

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2020-514746

(P2020-514746A)

(43) 公表日 令和2年5月21日 (2020.5.21)

(51) Int. Cl.	F I	テーマコード (参考)
<b>GO 1 N 33/543 (2006.01)</b>	GO 1 N 33/543 5 O 1 A	2 G O 4 5
<b>GO 1 N 33/53 (2006.01)</b>	GO 1 N 33/53 Z N A D	4 B O 6 3
<b>GO 1 N 33/50 (2006.01)</b>	GO 1 N 33/543 5 7 5	
<b>GO 1 N 33/15 (2006.01)</b>	GO 1 N 33/50 Z	
<b>C 1 2 Q 1/68 (2018.01)</b>	GO 1 N 33/15 Z	
審査請求 未請求 予備審査請求 未請求 (全 49 頁) 最終頁に続く		

(21) 出願番号 特願2019-550552 (P2019-550552)  
 (86) (22) 出願日 平成29年12月1日 (2017.12.1)  
 (85) 翻訳文提出日 令和1年7月12日 (2019.7.12)  
 (86) 国際出願番号 PCT/US2017/064322  
 (87) 国際公開番号 W02018/102759  
 (87) 国際公開日 平成30年6月7日 (2018.6.7)  
 (31) 優先権主張番号 62/429,063  
 (32) 優先日 平成28年12月1日 (2016.12.1)  
 (33) 優先権主張国・地域又は機関  
 米国 (US)  
 (31) 優先権主張番号 62/500,455  
 (32) 優先日 平成29年5月2日 (2017.5.2)  
 (33) 優先権主張国・地域又は機関  
 米国 (US)

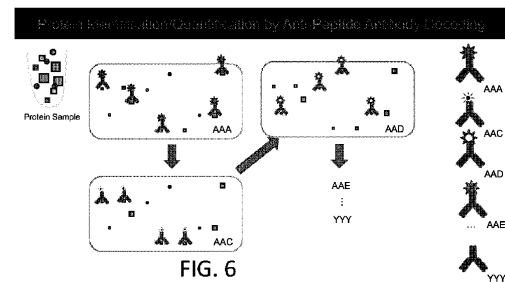
(71) 出願人 519196999  
 ノーティラス バイオテクノロジー イン  
 コーポレイテッド  
 アメリカ合衆国 94025 カリフォル  
 ニア州 メンロ パーク オブライエン  
 ドライブ 1165 スイート ビー  
 (74) 代理人 100102978  
 弁理士 清水 初志  
 (74) 代理人 100102118  
 弁理士 春名 雅夫  
 (74) 代理人 100160923  
 弁理士 山口 裕孝  
 (74) 代理人 100119507  
 弁理士 刑部 俊

最終頁に続く

(54) 【発明の名称】 タンパク質をアッセイする方法

## (57) 【要約】

本明細書において、試料中のタンパク質を同定するための方法およびシステムが提供される。そのいずれもがただ1つのタンパク質に特異的であることもただ1つのタンパク質ファミリーに特異的であることもない、抗体のパネルを入手する。加えて、パネル中の抗体の結合特性を決定する。さらに、タンパク質を抗体のパネルに反復的に曝露する。加えて、このタンパク質に結合する抗体のセットを決定する。抗体のセットをタンパク質の配列と照合するために、抗体の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いて、このタンパク質のアイデンティティを決定する。



**【特許請求の範囲】****【請求項 1】**

以下の工程を含む、タンパク質の特徴を決定する方法：

基材を得る工程であって、（分子レベルで）個々のタンパク質一部分がそれぞれ、光学的に解像可能な固有の空間的アドレスを有するように、1種または複数種のタンパク質の一部が該基材にコンジュゲートされている、工程；

1種または複数種の親和性試薬の第1～（順序づけられた）第nのセットを含む流体を該基材に適用する工程であって、該1種または複数種の親和性試薬のそれぞれは、該1種または複数種のタンパク質の一部のエピトープ（連続的なまたは非連続的なアミノ酸配列）1種に特異的であり、かつ1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は同定可能なタグに連結されている、工程；

1種または複数種の親和性試薬の第1およびそれに続く第nまでのセットの該基材への各適用の後で、以下の工程を実施する工程：

該同定可能なタグを観察する工程；

1つまたは複数の観察シグナルを有する、該基材の1つまたは複数の固有の空間的アドレスを同定する工程；

同定された固有の空間的アドレスを有する該1種または複数種のタンパク質の一部がそれぞれ、該1つまたは複数の観察シグナルに関連付けられる1種または複数種のエピトープを含むことを決定する工程；および

タンパク質一部分それぞれの特徴を、該1種または複数種のエピトープに基づいて決定する工程。

**【請求項 2】**

1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬が、個々のタンパク質に対して特異的であることも個々のタンパク質ファミリーに対して特異的であることもないが、1種または複数種の個々の識別可能なタンパク質の一部に対して特異的である、請求項1に記載の方法。

**【請求項 3】**

1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬が、2種以上のタンパク質中に存在する1種または複数種のエピトープのファミリーを認識する、請求項1に記載の方法。

**【請求項 4】**

1種または複数種のタンパク質の一部のエピトープが立体構造をとるかまたは直線状である、請求項1に記載の方法。

**【請求項 5】**

1種または複数種の親和性試薬が、対応するエピトープに特異的な、連続的なまたは非連続的なアミノ酸を含む、請求項1に記載の方法。

**【請求項 6】**

1種または複数種のタンパク質の一部のアイデンティティを、該一部分の中の決定された1種または複数種のエピトープに基づいて、正確度閾値に応じて決定する工程をさらに含む、請求項5に記載の方法。

**【請求項 7】**

1種または複数種の親和性試薬の第1～第nのセットが、100種超の親和性試薬を含む、前記請求項のいずれか一項に記載の方法。

**【請求項 8】**

ただ1つのタンパク質またはただ1つのタンパク質アイソフォームに結合する親和性試薬の使用をさらに含む、前記請求項のいずれか一項に記載の方法。

**【請求項 9】**

1種または複数種のタンパク質の一部のアイデンティティを、親和性試薬の結合のパターンに基づいて、正確度閾値に応じて決定する工程をさらに含む、前記請求項のいずれか一項に記載の方法。

## 【請求項 10】

基材がフローセルである、前記請求項のいずれか一項に記載の方法。

## 【請求項 11】

1種または複数種のタンパク質の一部分が、光活性化リンカーを用いて基材にコンジュゲートされている、前記請求項のいずれか一項に記載の方法。

## 【請求項 12】

1種または複数種のタンパク質の一部分が、光切断性リンカーを用いて基材にコンジュゲートされている、前記請求項のいずれか一項に記載の方法。

## 【請求項 13】

親和性試薬の少なくとも1つのセットの少なくとも一部分が、同定可能なタグにコンジュゲートされるよう修飾されている、前記請求項のいずれか一項に記載の方法。

10

## 【請求項 14】

同定可能なタグが蛍光タグである、前記請求項のいずれか一項に記載の方法。

## 【請求項 15】

同定可能なタグが核酸バーコードである、前記請求項のいずれか一項に記載の方法。

## 【請求項 16】

タンパク質の同定される一部分によって占有される空間的アドレスの数が、試料中のそのタンパク質のレベルを定量化するために計数される、前記請求項のいずれか一項に記載の方法。

## 【請求項 17】

20

1種または複数種のタンパク質の一部分のアイデンティティが、デコンボリューションソフトウェアを用いて決定される、前記請求項のいずれか一項に記載の方法。

## 【請求項 18】

1種または複数種のタンパク質の一部分のアイデンティティが、固有の空間的アドレスに関連付けられたエピトープの組み合わせをデコーディングすることによって決定される、前記請求項のいずれか一項に記載の方法。

## 【請求項 19】

1種または複数種のタンパク質の一部分を基材にコンジュゲートさせる前に、該1種または複数種のタンパク質を変性させる工程をさらに含む、前記請求項のいずれか一項に記載の方法。

30

## 【請求項 20】

基材に対する1種または複数種のタンパク質の一部分が、複数のタンパク質の複合混合物中に存在する、前記請求項のいずれか一項に記載の方法。

## 【請求項 21】

複数のタンパク質を同定するために使用される、前記請求項のいずれか一項に記載の方法。

## 【請求項 22】

以下の工程を含む、タンパク質を同定する方法：

そのいずれもがただ1つのタンパク質に特異的であることもただ1つのタンパク質ファミリーに特異的であることもない、抗体のパネルを入手する工程、

40

該パネル中の該抗体の結合特性を決定する工程、

タンパク質を該抗体のパネルに反復的に曝露する工程、

該タンパク質に結合する抗体のセットを決定する工程、および

該抗体のセットをタンパク質の配列と照合するために、該抗体の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いる工程であって、それにより該タンパク質のアイデンティティを決定する、工程。

## 【請求項 23】

同定されるべきタンパク質が、複数の異なるタンパク質を含む試料中で同定される、請求項22に記載の方法。

## 【請求項 24】

50

単一試料中の複数のタンパク質を同時に同定することが可能である、請求項22に記載の方法。

【請求項 2 5】

以下の工程を含む、タンパク質を同定する方法：

そのいずれもがただ1つのタンパク質に特異的であることもただ1つのタンパク質ファミリーに特異的であることもない、抗体のパネルを入手する工程、

該パネル中の該抗体の結合特性を決定する工程、

タンパク質を該抗体のパネルに反復的に曝露する工程、

該タンパク質に結合しない抗体のセットを決定する工程、および

該抗体のセットをタンパク質の配列と照合するために、該抗体の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いる工程であって、それにより該タンパク質のアイデンティティを決定する、工程。

10

【請求項 2 6】

m種の親和性試薬を用いて、タンパク質の混合物中のn種のタンパク質を一意的に同定しかつ定量化する方法であって、nはmより大きく、かつnおよびmは1より大きい正の整数であり、かつ該タンパク質は本来備わっている特性によって分離されていない、方法。

【請求項 2 7】

nがmよりもおよそ5倍大きい、請求項26に記載の方法。

【請求項 2 8】

nがmよりもおよそ10倍大きい、請求項26に記載の方法。

20

【請求項 2 9】

nがmよりもおよそ20倍大きい、請求項26に記載の方法。

【請求項 3 0】

m種の結合試薬を用いて、タンパク質の混合物中のn種のタンパク質を一意的に同定しかつ定量化する方法であって、nはmより大きく、かつ該タンパク質は無作為に配置される、方法。

【請求項 3 1】

タンパク質が、サイズに基づく分離方法でも電荷に基づく分離方法でも分離されていない、請求項26に記載の方法。

【請求項 3 2】

30

m種の親和性試薬を用いて、タンパク質分子の混合物中のn種のタンパク質単分子を一意的に同定しかつ定量化する方法であって、nはmより大きく、かつ該タンパク質単分子が基材にコンジュゲートされ、かつ個々のタンパク質分子がそれぞれ、光学的に解像可能な固有の空間的アドレスを有するように、該タンパク質単分子が空間的に分離される、方法。

【請求項 3 3】

親和性試薬のパネルを用いて、n種の可能性のあるタンパク質のプールからの未知のタンパク質1分子を、閾値を上回る確実性をもって同定する方法であって、該パネル中の親和性試薬の数がmであり、かつmがnの10分の1未満である、方法。

【請求項 3 4】

n種の可能性のあるタンパク質のプールから選択される未知のタンパク質を同定することができる、m種の親和性試薬のパネルを選択する方法であって、mがn-1未満である、方法。

40

【請求項 3 5】

n種の可能性のあるタンパク質のプールから選択される未知のタンパク質を同定することができる、m種の親和性試薬のパネルを選択する方法であって、mがnの10分の1未満である、方法。

【請求項 3 6】

4000種未満の親和性試薬のパネルが20,000種の異なるタンパク質のそれぞれを一意的に同定することができるように、該4000種未満の親和性試薬のパネルを選択する方法。

【請求項 3 7】

50

m種の結合試薬を用いて、タンパク質の混合物中のn種のタンパク質を一意的に同定しかつ定量化する方法であって、mはn-1未満であり、かつ各タンパク質は該m種の結合試薬のサブセットによる結合の独特なプロファイルにより同定される、方法。

【請求項 38】

以下の工程を含む、タンパク質1分子を一意的に同定する方法：

親和性試薬のパネルを得る工程；

該タンパク質1分子を、該パネル中の該親和性試薬のそれぞれに曝露する工程；

各親和性試薬が該タンパク質1分子に結合するかまたは結合しないかを決定する工程；  
および

該タンパク質1分子のアイデンティティを決定するために、収集された結合データを用いる工程であって、該タンパク質1分子のアイデンティティは、該親和性試薬のパネル中のいかなる個々の親和性試薬の結合データによっても決定することができない、工程。

10

【請求項 39】

タンパク質試料からヒトプロテオーム中のタンパク質の20%超を同定することができる方法であって、該タンパク質は処理中に実質的に破壊されない、方法。

【請求項 40】

4000種を超える親和性試薬は必要としない、請求項39に記載の方法。

【請求項 41】

100 mgを超えるタンパク質試料は必要としない、請求項39に記載の方法。

【請求項 42】

20

以下の工程を含む、タンパク質の特徴を決定する方法：

基材を得る工程であって、個々のタンパク質一部分がそれぞれ、光学的に解像可能な固有の空間的アドレスを有するように、1種または複数種のタンパク質の一部が該基材にコンジュゲートされている、工程；

1種または複数種の親和性試薬の1～n個のセットから選択された親和性試薬のセットを含む流体を、該基材に適用する工程；

1種または複数種の親和性試薬のセットの該基材への各適用の後で、以下の工程を実施する工程：

結合した親和性試薬を観察する工程；

1種または複数種の結合した親和性試薬を有する、該基材の1つまたは複数の固有の空間的アドレスを同定する工程；および

30

同定された固有の空間的アドレスを有する該1種または複数種のタンパク質の一部がそれぞれ、1つまたは複数の観察シグナルに関連付けられる1種または複数種のエピトープを含むことを決定する工程。

【発明の詳細な説明】

【技術分野】

【0001】

相互参照

本出願は2016年12月1日出願の米国特許仮出願第62/429,063号、および2017年5月2日出願の米国特許仮出願第62/500,455号の恩典を主張するものであり、これらの出願は参照により本明細書に組み入れられる。

40

【背景技術】

【0002】

発明の背景

タンパク質同定のための現行の技術は、典型的には、高度に特異的でかつ高感度の抗体の結合およびそれに続く情報の読み出し、または質量分析計からのペプチド読み取りデータ（典型的には12～30アミノ酸長ほど）のいずれかに依存する。

【発明の概要】

【0003】

本開示は、タンパク質をアッセイするための方法およびシステムを提供する。いくつか

50

の態様において本開示は、非常に不完全であってよく、かつ／または特定のタンパク質に対して特異的ではない一連の測定から、混合物中のタンパク質のアイデンティティ、すなわちそれらの配列を推測するアプローチを提供する。本明細書において記載される方法およびシステムはまた、タンパク質を含む生体高分子を特徴付けしかつ／または同定するために使用されてもよい。加えて、本明細書において記載される方法およびシステムは、質量分析計からのデータに依存するタンパク質同定のための技術よりも迅速にタンパク質を同定するために使用され得る。いくつかの例において、本明細書において記載される方法およびシステムは、少なくとも400個の異なるタンパク質を、少なくとも50%の正確性で、質量分析計からのデータに依存するタンパク質同定のための技術と比べて少なくとも10%より迅速に同定するために、使用され得る。いくつかの例において、本明細書において記載される方法およびシステムは、少なくとも1000個の異なるタンパク質を、少なくとも50%の正確性で、質量分析計からのデータに依存するタンパク質同定のための技術と比べて少なくとも10%より迅速に同定するために、使用され得る。

10

20

30

40

50

#### 【0004】

本発明の一つの局面は、タンパク質の特徴を決定する方法を提供する。該方法は、基材を得る工程を含み、個々のタンパク質一部分がそれぞれ、解像可能な固有の空間的地址を有するように、1種または複数種のタンパク質の一部が基材にコンジュゲートされている。いくつかの場合において、個々のタンパク質一部分はそれぞれ、光学的に解像可能な固有の空間的地址を有してよい。該方法は、1種または複数種の親和性試薬の第1～第nのセットを含む流体を基材に適用する工程をさらに含む。いくつかの態様において、親和性試薬は、同定可能なタグを含んでよく、または同定可能なタグに連結されてもよい。該方法は、1種または複数種の親和性試薬の第1～第nのセットの基材への各適用の後に、以下の工程を実施する工程を含む：親和性試薬または同定可能なタグを観察する工程；1つまたは複数の観察シグナルを有する、基材の1つまたは複数の固有の空間的地址を同定する工程；および、同定された固有の空間的地址を有する1種または複数種のタンパク質の一部がそれぞれ、1つまたは複数の観察シグナルに関連付けられる1種または複数種のエピトープを含むことを決定する工程。いくつかの例において、1種または複数種のタンパク質のコンジュゲートされた一部分のそれぞれは、基材上の固有の空間的地址に関連付けられる。いくつかの例において、1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は、個々のタンパク質に対して特異的であることも個々のタンパク質ファミリーに対して特異的であることもない。いくつかの例において、親和性試薬の結合エピトープは公知ではなく、個々のタンパク質に対して特異的であることも個々のタンパク質ファミリーに対して特異的であることもない。

#### 【0005】

いくつかの場合において、本開示の方法はまた、単一の位置に複数のタンパク質が結合された基材とともに使用されてよく、単一の位置におけるタンパク質の少なくとも約50%、60%、70%、80%、90%、または90%超は、共通のアミノ酸配列を含む。いくつかの場合において、本開示の方法はまた、単一の位置に複数のタンパク質が結合された基材とともに使用されてよく、単一の位置におけるタンパク質の少なくとも約50%、60%、70%、80%、90%、または90%超は、少なくとも95%のアミノ酸配列同一性を含む。

#### 【0006】

いくつかの態様において、1種または複数種のタンパク質は、1種のタンパク質単分子を含み得る。いくつかの態様において、1種または複数種のタンパク質は、バルクタンパク質を含み得る。いくつかの態様において、1種または複数種のタンパク質は、基材上の同じ固有の空間的地址にコンジュゲートされた、複数の同じタンパク質を含み得る。

#### 【0007】

いくつかの態様において、1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は、2種以上のタンパク質に存在する1種または複数種のエピトープのファミリーを認識する。いくつかの態様において、該方法は、1種または複数種のタンパク質の一部のアイデンティティを、該一部分の中の決定された1種または複数種のエピトープに基づ

き、正確度閾値に応じて決定する工程をさらに含む。いくつかの例において、1種または複数種の親和性試薬の第1～第nのセットは、100種超の親和性試薬を含む。いくつかの態様において、該方法は、ただ1つのタンパク質またはただ1つのタンパク質アイソフォームに結合する親和性試薬の使用をさらに含む。

【0008】

いくつかの態様において、該方法は、1種または複数種のタンパク質の一部分のアイデンティティを、親和性試薬の結合のパターンに基づき、正確度閾値に応じて決定する工程をさらに含む。いくつかの例において、基材はフローセルである。いくつかの例において、1種または複数種のタンパク質の一部分は、光活性化リンカーを用いて基材にコンジュゲートされる。いくつかの例において、1種または複数種のタンパク質の一部分は、光切断性リンカーを用いて基材にコンジュゲートされる。

10

【0009】

いくつかの例において、親和性試薬の少なくとも1つのセットの少なくとも一部分は、同定可能なタグにコンジュゲートされるよう修飾されている。いくつかの例において、同定可能なタグは蛍光タグである。いくつかの例において、同定可能なタグは磁気タグである。いくつかの例において、同定可能なタグは核酸バーコードである。いくつかの例において、同定可能なタグは親和性タグ（たとえばビオチン、Flag、myc）である。いくつかの例において、タンパク質の同定される一部分によって占有される空間的アドレスの数は、試料中のそのタンパク質のレベルを定量化するために計数される。いくつかの例において、1種または複数種のタンパク質の一部分のアイデンティティは、デコンボリューションソフトウェアを用いて決定される。いくつかの例において、1種または複数種のタンパク質の一部分のアイデンティティは、固有の空間的アドレスに関連付けられたエピトープの組み合わせをデコーディングすることによって決定される。いくつかの例において、該方法は、1種または複数種のタンパク質の一部分を基材にコンジュゲートさせる前に該1種または複数種のタンパク質を変性させる工程を、さらに含む。いくつかの例において、基材に対する1種または複数種のタンパク質の一部分は、複数のタンパク質の複合混合物中に存在する。いくつかの例において、該方法は複数のタンパク質を同定するために使用される。

