアパーチャにわたるＦＲＥＴ放出比のプロットである。全体として、ここでの実験は、異なるエキソペプチダーゼ及び異なる標識化ストラテジーを用いて単一分子レベルでリアルタイムにＮ末端アミノ酸切断を検出可能であることを実証する。

【０３３２】
実施例４．標識付認識分子による末端アミノ酸識別
Ｎ末端芳香族残基を検出するために、標識付認識分子として使用するための潜在的候補として、タンパク質分解経路に関与するアダプタータンパク質を同定した。α－プロテオバクテリア（Ａ．ツメファシエンス（Ａ．ｔｕｍｅｆａｃｉｅｎｓ））由来のアダプタータンパク質ＣｌｐＳ２を発現し、露出されたシステイン残基に標識した。図１９Ａは、スティックとして示される露出されたシステイン残基と共にＣｌｐＳ２タンパク質の結晶構造を示す。露出されたシステイン残基は、ローダミン色素（アトー（ＡＴＴＯ）５３２）で標識された。

【０３３３】
標識付ＣｌｐＳ２により単一分子レベルでＮ末端アミノ酸識別が可能であるかを試験するために、異なるＮ末端芳香族残基を有するペプチドを調製した。これらの実験の単一分子強度トレース例を図１９Ｂに示す。示されるように、シグナルトレースは、Ｎ末端フェニルアラニン残基（Ｆ、トップのシグナルトレース）、Ｎ末端チロシン残基（Ｙ、中間のシグナルトレース）、又はＮ末端トリプトファン残基（Ｗ、ボトムのシグナルトレース）のどれかを有するペプチドのＮ末端に可逆結合する標識付認識分子に対応する残基特異的オンオフ結合パターンを実証する。

【０３３４】
図１９Ｃ～１９Ｅに示される結果を用いて、単一分子トラジェクトリーのさらなる解析を行った。図１９Ｃは、標識付ＣｌｐＳ２により可逆結合されたときの３つのＮ末端残基間のディスクリミナントパルス持続時間（シグナルピークの持続時間）を示すプロットである。図１９Ｄは、３つのＮ末端残基間のディスクリミナント（ｄｉｓｃｒｉｍｉｎａｎｔ）パルス間持続時間（シグナルパルス間の持続時間）を示すプロットである。図１９Ｅは、ペプチドＮ末端のフェニルアラニン、チロシン、及びトリプトファン間のディスクリミナントパルス持続時間をさらに例示するプロットを示す。異なるＮ末端残基の平均パルス持続時間は、ヒストグラム（Ａ）～（Ｂ）及びレイヤードヒストグラム（Ｃ）により可視化される。

【０３３５】
ロイシン認識用の標識付認識分子として使用するために、サーモシノコッカス・エロンガタス（Ｔｈｅｒｍｏｓｙｎｏｃｈｏｃｃｕｓ　ｅｌｏｎｇａｔｕｓ）由来の他のアダプ

ータンパク質ＣｌｐＳを評価した。図１９Ｆ～１９Ｈに示される滞留時間解析から得られるデータは、標識付ｔｅＣｌｐＳタンパク質が０．７１秒間の平均パルス持続時間でポリペプチドの末端ロイシン残基との検出可能結合相互作用を生成することを実証した。これらの実験に使用されるｔｅＣｌｐＳタンパク質のアミノ酸配列は表１に示される。

【０３３６】

ロイシン認識用の潜在的試薬としてＡ．ツメファシエンス（Ａ．ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ１及びＳ．エロンガタス（ｅｌｏｎｇａｔｕｓ）ＣｌｐＳ２並びにプロリン認識用の潜在的試薬としてＧＩＤ４を評価するために、類似の実験を行った。図１９Ｉは、Ａ．ツメファシエンス（Ａ．ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ１によるフェニルアラニン、ロイシン、トリプトファン、及びチロシンの区別可能認識を示した滞留時間解析の結果例を示す。図１９Ｊは、Ｓ．エロンガタス（ｅｌｏｎｇａｔｕｓ）ＣｌｐＳ２によるロイシン認識を実証する滞留時間解析の結果例を示す。図１９Ｋ～１９Ｌは、ＧＩＤ４によるプロリン認識を実証する滞留時間解析の結果例を示す。

【０３３７】

さまざまなタイプの末端アミノ酸に結合する１つの認識分子の反応速度を評価するために、Ａ．ツメファシエンス（Ａ．Ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ２（ａｔＣｌｐＳ２－Ｖ１）のさまざまなペプチダーゼへの結合親和性を、結合偏光アッセイにて決定した。図１９Ｍには、ａｔＣｌｐＳ２－Ｖ１とさまざまなＮ末端アミノ酸（フェニルアラニン（左側プロット）、チロシン（中央プロット）、及びトリプロファン（右側プロット））を有するペプチドとのK_Dによる模範的結合曲線が示される。各ペプチドは、隣接したペナルティメート位置にアラニン残基を含有していた。アンサンブル測定に基づいて、ａｔＣｌｐＳ２－Ｖ１との決定されたK_D値は、以下：フェニルアラニンペプチドに対して７４３ｎＭ、チロシンペプチドに対して２０４９ｎＭ、及びトリプトファンペプチドに対して３５１０ｎＭのようであった。

【０３３８】

実施例５．分解時の認識によるポリペプチドシーケンシング

進行中の分解反応時のＮ末端アミノ酸認識によるペプチドシーケンシングを評価するために実験を行った。これらの実験の結果例は、図２０Ａ～２０Ｄに示されており、これらは、同一反応混合物で標識付ＣｌｐＳ２タンパク質及びアミノペプチダーゼを用いてリアルタイムで行われた２つの独立したポリペプチドシーケンシング反応により得られた単一分子強度トレースを示す。各反応では、配列ＹＡＡＷＡＡＦＡＤＤＤＷＫ（配列番号２３４）のポリペプチドは、チップ上にペプチド組成物（１０ｐＭ）を２０分間ロードすることによりＣ末端リシン残基を介してチップ表面に固定され、且つ固定されたペプチドは、標識付認識分子（５００ｎＭのアトー（ＡＴＴＯ）５４２標識付Ａ．ツメファシエンス（Ａ．ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ２－Ｖ１）及びアミノペプチダーゼ切断試薬（８μＭのＶＰｒ）の存在下でモニターされた。

【０３３９】

図２０Ａ及び２０Ｃは、２つの異なるシーケンシングランのシグナルトレースデータを示し、トップのパネル（図２０Ａのパネル１、図２０Ｃのパネル２）はフルトレースを示し、且つボトムのパネル（Ｙ、Ｗ、Ｆ）はフルトレース内の強調領域の各々に対応する拡大領域を示す。図２０Ｂ及び２０Ｄは、それぞれ、図２０Ａ及び２０Ｃに記された対応するパネルのトレースデータのパルス持続時間統計をヒストグラムで示す。各シーケンシングランのフルシグナルトレース（パネル１、２）に示されるように、反応過程全体にわたりシグナルパルスの３つの個別時間インターバルが観測された。拡大領域（パネルＹ、Ｗ、Ｆ）により強調されるように、３つのインターバルは、シグナルパルスのパターンの観測可能な差に基づいて互いに視覚的に区別可能である。

【０３４０】

シグナルパルスデータをさらに解析するために、各時間インターバルでパルス持続時間統計を決定した（図２０Ｂ及び２０Ｄ）。ＣｌｐＳ２を用いた定常状態オンチップ結合アッセイでこれらのアミノ酸に対して個別に観測されたものに対応するようにパルス持続時

間分布の差を決定したところ、シグナルパルス情報は、シーケンシングランからのインターバルと個別アミノ酸結合アッセイからのインターバルとの間で表現型に一貫性があった。

【０３４１】

シグナルパルス情報の解析により確認されるように、各シーケンシングランの進行で観測されるシグナルパルスの３つの時間インターバルは、それぞれ、Ｙ、Ｗ、及びＦの認識パターンに対応する（パネル１、２）。シグナルパルスパターン間の介在時間帯は、Ｎ末端アラニン残基に結合しないＣｌｐＳ２－Ｖ１の選択性に起因する。フルシグナルトレースにより例示されるように、第１のインターバルはＹ認識に対応し、それに続いて、ＶＰｒペプチダーゼがＹ及び２つアラニン残基をカットするので休止があり、続いてＷ認識に対応する第２のインターバルがあり、それに続いて、ＶＰｒペプチダーゼがＷ及び２つアラニン残基をカットするので他の休止があり、最後にＦ認識に対応する第３のインターバルがあり、その後、ＶＰｒペプチダーゼがＦをカットして除去し、残留するＡＤＤＤＷＫペプチドで停止する。進行中の分解反応時に末端アミノ酸認識により得られるパルス持続時間情報は、異なるタイプの末端アミノ酸を識別する特性パターンを決定するために使用可能であることが、これらの結果から示される。

【０３４２】

実施例６．標識付エキソペプチダーゼによる末端アミノ酸の同定及び切断

ペプチドの末端アミノ酸を同定することと、ペプチドから末端アミノ酸を切断することと、の両方が可能な単一試薬の可能性を調べるために、試験を実施した。単一試薬として、エキソペプチダーゼは、末端残基に対する切断活性を保持しつつペプチドに結合可能でなければならない。それゆえ、異なるエキソペプチダーゼの天然の表面露出アミノ酸に的を絞ることにより、伝統的標識化ストラテジーを利用する初期のアプローチを行った。これらの実験では、エキソペプチダーゼ標識化のためのロバストな方法であることが証明された蛍光色素で、表面露出システイン（－ＳＨ）又はリシン（－NH_2）残基を標識した。しかしながら、ある特定の場合には、このアプローチは、１つ以上の色素で標識されるタンパク質の不均一集団を生成した。

【０３４３】

エキソペプチダーゼ上のどこで標識化を行うかをより精密に制御して各エキソペプチダーゼ分子が単一蛍光色素で確実に標識されるようにするために（さらには色素のオフターゲット反応性を排除するために）、新しい標識化ストラテジーを検討した。これらの実験では、反応性官能基を含有する非天然アミノ酸がエキソペプチダーゼに導入される部位特異的標識化ストラテジーを用いて、標識付エキソペプチダーゼを調製した（たとえば、チン，Ｊ．Ｗ．（Ｃｈｉｎ，Ｊ．Ｗ．）ら著、ジャーナル・オブ・アメリカン・ケミカル・ソサイエティー（Ｊ　Ａｍ　Ｃｈｅｍ　Ｓｏｃ．）、２００２年８月７日、第１２４巻、第３１号、ｐ．９０２６～９０２７を参照されたい）。

【０３４４】

２８７位のリシン残基からｐ－アジドフェニルアラニン（ｐＡｚＦ）側鎖を有する残基への突然変異により、エルシニア・ペスティス（Ｙｅｒｓｉｎｉａ　ｐｅｓｔｉｓ）（ｙＰＩＰ）由来のプロリンイミノペプチダーゼを修飾した。図２１Ａは、スティックとして示されるＫ２８７側鎖で示されたｐＡｚＦの化学構造により示唆される突然変異を有するｙＰＩＰの結晶構造を示す。この突然変異部位は、この位置のαヘリックスにより提供される安定性に基づいて且つ新しいアジド官能基が確実に溶媒露出されるように選択される。

【０３４５】

突然変異体アミノｔＲＮＡシンテターゼとアミノ酸鎖中にｐＡｚＦを取り込むのに必要な突然変異体ｔＲＮＡとを含有するｐＥＶＯＬプラスミドを得た。次いで、クイックチェンジ（ＱｕｉｃｋＣｈａｎｇｅ）ＩＩ突然変異誘発キットを用いて、ｐＡｚＦの特異的組込みに必要なアンバー停止コドン（ＴＡＧ）をｃＤＮＡに導入した。次いで、ｃＤＮＡをシーケンスしてＴＡＧコドン位置を確認した。この後、ｙＰＩＰアンバー突然変異体を含

有するｐＥＴ２１ｂ＋プラスミドと、ｔＲＮＡがアンバーコドンに対してｐＡｚＦを仕込む細胞機構を含有するｐＥＶＯＬプラスミドと、の両方の共トランスフェクションを行った。次いで、共トランスフェクトされた細胞を０．８ＯＤＵまで成長させ、２ＬのＬＢ中の２ｍＭ　ｐＡｚＦの存在下で０．０２％アラビノース及び１ｍＭ　ＩＰＴＧにより誘導し、化学分解を用いて採取した。５ｍＬアフィニティークロマトグラフィーカラムを用いて精製し、１００ｍＭイミダゾールでタンパク質を溶出した。次いで、得られるタンパク質を５０ｍＭ　ＨＥＰＥＳ　ｐＨ８．０及び０．５Ｍ　ＫＣｌで透析して濃縮し、アリコートに分け、フラッシュ凍結し、その後、－２０℃で貯蔵した。

【０３４６】

精製タンパク質中のアジド基の存在を確認するために、ＤＢＣＯ－Ｃｙ３（２ｍＭ）とｐＡｚＦ－ｙＰＩＰ変異体（２２０μＭ）とを反応させた（反応条件：５０ｍＭ　ＨＥＰＥＳ　ｐＨ８．０、０．５ｍＭ　ＫＣｌ、及び２０％ＤＭＳＯ、３７℃で１０時間、室温で４８時間）。サイズ排除クロマトグラフィーによりタンパク質反応産物を精製したところ、得られたタンパク質はアジド反応性ＤＢＣＯ－Ｃｙ３試薬で１００％標識されたと判断されたことから（図２１Ｂ）、非天然アミノ酸のロバストな取込みが示唆される。

【０３４７】

非標識及び標識付ｐＡｚＦ変異体のＳＤＳ－ＰＡＧＥ解析により、最終産物のタンパク質標識化及び純度を確認した。図２１Ｃは、ｐＡｚＦ－ｙＰＩＰのＣｙ３標識化を確認するＳＤＳ－ＰＡＧＥゲルの写真を示す（ラダーを示すためにゲルのオーバー露出画像が図２１Ｄに示される）。図２１Ｅは、色素及びタンパク質の両方がコマイグレートして（ｃｏ－ｍｉｇｒａｔｅ）純粋であることを確認するクーマシー染色ゲルの写真を示す。

【０３４８】

標識化及び精製の後に酵素が依然として活性であることを確認するために、色素標識付ｐＡｚＦ－ｙＰＩＰ変異体を活性アッセイに使用した。図２１Ｆに示されるように、Ｃｙ３－ｐＡｚＦ－ｙＰＩＰは、ＨＰＬＣにより測定したとき、１０００倍過剰基質を用いて１時間でペプチド基質を１００％加水分解可能であった。これらの実験は、天然タンパク質構造／機能の撹乱を最小限に抑えてエキソペプチダーゼの部位特異的修飾及び標識化を可能にする方法を実証する。

【０３４９】

実施例７．ポリペプチドシーケンシングでの修飾アミノ酸の認識

特異的翻訳後修飾を含有するアミノ酸の認識を評価するために、実験を実施した。リン酸化チロシン残基に対する潜在的認識分子として、ＦｙｎチロシンキナーゼのＳｒｃ相同性２（ＳＨ２）ドメインの三重突然変異型変異体（Ｔ８Ｖ、Ｓ１０Ａ、Ｋ１５Ｌ）をペプチドシーケンシングで試験した。変異体タンパク質をサンプルウェルのボトムに固定し、Ｎ末端ホスホチロシンを含有する蛍光標識付ペプチドを添加し、単一分子シグナルトレースを収集した。図２２Ａの代表的トレースより示されるように、これらの実験時、固定タンパク質によるペプチド結合が検出された。これらの実験時に収集されたパルス持続時間データは、図２２Ｂに示される（トップ、ミドル、及びボトムのプロットは、それぞれ、図２２Ａのトップ、中間、及びボトムのトレースに対応する）。パルス持続時間及びパルス間持続時間の統計は、図２２Ｃに示される（それぞれ、トップ及びボトムのパネル）。

【０３５０】

Ｆｙｎタンパク質がリン酸化チロシンに特異的であることを確認するために、対照実験を実施した。３つの異なるペプチド：Ｎ末端非修飾チロシンを含有する第１のペプチド（Ｙ、図２２Ｄ）、Ｎ末端及びペナルティメートの非修飾チロシンを含有する第２のペプチド（ＹＹ、図２２Ｅ）、及びＮ末端ホスホセリンを含有する第３のペプチド（図２２Ｆ）の各々に対して、実験を繰り返した。示されるように、陰性対照実験に使用されるペプチドのいずれを用いても、結合は検出されなかった。

【０３５１】

実施例８．ポリペプチドシーケンシングでのペナルティメートアミノ酸の認識

Ａ．ツメファシエンス（Ａ．ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ２－Ｖ１のパルス持続

時間に及ぼすペナルティメートアミノ酸の影響を決定するために、実験を実施した。Ｎ末端にユニークジペプチド配列を含有する４９の異なる蛍光標識ペプチドを調製した。ただし、Ｎ末端アミノ酸は、Ｆ、Ｗ、又はＹであり、且つペナルティメート位置は、２０種の天然アミノ酸の１つであった。各実験では、ＣｌｐＳ２－Ｖ１をサンプルウェルのボトムに固定し、蛍光標識付ペプチドの１つを添加して一分子シグナルトレースを１０～２０分間収集した。各ペプチドに対しては最小でも５０のサンプルウェルで、パルス持続時間データを収集した。

【０３５２】

図２３は、５０ペプチドの各々に対するパルス持続時間の中央値を示しており、データ点は、異なる記号を用いて表されたペナルティメートアミノ酸（ｘ軸）とＮ末端アミノ酸とによりグループ化されている。

【０３５３】

実施例９．複数の認識分子による同時アミノ酸認識

１つ超の標識付認識分子による固定ペプチドの末端アミノ酸認識を実証するために、単一分子ペプチド認識実験を実施した。Ｎ末端フェニルアラニンを含有する単一ペプチド分子（ＦＹＰＬＰＷＰＤＤＤＹ（配列番号２３５））をチップのサンプルウェル内に固定した。各５００ｎＭのａｔＣｌｐＳ１（アグロバクテリウム・ツミファシエンス（Ａｇｒｏｂａｃｔｅｒｉｕｍ　ｔｕｍｉｆａｃｉｅｎｓ）ＣｌｐＳ１、表１に提供される配列）及びａｔＣｌｐＳ２－Ｖ１（アグロバクテリウム・ツミファシエンス（Ａｇｒｏｂａｃｔｅｒｉｕｍ　ｔｕｍｉｆａｃｉｅｎｓ）ＣｌｐＳ２変異体１、表１に提供される配列）を含有する緩衝液を添加した。ただし、ａｔＣｌｐＳ１及びａｔＣｌｐＳ２－Ｖ１は、それぞれ、Ｃｙ３及びＣｙ３Ｂで標識された。Ｃｙ３Ｂの強度はＣｙ３よりも高いので、ａｔＣｌｐＳ２－Ｖ１結合イベントは、ａｔＣｌｐＳ１結合イベントから容易に区別可であった。

【０３５４】

図２４Ａ～２４Ｃは、差次的標識付認識分子による単一分子ペプチド認識を示す実験の結果を示す。代表的トレースは図２４Ａに提示される。パルス持続時間分布は、各バインダーに対して識別可能であり（図２４Ｂ）、単一バインダーの実験で観測されるそれらの速度論的プロファイルに対応した。平均パルス持続時間は、ａｔＣｌｐＳ１では１．３秒間及びａｔＣｌｐＳ２－Ｖ１では１．０秒間であった（図２４Ｃ）。パルス速度もまた識別可能であり、ａｔＣｌｐＳ１では８．１パルス／ｍｉｎ及びａｔＣｌｐ２－Ｖ１では１４．１パルス／ｍｉｎであった（図２４Ｃ）。そのため、１つ超の認識分子が固定ペプチドの動的認識のために含まれるとき、各認識分子の結合特性（パルス持続時間、パルス間持続時間、及びパルス速度を含む）は、ペプチド配列に関する情報を同時に提供可能である。

【０３５５】

実施例１０．認識分子リンカーによる光安定性の増強

単一分子シーケンシング時の固定ペプチドの光安定性を評価するために、実験を実施した。アトー（ＡＴＴＯ）５３２からの放出により認識をモニターするために、５３２ｎｍの励起光の存在下で固定ペプチド基質を含有するサンプルウェルに実施例４に記載の色素標識付ａｔＣｌｐＳ２－Ｖ１を添加した。代表的トレースは図２５Ａに示される。トップパネルに示されるように、認識は、実験を始めてからおおよそ６００秒間で中止されることが観測された。ボトムパネルは、反応を始めてからおおよそ１８０～４３０秒間のシグナルパルスを示す拡大図である。

【０３５６】

図２５Ｂは、これらの実験に使用されるＣｌｐＳ２タンパク質の結晶構造の可視化を示す。示されるように、色素コンジュゲーション部位として機能するシステイン残基は、末端アミノ酸結合部位からおおよそ２ｎｍである。ペプチドに対する光損傷は、結合時にペプチドのＮ末端への色素の近接により引き起こされるという仮説を立てた。色素近接の潜在的光損傷作用を軽減するために、色素とペプチドのＮ末端との間の距離を１０ｎｍ超に

増加させるリンカーを介してＣｌｐＳ２タンパク質を色素標識した。リンカーは、ストレプトアビジンと二本鎖核酸とを含んでいた。２つのＣｙ３Ｂ色素分子で二本鎖核酸を標識し、ビスビオチン部分を介してストレプトアビジンに装着し、ビオチン部分を介してストレプトアビジン上の残留する２つの結合部位の各々にＣｌｐＳ２タンパク質を装着した。この色素シールドＣｌｐＳ２分子を用いた代表的トレースは図２５Ｃに示される。トップパネルに示されるように、認識時間は、実験を始めてからおおよそ６，０００秒間に延長された。ボトムパネルは、反応を始めてからおおよそ７５０～９３０秒間のシグナルパルスを示す拡大図である。

【０３５７】

２つのＣｙ３Ｂ色素分子で標識され、且つビスビオチン部分を介してストレプトアビジンに装着された二本鎖核酸を含有するリンカーを用いてＤＮＡ－ストレプトアビジン認識分子を生成し、ビスビオチン部分を介してストレプトアビジン上の残留する２つの結合部位に単一ＣｌｐＳ２タンパク質を装着した。この構築物は、単一分子ペプチドシーケンシング反応に使用され、これらの実験の代表的トレースは、図２６Ａ～２６Ｄに示される。

【０３５８】

反応条件を次のように変更して、すなわち、ＤＮＡ－ストレプトアビジンＣｌｐＳ２認識分子をｈＴＥＴアミノ酸切断試薬との組合せで使用して、実施例５に記載のシーケンシング実験を繰り返した。代表的シグナルトレースは図２７に示される。

【０３５９】

実施例１１．複数のエキソペプチダーゼによる分解時の認識によるシーケンシング

単一分子ペプチドシーケンシング反応混合物中での差次的切断特異性を有する複数のタイプのエキソペプチダーゼの使用を評価するために、実験を実施した。チップのサンプルウェル内にＣ末端リシン残基を介して単一ペプチド分子（ＹＡＡＷＡＡＦＡＤＤＤＷＫ（配列番号２３４））を固定した。アミノ酸認識用のａｔＣｌｐＳ２－Ｖ１とアミノ酸切断用のｈＴＥＴとを含有する緩衝液を添加した。代表的トレースは図２８Ａに提示され、パルスパターン領域の拡大図が図２８Ｂに示される。

【０３６０】

差次的特異性を有する２タイプのエキソペプチダーゼの存在下でシーケンシング反応を評価するために、実験を行った。チップのサンプルウェル内にＣ末端リシン残基を介して単一ペプチド分子（ＦＹＰＬＰＷＰＤＤＤＹＫ（配列番号２３６））を固定した。アミノ酸認識用のａｔＣｌｐＳ２－Ｖ１とアミノ酸切断用のｙＰＩＰ及びｈＴＥＴの両方とを含有する緩衝液を添加した。代表的トレースは図２８Ｃに提示され、パルスパターン領域の拡大図が図２８Ｄに示される。これらの反応条件の追加の代表的トレースは図２８Ｅに示される。

【０３６１】

差次的特異性を有する２タイプのエキソペプチダーゼの存在下でシーケンシング反応を評価するために、さらなる実験を行った。チップのサンプルウェル内にＣ末端リシン残基を介して単一ペプチド分子（ＹＰＬＰＷＰＤＤＤＹＫ（配列番号２３７））を固定した。一実験では、アミノ酸認識用のａｔＣｌｐＳ２－Ｖ１とアミノ酸切断用のｙＰＩＰ及びｈＴＥＴの両方とを含有する緩衝液を添加した。代表的トレースは図２８Ｆに提示され、パルスパターン領域の拡大図が図２８Ｇに示される。これらの反応条件の追加の代表的トレースは図２８Ｈに示される。さらなる実験では、アミノ酸認識用のａｔＣｌｐＳ２－Ｖ１とアミノ酸切断用のＰｆｕＴＥＴ及びｙＰＩＰの両方とを含有する緩衝液（５０ｍＭ　ＭＯＰＳ、６０ｍＭ　ＫＯＡｃ、２００μＭ　Ｃｏ（ＯＡｃ）$_2$）を添加した。代表的トレースは図２８Ｉに提示され、パルスパターン領域の拡大図が図２８Ｊに示される。

【０３６２】

実施例１２．新規ＣｌｐＳタンパク質の同定及び評価

Ｎ末端フェニルアラニン（Ｆ）、トリプトファン（Ｗ）、チロシン（Ｙ）、及びロイシン（Ｌ）に結合するＣｌｐＳタンパク質は、報告されている。以前に未知であったＮ末端結合性を有する潜在的な新規ＣｌｐＳホモログを発見するために、このタンパク質ファミ

リーに存在する配列多様性の全てを包含するおおよそ６０種から特徴付けられていないＣｌｐＳタンパク質の高度に多様なパネルを、相同性解析に基づいて設計した。

【０３６３】

ＣｌｐＳタンパク質のパネルを、ハイスループット発現および精製ワークフローを使用してスクリーニングした。ＣｌｐＳタンパク質を、１００ｍＬスケールでＥ．コリ（Ｅ．ｃｏｌｉ）細胞において過剰発現させ、同時発現させたビオチンリガーゼによってｉｎｖｉｖｏでビオチン化し、細胞溶解によって放出させ、ストレプトアビジンと複合体化し、コバルトアフィニティークロマトグラフィーによって精製した。ＳＤＳ－ＰＡＧＥクロマトグラフィーによる解析は、＞８５％純粋なＣｌｐＳタンパク質を示した（図２９Ａ）。

【０３６４】

各新規ホモログのアミノ酸結合プロファイルを、バイオレイヤー干渉法によって評価した。既知の結合プロファイルを有するＣｌｐＳタンパク質（ａｔＣｌｐＳ２－Ｖ１、ａｔＣｌｐＳ１、及びｔｅＣｌｐＳ）を対照として含めた。これらのアッセイでは、目的のＮ末端アミノ酸を含有するペプチドをバイオレイヤー干渉センサー表面上に固定化し、ＣｌｐＳタンパク質に結合させ、緩衝液中でインキュベートしてＣｌｐＳ分子を解離させた。結合反応（ｎｍ）及び解離オフ速度（ｋｄｉｓ、s^{-1}）を、速度論モデルを用いて測定した。各ＣｌｐＳタンパク質の測定を、アミノ酸Ｉ、Ｖ、Ｍ、Ｆ、Ｙ、及びＷ、並びに陰性対照としてのＤ、Ｅ、Ａ、及びＲについて３０℃で実施した。全てのペプチドは、ペナルティメートアラニン（Ａ）を含有し、配列ＸＡＫＬＤＥＥＳＩＬＫＱＫ（配列番号２３８）からなっていた。

【０３６５】

６１個のＣｌｐＳホモログ（５８個の新規ホモログ、３個の参照ホモログ）のバイオレイヤー干渉法スクリーニングからの結果を図２９Ｂに示す。１４個のＣｌｐＳホモログについての選択結果を、図２９Ｃに示されるプロットにまとめる。反応（ｙ軸）は、８つの異なるジペプチドについて解離速度の逆数に対してプロットされている（陰性対照ＤＡ、ＡＡ、及びＲＡについてのデータは示されていない）。これらのプロットにおいて、ｙ軸上のより高い値は、より強い会合反応に対応し、ｘ軸上のより高い値は、より遅い解離速度を示す。スクリーニングされたＣｌｐＳホモログの中で、ホモログ「ＰＳ３７２」は、ロイシン、イソロイシン、及びバリンのバインダーとして際立っていた。ＬＡ、ＩＡ、及びＶＡを用いたＰＳ３７２についての反応曲線を図２９Ｄに示す。ＰＳ３７２は、Ｆ、Ｙ、又はＷペプチドへの結合を示さなかった。結果は、ＰＳ３７２が、シーケンシング反応のための検出可能なオン及びオフ結合を示すバインダーに典型的な速度論的パラメータでＩＡおよびＬＡに結合することを示した。

