

MOLECULAR DOCKING CALCULATIONS UTILIZING DISCOVERY STUDIO & PIPELINE PILOT (October 21st, 2021)

Dr. Nikola Minovski Theory Department Laboratory for Cheminformatics National Institute of Chemistry 1000 Ljubljana, Slovenia e-mail: <u>nikola.minovski@ki.si</u>



Content

A quick introduction to Discovery Studio (DS) user interface	3
DS protocols for structure-based design (SBD)	4
Data accessibility and preparations for structure-based calculations	5
Parameterization of SBD protocols for molecular docking calculations	0
Analysis of docked poses (RMSD of re-docked ligand molecule)1	.0
A quick introduction to Pipeline Pilot (PP) user interface1	.3
Preparation of a ligands library for molecular docking calculations 1	.4
Molecular docking calculations using DS CDOCKER protocol 2	20
Analysis & selection of the most favorable docked poses	22



A quick introduction to Discovery Studio (DS) user interface

- 1. Tool panels (Macromolecule, Simulation, Receptor-Ligand Interactions, etc.)
- 2. Toolbars
- 3. Tool panel collections
- 4. Hierarchy tree view
- 5. Molecule window
- 6. Molecular detail and navigation panel
- 7. Protocol parameter window
- 8. Jobs explorer
- 9. Results window
- 10. Server status

DS protocols for structure-based design (SBD)

Discovery Studio is equipped with six different protocols for structure-based drug design:

- 1. CDOCKER (a grid-based molecular docking method that employs CHARMm forcefield)
- 2. GOLD (genetic algorithm for docking flexible ligands into a protein binding site)
- 3. LibDock (algorithm for docking small molecules into an active receptor site)
- 4. LigandFit (shape-based docking with Monte Carlo conformational sampling)
- 5. CDOCKER with structure-based pharmacophore restraints
- 6. Flexible Docking (CDOCKER-based method allowing protein side-chains flexibility)

These protocols can be accessed under **Protocols tab** > **Discovery Studio** > **Receptor-Ligand Interactions** > **Docking subdirectory**.



Note: The Protocols tab (DS collections) might be unavailable in case the DS Client is not connected to the AEP server (see AEP server status in the right-down corner of the DS-GUI) due to several reasons:

- The client-server connection is down because of technical issues.
- Another user is connected to the AEP server (licensing issues).

In case you are connected to the AEP server and the Protocols tab are not visible, you can enable it manually (**View** > **Explorers** > **Protocols**).

Data accessibility and preparations for structure-based calculations

In order to perform structure-based (molecular docking) calculations in DS, the accessibility to structural data (e.g., a protein, a ligand, or ideally a protein-ligand complex) of the biological system under consideration is necessary (e.g., protein database repository – <u>www.rcsb.org</u>). This can be done by downloading the X-ray, NMR or CryoEM structure manually from the <u>www.rcsb.org</u> repository or in an automated fashion directly through the DS interface.

To automatically fetch the protein structure from the PDB repository in the DS environment, follow these two steps (see picture below):

Di	scover	ry Studi	o 4.1 (lient	1.00												
File	Edit	View	Chen	nistry	Stru	cture	Seque	ence	Chart	Scr	ipts	Tools	Wir	ndow	He	lp	
Ma	acrom	olecul	es Si	mula	tion	Rece	eptor-l	ligan	d Inte	racti	ions	Pha	mac	opho	res	Sma	ill I
**	New	- 📄	8.		•	X	≫ 🗄	` 🗍	@	<u>1</u> , D	isplay	Style.	. ×	R	æ	÷\$⇒	÷
Ope	n URL	-	1		10	1				Ψ.						x	Ì
															_	_	
ht	tp://wv	ww.rcsb	.org/pd	b/files	/5csw.	pdb										-	
G	enerat	e URL us	sing —														
п): <u>5</u> c	SWV			2				5	Site:	Defa	ult PDB	Struc	tures	•	- I	
	_							_			_						
									Open			Cancel		н	lelp		
			_	_	_						_						J

- 1. Click on the **Open URL** button on the toolbar.
- 2. Write-down the PDB **ID** of the needed structure and click on the **Open** button.

In this way, the downloaded structure will be automatically displayed in the **Molecule window** as **solid-ribbon** (see picture below).



In the next step, the downloaded protein structure should be prepared for structure-based calculations. To do this, click on the **Macromolecule tool panel** and open the **Prepare Protein** tools collection. Under the **Automatic Preparation** section click on **Prepare Protein** button. This will open the **Prepare Protein** parameter dialog (see picture below).



