

Poster 32

蛋白質構造データバンク (PDBj) の高度化と統合的運用

工藤高裕⁽¹⁾、Gert-Jan Bekker⁽¹⁾、山下鈴子⁽¹⁾、鈴木博文⁽²⁾、横地政志⁽¹⁾、由良敬⁽²⁾、栗栖源嗣⁽¹⁾

(1) 大阪大学蛋白質研究所 (2) 早稲田大学先進理工学部

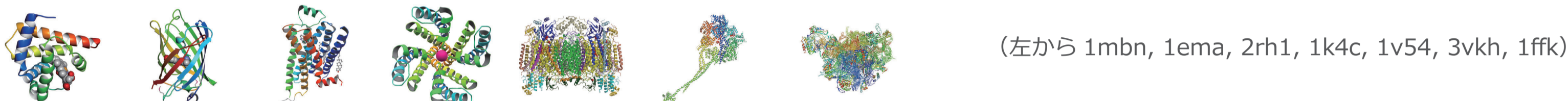


<https://pdbj.org/>

蛋白質構造データバンク (Protein Data Bank, PDB) は生体高分子の構造データを集めた世界で唯一のデータアーカイブ。1971年に7つの構造から始まり、2019年9月現在では150,000件以上のデータを提供している。PDBはPDBを維持するための国際組織「国際蛋白質構造データバンク」(worldwide Protein Data Bank, wwPDB)の下、4つのメンバー(米国のRCSB PDBとBMRB、欧州のPDBe、日本のPDBj)が協力して運営している。各メンバーの運営費用は各国の政府機関による研究費用でまかなわれている。

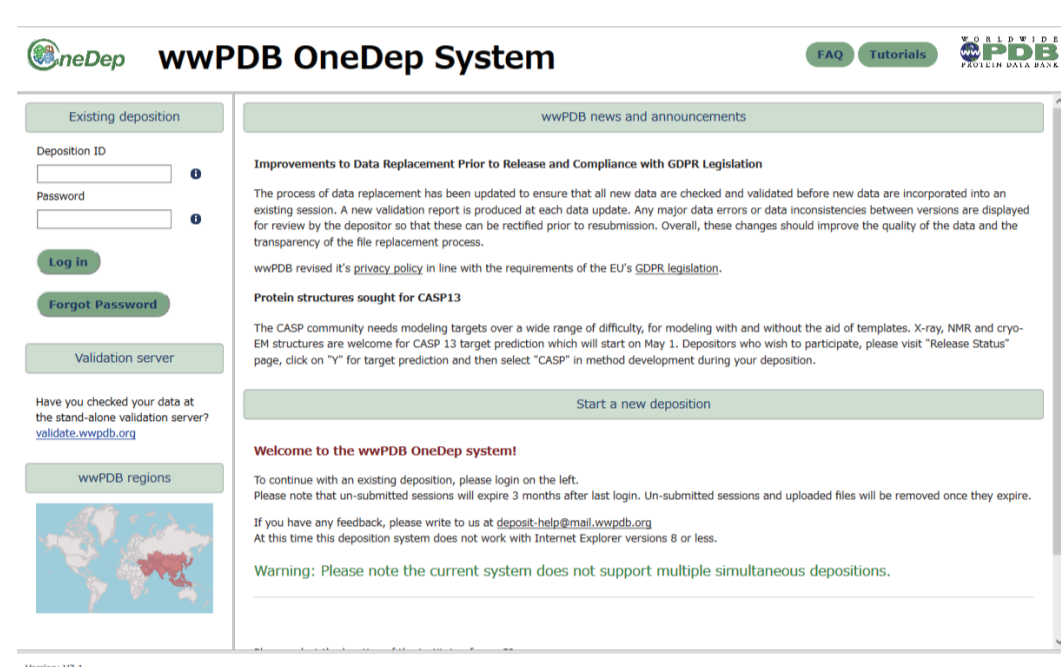


PDBjの活動



PDBデータの登録

<https://deposit-pdbj.wwpdb.org/deposition>



PDBjでは主にアジア地区からのPDB登録受付を行っている。登録された構造は決められた公開日になるとPDBの各拠点から同時に公開される。

Sequence Navigator : 類似配列検索

<https://pdbj.org/seq-navi>

あるアミノ酸配列と似た配列をPDB登録エントリーの中から探すBLAST検索サービス。アミノ酸配列または既にPDBに登録されている構造のPDBID+ChainIDで検索できる。