【０３６６】

ＣｌｐＳホモログを蛍光偏光によって評価して、ＦＩＴＣ標識付ペプチド（ＸＡＫＬＤＥＥＳＩＬＫＱＫ－ＦＩＴＣ（配列番号２３８））とＣｌｐＳ：ストレプトアビジン複合体との相互作用を測定した。Ｉ、Ｖ、Ｌ、Ｆ、Ｙ、及びＷについての偏光反応（ミリ偏光、ｍＰ）並びに全強度の測定を、ハイスループット３８４ウェルプレートフォーマットにおいて、４８０ｎｍ励起及び５３０ｎｍ発光波長を使用して実施し、２３℃で３０分間インキュベートした後に読み取り値を収集した。溶液中結合の結果は、ＰＳ３７２が、ロイシン、イソロイシン、及びバリンと、ＬＡ＞ＩＡ＞ＶＡの順の反応強度で結合することを確証した（図２９Ｅ）。ａｔＣｌｐＳ２－Ｖ１の結合反応はＦＡ＞ＹＡ＞ＷＡであり、ｔｅＣｌｐＳの結合反応はＬＡ＞ＦＡ＞ＹＡ＞ＷＡであった。

【０３６７】

ＰＳ３７２によるイソロイシン及びバリン結合に対するペナルティメートアミノ酸の効果を、バイオレイヤー干渉法によってＮ末端ＩＲ、ＩＱ、及びＶＲを有するペプチドへの結合を測定することによって評価した（ペプチド配列：ＸＸＫＬＤＥＥＳＩＬＫＱＫ（配列番号：２９１））。測定値は、比較のためにｔｅＣｌｐＳでも取得された。結果は、ＰＳ３７２が、ＩＲと比較してより高い親和性でＩＱ及びＩＡに結合することを示したが（

図２９Ｆ）、ｔｅＣｌｐＳは、ＩＲペプチドのみに対して非常に速い解離を伴う小さな反応を示した。ＰＳ３７２はＶＡとよく結合し、ＶＲとはごくわずかな反応しか示さず、ｔｅＣｌｐＳはＶペプチドとは反応を示さなかった。

【０３６８】

ＬＦ、ＬＡ、ＩＲ、ＩＡ、ＶＲ、及びＶＡペプチドを用いてＰＳ３７２についてバインダーオンチップ実験を行って、単一分子アッセイにおけるシグナルパルス検出による認識を評価した。ＰＳ３７２は、ＬＦ及びＬＡペプチドで長い平均パルス幅を示した（それぞれ１２０７ｍｓ及び８７６ｍｓ）。比較のために、Ｌ－バインダーｔｅＣｌｐＳで観測された平均パルス幅は７６８ｍｓであった（ＰＳ３７２よりも約３７％短い）。短いパルス持続時間でのパルス化が、ＩＲ（８３ｍｓ）及びＩＡ（７０ｍｓ）ペプチドの両方を有するＰＳ３７２について観測された。図２９Ｇは、ＩＲペプチド（トップ）及びＬＦペプチド（ボトム）を有するＰＳ３７２のパルス幅のヒストグラムを示し、５分ウィンドウ内の代表的トレースを、各ヒストグラムの右側に示す。

【０３６９】

これらの実験は、イソロイシン及びバリンに対して強い固有の親和性を示すことが知られている第１のＣｌｐＳタンパク質の同定および生化学的特徴付けを実証する。バイオレイヤー干渉法および蛍光偏光を使用する結合速度論の分析に基づいて、ＰＳ３７２は、Ｗ、Ｆ、Ｙ、及びＬの認識に限定される他の全ての公知のＣｌｐＳタンパク質とは異なり、Ｎ末端ロイシン、イソロイシン、及びバリンを認識する。さらに、単一分子レベルでのＩ及びＬとＰＳ３７２との観察可能な結合が、バインダーオンチップアッセイにおいて確認され、シグナルパルスデータは、シーケンシング反応に有利な性質を示した。

【０３７０】

実施例１３．新規ＣｌｐＳホモログの標識及びシーケンシングにおける使用

ＣｌｐＳホモログＰＳ３７２及びＢｉｒＡプラスミドの大規模（５Ｌ）バッチをＥ．コリ（Ｅ．ｃｏｌｉ）に形質転換し、０．４ｍＭ　ＩＰＴＧ及び１６０μＭビオチン（ｉｎｖｉｖｏビオチン化）で１４℃にて一晩発現させた。回収した細胞を５ｍＬのコバルトアフィニティークロマトグラフィーカラムで精製した。ビオチン化効率を、ＳＤＳ－ＰＡＧＥを使用して評価し（図３０Ａ）、６倍過剰のストレプトアビジン（ＳＶ）と混合した場合のＰＳ３７２バンドの完全なシフトに基づいて１００％であると決定した。

【０３７１】

ビオチン化ＰＳ３７２（ＰＳ３７２－Ｂｔ）を、透析を使用してＰＥＧ化のために適切な緩衝液に移した。次いで、タンパク質を、室温で２時間、ｍＰＥＧ４－ＮＨＳエステルを使用してＰＥＧ化し、続いて一晩透析して過剰のｍＰＥＧ４－ＮＨＳを除去した。ＰＥＧ化及びビオチン化ＰＳ３７２（ＰＳ３７２－Ｂｔ－ｍＰＥＧ４）を濃縮し、予め形成されたＳＶとビオチン化Ｃｙ３Ｂ標識付オリゴヌクレオチドとの１：１複合体とコンジュゲートさせ、ＨＰＬＣカラムで精製した。ＨＰＬＣプロファイルは、２つの主要なピークを示し（図３０Ｂ）、ピーク３が目的のピークであった（１：１（ＰＳ３７２：ＳＶ色素）複合体）。

【０３７２】

ＨＰＬＣ精製後、得られたピークを適切な濃度に濃縮し、チップ表面への非特異的結合に起因してオンチップアッセイにおいて干渉し得る標識付ＳＶ上の任意のフリーのビオチン結合部位について試験した。図３０Ｃに示すように、ＡｔｔｏＲｈｏ６で標識した遊離ビオチンを全ての試料に添加し、対照と共にＳＤＳ－ＰＡＧＥにかけた。Ｃｙ３チャネル（パネル（ａ）及び（ｃ））は、レーン６および７において１：１種（ＰＳ３７２：ＳＶ色素）の存在を示し、ＡｔｔｏＲｈｏ６の非存在（パネル（ｂ）及び（ｃ））は、標識付タンパク質上にフリーのビオチン結合部位がないことを確認した。

【０３７３】

標識付タンパク質をペプチドオンチップ認識アッセイで試験したところ、平均パルス幅７８ｍｓでＮ末端イソロイシンペプチド（ＩＡＡＬＡＡＶＡＡＤＤＤＷ（配列番号：２３９））に結合し、平均パルス幅９５７ｍｓでＮ末端ロイシンペプチド（ＬＡＡＩＡＡＦＡ

ＡＤＤＤＷ（配列番号：２９２））に結合することを示した。エキソペプチダーゼ（ｈＴＥＴ）の存在下での動的シーケンシングアッセイの結果は、Ｎ末端イソロイシンペプチドについてはＩからＬへ、Ｎ末端ロイシンペプチドについてはＬからＩへの明確な遷移を示した（図３０Ｄ）。

【０３７４】

進行中の分解反応時の色素標識付ＰＳ３７２でのＮ末端アミノ酸認識によるペプチドシーケンシングを評価するために実験を行った。リアルタイム動的ペプチドシーケンシングアッセイを、色素標識付ＰＳ３７２認識因子及びアミノペプチダーゼ切断試薬（ｈＴＥＴ）の存在において、表面固定化ポリペプチド（ＩＡＡＬＡＡＶＡＡＤＤＤＷ（配列番号：２３９））をモニタリングすることによって行った。リアルタイム動的ペプチドシーケンシングアッセイの例示データを図３０Ｅに示し、平均パルス幅を各パルスクラスターの下に列挙する。示されるように、パルス幅分布における差は、イソロイシン及びロイシン認識を容易に区別し、動的シーケンシング反応における色素標識付ＰＳ３７２によるこれらのアミノ酸の認識を実証した。

【０３７５】

実施例１４．新規なメチオニン結合ＣｌｐＳタンパク質の工学操作

アグロバクテリウム・ツメファシエンス（Ａｇｒｏｂａｃｔｅｒｉｕｍ　ｔｕｍｅｆａｃｉｅｎｓ）ＣｌｐＳ２の熱安定性変異体から出発して、１６０，０００個の変異体の潜在的多様性を有するＣｌｐＳ突然変異体のライブラリーを作製した（トゥルマン（Ｔｕｌｌｍａｎ）ら著、２０２０年、Ｂｉｏｃｈｅｍ．Ｅｎｊ．Ｊ．、第１５４巻、ｐ．１０７４３８を参照されたい）。タンパク質が酵母の表面上に提示されうるような突然変異ＣｌｐＳ遺伝子をＰＣＲ増幅するための縮重プライマーを使用して、酵母における相同組換えによってライブラリーを作製した。この方法を使用して、各タンパク質変異体は、酵母細胞の表面上に約１，０００コピーで提示され、クローンを単離することによって、提示されるタンパク質の遺伝子配列を決定しうる。これは、どのタンパク質変異体が目的の性質を有するかを決定するために必要な遺伝子型－表現型リンクを維持する。

【０３７６】

１００万を超えるクローンを得た後、変異体ライブラリーを提示する酵母と、配列ＭＲＦＶＧＥＣＫ－ビオチン（配列番号：２４０）を含有するペプチド、並びにフルオロフォアストレプトアビジン－ＰＥ及び抗ｍｙｃ　ＡｌｅｘａＦｌｕｏｒ６４７とを組み合わせることによって、蛍光活性化セルソーティング（ＦＡＣＳ）を介して選択を行った。ペプチドに結合し、且つＣｌｐＳタンパク質のＣ末端にｍｙｃタグを含有する細胞は、ＦＡＣＳによって得られたプロットの象限２（右上）に保持されると予想され、セルソーターは、これらを培養チューブに仕分けして別のラウンドの選択及び／又はシーケンシング用に増殖させ、捕捉された変異体を同定する。このライブラリーに対して２ラウンドのＦＡＣＳを実施したところ、図３１Ａに示されるプロットに見られる改善が得られた（左側プロット：０．５μＭのＭＲペプチドを用いた第１のラウンド、中央プロット：０．５μＭＭＲペプチドを用いた第２のラウンド、右側プロット：比較のために０．０５μＭのＭＲペプチドを用いた第２のラウンド、予想通りほとんど結合を示さなかった）。

【０３７７】

２ラウンドの選択後、酵母の単一クローンをプレーティングによって単離し、シーケンシングした。２つの候補が、「ＴＭＲ　Ｌ」（ＰＳ４９０）及び「ＴＡＦ　Ｋ」（ＰＳ４８９）として突然変異した４つの残基の配列を有して出現した。さらなる特徴付けの際に、酵母ディスプレイを使用して、タンパク質がＭＲ、ＭＱ、及びＩＲに結合するが、ＩＱ又はＥＧペプチドには結合しないことが示された。これらの２つのクローンを発現ベクターにサブクローニングして、バイオレイヤー干渉アッセイのための精製及びストレプトアビジン標識を容易にした。タンパク質を発現させ、精製し、アッセイして、元の熱安定性ａｔＣｌｐＳ２変異体（ＰＳ０２３）と共にさまざまなＮ末端アミノ酸との結合反応速度を測定した。図３１Ｂ～３１Ｆに示されるデータは、ＭＲペプチド、並びにＭＡ、ＬＡ、及びＦＡペプチドへのＰＳ４８９の結合を確認した。そのほか、ＰＳ４９０は、ＭＲ、Ｌ

Ａ、及びＦＡに結合する。

【０３７８】

実施例１５．設計された突然変異変異体ＣｌｐＳ認識分子

単一分子シーケンシング反応において有利でありうるより長いパルス幅を達成するように合理的に設計された活性部位置換を有するａｔＣｌｐＳ１の変異体を作製した。設計された変異体（ＰＳ２１８）は、野生型ａｔＣｌｐＳ１に対して以下の突然変異：Ｍ５１Ｆ、Ｅ５２Ｑ、Ｍ７３Ｔ、Ｔ７０Ｗを含有する。ＦＡ、ＬＦ、ＷＡ、及びＹＡペプチドを用いてＰＳ２１８についてバインダーオンチップ実験を行って、単一分子アッセイにおけるシグナルパルス検出による認識を評価した。表７の結果によって示されるように、野生型タンパク質に対する置換は、試験した各ペプチドでより長いパルス幅をもたらす。

【０３７９】

【表７－１】

表７．PS218及び野生型 atClpS1についてのバインダーオンチップ結果

バインダー	ペプチド	ウエルの数	平均パルス幅	PWの中央値
PS218	FA	212	1.28	0.78
PS218	LF	265	0.56	0.32
PS218	WA	180	1.00	0.58
PS218	YA	185	1.64	1.06
atClpS1 (WT)	FA	387	0.66	0.36

【０３８０】

【表７－２】

atClpS1 (WT)	LF	291	0.30	0.18
atClpS1 (WT)	WA	240	0.39	0.24
atClpS1 (WT)	YA	404	0.77	0.40

【０３８１】

実施例１６．シーケンシングに使用するためのＵＢＲボックスドメインホモログの評価

本発明者らは、酵母ＵＢＲ由来のＵＢＲボックスドメインが、Ｌ又はＩが後に続く場合にのみＲに対して強い結合親和性を示すことを見出した。単一分子ペプチドシーケンシングにおいて有用であるように、ペナルティメートアミノ酸の存在下でより広い範囲の高親和性Ｒ結合を有するＵＢＲボックスドメインを同定するために、ＵＢＲボックスドメインホモログのスクリーニングを行った。ＵＢＲボックスドメインホモログのアミノ酸結合性をバイオレイヤー干渉法によって評価した。Ｎ末端アルギニン（Ｒ）、続いてペナルティメート位置にさまざまなアミノ酸を含むペプチドをバイオレイヤー干渉センサー表面上に固定し、ＵＢＲボックスタンパク質に結合させ、緩衝液中でインキュベートしてＵＢＲボックス分子を解離させた。ペナルティメート位置のさまざまなアミノ酸の前にＮ末端Ｒを含有する１４のペプチダーゼと結合するＵＢＲボックスホモログＰＳ５３５（図３２Ａ）及びＰＳ５２２（図３２Ｂ）についての例示反応トラジェクトリーが示される。図３２Ｃは、ペナルティメート位置の様々なアミノ酸の前にＮ末端Ｒを含有する１４のポリペプチドに対する２４のＵＢＲボックスホモログを測定した結果を示すヒートマップである。結果は、ホモログＰＳ５３５、ＰＳ５２２、ＰＳ５２８、及びＰＳ５０５が、広範囲のペナルティメートアミノ酸の存在下でＲに対する高い結合親和性を維持することを実証する。

【０３８２】

ＵＢＲタンパク質は、保存されたＮ末端則経路の一部として塩基性Ｎ末端残基Ｒ、Ｋ、

及びＨに結合する。シーケンシングにおける使用に好ましい性質を有する高親和性ＵＢＲタンパク質を同定するために、４４種（ＰＳ５０１～ＰＳ５４４）由来の特徴付けられていないＵＢＲボックスドメインホモログの多様なパネルを設計した。ＵＢＲボックスドメインホモログを発現させ、精製し、バイオレイヤー干渉法によってＮ末端アミノ酸への結合について評価した。Ｒ、Ｋ及びＨ結合について、各タンパク質について３０℃で測定を行った。各新規ホモログのアミノ酸結合プロファイルを、まず、単一点スクリーニングアッセイを使用するバイオレイヤー干渉法によって評価した。このスクリーニングにおける全てのペプチドは、ＸがＲ、Ｋ、又はＨである場合、ペナルティメートアラニン（Ａ）を含有し、配列ＸＡＫＬＤＥＥＳＩＬＫＱＫ（配列番号２３８）からなっていた。

【０３８３】

ＲＡ／ＫＡ／ＨＡペプチド結合についてのバイオレイヤー干渉法反応測定を図３２Ｄにまとめる。ホモログＰＳ５２２、ＰＳ５２８、ＰＳ５３５、及びＰＳ５３８は、ＲＡペプチド及びＫＡペプチドに対する好ましい結合トラジェクトリーを示し、ＲＡペプチドに対して顕著に高い反応を示した。ＰＳ５２８は、より遅い解離トラジェクトリーを示し、ＨＡペプチドへの結合を示した。これらの結果に基づいて、１５のＵＢＲボックスドメインホモログ（ＰＳ６１４～ＰＳ６２９）の第２のパネルを、Ｋ．ラクティス（Ｋ．ｌａｃｔｉｓ）に関連するホモログの拡大されたセットを含めて設計した。ＰＳ６２１は好ましい結合特性を示し、結果はＰＳ５２８と同様であった（図３２Ｄ）。

【０３８４】

アルギニン結合に対するペナルティメートアミノ酸の影響を評価するために、ＵＢＲボックスドメインホモログを、Ｎ末端ＲＸを有する１４の異なるペプチド（ペプチド配列：ＲＸＧＧＧＤＤＤＦＦＫ（配列番号：２４１））への結合について評価した。選択されたＲＸジペプチドは、ヒトプロテオームにおいて最も頻繁に見出される１４種である。４０のＵＢＲボックスタンパク質についてのバイオレイヤー干渉反応測定のフルセットを図３２Ｅにまとめる。結果は、ＰＳ５３５、ＰＳ５２２、ＰＳ５２８、及びＰＳ６２１がＲＸジペプチドに対して広範な結合範囲を有することを示した。哺乳動物起源のＵＢＲ候補ＰＳ５２２及びＰＳ５３５は、より速い解離速度を有するＲジペプチドについて広範な認識を示した（図示せず）。Ｋｌｕｖｅｒｏｍｙｃｅｓ酵母ファミリー由来のホモログＰＳ５２８及びＰＳ６２１もまた、Ｒジペプチドの広範な認識を示した。

【０３８５】

次いで、選択された候補について単一点蛍光偏光アッセイを実施した。これらのアッセイでは、標識付ペプチド（ＸＡＫＬＤＥＥＳＩＬＫＱＫ－ＦＩＴＣ（配列番号：２３８））とＵＢＲストレプトアビジン複合体との相互作用を、これらの分子が溶液中で遊離している形式で測定した。ＲＡ、ＫＡ、及びＨＡペプチドについての偏光反応（ミリ偏光、ｍＰ）および総強度の測定を行った。結果は、ＰＳ５２８及びＰＳ６２１が強い反応でＲＡペプチドに結合することを示した（図３２Ｆ）。また、反応トラジェクトリーは、ＰＳ５２８及びＰＳ６２１が、強いＲＡバインダー（図示せず）であり、動的シーケンシング反応において検出可能なオン及びオフ結合を示すバインダーに典型的な速度論的パラメータで結合することを示した。

【０３８６】

次いで、選択されたＵＢＲボックスドメイン候補の結合親和性を、Ｎ末端Ｒ及びＫペプチドについてのK_D（解離定数）値を測定することによって評価した。ＲＡ及びＫＡペプチドについて、漸増濃度のＵＢＲタンパク質でアッセイを行った（図３２Ｇ、トッププロット）。結合滴定曲線からK_D値を決定した。図３２Ｇ（ボトムの棒グラフ）に示すように、ＲＡ及びＫＡペプチドについて結合が観測され、ＰＳ６２１及びＰＳ５２８は、ＲＡペプチドに対して比較的高い結合親和性を示した（それぞれ、４２０ｎＭ及び４６０ｎＭのK_D値）。

【０３８７】

Ｎ末端アルギニンペプチドについてのペプチドオンチップ認識アッセイを、ＰＳ６２１及びＰＳ５２８を用いて行った。ＰＳ５２８及びＰＳ６２１の両方は、Ｎ末端Ｒを認識し

、単一分子レベルでの色素標識付ＲＬペプチドへの結合が、約６０ｍｓの平均パルス幅で、これらの認識アッセイにおいて確認された。色素標識付ＲＬペプチドを用いたペプチドオンチップ認識アッセイにおけるＰＳ６２１の代表的トレースを図３２Ｈに示す。

【０３８８】

また、ＵＢＲタンパク質ＰＳ６２１又はＰＳ５２８をａｔＣｌｐＳ２－Ｖ１及びＰＳ５５７と組み合わせた３バインダー系を使用して、動的シーケンシングを行った。これらのセットのバインダーは、ヒトユビキチン由来の天然ペプチド断片（配列：ＤＱＱＲＬＩＦＡＧＫ（配列番号：２４２））のＲ、Ｌ、Ｉ、及びＦを認識することが示された。ＰＳ５２８、ａｔＣｌｐＳ２－Ｖ１、及びＰＳ５５７を使用する３－バインダー動的シーケンシングランからのこのペプチドについての例示的なシーケンシングトレースを図３２Ｉに示す。

【０３８９】

実施例１７．シーケンシングにおける使用のためのＰＳ３７２ホモログの同定及び評価

単一分子ペプチドシーケンシングにおいてイソロイシン及びバリン認識が可能なＣｌｐＳタンパク質を同定するために、ＰＳ３７２ホモログタンパク質のスクリーニングを行った。ＰＳ３７２ホモログのアミノ酸結合性を、実施例１２に記載されるようなバイオレイヤー干渉法によって評価した。例示の反応トラジェクトリーは、ペナルティメート位置のアラニンの前にさまざまなＮ末端アミノ酸（Ｉ、Ｖ、Ｌ、Ｆ）を含有する４つのポリペプチドと結合するＰＳ３７２（図３３Ａ）及びホモログＰＳ５４５（図３３Ｂ）、ＰＳ５５１（図３３Ｃ）、ＰＳ５５７（図３３Ｄ）、並びにＰＳ５５８（図３３Ｅ）について示される。図３３Ｆは、３４のＰＳ３７２ホモログを測定した結果を示すヒートマップである。

【０３９０】

実施例１８．多価アミノ酸バインダーの工学操作

フレキシブルペプチドリンカーをコードするセグメントによって末端同士で連結されたａｔＣｌｐＳ２－Ｖ１の２つのコピーをコードするセグメントを含有する単一コーディング配列から発現される直列融合分子として、多価アミノ酸バインダーを設計した。単一のコーディング配列の発現により、直列に配向された２つのＣｌｐＳタンパク質を有する単一の全長ポリペプチド（ビスａｔＣｌｐＳ２－Ｖ１）が生成された。それぞれ異なるリンカーを有する３つのビスａｔＣｌｐＳ２－Ｖ１バインダーを設計し、発現させた：ＰＳ６０９（「リンカー１」配列：ＧＧＧＳＧＧＧＳＧＧＧＳＧ（配列番号：２４３））、ＰＳ６１０（「リンカー２」配列：ＧＳＡＧＳＡＡＧＳＧＥＦ（配列番号：２４４））、及びＰＳ６１１（「リンカー３」配列：ＧＳＡＧＳＡＡＧＳＧＥＦＧＳＡＧＳＡＡＧＳＧＥＦＧＳＡＧＳＡＡＧＳＧＥＦ（配列番号：２４５））。

【０３９１】

ＰＳ６１０ポリペプチドをビオチン化し、上記のように色素標識付ストレプトアビジンとコンジュゲートさせた。色素標識付ＰＳ６１０をペプチドオンチップ認識アッセイにおいて試験し、Ｎ末端フェニルアラニンペプチド（ＦＡＡＡＹＰ（配列番号：２４６））への結合を示した。比較目的のために、１価バインダー（色素標識付ａｔＣｌｐＳ２－Ｖ１）及び直列バインダーＰＳ６１０をそれぞれ、同じバインダー濃度（５００ｎＭ）を用いて認識アッセイにおいて試験した。これらの実験からの代表的な単一分子パルストラジェクトリー及びパルス統計を図３４Ａに示す（左側パネル：１価バインダー、右側パネル：ＰＳ６１０）。示されるように、直列バインダーを使用して、パルス間持続時間の中央値は約５．８倍短く、パルス化の観測された速度は２．３倍速かったが、さまざまなバインダーを使用したパルス幅の中央値は、一貫していた。

【０３９２】

固定化インスリンＢ鎖（配列：ＦＶＮＱＨＬＣＧＳＨＬＶＥＡＬＹＬＶＣＧＥＲＧＦＦＹＴＰＫＡ（配列番号：２４７））のＮ末端の動的認識を、色素標識付ａｔＣｌｐＳ２－Ｖ１及びＰＳ６１０について、さまざまなバインダー濃度で独立して調べた。単一分子パルストラジェクトリーを２０分間取得した。これらの実験からの結果（表８に示す）を使

用して、バインダー濃度の関数としての平均パルス速度のプロットを生成した（図３４Ｂ）。示されるように、直列バインダーＰＳ６１０は、各濃度でより高い平均オンオフ結合速度（パルス速度）を示した。

【０３９３】

【表８】

表８．多価バインダーについての動的認識結果。

バインダー	トレース数	バインダー濃度 (nM)	パルス速度 (/min)
PS610	1696	100	6
PS610	519	50	4.8
PS610	594	100	8.4
PS610	1625	250	6.6
atClpS2V1	461	100	3.6
atClpS2V1	105	50	2.4
atClpS2V1	399	100	4.2
atClpS2V1	2314	250	3.6

【０３９４】

進行中の分解反応時の色素標識付ＰＳ６１０でのＮ末端アミノ酸認識によるペプチドシーケンシングを評価するために実験を行った。リアルタイム動的ペプチドシーケンシングアッセイを、色素標識付ＰＳ６１０（１００ｎＭ）、色素標識付ＰＳ３７２（２５０ｎＭ）、及びアミノペプチダーゼ切断試薬（ｈＴＥＴ及びｐｆｕＴＥＴ）の存在下において、グルカゴン様ペプチド１の表面固定化ペプチド断片（配列：ＥＦＩＡＷＬＶＫ（配列番号：２４８））をモニタリングすることによって行った。ＰＳ３７２を識別可能な色素で標識してＩ及びＬ認識のために提供し、固定化ペプチドのＮ末端アミノ酸の繰返し切断をアミノペプチダーゼによって実施した。リアルタイムシーケンシングアッセイからの例示トレースは、図３４Ｃに示される。ＰＳ６１０による特徴的なＦ及びＷ認識に対応するパルスの領域は、自動解析ワークフローを使用して正確に同定された。

【０３９５】

バインダーオンチップ実験は、ＰＳ６１０、又は比較目的のために１価バインダーについて、Ｎ末端ＦＡを含有する色素標識ペプチドで実施された。図３４Ｄには、１価バインダー（トップトレース）について及び直列バインダーＰＳ６１０（ボトムトレース）についての代表的トレースが示される。示されるように、ａｔＣｌｐＳ２の直列コピーからなる固定化ＰＳ６１０は、自由に拡散する色素標識付ペプチドへの結合におけるパルスオーバーパルスパターンを示す。これらの結果は、２つの連結されたＰＳ６１０のモノマーの各々が、独立して同時に、それらの標的ペプチドに結合可能であることを実証する。

【０３９６】

各ビスａｔＣｌｐＳ２－Ｖ１バインダー（ＰＳ６０９、ＰＳ６１０、及びＰＳ６１１）のアミノ酸結合プロファイルを、上述のようにバイオレイヤー干渉法によって評価した。１価ａｔＣｌｐＳ－Ｖ１を対照として行った。各バインダーの測定を、アミノ酸Ｆ、Ｙ、及びＷ、並びに陰性対照としてのＩ、Ｌ、Ｍ、及びＶについて実施した。全てのペプチドは、ペナルティメートアラニン（Ａ）を含有していた。１価バインダーについての反応曲線を図３５Ａに示す。直列バインダーＰＳ６０９、ＰＳ６１０、及びＰＳ６１１についての反応曲線を、それぞれ図３５Ｂ、３５Ｃ、及び３５Ｄに示す。

