

リン酸化プロテオミクスを用いた  
タンパク質キナーゼの基質プロファイルに関する研究

2013

金子（今村）春菜



## 目次

緒言 .....	4
第一章 NanoLC-MS/MS によるキナーゼ/基質のリン酸化プロファイリング .....	8
第一節 序 .....	8
第二節 実験デザイン .....	10
第三節 リン酸化部位同定 .....	10
第四節 モチーフ配列とその他の基質決定要素 .....	11
第五節 結語 .....	12
第二章 キナーゼ/基質プロファイルの細胞内シグナル伝達解析への応用 .....	14
第一節 序 .....	14
第二節 薬剤刺激によるリン酸化変動データの獲得 .....	16
第三節 <i>In vitro</i> 基質部位による PKA 標的基質の選別 .....	16
第四節 モチーフ配列による PKA 標的基質の選別 .....	17
第五節 結語 .....	17
結論 .....	19
発表論文 .....	20
謝辞 .....	21
引用文献 .....	22

## 緒言

地球上のすべての生物が保有する「ゲノム (genome)」は、生命を維持する為に必要な情報を記した「遺伝子 (gene)」の総体 (-ome) を意味し、生命の設計図に相当する。ヒトゲノム解析計画は 2003 年に終了が宣言され、約 30 億文字にのぼる全塩基配列が決定された。およそ  $10^{13}$  個あると言われている我々ヒト成人の体の細胞は、一部の例外を除き、同じゲノムのコピーを持っている。ところが、実際の細胞形態および機能は、時期や組織によって大きく異なる。この様な細胞の多様性、その集合体である組織別の機能、ひいては個としての生物種らしさを生み出す主要な高分子は、DNA から RNA を経て最終的に合成されるタンパク質である。生体内の活動を担うタンパク質の存在量や種類は、細胞種やその状態によって動的に変化し、互いを複雑に制御しあっている<sup>1,2</sup>。そこで、ゲノム解読の次段階であるポストゲノム時代では、生物学的機能への理解を更に深めるため、システムレベルでタンパク質の種類や数、またその相互作用などが研究対象とされはじめた。1995 年には、細胞内のタンパク質 (protein) の総体を意味する「プロテオーム (proteome)」という言葉が提唱され<sup>3</sup>、それらを扱う「プロテオミクス (proteomics)」は新たな研究分野として現在も注目されている<sup>4</sup>。

細胞機能の一つであるシグナル伝達は、様々な外部の状況に応じて細胞内部の機能や働きを変化させるための制御機構である。外界からのシグナルがリガンドとして細胞膜の受容体に結合することで、細胞は環境の変化を感知する。細胞外にある天然リガンドの中で、受容体に結合してこれを活性化するものには、ホルモン、神経伝達物質、サイトカイン、増殖因子などが含まれ、総じて第一メッセンジャーとよばれる。受容体の活性化により生産されるサイクリック AMP (cAMP) やカルシウムイオン ( $\text{Ca}^{2+}$ ) などの第二メッセンジャーによって、シグナルは更に分子から分子へと伝達され、最終的に核内で遺伝子発現量を制御することにより、細胞としてのふるまいに影響を及ぼす。この様なシグナルの流れを示すネットワークは「シグナル伝達経路」とよばれ、増殖や分化、アポトーシスなどといった主要な細胞活動に関わる重要な機構として働く<sup>5-7</sup>。

シグナル伝達の各段階を担う分子は上記の様に様々であるが、細胞表面受容体への刺激に対する応答の大半において、タンパク質の翻訳後修飾の一つであるリン酸化反応

を媒介とした伝達手段が利用されている<sup>8,9</sup>。一般に、第二メッセンジャーは特定の種類の酵素 (キナーゼと脱リン酸化酵素) に直接的または間接的に作用して、それらが標的タンパク質の特定部位のリン酸化または脱リン酸化を行う。タンパク質をリン酸化する酵素はプロテインキナーゼ (以下、キナーゼ) と呼ばれ、ヒトにおいては 518 ものキナーゼがリン酸化反応を制御することが予測されている<sup>10</sup>。真核細胞では、タンパク質のセリン (S)、トレオニン (T)、および、割合は少ないがチロシン (Y) へのリン酸基付加が知られている。これらのアミノ酸残基に負に帯電したリン酸基が付加されると、タンパク質の化学的性質は大きく変わる。分子の表面や調整タンパク質サブユニット間の接触面での化学的性質の変化は、基質の構造や結合タンパク質との相互作用の状態などに大きな影響を及ぼす。その結果として、基質タンパク質機能の活性化、あるいは不活性化が引き起こされる。細胞内のリン酸化と脱リン酸化のバランスが崩れると、生命にとって深刻な事態を招きかねない。例えば、チロシンキナーゼの恒常的な活性化は、ガン組織において広く見られる現象であり、ガン抑制、免疫抑制、炎症等の治療法に、リン酸化反応を促進または抑制する薬剤が広く利用されている。ゲノム解析からは翻訳後修飾であるリン酸化情報を得ることはできないため、リン酸化タンパク質の網羅的な解析「リン酸化プロテオミクス」による知見が、新薬の開発や病気の診断等において新たな展開をもたらすと期待される。

近年の質量分析 (MS) 関連技術の向上とゲノム情報の充実により、高速液体クロマトグラフィー (HPLC) とタンデム MS を組み合わせた LC-MS/MS システムによる「ショットガンプロテオーム解析法」が開発され、ハイスループットかつ網羅的なプロテオーム解析が現実的となった<sup>4</sup>。ショットガンプロテオミクスでは、タンパク質混合試料をトリプシン等のプロテアーゼで消化し、得られたペプチド断片を LC で分離しながら、タンデム MS へ導入する。タンデム MS ではペプチドの質量情報に加えて、プロダクトイオンスペクトルから部分アミノ酸配列情報および修飾基情報の取得が可能である。これらの情報と各ペプチド断片の質量情報を、タンパク質配列データベースに照合することで、タンパク質を一斉に同定する。ところが、リン酸化されているタンパク質はプロテオーム全体に対して存在量が非常に少なく、単純なプロテオーム解析からのリン酸化ペプチドの検出は困難であった。そこで、リン酸化ペプチドを選択的に濃縮する実験手法が開発され、これを LC-MS 導入前の試料に適用することで、リン酸化タンパク質および修飾部位の同定効率は飛躍的

