

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

林 孝紀^{1,2} 山本 悠生¹ 松崎 由理² 大上 雅史^{1,3} 秋山 泰^{1,2,3,a)}

概要: 我々は、タンパク質間相互作用 (Protein-Protein Interaction, PPI) の網羅的な理解に向けて、高速に PPI 予測計算を行うソフトウェアである MEGADOCK, および MEGADOCK による PPI 予測結果をより利活用しやすい形で提供するための PPI 予測統合データベース MEGADOCK-WEB を開発してきた。本研究では、アプリケーションと既存データベースとの通信回数を削減することで、MEGADOCK-WEB におけるデータ規模の大きさに起因する応答速度の遅さの問題について解決し、約 3600 万ペアに対する PPI 予測計算結果の閲覧を可能にした。さらに、未知のタンパク質構造に対するバックエンドでの PPI 予測計算を可能にするために、計算資源を動的に確保することが可能であるクラウド環境と連携した。クラウド計算環境を利用するための API を新規に開発し、それらの API を介してクラウドの仮想マシンと連携した。これにより、ユーザーがアップロードした任意のタンパク質の構造に対してオンデマンドに PPI 予測結果を提供することを可能にした。

キーワード: タンパク質間相互作用, PPI 予測, MEGADOCK, タンパク質ドッキング, クラウド計算環境, PPI 予測データベース

Improvement of MEGADOCK-WEB integrated database for predicted protein-protein interactions, and its coordination with cloud environment for on-demand dockings

TAKANORI HAYASHI^{1,2} YUKI YAMAMOTO¹ YURI MATSUZAKI² MASAHITO OHUE^{1,3}
YUTAKA AKIYAMA^{1,2,3,a)}

Abstract: We developed a protein-protein docking software, called MEGADOCK, to understand protein-protein interactions (PPIs) comprehensively. We are also developing a database of predicted PPIs and its web-based interface, named MEGADOCK-WEB, to make users access a vast amount of PPIs predicted by MEGADOCK easily. In this study, we solved the problem about latency caused by the huge data size by reducing the number of accesses to other databases. Thereby, approximately 36 million predicted PPIs can be browsed on MEGADOCK-WEB. Furthermore, we implemented on MEGADOCK-WEB a new feature of on-demand PPI prediction between two input structures which are not registered in the database yet. In order to facilitate docking calculation by MEGADOCK as the back-end on cloud computing environment, we developed dedicated APIs. Via these APIs, MEGADOCK-WEB can work together with the back-end MEGADOCK on cloud computing environment.

Keywords: protein-protein interaction (PPI), public cloud, protein-protein docking, MEGADOCK

¹ 東京工業大学 情報理工学院 情報理工学系,
Department of Computer Science, School of Computing,
Tokyo Institute of Technology
² 東京工業大学 情報生命博士教育院,
Education Academy of Computational Life Sciences, Tokyo

Institute of Technology
³ 東京工業大学科学 技術創成研究院 スマート創薬研究ユニット,
Advanced Computational Drug Discovery Unit, Institute of
Innovative Research, Tokyo Institute of Technology
a) akiyama@c.titech.ac.jp

1. 導入

生体内のタンパク質はお互いに相互作用しながら機能を発揮している [1]。この相互作用はタンパク質間相互作用 (Protein-Protein Interaction, PPI) と呼ばれており生命現象の中核を担っている。また、近年では疾患の要因とされる PPI を阻害することで効果を得る PPI 阻害薬の開発も進んでいる [2,3]。このため、PPI は生物学的に重要な研究対象であり、近年構造生物学や分子生物学において盛んに研究されている。しかし、タンパク質群に対して生化学的実験で網羅的に PPI を決定することは、時間や費用の面で困難である。したがって、計算機を用いて PPI 予測を行い、相互作用する可能性の高いタンパク質をあらかじめ選出することが有用である。

計算機によって PPI 予測を行う手法には、タンパク質の立体構造を用いる手法 [4] のほか、アミノ酸配列 [5] や共進化情報 [6] に基づく手法がある。この中で、タンパク質の立体構造を使用する方法の 1 つに本研究室で開発された MEGADOCK [7] がある。MEGADOCK はタンパク質ペアに対してドッキング計算を行い、その結果を利用して PPI の有無を予測をするソフトウェアである。また、簡潔でありながら精度を損なわない評価関数の設計や GPU 並列化などの工夫により高速に PPI 予測計算を行うことが可能である。

