

RCSB PDB – descrizione sintetica e utilità.

RCSB PDB (<http://www.rcsb.org/pdb/home/home.do>) è un risorsa di studio delle macromolecole affiliata alla worldwide Protein Data Bank.

Contiene informazioni riguardo a proteine la cui struttura è stata determinata sperimentalmente, acidi nucleici e complessi di macromolecole.

Nella Homepage, alla voce Featured Molecules (molecole in primo piano), si trova una delle peculiarità offerte dal RCSB PDB: suddivise nelle sei categorie di Protein Synthesis, Enzymes, Health and Disease, Biological Energy, Infrastructure and Communication e Biotechnology and Nanotechnology, a loro volta formate da sottocategorie più specifiche, si ritrovano presentate molecole selezionate di cui vengono riportate, oltre alla descrizione generale di struttura e funzione, anche approfondimenti tematici (relativi all'ambito della ricerca e delle applicazioni mediche ad esempio) e link di discussione, e ovviamente i relativi collegamenti alle voci "standard" della PDB. Si tratta quindi di un utile, e ben strutturato, sistema di approfondimento che evade i classici percorsi e strumenti di classificazione e descrizione delle macromolecole, offrendone una visione contestualizzata, sebbene ovviamente non sostitutiva dei comuni metodi di indagine (del resto si tratta di una selezione).

Sempre nella Homepage è possibile ritrovare la voce Latest Structures, tramite la quale è possibile accedere alla lista aggiornata delle ultime strutture caricate nel database.

Si possono inserire svariati tipi di Query, ma le sequenze non rientrano nelle possibilità di ricerca.

RCSB PDB è un database di strutture, molto ben organizzato, ma non svolge il lavoro di BLAST (ad esempio) motivo per cui, se partiamo da una sequenza ignota, dovremo scoprirne l'identità prima di poter effettuare una ricerca per la sua struttura attraverso questo strumento.

Può capitare, inoltre, che il database non contenga la struttura specifica che stiamo cercando, ma che al contempo possa contenerne altre ad essa affini e comunque utili per la nostra ricerca. Un accorgimento a tal proposito sarebbe quello di non effettuare ricerche in RCSB PDB che siano eccessivamente specifiche, a meno di non essere già in possesso del codice PDB di ciò che stiamo cercando (in tal caso avremmo la certezza dell'esistenza di una struttura della macromolecola nel db). Ad esempio se stiamo cercando la proteina X di H. sapiens, è più comodo ricercare (per nome) semplicemente la proteina X, e scremare poi i risultati ottenuti utilizzando le Query Refinements (Figura in basso).

504 Structure Hits | 32 Unreleased Structures | 201 Citations | 118 Ligand Hits | 94 Web Page Hits | 60 Hits | SCOP Hits | CATH Hits

Query Parameters:
Text Search for: HEMOGLOBIN

Query Refinements Hide

- Organism**
 - Homo sapiens (man) (204)
 - Scapharca inaequivalvis (39)
 - Lupinus luteus (17)
 - Equus caballus (equine) (17)
 - Physeter catodon (16)
 - Bos taurus (domestic cow) (11)
 - Mycobacterium tuberculosis (10)
 - Other (190)
- Taxonomy**
 - Eukaryota (465)
 - Bacteria (39)
 - Other (1)
- Experimental Method**
 - X-RAY (492)
 - Solution NMR (10)
 - Neutron Diffraction (2)
- X-Ray Resolution**
 - less than 1.5 Å (35)
 - 1.5 - 2.0 Å (192)
 - 2.0 - 2.5 Å (179)
 - 2.5 - 3.0 Å (64)
 - 3.0 and more Å (22)
 - more choices...
- Release Date**
 - before 2000 (137)
 - 2000 - 2005 (166)
 - 2005 - 2010 (161)
 - 2010 - today (40)
 - this month (7)
 - more choices...
- Polymer Type**
 - Protein (504)
- Enzyme Classification**
 - 3: Hydrolases (11)
 - 1: Oxidoreductases (1)
- SCOP Classification**
 - All alpha proteins (355)
 - All beta proteins (15)
 - Alpha and beta proteins (a+b) (6)
 - Multi-domain proteins (alpha an ... (2)
 - Alpha and beta proteins (a/b) (1)
 - Coiled coil proteins (1)
 - Small proteins (1)

Refine Query Remove Similar: Select Percent Similarity

Queste presentano le voci Organismo, Tassonomia, Metodo Sperimentale, Risoluzione X-Ray, Data di rilascio, Tipo di polimero, Classificazione Enzimatica, Classificazione SCOP. Sotto ognuna di esse, vengono visualizzate le varie sottoclassi ottenute dalla ricerca con il relativo numero di match fra parentesi.

Nell'esempio in figura è stata effettuata una ricerca per HEMOGLOBIN e sono state evidenziate le Query Refinements. L'emoglobina è praticamente impossibile da non ritrovare, perfino in una ricerca specifica, e addirittura una ricerca troppo generale per questa proteina è controproducente (se infatti le nostre intenzioni fossero quelle di ricercare l'emoglobina di *Mus musculus*, dovremmo andare a scartabellare fra le varie voci degli Organisms per recuperarla attraverso le Query Refinements). Tuttavia se fossimo interessati alla phosphatidylinositol transfer protein beta isoform di homo sapiens, e inserissimo "phosphatidylinositol transfer protein beta isoform homo sapiens" come query, troveremo un solo risultato, quello della phosphatidylinositol transfer protein alfa isoform umana. Effettuando una ricerca per il solo nome della proteina, invece, tra i tre risultati ottenuti risulterebbe la phosphatidylinositol transfer protein beta isoform di *Rattus norvegicus*, e trattandosi di una proteina estremamente conservata nelle varie specie, la struttura in questione farebbe comunque al caso nostro. In altre parole, una ricerca troppo specifica ci avrebbe portato a perdere un risultato interessante.

Ogni voce di RCSB PDB, oltre a permettere sia di visualizzare le strutture 3D direttamente dal sito, sia di scaricarne il relativo file PDB, presenta numerosi link, ben organizzati e soprattutto non ridondanti (è molto difficile perdersi utilizzando questo strumento). Per una proteina troveremo i link dei database relativi ai domini (External domain annotations) secondo SCOP, CATH, Pfam e GO. Inoltre la voce riporterà i ligandi quando presenti, e tutte le informazioni relative la struttura, la classificazione, l'organismo.

Da notarsi la distinzione tra i link veri e propri (indicati in blu e con una freccetta) e le ricerche rapide (anch'esse in blu, ma con una lente di ingrandimento a marcarle). Cliccando sui primi si arriva a siti esterni (quali quelli dei database di informazioni sui domini), mentre tramite le seconde si eseguono ricerche rapide sullo stesso RCSB PDB (cliccando sul nome di uno degli autori, ad esempio, si visualizzeranno tutte le strutture presenti nel db alla cui produzione quell'autore ha preso parte).