

# AADS

Automated Active Site Identification, Docking, and Scoring

## AADS web server Manual

Developed at Department of Chemistry, Supercomputing Facility for Bioinformatics & Computational Biology, and School of Biological Sciences, Indian Institute of Technology, Hauz Khas, New Delhi-110016, India

### Reference:

Singh T, Biswas D, and Jayaram B. (2011) **AADS-an automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors.** J Chem Inf Model. 24;51(10):2515-27. doi: 10.1021/ci200193z.

## Introduction

---

Prediction of the functional active site from the tertiary structure of the protein is one of the essential steps involved in drug discovery. The experimental evidence on the active site of the protein is quite valuable. In the absence of experimental data on the active site, several tools are available in the public domain to predict it. We present here an automated active site prediction server (AADS). AADS is a robust active site detection, docking, and scoring tool for proteins. It predicts the active site with 100% accuracy when the top ten binding pockets are considered. It is implemented on an 80-processor cluster and is freely available at <http://scfbio-iitd.res.in/AADS/>. The methodology takes 3-4 seconds on a single processor to predict active-site. Its accuracy was determined to be 100 % when the top ten cavities were assessed on 620 proteins with known active sites and amino acid lengths ranging from 100 to 600. AADS predicts cavities based on the fuzzy score, which utilizes the Physico-chemical properties of the functional groups lining the active site such as hydrogen bond donor, hydrogen bond acceptor, volume, ring structure, hydrophobic groups, etc. The cavity with the maximum volume, the maximum number of hydrogen-bond donor and acceptors and hydrophobic groups scores maximum utilizing the fuzzy score.

$$\text{Score}_j = \left( \sum_{i=1..n} (x_{ij}/x_i^{\max}) \right) / n$$

where, **Score<sub>j</sub>** is the fuzzy score that conveys the likeliness of the  $j^{\text{th}}$  cavity for a particular protein to be an actual ligand binding site,

**X<sub>ij</sub>** is any of the (n) properties considered such as approximate volume, number of hydrogen bond donors, number of hydrogen bond acceptors, or number of ring structures and number of hydrophobic groups in the  $j^{\text{th}}$  cavity of the protein,

**X<sub>i</sub><sup>max</sup>** is the maximum value for the corresponding parameter in the protein,

**n=5**; and the value of the score is in between 0-1, where 1.0 is the highest score to be called as ligand binding site.

Docking can also be done in an optional mode with the predicted top ten cavities in AADS. The docking protocol uses an all-atom energy-based Monte Carlo method. This method generates many random configurations of the ligand in six-dimensional space to predict the best possible conformation of the ligand in the binding pocket of the target macromolecule (protein). The combination of all-atom energy-based scoring with the Monte Carlo search technique appears to provide a reliable method for protein-ligand structure optimization and binding affinity prediction.

The following steps are involved in docking the candidate drug molecules at the reference points:

- a) Preparation of the protein and the candidate drug molecule,
- b) Translation of the drug molecule to the reference cavity points,
- c) Grid Generation,
- d) Generation of Monte Carlo configurations of the candidate drug molecule in the cavity points,
- e) Collection of eight low energy configurations for each reference cavity point.

Eight low energy docked structures of the candidate molecule are stored at each cavity point, giving a total of 80 docked structures. These 80 docked structures are energy minimized using AMBER's sander module, which is then ranked using an effective scoring function, and the top five structures are selected. The used scoring function in AADS is an all-atom energy based empirical scoring function comprising electrostatics, van der Waals, hydrophobicity,

and loss of conformational entropy of protein side chains upon ligand binding and is of the following:

$$\Delta G = \alpha(E_{el}) + \beta(E_{vdw}) + \left(\sum_{A=1}^{22} \sigma_A \Delta A_{LSA}\right) + \lambda(\Delta S_{CR}) + \delta$$