一方、MEGADOCK による PPI 予測結果をより利活用しやすい形で提供するために、我々はタンパク質間相互作用予測統合データベースである MEGADOCK-WEB [8] (<https://www.bi.cs.titech.ac.jp/megadock-web/>) を開発してきた。現在は Protein Data Bank (PDB) [9] 中のヒトに関する変異のない非冗長な約 6,000 タンパク質に対して、MEGADOCK による予測 PPI の検索機能、他データベースに収録されている既知 PPI の表示機能、クエリタンパク質に対する相互作用予測相手の生化学パスウェイ上での可視化機能、予測複合体構造の可視化機能の 4 つの機能を持つ。

従来の MEGADOCK-WEB には以下のような問題点が存在した。

問題点 (1) : 処理速度の問題から大量のタンパク質構造を扱うことができないため、ウェブサーバー上で稼働することが現実的でなかった。

問題点 (2) : ユーザーインターフェースの仕様上、3D ビューアを用いた複数の予測複合体構造の比較を同一画面で行うことができなかった。

問題点 (3) : ユーザーは PPI 予測データベースに登録されているタンパク質ペアの PPI 予測結果しか閲覧することができなかった。そのため、データベースに登録されていないタンパク質ペアについて新たに予測を行うことができなかった。

本研究では、以上の 3 つの問題点を解決し、MEGADOCK-WEB を改良することを目的とする。

2. 関連研究

予測 PPI 情報を閲覧できる既存のデータベースに Zhang らによって開発された PrePPI [10] がある。PrePPI は、PDB 中の PPI 複合体構造情報をテンプレートとし、このテンプレートやベイジアンネットワークを用いて予測された PPI の情報をデータベースとしてまとめたものであり、約 200 万件の PPI 情報を保有している。しかし、PrePPI のような既知の複合体構造を用いる予測では、既知 PPI と異なる様式の新規の PPI の発見が難しい。

3. MEGADOCK-WEB

MEGADOCK-WEB [8] は MEGADOCK による PPI 予測結果をより利活用しやすい形で提供するために長澤らによって開発され、現在も開発が進められているタンパク質間相互作用予測統合データベースである。MEGADOCK-WEB は主に以下の 4 つの機能を有する。

- 予測 PPI の検索機能
 - クエリタンパク質とデータベース中のタンパク質に対する予測 PPI 検索機能
 - 2 つのクエリタンパク質に対する予測 PPI 検索機能
- 既知 PPI の表示機能
- クエリタンパク質に対する相互作用予測相手の生化学 (生物の代謝やシグナル伝達) パスウェイ上での可視化機能 (本研究では、生物の代謝やシグナル伝達を閲覧可能な Kyoto Encyclopedia of Genes and Genomes (KEGG) [11] の KEGG パスウェイを用いている。以下では、パスウェイは KEGG パスウェイを指す)
- 予測複合体構造の可視化機能

以下では従来の MEGADOCK-WEB の問題点について述べる。

3.1 問題点 (1): データ規模による障害

処理速度の問題から大量のタンパク質構造を扱うことができないため、ウェブサーバー上で稼働することが現実的でなかった。具体的には以下の 2 つの問題点が存在した。

問題点 (1-a) : タンパク質の情報取得に大量の通信を必要とするため、ページの表示時にタイムアウトが発生していた。既存のシステムでは、既知 PPI を検索するため、ユーザーからのリクエストを受け付けるたびに Bioservices [12] を使用して、PPI 情報を集約した様々なデータベースに対する横断検索機能を提供する PSICQUIC [13] と通信を行っていた。このシステムは、常に最新の情報を表示できるという利点がある反面、ページの表示に必要なタンパク質の数が増えると PSICQUIC との通信コストが膨大になるという欠

点が存在した。実際に、データベースに格納するタンパク質が 100 個の場合、通信に 10 分以上要した。さらに、タンパク質が存在している生化学パスウェイの検索も、リクエストを受け付けるたびに通信を行っているため同様の問題が発生していた。

問題点 (1-b)：タンパク質の構造情報や PPI 予測結果を表すファイル群といった静的資源をアプリケーションの一部として jar 形式で管理していた。このため、静的資源を変更、追加する際にアプリケーションの再起動が必要となり、サービスが一時停止するという問題があった。

