

## *Additional File*

# **MEGADOCK-Web: an integrated database of high-throughput structure-based protein-protein interaction predictions**

Takanori Hayashi<sup>1,2</sup>, Yuri Matsuzaki<sup>2</sup>, Keisuke Yanagisawa<sup>1,2</sup>,  
Masahito Ohue<sup>1,3,4\*</sup>, Yutaka Akiyama<sup>1,2,3,4,5\*</sup>

<sup>1</sup> Department of Computer Science, School of Computing, Tokyo Institute of Technology, 2-12-1 W8-76 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

<sup>2</sup> Education Academy of Computational Life Sciences (ACLS), Tokyo Institute of Technology, 2-12-1 W8-93 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

<sup>3</sup> Advanced Computational Drug Discovery Unit (ACDD), Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsutacho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan

<sup>4</sup> AIST-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory (RWBC-OIL), National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8560, Japan

<sup>5</sup> Molecular Profiling Research Center for Drug Discovery (molprof), National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

\*Corresponding authors

Email addresses:

TH: hayashi@bi.c.titech.ac.jp

YM: matsuzaki@acsls.titech.ac.jp

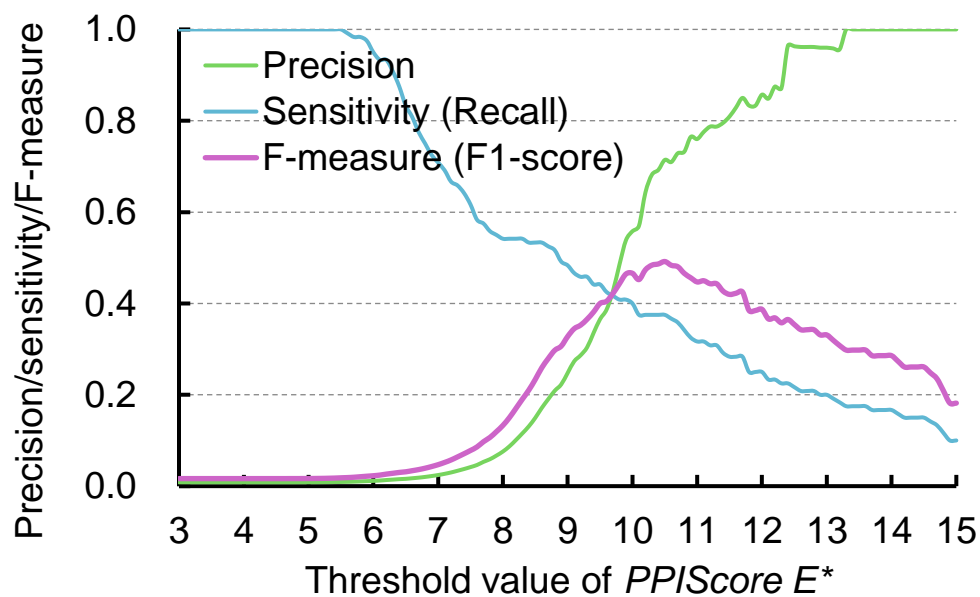
KY: yanagisawa@bi.c.titech.ac.jp

MO: ohue@c.titech.ac.jp

YA: akiyama@c.titech.ac.jp

## Supporting Figures

**Figure S1: Precision and recall at each threshold value with complete cross-docking on protein-protein docking benchmark version 4.0\***



\*Hwang H, Vreven T, Janin J, Weng Z. Protein-protein docking benchmark version 4.0. Proteins 2010;78:3111-4.

**Figure S2: MEGADOCK-Web top page**

TOP ABOUT USAGE STATISTICS CONTACT LOGIN

Free Word  Search

## MEGADOCK-Web

---

### Search Single Protein

Free Word

### Search Pair Proteins

Free Word

Free Word

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.

In the top page, there are two functions, i.e., “Search Single Protein” and “Search Pair Proteins”.

### Figure S3: Protein selection page

There are 3 matches for the query "1a12". Please click "View" to go to your protein page.

| UniProt AC             | Organism     | Protein Names                        | Gene Names | PDB ID               | Chain |                      |
|------------------------|--------------|--------------------------------------|------------|----------------------|-------|----------------------|
| <a href="#">P18754</a> | Homo sapiens | Regulator of chromosome condensation | RCC1 CHC1  | <a href="#">1A12</a> | A     | <a href="#">View</a> |
| <a href="#">P18754</a> | Homo sapiens | Regulator of chromosome condensation | RCC1 CHC1  | <a href="#">1A12</a> | B     | <a href="#">View</a> |
| <a href="#">P18754</a> | Homo sapiens | Regulator of chromosome condensation | RCC1 CHC1  | <a href="#">1A12</a> | C     | <a href="#">View</a> |

This page shows candidate protein structure hits after making a query. The table contains basic information on proteins, including PDB ID, chain, UniProt AC, gene name, and protein name for each protein.

