

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

略語表.....	4
緒言	7
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.....	12
第一節 序.....	12
第二節 実験デザイン	14
第三節 脱リン酸化反応効率の検証	16
第四節 リン酸化部位同定.....	18
第五節 リン酸化部位周辺のアミノ酸出現頻度	24
第六節 モチーフ配列とその他の基質決定要素	28
第七節 結語	38
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.....	39
第一節 序.....	39
第二節 薬剤刺激によるリン酸化変動データの獲得	41
第三節 <i>In vitro</i> 基質部位による PKA 標的基質の選別.....	44
第四節 同定情報を用いたモチーフ配列選択性の評価.....	49
第五節 モチーフ配列による PKA 標的基質の選別	52
第六節 結語	56
結論	57
実験の部.....	58
発表論文.....	66
謝辞	67
引用文献.....	68

略語表

ATP	: アデノシン三リン酸 (adenosine triphosphate)
cAMP	: 環状アデノシン一リン酸 (cyclic adenosine monophosphate)
CID	: 衝突誘起解離 (collision-induced dissociation)
DNA	: デオキシリボ核酸 (deoxyribonucleic acid)
DMEM	: ダルベッコ改変イーグル培地 (Dulbecco's modified Eagle's medium)
DTT	: ジチオトレイトール (dithiothreitol)
ERK1	: 細胞外シグナル調節キナーゼ 1 (extracellular signal-regulated kinase 1)
FBS	: ウシ胎児血清 (fetal bovine serum)
Fsk	: フォルスコリン (forskolin)
GSK-3	: グリコーゲン合成酵素 3 (glycogen synthase kinase 3)
GO	: 遺伝子オントロジー (gene ontology)
HAMMOC	: ヒドロキシ酸修飾酸化金属クロマトグラフィー (hydroxy acid-modified metal oxide chromatography)
HPLC	: 高速液体クロマトグラフィー (high performance liquid chromatography)

IAA	: ヨードアセトアミド (iodoacetamide)
MS	: 質量分析法/質量分析計 (mass spectrometry/mass spectrometer)
PKA	: cAMP 依存性 A キナーゼ (c-AMP regulated protein kinase A)
PBS	: リン酸塩緩衝生理食塩水 (phosphate buffer saline)
PVDF	: フッ化ポリビニリデン (polyvinylidene difluoride)
RNA	: リボ核酸 (ribonucleic acid)
SDB	: ポリスチレンジビニルベンゼン (poly-styrenedivinylbenzene)
SILAC	: 細胞培養液中のアミノ酸を用いた安定同位体ラベル (stable isotope labeling using amino acids in cell culture)
StageTip	: ステージ・チップ (stop-and-go-extraction tip)
TEAB	: 重炭酸トリエチルアンモニウムバッファー (triethylammonium bicarbonate buffer)
Tris	: トリス (ヒドロキシメチル) アミノメタン (tris(hydroxymethyl)aminomethane)
TSAP	: 熱失活性アルカリホスファターゼ (thermosensitive alkaline phosphatase)

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^{1,2}。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³、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⁴。

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 (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⁵⁻⁷。

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^{8,9}。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¹⁰。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⁴。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¹¹ (図 1)。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¹²⁻¹⁴。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¹⁵⁻²⁰ 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.ELM²¹、PHOSIDA²²、PhosphoSitePlus²³、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²⁴。

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²⁵。

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。

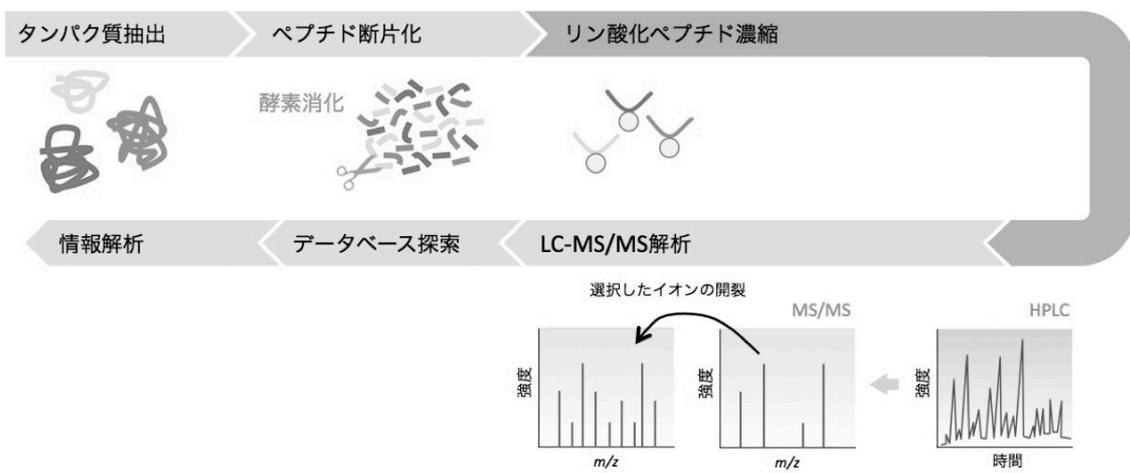


図 1: ショットガンリン酸化プロテオミクスの概略

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²⁶。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^{27,28}。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)²⁹、PhosphoSitePlus²³、また NetPhorest³⁰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^{17,25,31}。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³²。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³³⁻³⁵。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³⁵。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^{36,37}。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³⁸。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³⁹。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⁴⁰⁻⁴⁸。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得るために、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

実験手順を図2に示す。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(TiO₂)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)⁴⁹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。

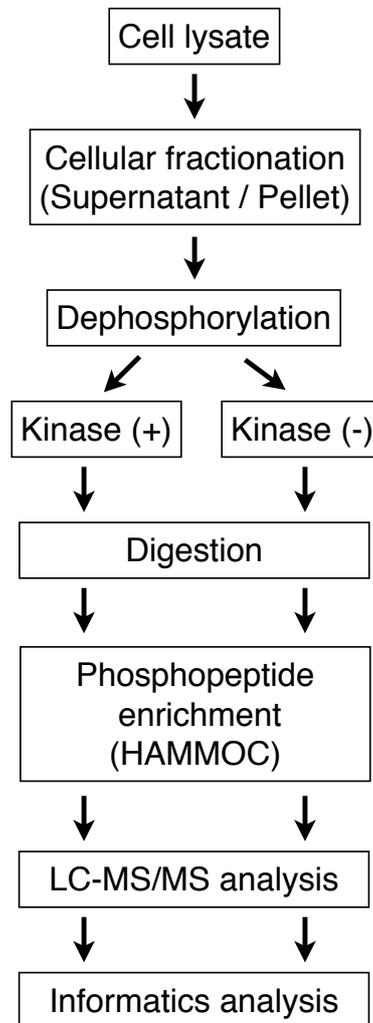


図 2: タンパク質キナーゼの基質プロファイル法

細胞からタンパク質を抽出し、脱リン酸化酵素にて脱リン酸化を行う。加熱により脱リン酸化酵素の不活性化を行った後に、リコンビナントキナーゼとタンパク質を反応させる。タンパク質を消化酵素によりペプチドに断片化した後に、濃縮したリン酸化ペプチドに対して LC-MS/MS 測定を行い、配列とリン酸化部位を同定する。

第三節 脱リン酸化反応効率の検証

脱リン酸化反応によって除去されなかったリン酸化残基は、コントロール (キナーゼ(-)) から同定される。*In vitro* キナーゼ反応によってリン酸化されたペプチドのみのリストを作成するためには、不完全な脱リン酸化反応由来のリン酸化ペプチドを除外する必要がある。そこで、まずジメチルラベル^{50,51}を用いた LC-MS/MS 分析にて脱リン酸化反応効率を定量的に評価した。

キナーゼ (-) の試料にはライト標識 (L)、キナーゼ (+) の試料にはヘビー標識 (H) を行った。この実験は三種 (PKA、ERK1、および AKT1) のキナーゼそれぞれに対して行った。全てのリン酸化ペプチドの H と L のピーク面積比 (H/L 比) を図 3 に示す。Log₂0 はキナーゼ (-) とキナーゼ (+) にてリン酸化ペプチドの量が同じであることを意味する。ピーク面積比が低いものからグラフを追っていくと、該当するリン酸化ペプチド数は log₂0 の後で局所的に下がる。ここで、定量の正確度を考慮し暫定的に二倍の違い (H/L 比=log₂1) を閾値と仮定すると、同定されたリン酸化ペプチドの内、キナーゼ (-) にてキナーゼ (+) より存在量が多いペプチドは 6.78 %であった。これらは、全体に占める、脱リン酸化反応によって除去されなかったリン酸化ペプチドの割合と考えられ、TSAP による脱リン酸化反応が高い効率で行われたことを表す。また、閾値より低い H/L 比のリン酸化ペプチドは、PKA で 27、ERK1 で 31、また AKT1 では 34 配列と同程度であったことから、キナーゼ (-) の重複にキナーゼ依存的な偏りが無いことを示している。これらの結果より、今後の実験では、非標識でキナーゼ (-) とキナーゼ (+) の LC-MS 測定を行い、同定結果を基にした比較法を採用することとした。

非標識での LC-MS 測定には複数の利点がある。まず、標識定量では一測定で H/L の二種を対象とするため、非標識試料と比較し試料中の複雑性が増大し、結果として最終的に得られる同定数が減少する傾向にある。また、標識反応の省略により、全体の実験過程が簡易化することから、将来の実験において多数のキナーゼに応用しやすくなる。最後に、同定結果のデータベース検索を行う際に、非標識同定は標識同定より考慮すべき修飾基が少なく、このことは同定数の向上に寄与する。

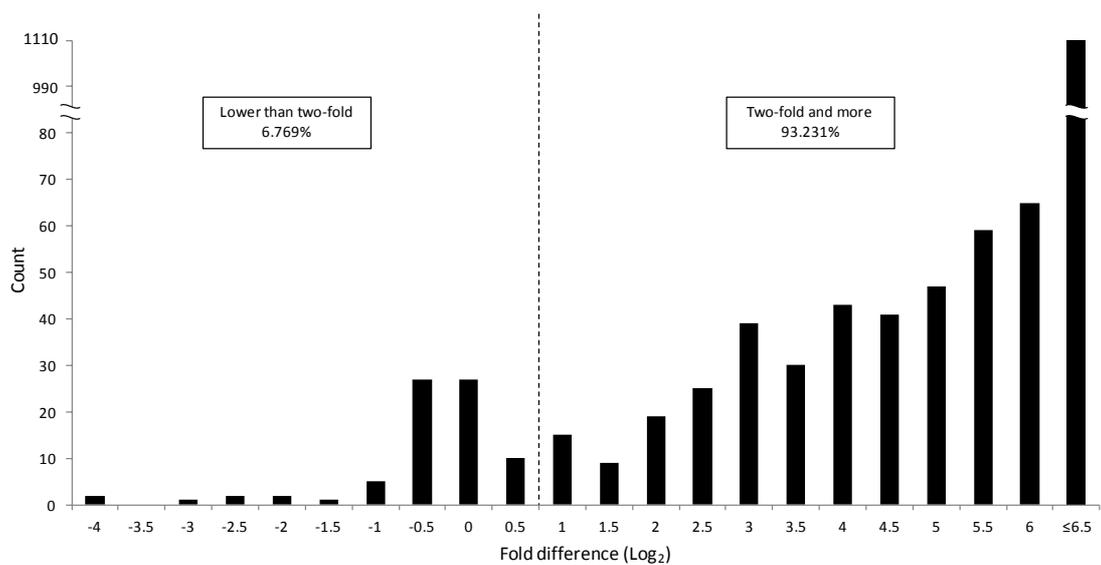


図 3: 脱リン酸化反応の定量評価

キナーゼと反応した試料とコントロール (キナーゼ (-) とキナーゼ (+)) はそれぞれ異なる質量を持つ安定同位体試薬によって標識した。それらを一つに混合した後、LC-MS/MS 解析を行い、定量情報を取得した。リン酸化ペプチドの変動比 (fold difference) は、キナーゼ (-) に対するキナーゼ (+) の面積比に底 2 の対数をとって算出した。変動比を $\log_2 0.5$ ずつに区切り、各区間に該当するリン酸化ペプチドの数を縦軸に表す。

第四節 リン酸化部位同定

In vitro 基質の同定数向上のために、以下の取り組みを行った。まず、試料に含まれるプロテオームの複雑性を低減させるため、タンパク質抽出の際に分画を行った。ノニデット P40 を用いて細胞膜を破碎した後、遠心分離にて上清とペレットの分画を行い、それぞれから抽出したタンパク質を実験に用いた。各画分から同定されたリン酸化部位の重複は、PKA で 29.71%、ERK1 で 27.63%、また、AKT1 で 24.92%と比較的低かった (図 4A)。遺伝子オントロジー (GO) ターム解析の結果から、細胞質タンパク質は上清分画に、核タンパク質がペレット画分に濃縮されており (表 1, 2)、各画分に存在するプロテオームは大きく異なることが示された。このことは、画分によりプロテオームの複雑性を緩和し、リン酸化同定数の向上に寄与したことを示唆している。二つ目に、各画分の試料に対して LC-MS 測定を 5 回繰り返し行い、結果として各キナーゼあたり計 10 回の MS 測定データを獲得した。複雑性の高い試料に対してデータ依存的なスキャン取得法 (DDA: data dependent acquisition) モードでの MS 分析を行った場合、現状のスキャン速度では MS² スキャンのトリガーとなるプレカーサーイオンの選択は確率的であり、一回の解析で十分な網羅性を示す結果を得ることは難しい。そのため、同定されるペプチド配列は各解析で異なり、解析を繰り返すことによる同定数の向上が見込まれる⁵²。実際に今回の複数回の解析結果を統合することにより、合計のリン酸化ペプチド数は増大した (図 4B)。

ここで、同定結果を基にキナーゼ (-) とキナーゼ (+) の比較を行ったところ、キナーゼ (-) で同定されたリン酸化ペプチドの多くは、キナーゼ (+) の同定結果と重複していた (上清, PKA: 133 (44.63%)、ERK1: 165 (55.37%)、AKT1: 181 (60.74%); ペレット, PKA: 1,099 (53.51%)、ERK1: 1,176 (57.25%)、AKT1: 1,466 (71.37%))。ピーク面積値の定量的な比較の結果、この重複中には、キナーゼ (-) と比べキナーゼ (+) で増加が見られたリン酸化ペプチドも存在した。しかし、ここでは第三節で述べた同定を基準とするより厳格なルールを採用しているため、上記の様な定量変化があったペプチドも基質リストからは除かれている。

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 (表 3)。これらのリン酸化ペプチドに対して、MS² 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