【０３９７】

多価バインダーＰＳ６５１を、フレキシブルペプチドリンカー（リンカー２）をコードするセグメントによって末端同士で連結されるａｔＣｌｐＳ－Ｖ１の３つのコピー（Ｔｒｉｓ－ａｔＣｌｐＳ－Ｖ１）をコードするセグメントを含む単一コーディング配列から発現する直列融合分子として設計した。蛍光偏光研究を、ＰＳ６５１、ビスＣｌｐＳ２－Ｖ

２バインダー（ＰＳ６０９、ＰＳ６１０、ＰＳ６１１）、及びモノマー形バインダーａｔＣｌｐＳ２－Ｖ１で実施して、ペナルティメート位置のアラニン残基とともにＮ末端フェニルアラニン残基を有するペプチドについて結合親和性を決定した。これらの研究から得られたＫ$_D$値を、以下の表９に報告する。これらの結果は、さまざまなリンカーを含む直列ａｔＣｌｐＳ２－Ｖ１構築物がＦ、Ｙ、及びＷへの結合を保持することを実証する。

【０３９８】

【表９】

表９．N末端フェニルアラニンペプチドに対する直列認識の結合親和性。

バインダー	FA ペプチド (K_D, nM)
atClpS2-V1	185
PS609（ビスatClpS2-V1、リンカー１）	221
PS610（ビスatClpS2-V1、リンカー２）	271
PS611（ビスatClpS2-V1、リンカー３）	137
PS651（トリスatClpS2-V1）	256

【０３９９】

追加の多価アミノ酸バインダーを、Ｎ末端認識について設計し、評価した。表１０は、設計されたさまざまな直列バインダー及び対応するポリペプチド配列のリストを含む。各発現構築物は、表３のＣ末端Ｈｉｓ／ビスビオチン化タグを含む。

【０４００】

【表１０－１】

表１０．多価バインダーの非限定的例。

名称	配列番号	配列
PS609（ビスatClpS2-V1、リンカー１）	249	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGGGSGGGSGGGSGMSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMS FVTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLG KEAGFPLMFTTEPEEGHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEW HEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS610（ビスatClpS2-V1、リンカー２）	250	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSF VTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGK EAGFPLMFTTEPEEGHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWH EGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS611	251	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE

【０４０１】

【表１０－２】

(ビスatClpS2-V1、リンカー3)		EGSAGSAAGSGEFGSAGSAAGSGEFGSAGSAAGSGEFMSDSPVDLKPKPKV KPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSEDTGRRVMMTAHRFG SAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPEEGHHHHHHHHHGG GSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWH E
PS612 (atClpS2-V1+ PS372、リンカー１)	252	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIE MLQKIFGFPPEKGFQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGPDPY LGRPCSGSMTCVIEPAVGGSHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQ KIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS613 (ビスPS372、リンカー２)	253	MAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIEMLQKIFGFPPEKG FQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGPDPYLGRPCSGSMTCVI EPAVGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTP MSFVTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATD LGKEAGFPLMFTTEPEEGHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKI EWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS614 (PS372 + atClpS2-V1、リンカー２)	254	MAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIEMLQKIFGFPPEKG FQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGPDPYLGRPCSGSMTCVI EPAVGSAGSAAGSGEFMAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRY VIEMLQKIFGFPPEKGFQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGP DPYLGRPCSGSMTCVIEPAVGGSHHHHHHHHHHGGGSGGGSGGGSGLNDFF EAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS637 (ビスPS557、リンカー１)	255	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGGSGGGSGGGSGGGSGMPTAASATESAIEDTPAPARP EVDGRTKPKRQPRYHVVLWNDDDHTYQYVVVMLQSLFGHPPERGYRLAKEV DTQGRVIVLTTTREHAELKRDQIHAFGYDRLLARSKGSMKASIEAEEGGSH HHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLN DFFEAQKIEWHE
PS638 (ビスPS557、リンカー２)	256	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMPTAASATESAIEDTPAPARPE VDGRTKPKRQPRYHVVLWNDDDHTYQYVVVMLQSLFGHPPERGYRLAKEVD TQGRVIVLTTTREHAELKRDQIHAFGYDRLLARSKGSMKASIEAEEGGSHH HHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLND FFEAQKIEWHE
PS639 (ビスPS557、リンカー３)	257	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFGSAGSAAGSGEFGSAGSAAGSG EFMPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQY VVVMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGY DRLLARSKGSMKASIEAEEGGSHHHHHHHHHHHGGGSGGGSGGGSGLNDFFE AQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS640 (atClpS2-V1 + PS557、リンカー2)	258	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVV LWNDDDHTYQYVVVMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAE LKRDQIHAFGYDRLLARSKGSMKASIEAEEGGSHHHHHHHHHHGGGSGGGS GGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS641 (PS557 + atClpS2-V1、リンカー2)	259	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPK LYKVMLLNDDYTPMSFVTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCE RDIAETKAKEATDLGKEAGFPLMFTTEPEEGGSHHHHHHHHHHGGGSGGGS GGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE

【０４０２】

【表１０－３】

PS651 (3×atClpS2-V1、リンカー2)	260	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSF VTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGK EAGFPLMFTTEPEEGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPKLYK VMLLNDDYTPMSFVTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDI AETKAKEATDLGKEAGFPLMFTTEPEEGHHHHHHHHHHGGGSGGGSGGGSG LNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS652 (4×atClpS2-V1、リンカー2)	261	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSF VTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGK EAGFPLMFTTEPEEGSAGSAAGSGEFMSDSPVDLKPKPKVKPKLERPKLYK VMLLNDDYTPMSFVTVVLKAVFRMSEDTGRRVMMTAHRFGSAVVVVCERDI AETKAKEATDLGKEAGFPLMFTTEPEEGSAGSAAGSGEFMSDSPVDLKPKP KVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSEDTGRRVMMTAHR FGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPEEGHHHHHHHHHH GGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIE WHE
PS653 (3×PS372、リンカー2) リンカー2	262	MAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIEMLQKIFGFPPEKG FQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGPDPYLGRPCSGSMTCVI EPAVGSAGSAAGSGEFMAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRY VIEMLQKIFGFPPEKGFQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGP DPYLGRPCSGSMTCVIEPAVGSAGSAAGSGEFMAFPARGKTAPKNEVRRQP PYNVILLNDDDHTYRYVIEMLQKIFGFPPEKGFQIAEEVDRTGRVILLTTS KEHAELKQDQVHSYGPDPYLGRPCSGSMTCVIEPAVGGSHHHHHHHHHHGG GSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWH E
PS654 (4×PS372, リンker 2)	263	MAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIEMLQKIFGFPPEKG FQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGPDPYLGRPCSGSMTCVI EPAVGSAGSAAGSGEFMAFPARGKTAPKNEVRRQPPYNVILLNDDDHTYRY VIEMLQKIFGFPPEKGFQIAEEVDRTGRVILLTTSKEHAELKQDQVHSYGP DPYLGRPCSGSMTCVIEPAVGSAGSAAGSGEFMAFPARGKTAPKNEVRRQP PYNVILLNDDDHTYRYVIEMLQKIFGFPPEKGFQIAEEVDRTGRVILLTTS KEHAELKQDQVHSYGPDPYLGRPCSGSMTCVIEPAVGSAGSAAGSGEFMAF PARGKTAPKNEVRRQPPYNVILLNDDDHTYRYVIEMLQKIFGFPPEKGFQI AEEVDRTGRVILLTTSKEHAELKQDQVHSYGPDPYLGRPCSGSMTCVIEPA VGGSHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGG GSGLNDFFEAQKIEWHE
PS655 (3×PS557, リンker 2)	264	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMPTAASATESAIEDTPAPARPE VDGRTKPKRQPRYHVVLWNDDDHTYQYVVVMLQSLFGHPPERGYRLAKEVD TQGRVIVLTTTREHAELKRDQIHAFGYDRLLARSKGSMKASIEAEEGSAGS AAGSGEFMPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDD HTYQYVVVMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQI HAFGYDRLLARSKGSMKASIEAEEGGSHHHHHHHHHHGGGSGGGSGGGSGL NDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS656 (4×PS557, リンker 2)	265	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMPTAASATESAIEDTPAPARPE VDGRTKPKRQPRYHVVLWNDDDHTYQYVVVMLQSLFGHPPERGYRLAKEVD TQGRVIVLTTTREHAELKRDQIHAFGYDRLLARSKGSMKASIEAEEGSAGS AAGSGEFMPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDD HTYQYVVVMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQI HAFGYDRLLARSKGSMKASIEAEEGSAGSAAGSGEFMPTAASATESAIEDT PAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVVVMLQSLFGHPPERGY

【０４０３】

【表１０－４】

		RLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDRLLARSKGSMKASIEA EEGGSHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSG GGSGLNDFFEAQKIEWHE
PS690 (ビスPS621、 リンカー１)	266	MHSKFSHAGRICGAKFKVGEPIYRCKECSFDDTCVLCVNCFNPKDHTGHHV YTTICTEFNNGICDCGDKEAWNHTLFCKAEEGGGGSGGGSGGGGSGMHSKFS HAGRICGAKFKVGEPIYRCKECSFDDTCVLCVNCFNPKDHTGHHVYTTICT EFNNGICDCGDKEAWNHTLFCKAEEGGGSHHHHHHHHHHGGGSGGGSGGGS GLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS691 (ビスPS621、 リンカー２)	267	MHSKFSHAGRICGAKFKVGEPIYRCKECSFDDTCVLCVNCFNPKDHTGHHV YTTICTEFNNGICDCGDKEAWNHTLFCKAEEGGSAGSAAGSGEFMHSKFSH AGRICGAKFKVGEPIYRCKECSFDDTCVLCVNCFNPKDHTGHHVYTTICTE FNNGICDCGDKEAWNHTLFCKAEEGGGSHHHHHHHHHHGGGSGGGSGGGSG LNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS692 (ビスPS621、 リンカー３)	268	MHSKFSHAGRICGAKFKVGEPIYRCKECSFDDTCVLCVNCFNPKDHTGHHV YTTICTEFNNGICDCGDKEAWNHTLFCKAEEGGSAGSAAGSGEFGSAGSAA GSGEFGSAGSAAGSGEFMHSKFSHAGRICGAKFKVGEPIYRCKECSFDDTC VLCVNCFNPKDHTGHHVYTTICTEFNNGICDCGDKEAWNHTLFCKAEEGGG SHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSG LNDFFEAQKIEWHE
PS693 (ビスPS528、 リンカー１)	269	MHSKFNHAGRICGAKFRVGEPIYRCKECSFDDTCVLCVNCFNPKDHVGHHV YTSICTEFNNGICDCGDKEAWNHELNCKGAEDGGGSGGGSGGGSGMHSKFN HAGRICGAKFRVGEPIYRCKECSFDDTCVLCVNCFNPKDHVGHHVYTSICT EFNNGICDCGDKEAWNHELNCKGAEDGGSHHHHHHHHHHGGGSGGGSGGGS GLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS694 (ビスPS528、 リンカー2)	270	MHSKFNHAGRICGAKFRVGEPIYRCKECSFDDTCVLCVNCFNPKDHVGHHV YTSICTEFNNGICDCGDKEAWNHELNCKGAEDGSAGSAAGSGEFMHSKFNH AGRICGAKFRVGEPIYRCKECSFDDTCVLCVNCFNPKDHVGHHVYTSICTE FNNGICDCGDKEAWNHELNCKGAEDGGSHHHHHHHHHHGGGSGGGSGGGSG LNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS695 (ビスPS528、 リンカー３)	271	MHSKFNHAGRICGAKFRVGEPIYRCKECSFDDTCVLCVNCFNPKDHVGHHV YTSICTEFNNGICDCGDKEAWNHELNCKGAEDGSAGSAAGSGEFGSAGSAA GSGEFGSAGSAAGSGEFMHSKFNHAGRICGAKFRVGEPIYRCKECSFDDTC VLCVNCFNPKDHVGHHVYTSICTEFNNGICDCGDKEAWNHELNCKGAEDGG SHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSG LNDFFEAQKIEWHE

【０４０４】

多価バインダーＰＳ６１２は、２つの異なるＣｌｐＳタンパク質モノマー（ａｔＣｌｐＳ２－Ｖ１及びＰＳ３７２）の直列コピーを含むポリペプチドである。ＰＳ６１２のアミノ酸結合プロファイルを、上述のようにバイオレイヤー干渉法によって評価した。１価バインダーａｔＣｌｐＳ２－Ｖ１及びＰＳ３７２を、対照として個別に行った。図３６Ａは、ａｔＣｌｐＳ２－Ｖ１（左側プロット）及びＰＳ３７２（右側プロット）の反応曲線を示す。各モノマー形バインダーについて以前観測されたように、ａｔＣｌｐＳ２－Ｖ１は、Ｎ末端Ｆ、Ｗ、及びＹに結合し、ＰＳ３７２は、Ｎ末端Ｉ、Ｌ、及びＶに結合する。直列バインダーＰＳ６１２の反応曲線は、図３６Ｂに示される。示されるように、結合は、Ｎ末端Ｆ、Ｗ、Ｙ、Ｉ、Ｌ、及びＶを有するＰＳ６１２について観測された。これらの結果は、２つの異なるＣｌｐＳモノマー（この例ではａｔＣｌｐＳ２－Ｖ１及びＰＳ３７２）の直列コピーからなる単一ポリペプチドが各親ＣｌｐＳタンパク質の完全なアミノ酸結合能力を示すことを実証する。

【０４０５】

多価バインダーＰＳ６１４は、直列に配向されたＣｌｐＳバインダーＰＳ３７２の２つのコピーを含むポリペプチドである。ＰＳ６１４のアミノ酸結合プロファイルを、上述のようにバイオレイヤー干渉法によって評価した。図３６Ｃは、１価バインダーＰＳ３７２についての反応曲線を示し、ＰＳ３７２がＮ末端Ｉ、Ｌ、及びＶに結合するという以前の

観察を裏付ける。直列バインダーＰＳ６１４についての反応曲線は、図３６Ｄに示される。示されるように、Ｎ末端Ｉ、Ｌ、及びＶを有するＰＳ６１４への結合が観測された。これらの結果は、直列ＰＳ３７２構築物がＬ、Ｉ、及びＶへの強い結合を保持することを実証する。

【０４０６】

多価バインダーＰＳ６３７、ＰＳ６３８、及びＰＳ６３９は、直列に配向され、それぞれリンカー１、リンカー２、及びリンカー３によって分離されるＣｌｐＳバインダーの２つのコピーを各々含むポリペプチドである。各バインダーのアミノ酸結合プロファイルを、上述のようにバイオレイヤー干渉法によって評価した。図３６Ｅは、１価バインダーＰＳ５５７についての反応曲線を示し、ＰＳ５５７がＮ末端Ｉ、Ｌ、及びＶに結合するという以前の観察を裏付ける。直列バインダーＰＳ６３７、ＰＳ６３８、及びＰＳ６３９についての反応曲線は、それぞれ図３６Ｆ、３６Ｇ、及び３６Ｈに示される。示されるように、Ｎ末端Ｉ、Ｌ、及びＶを有する直列バインダーへの結合が観測された。これらの結果は、さまざまなリンカーを含む直列ＰＳ５５７構築物がＩ、Ｌ、及びＶへの結合を保持することを実証する。

【０４０７】

上記の結果は、さまざまな多価ＣｌｐＳバインダーが、Ｆ、Ｙ、Ｗ、Ｌ、Ｉ、及びＶを含む末端アミノ酸への結合を保持することを示した。追加の多価アミノ酸バインダーを、フレキシブルペプチドリンカーをコードするセグメントによって末端どうして連結されたＵＢＲボックスバインダーの２つのコピーをコードするセグメントを含む単一コーディング配列から発現した直列融合分子として設計した。単一コーディング配列の発現により、直列に配向された２つのＵＢＲボックスホモログ（ビスＰＳ６２１、ビスＰＳ５２８）を有する単一全長ポリペプチドが生成される。

【０４０８】

３つの異なる構築物を、各多価ＵＢＲボックスバインダーについて設計して発現させた：ＰＳ６９０、ＰＳ６９１、及びＰＳ６９２（それぞれリンカー１、リンカー２、及びリンカー３を有するビスＰＳ６２１バインダー）、並びにＰＳ６９３、ＰＳ６９４、及びＰＳ６９５（それぞれリンカー１、リンカー２、及びリンカー３を有するビスＰＳ５２８）。蛍光偏光研究を実施して、ペナルティメート位置のアラニン残基とともにＮ末端アルギニン残基を有するペプチドについて、多価ＵＢＲボックスバインダー及び対応するモノマー形バインダーの各々の結合親和性を評価した。これらの研究から得られたK_D値を、以下の表１１に報告する。

【０４０９】

【表１１－１】

表１１．N末端アルギニンペプチドに対する直列認識の結合親和性。

バインダー	RA ペプチド (K_D, nM)
PS528	403
PS621	321
PS690 (ビスPS621、リンカー１)	354
PS691 (ビスPS621、リンカー２)	369
PS692 (ビスPS621、リンカー３)	478

【０４１０】

【表１１－２】

PS693（ビスPS528、リンカー１）	263
PS694（ビスPS528、リンカー2）	353
PS695（ビスPS528、リンカー3）	317

【０４１１】

これらの研究の結果は、多価バインダーが、複数のＣｌｐＳ又はＵＢＲボックスタンパク質ホモログをコードするセグメントを含む単一コーディング配列から発現した直列融合分子として得ることができることを実証する。そのほか、これらの多価アミノ酸バインダーは、Ｆ、Ｙ、Ｗ、Ｌ、Ｉ、Ｖ、及びＲを含む末端アミノ酸への結合を保持することを示される。

【０４１２】

実施例１９．光損傷軽減のためのシールド認識分子融合物

実施例１０で記載されるように、ＤＮＡストレプトアビジンシールディングエレメントの使用は、動的シーケンシング反応の間、固定されたペプチドの光安定性を向上させることが示された。追加のシールド認識分子を、フレキシブルペプチドリンカーによって末端同士で結合された、アミノ酸結合タンパク質及びタンパク質シールドの１つ又は２つのコピーをコードするセグメントを含む単一コーディング配列から発現した直列融合分子として設計した。こうして、シールド認識分子の結合成分及びシールディングエレメントは、単一発現構築物から生成されうる。表１２は、設計されたさまざまな融合構築物及び対応するポリペプチド配列のリストを提供する。

【０４１３】

【表１２－１】

表１２．融合構築物の非制限的例。

名称	配列番号	配列
PS696 (PS557-SNAP タグ融合物)	272	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMDKDCEMKRTTLDSPLGKLELS GCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQP EAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNP AATAAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHR LGKPGLGGSAGSAAGSGEFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQK IEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS697 (PS557-2x SNAP タグ融合物)	273	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMDKDCEMKRTTLDSPLGKLELS GCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQP EAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNP AATAAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHR LGKPGLGGMDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAV EVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFT RQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTALSGNPVPILIPCH RVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLGGSAGSAAGSGEFHHH HHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDF FEAQKIEWHE

【０４１４】

【表１２－２】

PS698 (PS557-SNAP タグ融合物２)	274	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIE WHEGGGSGGGSGGGSGLNDFFEAQKIEWHEGSAGSAAGSGEFMDKDCEMKR TTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLM QATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEV ISYQQLAALAGNPAATAAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGL AVKEWLLAHEGHRLGKPGLG
PS699 (PS557-EF 融合物)	275	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMIKKYTGDFETTTDLNDCRVWS WGVCDIDNVDNITFGLEIDSFFEWCEMQGSTDIYFHDEKFDGEFMLSWLFK NGFKWCKEAKEERTFSTLISNMGQWYALEICWNVKCTTTKTGKTKKEKQRT IIYDSLKKYPFPVKEIAEAFNFPIKKGEIDYTKERPIGYNPTDDEWDYLKN DIQIMAMALKIQFDQGLTRMTRGSDALGDYQDWVKTTYGKSRFKQWFPVLS LGFDKDLRKAYKGGFTWVNKVFQGKEIGEGIVFDVNSLYPSQMYVRPLPYG TPLFYEGEYKENIDYPLYIQNIKVRFRLKERHIPTIQVKQSSLFIQNEYLE SSVNKLGVDELIDLTLTNVDLDLFFEHYDILEIHYTYGYMFKASCDMFKGW IDKWIEVKNTTEGARKANAKGMLNSLYGKFGTNPDITGKVPYMGEDGIVRL TLGEEELRDPVYVPLASFVTAWGRYTTIITAQRCFDNIIYCDTDSIHLTGT EVPEAIEHLVDSKKLGYWKHESTFQRAKFIRQKTYVEEIDGELNVKCAGMP DRIKELVTFDNFEVGFSSYGKLLPKRTQGGVVLVDTMFTIKGSAGSAAGSG EFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGS GLNDFFEAQKIEWHE
PS700 (PS557-MBP 融合物)	276	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMKIEEGKLVIWINGDKGYNGLA EVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQ SGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLP NPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENG KYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMT INGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELA KEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKG EIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNN NNLGIEGRGSAGSAAGSGEFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQ KIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS701 (PS557-GST 融合物)	277	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMSPILGYWKIKGLVQPTRLLLE YLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIR YIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLS KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPK LVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPRGS PGIHRDGSAGSAAGSGEFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKI EWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS702 (PS557-GFP 融合物)	278	MPTAASATESAIEDTPAPARPEVDGRTKPKRQPRYHVVLWNDDDHTYQYVV VMLQSLFGHPPERGYRLAKEVDTQGRVIVLTTTREHAELKRDQIHAFGYDR LLARSKGSMKASIEAEEGSAGSAAGSGEFMSKGEELFTGVVPILVELDGDV NGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSR YPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIE LKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTA AGITHGMDELYKGSAGSAAGSGEFHHHHHHHHHHGGGSGGGSGGGSGLNDF FEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS703	279	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGT

【０４１５】

【表１２－３】

(atClpS2-V1-SNAP タグ融合物)		SAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVF QQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTALSGNPVP ILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLGGSAGSAAGS GEFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGG SGLNDFFEAQKIEWHE
PS704 (atClpS2-V1-2x SNAP タグ融合物)	280	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGT SAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVF QQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTALSGNPVP ILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLGMDKDCEMKR TTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLM QATAWLNAYFHQPEAIEEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEV ISYQQLAALAGNPAATAAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGL AVKEWLLAHEGHRLGKPGLGGSAGSAAGSGEFHHHHHHHHHHGGGSGGGSG GGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS705 (atClpS2-V1-SNAP タグ融合物２)	281	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSG LNDFFEAQKIEWHEGSAGSAAGSGEFMDKDCEMKRTTLDSPLGKLELSGCE QGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAI EEFPVPALHHPVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAAT AAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGK PGLG
PS706 (atClpS2-V1-EF 融合物)	282	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMIKKYTGDFETTTDLNDCRVWSWGVCDIDNVDNITFGL EIDSFFEWCEMQGSTDIYFHDEKFDGEFMLSWLFKNGFKWCKEAKEERTFS TLISNMGQWYALEICWNVKCTTTKTGKTKKEKQRTIIYDSLKKYPFPVKEI AEAFNFPIKKGEIDYTKERPIGYNPTDDEWDYLKNDIQIMAMALKIQFDQG LTRMTRGSDALGDYQDWVKTTYGKSRFKQWFPVLSLGFDKDLRKAYKGGFT WVNKVFQGKEIGEGIVFDVNSLYPSQMYVRPLPYGTPLFYEGEYKENIDYP LYIQNIKVRFRLKERHIPTIQVKQSSLFIQNEYLESSVNKLGVDELIDLTL TNVDLDLFFEHYDILEIHYTYGYMFKASCDMFKGWIDKWIEVKNTTEGARK ANAKGMLNSLYGKFGTNPDITGKVPYMGEDGIVRLTLGEEELRDPVYVPLA SFVTAWGRYTTITTAQRCFDNIIYCDTDSIHLTGTEVPEAIEHLVDSKKLG YWKHESTFQRAKFIRQKTYVEEIDGELNVKCAGMPDRIKELVTFDNFEVGF SSYGKLLPKRTQGGVVLVDTMFTIKGSAGSAAGSGEFHHHHHHHHHGGGS GGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGSGLNDFFEAQKIEWHE
PS707 (atClpS2-V1-MBP 融合物)	283	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTV EHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDK LYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKEL KAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKA GLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVN YGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEA VNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAV RTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRGSAGSAAG SGEFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGG GSGLNDFFEAQKIEWHE
PS708 (atClpS2-V1-GST 融合物)	284	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG DKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKER AEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHK TYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDK

【０４１６】

【表１２－４】

		YLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPRGSPGIHRDGSAGSAAGSG EFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGGGSGGGS GLNDFFEAQKIEWHE
PS709 (atClpS2-V1- GFP 融合物)	285	MSDSPVDLKPKPKVKPKLERPKLYKVMLLNDDYTPMSFVTVVLKAVFRMSE DTGRRVMMTAHRFGSAVVVVCERDIAETKAKEATDLGKEAGFPLMFTTEPE EGSAGSAAGSGEFMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT YGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMP EGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK LEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDG PVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYKGSAG SAAGSGEFHHHHHHHHHHGGGSGGGSGGGSGLNDFFEAQKIEWHEGGGSGG GSGGGSGLNDFFEAQKIEWHE

【０４１７】

表１２に示すように、各融合構築物は、以下のタンパク質シールド：リンカー２を介して融合されたＳＮＡＰタグタンパク質（「ＳＮＡＰタグ融合物」として示される）、直列に配向されたＳＮＡＰタグタンパク質の２つのコピー（「２ｘＳＮＡＰタグ融合物」として示される）、リンカー２及びＨｉｓ／ビスビオチンタグを含むセグメントを介して融合されたＳＮＡＰタグタンパク質（「ＳＮＡＰタグ融合物２」として示される）、ＤＮＡポリメラーゼ（「ＥＦ融合物として示される」）、マルトース結合タンパク質（「ＭＢＰ融合物」として示される）、グルタチオンＳトランスフェラーゼタンパク質（「ＧＳＴ融合物」として示される）、又は緑色蛍光タンパク質（「ＧＦＰ融合物」として示される）のうちの１つに融合したアミノ酸結合タンパク質（ＰＳ５５７又はａｔＣｌｐＳ２－Ｖ１）を含んだ。

【０４１８】

ストップトフローアッセイを実施して、ＰＳ５５７、ａｔＣｌｐＳ２－Ｖ１、及びタンパク質シールドのＣ末端付加によって駆動される融合タンパク質について、会合（k_{on}）及び解離（k_{off}）の速度を測定した。ストップトフローによってオン速度定数を測定するために、バインダーをＦＩＴＣペプチドと迅速に混合し、反応をリアルタイムで追跡した。アッセイ設計を例示する模式図を図３７Ａ（トップパネル）に示す。Ｎ末端ＬＡペプチドをＰＳ５５７派生物に使用し、Ｎ末端ＦＡペプチドをａｔＣｌｐＳ２－Ｖ１派生物に使用した。このアッセイでは、蛍光シグナルは、タンパク質結合時のクエンチングにより減少した。複数の濃度のバインダーで得られた平均トレースを減衰方程式にあてはめて、会合速度を導出した（図３７Ａ、中央パネル）。さまざまなバインダー濃度に対する会合速度のプロットの線形勾配は、k_{on}速度定数を与える（図３７Ａ、ボトムパネル）。

【０４１９】

ストップトフローによって解離速度（k_{off}）を測定するために、最適な濃度で標識付ペプチド基質に結合するバインダーの複合体を過剰非標識トラップペプチドと迅速に混合した。アッセイ設計を例示する模式図を図３７Ｂ（トップパネル）に示す。ペプチドＮ末端からのバインダー解離は、クエンチングの逆転により、蛍光シグナルの増加をもたらした。平均速度トレースを指数方程式に当てはめて、k_{off}を決定した（図３７Ｂ、ボトムパネル）。

【０４２０】

ストップトフローアッセイによって決定された速度は、表１３（ＰＳ５５７及びＰＳ５５７派生融合物）及び表１４（ａｔＣｌｐＳ２－Ｖ１及びａｔＣｌｐＳ２－Ｖ１派生融合物）に示される。

