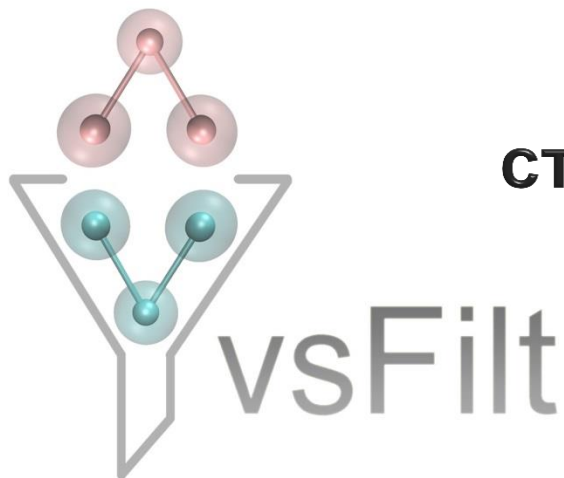




НИИ физико-химической биологии имени А.Н. Белозерского МГУ,
Научно-исследовательский вычислительный центр МГУ



Новый инструмент структурной фильтрации для виртуального скрининга



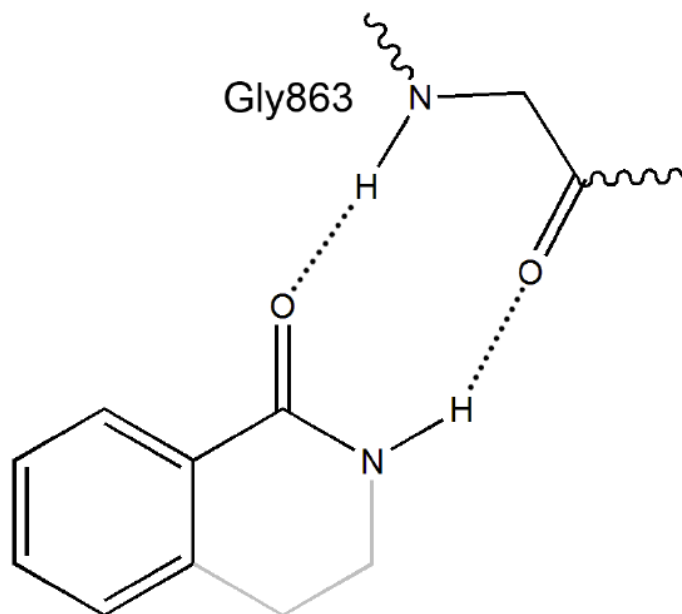
Дмитрий Нилов

К.Х.Н., С.Н.С.

E-mail: nilovdm@gmail.com

Structural filter is defined by a set of interactions:

- typically observed in available structures of protein–ligand complexes;
- considered to play a crucial role in ligand binding



JCIM
JOURNAL OF
CHEMICAL INFORMATION
AND MODELING

pubs.acs.org/jcim

Application Note

vsFilter: A Tool to Improve Virtual Screening by Structural Filtration of Docking Poses

Irina V. Gushchina,[§] Aleksandra M. Polenova, Dmitry A. Suplatov, Vytas K. Švedas, and Dmitry K. Nilov^{*,§}

Cite This: <https://dx.doi.org/10.1021/acs.jcim.0c00303>

Read Online

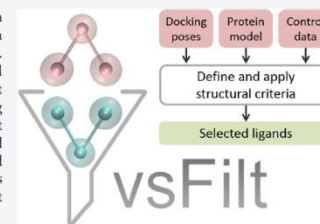
ACCESS |

Metrics & More

Article Recommendations

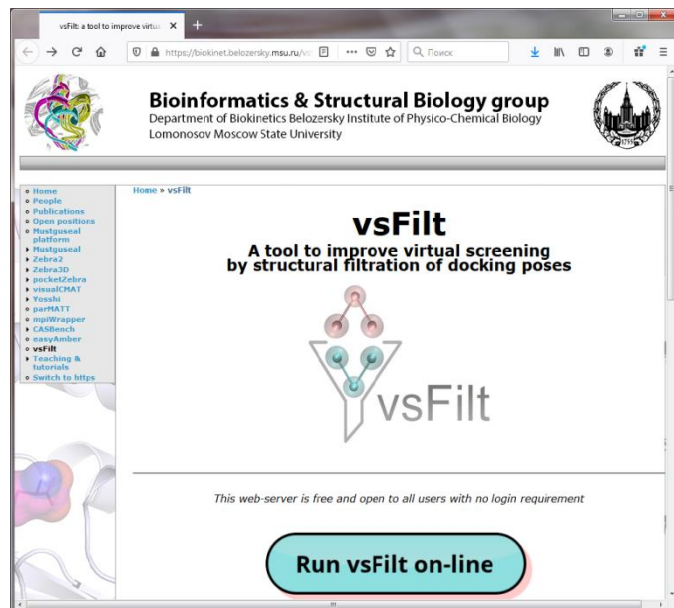
Supporting Information

ABSTRACT: The ability of ligands to form crucial interactions with a protein target, characteristic for the substrate and/or inhibitors, could be considered a structural criterion for identifying potent binders among docked compounds. Structural filtration of predicted poses improves the performance of virtual screening and helps in recovering specifically bound ligands. Here, we present vsFilter—a highly automated and easy-to-use Web server for postdocking structural filtration. The new tool can detect various types of interactions that are known to be involved in the molecular recognition, including hydrogen and halogen bonds, ionic interactions, hydrophobic contacts, π -stacking, and cation- π interactions. A case study for poly(ADP-ribose) polymerase 1 ligands illustrates the utility of the software. The Web server is freely available at <https://biokinet.belozersky.msu.ru/vsfilter>.

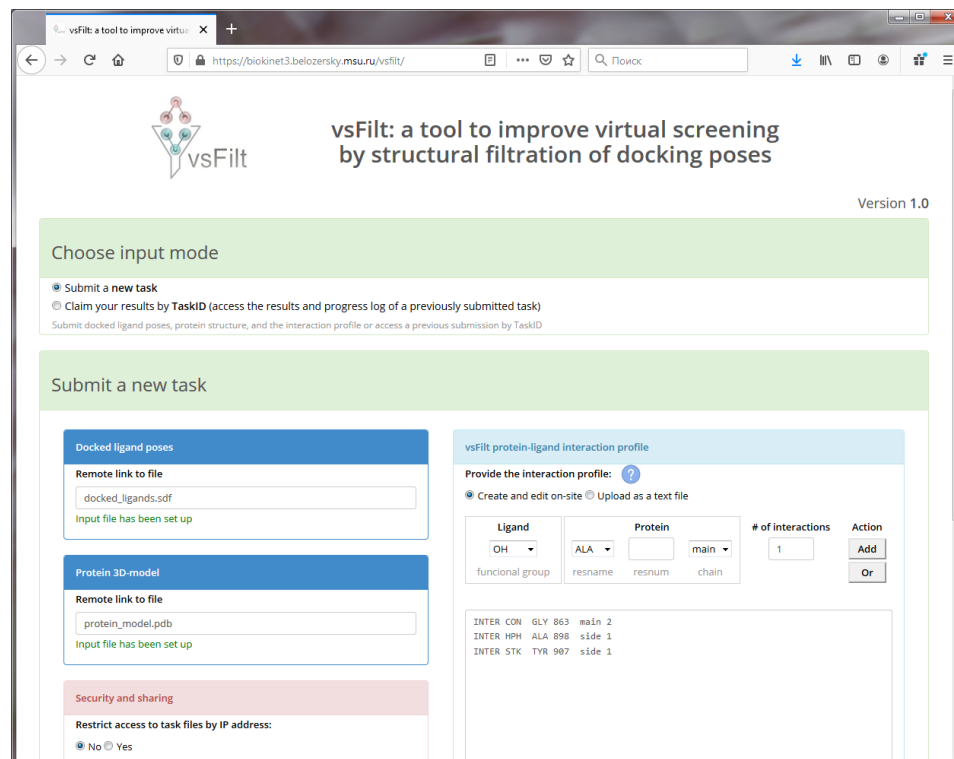


Gushchina et al. *J. Chem. Inf. Mod.* (2020) 60, 3692-2696

https://biokinet.belozersky.msu.ru/vsfilter



The screenshot shows the homepage of the vsFilter web application. At the top, it features the logo of the Bioinformatics & Structural Biology group and the text: "Bioinformatics & Structural Biology group, Department of Biokinetics Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University". The main heading reads "vsFilter: A tool to improve virtual screening by structural filtration of docking poses". Below this is a logo for vsFilter and the text "This web-server is free and open to all users with no login requirement". A prominent button at the bottom says "Run vsFilter on-line". A sidebar on the left contains a navigation menu with items like Home, People, Publications, Open positions, and various software tools.



