

ゲノムからプロテオームへ -プロテオーム解析の現状

中山 洋

(理化学研究所生体分子解析室)

ポストゲノム時代に向けて注目されるプロテオーム解析について概説した。まず、プロテオームの定義および重要性について述べた。次に、実際のプロテオーム解析の例として、(1)タンパク質間相互作用解析(2)翻訳後修飾解析(3)プロファイリングについて紹介した。

0. はじめに

生物学の現在を一言でいうならば「ゲノム生物学の時代」であろう。1996年に真核生物としては初めて出芽酵母(*Saccharomyces cerevisiae*)ゲノム上の全塩基配列が決定され、1998年末には多細胞生物としては初めて線虫(*Caenorhabditis elegans*)ゲノムの全塩基配列が決定された。さらに現在の予定ではヒトさえも今春中に全ゲノムの約90%の暫定的な塩基配列が決定出来るとされている。このようなポストゲノムを脱んだ状況下での生物学の大きな目標の一つは、ある特定の遺伝子が時間、空間的にどのように発現して細胞内にタンパク質として出現し、修飾されて機能を発現したのち、最終的に分解、代謝されていくのかを詳細に解析することであろう。しかも、これはゲノムシーケンシングの結果を最大限に利用した、大規模で網羅的なものとなるだろう。

本稿では、このようなタンパク質分子レベルでのポストゲノムシーケンシング解析であるプロテオーム解析の現状を概観し、その将来を展望する。

1. プロテオームとは?

1-1. プロテオームの定義

「ある生物のゲノムの産物の総和」すなわち、個体あるいは一つの細胞で発現しているタンパク質全ての動態を要素とする集合をプロテオーム(Proteome)と呼ぶ¹⁾。これは、ゲノムが「ある生物の遺伝情報の総和」であることに対応している。プロテオームという語自体も、ゲノムが gene と chromosome を組み合わせた造語であることに対応して、protein と genome を組み合わせた新造語である。このプロテオームを解析する研究をプロテオーム解析あるいはプロテオミクス(proteomics)と呼んでいる。

ポストゲノム解析は静的なゲノムデータを利用して、分子レベル(mRNA, タンパク質)、表現型レベルといった様々な観点からダイナミックな生物像を描き出すことを一つの目的としている。タンパク質分子レベルのポストゲノム解析であるプロテオーム解析では、従来のように一つのタンパク質に着目して解析するのではなく、プロテオーム全体の中であるタンパク質あるいはタンパク質群が共同的に生物活性を発現する様子を網羅的解析によって明らかにしていこうという立場を取る。図1にポ

ストゲノム解析の主要な課題と、プロテオーム解析の位置付けを示した。

現状では、二次元電気泳動法などをもちいたタンパク質としての発現分子種とその発現レベルの網羅的解析をプロテオーム解析と呼ぶことが多いようである(狭義のプロテオーム解析)。これは、それ以外のアプローチがまだ始められたばかりであることも原因である。私達は定義からポストゲノム解析のうちタンパク質を解析するもの全てをプロテオーム解析と呼ぶべきと考えている。以下では狭義のプロテオーム解析はプロファイリングと呼ぶ。

1-2. プロテオーム解析の重要性

ゲノムシーケンシングの結果は私達がいかにタンパク質の機能について知らないかを明らかにしてきた。たとえば、1998年にシーケンシングが完了し、機能研究へとシフトしている線虫の場合、実験的に機能が決定された遺伝子・タンパク質は全ゲノムの10%以下にすぎず、相同性から機能が推定されるものを加えても60%程度である³⁾。現在までに完了した多くの生物のゲノムでも同様な状況であり、今後、ヒトなどでも同様な結果となることは予想に難くない。したがって、来るべきポストゲノムシーケンシング時代の大きな課題の一つはこれら機能未知のタンパク質の実験的な機能解析であろう。

この機能解析のため、ポストゲノム解析としてトランスジェニック動物のような個体レベルの解析、mRNAレベルでのプロファイリングと並んで以下のようなプロテオーム解析が進められている。(1)タンパク質とタンパク質あるいは他の生体物質との相互作用解析。ほとんどのタンパク質の機能は、他の生体分子との相互作用による。ホルモン-レセプター、酵素-基質、抗原-抗体などはその典型的な例である。また、細胞内の情報伝達系が、リン酸化/脱リン酸化などの翻訳後修飾によって調節される分子間相互作用により実現されていることも明らかになってきた。したがって、分子間相互作用情報はタンパク質の機能を知る上で有用である。また、これらのデータを集積してタンパク質のリンケージマップをつくる試みも始まっている。これは将来的に主流になっていくであろう「複雑系としての生物」像を考える上で無くてはならない基礎データとなりうる。(2)翻訳後修飾の解析。多くのタンパク質の機能がリン酸化やプロセッシングなどの翻訳後修飾により発現、調節されることが知られている。