The default **Prepare Protein** settings usually give good results (e.g., missing loops will be added and minimized, crystal waters will be removed, the entire protein structure will be protonated, etc.).

As demonstrated on the picture above, **Keep Water** parameter is set to **None** by default that is good for most of the molecular docking calculations. Optionally, this parameter can be changed only in cases where some water molecules (usually surrounding the ligand in the binding pocket) are necessary for establishing ligand-protein interaction.

When all parameters are defined, just click on the **Run** button and check when Prepare Protein job is done (check the **job status** in **Jobs explorer**).

When Prepare Protein job is finished, double click the item in the Jobs explorer. This will open the **Results window** containing the job results. To visualize the prepared protein, click on the **Prepare Protein Structure** link or **View Results** button. The prepared protein structure will be opened in a **new Molecule window** (see next picture).



At this step, it is a good advice to save the prepared protein structure for future work (**File** > **Save As**...) as *.dsv or *.pdb format.

Once the protein structure is prepared (in this case a protein-ligand complex), the next step is to prepare it for structure-based (molecular docking) calculations. To easily access all the protein/ligand components, first close the **Results window**, click on the **Molecule window** where the protein structure is displayed and open the **Hierarchy tree view** by pressing **Ctrl + H** on the keyboard.



It should be noted however that during the Protein preparation step any other co-factors (e.g., metal ions) are not automatically removed from the structure and we should do it manually. This is somehow similar to the elimination of the water molecules described previously. Put differently, in the absence of a solid background to keep such co-factors and to use them in our calculations, we should remove them from the structure (see previous picture). In **Hierarchy tree view**, identify the unwanted co-factors, select them and press Delete (e.g., Cl⁻ ion in the above example). Moreover, in the **Hierarchy tree view**, click on **<Cell>** and go to main menu (**Structure > Crystal Cell > Remove Cell**). This will remove the crystal cell covering the entire crystal structure.

Now we should properly label the structure to be used for structure-based calculations. In the **Hierarchy tree view**, select the main structure tree (in our example it is named as 5CSW). Click on the **Receptor-Ligand Interactions** tool panel, go to **Tools** tab, open **Define and Edit Binding Site** tool collection and click on **Define Receptor** button. The structure label will appear in the **Hierarchy tree view** (see 5 on picture below).



While we are still in the **Hierarchy tree view**, we should also identify the ligand molecule in order to define the **binding site sphere** where docking calculations will be performed. The **binding site sphere** can be defined by selecting a ligand atom, a fragment or the entire ligand molecule in the **Hierarchy tree view**. Now under **Define Site** tools collection click on the **From Current Selection** button to create the **binding site sphere** (see next picture).



As demonstrated on the picture above, a new SBD_Site_Sphere item will be defined under the 5CSW tree in the Hierarchy tree view. Right-click on SBD_Site_Sphere item (blue sphere icon) and then click on Attributes of SBD_Site_Sphere. This will open a dialog named SphereObject Attributes where some default parameters parameters (XYZ coordinates and radius of the sphere) can be customized (see previous picture).

Now in **Hierarchy tree view** select again the ligand molecule and remove it from the structure. This can be done by a right click on the ligand in the **Hierarchy tree view** and click **Cut** on the shortcut menu. Open a new **Molecule window** (**File** > **New** > **Molecule Window**) and paste the ligand molecule here.

Visually inspect the ligand in the **Molecule window** for possible errors (e.g., atom valences), rename the ligand in **Molecular detail and navigation panel** (e.g., **5CSW_ref**) and save the ligand molecule as ***.mol** or ***.sdf file** (File > Save As...).

When everything is properly prepared, the system is ready for the next step (Parameterization of SBD protocols for molecular docking).

Parameterization of SBD protocols for molecular docking calculations

The parameterizations of the DS SBD protocols (e.g., CDOCKER, GOLD, etc.) for molecular docking calculations are mainly automatic. To open a SBD protocol switch to **Protocols tab** and navigate to the SBD protocol of your choice (**Protocols tab** > **Discovery Studio** > **Receptor-Ligand Interactions** > **Docking**). Double click on the selected SBD protocol (e.g. CDOCKER). The CDOCKER SBD protocol parameter window will open (see picture below).

