

## 分子動力学とデータベースによる統合解析

倭 剛 久<sup>1</sup>, 由 良 敬<sup>2</sup>

X線結晶解析などで明らかになるタンパク質の構造は、結晶中に存在する多数の分子の平均構造でしかない。得られた構造からタンパク質の機能がどのように発揮されるのかを知るためには、その動的構造を「見る」必要がある。分子動力学シミュレーションは、タンパク質のフェムト秒からマイクロ秒の動的構造を見るための道具である。タンパク質にはもう一つの動きが存在する。その動きは何億年もかけて積み上げられた変化である。タンパク質の分子進化の跡は、データベースに蓄積された共通祖先由来のアミノ酸配列を比較することで、明らかにすることができる。スケールのまったく異なるこれら二種類の「動的構造」を見ることによって、タンパク質が機能を発揮する機構の詳細を知ることができる。具体的なタンパク質をあげながら、これらの時間軸がどのように絡まり合っているのかを記す。

## 1. はじめに

ゲノム塩基配列を読み取るプロジェクトが、20世紀最後に目覚ましい速度で進展し、一匹の細菌がもつゲノム情報すべてが瞬く間に明らかになる時代が到来した。次世代シーケンサーとその次の世代のシーケンサーの登場は、真核生物のゲノム塩基配列でさえ、あっという間に読み取ってしまう時代をもたらすと言われている<sup>1,2)</sup>。このような技術革新をもたらした根底には、生物の(特にヒトの)ゲノム塩基配列すべてを読み取ることで、医学生化学の研究課題に俯瞰的な視点をもって挑戦することが可能となり、問題解決への道を加速するに違いないという考えがあ

る<sup>3,4)</sup>。確かにヒトゲノムの読み取りによって、多くのことが明らかになった。ゲノム塩基配列と、その後のトランスクリプトーム解析などにより、今まで知られていなかった転写の総体や、RNAの重要性が明らかになってきた<sup>5)</sup>。また、近縁種のゲノム塩基配列読み取り、あるいは同一種別個体のゲノム塩基配列読み取りが可能になったことで、ゲノム配列の多様性が明らかになり、一塩基置換(SNP)と疾病との関連が精力的に調べられるようになった<sup>6)</sup>。

ゲノムのどの部分にどのような情報が書き込まれているかを明らかにできたことは、ゲノム医学にとって非常に大きな一歩であった。読み取られた情報のわずかな違いが、表現型にどのような結果をもたらすかを知ること、疾病への新たな理解をもたらされた。しかし、これまでになされてきた多くの総体的理解は、生物がどのような情報をゲノムから生みだしているかの理解、または生みだされた情報と表現型との関連の理解にとどまっているように見える。これらの理解は、論理的な因果関係の理解にとどまっており、機械論的な理解が欠落している。卓近な例えをあげるならば、自動車において、アクセルを踏むと自動車は前進するという理解が因果関係に基づく理解であり、アクセルを踏むと、エンジンに霧状のガソリンが吹き込まれ、…となって自動車が前進するという理解が、機械論的な理解である。著者らは、このような機械論的理解があっ

<sup>1</sup>名古屋大学大学院理学研究科物理学教室 (〒464-8602 愛知県名古屋市千種区不老町)

<sup>2</sup>お茶の水女子大学生命情報学教育研究センター (〒112-8610 東京都文京区大塚 2-1-1)

Combinatorial approach of molecular dynamics simulations and database analyses for the studies of protein structure and function

<sup>1</sup>Takahisa Yamato (Graduate School of Science, Nagoya University, Furo-cho, chikusa-ku, Nagoya 464-8602, Japan) and <sup>2</sup>Kei Yura (Center for Informational Biology, Ochanomizu University, 2-1-1, Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan)

てこそ、SNPが酵素の機能低下をもたらす理由の理解や、ドラッグとしてはたらくことができる低分子の設計が可能になると考えている。

X線結晶解析やNMRによって、多くのタンパク質の立体構造が明らかになっている。しかし、これらの立体構造座標データを可視化して眺めていても、機械論的理解（構造機能相関）を得ることは期待できない。機能が特定されていないタンパク質の立体構造がわかったとしても、立体構造を見ただけでその機能がわかることはとうていあり得ない。我々が生活している普通の世界においては、物体の形を見ると、一つに限定できないまでも、その用途を絞り込むことはできるが、タンパク質の場合はそうはいかない。この違いは、我々がタンパク質構造の世界における「常識」を持ち合わせていないことに起因する<sup>7)</sup>。この常識を蓄積することができれば、タンパク質の構造を通して、ゲノム情報の機械論的理解を推し進めることができるはずである。

著者らは、タンパク質構造の世界における「常識」は、タンパク質の構造に時間軸を導入することで蓄積できると考えている。ただし、導入しなければならない時間軸は、 $10^{-15}$ 秒から $10^{15}$ 秒の範囲におよぶ広大な時間軸である（図1）。フェムト秒からナノ秒の範囲は、例えば、酵素が触媒する反応の時間ドメインであり、この時間ドメインは、現在のコンピュータでシミュレーションできる範囲である（タンパク質の大きさに依存することは言うまでもないが）。コンピュータを顕微鏡のように使うことで、タンパク質がどのようにして化学反応を触媒するのかを見ることは決して不可能ではない時代が到来している。ナノ秒から秒の範囲は、タンパク質を構成する部分構造の動きやタンパク質の立体構造形成、あるいはタンパク質の合成が進行

する時間ドメインである。この時間ドメインは、コンピュータでも実験的にも見ることが難しいドメインであり、多くの研究者が精力的に技術革新に挑戦しているところである。年以上の範囲は、タンパク質分子の進化による変異が蓄積する時間ドメインである。この時間ドメインは、比較ゲノム解析で明らかにできる範囲であり、それらの変異がタンパク質構造のどのような部位で発生しているかを知ることが、タンパク質の機能を知る上で重要になる。ゲノムに発生する変異は、タンパク質の機能によって選別され、機能しないタンパク質をもたらす変異は、進化の歴史から消去されてしまう。この結果として、機能を発現するために重要な部位は、ゲノムの進化過程で保存されることになる。つまり、長い時間ドメインで保存されている部位には、そのタンパク質が機能を発現するために必要な部位が含まれる。注目してほしいことは、ここにあげたいずれの時間ドメインの情報も、コンピュータを用いて抽出できる場所である。短い時間ドメインは、物理学を基礎とするシミュレーションが活躍する場であり、長い時間ドメインは、分子進化学を基礎とする情報処理が活躍する場である。これらがタンパク質の構造情報と融合することによって、タンパク質の構造機能相関が明らかになるはずである。

以上の考えに基づき、著者らは過去5年間にわたり、いろいろな手法を開発し、具体的なタンパク質にその手法を適用してきた。以下では、その中のいくつかを示す。これらの多くは、現在でも解析を進めているタンパク質である。

