

# タンパク質立体構造情報からのゲノム機能解読・進化の解析

●郷 通子<sup>1)</sup> ◆高橋 健一<sup>1)</sup> ◆白井 剛<sup>2)</sup> ◆曾田 邦嗣<sup>3)</sup> ◆由良 敬<sup>4)</sup>

1) 長浜バイオ大学バイオサイエンス学部 2) 生物分子工学研究所生命情報解析部門 3) 長岡技術科学大学工学部 4) 日本原子力研究所 計算科学技術推進センター

## ＜研究の目的と進め方＞

ゲノム情報から機能情報を抽出するために、タンパク質原子座標などの連続量を導入した新しい予測法を開発して、個々のタンパク質の立体構造を予測し、基質との相互作用やリガンド結合などの局所的機能部位を予測することが第1の目的である。また、ゲノム内部構築の生物間共通性と多様性について、およびゲノムの構築とプロテオームとの関わりについて明らかとすることが第2の目的である。個々の具体的な研究目的と進め方を以下にまとめる。

### (1) モジュールにもとづくゲノムの機能予測

現在200種以上の生物においてゲノム塩基配列が判明し、ゲノム塩基配列にコードされているタンパク質の40%から50%は、その機能が推定できないことがわかってきた。そこでゲノム塩基配列から推定されるタンパク質のアミノ酸配列にもとづき、タンパク質の機能を予測する方法を開発することを本研究の目的とする。タンパク質立体構造の部品であるモジュールの構造とモジュールに局在する機能の情報から、3Dキーノートの概念を構築し、この概念にもとづいてゲノムにコードされているタンパク質の機能をアミノ酸配列から予測する。さらに共同研究により予測を実験的に検証し、機能予測法の正確さを評価する(由良、郷)。

### (2) 経験則を利用したタンパク質—リガンド予測法の開発

ゲノム研究の成果を細胞機能に結びつけるには、遺伝子産物の構造から、それらの分子による相互作用を原子レベルで予測する方法論が必要である。この研究では、タンパク質立体構造データベースであるPDBから経験則を抽出し、未知の複合体構造を予測するバイオインフォマティクスツールの開発を目的とした。具体的には、複合体中のリガンド周辺タンパク質原子空間配置の経験則を利用し検体タンパク質内の原子配置と比較することで分子間相互作用を予測するプログラムを開発する。

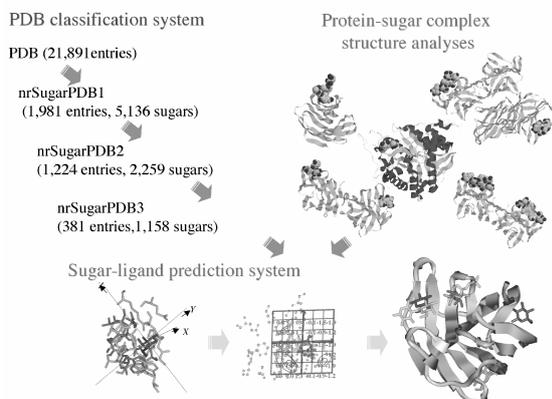


図1 PDB classification systemはPDBからの既知の構造情報のデータフローを表す。Protein-sugar complex structure analysisはX線結晶構造解析からのデータフローである。これら2つのデータフローから図下の経験的リガンド予測システムの完成を目指す。

また、計算機による方法論の検証も平衡して行うことを目的とする。これは、糖鎖結合タンパク質と糖鎖複合体の立体構造をX線結晶解析により多数明らかにし、プログラムによる予測と比較することでパラメータの改良を行い、また、プログラムが利用する経験則を供給することを目的とする(白井)。

### (3) タンパク質相互作用様式の解析と分子モデリング

タンパク質の立体構造情報と配列情報から相互作用部位・様式をどの程度予測することが可能かを知らるために、いくつかの種類のタンパク質について、相互作用部位と保存部位との関係や相互作用様式のパターンなどを明らかとし、相互作用様式の予測を試みる。また、ドメイン間相互作用様式を予測する手法を開発し、複数のドメインからなるタンパク質のモデリング手法の開発や、生物学的に重要な機能を予測するための基盤とする(郷)。

ゲノム解析や構造ゲノミクスプロジェクトはタンパク質機能に関する大量のデータを提供することになるが、現在でもPDBにはタンパク質と金属やヌクレオチド、糖との相互作用を研究するに十分なデータがあり、これらの相互作用の情報は、ゲノム配列からのタンパク質機能推定の基礎データとなる。しかしこれまで充分活用されているとは言えない状況である。そこで、使いやすい形に加工したデータベースとして公開し、それを利用して相互作用の特徴を抽出する(郷)。

### (4) 進化的に保存した立体構造上の特徴と機能の関連づけ

タンパク質立体構造情報からのゲノム機能解読を実現するには、立体構造と機能の相関を明らかにする必要がある。特に機能と密接に関係する構造的特徴を見つけることが重要である。タンパク質一般が示す最も基本的な機能は他の分子との相互作用とそれに伴う構造変化である。そこで、相互作用という点から見て重要と思われる構造的特徴として、タンパク質表面の形状や静電ポテンシャルの分布などが挙げられる。また、構造の柔らかさや大きな構造変化に関係があると思われる立体構造的な特徴として、タンパク質内部のすき間が考えられる。一方、機能的に重要なものは進化的に保存されているはずであるから、上記の構造的特徴がホモログなタンパク質において共通に存在するのかわかめることが重要である。そこで、機能とむすびつく構造的特徴を見つけるために、進化的に保存された立体構造上の特徴を抽出することを目的とする。特に、従来研究の未熟なすき間についてその保存性の有無と機能との関係を明らかとする(高橋)。

### (5) タンパク質の構造形成における疎水核の役割の解析

X線結晶構造解析法や多次元NMR溶液構造解析法の研究成果として、多くの蛋白質の立体構造が解明され、PDB等のデータベースに集積されている。これらの立体構造データを用いて、進化的に類縁関係にある蛋白質の立体構造を、ゲノム配列から予測する相同モデリング法が大きな成果を上げている。その一方で、相同モデリング法では、参照構造のない新規蛋白質については予測不能であり、その存在は未だ無視できない。蛋白質折り畳

みの分子機構の解明という純蛋白質科学の視点からも、折り畳みを単分子反応として捉え、最終生成物としての天然構造を物理法則のみに基づいて予測するab initio 予測能を強化する意義と必要性は高い。本研究の目的は、蓄積されている原子座標データベースを用いた蛋白質立体構造の網羅的な解析と、分子動力学模擬計算 (MDS) 法による蛋白質折り畳みの分子シミュレーションにより、ab initio 予測能の向上を支援できる情報を得ることである。

天然の球状蛋白質は、残基数< 20の小蛋白質を除いて、普遍的にその立体構造の内部に、非極性アミノ酸残基が凝集した疎水クラスタ (HPC) 構造を形成しており、その最重要部分は疎水核と呼ばれている (図2)。

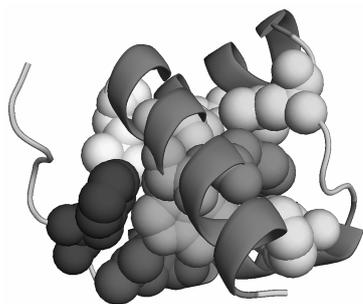


図2 Drosophila melanogaster 由来 Engrailed homeodomain の主鎖2次構造(リボン)と疎水クラスタ構造(球)。全53残基の内11残基が疎水クラスタを構成している。

従って、蛋白分子内部の構造的な特徴、特にHPCの構造特性の詳細を明らかにすることは、天然蛋白質の立体構造の設計指針を解釈する上で重要と考えられる。本研究の第1の目的は、上記の点を明らかにすることである。具体的には、蛋白質分子を構成する各原子が占有する空間の体積を、解析的、あるいは数値的に評価する方法を開発する。これを用いて、蛋白分子内部の原子充填状態を明らかにする。次に、与えられた蛋白質の立体構造に対して、そのHPC構成残基を自動同定するプログラムを開発する。これを立体構造データベースの多くの蛋白質に適用して、多数の蛋白質のHPCを網羅的に同定し、その構造特性、特にトポロジー (鎖の走り方) を明らかにする。