Dock Ligands (CDOCKER) 🔀	
Parameter Name	Parameter Value
Input Receptor	5CSW_prep:5CSW
Input Ligands	Molecule:All
Input Site Sphere	-15.8063, -1.20024, 13.495, 9.19538
Top Hits	3
Random Conformations	10
 Orientations to Refine 	0
 Simulated Annealing 	True
Advanced	
Parallel Processing	False

Parameterize the protocol for your needs and when everything is prepared just click **Run** (green play button) in the **Toolbar**.

When the calculation is finished, open the Results in exactly the same way as for the Prepare Protein job explained above.

Analysis of docked poses (RMSD of re-docked ligand molecule)

To analyze the docked poses click on **Docked Ligands** link in the Result window. This will open the docked poses in a new Molecule window. In order to compare how good our results are, we need to insert here the co-crystallized ligand we saved previously (**File** > **Insert From** > **File**). Rename the inserted original ligand molecules (e.g., **5CSW_ref_orig**) and enable visibility of all docked poses by putting check marks on visibility option (see picture below).

<mark>%</mark> 50	SCSW_prep 🛛 😨 Molecule (1) 🖾 🗸 🗸 🖓									
Ŧ	Index	Name	Visible	Tagged	Visibility Locked	TopHits	-CDOCKER_ENERGY	-CDOCKER_INTERACTION_ENERGY	POSE_NUMBER	Resolution I
_	1 1	5CSW_ref	✓ Yes	No	No	3	38.6522	66.2624	1	2.66
+	2 2	5CSW_ref	✓ Yes	No	No	3	37.9031	66.0536	2	2.66
-	3 3	5CSW_ref	✓Yes	No	No	3	36.3867	65.9034	3	2.66
•	4 4	5CSW_ref_orig	✓ Yes	No	No					2.66
Ŧ	inserte	ed original	co-crys	tallized	ligand (rena	med)				
.GT										
<u> </u>										

Now select the inserted original co-crystallized ligand (4th row in the previous picture) and label it as a reference molecule (main menu > **Structure** > **RMSD** > **Set Reference**) and open the **Molecule window**.

	<mark>%</mark> 5	csw	/_prep 🗵	😵 Molecule 🗵) 💽 Mol	ecule (1) 🔀]			click here to ope	n the Molecu	ile window	$\langle \bigcirc \rangle$	
	Ŧ		Index	Name	Visible	Tagged	Visibility Locked	TopHits	-CDOCKER_ENERGY	-CDOCKER_INTERACTION_ENERGY	POSE_NUMBER	Resolution	REMARK99	SEQRES
Ш	-	1	1	5CSW_ref	✓ Yes	No	No	3	38.6522	66.2624	1	2.66	REMARK 99 M REMARK 99 5	SEQRES
Ш	+	2	2	5CSW_ref	✓ Yes	No	No	3	37.9031	66.0536	2	2.66	REMARK 99 M REMARK 99 5	SEQRES
Ш		3	3	5CSW_ref	Yes	No	No	3	36.3867	65.9034	3	2.66	REMARK 99 M REMARK 99 5	SEQRES
Ш	•	4	4	5CSW_ref_orig	Yes	No	No					2.66	REMARK 99 M	SEQRES



In order to calculate the RMSD (it should be <= 2.0 Å), go to main menu (**Structure** > **RMSD** > **Heavy Atoms**). This will open the RMSD Report window where all RMSD comparisons are displayed.

5CSW_ref_orig - RMSD Report 🔀		
Heavy Atom RMSD to 5CSW_re	f_orig 4	
Name SCSW_ref 1 SCSW_ref 2 SCSW_ref 3 SCSW_ref_orig 4	Reference SCSW_ref_orig 4 SCSW_ref_orig 4 SCSW_ref_orig 4 SCSW_ref_orig 4	RMSD (A) 0.4246 0.5746 1.2021 0.0000

As demonstrated, all docked poses have RMSD below 2.0 Å (that is also visually true), i.e., the docking protocol is properly parameterized and ready to be used for performing molecular docking calculations on new compounds.

While still open, switch to **5CSW_prep** tab and save as a new *.dsv file (e.g., **5CSW_prep_final.dsv**).

