



Section 3. Life sciences

DOI:10.29013/EJBLS-26-2.3-20-37



SMALL MOLECULE USP 1 INHIBITORS FROM VIRTUAL SCREENING AS NEW CANCER THERAPIES

Bethany Ma¹

¹ Saint Francis High School, Mountain View, CA, USA

Cite: Bethany Ma. (2026). Small molecule USP1 inhibitors from virtual screening as new cancer therapies. *The European Journal of Biomedical and Life Sciences* 2026, No 2 – 3 <https://doi.org/10.29013/EJBLS-26-2.3-20-37>

Abstract

Ubiquitin-Specific Protease 1 (USP1) is a deubiquitinating enzyme that regulates DNA damage repair through the deubiquitination of key proteins, including PCNA and FANCD2-FANCI. USP1 supports DNA repair pathways that have been proven to be particularly important for the survival of certain cancer cells; So, its inhibition represents a promising strategy for developing new targeted cancer therapies. This study used a computational drug-discovery approach to identify potential small molecule USP1 inhibitors. Geometric and machine learning binding site prediction methods were used to identify and compare potential druggable pockets within USP1, resulting in the identification of several candidate binding sites. Pharmacophore modeling was then performed to characterize key molecular interactions within selected binding sites, including aromatic, hydrophobic, and hydrogen-bonding features. These pharmacophore models were used to screen compounds from the MCULE-ULTIMATE, Enamine, and MolPort chemical libraries, identifying 19 compounds with RMSD values below 0.2. The compounds were subsequently evaluated using SwissDock to assess predicted binding spontaneity, followed by SwissADME analysis to evaluate drug-likeness and Lipinski's Rule of Five and ProTox 3.0 analysis to assess predicted toxicity. Through this sequential screening process, MolPort-010-912-491 and Z1618187599 emerged as the most promising lead and backup compounds, respectively, based on their favorable docking results, low RMSD values, drug-like physicochemical properties, and predicted toxicity profiles. Overall, this study identified two potential USP1 inhibitors that warrant further investigation through molecular dynamics simulations, experimental binding assays, and in vitro cancer-cell testing.

Keywords: Cancer, drug discovery, small molecules, virtual screening, molecular docking

Introduction

Cancer is a major global health issue and has long been a significant cause of mor-

bidity and mortality worldwide. The term "cancer" serves as an umbrella term encompassing a diverse group of diseases charac-

terized by uncontrolled cell growth. Among the various types of cancer, lung cancer is the most frequently diagnosed, followed by cancers of the breast, colorectum, prostate, and stomach (Bray et al., 2024). Although advances in technology and improvements in healthcare systems have significantly reduced cancer mortality rates, it is still predicted that 2,114,850 new cancer cases and 626,140 cancer-related deaths will occur in the United States in 2026, corresponding to approximately 1,720 deaths per day (Siegel et al., 2026). In response to growing cancer cases worldwide, scientists and healthcare professionals have been working to develop and refine cancer therapies designed to reduce cancer progression and prevent tumor growth.

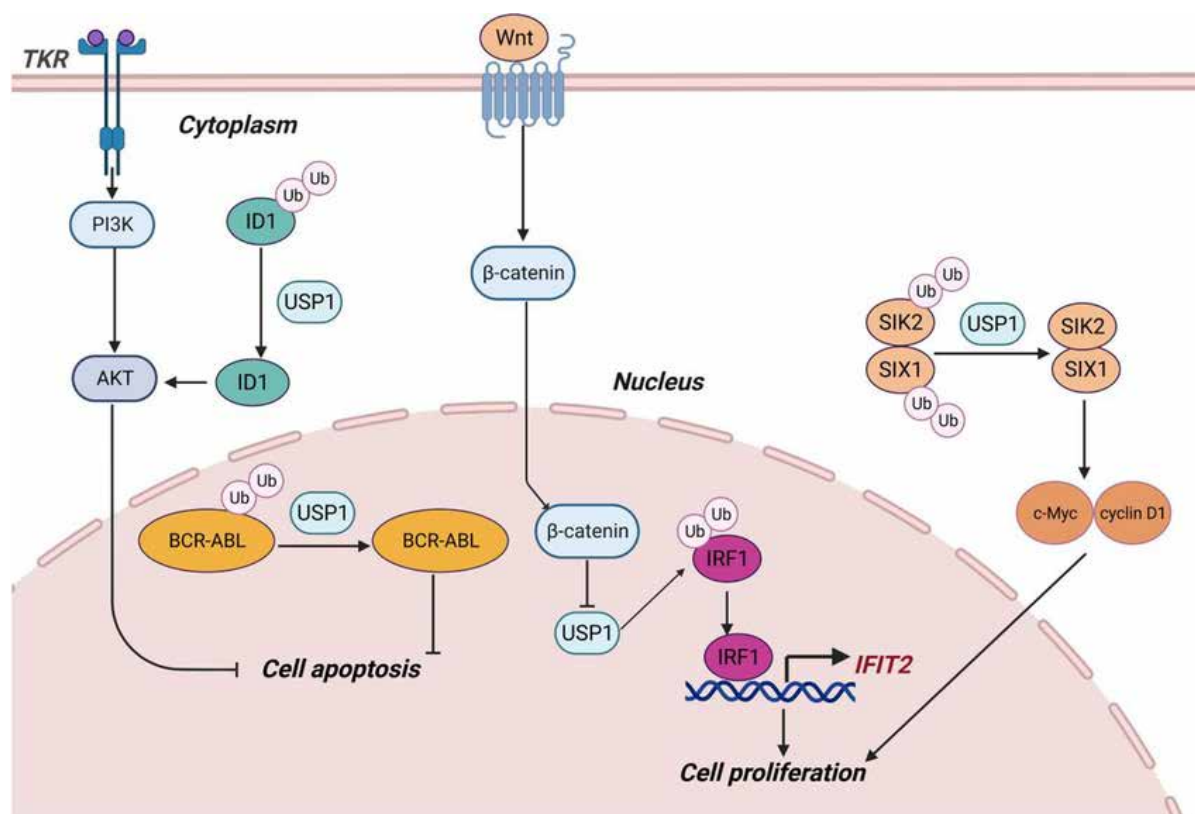
There is currently no cure for cancer, however, different cancer therapies have evolved to stop disease progression or shrink/remove tumors. Surgery, the first rational cancer treatment, is used for the removal of tumors surrounding cancerous tissue. In the beginning of the 20th century, radiation therapy emerged as a treatment for cancer through x-rays, and as time progressed, brachytherapy, proton beam therapy, and fast neutron therapy (Sonkin et al., 2024). Though radiation is cytotoxic towards cancerous cells, it must be implemented sparingly to minimize normal tissue radiation exposure. Chemotherapy is another form of cancer therapy which inhibits tumor growth and metastasis through disruption of DNA, RNA, and protein synthesis in rapidly proliferating cells (Amjad et al., 2023). Lastly, immunotherapy is designed to either enhance the host's immune system or to target a specific antigen on the tumor. By strengthening the host immune response, the immune system can more effectively recognize and eliminate malignant cells (Sahu & Suryawanshi, 2021).