The screenshot displays the main interface of the vsFilter web application. It includes the vsFilter logo and the title "vsFilter: a tool to improve virtual screening by structural filtration of docking poses" with "Version 1.0" on the right. The interface is divided into several sections:

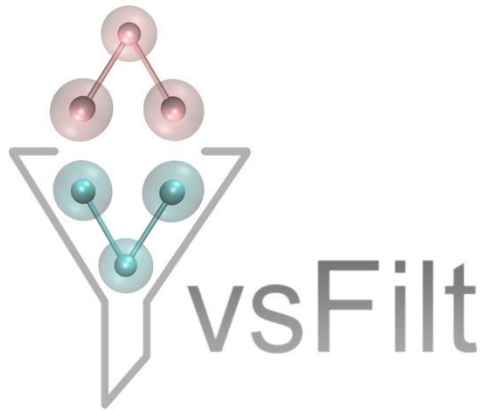
- Choose input mode:** Contains two radio buttons: "Submit a new task" (selected) and "Claim your results by TaskID". Below this is a note: "Submit docked ligand poses, protein structure, and the interaction profile or access a previous submission by TaskID".
- Submit a new task:** This section contains two main input areas:
 - Docked ligand poses:** A text box containing "docked_ligands.sdf" with the note "Input file has been set up".
 - Protein 3D-model:** A text box containing "protein_model.pdb" with the note "Input file has been set up".
- vsFilter protein-ligand interaction profile:** This section allows users to "Provide the interaction profile" by either "Create and edit on-site" (selected) or "Upload as a text file". It features a table for defining interactions:

Ligand	Protein	# of interactions	Action
OH	ALA	1	Add
functional group	resname resnum chain		Or

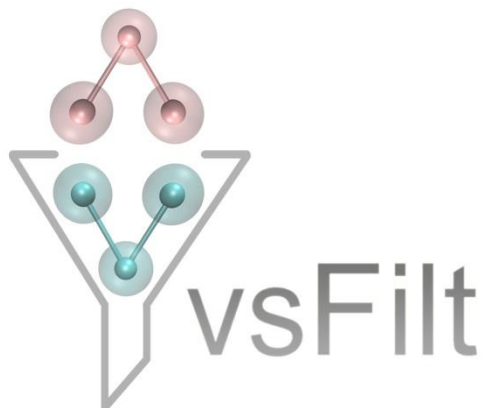
Below the table, a text area shows the resulting interaction profile:

```
INTER CON GLY 863 main 2
INTER HPH ALA 898 side 1
INTER STK TYR 987 side 1
```

- Security and sharing:** A section with a heading "Restrict access to task files by IP address:" and two radio buttons: "No" (selected) and "Yes".



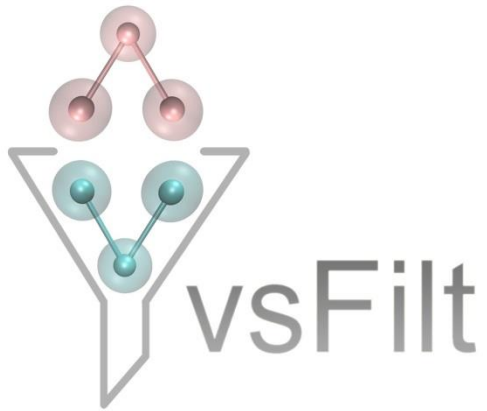
- hydrogen bonds;
- halogen bonds;
- ionic interactions;
- hydrophobic contacts;
- π -stacking;
- cation- π interactions



- hydrogen bonds;
- halogen bonds;
- ionic interactions;
- hydrophobic contacts;
- π -stacking;
- cation- π interactions

Analogs:

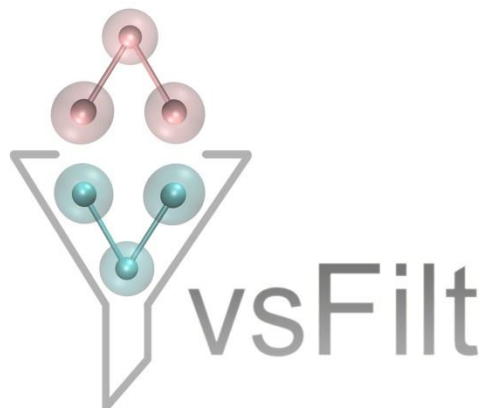
- in-house scripts;
- AutoDock/Raccoon (free);
- Schrödinger/Glide (commercial);
- nAPOLI server



	ligand group	protein residue		number of interactions	
INTER	CON	GLY	863	main	2
INTER	HPH	ALA	898	side	
INTER	STK	TYR	907	side	
OR					
...					

residue number residue chain

Figure 1. Example of control data used by vsFilt for structural filtration.

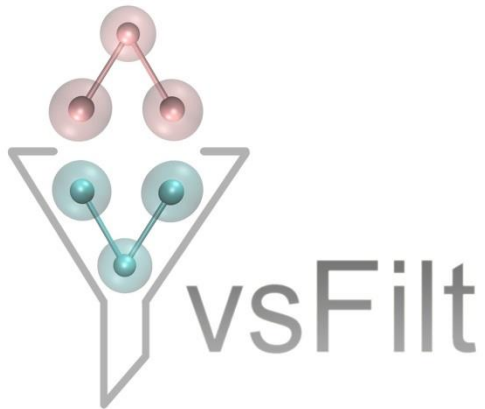


- automatically identifies atoms involved in the interaction;
- sets the corresponding criteria for filtering

	ligand group	protein residue		number of interactions	
INTER	CON	GLY	863	main	2
INTER	HPH	ALA	898	side	
INTER	STK	TYR	907	side	
OR					
...					

Annotations: Blue arrows point from 'ligand group' to 'CON', 'protein residue' to 'GLY', and 'number of interactions' to '2'. Another set of blue arrows points from 'residue number' to '907' and 'residue chain' to 'side'.

Figure 1. Example of control data used by vsFilt for structural filtration.



INTER CON GLY 863 main 2
OR
...
ligand group protein residue number of interactions
residue number residue chain

- recognizes the type of interaction (H-bonding);
- identifies interacting atoms (H-bond donors and acceptors) and their coordinates;
- applies the corresponding distance and angle criteria for structural filtration

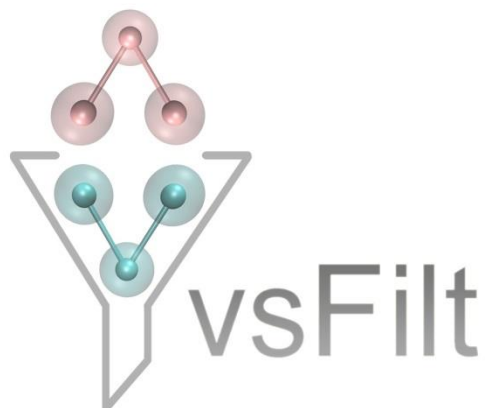
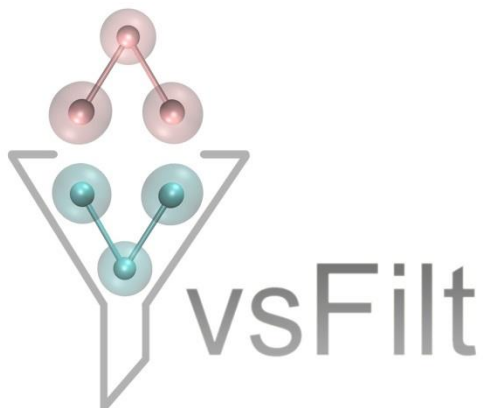


