

タンパク質立体構造のデータ 利用入門

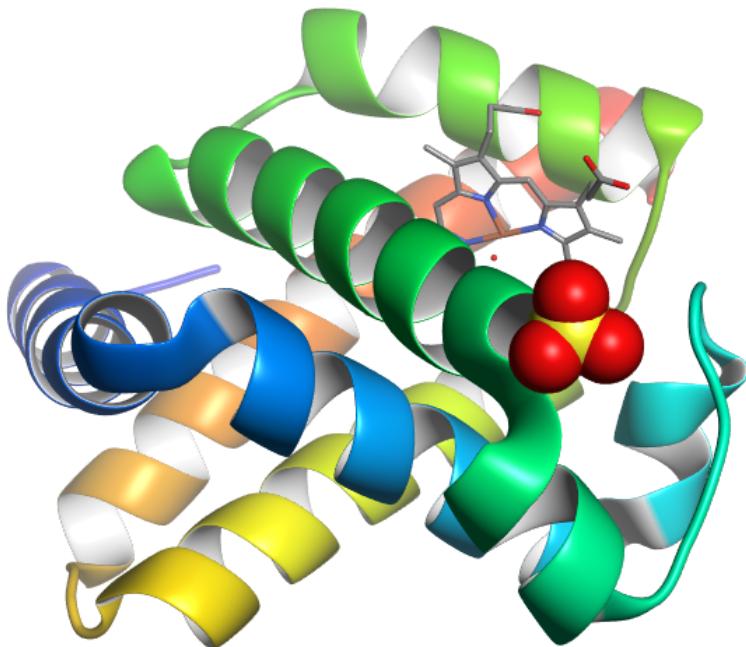
統合データベース講習会 : AJACS徳島

2019-06-06

大塚製薬株式会社 徳島研究所

鈴木 博文 (早稲田大学 先進理工)

どんなデータの話？



⇒ タンパク質などの生体分子の立体構造のデータ

構成原子の座標(X,Y,Z)と付随情報
(図: PDB-3vau ミオグロビン)

自己紹介

- 氏名: 鈴木 博文
- 研究のバックグラウンド
 - 電子顕微鏡を使った立体構造解析
- 2018年9月まで
 - 大阪大学 蛋白質研究所に所属
 - PDBjの開発スタッフ
 - 過去、 AJACS講習会の講師を数回
- 現在
 - 早稲田大学 先進理工
 - AMED-BINDS(創薬等先端技術支援基盤プラットフォーム)の雇用
 - バイオインフォマティクス・データベース
 - PDBjの外部スタッフ

この演習・この資料について

- 想定する聴衆
 - 構造生物学についての **専門外** ~ **初学者**
- 演習の形式
 - キーワード検索と構造ビューア操作が主
 - 解析などは行わない
 - 人数が多いので、動作の保証はできない
- この資料のアイコン
 - : 実施する手順 / : 確認すること
 - (うまく活用できていない)

PDBjを使ってみる

1. PDBjを使ってみる 
2. PDBの紹介
3. 「万見」で詳しく構造を見る
4. 遺伝子名で検索
5. 補足情報

PDBjトップページ

- 「pdbj」をWeb検索、あるいは[pdbj.org]とアドレスバーに入力
- ページを確認
 - ウインドウの大きさによって表示が異なる
 - 左メニューが表示されていない場合は、左上にメニューボタン≡
- 上部に検索ボックス

The screenshot shows the PDBj homepage with the following details:

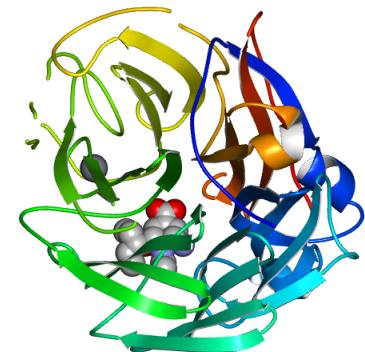
- Header:** PDBj 141415 (エントリー 2018-06-20)
- Search Bar:** pdbj.org 全体を検索 (日本語OK) with a magnifying glass icon.
- Navigation:** English, 日本語, 简体中文, 繁體中文, 한국어; links to JST-NBDC, 大阪大学蛋白質研究所, RCSB, BMRB, PDBe, Adv. Search, Search help.
- Content Area:**
 - 初めての利用者のためのガイド:** A section explaining how to use the site for beginners, mentioning the PDBj入門 link.
 - 必要なサービスを探す:** A section for searching services, with a note about using the キーワードボックス input field.
- Right Sidebar:**
 - 今月の分子:** 222: タンパク質とナノ粒子 (Proteins and Nanoparticles) with a small thumbnail image.
 - WORLDWIDE PROTEIN DATA BANK** logo.
 - BMRB検索** button.
 - パートナー:** IBI portal logo.
- Top Right:** 6EMX molecular structure visualization and 最新公開データベース (Latest published databases) link.

ページの構成

- ① 「初めての利用者のためのガイド」
 - 初心者向けへのガイド
 - 「対話型チュートリアル」
 - 「PDBj入門」 非専門家向けサービス
- ② 「必要なサービスを探す」
 - サービスを探すためのサービス
(たくさんのサービスがあるので)
 - キーワードか、カテゴリで検索

1つ目の主題「タミフル」

- ここではインフルエンザ治療薬・タミフルを主題とする
- 「タミフル」(成分名 オセルタミビル)は構造生物学の成果とされる
- インフルエンザウイルスの「ノイラミニダーゼ」というタンパク質に作用
- 今回見るデータ: PDB-IDが「3cl0」のエントリ
 - オセルタミビル・ノイラミニダーゼ複合体の構造
 - タミフル耐性変異 (リレンザ非耐性)、N1型のインフルエンザウイルス由来
 - 論文はnatureに掲載



検索してみる

- ① 画面上部の検索ボックスに「オセルタミビル」と入力
(実際には、データは英語のみなので、英語での検索のほうが確実)
- ② 検索結果を見てみる
 - ヒットしたエントリを見てみる
 - 「PDB」というタブ以外の内容も見てみる
- ③ タブ[ウェブページ]の「113: インフルエンザノイラミニダーゼ」をクリック



PDB: 1159 件 ウェブページ: 17 件 ステータス検索: 55 件 化合物検索: 0 件

インフルエンザ
変換クエリ: (influenza | grippe)

3BT6
↓ 3D

CRYSTAL STRUCTURE OF INFLUENZA B VIRUS HEMAGGLUTININ

分子名称:	2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE, Influenza B virus hemagglutinin, N-acetyl-D-glucosamine, ...
著者:	Wang, Q., Cheng, E., Lu, M., Tian, X., Ma, J., ...
登録日:	2007-12-27
公開日:	2008-05-27
最終更新日:	2011-07-13
実験手法:	X-RAY DIFFRACTION (2.8 Å)
主引用文献:	Crystal structure of unliganded influenza B virus hemagglutinin. J.Virology, 82, 2008

2RFT CRYSTAL STRUCTURE OF INFLUENZA B VIRUS

ノイラミニダーゼの解説ページ

113: インフルエンザ ノイラミニダーゼ (Influenza Neuraminidase)

このページはRCSB [の](#) David S. Goodsell博士による「[Molecule of the Month](#)」2009年5月の記事を日本語に訳したもので、転載・引用については[利用規約](#)をご覧下さい。

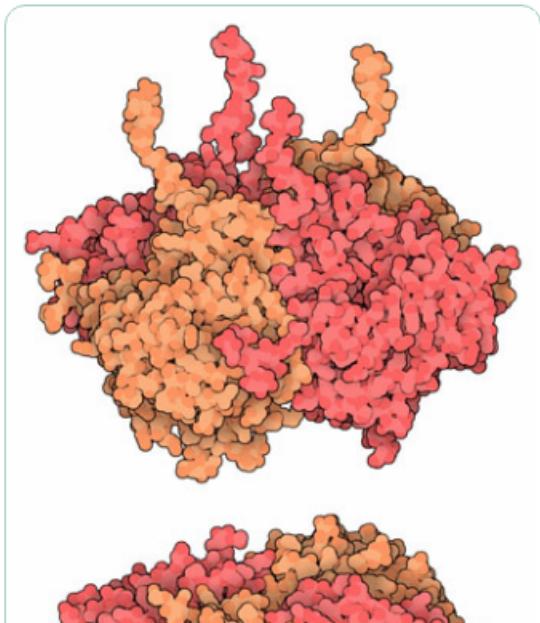
[「今月の分子」一覧に戻る](#) / [この記事のRCSBオリジナルサイト（英語）を見る](#)

: 翻訳 工藤高裕 (PDBj)

