

3. 膜タンパク質の構造解析

3.1 バクテリオロドプシンの構造変化の差マップによる評価

電子顕微鏡像からは振幅情報のみならず位相情報も直接入手できる。光駆動型プロトンポンプであるバクテリオロドプシンの中間体構造を模する三残基変異体(D96G/F171C/F219L)と天然体の基底状態それぞれの二次元結晶像から投影像における差マップを計算した。この差マップは構造変化におけるヘリックスFおよびGの大きなシフトを支持するがX線で構造が解かれている中間体は必ずしもこの大きな構造変化を捉えていない。特に三次元結晶において光照射により構造変化を行わせた場合には立体障壁が問題となり正しく構造変化していないと考えられることなどを議論した(Hirai T. & Subramaniam S., *PLoS ONE*, 2009, Hirai et al. *Curr. Opin. Struct. Biol.*, 2009)。

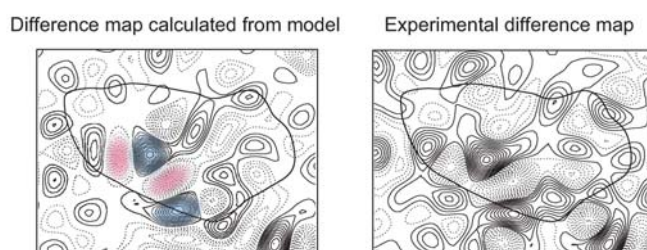


図3. バクテリオロドプシンの差マップ(投影像)

電顕(回折データを用いた分子置換法)で構造解析されたPDB座標から計算された差マップ(左)と実像のみから計算された差マップ(右)。この2つは良く一致した。

3.2 MFS型輸送体であるシュウ酸イオン輸送体の構造解析

シュウ酸イオン輸送体(OxiT)は基質の電気化学ポテンシャル駆動型輸送体の中のMFS型輸送体に属する。OxiTに関しては電子線結晶学により6.5 Åの三次元構造が報告されている(Hirai T. et al., *Nat. Struct. Biol.*, 2002、平井、*生物物理*、2007)。MFS型輸送体の輸送機構をより深く理解するために現在高分解能の構造解析に向けて結晶の改善を行っている。

3.3 ヒト赤血球膜蛋白質バンド3陰イオン輸送体の構造解析

バンド3は人の赤血球に存在する膜タンパク質で陰イオン交換輸送体として働く。具体的には HCO_3^- と Cl^- の交換を行うことにより、二酸化炭素によるpHの減少効果を増幅し、代謝が活発な組織においてヘモグロビンからの酸素の放出を促進することができる。バンド3の構造解析は以前から試みられているが10年以上前に20 Åの構造が報告されて以来(Wang D. et al., *EMBO J.*, 1994)、大きな進展はない。当研究チームに在籍した山口は長年このバンド3の二次元結晶化に取り組みらせん状結晶とチューブ状二次元結晶を得ることに成功した。らせん状結晶は単粒子解析法を応用したIHRSR法により20 Å程度で立体構造解析した(Yamaguchi T. et al., in preparation)。チューブ状二次元結晶はMRCプログラムを使用して7 Å程度で三次元構造解析を行った(Yamaguchi T. et al., in preparation)。いずれも投稿準備中である。

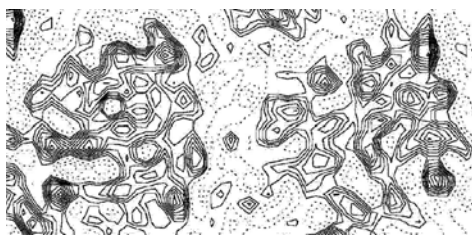


図3. バンド3投影像

バンド3の二次元結晶の電子顕微鏡像から得られた予備的な投影像。同じ結晶内でも左右異なるように若干の構造変化が識別される。

(7) 分子動力学計算による水銀の AQP1 抑制機序の解析

[慶應義塾大学大学院医学研究科]

水分子は生物が生きていく上で必須の物質である。体内にある水分子の含量は、生体組織や細胞レベルで巧妙かつ精密に調整されている。細胞は膜によって仕切られているが、膜の内部は疎水性の脂質から構成されているため、水分子は膜の内部を受動拡散により通過する。しかし、その速度は細胞内外の水分子の交換には充分ではない。例えば、赤血球や腎臓の尿細管上皮細胞の細胞膜の水分子透過性は受動拡散では説明できない。そのため、膜の中には水分子を通す「穴 (pore)」が存在し、その「穴」を介して分子の調整を行う膜タンパク質の存在が示唆されてきた。アクアポリンは水分子を選択的に透過する「穴」である。現在までの研究で、アクアポリンは全身に分布しており、ヒトを含めた哺乳類には 13 種類のアクアポリンのサブタイプが存在し、そのサブタイプ毎に細胞や組織に分布・局在していることが知られている。このことからアクアポリンは生命にとって非常に重要な働きをしていると考えられる。アクアポリンの身体における役割についてはこれまでに様々な研究が生化学的アプローチや理論的アプローチにより行われてきている。しかし、アクアポリンの分布・局在の理由や選択的水分子透過の制御機構に関しては未だ謎の部分が多い。

アクアポリンの機能を制御するような薬剤の発見は、その機能異常により起こる疾病の治療薬になると考えられる。しかしながら、これまでのところ、臨床応用可能なアクアポリンに対する阻害剤となりうる薬剤は見つかっていない。水銀などの金属は多くのアクアポリン中のシステイン残基に結合し、水分子透過を阻害することが生化学実験により知られている。しかし、その阻害機構について明確な答えはでていない。そこで、アクアポリン 1 中のシステイン残基に水銀が結合した系と結合していない系を構築して分子動力学計算を行い、両者の結果を比較することでアクアポリンの水銀による水分子透過阻害機構について解明することを研究の主な目的とした。

方法 分子動力学シミュレーションにより、アクアポリン 1 の膜中での構造ダイナミクスに関する研究を行った。分子動力学シミュレーションを行うためのモデル構造は次のように構築した。使用した結晶構造は bovine の AQP1 (PDB entry: 1J4N) である。アクアポリンの単量体の立体構造は 6 つのアルファ・ヘリックス (H1-H6) が膜を貫通している。さらに、2 つの短いアルファ・ヘリックス (HB, HE) が膜の中間付近で向き合っている。互いのアルファ・ヘリックス間は特に構造をもたないループで結ばれている。また、アクアポリン 1 は生体内では四量体として存在している。そこで、単量体構造から四量体構造を構築した。構築した四量体構造を脂質 (POPE) で構成された二重層膜に埋め込み、二重膜の上下に水分子 (TIP3P water) を発生させ、一辺がおよそ 110 Å の立方体セルを構築した (図 1)。このセルに含まれる原子数は約 10 万である。構築したモデル系について、まず、初期構造を構築するために温度 (310 K)、圧力 (1 atm) 一定 (NPT) の周期境界条件下で平衡状態の分子動力学シミュレーションを 2 ns 行った。2 ns の分子動力学シミュレーションを行い、得られた構造を初期構造とした。この初期構造から、水銀の結合した系 (Hg-AQP1) と結合していない系 (Free AQP1) を構築し、それぞれの構造について、さらに 40 ns の平衡状態の分子動力学シミュレーションをおこなった。分子動力学シミュレーション中は水素原子を含む結合長を平

平衡核間距離に保つために SHAKE 法を用いた。タイムステップは 1 fs とした。遠距離からの寄与は Particle Mesh Ewald (PME)法により計算した。計算プログラムは amber 8 を用いた。

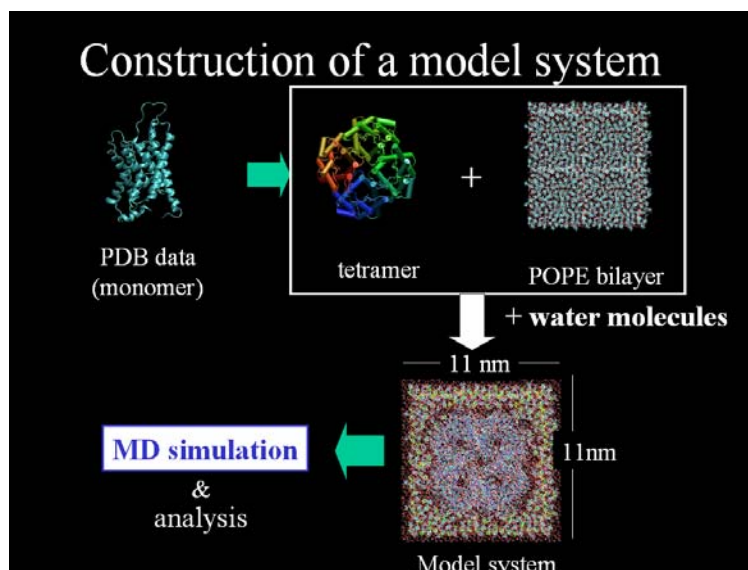


図1: アクアポリンの分子動力学計算のモデルの構築

結果および考察

Free AQP1 および Hg-AQP1 についてそれぞれ 40 ns の MD シミュレーションを行った。両者の全体構造について root mean square deviation (RMSD)を計算した。(図2)。Free AQP1 は 25 ns 以降、Hg-AQP1 は 30 ns 以降で RMSD の値が約 3.5 Å 付近に落ち着いている。各モノマー内部の RMSD を計算したところ、およそ 3 Å であったことから、全体構造の動きは、モノマー単位の移動ではなく、モノマー内部の動きであると予想される。また、2次構造をもつ H1-H6 の膜貫通ヘリックスと短い HB, HE ヘリックスに関して RMSD を計算したところ、全体のものよりも小さかった。つまり、全体構造のずれは、2次構造をもたない部位に起因すると考えられる。また、全体構造の場合と同様、Free AQP1 は 25 ns 以降、Hg-AQP1 は 30 ns 以降で RMSD の値が平衡に達している。ヘリックスの RMSD では、Free AQP1 と Hg-AQP1 に違いが見られた。このことから、両者は2次構造に違いがあると予測される。水銀が結合したシステインの周囲のアミノ酸残基を比較すると、両者に違いが見られた(図3)。水銀が結合することにより、主に水銀の結合したシステインの周囲のアミノ酸の側鎖の立体配置が変化することが分かった。水銀による水分子透過阻害機構は、水銀が pore を塞いでいるのではなく、水銀が結合し、周囲のアミノ酸の立体配置が変化し、pore の形が変形することによるが示唆された。pore の変形がどのような影響を与えるかを調べるために potential of mean force (PMF)により水分子通過の際の自由エネルギーを求めた(図4)。PMF の計算には WHAM 法を使用した。その結果、両者とも ar/R region を通過する際にエネルギー障壁が存在していることが明らかとなった。また、Hg-AQP1 は Free AQP1 よりも高いエネルギー障壁が存在した。このことから、水銀により誘導された pore の変形が水分子透過制御に影響を与えることが分かった。

今後の展望

PMF を求めることにより、分子・原子の結合により誘導された pore の変形が水分子透過制御に影響を与えることが分かった。この手法を応用することにより、タンパク質と基質(薬剤候補分子)の結合による機能への影響や結合部位の予測に役立つと期待される。

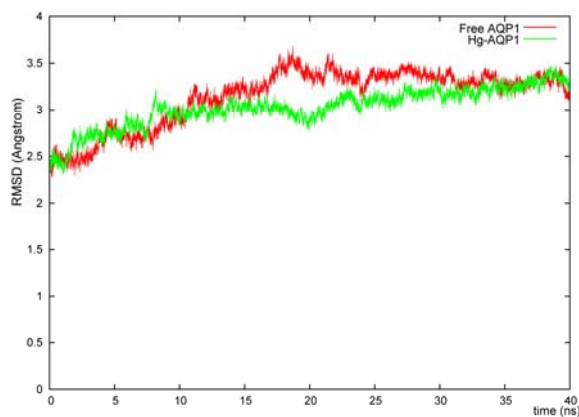


図2: 全体構造の C α について 40 ns 間の基準構造からのずれ(RMSD)を計算した。

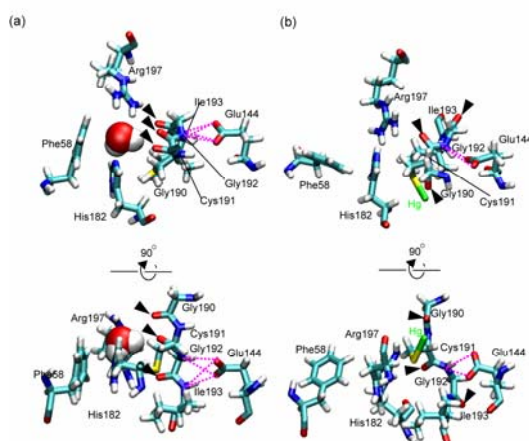


図3: 水銀が結合したシステインの周囲のアミノ酸残基の配座

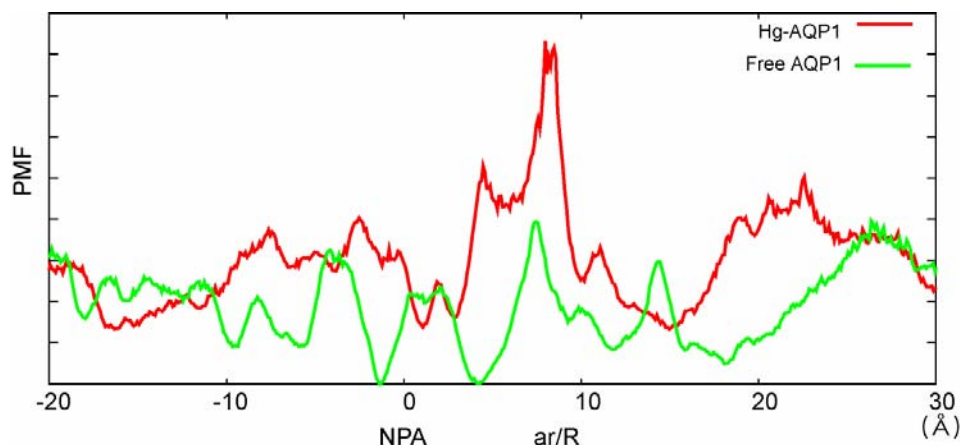


図4：水分子透過の際のPMF。赤線がHg-AQP1，緑線がFree AQP1を表す。両者ともar/R regionにエネルギー障壁が存在する。また、Hg-AQP1はFree AQP1よりも高いエネルギー障壁が存在する。水銀の結合により誘導されたporeの変形が水分子透過制御に影響を与えていることが分かる。

(8) 亜鉛によるアクアポリン4の水透過性抑制

〔慶應義塾大学大学院医学研究科〕

アクアポリン-4 (AQP4)は脳に最も多く存在する水チャネルであり、主にアストロサイトの終末足に発現している。脳における水の恒常性維持は生命活動において最も重要であり、AQP4が深く関わっていることが予想される。しかしながらその機能制御機構はほとんど明らかにされていない。ここで、われわれは、生体内微量元素である亜鉛がAQP4の水透過性を、プロテオリポソーム再構成系を用い可逆的に阻害することを見出したので報告する。塩化亜鉛のAQP4再構成リポソームに対する濃度応答効果を検討した結果、 IC_{50} 値は $287 \cdot M$ であり、この阻害された水透過性はAQP4再構成リポソームのWashoutやチオール還元剤との反応で完全に元に戻った。さらに、178番目のシステイン(Cys)をセリンに換えた変異体では塩化水銀の阻害効果がなかったことから、Cys178が塩化亜鉛の標的分子ということが示唆された。このCys178は細胞内に面するループDに存在するため、細胞内の亜鉛がAQP4の機能調節を担う可能性が示唆された。

背景

脳における水分子の恒常性維持は、生命維持に直結する最も重要な事象である。その恒常性の維持には水チャネルであるアクアポリン(AQP)の関与が予想される。AQPは哺乳動物では13種類が報告されており、そのいくつかは構造が決定されている。AQP4は脳において最も多く存在する水チャネルで、毛細血管を取り巻くアストロサイトの終末足に局在している。ノックアウトマウスの解析よりAQP4の脳浮腫への関与が強く示唆されているが、その機能調節に関してはほとんど明らかになっていない。最近、われわれはAQP4再構成リポソームを用いた系で、水銀がAQP4の水透過性を阻害すること、その標的分子が細胞内に面したCys178であることを見出した。このことは、AQP4の機能調節が細胞内から行われることを示唆した。しかしながら、生体内にお

いて水銀はほとんど存在しないので、他の Cys を標的とする生体内微量元素がこの役割を担っていると推測した。亜鉛は脳においてもっとも豊富に存在する生体微量元素の一つで、多くのタンパク質や酵素の活性に重要である。脳における亜鉛の濃度は $150 \cdot \text{M}$ と非常に高いことが知られているが、細胞内の遊離亜鉛イオンは picomolar と非常に少ない。これはメタロチオネインのような金属結合タンパク質が Cys を介し、細胞内の遊離亜鉛と結合、いわば貯蔵の役割をすることにより、細胞内の遊離亜鉛濃度を保っていることが考えられる。また、一酸化窒素などの酸化剤の存在下では亜鉛はこれらのタンパク質から遊離し、遺伝子発現やタンパク質の機能調節に働いていると考えられる。そこで、われわれは AQP4 が亜鉛によって機能制御されるという仮説を立て、AQP4 再構成リポソームを用いて実験を行った。

方法

1) AQP4の発現・精製、リポソームへの再構成 ラットAQP4遺伝子(野生型およびC178S変異体)を形質転換した *Saccharomyces cerevisiae*を発現用培地で培養し、回収した菌体から、膜画分を回収し、その膜を非イオン界面活性剤であるOctyl-glucoside(OG)で可溶化後、His-tagアフィニティカラムでAQPの精製を行った。AQP4、OG、脂質さらに水分子の指示薬として蛍光物質であるカルボキシフルオロセイン(CF)を混ぜた溶液を作製し、この溶液を、MOPS緩衝液で25倍希釈し、AQPを再構成したリポソームを作製した。

2) AQP4再構成リポソームの機能実験 水透過性はストップフロー装置を用いて、高張液とサンプル溶液を瞬時に混ぜ、リポソームの容積変化をCF quenching法にてミリ秒単位で追跡した。リポソーム内の水が外に出ることでリポソーム内の高濃度のCFが濃縮される。するとCFの蛍光の消失現象が起きる。この蛍光消失速度よりリポソームの容積変化速度を算出し、フィックの拡散方程式由来の水透過速度方程式に代入して、水透過速度を求めた。

結果

・ 塩化亜鉛の AQP4 再構成リポソームの水透過性への影響

生体内微量元素の AQP4 の水透過性に対する影響を調べるため、塩化物金属を AQP4 再構成リポソームに作用させ、水透過性を測定した。その結果、 $1000 \cdot \text{M}$ の塩化亜鉛と $100 \cdot \text{M}$ の塩化銅が AQP4 の水透過性を阻害した(Fig.1A)。また、水銀感受性の AQP1 ではその効果が認められなかった(Fig.1B)。また、AQP4 は His-tag を含むので、その影響を確認するために、His-tag を切断した AQP4 の水透過性も測定した(Fig. 1C)。その結果、これら亜鉛と銅による水透過性阻害効果は His-tag による影響はないと考えられた。

・ 塩化亜鉛と塩化銅の AQP4 水透過性に対する阻害効果の違い

塩化亜鉛と塩化銅の濃度応答効果を検討した結果(Fig. 2A)、 IC_{50} 値が塩化亜鉛では $287 \pm 58.5 \mu\text{M}$ であったのに対し、塩化銅では $7.7 \pm 2.5 \mu\text{M}$ であった。このことは、塩化亜鉛の AQP4 に対する親和性が塩化銅より低いことを示唆した。

これらの阻害効果の可逆性をプロテオリポソームを Washout すること、またチオール還元剤のジチオスレイトール(DTT)を添加することによって確認した(Fig2B,C)。その結果、塩化亜鉛添加ではいずれも AQP4 の水透過性が元に戻ったのに対し、塩化銅添加では Washout では戻らず、さらに DTT 添加でも部分的にしか阻害の回復が認められなかった。よって、亜鉛

の阻害効果は可逆的、銅の阻害効果は不可逆的ということがわかった。

- 亜鉛阻害効果の標的分子の同定

亜鉛の阻害効果の標的アミノ酸を同定するため水銀の標的である Cys178 の変異体の水透過性に対する影響を調べた(Fig.3C)。その結果、この変異体に対する塩化亜鉛の阻害効果は認められなかった。さらに、塩化銅でも同様の結果が得られた(Fig.3C)。よって、これらの金属の阻害は細胞内面に位置する Cys178 であることが示唆された。

- 酸化還元試薬の AQP4 に対する影響

Cys178 が AQP4 の機能阻害の標的アミノ酸であったため、Cys の修飾に関わる酸化及び還元剤添加時の AQP4 の水透過性に対する影響を diamide, S-nitroso- glutathione (GSNO), peroxyntirite (ONOO⁻), hydrogen peroxide (H₂O₂) および reduced glutathione (GSH)を 1 mM 添加し、検討した(Fig. 3A)。その結果、これらの試薬は AQP4 の水透過性に影響を与えなかった。また、diamide は Cys の-SH からプロトンを引き抜く作用がある。そこで、塩化亜鉛と併用実験を行ったところ塩化亜鉛は塩化銅と同等の阻害効果を示した(Fig.3B)。このことは、亜鉛は-SH からプロトンを引き抜き結合する、いわゆるメルカプチド結合で AQP4 に結合しているわけではないことが示唆された。このことは亜鉛の親和性が低く、可逆性を有する阻害形式をサポートする結果と考えられる。

Figure

Fig.1 塩化亜鉛及び塩化銅は AQP4 再構成リポソームの水透過性を減少させる。

(A) 1000 μ M MgCl₂, CaCl₂, MnCl₂, LiCl, CdCl₂, FeCl₃ and ZnCl₂ と 100 μ M CuCl₂ 処理(5 分間室温)した AQP4 再構成リポソームの水透過性 (n = 3-10, means \pm S.D.)

(B) 1000 μ M ZnCl₂ と 100 μ M CuCl₂ 処理(5 分間室温)した AQP1 再構成リポソームの水透過性 (n = 3, means \pm S.D.)

(C) His-tag を除去した AQP4 を再構成したリポソームの水透過性。(n = 3, means \pm S.D.)

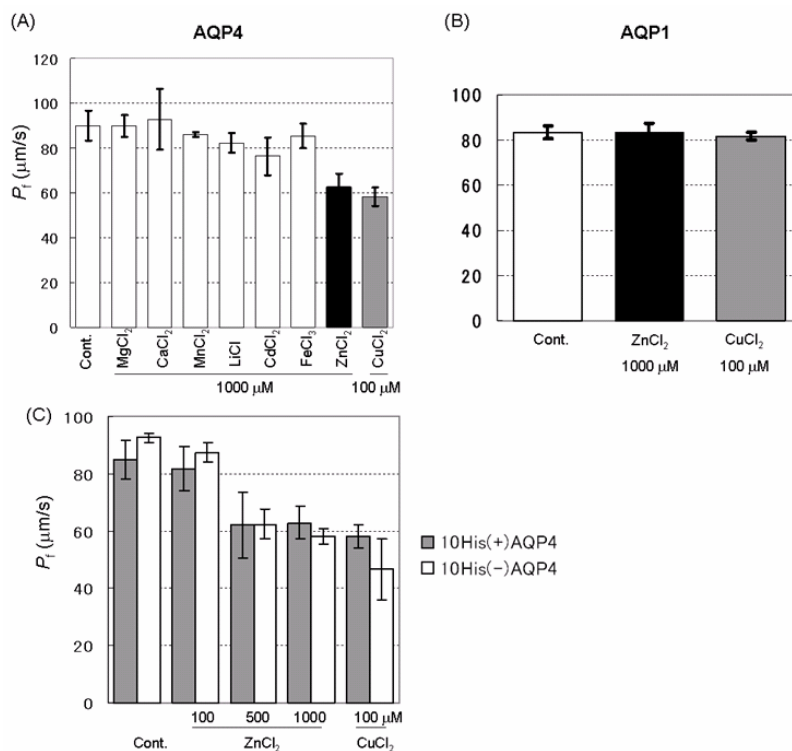


Fig.2 亜鉛と銅の AQP4 に対する阻害様式のの違い。
 (A) 亜鉛と銅の AQP4 に対する濃度応答効果。
 (n = 3-10, means ± S.D.) (B) 亜鉛及び銅を作用させた AQP4 再構成リポソームの Washout 効果。(n=3, means ± S.D.) (C) 亜鉛及び銅を作用させた AQP4 再構成リポソームの DTT 添加の効果 (n=3, means ± S.D.)

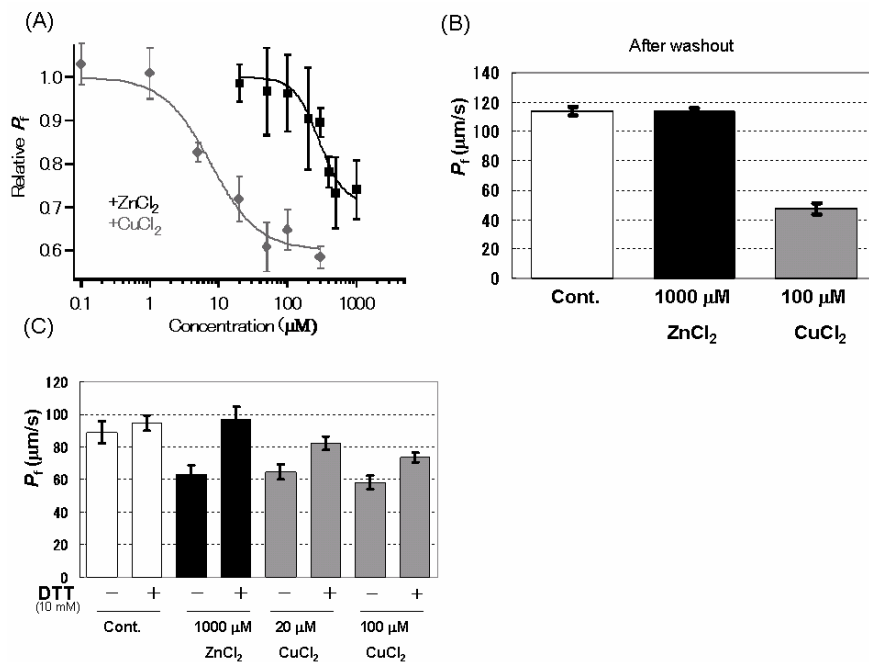
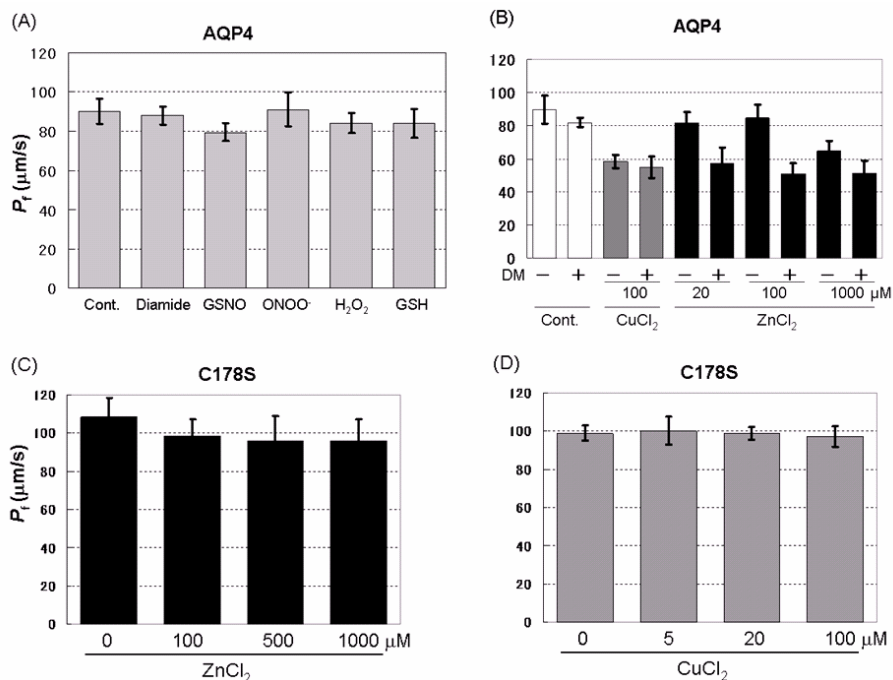


Fig.3 亜鉛による阻害効果の特性
 (A) 酸化及び還元剤の AQP4 水透過性に対する影響。反応時間は diamide、GSNO、H₂O₂ 及び GSH が 15 分、ONOO⁻ が 5 分で、それぞれ 1 mM を添加した。
 (B) 亜鉛と diamide(DM)の併用効果。
 (C, D) Cys178Ser 変異体に対する亜鉛(C)と銅(D)の効果



2.2 研究開発項目②「核磁気共鳴法等による膜タンパク質及びその複合体とリガンド分子の相互作用解析技術」

集中研究: 社団法人バイオ産業情報化コンソーシアム[バイオメディシナル情報研究センター(BIRC)、東京大学分室、三菱化学分室、第一三共分室、東レリサーチセンター分室、味の素分室、協和発酵キリン分室、エーザイ分室、東レ分室]

共同研究先: 独立行政法人産業技術総合研究所バイオメディシナル情報研究センター、国立大学法人東京大学大学院薬学系研究科

(1) 安定同位体標識タンパク質調製系の確立

(1)-1 NMR 解析を目指した新規酵母発現系の開発

[バイオメディシナル情報研究センター]

【序】

安定同位体標識タンパク質の発現には、容易かつ適切な標識コストにて大量発現が可能な大腸菌発現系が一般的に用いられる。しかし、ヒト由来のタンパク質や、高次構造形成に複雑なジスルフィド結合や翻訳後修飾が必要なタンパク質などの場合、大腸菌を含む原核細胞では発現が困難である場合も多い。高等な真核細胞である昆虫細胞や動物細胞を利用

した発現系では、活性を有するタンパク質発現に成功する確率が高いが、収量がそれほど高くないことに加え、最少培地を使用することができないため現実的なコストで安定同位体標識試料を調製することは困難になる。そのような場合、比較的大量に安定同位体標識タンパク質を発現できる真核細胞である酵母 *Pichia pastoris* (*P. pastoris*) 発現系が検討されることがあるが、炭素源として高価なメタノールを比較的大量に (> 15-20 mL/L) 必要とすることや、メタノール自体の細胞毒性のため、安定同位体標識試料調製系として汎用的に用いられているとは言い難い状況にある。(Table 1)。

そこで、我々はこれらの問題を克服するため、酵母 *Kluyveromyces lactis* (*K. lactis*) を活用した、新しい安定同位体標識タンパク質発現システムの構築を行った。その結果、本発現系を用いて大腸菌発現系と同等のコストで安定同位体標識タンパク質を大量発現させることに成功した。

【方法および結果】

1. *K. lactis* の炭素源および培養法の検討

本来、*K. lactis* は、タンパク質発現誘導剤および単一炭素源として、20 g/L のガラクトースを利用

Table 1. 従来のタンパク質大量発現系の特徴

	取り扱い 容易さ	安定同位体 標識のコスト	発現量	翻訳後修飾	哺乳類由来タンパク質/ 多重ジスルフィド結合形 成タンパク質発現
原核細胞					
大腸菌	◎	◎ <small>¹³C: グルコース (< 4g/L) ¹⁵N: ¹⁵NH₄Cl (1 g/L) 重水培養可能</small>	◎	×	△~×
真核細胞					
<i>Pichia pastoris</i> (酵母)	○	×	○(分泌◎)	○	○
		<small>誘導剤メタノールの毒性 同位体標識メタノールのコスト</small>			
Sf9, Sf21 (昆虫細胞)	△	×	△	○	○
CHO, HEK (動物細胞)	△	×	×	◎	◎

する。しかし、 ^{13}C 標識ガラクトースは高価 (^{13}C グルコースの 10 倍以上の価格) であるため、 ^{13}C 標識における障壁となりうる。そこで、グルコースでもタンパク質発現誘導は可能であるという報告¹⁾と、*K. lactis* の中でもグルコース資化性が比較的高いと言われている GG799 株 (New England Biolabs) を利用し、グルコースを用いて *K. lactis* に目的タンパク質を大量発現させる発現条件の検討を行った。モデルタンパク質として maltose binding protein (MBP) を使用し、ファーメンター培養で MBP を培地上清中へと分泌発現させた。しかし、グルコースで十分量のタンパク質発現を達成するには 20 g/L 必要であり (Fig 1, lane 1)、大腸菌発現系と同レベルの 4 g/L のグルコース量では十分な発現誘導は認められなかった (Fig 1, lane 2)。そこで、培養法を詳細に検討した結果、ファーメンター培養中に新鮮な培地をペリスタポンプで供給しつづける Fed-Batch 法により、大腸菌発現系とほぼ同レベルのグルコース量 (5 g/L) でも十分な発現誘導が達成されることを見出した (Fig 1, lane 3)。

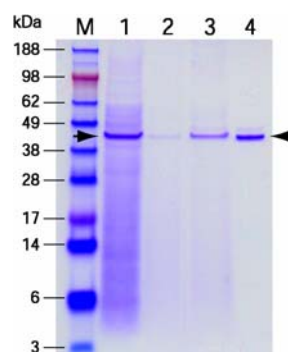


Fig 1. *K. lactis* 発現系で大量発現した MBP の SDS-PAGE 像
 レーン 1: 通常のファーメンター培養 (20 g/L グルコース使用)
 レーン 2: 通常のファーメンター培養 (4 g/L グルコース使用)
 レーン 3: Fed-Batch 培養 (5 g/L グルコース使用)
 レーン 4: 精製した MBP

2. 安定同位体標識効率の確認

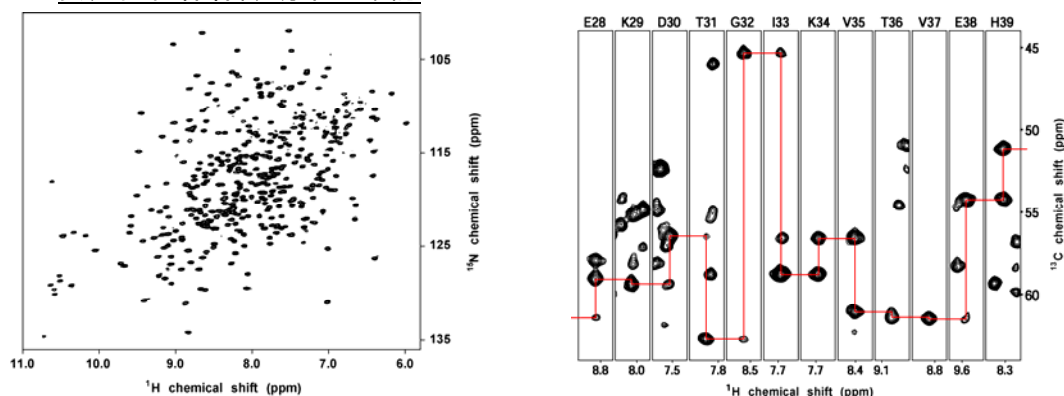


Fig 2. Fed-Batch 培養で大量発現させた均一 ^{13}C , ^{15}N 標識 MBP の NMR スペクトル
 パネル左: ^1H - ^{15}N HSQC スペクトル
 パネル右: 3 次元 HNCA スペクトルにおける各アミノ酸残基の ^1H - ^{13}C strip 表示 (赤線: 各アミノ酸残基間の ^1H - ^{13}C - ^{15}N シグナルの相関)

発現させた均一 ^{13}C , ^{15}N 標識 MBP の性質と標識効率を、 ^1H - ^{15}N HSQC および 3 次元 HNCA スペクトルの測定と MALDI TOF-mass 測定によって評価した。 ^1H - ^{15}N HSQC シグナルは良好に分散しており、発現した MBP が正しいフォールドを形成していることを確認できた (Fig. 2, パネル左)。3 次元 HNCA スペクトルもシグナル/ノイズ比が十分高く、主鎖シグナルの連鎖帰属を進めることができたことから、十分高い効率で ^{13}C , ^{15}N 標識されていることが示された (Fig. 2, パネル右)。さら

に、非標識、均一 ^{15}N 標識、および均一 ^{13}C , ^{15}N 標識 MBP について MALDI TOF-mass スペクトルの測定を行った結果、 ^{15}N および ^{13}C , ^{15}N の標識効率はいずれも約 93%、約 96%であることを確認した²⁾。

【考察】

酵母 *K. lactis* を使用し、かつ炭素源および培養法を工夫することにより、安定同位体標識のコストを大腸菌発現系と同レベルまで抑えた酵母発現系を構築することに成功した (Table 2)。

また *K. lactis* は、単一の炭素源を含む培地で培養し続けるだけで大量発現が可能であり、メタノールを添加する前に培地を交換する必要のある *P. pastoris* よりも作業上の手間が軽減される。そのような特長のため、*K. lactis* は取扱いや培養のスケールアップの容易さ、培養の再現性などにおいて優れており (Table 2)、目的タンパク質が酵母株で発現するかどうかの確認も比較的迅速にできるため、目的タンパク質を発現させる宿主を選択するうえで、大腸菌とともに第一選択肢になり得る発現系であると考えている。

	取り扱い 容易さ	安定同位体 標識のコスト	発現量	翻訳後修飾	哺乳類由来タンパク質/ 多重ジスルフィド結合形 成タンパク質発現
原核細胞					
大腸菌	◎	◎ <small>^{13}C: グルコース (< 4g/L) ^{15}N: $^{15}\text{NH}_4\text{Cl}$ (1 g/L) 重水培養可能</small>	◎	×	△~×
真核細胞					
<i>Kluyveromyces lactis</i> (酵母)	◎	◎ <small>^{13}C: グルコース (5g/L)</small>	◎ (分泌◎)	○	○
<i>Pichia pastoris</i> (酵母)	○	×	○ (分泌◎)	○	○
		<small>誘導剤メタノールの毒性 同位体標識メタノールのコスト</small>			
Sf9, Sf21 (昆虫細胞)	△	×	△	○	○
CHO, HEK (動物細胞)	△	×	×	◎	◎

Table 2. タンパク質大量発現系の特徴 (新たに *K. lactis* 発現系を加えた (赤字))

【今後の展望】

1. 重水素化タンパク質の大量発現法の確立

ヒト血清アルブミン (human serum albumin; HSA) は血漿中に含まれる総タンパク質の約 60% を占め、血管内外の水分の浸透圧調整に必須であるだけでなく、様々な薬物などの低分子の輸送タンパク質としても重要であることから、薬物と HSA の相互作用は、生体内での薬効発現にも深く関わってくるものである。しかし、ジスルフィド結合が 17 対存在し、分子量も比較的大きいため、native な立体構造の HSA を大腸菌で大量発現させることは不可能であった。

現在までの検討の結果、我々は native な立体構造の HSA を *K. lactis* で大量発現させることに成功している。HSA のような巨大分子を対象とする NMR 解析を進める場合、タンパク質の重水素化は必須となる。現在、*K. lactis* で重水素化タンパク質の大量発現を可能にする培養方法を探索中であり、重水素化 HSA を調製することができれば、これまで NMR 解析が困難であった薬物-HSA 複合体を対象とした相互作用解析を進展させることが可能になると考える。

2. アミノ酸選択的安定同位体標識法の開発

特定のアミノ酸を選択的に安定同位体標識する方法は、大腸菌ではほぼ確立されているが、*P. pastoris* をはじめとする酵母では未だ確立されていない。アミノ酸選択的標識は高分子量タンパク質の NMR 解析を進める上で必須の技術であり、重要な課題である。現在までの検討の結果、特定の安定同位体標識アミノ酸を Fed-Batch 法によって培養中に供給することで、そのアミノ酸のみを選択的かつ高効率で安定同位体標識した目的タンパク質を *K. lactis* で大量発現することに成功している。今後、他の安定同位体標識アミノ酸についても、その標識効率等を確認していくことで、*K. lactis* における体系的なアミノ酸選択的標識技術を確立する予定である。

3. 膜タンパク質の大量発現系の開発

一般的に、GPCR などの多数回膜貫通タンパク質は、大腸菌発現系では大量発現が困難である場合が多い。近年、酵母 *P. pastoris* 発現系で GPCR 等の大量発現の成功例が増えてきており、酵母発現系の有効性が指摘され始めている。今後、膜タンパク質の発現における *K. lactis* 発現系の有効性を検証していく予定である。

【参考文献】

- 1) Merico A et al. (2004) *J. Biotechnol.* **109**, 139–146
- 2) Sugiki T, et al., (2008) *J. Biomol. NMR* **42**, 159–162

(1)-2. NMR 試料の最適溶媒条件の新規スクリーニング法の開発

〔バイオメディシナル情報研究センター〕

【序】

生体高分子の溶液 NMR 解析において、試料分子間の非特異的な凝集は NMR シグナルの線幅の増大を引き起こし、構造情報取得の妨げとなる。したがって、試料が単分散の状態安定に存在する溶液条件を探索する必要がある。溶液条件の検討の手順として、様々な溶液組成の試料を調製して実際に NMR 測定を行うことで評価する手法がしばしば用いられるが、探索に必要な試料の量および費やす時間が膨大になるという難点があった。

我々は、溶液中における試料の並進拡散時間を少ない試料量で迅速に測定できる一分子蛍光分析法を用いることで、試料の凝集度と NMR スペクトルの質を迅速に評価できることを見出した。そこで、一分子蛍光分析法のハイスループット性を生かして、多種類の溶液条件を迅速、網羅的、体系的に探索する「溶液条件スクリーニングシステム」の構築を行った。

凝集性が高いために NMR 解析が困難であった ceramide trafficking protein (CERT) の PH ドメイン (CERT PH domain) の最適な溶液条件を本法にて探索した。その結果、凝集を改善させて良好なスペクトルを得られる溶液条件を一度のスクリーニング (従来の探索法の 7 倍以上の迅速性) で見出すことに成功し、CERT PH domain の溶液構造を決定することができたことから、本探索法の有効性が裏づけられた。

【方法と結果】

1. 一分子蛍光分析による凝集検出の検証

一般的にサブミリモル濃度という高濃度溶液である NMR 試料の場合、NMR 測定に必要とされる高濃度状態において特に凝集が激しくなる場合も多い。試料の凝集度を調べる既存の手法としては、微量透析法(目視で凝集の生成を検出する)、光散乱検出法、ゲル濾過法などが存在するが、しかしこれらの手法では、1) 見た目に沈殿が生じない凝集を検出できない、2) NMR 測定時と同等の高濃度状態で凝集度を見積もることが困難、といった問題があり、実際に NMR 測定を行ってみるまで NMR スペクトルの質を的確に判断できない場合も多い。その場合、凝集を解消する溶液条件を探索するには、様々な溶液組成の試料で NMR 測定を実際に行ってみるほかなく、したがって検討に必要な本数の NMR 試料を用意する必要があり、消費する試料量と NMR 測定時間は膨大なものとなる。

我々は、溶液中での試料の並進拡散時間を調べる手法である一分子蛍光分析法に着目した。そこでまず、一分子蛍光分析法が 1) NMR 試料のような高濃度試料の並進拡散時間を測定可能かどうか、2) 目視できない凝集を検出可能かどうか、を検証するため、pH の上昇によって凝集が進行し、それに伴って NMR スペクトルの質が低下するケモカイン MIP-1 α について一分子蛍光分析を行った。

一分子蛍光分析は、蛍光標識した試料が 1-10 nM 存在する溶液で可能である。そこで、NMR 測定に必要な高い試料濃度で一分子蛍光分析を行うため、80 μ M の MIP-1 α (非標識)溶液に蛍光標識 MIP-1 α が 1 nM となるように混和したものについて一分子蛍光分析を行った。

その結果、その状態で MIP-1 α の並進拡散時間が測定可能であっただけでなく、pH の上昇に伴う MIP-1 α の凝集亢進の様子を明瞭に検出することができた (Fig. 1e)。また、そこで検出された凝集の亢進は、NMR スペクトル

の質が低下していく様子に対応していた (Fig. 1)。すなわち、一分子蛍光分析を行うことによって、NMR 測定することなしに試料の凝集度と NMR スペクトルの質の良し悪しを判断することができることを見出した。この検討で、1つの NMR スペクトルの測定に 400 μ L の試料と約 3 時間の測定時間を必要としたため、合計 1.6 mL の試料と約 12 時間の測定時間を必要とした。一方、一分子蛍光分析は 1 条件あたり 30 μ L の試料と約 3 分の測定時間で済んだため、合計 270 μ L の試料で 9つの条件

を 30分弱程度で測定でき、NMR 測定を行うよりも多数の条件を少量の試料かつ短時間で検討することができた。また、一般的な NMR 試料のようなマイクロモル～ミリモルの高濃度試料溶液で凝集度が判断できたことから、本法は NMR 測定時と等しい条件下で凝集を判断することができる