Table 1. Ligand Functional Groups That Can Be Involved in Structural Filtration with vsFilt, and Their Role in the Interaction with Protein

	ligand group	interaction
	OH	hydroxyl
	CO	carbonyl
	COO	carboxyl
	COOC	ester
	COC	ether
	CON	amide
	NH	amino
	NAR	N aromatic
	SO3	sulfo
	SO2N	sulfonamide
	DON	hb donor
	ACC	hb acceptor
	HAL	halo
	HPH	hydrophobic
	STK	aromatic
	PIC	aromatic
		hb donor/acceptor
		hb acceptor
		hb acceptor, anion
		hb acceptor
		hb acceptor
		hb donor/acceptor
		hb donor
		hb acceptor
		hb acceptor, anion
		hb donor/acceptor
		hb donor
		hb acceptor
		halo bond donor
		hydrophobic
		stacking
		cation- π



ligand group protein residue number of interactions

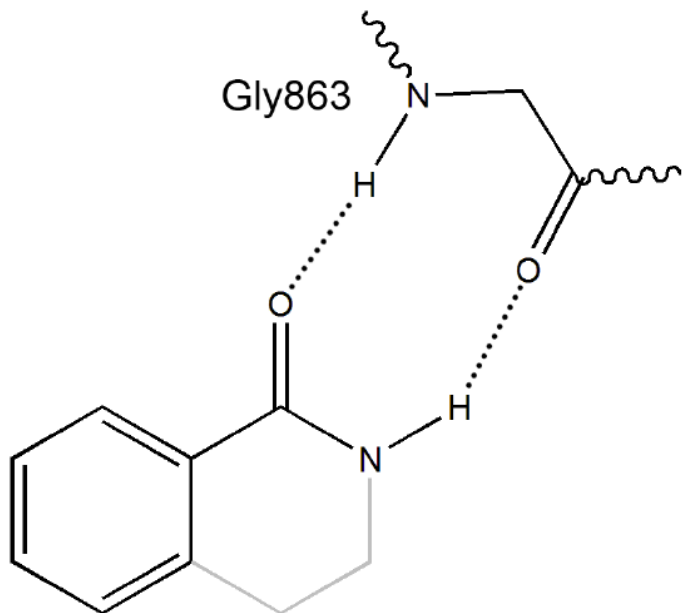
INTER CON GLY 863 main 2

OR

...

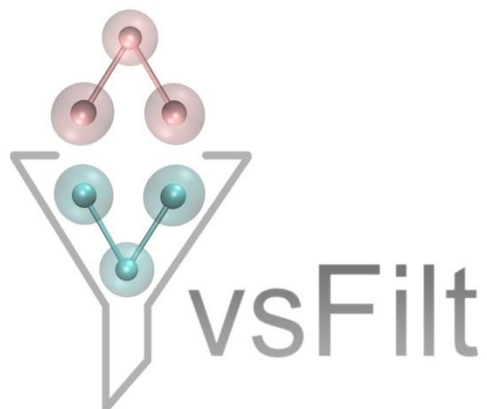
 residue number residue chain

- available residue names: ALA, ARG, ASN, ASP, CYS, GLN, GLU, GLY, HIS, ILE, LEU, LYS, MET, PHE, PRO, SER, THR, TRP, TYR, VAL, HOH (ordered water molecule), CA2, MG2, and ZN2 (Ca²⁺, Mg²⁺, and Zn²⁺ ions);
- main chain/side chain option;
- minimum number of interactions (default = 1) option



	ligand group	protein residue		number of interactions
INTER	CON	GLY	863	main 2
INTER	HPH	ALA	898	side
INTER	STK	TYR	907	side
OR				
...				
			residue number	residue chain

Figure 1. Example of control data used by vsFilt for structural filtration.




Hydrogen bonds


Maximum distances: upper 90% quantiles of H-bond distances obtained from structural database statistics + 0.5 Å

Bissantz et al. *J. Med. Chem.* (2010)

Ligand group	Ligand atom	Protein atom	Max distance, Å	
OH	hydroxyl	O:	3.5	
		carboxyl O:	3.3	
	hb donor O	aromatic N:	3.5	
		hb acceptor O	OH:	3.5
		NH:	3.6	
		CO	carbonyl	hb acceptor O
COO	carboxyl	hb acceptor O	OH:	3.3
		anion O	NH:	3.5
		CA2	2.9	
		MG2	2.6	
		ZN2	2.7	
		COOC	ester	hb acceptor O
		NH:	3.6	
		COC	ether	hb acceptor O
		NH:	3.6	
		CON	amide	hb donor N
carboxyl O:	3.5			
		aromatic N:	3.7	
		hb acceptor O	OH:	3.5
		NH:	3.6	
		NH	amino	hb donor N
		carboxyl O:	3.5	
		aromatic N:	3.7	
NAR	N aromatic	hb acceptor N	OH:	3.5
		hb acceptor O	NH:	3.7
SO3	sulfo	hb acceptor O	OH:	3.5
		anion O	NH:	3.6
		CA2	2.9	
		MG2	2.6	
		ZN2	2.7	
		SO2N	sulfonamide	hb donor N
carboxyl O:	3.5			
		aromatic N:	3.7	
		hb acceptor O	OH:	3.5
		NH:	3.6	


Additional vsFilt constraints for H-bonds


Apply angle constraints for h-bonds: No Yes 

Apply tight constraints for h-bonds: No Yes 

H-bond angles are constrained to be $\geq 130^\circ$


Additional vsFilt constraints for H-bonds


Apply angle constraints for h-bonds: No Yes 

Apply tight constraints for h-bonds: No Yes 

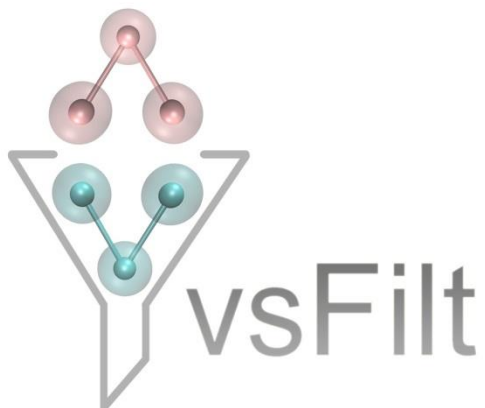
H-bond angles are constrained to be $\geq 130^\circ$

Additional vsFilt constraints for H-bonds

Apply angle constraints for h-bonds: No Yes 

Apply tight constraints for h-bonds: No Yes 

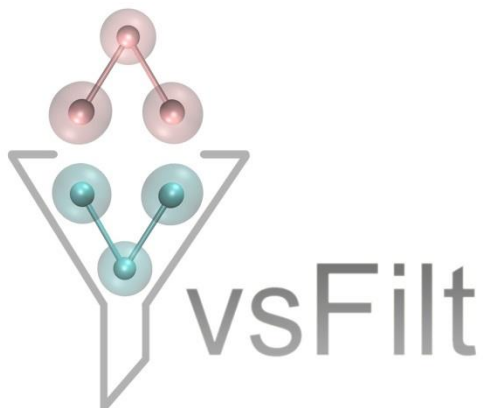
Distance \leq upper 90% quantile (without 0.5 Å increment), angle $\geq 150^\circ$



Hydrogen bonds

DON group can be used as an unspecified H-bond donor and ACC group as an unspecified acceptor:

DON	hb donor	hb donor	O: carboxyl O: aromatic N:	3.5–3.6 3.3–3.5 3.5–3.7
ACC	hb acceptor	hb acceptor	OH: NH:	3.3–3.5 3.5–3.7



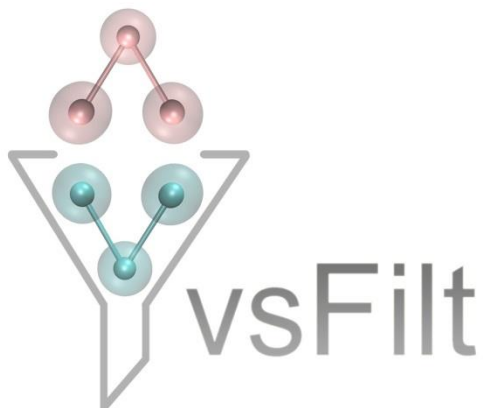
Ionic interactions

Maximum distance: the mean distance obtained from structural database statistics + 0.5 Å