インフルエンザウイルス(influenza virus)は継続的に変化し、10年程度ごとに危険な新種が現れて公衆衛生を脅威にさらす。今年、ブタインフルエンザ(swine flu)としてより知られるH1N1型インフルエンザの新種が流行している。H1N1という名前はウイルス表面を覆う2つの分子～赤血球凝集素(ヘマグルチニン、hemagglutinin)とノイラミニダーゼ(neuraminidase)～を意味している。これら2つの分子が協力してウイルスの感染を制御している。赤血球凝集素はウイルスが細胞に侵入する時に中心的な役割を果たしており、細胞表面にある多糖鎖に結合してウイルスゲノムを細胞内へと注入する。一方、ノイラミニダーゼはウイルスが感染した細胞から離れる時に重要な役割を果たす。ノイラミニダーゼはこの糖鎖の末端を切断することによって細胞表面に引っかかってしまわないようにしている。

糖の切断

右図上部に示したのはノイラミニダーゼ(PDBエントリー [1nn2](#))で、4つの同じサブユニットが四角く並んでいる。ウイルス表面は通常長いタンパク質の柄でつながっているが、ここに示した構造ではその部分が取り除かれている。なぜならその部分は非常に柔軟性が高く結晶構造が決定できないからである。活性部位は上側表面の深い窪みの中にある。ノイラミニダーゼは多糖鎖に結合し、末端の糖を切り取る。ノイラミニダーゼの表面はいくつかの多糖鎖で修飾されている



- 「今月の分子」のノイラミニダーゼの記事
 - ノイラミニダーゼ: 宿主細胞の糖鎖を切断する酵素
 - オセルタミビルなど: 酵素の活性部をブロックしてウイルスの増殖を抑制
- 米国 RCSB-PDBの David S. Goodsellによる「Molecule of the Month」の日本語訳

ノイラミニダーゼのPDBデータを見てみる

- ① ページ上部の検索ボックスに「3cl0(さん・シー・エル・ゼロ)」を入力
- ② PDBエントリ「3cl0」が表示される
 - 上部にタブがある、[概要]タブが選択されている
- ③ このデータの概要を見てみる

概要 構造情報 実験情報 機能情報 相同蛋白質 履歴 ダウンロード

3CL0

N1 Neuraminidase H274Y + oseltamivir

3CL0 の概要

関連するPDBエントリー [3CKZ](#) [3CL2](#)

分子名称	Neuraminidase, (3R,4R,5S)-4-(acetylamino)-5-amino-3-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid, CALCIUM ION, ... (4 entities in total)
機能のキーワード	n1, neuraminidase, h274y, oseltamivir, glycosidase, hydrolase, membrane, transmembrane, virion, viral protein
由来する生物種	Influenza A virus
ポリマー鎖数	1

分子構造ビューアを使ってみる

① 画面右側の分子構造の画像(「非対称単位」の下)をクリック

→ Molmilという分子構造ビューアが表示される

② 構造ビューアを操作してみる

- マウスのドラッグで分子の回転、ホイール回転(または右ボタンドラッグ)で拡大縮小、など

③ 分子の表示を変えてみる

- 上部の[表示様式]と[配色]のメニューを操作し、表示様式を変更してみる

④ 「生物学的単位」を表示してみる

- 元のウインドウの「生物学的単位」の下の画像をクリック
- Molmilのウインドウの中身が切り替わる
- このページでも、表示の切り替えが可能



「構造情報」タブの内容

① 元のウインドウ、構造情報タブをクリック

② 「エンティティ」という情報(構成要素の情報が表示されている)

- このデータの場合は、ノイライニダーゼ、長い名前の化合物、カルシウムイオン、水の項目がある
- 一つのデータエントリに複数の要素が含まれる、タンパク質以外の要素も

エンティティ	
鎖名	A
説明	Neuraminidase
種類	polymer
鎖長	385
分子量	42209.9
分子数	1
	UniProt (Q6DPL2) Pfam (PF00064)
	Influenza A virus
エンティティの一般名	
鎖名	
説明	(3R,4R,5S)-4-(acetylamino)-5-amino-3-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid
種類	non-polymer
鎖長	
分子量	284.4
分子数	1
	G39
エンティティの一般名	
鎖名	

③ 「配列ビューア」を見てみる

- アミノ酸配列と2次構造、結合部位の情報など

実験情報タブの情報

① [実験情報]タブをクリック

② 実験情報を見てみる

- X-RAY DIFFRACTION (X線回折)という方法で解析されたデータである
- このPDBというデータベースは実験結果のデータベース
 - 同じタンパク質・同じ試料でも、別の実験から得られたデータは別エントリ

精密化の統計情報			
実験手法:			X-RAY DIFFRACTION (2.2 Å)
格子定数 [Å]	115.066	115.066	63.321
格子定数 [度]	90.00	90.00	90.00
空間群			P 4 21 2
分解能 [Å] (低 - 高)	20.00	- 2.20	
最も高い分解能シェルの値			2.256 - 2.200
R因子	0.19561		
R-work	0.19323		
最も高い分解能シェルの値			0.272
R-free	0.24030		
最も高い分解能シェルの値			0.343
結合長の平均二乗偏差 (RMSD) [Å]	0.009		
結合角の平均二乗偏差 (RMSD) [度]	1.152		
回折データの統計情報			

その他のPDBjのサービス - 「PDBj入門」

- <https://numon.pdbj.org/>
- 教材などへのリンク集
- PDBjがアウトリーチ活動で利用しているサービスなど
 - スマートフォンでVRが見られるサービス
 - 赤青メガネで立体視
- 各種ドキュメント
- 「今月の分子」

PDBj の生体高分子学習ポータルサイト
PDBj入門

日本語 English

PDBj入門はPDBj(日本蛋白質構造データバンク)の一般向けサイトです。タンパク質などの生体分子について理解するためのサービスを提供しています。研究者向けサイトはこちら。

タンパク質とPDB
タンパク質とPDBについての紹介

VR分子ビュア
VR(仮想現実)で分子を見る

万見プライム
生体分子を立体的にいろんな角度から見て学ぶ

ペーパーモデル
紙を組み立ててタンパク質分子を作る

eProtS
PDBに登録されている生体高分子を取り上げ解説している蛋白質構造百科事典

今月の分子
月替わりの分子構造紹介記事

各種資料
過去の講習会などの資料

ゲーム
神經衰弱とスネークゲーム

[利用規約・プライバシーポリシー](#) [お問い合わせ](#)

その他のPDBjのサービス - PDBj Mine 詳細検索

- <https://pdbj.org/advanced-search>
- 多様な条件での絞り込み検索可能

Mine: 詳細条件検索

PDBID:

キーワード:

タイトル:

公開日:

 以降: - - 

 以前: - - 

登録日:

 以降: - - 

 以前: - - 

文献著者:

分解能: -

その他のPDBjのサービス - 更に高度な検索

- <https://pdbj.org/mine>
- **PDBj Mine**: PDBj検索エンジンの総称
- **XPath Search**: XML構造に基づく検索
- **CIF Query Service**: 関係モデルに基づくデータ構造を利用した検索
- **SQL検索**: SQL文で検索

The screenshot shows the 'PDBj Mine (PDBエントリーの検索)' interface. At the top is a search bar labeled 'Search PDBj Mine' with a magnifying glass icon. Below it is a section titled 'Mine quick search' with several buttons: '新規公開されたエントリー' (Newly published entry), '更新されたエントリー' (Updated entry), '最新の廃止/置換されたエントリー' (Latest withdrawn/replaced entry), '全ての廃止/置換されたエントリー' (All withdrawn/replaced entries), and '巨大構造エントリー' (Large structure entry). Further down is a 'Specialized search' section with buttons for '詳細条件検索' (Advanced search), 'スニータス検索' (SNTAS search), and 'Chemie search (化合物の検索)' (Chemie search (Compound search)). At the bottom is a 'More information' link.

PDBの紹介

1. PDBjを使ってみる
2. PDBの紹介 
3. 「万見」で詳しく構造を見る
4. 遺伝子名で検索
5. 補足情報

2つの側面

PDBは...