## 2. DNA フォトリアーゼの電子伝達経路と機能アノテーション

ゲノム塩基配列から推定されるアミノ酸配列を比較分類

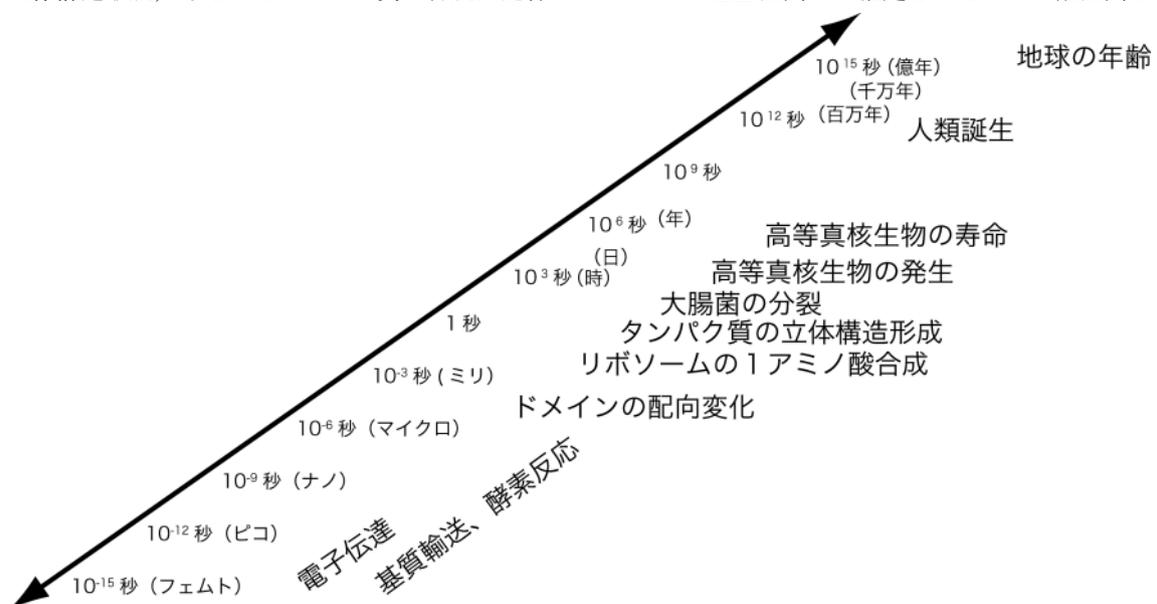


図1 タンパク質の動的構造解析をするために導入する時間スケール

すると、一つの生物種が複数の共通祖先由来遺伝子をもっていることが多々あることに気がつく。これらのパラログ遺伝子がすべて同じ機能を担っているのか、あるいは異なる機能を担っているのかは、アミノ酸配列の比較からだけでは容易に判別できない。これらの遺伝子産物が担っている機能を実験的に明らかにすることも、決して容易なことではない。見当もついていない機能を明らかにすることには、多くの困難がともなうことは言うまでもなく、さらには、ある機能を明らかにできたとしても、他の機能をもたないということにはならない。多重遺伝子族の各遺伝子がコードするタンパク質がどのような機能を担っているのかを、分子動力学とデータベースによる統合解析によって明らかにできないだろうか。

細菌ゲノムには、フォトリアーゼ青色光受容体ファミリーに属するタンパク質をコードする遺伝子が多く見つかる。シクロブタン型ピリミジン二量体(CPD)フォトリアーゼは、紫外線照射によりDNAに生じるチミン二量体(ダイマー)を切断し、DNAを修復する酵素である。この酵素がもつ還元型フラビンアデニンヌクレオチド( $\text{FADH}^-$ )が、紫外線を吸収して電子を放出することで、 $\text{FADH}^-$ からDNAへと電子移動反応が起こり、チミンダイマーの共有結合を切断する。ところが同じファミリーに属する青色光受容体は、青色光誘導遺伝子発現調整に関与することが知られているが、DNAを修復する機能はない。共通祖先由来のタンパク質において、この機能の違いはどのアミノ酸残基の違いによって発生したのだろうか。この疑問を明らかにするためには、このタンパク質の構造機能相関を理解することが不可欠である。アミノ酸残基の保存性を調べているだけでも、この疑問に対するヒントを得ることはできるが、疑問に解答を与えることはできない。保存されているアミノ酸残基すべてが機能に直接関与するわけではない。またファミリー内の配列を機能にもとづいてグループ化できないために、どのグループで保存されているアミノ酸残基が重要なのかを見いだすことができない。

そこで、著者らは分子動力学シミュレーション、電子状態計算、および分子進化の手法を組み合わせることでこの問題に取り組んだ<sup>8)</sup>。CPDフォトリアーゼとDNAとの複合体構造は、X線結晶構造解析により明らかにされている<sup>9)</sup>。この構造に基づく過去の理論的解析<sup>10,11)</sup>によると $\text{FADH}^-$ からチミンダイマーへの電子移動反応は、 $\text{FADH}^-$ のイソアロキシジン環からの電子移動と $\text{FADH}^-$ のアデニン部位を経由した電子移動の二つの経路が主要であるとされてきた。つまり電子移動経路にタンパク質は関与しないことになる。しかし、これらの研究では、電子のドナーとアクセプターを取り囲むタンパク質環境の揺らぎをあらわに考慮していない。そこで、著者らは複合体立体構造の分子動力学シミュレーションを行い、熱で揺らいでいる構造の瞬間を多

数推定し、これらの構造群における電子状態を計算し、電子移動反応経路を解析した。その結果、従来の電子移動経路に加え、メチオニン353を経由する経路の寄与も大きいことを見いだした(図2)。

メチオニン353がフォトリアーゼ活性に必要なことは、ピコ秒からナノ秒の時間ドメインを、分子動力学シミュレーションを用いて解析することで、はじめて明らかにできた。それでは、メチオニンはフォトリアーゼ青色光受容体ファミリーにおけるフォトリアーゼ活性をもつグループの中で保存されているだろうか。分子動力学シミュレーションに用いたアナシステイス・ニデュランス(*Anacystis nidulans*)由来のCPDフォトリアーゼと類似のアミノ酸配列をデータベースで検索すると、371本のタンパク質を見いだすことができた。これらのうち実験的にCPDフォトリアーゼとわかっている配列には、先のメチオニンが必ず存在した。また実験的にフォトリアーゼ活性が見いだされていない配列では、メチオニンが存在しないこともわかった。371本の配列で系統解析を行うと、201本は、フォトリアーゼ活性をもつことがわかっている配列を含むグループに分類でき、170本は、フォトリアーゼ活性をもたないグループに分類できた。この二つのグループわけと、メチオニンの有無とは完全に相関していることがわかった。すなわち、*A. nidulans*の353残基目に相当する部位におけるメチオニンの有無が、CPDフォトリアーゼ・サブファミリーとそれ以外のサブファミリーを識別する目印になることが明らかになった。億年の時間ドメインにおいて、このメチオニンを保存することが、フォトリアーゼ活性を維持するために必要であることを強く示唆する解析結果を得ることができた。