20

【0010】

本発明の追加の局面は、以下の工程を含むタンパク質を同定する方法を提供する：そのいずれもがただ1つのタンパク質に特異的であることもただ1つのタンパク質ファミリーに特異的であることもない、親和性試薬のパネルを入手する工程、パネル中の抗体の結合特性を決定する工程、タンパク質を抗体のパネルに反復的に曝露する工程、該タンパク質に結合する抗体のセットを決定する工程、および、抗体のセットをタンパク質の配列と照合するために、抗体の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いる工程であって、それにより該タンパク質のアイデンティティを決定する、工程。いくつかの例において、同定されるべきタンパク質は、複数の異なるタンパク質を含む試料中で同定される。いくつかの例において、該方法は、単一試料中の複数のタンパク質を同時に同定することが可能である。

30

【0011】

本発明の別の局面は、タンパク質を同定する方法を提供する。該方法は、そのいずれもがただ1つのタンパク質に特異的であることもただ1つのタンパク質ファミリーに特異的であることもない、抗体のパネルを入手する工程、パネル中の抗体の結合特性を決定する工程、タンパク質を抗体のパネルに反復的に曝露する工程、該タンパク質に結合しない抗体のセットを決定する工程、および、抗体のセットをタンパク質の配列と照合するために、抗体の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いる工程であって、それにより該タンパク質のアイデンティティを決定する、工程を含む。

40

【0012】

本発明の別の局面は、m種の親和性試薬を用いて、タンパク質の混合物中のn種のタンパク質を一意的に同定しかつ定量化する方法を提供し、ここでnはmより大きく、かつnおよ

50

び $m$ は1より大きい正の整数であり、かつタンパク質は本来備わっている特性によって分離されていない。いくつかの例において、 $n$ は $m$ よりおよそ5倍、10倍、20倍、50倍、100倍、500倍、1,000倍、5,000倍、または10,000倍大きい。

【0013】

本発明の別の局面は、 $m$ 種の結合試薬を用いて、タンパク質の混合物中の $n$ 種のタンパク質を一意的に同定しかつ定量化する方法を提供し、ここで $n$ は $m$ より大きく、かつタンパク質は無作為に配置される。いくつかの例において、タンパク質は、サイズに基づく分離方法でも電荷に基づく分離方法でも分離されていない。

【0014】

本発明の別の局面は、 $m$ 種の親和性試薬を用いて、タンパク質分子の混合物中の $n$ 種のタンパク質単分子を一意的に同定しかつ定量化する方法を提供する。該方法は、 $n$ は $m$ より大きいこと、および該タンパク質単分子が基材にコンジュゲートされ、かつ個々のタンパク質分子がそれぞれ、光学的に解像可能な固有の空間的アドレスを有するように、該タンパク質単分子が空間的に分離されていることを、さらに含む。

【0015】

本発明の別の局面は、 $n$ 種の可能性のあるタンパク質のプールからの未知のタンパク質1分子を、閾値を上回る確実性をもって同定する方法を提供する。該方法は親和性試薬のパネルを用いる工程を含み、ここでパネル中の親和性試薬の数は $m$ であり、かつ $m$ は $n$ の10分の1未満である。

【0016】

本発明の別の局面は、 $n$ 種の可能性のあるタンパク質のプールから選択される未知のタンパク質を同定することができる、 $m$ 種の親和性試薬のパネルを選択する方法を提供し、ここで $m$ は $n-1$ 未満である。

【0017】

本発明の別の局面は、 $n$ 種の可能性のあるタンパク質のプールから選択される未知のタンパク質を同定することができる、 $m$ 種の親和性試薬のパネルを選択する方法を提供し、ここで $m$ は $n$ の10分の1未満である。

【0018】

本発明の別の局面は、4000種未満の親和性試薬のパネルが20,000種の異なるタンパク質のそれぞれを一意的に同定できるように、該4000種未満の親和性試薬のパネルを選択する方法を、提供する。

【0019】

本発明の別の局面は、 $m$ 種の結合試薬を用いて、タンパク質の混合物中の $n$ 種のタンパク質を一意的に同定しかつ定量化する方法を提供し、ここで $m$ は $n-1$ 未満であり、かつ各タンパク質は、 $m$ 種の結合試薬のサブセットによる結合の独特なプロファイルにより同定される。

【0020】

いくつかの例において、該方法は、ヒトタンパク質試料からヒトプロテオーム中のタンパク質の20%超を同定することができ、ここでタンパク質は処理中に実質的に破壊されない。いくつかの例において、該方法は、利用可能なタンパク質配列データベースを有する任意の生物（たとえば酵母、大腸菌（*E. coli*）、線虫（*C. elegans*））のプロテオーム中のタンパク質の20%超を同定することができる。いくつかの例において、タンパク質配列データベースは、ゲノムの、エキソームの、および/またはトランスクリプトームの配列決定によって生成されてよい。いくつかの例において、該方法は4000種を超える親和性試薬を必要としない。いくつかの例において、該方法は100 mgを超えるタンパク質試料を必要としない。

【0021】

本発明の別の局面は、タンパク質1分子を一意的に同定する方法を提供する。該方法は、親和性試薬のパネルを得る工程、該タンパク質1分子を、パネル中の親和性試薬のそれぞれに曝露する工程、各親和性試薬が該タンパク質1分子に結合するかまたは結合しない



かを決定する工程、および該タンパク質1分子のアイデンティティを決定するために、収集された結合データを用いる工程を含む。加えて、いくつかの態様において、該タンパク質1分子のアイデンティティは、親和性試薬のパネル中のいかなる個々の親和性試薬の結合データによっても決定することはできない。いくつかの例において、重複した結合特徴を有する親和性試薬は、任意の特定の標的への親和性を強化するために使用されてよい。

#### 【0022】

本発明の別の局面は、タンパク質の特徴を決定する方法を提供する。該方法は、1種または複数種のタンパク質の一部を基材へコンジュゲートさせる工程を含み、ここで1種または複数種のタンパク質のコンジュゲートされた一部分はそれぞれ、基材上の固有の空間的アドレスに関連付けられる。いくつかの例において、固有の空間的アドレスは、タンパク質の特定の一部分に関連付けられる空間的アドレスであってよい。該方法はまた、1種または複数種の親和性試薬の第1～第nのセットを基材に適用する工程を含み、ここで1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は、1～10残基の長さのエピトープを認識し、かつ1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は、同定可能なタグに連結されている。加えて該方法は、1種または複数種の親和性試薬の第1～第nのセットの基材への各適用の後に、以下の工程が実施されることを含む：同定可能なタグを観察する工程；1つまたは複数の観察シグナルを有する、基材の1つまたは複数の固有の空間的アドレスを同定する工程；および同定された固有の空間的アドレスを有する1種または複数種のタンパク質の一部がそれぞれ、1つまたは複数の観察シグナルに関連付けられる1種または複数種のエピトープを含むことを決定する工程。

10

20

#### 【0023】

本発明の別の局面は、タンパク質の特徴を決定する方法を提供する。該方法は、1種または複数種のタンパク質の一部を基材にコンジュゲートさせる工程を含み、ここで1種または複数種のタンパク質のコンジュゲートされた一部分はそれぞれ、基材上の固有の空間的アドレスに関連付けられる。該方法はまた、1種または複数種の親和性試薬の第1～第nのセットを基材に適用する工程を含み、ここで1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は、1種または複数種のタンパク質に存在する1種または複数種のエピトープのファミリーを認識し、かつ1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は、同定可能なタグに連結されている。さらに、該方法は、1種または複数種の親和性試薬の第1～第nのセットの基材への各適用の後に、以下の工程が実施されることを含む：同定可能なタグを観察する工程；観察シグナルを有する、基材の1つまたは複数の固有の空間的アドレスを同定する工程；および同定された固有の空間的アドレスを有する1種または複数種のタンパク質の一部がそれぞれ、該エピトープを含むことを決定する工程。

30

#### 【0024】

本発明のさらなる局面は、タンパク質を同定する方法を提供し、該方法は以下の工程を含む：非特異性の度合いが既知である親和性試薬のパネルを入手する工程、パネル中の親和性試薬の結合特性を決定する工程、タンパク質を親和性試薬のパネルに反復的に曝露する工程、該タンパク質に結合する親和性試薬のセットを決定する工程、および、親和性試薬のセットをタンパク質の配列と照合するために、親和性試薬の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いる工程であって、それにより該タンパク質のアイデンティティを決定する、工程。

40

#### 【0025】

加えて、本発明の別の局面は、タンパク質を同定する方法を提供し、該方法は以下の工程を含む：非特異性の度合いが既知である親和性試薬のパネルを入手する工程、パネル中の親和性試薬の結合特性を決定する工程、タンパク質を親和性試薬のパネルに反復的に曝露する工程、該タンパク質に結合しない親和性試薬のセットを決定する工程、および、親和性試薬のセットをタンパク質の配列と照合するために、親和性試薬の既知の結合特性に基づく1つまたは複数のデコンボリューション方法を用いる工程であって、それにより該タンパク質のアイデンティティを決定する、工程。

50

## 【0026】

## 参照による組み入れ

本明細書において言及されるすべての刊行物、特許、および特許出願は、個々の刊行物、特許、または特許出願がそれぞれ、参照により組み入れられるように具体的にかつ個々に示された場合と同程度に、参照により本明細書に組み入れられる。

## 【図面の簡単な説明】

## 【0027】

本発明の新規な特徴は、添付の請求の範囲において詳細に説明される。本発明の特徴および利点のより良い理解は、本発明の原理が利用される例証的態様が説明される以下の詳細な説明、および添付される以下の図面の参照により得られるであろう。

10

## 【0028】

【図1】いくつかの態様にしたがった、抗ペプチド抗体デコーディングによるタンパク質定量化の第1の図解を示す。

【図2】いくつかの態様にしたがった、抗ペプチド抗体デコーディングによるタンパク質定量化の第2の図解を示す。

【図3】いくつかの態様にしたがった、フローセルコンジュゲーションを示す。

【図4】いくつかの態様にしたがった、フローセル上の固有の空間的アドレスの格子を示す。

【図5】いくつかの態様にしたがった、d-コード抗体と照合され得るペプチドのセットとしてのタンパク質の解体を示す。

20

【図6】いくつかの態様にしたがった、抗ペプチド抗体デコーディングによるタンパク質の同定/定量化の図解を示す。

【図7】いくつかの態様にしたがった、抗ペプチド抗体の第1のセットの観察を示す。

【図8】いくつかの態様にしたがった、抗ペプチド抗体の第2のセットの観察を示す。

【図9】いくつかの態様にしたがった、抗ペプチド抗体の第3のセットの観察を示す。

【図10】いくつかの態様にしたがった、抗体測定データのコンピューターによるデコーディングを示す。

【図11】いくつかの態様にしたがった、プロテオームの定量化を示す。

【図12】いくつかの態様にしたがった、例外リストの一例を示す。

【図13】態様にしたがった、定量化のために必要であり得る3マー-d-コード抗体サンプリングのカバレッジを示す。

30

【図14】本明細書において提供される方法を遂行するためにプログラムされた、または別の方法で設定された、コンピューター制御システムを示す。

【図15】本明細書における態様にしたがった、可同定性 対 プロテオームのカバレッジにおける、3マー-d-コードプローブ数の影響力の一例を示す。

【図16A】本明細書における態様にしたがった、基材にコンジュゲートされたタンパク質単分子を表す画像を示す。

【図16B】本明細書における態様にしたがった、矢印で示されるコンジュゲートされたタンパク質を有する、図16Aの示された区域の拡大部分を表す画像を示す。

40

【図17】本明細書の態様にしたがった、タンパク質の同定を示す。

【図18】本明細書の態様にしたがった、タンパク質の同定の図解を示す。

## 【発明を実施するための形態】

## 【0029】

## 発明の詳細な説明

いくつかの例において、アプローチは、以下の3つの局面を含み得る：1)タンパク質および/またはタンパク質断片がコンジュゲートできる、アドレス指定可能な基材；2)たとえば各親和性試薬が多様な特異性でペプチドに結合し得る、親和性試薬のセット；ならびに3)基材中の精確な空間的アドレスにおけるタンパク質のアイデンティティを推測するための、親和性試薬の結合の特徴についての事前知識、基材中の各アドレスにおける親和性試薬の結合の特異的パターン、および/または混合物（たとえばヒトプロテオーム）中の

50

タンパク質の可能性のある配列のデータベース、の組み合わせを使用することができる、ソフトウェア。いくつかの例において、精確な空間的地址は、固有の空間的地址であってよい。

#### 【0030】

##### 試料

試料は、タンパク質を含む任意の生物学的試料であってよい。試料は、組織もしくは細胞から、または組織もしくは細胞の環境から、採取されてよい。いくつかの例において、試料は、組織生検材料、血液、血漿、細胞外液、培養細胞、培地、破棄された組織、植物性物質、合成タンパク質、古細菌の、細菌の、および/もしくはウイルスの試料、菌糸組織、古細菌、または原生動物であり得る。いくつかの例において、タンパク質は、試料調製の間に、その一次供給源（細胞、組織、血液などの体液、環境試料等）から単離される。タンパク質は、その一次供給源から精製されてもよく、または精製されなくてもよい。いくつかの場合において、一次供給源は、さらなる処理の前にホモジナイズされる。いくつかの場合において、細胞は、RIPAバッファーなどの緩衝液を用いて溶解される。変性緩衝液もまた、この段階で使用されてよい。試料は、脂質および粒子状物質を除去するために、ろ過または遠心分離されてよい。試料はまた、核酸を除去するために精製されてよく、またはRNアーゼおよびDNアーゼで処理されてもよい。試料は、無傷のタンパク質、変性タンパク質、タンパク質断片、または部分的に分解したタンパク質を含んでよい。

10

#### 【0031】

試料は、疾患または障害を有する対象から採取されてよい。疾患または障害は、感染性疾患、免疫障害もしくは免疫疾患、がん、遺伝性疾患、変性疾患、生活習慣病、損傷、希少疾患、または加齢に関する疾患であり得る。感染性疾患は、細菌、ウイルス、菌類、および/または寄生生物によって引き起こされ得る。がんの非限定的な例は、膀胱がん、肺がん、脳がん、黒色腫、乳がん、非ホジキンリンパ腫、子宮頸がん、卵巣がん、結腸直腸がん、膵臓がん、食道がん、前立腺がん、腎臓がん、皮膚がん、白血病、甲状腺がん、肝臓がん、および子宮がんを含む。遺伝性疾患または障害のいくつかの例は、限定するものではないが、嚢胞性線維症、シャルコー・マリー・トゥース病、ハンチントン病、ポイツ・ジェガース症候群、ダウン症候群、関節リウマチ、およびテイ・サックス病を含む。生活習慣病の非限定的な例は、肥満、糖尿病、動脈硬化症、心臓病、脳卒中、高血圧、肝硬変、腎炎、がん、慢性閉塞性肺疾患（copd）、聴覚の問題、および慢性背痛を含む。損傷のいくつかの例は、限定するものではないが、擦過傷、脳損傷、挫傷、火傷、震とう症、うっ血性心不全、建築現場での傷害、脱臼、動揺胸、骨折、血胸、椎間板ヘルニア、ヒップポインター、低体温症、裂傷、神経が圧迫された状態（pinched nerve）、気胸、肋骨骨折、坐骨神経痛、脊髄損傷、腱・靱帯・筋膜の損傷、外傷性脳損傷、およびむち打ち症を含む。試料は、疾患または障害を有する対象の処置の前および/または後に採取されてよい。試料は、処置の前および/または後に採取されてよい。試料は、処置中または処置計画の間に採取されてよい。複数の試料が、処置の効果を経時的にモニターするために、対象から採取されてよい。試料は、診断用抗体が利用可能ではない感染性疾患を有することが分かっている、またはそれが疑われる対象から、採取されてよい。

20

30

#### 【0032】

試料は、疾患または障害を有することが疑われる対象から採取されてよい。試料は、疲労、悪心、体重減少、うずきおよび痛み、衰弱、または健忘などの、原因不明の症状を経験中である対象から、採取されてよい。試料は、原因が明らかな症状を有する対象から採取されてよい。試料は、家族の病歴、年齢、環境曝露、生活習慣上のリスク因子、もしくは他の公知のリスク因子の存在などの因子に起因する疾患または障害を発症するリスクのある対象から、採取されてよい。

40

#### 【0033】

試料は、胚、胎児、または妊婦から採取されてよい。いくつかの例において、試料は、母親の血漿から単離されたタンパク質を含んでよい。いくつかの例においては、母親の血液中の、循環する胎児細胞から単離されたタンパク質。

50

## 【0034】

タンパク質は、エピトープ結合を妨害し得る修飾を除去するために処理されてよい。たとえばタンパク質は、翻訳後グリコシル化を除去するためにグリコシダーゼ処理されてよい。タンパク質は、タンパク質中のジスルフィド結合を還元するために還元剤で処理されてよい。タンパク質は、リン酸基を除去するためにホスファターゼで処理されてよい。除去され得る翻訳後修飾の、他の非限定的な例は、アセテート、アミド基、メチル基、脂質、ユビキチン、ミリストイル化、パルミトイル化、イソプレニル化またはプレニル化（たとえばファルネソールおよびゲラニルゲラニオール）、ファルネシル化、ゲラニルゲラニル化、グリピエーション（glypiation）、リポイル化、フラビンモエティ付着、ホスホパンテティニル化、ならびにレチニリデンシッフ塩基形成を含む。試料はまた、翻訳後タンパク質修飾を保つために処理されてよい。いくつかの例において、ホスファターゼ阻害剤が試料に加えられてよい。いくつかの例において、ジスルフィド結合を保護するために酸化剤が加えられてよい。

10

## 【0035】

次に、タンパク質は完全にまたは部分的に変性されてよい。いくつかの態様において、タンパク質は完全に変性され得る。タンパク質は、界面活性剤、強酸もしくは強塩基、濃縮された無機塩、有機溶媒（たとえばアルコールもしくはクロロホルム）、放射線照射、または熱などの外部ストレスの適用によって変性されてよい。タンパク質は、変性緩衝液の添加により変性されてよい。タンパク質はまた、変性緩衝液中で、沈殿、凍結乾燥、および懸濁されてよい。タンパク質は、加熱により変性されてよい。タンパク質に化学的修飾を生じさせる可能性の低い変性方法が、選ばれ得る。

20

## 【0036】

試料のタンパク質は、より短いポリペプチドを産生するために、コンジュゲーションの前または後のいずれかで処理されてよい。残りのタンパク質は、断片を生成するためにプロテイナーゼKなどの酵素で部分的に消化してよく、または無傷のままにしておいてもよい。さらなる例において、タンパク質は、トリプシンなどのプロテアーゼに曝露されてよい。プロテアーゼの追加の例は、セリンプロテアーゼ、システインプロテアーゼ、スレオニンプロテアーゼ、アスパラギン酸プロテアーゼ、グルタミン酸プロテアーゼ、メタロプロテアーゼ、およびアスパラギンペプチドリアーゼを含んでよい。

## 【0037】

いくつかの場合において、極度に大きいおよび小さいタンパク質（たとえばタイチン）を除去することは有用となり得、そのようなタンパク質はる過または他の適切な方法によって除去され得る。いくつかの例において、極度に大きいタンパク質は、400 kD、450 kD、500 kD、600 kD、650 kD、700 kD、750 kD、800 kD、または850 kDを超えるタンパク質を含んでよい。いくつかの例において、極度に大きいタンパク質は、約8,000アミノ酸、約8,500アミノ酸、約9,000アミノ酸、約9,500アミノ酸、約10,000アミノ酸、約10,500アミノ酸、約11,000アミノ酸、または約15,000アミノ酸を超えるタンパク質を含んでよい。いくつかの例において、小さいタンパク質は、約10 kD未満、9 kD未満、8 kD未満、7 kD未満、6 kD未満、5 kD未満、4 kD未満、3 kD未満、2 kD未満、または1 kD未満のタンパク質を含んでよい。いくつかの例において、小さいタンパク質は、約50アミノ酸未満、45アミノ酸未満、40アミノ酸未満、35アミノ酸未満、または約30アミノ酸未満のタンパク質を含んでよい。極度に大きいまたは小さいタンパク質は、サイズ排除クロマトグラフィーにより除去され得る。極度に大きいタンパク質は、サイズ排除クロマトグラフィーによって単離され、中間サイズのポリペプチドを産生するためにプロテアーゼで処理され、試料の中間サイズのタンパク質と再度組み合わされてよい。

30

40

## 【0038】

いくつかの場合において、タンパク質はサイズ順に整列されてよい。いくつかの場合において、タンパク質は、タンパク質をマイクロウェル中に分類することによって整列されてよい。いくつかの場合において、タンパク質は、タンパク質をナノウェル中に分類することによって整列されてよい。いくつかの場合において、タンパク質は、タンパク質をSD