世界でたったひとつのデータベース

- 厳重に秘匿
- 厳格に管理

世界中にたくさんあるデータベース

- 容易に入手
- 自由に利用
- 勝手に公開

「用語」についてよもやま話 - Gmailの例

- 「Gmail」という名称が重複
 - サービス名
 - アプリ(メールクライアント)名
- それも良し悪し
 - 日常の利用では便利、覚えやすい
 - ちょっとマニアックな話をするときは混乱を招く
 - 例 「GmailでOutlookが使えるし、OutlookでGmailが使えるよ」

いろいろな「PDB」

実態	名称
データベース名 (アーカイブ名)	PDB (Protein Data Bank)
組織名	wwPDB, PDBj, RCSB-PDB, PDBe
サービス名	PDBjなど
ファイル形式 (旧・公式フォーマット)	PDB形式 (拡張子:ent/pdbなど)

- 現在はPDBx/mmCIFという形式が公式フォーマット

「用語」についてよもやま話 - 「データベース」

文脈で意味が変わる

- 辞書的には「整理されたデータの集まり」
- 検索などができるインターネット上のサービス
- 狹義では「データベース管理システム」を指すことも

データベースの種類

- 外部からのデータ登録で成り立っているデータベース運営を「**データバンク**」と呼ぶことも
- 検索できない、単なるファイルの集まりは「**データアーカイブ**」と呼ぶことも
- 一次データベース、派生データベース、統合データベースなどいろいろ
 - GenBank、PubMedで有名なNCBIのサイト、たくさんのデータベースがある
 - structureというデータベースがある <- 実体はPDB (+α)

PDBとは

ログインしていません トーク 投稿記録 アカウント作成 ログイン

ページ ノート 閲覧 編集 履歴表示 Wikipedia内を検索 検索



ウィキペディア
フリー百科事典

メインページ
コミュニティ・ポータル
最近の出来事
新しいページ
最近の更新
おまかせ表示
練習用ページ
アップロード (ウィキメディア・コモンズ)

ヘルプ
ヘルプ
井戸端
お知らせ
バグの報告
寄付
ウィキペディアに関するお問い合わせ

蛋白質構造データバンク

蛋白質構造データバンク (たんぱくしつこうぞうデータバンク、PDB; Protein Data Bank) は、蛋白質（タンパク質）と核酸の3次元構造の構造座標（立体配座）を蓄積している国際的な公共のデータベースである。PDBに蓄積されている構造データは、X線結晶解析法、NMR法（核磁気共鳴法）などによって実験的に決定されたデータである。なお、理論的な予測（蛋白質構造予測）で推定されたデータは蓄積していない。

世界中の生物学者や生化学者たちが、PDBに構造データを登録する。PDBに登録されたデータはパブリックドメインのもとで公開され、誰もが無償でアクセスすることができる。日本では大阪大学蛋白質研究所にその支所がある。

PDBは、生物学的構造データの中心的なデータベースである。構造生物学の研究で欠かせない情報源であり、また近年では構造ゲノミクスの研究でも重要なデータベースである。バイオインフォマティクスの研究でも、PDBに代表される3次元分子構造データベースは重要な研究対象である。PDBから派生したデータベースとプロジェクトは非常に多く、蛋白質の構造、機能、進化のそれぞれの側面から、PDBの構造データの統合や分類を行っている。

目次 [非表示]

- 1 歴史
 - 1.1 構造データの増加
- 2 構造データの内容
 - 2.1 統計情報

PDBとは

PDBとは

- (英)Protein Data Bank
- (日)蛋白質構造データバンク
- 「生物学的構造データの中心的なデータベース」

「Protein Data Bank」という名前だけど...

- タンパク質の情報を網羅的に集めたものではない
- タンパク質だけではない、核酸、糖鎖など
- 生体分子の立体構造(原子座標)のデータベース
 - 実験により測定された立体構造
 - 同じタンパク質でも別の実験のデータは別エントリ
 - 理論のみの構造モデルは受け入れていない

PDBの歴史

年	出来事	組織
1971	設立	ブラックヘブン国立研究所(米)
1998	移管	RCSB(米)
2003	wwPDB設立	RCSB-PDB/BMRB(米), PDBe(欧), PDBj(日)

- 非常に長い歴史
 - インターネット・データベースの「黎明期」から
- 運営組織の移管を経ている
- 国際的な運営組織、日米欧にメンバー
 - wwPDB ワールドワイドPDB
 - PDBj 日本のPDB運営組織、wwPDBのメンバー

データの受け入れ (登録・検証・品質管理)

- 世界中の研究者から集める
- 登録受付・検証・品質管理
 - PDBの3拠点で**分担**
 - 専任のアノテーターによる、共通ルールでの品質管理と注釈付
 - 全拠点で**共通のデータアーカイブ**を作成
- 新規データは**非公開**、公開日まで厳重に管理
 - 競争が激しい研究テーマが多い



PDBへの登録が義務

nature誌が論文を受理する条件

PDB登録データベース	
Macromolecular structure	Worldwide Protein Data Bank (wwPDB)
	Biological Magnetic Resonance Data Bank (BMRB)
	Electron Microscopy Data Bank (EMDB)

<http://www.nature.com/authors/policies/availability.html> より

他の多くの論文ジャーナルも同様の条件

→ PDBにデータを登録しないと、論文発表できない

wwPDBの「諮問委員会」



wwPDB Advisory Committee Meetings 2014

メンバーは世界中の構造解析・バイオインフォマティクスの専門家

<http://www.wwpdb.org/about/advisory>

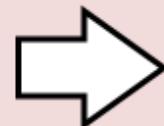
データの公開

- 世界同時公開
 - 毎週水曜0:00 (日本時間9:00)
- 公開後は**パブリックドメイン**
 - 使途は問わない、人類の財産
- 公開データ
 - 世界共通のデータアーカイブ
 - 各拠点独自のツール・独自のサービスを通じて公開





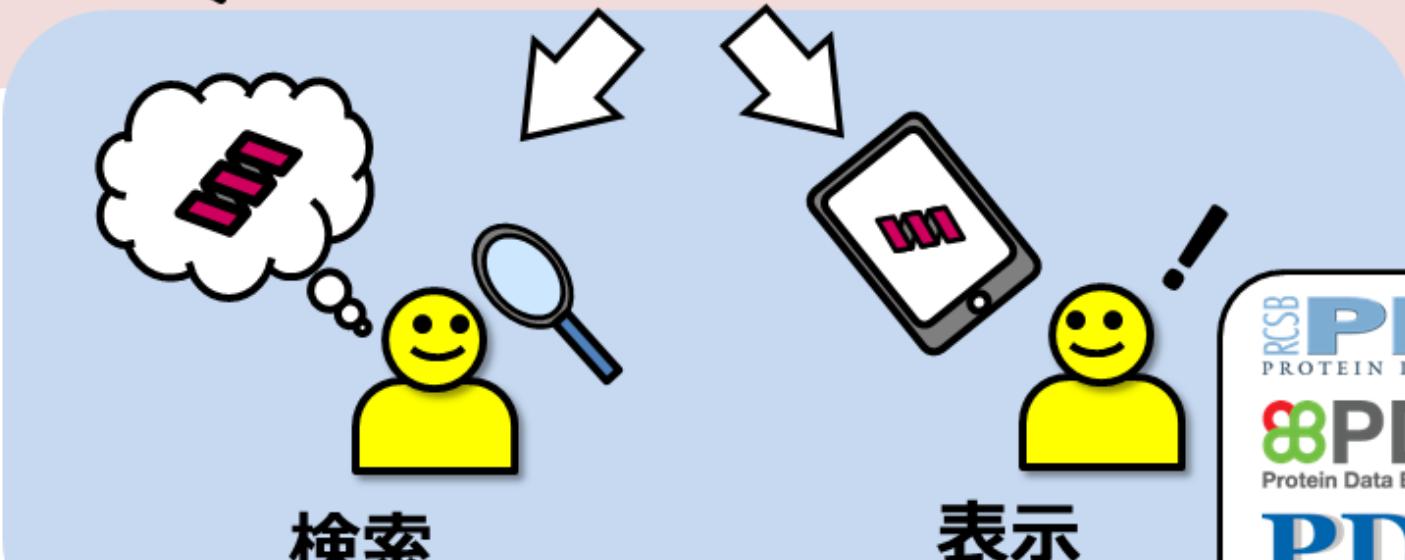
登録



PDB, EMDB,
& BMRB



公式アーカイブ
の公開



検索

表示



2つの側面

PDBは...

世界でたったひとつのデータベース

- 厳重に秘匿
- 厳格に管理

世界中にたくさんあるデータベース

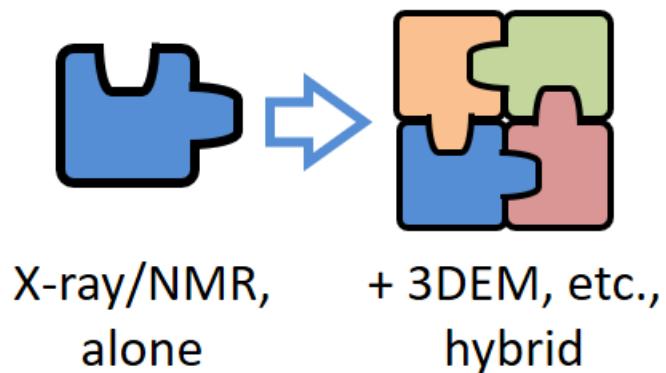
- 容易に入手
- 自由に利用
- 勝手に公開

3. 「万見」で詳しく構造を見る

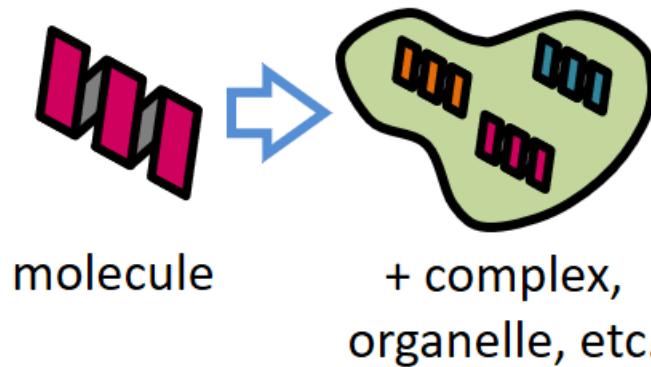
1. PDBjを使ってみる
2. PDBの紹介
3. 「万見」で詳しく構造を見る 
4. 遺伝子名で検索
5. 補足情報