Ubiquitin-Specific Protease 1 (USP1) is a deubiquitinating enzyme that removes ubiquitin from proteins which are involved in DNA repair or replication. Two crucial DNA repair pathways USP1 is involved in include the translesion synthesis (TLS) and Fanconi Anemia (FA) pathway (where interstrand crosslink (ICL) repair occurs) (Mazloumi et al., 2025). There, USP1 removes ubiquitin from PCNA and FANCD2-

FANCD1, respectively (Torrado et al., 2025). Without USP1, these proteins accumulate mass amounts of ubiquitin, which leads to defective DNA repair. Therefore, proper deubiquitination through USP1 is necessary for DNA repair completion and replication fork stability. Structurally, USP1 is a 7850amino-acid cysteine protease with a catalytic domain which contains Cys90, His593, and Asp752, which are responsible for its deubiquitinating activity. USP1 functions in complex with USP1-associated factor 1 (UAF1), a cofactor that enhances its catalytic activity and is required for efficient substrate deubiquitination (Lemonidis et al., 2022). Together, the two ensure regulation of ubiquitination levels during DNA replication and repair, allowing major repair pathways to be activated and reset whenever repair is completed (Kitamura, 2023).

USP1 is strongly linked to cancer due to many tumors' dependency on its DNA repair functions. Aside from BRCA1-deficient tumors, USP1 expression and activity are also increased in colorectal cancer, non-small cell lung cancer, gastric cancer, and platinum-resistant ovarian cancer. In this increased activity, USP1 promotes tumor growth and survival through stabilization of cancer-related proteins, as well as supports tumor progression by increasing cell proliferation and metastasis (Hu et al., 2025). Consequently, inhibition of USP1 has shown to result in suppression of cancer cell proliferation and epithelial-mesenchymal transition (EMT), a process that enables cancer cells to become more invasive and spread farther (Huang et al., 2022). Moreover, USP1 inhibition can disrupt multiple oncogenic signaling pathways that contribute towards tumor progression. So, by inhibiting USP1 in a cancer patient, not only can cancer cell proliferation be reduced, but cancer cell migration and survival. In BRCA1-deficient cancers, USP1 plays an especially important role due to the tumors lacking an effective homologous recombination repair pathway and therefore rely on alternative DNA repair mechanisms regulated by USP1. So, USP1 inhibition is a promising target for the development of new cancer therapies.

Figure 1. Models depicting the role of USP1 as a regulator for cancer cell fate through interactions with multiple substrates. Through deubiquitination of key substrates, USP1 regulates pathways including PI3K/AKT and Wnt/ β -catenin signaling, as well as transcriptional regulators such as ID1, BCR-ABL, SIX1, and IRF1. (TKR, tyrosine kinase receptor; PI3K, phosphoinositide 3-kinase; IRF1, interferon regulatory factor 1; IFIT2, interferon-induced protein with tetratricopeptide repeat 2; Ub, ubiquitin; SIK2, salt-inducible kinase 2; SIX1, sine oculis homeobox homolog 1) (12)



To therapeutically target USP1, researchers must determine the most effective type of drug molecule to inhibit its activity. In terms of size, USP1 is a relatively large protein enzyme, and it can be targeted by either large-molecule therapeutics or small-molecule ones. Though larger molecules have higher precision, as molecular weight of the molecule increases, its solubility in water or other fluids is decreased, making it more difficult for the patient's body to absorb and distribute the drug (Gu et al., 2023). Moreover, small molecules can selectively bind to specific cancer related targets, as seen in metarrestin, a small molecule which targets the perinucleolar compartment and inhibits cancer progression (Kashyap et al., 2024). With these benefits and its cheap production cost, small-molecule drugs have become a major component of modern targeted therapies, es-

pecially in cancer treatments (Pedre  ez et al., 2021).

Methodology

Geometric method

To obtain the model of USP1, the Protein Plus homepage must be opened up first by clicking the link, (<https://proteins.plus>). Once prompted for the PDB code, enter 7ZH4, which represents the structure of the USP1 inhibitor bound to FANCD2. After entering the code, launch DoGSiteScorer and keep all settings the same. This includes leaving analysis detail(s) to pockets, leaving chain set to all, and leaving the ligand section blank. Click the "calculate" option and soon after, and the 3d model should appear on the left of the screen with a table on the right. To show only all the high-druggability binding sites, make sure

to unhide the sites with a drug score over 0.5 (Ehrt et al., 2025).