この結果は、ゲノム塩基配列から推定されるタンパク質の機能アノテーションにも利用されている。著者らの解析が発表される前には、ここで問題となっているメチオニンの部位がバリンになっているシロガラシ(*Sinapis alba*)由来の配列がフォトリアーゼ活性をもつとアノテーションされていた。しかし、著者らの論文が発表された後にアノテーションが修正され、現在ではこの遺伝子は青色光受容体をコードするとアノテーションされている。

### 3. SNPをタンパク質の構造にみる

次世代シーケンサーの登場により、SNPの解析が加速し、SNPと疾患との相関解析が進められている。個人ゲノムに見つかるSNPに基づき、病気の発症リスクを推定する研究も行われている<sup>12)</sup>。これらの研究が、人類の健康増進に貢献することは間違いない。しかし、現在の多くの研究は因果関係に基づく理解が中心であり、機械論的理解をめざした研究は多くはない。あるSNPがある疾患に関係している理由を、構造機能相関の枠組みの中で理解す

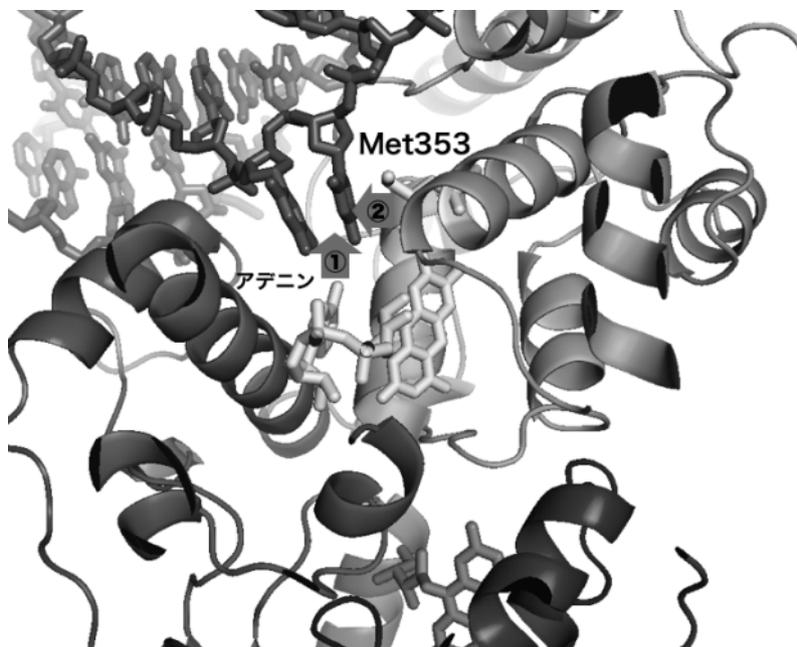


図2 CPD フォトリアーゼがチミンダイマーを含む DNA (上部) と相互作用している構造

タンパク質の中央にある FAD(白)がわかりやすいように、タンパク質の手前部分をすべて非表示とした。FAD のアデニン部分から直接チミンダイマーに電子が移動する経路①は以前から指摘されていた。今回新規に、Met353 を経由する経路②を見いだした。

ることができれば、疾患に直接作用する薬の合理的な創出も不可能ではないであろう。著者らに限らず誰でもがこのように考えているが、この問題の解析は容易ではないために、発表されている研究は少ない。SNP と疾患の関係が実験動物ではなく、ヒトで見られる現象であることから、検証実験が難しいことは言うまでもないが、SNP が生物進化の過程で、ある程度保持されていることから、SNP が種の保存のためには有害とは限らないことを意味しており、このことも構造機能相関の研究を難しくしていることの根幹にある。酵素などの触媒部位に変異が入っている場合には、その遺伝子をもつ個体の生存に問題が生じ、そのような変異は淘汰されてしまう。SNP と疾患とは、機械論的にはどのようなつながりがあるのかを見いだすのは、容易ではなさそうである。

著者らは、さまざまな SNP がタンパク質立体構造のどのような部位で発生しているのかを調べた<sup>13)</sup>。ここでの解析では、アミノ酸一残基に変異をもたらす SNP に限り、アミノ酸残基長に変化をもたらす SNP やアミノ酸残基に変化をもたらさない SNP は除外した。タンパク質立体構造解析技術の飛躍的な進歩のおかげで、現在ではゲノム塩基配列から推定されるタンパク質の半数以上が、その立体構造を推定することができるようになってきている<sup>14)</sup>。現在までに判明しているすべての SNP が登録されている dbSNP<sup>15)</sup> のデータ、つまり疾患との関係が明らかにされて

いない SNP も含めて、タンパク質の立体構造との関係を調べると、SNP によるアミノ酸残基の変化は、タンパク質の至る所で発生していることがわかる (図 3)。これらのうち、いずれの SNP が疾病と関係しているかが明らかになれば、タンパク質の構造を通して、SNP がタンパク質の機能におよぼす影響を機械論的に議論できるようになるはずである。

さまざまなグループが行っている研究により、疾患に関与する SNP は、タンパク質の表面のくぼみ部分に多く発生している<sup>16)</sup>、タンパク質の内部に多く発生している<sup>17)</sup>、進化の過程で保存されているアミノ酸に発生するケースが多い<sup>18)</sup>などの報告がなされている。著者らもがんに関係する SNP のみを集め、それらがタンパク質のどのような位置に存在するアミノ酸残基に変異をもたらしているかを調べた。その結果として、MSH6 遺伝子に存在する発がんに関連する SNP は、タンパク質の内部や基質結合部位に存在するケースが多いことを見いだした<sup>19)</sup>。これらの変異がタンパク質の安定性 (寿命) や基質との相互作用にどのような影響を与えるかは、分子動力学シミュレーションを用いて明らかにしていかなければならない。図 1 に示したナノ秒から秒の時間ドメインと、年から億年の時間ドメインの相互作用を見いだすことで、SNP の機械論的理解ができると考えている。

