

実践による超分子複合体モデリングシステムの開発

長浜バイオ大学バイオサイエンス学部

白井 剛

Practical development of the supramolecular modeling system

Tsuyoshi Shirai

Nagahama Institute of BioScience and Technology

The researchers in bioinformatics, biochemistry, X-ray crystallography, and electron microscopy have collaborated to develop the practical software toolbox for supramolecular modeling. The database tool named SIRD (Structure Interaction Relational Database) was designed to extract complex structures appropriate for modeling from known biomolecular structure database, and a package of tools to assemble these parts into a supramolecule was also developed. These developments have been executed under intensive collaboration with experimental determinations of novel 3R (stands for Replication, Repair, and Recombination of DNA) complex structures, namely, ligase/PCNA/DNA, polymerase/PCNA/DNA, and helicase/PCNA/DNA complexes, and the developed system was proven against the experimental structures.

1. はじめに

「生物学は積み木細工である」と指摘したのは、意外なことに理論物理学者の湯川秀樹博士であったという。現在この言葉は非常に的を射ている。なぜなら我々は今まさに「積み木」をしなくてはならないからである。この場合の積み木はタンパク質・DNA・RNAなどの生体分子である。生体分子は積み木のように寄り集まって超分子複合体という巨大な構造を形成して働き、生命活動を支えている。しかし、この超分子複合体を効率よく構造解析したり、計算機を使って構造予測する手段はまだない。

他の研究者から超分子構造に関するバイオインフォマティクスによる所見を求められてデータベースを精査した時に、なにかしら超分子相互作用のヒントになる構造に巡り合う機会が、特に最近増大していることが実感される。例えばA-B-C複合体の構造が知りたい時、A-B'とA'-C(A'、B'はA、Bに類似した分子)の構造が知られている様な場合である。

すなわち構造ゲノミクスの成果として、既に部品の情報はデータベースに豊富に蓄積されていると考えられる。問題はそれを生かすための道具(計算機ツール)の整備が遅れていることである。そこで我々は、バイオインフォマティクス研究と構造解析研究が丸となって超分子複合体のモデリングツールシステムを開発する事を目指した。

2. 研究開発の成果

図1は本報告の内容を模式的に示しており、以下にこの図を参照しつつ各部分を解説する。

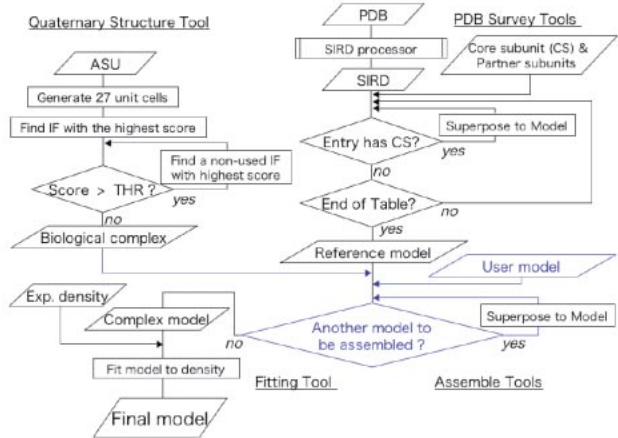
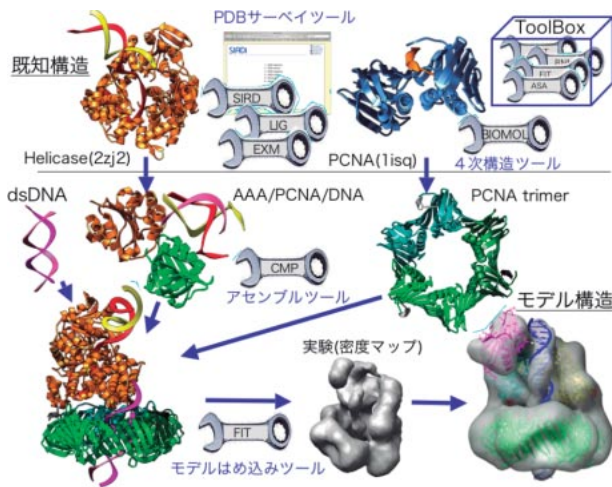