Machine learning method

To obtain a model of USP1, first head to the Prank Web homepage by clicking the link, (<https://prankweb.cz>) Once clicked, the user should arrive at the homepage where the words, «Prank Web: Ligand Binding Site Prediction» are at the top. Select the Experimental Structure option and provide the PDB code, 7ZH4. Be sure to select “Use original structure” as well as “Use conservation” before submitting. After submitting, the 3d model would appear on the left of the screen and a table of binding sites on the right. To only see the most probable binding sites, hide the sites with scores that are much lower compared to other sites, such as the sites with scores 1.15, 1.01, and 0.99 (Jendele et al., 2019).

Pharmacophore interactions model

To obtain a three-dimensional model that only consists of the interactions between a ligand and certain binding sites, the user must visit the pharmit homepage, (<https://pharmit.csb.pitt.edu>). Upon entering the homepage, select the empty box on the right of “start from PDB,” enter “7ZH4,” and hit submit. After hitting submit, select the first, fourth, seventh, ninth, and fifteenth option under the pharmacophore section. Also, be sure that the other options not listed above are deselected. When the “on” signal is next to each of the selected options, scroll down to the visualization section. There, select the box on the right of “Ligand” and click “none.” After, select the box on the right of “Receptor” and click “none.” Drag the slider of the “Receptor Surface Opacity” all the way to the left and you should have the desired results. For the additional compounds from other unique yet significant sets of interactions, feel free to select the third, fifth, tenth, eleventh, and thirteenth options under the pharmacophore sections. Another good set of pairings can be achieved by selecting the second, sixth, ninth, eleventh, and thirteenth options (Sunseri & Koes, 2016).

Overlaying the compounds with pharmacophore

After isolating the pharmacophore on pharmit, click on the top left button, “Search

MolPort.” After a couple seconds, a long list of results should appear on the right of the screen with different compounds from the library. To view the overlay, click on the name of the compound. For other chemical libraries, such as Enamine and MCULE-ULTIMATE, select the downward arrow on the right of “Search Molport” and click on the respective options.

Finding spontaneous interactions

To obtain SMILES of a compound, go to the website of your desired compound and input the compound name into the search bar. For the results in this paper, visit (<https://enaminestore.com/> for the Enamine library, <https://www.molport.com/shop/index>) for the MolPort library, or (<https://ultimateapp.mculc.com/search>) for the MCULE-ULTIMATE Library. Copy and paste the SMILES into the swissdock website (<https://www.swissdock.ch>). Once the SMILES is provided, enter “Prepare Ligand.” Make sure that a green checkmark appears next to the box after a couple seconds. Scroll down and enter “7ZH4” for the PDB code, then select “Select All” for the “Choose chain(s) to keep” section and tick the box next to “None” for the “Choose heteroatom(s) to keep” section. Select “Prepare target” and wait for a green checkmark to appear next to the box. Scroll further down to the “Define search space” section and put 40, 40, and 30 respectively for the three parameters in “Box size.” Be sure to leave the box center parameters as is. Next, scroll down to “Select Parameters.” Leave the number of RIC as is and hit the “Check parameters” box. Once a time is returned, start docking. After a bit of time, the results should be returned and a table will show, with SwissParam and AC values. Choose the most negative SwissParam value by selecting it. After selection, the three-dimensional model should update (Bugnon et al., 2024; Röhrig et al., 2023).

Evaluating absorbability, distributability, and excretability

To evaluate whether a compound fulfills Lipinski’s rule, use (<https://www.swissadme.ch>). First, enter the SMILES of the compound in the empty box on the right and click the red “Run!” button below it. Wait for the website

to load, and scroll to the bottom of the page. There, focus on the following sections: molecular weight, number of H-bond acceptors, number of H-bond donors, and Consensus Log P. If the compound's molecular weight is <500 g/mol, it has no more than 10 hydrogen bond acceptors, no more than 5 hydrogen bond donors, and a Log P value no more than 5, then the compound fulfills Lipinski's Rule. To easily check this, look at how many violations the compound has in the Lipinski section. If the compound has more than 0 violations, it does not fulfill the rule (Daina et al., 2017).

Evaluating toxicity

To evaluate the level of toxicity in a compound, first head to the Protox3 homepage, https://tox.charite.de/protox3/index.php?site=compound_search_similarity. Select the second box with "Tox prediction," and enter the SMILES of the compound in the "Canonical Smiles" section. Scroll down and select the "All" button on the bottom right. Once all the boxes are ticked, select "Start Tox-Prediction" (Banerjee et al., 2024).

Results and Discussion

Figure 2 and Table 1 are both generated through a geometric method; while the figure characterizes the binding sites visually, the table does so quantitatively. Geometric

methods generally detect surface concavities and clefts, which helps it stay unbiased when locating for possible druggable regions. In Figure 2, six different colored regions are marked, which represent the different possible binding sites. In Table 1, quantitative data is provided, including: surface area, volume, and drug score. Pockets P_0 and P_1 have higher drug scores, so will therefore be more favorable for ligand binding. Lower scoring sites such as P_3 and P_5 indicate less favorable binding sites, but can still serve as potential secondary sites.

Figure 2. Predicted binding sites of USP1 identified using geometric methods. Colored sections represent pockets with the highest drug scores, indicating that those areas are more favorable for the binding of drug-like molecules

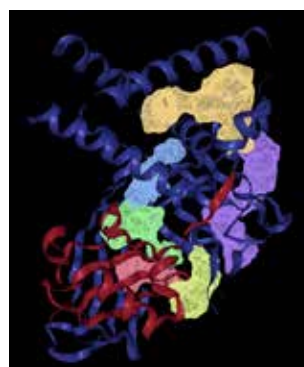


Table 1. Predicted binding pockets for USP1, including pocket surface area, volume, color designation for visualization, and drug score

Name	Color	Surface Area (Å ²)	Volume (Å ³)	Drug Score
P_0	Orange	1309.51	919.94	0.84
P_1	Purple	878.09	645.5	0.84
P_2	Green	381.12	359.94	0.71
P_4	Blue	229.39	242.56	0.64
P_3	Red	354.1	252.35	0.59
P_5	Yellow	383.81	229.95	0.58

Figures 3A and 3B, along with Table 2, were all generated through machine learning methods. Due to its rapid analysis of protein structures through a computational approach, machine learning methods for analyzing binding sites account for all circumstances, even mutations. In Figures 3A and 3B, the algorithm predicted and visualized the binding sites with varying levels of confidence. In Table 2, it pro-

vides the prediction scores and residue counts for each site. The score reflects the likelihood that a site can bind to drug-like molecules; the higher the score, the higher the likelihood. The residue numbers represent the amino acids that line each binding site and participate in the various interactions between the ligands. A greater amount of residues can also provide additional opportunities for molecular interactions.