【０４２１】

【表１３】

表１３．PS557融合物に対するLAペプチドk_{on}速度定数及びk_{off}速度

バインダー	LA kon (nM/s)	LA koff (/s)
PS557	0.0018	0.22
PS702 (PS557 GFP 融合物)	0.0025	0.1
PS696 (PS557-SNAP タグ融合物)	–	0.15
PS698 (PS557-SNAP タグ融合２)	0.004	0.19

【０４２２】
【表１４】

表１４．atClpS2-V1融合物に対するFAペプチドk_{on}速度定数及びk_{off}速度

バインダー	FA kon (nM/s)	FA koff (/s)
atClpS2-V1	0.0015	1.187
PS706 (atClpS2-V1-EF 融合物)	0.0014	0.58909
PS703 (atClpS2-V1-SNAP タグ融合物)	–	0.76741
PS704 (atClpS2-V1-2xSNAP タグ融合物)	–	0.6774
PS705 (atClpS2-V1-SNAPタグ融合物２)	–	1.2655
PS709 (atClpS2-V1-GFP 融合物)	–	0.9344

【０４２３】
結果は、タンパク質シールド構築物とＣｌｐＳタンパク質とのＣ末端融合物が、単一分子ペプチドシーケンシングアッセイでの使用に有用なＣｌｐＳ対照に類似の反応速度論プロフィールを有する完全に活性化したＮ末端認識剤を生成することを実証する。
【０４２４】
均等物及び範囲
請求項では、「ａ」、「ａｎ」、及び「ｔｈｅ」などの冠詞は、相反する指示がない限り又はとくに文脈から自明でない限り、１つ又は２つ以上を意味しうる。グループの１つ以上のメンバー間に「ｏｒ（又は）」を含む請求項又は説明は、相反する指示がない限り又はとくに文脈から自明でない限り、グループメンバーの１つ、２つ以上、又はすべてが所与のプロダクト又はプロセスに存在するか、利用されるか、又は他の形で関連する場合、満足されるとみなされる。本発明は、グループの厳密に１つのメンバーが所与のプロダクト又はプロセスに存在するか、利用されるか、又は他の形で関連する実施形態を含む。本発明は、グループメンバーの２つ以上又はすべてが所与のプロダクト又はプロセスに存在するか、利用されるか、又は他の形で関連する実施形態を含む。
【０４２５】
さらに、本発明は、列挙された請求項の１つ以上から他の請求項に１つ以上の限定、要素、文節、及び記述語が導入された、変形形態、組合せ形態、及び入替え形態をすべて包含する。たとえば、他の請求項に依存するいずれかの請求項は、同一基本請求項に依存するいずれかの他の請求項に見いだされる１つ以上の限定を含むように修正可能である。要素がリストとして、たとえば、マーカッシュグループ形式で提示された場合、要素の各サブグループもまた開示され、且つ任意の要素をグループから除外可能である。一般に、本発明又は本発明の態様が、特定の要素及び／又は特徴を含んで参照される場合、本発明の特定の実施形態又は本発明の態様は、そのような要素及び／又は特徴からなるか又は本質的にそれらからなると理解されるべきである。簡単にするために、それらの実施形態は、本明細書では、とくにこの通りの言葉で明記されていない。
【０４２６】

本明細書及び特許請求の範囲で用いられる「ａｎｄ／ｏｒ（及び／又は）」という語句は、そのように結合された要素の「一方又は両方」、すなわち、いくつかの場合には連結して存在し且つ他の場合には分離して存在する要素を意味するものと理解すべきである。「ａｎｄ／ｏｒ（及び／又は）」を用いて列挙された複数の要素は、同様に解釈されるべきである。すなわち、要素の「１つ以上」は、そのように結合される。任意選択的に、具体的に特定された要素に関連するかしないかにかかわらず、「ａｎｄ／ｏｒ（及び／又は）」という句により具体的に特定された以外の他の要素が存在していてもよい。したがって、たとえば、「Ａ及び／又はＢ」という参照語は、「ｃｏｍｐｒｉｓｉｎｇ（～を含む）」などのオープンエンドの言語と組み合わせて用いられる場合、一実施形態では、Ａのみ（任意選択的にＢ以外の要素を含めて）、他の実施形態では、Ａのみ（任意選択的にＡ以外の要素を含めて）、さらに他の実施形態では、Ａ及びＢの両方（任意選択的に他の要素を含む）などを意味しうる。

【０４２７】

本明細書及び特許請求の範囲で用いられる場合、「ｏｒ（又は）」は、以上に定義された「ａｎｄ／ｏｒ（及び／又は）」と同一の意味を有するものと理解すべきである。たとえば、リスト中の分離したアイテムの場合、「ｏｒ（又は）」又は「ａｎｄ／ｏｒ（及び／又は）」は、包括的なものと解釈されるものとする。すなわち、少なくとも１つを含むが、それだけでなく、いくつかの要素又は列挙された要素の２つ以上も含まれ、任意選択的に、追加の列挙されていないアイテムも含まれる。明らかに相反する指定がされている用語、たとえば、「ｏｎｌｙ　ｏｎｅ　ｏｆ（～のうちの１つのみ）」又は「ｅｘａｃｔｌｙ　ｏｎｅ　ｏｆ（～のうちの厳密に１つ）」又は特許請求の範囲で用いられる場合の「ｃｏｎｓｉｓｔｉｎｇ　ｏｆ（～からなる）」は、いくつかの要素又は列挙された要素のうち厳密に１つの要素の包含を意味するであろう。一般的には、本明細書で用いられる「ｏｒ（又は）」という用語は、「ｅｉｔｈｅｒ（いずれか一方）」、「ｏｎｅ　ｏｆ（～の１つ）」、「ｏｎｌｙ　ｏｎｅ　ｏｆ（～のうちの１つのみ）」、又は「ｅｘａｃｔｌｙ　ｏｎｅ　ｏｆ（～のうちの厳密に１つ）」などのように排他性の用語が先行する場合、排他的代替物（すなわち「一方又は他方ただし両方ではない」）を表すとのみ解釈されるものとする。特許請求の範囲で用いられる「ｃｏｎｓｉｓｔｉｎｇ　ｅｓｓｅｎｔｉａｌｌｙ　ｏｆ（～から本質的になる）」は、特許法の分野で用いられるその通常の意味を有するものとする。

【０４２８】

本明細書及び特許請求の範囲で用いられる場合、１つ以上の要素のリストを参照する「少なくとも１つ」という語句は、要素のリスト中の任意の１つ以上の要素から選択される少なくとも１つの要素を意味すると理解すべきであるが、必ずしも、要素のリスト内に具体的に列挙されたあらゆる要素の少なくとものの１つを含むとは限らず、要素のリスト中の要素の任意の組合せが除外されるとは限らない。また、この定義では、具体的に特定された要素に関連するかしないかにかかわらず、「少なくとも１つ」という語句が参照する要素のリスト内に具体的に特定された要素以外の要素が、任意選択的に存在してもよいことも許容される。そのため、限定されるものではないが例として、「ａｔ　ｌｅａｓｔ　ｏｎｅ　ｏｆ　Ａ　ａｎｄ　Ｂ（Ａ及びＢの少なくとも１つ）」（又は等価的に「ａｔ　ｌｅａｓｔ　ｏｎｅ　ｏｆ　Ａ　ｏｒ　Ｂ（Ａ又はＢの少なくとも１つ）」又は等価的に「ａｔ　ｌｅａｓｔ　ｏｎｅ　ｏｆ　Ａ　ａｎｄ／ｏｒ　Ｂ（Ａ及び／又はＢの少なくとも１つ）」）は、一実施形態では、Ｂが存在せずに（任意選択的にＢ以外の要素を含めて）、任意選択的に２つ以上のＡを含めて、少なくとも１つのＡ、他の実施形態では、Ａは存在せずに（任意選択的にＡ以外の要素を含めて）、任意選択的に２つ以上のＢを含めて、少なくとも１つのＢ、さらに他の実施形態では、任意選択的に２つ以上のＡを含めて、少なくとも１つのＡ、及び任意選択的に２つ以上のＢを含めて、少なくとも１つのＢ（任意選択的に他の要素を含めて）などを意味しうる。

【０４２９】

また、明らかに相反する指定がされていない限り、２つ以上の工程又は行為を含む本明

細書に特許請求されたいずれの方法でも、方法の工程又は行為の順序は、方法の工程又は行為が挙げられた順序に必ずしも限定されるとは限らないことを理解すべきである。

【０４３０】

特許請求の範囲では、以上の明細書の場合と同様に、「ｃｏｍｐｒｉｓｉｎｇ（～を含む）」、「ｉｎｃｌｕｄｉｎｇ（～を含む）」、「ｃａｒｒｙｉｎｇ（～を担持する）」、「ｈａｖｉｎｇ（～を有する）」、「ｃｏｎｔａｉｎｉｎｇ（～を含有する）」、「～を含む（ｉｎｖｏｌｖｉｎｇ）」、「～を保持する（ｈｏｌｄｉｎｇ）」、「ｃｏｍｐｏｓｅｄ　ｏｆ（～で構成される）」などの移行句はすべて、オープンエンドであることを理解すべきである。「ｃｏｎｓｉｓｔｉｎｇ　ｏｆ（～からなる）」及び「ｃｏｎｓｉｓｔｉｎｇ　ｅｓｓｅｎｔｉａｌｌｙ　ｏｆ（～から本質的になる）」という移行句だけは、米国特許庁特許審査手順マニュアル（Ｕｎｉｔｅｄ　Ｓｔａｔｅｓ　Ｐａｔｅｎｔ　Ｏｆｆｉｃｅ　Ｍａｎｕａｌ　ｏｆ　Ｐａｔｅｎｔ　Ｅｘａｍｉｎｉｎｇ　Ｐｒｏｃｅｄｕｒｅｓ）のセクション２１１１．０３に示されるように、それぞれ、クローズ又はセミクローズの移行句であるものとする。オープンエンドの移行句（たとえば、「ｃｏｍｐｒｉｓｉｎｇ（～を含む）」）を用いた本文書に記載の実施形態は、代替実施形態では、オープンエンドの移行句により記載される特徴について「ｃｏｎｓｉｓｔｉｎｇ　ｏｆ（～からなる）」及び「ｃｏｎｓｉｓｔｉｎｇ　ｅｓｓｅｎｔｉａｌｌｙ　ｏｆ（～から本質的になる）」として企図されることが認識されるべきである。たとえば、適用が「ＡとＢとを含む組成物」を記載する場合、適用はまた、代替実施形態の「ＡとＢとからなる組成物」及び「ＡとＢとから本質的になる組成物」も企図する。

【０４３１】

範囲が与えられる場合、端点が含まれている。さらに、とくに指定がないかぎり又はとくに文脈及び当業者の理解から自明でないかぎり、範囲として表現された値は、文脈上明らかに異なる場合を除いて、範囲の下限の１／１０単位まで、本発明のさまざまな実施形態で指定の範囲内の任意の特定の値又は部分範囲を仮定しうる。

【０４３２】

本出願では、さまざまな発行特許、公開特許出願、雑誌論文、及び他の出版物（それらはすべて参照により本明細書に組み込まれる）が参照される。組み込まれた参照のいずれかと本明細書との間で矛盾がある場合、本明細書が優先するものとする。それに加えて、先行技術に包含される本発明の特定の実施形態はいずれも、請求項の任意の１項以上から明示的に除外されうる。そのような実施形態は、当業者に公知であるとみなされるので、たとえ本明細書で明示的に除外が明記されていないとしても、それらは除外されうる。先行技術の存在に関係するかしないかにかかわらず、任意の理由で、本発明の任意の特定の実施形態を任意の請求項から除外しうる。

【０４３３】

当業者であれば、通常の実験の域を出ることなく、本明細書で説明した特定の実施形態に対する多くの等価形態が分かるであろう。又はそれらを確認できるであろう。本明細書に記載の実施形態の範囲は、以上の説明に限定されるものではなく、添付の特許請求の範囲に明記される通りである。以下の特許請求の範囲に規定される本発明の趣旨又は範囲から逸脱することなく、本明細書に各種変更及び修正を加えうることは当業者であれば分かるであろう。

【０４３４】

本明細書の変数のいずれかの定義での化学基のリストの列挙は、列挙された基のいずれかの単一基又は組合せとしてのその変数の定義を含む。本明細書の変数に対する実施形態の列挙は、いずれかの単一実施形態又はいずれかの他の実施形態若しくはその一部分との組合せとしての実施形態を含む。本明細書の実施形態の列挙は、いずれかの単一実施形態又はいずれかの他の実施形態若しくはその一部分との組合せとして実施形態を含む。

【図面】

【図１Ａ】

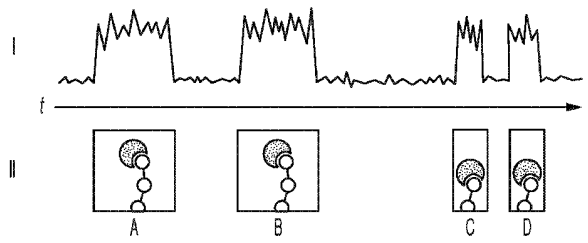


FIG. 1A

【図１Ｂ】

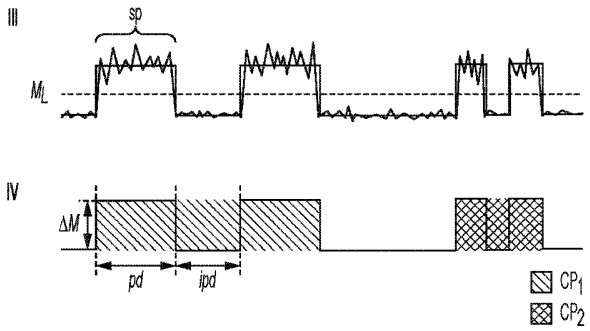


FIG. 1B

【図２】

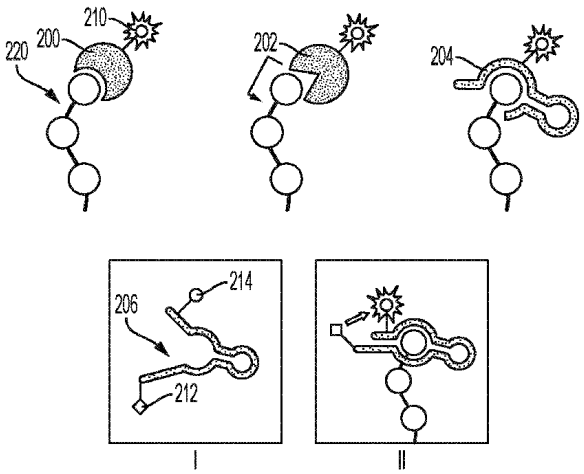


FIG. 2

【図３Ａ】

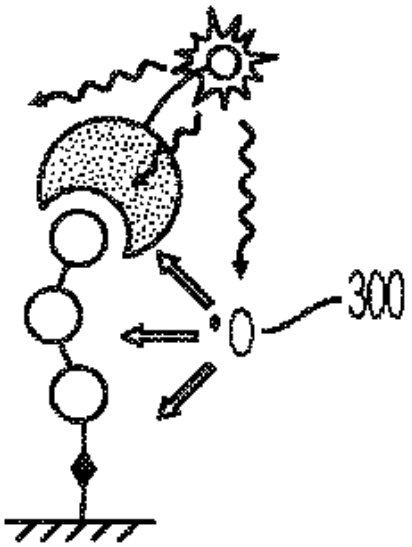


FIG. 3A

【図３Ｂ】

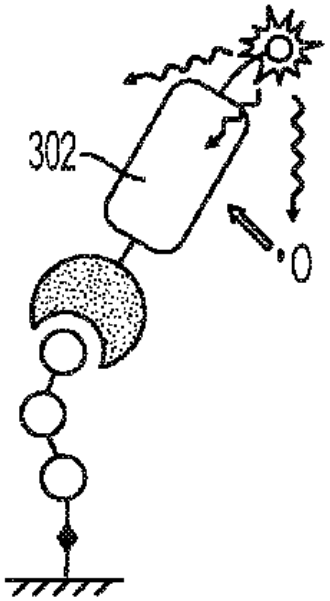


FIG. 3B

【図３Ｃ】

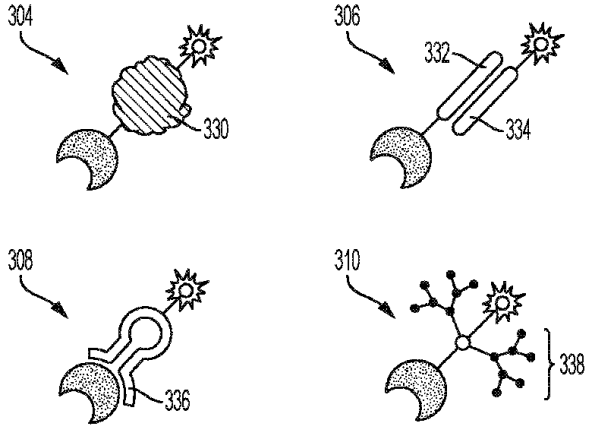


FIG. 3C

【図３Ｄ】

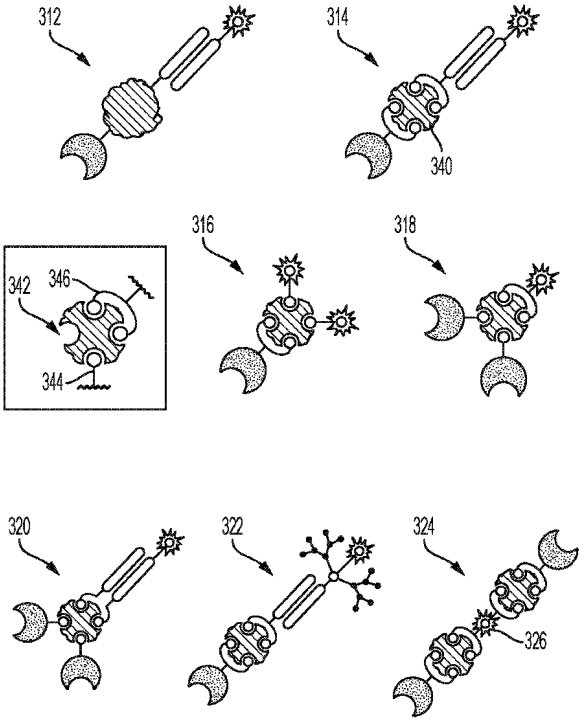


FIG. 3D

【図３Ｅ】

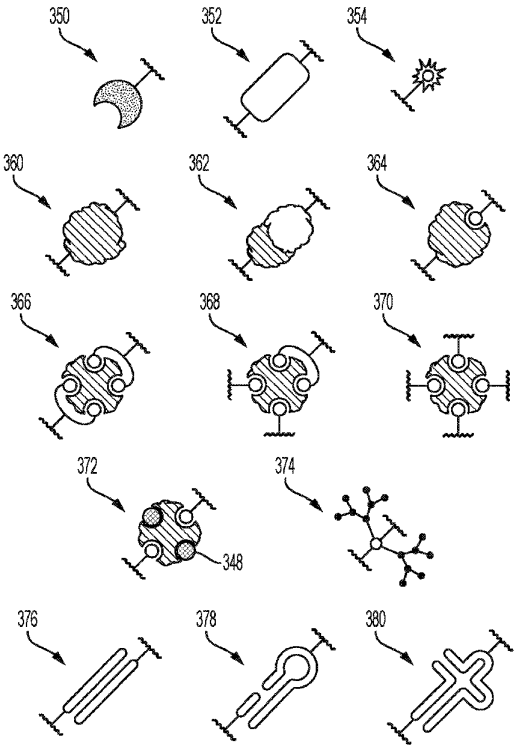


FIG. 3E

【図４】

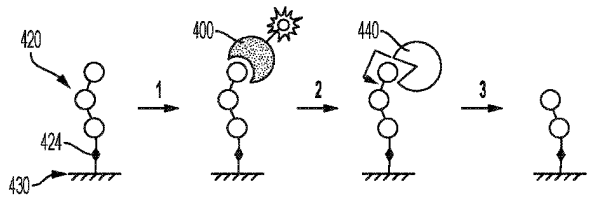
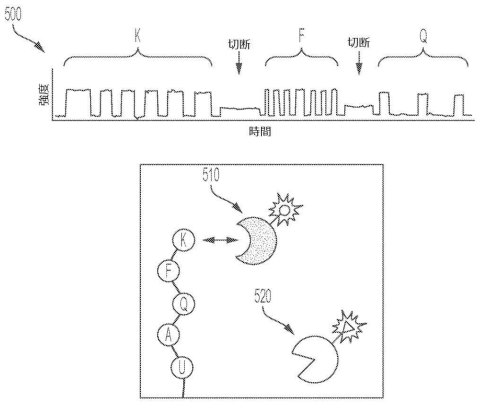


FIG. 4

【図５】



【図６】

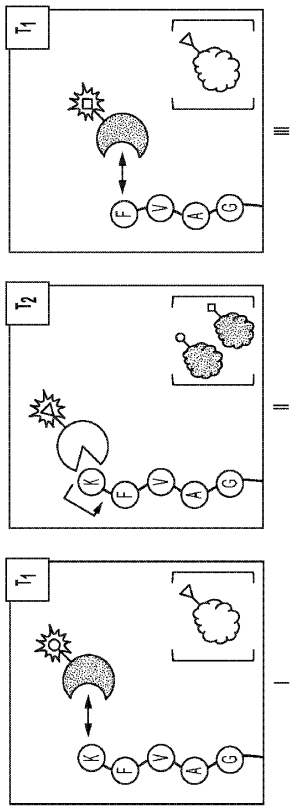
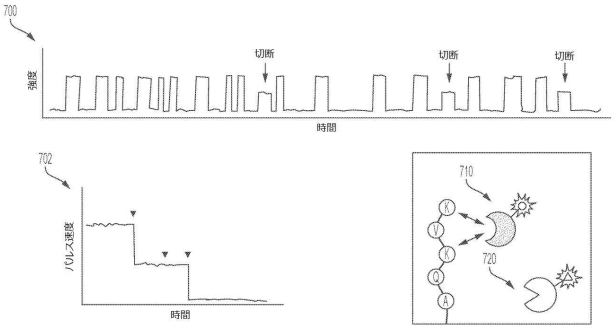


FIG. 6

【図７】



【図８】

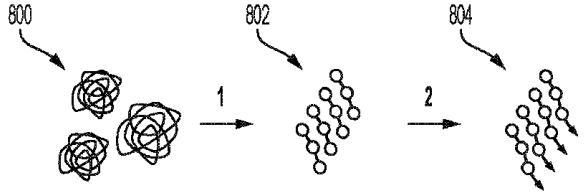


FIG. 8

【図９】

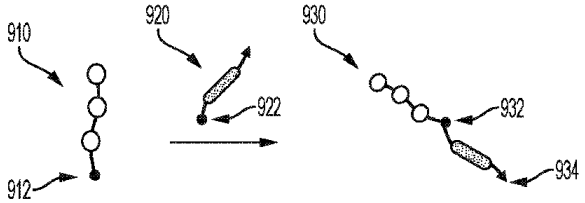


FIG. 9

【図１０】

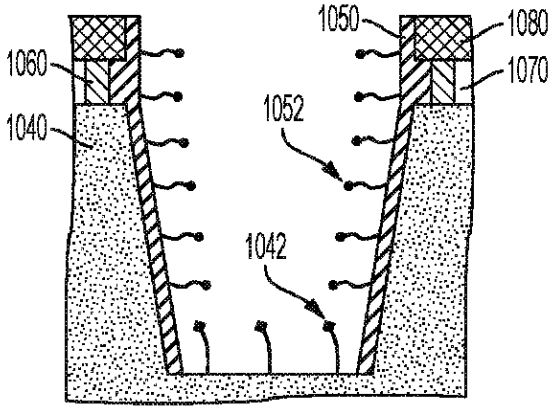
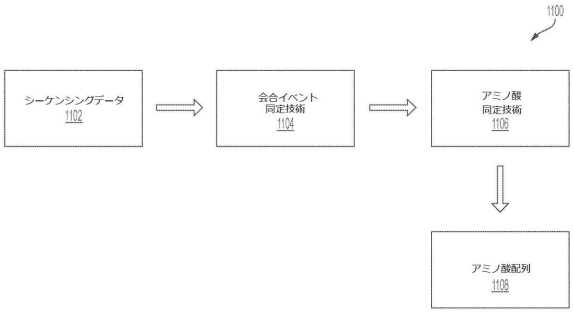
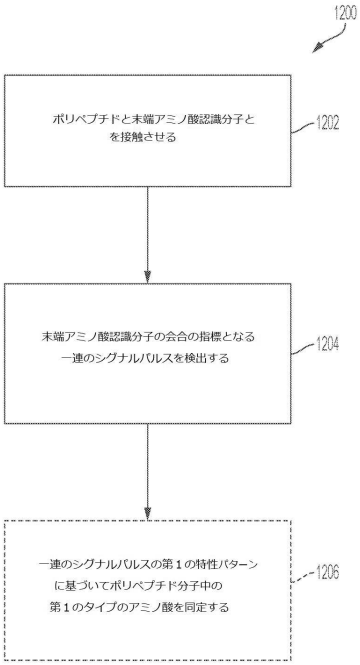