A quick introduction to Pipeline Pilot (PP) user interface

Pipeline Pilot Professional Client - (MOL2SDF)				8 8 0
<u>File Edit View Icols Window Help</u>		1		- <i>6</i> ×
🗋 🗟 🥝 🔗 X X 🕸 🖄 🖉 🤉 🖃 🗒	🕻 🕨 🕾 🖉 🗶 📾 📓 👭 📓 🖉 🖉 🖉	arch protocol database		
Components 🔗 🛙	Favoritas 🍙 Most Used			
2				
III Contract			3	
Contraction and Application Integration Contraction and Statistics	G A:=B>—G >—G			
Construction Reporting and Visualization	SD Booker Name - Add Budrasena 30 Core	diades Mainize Keen Properties 52 Witer		
Chemistry	SourceTag:	Malecule		
Data Access and Manipulation Converter				
B Co Filters				5
Manipulators		PilotScript		8 9 8
B CO 3D Methods		FS: Keywords and Functions F9: Proper	ty Names	
B CSR		P4: Calculable Properties (P6: Dy Category, P7: Refresh List		
Group Features		X 10 10 Q 4: 14 99 13 E D 1 16 2	10 80	
-C 2D Coords 2D Coords (Advanced)		Name (= SourceTag)		
Add Bond Orders				Cancel
- Add Hydrogens				
Aggregate Fragments				Apply
- Clean Molecule				Help
- Convert Fingerprint				
- Highlight Molecular Features				
- Ionize Molecule at pH				
- Keep Largest Fragment				
Molecular Pharmaconhare Features				
Normalize Structure (Cheshire)				
Perceive Hydrogen Bonds				
Protonate Acids Remove Hydroneous				
42 Separate Fragments				
- Standardize Molecule				
Translate Chemical Nome				
Readers				
🔅 😂 Viewers				
Writes Database Content				
Library Design				
🐵 😂 Molecular Toolkit				
Pharmacophore Pharmacophore				
B-D Quantum Chemistry				
- Con Reactions and Enumeration				
Receptor-Ligand Interactions		Initial Expression Expression Final Expression		Gude M
Search and similarity Search and similarity				
🗄 🎦 USRies				
is 😄 Biology				
in Controly				-
Documents and Text	- Bapsed Time			
🚴 nikolam 🕬 Protocols 🦲 Components	8 ⁰ MOL2SDF			4
lobs				3 Peremeters
Protocol Name		Status	Start Date 🗢 🌀	Bipension Name ::: SourceTag:
P MOL2SOF		Not Started		Initial Expression
Deck Ugends (CDOCKER)		Finished	Thu Mar 06 13:47:32 2018	Hind Expression A
Find Novel Molecules Example Devid Medawide Example		Finished	Wed Dec 06 14:51:40 2017	The Contract of the Contract o
Frank Structure Man Example		Finished	Wed Dec 06 11:26:33 2017	
Exact Structure Map Example		Finished	Wed Dec 06 08:00:30 2017	
■ 02 Filtering Reactive Molecules using SMARTS		Finished	Wed Dec 06 07:57:29 2017	
New Protocol4		Finished	Wed Dec 06 07:47:17 2017	
Pgp_SOM_hits_Extractor		Finished	Tue Dec 05 16:24:30 2017	
Jobs Help Auto Watch Debug Messages				
Ready				😞 inikelami localitoni 📒 92.0

- 1. Main PP tool panel
- 2. Components collections
- 3. PP protocol (assembled of piped components)
- 4. Protocol parameters
- 5. PilotScript interface
- 6. Jobs explorer

Preparation of a ligands library for molecular docking calculations

Regardless of the origin of ligands we would like to dock (e.g., commercially available compounds, ligands scrutinized from the literature, or even *de novo* virtually assembled ligands not commercially available), they should be properly pre-treated for molecular docking calculations. This means that all of them should be adequately prepared in a machine-readable format (e.g., *.mol, *.sd/*.sdf, SMILES, etc.). Comparing to the *.mol file format (commonly known as a connection table) that contains structural information (e.g., atoms, bonds and their connectivity) related to a single ligand, the *.sdf file format in a broader sense can be regarded as a database of significantly higher number of ligands. In fact, *.sdf format contains the same information as a single *.mol file, where each ligand (encoded in form of connection tables) are separated by "\$\$\$\$" delimiter. Moreover, these file formats can store various information related to the ligand(s) (e.g., various physico-chemical properties, chemical names, etc.). A snippet of these file formats (e.g., molecule of toluene) is given in the picture below.