図1 本報告の概要(左)とモデリングフローチャート(右)
 (右)は(左)をフローチャート化したものである

2.1 PDB サーベイツール群 (PDB Survey tools)

ここで目指したのは、蓄積された1次情報であるPDB (Protein Data Bank) に収められた分子構造および相互作用の情報を最大限活用するためのツールの開発である。これらの情報は多大な年月と資源を投資して蓄積されたものであり、利用し尽くさなければ「もったいない」。

しかし、PDBは分子間相互作用情報の宝庫であるが、その情報はユーザーが容易に抽出できる形になっていない。そのため1次データベース(PDB)を全自動で2次データベースにプロセスし、そこから相互作用情報を抽出するデータベースツール(SIRD: Structure Interaction Relational Database system)を整備した。

SIRDは類似したタンパク質の構造を分類しテーブル化する。また、高速グラフマッチアルゴリズムを用いて、非タンパク質(低分子リガンド)の構造も分類し、それらの分子間相互作用をテーブル化する^[1]。

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

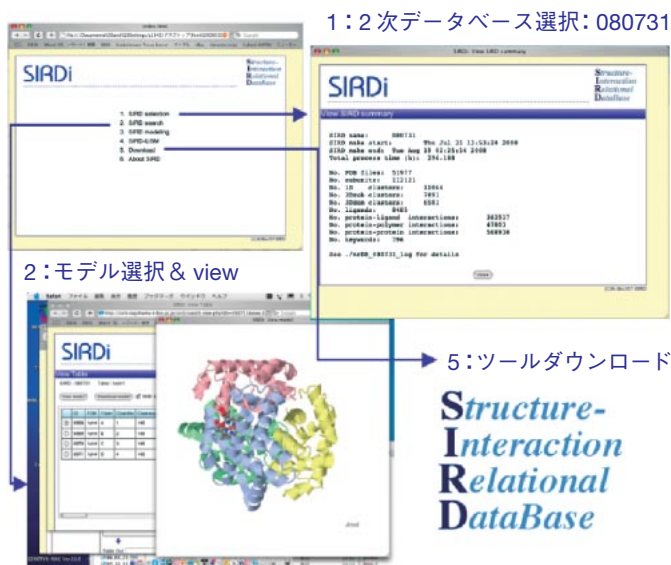


図2 SIRD インターフェース

これにより、ある分子と他の分子が相互作用している構造がPDB中に存在すれば、その座標データを取得することができる(図1の「既知構造」)。これらのモデルは超分子複合体構造モデリングのパーツとなるものである。

SIRDによるプロセスの結果、2008年7月31日現在のPDBには、51977エントリーの構造があり、それらは6581種のタンパク質ドメイン、8485種のリガンド構造を含み、そこから引き出すことのできる相互作用情報は、タンパク質間相互作用563,517件、タンパク質-ポリマー相互作用47,803件、タンパク質-低分子リガンド相互作用363,517件であった(図2 <http://sird.nagahama-i-bio.ac.jp/sird/>でβ版公開)。

2.2 4次構造ツール (Quaternary Structure Tool)

PDBの構造の95%は結晶構造であり、座標は非対称単位(図1のASU)についてのみ提示されている。超分子複合体に対称性が存在し、それが結晶学的対称と一致している場合は、複合体の一部しか見る

ことはできない。これは結晶学の知識のないユーザーにはストレスである。そこでASUから複合体構造を予測するツールBIOMOL (BIological MOLEcule)を開発した。

BIOMOLの特徴はローカル環境で実行することによりユーザーに試行錯誤を許容する点^[2]と、低分子リガンドを含めた全ての分子を扱える点にある。配列重複を排除した4,578個の結晶構造についてBIOMOLの性能試験を行ったところ、全体の89%についてアノテーション通りの複合体を取り出せる閾値が存在した。また、単一の先験的閾値で正解できる複合体は最大で全体の62%であった。

2.3 アセンブルツール群 (Assemble Tools)

アセンブルツールを使ってモデルを組み立てる。図1右はモデリングの手順を簡単に示したものである。この作業は主にCMP (CoMPare) というツールを使用する。これは2つの座標ファイルを与えると、両者の中で重ね合わせ可能なすべての分子構造を自動で探索・重ね合わせを行ってスコアを計算し、上位からユーザーが定義した個数の複合体モデルを出力する。

2.4 モデルはめ込みツール (Fitting Tool)