に上昇した<sup>11</sup>。現在でも LC および MS の性能向上や実験技術改良、またデータ解析基盤の開発により、リン酸化プロテオーム解析から得られる情報は増大し続けている<sup>12-14</sup>。ここ数年の間にヒトのみならず、マウス、酵母、ショウジョウバエ、植物、また細菌<sup>15-20</sup>などの多様な生物で大規模リン酸化データが測定されている。リン酸化部位同定データは、Phospho.ELM<sup>21</sup>、PHOSIDA<sup>22</sup>、PhosphoSitePlus<sup>23</sup>、UniProt などの公共データベースで公開されており、現在では細胞内の 70%を超えるタンパク質が、リン酸化修飾を受けると言われている<sup>24</sup>。

LC-MS/MS を用いたリン酸化プロテオミクスは、これまで発見することができなかった大量の細胞内リン酸化部位の一斉同定を可能にし、多くの基質検出に貢献してきた。一方で、これらの大規模データには、その基質へのリン酸化反応を直接触媒した「責任キナーゼ」を示す情報が含まれない。責任キナーゼと基質の対応付けは、キナーゼ、基質の両分子の機能類推のみならず、シグナル経路のネットワークの新たな発見につながる。さらに、キナーゼ/基質情報をリン酸化プロテオーム情報に照合することで、細胞内の活性化キナーゼ群および活性化シグナル経路の特定が可能である。先行研究では、異なる細胞株由来のリン酸化プロテオームデータに対してこの解析を行い、活性化シグナルにより薬剤感受性の違いが生ずる機構を解明している<sup>25</sup>。

この様にキナーゼ/基質の関連情報の有用性が認められているにも関わらず、既知のキナーゼ/基質情報は十分とは言えず、これまでに同定された基質リン酸化部位は責任キナーゼが未知のものが殆どである。複雑な制御機構をもつシグナル伝達経路では、生体内におけるキナーゼの直接的な基質同定が難しい。基質にキナーゼを対応させるための実験的な手法として *in vitro* でのキナーゼ反応が考えられるが、数十万をこえる既知リン酸化部位を対象とした一対一の検証は現実的でない。一方で、数の限られたキナーゼ一つ一つの基質プロファイルは、解析可能である。これら網羅的な基質情報から、更に、キナーゼの基質認識に関わる共通の要素を抽出することができれば、その法則は新規に同定された基質のキナーゼ予測に応用できる。こうした背景をうけ、本論文ではキナーゼの基質プロファイルの大規模な探索を行った。

第一章では、キナーゼ/基質のペア情報を大量に取得するため、*in vitro* にて細胞由来のタンパク質抽出物とキナーゼを反応させ、LC-MS/MS が得意とする大規模解析にて偏り

無くリン酸化部位を同定した。ここで得られた広域なリン酸化部位情報を俯瞰し、個々の事象に共通してみられる配列的な特徴である「モチーフ配列」を、キナーゼの基質選択性として抽出した。また、第二章では、第一章で得られたキナーゼの基質プロファイルを利用し、細胞内 (*in vivo*) におけるキナーゼの直接的なリン酸化基質を特定した。本研究で示した *in vitro* 基質部位情報に基づく一連の責任キナーゼ特定手法は、これまでに蓄積された様々な細胞状態の既知基質情報に適応可能である。今後新たなキナーゼ/基質の関連付けが進むとともに、複雑な細胞内シグナル伝達経路の解明が期待される。

## 第一章 NanoLC-MS/MS によるキナーゼ/基質のリン酸化プロファイリング

### 第一節 序

細胞内でキナーゼの基質となるタンパク質は多く存在するが、キナーゼはそれぞれが選択的に基質のリン酸化を行う。細胞中に何十万もの S/T/Y が存在する中で、キナーゼはどのようにして標的の基質リン酸化部位を認識するのだろうか。キナーゼの基質選択には、細胞内局在や、タンパク質との複合体形成、また、足場タンパク質への結合など、複数要素の関与が知られている<sup>26</sup>。その中でもより直接的な要素として、リン酸化部位周辺に位置する特徴的なアミノ酸の組み合わせが知られている<sup>27,28</sup>。キナーゼ活性部位と直接相互作用するこれらの配列は、疎水性や電荷、構造的な開溝の深さなどの物理化学的特性において相互に補完しあう関係にある。そのため、リン酸化部位周辺の配列は、責任キナーゼを予測する際に有用な情報となる。キナーゼによっては特有のモチーフ配列を持つことが知られており、HPRD (Human Protein Reference Database)<sup>29</sup>、PhosphoSitePlus<sup>23</sup>、また NetPhorest<sup>30</sup>などのデータベースに登録されている。

これまでに、リン酸化部位周辺配列を情報学的に利用したキナーゼ予測手法は、いくつかの報告がなされている。モチーフ配列を基にした配列の照合解析は容易に行えることから、基質同定ベースの大規模 MS 分析データから生物学的に有意義な議論を引き出す為に幅広く利用されている<sup>17,25,31</sup>。また、学習データに周辺配列情報を使用した機会学習による予測も、多数報告がある<sup>32</sup>。しかし、利用するモチーフ配列の選択性の低さや基質情報の少なさが欠点となり、高い予測精度の実現は難しいというのが現状である。上述の NetPhorest では登録されたモチーフ配列情報を基にしたキナーゼ予測を行っているが、これらの配列依存的な欠点を補う為に、タンパク質相互作用や遺伝子制御ネットワークなど、そのほかのオミックス情報を配列情報と組み合わせることで、より効果的なキナーゼ予測を試みている。