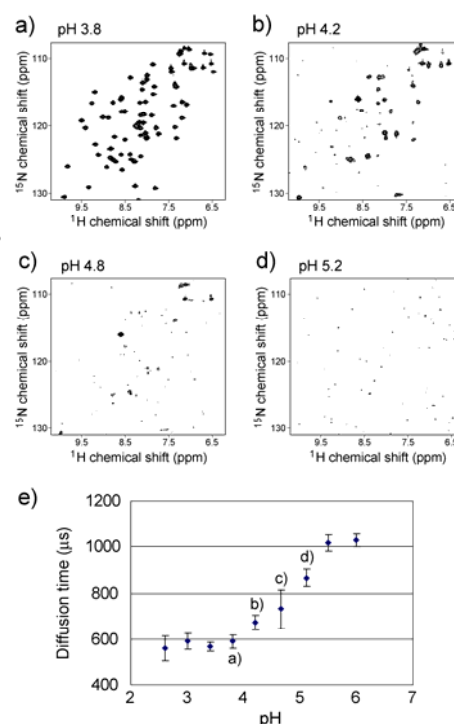


Fig. 1 MIP-1 α の ^1H - ^{15}N HSQC NMRと一分子蛍光分析。(a-d) 各 pH における MIP-1 α の ^1H - ^{15}N HSQC スペクトル。(e) 一分子蛍光分析で計測した各 pH における MIP-1 α の並進拡散時間。

点で、既存の手法に比べて信頼性が高いと思われる。

2. 溶液条件スクリーニングシステムの構築・実際の最適溶液条件の探索

試料の凝集が原因で NMR スペクトルの質が低く、解析が困難であった CERT PH domain について、凝集を解消する溶媒条件を一分子蛍光分析で探索した。また、探索にあたり、一分子蛍光分析は 384 穴プレートを使用し、必要な試料量が少ない(1 条件あたり 30 μL)、測定時間が短い(1 条件あたり 1~数分)という特長を生かし、多数の条件を網羅的かつ体系的に探索するためのシステムを構築した (Chart 1)。本システムは、目的タンパク質の蛍光色素標識と精製を行った後、まず pH と塩濃度および両者の組み合わせについて最適な条件を一分子蛍光分析で探索し、それで不十分であった場合は次の段階で様々な添加物(塩、ポリオール類、界面活性剤など)を検討するという、段階的にスクリーニングを行う戦略をとった (Chart 1)。

第一段階の探索の結果、pH 7.5 において最も並進拡散時間が低下した。しかし、その条件でも NMR スペクトルの改善は不十分であり、凝集は完全には解消されていないことがわかった (Fig. 2a, c)。そこで次に、第一段階の探索で最も良好であった pH 7.5 の溶液に 14 種類の化合物を別々に添加し、一分子蛍光分析を行った。その結果、化合物 4 (50mM Na_2SO_4) 添加時に最も小さな並進拡散時間を示し、この条件での NMR スペクトルは質が顕著に改善していた (Fig. 2b, d)。このスクリーニングで消費した CERT PH domain の総量は 450 μL (= NMR 測定 1-2 回分に相当)で、わずか 2 時間弱の一分子蛍光分析で 30 種類の溶液条件を 1 度に探索することができ、その結果 CERT PH domain の凝集を解消する条件を見出すことに成功した¹⁾。

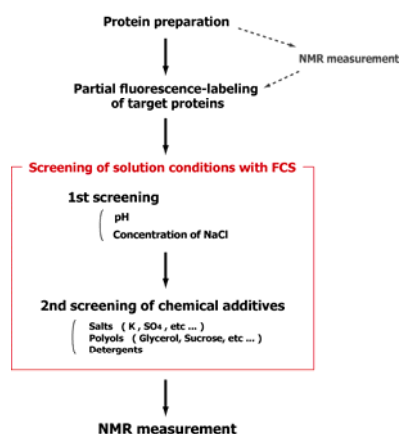


Chart. 1 一分子蛍光分析を用いて最適溶液条件を網羅的に探索するシステムの構築

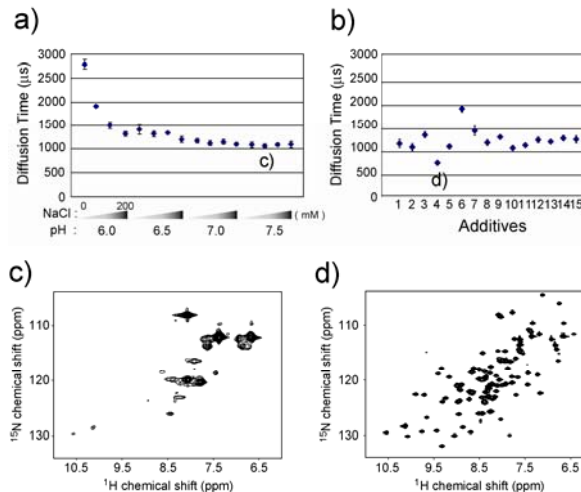
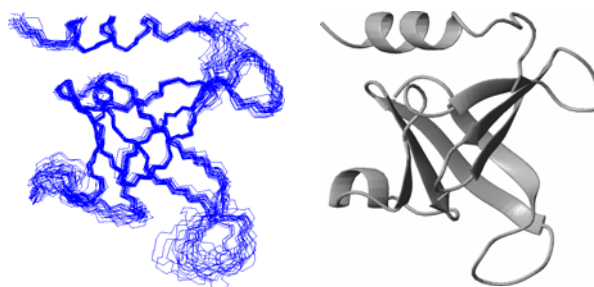


Fig. 2 一分子蛍光分析を用いて構築したシステムによる CERT PH domain の最適溶液条件の探索。(a) 最適な pH と塩濃度の探索。(b) pH7.5 における添加物の探索。(c) pH 7.5 における CERT PH domain の ^1H - ^{15}N HSQC スペクトル。(d) pH 7.5, 50 mM Na_2SO_4 における CERT PH domain の ^1H - ^{15}N HSQC スペクトル。

良好な溶液条件が決定できたことにより、CERT PH domain の NMR スペクトルの解析を進めることが可能になり、立体構造の決定を行うことができるまでに至った (Fig. 3)。

Fig. 3 構築したシステムを用いて CERT PH domain の最適溶液条件を見出した結果、NMR解析で立体構造解析を行うことが可能になった。
(左: 20 構造の重ね合わせ、右: 左図のリボン表示)



それに加えて、CERT PH domain のリガンド結合部位とその認識様式について NMR 分光法で解析を行うことも可能になり、CERT PH domain の機能的特徴の詳細と生物学的重要性を構造生物学的な見地から解明しつつある。

【考察】

一分子蛍光分析法は 1) 目視できない凝集(見た目には溶けているがマイクロに凝集しているもの)を検出できる、2) NMR 測定に必要な高濃度試料でも凝集度を判別できる という点で、NMR 測定時により近い状況で凝集が評価できる。また、一分子蛍光分析法で測定した試料の並進拡散時間が NMR スペクトルの質と対応することから、NMR 測定を行うことなしに一分子蛍光分析によって NMR スペクトルの質を評価できることを見出した。

一分子蛍光分析法のもつハイスループット性を生かし、凝集を解消する溶液条件を探索する体系的なスクリーニングシステムを構築した。凝集が原因で NMR スペクトルの質が悪く、解析が困難であった CERT PH domain について、実際に本システムを適用したところ、最適な溶媒条件を約 2 時間の一分子蛍光分析で見出すことに成功した。同様の探索を実際の NMR 測定で行う場合、約 15 時間程度を要すると考えられることから、本法は従来法に比べて 7 倍以上の迅速性を達成しているといえる。本法は NMR 試料についてだけでなく、結晶化のための溶液条件の探索等にも適用可能と考えられ、構造生物学の幅広い分野に貢献する手法となることが期待できる。

【今後の展望】

1. 必要試料量の少量化と検討条件数の拡大

現在の一分子蛍光分析装置では、384 穴プレートを使用し、1 穴あたりに必要な液量は 30 μL である。しかしながら観測領域である共焦点領域自体は極めて微小体積であることから、装置の改変を行うことで、より少量での測定を行うことは原理的には可能なはずである。1 穴あたりに必要な液量が 5 μL 程度である 1536 穴プレートが使用可能となれば、検討可能な条件数がさらに増え、より詳細な探索が可能になると期待される。

2. 試料調製の自動化・高速化

現在、多種類の溶液条件は実験者が 1 つ 1 つ手作業で調製しており、検討可能な溶液条件の数および実験の再現性に限界がある。そこで、タンパク質の結晶化条件の探索などで用いられている多種類溶液自動調製装置を本スクリーニング法と融合することにより、自動化、高速化、検討数の拡大が期待できる。

【参考文献】

1) Sugiki T, et al., (2009) *Protein Sci.* **18**, 1115–1120

(2) タンパク質複合体モデル構築を目指した NMR 測定法の開発

(2)-1. タンパク質複合体ドッキングモデル構築を可能とする新規 NMR 手法であるアミノ酸選択的交差飽和(ASCS)法、ならびに、分子動力的計算法の開発

〔東大分室(共同研究:中村チーム)〕

【序】

タンパク質の分子認識機構を立体構造の見地から解明することにより、生体内の多様なシステムである細胞認識や免疫機構、シグナル伝達などの重要な知見を得ることができる。タンパク質の結合界面を決定する NMR 手法として、化学シフト摂動法や H-D 交換法などが用いられるが、当研究室ではこれらの手法よりも正確な結合界面決定法として、交差飽和法(CS 法)の開発を行った。CS 法は目的分子(アクセプター)に結合している標的分子を交差飽和源(ドナー)とし、空間的に近接するアクセプター上の結合部位へと磁化飽和の移動を行う手法である。CS 法により、アクセプター上の結合界面を明らかとすることが可能であるが、ドナー全体を交差飽和源とするため、ドナー上のアクセプター結合部位に関する情報は得られない。しかしながら、より詳細な相互作用様式を明らかとするためには、両分子間の結合残基対の決定が必要となる。

そこで、ドナーに対して特定の amino 酸以外を ^2H とする、アミノ酸選択的 ^1H ラベルを行い、交差飽和源を 1 種類のアミノ酸に限定する、アミノ酸選択的

CS(ASCS)法を考案した(Fig.1)。本手法により、ドナー上においてアクセプターへ近接するアミノ酸残基に関する情報が得られ、近接残基対の同定が可能となる。

さらに、得られた構造情報を中村チームに提供し、ASCS 法により得られる構造情報を満たす複合体モデル構築を計算する分子動力的手法の開発を行った。

本手法を検証する相互作用系として、76 アミノ酸残基である yeast ubiquitin(Ub)と、その加水分解酵素であり、234 アミノ酸残基である yeast ubiquitin hydrolase C90S 変異体(YUH1)を選択した。両者の解離定数は 43nM であり、複合体の結晶構造は報告されている(PDB code : 1CMX)。Ub をアクセプターに、YUH1 をドナーとして本手法に適用した。

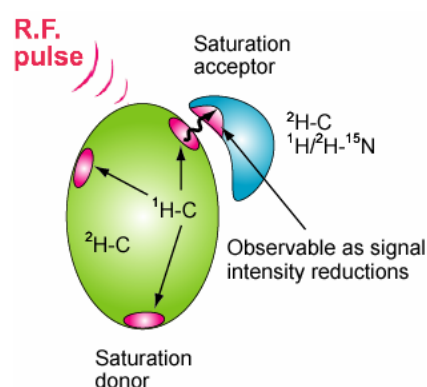


Fig. 1 アミノ酸選択的 CS 法

特定の amino 酸以外を ^2H 化したドナーをアクセプターに結合させる。ラジオ波照射に伴い、ドナーの特定 amino 酸の脂肪族 ^1H の磁化が飽和され、交差飽和現象により、近傍のアクセプター残基へと飽和が伝播する。これを NMR シグナルの強度減少として検出する。

【方法・結果】

1. アミノ酸選択的 ¹H 標識の選択性と標識率の定量的評価

本手法で用いるドナーは、特定のアミノ酸のみが選択的に ¹H 標識され、他のアミノ酸を均一に ²H 標識されている必要がある。このような標識法として、ドナーを発現する大腸菌を ²H 化 M9 最小培地で培養し、発現誘導時に、¹H ラベルを行う対象の ¹H アミノ酸を、培地に含まれる当該 ²H アミノ酸の 5 倍量添加し、6 時間後に集菌するという方法をとった。

大腸菌はアミノ酸を代謝する様々な酵素を保持しており、添加したアミノ酸と他のアミノ酸間の変換が予想される。本手法は添加した ¹H アミノ酸が目的のアミノ酸にのみ効率よく取り込まれることが必須であるため、各アミノ酸の代謝状況を NMR 法により各原子ごとに確認した。Asn, Glu を除く 18 種類のアミノ酸について調べた結果 (Table. 1)、タンパク質を発現させるタイムスケールにおいて、Asp, Gln, Ser 以外のアミノ酸が選択ラベル可能であることが示された。

Table 1 各アミノ酸の ¹H 標識の選択性と、部位ごとの標識率

amino acid	selectivity	percentage of ¹ H				
		H _α	H _β	H _γ	H _δ	others
Ala	selective	44	78			
Arg	selective	96	84	86	74	
His	selective	51	67			77 (H _β 1), 63 (H _β 1)
Ile	selective	<3	90		100	86(H _γ 12,H _γ 13), 100 (H _γ 2)
Leu	selective	10	100	100	97	
Lys	selective	72	93	93	93	75 (H _β)
Met	selective	40	71	94		80 (H _β)
Phe	selective	7	81		97	97 (H _β)
Pro	selective	54	79	79	56	
Trp	selective	10	45			75 (H _β 1), 77 (H _β 3), 77 (H _β 3), 79 (H _β 2)
Tyr	selective	<3	56	74	78	
Cys	selective	<3	30			
Gly	selective	37				
Thr	partially metamoblized	83	100	86		81(Ile H _β)
Val	partially metamoblized	11	70	68		40(Leu H _γ), 51(Leu H _β)
Asp	metabolized	54	N/A	N/A		
Gln	metabolized	<3	N/A			
Ser	metabolized	<3	N/A			

よく取り込まれることが必須であるため、各アミノ酸の代謝状況を NMR 法により各原子ごとに確認した。Asn, Glu を除く 18 種類のアミノ酸について調べた結果 (Table. 1)、タンパク質を発現させるタイムスケールにおいて、Asp, Gln, Ser 以外のアミノ酸が選択ラベル可能であることが示された。

2. ASCS 法の実施、結果のアミノ酸選択性の評価

1 に記述した方法で、目的のラベル状態を保持する YUH1(ドナー)を調製した。これを従来法により精製した。Ub(アクセプター)は従来の CS 法と同様に、[ul-²H,¹⁵N]体として調製した。Ub を完全に YUH1 結合型とすることで、アミノ酸選択的 CS 法に適用する NMR サンプルとした。

そこで、選択ラベルが可能な 7 種類のアミノ酸についてアミノ酸選択的 CS 実験を行った。Ala, Leu, Pro および Tyr 選択的 CS 実験の結果、Ub の L71, L73, R74, G75 の延べ 7 残基に大きな交差飽和が観測された (Fig. 2)。結晶構造において、これらの Ub 残基は ¹H ラベルを行った YUH1 のアミノ酸から 5 Å 以内に存在する。一方、5 Å 以上離れている残基には交差飽和が観測されなかった。このことは、近接アミノ酸からの交差飽和を観測する、アミノ酸選択的 CS 法が確立したことを示している。

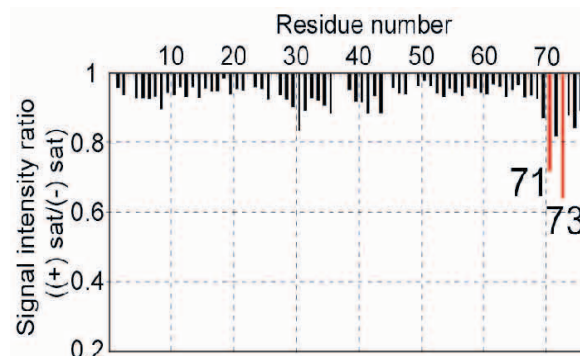


Fig. 2 Ala 選択的 ASCS 実験

横軸は Ub の残基番号、縦軸は交差飽和にとりまう NMR シグナルの強度減少率である。71 と 73 番目の Ub 残基に顕著な交差飽和が生じた。

3. ASCS 法による複合体中の分子間近接残基対の同定

交差飽和が観測された延べ 7 残基に対して、結晶構造に一致する交差飽和源の同定が実際に可能であるかの検証を行った。交差飽和源は、交差飽和が観測された Ub 残基から 5 Å 以内に存在すると考えられる。そこで、交差飽和を受けた Ub 残基のジオメトリーと一致する、YUH1 上で交差飽和源となったアミノ酸残基の組み合わせを検索した。まず、交差飽和源の種類と数を反映させて YUH1 から任意に延べ 7 残基を組み合わせた。これを最小二乗法により、交差飽和を受けた残基群と空間的に重ね合わせた。そして、対応する残基間の平均 RMSD をジオメトリーの一致度の指標とした。最小の RMSD を示す組み合わせにおいて、各 YUH1 残基と、それぞれにジオメトリーが近接する Ub 残基を 7 個の近接残基対とした結果、両者は実際に結晶構造中において 5 Å 以内に近接していた(Fig. 3)。よって、本手法によりタンパク質複合体の近接残基対の決定が可能であることが示された。

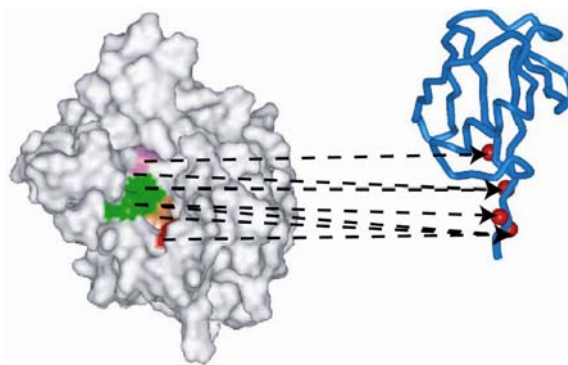


Fig. 3 ASCS 結果を満たす分子間近接残基対

赤の球で示した交差飽和を受けた Ub(右)の残基に対して、交差飽和源となった YUH(左)のアミノ酸残基を矢印で対応させた。

4. ASCS 結果を満たす複合体モデルを構築する分子動力学計算(中村チームとの共同研究)

次に、中村チームにより ASCS 結果を満たす複合体モデルを構築する分子動力学的(MD)計算法の確立が行われた。交差飽和の大きさは距離の 6 乗に反比例する。このことは、複合体の立体構造が明らかであれば、交差飽和の大きさを計算により算出することが可能であることを意味している。そこで、MD 計算を行った際に得られた複合体モデルについて、残基ごとに交差飽和の大きさを算出し、これを実際における ASCS 実験の交差飽和の大きさと比較することで、その差をエネルギー項 E_{sat} として取り扱った。これを通常行われる MD 計算におけるエネルギーに加算し、複合体のモデル構造を計算した(Fig. 4)。なお複合体モデル作成の計算プロトコール詳細は、研究開発項目③を参照のこと。このように、ASCS 実験結果を MD 計算に含めることで、詳細な複合体モデルが構築できることを示した。

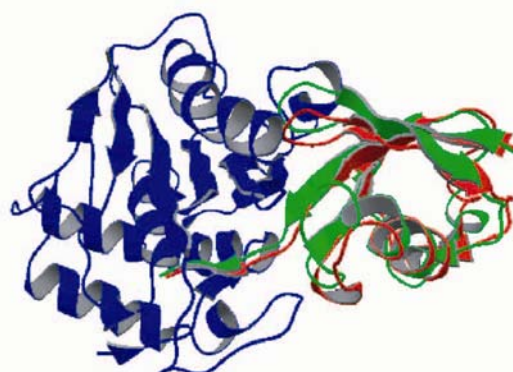


Fig.4 複合体モデルの構造

YUH(青)とUb(緑)で示した複合体の結晶構造に対して、YUH(青)とUb(赤)で示した複合体のモデル構造を重ね合わせた。

【まとめ】

本研究において我々は、タンパク質複合体の近接残基対を決定する手法として、アミノ酸選択的 CS 法を開発した。まず、本手法において選択ラベルが可能であるアミノ酸を明らかとした。そ

これらのアミノ酸選択的 CS 実験を行った結果、近接残基対間に選択的な交差飽和の検出に成功した。この実験結果から、結晶構造に一致する近接残基対の同定に成功した。さらに、MD 計算において、ASCS 実験におけるエネルギー項を含めることで、結晶構造における複合体構造と同等の複合体モデル構造を構築することに成功した。

従来の NMR 法を用いた近接残基対の決定には、複合体状態における両分子の側鎖水素原子の NMR シグナルおよび、分子間 NOE シグナルの帰属が必要であり、その同定には非常に多くの時間を費やす。また、分子量の増大に伴って、シグナルの帰属はより困難となることから、複合体として 30kDa 程度までの分子量に解析が限られていた。しかしながら、本手法においては非常に煩雑である上記の帰属は必要とせず、比較的容易に行える、複合体状態におけるアクセプター側の主鎖アミド水素原子のみの帰属が必要であるため、これまでに困難であった 30kDa 以上の複合体においても解析が可能であると予想される。

(3) 細胞膜複合体相互作用解析のための NMR 試料調製法の開発

(3)-1. 哺乳細胞内におけるタンパク質間相互作用を観測するための新規 In-cell NMR 測定法の開発

[東大分室]

【序】

細胞の機能発現は、細胞外シグナルに応答した細胞内生体分子の相互作用ネットワークの変化により達成される。細胞内では、膜タンパク質や、細胞骨格、細胞小器官などの構造体が、様々なタンパク質や補因子を保持しており、機能的な複合体としてその役割を担っている。そのため、生命現象の発動をタンパク質間相互作用の観点で解析する場合、機能を発揮する環境にある、よりインタクトに近い状態を構造生物学的対象とすることが望まれる。しかし、一般的な構造生物学的な解析では、高い純度のサンプルが解析対象となり、精製過程を必要とするため、構造体やタンパク質複合体を維持したまま調製することは困難である。また、膜タンパク質は、界面活性剤中では安定性が著しく低下するため試料調製が問題となるケースが多い。その一方で NMR 法は、生体試料に対し非侵襲測定が可能であること、安定同位体標識により解析対象のみを選択的に観測可能であるという他の構造生物学的解析法にはない特長を持っている。そのため、細胞をそのまま用いて、細胞内におけるタンパク質の立体構造・相互作用解析への適用が可能である。特に、動物細胞内における高次生命現象を、インタクトに近い状態にて構造的に解析することができれば、創薬へと応用する道も開かれる。そこで、本研究では、哺乳細胞を用いて、細胞内におけるタンパク質-タンパク質間の相互作用を検出するための NMR 測定法の開発を目的とした。

【方法】

細胞内にてアミノ酸残基毎の NMR シグナルを観測するためには、安定同位体標識を施した観測対象タンパク質を数 $10 \cdot M$ 程度細胞内に存在させる必要がある。これまでに、大腸菌内に過剰発現させた場合、および *Xenopus laevis* oocyte 内に microinjection した場合において、細胞内 NMR シグナルの観測が報告されているが、いずれの方法も哺乳細胞に対する適用は困難であるため、新たな手法が必要となる。そこで特別な装置を必要とせず、様々な種類の哺乳細胞に対して適用可能な汎用性の高い手法として、セミインタクト細胞の調製に用いられる streptolysin O (SLO) に注目した。SLO は、cholesterol と結合して形質膜上にポアを形成する

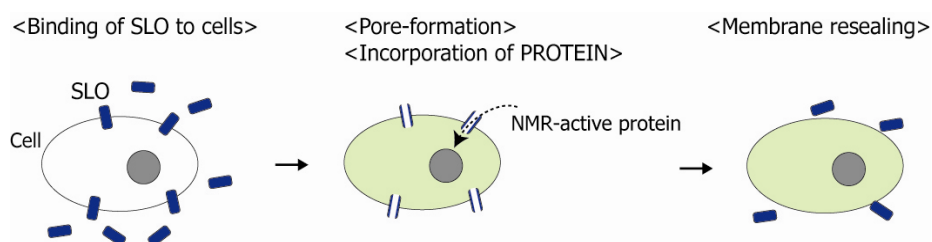


Fig.1 細胞調製方法の概略

細胞膜に対する SLO の結合、NMR 観測対象タンパク質 (あるいは、FITC 標識タンパク質) の導入、 Ca^{2+} 添加による細胞膜の再生 (resealing)、のステップから成る。

ため、細胞骨格や細胞内構造を保持した状態で、細胞質成分を自由に交換することができる。SLO が形成するポアの直径は約 30 nm であり、抗体 (150 kDa) や BSA (66 kDa) の導入が確認されているため、通常、NMR 観測対象となる 30 kDa 以下のタンパク質の導入は可能と考えられる。また、SLO により形成した形質膜上のポアは、細胞外に Ca^{2+} を添加することにより、再び塞がれる (reseal) ことが報告されている。そこで、SLO 処理と reseal を組み合わせることにより、細胞内部のみのタンパク質に由来する NMR シグナルを選択的に観測するための細胞調製方法の確立を目指した (Fig.1)。

細胞内にて NMR 観測および相互作用解析を行うためのモデル系として、細胞内に大量に存在する G-actin と相互作用する Thymosin-4 (T-4) を用いることとした。T-4 は、大腸菌により発現させ、精製したものをを用いた。細胞は、浮遊培養に適応させた 293F 細胞を用いた。

【結果および考察】

1. 細胞内 NMR 観測のための細胞試料の調製方法の確立

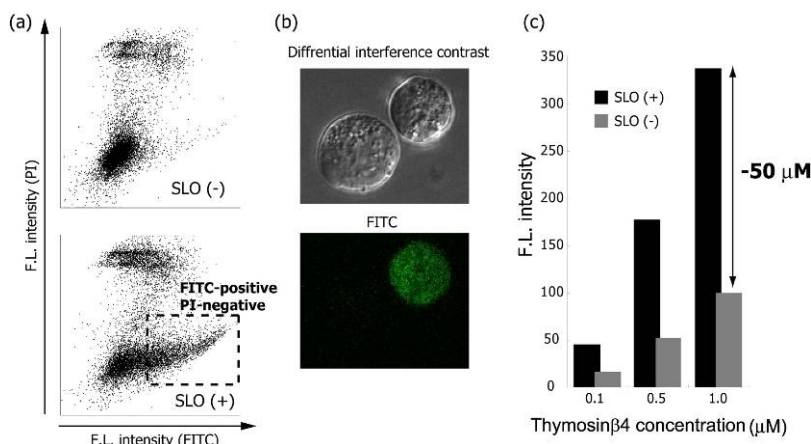


Fig.2 (a) 細胞調製時のフローサイトメトリーのプロファイル。SLO 非添加時(上)、SLO 添加時(下)。破線枠内の領域の細胞が、Tβ4 が導入された細胞集団を示す。(b) 細胞内に導入した Tβ4 の蛍光顕微鏡像。右上の細胞に FITC-Tβ4 が導入されている。(c) 細胞内 Tβ4 導入量の見積もり。横軸は、導入時に細胞外に添加した Tβ4 濃度を示す。SLO 非添加時と添加時の蛍光強度の差から、細胞内 Tβ4 導入量を算出した。

まず、 10^7 個程度の大量の細胞を SLO 処理し、NMR 観測可能な分子数の T・4 を導入するための細胞の調製条件を、Fluorescein isothiocyanate (FITC) にて蛍光標識した T・4 の導入を指標として検討した。Resealing の成否は、細胞調製後に、細胞膜非透過性の核染色試薬である Propidium iodide (PI) を添加して、染色されるか否かにて判断した。フローサイトメトリーにて FITC-positive/PI-negative の細胞集団が最も多く得られる条件を決定した (Fig.2a)。また、導入した FITC 標識 T・4 を共焦点蛍光顕微鏡にて観測した結果、細胞質内に分布していたことから、細胞内 actin と相互作用可能な状態にて存在することが示された (Fig.2b)。次に、細胞内 T・4 導入量の見積もりを行った。SLO 非添加時と添加時の FITC の蛍光強度 (蛍光波長 518 nm) の差より FITC-T・4 導入量を算出した結果、細胞内濃度の平均は $50 \cdot M$ 程度であったことから、NMR 測定に適用可能な量の T・4 が細胞内に導入されたと判断した (Fig.2c)。最適化した条件を用いて $[ul-^{15}N]$ T・4 を導入した細胞の $^1H-^{15}N$ HSQC スペクトルを測定した結果、12 時間の測定にて十分な S/N 比のスペクトルが観測された (Fig.3)。測定後の細胞をスピンドウンした後の上清成分の NMR スペクトルを測定した結果、細胞を用いた測定時と比較してシグナル面積が、1/10 程度であったことから、細胞を用いた場合に観測されたスペクトルは、細胞内に存在する T・4 に由来すると考えた。

2. 細胞内に導入した Thymosin・4 と内在性 G-actin との間の相互作用の検出

細胞内に導入した T・4 が、内在性の G-actin と相互作用しているか明らかにするため、細胞内 T・4 と緩衝液中 T・4 のスペクトル間における、各アミノ酸残基の化学シフト変化量を調べた

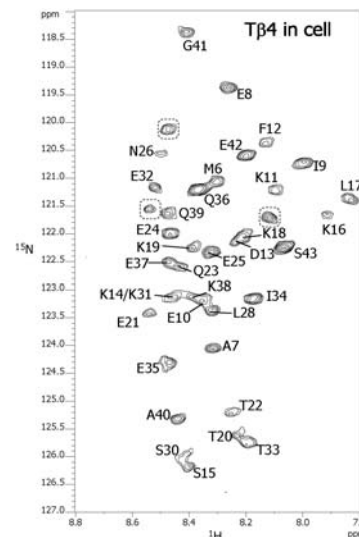


Fig.3 細胞内 Tβ4 の $^1H-^{15}N$ HSQC スペクトル

図中、破線にて囲んだシグナルは、緩衝液中にて観測されるシグナルと、顕著に異なる化学シフトを与えている。その他のシグナルについては、同程度の化学シフトであったため、帰属を入れている。

(Fig.4a)。両者間にて異なる化学シフトを与える原因として、G-actin との相互作用に加え、細胞質と緩衝液との pH・イオン強度の違いなど、さまざまな要因が考えられる。そこで、T・4 に対する G-actin の NMR 滴定実験を行い、各シグナルの化学シフト変化のパターンを比較した。生体内には ATP 型 G-actin (G_T)、ADP 型 G-actin (G_D) が存在しているため、それぞれを [^{15}N] T・4 に対して滴定した。

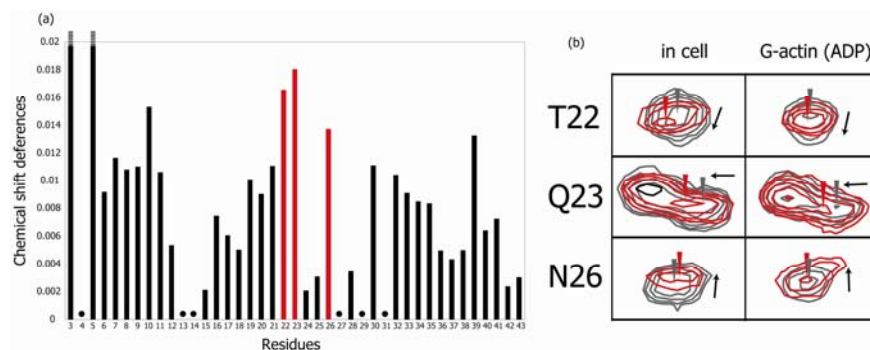


Fig.4 (a) 細胞内 T β 4 と緩衝液 (pH6.1)中 T β 4 間の化学シフト変化量。黒丸は、プロリン残基あるいは、シグナルが縮重したアミノ酸残基を示す。
(b) 細胞内にて観測された化学シフト変化と、ADP 型 G-actin 添加時に観測された化学シフト変化の様子の比較。G-actin 非存在下のシグナルを灰色にて示した。化学シフト変化の方向を矢印にて示した。

G_T の結合に伴う化学シフト変化は、NMR タイムスケールと比較して遅い交換であり、0.1 ppm 以上離れた位置に複合体に由来する広幅化したシグナルが新たに観測された。一方、 G_D の結合に伴う化学シフト変化は、NMR タイムスケールと比較して速い交換であり、50 \cdot M T・4 に対して 10 \cdot M G_D を滴定した結果、Q23, N26 にて最も大きな化学シフト変化が観測された。添加時に観測された化学シフト変化のパターンを調べた結果、細胞内 T・4 の NMR スペクトル中にて比較的大きな化学シフト変化が観測された T22, Q23, N26 にてその変化パターンがよく一致した (Fig.4b)。この結果は、細胞内に導入した T・4 が、T22-N26 を含む領域を介して G_D と相互作用していることを示す。一方で、T・4 の N 末端領域 (a.a. K3-K11) に集中して観測される 0.01 ppm 以上の比較的大きな化学シフト変化が、 G_D 滴定時には観測されていないことから、G-actin との結合以外の要因の存在が示唆される。

3. 細胞内に導入した Thymosin・4 の N 末端アセチル化修飾

生体内に発現した T・4 は、翻訳開始メチオニンが切断され、N 末端となったセリン残基がアセチル化修飾を受ける。そのため、細胞内に導入した T・4 の N 末端領域に観測された化学シフト変化が、細胞内における T・4 の N 末端アセチル化を反映している可能性を検証した。化学合成したアセチル化 T・4 の ^1H - ^{15}N HSQC スペクトルを測定し、細胞内にて測定した T・4 のスペクトルと比較した。その結果、両者に共通して K3, D5 の顕著な化学シフト変化と、1 個の新たなシグナルが観測された (Fig.5a)。この結果は、細胞内に導入した T・4 において N 末端のアセチル化が進行したことを示す。また、アセチル化による各アミノ酸残基の化学シフト変化量を調べた結果、アセチル化によって、E10 までの N 末端領域に比較的大きな化学シフト変化が誘起されており、細胞内にて観測された N 末端領域の化学シフト変化を説明する (Fig.5b)。

【総括】

本研究において、哺乳細胞内にタンパク質を導入し、細胞内にて NMR 解析するための細胞試料の調製方法を確立した。まず、細胞内に導入した T β 4 に由来する NMR シグナルの観測に成功し、T β 4 が内在性の ADP 型 G-actin と相互作用していることも示した。さらに、細胞内に導入した T β 4 が細胞内で N 末端のアセチル化修飾を受けることも判明した。以上の結果は、導入した T β 4 の NMR シグナルがインタクトに近い環境にて観測されていることを示す。したがって本手法を適用することにより、生体内の状態を保ったまま、膜タンパク質やタンパク質複合体とリガンドとの間の相互作用解析が可能と考える。また、今回観測されたアセチル化のみならず、リン酸化など翻訳後修飾をアミノ酸残基レベルにて経時的・定量的な解析を行う上でも有用な手法であると考え

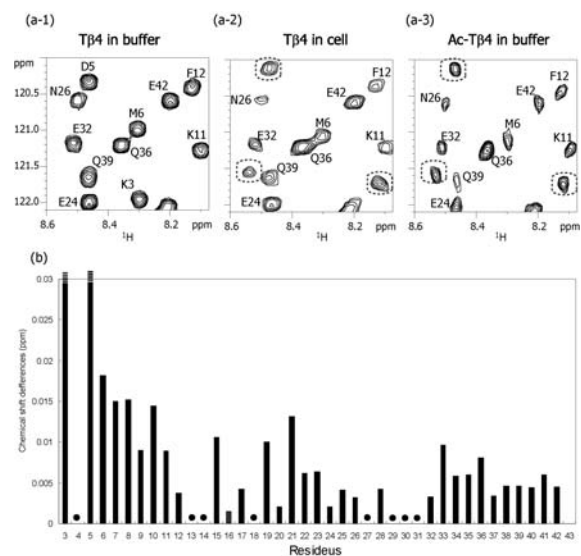


Fig.5 (a-1) 緩衝液中、(a-2) 細胞内における T β 4 のスペクトル。(a-3) 緩衝液中における N 末端アセチル化 T β 4 のスペクトル。破線枠内は、T β 4 を細胞内に導入した場合に、新たな化学シフト値に観測されたシグナルを示す。(b) N 末端アセチル化による化学シフト変化。黒丸は、プロリン残基あるいは、シグナルが縮重したアミノ酸残基を示す。

(3)-2. ケモカイン受容体 CCR5—リガンド間相互作用に関する構造生物学的解析法の確立

〔東大分室〕

【序】

CC-chemokine receptor 5 (CCR5) は、G タンパク質共役型受容体 (GPCRs) ファミリーに属する膜タンパク質である。CCR5 とそのリガンド (MIP-1 α , MIP-1 β , and RANTES など) 間相互作用は、G タンパク質を介したシグナル伝達を介して CCR5 を発現した白血球の遊走を引き起こす。加えて、CCR5 は HIV-1 の共受容体であることが報告されている。したがって、CCR5—リガンド間の相互作用機構を構造生物学的に解明することは、抗 HIV 薬の開発において重要である。しかし、その重要性にもかかわらず、CCR5 の構造生物学的解析は困難である。その原因は、CCR5 が生体膜から可溶化した状態において非常に不安定であり、生体内における機能を保った状態における大量調製法が確立していないためである。

CCR5 は可溶化した状態では不安定であるが、可溶化した状態から脂質二重膜中に再構成することにより安定性が向上する可能性がある。脂質二重膜再構成法の一つとして、近年報告された再構成 HDL (rHDL)がある。rHDL は、直径約 10 nm の脂質二重膜の周囲を、両親媒性 α -helix に富む apolipoprotein A-I (Apo A-I) 2 分子により囲まれた構造を持ち、脂質二重膜の内部に膜タンパク質を再構成することができる。rHDL は水溶性であり粒子径が均一であるという特徴から、構造生物学的解析にも適用可能であると考えられる。本研究では、rHDL を用いて溶液 NMR 解析に適した膜タンパク質再構成法の確立を行い、TCS 法を用いて CCR5 とそのリガンド間相互作用解析法の確立を目指した。

【結果】

1. rHDL 中に再構成した CCR5 の大量調製および性状解析

CCR5 はバキュロウイルス—昆虫細胞発現系を用いて Sf9 細胞により発現した。CCR5 を発現した細胞を超音波破碎した後、シヨ糖密度勾配遠心を用いて細胞膜画分を調製した。

CCR5 は、可溶化状態において急速に失活するため、可溶化した状態での精製は困難である。そこで、CCR5 が可溶化状態にある時間を短縮するため、rHDL による脂質二重膜再構成を導入し、以下の新規試料調製法を考案した (Fig. 1)。まず、CCR5 を発現した昆虫細胞膜画分を 1 % n-dodecyl-β-D-maltopyranoside (DDM) により可溶化した。これに Apo A-I を加え、DDM を除去することにより、CCR5 を脂質二重膜を含む rHDL 粒子中に再構成した。CCR5 を含む rHDL は、Ni アフィニティークロマトグラフィーおよび 1D4 抗体クロマトグラフィーにより精製した。

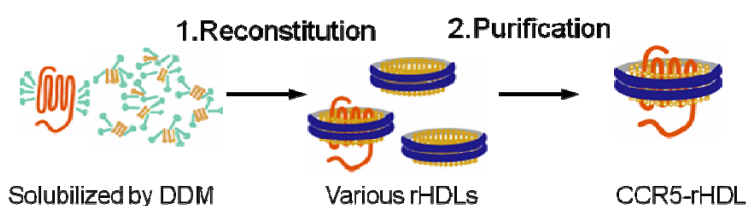


Fig. 1 rHDL 中に再構成した CCR5 新規調製法の概念図

界面活性剤による可溶化の直後に脂質二重膜再構成を行い、その後精製するという方法を考案した。

rHDL に再構成した CCR5 の純度は、SDS-PAGE 解析からおおよそ 80 % であった (Fig. 2)。また、rHDL に含まれる CCR5 の収量は、昆虫細胞 1 L 培養あたり約 10 · g であった。

再構成した CCR5 がリガンド結合活性を持つことを確認するため、プルダウンアッセイを行った。CCR5-rHDL を固定したカラム上に、MIP-1β が保持されたことから、CCR5-rHDL が MIP-1β 結合活性を持つことがわかった (Fig. 3)。さらに、再構成した CCR5 が精製した G_i タンパク質と共役することを確認するため、GDP-GTP 交換アッセイを行った。その結果、MIP-1β の濃度依存的に G_i サブユニット上における GDP-GTP 交換が促進されることがわかった (Fig. 4)。

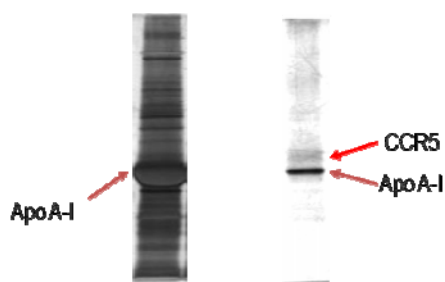


Fig. 2 CCR5-rHDL の精製前後の純度比較 (SDS-PAGE/銀染色)

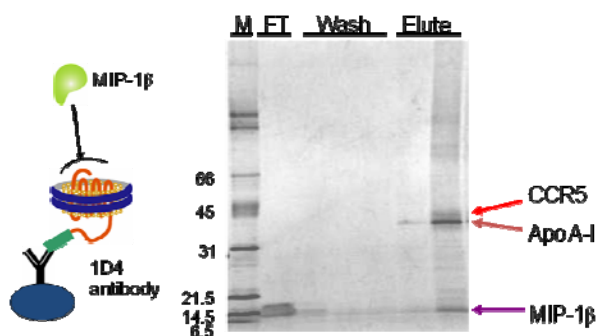


Fig. 3 CCR5-rHDL と MIP-1β を用いた pull down assay

(左) 実験概念図

(右) 各溶出フラクションに対する SDS-PAGE 解析 (銀染色)

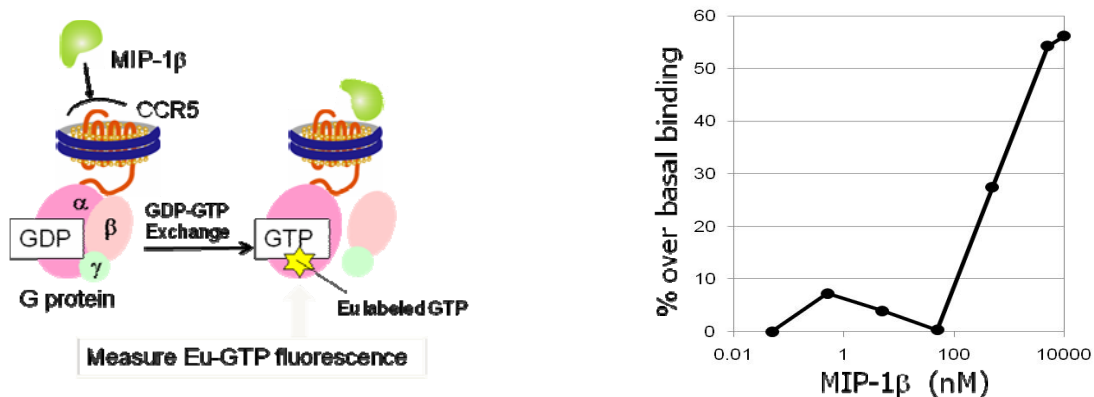


Fig. 4 rHDL に再構成した CCR5 のシグナル伝達活性に関する評価 (GDP-GTP 交換アッセイ)
 (左) 実験概念図 : MIP-1 \cdot 刺激を受けた CCR5 からのシグナル伝達により GTP 結合型となった G $\beta\gamma$ サブユニット量を、Eu 標識した GTP の蛍光強度をもとに評価した。(右) MIP-1 \cdot 濃度に対して、Eu 標識 GTP 蛍光強度の変化をプロットした。MIP-1 \cdot 濃度の増加に伴う、Eu-GTP 蛍光強度の増加を観測した。

加えて、rHDL の脂質二重膜中に再構成した CCR5 の安定性について、構造認識抗体 2D7 に対する結合活性を指標とした評価を行った。まず、昆虫細胞膜画分から 1 % DDM を用いて可溶化したミセル状態の CCR5 の安定性を調べたところ、4 $^{\circ}$ C 24 時間後で 2D7 結合活性が 10 % 未満に低下した。これに対し、rHDL 中に再構成した CCR5 では、4 $^{\circ}$ C 24 時間後にもとの活性の 85 % を保持していた。(Fig. 5)

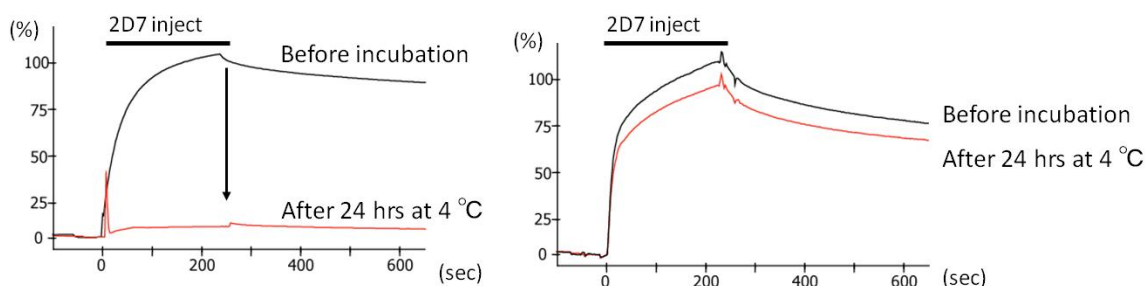


Fig. 5 DDM ミセルおよび rHDL 中における CCR5 の安定性比較
 センサーチップ上に CCR5 を固定し、2D7 抗体をアナライトとした際のレスポンスを示す。
 左) DDM ミセルにおける CCR5 は 24 時間後の 2D7 抗体結合量が 10 % 以下に減少した。
 右) rHDL 脂質二重膜中に再構成した CCR5 は 24 時間後に約 85 % の 2D7 抗体結合量を保持していた。

