

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 15 日現在

機関番号：32607

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26730151

研究課題名(和文)アミノ酸残基間ネットワーク比較によるタンパク質構造比較解析

研究課題名(英文)A Flexible Protein Structure Alignment Method Based on the Residue-Residue network comparison

研究代表者

寺師 玄記(Terashi, Genki)

北里大学・薬学部・講師

研究者番号：40383658

交付決定額(研究期間全体)：(直接経費) 2,000,000円

研究成果の概要(和文)：本研究は三次元のタンパク質構造情報をアミノ酸残基間におけるネットワークと捉え、タンパク質間の構造類似性の検出方法を新たに提唱するものである。タンパク質構造は本質的に柔軟な構造を持つが、タンパク質の構造が変化してもアミノ酸残基間のネットワークの多くが保持されるという性質に着目した研究である。この研究により三次元座標の比較では検出できない構造類似性を新たに検出できるようになった。本研究の成果であるコンピュータープログラムはhttp://www.pharm.kitasato-u.ac.jp/bmd/bmd/Download_files/CABalign/index.html で、公開されている。

研究成果の概要(英文)：In this project, we developed a new protein structure alignment method, contact area-based alignment (CAB-align), which uses the residue-residue contact area to identify regions of similarity. Our novel algorithm can identify homologous relationships at the residue level between related protein structures. The software is freely available to academic users as stand-alone software at <http://www.pharm.kitasato-u.ac.jp/bmd/bmd/Publications.html>.

研究分野：バイオインフォマティクス

キーワード：タンパク質立体構造 立体構造アライメント 構造比較 タンパク質三次元構造

1. 研究開始当初の背景

X線結晶解析や高解像度 cryo-EM 法などの実験手法が飛躍的に進歩した結果、タンパク質の立体構造情報は近年飛躍的に増大している。Protein Data Bank (PDB) に登録されているデータ数は、2014 年から 2017 年にかけて 10.5 万から 13 万に増加した。これは 3 年間で約 20% 以上の増加が見られたことを意味する。その結果、研究者があるタンパク質の類縁タンパク質を検索する際、従来のアミノ酸配列の比較・アライメントによって検索する手法だけでなくタンパク質の三次元構造情報に基づく検索も利用できるデータ数の観点から可能となった。

タンパク質の立体構造情報は、類縁タンパク質においてアミノ酸配列よりも強く保存される事が明らかとなっている。したがって、立体構造情報を使用した検索方法は従来では検出できなかった新たなタンパク質の類縁関係を見出すことができると期待されている。類縁タンパク質を検索することは、未知のタンパク質の機能推定や、タンパク質立体構造予測、タンパク質に作用する低分子化合物の開発において極めて重要な手法である。

一方で、タンパク質分子そのものは生体内で構造を柔軟に変化させ、この立体構造の柔軟性はタンパク質が機能を発揮する上で重要である事も明らかとなっている。このことから、2つのタンパク質分子を比較した際、三次元座標上では構造が異なる場合でもその機能が極めて近い類縁タンパク質であることがある。例えばカルシウム結合タンパク質カルモジュリンは球状の形のものもあれば、ダンベル型の構造を持つものも存在する。

現在広く用いられているタンパク質立体構造の比較法は、三次元空間上の類似性に基づく手法が殆どであり、タンパク質の柔軟性を考慮することができない。本研究はこの課題に取り組むものである。

2. 研究の目的

本研究は三次元のタンパク質構造情報をアミノ酸残基間におけるネットワークと捉え、タンパク質間の構造類似性の検出方法を新たに提唱するものである。タンパク質構造は本質的に柔軟な構造を持つが、タンパク質の構造が変化してもアミノ酸残基間のネットワークの多くが保持されるという性質に着目した研究である。この研究により三次元座標の比較では検出できない構造類似性を新たに検出できるようになると期待される。