FIG. 10

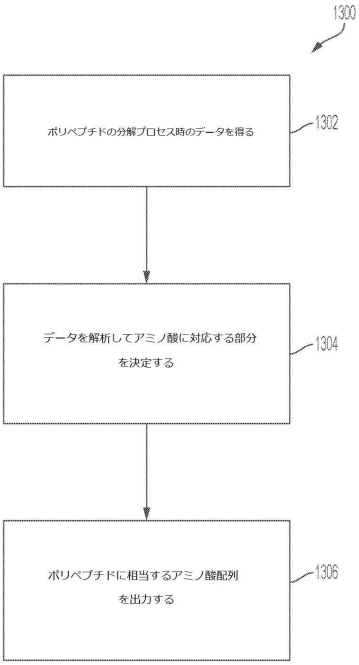
【図１１】



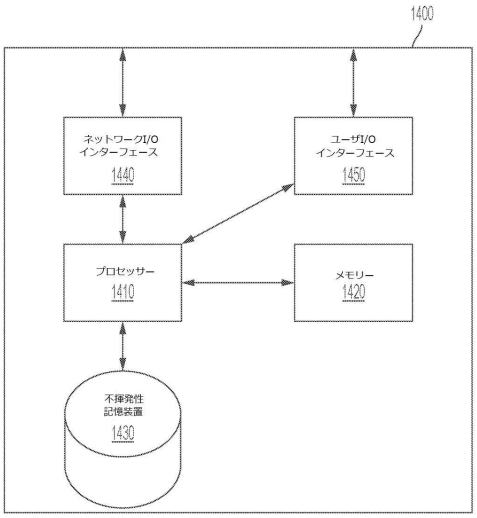
【図１２】



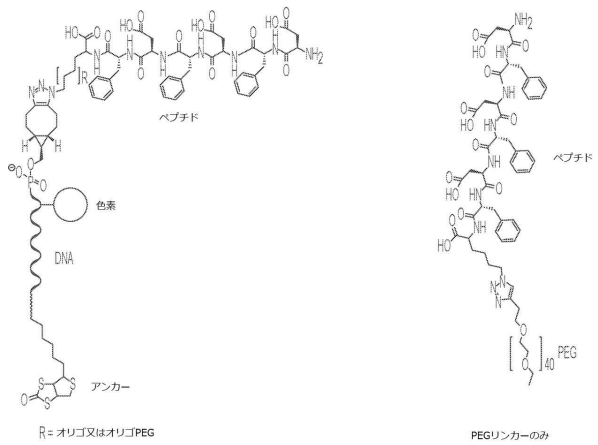
【図１３】



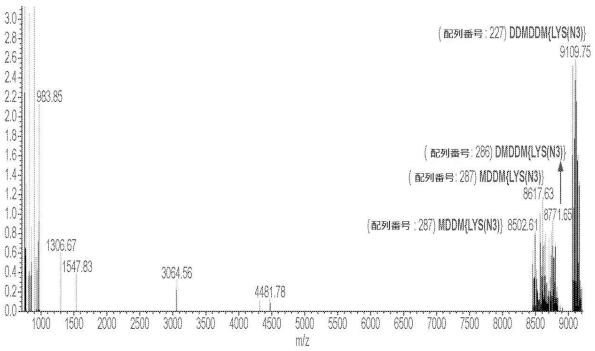
【図１４】



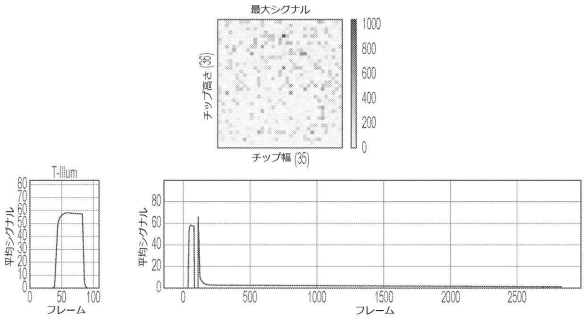
【図１５Ａ】



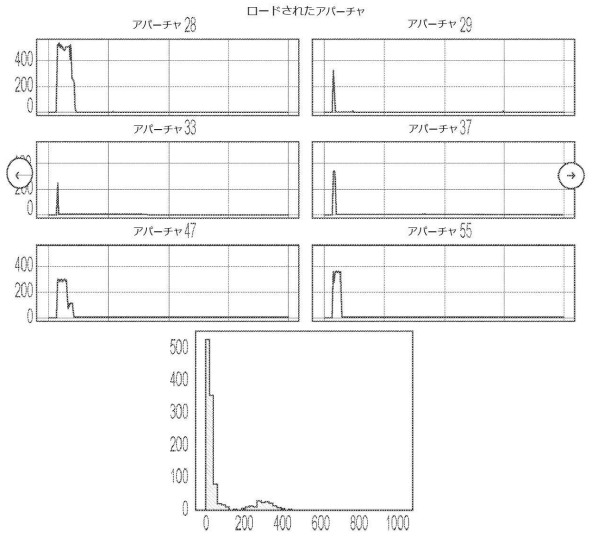
【図１５Ｂ】



【図１５Ｃ－１】



【図１５Ｃ－２】

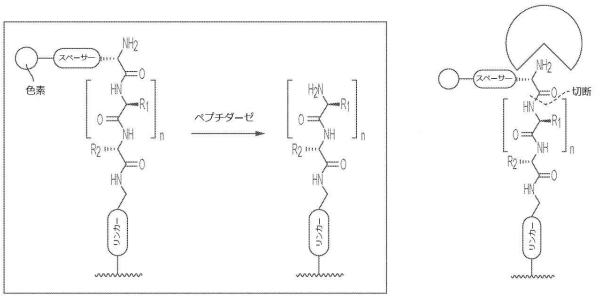


【図１６】



酵素	高い比活性を有する残基（K≧７５%、３０分 １μM基質、１μM酵素、２５℃での切断）	中程度の比活性 (75%≦X≧25%)	低い比活性 (25%≦X>0%)	活性なし
cvPr	R, A, F, Y, V, I, L, T, W, M	K, Y, Q, S	N, G, H	D, E, P (2つのPが続く場合), C
yPIP		P (2つのPが続く場合)		R, A, F, V, I, L, T, W, M, K, Y, Q, S, N, G, H, D, E, C
DIE APN		E, I	D	R, A, F, V, L, T, W, M, K, Y, Q, S, N, G, H, C, P
hTET	A, F, Y, V, I, L, S, T	N, Q, C, H, G	D, W	R, K, E, M
Pfu API	R, A, K, Y, D, E, N, Q, S, T, M, L	W, F, C	H	V, I, G

【図１７Ａ－１】



【図１７Ａ－２】

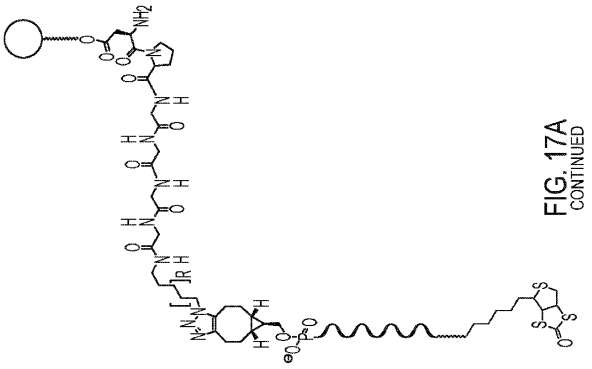
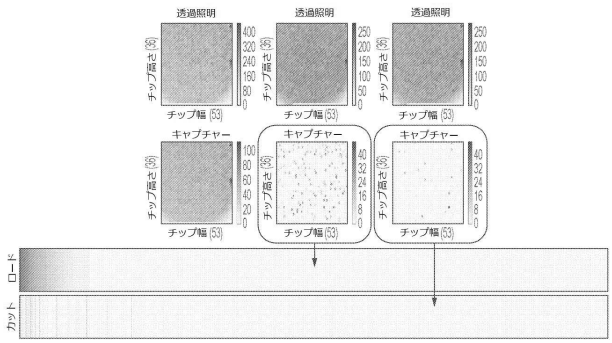
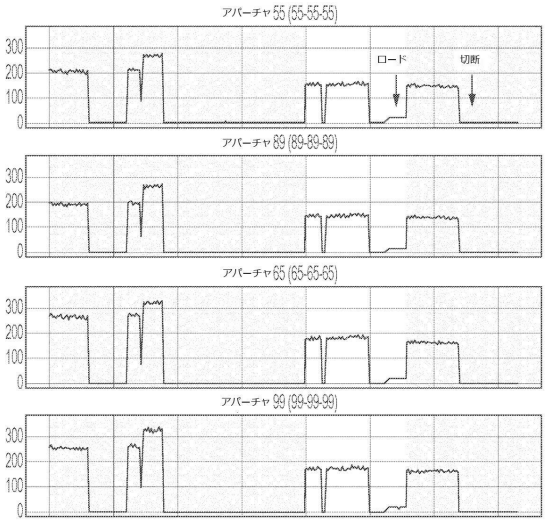


FIG. 17A
CONTINUED

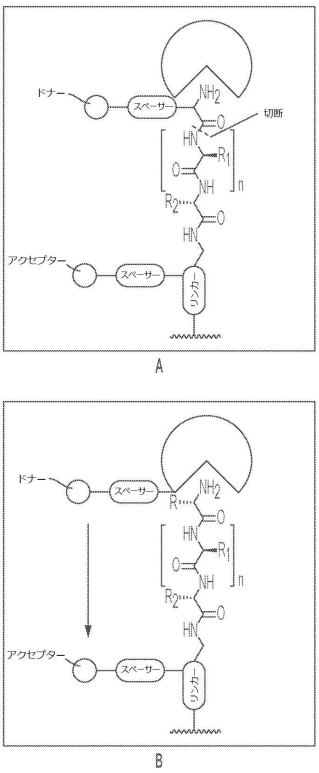
【図１７Ｂ】



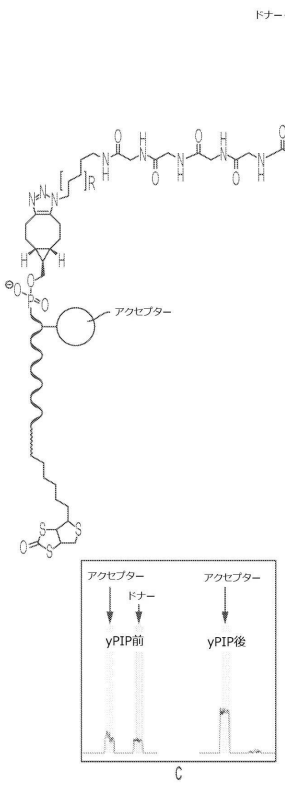
【図１７Ｃ】



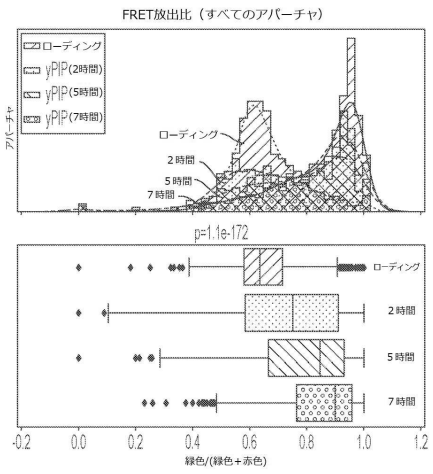
【図１８Ａ－１】



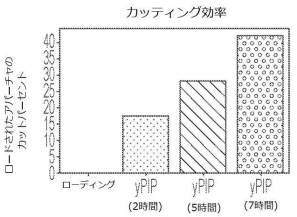
【図１８Ａ－２】



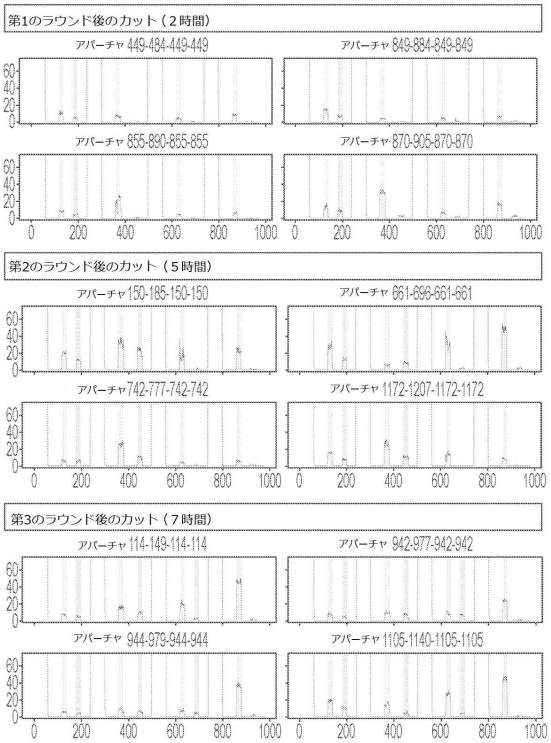
【図１８Ｂ】



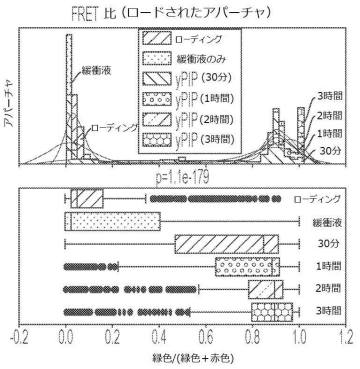
【図１８Ｃ】



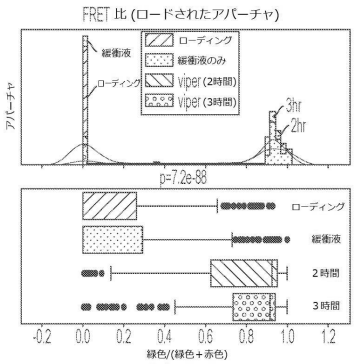
【図１８Ｄ】



【図１８Ｅ】



【図１８Ｆ】



【図１９Ａ】

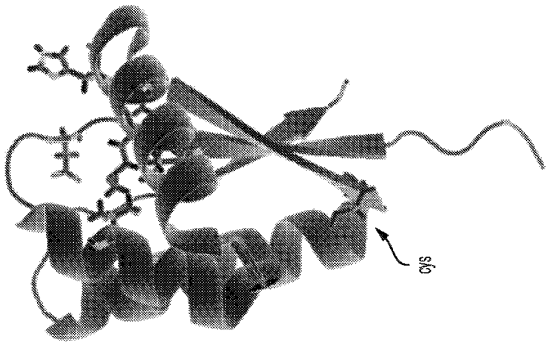


FIG. 19A

【図１９Ｂ－１】

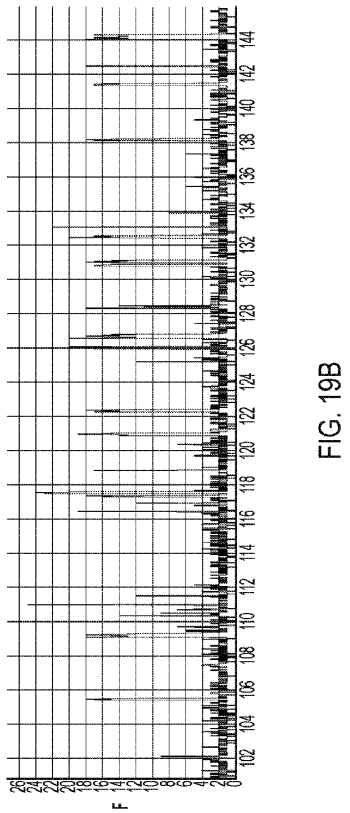


FIG. 19B

【図１９Ｂ－２】

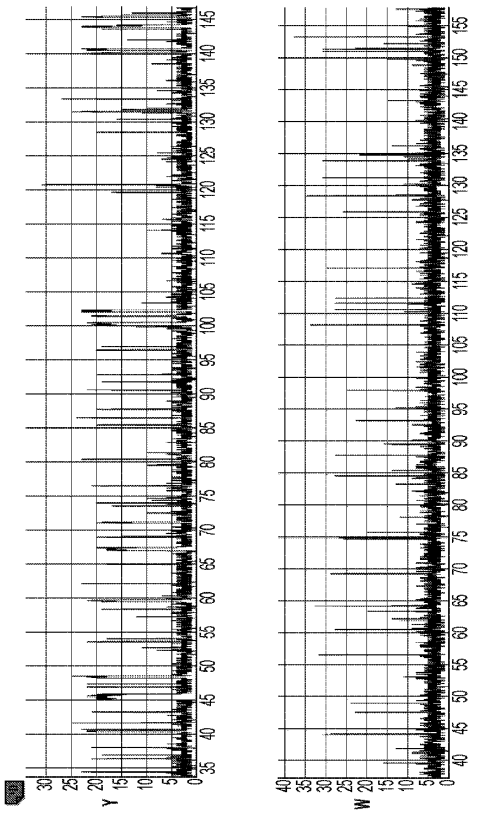
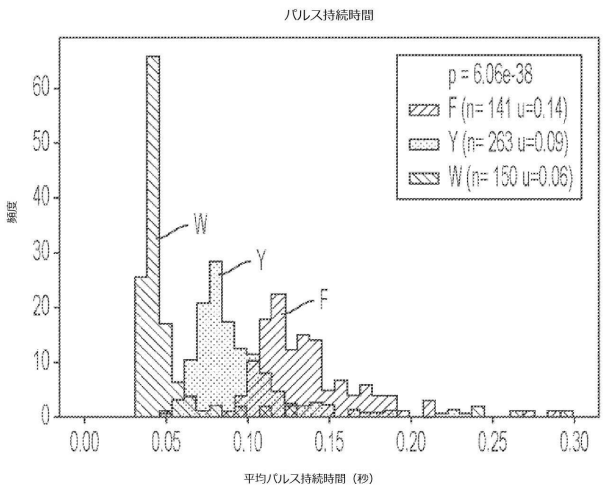
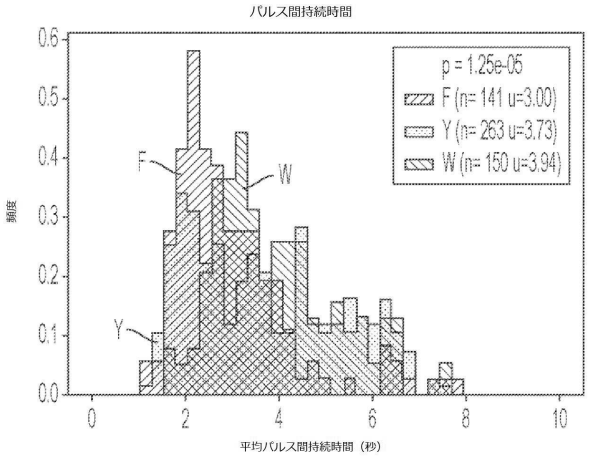


FIG. 19B
CONTINUED

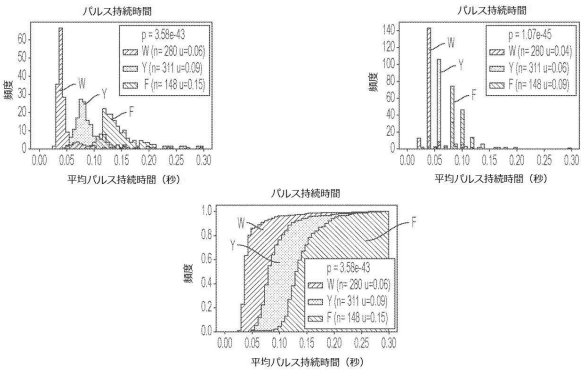
【図１９Ｃ】



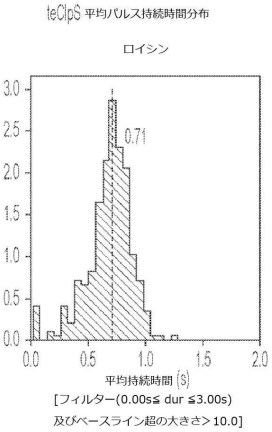
【図１９Ｄ】



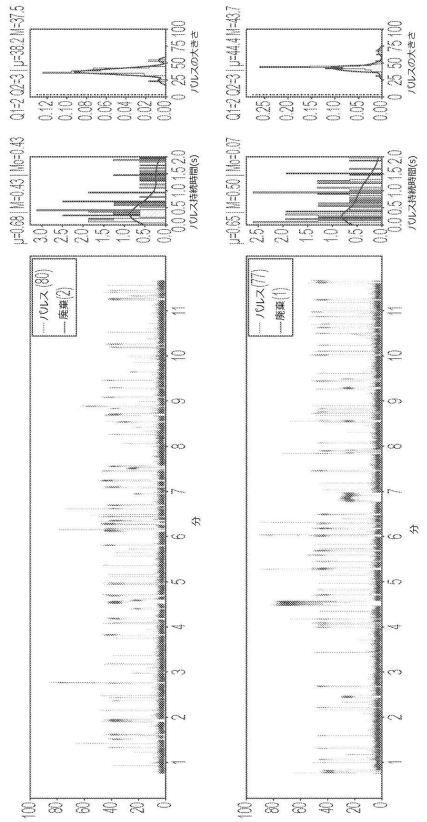
【図１９Ｅ】



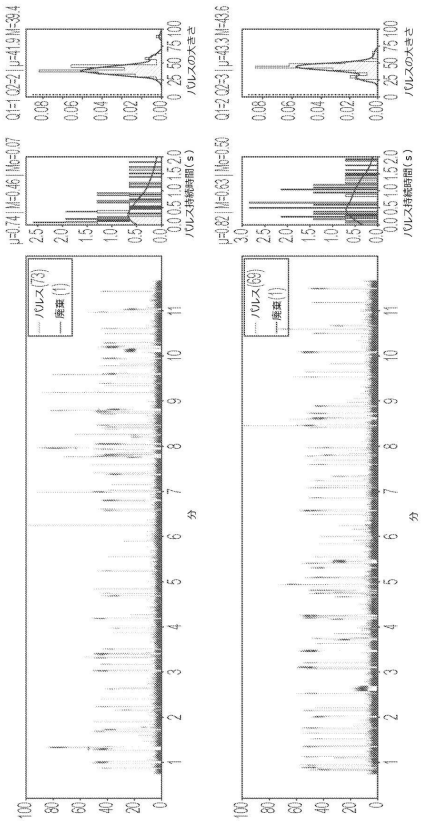
【図１９Ｆ】



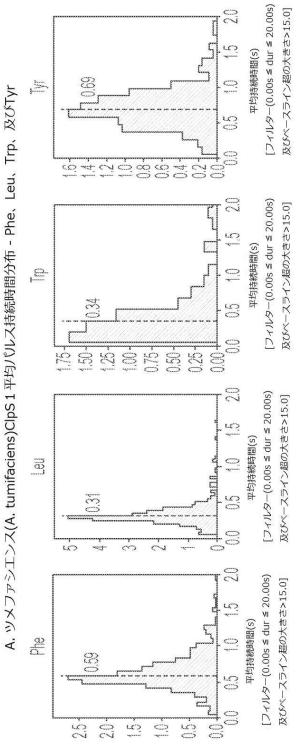
【図１９Ｇ】



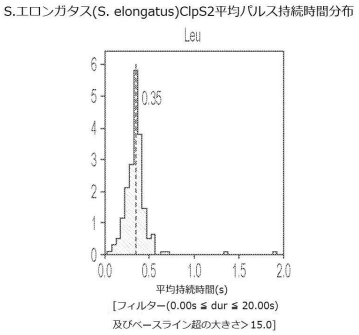
【図１９Ｈ】



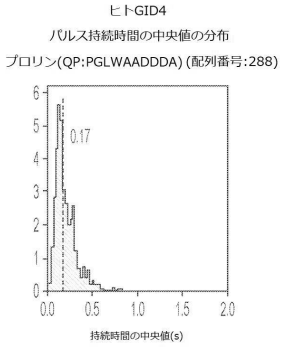
【図１９Ｉ】



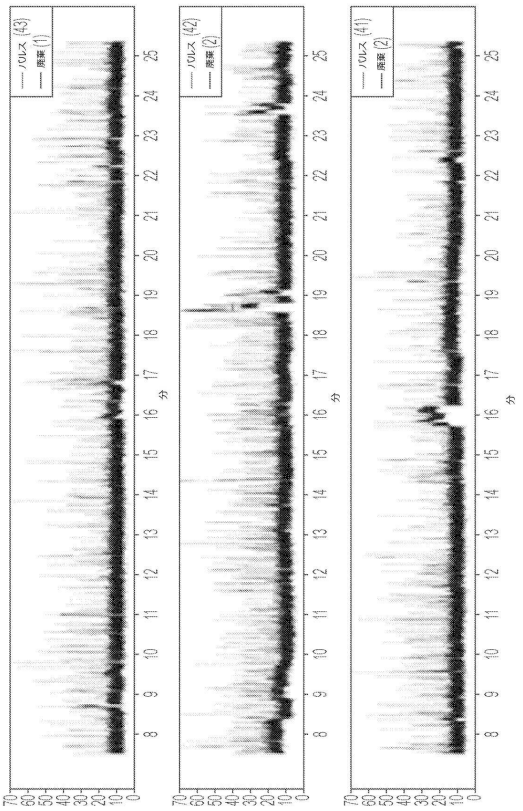
【図１９Ｊ】



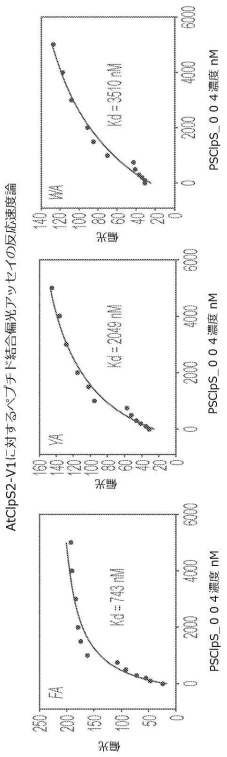
【図１９Ｋ】



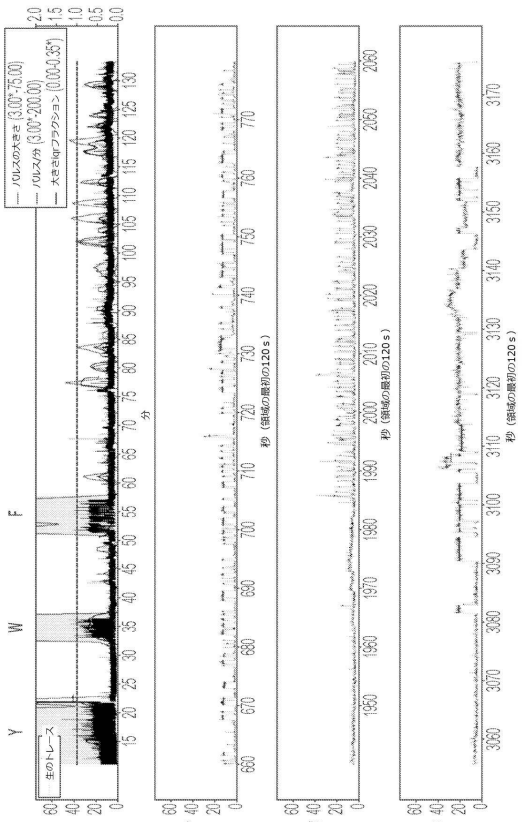
【図１９Ｌ】



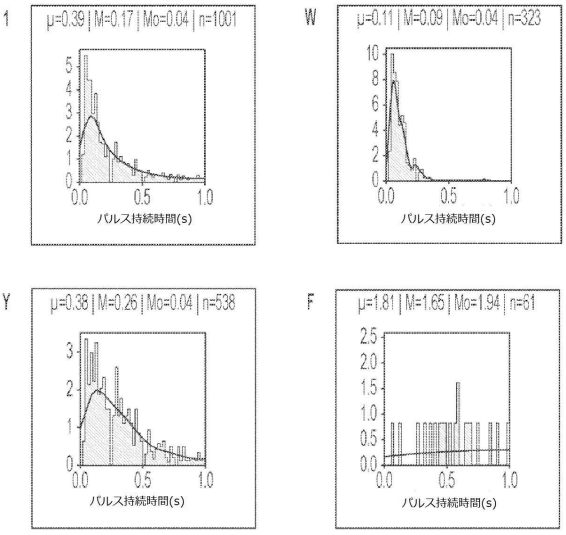
【図１９Ｍ】



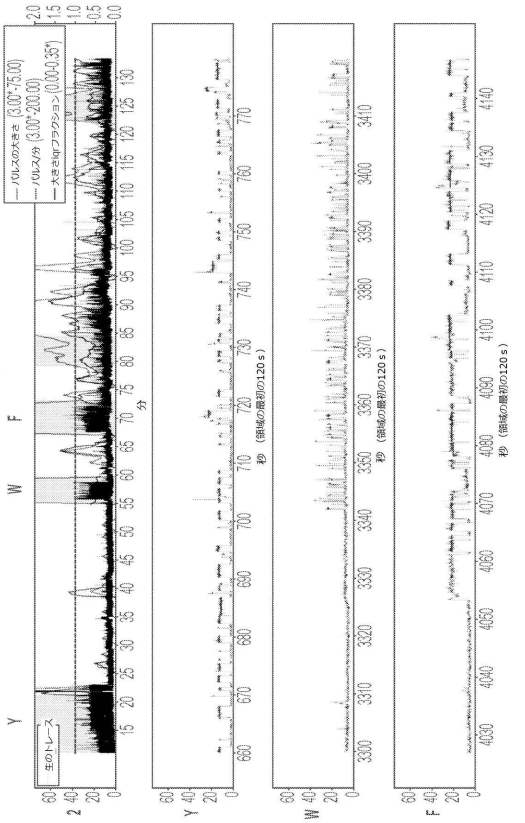
【図２０Ａ】



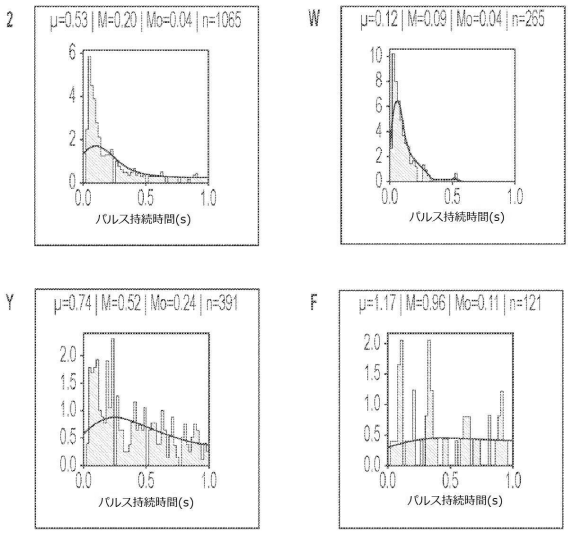
【図２０Ｂ】



【図２０Ｃ】



【図２０Ｄ】



【図２１Ａ】

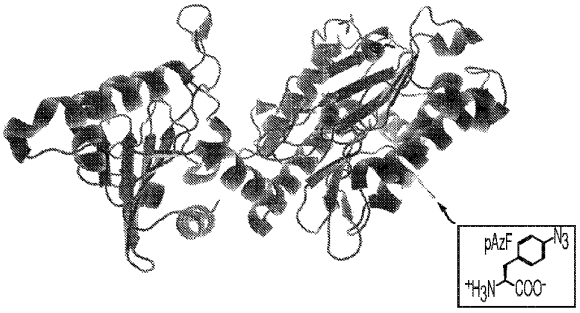
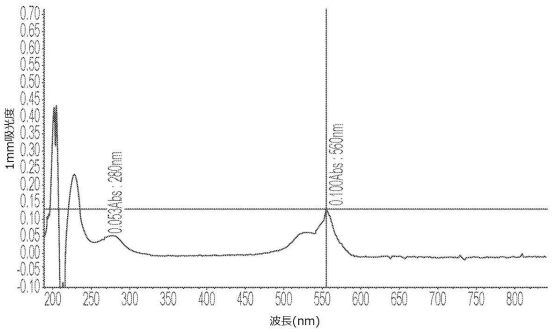
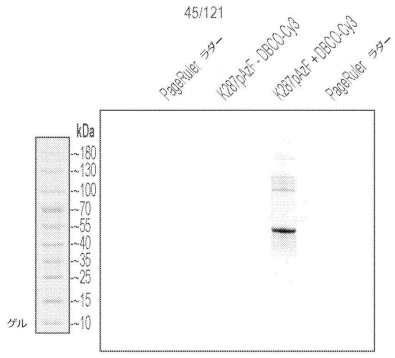