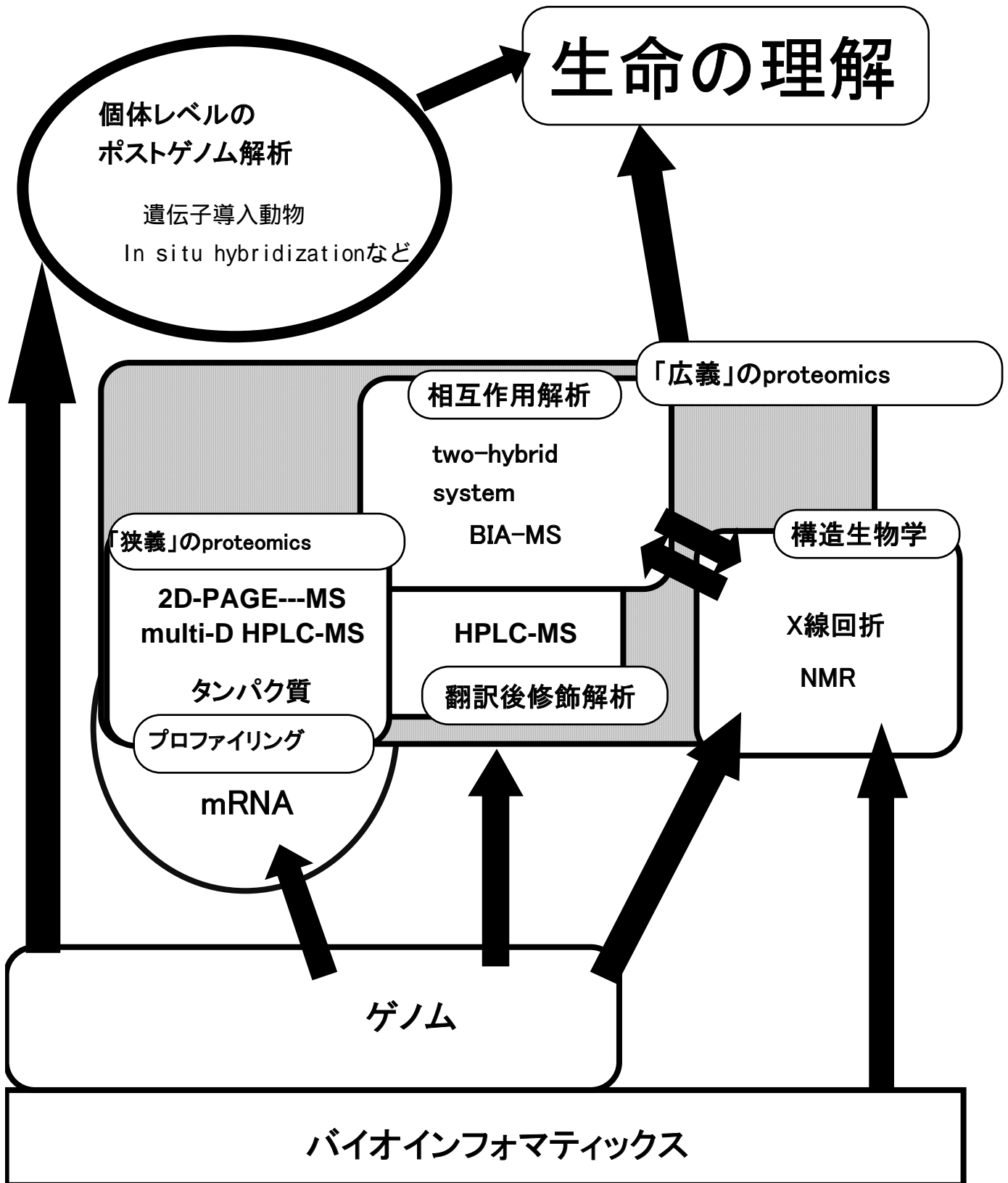


図1. プロテオーム解析の位置付け

この修飾の有無は遺伝子配列からは決定できないので、実際に生体内で働くタンパク質を解析することが必要である。(3)翻訳レベルでのプロファイリング。タンパク質が相互作用を通じて機能を発現するためには、分子同士の時間的、空間的発現プロファイルが一致している必要がある。mRNAとタンパク質の発現レベルは一般に異なることが知られているので、タンパク質自体の発現レベルを直接解析することが重要である。また、病態ではしばしばタンパク質の異常な発現や代謝が認められる。これらの情報は疾患の原因解明や診断のためにも重要である。

このようなプロテオーム解析の結果は膨大なものとなるので、それをデータベース化することが有効である。現在、ゲノムデータベースと相互リンクしたプロファイルデータベースやタンパク質機能データベースが試験的に構築されている。将来的には、これをもちいることで生命科学のみならず、医薬品の開発にも有用であると考えられている。例えば、今後さらに重要と考えられている生活習慣病などの疾患の大部分は複数の遺伝子が関与することが知られている。これらに対する医薬品の開発は、遺伝子産物であるタンパク質発現プロファイルやタンパク質同士のリンケージがデータベースとして利用できることで、大きく効率化されると考えられる。

2. プロテオーム解析の現状

ゲノム解析とプロテオーム解析の最大の違いは、ゲノムは膨大ではあるけれどもも有限の物質(モノ)として明確に規定出来るのに対してプロテオームは時間的(発生、分化、老化)、空間的(超分子レベル、オルガネラレベル、細胞レベル、個体レベル)に無限と言ってもよい広がりを持つなかで機能という漠然としたコトを対象とすることである。この性格から、プロテオームの全てを解析することは現実的に不可能であるし、またあまり意味も無い。したがって、ある程度独立したプロテオームの部分集合を研究対象に選ぶことが現実的である。

この部分集合を選ぶための指標としては、主として時間的あるいは空間的に解析対象を限定することが考えられる。発生・分化過程、あるいは薬物などの化学的刺激や熱ショック、電磁波などの物理的刺激に対する細胞応答などでは、時間的に限定した部分集合間での変化を解析することが有効である。一方、ある比較的独立した機能単位、例えば細胞内小器官、転写、翻訳装置などの超分子複合体を対象とする場合には、それらの機能単位を単離して、その構成成分全てを明らかにし、その間の相互作用を調べ、さらにはその機能単位を再構成することが有効である。これは、現在の構造生物学の延長線上にあるアプローチである。ウィルス粒子やリボソームなどで成功している。

このような指標に基づいた部分集合をもちいて、1-2.で述べたアプローチそれぞれについて解析法の構築と評価が進められている。以下ではこれらの現状について紹介する。

2-1. 相互作用解析

通常、相互作用情報には四つの階層に分類出来る。一番基本的なのは、そのタンパク質が問題とする時間、空間にどの程度の濃度で発現しているか?である。これは直接相互作用とは関係無いけれども、相互作用の生理的な意義を考える上で重要である。後述のプロファイリングにより得られる。次に、相互作用がある/ないの情報、すなわち、**特異性**である。免疫沈降法、融合タンパク質をもちいたアフィニティー法など多くの生化学的方法およびtwo-hybrid systemなど分子生物学的方法はこのレベルの情報を与える。さらに、定量的な情報として、平衡状態における結合の強さの指標である**平衡定数**および平衡に到達する速さの情報を与える**速度定数**がある。これらの定量的情報、特に速度論的な情報はその相互作用の生物学的意味を考える上で重要である。定量的情報は超遠心法、カロリメトリー法、アフィニティーキャピラリー電気泳動法および表面プラズモン共鳴検出器を備えたバイオセンサをもちいる方法(BIA)といった機器分析法により得ることが出来る。また、構造生物学から得られる原子レベルでの立体構造はこのような定量的な相互作用情報を化学的に裏付けるために最適なものである。