Multi-things going multi

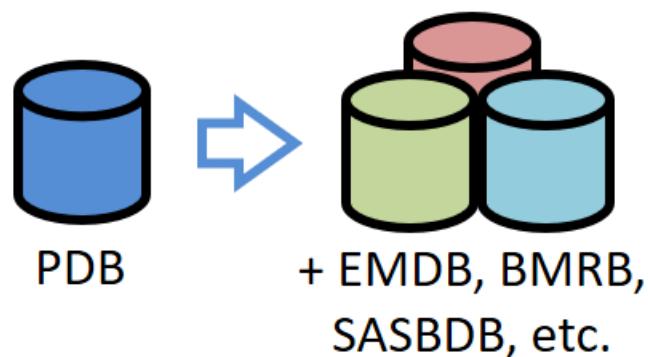
Methodology



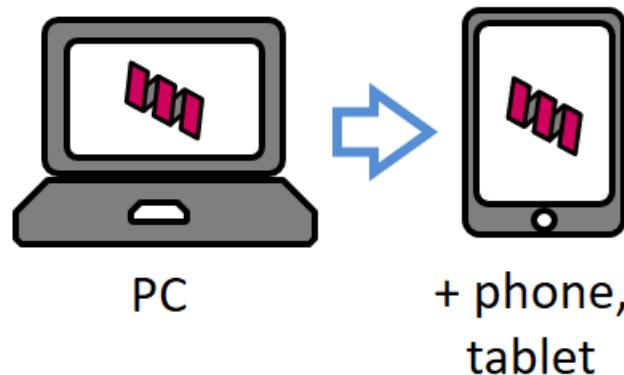
Scale



Databank

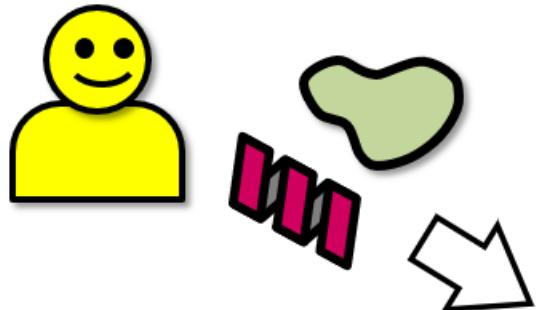


Browsing device

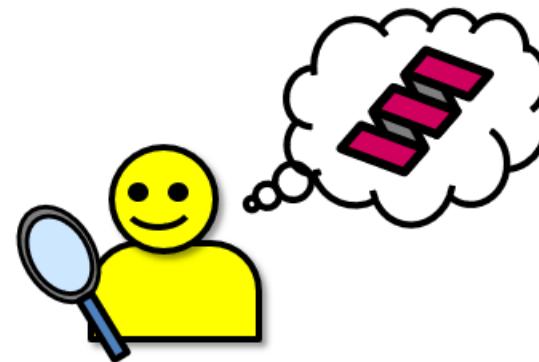


One for the multi

Deposition

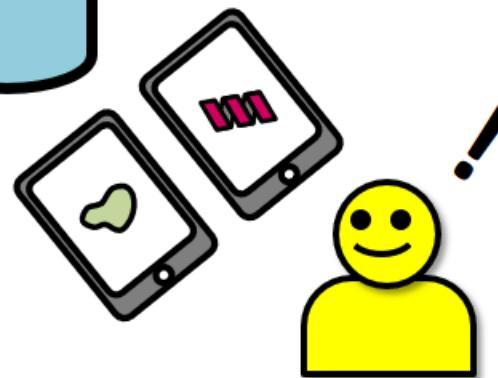


Search & browse

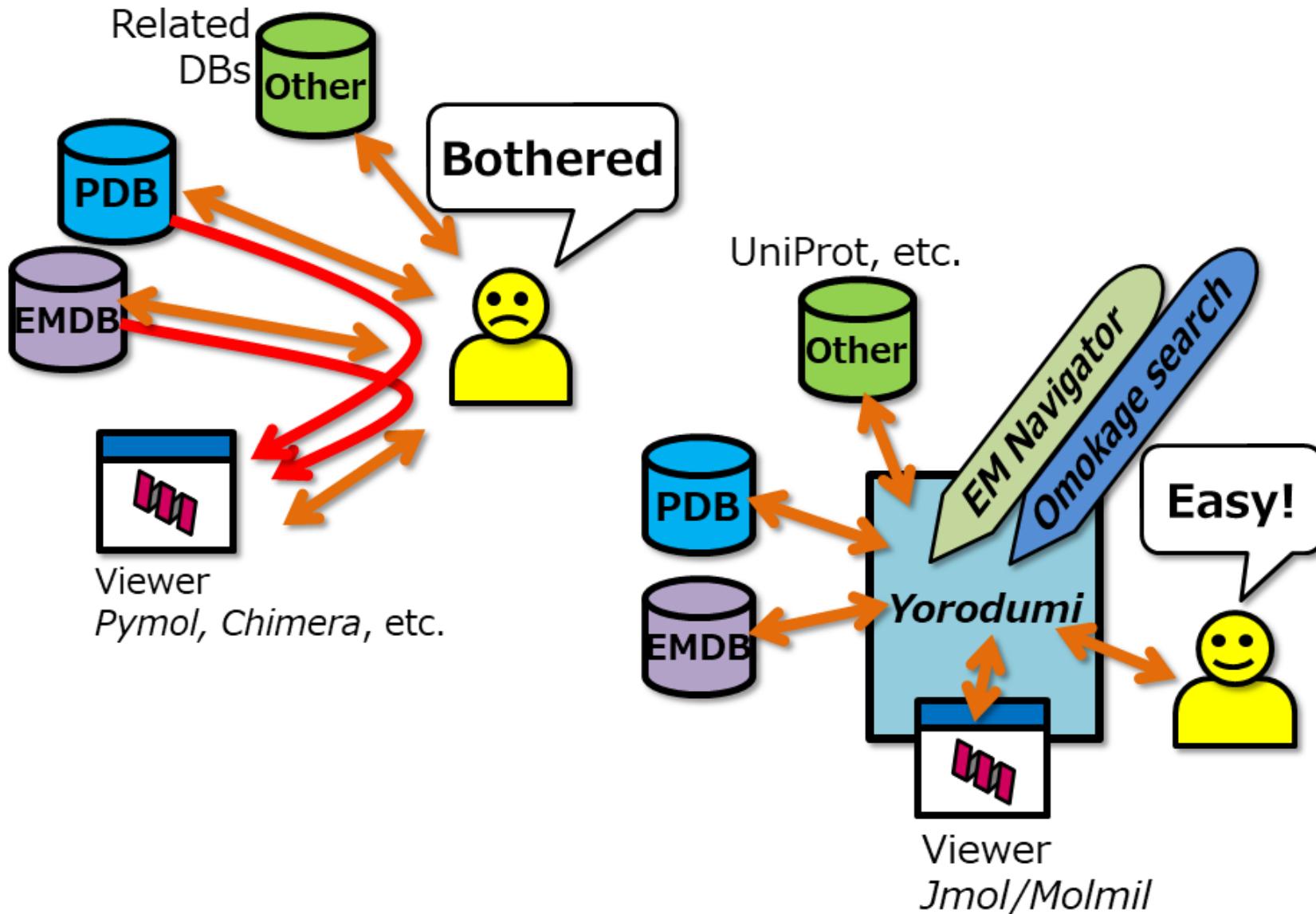


Comparison

Structure viewing



「万見」とは



「万見」でオセルタミビルのエントリを見てみる

① 「万見」のページを開く

- PDBエントリ3cl0のページに戻る
 - 閉じてしまった場合は、アドレスバーに "pdbj.org/pdb/3cl0"
- ページ右側の[他のデータベース情報]のボックスの中の3つめ、[Yorodumi]をクリック

万見 - PDB-3cl0: N1 Neuraminidase H274Y + oseltamivir -

+ データを開く
- 基本情報

エントリ情報	 データベース: PDB / ID: 3cl0 構造の表示 ダウンロードリンク
タイトル	N1 Neuraminidase H274Y + oseltamivir
構成要素	Neuraminidase 詳細
キーワード	VIRAL PROTEIN / HYDROLASE / N1 / neuraminidase / H274Y / oseltamivir / Glycosidase / Hydrolase / Membrane / Transmembrane / Virion (ウイルス)
機能・相同性	Sialidase, Influenza viruses A/B / Glycoside hydrolase, family 34 / Sialidase superfamily / Neuraminidase / exo-a-sialidase / exo-alpha-(2->3)-sialidase activity / exo-alpha-(2->6)-sialidase activity / exo-alpha-(2->8)-sialidase activity / viral budding from plasma membrane / host cell plasma membrane