抛に、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 (表 4)。*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⁴⁰。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⁴¹。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⁵³。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。

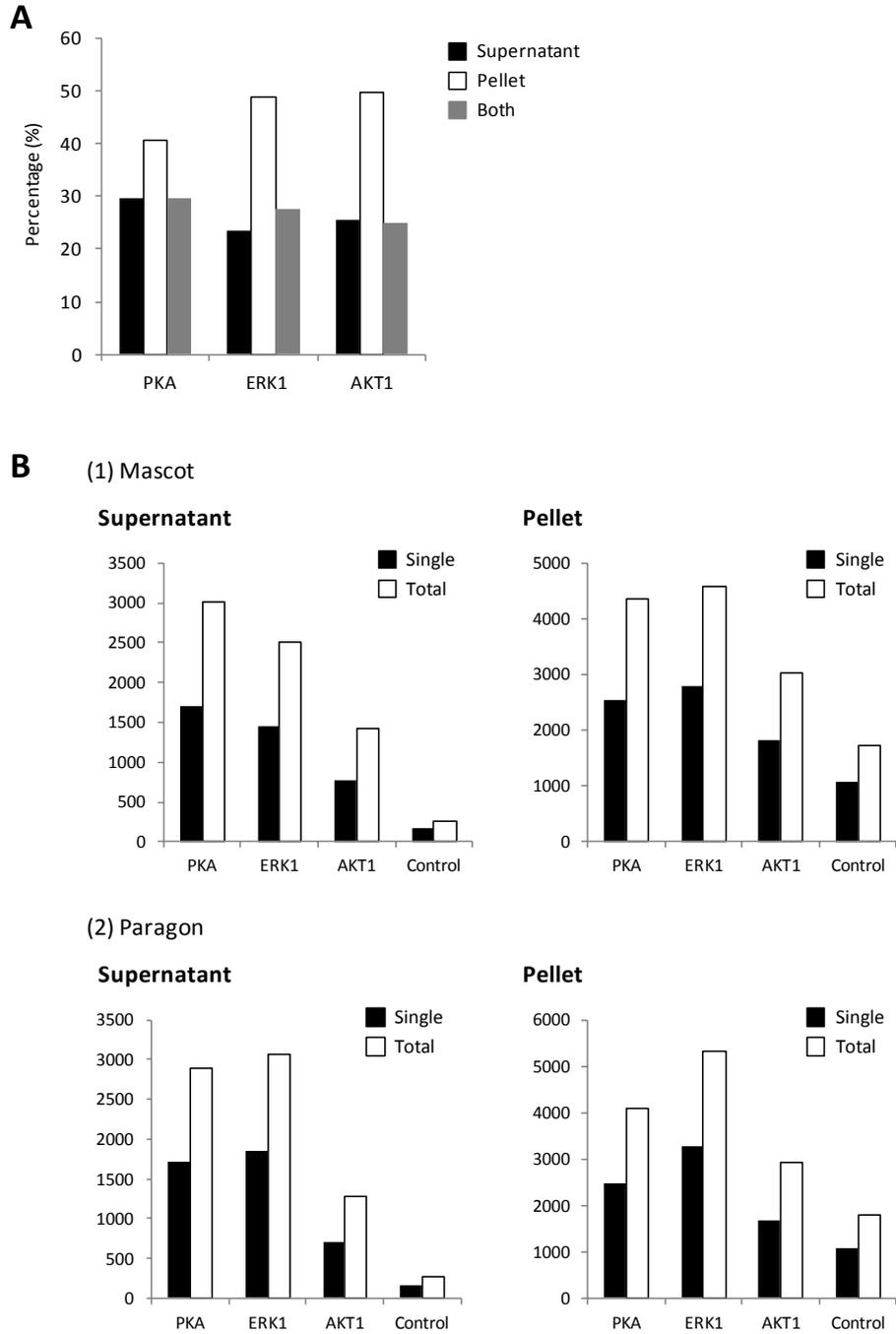


図 4: リン酸化ペプチド同定数

A. 各画分で同定されたリン酸化部位数の比率 (*black*: 上清のみ, *white*: ペレットのみ, *gray*: 両方) **B.** リン酸化ペプチド数の向上: 各画分 (上清/ペレット) から得られた試料を LC-MS/MS にて 5 回測定し、同定数向上のため 2 種の探索エンジン (Mascot/Paragon) を用いてデータベース検索を行った。LC-MS/MS 測定を繰り返すことで同定数が向上した (*black*: 一回の解析での同定数, *white*: 合計数)。

表 1: 上清画分に濃縮されている GO ターム上位 20 種

Term	Count	PValue
GO:0005829~cytosol	189	6.91E-43
GO:0005856~cytoskeleton	120	1.48E-09
GO:0043232~intracellular non-membrane-bounded organelle	191	6.14E-09
GO:0043228~non-membrane-bounded organelle	191	6.14E-09
GO:0015630~microtubule cytoskeleton	61	9.02E-09
GO:0005739~mitochondrion	95	9.15E-08
GO:0005852~eukaryotic translation initiation factor 3 complex	9	1.82E-07
GO:0031967~organelle envelope	60	1.46E-06
GO:0031975~envelope	60	1.63E-06
GO:0019898~extrinsic to membrane	51	1.69E-06
GO:0048770~pigment granule	18	1.71E-06
GO:0042470~melanosome	18	1.71E-06
GO:0044445~cytosolic part	24	2.01E-06
GO:0005625~soluble fraction	37	2.88E-06
GO:0044430~cytoskeletal part	78	1.76E-05
GO:0031090~organelle membrane	87	1.77E-05
GO:0031252~cell leading edge	21	1.78E-05
GO:0000267~cell fraction	86	1.94E-05
GO:0005768~endosome	35	2.11E-05
GO:0005874~microtubule	30	1.20E-04
GO:0044429~mitochondrial part	52	1.23E-04
GO:0005938~cell cortex	20	1.28E-04
GO:0044455~mitochondrial membrane part	18	1.67E-04
GO:0031982~vesicle	56	2.03E-04
GO:0033279~ribosomal subunit	18	2.23E-04
GO:0048471~perinuclear region of cytoplasm	30	2.84E-04
GO:0005740~mitochondrial envelope	39	3.02E-04
GO:0012505~endomembrane system	62	3.69E-04
GO:0005813~centrosome	25	3.78E-04
GO:0005815~microtubule organizing center	27	4.22E-04

表 2: ペレット画分に濃縮されている GO ターム上位 20 種

Term	Count	PValue
GO:0031981~nuclear lumen	321	1.54E-124
GO:0070013~intracellular organelle lumen	348	5.69E-120
GO:0031974~membrane-enclosed lumen	355	6.99E-120
GO:0043233~organelle lumen	349	1.36E-117
GO:0043232~intracellular non-membrane-bounded organelle	351	5.90E-72
GO:0043228~non-membrane-bounded organelle	351	5.90E-72
GO:0005654~nucleoplasm	194	1.72E-68
GO:0005730~nucleolus	165	4.45E-62
GO:0044451~nucleoplasm part	125	9.60E-44
GO:0005694~chromosome	110	6.99E-41
GO:0030529~ribonucleoprotein complex	114	5.90E-39
GO:0005681~spliceosome	58	4.86E-37
GO:0044427~chromosomal part	88	7.45E-31
GO:0016604~nuclear body	48	3.28E-21
GO:0000785~chromatin	49	1.31E-18
GO:0000228~nuclear chromosome	44	1.46E-18
GO:0016585~chromatin remodeling complex	27	2.13E-15
GO:0016607~nuclear speck	32	2.43E-15
GO:0000793~condensed chromosome	35	8.80E-15
GO:0000775~chromosome, centromeric region	31	3.77E-12
GO:0044454~nuclear chromosome part	30	1.38E-11
GO:0000123~histone acetyltransferase complex	18	6.06E-10
GO:0000792~heterochromatin	16	1.94E-09
GO:0030530~heterogeneous nuclear ribonucleoprotein complex	11	3.39E-09
GO:0000776~kinetochore	21	4.00E-09
GO:0000790~nuclear chromatin	17	2.49E-08
GO:0034399~nuclear periphery	17	1.24E-07
GO:0017053~transcriptional repressor complex	14	1.67E-07
GO:0016363~nuclear matrix	16	2.28E-07
GO:0030532~small nuclear ribonucleoprotein complex	11	2.40E-07

表 3: リン酸化ペプチド数

Kinase	Total	# of phosphate			
		1	2	3	4
PKA	5,625	5,042	564	18	1
ERK1	6,455	4,616	1,681	153	5
AKT1	2,785	2,499	270	16	0

ペプチド数は修飾情報と配列に基づいて冗長を除いた数を用いた

表 4: リン酸化部位数

Kinase	Total	Phosphorylated residue		
		S	T	Y
PKA	3,585	3,147	393	45
ERK1	4,347	2,947	1,347	53
AKT1	1,778	1,388	378	12

第五節 リン酸化部位周辺のアミノ酸出現頻度

キナーゼファミリーは配列相同性によって分類されていることから⁵⁴、同じファミリーに属するキナーゼは類似したモチーフ配列を持つ。今回使用したキナーゼでは、PKA と AKT1 は同じ AGC ファミリー (PKA、PKG、PKC ファミリーを含む) に所属し、ERK1 は CMGC ファミリー (CDK、MAPK、GSK3、CLK ファミリーを含む) とされている。

リン酸化部位前後 7 残基に出現する特徴的なアミノ酸を観察するため、各位置でのアミノ酸出現頻度を計算した。中心とする残基のリン酸化以外の影響をできるだけ排除する為に、今後の配列解析ではリン酸基の付与が一カ所であるモノリン酸化ペプチド情報のみを利用した (PKA: 2,840 配列、ERK1: 2,944 配列、AKT1: 1,385 配列)。図 5 のアミノ酸頻度プロットで、記号の縦幅は各残基の各位置での頻度を意味している。

ヒトキナーゼには、大きく大別して Y のみをリン酸化するチロシンキナーゼ (Y キナーゼ) と、S、T の二残基がリン酸基受容残基 (phosphoacceptor) となるセリン-トレオニンキナーゼ (ST キナーゼ) が存在する。また、ST キナーゼの中にも、S あるいは T のどちらかにより強い選択性を示すキナーゼが報告されている^{55,56}。今回の結果では、全てのキナーゼが基本的に T に比べて S に強い選択性を示しているものの、その選択性の強さに偏りが見られた。特に、ERK1 はその他のキナーゼと比較するとより強く T をリン酸基受容残基として選択する (図 5, 表 4, PKA=S: 87.78%, T: 10.96%; ERK1=S: 67.79%, T: 30.99%; AKT1=S: 78.07%, T: 21.26%)。χ² 検定によって、ファミリー分類に関わらず、各キナーゼの選択性の偏りには有意な差があることが支持された (PKA と ERK1: $p < 2.2e-16$, ERK1 と AKT1: $p = 8.81e-15$, PKA と AKT1: $p < 2.2e-16$)。これまでも酵母のキナーゼを用いたペプチドアレイで S/T 選択性が報告されており、オルソログのキナーゼと今回の選択性の傾向は一貫した結果であった^{35,56}。

周辺のアミノ酸に関しては、AGC ファミリーの PKA と AKT1 は塩基性のアミノ酸選択的であり、CMGC ファミリーの ERK1 はプロリン指向性があるという、先行研究と一致した結果が得られた (図 5 左)³⁵。リン酸化部位のうち、一つのキナーゼのみから同定された部位は PKA で 2,433 カ所、ERK1 で 3,904 カ所、また AKT1 で 695 カ所であった (図 6)。ERK1 と比較すると PKA と AKT1 の重複率は高く、これは配列選択性の類似に起因していると考えた。PhosphoSitePlus から得られた既知の *in vitro* 基質情報を用いて、上述と同様に

アミノ酸頻度を計算したところ、全体的なアミノ酸の選択性は本実験によって得られた結果と類似した傾向にあった (図 5 右)。しかし、本実験によって得られた大規模データでの各残基の割合は従来と異なり、また、新たな残基の出現も見られた。これらの基質プロファイルの変化は、大規模データ獲得による新たな基質選択性の同定が可能であることを示す。

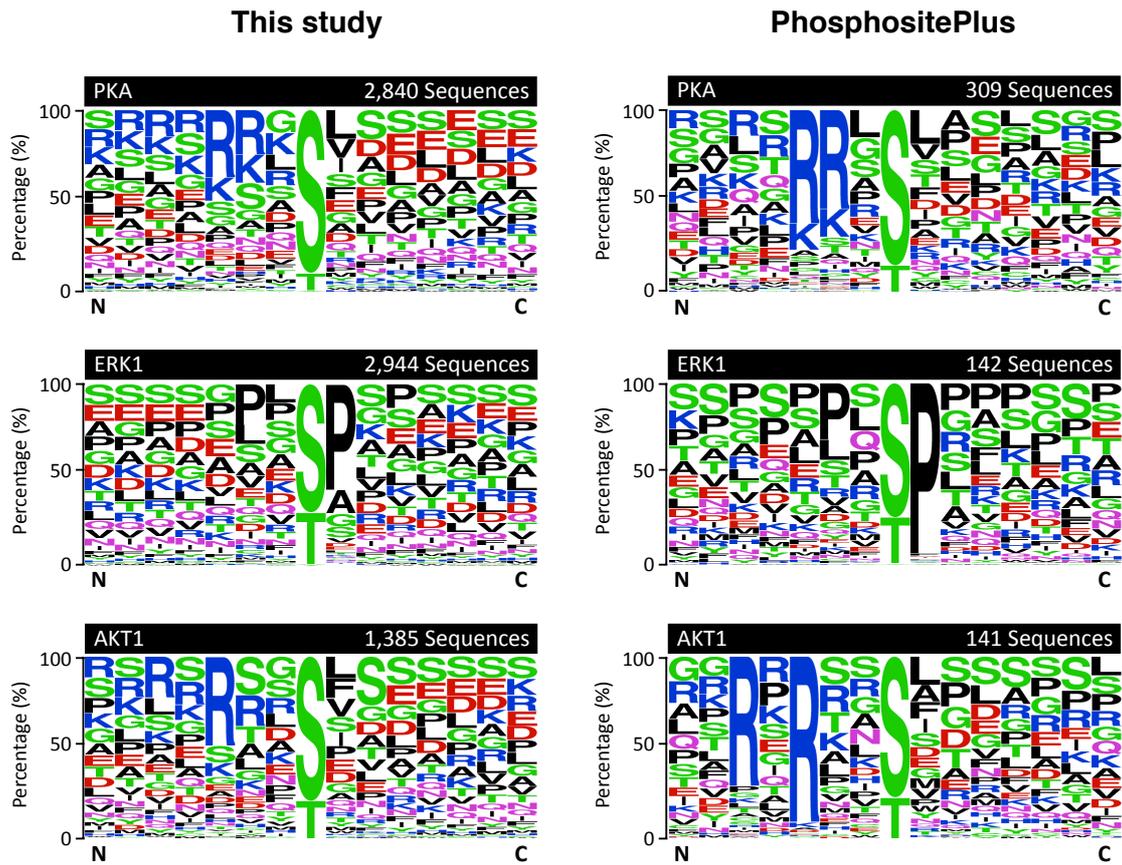


図 5: リン酸化部位周辺アミノ酸の頻度プロット

リン酸化部位を中心とし、前後7残基のアミノ酸出現頻度を記号の高さとして描画した。本研究によって得られたリン酸化部位と、PhosphoSitePlus に登録のある *in vitro* リン酸化部位の周辺配列を比較した。