ゲノムスケールでの分析を考えた場合には、ライブラリーをスクリーニングでき、検出した相互作用パートナーの同定が簡単で、高感度なtwo-hybrid system法および自動化装置が市販されており、定量的な相互作用情報を得ることが出来、高感度であるBIAが優れている。

すでにウィルスや酵母などのモデル生物では、two-hybrid systemによるゲノムスケールでのタンパク質リンケージマップ作成が試みられている。ところが、この方法では、偽陽性が多く、スクリーニングの結果は他の方法で確認することが重要である。このために従来免疫沈降法などの生化学的方法がもちいられる。しかし、これらの方法では安定な遅い相互作用の特異性は確認できても、速い反応は見落とす可能性が高い。また逆に、two-hybrid systemは多対多あるいは翻訳後修飾により制御されている相互作用の検出が苦手(偽陰性)である。実際、酵母ゲノムをtwo-hybrid systemでスクリーニングした結果、約700組のタンパク質?タンパク質ペアしか検出出来なかった³⁾。このため上述した機能未知タンパク質の解析のためには、より一般的な相互作用検出・解析法が必要であることが明らかになってきた。

一方、BIAは、速度論的な情報を含めた定量的な相互作用情報を得ることが出来るため、スクリーニング的な目的だけでなく、two-hybrid systemの結果の確認や相互作用の詳細な解析に向いている。ところが、もちいる試料が微量であるためもあり、多くの場合にBIAで検出したリガンドを同定・一次構造解析することは困難であった。シグナル伝達系などで重要な翻訳後修飾による相互作用の調節を解析する場合には、相互作用した分子が実際に修飾を受けているかどうかの解析が必須である。このようなキャラクタライズを目的として、BIAとMSを統合したシステムを構築しようという研究が盛んに進められている。

NelsonらはBIAのセンサーチップにイオン化のためのマトリックスを直接加えてMALDI/TOF MSでタンパク質の分子量を測定することに成功した^{4,5)}。この方法はMALDI/TOF MSをもちいているため、高感度ではある。しかし、タンパク質自体の分子量を測定するのでそのタンパク質の一次構造情報は得られず、同定も出来ない。BIAで相互作用解析したタンパク質の一次構造情報を得るために、BIAとMS/MS法を組み合わせる試みもなされてきた。このために、タンパク質をセンサーチップから回収した後、SDS電気泳動法で再度分離精製したタンパク質をゲル中で酵素消化してMS法で分析してきた。しかし、この方法では、試料の損失が多く、現実的にBIAでもちいる微量の試料には適用するのが困難なため、何回分もの分析で試料を集めて、ようやく一次構造解析が可能である。このように、手順が煩雑で時間がかかるため自動分析であるBIAの利点を生かせない。

そこで、著者らは、BIAとHPLC-MS/MSを統合したオンラインシステムを提案・構築した。(Natsume, T. et al. 投稿中)これは、(1)チップ上でタンパク質をプロセスし、(2)オンラインで回収した消化ペプチドを(3)ESIカラム法をもちいた高感度HPLC-MS/MSで一次構造を解析するシステムである。(図2)以下で簡単に説明する。(1)BIAで相互作用解析したタンパク質はそのままではMS/MS法で一次構造解析出来ない。そこで、タンパク質分解酵素をBIAのセンサーチップに導くことで、センサーチップ上でタンパク質をプロセスする。(2)BIAで通常もちいるタンパク質はfmol量であり、この量のタンパク質・ペプチドは、例えば、分取してポリプロピレンチューブに一晩保存しておくだけで器壁への吸着で失われてしまう。したがって、試料はオンラインで迅速に分析計に導入したい。私たちのシステムでは微小な逆相プレカラム(60nL)にペプチドをオンラインで吸着させ、洗浄した後HPLC-MS/MS法で分析する。(3)上述のように調製したペプチドはスプレーチップ自体に充填剤を詰めたESIカラムをもちいたオンラインHPLC-MSで分子量を測定し、MS/MSにより一次構造解析する。一般にオンラインシステムでは各成分を分離してからイオン化するので、相対的にイオン化しにくい成分を検出しよう。特に、この方法は分離カラム以降の死体積が非常に小さいため、カラムの分離能を最大限に生かせる。(Nakayama, H. et al. 投稿準備中)このため、タンパク質複合体から生じる複雑なペプチド混合物の分析や翻訳後修飾の検索に適している。

このシステムでは、大腸菌による発現系をもちいたモデル系の場合、His-tag/Ni-NTAのような安定な相互作用だけでなく、イノシトール三リン酸とそのレセプター(解離速度定数 $1s^{-1}$)のような速い相互作用をもセンサー上で検出した後に、そのfmol量のタンパク質をMS/MSにより一次構造解析し同定出来る。

これはBIAで相互作用したタンパク質を直接一次構造解析した最初の例であり、まだプロテオーム解析用にルーチンで使うまでには多くのハードルがある。しかし上述したBIAの利点を生かした定量的速度論データとMSに