万見の3cl0のページ

「生物学的」集合体(構造単位)の確認

① 画面を「集合体」の部分まで下にスクロール

② 各種集合体の詳しい情報が表示される

- 「登録構造単位」と「1」の情報を確認

③ [ビューアーで表示]ボタンをクリックするとMolmilで表示される

- 双方の構造を表示してみる

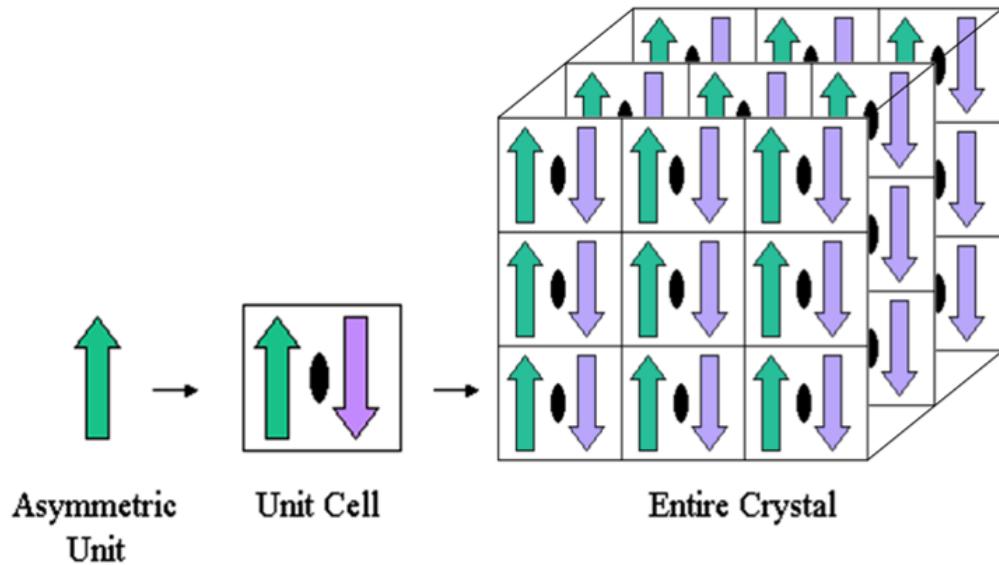
The screenshot shows the PDB database interface. At the top, there's a search bar with the text '集合体' (Assembly). Below it, a navigation bar has tabs for '概要' (Summary) and '構成要素の詳細' (Detailed components). A 3D molecular model is displayed with a green ribbon representation. Below the model, there are two sections:

- 登録構造単位**: Shows a summary for entry '1'. It includes a 'ビューアーで表示' (View in viewer) button, a 'Length a, b, c (Å)' value of 115.066, 115.066, 63.321, and a 'Space group name H-M' value of P 4 21 2.
- 1**: Shows detailed information for entry '1'. It includes a 'ビューアーで表示' (View in viewer) button, tabs for '概要', '構成要素の詳細', '対称操作', and '計算値', and a 'Length a, b, c (Å)' value of 115.066, 115.066, 63.321. It also features a '構造データをダウンロード' (Download structure data) button and a 'Omokage検索' (Search) link.

At the bottom right, there's a table with the following data:

Length a, b, c (Å)	115.066, 115.066, 63.321
Angle α, β, γ (deg.)	90.00, 90.00, 90.00
Int Tables number	90
Space group name H-M	P 4 21 2

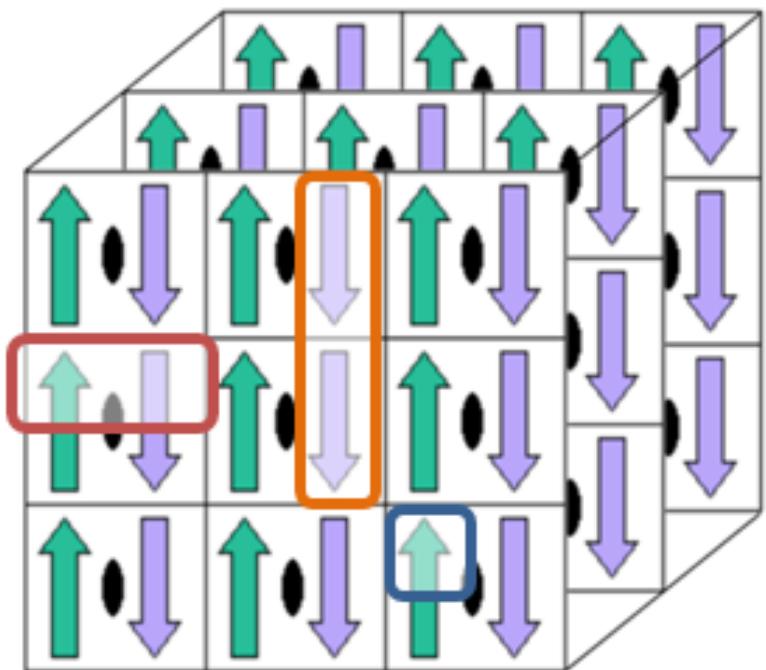
非対称単位



- PDBに登録されているデータのほとんどは結晶構造
- 結晶構造では非対称単位 (Asymmetric unit)の原子配置とその対称性情報が登録されている
- らせん対称や、正20面体型対称など、結晶学以外の非対称単位もある

生物学的単位（生物学的集合体）

- 生体内での構造や溶液構造に近いと考えられる構造単位を、結晶構造から切り出したもの
- データ登録者が定義したものと、計算によって見積もられたものが登録されている
- 非対称単位との関係（大きい・小さい）は、データごとに異なる



構成要素を確認

画面を「構成要素」の部分まで

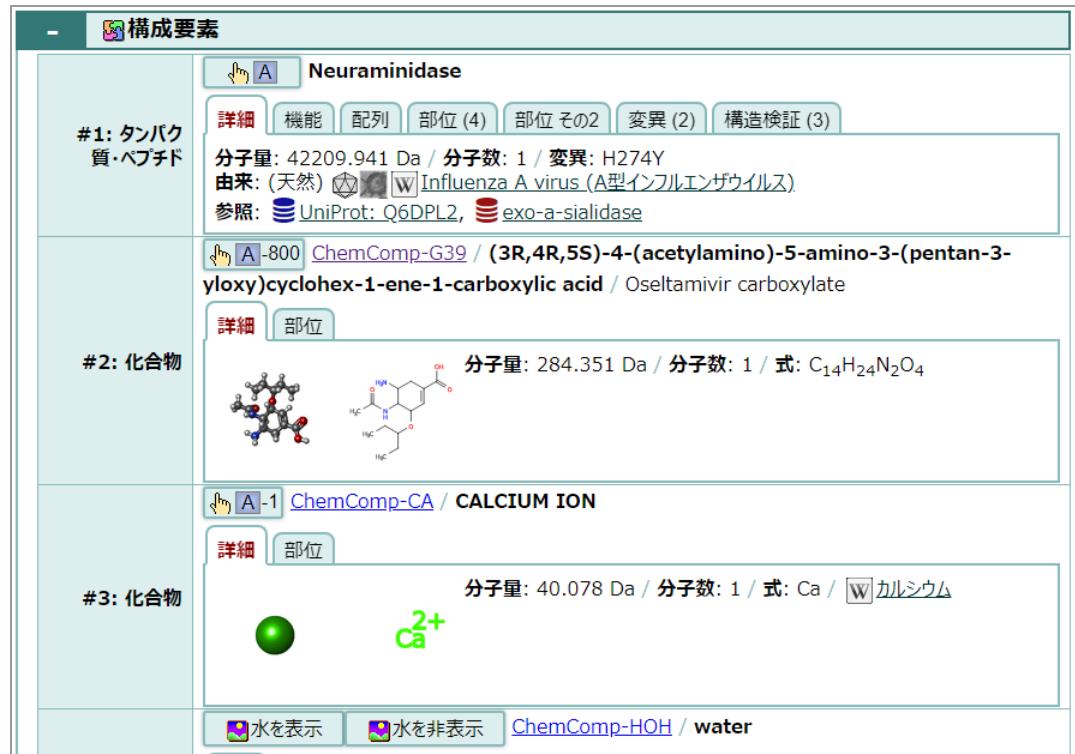
下にスクロール

各種構成要素の情報を確認

- 要素#2がオセルタミビル、「G39」というIDを持つ化合物

各要素の選択ボタンを押して、位置を確認

- 指アイコンのついたボタンは、ビューア上でその要素を選択するボタン
- 「選択」された要素は黄色く光る



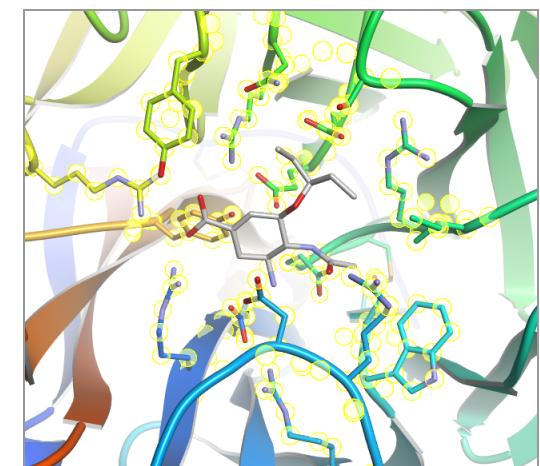
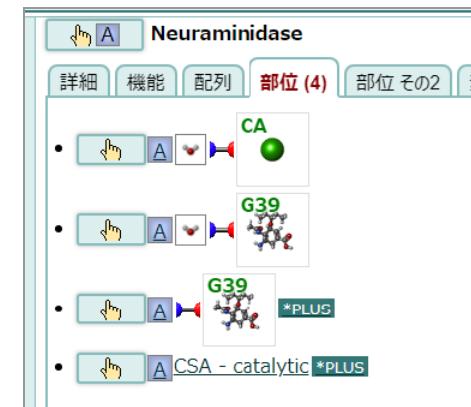
オセルタミビルの結合部位