50

S-PAGEゲルなどのゲルを通過させることによって整列されてよい。いくつかの場合において、タンパク質は、サイズによる他の分画方法によって整列されてよい。いくつかの場合において、タンパク質は電荷に基づいて分離されてよい。いくつかの場合において、タンパク質は疎水性に基づいて分離されてよい。いくつかの場合において、タンパク質は、他の物理的な特徴に基づいて分離されてよい。いくつかの場合において、タンパク質は変性条件下で分離されてよい。いくつかの場合において、タンパク質は非変性条件下で分離されてよい。いくつかの場合において、分画されたタンパク質の異なる画分は、基材の異なる領域に配置されてよい。いくつかの場合において、分離されたタンパク質の異なる一部分は、基材の異なる領域に配置されてよい。いくつかの場合において、タンパク質試料は、SDS-PAGEゲルにおいて分離され、そして、タンパク質が連続体においてサイズによって分類されるように、SDS-PAGEゲルから基材へと転写されてよい。いくつかの場合において、タンパク質試料は、サイズに基づいて3つの画分に分類されてよく、かつ該3つの画分はそれぞれ、基材上の第1、第2、および第3の領域に適用されてよい。いくつかの場合において、本明細書において記載されるシステムおよび方法において使用されるタンパク質は、分類されていてよい。いくつかの場合において、本明細書において記載されるシステムおよび方法において使用されるタンパク質は、分類されていなくてよい。

10

#### 【0039】

タンパク質は、試料を多重化することを可能にするために、たとえば同定可能なタグで、タグ付けされてよい。同定可能なタグのいくつかの非限定的な例は、以下を含む：フルオロフォアまたは核酸バーコード化された塩基リンカー。使用されるフルオロフォアは、GFP、YFP、RFP、eGFP、mCherry、tdtomato、FITC、Alexa Fluor 350、Alexa Fluor 405、Alexa Fluor 488、Alexa Fluor 532、Alexa Fluor 546、Alexa Fluor 555、Alexa Fluor 568、Alexa Fluor 594、Alexa Fluor 647、Alexa Fluor 680、Alexa Fluor 750、Pacific Blue、クマリン、BODIPY FL、Pacific Green、Oregon Green、Cy3、Cy5、Pacific Orange、TRITC、Texas Red、R-フィコエリスリン、アロフィコシアニン (Allophycocyanin) などの蛍光タンパク質、または当技術分野において公知の他のフルオロフォアを含んでよい。

20

#### 【0040】

任意の数のタンパク質試料が多重化され得る。たとえば多重化された反応は、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、約20、約25、約30、約35、約40、約45、約50、約55、約60、約65、約70、約75、約80、約85、約90、約95、約100、または100超の初期試料からのタンパク質を含んでよい。同定可能なタグは、各タンパク質を、その由来の試料に関して調べる手段を提供してよく、または異なる試料からのタンパク質を、固体支持体上の異なる区域に隔離するように誘導してもよい。

30

#### 【0041】

##### 基材

いくつかの態様において、タンパク質はその後、タンパク質を基材に化学的に付着させるために、機能化された基材に適用される。いくつかの場合において、タンパク質は、ビオチンの付着を介して基材に付着されてよい。いくつかの場合において、タンパク質は、核酸の付着を介して基材に付着されてよい。いくつかの態様において、タンパク質には仲介物質が適用されてよく、その後仲介物質は基材に付着する。いくつかの場合において、タンパク質は、その後に表面（たとえばチオール化された表面）の上に捕捉され得るビーズ（たとえば金ビーズ）にコンジュゲートされてよい。いくつかの場合において、タンパク質1つが各ビーズにコンジュゲートされてよい。いくつかの場合において、タンパク質はビーズにコンジュゲートされてよく（たとえばビーズ1つにつきタンパク質1つ）、かつビーズは表面上に捕捉されてよい（たとえばマイクロウェルおよび/またはナノウェル中で）。

40

#### 【0042】

基材は、固体支持体を形成することができる任意の基材であってよい。本明細書において使用されるように、基材、または固体基材とは、タンパク質が共有結合的または非共有

50

結合的に付着できる、任意の固体表面を指し得る。固体基材の非限定的な例は、粒子、ビーズ、スライド、装置の構成要素の表面、膜、フローセル、ウェル、チャンバー、マイクロ流体チャンバーを含み、平面であるかもしくは湾曲しているか、または他の形状を有し得、かつ平滑であり得るかもしくは凹凸を有し得る。いくつかの場合において、基材の表面はマイクロウェルを含んでよい。いくつかの場合において、基材の表面はナノウェルを含んでよい。いくつかの場合において、基材の表面は、1つまたは複数のナノウェルと組み合わせた1つまたは複数のマイクロウェルを含んでよい。いくつかの態様において、基材は、ガラス、デキストランなどの糖類、ポリスチレンもしくはポリプロピレンなどのプラスチック、ポリアクリルアミド、ラテックス、シリコン、金などの金属、またはセルロースから構成され得、かつオリゴヌクレオチドの共有結合的もしくは非共有結合的付着を可能にするまたは増強するために、さらに修飾されてよい。たとえば、基材の表面は、マレイン酸モエティもしくはコハク酸モエティなどの特定の官能基での修飾によって機能化されてよく、またはアミノ基、チオール基、もしくはアクリレート基などの化学的な反応基での修飾によって、たとえばシラン化によって、誘導体化されてもよい。適切なシラン試薬は、アミノプロピルトリメトキシシラン、アミノプロピルトリエトキシシラン、および4-アミノブチルトリエトキシシランを含む。基材は、N-ヒドロキシスクシンイミド（NHS）官能基で機能化されてよい。ガラス表面はまた、たとえばエポキシシラン、アクリレートシラン、またはアクリルアミドシランを用いて、アクリレートまたはエポキシなどの他の反応基で誘導体化され得る。オリゴヌクレオチド付着のための基材および処理は、好ましくは、繰り返される結合、洗浄、画像化、および溶出の各工程に対して安定である。いくつかの例において、基材はスライドまたはフローセルであってよい。

10

20

#### 【0043】

官能基の規則正しいアレイは、たとえばフォトリソグラフィ、ディップペンナノリソグラフィ、ナノインプリントリソグラフィ、ナノスフェアリソグラフィ、ナノボールリソグラフィ、ナノピラーアレイ、ナノワイヤリソグラフィ、走査型プローブリソグラフィ、熱化学リソグラフィ、熱走査型プローブリソグラフィ、局所酸化ナノリソグラフィ、分子自己集合、ステンシルリソグラフィ、または電子線リソグラフィによって作製されてよい。規則正しいアレイ中の官能基は、官能基がそれぞれ、任意の他の官能基から200ナノメートル（nm）未満であるように、または約200 nm、約225 nm、約250 nm、約275 nm、約300 nm、約325 nm、約350 nm、約375 nm、約400 nm、約425 nm、約450 nm、約475 nm、約500 nm、約525 nm、約550 nm、約575 nm、約600 nm、約625 nm、約650 nm、約675 nm、約700 nm、約725 nm、約750 nm、約775 nm、約800 nm、約825 nm、約850 nm、約875 nm、約900 nm、約925 nm、約950 nm、約975 nm、約1000 nm、約1025 nm、約1050 nm、約1075 nm、約1100 nm、約1125 nm、約1150 nm、約1175 nm、約1200 nm、約1225 nm、約1250 nm、約1275 nm、約1300 nm、約1325 nm、約1350 nm、約1375 nm、約1400 nm、約1425 nm、約1450 nm、約1475 nm、約1500 nm、約1525 nm、約1550 nm、約1575 nm、約1600 nm、約1625 nm、約1650 nm、約1675 nm、約1700 nm、約1725 nm、約1750 nm、約1775 nm、約1800 nm、約1825 nm、約1850 nm、約1875 nm、約1900 nm、約1925 nm、約1950 nm、約1975 nm、約2000 nm、もしくは2000 nm超であるように、配置されてよい。無作為な間隔の官能基は、官能基が、任意の他の官能基から平均して少なくとも約50 nm、約100 nm、約150 nm、約200 nm、約250 nm、約300 nm、約350 nm、約400 nm、約450 nm、約500 nm、約550 nm、約600 nm、約650 nm、約700 nm、約750 nm、約800 nm、約850 nm、約900 nm、約950 nm、約1000 nm、または100 nm超であるような密集状態で、提供されてよい。

30

40

#### 【0044】

基材は、間接的に機能化されてよい。たとえば、基材はPEG化されてよく、かつ官能基がPEG分子のすべて、またはPEG分子のサブセットに適用されてよい。加えて、上述のように、いくつかの場合において、ビーズ（たとえば金ビーズ）がコンジュゲートされてよく、かつその後ビーズは、表面（たとえばチオール化された面）上に捕捉されてよい。いくつかの場合において、タンパク質1つが各ビーズにコンジュゲートされてよい。いくつかの場合において、タンパク質はビーズにコンジュゲートされてよく（たとえばビーズ1つ

50

につきタンパク質1つ)、かつピーズは表面上に捕捉されてよい(たとえばマイクロウェルおよび/またはナノウェル中で)。

【0045】

基材は、マイクロスケールのまたはナノスケールの構造(たとえばマイクロウェル、ナノウェル、マイクロピラー、単一分子アレイ、ナノボール、ナノピラー、もしくはナノワイヤなどの規則正しい構造)のために適した技術を用いて、機能化されてよい。いくつかの場合において、基材は、異なるサイズのマイクロウェルを有してよい。いくつかの場合において、マイクロウェルは、1マイクロメートル( $\mu\text{m}$ )であってよく、約2  $\mu\text{m}$ 、約3  $\mu\text{m}$ 、約4  $\mu\text{m}$ 、約5  $\mu\text{m}$ 、約6  $\mu\text{m}$ 、約7  $\mu\text{m}$ 、約8  $\mu\text{m}$ 、約9  $\mu\text{m}$ 、約10  $\mu\text{m}$ 、約15  $\mu\text{m}$ 、約20  $\mu\text{m}$ 、約25  $\mu\text{m}$ 、約30  $\mu\text{m}$ 、約35  $\mu\text{m}$ 、約40  $\mu\text{m}$ 、約45  $\mu\text{m}$ 、約50  $\mu\text{m}$ 、約55  $\mu\text{m}$ 、約60  $\mu\text{m}$ 、約65  $\mu\text{m}$ 、約70  $\mu\text{m}$ 、約75  $\mu\text{m}$ 、約80  $\mu\text{m}$ 、約85  $\mu\text{m}$ 、約90  $\mu\text{m}$ 、約95  $\mu\text{m}$ 、約100  $\mu\text{m}$ 、約105  $\mu\text{m}$ 、約110  $\mu\text{m}$ 、約115  $\mu\text{m}$ 、約120  $\mu\text{m}$ 、約125  $\mu\text{m}$ 、約130  $\mu\text{m}$ 、約135  $\mu\text{m}$ 、約140  $\mu\text{m}$ 、約145  $\mu\text{m}$ 、約150  $\mu\text{m}$ 、約155  $\mu\text{m}$ 、約160  $\mu\text{m}$ 、約165  $\mu\text{m}$ 、約170  $\mu\text{m}$ 、約175  $\mu\text{m}$ 、約180  $\mu\text{m}$ 、約185  $\mu\text{m}$ 、約190  $\mu\text{m}$ 、約195  $\mu\text{m}$ 、約200  $\mu\text{m}$ 、約205  $\mu\text{m}$ 、約210  $\mu\text{m}$ 、約215  $\mu\text{m}$ 、約220  $\mu\text{m}$ 、約225  $\mu\text{m}$ 、約230  $\mu\text{m}$ 、約235  $\mu\text{m}$ 、約240  $\mu\text{m}$ 、約245  $\mu\text{m}$ 、約250  $\mu\text{m}$ 、約255  $\mu\text{m}$ 、約260  $\mu\text{m}$ 、約265  $\mu\text{m}$ 、約270  $\mu\text{m}$ 、約275  $\mu\text{m}$ 、約280  $\mu\text{m}$ 、約285  $\mu\text{m}$ 、約290  $\mu\text{m}$ 、約295  $\mu\text{m}$ 、約300  $\mu\text{m}$ 、約305  $\mu\text{m}$ 、約310  $\mu\text{m}$ 、約315  $\mu\text{m}$ 、約320  $\mu\text{m}$ 、約325  $\mu\text{m}$ 、約330  $\mu\text{m}$ 、約335  $\mu\text{m}$ 、約340  $\mu\text{m}$ 、約345  $\mu\text{m}$ 、約350  $\mu\text{m}$ 、約355  $\mu\text{m}$ 、約360  $\mu\text{m}$ 、約365  $\mu\text{m}$ 、約370  $\mu\text{m}$ 、約375  $\mu\text{m}$ 、約380  $\mu\text{m}$ 、約385  $\mu\text{m}$ 、約390  $\mu\text{m}$ 、約395  $\mu\text{m}$ 、約400  $\mu\text{m}$ 、約405  $\mu\text{m}$ 、約410  $\mu\text{m}$ 、約415  $\mu\text{m}$ 、約420  $\mu\text{m}$ 、約425  $\mu\text{m}$ 、約430  $\mu\text{m}$ 、約435  $\mu\text{m}$ 、約440  $\mu\text{m}$ 、約445  $\mu\text{m}$ 、約450  $\mu\text{m}$ 、約455  $\mu\text{m}$ 、約460  $\mu\text{m}$ 、約465  $\mu\text{m}$ 、約470  $\mu\text{m}$ 、約475  $\mu\text{m}$ 、約480  $\mu\text{m}$ 、約485  $\mu\text{m}$ 、約490  $\mu\text{m}$ 、約495  $\mu\text{m}$ 、約500  $\mu\text{m}$ 、または500  $\mu\text{m}$ 超であってもよい。いくつかの場合において、基材は、サイズが5  $\mu\text{m}$ ~500  $\mu\text{m}$ の範囲であるマイクロウェルを有してよい。いくつかの場合において、基材は、サイズが約5  $\mu\text{m}$ ~約500  $\mu\text{m}$ の範囲であるマイクロウェルを有してよい。いくつかの場合において、基材は、サイズが10  $\mu\text{m}$ ~100  $\mu\text{m}$ の範囲であるマイクロウェルを有してよい。いくつかの場合において、基材は、サイズが約10  $\mu\text{m}$ ~約100  $\mu\text{m}$ の範囲であるマイクロウェルを有してよい。いくつかの場合において、基材は、異なるサイズのタンパク質が異なるサイズのマイクロウェル中に分類され得るように、ある範囲の異なるサイズのマイクロウェルを有してよい。いくつかの場合において、基材中のマイクロウェルは、サイズによって分布されてよい(たとえば、より大きいマイクロウェルが第1の領域に分布され、かつより小さいマイクロウェルが第2の領域に分布される)。いくつかの場合において、基材は、約10の異なるサイズのマイクロウェルを有してよい。いくつかの場合において、基材は、約20の異なるサイズ、約25の異なるサイズ、約30の異なるサイズ、約35の異なるサイズ、約40の異なるサイズ、約45の異なるサイズ、約50の異なるサイズ、約55の異なるサイズ、約60の異なるサイズ、約65の異なるサイズ、約70の異なるサイズ、約75の異なるサイズ、約80の異なるサイズ、約85の異なるサイズ、約90の異なるサイズ、約95の異なるサイズ、約100の異なるサイズ、または100超の異なるサイズの、マイクロウェルを有してよい。

【0046】

いくつかの場合において、基材は、異なるサイズのナノウェルを有してよい。いくつかの場合において、ナノウェルは、約100ナノメートル( $\text{nm}$ )、約150  $\text{nm}$ 、約200  $\text{nm}$ 、約250  $\text{nm}$ 、約300  $\text{nm}$ 、約350  $\text{nm}$ 、約400  $\text{nm}$ 、約450  $\text{nm}$ 、約500  $\text{nm}$ 、約550  $\text{nm}$ 、約600  $\text{nm}$ 、約650  $\text{nm}$ 、約700  $\text{nm}$ 、約750  $\text{nm}$ 、約800  $\text{nm}$ 、約850  $\text{nm}$ 、約900  $\text{nm}$ 、約950  $\text{nm}$ であるか、または950  $\text{nm}$ ~1マイクロメートルの間であってよい。いくつかの場合において、基材は、サイズが100  $\text{nm}$ ~1マイクロメートルの範囲であるナノウェルを有してよい。いくつかの場合において、基材は、サイズが100  $\text{nm}$ ~500  $\text{nm}$ の範囲であるナノウェルを有してよい。いくつかの場合において、基材は、異なるサイズのタンパク質が異なるサイズのナノウェル中に分類され得るように、ある範囲の異なるサイズのナノウェルを有してよい。いくつかの場合において、基材中のナノウェルは、サイズによって分布されてよい(たとえば、より大

きいナノウェルが第1の領域に分布され、かつより小さいナノウェルが第2の領域に分布される)。いくつかの場合において、基材は、約10の異なるサイズのナノウェルを有してよい。いくつかの場合において、基材は、約20の異なるサイズ、または30超の異なるサイズのナノウェルを有してよい。

#### 【0047】

いくつかの場合において、基材は、異なるサイズのタンパク質が異なるサイズのナノウェルおよび/またはマイクロウェルに分類され得るように、ある範囲の異なるサイズのナノウェルおよび/またはマイクロウェルを有してよい。いくつかの場合において、基材中のナノウェルおよび/またはマイクロウェルは、サイズによって分布されてよい(たとえば、より大きいマイクロウェルが第1の領域に分布され、かつより小さいナノウェルが第2の領域に分布される)。いくつかの場合において、基材は、約10の異なるサイズのナノウェルおよび/またはマイクロウェルを有してよい。いくつかの場合において、基材は、約20の異なるサイズ、約25の異なるサイズ、約30の異なるサイズ、約35の異なるサイズ、約40の異なるサイズ、約45の異なるサイズ、約50の異なるサイズ、約55の異なるサイズ、約60の異なるサイズ、約65の異なるサイズ、約70の異なるサイズ、約75の異なるサイズ、約80の異なるサイズ、約85の異なるサイズ、約90の異なるサイズ、約95の異なるサイズ、約100の異なるサイズ、または100超の異なるサイズの、ナノウェルおよび/またはマイクロウェルを有してよい。

10

#### 【0048】

基材は、金属、ガラス、プラスチック、セラミック、またはそれらの組み合わせを含む、任意の素材を含んでよい。いくつかの好ましい態様において、固体基材はフローセルであり得る。フローセルは、単一の層、または複数の層から構成され得る。たとえばフローセルは、基部層(たとえばホウケイ酸ガラス製のもの)、基部層を覆うチャンネル層(たとえばエッチングされたシリコン製のもの)、およびカバー層または最上部層を含み得る。それらの層が共に組み立てられると、囲われたチャンネルが形成され得、該チャンネルは、カバーを通して両端に入口/出口を有する。それぞれの層の厚さは可変であるが、好ましくは約1700  $\mu$  未満である。層は、限定するものではないが、感光性ガラス、ホウケイ酸ガラス、溶融シリケート(fused silicate)、PDMSまたはシリコンを含む、当技術分野において公知の任意の適した素材から構成され得る。異なる層は、同じ素材から、または異なる素材から、構成され得る。

20

30

#### 【0049】

いくつかの態様において、フローセルは、フローセルの底部にチャンネルのための開口部を含み得る。フローセルは、別々に視覚化され得る位置において、何百万もの付着された標的コンジュゲーション部位を含み得る。いくつかの態様において、本発明の態様で使用されるさまざまなフローセルは、異なる数のチャンネル(たとえば1チャンネル、2以上のチャンネル、3以上のチャンネル、4以上のチャンネル、6以上のチャンネル、8以上のチャンネル、10以上のチャンネル、12以上のチャンネル、16以上のチャンネル、または16超のチャンネル)を含み得る。さまざまなフローセルは、異なる深さまたは幅のチャンネルを含み得、これらは1つのフローセル中のチャンネルの間で異なっていてよく、または異なるフローセルのチャンネルの間で異なっていてよい。1つのチャンネルはまた、深さおよび/または幅が変化し得る。たとえば1つのチャンネルは、チャンネル中の1つまたは複数の場所で、約50  $\mu$  未満の深さ、約50  $\mu$  の深さ、約100  $\mu$  未満の深さ、約100  $\mu$  の深さ、約100  $\mu$  から約500  $\mu$  の深さ、約500  $\mu$  の深さ、または約500  $\mu$  超の深さであり得る。チャンネルは、限定するものではないが、円形の、半円形の、長方形の、台形の、三角形の、または卵形の断面を含む、任意の断面形状を有し得る。

40

#### 【0050】

タンパク質は、基材に、スポットされてよく、滴下されてよく、ピペットされてよく、注がれてよく、洗浄されてよく、または他の方法で適用されてよい。NHSエステルなどのモエティで機能化された基材の場合には、タンパク質の修飾は必要とされない。代替の

50



モエティ（たとえばスルフヒドリル、アミン、またはリンカー核酸）で機能化された基材の場合には、架橋試薬（たとえばスベリン酸ジスクシンイミジル、NHS、スルホンアミド）が使用されてよい。リンカー核酸で機能化された基材の場合には、試料のタンパク質は、相補的な核酸タグで修飾されてよい。

