

## 博士論文

論文題目     NMR分析と計算科学的手法に基づく分子間相互作用  
部位の迅速な解析法の開発

氏     名     小玉 優哉

NOTICE:

This document contains the Accepted Manuscript version of a Published Work that appeared in final form in Journal of Medicinal Chemistry, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see <http://pubs.acs.org/articlesonrequest/AOR-ufIYPjrmngFcJBdJ3I2J>.

This document also contains the author's version of a work that was accepted for publication in Journal of Structural Biology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms, may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Structural Biology, 174, 3, June 2011, DOI: 10.1016/j.jsb.2011.04.001.



# 目次

目次	1
略語	3
<b>第1章 序論</b>	
1.1 背景および先行技術とその問題点	4
1.2 本研究の概要	6
<b>第2章 タンパク質—タンパク質間相互作用部位の迅速解析法の開発</b>	
2.1 原子特異的標識の利用による迅速解析法の実現	7
2.2 二種の異種核相関を同時に観測するパルスシーケンスの開発	10
2.3 相互作用に伴う化学シフト変化の大小判定におけるバイアスの排除	10
2.4 ユビキチンおよびユビキチン加水分解酵素を用いた本手法の有効性の実証	15
<b>第3章 計算科学的手法の組み合わせによる手法の高度化</b>	
3.1 相互作用部位同定の自動化と任意性の排除	19
3.2 既知のタンパク質複合体構造を用いたシミュレーションによる手法の有効性検証	20
3.3 プログラムを用いた YUH および Y-Ub の相互作用界面の同定	20
3.4 手法の頑健性の検証	27
3.5 実験結果を反映させた確からしいタンパク質複合体モデル構造の構築	28
<b>第4章 タンパク質—低分子間相互作用部位解析への手法の拡張</b>	
4.1 タンパク質—タンパク質間相互作用部位解析法の応用と課題	30
4.2 三次元空間情報を利用したタンパク質—低分子間相互作用部位解析法の開発	30
4.3 手法の頑健性向上	33
4.4 既知の複合体構造を用いたシミュレーションによる手法の有効性検証	33
4.5 p38 $\alpha$ mitogen-activated protein kinase と阻害剤を用いた本手法の有効性の実証	35
<b>第5章 天然リガンドの相互作用部位同定への応用</b>	
5.1 MAP kinase insert region および脂質に関する先行知見と課題	39
5.2 p38 $\alpha$ mitogen-activated protein kinase の脂肪酸相互作用部位の同定	39
5.3 他のアプローチとの比較	40
<b>第6章 総括および今後の展望</b>	

6.1 タンパク質—タンパク質間相互作用部位の迅速解析法と課題	43
6.2 タンパク質—低分子間相互作用部位の迅速解析法と課題	45
6.3 結論	47

## 第7章 材料と方法

7.1 試薬および立体構造表示ソフトウェア	48
7.2 ユビキチンおよびユビキチン加水分解酵素の発現、精製および安定同位体標識	48
7.3 ユビキチンとユビキチン加水分解酵素との相互作用実験	49
7.4 ドッキングシミュレーションによる複合体モデル構造の構築	50
7.5 p38 $\alpha$ mitogen-activated protein kinase の合成、精製および変異導入	50
7.6 p38 $\alpha$ mitogen-activated protein kinase と阻害剤、脂肪酸との相互作用実験	51
7.7 タンパク質周辺へのグリッドの発生	52

参考文献	53
------	----

謝辞	64
----	----

付録	65
----	----

## 略語

$\beta$ -OG: *n*-octyl- $\beta$ -D-glucopyranoside

BMRB: Biological Magnetic Resonance Bank

DTT: 1,4-dithiothreitol

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPLC: High performance liquid chromatography

HSQC: Heteronuclear single quantum coherence

IPTG: Isopropyl- $\beta$ -D-thiogalactopyranoside

LC-MS: Liquid chromatography – Mass spectrometry

MAP: mitogen-activated protein

MAPK: mitogen-activated protein kinase

MD: Molecular dynamics

NMR: Nuclear magnetic resonance

PCR: Polymerase chain reaction

PDB: Protein Data Bank

PIA: Phosphatidylinositol ether lipid analog

PMSF: Phenylmethylsulfonyl fluoride

PPI: Protein-protein interaction

RDC: Residual dipolar coupling

SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

TEV: Tobacco etch virus

Tris: Tris(hydroxymethyl)aminomethane

TROSY: Transverse relaxation optimized spectroscopy

Ub: Ubiquitin

Y-Ub: Yeast ubiquitin

YUH: Yeast ubiquitin hydrolase

## 第 1 章 序論

### 1.1 背景および先行技術とその問題点

タンパク質—タンパク質間相互作用 (PPI) は、生体の代謝やシグナル伝達をはじめとした種々の現象に関与していることから、相互作用界面に関する情報の取得は様々な生物学的、生化学的現象の理解、ひいては創薬研究において極めて重要である。近年の構造ゲノム研究の進展により、タンパク質の立体構造解析数は飛躍的に増加したものの、その多くはタンパク質単体での構造であり、興味のあるタンパク質複合体について各単体の構造は解析されていてもその複合体構造が解析されていることは少なく、相互作用様式は不明なままである場合が多い。

一般に相互作用界面に関する情報を得るためには、例えば X 線結晶構造解析においては結晶化の作業、NMR 解析においては NMR シグナルを帰属するための測定、解析作業が必要となる。これらは、いずれも多大な労力と時間を必要とし、例えば PPI に関する情報から創薬研究を展開するうえで課題となっていた。特に NMR は、溶液状態での相互作用情報を原子レベルで取得できる強力な手法であるが<sup>1-6</sup>、近年、解析対象とするタンパク質の分子量が大きくなるにつれ、解析に必要な時間は加速度的に増加していた。この問題の解決策の 1 つとして、解析対象のタンパク質単体の立体構造が既知であれば、NMR シグナル未帰属のまま相互作用界面の同定が可能となる手法が Reese らによって報告<sup>7</sup>された。その概要は、以下の通りである (図 1)。

- (1) 1 種類のアミノ酸残基種のみ <sup>15</sup>N 標識されたタンパク質を、標識するアミノ酸残基種を変えて多種類調製する。
- (2) 各試料について、相互作用対象を滴定し相互作用実験を行う。
- (3) 化学シフト変化を観測し、有意な化学シフト変化が観測されたシグナルの個数を、アミノ酸残基種ごとに数える。これらの化学シフト変化は、相互作用に伴う化学的環境の変化に起因していることから、各アミノ酸残基種について、先に数えた個数が相互作用界面に存在していると考えられる。
- (4) タンパク質立体構造の表面において、実験結果に最も一致する箇所が相互作用界面として同定される。

上記の手法は、シグナルの帰属を必要とする従来の解析法に比べ簡便ではあるものの、多数の標識サンプルと大量の相互作用対象のタンパク質を調製する必要があり、また相互作用実験を多数行う必要があった。したがって、結果として未だ多くの時間と労力を必要としていた。さらに、化学シフト変化の大小を判断する際に用いる閾値の設定には、観測対象とするアミノ酸残基種を限定することによるバイアスが含まれていた。加えて、実験結果を基にした相互作用界面の探索は目視で行われており、得られる結果は解析者の主観に依るところが大きかった。

PPI 以外にもタンパク質—化合物間相互作用の情報も極めて重要であり、例えば、天然リガンドの生化学的意義を分子レベルから理解する際はもちろんのこと、創薬においては化合物設計やバリデーションにおいて重要な役割を果たしている。特に、近年のハイスループットスクリーニング技術と *in silico* 技術の進展により、巨大なライブラリから活性化合物を迅速かつ効率的に見出だ

することが可能となったが<sup>8,9</sup>、化合物と標的タンパク質との相互作用、特に結合部位に関するバリデーションは、依然として別の実験により行う必要がある。NMR は、先に挙げた点からバリデーション法としても非常に有用である一方、やはり前述の通り、NMR シグナルの帰属が課題であり、特にタンパク質の分子量が大きい場合において顕著であった。また、NMR 以外の解析手法においても、相互作用部位を同定するための簡便な手法は未だ開発されていなかった。

## 1.2 本研究の概要

本研究では、前述の課題を解決するための新たな方法論を開発した。すなわち、解析対象とするタンパク質単体の立体構造が既知であれば、NMRシグナルを帰属することなく、1つのサンプルに関する実験のみからPPI界面の同定を可能とした。その際、1.1で挙げた閾値の設定について、

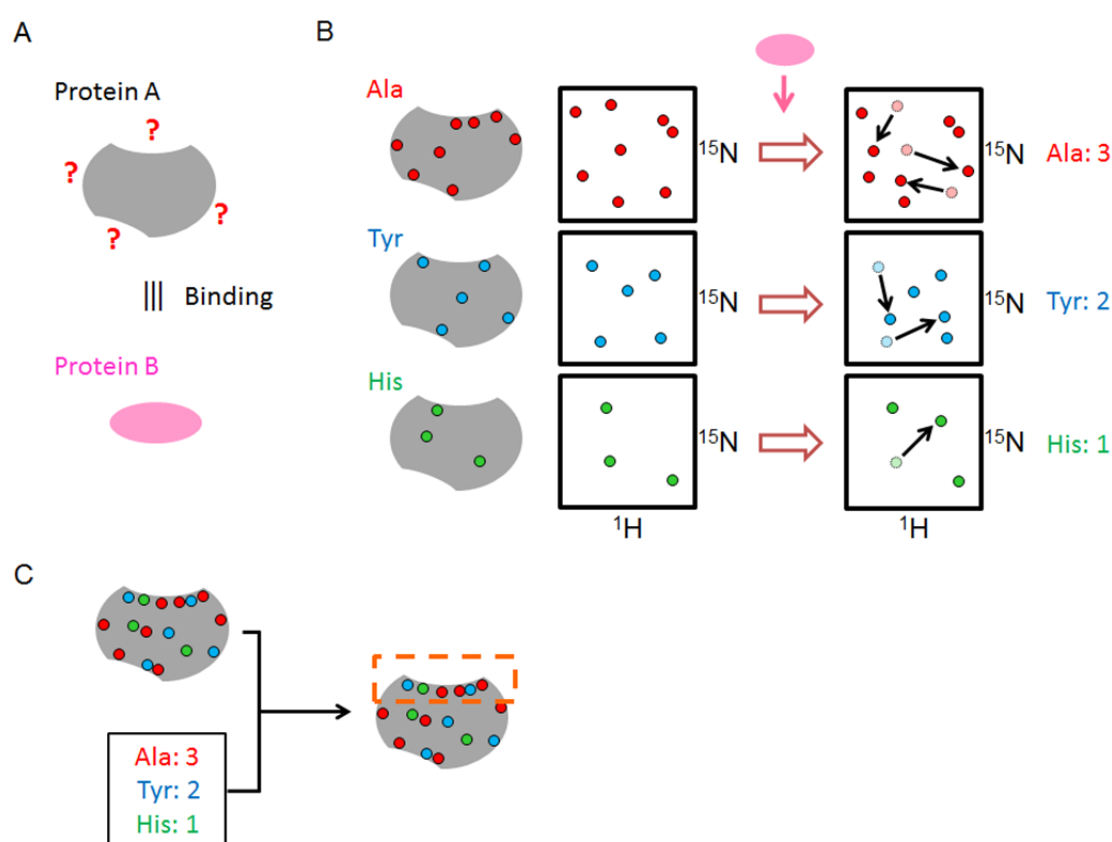


図 1. 既報のタンパク質—タンパク質間相互作用部位解析法の概要。立体構造が既知である Protein A の、Protein B との相互作用界面が不明である場合を例とする(A)。まず、Protein A について、例えば Ala, Tyr, His のいずれか 1 つのみを  $^{15}\text{N}$  で標識したサンプルを調製する(B 左)。次に、各サンプルについて Protein B の滴定実験を行い、化学シフト変化を観測し、有意な化学シフト変化が観測されたシグナルの個数をカウントする(B 右)。Protein A の立体構造表面において、実験結果に最も一致する、つまり各アミノ酸残基種がこれらの個数存在している箇所が相互作用界面と同定される(C)。

観測対象とするアミノ酸残基種によるバイアスを除いた値が一意に決まるよう、設定法を規定した。また、実験結果とタンパク質立体構造情報に基づき PPI 界面を自動で決定するプログラムを開発することで、解析者の主観を排除し、一義的に PPI 界面を同定することが可能になった。さらに、新たなパルスシーケンスを開発し本手法に適用することで、測定時間を短縮し、かつ高解像度のスペクトルによる解析が可能となった。加えて、得られた PPI 界面の情報をを用いたドッキングシミュレーションにより、実験結果に基づく確からしい複合体モデル構造を迅速に取得する方法を構築した。これらの方法論の有効性を、複合体構造が既知である yeast ubiquitin hydrolase (YUH)と ubiquitin (Ub)を用いて実証した。

一方で、PPI とタンパク質—化合物間相互作用とではその様式に差異があるため、前記の方法論では化合物の相互作用部位を同定することは困難であった。そのため、3 次元空間情報を用いた方法論へと手法を拡張し、タンパク質—化合物間相互作用部位の解析法を新たに開発した。本方法論が正しく機能することを、複合体構造が既知である p38 $\alpha$  とその阻害剤を用いて実証した他、p38 $\alpha$  の天然リガンドである可能性が報告されているデカン酸との相互作用部位を初めて明らかにし、本手法の有効性を示した。

## 第2章 タンパク質—タンパク質間相互作用部位の迅速解析法の開発

### 2.1 原子特異的標識の利用による迅速解析法の実現

#### 2.1.1 原子特異的標識

Reese らにより開発された手法<sup>7</sup>では、1つのアミノ酸残基種のみを選択的に標識したサンプルを調製する。これは、シグナル未帰属のまま特定のアミノ酸残基種のみに関する情報の取得を可能にすると同時に、シグナルの縮重を回避し、大きなタンパク質の解析を可能とする点で有効なアプローチである。一方で、観測したいアミノ酸残基種の数だけサンプルが必要となるためサンプル調製に大きな労力を必要とし、また、複数のサンプルについての実験が必要なため、相互作用対象のタンパク質も大量に調製する必要がある。したがって、発現量が少ない等、サンプル調製が実験の進展において大きな律速となる場合は、事実上適用が不可能である。さらに、複数のサンプルについてそれぞれ滴定実験を行い、結果を解析する必要がある。このように、既報の手法を適用するにあたっては、実際には未だ多くの時間と労力が必要である。

既報における上記の課題は、複数のサンプルを必要とするという点に集約される。したがって、課題解決のためには、1つのサンプルのみで同様の情報を取得できればよい。既報において複数のサンプルを必要とした理由は、複数のアミノ酸残基種に由来するシグナルが $^1\text{H}$ - $^{15}\text{N}$  相関スペクトル上で同時に観測されると、各シグナルがどのアミノ酸残基種に由来するか判別が不可能なためである。そこで本研究においては、 $^1\text{H}$ - $^{15}\text{N}$  の相関ではなく、 $^1\text{H}$ - $^{13}\text{C}$  の相関の利用に着目した。各アミノ酸残基の $^{13}\text{C}$  およびその $^{13}\text{C}$  と共有結合を有する $^1\text{H}$ の化学シフト値は、それぞれ一定の

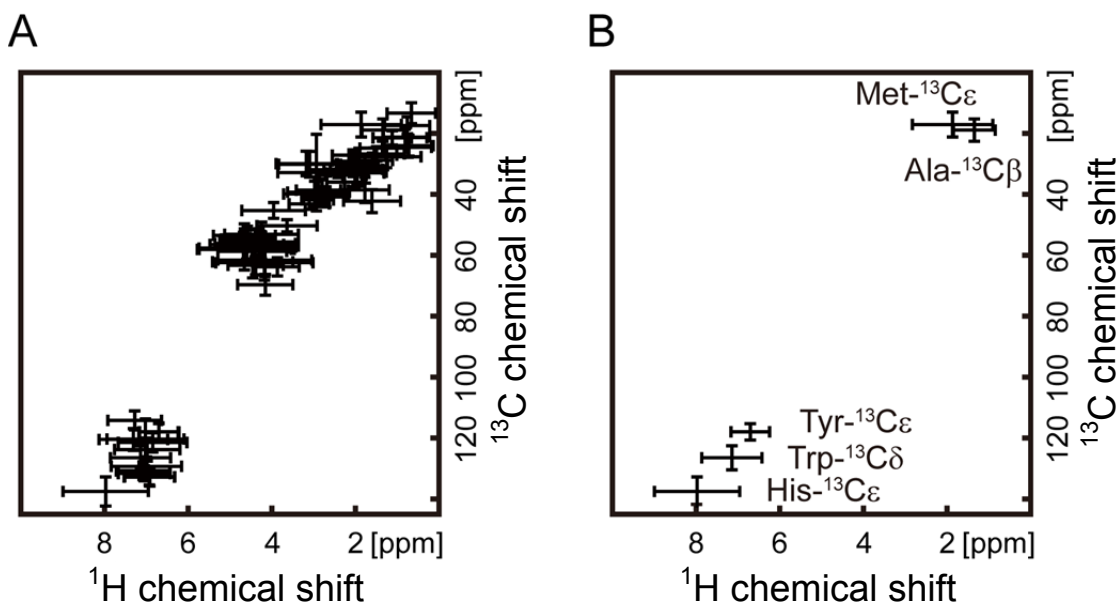


図 2.1. タンパク質中の各  $^1\text{H}$ 、 $^{13}\text{C}$  がとり得る化学シフト値の分布。既報の統計値を基に、平均値 $\pm 2 \times$ 標準偏差で表した。(A) 全  $^1\text{H}$ - $^{13}\text{C}$  相関シグナルの分布。(B) 本研究にて用いた標識の組み合わせ。

範囲に局在している。したがって、 $^1\text{H}$ - $^{13}\text{C}$  相関スペクトル上において、各シグナルは各アミノ酸残基種の原子タイプ ( $^{13}\text{C}_x$  ( $x=\alpha, \beta, \gamma, \dots$ )) ごとにスペクトル上に局在するため、化学シフト値の分布、つまりスペクトル上においてシグナルが観測される範囲が重ならなければ、各シグナルがどのアミノ酸残基種に由来するかが同一スペクトル上で判別できる。このような組み合わせにて安定同位体標識されたタンパク質を調製することで、1 サンプルのみを調製し、その 1 サンプルについての実験のみから必要な情報の取得が可能となる。具体的にはまず、原子タイプごとの化学シフト値の分布を、Biological Magnetic Resonance Bank (BMRB, <http://www.bmrb.wisc.edu/>) より入手した既報の化学シフト値に関する統計値を基に、平均値  $\pm 2 \times$  標準偏差とした (図 2.1A)。これらの中から化学シフト値の分布が重ならず、また PPI への寄与が高い、いわゆる hot spot である傾向<sup>10, 11</sup>と相互作用界面への存在傾向<sup>12, 13</sup>の高さを指標に、標識の組み合わせを選択した。さらに、実用性を鑑み、市販されている安定同位体標識アミノ酸のみを用いることにした。以上の観点から、標識の組み合わせとして [ $^{13}\text{C}_\epsilon$ ] Met、[ $^{13}\text{C}_\epsilon$ ] His、[ $^{13}\text{C}_\epsilon$ ] Tyr、[ $^{13}\text{C}_\delta$ ] Trp を選択した (図 2.1B)。一方で、これらのアミノ酸残基種はタンパク質中の存在数が比較的少なく、タンパク質表面全体をカバーするには不足があるため、タンパク質表面に広く分布する Ala 由来のシグナルも観測することにした。なお、メチル基はその化学的性質から、 $^1\text{H}$ - $^{13}\text{C}$  相関シグナルがシャープで一般にスペクトル上での分散がよいため、Ala の標識部位としては [ $^{13}\text{C}_\beta$ ] が望ましい。しかしながら、Ala のメチル基に由来する  $^1\text{H}$ - $^{13}\text{C}$  相関シグナルの化学シフト値の分布は、Met のメチル基でのそれと重なっており、スペクトル上で区別できない可能性がある (図 2.1B)。そのため、Ala の標識を [ $^{13}\text{C}_\alpha$ / $^{13}\text{C}_\beta$ ] Ala とすることで、カップリングにより Met 由来のシグナルと区別できるようにし、シグナル数が多く解析に支障をきたす場合には  $^{13}\text{C}_\alpha$  をデカップリングした。また、Constant-time  $^1\text{H}$ - $^{13}\text{C}$  HSQC 測定の実用によっても、Met 由来と Ala 由来との区別が可能である (図 2.2)。

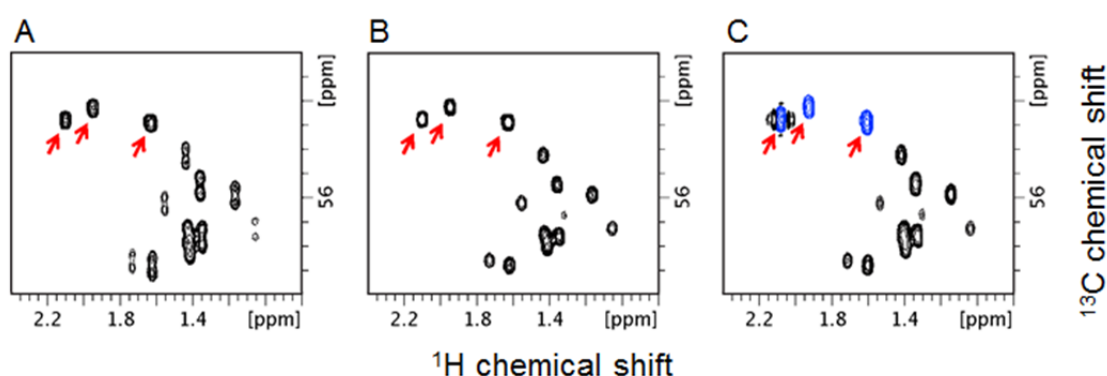


図 2.2. カップリングまたは Constant-time 測定による Ala 由来シグナルと Met 由来の区別の例。(A)  $^1\text{H}$ - $^{13}\text{C}$  HSQC (デカップリングなし)。(B)  $^1\text{H}$ - $^{13}\text{C}$  HSQC (デカップリングあり)。(C) Constant-time  $^1\text{H}$ - $^{13}\text{C}$  HSQC。黒と青は位相が  $180^\circ$  異なる。YUH のスペクトルを例に用いた。赤矢印は Met 由来シグナルを示す。



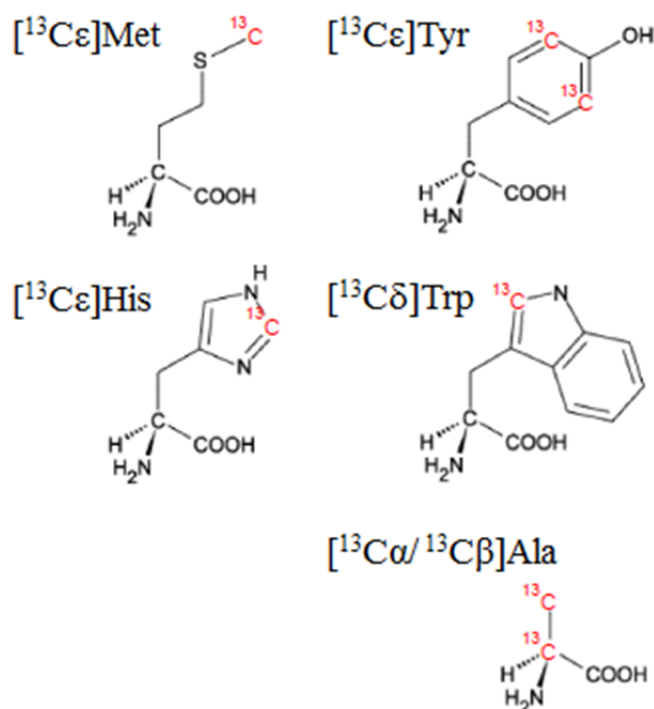


図 2.3. 本研究にて用いた安定同位体標識アミノ酸。これらの他に、アミノ基が  $^{15}\text{N}$  標識された Arg を用いた。

前記では  $^1\text{H}$ - $^{13}\text{C}$  相関シグナルの観測について述べたが、1 種類のアミノ酸残基については、 $^1\text{H}$ - $^{15}\text{N}$  相関シグナルの観測による情報の取得が可能である。本研究では、観測プローブのタイプにバリエーションをもたせるため、荷電性側鎖を有するアミノ酸残基を  $^{15}\text{N}$  標識の対象とし、hot spot である傾向と相互作用界面への存在傾向の高さから、最終的に Arg を選択した。なお、 $^{15}\text{N}$  標識するアミノ酸残基種には制限がないため、これまでに得られている知見や解析対象の系に合わせて、任意のアミノ酸残基種を選択してもよい。

以上のように、原子特異的標識とその組み合わせを利用することで、 $^{13}\text{C}$  標識したアミノ酸残基 5 種(図 2.3)と  $^{15}\text{N}$  標識したアミノ酸残基 1 種(本研究では Arg)の、計 6 種類のアミノ酸残基種に関する情報を 1 サンプルのみで取得可能となった。

## 2.1.2 セルフリー合成系を用いた安定同位体標識タンパク質の調製

安定同位体で標識されたタンパク質を調製する際には、一般に大腸菌や昆虫細胞などを用いた組換え発現系が利用されてきたが、近年はセルフリー合成系の利用<sup>14-25</sup>も盛んである。セルフリー合成系は、反応液、特にライセートの調製に技術と労力を要していたが、各社から合成キットが市販されるようになり、非常に簡便ながら高効率でタンパク質を合成することが可能になった。特に、小麦胚芽由来のライセートを用いることで、これまでは調製が困難であった哺乳類動物由

来タンパク質なども大量調製が可能となり、様々な研究に用いられている。組換え発現系と比較した際のセルフリー合成系の利点としては、(1)アミノ酸の代謝に伴う標識のスクランブリングが起りにくい、(2)必要とするアミノ酸の量が少ないことが挙げられる。(1)は、特異的標識を行う上で極めて重要であり、多くのスクランブリングが確認されている組換え発現系では実現不可能な利点である。(2)は、標識アミノ酸、特に原子特異的標識アミノ酸は高価であるため、実用上の観点から重要である。これらの利点から、本研究ではセルフリー合成系を用いて標識アミノ酸を調製した。反応液に使用するライセートは、目的タンパク質の合成が可能でかつ最も合成の効率が高い、大腸菌由来のものを使用した。