where,  $\Delta G$  is the binding free energy (in kcal/mol),

$E_{el}$  is the electrostatic component,

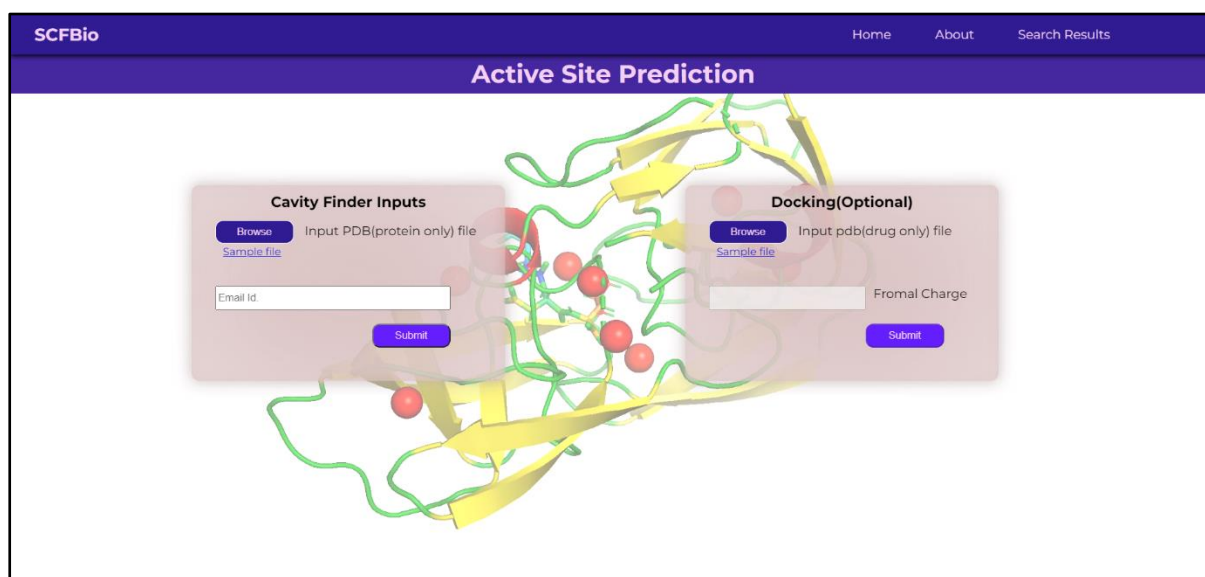
$E_{vdw}$  is the van der Waals component,

$\Delta A_{LSA}$  is the loss in surface area of the atom type,

$\sigma_A$  is the atomic desolvation parameter in kcal/mol/Å<sup>2</sup> for an atom type,

$\Delta S_{CR}$  is the loss in conformational entropy, and  $\delta$  is a constant.

**Availability - <http://scfbio-iitd.res.in/AADS/>**



## Format Specification

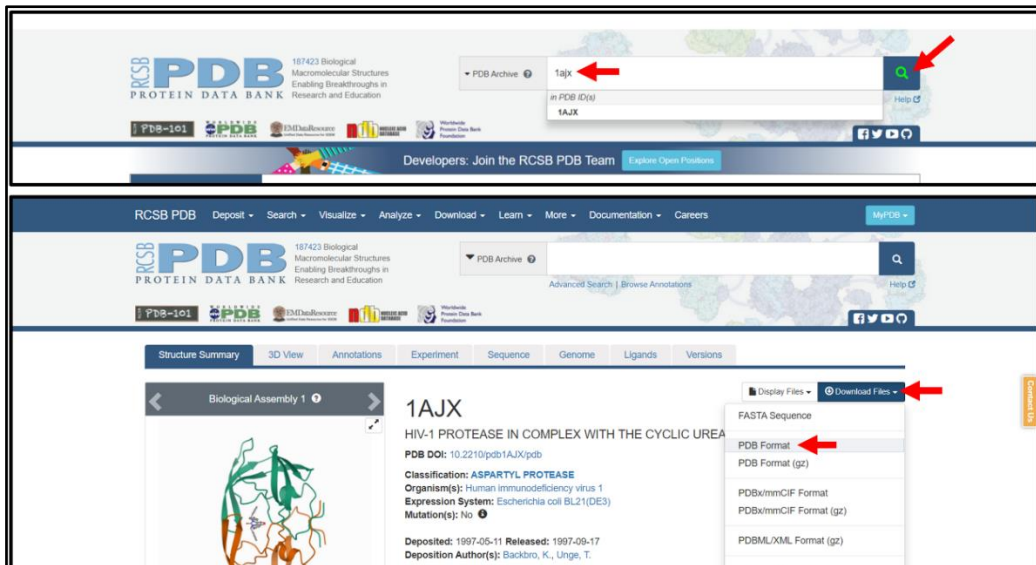
### Protein:

Input proteins must be in standard PDB format and have a .pdb extension. The term PDB stands for Protein Data Bank. One can download a PDB file from the RCSB website (<https://www.rcsb.org>) or generate the homology/ab-initio based tertiary structure by protein modelling tools. To download the protein from RCSB, follow the below steps:

- 1) Go to the site <https://www.rcsb.org/>
- 2) Type the 4-character PDB code (eg., 1AJX) of the protein you want to download on the search bar and click at the search icon 🔍.
- 3) Now click at the drop-down arrow present at the right-side of “Download Files”
- 4) Now select PDB format from the drop-down menu. The file will download.

### Ligand:

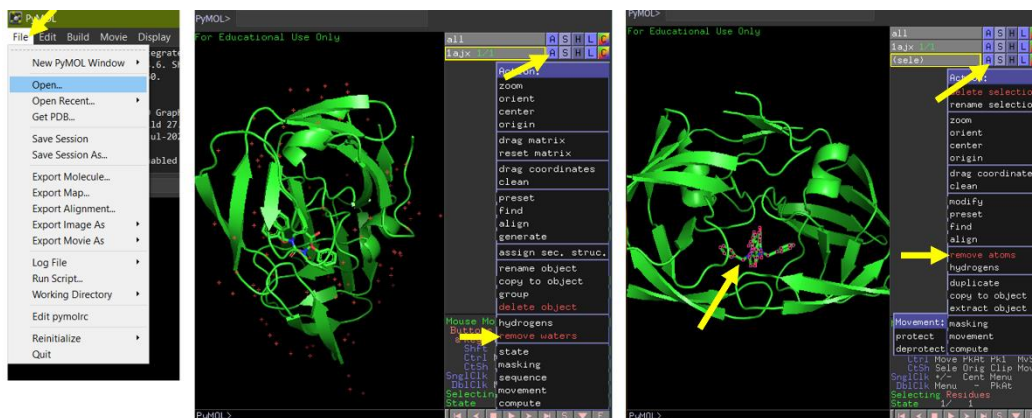
For docking, AADS only takes the input ligand in PDB format. **Hydrogen must be added to the ligand PDB file and must have a .pdb extension.**



## File Preparation:

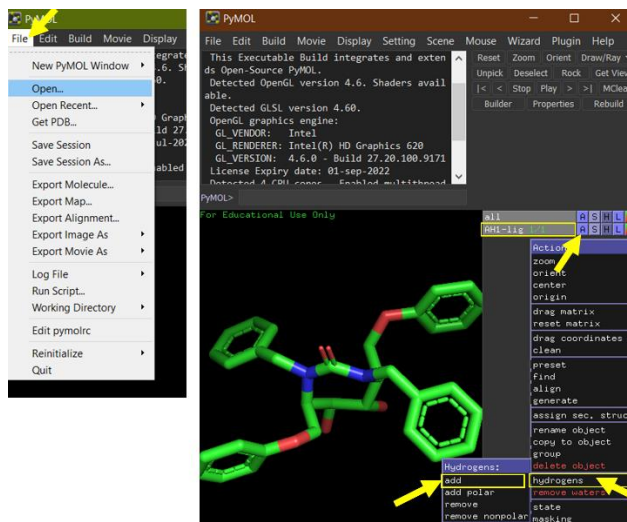
**Protein:** It is advisable to clean the input protein file before submitting it to the AADS. There are several ways to clean a PDB file. One is given below using PyMol (a molecular visualization tool)-

- Open PyMol and click on **File** tab.
- Select **Open**.
- Select the path and type the name of your protein file.
- **To remove water molecule from protein**, click on menu **Action (A)** located at the right corner and select **remove waters**.
- **To remove heteroatoms from protein**, select the heteroatoms by using left-click button of mouse. Now click on the menu **Action (A)** of section object and select **remove atoms**.