① ノイラミニダーゼ(Neuraminidase)の[配列]タブをクリック

- 配列ビューアが表示される
- 配列の一部をマウスで選択すると、対応する部位がビューアでフォーカスされる

② ノイラミニダーゼの結合部位を観察

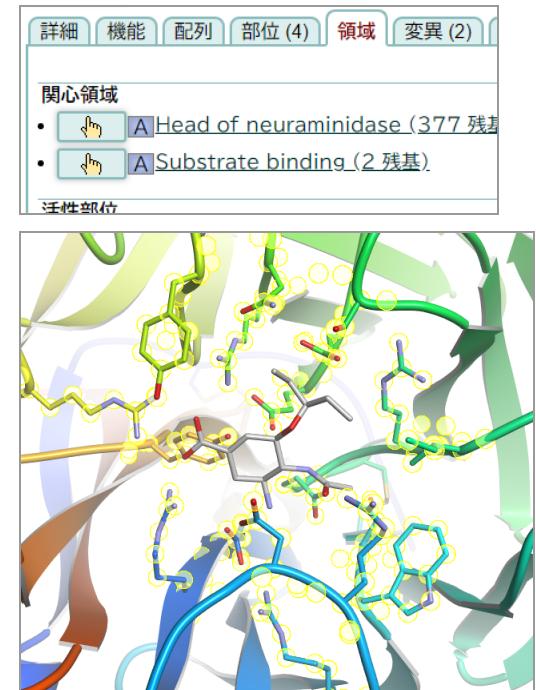
- ノイラミニダーゼの[部位]タブをクリック
(PDBデータに登録されている活性部位などの情報が表示されるタブ)
- G39の選択ボタンをクリック → 結合部位の残基の側鎖が表示される
- ビューア中の原子を右クリックすると、原子の情報とスタイルのコントロールメニューが表示される



その他、配列領域情報を確認

④ その他の「領域」を観察
(UniProtデータから抽出した情報)

- ノイラミニダーゼの[領域]タブをクリック
- 各種「領域」の選択ボタンをクリックし構造ビューアで確認
 - 活性部位 - Proton donor/acceptor や Nucleophile
- 要素名の文字列をクリックすると、その根拠となるデータへのリンク
- 他のタンパク質では、ドメイン情報なども表示される場合もある



変異残基を確認

⌚ タミフル耐性を引き起こした変異残基を確認

- 要素ノイラミニダーゼの[変異]タブをクリック
2つの変異が登録されているが、上はCONFLICT(配列データベースとの相違)
- 「設計変異」と書かれている残基274の選択ボタンをクリック

👀 構造ビューアで位置を確認

- 変異残基は結合残基には含まれていない、オセルタミビルから少し離れている
- 加水分解活性を失わず、オセルタミビル結合能を下げる変異

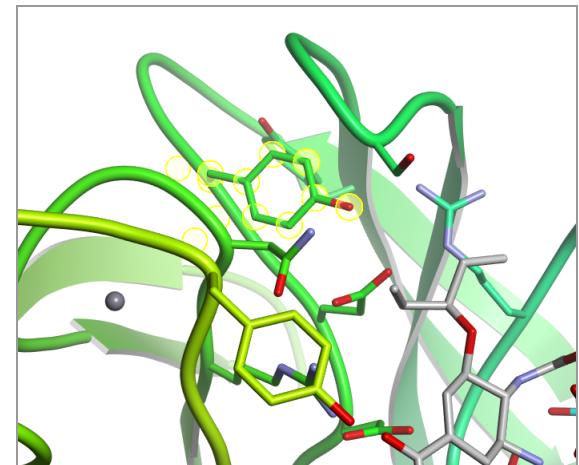
詳細 機能 配列 部位 (4) 部位 その2 変異 (2)

CONFLICT

- A-252: HIS → TYR

ENGINEERED

- A-274: HIS → TYR



更に知りたいときは

- 以前の講習会では、UCSF Chimeraを利用して、リレンザ結合構造・野生型の構造との比較を実施

<https://pdbj.org/workshop/20150613/suzuki.pdf>

4. 遺伝子名で検索

1. PDBjを使ってみる
2. PDBの紹介
3. 「万見」で詳しく構造を見る
4. 遺伝子名で検索 
5. 補足情報

PDBは遺伝子名での検索に向いていない

データの問題

- PDBでは遺伝子名情報は厳格に管理されていない
- 天然試料の場合、遺伝子名は記載されない

サービスの問題

- 近日中に遺伝子検索に対応したサービスを公開する予定
- キーワード検索でヒットするが、ノイズが多い

UniProtで遺伝子名検索

順序としては、UniProtから

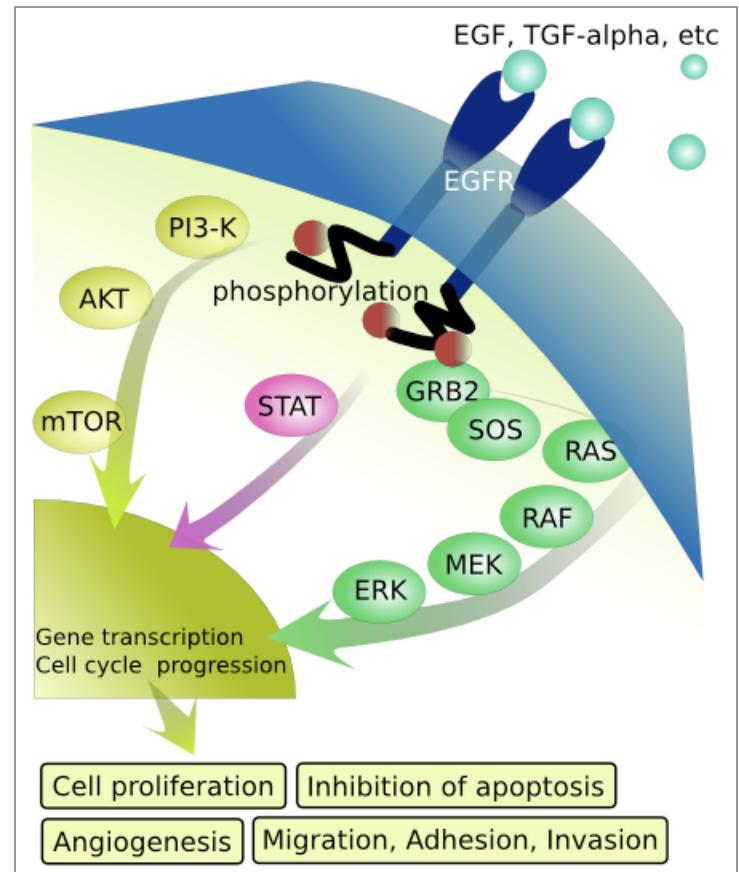
- UniProtはタンパク質情報を網羅
- PDBは「実験データ」のデータベース

ドメイン情報

- ドメイン情報もUniProtで見るのが便利
- **InterPro**: ドメイン、相同性の統合データベース
 - 今回はデータを利用するのみ

主題: ヒトEGFR遺伝子

- 上皮成長因子受容体(Epidermal Growth Factor Receptor)をコード
- 上皮成長因子: 細胞の増殖や成長を制御するシグナル
- EGFR産物は、その受容体
- 細胞膜を貫通、2量体を形成(?)
- EGFR遺伝子の変異が、肺がんなどに関与
- 細胞内にATPase部位、チロシンキナーゼ
- 抗癌剤ゲフィチニブ・アファチニブの標的



ウィキメディア・コモンズより
(パブリックドメイン)

UniProtで遺伝子名検索

① UniProtのサイトを開く (<https://uniprot.org>)

② UniProtでEGFRを検索キーワード検索

- 上部の入力ボックスに「egfr」と入力、



(Advancedボタンをクリックするとより厳密な検索が可能だが、今回は「ゆるい」使い方を実施)

UniProtでの検索結果

00結果を見る

- 検索結果は表敬式で、Gene namesというカラムがある
- EGFRという遺伝子名を持つアイテムが複数みつかる
- ヒト以外のEGFRがある

Filter by:

BLAST Align Download Add to basket Columns < 1 to 25 of 578 >

	Entry	Entry name		Protein names	Gene names	Organism
Reviewed (385) Swiss-Prot	P00533	EGFR_HUMAN		Epidermal growth factor receptor	EGFR ERBB, ERBB1, HER1	Homo sapiens (Human)
Unreviewed (193) TrEMBL	Q96B97	SH3K1_HUMAN		SH3 domain-containing kinase-bindin...	SH3KBP1 CIN85	Homo sapiens (Human)
Popular organisms	Q12802	AKP13_HUMAN		A-kinase anchor protein 13	AKAP13 BRX, HT31, LBC	Homo sapiens (Human)
Human (370)	Q96AW1	VOPP1_HUMAN		Vesicular, overexpressed in cancer,...	VOPP1 ECOP, GASP	Homo sapiens (Human)
Mouse (39)	P04412	EGFR_DROME		Epidermal growth	Egfr c-erbB, DER, top,	Drosophila

ヒト由来のEGFRを探す

④検索ボックスに" human"を追加し検索し直す

⑤結果を確認

- 2つ目のカラムがUniProtのID、3つめのカラムがエントリ名
- P00533、EGFR_HUMANが「それっぽい」ことがわかる

Entry	Entry name	Protein names	Gene names	Organism
P00533	EGFR_HUMAN	Epidermal growth factor receptor	EGFR ERBB, ERBB1, HER1	Homo sapiens (Human)
Q96B97	SH3K1_HUMAN	SH3 domain-containing kinase-bindin...	SH3KBP1 CIN85	Homo sapiens (Human)
Q12802	AKP13_HUMAN	A-kinase anchor protein 13	AKAP13 BRX, HT31, LBC	Homo sapiens (Human)

⑥ IDの部分をクリックし、個別エントリページを開く

ヒト由来のEGFRのページ

① EGFRのページを確認

- 非常に多様な情報
- ページ左にナビゲータークリックすると、その部分へ移動
 - Family & Domainsなどを確認

The screenshot shows the UniProtKB entry page for P00533 (EGFR_HUMAN). The main content area displays the following information:

- Protein:** Epidermal growth factor receptor
- Gene:** EGFR
- Organism:** Homo sapiens (Human)
- Status:** Reviewed - Annotation score: 5/5 - Experimental evidence at protein level¹

On the left, there is a sidebar titled "Display" with several tabs:

- Entry (selected)
- Publications
- Feature viewer
- Feature table

Below these tabs is a list of checkboxes for navigating to different protein features:

- Function (selected)
- Names & Taxonomy
- Subcellular location
- Pathology & Biotech
- PTM / Processing
- Expression
- Interaction
- Structure

At the bottom of the sidebar, there are two buttons: "By similarity" and "20 Publications".

The main content area also contains a detailed description of the protein's function and its role in various signaling pathways.

トポロジーを確認

① トポロジーを確認

- ナビゲーター - Subcellular locationをクリック
- 「Topology」というセクションを探す

Topology						
Feature key	Position(s)	Description	Actions	Graphical view	Length	
Topological domain ⁱ	25 – 645	Extracellular  Sequence analysis	 Add  BLAST		621	
Transmembrane ⁱ	646 – 668	Helical  Sequence analysis	 Add  BLAST		23	
Topological domain ⁱ	669 – 1210	Cytoplasmic  Sequence analysis	 Add  BLAST		542	

② 確認

- N末端側 25 – 645 が**細胞外**、C末端側 669-1210 が**細胞内**

構造データへのリンク

構造データを確認

- ナビゲーター - Structureをクリック
- 多数の構造データがあることがわかる

Structure



PDB Entry	Method	Resolution	Chain	Positions	Links
3W33	X-ray	1.70 Å	A	696-102	PDBsu...
4G5J	X-ray	2.80 Å	A	696-102	PDBe RCSB ... PDBj PDBsu...
4G5P	X-ray	3.17 Å	A/B	696-102	PDBe RCSB ... PDBj PDBsu...
4HJO	X-ray	2.75 Å	A	696-102	PDBe RCSB ... PDBj PDBsu...
4I1Z	X-ray	3.00 Å	A	695-102	PDBe RCSB ... PDBj PDBsu...
4I20	X-ray	3.34 Å	A	695-102	PDBe

- C末端配列側の構造、PDBエントリ「4g5p」を探し、PDBjへのリンクをクリック

目的のタンパク質の構造データがない場合は？

Search by PDBID Search by sequence

Sequence navigator

類似配列を持つPDBエントリーを検索します。

Sequence:

Clustering: No clustering



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

PDBjのSequence Navigator

アミノ酸配列でPDBデータに対しBLAST検索実行が可能
類似タンパク質の構造データの検索・配列アライメント

YorodumiでEGFRデータを開く

① 先程と同様の手順でPDBjのページから万見のページへ移動

② 集合体の情報を確認

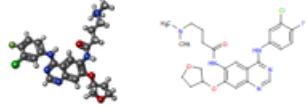
The screenshot shows a protein structure on the left with a blue ribbon model. To the right, there is a list of properties:

- 登録者が定義した集合体
- 38.1 kDa, 1 ポリマー
- [構造データをダウンロード](#)
- [Omokage検索でこの集合体の類似形状データを探す \(詳細\)](#)

- 登録者はモノマーであると主張
- ソフトウェアではダイマーも安定である可能性を示唆

構成要素を確認

⑥ 同様の手順で構成要素の内容を確認

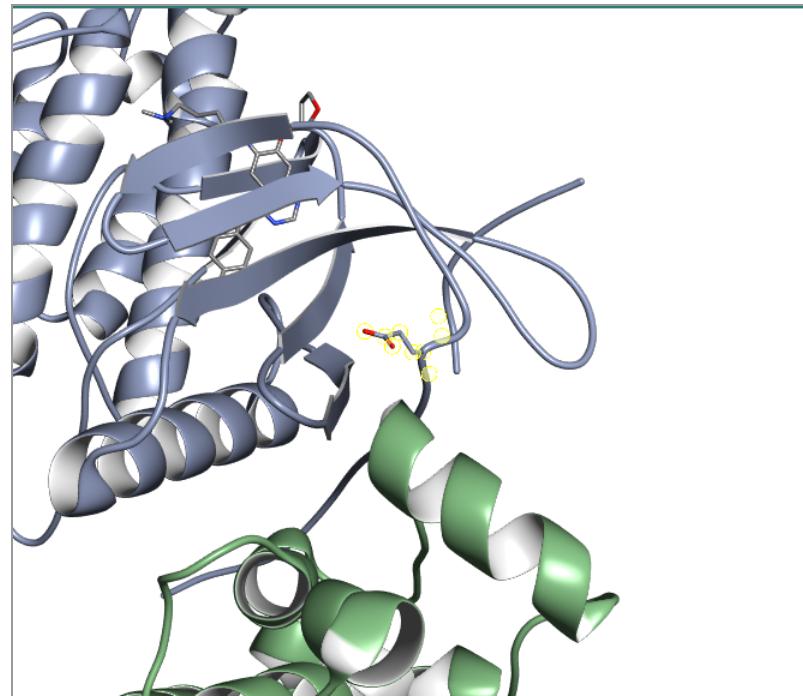
#1: タンパク質・ペプチド	<p>Epidermal growth factor receptor / 上皮成長因子受容体 / Proto-oncogene c-ErbB-1 / Receptor tyrosine-protein kinase erbB-1</p> <p>詳細 機能 配列 部位 (12) 領域 変異 (8) 構造検証 (1)</p> <p>分子量: 37609.547 Da / 分子数: 2 / 断片: Kinase domain, UNP residues 696-1022 / 由来タイプ: 組換発現 / 由来: (組換発現)  Homo sapiens (ヒト) 参照:  UniProt: P00533,  受容体型チロシンキナーゼ</p> <p>JSON: struct_site, struct_site_gen, plus:struct_site, plus:struct_site_gen</p>
#2: 化合物	<p>A-1101 ChemComp-OWN / N-{4-[(3-chloro-4-fluorophenyl)amino]-7-[(3S)-tetrahydrofuran-3-yloxy]quinazolin-6-yl}-4-(dimethylamino)butanamide / Afatinib, bound form</p> <p>詳細 部位</p> <p>分子量: 487.954 Da / 分子数: 1 / 由来タイプ: 合成 / 式: C₂₄H₂₇ClFN₅O₃ / アファチニブ</p> 

- EGFRタンパク質
- アファチニブ（抗がん剤）

変異部位などを確認

👉 先程と同様の手順で領域情報などを確認

- アファチニブはATP結合部位の近傍に
- [領域]タブの[配列バリアント]セクションのところに肺癌試料から見つかった変異が多数
- "Important for dimerization"とされる領域、など



配列バリアント

- A Found in a lung cancer sample; more sensitive to gefitinib; dbSNP:rs397517085. (1 残基)
- A Found in a lung cancer sample; constitutively activated higher levels of basal autophosphorylation; more sensitive to gefitinib; dbSNP:rs397517085. (1 残基)
- A Found in a lung cancer sample; dbSNP:rs727504256. (1 残基)
- A Found in a lung cancer sample; dbSNP:rs121913428. (1 残基)

