

## Vizualizacija proteina

Protein Data Bank je kristalografska baza podataka trodimenzionalnih strukturalnih podataka velikih bioloških molekula poput bjelančevina i nukleinskih kiselina. Podatci su obično dobiveni kristalografijom X-zrakama, spektroskopijom NMR-om, ili u zadnje vrijeme elektronskom mikroskopijom.

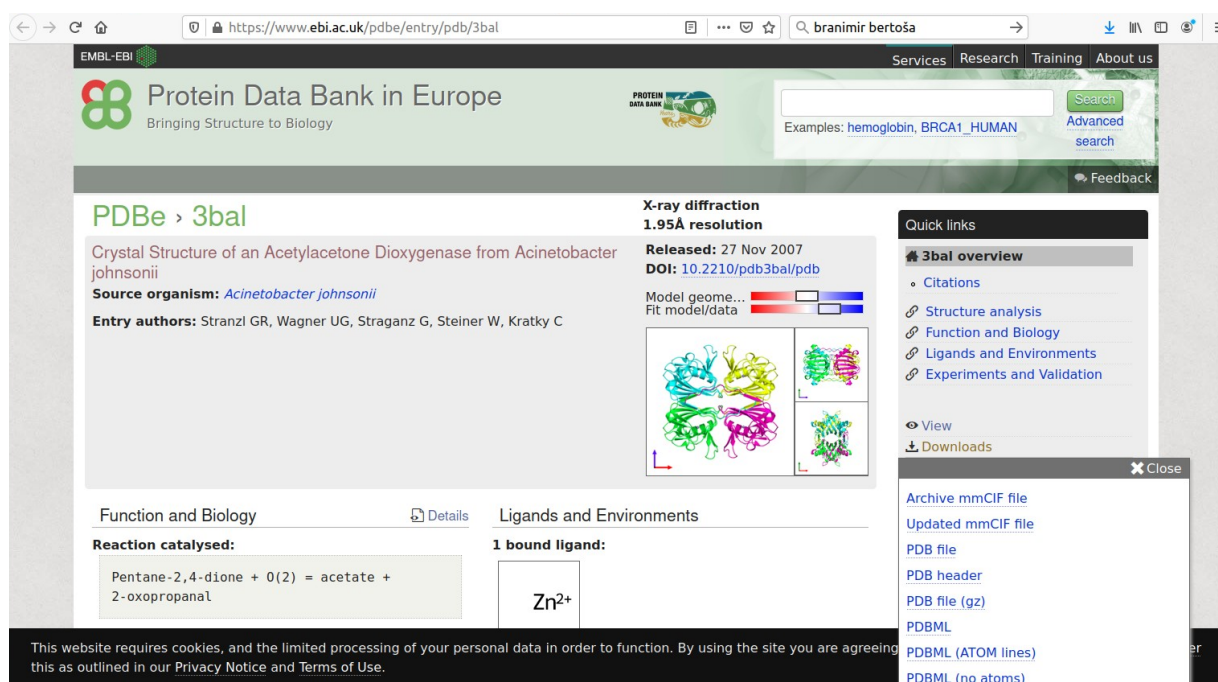
Tri su dostupne proteinske baze podataka:

<https://www.ebi.ac.uk/pdbe/>

<http://www.rcsb.org/pdb>

<https://pdj.org/>

do 28. 4. 2021. u proteinskim bazama podataka spremljene su 177 219 stukture proteina i DNA molekula. Svaka molekula ima svoj jedinstveni kod koji se sastoji od 4 znaka. Npr. 3BAL predstavlja strukturu *Acetil aceton dioksigenaze iz Acinetobacter Johnsonii*. Upisom u tražilicu na bilo kojoj proteinskoj bazi podataka može se dobiti informacija o pohranjenoj strukturi molekule. Klikom na četveroznačni kod strukture mogu se dobiti informacije o strukturalnoj analizi, funkciji i biologiji molekule, ligandima, okruženju eksperimentima i validaciji. Klikom na poveznicu "DOWNLOADS" koja se nalazi na desnoj strani stranice može se preuzeti pdb (.ent) zapis strukture molekule (slika 1).