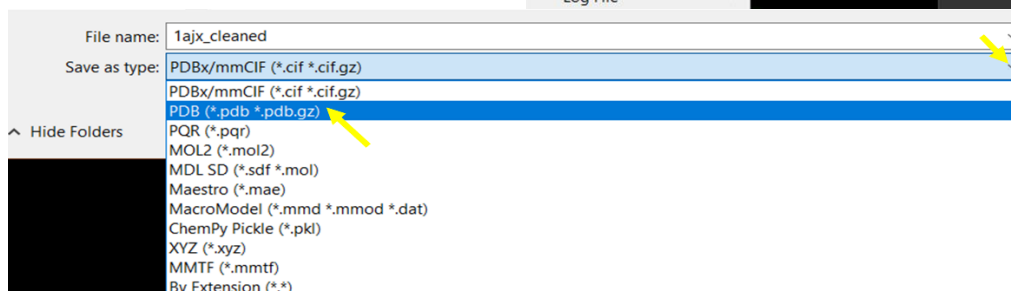
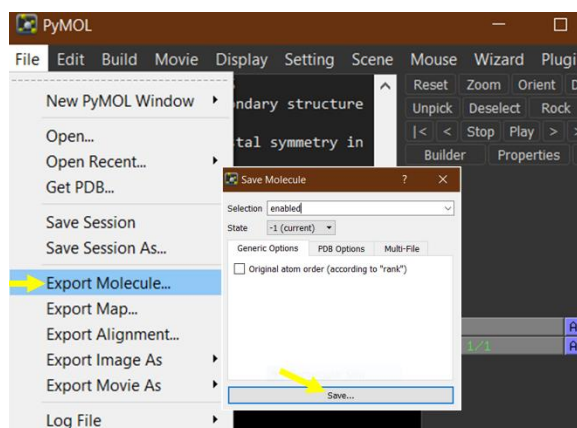
**Ligand:** There are several ways to add hydrogen to the ligand molecule. One is given below using PyMol-

- Open PyMol and click on the **File** tab.
- Select **Open**.
- Select the path and type the name of your ligand file.
- **To add hydrogen molecule in ligand**, click on menu **Action (A)**, select **hydrogens**, and click on **add**.



### Now, to save cleaned protein and ligand file from pymol –

- Click on the **File** tab and select **Export Molecule**.
- Click on **save**
- Select the path and type a new file name in the **File Name:** text box.
- Select **PDB (\*.pdb \*.pdb.gz)** in the **Save as type:** list box.
- Click on the **Save** button.



# Job Submission

To Submit a job in AADS, follow the steps below- For active site prediction:

1. Upload the three-dimensional structure of protein biomolecule in **PDB** format.
2. Input the email address in the text box to retrieve the result after the completion of the job.
3. Click on Submit button.

For docking: After active site detection, the user can perform docking of the ligand in the top ten predicted binding pockets.

1. Click on the **Docking** option provided at the end of the result page.
2. Upload the ligand molecule in PDB format.
3. Verify the formal charge of the ligand that was automatically calculated by the server.
4. Click on Submit button to proceed.

The screenshot shows the SCFBio website's 'Active Site Prediction' page. It features a central 3D protein structure with highlighted cavities. On the left, a red bracket labeled 'For active site prediction' encompasses three steps: 1. Clicking 'Browse' to upload a file named '2aps.pdb'. 2. Entering an email address 'skankshakeshanani44@gmail.com' in the text box. 3. Clicking the 'Submit' button. On the right, the 'Docking(Optional)' section is visible, with a 'Browse' button and a 'Formal Charge' input field. A status bar at the bottom indicates the job is submitted with ID '1646025254' and provides a link to check results.

The screenshot displays the 'Cavities Results' page. It lists 36 cavities from Cavity\_1 to Cavity\_36, each with a corresponding link. A 'Download Final Result' button is present. A red arrow labeled '4' points to the 'Docking (Optional)' button. A red text box at the bottom says 'Click here to proceed for docking. User will be navigated here'. A 3D protein structure with highlighted cavities is shown on the right.