- 1 Molecule name: toluene
- 2 Number of atoms and bonds in the molecules: 7 atoms and 7 bonds
- 3 Format limitation: maximum 999 characters
- 4 Mol file format: V2000 or V3000 (for larger structures)
- 5 Cartesian coordinates: 2D structures (X and Y) or 3D structures (X, Y, and Z)
- 6 Atom symbol: C, H, ...
- 7 Connectivity information: how the atoms in the molecule are inter-connected
- 8 End of the connection table: M END
- 9 Property name and value: CLOGP = 2.588
- 10 Delimiter "\$\$\$\$": without (*.MOL file; one molecule per file), with (*.SD/*.SDF file format; one or more molecules per file)

Assuming that our ligands are represented as 2D structures (see previous picture) it should be stressed that such a representation is not very recommended for performing molecular docking calculations. Put differently, prior to perform molecular docking calculations, all ligands should be:

- Converted to 3D structures,
- All hydrogen atoms should be added (2D structures are lacking of H-atoms),

and optionally

• They should be geometrically minimized (e.g., to remove any van der Walls clashes).

Finally and also preferably, all ligands should be assembled as one database in an appropriate molecular machine-readable formed (e.g., *.sdf). In order to do all these crucial steps in a concomitant fashion, we will construct a Pipeline Pilot (PP) protocol (see picture below).



Let's assume that our ligands we scrutinized from different sources (e.g., literature, online structural databases, etc.) represent separate 2D molecules, which are either in *.mol or *.sdf format. In order to construct a PP protocol (like the one represented above) that will automate all the steps described above, we will use PP interface.

So, here is the step-by-step guide:

As a very first step, we should read all the ligands we collected (*.mol or *.sdf files). Since, both file formats are practically very similar, we will use the so-called **SD Reader** PP component, which is capable to read both formats.

To do this, in the **Components Collection** tab, locate **SD Reader** component (stored in **Chemistry > Data Access and Manipulation > Readers**) and double-click on the **SD Reader** component to put it on the working layout (see picture below). At first, the **SD Reader** component is red, which means that we should add some data to it. Select the red **SD Reader** component on the working layout and in the **Parameters** window, under **Source** click the corresponding button as depicted in the picture to open the file browser dialog.



Navigate to the directory where your ligands are stored, select all ligands, and click on the **Select** button to select them. In this way, the **SD Reader** component will become blue, while in the **Parameters** window all *.mol and *.sdf files we selected will be separated by commas. The parameters for a particular component in the **Parameters** window will be visible only in case the target component is selected in the working layout!

As we can see on the picture above, the filenames of our ligands are actually their names, so it is not a bad idea to use these information in the final *.sdf file (the ligand library we are constructing) for easier ligand's recognition. While we have still open the **Parameters** for the selected **SD Reader** component, locate the **SourceTag** parameter and from the corresponding list select **Filename** (see picture below).

Parameters		
Source	C:\Users\nikolam\Desktop\CDOCKER\lib\Encorafenib	.sdf,C:\Users\nikolam\Desktop\
Maximum		
SourceTag	Filename	-
Keep Properties		
UTF-8 Auto Detect	False	•
Additional Options		

Now, navigate to the **Components Collection** tab, locate the **Custom Manipulator (PilotScript)** component (stored in **Data Access and Manipulation** > **Manipulators**) and double-click it to automatically pipe it to the **SD Reader** component (see picture below).



Once piped, select the **Custom Manipulator (PilotScript)** component and in the **Parameters** window click the **Expression** button to open the **PilotScript** interface. Here, define the **Name** for each ligand as demonstrated on the picture above (**Name := SourceTag**). In this way we constructed a new property named "**Name**" that contain names of our ligands. Check if the **PilotScript** is correct (5), click on the **Apply** button (6) and finally click on the **OK** button (7) to close the **PilotScript** interface.

Now navigate again to the **Components Collection** tab and find the following components: **3D Coords** (located in **Chemistry > Data Access and Manipulation > Manipulators > 3D Methods**), **Add Hydrogens** (located in **Chemistry > Data Access and Manipulation > Manipulators**) and **Minimize Molecule** (located in **Chemistry > Data Access and Manipulation > Manipulators**) and **3D Methods**). Double-click each of these components in the same order (**3D Coords > Add Hydrogens > Minimize Molecule**) to pipe them automatically (see next picture). So, the **3D Coords** component will update the 2D connection table of the input ligands with Z-coordinate, the **Add Hydrogens** component will also update each input ligand by adding H-atoms, while **Minimize Molecule** will minimize the energy of all input ligands.