2. CCR5-rHDL と MIP-1 \cdot 変異体を用いた TCS 実験

CCR5—MIP-1 \cdot 間相互作用に関する構造生物学的知見を得るため、TCS 実験を行った (Fig. 6)。低濃度の CCR5-rHDL を用いて結合界面を同定

するために、高感度のメチル検出型 TCS 実験を適用した。Ile \cdot Leu \cdot Val のメチル基に選択的に 1 H, 13 C が導入された重水素化 MIP-1 \cdot (P8A/F13Y/E67Q) 10 \cdot M に対し、約 1 \cdot M の非標識の CCR5-rHDL を添加した試料を調製した。CCR5-rHDL の芳香族プロトンラジオ波照射した場合に、CCR5-rHDL との結合に起因する MIP

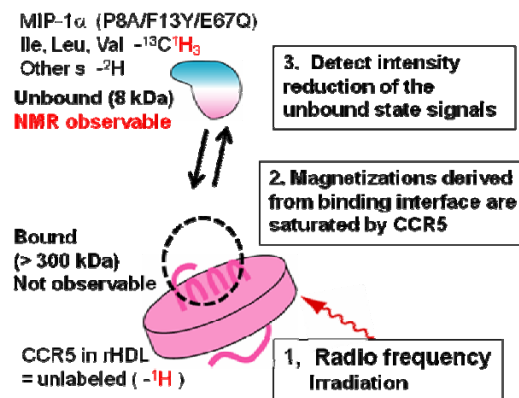


Fig. 6 CCR5-rHDL—MIP-1 α 間相互作用解析における TCS 実験の概念図
 ラジオ波の照射を照射した非標識の CCR5-rHDL に結合することにより、飽和が伝播し、MIP-1 α の NMR シグナル強度が減少する。

調べた。rHDL への非特異的結合の影響を見積もるコントロール実験として、CCR5 を含まない rHDL を用いた TCS 実験も行った。その結果、MIP-1 \cdot V59 の H \cdot シグナルにおいて、ラジオ波照射時に顕著なシグナル強度減少が観測された。この結果は、コントロール実験と比較して有意であり、V59 が CCR5 との結合界面に存在することを示している。

(Fig. 7)

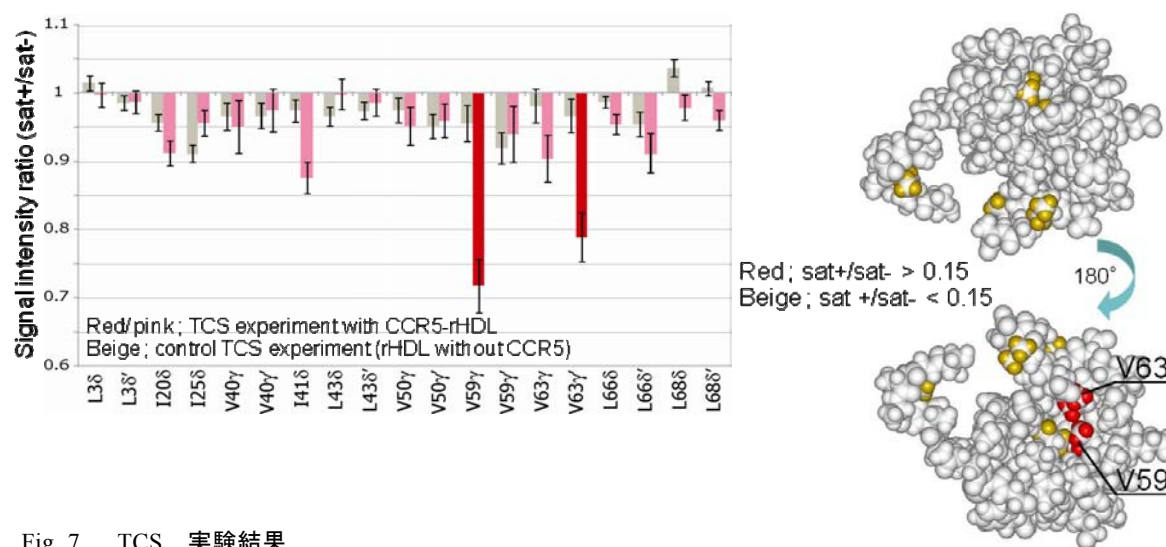


Fig. 7 TCS 実験結果

(左) 観測対象としたメチルプロトンの、ラジオ波照射によるシグナル強度減少比のプロット

(右) MIP-1 α 分子構造上における、シグナル強度減少が観測されたメチルプロトン(V59・V63) のマッピング

【考察】

本研究では、CCR5 の Sf9 細胞による大量発現および rHDL 脂質二重膜中への再構成を行った。rHDL に再構成した CCR5 は、リガンド結合活性およびシグナル伝達活性を持つことを示した。したがって、rHDL に再構成した CCR5 は生体膜中と同様の環境にあるといえる。

TCS 実験の結果から、MIP-1 \cdot の V59 が CCR5 との結合界面に含まれることが新たに示唆された。この領域については、CCR5 との相互作用に関する変異体解析が報告されていない。TCS 実験は、変異体解析が困難な疎水性領域に対して相互作用情報を得る手段としても有効であると考えられる。

以上から、本研究で確立した試料調製法と TCS 実験を組み合わせることにより、CCR5 とリガンド間の相互作用において構造生物学的解析が可能となった。この手法は、他の様々な GPCR においても、リガンドおよび細胞内エフェクター分子との相互作用解析へと応用が期待できる。

(4) 細胞膜複合体相互作用解析のための NMR 解析法の開発

(4)-1. HR-MAS を用いた膜蛋白質複合体における分子間相互作用の新規構造生物学的解析法の開発

[東大分室]

【序】

膜蛋白質の機能発現メカニズムを明らかにするためには、膜タンパク質・リガンド相互作用の解

明が必要である。従来の構造生物学的手法では、膜蛋白質は界面活性剤により可溶化された状態で解析されてきた。しかし、膜蛋白質は界面活性剤ミセル中でしばしばその高次構造が不安定化することが知られており、活性を保持した状態での試料調製法が必要とされる。

これまでに我々は、膜蛋白質をアフィニティビーズに固定化した状態で脂質二重膜中に再構成する”bead-linked proteoliposome (BPL)”を開発し、NMRを用いた相互作用解析法である転位交差飽和(TCS)法との併用により、カリウムチャンネルとその阻害毒素の複合体解析に成功した(Yokogawa M. *et al.* J. Am. Chem. Soc 2005)。しかし、ビーズなどの不溶性成分の混在により局所磁場が不均一となり、NMRシグナルが大きく広幅化するため、観測対象蛋白質は40残基程度に制限されていた。また、脂質二重膜中の膜蛋白質のように運動性の低下した分子においては双極子-双極子(DD)相互作用が亢進しており、TCS法において結合界面に観測される交差飽和がその周辺に拡散し、結合界面が明確に同定できない可能性が懸念されていた。

このような問題に対し、局所磁場の不均一性を解消しDD相互作用を減弱するマジック角高速回転(MAS, Fig.1 左)の適用により、スペクトルを高分解能・高感度化すると同時に、スピン拡散の抑制による明確な結合界面の同定が可能になると考えた。

そこで本研究では、MAS条件下においてビーズに固定化した蛋白質と溶液中の蛋白質の間(固液界面)における相互作用解析を可能とするNMR手法を確立することを目的とし、試料調製法の開発および交差飽和プロファイルに対するMASの効果の評価を行なった。

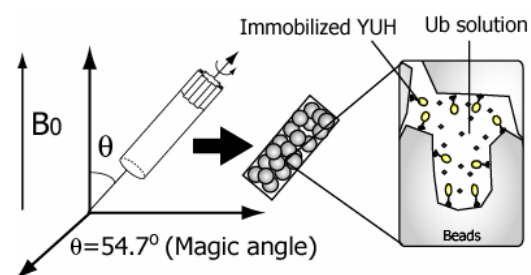


Fig. 1 MAS および MAS 試料内部の模式図

左：マジック角高速回転(MAS)とは、静磁場に対し、約 54.7°傾けた軸の周りに試料を高速回転することにより高分解能化を図る NMR 手法である。
右：YUH (黄色楕円) を固定化した多孔性シリカビーズを Ub (黒丸) 溶液に懸濁し、MAS 試料管に詰めた。

【方法】

解離定数(10 μ M)および複合体の立体構造が明らかとなっている Yeast ubiquitin (Ub: 8.6kDa)と Yeast ubiquitin hydrolase 1(YUH: 26kDa)の相互作用系を用いた。

1. 固定化担体・固定化方法の検討

固定化担体の材質として膨潤性セファロースと多孔性シリカ、固定化方法として His タグを介した Zn アフィニティビーズへの固定化法と、担体上のカルボキシメチル(CM)基に対してアミノ基を介して共有結合させる固定化法を検討し、担体の空隙率、非特異吸着の有無、YUH の固定化量、固定化された YUH の Ub 結合活性を解析した。

2. MAS による高感度・高分解能化の評価

ビーズと Ub の懸濁液の ^1H - ^{15}N HSQC スペクトルを MAS 適用時・非適用時に取得し、Ub 溶液単独の ^1H - ^{15}N HSQC スペクトルと比較することで、感度と分解能を評価した。

3. MAS 条件下の TCS 実験

2mM 均一 ^2H , ^{15}N 標識 Ub 溶液(50mM NaCl, 50mM NaPi, pH6.5, 80%D₂O)にて YUH 固定化ビーズを平衡化し、NMR 試料とし、25°C、回転速度 7000Hz にて TCS 実験を行ない、交差飽和プロファ

イルを取得した。

【結果】

1. MAS 条件下の TCS 実験に適した固定化担体・固定化方法を見出した。

固定化担体に関して、セファロースビーズはその膨潤性のために MAS の遠心力により円筒状の固相と Ub 溶液からなる内腔に分離した。一方、シリカビーズは MAS 回転後も均一に充填された状態を保持していた。さらに大きな比表面積により相互作用の場が確保されること、および 80% の高い空隙率により高感度測定が可能であることを見出し、固定化担体として多孔性シリカを採択した。

固定化方法に関して、Zn に対する His タグを介した固定化法では、YUH の Ub 結合活性の低下や Ub が Zn に対して非特異吸着することによる測定感度の低下および TCS 結果にアーティファクトが見られたのに対し、CM 基と YUH を共有結合させる固定化法によりこれらの問題が克服できることを見出し、固定化法として後者を採択した。

2. MAS の適用により NMR スペクトルの感度・分解能が約 6 倍向上した。

Ub の ^1H - ^{15}N HSQC スペクトルは、Ub 単独では ^1H 方向の線幅は 13Hz 程度であるのに対し、MAS 非適用時のシリカビーズ混在下では、局所磁場が不均一になることで ^1H の方向の線幅が 108Hz 程度まで広幅化し、観測可能な 72 残基中 42 残基が縮重していた (Fig. 2 赤)。一方、MAS の適用によりシグナルの ^1H の方向の線幅が 18Hz にまで先鋭化し、感度・分解能が約 6 倍向上した結果、縮重したシグナルは 6 残基にまで減少した (Fig. 2 青)

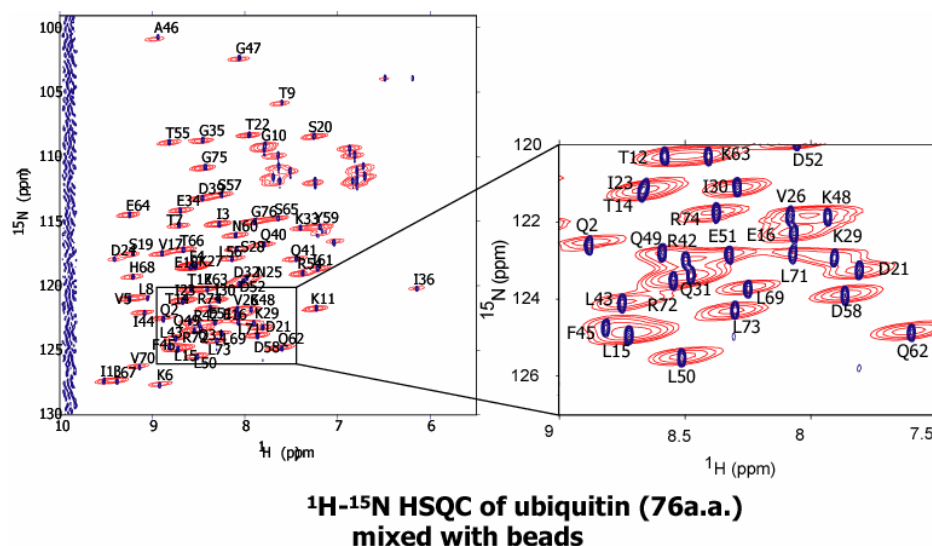


Fig. 2 silica beads 中の Ub の ^1H - ^{15}N HSQC スペクトルの重ね合わせ

(赤：非 MAS 条件下 青：MAS 条件下)

MAS の適用により ^1H 方向にシグナルが先鋭化し、縮重が解消されたことがわかる。

3. MAS 条件下の TCS 実験において、結合界面特異的な交差飽和の観測に成功した。

7000Hz の回転速度で TCS 実験を行ったところ、Ub の YUH 結合界面に顕著な交差飽和が観測され (Fig.3 左)、Ub 上の YUH 結合界面を同定することに成功した (Fig. 3 右)。一部、結合界面周辺の残基への交差飽和も観測されている。蛋白質のビーズへの固定化に伴う YUH の運動性低下

により、複合体形成時に分子間での DD 相互作用が亢進しスピン拡散による強度減少が結合界面周辺残基に及んだと考えている。

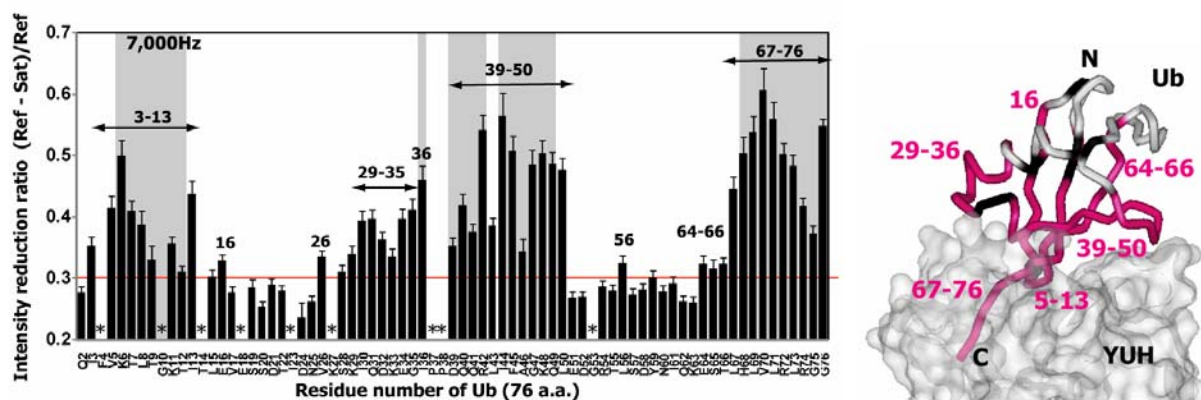


Fig. 3 MAS 回転速度 7kHz における交差飽和プロファイルと、大きな交差飽和が観測された残基の Ub-YUH 複合体の立体構造上へのマッピング。

左：横軸は Ub の残基番号、縦軸は YUH 非飽和時に対する YUH 飽和時の Ub のシグナル強度を表しており、理論上、交差飽和が観測される YUH から 5Å 以内に存在する Ub の残基をグレーの背景で示した。bar が長いほど、YUH からの交差飽和が大きいことを表している。(* は解析対象外の残基)

右： Ub-YUH 複合体の結晶構造に対するマッピング。

Ub を主鎖 tube 表示、YUH を表面表示にて示し、35%以上のシグナル強度減少を示した残基をマゼンタで示した。Ub の YUH に近接した残基が適切に同定されたことが分かる。(黒は解析対象外の残基)

【まとめ】

本研究により、固液界面の TCS 実験に対して MAS を適用することで、スペクトルが高感度・高分解能化するとともに、相互作用界面が明確に同定できることが示された。本手法は、BPLとして再構成した膜蛋白質の相互作用系に適用可能であり、脂質二重膜中の膜蛋白質の分子認識機構を解明する手法として期待される。

【今後の課題】

7000Hz の MAS 条件下の TCS 実験にて相互作用界面に特異的な交差飽和の観測に成功した一方で、MAS が DD 相互作用に与える影響に関してはより詳細に解析する必要がある。今後、静止条件下やより高速の回転速度にて TCS 実験を行ない、回転速度が DD 相互作用に与える影響を解析する。

(5) 創薬標的タンパク質の個別解析例

(5)-1. 転移交差飽和法を用いたケモカイン SDF-1 とその受容体 CXCR4 との相互作用解析

〔東大分室〕

【序】

ケモカインの1つである SDF-1 とその受容体 CXCR4 は、白血球の動態制御、造血、器官形成などの生理機能を有しており、また、ガンの転移や HIV-1 の感染など、多くの病態にも関係している。CXCR4 と SDF-1 との相互作用を特異的に阻害する薬剤を合理的に設計するには、両者の認識機構に関する原子レベルでの知見が必要となる。

先行研究における SDF-1 の変異体解析からは、SDF-1 の N 末端がシグナル伝達活性に必須である一方、受容体の親和性には SDF-1 の N 末端を除くコア領域が重要であることがわかっている。また、CXCR4 については、細胞外領域がリガンドとの親和性に重要である一方、膜貫通(TM) 領域がシグナル伝達に関与すると考えられている。このことから、CXCR4 と SDF-1 の間には、結合のみを担う相互作用 (SDF-1 コア領域-CXCR4 細胞外領域) と、シグナル伝達を担う相互作用 (SDF-1 N 末端-CXCR4 TM 領域) の、2 つの独立した相互作用が存在するという、2 段階仮説が提唱されている。しかし、この2種類の結合状態を示した知見はなく、これらの相互作用に関する原子レベルでの情報も得られていない。さらに、CXCR4 の N 末端に相当するペプチドと SDF-1 との相互作用解析はなされているものの、CXCR4 全長を用いた構造生物学的知見は皆無である。これは、CXCR4 が7回膜貫通構造をもつ G タンパク質共役型受容体 (GPCR) の1つであり、リコンビナント発現系や精製法の確立および結晶化がいずれもきわめて困難であることによる。

そこで、本研究では、SDF-1 上の CXCR4 相互作用部位の同定を行い、さらに、SDF-1 と CXCR4 との間に2つの独立した相互作用が存在するかを、構造生物学的に解明することを目的とした。

【方法および結果】

1. CXCR4 の発現・精製と性状解析

CXCR4 はバキュロウイルス-昆虫細胞発現系により発現し、ドデシルマルトシド (DDM) による可溶化後、C 末端に付加したエピトープタグを利用した抗体アフィニティー精製により精製した。SDS-PAGE 解析から、得られた CXCR4 の純度は 80 % 以上と見積もった。表面プラズモン共鳴 (SPR) 法による構造認識抗体 12G5 結合活性の解析から、得られた CXCR4 のうち約 50 % が正しくフォールドしていることを明らかとした。12G5 の結合量から見積もった、正しくフォールドした CXCR4 量は、1.5 L 培養あたり約 100 \cdot g であった。得られた CXCR4 が SDF-1 結合活性を有していることは、プルダウンアッセイにより確認した。さらに、SPR 法における 12G5 結合活性の解析から、低温に保持し、かつ glycerol を添加することが CXCR4 の安定性に必須であることが判明した。

2. 転移交差飽和法を用いた SDF-1 の CXCR4 相互作用部位同定

SDF-1 の CXCR4 相互作用部位を同定するため、安定同位体標識 SDF-1 (100 \cdot M) に対し、非標識 CXCR4 (10 \cdot M) を添加した条件で、転移交差飽和 (TCS) 法による解析をおこなった。低濃度の試料、低温、glycerol 存在下での解析を可能とするため、高感度検出に適したイソロイシン・

ロイシン・バリンのメチル基を選択的に¹H, ¹³C 標識してプローブとする、methyl-TCS法を採用した。結合・解離の交換を促進し、解離状態の SDF-1 へ飽和を効率よく伝播させるため、シグナル伝達活性を保持している一方、野生型よりも CXCR4 に対する親和性が低い SDF-1 R8A/R12A 変異体を利用した。さらに、コントロール実験として、活性のある CXCR4 とほぼ等量の均一²H 標識野生型 SDF-1 を加え、特異的な相互作用のみを阻害した条件で TCS をおこなった。得られた結果をもとに、変性した CXCR4・DDM ミセル・不純物に由来する非特異的相互作用の影響を見積もった。

TCS 実験の結果、非特異的相互作用の影響を除いたシグナル強度減少幅 ($\Delta_{\text{reduction ratio}}$) が大きかった残基を SDF-1 構造上にマッピングした (Fig.1A)。シグナル強度減少幅が大きかった残基は、SDF-1 の N 末端 (V3, L5) に加えて、コア領域の半周にもおよぶ帯状の広い範囲 (V18, V23, L26, I28, L29, V39, V49, L55) に分布していた。したがって、この領域が SDF-1 の CXCR4 相互作用部位であると結論した。

3. SDF-1 の変異体解析

上述の TCS 解析結果は、先行研究における変異体解析から推測された領域に加え、コア領域の L55 を含むより広い範囲が、CXCR4 結合部位を形成することを示している。そこで、あらたに同定された結合部位の1つである L55 の近傍に変異導入し、THP-1 細胞を用いたケモタキシスアッセイにより評価した。その結果、L55 の近傍に位置する D52 への変異体 (D52A, D52S) は、受容体に対する親和性が低下していることが示された。

4. AMD3100 の SDF-1-CXCR4 相互作用に与える影響

さらに、SDF-1 と CXCR4 との相互作用様式を明らかにするため、CXCR4 の特異的なアンタゴニストである AMD3100 を用いた解析をおこなった。AMD3100 は CXCR4 の TM 領域に結合することが示されており、その SDF-1-CXCR4 相互作用に与える影響を解析することで、CXCR4 TM 領域をブロックした状態でも SDF-1 が結合しうるかを解析することができる。

まず、AMD3100 存在下にて上述と同様の TCS 解析をおこなった (Fig.1B)。その結果、SDF-1 のコア領域では AMD3100 添加前と同等のシグナル強度減少が観測されたものの、N 末端 (L3, V5)

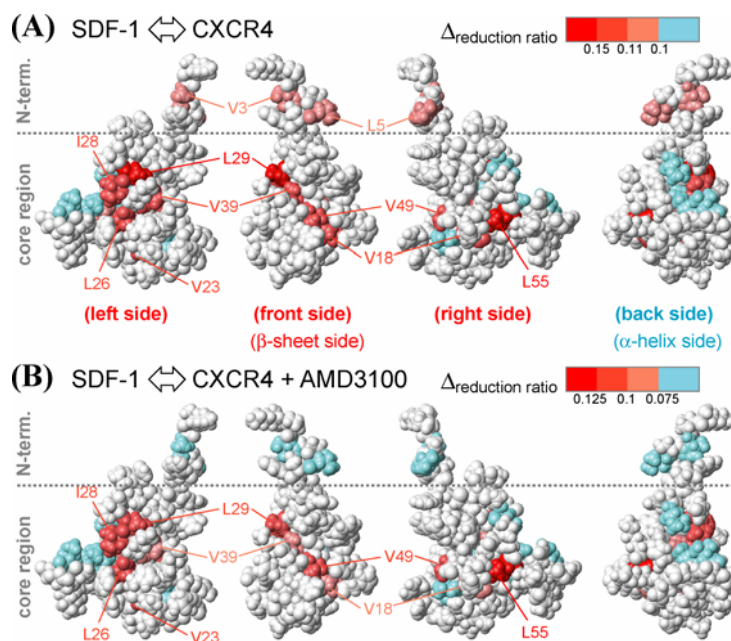


Fig.1 TCS 解析の結果

(A) SDF-1 の CXCR4 結合部位同定を目的としておこなった TCS 解析の結果。SDF-1 の構造を CPK 表示し、シグナル強度減少幅 ($\Delta_{\text{reduction ratio}}$) が大きいものを濃い赤～薄い赤、小さいものを青にて、残基ごとに色付けした。いずれも N 末端を上側に向け、コア領域の β -sheet を手前に向けたものを正面側、それとは逆の C 末端 α -helix を手間に向けたものを背面側とした。正面側から、左右に 90 度ずつ回転したものをそれぞれ、右面側および左面側とした。

(B) 1 mM AMD3100 存在下での TCS 解析の結果。(A) と同様に、SDF-1 の構造上に色付けして表示した。

のシグナル強度減少はほとんど観測されなくなった。このことは、AMD3100 存在下では SDF-1 N 末端は CXCR4 から解離しているものの、SDF-1 コア領域は依然として CXCR4 に結合しうることを意味する。

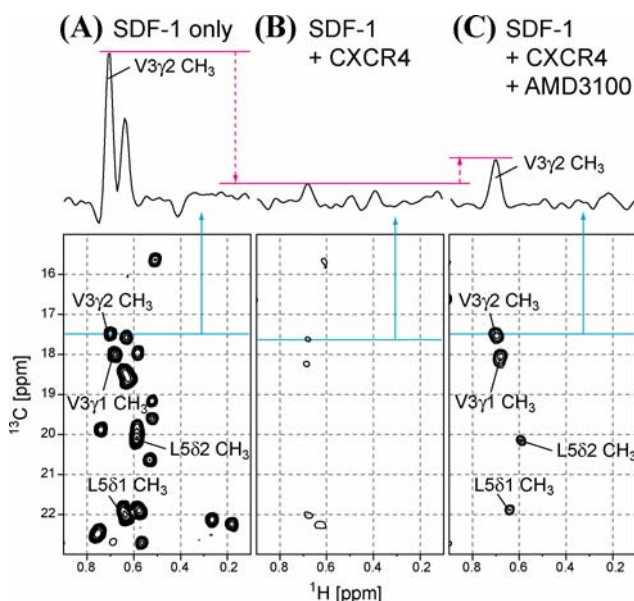


Fig.2 過剰量 CXCR4 存在下での

SDF-1 の NMR 解析

単独 (A)、過剰量 CXCR4 (20 μ M) 存在下 (B)、過剰量 CXCR4 および 1 mM AMD3100 存在下 (C) のロイシン・バリン選択的 ^1H , ^{13}C 標識 SDF-1 (10 μ M) の ^1H - ^{13}C HMQC スペクトル。スペクトルの上部には、それぞれのスペクトルの青線にて示した 1 次元の切り出しを示す。

さらに、過剰量の CXCR4 存在下で SDF-1 の NMR 測定をおこない、AMD3100 の影響を解析した。ロイシン・バリンのみを選択的に ^1H , ^{13}C 標識した遊離の SDF-1 の ^1H - ^{13}C HMQC スペクトル上では、13 個のロイシン・バリン残基に由来する 26 個のシグナルが観測された (Fig.2A)。これに過剰量の CXCR4 を添加したところ、すべてのシグナルが著しく強度減少した (Fig.2B)。このスペクトル変化は、SDF-1 が CXCR4 との複合体形成にともない、高分子量化したことを示している。さらに AMD3100 を添加したところ、N 末端 (V3, L5) に由来するシグナルのみ強度が回復して観測された (Fig.2C)。このことは、AMD3100 の添加にともない、SDF-1 の N 末端が CXCR4 から解離し運動性が上昇したことを示している。一方で、SDF-1 の N 末端を除くコア領域のシグナルは十分な強度で観測されず、実際に AMD3100 添加後も SDF-1 のコア領域は CXCR4 と相互作用していることが明らかとなった。

【考察】

1. SDF-1 上の CXCR4 相互作用部位

TCS 解析の結果から、SDF-1 の CXCR4 相互作用部位として、コア領域の広い範囲が寄与していることが明らかとなった。これらの領域には、塩基性残基 (R12, R47 など)、酸性残基 (E15, D52)、疎水性残基 (L29, V39, V49, L55 など) が分子表面に側鎖を向ける形で存在している。SDF-1 は他の多くのケモカインと異なり、受容体選択性が高いことが知られている。CXCR4 は、これら異なる性質をもつ多くの残基を組み合わせることで高い特異性を発揮していると考えた。

2. 2 段階結合モデル

今回の NMR 解析からは、AMD3100 は、SDF-1 の N 末端を CXCR4 から解離させるものの、SDF-1 コア領域の CXCR4 への結合への影響は小さいことが示された。AMD3100 が CXCR4 の TM 領域に結合することを考慮すると、SDF-1 の N 末端が CXCR4 の TM 領域に、SDF-1 のコア領域が CXCR4 の細胞外領域に、それぞれ独立に結合すると考えることが妥当である。この結果は、先

行研究にて推測された 2 段階の相互作用が実際に存在していることを、明確に示す知見といえる。

近年、複数の GPCR の X 線結晶構造が報告されており、低分子リガンドは、GPCR の TM 領域に存在する入り口の制限された cavity に結合することが示されている。同様の cavity 構造は、CXCR4 においても存在することが推測できるが、このような構造はケモカインのような比較的分子量の大きいリガンドの結合には不利である。上述の 2 段階の相互作用は、このような性質をもつケモカイン-ケモカイン受容体の相互作用に適しているといえる (Fig.3)。SDF-1 コア領域のみの独立した相互作用は、CXCR4 細胞外領域への迅速な結合とともに、効率よく SDF-1 を CXCR4 上につなぎとめておくことを可能とする。SDF-1 の N 末端は、この状態においても運動性が高く、効率よく結合空間を探索し、比較的狭い TM 領域の cavity への結合を促進する。

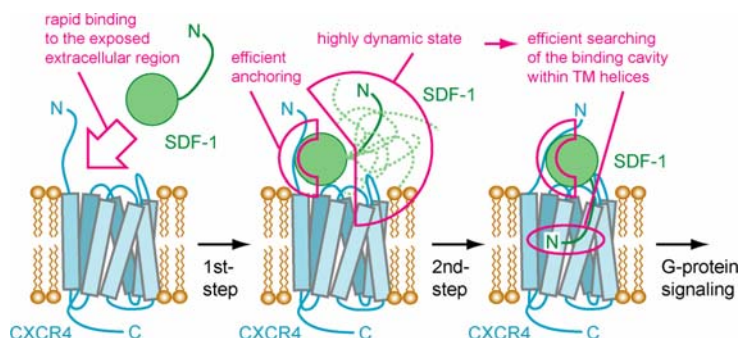


Fig.3 SDF-1 と CXCR4 の 2 段階相互作用
 2 段階相互作用モデルにおいては、まず SDF-1 のコア領域が CXCR4 の細胞外領域に結合する (1 段階目)。その後、運動性の高い SDF-1 の N 末端が、CXCR4 TM 領域に存在する cavity へ結合する (2 段階目)。さらに、TM 領域に構造変化が起こり、細胞内の G タンパク質へとシグナルが伝達される。

(5)-2. 白血球のローリングにおける CD44 リガンド結合ドメインの構造平衡

[東大分室]

【序および目的】

CD44 は細胞外マトリックスを構成するヒアルロン酸 (HA) に対する主要な受容体で、HA との相互作用を介して、内皮細胞上におけるローリングなどのリンパ球の動態制御に重要な役割を果たしている。CD44 は一回膜貫通型のタンパク質で、細胞外 N 末端側にある HA 結合ドメイン (HABD) を介してヒアルロン酸を認識する。HABD は、ヒアルロン酸結合タンパク質の間で保存された LINK モジュールとその N および C 末端に位置する付加配列領域からなり、その立体構造が NMR および X 線結晶構造解析から明らかとなっている (Teriete et al, *Mol Cell* 2004)。我々は HA 結合に伴う HABD の構造変化を解析し、HABD は C 末端領域が一定の構造を形成しない partially disordered (PD) form へと変化することを示した (Fig.1b, Takeda et al, *J Biol Chem* 2006)。しかし、2007 年に発表された HA 結合状態の X 線結晶構造では、C 末端領域が一定の構造を形成した ordered (O) form を

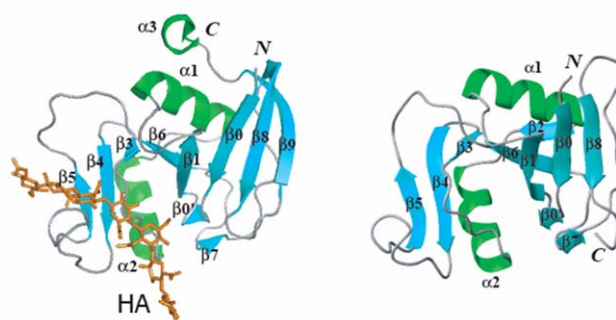


Fig.1 HA 結合状態の CD44 HABD の立体構造
 (a) HA8 が結合した状態の HABD の結晶構造 (PDB code: 2JCR)。
 (b) HA6 が結合した状態の NMR 構造 (PDB code: 2I83)。

形成しており (Fig.1a, Benerji et al, *Nat Struct Mol Biol.*, 2007)、HA 結合に伴って CD44 に構造変化が起こるか否かは不明である。そこで、本研究では CD44 の HA 結合に伴う構造変化についてさらなる解析を行い、CD44 の細胞接着能および血流下のローリング活性の制御機構を HABD の HA 認識に基づいて明らかにすることを目的とした。

【結果と考察】

1. 野生型の CD44 HABD には、リガンド結合の有無にかかわらず 2 状態の平衡にある。

HA 非結合時の CD44 HABD の ^1H - ^{15}N HSQC スペクトルには、O form に由来するシグナルに加えて、PD form に対応するマイナーなシグナルも同時に観測された (Fig.2a)。逆に HA 存在下の HABD のスペクトル上には PD form 由来のシグナルに加え、ordered form 由来のシグナルが観測された (Fig.2b)。さらに、2 状態間の交換をクロスピークとして観測する N_{zz} exchange 実験において、両シグナル間にクロスピークが観測された (Fig.2c)。よって、CD44 HABD は、リガンド非結合時において 2 状態間の平衡にあり、HA の結合は 2 状態間の平衡を PD form へシフトさせることが明らかとなった。以上の結果より、HA 結合状態の HABD の結晶構造は、溶液中にわずかに存在する O 状態を反映しており、生理的条件下の HABD は主に NMR 構造にて明らかとなった PD form を形成することが明らかとなった (Fig.2d)。

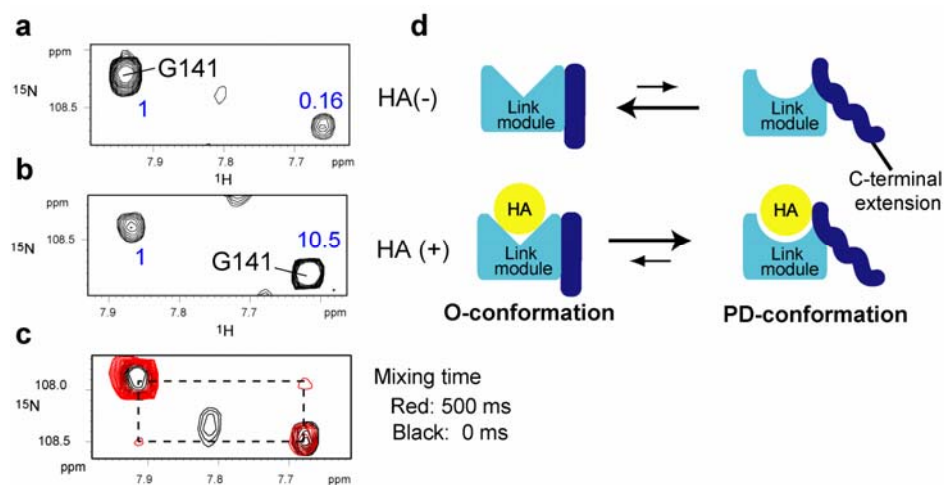


Fig.2 溶液中における HABD の 2 状態間の構造平衡

(a) リガンド非存在状態および(b)存在状態の HABD の G151 由来シグナル
(c) N_{zz} exchange 実験 (d) CD44 の HA 存在・非存在状態における構造平衡

2. Y161A 変異体は恒常的に PD form を形成し、HA に対して高親和性である。

O 状態の結晶構造には、リガンド結合部位から PD form でランダムコイル化する C 末端領域にかけて、水素結合のネットワークが形成されている (Fig.3a)。そこで、この相互作用の維持に特に重要な役割を果たすと予測される Y161 をアラニンに置換した変異体を作製し、その構造平衡がどのように変化するか調べた。HA 非存在下の Y161A の ^1H - ^{15}N HSQC スペクトルは、野生型と比較し大きな変化が観測されたことから、Y161A では分子全体にわたる構造変化が生じていることが明らかとなった (Fig.3b)。また、HA 存在下の野生型のスペクトルに観測される PD 状態由来のシグナルと Y161A のシグナルはよく一致した (Fig.3c)。また、Y161A のスペクトルには O 状態由来のシグナルが観測されないことから、平衡が完全に PD form にシフトしていると結論した。

次に Y161A HABD と野生型 HABD の HA 結合活性を SPR 法により解析した結果、Y161A は、野生型と比較して、7 倍高い HA 結合活性を有することが明らかとなった。野生型は、2 状態の平衡として存在するのに対して、Y161A 変異体は PD 状態としてのみ存在することから、O 状態よりも PD 状態の方が HA に対する親和性が高いことが示された。

3. Static な条件では細胞に発現した Y161A CD44 は野生型 CD44 と同程度の HA 結合活性を示す。

CD44 の構造平衡の生理的意義を明らかにするため、野生型 CD44 および Y161A CD44 の安定形質転換株を VMRC-LCD 細胞 (肺癌由来) を用いて調製した。フローサイトメトリーにより、同程度の発現量の細胞集団を選別し (Fig.4a)、ウェスタンブロッティングにより野生型と Y161A CD44 で糖鎖修飾などがほぼ同一であることを確認した (Fig.4b)。次に蛍光標識を施した HA との親和性を調べたところ、野生型、Y161 変異体ともに同程度の HA 結合能を示した (Fig.4c)。また、HA を固定化したプレートに対する、野生型および Y161A 変異体 CD44 発現細胞の接着を調べたところ、両者に有意な差は観測されなかった。よって、野生型と Y161A 変異体を発現する細胞の HA 接着能は、Static な条件では同等であると結論した (Fig.4d)。

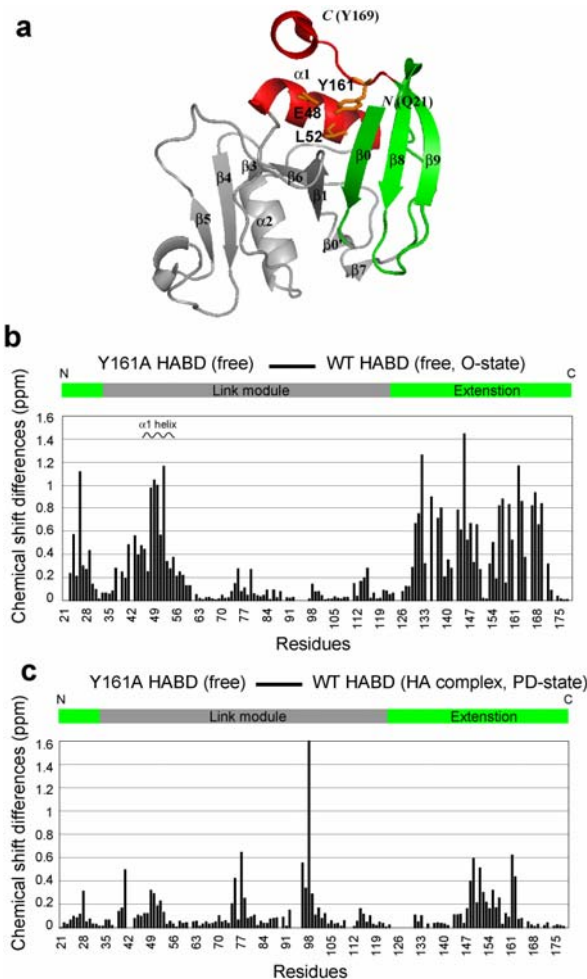


Fig.3 Y161A 変異体の構造

(a) O 状態の HABD における Y161A の位置 (b) リガンド非存在下および (c) 存在下の野生型 HABD と Y161A のシグナルの化学シフト変化

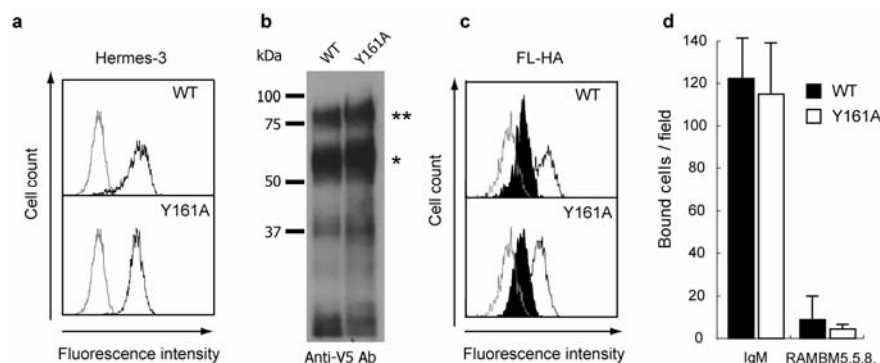


Fig.4 CD44 発現細胞の性質

(a) VMRC-LCD 細胞における野生型および Y161A 変異体 CD44 の発現量の比較。(b) ウェスタンブロッティングによる解析 (c) CD44 発現細胞の蛍光標識 HA に対する親和性。(d) HA 固定化プレートに対する CD44 発現細胞の接着能。

4. Shear stress 存在下において Y161A 変異体はローリング活性が損なわれる。

CD44 の構造平衡が、shear stress が存在する条件における細胞のローリング活性に影響を及ぼすか調べた。様々な shear stress 下において、HA を固定化したキャピラリーに CD44 発現細胞を還流させ、HA との相互作用によって生じる細胞のローリングの振る舞い(ローリング速度・細胞数)を調べた(Fig.5a)。その結果、野生型 CD44 発現細胞では、細胞のローリングが観測されたのに対し、Y161A 発現細胞では、ほとんどの細胞は接着後にローリングを示さず強い接着(Firm adhesion)のみを示した。また、ローリングや強い接着を示す細胞数は、すべての shear stress において Y161A 変異体よりも野生型 CD44 を発現する細胞のほうが多いことも明らかとなった(Fig.5a)。

5. Y161A 変異体は野生型よりも解離速度が遅く、Tether 形成効率が低い。

次に、細胞上の 1 分子の CD44 と HA との相互作用を調べるため、Transient tethering 実験を行った。Transient tether 実験では、低濃度の HA を固層化した条件において細胞を還流させるため、ローリングが観測されず、細胞上の 1 分子の CD44 と HA との相互作用に由来する一過的な接着の形成(transient tether)が観測される。一定時間における Tether 形成頻度を調べた結果、Y161A と比較して野生型の方が有意に tether 形成頻度が高いことが明らかとなった(Fig.5b)。また、tethering を示した細胞の滞在時間を tethering 細胞数の自然対数に対してプロットして、一次の解離速度定数を算出した。その結果、野生型と比較して Y161A 変異体では cellular off rate が低下していることが示された(Fig.5c)。野生型が主に形成する O 状態は、shear stress 下における新しい tether 形成に有利であり、また HA 結合状態においてもローリングに必要な速やかな解離に有利であると考えられる。O 状態を形成しない Y161A でローリングが観測されなかったことから、HA 結合状態における O 状態と PD 状態間の平衡が CD44 のローリングに重要な役割を果たしていると結論した。

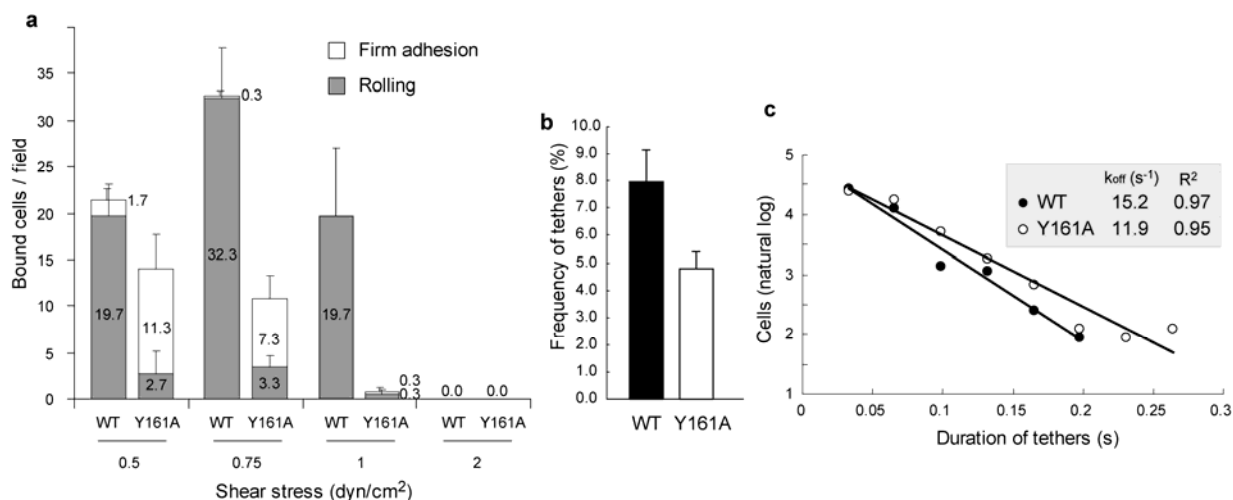


Fig.5 CD44 発現細胞のローリング活性の比較

(a) 野生型および Y161A 変異体 CD44 発現細胞のローリング活性とその shear 依存性。 (b) Transient tether の形成頻度。 (c) Transient tether の滞在時間に基づいた cellular off rate の見積もり

【参考文献】

Takeda M., Ogino S., Umemoto R., Sakakura M., Kajiwarra M., Sugahara KN, Kawashima H., Miyasaka M., Terasawa H., *Shimada I. (2006) Ligand-induced structural change of the CD44 hyaluronan-binding domain revealed by NMR. *J Biol Chem.* 281, 40089–40095.

