

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 2 日現在

機関番号：14301

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26670015

研究課題名(和文) 高深度プロテオミクスによるヒトプロテオーム完全解析

研究課題名(英文) Complete Analysis of Human Proteome by Deep Proteomics

研究代表者

石濱 泰 (Ishihama, Yasushi)

京都大学・薬学研究科(研究院)・教授

研究者番号：30439244

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：ヒト試料中に発現しているタンパク質すべての同定・定量を可能にする高深度ヒトプロテオーム解析システムの開発を目指し、研究を行った。メートル長シリカモノリスカラムを用いる超高分離能LC-MSシステムに、メートル長HILICカラムを用いた超高分離能分画・間引き法を組み合わせた高深度プロテオーム解析システムを開発した。その結果、間引き法を使った場合には、従来法に比べ、タンパク質同定数を46%増加させることに成功した。また、分画法を用いた場合には、従来法とほぼ同様に約10,000種のタンパク質同定が可能となり、従来法に比べ出発試料量を200分の1に、分析時間を6分の1に抑えることができた。

研究成果の概要(英文)：We aimed to develop a deep proteomics analytical system to identify and quantify complete human proteome expressed in human cells. We developed ultrahigh resolution LC-MS systems using meter-scale monolithic silica capillary columns in combined with meter-scale HILIC columns for fractionating or slicing samples. As a result, when using the slicing mode, the number of identified proteins increased by 46%, compared with the conventional method. Further, using the fractionation method, we successfully identified 10,000 human proteins with 200-fold sensitivity and 6-times throughput.

研究分野：プロテオミクス

キーワード：プロテオミクス LC-MS

## 1. 研究開始当初の背景

ヒト細胞試料中には、20,000-25,000 遺伝子由来のタンパク質が、濃度差  $10^6$  倍の範囲で存在している。遺伝子のような PCR 増幅もできないことから、これを一斉に解析することには大きな困難が伴う。近年の質量分析 (MS) の急速な発展により、タンパク質の網羅的解析法 (プロテオミクス) は飛躍的に進歩し、2,000-3,000 種のタンパク質程度であれば、数時間の LC-MS 測定で同定・定量することが可能となった。しかし、依然として存在量の少ないタンパク質群を検出することには困難が伴い、遺伝子解析ではその発現が確認されていてもタンパク質としていまだに発現が確認されていないもの (いわゆる missing protein) が全遺伝子の 30% 程度存在している。HUPO (ヒトプロテオーム機構) を中心に国際プロジェクト「Human Proteome Project」が 2011 年に発足し、さまざまな試料と測定法を用いてこの missing protein を含むすべてのタンパク質 (プロテオーム) を明らかにする計画が立てられているが、いまだに一つの試料中に含まれる全発現タンパク質を測定する方法は確立されていない [Legrain et al., Mol Cell Proteomics, 2011]。

代表者らは、今までに、メートル長モノリスシリカカラムと緩勾配溶出による逆相 LC を用い、世界で初めて大腸菌の発現プロテオーム一斉解析を成功させた [Iwasaki et al., Anal Chem 2010.]。この手法をヒト iPS 細胞プロテオームにも適用したが、9,500 タンパク質を同定するにとどまった [Yamana et al., J. Proteome Res., 2013]。Nagaraj らはイオン交換クロマトによる前分画と 3 種類の消化酵素産物を組み合わせ、12 日間測定を行ったが 9,207 遺伝子産物同定にとどまり、RNA-seq 法による同定遺伝子 11,936 のうち 3000 あまりの遺伝子についてはタンパク質としての発現を確認できなかった [Nagaraj et al., Mol Syst. Biol., 2011]。イオン交換分画法では分離能が低いと十分に複雑性を軽減できない。また、分画法を用いる限り、測定時間の増加は避けられない。そこで逆相 LC と分離選択性の直交性が高く、しかも分離効率のよい方法を用いて試料を分画し、ある一画分のみを間引いてメートル長カラム LC-MS で測定を行うことにより、ヒトプロテオーム一斉解析を実現することを着想するにいたった。

## 2. 研究の目的

ヒト試料中に発現しているタンパク質すべての同定・定量を可能にする高深度ヒトプロテオーム解析システムの開発を提案する。最新鋭の質量分析システムと多次元分離を用いた試料前分画法を用いても現在のとこ

ろ世界中で誰も成功していない「ヒトプロテオーム完全解析」という課題に対して、超高分離能 LC-MS と超高分離能分画 間引き法を組み合わせた高深度プロテオーム解析システムを開発することにより挑戦する。真のオミクス規模でのヒトプロテオーム解析を可能にすることにより、遺伝子発現解析からタンパク発現量を推測するのではなく、タンパク発現量を直接解析することによって細胞機能制御の分子メカニズムを正確・精密に解き明かし、我が国のライフサイエンス研究を強力に加速させるための基盤技術を提供する。