## 2.2 二種の異種核相関を同時に観測するパルスシーケンスの開発

2.1 に記載したように、本手法では  $^1\text{H}$ - $^{13}\text{C}$  相関シグナルと  $^1\text{H}$ - $^{15}\text{N}$  相関シグナルの両方を観測するため、一般には2種類の測定を行う必要がある。一方で、time-sharing 実験により両相関シグナルを1つのスペクトル上で同時に観測できることが報告されている<sup>26</sup> ため、本研究へも適用し測定時間の短縮を図ることにした。また、大きなタンパク質を解析対象とする場合、芳香族アミノ酸残基に由来するシグナルの広幅化が間接観測軸方向において特に著しく<sup>27</sup> 解析が困難になることから、transverse relaxation optimized spectroscopy (TROSY)法の適用が必須である。そこで、 $^1\text{H}$ - $^{13}\text{C}$  相関シグナルと  $^1\text{H}$ - $^{15}\text{N}$  相関シグナルを同スペクトル上で同時に観測でき、かつ芳香環の  $^1\text{H}$ - $^{13}\text{C}$  相関シグナルおよびタンパク質主鎖の  $^1\text{H}$ - $^{15}\text{N}$  相関の間接観測軸方向について TROSY 法を適用した、新たなパルスシーケンスを開発した(図 2.4)。なお、 $^1\text{H}$ - $^{13}\text{C}$  相関シグナルと  $^1\text{H}$ - $^{15}\text{N}$  相関シグナルとで位相を反転させているため、同スペクトル上で互いを容易に区別可能である。この測定手法を適用することで、 $^1\text{H}$ - $^{13}\text{C}$  相関シグナルと  $^1\text{H}$ - $^{15}\text{N}$  相関シグナルの同時測定を高分解能で行うことが可能となった。

## 2.3 相互作用に伴う化学シフト変化の大小判定におけるバイアスの排除

### 2.3.1 化学シフト値の変化量のスケールリング

NMR を利用した分子間相互作用解析において、相互作用により誘起される化学シフト変化の観測は多用される実験手法の 1 つであり、各研究において化学シフト値の変化量は定量的に解析され、その変化量の大小から相互作用についての議論が行われている。化学シフト値の変化量は、核種によりそのスケールが異なるため、異種核相関スペクトル上の各シグナルの変化量を解析する際には、一般に変化量のある定数(例えば各核の磁気回転比)で除算しスケールリングすることでその寄与率を揃えている。本研究では既報のスケールリング手法<sup>28</sup>を採用し、タンパク質中の各アミノ酸残基種の原子ごとに、BMRB より入手した化学シフト値の分布の標準偏差(表 1)を除算に用いる定数として使用した。これらのスケールリングした値の二乗平均平方根を、異種核相関スペクトルにおける各シグナルの化学シフト値の変化量  $\omega_{\text{RMS}}$  として、式 1 に従い求めた。

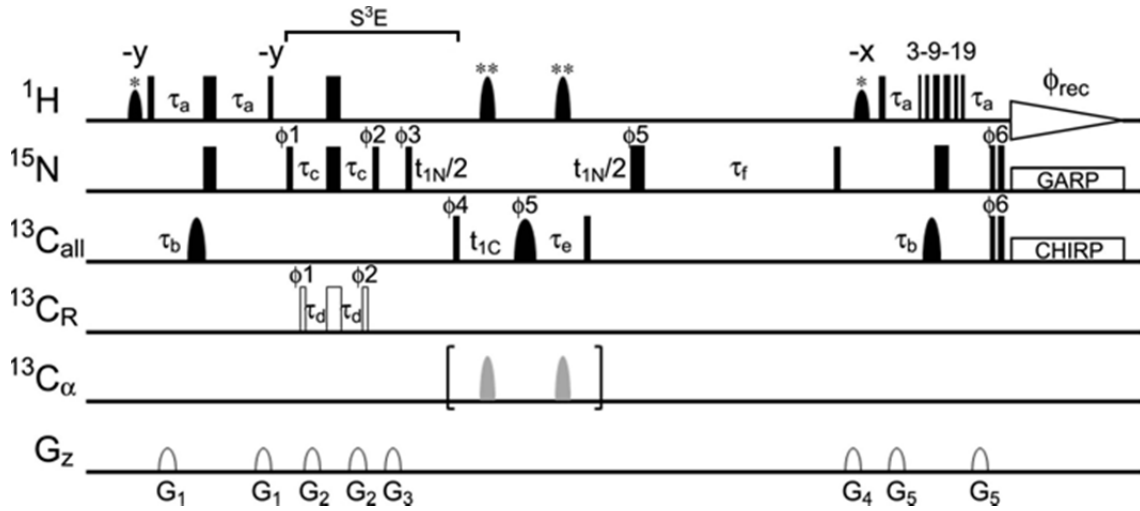


図 2.4. 新たに開発した time-shared  $[^1\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}]-\omega_1\text{-TROSY-HSQC}/[^1\text{H}, ^{13}\text{C}_{\text{methyl}}]\text{-HSQC}$  実験用パルスシーケンスのダイアグラム。特に記載がない限り、幅が狭い(広い)四角は、それぞれ X 軸方向に沿って回転させる  $90^\circ(180^\circ)$  パルスを示す。 $^1\text{H}$ 、 $^{15}\text{N}$ 、 $^{13}\text{C}_{\text{all}}$ 、 $^{13}\text{C}_{\text{R}}$ 、および  $^{13}\text{C}_{\alpha}$  の中心周波数は、それぞれ 4.7、118、55、128、および 55ppm とした。アスタリスクで示すパルスは、水分子の  $^1\text{H}$  選択的な sinc 関数型の  $90^\circ$  パルスであり、照射時間は 1 ms である。2 つのアスタリスクで示すパルスは、メチル基の  $^1\text{H}$  選択的な反転を行うための re-burp パルス<sup>29</sup>を示し、中心周波数は -1ppm、照射時間 1 ms である。 $^{15}\text{N}$  用の  $90^\circ$  および  $180^\circ$  パルス、 $^{13}\text{C}_{\text{all}}$  の  $90^\circ$  パルスの field strength 値は、それぞれ 5.6 kHz および 16.6 kHz とした。 $\tau_b$  の後の  $180^\circ$  シェイプドパルスは照射時間 500  $\mu\text{s}$  の断熱パルス<sup>30</sup>、 $\phi_5$  フェーズにある  $180^\circ$  シェイプドパルスは照射時間 2 ms の複合断熱リフォーカスパルス<sup>31</sup>である。白四角で示す  $^{13}\text{C}_{\text{R}}$  の  $90^\circ(180^\circ)$  パルスの field strength 値は、それぞれ  $\Delta/\sqrt{15}$  ( $\Delta/\sqrt{3}$ ) とした<sup>32</sup> ( $\Delta$  は、 $^{13}\text{C}_{\alpha}$  (55ppm) と  $^{13}\text{C}_{\text{R}}$  (128 ppm) の化学シフトの差(単位: Hz)を示す)。 $^{13}\text{C}_{\alpha}$  のデカップリングが必要な場合は、灰色で示す  $180^\circ$  シェイプドパルスとして G3 パルスを 750  $\mu\text{s}$  照射する。測定時に  $^{13}\text{C}$  および  $^{15}\text{N}$  をデカップリングするために CHIRP<sup>33, 34</sup> および GARP-1<sup>35</sup> を適用し、field strength 値はそれぞれ 4.4 kHz、1.0 kHz とした。各ディレイは、それぞれ  $\tau_a = 2.25$  ms、 $\tau_b = 1.50$  ms、 $\tau_c = 1.37$  ms、 $\tau_d = 0.78$  ms、 $\tau_e = t_{1C}(0)$ 、 $\tau_f = t_{1N}(0) + 2t_{1C}(0) + \text{pw}(^{13}\text{C})$  とした( $t_{1C}(0)$  は最初の  $t_{1C}$  ディレイ、 $t_{1N}(0)$  は最初の  $t_{1N}$  ディレイ、 $\text{pw}(^{13}\text{C})$  は  $t_{1C}$  で展開する際に関連する  $^{13}\text{C}$  パルスの総時間を意味する)。各位相サイクルは、それぞれ  $\phi_1 = [\pi/4, 5\pi/4]$ 、 $\phi_2 = [-y, y]$ 、 $\phi_3 = [2(-x), 2(x)]$ 、 $\phi_4 = [2(x), 2(-x)]$ 、 $\phi_5 = [4(x), 4(y)]$ 、 $\phi_6 = [8(x), 8(-x)]$ 、 $\phi_{\text{rec}} = [2(x), 4(-x), 2(x)]$  とした。各磁場勾配パルスの照射時間と強度は、それぞれ  $G_1 = (1$  ms, 8.6 G/cm)、 $G_2 = (300$   $\mu\text{s}$ , 11.8 G/cm)、 $G_3 = (1$  ms, 34.8 G/cm)、 $G_4 = (1$  ms, -19.3 G/cm)、 $G_5 = (1$  ms, 8.6 G/cm) とし、照射後に 200  $\mu\text{s}$  の回復時間を設けた。時間増分は、 $t_{1C} = 1/\text{SW}(\text{C})$ 、 $t_{1N} = 1/\text{SW}(\text{N}) - 1/\text{SW}(\text{C})$  とした。 $t_1$  展開期の位相  $\phi_3$  および  $\phi_4$  におけるクワドラチャ検波には、States-TPPI<sup>36</sup> を適用した。

表 1. タンパク質中のアミノ酸残基種の原子が示す化学シフト値の分布(標準偏差)。

Labeled atom	d (ppm)	
	$^1\text{H}$	$^{13}\text{C}$ or $^{15}\text{N}$
Ala- $^{13}\text{C}\beta$	0.25	1.80
Met- $^{13}\text{C}\epsilon$	0.45	1.43
His- $^{13}\text{C}\epsilon$	0.48	2.16
Trp- $^{13}\text{C}\delta$	0.35	1.83
Tyr- $^{13}\text{C}\epsilon$	0.23	1.21
Arg- $\alpha$ - $^{15}\text{N}$	0.61	3.73
Phe- $\alpha$ - $^{15}\text{N}$	0.72	4.16

$$\omega_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_i \left( \frac{\omega_i}{d_i} \right)^2} \quad (1)$$

ここで、 $\omega$  は相互作用対象の添加前後における化学シフト値の差、 $d$  はスケーリングに用いる化学シフト値の分布の標準偏差(表 1)、 $N$  は核種の数(本研究では  $N=2$ )である。また、 $i$  は各核種に対応し、例えば  $^1\text{H}$ 、 $^{13}\text{C}$ (または  $^{15}\text{N}$ )について  $i=1, 2$  とおき、計算する。

### 2.3.2 バイアスを排除した閾値の設定

化学シフト値の変化量の大小から相互作用に関する議論を行う際、その大小を区別するための閾値としては、例えば化学シフト値の変化量の相加平均<sup>37, 38</sup>等、実験結果を基に経験的に適切と考えられる値が選択されている。一方で、アミノ酸残基の種類により、相互作用界面における存在傾向は異なるため、特定のアミノ酸残基のみを観測プローブとした場合、変化量の相加平均などを閾値とするとバイアスが含まれてしまう。例えば、相互作用界面への存在傾向が高いアミノ酸残基を多数、観測プローブとした場合、その相加平均は大きくなり、その逆においては小さくなる(図 2.5)。しかしながら、特定のアミノ酸残基種に関する実験結果の解析において、この点はこれまで留意されていなかった。本手法でも 2.1 に挙げた特定のアミノ酸残基種のみを観測するため、一般的な相加平均等の値を用いた場合、閾値にバイアスが含まれてしまう。そこで、相互作用界面への存在傾向を考慮した統計情報に基づく重みづけ係数を設定し、変化量の加重平均を閾値として用いることにした。この重み  $w_j$  ( $j$  は Ala、Arg など各アミノ酸残基種に対応する) は、式 2 に従い求められる。

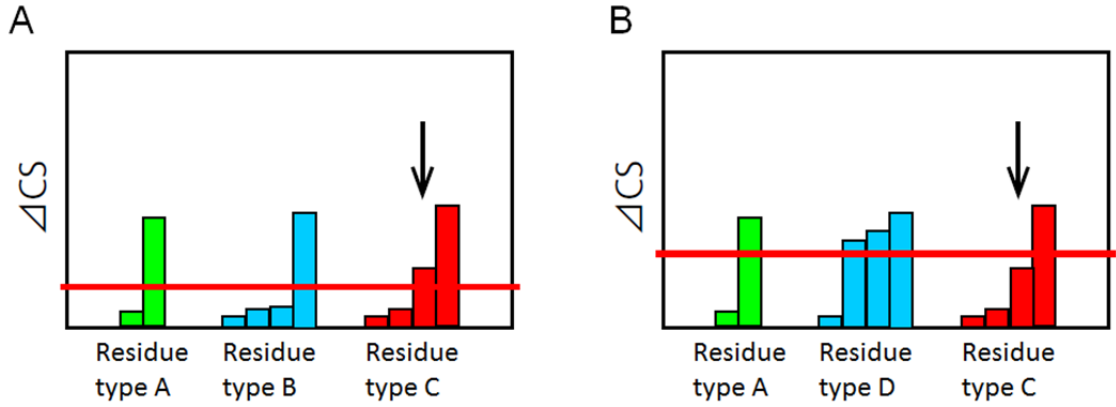


図 2.5. 観測プローブとするアミノ酸残基種が変わることで、閾値が変化することを示す概念図。 $\Delta CS$  は化学シフト値の変化量、赤い横棒は $\Delta CS$  の相加平均の値を意味する。(A) 相互作用界面への存在傾向が低いアミノ酸残基種 B を観測プローブとした場合。(B) アミノ酸残基種 B の代わりに、相互作用界面への存在傾向が高いアミノ酸残基種 D を観測プローブとした場合。観測プローブとするアミノ酸残基種が異なると、黒矢印で示す化学シフト値の変化量の大小判定が変わってしまう。

$$w_j = \frac{\frac{\sum_j n_j}{N_{\text{all}}}}{\frac{n_j}{N_j}} = \frac{\frac{N_j}{N_{\text{all}}}}{\frac{n_j}{\sum_j n_j}} = \frac{P_j}{S_j} \quad (2)$$

ここで、ある仮想タンパク質を仮定し、 $n_j$  はその仮想タンパク質の相互作用界面に存在するアミノ酸残基種  $j$  の数、 $N_{\text{all}}$  は仮想タンパク質の全残基数、 $N_j$  は仮想タンパク質を構成するアミノ酸残基種  $j$  の数を示す。この仮想タンパク質は、構成するアミノ酸残基の種類と数、および PPI 界面に関するそれらについて、全てのタンパク質を考慮した平均値を有する（言い換えれば、全てのタンパク質を代表する）。ただし、真の平均値は不明であるため、下記に示す既知の統計データを基にした標本平均で代用する。式 2 に示すように、 $w_j$  は最終的に  $P_j$  および  $S_j$  で表される。ここで、 $P_j$  は先の仮想タンパク質において、構成アミノ酸残基のうち残基種  $j$  が占める割合、 $S_j$  は相互作用界面に存在する残基のうち、アミノ酸残基種  $j$  が占める割合を意味する。 $P_j$  は UniProtKB (<http://www.uniprot.org>) から、 $S_j$  は文献<sup>12</sup>から値の入手が可能である。これらの値を、 $w_j$  と共に表 2 に示した。この重み  $w_j$  を用いた  $\omega_{\text{RMS}}$  の加重平均値  $\omega_t$  は、下記の式 3 より求められる。

$$\omega_t = \frac{\sum_j w_j a_j}{\sum_j w_j k_j} \quad (3)$$

ここで、 $a_j$ はアミノ酸残基種  $j$  の  $\omega_{\text{RMS}}$  の合計値、 $k_j$ はアミノ酸残基種  $j$  の数を示す。

このようにして求めた重みを用いて加重平均値を求めることで、相互作用界面への存在傾向に起因するバイアスを除いた閾値の設定が可能となった。本研究では、この閾値  $\omega_t$  より大きな  $\omega_{\text{RMS}}$  を有するシグナルを相互作用界面に存在するアミノ酸残基に由来するものとして扱い、アミノ酸残基種ごとにその個数をカウントすることで、相互作用界面に存在すると考えられる個数についての情報を取得した。

表 2. 各アミノ酸残基種の  $P$ 、 $S$ 、および  $w$ 。

Amino acid residue	$P$	$S$	$w$
Ala	0.086	0.039	2.2
Arg	0.055	0.064	0.9
Asn	0.042	0.059	0.7
Asp	0.053	0.066	0.8
Cys	0.013	0.035	0.4
Gln	0.039	0.037	1.1
Glu	0.061	0.065	0.9
Gly	0.071	0.081	0.9
His	0.022	0.034	0.7
Ile	0.060	0.036	1.7
Leu	0.098	0.050	2.0
Lys	0.053	0.057	0.9
Met	0.024	0.020	1.2
Phe	0.040	0.035	1.2
Pro	0.048	0.038	1.3
Ser	0.068	0.079	0.9
Thr	0.056	0.062	0.9
Trp	0.013	0.028	0.5
Tyr	0.031	0.068	0.4
Val	0.067	0.045	1.5
All	0.998	0.998	-

## 2.4 ユビキチンおよびユビキチン加水分解酵素を用いた本手法の有効性の実証

### 2.4.1 原子特異的標識によるアミノ酸残基種の判別

2.1-2.3 で述べた手法が有効に機能することを、YUH と yeast ubiquitin (Y-Ub)<sup>39-41</sup> の相互作用実験を例に確認した。YUH は、236 残基から成る分子重量 26.4 kDa の酵母由来ユビキチン加水分解酵素であり、ユビキチンの C 末端を加水分解し、脱ユビキチン反応を触媒する。Y-Ub は、76 残基からなる分子重量 8.6 kDa の酵母由来ユビキチンである。YUH と Y-Ub との複合体について立体構造解析は行われていないが、YUH とヒト由来 Ub との複合体については既に立体構造が解かれており、その相互作用様式が明らかとなっている<sup>42</sup> (PDB ID: 1CMX)。Y-Ub とヒト由来 Ub は、YUH との相互作用界面から離れた 3 残基のみが異なり (ヒト由来 Ub の Pro19、Glu24、Ala28 が酵母由来ではそれぞれ Ser、Asp、Ser)、両者の YUH との複合体構造は非常に良く似ていることが過去の研究からも示唆されているため<sup>41, 43, 44</sup>、本研究では、相互作用界面が正しく同定されたか確認するための比較対象として、上記の複合体立体構造を用いた。

検証を行うにあたり、まず、2.1 に記載した組み合わせで安定同位体標識を施した YUH を 7.2 に記載の手順にて調製した。次に、2.2 に記載の NMR パルスシーケンスを用いて、標識 YUH の time-shared  $[^1\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}]-\omega_1\text{-TROSY-HSQC}$  /  $[^1\text{H}, ^{13}\text{C}_{\text{methyl}}]\text{-HSQC}$  実験を行ったところ、図 2.6 に示すスペクトルが得られた。スペクトル上で観測された各シグナルは、その化学シフト値からそれぞれ  $[^{13}\text{C}_\epsilon]$  Met、 $[^{13}\text{C}_\beta]$  Ala、 $[^{13}\text{C}_\epsilon]$  His、 $[^{13}\text{C}_\epsilon]$  Tyr、 $[^{13}\text{C}_\delta]$  Trp のいずれに由来するか、容易に判別できた。Met と Ala の区別については、図 2.2 に示す通り  $^{13}\text{C}_\alpha$  のカップリングによるシグナルの分裂からも可能であった。また、 $^1\text{H}-^{13}\text{C}$  相関シグナルと  $^1\text{H}-^{15}\text{N}$  相関シグナルは位相を反転さ

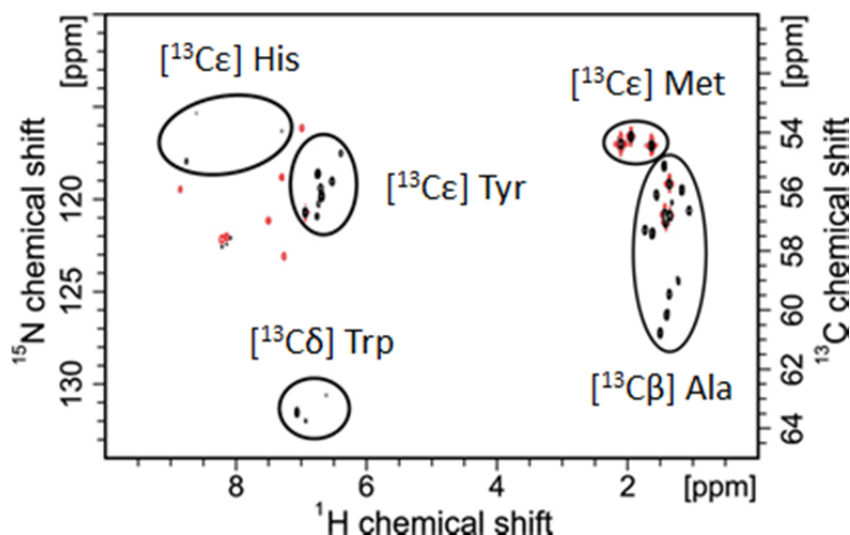


図 2.6. 標識 YUH の time-shared  $[^1\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}]-\omega_1\text{-TROSY-HSQC}$  /  $[^1\text{H}, ^{13}\text{C}_{\text{methyl}}]\text{-HSQC}$  スペクトル。黒と赤は位相が  $180^\circ$  異なり、赤のシグナルは  $[\alpha\text{-}^{15}\text{N}]$  Arg に由来する。スペクトルの折り返しのため、見かけ上、各シグナルの  $^{13}\text{C}$  の化学シフト値は図 2.1B で示す範囲とは異なる値になっている。

せているため、 $[\alpha\text{-}^{15}\text{N}]$  Arg に由来するシグナルも容易に判別が可能であった。各シグナルの個数は、Met: 3、Ala: 14、His: 3、Tyr: 8、Trp: 3、Arg: 7 であり、それぞれ観測が期待される個数が確認された。以上のように、注意深く検討された原子特異的標識の組み合わせを利用することで、同一スペクトル上の各シグナルがどのアミノ酸残基種に由来するか、判別が可能であることが確認された。なお、スペクトルの折り返しのため、各シグナルの  $^{13}\text{C}$  の化学シフト値は見かけ上、図 2.1B で示す範囲と異なる。特に、 $[\text{}^{13}\text{C}_\beta]$  Tyr 由来と  $[\text{}^{13}\text{C}_\beta]$  His 由来とで、シグナルの  $^{13}\text{C}$  の化学シフト値が類似し判別を誤りかねないように見えるが、実際の化学シフト値は大きく異なっており、両者を明確に区別することが可能である。

#### 2.4.2 相互作用に伴う化学シフト変化の観測と PPI 界面の同定

観測対象の各アミノ酸残基種について、PPI 界面に存在すると考えられる個数を調べるため、7.2 に記載の手順で調製した未標識の Y-Ub を前記の標識 YUH へ滴定し、相互作用に伴う化学シフト変化を観測した(図 2.7)。モル比で YUH : Y-Ub = 1:4 となるまで Y-Ub を添加し、その際の化学シフト値の変化量を、2.3 に示す方法により規格化およびスケーリングされた  $\omega_{\text{RMS}}$  として求めた(図 2.7A)。2.3 の記載に従いこれらの加重相加平均値  $\omega_t$  を求め、化学シフト値の変化量の大小を規定するための閾値として使用した。結果、化学シフト変化が大きい、つまり PPI 界面に存在すると考えられるシグナルの個数は、Met: 1、Ala: 5、His: 1、Tyr: 3、Trp: 0、Arg: 0 であった。

次に、Y-Ub の分子量から PPI 界面の大きさを大まかに推定し、YUH の表面(図 2.8)上で先の実験結果に最も一致する箇所を目視で探索した。結果、図 2.8 に示す箇所が PPI 界面である可能性が最も高いと判断された。この箇所は、Ub との複合体結晶構造で確認される PPI 界面であった。以上より、本手法を適用することで、1 サンプルについての実験のみから、シグナル未帰属のまま PPI 界面の同定が可能であることが実証された。



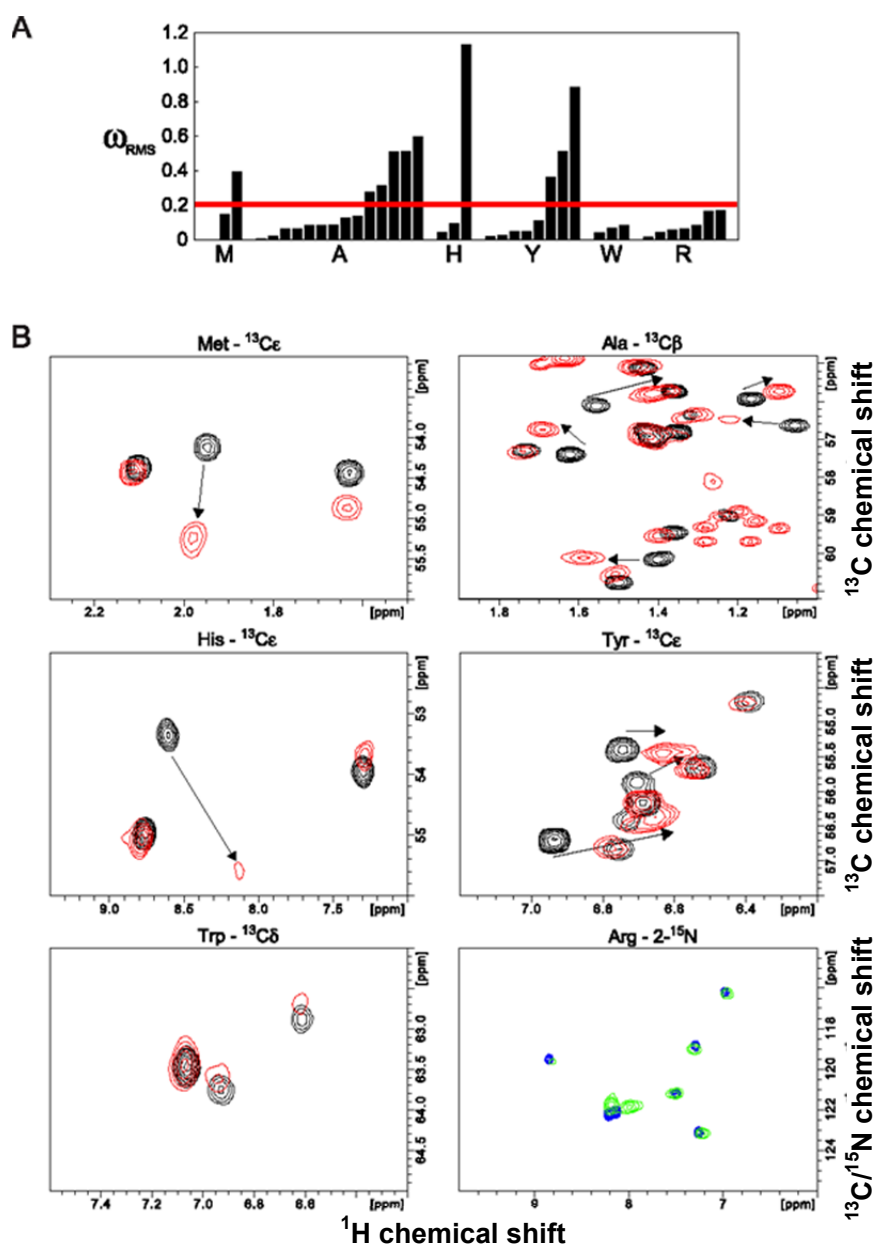


図 2.7. 標識 YUH への Y-Ub 添加実験。(A) 標識 YUH の各シグナルの  $\omega_{RMS}$ 。赤線は、化学シフト変化量の大小を判定するための閾値である  $\omega_t$  を示す。(B) 標識 YUH の time-shared  $[^1H, ^{13}C_{aromatic}/^{15}N]-\omega_1$ -TROSY-HSQC /  $[^1H, ^{13}C_{methyl}]-HSQC$  スペクトル。Y-Ub 滴定前後のスペクトルを重ね、着目する化学シフト範囲を拡大して示す。黒・赤と、青・緑は、位相が  $180^\circ$  異なる。黒および青は滴定前、赤および緑は滴定後のスペクトルであり、矢印は(A)に示す  $\omega_{RMS}$  が  $\omega_t$  より大きいシグナルの摂動先を示す。Y-Ub 添加後にメチル領域に出現したシグナルは、Y-Ub 中に天然に存在する  $^{13}C$  に由来する。