Pipeline Pilot Professional Client - [New Protocol2*]								• 🕺
File Edit View Tools Window Help								- 7 X
	N (C) (A) (C) (A) (C) (A)		4					
🗍 🗖 🗑 🙆 📐 V 😤 🖻 🖻 🏡 🖓 📷 🖼		🔽 🗗 🙆 🚕	Search protoc	ol database				
Components 😥 🗵	Favorites 🙀 Most Used							
1								*
Data Access and Manipulation								
Converters				ſL				
🕀 💼 Filters		r 4 1	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	ľ				
🚊 🗁 Manipulators	SD Reader Name	3D Coordina	tee Add Hydropens	linimize				
🖨 🗁 3D Methods	SourceTag			lolecule				
- 3D Conformations		_						=
3D Coords								
Minimize Molecule 3								
H-C Salts								
E-CSR								
B-G Sgroup reatures								
2D Coords (Advanced)								
Add Bond Orders								
Add Hydrogens 2								
Aggregate Fragments								
Center Molecule								
Clean Molecule								
Convert Fingerprint								
Deprotonate Bases								
82 Generate Fragments								
Highlight Molecular Features								
Ionize Molecule at pH								
Merge Molecules								
Molecular Pharmacophore Features								
Derseine Hudrogen Ronds								
Perceive Hydrogen bolids Protonate Acide								
Remove Hydrogens								
Separate Fragments								
Tile Fragments								-
Translate Chemical Name					III			•
🔹 🗁 Readers 🔹 👻		Elapsed Time:						
Anikolam & Protocols Components	SP MOL2SDF SP N	New Protocol2*						
lohs			X	Parameters				
				T of officers				
Protocol Name	5	status	Start Date 🗸 🏠					
30 MOL2SDF	•	Not Started						
P New Protocol2	N	Not Started				There are no parameters defined.		
New Protocol1	F	inished	Tue Mar 13 08:52:		TO	create parameters, right-click in this pane and select "Edit".		
Dock Ligands (CDOCKER)	F	inished	Thu Mar 08 13:47: =					
 Find Novel Molecules Example 	F	inished	Wed Dec 06 14:51:					
 Find Novel Molecules Example 	F	inished	Wed Dec 06 13:31:					
Exact Structure Map Example	F	inished	Wed Dec 06 11:26:					
Exact Structure Map Example	F	inished	Wed Dec 06 08:00:					
	F	inished	Wed Dec 06 07:57:					
New Posts and Jaka Halp Auto Watch Debug Merrager			WILL D 06 07.47.	Decements	Implementation Web Card			
Jobs Help Auto Watch Debug Messages				Parameters	amplementation web Servi	e		
Ready							😹 nikolam localhost 📒	9.2.0

Before to finalize the protocol, we should keep in mind that the final *.sdf library can contain some residual properties (e.g., SourceTag, Minimized Energy) that will be generated when the protocol will be run.

Except the "Name" property we created, we actually don't need any other property. In order to keep just the "Name" property, we will add one more component (double-click on it) named **Keep Properties** (located in **Data Access and Manipulation** > **Manipulators**) to our protocol (see next picture).

After adding the **Keep Properties** component, select it in the working layout and in the **Parameters** window type the name of the property (e.g., **Name**) for the needed **PropertyList** parameter (see next picture). In this way our *.sdf library will contain only one property, i.e., "**Name**".

Pipeline Pilot Professional Client - [New Protocol2*]	
<u>Eile Edit View Tools Window H</u> elp	_ # ×
🕒 🖬 🥩 🥹 🦻 🗶 🐰 🗅 🖏 🌮 🔿 100 🗨 🖽 🎇	🕨 🛞 🛆 🔲 🛤 📓 100 🤜 🕀 🗇 🗇 🗇 Search protocol database
Components 🖉 🗵	Favorites 🍘 Most Used
	A 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Data Access and Manipulation Convertes Grovertes Grover	2 SD Reser Name * 30 Coordinates Add Hydrogens Minicica Boorder Boorder * Bo
Database and Application Integration Analysis and Statistics	
Reporting and Visualization	۲
🗄 🧁 Chemistry 💌	Finished Elapsed Time: 14 Seconds
Anikolam Protocols Components	3 ²⁰ MOL2SDF 3 ²⁰ New Protocol2*
Jobs	Parameters 3
Protocol Name	Status Statt Date 🧮 All PropertyList Name
BB MOL2SDF	Not Started

Finally, we will finish our protocol by adding the last component named **SD Writer** (located in **Chemistry > Data Access and Manipulation > Writers**).