The screenshot shows the PDB entry page for 3BAL. The main content area includes the title 'Crystal Structure of an Acetylaceton Dioxygenase from Acinetobacter johnsonii', the source organism 'Acinetobacter johnsonii', and the authors 'Stranzi GR, Wagner UG, Straganz G, Steiner W, Kratky C'. It also displays the X-ray diffraction resolution of 1.95 Å and the release date of 27 Nov 2007. A 'Downloads' menu is open, showing various file formats available for download, including mmCIF, PDB, and PDBML files.

Slika 1. prikaz zapisa molekule 3BAL u PDBe bazi

Zapis PDB molekule može se otvoriti u bilo kojem softveru za editiranje, a nakon zaglavlja u kojem se nalaze osnovni podaci o načinu i datumu kristalizacije, modificiranima aminokiselinama, rezoluciji, ligandima, metalima u aktivnom mjestu itd... slijedi niz znakova koji su ključni za vizualizaciju:

```
ATOM 1 N MET A 1 45.464 -13.560 33.175 1.00 34.09 N
ATOM 2 CA MET A 1 46.503 -12.539 33.454 1.00 31.18 C
ATOM 3 C MET A 1 46.012 -11.159 33.022 1.00 28.68 C
ATOM 4 O MET A 1 44.951 -10.714 33.466 1.00 27.01 O
ATOM 5 CB MET A 1 47.815 -12.899 32.735 1.00 36.07 C
```