## 3. 研究の方法

プロテオミクスで現在最も汎用されている測定法は LC-MS である。近年の MS の高性能化は著しいが、現在でもプロテオミクス試料に対する最大の弱点はスキャンスピードの遅さとイオン化抑制である。すなわち、試料成分数が多すぎるため共溶出することになり、スキャンが追い付かず、測定対象から漏れてしまう。さらに、濃度の低い試料成分が高濃度成分と共存すると、イオン化抑制現象が起き、低発現タンパクの検出が極端に困難になる。この解決法としては、LC-MS の前に試料を前分画する方法が一般的で、LC-MS で用いられる逆相モードと分離選択性が直交するイオン交換モードが用いられる。しかし分離効率が低く、一つの成分が多数の画分にまたがってしまうため、その効果には限界がある。そこで、直交性は低いが高分離効率の高い逆相モード (塩基性条件) が提案され、イオン交換分画よりも高い効果が認められているが、劇的な改善にはつながっていない。すなわち、逆相モードと直交した分離選択性を持ち、逆相モードに匹敵する超高分離能をもつ分離系を開発すればプロテオミクスにとって理想的な前分画法となる。この課題に対し、2つの方法を検討する。一つは、LC-MS だけではなく分画法にもメートル長を有するシリカモノリスキャピラリーを用いることである。これにより、従来法とはけた違いの高分離効率が期待できる。二つ目は、間引き分画である。理想的な分画であっても画分数が増えれば増えるほど測定時間が増大する。ショットガンプロテオミクス法では 1 つのタンパク質から数十のペプチドが生成されるため、理想的分画法ではすべての画分にすべてのタンパクに由来するペプチドが等しく分布することになり、ある画分のみを LC-MS 測定すれば、すべてのタンパクに対する発現情報が得られることとなる。このような画分の間引きを行うことにより、タンパク成分数は変化させず、試料の複雑性のみを軽減することにより測定時間の劇的な短縮を

はかる。直交高分離能分画法と合わせて間引き分画法を開発し、メートル長 C18 シリカモノリスカラムと緩勾配溶出を用いた高深度 LC-MS と組み合わせる。

#### 4. 研究成果

モノリス型シリカカラムのシラノール基に対し、3-(Trimethoxysilyl)propyl methacrylate( $\gamma$ -MAPS)を脱水縮合により固定化させた後、[2-(Methacryloyloxy)ethyl]dimethyl-(3-sulfopropyl)ammonium hydroxide(MEDSA)をラジカル重合させ、両性イオン型官能基を表面修飾した HILIC モノリス型シリカカラムを作製した。作製した 4 m の HILIC モノリス型シリカカラムを用いて 12 時間のグラジエントで測定を行った際の 3 種の合成ペプチドのピーク幅は平均で 2.14 分であった。グラジエント分離のカラム性能評価指標となるピークキャパシティは 336 と算出され、高分離な 2 m の RPLC モノリス型シリカカラムと同等の性能(ピークキャパシティ:317)を發揮した(図 1)。

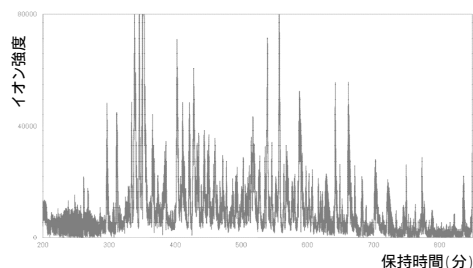


図 1 両性イオン型 HILIC モノリスカラムを用いた HeLa 消化ペプチドのベースピーククロマトグラム

また、作製した HILIC カラムと従来の RPLC カラムにおけるペプチド保持時間の相関係数は -0.481 であり相関性は低かった(図 2)。したがって、本カラムは高い分離性能と RPLC に対する高い直交性を併せ持っており、

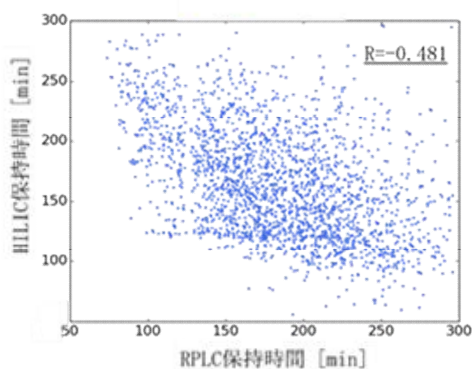


図 2 RPLC と HILIC の分離選択性

前分画のカラムとして適していることがわかった。

作製した 4 m の HILIC モノリス型シリカカラムにより前分画を行った試料について、2 m の C18 モノリス型シリカカラムを用いて 4 時間のグラジエントで LC-MS/MS 測定を行ったところ、5080 タンパク質(17694 ペプチド)が同定された。一方、前分画を行わずに同条件で測定を行った場合の同定数は 3484 タンパク質(21598 ペプチド)にとどまった(図 3)。