Figure 3. (A) Predicted binding sites of USP1 identified using machine learning methods. Colored sections represent pockets with the highest ranks, indicating that those areas are more favorable for the binding of drug-like molecules. This figure contains a semi-transparent molecular surface to allow for users to see the inner molecule (B) Showing only the outer surfaces of the molecule, three of the four high ranked binding sites are shown

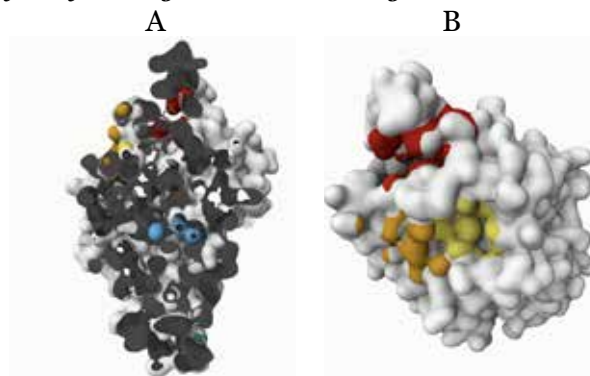


Table 2. Predicted ligand-binding pockets identified in USP1, including rank visualization color, prediction score, and number of residues in each binding site

Name	Rank	Color	Score	# Of Residues
B_0	1	Red	9.01	16
B_1	2	Yellow	5.75	11
B_2	3	Orange	3.28	9
B_3	4	Blue	3.05	10

Figure 4. (A) Three-dimensional pharmacophore model where purple represents aromatic features, yellow represents a hydrogen acceptor, and green represents hydrophobic features. This figure contains a close set of interactions with two aromatic interactions, two hydrophobic features, and one hydrogen acceptor. (B) This figure contains two aromatic interactions, two hydrogen acceptors, and one hydrophobic feature. (C) This figure contains slightly more dispersed features, with two aromatic interactions, two hydrogen acceptors, and one hydrophobic feature

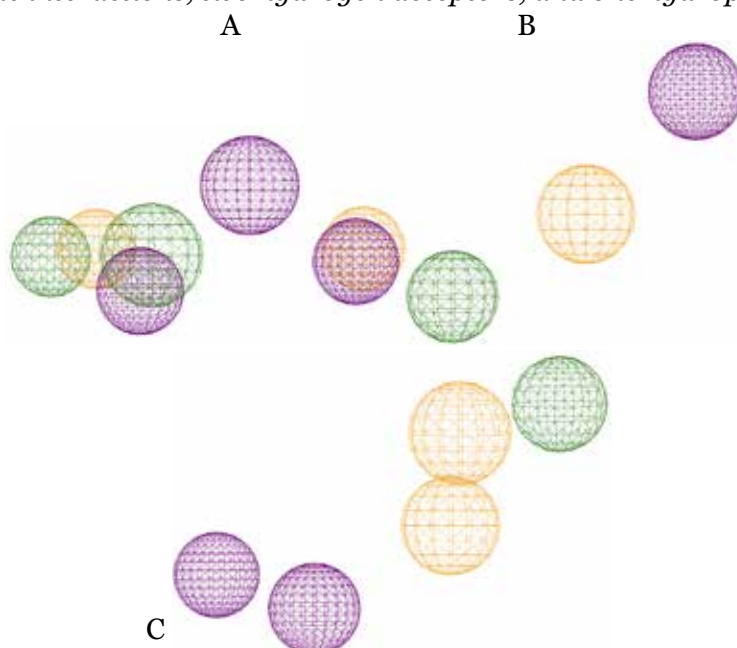
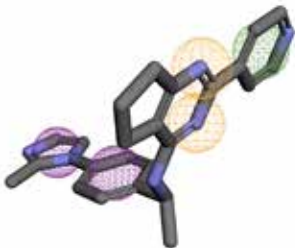
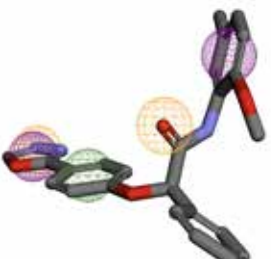
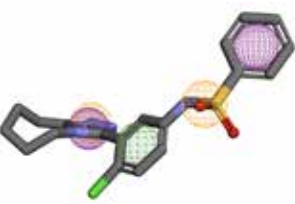
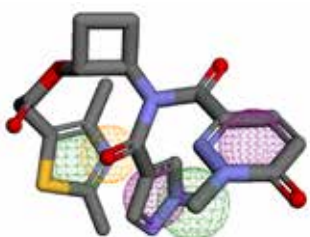
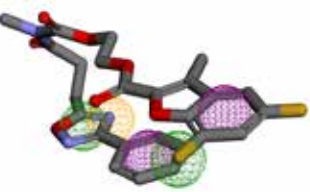
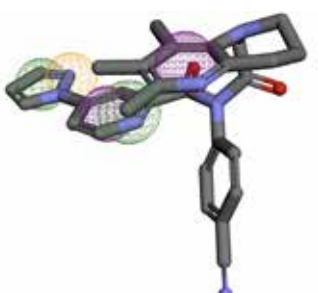
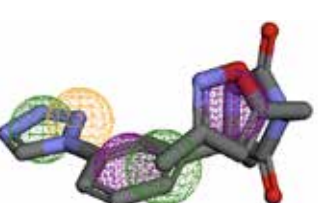
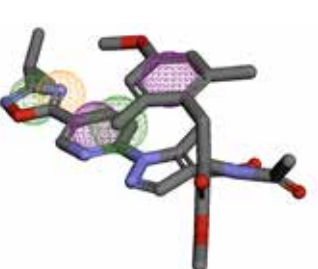
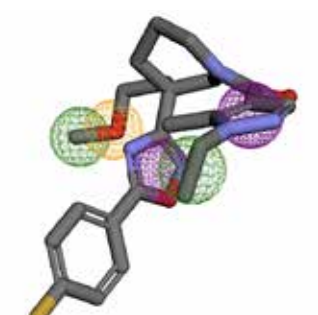
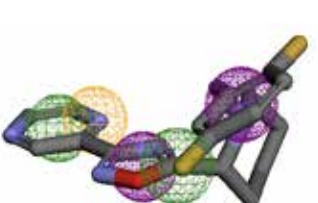
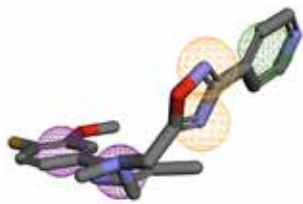
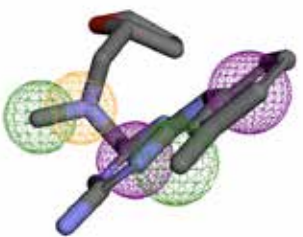
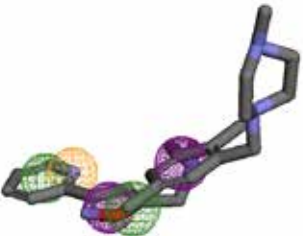
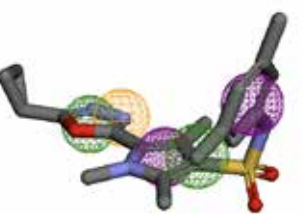
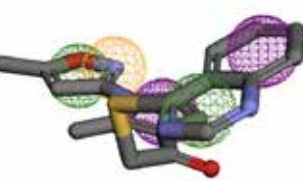