COO	carboxyl	hb acceptor O	OH:	3.3
			NH:	3.5
		anion O	CA2	2.9
			MG2	2.6
			ZN2	2.7

SO3	sulfo	hb acceptor O	OH:	3.5
			NH:	3.6
		anion O	CA2	2.9
			MG2	2.6
			ZN2	2.7

Zheng et al. *J. Inorg. Biochem.* (2010)

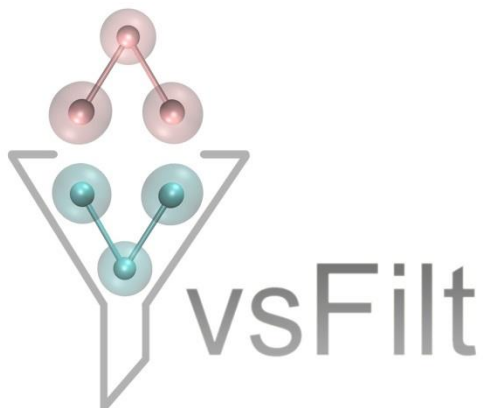


Halogen bonds

Maximum distances: upper 90% quantiles obtained from structural database statistics + 0.5 Å

HAL	halo	Cl	carbonyl O:	3.9
		Br, I	carbonyl O:	4.0

Bissantz et al. *J. Med. Chem.* (2010)

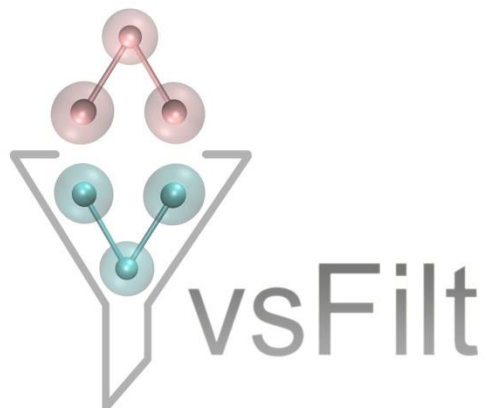


Hydrophobic interactions

Maximum distances: upper 90% quantiles obtained from structural database statistics + 0.5 Å

HPH	hydrophobic	aliphatic C	aliphatic C	4.9
			aromatic C	4.9
		aromatic C	aliphatic C	4.9
			aromatic C	4.3
		F	aliphatic C	4.4
		Cl	aliphatic C	4.8

Bissantz et al. *J. Med. Chem.* (2010)



Stacking interactions

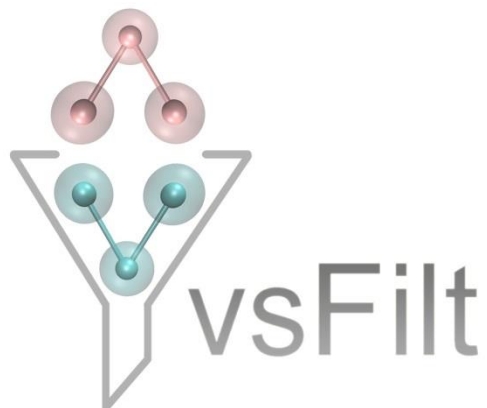
Maximum distance between centroids of stacked rings: $\leq 4.5 \text{ \AA}$

STK	aromatic	centroid*	centroid**	4.5
-----	----------	-----------	------------	-----

*Geometric center of ligand aromatic rings

**Geometric center of Phe/Tyr/His/Trp aromatic rings

Gonzalez et al. *J. Phys. Chem.* (2000)

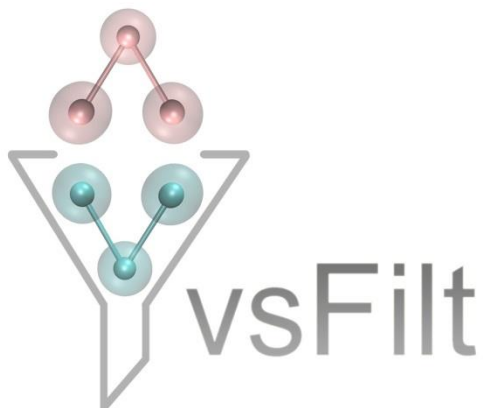


Cation – π Interactions

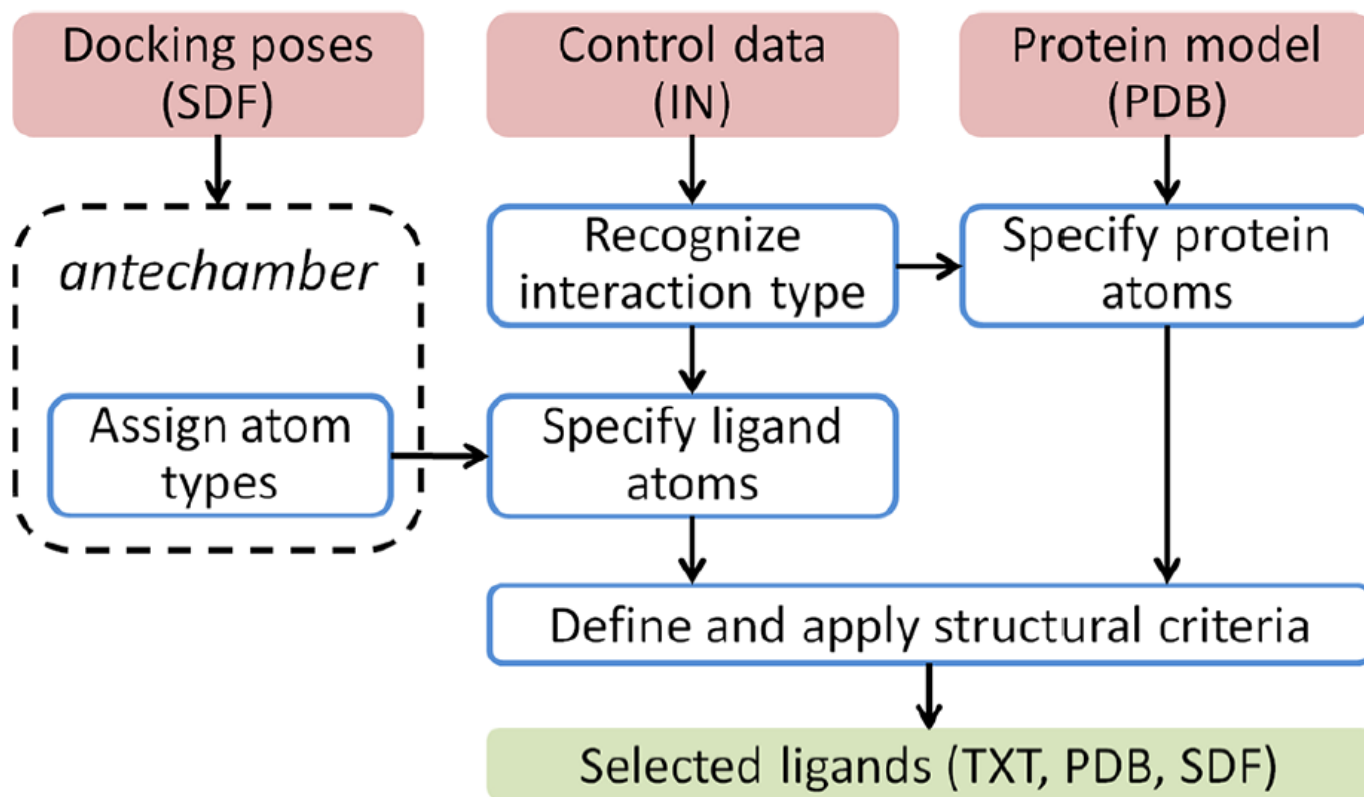
Maximum distances: upper 90% quantiles obtained from structural database statistics + 0.5 Å

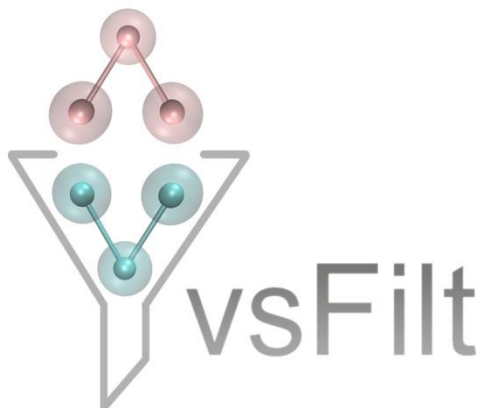
PIC	aromatic	aromatic C	guanidinium C	4.5
-----	----------	------------	---------------	-----

Bissantz et al. *J. Med. Chem.* (2010)



Workflow





<https://biokinet.belozersky.msu.ru/vsfilt>

vsFilt protein-ligand interaction profile

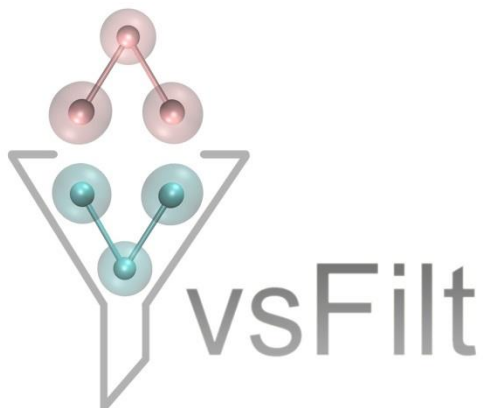
Provide the interaction profile: ?