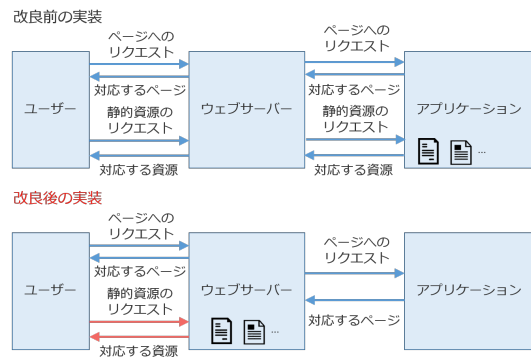


図 1 改良前と改良後のリクエストの流れの比較

3.2 問題点 (2): 単一の予測複合体のみ閲覧可能

MEGADOCK-WEB では、1 つの予測 PPI について複数の予測複合体構造を取得できる。しかし、3D 分子ビューアで予測複合体構造を閲覧する際、予測複合体構造の表示が 1 つの構造のみであった。このため、3D 分子ビューアを用いた複数の予測複合体構造の比較を同一画面で行うことができなかった。

3.3 問題点 (3): 新規構造への未対応

既存の MEGADOCK-WEB は事前に計算した PPI の予測結果のみを表示していた。したがって、ユーザーはデータベースに登録されていないタンパク質ペアについて新たに予測を行うことができず、PPI 予測結果を閲覧することができなかった。

4. MEGADOCK-WEB の改良

4.1 データベース設計の改良

3.1 節で示した問題点 (1-a), (1-b) をそれぞれ解決し、大量のタンパク質構造に対応させるために以下の改良を行った。

- データベース作成の際、あらかじめ KEGG パスウェイ、既知 PPI の情報をデータベースに登録しておくことで、ユーザーからのリクエストを受けた際に行われていた各データベースとの通信をすることなく、ウェブページの表示を可能にした。これにより、数秒程度でユーザーにウェブページを送信することを可能になった。
- 静的資源をアプリケーションの管理下に置かずウェブサーバーの管理下に置き、静的資源へのリクエストをアプリケーションではなくウェブサーバーが処理するように実装を行った (図 1)。これにより、サービスを一時停止させることなく静的資源の追加、変更を可能にした。

4.2 ユーザーインターフェースの改良

3.2 節で示した問題点 (2) を解決するために、ユーザーイ

	Protein Names	Gene Names	PDB ID
Receptor	C-1-tetrahydrofolate synthase cytoplasmic	MTHFD1 MTHFC MTHFD	1A4I (Chain B)
Ligand	C-1-tetrahydrofolate synthase cytoplasmic	MTHFD1 MTHFC MTHFD	1A4I (Chain B)

Rank	Score	Model	Download
11	3855.67		1A4I_B_1A4I_B_rank11.pdb
12	3834.36		1A4I_B_1A4I_B_rank12.pdb

図 2 改良前の予測複合体構造一覧表示ページ；静止画が埋め込まれており、クリックすることにより JSmol のページに遷移しなければ、構造を回転、拡大できない

MEGADOCK-WEB

Docking Information
PPIScore: 8.1506 Show KEGG Pathway

Protein1	Protein Names	Gene Names	PDB ID
Protein1	Adenomatous polyposis coli protein	APC DP2.5	1DEB (Chain B)
Protein2	Ataxin-1	ATXN1 ATX1 SCA1	1Q48 (Chain D)

Docking results from rank 1 to 10
Download complex in this page

Rank	Score	Model	Download
1	3508.92		1DEB_B-1Q48_D_decoy1.pdb
2	3306.45		1DEB_B-1Q48_D_decoy2.pdb

図 3 改良後の予測複合体構造一覧表示ページ；それぞれのカラムに 3D 分子ビューアが埋め込まれており、構造の回転、拡大が可能

ンターフェースの改良を行った。既存の予測複合体構造一覧表示ページ (図 2) では「Model」のカラムに画像を表示し、その画像をクリックすることにより対応する予測複合体構造のみを JSmol [14] で閲覧することが可能であった。

しかし、同一画面に複数の分子ビューアが埋め込まれていないため、複数の予測複合体構造を動かしながら比較することが不可能であった。JSmolは機能が豊富ではあるが動作が遅く、同一のページに10個を埋め込んだ場合は画面表示に数分程度の時間を要するため、図2の画面上に直接埋め込むとレイテンシーが増大してしまう。そこで、本研究では予測複合体構造一覧表示ページで用いる分子ビューアとして Bekker らによる Molmil [15] を採用した。Molmilの動作は軽量で、最大でも数十秒程度で10個の3D分子ビューアを埋め込んだページ(図3)を表示することが可能になった。