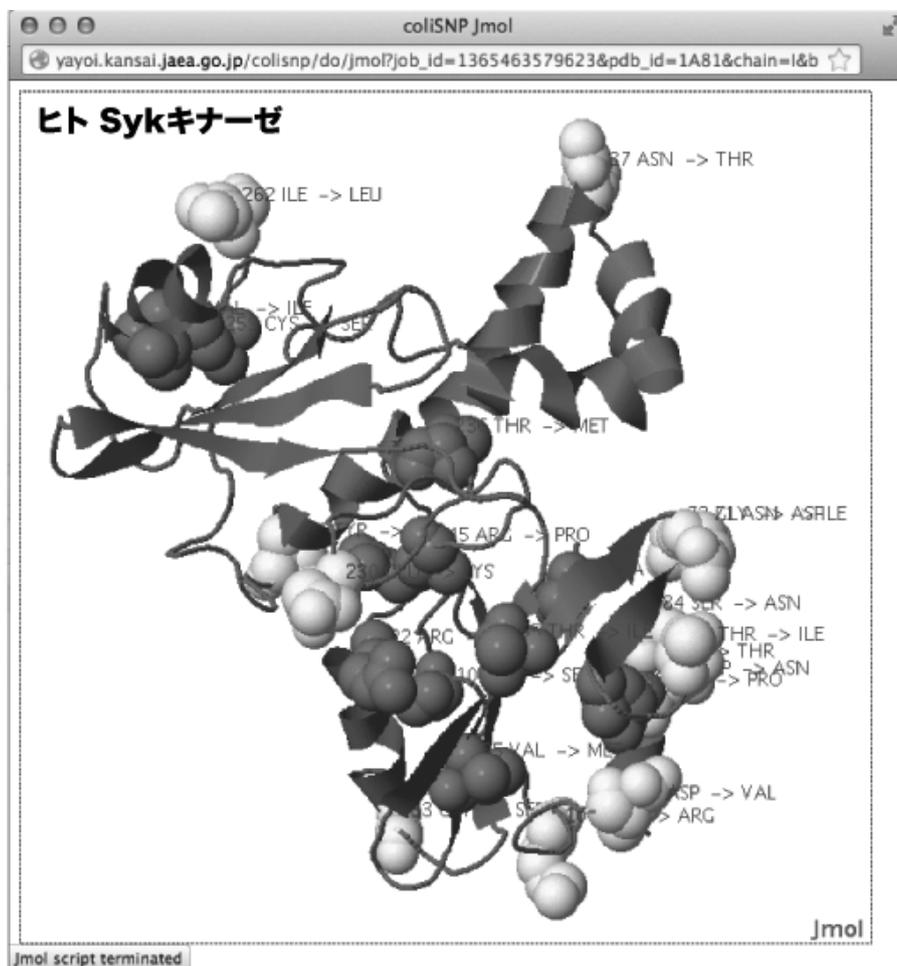


図3 ヒト Syk キナーゼに見られるヒトの nsSNP の位置

タンパク質の主鎖をリボンで示し、SNPにより変化するアミノ酸残基のみを空間重点モデルで示した。灰色の残基は実際にはタンパク質の内部に埋まっており、外部から見ることができない（水とはほとんど接触しない）。白の残基は、実際にはタンパク質の表面に位置し、水と接触することができる。これらの nsSNP が疾病と関与しているかは明らかになっていない。

#### 4. タンパク質の分子機能における 5W1H—いつ、どこで、どの原子/アミノ酸残基が、何を、どうするのか—

分子動力学シミュレーションを用いると、タンパク質の性質を物理学の言葉で表現することができるようになる。一方、生物学ではタンパク質の性質は遺伝子の言葉で表現される場合が多い。この二つの表現が結びつければ、タンパク質が機能する分子機構をより深く理解できるはずである。分子動力学シミュレーションは、いわば高い時間空間分解能を持った「分子顕微鏡観察」であるといえる。この特徴を生かして、「いつ、どこで (XYZ 座標)、どの原子/アミノ酸残基が、何を、どうするのか」をできる限り正確に表現すれば、個々のアミノ酸残基がそこにあるべき生物学的な根拠を物理的に裏付けることができる。

ここで、タンパク質内部の「空間」に特に着目し、時間分解 X 線結晶解析と分子動力学シミュレーションを組み合わせ分子機能を精密に解析した研究を紹介する。2009 年、低温条件下でのフラッシュフォトリス実験により、ミオグロビンのダイナミクスが高い空間分解能で追跡された<sup>20)</sup>。ミオグロビン内部では疎水性空洞以外の領域は原子が密に充填されているため、ガス分子が内部をどのように移動するのが明らかにされていなかった。そこで、著者の一人らは一酸化炭素分子の移動に対するミオグロビンの構造変化と熱揺らぎの効果を調べるため、ミオグロビン中に一酸化炭素を配置する統計力学的な平均力ポテンシャルを計算した<sup>21)</sup>。生体高分子の分子動力学シミュレーションを行う代表的なプログラムの一つである GROMACS<sup>22)</sup>を用いて、92.0 ナノ秒の分子動力学シミュレーションを実行した。そして、ミオグロビン分子内部に格子点を取り、各

点における一酸化炭素とミオグロビンの相互作用を評価した。フラッシュ照射によって、一酸化炭素分子がミオグロビン分子内空洞で熱揺らぎをすると、両者の相互作用によってミオグロビンの分子構造の統計集団は変化する。その結果、タンパク質の構造変化に応じて平均力ポテンシャルが変化している様子が明らかになった(図4)。図4aに示す状態は、ミオグロビン分子のヘム直下にあるXe1空洞の平均力ポテンシャルである。図4bの平均力ポテンシャルと明らかな差異が認められる。これは、Xe1空洞に一酸化炭素分子が入ると、ミオグロビン分子の外部に一酸化炭素分子が放出されやすくなることを意味する。つまりミオグロビンと一酸化炭素の相互作用によって、疎水性空洞の出入り口が開く機構(自己開閉モデル)が存在していることを意味する。このように、分子動力学シミュレーションにより、タンパク質構造の熱揺らぎによる統計集団

の分布を明らかにすることで、タンパク質の構造データに分子機能の機構に関する知見を付加することができる。ここで示された空洞が、億年の時間ドメインでどのような変化をともなっているのかは、これからの研究課題である。なお、ミオグロビンと一酸化炭素の相互作用に関する平均力ポテンシャルの地図(図4右)のデータは以下のURLからダウンロードでき、例えば、タンパク質可視化ツールVMD<sup>23)</sup>を用いてPC上で可視化することができる(<http://www.tb.phys.nagoya-u.ac.jp/~yamato/mysite1/yamato/PMF3D/PMF3D.htm>)。