Pipeline Pilot Professional Client - [New Protocol2"]				
Eile Edit View Iools Window Help	4			_ # ×
🗋 🖬 🥩 😢 🕫 X X 🖻 🗟 🖉 🤉 100 💌 🖾 🔀	🕨 🖉 🛄 👫 📓 100 🔽 🕄 📚	🤃 🖒 Search protoc	ol database	
Components 🥥 🗵	Favorites 🍙 Most Used			
]			<u>^</u>
Data Access and Manipulation Database and Application Integration Analysis and Statistics Analysis and Statistics Deta Access and Manipulation Chemistry Data Access and Manipulation Deta Access and	SD Reader Name := 30 Coordin	tes Add Hydrogens N	Instruct Kep Projection Projection	
Database Content				
🖶 🗁 Library Design				
🕀 🧰 Molecular Toolkit				•
Pharmacophore	Finished Flanced Times	14 Seconds		,
nikolam 20 Property Calculators	Pinished Elapsed Time:	14 Seconds		
Components	B [®] MOL2SDF B [®] New Protocol2^			
Jobs		×	Parameters	3
Protocol Name	Status	Start Date 🧠 🔺	Destination	C:\Users\nikolam\Desktop\b-raf_inhibitors.sdf
8 MOL2SDF	Not Started		SD Format Options	Write Structure Write Properties Write Date Write Query Info Wrap Long Property
🕫 🖪 New Protocol2	Finished	Tue Mar 13 13:23:	Additional Options	
New Protocol1	Finished	Tue Mar 13 08:52:		
Dock Ligands (CDOCKER)	Finished	Thu Mar 08 13:47:		
Find Novel Molecules Example	Finished	Wed Dec 06 14:51:		
Find Novel Molecules Example	Finished	Wed Dec 00 13:31:		
Exact Structure Map Example	Finished	Wed Dec 06 08:00:		
02 Filtering Reactive Molecules using SMARTS	Finished	Wed Dec 06 07:57:		
Marco Data and	Protocolo d	W-J D 06 07.47. T		
Jobs Help Auto Watch Debug Messages Ready				😹 nikolam localhost 📕 9.2.0 🔮

Select **SD Writer** component in the working layout and in the **Parameters** window, under the **Destination** parameter, type the full path and file name with extension (e.g. C:\Users\username\Desktop\library_filename.sdf), where username is your username and library_filename.sdf is the name of the ligand library in *.sdf format (e.g., in this case b-raf_inhibitors.sdf). Finally, run the protocol by clicking the green play button from the main PP tool panel (3) and wait for it to finish. Moreover, save the protocol for further use (goto File > Save Protocol As...) in your home directory.

Molecular docking calculations using DS CDOCKER protocol

The procedure of performing molecular docking calculations for our previously constructed ligand library is actually very similar to the re-docking of the co-crystallized ligand conformation within its binding site (see pages 9-10). Previously, we saved our molecular docking settings as a *.dsv file named **5CSW_prep_final.dsv**, which contains all the important information we obtained from the re-docking validation step (e.g., prepared protein, defined appropriate calculation sphere as well as the corresponding calculation parameters).

In order to perform molecular docking calculations for our previously constructed ligand library, in Discovery Studio first we will open our well-validated system (goto **File > Open**... and then locate and load the file **5CSW_prep_final.dsv**). After this, we will open our prepared ligand library as *.sdf format (**b-raf_inhibitors.sdf**) in a new **Molecule window** (see picture below).

N Discovery Studio 4.1 Client							
File Edit View Chemistry Structure Sequence Cha	rt Script	s Tools	Window Help				
Macromolecules Simulation Receptor-Ligand In	teractio	s Pharm	acophores S	mall Molecu	ules X-ray	My Tools	
🕴 🍕 New 🔹 📄 🤮 🔚 🦘 🥐 💢 % 📅 📋 🐵	🛃 Disp	lay Style	- i k & +	∲ 😪 🏠	💥 💥 🗋	n i Non-bond Inter	actions
Tools 🛛 🛛 Protocols 💟 🛛 Files 💟 📮	😵 5CS	W_prep_fina	al 🗵 🛛 😵 b-ra	af_inhibitors [3		
🗄 🧿 Discovery Studio					_		
🗄 🧕 nikola.minovski's protocols	Ŧ	Index	Name	Visible	Tagged	Visibility Locked	
		1 1	Encorafenib	✓Yes	No	No	
	+	2 2	GDC-0879	No	No	No	
	- 1	3 3	Lifirafenib	No	No	No	
	•	4 4	PLX-4720	No	No	No	
	±	5 5	Sorafenib	No	No	No	
		5 6	Vemurafenib	No	No	No	
	50						

Next, switch back on the protein (click on **5CSW_prep_final** tab). From **Protocols tab** navigate to the SBD protocol of your choice (**Protocols tab** > **Discovery Studio** > **Receptor-Ligand**

Interactions > **Docking**). Double click on the selected SBD protocol (e.g. CDOCKER). The CDOCKER SBD protocol parameter window will open (see next picture).