4.3 クラウド計算環境との連携

3.3節で示した問題点(3)を解決するために、クラウド環境との連携によりオンデマンドにPPI予測を行うシステムを開発した。本研究では、Microsoft Azure [16] (以下、Azure) 上の仮想マシン (virtual machine, VM) で MEGADOCK を実行するシステムと連携するためのAPIを作成した。これにより、外部から Azure 上に VM を作成してドッキング計算を行うことができるようになった。MEGADOCK-WEB ではこのAPIを通してウェブ上からドッキング計算を行うシステムを構築した。

Azure 上で動作する MEGADOCK とのAPIを通じた連携によって、ユーザーがウェブページからアップロードした2つのタンパク質構造ファイル群(便宜的にレセプター、リガンドと呼ぶ)に対して網羅的なドッキング計算をオンデマンドに行うことが可能になった。具体的には、ユーザーはアカウントを作成した後、PPI予測を行いたい2つのPDBファイル群をそれぞれzip形式に圧縮したものをユーザー画面にアップロードすることによりPPI予測結果を閲覧することができる。たとえば、レセプターとしてタンパク質 R1, R2, R3 のPDBファイル群をzip形式で圧縮したもの、リガンドとしてタンパク質 L1, L2 のPDBファイル群をzip形式に圧縮したものをそれぞれアップロードしたとする。このとき、ユーザーは (R1, L1), (R1, L2), (R2, L1), (R2, L2), (R3, L1), (R3, L2) の全6通りについてのPPI予測結果を受け取ることができる。

また、内部的な処理としては以下のような処理を行っている(図4)。

- (1) 新規ジョブの作成リクエストを送信し、一意に定まるIDとtokenを取得する。
- (2) レセプター、リガンドの各ファイルをアップロードする。
- (3) PPI予測したい構造の組み合わせについて書かれたテーブルファイルをアップロードする。
- (4) ジョブの実行要求を送信する。この際に、ジョブの終了時にアクセスされるURLを指定する。
- (5) 上で指定したURLにAzure上のシステムからのアク

セスがあった場合、計算結果を取得する。

- (6) ユーザーがメールに指定されたアドレスにアクセスすることで、ドッキング計算の結果をダウンロードすることができるように計算結果の加工をする。
- (7) アカウント作成時に登録されたメールアドレスにジョブの終了を知らせるメールを送信する。

なお、このシステムはセキュリティ上の問題から現在は外部のブラウザ上からのアクセスを許可せずに、内部的な利用にとどめている。今後、セキュリティを強化してシステムを公開することを目指す。

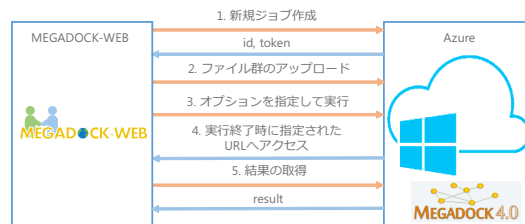


図4 MEGADOCK-WEBとAzureのAPIを介した通信

5. 改良後の MEGADOCK-WEB の構成と利用例

5.1 改良後の MEGADOCK-WEB の構成

改良後の MEGADOCK-WEB の画面遷移図を図5に示す。改良後の MEGADOCK-WEB のページ構成は以下の通りである。

- Top ページ: MEGADOCK-WEB のトップページでここからPPI予測検索機能を使用できる(図6(a))
- タンパク質選択ページ: 単一のタンパク質の検索結果ページ(図6(b))
- タンパク質ペア選択ページ: タンパク質ペアの検索結果ページ(図6(c))
- PPIスコア一覧表示ページ: クエリタンパク質とデータベース中のタンパク質とのPPI予測スコア一覧の表示ページ(図6(d))
- パスウェイページ: クエリタンパク質とPPIスコアが一定以上のタンパク質が属するKEGGパスウェイへのリンク表示ページ(図6(e))
- 予測立体構造一覧表示ページ: 選択したペアの予測立体構造ビューアを表示するページ(図6(f))
- ログインページ: アカウント作成後のユーザーがログインするためのページ
- ドッキングリクエスト送信ページ: ドッキング計算のリクエストを送信するページ

5.2 予測 PPI 検索機能の利用

MEGADOCK-WEB の予測 PPI 検索機能としての利用例を以下に示す。

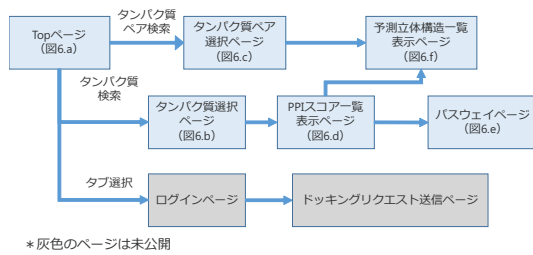


図 5 改良後の MEGADOCK-WEB の遷移図

Adenomatous polyposis coli (APC) タンパク質はガンの抑制に関連するタンパク質である [17]. 今回はこの APC をクエリとして使用し, APC の相互作用相手の候補のうちガンに関連していると思われるものを調べる. 検索の流れは以下になる.