超分子複合体の構造解析では、電子顕微鏡単粒子解析などから得られる比較的低分解能の密度マップにモデルを当てはめる必要がある。FIT (FITting) はモデル構造をサーチモデルとして電顕密度マップに対して“分子置換法”を適用するツールである。回転サーチと並進サーチによりモデルを密度マップに位置づけることができる。

2.5 3R 複合体超分子構造解析

超分子複合体の構造生物学には、生化学、結晶構造解析、電子顕微鏡解析、バイオインフォマティクスの協力が必要であるが、本研究開発チームはこのすべてを備えている。

我々が対象としたのは、主にPCNA (DNA クランプ) を中心とした古細菌 3R 系超分子複合体である。PCNA は生物界に普遍的に存在する環状のタンパク質分子であり、DNA を中央の穴に捉え(クランプ)、3R (Replication、Repair、Recombination) ファクターのプラットフォームとして中心的な役割を果たす。古細菌 3R 系はヒトなどの高等生物との類似性が高く格好のモデルシステムである。しかしながら 3R 系複合体の機能解析と構造解析は、現時点においてもそれぞれ自体が基礎研究として挑戦的なテーマである。

2.6 モデリングシステムの検証

本研究開発の成果を総合し、モデリングシステムの検証を行った結果を図3に示す。上段の最終モデルは上位15位(rank)までのものである。これらは下段の電顕密度マップとよく似ており、システムの有効性を示している。

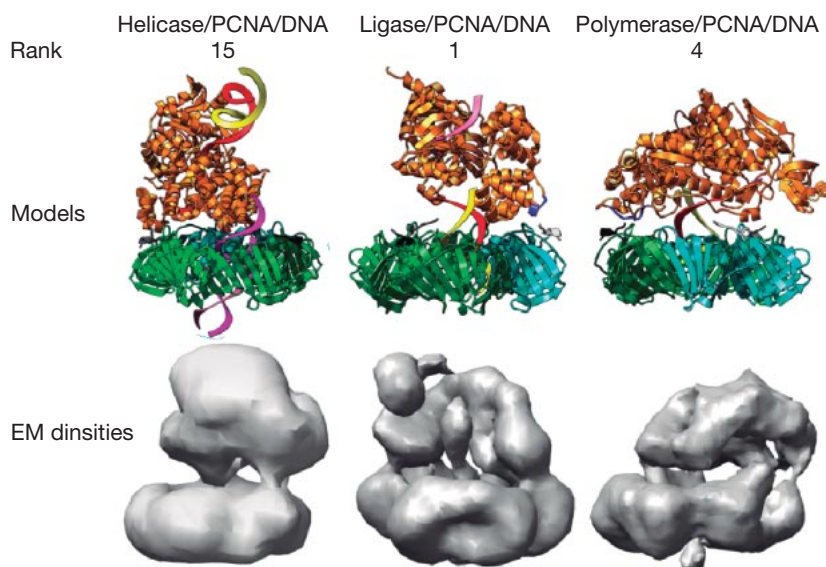


図3 モデル(上)と実験結果(下)の比較

得られたモデルを複合体の機能の推定に用いることも可能である。Helicase/PCNA/DNA 複合体 (図 4 左) は、停止した DNA 複製フォークを一時後退させる役割を持つが、得られたモデルは PCNA にクランプされた 2 重鎖 DNA と helicase に結合した (ほどかれる) 2 重鎖 DNA を的確にモデリングできている。Ligase/PCNA/DNA 複合体 (図 4 右) はニック (DNA の切れ目) をふさぐことで DNA 複製の仕上げをするが、得られたモデルでは PCNA にクランプされた DNA を ligase が取り囲んでいる。この ligase は PDB から抽出された反応終了後の構造であり、厳密には実験で得られた密度マップとドメインの相対配置が異なっていた。このことから、ligase は反応が終了するまでは PCNA の結合部位を押さえることで他のファクターの接近を防ぎ、反応終了に同期して結合部位を解放する「反応の交通整理」機構 (図 4 右) が推定できる。