#### 【0051】

いくつかの場合において、タンパク質は核酸にコンジュゲートされてよい。核酸を用いて核酸ナノボールが形成され得、それによりタンパク質に核酸ナノボールを連結させる。核酸ナノボールが基材に付着する場合、核酸に付着したタンパク質は、核酸ナノボールを介して基材に付着する。DNAナノボールは、基材に（たとえば吸着によってまたはコンジュゲーションによって）付着し得る。基材は、核酸ナノボールが付着できる、アミンで機能化された表面を有してよい。

10

#### 【0052】

いくつかの場合において、核酸ナノボールは、機能的に活性な末端（たとえばマレイミド、NHS-エステル等）を有して形成されてよい。タンパク質はその後ナノボールにコンジュゲートされてよく、それによりタンパク質に核酸ナノボールを連結させる。核酸ナノボールが基材に付着する場合、核酸に付着したタンパク質は、核酸ナノボールを介して基材に付着する。DNAナノボールは、基材に（たとえば吸着によってまたはコンジュゲーションによって）付着し得る。基材は、核酸ナノボールが付着できる、アミンで機能化された表面を有してよい。

20

#### 【0053】

光活性化架橋剤は、試料の架橋を基材上の特定の区域に指向させるために使用されてよい。光活性化架橋剤は、各試料を基材の既知の領域に付着させることによってタンパク質試料の多重化を可能にするために、使用されてよい。光活性化架橋剤は、たとえばタンパク質を架橋する前に蛍光タグを検出することによって、成功裏にタグ付けされたタンパク質の特異的な付着を可能にし得る。光活性化架橋剤の例は、限定するものではないが、N-5-アジド-2-ニトロベンゾイルオキシスクシンイミド、スルホスクシンイミジル6-(4'-アジド-2'-ニトロフェニルアミノ)ヘキサノアート、スクシンイミジル4,4'-アジペンタノアート、スルホスクシンイミジル4,4'-アジペンタノアート、スクシンイミジル6-(4,4'-アジペンタンアミド)ヘキサノアート、スルホスクシンイミジル6-(4,4'-アジペンタンアミド)ヘキサノアート、スクシンイミジル2-((4,4'-アジペンタンアミド)エチル)-1,3'-ジチオプロピオナート、およびスルホスクシンイミジル2-((4,4'-アジペンタンアミド)エチル)-1,3'-ジチオプロピオナートを含む。

30

#### 【0054】

試料はまた、各試料の結合を基材上の別々の区域に制限することによって、多重化されてよい。たとえば、基材はレーンに編成されてよい。多重化のための別の方法は、試料を反復的に基材の全域に適用することであり、各試料適用の後、非特異的なタンパク質結合試薬または色素を利用するタンパク質検出工程が続く。いくつかの場合において、色素の例は、SYPRO（登録商標）Ruby、SYPRO（登録商標）Orange、SYPRO（登録商標）Red、SYPRO（登録商標）Tangerine、およびCoomassie（商標）Fluor Orangeなどの、蛍光タンパク質ゲル染色剤を含んでよい。

40

#### 【0055】

試料の各添加後にすべてのタンパク質の位置を追跡することにより、基材上の各位置が最初にタンパク質を含んだ段階を決定することが可能であり、したがって当該タンパク質が由来する試料を決定することが可能である。この方法はまた、試料の各適用後の基材の飽和状態を決定し得、かつ基材上のタンパク質結合の最大化を可能にする。たとえば、機能化された位置の30%のみが、試料の第1の適用後にタンパク質によって占有されるとすると、その後、同じ試料の第2の適用、または異なる試料の適用のいずれかが、なされてよい。

#### 【0056】

ポリペプチドは、もう1つの残基によって基材に付着されてよい。いくつかの例におい

50

て、ポリペプチドは、N末端、C末端、両方の末端を介して、または内部の残基を介して、付着されてよい。

【0057】

永続的な架橋剤に加えて、いくつかの適用に関して光切断性リンカーを使用すること、およびそうすることによって分析後に基材からタンパク質を選択的に抽出することを可能にすることは、適切であり得る。いくつかの場合において、光切断性架橋剤は、いくつかの異なる多重化試料に使用されてよい。いくつかの場合において、光切断性架橋剤は、多重化された反応中の1つまたは複数の試料から使用されてよい。いくつかの場合において、多重化された反応は、永続的な架橋剤を介して基材に架橋された対照試料、および光切断性架橋剤を介して基材に架橋された実験試料を含んでよい。

10

【0058】

コンジュゲートされた各タンパク質は、コンジュゲートされた各タンパク質が光学的に解像可能であるように、コンジュゲートされた他のタンパク質それぞれから空間的に分離されてよい。タンパク質はしたがって、固有の空間的アドレスを用いて、個々にラベルされてよい。いくつかの態様において、これは、各タンパク質分子が他のタンパク質分子それぞれから空間的に分離されるように低い濃度のタンパク質および基材上の低い密度の付着部位を用いるコンジュゲーションによって、達成され得る。例として、光活性化架橋剤が使用される場合、タンパク質があらかじめ決定されている位置に付加されるように、光パターンが使用されてよい。

20

【0059】

いくつかの方法において、精製されたバルクタンパク質は、精製されたタンパク質を同定するために、基材にコンジュゲートされ、かつ本明細書において記載される方法を用いて処理されてよい。バルクタンパク質は、共に収集された、精製されたタンパク質を含んでよい。いくつかの例において、バルクタンパク質は、コンジュゲートされた各タンパク質または各バルクタンパク質が光学的に解像可能であるように、コンジュゲートされた他のタンパク質それぞれまたは他のバルクタンパク質それぞれから空間的に分離された位置にコンジュゲートされてよい。タンパク質またはバルクタンパク質は、したがって、固有の空間的アドレスを用いて個々にラベルされてよい。いくつかの態様において、これは、各タンパク質分子が他のタンパク質分子それぞれから空間的に分離されるように低い濃度のタンパク質および基材上の低い密度の付着部位を用いるコンジュゲーションによって、達成され得る。例として、光活性化架橋剤が使用される場合、1つまたは複数のタンパク質があらかじめ決定されている位置に付加されるように、光パターンが使用されてよい。

30

【0060】

いくつかの態様において、各タンパク質は、固有の空間的アドレスに関連付けられてよい。たとえば、タンパク質が空間的に分離される位置において基材に付着すると、各タンパク質には、座標などによる、インデックス付きアドレスが割り当てられ得る。いくつかの例において、あらかじめ割り当てられる固有の空間的アドレスの格子が、あらかじめ決定されていてよい。いくつかの態様において、各タンパク質の配置が基材上の固定マークに対して決定され得るように、基材は、容易に同定可能な固定マークを含んでよい。いくつかの例において、基材は、表面上に永続的に記された、格子線および/もしくはおよび「起点」または他の基準を有してよい。いくつかの例において、基材の表面は、架橋されたタンパク質の位置を特定するための基準を提供するために、永続的にまたは半永続的にマークが付けられていてよい。コンジュゲートされたポリペプチドの外縁などのパターンそれ自体の形状もまた、各スポットの固有の位置を決定するための基準として使用されてよい。

40

【0061】

基材はまた、コンジュゲートされたタンパク質標準物および対照を含んでよい。コンジュゲートされたタンパク質標準物および対照は、既知の位置にコンジュゲートされた、既知の配列のペプチドまたはタンパク質であってよい。いくつかの例において、コンジュゲートされたタンパク質標準物および対照は、アッセイにおける内部対照として役立ち得る

50

。該タンパク質は、精製されたタンパク質ストックから基材へと適用されてよく、または核酸プログラマブルタンパク質アレイ (Nucleic Acid-Programmable Protein Array) (NAPPA) などの処理によって基材上で合成されてよい。

#### 【0062】

いくつかの例において、基材は蛍光標準物を含んでよい。これらの蛍光標準物は、アッセイ間の蛍光シグナルの強度を校正するために使用されてよい。これらの蛍光標準物はまた、蛍光シグナルの強度を、ある区域に存在するフルオロフォアの数と相関させるために、使用されてもよい。蛍光標準物は、アッセイにおいて使用される異なる種類のフルオロフォアのいくつかまたはすべてを含んでよい。

#### 【0063】

#### 親和性試薬

基材に試料からのタンパク質がコンジュゲートされると、複数の親和性試薬測定を実施することが可能である。本明細書において記載される測定のプロセスは、さまざまな親和性試薬を利用してよい。

#### 【0064】

親和性試薬は、タンパク質またはペプチドに再現性のある特異性をもって結合する、任意の試薬であり得る。たとえば親和性試薬は、抗体、抗体断片、アプタマー、またはペプチドであってよい。いくつかの例において、モノクローナル抗体が選ばれ得る。いくつかの例において、Fab断片などの抗体断片が選ばれ得る。いくつかの場合において、親和性試薬は、市販の抗体などの市販の親和性試薬であってよい。いくつかの場合において、望ましい親和性試薬は、有用な特徴を有するものを同定するために市販の親和性試薬をスクリーニングすることによって、選択されてよい。いくつかの場合において、親和性試薬は、ただ1つのタンパク質に結合するそれらの能力に関してスクリーニングされてよい。いくつかの場合において、親和性試薬は、エピトープまたはアミノ酸配列に結合するそれらの能力に関してスクリーニングされてよい。いくつかの場合において、親和性試薬のグループは、差異のある結合によって、類似のタンパク質 (たとえば高度に類似の配列を有するもの) を集合的に区別する、それらの能力に関して、スクリーニングされてよい。いくつかの場合において、親和性試薬は、特定のタンパク質に対する結合特異性を増加させるため、重複した結合特徴を有するようスクリーニングされてよい。親和性試薬のスクリーニングは、多様な異なる様式で実施されてよい。一例は、NAPPAまたはエピトープタイリングアレイに対して親和性試薬をスクリーニングすることであり得る。いくつかの場合において、タンパク質標的に結合するよう設計されたタンパク質特異的親和性試薬が使用されてよい (たとえば市販の抗体またはアプタマー)。いくつかの場合において、複数のタンパク質特異的親和性試薬は、結合測定の前に混合されてよい。たとえば、結合測定の工程それぞれのために、タンパク質特異的親和性試薬の新しい混合物が、完全なセットから無作為に選択された、利用可能な親和性試薬のサブセットを含むように選択されてよい。たとえば、後続の混合物それぞれは、親和性試薬の多くが2つ以上の混合物中に存在することを期待して、同じ無作為な様式で生成されてよい。いくつかの場合において、タンパク質の同定は、タンパク質特異的親和性試薬の混合物を用いて、より速やかに達成されてよい。いくつかの場合において、タンパク質特異的親和性試薬のそのような混合物は、任意の個々の工程において、親和性試薬が結合する未知のタンパク質のパーセンテージを増加させ得る。親和性試薬の混合物は、すべての利用可能な親和性試薬の、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、またはそれより多くからなり得る。

#### 【0065】

親和性試薬は、高い、中間の、または低い特異性を有してよい。いくつかの例において、親和性試薬は、いくつかの異なるエピトープを認識してよい。いくつかの例において、親和性試薬は、2つ以上の異なるタンパク質に存在するエピトープを認識してよい。いくつかの例において、親和性試薬は、多くの異なるタンパク質に存在するエピトープを認識してよい。いくつかの場合において、本開示の方法において使用される親和性試薬は、ただ1つのエピトープに対して高度に特異的であってよい。いくつかの場合において、本開

10

20

30

40

50

示の方法において使用される親和性試薬は、翻訳後修飾を含むただ1つのエピトープに対して高度に特異的であってよい。

【0066】

いくつかの態様において、標的アミノ酸配列を同定することに指向された親和性試薬は、本明細書において記載される方法において使用されるように、互いに区別されることも互いに識別可能であることもない異なる構成要素のグループを実際に含んでよい。特に、同じ標的アミノ酸配列を同定するために使用され得る異なる構成要素は、同じ標的アミノ酸配列を同定するために同じ検出モエティを使用してよい。たとえば、隣接配列にかかわらず三量体アミノ酸配列(AAA)に結合する親和性試薬は、結合隣接配列からのいかなる影響も無く三量体AAA配列に結合するただ1つのプローブを含んでもよいし、または、AA A (式中 および は任意のアミノ酸であってよい)という形式の異なる5アミノ酸エピトープにそれぞれが結合する、400種のプローブのグループを含んでもよい。第2の場合のいくつかの場合においては、400種のプローブは、そのそれぞれが等量で存在するように組み合わせられてよい。第2の場合のいくつかの場合においては、400種のプローブは、どの所与の5アミノ酸エピトープも結合する確率が等しくなるよう、各プローブの特徴的な結合親和性によって各プローブの量が重みづけされ得るように、組み合わせられてよい。

【0067】

新規な親和性試薬は、当技術分野において公知の任意の方法によって作り出されてよい。親和性試薬を開発する方法は、SELEX、ファージディスプレイ、および接種を含む。いくつかの例において、親和性試薬は、構造に基づく薬剤設計方法を用いて設計されてよい。構造に基づく薬剤設計(または直接的な薬剤設計)は、関心対象のエピトープおよび親和性試薬の結合部位の三次元構造の知見を利用する。

【0068】

いくつかの場合において、親和性試薬は核酸バーコードで標識されてよい。いくつかの例において、核酸バーコードは、使用後に親和性試薬を精製するために使用されてよい。いくつかの例において、核酸バーコードは、繰り返しの使用のため親和性試薬を分類するために、使用されてよい。いくつかの場合において、親和性試薬は、使用後に親和性試薬を分類するために使用され得るフルオロフォアで標識されてよい。

【0069】

いくつかの場合において、核酸バーコードで標識された複数の親和性試薬は、多重化されてよく、その後相補的核酸プローブを用いて検出されてよい。親和性試薬の多重化されたグループは、別個の検出モエティを有する複数の相補的核酸を用いる単一のサイクルにおいて、検出されてよい。いくつかの場合において、親和性試薬の多重化されたグループは、検出モエティにコンジュゲートされた単一の相補的核酸を用いる複数のサイクルにおいて、検出されてよい。いくつかの場合において、親和性試薬の多重化されたグループは、別個の検出モエティにそれぞれコンジュゲートされた複数の相補的核酸を用いる複数のサイクルにおいて、検出されてよい。いくつかの場合において、親和性試薬の多重化されたグループは、検出モエティの別個のグループにそれぞれコンジュゲートされた複数の相補的核酸を用いる複数のサイクルにおいて、検出されてよい。

【0070】

いくつかの場合において、核酸バーコードで標識された1種または複数種の親和性試薬は、結合されたタンパク質に架橋されてよい。1種または複数種の親和性試薬がタンパク質に架橋されると、架橋された親和性試薬のアイデンティティを決定するために、バーコードが配列決定され得る。いくつかの場合において、複数の結合されたタンパク質が、1種または複数種の親和性試薬に曝露されてよい。いくつかの場合において、複数の結合されたタンパク質が、1種または複数種の親和性試薬と架橋される場合、結合した親和性試薬に関連付けられるバーコードは、複数の結合されたタンパク質のそれぞれに関連付けられる、架橋された親和性試薬のアイデンティティを決定するために、配列決定され得る。

【0071】

親和性試薬のファミリーは、親和性試薬の1つまたは複数の種類を含んでよい。たとえ

10

20

30

40

50

ば、本開示の方法は、抗体、抗体断片、Fab断片、アプタマー、ペプチド、およびタンパク質のうち1つまたは複数を含む親和性試薬のファミリーを使用してよい。

【0072】

親和性試薬は修飾されてよい。修飾は、限定するものではないが、検出モエティの付着を含む。検出モエティは、直接的にまたは間接的に付着されてよい。たとえば検出モエティは、親和性試薬に直接的に共有結合的に付着されてよく、またはリンカーを介して付着されてよく、または相補的な核酸タグもしくはビオチン・ストレプトアビジン対などの親和性反応を介して付着されてよい。軽い洗浄および親和性試薬の溶出に耐えることのできる付着方法が、選ばれ得る。

【0073】

検出モエティは、限定するものではないが、フルオロフォア、生物発光タンパク質、不変の領域およびバーコード領域を含む核酸セグメント、または磁気粒子などのナノ粒子を連結するための化学的テザーを含む。検出モエティは、励起または発光の異なるパターンを有する、いくつかの異なるフルオロフォアを含んでよい。

【0074】

検出モエティは、親和性試薬から切断可能であってよい。これは、もはや関心対象ではない親和性試薬から検出モエティが除去される工程によってシグナル混入を減少させることを可能にし得る。

【0075】

いくつかの場合において、親和性試薬は未修飾である。たとえば、もし親和性試薬が抗体であるならば、抗体の存在は原子間力顕微鏡によって検出されてよい。親和性試薬は未修飾であってよく、かつ、たとえば親和性試薬の1種または複数種に対して特異的な抗体を手に入れることによって、検出されてよい。たとえば、もし親和性試薬がマウス抗体であるならば、マウス抗体は、抗マウス二次抗体を用いて検出されてよい。交代で、親和性試薬は、アプタマーに特異的な抗体によって検出されるアプタマーであってよい。二次抗体は、上述のように検出モエティで修飾されてよい。いくつかの場合において、二次抗体の存在は、原子間力顕微鏡によって検出されてよい。

【0076】

いくつかの例において、親和性試薬は、同じ修飾、たとえばコンジュゲートされた緑色蛍光タンパク質を含んでよく、または異なる2種類以上の修飾を含んでよい。たとえば、各親和性試薬は、異なる励起波長または発光波長をそれぞれ有する、いくつかの異なる蛍光モエティの1つにコンジュゲートされてよい。いくつかの異なる親和性試薬は組み合わせられ得、かつ/または識別され得るので、これは、親和性試薬の多重化を可能にし得る。1つの例において、第1の親和性試薬は緑色蛍光タンパク質にコンジュゲートされてよく、第2の親和性試薬は黄色蛍光タンパク質にコンジュゲートされてよく、かつ第3の親和性試薬は赤色蛍光タンパク質にコンジュゲートされてよく、したがってこれら3つの親和性試薬は多重化され得、かつそれらの蛍光によって同定され得る。さらなる例において、第1、第4、および第7の親和性試薬は緑色蛍光タンパク質にコンジュゲートされてよく、第2、第5、および第8の親和性試薬は黄色蛍光タンパク質にコンジュゲートされてよく、かつ第3、第6、および第9の親和性試薬は赤色蛍光タンパク質にコンジュゲートされてよい；この場合、第1、第2、および第3の親和性試薬はともに多重化され得、一方で第2、第4、および第7の、ならびに第3、第6、および第9の親和性試薬は、2つのさらなる多重化反応を形成する。ともに多重化され得る親和性試薬の数は、それらを区別するために使用される検出モエティ次第で変わり得る。たとえば、フルオロフォアで標識された親和性試薬の多重化は、利用可能な独特のフルオロフォアの数によって制限され得る。さらなる例としては、核酸タグで標識された親和性試薬の多重化は、核酸バーコードの長さによって決定されてよい。

【0077】

各親和性試薬の特異性は、アッセイにおける使用の前に決定され得る。親和性試薬の結合特異性は、既知のタンパク質を用いる対照実験において決定され得る。任意の適切な実

10

20

30

40

50

験方法が、親和性試薬の特異性を決定するために使用されてよい。一例では、基材に、既知のタンパク質標準物を既知の位置に載せて、複数の親和性試薬の特異性を評価するために使用してよい。別の例においては、各親和性試薬の特異性が対照および標準物への結合から算出され得、その後実験試料を同定するために使用され得るように、基材は、実験試料ならびに対照および標準物のパネルの両方を含んでよい。いくつかの場合において、未知の特異性を有する親和性試薬が、既知の特異性の親和性試薬とともに含まれてよく、既知の特異性の親和性試薬からのデータが、タンパク質を同定するために使用されてよく、かつ未知の特異性の親和性試薬の、同定されるタンパク質への結合のパターンが、それらの結合特異性を決定するために使用されてよい。どのタンパク質が個々の親和性試薬と結合したのかを評価するために、他の親和性試薬の既知の結合データを用いて、任意の個々の親和性試薬の特異性を再確認することもまた、可能である。したがって、親和性試薬パネルの複数の使用により、親和性試薬の特異性は各反復のたびにますます洗練され得る。特定のタンパク質に対して一意的に特異的な親和性試薬が使用され得るが、本明細書において記載される方法は、それらを必要としなくてもよい。加えて方法は、ある範囲の特異性において効果的であり得る。いくつかの例において、本明細書において記載される方法は、親和性試薬が、いかなる特定のタンパク質に対しても特異的ではないが、代わりに、アミノ酸モチーフ（たとえばトリペプチドAAA）に対して特異的である場合に、特に有効であり得る。

10

**【0078】**

いくつかの例において、1種または複数種の親和性試薬は、所与の長さ、例えば2、3、4、5、6、7、8、9、10アミノ酸の、または10アミノ酸を超えるアミノ酸モチーフに結合するように、選ばれてよい。いくつかの例において、1種または複数種の親和性試薬は、2アミノ酸から40アミノ酸までの異なる長さの範囲のアミノ酸モチーフに結合するように、選ばれてよい。