### 5. トリプトファン合成酵素の基質輸送機構

ここまでに示した分子動力学シミュレーションでは、おもにピコ秒の時間ドメインの瞬間構造を利用することで明らかにできる、タンパク質機能の機械論的理解を示してき

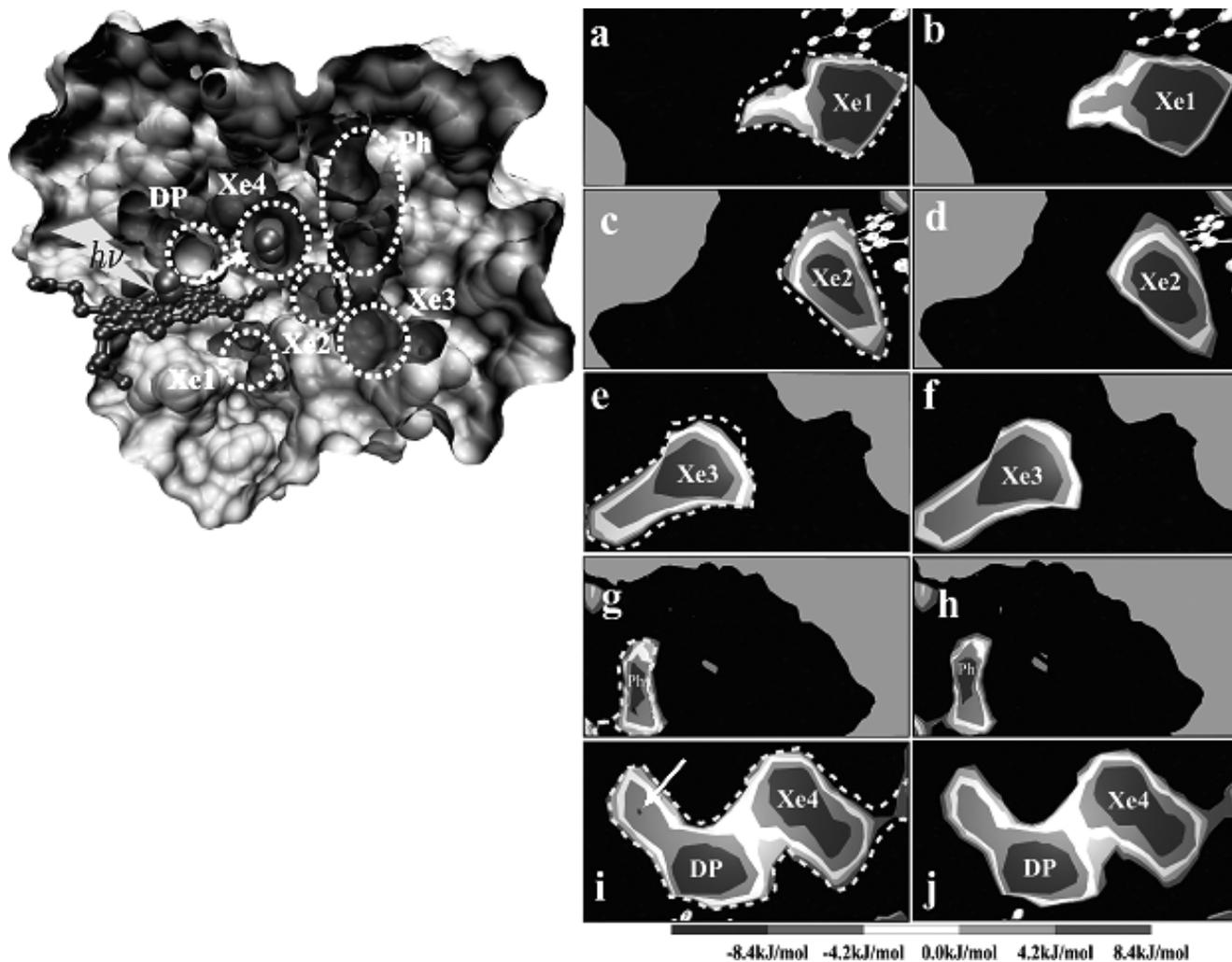


図4 ミオグロビン分子と一酸化炭素の相互作用

ミオグロビンの分子内空洞(左)と、ミオグロビンの分子内の一酸化炭素分子を配置する統計力学的平均力ポテンシャル(右)(文献21より一部転載)。ヘム鉄と結合した一酸化炭素を光で励起し、一酸化炭素がミオグロビン分子内のどこを漂うかをシミュレーションした。左図に白で示した空洞の名称と、右図の白文字の名称が対応する。右図の白文字周辺に、一酸化炭素分子が安定に滞在できる空洞が存在している。

た。しかし、分子動力学シミュレーションから得られるデータとして広く認識されているのは、時間発展による構造変化であろう。

トリプトファン合成酵素は、 $\alpha$ サブユニットと $\beta$ サブユニットのヘテロ二量体が二つ相互作用したヘテロ四量体構造である。トリプトファンの合成は二段階の化学反応で進行し、これらの化学反応は、ヘテロ二量体構造の中で完結する。 $\alpha$ サブユニットにおいて基質よりインドール環が切り出され、 $\beta$ サブユニットにおいてインドール環がセリンと結合する。1ミリ秒の時間ドメインで起こるこの酵素反応の過程において、 $\alpha$ サブユニットで切り出されたインドール環が、分子内に存在する25Åにわたる疎水性基質トンネルを通過して、 $\beta$ サブユニットの活性部位に移動するとされている<sup>24)</sup>。しかし、トリプトファン合成酵素のX線結晶解析構造を見てみると、基質トンネルは空間的に非常に狭い。そのため、酵素分子の熱揺らぎや酵素-基質相互作用が、インドール環の輸送に重要な役割を果していると考えられる。そこで著者らは、インドール環の輸送機構を明らかにするため、アンブレラサンプリング法とよばれる特別な分子動力学シミュレーション法を用いて、インドール環のトンネル内輸送の自由エネルギー曲面を解析中である(Tsuduki, T., Yano, M., Mitsutake, A., Yura, K., Okamoto, Y., & Yamato, T. 投稿準備中)。インドール環はトンネル内を $\alpha$ サブユニット側から $\beta$ サブユニット側へ移動する。よって移動経路は一次元で表現することができる。そこでこの移動経路を「反応座標」と命名し、シミュレーションの途中における基質の位置を、「基質と始点からの距離」-「基質と終点からの距離」で表現した。反応座標上のさまざまな位置で、基質がどの程度の位置エネルギーをもっているかがわかれば、基質の通り方を明らかにできる。

図5左に示すトリプトファン合成酵素を初期構造として、水中における分子動力学シミュレーションを先に示し

たソフトウェア GROMACS<sup>22)</sup>を用いて行った。シミュレーションを行った系には、総計75,112個の原子が含まれており、この原子間にはたらく力をすべて計算しながら時間発展させていった。最初の120ナノ秒は、トリプトファン合成酵素と、基質および水を含む系全体を平衡状態にするために費やした。次に基質の移動を、分子動力学シミュレーションが実行できる計算時間内に完了させるために、反応座標を56分割し、始点から順に各区分に基質が移動するように、基質に外力を加えた(アンブレラサンプリング法)。この方法を用いることで、実際には、ミリ秒の時間ドメインに属するであろう反応を、ナノ秒の時間ドメインの計算で実現することができ、反応自由エネルギー曲面を得ることができた(図5右)。驚くべきことに、反応自由エネルギーはなだらかではなく、大きな極大値と極小値をもっていた。このシミュレーション結果は、別の方法で行われた計算とも定性的に一致していることがわかった<sup>25)</sup>。ここで行った分子動力学シミュレーションは、まだ予備的な計算であり、その精度に問題がある。そこで、著者らは現在、レプリカ交換アンブレラサンプリング法を用いて、自由エネルギー曲面の精密化を進めている。さらにこの時間ドメインで明らかになってきている基質の停留ポイントが、億年の時間ドメインではどのようにしているのかを明らかにすることも、構造機能相関を知る上で重要になる。このような基質の停留ポイントが、トリプトファンの生合成反応において、本質的なことなのか、それともある限られた生物種のトリプトファン合成酵素でたまたま発生していることなのかは、この酵素の分子進化の過程を解析することで明らかにすることができる。

