

タンパク質複合体の立体構造と機能の予測

●木寺 詔紀¹⁾ ◆池口 満徳¹⁾ ◆北尾 彰朗²⁾

1) 横浜市立大学大学院国際総合科学研究科 2) 東京大学分子細胞生物学研究所

＜研究の目的と進め方＞

タンパク質の立体構造情報からその分子機能を理解・予測することを最終的な目的として、立体構造と分子機能の相関を明らかにする研究を行った。本課題では、タンパク質の分子認識と基質結合に伴う構造変化等について、(1) データベース解析、(2) 分子シミュレーション解析、(3) 実験情報解析の3つの方法を用いた研究を行った。

データベース解析は、木寺が担当し、基本的には立体構造データベースProtein Data Bank (PDB) にある座標情報を帰納的な立場から網羅的に観察することで機能情報を読み取る方法論の開発を行った。そこで用いる観点は比較・分類によって、同じグループに分類されたものに共通する情報、異なったグループからは差異情報を取り出すという、古典的な分類学の手法を用いた。

分子シミュレーション解析は、木寺、池口が担当し、原子レベルの詳細なモデルを計算機中に構築し、長時間の分子動力学シミュレーションを行うことによって、そこに観察されるダイナミクスから、詳細な機能情報を取り出すという研究手法である。ここでは、その方法としての高速に大規模な系のシミュレーションを行うための、超並列計算機用ソフトウェアの開発から始めた。次いで、そこで完成したソフトウェアを用いて巨大系のシミュレーションを行った。その結果を理論のレベルまで抽象化し、分子機能が発現する際の一般論を構築した。

実験情報の解析は、北尾が担当し、タンパク質の立体構造とそのダイナミクスに関わる実験情報（ここではNMRと中性子散乱に注目する）から、機能に関わる情報を取り出す方法論の開発とその応用を試みた。NMRでは、距離情報とオーダーパラメータによるダイナミクスの情報を同時に扱う方法の応用を、中性子散乱では、非干渉性散乱からダイナミクスの情報を抽出する問題を扱った。

の方法によって研究を分担して行う。

立体構造のデータベース解析（木寺担当）：蛋白質の立体構造情報からその分子機能を理解・予測することを最終的な目的として、帰納的に立体構造と分子機能の相関を明らかにする研究を行う。第一に、基質認識部位の原子レベルの物理化学的特性の共通性を帰納的に抽出する試みを行う。基質としては、立体構造データの豊富な低分子化合物であるモノヌクレオチド、糖鎖を取り上げる。第二に、立体構造分類の基盤的方法論の確立である。立体構造と分子機能との関係を帰納的に探るためには、立体構造類似性の記述・分類が必須である。そのために新たな方法論を構築することを目指す。

分子機能のシミュレーション解析（池口担当）：上記のデータベース解析に対して、理論・シミュレーションの技術を用いて、より演繹的に分子機能発現過程の記述を試みる。その目的のために、高速に巨大なタンパク質系の分子動力学計算を行うことが必要となる。そこで超並列計算機用分子動力学計算プログラムを開発することから始める。次に、そのソフトウェアを用いて、実際に機能発現の現場をシミュレートする。さらに、それらで得られたシミュレーション結果をより一般的な形に集約し、分子機能を表現する理論を構築する。

立体構造ダイナミクスの実験的解析（北尾担当）：上の担当では、シミュレーションによってタンパク質複合体の動的構造を見ているが、ここではシミュレーションの情報に併せて、実験的なダイナミクスの情報、核磁気共鳴(NMR)、中性子散乱のデータからダイナミクスの情報を抽出する。NMR情報においては、これまでに開発してきた動的立体構造精密化法を用いて新たな系の動的構造を明らかにする。中性子散乱については、その非干渉性散乱が含むダイナミクスの意味を明らかにすることから始める。

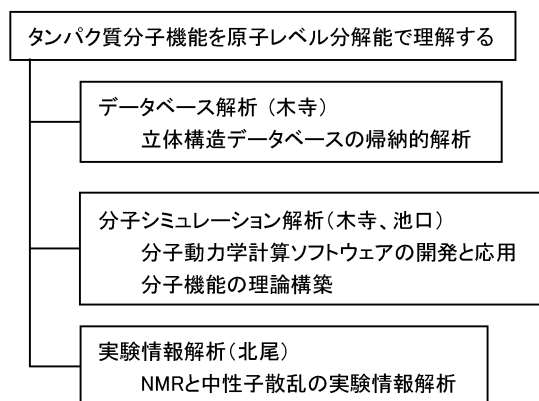


図1 研究体制

＜研究開始時の研究計画＞

本研究では、タンパク質の立体構造という実像に基づいて、分子機能発現のメカニズムを原子・分子のレベルで明らかにする理論的研究を行う。ここでは以下の三つ

＜研究期間の成果＞

立体構造のデータベース解析

(1) All β タンパク質の立体構造分類 (Ref.11)

タンパク質立体構造の分類は、代表的なSCOP、CATHに見られるように、基本的には2次構造 (α -helix、 β -sheet) の空間配置の類似性によってなされ、相同性の検出、即ち機能の推定に大きな役割を果たしている。しかしながら、ここで注目するAll β タンパク質は、95個のフォールドに分割され、そこからprotein worldの全体像を知ることはいかなる。ここで、2次構造の空間配置による分類の上位をなす分類を試みた。この分類結果にも続いてprotein worldにある立体構造の多様性の意味について考察した。

分類の方法論は、まず、"Ring"構造の存在の有無で2分する。"Ring"構造とは、N末端とC末端が十分に近い位置にあるかどうかである。さらに"Ring"構造がある場合には、"Zipper"型コンタクトの有無で2分する。"Zipper"型コンタクトとは、図2に示したように、いわゆるコンタクトマップを逆対角線に入るコンタクトの形である。