Create and edit on-site Upload as a text file

Ligand	Protein	# of interactions	Action
OH	ALA	1	Add
			Or

Add a filtering rule

Paste/create/edit your interaction profile (at most 5000 characters)



<https://biokinet.belozersky.msu.ru/vsfilt>

vsFilt protein-ligand interaction profile

Provide the interaction profile: ?

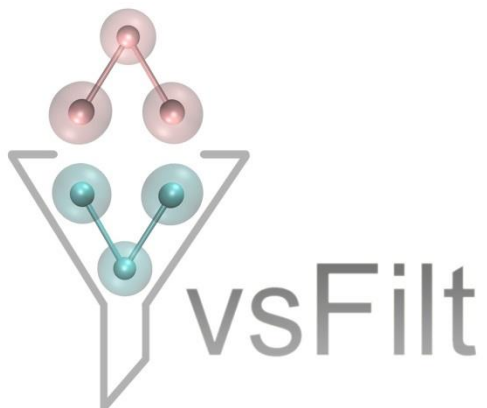
Create and edit on-site Upload as a text file

Ligand	Protein	# of interactions	Action
OH functional group	<input type="text"/> resnum main chain	<input type="text" value="1"/>	<input type="button" value="Add"/> <input type="button" value="Or"/>

Add at least one filtering criterion

Paste/create/edit your interaction profile (at most 5000 characters)

- ALA
- ALA
- ARG
- ASN
- ASP
- CYS
- GLN
- GLU
- GLY
- HIS
- ILE
- LEU
- LYS
- MET
- PHE
- PRO
- SER
- THR
- TRP
- TYR
- VAL



<https://biokinet.belozersky.msu.ru/vsfilt>

vsFilt protein-ligand interaction profile

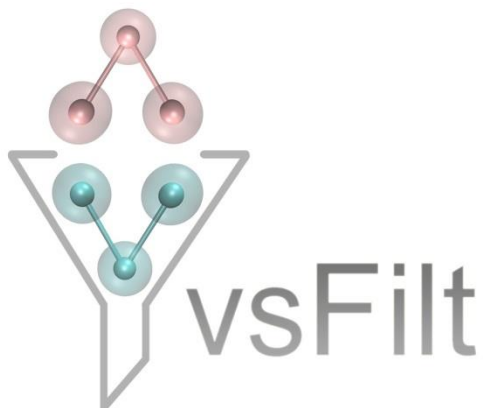
Provide the interaction profile: ?

Create and edit on-site Upload as a text file

Ligand	Protein		# of interactions	Action
OH ▾ functional group	ALA ▾ resname	<input type="text"/> resnum	<input type="text" value="1"/>	<input type="button" value="Add"/> <input type="button" value="Or"/>
		main ▾ main side		

Add at least one filtering rule

Paste/create/edit your interaction profile (at most 5000 characters)



<https://biokinet.belozersky.msu.ru/vsfilt>

vsFilt protein-ligand interaction profile

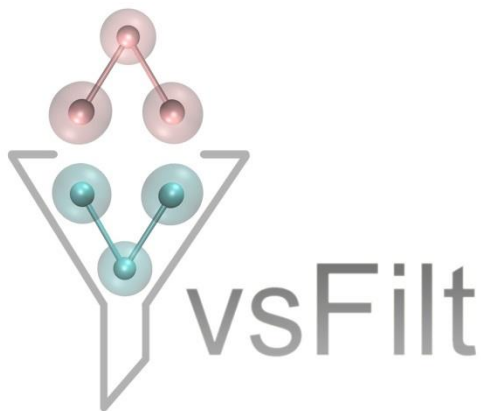
Provide the interaction profile: ?

Create and edit on-site Upload as a text file

Ligand	Protein			# of interactions	Action
<input type="text" value="OH"/>	<input type="text" value="ALA"/>	<input type="text"/>	<input type="text" value="main"/>	<input type="text" value="1"/>	<input type="button" value="Add"/>
funcional group	rename	resnum	chain		<input type="button" value="Or"/>

```
INTER CON  GLY 863  main 2
INTER HPH  ALA 898  side
INTER STK  TYR 907  side
```

Paste/create/edit your interaction profile (at most 5000 characters)



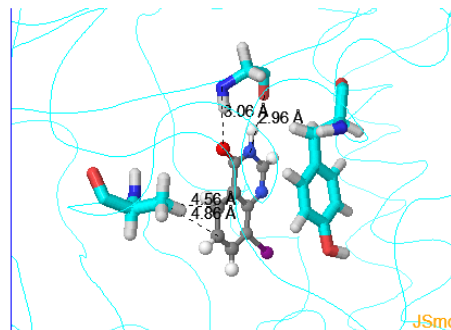
Online analysis

Analysis of the vsFilt results

Basic operations with the 3D-viewer: *Left-click-and-hold* and then *move your mouse* to rotate the structure, *Shift + Left-click-and-hold + Mouse Up/Down* to zoom in and out, *Ctrl + Right-click-and-hold + Mouse Up/Down/Left/Right* to move the structure in the viewer, *Right-click* for more options. Hold mouse pointer over selected amino acid for one second to view the label. Double click on a selected atom to activate the distance/angle measurement feature. For more refer to the [JSmol manual](#).

Quick hints: Each ligand and its interactions with the protein residues that comply with the filtering rules can be visualized individually by clicking on the respective checkbox, or loaded all at once by using the buttons **Toggle all ligands** (i.e., all ligands that passed the structural filtration will be shown in the 3D-viewer), **Toggle residues** (i.e., all residues involved in accommodation of ligands that passed the structural filtration will be shown in the 3D-viewer) and **Toggle interactions** (i.e., all interactions will be shown between ligands and residues previously enabled in the 3D-viewer). Click on a "cross" icon (✖) to hide the info for particular ligand (i.e., row) from the table. To restore all rows use the **Restore all rows** button. To highlight a ligand (row) in the table click on any cell (i.e., this feature can help to work with large tables).

