

平成 30 年 6 月 13 日現在

機関番号：12608

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K16081

研究課題名(和文) 中規模ペプチドの構造的相互作用解析基盤技術の開発

研究課題名(英文) Development of a computational peptide-protein interaction prediction method based on their tertiary structures

研究代表者

大上 雅史 (Ohue, Masahito)

東京工業大学・情報理工学院・助教

研究者番号：50743209

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：中分子創薬において重要とされるペプチドの標的を効率良く探索し、かつ標的との結合性を予測できる手法の開発が目的である。立体構造情報に基づくタンパク質間相互作用予測の大規模利用のための高度化と並列実行を実施した。また、代表的なヒトタンパク質構造に対して総当たり予測計算を実施し、結果をウェブブラウザを介して閲覧できるようにインターフェースを構築した。さらに、ペプチド分子とタンパク質との結合性解析計算を行う手法、2部グラフ予測を主体とした機械学習による分子標的予測手法であるLIK法、およびランク学習を主体とした機械学習による分子標的予測手法であるPKRank法を開発した。

研究成果の概要(英文)：The purpose of this study is to efficiently search for peptide targets, which is considered important in middle molecule drug discovery, and to develop a prediction method reveals binding target. I have performed large-scale parallelization to predict protein-protein interactions based on three-dimensional structure information. In addition, an all-to-all prediction was performed on representative human protein structures, and a web-interface was constructed to find the prediction results via a web browser. In addition, I have developed computational methods that conducting binding analysis of peptide molecules and proteins, LIK method which is a molecular target prediction method based on bipartite graph prediction by machine learning, and another molecular target prediction method named PKRank based on learning-to-rank.

研究分野：バイオインフォマティクス

キーワード：ペプチド創薬 タンパク質間相互作用 バイオインフォマティクス 分子ドッキング 標的予測 カーネル法 ランク学習

1. 研究開始当初の背景

低分子創薬からの脱却を目指して、ペプチド分子が創薬研究で注目されている。機能未知のペプチドが細胞内で数多く翻訳されていることも分かっており、ペプチドの解析技術の開発が急務である。特に立体構造の観点から、ペプチドが結合するタンパク質等の分子との結合性に関して計算機で解析可能な技術開発が必要とされている。本研究は、特に 10 残基以上の中規模ペプチドの構造解析を行うための、バイオインフォマティクス技術を開発する。

2. 研究の目的

本研究では、特に中分子創薬において重要とされるペプチドの標的について、効率良く標的を探索し、かつ標的との結合性を予測できる手法の開発を目的とした。当初はペプチドの機能部位（結合部位やシグナル配列など）の推定を行った上で、標的との結合性についての予測技術を開発する計画であったが、本研究実施期間中に直鎖状ペプチド以外のペプチドや天然化合物等への応用可能性について検討を加え、標的探索のための技術開発と、標的との結合性の予測技術開発の 2 項目を柱として研究を実施した。

3. 研究の方法

研究代表者らが開発してきた立体構造情報に基づくタンパク質間相互作用予測ソフトウェア MEGADOCK について、大規模利用のための高度化と並列実行を実施した。また、代表的なヒトタンパク質を Protein Data Bank から収集し、それらの総当たりタンパク質間相互作用予測計算を実施し、結果をアーカイブ化した。予測結果はウェブブラウザを介して閲覧できるように分子ビューワ等を用いてインターフェースを構築した。

また、上記の MEGADOCK ソフトウェアを応用し、ペプチド分子とタンパク質との結合性解析計算を行う手法を開発した。本手法はペプチドの配座生成、タンパク質とのドッキング計算、候補構造群からの統計解析の 3 段階を経て予測を行うものである (図 1)。

さらに直鎖状ペプチド以外の分子の標的探索技術として、機械学習による分子標的予測手法の開発を実施した。カーネル法の枠組みを利用し、2 部グラフ予測を主体とした LIK 法、およびランク学習を主体とした PKRank 法を開発した。