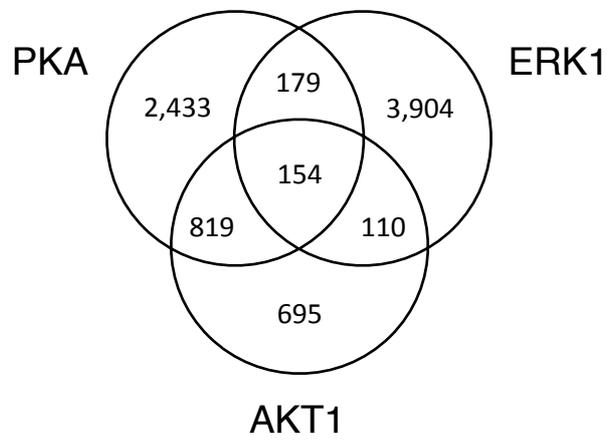


図 6: キナーゼ間でのリン酸化ペプチド数の重複

第六節 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、平均的な傾向を示した頻度プロットから、既存のモチーフ抽出ソフトウェアである motif-X⁵⁷ を用いて独立したモチーフ配列へと分解した (図 7AB)。

Motif-X 解析では、修飾を受ける残基を一種指定し (S または T)、解析対象とする前景 (foreground) と背景 (background) の二種のデータセットを用意する必要がある。ここでは、前景としては該当する残基のリン酸化部位周辺配列を、背景にはタンパク質配列データベースから指定残基の周辺配列を用いた。その後、この二つのデータセットを比較し、前景に特徴的に見られるアミノ酸の組み合わせを取り出し、モチーフ配列とする。ひとたびモチーフ配列が決定すると、そのモチーフに該当する配列を前景と背景のセットから取り除き、再度二つのデータセットの比較およびモチーフ配列抽出を行う。この過程を、モチーフ配列が全て抽出できるまで繰り返す。各モチーフ配列に対して、モチーフ配列に適合した配列の前景中での割合を背景中での割合で割って、増倍比 (fold increase) が算出される。本手法で得られたリン酸化部位を用いてモチーフ解析を行ったところ、31 種、21 種および 18 種のモチーフ配列が抽出された (増倍比 \geq 3)。これらを既知情報由来のモチーフ配列と比較したところ、本実験によるモチーフ配列には、既知情報で得られたモチーフ全てに加え、新規モチーフも含まれていた (図 7A-D)。

ここで、既知情報と本研究で共通に同定されたモチーフ配列に関して、増倍比の比較を行った (図 7E)。増倍比は、前景および背景の数はモチーフ配列が決定するたびに変わる為、増倍比はインプットの前景データの種類や数のみならず、モチーフ抽出の順番によっても影響を受ける。各モチーフ配列に適合する配列の比率が両前景データ内で一様であれば、増倍比に高い相関が見られることになる。図 7E に示すように既知情報と本研究に共有して同定されたモチーフの増倍比には高い相関が認められ、本研究で得られた基質群においてもこれまでのデータ間に見られていた相対的な配列選択性が維持されていることが示された。また、今回のモチーフ配列抽出には、図 5 で用いた PKA、ERK1、および AKT1 の全配列データの内、それぞれ 81.20% (2,306 配列)、73.23% (2,156 配列)、および 72.71% (1,007 配列) が利用された。このことは大規模データで同定されたリン酸化部位の大部分は周囲にモチーフ配列を含有することを意味する。つまり、アミノ酸頻度プロットで見られた多様

なアミノ酸の増加は、非特異的なリン酸化部位の混入ではなく、情報量の増加に起因することが裏付けられた。

実験報告の多い PKA に関して、先行研究で議論されたキナーゼ選択性との相違を考察した。ペプチドアレイを用いた先行研究によると、リン酸化部位の N 末側の塩基性残基と +1 の位置に疎水性残基が位置することは確認されていた^{33,41}。この知見に加えて、*in vitro* 実験系を用いた Douglas らの報告では-1 にグリシン (G) が位置することも言及されている⁴¹。これらの知見と同じく、本研究のモチーフ解析からは、N 末側の R または K と組み合わせたモチーフが同定されている (RXXXG[pS]: 増倍比=8.88, RXXG[pS]: 増倍比=7.29, RG[pS]: 増倍比=6.52)。以上をふまえると、今回の結果は過去の実験結果によって得られたキナーゼ選択性と一致した結果であった。

上記のモチーフ配列に加えて、本研究データからは、これまでの限られた基質プロフィールからは同定できなかった、より多様なモチーフ配列が多数同定された。

(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。

PKA と AKT1 それぞれに特異的なモチーフを検出する為には、より直接的に二つのリン酸化部位周辺配列を比較する必要がある。そのため、motif-X を用いた追加の比較解析を行い、各キナーゼの選択性を生み出す特徴を抽出した。具体的には、PKA 選択的なモチーフ配列を検出するために、AKT1 のリン酸化部位周辺配列を背景、PKA のリン酸化部位周辺配列を前景として用いた。また、AKT1 に特徴的な配列を抽出する為には、逆の操作

を行い、これらの解析により取り出された特徴を「識別要素」とした (表 5)。次に、通常のもチーフ解析結果 (図 7A) に識別要素を照らし合わせ、該当したもチーフ配列を、二者の識別に利用できる「キナーゼ特異もチーフ配列」と推定した。この評価の信頼性を確かめる為に、各特異もチーフ配列に該当するリン酸化部位が、PKA あるいは AKT1 の配列中ほどの程度存在するのか検証した。その結果、該当する PKA 特異もチーフ 12 種のうち 11 種のもチーフ配列が、また、該当する AKT1 特異もチーフの 8 種全てが、期待した結果を示した。つまり、該当するキナーゼの配列中でこれらもチーフ配列に適合するリン酸化部位の割合は、もう一方の配列データ中よりも高かった (表 6)。この特徴的な PKA 特異もチーフの例として、KKX[pS] というもチーフは PKA のリン酸化部位で 76 配列中に存在したが、AKT1 のリン酸化部位データ中では 2 配列にしか適合しなかった。また、別の PKA 特異もチーフである KK[pS] は PKA では 72 配列、AKT1 では 5 配列に照合し、高い PKA 選択性を示した。逆に、唯一の例外となったもチーフ配列は、-3 と -2 に R を持つ RRX[pS] であった。通常のもチーフ解析の内、大規模基質プロファイルでは PKA、AKT1 の両方で、このもチーフ配列は同定された (図 7A)。一方、既知リン酸化部位のみを利用した同様の解析では、このもチーフは PKA のみでの同定であった (図 7C)。この結果に従いこのもチーフをキナーゼ予測に利用した場合、AKT1 基質を PKA 基質として評価することにより、偽陽性が生じる。最後に、表 5 に示した識別要素には、該当するもチーフ配列を持たない配列も存在した。その理由としては、これらの特徴を持つリン酸化部位の同定数が、もチーフ配列抽出に利用されるほど多数でないことがある。しかし、PKX[pS] や KX[pS]XXXXXXXXP の様に、AKT1 には対象リン酸化部位が存在せず、PKA のみに高い特異性を示す要素も含まれており、これらは PKA と AKT1 の基質を識別する際に有力な手がかりとなる。

基質にタンパク質を使用した本実験の利点として、上述の PKA や AKT1 の様な好塩基性もチーフ配列の同定数向上が期待できる。先行研究の中には、タンパク質をペプチドに消化した後にキナーゼ反応を行う手法の報告がある^{42,44,55}。この場合、消化に用いた酵素の選択によっては、もチーフ配列を切断し、得られる基質プロファイルに影響を及ぼす恐れが有る。例えば、ショットガンプロテオミクス解析でよく用いられる消化酵素である lys-C とトリプシンの酵素消化は、リジンおよびアルギニンの C 末側を開裂する。これら塩基性アミノ酸は上述の通りもチーフ配列中に頻出するアミノ酸である。実際に、lys-C と

トリプシンで消化したペプチドを基質に用いて PKA の *in vitro* キナーゼ反応を行い、LC-MS 測定を行ったところ、基質をタンパク質とした時より得られるモチーフの種類は減少した (図 8)。

(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 (図 7BD)。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 (表 3)。モノリン酸化ペプチドの比率は PKA、ERK1、また AKT1 にてそれぞれ、89.64%、71.51%、および 89.73%であった。 χ^2 検定を用いた統計解析から、ERK1 はマルチリン酸化ペプチドに対して有意な偏り ($p < 0.01$) が見られた一方で、PKA と AKT1 ではその差が認められなかった (PKA vs ERK1: $p < 2.2e-16$, ERK1 vs AKT1: $p < 2.2e-16$, PKA vs AKT1: $p = 0.9228$)。

この原因として、以下の二つを考えた。一つは、ERK1 がリン酸基を既に持つ基質に対して、より高い親和性を有すること、もう一つは、ERK1 が標的とするモチーフ配列は、そもそもタンパク質配列上に局在していることである。後者の可能性を検証する為に、ERK1 が標的とする部位の出現頻度を計算した。典型的な ERK1 のモチーフの “[pS]P” をもつリン酸化部位を利用し、本実験から得られた全ての [pS]P 間の残基数を距離として計算した。また、リン酸化部位を同一タンパク質中でランダムに別の SP へと再配置し直したデータをコントロールとして作成し、実データと比較した。その結果、実際のリン酸化部位は、ランダムなデータ中と比べ、互いに隣接して存在することが分かった (図 7F)。つまり、ERK1 によるリン酸化は、SP 配列の局在している部分に選択的に起きており、リン酸基の存在が更なるリン酸化をリクルートする可能性が支持された。同様の結果は、TP のみ、または SP と TP を対象配列とした解析からも得られた (図 9)。

上記の仮説を更に肯定する結果として、モノリン酸化ペプチドのみを用いた

ERK1 のモチーフ配列解析結果から、グルタミン酸 (E) を C 末側にもつモチーフ配列 (EPX[pS]P) が高い倍増比 (40.64) で検出された (図 7B)。アスパラギン酸 (D) と E は酸性残基であり、生化学的な実験でリン酸基を模倣した残基として利用される。このモチーフ配列でも示される様に、もし ERK1 が負に帯電した残基を選択的に認識するとしたら、リン酸基も同じ様に ERK1 の基質親和性を高める為に機能していると考えられる。

実際に、同じ CMGC ファミリーのグリコーゲン合成酵素 3 (GSK3) では、標的部位のリン酸化に先立ち、4 残基離れた位置にリン酸基が付加された S または T を必要とする。この GSK3 によるリン酸化の引き金 (priming) となる反応自体は、GSK3 以外のキナーゼによって触媒されるが、GSK3 の活性部位の隣の正電荷のポケットに結合することで、GSK3 と基質間の親和性を高めることが知られている⁵⁸。ERK1 リン酸基選択性に関するより詳細な議論を行うには、更なる実験が必要である。最後に、マルチリン酸化ペプチドは低いイオン化効率のため、通常の LC-MS 分析では検出されにくい。今回の大規模実験により多数のマルチリン酸化ペプチドが同定されたことにより、この様な議論が初めて可能となった。

A

N	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	PKA	AKT1
--	R	--	--	--	S	--	--	--	--	--	--	--	--	--	3.01	3.46	
--	K	--	--	--	S	--	--	--	--	--	--	--	--	--	3.10	n.a.	
--	R	--	--	--	S	--	--	--	--	--	--	--	--	--	3.10	3.31	
--	R	--	--	--	T	--	--	--	--	--	--	--	--	--	3.83	n.a.	
--	R	--	R	--	S	--	--	--	--	--	--	--	--	--	10.77	25.18	
--	R	--	R	--	T	--	--	--	--	--	--	--	--	--	n.a.	35.74	
--	--	--	K	--	S	--	--	--	--	--	--	--	--	--	3.21	n.a.	
--	--	--	R	--	S	--	--	--	--	--	--	--	--	--	4.70	4.95	
--	--	--	R	--	T	--	--	--	--	--	--	--	--	--	3.42	6.32	
--	--	--	K	K	S	--	--	--	--	--	--	--	--	--	12.86	n.a.	
--	--	--	R	K	S	--	--	--	--	--	--	--	--	--	15.12	n.a.	
--	--	--	R	K	T	--	--	--	--	--	--	--	--	--	30.22	n.a.	
--	--	--	R	N	S	--	--	--	--	--	--	--	--	--	12.44	n.a.	
--	--	--	K	R	S	--	--	--	--	--	--	--	--	--	11.69	n.a.	
--	--	--	R	R	S	--	--	--	--	--	--	--	--	--	12.31	18.58	
--	--	--	R	R	T	--	--	--	--	--	--	--	--	--	41.51	36.79	
--	--	--	K	S	S	--	--	--	--	--	--	--	--	--	9.11	n.a.	
--	--	--	R	S	S	--	--	--	--	--	--	--	--	--	9.83	13.41	
--	--	--	R	T	S	--	--	--	--	--	--	--	--	--	n.a.	16.55	
--	--	--	K	S	--	--	--	--	--	--	--	--	--	--	3.71	n.a.	
--	--	--	R	S	--	--	--	--	--	--	--	--	--	--	4.14	3.11	
--	--	--	R	T	--	--	--	--	--	--	--	--	--	--	3.89	3.61	
--	R	--	--	G	S	--	--	--	--	--	--	--	--	--	8.88	n.a.	
--	R	--	--	G	S	--	--	--	--	--	--	--	--	--	7.29	n.a.	
--	--	--	R	G	S	--	--	--	--	--	--	--	--	--	6.52	n.a.	
--	--	--	K	K	S	--	--	--	--	--	--	--	--	--	17.33	n.a.	
--	--	--	R	K	S	--	--	--	--	--	--	--	--	--	14.43	n.a.	
--	--	--	R	R	S	--	--	--	--	--	--	--	--	--	8.86	n.a.	
--	--	--	G	S	--	--	--	--	--	--	--	--	--	--	2.65	3.15	
--	--	--	K	S	--	--	--	--	--	--	--	--	--	--	3.25	n.a.	
--	--	--	R	S	V	I	--	--	--	--	--	--	--	--	11.33	n.a.	
--	--	--	K	S	T	I	--	--	--	--	--	--	--	--	13.98	n.a.	
--	--	--	R	S	I	I	--	--	--	--	--	--	--	--	15.85	n.a.	
--	--	--	R	S	L	--	--	--	--	--	--	--	--	--	10.69	n.a.	
--	--	--	R	S	S	--	--	--	--	--	--	--	--	--	n.a.	9.43	
--	--	--	R	T	S	S	--	--	--	--	--	--	--	--	n.a.	36.21	
--	--	--	R	S	S	--	--	--	--	--	--	--	--	--	n.a.	18.34	
--	--	--	S	F	--	--	--	--	--	--	--	--	--	--	n.a.	3.96	
--	--	--	S	D	--	--	--	--	--	--	--	--	--	--	2.86	3.94	

C

N	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	PKA	AKT1
--	R	--	R	--	S	--	--	--	--	--	--	--	--	--	n.a.	112.74	
--	R	--	R	--	T	--	--	--	--	--	--	--	--	--	n.a.	189.44	
--	--	--	R	--	S	--	--	--	--	--	--	--	--	--	9.04	13.04	
--	--	--	R	--	T	--	--	--	--	--	--	--	--	--	12.39	n.a.	
--	--	--	R	K	S	--	--	--	--	--	--	--	--	--	42.28	n.a.	
--	--	--	R	R	S	--	--	--	--	--	--	--	--	--	68.35	n.a.	
--	--	--	R	--	S	--	--	--	--	--	--	--	--	--	7.39	n.a.	