HPC残基の間には、折り畳みの途上で疎水相互作用による凝集力が働き、所謂、疎水凝集効果を生成することが予想される。そこでHPC残基を、残基番号順に仮想的なバネで繋いだ擬ペプチド鎖を計算機内に生成させ、MDS法によりその動力的挙動を解析する。これによって、折り畳み過程において重要な役割を果たすことが示唆されている疎水凝集効果の分子の実態と、折り畳み過程での分子鎖トポロジーの決定機構に関わる因子を解明することが、本研究の第2の目的である。

もし第1の目的が良く達成され、その結果から蛋白質の一次構造の特徴とHPC残基の分布の間に、何らかの規則性が見出されるならば、アミノ酸配列からのHPC残基の予測可能性が生まれる。更に第2の目的が良く達成されれば、上と組み合わせることにより、疎水クラスタの予測からトポロジー予測へ進む可能性が期待される。この可能性を調べることが、本研究の第3の目的である。(曾田)

#### (6) ゲノム構造とプロテオームとの関わり

ゲノム構造の成り立ちは生物種によって、どこまでが共通であり、あるいは多様であるのか。ゲノムの構築はプロテオームにどう反映しているのか。真核生物遺伝子

のエクソン・イントロン構造とプロテオームのタンパク質間相互作用、機能、進化とはどのように関わっているのか。真核生物のゲノムのイントロンの起源と機能を知ることは、ゲノムの構築原理を明らかとする第1ステップである。それに向けてタンパク質の構造部品モジュールとエクソンとの関係をゲノム規模で明らかにする (郷、由良)。遺伝子のエクソン・イントロン構造の存在によって可能となる選択的スプライシングに関して、ヒトでは遺伝子数が当初の想定よりもずっと少なく、限られた数の遺伝子から選択的スプライシングによりタンパク質の多様性を増している可能性が示唆されている。選択的スプライシングによりタンパク質の立体構造がどのように多様化するのか、それによりタンパク質の機能の多様化、相互作用の多様性獲得へどのようにつながるのかをタンパク質立体構造情報に基づいて明らかにする (郷、高橋)。

#### <研究開始時の研究計画>

- (1) タンパク質の部品であるモジュールをタンパク質の立体構造から同定し、それぞれのモジュールにどのような機能があるかを、タンパク質と低分子との相互作用情報から抽出する。低分子との相互作用構造座標にもとづきモジュールの機能部位を同定し、モジュールの立体構造とアミノ酸配列及び機能との関係から、3Dキーノートの概念を構築する。この概念にもとづいて、いろいろなモジュールの3Dキーノートを構築する。さらに、アミノ酸配列から3Dキーノートに相当する部位を予測する手法を開発し、判明しているすべてのゲノム塩基配列にたいして、3Dキーノートによるタンパク質の機能予測を行う。それらの予測を計算生物学的に評価するとともに、共同研究により実験的にも検証し、機能予測法の正確さを評価する (由良、郷)。

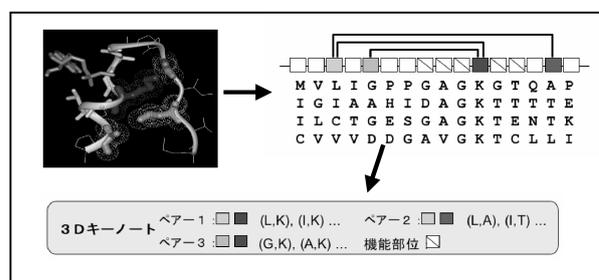


図3 タンパク質モジュールの構造・機能情報をパターン化して3Dキーノートを作成

- (2) (i) PDB経験則による相互作用予測法の開発。PDBの立体構造情報を処理し、原子分布統計をとるプログラムを作成する。また統計を元にタンパク質構造をスキャンニングし、リガンド結合位置を予測するアプリケーションを作成する。(ii) リガンド予測システムについて実験的検証を同時に進める。具体的には、コンジュリン (糖鎖結合タンパク質) や糖代謝酵素などとリガンドの複合体の立体構造をX線結晶解析により明らかにし、予測システムの検証、およびシステムの入力データを補強する (白井)。
- (3) プロテオームにおけるタンパク質間相互作用の実態を明らかにするために、立体構造既知の蛋白質におけるサブユニット間相互作用をグロビンファミリーで解析する。グロビンファミリーは、由来する生物種により様々な四次構造を形成することが知られている。これは、進化の過程で起きたタンパク質サブユニット表面のアミノ酸残基の置換が、サブユニッ

ト間の安定な接触様式を変えたことによると考えられる。そこで、サブユニット接触面とそれ以外のサブユニット表面を構成するアミノ酸残基を比較することで、サブユニット接触様式が変化するメカニズムの知見を得ることをめざす(郷)。

ドメイン間相互作用様式を予測する新しい手法を開発し、その手法をバーシカンと呼ばれる糖タンパク質のヒアルロン酸結合セグメント、Bドメイン+B'ドメインにおけるドメインの相互作用予測に適用する(郷)。

アミノアシルtRNA合成酵素(ARS)について立体構造情報と配列情報からRNAとの相互作用部位・相互作用様式の特徴を把握し、また、特異的相互作用と非特異的相互作用における違いを見つける(高橋)。

PDBに存在する、タンパク質とリガンド(金属やヌクレオチド、糖など)との相互作用の情報を活用するために、リガンドから検索ができ、タンパク質との相互作用様式を立体構造上に表現するデータベースを公開し、またそれを利用して各リガンドとの相互作用に関わるアミノ酸残基を解析する(郷)。

立体構造情報に基づき、核酸とタンパク質の相互作用様式をスコア化し、相互作用部位の予測能を評価する(白井、郷)。

(4) タンパク質の機能に関する立体構造上の特徴を抽出し、タンパク質の機能予測に利用することを目指すために、タンパク質内部のすき間に注目し、すき間と機能との関係を明らかにする。すき間の検出やすき間の位置の比較など、共通のすき間の有無を調べるための一連のプログラムを作成し、多くの生物種から立体構造の決定がなされているタンパク質ファミリーを選び、共通のすき間が存在するか解析する。存在する場合、機能部位との関係、タンパク質の種類との関係を解析する(高橋)。

(5) 目的1、2の各々に対応して、以下の研究計画を立てた:

### 1. 蛋白質の疎水クラスタの解析

#### (1.1) 蛋白質分子の体積量評価法の開発

分子内の各原子、残基、二次構造セグメント等、階層構造単位の原子充填状態を定量的に評価する為に必要な、各種体積の計算法を開発する。具体的には、

- (a) 原子単位での、van der Waals 体積、排除体積の計算プログラムを開発する。
- (b) 各原子への寄与に分割可能な、分子体積の解析的、数値的計算法を開発する。
- (c) 蛋白質分子体積の各原子への空間分割による原子占有体積の計算法を開発する。

#### (1.2) 疎水クラスタの構造特性の解析

- (a) 与えられた蛋白質の構造に対して、そのHPC残基を自動同定するプログラムを開発する。
- (b) (a) を用いて、リガンドを持つ蛋白質も含めて、SCOP folds の全蛋白質ファミリーが少なく共1個の代表蛋白質を持つように、基礎構造のデータと「疎水クラスタ残基データベース」を構築する。
- (c) (a) の体積計算法と(b)の「データベース」を用いて、約1300種の蛋白質のHPCについて、原子充填状態と構造特性を解析する。
- (d) (c) の結果に基づいて、HPC残基と蛋白質の立体構造の相関、及びHPCのトポロジー特性を明らかにする。
- (e) 以上の結果と、2次構造予測などの情報を併用して、1次配列からHPC鎖のトポロジーを予測する可能性を検討する。