Dock Ligands (CDOCKER) 🗵									
Parameter Name	Parameter Value								
Input Receptor	5CSW_prep_final:5CSW								
Input Ligands	b-raf_inhibitors:All								
Input Site Sphere	-15.8061, -1.20006, 13.495, 10								
• Top Hits	3								
 Random Conformations 	10								
 Orientations to Refine 	0								
Simulated Annealing	True								
Advanced									
Parallel Processing	False								

Parameterize the protocol as demonstrated on the picture above and click **Run** (green play button) in the **Toolbar** to start the docking calculation for our ligands.

When the calculation is finished, in the **Jobs** tab double-click to open the **Results** (2) and then click on the **View Results** button (3). This will open a new **Molecule window** (4) containing our protein structure and the docked poses (5) for the ligands (see picture below).



Analysis & selection of the most favorable docked poses

For the selection of the most favorable docked poses, first we will focus on the values for the scoring functions (in our case, -CDOCKER_ENERGY and -CDOCKER_INTERACTION_ENERGY) calculated by the CDOCKER protocol. These scoring functions demonstrating the binding affinity of our ligands in kcal/mol give a first insight about the quality of the results (see picture below).

: [[Index	Name	Visible	Tagged	Visibility Locked	POSE_NUMBER	-CDOCKER_ENERGY	-CDOCKER_INTERACTION_ENERGY	LoopList	Resolution	REMARK99	SEQRES	PartialChargeMethod	UseResidueTemplateCharge	Forcefield	ForcefieldBase	Initial RMS Gradie
1	1	5CSW	✓Yes	No	✓Yes				0:5CSW:A:P.	2.66	DEMARK 99 M.	SEQRES 1.	MomanyRone	True	CHARMm	CHARMm	2.46934e+10
1 2	2	Encorafenib	No	No	□No	1	20.4381	47.8047									
3	3	Encorafenib	No	No	□No	2	19.1216	51.0081									
4	4	Encorafenib	No	No	□No	3	-153.575	1.1564									
2 5	5	GDC-0879	No	No	No	1	-16.8529	42.5307									
6	6	GDC-0879	No	No	No	2	-18.0384	38.6495									
Ĵ 7	7	GDC-0879	No	No	□No	3	-18.1339	40.8887									
8	8	Lifrafenib	No	No	□No	1	-34.2423	42.8188									
9	9	Lifirafenib	No	No	No	2	-61.6088	27.3291									
1	.0 10	Ufirafenib	No	No	No	3	-61.7708	27.3665									
1	1 11	PLX-4720	No	No	No	1	22.3926	43.2849									
1	2 12	PLX-4720	No	No	□No	2	21.5333	40.3133									
1	3 13	PLX-4720	No	No	No	3	21.2365	40.241									
1	4 14	Sorafenib	No	No	No	1	16.7086	40.2395									
1	5 15	Sorafenib	No	No	No	2	15.8127	40.5167									
1	6 16	Sorafenib	No	No	□No	3	15.1815	39.458									
1	7 17	Vemurafenib	No	No	No	1	16.3346	45.2359									
1	.8 18	Vemurafenib	No	No	No	2	15.2932	43.5978									
1	.9 19	Venurafenib	No	No	No	3	12.304	42.1518									
									_								
tala a da	ji Ni ang Angar An																

Here, it is important to take into account the pre-defined "minus" sign of these scoring functions (see above picture). Put differently, if -CDOCKER_ENERGY and -CDOCKER_INTERACTION_ENERGY for a docked pose are both positive values, then we are talking about an energetically favorable ligand pose (ligand poses framed in green). Conversely, if negative values are obtained, then these docked poses are energetically unfavorable (ligand poses framed in red) and consequently we can reject them for further assessments.

Since, we are interested solely for the energetically favorable docked poses (framed in green), we can extract them and further analyze visually within the binding pocket (e.g., ligand spatial orientation, interaction of the ligands with the protein's key amino acid residues, etc.).