

5793-5798.

- 12) Hicke, L., Schubert, H.L., & Hill, C.P. (2005) *Nat. Rev. Mol. Cell Biol.*, 6, 610-621.
- 13) Lin, C.H., MacGurn, J.A., Chu, T., Stefan, C.J., & Emr, S.D. (2008) *Cell*, 135, 714-725.

十島 純子, 十島 二郎

(東京理科大学基礎工学部生物工学科)

Regulation of G protein-coupled receptor endocytosis via ubiquitination

Junko Y. Toshima and Jiro Toshima (Department of Biological Science and Technology, Tokyo University of Science, Yamazaki 2641, Noda, Chiba 278-8510, Japan)

投稿受付:平成21年10月14日

タンパク質モチーフ活性の予測とシグナル伝達解析

はじめに

タンパク質の機能は、タンパク質の修飾、分解、細胞内局在などの翻訳後制御によって調節されている。これらを制御するタンパク質は、標的タンパク質内の短いモチーフ配列を目印として認識、相互作用することでその機能を発揮する。したがって、タンパク質が特異的モチーフを持つかどうか、翻訳後制御を受けるかどうかの決定因子となる。一般的に、単離されたモチーフはペプチドとして独立して機能することができ、モチーフを認識するタンパク質はモチーフの配列のみを認識して結合する。モチーフ配列は、3~10 アミノ酸からなる短い配列で、通常その中の2~3 アミノ酸のみがモチーフ機能に大きな働きをしている。モチーフは明確な高次構造をとらず（このため linear motif と呼ばれる）、モチーフの多くは、タンパク質の非構造領域 (intrinsically disordered region) に存在することが観察されている¹⁾。

モチーフとは対照的に、タンパク質ドメインは、30 アミノ酸以上からなる進化的によく保存された高次構造をとる配列単位である。一般的に、ドメイン-ドメイン間の相互作用は強いが、ドメイン-モチーフ間の相互作用は弱いものが多い。このことは、一過的なタンパク質-タンパク質相互作用が中心となるシグナル伝達ネットワークにおいて、ドメイン-モチーフ相互作用が有利に働くことを示している。タンパク質相互作用ネットワークにおいて、多く

のタンパク質との相互作用を持つネットワークハブとして働くタンパク質は、多くの相互作用モチーフを持つことが知られている。

進化的に保存されたドメインは、アミノ酸配列の相同性から容易に類似ドメインを同定できるが、モチーフをその配列から正確に予測することは困難である。その理由は、モチーフの長さが高い配列多様性を生ずることにある²⁾。このような多様な配列候補の中から、タンパク質が特異的モチーフを選択し、特異的機能を発揮する機構はどのようなものであろうか？ 現在考えられるタンパク質モチーフ認識機構は主に二つある。一つは、その認識に複数のモチーフが関わることで、結合親和性、特異性を高めていること。もう一つは、現在知られているモチーフ配列よりもより広い領域において、モチーフ内の各アミノ酸が様々なレベルでタンパク質相互作用に貢献していることによる。これらの理由から、モチーフ認識の特異性は現在の単純なコンセンサス配列では表すことはできない。まずこれら二つのモチーフ認識機構について簡単に解説し、モチーフ予測法の現状と、最近筆者らによって開発された定量的モチーフ予測法について概説する。

1. 複数のモチーフを介した相互作用

タンパク質リン酸化酵素は、リン酸化部位であるセリン、スレオニン、またはチロシン残基とその周辺配列を認識することで基質タンパク質を特異的にリン酸化する。したがって殆どのタンパク質リン酸化酵素は、特異的リン酸化配列を含むペプチドを基質にすることができる。一方、mitogen-activated protein kinase (MAPK) ファミリーやサイクリン依存性キナーゼ (CDK) ファミリーなどは、共に最小配列として (S/T)P を含む配列をリン酸化するが、これら酵素はリン酸化部位とは異なる基質上の配列ドッキングモチーフを認識することにより基質特異性を高めていることが知られている³⁾。例えば、MAPK が認識するドッキングモチーフの一つである D-site は、コンセンサス配列 (K/R)₂₋₃-X₁₋₆-φ-X-f (φ:疎水性アミノ酸) を持ち、ERK, JNK を含む多くの MAPK 基質がこのモチーフを含んでいる。一方、CDK は二量体として結合するサイクリンの種類によってリン酸化配列の特異性を変えることが観察されている。これは、サイクリンが基質に存在するドッキングモチーフ (RXL モチーフ) を認識することによる⁴⁾。

