

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-189241  
(P2005-189241A)

(43) 公開日 平成17年7月14日(2005.7.14)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
GO 1 N 33/48	GO 1 N 33/48	2 GO 4 5
GO 1 N 27/447	GO 1 N 27/62	
GO 1 N 27/62	GO 1 N 27/62	
GO 1 N 30/72	GO 1 N 30/72	
GO 1 N 33/483	GO 1 N 33/483	

審査請求 有 請求項の数 23 O L 外国語出願 (全 41 頁) 最終頁に続く

(21) 出願番号	特願2004-367043 (P2004-367043)	(71) 出願人	591003013
(22) 出願日	平成16年12月20日 (2004.12.20)		エフ. ホフマン-ラ ロシュ アーゲー
(31) 優先権主張番号	03104775.6		F. HOFFMANN-LA ROCH
(32) 優先日	平成15年12月18日 (2003.12.18)		E AKTIENGESELLSCHAF
(33) 優先権主張国	欧州特許庁 (EP)		T
			スイス・シーエイチ-4070バーゼル・
			グレンツアーヘルストラツセ124
		(74) 代理人	100078662
			弁理士 津国 肇
		(74) 代理人	100075225
			弁理士 篠田 文雄
		(72) 発明者	ペーター・ベルント
			スイス国、ツェーハー-4056 バーゼ
			ル、ミットレル・シュトラーセ 138

最終頁に続く

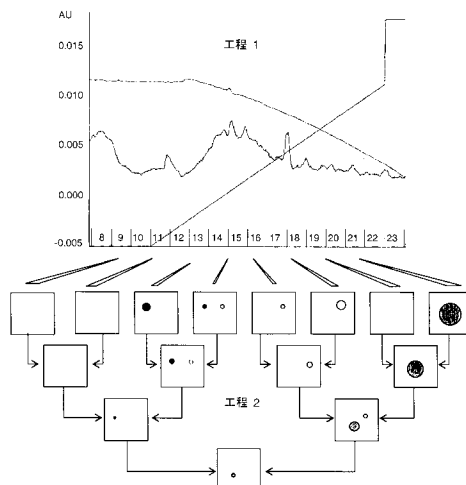
(54) 【発明の名称】 連続組合せ希釈による定量法

(57) 【要約】 (修正有)

【課題】 生体分子の複合混合物中のある生体分子の定量方法を提供する。

【解決手段】 少なくとも1つの識別できる成分を含む、少なくとも2つの画分を与える、生体分子の混合物の画分を含むの画分は、次に組合せ連続希釈に付される。続いて生体分子は、感度閾値及び同一性情報を与える方法により、画分中で検出及び同定される。生体分子の量は、それぞれの希釈係数を考慮して各希釈レベルで各画分中の生体分子の同定の数を集約することにより測定する。正規化を目的として、この合計は、全ての希釈レベルでの全ての画分中の全ての生体分子の同定の総数で割ることができる。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

生体分子の複合混合物中のある生体分子の定量方法であって、

- a . それぞれ少なくとも 1 つの識別できる成分を含む、生体分子の混合物の分画の少なくとも 2 つの画分を用意すること、
  - b . 画分を組合せ連続希釈に付すこと、
  - c . 各原画分及び各希釈画分中の生体分子を、感度閾値及び同一性情報を提供する方法により、検出及び同定すること、並びに
  - d . それぞれの希釈係数を考慮して各希釈レベルで各画分中の生体分子の同定の数を集約することにより、複合混合物中の生体分子を定量すること
- を含む方法。

10

**【請求項 2】**

d ) の合計を、全ての希釈レベルでの全ての画分中の全ての生体分子の同定の総数で割る、請求項 1 記載の方法。

**【請求項 3】**

a ) の画分の数、生体分子の混合物の複雑さに依存する、請求項 1 又は 2 記載の方法。

**【請求項 4】**

a ) の画分の数、生体分子の複合混合物中の異なる生体分子の濃度に依存する、請求項 1 ~ 2 のいずれか 1 項記載の方法。

20

**【請求項 5】**

a ) の画分の数、c ) の検出及び同定方法に依存する、請求項 1 ~ 2 のいずれか 1 項記載の方法。

**【請求項 6】**

b ) の希釈レベルの数、画分中の生体分子の出発濃度に依存する、請求項 1 ~ 5 のいずれか 1 項記載の方法。

**【請求項 7】**

b ) の希釈レベルの数、分画後の画分の数に依存する、請求項 1 ~ 5 のいずれか 1 項記載の方法。

**【請求項 8】**

b ) の希釈レベルの数、検出及び同定方法の検出限界に依存する、請求項 1 ~ 5 のいずれか 1 項記載の方法。

30

**【請求項 9】**

生体分子が、( n - 1 ) 個以下の画分に存在する ( ここで、n は、画分の総数であり、かつ n は、2 以上である )、請求項 1 ~ 8 のいずれか 1 項記載の方法。

**【請求項 10】**

生体分子が、2 つの画分に存在する、請求項 1 ~ 8 のいずれか 1 項記載の方法。

**【請求項 11】**

生体分子が、1 つの画分に存在する、請求項 1 ~ 8 のいずれか 1 項記載の方法。

**【請求項 12】**

c ) の検出及び同定方法が、二次元ゲル電気泳動を含む、請求項 1 ~ 11 のいずれか 1 項記載の方法。

40

**【請求項 13】**

c ) の検出及び同定方法が、更に質量分析を含む、請求項 12 記載の方法。

**【請求項 14】**

c ) の検出及び同定方法が、免疫検定法を含む、請求項 1 ~ 11 のいずれか 1 項記載の方法。

**【請求項 15】**

c ) の検出及び同定方法が、質量分析と組合せたガスクロマトグラフィーを含む、請求項 1 ~ 11 のいずれか 1 項記載の方法。

50

## 【請求項 16】

c) の検出及び同定方法が、特異的に標識した分子の実体による電気泳動を含む、請求項 1 ~ 11 のいずれか 1 項記載の方法。

## 【請求項 17】

a) の画分が、クロマトグラフィー分画法、超遠心分離法、タンパク質沈降法、又は免疫沈降法により分離される、請求項 1 ~ 16 のいずれか 1 項記載の方法。

## 【請求項 18】

生体分子が、ポリペプチドである、請求項 1 ~ 17 のいずれか 1 項記載の方法。

## 【請求項 19】

生体分子が、ポリヌクレオチドである、請求項 1 ~ 17 のいずれか 1 項記載の方法。 10

## 【請求項 20】

生体分子が、炭水化物である、請求項 1 ~ 17 のいずれか 1 項記載の方法。

## 【請求項 21】

生体分子が、脂質である、請求項 1 ~ 17 のいずれか 1 項記載の方法。

## 【請求項 22】

生体分子が、糖タンパク質である、請求項 1 ~ 17 のいずれか 1 項記載の方法。

## 【請求項 23】

生体分子が、リポタンパク質である、請求項 1 ~ 17 のいずれか 1 項記載の方法。

## 【発明の詳細な説明】

## 【技術分野】

20

## 【0001】

本発明は、生体分子の複合混合物中のある生体分子の定量方法であって、画分への複合混合物の分画と、これに続く画分の組合せ連続希釈、並びに明確な感度閾値及び同定能力を持つ方法による各原画分及び各希釈画分中の生体分子の検出を含む方法に関する。

## 【0002】

生体分子（例えば、タンパク質）の現行の検出方法は、二次元ゲル電気泳動と、これに続く染色ゲルの容量分析である。しかし、分析される生体分子の量を測定することは、特に異なる試料中のその量を比較しようとするると困難である。生体分子濃度における試料間変動を明らかにするには、ゲルを並行して処理し、かつゲル間のマッチングを行わなければならない。 30

## 【0003】

生体分子の複合混合物中の生体分子の定量のための簡易な方法が求められている。

## 【0004】

本発明は、生体分子の複合混合物中のある生体分子の定量方法であって、

a. それぞれ少なくとも 1 つの識別できる生体分子成分を含む、生体分子の複合混合物の分画から生じる少なくとも 2 つの画分を用意すること、

b. 画分を組合せ連続希釈工程に付すこと、

c. 各原画分及び各希釈画分中の生体分子を、安定かつ十分に明確な感度閾値及び同一性情報を持つ方法により、検出及び同定すること、並びに

d. それぞれの希釈係数を考慮して各希釈レベルで各画分中の生体分子の同定の数を集約することにより、生体分子の複合混合物中の生体分子を定量することを含む方法を提供する。 40

## 【0005】

正規化を目的として、d) の合計は、全ての希釈レベル（原画分及び希釈画分の希釈レベル）での全ての画分中の全ての生体分子の同定の総数で割ることができる。

## 【0006】

生体分子の定量のための本発明の方法は、1 つのソースからの生体分子の複合混合物中の 1 つ以上の生体分子の、他のソースからの複合混合物中のそれぞれの生体分子に比較した相対的定量法を提供する。

## 【0007】

50

生体分子の複合混合物は、細胞、細胞培養上清、生物学的流体（血清、血漿、尿、気管支洗浄液、喀痰、生検材料（脳脊髄液など））を含む生物学的ソースを含む、任意のソースに由来してよい。生体分子の複合混合物は、少なくとも2つの異なる生体分子を含む。本発明における生体分子は、ポリヌクレオチド、ポリペプチド、炭水化物、脂質、糖タンパク質、リボタンパク質あるいは他の修飾型又はこれらの代謝物を含む、任意の生体分子であってよい。検出及び同定方法は、単一型の生体分子に限定されていてもよいし、あるいは一度に幾つかの分類の生体分子を検出及び分析することもできる。

**【0008】**

本発明の方法において使用される分画方法は、生体分子の複合混合物を識別できる画分に効率的に分離するものである。好ましくは、生体分子の複合混合物は、各異なる生体分子が $(n - 1)$ 個以下だけの画分に存在する（ここで、 $n$ は、画分の総数であり、かつ $n$ は、2以上である）、識別できる画分に分画される。好ましくは、異なる生体分子は、2つの異なる画分に、更に好ましくは1つの画分に存在する。本発明の方法において使用することができる分画方法は、標的とする型の生体分子の複合混合物の分離に適した任意の方法から選択することができる。本発明の方法において使用することができる分画方法は、吸着、重力若しくは沈降速度に基づく分画、電気泳動分画又はこれらの方法論の組合せを含む群から選択することができる。例えば、標的分子としてタンパク質の場合には、この群は、クロマトグラフィー分画法、超遠心分離法、タンパク質沈降法、親和性精製法、又は免疫沈降法を非限定的に含む。標的分子としてペプチド（例えば、タンパク質消化物から得られる）の場合には、この群は、高圧液体クロマトグラフィー（HPLC）を非限定的に含む。

10

20

**【0009】**

画分は次に、組合せ連続希釈に付される。組合せ連続希釈には、そこから始めるための少なくとも2つの画分が必要である。好ましくは、生体分子の複合混合物は、続いての検出工程において十分な数の異なる標的分子の検出及び同定が可能となるのに必要な数の画分に分画される。好ましくは、原画分の数には素数ではなく、更に好ましくは、原画分の数には偶数であり、そして好ましくは、各最初の画分は、少なくとも1つの識別できる成分を含む。

**【0010】**

組合せ連続希釈を始めるための画分の数には、生体分子の混合物の複雑さ、生体分子の複合混合物中の個々の生体分子の濃度、分離方法の効率、並びに分画及び組合せ連続希釈後に使用される検出及び同定方法に依存する。

30

**【0011】**

組合せ連続希釈には、少なくとも1つの異なる生体分子を含む、少なくとも2つの異なる画分が組合せられる。好ましくは、2つの画分が組合せられる。これは、各画分中の生体分子の濃度と各画分の容量とのドット積を全てを組合せた画分の総容量で割った商によって、プールした画分中の生体分子の濃度を変化させる。一般に、これは、希釈画分中のどの生体分子成分でも、原画分中のそれぞれの生体分子の最大濃度に比較して、その濃度を低下させる。以下の希釈工程において、個々のタンパク質の濃度は、次の検出及び同定方法の感度閾値を下まわるまで低下する。

40

**【0012】**

希釈工程の数は、原画分中の生体分子の出発濃度、分画後の原画分の数、並びに検出及び同定方法の検出限界に依存する。

**【0013】**

本発明の方法は更に、検出及び同定方法を含む。この検出方法は、明確な感度閾値を特徴とし、また検出生体分子に関する同一性情報を提供する必要がある。それによって、特異的な生体分子の存在又は非存在を、原画分又は希釈画分において決定することができる。本発明の検出及び同定方法は、生体分子の化学組成、構造、又は配列、並びにそれから生じる物理化学的又は酵素的性質に依存することがある。これらは、特異的プローブとのハイブリダイゼーション、特異的抗体又はレクチンとの反応、特異的分子プローブとの酵

