

タンパク質立体構造モデリングにおける 修飾アミノ酸の扱いについて

工藤麻由^{†1} 佐藤亘^{†1} 梅山秀明^{†2} 岩館満雄^{†2}

この研究はタンパク質立体構造データベース中において近年増加しつつあり、無視できない割合を占める「修飾アミノ酸」のホモロジーモデリングについて取り扱いについての方法論である。従来、参照してこなかった修飾アミノ酸の座標情報のうち、93%を20種類の天然アミノ酸としてその原子座標情報を利用した自動的ホモロジーモデリング法とその有用性について述べる。

Protein Structure Homology Modeling Method Considering Modified Amino Acid

MAYU KUDO^{†1} WATARU SATO^{†1}
HIDEAKI UMEYAMA^{†2} MITSUO IWADATE^{†2}

We have developed a new FAMS program including the respond guideline for non-natural amino acids using the former FAMS program. Thus, we created the automatic homology modeling soft program (NN FAMS) devised newly in inclusion of many non-natural amino acids in the PDB data.

1. はじめに

近年、タンパク質の結晶化を容易にする目的等で、人工的に修飾された非天然型アミノ酸を含む多くのタンパク質がPDB (Protein Data Bank) に登録されている。しかしながら、我々の開発してきたホモロジーモデリングシステムFAMSでは、これら非天然型アミノ酸を利用することについて、合理的な対応ができていない。

本研究では、2012年6月時点での全PDBに対して、総当たりで作成したアライメントの内、非天然型アミノ酸を含むタンパク質を鋳型として、非天然型アミノ酸に対応・非対応それぞれの場合でホモロジーモデリングを行い、その精度を比較した。

プログラムの中で、今回対応させた非天然型アミノ酸は36種、ABA, ALY, BB9, CAS, CME, CMH, CSO, CSS, CSW, CSX, CXM, FME, HIC, HYP, KCX, M3L, MEN, MHO, MLY, MLZ, MME, MP8, MSE, NLE, OCS, ORN, PTR, SAC, SEP, SLZ, TPO, TRN, TRQ, TYS, YCM, YOF となっている。

図1は全PDBに含まれる非天然型アミノ酸の内訳で、合計111103残基が非天然型アミノ酸として存在し、種類とし

非天然型アミノ酸の内訳

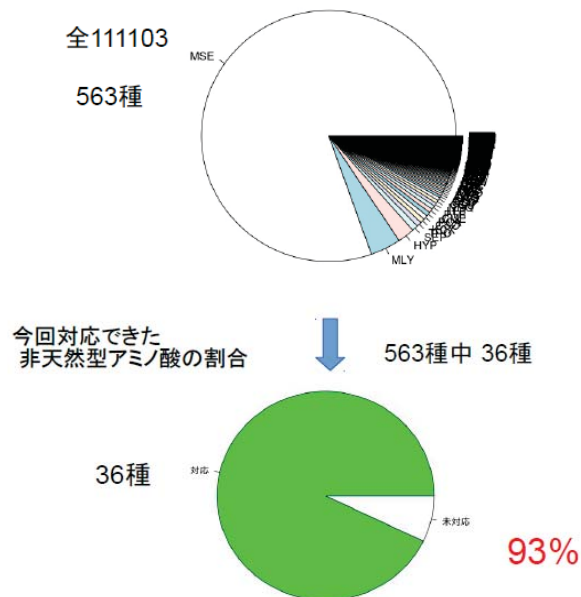


図1. 非天然型アミノ酸の内訳

ては563種、その割合は上の円グラフが表している。そしてその下にある円グラフが、今回36種の非天然型アミノ酸に対応したことにより93%の非天然型アミノ酸残基をカバーできるようになったことを示している。

†1 中央大学大学院理工学研究科生命科学専攻
Department of Biological Sciences,
Graduated School of Science and Engineering, Chuo University

†2 中央大学理工学部生命科学科
Department of Biological Sciences,
Faculty of Science and Engineering, Chuo University