## 2. 疎水凝集効果の特性と寄与の解析

折り畳み過程でHPC残基間に働く、疎水相互作用による分子内の大局的な引力としての疎水凝集効果の特性と寄与の大きさを調べる。

- (a) HPC残基を、特性を指定した仮想的なバネで残基番号順に繋いで生成させた擬ペプチド鎖に、MDS法を適用して、疎水凝集効果の寄与を含む原子座標値の時間変化データを得る。
- (b) 対照として、(a)と同様のMDSを、極性残基に置換した擬ペプチド鎖について行う。
- (c) (a)と(b)で生成する構造のクラスタ解析と両者の比較により、天然構造の分子鎖トポロジーの探索に果たす、分子内疎水凝集効果の役割と寄与を明らかにする。
- (d) 以上の結果に基づいて、蛋白質折り畳みに果たすHPC残基とそれ以外の残基の役割、協同効果を考察する。(曾田)
- (6) 線虫、ショウジョウバエ、アラビドプシス、マウス、ヒトなど多くの多細胞真核生物のゲノム解読が終了し、膨大なイントロン情報が蓄積しており、それらを利用してゲノムワイドのスケールで、真核生物ゲノムで明らかとなってきたイントロンとモジュールの相関の普遍性を調べる(郷、由良)。

ヒトの完全長cDNAの情報を利用して、大規模にスプライシング・バリエーションの同定を行い、選択的スプライシングにより変化(欠失、挿入、変異)のある領域を同定する。ホモログの立体構造情報を基に、選択的スプライシングが立体構造レベルでどういった効果を及ぼすのかを調べる。特に疎水性コアへ与える影響や他分子との相互作用部位へ与える影響に注目して解析する。これら立体構造上の変化がタンパク質の機能の変化、相互作用の変化にどうつながるのか推定する(郷、高橋)。

### <研究期間の成果>

(1) タンパク質のサブユニット間相互作用に関与する部位が、フィコシアニンの場合は1つないしは2つのモジュールに局在していることがわかった(2)。さらに酵素タンパク質において、タンパク質の機能部位がモジュールを単位として形成されていることが実験的に示すことができた(1、3、20)。これらの結果を受けて、3Dキーノートを構築するデータ収集のために、タンパク質の機能部位を立体構造情報から自動的に同定する方法の開発に着手した。タンパク質立体構造データベースに格納されているデータから低分子の情報を抽出整理し(21)、それら低分子とモジュールとの相互作用座標情報から、モジュールの機能部位を自動同定できるようにした。その結果として例えば、金属イオンを配位するモジュールの3Dキーノートを立体構造の情報から自動生成することができるようになった。DNAのリン酸基と相互作用するモジュールの3Dキーノートをきっかけにして、らん藻の形質転換の初期段階を担うと予測された遺伝子slr0197に、予測通り形質転換の機能があることを、領域「ゲノム生物学」の班員である池内昌彦、大森正之両氏との共同研究により実証することができた(4)。モジュール3Dキーノートによる機能予測が有効であることが判明したことを受けて、2001年度には各種機能の3Dキーノートの自動作成方法の確立を開始し、196種類の3Dキーノート作成に成功した。それらの予測精度を評価する方法を考案し適用した結果、これらの3Dキーノートは平均約85%の精度を持つことがわかった。

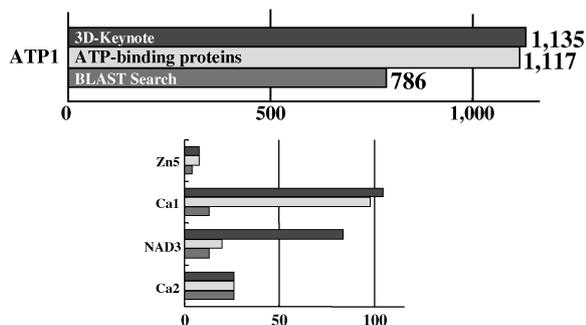


図4 3Dキーノートの精度を評価。機能既知(SwissProt)タンパク質の配列に3Dキーノートを適用し、本来の機能を予測できるか調べた。全 SwissProt 配列(99,134 本)において 8,414箇所まで3Dキーノートと一致。うち少なくとも7,186個(85%)は正しい。また、1,252 個(15%)はホモロジー検索では検出できない。

らん藻全ゲノムに3Dキーノートを適用したところ、機能未知のORFの約12%に機能を推定する手がかりを与えた。さらにヒト転写産物から推定されるタンパク質のアミノ酸配列に対して3Dキーノートを適用したところ、350個の機能不明cDNAに機能を推定する手がかりを与えることができた(19) (由良、郷)。

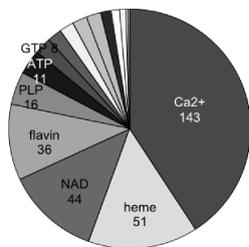


図5 ヒトの hypothetical protein の機能予測。未知の多様なカルシウム結合タンパク質の存在が予想される。

(2) PDBに含まれるタンパク質複合体の情報から、特定のリガンド周辺のタンパク質原子分布の統計を求め、それをスコアシステムとしてリガンド結合予測を行うプログラムを作成し、報告した(12)。これは、リガンド分子中でリジッドな原子団(プロトグループ)の座標を元に重複のないタンパク質を重ね合わせ、ある位置で観測されたタンパク質原子数の平均値からの標準偏差で計ったずれを得点として、特定のリガンド分子の安定性を見積もる方法である。

このプログラムは様々な種類のリガンドに対応することができるが、研究後半では実験的検証とも関連して、ポストゲノムで重要な「第3の生命鎖」として注目される糖鎖とタンパク質の相互作用を主に扱った。開発した方法の、糖鎖-タンパク質相互作用予測システムとしての性能をテストした結果、予測効率を厳密評価で19%(図6のStrict-Site base、予測第1位の糖結合部位から3A以内で実際の糖残基が見つかる割合)、ゆるやかな評価で69%(図6のtolerant-chain base、予測第3位までの結合部位から3A以内に実際の糖鎖が見つかる割合)であり、既存の方法よりかなり向上していた。

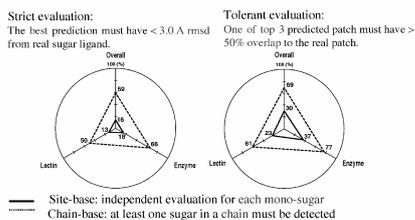


図6 リガンド結合予測プログラムの糖鎖予測結果。予測システムの結果(正解率)を円グラフで表したものの、strict evaluation はトップの予測のみを使い、tolerant evaluation は上位3位までの予測を用いる。それぞれ、レクチン、酵素を別々に評価した場合と、総合的中率を示している。

さらに、予測システムの実験的検証の為、糖鎖結合タンパク質レクチン(7,8,12,14,15,22,27,28)やセルラーゼなどの糖代謝酵素(5,13)の糖鎖複合体の結晶構造解析を行った。特にβガラクトシド特異的レクチンであるコンジェリンについては、20種を超える様々な糖鎖との結合定数のデータをもとに、網羅的な複合体構造解析を試みた。これは、結合定数-立体構造-予測プログラムによる安定性評価を対応させることで、より効率的に予測プログラムのパラメータをfittingできると考えたからである。

コンジェリンIIについては、7種(lactose、lacto-N-fucopentaose I、II、III、2'-fucosyllactose、lacto-N-tetraose、lacto-N-neotetraose)の新規複合体の構造解析を行った(図7)。

コンジェリンIIについては、主要な糖鎖(ガラクトース)結合部位が予測プログラムで最上位に予測されること、及び、2位以下の予測部位がこれまで指摘されていない位置に予測されたが、それらの予測結合部位が実際に延長結合部位として利用されていることを実験的に示すことに成功した。よって、当初計画した実験による予測システムの検証はある程度達成できた(白井)。