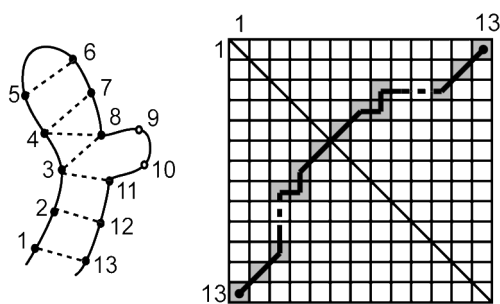


図2 Zipper型コンタクトのイメージとそのコンタクトマップ

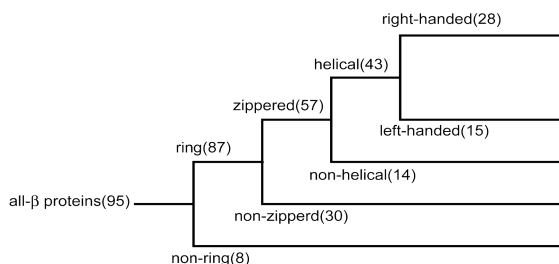


図3 All β タンパク質の分類結果

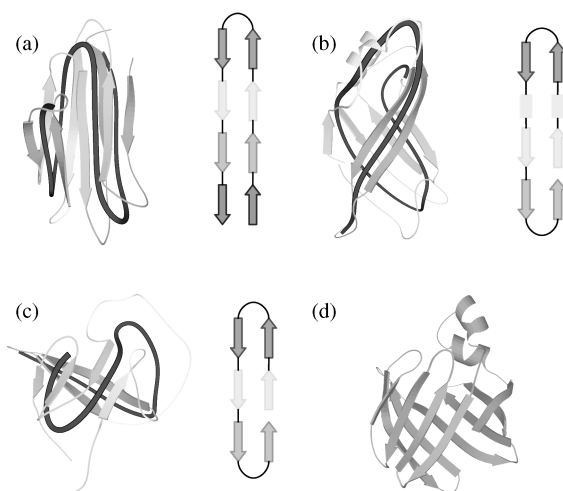


図4 分類結果の例 (a) ring/zippered/helical/right-handed、(b) ring/zippered/helical/left-handed、(c) ring/zippered/non-helical、(d) ring/non-zippered

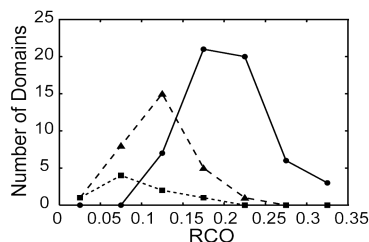


図5 All β タンパク質の Relative Contact Order (RCO)

これが、図2にあるような"Zipper"を形づくることから、そのように呼ぶ。また、いわゆるGreek-key構造を同様に"Zipper"と分類するために、コンタクトマップに周期境界条件をつけて計算を行った。

この分類結果を図3にまとめる。これによれば、All β タンパク質の95個のフォールドは、"Ring"構造を持つ多数のフォールドからなり、"Zipper"はその中でも多数を

形成していることがわかる。さらにその後に、"Zipper"をひとつの鎖と見ることによって、その鎖の形状について、helicalか否か、helicalであれば巻き方向は右か左かで分類される。それらの例を図4に上げる。

ここで、"Zipper"型が多数であることの意味を考える。図5のRCOというindexは、相互作用が鎖に沿って近いものが多いときに小さな値を、遠いときに大きな値をとる、フォールドの複雑さの尺度であり、生状態へのフォールドのしやすさを示しているといわれている。この図に明らかなように、多数を占める"Zipper"型構造は大きなRCOを持ち、複雑な構造であることを示している。するとそれらは、生状態へフォールドしにくいこととなる。しかし、細胞中でそのようなフォールドのしにくさは、そのタンパク質にとって不利な性質である。しかし、仮に、DNAのスーパーヘリックスのように、"Ring"構造を作った後に鎖のねじれのみによって高次構造を形成するような仕組みが"Zipper"にあれば、その問題が解決されていることになる。逆に、最もフォールドのしやすい"non-Ring"構造は、その数が少ない。換言すれば、そのような構造では構造多様性を生みにくいことを意味している。その意味で、"Zipper"は多様性を生みつつ、フォールドしやすさを維持するprotein worldの戦略であると見ることができる。

(2) 確率的アラインメント法 (Ref.14)

現時点での蛋白質の機能同定におけるインフォーマティクスの分野における最も強力な方法は、PSI-BLASTに代表される配列のプロファイル比較法である。これは、multiple alignmentにより定義される各ローカスに対して、アミノ酸の出現頻度をプロファイルとし、そのプロファイルを用いて、データベースを検索する。さらに、そこで得られた相同配列を加えてさらにプロファイルを更新し、検索を行い、プロファイルが収束するまで行う。これによって、微弱な配列相同性をも検出することができる。

しかし、ここで基本となっている方法はあくまでもdynamic programmingという最適化法であり、それによる限界を超えることはできない。そこで、ひとつの試みとして、我々は確率的アラインメント法 (probabilistic alignment method) という相同性検索の方法を開発した。この方法は、比較において、ふたつの配列 (もしくは立体構造) のみを用い、相同配列を用いることがないにもかかわらず、dynamic programmingの標準法であるBLASTをはるかに超える検出感度を与えるものとなった。これは、dynamic programmingを有限温度に拡張し、すべてのalignment α を確率の形で表現することによって実現される

$$Z = \sum_{\alpha} \exp[-E(\alpha)/T]$$

として最も認識に適した温度Tのもとで、分配関数Zを表現し、

$$P(p,q) = \frac{1}{Z} \sum_{\alpha}^{(p,q)} \exp[-E(\alpha)/T]$$

のようにして、ふたつのアミノ酸p, q間の類似度に変えて、p, qを対応させる確率を用いる。

この方法には解決しなければならない問題がある。それは確率にあるエントロピー項の問題である。図6に示すように、対角部分にあるアミノ酸対は、多数のalignmentに含まれ得るため、それだけで確率が高くなるという問題が生ずる。このエントロピーの問題は、人工的にN末端とC末端を結ぶことによる、周期境界条件を科するこ