B

N	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	ERK1
D	--	--	--	--	S	P	--	--	--	--	--	--	--	--	16.23	
--	K	--	--	--	T	P	--	--	--	--	--	--	--	--	20.56	
--	--	E	P	--	S	P	--	--	--	--	--	--	--	--	40.64	
--	--	P	--	S	P	--	--	--	--	--	--	--	--	--	14.70	
--	--	P	--	T	P	--	--	--	--	--	--	--	--	--	16.63	
--	--	P	--	S	A	--	--	--	--	--	--	--	--	--	14.41	
--	--	P	--	T	A	--	--	--	--	--	--	--	--	--	15.32	
--	--	L	--	S	A	--	--	--	--	--	--	--	--	--	10.20	
--	--	--	K	S	A	--	--	--	--	--	--	--	--	--	14.15	
--	--	--	--	S	A	--	--	--	--	--	--	--	--	--	3.63	
--	--	--	--	T	A	--	--	--	--	--	--	--	--	--	3.37	
--	--	--	--	S	P	--	--	--	--	--	--	--	--	--	7.32	
--	--	--	--	T	P	--	--	--	--	--	--	--	--	--	7.12	
--	--	--	--	S	G	K	--	--	--	--	--	--	--	--	10.92	
--	--	--	--	T	P	K	--	--	--	--	--	--	--	--	20.23	
--	--	--	--	S	S	P	--	--	--	--	--	--	--	--	7.40	
--	--	--	--	S	P	--	K	--	--	--	--	--	--	--	16.73	
--	--	--	--	S	P	--	K	--	--	--	--	--	--	--	16.70	
--	--	--	--	T	P	--	K	--	--	--	--	--	--	--	22.07	
--	--	--	--	S	P	--	E	--	--	--	--	--	--	--	15.15	
--	--	--	--	S	P	--	K	--	--	--	--	--	--	--	16.04	

D

N	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	ERK1
--	--	--	--	--	S	P	--	--	--	--	--	--	--	--	12.12	
--	--	--	--	--	T	P	--	--	--	--	--	--	--	--	13.75	

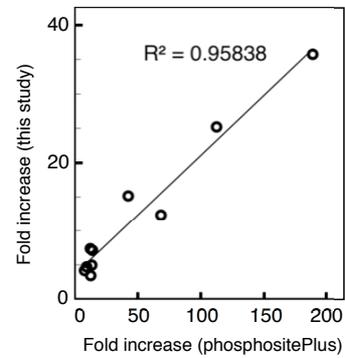
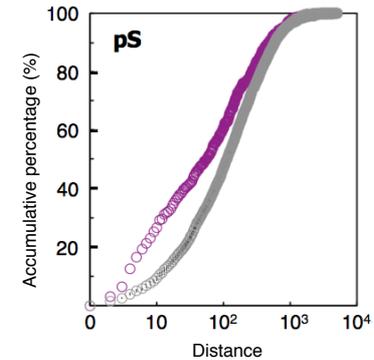
E**F**

図 7: motif-X によるモチーフ解析

ABCD. Motif-X により得られたモチーフ配列 (増倍比 ≥ 3) 結果を本研究 (**A:** PKA と AKT1, **B:** ERK1) と既知リン酸化部位 **C:** PKA と AKT1, **D:** ERK1) から得られた配列で比較した。リン酸化部位を中心とし、固定残基が N 末側および C 末側に見られたモチーフ (*red*)、末側のみ (*green*) また、C 末側のみ (*blue*) に分類し表示する。配列中の横棒 (-) は特定の残基が指定され無かった場所を意味する。数字は増倍比を示し、グレースケールは同一のデータセット内で標準化したスケールに準ずる (*eg.* 本研究で同定した基質、または、既知基質)。 **E.** 本研究と既知基質から共通に同定された PKA、ERK1、および AKT1 モチーフ配列の増倍比。 **F.** +1 位置にプロリンを持つリン酸化セリン (*i.g.*, [pS]P) のペア間のアミノ酸の数を距離として計算した。距離に対する累積数の割合をプロットし、リン酸化部位をランダムに再配置して作成したコントロールと比較した (*purple:* 実データ, *gray:* ランダム)。

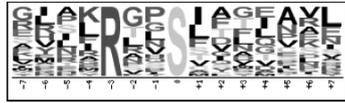
表 5: PKA および AKT1 の識別要素

#	Motif	Motif Score	Foreground Matches	Foreground Size	Background Matches	Background Size	Fold Increase
1: Foreground: PKA, Background AKT1, Centered residue: S							
1R.SR.....	26.86	22	2,546	1	1,096	9.47
2R.S..E....	25.88	64	2,524	8	1,095	3.47
3KK.S.....	32	76	2,460	2	1,087	16.79
4PK.S.....	32	29	2,384	0	1,085	Inf
5R.SG.....	23.69	36	2,355	4	1,085	4.15
6K.S.....P	32	25	2,319	0	1,081	Inf
7	.L...K.S.....	25.78	25	2,294	1	1,081	11.78
8KSL.....	31.48	62	2,269	5	1,080	5.9
9K.SA.....	23.32	20	2,207	1	1,075	9.74
10K.S.....	16	212	2,187	50	1,074	2.08
11R.S.....	16	483	1,975	167	1,024	1.5
12Q.SL.....	23.13	23	1,492	1	857	13.21
13KN.S.....	15.61	22	1,469	2	856	6.41
14RD.S.....	23.62	54	1,447	5	854	6.37
15KG.S.....	14.17	21	1,393	2	849	6.4
16Q.S.....	6.37	84	1,372	29	847	1.79
17	..G....S.....	6.1	77	1,288	27	818	1.81
2: Foreground: PKA, Background AKT1, Centered residue: T							
1K.T.....	6.87	40	271	17	285	2.47
3: Foreground: AKT1, Background PKA, Centered residue: S							
1S.S.....	16	312	1,096	418	2,546	1.73
2SP.....	16	68	784	49	2,128	3.77
3T.S.....	13.75	77	716	84	2,079	2.66
4	..R.R..S.....	22.47	92	639	103	1,995	2.79
5S.S.....	8.11	121	547	250	1,892	1.67
6L.S.....	7.81	47	426	73	1,642	2.48
7A.S.....	6.14	47	379	90	1,569	2.16
4: Foreground: AKT1, Background PKA, Centered residue: T							
1	..R.R..T.....	14.33	42	285	12	271	3.33

表 6: キナーゼ特異モチーフ配列

Motif sequence	#PKA	#AKT	%PKA	%AKT	Fold
Motif sequences from PKA					
....KK.S.....	76	2	2.99	0.18	16.36
.....KKS.....	72	5	2.83	0.46	6.20
.....K.S.....	400	54	15.71	4.93	3.19
.....RKS.....	86	17	3.38	1.55	2.18
....KR.S.....	77	16	3.02	1.46	2.07
....RK.T.....	21	12	7.75	4.21	1.84
....RK.S.....	131	32	5.15	2.92	1.76
.....RGS.....	55	15	2.16	1.37	1.58
.....R.S.....	617	181	24.23	16.51	1.47
.....RRS.....	75	23	2.95	2.10	1.40
.....R.SV.....	67	22	2.63	2.01	1.31
....RR.S.....	149	85	5.85	7.76	0.75
Motif sequences from AKT1					
..R.R..T.....	12	42	4.43	14.74	3.33
.....SP.....	59	83	2.32	7.57	3.27
....RT.S.....	56	66	2.20	6.02	2.74
..R.R..S.....	144	149	5.66	13.59	2.40
....R..S.S.....	169	150	6.64	13.69	2.06
....RS.S.....	154	120	6.05	10.95	1.81
.....S.S.....	418	312	16.42	28.47	1.73
.....R.S.S.....	99	61	3.89	5.57	1.43

Peptide substrate



Protein substrate

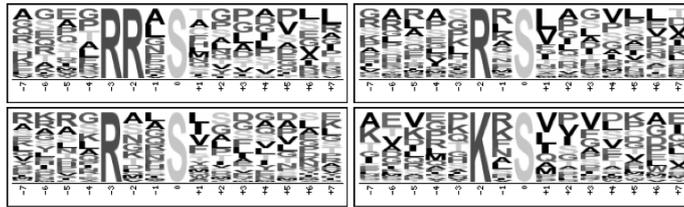


図 8: 基質フォーム別のモチーフ配列解析の結果

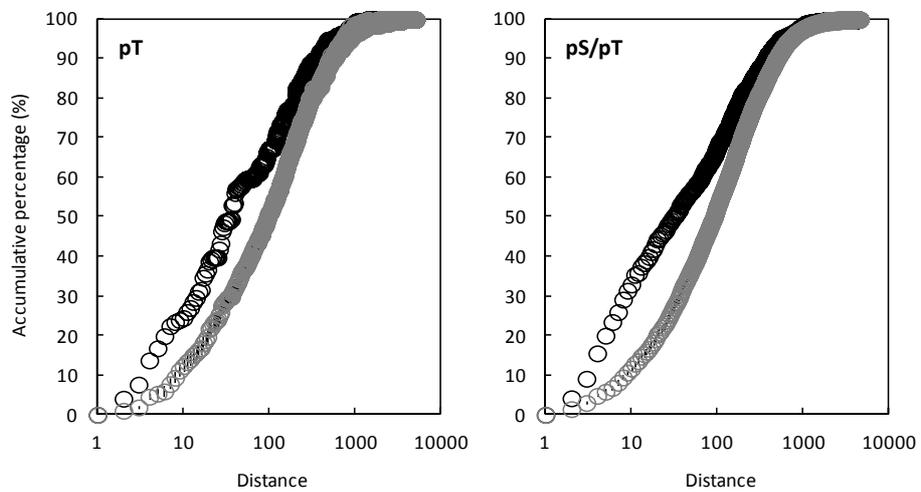


図 9: リン酸化部位間の距離

距離に対する累積数の割合をプロットし、リン酸化部位をランダムに再配置して作成したコントロールと比較した (*black*: 実データ, *gray*: ランダム)。

第七節 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⁵⁹⁻⁶¹。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⁶²。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⁶³。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 (図 10C)。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⁶⁴⁻⁶⁶ (図 10D)。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。

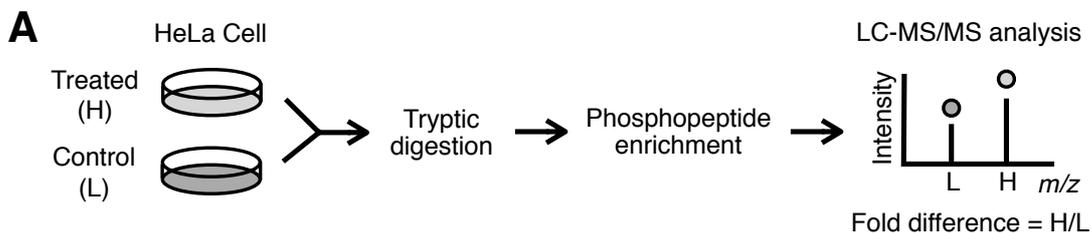
第二節 薬剤刺激によるリン酸化変動データの獲得

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 (図 10A)。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、上清画分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。

高い定量精度を実現するためには、標識効率の高さ、および実験操作により試料間に生じる誤差が重要となる。第一に、本実験での標識効率を検証するため、H アミノ酸を含む SILAC 培地培養した細胞からタンパク質を抽出し、LC-MS 分析法を用いたショットガンプロテオーム解析を行った。得られたペプチドの総面積値に占める標識ペプチド面積の割合を算出し、標識効率とした。SILAC 培養は二回繰り返し行い、それぞれに対して二回の LC-MS 測定を行った。その結果、得られた標識効率の平均値は、 $99.06 \pm 0.04\%$ であり、SILAC による高い標識効率を示した。第二に、試料間に生じる誤差を検証するため、H/L 試料の混合比を求めた。現在広く用いられている標識手法は SILAC の様に培地中で標識を行う方法の他に、アミノ基の対する還元的ジメチル化反応を利用した標識法 (ジメチル標識法) などの消化したペプチドに対する化学的標識法がある。後者の場合、比較する試料の混合は抽出タンパクを消化、および精製した後となり、実験誤差が生じ易い。その点、SILAC は細胞回収後という実験手順の上流の過程で比較試料の混合が可能であり、実験誤差を最小限に留められる。ここで、本実験での H および L 標識試料の混合比を検証するため、細胞回収後に H/L を混合した四種の試料に対して、各二回の LC-MS 測定を行った。各ペプチドの H/L 比から中央値を計算したところ、全ての混合試料で 1 に近い結果が得られ、各試

料の混合比はほぼ等しいことが示唆された (図 10B)。以下の実験では、この H/L 混合試料からリン酸化ペプチドを濃縮し、LC-MS 分析を用いて薬剤刺激による変動リン酸化プロテオームデータを取得した。