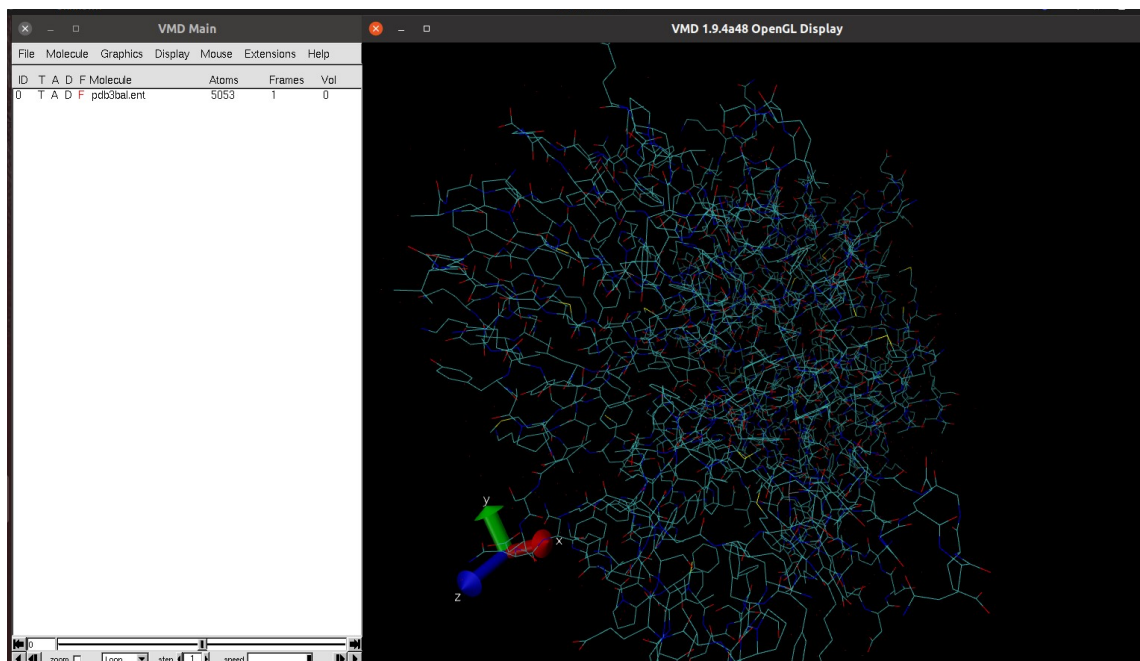
ATOM	6	CG	META	1	47.744	-12.947	31.206	1.00	43.73	C
ATOM	7	SD	META	1	49.396	-13.050	30.403	1.00	54.66	S
ATOM	8	CE	META	1	49.571	-14.846	30.138	1.00	50.02	C
ATOM	9	N	ASPA	2	46.768	-10.494	32.148	1.00	24.97	N
ATOM	10	CA	ASPA	2	46.407	-9.158	31.692	1.00	22.84	C
ATOM	11	C	ASPA	2	46.213	-9.043	30.178	1.00	20.65	C
ATOM	12	O	ASPA	2	46.550	-9.954	29.428	1.00	19.78	O
ATOM	13	CB	ASPA	2	47.478	-8.156	32.136	1.00	24.76	C
ATOM	14	CG	ASPA	2	47.782	-8.247	33.621	1.00	27.81	C
ATOM	15	OD1	ASPA	2	46.837	-8.168	34.439	1.00	28.75	O
ATOM	16	OD2	ASPA	2	48.970	-8.397	33.972	1.00	29.75	O
ATOM	17	N	TYRA	3	45.664	-7.907	29.749	1.00	18.68	N
ATOM	18	CA	TYRA	3	45.423	-7.627	28.333	1.00	16.67	C
ATOM	19	C	TYRA	3	46.731	-7.780	27.548	1.00	17.97	C
ATOM	20	O	TYRA	3	47.786	-7.327	28.003	1.00	16.42	O
ATOM	21	CB	TYRA	3	44.887	-6.195	28.173	1.00	16.56	C
ATOM	22	CG	TYRA	3	44.797	-5.713	26.739	1.00	16.02	C
ATOM	23	CD1	TYRA	3	43.771	-6.145	25.899	1.00	15.45	C
ATOM	24	CD2	TYRA	3	45.785	-4.886	26.204	1.00	15.12	C
ATOM	25	CE1	TYRA	3	43.736	-5.776	24.555	1.00	14.74	C
ATOM	26	CE2	TYRA	3	45.763	-4.511	24.865	1.00	15.64	C
ATOM	27	CZ	TYRA	3	44.740	-4.964	24.043	1.00	15.78	C
ATOM	28	OH	TYRA	3	44.756	-4.650	22.699	1.00	15.97	O
ATOM	29	N	CYSA	4	46.661	-8.408	26.371	1.00	17.20	N
ATOM	30	CA	CYSA	4	47.847	-8.612	25.531	1.00	17.07	C
ATOM	31	C	CYSA	4	47.937	-7.582	24.408	1.00	16.52	C
ATOM	32	O	CYSA	4	47.078	-7.549	23.527	1.00	14.48	O
ATOM	33	CB	CYSA	4	47.830	-10.014	24.912	1.00	17.90	C
ATOM	34	SG	CYSA	4	47.792	-11.384	26.104	1.00	24.53	S
ATOM	35	N	ASNA	5	48.976	-6.749	24.437	1.00	15.86	N
ATOM	36	CA	ASNA	5	49.171	-5.736	23.403	1.00	17.51	C
ATOM	37	C	ASNA	5	49.332	-6.409	22.039	1.00	17.63	C
ATOM	38	O	ASNA	5	49.945	-7.469	21.938	1.00	15.82	O

.....

gdje nakon oznake „ATOM” slijedi broj atoma, zatim oznaka atoma (npr. N, C, CA, O....), oznaka aminokiselinskog ostatka, (npr. CYS, ARG, ALA...), broj aminokiselinskog ostatka, koordinate svakog pojedinog atoma u Kartezijevom (3D) koordinatnom sustavu...