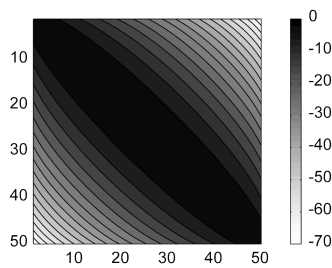


図 6 50 残基のまったく同一の配列の比較をしたときの確率マップ。周期境界条件を使用していない。

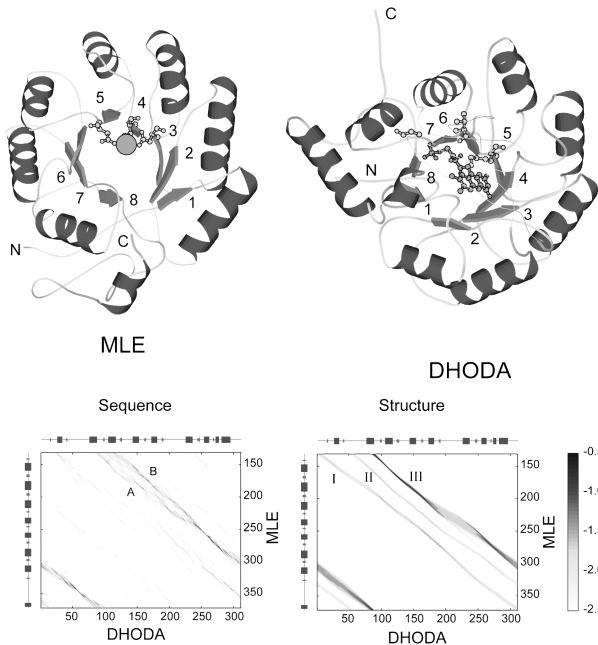


図 7 TIM-Barrel 構造 MLE と DHODA の確率アラインメント法による比較。確率マップをそれぞれ下に示す。

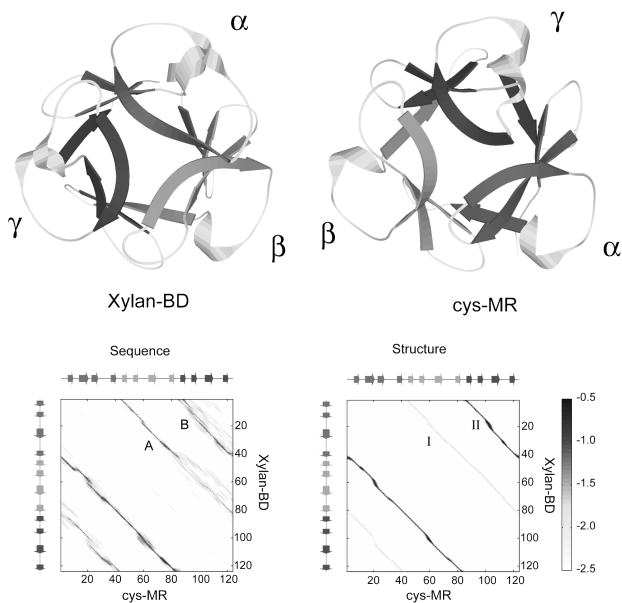


図 8 β -Trefoil 構造 Xylan-BD と Cys-MR の確率アラインメント法による比較。確率マップをそれぞれ下に示す。

とによって解決することができた。そのとき、周期境界条件における分配関数は、以下の表式として得られる。

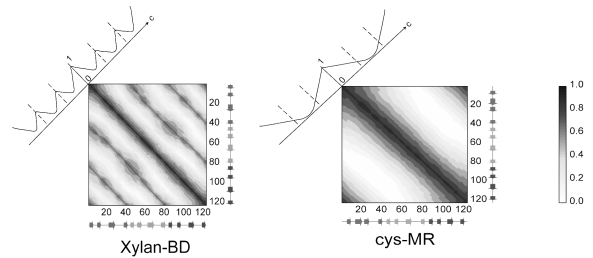


図 9 β -Trefoil の配列の自己比較

$$Z(E) = \sum_{k=1}^{L_H} \{Z[A(p, q_k)] - Z[A(p, q_k^{k-2})]\}$$

ここで、 $A(p, q)$ は、配列 p, q を結ぶalignment全体の集合である。

確率的アラインメント法を用いた円順列形の対応関係について2例示す。まず、C-terminal domain of muconate lactonizing enzyme (MLE)と dihydroorotate dehydrogenase A (DHODA)というふたつのTIM-Barrel構造の比較をした。TIM-Barrelはもともと、遺伝子重複によってできあがった擬8回対称性の立体構造を持っている。これらの間の相同性は、標準的にPSI-BLASTを用いだけでは検出できない。しかし、確率的アラインメント法によれば、図7に示したように、AとBという配列相同性、また、I、II、IIIという円順列形の構造類似性が検出された。ここで、立体構造に適用する際には、距離情報を用いなければならないが、分配関数を計算可能とするため、距離情報を自己無撞着平均場として扱った。この配列相同性と構造類似性は、活性残基の対応関係を反映する確からしいものとなった。

次に、xylan binding domain (xylan-BD)と cysteine-rich domain of mannose receptor (cys-MR)というふたつの β -Trefoil構造の比較をした。これらはもともと相同であるケースであり、擬3回対称性の立体構造に沿った、円順列形相同性が検出対象である。確率的アラインメント法によれば、図8に示したような、AとBという配列相同性があり、これらは、IとIIという立体構造類似性に対応していることが示された。この様な相同性は、 β -Trefoilが遺伝子の3重化によって生じたことを意味している。これについては、図9に示した自己相同性からも伺うことができる。xylan-BDははっきりと3回対称の配列相同性を持つが、cys-MRには見えない。このことは、3重化以降の進化の時間に対応しているものと推定される。

現在、確率的アラインメント法は、網羅的なROC解析によって、その効率が確からしいものであること、さらに、profile法への拡張において、dynamic programmingを基礎とした方法の効率の改善を図れることを確認した。

分子機能のシミュレーション解析

(1)超並列高速分子動力学計算ソフトウェアの開発: MARBLE (Ref.9)

分子動力学シミュレーションは、立体構造から、機能情報を原子レベルで研究する最も優れた方法である。しかし、生体内で重要な機能を果たす生体分子は、通常、巨大複合体を形成しており、原子レベルで分子動力学シミュレーションを実行するのは、系が巨大すぎるため計算能力上の限界があった。一方、ハイパフォーマンスコンピューティングの分野では、多数のCPUを用いて同期

をとりながら並列に計算していくという計算手法が、主流になりつつあった。その超並列計算手法を、分子動力学シミュレーションに適用しようとした場合、CPUごとの負荷が一定にならないためCPUの待ち時間が生じ、結果として計算効率が上がらないという問題があった。そこで、MARBLEでは、実行時に計算時間を計算しながら、動的に負荷分散を行うシステムを組み込み、並列化を行った。その結果、図10にあるように128CPU程度まで効率よく加速することができるようになった。これにより、図10中にある30万原子からなる超分子複合体のような大規模系の分子動力学シミュレーションが可能となった。