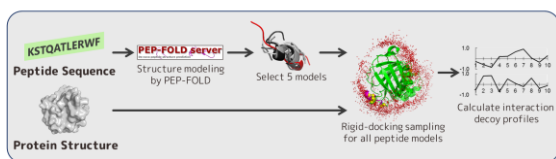


図 1 ペプチドの結合性解析計算手法の開発における研究方法の概要図。

4. 研究成果

4. 1 MEGADOCK の高度化と並列実行

MEGADOCK について、大規模利用のための高度化と並列実行を実施した。特にパブリッククラウド計算環境を高効率に活用するための実装を実施した。Microsoft Azure の GPU インスタンスを用いた多数の Virtual Machine を利用した並列分散環境において、高い並列化効率でのタンパク質間相互作用予測計算を実現した (図 2)。

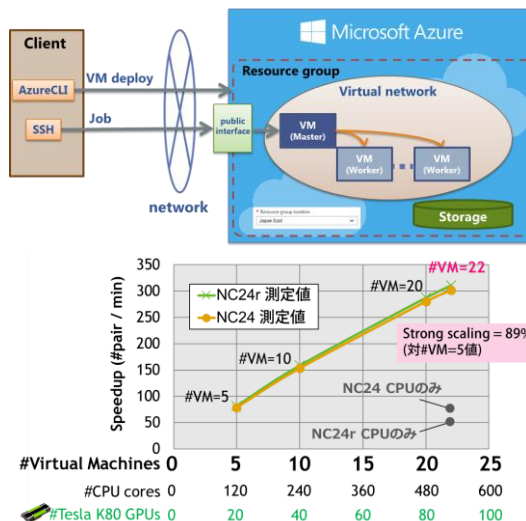


図 2 Microsoft Azure 上の並列計算環境の実装 (上) と並列計算性能評価 (下)。

4. 2 MEGADOCK-Web の開発

ペプチドが標的とするタンパク質間相互作用を列挙するためのツール整備として、MEGADOCK によって予測されたタンパク質間相互作用情報を蓄積したデータベース MEGADOCK-Web を構築した (図 3)。MEGADOCK-Web はおよそ 3000 万件のタンパク質間相互作用予測の結果が収録されており、Molmil 分子ビューワによるインタラクティブな構造可視化、KEGG パスウェイとの関連付けなどによって標的候補分子を効率よく検索することが可能である。MEGADOCK-Web は以下の URL にフリーで公開しており、ウェブブラウザからアクセスできる。

<http://www.bi.cs.titech.ac.jp/megadock-web>

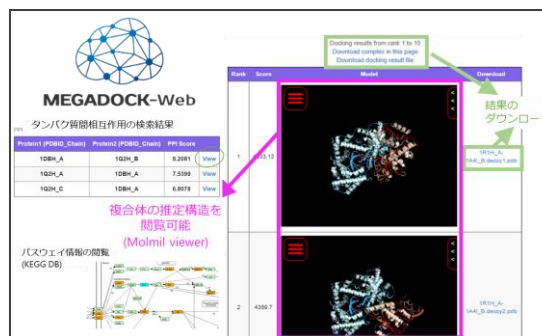


図 3 MEGADOCK-Web の各種機能。

4. 3 ペプチドの結合性解析計算

ペプチド分子とタンパク質の結合に係る物理化学的性質の予測に焦点を当てた。既存のモデリングツールでは 15 残基程度のアミノ酸配列情報からペプチドの構造 (配座) を正しく再現することは困難であることを確認しているが、一方でそれらのペプチドモデル構造を複数使い、対象のターゲットとなるタンパク質に対して MEGADOCK ソフトウェアによる相互作用計算を実施したところ、正解配座との距離基準を満たす部位へ相互作用するモデルが効率よく得られることがわかった (図 4)。そのようなモデルを抽出するための方法として相互作用プロファイルベクトルによる解析手法を提案し、本研究期間中に複合体 10 例における検証を完了した (図 5)。これらの結果から、一定の構造集団 (構造サンプリング) を用いることで、完全に正しい構造モデルが得られなくとも、相互作用親和性と相関するパラメータが得られる可能性があることを示した。

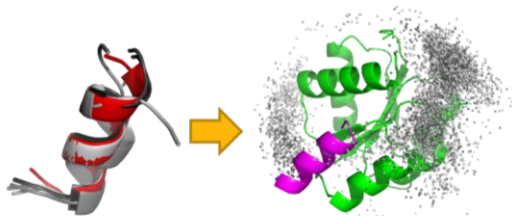


図 4 ペプチドドッキング計算の結果。灰色の点は結合部位候補の重心位置を示している。このケースではペプチドの配座は真の結合配座 (赤色) とほぼ同じ構造が予測された。