キナーゼの活性ドメインが広く保存されていることを考えると、配列依存的に決定されたキナーゼファミリー分類が同じキナーゼは、類似したモチーフ配列を好む特性をもつ。しかし、現存のキナーゼ/基質関連情報の少なさをふまえると、基質情報の欠如こそがモチーフ配列の選択性を限定的なものにしている第一の原因として考えられる。現に、

およそ 209,000 事例のリン酸化部位情報を含有する公共データベースの PhosphoSitePlus でも、対応するキナーゼが判明している事例はわずか 6.6% (13,751 部位) である。そこで、キナーゼの配列選択性に関するより詳細な情報を取得するためには、より大規模な基質プロファイリングを行う必要があると考えた。

試験管内 (*in vitro*) での環境は、生体内 (*in vivo*) で見られるタンパク質相互作用や細胞内局在といった生物学的な制約がない。つまり、キナーゼは、自身の物理化学的な性質に基づき基質と反応するため、モチーフ配列の検出に適した実験系と言える。そこで、これまでプレートに合成ペプチドを結合した「ペプチドアレイ」に対してキナーゼ反応を行い、基質プロファイルを得る手法が利用されてきた<sup>33-35</sup>。ペプチドアレイを用いた大規模な先行研究として、酵母がもつ 122 のタンパク質キナーゼのうちの 61 種を用いて実験が行われ、好塩基性や好酸性といったキナーゼのアミノ酸選択性が報告された<sup>35</sup>。ペプチドアレイは、 $\gamma^{32}\text{P}$ -標識 ATP を用いることで、リン酸化部位の高感度な検出を可能とした一方で、以下の三つの欠点がある。第一に、アレイ上に配置される配列の多様性が限定的な点である。ペプチドアレイでは、S/T/Y のリン酸化される残基と任意の二カ所の残基が固定された配列を持つペプチドが合成される。そのため、探索対象は限定され、細胞内に存在し得る幅広いアミノ酸配列の組み合わせを検証することはできない。第二に、結果として得られる情報はキナーゼの選択性を平均化したものになり、意義のある組み合わせを抜き出すことができない。最後に、結合に用いる化学物質や結合方法によって、ペプチドの固定化法が異なり、反応効率に影響を及ぼす<sup>36,37</sup>。ペプチドアレイの代替手法として、タンパク質全长をアレイに固定化した「タンパク質アレイ」がある<sup>38</sup>。近年のアレイを用いた最も大規模な研究として、Newman らによる報告では 289 のキナーゼをタンパク質アレイと反応させ基質プロファイルを取得し、得られた結果から同定したモチーフ配列を基にキナーゼ/基質のネットワーク構築を行った<sup>39</sup>。しかし、ペプチドアレイと同様に、タンパク質アレイは上述の問題を含有し、また、タンパク質中のリン酸化部位を同定するためには、更なる解析が必要である。

一方、LC-MS/MS の分析技術は、アミノ酸残基レベルでのリン酸化反応を同定する際の有力なツールである。そのため、近年では LC-MS/MS 分析と溶液中での *in vitro* キナーゼ反応を組み合わせた手法が、キナーゼ/基質関連情報を取得するために広く用いられて

いる<sup>40-48</sup>。この手法には、MSが配列の同定とリン酸化部位の決定を同時に行えるというメリットがある。また、溶液中でのキナーゼ反応は、基質をアレイに固定化する必要がないため、細胞抽出物を反応対象として利用できる。後者の利点により、細胞内プロテオーム中のアミノ酸含有比および組み合わせを保ったペプチド配列という、幅広い基質候補を用いた基質プロファイルの取得が可能となった。

本章では、これまでに同定されなかった基質を含む、より広域かつ詳細な基質プロファイルを同定するために、LC-MS/MSと*in vitro*リン酸化反応の組み合わせ実験系を構築した。cAMP依存性Aキナーゼ(PKA)、細胞外シグナル調節キナーゼ1(ERK1)、および、RAC-alpha serine/threonine-protein kinase(AKT1)の3種のセリン-トレオニンキナーゼを用いてこの実験系を検証した。その結果、各キナーゼの基質プロファイルから、モチーフ配列をはじめとするキナーゼの基質選択性を示す特異的要素が特定された。

## 第二節 実験デザイン

まず、細胞からタンパク質を抽出し、上清とペレットに分画した。*In vitro*キナーゼ反応の前に、内在性のリン酸基を除くため、脱リン酸化反応を行った。脱リン酸化酵素として、穏やかな加熱により非可逆的に不活性化可能な、熱失活性アルカリホスファターゼ(TSAP)を利用した。TSAP反応は特別なバッファーと追加の金属を必要としないことから、脱リン酸化反応の後に続けてキナーゼ反応を行うことができ、時間および試料の損失を防ぐことができる。各キナーゼと反応した試料(キナーゼ(+))と蒸留水を添加したコントロール試料(キナーゼ(-))を用意した後、それぞれを消化酵素により断片化した。その後、酸化チタン(TiO<sub>2</sub>)を用いたヒドロキシ酸修飾酸化金属クロマトグラフィー(HAMMOC)<sup>49</sup>にて、リン酸化ペプチドのみを濃縮し、LC-MS/MS測定に供した。

## 第三節 リン酸化部位同定

最終的に、PKA、ERK1、およびAKT1の基質リン酸化ペプチドとして、5,625、6,455、および2,785配列が同定された。これらのリン酸化ペプチドに対して、MS<sup>2</sup>スペクトル情報から算出されたPTMスコア、あるいはリン酸化部位決定イオンの同定を根拠に、信頼性の高いリン酸化部位のみを抽出した。その結果、3,585、4,347、および1,778のリン酸化部位情