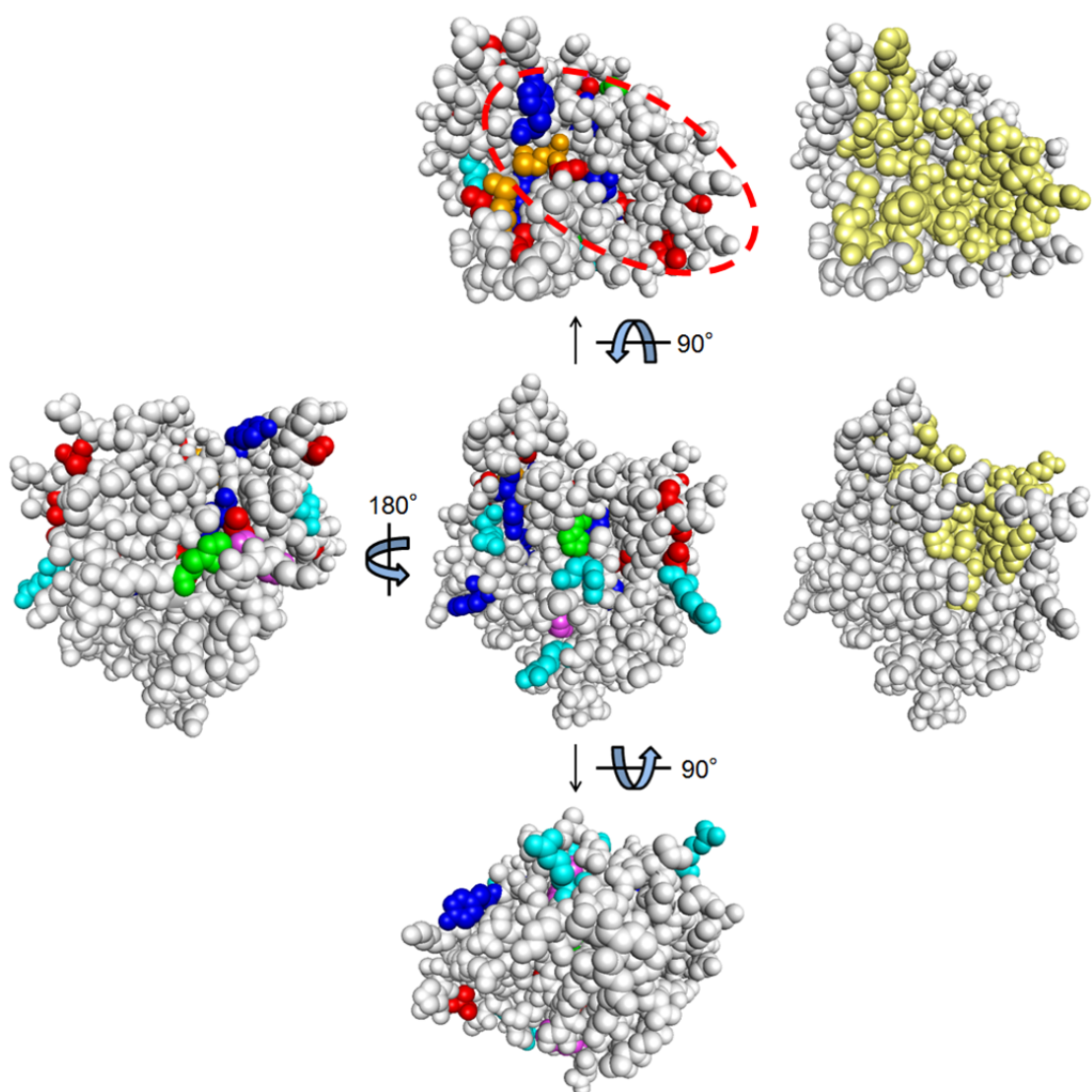


図 2.8. YUH の立体構造表面。全原子を CPK モデルで表示した。各色は、橙: Met、赤: Ala、青: Tyr、緑: His、マゼンタ: Trp、シアン: Arg、黄: Ub との複合体結晶構造で確認される相互作用界面 (Ub から 5 Å 以内に重原子が存在するアミノ酸残基) を示す。目視で実験結果に最も一致する箇所を探索し相互作用界面と同定した箇所を、赤い破線で囲んだ。

## 第 3 章 計算科学的手法の組み合わせによる手法の高度化

### 3.1 相互作用部位同定の自動化と任意性の排除

第 2 章にて、本研究で開発した手法の適用により、PPI 界面を迅速、簡便に同定できることが示された。この手法では、滴定実験の結果を解析する手順が規定され、実験者によらず同じ解析結果を得られるようにしている。一方、その解析結果に最も一致する箇所をタンパク質立体構造の表面から探索する際は、解析者が目視で行う必要がある。したがって、最終的に得られる結果は解析者の主観に依るところが大きく、また簡便であるとは言え、探索には一定の時間と労力を要してしまう。そこで、解析者が行う作業の代わりとなる合理的な手順を規定し、タンパク質の立体構造情報と実験結果を基に PPI 界面を探索するプログラムを開発した。プログラムの処理内容は、下記の通りである。

- (1) プログラムへの入力情報を準備するための前処理として、解析対象のタンパク質立体構造をプログラム CNS<sup>45</sup> で解析し、表面に露出しかつ重原子が 3.5 Å 以内にある近接したアミノ酸残基のペアのリストを作成する。
- (2) 上記のリストを基に、タンパク質表面に沿った近接残基間を結ぶ全経路を探索し、相互作用ネットワークを示すグラフを作成する。この際、各経路の距離は、終点までに経由するアミノ酸残基の個数(終点も数える)として表される。
- (3) 先のリストに挙げたアミノ酸残基の位置を、タンパク質の重心を中心として半径  $p_{\max}/2\pi$  の球表面に投影する。ここで  $p_{\max}$  は、先のグラフにおける最長の経路距離である。
- (4) 先のグラフにおける経路距離と投影後の残基間距離の比を各経路について求め、それらが最も類似の値をとるよう、投影後の各アミノ酸残基の位置を調整する。その後、球の半径を 3 にスケールする。これら一連の作業により、各アミノ酸残基間の相対的な位置関係が単純化される。
- (5) 実験結果から得られた、PPI 界面に存在すると考えられるアミノ酸残基種とその個数に一致するアミノ酸残基の組み合わせ全てを相互作用界面の候補としてリストアップした後、各候補について凸包問題を解く。
- (6) 各候補について、下記の式に従いスコア  $S_c$  を計算する。

$$S_c = S_{\text{area}} + N_{\text{imp}} \quad (4)$$

ここで、 $S_{\text{area}}$  は凸包の面積、 $N_{\text{imp}}$  は凸包内に存在する不適切なアミノ酸残基種の数を表す。この式は、候補のアミノ酸残基から成る面の面積がなるべく小さく、かつ実験結果から PPI 界面に存在しないと考えられるアミノ酸残基種が上記の面内に存在しないほど、PPI 界面である可能性が高いことを意味する。

- (7)  $S_c$  が小さい順に 2000 個の候補を抽出し、アミノ酸残基が 50% 以上一致していることを基準にクラスタリングし、それらをマージする。マージ後のアミノ酸残基の組み合わせについて凸包問題を解き、再度  $S_c$  を求める。この値が最も小さいアミノ酸残基の組み合わせを、PPI 界面

に存在するアミノ酸残基(の一部)とする。

滴定実験の際、相互作用時の周辺の化学的環境によっては、PPI 界面に存在しながらも結果的に化学シフト変化が小さいシグナルが存在しうる等、実験結果から得られる個数の情報は、全ての系において必ずしも絶対的な指標ではない。このような若干の曖昧さが含まれている情報を扱う際、これまでは解析者が最適と考える解を模索し、それゆえ結果に任意性が含まれていた。本研究では、主に(6)のスコアリングと(7)のマージにて、多少の曖昧さを許容しつつも確からしいアミノ酸残基を選び出す手順を規定して任意性を排除し、それらをプログラム化することで解析者に依らず同じ妥当な結果を得られるようにした。

### 3.2 既知のタンパク質複合体構造を用いたシミュレーションによる手法の有効性検証

3.1 で述べたプログラムにより PPI 界面を正しく同定できるか確認するため、既知のタンパク質複合体構造を用いたベンチマークシミュレーションを行った。複合体構造は、統計解析の結果から区分される 3 種類の複合体形成様式<sup>13</sup>について 1 つずつ (PDB ID: 1E3S<sup>46</sup>、1AOH<sup>47</sup>、1K4D<sup>48</sup>) と、大きな PPI 界面を有する複合体 1 つ (PDB ID: 1E6V<sup>49</sup>) の計 4 個を用いた。相互作用対象のタンパク質から 5 Å 以内に重原子が存在するアミノ酸残基を PPI 界面の残基とし、Ala、Arg、His、Met、Tyr、Trp についてその個数を数え、タンパク質単体の立体構造と共に 3.1 のプログラムに入力した。結果を図 3.1 および表 3 に示す。PPI 界面に存在するアミノ酸残基の 73%以上、平均で 88%が同定され、これら 4 つの構造でシミュレートした限りでは、相互作用界面を同定する上で十分な精度の結果が得られることが確認された。

### 3.3 プログラムを用いた YUH および Y-Ub の相互作用界面の同定

3.2 より、プログラムを用いることで十分な精度で PPI 界面の同定が可能であることが確認されたため、次に YUH および Y-Ub について、滴定実験の結果を基に PPI 界面の同定を行った。

2.4 で得られた YUH に関する実験結果と、複合体構造のデータより準備した YUH の立体構造座標をプログラムへ入力し計算した結果、Tyr33、Tyr37、Ala156、Ala159、Ala161、His166、Tyr167、Tyr214、Ala218、Ala227、Met228 が相互作用界面の残基として判断された。これらの残基を複合体構造へマッピングした結果、Y-Ub との相互作用界面が同定されていることが確認された(図 3.2A)。

次に、Y-Ub についても PPI 界面の同定を行った。Y-Ub について、7.2 に記載の手順で標識 Y-Ub と未標識 YUH を調製し、モル比で Y-Ub : YUH = 1: 4 となるまで YUH を添加した後、YUH の場合と同様に化学シフト変化が大きいシグナルの個数を各アミノ酸残基種について求めた。結果、PPI 界面に存在すると考えられるアミノ酸残基の数は、Met: 0、Ala: 0、His: 1、Tyr: 0、Trp: 0、Arg: 3 であった(図 3.3)。複合体構造のデータより Ub の立体構造座標を準備し、先のデータと共にプログラムへ入力した結果、Arg42、His68、Arg72、Arg74 が PPI 界面の残基として判断され、YUH との

相互作用界面が同定されていることが確認された(図 3.2B)。

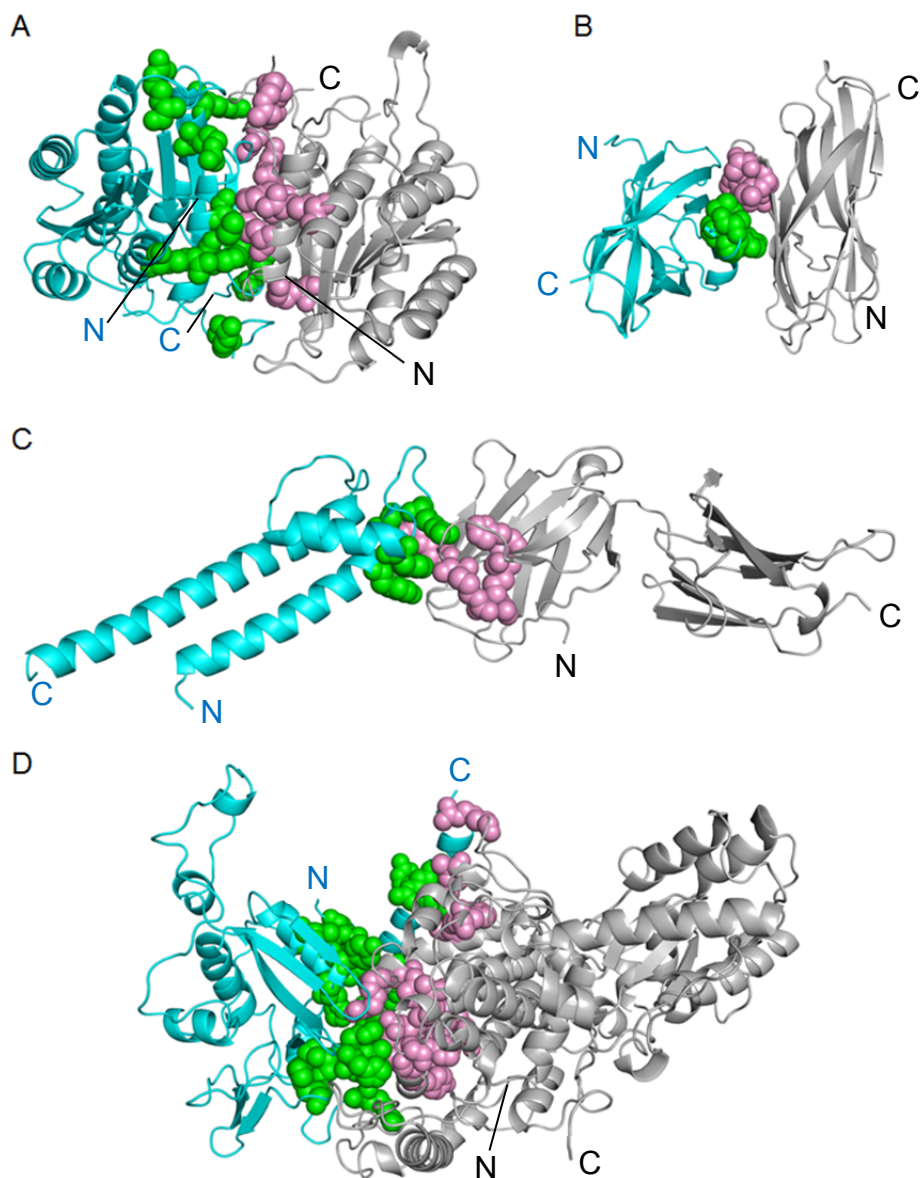


図 3.1. 既報の複合体構造を用いたシミュレーションの結果。プログラムにより同定された PPI 界面の Ala、Arg、His、Met、Tyr、Trp を、CPK モデル(緑またはピンク)で示す。(A) 3-hydroxyacyl-coa dehydrogenase(241 残基、homodimer) (PDB ID: 1E3S)。 (B) type I cohesin domain(143、147 残基、homodimer) (PDB ID: 1AOH)。 (C) Potassium channel KcsA のフラグメント(103 残基、シアン)と抗体の Fab 重鎖(219 残基、灰色) (PDB ID: 1K4D)。 (D) Methyl-coenzyme m reductase  $\gamma$  サブユニット(248 残基、シアン)と  $\alpha$  サブユニット(545 残基、灰色) (PDB ID: 1E6V)。

表 3. 既報の複合体構造を用いたシミュレーション結果の詳細。

PDB ID	Interface residues of crystal structure <sup>†</sup>	Identified interface residues <sup>‡</sup>	Rate of correctness of identified residues <sup>§</sup>
1E3S A (241 residues)	A183		0.86
	R184		
	A187	R29	
	R192	R192	
	R226	R226	
	H235	H235	
	M239	M239	
	R252	R252	
	A256	R258	
	R258		
1E3S C (241 residues)	A183	R29	0.78
	R184		
	A187	R147	
	R192	R184	
	R226	R192	
	H235	R226	
	M239	H235	
	R252	M239	
	A256	R252	
	R258	R258	
1AOH A (143 residues)	A72	A72	1.00
	Y74	Y74	
	R77	R77	
	M79	M79	
1AOH B (147 residues)	A72	A72	1.00
	Y74	Y74	
	R77	R77	
	M79	M79	

1K4D A (219 residues)	W33		
	H35	H35	
	Y55	R57	0.75
	R57	R98	
	R100	R100	
1K4D C (103 residues)	Y45	Y45	
	A50	A50	1.00
	R52	R52	
	Y62	Y62	
	Y25		
	Y29		
	R70	Y383	
	M370	Y388	
	Y383	R390	
	Y388	Y391	
	R390	M395	
	A393	H398	
	M395	R404	
	H398	A405	0.73
	R404	Y431	
1E6V A (545 residues)	Y431	H437	
	H437	R443	
	R443	A457	
	Y447	A458	
	A450	R464	
	A457	Y481	
	A458		

	Y9		
	R19	Y9	
	Y96	R19	
	H101	R20	
	Y103	Y96	
	R107	H101	
	R151	Y103	
	H160	R107	
	H162	R151	
	A163	H160	
	R165	H162	
1E6V C	M172	A163	0.90
(248 residues)	H174	R165	
	A175	M172	
	R177	H174	
	Y221	Y221	
	R222	R222	
	Y228	Y228	
	R229	R229	
	H242	R240	
	R245	H242	
	Y250	R245	
	R251		

---

<sup>†</sup>PPI 界面の残基 (相互作用対象のタンパク質から 5 Å 以内に重原子が存在するアミノ酸残基)。

<sup>‡</sup>プログラムにより PPI 界面に存在すると判断されたアミノ酸残基。

<sup>§</sup>PPI 界面に存在するアミノ酸残基のうち、プログラムにより同定された割合。



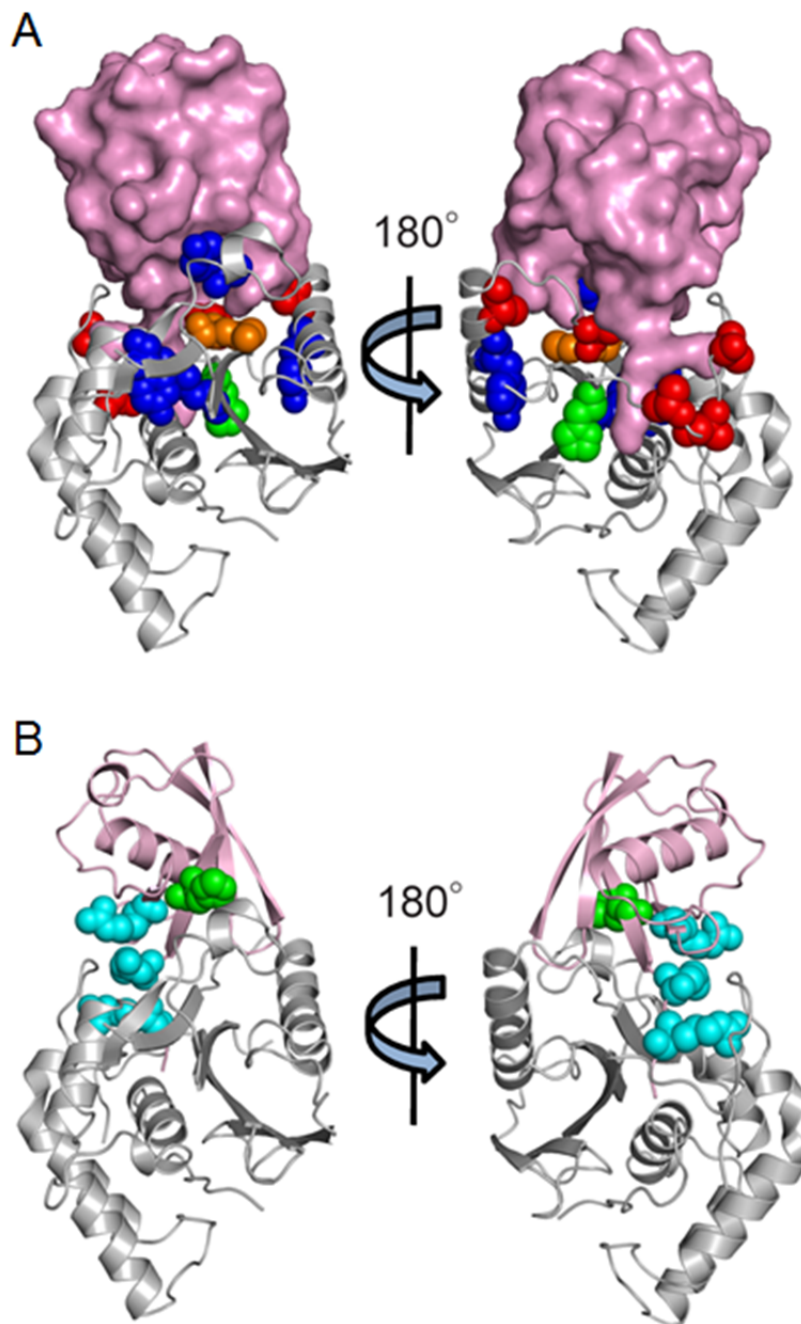


図 3.2. プログラムを用いて実験結果と立体構造座標から同定した、PPI 界面のアミノ酸残基。  
各残基を CPK モデル(赤:Ala、橙:Met、緑:His、青:Tyr、シアン:Arg)で示す。(A) YUH  
上の PPI 界面残基。YUH を灰色のリボン表示で、Ub をピンクの surface 表示で表した。  
(B) Ub 上の PPI 界面残基。YUH、Ub をそれぞれ灰色、ピンクのリボン表示で示した。

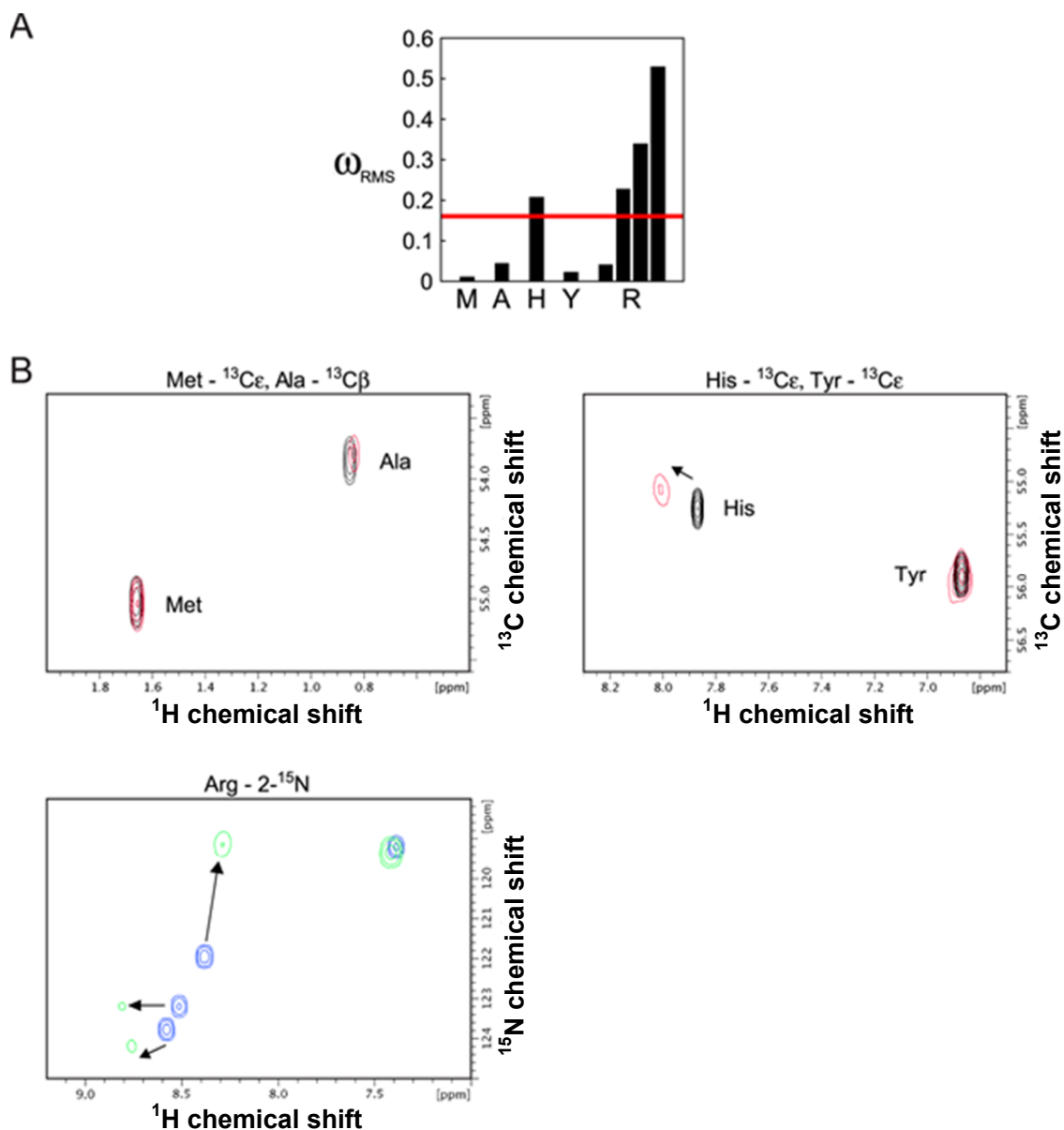


図 3.3. 標識 Y-Ub への YUH 添加実験。(A) 標識 Y-Ub の各シグナルの  $\omega_{\text{RMS}}$ 。赤線は、化学シフト変化量の大小を判定するための閾値である  $\omega_t$  を示す。(B) 標識 Y-Ub の HSQC スペクトル。YUH 滴定前後のスペクトルを重ね、着目する化学シフト範囲を拡大して示す。黒・赤と青・緑は、位相が  $180^\circ$  異なる。黒および青は滴定前、赤および緑は滴定後のスペクトルであり、矢印は(A)に示す  $\omega_{\text{RMS}}$  が  $\omega_t$  より大きいシグナルの摂動先を示す。観測されるシグナルの数が少なく、高分解能スペクトルを取得する必要がないため、既報の time-shared HSQC 測定用のパルスシーケンス<sup>26</sup>を適用した。