- (1) “APC” をクエリとして検索する (図 7 (a)).
- (2) 検索の結果, 構造として 3 つの構造が表示される (図 7 (b)).
- (3) 上で得られた構造のうち, たとえば PDB ID が 1DEB のチェーン A の構造を選択すると PPI スコア一覧表示ページに相互作用予測相手と PPI スコアが表示される (図 7 (c)).
- (4) PPI スコア一覧表示ページにおいて, 相互作用予測相手の属するパスウェイを閲覧するために, PPI スコアが 8.0 以上のタンパク質が属する KEGG のパスウェイ一覧を表示する (図 7 (d)).
- (5) 今回はガンに関するパスウェイについて調べたいので, “cancer” という単語によるフィルターをかけて出てきたパスウェイである breast cancer と書いてあるパスウェイ (hsa05224) を選択する.
- (6) クエリタンパク質である APC (青色で表示) の横に予測相互作用相手である Axin (赤色で表示) が存在する (図 7 (e)).
- (7) APC と Axin についての相互作用を調べるために, APC と Axin をクエリとして検索する.
- (8) MEGADOCK-WEB の検索結果として複数の構造がヒットする. この中から, 1DEB のチェーン A の構造と 1DK8 のチェーン A の構造を選択する (図 7 (f)).
- (9) 予測複合体をビューアで閲覧し, 相互作用の可能性を判断する (図 7 (g)). また, 必要に応じて予測複合体構造の PDB ファイルもダウンロード可能である.

実際に, APC は Axin と相互作用すると報告されている [18]. 今回は PPI が既知であるタンパク質ペアを例に用いて説明したが, 同様の手順で未知の PPI の発見にも応用できると考えられる.

6. まとめ

本研究では, 以下のように MEGADOCK-WEB の改良を行った.

改良 (1): MEGADOCK-WEB のデータ規模に起因する問題について解決した.

改良 (1-a): 既知 PPI 情報と KEGG パスウェイの情報をあらかじめデータベースに格納しておくことにより, 他のデータベースとの通信を不要にした. これにより, 数秒程度の時間でウェブページを表示することを可能にした.

改良 (1-b): 静的資源をウェブアプリケーション本体から分離し, ウェブサーバーの管理下に配置した. これにより, アプリケーションの再起動を行うことなく, 静的資源の追加, 変更が可能になった.

改良 (2): Molmil を使用することによって, 同一画面における 3D ビューアを用いた複数の予測複合体構造の比較を可能にした.

改良 (3): クラウド上で動作する MEGADOCK と連携してオンデマンドにドッキング計算を行うことを可能にした.

今後の課題としては, 以下の 4 点があげられる.

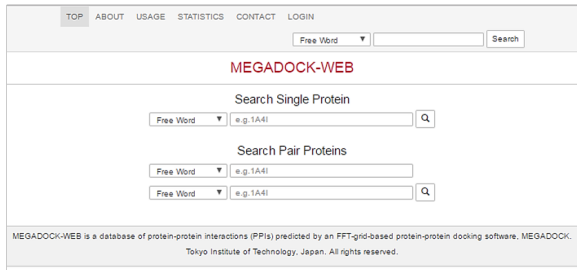
- 現在の MEGADOCK-WEB はまだユーザー視点の意見を得るまでには至っていないため, 構造生物学や生化学を専門とする研究者に利用してもらい, 表示系などに対する意見のフィードバックを行う必要がある.
- 現在収録しているタンパク質のデータはヒトに関するタンパク質のみであるため, 対応する生物種を増やす必要がある. また, 既知 PPI 収集に用いているデータベースの種類を増やすことにより, 研究者の利便性がさらに向上することが考えられる.
- 現在はクエリタンパク質と相互作用予測タンパク質を KEGG パスウェイの表示画面上に着色することで MEGADOCK における PPI 予測の結果の評価に使用できる判断材料を提供した. しかし, KEGG パスウェイとは別の視点の PPI 予測結果への評価に用いる判断材料を提供し, PPI に対してより多くの見識を得られるようにすることが望まれる.
- Azure との連携部分は, セキュリティ上の問題を抱えている (使用制限がない, 通信が暗号化されていないなど) ため, 内部利用のみとなっている. このため, セキュリティを強化して外部公開をすることでさらなる利便性の向上が見込める.