よる一次構造情報が同時に得られることで、ゲノムスケールでのタンパク質機能解析を加速しよう。

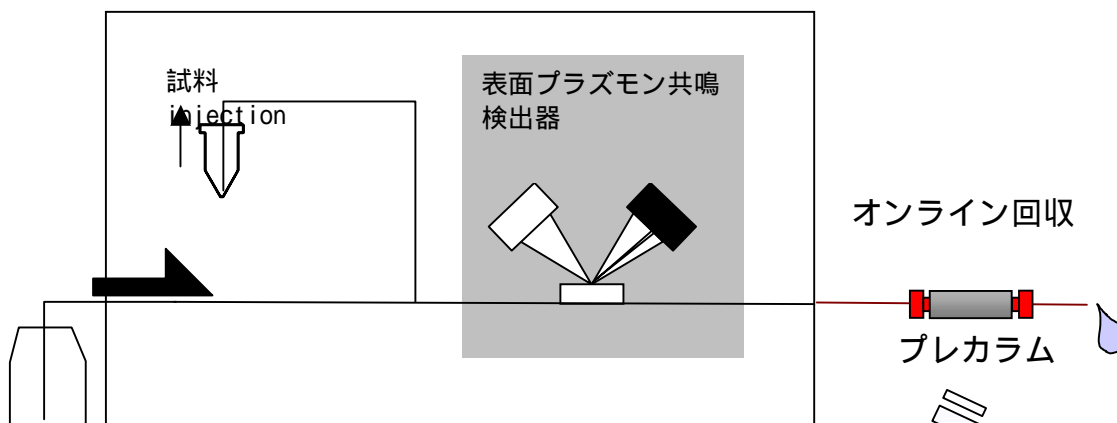
2-2. 翻訳後修飾解析

タンパク質は、そのアミノ酸配列自体の情報や、タンパク質分解酵素による限定的なプロセッシング、糖鎖の付加などによって選別され、細胞内外での局在が決定されている。その後、その部位で機能を発現するために、多くのタンパク質が修飾によりその機能を発現する。例えば、プロセッシングによる細胞内局在部位の移動および活性発現⁶⁾やリン酸化/脱リン酸化やその他の化合物の可逆的な付加/脱離の場合には、被修飾残基の性質が変化することで時には大きな立体構造変化とそれに伴う相互作用の変化が生じ、タンパク質の機能を発現・調節することが知られている。これらの修飾部位はコンセンサス配列と呼ばれるアミノ酸配列に基づいて規定出来るが、そのコンセンサス配列がある時点のある細胞、臓器で実際に修飾されているかどうかは、現状では配列情報だけでは決定出来ない。これら生体内での翻訳後修飾分子種は非常に速いものでは秒単位の寿命であり、多くのタンパク質では、その一部のみが修飾される。また、リン酸化などの場合には一つのタンパク質中の複数の部位が修飾されることがあり、その修飾部位および状態により機能が異なる例が知られている。したがって、修飾残基およびその部位を特定できる高感度な分析法が必要である。

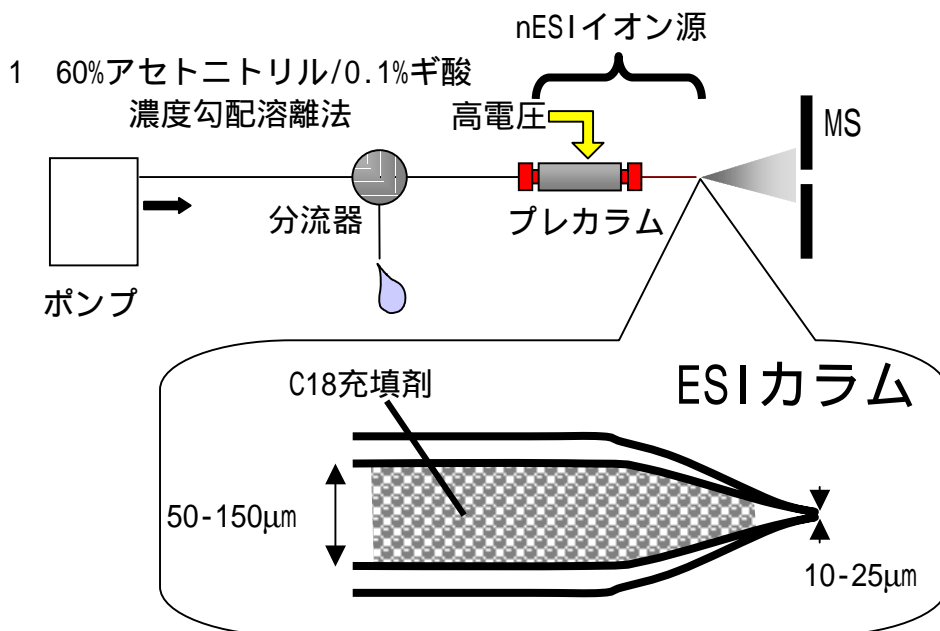
従来、翻訳後修飾の解析法としてはRIラベルした基質分子を細胞や組織切片などに取り込ませ、生合成過程でタンパク質分子をラベルしてから分析する方法や、分離したタンパク質の修飾部分を何らかの方法で検出した後に単離し、化学修飾して解析する方法などが用いられてきた。これらの方法は、修飾の種類により異なる、いわば個別的な解析法であり、修飾の種類が予測できないときには適用が難しい。またRI法は高感度ではあるが動物個体などに対してのin vivoの解析が難しいこと、化学修飾法は手順が煩雑であり、必ずしも高感度ではないなどの理由から、より一般的な翻訳後修飾の検出・解析法の発展が望まれていた。

MS法は分子量という最も基本的なパラメータを直接測定するため、リン酸化、脂肪酸付加、糖鎖付加、プロセッシングなど翻訳後修飾の種類を問わない一般的な翻訳後修飾の検出、解析法として有望である。タンパク質をペプチドに分解して質量を高精度(5mu程度)に決定できれば、修飾残基の組成式を算出し、新規修飾の化学構造をも推定しよう。MS/MS法をもちいれば修飾の種類だけでなく、その位置まで決定できる。さらにHPLC-MS法により酵素消化ペプチドを分離して詳細に分析すれば、タンパク質全体の翻訳後修飾の状態を、高感度に明らかに出来る。リン酸化など一つのタンパク質内で多くの部位への修飾の可能性がある場合にはこれは重要である。