FIG. 21A

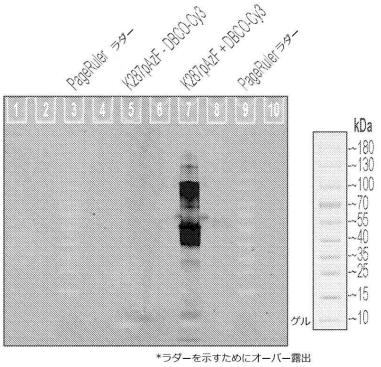
【図２１Ｂ】



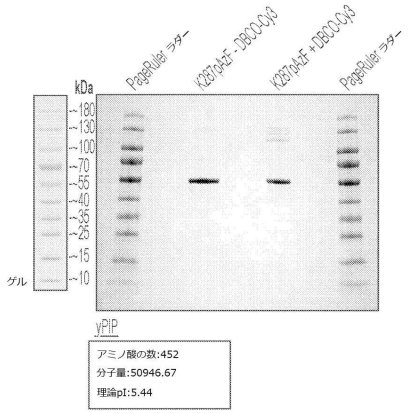
【図２１Ｃ】



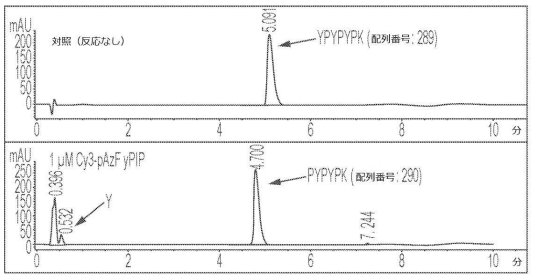
【図２１Ｄ】



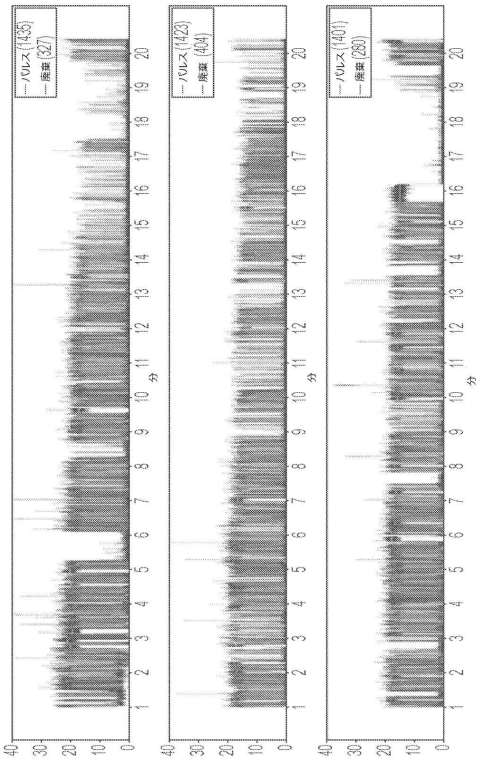
【図２１Ｅ】



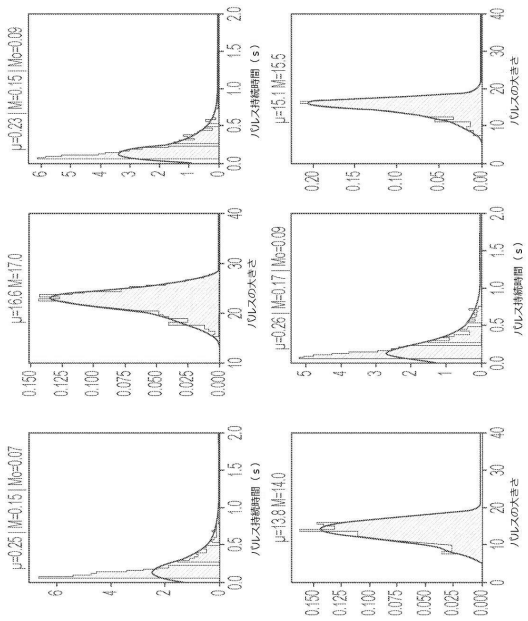
【図２１Ｆ】



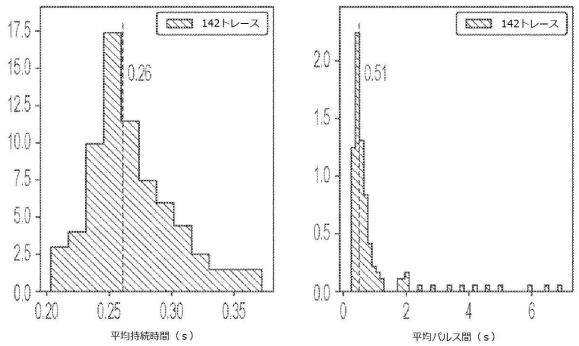
【図２２Ａ】



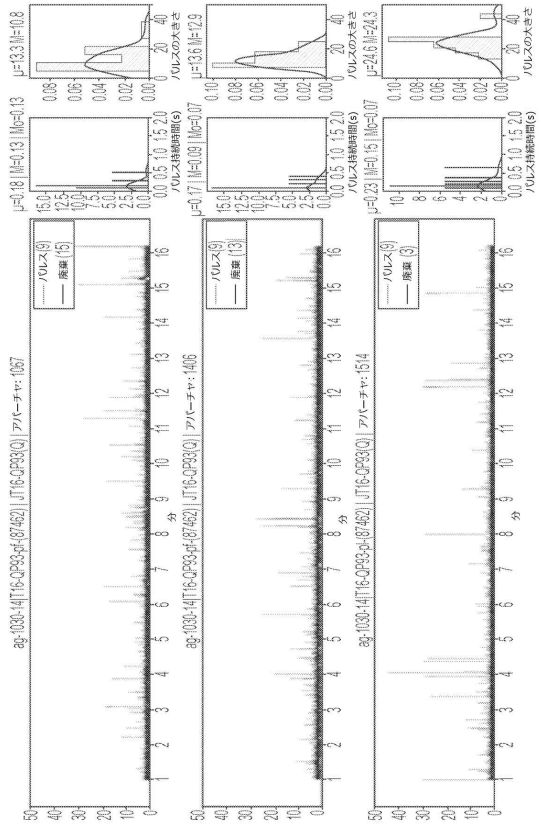
【図２２Ｂ】



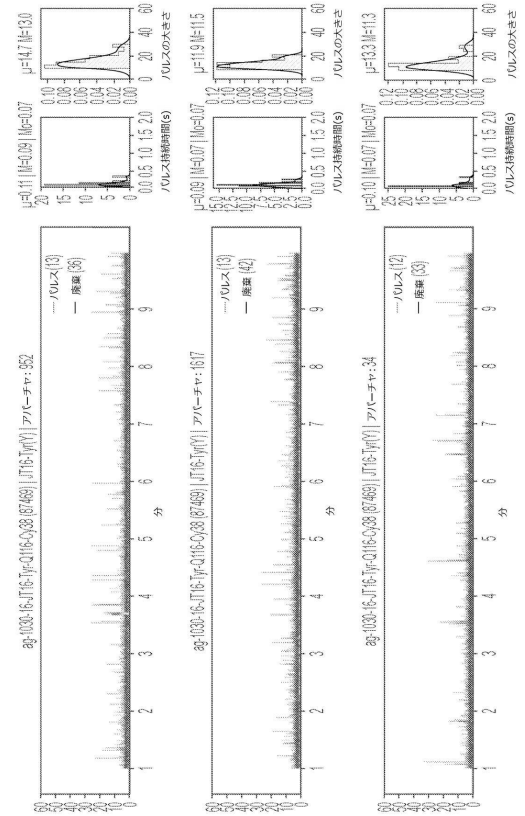
【図２２Ｃ】



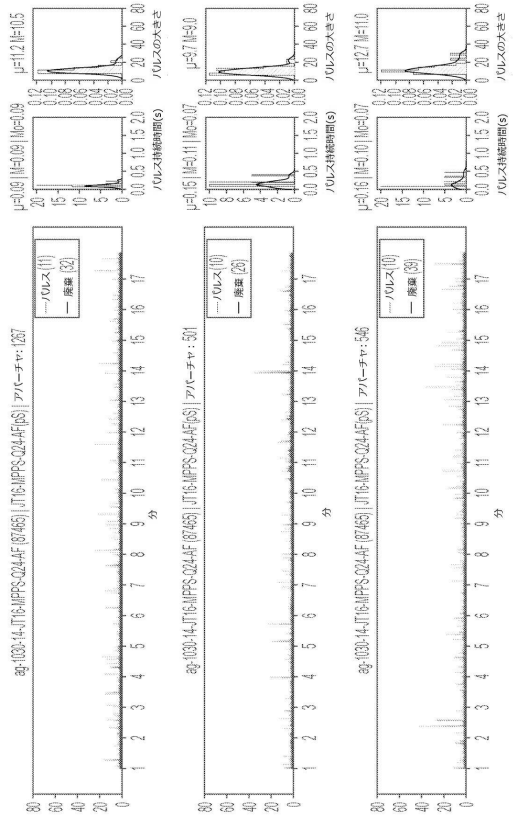
【図２２Ｄ】



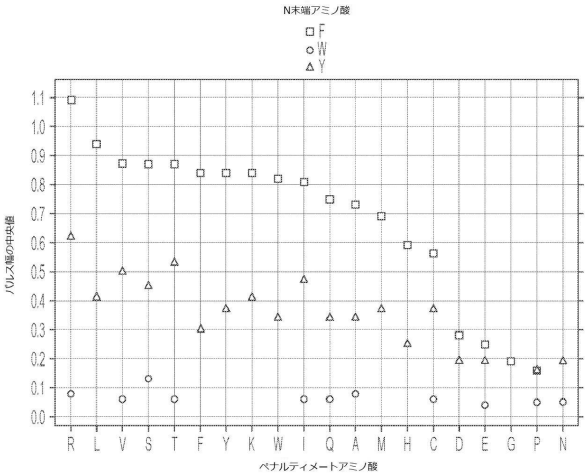
【図２２Ｅ】



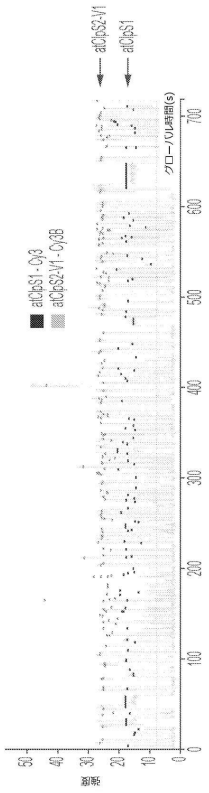
【図２２Ｆ】



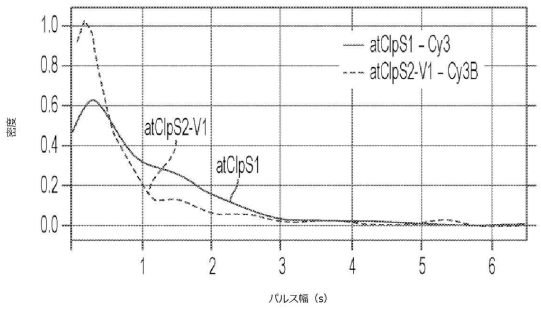
【図２３】



【図２４Ａ】



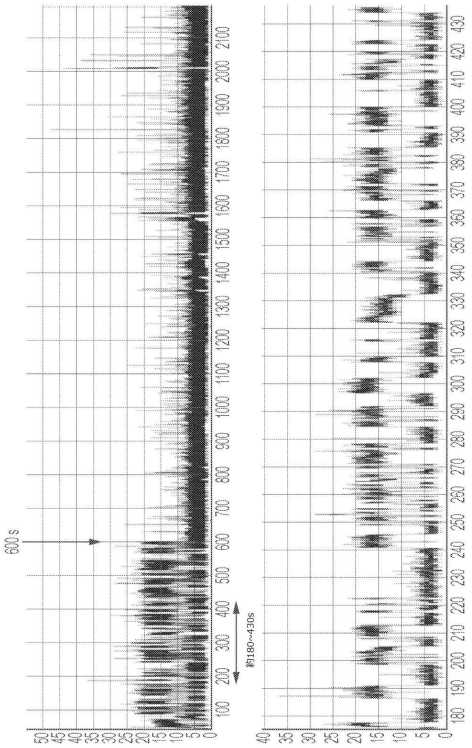
【図２４Ｂ】



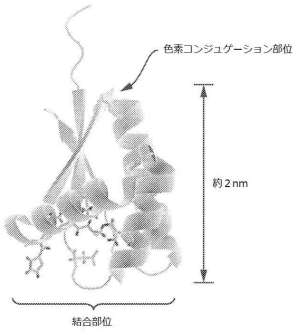
【図２４Ｃ】

	N パルス	平均パルス幅(s)	平均パルス幅(s)	パルス速度(パルス/分)
atClpS1	98	1.3	0.63	8.1
atClpS2-V1	169	1.0	0.35	14.1

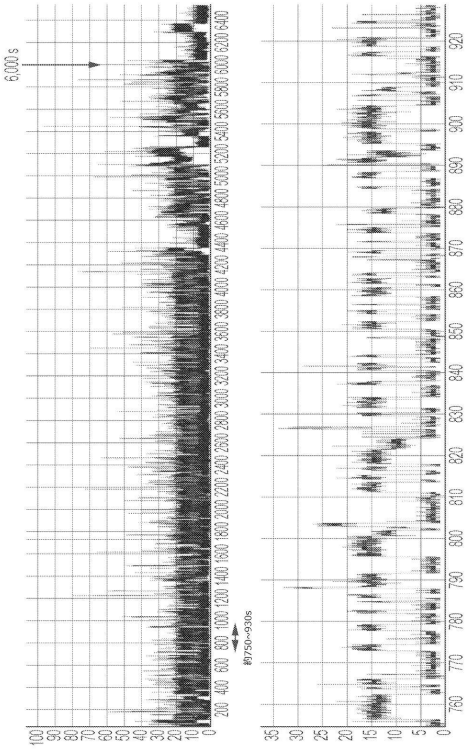
【図２５Ａ】



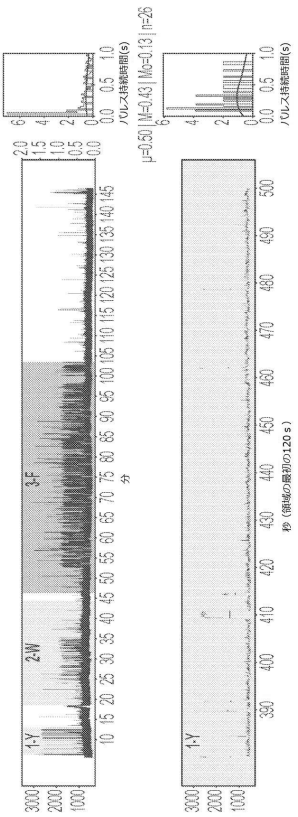
【図２５Ｂ】



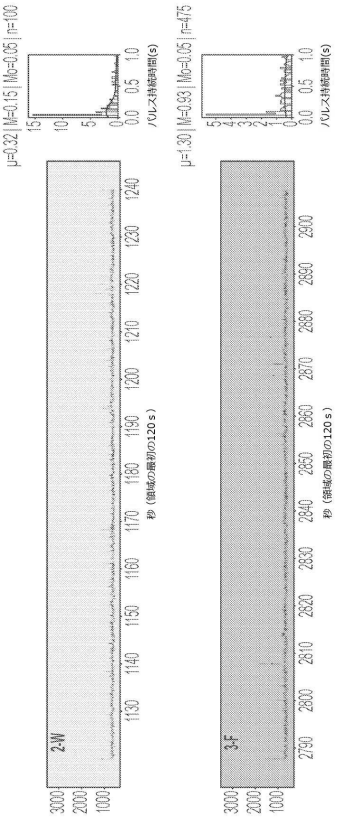
【図２５Ｃ】



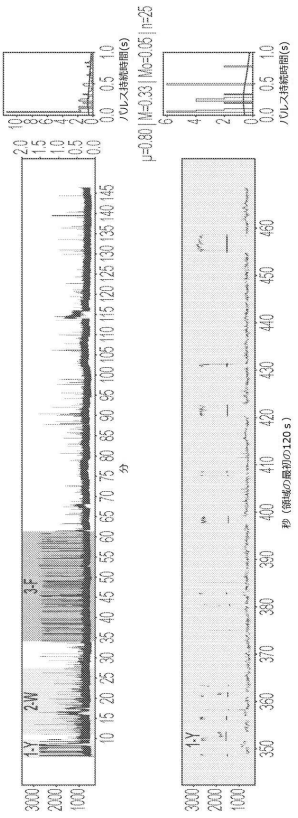
【図２６Ａ－１】



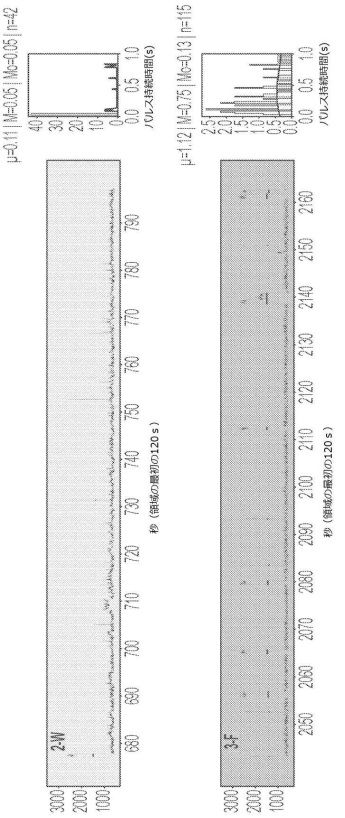
【図２６Ａ－２】



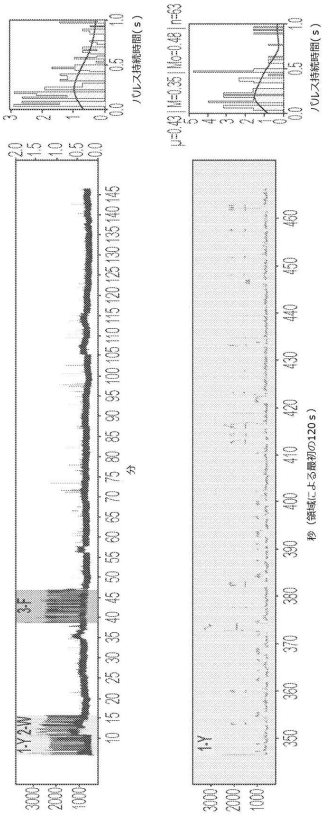
【図２６Ｂ－１】



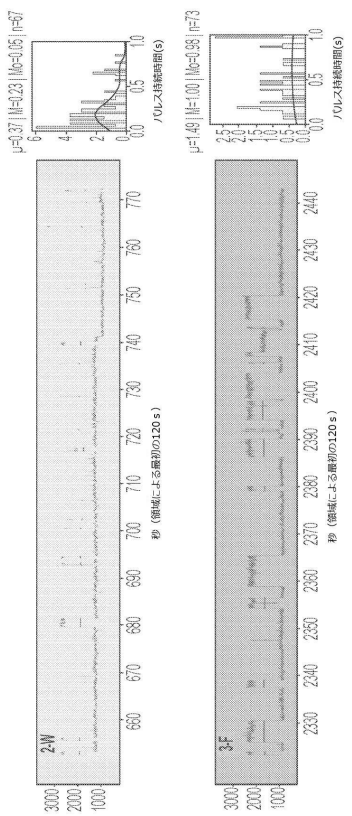
【図２６Ｂ－２】



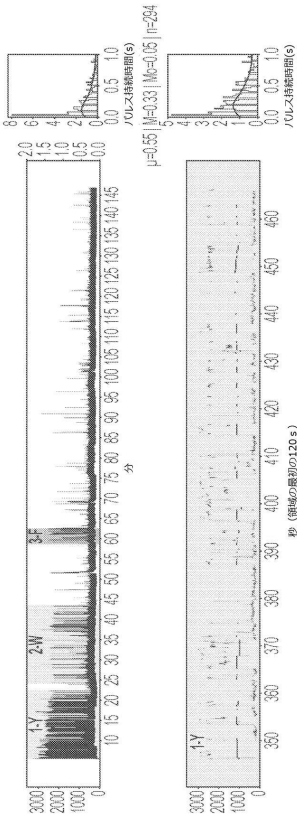
【図２６Ｃ－１】



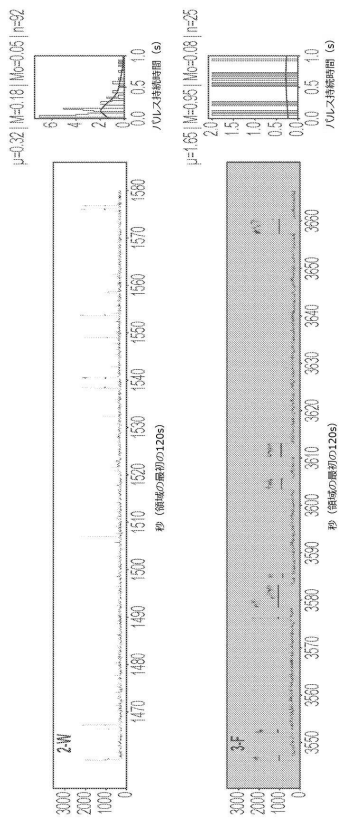
【図２６Ｃ－２】



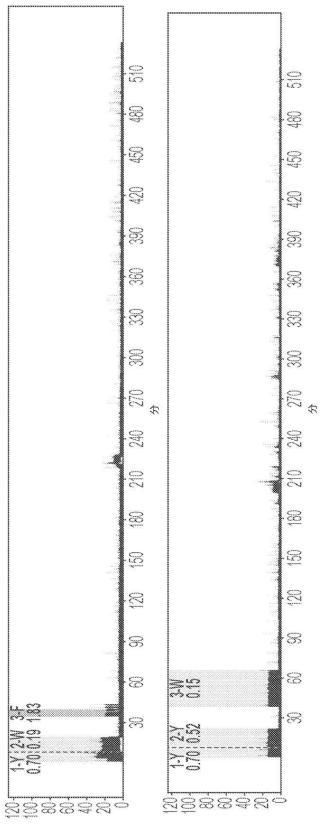
【図２６Ｄ－１】



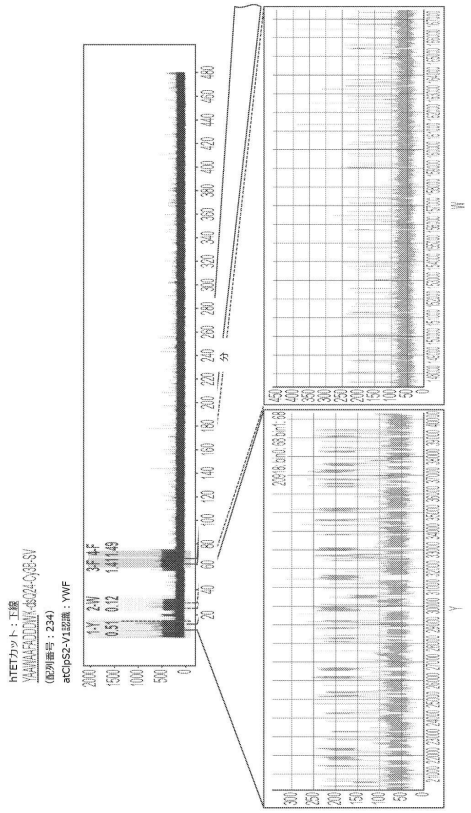
【図２６Ｄ－２】



【図２７】



【図２８Ａ－１】



【図２８Ａ－２】

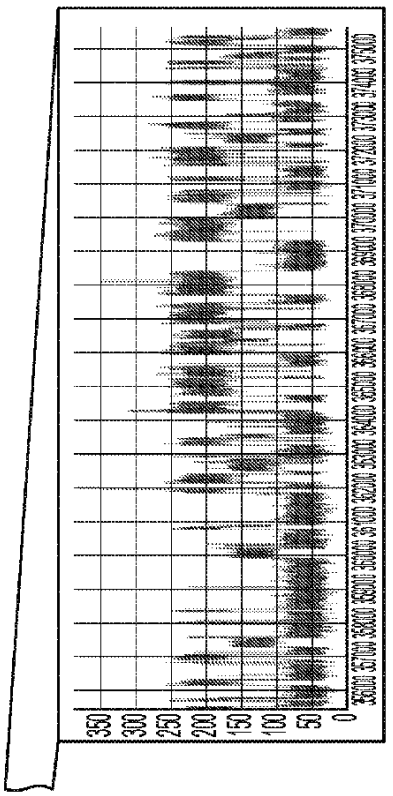


FIG. 28A
CONTINUED

【図２８Ｂ】

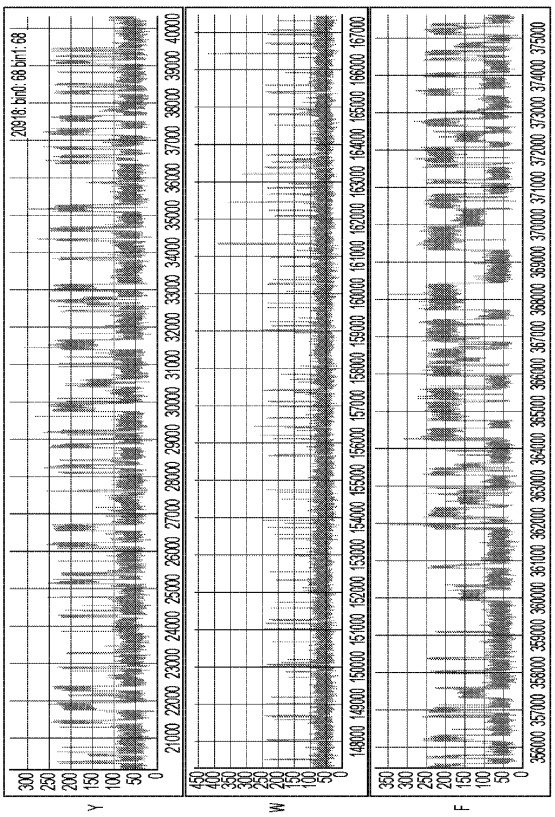
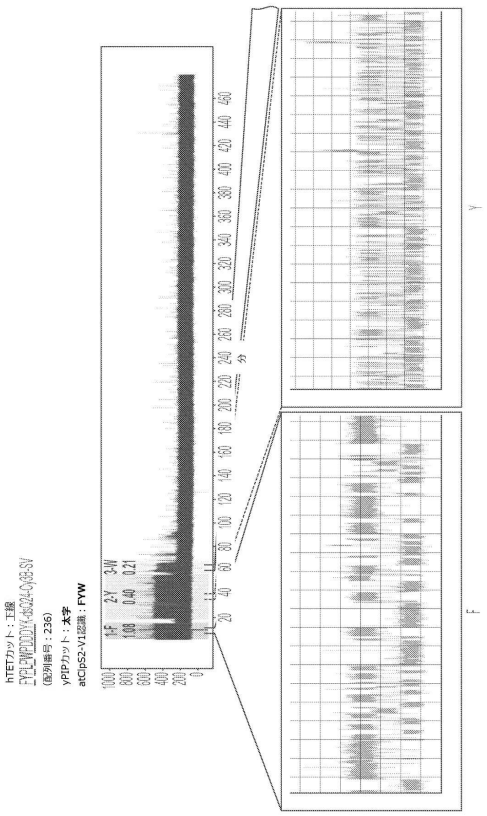