50

素又は化学反応、等電点、分子量、生体分子の酵素消化により生じる断片の分子質量、NMRスペクトルあるいはこれらの組合せを含む。タンパク質定量の例では、本発明の検出及び同定方法は、一次元若しくは二次元ゲル電気泳動と質量分析との組合せ、免疫検定法（例えば、ウェスタンブロット）、質量分析と組合せたガスクロマトグラフィー（GS/MS）又は特異標識した分子の実体（例えば、蛍光、化合物（例えば、ビオチン）、又は放射活性タグ）による電気泳動よりなる群から選択することができる。

【0014】

生体分子の定量法を導くために、各希釈工程のそれぞれの希釈係数が考慮される、各希釈工程の画分の同定又は特定の指紋（ペプチド質量指紋）が計算される。得られる生体分子の同定の数は、全ての希釈レベルについて集約される。正規化を目的として、この合計は、全ての画分（原画分及び希釈画分）中の全ての生体分子の同定の総数で割ることができる。

10

【0015】

【数1】

$$\text{相対量 (q)} = \frac{\sum (d_i \times N_i)}{N_{\text{total}}}$$

【0016】

[ 式中、 $N_i$  は、希釈レベル  $i$  での個々の生体分子の同定の数  $N$  であり、 $d_i$  は、それぞれの希釈レベル  $i$  の希釈係数  $d$  であり、そして  $N_{\text{total}}$  は、全ての希釈レベルの全ての画分中の全ての生体分子の同定の総数  $N$  である ]。

20

【0017】

本方法は、種々の生体分子の性質とは独立である。ポリヌクレオチドは、ポリペプチド又は炭水化物と同様に処理することができる。本方法の更に別の利点は、生体分子定量を目的とする更なる努力を必要とすることなく、定量法を簡単に生体分子の同定と組合せることである。更に、1つのソース由来の生体分子の量を、別のソースのものと比較しようとするならば、生体分子の混合物を相互に別々に処理することができる。

【0018】

これまで本発明を一般的に記述してきたが、同発明は、以下の図面と一緒に、説明だけを目的として本明細書に含まれるものであり、かつ特に断りない限り限定するという意図のない、特定の実施例を参照することにより更によく理解されよう。

30

【0019】

図1は、本発明の方法を図解する：第1の工程において、複合混合物を異なる画分に分画する。これらの画分は、次に組合せ連続希釈に付す。第2の工程において、生体分子は、例えば、試料プールでの二次元ゲル電気泳動と、これに続く質量分析での同定により検出する。（AU：吸収単位；8～23：画分）

【0020】

図2は、生体分子の相対量の計算を示す。 $q$ ：生体分子の相対量； $N_i$ ：希釈レベル  $i$  の個々の生体分子の同定の数  $N$ ； $d_i$ ：それぞれの希釈レベル  $i$  の希釈係数  $d$ ； $N_{\text{total}}$ ：全ての希釈レベルの全ての画分中の全ての生体分子の同定の総数  $N$ 。（スキーム：N1：非希釈、N2：2倍希釈、N3：4倍希釈、N4：8倍希釈）

40

【0021】

図3は、レベル1（希釈なし）、レベル2（2倍希釈）、レベル3（4倍希釈）、及びレベル4（8倍希釈）からの二次元電気泳動ゲルにおけるタンパク質、グリコーゲンホスホリラーゼ（a）、ピメンチン（b）及び熱ショックタンパク105（c）に関する同定の数を示す。値は、三重測定で行われた実験から合計した。（対照：5mMグルコース；高グルコース：10mM）

【0022】

50

## 実施例

実施例において参照される市販の試薬は、特に断りない限り製造業者の指示書にしたがい使用した。

### 【0023】

#### 実施例1：

##### 細胞培養

I N S - 1 細胞 (Asfari, Janjicら 1992) は、10% FCS (インビトロジェン (Invitrogen)、熱不活化)、10 mM ヘペス溶液 (インビトロジェン)、1 mM ピルビン酸 Na (シグマ (Sigma)) ; 50  $\mu$ M ベータ - メルカプトエタノール (シグマ)、1% ペニシリン / ストレプトマイシン溶液 (シグマ)、及び低濃度 (5 mM) 又は高濃度 (10 mM) のグルコース (シグマ) を補足した RPMI 1640 培地 (インビトロジェン) で培養した。細胞は大体、低グルコース濃度で培養した。予備培養のために、細胞を分割し、次に低グルコース培地で細胞がコンフルエントになるまでインキュベートした。次いで培地を低グルコース又は高グルコース培地に替えて、インキュベーションを4日間続けた。回収のために、細胞を最初にハンス平衡塩溶液 (HBSS、インビトロジェン) で1回洗浄し、次に細胞が丸みを帯び、瓶表面から離れるまでトリプシン / EDTA 溶液で1~2分間覆った。トリプシン / EDTA 溶液を廃棄して、細胞をトリプシンインヒビター溶液 (シグマ) に懸濁し、遠心分離管に移して、1200  $\times$  g で5分間遠心分離した。この後、再度同じ遠心分離機パラメーターを用いて、細胞をHBSS中で3回洗浄した。上清を吸引して、廃棄し、そしてサイトゾルの調製に使用するまでペレットを - 80 で凍結貯蔵した。

### 【0024】

#### サイトゾルの調製

全ての溶液 (HBSS を除く) を4 に冷却し、全ての工程は、冷却環境 (氷浴) 中で行った。約  $10^8$  個の細胞を細胞均質化培地 (CHM ; 150 mM MgCl<sub>2</sub>、10 mM KCl、10 mM トリス、0.25 M グルコース、1 mM EDTA、pH 7.4) に再懸濁して、氷上に2分間放置した。この細胞懸濁液をポッター・エルヴェージェム (Potter-Elvehjem) ホモジェナイザーに移した。ポッター・エルヴェージェム ホモジェナイザーの冷ペッスルをオーバーヘッド高トルク電動機に取り付け、細胞を1000 rpm で10ストロークを用いてホモジェナイズした。均質化の効率 (> 90% の破壊細胞) は、位相差顕微鏡により確認した。細胞破砕物及び核は、1000  $\times$  g で5分間遠心分離することにより除去した。ミトコンドリアは、5000  $\times$  g での遠心分離により分離した。濃縮サイトゾル画分は、最後に20000  $\times$  g で遠心分離し、上清を清浄な管に移すことにより回収した。調製物中の最終タンパク質濃度は、2.5 ~ 5.0 mg/ml であった。

### 【0025】

#### クロマトグラフィー分画

全ての分画工程は、AKTA エクスプローラー 10 (AKTAexplorer 10) クロマトグラフィーシステム (アマシャム (Amersham)) を用いて室温で行った。サイトゾル調製物 (10 mg の総タンパク質) を0.45  $\mu$ m ミレックス - HV (Millex-HV) シリンジ駆動フィルターユニットに通して、脱塩カラム (直列に連結された3個の5 ml ハイトラップ (HiTrap) 脱塩カラム、アマシャム) に載せた。タンパク質は、緩衝液 A (25 mM NaHPO<sub>4</sub>、pH 7.5 ; 1 mM EDTA ; 0.5 mM ジチオエリトリトール ; 1  $\times$  コンプリート EDTA - フリー (Complete EDTA-free) (ロシュ・ダイアグノスティクス (Roche Diagnostics) 製のプロテアーゼインヒビターカクテル錠 ; pH は7.5 に調整) を用いて、1.5 ml/分の流量を用いて溶出した。タンパク質は、UV 吸収 (280 nm) の上昇と導電率の最低値をタンパク質画分の境界として利用して、20 ml 注入ループに回収した。次にタンパク質は、TSK DEAE - 5PW 7.5 cm  $\times$  7.5 mm カラム (トーソー・バイオセップ (TOSOH BIOSEP)) を1 ml/分の流量で用いるアニオン交換クロマトグラフィーにより分離した。緩衝液 A は、結合緩衝液として使用し、緩衝液 B (25 mM NaHPO<sub>4</sub>、pH 7.5 ; 1 mM EDTA ; 0.5 mM ジチオエリトリトール ; 1  $\times$  コンプリート E

D T A - フリー (ロシュ・ダイアグノスティクス製のプロテアーゼインヒビターカクテル錠; 1 M NaCl、pHは7.5に調整)は、溶離緩衝液として使用した。試料をカラムに載せて、非結合物質を7カラム容量(CV)の緩衝液Aで洗い流した。次に結合タンパク質を3セグメントの勾配(第1セグメント: 3CV中に0~11%緩衝液B; 第2セグメント: 10CV中に11~30%緩衝液B; 1.5CV中に30~50%緩衝液B)により溶出した。最後に、カラムを5CVの50%緩衝液Bで洗浄した。1mlの画分を回収し、合わせて8個のプールと流出液を生成した。導電率境界は以下とした: FT: 導電率の上昇までのUV280上昇; 1(導電率上昇開始~12mS); 2(12~15mS); 3(15~18mS); 4(18~21mS); 5(21~24mS); 6(24~27mS); 7(27~30mS); 8(30~40mS)。

10

## 【0026】

## 二次元電気泳動

画分を濃縮して、POROS 20 R1材料(パーセプティブ・バイオシステムズ(PerSeptive Biosystems))100mgで充填した自動充填シリンジ駆動ミニカラム(MoBio Tec M1002)を用いる逆相クロマトグラフィーにより脱塩した。カラムを0.1%トリフルオロ酢酸(TFA)10ml及び70%アセトニトリル/0.1%TFAで洗浄した。試料を装填後、カラムを0.1%TFA 10mlで洗浄して、70%アセトニトリル/0.1%TFA 2mlで溶出した。次に溶出液をスピードヴァク(SpeedVac)エバポレーター中で乾燥して、IEF試料緩衝液(7M尿素、2Mチオ尿素、50mMトリス、pH7.5、2%(w/v)CHAPS、0.4%(w/v)ジチオエリトリール、0.5%(w/v)両性電解質)にとった。タンパク質0.5mgを含むアリコート各画分からとっておき、試料1~8として標識した。画分の残りから以下の試料を調製した: 試料9: 0.25mg画分1+0.25mg画分2; 試料10: 0.25mg画分3+0.25mg画分4; 試料11: 0.25mg画分5+0.25mg画分6; 試料12: 0.25mg画分7+0.25mg画分8; 試料13: 0.125mg画分1+0.125mg画分2+0.125mg画分3+0.125mg画分4; 試料14: 0.125mg画分5+0.125mg画分6+0.125mg画分7+0.125mg画分8; 試料15: 0.0625mg画分1+0.0625mg画分2+0.0625mg画分3+0.0625mg画分4+0.0625mg画分5+0.0625mg画分6+0.0625mg画分7+0.0625mg画分8。よって試料1~8は、タンパク質画分0.5mgを含み、試料9~12は、それぞれこれらの試料の2倍希釈に相当し、試料13及び14は、4倍希釈に、そして試料15は、これら原画分の8倍希釈に相当する。等電点電気泳動は、20のプロテアンIEFセル(Protean IEF Cell)(バイオラッド(BioRad))中で3~10のpH範囲の固定化pH勾配(IPG)ストリップ(IPG 3-10L; アマシャム)を用いて実行した。乾燥ストリップは、7M尿素、2Mチオ尿素、2%(w/v)CHAPS、0.4%(w/w)ジチオエリトリール、及び0.5%(w/v)両性電解質を含む溶液に再水和した。このタンパク質画分をストリップの陰極末端にカップ装填した。電圧は、8時間にわたり5000Vまで直線的に上昇し、続いて10時間5000Vで安定した。このストリップを、平衡化溶液I(6M尿素、50mMトリスpH7.5、30%グリセロール、2.0%SDS、30mMジチオエリトリール)及び平衡化溶液II(6M尿素、50mMトリスpH8.8、30%グリセロール、2.0%SDS、0.23Mヨードアセトアミド)中でそれぞれ10分間連続洗浄することにより、平衡化及びアルキル化した。このストリップを11%アクリルアミド/PDA(37:1)勾配ゲル(240mm×200mm×1.5mm)に載せた。タンパク質は、常に冷却(20)しながらのETTAN Dalt電気泳動装置(アマシャム)で80V O/Nでの電気泳動により分離した。

20

30

40

## 【0027】

## ゲル染色及び加工

ゲルは、50%メタノール/10%酢酸中で固定して、クーマシーブルー(コロイダル・ブルー(Colloidal Blue)、インビトロジェン、カールズバッド(Carlsbad)、カリフォルニア州)で一晩染色し、続いて超純水で合計7時間何回も洗浄した。ゲルを走査して

