



HOKKAIDO UNIVERSITY

Title	The development of model fitting and completion program for automated iterative nucleic acid refinement [an abstract of dissertation and summary of dissertation review]
Author(s)	山下, 恵太郎
Degree Grantor	北海道大学
Degree Name	博士(生命科学)
Dissertation Number	甲第11050号
Issue Date	2013-06-28
Doc URL	https://hdl.handle.net/2115/53235
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Keitaro_Yamashita_abstract.pdf, 「論文内容の要旨」



学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称 博士（生命科学） 氏名 山下 恵太郎

学 位 論 文 題 名

The development of model fitting and completion program for automated iterative
nucleic acid refinement

（核酸構造の精密化を自動的に行うための
フィッティングおよびモデル構築プログラムの開発）

X 線結晶構造解析は、生体高分子の 3 次元構造をオングストローム (10^{-10} m) のオーダーで決定することのできる強力な手法である。これまでに数多くの蛋白質・核酸分子の構造が決定され、生物学的機能の理解に多く貢献してきた。このうち、RNA または DNA の核酸分子は、生命現象における中心的な役割を果たし、情報分子としての役割だけでなく、単体で触媒分子としての機能を持つものや、蛋白質と協同して働く機能分子としての役割を担うものなどがある。近年、核酸分子を含む数多くの結晶構造が決定されてきたが、その多くに共通する問題として、高い分解能が得られにくいということが挙げられる。現在 PDB (Protein Data Bank) に登録されている構造を見てみると、蛋白質単体では平均的に 2.1 Å 程度の分解能を持つのにに対して、核酸単体では半数が 2.5 Å 以下であり、さらに蛋白質・核酸複合体になると半数が 3 Å 以下になる。分解能が低くなるほど、構造解析の難易度は上昇する。電子密度図の明瞭さが悪化するので、その解釈は困難になり、構造を決定するためには X 線結晶学における深い専門知識と、核酸構造についての理解が必要になる。

X 線結晶構造解析のステップは、結晶作製後、X 線回折データ収集、位相決定、モデル構築および精密化からなる。X 線回折実験では、電子密度分布の Fourier 変換である構造因子の絶対値しか観測することができないため、即座に正確な電子密度図を Fourier 逆変換によって得ることができない。位相情報は、重原子を導入する方法や、分子置換法などによって得る必要があるが、初期段階で得られる位相情報は、大きな誤差を持つ。分子モデルは電子密度に従って構築されるが、その電子密度の計算に必要な位相情報が誤差を持つため、正確なモデルを一度に得ることができない。分子モデルの構築がある程度まで進んだとき、精密化と呼ばれる計算が行われる。精密化によって構築された分子モデルは実験データとよく一致するように修正され、その結果得られるモデル由来の位相情報によって、より正確な電子密度図を描くことが可能になる。しかしながら、モデルの持つ多くの誤差は精密化計算によって除くことができないので、更新された電子密度に対して手動によるモデル修正および欠落部分の構築を行う必要がある。すなわち、分子モデルの構築は、精密化というプロセスの中で反復的に行われ、実験データとの一致が十分良くなるまで繰り返される必要がある。

本研究では、核酸構造のフィッティングおよび伸長を行うプログラム *NAFIT* および

NABUILD を開発した。これらのプログラムは、独立なプログラムとしてだけでなく、自動精密化システム *LAFIRE* のモジュールとしても利用できる。*LAFIRE* は、人の手による構造精密化プロセスを自動的に行うプログラムで、初期モデルをもとに、モデルの修正・伸長を、精密化プログラムと組み合わせて反復的に行うよう設計されている。

核酸は主鎖に多くの回転可能な結合を持つが、実際にはこれらの取りうる範囲は制限されている場合が多い。しかしながら中～低分解能の電子密度では、その配座を正確に決定することは困難であり、立体化学的に不安定な構造を作ってしまうがちな。 *NAFIT* は、従来の基本的な実空間精密化法に対して経験的なねじれ角制約関数と、水素原子を含めた非結合原子間の反発関数を加え、空間的に近接する残基同士を同時にフィッティングするグループ化フィッティングを、逐次的に行うよう設計した。経験的な制約関数は、PDB 登録された核酸構造から、比較的信頼できるものを抽出し、頻度の負の対数として設計した。分解能 4.2 Å という低分解能のサンプルデータに対して *NAFIT* を実行した結果、電子密度とモデルの一致度およびモデルの幾何学的な質を著しく向上させることができた。

NABUILD は、はじめに電子密度図を細線化して主鎖の経路を見つけ、粗視化モデルを構築した後、*NAFIT* によって原子モデルに変換する戦略を取った。主鎖をたどる際、電子密度の断絶によってトレースが止まってしまうのを防ぐため、細線化の際、断裂した電子密度図を連結させるようにし、成功率を向上させた。これによって偽の主鎖候補も増加するが、核酸の構造的特徴および電子密度の評価を用いたスコアリングを行うことによって、偽の候補の増加を抑えた。先にも述べたように、モデル構築は精密化とともに反復的に行うことによって効果的に進行することが期待できる。*NABUILD* を精密化と組み合わせるときには、2 サイクル目以降では電子密度図との一致が悪い部分を削除した上で構築を行うようにした。PDB に登録された分解能 2.1 から 3.1 Å のデータを用いてテストを行った結果、*NABUILD* は反復的精密化と組み合わせることによって高い成功率でモデルを伸長させ、結晶構造の信頼度因子である R_{free} 因子を有意に向上させることが示された。

以上のように、核酸のモデル修正・伸長を実験者の介入無しで行うアルゴリズムを考案し、その有用性を示すことができた。