### 3.4 手法の頑健性の検証

滴定実験の結果の解析からプログラムによる PPI 界面の同定まで、本手法が正しく機能することが、これまでの結果より示された。そこで次に、外乱による影響が PPI 界面の同定結果にどの程度影響するか確認し、本手法の頑健性を評価した。手法を構成する各要素のうち外乱の影響を受ける箇所は、化学シフト値の変化量を取得するための滴定実験であり、PPI 界面に存在すると考えられるアミノ酸残基の個数を取得する重要な過程である。そこで例として、YUH の実験結果について、 $\omega_t$  の値を 20%低下させた場合にどのような影響が生じるか確認した。

上記の通り  $\omega_t$  の値を変えることで、PPI 界面に存在すると考えられるアミノ酸残基の個数が Met: 1、Arg: 2 追加された。このデータを用いて 3.3 と同様にプログラムで計算した結果、Arg6、Arg65、および Met215 が新たに追加された、Arg6、Tyr37、Arg65、Ala156、Ala159、Ala161、His166、Tyr167、Tyr214、Met215、Ala218、Ala227、Met228 が YUH の PPI 界面にある残基として判断された(図 3.4)。この結果は、正しい入力データから得られる結果と非常に類似しており、Met215 は 3.3 で同定された相互作用界面のすぐ隣に位置し、また Arg6 と Arg65 が実際の相互作用界面から離れたところに存在しているものの、それ以外のアミノ酸残基からなるパッチの外側に孤立して存在していることから、これらは無視してよいものと理解できる。結果、外乱がない場合の結果とほぼ同一の箇所が、PPI 界面として同定された。

以上のように、実験のエラー等、結果に影響を及ぼす何らかの要因が生じていた際にも、プログラムの出力結果を注意深く評価することで PPI 界面を同定できることが確認され、本手法の頑健性を支持する結果となった。

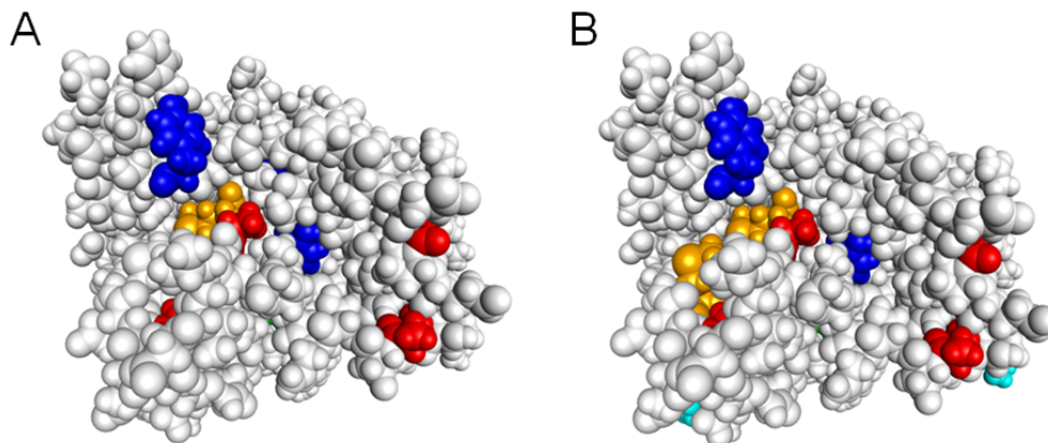


図 3.4. YUH の相互作用界面として同定される残基の比較。各原子を CPK モデルで表示した。各色は、プログラムにより相互作用界面として同定されたアミノ酸残基であり、橙: Met、赤: Ala、青: Tyr、緑: His、マゼンタ: Trp、シアン: Arg を示す。(A) 3.3 で同定した相互作用界面に存在するアミノ酸残基。(B)  $\omega_t$  の値を 20%低下させた際に相互作用界面に存在すると判断されたアミノ酸残基。下部に確認される 2 つの Arg は Arg6 と Arg65 である。

### 3.5 実験結果を反映させた確からしいタンパク質複合体モデル構造の構築

PPI 界面は相互作用に関する知見として非常に重要であるが、構造生物学的あるいは創薬の観点から、複合体としてのより詳細な知見の取得が望ましいことは言うまでもない。この場合、ドッキングシミュレーションなどの計算化学的手法による複合体構造の予測が有用であり、特にシミュレーションを行う際に実験により得られた構造情報を反映させることで、高精度の複合体モデル構造の構築が期待できる。しかしながら、それらの実験的な情報の取得には時間と労力を要することが多かった。一方、今回開発した相互作用部位解析法は、PPI 界面に存在するアミノ酸残基の一部を迅速かつ簡便に同定でき、これらの情報は複合体構造の予測において非常に有用であると考えられる。そこで、本手法で得られた情報を利用したドッキングシミュレーションを行い、得られた複合体モデル構造の精度を評価することで、複合体構造の予測におけるそれらの情報の有用性を確認した。

これまでに述べたように、本手法で得られる相互作用界面の情報には多少の曖昧さが含まれている。そこで本研究では、このような情報を利用したドッキングシミュレーションに適したプログラムである HADDOCK<sup>50, 51</sup> を使用した。シミュレーションに使用できる情報は相互作用界面の一部の残基に関してのみであるため、7.4 に記載したように、これらの情報から適切な距離拘束条件を設定する手順を規定した。

例として、YUH と Ub に関する 3.3 の結果を用いて、7.4 に記載の手順でドッキングシミュレーションを行った。Ub の C 末端は YUH-Ub の相互作用系において重要な領域であり、非常にフレキシブルでありながら YUH の活性部位である狭く深いポケットに結合する。そのため、複合体モデル構造の構築において再現が重要な箇所でありながら、一般にドッキングシミュレーションでの再現が難しい。しかしながら、シミュレーションの結果得られた最も確からしいとされた複合体構造において、結晶構造に非常に類似の結合様式が再現されていた(図 3.5)。次に、得られた YUH と Ub の複合体モデル構造の精度を、世界的なタンパク質複合体構造予測コミュニティとして知られる CAPRI<sup>52-54</sup> の基準を適用して、複合体結晶構造と比較し評価した。結果、“fraction of native contacts”(分子間で 5 Å 以内に存在する残基ペアの再現率)、“interface RMSD”(相互作用対象より 10 Å 以内に存在する残基の主鎖に関する、結晶構造との RMSD)、“ligand RMSD”(YUH について主鎖重原子でフィッティングした際の Ub 主鎖重原子に関する、結晶構造との RMSD)は、それぞれ 0.70、1.84 Å、3.12 Å であり、これらの結果の総評としては、2 番目に良いカテゴリに分類される精度の複合体モデル構造であった。また、最も確からしいとされた予測構造を含むクラスタの各構造に関する pairwise RMSD は表 4 に示す通りであり、これらの mean global backbone RMSD は 2.02 Å であることから、各構造に大きな差異はなく、確からしい構造を複数、安定して得られていることが確認された。

以上のように、複合体を構成する各タンパク質の構造が既知であれば、本手法により同定された PPI 界面の残基に関する情報を基に、実験結果を反映させた確からしい複合体モデル構造の構築が可能であることが示された。

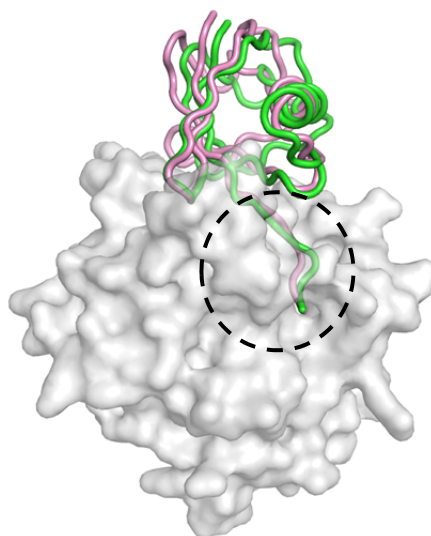


図 3.5. ドッキングシミュレーションにより得られた複合体モデル構造と複合体結晶構造の重ね合わせ。YUH の主鎖重原子についてフィッティングを行った。結晶構造の YUH を surface 表示(モデル構造の YUH は非表示)、Ub を tube 表示(ピンク:結晶構造、緑:モデル構造)で表し、YUH の活性部位を破線で囲んだ。

表 4. ドッキングシミュレーションで得られた複合体モデル構造 4 個と結晶構造の pairwise RMSD (Å)。

	Crystal structure	Model 1	Model 2	Model 3	Model 4
Crystal structure	–	1.65	1.86	1.30	3.08
Model 1	2.17	–	1.52	1.22	2.25
Model 2	2.22	1.99	–	1.46	2.04
Model 3	1.85	1.65	1.80	–	2.36
Model 4	3.42	2.57	2.31	2.66	–

右上は主鎖原子の RMSD を、左下は全重原子の RMSD を示す。これらの値は、MOLMOL 2K.2<sup>55</sup> を使用して求めた。

## 第 4 章 タンパク質—低分子間相互作用部位解析への手法の拡張

### 4.1 タンパク質—タンパク質間相互作用部位解析法の応用と課題

これまで、PPI 界面の情報取得を目的とした手法を開発してきたが、一方で、生体内の天然リガンドや薬様化合物などの、いわゆる低分子化合物とタンパク質との相互作用に関する情報もまた、種々の生化学的現象の理解や創薬において極めて重要である。例えば、創薬研究の初期においては、スクリーニングにより得られた活性化合物群に関する特異的結合の確認と偽陽性の排除、さらには構造情報に指南された効率的な化合物設計などを行うにあたり、相互作用の部位やその様式に関する情報が必須となる。特に近年盛んな、フラグメントと呼ばれる分子量が比較的小さな化合物から出発し効率的に合成展開を進めていく Fragment-based drug discovery<sup>56-63</sup> のアプローチでは、その思想から、特に相互作用部位に関する情報の重要性が顕著である。

相互作用部位に関する情報を取得する際には、1.1 で述べた通り幾つかの手法が利用できるものの、いずれも時間と労力を要するため、例えば創薬研究においては開発のタイムスケジュールとの乖離が問題となることが多かった。そこで、第 1 章から第 3 章までに述べた PPI 界面の迅速解析法のアプローチを応用した、タンパク質—低分子間相互作用部位の迅速解析法を開発した。

タンパク質—低分子間相互作用への応用において、根本的な概念は PPI 界面の解析法と共通しており、相互作用部位に存在するアミノ酸残基の種類とその個数の取得には第 2 章で示すアプローチが適用できると考えられる。一方で、その後のデータ解析の自動化においては、低分子化合物とタンパク質との相互作用様式の違いにより、同じ手法をそのまま適用することは困難である。タンパク質間相互作用部位は連続した面を構成するため、3.1 で述べた通り、得られた実験結果を満たすアミノ酸残基の組み合わせで構成される面の面積がなるべく小さくなることを指標の 1 つとして、適切な組み合わせを選択した。しかしながら、低分子化合物はタンパク質表面のポケットの途中にあるアミノ酸残基と結合することが多々あり、相互作用に関与する残基は必ずしも連続する面を構成するわけではなく、タンパク質表面に沿った距離において近い位置にない。したがって、3.1 に示すような”面”を考慮したアプローチは、低分子相互作用部位の同定に適用できない。

### 4.2 三次元空間情報を利用したタンパク質—低分子間相互作用部位解析法の開発

4.1 の課題を解決するため、実験結果に最も一致する”面”ではなく、”空間”を探索する手法を構築することにした。3 次元空間を計算科学的手法により探索するためには、空間を量子化する必要がある。そこで、空間を格子状に区切り、各格子点(グリッド)について、周囲に存在するアミノ酸残基の種類と数を実験結果と比較し一致度を評価することにした。このアプローチを適用した手法(図 4.1)を、以下に示す。

- (A) 2.1 と同様の方針で選択したアミノ酸残基種の特定の原子を安定同位体で標識した、解析対象のタンパク質を調製する。この際、PPI 界面への存在傾向の代わりに、低分子との相互作用部位への存在傾向<sup>64</sup>を考慮する(図 4.2)。本研究では、標識の組み合わせとして

[<sup>13</sup>Cε] Met、[<sup>13</sup>Cα/<sup>13</sup>Cβ] Ala、[<sup>13</sup>Cε] His、[<sup>13</sup>Cε] Tyr、[<sup>13</sup>Cδ] Trp、[α-<sup>15</sup>N] Phe を選択した。なお、標識タンパク質の調製には、2.1 と同様にセルフリー合成系を用いる。

- (B) 前記の標識サンプルを用いて、低分子化合物の滴定実験を行う。観測される各シグナルの化学シフト値の変化量は、2.3.1 と同様にスケールリングされた  $\omega_{\text{RMS}}$  で表される。各シグナルの  $\omega_{\text{RMS}}$  を用いて、化学シフト変化の大きさを規定するための閾値  $\omega_{\text{tp}}$  を下記の式に従い求める。

$$\omega_{\text{tp}} = m + s \quad (5)$$

ここで、 $m$  は全  $\omega_{\text{RMS}}$  の相加平均、 $s$  は全  $\omega_{\text{RMS}}$  から求めた標準偏差である。これらの値を用いて、化学シフト変化が大きいアミノ酸残基種とその個数に関する情報を取得する。

- (C) 解析対象のタンパク質の立体構造座標を準備する。  
 (D) プログラム EasyMIFs および SiteHound<sup>65-67</sup> を使用して、タンパク質表面の空間を格子状に区切り、エネルギー的観点から疎水性官能基と相互作用する可能性が高いグリッドを抽出する。  
 (E) (D) で抽出した各グリッドについて、実験結果(B)との不一致度を評価する。まず、下記に示す仮想球の半径  $r_v$  を求める。

$$r_v = r_{\text{seb}} + r_{\text{csp}} \quad (6)$$

ここで、 $r_{\text{seb}}$  は化合物の最小包含球の半径、 $r_{\text{csp}}$  は化学シフト変化の誘起が期待される化合物からの距離を意味する。 $r_{\text{seb}}$  は、PubChem3D (<http://pubchem.ncbi.nlm.nih.gov>) より入手した化合物の構造を用いて、既報のアルゴリズム<sup>68</sup>により求めた。なお、 $r_{\text{seb}}$  を求めるためのスクリプトを付録に記載した。また、 $r_{\text{csp}}$  は、本研究では 4 Å とした。このようにして求めた仮想球の中心を各グリッドにおき、(A)に記載の各アミノ酸残基種について、仮想球内に存在する個数と(B)における実験結果との不一致度を示すペナルティスコア  $S_{\text{tot}}$  を下記の式で求めた。

$$S_{\text{tot}} = \sum_a |V_a - R_a| \quad (7)$$

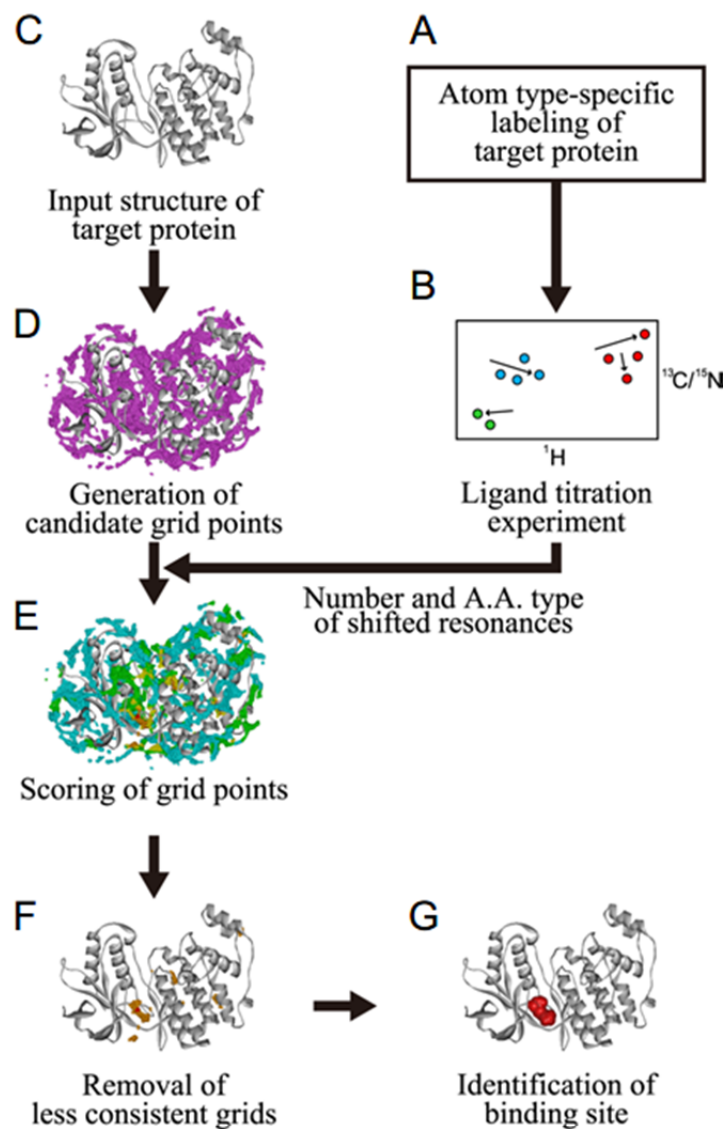
ここで、 $a$  は標識された各アミノ酸残基 ( $a$  = Ala, Met, ……),  $V_a$  は仮想球内に存在するアミノ酸残基  $a$  の個数、 $R_a$  は(B)において  $\omega_{\text{tp}}$  より  $\omega_{\text{RMS}}$  が大きかったアミノ酸残基  $a$  の個数であり、 $\Sigma$  は各アミノ酸残基に関して差の絶対値の和を取ることを意味する。 $S_{\text{tot}}$  の値が小さいほど、実験結果と一致度が高い、すなわち相互作用部位に存在している可能性が高いグリッドであることを意味する。

- (F) 化合物の体積以上のグリッド塊が出現するまで、 $S_{\text{tot}}$  の値が小さいグリッドから順に表示していく。なお、グリッドに炭素原子を置き、その van der Waals 半径内の空間をグリッド塊の体積とする。



(G) 最終的に、 $S_{\text{tot}}$  の値が小さいグリッドが化合物の体積以上の塊を形成している箇所が、相互作用部位として同定される。

上記手順に従い、実験結果に一致する空間を探索することで、低分子相互作用部位の同定が可能となった。



Adapted with permission from J. Med. Chem., 2013, 56 (22), pp 9342–9350. Copyright 2013 American Chemical Society.

図 4.1. 低分子相互作用部位解析法のストラテジー。標識タンパク質を調製し(A)、化合物の滴定実験を行うことで、化学シフト変化が大きいシグナルの個数をアミノ酸残基種ごとに調べる(B)。既知のタンパク質立体構造(C)の表面から、エネルギー的に疎水性官能基が相互作用しやすい箇所を探索しグリッドで表現する(D)。各グリッドについて、実験結果との不一致度を評価し(E)、一致度が高く(F)化合物の体積以上のグリッド塊が存在する箇所が相互作用部位と考えられる(G)。



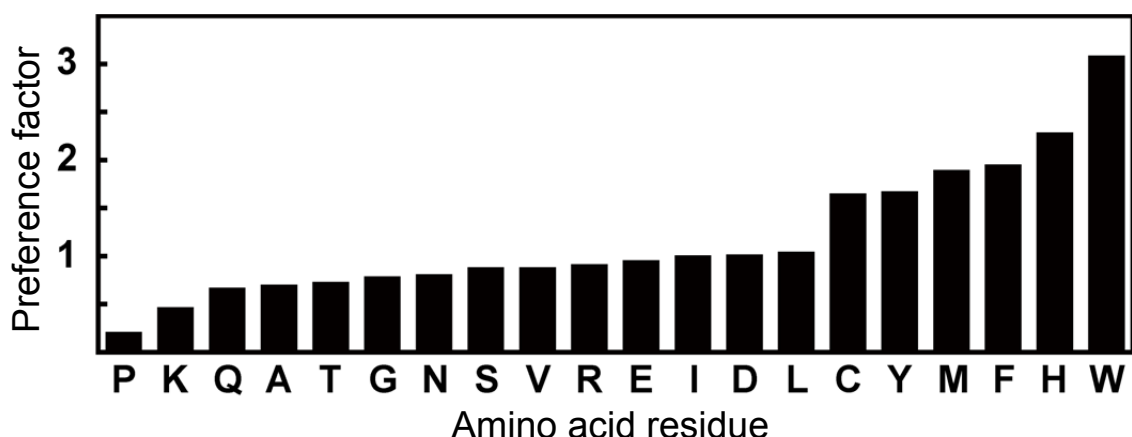


図 4.2. 各アミノ酸残基種の、低分子相互作用部位への存在傾向。アミノ酸残基種  $a$  の Preference factor  $P_a$  は、低分子相互作用部位に占めるアミノ酸残基種  $a$  の割合  $C_a$  と、タンパク質表面に占める割合  $J_a$  を用いて、 $P_a = C_a/J_a$  で求められる。全てのデータは、4.2 に記載の既報のものを用いた。

#### 4.3 手法の頑健性向上

4.2 に記載の手法では、標識されたアミノ酸残基種について全てのシグナルが観測されることを前提としている。しかしながら、化学交換によるシグナルの広幅化やシグナルの縮重等により、全てのシグナルが観測されない場合も多い。そこで、そのような場合においても適切な解析が可能なるよう、 $S_{\text{tot}}$  を求める際には全ての可能性を考慮するようにし、観測されない各シグナルについて、それぞれ化学シフト変化が大きい場合と小さい場合の両方を考慮することにした。例えば、化学シフト変化が大きいシグナルの個数が His と Phe でそれぞれ 1、3 であり、さらに His と Phe のいずれにおいても観測されないシグナルが 1 個ずつある場合、化学シフト変化が大きいシグナルの個数として (His, Phe) = (1, 3), (2, 3), (1, 4), (2, 4) の 4 通りを考え、各場合における  $S_{\text{tot}}$  を求めその平均値を算出し、この値を各グリッドの  $S_{\text{tot}}$  とした。このように処理することで、観測されないシグナルがある場合においても、それらによるバイアスを除いた解析が可能となった。この処理を加えた 4.2 の解析が可能なスクリプトを、付録に記載した。

#### 4.4 既知の複合体構造を用いたシミュレーションによる手法の有効性検証

4.2、4.3 で示した手法の有効性を確認するため、既知のタンパク質—低分子複合体構造を用いたシミュレーションを行った。まず、低分子の周囲 5 Å 以内に重原子が存在するアミノ酸残基を相互作用に関与する残基とし、4.2 に記載の標識したアミノ酸残基種 (Ala、His、Met、Phe、Tyr、Trp) についてその個数を数え、実験結果の代わりとした。次に、*apo* 体について構造決定された座標データと Pubchem 3D より入手した化合物構造を用いて 4.2、4.3 の手法により解析し、

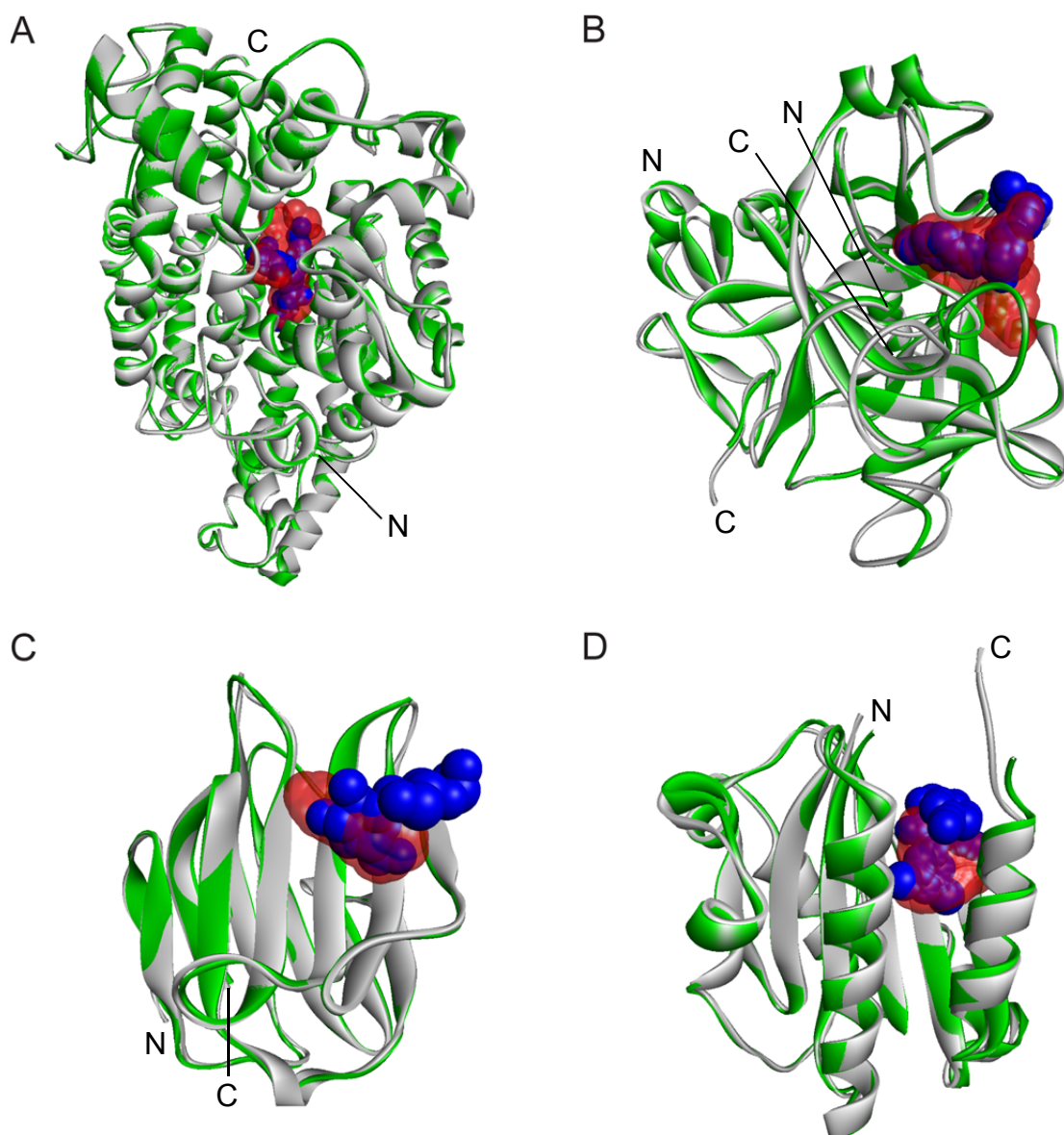


図 4.3. 既報の構造を用いたシミュレーションの結果。*apo* 体の構造を灰色のリボンモデル、*holo* 体の構造を緑のリボンモデル、低分子を青の CPK モデル、*apo* 体の構造を用いて同定された相互作用部位を赤で表した。(A) アンジオテンシン I 変換酵素の N ドメインと阻害剤 lisinopril (PDB ID: 2C6F (*apo*), 2C6N (*holo*))。 (B) Factor Xa と阻害剤 (PDB ID: 1C5M (*apo*), 2J2U (*holo*))。 (C) galectin-9 の N ドメインとラクトース (PDB ID: 3LSD (*apo*), 3LSE (*holo*))。 (D) LFA-1 (CD11a/CD18) の N ドメインと LFA-1/ICAM-1 相互作用の阻害剤と予想されている化合物 (PDB ID: 1ZOP (*apo*), 2O7N (*holo*))。