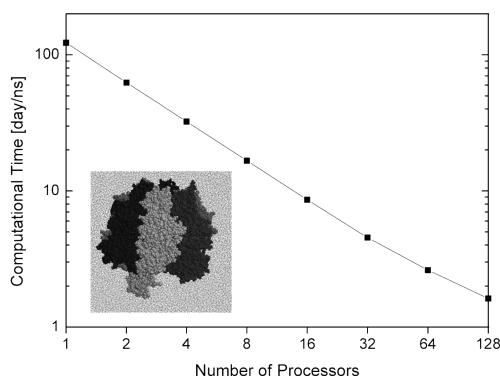


図10 分子動力学計算の並列化効率

また、部分剛体分子動力学法を開発し、MARBLEに組み込んだ。この部分剛体分子動力学法では、メチル基のような水素原子を含む原子団を剛体として扱うことで、速く変化する自由度を扱うことなく、時間発展を計算する。このことにより、より幅広く時間刻みをとることができ、計算の効率向上が図られる。剛体の時間発展を記述するには、オイラー角など配向のパラメータも用いなければならない。運動方程式は複雑になる。そのような運動方程式に対しても、安定に時間発展の計算が可能なシンプレクティックという性質を持つ時間積分法を組み込んだ。分子動力学計算の安定性の指標としては、全エネルギーの保存がよい指標となる。従来の時間積分法では、長時間の分子シミュレーションを行うと、全エネルギーがずれてしまい、安定なシミュレーションが困難な場合があった。図11は、MARBLEによる部分剛体分子動力学法の安定性を示している。ポテンシャルエネルギーのゆらぎや緩和に比べて、全エネルギーはよく保存しており、この方法の精度の高さを示している。

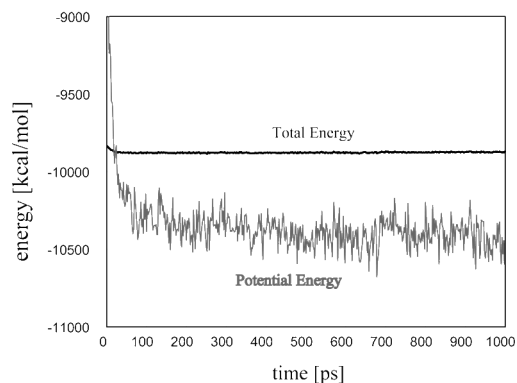


図11 部分剛体分子動力学法の安定性

生体中で、重要な役割を果たしているタンパク質は、水溶性のものばかりではなく、生体膜中にある膜タンパク質もある。これまで、膜タンパク質は立体構造決定が困難であったために、報告された立体構造は数少なく、

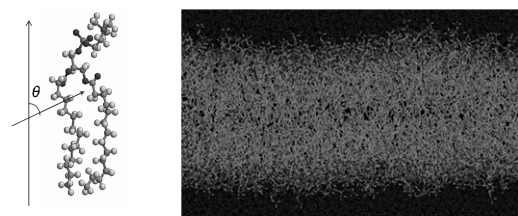


図12 生体膜を構成するリン脂質分子(左)とリン脂質が脂質二重膜を形成している様子(右)

分子動力学シミュレーションによる解析もあまり行われてこなかった。しかし、近年、実験技術の進歩により、膜タンパク質の立体構造も次第に決定されるようになっており、膜タンパク質の立体構造—機能相関解析も行われるようにきた。とくに、膜タンパク質は、チャネルやポンプなど、膜内外の輸送にかかわるものが多く、その機能発現において、タンパク質の動的性質が決定的に重要な役割を持つ。したがって、膜タンパク質の分子動力学シミュレーションは、機能解析において、重要な意味を持つのである。このような膜タンパク質の分子動力学計算をしようとする、まずは、その土台となる生体膜の分子動力学シミュレーションを十分精密に行うことのできる技術を確認しなくてはならない。図12に、生体膜の様子を示した。この図からわかるように、生体膜は、膜の接線方向と法線方向で物理的性質が大きく異なり、異方性の高い系である。この点が、通常の水溶液の系と大きく異なる。このような異方性の高い系の分子シミュレーションを行うには、通常のNPTアンサンブルの計算法に加えて、生体膜計算専用のアンサンブル計算法を導入せねばならない。MARBLEでは、NPATアンサンブル(膜法線方向圧力・膜表面積一定)、NPgTアンサンブル(膜法線方向圧力・膜表面張力一定)という2つの膜専用アンサンブルを、拡張系の方法で定式化し、部分剛体分子動力学法と結合させるような形で組み込んだ。

図13は、リン脂質DPPC(dipalmitoylphosphatidylcholine)の脂質二重膜におけるNPATアンサンブルによる分子動力学シミュレーションの結果を示している。縦軸のOrder Parameterは、図12(左)におけるCHベクトルと膜の法線のなす角 q から計算される量で、脂質二重膜中での脂肪鎖の乱雑さを表している。横軸は、リン脂質の脂肪鎖における炭素の番号である。この図からわかるように、MARBLEによる分子動力学シミュレーションの結果は、NMR実験による測定結果とよく一致しており、MARBLEによる生体膜の分子シミュレーションの妥当性が示されたといえる。

以上のように、生体系、特に、生体膜中の大規模系に対し、分子動力学シミュレーションを精度よく、かつ、効率的に実行することができるソフトウェアMARBLEを開発することができた。引き続いて、このMARBLEを、実際に、生体膜中の膜タンパク質の機能解析へ応用した。

(2)水チャネルアクアポリンファミリーの分子動力学シミュレーション (Ref.17)

アクアポリンは細菌から植物、哺乳類に至る幅広い生物種に見られる膜チャネル蛋白質である。アクアポリンファミリーには大きく分けて2種類あり、水分子しか通さないアクアポリンと、水およびグリセロールを通すアクアグリセロポリンがある。このどちらの種類も他の小分子・プロトン・イオンの透過を阻害する。ヒトでは10種類を超えるアクアポリンおよびアクアグリセロポリンが発見された。これらヒトのアクアポリンは様々な組織で発現している。たとえば、AQP1は腎臓・赤血球・呼吸

器ほかで発現している。アクアポリン遺伝子の変異は遺伝的疾患に関わっていることが知られている。例えば、AQP0の変異体は白内障にかかわっており、AQP2からも多くの尿崩症原因遺伝子が見つかった。