## 6. $\beta 2$ アドレナリン受容体内のアミノ酸残基間コミュニケーション

Gタンパク質共役受容体(GPCR)は、光刺激やリガン

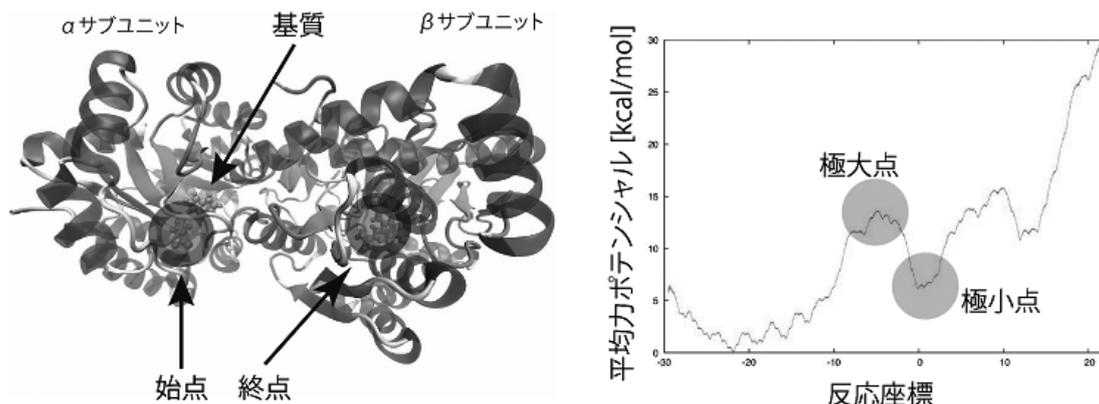


図5 トリプトファン合成酵素の立体構造とインドール環の自由エネルギー曲面  
タンパク質立体構造データベース PDB<sup>31)</sup>にある二つのトリプトファン合成酵素構造データ(2J9X<sup>32)</sup>と3CEP<sup>33)</sup>)をもとにして、分子動力学シミュレーションの初期構造を作成した。左に基質移動反応の始点と終点を示した。右にシミュレーションの結果得られた自由エネルギー曲面を示す。