BIA -molecular interaction



HPLC-MS -primary structure



上述のようにMS法は多くの優れた点を持つため、ポストゲノム時代の翻訳後修飾の解析法として最適と考える。

ここでは、鉄型ニトリルヒドラーゼ⁷⁾の活性に重要な役割を果たしていると考えられる翻訳後修飾の分析例を示す。この酵素の活性中心は、トリプシン消化した場合一つのペプチド上にあるため、それをMS/MS分析した。(図3)その結果、遺伝子からの予想アミノ酸がシステインである α 鎖の112番は、32[u]システインより大きかった。これは酸素2原子(31.990)あるいはイオウ1原子(31.972)の付加が考えられる。この測定にもちいた四重極/飛行時間ハイブリッド型質量分析計は、外部標準による校正での質量精度は50-100ppm程度なので、この段階では上記のどちらかを判定できない。次に、修飾残基を含まない部分の配列情報をあたえるシグナル(y_8)を内部標準としてもちい、再度質量を校正した。これによりm/z750-1000程度では、修飾を含まないyイオン全てを3[mu]以内の精度で決定できた。この精度があれば十分に酸素付加かイオウ付加かを判定できるので、修飾部分の残基質量を計算すると135.000[u]であった。この結果から、修飾残基は酸素付加すなわちシステインスルフィン酸と同定できる。これは、試料をジチオスレイトールで還元しているため化学的にもイオウの可能性はほぼないこと、エドマン法をもちいたアミノ酸配列分析、アミノ酸組成分析、各種のMS法をもちいた分子量測定、そしてX線結晶構造解析の結果を裏付けるものであった。

このように翻訳後修飾の解析にMS法は非常に有用であるが、現状ではゲノムスケールでの翻訳後修飾の検索の試みはほとんどなされていない。これは翻訳後修飾の解析が極めて煩雑であることと、いまだ続々と新しい翻訳後修飾が検出されていること、そしてこれらの修飾と機能の関係を証明することは極めて困難であることが原因であると考えられる。

2-3. プロファイリング

ゲノムシークエンスの結果、遺伝子産物の種類は、大腸菌など細菌で約4000種類、単細胞真核細胞生物である酵母の場合には約6000、多細胞生物である線虫では約19000であると見積られている。この遺伝子産物であるタンパク質は上述のように翻訳後に各種の修飾により成熟して機能を発現することが知られている。これらを区別すると、さらに多くの種類のタンパク質分子種がプロファイリングの対象となる。しかし、これらが全て一時期の一つの細胞で発現しているのではなく、ある一時期に、ある一つの細胞では数千のタンパク質が発現していると考えられている。したがって、プロテオームの部分集合をうまく選べば数百?数千程度のタンパク質を分離すればよい。

これらのタンパク質を網羅的に分析するためには、種々の分離分析法を多次的に複合化した複合化分離分析法が有効である。たとえば、目的のために10000種類のタンパク質を分離しなければならないとしよう。一つの方法のみをもちいた場合には、高速液体クロマトグラフィーや電気泳動法でも現状では多く見積っても100

種類程度を分離出来るに過ぎない。また、これらの方法をいくら改良しても、せいぜい数倍程度の分解能向上が限界であろう。もちろん、これではプロファイリングには分解能が足りない。これに対して、独立したパラメータに基づく方法の組み合わせでは、それぞれは100種類にしか分離出来なくても100x100x100x...と効率的に分解能を上げられる。10000種類なら独立した2つのパラメータすなわち2次的分離で充分である。タンパク質は電荷と大きさという独立した物理化学的パラメータと活性または親和性という生物学的パラメータを持つのでこれらを組み合わせて複合化分離分析法をデザイン出来る。以下では現在最もよくもちいられている複合化分離分析法である2D PAGE(-MS)法によるプロファイリングを紹介する。

2D PAGEは1次元目にタンパク質の電荷で分離するIEF、2次元目にタンパク質の大きさ(=分子量)で分離するゲル電気泳動法を組み合わせた複合化分離法である。この方法ではタンパク質を生理的な条件で分離するか変性条件で分離するかによりいくつかの変法がある。この中でプロテオーム解析で最もよく用いられているのは、変性-変性のいわゆるO'Farrellの方法である⁸⁾。この方法のプロファイリングにおける利点は以下の点である:5000種類以上のタンパク質を一回に分離することが出来るとされ、現在でも最も高分解な方法である。ゲノムに対応する遺伝子産物を直接分析することが出来る。そして、その遺伝子産物を2次元平面上に一覧出来る(differential display法に適する)。

この方法では、基本的には、タンパク質は一本鎖のポリペプチド(=遺伝子産物)として分離される。実際、大腸菌の2次元電気泳動タンパク質スポットの数から見積った大腸菌タンパク質数は大腸菌ゲノムから推定したタンパク質総数とよく一致している。したがって、翻訳後修飾の少ない原核生物では2次元電気泳動マップ上にほとんど全ての発現遺伝子が検出できる。すなわち、2次元電気泳動マップ=プロテオームマップである。(プロテオームという概念が微生物学をバックグラウンドとした研究者から提案されたのも偶然ではない。)しかし、真核生物では、遺伝子数が多いことと、多くのタンパク質が翻訳後修飾されていることから、上述の時間・空間的なサブセットに注目して解析する必要がある。

Taokaらは齧歯類の小脳が形成される生後の三週間というサブセットを選び、生後日齢の異なる2Dゲルをdifferential displayして、小脳形成期に特異的に発現が亢進するタンパク質をMS法で同定した⁹⁾。同定されたタンパク質の多くは細胞内情報伝達系のタンパク質であり細胞分化や発生に関わることが知られている因子であった。一方、小脳が成熟するにしたがって発現が上昇し、成体で発現量が最大となるようなタンパク質の多くは糖代謝系酵素、細胞骨格系タンパク質など神経細胞の維持に関わる因子であった。したがって、このような時間軸を指標としたdifferential displayは発生、老化などの複雑な現象に関わるタンパク質を選択的に検出出来る有用な方法であると考えられる。