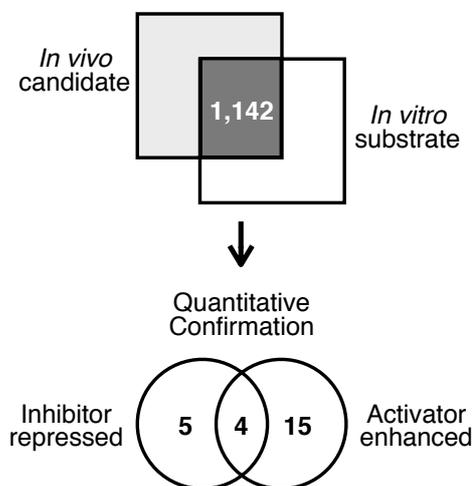
本実験では、一度の LC-MS 測定でより高い同定数を得るため、比較的長いカラムと長いグラジエント時間によるペプチド分離を行った。第一章では 10 cm のカラムに対して 60 分のグラジエント時間をかけてペプチド分離を行った。この分離条件では、タンパク質 100 μg をリン酸化濃縮した試料の注入が適量とされており、HeLa 細胞の消化物からは約 1,000 のリン酸化ペプチド同定が期待された⁴⁹。本実験では、25cm のカラムに、400 μg のペプチド消化物を濃縮して試料を注入し、グラジエント時間を 180 分の分離条件で LC-MS 測定を行った。その結果、平均 3,749 のリン酸化ペプチドが一度の LC-MS 測定で同定された (図 10B)。このようにして得られた定量リン酸化プロテオーム情報に対して二種類のフィルター手法を用い、信頼性の高い PKA 直接基質の同定を試みた (図 10CD)。



B

	Biological replicate	Treated	Control	SILAC ratio	Technical replicate	ID
Activator	1	Heavy	Light	1.020	1	3,543
					2	3,659
	2	Light	Heavy	0.984	1	4,174
					2	3,670
Inhibitor	1	Heavy	Light	1.055	1	3,848
					2	3,328
	2	Light	Heavy	1.076	1	4,121
					2	3,649

C Filtering by *in vitro* substrate



D Filtering by Motif sequence

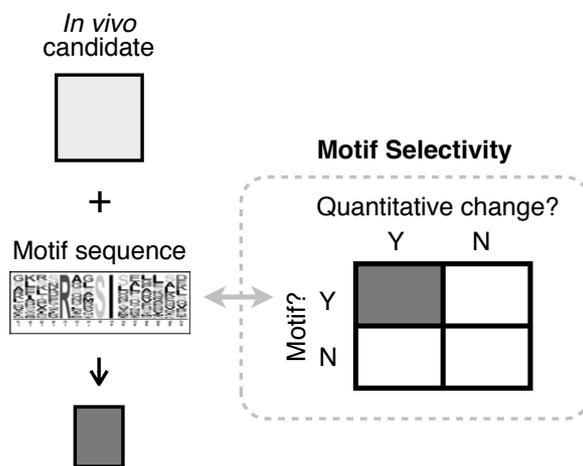


図 10: 実験概要

A. 薬剤刺激によるリン酸化変動を LC-MS 法にて定量的に評価した。B. 同位体標識 (H/L) を交換し、細胞実験を繰り返した (biological replicate=2)。SILAC ratio: 二測定間の H/L 混合比の平均 (technical replicate=2)。ID: リン酸化ペプチド同定数。C,D. 二通りの直接基質選別法

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 (図 10C)。 *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。

第一章でも述べた様に、LC-MS 測定にて MS² 情報が取られるペプチドは確率論的であり、全ての解析結果から同一の同定ペプチドリストは得られない。そのため、ある LC-MS 測定で同定されたペプチドが、同じ検体に対する別の測定で同定されないことがある。その場合、すでに同定されたピークの保持時間と質量電荷比を基にした LC-MS 測定情報の再検索により、別ファイル中でピークの再同定ができる。本手法では、まず、各 LC-MS 測定データ (8 ファイル) 中の同定リン酸化部位から、*in vitro* リン酸化部位との重複があるペプチド情報を抽出した。選択した全ペプチドを統合し冗長なデータを除いたところ、1,142 リン酸化ペプチドが再検索の対象候補となった。誤同定を除くため、改めて全 8 ファイルの対象ピークに対応するスペクトルを目視確認した。その後、ピークの積分により面積値を算出し、コントロールと薬剤添加試料の比からリン酸化変動値を計算した。この内、二回の繰り返し実験で同じリン酸化変動傾向を示したペプチドを、薬剤添加に応答するペプチドとした (表 7)。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 (表 8)。

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⁶⁷。S2152 部位は Filamin-A の C 末端に位置し、自身の安定性に関与することが知られている⁶⁸。Filamin-A の S2152 へのリン酸化を触媒する上流キナーゼとしては、その他に p90 リボソーム S6 キナーゼ (p90RSK)⁶⁹ と p21-activated kinase1 (PAK1)⁷⁰ の報告がある。リン酸化 Filamin-A はアクチン細胞骨格の再編成をおこない、細胞膜動態を変化する⁶⁷。残りの直接基質候補は PKA に関与する先行研究の報告は無いが、それぞれの分子機能は PKA の機能と関連性がある。例えば、Microtubule-associated protein 4 (MAP4_HUMAN) は細胞骨格の中でも微小管に関連するタンパク質であり、Vesicle-trafficking protein SEC22b (SEC22B_HUMAN) は小胞体とゴルジ体

間の輸送に関与し膜融合を調節する SNARE タンパク質である。Sodium-coupled neutral amino acid transporter 1 (S38A1_HUMAN) はナトリウムのトランスポーターの一種で、S137 は細胞質に露呈されている部位となる。

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 (図 11)。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。

表 7: *In vitro* 基質と重複する *in vivo* 基質のリン酸化変動

Category	Swissprot ID	Protein Discription	Identified sequencece and modifications	Site	Sequences	Fold change			
						a1	a2	i1	i2
A	FLNA_HUMAN	Filamin-A	RAPSVANVGSHCDLSLK:79.96633052075@S:4	2152	TRRRRAPSVANVGSH	0.79	0.71	-0.40	-0.55
A	MAP4_HUMAN	Microtubule-associated protein 4	KCSLPAEEDSVLEK:79.96633052075@S:3	636	TGTGKKCSLPAEEDS	0.87	0.79	-0.17	-0.10
A	S38A1_HUMAN	Sodium-coupled neutral amino acid transporter 1	ESRRSLTNSHLEK:79.96633052075@S:5	52	SDRESRRSLTNSHLE	0.74	0.37	-0.91	-0.49
A	SC22B_HUMAN	Vesicle-trafficking protein SEC22b	NLGSINTELQDVQR:79.96633052075@S:4	137	CARRNLGSINTELQD	3.09	3.35	-0.59	-0.41
B	AHNB_HUMAN	Neuroblast differentiation-associated protein AHNAK	LKSEDEVGDLGETQSR:79.96633052075@S:3	135	KIKPRLKSEDEVGEGD	-0.60	-1.32	2.84	1.74
B	CD2AP_HUMAN	CD2-associated protein	SVDFDSLTVR:79.96633052075@S:1	458	QLPLRPKSVDFDSLTL	-1.67	-0.35	1.62	1.88
C	CLIP1_HUMAN	CAP-Gly domain-containing linker protein 1	KISGTALQEALK:79.96633052075@S:3	348	SRYARKISGTALQE	3.46	4.03	n.d.	n.d.
C	LASP1_HUMAN	LIM and SH3 domain protein 1	MGPSGGEGMEPERRDSQDGSYSYR:79.96633052075@S:16	146	MEPERRDSQDGSYSYR	2.29	1.59	n.d.	n.d.
C	NRDC_HUMAN	Nardilysin	RGSLNAGDPEIVK:79.96633052075@S:3	94	EEEGRRGSLNAGDP	2.27	2.46	n.d.	n.d.
C	PDE3B_HUMAN	cGMP-inhibited 3',5'-cyclic phosphodiesterase	RRSSCVSLGETAASYGSK:79.96633052075@S:4,79.96633052075@S:7	296 299	IRPRRRSSCVSLGET RRRSCVSLGETAAS	1.94	3.26	n.d.	n.d.
C	RMP_HUMAN	Unconventional prefoldin RPB5 interactor	RKNSTGSGHSAQELPTIR:79.96633052075@S:7	375	KRKNSTGSGHSAQEL	0.78	1.26	n.d.	n.d.
C	SC16A_HUMAN	Protein transport protein Sec16A	RSSLSSHQSQIYR:79.96633052075@S:3	1191	SLASRRSSLSSHSHQ	0.89	0.46	n.d.	n.d.
C	TENS3_HUMAN	Tensin-3	KLSLGQYDNDAGQLPFSK:79.96633052075@S:3	776	GGRLRKLKSLGQYDND	1.36	1.06	n.d.	n.d.
C	TPI3_HUMAN	Triosephosphate isomerase	KQSLGELIGLINAAK:79.96633052075@S:3	58	KMNGRKQSLGELIGT	2.56	2.42	n.d.	n.d.
C	UB2J1_HUMAN	Ubiquitin-conjugating enzyme E2 J1	RLSTSPDVIQGHQPR:79.96633052075@S:3	266	QQSQRRLLSTSPDVIQ	1.85	1.53	n.d.	n.d.
C	YTDC2_HUMAN	Probable ATP-dependent RNA helicase YTHDC2	KSSADTEFDECTTAER:79.96633052075@S:3	1202	SNNSRKSSADTEFSD	2.51	1.66	n.d.	n.d.
C	IRS2_HUMAN	Insulin receptor substrate 2	RVSGDAAQLDLR:79.96633052075@S:3	560	GRSYRRVSGDAAQDL	1.41	1.71	-0.81	0.16
C	LRFN4_HUMAN	Leucine-rich repeat and fibronectin type-III domain-containing protein 4	SCSLDLGDAGCYGYAR:79.96633052075@S:3	585	PRPQRCSLDDLGDAG	1.02	0.89	-0.63	0.38
C	PA24A_HUMAN	Cytosolic phospholipase A2	QNPSRCSVSLNVEAR:79.96633052075@S:7,79.96633052075@S:9	727 729	RQNPSRCSVSLNVE NPSRCSVSLNVEAR	0.44	1.05	-0.91	0.50
C	SQSTM1_HUMAN	Sequestosome-1	RFSFCCSPEPEAEAAAGPGPCR:79.96633052075@S:3	24	AREIRRFSCCSP	2.11	1.71	0.03	-0.42
C	TB182_HUMAN	182 kDa tankyrase-1-binding protein	RFSEGVLSQSPQDQEK:79.96633052075@S:3	429	SLVQRRFSEGVLSQSP	1.78	1.36	-0.25	0.01
D	DKC1_HUMAN	H/ACA ribonucleoprotein complex subunit 4	SLPEEDVAIEIQAEEFLIKPESK:79.96633052075@S:1	21	KKKKERKSLPEEDVA	-0.88	-0.99	-0.02	-0.11
D	RAI3_HUMAN	Retinoic acid-induced protein 3	EFSIPR:79.96633052075@S:3	336	QPPQKEFSIPRAHAW	-0.76	-0.35	0.29	-0.19
D	ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3	SSGSPYGGYGSGGGSGGYGSR:79.96633052075@S:12	366	PYGGYGSGGGSGGY	-0.61	-0.35	-0.08	0.07
D	ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3	SSGSPYGGYGSGGGSGGYGSR:79.96633052075@S:2	356	GSFGGRSSGSPYGGG	-0.33	-0.28	0.06	0.00
D	SRSF9_HUMAN	Serine/arginine-rich splicing factor 9	GSPHYFSPFRPY:79.96633052075@S:2,79.96633052075@S:7	211 216	SPYQSRGSPHYFSPF RGSPHYFSPFRPY--	-0.71	-0.38	-0.05	0.05
D	STMN1_HUMAN	Stathmin	RASGQAFELILSPR:79.96633052075@S:3,79.96633052075@S:12	16 25	KELEKRASGQAFELI QAFELILSPRSKESV	-1.33	-1.43	-0.18	0.08
D	VIME_HUMAN	Vimentin	SLYASSPGGVYATR:79.96633052075@S:1	51	LRPSTSRSLYASSPG	-0.92	-0.52	0.03	0.13

(続く)

(承前)

Category	Swissprot ID	Protein Discription	Identified sequence and modifications	Site	Sequences	Fold change			
						a1	a2	i1	i2
E	DNJC2_HUMAN	DnaJ homolog subfamily C member 2	NASASFQLEDKK:79.96633052075@S:3,79.96633052075@S:5	47 49	KRRNRNASASFQELE RNRNASASFQLEDK	0.22	-0.31	-0.79	-0.98
E	ODPA_HUMAN	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	YHGHSMDPGVSYR:79.96633052075@S:7,79.96633052075@S:12	295 300	RYHGHSMDPGVSYR SMSDPGVSYRTREEI	-0.17	-0.02	-0.67	-0.63
E	RS6_HUMAN	40S ribosomal protein S6	RLSSLR:79.96633052075@S:3,79.96633052075@S:4	235 236	IAKRRRLSSLRASTS AKRRRLSSLRASTSK	-0.97	0.62	-1.56	-0.67
E	RS6_HUMAN	40S ribosomal protein S6	RLSSLR:79.96633052075@S:3,79.96633052075@S:4	235 236	IAKRRRLSSLRASTS AKRRRLSSLRASTSK	0.04	0.22	-1.32	-0.70
E	RS6_HUMAN	40S ribosomal protein S6	RLSSLR:79.96633052075@S:3,79.96633052075@S:4,79.96633052075@S:8	235 236 240	IAKRRRLSSLRASTS AKRRRLSSLRASTSK RLSSLR:79.96633052075@S:8	0.06	0.25	-1.76	-1.51
F	DNJC2_HUMAN	DnaJ homolog subfamily C member 2	NASASFQLEDKK:79.96633052075@S:3	47	KRRNRNASASFQELE	0.04	0.36	0.46	0.87
G	HN1_HUMAN	Hematological and neurological expressed 1 protein	RNSSEASSGDFDLK:79.96633052075@S:3	87	SGLQRNSSEASSGD	0.87	0.64	0.19	0.23
G	KI21A_HUMAN	Kinesin-like protein KIF21A	KLSSSDAPAQDTGSSAAAVETDASR:79.96633052075@S:4	854	KVTRKLSDDAPAQD	1.68	1.35	0.38	0.63
G	PLCL2_HUMAN	Inactive phospholipase C-like protein 2	SLEVIPEK:79.96633052075@S:1	1113	DVQKPRRSLEVIPEK	0.29	0.33	0.44	0.42
G	PKN2_HUMAN	Serine/threonine-protein kinase N2	ASSLGEIDSESLR:79.96633052075@S:3	583	PAPPRASSLGEIDES	0.23	0.75	0.41	1.00
H	PDCD4_HUMAN	Programmed cell death protein 4	DSGRGDSVSDSGDALR:79.96633052075@S:7	76	RDSGRGDSVSDSGSD	-0.49	-0.46	-0.87	-1.34
H	PDCD4_HUMAN	Programmed cell death protein 4	NSSRDSGRGDSVSDSGDALR:79.96633052075@S:11	76	RDSGRGDSVSDSGSD	-0.67	-0.78	-0.85	-0.59
H	PRC2B_HUMAN	Protein PRC2B	VASETHSEGSEYEELPK:79.96633052075@S:3,79.96633052075@S:7,79.96633052075@S:10	1132 1136 1139	KPRRRVASETHSEGS RVASETHSEGSEYEE SETHSEGSEYEELPK	-0.39	-0.36	-0.83	-0.77
H	RS10L_HUMAN	Putative 40S ribosomal protein S10-like	KAEAGAGSATEFQFR:79.96633052075@S:8	157	KAEAGAGSATEFQFR	-0.75	-0.58	-0.10	-0.18