報を獲得した。*In vitro* キナーゼ反応と LC-MS 分析を組み合わせた手法は、以前にも数種のキナーゼに対して行われている。特に代表的なキナーゼである PKA を研究対象とした報告は多く、初期の報告としては Huang らがラットの子宮を用いた実験をおこない、PKA の *in vitro* リン酸化ペプチドを 61 種報告している<sup>40</sup>。続いて、Douglas らが類似した手法を用いて、ラットの複数種の組織から合わせて 934 の *in vitro* 基質ペプチドを同定した<sup>41</sup>。Chou らは大腸菌内でヒトキナーゼを過剰発現させることにより、菌内のプロテオームをタンパク質ライブラリーとして利用し、*in vivo* キナーゼ反応を行うというユニークな手法を確立した<sup>50</sup>。リン酸化部位を同定した後、その情報を基にモチーフ配列を抽出し、ヒトタンパク質に対する *in silico* リン酸化部位予測に活用した。この研究からは、806 の PKA リン酸化部位が報告された。これらの報告と比較すると、今回のリン酸化部位同定数は数倍から数百倍の規模であり、この研究から今までにない幅広いリン酸化部位の検出が行われたことを示している。

#### 第四節 モチーフ配列とその他の基質決定要素

より詳細なキナーゼの基質選択性同定に向けて、既存のモチーフ抽出ソフトウェアである motif-X<sup>51</sup> を用いて独立したモチーフ配列へと分解した。

##### (1) PKA と AKT1 に関する結果と議論

既知リン酸化部位情報を用いたモチーフ解析からは、PKA と AKT1 のどちらからも N 末側に塩基性残基を持つモチーフ配列が抽出された。一方、本研究では、N 末側のみならず、C 末側のみ、あるいは N および C 末側に特徴的なアミノ酸を持つモチーフが検出された。例えば、N 末側の塩基性残基と+1 の疎水性残基の組み合わせは PKA モチーフとして同定され (増倍比=10.69-15.85)、AKT1 モチーフの中では N 末側の塩基性残基と+2 の S の組み合わせが同定された (増倍比=9.43-36.21)。増倍比は比較的低い (3.96) が、フェニルアラニン (F) を+1 に持つモチーフ配列も AKT1 で同定された。これらの結果は、限られた基質では有意な差として捉えられなかった微細な基質選択性が、大規模基質同定により検出できるようになったことを示す。

## (2) ERK1 に関する結果と議論

アミノ酸頻度プロットと同様に、既知情報および大規模情報の両方のモチーフ配列解析から ERK1 の+1 プロリン (P) への選択性を示す配列が同定された。さらに、大規模基質プロファイルを基にしたモチーフ解析では、PKA、AKT1 と同じく ERK1 でも多様なモチーフ配列が検出できた。これらには、C 末側の塩基性の残基をもつモチーフ配列などが含まれた。

ERK1 特徴的な選択性として、リン酸化部位を複数持つペプチド (マルチリン酸化ペプチド) が、他の二種のキナーゼより多く見つかった。この結果から、リン酸基の存在が更なる ERK1 のリン酸化を促進するのではないかと考えた。

## 第五節 結語

本章では、*in vitro* キナーゼ反応と LC-MS/MS 分析により、大規模な基質のプロファイルを同定した。*In vitro* の実験系はタンパク質の細胞局在や分子間相互作用などの生体内での制約に囚われず、キナーゼが標的とするリン酸化部位とその周辺配列情報を獲得するための直接的な手法である。

この実験系を PKA、ERK1、および AKT1 の 3 種のキナーゼに対して適用し、大規模な基質プロファイルを獲得した。得られたデータの情報量の豊富さゆえ、公共データベース登録のあった既知リン酸化部位情報と比べると、より詳細なモチーフ配列の検出が可能となった。大規模基質プロファイルから抽出したモチーフ配列は、既知リン酸化部位より同定されたモチーフ配列のみならず、キナーゼ特異的なモチーフを含む新規のモチーフ配列を多数同定した。同じキナーゼファミリーである PKA と AKT1 のモチーフ配列間で選択性の評価を行い、既知基質情報の欠如が選択性の低いモチーフ配列の同定の原因となるという具体的な事例を示した。更に、モチーフ配列のみならず、リン酸基受容体 (S または T) やリン酸基自身に対するキナーゼの選択性といった、その他の基質認識要素を示唆した。

現在、新たなリン酸化部位が次々と同定され、データベースに蓄積されていく中で、殆どのリン酸化部位はキナーゼと結びつきの無い“オーファン基質”として捨て置かれている。本研究で明らかとなった、モチーフ配列をはじめとする基質認識要素は、より高い精度を持った基質/キナーゼの予測に貢献する。また、本研究の実験系は簡易かつ汎用的で

あり、今回利用した 3 種のキナーゼ以外にも広く応用可能である。ここに示した基質プロファイルの取得が複数のキナーゼに対して行われるにつれ、キナーゼの選択性がより明確となり、リン酸化を土台としたシグナル経路の探索を更に促進すると考えられる。

## 第二章 キナーゼ/基質プロファイルの細胞内シグナル伝達解析への応用

### 第一節 序

現在、対象生物・組織の別、阻害薬や増殖因子の様な刺激による影響など、様々な状況におけるリン酸化プロテオームが同定されている<sup>52-54</sup>。In vivo 実験から得られるリン酸化プロテオーム情報は、時間変化や細胞局在、分子間相互作用などの生物学的な制御機構を反映しており、細胞内で上流から下流へと流れるシグナル経路を理解するための直接的な情報となる。なかでも、細胞内におけるキナーゼの標的基質を同定するには、薬剤による刺激実験や、標的キナーゼの高発現細胞、または RNA 干渉による阻害実験が用いられ、キナーゼ活性、あるいは阻害に応答した細胞内のリン酸化変動から、キナーゼが標的とする基質が特定される。その際に、リン酸化タンパク質同定の網羅性は、効果的な直接基質の特定を行う上で肝要となる。LC-MS/MS を用いたリン酸化プロテオーム解析はこの目的にかなった手法であり、この種の実験と合わせて多用されている。