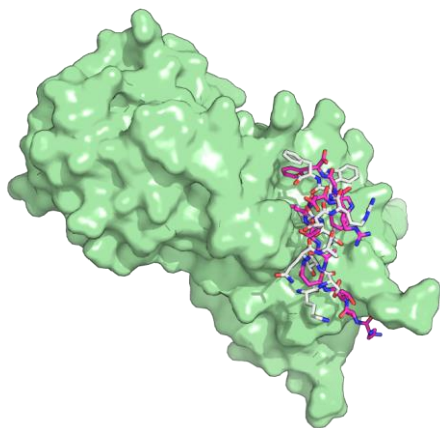


図 5 ペプチドの結合構造予測の一例。PDB ID は 1RXZ である。ペプチド部分における結晶構造との差異 (RMSD) は 5.4 Å であった。

4. 4 機械学習による分子標的予測

実験情報が豊富な低分子のアッセイデータを用いて、機械学習によって分子標的を予測する手法を開発した。本研究ではカーネル法の枠組みに着目し、アッセイデータが多数の類似化合物と標的分子において得られている状況下での 2 部グラフ上のリンク予測問題を解く LIK 法、および標的分子と類似の

分子群のアッセイデータが活用できる状況下でのランク学習問題を解く PKRank 法を新規に提案した。

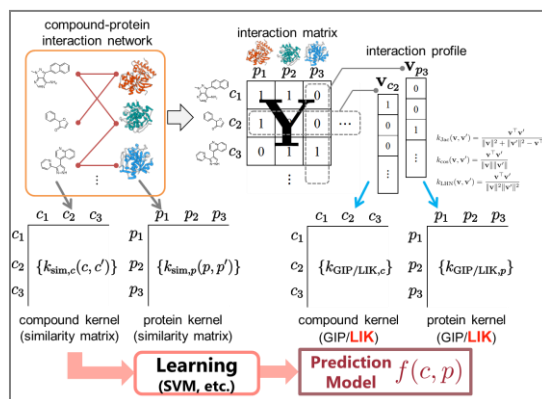


図 6 新規カーネル関数に基づく分子標的予測法 LIK の概要図。

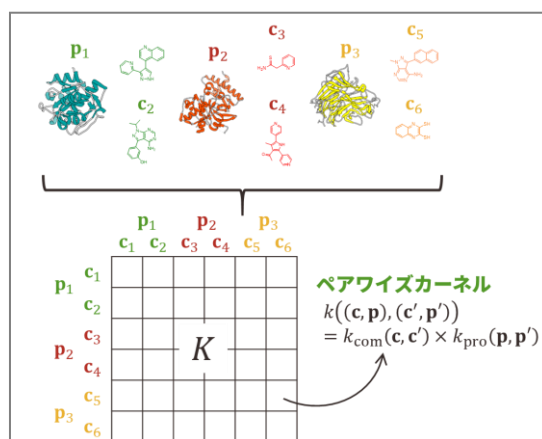


図 7 ランク学習と新規カーネル法に基づく分子標的予測法 PKRank の概要図。

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 11 件)

- ① Masahiro Mochizuki, Shogo D. Suzuki, Keisuke Yanagisawa, Masahito Ohue, Yutaka Akiyama. **QEX: Target-specific druglikeness filter enhances ligand-based virtual screening**, *Molecular Diversity*, in press. 査読有
- ② Keisuke Yanagisawa, Shunta Komine, Rikuto Kubota, Masahito Ohue, Yutaka Akiyama. **Optimization of memory use of fragment extension-based protein-ligand docking with an original fast minimum cost flow algorithm**, *Computational Biology and Chemistry*, in press. 査読有
doi:
10.1016/j.compbiolchem.2018.03.013
- ③ Takanori Hayashi, Yuri Matsuzaki, Keisuke Yanagisawa, Masahito Ohue, Yutaka Akiyama. **MEGADOCK-Web: an**