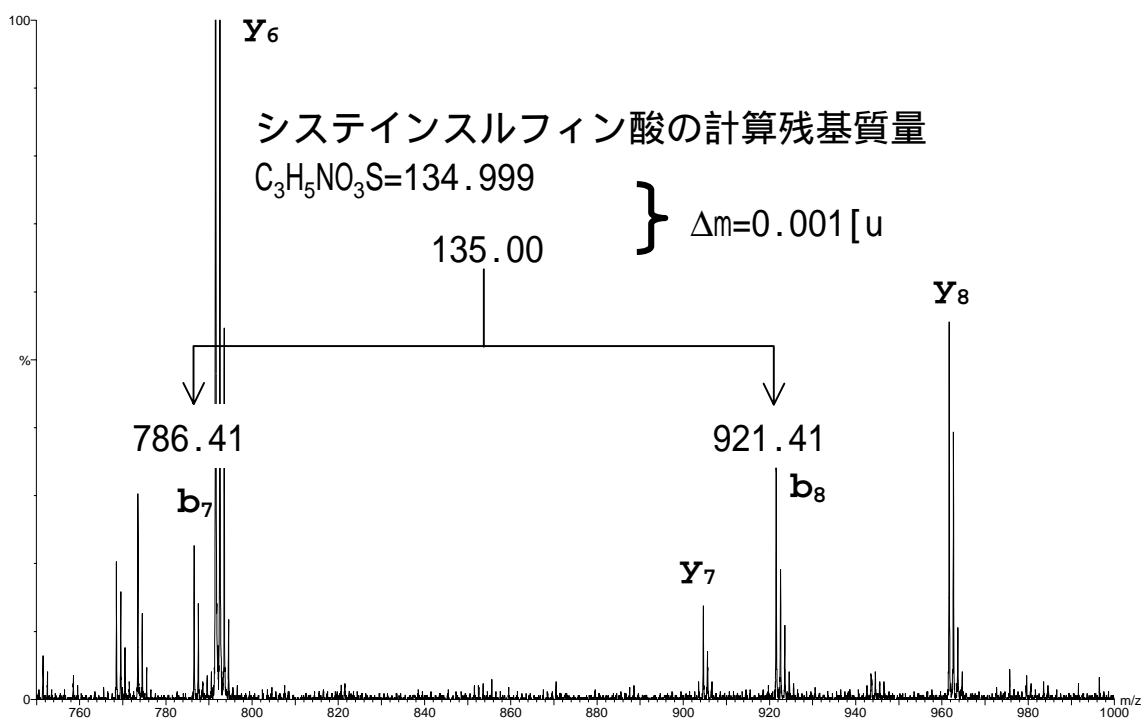


図3. MSによる翻訳後修飾の解析

ニトリルヒドラーゼ活性中心ペプチドのMS/MSスペクトル(部分)
還元カルボキシアミドメチル化後にトリプシン消化した活性中心ペプチドの2価イオンを、オフラインLC-nanoESI/Q/TOFでMS/MSした。b₇, 8イオンから鎖112番の実測残基質量は135.000であり、システイン(計算質量103.009)がシステインスルフィン酸(134.999)に酸化していることがわかる。y₈イオンを内部標準として質量を再校正した。

しかし、二次元電気泳動法には、(1)疎水性のタンパク質や高分子量のタンパク質は分析が困難であること¹⁰⁾、(2)操作が複雑なためデータの再現性が低いこと、(3)ゲル上のタンパク質の定量法がないという問題点がある。これらは多くの研究者の十年以上にわたる努力にも関わらず、あまり改善は見られていない。このため、他のプロファイリング法、特によりMS分析に重点を置いた方法が提案されている。

例えば、安定同位体ラベル法をもちいた定量的な differential display 法である。

Odaらは安定同位体ラベルした試料をもちいてタンパク質の発現量を定量的に比較する方法を提案した¹¹⁾。まず比較したい2つの条件(ミュータント/ワイルドタイプ、刺激/未刺激など)の片方のタンパク質のみを¹⁵Nを含む培地で培養する。次にこれらの異なる同位体組成を持つ試料を混合して、例えば二次元電気泳動法で分離する。そして、タンパク質スポットをゲル内消化してMSにより分析すれば、2つの条件下でのタンパク質の発現量比は¹⁴N由来と¹⁵N由来のシグナル強度比で示される。この方法は、二次元電気泳動法の大きな問題点の一つである定量性の低さを解決出来る。

一方、Aebersoldらは、タンパク質内のシステイン残基を通常の試薬あるいは安定同位体ラベルした試薬で修飾して、その両者を混合して分析する方法を提案した¹²⁾。この方法は、代謝ラベルが困難な動物組織などにも適用でき、より一般的である。また、修飾基をアフィニータグとしておくことで、修飾したタンパク質・ペプチドのみを簡便に分離・濃縮することができるため有用である。

これらの方法は、従来の二次元電気泳動法と組み合わせることでその欠点を補う方法としても使えるが、むしろ、細胞や組織の抽出液をそのままタンパク質分解酵素で分解し生じたペプチド混合物を直接多次元化したHPLC-MSシステムで分析することが有用と考える。このアプローチは、分析システムを自動化できる点、種類によって分子量や可溶性などの性質が全く異なるタンパク質を扱わないですむ点で優れている。

3. おわりに

プロテオーム解析はまだ方法模索の段階である。現在、プロテオームといえは二次元電気泳動で分けたタンパク質をMSで同定するものと単純に結び付けられているが、このようなプロファイリングだけから得られる情報は少ない。プロファイリングはプロテオーム解析のごく一部分でしかなく、それに続く機能解析がプロテオーム解析の中心的な課題となる。この機能解析は、上述してきたように、始まったばかりである。今後新しい方法の開発によりその流れは大きく変わりうる。そこで最後に、将来のプロテオーム解析システムに最も有用と思われるマイクロ流路技術について述べたい。