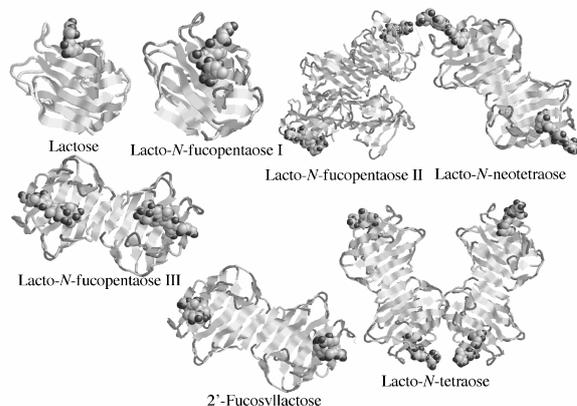


図7 コンジェリンII糖鎖複合体。コンジェリンの構造をリボンモデルで、結合した糖鎖の構造を空間充填モデルで示した。

(3) 生物種によりサブユニット間相互作用が様々なに変化しているグロビンファミリーについて、サブユニット間相互作用の特徴を解析し、相互作用の形成に必要な因子の解明を試み、サブユニット接触様式が変化するメカニズムの知見を得た。20種類のアミノ酸を親水性・中性・疎水性の3種類に分類し、ヒト、ナマコ、アカガイ、ユムシの四量体もしくは二量体を形成するヘモグロビンについて、サブユニット接触面とその他の表面における出現頻度を求めた。その結果、サブユニット接触面では、その他の表面に比べて主に中性残基の出現頻度が減り、その分疎水性残基の出現頻度が約10%増加していた。この接触面における疎水性残基の増加は、サブユニット間の接触をより安定化するのに寄与していると考えられた。ただし、解析したヘモグロビンにおいて接触面を形成するアミノ酸残基の数は平均約13残基と少なく、観察された疎水性残基の割合の増加はサブユニット接触面における1から2個の中性残基の疎水性残基への置換により説明できた。多くの中立的な変異はタンパク質表面において起こるため、この程度のサブユニット表面における疎水性残基へのアミノ酸残基置換は進化の過程で偶然にでも十分起こり得ると考えられる。実際に、動物の進化の過程でヘモグロビンの表面残基はほとんどすべて置換されており、例えばヒトヘモグロビンαサブユニットとア

カガイヘモグロビン I のサブユニット表面では、約91%のアミノ酸残基が異なっていた。一方、サブユニット接触面においても約91%のアミノ酸残基が異なっており、サブユニット接触面において特にアミノ酸残基の置換の割合が高いということは無かった。これらのことから、サブユニット表面において起きた中立的な変異が局所的に疎水性の高い面が生じるさせることが、新たなサブユニット接触面の形成と多様な接触様式を形成するための重要なメカニズムだと考えられた。以上のことは、ある生物種において観察されたサブユニット間やタンパク質間の相互作用様式が、他の生物種における同祖タンパク質において共通しているとは限らず、ホモロジーモデリングなどによるタンパク質複合体の立体構造推定には注意を必要とすることを示している(6) (郷)。

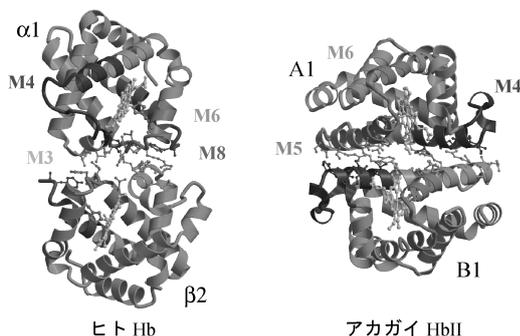


図8 ヒトとアカガイのヘモグロビンにおけるサブユニット間相互作用の違い。サブユニット間接触に関わるモジュールはヒトとアカガイのヘモグロビンで異なるが、ヘムを結合するモジュールであることは共通している。

ドメイン間相互作用様式を予測する新しい手法を開発し、その手法をリンクモジュールと呼ばれるドメインの相互作用予測に適用した。パーシカンと呼ばれる糖タンパク質は、心臓や血管、神経系などの広い範囲の組織に発現し、細胞表面にあるヒアルロン酸と結合する。これによって作られるタンパク質糖鎖集合体が、細胞間マトリックスの形成に重要であると考えられている。パーシカンのN末端にはA、B、B'の3つのサブドメインから構成されるG1ドメインが存在する。特にBとB'は、ヒアルロン酸結合タンパク質であるTSG-6やCD44のリンクモジュールと相同性がある重複したドメインであり、パーシカンのヒアルロン酸結合能は、BおよびB'により担われていると考えられる。実際にBとB'のみからなるポリペプチド(B-B'セグメント)がヒアルロン酸を結合することが実験的に示されている。そこで、B-B'セグメントの立体構造を推定し、B-B'セグメントがどのような様式でヒアルロン酸と結合するか推定した。まず、立体構造既知であるTSG-6由来のリンクモジュールを鋳型としてホモロジーモデリングによりBとB'それぞれの立体構造を推定した。さらに、TSG-6とCD44のリンクモジュールで調べられているヒアルロン酸結合に関わる残基の情報を用いて、BおよびB'それぞれにおけるヒアルロン酸結合部位を推定した。次に、BおよびB'を含むリンクモジュールファミリーのマルチプルアライメントおよび系統樹を作成してBとB'の重複後に特徴的に見られる変化を探することで、B-B'セグメントにおけるドメイン間相互作用部位の推定を行った。その結果、単独で存在するリンクモジュールと比較して、Bには特有の挿入、B'には特有の欠失が明らかとなった。特にB'特有の欠失がヒアルロン酸結合に関わると考えられる座位と重なることから、

これらの挿入・欠失は単独のリンクモジュールが重複した後に、ドメイン同士の接触のために起きた可能性が高いと考えられた。最後に、1)BおよびB'に存在する挿入・欠失がドメイン間相互作用部位の一部であると考えられること、2)BとB'間にはリンカー配列がないこと、3)BとB'の両方が直鎖状のヒアルロン酸と結合するために、ヒアルロン酸結合部位が直線上にあると考えられること、の3点からBとB'の相互作用の配向を推定し、B-B'セグメント全体の立体構造の推定を行った。

推定した立体構造上に直鎖状のヒアルロン酸分子を配置したところ、ヒアルロン酸結合面には10個の糖からなるヒアルロン酸が結合すると考えられた。これは実験的に示されているB-B'セグメントのヒアルロン酸最小結合単位と一致しており、推定したB-B'セグメントのモデルが妥当であることを意味している。また、作成したモデルにおいてヒアルロン酸と近接しているアミノ酸残基が実際のヒアルロン酸結合残基である可能性が高く、実験による検証が可能となった(18) (郷)。

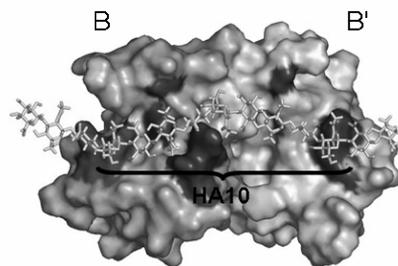


図9 パーシカン B-B'セグメントとヒアルロン酸複合体のモデル構造。ホモログスなドメインの機能部位、および、ドメイン重複の際に生じたと考えられる挿入・欠失部位の情報を用いてドメイン間相互作用部位の予測を行った。