In vivo でのキナーゼ阻害/活性実験の大きな問題として、変動リン酸化部位に含まれる偽陽性基質がある。この主な原因としては、以下の二点が考えられる。一つ目は、関連分子への影響である。細胞内で起こるシグナル伝達は、連鎖的なリン酸化反応を含む。特定のキナーゼへの阻害、あるいは活性は、標的キナーゼのみならず、下流のリン酸化状態にも間接的な影響を及ぼす。すなわち、LC-MS 分析により同定されたリン酸化プロテオーム情報には、標的キナーゼの直接的な基質とともに、その下流分子も含まれる。二つ目の理由として、標的キナーゼ以外のキナーゼに対して薬剤の効果がみられる、「オフターゲット効果」がある。キナーゼの阻害薬の多くは触媒活性部位を標的とし、ATP に競合的な阻害をする様に設計されている。ところが、キナーゼの活性部位の配列は保存性が高く、互いに類似している。そのため、特定のキナーゼに対して選択性の高い阻害薬を作成することは難しく、薬剤の選択性によっては、ファミリー分類が同じキナーゼには一様の効果を示すものもある。H89 は PKA に高い選択性を示す阻害薬として知られているが、より高い濃度では CAM キナーゼ II やカゼインキナーゼ 1、ミオシン軽鎖キナーゼ、PKC、II 型 Rho キナーゼに対しても阻害効果を示す<sup>55</sup>。また、PKA の活性化薬として知られているフォルスコリン (Fsk) は、アデニル酸シクラーゼを活性化させることで cAMP の細胞内濃度

を上昇し、PKA を活性化する<sup>56</sup>。そのため、cAMP 依存的に活性を持つ全ての酵素も同様に活性化されることとなる。

以上のことから、*in vivo* リン酸化プロテオミクスにより得られる多数の基質候補の中から、直接的な基質を選別するには何らかの有効なフィルターが必要となる。ここで、既知の *in vitro* 基質情報、あるいはモチーフ配列が有力な選別手段となる。*In vitro* キナーゼ実験は、あるキナーゼと基質候補の間におこる直接的な相互作用の検証に用いられる。そのため、既知の *in vitro* 基質のリン酸化部位の直接的な照合は、信頼性の高い部位情報の抽出を行う上で、最も単純かつ簡易な手法である。しかし、この手法にて、より効果的なスクリーニングを行うためには、*in vitro* 基質プロファイルの網羅性が制約となる。既知の基質情報が少ない、あるいは、*in vitro* 実験を行った基質のタンパク質プロファイルが *in vivo* 実験で用いる試料と一致していない場合、二つの母集団の重複により得られる *in vivo* 基質候補の数も必然的に限られてくる。一方、モチーフ配列は具体的なリン酸化部位からの特徴抽出により一段階抽象化された情報であり、発現タンパク質プロファイルの一致に関わらず汎用的に利用できる<sup>57-59</sup>。ただし、モチーフ配列の選択性が低い場合、その結果に偽陽性が多く含まれることとなる。

本章では第一章で獲得した基質プロファイルを用いて、*in vivo* 細胞内の PKA 直接基質の特定を試みた。PKA は幅広いシグナル経路に関わる基質を標的とし、細胞骨格の制御やカルシウムチャネルの調整、細胞の運動性などの細胞機能に影響を及ぼす。新規の生体内基質の同定は、PKA の機能推定および関連シグナル経路に新たな知見をもたらすことが期待される。

## 第二節 薬剤刺激によるリン酸化変動データの獲得

リン酸化変動データを獲得するために、SILAC を用いた定量解析を行った。SILAC とは、 $^{12}\text{C}_6$  または  $^{13}\text{C}_6$  L-リジンのような軽いまたは、重い必須アミノ酸を含む培地で細胞を培養する標識法である。細胞は培地中のアミノ酸を取り込み、同位体標識の異なるタンパク質を合成する。そのため、混合試料をショットガンプロテオミクスで解析すると、各試料のペプチドは一定の質量電荷比の差をもって質量分析計で検出される。例えば H (heavy) 標識した細胞に阻害薬を添加し、L (light) 標識した細胞をコントロールとした場合、ピークの面積比 (H/L 比) から阻害薬によるリン酸化変動比が取得できる。今回の実験では PKA の活性化薬としてフォルスコリン、阻害薬として H89 を用いて、薬剤刺激によるリン酸化変動を定量した。H/L 細胞の混合試料からタンパク質抽出を行った後に細胞分画をおこない、上清分画のみを実験に利用した。阻害薬または活性化薬による刺激実験は、H/L 標識を取り替えて二回繰り返し、計四種の細胞試料に対して、それぞれ二回ずつ MS 測定をおこなった。

## 第三節 *In vitro* 基質部位による PKA 標的基質の選別

薬剤刺激による *in vivo* リン酸化プロテオームから、信頼性の高い PKA 直接基質を選別することを目的に、第一章で得た *in vitro* PKA リン酸化部位との照合を行った。*In vitro* リン酸化部位で同定され、かつ、薬剤添加によって期待される定量変化を示した部位を PKA の直接基質候補とした。

その結果、24 リン酸化ペプチドにて活性化薬による促進、あるいは阻害薬による抑制が見られた。その内、活性化薬による促進は 19 種、阻害薬による抑制は 9 種、両方に共通して含まれたリン酸化ペプチドは 4 種であった。

この内、上流キナーゼが PKA であると既に *in vivo* 実験を用いて推定されているリン酸化部位は、Filamin-A (FLNA\_HUMAN) の S2152 が該当した<sup>60</sup>。SILAC による定量結果を検証するため、リン酸化抗体を用いたウエスタンブロッティングを行った。得られた基質候補のうち唯一抗体が入手可能であった、Filamin-A の S2152 を実験対象とした。泳動タンパク量のコントロールとして、アクチンの検出を行い、コントロールに対する強度比を最終的な変動値として算出した。その結果、Filamin-A のリン酸化量は活性化薬により増加、および、阻害薬によって減少し、LC-MS 法による定量結果と一致した。

#### 第四節 モチーフ配列による PKA 標的基質の選別

選択性の高いモチーフ配列トップ 5 位の配列を用いて、PKA の基質候補を特定した。リン酸化部位のうち、モチーフ配列に該当し、かつ、定量値に変動があったリン酸化部位は全部で 19 部位であった。この結果には、既知の報告で上流キナーゼが PKA と推定されているリン酸化部位が 4 カ所含まれた。