**Figure S4: *PPIScore* list page**

**Protein Information**



Your Query: 1oa8

|                      |                 |
|----------------------|-----------------|
| <b>Protein Names</b> | Ataxin-1        |
| <b>Gene Names</b>    | ATXN1 ATX1 SCA1 |
| <b>PDB ID</b>        | 1OA8 (Chain A)  |
| <b>UniProt AC</b>    | P54253          |

**Predicted PPI Information**

Show KEGG pathways in candidates

PPIScore >  [Show Pathway](#)

| UniProt AC | Organism     | Protein Names   | Gene Names             | PDB ID         | PPIScore | View                 | Databases   |
|------------|--------------|---|------------------------|----------------|----------|----------------------|---|
| P54253     | Homo sapiens | Ataxin-1  | ATXN1 ATX1 SCA1        | 1OA8 (Chain B) | 20.3817  | <a href="#">View</a> |  |
| P54253     | Homo sapiens | Ataxin-1  | ATXN1 ATX1 SCA1        | 1OA8 (Chain D) | 18.2159  | <a href="#">View</a> |  |
| P21583     | Homo sapiens | Kit ligand [Cleaved into: Soluble KIT ligand ]                          | KITLG MGF SCF          | 1EXZ (Chain C) | 10.8461  | <a href="#">View</a> |   |
| P01130     | Homo sapiens | Low-density lipoprotein receptor  | LDLR                   | 1IJQ (Chain B) | 10.5583  | <a href="#">View</a> |   |
| Q15797     | Homo sapiens | Mothers against decapentaplegic homolog 1                               | SMAD1 BSP1 MADH1 MADR1 | 1KHU (Chain A) | 10.4384  | <a href="#">View</a> |   |
| O14936     | Homo sapiens | Peripheral plasma membrane protein CASK                                 | CASK LIN2              | 1KWA (Chain B) | 10.3444  | <a href="#">View</a> |   |
| P23381     | Homo sapiens | Tryptophan--tRNA ligase; cytoplasmic [Cleaved into: T1-TrpRS; T2-TrpRS] | WARS IFI53 WRS         | 1ULH (Chain A) | 9.9297   | <a href="#">View</a> |   |
| P04234     | Homo sapiens | T-cell surface glycoprotein CD3 delta chain                             | CD3D T3D               | 1XIW (Chain B) | 9.9288   | <a href="#">View</a> |   |
| Q9HAN9     | Homo sapiens | Nicotinamide/nicotinic acid mononucleotide adenylyltransferase 1        | NMNAT1 NMNAT           | 1KQN (Chain B) | 9.9102   | <a href="#">View</a> |   |
| Q9UKK6     | Homo sapiens | NTF2-related export protein 1   | NXT1                   | 1JKG (Chain A) | 9.9045   | <a href="#">View</a> |   |

This page shows candidate proteins interacting with the query protein on the basis of *PPIScore*. Specifying the threshold of *PPIScore* and pressing the “Show Pathway” button at the top of the page results in transition to the pathway selection page, whereas clicking the “View” link in the table takes the viewer to the predicted complex page.

**Figure S5: Predicted complex page**



In this page, complex structures predicted by MEGADOCK for the two proteins are displayed with the 3D molecular viewer, Molmil. A PDB-format file can be downloaded for each predicted complex, and 10 PDB files can be compressed into zip format from the “Download complexes in this page” link at the top of the page.

## Figure S6: Pathway selection page

There are 210 pathways that contains proteins interacted with your query protein.

---

Pathway Filtering

filter

---

- [hsa01100 Metabolic pathways - Homo sapiens \(human\)](#) (19)
- [hsa05200 Pathways in cancer - Homo sapiens \(human\)](#) (12)
- [hsa04060 Cytokine-cytokine receptor interaction - Homo sapiens \(human\)](#) (9)
- [hsa05166 HTLV-I infection - Homo sapiens \(human\)](#) (9)
- [hsa05203 Viral carcinogenesis - Homo sapiens \(human\)](#) (8)
- [hsa04657 IL-17 signaling pathway - Homo sapiens \(human\)](#) (8)
- [hsa04062 Chemokine signaling pathway - Homo sapiens \(human\)](#) (8)
- [hsa04668 TNF signaling pathway - Homo sapiens \(human\)](#) (8)
- [hsa01200 Carbon metabolism - Homo sapiens \(human\)](#) (8)
- [hsa05205 Proteoglycans in cancer - Homo sapiens \(human\)](#) (8)
- [hsa01522 Endocrine resistance - Homo sapiens \(human\)](#) (8)
- [hsa04151 PI3K-Akt signaling pathway - Homo sapiens \(human\)](#) (8)
- [hsa04630 Jak-STAT signaling pathway - Homo sapiens \(human\)](#) (8)
- [hsa05152 Tuberculosis - Homo sapiens \(human\)](#) (8)
- [hsa05224 Breast cancer - Homo sapiens \(human\)](#) (8)
- [hsa03040 Spliceosome - Homo sapiens \(human\)](#) (5)
- [hsa05202 Transcriptional misregulation in cancer - Homo sapiens \(human\)](#) (5)
- [hsa04010 MAPK signaling pathway - Homo sapiens \(human\)](#) (5)
- [hsa05160 Hepatitis C - Homo sapiens \(human\)](#) (5)
- [hsa05168 Herpes simplex infection - Homo sapiens \(human\)](#) (5)
- [hsa05169 Epstein-Barr virus infection - Homo sapiens \(human\)](#) (5)
- [hsa04144 Endocytosis - Homo sapiens \(human\)](#) (5)
- [hsa04390 Hippo signaling pathway - Homo sapiens \(human\)](#) (5)
- [hsa04722 Neurotrophin signaling pathway - Homo sapiens \(human\)](#) (5)
- [hsa04611 Platelet activation - Homo sapiens \(human\)](#) (5)
- [hsa05418 Fluid shear stress and atherosclerosis - Homo sapiens \(human\)](#) (5)
- [hsa04068 FoxO signaling pathway - Homo sapiens \(human\)](#) (5)
- [hsa04550 Signaling pathways regulating pluripotency of stem cells - Homo sapiens \(human\)](#) (5)
- [hsa05161 Hepatitis B - Homo sapiens \(human\)](#) (5)
- [hsa05164 Influenza A - Homo sapiens \(human\)](#) (5)