タンパク質の主要な核輸送体である importin α は、分子内に二つの核移行シグナル (NLS) 結合部位を持ち、それぞれ異なるクラスの monopartite NLS (クラス 1/2 および

クラス3)が結合する^{5,6)}。二つの塩基性アミノ酸クラスターを持つ bipartite NLS の場合には、二つの塩基性クラスターが、それぞれ不完全なクラス3およびクラス1/2 NLS を構成しており、両者が協調して importin α への結合親和性を高めている^{5,6)}。一方、importin β が認識する NLS (PY-NLS) は、三つの保存されたセグメントからなる比較的長い領域を必要とするが、高次構造をとらない linear motif としての性質を持っている。importin α のように PY-NLS の三つのセグメントのうちの一つだけで NLS として機能する例は未だ見出されていないが、それぞれのセグメントは importin β に対して幅広い結合親和性を持ち、このことが単純なコンセンサス配列では表しきれない PY-NLS の高い配列多様性を生む要因となっている⁷⁾。

タンパク質の核外輸送は、CRM1 (exportin) が核外移行シグナル (NES) を認識することによって仲介される。NES は疎水性アミノ酸の規則的な繰り返しからなる 8-10 アミノ酸のモチーフで、三つのクラスのコンセンサス配列が同定されている⁸⁾。通常これらの NES 配列は核外移行活性に十分であるが、Snurportin 1 に存在する NES はより広い領域を必要とする。CRM1-Snurportin 1 複合体の立体構造から明らかとなったのは、Snurportin 1 の NES とは別の領域が CRM1 との相互作用に関わっていることであった^{9,10)}。すなわち、不完全な親和性の弱い NES が他のモチーフと協調して CRM1 と結合することによって結合親和性を高めていると考えられる。

このように、複数のモチーフがモチーフ相互作用に関わる認識機構は、相互作用の親和性、特異性を高める働きをすると同時に、認識に関わるそれぞれのモチーフが至適配列を持つ必要性を減らし、より多様な配列を許容する結果となる。

2. モチーフ内の配列多様性

多くの機能的モチーフは、ドッキングモチーフ等を必要とせず単一の短いモチーフで機能を発揮する。一般的にコンセンサス配列で表されるモチーフのアミノ酸は、相互作用するタンパク質に存在するモチーフ結合に必須のホットスポットと呼ばれる領域と結合するが、ホットスポット以外の領域もモチーフとの弱い相互作用に関わっている。

我々が行った NLS の包括的アミノ酸置換解析の結果は、NLS コア配列 (K[K/R]X[K/R]) の周辺配列が配列特異的に活性へ大きな影響を与えていることを示している^{6,11)}。NLS の周辺配列のアミノ酸特異性は、コア配列のアミノ酸特異性よりも低い、コア配列だけでは NLS として機

A		Loc	Score	B	
PPPKRPRLD	Nuc	>10		RHA-NLS	
****L****	Nuc	8		RPPKMARYD	
****M****	Nuc	9		NS5A-NLS	
****C****	Nuc	7		PPRKKRTVV	
****A****	N+C	5			C
****G****	N+C	4			Score
*****L**	Nuc	7		LAARKLRST	2.5
*****F**	Nuc	8		*****T*	3.5
*****A**	Nuc	7		*****P***	3.5
*****C**	Nuc	10		*****P*T*	5.0
*****G**	Cyt	2		R*****	7.0
PAARKLRIT	Nuc	9		R*****T*	8.5
*****L**	Cyt	1		R*****P***	8.0
*****F**	Cyt	1		R*****P*T*	9.5