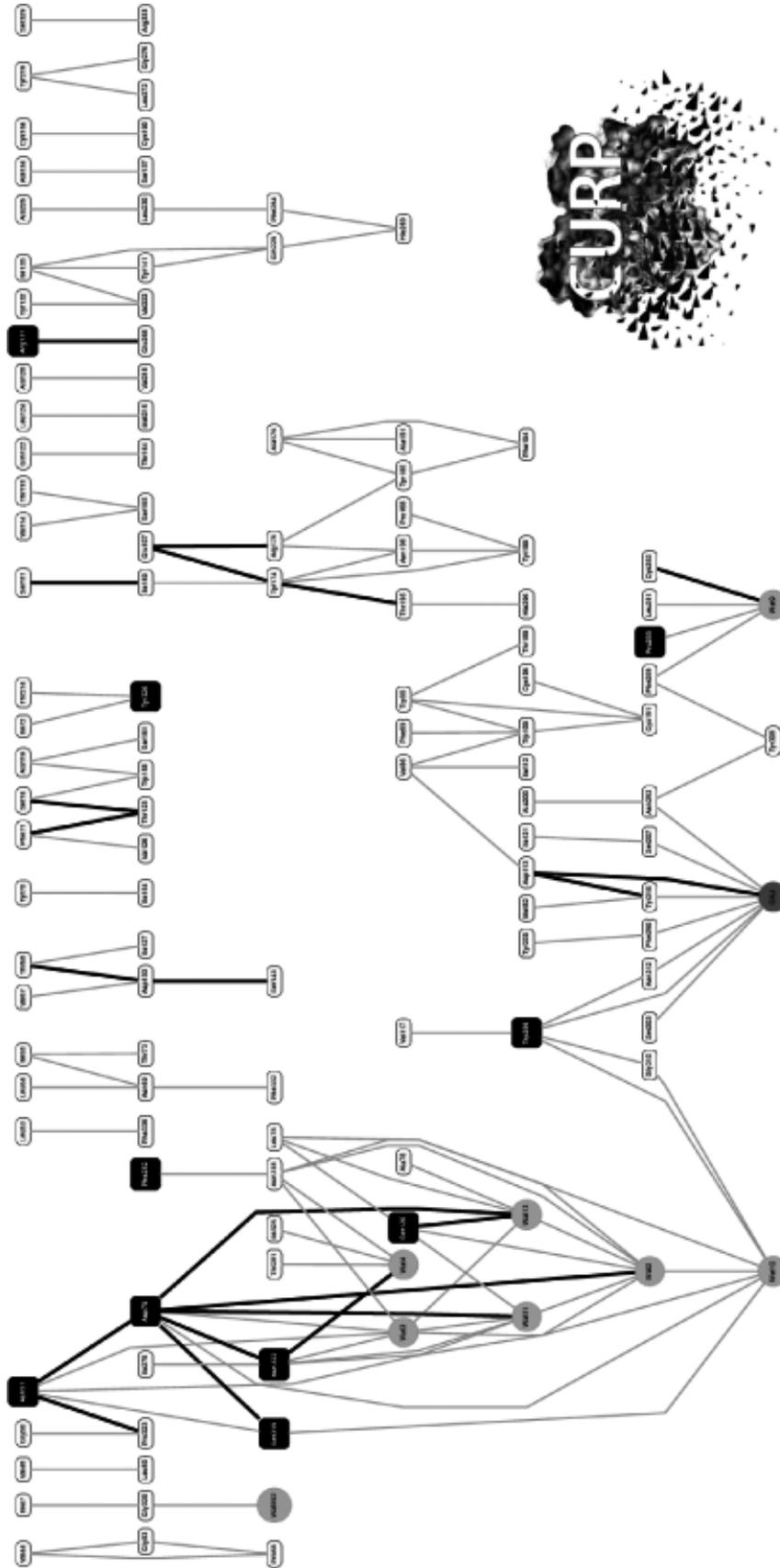


図6 ヒトβ2アドレナリン受容体の分子内情報伝達ネットワーク  
 ネットワークの各ノードは、アミノ酸残基/水分子/リガンドを表す。進化トレース法<sup>30)</sup>により、系統的に離れたタンパク質においても非常によく保存されていると同定されたアミノ酸残基のノードを黒、タンパク質分子の内部に安定に存在する水分子を灰色、その他のアミノ酸残基を白で着色した。リガンドであるカラゾロールは濃い灰色で表した。アミノ酸残基間エネルギー伝導度は、ノードをつなぐエッジの色で表した。エネギー伝導度にRT (R: 気体定数, T: 常温=310K) をかけた値が、0.1 (kcal/mol)<sup>2</sup>/fs以上1.0 (kcal/mol)<sup>2</sup>/fs未満の場合はうすい灰色の、1.0 (kcal/mol)<sup>2</sup>/fs以上の場合は黒の線を用いた。図には、溶媒露出表面積が0.3以上のアミノ酸残基、ペプチド結合したアミノ酸残基ペア、および膜貫通ヘリックス内にあって主鎖間で水素結合しているアミノ酸残基のペアは表示していない。

ド受容に応答して細胞内に情報を伝達する分子機械である。このような情報伝達タンパク質の機能は、因果関係の連鎖として理解されることが多い。タンパク質 A とタンパク質 B が相互作用することで「情報」が A から B に伝わり、B が別のタンパク質と相互作用し…、という具合である。しかし、各タンパク質には、相互作用をするたびに機械的な変化が発生しているはずであり、その変化の理解があつてはじめて、情報伝達タンパク質の構造機能相関を理解できたことになる。著者らは考えている。こういう意味では、GPCR の分子機構にはまだまだ明らかにしなければならないことが多く存在する。そこで著者の一人は、GPCR タンパク質のメンバーである光受容タンパク質のアミノ酸残基間に、どのような相互作用があるかを調べる新しい手法を開発し、分子内情報伝達経路を調べてきた<sup>26,27</sup>。従来、タンパク質内の相互作用は、分子動力学シミュレーション計算中に発生する原子間接触の頻度や、静的構造における原子間距離の近さで調べてきた。しかしここで開発した手法は、アミノ酸残基間のエネルギー伝導度（単位時間あたりにアミノ酸残基間でやり取りするエネルギー、irEC）を用いる点に特徴がある。さらに、タンパク質内のエネルギーの流れを容易に計算できるようにするため、独自のソフトウェア CURP (CURrent calculations for Proteins) を開発し<sup>28</sup>、公開の準備を進めている (URL: <http://www.tb.phys.nagoya-u.ac.jp/~yamato/CURP.htm>)。CURP プログラムは Python と FORTRAN 言語で書かれており、一般的

な分子動力学シミュレーションソフトウェアである AMBER<sup>29</sup>などのプログラムが作成する出力データを読み込み、分子内エネルギー（情報）伝達経路網を解析することができる。解析の結果は、二次元 (図 6) または三次元 (図 7) の模式図で出力することができる。

著者の一人は、この新規手法を用いて、 $\beta_2$  アドレナリン受容体の分子内情報伝達網を調べた。今回は分子動力学シミュレーションを AMBER プログラム<sup>29</sup>で実行した。ジパルミトイルホスファチジルコリン (DPPC) 脂質に埋め込んだ  $\beta_2$  アドレナリン受容体が、受容体の逆アゴニストであるカラゾロールと相互作用した際の分子内の情報伝達ネットワークは、図 6 のようになることがわかった。タンパク質の進化の過程において変異しないアミノ酸残基と、分子動力学シミュレーションの過程で分子内に安定に存在する水分子とが、ネットワーク内の情報伝達に重要な役割を果たしていることは興味深い。なお、リガンド (濃い灰色) は最も多くの結合をもっており、ネットワークのハブとして機能していることがわかる。この情報伝達ネットワークを  $\beta_2$  アドレナリン受容体の立体構造上に表示すると、分子内情報伝達機構が非常にわかりやすい (図 7) (Ishikura, T., Wilkins, A.D., Lichtarge, O., & Yamato, T. 投稿準備中)。 $\beta_2$  アドレナリン受容体の立体構造は、他の結晶状態でも決定されている。これらの構造についても同様の解析を行い、分子内情報伝達経路の違いの比較や、他の GPCR の解析結果との比較によって、さらに広い時間ドメインの

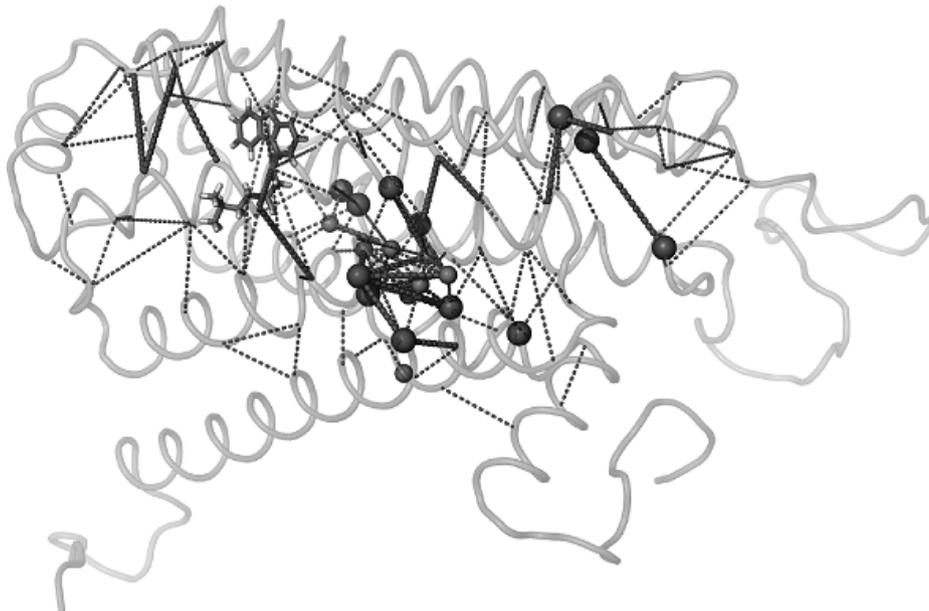


図 7 ヒト  $\beta_2$  アドレナリン受容体の構造<sup>35</sup>と情報伝達ネットワーク

進化トレース法<sup>30</sup>により、系統的に離れたタンパク質においても非常によく保存されていると特定されたアミノ酸残基を黒の、タンパク質分子の内部に安定に存在する水分子を灰色の球で示している。リガンドであるカラゾロールはスティックモデルで示した。エネルギー伝導度に  $RT$  ( $R$ : 気体定数,  $T$ : 常温 = 310K) をかけた値が、 $0.1 \text{ (kcal/mol)}^2/\text{fs}$  以上  $1.0 \text{ (kcal/mol)}^2/\text{fs}$  未満の場合はうすい灰色の、 $1.0 \text{ (kcal/mol)}^2/\text{fs}$  以上の場合は、黒の線を用いた。イオニックロックを形成するアミノ酸残基は濃い灰色の球で示した。