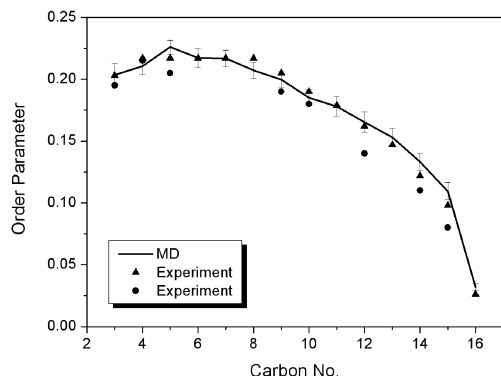


図13 リン脂質 DPPC の脂質二重膜の Order Parameter についての分子動力学シミュレーションと NMR による実験結果の比較

アクアポリンの水透過のスピードは速く、毎秒109分子程度を透過する。またアクアポリンの透過の選択性は厳密で、特にプロトン透過阻害は、膜内外のpH勾配の維持に重要な役割を果たしている。通常、プロトンは水分子の水素結合ネットワークによって運ばれるため、水分子は通すがプロトンを通さないメカニズムは長年の疑問であった。このような効率的な水透過、プロトンの透過阻害は、立体構造を元にして、物理化学的に理解できる問題であり、分子動力学シミュレーションの格好のテーマである。そこで本研究では、アクアポリンの分子動力学シミュレーションを行い、その機能発現のメカニズムを探った。

図14は、脂質二重膜に埋めたアクアポリン四量体の分子動力学シミュレーション中のスナップショットである。膜の上下に多数の水分子とイオンも配置しており、計15万原子以上からなる系の分子動力学計算を行った。対象としたアクアポリンは、研究を行った当時立体構造が決定されていたものすべてを用いた。具体的には、水分子しか通さないAQP1とAQPZ、水透過性の低いAQP0、グリセロールも通すGlpFの4種類に対して、それぞれ分子動力学シミュレーションを行い、その結果を比較した。

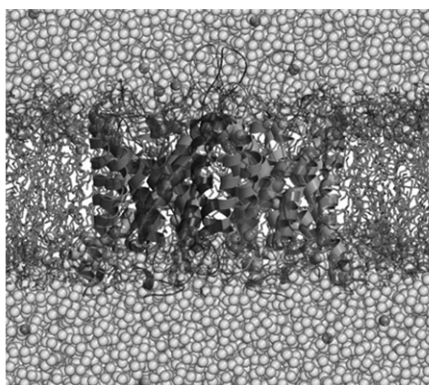


図14 脂質二重膜中のアクアポリン4量体。15万原子からなる系の分子動力学シミュレーションを行った。

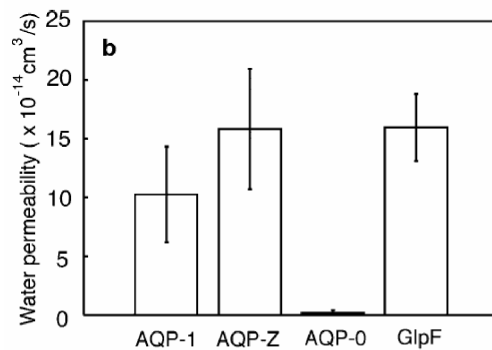


図15 各アクアポリンの水透過効率の計算結果

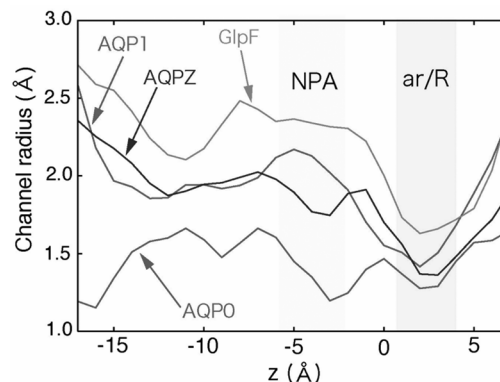


図16 チャンネルの平均半径

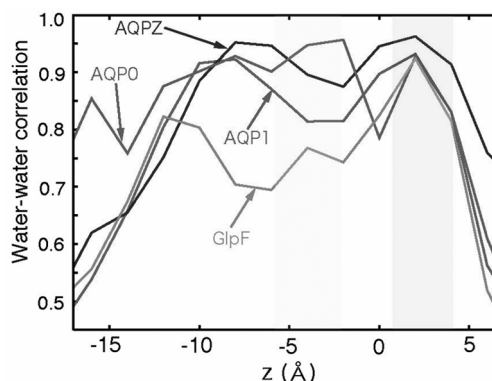


図17 水分子間の運動の相関

図15に各アクアポリンに対する水透過効率pfを比較した。これらの値は、実験値とよく一致しており、本シミュレーションの妥当性を示している。これらのアクアポリンのチャンネル部分のアミノ酸配列の一致度は45-55%程度であり、さらに各立体構造のチャンネル部分におけるC α 原子の平均2乗根距離は、0.7-1.3 Å程度と小さく、ほとんど立体構造が類似であることを考えると、この水透過効率の違いの原因は、チャンネルに面している少数の残基の側鎖の違いにあると考えられる。

図16には、分子動力学シミュレーション中のチャンネルの平均半径を示している。グリセロールも通すGlpFでは、チャンネル全域を通じて半径は大きく、そのために水の透過効率も高いものと考えられる。また、AQP0では、チャンネル全域で半径が小さく、水透過効率が非常に低いことと対応している。問題は、AQP1とAQPZの違いである。AQP1とAQPZでは、チャンネル半径がほとんど変わらないにもかかわらず、図15の水透過効率が大きく違う。この違いの原因は、水分子間の運動の相関(図17)に現れている。この図では、相関が1に近いほど、水分子は協同して動い

ていることを示している。AQPZとAQP1では、チャンネル中心のNPA部分で、チャンネル半径に違いが見られる(図16)。このチャンネル半径の違いによって、水分子間の水素結合ネットワークに差異が生じて、水分子の相関運動に違いが生じ、結果として、水透過効率の違いを生むのである。

以上のように、アクアポリンファミリーの分子動力学シミュレーションを行った。類縁タンパク質の系統的な分子動力学シミュレーションにより、その機能の違いが浮き彫りになった。この計算では、タンパク質の機能を直接観察することが可能になったのである。

(3)分子結合に伴うタンパク質構造変化の線形応答理論 (Ref. 16)