FIG. 28B

【図２８Ｃ－１】



【図２８Ｃ－２】

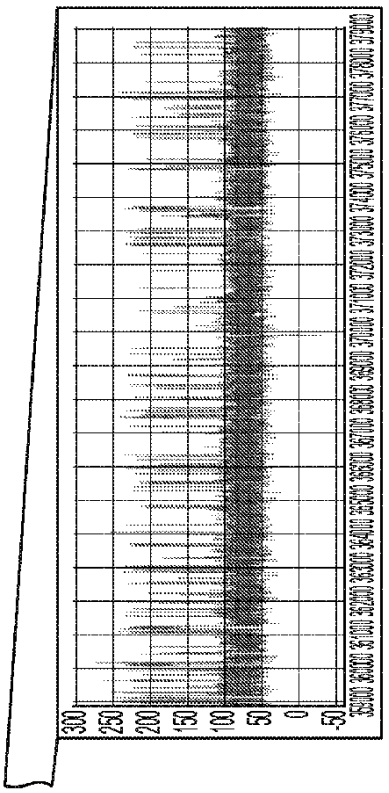


FIG. 28C
CONTINUED

【図２８Ｄ】

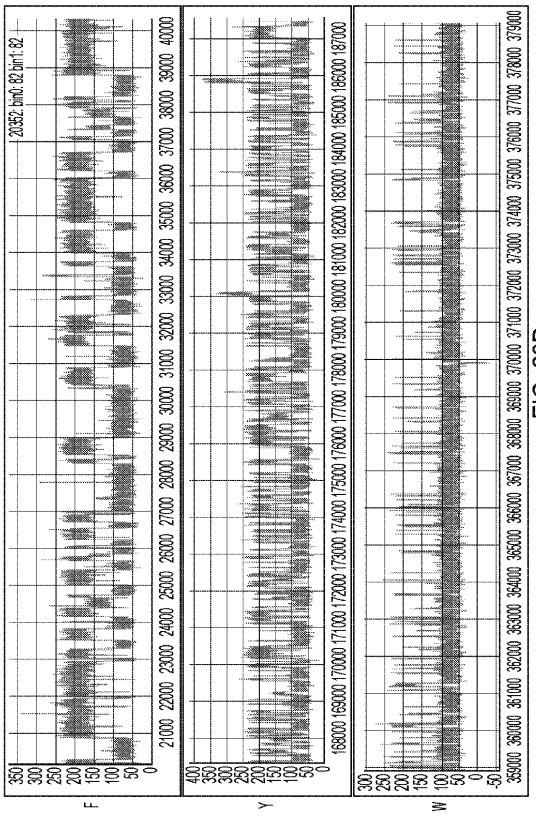
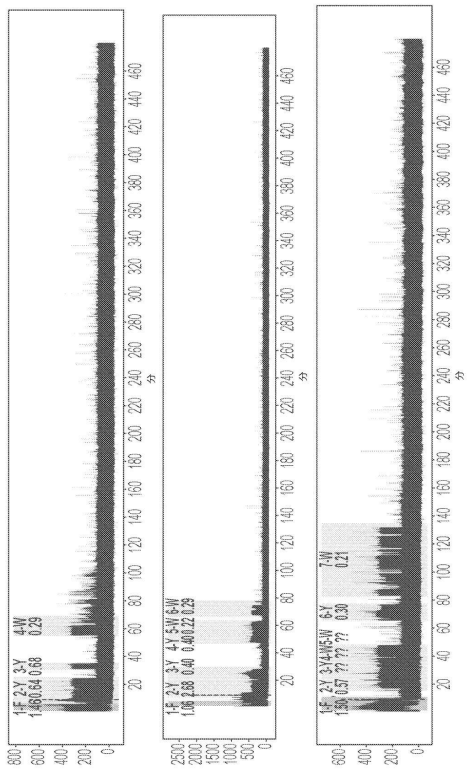
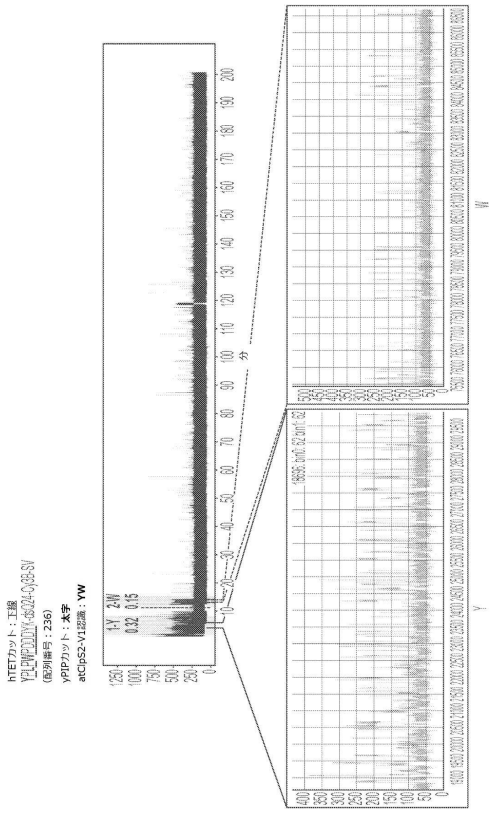


FIG. 28D

【図２８Ｅ】



【図２８Ｆ】



【図２８Ｇ】

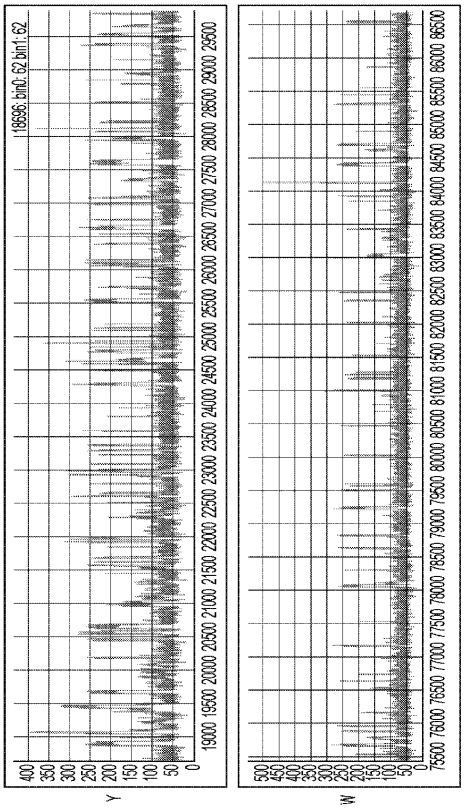
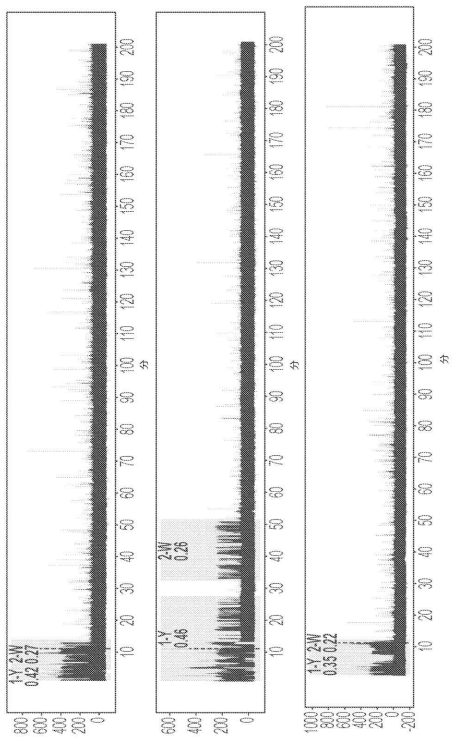
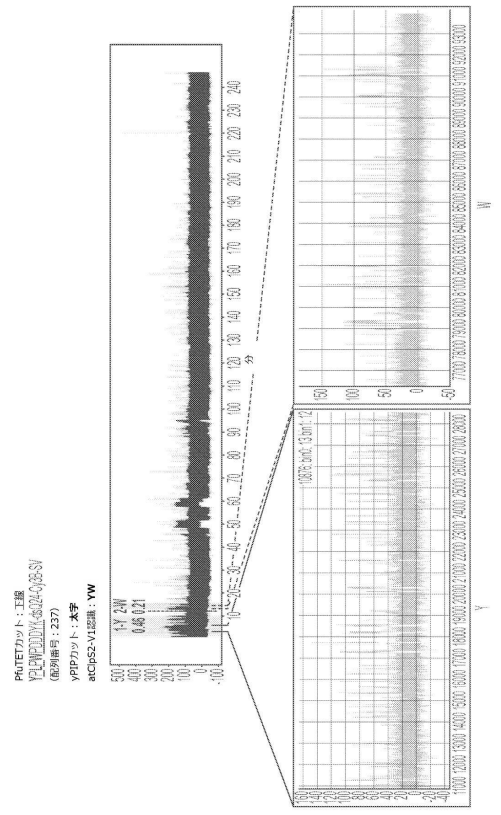


FIG. 28G

【図２８Ｈ】



【図２８Ｉ】



【図２８Ｊ】

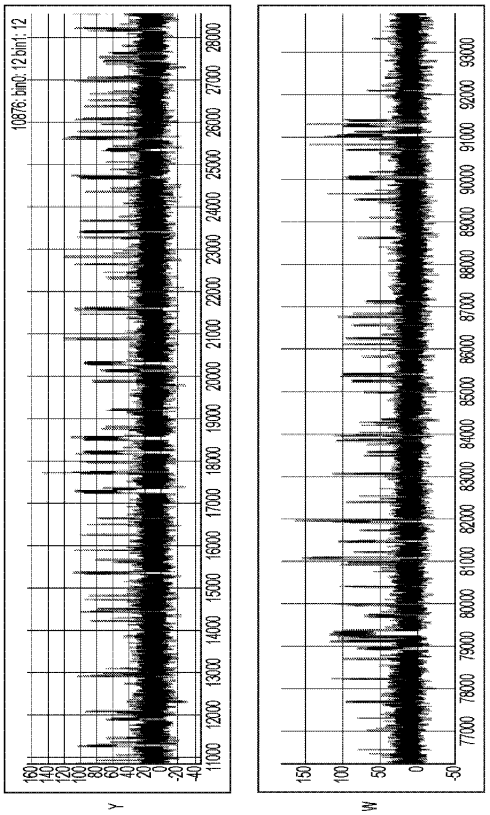
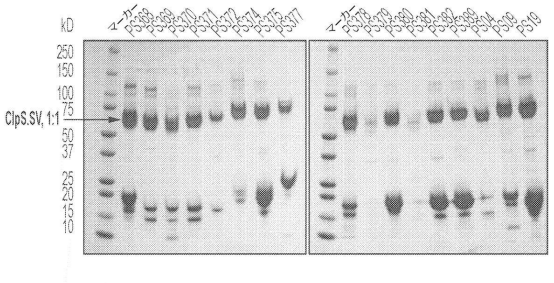
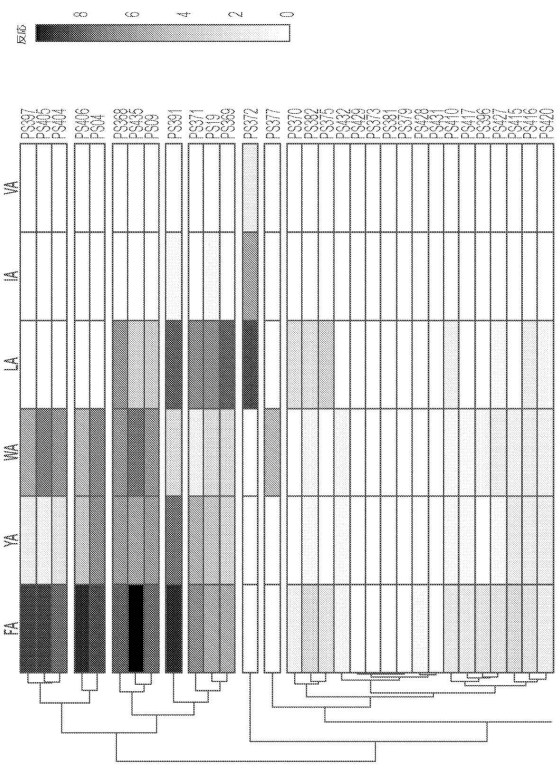


FIG. 28J

【図２９Ａ】



【図２９Ｂ－１】



【図２９Ｂ－２】

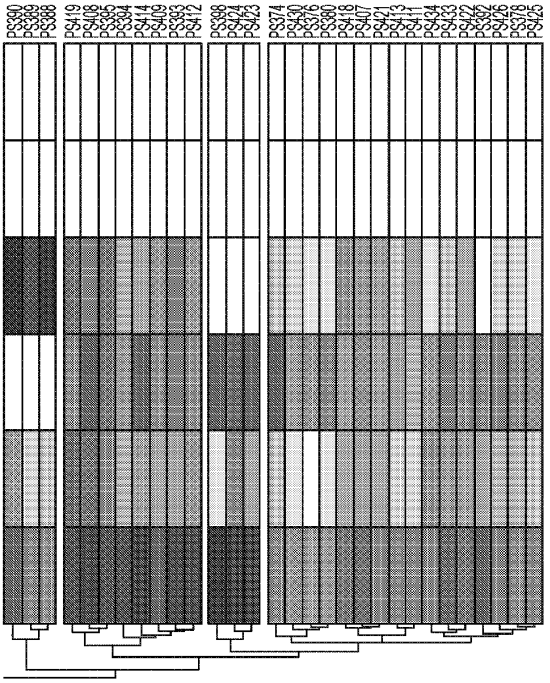
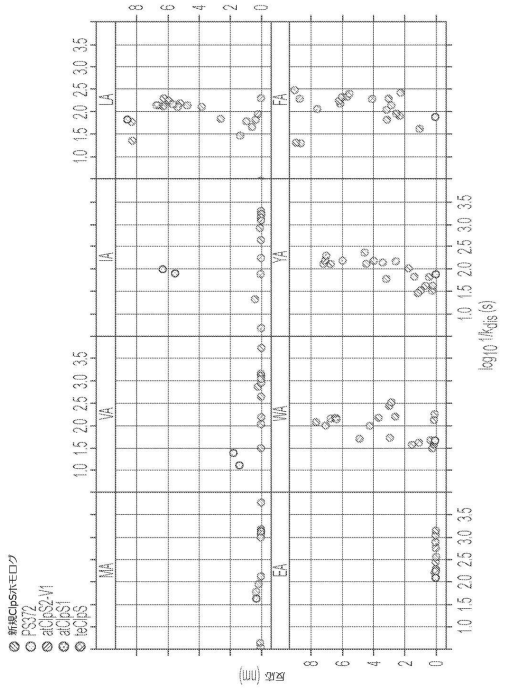
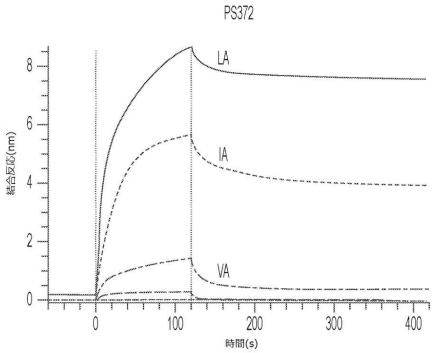


FIG. 29B
CONTINUED

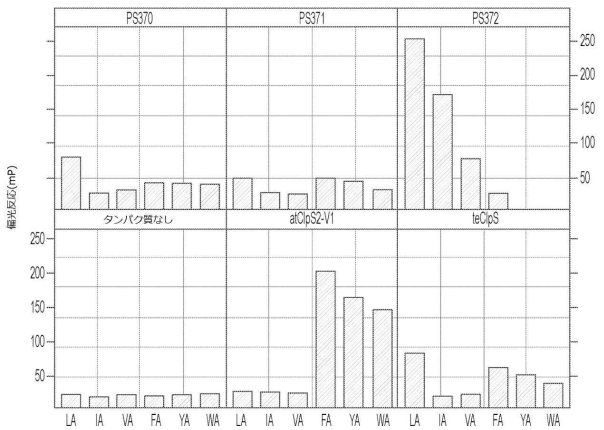
【図２９Ｃ】



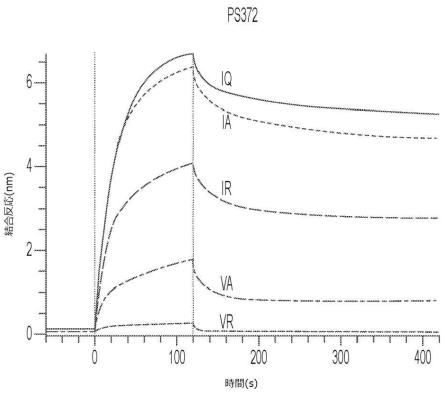
【図２９Ｄ】



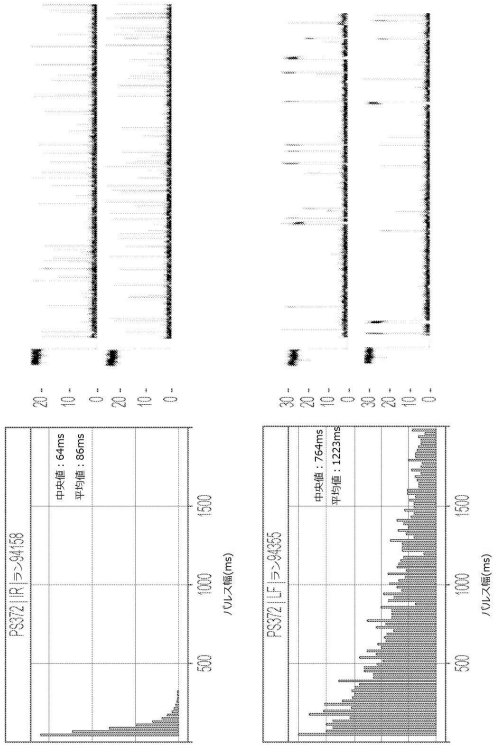
【図２９Ｅ】



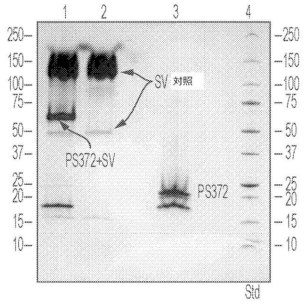
【図２９Ｆ】



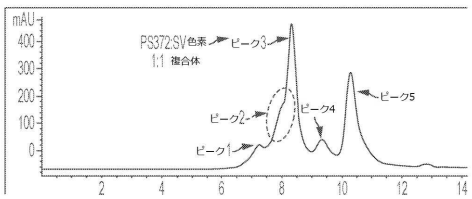
【図２９Ｇ】



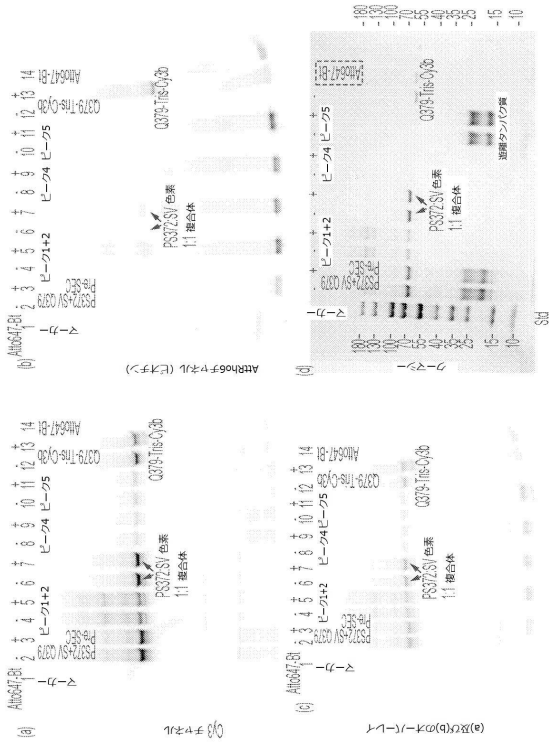
【図３０Ａ】



【図３０Ｂ】



【図３０Ｃ】



【図３０Ｄ】

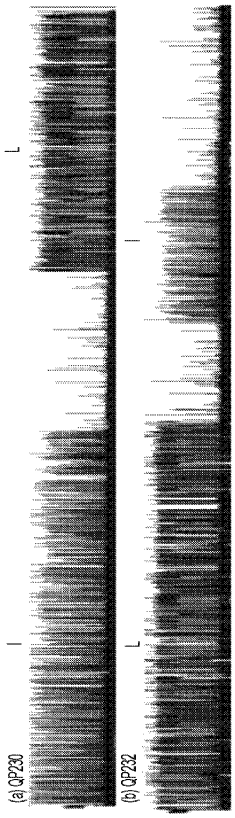
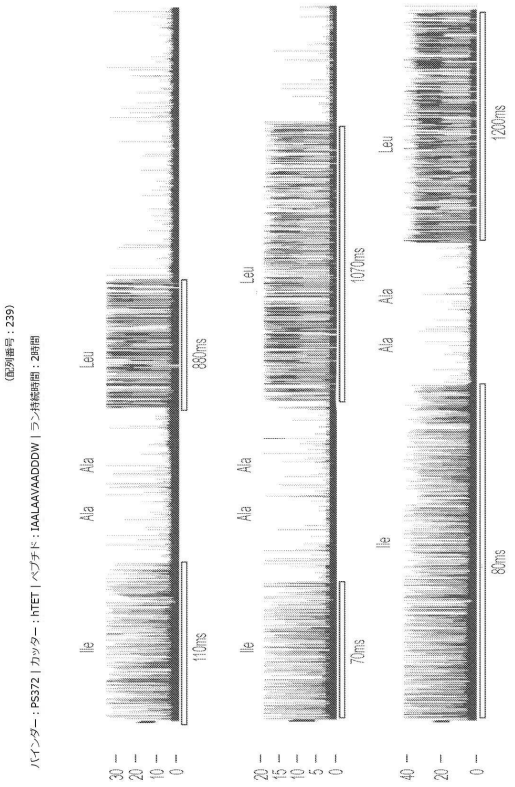
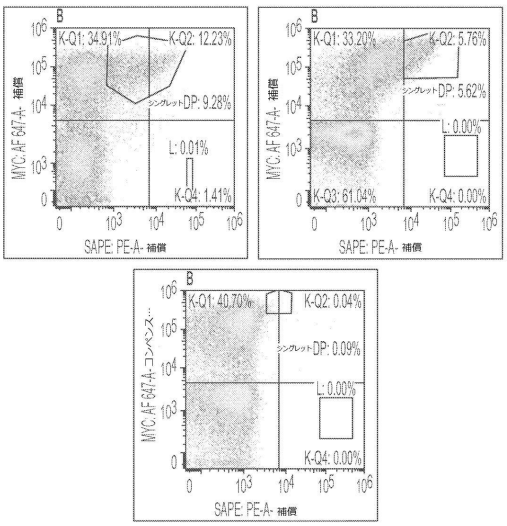