Structure Navigator : 類似立体構造検索

<https://pdbj.org/struc-navi>

立体構造が似た分子をPDBに登録された分子の中から検索するサービス。検索条件はPDBID+ChainID、またはPDB形式の構造情報ファイルで指定。

Spanner/SFAS : 配列から立体構造を予測

<https://sysimm.org/spanner/>
<https://sysimm.org/pipeline7/>

Spanner (スパンナー) はアミノ酸配列情報からタンパク質の立体構造を予測するサービス。似た配列を持ち構造が分かっている鑄型構造を別途指定し、それに基づいて予測を行う。鑄型構造をPDBエントリーから探し出す過程をつなげたパイプラインサービスSFAS (Sequence Function Annotation Service) も提供している。

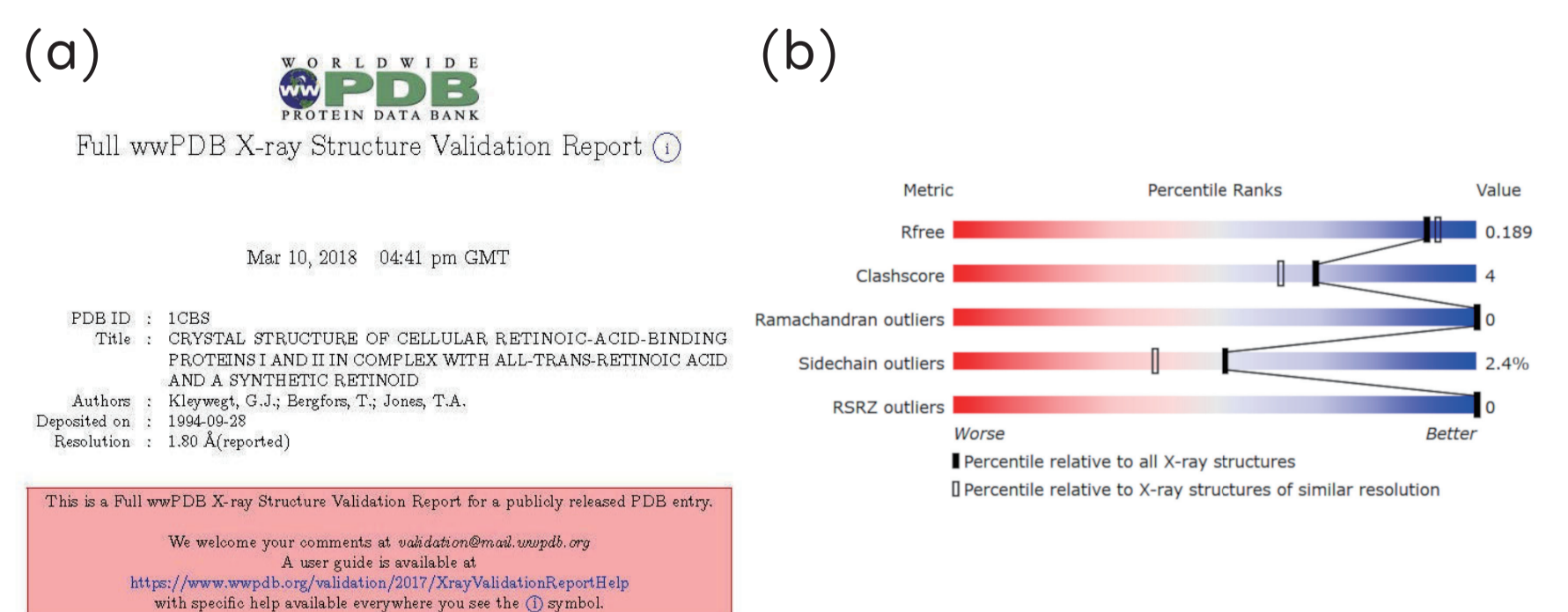
wwPDB検証レポート

PDBアーカイブの検証レポート (wwPDB validation reports) は、X線結晶回折、核磁気共鳴、電子顕微鏡の構造決定法毎の専門委員会 (Validation Task Force) の意見に基づいて、広く受け入れられている評価基準を使って、PDBに登録された立体構造の品質を評価する文書。特に論文の投稿・査読過程における利用が推奨されている。