▽ 残り16件を表示

5. 補足

1. PDBjを使ってみる
2. PDBの紹介
3. 「万見」で詳しく構造を見る
4. 遺伝子名で検索
5. 補足情報 

HOMCOS : 相同複合体の検索・モデリングサーバ

HOMCOS

HOMCOS : 相同複合体の検索・モデリングサーバ

[Go to English page]  ヘルプページ

HOMCOS(HOMology modeling of COnplex Structure) は、PDBに収納されている複合体の立体構造データを利用して、分子の類似性・相同性から、構造未知の分子ペアの構造を予測するためのサーバです。アミノ酸配列や化学構造から、PDB内から複合体の立体構造が解けている他の分子を探査したり、PDB内の構造を鑄型にして複合体立体構造を予測することができます。タンパク質の配列類似性検索にはBLASTを、化合物の構造類似性検索にはKCOMBUを用いています。

サービス	クエリ(問い合わせ)	説明
結合分子の検索	 タンパク質に対する結合分子の検索	アミノ酸配列 1本のアミノ酸配列をクエリとして、それと類似したタンパク質と結合している分子を検索します
	 化合物に対する結合タンパク質の検索	化学構造 1つの化合物構造をクエリとして、それと類似した化合物と結合しているタンパク質を検索します
複合体立体構造のモデリング	 ホモ多量体のモデル	アミノ酸配列 1本のアミノ酸配列をクエリとして、そのホモ多量体の立体構造をホモロジーモデリングによって予測します
	 ヘテロ多量体のモデル	2本のアミノ酸配列 2本のアミノ酸配列をクエリとして、そのヘテロ多量体の立体構造をホモロジーモデリングによって予測します
	 化合物-タンパク質複合体のモデル	アミノ酸配列と化学構造 1本のアミノ酸配列と1つの化学構造をクエリとして、その複合体の立体構造をホモロジーモデリングによって予測します

LastUpdate of PDB:20190116, Release of UniProt:2019_01 [Download data(PDB:20190116)]

<http://homcos.pdbj.org/>

HOMCOSを利用した過去の講習会

1. 構造が未知の目的タンパク質
 2. 相同タンパク質の複合体の検索
 3. 複合体立体構造へ配列アラインメント
 4. リガンド結合部位の予測
- など

川端 猛（大阪大・蛋白研）：

ホモロジー・モデリングと複合体の検索・モデリングサーバHOMCOSの紹介

https://pdbj.org/workshop/20180220/kawabata_20180220.pdf

EM Navigator - 3次元電子顕微鏡データブラウザ

EM Navigator - 3次元電子顕微鏡データナビゲーター -

- 3DEMデータを見る

- キーワード / EMDB-ID / PDB-ID 送信
- 詳細な検索、表形式による表示など
- Omokage検索 - 形状が似ている構造データを探す
- 万見 - EMDB/PDBの付随情報と構造ビューアを統合
- ギャラリー 統計 3DEM文献 生物種

- 最近公開されたデータ

- 2019年5月29日 (EMDB: 46 / PDB: 29)

EM Navigator: 電子顕微鏡構造データの検索・閲覧・統計情報など

<https://pdbj.org/emnavi/>

EM Navigatorを活用した過去の講習会

<https://pdbj.org/news/20180105>

鈴木博文 「クライオ電顕の密度マップと原子モデルの可視化」

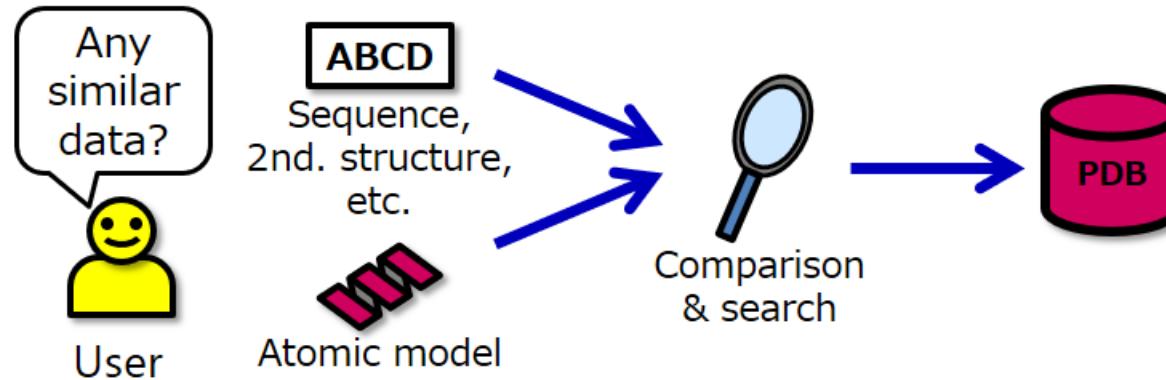
https://pdbj.org/workshop/20180220/suzuki_lecture_20180220.pdf

https://pdbj.org/workshop/20180220/suzuki_exercise_20180220.pdf

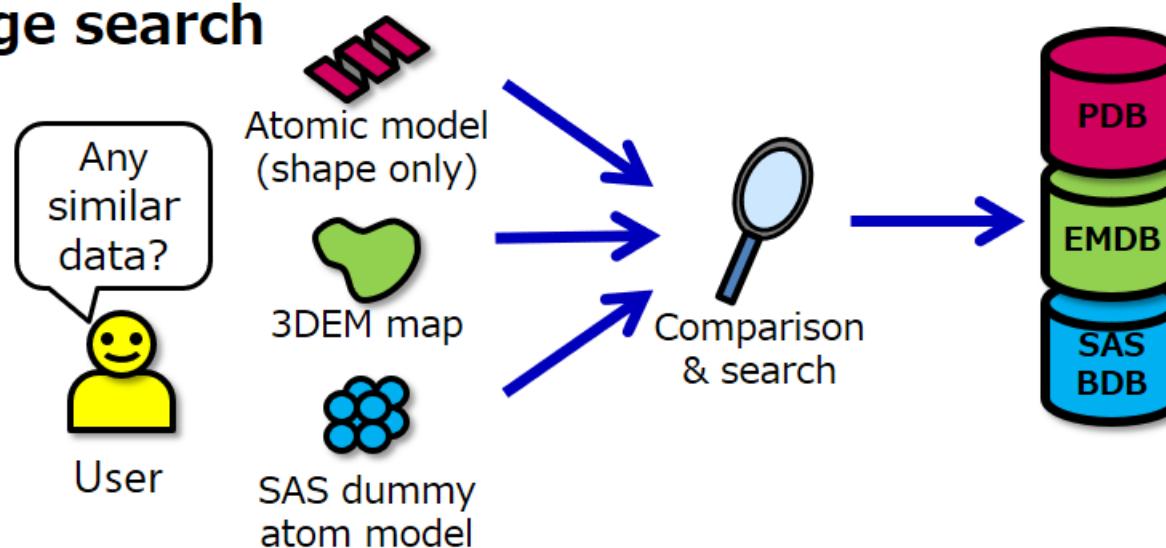
- 本当にリガンドが見えるの？など

「Omokage検索」

Typical structure search (e.g. Structure Navigator of PDBj)



Omokage search



URL: <https://pdbj.org/yorodumi>

AMED-BINDS

The screenshot shows the homepage of the AMED-BINDS website. At the top, there is a navigation bar with links for Japanese (selected), English, Googleカスタム検索 (Custom Search), and search functionality. Below the navigation bar are links forよくある質問 (FAQ), コンサルティング支援申請 (Consulting Support Application), and 内部向け (Internal). A horizontal menu bar includes links for トップ (Top), 運営 (Operation), 研究課題 (Research Topics), 支援者 (Supporters), 支援の流れ (Support Flow), ニュース (News), 成果 (Achievements), 用語集 (Glossary), お問い合わせ (Contact), and リンク集 (Link Collection). The main content area features three images: a large aerial view of a circular research facility in a green landscape, a molecular structure model, and a server room. A banner below these images reads "創薬やライフサイエンス研究を支援するプロジェクト" (Project supporting pharmaceutical and life sciences research). Below the banner, the Binds logo and the full name "創薬等先端技術支援基盤プラットフォーム" are displayed, along with the English translation "Basis for Supporting Innovative Drug Discovery and Life Science Research". The bottom section of the page shows images of a laboratory, a molecular structure, and a close-up of a sample.

URL: <https://binds.jp>

電子顕微鏡構造解析・化合物ライブラリーなど、創薬に係る研究を支援します

データと画像の利用について

権利の話

- PDBデータはフリー（パブリックドメイン）
- 各ウェブサイトの画像などについてはそれぞれのドキュメントを参照
 - PDBjの利用規約: <https://pdbj.org/info/terms-conditions>

科学の文脈でのマナー

- データの出処は明記する、PDBの場合はPDB-ID

ソフトウェアのライセンスにも注意

- 文献引用の義務・要請がある場合も