

研究開発課題別事後評価結果

1. 研究開発課題名

実践による超分子複合体モデリングシステムの開発

2. 代表研究者名

白井 剛 (長浜バイオ大学バイオサイエンス学部 教授)

3. 代表研究者による成果開発概要

3-1. 研究目的

細胞内の多くのタンパク質は超分子複合体を形成して協調的に機能する。従って今後、構造生物学研究は複合体解析に主眼を移さなければ生命現象の本質を理解できない。残念ながらそのための技術は発展途上にあり、バイオインフォマティクスにおいても実験技術においても多大の努力が要求される。

タンパク質構造モデリングの信頼性は未だ十分とは言えない。さらに深刻な問題として、現在利用可能なアプリケーションは実際に構造解析を行う研究者が必要とする情報・機能が十分に提供されていないという問題がある。問題解決にはバイオインフォマティクスと構造解析実験の緊密な協力が必要である。

本課題では、情報・構造・機能生物学を専門とする研究者が一体となってチームを形成し、超分子複合体モデリングに必要なとされる toolbox system を実践的に開発する。未解明の複合体の多い古細菌 3R (DNA Replication, Recombination, Repair) 系をモデル系として構造・機能解析実験を行いつつ、その過程で必要となるアルゴリズムやソフトウェアツールを同時並行して整備・開発することを目的とする。

3-2. 研究開発概要

超分子複合体モデリングに求められるのは、多岐にわたるモデリングの各段階で必要になる手法が不足なく統一された仕様・フォーマットで存在することである。すなわち、分子間インターフェースの予測やモデルのアセンブリなどのステップは個別のツールとして実行可能であり、それらツールの間でのデータの入出力フォーマットが統一されている toolbox system が理想である。本課題では 1) データベースサーベイ、2) インターフェース予測、3) サブユニットアセンブリ、4) 実験結果とのフィッティングなどのツール群を、複合体構造解析実験との緊密な連携のもとで開発する。

データベースサーベイトールとして、蓄積された 1 次情報である PDB (Protein Data Bank) から分子構造および相互作用の情報を全自動で 2 次データベース化するシステムを開発する。超分子モデリングはこの 2 次データベースから引き出されたモデルを、配列・立体構造情報からの経験的インターフェース予測情報と高度に自動化されたサブユニットアセンブリツール群を用いて集合させることで実現する。

同時に、ゲノムの複製や維持に必須の 3R 超分子複合体である PCNA (DNA クランプ) 複合体の構造解析実験を進行させる。生化学解析により複合体の機能解析 (および構造解析実験に対するサンプル提供) を、X 線結晶解析により複合体サブユニットの精密構造を、電子顕微鏡単粒子解析により複合体の構造を求め、その結果をツールによるモデリングの結果とフィッティングツールを用いて比較する。具体的な対象として、古細菌 *P. furiosus* の liga

se/PCNA/DNA、polymerase/PCNA/DNA、helicase/PCNA/DNA複合体を構造解析する。

3-3. 研究成果概要

本課題の成果として、超分子モデリングの基本過程を完遂できるシステムが開発された。データベースサーベイツールとして SIRD (Structure Interaction Relational Database) system を作成した。これは、PDB を全自動で構造類似性・相互作用類似性に基づいて分類することができる。このツールを使えば、ある構造を持ったタンパク質が別の分子(タンパク質、DNA などのポリマー、低分子リガンド)と結合したモデルを取得できる。ここから選抜されたモデルを集合させ超分子複合体モデルを構築するために、モデルアセンブリーツール群を開発した。主要なツールとしては、2つの座標ファイルから類似した分子を捜して重ね合わせを行うツールや、結晶構造の非対称単位から複合体を再現するツールがある。また、得られたモデルを実験結果に当てはめるフィッティングツールを開発した。これは電子顕微鏡実験などから得られた低分解能マップに対してモデル座標を分子置換するものである。

また、平行して行った生化学実験・構造解析実験から 3R 複合体の機能・構造の新たな知見が多数蓄積された。これらの機能解析の結果はモデルの妥当性の検証に利用された。X線結晶解析では、ligase、helicase、PCNA 複合体などの構造解析に成功し、これらはモデリングの精密部品構造として役立った。電子顕微鏡解析では ligase/PCNA/DNA、helicase/PCNA/DNA、polymerase/PCNA/DNA 複合体の解析に成功した。これらは機能状態の PCNA 超分子複合体解析としては世界初の成果となった。

最終的に、作成されたモデルとこれらの実験結果を比較検証する段階までたどり着いた。計算機上で作成されたモデルは電子顕微鏡実験から得られたマップとよく一致しており、他の実験事実とも符合するものであることが確認された。

本研究開発課題により、モデリングシステムが生み出す超分子複合体モデルが有効であることが示された。これは同時に、データベースに蓄積された情報を整理・活用するストラテジーの有効性を示している。データベースから抽出できる情報はまだ多く残されているので、今後も本システムを発展させることで超分子モデリングの性能を向上させられると期待される。

4. 事後評価結果

4-1. 当初計画の達成度

○当初の研究開発計画から見た進捗状況や達成度等

超分子複合体モデリングシステムを整備し、データベースサーベイツールからフィッティングツールまでを開発した。複数のタンパク質複合体の構造解析の実験結果との比較により、モデリングに成功したことから、モデリング手法の有効性が示された。