解析ができるようになると考えている。

## 7. おわりに

コンピュータを用いたタンパク質の動的構造シミュレーションの研究は1970年代後半に登場し、高速演算を可能とするコンピュータの出現により、シミュレーションできるタンパク質の大きさと時間を飛躍的に延ばしている<sup>30)</sup>。コンピュータを用いたタンパク質の遺伝情報解析に基づく生物学は、ゲノム塩基配列データが登場した20世紀後半に、大きく躍進したと言える。これらに、高速シーケンサーとタンパク質立体構造解析の技術革新による大量のデータが加わることで、ここまで記してきた、さまざまな研究ができるようになってきた。

次に進むべき道は実験家との協力による研究の発展だと、著者らは考えている。ここに記した構造機能相関は、ピコ秒からナノ秒の時間ドメインと億年の時間ドメインの融合によりわかってきたことである。この中には、分や時の時間ドメインに対する知見が含まれていない。しかし、この時間ドメインこそが、生物学において最も関心がもたれるところであろう。コンピュータ(ドライ)によるこの時間ドメインの解析は、精力的に挑戦はされているが、簡単なことではない。この時間ドメインの解析は、むしろウェットな実験が最も得意なところであろう。ピコ秒の時間ドメインを積み上げた先にある、この時間ドメインの解析へぜひとも発展していきたいと思っている。その結果として全時間ドメインの知見が接続され、生物を俯瞰的に見ることができるようになるのだと考えている。この文章がきっかけとなって、タンパク質の機能構造相関の研究が多角的に進展すれば何よりである。

## 文 献

- Shendure, J. & Ji, H. (2008) *Nat. Biotech.*, **26**, 1135–1145.
- Gupta, P.K. (2008) *Trends Biotech.*, **26**, 602–611.
- International Human Genome Sequencing Consortium (2001) *Nature*, **409**, 30–91.
- Lander, E.S. (2011) *Nature*, **470**, 187–197.
- The ENCODE Project Consortium (2012) *Nature*, **489**, 57–74.
- Lehner, B. (2013) *Nat. Rev. Genet.*, **14**, 168–178.
- 由良 敬 (2010) 化学と生物, **48**, 43–50.
- Miyazawa, Y., Nishioka, H., Yura, K., & Yamato, T. (2008) *Biophys. J.*, **94**, 2194–2203.
- Mees, A., Klar, T., Gnau, P., Hennecke, U., Eker, A.P.M., Carell, T., & Essén, L.-O. (2004) *Science*, **306**, 1789–1793.
- Medvedev, D. & Stuchebrukhov, A.A. (2001) *J. Theor. Biol.*, **210**, 237–248.
- Prytkova, T.R., Beratan, D.N., & Skourtis, S.S. (2007) *Proc. Natl. Acad. Sci. USA*, **104**, 802–807.
- Cooper, G.M. & Shendure, J. (2011) *Nat. Rev. Genet.*, **12**, 628–640.
- Kono, H., Yuasa, T., Nishiue, S., & Yura, K. (2008) *Nucleic Acids Res.*, **36**, D409–D413.
- Yura, K., Yamaguchi, A., & Go, M. (2006) *J. Struct. Funct. Genomics*, **7**, 65–76.
- Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., & Sirotkin, K. (2001) *Nucleic Acids Res.*, **29**, 308–411.
- Stitzel, N.O., Tseng, Y.Y., Pervouchine, D., Goddeau, D., Kasif, S., & Liang, J. (2003) *J. Mol. Biol.*, **327**, 1021–1030.
- Wang, Z. & Moul, J. (2001) *Hum. Mutat.*, **17**, 263–270.
- Yates, C.M. & Sternberg, M.J.E. (2013) *J. Mol. Biol.*, **425**, 1274–1286.
- Terui, H., Akagi, K., Kawame, H., & Yura, K. (2013) *J. Bio-med. Sci.*, **20**, 25.
- Tomita, A., Sato, T., Ichiyanagi, K., Nozawa, S., Ichikawa, H., Chollet, M., Kawai, F., Park, S.Y., Tsuduki, T., Yamato, T., Koshihara, S., & Adachi, S. (2009) *Proc. Natl. Acad. Sci. USA*, **106**, 2612–2616.
- Tsuduki, T., Tomita, A., Koshihara, S., Adachi, S., & Yamato, T. (2012) *J. Chem. Phys.*, **136**, 165101.
- Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008) *J. Chem. Theory Comp.*, **4**, 435–447.
- Humphrey, W., Dalke, A., & Schulten, K. (1996) *J. Molec. Graphics*, **4**, 33–38.
- Rauschel, F.M., Thoden, J.B., & Holden, H.M. (2003) *Acc. Chem. Res.*, **36**, 539–548.
- Fatmi, M.Q. & Chang, C.-e A. (2010) *PLoS Comput. Biol.*, **6**, e1000994.
- Yamato, T. (2009) in *Proteins: Energy, Heat, and Signal Flow* (Leitner, D. & Straub, J. eds.), pp. 129–147, Taylor and Francis/CRC Press.
- Ishikura, T. & Yamato, T. (2006) *Chem. Phys. Lett.*, **432**, 533–537.
- Ishikura, T., Hatano, T., & Yamato, T. (2012) *Chem. Phys. Lett.*, **539–540**, 144–150.
- Case, D.A., Darden, T.A., Cheatham, III, T.E., Simmerling, C. L., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A.W., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Wolf, R.M., Liu, J., Wu, X., Brozell, S.R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D.R., Mathews, D.H., Seetin, M.G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., & Kollman, P.A. (2012) AMBER 12, University of California, San Francisco.
- Sanbonmatsu, K.Y. & Tung, C.-S. (2007) *J. Struct. Biol.*, **157**, 470–480.
- Berman, H.M., Henrick, K., & Nakamura, H. (2003) *Nat. Struct. Biol.*, **10**, 980.
- Ngo, H., Kimmich, N., Harris, R., Niks, D., Blumenstein, L., Kulik, V., Barends, T.R., Schlichting, I., & Dunn, M.F. (2007) *Biochemistry*, **46**, 7740–7753.
- Barends, T.R., Domratcheva, T., Kulik, V., Blumenstein, L., Niks, D., Dunn, M.F., & Schlichting, I. (2008) *Chembiochem*, **9**, 1024–1028.
- Lichtarge, O., Bourne, H.R., & Cohen, F.E. (1996) *J. Mol. Biol.*, **257**, 342–358.
- Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S. G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I., Kobilka, B.K., & Stevens, R.C. (2007) *Science*, **318**, 1258–1265.