#### 第五節 結語

薬剤刺激によるキナーゼの活性変化は、細胞内リン酸化プロテオームを変動させる。この情報は、キナーゼが標的とする直接基質を特定するために有用であるが、変動するリン酸化部位には下流分子も含まれる。そこで、本章では、PKA の *in vitro* 基質プロファイルを用いた二つの手法を適用し、細胞内での PKA 基質同定を試みた。

第一に、より簡便な方法として、基質リン酸化部位の直接的な比較を行った。その結果、既知 PKA 基質を含む 4 リン酸化部位が PKA の基質候補として同定された。次に、モチーフ配列による基質選択を行うため、薬剤刺激によるリン酸化変動情報を利用し PKAモチーフ配列の選択性を評価した。この評価で上位にランク付けされたモチーフは、第一章で PKA 選択性が高いと同定された配列を多く含んだ。最後に、選択性の高いモチーフ配列を含有し、かつ、薬剤刺激に応じたリン酸化変動を示したリン酸化部位を PKA の直接基質候補として抽出したところ、19 リン酸化部位が選別された。その内の、4 カ所は既知の報告から、細胞中で PKA が責任キナーゼとして機能すると実験的に検証されている。また、11 カ所は PKA の *in vitro* 基質として同定されていた。以上の結果をまとめると、本手法によって、5 カ所の既知リン酸化部位を含む、22 のリン酸化部位が PKA の直接基質として同定された。この PKA の *in vivo* 既知基質を多く含んだ選別結果は、*in vitro* の基質認識要素が、*in vivo* 基質の探索に効果的であることを裏付けている。一方で、偽陽性と偽陰性はトレードオフの関係にある。本手法で選別した基質候補の偽陽性の低減を主眼としており、今回の直接基質選別で同定できなかったリン酸化部位に PKA の *in vivo* 基質が含まれている可能性は十分にある。

本章で示した *in vitro* 基質を用いた選別手法は、大規模な基質プロファイルが取得されているキナーゼの基質探索に広く応用可能である。また、現在、時間経過や刺激に

応じた様々な細胞状態のリン酸化プロテオーム情報が、蓄積され続けている。この様な多様な *in vivo* 情報に本手法を適用し、時間や細胞局在などの特定の状態で同定されるキナーゼ/基質情報の比較解析を行うことで、より動的な生体内シグナル伝達経路の考察が期待できる。

## 結論

タンパク質キナーゼの基質プロファイルの同定に向けて、リン酸化プロテオミクスを用いた大規模な基質同定を行い、以下の知見を得た。

第一章では、特定のキナーゼの大規模な基質プロファイルを獲得するため、*in vitro* キナーゼ反応と LC-MS/MS 分析を合わせた実験手法を構築した。本手法を PKA、ERK1、および AKT1 の三種のキナーゼに応用し、大規模な基質タンパク質、およびリン酸化部位の同定を試みた。その結果、いずれのキナーゼからも、既存報告よりはるかに大規模の基質同定に成功した。次に、この実験で得られた基質プロファイルから、キナーゼの基質認識要素を抽出した。その結果、キナーゼが基質認識に利用する要素として、リン酸化部位の周辺に存在する特徴的なアミノ酸の組み合わせであるモチーフ配列が、新規の配列を含め、多数同定された。また、リン酸基受容体の選択性、および、リン酸基自身への選択性が基質認識に関与することを見出した。第二章では、生体内のシグナル伝達解明のため、*in vitro* 基質プロファイルを *in vivo* での基質同定に利用する二通りの解析手法を構築し、PKA の直接基質同定に応用した。第一の方法では、第一章で構築した *in vitro* 基質プロファイル同定実験で獲得した PKA の *in vitro* リン酸化部位を用いた。このリン酸化部位と、*in vivo* で薬剤添加依存的にリン酸化が変動した部位を照合した。その結果、1つの既知基質を含む4つの PKA 標的基質が同定された。第二の方法では、より汎用的な手法としてモチーフ配列を利用した。リン酸化変動部位の内、選択性の高いモチーフ配列を持つ部位を PKA の直接基質として選別した。その結果、4つの既知 PKA 基質を含む、19の PKA 標的基質が同定された。以上の結果、PKA の *in vitro* 基質プロファイルを基にした本法を用いることで、既知基質を多く含む信頼性の高い直接基質が特定できた。

以上、キナーゼ/基質関連情報の獲得に向けて、リン酸化プロテオミクスを基盤とする大規模解析系を構築した。現在、新たなリン酸化部位が次々と同定される一方で、殆どのリン酸化部位は責任キナーゼとの関連付けがなされていない。第一章で示したキナーゼのより詳細な基質プロファイルの獲得、および、第二章で示した基質認識要素を用いた生体内でのキナーゼ/基質情報の特定は、この問題を解決し、細胞内シグナル伝達ネットワーク全体のダイナミクスの俯瞰を実現する。このことは、細胞機能の解明や、創薬・医療の発展のみならず、生命科学全般の進歩に大きく寄与するものと期待される。

## 発表論文

### 第一章:

[1] Phosphoproteomic profile of kinase substrates by nanoLC-MS/MS

Haruna Imamura, Naoyuki Sugiyama, Masaki Wakabayashi, Yasushi Ishihama

投稿予定

### 第二章:

[2] Towards uncovering intracellular signaling based on kinase/substrate profile

Haruna Imamura, Naoyuki Sugiyama, Masaki Wakabayashi, Yasushi Ishihama

投稿予定

### その他:

[3] Towards the systematic discovery of signal transduction networks using phosphorylation dynamics data

Haruna Imamura, Nozomu Yachie, Rintaro Saito, Yasushi Ishihama, Masaru Tomita

平成 22 年 5 月発行 *BMC Bioinformatics* 第 11 巻 232

[4] Analytical strategies for shotgun phosphoproteomics: Status and prospects

Haruna Imamura, Masaki Wakabayashi, Yasushi Ishihama

平成 24 年 10 月発行 *Seminars in Cell & Developmental Biology* 第 23 巻 8 号 836-842 頁

[5] Temporal Profiling of Lapatinib-suppressed Phosphorylation Signals in EGFR/HER2 Pathways