Table 3. *Overlays between the pharmacophore and nineteen different compounds drawn from three chemical libraries: MCULE-ULTIMATE, Enamine, and MolPort. Each compound is selected based on low root mean square deviation (RMSD) and all compounds selected fall under 0.2 RMSD, which means that the compound has a closer alignment with the pharmacophore model*

Overlay Between Pharmacophore and Compound	Name of Compound	Rmsd	Chemical Library	Figure #
	Z1359373132	0.100	Enamine	Figure 4C
	Z54926392	0.103	Enamine	Figure 4B
	Z299686082	0.118	Enamine	Figure 4B
	SKQBPFUHHTWY- DZ-UHFFFAOYSA-N	0.123	MCULE-UL- TIMATE	Figure 4A
	FUNUR- ESQIKQOAP-UHFFFAOYSA-N	0.128	MCULE-UL- TIMATE	Figure 4A
	MolPort-005-108-116	0.132	MolPort	Figure 4A

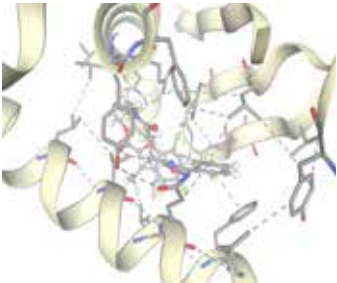
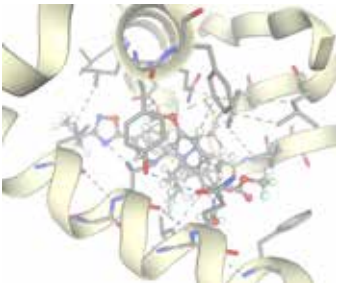
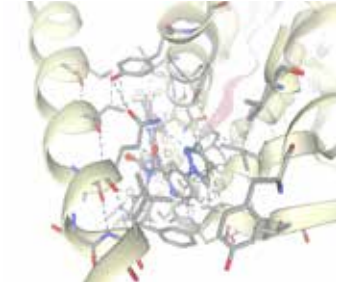
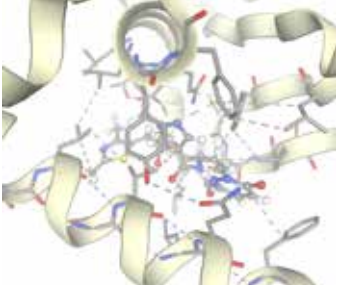
Overlay Between Pharmacophore and Compound	Name of Compound	Rmsd	Chemical Library	Figure #
	VIBCQN- PAWUBELN-UHFFF- AOYSA-N	0.133	MCULE-UL- TIMATE	Figure 4A
	HUNPSCXefd- CREC-UHFFFAOYSA-N	0.133	MCULE-UL- TIMATE	Figure 4A
	Z134661322	0.134	Enamine	Figure 4A
	BUDVOCFTH- VHXBT-UHFFFAOY- SA-N	0.138	MCULE-UL- TIMATE	Figure 4A
	MolPort-046–059–419	0.147	MolPort	Figure 4A
	MolPort-010–950–419	0.169	MolPort	Figure 4A