- integrated database of high-throughput structure-based protein-protein interaction predictions, *BMC Bioinformatics*, 19(Suppl 4): 62, 2018. 査読有 doi: 10.1186/s12859-018-2073-x
- ④ Tomohiro Ban, Masahito Ohue, Yutaka Akiyama. **Multiple grid arrangement improves ligand docking with unknown binding sites: Application to the inverse docking problem**, *Computational Biology and Chemistry*, 73: 139-146, 2018. 査読有 doi: 10.1016/j.compbiolchem.2018.02.008
- ⑤ Shogo D. Suzuki, Masahito Ohue, Yutaka Akiyama. **PKRank: A novel learning-to-rank method for ligand-based virtual screening using pairwise kernel and RankSVM**, *Artificial Life and Robotics*, 23(2): 205-212, 2018. 査読有 doi: 10.1007/s10015-017-0416-8
- ⑥ Masahito Ohue, Takuro Yamazaki, Tomohiro Ban, Yutaka Akiyama. **Link Mining for Kernel-based Compound-Protein Interaction Predictions Using a Chemogenomics Approach**, In *Proceedings of the Thirteenth International Conference On Intelligent Computing (ICIC2017) (Lecture Notes in Computer Science)*, 10362, 549-558, 2017. 査読有 doi: 10.1007/978-3-319-63312-1_48
- ⑦ Naoki Wakui, Ryunosuke Yoshino, Nobuaki Yasuo, Masahito Ohue, Masakazu Sekijima. **Exploring the selectivity of inhibitor complexes with Bcl-2 and Bcl-XL: a molecular dynamics simulation approach**, *Journal of Molecular Graphics and Modelling*, 79: 166-174, 2018. 査読有 doi: 10.1016/j.jmgm.2017.11.011
- ⑧ Tomohiro Ban, Masahito Ohue, Yutaka Akiyama. **Efficient Hyperparameter Optimization by Using Bayesian Optimization for Drug-Target Interaction Prediction**, In *Proceedings of the 7th IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICABS 2017)*, 8 pages, 2017. 査読有 doi: 10.1109/ICABS.2017.8114299,
- ⑨ Keisuke Yanagisawa, Shunta Komine, Shogo D. Suzuki, Masahito Ohue, Takashi Ishida, Yutaka Akiyama. **Spresso: An ultrafast compound pre-screening method based on compound decomposition**, *Bioinformatics*, 33(23): 3836-3843, 2017. 査読有 doi: 10.1093/bioinformatics/btx178
- ⑩ Yuri Matsuzaki, Nobuyuki Uchikoga, Masahito Ohue, Yutaka Akiyama. **Rigid-docking approaches to explore protein-protein interaction space**, *Advances in Biochemical Engineering/Biotechnology (Network Biology)*, 160: 33-55, 2017. 査読有 doi: 10.1007/10_2016_41
- ⑪ Nobuyuki Uchikoga, Yuri Matsuzaki, Masahito Ohue, Yutaka Akiyama. **Specificity of broad protein interaction surfaces for proteins with multiple binding partners**, *Biophysics and Physicobiology*, 13: 105-115, 2016. 査読有 doi: 10.2142/biophysico.13.0_105
- [学会発表] (計32件)
- ① Masahito Ohue. "MEGADOCK: a supercomputing bioinformatics application for protein-protein interaction prediction", 1st RWBC-OIL Workshop, Tokyo Institute of Technology, Tokyo Japan, May 8, 2018.
- ② 大上雅史, "出会い系タンパク質を探す旅", 情報処理学会第80回全国大会, IPSJ-ONE (4), 早稲田大学西早稲田キャンパス, 3/15, 2018.
- ③ 山本悠生, 大上雅史, 秋山泰:"クラウド上の分散GPU環境におけるタンパク質間相互作用予測計算フレームワークの開発", 情報処理学会研究報告, 2018-BI0-53(8), 1-8, 3/10 2018.
- ④ 大上雅史, "MEGADOCK-Web: タンパク質間相互作用予測の計算結果を収録したデータベースとWebインターフェース", 第22回オープンバイオ研究会, 北陸先端科学技術大学院大学, 3/9 2018.
- ⑤ Masahito Ohue, "Prediction of Protein-Protein Interactions with MEGADOCK: Parallelization, Application, and Open Database", 4th IIT Madras-Tokyo Tech Joint Symposium on Frontiers in Bioinformatics: Large Scale Data Analysis, Resources and Drug Design, Indian Institute of Technology Madras, Chennai India, Nov 10-11, 2017.
- ⑥ Masahito Ohue, Yuki Yamamoto, Kento Aoyama, Yutaka Akiyama. "MEGADOCK-Azure: high-performance protein-protein docking on Microsoft Azure HPC", 4th IIT Madras-Tokyo Tech Joint Symposium on Frontiers in Bioinformatics: Large Scale Data Analysis, Resources and Drug Design, Indian Institute of Technology Madras, Chennai India, PP35, Nov 10-11, 2017.