謝辞 本研究の一部は JSPS 科研費 (15K16081, 17H01814), JST CREST 「EBD: 次世代の年ヨッタバイト処理に向けたエクストリームビッグデータの基盤技術 (課題番号 JPMJCR1303)」の支援によって行われた.

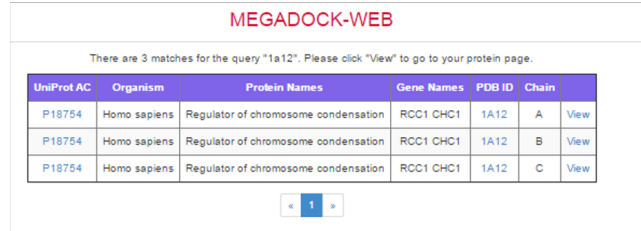
参考文献

- [1] Stelzl, U. *et al.* A human protein-protein interaction network: A resource for annotating the proteome. *Cell* 122,

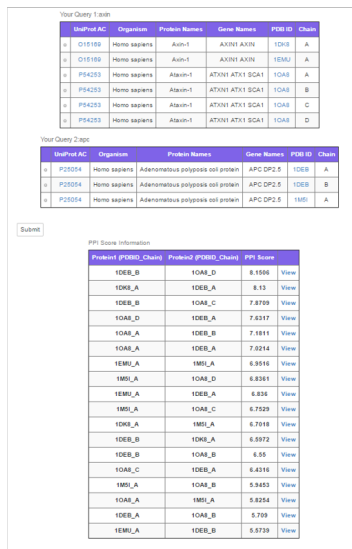
(a)



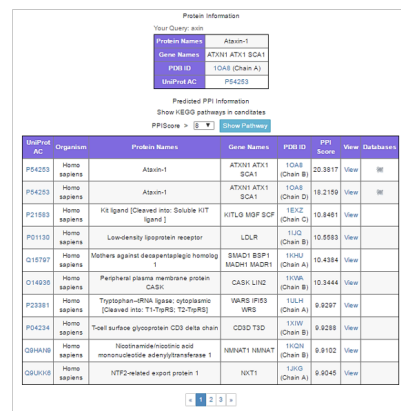
(b)



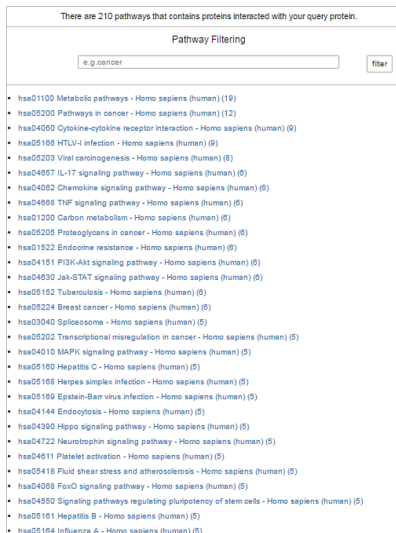
(c)



(d)



(e)



(f)



図 6 MEGADOCK-WEB ページ一覧 ; (a): Top ページ, (b): タンパク質選択ページ, (c): タンパク質ペア選択ページ, (d): PPI スコア一覧表示ページ, (e): パスウェイページ, (f): 予測立体構造一覧表示ページ

(a)

MEGADOCK-WEB is a database of protein-protein interactions (PPIs) predicted by an FFT-grid-based protein-protein docking software, MEGADOCK. Tokyo Institute of Technology, Japan. All rights reserved.

(b)

There are 3 matches for the query "APC". Please click "View" to go to your protein page.