Koshi Imami, Naoyuki Sugiyama, Haruna Imamura, Masaki Wakabayashi, Masaru Tomita,

Masatoshi Taniguchi, Takayuki Ueno, Masakazu Toi, Yasushi Ishihama

平成 24 年 12 月発行 *Molecular & Cellular Proteomics* 第 11 巻 1741-1757 頁

## 謝辞

本研究を行うに当たり、終始ご指導を賜りました、京都大学大学院薬学研究科 石濱泰教授に深謝申し上げます。また、本研究の遂行を基礎から指導して頂きました、京都大学大学院薬学研究科 杉山直幸准教授、並びに、若林真樹助教に深く感謝申し上げます。

本研究を進めるに当たり、有益なご助言を賜りました富山大学大学院医学薬学研究部 生体界面化学研究室 中野実教授、慶応義塾大学 先端生命科学研究所 冨田勝教授、並びに金井昭夫教授に感謝申し上げます。

本研究を行うに当たり、御討議を頂き大変お世話になりましたエーザイ株式会社 バイオマーカー&パーソナライズドメディスン機能ユニット 田畑剛氏、アジレント・テクノロジー株式会社 ライフサイエンス・化学分析本部 京野完博士、ハーバード大学 血液学部門 ブリガム・アンド・ウィメンズ病院 増田豪博士、ブリティッシュコロンビア大学 ハイスループトバイオロジーセンター 今見考志博士、カルフォルニア大学 サンフランシスコ校 篠田幸作博士、京都大学 iPS 細胞研究所 岩崎未央博士、慶応義塾大学 先端生命科学研究所 玉木聡志氏、京都大学大学院薬学研究科 製剤機能解析学分野の卒業生並びに在学生の皆様に感謝致します。本研究の一部は日本学術振興会 (DC1) の援助によるものであります。

最後に、これまでの研究生生活を支えてくれた家族に感謝致します。

## 引用文献

1. Krüger, M. *et al.* Dissection of the insulin signaling pathway via quantitative phosphoproteomics. *Proc Natl Acad Sci USA* **105**, 2451–2456 (2008).
2. Lundberg, E. *et al.* Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol* **6**, 450 (2010).
3. Wasinger, V. C. *et al.* Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* **16**, 1090–1094 (1995).
4. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
5. Ang, X. L. & Wade Harper, J. SCF-mediated protein degradation and cell cycle control. *Oncogene* **24**, 2860–2870 (2005).
6. Cohen, P. The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends Biochem Sci* **25**, 596–601 (2000).
7. Schlessinger, J. & Lemmon, M. A. SH2 and PTB domains in tyrosine kinase signaling. *Sci STKE* **2003**, RE12 (2003).
8. Pawson, T. & Scott, J. D. Protein phosphorylation in signaling--50 years and counting. *Trends Biochem Sci* **30**, 286–290 (2005).
9. Beltrao, P., Bork, P., Krogan, N. J. & van Noort, V. Evolution and functional cross-talk of protein post-translational modifications. *Mol Syst Biol* **9**, 714 (2013).
10. Manning, G., Plowman, G. D., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* **27**, 514–520 (2002).
11. Imamura, H., Wakabayashi, M. & Ishihama, Y. Analytical strategies for shotgun phosphoproteomics: Status and prospects. *Semin Cell Dev Biol* (2012).  
doi:10.1016/j.semcdb.2012.05.007
12. Gomase, V. S., Kale, K. V., Tagore, S. & Hatture, S. R. Proteomics: technologies for protein analysis. *Curr Drug Metab* **9**, 213–220 (2008).

13. Olsen, J. V. & Mann, M. Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry. *Mol Cell Proteomics* **12**, 3444–3452 (2013).
14. Choudhary, C. & Mann, M. Decoding signalling networks by mass spectrometry-based proteomics. *Nat Rev Mol Cell Biol* **11**, 427–439 (2010).
15. Chi, A. *et al.* Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proc Natl Acad Sci USA* **104**, 2193–2198 (2007).
16. Hilger, M., Bonaldi, T., Gnad, F. & Mann, M. Systems-wide analysis of a phosphatase knock-down by quantitative proteomics and phosphoproteomics. *Mol Cell Proteomics* **8**, 1908–1920 (2009).
17. Olsen, J. V. *et al.* Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Science signaling* **3**, ra3 (2010).
18. Ravichandran, A., Sugiyama, N., Tomita, M., Swarup, S. & Ishihama, Y. Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic *Pseudomonas* species. *Proteomics* **9**, 2764–2775 (2009).
19. Sugiyama, N. *et al.* Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. *Mol Syst Biol* **4**, 193 (2008).
20. Villén, J., Beausoleil, S. A., Gerber, S. A. & Gygi, S. P. Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci USA* **104**, 1488–1493 (2007).
21. Diella, F., Gould, C. M., Chica, C., Via, A. & Gibson, T. J. Phospho.ELM: a database of phosphorylation sites--update 2008. *Nucleic Acids Res* **36**, D240–4 (2008).
22. Gnad, F. *et al.* PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol* **8**, R250 (2007).
23. Hornbeck, P. V. *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* **40**, D261–70 (2012).
24. Imai, K. & Yau, S. L. F. *Quantitative Proteome Analysis*. (CRC Press, 2013).