- ⑦ Masahito Ohue, Takanori Hayashi, Keisuke Yanagisawa, Yuri Matsuzaki, Yutaka Akiyama: "MEGADOCK-Web: an integrated database of high-throughput structure-based protein-protein interaction predictions", Informatics in Biology, Medicine and Pharmacology 2017 (IIBMP2017), P57, 2017.
- ⑧ 柳澤溪甫, 小峰駿汰, 久保田陸人, 大上雅史, 秋山泰: "フラグメント伸長型化合物ドッキング計算のための重み付きオフラインキャッシュ問題の厳密解アルゴリズム", 情報処理学会研究報告, 2017-BIO-50(39), 1-8, 6/25 2017.
- ⑨ 林孝紀, 山本悠生, 松崎由理, 大上雅史, 秋山泰: "タンパク質間相互作用予測統合データベース MEGADOCK-WEB の改良とクラウド計算環境との連携", 情報処理学会研究報告, 2017-BIO-50(40), 1-8, 6/25 2017.
- ⑩ 大上雅史, 山本悠生, 秋山泰: "Microsoft Azure 上でのタンパク質間相互作用予測システムの並列計算と性能評価", 情報処理学会研究報告, 2017-BIO-49(4), 1-3, 3/23 2017.
- ⑪ 青山建人, 山本悠生, 大上雅史, 秋山泰: "コンテナ型仮想化による分散計算環境におけるタンパク質間相互作用予測システムの性能評価", 情報処理学会研究報告, 2017-BIO-49(3), 1-8, 3/23 2017.
- ⑫ Masahito Ohue, Yuki Yamamoto, Takanori Hayashi, Yuri Matsuzaki, Yutaka Akiyama. "Cloud Computing for All-to-All Protein-Protein Docking on Azure HPC", Biophysical Society 61st Annual Meeting, 2218-Pos, Ernest N. Morial Convention Center, New Orleans, LA, USA, 2/14 2017.
- ⑬ 長澤一輝, 松崎由理, 大上雅史, 秋山泰: "タンパク質間相互作用予測結果データベースと表示系の構築", 情報処理学会研究報告, 2016-BIO-45(3): 1-4, Mar 2016.
- ⑭ 山崎卓朗, 大上雅史, 秋山泰: "タンパク質-化合物相互作用ネットワークのリンクマイニング", 情報処理学会研究報告, 2016-BIO-45(12): 1-7, Mar 2016.
- ⑮ Masahito Ohue, Yuri Matsuzaki, Nobuyuki Uchikoga, Takashi Ishida, Yutaka Akiyama. "MEGADOCK 4.0: an ultra-high-performance protein-protein docking software for heterogeneous supercomputers", Biophysical Society 60th Annual Meeting, Los Angeles, CA, USA, Mar 1, 2016.
- ⑯ Masahito Ohue, Yutaka Akiyama: "A rescoring method of protein-peptide docking prediction with amino acid profiles of rigid-body search results",

The Fourteenth Asia Pacific Bioinformatics Conference (APBC2016), P188, San Francisco Bay Area, CA, USA, Jan 11-13 2016.

- ⑰ 大上雅史, "計算で明らかにするタンパク質の出会いとネットワーク", FIT2016 第15回情報科学技術フォーラム, 富山大学五福キャンパス, 9/9 2016.
- ⑱ Masahito Ohue, Takehiro Shimoda, Shuji Suzuki, Takashi Ishida, Yutaka Akiyama: "A comparative study of GPU and MIC accelerations in fast Fourier transform-based protein-protein docking", 3rd IIT Madras - Tokyo Tech Joint Symposium on Algorithms and Applications of Bioinformatics, P26, Nov 2015.

他

[その他]

- ① MEGADOCK-Web
<http://www.bi.cs.titech.ac.jp/megadock-web/>
- ② MEGADOCK
<http://www.bi.cs.titech.ac.jp/megadock/>
- ③ 大上雅史 個人ホームページ
<http://www.bi.cs.titech.ac.jp/~ohue/>

6. 研究組織

(1) 研究代表者

大上 雅史 (Ohue Masahito)
東京工業大学・情報理工学院・助教
研究者番号: 50743209