<https://biokinet.belozersky.msu.ru/vsfilt>

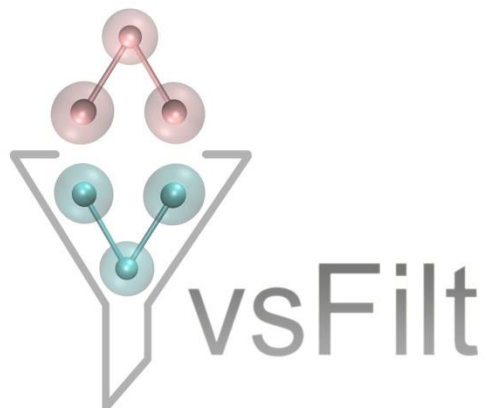


Viewport: 420x300, 840x600, 1260x900, 1680x1200. **Rendering of static image:** *antialias on (slower), antialias off (faster). Rendering of dynamic image:* *all features (slower), no antialiasing, no translucency, surfaces dotted, cartoons as trace, geosurfaces as dots, ellipsoids as dots, wireframe only (faster).*

Last action with the 3D-viewer: you have **enabled** ligand ZINC26894394 ranked #1

Operate PDB Heteatoms: Show/Hide water Show/Hide ligands Show/Hide ions
 Operate Ligands: Toggle all ligands Toggle residues Toggle interactions Restore all rows

Show/Hide	Rank	Ligand ID	Score	List of interactions	Hide row
<input checked="" type="checkbox"/>	1	ZINC26894394	-9.968	GLY:863 <i>main</i> O ... N <i>amide</i> CON GLY:863 <i>main</i> N ... O <i>amide</i> CON ALA:898 <i>side</i> CB ... C <i>aromatic</i> HPH ALA:898 <i>side</i> CB ... C <i>aromatic</i> HPH TYR:907 <i>side</i> centroid ... centroid <i>aromatic</i> STK	<input checked="" type="checkbox"/>
<input type="checkbox"/>	2	ZINC19522823	-9.364	GLY:863 <i>main</i> N ... O <i>amide</i> CON GLY:863 <i>main</i> N ... O <i>amide</i> CON ALA:898 <i>side</i> CB ... C <i>aliphatic</i> HPH ALA:898 <i>side</i> CB ... C <i>aliphatic</i> HPH TYR:907 <i>side</i> centroid ... centroid <i>aromatic</i> STK	<input checked="" type="checkbox"/>

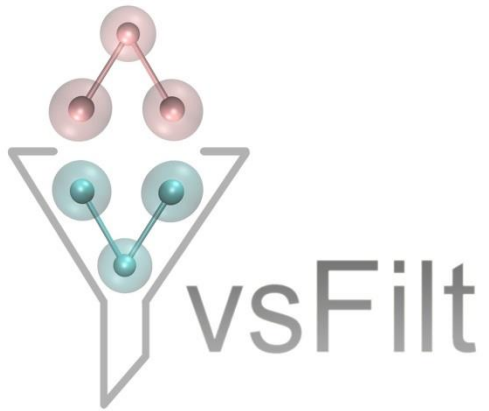


- free (no login required);
- can be combined with any type of docking software;
- implemented using HTML 5;
- processes an SDF library of up to 150 000 docked ligand poses

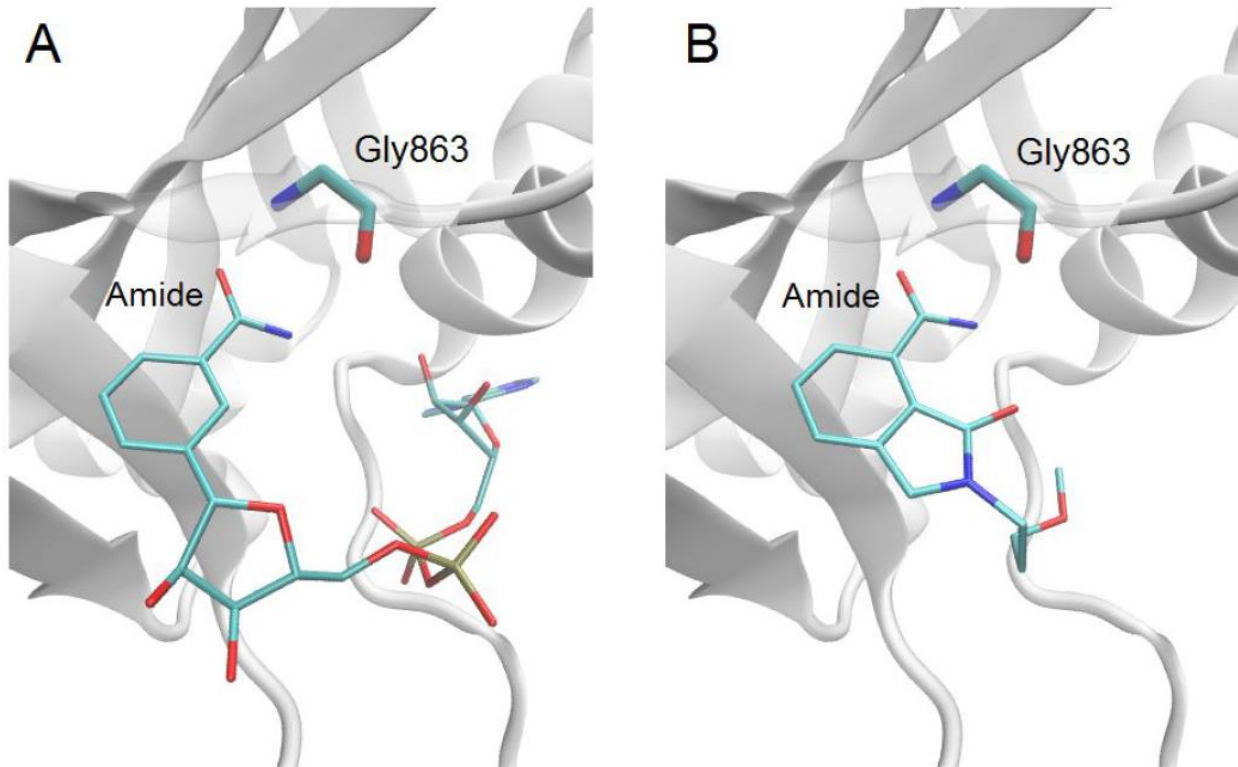
User data is protected by:

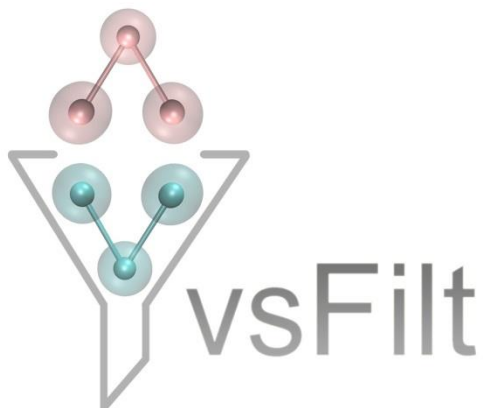
- unique access code (TaskID);
- use of HTTPS protocol;
- optional IP/password-based authentication

DEMO mode: ~9000 ligands, takes ~2 min



Illustrative example: PARP ligands





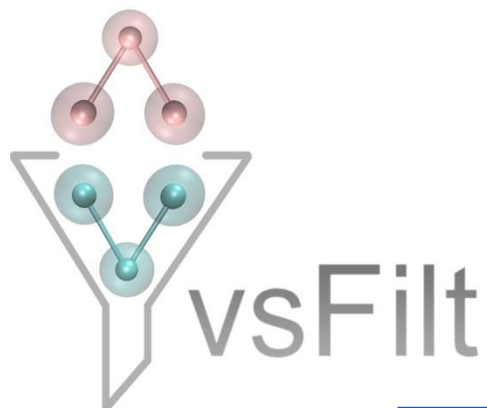
Illustrative example: PARP ligands

- 236 026 compounds containing a benzamide substructure (classical PARP-1 inhibitor scaffold) from the ZINC12 library;
- Docking with Lead Finder;
- Structural filtration with vsFilt

	ligand group	protein residue		number of interactions	
INTER	CON	GLY	863	main	2
INTER	HPH	ALA	898	side	
INTER	STK	TYR	907	side	
OR					
...					

Annotations: Blue arrows point from 'ligand group' to 'CON', 'protein residue' to 'GLY', 'number of interactions' to '2', 'residue number' to '907', and 'residue chain' to 'side'.

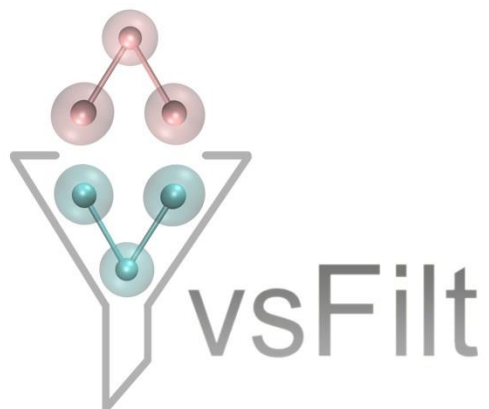
Figure 1. Example of control data used by vsFilt for structural filtration.