タンパク質の立体構造情報と配列情報からRNAとの相互作用部位・相互作用様式の特徴を把握するために、ARSとtRNAとの複合体の立体構造情報、およびARS、tRNAの配列情報を組み合わせた解析を行い、以下の知見が得られた。ARSの立体構造上の配列保存部位の分布に局在性が見られた。その部位はtRNAとの相互作用部位に含まれるが、保存部位以外にもtRNAとの相互作用部位は存在した。しかしアイデンティティ決定に重要なtRNA上の部位との相互作用部位(特異的相互作用部位)とよい対応関係があった。特異的相互作用様式の特徴として、基本的には相互作用面の形状や物理化学的性質の相補性が見られ、いわゆる鍵と鍵穴のモデルに合う特徴が見られた。しかし、そのように単純には捉えられないものとして、複数の塩基を他の複数の塩基から識別する部位がいくつか存在した。たとえば、大腸菌のイソロイシルtRNA合成酵素において、同じアンチコドン結合部位を使って、グアニンとリジン塩基をそれ以外の塩基から識別している。その複雑な識別を実現するための相互作用様式については未解明であったが、立体構造上の特徴(相補性)や保存部位の分布、および分子モデリングの手法により、その相互作用様式を推定することができた(高橋)。

タンパク質に結合したヘテロアトムをPDBから検索し、3次構造上に表現するデータベースHetPDB-Naviを公開し毎月更新した(24,25)。検索機能の強化、および、ユーザーインターフェースの改良も進めた。このデータベースは、タンパク質とリガンの相互作用を解析し、さらに予測するための基盤として有用である。

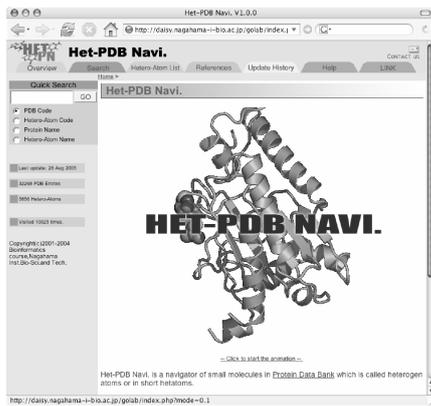


図10 リガンドデータベース HetPDB=Navi  
URL は <http://daisy.nagahama-i-bio.ac.jp/golab/hetpdbnavi.html>

PDB中には35,557の低分子が含まれており、3文字コードで4,189グループに分類されている(2003.4)。このうち29種の低分子でPDB中の低分子の約64%を占める。この中には、タンパク質の結晶化のために使用されたものも含まれているが、結晶化のために加えられた硫黄イオンが、タンパク質と結合したDNA分子のリン酸基と同じ位置をとる例などが知られており、人工的に加えられた分子の結合部位も生物学的に重要な部位であると考えられる。上位29種の内11種はカルシウム、亜鉛、マグネシウムなどの金属イオンであり、その次に糖分子やヌクレオチドが多く含まれることがわかった。各低分子とタンパク質の相互作用では、相互作用部位に位置する残基に異なる傾向が見られた。まず金属イオンについて：マグネシウムイオンとの相互作用相手は水分子が最も多く、ついでヌクレオチドが多い。アミノ酸残基ではAsp、Gluが多い。カルシウムイオンはアミノ酸側鎖との相互作用はAsp、Gluが多いが、主鎖のカルボニル基との相互作用が全体の1/4を占める。亜鉛イオンではCys、Hisが多くマグネシウムイオンやカルシウムイオンとは異なり銅イオンと同傾向であった。銅イオンは酸素原子との結合はまれだと考えられていたが、全体の1/4ほどがAsp、Glu、水の酸素原子であるという結果が得られた。次にヌクレオチドについて：ATP、GTP 結合する残基上位は以前から報告があったようにArg、Lys、Glyで共通だが、その次に多いのがATPではPheや脂肪族の残基であるのに対し、GTPではTrp、Asp、GluとATPと異なる傾向があった。NAD<sup>+</sup>、NADPHについて：ATPと同様の傾向を示すとの期待どおりGly、Argが上位だったが、Lys、PheはATPと異なり観察されなかった。PLPについて：Gly、Ser、Tyrが上位。Ser、TyrがPLP結合部位の特徴。糖について：マンノースは、Aspの側鎖との共有結合が知られているが、特にAspと相互作用しているといった傾向は無く、相互作用するアミノ酸残基に明確な傾向が無かった。グルコースは、マンノースと異なり負電荷や芳香環をもつ残基と相互作用する傾向があった。相互作用するタンパク質の機能の違い(マンノースは糖タンパク質、グルコースは細胞呼吸や代謝経路)と関連していると考えられる(21)(郷)。

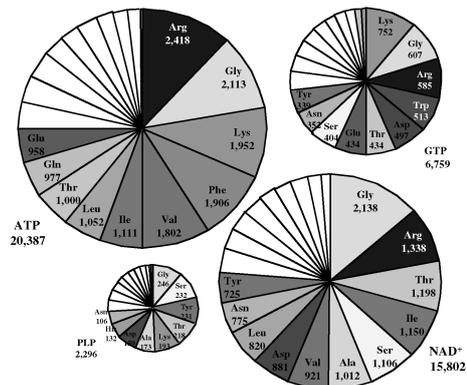


図11 リガンド(ATP、GTP、NAD<sup>+</sup>、PLP)結合部位を占めるアミノ酸残基の頻度

複合体の立体構造情報を利用して、核酸-タンパク質相互作用の経験的スコアを作成した。それにより従来知られている核酸結合モチーフとは異なる構造モチーフを見いだした。核酸結合部位の予測能を評価したところ、アデニン、グアニン、ニコチンアミド、フラビンのそれぞれについて、31、29、32、40%であった(23)(白井、郷)。(4)機能に関する立体構造上の特徴として蛋白質内部のすき間に注目し、その進化的保存性の解析を行った。118タンパク質ファミリーを解析し、少なくとも65%のタンパク質ファミリー内に共通のすき間を発見した。酵素と非酵素に分けて見た場合、すき間の保存性が見られるファミリーは酵素では77%、非酵素では52%と多少差があった。保存したすき間が触媒部位の近傍に存在する酵素ファミリーがあったが数例に留まった。機能との関係づけにはダイナミクス解析が不可欠であることが明らかとなった(高橋)。

#### (5) 1. 蛋白分子内部の原子充填率の解析

前項「研究計画」の(1.1)蛋白質分子の体積量評価法で述べた、全ての体積計算プログラムを開発した。有機小分子、アミノ酸、ジペプチド、球状蛋白質など、サイズの大きく異なる分子の体積計算にこれを適用して、実測が良く再現されることを確認した。更に、この計算法を蛋白質分子内部に適用し、側鎖の平均原子充填率が~0.62と、有機液体のそれより有意に高いことを見出した。この結果は、蛋白質の分子内部が固体的に密に原子充填されていることを示しており、HPC残基の空間配置が、蛋白質の立体構造安定性と共に、折り畳み構造自体の決定に関与している可能性を強く示唆する。

#### 2. 疎水クラスタのデータベース構築と構造特性

##### (2.1) 疎水クラスタのデータベースの構築

蛋白質の原子座標データから、そのHPC残基を同定するアルゴリズムを考案し、判定プログラムを開発した。これを用いて、SCOP foldsの全ファミリーの代表蛋白質を含む、2,000種以上の蛋白質のHPCを同定し、「疎水クラスタ残基データベース」を構築すると共に、1次構造上の分布、2次構造との相関等を解析した。

##### (2.2) 疎水クラスタ残基の分布

代表蛋白質の多くは1個のHPCを持つが、全く持たないものや、2個以上持つものもあり、綱種に依存する。数十残基以下の小さい蛋白質を除くと、HPC残基数は、アミノ酸残基数にほぼ比例する。蛋白分子鎖を、郷通子代表らが考案したcentripetal profile法を用いてモジュールに分割すると、HPC残基はモジュールの境界に現れる傾向を示すことが見出された。これは、モジュール構造が遺伝子の分断構造と対応しているときは、ゲノム配列からHPC残基が予測できる可能性を示している。HPC残基