50

、1.2 mmの直径を持つスポットを自動スポットピッキング装置を用いて切り出した。スポットは、100 mM炭酸水素アンモニウム及び30%アセトニトリルを含む溶液中で脱染色した。乾燥した脱染色ゲル断片を10 µg/mlトリプシン溶液（ロシュ・ダイアグノスティクス）5 µl中で一晩室温で消化した。超純水10 µlの添加後、タンパク質は、75%アセトニトリル及び0.3% (v/v) TFAを含む溶液5 µlで抽出した。このペプチド溶液を、MALDI標的にマトリックスとしてのシアン-4-ヒドロキシケイ皮酸と一緒にスポットした。

#### 【0028】

質量分析及びタンパク質同定

ペプチド質量は、ブルーカー・ウルトラフレックス測定器（Bruker Ultraflex Instrument）（ブルーカー（Bruker）、ブレーメン、ドイツ）で、ACTH及びブラジキニンを内部質量標準として使用して測定した。以下に説明されるように、モノアイソトピックペプチド質量は、質量スペクトルから自動検出して、タンパク質配列のデータベース（例えば、スイスプロット（SwissProt）、又はNCBIラットゲノムドラフト）からの全てのタンパク質のコンピュータ内（in-silico）トリプシン消化から誘導されるペプチドの理論質量と比較した。

#### 【0029】

MALDI質量スペクトルに関するピーク注釈

質量分析データは、測定器のベースラインを決定するために、低域中央パラメトリックスプラインフィルターで2回フィルターにかけた。ベースラインからの平滑化残差平均標準偏差を、データにおける測定器ノイズレベルの推定値として使用した。

#### 【0030】

ベースライン補正、及びノイズを超えるレベル座標におけるデータの修正後、ベースラインから最大偏差のデータ点を使用して、予想ペプチドピークを検出するための非線形（Levenberg-Marquardt）データフィッティング手順の種を蒔いた。具体的には、このフィット手順は、ピーク高、分解能、及びモノアイソトピック質量によりパラメーター化された最良のフィッティングの平均理論ペプチド同位体分布を作り出すことを試みる。有意なフィットへの収斂は、シグマ値を追跡することにより常法で求められる。

#### 【0031】

収斂の成功後、データ点の1/3をランダム交換しての16回反復を用いるブートストラップ手順を利用して、求めたパラメーターの誤差を推定する。

#### 【0032】

得られるフィットは、データから引き算して、フィット付近のノイズレベルを、外挿ノイズレベル及びピークフィットからの偏差の合計に調整し、そしてノイズのレベルを5倍を超えて上まわる候補ピークを見い出すことができるまで、次のピークを見い出すため、このプロセスを繰り返す。このプロセスは、50を超えるデータピークが見い出されたら終える。

#### 【0033】

ゼロ次及び一次の質量への飛行時間の変換は、検出内部標準ピークからの直線外挿を用いて補正し、そしてモノアイソトピック質量値の信頼区間は、ピークと標準の質量の正確さから推定する。

#### 【0034】

コンピュータ内タンパク質消化物へのスペクトルピークの確率的マッチング

質量スペクトルのピーク質量リストは、タンパク質全体の配列データベースについての理論消化物と直接比較する。各理論消化物について、 $[1 - (1 - NP(p_i))]^{c \text{ Matches}}$ （ここで、Nは、消化物中のペプチドの数であり、P(p\_i)は、配列データベース中の全てのペプチドのカウントにより割ったピークのモノアイソトピック質量に関する信頼区間と一致するペプチドの数であり、そしてc Matchesは、消化物と質量スペクトルの間の一致の数である）を計算する。この値は、消化物とスペクトルの間の擬陽性一致が得られる確率に比例することを証明することができる。確率値は、スペクトルピー

クから高い有意性で一致が得られるように更にフィルターにかける。同定の第1ラウンド後、同一条件下で獲得される質量スペクトルについての同定の偏差を使用することにより、質量への飛行時間の変換の第2及び第3次条件を補正する。得られる質量値は、ほぼ10 ppm未満の絶対偏差を持つ。次にこれらの質量値は、最終ラウンドのマッチングに使用するが、ここでは、 $0.01 / N \text{ Proteins}$  (ボンフェローニ (Bonferroni) 補正による1%有意レベル) 未満の  $P \text{ mis}$  を持つ全ての一致が許容される。

【0035】

データベース分析

データベース中の各タンパク質について、本試験において分析した2D-PAGEゲル当たりの同定の数をカウントした。本実施例では、希釈レベル1を対照として設定した。次に、以下の値を導いた： 10

- ・ 希釈レベル1における同定の数 (非希釈試料、試料1~8) =  $N_1$
- ・ 希釈レベル2における同定の数 (2倍希釈、試料9~12) =  $N_2$
- ・ 希釈レベル3における同定の数 (4倍希釈、試料13、14) =  $N_3$
- ・ 希釈レベル4における同定の数 (8倍希釈、試料15) =  $N_4$

【0036】

予想されるとおり、多くのタンパク質ではN値は、層から層へおおよそ2倍低下した。希釈係数を明らかにするために、及び各タンパク質についての大ざっぱな絶対量を導くために、定量値  $q$  を以下のとおり計算した：

$q = (N_1 + 2 \times N_2 + 4 \times N_3 + 8 \times N_4) / \text{全ての希釈レベルでの同一ソースの全ての試料についての同定タンパク質スポットの総数}$  20

【0037】

同一ソースの全ての試料についての同定の総数で割ることは、タンパク質濃度における試料間変動を明らかにするために導入した。

【0038】

各タンパク質について、両方の混合試料 (高及び低グルコース) の  $q$  値を計算して比較した。

【0039】

以下の3種のタンパク質を、本定量方法の実行可能性の説明のための実施例として選択した：グリコーゲンホスホリラーゼ (肝臓型) ; ビメンチン、及び熱ショックタンパク105 (表1、図3)。 30

【0040】

【表 1】

	q (5mMグルコース) x 10 <sup>-5</sup>	q (10 mMグルコース) x 10 <sup>-5</sup>
グリコーゲンホスホリラーゼ		
実験 1	112	0
実験 2	8	0
実験 3	124	44
ビメンチン		
実験 1	0	2130
実験 2	80	305
実験 3	0	1758
熱ショックタンパク 105		
実験 1	17	13
実験 2	121	39
実験 3	200	89

10

20

## 【0041】

表 1 : 3 種の実施例タンパク質の 3 回の実験で得られたサイトゾル中に存在するタンパク質の相対量 ( q 値 )

30

## 【0042】

実施例 2 : コラーゲンアルファ I ( IV )

インスリン抵抗性患者 3 名とインスリン感受性患者 3 名 ( 白人、女性 ) からの血清試料を後述のとおり分画した。肥満度指数 ( Body Mass Index ) ( BMI ) と、オイグリセミック - 高インスリン血症クランプ ( Euglycemic-Hyperinsulinemic Clamp ) 方法 ( Garvey ら, Diabetes 34 (1985) 222-234 ) により求めたグルコース廃棄速度 ( Glucose Disposal Rate ) ( GDR ) を表 2 に示す。組合せ連続希釈は、本特許出願に記載されるように実施して、生じた試料を後述のとおり二次元 SDS - ポリアクリルアミドゲル電気泳動 ( 2D - PAGE ) に付した。検出可能なタンパク質スポット全てを各ゲルから切り出した。タンパク質をトリプシンで消化して、生じたペプチドをマトリックス支援レーザーイオン化飛行時間質量分析 ( Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry ) ( MALDI - MS ) に付した。タンパク質同定は、後述のようにペプチド質量指紋分析により実現し、そしてタンパク質リストを実施例 1 に記載されたように比較した。

40

## 【0043】

【表 2】

患者	インスリン感受性 (IS)		患者	インスリン抵抗性 (IR)	
	BMI	GDR		BMI	GDR
IS1	22.4	21.9	IR1	31.3	10.2
IS2	22.4	19.7	IR2	33	11.65
IS3	29.5	20.4	IR3	33.1	8.0

10

## 【0044】

表 2：被験者 6 名の肥満度指数 (BMI) 及びオグリセミック高インスリン血症クランプ方法により求めたグルコース廃棄速度 (GDR)。15 を超える GDR が、インスリン抵抗性の決定の区切点と考えられるため、パネルの左側の患者は、インスリン感受性 (IS) として、そして右側はインスリン抵抗性 (IR) として分類される。これらの個人からの血漿は、組合せ連続希釈と、続く 2 D - PAGE、スポット切り出し、トリプシン消化、MALDI - MS そして最終的にペプチド質量指紋比較法によるタンパク質同定によって分析した。

20

## 【0045】

## 試料調製

プロテオミクス技術を適用することにより、ヒト血漿中のインスリン抵抗性 (Insulin Resistance) マーカーを探索するための方法を確立した。血漿は、約 10 個の高存在量タンパク質を含み、これらが、総タンパク質量の約 98% に相当するため、プロテオミクス技術により分析することが困難である。高存在量タンパク質である、アルブミンと抗体鎖は、クロマトグラフィー法を適用することにより除去して、流出画分をイオン交換体で分画した。説明したスキームは、3 つのクロマトグラフィー工程のマトリックスブルー、プロテイン G 及びイオン交換を含み、そして再現性が高い。全てのクロマトグラフィー工程は、FPLC システム (ファルマシア (Pharmacia)) で実行した。

30

## 【0046】

ミメティックブルー (Mimetic blue) での親和性クロマトグラフィーによるアルブミンの除去及びプロテイン G での親和性クロマトグラフィーによるイムノグロブリンの除去

ヒト血漿を、対照となる個人 3 名及び II 型糖尿病の患者 3 名から採取した。プロテアーゼインヒビターカクテル (ロシュ・ダイアグノスティクス、マンハイム、ドイツ) を血漿に加えた (50 ml に 1 錠)。血漿は、25 mM MES、pH 6.0 で 3 倍希釈することにより塩濃度を低下させ、pH を約 6.0 に調整した。2 つのカラム、ミメティックブルー SA P 6 X L (50 ml、プロメティック・バイオサイエンス社 (ProMetic BioSciences Ltd.)) 及びハイトラップ・プロテイン G HP (HiTrap Protein G HP) (5 ml、アマシャム・バイオサイエンス (Amersham Biosciences)) を連続してつなぎ、25 mM MES、pH 6.0 で平衡化した。血漿タンパク質約 1 g に対応する容量 (15 ml、66 mg/ml) を 0.22 µm フィルターにより濾過して、ミメティックブルーカラムに 5 ml / 分で適用した。このカラムの流出液が直接プロテイン G カラムに載るため、後のカラムからの流出画分を集めた (約 120 mg)。2 つのカラムを 25 mM MES、pH 6.0 100 ml で洗浄し、次にこれらを分離した。ミメティックブルーカラムは、50 mM トリス - HCl、pH 7.5 中の 2 M NaCl の段階的勾配で溶出し、プロテイン G は、100 mM グリシン - HCl、pH 2.8 で溶出し、そして溶出液は 1 M トリス塩基で中和した。流出画分と 2 つの溶出液は、二次元ゲルにより分析して、タンパク質は MALDI - MS により同定した。ミメティックブルーからの溶出液では、主として完全長及び断片化アルブミンを検出した。プロテイン G カラムからの溶出液では、主として重鎖及び軽鎖 Ig

40

50

鎖を検出した。他の血漿タンパク質の多くは、流出画分中に回収した。

【0047】

イオン交換クロマトグラフィーによるタンパク質分画

ミメティックブルー及びプロテインGカラムからの流出及び洗浄画分を合わせて、2 M トリス塩基でpH 8.0に調整し、ハイトラップQ HPカラム (HiTrap Q HP column) (5 ml、アマシャム・バイオサイエンス) に適用して、50 mM トリス - HCl、pH 8.0で5 ml/分で平衡化した。このカラムを、50 mM トリス - HCl、pH 7.5中の0 ~ 1 M NaClの上昇塩濃度の直線勾配で溶出した。5 ml画分を回収して、1-Dゲルにより分析した。総タンパク質約50 mgをこのカラムから回収した。ゲル分析に基づいて、画分をプールすることにより、各プールが総タンパク質約5 mgを含むように、8個のプールを形成した。このプールをウルトラフリー-15遠心分離フィルター (Ultrafree-15 Centrifugal Filter) (5 k MWCO、ミリポア (Millipore)) で濃縮して、8個のプールのそれぞれを2-Dゲルにより分析した。各ゲルから約400スポットを切り出して、MALDI-MSにより分析した。

10

【0048】

2D-PAGE

固定化pH勾配 (IPG) ストリップは、アマシャム・バイオサイエンス (ウプサラ、スウェーデン) から購入した。アクリルアミドは、バイオソルブ (Biosolve) (ファルケンスワード (Valkenswaard)、オランダ) から入手し、そしてポリアクリルアミドゲル調製のための他の試薬は、バイオラッド・ラボラトリーズ (Bio-Rad Laboratories) (ハーキュリーズ (Hercules)、カリフォルニア州、米国) から入手した。CHAPSはロシュ・ダイアグノスティクス (マンハイム、ドイツ) から、尿素はアプリーケム (Applische m) (ダルムシュタット (Darmstadt)、ドイツ) から、チオ尿素はフルカ (Fluka) (ブックス (Buchs)、スイス) から、そしてジチオエリトリトールはメルク (Merck) (ダルムシュタット) から入手した。