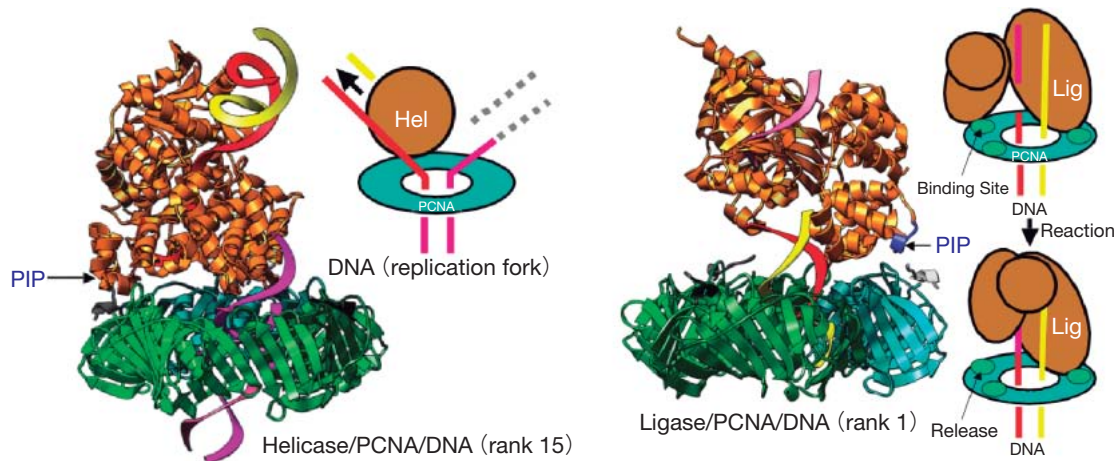


図 4 Helicase/PCNA/DNA 複合体 (左) と ligase/PCNA/DNA 複合体モデル
青で示した PIP (PCNA interacting peptide) はファクターが PCNA に結合する部位である。これらが PCNA に近接していることからモデルの有効性が示される

3. まとめ

本研究開発で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。同時にこれは、データベースに蓄積された情報を整理・活用することの有効性を示している。分子の構造変化など、データベースから抽出できる情報はまだ残されている。今後も本システムをモデリングシステムおよび 2 次データベース生成システムとして発展させることで、超分子モデリングの効率をいっそう向上させられると期待している。

4. 研究開発実施体制

代表研究者 白井 剛 (長浜バイオ大学バイオサイエンス学部)

研究開発題目

(1) 超分子複合体モデリングシステム開発グループ

グループリーダー 白井 剛 (長浜バイオ大学バイオサイエンス学部)

(2) 古細菌 3R 複合体解析グループ

グループリーダー 石野 良純 (九州大学大学院農学研究院)

(3) 超分子複合体構造解析グループ (X 線)

グループリーダー 大山 拓次 (大阪大学蛋白質研究所)

(4) 超分子複合体構造解析グループ (電顕)

グループリーダー 真柳 浩太 (九州大学生体防御医学研究所)

5. 参考文献

- [1] 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).
- [2] 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).
- [3] 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.* *in press* (2008).
- [4] F. Matsunaga, A. Glatigny, M.H. Mucchielli-Giorgi, N. Agier, H. Delacroix, M. Marisa, P. Durosay, Y. Ishino, L. Aggerbeck, and P. Forterre, Genomewide and Biochemical Analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in *Pyrococcus* sp. *Nucleic Acids Res.* **35**, 3214-3222 (2007).
- [5] 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).
- [6] K. Imamura, K. Fukunaga, Y. Kawarabayasi, and Y. Ishino, Specific interactions of three PCNAs with replication-related proteins in *Aeropyrum pernix*. *Mol. Microbiol.* **64**, 308-318 (2007).
- [7] S. Kiyonari, K. Takayama, H. Nishida, and Y. Ishino, Identification of a novel binding motif in *Pyrococcus furiosus* DNA ligase for the functional interaction with proliferating cell nuclear antigen. *J. Biol. Chem.* **281**, 28023-28032 (2006).
- [8] H. Nishida, S. Kiyonari, Y. Ishino, and K. Morikawa, The closed structure of an archaeal DNA ligase from *Pyrococcus furiosus*. *J. Mol. Biol.* **360**, 956-967 (2006).
- [9] Y. Fujiwara, K. Mayanagi, and K. Morikawa, Functional significance of octameric RuvA for a branch migration complex from *Thermus thermophilus*. *BBRC* **366**, 426-431 (2008).
- [10] 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).