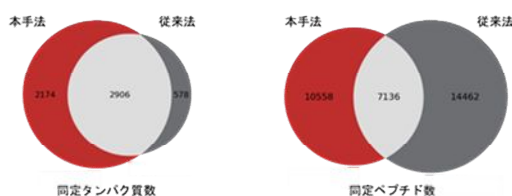


図 3 間引き法および従来法におけるペプチドおよびタンパク質同定数の比較

つまり、従来法(前分画を行わない方法)と比較して本手法の同定ペプチド数は 18.1% 減少していたにも関わらず、同定タンパク質数は 45.8% 増加していた。グラジエント分析 1 時間あたりのタンパク質同定効率は 871 から 1270 へ向上した。

また、手法間で同定されたペプチドのオーバーラップは 22.2% と非常に低く(図 3 右)かつ本手法で同定されたペプチドの検出ダイナミックレンジが拡大していたことから(図 4)、イオン化抑制の軽減により従来法では検出不可能であったペプチドが検出できるようになった可能性が示唆された。

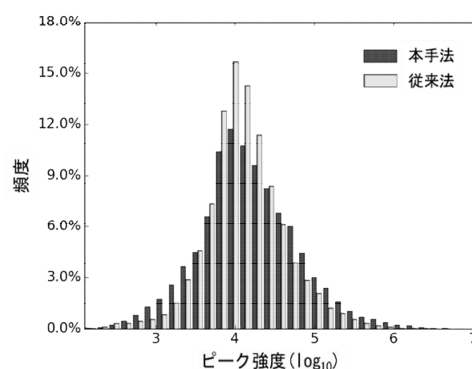


図 4 同定ペプチドのピーク強度分布

さらに、Nagaraj らの先行研究[Nagaraj et al., Mol Syst. Biol., 2011] で報告されている HeLa 細胞株のタンパク質発現量を参考とし、本手法と従来法で同定されたタンパク質について比較を行った(図 5)。高発現量側(100 fmol 以上)では手法間でほとんど差が見られないのに対し、低発現量側(100 fmol 以下)では顕著な差が見られ、本手法の優位性が示さ

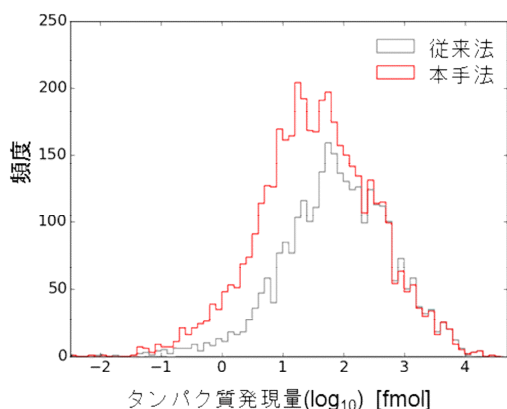


図5 同定タンパク質発現量の分布

れた。1 タンパク質あたりのペプチドを万遍なく間引く本手法によって、高発現量タンパク質の同定数を損なうことなく中・低発現量タンパク質の同定数を向上させることに成功した。

以上の結果から、本手法を用いることにより、従来法よりも効率的かつ高深度なプロテオーム解析が可能となることがわかった。

本二次元ナノ LC システムを用い、50  $\mu$ g の HeLa 細胞トリプシン消化ペプチドを出発試料として、単一試料から 6 つの画分を取得した後、それぞれの画分について長さ 2 m の RPLC モノリス型シリカカラムを用いて 8 時間のグラジエント条件下で LC-MS/MS を行った。本手法のプロテオーム解析の深度を評価するにあたって、Nagaraj らの大規模解析との比較を行った。なお、解析手法を統一するため、本手法と Nagaraj らのデータで同定されたペプチドをリストアップし、BLAST によって SwissProt データベースに対する配列再検索を実施した。その結果、同定されたタンパク質数は Nagaraj らと比較してやや少なかったものの、本手法では出発試料量を 200 分の 1 に、分析時間を 6 分の 1 に抑えることができた (表 1)。