20

【0049】

総タンパク質0.5 mgの試料を、試料カップ中その塩基性及び酸性末端で3-10 N L IPGストリップに適用した。等電点電気泳動は200 Vで開始し、電圧は、コンピュータ制御電源を用いて3 V/分で5000 Vまで徐々に上昇させて、更に6時間一定に保持した。第2の次元の分離は、12%一定SDSポリアクリルアミドゲル (180 x 200 x 1.5 mm) 上でゲル当たり40 mAで実行した。5%リン酸を含む40%メタノール中で12時間のタンパク質固定後、ゲルはコロイド状クーマシーブルー (ノーベックス (Novex)、サンジエゴ、カリフォルニア州、米国) で24時間染色した。過剰の染料は、H<sub>2</sub>Oでゲルから洗浄し、そしてゲルは、アグファ・デュオスキャン (Agfa DUOSCAN) デンシトメータ (分解能400) で走査した。ゲルの電子画像をフォトショップ (Photoshop) (アドビ (Adobe)) ソフトウェアで記録した。画像は、tiff (約5 Mバイト/ファイル) 及びjpeg (約50 Kバイト/ファイル) フォーマットで保存した。ゲルは、MS分析に使用するまで4に保持した。

30

【0050】

MALDI-MS

選択した1.2 mm直径のスポットを、手製のスポットピッカー (ヨーロッパ特許出願EP 1,384,994に記載) で切り出し、96ウェルマイクロタイタープレートに入れて、各ゲル片を、サイバイ (登録商標) - ウェル (CyBi<sup>TM</sup>-Well) 装置 (サイビオ社 (Cybio AG)、イエーナ (Jena)、ドイツ) 中で50 mM重炭酸アンモニウム中の30%アセトニトリル100 µlにより脱染色した。脱染色後、ゲル片をH<sub>2</sub>O 100 µlで5分間洗浄し、スピードヴァク (speedvac) エバポレーター中で加熱することなく45分間乾燥した。各乾燥ゲル片は、トリプシン (ロシュ・ダイアグノスティクス、マンハイム、ドイツ) 50 ngを含む1 mM重炭酸アンモニウム5 µlで再水和した。室温で16時間後、0.3%トリフルオロ酢酸を含む50%アセトニトリル20 µlを各ゲル片に加えた。ゲル片は、絶えず振盪しながら15分間インキュベートした。ペプチド混合物 (1.5 µl) は、0.025

40

50

% - シアノ - 4 - ヒドロキシケイ皮酸 (シグマ (Sigma)) からなり、かつ標準ペプチドの脱 Arg - ブラジキニン (シグマ、20 nM、904.4681 Da) 及び副腎皮質刺激ホルモン断片 18 ~ 39 (シグマ、20 nM、2465.1989 Da) を 65 % エタノール、32 % アセトニトリル、及び 0.03 % トリフルオロ酢酸中に含むマトリックス溶液 1  $\mu$ l と同時に、アンカーチップ (AnchorChip) (登録商標) に適用した。試料の適用は、サイバイ - ウェル装置により実行した。試料は、飛行時間質量分析機 (ウルトラフレックス (Ultraflex) TOF - TOF、ブルーカー・ダルトニクス (Bruker Daltonics)) でリフレクトロンモードで分析した。20 kV の加速電圧を使用した。タンパク質は、ペプチド - 質量マッチングに基づいて同定した。

#### 【0051】

10

MALDI 質量スペクトルのピーク注釈

質量分析データは、測定器のベースラインを決定するために、低域中央パラメトリックスプラインフィルターで 2 回フィルターにかける。ベースラインからの平滑化残差平均標準偏差を、データにおける測定器ノイズレベルの推定値として使用する。ベースライン補正、及びノイズを超えるレベル座標におけるデータの修正後、ベースラインから最大偏差のデータ点を使用して、予想ペプチドピークを検出するための非線形 (Levenberg-Marquardt) データフィッティング手順の種を蒔いた。具体的には、このフィット手順は、ピーク高、分解能、及びモノアイソトピック質量によりパラメーター化された最良のフィッティングの平均理論ペプチド同位体分布を作り出すことを試みる。有意なフィットへの収斂は、シグマ値を追跡することにより常法で求められる。収斂の成功後、データ点の 1/3 をランダム交換しての 16 回反復を用いるブートストラップ手順を利用して、求めたパラメーターの誤差を推定する。得られるフィットは、データから引き算して、フィット付近のノイズレベルを、外挿ノイズレベル及びピークフィットからの偏差の合計に調整し、そしてノイズのレベルを 5 倍を超えて上まわる候補ピークを見い出すことができるまで、次のピークを見い出すため、このプロセスを繰り返す。このプロセスは、50 を超えるデータピークが見い出されたら終える。ゼロ次及び一次の質量への飛行時間の変換は、検出内部標準ピークからの直線外挿を用いて補正し、そしてモノアイソトピック質量値の信頼区間は、ピークと標準の質量の正確さから推定する。

20

#### 【0052】

コンピュータ内タンパク質消化物へのスペクトルピークの確率的マッチング

30

質量スペクトルのピーク質量リストは、タンパク質全体の配列データベースについての理論消化物と直接比較する。各理論消化物について、 $[1 - (1 - NP(p_i))]^{cMatches}$  (ここで、N は、理論消化物中のペプチドの数であり、P(p\_i) は、配列データベース中の全てのペプチドのカウントにより割ったピークのモノアイソトピック質量に関する信頼区間と一致するペプチドの数であり、そして cMatches は、消化物と質量スペクトルの間の一致の数である) を計算する。この値は、消化物とスペクトルの間の擬陽性一致が得られる確率に比例することを証明することができる。確率値は、スペクトルピークから高い有意性で一致が得られるように更にフィルターにかける。同定の第 1 ラウンド後、同一条件下で獲得される質量スペクトルについての同定の偏差を使用することにより、質量への飛行時間の変換の第 2 及び第 3 次の条件を補正する。得られる質量値は、ほぼ 10 ppm 未満の絶対偏差を持つ。次にこれらの質量値は、最終ラウンドのマッチングに使用するが、ここでは、 $0.01 / NProteins$  (ボンフェローニ補正による 1 % 有意レベル) 未満の  $P_{miss}$  を持つ全ての一致が許容される。

40

#### 【0053】

結果

コラーゲンアルファ I (IV) (コラーゲン IV; スイスプロット (Swissprot) アクセション番号 P12109; O00117; O00118; Q14040; Q14041; Q16258) は、専らインスリン抵抗性患者 2 名 (IR2 及び IR3、表 3 を参照のこと) において検出された。1 名の患者 (IR2) では、スポットは第 2 レベル (2 倍希釈試料) で検出されたが、一方第 2 の患者 (IR3) では、タンパク質は、第 4 レベル (8

50

倍組合せ希釈)で2回検出された。同定の数は、希釈係数と共に増加し(この場合には、それぞれ1と4)、そしてそれぞれ試料で同定されたタンパク質スポットの総数について補正した。

【0054】

コラーゲンIVレベルはまた、免疫検定法(バイオトリン(Biotrin)コラーゲンIV EIA;カタログ番号BIO82;バイオトリン(Biotrin)、ダブリン、アイルランド)を供給業者のプロトコールにより使用することによって測定した。2つの検定法からの結果は、表3において比較する。

【0055】

【表3】

患者	IS1	IS2	IS3	IR1	IR2	IR3
組合せ連続希釈	0	0	0	0	10	1
コラーゲンIV EIA (ng/ml)	108	111	139	86	208	158

10

【0056】

表3: 組合せ連続希釈からの結果と免疫検定法からの結果との比較。

IS = インスリン感受性患者、IR = インスリン抵抗性患者。

組合せ連続希釈: 同定の数は、希釈係数と総スポットカウントについて調整した。

免疫検定法(コラーゲンIV EIA): コラーゲンIVレベルは、バイオトリン・コラーゲンIV EIAにより測定した。与えられた結果は、二重反復測定の平均である。

20

【0057】

記載されたプロテオミク方法論の検出の限界は、約150 ng/mlの免疫検定法の限界を上まわる。このレベルより上で、タンパク質を検出することができ、記載された同定方法と対にした組合せ連続希釈により粗く定量することができる。絶対定量化は観察されていないが、幾つかの順位相関が存在する、即ち、最高レベルと2番目に高いレベルの試料が正しく同定された。

【0058】

プロテオミクス型大規模タンパク質同定と組合せた、組合せ連続希釈法は、並行して数百のタンパク質を定量するための、及び(例えば、バイオマーカーの同定研究のために)ディファレンシャルタンパク質発現分析において使用することができる、濃度が顕著に異なるタンパク質を同定するための、効率的な手段である。

30

【0059】

本発明の実施態様は以下のとおりである。

1. 生体分子の複合混合物中のある生体分子の定量方法であって、
  - a. それぞれ少なくとも1つの識別できる成分を含む、生体分子の混合物の分画の少なくとも2つの画分を用意すること、
  - b. 画分を組合せ連続希釈に付すこと、
  - c. 各原画分及び各希釈画分中の生体分子を、感度閾値及び同一性情報を提供する方法により、検出及び同定すること、並びに
  - d. それぞれの希釈係数を考慮して各希釈レベルで各画分中の生体分子の同定の数を集約することにより、複合混合物中の生体分子を定量することを含む方法。
2. d)の合計を、全ての希釈レベルでの全ての画分中の全ての生体分子の同定の総数で割る、上記1の方法。
3. a)の画分の数、生体分子の混合物の複雑さに依存する、上記1又は2の方法。
4. a)の画分の数、生体分子の複合混合物中の異なる生体分子の濃度に依存する、上記1~2のいずれか1の方法。
5. a)の画分の数、c)の検出及び同定方法に依存する、上記1~2のいずれか1の

50

方法。

6. b) の希釈レベルの数が、画分中の生体分子の出発濃度に依存する、上記 1 ~ 5 のいずれか 1 の方法。

7. b) の希釈レベルの数が、分画後の画分の数に依存する、上記 1 ~ 5 のいずれか 1 の方法。

8. b) の希釈レベルの数が、検出及び同定方法の検出限界に依存する、上記 1 ~ 5 のいずれか 1 の方法。

9. 生体分子が、(n - 1) 個以下の画分に存在する(ここで、n は、画分の総数であり、かつ n は、2 以上である)、上記 1 ~ 8 のいずれか 1 の方法。