20

**【0079】**

いくつかの例において、親和性試薬は、高い、中間の、または低い結合親和性を有するように選ばれてよい。いくつかの場合において、低いまたは中間の結合親和性を有する親和性試薬が選ばれ得る。いくつかの場合において、親和性試薬は、約 $10^{-3}$  M、 $10^{-4}$  M、 $10^{-5}$  M、 $10^{-6}$  M、 $10^{-7}$  M、 $10^{-8}$  M、 $10^{-9}$  M、 $10^{-10}$  Mの、またはそれより小さい解離定数を有してよい。いくつかの場合において、親和性試薬は、約 $10^{-10}$  M、 $10^{-9}$  M、 $10^{-8}$  M、 $10^{-7}$  M、 $10^{-6}$  M、 $10^{-5}$  M、 $10^{-4}$  M、 $10^{-3}$  M、 $10^{-2}$  Mを超える、またはそれより大きい解離定数を有してよい。

30

**【0080】**

親和性試薬のいくつかは、リン酸化されたまたはユビキチン化されたアミノ酸配列などの修飾されたアミノ酸配列に結合するために、選ばれてよい。いくつかの例において、1種または複数種の親和性試薬は、1種または複数種のタンパク質によって含まれ得るエピトープのファミリーに対し広く特異的であるように、選ばれてよい。いくつかの例において、1種または複数種の親和性試薬は、2つ以上の異なるタンパク質に結合してよい。いくつかの例において、1種または複数種の親和性試薬は、それらの1つまたは複数の標的に弱く結合してよい。たとえば親和性試薬は、10%未満、10%未満、15%未満、20%未満、25%未満、30%未満、35%未満、または35%未満が、それらの1つまたは複数の標的に結合し得る。いくつかの例において、1種または複数種の親和性試薬は、それらの1つまたは複数の標的に中程度にまたは強固に結合してよい。たとえば親和性試薬は、35%超、40%超、45%超、60%超、65%超、70%超、75%超、80%超、85%超、90%超、91%超、92%超、93%超、94%超、95%超、96%超、97%超、98%超、または99%超が、それらの1つまたは複数の標的に結合し得る。

40

**【0081】**

弱い結合を補うために、過剰な親和性試薬が、基材に適用されてよい。親和性試薬は、約1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1または10:1で、試料タンパク質に対して過剰に適用されてよい。親和性試薬は、約1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、

50

9:1、または10:1で、試料タンパク質におけるエピトープの予想される出現率に対して過剰に適用されてよい。

【0082】

親和性試薬はまた、磁気を帯びた構成要素を含んでよい。磁気を帯びた構成要素は、いくつかのまたはすべての結合親和性試薬を同じ画像面またはzスタックで操作するために、有用であり得る。いくつかのまたはすべての親和性試薬を同じ画像面で操作することは、画像化データの質を改善し得、システム中のノイズを減少させ得る。

【0083】

結合測定

修飾された親和性試薬のセットおよびコンジュゲートされた基材が与えられると、親和性試薬は、基材に反復的に適用されてよい。各測定サイクルは、いくつかの段階からなる。第1の段階において、親和性試薬が基材に適用され、該基材において親和性試薬はコンジュゲートされたタンパク質に吸着し得る。

10

【0084】

次に、非特異的結合を除去するために、基材は軽く洗浄され得る。この洗浄工程は、固定化されたタンパク質に結合した親和性試薬が溶出しない条件下で実施され得る。この工程で使用され得る緩衝液のいくつかの例は、リン酸緩衝生理食塩水、トリス緩衝生理食塩水、Tween20を添加したリン酸緩衝生理食塩水、およびTween20を添加したトリス緩衝生理食塩水を含む。

【0085】

20

吸着に続き、修飾された各親和性試薬の結合アドレスが、親和性試薬に直接的にコンジュゲートされているフルオロフォアの測定、または親和性試薬にコンジュゲートされている核酸鎖に対する相補核酸にコンジュゲートされているフルオロフォアの測定などによって決定される。検出方法は、検出モエティの選択によって決定される。フルオロフォアおよび生物発光モエティは光学的に検出され得、いくつかの場合において、二次検出試薬が必要とされる。基材上で固定化された各タンパク質の固有のアドレスは、結合測定の前に決定されてよく、または固定化されたタンパク質を含むアドレスのリストが、結合測定によって生成されてもよい。

【0086】

次に、親和性試薬は、よりストリンジェントな洗浄によって脱離され得る。この洗浄工程は、いくつかのまたはすべての親和性試薬を、固定化基材から除去し得る。いくつかの場合において、親和性試薬は除去を促進する目的で、低い～中間の結合親和性を有するように選ばれていてよい。使用された親和性試薬は、再使用のために再捕捉されてよく、または破棄されてもよい。例として、切断可能な検出モエティを有する親和性試薬が使用される場合、検出モエティは、この段階で切断され、かつ除去されてよい。ストリンジェントな洗浄に続き、いくつかの例において、残存する蛍光はどれも消光され得、さらによりストリンジェントな洗浄が、残りの親和性試薬を除去するために適用され得る。持ち越し/混入は、次の親和性試薬を適用する前に基材を再画像化することによって検出され得る。混入はまた、繰り返し生じるシグナルについて画像を連続してモニタリングすることによって検出されてもよい。これにより、分析の1サイクルが完結する。

30

40

【0087】

いくつかの態様において、蛍光タグ付けされた親和性試薬は、活性化波長の強い光への長時間の曝露により、消光されてよい。蛍光タグの消光は、親和性試薬を除去するための洗浄工程と置き換えてよい。いくつかの態様において、どのシグナルが先のn-1回のサイクルに由来するものだったのかを識別するためにn種のフルオロフォアを反復させることは、望ましいことであり得る。

【0088】

各親和性試薬のために、またはその多重化のために、サイクルを継続する。測定段階の結果は、各親和性試薬の結合座標、または各座標位置で結合した親和性試薬を列挙する、非常に大きな表であり、たとえば図10を参照のこと。

50

## 【0089】

## 分析

タンパク質同定の最後の工程は、座標に結合した親和性試薬についての情報から、基材の各座標における各タンパク質の、最も可能性が高いアイデンティティを決定するためのソフトウェアツールを含んでよい。ソフトウェアは、各親和性試薬の結合の特徴についての情報を利用してよい。たとえば、ある所与の親和性試薬が、トリペプチドエピトープAA Aを含むタンパク質に選択的に結合するとする。各親和性試薬の結合の特徴、試料中のタンパク質のデータベース、および結合した座標のリスト、結合のパターンについての情報が与えられると、ソフトウェアツールは、アイデンティティの信頼性ととも、蓋然的なアイデンティティを各座標に割り当てる。親和性試薬とタンパク質との間の精確な1-1マッピングという極端な場合には、これは、単純なルックアップテーブルで達成され得る。しかしながら、結合がより複雑な場合には、これは適切な充足問題を解くことによって実施されてよい。結合の特徴が高度に複雑な場合には、期待値最大化アプローチが採用されてよい。

10

## 【0090】

ソフトウェアはまた、各親和性試薬が結合しなかったいくつかのまたはすべての位置のリストを利用し得、存在するタンパク質を決定するために、エピトープの非存在についてのこの情報を使用し得る。ソフトウェアはまた、親和性試薬が各アドレスに結合したこと、および結合しなかったことについての情報を利用してよい。したがってソフトウェアは、エピトープが存在していたこと、およびエピトープが存在していなかったことの両方についての情報を使用し得る。ソフトウェアはデータベースを含んでよい。データベースは、試料が得られた種におけるいくつかのまたはすべての既知のタンパク質の配列を含んでよい。たとえば、試料がヒト由来のものであることが既知であるとする、いくつかのまたはすべてのヒトタンパク質の配列を有するデータベースが使用されてよい。試料の種が未知であるならば、いくつかのまたはすべてのタンパク質配列のデータベースが使用されてよい。データベースはまた、いくつかのまたはすべての既知のタンパク質変種および変異体タンパク質の配列、ならびにDNAフレームシフト変異から生じ得るいくつかのまたはすべての可能性のあるタンパク質の配列を含んでよい。データベースはまた、未成熟終止コドンからまたは分解から生じ得る、可能性のある切断型タンパク質の配列を含んでよい。

20

## 【0091】

ソフトウェアは、機械学習、深層学習、統計的学習、教師あり学習、教師なし学習、クラスタリング、期待値最大化、最尤推定、ベイズ推論、線形回帰、ロジスティック回帰、二項分類、多項分類、または他のパターン認識アルゴリズムなどの、1つまたは複数のアルゴリズムを含んでよい。たとえばソフトウェアは、(i)各親和性試薬の結合の特徴の情報、(ii)試料中のタンパク質のデータベースの情報、(iii)結合座標のリストの情報、および/または(iv)親和性試薬のタンパク質への結合のパターンの情報を(たとえば1つまたは複数のアルゴリズムの入力として)分析する1つまたは複数のアルゴリズムを、(a)各座標の蓋然的なアイデンティティならびに/または(b)信頼性(たとえばアイデンティティの信頼水準および/もしくは信頼区間)を(たとえば1つまたは複数のアルゴリズムの出力として)生成するまたは割り当てる目的のために、実施してよい。機械学習アルゴリズムの例は、サポートベクターマシン(SVM)、ニューラルネットワーク、畳み込みニューラルネットワーク(CNN)、深層ニューラルネットワーク、カスケードニューラルネットワーク、k-近傍(k-NN)分類、ランダムフォレスト(RF)、ならびに他の種類の分類木および回帰木(CART)を含んでよい。

30

40

## 【0092】

ソフトウェアは、各アドレスにおけるタンパク質のアイデンティティがあらかじめ決定されている基材上で本開示の方法を実施することにより、訓練されてよい。たとえばソフトウェアは、核酸プログラマブルタンパク質アレイ(Nucleic Acid-Programmable Protein Array)またはエピトープタイリングアレイを訓練用データセットとして用いて、訓練されてよい。

50



## 【0093】

試料の特徴の決定

デコーディングが完了すると、各アドレスにコンジュゲートされたタンパク質の蓋然的なアイデンティティが定義される。その結果として、混合物中のそれらの存在量が、計数観察により推定され得る。したがって、混合物中に存在する各タンパク質のリスト、およびそのタンパク質の観察数が、収集され得る。

## 【0094】

さらに、光切断性リンカー、または他の型の特異的に切断可能なリンカーが、タンパク質を基材に付着させるために使用されたならば、関心対象の特定のタンパク質は、その後基材から外され、さらなる調査のために収集されてよい。たとえば特定のタンパク質は、同定され、さらなる調査のために溶出されてよい。本開示の方法はまた、混合物から望ましいタンパク質を精製および/または単離する手段として役立ち得る。いくつかの場合において、該方法は、特定のアイソタイプまたは翻訳後修飾されたタンパク質を精製および/または単離することが可能であり得る。可能性のあるタンパク質および関連付けられる配列の完全なリストが利用可能ではない試料においては、本方法は、識別されるタンパク質グループ中の、異なるタンパク質を識別することが可能であり得、これらはその後さらなる調査のために溶出することができる。たとえば、腸マイクロバイオーーム試料などの、多くの未知のタンパク質を含む高度に複雑な試料については、本明細書において記載される方法は、質量分析の前に試料を分画するために使用されてよい。いくつかの場合において、タンパク質は、いったんそれらのアイデンティティが判定できたら、基材から溶出されてよい。タンパク質が同定されたときにそれらを基材から除去することにより、そのアイデンティティがまだ判定できないタンパク質のためにその後の親和性試薬結合のラウンドを続けることが可能となり、かつバックグラウンドノイズおよび残りのラウンドについての標的外のシグナルが低下し得る。いくつかの例において、特定のタンパク質に特異性を有する1種または複数種の親和性試薬が、血液試料中の血清アルブミンまたは免疫グロブリンなどの多量に存在するタンパク質を同定するために、第1ラウンドとして使用されてよく、これらの多量に存在するタンパク質は、処理の初期に除去されてよい。いくつかの場合において、基材上のタンパク質のサブセットは、親和性試薬結合の各ラウンドの後に、または親和性試薬結合の2、3、4、5、6、7、8、9、10、15、20ラウンドごと、もしくは20を超えるラウンドごとの後に、除去されてよい。信号対雑音比は、タンパク質溶出の各ラウンドの後に増加し得る。

## 【0095】

いくつかの場合において、未同定のタンパク質は、それらの結合パターンに基づき、グループ化またはクラスター化されてよい。たとえば、いくつかの場合において、試料中に存在するタンパク質は、配列データベース中に示されていないことがある。未同定のタンパク質は、各グループが試料中の同じ配列を有する未知のタンパク質のセットを含むことを目的として、親和性プローブへのそれらの結合パターンに基づいてグループへとクラスター化されてよい。タンパク質の量は各グループについて推定されてよく、かつ、限定するものではないが、健康な状態と疾患状態との間での差次的な定量化、縦断分析、またはバイオマーカー発見を含む定量的分析に含めてよい。いくつかの場合において、未同定のグループは、質量分析による同定のために、基材から選択的に外されてよい。他の場合においては、未同定のグループは、信頼性のある同定を生成するために特異的に設計されたさらなる結合親和性測定実験を実施することによって、同定されてよい。

## 【0096】

いくつかの場合において、タンパク質またはタンパク質のセットが除去された後で、基材に追加の試料を加えることが可能であってよい。たとえば血清アルブミンは、試料中のタンパク質の約半分を占め得る、血清中に多量に存在するタンパク質であり、親和性試薬結合の第1ラウンド後に血清アルブミンを除去することは、基材にさらなる血液試料を添加することを可能にし得る。いくつかの態様において、基材に試料を固定化する前に、たとえば免疫沈降または親和性カラム精製を介して、多量に存在するタンパク質を除去する

ことが選ばれ得る。

【0097】

タンパク質修飾は、本開示の方法を用いて同定され得る。たとえば翻訳後修飾は、酵素処理（たとえばホスファターゼ処理）を取り入れた、修飾に特異的な検出試薬を用いる検出の反復サイクルによって、同定されてよい。異なる修飾に特異的な親和性試薬が、固定化されたタンパク質におけるそのような修飾の非存在の存在を決定するために、使用されてよい。該方法はまた、各タンパク質が所与の修飾を有する例および有さない例の数の定量化を可能にする。

【0098】

タンパク質中の変異は、試料タンパク質の結合パターンと、予測されるタンパク質のアイデンティティとの間の不一致を照合することにより、検出されてよい。たとえば、1つの親和性試薬の結合を除いて既知のタンパク質の親和性試薬結合プロファイルと一致する、基材上の固定化されたタンパク質またはポリペプチドは、1アミノ酸置換を有し得る。親和性試薬は重複したエピトープを有し得るので、固定化されたタンパク質は、1アミノ酸置換だけを有していても、予測される親和性結合パターンとのいくつかの不整合を有し得る。未成熟終止コドンのフレームシフトを引き起こすDNA変異もまた、検出され得る。

【0099】

必要な親和性試薬の数は、試料中に存在するエピトープの総数より少なくてもよい。たとえば、仮に、各親和性試薬が1つの独特な3ペプチドエピトープを認識するように親和性試薬が選択された場合、試料中の、可能性のあるエピトープすべてを認識するための親和性試薬の総セットは、 $20 \times 20 \times 20 = 8000$ 種となる。しかしながら本開示の方法は、それら親和性試薬の、約100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、5500、または6000種だけを、必要とし得る。いくつかの場合において、該方法は、約500、1000、2500、3000、3500、4000、4500、5000、5500、または6000種未満の親和性試薬だけを必要とし得る。図13は、各親和性試薬の結合能力の関数としての、独特なアミノ酸3マーに特異的なx種の親和性試薬のセットが与えられた場合に同定され得る公知のヒトタンパク質のパーセンテージを実証するシミュレーションの結果を示す。図13に示されるように、ヒトタンパク質の98%が、8000種の3マー親和性試薬、および10%の結合可能性で、一意的に同定され得る。

【0100】

本開示の方法は、高度に正確であり得る。本開示の方法は、約90%、91%、92%、93%、94%、95%、96%、97%、98%、98.5%、99%、99.5%、99.9%、または99.9%超の正確性をもって各タンパク質を同定することが可能であり得る。

【0101】

本開示の方法は、約90%、91%、92%、93%、94%、95%、96%、97%、98%、98.5%、99%、99.5%、99.9%、または99.9%超の信頼性をもって、各タンパク質のアイデンティティを予測することが可能であり得る。信頼度は、試料中の異なるタンパク質で異なっていてよい。たとえば、独自性の非常に高い配列を有するタンパク質は、他のタンパク質に高度に類似したタンパク質と比べ、より高い信頼性をもって同定され得る。いくつかの場合において、タンパク質は、あるタンパク質ファミリーの一員として、高い信頼性をもって同定され得るが、しかしながら該タンパク質の厳密なアイデンティティは、より低い信頼性をもって予測されてよい。いくつかの場合において、極度に大きいまたは極度に小さいタンパク質は、より中程度のサイズのタンパク質と比べ、より低い信頼性をもって予測されてよい。

【0102】

いくつかの場合において、タンパク質は、あるタンパク質ファミリーの一員として、高い信頼性をもって同定され得るが、しかしながら該タンパク質の厳密なアイデンティティは、より低い信頼性をもって予測されてよい。たとえば、1つのアミノ酸変異を含むタンパク質は、当該タンパク質の標準型から、高い信頼性をもって区別することが困難であり得る。この場合は、標準配列も1つのアミノ酸変異を含む型も高い信頼性を有さない可能性があるが、しかし、未知のタンパク質が、両方の配列を含むタンパク質グループの一員

であることについて、高い信頼性が評価され得る。類似の状況が、タンパク質が、類似の配列を有する複数の関連するアイソフォームを有し得る例において、生じ得る。

【0103】

本開示の方法は、所与の試料中のいくつかのまたはすべてのタンパク質を同定することが可能であり得る。本開示の方法は、試料中のタンパク質の、約90%、91%、92%、93%、94%、95%、96%、97%、98%、98.5%、99%、99.5%、99.9%、または99.9%超を同定することが可能であり得る。

【0104】

本開示の方法は、試料中のタンパク質を速やかに同定することが可能であり得る。本開示の方法は、1日あたり、フローセル1つにつき約100個超、約1000個超、約5000個超、約10000個超、約20,000個超、約30,000個超、約40,000個超、約50,000個超、約100,000個超、1,000,000個超、約10,000,000個超、約100,000,000個超、約1,000,000,000個超、約10,000,000,000個超、約100,000,000,000個超、約1,000,000,000,000個超のタンパク質を同定することが可能であり得る。本開示の方法は、1日あたり、フローセル1つにつき約 $10^{10}$ 個超、 $10^{11}$ 個超、 $10^{12}$ 個超、 $10^{13}$ 個超、 $10^{14}$ 個超、 $10^{15}$ 個超、 $10^{16}$ 個超、 $10^{17}$ 個超、または約 $10^{17}$ 個超のタンパク質を同定することが可能であり得る。本開示の方法は、1日あたり、フローセル1つにつき約 $10^{10} \sim 10^{12}$ 個、 $10^{11} \sim 10^{14}$ 個、 $10^{12} \sim 10^{16}$ 個、または $10^{13} \sim 10^{17}$ 個のタンパク質を同定することが可能であり得る。本開示の方法は、1日あたり、フローセル1つにつき、約10 pg、約20 pg、約30 pg、約40 pg、約50 pg、約60 pg、約70 pg、約80 pg、約90 pg、約100 pg、約300 pg、約300 pg、約400 pg、約500 pg、約600 pg、約700 pg、約800 pg、約900 pg、約1 ng、約2 ng、約3 ng、約4 ng、約5 ng、約6 ng、約7 ng、約8 ng、約8 ng、約10 ng、約10 ng、約20 ng、約30 ng、約40 ng、約50 ng、約60 ng、約70 ng、約80 ng、約90 ng、約100 ng、約300 ng、約300 ng、約400 ng、約500 ng、約600 ng、約700 ng、約800 ng、約900 ng、約1  $\mu$ g、約2  $\mu$ g、約3  $\mu$ g、約4  $\mu$ g、約5  $\mu$ g、約6  $\mu$ g、約7  $\mu$ g、約8  $\mu$ g、約8  $\mu$ g、約10  $\mu$ g、約10  $\mu$ g、約20  $\mu$ g、約30  $\mu$ g、約40  $\mu$ g、約50  $\mu$ g、約60  $\mu$ g、約70  $\mu$ g、約80  $\mu$ g、約90  $\mu$ g、約100  $\mu$ g、約300  $\mu$ g、約300  $\mu$ g、約400  $\mu$ g、約500  $\mu$ g、約600  $\mu$ g、約700  $\mu$ g、約800  $\mu$ g、約900  $\mu$ gの、または約1 mg超のタンパク質中の、95%超のタンパク質を同定することが可能であり得る。