Category: 薬剤によるリン酸化変動の分類

- (1) 両方で変化があった同定ペプチド (A: 阻害薬↓活性化薬↑, B: 阻害薬↑活性化薬↓, G: 両方↑, H: 両方↓),
- (2) 活性化薬のみで変化があった同定ペプチド (C: ↑, D: ↓),
- (3) 阻害薬のみで変化があった同定ペプチド (E: ↓, F: ↑)

Sequences: リン酸化部位を中心とした周辺配列

Fold change: コントロールに対する薬剤添加試料のピーク面積比 (a1, a2: 活性化薬, i1, i2: 阻害薬)。二回の反復測定
の平均値を対数に変換した (底は 2)。

表 8: *In vitro* 基質情報による PKA 直接基質の選別

#	UniProt ID	Protein Name	Site	Sequence	Known PKA association?	Known phosphorylation function
1	FLNA_HUMAN	Filamin-A	2152	TRRRRAPsVANVGS <u>H</u>	Yes, putative upstream kinase is PKA	Effect on FLNS: Protein stabilization; Effects of modification on biological processes; cytoskeletal reorganization
2	MAP4_HUMAN	Microtubule-associated protein 4	636	TGTGK <u>K</u> CsLPAAEEDS	No	na
3	S38A1_HUMAN	Sodium-coupled neutral amino acid transporter 1	52	SDRESRR <u>s</u> LTNSHLE	No	na
4	SC22B_HUMAN	Vesicle-trafficking protein SEC22b	137	CARRNL <u>G</u> sINTELQD	No	na

Sequence: リン酸化部位を中心とした前後7残基のアミノ酸配列。下線は対応するモチーフ配列 (第一章) を示した

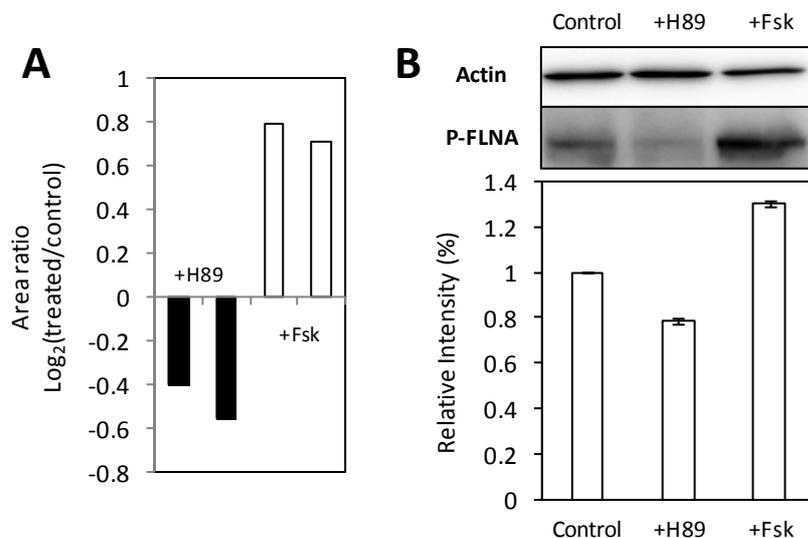


図 11: LC-MS およびウエスタンブロッティングによる定量評価

A. MS 定量解析結果。薬剤刺激による変動値を底 2 の対数を取った (*black*: H89, *white*: Fsk)。 **B.** (上) ウエスタンブロッティング結果 (下) コントロール (+DMSO) に対する薬剤刺激 (+H89 または +Fsk) によるリン酸化フィラミン A の変動比。リン酸化フィラミン A の検出値はアクチンで正規化し、二回の実験繰り返しの平均値を示した。

第四節 同定情報を用いたモチーフ配列選択性の評価

モチーフ配列は、汎用性が高くキナーゼの直接基質の選別に利用できる。しかし、結果に含まれる偽陽性を低減するためには、そのモチーフを標的とするキナーゼの選択性が重要となる。そこで、モチーフ配列を用いた PKA 基質の推定を行う前に、薬剤刺激によるリン酸化変動データを用いて、モチーフ配列の選択性を評価した。

薬剤刺激を行ったリン酸化プロテオームの同定修飾部位のうち、その存在量が活性化薬によって増加、あるいは、阻害薬によって減少したものは、PKA の直接基質である蓋然性が高い。その一方で、モチーフ配列による PKA の基質選択が正当であれば、リン酸化量に変化のあったリン酸化部位は、変化の無いものに比べて高い割合でモチーフ配列を含有すると考えられる。そのため、ここでは変動のあったリン酸化部位に占めるモチーフの含有率を、変動の無かったリン酸化部位でのモチーフ含有率で除し、モチーフの選択性を評価するスコア (増倍比) とした。

全同定ペプチドの定量値を計算するため、ソフトウェアにてピークを検出し、H/L 面積比の自動定量を行い、ピークの相似比を示す QuanScore が高いピークペアのみを採用した (QuanScore>0.8)。リン酸化部位が複数のペプチドから同定された場合、対応するペプチド面積比の平均値を各部位のリン酸化変動値とした。この変動値が、活性化薬で 2 倍以上あるいは、阻害薬で 0.5 倍以下の変動を示したものを「変動があった部位」、それ以外を「変動が無い部位」とした。ここでは、変化量が比較的大きい部位のみを「変動あり」としているため、二回の繰り返し実験のうち一方で変化が認められたリン酸化部位は全て採用した。この変動基質群に対して、第一章で同定した PKA のモチーフ配列の有無を検証し、活性化薬と阻害薬によるリン酸化変動情報に対してそれぞれ増倍比を計算した。増倍比の合計値が高い配列から順に、選択性が高いモチーフ配列とした (表 9)。この内、上位 5 のモチーフ配列を、この後の PKA の直接基質の推定に利用した。

選択性の高いモチーフとして同定されたトップ 5 位のうち、RKX[pS]、KKX[pS]、および RX[pS]V は第一章の AKT1 との比較モチーフ解析で PKA に選択性の高いモチーフとして同定された。一方で、RRX[pS]は AKT1 と PKA 間の比較では PKA への選択性が低いと評価されており、そのほかのモチーフ配列と比べると PKA 以外の基質候補を含む蓋然性が高い。

活性化薬と阻害薬を比較した時、その増倍比が大きく異なるモチーフ配列が存在する。この理由としては、一つ目に細胞の薬剤応答性が考えられる。今回実験に用いた活性化薬および阻害薬の濃度は一意である。第三節にて LC-MS およびウエスタンブロッティングにて既知基質の定量変化が確認され、基質候補を抽出する目的に対して、今回の薬剤添加濃度は十分であると検証された。しかし、そのリン酸化変動の強さは薬剤の濃度にある程度依存するため、より高濃度の阻害薬添加は増倍比の更なる上昇に寄与すると考えられる。また、二つ目の理由として、MS 分析の検出感度がある。リン酸化の減少を検出する阻害薬と比較して、対象基質のリン酸化が上昇する活性化薬はより MS 測定により検出されやすくなる。この検出状態の偏りも、活性化薬でより高い増倍比を示すモチーフが多い原因となり得る。

表 9: 変動値によるモチーフ配列選択性の評価

#	Motif sequence	Activator	Inhibitor
1RK.S.....	14.59	14.75
2R..SL.....	14.92	10.67
3RR.S.....	19.70	3.69
4KK.S.....	13.67	0.00
5R.SV.....	10.92	0.00
6RN.S.....	0.00	9.60
7	..R.R..S.....	4.64	4.88
8K.S.....	4.76	4.17
9R..S.....	4.51	2.65
10R.S.....	4.53	2.20
11RS.S.....	3.80	1.45
12RGS.....	4.97	0.00
13	..R....S.....	1.89	1.81
14	...R..GS.....	3.65	0.00
15K..S.....	1.81	1.59
16KS.....	1.85	1.33
17	..R...GS.....	2.88	0.00
18	...R...S.....	1.46	1.31
19	...K...S.....	0.90	0.00

第五節 モチーフ配列による PKA 標的基質の選別

第四節で評価された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 (図 7C)。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 (表 10)。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。PKA の既知基質が関連する機能は多岐にわたっている。以下で今回の報告に含まれる既知基質の機能情報を述べる。

(1) CD44 antigen (CD44_HUMAN)

CD44_HUMAN は細胞-細胞または細胞-細胞間基質を接着させる接着分子であり、S697 部位は細胞質側に位置している。細胞接着には PKC (タンパク質キナーゼ C) の関与が知られているが、この部位のリン酸化は PKC ではなく PKA に触媒されており、細胞の運動性および方向決定を制御することが報告されている⁷¹。この報告中ではアイソフォームを実験に用いており、本文中のリン酸化部位は S316 と表記されているが、SwissProt に登録のある CD44_HUMAN では S697 に該当し、これらの周辺配列は保存されている。

(2) Inositol 1,4,5-trisphosphate receptor type 3 (ITPR3_HUMAN)

ITPR3_HUMAN は IP₃ (イノシトール 1,4,5-三リン酸) タイプ 3 受容体であり、小胞体の膜表面で四量体を形成し Ca²⁺チャネルの働きをする。表面に曝露されたペプチドループに位置する S916, S934, S1832 のリン酸化が Ca²⁺チャネルの活性に影響を及ぼす。この 3 カ所中でも特に S934 が PKA による基質部位として報告されている⁷²。

(3) LIM and SH3 domain protein 1 (LASP1_HUMAN)

LASP1_HUMAN は SH3 ドメインと LIM ドメインを持つタンパク質であることから、アダプター分子としての機能が知られている。SH3 ドメインはプロリンに富んだ短いアミノ酸配列が結合するドメインである。また、LIM ドメインは、2 個のジンクフィンガーが 2 アミノ酸で連結された構造を持ち、同じくタンパク質間の相互作用を担う。LIM ドメインは転写因子以外にも、細胞骨格結合タンパク質や、タンパク質キナーゼなどのタンパク質にも

見いだされており、細胞骨格の制御や細胞内情報伝達に関わることが知られている。S146 への PKA あるいは PKG (c-GMP 依存性タンパク質リン酸化酵素) によるリン酸化は、LASP1 のアクチン結合アフィニティーを弱めるだけでなく、細胞局在の移行および細胞移動を変化するとの報告がある^{73,74}。

(4) cGMP-inhibited 3',5'-cyclic phosphodiesterase A (PDE3A_HUMAN)

PDE3A_HUMAN は PKCA (タンパク質キナーゼ α タイプ)⁷⁵ と PKA⁷⁶ が責任キナーゼとして知られている。複数のリン酸化部位を持つ PDE3A は環状リン酸ジエステルを加水分解するホスフォジエステラーゼであり、cAMP および cGMP のどちらにも選択性を示す。p42/p44 MAPK、SAPK2/p38、および PKC を活性化する PMA (horbol 12-Myristate 13-acetate) 刺激の結果 S428 がリン酸化され、14-3-3 に結合する。一方で、H89 および Fsk を用いた細胞内実験で、S312 のリン酸化が誘発されることは確認されている。しかし、このリン酸化により自身の酵素活性は上昇するが、14-3-3 への結合は無い^{75,76}。

また、*in vivo* の実験にて責任キナーゼが PKA であることは確認されていないが、40S ribosomal protein S6 (RS6_HUMAN) および、Cytosolic phospholipase A2 (PA24A_HUMAN) は、今回同定された部位へのリン酸化修飾に伴い、特定の細胞内機能を獲得することが知られている。RS6_HUMAN は転写制御に重要な機能を持つリボソームタンパク質であり、複数のキナーゼの間接的または直接的な制御が S235 および S236 への修飾に関わる。その中でも直接的な上流キナーゼとしては、p70 リボソーム S6 キナーゼ (p70RSK) と p90RSK の報告がある^{77,78}。この先行研究では、RAS /ERK シグナルの下流に位置する RSK が S235/236 をリン酸化することにより、翻訳制御を行う機構を明らかにしている⁷⁸。一方、PA24A_HUMAN の S727 リン酸化はタンパク質自身を活性化し、その結果、細胞局在や分子相互作用を制御⁷⁹することが明らかとなっている。

上記の直接基質を除いた残りの基質候補 15 カ所のうち、11 カ所は第一章の基質プロファイルで同定された (表 10)。MAP4_HUMAN に関しては、基質およびモチーフ配列による両選抜法により直接基質と特定された。ソフトウェアを用いた自動定量の結果、MAP4_HUMAN の S636 のリン酸化は、阻害薬にてリン酸化が減少、および、活性化薬によ

り増加していた。また、RS6_HUMAN (S235/S236) のリン酸化は、阻害薬による減少が確認された。一方、活性化薬によってリン酸化の増加が観察された部位は、上述の既知基質である LASP1_HUMAN (S146)と PA24A_HUMAN (S727) の 2 カ所に加えて、Insulin receptor substrate 2 (IRS2_HUMAN) (S560)、Nardilysin (NRDC_HUMAN) (S94)、Sequestosome-1 (SQSTM_HUMAN) (S24)、182 kDa tankyrase-1-binding protein (TB182_HUMAN) (S429)、Tensin-3 (TENS3_HUMAN) (S776)、Triosephosphate isomerase (TPIS_HUMAN) (S58)、および、Ubiquitin-conjugating enzyme E2 J1 (UB2J1_HUMAN) (S266) の 7 カ所であった。PKA による直接的な触媒が *in vitro* で確認され、かつ薬剤添加によりリン酸化変動を起こす上記の基質部位は、*in vivo* にて PKA に触媒を受ける新規直接基質であることが強く示唆される。