○当初計画では想定されていなかった新たな展開

研究開発を通して、システムを構造変化や相互作用変化に十分に利用できるよう、改良する糸口を見つけた。

4-2. 知的財産権、外部発表（論文等）等研究開発成果の状況

○知的財産権

開発されたモデリングシステムは新規性が高く、機能的に重要な構造の同定に結びつく

可能性がある。

○外部発表（論文等）

比較的出ている。

4-3. 研究開発成果の公開による波及効果

○データベース、プログラムや論文、学会発表など、これまでの成果公開によるバイオインフォマティクス領域およびライフサイエンス分野全体への影響や効果について

開発されたツール群は公開されて間もないので、実際の効果については評価できないが、研究開発成果報告会での発表を通じ、既に構造生物学に関心のある研究者には大きな影響や効果を与えたと思われる。

○国内外の類似研究開発の動向と比較して本課題の意義

独自のツールを活用したモデリングシステムの開発と実験による実証を組み合わせた点はユニークである。また、バイオインフォマティクスと実験解析の両方から活用できるデータベースからの情報抽出法を開発しており、従来の構造分類データベースが構造データの急激な増加に対応できていない中、今後、構造データベースのデータ解析の一つのモデルケースになれば、大きな意義が出てくる。

4-4. 成果の実用化の可能性及び成果から予想される波及効果

○今後、成果の生命科学、研究基盤、産業への貢献度

実験的検証によるフィードバックを含んだシステムであり、実験系研究者からの信頼感が高い。本課題で開発された方法を他の一般の複合体にも適用していくにはまだ研究が必要と思われるが、分子複合体のモデリングの研究分野での研究基盤になる可能性がある。さらに期待するならば、超分子構造をとる分子間の相互作用のメカニズム解明にまで踏み込むデータベースの開発を達成できれば、きわめて信頼性の高いものになるであろう。

4-5. 総合評価

研究開発計画通りの成果が見られ、バイオインフォマティクス研究の進展に貢献した。 バイオインフォマティクスと実験解析をうまく組み合わせることにより、実際に複数の複合体のモデリングに成功しており、研究成果自体も意義が大きい。開発されたデータベースやツールがより普及することが望ましい。今後の構造データの急激な増加にも対応できるだけの高性能な自動解析方法が確立されれば大きな意義がある。

5. 主な論文発表等

T. Shirai, V.-S. Hung, K. Morinaka, T. Kobayashi, and S. Ito, Crystal structure of GH13 α -glucosidase GSJ from one of the deepest sea bacteria. *Proteins*, 73, 126-133 (2008)

S. Kiyonari, M. Uchimura, T. Shirai & Y. Ishino, Physical and Functional Interactions between Uracil-DNA glycosylase and proliferating cell nuclear antigen from the euryarchaeon *Pyrococcus furiosus*. *J. Biol. Chem.* 283, 24185-24193 (2008)

T. Yoshimochi, R. Fujikane, M. Kawanami, F. Matsunaga, and Y. Ishino, The GINS complex from

Pyrococcus furiosus stimulates the MCM helicase activity. *J. Biol. Chem.* 283, 1601–1609 (2008)

K. Mayanagi, Y. Fujiwara, T. Miyata, and K. Morikawa, Electron microscopic single particle analysis of a tetrameric RuvA/RuvB/Holliday junction DNA complex. *BBRC* 365, 273–278, (2008)

K. Sasaki, T. Ose, N. Okamoto, K. Maenaka, T. Tanaka, H. Masai, M. Saito, T. Shirai, D. Kohda, Structural basis of the 3' -end recognition of a leading strand in stalled replication forks by PriA., *The EMBO Journal*, 26, 2584–2593 (2007)

参考

論文発表 国内 5 件、海外 21 件

口頭発表 国内 15 件、海外 4 件

ポスター発表 国内 55 件、海外 7 件

公開されている成果プログラムおよびデータベース等

・ SIRD

<http://sird.nagahama-i-bio.ac.jp/sird/>

自動で PDB の構造と相互作用の情報を分類するプログラム群 SIRDp (processor) とユーザーインタフェースとウェブサイトとを兼ねる SIRDi (interface) からなる。SIRDi は HP 形式で利用できるほか、ユーザーがローカルに GUI としても用いる。本研究開発のモデリングツールは SIRD データベースをモデリングテンプレートとして利用する。最新版 080909 は 309, 286 エントリーの情報を含む。

・ BIOMOL

<http://sird.nagahama-i-bio.ac.jp/sird/> の 3. SIRD modeling – 3. 1: Make quaternary complex より Web ツールとして利用可能

BIOMOL は PDB 形式の結晶構造 (非対称単位の原子座標が含まれる) を元に、生理学的な複合体構造 (4 次構造) を推定するツール。BIOMOL はデフォルト内部パラメータで自動推定を行えるほか、ユーザー指定のパラメータ/モードで 4 次構造推定ができる。

・ EXT

<http://sird.nagahama-i-bio.ac.jp/sird/> の 3. SIRD modeling – 3. 1: Make complex より Web ツールとして利用可能

SIRD を利用し、クエリ構造と類似した既知タンパク質と相互作用する分子をクエリ構造に重ね合わせて示すプログラム。モデリングのテンプレートを提供する。相互作用する分子として i) タンパク質ドメイン、ii) 非タンパク質分子 (DNA/RNA/糖鎖などのポリマー、低分子化合物)、iii) 両方を選択可能である。類似した位置に存在する類似した分子のクラスタリングにより代表モデルを選ぶ機能がある。

・ CMP

<http://sird.nagahama-i-bio.ac.jp/sird/> の 3. SIRD modeling – 3. 2: Assemble complex より Web ツールとして利用可能

PDB フォーマットの分子座標ファイルを 2 つ与えると、重ね合わせ可能な分子 (タンパク

質、ペプチド、DNA、RNA、糖鎖などのポリマー、および低分子リガンドを含む)を検索し、それらを重ね合わせることによって複合体分子モデルを構築する。