(5)-3. Discoidin domain receptor 2 (DDR2) によるコラーゲン認識と活性化機構の解明

[東大分室]

【序】

Discoidin domain receptor 2 (DDR2) は、細胞外マトリックスの主要な構成成分であるコラーゲン線維をリガンドとする受容体型チロシンキナーゼ (RTK) であり、細胞の増殖、遊走などの生理機能や腫瘍細胞の転移、アテローム性動脈硬化などの疾患に関与している。DDR2 の活性化は、細胞外のディスコイジン (DS) ドメインとリガンドであるコラーゲンが結合することで誘起される (Fig.1)。我々はこれまでの研究において、DS ドメインの立体構造を解明し、転移交差飽和 (TCS) 法によりコラーゲン線維に対する結合残基を明らかとすることに成功した。しかしながら、DDR2 のコラーゲン結合様式を理解するためには、これらの知見に加えて DDR2 とコラーゲン複合体構造の解明が不可欠である。

また、DS ドメインへのコラーゲン結合により細胞内へとシグナルが伝達される機構は不明である。DDR2 は一般的な RTK とは異なり、リ

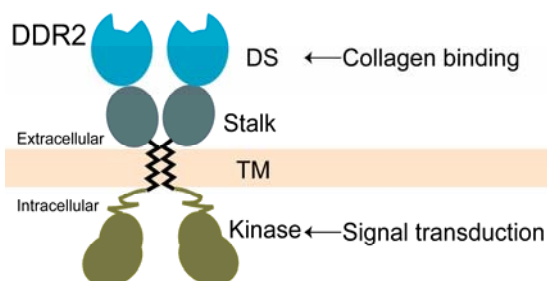


Fig.1 DDR2 ドメイン構造の模式図と各ドメインの機能
 DDR は DS ドメイン (アミノ酸番号: 30-185) によりコラーゲンと結合する。また、膜貫通領域 (TM) はコラーゲン非結合状態で二量体を形成している。

バンド非結合状態で膜貫通領域 (TM) が二量体を形成していることが報告されているため (Fig.1)、DDR2 の活性化機構の解明にはコラーゲン結合に伴う TM の構造情報を得ることが必要である。

本研究では、NMR 法により DDR2 とコラーゲンの複合体構造を解明し、システインスクランニング法により DDR2 の TM の二量体構造の解析を行った。

【結果と考察】

1.コラーゲン模倣ペプチドを用いたコラーゲン結合様式の解析

DDR2 はコラーゲンタイプ II の $^{394}\text{G}-^{405}\text{O}$ (O: 4-水酸化プロリン) 領域を認識することが報告されている。DDR2 の複合体構造解析には DDR2 結合配列を GPO の繰り返し配列で挟むことによりトリプルヘリックス構造を形成させたコラーゲン模倣ペプチド $(\text{GPO})_4$ - $^{394}\text{GPRGQOGVMGF}^{405}\text{O}-(\text{GPO})_4$ を用いた。SPR 法により DDR2-DS ドメインとの結合活性を解析したところ、野生型コラーゲンタイプ II と同程度の解離定数 $29 \cdot \text{M}$ を得た。また、TCS 法により同定した合成ペプチドに対する DS ドメインの結合界面は野生型コラーゲンタイプ II に対する結合界面とよく一致した。以上の結果から、DS ドメインはコラーゲン線維のうち一本のトリプルヘリックス鎖を認識することが判明した。そこで、ペプチドを用いて複合体の解析を以降行うこととした。

2. DS ドメインとコラーゲン複合体における距離情報の抽出

コラーゲンに対して DS ドメインが結合する配向を常磁性緩和促進 (PRE) 実験により決定した。DDR2 結合配列を含むコラーゲンペプチドの C 末端に、三量体構造を安定化させる foldon を付加したコンストラクト (FFCP: foldon fusion collagen peptide) を大腸菌にて調製後、Cys 残基へスピンラベル試薬 MTSL を導入した。スピンラベルなどのラジカル分子は近接する核スピンの緩和を促進し、周囲およそ 15\AA 以内に位置する原子の NMR シグナル強度を減少させる。DDR2 結合配列よりも N 末端側にスピンラベルを導入した場合にシグナル強度が減少した残基は、Fig. 2A 下側の領域に集中して存在した。一方、C 末端側にスピンラベルを導入した場合には、Fig. 2B 上側に存在する残基のシグナルが強度減少した。以上の結果から、DS ドメインの W52-G70 領域に対してコラーゲン

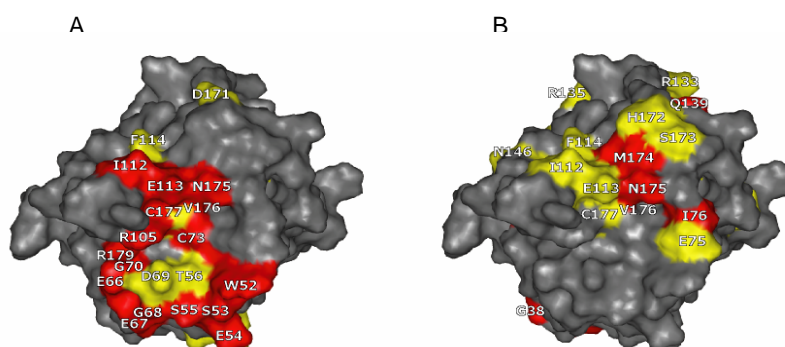


Fig. 2 A. N 末端側に MTSL 修飾した FFCP を用いた PRE 実験の結果。アスコルビン酸による還元前後のシグナル強度比に基づき DS ドメインの表面構造にマッピングした。赤: 強度比 0.74 以下, 黄: 0.74-0.8。B. C 末端側に MTSL 修飾した FFCP を用いた PRE 実験の結果。赤: 強度比 0.8 以下, 黄: 0.8-0.9。

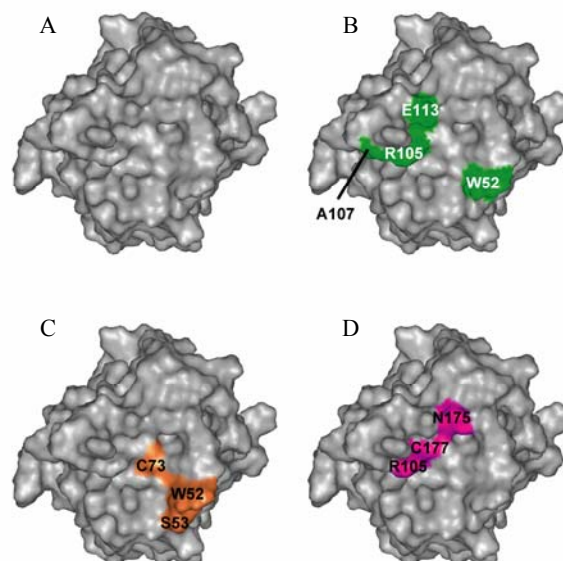


Fig. 3 Arg (A), Val (B), Met (C), Phe (D) 選択的 ^1H FFCP を用いた ASTCS 実験の結果。DS ドメインの表面構造に 0.2 以上のシグナル強度減少率を示した残基を色付けして示した。

が N 末端側を向けた配向で結合することが明らかとなった。

続いて、コラーゲンと DS ドメインの近接残基対をアミノ酸選択的転移交差飽和 (ASTCS) 法により決定した。ASTCS 法では、特定のアミノ酸のみに ^1H 標識を施した FFCP を用いて TCS 実験を行うことにより、当該アミノ酸からおよそ 5 Å 以内に存在する DS ドメイン上の残基を決定することが可能である。Arg, Val, Met, Phe についてアミノ酸選択的に ^1H 標識を施した FFCP を調製後、ASTCS 実験を行った。その結果、Arg 標識体については顕著な強度減少を示す残基は観測されなかったものの、Val 標識体では、W52 側鎖、A57、R105、R105 側鎖、A107、E113 が、Met 標識体では W52、W52 側鎖、S53、C73 が、Phe 標識体では R105 側鎖、N175 側鎖、C177 が 20% 以上顕著に強度減少した (Fig. 3)。 ^1H 標識したアミノ酸ごとに別々のコラーゲン結合残基のシグナルが強度減少したことから、アミノ酸選択的な距離情報を抽出できたと判断した。

3. DS ドメインとコラーゲンの複合体モデルの構築

次に、得られた実験結果に基づいて DS ドメインとコラーゲンの複合体モデル構造を中村チームによって開発されたドッキングプログラムにて作製した。DDR2 の 20 構造に対して、コラーゲンペプチドをドッキングさせ、各 300 の複合体構造を作成した。この 20x300 構造から ASTCS 実験の結果と定性的に合致する 119 構造を選び、疑似エネルギー項 E_{sat} を加えた分子動力学計算を行った。119 個の計算でサンプルされた構造のうち、 E_{sat} が小さく磁化変化 η が実験値とよく一致する構造を最終的な複合体構造とした。

得られた複合体構造において (Fig.4)、DDR2 結合への重要性が示されているコラーゲンの Met が、W52 と C73 で形成される疎水性領域と相互作用している。W52 に Ala 変異を導入すると、コラーゲンに対する結合親和性が 30 分の 1 に低下するため、W52 と Met の疎水性相互作用は重要であると考えた。同じく、相互作用に重要な Phe の側鎖が R105、C177 などの疎水性残基で形成されるくぼみに位置しており、嵩高で疎水性の芳香環が相互作用するのに適している。

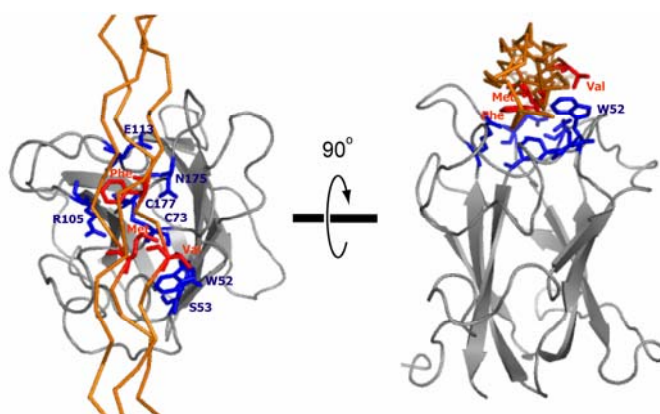


Fig.4 DS ドメインとコラーゲンペプチド複合体のモデル構造。相互作用に重要な残基について側鎖を表示した。

4. 膜貫通領域の二量体構造解析

膜貫通領域 (TM) の近接残基対は全長の DDR2 を用いて脂質二重膜中で解析する必要があるため、DDR2 を 293T 細胞に発現させ、システインスキャンニング法により解析した。システインスキャンニング実験では、DDR2 の特定の残基に Cys を導入し、SH 基を介したジスルフィド結合形成の有無をみることで、二量体中で当該残基が近接しているかを調べる。まず、

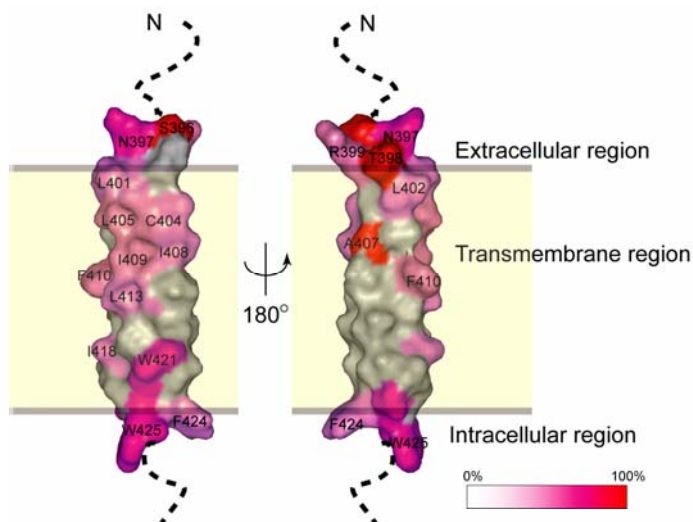


Fig. 5. DDR2 の TM とその近傍領域のシステインスキャンニング実験結果。 α -ヘリックスモデル構造上に、ジスルフィド形成効率が高い残基ほど濃い赤色にて示した。

DDR2 の内在性の遊離 Cys 残基に Ser 変異を導入した C288S/C404S 変異体に対して、TM とその近傍の残基について一残基ずつ Cys 変異を導入し、コラーゲン刺激によるリン酸化活性を確認した。コラーゲン非存在下におけるシステインスクヤニング実験の結果、細胞外領域 (P389-T398)、細胞内領域(Q423-W425) に関しては、K392C を除きすべての変異体がジスルフィド結合を形成した一方で、TM 領域の変異体は周期的にジスルフィド結合を形成した。これらの残基は TM ヘリックスモデル構造上で片側の面に局在したため、この面で二量体を形成すると考えた (Fig. 5)。一方、リガンド存在下において同様にシステインスクヤニング実験を行うと、TM 領域における周期性が消失した。この結果は、コラーゲン結合に伴って DDR2 の TM 領域に構造変化が生じていることを示し、コラーゲン非存在下において形成される TM 領域の二量体構造は、DDR2 の活性化を抑制していることが示唆される。

【総括】本研究では、NMR 法を用いた常磁性緩和促進実験およびアミノ酸選択的転移交差飽和実験により、コラーゲンと DDR2 の複合体構造を解明した。また、不活性状態における DDR2 膜貫通領域の二量体構造情報とコラーゲン結合に伴う構造変化を抽出することに成功した。今後、本研究において明らかとなった複合体構造に基づくコラーゲン結合阻害剤の設計や、TM 領域の二量体構造を安定化する薬剤の探索により、DDR2 活性化を阻害する抗腫瘍薬の開発へとつながっていくことを期待する。

【謝辞】

DDR2-DS ドメインとコラーゲンペプチド複合体のモデル構造の構築は、大阪大学・蛋白質研究所の中村春木教授、産総研・BIRC の金森英司研究員により行われた。

【参考文献】Ichikawa, O., Osawa, M., Nishida, N., Goshima, N., Nomura, N., Shimada, I. (2007) *EMBO J.*, **26**, 4168-4176.

2.3 研究開発項目③「高精度in silico スクリーニング等のシミュレーション技術」

集中研究: 社団法人バイオ産業情報化コンソーシアム[バイオメディシナル情報研究センター(BIRC)、大阪大学分室、情報数理研究所分室、塩野義製薬分室、アステラス製薬分室、三井化学アグロ分室]

共同研究先: 独立行政法人産業技術総合研究所バイオメディシナル情報研究センター、国立大学法人大阪大学蛋白研究所

【中期目標】

高精度のインシリコスクリーニングを実現するため、以下の技術を開発する。さらに、①、②の技術開発との連携により、産業上有用な化合物を5個以上取得する。

- 1) タンパク質の動的性質を正しく評価し、タンパク質受容体への基質結合能を高い精度で計算できる新しい計算科学手法の開発し、インシリコスクリーニングの効率を従来法に比べ5倍程にあげる。
- 2) タンパク質と化合物とのドッキング計算手法の精度を高め、ターゲット選択性能を従来法に比べ5倍程度上げる
- 3) タンパク質間相互作用を阻害・制御する低分子化合物の選択・設計の技術開発を行う。

【達成度】

高精度のインシリコスクリーニング技術を開発し、国際的な業界標準であるソフトウェアに対する優位性を確認した。その手法により、 μ オピオイド受容体アゴニスト(48ヶ)や農薬のシードとなる化合物(23ヶ)など70ヶを超えるヒット化合物を得た。このうち構造上、前者では少なくとも2ヶ、後者では16ヶについて新規性が高く産業上有用と考えられ、また、横浜市大との共同研究によるインフルエンザウイルスRNAポリメラーゼPA-PB1複合体阻害剤として3ヶの化合物の発見など、将来有用となるもの、有用になりうる化合物を、総計20個以上を得ることができた。

→ 目標を達成できた。

1) タンパク質のドッキング効率を上げる新たな手法や、分子動力学計算に基づく結合自由エネルギー計算の新規な方法を開発した。また、 μ オピオイド受容体に対する*in-silico*スクリーニング法によって200万化合物から選択された399化合物に対して、ヒット化合物48ヶ(12%)が見出され、また膜タンパク質に対する新規阻害剤が187ヶの評価化合物中23ヶ(12.3%)見出された。これらのヒット率は、通常の0.1%~1%のオーダーのヒット率に比較すると、少なくとも10倍以上のスクリーニング効率となっている。

→ 目標を達成できた。

2) データベースエンリッチメント・カーブにおいて上位1%における選択的なヒット率は、ssDSM法では通常のドッキング手法の5.1倍(45.8/8.9%)、標的タンパク質の動的モデルによるMTS法ではGOLDに比べ20倍(66/3.3)、機械学習法とMTS & DSIを組み合わせた方法ではS社のドッキングソフトに比べて18.0倍(22.1/1.2)を得た。GOLDも後者のソフトも国際的に広く利用されている標準的な

ッキング・ソフトウェアである。一方、不整脈等の原因となるhERGチャネル阻害活性予測を行う新規手法の開発においても、市販プログラムに搭載されている従来法の7.9倍の選択性を得た。

→ 目標を達成できた。

3) 類似化合物探索手法としてMD-MVO法を開発したところ、代表的なソフトウェアROCSより有意に高いヒット率を示し、ペプチド・リガンドからの低分子探索が可能であることを示した。

→ 目標を達成できた。

(1) in silicoドッキング計算の高精度化

【研究内容】

創薬プロセスにおけるin-silico ドッキング計算において、タンパク質の動的性質を正しく評価するため、動的性質を抽出する手法を開発するとともに、ドッキングスコアの精度を高め、タンパク質および低分子リガンドの動的構造、及び各種相互作用を考慮した複合体の構造予測法及び結合エネルギー算出法の開発を行う。

【研究成果】

(i) ドッキングスクリーニング法の改良: MTS 法の改良

[BIRC 集中研]

我々は今までに、蛋白質—化合物ドッキングソフト、複数の in silico スクリーニング手法を開発してきた。これらスクリーニング手法は、蛋白質—化合物相互作用行列に基づく手法であり、標的蛋白質の3D 構造をもとにした Multiple target screening method (MTS 法)、標的の構造を用いずに、活性化合物の情報のみを用いる Docking score index method (DSI 法) である。ドッキングスコアの精度そのものを改善する手法も提案してきた。1つは、蛋白質—化合物相互作用行列に基づく手法である Direct score modification method (DSM 法)、もうひとつは、活性化合物の情報を用いる Machine-learning score modification method (MSM 法) である。

標的蛋白質の3D 構造が既知であり、かつ活性化合物の情報もあれば、Machine-learning score modification method と MTS 法を組み合わせることで、安定して高いヒット率を出すことができる (MSM-MTS 法)。しかし、活性化合物の情報がない場合、Direct score modification method と MTS 法を組み合わせることになるが (DSM-MTS 法)、標的蛋白質の種類によってヒット率にばらつきがあり、またヒット率は、MSM-MTS 法に比べてかなり低いのが問題だった。

【sequence-based DSM 法の開発】

本年度では、DSM-MTS 法の欠点を修正し、標的蛋白質の3D 構造が既知である場合、従来より高いヒット率を、標的蛋白質の種類に関わらずに安定して高いヒット率を出すことを目指す。スコアの修正を下記の式で行う。つまり蛋白質 a、化合物 i のドッキングスコア $s_a^{\text{mod } i}$ は、蛋白質 b、化合物 i のドッキングスコア s_b^i の線形結合で近似する。 M_a^b は、パラメーターである。

$$s_a^{\text{mod } i} = \sum_b s_b^i M_a^b \quad \text{式(1)}$$

DSM 法では、蛋白質—化合物相互作用行列に基づいて蛋白質 a-b の相関係数 R_a^b を計算し、これを M_a^b として用いる。

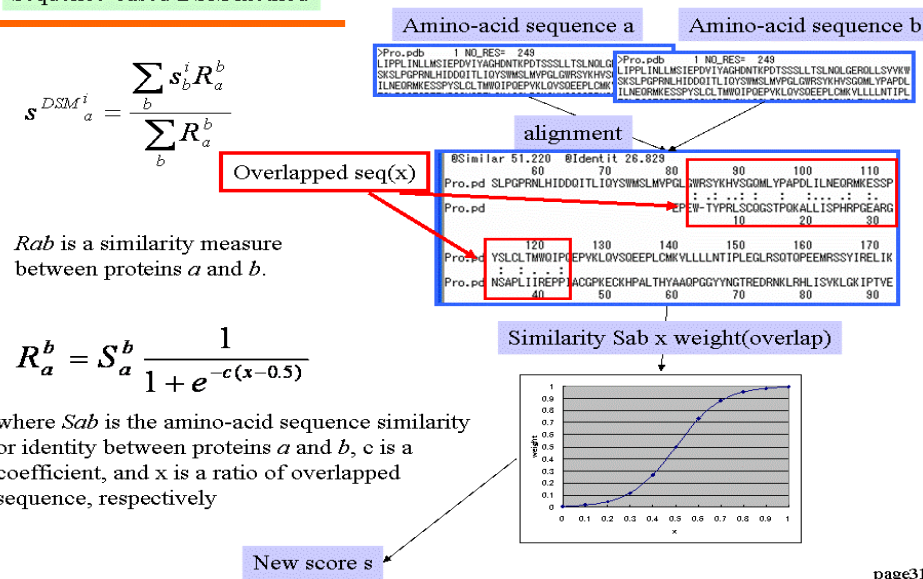
$$s^{DSM^i}_a = \frac{\sum_b s_b^i R_a^b}{\sum_b R_a^b} \quad \text{式(2)}$$

我々は今年度、新たに sequence-based score modification を考案し、 M_a^b を蛋白質 a-b のアミノ酸配列相同性でみるもることにした。

$$R_a^b = S_a^b \frac{1}{1 + e^{-c(x-0.5)}} \quad \text{式(3)}$$

に置く。 S_a^b は、fasta version 3.4 での相同性(%)、 c はパラメーターで、 x は、(a-b の重複する配列の長さ)/(標的蛋白質 a の全長)である。Fasta34 は、部分配列の相同性を計算するようになっている。そのため、数残基(5残基以下)で100%相同性といった部分配列を検出してしまふ。相同性がやや低くても、全長で50-60%の相同性があれば同じ構造の蛋白質である。したがって、fasta34 使用時には、蛋白質の相同性を考えるとき、相同性と重複配列の長さの両方を考える必要がある。そこで、配列相同性だけでなく、アミノ酸配列の相同性が見られる領域の長さも考慮するように式を構成した(分母の $1 + e^{-c(x-0.5)}$ の部分)。

Sequence-based DSM method

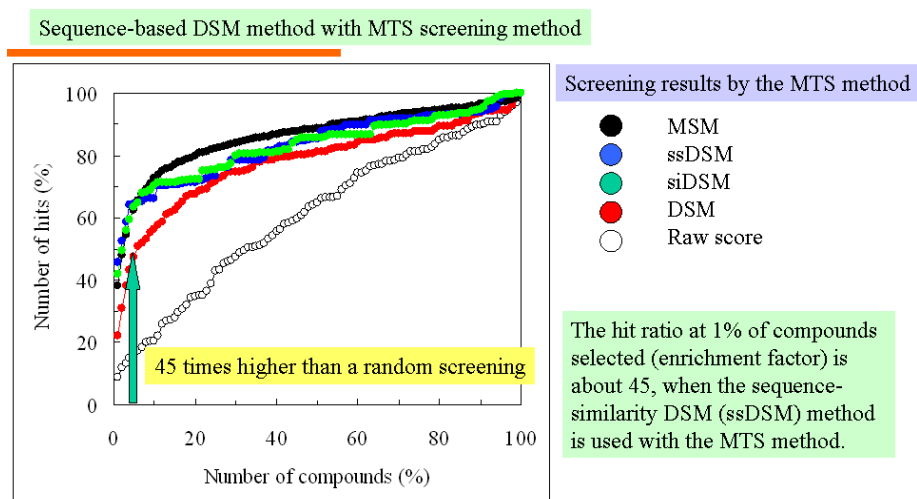


図は、この手法の計算手順を模式図としたものである。

上式において、 S_a^b をたんぱく質 a,b の similarity で置き換えたものを sequence-similarity DSM (ssDSM), identity で置き換えたものを sequence-identity DSM (siDSM)と呼ぶことにした。

ssDSM/siDSM を既存のオリジナルの MTS,DSM-MTS, MSM-MTS と比較してみた。標的蛋白質として、COX-2, サーマライシン, MIS, グルタチオン-S-転移酵素, HIV プロテアーゼ1の5種を選び、これらの活性化化合物を文献より収集、さらに、約1万化合物をランダムライブラリー

より不活性化化合物として収集した。180 蛋白質x約1万化合物の相互作用行列を SievGene を用いて作成し、ssDSM/siDSM、オリジナルの MTS,DSM-MTS, MSM-MTS を適用して、データベースエンリッチメント曲線を計算したのが下図である。MSM-MTS 法は、既知活性化化合物が発見される確率を最大化するようにパラメータを調整する手法なので、事実上の上限を与えている。ssDSM/siDSM は、オリジナルの MTS、DSM-MTS 法の結果を大きく上回り、MSM-MTS 法に近い結果を与えた。上位1%の化合物を採択した場合のヒット率は45%に達し、ヒット率を向上させる初期の目的は果たされている。



ただし、ssDSM/siDSM は、通常の MTS 法でヒット率が高かった標的に対し、より高いヒット率を示す反面、通常の MTS 法でヒット率が低かった標的に対しては、より低いヒット率を与えてしまう。平均すれば、ヒット率を完全しているが、標的による結果のばらつきが大きくなるのが欠点であることが分かったので、適用に際しては注意が必要である。

(ii) 膜蛋白質の分子動力学計算による動的性質の抽出

〔阪大分室〕

創薬プロセスにおけるin-silico ドッキング計算において、タンパク質の動的性質を正しく評価し動的性質を抽出するため、分子動力学計算結果における蛋白質分子の各原子のトラジェクトリーから、主成分解析(PCA: Principal Components Analysis)を行って動的構造を解析する手法を開発・確立した。また、主鎖原子間をバネでつなぐモデルである弾性ネットワーク・モデル(ENM: Elastic Network Model)についても、その解析法を開発し、上記PCA解析と同様に利用できる仕組みとした。

まず、膜蛋白質であるモノアミン酸化酵素-A(MAO-A)の膜分子と水分子をあわらに取り扱った計81,647原子からなる系(水分子:17125, POPC

(1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) リン脂質分子: 163, Na⁺: 76, Cl⁻: 78)に対し、myPresto/cosgeneを用いてNPTアンサンブル(Nosé-Anderson法)による1気圧、310Kでの分子動力学計算を、1 step = 2 fs として、20 nsの長さで3回初期速度を変えて行った(周期境界条件: 72 Å x 94 Å x 124 Å)。同様に、膜貫通ヘリックスを削除し、膜分子のない水溶液系(計

63,605原子, 周期境界条件: 72 Å x 92 Å x 96 Å)も同時に3回の分子動力学計算を行い、トータル6ヶのトラジェクトリーを比較・解析した。

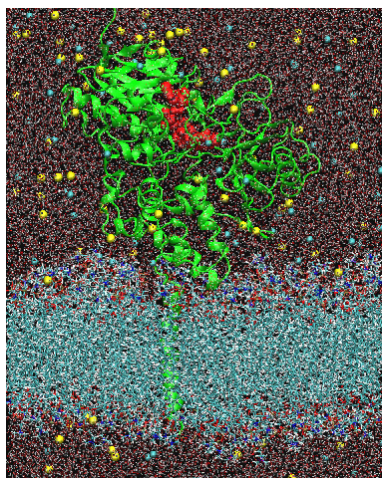


図1: 膜蛋白質 MAO-A に対するシミュレーション計算の系。MAO-A (緑のリボン) と、膜分子 (POPC:水色)、溶媒の水 (赤)、Na⁺イオン (黄色)、Cl⁻イオン (水色の球)。

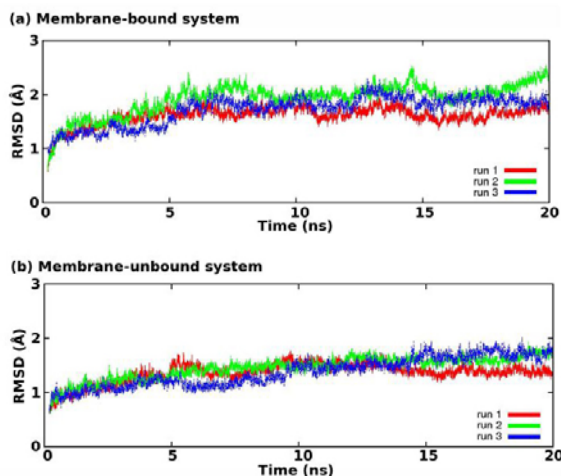


図2: (a)膜貫通ヘリックスがある場合のトラジェクトリにおける初期構造からのズレ。(b)水溶液系のシステムにおけるトラジェクトリ。

この時、特に、MAO-Aの内部のドメイン構造の動的性質を明らかにするため、ゆらぎの相関を

$$Corr_{ij} = \frac{\langle (\vec{r}_i - \langle \vec{r}_i \rangle) \cdot (\vec{r}_j - \langle \vec{r}_j \rangle) \rangle}{\sqrt{\langle (\vec{r}_i - \langle \vec{r}_i \rangle)^2 \rangle \langle (\vec{r}_j - \langle \vec{r}_j \rangle)^2 \rangle}} \quad (\vec{r}_i: i \text{ 番目の残基の } C_{\alpha} \text{ 原子の座標})$$

と定義して、全ての60 ns ずつのトラジェクトリを計算し、膜貫通ヘリックスが膜にアンカーしている状態のもの(左上)と、膜貫通ヘリックスを含まず水溶液中のもの(右下)とを比較した。

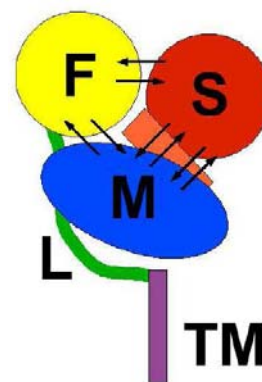
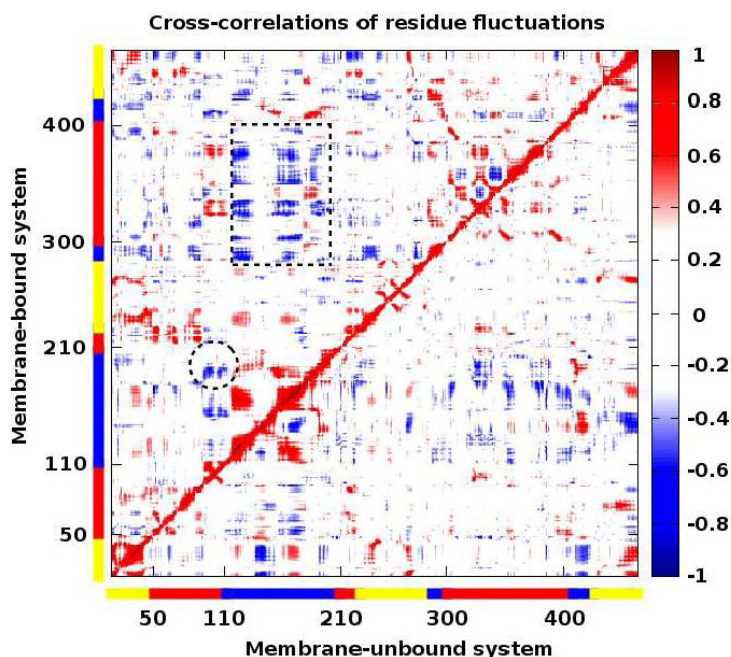


図3: i 番目の残基と j 番目の残基お C_{α} 原子間のゆらぎの相関 $Corr_{ij}$ 。左上は膜貫通ヘリックスを含む系であり、右下は水溶液系。残基部分の色は、右の概略図におけるドメインに対応する。オレンジ色はリガンド結合部位。

明らかに、点線で囲む負の相関(青色で示される領域)が、膜貫通ヘリックスを持つ場合には明瞭であるのに対し、水溶液中のトラジェクトリーには見られないのがわかる。これらは、図3で示すSDメインとMDメイン間の負の相関であり、リガンド結合部位の開閉が、膜貫通ヘリックスがある場合には拡大されることがわかる。

さらに、どのような運動かを詳細にみるため、主成分解析とENM解析とを比較した(図4)。

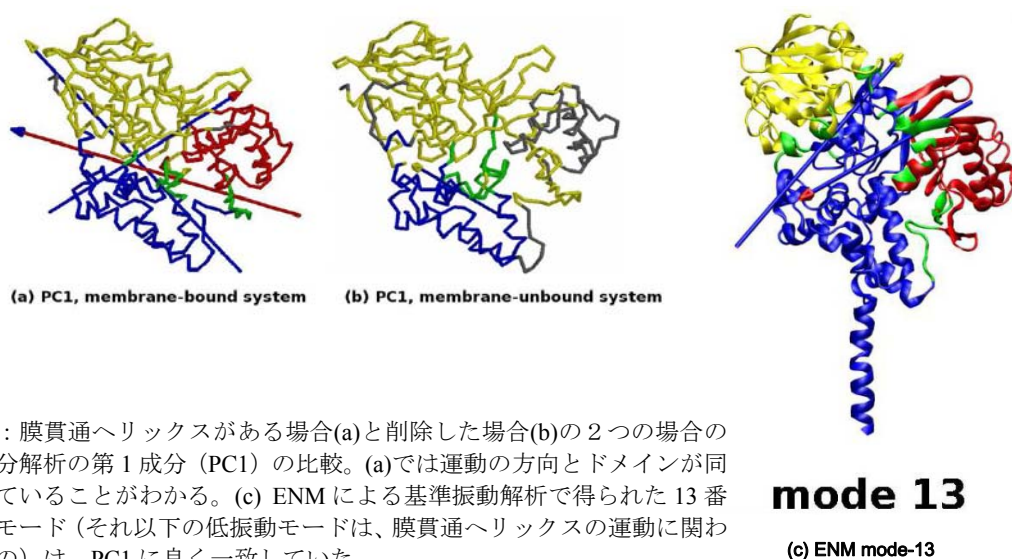


図4：膜貫通ヘリックスがある場合(a)と削除した場合(b)の2つの場合の主成分解析の第1成分(PC1)の比較。(a)では運動の方向とドメインが同期していることがわかる。(c) ENMによる基準振動解析で得られた13番目のモード(それ以下の低振動モードは、膜貫通ヘリックスの運動に関わるもの)は、PC1に良く一致していた。

ENM解析で得られる基準振動で、主成分解析で得られる主な運動が再現されていることがわかり、分子動力学計算と主成分解析によって得られる振動モードは、この蛋白質分子がもともと持っている運動であることがわかる。すなわち、リガンド結合に関係すると思われるドメインの開閉運動は、もともとのMAO-A分子が持っている振動モードであったが、膜貫通ヘリックスが膜にアンカーリングすることによって、テコのように働き、このドメイン間の開閉運動がさらに強調されたものと考えられる。これから、図5のようなリガンドの入出に伴う道筋(トンネル)が推測される。特に右側のピンク色のトンネルについては、リガンド選択性をつかさどるアミノ酸がそのトンネル中に位置することから、この推測の正当性が示唆されている。

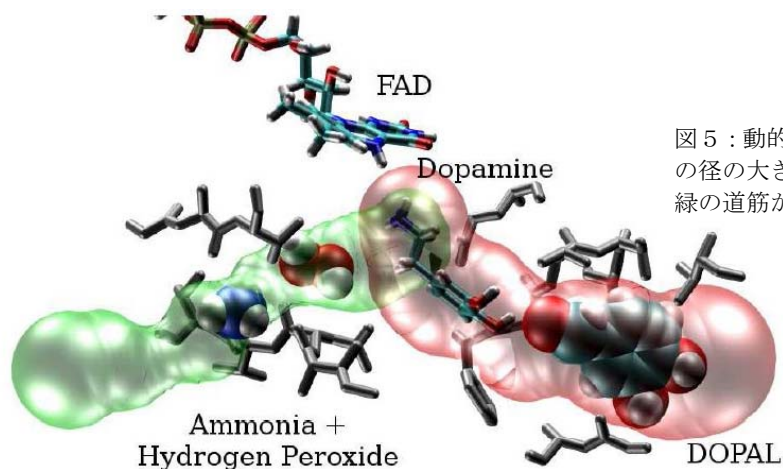


図5：動的なリガンド結合の道筋。トンネルの径の大きなピンクの道筋と、若干小さい緑の道筋が可能なものとして考えられる。

以上、MAO-Aの静止状態の構造からはリガンド結合の道筋が不明であったのが、このようなシミ

ュレーション計算によって、ダイナミックなリガンド結合の仕組みを明らかにすることができた。
(Apostolov et al. (2009) *Biochemistry*, 48, 5864–5873)

(iii) 新しいリガンド結合エネルギー計算手法 (Smooth Reaction Path Generation 法) の開発

[BIRC 集中研]

医薬分子の標的への結合活性は、結合自由エネルギーで測ることができる。医薬品の最高活性は、結合自由エネルギーで最大 15 kcal/mol 程度であるとされている。通常、医薬品の平均活性は $IC_{50} = 30$ nM であり、70% の化合物が IC_{50} が 50 nM より強い活性をしめす。 IC_{50} が nano-, pico-M である場合、結合自由エネルギーは 12.6 – 16.8 kcal/mol となる。

結合自由エネルギー (ΔG) の計算は、計算化学の究極の目的の一つだが、計算は難しい。蛋白質—化合物ドッキングソフトによる見積もりでは、様々なソフトが開発されてきたものの、実測 ΔG と計算値の相関は、相関係数 0.6~0.7、平均誤差 2.5 kcal/mol となる。この誤差は、ヒット化合物 (シードレベル) の ΔG に相当するエネルギーで、その結果、計算によるヒット化合物探索の精度は低い。MMPBSA、MMGBSA、LIE (Linear Interaction Energy)、COMBINE 法などいろいろな半経験的 ΔG 推算方法が開発されてきたが、精度がドッキングソフトに比べて飛躍的に高いわけでもなく、既知活性化合物の ΔG がないと推算ができなかったりする欠点があった。 ΔG の推算誤差を 1 kcal/mol 程度にすることができれば、活性の推算は、従来の 10 倍程度に改善されたといえるので、この値を最終目標とする。なお、実験での pH、塩濃度によって活性値は 10 倍程度変化するので、誤差 1kcal/mol 以上の精度は、期待できない。

分子動力学シミュレーションで ΔG が計算できれば、既知の ΔG を必要とせず、理想的であるが、その計算は、極めて計算量が多く、困難である。我々は、2003 年に filling potential (FP) 法を開発した。この手法では、蛋白質—化合物複合体から、人為的なポテンシャルを用いて、化合物を解離させ、その経路に沿った自由エネルギー面を計算することで、少ない計算時間で ΔG を求める方法である。この FP 法を用いた全原子シミュレーションによって、酵素—阻害剤の ΔG の高精度の計算に成功した。しかし、FP 法は、化合物が水中ではなかなか解離しない、ヒストグラム法という計算方法を使うため、慎重な計算が必要で、計算 5 回~10 回に 1 回しか成功しないという扱いの難しい方法だった。そこで、我々は、より簡便に、高い確率で ΔG が計算できる方法を開発することにした。

開発した計算手法: Smooth Reaction Path Generation (SRPG) 法

蛋白質と化合物の複合体構造を最初に準備し、蛋白質から化合物を、真空中で簡単に解離させ、その解離の経路をルジャンドル多項式を用いて、滑らかな経路で近似する。この滑らかな経路に沿って、化合物が様々な位置に存在する状態を作り、タンパク質、水、イオンと合わせて系を作り、各系で MD を行い、リガンドの対象原子に働く平均力を計算する。滑らかな経路に沿って、力を積分することで自由エネルギー面を計算し、結合状態と解離状態の存在確率を求めて、その比率から ΔG を計算する。各状態をパラレルに計算することができることで計算時間を大き

く短縮させることができるという利点がある。

SRPG 法による ΔG 計算では、

$$\Delta G = -k_B T \ln \frac{P_B}{P_U} \quad \text{eq. 1}$$

結合状態の規格化されていない確率 P_B と非結合状態の規格化されていない確率 P_U を用いて ΔG を計算する。このとき、

$$P_B = \int_{R_B} \exp(-\beta G(r)) dr \quad \text{eq. 2}$$

$$P_U = \int_{R_U} \exp(-\beta G(r)) dr \quad \text{eq. 3}$$

ここで $G(r)$, \cdot , R_B , and R_U は、座標 r での自由エネルギー(PMF), $1/k_B T$, リガンドの結合状態の領域、リガンドの非結合状態の領域である。

もし、結合状態の確率 P_B が、PMFの最安定座標の周辺(r_0),に限定され、 $G(r)$ がパラボラポテンシャルで式4のように近似されるなら、

$$G(\vec{r}) = G(\vec{r}_0) + \left(\frac{k_x}{2} \Delta x^2 + \frac{k_y}{2} \Delta y^2 + \frac{k_z}{2} \Delta z^2 \right) \quad \text{eq. 4}$$

ここで

$$\vec{r} = \vec{r}_0 + \vec{\Delta r}, \quad \vec{\Delta r} = (\Delta x, \Delta y, \Delta z) \quad \text{eq. 5}$$

結合状態の確率 P_B は、式6で与えられる。

$$P_B = \int_{V_B} \exp(-\beta(G(r_0) + \frac{k_x}{2} x^2 + \frac{k_y}{2} y^2 + \frac{k_z}{2} z^2)) dx dy dz \quad \text{eq. 6}$$

PMF の原点を解離状態に置く($G(\infty) = 0$). もし $G(r)$ の値が、0 kcal/mol に近いなら、

$$P_U = \int_0^R 4\pi r^2 \exp(-\beta G(r_\infty)) dr = \frac{4\pi}{3} R^3 \exp(-\beta G(r_\infty)) = V_0 \exp(-\beta G(r_\infty)) \quad \text{eq. 7}$$

化合物の密度が 1M density だと、1 個の化合物分子の占める体積 V_0 は 1661 \AA^3 となる。

以上、まとめると、 ΔG は式8で与えられる。

$$\Delta G = G(r_0) - G(r_\infty) - k_B T \ln \frac{\int_{-\infty}^{\infty} \exp(-\beta \frac{k_x}{2} x^2) dx \int_{-\infty}^{\infty} \exp(-\beta \frac{k_y}{2} y^2) dy \int_{-\infty}^{\infty} \exp(-\beta \frac{k_z}{2} z^2) dz}{V_0} \quad \text{eq. 8}$$

式8の第3項の分子は、ガウス積分なので、解析的に解くことができる。

$$\Delta G = G(r_0) - G(r_\infty) - k_B T \ln \frac{\sqrt{\pi}}{\sqrt{\beta k_x / 2}} \frac{\sqrt{\pi}}{\sqrt{\beta k_y / 2}} \frac{\sqrt{\pi}}{\sqrt{\beta k_z / 2}} \quad \text{eq.9}$$

1661

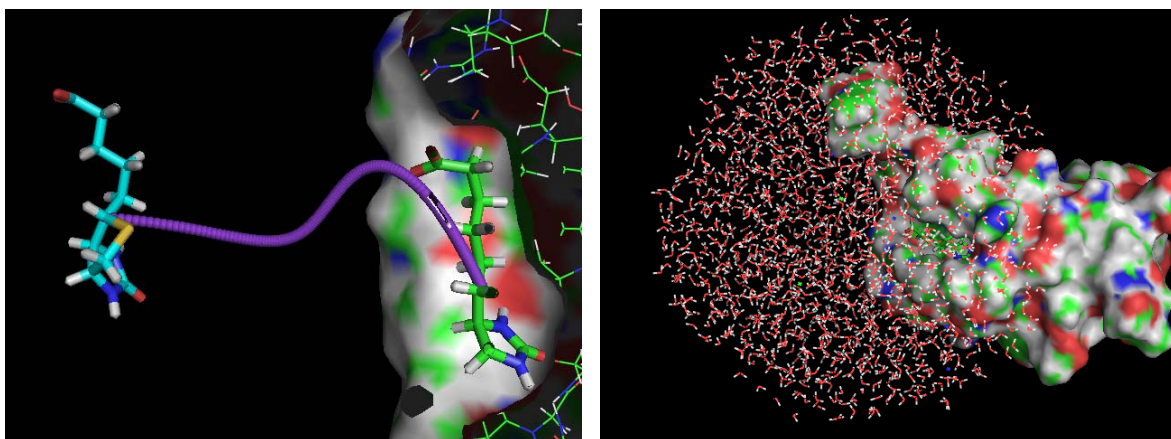
解離の経路 $p(t)$ は、ルジャンドル多項式として、

$$\begin{cases} p_x(t) = \sum_{i=0}^L c_x^i P_i(t) \\ p_y(t) = \sum_{i=0}^L c_y^i P_i(t) \\ p_z(t) = \sum_{i=0}^L c_z^i P_i(t) \end{cases} \quad \text{eq. 10}$$