第 3 節で検討した、*in vitro* PKA 基質であり、更に薬剤添加により *in vivo* でリン酸化が変動する基質 (表 7) のうち、キナーゼ選択性の高いモチーフ配列 (表 9) を含有するが、今回基質候補として選択されなかった基質が存在する (活性化薬でリン酸化が促進する 4 カ所、および阻害薬で抑制する 1 カ所)。これらは本節で用いた厳しい変化量の検索条件を満たさず、基質候補として選出されなかったためである。

表 10:モチーフ配列による PKA 直接基質の選別

#	UniProt ID	Protein Name	Site	Sequence	Matched motif	Known PKA association?	Known phosphorylation function
1	CD44_HUMAN	CD44 antigen	697	AVED <u>R</u> KPsGLNGEAS	1	Yes, putative upstream kinase	Cell motility, altered
2	FA83H_HUMAN	Protein FAM83H	881	AYPER <u>K</u> GsPTPGFST	1	No	na
3	HERC1_HUMAN	Probable E3 ubiquitin-protein ligase HERC1	2720	DEVG <u>R</u> RRQsLTS ² PDSQ	2,3	No	na
4	IF2P_HUMAN	Eukaryotic translation initiation factor 5B	107	GQK <u>G</u> KKQsFDDNDSE	4	No	na
5	IRS2_HUMAN	Insulin receptor substrate 2	560	GRSY <u>R</u> RVsGDAAQDL	3	Yes, confirmed in vitro	na
6	ITPR3_HUMAN	Inositol 1,4,5-trisphosphate receptor type 3	934	MVLS <u>R</u> KQsVFSAPSL	1	Yes, putative upstream kinase (rat)	na
7	LASP1_HUMAN	LIM and SH3 domain protein 1	146	MEPE <u>R</u> RDsQDGSSYR	3	Yes, putative upstream kinase and confirmed in vitro	Intracellular localization; molecular association, regulation
8	MAP4_HUMAN	Microtubule-associated protein 4	636	TGTG <u>K</u> KCsLPAEEDS	4	Yes, confirmed in vitro	na
9	NRDC_HUMAN	Nardilysin	94	EEEG <u>R</u> RGsLSNAGDP	2,3	Yes, confirmed in vitro	na
10	PA24A_HUMAN	Cytosolic phospholipase A2	727	RQNPS <u>R</u> CsVLSNVE	5	Yes, confirmed in vitro	Enzymatic activity, induced; intracellular localization; molecular association, regulation
11	PDE3A_HUMAN	cGMP-inhibited 3',5'-cyclic phosphodiesterase A	312	SKSH <u>R</u> RTsLPCIPRE	2,3	Yes, putative upstream kinase	Enzymatic activity, induced
12	PUM1_HUMAN	Pumilio homolog 1	709	GSGS <u>R</u> RDsLTGSSDL	2,3	No	na
13	RS6_HUMAN	40S ribosomal protein S6	235	IAK <u>R</u> RRLsSLRASTS	3	Yes, confirmed in vitro	Molecular association, regulation
14	RS6_HUMAN	40S ribosomal protein S6	236	AK <u>R</u> RRLsLRASTSK	2	Yes, confirmed in vitro	Molecular association, regulation
15	SQSTM_HUMAN	Sequestosome-1	24	AREI <u>R</u> RFsFCCSPEP	3	Yes, confirmed in vitro	na
16	TB182_HUMAN	182 kDa tankyrase-1-binding protein	429	SLVQ <u>R</u> RFsEGVLQSP	3	Yes, confirmed in vitro	na
17	TENS3_HUMAN	Tensin-like SH2 domain-containing protein 1	776	GGRL <u>R</u> KLsLGQYDND	1,2	Yes, confirmed in vitro	na
18	TPIS_HUMAN	Triosephosphate isomerase	58	KMNG <u>R</u> KQsLGELIGT	1,2	Yes, confirmed in vitro	na
19	UB2J1_HUMAN	Ubiquitin-conjugating enzyme E2 J1	266	QQSQ <u>R</u> RRLsTSPDVIQ	3	Yes, confirmed in vitro	na

Sequence: リン酸化部位を中心とした前後 7 残基のアミノ酸配列。下線は対応するモチーフ配列 (表 9) を示した

第六節 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。

実験の部

1. 材料

ダルベッコ・フォークト変法イーグル最小必須培地 (DMEM)、カナマイシン、リン酸塩緩衝生理食塩水 (PBS)、ジチオトレイトール (DTT)、ヨードアセトアミド (IAA)、エンドプロテイナーゼ lys-C、V8 プロテアーゼ、ピペリジン、リジン、およびアルギニンは和光純薬工業 (大阪) より購入した。ウシ胎児血清 (FBS) は Life Technologies (Grand Island, NY, USA) より購入した。プロテアーゼ阻害薬、重炭酸トリエチルアンモニウムバッファー (TEAB)、ホルムアルデヒド、シアノ水素化ホウ素ナトリウム、フォルスコリン、および H-89 二塩酸塩水和物はシグマ アルドリッチ ジャパン (東京) より購入した。BCA タンパク定量キットは、サーモフィッシャーサイエンティフィック株式会社 (横浜) より購入した。トリプシン、および TSAP はプロメガ株式会社 (東京) より購入した。アミコンウルトラ 10K はメルクミリポア (Billerica, MA, USA) より購入した。リコンビナントキナーゼは、カルナバイオサイエンス株式会社 (神戸) より供与された。PVDF 膜、およびウエスタンブロットイング検出試薬 (ECL Prime western blotting detection reagent) は GL Healthcare (東京) より購入した。ブロッキング試薬はブロッキングワン P をナカライテスク (京都) より購入した。ウエスタンブロットイングの検出用抗体は全てウサギ由来のものを CST ジャパン株式会社 (東京) より購入した。

2. 細胞培養

HeLa S3 細胞の培養には 10%FBS と 100 µg/ml カナマイシンを添加した DMEM 培地を用いた。直径 15 cm の培養皿上にて、およそ 80%コンフルエントの状態まで HeLa 細胞を培養した。氷冷した PBS にて細胞を洗浄した後に、PBS 液中でスクレイパーにて剥離した。PBS 溶液ごと細胞をチューブに回収した後、4°C、500 g にて 10 分間遠心分離を行い、ペレットのみを使用まで -80 °C にて保存した。

3. タンパク質抽出

細胞溶解バッファー A (20 mM HEPES (pH 7.5)、250 mM スクロース、1.5 mM MgCl₂、10 mM

KCl、0.5%ノニデット P-40、および 1%プロテアーゼ阻害薬)にて細胞を再懸濁し、氷上にて 5 分間静置した後に 5 分間超音波処理した。この細胞溶解液を 4°C、1,500 g にて 5 分間遠心分離した後の上清を、本文中の「上清画分」として取得した。ペレット画分の取得は、おもに Dignam らの論文を参考にした⁸⁰。手順を簡潔に述べると、細胞溶解バッファーB (10 mM HEPES (pH 7.9)、1.5 mM MgCl₂、10 mM KCl、および 0.5 mM DTT) に細胞を懸濁した後、氷上に 10 分間静置した。溶液を 4°C、1,000 g にて 10 分間遠心分離した後、再度、ペレットを溶解バッファーB に懸濁し、コンテス ダウンス ホモジナイザー (B タイプ) を用いて細胞を粉碎した。細胞膜の粉碎を顕微鏡で確認した後、細胞溶解液を 4°C、1,000 g にて 10 分間遠心分離した。上清を除いた後に、ペレットのみを再度 4°C、25,000 g にて 20 分間遠心分離した。得られたペレットを、溶解バッファーC (20 mM HEPES (pH 7.9)、25% (v/v) glycerol、0.42 M NaCl、1.5 mM MgCl₂、0.2 mM EDTA、0.5 mM DTT、および 1%プロテアーゼ阻害薬) に溶解した。溶液中にて、再度ダウンス ホモジナイザー (B タイプ) を用いて核膜の粉碎を行った後、スターラーで攪拌しながら氷上に 10 分間静置した。その後、溶液を 4°C、25,000 g にて 30 分間遠心分離し、得られた上清を本文中の「ペレット画分」とした。キナーゼ反応を行う際は、アミコンウルトラ 10K を用いて 4°C、14,000 g の遠心分離にて細胞溶液を濃縮し、40 mM Tris-HCl (pH 7.5) にて再度希釈することで、キナーゼ反応溶液へ置換した。この濃縮希釈の手順は二回繰り返した。BCA タンパク定量キットを用いて、溶液中のタンパク質量を測定した。

4. 脱リン酸化反応およびリン酸化反応

脱リン酸化反応は、1 mg タンパク質あたり 10 units の TSAP を添加し、37°C にて 1 時間加熱することにより行った。その後、TSAP を不活性化するために 75°C にて 30 分間加熱し、氷上に静置した。続くキナーゼ反応は、脱リン酸化タンパク質を 100 µg ずつ分注した後、キナーゼ反応溶液 (40 mM Tris-HCl (pH 7.5)、20mM MgCl₂、および 1 mM アデノシン三リン酸 (ATP)) 中にて、リコンビナントキナーゼまたは精製水を添加し、37°C で 3 時間加熱することにより行った。95°C で 5 分間加熱しキナーゼ反応を停止した後、5 分間氷上に静置した。最終濃度が 100 mM Tris-HCl (pH 9.0)、8M Urea になるように溶液を希釈し、タンパク質を変性した。

5. リン酸化ペプチド試料調整

消化前のタンパク質の還元アルキル化、および酵素消化は、今見らの報告⁸¹に従って行った。具体的には、DTT および IAA の添加によりタンパク質の還元アルキル化を行った後、lys-C およびトリプシンを用いてタンパク質を断片化した。リン酸化プロテオーム解析用の試料は、ヒドロキシ酸修飾酸化金属クロマトグラフィーを応用した HAMMOC カラム⁴⁹ によるリン酸化ペプチド濃縮法を用いて、ペプチド消化物からリン酸化ペプチドの濃縮を行った。HAMMOC カラムからの溶出には 0.5%ピペリジンを用いた。その後、SDB-XC を充填した StageTip⁸²を用いてペプチドの脱塩・濃縮を行い、LC-MS/MS に導入する試料とした。ペプチド基質は、タンパク質の消化および StageTip による脱塩により調製した。溶液中のペプチドを遠心エバポレーターにて乾固し、再度キナーゼ反応溶液に溶解した後に、上述と同じ手順でキナーゼ反応を行った。基質フォーム別のキナーゼ反応比較に用いた試料は、lys-C、トリプシン、V8 プロテアーゼによって消化した後、LC-MS/MS 測定を行った。

6. 同位体ラベル

6.1. ジメチルラベル

ジメチルラベルは、Boersema らの報告に変更を加えて行った⁵¹。手順を簡潔に述べると、脱塩後の溶液中のペプチドを乾固した後に、100 μg タンパク質を 200 μl の 200 mM TEAB に再溶解した。コントロールの試料とキナーゼ添加した試料それぞれに、8 μl の質量の異なる 8%ホルムアルデヒド同位体 (CH_2O または $^{13}\text{CD}_2\text{O}$) を添加した。その後、1.2 M シアノ水素化ホウ素ナトリウム (NaBH_3CN) をそれぞれの試料に 8 μl ずつ添加し、室温にて一時間静置した。各試料に 2%アンモニア溶液を 32 μl 添加した後、ギ酸を添加して溶液を酸性化し反応を止めた。その後、二つの試料を混ぜ、StageTip にてペプチドを精製した。上述の通りリン酸化ペプチド濃縮を行い、LC-MS/MS を用いて分析した。

6.2. SILAC ラベル

通常の DMEM からリジンおよびアルギニンを除いたカスタム培地に、通常のリジンとアルギニンを添加した L (light) 標識培養液と、 $^{13}\text{C}_6$, $^{15}\text{N}_2$ -リジンと $^{13}\text{C}_6$, $^{15}\text{N}_4$ -アルギニンを添加した H (heavy) 標識培養液を調製した。それぞれに 100 $\mu\text{g}/\text{ml}$ カナマイシン、および 10% FBS

を添加して、HeLa S3 細胞の培養に用いた。Ong らの報告⁸³に従い、細胞内のアミノ酸が安定同位体アミノ酸に置き換わるまで細胞を 5 倍化期間以上培養し、実験に使用した。

同数の細胞をまいた培養皿を用意し、薬剤刺激を行う 3 時間前にあらかじめ新たな DMEM へ培地交換を行った。それぞれ DMSO に溶解した 30 μ M H89 または 50 μ M フォルスコリンを添加し、細胞インキュベーター (37°C、5% CO₂) にて一時間インキュベートした。コントロールには同濃度の DMEM を添加した培養皿を用意した。DMEM を回収した後、氷冷した PBS にて細胞を洗浄し、PBS 溶液中にてスクレイパーを用いて細胞を剥離した。PBS 溶液ごと細胞を回収した後、再度 PBS にて細胞皿を共洗いして溶液を回収した。細胞液を 4°C、500 g にて 5 分間遠心分離し、ペレットのみを -80 °C にて使用まで保存した。タンパク質抽出前に薬剤刺激試料とコントロール試料を混合し、タンパク質抽出、消化、リン酸化ペプチド濃縮を上述の手順で行い、LC-MS/MS 法にて分析した。

7. NanoLC-MS/MS システム

LC-MS/MS 装置は HTC-PAL (CTC Analytics, Zwingen, Switzerland) および Ultimate 3000 pump (DIONEX, Germering, Germany) を装備した、TripleTOF 5600 (AB Sciex, Foster City, CA, USA)、または、LTQ-Orbitrap XL (Thermo Fisher Scientific, Rockwell, Illinois, USA) を使用した。ペプチドの分離には、粒子径 3 μ m の ReproSil-Pur C18-AQ (Dr. Maisch, Ammerbuch, Germany) を窒素ガスにて自家充填したスプレーチップ一体型の nano LC カラム (第一章: 100 mm length x 100 μ m i.d.、第二章: 250 mm length x 100 μ m) を使用した。試料注入量は 5 μ l、移動相流速は 500 nL/min に設定した。移動相には (A) 0.5% 酢酸水溶液、および (B) 0.5% 酢酸含有 80% アセトニトリル水溶液を用いた。グラジエント条件は 5-10%B (5 分間) \rightarrow 10-40%B (60 分間) \rightarrow 40-100%B (5 分間) \rightarrow 100% B (10 分間)、または 5-10%B (5 分間) \rightarrow 10-40%B (180 分間) \rightarrow 40-100%B (5 分間) \rightarrow 100% B (10 分間) を適用した。TripleTOF 5600 では、スプレー電圧は 2,300 V (ポジティブイオンモード) とし、スキャン範囲を m/z 300-1,500 に設定した。また、high-sensitivity モードで 0.25 秒間 MS スキャンを行った。通常のプロテオーム解析では、MS スペクトル中最も強度の強い 10 個のプレカーサーイオンに対して、データ依存的に 0.1 秒間 MS² スペクトルを取得した。リン酸化プロテオーム解析の際は、MS スペクトル中最も強度の強い 7 個のプレカーサーイオンに対して、データ依存的に 0.143 秒間 MS² スペ