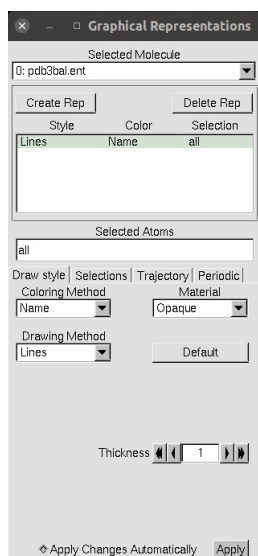
Postoji cijeli niz programa koji služe sa vizualizaciju, od kojih su najpoznatiji pymol (<https://pymol.org/2/>) i VMD (<https://www.ks.uiuc.edu/Research/vmd/>). U nastavku će biti opisana vizualizacija molekule 3BAL pomoću softvera VMD.

Nakon preuzimanja .pdb zapisa molekule, istu je potrebno otvoriti pomoću softvera VMD (prethodno je softver potrebno preuzeti s gornje poveznice i instalirati na računalo). Prikaz koji se dobije otvaranjem molekule u softveru VMD dan je na slici 2.



Slika 2. prikaz molekule nakon otvaranja u softveru VMD

Ovakav prikaz nije pregledan, stoga ga je potrebno prilagoditi. To se može napraviti klikom na isbornik *Graphics-> Representations* gdje se dobije izbornik prikazan na slici 3.

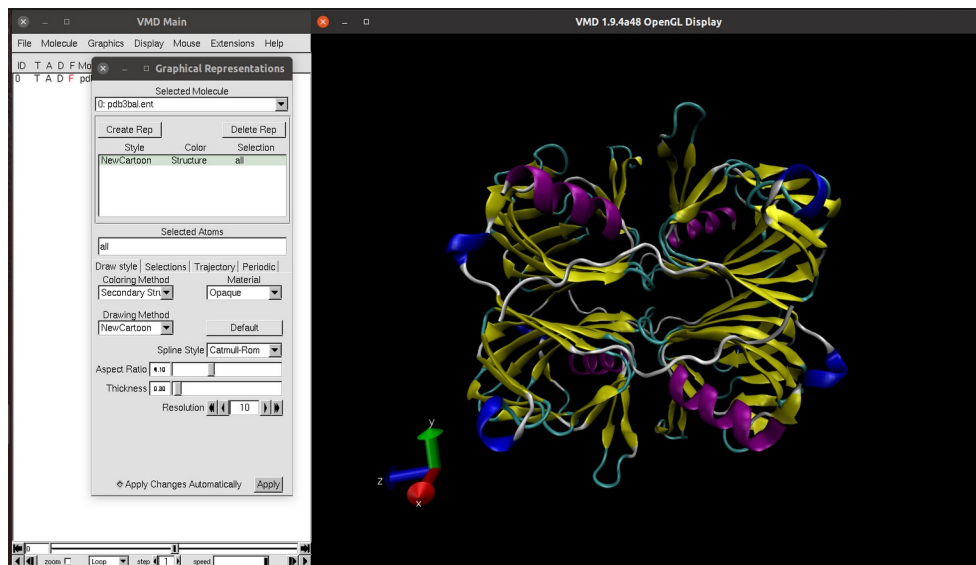


Slika 3. Izbornik

Na izborniku *Drawing method* može se odabrati cijeli niz prikaza proteina, kao npr. *CPK* (Corey, Pauling and Koltun) reprezentacija pomoću štapića i kuglica, samo štapićasta (*Licorice*) reprezentacija, preprezentacija van der Waalsovih kuglica (*VDW*) ili reprezentacija sekundarnih struktura (*New cartoon*). Pojedine reprezentacije zahtjevaju značajne računalne resurse, pa postoji mogućnost sporijeg prikaza molekule.