FIG. 30D

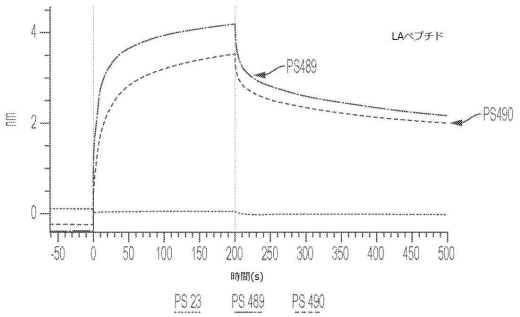
【図３０Ｅ】



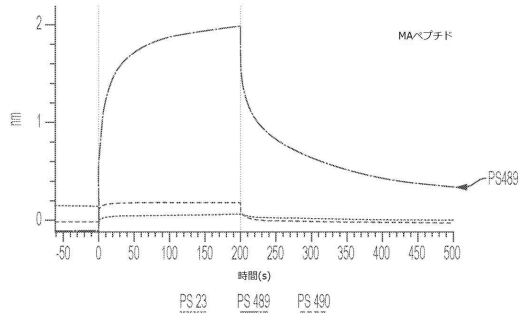
【図３１Ａ】



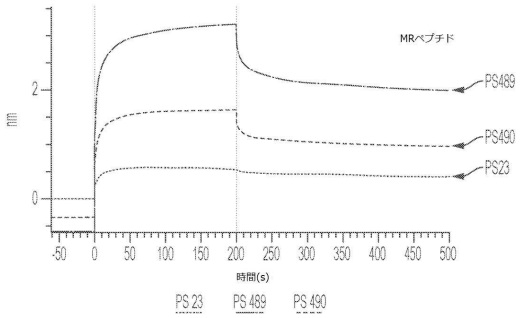
【図３１Ｂ】



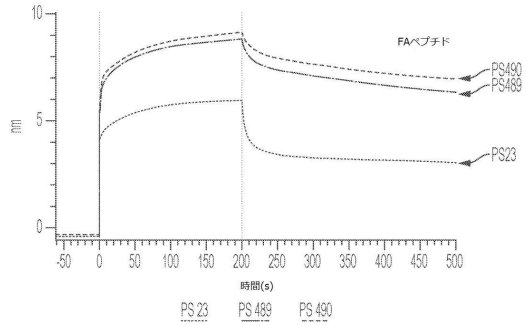
【図３１Ｃ】



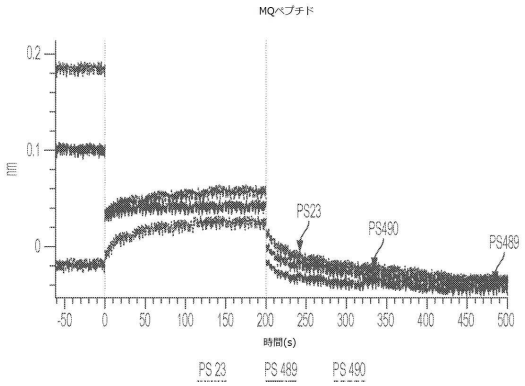
【図３１Ｄ】



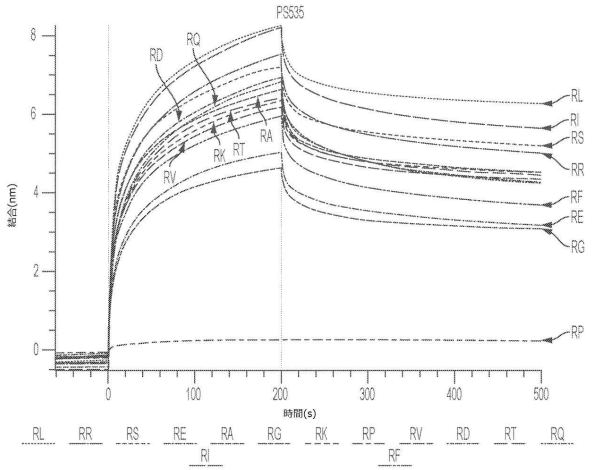
【図３１Ｅ】



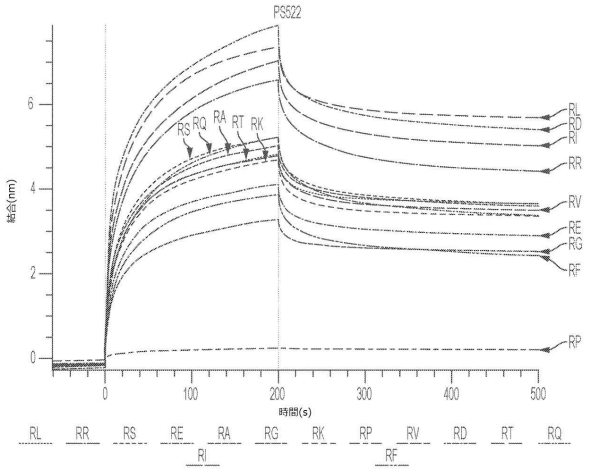
【図３１Ｆ】



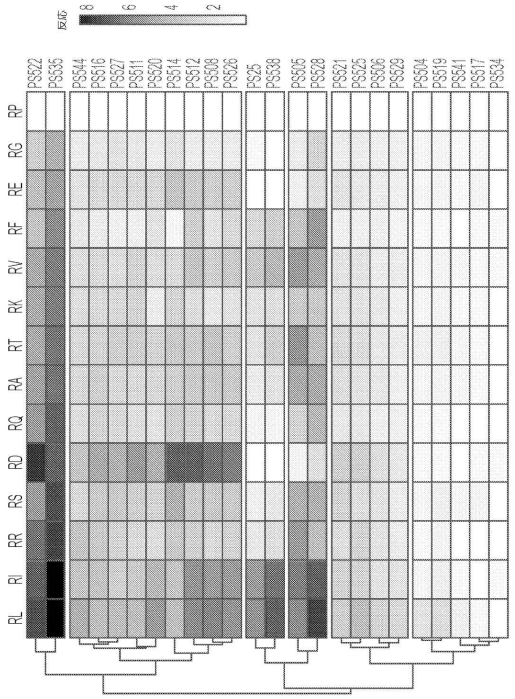
【図３２Ａ】



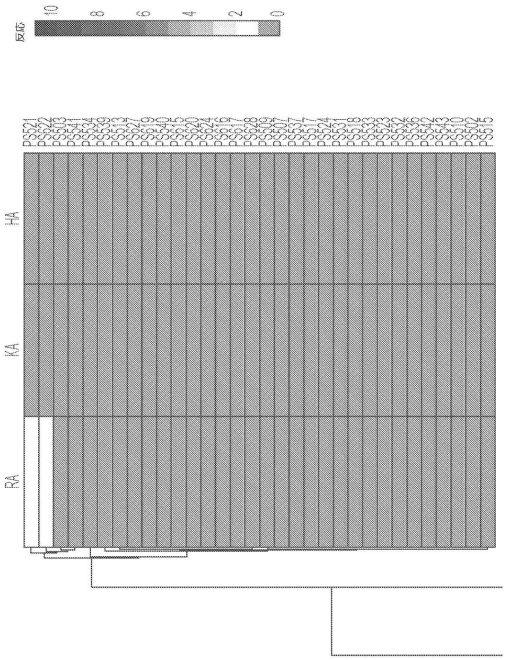
【図３２Ｂ】



【図３２Ｃ】



【図３２Ｄ－１】



【図３２Ｄ－２】

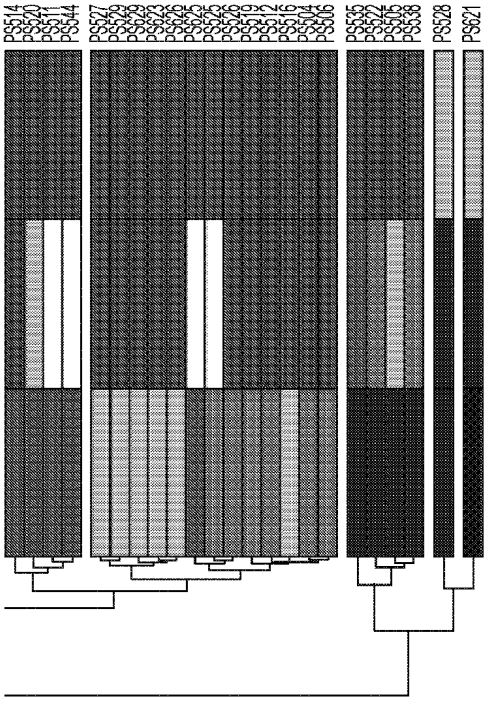


FIG. 32D
CONTINUED

【図３２Ｅ－１】

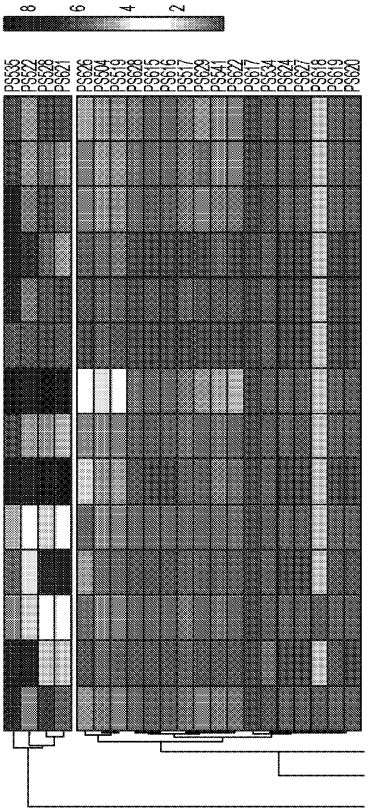


FIG. 32E

【図３２Ｅ－２】

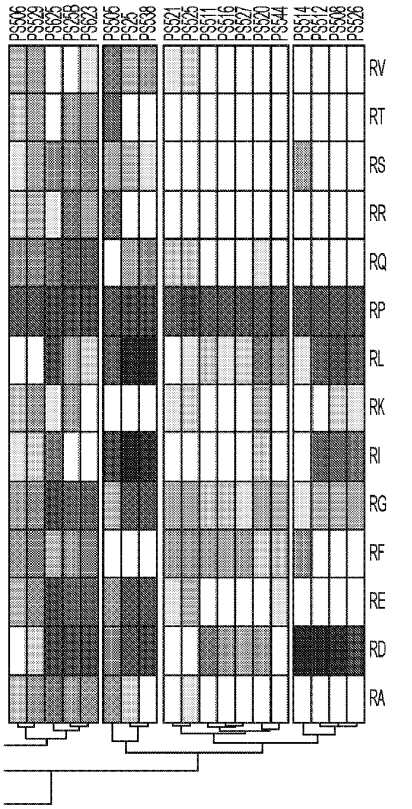
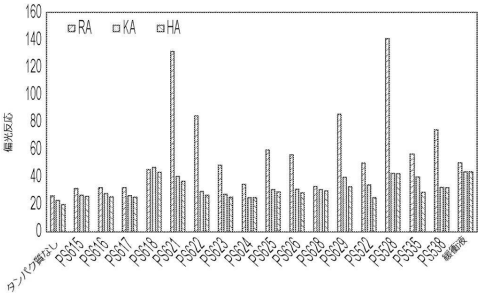
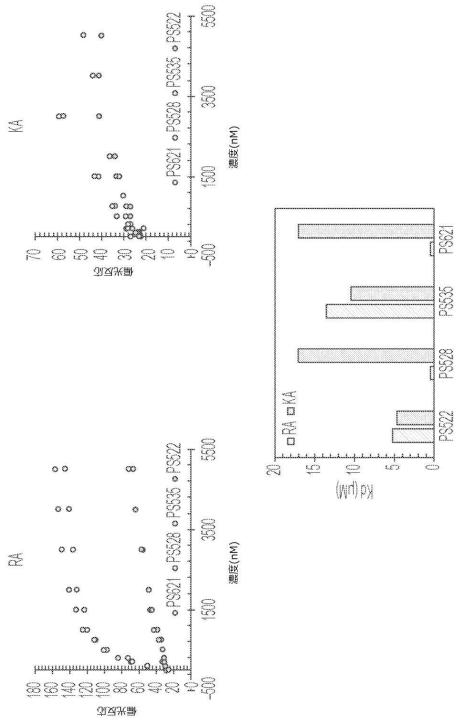


FIG. 32E
CONTINUED

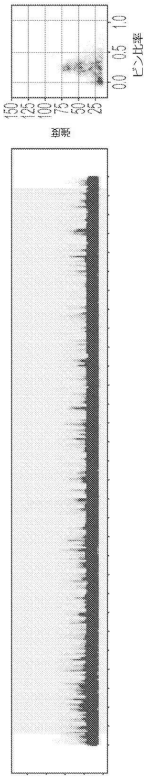
【図３２Ｆ】



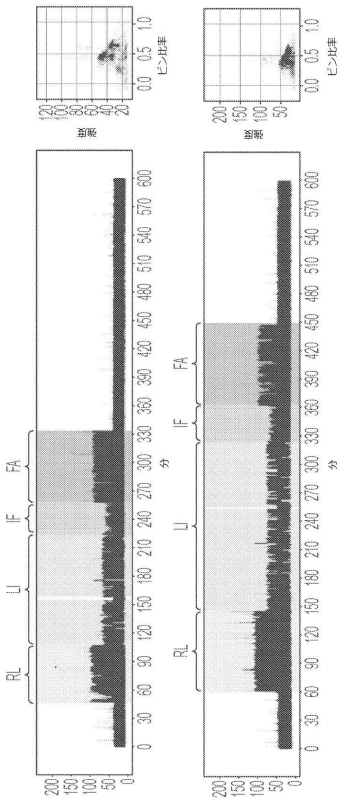
【図３２Ｇ】



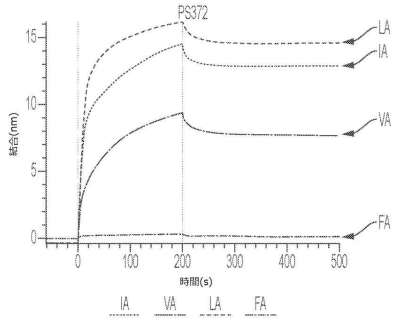
【図３２Ｈ】



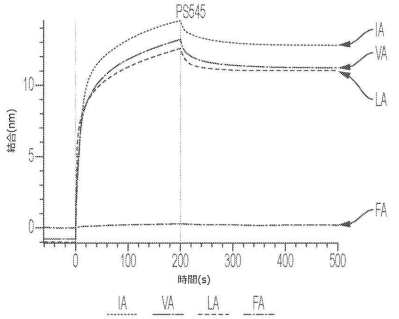
【図３２Ｉ】



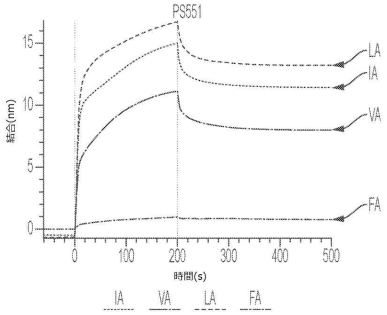
【図３３Ａ】



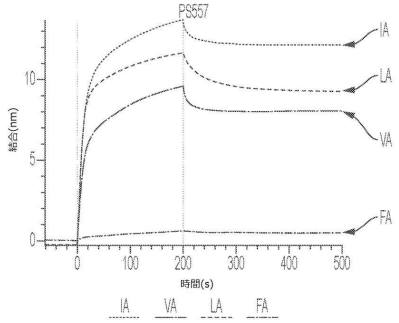
【図３３Ｂ】



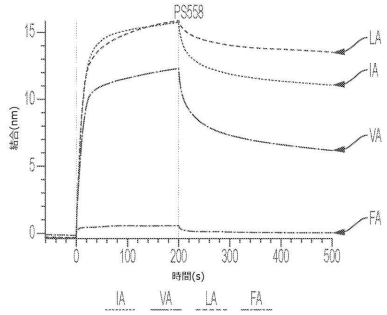
【図３３Ｃ】



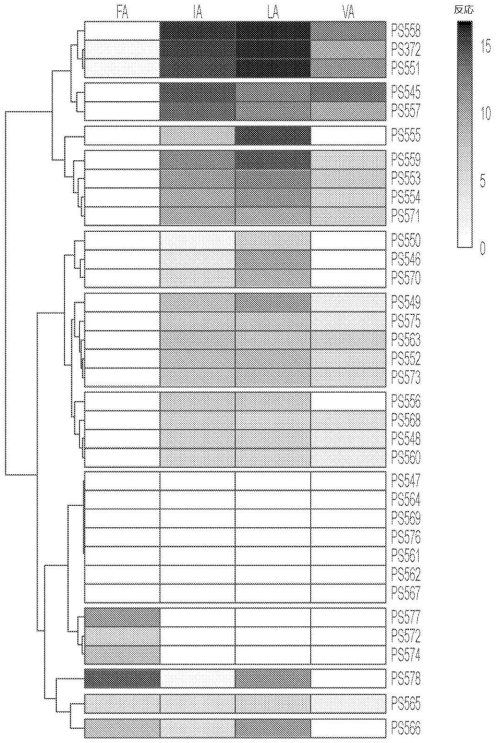
【図３３Ｄ】



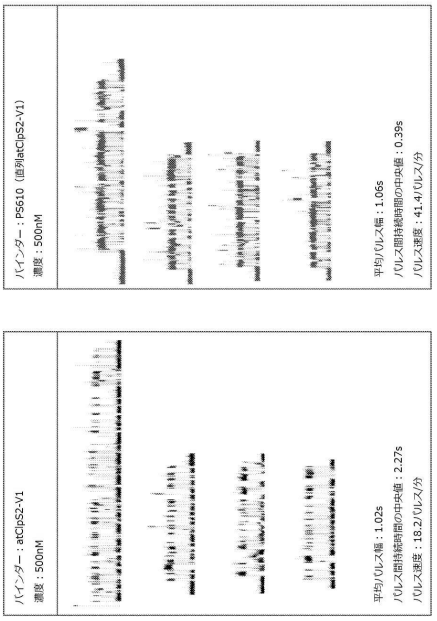
【図３３Ｅ】



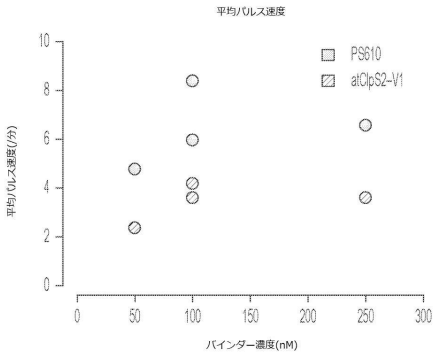
【図３３Ｆ】



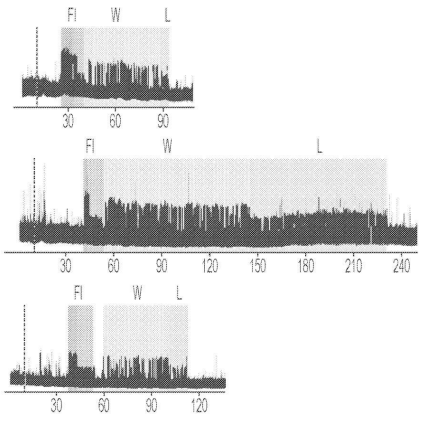
【図３４Ａ】



【図３４Ｂ】



【図３４Ｃ】



ペプチド：GLP1(EFIAWLVK)(配列番号：248)
F/Y/W認識：PS610（直列atClpS2-V1）
L/I認識：PS372
カッター：hTET、pfuTET

【図３４Ｄ】

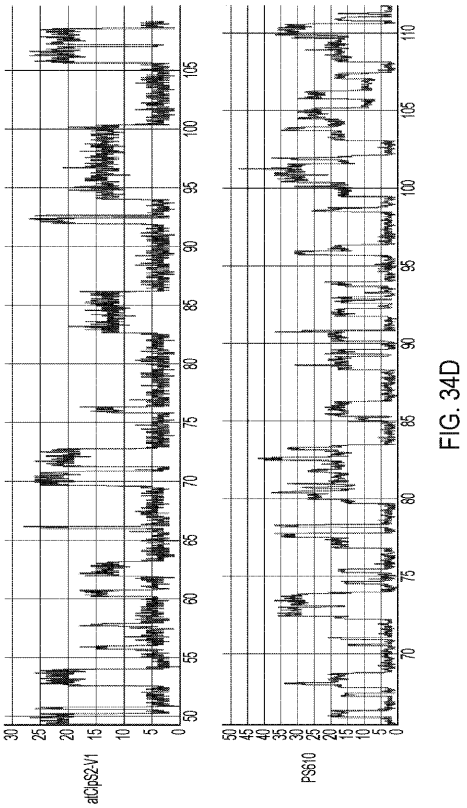
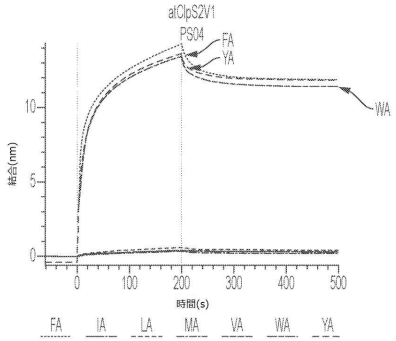
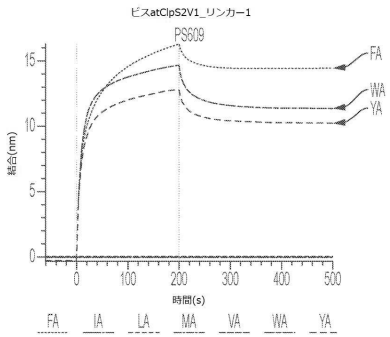


FIG. 34D

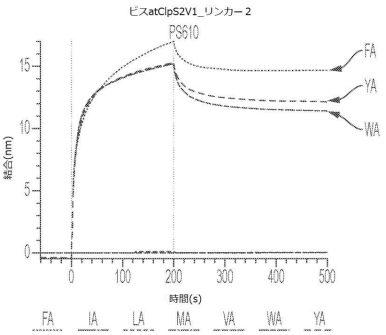
【図３５Ａ】



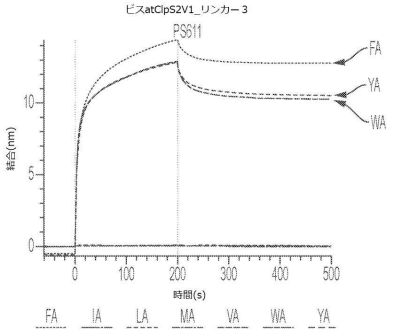
【図３５Ｂ】



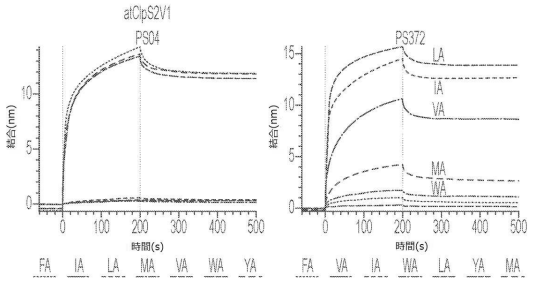
【図３５Ｃ】



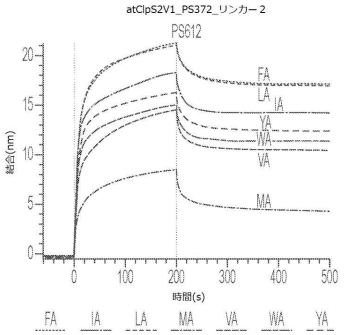
【図３５Ｄ】



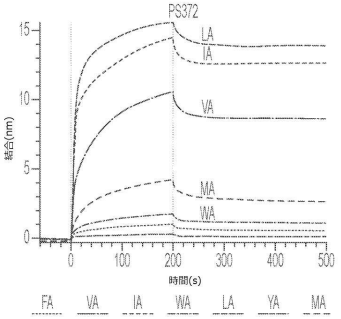
【図３６Ａ】



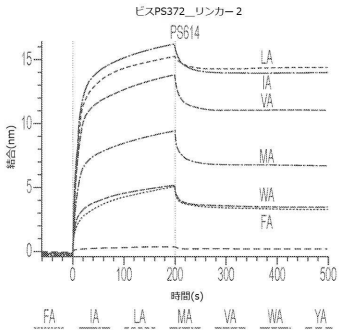
【図３６Ｂ】



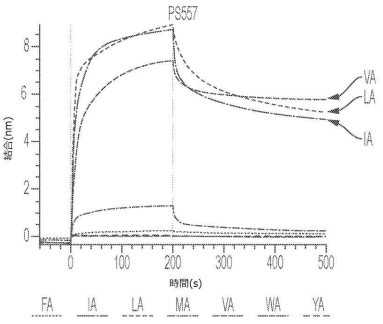
【図３６Ｃ】



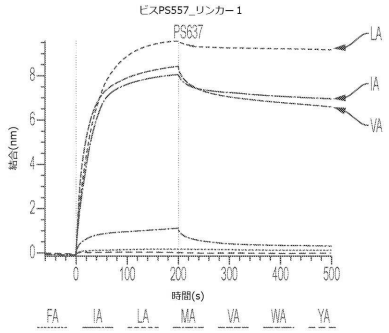
【図３６Ｄ】



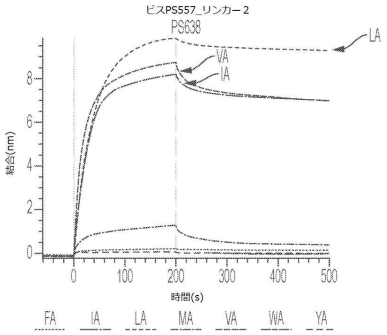
【図３６Ｅ】



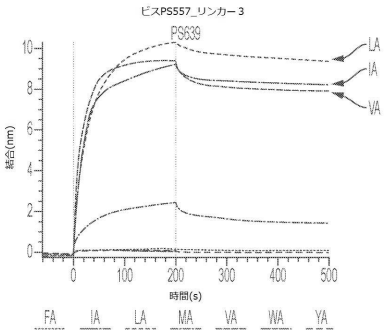
【図３６Ｆ】



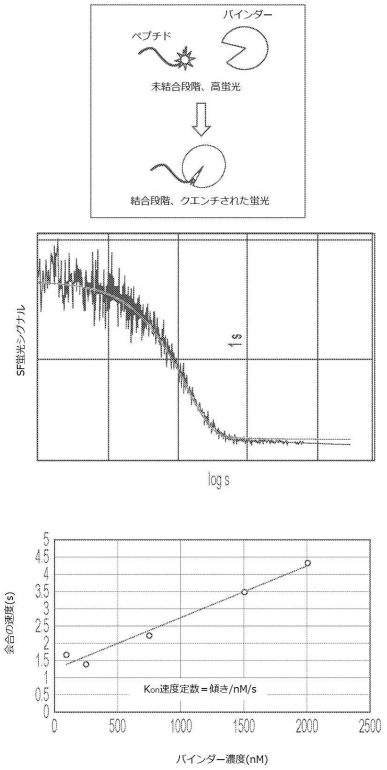
【図３６Ｇ】



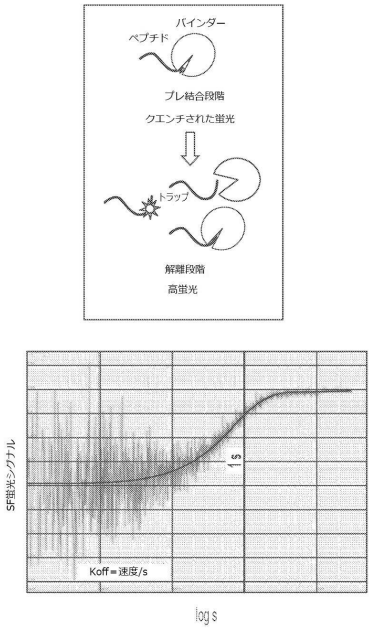
【図３６Ｈ】



【図３７Ａ】



【図３７Ｂ】



【配列表】

２０２３５２７７６４０００００１．ａｐｐ

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2021/033493

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/58 C07K14/00 G01N33/68
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2018/045186 A1 (HARVARD COLLEGE [US]) 8 March 2018 (2018-03-08) claims 60-65; par. 66-67 -----	1-24, 150,151
Y	TULLMAN JENNIFER ET AL: "A ClpS-based N-terminal amino acid binding reagent with improved thermostability and selectivity", BIOCHEMICAL ENGINEERING JOURNAL, ELSEVIER, AMSTERDAM, NL, vol. 154, 17 November 2019 (2019-11-17), XP085969363, ISSN: 1369-703X, DOI: 10.1016/J.BEJ.2019.107438 [retrieved on 2019-11-17] abstract; section 2.1; table 1 ----- -/--	1-24, 150,151

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 September 2021	30/11/2021

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Hohwy, Morten

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2021/033493

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CALLAHAN NICHOLAS ET AL: "Strategies for Development of a Next-Generation Protein Sequencing Platform", TRENDS IN BIOCHEMICAL SCIENCES, ELSEVIER, AMSTERDAM, NL, vol. 45, no. 1, 30 October 2019 (2019-10-30), pages 76-89, XP085968588, ISSN: 0968-0004, DOI: 10.1016/J.TIBS.2019.09.005 [retrieved on 2019-10-30] Fig. 5; scheme B -----	1-24, 150,151
Y	US 9 566 335 B1 (EMILI ANDREW [CA] ET AL) 14 February 2017 (2017-02-14) claim 1; Figs. 1, 9-12 and captions -----	1-24, 150,151
Y	DATABASE Uniprot [Online] 11 November 2015 (2015-11-11), Zhang J Y: "ATP-dependent Clp protease adapter protein ClpS", XP055836295, Database accession no. A0A0K1RWM1 abstract -----	1-24, 150,151

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2021/033493

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed:

      [X] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:

      [ ] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2021/033493

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-24, 150, 151(all partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (2)) (April 2005)

International Application No. PCT/US2021/033493

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-24, 150, 151(all partially)

A recombinant amino acid binding protein comprising a label, a composition comprising such, a sequencing reaction composition comprising such, and a method of sequencing using said amino acid binding protein, wherein said amino acid binding protein is having a sequence at least 80% identical to sequence SEQ. 1 from Table 1.

2-189. claims: 1-24, 150, 151(all partially)

A recombinant amino acid binding protein comprising a label, a composition comprising such, a sequencing reaction composition comprising such, and a method of sequencing using said amino acid binding protein, wherein said amino acid binding protein is having a sequence at least 80% identical to sequence SEQ. n from Table 1 or 2 for invention n.

---

190. claims: 25-57

A polypeptide sequencing reaction mixture comprising an amino acid binding protein and a peptidase.

---

191. claims: 58-64, 90-101

A substrate comprising an array comprising a cleaving means and a binding means, wherein the the binding means and the cleaving means are configured to achieve at least 10 associations events between the binding means and a terminal amino acid on the polypeptide prior to removal.

---

192. claims: 65-75

An amino acid recognition molecule according to formula I of claim 65.

---

193. claims: 76-89

An amino acid recognition molecule joined to a labeled protein via a linker.

---

194. claims: 102-125

A method of sequencing wherein a polypeptide is degraded and a signal pulse duration of between 1 millisecond and 10

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

seconds is used.

---

195. claims: 126-142

An amino acid recognition molecule according to formula II of claim 126.

---

196. claims: 143-149

An amino acid recognition molecule according to formula III of claim 143.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2021/033493

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018045186 A1	08-03-2018	CN 109923216 A	21-06-2019
		EP 3507385 A1	10-07-2019
		GB 2569252 A	12-06-2019
		JP 2019531715 A	07-11-2019
		US 2019194709 A1	27-06-2019
		WO 2018045186 A1	08-03-2018
US 9566335 B1	14-02-2017	US 9566335 B1	14-02-2017
		US 2017212126 A1	27-07-2017
		US 2018372752 A1	27-12-2018
		US 2020141944 A1	07-05-2020

Form PCT/ISA/210 (patent family annex) (April 2005)

フロントページの続き

(51)国際特許分類			FI			テーマコード（参考）
Ｃ１２Ｎ	*9/50*	*(2006.01)*	Ｃ１２Ｎ	9/50		
Ｃ１２Ｑ	*1/37*	*(2006.01)*	Ｃ１２Ｑ	1/37		
Ｇ０１Ｎ	*33/68*	*(2006.01)*	Ｇ０１Ｎ	33/68		
Ｃ１２Ｎ	*15/31*	*(2006.01)*	Ｃ１２Ｎ	15/31	ＺＮＡ	

,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

弁理士　中村 美樹

(72)発明者　リード、ブライアン

アメリカ合衆国 ０６４４３ コネチカット州 マディソン ハル ロード ９４

(72)発明者　パンディ、マンジュラ

アメリカ合衆国 ０６４３７ コネチカット州 ギルフォード ホワイト バーチ ドライブ ４６０

Ｆターム（参考）　2G045 AA34 DA36 FB07 FB12
4B029 AA07 BB15 BB16 CC03 CC08 GA03 GB10
4B050 CC05 EE10 LL03
4B063 QA13 QQ13 QQ36 QQ79 QR16 QR48 QR80 QR82 QR84 QS32 QS36 QS38 QS39 QX02
4H045 AA10 AA20 AA30 BA10 BA40 BA70 CA11 CA40 DA89 EA50 FA74