【0105】

本開示の方法は、実験的処置の後でプロテオームを評価するために使用され得る。本開示の方法は、治療的介入の効果を評価するために使用され得る。

【0106】

本開示の方法は、バイオマーカーの発見のために使用され得る。疾患を有する対象および疾患を有さない対象においてプロテオーム発現をモニターすることにより、バイオマーカーを同定し得る。疾患を発症する前の対象において、または疾患を発症するリスクのある対象において、プロテオーム発現をモニターすることにより、リスクを予測するバイオマーカーを同定し得る。対象のプロテオーム発現を評価することにより、対象の健康状態、またはある疾患もしくは障害を発症するリスクが示され得る。本開示の方法は、治療法を評価するために、または薬剤/治療法への応答者を非応答者から区別するために、使用され得る。本開示の方法は、個別化医療向けの特定の使用のためのものであり得る。

【0107】

本開示の方法は、疾患を診断するために使用され得る。異なる疾患または疾患段階は、タンパク質発現の異なるパネルに関連付けられてよい。タンパク質発現の異なるパネルが、所与の処置それぞれの異なる処置成果に関連付けられてよい。対象のプロテオーム発現データは、対象を診断するために、および/または最も適切な治療法を選択するために、使用され得る。

【0108】

本開示の方法は、試料が由来する個体または種を同定するために使用され得る。たとえば本開示の方法は、試料が、実際に主張される種または供給源に由来するのかどうかを決

定するために使用され得る。本明細書において記載される方法は、タンパク質が豊富だが核酸に乏しい試料において、PCRに基づく方法に対して優位性を有し得る。たとえば、蜂蜜試料の由来の同定。さらなる例としては、本開示の方法は、食品の安全性および食品の品質制御を評価するために使用され得る。

#### 【0109】

本開示の方法は、可能性のあるタンパク質の数よりも少ない親和性試薬を用いて、タンパク質分子のプールからの任意のタンパク質1分子を同定するために、使用され得る。たとえば該方法は、親和性試薬のパネルを用いて、n種の可能性のあるタンパク質のプールからの未同定のタンパク質1分子を、閾値を上回る確実性をもって同定してよく、ここでパネル中の親和性試薬の数はmであり、かつmはn未満である。未同定のタンパク質は、既知のタンパク質配列および遺伝子配列に対応する既知のタンパク質であってよく、または既知のタンパク質配列も遺伝子配列もない未知のタンパク質であってもよい。未知のタンパク質の場合、この方法は、未知のタンパク質のシグネチャーを同定し得、したがって、アミノ酸配列ではなく、未知のタンパク質の存在および量を同定し得る。本開示の方法は、n種の可能性のあるタンパク質のプールから選択される未同定のタンパク質を同定することができる、m種の親和性試薬のパネルを選択するために使用され得る。本明細書において開示される方法はまた、m種の結合試薬を用いて、タンパク質の混合物中のn種のタンパク質を一意的に同定しかつ定量化することができ、かつここで各タンパク質は、m種の結合試薬のサブセットによる結合の独特なプロファイルによって同定される。さらに、mは、nの約2分の1未満、3分の1未満、4分の1未満、5分の1未満、6分の1未満、7分の1未満、10分の1未満、20分の1未満、50分の1未満、または100分の1未満であってよい。さらなる例としては、本開示は、親和性試薬のパネルが、少なくとも約100、500、1000、2000、3000、4000、5000、6000、7000、8000、9000、10,000、12,000、14,000、16,000、18,000、20,000、25,000、30,000、35,000、40,000、50,000、100,000、200,000、300,000、400,000、500,000、600,000、700,000、800,000、900,000、1,000,000、2,000,000、3,000,000、4,000,000、または5,000,000個の異なるタンパク質のそれぞれを一意的に同定することができるよう、約100種未満、200種未満、300種未満、400種未満、500種未満、600種未満、700種未満、800種未満、900種未満、1000種未満、1500種未満、2000種未満、2500種未満、3000種未満、3500種未満、または4000種未満の親和性試薬のパネルを選択するために、使用されてよい。

#### 【0110】

本開示の方法は、プロテオーム中のタンパク質の大部分を同定することが可能であり得る。本開示の方法は、哺乳類の、鳥類の、魚類の、両生類の、爬虫類の、脊椎動物の、無脊椎動物の、植物の、菌類の、細菌の、または古細菌のプロテオーム中の、タンパク質の大部分を同定することが可能であり得る。本開示の方法は、プロテオーム中のタンパク質の、約5%超、10%超、15%超、20%超、25%超、30%超、35%超、40%超、45%超、50%超、55%超、60%超、65%超、70%超、75%超、80%超、85%超、90%超、95%超、96%超、97%超、98%超、99%超を同定することが可能であり得る。

#### 【実施例】

#### 【0111】

#### 実施例1：独特な3マーペプチドに結合する抗体を用いるタンパク質の同定

プロテオームにおけるすべてのエピトープのセットのカバレッジのパーセンテージと、本開示の方法を用いて同定され得るプロテオームのパーセンテージとの間の関係性を決定するために、コンピューターによる実験を実施した。この実験のため、すべての3マーアミノ酸エピトープのセットを選択した。タンパク質修飾は考慮しなかった。20種の天然のアミノ酸が存在するので、すべての3マーエピトープの総セットは、 $20 \times 20 \times 20 = 8000$ 種の、可能なエピトープとなる。シミュレーションのため、xを、1つの実験でスクリーニングされるエピトープの数として設定し、1から8000までのxの各値について、x種のエピトープのセットを無作為に選択し、同定され得るプロテオームのパーセンテージを算出した。図13は、このシミュレーションの結果を示す。

## 【 0 1 1 2 】

実施例2：独特な3マーペプチドに結合する抗体を用いるタンパク質の同定

可同定性およびカバレッジにおける親和性試薬の数の影響力を決定するために、コンピューターによるさらなる実験を実施した。タンパク質の、可能性のあるカバレッジそれぞれについて、同定され得るプロテオームのパーセンテージ（y軸）を示すために、親和性試薬プールのサイズの範囲についてデータ系列を算出した。その結果は表1に示される。たとえば、100アミノ酸のタンパク質は、98個の3マーアミノ酸エピトープ「着地点」を有し、仮にそれら3マーアミノ酸エピトープの20%が結合される場合に、それはそのタンパク質を同定するために十分であり得るかまたは十分ではない可能性がある。図15に示すように、250種の3マー特異的親和性試薬の親和性試薬プールにより、仮に各タンパク質の着地点の20%が結合される場合、プロテオームのわずかに約7%が同定され得る。8000種の親和性試薬の親和性試薬プールの場合には、20%の着地点が結合されると、プロテオームの約98%が同定され得る。

## 【 0 1 1 3 】

（表1）可同定性 対 プロテオームのカバレッジにおける、3マー-d-コードプローブ数の影響力

	8000	7000	6000	5000	4000	3000	2000	1000	500	250
1.00%	0.1825	0.135	0.0845	0.072	0.042	0.0125	0.0035	0	0	0
2.00%	0.492	0.41	0.3515	0.26	0.156	0.0985	0.037	0.0035	0.0005	0
3.00%	0.677	0.614	0.55	0.455	0.344	0.2175	0.0985	0.015	0.0005	0
4.00%	0.786	0.745	0.676	0.604	0.472	0.334	0.176	0.029	0.003	0
5.00%	0.843	0.811	0.765	0.7005	0.61	0.4765	0.269	0.054	0.0075	0
6.00%	0.9025	0.852	0.809	0.7645	0.6815	0.569	0.3485	0.092	0.012	0.0015
7.00%	0.9005	0.877	0.8435	0.81	0.7285	0.626	0.4345	0.1395	0.022	0.0025
8.00%	0.9275	0.9025	0.8875	0.835	0.782	0.678	0.491	0.192	0.034	0.002
9.00%	0.9415	0.923	0.898	0.8725	0.814	0.728	0.5495	0.221	0.0415	0.0065
10.00%	0.9575	0.941	0.919	0.8835	0.8535	0.751	0.601	0.261	0.0715	0.007
12.00%	0.9635	0.957	0.946	0.913	0.8825	0.81	0.663	0.3445	0.0955	0.0145
15.00%	0.978	0.969	0.962	0.9505	0.9185	0.8605	0.7675	0.443	0.1585	0.0295
17.00%	0.981	0.9765	0.9645	0.9575	0.927	0.884	0.8005	0.503	0.1915	0.0435
20.00%	0.9885	0.986	0.9725	0.9635	0.9575	0.9105	0.847	0.584	0.2525	0.0775
25.00%	0.99	0.9865	0.9785	0.9745	0.966	0.9445	0.8915	0.6955	0.357	0.1165
30.00%	0.9865	0.9895	0.985	0.9825	0.973	0.9625	0.9245	0.76	0.4355	0.1665
50.00%	0.9915	0.9915	0.9935	0.9895	0.9855	0.978	0.967	0.89	0.691	0.374

## 【 0 1 1 4 】

実施例3：基材上にコンジュゲートされた発光タンパク質分子

蛍光タンパク質試料であるフィコエリスリンを、インキュベーションチャンバー内で4度で4時間、NHS-エステルでコートされたカバースリップに直接的にコンジュゲートした。蛍光タンパク質試料をその後、Hamamatsu orca flash 4.0カメラを備えたLeica DMI8上で、300 msの露光時間を用いて画像化した。図16Aおよび16Bは、結果として得られたキャプチャー画像を示す（明確性のため色は反転した）。図16Aおよび16Bに示されるように、暗いスポットはそれぞれ、タンパク質の存在を示す蛍光シグナルの区域を表す。図16Bは図16Aの拡大図である。図16B中の矢印は、バックグラウンドノイズから明確に識別可能な、タンパク質を表すシグナルを示す。

## 【 0 1 1 5 】

第2のタンパク質試料である緑色蛍光タンパク質を変性させ、インキュベーションチャンバー内で4度で4時間、NHS-エステルでコートされたカバースリップに直接的にコンジュゲートした。最初の画像化は、ベースラインの残存蛍光を示さず、これは緑色蛍光タンバ

ク質の完全な変性を示した。タンパク質をその後、Alexa-Fluor 647が付着している抗ペプチド抗体とインキュベートした。抗ペプチド抗体をその後、0.1% Tween-20でリンスした。これをその後、Andor NEO sCMOSカメラを備えたNikon Eclipse TiでTIRFを用いて画像化した。図17は、結果として得られたキャプチャー画像を示す（明確性のため色は反転した）。

【0116】

#### 実施例4：タンパク質の同定

緑色蛍光タンパク質、RNASE1、LTF、およびGSTM1という、可能性のある4つのタンパク質のプロテオームを、図18に表わす。この例においては、このプロテオームからの未知のタンパク質1分子を、基材上のある場所にコンジュゲートする。この未知のタンパク質を、9つの異なる親和性試薬のパネルによって順次調べる。該9つの異なる親和性試薬のそれぞれは、異なるアミノ酸三量体 [ AAA、AAC、AAD、AEV、GDG、QSA、LAD、TRK、DGD ] を認識し、かつそれぞれは蛍光色素で標識されている。未知のタンパク質には親和性試薬DGD、AEV、LAD、GDG、およびQSAが結合することが決定される。このプロテオームの4つのタンパク質の配列の分析により、GFPのみが、これらの3アミノ酸モチーフの5つすべてを含むことが示され、これらのモチーフには、図18の配列において下線が引かれている。したがって、未知のタンパク質1分子がGFPタンパク質であることが決定される。

【0117】

#### コンピューター制御システム

本開示は、本開示の方法を遂行するためにプログラムされたコンピューター制御システムを提供する。図14は、タンパク質などの生体高分子を特徴付けし同定するためにプログラムされた、または他の方法で設定された、コンピューターシステム1401を示す。コンピューターシステム1401は、たとえば基材の固有の空間的アドレスでのシグナルを観察する工程；観察されたシグナルに基づき、固有の空間的アドレスにおける、生体高分子の一部分に連結された同定可能なタグの存在を決定する工程；生体高分子の一部分の特徴を決定するために、決定された同定可能なタグを、生体高分子の配列のデータベースと比較して評価する工程などの、試料を評価および分析する本開示のさまざまな局面を統制し得る。コンピューターシステム1401は、ユーザーの電子装置、または電子装置に対して遠隔に位置するコンピューターシステムであり得る。電子装置は携帯型電子装置であり得る。

【0118】

コンピューターシステム1401は中央処理装置（CPU、同じく本明細書における「プロセッサ」および「コンピュータープロセッサ」）1405を含み、これはシングルコアもしくはマルチコアプロセッサ、または並列処理のための複数のプロセッサであり得る。コンピューターシステム1401はまた、メモリーまたはメモリーロケーション1410（たとえばランダムアクセスメモリー、読み出し専用メモリー、フラッシュメモリー）、電子ストレージユニット1415（たとえばハードディスク）、1つまたは複数の他のシステムとの通信のための通信インターフェース1420（たとえばネットワークアダプター）、ならびにキャッシュ、他のメモリー、データストレージおよび/もしくは電子ディスプレイアダプターなどの周辺装置1425を含む。メモリー1410、ストレージユニット1415、インターフェース1420、および周辺装置1425は、マザーボードなどの通信バス（実線）を経由して、CPU 1405と通信する。ストレージユニット1415は、データを記憶するためのデータストレージユニット（またはデータレポジトリ）であり得る。コンピューターシステム1401は、通信インターフェース1420の補助を受けて、コンピューターネットワーク（「ネットワーク」）1430に機能的に連結され得る。ネットワーク1430は、インターネット（the Internet）、インターネット（an internet）および/もしくはエクストラネット、またはインターネット（the Internet）と通信するイントラネットおよび/もしくはエクストラネットであり得る。ネットワーク1430は、いくつかの場合において、電気通信および/またはデータネットワークである。ネットワーク1430は、クラウドコンピューティングなどの分散コンピューティングを可能にし得る1つまたは複数のコンピューターサーバーを含み得る。ネットワーク1430は、いくつかの場合において、コンピューターシステム1401の補助を受けて、コ

ンピューターシステム1401に連結される装置がクライアントまたはサーバーとして機能することを可能にし得るピアトゥピアネットワークを実装し得る。

【0119】

CPU 1405は、プログラムまたはソフトウェアの形で具現化され得る、機械で読み取り可能な指示のシーケンスを実行することができる。指示は、メモリー1410などのメモリーロケーションにおいて記憶されてよい。指示はCPU 1405に向けられ得、これは次いで、本開示の方法を遂行するために、CPU 1405をプログラムし得るまたは他の方法で設定し得る。CPU 1405により実施される作業の例は、フェッチ、デコード、実行、およびライトバックを含み得る。

【0120】

CPU 1405は、集積回路などの回路の一部であり得る。システム1401の1つまたは複数の他の構成要素は、該回路に含まれ得る。いくつかの場合において、該回路は特定用途向け集積回路（ASIC）である。

【0121】

ストレージユニット1415は、ドライバー、ライブラリー、および保存されたプログラムなどのファイルを記憶可能である。ストレージユニット1415は、たとえば、ユーザープリファレンスおよびユーザープログラムといったユーザーデータを記憶可能である。コンピュータシステム1401は、いくつかの場合において、イントラネットまたはインターネット（the Internet）を経由してコンピュータシステム1401と通信する遠隔のサーバーに位置するものなどの、コンピュータシステム1401の外部の、1つまたは複数の追加のデータストレージユニットを含み得る。

【0122】

コンピュータシステム1401は、ネットワーク1430を経由して、1つまたは複数の遠隔のコンピュータシステムと通信可能である。たとえば、コンピュータシステム1401は、ユーザーの遠隔のコンピュータシステムと通信可能である。遠隔のコンピュータシステムの例は、パーソナルコンピュータ（たとえば携帯型PC）、スレート型もしくはタブレット型PC（たとえばApple（登録商標）iPad、Samsung（登録商標）Galaxy Tab）、電話機、スマートフォン（たとえばApple（登録商標）iPhone、Androidが作動可能な装置、Blackberry（登録商標））、またはパーソナルデジタルアシスタントを含む。ユーザーは、ネットワーク1430を介してコンピュータシステム1401にアクセス可能である。

【0123】

本明細書において記載されるような方法は、たとえばメモリー1410上または電子ストレージユニット1415上などのコンピュータシステム1401の電子ストレージロケーション上に記憶された、機械（たとえばコンピュータプロセッサ）で実行可能なコードによって、遂行され得る。機械で実行可能なまたは機械で読み取り可能なコードは、ソフトウェアの形で提供され得る。使用の間に、コードはプロセッサ1405によって実行され得る。いくつかの場合において、コードはストレージユニット1415から引き出され得、そしてプロセッサ1405の速やかなアクセスのためにメモリー1410に記憶され得る。いくつかの状況においては、電子ストレージユニット1415は除外され得、かつ機械で実行可能な指示はメモリー1410に記憶される。

【0124】

コードはプリコンパイルされ得、かつコードを実行するために適合されたプロセッサを有する機械での使用のために設定され得るか、または実行時の最中にコンパイルされ得る。コードは、プリコンパイルされる様式または実行時にコンパイルされる（as-compiled）様式でコードを実行することを可能にするように選択され得るプログラミング言語で供給され得る。

【0125】

コンピュータシステム501などの、本明細書において提供されるシステムおよび方法の局面は、プログラミングの形で具現化され得る。技術のさまざまな局面が、機械で読み取り可能な媒体の一種類で運ばれるもしくは具現化される、典型的には機械（もしくはプ

10

20

30

40

50

ロセッサ)で実行可能なコードおよび/もしくは関連付けられるデータの形である「物品」または「製品」として、みなされてよい。機械で実行可能なコードは、メモリー(たとえば読み出し専用メモリー、ランダムアクセスメモリー、フラッシュメモリー)またはハードディスクのような、電子ストレージユニット上に記憶され得る。「ストレージ」タイプの媒体は、コンピューター、プロセッサ、もしくは同様のものの有形メモリー、もしくはさまざまな半導体メモリー、テープドライブ、ディスクドライブおよび同様のものなどの関連するそのモジュールの、任意のものまたはすべてを含み得、これらはソフトウェアプログラミングの際にいつでも非一過性ストレージを提供し得る。ソフトウェアのすべてまたは一部分は、インターネット(the Internet)またはさまざまな他の電気通信ネットワークを経由して、時々通信してよい。そのような通信は、たとえば、1つのコンピューターまたはプロセッサから別のものへ、たとえば、マネージメントサーバーまたはホストコンピューターからアプリケーションサーバーのコンピュータープラットフォームへ、ソフトウェアを読み込ませることを可能にし得る。したがって、ソフトウェア要素を運び得る別の種類の媒体は、ローカル装置間の物理的なインターフェースを通過して使用されるもの、有線のおよび光学の固定電話ネットワークを経由して使用されるもの、ならびにさまざまなエアリンクを通じて使用されるものなどの、光波、電波、ならびに電磁波を含む。有線もしくは無線リンク、光学リンク、または同様のものなどの、そのような波を運ぶ物理的な要素もまた、ソフトウェアを運ぶ媒体としてみなされてよい。本明細書において使用されるように、非一過性で有形の「ストレージ」媒体に限定されない限り、コンピューターでまたは機械で「読み取り可能な媒体」などの語は、実行のためのプロセッサへの指示の提供に関与する、任意の媒体を指す。

10

20

#### 【0126】

したがって、コンピューターで実行可能なコードなどの機械で読み取り可能な媒体は、限定するものではないが、有形のストレージ媒体、搬送波媒体、または物理的な伝達媒体を含む、多くの形をとり得る。不揮発性のストレージ媒体は、たとえば、任意のコンピューターまたは同様のものにおける任意のストレージ装置などや、図面において示されるデータベース等を実装するために使用され得るものなどの、光学もしくは磁気ディスクを含む。揮発性ストレージ媒体は、コンピュータープラットフォーム等のメインメモリーなどの、ダイナミックメモリーを含む。有形の伝達媒体は、コンピューターシステム中のバスを含む配線を含む、同軸ケーブル；銅線、および光ファイバーを含む。搬送波伝達媒体は、電気信号もしくは電磁信号の形、または無線周波(RF)での、および赤外線での(IR)データ通信の間に生成されるものなどの、音波もしくは光波の形をとってよい。コンピューターで読み取り可能な媒体の一般的な形態は、したがって、たとえば以下を含む：フロッピーディスク、フレキシブルディスク、ハードディスク、磁気テープ、任意の他の磁気媒体、CD-ROM、DVDもしくはDVD-ROM、任意の他の光学媒体、パンチカード紙テープ、穿孔のパターンを有する任意の他の物理的なストレージ媒体、RAM、ROM、PROMおよびEPROM、FLASH-EPROM、任意の他のメモリーチップもしくはカートリッジ、データもしくは指示を運ぶ搬送波、そのような搬送波を運ぶケーブルもしくはリンク、またはコンピューターが読み込み得るプログラミングコードおよび/もしくはデータからの、任意の他の媒体。コンピューターで読み取り可能な媒体のこれらの形態の多くが、実行のためにプロセッサへ1つまたは複数の指示の1つまたは複数のシーケンスを運ぶことに関与し得る。