とモジュール境界との相関は、進化による蛋白質の立体構造設計における、モジュール構造の物理的な意義を確認するものと考えられる。

### (2.3) 疎水クラスタ鎖のトポロジー

HPC残基は、1次配列上に偏在せず、ほぼ一様に分布していることも見出された。これは、HPC残基を繋いだ鎖が、蛋白分子鎖と実質的に同じトポロジー（空間的な走り方）をもつこと、すなわち鎖トポロジーの表現において、HPC鎖が分子鎖全体の簡約表現になっていることを意味している。この事実は、HPCが折り畳み過程で、極めて重要な役割を果たす可能性を示している。それは、HPC鎖のトポロジーの正否が、折り畳みの正否、つまり正しいトポロジーを持つ天然構造の形成を決定する可能性があるからである。また上の事実に基づいて、異なる蛋白質分子間のHPCトポロジーの類似性を評価することにより、立体構造の相同性を自動判定できる可能性も予測される。

### 3. 疎水クラスタ鎖の凝集効果とモジュール・ペプチド鎖の動力学

3種の蛋白質 (1abo: Abl tyrosine kinase Sh3 domain, 1enh: DNA binding protein (engrailed homeodomain), 1csp: bacillus subtilis cold shock protein) についてHPC残基からなる擬ペプチド鎖の、またbarnaseについてHPC残基間を連結するモジュール・ペプチド鎖のMDSを実行し、折り畳みにおける疎水効果の役割と構造形成傾向を調べた。

#### (3.1) 疎水クラスタ鎖の凝集効果

HPC残基を繋いで構築した擬ペプチド鎖に対するMDSの解析から、以下が明らかになった：

HPC残基は、非特異的な疎水凝集力によって分子鎖をコンパクト化し、変性状態にある蛋白質分子内に大局的な凝集力を生む。これは、空間的な広がり小さい構造集団を生成させて探索空間を大幅に減少させることにより、折り畳み過程を加速する役割を果たす可能性が予測される。然し疎水凝集した構造の安定性は低く、多くの異なる構造間を急速に渡り歩く。これは、局所的な安定構造にトラップされることなく、凝集状態を保った儘での構造空間の探索を可能にする。

#### (3.2) モジュール・ペプチド鎖の動力学

モジュール・ペプチド鎖の多くも、多様な構造間を急速に渡り歩くことが見出された。その一方で、単独でも天然構造に類似の局所構造を取り易いモジュール鎖は、折り畳みの鍵残基を提供する可能性があることが明らかになった。これから、蛋白質分子の折り畳みでは、大局的な疎水凝集力と、局所的・特異的な極性相互作用の間の協同効果が重要であることが、数種の擬ペプチド鎖とモジュール・ペプチド鎖の解析から強く示唆された。(曾田)

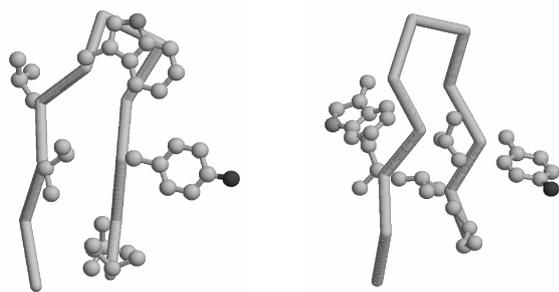


図12 barnase のモジュール# 5 の (a) 天然 b-ヘアピン構造と (b) モジュール・ペプチドの MDS で生成された天然類似構造

(6) ゲノム上のイントロン位置とタンパク質のモジュール境界との相関に関して、ゲノム情報に基づく大規模な解析を行ったところ、統計的に有意な相関があること、および、イントロンの数が増えるに従って、その相関が一層明確になることがわかった(郷、由良)。

選択的スプライシング (AS) 同定法を独自に開発し、ヒト完全長cDNAデータに対して網羅的に解析した結果、ASにより欠失、挿入、配列置換等の変化を受ける領域 (AS領域) の大多数はタンパク質の一般的なドメイン・サイズ (150アミノ酸残基程度) よりも短いことがわかった。次に、それ自身の、またはホモログの立体構造情報を基にAS領域をタンパク質立体構造上にマッピングし、疎水性コアとの関係、および他分子との相互作用部位との関係を解析した。その結果、立体構造上にマッピングできた243のAS領域のうち、177が疎水性コアに大きく影響を与えられられるものであり、139は相互作用部位を変化させるものであった。これらを合わせると、AS領域の83%がタンパク質の構造・機能上影響を与えられられるものとなる。また相互作用部位について、相互作用の相手により分類すると、約半数はタンパク質-タンパク質相互作用に関わる部位であった。

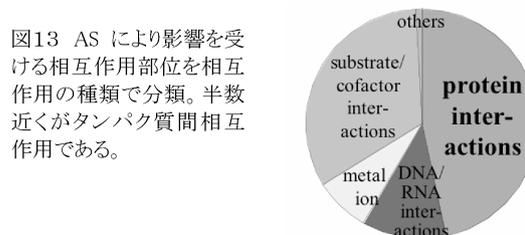


図13 AS により影響を受ける相互作用部位を相互作用の種類で分類。半数近くがタンパク質間相互作用である。

たとえば、G蛋白質共役受容体にASが起こることによりG蛋白質  $\alpha$  サブユニットとの相互作用部位が欠失し、また、G蛋白質  $\beta$  サブユニットにASが起こることによりG蛋白質  $\alpha$  サブユニットとの相互作用部位が欠失する例などがあった。この場合、ASにより、シグナル伝達経路のタンパク質間相互作用が変化することにより、シグナルの経路や伝わり方が変化することが考えられる。

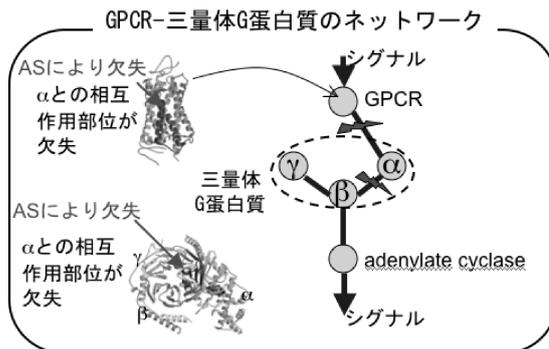


図14 タンパク質ネットワークに影響を及ぼす AS

これらのことから、ASによる変化はタンパク質の立体構造の不安定化または相互作用部位の消失などにつながり、それによりタンパク質間相互作用を変化させることが考えられ、ASがタンパク質ネットワークの制御に関わっている可能性が見出された(郷、高橋)。

### <国内外での成果の位置づけ>

(1) ゲノム塩基配列から推定される機能未知タンパク質の機能を予測する方法は、現在でも確固としたものはない。また機能予測方法開発においては、予測の実験的検証がなされていない場合が多い。このような状況におい

て、3Dキーノートによる機能予測の方法は他には存在しない独自のものであり、また、モジュールの情報の集積により開発された、日本のオリジナルな成果である。ヒトゲノムの機能未知のORFの機能を予測するためにも使われた(由良、郷)。

(2) 糖鎖研究のためのバイオインフォマティクスは整備が遅れているため、特に糖鎖に対象を絞ったことにより、本研究で開発したりガンド予測システムの評価は対外的にも良好であったと思われる。本課題実施後を含めて3件の海外のシンポジウムで、この成果について招待講演により紹介したが、反響は良好であった。プログラムについてはネットでの公開は出来ていないが、リクエストに応じて配布しており、現在までに数件の譲渡要求に応じた(白井)。

(3) 国内のウェット系研究者に興味を持ってもらい、推定した相互作用様式を検証するための共同研究を開始した(高橋)。

タンパク質に結合したヘテロアトムをPDBから検索し、3次構造上に表現するデータベースHetPDB-Naviiは、創薬のためのリガンド検索などに使われている。世界中から多くの研究者が高頻度にアクセスしており、Protein Data Bank (PDB) からの要望によりPDBにリンクされた。また Nucleic Acids Research のデータベース集に登録された(郷)。