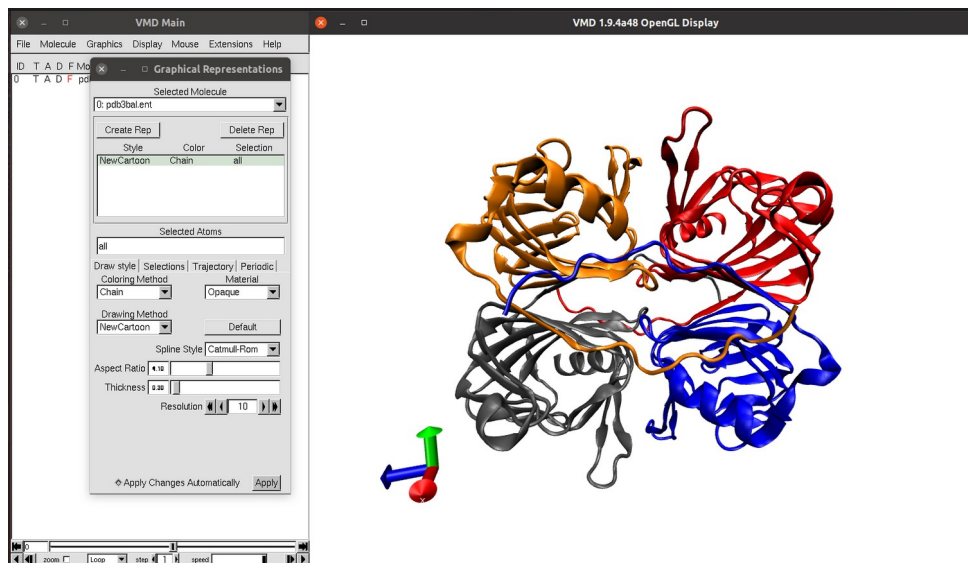
Osim metode crtanja može se odabrati i metoda bojanja u padajućem izborniku *Coloring Method*. Tako bojanje može biti prema imenu, vrsti, lancu, Sekundarnoj strukturi...

U izborniku *Selected Atoms* upisuje se koje atome želimo odabrati. Ako je u izbornik upisano *all*, prikazati će se svi atomi izlistani u PDB datoteci (slika 4 prikazuje sve atome u molekuli, metodom *New Cartoon*, i bojanjem sekundarnih stuktura). Slika 5 prikazuje sve atome obojane prema lancima. Ako se izbornik upiše npr. *resid 62 64 104* u sva četiri lanca prikazati će se aminokiselinski ostatci 62, 64 i 104 koji tvore aktivno mjesto u molekuli 3BAL.

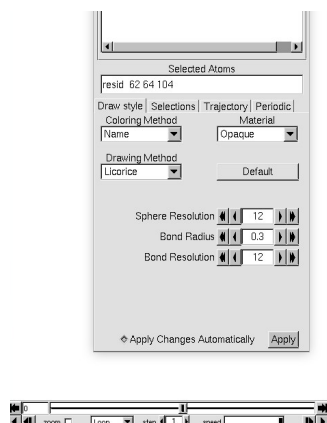


Slika 4. Svi atomi u molekuli metodom New Cartoon, i bojanjem sekundarnih stuktura

Boja pozadine može se promijeniti klikom na izbornik *Graphics* -> *Colors*, i odabrati *Categories – Display, Names – Background* i *Colors – 8 white* (ili prema vlastitoj želji). Bolji kontrast slike može se dobiti gašenjem gumba „*Depth Cueing*” u izborniku *Display*.