表 1 HeLa 細胞大規模解析における本手法と従来法の比較

	本手法	Nagaraj <i>et al.</i>
試料量	50 $\mu$ g	10 mg
分画(タンパク質レベル)	なし	ゲル濾過(4分画)
使用消化酵素	Trypsin	Trypsin, Lys-C, Glu-C
分画(ペプチドレベル)	HILIC(6分画)	SAX(6分画)
グラジエント時間	2日(8時間 $\times$ 6)	12日(4時間 $\times$ 4 $\times$ 3 $\times$ 6)
タンパク質同定数	9,567 (9,255 genes)	10,160 (9,843 genes)
ペプチド同定数	85,014	117,330

この分析時間の短縮には、強アニオン交換クロマトグラフィー(SAX)と比較して、

RPLC に対しペプチドの分離選択性において高い直交性を持つ HILIC を一次元目に用いたこと、また、高分離能モノリス型シリカカラムを用いたことが関与しているものと考えられる。本研究と先行研究で同定されたタンパク質について比較したところ(図 6)、それぞれでユニークに同定されたタンパク質が存在していたことから、本手法においても十分高深度なプロテオーム解析を行うことができたと考えられる。

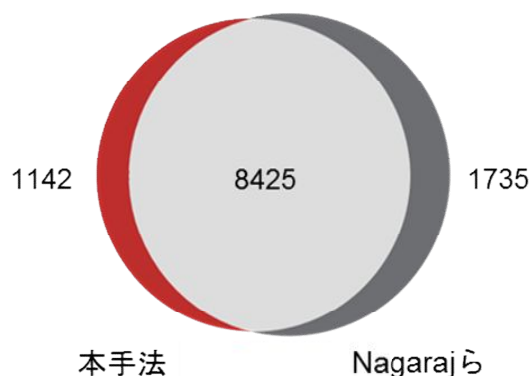


図 6 HeLa 細胞大規模解析における本手法と従来法による同定タンパク質の比較

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1 件)

Miao-Hsia Lin, Naoyuki Sugiyama, Yasushi Ishihama, Science Signaling, 査読有, 8, 2015, rs10.

Doi: 10.1126/scisignal.aaa3117

〔学会発表〕(計 7 件)

Kanta Horie, Takeo Kamakura, Suguru Ichihara, Masaki Wakabayashi, Nobuo Tanaka, Yasushi Ishihama, High Resolution HILIC for Proteomic LC-MS, 62nd ASMS Conference on Mass Spectrometry and Allied Topics, Baltimore, MD, USA, 2014/06/15-19.

鎌倉健雄、市原駿、堀江勘太、若林真樹、杉山直幸、石濱泰、メートル長モノリス型シリカカラムを用いた LC-MS/MS システムによるヒトプロテオーム解析、日本プロテオーム学会 2014 年会、つくば国際会議場(茨城県つくば市) 2014/07/17-18

市原駿、鎌倉健雄、堀江勘太、若林真樹、杉山直幸、石濱泰、ナローボアメー

トル長モノリス型シリカカラムを用いたプロテオミクス LC-MS 高感度システムの構築、第 27 回バイオメディカル分析科学シンポジウム、帝京大学板橋キャンパス（東京都板橋区）2014/08/20-21

木村 迪子，橋本 貴行，若林 真樹、杉山 直幸、石濱 泰、タンパク質の N 末端大規模解析に向けた新規 N 末端ペプチド濃縮法の開発、日本分析化学会第 63 年会、広島大学東広島キャンパス（広島県東広島市）2014/09/17-19

鎌倉健雄、市原駿、堀江勘太、若林真樹、杉山直幸、石濱泰、メートル長 C18 モノリス型シリカカラムを用いた高分離能二次元ナノ LCMS プロテオミクス、第 25 回クロマトグラフィー科学会議、京都大学桂キャンパス（京都市）2014/12/07-12

鎌倉健雄、市原駿、若林真樹、杉山直幸、石濱泰、ヒトショットガンプロテオーム解析のための高分離能二次元ナノ LC、第 26 回クロマトグラフィー科学会議、九州大学医系キャンパス（福岡市）2015/11/11-13

Yasushi Ishihama, Slicing the Human Proteome: Chromatographic Challenges Facing Complete Human Proteome & Phosphoproteome Analysis, Keystone Symposia meeting on “The Human Proteome”, Stockholm, Sweden, 2015/04/24-29

## 6. 研究組織

### (1) 研究代表者

石濱 泰 (ISHIHAMA Yasushi)  
京都大学・大学院薬学研究科・教授  
研究者番号：30439244

### (3) 連携研究者

杉山 直幸 (SUGIYAMA Naoyuki)  
京都大学・大学院薬学研究科・准教授  
研究者番号：50545704

若林 真樹 (WAKABAYASHI Masaki)  
京都大学・大学院薬学研究科・助教  
研究者番号：70552024