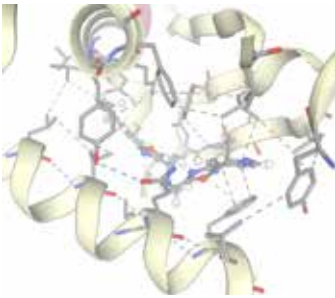
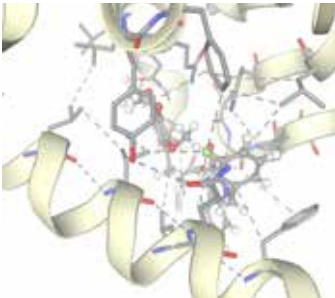
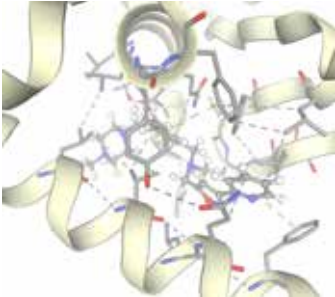
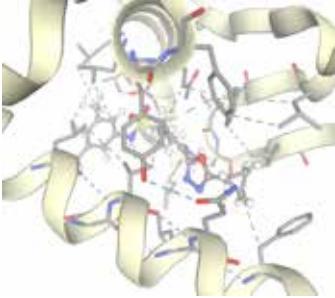
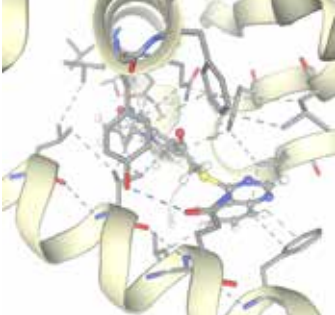
Overlay Between Pharmacophore and Compound	Name of Compound	Rmsd	Chemical Library	Figure #
	MolPort-005-163-256	0.170	MolPort	Figure 4C
	Z1618187599	0.185	Enamine	Figure 4A
	MolPort-044-685-945	0.186	MolPort	Figure 4A
	MolPort-044-710-584	0.193	MolPort	Figure 4A
	MolPort-045-965-232	0.193	MolPort	Figure 4A
	MolPort-010-912-491	0.196	MolPort	Figure 4A
	Z19316285	0.196	Enamine	Figure 4A

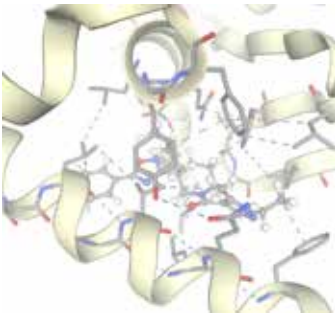
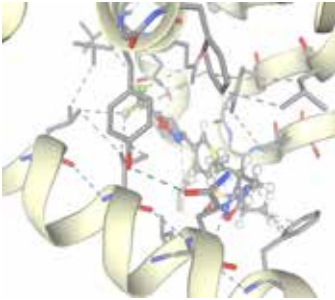
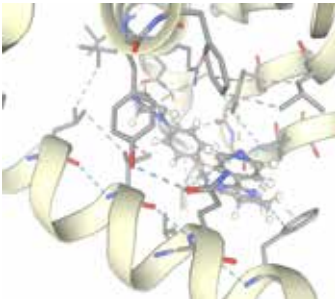
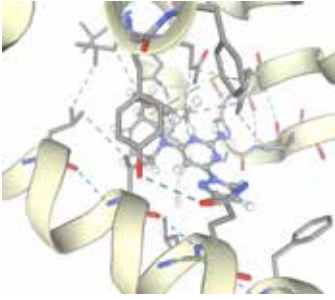
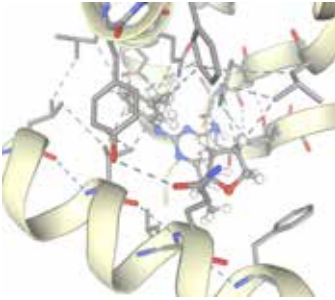
With the unique sets of interactions from Figure 4, Table 3 consists of potential compounds that fulfill the unique pharmacophore maps. The RMSD value indicates the severity of deviation, where the closer the RMSD value is to 0, the less deviation the

compound has with its set of pharmacophores. The RMSD values are arranged in ascending order, therefore the compounds with the least deviation are sorted to the top, and increasingly more deviated compounds are towards the lower end of the table.

Table 4. Nineteen compounds taken to evaluate spontaneity by a method of attracting cavities. For each compound, the SwissParam value of the most spontaneous cluster was taken, as well as a three dimensional protein- ligand model of the hydrogen bonds (blue), ionic interactions (yellow), cation interactions (orange), hydrophobic contacts (gray), and π -stacking interactions (white). The table is sorted from most spontaneous (top) to least spontaneous (bottom)

Name of Compound	Spontaneity/ Swissparam	Structural Model
FUNURESQIKQOAP-UHFFFAOY- SA-N	−9.5550	
BUDVOCFTHVHXB-T-UHFFFAOY- SA-N	−9.5074	
HUNPSCXEFDCREC-UHFFFAOY- SA-N	−9.4277	
SKQBPFUHTWYDZ-UHFFFAOY- SA-N	−9.4168	

Name of Compound	Spontaneity/ Swissparam	Structural Model
MolPort-005–163–256	–9.0517	
VIBCQNPAWUBELN-UHFFFAOY- SA-N	–9.0194	
MolPort-045–965–232	–8.9489	
MolPort-010–912–491	–8.8943	
Z19316285	–8.8594	

Name of Compound	Spontaneity/ Swissparam	Structural Model
MolPort-046-059-419	−8.8570	
MolPort-044-685-945	−8.7693	
Z1359373132	−8.6988	
Z1618187599	−8.6578	
MolPort-044-710-584	−8.5538	