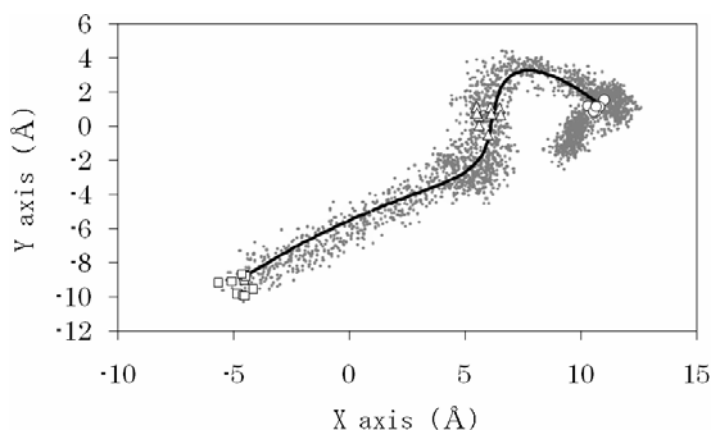
FP法で解離させた経路 p^0 に対して、 S を最小化するように、係数 c をモンテカルロ計算で最適化するように決定した。

$$S = \sum_m^M \sum_n^N D(\vec{p}(n/N), \vec{p}^0(m))^2 \quad \text{eq. 11}$$

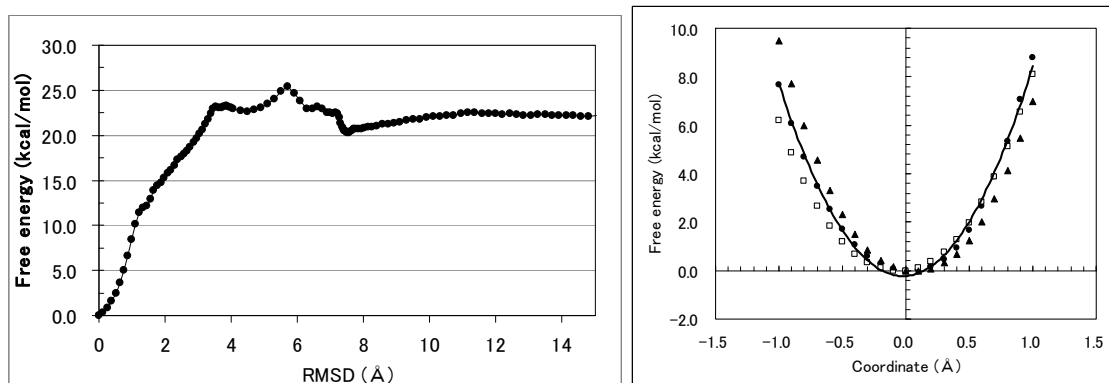
蛋白質—化合物の ΔG の計算: ストレプトアビジンとビオチンの系について計算を行った。



図にストレプトアビジンからビオチンが解離していく経路を示す。上図左のように、真空中で解離させ、右のように、後から溶媒水を加えて、経路にそって、150 個の系を作成した。



FP 法によって、生成した解離していくリガンドの座標を点で表し、それに対してルジャンドル多項式で生成した滑らかな経路を実線で示す。



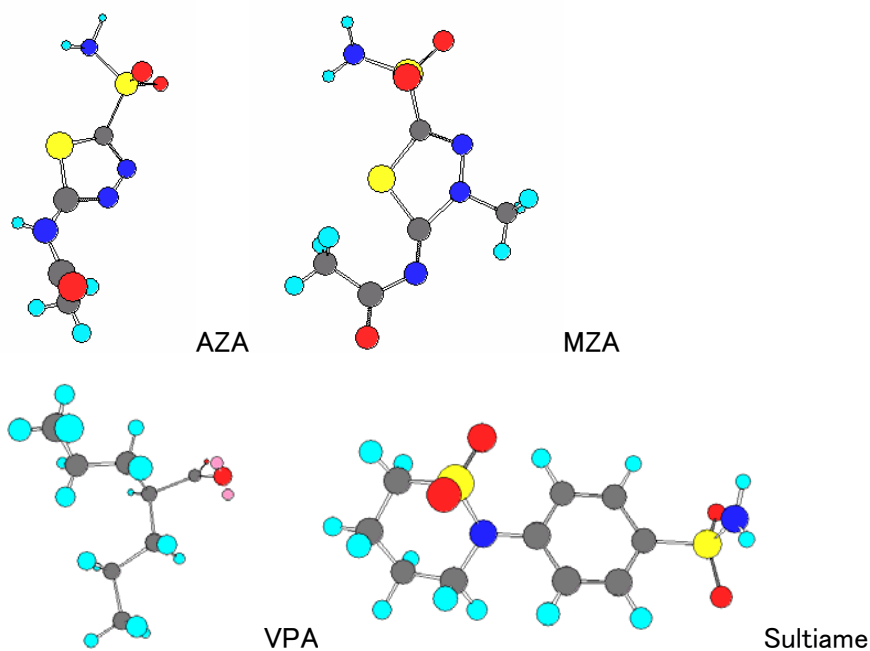
この経路に沿った自由エネルギー面は上左図のように得られた。リガンドの結合状態周辺の自由エネルギー面は、上右図のようであり、2次曲線で近似でき、理論が適用できることが示された。

	SRPG 法	実験値
$\Delta G(\text{kcal/mol})$	-16.5	-18.27

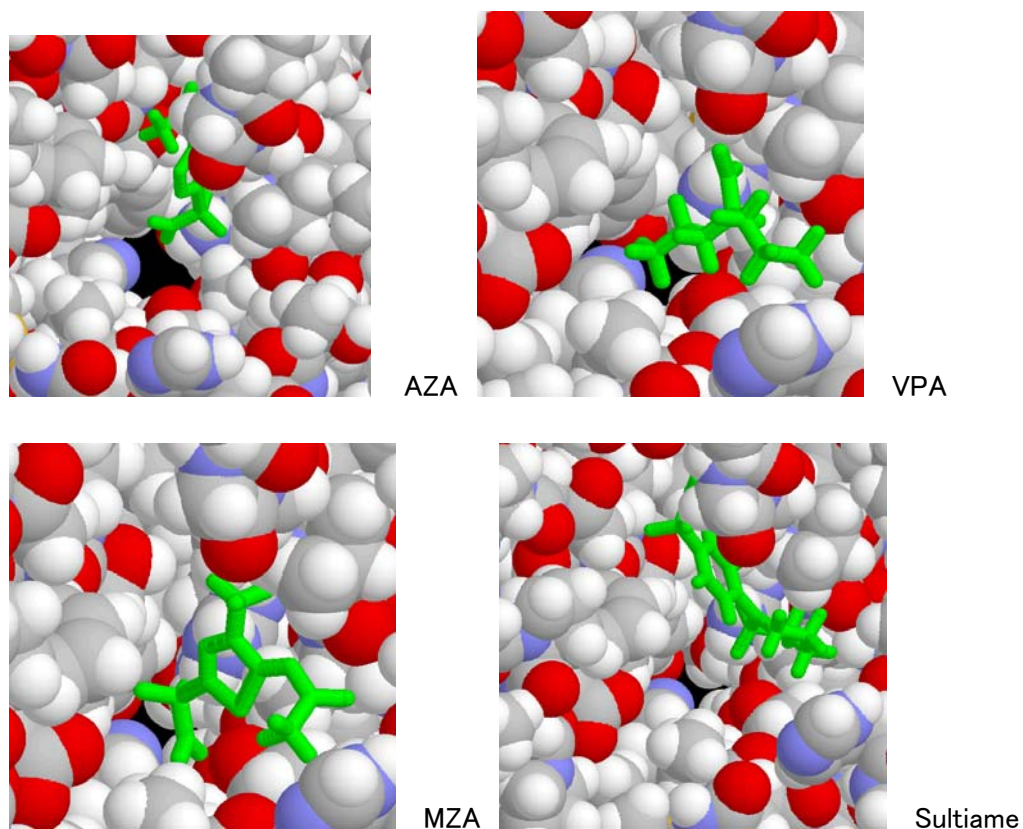
こうしてSRPG法では、誤差1.8kcal/mol(9.8%)という高い精度で、 ΔG を計算することができた。一例に過ぎないが、ドッキングソフトでの ΔG 平均誤差を1kcal/mol近く上回り、活性値で10倍の精度に達している、ともいえる。

アクアポーリン4(AQP4)への応用

藤吉チームで研究している AQP4 阻害剤探索に対して、SRPG法を適用した。試した化合物候補は、藤吉チームと同じ以下、4化合物である。



Sievgene でこれら4化合物を AQP4 へドッキングした。ドッキングポーズは、以下のものである。

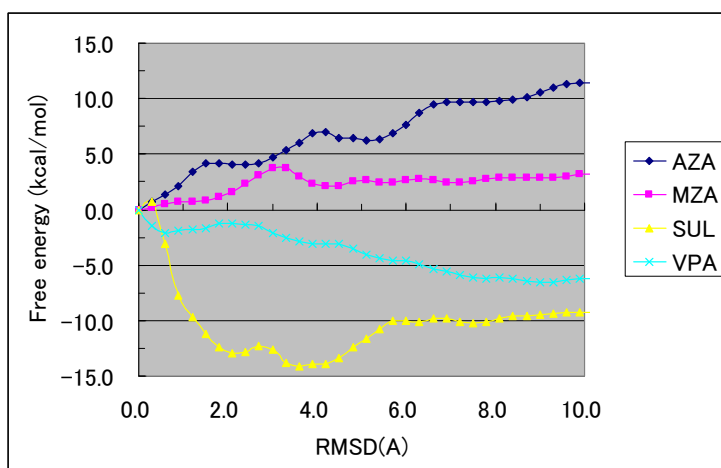


水中でのMD

SRPG 法を適用した。2 nsec の水中でのMDを実施した。膜は作成せず、通常の球状蛋白質と同

じようにCAP水に入れて計算した。各化合物を(2)でドッキングした位置から、10 Å 引き離れた位置まで解離させることにする。0.3 Å 刻みに解離経路上に化合物を配置し、化合物に加わる力を計算することで、自由エネルギー一面を計算する。

計算された自由エネルギー一面



自由エネルギー一面

自由エネルギー一面での結合状態と解離状態の自由エネルギー差から約 5kcal/mol を差し引いた値が ΔG に相当する。AZA は、結合しやすいと推測される。AZA と MZA は、分子構造ではメチル基1つの違いしかないが、自由エネルギー一面には大きな違いが現れた。Sultiamine, VPA は、結合状態が不安定で、結合しえない可能性が高い。この傾向は、実験と一致することが示された。

(iv) McMD (マルチカノニカル分子動力学法) 計算法による蛋白質-阻害剤の結合自由エネルギー計算 [阪大分室]

結合自由エネルギーを高い精度で算出するため、溶媒をあらわな水分子モデルで扱った系における、長時間のマルチカノニカル計算によるドッキング・シミュレーション手法を開発した。この手法により、リゾチーム蛋白質とその阻害分子 (NAG (N-acetyl-D-glucosamin) trimer) との結合についての、ab-initio 的なシミュレーションを行った。643 の TIP3P 水分子を CAP 境界として部分的に付加し、マルチカノニカル分子動力学 (McMD) 計算法によって、700K から 300K までをカバーするアンサンブルを得た (図6)。この時、蛋白質側は距離拘束条件を付加することによって、高い温度でも大きな構造変化がおきないように工夫した (図7)。

最終的に 72 ns (7.2×10^7 steps) の McMD ランを行い、その後、300K、500K、700K のカノニカル・アンサンブルを再構築し、阻害剤 (NAG)₃ の中心位置の分布をプロットした (図8)。面白いことに、図8の (a) 700K や (b) 500K では、阻害剤の位置は蛋白質から離れているのに対し、(c) 低温 (300K) になると蛋白質に近い所の存在確率が大きくなっている。

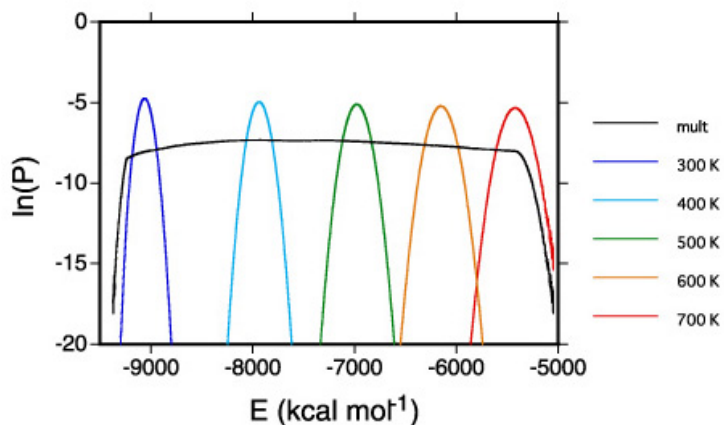
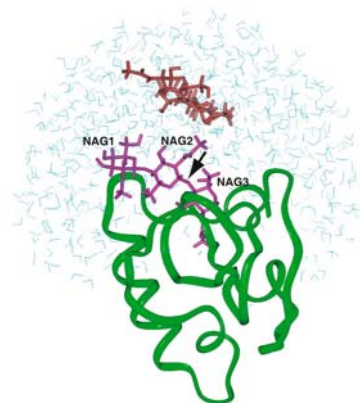


図6：MeMD アンサンブルと再構築したカノニカル・アンサンブルのエネルギー分布。



Kamiya et al. Figure 1

図7：対象としたリゾチームとその阻害剤 ((NAG)₃) および CAP の水分子。阻害剤は、初期の位置を赤で示し、紫は結晶構造。

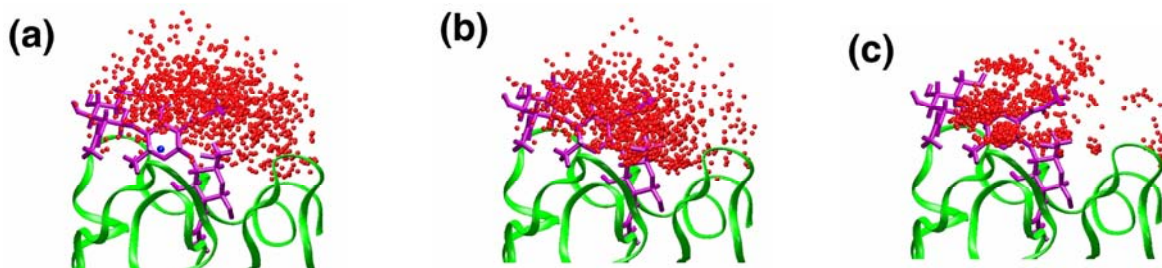


図8：(a) 700K, (b) 500K, (c) 300K として再構築した系のカノニカル・アンサンブル中の、リゾチーム (緑のリボンで示される主鎖構造) に対する阻害剤の重心の位置の分布。紫の棒モデルは、この阻害剤のリゾチームとの複合体の X 線結晶構造中の構造を参照構造として示している。

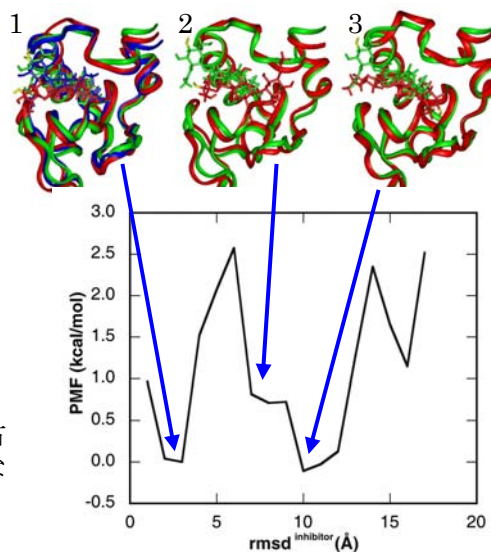
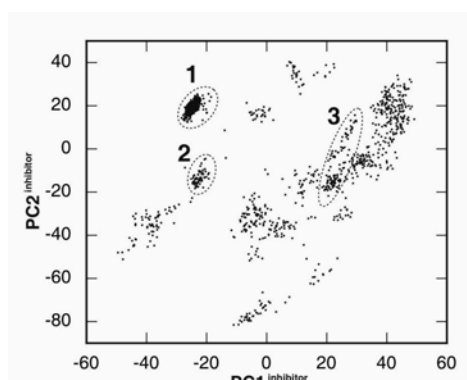


図9：自由エネルギー地形 (左上と右下) と、リゾチームと阻害剤(NAG)₃との複合体構造(右上) (緑：X 線結晶構造, 青と赤とは典型的なクラスター1, 2, 3 中の構造)

このカノニカル・アンサンブルから、阻害剤の座標に対する主成分解析を行って、自由エネルギー地形を描いた(図9)。複数の結合様式の中で、天然に観測される結合状態が安定となることが理解された。(Kamiya et al. (2008) *PROTEINS*, 70, 41-53)

(2) 構造生理学アプローチによるタンパク質間相互作用解析

【研究内容】

タンパク質間相互作用及び超分子複合体の構造情報まで含めた詳細な解析「構造インタラクーム」に踏み込み、タンパク質間相互作用の阻害等の創薬において有用な機能を有するものの活性の維持等の観点から医薬品化が困難な生理活性ペプチドから、医薬品となりやすい非ペプチド性の低分子化合物等へ展開するため、ペプチドと同様あるいはそれ以上の強い結合性を有する非ペプチド性化合物(低分子化合物等)を探索・設計する新しい手法の開発に着手する。

【研究成果】

(i) タンパク質複合体構造予測法の開発 [BIRC集中研、阪大分室]

タンパク質複合体構造予測法の開発を行い、複合体構造予測のブラインドコンテストであるCAPRI(Critical Assessment of Predicted Interactions)に参加してその性能を評価した。

構造プロテオミクスの発展に伴い、多くのタンパク質の立体構造が公開されるようになったが、それらからタンパク質の複合体構造を予測することは、タンパク質の機能を理解し制御するために重要であるが、尚困難な課題の一つである。

タンパク質複合体構造の予測においては、1)タンパク質間相互作用部位の予測、2)候補構造の発生と評価、が重要なポイントである。我々は進化系統樹を利用した機能部位予測法であるET(Evolutionary Trace)法とタンパク質形状の相補性を利用した独自の構造探索エンジン・構造評価関数を組み合わせた手法を開発した。

この手法をCAPRIの問題に適用したところ、いくつかの例で、高精度の複合体構造をサンプルすることができた。

下図は予測結果の一例である、ARF-1とARF-BDの複合体の予測(T25)では、44番目の候補として、結晶構造とのRMSD=2.59 Åの構造が得られた。また、ToIBとPalの複合体の予測(T26)では14番目の候補として、結晶構造とのRMSD=2.55 Åの構造を得ることができた。

今後は相互作用部位予測法と評価関数の改良により、さらに高精度の予測を目指す。

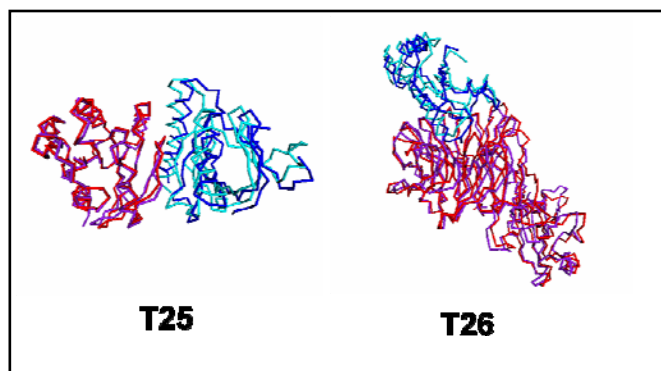


図) 予測結果の例

結晶構造のサブユニットを赤及び青、予測構造を紫及び薄青で示す。

(ii) アミノ酸選択的交差飽和法の実験データを用いたタンパク質複合体構造の構築

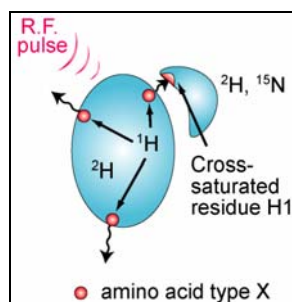
タンパク質-タンパク質間の相互作用は、シグナルの伝達など生体内で重要な役割を果たす。近年、タンパク質複合体の形成を妨げることで効果を発揮する医薬品の研究開発にも重点が置かれるようになってきている。

このような医薬品の開発において、タンパク質の複合体の立体構造を知ることは、薬剤を結合させる部位を決定するために重要であるが、X線結晶構造解析等の手法を用いて実験的に構造を決定することは、単体の状態の立体構造を決定するよりもさらに困難であることが知られている。

そこで今回、嶋田グループが開発した、NMRによる実験手法である、アミノ酸選択的交差飽和法(ASCS法)の実験データと、分子動力学計算とを組み合わせることにより、それぞれのタンパク質の単体の構造が既知の場合に、それらが形成するタンパク質複合体の立体構造を高精度に構築する方法を開発した。

方法

ASCS法は、一方のタンパク質の特定のアミノ酸種を標識して(交差飽和ドナー)、他方のタンパク質での交差飽和(η)を観測する実験手法で(下図)、強い交差飽和が観測された残基の近傍に、標識したアミノ酸種が存在するという情報が得られる。



複数種のアミノ酸を標識したデータを組み合わせることで、ドナー側の残基とアクセプター側の残基

の位置関係についてある程度特定はできるが、原子レベルでの立体構造を構築することはできない。

そこで、この実験データに分子動力学計算を組み合わせることで複合体構造を構築することを試みた。実験データを再現するほどエネルギーが低下する、次の疑似エネルギー項を加えた分子動力学計算を行う。

$$E_{sat} = \frac{1}{2} W_{sat} \sum_{all\ exp.} \sum_{acceptor\ residues} (\eta_{calc} - \eta_{exp.})^2$$

ここで W_{sat} : 重み、 η_{calc} : 計算値、 $\eta_{exp.}$: 実験値である。

結果

(a) Ubiquitin(Ub)とYUHの複合体構造構築における本手法の検証

計算には嶋田グループで実施した5種類のアミノ酸を標識した実験データを用いた。

当該タンパク質複合体の立体構造は、PDBコード:1cmxとして結晶構造が知られているものであり、UbとYUHを分離した状態から、本方法を用いて結晶構造と同様の構造を得ることを試みた。

初期構造

交差飽和が観測された残基と交差飽和源の残基が近接するという条件から、初期構造(2つのタンパク質の相対位置)を作成した。可能性の高い42構造を初期構造として各初期構造につき2例ずつ(初期速度の乱数が異なる)分子動力学計算を実行した。

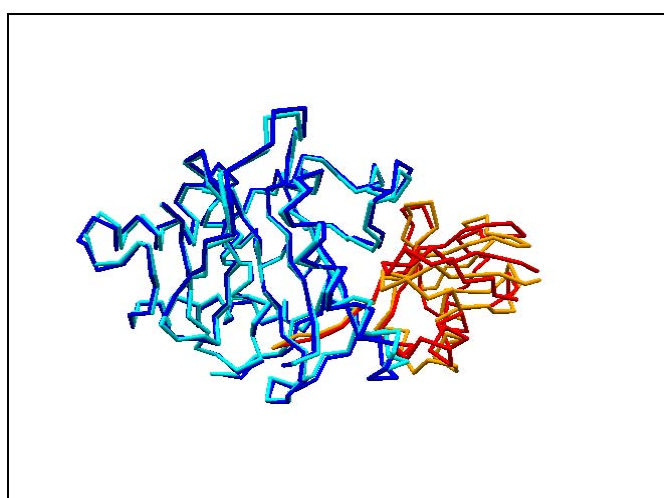
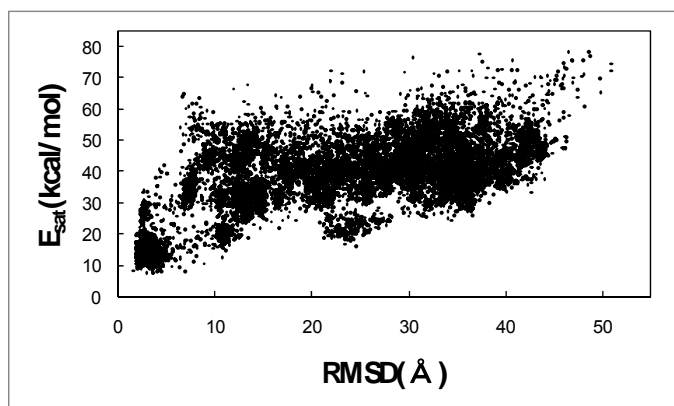
計算条件

分子動力学の計算条件を下表のとおりとし、真空中で構造を維持するためにUbとYUHの双方に、原子間の距離拘束を適用した。

time step	1.0 fsec
shake	H only
time	150 psec
温度コントロール	500K→100K (150psecで一定の割合で冷却)
誘電率	4.0R
カットオフ長	20 Å(テーブル更新 10step)

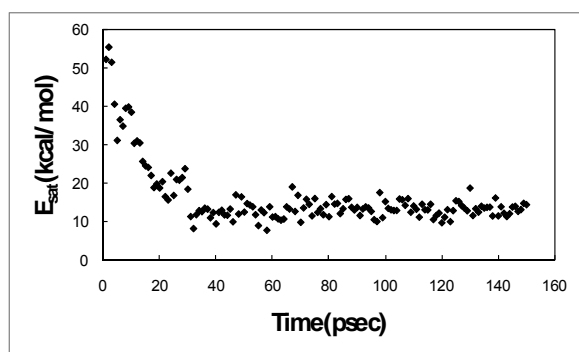
42 x 2個の計算のうち最終構造のRMSD(結晶構造とYUH側で重ね合わせた際のUb側の主鎖のRMSD。以下同じ)が5 Å以下の構造となったものは9個であった。

42x2個の分子動力学計算中に、サンプルした構造のRMSDとEsatをプロットすると、下図のようになり、Esatが小さい構造(実験データをよく説明する構造)を選択するとRMSDが小さい構造を選択できることがわかった。

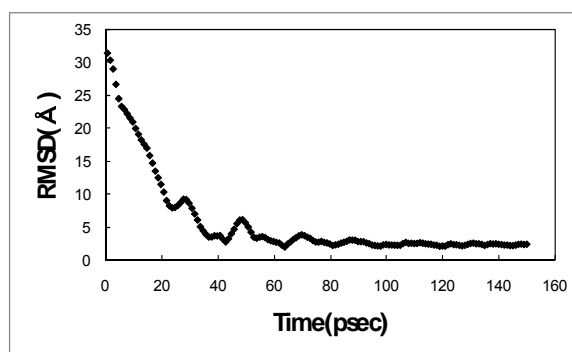


Esat の最小構造（青：YUH1、赤：Ub）（ $E_{\text{sat}}=7.88\text{kcal/mol}$ 、 $\text{RMSD}=3.05\text{Å}$ ）、参照構造（シアン：YUH1、オレンジ：Ub）

また本構造をサンプルした際の E_{sat} と RMSD の時間変化は以下の通りであった。

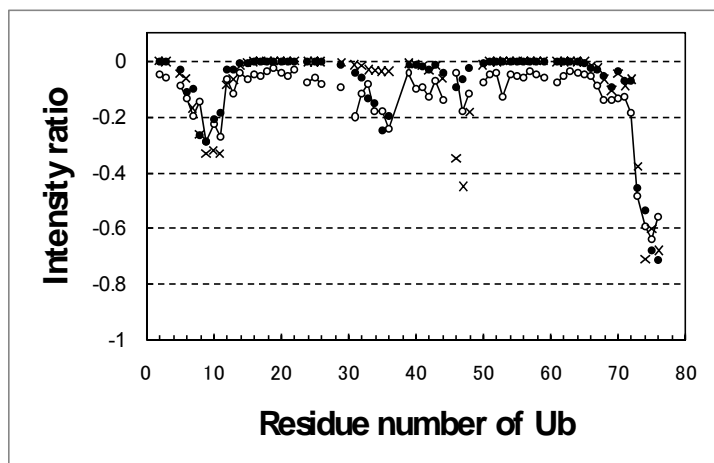


E_{sat} の時間変化



RMSD の時間変化

Esat 最小構造での η の計算値と実験値を比較すると下図 (LEU 標識データ) のようになり、 η_{exp} と η_{calc} (Esat 最小構造) が十分一致していることがわかる。



○・実線： η_{exp} 、●： η_{calc} (Esat 最小構造)、×： η_{calc} (結晶構造)

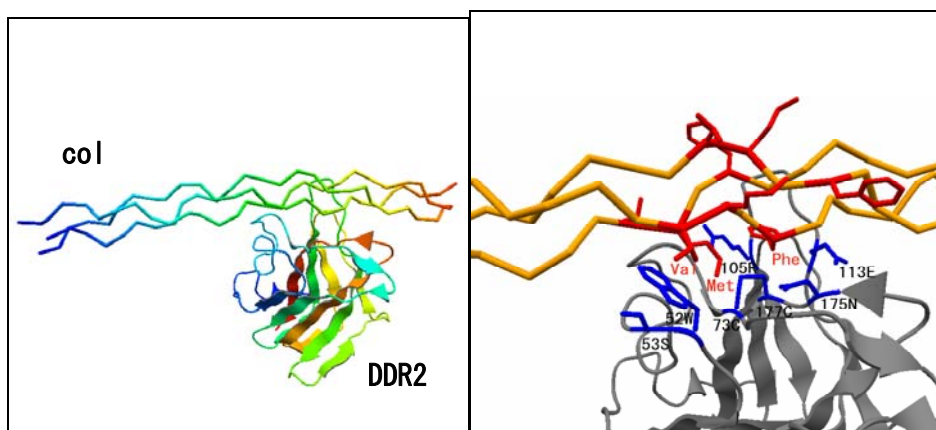
さらに、本構造を CAP 水中で Esat 項を外して再度 simulated annealing を行うと、RMSD は、0.67 Å まで改善される。これは、真空中で Esat 項を用いて計算したことにより生じた構造のゆがみが解消されたためと考えられる。

(b) DDR2(Disocoidin domain receptor2)-コラーゲンペプチドの複合体モデルの構築

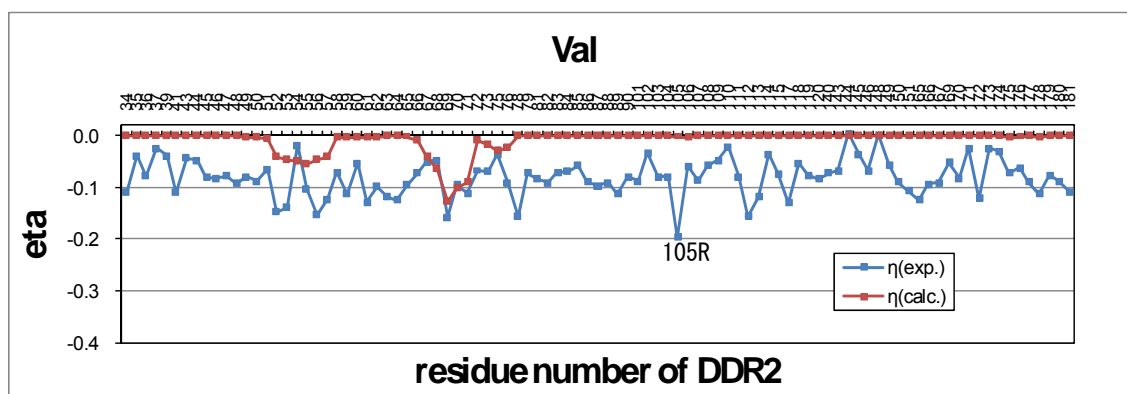
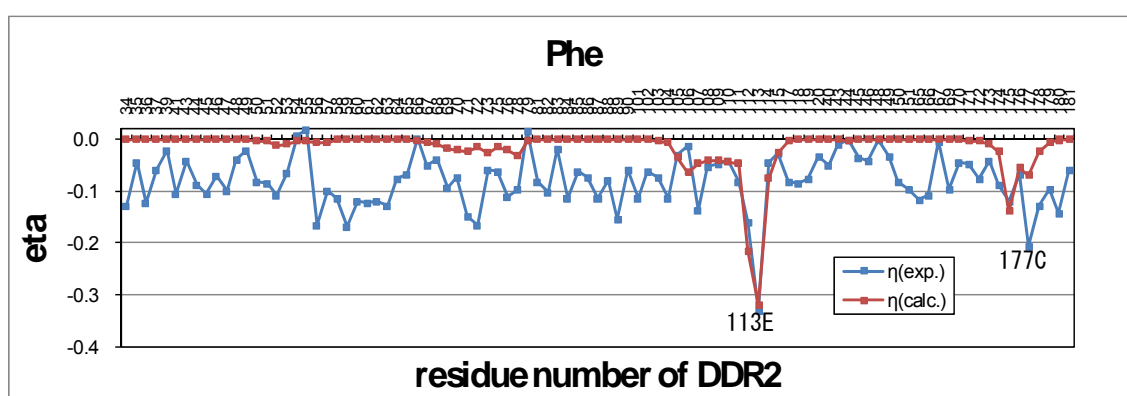
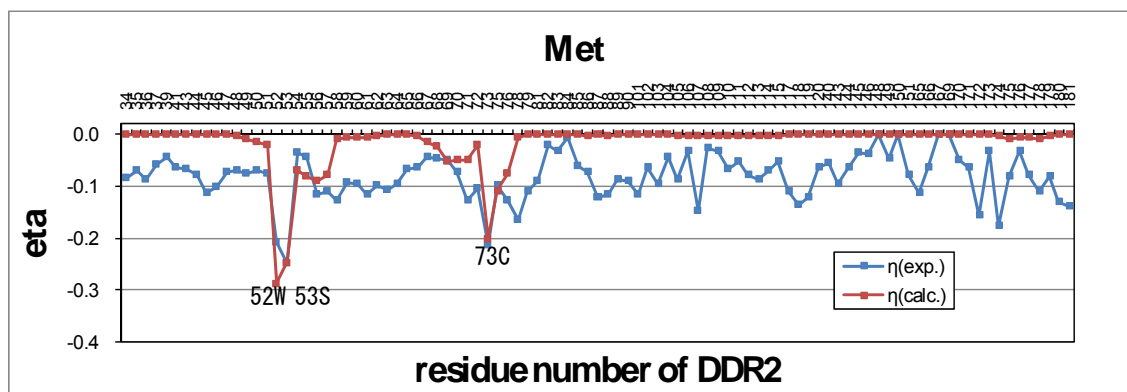
それぞれの単体構造として、DDR2 については嶋田チームが NMR で決定した (pdb code:2z4f) 20 構造を使用し、コラーゲンペプチドについてはテンプレートとして、PDB データベースに登録されているコラーゲンの X 線結晶構造 (pdb code:1k6f) に基づくホモロジー・モデルを用いた。

DDR2 の 20 構造に対して、我々が開発したドッキングソフト(上記「①タンパク質複合体構造予測法の開発」参照)を用いてコラーゲンペプチドをドッキングさせ、各 300 の複合体構造を作成した。嶋田チームがこの系に対して得た ASCS 実験の観測値を用い、この 20x300 構造から ASCS 実験結果と定性的に合致する 119 構造を選び、疑似エネルギー項 Esat を加えた分子動力学計算を行った。

119 個の計算でサンプルされた構造のうち、Esat が小さく磁化変化 η が ASCS 実験値とよく一致する構造の 1 つが下図の構造であった。



本構造の磁化変化 η の計算値は下図のようになり、主要なピークのほとんどについては実験値とよい整合性を示した。実験値と一致しない点については、側鎖の精密なモデリングの必要性や、複数の複合体ポーズの存在などが考えられる。



結論

以上の結果から、今回開発した ASCS の実験データを活用した分子動力学計算は、それぞれの単体の立体構造が既知の場合に、それらが形成するタンパク質複合体構造を高精度に構築する手法として、極めて有用であることが示唆された。これは、界面という2次元の情報を与える ASCS 実験データから、複合体という3次元の情報を得ることができることを示している。

本手法は、NMRの実験データを利用する方法であるため、複合体の結晶が得られない場合や、相互作用が弱い場合にも適用可能な手法である。また、ASCS 実験では、同位体ラベルした残基の界面情報しか得られないが、本手法では、複合体という3次元の情報を得ることで、ラベルされてい

ない残基も含めた蛋白質界面の情報を得ることができる。これにより、蛋白質複合体化を阻害する物質を設計する部位を探索することができる。

今回の研究は、嶋田チームの発表前の NMR 実験値を利用することができ、手法の検証に加えて新たな構造解析にも適用し、その有効性を示すことができた。チームを越えた協力研究としての成果を挙げる事ができたと考えている。

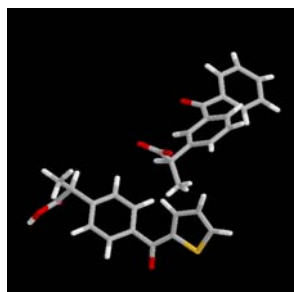
(iii) ペプチドと同様の結合性を有する非ペプチド性化合物を探索・設計する新しい手法の開発

(a) 新規で高精度な類似化合物探索ソフトウェアの開発 [BIRC集中研]

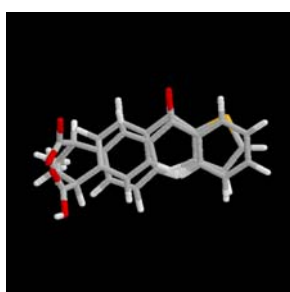
(a-1) molecular dynamics maximum volume overlap (MD-MVO)法—の開発と実施

類似化合物探索ソフト molecular dynamics maximum volume overlap (MD-MVO)法を、開発し薬物スクリーニングを開始した。

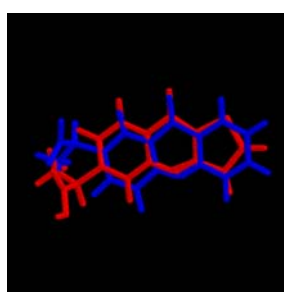
MD-MVO法とは、既知活性化合物に対して、データベースから選んだ化合物を、原子電荷も考慮しながら2化合物の体積重なりが最大になるようにMD計算によって、分子の配座探索と同時に分子重ね合わせをする手法である。この分子重ね合わせでは、周辺の環境である標的蛋白質を考慮した計算を行うこともできる。また、既知活性化合物を、標的蛋白質に結合していると期待されるもっともらしい構造に固定して、データベースから選んだ化合物のみを自由に運動させて2分子の重ね合わせを行うこともできる。



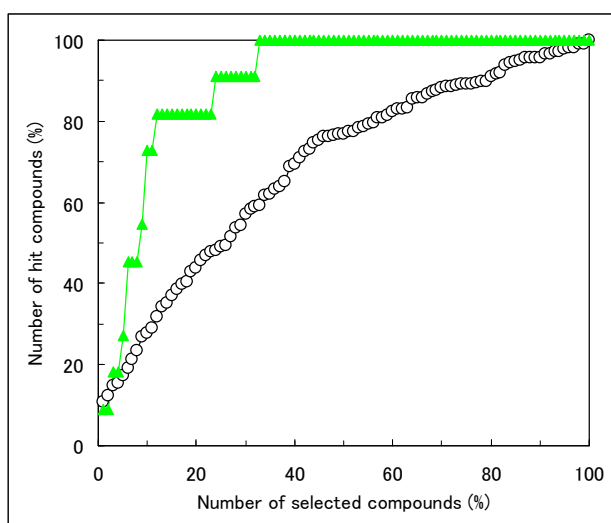
2分子の初期座標



MD-MVO法による重ね合わせの結果



μ オピオイド受容体のペプチド性リガンドであるエンドモルフィンを既知活性化合物として選択した。 μ オピオイド受容体について、既知の低分子リガンド(モルヒネ類など)を11種類文献より収集した。そして、エンドモルフィンを既知活性化合物として、1万化合物を含む化合物データベースに既知の低分子リガンド11種類を混ぜ、これらの既知の低分子リガンドを発見できるかテストしてみた。その結果が、下のデータベースエンリッチメントカーブである。我々が以前に開発した、ドッキングソフトを用いた ligand-based drug screening法である Docking score index (DSI)法と、今回の MD-MVO法の結果を比較した。



緑▲がMD-MVO法の結果、白○がDSI法の結果である。ペプチドを活性化化合物とするスクリーニングは困難であり、DSI法の結果がランダムスクリーニングに近くなっていることから、その困難さが伺える。明らかに、MD-MVO法はペプチドを活性化化合物とする場合も、優れたスクリーニング結果を示している。また、世界でもっとも広く使われている類似化合物探索ソフトROCSやChemFinderとの比較テストを複数の標的で行ったが、いずれもMD-MVO法がヒット率において、より優れていることが示された。

		MD-MVO 法		ROCS	
		AUC	ヒット率	AUC	ヒット率
COX2	1cx2	85.8	28.6	52.8	23.1
	1pxx	98.5	50.0	64.5	30.8
	3pgh	85.3	33.3	64.8	30.8
	4cox	74.0	33.3	68.5	15.4
HIV	1aid	59.8	5.0	29.2	4.5
	1bv7	93.3	40.0	75.8	36.4
	1hte	70.3	5.0	59.3	13.6
	1mes	88.2	35.0	77.0	36.4
	Average	81.9	28.8	61.5	23.9
	上位	MD-MVO	ChemFinder		
COX2	9%	38.20%	9.80%		
HIV	22%	51.40%	47.50%		

MD-MVO法とROCSによるデータベースエンリッチメントカーブの下の面積(AUC)と、上位1%化合物を採択したときのヒット率を表に示す。AUCが80%以上が緑色、ヒット率で30倍以上を空色で示

したが、MD-MVO法がROCSに対し優位であることが分かる。また、ChemFinderとの比較では、COX2阻害剤について上位9%、HIVプロテアーゼ1阻害剤に対し上位22%の化合物を採取したときのヒット率を表にしたが、やはりMD-MVO法はChemFinderに対しても優位性を示した。

MD-MVO法は、既知活性化合物とクエリー化合物の重ね合わせにおいて、周辺環境である標的蛋白質を考慮した計算を行うこともできることは、従来の構造重ね合わせ手法と全く異なる優れた点である。シクロオキシゲナーゼ2 (COX-2) に、MD-MVO法を適用して、その効果を見てみた。まず、PDB ID:4coxのリガンドである活性化合物compound aと、非活性化合物であるcompound bを用意した。COX-2蛋白質としてPDB ID:1cx2の構造と、既知活性化合物として1cx2のリガンドを選んだ。この結果を表にまとめた。

1cx2に対してcompound a/bを我々の蛋白質—化合物ドッキングソフトsievgeneでドッキングし、スコア上位の構造5つを選んだ。これらのスコア上位5構造に対して、1cx2の構造を考慮したMD-MVO法を適用した。Compound aのMD-MVO法のスコア(MVOscore)は、compound bのスコア0.46より大きく、活性化合物aが、非活性化合物bより、予測活性が高いという合理的な結果になった。

また、化合物aのドッキング構造を正解の構造である4coxの構造に重ね合わせてみると、ドッキング予測構造(蛋白質—リガンド複合体構造)が、極めて正しく予測されることがわかった。Sievgeneによるドッキングでは、compound aの予測複合体構造の座標の誤差は、2~6 Åと大きかったが、MD-MVO法での最終座標では、正解の4cox結晶構造に0.2~1.5 Åとほぼ一致した。このように、標的蛋白質構造を環境として考慮することは、ドッキングソフトによる予測複合体構造の精度向上に貢献することが示された。

ドッキングスコアでの順位	compound a				compound b	
	ドッキングスコア	RMSD (ドッキング計算) Å	RMSD (MD-MVO) Å	MVOscore	ドッキングスコア	MVOscore
1	-3.57	2.25	1.06	0.47	-3.53	0.45
2	-3.50	2.40	0.26	0.45	-3.48	0.45
3	-3.43	1.97	0.99	0.45	-3.46	0.46
4	-3.38	2.59	1.13	0.44	-3.40	0.45
5	-3.34	6.78	1.56	0.40	-3.38	0.37

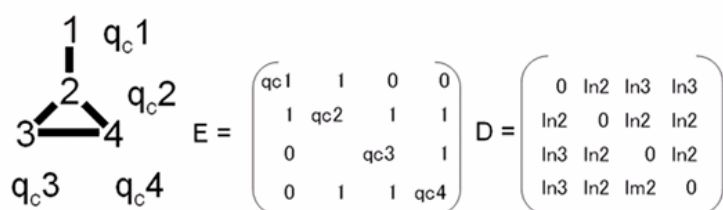
MD-MVO法は、MD計算の一種であるため、計算時間が長い。蛋白質を考慮しない状態でも1 CPUで、1日に約300化合物しかスクリーニングすることができない。従って、蛋白質を考慮しない状態で運用するしかない。現在保有する300コアを使用すると、2週間で100万化合物をスクリーニングすることができるため、 μ オピオイドに対して、エンドモルフィンを既知化合物としたスクリーニング計算を行い、塩野義製薬でアッセイを行っている。