Illustrative example: PARP ligands

Table 2. Number of Ligands Selected by vsFilt among 236 026 Benzamide Derivatives by Applying Structural Criteria with “Angle Constraints” and “Tight Constraints” Options

interaction	angle constraints	tight constraints	number of ligands
CON...Gly863	—	—	8814
CON...Gly863	+	—	8199
CON...Gly863	+	+	1857
CON...Gly863	+	+	604 (0.26 %)
HPH...Ala898			
STK...Tyr907			

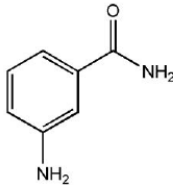
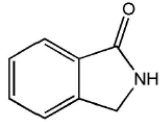
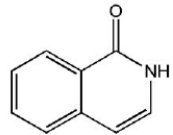
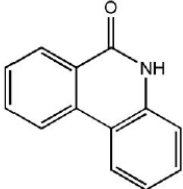
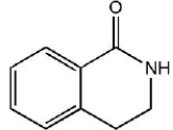
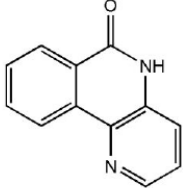
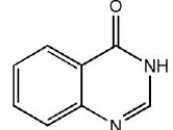
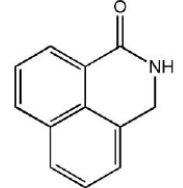


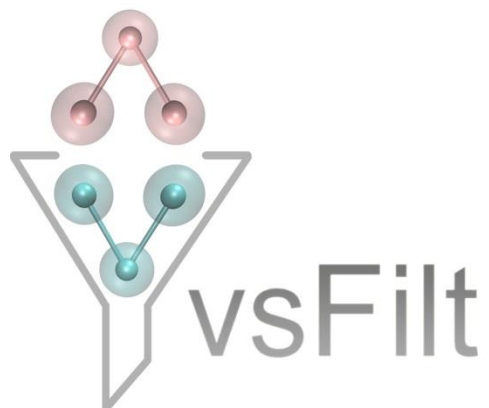
Illustrative example: PARP ligands

	ligand group	protein residue	number of interactions	
INTER	CON	GLY	863	main 2
INTER	HPH	ALA	898	side
INTER	STK	TYR	907	side
OR				
...				

residue number
residue chain

Figure 1. Example of control data used by vsFilt for structural filtration.

Scaffold	Number of ligands	Scaffold	Number of ligands
	43		8
	28		6
	8		1
	315		6



Illustrative example: PARP ligands

	ligand group	protein residue	number of interactions	
INTER	CON	GLY 863	main	2
INTER	HPH	ALA 898	side	
INTER	STK	TYR 907	side	
OR			residue number	residue chain
...				

Figure 1. Example of control data used by vsFilt for structural filtration.

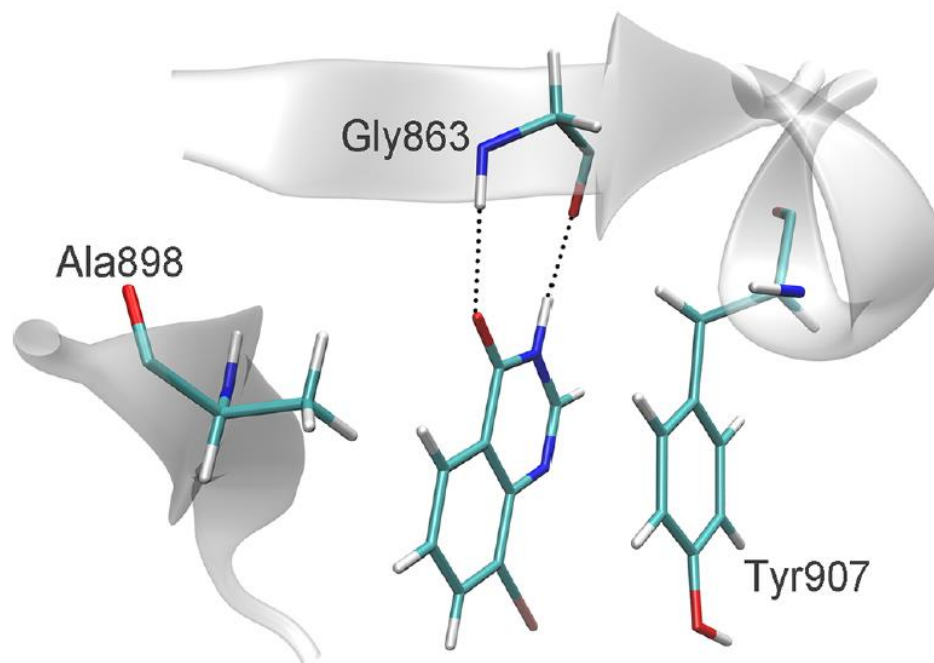
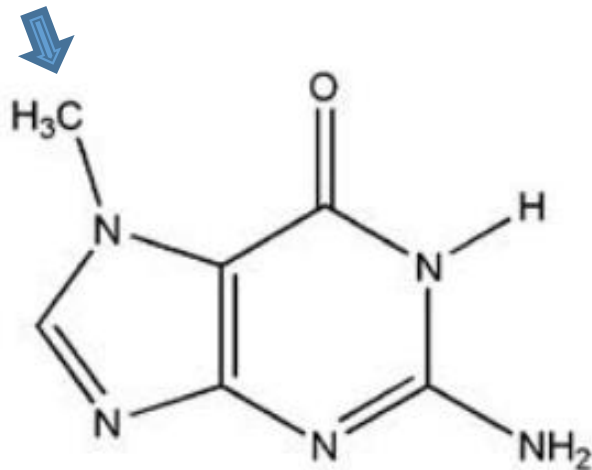


Figure 3. Interactions of ZINC26894394 ligand with the active site residues in the PARP-1 model: H-bonds with Gly863, hydrophobic contact with Ala863, and π -stacking with Tyr907. The ligand pose was selected by vsFilt and visualized using VMD 1.9.2.²⁶

Пример: 7-Метилгуанин (7-МГ):



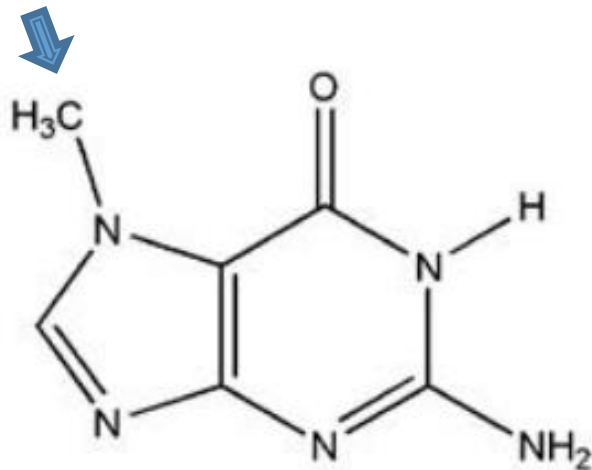
- Метаболит РНК и ДНК
- В небольшой концентрации обнаруживается в моче
- Не используется для синтеза нуклеотидов и не встраивается в ДНК

Weissmann et al. *J. Biol. Chem.* (1957) 224, 407-422

Kaina et al. *Mutat. Res.* (1983) 108, 279-292

Svoboda et al. *Anal. Biochem.* (2004) 334, 239-250

Пример: 7-Метилгуанин (7-МГ):



- Метаболит РНК и ДНК
- В небольшой концентрации обнаруживается в моче
- Не используется для синтеза нуклеотидов и не встраивается в ДНК
- Ингибитор ПАРП-1

Weissmann et al. *J. Biol. Chem.* (1957) 224, 407-422
Kaina et al. *Mutat. Res.* (1983) 108, 279-292
Svoboda et al. *Anal. Biochem.* (2004) 334, 239-250
Nilov et al. *Int. J. Mol. Sci.* (2020) 21, 2159

Биохимические исследования

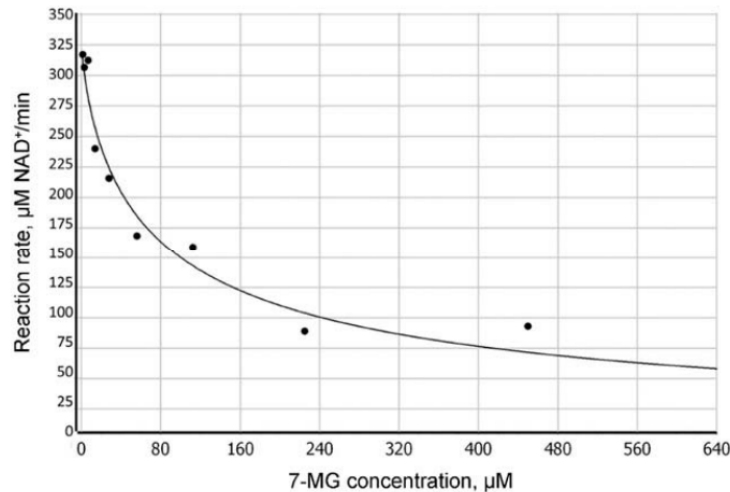


Figure 2. Dependence of the PARP-1-catalyzed reaction rate on the concentration of 7-MG inhibitor determined by fluorescence anisotropy (100 μM NAD⁺ concentration).