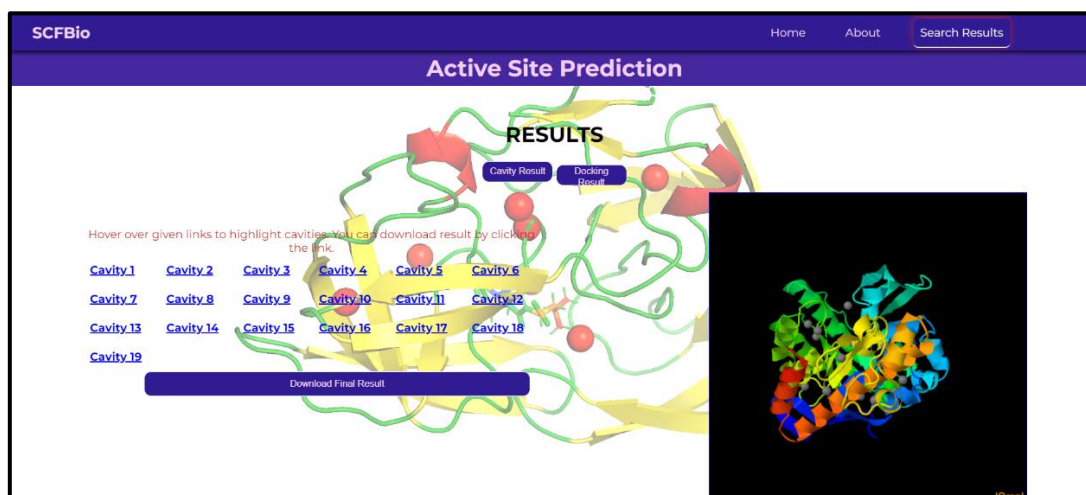
This screenshot shows the 'Active Site Prediction' page again, but with the 'Docking(Optional)' section highlighted. A yellow arrow from the previous page points to step 5: clicking 'Browse' to upload a ligand file named '1sh\_NEW-lif.pdb'. Step 6 shows the 'Formal Charge' input field set to '1'. Step 7 shows the 'Submit' button being clicked. The status bar at the bottom shows the job ID '1646025254'.

# Results

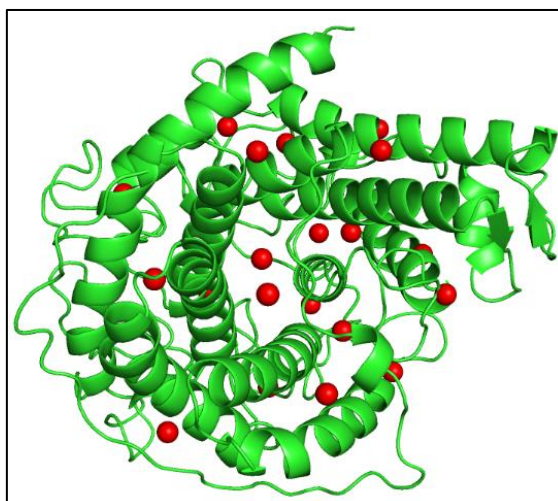
## Active-site result:

AADS predicts an active site in 3-4 seconds and performs a docking with the top 10 predicted active-site in 30 minutes. The docking time may increase depending on the input file's size and the queue's availability at the server site. To retrieve the result, click on the link mailed to you via the [SCFBio@scfbio-iitd.res.in](mailto:SCFBio@scfbio-iitd.res.in) account with the subject "AADS Cavity Results". Now click on the provided link to see the status of the result. The result can also be retrieved by the **Job id** provided after each job submission. The AADS finds all the possible cavity points in a protein and gives a list of the residues (in single letter code) lining the respective cavity points and names them as Cavity1, Cavity2, etc. Users can highlight the cavity to see its location in the protein by hovering over each link one by one. These cavity points can be downloaded just by clicking on their name. Click on the '**Download Final Result**' link to download the result. Active-site information is provided at the end of the PDB file with the DRG identifier.

A result page is shown below for more clarity-



The file can be graphically viewed by using certain software like Viewerlite, PyMol, Mercury, Swissmol PDBviewer (spdv). Below is shown a result visualized by PyMol. The red dots are representing the predicted active sites.