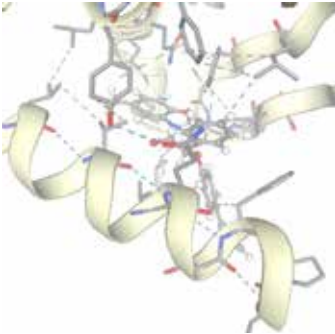
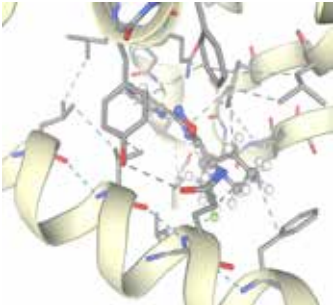
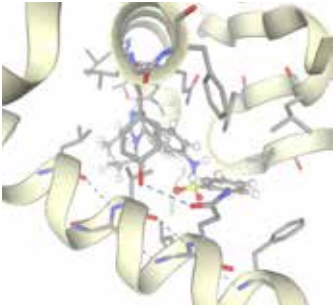
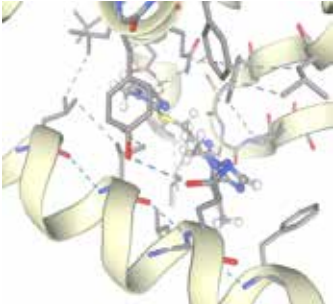
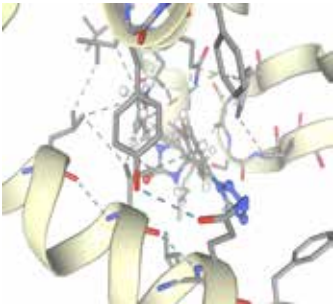
Name of Compound	Spontaneity/ Swissparam	Structural Model
Z54926392	−8.5050	
MolPort-010–950–419	−8.4912	
Z299686082	−8.3999	
MolPort-005–108–116	−8.3253	
Z134661322	−8.2139	

Table 5. Eight compounds with its molecular weight, number of hydrogen bond acceptors, number of hydrogen bond donors, average log p value, and solubility class taken into account. These factors are used to determine whether it fulfills Lipinski's rule; If a compound does in fact fulfill Lipinski's rule, it is a compound that can be successfully absorbed, distributed and excreted in the human body

Compound Name	Molec. Weight (G/Mol)	# Of Hydrogen Bond Acceptors	# Of Hydrogen Bond Donors	Log P Value	Solubility Class	Lipinski's Rule Applies
FUNUR-ESQIKQOAP-UHFF-FAOYSA-N	513.45	11	0	4.25	Moderately soluble	No
BUDVOCFTH-VHXB-T-UHFFFAOYSA-N	574.63	10	0	3.62	Moderately soluble	No
HUNPSCXEFD-CREC-UHFFFAOYSA-N	505.57	6	0	2.99	Moderately soluble	No
Mol-Port-010-912-491	400.49	5	1	2.45	Moderately soluble	Yes
VIBCQN-PAWUBELN-UHFFFAOYSA-N	478.51	6	0	3.98	Moderately soluble	Yes
Z19316285	378.45	5	0	2.56	Moderately soluble	Yes
Z1618187599	340.38	6	1	0.74	Soluble	Yes
Mol-Port-046-059-419	441.50	7	0	2.99	Moderately soluble	Yes

From Table 9, three compounds do not fulfill Lipinski's rule while five others do. Lipinski's rule relies mainly on four categories: the molecular weight, number of hydrogen bond acceptors, number of hydrogen bond donors, and the average LogP value. To fulfill the weight category, the compound must be under 500 Da; To fulfill the LogP category the compound's value must be under 5; To fulfill the hydrogen bond donors (HBDs) category, the compound must contain less than or equal to five hydrogen bond donors; Finally, to fulfill the hydrogen bond acceptors (HBAs) category, the compound must contain less than or equal to ten hydrogen bond acceptors. Based on the data in the table, the compounds that do not pass Lipinski's Rule of Five are from MCULE-ULTIMATE. Generally, compounds that violate two or more of Lipinski's rule are statistically more likely to have poor membrane permeability. However,

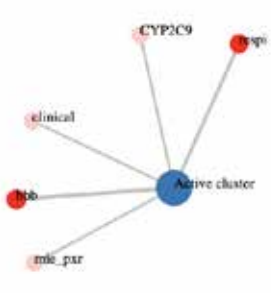
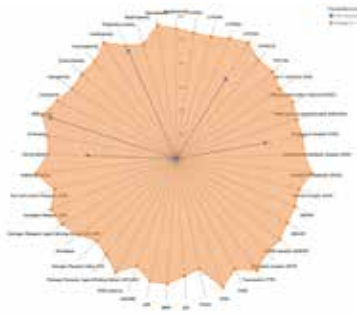
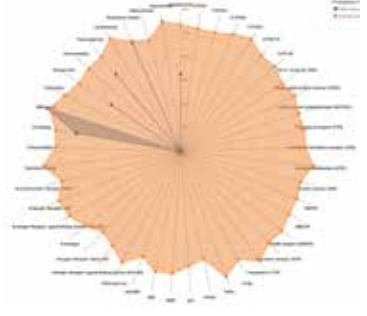
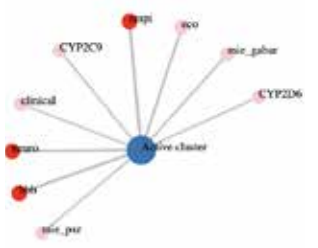
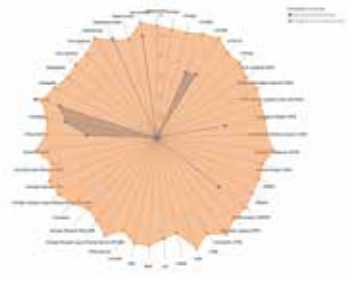
in the table, as long as the compound violates one category, it is considered to fail the Lipinski's Rule.