タンパク質の立体構造変化は、さまざまな生命現象で重要な

役割を担っている。例えば、モーター蛋白質では、基質結合・反応などの外部からの刺激を立体構造変化に変換し、力学的力を発生する。また、シグナル伝達蛋白質では、リン酸化などの外部からの刺激を受けて、その立体構造を変化させ、活性を調節する。このようなタンパク質の機能で共通なのは、分子結合などの外部からの刺激(摂動)に対して、タンパク質が応答して特定の立体構造変化を引き起こすということである。このような構造変化を理論的に扱おうとしたとき、まず考えられるのは分子動力学シミュレーションにより構造変化を計算機中で観察しようということであるが、タンパク質の構造変化は長時間にわたる現象であるため、現時点では分子動力学シミュレーションで扱える時間範囲を逸脱している。その点が、非常に速い機能を持つアクアポリン(上記)と大きな違いである。そこで、本研究では、分子動力学シミュレーションと統計力学理論の組み合わせで、この問題を解決しようと考えた。

本研究では、非平衡統計力学理論である線形応答理論をタンパク質構造変化に応用した。この理論では、結合する分子からタンパク質原子に力ベクトル f がかかるとすると、そのタンパク質の立体構造変化ベクトル Δr は

$$\Delta r = \frac{1}{k_B T} \Sigma_0 f$$

と表すことができる。ここで、 Σ_0 はタンパク質の平衡ゆらぎを表す分散共分散行列で、この計算に分子動力学シミュレーションや基準振動解析を用いる。添え字の0が重要で、この平衡揺らぎは非結合状態で得られるものを指している。

ここで構築した線形応答理論の適用例として、ペリプラズム鉄結合タンパク質の例を挙げる。このタンパク質は、すでにX線結晶構造解析により、非結合状態と鉄イオン結合状態の立体構造が決定されている。このタンパク質は309残基からなり、鉄イオンの結合に際し、片側のドメインが約20度回転し、結合部位を閉じる。このタンパク質に対し上式を適用するために、まず、非結合状態の分子動力学シミュレーションを行い、平衡ゆらぎを表す分散共分散行列 S_0 を求めた。この分子動力学シミュレーションには、上記のMARBLEを用い、水分子を陽に導入することで溶媒効果を考慮に入れた。分散共分散行列 S_0 から、非結合状態の平衡ゆらぎですでに二つのドメインがそれぞれ固まりとなって相対運動していることがわかった。結合する鉄イオンとの相互作用として、結合時に配位するGlu58の酸素原子に対し、鉄イオン方向に力

ベクトル f をかけた。図18に理論による構造変化の予測値と実験値を示した。両者の構造変化ベクトルの相関係数は0.95で、よく一致している。図19は、構造変化の大きさを示した図で、理論による予測値と実験値の間の相関係数は0.98と非常に一致を示している。



図18 鉄結合タンパク質における構造変化。非結合状態の主鎖構造を表示している。黒矢印が理論により予測された構造変化を示し、灰色の矢印が実験値を示す。中央の球は結合する鉄イオンである。中央の太い矢印はタンパク質にかかる力のモデルを示している。

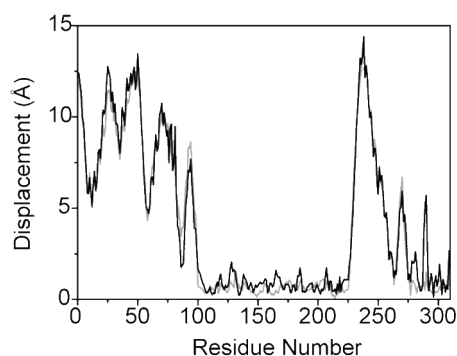


図19 鉄結合タンパク質における構造変化の大きさ。黒線が理論により予測された構造変化の大きさと、灰色の線が実験値である。

次に、力ベクトル f による結果の依存性について考察した。まず、先と同じGlu58の酸素原子に360度、様々な方向に摂動力をかけ、結果を比較した。上式から明らかに、反転した摂動力をかけると結果となる構造変化も反転する。図20には、構造変化の実験値との相関係数が0.8以上となる摂動力の方向を表した。大半の方向に摂動力をかけてもタンパク質は実際と同じような構造変化を応答することを示している。また、力ベクトル f の場所依存性についても検討した。図20の鎖の濃い色の部分は、その $C\alpha$ 原子から鉄イオン方向に力ベクトルをかけたとき構造変化の実験値との相関係数が0.8以上になる部分である。移動ドメイン側であれば大半の場所に力をかけても実際と同じような構造変化を引き起こすことを示している。

このような摂動力の位置・方向に対するタンパク質の構造変化のロバストネスは、どこから来ているのだろうか。これには、タンパク質の平衡ゆらぎの異方性が深くかかわっている。平衡ゆらぎを表す分散共分散行列について、仮に $C\alpha$ 原子のみを考えたとしても、鉄結合蛋白質で927の自由度がある。この行列を対角化し主成分を取り出すと、ゆらぎの大きい方から上位10自由度だけで全平衡ゆらぎの約90%近くを占める。つまり、タンパク質は多数の原子からなる大自由度系なのだが、実際に柔らかな方向は極めて限定されているのである。そして、力

ベクトルは限定された柔らかな方向を選択し、その方向に構造変化を増幅する役割を担っているということがわかった。

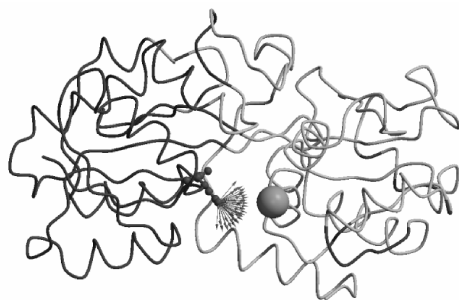


図 20 力ベクトルの方向・位置依存性の検証。Glu58 の酸素原子上に全方向の力を試し、理論値と実験値の相関係数が 0.8 以上になる方向を矢印で表示した。また、黒色の鎖の C α 原子の位置から、鉄イオン方向に力をかけても、理論値と実験値の相関係数が 0.8 以上になった。このことは力ベクトルの方向・位置に対し、タンパク質の構造変化の依存性が少ないことを示している。