図1 NLSモチーフ内アミノ酸の核移行活性への相加的寄与 (A) 疎水性アミノ酸によって交換可能な古典的 NLS の塩基性コア配列。古典的 monopartite NLS (クラス2 NLS) は、周辺配列に下線で示される活性の高いアミノ酸を含むとき、コア配列 (K[K/R]X[K/R]) における+2および+4位の塩基性アミノ酸が疎水性アミノ酸によって置換可能となる。GUS-GFP レポーターを用いて測定された NLS 活性を score で示す。便宜的にレポーターの細胞内局在性は、score ≥ 7 が核 (Nuc)、7~3 が核と細胞質 (N+C)、 < 3 が細胞質 (Cyt) であることを示す。(B) コンセンサスコア配列を含まないクラス2 NLS: human RNA helicase A (RHA), hepatitis C virus nonstructural protein (NS5A) から同定された NLS は、活性の高い周辺配列と不完全な塩基性コア配列を含む。(C) NLS 内アミノ酸の独立的、相加的な活性への寄与。クラス2 NLS の三箇所の異なる位置に重複して導入されたアミノ酸置換の効果は、それぞれのアミノ酸置換効果のほぼ総和として表れる。文献6), 11) からの図を改変。

能するには不十分なため、核移行活性に必要な NLS-importin 結合親和性の閾値を超えるためには、コア配列の他に周辺配列の適切なアミノ酸の組み合わせが必要となる。特筆すべきは、周辺配列が活性の高い配列を持つとき、コア配列の塩基性アミノ酸が他のアミノ酸 (疎水性アミノ酸) と置換可能になることであり、モチーフはコア配列を必ずしも含む必要がないことを示している⁶⁾ (図1A, B)。これらの結果は、周辺配列とコア配列が組み合わされた様々な配列パターンがモチーフの配列多様性を生む大きな要因となっていることを物語っている。

3. モチーフの予測

(a) 現在用いられている予測法

現在、モチーフ配列の予測は、同定されたモチーフをアラインメントしたときのモチーフ内各位置におけるアミノ酸出現頻度をプロファイリングして予測に反映させる手法

が主流である¹²⁾。この手法は、ScansiteやPrositesでのモチーフ予測に用いられており、コンセンサス配列を基にした単純な配列探索法よりも優れている。しかし、この方法の欠点は、プロファイル作製に十分量のモチーフ配列のデータが用いられないと、統計的に有為なアミノ酸出現頻度を算出できないことと、モチーフ機能に負に影響を与えるアミノ酸を考慮することが困難なことである。さらに、文献情報から収集したモチーフ配列を基にしたものでは、強弱様々な活性を持つ配列を同列に扱わなければならない欠点があり、Scansiteのようにペプチドライブラリーから選択した配列を基にしたものであっても、強い相互作用を示す配列パターンが優先的に選択されて配列にバイアスがかかってしまうなどの欠点がある。これらの予測手法の他に、コンピューターによる学習理論を用いた手法が用いられているが、その予測の判断基準がブラックボックスとなってしまう欠点があり、従来法同様、予測プログラム作製に用いるモチーフ配列を文献情報に依存している点や、配列比較を基にしている点から、大きな予測精度の向上は得られていないのが現状である¹²⁾。

(b) 体系的アミノ酸置換解析に基づいた新たなモチーフ予測法

我々は、NLSのアミノ酸置換解析を進めていく過程で、NLSモチーフの各位置に導入したアミノ酸変異の影響がほぼ独立に、相加的に全体のNLS活性に及んでいることを見出した¹¹⁾(図1C)。このアミノ酸の機能的独立性は、モチーフが高次構造をとらないため、アミノ酸変異の影響がモチーフ全体に及ばないことが一つの要因と考えられる。このモチーフの性質は、モチーフ配列の各位置における20種類のアミノ酸それぞれの活性影響度(貢献度)を示すプロファイルを作製し、各位置の活性影響度を加算することにより、任意の配列のモチーフ活性を計算できることを示している。このモチーフ構成単位の活性を相加的に評価する手段は、タンパク質に結合する低分子薬剤の生物活性を予測する手段として用いられているQSAR(定量的構造活性相関)でも使われている。しかし、アミノ酸側鎖の物理化学的性質または結合タンパク質の立体構造情報から、モチーフ活性を定量化するのは現状では困難であり、モチーフに応じて各アミノ酸の活性効果を実験的に計測しなければならない。