結果を複合体構造と比較した。シミュレーションには種類が異なる 4 つのタンパク質として、酵素とその阻害剤(PDB ID: 1C5M, 2J2U)<sup>69, 70</sup>、大きな酵素とその阻害剤(PDB ID: 2C6F, 2C6N)<sup>71</sup>、糖結合タンパク質と糖(PDB ID: 3LSD, 3LSE)<sup>72</sup>、PPI を機能として有するタンパク質と PPI 阻害剤と予想されている化合物(PDB ID: 1ZOP, 2O7N)<sup>73, 74</sup>を用いた。結果を図 4.3 に示す。これら 4 つのタンパク質について、いずれも本手法により相互作用部位を正しく同定することができた。糖結合タンパク質についても相互作用部位を正しく同定できたことから、本手法が薬様化合物以外についても適用できることが確認された。また、この複合体の相互作用部位はいわゆるポケット様の形状をしておらず、タンパク質の形状から目視あるいは計算科学的手法により相互作用部位を予測することは困難であるが、本手法で必要とする僅かな情報さえあれば同定が可能であることが示された。

#### 4.5 p38 $\alpha$ mitogen-activated protein kinase と阻害剤を用いた本手法の有効性の実証

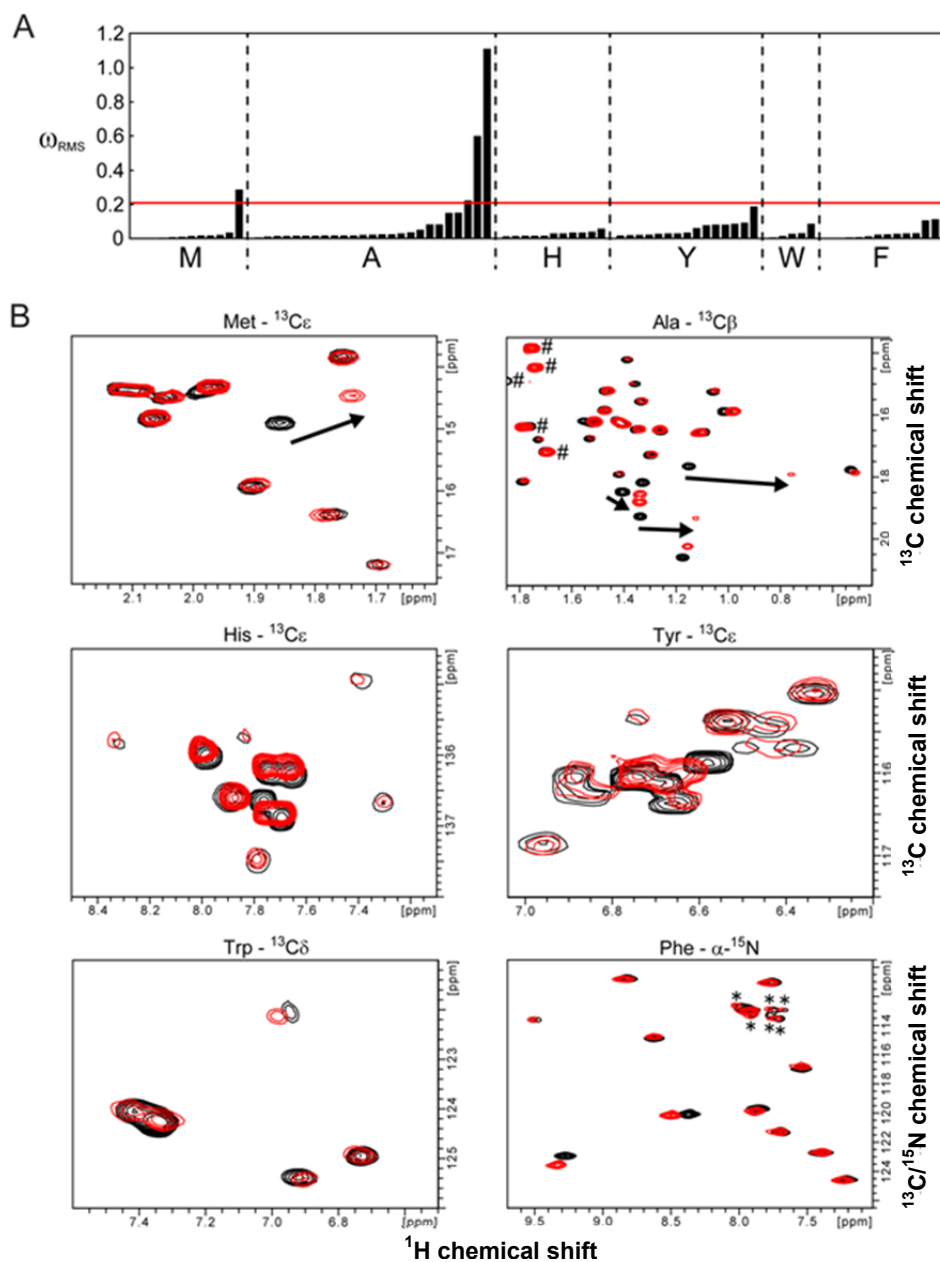
4.4 にて、本手法が正しく機能することをシミュレーションにより確認したため、次に、実験結果から相互作用部位が正しく同定されることを実証した。実験に使用するタンパク質としては、mitogen-activated protein kinase (MAPK)とよばれる広く真核生物に保存された serine/threonine kinase の 1 つである、p38 $\alpha$ (別名: MAPK14、分子重量: 41 kDa)を用いた。MAPK は、MAPK kinase によりリン酸化され活性化することで核内へと移行し、細胞外から核内へのシグナル伝達において重要な役割を果たす kinase である。現在、4 種類の MAPK カスケードが知られており、このうち MAPK として p38 $\alpha$ / $\beta$ / $\gamma$ / $\delta$  が関与する p38MAPK カスケードは、多くの細胞種の生存や増殖の制御に加え、サイトカインやストレスへの細胞応答をコントロールしていると考えられている<sup>75-78</sup>。中でも p38 $\alpha$  は、4 種のうち最も広範に発現しており、創薬のターゲットタンパク質として広く認知され性質がよく調べられている<sup>79, 80</sup>。p38 $\alpha$  と相互作用する低分子としては、p38 $\alpha$  との複合体構造が既知であるフラグメントと呼ばれるサイズの阻害剤 2-amino-3-benzoyloxypyridine<sup>81, 82</sup>(分子量: 200)を用いた。

最初に、4.2 に記載の安定同位体標識を施した p38 $\alpha$  を、セルフリー合成系を用いて 7.5 に記載の手順で調製した。2.2 に記載の NMR パルスシーケンスを用いた time-shared [<sup>1</sup>H, <sup>13</sup>C<sub>aromatic</sub>/<sup>15</sup>N]- $\omega_1$ -TROSY-HSQC / [<sup>1</sup>H, <sup>13</sup>C<sub>methyl</sub>]-HSQC 実験を行ったところ、観測されたシグナルの個数は Met: 11/11、Ala: 25/25、His: 11/12、Tyr: 15/15、Trp: 5/5、Phe: 12/13 であり、His および Phe は期待される数よりも 1 つ少なかった。このサンプルへ阻害剤をモル比で 5.5 倍量になるまで添加し、相互作用に伴う化学シフト変化を観測した(図 4.4)。その際の化学シフト値の変化量をスケールリングされた  $\omega_{\text{RMS}}$  として求め、 $\omega_{\text{tp}}$  より大きいシグナルの個数を求めたところ、Met: 1、Ala: 3、His: 0、Tyr: 0、Trp: 0、Phe: 0 となり観測された全シグナル数の 5%が抽出された(図 4.4A)。ここで、His および Phe については、観測されたシグナルの数が期待される数よりも 1 つ少なかったことから、4.3 に記載の通り全ての可能性を考慮し、結果、相互作用部位周辺に存在する個数は Met: 1、Ala: 3、His: 0 or 1、Tyr: 0、Trp: 0、Phe: 0 or 1 となった。

次に、Pubchem 3D から入手した阻害剤の構造を用いて  $r_{\text{seb}}$  を求めたところ、4.497 Å であった。

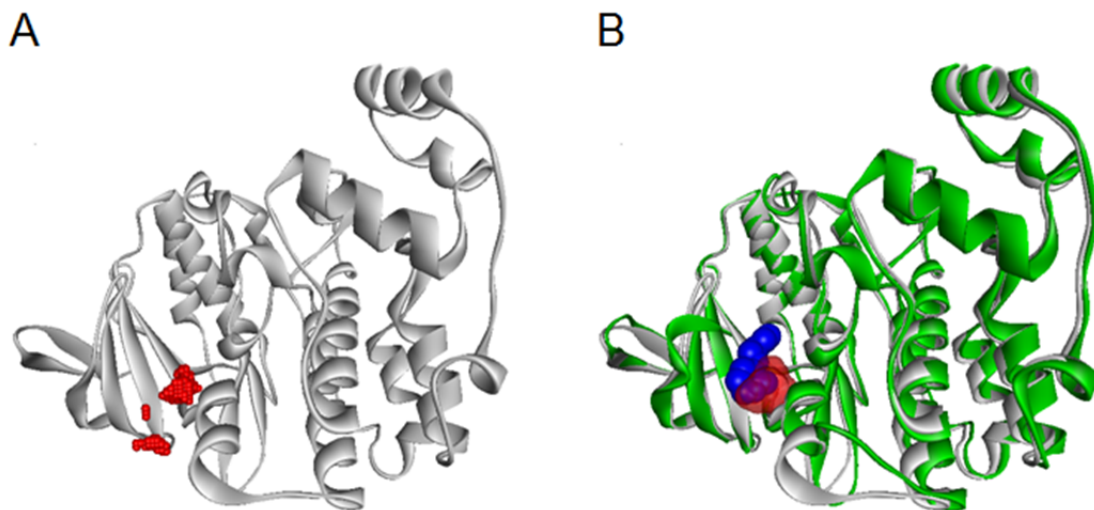
この値と 7.7 に記載の手順で準備した p38 $\alpha$  の apo 体の構造 (PDB ID: 1R39)<sup>83</sup> および先の実験結果を基に、7.7 に記載のようにしてグリッドを発生させ、各グリッドについて  $S_{\text{tot}}$  を求めた。結果、実験結果と矛盾がない  $S_{\text{tot}}=1$  (観測されなかったシグナルが 2 つあるため、 $S_{\text{tot}}=1$  が最小値) であるグリッド塊が 3 つ存在し、これらの体積はそれぞれ 176、75、および 26 Å<sup>3</sup> であった (図 4.5A)。阻害剤の体積が 153 Å<sup>3</sup> であることから、2 つのグリッド塊は明らかに小さ過ぎるため、最終的に 1 つのグリッド塊のみが相互作用部位として求められた。このグリッド塊は、最初に発生させた全グリッド、つまり相互作用エネルギーの観点のみから抽出した相互作用部位の候補の 1.0% にあたる 170 個のグリッドから構成されており、本解析法により相互作用部位の候補となる箇所を大きく絞り込んでいることが確認できる。この結果を既知の holo 体の構造 (PDB ID: 1W7H)<sup>82</sup> と比較したところ、グリッド塊が存在する場所は holo 体の構造で確認される相互作用部位に極めて近い場所であった (図 4.5B)。このグリッド塊と阻害剤の重心間距離 ( $D_c$ )、およびグリッド塊の重心と阻害剤の重原子との最短距離 ( $D_{\text{min}}$ ) は、それぞれ 5.32 Å および 2.81 Å であった。相互作用部位を解析する際、apo 体の構造の代わりに holo 体の構造を使用したところ、 $D_c$  および  $D_{\text{min}}$  はそれぞれ 4.60 Å および 1.10 Å となり、相互作用部位とされるグリッド塊の位置は、holo 体の構造で確認される阻害剤の相互作用部位により近づいた。これは、apo 体と holo 体とでは相互作用部位周辺の構造が変化しているために、holo 体の構造を用いた方がより良好な結果が得られたものと考えられる。

以上のように、本手法により、一部のシグナルが観測されない実験結果からでも、低分子との相互作用部位の同定が可能であることが示された。



Adapted with permission from J. Med. Chem., 2013, 56 (22), pp 9342–9350. Copyright 2013 American Chemical Society.

図 4.4. 安定同位体で標識した p38 $\alpha$  への 2-amino-3-benzoyloxypyridine 添加実験。(A) 標識 p38 $\alpha$  の各シグナルの  $\omega_{RMS}$ 。赤線は、化学シフト変化の大小を判定するための閾値である  $\omega_{tp}$  を示す。(B) 標識 p38 $\alpha$  の time-shared [ $^1H, ^{13}C_{aromatic}/^{15}N$ ]- $\omega_1$ -TROSY-HSQC / [ $^1H, ^{13}C_{methyl}$ ]-HSQC スペクトル。阻害剤の滴定前後のスペクトルを重ね、着目する化学シフト範囲を拡大して示す。黒は滴定前、赤は滴定後のスペクトルであり、矢印は(A)に示す  $\omega_{RMS}$  が  $\omega_{tp}$  より大きいシグナルの摂動先を示す。#は[ $^{13}C_{\epsilon}$ ]Met 由来のシグナル、\*は time-shared 実験により同じ領域に現れてしまった[ $^{13}C_{\epsilon}$ ]His の anti-TROSY 成分である。



Adapted with permission from J. Med. Chem., 2013, 56 (22), pp 9342–9350. Copyright 2013 American Chemical Society.

図 4.5. 本手法により同定した 2-amino-3-benzoyloxypyridine の相互作用部位。(A) *apo* 体の構造を用いて解析した結果  $S_{\text{tot}}=1$  であったグリッドを赤で、*apo* 体の構造を灰色のリボンモデルで表示した。(B) *holo* 体の構造を主鎖の重原子でフィッティングし重ね合わせた。2-amino-3-benzoyloxypyridine を青の CPK モデル、*holo* 体の p38α を緑のリボンモデル、*apo* 体の構造を用いて同定された相互作用部位を赤で表した。

## 第 5 章 天然リガンドの相互作用部位同定への応用

### 5.1 MAP kinase insert region および脂質に関する先行知見と課題

近年、幾つかの脂肪酸やアラキドン酸等の脂肪酸誘導体が、p38MAPK カスケードを活性化すると報告がある<sup>84-87</sup>。また、Diskin らの報告<sup>88</sup>では、*n*-octyl- $\beta$ -D-glucopyranoside ( $\beta$ -OG)が MAP kinase insert region と呼ばれる領域へ結合することが共結晶の構造解析により確認されており、アラキドン酸が同じ領域へ結合することを示唆する変異実験の結果も得られていることから、MAP kinase insert region への脂質の結合が p38 $\alpha$  の機能へ影響を及ぼす可能性についても議論されている。さらに、PIA23、PIA24、perifosine との複合体構造<sup>89</sup>とそれらが p38 $\alpha$  に及ぼす影響に関する生化学的知見<sup>90</sup>が報告されているほか、p38 $\alpha$  の活性化を調節している kinase の 1 つ PTPN5 が、MAP kinase insert region 付近に相互作用することが示唆されている<sup>91</sup>。これらの知見は、MAP kinase insert region が p38 $\alpha$  の機能に影響を及ぼす新規のアロステリックサイトであり、脂肪酸がその領域へ結合する天然リガンドである可能性を示唆しているものの、直接的な証拠は未だ報告されていなかった。溶液状態での相互作用を直接確認する手段としては NMR による分析が有用であるが、分子量が大きく、また p38 $\alpha$  のダイナミクスに関連していると予想される化学交換により、着目している領域の残基に由来するシグナルの帰属が困難であった<sup>92</sup>。そこで、第 4 章で述べた手法を適用することで、p38MAPK カスケードを活性化すると報告されているデカン酸が p38 $\alpha$  へ結合するか確認し、結合する場合にはその相互作用部位を同定することにした。

### 5.2 p38 $\alpha$ mitogen-activated protein kinase の脂肪酸相互作用部位の同定

デカン酸の相互作用部位の同定は、阻害剤の相互作用部位を同定した際と同様の手順にて行った。まず、先の実験と同様に調製した標識 p38 $\alpha$  を用意し、7.6 に記載の通りモル比で 3 倍量になるまでデカン酸ナトリウムを滴定し、各シグナルの化学シフト変化を time-shared [ $^1\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}$ ]-TROSY-HSQC / [ $^1\text{H}, ^{13}\text{C}_{\text{methyl}}$ ]-HSQC 実験により観測した。化学シフト値の変化量から各シグナルの  $\omega_{\text{RMS}}$  と  $\omega_{\text{tp}}$  を求め、 $\omega_{\text{tp}}$  より大きいシグナルの個数を求めたところ、Met: 1、Ala: 1、His: 1、Tyr: 0、Trp: 1、Phe: 0 であった(図 5.1)。4.5 と同様、His および Phe については観測されたシグナルの数が期待される数よりも 1 つ少ないため、最終的に相互作用部位周辺に存在する個数は Met: 1、Ala: 1、His: 1 or 2、Tyr: 0、Trp: 1、Phe: 0 or 1 となった。阻害剤滴定実験の結果に比べ化学シフト変化量が小さいが、これはデカン酸の濃度が高くなるにつれてスペクトルの質が低下し解析が困難になるため、デカン酸ナトリウムを多量に添加できず、したがって系中の結合状態の存在比が小さいためである。また、環電流の影響により周囲の原子に大きな化学シフト変化を誘起する芳香環を脂肪酸が有していないことも、原因の 1 つと考えられる。

次に、Pubchem 3D から入手したデカン酸の構造を用いて  $r_{\text{seb}}$  を求めたところ、7.151 Å であった。p38 $\alpha$  の構造として、*apo* 体の構造 (PDB ID: 1R39)、阻害剤との *holo* 体の構造 (PDB ID: 1W7H)、および  $\beta$ -OG との *holo* 体の構造 (PDB ID: 2NPQ) を使用して解析した結果、実験結果に最も一致

しデカン酸以上の体積を有するグリッド塊は、 $\beta$ -OG との相互作用部位にのみ見られた(図 5.2A, B)。この結果は、溶液中においてデカン酸が  $\beta$ -OG と同様に MAP kinase insert region へ相互作用することを強く示唆している。

$\beta$ -OG(およびデカン酸)との相互作用部位には、脂質との相互作用において重要と考えられている Trp197 が存在するため、 $\beta$ -OG またはデカン酸がこの箇所へ相互作用するならば、Trp197 由来のシグナルの化学シフトが大きく変化するはずである。そこで、 $\beta$ -OG の滴定実験を行ったところ、デカン酸ナトリウムの滴定実験結果と同様の化学シフト変化が Trp 由来のシグナルで観測された(図 5.2C, D)。この化学シフト変化が観測されたシグナルが Trp197 に由来するか、Trp197 を Phe に置換した W197F 変異体を調製して NMR 測定を行い確認した。結果、上記のシグナルは W197F を測定したスペクトルでは消失しており、すなわちこのシグナルは Trp197 に由来することが確認された(図 5.2E)。したがって、デカン酸の相互作用部位は、本研究で開発した手法で同定された場所と同じであることが強く示唆された。また、デカン酸が MAP kinase insert region へ相互作用することが初めて示された。

上記結果から、薬様化合物だけでなく、天然リガンド等の相互作用部位を同定する際にも本手法を活用できることが示された。また、化学シフト変化が小さい化合物を対象とする場合や、結合状態の存在比が小さい場合においても適用できるなど、本手法の頑健性が示された。

### 5.3 他のアプローチとの比較

近年、計算法学的手法による低分子結合部位の予測法は著しい発展を遂げており<sup>64, 93-121</sup>、これらのアプローチは、いわゆる”結合ポケット”と見なされるような形状の深い窪みから成る結合部位については、良好な予測精度を示す。例えば、最近開発された MolSite<sup>122</sup> と呼ばれる手法は、計算量が多いものの 80~90%という高い予測率を誇る。しかしながら、明らかに結合ポケットと見なされるような深い窪みではない相互作用部位については、未だ予測が困難である。一方、本手法は実験データを基にしたアプローチのため、4.4 に示すように、そのような相互作用部位についても問題なく同定することが可能である。

他の既報のアプローチとしては、NMR データと計算法学のアプローチを組み合わせた McCoy らによる報告<sup>123</sup>があり、シグナル未帰属での相互作用部位同定へ応用できる可能性についても議論されている。しかしながら、彼らの方法は複数の <sup>15</sup>N 標識サンプルとそれらを用いた滴定実験を必要とする他、低分子の芳香環の環電流により誘起される化学シフト変化のみを想定したものとなっている。一方で、本手法は 1 サンプルについての実験のみを必要とし、低分子化合物の化学構造に左右されない、汎用性の高い手法である。

以上のように、本手法は迅速性、簡便性、汎用性の観点から、既報の手法に比べ有用であると考えられる。なお、本手法ではタンパク質を観測対象とするが、低分子を観測対象とする SOS-NMR<sup>124</sup> や INPHARMA<sup>125</sup>、DIRECTION 法<sup>126</sup>等の手法による解析と互いに補う形での利用がより有効である。



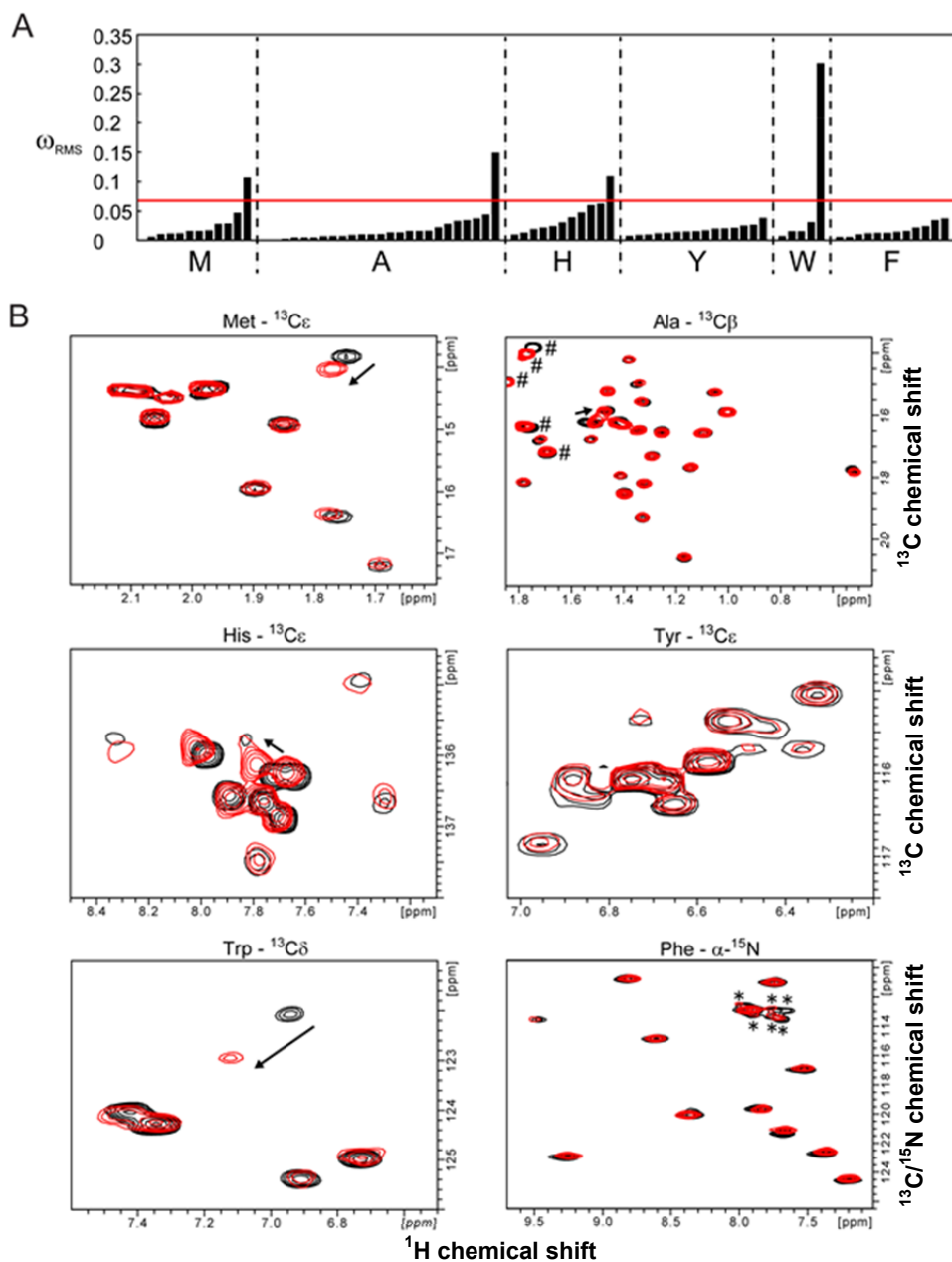
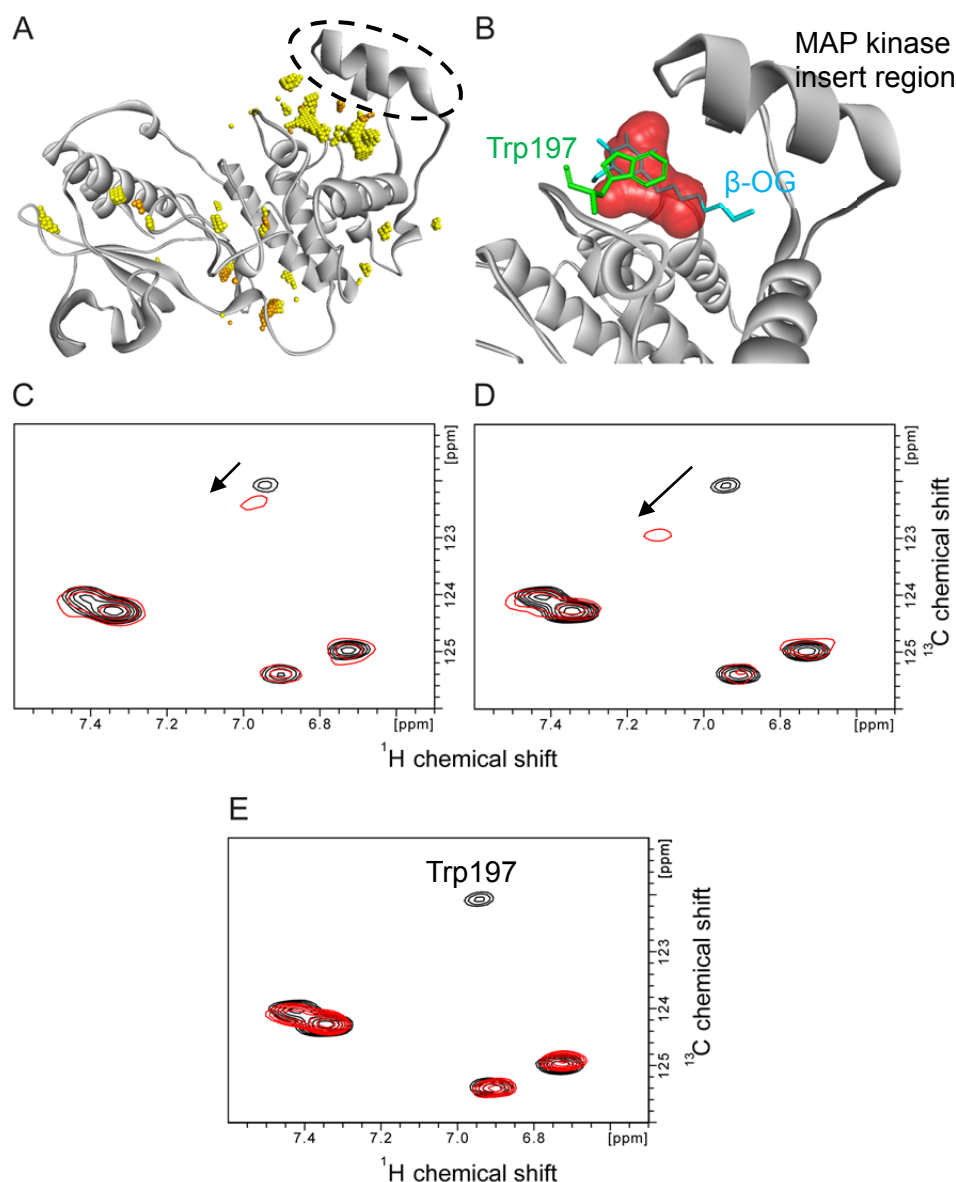