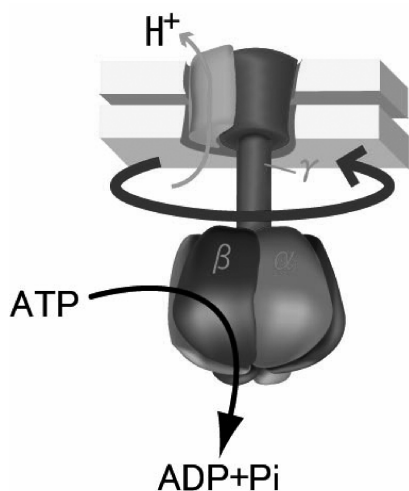


図 21 F₀F₁-ATP 合成酵素の概念図。
下の水溶性部分が分子モーター F₁-ATPase である。

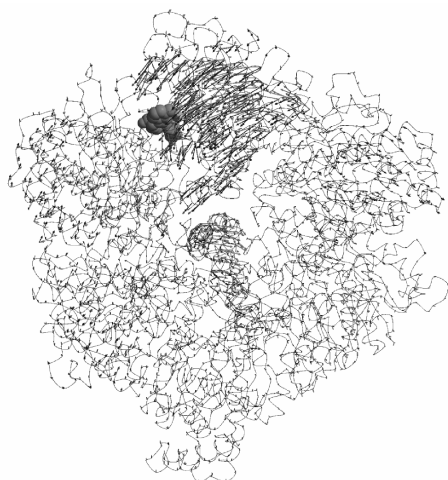


図 22 F₁-ATPase に線形応答理論を適用した図。黒の矢印が理論値で、灰色の矢印が実験値である。 β_E サブユニットの構造変化の理論値と実験値の相関係数は 0.84 である。

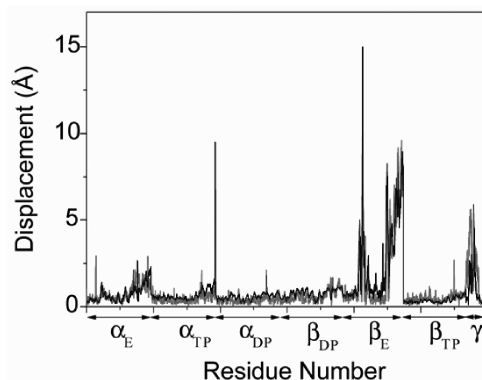


図 23 F₁-ATPase の構造変化の大きさ。黒線が理論値で、灰色の線が実験値である。理論値と実験値の相関係数は 0.85 である。

さらに、このタンパク質構造変化の線形応答理論を、より複雑な構造変化をする分子モーター F₁-ATPase に適用した。図 21 に F₀F₁-ATP 合成酵素の概要図を示す。F₀F₁-ATP 合成酵素は、膜内外のプロトン勾配の駆動力を利用して、生体内のエネルギー源である ATP を合成する酵素であり、生命現象に欠くことのできないタンパク質超分子複合体である。膜中にある部分を F₀ と呼び、水溶性部分を F₁-ATPase と呼ぶ。水溶性部分だけでも存在でき、ATP を ADP に加水分解することで、中心軸である γ サブユニットを回転させることが観察されている。 γ サブユニットの回転の初動作は、ATP の結合によって誘起されることが知られている。そこで、F₁-ATPase の ATP 結合による構造変化に線形応答理論を適用可能かどうか検証した。

F₁-ATPase に適用した力ベクトルは、リン酸に結合する Gly161 から結合するヌクレオチドのリン酸方向に向かうベクトルである。図 22, 23 は、線形応答理論の結果である。線形応答理論で予測される構造変化は、超分子複合体全域にわたって、実験値とよく対応しており、ここで構築した構造変化の線形応答理論が、分子モーターのような複雑な構造変化にも適用可能であることがわかった。このことは、線形応答理論による、構造変化の理解の枠組みの妥当性と、適用範囲の広さを示しているといえる。

立体構造ダイナミクスの実験的解析

(1) 中性子散乱スペクトルの解析方法の開発 (Ref.15)

当初の X 線結晶解析と NMR に加えて、途中年度から中性子散乱スペクトルを解析するための方法論の研究も開始し、スペクトルに見られるボゾンピークが水和に起因することを明らかにした。ボゾンピークは、約 200K 以下の低温で測定されたガラス様物質の中性子非弾性散乱スペクトルに観測される約 25 cm⁻¹ 付近のブロードなピークであるが、その起源は謎であった。200K 付近はガラス転移温度付近であり、その温度以上でガラス転移が現れ蛋白質が分子機能を発揮し、その温度以下でボゾンピークが現れるということは、機能とダイナミクスの密接な関係を示唆する。我々は、ニワトリ卵白リゾチーム蛋白質を用いて、基準振動解析、ランジュバンモード解析、分子動力学計算(水中、真空)等の分子シミュレーションを 100K-300K のさまざまな温度で行い、その結果を用いて中性子非干渉性非弾性散乱スペクトルを計算した。計算されたスペクトルを比較すると、水中の分子動力学計算(< 200K)の結果からのみ約 25 cm⁻¹ 付近にブロードなピークが観測される。我々は JAM モデルに基づき、蛋白質のボゾンピークの起源を次のように明らかにした。水和していない状態の蛋白質のエネルギー面は滑らかだが、水和