UniProt AC	Organism	Protein Names	Gene Names	PDB ID	Chain	
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1DEB	A	View
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1DEB	B	View
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1M5I	A	View

(c)

Your Query: APC

Protein Names	Adenomatous polyposis coli protein
Gene Names	APC DP2.5
PDB ID	1DEB (Chain A)
UniProt AC	P25054

Predicted PPI Information
Show KEGG pathways in candidates
PPIScore > 8.0

UniProt AC	Organism	Protein Names	Gene Names	PDB ID	PPI Score	View	Databases
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1DEB (Chain B)	13.4563	View	
Q8N4E7	Homo sapiens	Ferritin, mitochondrial	FTMT	1R03 (Chain A)	12.2521	View	
Q09013	Homo sapiens	Myosin-protein kinase	DMPK DMPFK MDPK	1W76 (Chain D)	11.9903	View	
P08670	Homo sapiens	Vimentin	VIM	1GK4 (Chain B)	11.5430	View	
P61244	Homo sapiens	Protein max	MAX MMAX BHLHD4	1NKP (Chain E)	11.1648	View	
P08670	Homo sapiens	Vimentin	VIM	1GK6 (Chain A)	11.0672	View	
Q00161	Homo sapiens	Synaptosomal-associated protein 23	SNAP23	1NHL (Chain A)	10.8497	View	
P08670	Homo sapiens	Vimentin	VIM	1GK4	10.8067	View	

(d)

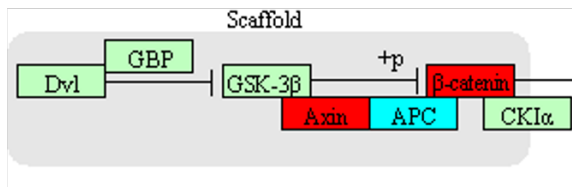
There are 197 pathways that contains proteins interacted with your query protein.

Pathway Filtering

cancer

- hsa05200 Pathways in cancer - Homo sapiens (human) (13)
- hsa01100 Metabolic pathways - Homo sapiens (human) (11)
- hsa05202 Transcriptional misregulation in cancer - Homo sapiens (human) (9)
- hsa05418 Fluid shear stress and atherosclerosis - Homo sapiens (human) (9)
- hsa05166 HTLV-1 infection - Homo sapiens (human) (8)
- hsa05162 Measles - Homo sapiens (human) (7)
- hsa04151 PI3K-Akt signaling pathway - Homo sapiens (human) (7)
- hsa05152 Tuberculosis - Homo sapiens (human) (7)
- hsa05169 Epstein-Barr virus infection - Homo sapiens (human) (6)
- hsa05222 Small cell lung cancer - Homo sapiens (human) (6)
- hsa04550 Signaling pathways regulating pluripotency of stem cells - Homo sapiens (human) (6)
- hsa05164 Influenza A - Homo sapiens (human) (6)
- hsa05210 Colorectal cancer - Homo sapiens (human) (6)
- hsa05213 Endometrial cancer - Homo sapiens (human) (6)
- hsa05142 Chagas disease (American trypanosomiasis) - Homo sapiens (human) (6)
- hsa04510 Focal adhesion - Homo sapiens (human) (5)
- hsa04015 Rap1 signaling pathway - Homo sapiens (human) (5)
- hsa04210 Apoptosis - Homo sapiens (human) (5)
- hsa05160 Hepatitis C - Homo sapiens (human) (5)
- hsa04390 Hippo signaling pathway - Homo sapiens (human) (5)
- hsa04014 Ras signaling pathway - Homo sapiens (human) (5)
- hsa04022 cGMP-PKG signaling pathway - Homo sapiens (human) (5)
- hsa04933 AGE-RAGE signaling pathway in diabetic complications - Homo sapiens (human) (5)

(e)



(f)

Your Query 1: apc

UniProt AC	Organism	Protein Names	Gene Names	PDB ID	Chain
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1DEB	A
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1DEB	B
P25054	Homo sapiens	Adenomatous polyposis coli protein	APC DP2.5	1M5I	A

Your Query 2: axin

UniProt AC	Organism	Protein Names	Gene Names	PDB ID	Chain
Q15169	Homo sapiens	Axin-1	AXIN1 AXIN	1DK8	A
Q15169	Homo sapiens	Axin-1	AXIN1 AXIN	1EMU	A
P54253	Homo sapiens	Axin-1	ATXN1 ATX1 SCAT1	1OA8	A
P54253	Homo sapiens	Axin-1	ATXN1 ATX1 SCAT1	1OA8	B
P54253	Homo sapiens	Axin-1	ATXN1 ATX1 SCAT1	1OA8	C
P54253	Homo sapiens	Axin-1	ATXN1 ATX1 SCAT1	1OA8	D