図 5.1. 安定同位体で標識した p38 $\alpha$  へのデカン酸ナトリウム添加実験。(A) 標識 p38 $\alpha$  の各シグナルの  $\omega_{\text{RMS}}$ 。赤線は、化学シフト変化の大小を判定するための閾値である  $\omega_{\text{tp}}$  を示す。(B) 標識 p38 $\alpha$  の time-shared  $[\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}]-\omega_1\text{-TROSY-HSQC}$  /  $[\text{H}, ^{13}\text{C}_{\text{methyl}}]-\text{HSQC}$  スペクトル。デカン酸ナトリウムの滴定前後のスペクトルを重ね、着目する化学シフト範囲を拡大して示す。黒は滴定前、赤は滴定後のスペクトルであり、矢印は(A)に示す  $\omega_{\text{RMS}}$  が  $\omega_{\text{tp}}$  より大きいシグナルの摂動先を示す。#は  $^{13}\text{C}_\epsilon$ Met 由来のシグナル、\*は time-shared 実験により同じ領域に現れてしまった  $^{13}\text{C}_\epsilon$ His の anti-TROSY 成分である。



Adapted with permission from J. Med. Chem., 2013, 56 (22), pp 9342–9350. Copyright 2013 American Chemical Society.

図 5.2. p38α のデカン酸結合部位の同定。(A) β-OG との *holo* 体の構造 (PDB ID: 2NPQ、灰色のリボンモデル) を用いた解析結果。橙色:  $S_{\text{tot}}=2$ 、黄色:  $S_{\text{tot}}=3$  のグリッドを示す。 $S_{\text{tot}}=1$  のグリッドはなかった。MAP kinase insert region を破線で囲んだ。(B) MAP kinase insert region 付近の拡大図。*holo* 体の構造中の β-OG をシアンスティックモデル、Trp197 を緑スティックモデル、本手法により同定された相互作用部位を赤で表した。(C, D) β-OG(C) またはデカン酸ナトリウム(D) 添加前後の標識 p38α の time-shared [<sup>1</sup>H, <sup>13</sup>C<sub>aromatic</sub>/<sup>15</sup>N]-ω<sub>1</sub>-TROSY-HSQC / [<sup>1</sup>H, <sup>13</sup>C<sub>methyl</sub>]-HSQC スペクトルを重ね、Trp 由来シグナルが観測される化学シフト範囲を拡大して示した。黒は滴定前、赤は滴定後のスペクトルである。(E) WT (黒) と W197F 変異体 (赤) のスペクトルを重ねて表示した。W197F のスペクトルで消失したシグナルが Trp197 に由来することが分かる。

## 第 6 章 総括および今後の展望

### 6.1 タンパク質—タンパク質間相互作用部位の迅速解析法と課題

#### 6.1.1 タンパク質—タンパク質間相互作用部位の迅速解析法

本研究で開発した PPI 界面の迅速解析法を適用することで、解析対象のタンパク質の立体構造さえ既知であれば、1 サンプルについて簡単な滴定実験と NMR 実験を行い、シグナル未帰属のまま化学シフト変化を観測するだけで、迅速かつ簡便に PPI 界面を同定することが可能である。この手法の実現においては、(1)注意深く選択された原子特異的安定同位体標識の組み合わせ、(2)新規に開発した  $^1\text{H}$ - $^{13}\text{C}$  相関シグナルと  $^1\text{H}$ - $^{15}\text{N}$  相関シグナルを同スペクトル上で高分解能に観測する測定法、(3)化学シフト変化の大きさを判定するための閾値の設定法、(4)実験結果から PPI 界面の残基を同定するプログラムの 4 点が鍵となっている。

(1)については幾つかの組み合わせが考えられるが、本研究で用いた組み合わせ (Met, Ala, His, Tyr, Trp, Arg) を使用することで、統計的には PPI 界面に存在するアミノ酸残基の 34% がカバーされており<sup>12</sup>、大半の PPI 界面を探索するうえで十分な情報が得られるものと考えられる。もし、先の組み合わせでは情報が不足していると予想される場合、あるいはより多くのアミノ酸残基種についても情報が必要な場合には、観測プローブとするアミノ酸残基の種類をさらに増やせばよい。例えば、 $[1-^{13}\text{C}]$  標識されたアミノ酸も取り込ませカルボニル  $^{13}\text{C}$  を直接観測する、あるいは  $[^{13}\text{C}\alpha, \alpha-^{15}\text{N}]$  標識されたアミノ酸も取り込ませカップリングにより  $[\alpha-^{15}\text{N}]$  標識アミノ酸と区別する、等が考えられる。また、コストや時間の観点からより現実的な手法とするため、本研究では一般に入手が容易な標識アミノ酸を用いたが、例えば  $^{13}\text{C}\delta$  標識された Ile を合成して用いる、あるいは SAIL アミノ酸<sup>127</sup>等を利用することで、観測プローブとするアミノ酸残基種のさらなる追加が可能である。原子特異的標識の恩恵としてスペクトル上のシグナル数が少ないために縮重が起こりにくく、それゆえ本手法では比較的大きなタンパク質への適用が可能となっているが、分子重量 50 kDa 程度以上のより大きなタンパク質を対象とする場合には、横緩和を抑えるために部分的に重水素化した標識アミノ酸の利用が望ましく<sup>128, 129</sup>、先に挙げた SAIL アミノ酸等の利用も有用である。

今回、(2)の新規測定法を開発し高分解能スペクトルの取得を可能にすることで、比較的大きなタンパク質を扱う際にも time-shared 実験の利点である測定時間の短縮を可能としている。YUH を測定した結果を見る限りでは、標準的な time-shared 実験<sup>26</sup> に比べ平均で約 80% の感度であり、多少の感度低下はあるものの、本研究で確認された通り、通常は問題とならないと考えられる。

(3)に関しては、過去の研究において議論されてこなかった点である。近年、情報取得における迅速性の要求が高まる一方で、解析対象となるタンパク質の分子量は大きくなり、また膜タンパク質など見かけの分子量が大きいものを対象とした研究も多い。そのような場合において、スペクトルの複雑さや緩和に起因する解析困難な状況を回避する際、アミノ酸選択的標識は有力なアプローチの 1 つであり多用されていることから、今回示すバイアスを考慮した解析法が、実験結果をより正しく理解するうえで有効なケースは多いと考える。

(4)に関して、専用の解析プログラムを開発し適用することで、解析者に依らず同じ結果を迅速

に取得することが可能となった。実験結果の曖昧さを許容しつつ妥当な結果を得るために、3.2 に示す通り、全ての PPI 界面の残基が正しく選択されるわけではない。しかしながら、平均で約 90% の残基が同定されており、PPI 界面を同定するうえで十分な精度の結果を得られる。

#### 6.1.2 タンパク質複合体モデル構造の構築

PPI 界面の迅速解析法では、相互作用対象のタンパク質の構造は未知でよいが、もし既知の場合には、両タンパク質について PPI 界面に関与する残基を同定し、その情報を 3.5 に示すように拘束条件として使用することで、実験結果を反映させた確からしい複合体モデル構造の取得が可能である。本研究で得られた情報から入力可能な拘束条件は、相互作用界面に関与する残基の指定のみであるため、タンパク質間の相対的な配向の予測は、ドッキングシミュレーションで使用するスコアリング関数の性能に委ねられている。本研究で確認した限りでは、3.5 に記載の通り高精度の複合体モデル構造の構築が可能であったが、さらに精度を高める必要がある場合は、residual dipolar coupling (RDC)<sup>130, 131</sup> などで実験的に決定された配向情報の利用が考えられる。

#### 6.1.3 タンパク質—タンパク質間相互作用部位の迅速解析法における課題

本手法では、化学シフト変化に関する情報を用いているが、周辺の化学的環境の変化に応じて化学シフトが変化しているため、相互作用以外に構造変化によっても化学シフト変化が起こる。化学シフト変化がどちらに由来するものか、滴定実験の結果だけから区別することは困難であるため、実験結果に不適切な情報が含まれている可能性は否定できない。そのような観点からも、得られる実験結果には曖昧さが含まれるが、3.1 でも述べたように、解析において立体構造という曖昧性が極めて低い情報を用い、専用の解析プログラムによりもっとも妥当な結果を得られるようにすることで、可能な限り問題を回避している。また、区別するための情報を実験的に取得するためには、例えば常磁性添加物を用いた溶媒露出度に関する情報の取得<sup>44, 132–136</sup>が有用であると考えられる。また、過去の研究例において、PPIにより大きな構造変化が起こる系を対象に既報の迅速解析法を適用した例<sup>137</sup>があるが、得られた結果を基に確からしい複合体モデル構造の構築に

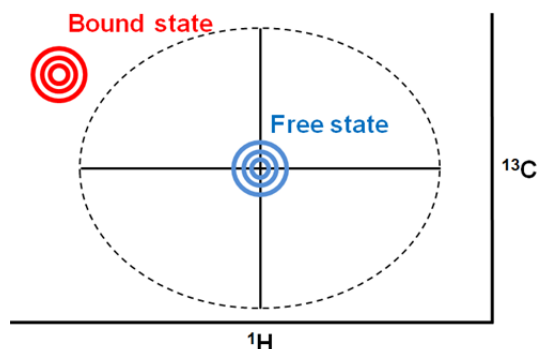


図 6.1. 強い相互作用系での解析例。相互作用対象を滴定した後のシグナルが一定の範囲内に観測されない場合、そのシグナルの化学シフト変化は大きいと見做す。

成功している。さらに、同様に構造変化が大きい系について、HADDOCKを用いてドッキングシミュレーションを行った研究結果においても、良好な結果が得られている<sup>51</sup>。以上から、PPIに伴い大きな構造変化が起こる系に対しても、本手法の適用は決して不可能ではなく、構造変化に伴う情報の不確実性をいかにして時間と労力をかけずに除くかが、本手法を今後発展させるうえで重要となる。

別の課題としては、いわゆる強い相互作用系において、相互作用対象の滴定前後におけるシグナルの対応を確認することが容易ではない点が挙げられる。この場合、相互作用に伴う各シグナルの化学シフト変化量は不明であるため、本研究で用いたような定量的な解析によるシグナルの抽出はできない。また厳密には、滴定前後で同じ化学シフト値をもつシグナルが観測されたとしても、それらのシグナルが同じアミノ酸残基に由来しているか否かは不明である。このような場合においては、例えば変異の導入により相互作用を弱くするなどの対応が考えられるが、適切な変異の探索と確認には時間と労力を要するため、より良い別のアプローチを検討する必要があるかもしれない。実践的な観点からは、例えば滴定前のシグナルの化学シフト値より一定範囲内に滴定後のシグナルが存在しなければ、そのシグナルは化学シフトが大きく変化したと見做す、などの解析方法が考えられる(図 6.1)。まだ検討段階ではあるが、この解析方法を適用することで、Y-Ub との相互作用が強い YUH の変異体 C90S<sup>43,44</sup>を用いた実験において、相互作用界面の同定に成功している(データ未掲載)。

他の課題として、シグナルの広幅化のために一部のシグナルが観測されない場合の解析法が規定されていないが、3.4 で頑健性を検証した際の結果から、少なくとも全残基数の 1%程度が観測されない場合においても、それらの残基を相互作用界面に存在すると仮定しても存在しないと仮定しても、PPI 界面の同定に大きな影響はないと予想される。

## 6.2 タンパク質—低分子間相互作用部位の迅速解析法と課題

### 6.2.1 タンパク質—低分子間相互作用部位の迅速解析法

先に述べた PPI 界面を迅速に解析する手法を基に改良することで、低分子の相互作用部位を迅速に同定する手法を開発した。本手法を構成する要素のうち、(1)実験結果に最も一致する空間の探索法、(2)観測されないシグナルがある場合の適切な対処法の 2 点が特に重要である。

(1)については、各グリッドの評価法を、周囲に存在するアミノ酸残基の種類とその数がどれだけ実験結果に一致しているかとし、多少の不一致も許容することで、実験結果の曖昧さやタンパク質の多少の構造変化にも対応できるようにした。実際、4.5 において、相互作用部位に構造変化がある *holo* 体の構造を使用した方が良好な結果が得られているものの、*apo* 体の構造を用いた場合においても十分な精度で相互作用部位が同定されている。各グリッドの評価においては、低分子の構造を基に仮想球を設定しているが、低分子の長辺方向の長さから仮想球の半径が規定されるため、形状の異方性が高い場合は低分子の短辺方向について、本来化学シフトへ影響を及ぼさない距離にあるアミノ酸残基もカウントしてしまうことになる。そのため、形状の異方性が高

い場合には不適切なグリッドが残り易く、また全体的にグリッドの不一致度が高くなる。5.2 でデカン酸の相互作用部位を同定した際には、その影響が表れていると考えられる。形状の異方性に合わせて仮想球のサイズを調整する等、今後何らかの対処を行うことで、結果がより改善される可能性が考えられる。

(2)に関しては、観測されないシグナルが相互作用部位に存在する場合と存在しない場合の全てを考慮することで、解析結果へバイアスがかからないようにした。酵素など創薬の対象となる可能性が高くかつ柔軟性が高いタンパク質において、p38 $\alpha$  のように一部のシグナルが観測されないケースが頻出すると考えられるが、本解析法によりそのような場合においても適用を可能にしている。

#### 6.2.2 タンパク質—低分子間相互作用部位の迅速解析法における課題

課題については 6.1.3 に記載の内容とほぼ同様であり、構造変化がある場合の解析法と、いわゆる強い相互作用系における解析法が挙げられる。このうち、構造変化への対処としては、先に述べた通り実験結果に一致する空間の探索法で多少の不一致を許容している他、6.1.3 と同様、溶媒露出度に関する情報の取得が有用と考えられる。一方で、*apo*体では結合部位が埋もれているような、相互作用に伴う大きな構造変化が起こるタンパク質の場合、本手法をそのまま適用するのは困難である。事実、5.2 においてデカン酸結合部位を同定する際、結合部位の空間が形成されていた  $\beta$ -OG との *holo* 体の構造以外を用いると、適切なグリッド塊が現れず結合部位が同定されなかった。このような一過性の結合ポケットの同定に関しては、6.2.3 に述べるアプローチが有効かもしれない。

強い相互作用系における問題に関しては、タンパク質—タンパク質間相互作用部位の場合と同様のアプローチが考えられるが、一方で本手法の活用が期待される創薬研究の初期段階、とくにフラグメントと呼ばれる化合物を対象とした研究において、対象となる化合物との相互作用は強い場合が多いため、実際に手法を適用する際に問題となるケースは少ないと予想される。

#### 6.2.3 一過性ポケットの同定および低分子との複合体モデル構造の構築へ向けた展望

6.2.2 に述べた通り、本手法をそのまま適用し、一過性ポケットの同定を *apo* 体の構造から行うのは困難である。一方で、p38 $\alpha$  について *apo* 体の構造と  $\beta$ -OG との *holo* 体の構造を比較すると、一過性ポケット周辺の主鎖構造は両者であまり差がない一方で、側鎖のコンフォメーションが大きく異なることが確認された。したがって、*apo* 体の主鎖構造を用いて類似の解析を行うことで、一過性ポケットである  $\beta$ -OG の相互作用部位を同定できる可能性が考えられる。また、p38 $\alpha$  に限らず、他のタンパク質の一過性ポケットについても同様のアプローチが有効である可能性がある。その他アプローチとしては、既報の計算化学的手法による一過性ポケットが開いた構造の予測<sup>138-142</sup> が有効かもしれない。また、これまでに示したように本手法はある程度の曖昧性を許容するため、そもそもタンパク質の構造が不明な場合には、モデル構造の利用が可能かもしれない。もっとも、その場合には変異実験などによるバリデーションの実施が望ましい。

今回開発した手法では相互作用部位の同定までを行うが、相互作用についてより詳細な解析と予測を行うためには、複合体モデル構造の構築も行うことが望ましい。本手法では、NMR 実験において各シグナルの化学シフト値の変化量に着目しているが、各原子の化学シフト値は、相互作用に伴う周囲の化学的環境の変化を反映した結果として、低磁場側または高磁場側に変化している。したがって、NMR 実験の結果には、相互作用に関する構造的観点からの情報が含まれている。そこで、これらの情報を活用した複合体モデル構造の構築法として、下記のようなアプローチが考えられる。

- (1) 同定した相互作用部位を標的に、低分子化合物とのドッキングシミュレーションを行う。
- (2) 各ドッキング結果について、低分子の周囲に存在する観測プローブとしているアミノ酸残基種の標識原子へ、低分子が及ぼす化学シフト変化を計算から求める。
- (3) 計算から求めた化学シフト変化のパターンと、実験結果から確認されるパターンがなるべく近い構造を複数、バリエーションを持たせて抽出する。
- (4) 抽出した各構造の全原子について molecular dynamics (MD)計算を行い、一定時間おきにスナップショットを撮っていく。
- (5) 各スナップショットの構造について、低分子やアミノ酸残基など標識原子の周辺環境が及ぼす化学シフト変化の影響を計算し、その変化のパターンと実験結果から確認されるパターンがもっとも一致する構造を複合体構造とする。

化学シフト変化の影響を計算するためにはある程度の計算量を必要とするが、必要に応じて環電流の影響のみを考慮するなどして、現実的な計算量に落とし込めばよい。このアプローチにより、タンパク質側の構造も最適化した複合体構造の取得が期待される。

## 6.3 結論

解析対象のタンパク質の立体構造さえ既知であれば、本研究で開発した手法を適用することで、1 サンプルについて簡単な滴定実験と NMR 実験を行いシグナル未帰属のまま化学シフト変化を観測するだけで、迅速かつ簡便に PPI 界面と低分子相互作用部位を同定することが可能である。さらには、実験結果を反映させた、確からしいタンパク質複合体モデル構造の迅速な構築も可能である。本手法は、種々の生化学的知見を迅速かつ簡便に取得することが可能であり、創薬研究の初期段階において構造情報に指南された化合物設計を行うため、あるいはタンパク質工学による有用タンパク質創出のための構造情報の取得において、特に有用である。

なお、本研究成果は、Elsevier より出版された Journal of Structural Biology 2011 年 174 巻 434 頁から 442 頁 (<http://dx.doi.org/10.1016/j.jsb.2011.04.001>、本論文中の図 2.2、2.4、2.6、2.7、3.2、3.3、3.5 および表 2、3、4 は当該論文の著者原稿よりそのまま、またはその一部を改変して使用) ならびに ACS より出版された Journal of Medicinal Chemistry 2013 年 56 巻 9342 頁から 9350 頁 (<http://pubs.acs.org/articlesonrequest/AOR-ufIYPjrmngFcJBdJ3I2J>、DOI: 10.1021/jm4014357) に掲載された。

## 第 7 章 材料と方法

### 7.1 試薬および立体構造表示ソフトウェア

#### 7.1.1 試薬

各実験で使用したオリゴ DNA は、Life Technologies Corporation にて合成した。安定同位体標識された各試薬は、Cambridge Isotope Laboratories Inc. または Icon Services Inc. より購入した。他の各試薬は、特に記載がない限りナカライテスク株式会社製の特級試薬を用いた。

#### 7.1.2 立体構造の表示

各立体構造の表示と画像の作成には、PyMOL 1.5.0.3 (Schrödinger, LLC) または DS Visualizer 3.0 (Accelrys, Inc.) を用いた。

### 7.2 ユビキチンおよびユビキチン加水分解酵素の発現・精製および安定同位体標識

#### 7.2.1 安定同位体標識されたユビキチン加水分解酵素 (YUH) のセルフリー合成系および精製系

安定同位体で標識された YUH は、下記の方法にて調製した。まず、*Saccharomyces cerevisiae* 由来 *YUH1* 遺伝子配列を鋳型に、下記のオリゴ DNA を用いて常法に従い PCR 反応を行うことで、5' 末端側に tobacco etch virus (TEV) protease 認識サイトの配列を付加した *YUH1* 遺伝子配列の DNA 断片を取得した。

```
TAATTGCGGCCGCGAAAAACCTGTACTTCCAGGGTATGAGCGGAGAAAATCGTGCT  
TAATACTCGAGTCATTCCCAATTAGGGCCCA
```

得られた DNA 断片を常法に従い精製し、制限酵素 NotI (Takara Bio Inc.) および XhoI (Takara Bio Inc.) で消化し精製した後、同制限酵素で消化し精製した pIVEX2.4d プラスミド (Roche Diagnostics GmbH) と、Ligation high Ver.2 (Toyobo Co., Ltd.) を用いてライゲーションした。このライゲーション産物で *Escherichia coli* DH5 $\alpha$  を形質転換し、アンピシリン耐性を指標にプラスミドを有する形質転換体を取得した後、常法に従い形質転換体を培養しプラスミドを精製した。その後、DNA シーケンサーにより塩基配列を確認し、目的とする YUH 合成用プラスミド (pIVEX2.4d-YUH) を取得した。

次に、RTS Amino Acid Sampler (Roche Diagnostics GmbH) 付属のプロトコルに従い、各標識アミノ酸 (Met- $^{13}\text{C}_\epsilon$ 、Ala- $^{13}\text{C}_\alpha$ 、 $^{13}\text{C}_\beta$ 、His- $^{13}\text{C}_\epsilon$ 、Tyr- $^{13}\text{C}_\epsilon$ 、Trp- $^{13}\text{C}_\delta$ 、Arg- $\alpha$ - $^{15}\text{N}$ ) のストック溶液を調製後、*Escherichia coli* をベースとしたセルフリータンパク質合成キット RTS 500 ProteoMaster E. coli HY Kit (Roche Diagnostics GmbH) と pIVEX2.4d-YUH を用いて、付属のプロトコルに従い 303 K で目的タンパク質を合成した。

得られた合成産物は、下記の手順にて精製し、安定同位体で標識された実験用サンプルを調製した。まず、HisTrap HP column (GE Healthcare UK Ltd.) を用いたアフィニティークロマトグラフィーで常法に従い精製後、ヒスチジンタグが付加された TEV protease である AcTEV Protease



(Life Technologies Corp.)を用いてタグを切断した。続いて、HisTrap HP columnを用いて再度アフィニティー精製を行い、AcTEV Protease、切断したタグおよび最初のアフィニティー精製でカラム充填物との非特異的結合により残留していた不純物タンパク質を除去した。最後に、Superdex 75 10/300 GL (GE Healthcare)と溶出バッファ(50 mM リン酸ナトリウム、100 mM 塩化ナトリウム、2 mM DTT、pH 6.0)を用いたサイズ排除クロマトグラフィーで精製した。

### 7.2.2 未標識ユビキチン加水分解酵素(YUH)の発現系および精製系

未標識 YUH は、下記の方法にて調製した。まず、*YUH1* 遺伝子の配列を鋳型に、下記のオリゴ DNA を用いて常法に従い PCR 反応を行うことで、5' 末端側に TEV protease 認識サイトの配列を付加した *YUH1* 遺伝子配列の DNA 断片を取得した。

TAATTGCTAGCGAAAACCTGTACTTCCAGGGTATGAGCGGAGAAAATCGTGC

TAATAAAGCTTTTCATTCCCAATTAGGGCCCAATC

得られた DNA 断片を常法に従い精製し、制限酵素 *NheI* (Takara Bio Inc.)および *HindIII* (Takara Bio Inc.)で消化し精製した後、同制限酵素で消化し精製した pET-28b(+)プラスミド (Merck KGaA)と、Ligation high Ver.2 (Toyobo Co., Ltd.)を用いてライゲーションした。このライゲーション産物で *Escherichia coli* DH5 $\alpha$  を形質転換し、カナマイシン耐性を指標にプラスミドを有する形質転換体を取得した後、常法に従い形質転換体を培養しプラスミドを精製して、DNA シーケンサーにより塩基配列を確認することで、目的とする YUH 発現用プラスミド(pET28b-YUH)を取得した。