本研究手法は、構造の類似性を検出するだけでなく、タンパク質分子のどの部分が類似しているかをアミノ酸残基単位で同定することの可能である。すなわち、2分子間の構造アライメントを行う。

本研究で開発された手法を一般のパソコンで使用できるプログラムに実装し公開することで、多くの研究者に自由に利用しても

らいタンパク質構造の研究推進に寄与することを目的とする。

3. 研究の方法

本研究は主に以下の3つの項目で遂行された。

(1) 理論の実装・プログラミング

本研究の提案手法をコンピュータプログラムとして実装した。また研究期間中に計算アルゴリズムの改良を行い、実行時間の大幅に短縮を達成した。プログラム中で使用されるパラメーターをタンパク質ドメイン構造から作成された25万データによって最適化行なった。これにより本研究論文で報告されているように、高い精度で類縁タンパク質を検索することが可能となった。

(2) 比較検証

大規模検証データを使い、従来法との比較を行い、その精度検証を実施した。比較検証のために用いた従来法は Dali, TMalign, FATCAT, HAlign の4つである。精度検証は、タンパク質ドメインデータベース SCOPe を使用した。アミノ酸配列の類似性、機能の類似性に基づきタンパク質構造を階層的に分類したデータベースである。

(3) 研究成果発表

欧文論文雑誌への投稿を行なった²。また研究成果を学会などで発表した。詳細は次項目「4. 研究成果」にて報告する。

引用文献:

- 1 <http://www.rcsb.org/pdb/home/home.do>
- 2 Terashi, Genki, and Mayuko Takeda-Shitaka. "CAB-Align: A Flexible Protein Structure Alignment Method Based on the Residue-Residue Contact Area." *PLoS one* 10.10 (2015): e0141440.

4. 研究成果

本研究の成果は以下の4つである。

(1) 新たな構造アライメント手法の提案

本研究により、従来の構造比較法では対応関係を示すことが困難であったタンパク質構造の類似性を検出し視覚的に示すことが可能となった。本研究で開発された手法の適用例を以下に示す。

下図1, 2は、どちらもカルシウム結合タンパク質 (PDBID 1ncx, 2sas) である。三次元空間上、この2つのタンパク質分子の類縁性は一部分に限られる。そのため従来手法では正しく構造アライメントを計算することができない。しかし、本研究の手法を適用すると、図1, 2の点線で示された部分で、共通したアミノ酸残基間のネットワークが存在していることが示される。このことにより、本研究手法ではこの2つのタンパク質分子はほぼ全域に渡って構造上類似している判

断される。この結果は、配列アライメント法の結果と一致している。

この例からもわかるように、本研究手法により三次元空間配置に依存しない新たなタンパク質構造アライメント法が提案された。

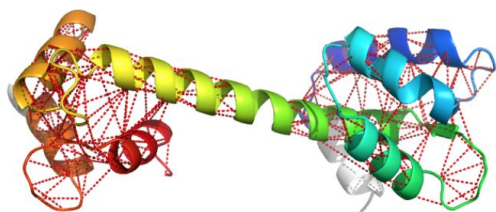


図1 PDBID:1ncx

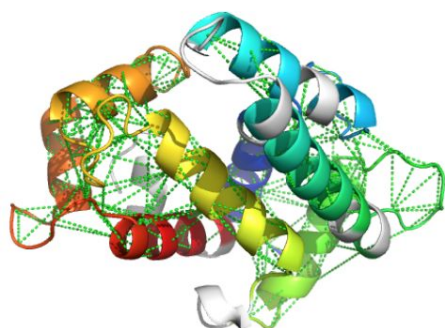


図2 PDBID:2sas

下図3は、ドメインデータベースで分類されている Superfamily (構造および機能上タンパク質が類縁関係にあるもの) を正しく判別できるかについての ROC 曲線である。