Submit

PPI Score Information

Protein1 (PDBID_Chain)	Protein2 (PDBID_Chain)	PPI Score	
1DEB_B	1OA8_D	8.1506	View
1DK8_A	1DEB_A	8.13	View
1DEB_B	1OA8_C	7.8709	View
1OA8_D	1DEB_A	7.6317	View
1OA8_A	1DEB_B	7.1811	View
1OA8_A	1DEB_A	7.0214	View
1EMU_A	1M5I_A	6.9516	View
1M5I_A	1OA8_D	6.8361	View
1EMU_A	1DEB_A	6.836	View
1M5I_A	1OA8_C	6.7529	View
1DK8_A	1M5I_A	6.7018	View
1DEB_B	1DK8_A	6.5972	View
1DEB_B	1OA8_B	6.55	View

(g)

MEGADOCK-WEB

Docking Information
PPIScore: 8.13 Show KEGG Pathway

Protein1	Protein Names	Gene Names	PDB ID
Protein1	Axin-1	AXIN1 AXIN	1DK8 (Chain A)
Protein2	Adenomatous polyposis coli protein	APC DP2.5	1DEB (Chain A)

Docking results from rank 1 to 10
Download complex in this page

Rank	Score	Model	Download
1	3090.43		1DK8_A-1DEB_A_dock1.pdb
2	3076.25		1DK8_A-1DEB_A_dock2.pdb

図 7 MEGADOCK-WEB の検索機能利用例 ; (a): 検索画面, (b): “APC” に対する検索結果, (c): 1DEB のチェーン A に対する PPI スコア一覧表示画面, (d) 1DEB のチェーン A に対して PPI スコアが 8.0 以上のタンパク質が属するパスウェイ一覧, (e): hsa05224 パスウェイの一部, (f): “APC”, “Axin” の 2 つのクエリに対する検索結果, (g): 1DEB のチェーン A と 1DK8 のチェーン A の予測複合体構造一覧表示ページ

- 957–968 (2005).
- [2] Oltersdorf, T. *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* 435, 677–681 (2005).
- [3] Popowicz, G. M. *et al.* Structures of low molecular weight inhibitors bound to MDMX and MDM2 reveal new approaches for p53-MDMX/MDM2 antagonist drug discovery. *Cell Cycle* 9, 1104–1111 (2010).
- [4] Chen, R. *et al.* ZDOCK: An initial-stage protein-docking algorithm. *Proteins* 52, 80–87 (2003).
- [5] Murakami, Y. *et al.* Homology-based prediction of interactions between proteins using Averaged One-Dependence Estimators. *BMC Bioinformatics* 15, 213 (2014).
- [6] Zhou, H. *et al.* Predicting protein-protein interaction by the mirrortree method: Possibilities and limitations. *PLOS ONE* 8, (2013).
- [7] Ohue, M. *et al.* MEGADOCK 4.0: An ultra-high-performance protein–protein docking software for heterogeneous supercomputers. *Bioinformatics* 30, 3281–3283 (2014).
- [8] 長澤 一輝, 他. タンパク質間相互作用予測結果データベースおよび表示系の構築. 情報処理学会研究報告 2016-BIO-45(3), 1–4 (2016).
- [9] Rose, P. W. *et al.* The RCSB Protein Data Bank: Views of structural biology for basic and applied research and education. *Nucleic Acids Research* 43, D345–D356 (2015).
- [10] Zhang, Q. C., *et al.* PrePPI: A structure-informed database of protein-protein interactions. *Nucleic Acids Research* 41, D828–D833 (2013).
- [11] Kanehisa, M., *et al.* KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* 45, D353–D361 (2017).
- [12] Cokelaer, T. *et al.* BioServices: A common Python package to access biological Web Services programmatically. *Bioinformatics* 29, 3241–3242 (2013).
- [13] Aranda, B. *et al.* PSICQUIC and PSIScore: accessing and scoring molecular interactions. *Nature Methods* 8, 528–529 (2011).
- [14] JSmol: an open-source HTML5 viewer for chemical structures in 3D. <http://www.jmol.org>.
- [15] Bekker, G. J., *et al.* Molmil: a molecular viewer for the PDB and beyond. *Journal of Cheminformatics* 8, 42 (2016).
- [16] Microsoft Azure <http://azure.microsoft.com/ja-jp/>.
- [17] Morin, P. J. Activation of beta-Catenin-Tcf Signaling in Colon Cancer by Mutations in beta-Catenin or APC. *Science* 275, 1787–1790 (1997).
- [18] Behrens, J. Functional Interaction of an Axin Homolog, Conductin, with -Catenin, APC, and GSK3. *Science* 280, 596–599 (1998).