次に、pET28b-YUH で形質転換した *Escherichia coli* BL21(DE3)を LB 培地で 310 K にて振盪培養し、波長 660 nm における濁度が約 0.6 となった時点で終濃度 1 mM の isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG)を添加し、続けて 16 時間、振盪培養した。集菌後、菌体を破碎用バッファ(50 mM Tris-HCl、2 mM DTT、500 mM phenylmethylsulfonyl fluoride (PMSF)、pH 8.0)にて懸濁し、超音波破碎して遠心後、その上清を Millex-GV 0.22  $\mu$ m PVDF Durapore membrane (Millipore Corp.)で濾過した。この濾液から、7.2.1 に記載の手順にて YUH を精製した。

### 7.2.3 未標識および安定同位体標識されたユビキチン(Y-Ub)の調製

未標識 Y-Ub は、既報<sup>40</sup>の手順に従い調製した。安定同位体で標識された Y-Ub は、未標識 Y-Ub の調製に用いられたプラスミドを使用して、7.2.1 に記載した YUH の合成と同じ手順にて合成し、未標識 Y-Ub の調製と同じ方法で精製した。

## 7.3 ユビキチンとユビキチン加水分解酵素との相互作用実験

滴定実験は、サンプルバッファ(50 mM リン酸ナトリウム、100 mM 塩化ナトリウム、2 mM 重水素化 DTT、5% 重水、pH 6.0)に溶解した 0.19 mM の標識 YUH および 0.14 mM の標識 Y-Ub を調製し、同バッファに溶解した未標識の相互作用対象をモル比 4 倍量まで添加することで行った。

NMR 測定は、クライオ型 TCI プローブを装着した AVANCE600 (Bruker BioSpin GmbH)を用いて、

303 K にて行った。標識 YUH への Y-Ub 滴定実験では、2.2 に記載した新開発のパルスシーケンスを用いて、time-shared  $[^1\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}]-\omega_1$ -TROSY-HSQC/ $[^1\text{H}, ^{13}\text{C}_{\text{methyl}}]$ -HSQC 実験によりスペクトルを取得した。一方、標識 Y-Ub への未標識 YUH 滴定実験においては、観測され得るシグナル数が少なくシグナルの縮重等に伴う解析上の不具合が生じないため、TROSY 法を適用しない既報の time-shared 実験によりスペクトルを取得した。いずれの実験においても、測定時間の短縮とスペクトルの高解像度化のため、各シグナルが重ならないよう十分に配慮した上で、間接測定軸の折り返しを活用した。スペクトル幅は  $3018.2 \text{ Hz } (^{13}\text{C})/1945.5 \text{ Hz } (^{15}\text{N}) \times 8389.3 \text{ Hz } (^1\text{H})$ 、中心周波数は 55ppm ( $^{13}\text{C}$ )、118ppm ( $^{15}\text{N}$ ) および 4.7ppm ( $^1\text{H}$ )、間接軸観測方向および直接観測軸方向のサンプリングポイント数は 112 および 512、リサイクル遅延時間を 1.0 s とし、32 回積算した。

#### 7.4 ドッキングシミュレーションによる複合体モデル構造の構築

3.3 にて行った、CNS1.2<sup>45</sup> と HADDOCK 2.0<sup>50, 51</sup> を用いたドッキングシミュレーションの詳細は下記の通りである。まず、YUH の立体構造座標を、YUH-Ub 複合体の結晶構造<sup>42</sup> (PDB ID: 1CMX) より準備した。欠損している Glu63-Tyr77 の座標は MODELLER<sup>143, 144</sup> を用いて生成し、amber99 力場パラメーターを適用して cosgene<sup>145, 146</sup> を用いた MD シミュレーションにより最適化した。シミュレーション中は、Phe59-Trp81 の全原子および Phe59-Trp81 より 5 Å 以内にある側鎖の原子をフレキシブルに、その他の原子は、結晶構造での座標周辺に位置拘束のポテンシャルを設定して位置を拘束した。シミュレーションは、TIP3P モデル<sup>147</sup> を用いて水分子の影響を陽に取り入れ、300 K で 500 ps 行い、得られた YUH の最終構造を使用した。Ub の立体構造座標には、単体での立体構造として解かれているデータ<sup>148</sup> (PDB ID: 1UBQ) を使用した。

ドッキングシミュレーションは、下記の設定と手順にて行った。まず、3.3 で得られた候補の残基を、HADDOCK を用いたシミュレーションにおける "active residue"、active residue から 4 Å 以内に重原子が存在する残基を "passive residue" に設定した。次に、"ambiguous interaction restraint" を、全ての active residue と、相互作用対象タンパク質の全ての active residue および passive residue との組み合わせとして設定した。フレキシブル領域の設定は、"semiflexible region" は自動設定、"fully flexible region" は末端部を含むループ領域を設定した。また、"scaling factor" を 0.01 に、使用するデータの 25% を無作為に除くようにして、シミュレーションを行った。候補構造のサンプリングについては、it0 ステップで 2000 構造を計算したうち、エネルギー値が最も低い 400 個を it1 ステップとその後に続く水分子を配置した MD によるリファインメントのステップに用いた。リファインメント後は、"ligand RMSD" のカットオフ値を 7.5 Å に設定してクラスタリングを行い、HADDOCK score の値が最も低いクラスタに含まれる、HADDOCK score の値が最も低い構造を複合体モデル構造として抽出した。

#### 7.5 p38 $\alpha$ mitogen-activated protein kinase の合成、精製および変異導入

安定同位体で標識された p38 $\alpha$  は、下記の方法にて調製した。まず、ヒト由来 *MAPK14* 遺伝子の配列を鋳型に、下記のオリゴ DNA を用いて常法に従い PCR 反応を行うことで、5' 末端側に TEV protease 認識サイトの配列を付加した *MAPK14* 遺伝子配列の DNA 断片を得た。

TAATTGCGGCGCGCGAAAACCTGTACTTCCAGGGTATGTCTCAGGAGAGGCCACGTTG  
TAATACTCGAGTCAGGACTCCATCTCTTCTTGG

得られた DNA 断片を常法に従い精製し、制限酵素 NotI および XhoI で消化し精製した後、同制限酵素で消化し精製した pIVEX2.4d プラスミドと Ligation high Ver.2 を用いてライゲーションした。このライゲーション産物で *Escherichia coli* DH5 $\alpha$  を形質転換し、アンピシリン耐性を指標にプラスミドを有する形質転換体を取得した後、常法に従い形質転換体を培養しプラスミドを精製して、DNA シーケンサーにより塩基配列を確認することで、目的とする p38 $\alpha$  合成用プラスミド (pIVEX2.4d-p38 $\alpha$ ) を取得した。

次に、RTS Amino Acid Sampler 付属のプロトコルに従い、各標識アミノ酸 (Met- $^{13}\text{C}_6$ , Ala- $^{13}\text{C}_6$ ,  $^{13}\text{C}_6$ , His- $^{13}\text{C}_6$ , Tyr- $^{13}\text{C}_6$ , Trp- $^{13}\text{C}_6$ , Phe- $\alpha$ - $^{15}\text{N}$ ) のストック溶液を調製後、RTS 500 ProteoMaster *E. coli* HY Kit と pIVEX2.4d-p38 $\alpha$  を用いて、付属のプロトコルに従い 303 K で目的タンパク質を合成した。

得られた合成産物は、下記の手順にて精製し、安定同位体で標識された実験用サンプルを調製した。まず、Ni-NTA Sepharose 6 Fast Flow column (GE Healthcare UK Ltd.) を用いたアフィニティークロマトグラフィーで常法に従い精製後、ヒスチジンタグが付加された TEV protease である TurboTEV Protease (Accelagen Inc.) を用いてタグを切断した。続いて、HisTrap HP column を用いて再度アフィニティー精製を行い、TurboTEV Protease、切断したタグおよび最初のアフィニティー精製でカラム充填物との非特異的結合により残留していた不純物タンパク質を除去した。最後に、Superdex 200 10/300 GL (GE Healthcare) と溶出バッファ (50 mM リン酸ナトリウム、150 mM 塩化ナトリウム、5 mM DTT、pH 6.8) を用いたサイズ排除クロマトグラフィーで精製した。

変異の導入は、QuikChangeII Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.) を用いて、付属のプロトコルに従い行った。

## 7.6 p38 $\alpha$ mitogen-activated protein kinase と阻害剤、脂肪酸との相互作用実験

滴定実験は、サンプルバッファ (5 mM 重水素化 HEPES、150 mM 塩化ナトリウム、5 mM 重水素化 DTT、5% 重水、pH 6.8) に溶解した標識 p38 $\alpha$  を調製し、同バッファに溶解した阻害剤 (2-amino-3-benzoyloxypyridine)、 $\beta$ -OG、デカン酸ナトリウムを各実験にて添加し行った。各サンプルの終濃度は、阻害剤添加時で 0.29 mM p38 $\alpha$  および 1.6 mM 阻害剤 (モル比: 5.5)、 $\beta$ -OG 添加時で 0.43 mM p38 $\alpha$  および 0.29 mM  $\beta$ -OG (モル比: 0.67)、デカン酸添加時で 0.25 mM p38 $\alpha$  および 0.75 mM デカン酸ナトリウム (モル比: 3.0) であった。

NMR 測定は、クライオ型 TCI プローブを装着した AVANCE600 を用いて、303 K にて行った。この際、2.2 に記載した新開発のパルスシーケンスを用いて、time-shared

$[^1\text{H}, ^{13}\text{C}_{\text{aromatic}}/^{15}\text{N}]-\omega_1\text{-TROSY-HSQC}/[^1\text{H}, ^{13}\text{C}_{\text{methyl}}]\text{-HSQC}$  実験によりスペクトルを取得した。また、測定時間の短縮とスペクトルの高解像度化のため、間接測定軸は折り返しを活用した。

## 7.7 タンパク質周辺へのグリッドの発生

第4章および第5章で使用した p38 $\alpha$  の立体構造座標は、下記のようにして準備した。まず、*apo* 体の構造、阻害剤との *holo* 体の構造、および  $\beta$ -OG との *holo* 体の構造を、それぞれ結晶構造 (PDB ID: 1R39<sup>83</sup>、1W7H<sup>82</sup>、および 2NPQ<sup>88</sup>) から準備した。この際、欠損している残基の座標は、Swiss-PDB Viewer<sup>149</sup> 4.0.1 の "Scan Loop Database" 機能を用いて生成した後、該当の残基とその周辺を minimization した中で最もエネルギーが低い構造のものを用いた。

次に、準備した構造の座標と EasyMIFS<sup>65-67</sup> を用いて、阻害剤を対象とした場合は  $-\text{CH}_3$  プローブ、その他の場合は  $-\text{CH}_2$  プローブを用いて、相互作用する可能性が高い箇所を検索しグリッドを発生させた。その際、グリッドの間隔は 0.5 Å とした。その後、SiteHound<sup>65-67</sup> を用いてエネルギー的観点からより適切なグリッドのみを残し、他は削除した。その際の各パラメーターは、"energy threshold" が -8.0、他は標準の値とした。

## 参考文献

1. del Carmen Fernandez-Alonso, M.; Diaz, D.; Berbis, M. A.; Marcelo, F.; Canada, J.; Jimenez-Barbero, J. Protein-carbohydrate interactions studied by NMR: from molecular recognition to drug design. *Current protein & peptide science* **2012**, 13, 816-30.
2. Kofuku, Y.; Ueda, T.; Okude, J.; Shiraishi, Y.; Kondo, K.; Maeda, M.; Tsujishita, H.; Shimada, I. Efficacy of the beta(2)-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nature communications* **2012**, 3, 1045.
3. Moumne, R.; Catala, M.; Larue, V.; Micouin, L.; Tisne, C. Fragment-based design of small RNA binders: promising developments and contribution of NMR. *Biochimie* **2012**, 94, 1607-19.
4. Hiroaki, H. Recent applications of isotopic labeling for protein NMR in drug discovery. *Expert opinion on drug discovery* **2013**, 8, 523-36.
5. Wagstaff, J. L.; Taylor, S. L.; Howard, M. J. Recent developments and applications of saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy. *Molecular bioSystems* **2013**, 9, 571-7.
6. Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., 3rd.; Skelton, N. J.; Rance, M. *Protein NMR Spectroscopy, Second Edition*. Academic Press: London, 2006.
7. Reese, M.; Dötsch, V. Fast mapping of protein-protein interfaces by NMR spectroscopy. *J Am Chem Soc* **2003**, 125, 14250-1.
8. Mayr, L.; Bojanic, D. Novel trends in high-throughput screening. *Curr Opin Pharmacol* **2009**, 9, 580-8.
9. Villoutreix, B.; Eudes, R.; Miteva, M. Structure-based virtual ligand screening: recent success stories. *Comb Chem High Throughput Screen* **2009**, 12, 1000-16.
10. Bogan, A.; Thorn, K. Anatomy of hot spots in protein interfaces. *J Mol Biol* **1998**, 280, 1-9.
11. Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Hot spots--a review of the protein-protein interface determinant amino-acid residues. *Proteins* **2007**, 68, 803-12.
12. Chakrabarti, P.; Janin, J. Dissecting protein-protein recognition sites. *Proteins* **2002**, 47, 334-43.
13. Keskin, O.; Tsai, C. J.; Wolfson, H.; Nussinov, R. A new, structurally nonredundant, diverse data set of protein-protein interfaces and its implications. *Protein Sci* **2004**, 13, 1043-55.
14. Morita, E. H.; Shimizu, M.; Ogasawara, T.; Endo, Y.; Tanaka, R.; Kohno, T. A novel way of amino acid-specific assignment in (1)H-(15)N HSQC spectra with a wheat germ cell-free protein synthesis system. *J Biomol NMR* **2004**, 30, 37-45.

15. Torizawa, T.; Shimizu, M.; Taoka, M.; Miyano, H.; Kainosho, M. Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J Biomol NMR* **2004**, 30, 311-25.
16. Ozawa, K.; Dixon, N. E.; Otting, G. Cell-free synthesis of  $^{15}\text{N}$ -labeled proteins for NMR studies. *IUBMB life* **2005**, 57, 615-22.
17. Ozawa, K.; Jergic, S.; Crowther, J. A.; Thompson, P. R.; Wijffels, G.; Otting, G.; Dixon, N. A. Cell-free protein synthesis in an autoinduction system for NMR studies of protein-protein interactions. *J Biomol NMR* **2005**, 32, 235-41.
18. Tyler, R. C.; Aceti, D. J.; Bingman, C. A.; Cornilescu, C. C.; Fox, B. G.; Frederick, R. O.; Jeon, W. B.; Lee, M. S.; Newman, C. S.; Peterson, F. C.; Phillips, G. N., Jr.; Shahan, M. N.; Singh, S.; Song, J.; Sreenath, H. K.; Tyler, E. M.; Ulrich, E. L.; Vinarov, D. A.; Vojtik, F. C.; Volkman, B. F.; Wrobel, R. L.; Zhao, Q.; Markley, J. L. Comparison of cell-based and cell-free protocols for producing target proteins from the *Arabidopsis thaliana* genome for structural studies. *Proteins* **2005**, 59, 633-43.
19. Klammt, C.; Schwarz, D.; Lohr, F.; Schneider, B.; Dotsch, V.; Bernhard, F. Cell-free expression as an emerging technique for the large scale production of integral membrane protein. *The FEBS journal* **2006**, 273, 4141-53.
20. Koglin, A.; Klammt, C.; Trbovic, N.; Schwarz, D.; Schneider, B.; Schafer, B.; Lohr, F.; Bernhard, F.; Dotsch, V. Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins. *Magn Reson Chem* **2006**, 44 Spec No, S17-23.
21. Ozawa, K.; Wu, P. S.; Dixon, N. E.; Otting, G. N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes. *The FEBS journal* **2006**, 273, 4154-9.
22. Staunton, D.; Schlinkert, R.; Zanetti, G.; Colebrook, S. A.; Campbell, I. D. Cell-free expression and selective isotope labelling in protein NMR. *Magn Reson Chem* **2006**, 44 Spec No, S2-9.
23. Vinarov, D. A.; Loushin Newman, C. L.; Markley, J. L. Wheat germ cell-free platform for eukaryotic protein production. *The FEBS journal* **2006**, 273, 4160-9.
24. Vinarov, D. A.; Newman, C. L.; Tyler, E. M.; Markley, J. L.; Shahan, M. N. Wheat germ cell-free expression system for protein production. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **2006**, Chapter 5, Unit 5 18.
25. Matsuda, T.; Koshiba, S.; Tochio, N.; Seki, E.; Iwasaki, N.; Yabuki, T.; Inoue, M.; Yokoyama, S.; Kigawa, T. Improving cell-free protein synthesis for stable-isotope labeling. *J Biomol NMR* **2007**, 37, 225-9.
26. Sattler, M.; Maurer, M.; Schleucher, J.; Griesinger, C. A simultaneous  $^{15}\text{N}$ ,  $^1\text{H}$ - and

<sup>13</sup>C,<sup>1</sup>H-HSQC with sensitivity enhancement and a heteronuclear gradient echo. *J Biomol NMR* **1995**, 5, 97-102.

27. Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in <sup>13</sup>C-Labeled Proteins. *J Am Chem Soc* **1998**, 120, 6394–6400.

28. Vallurupalli, P.; Hansen, D.; Kay, L. Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proc Natl Acad Sci U S A* **2008**, 105, 11766-71.

29. Geen, H.; Freeman, R. Band-selective radiofrequency pulses. *J Magn Reson* **1991**, 93, 93-141.

30. Böhlen, J.; Bodenhausen, G. Experimental Aspects of Chirp NMR Spectroscopy. *J Magn Reson A* **1993**, 102, 293-301.

31. Hwang, T.; van Zijl, P.; Garwood, M. Broadband adiabatic refocusing without phase distortion. *J Magn Reson* **1997**, 124, 250-4.

32. Kay, L.; Ikura, M.; Tschudin, R.; Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* **1990**, 89, 496-514.

33. Fu, R.; Bodenhausen, G. Broadband decoupling in NMR with frequency-modulated 'chirp' pulses. *Chem Phys Lett* **1995**, 245, 415-20.

34. Fu, R.; Bodenhausen, G. Ultra-Broadband Decoupling. *J Magn Reson A* **1995**, 117, 324-5.

35. Shaka, A.; Barker, P.; Freeman, R. Computer-optimized decoupling scheme for wideband applications and low-level operation. *J Magn Reson* **1985**, 64, 547-52.

36. Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J Magn Reson* **1989**, 85, 393-9.

37. Chang, Y.; Song, A.; Gao, Y.; Shi, Y.; Lin, X.; Cao, X.; Lin, D.; Hu, H. Solution structure of the ubiquitin-associated domain of human BMSC-UbP and its complex with ubiquitin. *Protein Sci* **2006**, 15, 1248-59.

38. Gutiérrez, P.; Kozlov, G.; Gabrielli, L.; Elias, D.; Osborne, M.; Gallouzi, I.; Gehring, K. Solution structure of YaeO, a Rho-specific inhibitor of transcription termination. *J Biol Chem* **2007**, 282, 23348-53.

39. Liu, C.; Miller, H.; Kohr, W.; Silber, J. Purification of a ubiquitin protein peptidase from yeast with efficient in vitro assays. *J Biol Chem* **1989**, 264, 20331-8.

40. Rajesh, S.; Sakamoto, T.; Iwamoto-Sugai, M.; Shibata, T.; Kohno, T.; Ito, Y. Ubiquitin binding interface mapping on yeast ubiquitin hydrolase by NMR chemical shift perturbation. *Biochemistry* **1999**, 38, 9242-53.

41. Sakamoto, T.; Tanaka, T.; Ito, Y.; Rajesh, S.; Iwamoto-Sugai, M.; Koder, Y.;

- Tsuchida, N.; Shibata, T.; Kohno, T. An NMR analysis of ubiquitin recognition by yeast ubiquitin hydrolase: evidence for novel substrate recognition by a cysteine protease. *Biochemistry* **1999**, 38, 11634-42.
42. Johnston, S.; Riddle, S.; Cohen, R.; Hill, C. Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J* **1999**, 18, 3877-87.
43. Igarashi, S.; Osawa, M.; Takeuchi, K.; Ozawa, S.; Shimada, I. Amino acid selective cross-saturation method for identification of proximal residue pairs in a protein-protein complex. *J Am Chem Soc* **2008**, 130, 12168-76.
44. Moriya, J.; Sakakura, M.; Tokunaga, Y.; Prosser, R.; Shimada, I. An NMR method for the determination of protein binding interfaces using TEMPOL-induced chemical shift perturbations. *Biochim Biophys Acta* **2009**, 1790, 1368-76.
45. Brunger, A. Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2007**, 2, 2728-33.
46. Powell, A. J.; Read, J. A.; Banfield, M. J.; Gunn-Moore, F.; Yan, S. D.; Lustbader, J.; Stern, A. R.; Stern, D. M.; Brady, R. L. Recognition of structurally diverse substrates by type II 3-hydroxyacyl-CoA dehydrogenase (HADH II)/amyloid-beta binding alcohol dehydrogenase (ABAD). *J Mol Biol* **2000**, 303, 311-27.
47. Tavares, G. A.; Beguin, P.; Alzari, P. M. The crystal structure of a type I cohesin domain at 1.7 Å resolution. *J Mol Biol* **1997**, 273, 701-13.
48. Zhou, Y.; Morais-Cabral, J. H.; Kaufman, A.; MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K<sup>+</sup> channel-Fab complex at 2.0 Å resolution. *Nature* **2001**, 414, 43-8.
49. Grabarse, W.; Mahlert, F.; Shima, S.; Thauer, R. K.; Ermler, U. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation. *J Mol Biol* **2000**, 303, 329-44.
50. Dominguez, C.; Boelens, R.; Bonvin, A. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **2003**, 125, 1731-7.
51. de Vries, S.; van Dijk, A.; Krzeminski, M.; van Dijk, M.; Thureau, A.; Hsu, V.; Wassenaar, T.; Bonvin, A. HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* **2007**, 69, 726-33.
52. Méndez, R.; Leplae, R.; De Maria, L.; Wodak, S. Assessment of blind predictions of protein-protein interactions: current status of docking methods. *Proteins* **2003**, 52, 51-67.
53. Méndez, R.; Leplae, R.; Lensink, M.; Wodak, S. Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures. *Proteins* **2005**, 60, 150-69.
54. Lensink, M.; Méndez, R.; Wodak, S. Docking and scoring protein complexes: CAPRI



3rd Edition. *Proteins* **2007**, 69, 704-18.

55. Koradi, R.; Billeter, M.; Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **1996**, 14, 51-5, 29-32.
56. Zartler, E. R.; Shapiro, M. J. Fragonomics: fragment-based drug discovery. *Current opinion in chemical biology* **2005**, 9, 366-70.
57. Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nature reviews. Drug discovery* **2007**, 6, 211-9.
58. Bower, J. F.; Pannifer, A. Using fragment-based technologies to target protein-protein interactions. *Current pharmaceutical design* **2012**, 18, 4685-96.
59. Kumar, A.; Voet, A.; Zhang, K. Y. Fragment based drug design: from experimental to computational approaches. *Current medicinal chemistry* **2012**, 19, 5128-47.
60. Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **2012**, 51, 4990-5003.
61. Winter, A.; Higuero, A. P.; Marsh, M.; Sigurdardottir, A.; Pitt, W. R.; Blundell, T. L. Biophysical and computational fragment-based approaches to targeting protein-protein interactions: applications in structure-guided drug discovery. *Quarterly reviews of biophysics* **2012**, 45, 383-426.
62. Caliendo, R.; Belviso, D. B.; Aresta, B. M.; de Candia, M.; Altomare, C. D. Protein crystallography and fragment-based drug design. *Future medicinal chemistry* **2013**, 5, 1121-40.
63. Sancineto, L.; Massari, S.; Iraci, N.; Tabarrini, O. From small to powerful: the fragments universe and its "chem-appeal". *Current medicinal chemistry* **2013**, 20, 1355-81.
64. Soga, S.; Shirai, H.; Kobori, M.; Hirayama, N. Use of amino acid composition to predict ligand-binding sites. *J Chem Inf Model* **2007**, 47, 400-6.
65. Ghersi, D.; Sanchez, R. EasyMIFS and SiteHound: a toolkit for the identification of ligand-binding sites in protein structures. *Bioinformatics* **2009**, 25, 3185-6.
66. Hernandez, M.; Ghersi, D.; Sanchez, R. SITEHOUND-web: a server for ligand binding site identification in protein structures. *Nucleic Acids Res* **2009**, 37, W413-6.
67. Ghersi, D.; Sanchez, R. Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites. *Proteins* **2009**, 74, 417-24.
68. Ishihata, K. Smallest Enclosing Sphere(<Series>Program Promenade). *IPSJ Magazine* **2002**, 43, 1009-1015.
69. Katz, B.; Mackman, R.; Luong, C.; Radika, K.; Martelli, A.; Sprengeler, P.; Wang, J.; Chan, H.; Wong, L. Structural basis for selectivity of a small molecule, S1-binding, submicromolar inhibitor of urokinase-type plasminogen activator. *Chem Biol* **2000**, 7, 299-312.