Here is shown the downloaded result file in a notepad having active site info -

ATOM	6884	C	ALA	470	28.084	31.534	41.137	1.00	0.00
ATOM	6885	O	ALA	470	29.242	31.895	41.428	1.00	0.00
ATOM	6886	OXT	ALA	470	27.136	32.290	40.830	1.00	0.00
TER									
ATOM	99800	He	DRG	9900	32.399	17.910	7.713	1.00	0.00
ATOM	99801	He	DRG	9901	51.993	33.081	25.742	1.00	0.00
ATOM	99802	He	DRG	9902	24.203	28.346	26.032	1.00	0.00
ATOM	99803	He	DRG	9903	39.749	17.687	35.260	1.00	0.00
ATOM	99804	He	DRG	9904	52.686	36.609	10.499	1.00	0.00
ATOM	99805	He	DRG	9905	38.618	32.768	28.892	1.00	0.00
ATOM	99806	He	DRG	9906	45.928	7.148	30.942	1.00	0.00
ATOM	99807	He	DRG	9907	24.450	15.218	19.149	1.00	0.00
ATOM	99808	He	DRG	9908	58.950	17.527	25.847	1.00	0.00
ATOM	99809	He	DRG	9909	39.764	6.678	14.328	1.00	0.00
ATOM	99810	He	DRG	9910	39.067	21.392	22.110	1.00	0.00
ATOM	99811	He	DRG	9911	34.996	36.829	11.361	1.00	0.00
ATOM	99812	He	DRG	9912	53.864	8.599	16.645	1.00	0.00
ATOM	99813	He	DRG	9913	57.653	21.974	12.274	1.00	0.00
ATOM	99814	He	DRG	9914	31.565	27.261	12.790	1.00	0.00
ATOM	99815	He	DRG	9915	28.771	2.430	23.613	1.00	0.00
ATOM	99816	He	DRG	9916	44.850	16.019	6.060	1.00	0.00
ATOM	99817	He	DRG	9917	45.882	24.587	3.600	1.00	0.00
ATOM	99818	He	DRG	9918	27.928	19.234	32.532	1.00	0.00
ATOM	99819	He	DRG	9919	47.974	13.455	22.578	1.00	0.00
ATOM	99820	He	DRG	9920	41.848	34.521	16.427	1.00	0.00
ATOM	99821	He	DRG	9921	49.282	24.485	13.960	1.00	0.00

Predicted active sites

Here is shown the cavity file when opened in a notepad-

The volume of the cavity is in the Angstrom cube. The fourth column represents the amino acid surrounding the cavity, and the fifth their corresponding number.

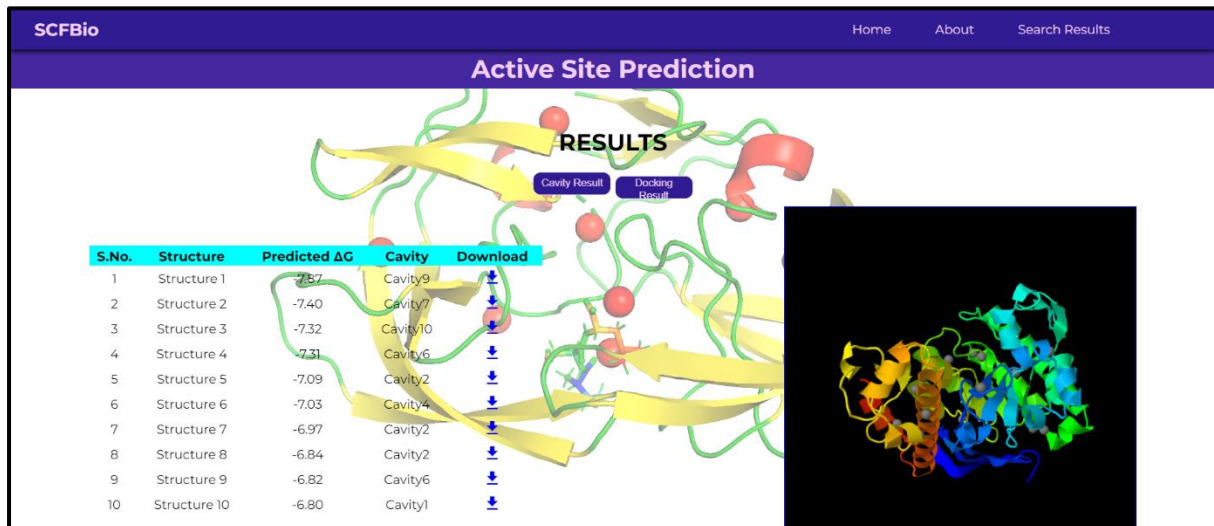
```
REMARK File Name = cavity_1_YDFSWTKPVEAGLQR
REMARK Cavity point 32.399 17.910 7.713
REMARK Volume of the Cavity in Angstrom cube = 513
REMARK
```