10. 生体分子が、2 つの画分に存在する、上記 1 ~ 8 のいずれか 1 の方法。

10

11. 生体分子が、1 つの画分に存在する、上記 1 ~ 8 のいずれか 1 の方法。

12. c) の検出及び同定方法が、二次元ゲル電気泳動を含む、上記 1 ~ 11 のいずれか 1 の方法。

13. c) の検出及び同定方法が、更に質量分析を含む、上記 12 の方法。

14. c) の検出及び同定方法が、免疫検定法を含む、上記 1 ~ 11 のいずれか 1 の方法。

15. c) の検出及び同定方法が、質量分析と組合せたガスクロマトグラフィーを含む、上記 1 ~ 11 のいずれか 1 の方法。

16. c) の検出及び同定方法が、特異的に標識した分子の実体による電気泳動を含む、上記 1 ~ 11 のいずれか 1 の方法。

20

17. a) の画分が、クロマトグラフィー分画法、超遠心分離法、タンパク質沈降法、又は免疫沈降法により分離される、上記 1 ~ 16 のいずれか 1 の方法。

18. 生体分子が、ポリペプチドである、上記 1 ~ 17 のいずれか 1 の方法。

19. 生体分子が、ポリヌクレオチドである、上記 1 ~ 17 のいずれか 1 の方法。

20. 生体分子が、炭水化物である、上記 1 ~ 17 のいずれか 1 の方法。

21. 生体分子が、脂質である、上記 1 ~ 17 のいずれか 1 の方法。

22. 生体分子が、糖タンパク質である、上記 1 ~ 17 のいずれか 1 の方法。

23. 生体分子が、リポタンパク質である、上記 1 ~ 17 のいずれか 1 の方法。

【図面の簡単な説明】

【0060】

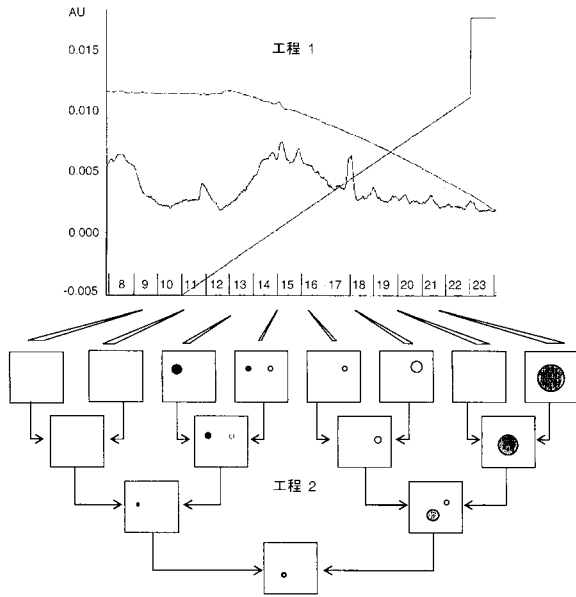
30

【図1】本発明の方法を示す。

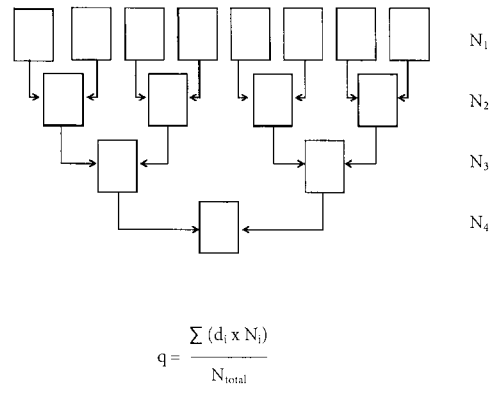
【図2】生体分子の相対量の計算を示す。

【図3】タンパク質、グリコーゲンホスホリラーゼ(a)、ビメンチン(b)及び熱ショックタンパク105(c)に関する同定の数を示す。

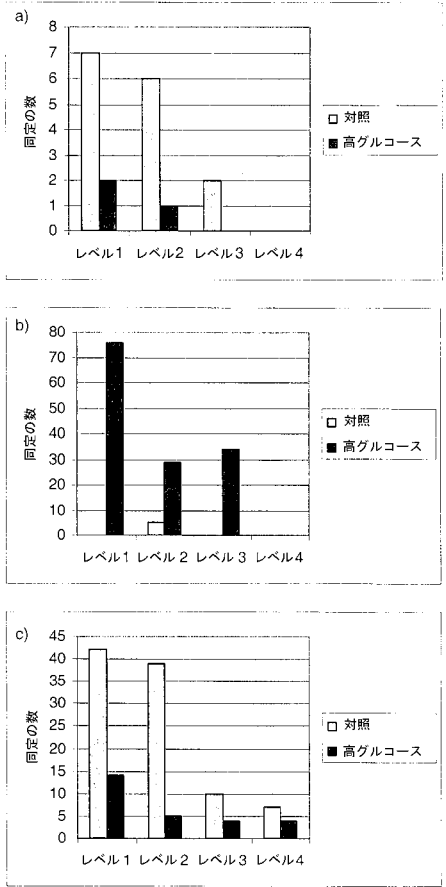
【 図 1 】



【 図 2 】



【 図 3 】



## フロントページの続き

(51) Int.Cl. <sup>7</sup>	F I	テーマコード(参考)
G 0 1 N 33/50	G 0 1 N 33/483	Z
G 0 1 N 33/566	G 0 1 N 33/50	P
G 0 1 N 33/68	G 0 1 N 33/566	
	G 0 1 N 33/68	
	G 0 1 N 27/26	3 1 5 H

(72)発明者 シュテファン・エヴァース

ドイツ国、デー - 7 9 3 7 9 ミュルハイム、パウラ - ヘレンヴェッガー - シュトラーセ 2 3

(72)発明者 ハノー・ランゲン

ドイツ国、デー - 7 9 5 8 5 シュタイネン、イム・ヴェルフィッシュビュール 2 7 / 4

Fターム(参考) 2G045 BB07 BB10 BB20 BB41 CA26 CB01 CB03 CB08 DA12 DA13

DA30 DA36 DA44 DA60 DA62 FA36 FB02 FB03 FB05 FB06

JA01

## 【外国語明細書】

## 1. Title of Invention

Quantitation by serial combinatorial dilution

## 2. Detailed Description of Invention

The present invention relates to a method for the quantification of a biomolecule in a complex mixture of biomolecules comprising fractionation of the complex mixture into fractions with subsequent serial combinatorial dilution of the fractions and detection of the biomolecules in each original fraction and each diluted fraction by a method with a defined sensitivity threshold and identification capabilities.

A current method for detection of a biomolecule (for example a protein) is the two dimensional gel electrophoresis with subsequent volumetric analysis of the stained gel. However, it is difficult to determine the quantity of the analyzed biomolecule, especially if its quantities in different samples shall be compared. To account for the inter-sample variation in biomolecule concentrations the gels have to be processed in parallel and a gel-to-gel-matching has to be done.

There is a need of a simple method for quantification of biomolecules in a complex mixture of biomolecules.

The present invention provides a method for the quantification of a biomolecule in a complex mixture of biomolecules comprising

- a. providing at least two fractions derived from the fractionation of the complex mixture of biomolecules comprising each at least one distinct biomolecule component,
- b. subjecting the fractions to a serial combinatorial dilution step,
- c. detecting and identifying the biomolecule in each original fraction and each diluted fraction by a method with a stable and well defined sensitivity threshold and identity information, and
- d. quantifying the biomolecule in the complex mixture of biomolecules by summarizing the number of identifications of the biomolecule in each

fraction on each dilution level in consideration of the respective dilution factor.

For the purpose of normalization the sum of d) may be divided by the total number of identifications of all biomolecules in all fractions on all dilution levels (dilution levels of original fractions and diluted fractions).

The method of the present invention for the quantification of a biomolecule provides a relative quantification of one or more biomolecules in a complex mixture of biomolecules from one source compared to the respective biomolecules in a complex mixtures from other sources.

The complex mixture of biomolecules may be derived from any source comprising biological sources comprising cells, cell culture supernatants, biological fluids such as serum, plasma, urine, bronchial lavage fluid, sputum, biopsies like cerebrospinal fluid. The complex mixture of biomolecules comprises at least two different biomolecules. The biomolecule in the present invention may be any biomolecule comprising polynucleotides, polypeptides, carbohydrates, lipids, glycoproteins, lipoproteins or other modified forms or metabolites thereof. The detection and identification method can be restricted to a single type of biomolecules, or can detect and analyze several classes of biomolecules at one time.

The fractionation method used in the method of the present invention should efficiently separate the complex mixture of biomolecules into distinct fractions. Preferably, the complex mixture of biomolecules is fractionated into distinct fractions with each different biomolecule only being present in not more than  $n$  minus one fractions wherein  $n$  is the total number of fractions and  $n$  is equal or higher than two. Preferably, the different biomolecules are present in two different fractions, more preferably in one fraction. The fractionation method which may be used in the method of the present invention may be selected from any method suitable for separation of a complex mixture of the targeted type of biomolecules. The fractionation method which may be used in the method of the present invention may be selected from the group comprising fractionation based on adsorption, gravity or sedimentation velocity, electrophoretic fractionation or combinations of these methodologies. For example, in the case of proteins as the target molecule the group includes but is not limited to chromatographic fractionation, ultracentrifugation, protein precipitation, affinity purification, or immunoprecipitation. In the case of peptides (for example obtained from proteolytic digests) as the target molecule the group includes but is not limited to high pressure liquid chromatography (HPLC).

The fractions are then subjected to a serial combinatorial dilution. The serial combinatorial dilution requires at least two fractions to start with. Preferably, the complex mixture of biomolecules is fractionated in as many fractions as necessary to allow a detection and identification of a sufficient number of different targeted biomolecules in the subsequent detection step. Preferably, the number of original fractions is not a prime number, more preferably the number of original fractions is even, and preferably, each initial fraction comprises at least one distinct component.

The number of the fractions to start with the serial combinatorial dilution is dependent on the complexity of the mixture of biomolecules, on the concentration of the individual biomolecules in the complex mixture of biomolecules, the efficiency of the separation methodology, and on the detection and identification method used after the fractionation and the serial combinatorial dilution.

For the serial combinatorial dilution at least two different fractions containing at least one different biomolecule are combined. Preferably, two fractions are combined. This will change the concentration of a biomolecule in the pooled fraction according to the quotient of the dot product of the concentration of the biomolecule in each fraction with the volume of each fraction by the total volume of all combined fractions. In general, this will result in a smaller concentration of any biomolecular component in the diluted fraction compared with the maximum concentration of the respective biomolecule in the original fractions. In the following dilution steps, the concentrations of the individual proteins decrease till they fall below the sensitivity threshold of the subsequent detection and identification method.

The number of dilution steps depends on the starting concentration of the biomolecules in the original fractions, the number of original fractions after fractionation and the detection limit of the detection and identification method.

The method of the present invention further comprises a detection and identification method. The detection method has to feature a defined sensitivity threshold and to provide identity information about the detected biomolecule. Thereby, the presence or absence of a specific biomolecule can be determined in an original fraction or a diluted fraction. The detection and identification method of the present invention may rely on the chemical composition, structure, or sequence of the biomolecule and the physico-chemical or enzymatic properties resulting therefrom. These include hybridization with a specific probe, reaction with a specific antibody or lectin, enzymatic or chemical reaction with a specific molecular probe, isoelectric point, molecular weight, molecular masses of fragments resulting from enzymatic digestion of the biomolecule, NMR spectrum or combinations thereof. For the example of protein

quantitation, the detection and identification method of the present invention may be selected from the group comprising combinations of one- or two-dimensional gel electrophoresis with mass spectrometry, immunoassays (e.g. western blot), gas chromatography combined with mass spectrometry (GS/MS) or electrophoresis with specifically labeled molecular entities, e.g. fluorescent, chemical (e.g. biotin), or radioactive tags.

To derive the quantitation of a biomolecule the identifications or the specific fingerprints (peptide mass fingerprints) of the fractions of each dilution step are calculated whereby the respective dilution factor for each dilution step is considered. The resulting number of identifications of the biomolecule is summarized for all dilution levels. For the purpose of normalization this sum may be divided by the total number of identifications of all biomolecules in all fractions (original fractions and diluted fractions).

$$\text{Relative Quantity (q)} = \frac{\sum (d_i \times N_i)}{N_{\text{total}}}$$

wherein  $N_i$  is the number  $N$  of identifications of an individual biomolecule at dilution level  $i$ ,  $d_i$  the dilution factor  $d$  of the respective dilution level  $i$  and  $N_{\text{total}}$  the total number  $N$  of identifications of all biomolecules in all fractions on all dilution levels.

This method is independent of the properties of the various biomolecules. Polynucleotides may be processed as well as polypeptides or carbohydrates. A further advantage of this method is that it combines quantification with the identification of a biomolecule in a simple manner without the need for additional efforts targeted at biomolecule quantitation. Moreover, if the quantity of a biomolecule derived of one source shall be compared with the one of another source the mixtures of biomolecules may processed separately of each other.

Having now generally described this invention, the same will become better understood by reference to the specific examples, which are included herein for purpose of illustration only and are not intended to be limiting unless otherwise specified, in connection with the following figures.

Figure 1 illustrates the method of the present invention: In a first step the complex mixture is fractionated into different fractions. These fractions are then subjected to a serial combinatorial dilution. In a second step a biomolecule is detected by for example two dimensional gel electrophoresis on the sample pools with subsequent mass spectrometric identification. (AU: Absorption Unit; 8 to 23: Fractions)

Figure 2 shows the calculation the relative quantity of a biomolecule.  
 $q$ : relative quantity of a biomolecule;  $N_i$ : the number  $N$  of identifications of an individual biomolecule on dilution level  $i$ ;  $d_i$ : the dilution factor  $d$  of the respective dilution level  $i$ ;  $N_{total}$ : the total number  $N$  of identifications of all biomolecules in all fractions on all dilution levels. (Scheme: N1: undiluted, N2: 2-fold dilution, N3: 4-fold dilution, N4: 8-fold dilution)

Figure 3 shows the number of identifications for the proteins glycogen phosphorylase (a), vimentin (b) and the heat shock protein 105 (c) in two dimensional electrophoresis gels from level 1 (no dilution), level 2 (2-fold dilution), level 3 (4-fold dilution), and level 4 (8-fold dilution). The values were added up from experiments carried out in triplicate. (Control: 5mM Glucose; high Glucose: 10mM)

## Examples

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated.