マイクロ流路技術は半導体の加工技術などを利用して、ガラスやシリコンチップ上に μm オーダーの溝または細管を作る技術である。その微小空間を利用して、極微量

の試料を、迅速、人手を介さずに分析しうるため、注目されている。ゲノム解析からのニーズと技術的な容易さから、現状ではDNAをマイクロ流路中で電気泳動分析することが中心であった。ところがタンパク質解析のための要素技術(試料濃縮や逆相HPLCのためのカラムや濃度勾配分離システム)は技術的な難しさから遅れていた。最近になってようやく、超低流速域(nL/min)での濃度勾配分離法¹³⁾やチップ上から直接ナノフローエレクトロスプレーするインターフェイスの開発¹⁴⁾などタンパク質分析のための基礎検討も進められている。また、これらを統合したシステムも現在盛んに研究されている。システム化することにより、環境?特にタンパク質の塊であるヒト?の影響を受けずに微量分析が可能であり、容易にゲノムスケールの大量処理に必須な自動分析装置を構築出来る。

参考文献

- 1) Wasinger, V.C., Cordwell, S.J., Cerpa-Poljak, A., Yan, J.X., Gooley, A.A., Wilkins, M.R., Duncan, M.W., Harris, R., Williams, K.L., Humphery-Smith, I.: *Electrophoresis* 16, 1090-1094 (1995)
- 2) WormPD(<http://www.proteome.com/databases/index.html>)
- 3) Uetz, P., et al.: *Nature* 403, 623 - 627 (2000)
- 4) Nelson, R.W., Krone, J.R., Jansson, O.: *Anal. Chem.*, 69, 4369-4374 (1997)
- 5) Nelson, R.W., Krone, J.R., Jansson, O.: *Anal. Chem.*, 69, 4363-4368
- 6) Okamoto, T., Nakayama, H., Seta, K., Isobe, T., Minamikawa, T.: *FEBS Letters*, 351, 31-34(1994)
- 7) Nakasako, M., Dohmae, N., Tsujimura, M., Takio, K., Odaka, M., Yohda, M., Kamiya, N., Endo, I.: *Nat. Struct. Biol.*, 5, 347-351 (1998)
- 8) O'Farrell, P.H.: *J. Biol. Chem.*, 250, 4007-4021 (1975)
- 9) Taoka, M., Wakamiya, A., Nakayama, H., Isobe, T.: *Electrophoresis in press*
- 10) Wilkins, M.R., Gasteiger, E., Sanchez, J.C., Bairoch, A., Hochstrasser, D.F.: *Electrophoresis*, 19, 1501-1505 (1998)
- 11) Oda, Y., Huang, K., Cross, F.R., Cowburn, D., Chait, B.T.: *Proc. Natl. Acad. Sci. USA*, 96, 6591-6596 (1999)
- 12) Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., Aebersold, R.: *Nat. Biotechnol.*, 17, 994-999 (1999)
- 13) Xue, Q., Dunayevskiy, Y.M., Foret, F., Karger, B.L.: *Rapid Commun. Mass Spectrom.*, 11, 1253-1256 (1997)
- 14) Figeys, D., Aebersold, R.: *Anal. Chem.*, 70, 3721-3727 (1998)

付録 MS によるタンパク質の同定

プロテオームの現状で述べた三つのアプローチのいずれでもタンパク質の同定が重要なステップである。タンパク質を同定するためには一次構造情報をもちいるのが最も確実である。MS をもちいた場合には大きく分けて二つの方法がある。

(1) 酵素消化ペプチドの分子量データをもちいる方法と
(2) MS/MS 法により得られたシークエンシオン情報をもちいる方法である。いずれの場合でも溶液中あるいはゲル中や PVDF に転写し膜上でタンパク質を消化して生じたペプチドを分析する。

(1) 分子量データ

タンパク質から調製したペプチド混合物を MS でペプチドマッピングすることでタンパク質の同定が可能である。タンパク質を酵素的な方法などで限定的に分解したペプチドの分子量のセットをもちいて核酸/タンパク質配列データベースを検索する^{A-1)}。この方法は、peptide mass fingerprint 法と呼ばれる。図 A-1 にこの方法の概念を示した。簡便、迅速に一次構造に基づいた結果が得られるために、二次元電気泳動ゲル上のスポットを網羅的に同定するなど大規模な分析には、自動化された MALDI/TOF と組み合わせが有用である。ただし、完全に精製したタンパク質のみにしか適用できない点、また、現状では、データベースや検索アルゴリズムの整備が不十分である点、および MALDI/TOF でのペプチドの振るまい (eg. イオン化時に反応が起こるなど、筆者ら未発表) が完全に理解されているわけではない点から、複数のデータベースで検索する、検索にもちいなかったピークを計算分子量と比較するなど検索結果の確認・評価が必須である。