PDBjでは、検証レポートに含まれる評価指標を検索しやすくするため、PDBML-validation (XML形式) とwwPDB/RDF-validation (RDF形式) に変換した検証レポートファイルを開発した。これらのフォーマットは機械処理に向いておりPDBアーカイブのマスターフォーマットとなっているPDBx/mmCIF形式と高い互換性があるという特徴を備えている。

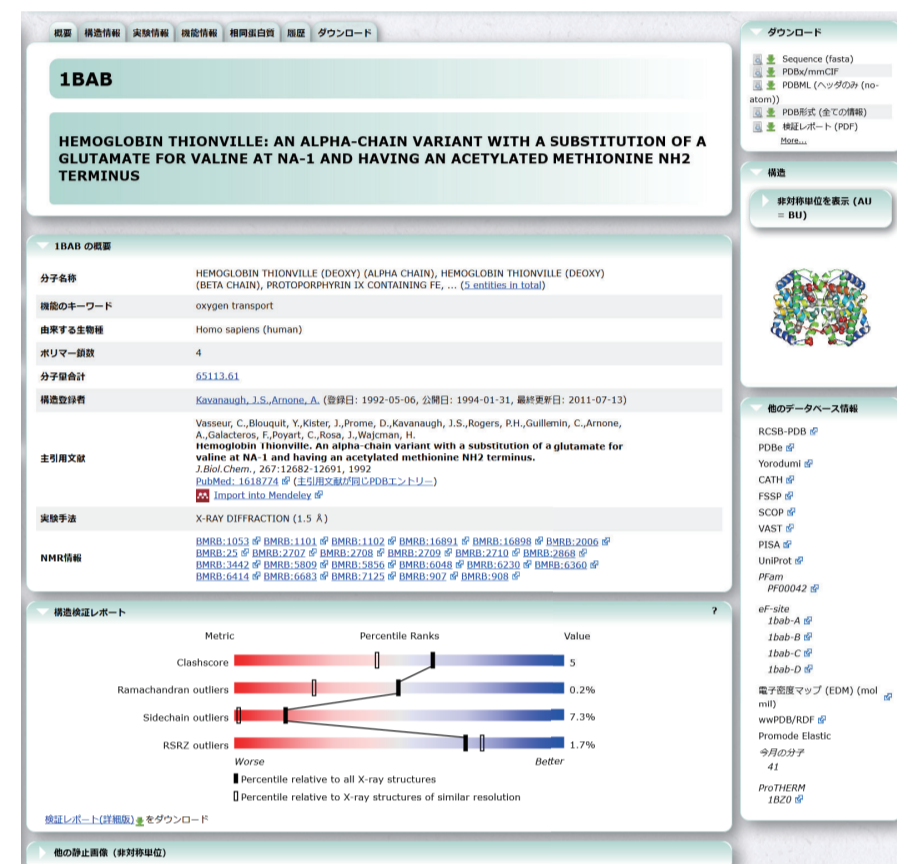
これらを利用し、検証レポートの関係データベース (PostgreSQL) およびSPARQLエンドポイントを公開している。これらは立体構造に基づくバイオインフォマティクス、創薬の分野において活用が期待される。

検証レポートの例



PDBj Mine (PDB構造検索)

<https://pdbj.org/mine>



PDBj MineはPDBjが提供するPDBエントリー検索サービス。キーワード検索だけでなく、詳細検索やSQL検索を使ってより細かな条件を指定し検索することもできる。また、この検索ではPDBで規定した化合物、PDBサイトの記事内容 (ニュース、ヘルプなど)、未公開エントリーの状態 (公開日や現在の処理段階) など合わせて横断検索することができる。日本語ページでは日本語でも検索できる。

EM Navigator・Omokage検索・万見

<https://pdbj.org/emnavi/>
<https://pdbj.org/emnavi/omo-search.php>
<https://pdbj.org/yorodumi/>