In this page, the pathways to which the predicted binders belong are shown as a list. The pathways can be filtered according to a word or phrase typed in the “Pathway Filtering” box at the top of the page. The number in parentheses next to the pathway name indicates the number of predicted binders included in the pathway.

**Figure S7: Protein pair selection page**

Your Query 1: axin

|                       | UniProt AC             | Organism     | Protein Names | Gene Names      | PDB ID               | Chain |
|-----------------------|------------------------|--------------|---------------|-----------------|----------------------|-------|
| <input type="radio"/> | <a href="#">O15169</a> | Homo sapiens | Axin-1        | AXIN1 AXIN      | <a href="#">1DK8</a> | A     |
| <input type="radio"/> | <a href="#">O15169</a> | Homo sapiens | Axin-1        | AXIN1 AXIN      | <a href="#">1EMU</a> | A     |
| <input type="radio"/> | <a href="#">P54253</a> | Homo sapiens | Ataxin-1      | ATXN1 ATX1 SCA1 | <a href="#">1OA8</a> | A     |
| <input type="radio"/> | <a href="#">P54253</a> | Homo sapiens | Ataxin-1      | ATXN1 ATX1 SCA1 | <a href="#">1OA8</a> | B     |
| <input type="radio"/> | <a href="#">P54253</a> | Homo sapiens | Ataxin-1      | ATXN1 ATX1 SCA1 | <a href="#">1OA8</a> | C     |
| <input type="radio"/> | <a href="#">P54253</a> | Homo sapiens | Ataxin-1      | ATXN1 ATX1 SCA1 | <a href="#">1OA8</a> | D     |

Your Query 2: apc

|                       | UniProt AC             | Organism     | Protein Names                      | Gene Names | PDB ID               | Chain |
|-----------------------|------------------------|--------------|------------------------------------|------------|----------------------|-------|
| <input type="radio"/> | <a href="#">P25054</a> | Homo sapiens | Adenomatous polyposis coli protein | APC DP2.5  | <a href="#">1DEB</a> | A     |
| <input type="radio"/> | <a href="#">P25054</a> | Homo sapiens | Adenomatous polyposis coli protein | APC DP2.5  | <a href="#">1DEB</a> | B     |
| <input type="radio"/> | <a href="#">P25054</a> | Homo sapiens | Adenomatous polyposis coli protein | APC DP2.5  | <a href="#">1M5I</a> | A     |

Submit

PPI Score Information

| Protein1 (PDBID_Chain) | Protein2 (PDBID_Chain) | PPI Score |                      |
|------------------------|------------------------|-----------|----------------------|
| 1DEB_B                 | 1OA8_D                 | 8.1506    | <a href="#">View</a> |
| 1DK8_A                 | 1DEB_A                 | 8.13      | <a href="#">View</a> |
| 1DEB_B                 | 1OA8_C                 | 7.8709    | <a href="#">View</a> |
| 1OA8_D                 | 1DEB_A                 | 7.6317    | <a href="#">View</a> |
| 1OA8_A                 | 1DEB_B                 | 7.1811    | <a href="#">View</a> |
| 1OA8_A                 | 1DEB_A                 | 7.0214    | <a href="#">View</a> |
| 1EMU_A                 | 1M5I_A                 | 6.9516    | <a href="#">View</a> |
| 1M5I_A                 | 1OA8_D                 | 6.8361    | <a href="#">View</a> |
| 1EMU_A                 | 1DEB_A                 | 6.836     | <a href="#">View</a> |
| 1M5I_A                 | 1OA8_C                 | 6.7529    | <a href="#">View</a> |
| 1DK8_A                 | 1M5I_A                 | 6.7018    | <a href="#">View</a> |
| 1DEB_B                 | 1DK8_A                 | 6.5972    | <a href="#">View</a> |
| 1DEB_B                 | 1OA8_B                 | 6.55      | <a href="#">View</a> |

In this page, a list of candidate hits for the two queries is displayed. In addition, at the bottom of the page, a table showing the all-to-all *PPI Score* of the protein hit by query 1 and the protein hit by query 2 is displayed.