MS² スペクトルを取得した。どちらの場合も、同一イオンに対する MS² スキャンを防ぐため、すでにスキャンされたイオンは 12 秒間 MS² スキャンの対象から除外する様に設定した。Analyst TF1.6 のローリング衝突誘起解離 (CID) 機能を採用し、CID エネルギーは自動的に調節する様にした。LTQ-Orbitrap XL では、スプレー電圧は 2,400 V (ポジティブイオンモード) とし、スキャン範囲を *m/z* 350-1,500 に設定した。また、MS スペクトル中最も強度の強い 10 個が Orbitrap (分解能 R=60,000, FWHM@*m/z* 400) によって選択され、イオントラップにてデータ依存的に MS² スキャンを行う様に設定した。CID エネルギーは 35.0 とした。dynamic exclusion time は 20 秒間とし、繰り返しのカウント数は 1、また exclusion list のサイズは 500 に設定した。また、質量分析計による測定精度を保つため lock mass 機能を採用した。LTQ-Orbitrap XL は、第一章の脱リン酸化反応効率の定量評価の試料測定に用い、その他の測定は全て TripleTOF 5600 を用いた。

8. タンパク質同定のためのデータベースサーチ

TripleTOF によって得られた解析データは、AB Sciex MS Data Converter を用いてピークリストを作成した。データベース検索は基本的には Mascot v. 2.4 (Matrix Sciences, London, U.K.) を利用し、一部、AB SCIEX ProteinPilot (v. 4.5) の Paragon を用いて検索を行った (詳細は以下で述べる)。どちらの場合も、データベースには SwissProt Database (バージョン: 第一章 2013-03、第二章 2013-11) を使用した。検索条件は既報⁸⁴と同様に設定したが、Mascot のサーチ条件は、計算質量との許容誤差範囲をプレカーサーイオンで 10 ppm、プロダクトイオンで 0.1 Da と変更した。また、試料により以下の点は Mascot の検索条件に変更を加えた。三種類の酵素を用いて消化を行った試料は、制限酵素にトリプシンおよび V8 を指定した。SILAC ラベル試料には、¹³C₆, ¹⁵N₂-リジンと¹³C₆, ¹⁵N₄-アルギニン、ジメチルラベル試料には、リジンの側鎖及びペプチドの N 末端の¹H₄, ¹²C₂-または ²H₄, ¹³C₂-ジメチル化を variable modification として追加した。LTQ-Orbitrap XL によって得られた解析データは、Mass Navigator v. 1.2 (三井情報株式会社、東京) を用いてピークリストの抽出を行った。その際は Mascot エンジンのみを検索に利用し、計算質量との許容誤差範囲はプレカーサーイオンで 3 ppm、プロダクトイオンで 0.8 Da に設定した。同定結果の擬陽性率 (false discovery ratio) は、タンパク質配列をランダムに並び替えた偽データベース (decoy database) に対してサーチ

を行った結果と比較して計算し、1%未満であった。第一章で記した *in vitro* kinase 反応実験の LC-MS/MS 分析結果に対してのみ、Mascot および Paragon の二つのサーチエンジンによる検索を適用した。データの統合を行う際に、同一のスキャンから異なる同定結果が得られた場合には Mascot による同定結果を優先的に採用した。リン酸化部位の確認には、PTM スコア⁶⁰による計算 (≥ 0.75 を採用) と、既報^{49,85}で述べられているリン酸化部位の決定を支持する b-イオンおよび y-イオンの組み合わせを確認する二手法を用いた。ここで、MS² スキャンからリン酸化部位が決定できた部位のみを、最終的なリン酸化部位数として考慮した。リン酸化ペプチドを再度タンパク質配列上にマッピングし、同定タンパク質のリストを作成した。ペプチド配列が複数のタンパク質に一致した場合、任意に一つのタンパク質を割り振った。また、*in vitro* キナーゼ反応では、リコンビナントキナーゼ自身のリン酸化部位は最終結果から除いている。

9. 情報解析

9.1. 共通部

基本的な配列処理、および情報解析は、Perl 言語 (v. 5.16.3) で作成したプログラムを利用した。

9.2. 第一章

DAVID (v. 6.7)⁸⁶による機能解析には、上清画分または核画分から同定された SwissProt ID を解析対象データ (foreground) として使用した。バックグラウンド (background) データとして、web 上で用意されている全てのヒトタンパク質リストの ‘Homo sapiens’ を選択した。結果として得られた機能同定チャート (functional annotation chart) から、GO タームのひとつである ‘cellular component’ を p-値 (<0.01) が昇順になるようにリストした。 χ^2 検定は R (v2.15.1) にて行った。アミノ酸頻度プロットとモチーフ解析は、リン酸基の付与が一カ所であったリン酸化ペプチドのみを使用した。また、複数のタンパク質に照合したペプチド配列は除外した。PKA、ERK1、および AKT1 の PhosphoSitePlus に登録のあるリン酸化部位とタンパク質配列データは、web サイト (<http://www.phosphosite.org>) より取得した (ver. last modified on 3rd May, 2013). 全てのデータのから、基質が検出された生物種が ‘human’ か

つ 'in vitro' 実験のリン酸化部位を解析に利用した。本研究で得られたリン酸化部位、および PhosphoSitePlus に登録のあるリン酸化部位情報とタンパク質配列を照合し、リン酸化部位を中心とした前後 7 残基 (全 15 残基) を取得した。リン酸化部位が N 末あるいは C 末付近にあり、15 残基全長が獲得できない場合、その配列は解析に利用しなかった。ジメチルラベルにより定量評価した試料は、H/L 比が 2.0 より大きいもののみを配列解析に利用した (H: キナーゼ反応試料, L: コントロール試料)。アミノ酸頻度プロットは WebLogo (v. 2.8.2)⁸⁷ を用いて描写した。モチーフ解析には Motif-X (v. 1.2) を用いた。その際、本文中で特筆しているもの以外はデフォルトのパラメータを利用した (width=15, occurrence=20, significance=0.000001, background=IPI HUMAN Proteome)。

9.3. 第二章

ペプチドの面積は Mass Navigator v1.2 により積分して算出し、H/L 比を計算した。第三節では、大規模 *in vitro* 基質プロファイルの重複のあったリン酸化部位を取り出した。ピークが一部のファイルのみで同定された場合、ペプチドが同定されていないファイル中にて、標的のペプチドの *m/z* 及び保持時間を基にピークの再同定を行った。これらのピーク形状およびスペクトルは、全て手動で確認した後、それぞれの面積から H/L 比を計算した。この内、阻害薬による減少、および、活性化薬による増加が見られたリン酸化部位を *in vivo* 基質候補とした。第四節では、モチーフ配列の選択性を計算するため、各 LC-MS 測定で同定された全てのペプチド面積を自動定量した。その内、H と L のピーク形状の相似を示す QuanScore が 0.8 以上のペプチドのみを採用し、阻害薬での減少、または、活性化薬での増大が 2 倍より大きなペプチドを、変動のあったリン酸化ペプチドとし、その他のペプチドは変動なしとした。一リン酸化部位に対して定量値が複数対応する場合、その平均値をとり変動値とした。

10. ウェスタンブロッティング

HeLa S3 細胞は通常の DMEM で培養し、6 穴プレートの各ウェル中に 2.5×10^5 細胞をにまいた。上述と同じ手順で、薬剤刺激をあたえた。DMEM ごと薬剤を取り除いた後、SDS バッファー (62.5 mM Tris-HCl, 5% (v/v) スクロース、10% (v/v) 2-メルカプトエタノール、2%

(w/v) SDS、および 0.002% ブロモフェノールブルー) にて細胞を溶解した。等量の細胞溶解液を SDS ゲルにロードし、ランニングバッファー (25 mM Tris、192 mM グリシン、および 0.1% SDS) 中で 1 ゲルあたり 20 mA にて泳動した。ウエスタンブロッティング用の PVDF 膜はメタノールに浸した後、蒸留水に浸し、最後に転写バッファー (25 mM Tris、192 mM グリシン、および 20% (v/v) メタノール) に浸して親水化した。タンパク質を転写した後、ブロッキング試薬に浸漬して膜表面をブロッキングした。一次抗体と PVDF 膜をインキュベートした後に、膜を洗浄バッファー (20 mM Tris-HCl および 137 mM NaCl, pH 7.6) で洗浄した。続いて、西洋ワサビペルオキシダーゼ (horseradish peroxidase, HRP) が結合した抗ウサギ IgG 二次抗体を用いてインキュベートを行い、膜を洗浄した。その後、ルミノールを含む検出液を用いて、抗体結合部位を検出した。CCD (Charged Coupled Device) カメラ (ImageQuant LAS 4000) にて撮影後、ImageJ (1.46r, National Institute of Health) を用いた画像解析により発光強度を定量化した。

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.semdb.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. Hsu, J.-L., Huang, S.-Y., Chow, N.-H. & Chen, S.-H. Stable-Isotope Dimethyl Labeling for Quantitative Proteomics. *Anal Chem* **75**, 6843–6852 (2003).
51. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4**, 484–494 (2009).
52. Ahrens, C. H., Brunner, E., Qeli, E., Basler, K. & Aebersold, R. Generating and navigating proteome maps using mass spectrometry. *Nat Rev Mol Cell Biol* **11**, 789–801 (2010).
53. Chou, M. F. *et al.* Using bacteria to determine protein kinase specificity and predict target substrates. *PLoS ONE* **7**, e52747 (2012).
54. Manning, B. D. & Cantley, L. C. Hitting the target: emerging technologies in the search for kinase substrates. *Sci STKE* **2002**, PE49 (2002).
55. Kettenbach, A. N. & Gerber, S. A. Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments. *Anal Chem* **83**, 7635–7644 (2011).
56. Chen, C. *et al.* Identification of a Major Determinant for Serine-Threonine Kinase Phosphoacceptor Specificity. *Mol Cell* (2013). doi:10.1016/j.molcel.2013.11.013
57. 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).
58. Biondi, R. M. & Nebreda, A. R. Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem J* **372**, 1–13 (2003).
59. 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).
60. Olsen, J. V. *et al.* Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **127**, 635–648 (2006).
61. Van Hoof, D. *et al.* Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell* **5**, 214–226 (2009).
62. 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).
63. 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).
 64. 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).
 65. Moritz, A. *et al.* Akt-RSK-S6 Kinase Signaling Networks Activated by Oncogenic Receptor Tyrosine Kinases. *Science signaling* **3**, ra64 (2010).
 66. Courcelles, M. *et al.* Phosphoproteome dynamics reveal novel ERK1/2 MAP kinase substrates with broad spectrum of functions. *Mol Syst Biol* **9**, 669 (2013).
 67. 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).
 68. García, E., Stracher, A. & Jay, D. Calcineurin dephosphorylates the C-terminal region of filamin in an important regulatory site: a possible mechanism for filamin mobilization and cell signaling. *Arch. Biochem. Biophys.* **446**, 140–150 (2006).
 69. Woo, M. S., Ohta, Y., Rabinovitz, I., Stossel, T. P. & Blenis, J. Ribosomal S6 kinase (RSK) regulates phosphorylation of filamin A on an important regulatory site. *Mol Cell Biol* **24**, 3025–3035 (2004).
 70. Vadlamudi, R. K. *et al.* Filamin is essential in actin cytoskeletal assembly mediated by p21-activated kinase 1. *Nat Cell Biol* **4**, 681–690 (2002).
 71. Tzircotis, G., Thorne, R. F. & Isacke, C. M. Directional sensing of a phorbol ester gradient requires CD44 and is regulated by CD44 phosphorylation. *Oncogene* **25**, 7401–7410 (2006).
 72. Soulsby, M. D. & Wojcikiewicz, R. J. H. The type III inositol 1,4,5-trisphosphate receptor is phosphorylated by cAMP-dependent protein kinase at three sites. *Biochem J* **392**, 493–497 (2005).

73. Butt, E. *et al.* Actin binding of human LIM and SH3 protein is regulated by cGMP- and cAMP-dependent protein kinase phosphorylation on serine 146. *J Biol Chem* **278**, 15601–15607 (2003).
74. Mihlan, S. *et al.* Nuclear import of LASP-1 is regulated by phosphorylation and dynamic protein-protein interactions. *Oncogene* **32**, 2107–2113 (2013).
75. Hunter, R. W., Mackintosh, C. & Hers, I. Protein kinase C-mediated phosphorylation and activation of PDE3A regulate cAMP levels in human platelets. *J Biol Chem* **284**, 12339–12348 (2009).
76. Pozuelo Rubio, M., Campbell, D. G., Morrice, N. A. & Mackintosh, C. Phosphodiesterase 3A binds to 14-3-3 proteins in response to PMA-induced phosphorylation of Ser428. *Biochem J* **392**, 163–172 (2005).
77. Lal, L. *et al.* Activation of the p70 S6 kinase by all-trans-retinoic acid in acute promyelocytic leukemia cells. *Blood* **105**, 1669–1677 (2005).
78. Roux, P. P. *et al.* RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *J Biol Chem* **282**, 14056–14064 (2007).
79. Tian, W. *et al.* Mechanism of regulation of group IVA phospholipase A2 activity by Ser727 phosphorylation. *J Biol Chem* **283**, 3960–3971 (2008).
80. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11**, 1475–1489 (1983).
81. Imami, K. *et al.* Temporal Profiling of Lapatinib-suppressed Phosphorylation Signals in EGFR/HER2 Pathways. *Mol Cell Proteomics* **11**, 1741–1757 (2012).
82. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896–1906 (2007).
83. Ong, S.-E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc* **1**, 2650–2660 (2006).

84. Yamana, R. *et al.* Rapid and deep profiling of human induced pluripotent stem cell proteome by one-shot NanoLC-MS/MS analysis with meter-scale monolithic silica columns. *J Proteome Res* **12**, 214–221 (2013).
85. Nakagami, H. *et al.* Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol* **153**, 1161–1174 (2010).
86. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13 (2009).
87. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190 (2004).