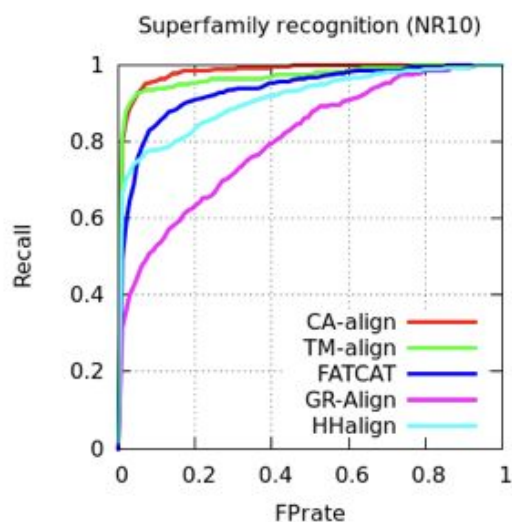


図3 25万データにおける Superfamily 分類

本研究手法を表す CA-align は、曲線下の面積(AUC)が最も大きいことから、最も高い精

度を示している。このことから、本研究手法は、従来法よりも正しくタンパク質間の類縁関係を検出可能であることが示された。詳細なデータ・計算手法・アルゴリズムは、発表論文にて報告されている。

(2) 論文発表

オープンアクセスの欧文雑誌 PlosOne で本研究の内容を発表した。

Terashi, Genki, and Mayuko Takeda-Shitaka. "CAB-Align: A Flexible Protein Structure Alignment Method Based on the Residue-Residue Contact Area." *PloS one* 10.10 (2015): e0141440.

(3) 学会などでの研究成果発表

日本国内では JSBi2014 (仙台) でポスタ発表、第42回構造活性相関シンポジウム(熊本)で口頭発表を行なった。特に第42回構造活性相関シンポジウムの口頭発表は SAR Presentation Award を受賞した。

米国インディアナ州 Purdue 大学での国際学会 GLBI02015、およびフロリダ州オーランドでの国際学会 ISMB-3DSIG においてポスター発表を行った。

発表した論文は Indiana Japan Chamber より 2015 Outstanding Research Paper Award を受賞し、受賞講演を行った。

本研究の成果であるコンピュータープログラムは

http://www.pharm.kitasato-u.ac.jp/bmd/bmd/Download_files/CABalign/index.html

で、公開され自由に使用することができる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計1件)

Terashi, Genki, and Mayuko Takeda-Shitaka. "CAB-Align: A Flexible Protein Structure Alignment Method Based on the Residue-Residue Contact Area." *PloS one* 10.10 (2015): e0141440. (査読有り)

[学会発表](計4件)

Genki Terashi, Daisuke Kihara, Mayuko Takeda-Shitaka, CAB-align: Novel protein structure alignment method based on the residue-residue contact area. 3DSIG, Orlando, Florida. US. July 8-9/2016

Genki Terashi, and Mayuko Takeda-Shitaka, Flexible protein structure alignment method based on the residue-residue interaction. GLBI02015, Purdue University US. May 18, 2015

寺師 玄記. 残基間相互作用に基づく新たなタンパク質構造アライメント法の開発
第 42 回構造活性相関シンポジウム(熊本),
2014.11.13-14

寺師 玄記, Novel protein structure alignment method based on the inter-residue contact area, JSBi2014・第 3 回生命医薬情報学連合大会(仙台).
2014-10-02 - 2014-10-04

〔その他〕

(1) ホームページ

http://www.pharm.kitasato-u.ac.jp/bmd/bmd/Download_files/CABalign/index.html

(2) 本研究成果による受賞

2015 年 Outstanding Research Paper Award (Indiana Japan Chamber) 本研究の研究成果論文に対する受賞

2015 年 SAR Presentation Award 本研究についての口頭発表に対する受賞

6 . 研究組織

(1)研究代表者

寺師 玄記 (TERASHI, GENKI)

北里大学薬学部 講師

研究者番号 : 4 0 3 8 3 6 5 8