ATOM	1	3HB	TYR	115	22.517	12.649	10.538	1.00	1.00
ATOM	2	CG	TYR	115	21.910	13.897	8.886	1.00	1.00
ATOM	3	CB	TYR	115	21.707	13.335	10.283	1.00	1.00
ATOM	4	OD2	ASP	47	21.627	20.871	5.050	1.00	1.00
ATOM	5	CG	TYR	50	22.649	19.523	10.336	1.00	1.00
ATOM	6	CB	TYR	50	23.258	19.740	11.700	1.00	1.00
ATOM	7	HE2	TYR	50	22.991	20.567	7.110	1.00	1.00
ATOM	8	CE2	TYR	50	22.605	20.023	7.958	1.00	1.00
ATOM	9	CD2	TYR	50	23.184	20.197	9.223	1.00	1.00
ATOM	10	CB	ASP	47	22.867	22.718	4.222	1.00	1.00
ATOM	11	3HB	ASP	47	23.429	21.934	3.712	1.00	1.00
ATOM	12	H	PHE	49	22.720	23.813	8.383	1.00	1.00
ATOM	13	N	PHE	49	23.006	23.860	9.356	1.00	1.00
ATOM	14	2HB	SER	118	23.899	10.386	1.572	1.00	1.00
ATOM	15	H	PHE	108	23.719	16.347	14.829	1.00	1.00
ATOM	16	3HB	SER	118	25.084	9.101	1.888	1.00	1.00
ATOM	17	C	SER	118	24.801	10.547	4.156	1.00	1.00
ATOM	18	O	SER	118	24.731	11.745	3.883	1.00	1.00
ATOM	19	HE3	TRP	119	25.251	14.236	5.359	1.00	1.00
ATOM	20	HZ3	TRP	119	25.375	16.684	4.940	1.00	1.00
ATOM	21	O	THR	51	24.888	17.629	15.476	1.00	1.00
ATOM	22	C	THR	51	25.298	18.781	15.594	1.00	1.00
ATOM	23	H	TRP	119	25.759	9.071	5.166	1.00	1.00
ATOM	24	N	TRP	119	25.732	10.063	4.976	1.00	1.00
ATOM	25	CE3	TRP	119	25.984	14.834	5.883	1.00	1.00
ATOM	26	HA	LYS	107	25.844	15.590	14.674	1.00	1.00
ATOM	27	CZ3	TRP	119	26.054	16.222	5.648	1.00	1.00
ATOM	28	HA	PRO	46	26.314	24.853	1.902	1.00	1.00
ATOM	29	C	VAL	38	25.878	24.817	15.661	1.00	1.00
ATOM	30	2HB	TYR	48	26.323	25.590	8.200	1.00	1.00



## Docking Result:

The docking results will be emailed back to you in the form of ten docked structures and their corresponding binding free energies in kcal/mol. The ten docked structures represent the different poses of the candidate drug molecule in the active site. User can download each docked complex by clicking at [↓](#) icon.



SCFBio Home About Search Results

### Active Site Prediction

#### RESULTS

Cavity Result Docking Result

S.No.	Structure	Predicted $\Delta G$	Cavity	Download
1	Structure 1	-7.57	Cavity9	↓
2	Structure 2	-7.40	Cavity7	↓
3	Structure 3	-7.32	Cavity10	↓
4	Structure 4	-7.31	Cavity6	↓
5	Structure 5	-7.09	Cavity2	↓
6	Structure 6	-7.03	Cavity4	↓
7	Structure 7	-6.97	Cavity2	↓
8	Structure 8	-6.84	Cavity2	↓
9	Structure 9	-6.82	Cavity6	↓
10	Structure 10	-6.80	Cavity1	↓