- Конкурентный ингибитор
- $K_i \approx 10 \text{ мкМ}$

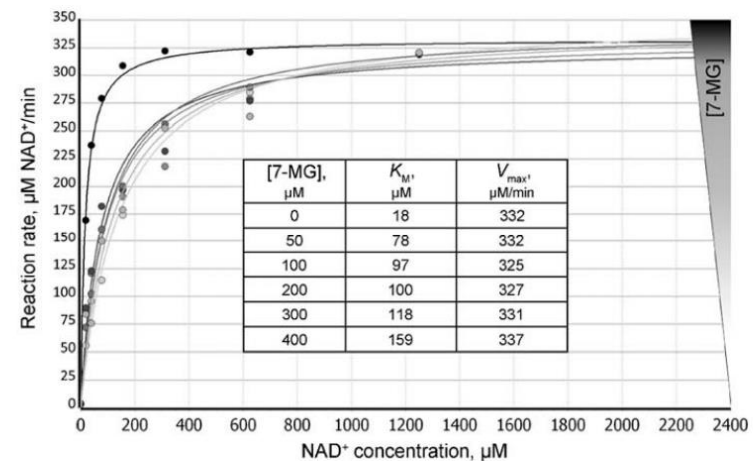
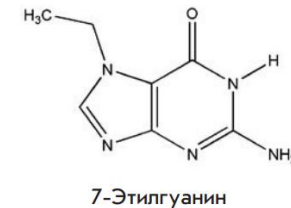
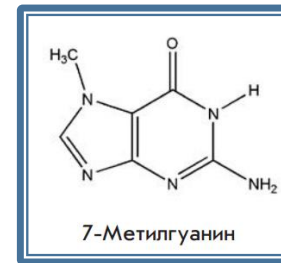


Figure 3. Dependence of the PARP-1-catalyzed reaction rate on the NAD⁺ concentration at different concentrations of 7-MG added to the reaction mixture. Insert: calculated K_M^{APP} values increase with increasing 7-MG concentrations, thus demonstrating the competitive inhibition mechanism.

Молекулярное моделирование

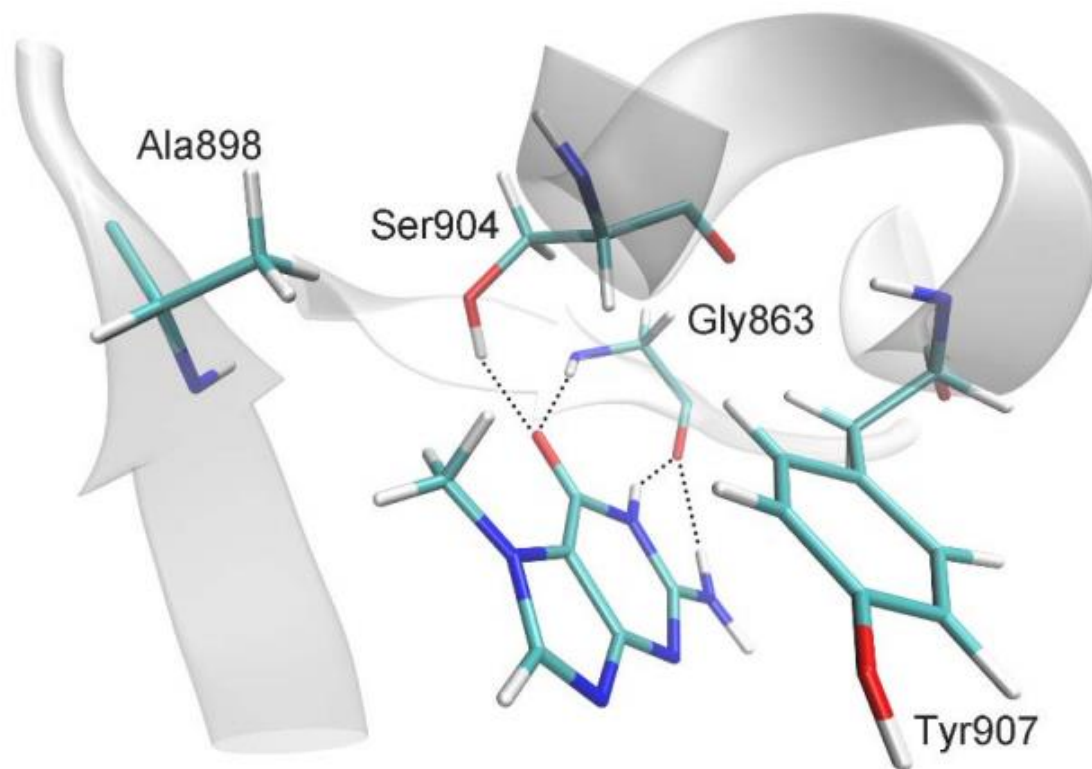


Figure 1. Interactions of 7-MG molecule in the PARP-1 active site revealed by MD simulation: hydrogen bonds with Gly863 and Ser904, π -stacking of purine rings with Tyr907, and hydrophobic contact between the 7-MG methyl group and Ala898.

Нилов и соавт. *Acta Naturae* (2016) 8, 120-128

Nilov et al. *Int. J. Mol. Sci.* (2020) 21, 2159

Manasaryan et al. *Cancers* (2021) 12, 1201



Дмитрий Нилов

E-mail: nilovdm@gmail.com

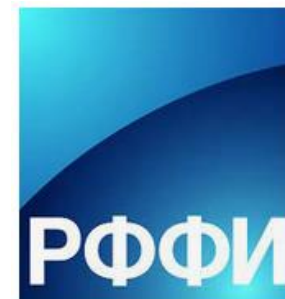
<https://istina.msu.ru/profile/nilovdm/>

Ирина Гущина (ФББ МГУ)
Дмитрий Суплатов (НИИ ФХБ МГУ)
Элина Смолкина (ФФМ МГУ)
Александра Поленова (ФФМ МГУ)
Сергей Евтеев (ФФМ МГУ)
Кристина Мотузенко (ФФМ МГУ)
Ксения Яшина (ФББ МГУ)
Витас Швядас (НИВЦ МГУ)



Российский научный фонд

№ 21-71-30003





Дмитрий Нилов

E-mail: nilovdm@gmail.com

<https://istina.msu.ru/profile/nilovdm/>

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING
pubs.acs.org/jcim Application Note

vsFilt: A Tool to Improve Virtual Screening by Structural Filtration of Docking Poses

Irina V. Gushchina,[§] Aleksandra M. Polenova, Dmitry A. Suplatov, Vytas K. Švedas, and Dmitry K. Nilov^{*§}

Cite This: <https://dx.doi.org/10.1021/acs.jcim.0c00303> Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The ability of ligands to form crucial interactions with a protein target, characteristic for the substrate and/or inhibitors, could be considered a structural criterion for identifying potent binders among docked compounds. Structural filtration of predicted poses improves the performance of virtual screening and helps in recovering specifically bound ligands. Here, we present vsFilt—a highly automated and easy-to-use Web server for postdocking structural filtration. The new tool can detect various types of interactions that are known to be involved in the molecular recognition, including hydrogen and halogen bonds, ionic interactions, hydrophobic contacts, π -stacking, and cation- π interactions. A case study for poly(ADP-ribose) polymerase 1 ligands illustrates the utility of the software. The Web server is freely available at <https://biokinet.belozersky.msu.ru/vsfilt>.

<https://biokinet.belozersky.msu.ru/vsfilt>