(a-2) 分子グラフを用いた高速な類似化合物探索法

〔BIRC集中研〕

MD-MVO法では、1CPUにて1日当たり300化合物しかスクリーニングできない。このままでは、応用範囲が極めて限られてしまうため、1時間当たり1000万分子以上をスクリーニングできる高速なスクリーニング手法を開発した。

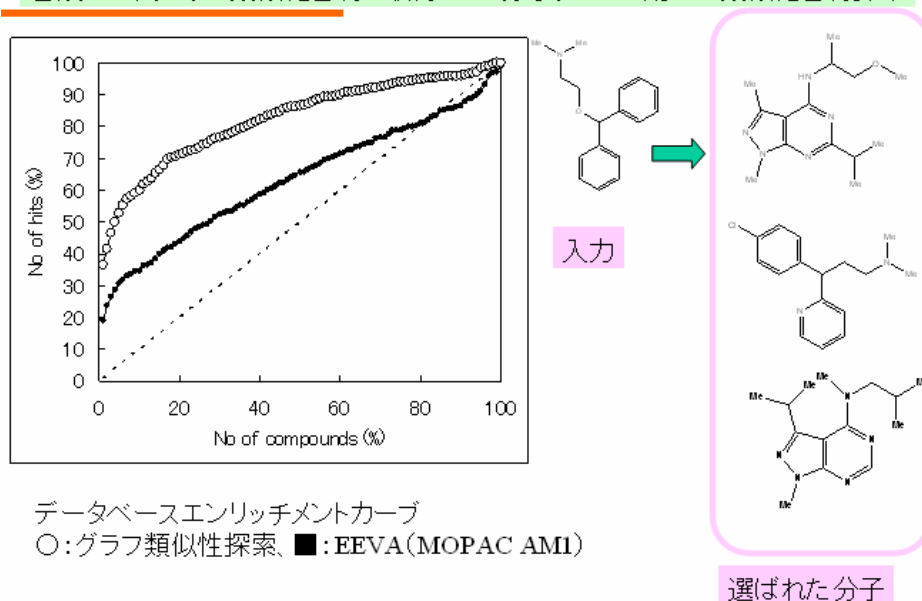
合成する代わりに類似化合物を取得する: 分子グラフを用いた類似化合物探索



分子グラフのグラフ不変量(行、列の入れ替えに対して不変な量)を用いる。結合行列、距離行列の行列固有値を分子記述子とする。

この手法では、分子をグラフとして捕らえ、その結合行列と距離行列を作成し、この行列固有値をグラフ不変量とする。この固有値を幅を持たせたヒストグラムに変換し、2分子の比較の場合、2つのヒストグラムの重なりによって、分子の類似性を評価する。

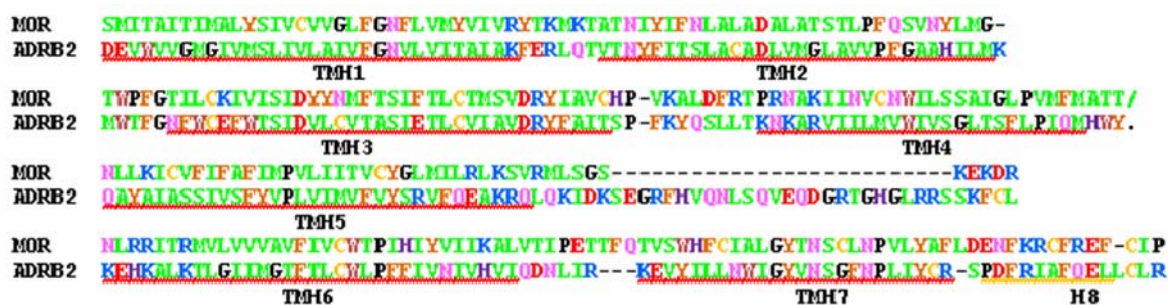
合成する代わりに類似化合物を取得する: 分子グラフを用いた類似化合物探索



この手法を実際に適用したのが、上の図である。グラフ類似性探索は、1時間1300万分子を探索するという簡便さのわりには良好なデータベースエンリッチメントを示し、実際、上位に選ばれた分子は、既知活性化合物と構造が似ていることが確かめられた。

(b) 蛋白質の動的構造を反映した薬物スクリーニングー GPCR を標的としたドッキングシミュレーション
〔BIRC 集中研〕

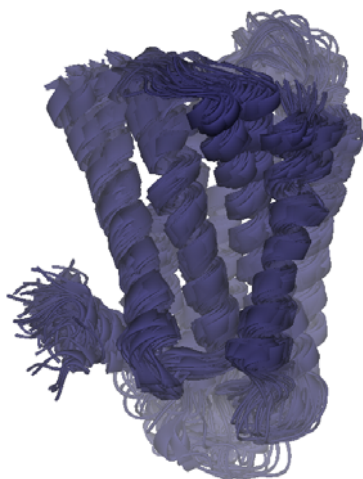
Gタンパク質共役受容体(GPCR)は重要な創薬標的であるが、その立体構造はほとんど解明されていない。そのため、GPCRを標的としてin-silicoスクリーニングを行うにはホモロジーモデリングにより標的タンパク質の立体構造を拘置する必要がある。また、induced-fitやタンパク質の動的挙動を考慮するには分子動力学(MD)計算が有効かもしれない。モデリング構造の信頼性やMDによる構造サンプリングの有用性を検証するために、まず、リガンド分子の構造活性相関が最もよく研究されているGPCRの一つであるヒト μ オピオイド受容体(MOR)のドッキングシミュレーションを行った。



μ オピオイド受容体(MOR)と β_2 アドレナリン受容体(ADRB2)の配列アライメント

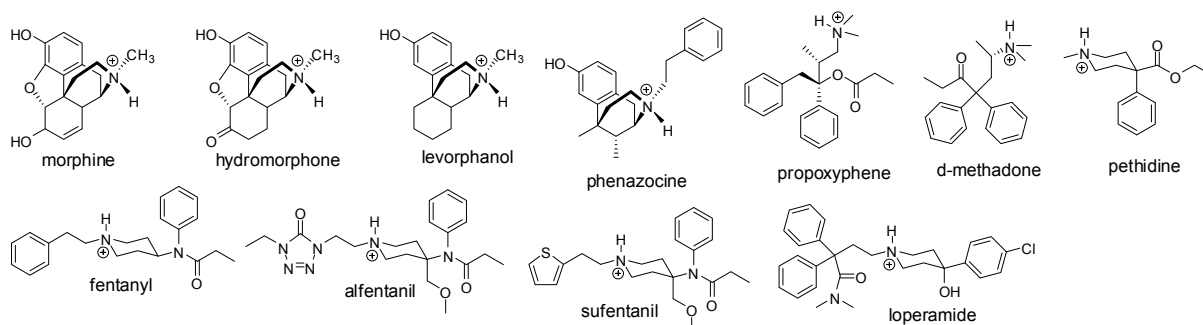
ヒト由来GPCRの β_2 アドレナリン受容体(PDB entry: 2rh1)をテンプレートとして、3種類のモデリングソフト(MOE, Modeller, Prime)を用いて、MORモデルを21構造作成した。

作成したホモロジーベースモデルそれぞれに対して周囲に水分子と電荷中和のためのイオンを配置し、エネルギー最小化と平衡化計算の後、周期境界条件で1n秒のNVT計算を実行し合計462種類のMOR構造をサンプリングした。



【図: MD計算により生成されたMORの構造】

ドッキングシミュレーション評価用化合物ライブラリは、Coelacanth社化合物ライブラリから抽出した18190化合物(デコイセットA)と3次元化合物ライブラリLigandBoxから抽出した11489化合物(デコイセットB)の2種類を準備した。また、MORの既知リガンド(アゴニスト)としてモルヒネやフェンタニルを含む11化合物を用意した。



μ オピオイド受容体の既知アゴニスト

462の標的MOR構造それぞれに対して、2つのデコイセット及び既知化合物のドッキング計算をSievGene/myPrestoで実行し、Multiple Target Screening(MTS)法による統計的スクリーニングを適用した。Area Under the Curve(AUC)として算出したヒット率の値を以下に示す。AUCとは、ヒット率(スクリーニングの精度)を定量的に表す指標で、0%~100%の範囲の値をとり、ランダムスクリーニングでは50%、理想的な高精度のスクリーニングならば100%に近い値を取り、逆に50%を下回る数字は、ランダムスクリーニングより悪い(すなわち計算しないほうがよい)結果を表す。

	MODELLER										MOE										PRIME
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	
initial	33.6	50.9	51.3	51.3	71.8	45.2	47.4	46.8	43.7	28.8	48.7	31.0	32.6	45.2	49.7	34.4	52.6	34.8	53.5	36.6	58.5
minimized	28.5	41.5	55.8	54.2	75.9	57.6	56.8	46.9	74.2	60.5	66.2	79.5	73.5	30.5	32.7	60.1	61.7	64.6	68.7	68.3	69.1
50 ps	68.3	60.5	47.4	52.9	70.6	51.6	67.5	71.8	69.5	44.7	66.3	33.6	65.8	36.9	18.6	50.4	62.2	74.9	75.0	35.7	69.4
100 ps	53.9	47.2	45.2	49.3	71.5	64.0	50.0	55.7	78.1	61.6	61.5	47.0	66.9	65.5	43.3	48.6	69.6	57.6	84.0	38.4	56.8
150 ps	50.9	47.8	45.0	73.8	63.4	43.5	60.0	69.7	86.5	67.4	70.5	59.4	70.6	70.8	28.6	44.3	86.5	60.1	72.6	42.4	57.8
200 ps	41.7	43.1	54.5	77.6	53.4	72.2	47.0	29.1	76.5	38.2	54.5	47.2	63.9	76.9	58.8	43.5	84.7	80.7	87.0	55.1	43.1
250 ps	37.2	39.1	54.0	76.0	57.6	42.5	60.1	43.7	70.9	49.3	64.9	48.8	63.8	51.7	33.9	38.9	80.3	58.0	59.5	64.5	52.8
300 ps	39.7	46.9	54.4	81.0	59.3	67.8	43.9	51.5	82.1	32.5	54.1	39.0	42.4	45.4	47.9	55.2	65.6	72.4	86.7	58.6	73.8
350 ps	53.1	59.9	52.5	70.2	63.8	51.8	45.2	57.5	70.0	41.4	75.4	42.1	77.0	55.2	64.9	62.4	83.0	60.0	61.5	53.8	69.2
400 ps	50.8	56.7	69.2	67.9	67.3	33.5	57.5	57.0	59.2	63.0	48.6	32.4	62.5	43.6	48.7	57.3	77.2	69.9	78.2	48.0	76.6
450 ps	50.5	50.4	52.4	63.4	56.8	55.6	52.8	32.9	80.9	55.8	56.6	38.2	51.2	47.4	75.0	46.7	40.1	67.4	77.1	30.8	56.6
500 ps	37.4	49.3	56.5	79.2	53.5	48.3	45.3	59.5	70.3	48.5	60.3	37.8	59.5	55.8	61.4	43.0	73.9	90.3	75.6	64.0	51.4
550 ps	50.0	67.7	46.3	67.1	62.0	53.6	43.2	47.5	65.5	57.9	53.1	51.8	51.1	46.6	62.3	46.9	74.3	82.6	74.3	64.0	57.1
600 ps	51.1	44.0	54.0	73.7	48.9	73.2	31.3	31.5	52.5	71.5	66.3	45.5	52.0	31.3	65.8	59.6	49.9	70.0	81.6	48.5	64.9
650 ps	44.7	42.4	34.9	66.8	41.4	65.7	34.7	56.5	66.5	66.0	62.8	44.5	53.2	38.8	23.6	65.7	92.5	67.3	70.7	53.3	66.9
700 ps	49.3	62.8	30.6	88.6	70.4	56.5	47.9	48.9	70.4	59.8	48.9	48.4	53.9	55.2	59.6	46.4	85.6	67.4	40.9	63.3	69.2
750 ps	40.2	46.2	61.4	76.8	58.5	39.3	37.6	71.6	39.6	59.3	62.7	38.6	60.1	45.0	62.6	58.6	81.9	49.1	89.0	58.6	54.4
800 ps	45.5	41.3	54.5	73.0	56.9	57.7	54.2	67.3	64.6	62.5	57.9	48.2	42.1	58.2	44.3	48.5	88.9	83.5	70.1	34.4	55.0
850 ps	59.0	54.5	20.5	70.5	50.3	59.7	52.6	54.1	61.0	72.1	76.2	39.9	63.5	59.1	64.7	44.9	60.4	59.5	83.3	44.4	76.0
900 ps	66.9	31.5	60.3	78.8	41.5	50.3	55.4	57.6	59.2	37.2	44.3	50.5	63.4	57.5	41.2	49.3	63.5	73.3	56.1	54.1	80.0
950 ps	31.5	45.0	32.0	70.9	51.5	68.5	47.4	58.9	72.8	46.3	44.8	55.0	48.1	41.2	53.0	62.6	85.1	64.5	60.6	55.9	54.7
1000 ps	41.5	39.0	61.4	82.6	51.4	66.6	65.5	36.4	56.7	35.5	40.3	52.2	61.1	53.0	42.5	54.1	57.5	66.6	47.6	63.9	53.1

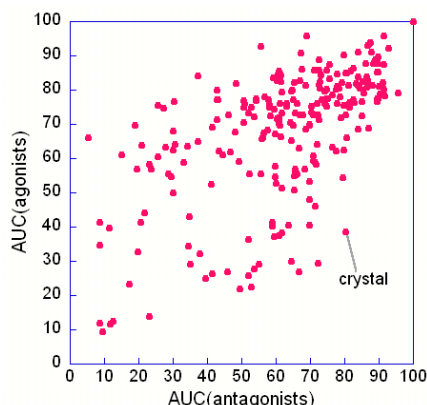
デコイセットAの462の標的構造に対するAUCの値。

	MODELLER										MOE										PRIME
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	
initial	67.4	61.5	74.9	71.2	82.4	65.8	64.5	58.5	54.3	56.9	65.6	38.9	33.8	66.0	44.5	44.7	67.1	48.7	40.5	60.6	59.4
minimized	58.2	64.3	45.5	67.3	85.7	69.1	77.8	62.2	71.8	77.7	65.5	56.2	56.1	42.5	33.7	50.2	64.2	69.9	77.4	63.6	59.9
50 ps	79.6	75.3	29.5	41.2	63.8	32.8	78.4	65.9	72.4	68.3	76.5	44.7	36.4	44.3	31.0	47.9	46.5	67.7	58.7	53.6	49.4
100 ps	65.6	68.0	54.5	46.0	76.4	42.6	56.8	68.3	72.0	82.5	76.4	66.7	46.9	78.2	60.9	56.4	34.9	63.8	77.2	41.0	38.1
150 ps	68.4	52.1	23.5	64.8	69.7	40.5	53.9	74.5	80.9	86.0	70.6	61.6	54.6	76.2	45.4	44.2	66.3	53.6	70.3	42.8	30.7
200 ps	70.8	71.2	52.0	81.4	63.0	50.5	46.3	45.2	80.9	67.9	63.7	59.0	56.0	79.7	46.5	60.2	64.4	66.9	78.5	67.8	17.0
250 ps	74.8	58.5	57.1	61.7	57.0	49.2	57.7	58.5	72.4	71.1	76.5	62.0	67.3	44.9	26.9	56.7	60.8	40.7	53.6	66.9	26.6
300 ps	62.6	55.2	59.5	65.1	63.6	53.5	34.1	47.9	77.3	48.5	59.1	61.5	40.6	41.2	71.5	60.5	49.6	57.5	71.4	59.5	42.5
350 ps	74.7	65.3	66.5	69.1	59.5	45.5	34.8	73.5	65.8	64.3	76.4	57.3	65.5	46.0	66.6	76.6	72.0	41.5	49.3	49.0	37.9
400 ps	79.6	68.5	72.6	67.2	59.6	33.4	48.5	71.8	52.8	84.4	55.8	54.7	60.5	40.4	31.1	74.0	48.6	63.5	75.0	45.0	53.5
450 ps	64.1	70.0	51.7	52.7	55.9	36.6	54.7	58.5	87.0	69.9	58.4	44.5	58.2	46.8	69.5	58.5	24.3	69.0	66.5	29.0	42.9
500 ps	68.2	63.2	61.2	79.9	61.6	34.3	35.8	68.3	77.9	69.3	61.5	60.5	37.8	43.9	58.6	40.2	59.8	85.0	58.8	62.0	27.2
550 ps	70.8	68.5	60.9	63.1	68.1	45.0	27.3	66.2	74.5	69.6	66.0	38.2	47.8	51.2	51.4	52.0	52.2	79.6	72.3	70.8	28.5
600 ps	64.5	58.0	57.9	55.0	49.4	69.6	32.6	59.0	42.5	78.5	75.6	58.3	51.2	53.5	57.8	72.6	33.6	68.5	80.4	56.4	37.4
650 ps	62.2	56.6	43.5	60.5	51.0	63.7	21.5	62.2	69.3	67.1	80.4	49.2	63.7	42.9	45.1	78.6	78.1	72.2	47.7	63.8	48.8
700 ps	61.7	65.1	39.7	75.4	79.4	51.1	50.7	62.3	76.7	76.2	70.8	57.6	63.7	65.8	57.6	53.9	78.4	65.4	28.1	64.5	44.5
750 ps	59.5	51.9	51.5	71.2	59.5	19.6	26.2	70.5	41.2	57.0	72.6	62.8	58.4	52.3	52.7	81.7	59.5	55.1	85.5	64.7	23.3
800 ps	63.5	49.5	35.3	61.5	63.0	42.5	44.8	78.2	67.1	73.4	66.4	61.6	21.7	60.8	44.2	68.2	72.5	65.7	74.5	34.3	25.9
850 ps	77.4	53.8	18.4	59.8	62.3	35.8	49.9	73.8	70.4	80.0	75.0	56.3	47.5	71.5	39.3	58.5	45.4	70.6	84.6	48.1	40.4
900 ps	85.3	45.2	32.2	49.5	68.9	29.3	56.1	66.4	71.0	45.3	50.9	57.2	48.4	61.3	22.5	70.3	45.2	71.3	62.4	48.4	44.5
950 ps	34.4	48.1	24.4	41.5	48.4	61.8	39.5	64.5	70.5	67.9	63.6	54.0	45.9	56.5	28.8	71.9	71.5	62.5	61.2	63.8	28.2
1000 ps	55.5	55.1	28.8	50.8	59.4	66.8	57.1	56.1	57.2	49.1	66.7	61.2	46.7	65.3	35.0	69.7	52.4	65.3	35.7	62.7	26.7

デコイセットBの462の標的構造に対するAUCの値。

MD計算における構造サンプリングした時間によってAUC値はさまざまで、ランダムスクリーニングよりも悪い結果を与えるもの(図の青色部分)もあるものの、80以上の高い値を示す標的構造(赤色部分)も数多く見つかった。膜を頭で考慮していないにもかかわらず、1n秒程度のMD計算で高いヒット率を示す標的構造をサンプリングすることができた。

同様のスクリーニング計算を、b₂アドレナリン受容体(ADRB2)、A_{2A}アデノシン受容体、ロイコトリエン受容体(CysLT1)およびb₃アドレナリン受容体の4種類のGPCRにも適用した。ADRB2とA_{2A}の結晶構造は既知であるが、Modellerにより側鎖配座の異なる初期構造をいくつか用意してMDによる構造サンプリングを実行した。いずれのGPCRの場合も、AUC値が80~90を超える標的構造が多く見つかった。また、アゴニストとアンタゴニストの両方のスクリーニング計算を実行してみると、それぞれに選択性のある標的構造も見出すことができることもわかった。



ADRB2 サンプリング構造を標的として、アゴニストとアンタゴニストのスクリーニング計算を実行したときのAUC値の相関。

以上の結果から、GPCRIに対して、標的タンパク質構造のホモロジー・モデリングとMDシミュレーション計算、およびMTS法による統計的スクリーニングによって、既知リガンドがある場合には、ヒット率が高いモデルをあらかじめ選択できることが示唆される。すなわち、シミュレーション計算に用いる出発構造は、たとえX線結晶構造でないホモロジー・モデルによるものでも、ヒット率としては高いものが得られることがわかった。ここで選択された構造を基に、実証研究を試みることにした。

さらに、agonist/antagonistの結合の違いについても、差別化した探索の可能性も示唆された。すなわち、GPCRIはagonist結合型とantagonist結合型の間の変化する構造を行うが、それら構造変化のうちリガンド結合サイトについては、シミュレーション計算によってサンプリングされる構造が類似の構造をある程度再現できているとも考えられる。複数のagonist/antagonistが既知であれば、この手法によって、未知の化合物が結合するモデルのタイプに依存して、その化合物がagonist型かantagonist型からの推定ができよう。さらに、リガンド結合サイトについて、agonists/antagonist結合型の特徴に基づく分類がなされれば、McMD法やSRPG法との組み合わせによる高効率の構造サンプリング手法と合わせて、より効果的なドッキング計算も期待される。

(c)生理活性ペプチドから非ペプチド性化合物(低分子化合物等)を探索・設計する手法の開発

[塩野義製薬分室]

【事前調査】

PDBより、蛋白質と複合体を形成しているペプチドと、同じドッキングポケットに結合している他の低分子のデータを約20対収集した。このデータを用い生理活性ペプチドから非ペプチド化合物が探索できるか、ドッキング計算を基にした類似化合物探索手法(DSI法)で探索する実験を試みたが、ヒット率はランダムスクリーニングと大差なかった。理由としては、ペプチドは自由度が大きくdocking計算に向かないのでDSI法が使いにくいと考えられる。そこで新しい類似化合物探索手法を開発することで、生理活性ペプチドから非ペプチド低分子化合物が探索できるかどうか検討を進めた。

【研究内容】

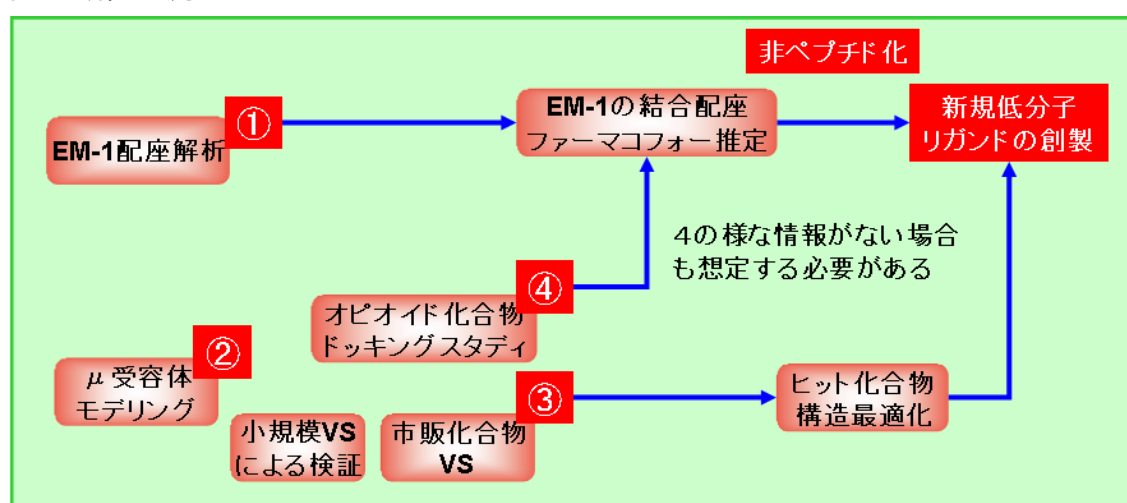
生理活性ペプチドから非ペプチド性化合物(低分子化合物等)を探索・設計する手法の開発の一環として、内在性の μ -オピオイドの一種であるEndomorphine-1 (Tyr-Pro-Trp-PheNH₂: 以下、EM-1と略す)から非ペプチド性の新規 μ 受容体リガンドを探索するための検討を行った。具体的な研究方針を以下に示す。

EM-1の非ペプチド化を行う上では、その結合配座およびファーマコフォーが有効な情報になると考えられる。そこで、これらの情報を得るための解析に必要となるEM-1の立体配座集団および μ 受容体モデルを得るべく、EM-1の立体配座解析(後述の取り組み1)および μ 受容体のホモロジーモデリング(取り組み2)を行った。

次に、取り組み2で得られたモデルの検証を行うため、取り組み3において、小規模バーチャルスクリーニング(VS)による検証を行った。また、実証実験として市販化合物を対象としたVSについても検討を行い、新規 μ 受容体リガンドの探索を試みた。

また、 μ 受容体については既知のアゴニストリガンドが複数報告されており、これらのファーマコフォー情報についてもEM-1の結合配座およびファーマコフォーを推定する上で有効と考え、取り組み4において、オピオイドリガンドのドッキングスタディについて検討を行った。ただし、既知リガンドがなく生理活性ペプチドの情報しかない場合も想定されるため、このようなケースに対応する方法についても今後検討する必要があると考えている。以下に全体の研究の流れを示す。

図1. 研究の流れ



【研究成果】

以下にこれまでの取り組みおよび得られた成果を挙げる。

1. EM-1の立体配座解析

ターゲットである μ 受容体への結合配座の推定やファーマコフォー抽出の解析に活用することを目的に、立体配座の網羅的な探索を行った。

この検討には、myPrestoのcosgeneによる溶液中でのForce-biased Multicanonical Molecular Dynamics (F.B.McMD)シミュレーションにより解析を行った。解析の結果、2318個の立体配座集団が得られた。これら得られた配座集団については、今後のEM-1のドッキングスタディの取り組みにおいて活用したいと考えている。

2. μ 受容体のホモロジーモデリング

上記1の取り組みと合わせたEM-1の結合配座の推定や、バーチャルスクリーニングによる新規 μ 受容体リガンドの探索に活用することを目的に、 μ 受容体の立体構造を予測することとした。

human β 2 adrenergic G-protein-coupled receptor (β 2AR)を鋳型構造として、MOE (CCG社)

とPrime (SCHRODINGER社)を用いて計11個の μ 受容体モデルを構築した。また、これらはモデル構造であることから、構造の精密化およびタンパク質の動的挙動を考慮するため、cosgeneを用いたエネルギー極小化およびMDシミュレーションによる構造サンプリングを行った。解析の結果、最終的に132個の μ 受容体のサンプリング構造を得た。

3. MTS法スクリーニングによる μ 受容体構造の検証および新規 μ 受容体リガンドの探索

はじめに、上記2の取り組みにより得られた μ 受容体のサンプリング構造集団を用いて、MTS法(Multiple target screening法)による小規模スクリーニングの検証を行った。検証セットとしては、モルヒネやフェンタニル等の既知のオピオイドリガンド11個とデコイリガンド11,479個を用いた。この検証の結果、最も高いヒット率で既知リガンドを抽出できる構造を選抜し、この構造を用いて市販化合物に対するMTS法およびDSI法(Docking Score Index法)スクリーニングを行った(この計算についてはJBICで実施)。スクリーニング結果および化合物の在庫状況の確認の後、最終的に399個の候補化合物が選抜された。そしてこれらについて、 μ 受容体に対する結合親和性試験を実施したところ、複数個の活性化合物が確認された。

4. オピオイドリガンドのドッキングスタディ

EM-1の結合配座の推定やファーマコフォー抽出の解析に活用することを目的に、既知のオピオイドリガンドのドッキングスタディを行った。この際、既知リガンドとしては、モルヒネやフェンタニル等10個のリガンドを用いた。また μ 受容体モデルとしては、Mosberg等により報告されている活性型の μ 受容体モデルを用いた。cosgeneを用いたMDシミュレーションにより200個の構造をサンプリングした後、Sievene, GOLD(CCDC社), S社のドッキングソフトを用いてアンサンブルドッキングを行い、多数のドッキングポーズを得た。そして、得られたドッキングポーズをエネルギー極小化により精密化した後、MM-GBSA手法により簡易的な結合自由エネルギー(ΔG_{bind})を算出した(エネルギー極小化, MM-GBSA計算については、AMBER 8を使用)。最後に、各リガンド毎で ΔG_{bind} が良好なドッキングモデルを比較し、推定の共通ファーマコフォー情報を得た。

以下、各項目の詳細な検討内容について記載する。

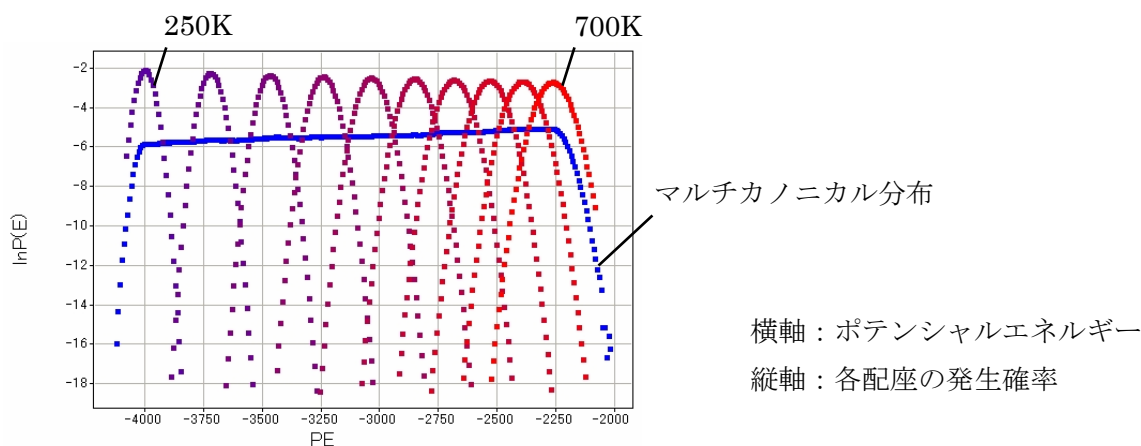
1. EM-1の立体配座解析

EM-1分子が単独で取り得る立体配座の算出を試みた。EM-1は4残基と比較的短いペプチドであること、および2番目のアミノ酸として自由度の少ないProを含んでいるため、その立体配座の数は通常の生理活性ペプチドよりも限定されるものと期待される。

立体配座解析の手法としては、先の生体高分子プロジェクトの成果物であるmyPrestoのcosgeneによる溶液中でのF.B.McMDシミュレーションによる解析を行った。通常、配座解析手法としては分子動力学法や二面角の系統的探索法があるが、ペプチド分子のように回転可能結合が多い分子では、全配座空間を網羅的に探索することは難しく、一般的な分子動力学法による配座解析では局所空間に留まってしまう恐れがある。また、二面角の系統的探索法では、多くの計算時間を要し、また実際には優位に取り得ない配座をサンプリングしてしまう恐れもある。そこで、F.B.McMDシミュレーションを用いることで、網羅的かつ効率的な配座探索を試みた。

F.B.McMDシミュレーションの検討の結果、下図で示すような高いエネルギーから低いエネルギーを持つ配座までのマルチカノニカル分布を示す配座集団が得られた。そして、これらの集団について、myPrestoのreweightFBツールを用いて解析を行ったところ、下記で示すような各温度に対応するカノニカル分布を示す配座集団を得ることができた(50K毎に表示)。そして、これらの結果の内、300Kに該当する配座集団を抽出したところ、2318個のEM-1の立体配座が得られた。

図2. F.B.McMDシミュレーション結果

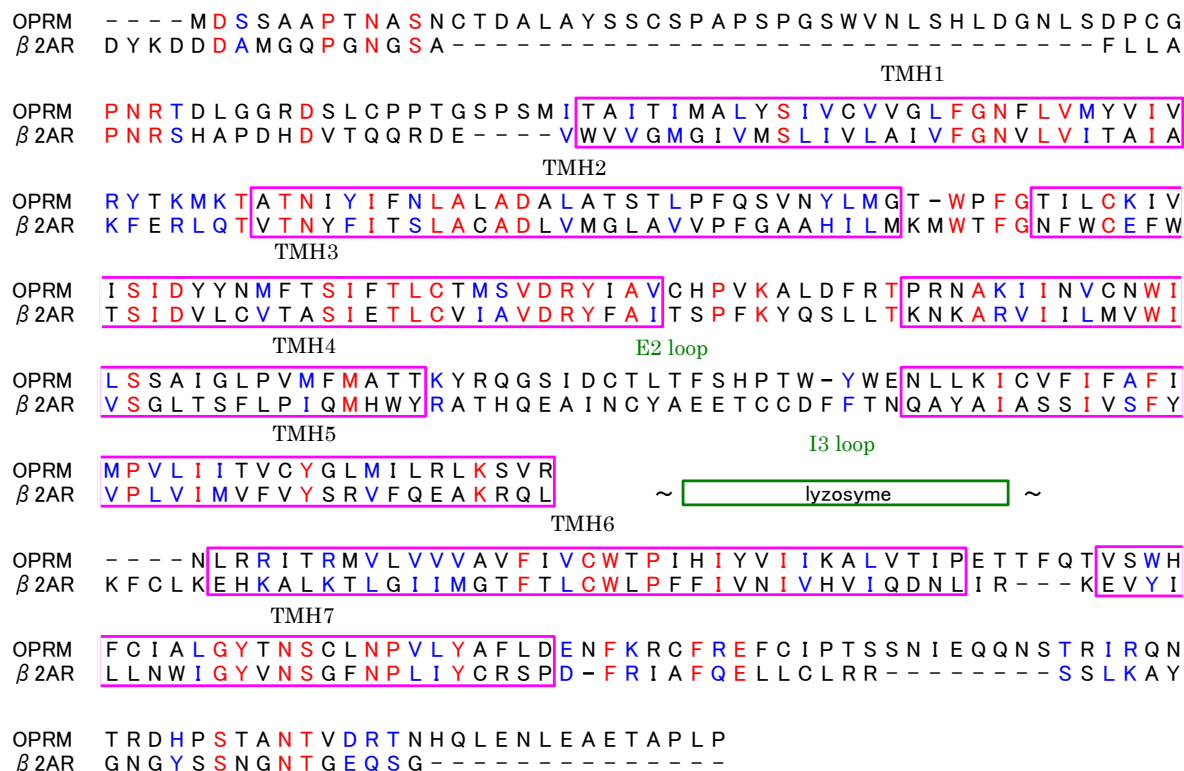


2. μ 受容体のホモロジーモデリング

μ 受容体は、7回膜貫通型のGPCRであり、その立体構造は明らかにされていない。本検討では、ホモロジーモデリング手法により、その立体構造の予測を試みた。

実際のホモロジーモデリングの作業としては、はじめに、clustalwを用いて、アミノ酸配列の multiple alignmentを行った。用いたアミノ酸配列としては、GPCRの中でRhodopsin-likeファミリーの内、peptideサブファミリーに属するOpioidやAngiotensin、Somatostatin受容体等の38配列、そして結晶構造が明らかにされているRhodopsin受容体(Bovine)および β 2ARの2配列の計40配列である。以下に、multiple alignmentの結果の中から、 μ 受容体と β 2AR(結晶構造PDB ID:2rh1.pdbのアミノ酸配列)のアミノ酸アライメントを抜粋したものを示す。

図3. アミノ酸アライメント



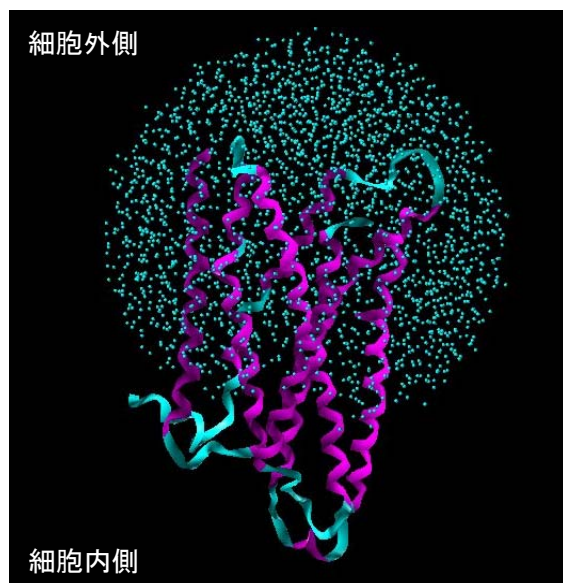
上段： μ 受容体
下段： $\beta 2$ アドレナリン受容体
赤字：一致アミノ酸残基
青字：類似アミノ酸残基

2rh1.pdbでは、 $\beta 2AR$ -lyzosymeのfusion proteinが結晶化用タンパク質として用いられており、TMH5と6の間にlyzosymeが存在している。

上記のアライメントを基に、 $\beta 2AR$ の結晶構造を鋳型構造として、MOEおよびPrimeを用いて、ホモロジーモデリングを行った。なおこの際、E2ループ(Extracellular loop2, TMH4と5の間にあるループ構造)部分については、 μ 受容体と $\beta 2AR$ 間で相同性が低く且つ非常に長いループ構造となっていることから、モデリングが困難であったので除外した。以上のホモロジーモデリングの結果、MOEとPrimeから各々10個と1個の計11個の μ 受容体モデルが得られた。

続いて、構造の精密化および構造のサンプリングを目的に、cosgeneを用いたエネルギー極小化およびMDシミュレーションを行った。この際、結合ポケットが存在すると予想される細胞外側の周辺に下図のようにCAP水をあらわに配置した。また、構造が崩れるのを防ぐために、主鎖にposition restrainをかけて計算を行った。詳細な設定については、以下の通りである。

図4. MDシミュレーションの条件

System

Method : canonical(NVT)

11362 atoms(264 protein residues,
water)

Cut off : 10Å

CAP : 30Å

Time step : 1.0fs SHAKE(H only)

Temperature : 300K

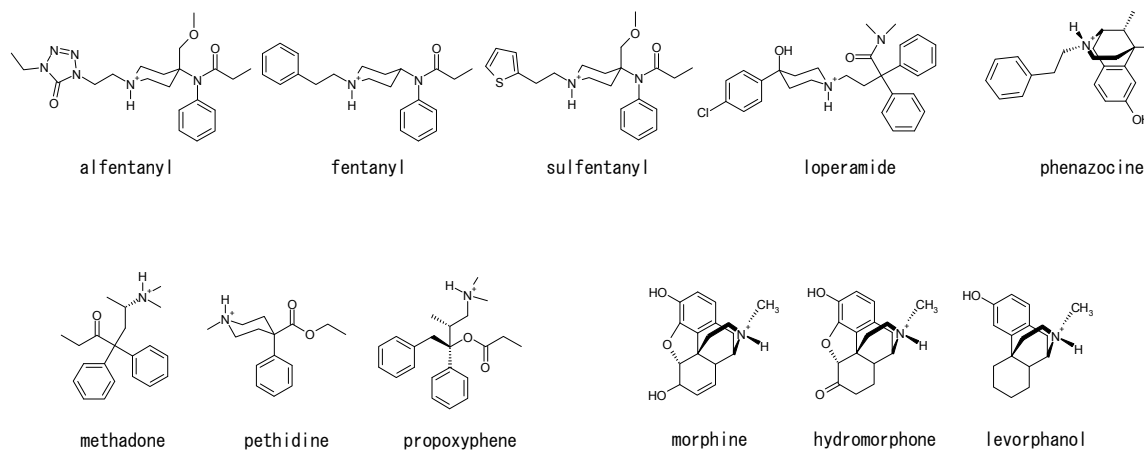
Position restrain : 10(main chain)

このシミュレーションについて、11個の μ 受容体モデルの各々について4nsecずつ行った。その結果最終的に、各 μ 受容体モデル構造毎に、初期構造:1個、エネルギー極小化構造:1個、MDサンプリング構造:10個(構造が平衡化した2nsec~4nsec間で200psec毎にサンプリング)の構造が得られ、合計で132個のサンプリング構造が得られた。

3. MTS法スクリーニングによる新規 μ 受容体アゴニストの探索

上記2の取り組みにより得られた μ 受容体のサンプリング構造集団を用いて、MTS法による小規模バーチャルスクリーニングの検証を行った。検証セットとしては、以下に示すオピオイドリガンド(アゴニスト)11個とデコイリガンド11,479個を用いた。デコイリガンドについては、MTS法スクリーニングシステム内の化合物データベース(Namiki, 約200万化合物)の中から、オピオイドリガンドと構造は似ていないが、物理化学的特性(AlogP、MW、H_Donor、H_Acceptor)が似ている化合物集団を選抜した。

図5. 既知のオピオイドリガンド



以下にMTS法スクリーニングの解析結果を示す。スクリーニング精度の指標としては、AUC (the Area Under the Curve)を用いた。AUCは、データベースエンリッチメントカーブ下の面積に相当するもので、0%~100%の範囲の値をとり、ランダムスクリーニングでは50%となり、値が高い程スクリーニング精度が良いことを示す。

図6. 小規模MTS法スクリーニングの結果

	moe										prime
	1	2	3	4	5	6	7	8	9	10	
ini.	79.2	63.3	47.6	62.6	56.5	50.5	64.3	47.7	35.5	63.3	49.3
min.	54.3	48.0	50.1	65.6	64.4	69.2	60.7	47.7	51.6	71.9	62.3
2100 ps	76.6	69.4	61.8	59.1	56.1	60.0	60.1	45.7	65.0	73.4	59.2
2300 ps	66.7	63.2	35.2	57.2	50.2	49.2	82.8	34.9	66.7	48.7	60.3
2500 ps	68.8	71.1	51.9	66.2	64.8	36.2	75.6	54.7	65.8	63.9	59.7
2700 ps	48.8	60.3	44.5	62.7	65.8	47.5	63.0	61.2	60.9	54.1	58.0
2900 ps	74.2	64.8	56.7	62.2	71.9	56.9	53.6	57.2	62.4	37.3	50.7
3100 ps	87.7	65.4	70.9	57.2	60.1	63.7	67.3	53.3	53.0	58.6	46.3
3300 ps	73.7	68.8	57.9	52.6	56.8	59.1	65.2	54.3	41.2	64.6	60.3
3500 ps	64.7	40.4	36.8	57.8	58.0	58.2	57.1	53.2	55.7	49.0	60.6
3700 ps	76.9	61.9	48.9	56.1	57.5	37.3	70.0	60.9	66.2	32.5	55.2
3900 ps	75.8	64.3	54.6	45.0	52.2	49.3	69.6	47.3	63.3	44.9	60.4

iniは初期構造、minはエネルギー極小構造、xxxxpsはMDシミュレーションの実行時間を表す。赤枠の部分だけ色分けをした。

上記スクリーニングの結果、132個の構造の中で、MOE MODEL1の3900psのMD構造が最も精度良くオピオイドリガンドを抽出できたことが分かった。この結果を受けて、MOE MODEL1の3900psのMD構造を用いて市販化合物(Namiki, 約200万化合物)を対象としたMTS法スクリーニングを行った。この際、図4に示す既知のオピオイドリガンドを用いたDSI法についても検討を行った(これらの計算についてはJBICで実施)。検討の結果、各手法から以下の数の化合物が選抜された。

MTS法 : 1000化合物(98化合物)

DSI法 : 1000化合物(54化合物)

MTS法/DSI法共通 : 1234化合物(247化合物)

そしてこれら化合物について、在庫状況等を勘案したところ、最終的に上記で示す括弧内の数(計399化合物)が選抜された。

選抜した化合物を評価するために、 μ 受容体に対する結合親和性試験を実施した(この試験については外部機関で実施)。測定条件としては、Diprenorphineに対する競合阻害実験を化合物濃度50 μ M、1点の測定で行った。その結果、以下に示すように複数の活性化化合物が確認された。仮に、試験条件である化合物濃度50 μ Mにおける%inhibition \geq 50をヒット化合物とすれば、ヒット率は12%であった。

今回は、暫定的に%inhibition \geq 50を仮のヒット化合物とした。今後、濃度依存性試験が得られた後に、再度、活性化合物の判定を行いたいと考えている。

表1. 結合親和性試験結果

%inhibition	化合物数
$90 \leq x \leq 100$	8
$80 \leq x < 90$	7
$70 \leq x < 80$	12
$60 \leq x < 70$	11
$50 \leq x < 60$	10
$x < 50$	351
計	399

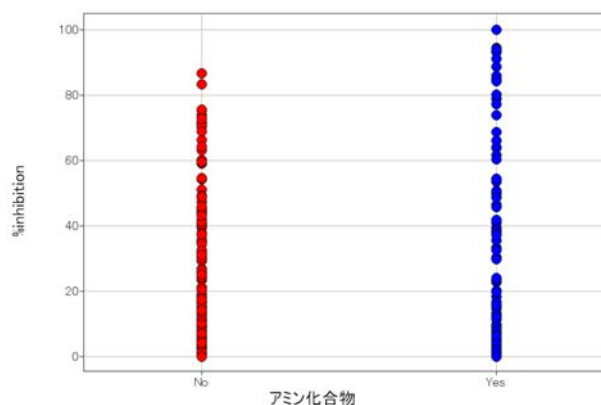


図7. アミン化合物を分類した場合の%inhibitionとの関係

また上記結果について、分子内にN+(カチオン性の窒素原子)を含むか否かで化合物を分類し、ヒット率を比較した。 μ 受容体をターゲットとしたリガンドでは、N+のファーマコフォーが重要であることが示唆されているが、今回のVSでは、そのような条件は加味していない。そこで、もし、N+を含む分子だけを選択していた場合には、どれだけのヒット率が得られたのか、そして、N+を含まない分子でもヒット化合物が得られているのか確認を行うことにした。ここで、N+を含む分子を「アミン化合物:Yes」、N+を含まない分子を「アミン化合物:No」とした。

表2. アミン化合物を分類した場合の結合親和性試験結果

(単位:個数)