(4) 国内外ともにあまり注目されていないタンパク質内部のすき間の重要性を明かにした(高橋)。

(5) 蛋白質の立体構造に果たす疎水相互作用の役割を調べた研究は多数ある。天然及び中間状態での疎水核に関する実験的研究も、数多い。然し蛋白質の立体構造安定性に関連した疎水核の構造解析に関する研究は、国外では2,3あるが、国内には未だない。また折り畳み過程に果たす疎水クラスタの役割の解明をめざす研究は皆無であり、本研究のユニークさはこの点にある。

蛋白質の天然構造は安定な特異的構造でなければならないが、これが、蛋白分子鎖内の大局的な凝集力である疎水相互作用を生成する疎水クラスタと、それに隣接する極性相互作用の協同効果で実現される可能性に着目している点が、本研究の独創的な点であり、本研究を進める意義はこの点にある。

蛋白質の立体構造の形成は、分子形に基づく斥力と多くの極性、非極性相互作用の集積の結果であるが、その中で、実験による解析が困難な疎水相互作用の役割、特に極性相互作用と協調作用する機構の分子的詳細を明らかにしていることの意義は大きい。蛋白質分子を、疎水クラスタ残基と連結モジュール・ププチド鎖に分解し、折り畳みに果たすそれらの役割を評価し、それらの協同として折り畳みの分子機構の解明と立体構造予測を目指す研究は殆ど無い(曾田)。

(6) モジュールは日本発のタンパク質の構造・機能概念であるため、イントロンとモジュールの相関は、国内外を通して全くオリジナルな成果である(郷、由良)。

大規模なcDNA配列解析などによりスプライシング・バリエーションが大量に見つかってきており、配列レベルの情報を集めたデータベースもいくつか存在するが、バリエーションの役割についてはわかっていないものばかりであり、今後多くの機能解析実験が必要である。そのような状況で、ASによりタンパク質の立体構造上のような効果が及ぶのかを解析した本研究のデータは、ASの役割を解明していくための基盤となり、また今後どのバリエーションに注目して機能解析実験を進めていくべきか指針を与えるものである(郷、高橋)。

#### <達成できなかったこと、予想外の困難、その理由>

(1) ゲノム塩基配列の決定速度は爆発的に向上し、それらに対する3Dキーノートの適用に時間的な困難が生じてきている。さらに、増大しているタンパク質立体構造からの3Dキーノート作成も多くの時間を要すようになってきている。これら両工程の完全自動化と並列化が必要となってきている(由良)。

(2) コンジェリンII複合体の解析は、予想外の結晶化の困難さから7種の解析にとどまった。これは、図7に示された複合体構造がすべて異なる結晶系から得られた事による。コンジェリンIIについては高分解能系を含め幾つかの結晶系が得られていたので、20数種類に及び複合体解析はそれほど困難でないと思われたが、上記の理由のため、それぞれの糖鎖複合体について結晶化条件の探索を繰り返す必要が生じ、結果として予定した複合体全てを解析することが出来なかった(白井)。

(4) 進化的に保存したすき間を見いだすことができたが、その機能的な意味付けにはダイナミクスの解析に踏み込む必要が生じた。新しい発見であるだけに、今後の解析が必須である(高橋)。

(5) (i) 他の予測法との結合など、1次配列のみからHPCを予測するのに必要な準備作業が未着手である。これは、計算量を削減する手法の開発がネックになっている。

(ii) 蛋白質立体構造のab initio予測における最大の問題は、要探索空間の大きさである。従って、ab initio予測能を上げるためには、探索空間を最小にする手法の開発が必須である。本研究は、HPCが立体構造安定化だけでなく、折り畳み過程でも重要な役割を果たすことを示した。もし一次構造からHPC残基が予測できれば、探索空間の削減に寄与できる可能性は高いが、残念乍ら未だ有効な手法が開発できていない。これについては、HPC残基の序列化による分類、2次構造予測との組み合わせ、MDS法の併用等により、他の可能性を追求している。

(iii) 同定されたHPC残基を、ランク付けして更に分類し、折り畳み過程で重要な $\phi$ 値の高い残基を決める因子が何かを予測する規則が見出されていない。

(iv) MDSで得られる擬ペプチド鎖の構造クラスタの多様性が高く、2次構造予測との組み合わせによる天然トポロジーの構造絞り込みが困難である。(曾田)

#### <今後の課題>

(1) ゲノムにコードされたすべてのORFのうち約半数は機能不明である。3Dキーノートを用いて機能を予測したORFの機能を検証するための共同研究を、すべての解読済みのゲノムに関して展開していきたい。多くの生物学者は特定の遺伝子やタンパク質については実現可能だが、網羅的な予測機能の検証については、今後、大規模なプロジェクトを組むなどして、共同研究により機能を検証することが今後の課題である(由良、郷)。

(2) リガンド予測システムについては、いまだ改良の余地があり(例えば分子動力学計算による予測結合構造の精密化など)、インターネットでの公開には至っていない。今後改良を進めると同時に、日本からの公開を目指している糖鎖構造データベースの一部として公開する案が浮上しているため、最終的には公開を目指したい。

コンジェリンIIと糖鎖複合体の網羅的構造解析については、目標を十分に達成できなかったが、解析は今後も継続して行い、当初目的の立体構造と予測構造の対応を解明したいと考えている(白井)。

(4) 本研究で見いだした保存性の見られるすき間について

て、タンパク質のダイナミクスとの関係を解析することにより、その機能的な重要性を明らかとしていきたい（高橋）。

(5) 本研究は *ab initio* 構造予測能の向上を支援することを当初からの目的としてきた。立体構造未知の蛋白分子鎖にMDS法を適用し、その到達構造から天然構造を予測することは、数十残基程度の小さい蛋白質を除いて、実際上不可能に近い。それは、蛋白分子鎖が取り得るコンホメーション空間が膨大なために、現在の計算機の能力を以てしても、折り畳みの全過程を再現することが困難だからである。一方、現実の蛋白分子は、原理的に取り得る構造空間に対して、無限小に近い空間しか探索していない。従って、この微小空間の実体を解明し、そのみを探索する方法を如何にして見出すかが、*ab initio* 法に課せられた大きな課題である。

本研究により、HPC残基間に作用する疎水凝集効果は、探索空間の大幅な削減を実現して、蛋白質の折り畳みを加速する効果を持つと共に、天然構造の鎖トポロジーの選別においても重要な役割も果たし得ることが明らかにされた。従って、HPC残基が与えられれば、天然構造を特定はできなくても、少数の候補構造に絞ることが可能になり、*ab initio* 予測能の向上に大きく貢献できる可能性がある。非極性残基の中には、分子表面に露出しているものも多い。本研究でも幾つかの可能性は試したが、一次構造情報からHPC残基候補を選び出す手法の開発が今後の課題となっている。(曾田)

(6) ASがタンパク質の立体構造へ与える影響について、分子モデリングなどにより精密に評価する方法を開発して、より精度の高い情報を発信していきたい。一方、実験からのフィードバックを得られるような共同研究体制を作りたいと考えている（郷、高橋）。

#### <研究期間の全成果公表リスト>

##### 1) 論文

1.0602071826

Kaneko, S., Iwamatsu, S., Kuno, A., Fujimoto, Z., Sato, Y., Yura, K., Go, M., Mizuno, H., Taira, K., Hasegawa, T., Kusakabe, I., Hayashi, K.: Module shuffling of a family F/10 xylanase: replacement of modules M4 and M5 of the FXYN of *Streptomyces olivaceoviridis* E-86 with those of the Cex of *Cellulomonas fimi*. *Protein Engng.* 13: 873-879 (2000).

2.0602071829

Kikuchi, H., Wako, H., Yura, K., Go, M. and Mimuro, M.: Significance of a two-domain structure in subunits of phycobiliproteins revealed by the normal mode analysis. *Biophysical Journal*, 79: 1587-1600 (2000).

3. 0602071755

Ahsan M.M., Kaneko, S., Wang, Q., Yura, K., Go, M. and Hayashi K.: Capacity of *Thermomonospora alba* XylA to impart thermostability in family F/10 chimeric xylanases. *Enzyme Microb. Technol.*, 28: 8-15 (2001).