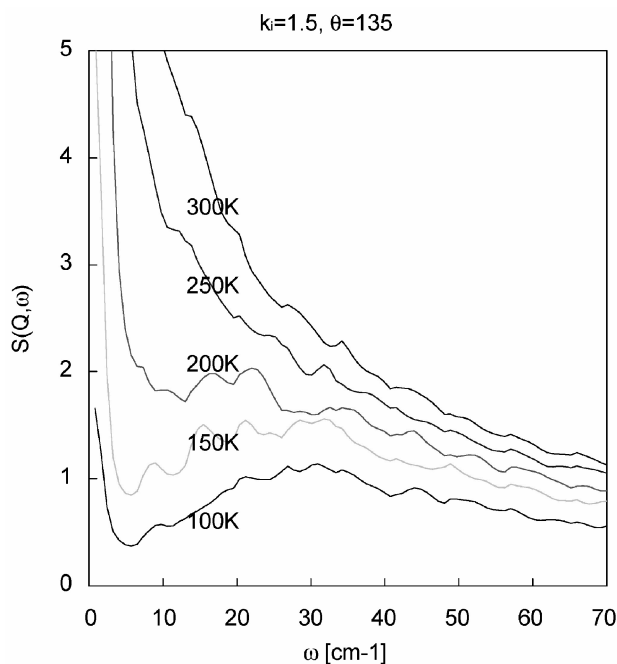


図 24 シミュレーションで得られたボゾンピーク

により蛋白質のエネルギー地形に微細構造が出現する。常温では、微細構造のエネルギー障壁を乗り越えることが比較的容易であり、極小点間をジャンプする拡散的運動が、 $5\text{cm}^{-1} \sim 30\text{cm}^{-1}$ の準弾性散乱スペクトルとして観測される。しかし、 200K 以下の低温になると、蛋白質の立体構造ダイナミクスが、微細構造の極小点にトラップされる。これにより、高温で観測された拡散的運動がより高振動領域の振動的運動に取って代わり、 $5\text{cm}^{-1} \sim 30\text{cm}^{-1}$ の中性子散乱強度が減少し、約 25cm^{-1} 付近のボゾンピークが観測されるものと考えられる。

〈国内外での成果の位置づけ〉

いずれの研究も極めて独自の観点で行われた研究である。データベース解析においては、確率的アラインメントは、今後のprofile法への応用によって初めてその真価が認められるものと期待される。シミュレーション研究においては、技術的には世界的なトップレベルに到達したと言えるだろう。AQPのシミュレーションは厳しい競争があるが、ここでの成果は一步抜け出たものとして評価されている。線形応答理論は、その枠組みが世界的に高く評価されている。中性子散乱解析は、今後、J-PARC（大強度陽子加速器）において重要な意義を持つものとなると考えられる。

〈達成できなかったこと、予想外の困難、その理由〉

データベース解析において、ヌクレオチドと糖の結合についての研究は、研究体制の不備からデータベース構築以降大きな遅れを来し、現在ようやく集約の段階に達している。また、構造変化データベースについても進捗が遅れていたが、その後新たな体勢で進捗を開始している。実験情報解析において、NMRによる構造精密化ではテストデータの作成だけに終わり具体的な方法論の改良まで達成できなかった。中性子散乱に関しては、測定装置の建設が遅れているため実験データとの比較が十分できなかった。

〈今後の課題〉

構造変化データベースの構築、タンパク質間相互作用の理論構築、巨大な超分子複合体の動作機序の解明、構

造変化シミュレーション、中性子散乱データの解釈、等が、上記の成果の延長線上にある課題である。

〈研究期間の全成果公表リスト〉

- 0202272021
Chong, S.-H., Joti, Y., Kidera, A., Go, N., Ostermann, A. Gassmann, A. and Parak, F. Dynamical transition of myoglobin in a crystal. comparative studies of X-ray crystallography and Mössbauer Spectroscopy, European Biophysics J. 30, 319-329 (2001).
- 0202272036
Yamashita, H. and Kidera, A., Environmental influence on electron scattering from a molecule, Acta Cryst. Sect A57, 518-525 (2001).
- 0602091359
Yamashita, H., Endo, S., Wako, H. and Kidera, A. Sampling efficiency of molecular dynamics and Monte Carlo method in protein simulation, Chem. Phys. Lett. 342, 382-386 (2001).
- 0202272059
Tanaka, A., Ikeda, Y., Abe-Dohmae, S., Arakawa, R., Sadanami, K., Kidera, A., et al., Human ABCA1 contains a large amino-terminal extracellular domain homologous to an epitope of sjogren's syndrome, Biochem Biophys Res Commun., 283,1019-25 (2001).
- 0303271346
Kitao A. and Takeda M., The Effects of Solvent and Anharmonicity on Incoherent Inelastic Neutron Scattering Spectra of Proteins. Journal of Neutron Research 10 (3-4), 143-7 (2002)
- 0202272108
Terada, T. and Kidera, A., A generalized form of the conserved quantity in constant-temperature molecular dynamics, J. Chem. Phys 116, 33-41 (2002).
- 0303271104
Joti, Y., Nakasako, M., Kidera, A. and Go, N., Nonlinear temperature dependence of crystal structure of Lysozyme: Correlation between coordinate shifts and thermal factors, Acta Cryst. Sect. D58, 1421-1432 (2002).
- 0311121726
Funahashi, J., Sugita, Y., Kitao, A., and Yutani, K. How can free energy component analysis explain the difference in protein stability caused by amino acid substitutions? effect of three hydrophobic mutations at the 56th residue on the stability of human lysozyme Protein Engineering 16(9), 665-571 (2003)
- 0404022048
Ikeguchi, M., Partial rigid-body dynamics in NPT, NPAT and NPgT ensembles for proteins and membranes, J. Comput. Chem. 25, 529-541 (2003).
- 0404022024
Terada, T., Matsuo, Y. and Kidera, A., A method for evaluating multicanonical potential function without iterative refinement: Application to conformational sampling of a globular protein in water, J. Chem. Phys 118, 4306-4311 (2003).
- 0404022028
Koike, R., Kinoshita, K. and Kidera, A., Ring and zipper formation is the key to understanding the structural variety in all beta proteins, FEBS Lett. 533, 9-13 (2003).
- 0404022035

- Moritsugu, K. , Miyashita, O. and Kidera, A., Temperature dependence of vibrational energy transfer in a protein molecule, *J. Phys. Chem. B* 107, 3309-3317 (2003).
13. 0404022039
Moritsugu, K. and Kidera, A., Protein motions represented in moving normal mode coordinates, *J. Phys. Chem. B* 108, 3890-3898 (2004).
14. 0602091441
Koike, R. , Kinoshita, K. and Kidera, A., Probabilistic description of protein alignments for sequences and structures, *Proteins: Structure, Function, and Bioinformatics*, 56, 157-166 (2004).
15. 0602080955
Joti, Y., Kitao, A., and Go, N., Protein Boson Peak Originated from Hydration-Related Multiple Minima Energy Landscape. *J. Am. Chem. Soc.* 127, 9705-8709 (2005).
16. 0602081632
Ikeguchi, M., Ueno, J., Sato, M. and Kidera, A., Protein structural change upon ligand binding: Linear response theory, *Phys. Rev. Lett.* 94, 078102 1-4 (2005).
17. 0602081637
Hashido, M. , Ikeguchi, M. and Kidera, A., Comparative simulations of aquaporin family: AQP1, AQPZ, AQP0 and GlpF, *FEBS Lett.* 579, 5549-5552 (2005).