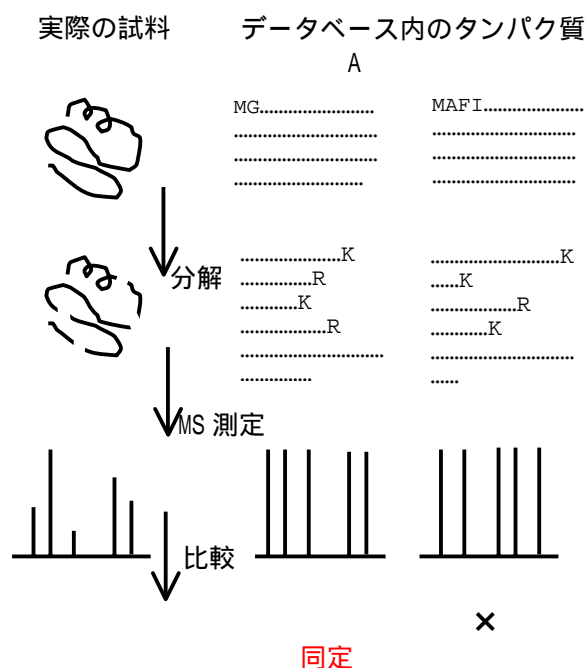


図 A-1 peptide mass fingerprinting 法の概略

(2) MS/MS データ

MS/MS 法はペプチドを分析計中で希ガス分子と衝突させたときに生ずる断片化したイオン (シークエンシオン) を質量分析することにより一次構造情報を得る方法である。(図 A-2) 従来のエドマン法に比べて高感度だが、シークエンシオンの一義的な帰属が不可能な場合があるため不確実とされる。MS/MS 情報のみからアミノ酸配列を決定するために、シグナル帰属を容易にするための試みがなされてきた。例えば、高尾らはペプチドの C 末端を酸素の安定同位体 ¹⁸O でラベルし、各末端由来のシグナルを区別する方法を提案した^{A-2)}。この概略を図 A-3 に示す。タンパク質を加水分解するとき水由来の酸素がペプチドの C 末端に取り込まれる。このとき H₂¹⁸O を半分含む溶液をもちいれば、半分のペプチド C 末端が ¹⁸O でラベルされる。これを MS/MS 分析すると N 末端由来のイオン (b 系列) は通常と同位体分布だが、C 末端由来 (y 系列) は 2[u] 違いのダブルピークになる。したがって、シークエンシオンの同位体ピークを分離できる高分解能装置をもちいた MS/MS では、ピーク形状によりシグナルの帰属が可能である。

この方法で、7 残基程度、部分配列を決定できれば、遺伝子・タンパク質配列データベースを検索して、タンパク質を同定可能である。また、それ以下でも、(1) 消化酵素名、(2) 元のペプチドの分子量、(3) 部分配列あるいはシークエンシオンの分子量セットをもちいてデータベース検索が可能である^{A-3)}。この方法は、一つ一つのペプチドを同定しうるので、上述の peptide mass fingerprinting 法とは異なり、あるタンパク質と相互作用するタンパク質の検索・同定などタンパク質混合物の分析に威力を発揮する。

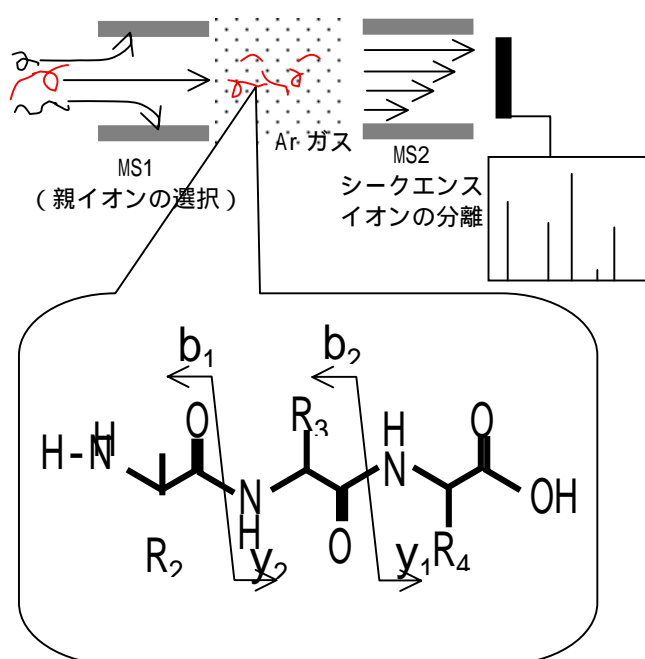
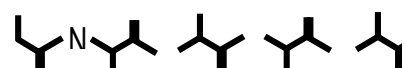


図 A-2 MS/MS 法の概略とペプチドの開裂パターン
酵素消化



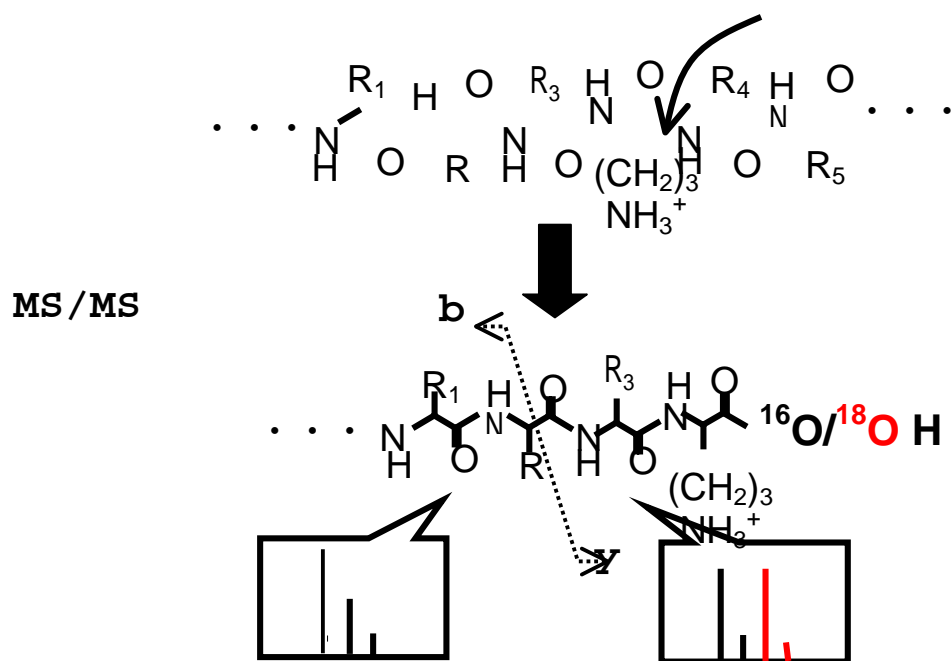


図 A-3 ^{18}O ラベル法によるシーケンスイオンの同定

参考文献

- A-1) MS Fit (<http://prospector.ucsf.edu/ucsfhtml3.2/msfit.htm>)
 A-2) Takao, T., Hori, H., Okamoto, K., Harada, A., Kamachi, M., Shimonishi, Y.: *Rapid Commun. Mass Spectrom.*, **5**, 312-315(1991)
 A-3) MS Tag (<http://prospector.ucsf.edu/ucsfhtml3.2/mstagfd.htm>)