EM Navigatorは、電子顕微鏡で得られた電子密度マップのデータベースEMDBのデータを検索・閲覧するためのウェブサービス。

Omokage検索は、PDBの原子座標モデル、EMDBの電子密度マップ、SASBDB (生体試料小角散乱データバンク) の分子モデルの中から横断的に検索し、似た「形状」の分子を探すサービス。

以上のサービスは、専門的な知識や経験がなくても気軽に分子を閲覧できるようつくられた共通の分子構造閲覧ウェブサービス「万見」(よろづみ) を使っている。

wwPDB/RDF-validationのSPARQL検索例

```

PREFIX PDBov: <https://rdf.wwpdb.org/schema/pdbx-validation-v1.owl#>
SELECT ?PDB_ID ?enzyme ?ligand ?comp_id MIN(?RSR AS ?minRSR)
FROM <https://rdf.wwpdb.org/pdb-validation>
WHERE {
  ?entity PDBov:link_to_enzyme ?link_to_enzyme ;
  PDBov:entity_pdbx_description ?enzyme ;
  PDBov:of_datablock ?datablock .
  BIND (SUBSTR(STR(?datablock),38,4) AS ?PDB_ID)
  BIND (IRI (CONCAT(?datablock, "/pdbx_entity_nonpolyCategory")) AS ?entity_nonpoly_category)
  ?entity_nonpoly_category PDBov:has_pdbx_entity_nonpoly ?entity_nonpoly .
  ?entity_nonpoly PDBov:entity_name ?ligand ;
  PDBov:entity_nonpoly.entity_id ?entity_id ;
  PDBov:entity_nonpoly.comp_id ?comp_id .
  FILTER (?ligand!="water" && !STRENDS(?ligand, " ION"))
  BIND (IRI (CONCAT(?datablock, "/pdbx_nonpoly_schemeCategory")) AS ?nonpoly_scheme_category)
  ?nonpoly_scheme_category PDBov:has_pdbx_nonpoly_scheme ?nonpoly_scheme .
  ?nonpoly_scheme PDBov:entity_name ?seq_id ;
  PDBov:entity_nonpoly_scheme.pdb_strand_id ?asm_id ;
  PDBov:entity_nonpoly_scheme.pdb_seq_num ?seq_id ;
  PDBov:entity_nonpoly_scheme.entity_id ?entity_id ;
  PDBov:entity_nonpoly_scheme.mon_id ?comp_id .
  BIND (IRI (CONCAT(?datablock, "/pdbx_dcc_mapCategory")) AS ?dcc_map_category)
  ?dcc_map_category PDBov:has_pdbx_dcc_map ?dcc_map .
  ?dcc_map PDBov:entity_name ?asm_id ;
  PDBov:entity_nonpoly_scheme.pdb_strand_id ?asm_id ;
  PDBov:entity_nonpoly_scheme.pdb_seq_num ?seq_id ;
  PDBov:entity_nonpoly_scheme.entity_id ?entity_id ;
  PDBov:entity_nonpoly_scheme.mon_id ?comp_id .
  FILTER (xsd:float(?RSR) < 0.1)
} GROUP BY ?PDB_ID ?enzyme ?ligand ?comp_id
  
```

← 酵素を選択

← 水・イオンを除く非ポリマーのリガンドを選択

← RSR<0.1を選択

リガンドのRSR (Real space R-factor, 残基毎に実験的に求めた電子密度と構造モデルの電子密度の残差の割合) が10%より小さい全ての酵素-リガンド複合体を選択

日本蛋白質構造データバンク (PDBj)
〒565-0871 大阪府吹田市山田丘3-2
大阪大学蛋白質研究所・構造解析研究棟
蛋白質データベース開発研究室
TEL:06-6879-4311 (全般)
06-6879-8634 (PDB登録)
Web問い合わせ: <https://pdbj.org/contact>

NBDC INSTITUTE for PROTEIN RESEARCH OSAKA UNIVERSITY

日本蛋白質構造データバンク (PDBj: Protein Data Bank Japan) は、「JST-NBDC」「大阪大学蛋白質研究所に措置された共同利用・共同研究拠点経費 (文部科学省)」の支援を受け、米国 RCSB、BMRB、および欧州 PDBe と協力して、生体高分子の立体構造データベースを国際的に統一化された PDB アーカイブとして運営するとともに、様々な解析ツールを提供しています。