### Example 1:

#### *Cell Culture*

INS-1 cells (Asfari, Janjic et al. 1992) were cultured in RPMI 1640 medium (Invitrogen) supplemented with 10% FCS (Invitrogen, heat inactivated) 10 mM HEPES solution (Invitrogen), 1 mM Na pyruvate (Sigma); 50  $\mu$ M beta-mercaptoethanol (Sigma), 1% Penicillin / Streptomycin solution (SIGMA), and low (5 mM) or high (10 mM) concentrations of glucose (SIGMA). Cells were generally cultivated at low glucose concentrations. For preparative culture, the cells were split and then incubated in low-glucose medium until cells were confluent. The medium was then changed to either low-glucose or high-glucose medium and incubations were continued for four days. For

harvesting, cells were first washed once with Hanks Balanced Salt Solutions (HBSS, Invitrogen) and then covered with Trypsin/EDTA solution for 1 – 2 min until cells become rounded and detach from the bottle surface. The Trypsin/EDTA solution was discarded and the cells were suspended in Trypsin Inhibitor Solution (SIGMA), transferred to centrifuge tubes and centrifuged at 1200 x g for 5 min. After this, the cells were washed three times in HBSS, again using the same centrifuge parameters. The supernatant was aspirated and discarded and the pellet was stored frozen at -80°C until used for the preparation of cytosol.

### *Preparation of Cytosol*

All solutions were cooled to 4°C (except for HBSS) and all steps were carried out in a cooled environment (ice bath). Ca  $10^8$  cells were resuspended in cell homogenization medium (CHM; 150 mM MgCl<sub>2</sub>, 10 mM KCl, 10 mM Tris, 0.25 M glucose, 1 mM EDTA, pH 7.4) and left on ice for 2 min. The cell suspension was transferred to a Potter-Elvehjem homogenization vessel. The cold pestle of a Potter-Elvehjem homogenizer was attached to an overhead high-torque electric motor and the cells were homogenized using 10 strokes at 1000 rpm. The efficiency of the homogenization (> 90% of broken cells) was confirmed by phase-contrast microscopy. Cell debris and nuclei were removed by centrifuging for 5 min at 1000 x g. The mitochondria were separated by centrifugation at 5000 x g. The enriched cytosolic fraction was finally recovered by centrifuging at 200000 x g and by transferring the supernatant to a clean tube. The final protein concentration in the preparation was 2.5 – 5.0 mg/ml.

### *Chromatographic Fractionation*

All fractionation steps were carried out using an ÄKTAexplorer 10 chromatography system (Amersham) at room temperature. The cytosol preparations (10 mg of total protein) were passed through a 0.45 µm Millex-HV syringe-driven filter unit and the loaded onto desalting columns (three 5 ml HiTrap desalting columns connected in series, Amersham). The proteins were eluted using Buffer A (25 mM NaHPO<sub>4</sub><sup>-</sup> pH 7.5; 1 mM EDTA; 0.5 mM dithioerythritol; 1x Complete EDTA-free (Protease inhibitor cocktail tablets from Roche Diagnostics; pH adjusted to 7.5) using a flow rate of 1.5 ml/min. Proteins were recovered in a 20 ml injectionloop using the increase in UV absorption (280 nm) and the minimum in conductivity as boundaries for the protein fraction. The proteins were then separated by anion exchange chromatography using a TSK DEAE-5PW 7.5cmx7.5mm column (TOSOH BIOSEP) at a flow rate of 1 ml/min. Buffer A was

used as the binding buffer, buffer B (25 mM NaHPO<sub>4</sub><sup>-</sup> pH 7.5; 1 mM EDTA; 0.5 mM dithioerythritol; 1x Complete EDTA-free (Protease inhibitor cocktail tablets from Roche Diagnostics; 1 M NaCl, pH adjusted to 7.5) as the elution buffer. The sample was loaded onto the column and unbound material was washed off with 7 column volumes (CV) of Buffer A. The bound proteins were then eluted by three-segment gradient (1<sup>st</sup> segment: 0 – 11% Buffer B in 3 CV; 2<sup>nd</sup> segment: 11 – 30% Buffer B in 10 CV; 30 – 50% Buffer B in 1.5 CV). Finally, the column was washed with 5 CV of 50% Buffer B. Fractions of 1 ml were collected and combined to form eight pools plus the flow-through. The conductivity boundaries were: FT: UV280 increase to increase in conductivity; 1 (start of conductivity-increase to 12 mS); 2 (12 to 15 mS); 3 (15 to 18 mS); 4 (18 to 21 mS); 5 (21 to 24 mS); 6 (24 to 27 mS); 7 (27 to 30 mS); 8 (30 to 40 mS).

### *Two-Dimensional Electrophoresis*

The fractions were concentrated and desalted by reversed phase chromatography using self-packed syringe-driven minicolumns (MoBiTec M1002) filled with 100 mg of POROS 20 R1 material (PerSeptive Biosystems). The columns were washed with 10 ml of 0.1% Trifluoroacetic Acid (TFA) and with 70% Acetonitrile/ 0.1% TFA. After loading the sample, the columns were washed with 10 ml of 0.1% TFA and eluted with 2 ml of 70% Acetonitrile/ 0.1% TFA. The eluate was then dried in a SpeedVac evaporator and taken up in IEF Sample Buffer (7 M Urea, 2 M Thiourea, 50 mM Tris pH 7.5, 2 % (w/v) CHAPS, 0.4 % (w/v) Dithioerythritol, 0.5% (w/v) ampholytes). Aliquots containing 0.5 mg of protein were set aside from each fraction and labeled as Sample 1 to 8. The following samples were prepared from the remainder of the fractions: Sample 9: 0.25 mg fraction 1 + 0.25 mg fraction 2; Sample 10: 0.25 mg fraction 3 + 0.25 mg fraction 4; Sample 11: 0.25 mg fraction 5 + 0.25 mg fraction 6; Sample 12: 0.25 mg fraction 7 + 0.25 mg fraction 8; Sample 13: 0.125 mg fraction 1 + 0.125 mg fraction 2+ 0.125 mg fraction 3 + 0.125 mg fraction 4; Sample 14: 0.125 mg fraction 5 + 0.125 mg fraction 6+ 0.125 mg fraction 7 + 0.125 mg fraction 8; Sample 15: 0.0625 mg fraction 1 + 0.0625 mg fraction 2 + 0.0625 mg fraction 3+ 0.0625 mg fraction 4+ 0.0625 mg fraction 5+ 0.0625 mg fraction 6+ 0.0625 mg fraction 7+ 0.0625 mg fraction 8. Thus, samples 1 – 8 contain 0,5 mg of protein fractions, samples 9 – 12 each correspond to a two-fold dilution of these samples, samples 13 and 14 to a four-fold, and sample 15 to an eight-fold dilution of these original fractions. Isoelectric Focusing was performed using immobilized pH gradient (IPG) strips with a pH range from 3 to 10 (IPG 3-10L; Amersham) in a Protean IEF Cell (BioRad) at 20°C. The dried strips were re-hydrated in a solution containing 7 M Urea, 2M Thiourea, 2 % (w/v) CHAPS, 0.4 % (w/w) Dithioerythritol, and 0.5 % (w/v) ampholytes. The protein fractions were cup-loaded at the cathodic end of the strip. The

voltage was linearly increased to 5000V over 8h, followed by a 5000 V plateau for 10 h. The strips were equilibrated and alkylated by successive washes in Equilibration Solution I (6 M Urea, 50 mM Tris pH 7.5, 30 % Glycerol, 2.0 % SDS, 30 mM Dithioerythritol) and Equilibration Solution II (6 M Urea, 50 mM Tris pH 8.8, 30 % Glycerol, 2.0 % SDS, 0.23 M Iodoacetamide) for 10 min each. The strips were loaded onto 11% Acrylamide/PDA (37:1) gradient gels (240 mm x 200 mm x 1.5 mm). The proteins were resolved by electrophoresis at 80V O/N in an ETTAN Dalt Electrophoresis apparatus (Amersham) with constant cooling (20°C).

#### *Gel Staining and Processing*

The gels were fixed in 50% methanol/ 10% acetic acid and stained with Coomassie Blue (Colloidal Blue, Invitrogen, Carlsbad, CA) overnight followed by multiple washes in ultra-pure water for 7h total. The gels were scanned and spots with a diameter of 1.2 mm were excised using an automatic spot picking device. The spots were de-stained in a solution containing 100 mM Ammonium hydrogen carbonate and 30 % Acetonitrile. The dried de-stained gel pieces were digested in 5  $\mu$ l of a 10  $\mu$ g/ml Trypsin solution (Roche Diagnostics) overnight at room temperature. After addition of 10  $\mu$ l of ultra-pure water, proteins were extracted with 5  $\mu$ l of a solution containing 75% Acetonitrile and 0.3% (v/v) TFA. The peptide solution was spotted onto a MALDI target together with  $\alpha$ -Cyano-4-hydroxycinnamic acid as matrix.

#### *Mass Spectrometry and Protein Identification*

Peptide masses were measured on a Bruker Ultraflex Instrument (Bruker, Bremen, Germany), using ACTH and Bradykinin as internal mass standards. As explained below, monoisotopic peptide masses were automatically detected from the mass spectra and compared to theoretical masses of peptides derived from an in-silico tryptic digest of all proteins from a database of protein sequences (e.g. SwissProt, or NCBI rat genome draft).

#### *Peak annotation for MALDI mass spectra*

The mass spectrometric data is two times filtered with a low-pass median parametric spline filter in order to determine the instrument baseline. The smoothed

residual mean standard deviation from the baseline is used as an estimate of the instrument noise level in the data.

After baseline correction and rescaling of the data in level-over-noise coordinates, the data point with the largest deviation from the baseline is used to seed a non-linear (Levenberg-Marquardt) data fitting procedure to detect possible peptide peaks. Specifically, the fit procedure attempts to produce the best fitting average theoretical peptide isotope distribution parameterized by peak height, resolution, and monoisotopic mass. The convergence to a significant fit is determined in the usual way by tracking sigma values.

After a successful convergence, an estimate for the errors of the determined parameters is produced using a bootstrap procedure using sixteen repeats with a random exchange of 1/3 of the data points.

The resulting fit is subtracted from the data, the noise level in the vicinity of the fit is adjusted to the sum of the extrapolated noise level and the deviation from the peak fit, and the process is iterated to find the next peak as long as a candidate peak more than five times over level of noise can be found. The process is stopped when more than 50 data peaks have been found.

The zero and first order of the time-of flight to mass conversion are corrected using linear extrapolation from detected internal standard peaks, and confidence intervals for the monoisotopic mass values are estimated from the mass accuracies of the peaks and standards.

#### *Probabilistic matching of spectra peaks to in-silico protein digests*

Peak mass lists for mass spectra are directly compared to theoretical digests for whole protein sequence databases. For each theoretical digest,  $[1 - \prod(1 - N P(\pi))]^{cMatches}$  is calculated, where  $N$  is the number of peptides in the digest,  $P(\pi)$  is the number of peptides that match the confidence interval for the monoisotopic mass of the peak divided by the count of all peptides in the sequence database, and  $cMatches$  is the number of matches between digest and mass spectrum. It can be shown that this value is proportional to the probability of obtaining a false positive match between digest and spectrum. Probability values are further filtered for high significance of the spectra peaks that produce the matches. After a first round of identifications, deviations of the identifications for mass spectra acquired under identical conditions are used to correct the second and third order terms of the time-of-flight to mass conversion. The resulting

mass values have mostly absolute deviations less than 10ppm. These mass values are then used for a final round of matching, where all matches having a  $P_{\text{mism}}$  less than  $0.01/N_{\text{Proteins}}$  (1% significance level with Bonferoni correction) are accepted.

### *Database analysis*

For each protein in the database, the number of identifications per 2D-PAGE gel analyzed in this study was counted. In this example the dilution level 1 was set as reference. Then, the following values were derived:

- Number of identifications in dilution level 1 (undiluted samples, samples 1 – 8) =  $N_1$
- Number of identifications in dilution level 2 (2-fold dilution, samples 9 – 12) =  $N_2$
- Number of identifications in dilution level 3 (4-fold dilution, samples 13,14) =  $N_3$
- Number of identifications in dilution level 4 (8-fold dilution, sample 15) =  $N_4$

As expected, for most proteins the  $N$  values decreased roughly two-fold from layer to layer. To account for the dilution factors and to derive a rough absolute quantity for each protein, a quantity value  $q$  was calculated as follows:

$$q = \frac{(N_1 + 2 \times N_2 + 4 \times N_3 + 8 \times N_4)}{\text{total number of identified protein spots for all samples of the same source on all dilution levels}}$$

The division by the total number of identification for all samples of the same source was introduced to account for inter-sample variations in protein concentration.

For each protein, the  $q$  values for both mixture samples (high and low glucose) were calculated and compared.