この手法を用いて、全てのクラスのimportin α 依存的NLSについてアミノ酸置換解析に基づいたプロファイルを作製し、プロファイルのスコアを取り込んでペプチド断片のNLS活性を計算するプログラムcNLS Mapper (predic-

tor for the importin α/β pathway-specific NLSs: <http://nls-mapper.iab.keio.ac.jp/>で利用可能)が作製された¹¹⁾。各NLSクラスのテスト配列を用いて求めたcNLS Mapperの予測精度(約90%)は、既存のNLS予測ツール(PSORT II, PredictNLS)の予測精度を大きく上回り、酵母タンパク質から単離したNLSについて予測したスコアは、実測スコアに近いものであった¹¹⁾。このモチーフ活性予測計算手法は、(1)負に作用するアミノ酸の影響度を決定できることや、(2)定量的にモチーフ活性を計算できる、などの点で従来のアミノ酸出現頻度に基づいたプロファイルを用いる予測手法よりも、高い予測精度の達成を可能にしていると考えられる。既存のモチーフ予測ツールの多くは、モチーフとして機能するかどうかの判定を行うのみである。しかし、モチーフの活性レベルは様々であり、特に複数の異なるモチーフ活性が一つのタンパク質機能を制御する場合、それぞれのモチーフ活性のレベルがモチーフ間の機能バランスに大きく影響する。このため、多くのモチーフ活性に関わるシグナル伝達機構の解析を行うには、モチーフ活性を定量的に予測、計算することが重要となる。

(c) リン酸化によって制御されるNLSの予測

NLSは周辺配列のリン酸化によってimportin α との結合親和性が低下し、核移行活性が失われる例が多く報告されている。我々が作製したNLSプロファイルでは、塩基性コア周辺配列の酸性アミノ酸が位置依存的に活性を抑制する(bipartite NLSのリンカー領域では逆に活性化すること)を示している。このため、酸性アミノ酸によって活性が低下あるいは上昇する位置でのセリン、スレオニン、またはチロシンのリン酸化はNLS活性に大きな影響を及ぼすことが考えられる。そこで著者らはCDKによるリン酸化によって核移行制御を受けるタンパク質を酵母から予測する試みを行った。細胞周期を制御するCDKによるNLSのリン酸化は、細胞周期特異的に核-細胞質間の移行を制御することが知られており、出芽酵母では現在まで五つのタンパク質がCDKによって細胞周期依存的核移行制御を受けることが報告されている。CDKのリン酸化部位の至適コンセンサス配列は(S/T)PX(K/R)であるが、NLS配列内でこのコンセンサス配列を重複して含むNLSは、各NLSクラスで数パターンに限られてくる。これらのNLSパターンを持つ出芽酵母のタンパク質を検索し、実験的検証を行った結果、予測によりヒットしたタンパク質の約30%がCDKによって核移行制御されるタンパク質であった¹¹⁾。同定したタンパク質は、すでに報告されている六つのタンパク質全てと五つの新規タンパク質を含んでいた。

CDK リン酸化部位の予測をより精度高く行うことができれば、今回の CDK 制御 NLS の予測もより精度高く実施できると考えられる。以上の結果は、モチーフを高精度で予測することによって、複数の異なるシグナル伝達経路により制御されるモチーフ、タンパク質の予測が可能になることを示すものである。

(d) モチーフプロファイルを用いたペプチド阻害剤のデザイン

モチーフ内各アミノ酸が独立に、相加的にモチーフ活性に影響を与えるという性質は、アミノ酸置換解析に基づいて作製したプロファイルを用いることによって、モチーフ受容体に対して高い結合能を持つペプチド配列をデザインできることを示している。bipartite NLS プロファイルに示される、各位置において高い活性スコアを示すアミノ酸を各々選択することで、二つのペプチド配列がデザインされた¹³⁾。これらのペプチドは、importin α Δ IBB に対して約 2×10^{-12} M および 2×10^{-14} M の解離定数を持ち、プロファイル作製に用いた鋳型 NLS 配列の解離定数 (約 10^{-7} M) に比べ、それぞれ 5 万および 500 万倍高い結合親和性を示した。実際にこれらのペプチドを融合タンパク質として酵母および哺乳類培養細胞で発現させると、importin α/β 経路特異的な核移行活性の阻害、および細胞増殖の抑制が観察された。タンパク質リン酸化や分解に関わるタンパク質は、重要な創薬の標的となっていることから、関連するモチーフに対する本手法を用いたペプチド阻害剤の開発は、有用な治療薬の開発に繋がるかもしれない。