%inhibition	アミン化合物		計
	Yes	No	
$90 \leq x \leq 100$	8	0	8
$80 \leq x < 90$	5	2	7
$70 \leq x < 80$	4	8	12
$60 \leq x < 70$	6	5	11
$50 \leq x < 60$	3	7	10
$x < 50$	99	252	351
計	125	274	399

仮に、%inhibition \geq 50をヒットとすれば、

ヒット率	20.8%	8.0%	12.0%
------	-------	------	-------

上記結果より、「アミン化合物:Yes」の場合では、20%と高いヒット率であることが分かった。また、

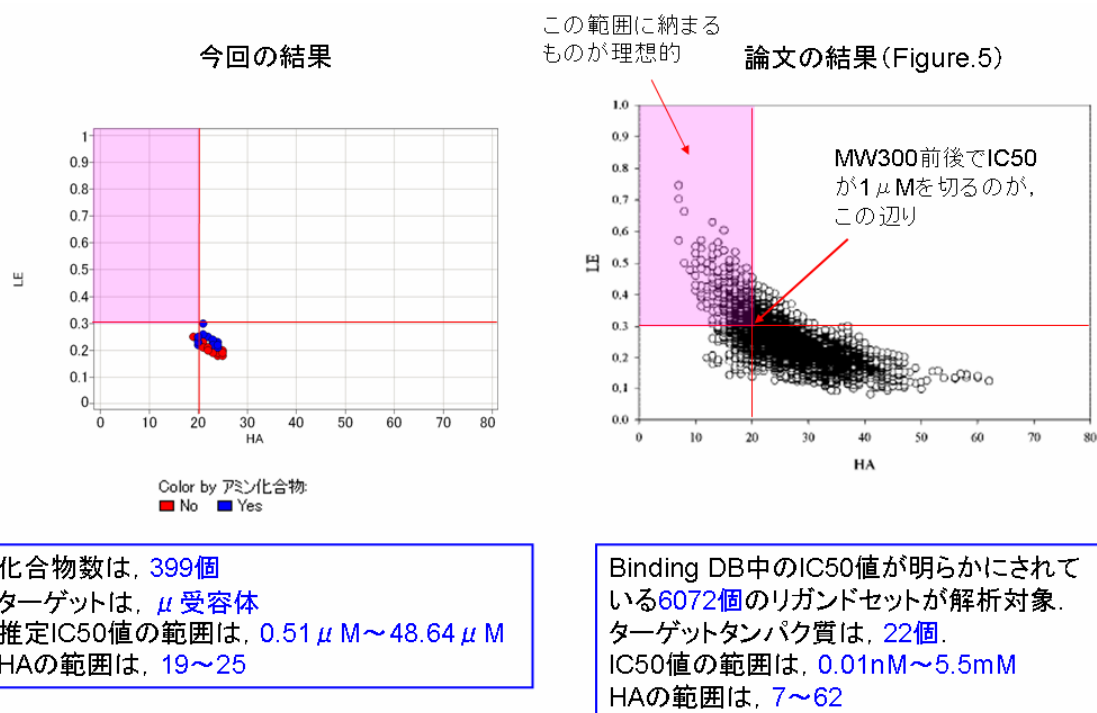
「アミン化合物:No」の中でも、%inhibitionが80~70と高活性が期待されるものが幾つか見出されていることが分かった。

次に、ヒット化合物の推定のLE (Ligand Efficiency) 値を算出し、リード化合物の候補として有望であるか確認を行った。解析手順としては、以下のとおりである。

1. 各化合物の%inhibitionの値を基に、ロジット変換により推定のIC50値を算出した。
2. 次の式に従い、推定のLE値を算出した。 $LE = pIC_{50} / \text{Heavy atom (HA)}$ の数
3. LE解析が実施された次の論文の結果と比較を行った。この論文では、様々なターゲットタンパク質(22種のタンパク質)に対するIC50値が明らかにされている6072個のリガンドセットを解析対象としている。J. Med. Chem. 2008, 51, 2432-2438

上記手順に従って解析を行った。以下に解析結果を示す。

図8. LE解析結果



一般的に、 $HA \leq 20$ かつ $LE \geq 0.3$ の化合物が理想的なリード化合物と言われる。今回の結果を見ると、この条件を満たす化合物が2つと、この範囲に隣接しているものが幾つかあることから、リード化合物の候補として良好なものがヒットしてきているものと思われる。

上記結果は、推定のIC50値を用いて解析を行ったので、今後、濃度依存性試験が得られた後に、再度解析を行いたいと考えている。

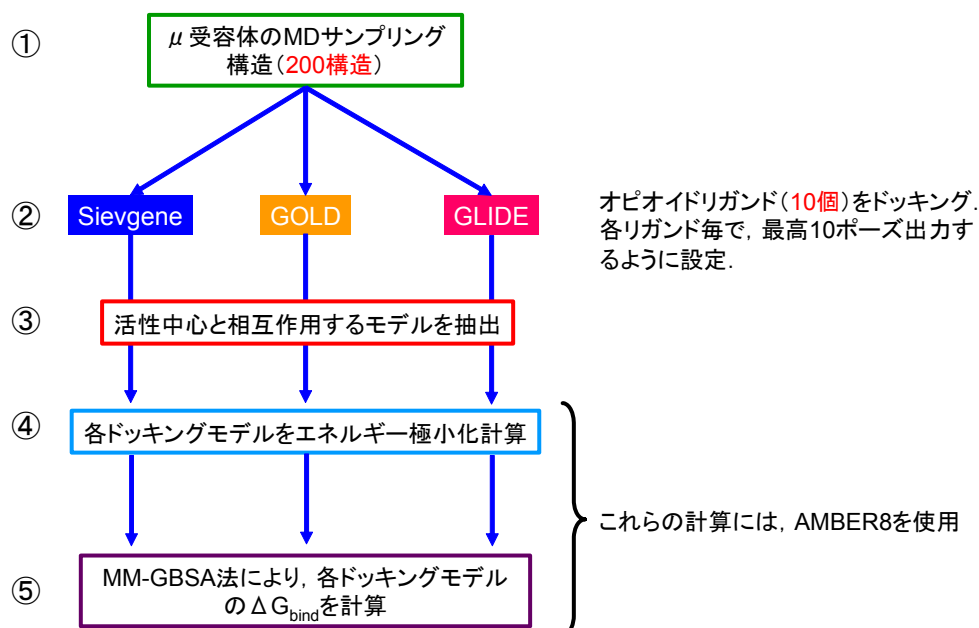
4. オピオイドリガンドのドッキングスタディ

図4で示したオピオイドリガンドを用いて、ドッキングスタディを行った(この内alfentanylについては、1,4-Dihydro-tetrazol-5-oneの部分構造のパラメーターがAMBER内になかったので除外した)。また、 μ 受容体モデルとしては、Mosberg等により報告されている活性型の μ 受容体モデルを用いた。このモデルについては、次のウェブサイトから入手した

(<http://mosberglab.phar.umich.edu/projects/proj2.php>)。GPCRにおいては、アゴニストが結合することで受容体が活性化され、構造変化が起こることが示唆されており、Mosbergモデルではこの説を志向したモデルとなっている。本研究では、上記2. の取り組みにおいて、ホモロジーモデリングにより μ 受容体モデルの構築を行ったが、鑄型構造に用いた β 2ARはアンタゴニストが結合した不活性型モデルであり、構築した μ 受容体モデルも不活性型であると想定される。また、検証の結果最も精度が高かったMOE MODEL1の3900psのMD構造に対するオピオイドリガンドのドッキングポーズも確認したが、活性中心と思われるAsp149(TNH3)との相互作用が見られない等、妥当なドッキングモデルが得られなかった。

以上を踏まえ、 μ 受容体アゴニストであるオピオイドリガンドのドッキングスタディを行う上では、活性型の μ 受容体モデルを用いる方がより適切と判断し、Mosbergモデルを用いて検討を行うこととした。本検討の流れとしては以下の形で行った。

図9. ドッキングスタディの流れ



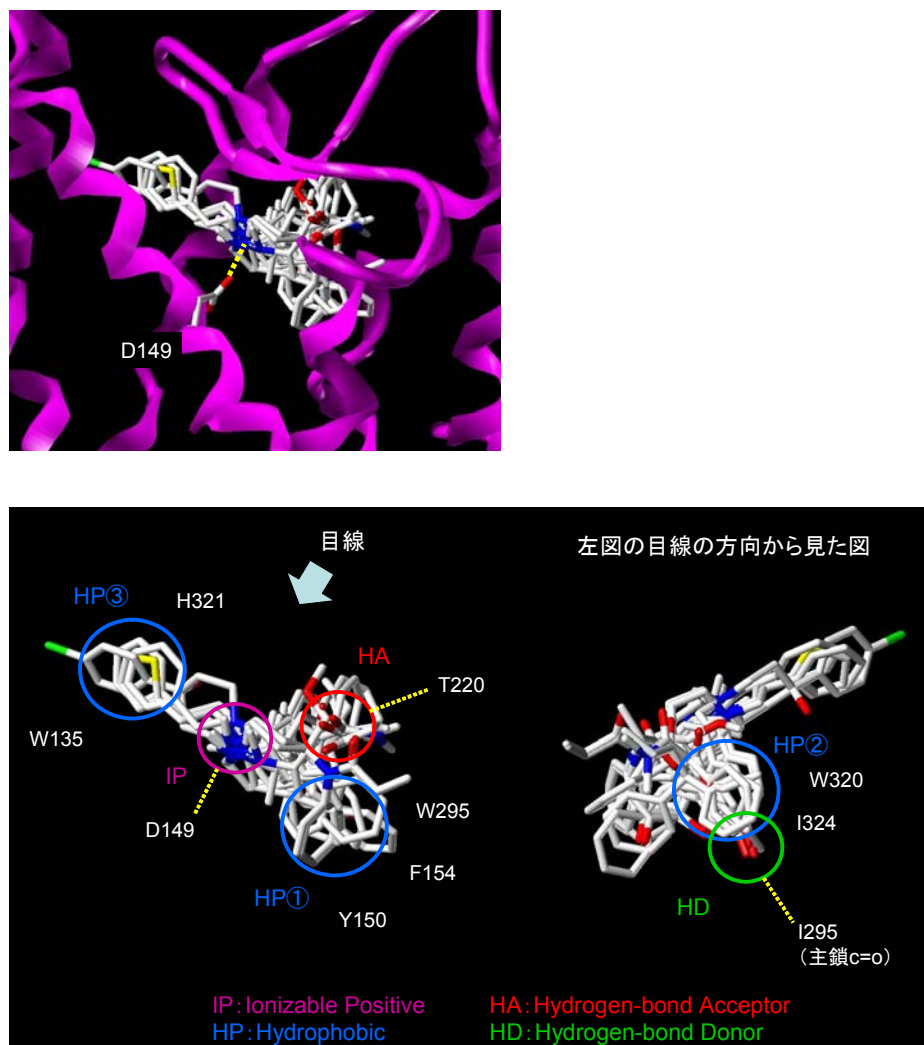
- ① タンパク質の動的挙動を考慮するため、上記2の取り組み(図3)と同様の条件でMDシミュレーションを行った。4nsecのシミュレーションを行い、構造が平衡化した2nsecから4nsecの間で10psec毎にサンプリングを行い、計200構造をサンプリングした。
- ② ドッキングソフトウェアとして、SievGene、GOLD、S社のドッキングソフトを用いて、①で得られ

た200構造の μ 受容体に対して、オピオイドリガンドをドッキング。多数のドッキングポーズを得た。

- ③ ②のドッキングポーズの内、 μ 受容体の活性中心と思われるAsp149(TM3)と相互作用を形成するドッキングモデルを抽出した。
- ④ ③で抽出されたドッキングモデルについて、AMBER8のsanderモジュールを用いて、エネルギー極小化を行い、構造を精密化した。
- ⑤ ④の精密化後の構造を用いて、MM-GBSA法により各ドッキングモデルの簡易的な結合自由エネルギー(ΔG_{bind})を算出した。

上記検討の結果を基に、各リガンド毎で ΔG_{bind} が良好なドッキングモデルを比較し、以下に示すような共通ファーマコフォーを推定した。

図10. オピオイドリガンドのドッキングモデルおよび推定ファーマコフォー



以上、1～4の取り組みの結果、以下の成果が得られた。

1. EM-1立体配座集団

2と3. μ 受容体に対して結合活性を有するリガンド

4. オピオイドリガンドの推定ファーマコフォー

今後の取り組みの予定として、1と4の情報を基に、EM-1の結合配座およびファーマコフォーの推定に関する検討を進め、EM-1の非ペプチド化のための具体的な検討を進めていきたいと考えている。また、2と3の検討により得られた活性リガンドについても、4の検討で得られた情報と合わせて、今後構造最適化を行い、より μ 受容体に親和性の高いリガンドの創製に望みたいと考えている。また、既知リガンドがなく生理活性ペプチドの情報しかない場合を想定して、4の検討の情報を使わずにEM-1の結合配座およびファーマコフォーを推定する手法についても検討を行いたいと考えている。

(3)創薬開発への応用促進に向けた技術開発

【研究内容】

計算の高効率化と高速性が発揮できるアルゴリズムとプログラムの実装法の開発、専用ボードの利用、リガンド・データベースや計算結果を整理したデータベース等の開発に着手する。

また、本開発研究のチーム間だけでなく、創薬メーカーと研究協力を行って具体的な創薬実証研究を検討する。

【研究成果】

(i) 分子シミュレーション高速化技術の開発 [BIRC 集中研、阪大分室、情報数理研分室]

我々は、分子シミュレーションソフト myPresro を開発してきたが、計算の高効率化と高速性が発揮できるアルゴリズムとプログラムの実装法の開発、専用ボードの利用を行った。

我々の開発してきたソフトウェア一式は、大阪大学、及び、2008年に整備された経済産業省ライフサイエンス統合データベースポータルサイト「medals」において、一般に公開している。当サイトでは、経済産業省関連機関によるライフサイエンス分野のプロジェクトの成果物を集めたものだが、ダウンロードできる成果物としては、myPresto を含む2つのみである。

MEDALS: 経済産業省ライフサイエンス統合データベースポータルサイト

MEDALS データ更新で2件を新たに追加し、プロジェクト成果報告書ダウンロードページを公開しました。[08.05.20]

経済産業省ライフサイエンス統合データベースポータルサイト

このウェブサイトは経済産業省統合データベースプロジェクトのポータルサイトです。"MEDALS"は、METI Database portの略であり、METIとは経済産業省 (Ministry of Economy, Trade and Industry) の意味です。これまで経済産業省関連機関のライフサイエンス分野の研究開発プロジェクトが実施されてきました。それらの成果物が効果的に利用されることを目的に、果物(データベース、ツール等)の情報を整理・提供していきます。今後も内容を更新・追加していきます。●[トップ](#)

データベース一覧 全リスト(40件)
 ・様々なデータベースの情報がわかります
 DNA-グラム(17件) 蛋白(14件)
 蛋白質(22件) その他(22件)

統合データベース
 ・各種のデータも統合したDBです
 ・遺伝子統合データベース (H-keiDB)
 文庫統合データベースプロジェクト (WDB)

ツール一覧 全リスト(33件)
 サイト内検索

MEDALS METI database portal for life science

Home > ダウンロードページ

全ダウンロードリスト (2 / 2 件)

名称(サイトへのリンク) 解説ページへのリンク	主な対象データ	生物種
PubMedScan 解説ページ	PubMedに登録されている文献	なし
myPresto 解説ページ	タンパク質-立役構造	全生物種

myPresto

myPresto (Medicinally Yielding PProtein Engineering SimulaTOR) は、医薬品開発支援のために作成された分子シミュレーション計算のプログラム集です。このウェブサイトは、フリープログラム **myPresto** のダウンロードを行うためのページです。

A Program suite for SGDD (Structure Guided Drug Development)

→ [DOWNLOAD](#)

cosgene: COnformation SamplinG ENginE (分子構造探索エンジン)

このエンジンは、マルチカノニカル分子力学等、最新の機能を持つ分子力学 (MD) ツールによって、膜蛋白質や低分子化合物との複合体を含む生体高分子系に対して、高能率の構造探索や自由エネルギー計算、結合自由エネルギーの算出を行うためのものです。(Fukunishi, Y., Mikami, Y., Nakamura, H. (2003) J. Phys. Chem. B. 107, 13201-13210)

【研究成果】

①目標・課題

これまで CPU の高速化はクロックアップで実現してきたが、既に原理的な限界に近づいてきている。そこで浮動小数点演算に優れたコアを複数持つ(マルチコアの) APs (Accelerator Processors) が登場した。APs は特定の数値計算においては 100 倍の高速化を実現し、GPU (Graphics Processing Unit) の様な家庭向けゲーム機により低価格化を実現している。

その一方で、APs のアプリケーションは新しいノウハウと新しいアーキテクチャに合わせたプログラム開発を要求する。GPGPU (General Purpose GPU) への移植における既存アプリケーションのリファクタリング(コードの並列化対応)は多大なコストがかかるが、今後 CPU 等のマルチコア化が進むにつれて並列対応アプリケーション数も増加するものと予想される。したがって、今後シングル CPU では到達できない高速環境が実現する事が期待できる。

新薬の開発には多大なコストが掛かり失敗も多い為、コンピュータによる薬物候補探索の省力化は利益が大きい。数ある手法の中でも、分子ダイナミクスや熱力学的性質の計算が可能な分子シミュレーションは最も有用な手法である。例えば、数百のタンパク質-リガンド系の結合自由エネルギーを低コストで計算できる事は、創薬への応用に大きな利益をもたらす。そういった意味でドッキングプログラムなどの APs への移植は期待が大きい。

②Sievgene の高速化

したがって、APs の利用によるドッキング計算の高速化が目標となる。そこで薬物ドッキングプログラム—myPresto (Sievgene)—に対して APs の導入をはかり、超高速の薬物ドッキング計算が可能なシステムを構築する事を検討した。導入する APs は、汎用アクセラレータである GRAPE-DR (SING チップ)、GRAPE-DR に比して速度性能は低下するがより自由度の高い GPGPU (NDIVIA 社製 GeForce, Tesla シリーズ) とし、其々のプロトタイプを開発して動作検証を行った。

③ GRAPE-DR への移植

GRAPE-DR (SING) は、東京大学で開発された 1 チップで 512G FLOPS の浮動小数点演算が可能な AP であり、その性能は世界最高を達成している (2006 年 11 月 6 日時点)。GRAPE-DR は SING チップと制御プロセッサ、および外部メモリで構成されている。SING チップにはインストラクションを同時実行する 512 個の PE (Processing Element) が存在し、各 PE は 4 データを同時演算可能 (SIMD) である。

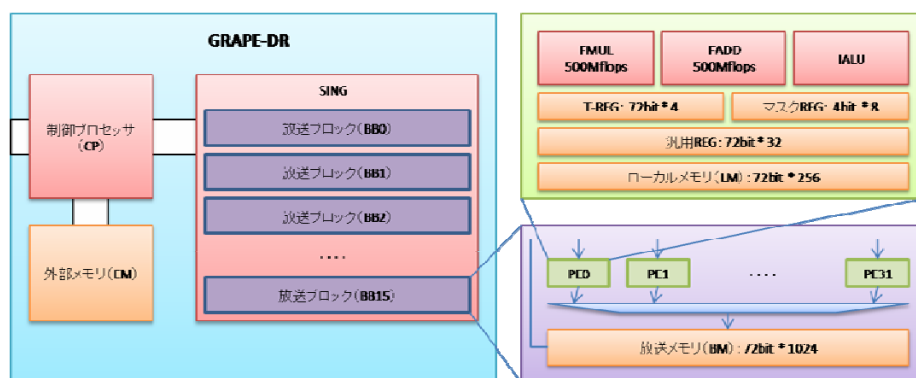


図 1 GRAPE-DR

SING チップに計算させるには、SING チップに送り込む命令コード列の作成 (アセンブラで記述)、制御プロセッサの API を介して上記命令コード列を呼び出すホスト上のアプリケーションの作成が必要である。なお、粒子相互作用型のアプリケーションについては、一部フレームワーク化されている。基本的な並列型アプリケーションのフローは下記の通りである。

- Host 側でデータを用意しておき、EM、BB0 の BM および各 PE の LM に送信する
- BB0 の各 PE 上で 4 個のデータに対する処理が同時に行われる (SIMD 演算)
- 処理結果は BB0 の BM 経由で Host 側に返信される

※転送先が BB0 に限定されているのは GRAPE-DR の仕様 (単なる並列計算の場合)

移植対象は、Sievgene から最も単純なアルゴリズムである三原子マッチ、それよりも複雑な 4×4 実対称行列の三重対角化処理である Householder 法の 2 つを選択した。これらのアルゴリズムはフレームワークが用意された粒子相互作用型と異なる為、1PE で 1 行列を処理する単純な並列化として実装した。

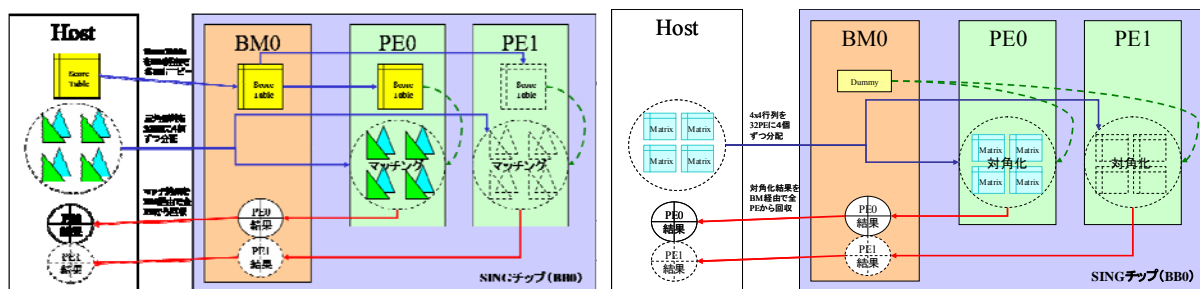


図 2 三原子マッチおよび Householder 法(4×4 実対称行列)

これら Sievgene の一部を GRAPE-DR に移植し、実機で稼働させる事ができた(Opteron 3GHz + GRAPE-DR ボード:1chip 搭載、OS:Linux)。しかしながら、三原子マッチでは約 100 倍、Householder 法では約 34 倍オリジナル(CPU)と比較して遅くなった。これは、GRAPE-DR の仕様による所が大きく、1) 狭いメモリバンド幅、2) 共有メモリ量不足、3) 実機での命令数制限(480 命令以内)で、用意された 512PE を殆ど活用できない事が原因である。特に 3)の問題は大きく、Householder 法以上の規模のアルゴリズムを移植するのは難しい事が判明した。

④GPGPU(General Purpose GPU)への移植

GPGPU(NDIVIA 社製 GeForce, Tesla シリーズ)のフレームワーク CUDA は、NDIVIA 社が提唱する GPU を汎用ベクトルプロセッサとして活用するためのハードウェアとソフトウェアの統合環境である。ハードウェアには汎用コンピューティングが可能な GPU のみならずそれを支援する周辺機構、ソフトウェアにはそのハードウェア上で利用可能な C コンパイラを中核とした開発環境およびドライバソフトが含まれる。

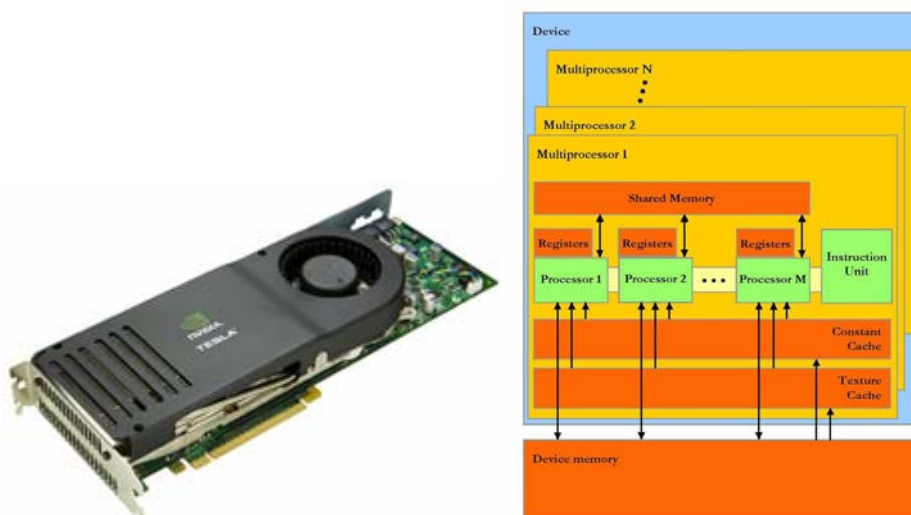


図 3 Tesla C870(外装および模式図)

CUDA でプログラムを実行する際、GPU は非常に多数のスレッドを並列実行できる計算デバイスとして扱われ、CPU(Host)のコプロセッサとして動作する。その際、スレッド群はブロックとして管理され、ブロック群はグリッドとして管理される。なお、共有メモリは同ブロック内でのみ有効である。基本的な並列型アプリケーションのフローは下記の通りである。

- a) Host 側でデータを用意しておき、デバイスメモリおよび共有メモリに送信する
- b) 各スレッドが共有メモリ上のデータに対して処理を行う(ブロック単位で同期可能)
- c) 処理結果はデバイスメモリ経由で Host 側に返信される

移植対象は、Sievgene から Householder 法、Bisection 法、Inverse iteration 法、およびこれら 3 関数からなる固有値計算を選択した。アルゴリズムは、Householder 法では 1 スレッド当たり 1 行列、Bisection 法では 1 スレッド当たり 1 固有値、Inverse iteration 法は 1 スレッド当たり 1 固有ベクトルが計算される様に並列化した。また、上記関数での共有メモリ配置についても、(オリジナルと比較して)メモリ使用量が少なくなる様に構成した。

固有値計算は、一旦デバイスメモリへ入力行列をデータ転送した後で、上記の 3 関数を逐次処理する様に実装した。その為、デバイスメモリ上の行列数が一定である様に各関数のスレッド数、ブロック数を調整する必要がある。今回は指定したスレッド数からブロック数を算出する様に実装した。

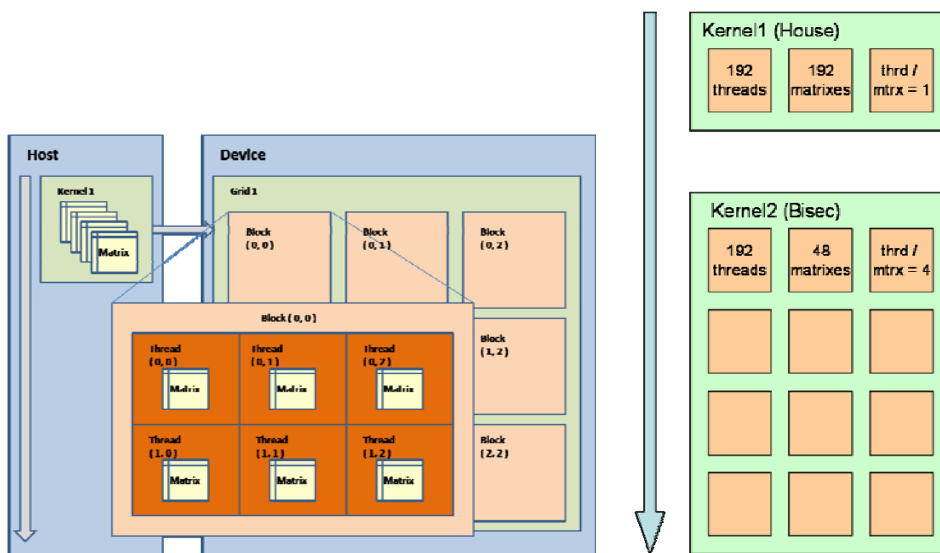
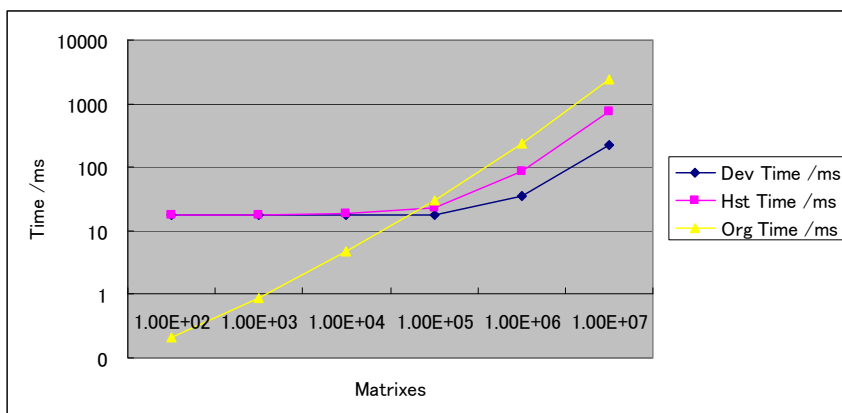
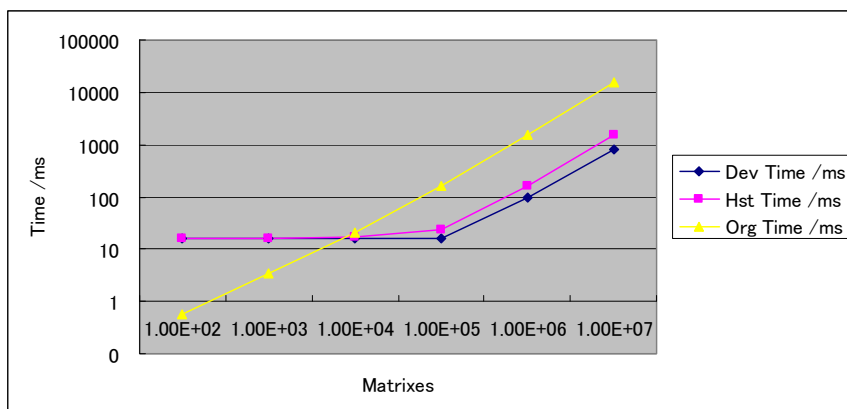


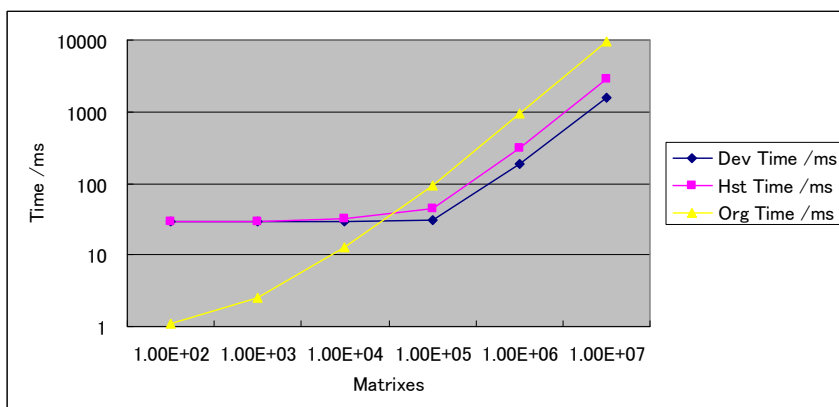
図 4 Householder 法(左)および固有値計算(右) (4 × 4 実対称行列)



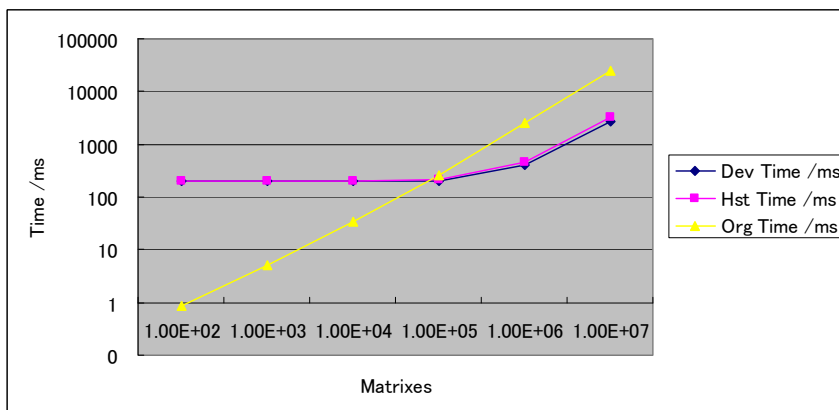
Householder 法



Bisection 法



Inverse iteration 法



固有値計算

図 5 処理時間計測(実行時間-4x4 実対称行列数:192threads, 4096blocks)

これら Sievgene の一部を GPGPU に移植し、実機で稼働させることができた(CPU: Xeon E5462 2.8GHz Quadcore + GPU: Tesla C870、CentOS 5.1/64bit)。Householder 法では約 3.3 倍、Bisection 法では約 10.2 倍、Inverse iteration 法では約 3.4 倍、これらを統合した固有値計算では約 7.8 倍オリジナル(CPU)と比較して高速化する事ができた(入力行列数が $1e7$ (1千万) 個の場合)。高速化の程度が最大で約 10 倍程度に留まっているのは、1) 1 スレッド当たりの共有メモリ/レジスタ使用量が多い為に立ち上げ可能なスレッド数が抑えられる事、2) スレッド当たりのデータが多い為にメモリバンド幅を圧迫する事、が原因であると考えられる。

【結論】

蛋白質—化合物ドッキングソフト(sievgene)の一部を専用計算機(マルチコアアクセラレータ: APs)である GRAPE-DR 及び GPGPU に移植し、高速性能を発揮することができた。しかし、蛋白質—化合物ドッキングソフトは、多種類の複雑な演算から構成され、一部分を高速化しても全体が速くならないこと、複雑な演算の中には、APs になじまないものも多く含まれ、既存ソフトの単純な移植では非効率であることが分かった。今後は、計算機創薬過程全体を考えたときに、どう APs による高速化を図るかを検討していきたい。

(ii) データベースの構築と標準化、高度化 [BIRC集中研]

(a) 化合物データベースの構築、標準化、高度化

今まで、AMBER の力場で低分子の立体構造を予測してきたが、スルホンアミドなどいくつかの構造で、間違っただけの立体構造が得られることがわかった。問題が生じるのは、POOH や SO₂ 基に N が結合している構造が平面にならない、N を含む環構造が平面にならないなど、N 原子の π 軌道と隣接する原子の π 軌道の相互作用を過小評価している例が多く見られた。これは、他の多くの市販分子力場ソフトでも同様の傾向が見られた。

新しい力場の追加は、影響範囲が大きいので、力場の割り当て方の改良などを加え、正しい分子構造が得られるように化合物データベースの構築手法を改良した。現在はこれらの改良ですべての分子が正しく計算されるかテストをしているところである。2007年に市販された化合物400万件の3D 化を進める予定であったが、正しい構造が得られるように手法を改良することに重点を置いた。

また、電子カタログで配給される化合物データベースの元のデータには、水素原子が記載されていないので、水素を付加して完全な分子を組み立てる必要があるが、水素の付加には多少の任意性があり、我々のソフトウェアは、一般に用いられるツールよりも精度は高いものの従来の手法では、水素が誤って付加されることが、しばしば見られた。このような点を改善した。

(b) 蛋白質の動的構造データベースの構築

蛋白質は水中で動的に構造変化しており、またリガンドの結合によってもその立体構造が変化する。ドッキング計算の精度を上げるためには、この動的構造変化を効果的に取り込むことがかぎになる。我々は、様々な構造を効率的にサンプリングし、種々の温度でのカノニカル・アンサンブルを再構築できる Force-biased Multicanonical Molecular Dynamics (FB-McMD)法を既に関し、myPresto プログラム・パッケージに加えている。この手法によって蛋白質の動的構造を正しくとらえようとする場合、主成分分析を利用してその動的特徴を抽出する手法が有効である。ここでは、単に主成分分析による動的構造の解析をするだけでなく、全体的な動きを反映する固有値の大きな主成分を抽出し、それを基にして蛋白質座標データを圧縮・復元する手法の開発を行った。開発したプログラムでは、座標情報をもとの7%以下の容量の情報に圧縮することができた。主成分の一部のみを

利用するため、構造全体としての座標復元精度は悪くないものの、側鎖先端部分に位置する原子座標の復元精度は低下する傾向がある。しかし、全体の動的情報を保持しつつ、アンサンブル・データの圧縮が可能であり、次に述べる動的構造を反映したドッキング計算などに有効な手法と考えられる。

(iii)創薬の為の水溶解度予測

〔BIRC 集中研、アステラス製薬分室〕

本研究では、BIRC 集中研、アステラス製薬分室で、下記のように役割を分担しながら連携して行った。

- ・溶解度(logS)データの収集、入力分子に対し、類似の化合物の logS データを推算の学習セットとして用いる手法(SimQSSR)をアステラス製薬分室で開発した。
- ・物理化学的記述子の開発、SimQSSR 法を発展させ、入力分子に対して化合物の類似性に応じて logS データに重みをつけて学習セットとして用いる手法を BIRC 集中研で開発した。この手法はアステラス製薬分室で、実地に検証、問題点を BIRC 集中研にフィードバックしながら手法の改良を行った。
- ・非特異的に相互作用する分子(アグリゲーター)のデータ収集、logS を用いたこれらの解析を BIRC 集中研で行った。

(a) 創薬現場で有効な水溶解性の新規予測手法(SimQSSR 法)の開発

〔アステラス製薬分室〕

創薬を指向する以上、化合物は薬剤として機能できる物性条件を満たす必要がある。物性条件の中でも重要な溶解度に監視、医薬品に求められる測定域での予測技術を開発した。

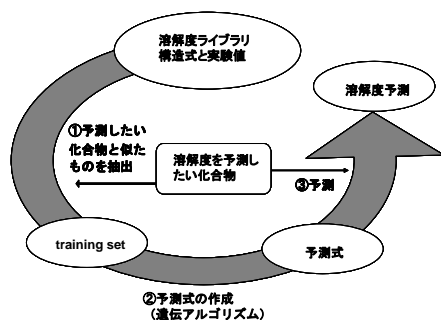
【研究成果】

1) 課題

創薬の現場では、「10ug/L以上の水溶解性の有無を確実に予測する」か、「1ug/L~100ug/L付近の溶解性を持つ化合物について、溶解度の相対比較を行う」ことが必要である。上記を満たす予測計算のアルゴリズムを考案し、有効性を確認、実用化する。

2) SimQSSR法の考案

一般的なQSPR (Quantitative Structure Property Relationship) アプローチでは、上記に必要な予測精度は得られない。QSPRをベースに化学構造上似ている化合物群で予測式を作成することが有効と考え、Similarity Quantitative Structure Solubility Relationship = SimQSSR法を考案した。



SimQSSR法の概念図

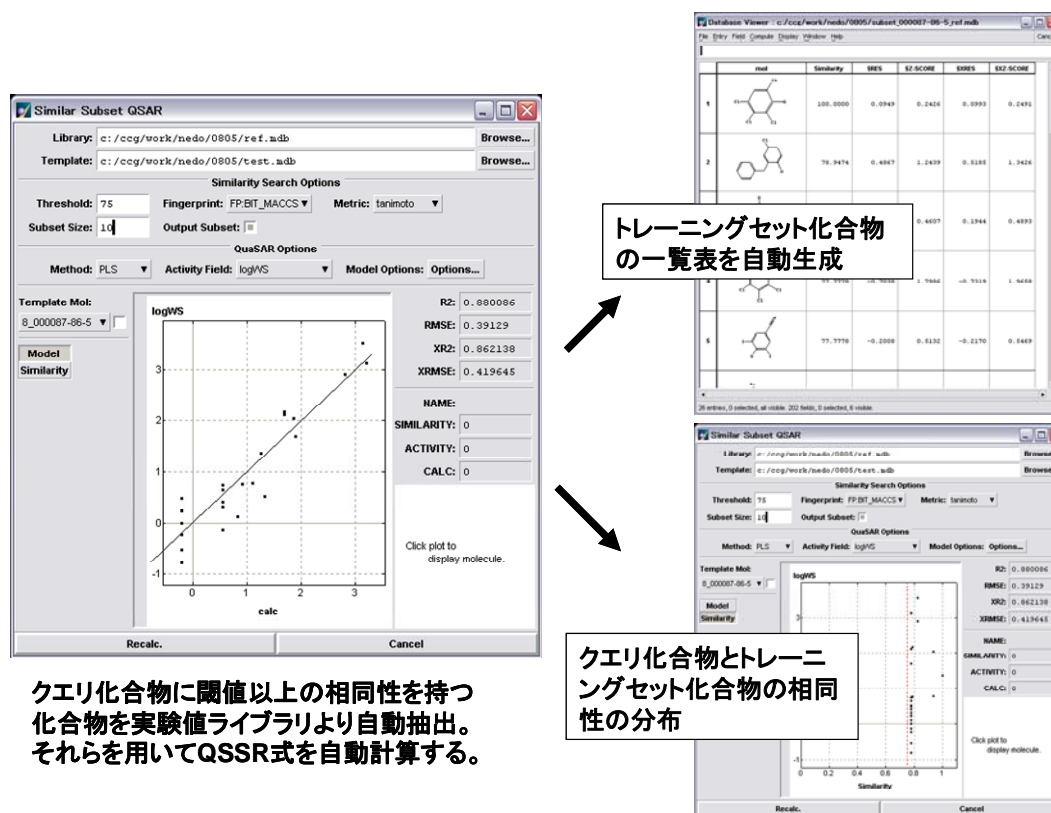
3) SimQSSR法の有効性

溶解性予測の方法のうち、Hou, T.Jら(2004)の簡易計算式や一般的なQSPR式による予測精度との比較を行った。一例としてHouらの方法(MOEに実装)との比較結果を示す。SimQSSR法は、予測式の立式に十分な実験値が得られるケースにおいては、有意に優れた予測結果を示した。

4) SimQSSR法のプログラム実装

CCG社のMOE環境上に、SimQSSR計算の実装を行った。予測式作成に用いる化合物セットは、閾値を与えることで自動抽出され、化合物記述子は遺伝的アルゴリズムで選択する。また、アルゴリズムの性質上、予測式を作成するトレーニングセットの数、予測対象化合物との相同性分布、実験値の幅により、予測精度が大きく変動することから、アプリケーションプログラムの開発に於いては、トレーニングセットの内容を簡便に確認するための機能を実装した。

5) 現時点での研究の効果



化学構造に関連の深い実測値データが豊富にあれば、それを自動検索して精度の高い溶解性予測を行うことが出来るようになった。また、このアルゴリズムは化合物の構造式との関連が予想される物性値予測にも応用が可能である。つまり、SimQSPR (Similarity Quantitative Structure Property Relationship)法として拡張が可能なアルゴリズムを実用化することができた。

6)研究の今後

溶解度予測技術の開発を終了し、本法を化合物データベースの充実に活用する。

(b) 化合物の水への溶解度の予測

〔BIRC 集中研〕

【研究内容】

化合物情報で重要なものは、溶解度(logS)と、水—オクタノール平衡定数(logP)である。薬物は水に溶けなければ、体に吸収できず、排泄できず、また水に溶けにくく脂溶性の極度に高いものは、毒性・副作用を示す可能性が高い。水—オクタノール平衡定数も、吸収を左右する要素である。logS の予測計算手法は、数多く提案されているが、その予測値の精度は、低い。これは、通常の有機化合物は常温で結晶なのだが、結晶構造の予測は極端に困難で、いまだ、CO₂ とベンゼン以外の物質の結晶構造は予測できないほどである。そのため、多くのソフトウェアでは、分子の構造を部分構造に分割し、各部分構造の logS への寄与を実験値から統計的に決めておいて、部分構造の寄与を加算することで logS を推算する。我々は、この手法を基本にして、より物理化学的根拠に基づいて、推算値を補正することを試みた。固体から水へ溶ける過程を、固体から有機溶媒へ解ける過程と、有機溶媒から水へ溶ける過程に分割し、これらの過程でのエネルギー変化を、原子電荷の計算と myPresto に組み込んだ GBSA 法によって、物理化学的に計算し、このエネルギーで、従来手法の推算値を補正するようにした。具体的な計算方法は、人工知能の一種であるエラーバックプロパゲーション法を利用した。現在も開発を続けているが、我々の補正により、logS の推算値の精度の向上と、最大誤差の大幅な減少が見られ、実用性の向上が見られている。