The following three proteins were chosen as examples for the illustration of the feasibility of this quantification method: Glycogen Phosphorylase (liver form); Vimentin, and Heat shock protein 105 (Table 1, Fig. 3).

	q (5mM Glucose) x 10 <sup>-5</sup>	q (10 mM Glucose) x 10 <sup>-5</sup>
<b>Glycogen phosphorylase</b>		
Experiment 1	112	0
Experiment 2	8	0
Experiment 3	124	44
<b>Vimentin</b>		
Experiment 1	0	2130
Experiment 2	80	305
Experiment 3	0	1758
<b>Heat shock protein 105</b>		
Experiment 1	17	13
Experiment 2	121	39
Experiment 3	200	89

Table 1: relative Quantity (q Values) of the proteins present in the cytosol obtained for the three experiments for three example proteins

### Example 2: Collagen alpha I (IV)

Serum samples from three insulin-resistant and three insulin sensitive patients (Caucasian, female) were fractionated as described below. The Body Mass Index (BMI) and the Glucose Disposal Rate (GDR) as determined by the Euglycemic-Hyperinsulinemic Clamp method (Garvey et al. *Diabetes* 34 (1985) 222-234) are indicated in Table 2. Combinatorial serial dilution was performed as described in the patent application and the resulting samples were subjected to Two-Dimensional-SDS-Polyacrylamide Gel Electrophoresis (2D-PAGE) as described below. All detectable protein spots were excised from each gel. The proteins were digested with trypsin and the resulting peptides subjected to Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-MS). Protein identification was achieved by peptide mass fingerprint analysis as described below and protein lists were compared as described in Example 1.

Patient	Insulin-Sensitive (IS)		Patient	Insulin-Resistant (IR)	
	BMI	GDR		BMI	GDR
IS1	22.4	21.9	IR1	31.3	10.2
IS2	22.4	19.7	IR2	33	11.65
IS3	29.5	20.4	IR3	33.1	8.0

Table 2: Body Mass Index (BMI) and Glucose Disposal Rates (GDR) determined by the euglycemic hyperinsulinemic clamp method of six subjects. As GDRs above 15 are considered as the breakpoint for the determination of insulin resistance, the patients on the left side of the panel are classified as insulin sensitive (IS) and those on the right side as insulin resistant (IR). Plasma from these individuals was analyzed by serial combinatorial dilution followed by 2D-PAGE, spot excision, tryptic digest, MALDI-MS and finally protein identification by peptide mass fingerprint comparison.

### *Sample Preparation*

A method was established to search for Insulin Resistance markers in human plasma by applying proteomics technologies. Plasma is a difficult to analyze by Proteomics

techniques because it includes ca. ten high-abundance proteins, which represent approximately 98% of the total protein mass. The high-abundance proteins, albumin and antibody chains were removed, by applying chromatographic techniques and fractionated the flow through fraction over an ion exchanger. The scheme described comprises three chromatography steps, matrix blue, protein G and ion exchange, and is highly reproducible. All chromatographic steps were performed on an FPLC System (Pharmacia).

*Removal of albumin by affinity chromatography on Mimetic blue and removal of immunoglobulins by affinity chromatography on Protein G*

Human plasma was received from three control individuals and three patients with diabetes type II. Protease inhibitors cocktail (Roche Diagnostics, Mannheim, Germany) was added to the plasma (one tablet to 50 ml). Plasma was diluted three-fold with 25 mM MES, pH 6.0, to reduce the salt concentration and adjust the pH to about 6.0. The two columns, Mimetic blue SA P6XL (50 ml, ProMetic BioSciences Ltd.) and HiTrap Protein G HP (5 ml, Amersham Biosciences) were connected in series and equilibrated with 25 mM MES, pH 6.0. The volume corresponding to approximately one g of plasma protein (15 ml, 66 mg/ml) was filtered through a 0.22 µm filter and applied onto the Mimetic blue column at 5 ml/min. The flow through of this column was directly loaded onto the Protein G column and the flow-through fraction from the latter column was collected (about 120 mg). The two columns were washed with 100 ml of 25 mM MES, pH 6.0 and then they were separated. The Mimetic blue column was eluted with a step gradient of 2 M NaCl in 50 mM Tris-HCl, pH 7.5 and the Protein G was eluted with 100 mM glycine-HCl, pH 2.8 and the eluate was neutralized with 1 M Tris base. The flow through fraction and the two eluates were analyzed by two-dimensional gels and the proteins were identified by MALDI-MS. In the eluate from Mimetic blue, mainly full-length and fragmented albumin were detected. In the eluate from the Protein G column, mainly heavy and light Ig chains were detected. Most of the other plasma proteins were recovered in the flow through fraction.

*Protein fractionation by ion exchange chromatography*

The flow through and the wash fractions from the Mimetic blue and Protein G columns were combined, adjusted to pH 8.0 with 2 M Tris base and were applied onto a HiTrap Q HP column (5 ml, Amersham Biosciences), equilibrated with 50 mM Tris-HCl, pH 8.0 at

5 ml/min. The column was eluted with a linear gradient of increasing salt concentration from 0 to 1 M NaCl in 50 mM Tris-HCl, pH 7.5. Five-ml fractions were collected and analyzed by 1-D gels. Approximately 50 mg of total protein were recovered from this column. On the basis of the gel analysis, the fractions were pooled to form eight pools, so that each pool included about 5 mg of total protein. The pools were concentrated with Ultrafree-15 Centrifugal Filter (5k MWCO, Millipore) and each of the eight pools was analyzed by 2-D gels. About 400 spots from each gel were excised and analyzed by MALDI-MS.

### *2D-PAGE*

Immobilized pH gradient (IPG) strips were purchased from Amersham Biosciences (Uppsala, Sweden). Acrylamide was obtained from Biosolve (Valkenswaard, The Netherlands) and the other reagents for the polyacrylamide gel preparation were from Bio-Rad Laboratories (Hercules, CA, USA). CHAPS was from Roche Diagnostics (Mannheim, Germany), urea from Applichem (Darmstadt, Germany), thiourea from Fluka (Buchs, Switzerland) and dithioerythritol from Merck (Darmstadt).

Samples of 0.5 mg total protein were applied on 3-10 NL IPG strips, in sample cups at their basic and acidic ends. Focusing started at 200 V, and the voltage was gradually increased to 5000 V at 3 V/min, using a computer-controlled power supply and was kept constant for a further 6 h. The second-dimensional separation was performed either on 12% constant SDS polyacrylamide gels (180x200x1.5 mm) at 40 mA per gel. After protein fixation for 12 h in 40% methanol that contained 5% phosphoric acid, the gels were stained with colloidal Coomassie blue (Novex, San Diego, CA, USA) for 24 h. Excess dye was washed from the gels with H<sub>2</sub>O, and the gels were scanned in an Agfa DUOSCAN densitometer (resolution 400). Electronic images of the gels were recorded with Photoshop (Adobe) software. The images were stored in tiff (about 5 Mbytes/file) and jpeg (about 50 Kbytes/file) formats. The gels were kept at 4 °C until used for MS analysis.

### *MALDI-MS*

Selected spots of 1.2 mm diameter were excised with a homemade spot picker (described in European Application EP 1 384 994), placed into 96-well microtiter plates and each gel piece was destained with 100 µl of 30% acetonitrile in 50 mM ammonium bicarbonate in

a CyBi™-Well apparatus (Cybio AG, Jena, Germany). After destaining, the gel pieces were washed with 100  $\mu$ l of H<sub>2</sub>O for 5 min, and dried in a speedvac evaporator without heating for 45 min. Each dried gel piece was rehydrated with 5  $\mu$ l of 1 mM ammonium bicarbonate, that contained 50 ng trypsin (Roche Diagnostics, Mannheim, Germany). After 16 h at room temperature, 20  $\mu$ l of 50% acetonitrile, that contained 0.3% trifluoroacetic acid was added to each gel piece. The gel pieces were incubated for 15 min with constant shaking. A peptide mixture (1.5  $\mu$ l) was simultaneously applied with 1  $\mu$ l of matrix solution, that consisted of 0.025%  $\alpha$ -cyano-4-hydroxycinnamic acid (Sigma), and that contained the standard peptides des-Arg-bradykinin (Sigma, 20 nM, 904.4681 Da) and adrenocorticotrophic hormone fragment 18-39 (Sigma, 20 nM, 2465.1989 Da) in 65% ethanol, 32% acetonitrile, and 0.03% trifluoroacetic acid, to the AnchorChip™. The sample application was performed with a CyBi-Well apparatus. Samples were analyzed in a time-of-flight mass spectrometer (Ultraflex TOF-TOF, Bruker Daltonics) in the reflectron mode. An accelerating voltage of 20 kV was used. Proteins were identified on the basis of peptide-mass matching.

#### *Peak annotation for MALDI mass spectra*

Mass spectrometric data is two times filtered using a low-pass median parametric spline filter in order to determine the instrument baseline. The smoothed residual mean standard deviation from the baseline is used as an estimate of the instrument noise level in the data. After baseline correction and rescaling of the data in level-over-noise coordinates, the data point with the largest deviation from the baseline is used to seed a non-linear (Levenberg-Marquardt) data fitting procedure to detect possible peptide peaks. Specifically, the fit procedure attempts to produce the best fitting average theoretical peptide isotope distribution parameterized by peak height, resolution, and monoisotopic mass. The convergence to a significant fit is determined in the usual way by tracking sigma values. After a successful convergence, an estimate for the errors of the determined parameters is produced using a bootstrap procedure using sixteen repeats with a random exchange of 1/3 of the data points. The resulting fit is subtracted from the data, the noise level in the vicinity of the fit is adjusted to the sum of the extrapolated noise level and the deviation from the peak fit, and the process is iterated to find the next peak as long as a candidate peak more than five times over level of noise can be found. The process is stopped when more than 50 data peaks have been found. The zero and first order of the time-of flight to mass conversion are corrected using linear extrapolation from detected internal standard peaks, and confidence intervals for the

monoisotopic mass values are estimated from the mass accuracies of the peaks and standards.

#### *Probabilistic matching of spectra peaks to in-silico protein digests*

Peak mass lists for mass spectra are directly compared to theoretical digests for whole protein sequence databases. For each theoretical digest,  $[1 - \prod(1 - N P(pi))]^{cMatches}$  is calculated, where  $N$  is the number of peptides in the theoretical digest,  $P(pi)$  is the number of peptides that match the confidence interval for the monoisotopic mass of the peak divided by the count of all peptides in the sequence database, and  $cMatches$  is the number of matches between digest and mass spectrum. It can be shown that this value is proportional to the probability of obtaining a false positive match between digest and spectrum. Probability values are further filtered for high significance of the spectra peaks that produce the matches. After a first round of identifications, deviations of the identifications for mass spectra acquired under identical conditions are used to correct the second and third order terms of the time-of-flight to mass conversion. The resulting mass values have mostly absolute deviations less than 10ppm. These mass values are then used for a final round of matching, where all matches having a  $P_{mism}$  less than  $0.01/NProteins$  (1% significance level with Bonferoni correction) are accepted.

#### *Results*

Collagen alpha I (IV) (Collagen IV; Swissprot accession numbers P12109; O00117; O00118; Q14040; Q14041; Q16258) was exclusively detected in two insulin resistant patients (IR2 and IR3, see Table 3). In one patient (IR2), the spots were detected at the second level (two-fold diluted sample), whereas in the second patient (IR3), the protein was detected twice at the fourth level (eightfold combinatorial dilution). The number of identifications was multiplied with the dilution factor (in this case, one and four, respectively) and corrected for the total number of protein spots identified for the respective sample.

Collagen IV levels were also measured using an immunoassay (Biotrin Collagen IV EIA; Catalogue Number NoBIO82; Biotrin, Dublin, Ireland) following the supplier's protocol. The results from the two assays are compared in Table 3.

Patient	IS1	IS2	IS3	IR1	IR2	IR3
Serial combinatorial dilution	0	0	0	0	10	1
Collagen IV EIA (ng/ml)	108	111	139	86	208	158

Table 3: Comparison of the results from the serial combinatorial dilution with the results from the immunoassay.

IS = Insulin-sensitive patient, IR= Insulin-resistant patient.

Serial combinatorial dilution: The number of identifications were adjusted for dilution factor and total spot count.

Immunoassay (Collagen IV EIA): The Collagen IV levels were determined by the Biotrin Collagen IV EIA was used. The presented results are the mean of duplicate measurements.

The limit of detection of the described proteomic methodology lies above that for the immunoassay at approx. 150 ng/ml. Above that level, proteins can be detected and coarsely quantified by serial combinatorial dilution coupled to the described identification method. Although no absolute quantification is observed, there is some rank correlation, i.e. the samples with the highest and second highest levels were correctly identified.

Serial combinatorial dilution combined with Proteomics type large scale protein identification is an efficient tool to quantify hundreds of proteins in parallel and to identify proteins with marked differences in concentration which can be used in differential protein expression analysis, e.g. for biomarker identification studies.

The embodiment of the present invention is as follows.