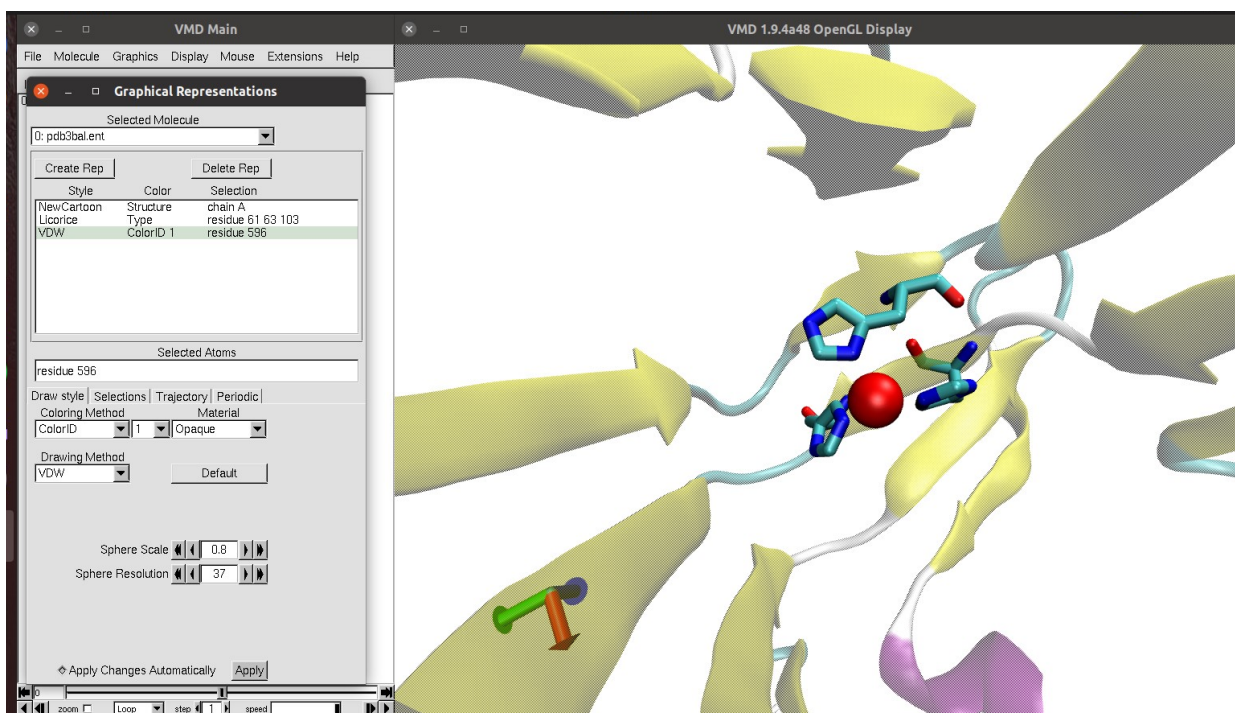


Slika 5. Prikaz tetramera onojanog prema lancima



Slika 6. Aminokiselinski ostatci u aktivnom mjestu

Ako se želi prikazati više značajki istovremeno, klikom na tipku *Create Rep* dodaje se nova reprezentacija, na taj je način moguće prikazati smještaj aktivnog mjesta unutar sekundarne strukture, prikazati aminokiselinke ostatke u veznom mjestu i ligand vezan u to mjesto, kao na slici 7.



Slika 7. Lanac A enzima prikazan je u New Cartoon reprezentaciji, aminokiselinski ostatci koji tvore vezno mjesto prikazani su štapićasto i obojeni prema imenu, dok je atom cinka koji se nalazi u veznom mjestu obojan crvenom kuglom, prikazanom u VDW reprezentaciji. U ovom prikazu je umjeto riječi resid koristena riječ residue, pomoću koje se prikazuje redni broj aminokiselinskog ostatka, kojem brojanje započinje od 0, stoga su brojevi aminokiselinskih ostataka umanjeni za 1.

**Zadatak:** u proteinskoj bazi podataka pronaći molekulu prema izboru, i vizualizirati je pomoću jednog od softvera za vizualizaciju. Prikazati njenu strukturu, kao i strukturu aktivnog odnosno veznog mjesta. Izvještaj mora sadržavati:

1. Naziv proteina i njegovi četveroznačnu oznaku
2. Informacije o primarnoj, sekundarnoj, tercijalnoj i kvartalnoj strukturi
3. Informaciju o rezoluciji na kojoj je struktura dobivena,
4. Informacije o ligandima (metalima, inhibitorima) koji su važni za funkciju enzima
5. Prikaz proteina i prikaz aktivnog mjesta proteina (samo jednog ako je protein kristaliziran kao dimer ili tetramer)

Svaki student mora prikazati različiti protein,protein prikazan u primjeru je isključen iz odabira.