30

40

#### 【0127】

コンピューターシステム1401は、ユーザーインターフェース(UI)1440を含む電子ディスプレイ1435を含み得る、またはこれと通信し得る。UIの例は、限定するものではないが、グラフィカルユーザーインターフェース(GUI)およびウェブベースのユーザーインターフェースを含む。

#### 【0128】

本開示の方法およびシステムは、1つまたは複数のアルゴリズムによって遂行され得る。アルゴリズムは、中央処理装置1405による実行に際し、ソフトウェアによって遂行され得る。アルゴリズムは、たとえばタンパク質の一部分などの、生体高分子の一部分の特徴

50



および/またはアイデンティティを決定することができる。たとえばアルゴリズムは、候補のタンパク質の一部などの候補の生体高分子の一部の最も可能性のあるアイデンティティを決定するために、使用されてよい。

【0129】

いくつかの態様において、多くの異なるタンパク質中に存在する短いエピトープを認識するアプタマーまたはペプタマーは、デジタルアプタマーまたはデジタルペプタマーと称されてよい。本発明の1つの局面は、デジタルアプタマーまたはデジタルペプタマーのセットを提供し、ここで該セットは少なくとも約15種のデジタルアプタマーまたはデジタルペプタマーを含み、該15種のデジタルアプタマーまたはデジタルペプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーまたはデジタルペプタマーは、該デジタルアプタマーまたはデジタルペプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する。いくつかの態様において、デジタルアプタマーまたはデジタルペプタマーのセットは、連続した3アミノ酸からなるエピトープに結合する、100種のデジタルアプタマーまたはデジタルペプタマーを含む。いくつかの態様において、デジタルアプタマーまたはデジタルペプタマーのセットは、連続した4アミノ酸からなるエピトープに結合する、100種のデジタルアプタマーをさらに含む。いくつかの態様において、デジタルアプタマーまたはデジタルペプタマーのセットは、連続した5アミノ酸からなるエピトープに結合する、100種のデジタルアプタマーまたはデジタルペプタマーをさらに含む。いくつかの場合において、デジタル親和性試薬は、抗体、アプタマー、ペプタマー、ペプチド、またはFab断片であってよい。

【0130】

いくつかの態様において、デジタルアプタマーのセットは少なくとも約20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、または1000種のデジタルアプタマーを含む。いくつかの態様において、デジタルアプタマーのセットは、連続した4アミノ酸からなるエピトープに結合する、少なくとも1000種のデジタルアプタマーを含む。いくつかの態様において、デジタルアプタマーのセットは、連続した5アミノ酸からなるエピトープに結合する、少なくとも100種のデジタルアプタマーをさらに含む。デジタルアプタマーのセットは、連続した3アミノ酸からなるエピトープに結合する、少なくとも100種のデジタルアプタマーをさらに含む。いくつかの態様において、デジタルアプタマーのセットは表面上に固定化される。いくつかの態様において、表面はアレイである。

【0131】

別の局面において、本発明は、複数の異なるタンパク質を含む試料のタンパク質結合プロファイルを生成するための方法を提供し、該方法は、以下の工程を含む：結合を可能にする条件下で、試料を、デジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；任意で、未結合のタンパク質を除去する工程；およびタンパク質と該デジタルアプタマーとの結合を検出する工程であって、それによって該試料のタンパク質結合プロファイルが生成される、工程。

【0132】

いくつかの態様において、該方法は、結合を可能にする条件下で試料をデジタルアプタマーのセットに接触させる工程(a)の前に、試料をタンパク質切断剤で処理する工程を、さらに含む。

【0133】

別の局面において、本発明は、それぞれが複数のタンパク質を含む2つ以上の異なる試料のためのタンパク質結合プロファイルのライブラリーを含み、該方法は以下の工程を含

む：結合を可能にする条件下で、試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；任意で、未結合のタンパク質を除去する工程；タンパク質とデジタルアプタマーとの結合を検出することによって、試験されている試料のタンパク質結合プロファイルを生成する工程であって、それによってタンパク質結合プロファイルが生成される、工程；および少なくとも2つの試料を用いて上記工程を繰り返す工程。

【0134】

いくつかの態様において、該方法は、結合を可能にする条件下で試料をデジタルアプタマーのセットに接触させる工程の前に、試料をタンパク質切断剤で処理する工程をさらに含む。

【0135】

別の局面において、本発明は、試験試料を特徴付けするための方法を含み、該方法は以下の工程を含む：結合を可能にする条件下で、試験試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；任意で、未結合のタンパク質を除去する工程；タンパク質とデジタルアプタマーとの結合を検出することによって、試験試料のタンパク質結合プロファイルを生成する工程；および試験試料を特徴付けするために、試験試料の生成されたタンパク質結合プロファイルを、参照試料のタンパク質結合プロファイルと比較する工程。

【0136】

別の局面において、本発明は、試験試料中の細菌、ウイルス、または細胞の有無を決定するための方法を含み、該方法は以下の工程を含む：結合を可能にする条件下で、試験試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；任意で、未結合のタンパク質を除去する工程；タンパク質とデジタルアプタマーとの結合を検出することによって、試験試料のタンパク質結合プロファイルを生成する工程であって、それによってタンパク質結合プロファイルが生成される、工程；および試験試料のタンパク質結合プロファイルを参照試料のタンパク質結合プロファイルと比較する工程であって、それによって試験試料中の細菌、ウイルス、または細胞の有無が比較により決定される、工程。

【0137】

別の局面において、本発明は、試料中の試験タンパク質を同定するための方法を含み、該方法は以下の工程を含む：試験タンパク質を含む、または試験タンパク質を含むことが疑われる試料を、少なくとも約15種のデジタルアプタマーを含むデジタルアプタマーのセットに接触させる工程であって、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；および試験タンパク質とデジタルアプタマーのセットとの結合を検出することによって、試験タンパク質のアイデンティティを決定する工程であって、少なくとも約6種のデジタルアプタマーが試験タンパク質に結合し；かつ結合の存在が、試験タンパク質中の少なくとも約6種のエピトープの存

10

20

30

40

50

在を示し、少なくとも約6種のエピトープのアイデンティティが、試験タンパク質を同定するために使用される、工程。

【0138】

本発明の好ましい態様は本明細書において示されかつ記載されているが、そのような態様が、単なる例として提供されていることは、当業者には明らかである。本発明が、本明細書中に提供される特定の例によって限定されることを目的としているわけではない。本発明は上述の明細書を参照して記載されているが、本明細書における態様の説明および図面は、制限的な意味に解釈されるべきではない。無数の変更、改変、および置き換えが、本発明から逸脱することなく、今や当業者に想起されるであろう。さらに、本発明のすべての局面は、多様な条件および変数に左右される、本明細書において説明される特定の描写、構成、または相対的比率に限定されないことが、理解されるべきである。本明細書において記載される本発明の態様のさまざまな代替物が、本発明を実践する際に採用され得ることが、理解されるべきである。したがって、本発明はまた、任意のそのような代替物、修正形態、変更形態、または等価物を包含することが、企図される。添付の特許請求の範囲は、本発明の範囲を定義し、かつこれらの特許請求の範囲の範囲内の方法および構成、ならびにそれらの等価物がそれにより包含されることを、意図する。

【0139】

添付の特許請求の範囲にかかわらず、本明細書において説明される開示はまた、以下の条項によって定義される：

1. 少なくとも約15種のデジタルアプタマーを含む、デジタルアプタマーのセットであって、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、デジタルアプタマーのセット。 20
2. 連続した3アミノ酸からなるエピトープに結合する100種のデジタルアプタマーを含む、条項1にしたがうデジタルアプタマーのセット。
3. 連続した4アミノ酸からなるエピトープに結合する100種のデジタルアプタマーをさらに含む、条項1にしたがうデジタルアプタマーのセット。
4. 連続した5アミノ酸からなるエピトープに結合する100種のデジタルアプタマーをさらに含む、条項3にしたがうデジタルアプタマーのセット。 30
5. 少なくとも約20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、または1000種のデジタルアプタマーを含む、条項1にしたがうデジタルアプタマーのセット。
6. 連続した4アミノ酸からなるエピトープに結合する少なくとも1000種のデジタルアプタマーを含む、条項1にしたがうデジタルアプタマーのセット。
7. 連続した5アミノ酸からなるエピトープに結合する少なくとも100種のデジタルアプタマーをさらに含む、条項6にしたがうデジタルアプタマーのセット。
8. 連続した3アミノ酸からなるエピトープに結合する少なくとも100種のデジタルアプタマーをさらに含む、条項7にしたがうデジタルアプタマーのセット。
9. デジタルアプタマーが表面上に固定化される、条項1～8のいずれかにしたがうデジタルアプタマーのセット。 40
10. 表面がアレイである、条項9にしたがうデジタルアプタマーのセット。
11. 以下の工程を含む、複数の異なるタンパク質を含む試料のタンパク質結合プロファイルを生成するための方法：
  - a) 結合を可能にする条件下で、試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程； 50

- b) 任意で、未結合のタンパク質を除去する工程；および
- c) タンパク質と該デジタルアプタマーとの結合を検出する工程であって、それによって該試料のタンパク質結合プロファイルが生成される、工程。
12. 結合を可能にする条件下で試料をデジタルアプタマーのセットに接触させる工程(a)の前に、該試料をタンパク質切断剤で処理する工程をさらに含む、条項11の方法。
13. 以下の工程を含む、それぞれが複数のタンパク質を含む2つ以上の異なる試料のためのタンパク質結合プロファイルのライブラリーを生成するための方法：
- a) 結合を可能にする条件下で、試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；
- b) 任意で、未結合のタンパク質を除去する工程；
- c) タンパク質とデジタルアプタマーとの結合を検出することによって、試験されている試料のタンパク質結合プロファイルを生成する工程であって、それによってタンパク質結合プロファイルが生成される、工程；および
- d) 少なくとも2つの試料を用いて工程(a)～(c)を繰り返す工程。
14. 結合を可能にする条件下で試料をデジタルアプタマーのセットに接触させる工程(a)の前に、該試料をタンパク質切断剤で処理する工程をさらに含む、条項13の方法。
15. 条項13の方法を用いて調製されるタンパク質結合プロファイルのライブラリー。
16. 以下の工程を含む、試験試料を特徴付けするための方法：
- a) 結合を可能にする条件下で、試験試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；
- b) 任意で、未結合のタンパク質を除去する工程；
- c) タンパク質とデジタルアプタマーとの結合を検出することによって、試験試料のタンパク質結合プロファイルを生成する工程；および
- d) 試験試料を特徴付けするために、試験試料の生成されたタンパク質結合プロファイルを、参照試料のタンパク質結合プロファイルと比較する工程。
17. 以下の工程を含む、試験試料中の細菌、ウイルス、または細胞の有無を決定するための方法：
- a) 結合を可能にする条件下で、試験試料をデジタルアプタマーのセットに接触させる工程であって、デジタルアプタマーのセットは少なくとも約15種のデジタルアプタマーを含み、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；
- b) 任意で、未結合のタンパク質を除去する工程；
- c) タンパク質とデジタルアプタマーとの結合を検出することによって、試験試料のタンパク質結合プロファイルを生成する工程であって、それによってタンパク質結合プロファイルが生成される、工程；および
- d) 試験試料のタンパク質結合プロファイルを参照試料のタンパク質結合プロファイルと比較する工程であって、それによって、試験試料中の細菌、ウイルス、または細胞の有無が比較により決定される、工程。
18. 以下の工程を含む、試料中の試験タンパク質を同定するための方法：
- a) 試験タンパク質を含む、または試験タンパク質を含むことが疑われる試料を、少な

くとも約15種のデジタルアプタマーを含むデジタルアプタマーのセットに接触させる工程であって、該15種のデジタルアプタマーのそれぞれは、連続した3または4または5アミノ酸からなる異なるエピトープに特異的に結合することが特徴付けられており、かつ各デジタルアプタマーは、該デジタルアプタマーが結合する同じエピトープを含む、複数の別個の異なるタンパク質を認識する、工程；および

b) 試験タンパク質とデジタルアプタマーのセットとの結合を検出することによって、試験タンパク質のアイデンティティを決定する工程であって、少なくとも約6種のデジタルアプタマーが試験タンパク質に結合し；かつ結合の存在が、試験タンパク質中の少なくとも約6種のエピトープの存在を示し、少なくとも約6種のエピトープのアイデンティティが、試験タンパク質を同定するために使用される、工程。

10

19. 以下の工程を含む、タンパク質の特徴を決定する方法：

基材を得る工程であって、（分子レベルで）個々のタンパク質一部分がそれぞれ、光学的に解像可能な固有の空間的アドレスを有するように、1種または複数種のタンパク質の一部分が基材にコンジュゲートされている、工程；

1種または複数種の親和性試薬の第1～（順序づけられた）第nのセットを含む流体を、基材へ適用する工程であって、1種または複数種の親和性試薬のそれぞれは、1種または複数種のタンパク質の一部分のエピトープ（連続的なまたは非連続的なアミノ酸配列）1種に特異的であり、かつ1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は同定可能なタグに連結されている、工程；

1種または複数種の親和性試薬の第1およびそれに続く第nまでのセットの基材への各適用の後で、以下の工程を実施する工程：

20

同定可能なタグを観察する工程；

1つまたは複数の観察シグナルを有する、基材の1つまたは複数の固有の空間的アドレスを同定する工程；

同定された固有の空間的アドレスを有する1種または複数種のタンパク質の一部分がそれぞれ、1つまたは複数の観察シグナルに関連付けられる1種または複数種のエピトープを含むことを決定する工程；ならびに

タンパク質一部分それぞれの特徴を、1種または複数種のエピトープに基づいて決定する工程。

20. 以下の工程を含む、タンパク質の特徴を決定する方法：

30

基材を得る工程であって、各位置が、タンパク質1つ、またはタンパク質の少なくとも60%が同じアミノ酸配列を共有する該タンパク質のプールを含む、複数の位置を基材が有するように、1種または複数種のタンパク質の一部分が該基材にコンジュゲートされる、工程；

1種または複数種の親和性試薬の第1～（順序づけられた）第nのセットを含む流体を基材に適用する工程であって、1種または複数種の親和性試薬のそれぞれは、1種または複数種のタンパク質の一部分のエピトープ（連続的なまたは非連続的なアミノ酸配列）1種に特異的であり、かつ1種または複数種の親和性試薬の第1～第nのセットの各親和性試薬は同定可能なタグに連結されている、工程；

1種または複数種の親和性試薬の第1およびそれに続く第nまでのセットの基材への各適用の後で、以下の工程を実施する工程：

40

同定可能なタグを観察する工程；

1つまたは複数の観察シグナルを有する、基材の1つまたは複数の固有の空間的アドレスを同定する工程；

同定された固有の空間的アドレスを有する1種または複数種のタンパク質の一部分がそれぞれ、1つまたは複数の観察シグナルに関連付けられる1種または複数種のエピトープを含むことを決定する工程；ならびに

タンパク質一部分それぞれの特徴を、1種または複数種のエピトープに基づいて決定する工程。

21. 少なくとも400個の異なるタンパク質を、質量分析計からのデータに依存するタンバ

50

ク質同定のための技術と比べて少なくとも10%より迅速に同定するために使用され得る、条項19または20の方法。

22. 少なくとも400個の異なるタンパク質を少なくとも50%の正確性で同定する、条項21の方法。

23. 方法が、ある特定のタンパク質それ自体を10%超の信頼度閾値内で同定するかどうかという点とは無関係に、該特定のタンパク質をある特定のタンパク質ファミリーのメンバーとして同定する、条項22の方法。

24. 1種または複数種のタンパク質の一部が、タンパク質のサイズに基づいて基材上で分離されている、条項19または20の方法。

25. 1種または複数種のタンパク質の一部が、タンパク質の電荷に基づいて基材上で分離されている、条項19または20の方法。

26. 基材がマイクロウェルを含む、条項19または20の方法。

27. 基材が異なるサイズのマイクロウェルを含む、条項19または20の方法。

28. タンパク質が、ピオチン付着を介して基材に付着される、条項19または20の方法。

29. タンパク質が、核酸を介して基材に付着される、条項19または20の方法。

30. タンパク質が、核酸ナノボールを介して基材に付着される、条項29の方法。

31. タンパク質が、ナノビーズを介して基材に付着される、条項19または20の方法。

32. 1種または複数種のタンパク質の一部が結合する基材を得る工程が、官能基の規則正しいアレイを有する基材を得る工程、および各官能基が試料からのタンパク質1分子だけにコンジュゲートするように、タンパク質試料を適用する工程を含む、条項19または20の方法。

33. 官能基の規則正しいアレイを有する基材を得る工程が、フォトリソグラフィ、ディップペンナノリソグラフィ、ナノインプリントリソグラフィ、ナノスフェアリソグラフィ、熱走査型プローブリソグラフィ、局所酸化ナノリソグラフィ、分子自己集合、ステンスルリソグラフィ、および電子線リソグラフィからなる群から選択される方法を用いる工程を含む、条項32の方法。

34. 各官能基が、他の官能基それぞれから少なくとも約300 nm離れて配置される、条項32の方法。

35. 基材が、異なるサイズのマイクロウェルの規則正しいアレイを含む、条項19または20の方法。

36. 基材を得る工程が、タンパク質の第1の試料を基材にコンジュゲートする工程、第1の試料からのタンパク質が結合した各位置を検出するためにタンパク質色素を用いる工程、第2の試料をコンジュゲートする工程、および第2の試料からのタンパク質が結合した各位置を検出するためにタンパク質色素を用いる工程を含む、条項19または20の方法。

37. 基材を得る工程が、タンパク質の第1の試料を基材にコンジュゲートする工程、第1の試料からのタンパク質が結合した各位置を検出するためにタンパク質色素を用いる工程、結合したタンパク質の数から、タンパク質が結合していない基材上の官能基の割合を決定する工程を含む、条項19または20の方法。

38. 少なくとも1つの構成要素が、隣接配列を問わないコア配列のいずれの例の結合閾値をも上回る結合親和性を有するように、親和性試薬が、異なる隣接配列を有する同じコア配列に結合する構成要素のプールを含み得る、条項19または20の方法。