2. 方法

2012年6月時点で、PDBに存在しているタンパク質の内、非天然型アミノ酸を含むタンパク質のアミノ酸配列を集め、それらをクエリー (query) として、PDBの95%冗長性排除したデータベースに対してホモロジー検索を実行し、アライメントを得た。

作成されたアライメントのクエリーとテンプレート (template) を入れ替え、ホモロジーモデリングソフトウェア FAMS を使い立体構造を構築した。

FAMSにおいて非天然型アミノ酸を読み込む際は以下のルールに従っている。

- ・複数の天然型アミノ酸が結合して一つの非天然型アミノ酸を形成している場合は欠損として扱う。
- ・PDBのMODRES行に記載されている天然型・非天然型の置換データに基づく。
- ・複数の置換データがあった場合は、非天然型アミノ酸に似ている天然型アミノ酸に置換。

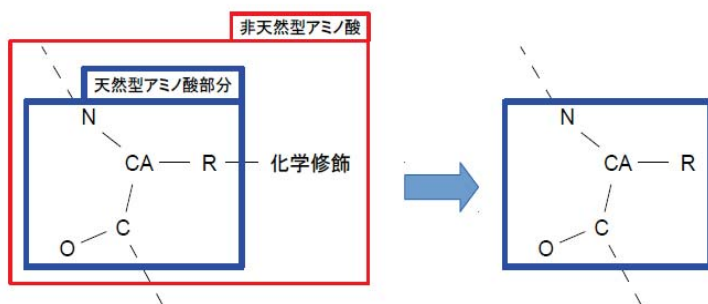


図2. 非天然型アミノ酸の置換方法

図2は非天然型アミノ酸を天然型アミノ酸に置換する方法を示している。非天然型アミノ酸は天然型の側鎖に化学修飾が施されたものが大半を占め、そのようなものは天然型の部分だけを参照できるように置換する。そうでない場合は、天然型に類似する部分だけを置換するように試みた。

上記のように構築したモデルを正解構造と GDT_TS と RMSD を用いて比較した。

3. 結果

まず、GDT_TS を用いた図3の比較のグラフを見ると、126,892個のモデルを対象として非天然型アミノ酸対応後と対応前の平均の差が0.132となり、モデル全体での構造が僅かに良くなっていることがわかる。

つまり、ホモロジーモデリング過程において、非天然型アミノ酸の座標情報を置換してでも参照したほうが、参照しない場合よりも、立体構造モデルの精度が高いというこ

とがわかった。

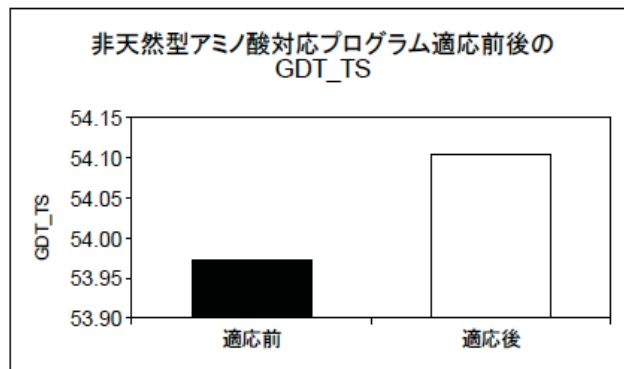


図3. 非天然型アミノ酸対応プログラム適応前後の GDT_TS (平均)

ホモロジーモデリングソフトウェアとして、我々が使用している FAMS に対して、アライメントの結果を使って参照タンパク質の立体構造を使うプロセスに、この論文の結果を反映するように改良を加えた。

参考文献

- 1) Ogata, K. and Umeyama, H. (2000). An automatic Homology modeling method consisting of database searches and simulated annealing. *J Mol Graph Model.*, 18(3):258-72, 305-6.
- 2) Umeyama, H. and Iwadate, M. (2004). FAMS and FAMSBASE for protein structure. *Current Protocols in Bioinformatics.*, Chapter 5[Unit 5.2].