1. A method for the quantification of a biomolecule in a complex mixture of biomolecules comprising
  - a. providing at least two fractions of a fractionation of a mixture of biomolecules comprising each at least one distinct component,
  - b. subjecting the fractions to a serial combinatorial dilution,
  - c. detecting and identifying the biomolecule in each original fraction and each diluted fraction by a method providing a sensitivity threshold and identity information and
  - d. quantifying the biomolecule in the complex mixture by summarizing of the number of identifications of the biomolecule in each fraction on each dilution level in consideration of the respective dilution factor.
2. A method according to the above 1, wherein the sum of d) is divided by the total number of identifications of all biomolecules in all fractions on all dilution levels.
3. A method according to the above 1 or 2 wherein the number of fractions of a) depends on the complexity of the mixture of biomolecules.
4. A method according to any one of the above 1 to 2 wherein the number of fractions of a) depends on the concentration of the different biomolecules in the complex mixture of biomolecules.
5. A method according to any one of the above 1 to 2 wherein the number of fractions of a) depends on the detection and identification method of c)
6. A method according to any one of the above 1 to 5 wherein the number of dilution levels of b) depends on the starting concentration of the biomolecules in the fractions.
7. A method according to any one of the above 1 to 5 wherein the number of dilution levels of b) depends on the number of fractions after fractionation.
8. A method according to any one of the above 1 to 5 wherein the number of dilution levels of b) depends on the detection limit of the detection and identification method.

9. A method according to any one of the above 1 to 8 wherein the biomolecule is present in not more than  $n$  minus one fractions wherein  $n$  is the total number of fractions and wherein  $n$  is equal or higher than two.
10. A method according to any one of the above 1 to 8 wherein the biomolecule is present in two fractions.
11. A method according to any one of the above 1 to 8 wherein the biomolecule is present in one fraction.
12. A method according to any one of the above 1 to 11 wherein the detecting and identifying method of c) comprises a two dimensional gel electrophoresis.
13. A method according to the above 12 wherein the detecting and identifying method of c) comprises additionally mass spectrometry.
14. A method according to any one of the above 1 to 11 wherein the detecting and identifying method of c) comprises immunoassays.
15. A method according to any one of the above 1 to 11 wherein the detecting and identifying method of c) comprises gas chromatography combined with mass spectrometry.
16. A method according to any one of the above 1 to 11 wherein the detecting and identifying method of c) comprises electrophoresis with specifically labeled molecular entities.
17. A method according to any one of the above 1 to 16 wherein the fractions of a) are separated by chromatographic fractionation, ultracentrifugation, protein precipitation, or immunoprecipitation.
18. A method according to any one of the above 1 to 17 wherein the biomolecule is a polypeptide.
19. A method according to any one of the above 1 to 17 wherein the biomolecule is polynucleotide.
20. A method according to any one of the above 1 to 17 wherein the biomolecule is a carbohydrate.
21. A method according to any one of the above 1 to 17 wherein the biomolecule is a lipid.

22. A method according to any one of the above 1 to 17 wherein the biomolecule is a glycoprotein.
23. A method according to any one of the above 1 to 17 wherein the biomolecule is a lipoprotein.

### 3. Brief Description of Drawings

Figure 1 illustrates the method of the present invention.

Figure 2 shows the calculation the relative quantity of a biomolecule.

Figure 3 shows the number of identifications for the proteins glycogen phosphorylase (a), vimentin (b)and the heat shock protein 105 (c).

### 1. Claims

1. A method for the quantification of a biomolecule in a complex mixture of biomolecules comprising
  - a. providing at least two fractions of a fractionation of a mixture of biomolecules comprising each at least one distinct component,
  - b. subjecting the fractions to a serial combinatorial dilution,
  - c. detecting and identifying the biomolecule in each original fraction and each diluted fraction by a method providing a sensitivity threshold and identity information and
  - d. quantifying the biomolecule in the complex mixture by summarizing of the number of identifications of the biomolecule in each fraction on each dilution level in consideration of the respective dilution factor.
2. A method according to claim 1, wherein the sum of d) is divided by the total number of identifications of all biomolecules in all fractions on all dilution levels.
3. A method according to claim 1 or 2 wherein the number of fractions of a) depends on the complexity of the mixture of biomolecules.
4. A method according to any one of the claims 1 to 2 wherein the number of fractions of a) depends on the concentration of the different biomolecules in the complex mixture of biomolecules.
5. A method according to any one of the claims 1 to 2 wherein the number of fractions of a) depends on the detection and identification method of c)
6. A method according to any one of the claims 1 to 5 wherein the number of dilution levels of b) depends on the starting concentration of the biomolecules in the fractions.
7. A method according to any one of the claims 1 to 5 wherein the number of dilution levels of b) depends on the number of fractions after fractionation.
8. A method according to any one of the claims 1 to 5 wherein the number of dilution levels of b) depends on the detection limit of the detection and identification method.

9. A method according to any one of the claims 1 to 8 wherein the biomolecule is present in not more than  $n$  minus one fractions wherein  $n$  is the total number of fractions and wherein  $n$  is equal or higher than two.
10. A method according to any one of the claims 1 to 8 wherein the biomolecule is present in two fractions.
11. A method according to any one of the claims 1 to 8 wherein the biomolecule is present in one fraction.
12. A method according to any one of the claims 1 to 11 wherein the detecting and identifying method of c) comprises a two dimensional gel electrophoresis.
13. A method according to claim 12 wherein the detecting and identifying method of c) comprises additionally mass spectrometry.
14. A method according to any one of the claims 1 to 11 wherein the detecting and identifying method of c) comprises immunoassays.
15. A method according to any one of the claims 1 to 11 wherein the detecting and identifying method of c) comprises gas chromatography combined with mass spectrometry.
16. A method according to any one of the claims 1 to 11 wherein the detecting and identifying method of c) comprises electrophoresis with specifically labeled molecular entities.
17. A method according to any one of the claims 1 to 16 wherein the fractions of a) are separated by chromatographic fractionation, ultracentrifugation, protein precipitation, or immunoprecipitation.
18. A method according to any one of the claims 1 to 17 wherein the biomolecule is a polypeptide.
19. A method according to any one of the claims 1 to 17 wherein the biomolecule is polynucleotide.
20. A method according to any one of the claims 1 to 17 wherein the biomolecule is a carbohydrate.
21. A method according to any one of the claims 1 to 17 wherein the biomolecule is a lipid.

22. A method according to any one of the claims 1 to 17 wherein the biomolecule is a glycoprotein.
23. A method according to any one of the claims 1 to 17 wherein the biomolecule is a lipoprotein.

#### 1. Abstract

The invention provides a method for the quantification of a biomolecule in a complex mixture of biomolecules which comprises a fractionation of the mixture of biomolecules providing at least two fractions with at least one distinct component each. These fractions are then subjected to serial combinatorial dilution. Subsequently, the biomolecule is detected and identified in the fractions by a method providing a sensitivity threshold and identify information. The quantity of the biomolecule is determined by summarizing the number of identifications of the biomolecule in each fraction on each dilution level in consideration of the respective dilution factor. For purpose of normalization this sum may be divided by the total number of identifications of all biomolecules in all fractions on all dilution levels.

#### 2. Representative Drawing

None

Fig. 1

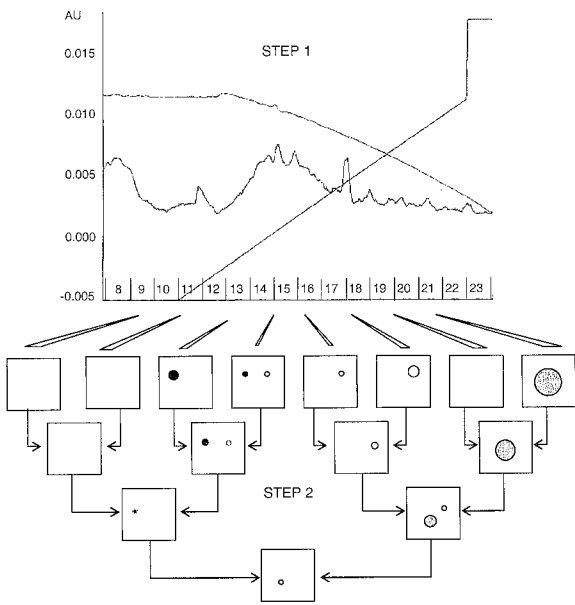


Fig. 2

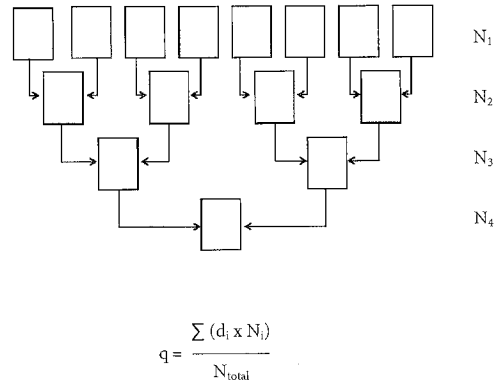
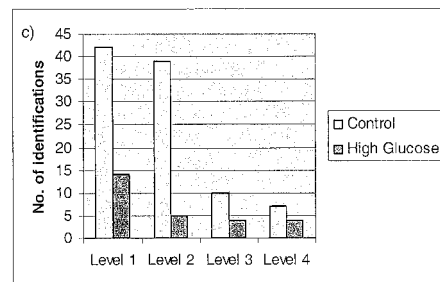
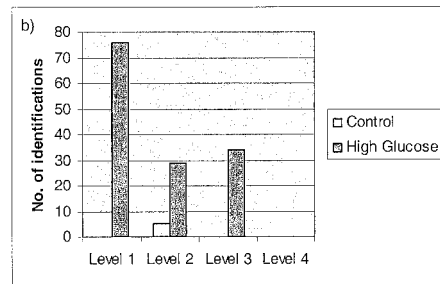
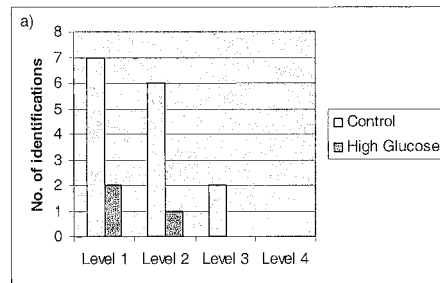


Fig. 3



专利名称(译)	通过连续组合稀释的定量方法		
公开(公告)号	<a href="#">JP2005189241A</a>	公开(公告)日	2005-07-14
申请号	JP2004367043	申请日	2004-12-20
申请(专利权)人(译)	F.霍夫曼 - 罗氏公司		
[标]发明人	ペーターベルント シュテファンエヴァース ハノーランゲン		
发明人	ペーターベルント シュテファンエヴァース ハノーランゲン		
IPC分类号	G01N27/62 C12Q1/06 C12Q1/68 G01N1/38 G01N27/447 G01N30/02 G01N30/72 G01N33/48 G01N33/483 G01N33/50 G01N33/53 G01N33/566 G01N33/68 G01N33/92 G01N35/00		
CPC分类号	G01N33/6803		
FI分类号	G01N33/48.A G01N27/62.C G01N27/62.V G01N30/72.A G01N33/483.F G01N33/483.Z G01N33/50.P G01N33/566 G01N33/68 G01N27/26.315.H G01N27/447.315.H		
F-TERM分类号	2G045/BB07 2G045/BB10 2G045/BB20 2G045/BB41 2G045/CA26 2G045/CB01 2G045/CB03 2G045/CB08 2G045/DA12 2G045/DA13 2G045/DA30 2G045/DA36 2G045/DA44 2G045/DA60 2G045/DA62 2G045/FA36 2G045/FB02 2G045/FB03 2G045/FB05 2G045/FB06 2G045/JA01 2G041/CA01 2G041/DA04 2G041/EA03 2G041/EA06 2G041/EA11 2G041/EA13 2G041/FA10 2G041/FA11 2G041/FA12 2G041/GA06 2G041/GA07 2G041/GA09 2G041/GA13 2G041/HA01 2G041/HA02 2G041/JA02 2G041/JA07 2G041/JA08 2G041/KA01 2G041/LA01 2G041/LA07 2G041/LA08 2G041/LA10 2G041/LA12		
代理人(译)	津国 肇 筱田文雄		
优先权	2003104775 2003-12-18 EP		
其他公开文献	JP3974130B2		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

要解决的问题：提供一种定量复杂生物分子混合物中生物分子的方法。然后对包含给出至少两个包含至少一种可区分组分的级分的生物分子混合物的级分的级分进行组合系列稀释。随后，通过提供灵敏度阈值和身份信息的方法在级分中检测和鉴定生物分子。考虑到每种稀释因子，通过汇总每个稀释水平下每个馏分中生物分子身份的数量来测量生物分子的量。为了进行标准化，可以将总和除以所有稀释水平下所有馏分中所有生物分子的鉴定总数。 [选型图]图1