10

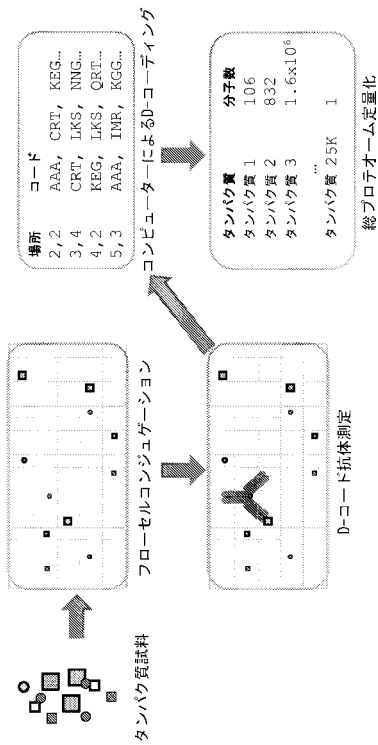
20

30

40

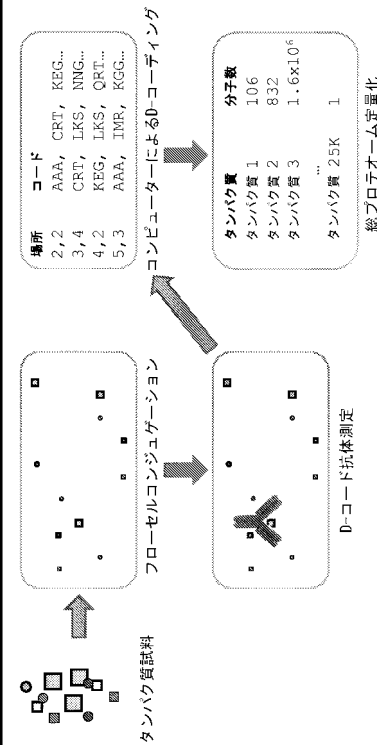
【図 1】

抗ペプチド抗体デコーディングによるタンパク質定量化



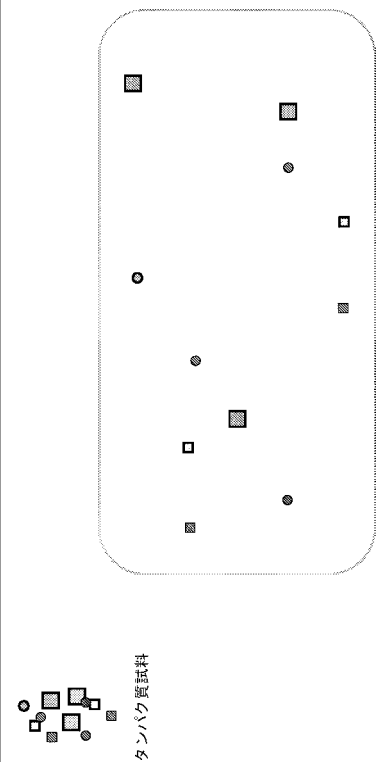
【図 2】

抗ペプチド抗体デコーディングによるタンパク質定量化



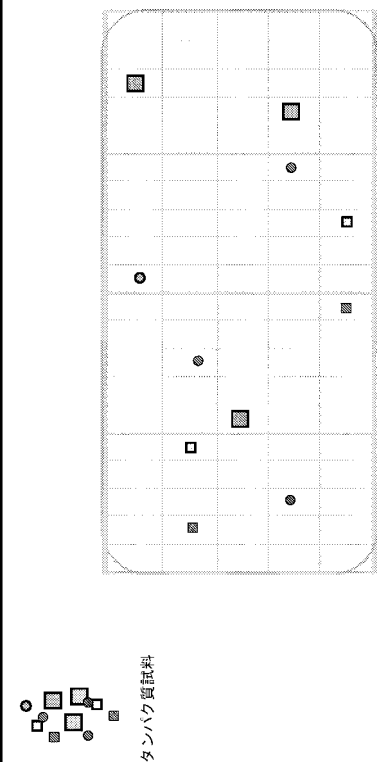
【図 3】

工程2：フローセルコンジュゲーション



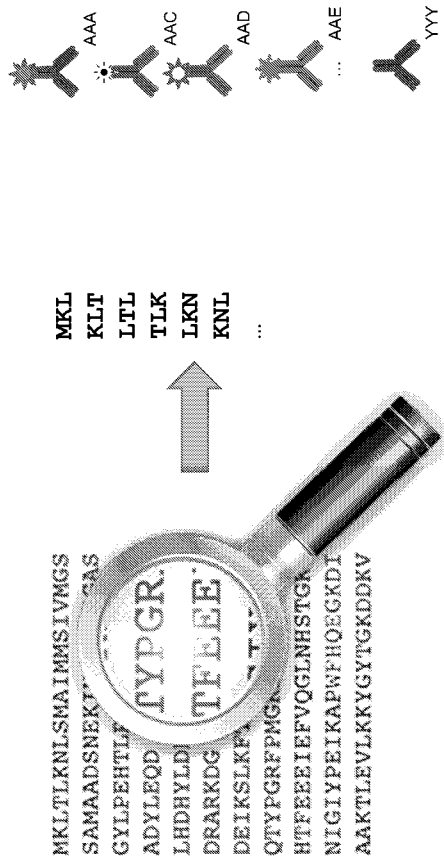
【図 4】

抗ペプチド抗体デコーディングによるタンパク質定量化

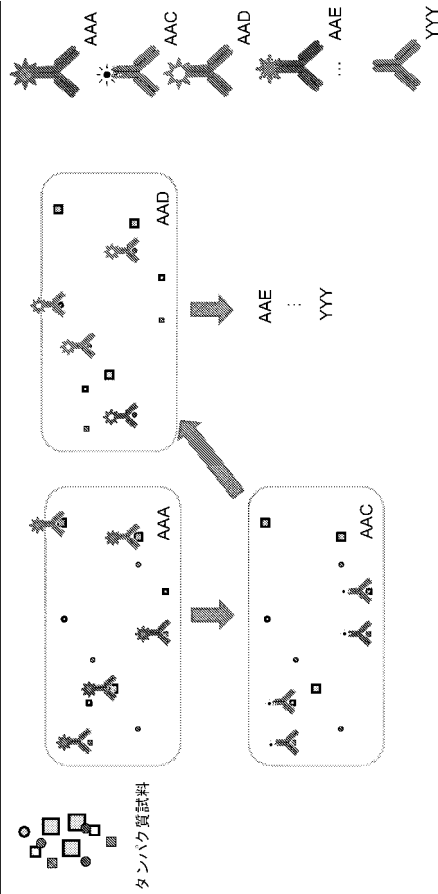


α-コラーゲン抗体と照合され得るペプチドのセットとしての  
タンパク質の解体

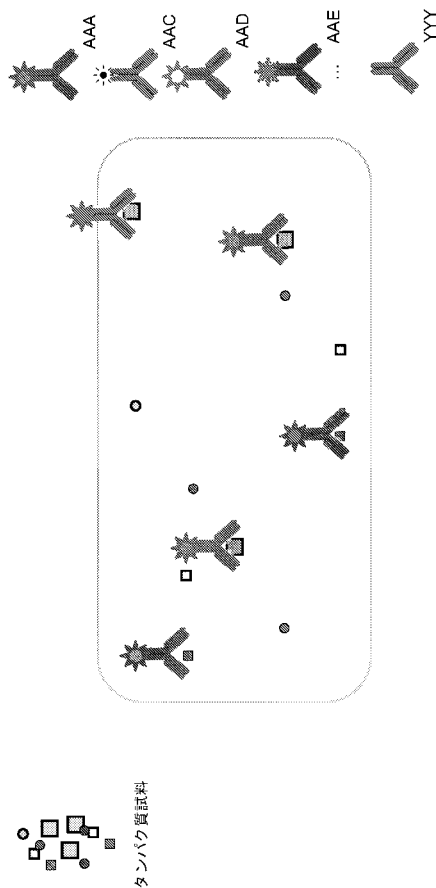
【図 5】



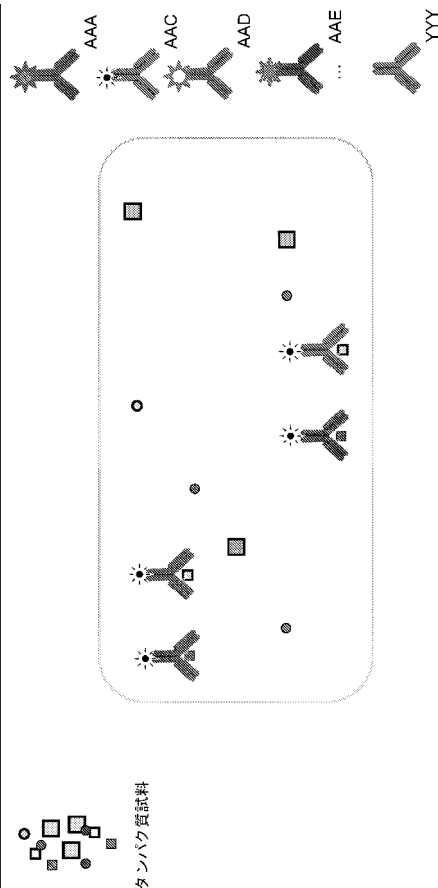
【図 6】



【図 7】



【図 8】





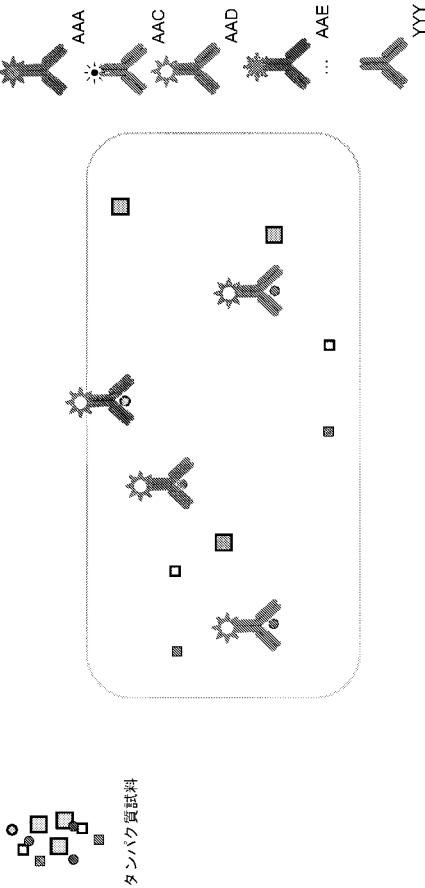
総プロテオーム定量化：全タンパク質の計数

【図 1 1】

タンパク質	分子数	d-コードのヒット数
タンパク質 1	106	10
タンパク質 2	832	6
タンパク質 3	$1.6 \times 10^6$	30
...		
タンパク質 25K	1	22

抗ペプチド抗体デコーディングによるタンパク質同定／定量化

【図 9】



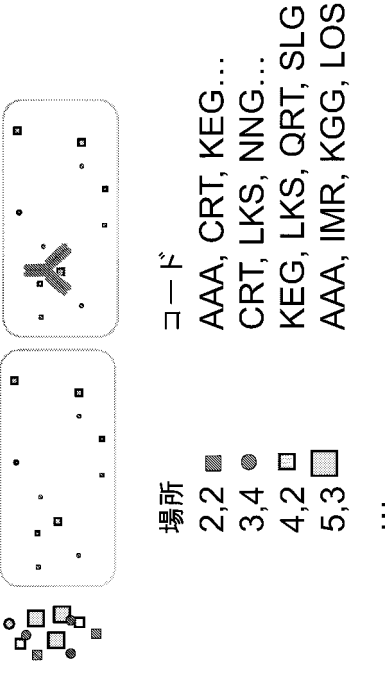
例外リスト：変異体および修飾されたタンパク質である可能性のあるもの

【図 1 2】

位置	可能性のあるタンパク質	余分なコードのヒット
スポット 1, 2	タンパク質 1	TRT, Y', ...
スポット 12, 36	タンパク質 2	RBM, ...
...		

抗体測定データのコンピューターによるデコーディング

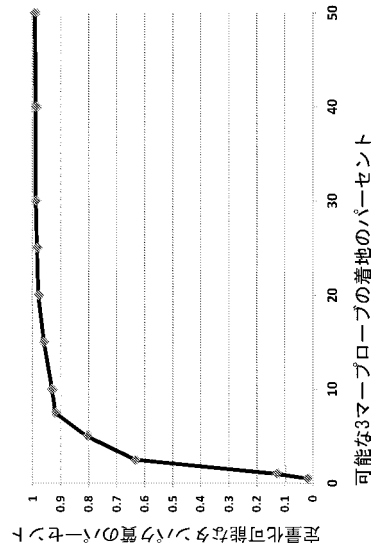
【図 1 0】



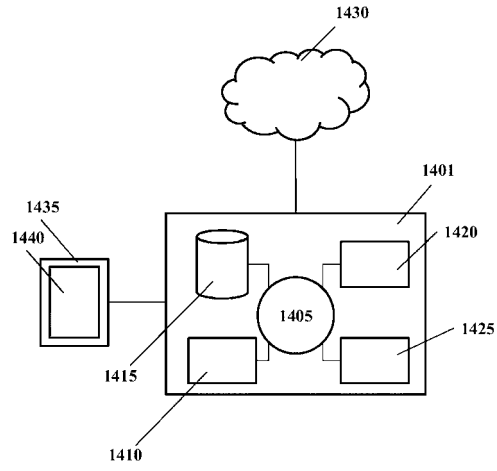
場所	コード
2,2	AAA, CRT, KEG...
3,4	CRT, LKS, NNG...
4,2	KEG, LKS, QRT, SLG
5,3	AAA, IMR, KGG, LOS
...	

【図 13】

定量化に必要な3マーカーコード抗体サブリングのカパレツジ

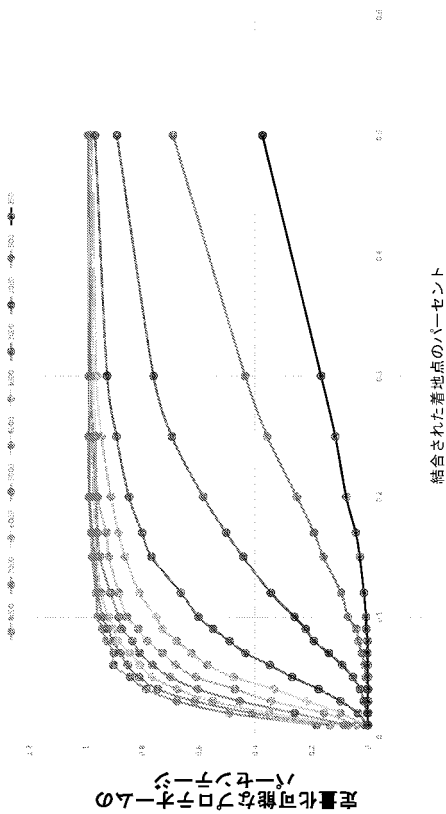


【図 14】



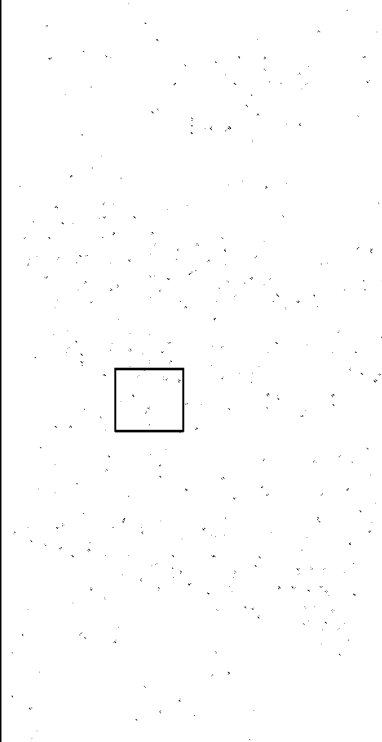
【図 15】

可定性 対 カパレツジにおける、プローブ数の影響力

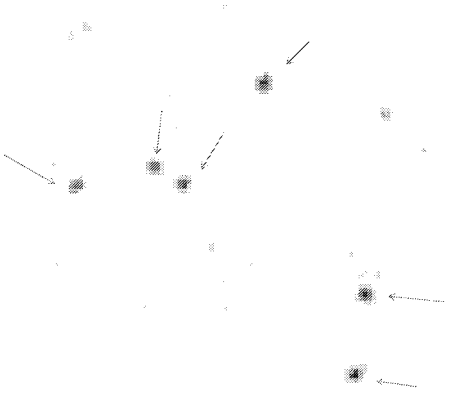


【図 16 A】

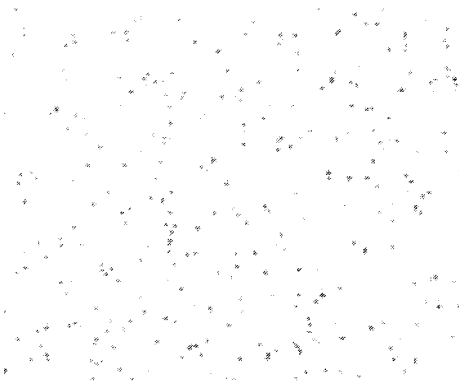
タンパク質単分子の蛍光画像化



【図 16 B】



【図 17】



【図 18】

# タンパク質の同定

タンパク質の同定は、ゲノムデータから得られた配列をデータベースと比較することで行われる。ここでは、ヒトゲノムデータベース（NCBI）と、タンパク質データベース（UniProt）を用いて、ゲノムデータから得られた配列を同定する。同定されたタンパク質の機能や構造に関する情報は、タンパク質データベースから取得される。同定されたタンパク質の機能や構造に関する情報は、タンパク質データベースから取得される。

【手続補正書】

【提出日】令和1年8月1日(2019.8.1)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】配列表

【補正方法】追加

【補正の内容】

【配列表】

2020514746000001.app

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No.  PCT/US 17/64322
<b>Box No. I</b>	<b>Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)</b>	
<p>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:</p> <p>a. <input type="checkbox"/> forming part of the international application as filed:</p> <div style="margin-left: 20px;"><input type="checkbox"/> in the form of an Annex C/ST.25 text file. <input type="checkbox"/> on paper or in the form of an image file.</div> <p>b. <input type="checkbox"/> furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.</p> <p>c. <input checked="" type="checkbox"/> furnished subsequent to the international filing date for the purposes of international search only:</p> <div style="margin-left: 20px;"><input checked="" type="checkbox"/> in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)). <input type="checkbox"/> on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).</div> <p>2. <input checked="" type="checkbox"/> 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.</p> <p>3. Additional comments:</p>		

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/64322

**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.: 8-21  
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:

----- please see extra 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 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.:  
1-7, 22-33, 37-38, 42

**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.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 17/64322

A. CLASSIFICATION OF SUBJECT MATTER  
IPC(8) - C40B 30/04, G01N 33/48, G01N 33/53, G01N 33/68 (2018.01)  
CPC - C40B 30/04, G01N 33/6803, G01N 33/6845, G06F 19/18, G01N 33/50

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)

See Search History Document

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

See Search History Document

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

See Search History Document

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/0054408 A1 (RAVI et al.) 20 Mach 2003 (20.03.2003) para [0007], [0009], [0030], [0031], [0032], [0034], [0038], [0041], [0042], [0043], [0047], [0053], [0054], [0055], [0057], [0069], [0078], [0080], [0082], [0084], [0110], [0115], [0130], [0131], [0132]	1-7, 22-33, 37-38, 42

☐ Further documents are listed in the continuation of Box C. ☐ 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

"G" document member of the same patent family

Date of the actual completion of the international search

4 April 2018

Date of mailing of the international search report

25 APR 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents  
P.O. Box 1450, Alexandria, Virginia 22313-1450  
Facsimile No. 571-273-8300

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300  
PCT OSF: 571-272-1774

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/64322

Continuation of Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I: Claims 1-7, 22-33, 37-38, 42, drawn to methods of determining the identity of a protein and quantifying said protein.

Group II: Claims 34-36, drawn to methods to select a panel of affinity reagents.

Group III: Claims 39-41, drawn to a method that is capable of identifying more than 20% of proteins in the human proteome from a protein sample, wherein the proteins are not substantially destroyed in the process.

The inventions listed as Groups I, II, III do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

## Special Technical Features

Group I requires method steps for methods of determining the identity of a protein and quantifying said protein, wherein some embodiments require use of a substrate having unique optically resolvable spatial address, not required by Groups II and III.

Group II requires method steps for selecting a panel of affinity reagents, not required by Groups I and III.

Group III requires identifying more than 20% of proteins in the human proteome from a protein sample, and also requires that the proteins are not substantially destroyed in the process, not required by Groups I and II.

## Common Technical Features

The feature shared by Groups I and II is a panel of m affinity reagents capable of identifying an unknown protein selected from a pool of n possible proteins, wherein n is larger than m.

The feature shared by Groups I and III is identifying proteins in a protein sample.

However, these shared technical features do not represent a contribution over prior art, because the shared technical features are taught by US 2003/0054408 A1 to Ravi et al. (hereinafter 'Ravi').

Ravi discloses identifying proteins in a protein sample (para [0007] - "The disclosed methods and system include a method for identifying a protein, the method including cleaving the protein with a proteolytic agent to produce peptide fragments, providing an array comprising a solution set of binding reagents, contacting the peptide fragments with the array to promote specific interactions between the fragments and the array, detecting the binding pattern of the peptide fragments on the array; and comparing the binding pattern of the peptide fragments to a reference set").

Ravi discloses a panel of m affinity reagents capable of identifying an unknown protein selected from a pool of n possible proteins, wherein n is larger than m (para [0032] - "As used herein, the term 'solution set' refers to an optimum set of binding reagents . . . In certain embodiments, proteins in a given protein mixture may be identified based solely on the binding pattern to the solution set. In other embodiments, proteins may be identified based on the binding pattern"; para [0047] - "a goal in determining a solution set of binding reagents is not that the binding reagents each recognize distinct epitopes within the protein mixture (e.g., a one-to-one correlation of binding reagents to proteins in a given mixture), but that many or all of the binding reagents recognize epitopes common to peptide fragments produced from cleavage of a plurality of proteins within a given mixture. A concurrent goal is that the minimum set produces a unique binding pattern for the peptide fragments of each protein within the mixture. Thus, in one embodiment, a solution set of binding reagents provides for identification of a maximum number of proteins within a given protein mixture using the fewest number of binding reagents possible"; para [0084] - "choosing the minimum number of binding reagents or antibodies ('selected set') based on varying constraints, such that the binding reagent binding pattern for a given protein is unique for each protein to be identified and/or distinguished").

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Groups I, II, III therefore lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

## フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

C 1 2 Q 1/68

(81)指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, 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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT

(特許庁注: 以下のものは登録商標)

1 . T W E E N

2 . A N D R O I D

(74)代理人 100142929

弁理士 井上 隆一

(74)代理人 100148699

弁理士 佐藤 利光

(74)代理人 100128048

弁理士 新見 浩一

(74)代理人 100129506

弁理士 小林 智彦

(74)代理人 100205707

弁理士 小寺 秀紀

(74)代理人 100114340

弁理士 大関 雅人

(74)代理人 100114889

弁理士 五十嵐 義弘

(74)代理人 100121072

弁理士 川本 和弥

(72)発明者 マリック パラグ

アメリカ合衆国 9 4 0 2 5 カリフォルニア州 メンロ パーク オブライエン ドライブ 1  
1 6 5 スイート ビー

Fターム(参考) 2G045 AA13 CA25 CA26 CB01 CB21 DA13 DA14 DA36 FB12

4B063 QA18 QQ42 QQ79 QR32 QR48

【要約の続き】



# 抗ペプチド抗体デコーディングによるタンパク質同定／定量化