4. 0602071800

Yoshihara, S., Geng, X.X., Okamoto, S., Yura, K., Murata, T., Go, M., Ohmori, M., and Ikeuchi, M.: Mutational analysis of genes involved in pilus structure, motility and transformation competency in the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.*, 42: 63-73 (2001).

5.0601302010

Shirai, T., Ishida, H., Noda, J., Yamane, T., Ozaki, K.,

Hakamada, Y. and Ito, S.: Crystal structure of alkaline cellulase K: insight into the alkaline adaptation of an industrial enzyme. *J. Mol. Biol.*, 310: 1079-1087 (2001).

6. 0602080950

Shionyu, M., Takahashi, K., Go, M.: Variable subunit contact and cooperativity of hemoglobins. *J. Mol. Evol.*, 53: 416-429 (2001).

7. 0303262132

Shirai, T., Matsui, Y., Shionyu-Mitsuyama, C., Yamane, T., Kamiya, H., Ishii, C., Ogawa, T., and Muramoto, K.: Crystal structure of a conger eel galectin (congerin II) at 1.45 Å resolution: Implication for the accelerated evolution of a new ligand-binding site following gene duplication. *J. Mol. Biol.* 322: 879-889 (2002).

8. 0601301959

Ogawa, T., Shirai, T., Yamane, T. and Muramoto, K.: Adaptive evolution of conger eel galectins. *Trend Glyco.* 14: 177-187 (2002).

9. 0602071705

Hiragi, Y., Seki, Y., Ichimura, K., and Soda, K.: "Direct Detection of the Protein Quaternary Structure and Denatured Entity by a Small-Angle Scattering Method: Guanidine Hydrochloride Denaturation of Chaperonin Protein GroEL", *J. Appl. Cryst.*, 35: 1-7 (2002).

10. 0602071708

Seki, Y., Tomizawa, T., Khechinashvili, N. N., and Soda, K.: "Contribution of Solvent Water to the Solution X-ray Scattering Profile of Proteins", *Biophys. Chem.*, 95: 235-252 (2002).

11. 0304291919/0304282223

Yamaguchi, A., Iwadate, M., Suzuki, E., Yura, K., Kawakita, S., Umeyama, H., and Go, M.: Enlarged FAMSBASE: protein 3D structure models of genome sequences for 41 species. *Nucleic Acids Research*, 31: 463-468 (2003).

12. 0308241341

Shionyu-Mitsuyama, C., Shirai, T., Ishida, H., Yamane, T.: An empirical approach for structure-based prediction of carbohydrate-binding sites on proteins. *Protein. Engng.*, 16: 467-478 (2003).

13. 0601302004

Shirai, T., Hung, V.-S., Akita, M., Hatada, Y., Ito, S., Horikoshi, K.: Crystallization and preliminary X-ray study of  $\alpha$ -glucosidase from *Geobacillus* sp. strain HTA-462, one of the deepest sea bacteria. *Acta Cryst. D*, 59: 1278-1279 (2003).

14. 0303262139

Shiba, K., Shirai, T., Homma, T., and Noda, T.: Translated products of tandem microgene repeats exhibit diverse properties also seen in natural proteins. *Protein. Engng.*, 16: 57-63 (2003).

15. 0404021214

Ogawa, T., Shirai, T., Shionyu-Mitsuyama, C., Yamane, T., Kamiya, H. and Muramoto, K.: The speciation of conger eel galectins by rapid adaptive evolution. *Glyco. J.*, 19: 451-458 (2003).

16. 0602071711

Higurashi, T., Hiragi, Y., Ichimura, K., Seki, Y., Soda, K., Mizobata, T. and Kawata, Y.: "Structural Stability and Solution Structure of Chaperonin GroES Heptamer Studied by Synchrotron Small-angle X-ray Scattering",

- J. Mol. Biol., 333: 605-620 (2003).
17. 0404101844  
Yada, T., Gotoh, M., Sato, T., Shionyu, M., Go, M., Kaseyama, H., Iwasaki, H., Kikuchi, N., Kwon, Y. D., Togayachi, A., Kudo, T., Watanabe, H., Narimatsu, H. and Kimata, K.: Chondroitin sulfate synthase-2. Molecular cloning and characterization of a novel human glycosyltransferase homologous to chondroitin sulfate glucuronyltransferase, which has dual enzymatic activities. *J. Biol. Chem.*, 278(32): 30235-30247 (2003).
18. 0404101713  
Matsumoto, K., Shionyu, M., Go, M., Shimizu, K., Shinomura, T., Kimata K. and Watanabe, H.: Distinct Interaction of Versican/PD-M with Hyaluronan and Link Protein. *J. Biol. Chem.*, 278(42): 41205-41212 (2003).
19. 0408232141  
Imanishi T et al.: Integrative annotation of 21,037 human genes validated by full-length cDNA clones, *PLoS Biology*, 2(6): 856-75 (2004).
20. 0408232142  
Kaneko, S., Ichinose, H., Fujimoto, Z., Kuno, A., Yura, K., Go, M., Mizuno, H., Kusakabe, I., and Kobayashi, H.: Structure and function of a family 10 beta-xylanase chimera of *Streptomyces olivaceoviridis* E-86 FXYN and *Cellulomonas fimi* Cex. *J. Biol. Chem.*, 279: 26619-26626 (2004).
21. 0403290830/0404112027  
Yamaguchi, A., Iida, K., Matsui, N., Tomoda, S., Yura, K., and Go, M.: Het-PDB Navi. : A database for protein-small molecule interactions. *J. Biochem. (Tokyo)* 135:79-84 (2004).
22. 0503071957  
Shionyu-Mitsuyama, C., Ito, Y., Konno, A., Miwa, Y., Ogawa, T., Muramoto, K., Shirai, T.: In vitro evolutionary thermostabilization of congerin II: a limited reproduction of natural protein evolution by artificial selection pressure. *J. Mol. Biol.*, 347: 385-397 (2005).
23. 0602081023  
Saito, M., Go, M., Shirai, T.: An empirical approach for detecting nucleotide-binding sites on proteins., *Protein Eng Des Sel.* 19(2):67-75 (2006).
- 2) データベース/ソフトウェア
24. 030209042  
データベース:HetPDB-Navi., Yamaguchi A., Tomoda, S., Yuka, K., Go, M.  
URL=<http://famsbase.bio.nagoya-u.ac.jp/golab/hetpdbnavi.html>
25. 0304301957  
データベース (更新) :Het-PDB Navi. , Go, M., Yamaguchi, A., Tomoda, S., Yura, K.  
<http://daisy.bio.nagoya-u.ac.jp/golab/hetpdbnavi.html>
- 3) 特許など
- 4) その他顕著なもの
26. 0304282230  
由良 敬, 郷 通子, 「モジュールに基づく機能予測-3D キーノート」, 蛋白質核酸酵素, 47(8), 1090-1096, (2002)
27. 0303262142  
白井 剛, 「タンパク質の構造予測とフォールド進化」, 日本結晶学会誌, 44, 20-24 (2002).
28. 0404021226  
白井 剛, 塩生-光山くらはら, 小川智久, 「加速塩基置換による蛋白質の機能・構造の進化」, 蛋白質核酸酵素, 48, 1913-1919, (2003).
29. 0602071722  
曾田 邦嗣, “ポストゲノム時代と蛋白質分子熱力学の重要性”, 熱測定 32, 7-12 (2005).
30. 0602071724  
曾田 邦嗣, “4.1 物理的相互作用”, タンパク質科学, 後藤祐児・桑島邦博・谷澤克行編, 化学同人, pp. 171-183 (2005)
31.  
曾田 邦嗣, “水の構造と特性”, “水分子のモデル”, “生体分子の水和”, “非天然状態の構造解析”, “揺らぎと実験観測量”, バイオインフォマティクス事典, 日本バイオインフォマティクス学会編, 共立出版 (2006, in press)