おわりに

アミノ酸出現頻度に基づいた従来のモチーフ予測法も、モチーフ内の各アミノ酸が独立に働くことを仮定している。このことから、アミノ酸活性に基づいた予測法は、リン酸化モチーフやドッキングモチーフを含めた、高次構造をとらない他の多くの一般的なモチーフに適用可能であると考えられる。モチーフの予測精度をさらに上げるためには、いくつかの改良が必要である。モチーフの多くがタンパク質の非構造領域に存在することから、非構造領域の予測を取り入れることでより精度を上げることができると考えられる。さらに、NES や SUMO 相互作用モチーフのように疎水性アミノ酸に富んだモチーフは、タンパク質の折り畳まれた内部領域や膜貫通ドメインに擬陽性配列を含む確率が高くなるため、これら領域の予測法も組み合わせる必要がある。

モチーフ活性を定量的に高精度で予測することが可能に

なると、モチーフの主な機能であるタンパク質修飾、分解や細胞内移行などを単に予測するだけでなく、それらが組み合わされて起こるシグナル伝達制御の定量的解析、モデリングが可能となる。タンパク質相互作用、タンパク質修飾、転写産物などについての網羅的実験データを組み合わせることにより、より包括的かつ高精度なタンパク質機能制御機構の予測が可能となるであろう。最終的に実験検証が必須であっても、可能性のある制御機構を推定できるのは、一般的な網羅的解析手段に比べてはるかに効率のよい解析手段である。さらに、モチーフは進化的保存度が低く、1塩基の変異によって容易にモチーフ機能が失われると同時に新たなモチーフが形成される潜在性を持っている。ゲノム情報や一塩基多型情報を基にして機能モチーフを高精度に予測することは、ヒトの疾患や動植物の量的形質にかかわる遺伝子、タンパク質を同定するための有効な手段となるかもしれない。

- 1) Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G., & Gibson, T.J. (2008) *Front. Biosci.*, **13**, 6580-6603.
- 2) Neduva, V. & Russell, R.B. (2005) *FEBS Lett.*, **579**, 3342-3345.
- 3) Reményi, A., Good, M.C., & Lim, W.A. (2006) *Curr. Opin. Struct. Biol.*, **16**, 676-685.
- 4) Ubersax, J.A. & Ferrell, J.E., Jr. (2007) *Nat. Rev. Mol. Cell Biol.*, **8**, 530-541.
- 5) Conti, E., Uy, M., Leighton, L., Blobel, G., & Kuriyan, J. (1998) *Cell*, **94**, 193-204.
- 6) Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., & Yanagawa, H. (2009) *J. Biol. Chem.*, **284**, 478-485.
- 7) Süel, K.E., Gu, H., & Chook, Y.M. (2008) *PLoS Biol.*, **6**, e137.
- 8) Kosugi, S., Hasebe, M., Tomita, M., & Yanagawa, H. (2008) *Traffic*, **9**, 2053-2062.
- 9) Dong, X., Biswas, A., Süel, K.E., Jackson, L.K., Martinez, R., Gu, H., & Chook, Y.M. (2009) *Nature*, **458**, 1136-1141.
- 10) Monecke, T., Güttler, T., Neumann, P., Dickmanns, A., Görlich, D., & Ficner, R. (2009) *Science*, **324**, 1087-1091.
- 11) Kosugi, S., Hasebe, M., Tomita, M., & Yanagawa, H. (2009) *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 10171-10176.
- 12) Fox-Erlich, S., Schiller, M.R., & Gryk, M.R. (2009) *Front. Biosci.*, **14**, 1143-1151.
- 13) Kosugi, S., Hasebe, M., Entani, T., Takayama, S., Tomita, M., & Yanagawa, H. (2008) *Chem. Biol.*, **15**, 940-949.

小杉 俊一^{1,3}, 柳川 弘志²
(¹ 奈良先端科学技術大学院大学
バイオサイエンス研究科)
(² 慶應義塾大学理工学部生命情報学科)
(³ 現所属先: 岩手生物工学研究所)

School of Biological Sciences, Nara Institute of Science and
Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192,
Japan; ²Department of Biosciences and Informatics, Faculty
of Science and Technology, Keio University, 3-14-1 Hi-
yoshi, Kohoku-ku, Yokohama 223-8522, Japan; ³Present af-
filiation: Iwate Biotechnology Research Center, 22-174-4
Narita, Kitakami, Iwate 024-0003, Japan)
投稿受付: 平成 21 年 12 月 1 日

Prediction of protein motif activities and signal transduction
research

Shunichi Kosugi^{1,3} and Hiroshi Yanagawa² (¹Graduate