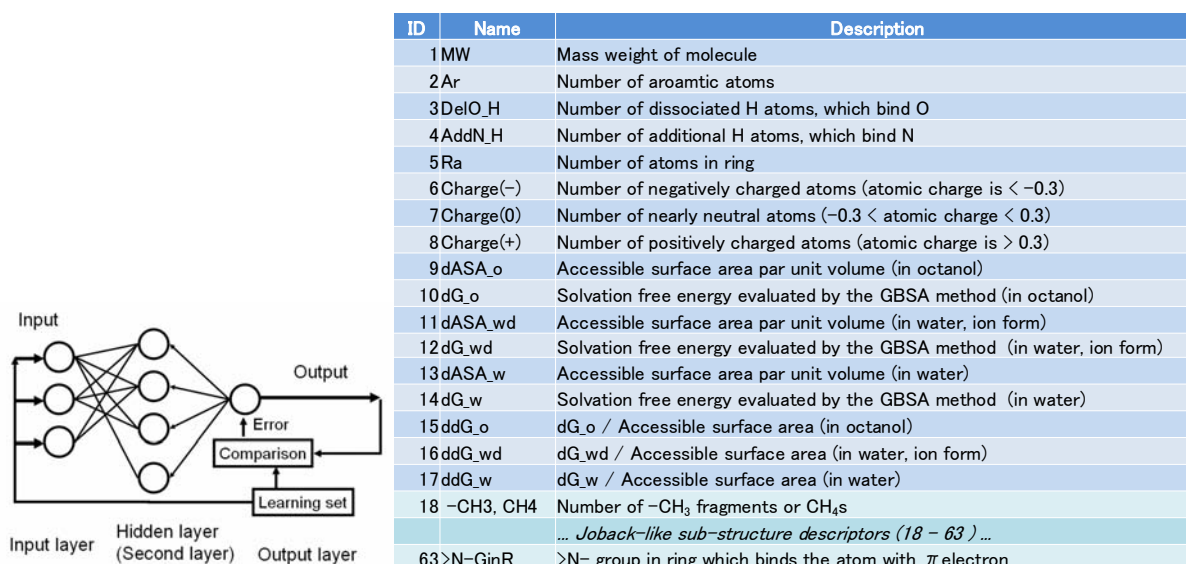
【研究成果】

薬物スクリーニングを行うと、10—30%の化合物が、非特異的に蛋白質に結合することが見出される。そのため、活性測定は、濃度依存性、他の蛋白質への結合性などを測定しなければならない。これら非特異的結合活性を示す化合物は、frequent hitter/aggregator と呼ばれるが、近年、水中でミセルを形成することが報告された。この実験的事実から、我々は、非特異的結合活性を示す化合物は水への溶解度が低いのではないかと、逆に溶解度から、非特異的結合活性を示す化合物を予測・除外できるのではないかと考えた。溶解度(LogS)は Bioavailability や ADME-Tox property に係る重要な指標でもある。

溶解度の予測は難しく、一般的に経験的手法—分子記述子に基づくある種の回帰(重回帰、PLS、NN 等)—によって行われている。現存する予測ソフトの精度は決定係数 0.8—0.9 程度に達しているが、医薬品となる分子の殆どが水に難溶性であり、溶け難い分子の間でのわずかな溶解度の違いが問題となる為、創薬現場では更に高い推算精度が要求されている。そこで精度の高い溶解度推算法として、ニューラルネットを使った新しい予測法(分子記述子および学習法)を開発した。

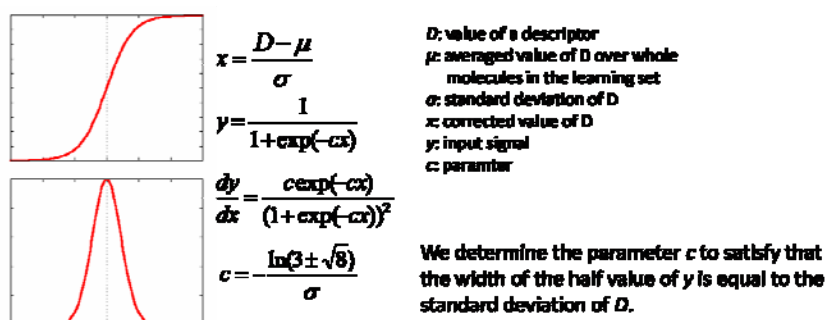
通常、溶解度は、化合物にどのような官能基が含まれるかを数え上げて、これら官能基からの寄与を足し算することで分子全体の溶解度を推算する(原子団寄与法)。しかし、原子団寄与法の推算限界は、既に明らかとなっており、我々はこの限界を超えなければならない。溶解は、物理的な現象だから、物理的なプロセスに分解できる。結晶を個々の分子にばらばらにする場合、同じ体積の分子であっても分子表面積の大きい分子は、結晶が安定である。ばらばらの分子を水に溶かす場合、その溶解の自由エネルギーは分子シミュレーションで計算可能である。そこで、分子シミュレーションで計算できるエネルギー項を原子団寄与法に追加した新しい分子記述子を開発した。また、エネルギーなどの非整数となる記述子を加えたことで、それに適した推算アルゴリズムも新しく開発した。

新しい分子記述子と推算法(学習法)の開発は、MolWorks (Beyond Computing Co. Ltd.)をベースに実施した。MolWorks は、分子記述子として Joback 記述子(原子団寄与法)を使用し、ニューラルネットワーク(NN)を使って推算を行うシステムである。また、学習データセットとして 1300 分子の LogS データを用意した。



エラーバックプロパゲーション(左)および新しい分子記述子

開発した新しい学習法は、学習入力信号の非線形スケーリング (Sigmoid scaling method) 及び重み付き学習法 (Weighted learning (WL) method) である。学習入力信号の非線形スケーリングは、sigmoid scaling curve を使う事により $-\infty \sim +\infty$ の外部入力値を $0 \sim 1$ の NN 入力値にスケールし、学習効率を向上する。これと上記新記述子により、既存システムでは医薬様分子について決定係数 0.4-0.5 程度であったものが、0.8-0.9 程度まで精度向上した。

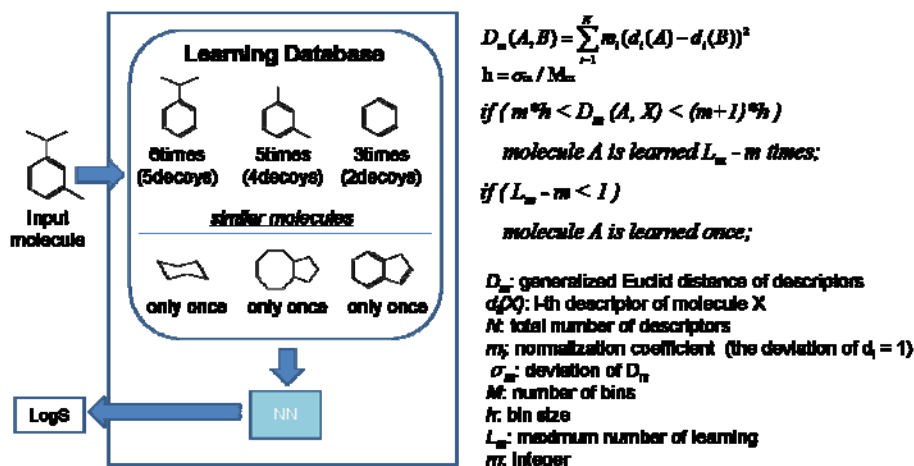


Prediction accuracies for the learning set and the test set (non-WL)

	All descriptors		Joback only (18-63)		no-Joback (1-17)	
	learning	test	learning	test	learning	test
R ²	0.949	0.906	0.938	0.888	0.912	0.811
Ave. Err	0.458	0.629	0.472	0.670	0.599	0.913
Max. Err	1.880	3.363	1.812	5.093	2.702	9.120

Sigmoid scaling(上)および JackKnife-test: [100tests / 1200learnings] x 10trails(下)

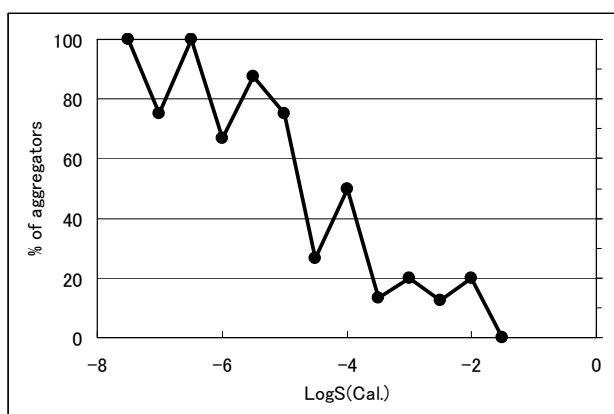
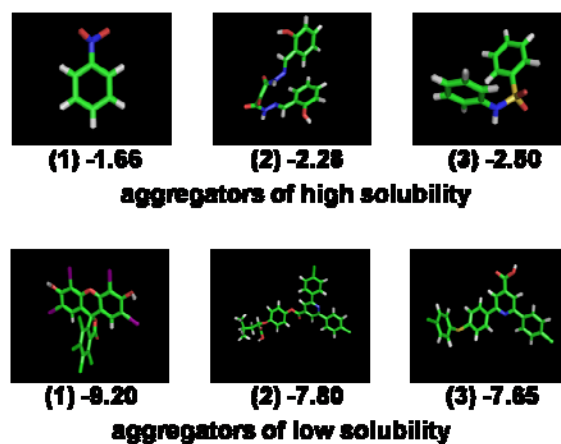
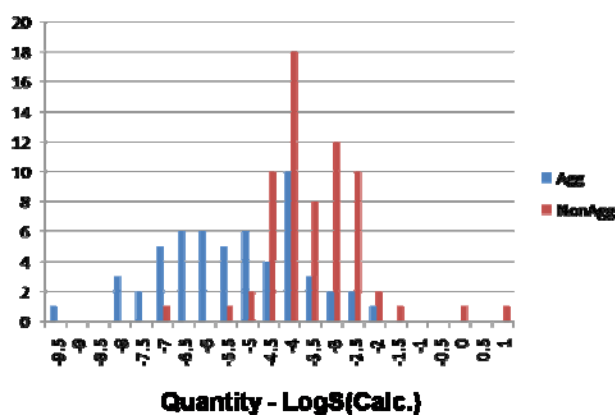
溶解度を計算したい分子に対し、構造の類似した分子の溶解度データは、構造の全く似ていない分子の溶解度データよりも、参考になるはずである。そこで、計算したい分子に類似の分子の溶解度データに重みをつけて学習する手法を開発した。重み付き学習法は、推算対象との類似度を計算して類似度に応じて学習回数を増やす(似ていなくても1回は学習する)という高精度推算法である。なお、類似度は分子記述子の一般ユークリッド距離で計算する。これにより、推算精度を決定係数 0.8-0.9 程度から 0.92-0.93 程度(ほぼ予測限界)まで引き上げる事に成功した。この手法は、従来の原子団寄与法とNNを用いた推算手法では実現できない。



Prediction accuracies for the learning set and the test set (WL)

	Without WL		WL
	learning	test	L _m = 6, M _m = 15
R ²	0.965	0.899	0.934
Ave. Err	0.265	0.480	0.545
Max. Err	0.445	0.583	0.665

Weighted learning(上)および JackKnife-test: [1tests / 1299learnings] x 30trails(下)



アグリゲータの溶解度分布(左)および溶解度上位・下位 3 分子の分子構造(右)

学習データセット(1300 分子の構造データ及び LogS 実験値)を使用して、アグリゲータ 56 分子及び非アグリゲータ 67 分子の溶解度推算を重み付き学習法で行った。アグリゲータは、スクリーニングにおいて非特異的にヒットする低分子化合物の事である。溶解度を横軸としたときのアグリゲータ・非アグリゲータの分布と、割合を図に示す。アグリゲータを除外できればスクリーニング精度の更なる向上・選択性向上が期待できる。推算の結果、アグリゲータと非アグリゲータは、溶解度によって判別することが可能で、その閾値は $\text{LogS} = -5$ 程度と見積もる事ができた。

最後に、代表的な市販ソフト MOE、PipelinePilot と、本手法(重み付き学習なし)の比較を示すが、我々の手法は市販ソフトに対して優位である。重み付き学習の精度は更に高いが、推算速度が遅く多数の分子の推算をしていないため、データは割愛した。

手法	R ²	average error	maximum error
線形重回帰	0.863	0.765	2.48
MOE	0.884	2.26	5.54
PipelinePilot	0.847	2.12	4.97
本手法	0.900	0.637	2.87

(iv) COMBINE 法を用いた薬物の in silico hERG 阻害活性予測モデルの確立 [東レ分室]

【研究内容】

膜蛋白質であるhERGチャネルの構造モデルに基づき、薬物のhERGチャネル阻害活性を合理的に予測する新手法を開発する。

近年、医薬品の開発において薬物の心室再分極および催不整脈リスクに及ぼす影響の評価が活発な研究対象となっている。その背景には、一般に使用されている非循環器用薬剤を服用中に患者が失神もしくは突然死する報告が知られ、その症候は薬剤による致死的な不整脈の惹起が原因であることが明らかにされてきている。さらにその原因となる生体内タンパク質は hERG と呼ばれる心臓の K⁺チャネルであることが分かっている。その結果、催不整脈作用の因果関係が明らかになり承認後に市場から撤退を余儀なくされる、或いは開発が中止されるケースもある。それに対し ICH によるガイドラインが作成されており、今後は新規化合物の開発、或いは既知薬剤の適用拡大において非臨床の評価による催不整脈リスク回避、即ち化合物の hERG 阻害活性の検証は不可避である。現在、化合物の hERG 阻害活性検証の実験的手法は幾種類か存在するものの、throughput 性や信頼性に問題がある(1-3)。

一方、計算化学的手法を用いた分子設計においても hERG に関する問題提起はまだ新しく、徐々に報告が増えてきている状況である。これまで ADMET 予測のための方法論は多数展開されてきたが、特定の毒性に関する評価ではなく、経験則に基づく総合的な評価方法がほとんどである。まして催不整脈リスクの回避、hERG 阻害活性の低減という命題は分子設計の世界では非常に少ない。即ち現状では、適切な hERG 阻害活性検証法は、実験的手法、計算科学的手法の両者ともに充実しているとは言い難い。

従って薬剤開発における分子設計の方針決定やライブラリデザインに際して、計算化学的手法は cost-effective であり果たすべき役割は大きい。このような方法論の開発は急務である。これまでに我々は社内で市販ソフトウェアによる hERG 阻害活性のドッキングシミュレーションによる予測法を確立している。hERG はX線結晶座標が存在しないため、hERG 分子のモデルを用いている。しかし精度及び計算コストの点において問題があると考えている。そこで、本プロジェクトにおいてそ

の問題点を解決すべく、新手法の開発に取り組むものである。

【研究成果】

I. 事前調査

プロジェクトでのディスカッションにより、モデルを用いたドッキングシミュレーションによる予測法は、精度において限界に達しているとの指摘を受けた(図1)。それはモデルの精度の限界によると考えられる。そこで我々は予測法の精度向上のためには、モデルの精度を補完することが必要であると考え、ドッキングシミュレーション以外の種々の方法を導入を検討した結果、COMBINE 法(4)に着目しその検討を開始することとした。

COMBINE 法(Comparative Binding Energy Analysis)はドッキングソフト等を用いて作成した化合物-タンパク質複合体のポーズにおいて、両者の間の相互作用エネルギーを力場計算に基づいて算出し、それを注目するタンパク質の残基毎に振り直し、各値を記述子として定量的構造活性相関(QSAR)解析を行う方法である。所謂 QSAR 解析は、共通骨格を有する化合物群に対して適用するが、本法は共通骨格がない化合物群に対しても適用可能であることと、副次的に複合体ポーズの妥当性を評価できること、という利点があると考えられた。COMBINE 法は活性値を以下の式で定義され、

$$\text{bioactivity} = \sum w_i^{\text{vdw}} u_i^{\text{vdw}} + \sum w_i^{\text{ele}} u_i^{\text{ele}} + C$$

(u_i^{vdw} 、 u_i^{ele} はそれぞれ化合物と i 番目のアミノ酸残基との vdw 相互作用エネルギー、静電相互作用エネルギー)それぞれの係数 w_i^{vdw} 、 w_i^{ele} 、 C を PLS 回帰によって求める。これらの係数を比較することで、どのアミノ酸残基との相互作用が活性発現に重要であるかが特定でき、またその情報は新たな分子設計の指標となる。

II. モデルケースによる COMBINE 法の実施

1. 方法

将来的に myPresto の機能として COMBINE 法を組み込むことを見据えて、適用可能な複数の市販プログラムを用いて COMBINE 法の理解を深める。併せて、市販プログラムよりも精度の高いものを追及する。まず始めに、COMBINE 法の計算科学的手法としての有用性を確認するために、文献を参照し、AchE とその阻害剤をモデルケースとして計算を行った。

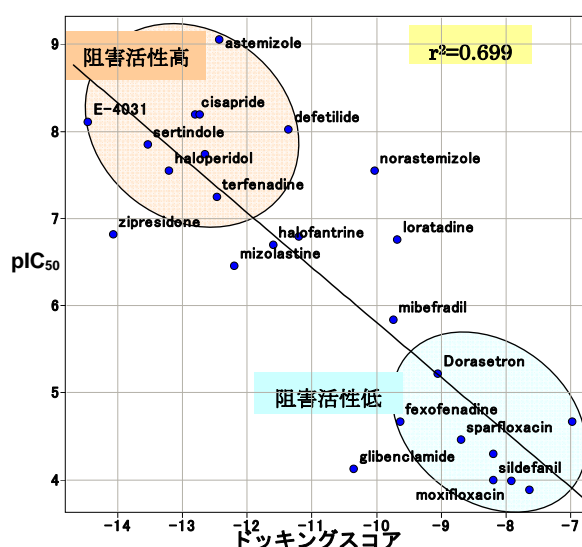


図1：ドッキングシミュレーションによる予測精度

A. 計算の実行

文献(5)にあるタンパク質と化合物群を参照し、以下のフローに従って計算を試行した。(図2)。まず、図2Bの化合物を三次元化し電荷を付与後、ドッキングし化合物-タンパク質複合体を作成した。次に複数のドッキングポーズの候補から適切なポーズを各化合物から一つ選び、複合体の相互作用エネルギーを計算した。さらに PLS 解析による回帰分析を行った。各々の工程は図2Aに示すプログラムで実行した。

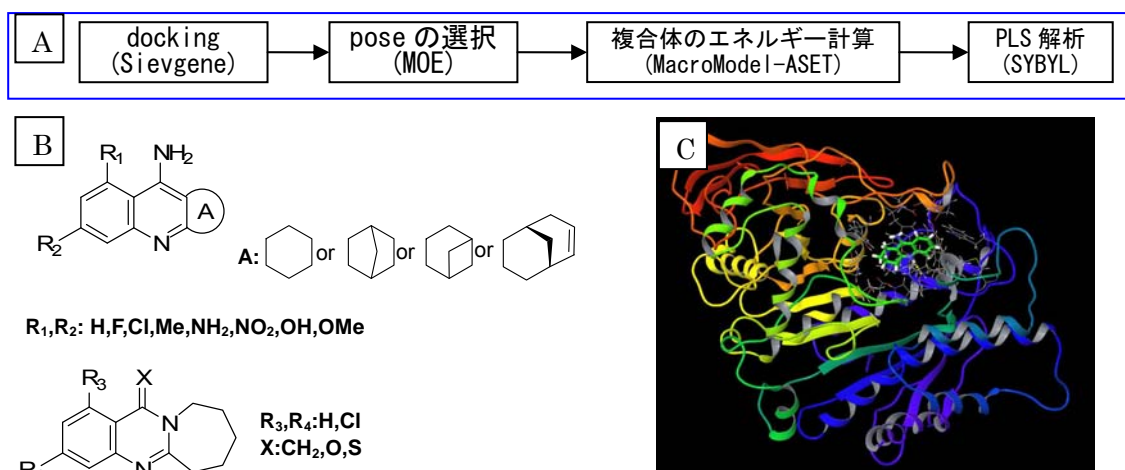


図2 : A 計算のフロー、B 化合物群の一般式、C AChE の X 線共結晶 (1ACJ)

1ACJ の X 線共結晶中の化合物から半径 5.0 Å 以内にあるアミノ酸残基 (一部でも) との相互作用エネルギーを計算し各残基の寄与へ breakdown した。力場は OPLS2005 を用いた。対象残基は以下の 19 残基である。TYR 130、ILE 444、SER 200、HIS 440、TYR 330、SER 122、LEU 333、TYR 334、PHE 331、MET 436、ASP 72、GLY 119、TYR 442、SER 81、PHE 288、GLY 441、PHE 290、GLY 118、PHE 120。

2. 結果

6 個の化合物についてはポーズ全てが共結晶の結合モードとは大きく異なったため除外し、残り 29 化合物について相互作用を計算した。その後、各残基毎の相互作用エネルギーを記述子として pIC_{50} に対して PLS 解析を行った。その結果、 $q^2=0.921$ 、 $r^2=1.000$ という非常に高い相関が得られた。また回帰式の規準化した係数を図3に示す。

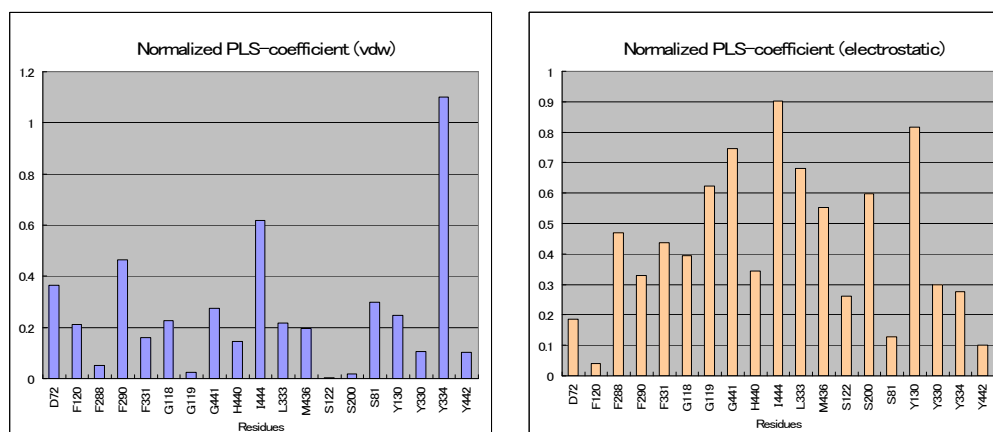


図3 各アミノ酸残基の各エネルギー項の係数（規準化）

3. 考察

A. COMBINE 法の特性と効果について

sievgene によるドッキングポーズから得られた相互作用エネルギー計算値を用いて COMBINE 法を試行した。水分子や脱溶媒和を考慮する方法、また各複合体についてタンパク質残基の側鎖を動かして最適化している方法の報告があるが、今回は溶媒や水分子を考慮せず、またタンパク質は rigid のまま pose を作成し、最もシンプルな系で計算を行った。その結果 QSAR として非常に高い相関が見られた。本ケースでは、用いたタンパク質の結合サイトが比較的狭いものであったこと、用いた化合物が平面的かつ類似性の高いものであったため、考慮すべき対象アミノ酸残基が効率よく網羅され、高い相関が得られたと思われる。図3において、高い値を示す残基は化合物から空間的に比較的近い位置にあるが、それらよりもさらに空間的に近い残基も存在する(図4A)。このことは、必ずしも空間的に近い残基のみを考慮すべきではなく、かつ計算結果から相互作用を考慮する上で重要なアミノ酸残基を抽出されたことを示している。また、 pIC_{50} 値とドッキングスコアとの相関係数 r^2 は約 0.4 であり(図4B)、COMBINE 法を適用することではるかに高い相関が得られること分かった。このことから、COMBINE 法によりドッキングポーズの再評価が行われ、活性予測の精度を向上させることができると期待される。通常ドッキングスコアは無次元の離散的な値であり、エネルギー値から直接導かれるものではない。従って生物活性の値とドッキングスコアとの間に極めて高い相関が見出される必然性はない。しかし、化合物がどのように標的タンパク質に結合するかを推定する必要度は高く、両者の間の相関は高い方が望ましい。そのためにはドッキングポーズのモデルの精度や妥当性の判断を行う方法論が必須である。その点において、今回検証した COMBINE 法はモデルの精度・妥当性を判断する有用な方法論である。

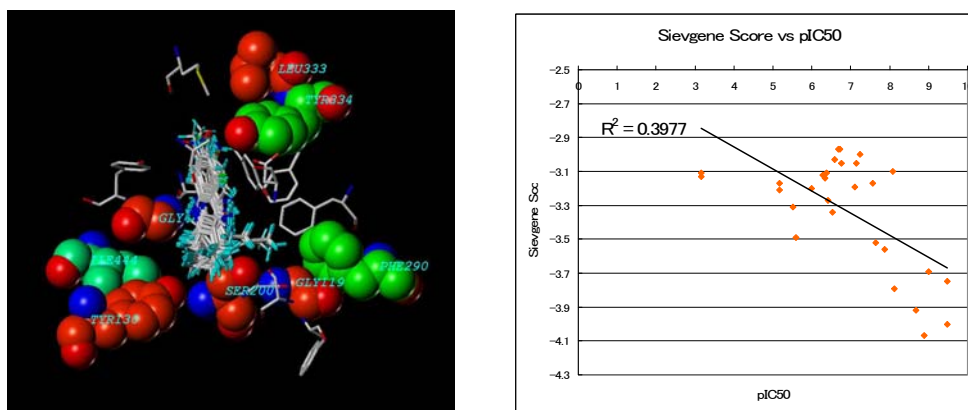


図4 A 緑：vdw 項の相互作用の大きい残基。赤：electrostatic 項の相互作用の大きい残基。青緑：両者の相互作用の大きい残基 図4 B：ドッキングスコアと活性値のプロット

B. 他の方法論との比較—CoMFA との比較—

CoMFA (Comparative Molecular Field Analysis) は、化合物分子の立体構造を直接反映した QSAR 解析を行うために開発された計算手法であり、Tripos 社 SYBYL[®] の QSAR 手法として最も汎用されているものである。化合物のみを取り扱い、標的となるタンパク質は扱わない。前項の COMBINE 法との比較を行うために、同じ化合物群を用いて CoMFA の計算を行った。結果の比較を表3に示す。

両者の q^2 、 r^2 、SE の値を比較すると、COMBINE 法の統計上の妥当性は CoMFA と同等以上に高いことが分かった。このことから、COMBINE 法は汎用されている QSAR である CoMFA 以上の予測精度を有すると考えられる。

表3 COMBINE 法と CoMFA との比較

	q^2	r^2	SE
COMBINE	0.921	1.000	0.000
CoMFA	0.874	0.991	0.163

4. 結論

以上の検討から、COMBINE 法は myPRESTO に組み込むことで機能向上を図ることが期待される有用な方法論である。

III. COMBINE 法による、hERG 立体構造モデル(ホモロジーモデル)を用いた化合物の阻害活性予測法の確立

1. 方法

タンパク質は社内でホモロジーモデリングにより既に構築済みの hERG の立体構造モデルを利用した。化合物は文献調査により、なるべく構造が偏らないように、かつ、 pIC_{50} 値が確定しているもの(文献によってばらつきがない)を 100 個選定した。作業は可能ならば市販プログラムを用いて行った。計算全体の流れを図5に示す。

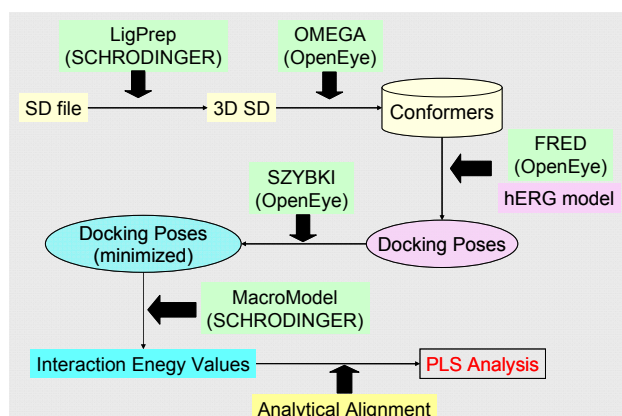


図5 計算全体の流れ。矢印は作業の流れを示す。

緑色囲いは市販プログラム名を示す。

A. hERG 立体構造モデル

構築したモデルを図6に示す。hERGは4回対称の四両体から成るが、ホモロジーモデリングにより構築した単量体を回転・並進行列処理により四両体を構築した。

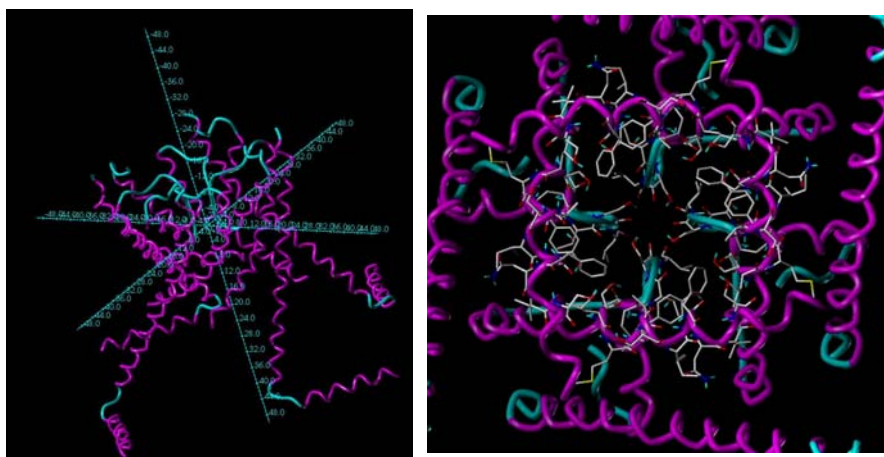


図6右：構築した hERG モデル。中心を原点とし xyz 座標軸上に表示した。

左：同モデルを z 軸負側から見た図。

B. 化合物の準備

文献から収集した100化合物を市販プログラムLigPrepにて三次元化し、適宜電荷を振った。阻害活性値 pIC_{50} は1.59~9.00、分子量は180~750の範囲である。構造式と阻害活性値は章末の附表1にまとめた。さらにOpenEye-OMEGA(2.3.2)TMにてコンフォーマーDB作成した(MMFF力場)。各化合物のコンフォーマーは最大で400個生成するようにした。

C. ドッキングシミュレーションによる化合物-hERG・化合物複合体の作成

市販プログラムFREDにより、上記hERGモデルを用いてドッキングシミュレーションを行った。ドッキングの際のグリッドは、4つの各chainのS621~S660のループの近傍に設定した。また複合体は、市販プログラムSZYBKIを用いて、構造最適化を行った(MMFF94力場)。構造

最適化は、複合体中の化合物及び化合物から半径 6 Å 以内に含まれるアミノ酸残基について行った。

D. 化合物-タンパク質間の相互作用エネルギーの計算

前章と同様、Macromodel-ASET を用いて相互作用エネルギーを計算した。計算対象としたアミノ酸残基は以下の 14 残基とした。

S621,L622,T623,S624,V625,S649,L650、
Y652,A653,S654,I655,F656,G657,S660。これらのアミノ酸残基の位置関係について図7に示す。

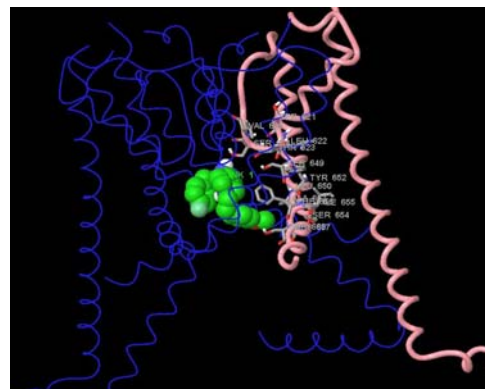


図7 計算対象とした残基の位置関係。4つの chain のうち一つを赤色で表した。中央の緑色は Astemizole。

E. Analytical Alignment について

化合物と残基あたりの相互作用エネルギー値の偏差平方和が最大となる chain を『chain A』とし、hERG モデルの Z 軸負側から見て右回りに chain B、C、D とした(図8)。

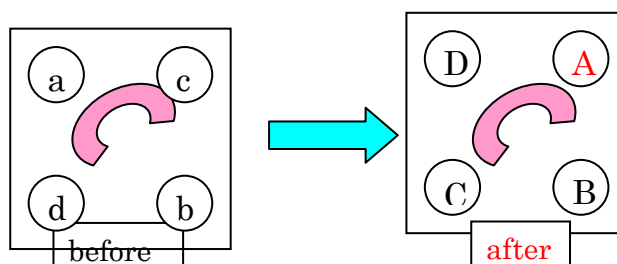


図8 丸印アルファベットは各 chain、中央の曲線は化合物を示す。計各々の化合物-hERG 複合体において、便宜上付した chain 名 a, b, c, d について各 chain の相互作用エネルギーを計算し、その偏差平方和を求めてその値が最大の chain を新たに chain A とする。次に、モデルを Z 軸負側から見て右回りに chain B、C、D とする。

G. PLS 回帰について

各残基毎の相互作用エネルギーを記述子として pIC_{50} 値に対して PLS 解析を行った。記述子は、対象残基 14 残基 \times chain 数 4 \times 2 (vdw 項、electrostatic 項) = 112 個になる。また、100 化合物中、任意に選んだ 80 化合物を訓練セット、残り 20 化合物をテストセットとして各々の 5 組を計算に付した。

2. 結果

A. PLS 解析結果について

各残基毎の相互作用エネルギーを記述子として PLS 解析を行った。5 組の訓練セットから得られた回帰モデルを用いてテストセットに適用した(表6、図10)。

表6 PLS 解析結果

Set	トレーニングセット			テストセット	
	q ²	r ²	SE	r ²	SE
1	0.641	0.862	0.617	0.829	0.923
2	0.601	0.839	0.691	0.814	0.696
3	0.637	0.839	0.691	0.864	0.576
4	0.577	0.809	0.739	0.809	0.877
5	0.620	0.828	0.694	0.811	0.846
Average	0.617	0.835	0.686	0.824	0.769

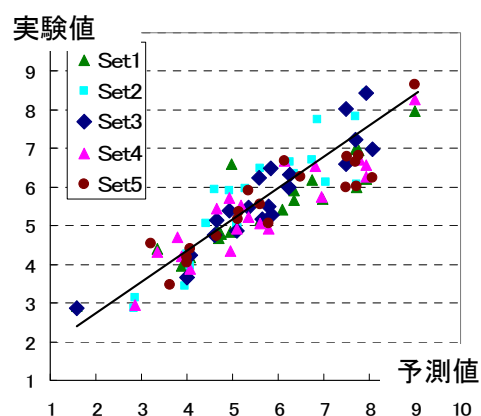


図10 テストセットの実験値-予測値のプロット

いずれの訓練セットにおいても、 $q^2 > 0.57$ 、 $r^2 > 0.8$ であり、高い統計的妥当性と高い相関を有する回帰モデルが得られた。また、5組全てのテストセットにおいても $r^2 > 0.8$ であり、高い汎用性と予測能を有するモデルであることが分かった。

C. 社内化合物への応用

知財上構造は明らかにできないが、前項で得られた回帰モデルを hERG 阻害活性値の分かっている社内化合物をテストセットとして5組の回帰モデルから阻害活性値の予測を試みた。5組のモデルから得られた予測値の平均値と実験値のプロットを図11に示す。社内化合物においても、上記モデルは $r^2 = 0.73$ と良好な予測能を示した。

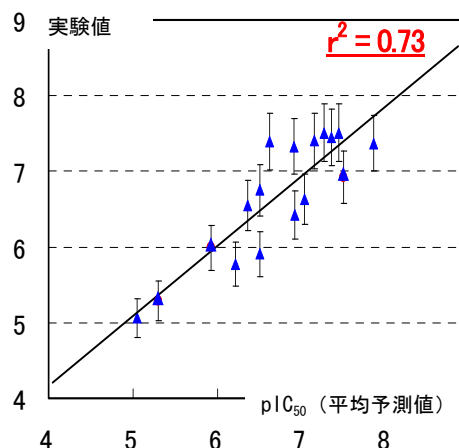


図11 社内化合物をテストセットとしたときの実験値-予測値のプロット

3. 考察

A. COMBINE 法の効果

ドッキングスコア(市販プログラム)と pIC_{50} 値を以下にプロットする(図12)。図に示すように、 $r^2 = 0.282$ であり、ドッキングスコアと阻害活性値の相関は低く、活性値の予測は困難である。即ち、COMBINE 法はX線結晶ではない hERG モデルの場合でも、生物活性の予測を可能にし、精度向上に大きな効果を有すると考えられる。

COMBINE 法は複合体モデル自身の座標を調整するのではなく、ある一定の精度を有するモデルのエネルギーを計算し、統計処理することによりモデルの再評価を行っている。即ち、モデル自身の

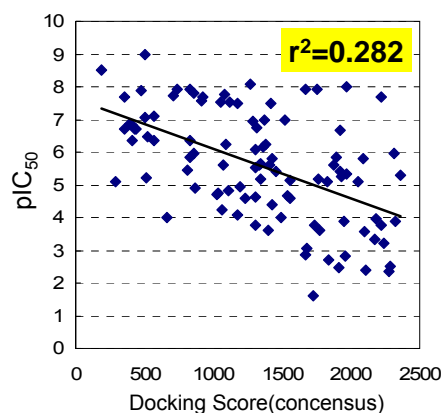


図12 スコアと阻害活性値のプロット

精度を向上させるために計算資源をかけるのではなく、複数のモデル間の有意な関係を見出すことによる構造活性相関解析である。市販プログラムの中には、ドッキングシミュレーションの工程とその評価の工程に多大な計算資源を投入するものがあるが、COMBINE 法をドッキングに組み合わせる本法の方がコストパフォーマンスが高いと考えられる。

B. 市販プログラムとのパフォーマンスの比較

COMBINE 法と市販プログラムとの活性予測法としてのパフォーマンス比較のために、前章の5組のテストセット(各セット20化合物)について両者の方法における、Sensitivity(感度;真の陽性率)と Specificity(特異性;真の陰性率)を算出し、それらからパフォーマンスの指標である陽性尤度比(LR)を求めた。Sensitivity と Specificity は以下のように定義される。

$$\text{Sensitivity}(\%) = A / (A + C)$$

$$\text{Specificity}(\%) = D / (B + D)$$

ただし、実験値での陽性の化合物数=A

実験値での陰性の化合物数=B

予測値での陽性の化合物数=C

予測値での陰性の化合物数=D

今回は hERG 阻害活性値 $pIC_{50} > 5.0$ を陽性とした。また、市販プログラムから計算されたドッキングスコアについては、スコア > 1250 を予測値での陽性とした。5セットの結果の平均値を下表に示す(表7)。

表7 Sensitivity と Specificity の比較(5セットの平均)

	Sensitivity(%)	Specificity(%)	陽性尤度比
COMBINE 法	94%	95%	9.30
市販プログラム	51%	62%	1.18

以上のように、COMBINE 法は市販プログラムに比較して、Sensitivity、Specificity ともに大幅に上回っていることが分かった。さらに陽性尤度比を比較すると、COMBINE 法は市販プログラムの 7.9 倍高いパフォーマンスを示すことがわかった。即ち、COMBINE 法は検査法として市販プログラムよりも優れた方法であると言える。

C. 他の hERG 阻害活性予測法との比較(表8)

hERG の阻害活性予測の方法論はいくつか報告があるが、現在その多くは IC_{50} 値が 1 μ M 以上か以下か、といったバイナリー型のものである。バイナリー型の予測法は一度に多くの化合物の処理が可能であるが、汎用性に課題があると思われる。また、創薬の現場において、主

活性については有望であるが、hERG 阻害活性が高いため構造変換をして類縁化合物を作る必要がある、といった局面がある。そのような場合は、バイナリー型の予測値よりは本法のように具体的な値を予測できる方が有用である。このように複数の観点から見て、COMBINE 法は優れた方法論であると言える。

表8 hERG 阻害活性予測法の比較

方法	計算時間	汎用性	予測精度	学習能力	分子設計指針
精密ドッキング	×	△	○	×	×
化合物の 3D-QSAR	△	×	◎	△	◎
COMBINE 法	○	○	◎	○	△
ニューラルネット	△	○	○*	○	×
決定木	○	△	△*	○	×

*バイナリー型の予測法

D. PLS 解析から得られる情報

PLS 解析では、Loading Plot により回帰モデルにおける critical な項を抽出することが可能である。即ち、本法においては、hERG 内の阻害活性に対する影響の大きいアミノ酸残基を見出すことが可能である。図13に Loading Plot を示す。Loading Plot においては原点からの距離が大きいものが影響の大きい項である。上図のように、vdw 項、electrostatic 項それぞれについて、活性に影響の大きいアミノ酸残基が見出されることが分かる。この値を hERG モデルにマッピングしたものを図14に示す。図に示したアミノ酸残基は、文献で活性発現に重要であることが報告されているものと一致していることが分かった。即ち、本法における PLS 解析から見出された情報は妥当であると言える。

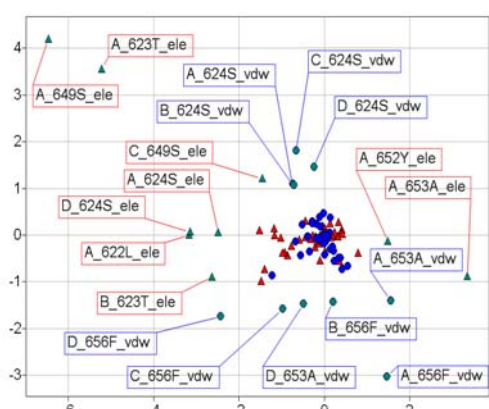


図 13 PLS Loading plot

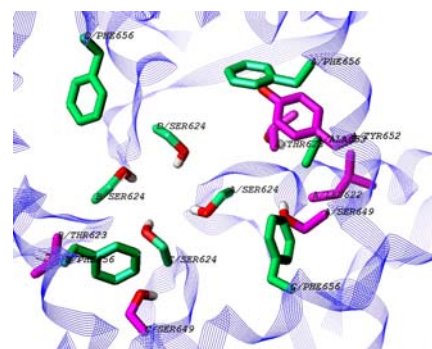


図 14 Loading Plot により抽出された残基
 緑色：vdw 項の影響の大きい残基。
 赤色：electrostatic 項の影響が大きい残基

E. Analytical Alignment の効果について

統計処理を行う以上、記述子となるアミノ酸残基は全ての複合体モデルにおいて統一されていなくてはならない。しかし化合物の構造がバラエティに富んでおり、さらにタンパク質側鎖を含めて構図最適化しているため、側鎖の座標が複合体ごとに異なる。また、用いたモデル

は4回対称の四量体であるため、化合物と複雑な関係が生じている。そのため、計算対象の残基を化合物からの一定距離内にあるもの、といった方法では、対象残基の組合せが統一されない。そこで、何らかの方法で全ての複合体モデルでの対象アミノ酸残基を統一しなくてはならない。はじめに複合体モデルを geometric にアライメントを試みたが、残基を統一することはできなかった。次に、あらかじめ各 chain における同じ組合せのアミノ残基を計算対象とし、相互作用エネルギー値の偏差平方和を回帰分析上の情報量として考え、基準となる chain を決めることで記述子を一致させることを試みた (Analytical Alignment)。その効果を表9に示す。

表9 Analytical Alignment の効果

	Q ²	R ²	SE
アライメント後	0.637	0.839	0.691
アライメント前	0.198	0.615	0.933
chain を区別しないとき	0.367	0.537	1.024

Analytical Alignment を行わない場合は統計的に妥当な回帰モデルは確立できないことが分かった。また4本の chain を区別せず、同じ配列番号の4残基を一つの残基述子とした場合も有意なモデルは見出されなかった。即ち、Analytical Alignment は極めて重要かつ有効な考え方であると言える。またこの方法は四量体に限らず、四次構造をとる全てのタンパク質に適用可能であると考えられる。

4. 結論

ホモロジーモデリングにより構築した hERG 立体構造モデルのドッキングシミュレーションと COMBINE 法の適用により、実用的な hERG 阻害活性予測モデルを確立した。また、エネルギー値に基づく analytical な alignment は、四次構造を有するタンパク質の化合物複合体に対して応用可能である。活性値の範囲が広い場合は、ドッキングスコアのみでは相関関係を見出すには困難であり、その際 COMBINE 法はドッキングポーズの再評価法としての可能性を有している。さらに、立体構造が明らかになっていないタンパク質のモデルを用いたドッキングシミュレーションの有用性が広がることが期待される。

以上の検討から、COMBINE 法は myPRESTO に組み込むことで機能向上を図ることが期待される有用な方法論であると言える。

1. Finlayson, K., Witchel, H. J., McCulloch, J., and Sharkey, J. (2004) *Eur J Pharmacol* 500, 129-42.
2. Rezazadeh, S., Hesketh, J. C., and Fedida, D. (2004) *J Biomol Screen* 9, 588-97.
3. Shimizu, H., Toyoshima, C., and Oiki, S. (2003) *Jpn J Physiol* 53, 25-34.
4. Ortiz, A. R., Pisabarro, M. T., Gago, F., and Wade, R. C. (1995) *J Med Chem* 38, 2681-91.
5. Martin-Santamaria, S., Munoz-Muriedas, J., Luque, F. J., and Gago, F. (2004) *J Med Chem* 47,

- 4471–82.
- 6 Ortiz, A.R., et al., *J Med Chem*, 1995. 38,14,2681–91.
- 7 Peters, M.B. and J. Kenneth M. Merz, *J. Chem. Theory Comput.*, 2006,2,383–99.
- 8 Cavalli, A., et al., *J Med Chem*, 2002. 45,18,3844–53.
- 9 Sherman, W., et al., *J Med Chem*, 2006. 49,2,534–53.
- 10 Atsushi Inanobe, Narutoshi Kamiya, Shingo Murakami, Yoshifumi Fukunishi, aruki Nakamura, Yoshihisa Kurachi, “In silico prediction of chemical block of cardiac repolarization reserve K⁺ current”, *Journal of Physiological Science*, 2008, 58, 459–470.
- 11 Yukio Hosaka, Miki Iwata, Narutoshi Kamiya, Mitsuhiko Yamada, Kengo Kinoshita, Yoshifumi Fukunishi, Kenji Tsujimae, Hiroshi Hibino, oshihusa Aizawa, tsushi Inanobe, Haruki Nakamura, Yoshihisa Kurachi “Mutational Analysis of Block and Facilitation of HERG Current by A class III Anti-Arrhythmic Agent, Nifekalant”, *CHANNELS*, 1, 198–208, (2007).