25. Casado, P. *et al.* Kinase-Substrate Enrichment Analysis Provides Insights into the Heterogeneity of Signaling Pathway Activation in Leukemia Cells. *Science signaling* **6**, rs6 (2013).
26. Ubersax, J. A. & Ferrell, J. E. Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol* **8**, 530–541 (2007).
27. Pearson, R. B. & Kemp, B. E. Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Meth Enzymol* **200**, 62–81 (1991).
28. Pinna, L. A. & Ruzzene, M. How do protein kinases recognize their substrates? *Biochim Biophys Acta* **1314**, 191–225 (1996).
29. Amanchy, R. *et al.* A curated compendium of phosphorylation motifs. *Nat Biotechnol* **25**, 285–286 (2007).
30. Miller, M. L. *et al.* Linear Motif Atlas for Phosphorylation-Dependent Signaling. *Science signaling* **1**, ra2 (2008).
31. Sudhir, P.-R. *et al.* Phosphoproteomics identifies oncogenic Ras signaling targets and their involvement in lung adenocarcinomas. *PLoS ONE* **6**, e20199 (2011).
32. Trost, B. & Kusalik, A. Computational prediction of eukaryotic phosphorylation sites. *Bioinformatics* **27**, 2927–2935 (2011).
33. Songyang, Z. *et al.* Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol* **4**, 973–982 (1994).
34. Hutti, J. E. *et al.* A rapid method for determining protein kinase phosphorylation specificity. *Nat Methods* **1**, 27–29 (2004).
35. Mok, J. *et al.* Deciphering Protein Kinase Specificity Through Large-Scale Analysis of Yeast Phosphorylation Site Motifs. *Science signaling* **3**, ra12 (2010).
36. Inamori, K. *et al.* Detection and quantification of on-chip phosphorylated peptides by surface plasmon resonance imaging techniques using a phosphate capture molecule. *Anal Chem* **77**, 3979–3985 (2005).
37. Inamori, K. *et al.* Optimal surface chemistry for peptide immobilization in on-chip phosphorylation analysis. *Anal Chem* **80**, 643–650 (2008).

38. Mok, J., Im, H. & Snyder, M. Global identification of protein kinase substrates by protein microarray analysis. *Nat Protoc* **4**, 1820–1827 (2009).
39. Newman, R. H. *et al.* Construction of human activity-based phosphorylation networks. *Mol Syst Biol* **9**, 655 (2013).
40. Huang, S.-Y., Tsai, M.-L., Chen, G.-Y., Wu, C.-J. & Chen, S.-H. A systematic MS-based approach for identifying in vitro substrates of PKA and PKG in rat uteri. *J Proteome Res* **6**, 2674–2684 (2007).
41. Douglass, J. *et al.* Identifying protein kinase target preferences using mass spectrometry. *Am J Physiol, Cell Physiol* **303**, C715–27 (2012).
42. Xue, L. *et al.* Sensitive kinase assay linked with phosphoproteomics for identifying direct kinase substrates. **109**, 5615–5620 (2012).
43. Xue, L., Geahlen, R. L. & Tao, W. A. Identification of Direct Tyrosine Kinase Substrates Based on Protein Kinase Assay-Linked Phosphoproteomics. *Mol Cell Proteomics* (2013). doi:10.1074/mcp.O113.027722
44. Knight, J. D. *et al.* A novel whole-cell lysate kinase assay identifies substrates of the p38 MAPK in differentiating myoblasts. *Skeletal Muscle* **2**, 5 (2012).
45. Kettenbach, A. N. *et al.* Rapid Determination of Multiple Linear Kinase Substrate Motifs by Mass Spectrometry. *Chem Biol* **19**, 608–618 (2012).
46. Hennrich, M. L. *et al.* Universal Quantitative Kinase Assay Based on Diagonal SCX Chromatography and Stable Isotope Dimethyl Labeling Provides High-definition Kinase Consensus Motifs for PKA and Human Mps1. *J Proteome Res* **12**, 2214–2224 (2013).
47. Wang, C. *et al.* Determination of CK2 Specificity and Substrates by Proteome-Derived Peptide Libraries. *J Proteome Res* (2013). doi:10.1021/pr4002965
48. Bian, Y. *et al.* Global screening of CK2 kinase substrates by an integrated phosphoproteomics workflow. *Sci. Rep.* **3**, 3460 (2013).
49. Sugiyama, N. *et al.* Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics* **6**, 1103–1109 (2007).

50. Chou, M. F. *et al.* Using bacteria to determine protein kinase specificity and predict target substrates. *PLoS ONE* **7**, e52747 (2012).
51. Schwartz, D. & Gygi, S. P. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol* **23**, 1391–1398 (2005).
52. Kosako, H. *et al.* Phosphoproteomics reveals new ERK MAP kinase targets and links ERK to nucleoporin-mediated nuclear transport. *Nat Struct Mol Biol* **16**, 1026–1035 (2009).
53. Olsen, J. V. *et al.* Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **127**, 635–648 (2006).
54. Van Hoof, D. *et al.* Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell* **5**, 214–226 (2009).
55. Chijiwa, T. *et al.* Inhibition of forskolin-induced neurite outgrowth and protein phosphorylation by a newly synthesized selective inhibitor of cyclic AMP-dependent protein kinase, N-[2-(p-bromocinnamylamino)ethyl]-5-isoquinolinesulfonamide (H-89), of PC12D pheochromocytoma cells. *J Biol Chem* **265**, 5267–5272 (1990).
56. de Souza, N. J., Dohadwalla, A. N. & Reden, J. Forskolin: a labdane diterpenoid with antihypertensive, positive inotropic, platelet aggregation inhibitory, and adenylate cyclase activating properties. *Med Res Rev* **3**, 201–219 (1983).
57. Andersen, J. N. *et al.* Pathway-based identification of biomarkers for targeted therapeutics: personalized oncology with PI3K pathway inhibitors. *Sci Transl Med* **2**, 43ra55 (2010).
58. Moritz, A. *et al.* Akt-RSK-S6 Kinase Signaling Networks Activated by Oncogenic Receptor Tyrosine Kinases. *Science signaling* **3**, ra64 (2010).
59. Courcelles, M. *et al.* Phosphoproteome dynamics reveal novel ERK1/2 MAP kinase substrates with broad spectrum of functions. *Mol Syst Biol* **9**, 669 (2013).
60. Jay, D., García, E. J. & la Luz Ibarra, de, M. In situ determination of a PKA phosphorylation site in the C-terminal region of filamin. *Mol Cell Biochem* **260**, 49–53 (2004).