70. Senger, S.; Convery, M. A.; Chan, C.; Watson, N. S. Arylsulfonamides: a study of the relationship between activity and conformational preferences for a series of factor Xa inhibitors. *Bioorg Med Chem Lett* **2006**, 16, 5731-5.
71. Corradi, H. R.; Schwager, S. L.; Nchinda, A. T.; Sturrock, E. D.; Acharya, K. R. Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a structural basis for domain-specific inhibitor design. *J Mol Biol* **2006**, 357, 964-74.
72. Solis, D.; Mate, M. J.; Lohr, M.; Ribeiro, J. P.; Lopez-Merino, L.; Andre, S.; Buzamet, E.; Canada, F. J.; Kaltner, H.; Lensch, M.; Ruiz, F. M.; Haroske, G.; Wollina, U.; Kloor, M.; Kopitz, J.; Saiz, J. L.; Menendez, M.; Jimenez-Barbero, J.; Romero, A.; Gabius, H. J. N-domain of human adhesion/growth-regulatory galectin-9: preference for distinct conformers and non-sialylated N-glycans and detection of ligand-induced structural changes in crystal and solution. *The international journal of biochemistry & cell biology* **2010**, 42, 1019-29.
73. Qu, A.; Leahy, D. J. The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. *Structure* **1996**, 4, 931-42.
74. Dodd, D. S.; Sheriff, S.; Chang, C. J.; Stetsko, D. K.; Phillips, L. M.; Zhang, Y.; Launay, M.; Potin, D.; Vaccaro, W.; Poss, M. A.; McKinnon, M.; Barrish, J. C.; Suchard, S. J.; Murali Dhar, T. G. Design of LFA-1 antagonists based on a 2,3-dihydro-1H-pyrrolizin-5(7aH)-one scaffold. *Bioorg Med Chem Lett* **2007**, 17, 1908-11.
75. Rouse, J.; Cohen, P.; Trigon, S.; Morange, M.; Alonso-Llamazares, A.; Zamanillo, D.; Hunt, T.; Nebreda, A. R. A novel kinase cascade triggered by stress and heat shock that stimulates MAPKAP kinase-2 and phosphorylation of the small heat shock proteins. *Cell* **1994**, 78, 1027-37.
76. Han, J.; Lee, J. D.; Bibbs, L.; Ulevitch, R. J. A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. *Science* **1994**, 265, 808-11.
77. Lee, J. C.; Laydon, J. T.; McDonnell, P. C.; Gallagher, T. F.; Kumar, S.; Green, D.; McNulty, D.; Blumenthal, M. J.; Heys, J. R.; Landvatter, S. W.; et al. A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature* **1994**, 372, 739-46.
78. Freshney, N. W.; Rawlinson, L.; Guesdon, F.; Jones, E.; Cowley, S.; Hsuan, J.; Saklatvala, J. Interleukin-1 activates a novel protein kinase cascade that results in the phosphorylation of Hsp27. *Cell* **1994**, 78, 1039-49.
79. Fischer, S.; Koeberle, S. C.; Laufer, S. A. p38alpha mitogen-activated protein kinase inhibitors, a patent review (2005-2011). *Expert opinion on therapeutic patents* **2011**, 21, 1843-66.
80. Kong, T. T.; Zhang, C. M.; Liu, Z. P. Recent developments of p38alpha MAP kinase inhibitors as antiinflammatory agents based on the imidazole scaffolds. *Current medicinal*

*chemistry* **2013**, 20, 1997-2016.

81. Gill, A.; Frederickson, M.; Cleasby, A.; Woodhead, S.; Carr, M.; Woodhead, A.; Walker, M.; Congreve, M.; Devine, L.; Tisi, D.; O'Reilly, M.; Seavers, L.; Davis, D.; Curry, J.; Anthony, R.; Padova, A.; Murray, C.; Carr, R.; Jhoti, H. Identification of novel p38alpha MAP kinase inhibitors using fragment-based lead generation. *J Med Chem* **2005**, 48, 414-26.
82. Hartshorn, M.; Murray, C.; Cleasby, A.; Frederickson, M.; Tickle, I.; Jhoti, H. Fragment-based lead discovery using X-ray crystallography. *J Med Chem* **2005**, 48, 403-13.
83. Patel, S.; Cameron, P.; Frantz-Wattley, B.; O'Neill, E.; Becker, J.; Scapin, G. Lattice stabilization and enhanced diffraction in human p38 alpha crystals by protein engineering. *Biochim Biophys Acta* **2004**, 1696, 67-73.
84. Reddy, M.; Thimmalapura, P.; Lanting, L.; Nadler, J.; Fatima, S.; Natarajan, R. The oxidized lipid and lipoxygenase product 12(S)-hydroxyeicosatetraenoic acid induces hypertrophy and fibronectin transcription in vascular smooth muscle cells via p38 MAPK and cAMP response element-binding protein activation. Mediation of angiotensin II effects. *J Biol Chem* **2002**, 277, 9920-8.
85. Collins, Q.; Xiong, Y.; Lupo, E. J.; Liu, H.; Cao, W. p38 Mitogen-activated protein kinase mediates free fatty acid-induced gluconeogenesis in hepatocytes. *J Biol Chem* **2006**, 281, 24336-44.
86. Nony, P.; Kennett, S.; Glasgow, W.; Olden, K.; Roberts, J. 15S-Lipoxygenase-2 mediates arachidonic acid-stimulated adhesion of human breast carcinoma cells through the activation of TAK1, MKK6, and p38 MAPK. *J Biol Chem* **2005**, 280, 31413-9.
87. Kong, J.; Klassen, S.; Rabkin, S. Ceramide activates a mitochondrial p38 mitogen-activated protein kinase: a potential mechanism for loss of mitochondrial transmembrane potential and apoptosis. *Mol Cell Biochem* **2005**, 278, 39-51.
88. Diskin, R.; Engelberg, D.; Livnah, O. A novel lipid binding site formed by the MAP kinase insert in p38 alpha. *J Mol Biol* **2008**, 375, 70-9.
89. Tzarum, N.; Eisenberg-Domovich, Y.; Gills, J. J.; Dennis, P. A.; Livnah, O. Lipid molecules induce p38alpha activation via a novel molecular switch. *J Mol Biol* **2012**, 424, 339-53.
90. Gills, J. J.; Castillo, S. S.; Zhang, C.; Petukhov, P. A.; Memmott, R. M.; Hollingshead, M.; Warfel, N.; Han, J.; Kozikowski, A. P.; Dennis, P. A. Phosphatidylinositol ether lipid analogues that inhibit AKT also independently activate the stress kinase, p38alpha, through MKK3/6-independent and -dependent mechanisms. *J Biol Chem* **2007**, 282, 27020-9.
91. Francis, D. M.; Kumar, G. S.; Koveal, D.; Tortajada, A.; Page, R.; Peti, W. The differential regulation of p38alpha by the neuronal kinase interaction motif protein tyrosine

- phosphatases, a detailed molecular study. *Structure* **2013**, 21, 1612-23.
92. Vogtherr, M.; Saxena, K.; Grimme, S.; Betz, M.; Schieborr, U.; Pescatore, B.; Langer, T.; Schwalbe, H. NMR backbone assignment of the mitogen-activated protein (MAP) kinase p38. *J Biomol NMR* **2005**, 32, 175.
  93. Laskowski, R. A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* **1995**, 13, 323-30, 307-8.
  94. Laskowski, R. A.; Luscombe, N. M.; Swindells, M. B.; Thornton, J. M. Protein clefts in molecular recognition and function. *Protein Sci* **1996**, 5, 2438-52.
  95. Peters, K. P.; Fauck, J.; Frommel, C. The automatic search for ligand binding sites in proteins of known three-dimensional structure using only geometric criteria. *J Mol Biol* **1996**, 256, 201-13.
  96. Hendlich, M.; Rippmann, F.; Barnickel, G. LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *J Mol Graph Model* **1997**, 15, 359-63, 389.
  97. Liang, J.; Edelsbrunner, H.; Woodward, C. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci* **1998**, 7, 1884-97.
  98. Brady, G. P., Jr.; Stouten, P. F. Fast prediction and visualization of protein binding pockets with PASS. *J Comput Aided Mol Des* **2000**, 14, 383-401.
  99. Lichtarge, O.; Sowa, M. E. Evolutionary predictions of binding surfaces and interactions. *Curr Opin Struct Biol* **2002**, 12, 21-7.
  100. Joughin, B. A.; Tidor, B.; Yaffe, M. B. A computational method for the analysis and prediction of protein:phosphopeptide-binding sites. *Protein Sci* **2005**, 14, 131-9.
  101. Laurie, A. T.; Jackson, R. M. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics* **2005**, 21, 1908-16.
  102. Coleman, R. G.; Salzberg, A. C.; Cheng, A. C. Structure-based identification of small molecule binding sites using a free energy model. *J Chem Inf Model* **2006**, 46, 2631-7.
  103. Glaser, F.; Morris, R. J.; Najmanovich, R. J.; Laskowski, R. A.; Thornton, J. M. A method for localizing ligand binding pockets in protein structures. *Proteins* **2006**, 62, 479-88.
  104. Huang, B.; Schroeder, M. LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Struct Biol* **2006**, 6, 19.
  105. Chen, B. Y.; Bryant, D. H.; Fofanov, V. Y.; Kristensen, D. M.; Cruess, A. E.; Kimmel, M.; Lichtarge, O.; Kavraki, L. E. Cavity scaling: automated refinement of cavity-aware motifs in protein function prediction. *J Bioinform Comput Biol* **2007**, 5, 353-82.
  106. Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S. Structure-based maximal affinity model predicts

small-molecule druggability. *Nat Biotechnol* **2007**, 25, 71-5.

107. Kahraman, A.; Morris, R. J.; Laskowski, R. A.; Thornton, J. M. Shape variation in protein binding pockets and their ligands. *J Mol Biol* **2007**, 368, 283-301.

108. Kawabata, T.; Go, N. Detection of pockets on protein surfaces using small and large probe spheres to find putative ligand binding sites. *Proteins* **2007**, 68, 516-29.

109. Pettit, F. K.; Bare, E.; Tsai, A.; Bowie, J. U. HotPatch: a statistical approach to finding biologically relevant features on protein surfaces. *J Mol Biol* **2007**, 369, 863-79.

110. Weisel, M.; Proschak, E.; Schneider, G. PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem Cent J* **2007**, 1, 7.

111. Xie, L.; Bourne, P. E. A robust and efficient algorithm for the shape description of protein structures and its application in predicting ligand binding sites. *BMC Bioinformatics* **2007**, 8 Suppl 4, S9.

112. Zhong, S.; MacKerell, A. D., Jr. Binding response: a descriptor for selecting ligand binding site on protein surfaces. *J Chem Inf Model* **2007**, 47, 2303-15.

113. Brylinski, M.; Skolnick, J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc Natl Acad Sci USA* **2008**, 105, 129-34.

114. Harris, R.; Olson, A. J.; Goodsell, D. S. Automated prediction of ligand-binding sites in proteins. *Proteins* **2008**, 70, 1506-17.

115. Kalidas, Y.; Chandra, N. PocketDepth: a new depth based algorithm for identification of ligand binding sites in proteins. *J Struct Biol* **2008**, 161, 31-42.

116. Kim, D.; Cho, C. H.; Cho, Y.; Ryu, J.; Bhak, J.; Kim, D. S. Pocket extraction on proteins via the Voronoi diagram of spheres. *J Mol Graph Model* **2008**, 26, 1104-12.

117. Ming, D.; Cohn, J. D.; Wall, M. E. Fast dynamics perturbation analysis for prediction of protein functional sites. *BMC Struct Biol* **2008**, 8, 5.

118. Capra, J. A.; Laskowski, R. A.; Thornton, J. M.; Singh, M.; Funkhouser, T. A. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. *PLoS Comput Biol* **2009**, 5, e1000585.

119. Huang, B. MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS* **2009**, 13, 325-30.

120. Coleman, R. G.; Sharp, K. A. Protein pockets: inventory, shape, and comparison. *J Chem Inf Model* **2010**, 50, 589-603.

121. Kawabata, T. Detection of multiscale pockets on protein surfaces using mathematical morphology. *Proteins* **2010**, 78, 1195-211.

122. Fukunishi, Y.; Nakamura, H. Prediction of ligand-binding sites of proteins by molecular docking calculation for a random ligand library. *Protein Sci* **2011**, 20, 95-106.

123. McCoy, M. A.; Wyss, D. F. Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J Am Chem Soc* **2002**, 124, 11758-63.
124. Hajduk, P. J.; Mack, J. C.; Olejniczak, E. T.; Park, C.; Dandliker, P. J.; Beutel, B. A. SOS-NMR: a saturation transfer NMR-based method for determining the structures of protein-ligand complexes. *J Am Chem Soc* **2004**, 126, 2390-8.
125. Orts, J.; Tuma, J.; Reese, M.; Grimm, S. K.; Monecke, P.; Bartoschek, S.; Schiffer, A.; Wendt, K. U.; Griesinger, C.; Carlomagno, T. Crystallography-independent determination of ligand binding modes. *Angew Chem Int Ed Engl* **2008**, 47, 7736-40.
126. Mizukoshi, Y.; Abe, A.; Takizawa, T.; Hanzawa, H.; Fukunishi, Y.; Shimada, I.; Takahashi, H. An accurate pharmacophore mapping method by NMR spectroscopy. *Angew Chem Int Ed Engl* **2012**, 51, 1362-5.
127. Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Mei Ono, A.; Güntert, P. Optimal isotope labelling for NMR protein structure determinations. *Nature* **2006**, 440, 52-7.
128. LeMaster, D. Deuteration in protein proton magnetic resonance. *Methods Enzymol* **1989**, 177, 23-43.
129. Fiaux, J.; Bertelsen, E.; Horwich, A.; Wüthrich, K. NMR analysis of a 900K GroEL GroES complex. *Nature* **2002**, 418, 207-11.
130. Miao, X.; Mukhopadhyay, R.; Valafar, H. Estimation of relative order tensors, and reconstruction of vectors in space using unassigned RDC data and its application. *J Magn Reson* **2008**, 194, 202-11.
131. Mukhopadhyay, R.; Miao, X.; Shealy, P.; Valafar, H. Efficient and accurate estimation of relative order tensors from lambda-maps. *J Magn Reson* **2009**, 198, 236-47.
132. Arumugam, S.; Hemme, C.; Yoshida, N.; Suzuki, K.; Nagase, H.; Berjanskii, M.; Wu, B.; Van Doren, S. TIMP-1 contact sites and perturbations of stromelysin 1 mapped by NMR and a paramagnetic surface probe. *Biochemistry* **1998**, 37, 9650-7.
133. DellaVecchia, M.; Merritt, W.; Peng, Y.; Kirby, T.; DeRose, E.; Mueller, G.; Van Houten, B.; London, R. NMR analysis of [methyl-<sup>13</sup>C]methionine UvrB from *Bacillus caldotenax* reveals UvrB-domain 4 heterodimer formation in solution. *J Mol Biol* **2007**, 373, 282-95.
134. Deschamps, M.; Pilka, E.; Potts, J.; Campbell, I.; Boyd, J. Probing protein-peptide binding surfaces using charged stable free radicals and transverse paramagnetic relaxation enhancement (PRE). *J Biomol NMR* **2005**, 31, 155-60.
135. Liepinsh, E.; Baryshev, M.; Sharipo, A.; Ingelman-Sundberg, M.; Otting, G.; Mkrtchian, S. Thioredoxin fold as homodimerization module in the putative chaperone

- ERp29: NMR structures of the domains and experimental model of the 51 kDa dimer. *Structure* **2001**, 9, 457-71.
136. Molinari, H.; Esposito, G.; Ragona, L.; Pegna, M.; Niccolai, N.; Brunne, R.; Lesk, A.; Zetta, L. Probing protein structure by solvent perturbation of NMR spectra: the surface accessibility of bovine pancreatic trypsin inhibitor. *Biophys J* **1997**, 73, 382-96.
  137. Reese, M.; Dakoji, S.; Bredt, D.; Dötsch, V. The guanylate kinase domain of the MAGUK PSD-95 binds dynamically to a conserved motif in MAP1a. *Nat Struct Mol Biol* **2007**, 14, 155-63.
  138. de Groot, B. L.; van Aalten, D. M.; Scheek, R. M.; Amadei, A.; Vriend, G.; Berendsen, H. J. Prediction of protein conformational freedom from distance constraints. *Proteins* **1997**, 29, 240-51.
  139. Seeliger, D.; Haas, J.; de Groot, B. L. Geometry-based sampling of conformational transitions in proteins. *Structure* **2007**, 15, 1482-92.
  140. Eyrisch, S.; Helms, V. Transient pockets on protein surfaces involved in protein-protein interaction. *J Med Chem* **2007**, 50, 3457-64.
  141. Eyrisch, S.; Helms, V. What induces pocket openings on protein surface patches involved in protein-protein interactions? *J Comput Aided Mol Des* **2009**, 23, 73-86.
  142. Seeliger, D.; de Groot, B. L. Conformational transitions upon ligand binding: holo-structure prediction from apo conformations. *PLoS Comput Biol* **2010**, 6, e1000634.
  143. Sali, A.; Blundell, T. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **1993**, 234, 779-815.
  144. Fiser, A.; Do, R.; Sali, A. Modeling of loops in protein structures. *Protein Sci* **2000**, 9, 1753-73.
  145. Gil Kim, J.; Fukunishi, Y.; Kidera, A.; Nakamura, H. Determination of multicanonical weight based on a stochastic model of sampling dynamics. *Phys Rev E Stat Nonlin Soft Matter Phys* **2003**, 68, 021110.
  146. Kim, J.; Fukunishi, Y.; Nakamura, H. Multicanonical molecular dynamics algorithm employing an adaptive force-biased iteration scheme. *Phys Rev E Stat Nonlin Soft Matter Phys* **2004**, 70, 057103.
  147. Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, 79, 926-35.
  148. Vijay-Kumar, S.; Bugg, C.; Cook, W. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* **1987**, 194, 531-44.
  149. Guex, N.; Peitsch, M. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, 18, 2714-23.

## 謝辞

本研究に携わらせて頂く機会と素晴らしい研究環境を与えて下さり、研究の方向性を含め様々な御指導下さいました、東京大学大学院薬学系研究科教授 嶋田一夫博士に深く感謝の意を表します。

本研究における NMR 測定をはじめとした各種実験に関する御指導の他、特に新規測定パルスシーケンスの開発において多大なる御助力を下さいました、現横浜市立大学大学院生命医科学研究科教授 高橋栄夫博士、並びに、日々の御指導の他、各種実験やシミュレーションの結果および手法改善の展望に関する適切な御助言を下さいました、独立行政法人産業技術総合研究所現創薬分子プロファイリング研究センター主任研究員 竹内恒博士に、心より感謝致します。

タンパク質—タンパク質間相互作用部位を同定するプログラムの開発に関し本研究へコラレーション下さいました、現テキサス大学サウスウエスタン医学センター准教授 Michael L. Reese 博士、並びに、化学シフト値の統計データの母集団に関する適切な御助言を下さいました、フランクフルト大学教授 Volker Dötsch 博士に深く感謝致します。また、ドッキングシミュレーションに用いるタンパク質立体構造の準備に関し御助力下さいました、独立行政法人産業技術総合研究所現創薬分子プロファイリング研究センター研究チーム長 福西快文博士、同センター 小野克輝博士、現静岡県立大学薬学部准教授 野口修治博士、および、株式会社日立ソリューションズ金森英司博士に心より感謝申し上げます。

本研究の要所において適切な御助言を下さいました、本務先である味の素株式会社の鈴木榮一郎博士(上席フェロー)、榛葉信久博士、石川弘紀博士に、深く感謝申し上げます。特に鈴木上席フェローからは、本研究成果をまとめるにあたって数多くの貴重な御意見と御助力を頂戴しました。

産学官連携である本研究の推進においては、嶋田教授、高橋教授、竹内主任研究員、鈴木上席フェローの他、一般社団法人バイオ産業情報化コンソーシアム 榊原信雄様、独立行政法人産業技術総合研究所 須貝潤一研究参与、並びに、味の素株式会社の本研究主管研究グループ長の宮野博、石川弘紀、水越利巳各博士をはじめとした各所の皆様に御助力頂きました。この場をお借りして厚く御礼申し上げます。また、本研究に不可欠な YUH と Y-Ub の遺伝子を御提供下さいました北里大学医学部 河野俊之先生、p38 $\alpha$  の遺伝子を御提供下さいました第一三共株式会社 半沢宏之博士に厚く御礼申し上げます。

最後に、日々の研究を進めるにあたり数多くの御助言、御助力を下さいました、独立行政法人産業技術総合研究所 水越弓子博士、株式会社東レリサーチセンター 鬼塚拓男博士をはじめ、私が本研究を行わせて頂いた独立行政法人産業技術総合研究所の研究室の皆様に、御礼申し上げます。

なお、本研究は独立行政法人新エネルギー・産業技術総合開発機構(NEDO)および経済産業省(METI)の支援、ならびに文部科学省(MEXT)および日本学術振興会(JSPS)からの科学研究費補助金(No. 24370048、24102524、25121743)の一部支援の下、行われました。



## 付録

### 1. 最小包含球を求めるスクリプト

PDB フォーマットの構造を入力ファイルとして、下記のように実行する。なお、出力結果の radius が、最小包含球の半径にあたる。

```
script.py input.pdb

-----

#!/usr/bin/env python
import sys
infile=sys.argv[1]

xa=[]          #coordinate of x
ya=[]          #coordinate of y
za=[]          #coordinate of z

num_atom=0

x=0.0          #start point
y=0.0
z=0.0
mobility=0.5  #initial value of mobility weight

#get coordinates
fin=open(infile,'r')
for line in fin:
    head=line[0:6]
    chain=line[21:22]
    if ((head.find('ATOM') != -1) or (head.find('HETATM') != -1)):
        xa.append(float(line[30:38]))
        ya.append(float(line[38:46]))
        za.append(float(line[46:54]))

    num_atom += 1

fin.close
```

```

#move temporary center
while mobility > 10.0e-8:
    for k in range(100):
        max_dist=0.0
        for i in range(num_atom):
            d = (xa[i]-x)**2 + (ya[i]-y)**2 + (za[i]-z)**2
            if (d > max_dist):
                max_dist = d
                max = i

        x += (xa[max]-x)*mobility
        y += (ya[max]-y)*mobility
        z += (za[max]-z)*mobility

    mobility = mobility/2.0

#output
print "number of atoms: %d" %(num_atom)
print "radius: %f" %(max_dist**0.5)
print "center: x=%f y=%f z=%f" %(x,y,z)
-----

```

## 2. 低分子相互作用部位同定用の解析スクリプト

入力(input.pdb)および出力(output.pdb)ファイルのフォーマットは PDB である。入力ファイルにおけるタンパク質の chain ID は A、生成したグリッドの chain ID は B とする。スクリプトの実行は、例えば下記のようにして行う。

```
script.py input.pdb output.pdb radius probe1 min max probe2 min max ...
```

アウトプットデータには、C3～C9 の名称がつけられた炭素原子がグリッドの座標におかれており、C9 が最も実験結果に一致するグリッド、C3 が最も実験結果に一致しないグリッドである。上記”radius”には、仮想球の半径  $r_v$  を入力する。”Probe”は、観測プローブとするアミノ酸残基種を 3

文字表記で、“min”と“max”は、それぞれ化学シフト変化が大きいシグナルの個数の最小値と最大値を入力する。例えば、p38 $\alpha$  と阻害剤の実験結果から相互作用部位を解析する際には、下記のように入力する。なお、このスクリプトは、観測プローブとするアミノ酸残基種の数に 6 種類の場合にのみ対応し、それ以外の場合は修正が必要である。

```
script.py 1r39.pdb output.pdb 8.497 ALA 3 3 MET 1 1 TYR 0 0 HIS 0 1 TRP 0 0 PHE 0 1

-----

#!/usr/bin/env python
import sys

infile=sys.argv[1]
outfile=sys.argv[2]
th=sys.argv[3]          #threshold (angstrom)

probe0=sys.argv[4]
expdata0s=int(sys.argv[5])
expdata0e=int(sys.argv[6])
probe1=sys.argv[7]
expdata1s=int(sys.argv[8])
expdata1e=int(sys.argv[9])
probe2=sys.argv[10]
expdata2s=int(sys.argv[11])
expdata2e=int(sys.argv[12])
probe3=sys.argv[13]
expdata3s=int(sys.argv[14])
expdata3e=int(sys.argv[15])
probe4=sys.argv[16]
expdata4s=int(sys.argv[17])
expdata4e=int(sys.argv[18])
probe5=sys.argv[19]
expdata5s=int(sys.argv[20])
expdata5e=int(sys.argv[21])

th=(float(th))*2
```

```

lineb=[]
heada=[]
headb=[]
restypea=[] #residue type of A
resnoa=[] #residue No. of A
xa=[] #coordinate of x of A
ya=[] #coordinate of y of A
za=[] #coordinate of z of A
xb=[] #coordinate of x of B
yb=[] #coordinate of y of B
zb=[] #coordinate of z of B


#get each parameter from input.pdb

fin=open(infile,'r')
for line in fin:
    head=line[0:6]
    chain=line[21:22]
    if ((head.find('ATOM') != -1) and (chain.find('A') != -1)):
        restypea.append(line[17:21])
        resnoa.append(line[22:30])
        xa.append(float(line[30:38]))
        ya.append(float(line[38:46]))
        za.append(float(line[46:54]))

    elif (((head.find('ATOM') != -1) or (head.find('HETATM') != -1)) \
        and (chain.find('B') != -1)):
        lineb.append(line[0:80])
        headb.append(line[0:6])
        xb.append(float(line[30:38]))
        yb.append(float(line[38:46]))
        zb.append(float(line[46:54]))

fin.close

```

```

#estimation and pick residues

tempstock=["" for i in range(len(xb))]
for i in range(len(xb)):
    for j in range(len(xa)):
        d = (xb[i]-xa[j])**2 + (yb[i]-ya[j])**2 + (zb[i]-za[j])**2
        if (d<=th):
            if(tempstock[i].find(resnoa[j]) == -1):
                tempstock[i] = tempstock[i] + restypea[j] \
                    + resnoa[j]

#count and output

fout=open(outfile,'w')
for i in range(len(xb)):
    count=0
    for l in range(expdata0s, expdata0e+1, 1):
        for m in range(expdata1s, expdata1e+1, 1):
            for n in range(expdata2s, expdata2e+1, 1):
                for o in range(expdata3s, expdata3e+1, 1):
                    for p in \
                        range(expdata4s, expdata4e+1, 1):
                        for q in \
                            range(expdata5s, expdata5e+1, 1):
                            count = count \
                                + abs(l - tempstock[i].count(probe0)) \
                                + abs(m - tempstock[i].count(probe1)) \
                                + abs(n - tempstock[i].count(probe2)) \
                                + abs(o - tempstock[i].count(probe3)) \
                                + abs(p - tempstock[i].count(probe4)) \
                                + abs(q - tempstock[i].count(probe5))

    if 0 <= count <= 4:
        lineb[i] = lineb[i].replace("C3","C9",1) + "\n"
    elif 5 <= count <= 8:
        lineb[i] = lineb[i].replace("C3","C8",1) + "\n"

```

```

elif 9 <= count <= 12:
    lineb[i] = lineb[i].replace("C3", "C7", 1) + "\n"
elif 13 <= count <= 16:
    lineb[i] = lineb[i].replace("C3", "C6", 1) + "\n"
elif 17 <= count <= 20:
    lineb[i] = lineb[i].replace("C3", "C5", 1) + "\n"
elif 21 <= count <= 24:
    lineb[i] = lineb[i].replace("C3", "C4", 1) + "\n"
else:
    lineb[i] = lineb[i].replace("C3", "C3", 1) + "\n"
fout.write(lineb[i])

fout.write("TER\nEND\n")
fout.close
-----

```