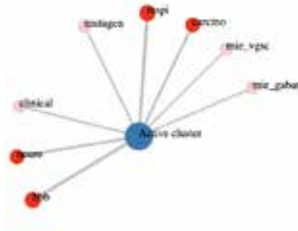
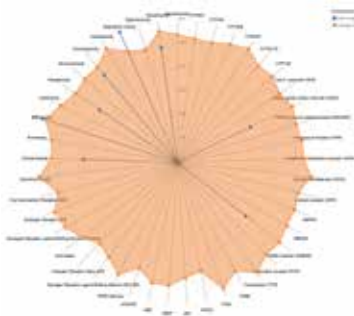
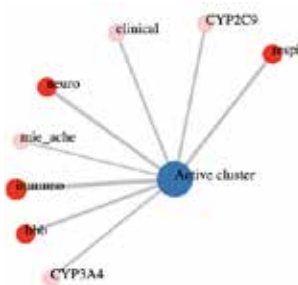
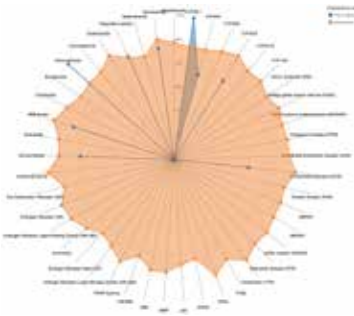
Out of the five compounds taken for evaluation of toxicity in Table 10, only one compound, MolPort-010-912-491, has a probability where all its main activities were less than average for its toxicity class. Three other compounds, Z1618187599, MolPort-046-059-419, and Z19316285 all have one average toxicity level slightly over the average. Though it is over average, the three compounds still pass as possible compounds that can be used for future medicinal developments, only with a limitation for the activity that it is over average for. VIBCQN-PAWUBELN-UHFFFAOYSA-N cannot be used at all since it has two activities where it is way above average for its respective toxicity class. This is illustrated in the toxicity radar chart, where the large orange-shaded region

symbolizes the average toxicity of its class for a certain activity, and the blue dots with gray lines symbolize the toxicity of the compound for the given activities. For some compounds, there are gray regions which represent the toxicity for the compound, and only occurs when the activities of the compound are adjacent in the chart. This includes respiratory toxicity, BBB-barrier, clinical toxicity, CYP2C9, and pregnane X receptor. If one of the compound's main activities were higher than average, then the compound is highly toxic for that activity. For the network chart in the

table, the lighter red colors activities indicate that the probability for that activity is <0.7 while the brighter red colors indicate a probability of ≥ 0.7 . There are only six toxicity classes, 1,2,3,4,5, and 6. If a compound's toxicity class is 3 and higher, then it is considered a safe compound for drug usage. For Table 10, all compounds are above toxicity class 3. All chemicals should have an LD50 of over 300mg/kg; if it does not go beyond that threshold, it is classified as much more toxic due to its ability to cause significant harm, or even death, in small amounts.

Table 6. Five compounds that passed Lipinski's Rule of Five taken to evaluate level of toxicity. Factors taken into account include the Lethal Dose 50% (LD50), which represents the single dose required to kill 50% of the population; toxicity class, a numerical class represented from 1–6; network chart of the active cluster, which ties the compound to potential toxicity targets; the toxicity radar chart, which illustrates the confidence of positive toxicity compared to the average of its class

Name	Ld50 (Mg/Kg)	Tox-icity Class	Network Chart of Active Cluster	Toxicity Radar Chart
Mol-Port-010-912-491	550	4		
Z19316285	1000	4		
Mol-Port-046-059-419	1250	4		

Name	Ld50 (Mg/Kg)	Toxicity Class	Network Chart of Active Cluster	Toxicity Radar Chart
Z1618187599	3000	5		
VIBCN-PAWUBELN-UHFF-FAOYSA-N	6000	6		

Conclusion

Inhibition of USP1 will disrupt DNA repair processes and increase the cancer cell susceptibility to DNA damage. To identify potential compounds capable of binding USP1, a combination of structural, machine learning, pharmacophore, and virtual screening approaches was used to characterize potential binding sites and prioritize candidate molecules. Geometric analysis identified six potential druggable binding pockets, with P_0 and P_1 having the highest drug scores of 0.84, while machine learning analysis identified four predicted binding sites, with B_0 receiving the highest prediction score of 9.01. These results provided multiple potential regions for ligand binding and guided the development of three pharmacophore models containing combinations of aromatic, hydrophobic, and hydrogen-bonding interactions. Screening of the Enamine, MolPort, and MCULE-ULTIMATE libraries identified 19 compounds with RMSD values below 0.2, with Z1359373132 showing the lowest RMSD of 0.100. SwissDock analysis then evaluated the predicted spontaneity of these compounds, with FUNURESQIKQOAP-UHFF-

FAOYSA-N producing the most negative SwissParam value of -9.5550.

Further screening based on drug-likeness and toxicity identified MolPort-010-912-491 as the lead compound and Z1618187599 as the backup compound. MolPort-010-912-491 had an RMSD of 0.196 and a SwissParam value of -8.8943, with a molecular weight of 400.49 g/mol, five hydrogen-bond acceptors, one hydrogen-bond donor, and a LogP of 2.4, allowing it to fulfill Lipinski's Rule of Five. It also had an LD50 of 550 mg/kg and toxicity class 4, with all of its main predicted toxicity activities below the average for its toxicity class. Z1618187599 had an RMSD of 0.185 and a SwissParam value of -8.6578, along with a molecular weight of 340.38 g/mol, six hydrogen-bond acceptors, one hydrogen-bond donor, and a LogP of 0.74. It also has a LD50 of 3000 mg/kg and toxicity class 5. Although other compounds, including Z19316285 and MolPort-046-059-419, also fulfilled Lipinski's Rule of Five, MolPort-010-912-491 was the strongest overall candidate because it was the only compound whose main predicted toxicity activities were all below the average of its toxicity class.

Despite these favorable computational results, this study has several limitations. The virtual screening process evaluated only the compounds that were selected through the pharmacophore and docking workflow, meaning that additional potential USP1 inhibitors may exist among the larger chemical libraries that were not investigated. Furthermore, the ADME and toxicity results were predictions and cannot fully represent how these compounds will behave in a living organism. The toxicity radar charts also showed that some compounds had individual pre-

dicted activities above the average for their respective toxicity classes, demonstrating that additional safety concerns may remain even among compounds with favorable overall toxicity classifications. Future research should therefore screen a larger number of compounds and experimentally validate the lead and backup candidates. Overall, this study successfully narrowed thousands of possible small molecules to two promising USP1 inhibitor candidates, providing a computational starting point for the development of new targeted cancer therapies.

References

- Amjad MT, Chidharla A, Kasi A. (2023). Cancer Chemotherapy. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/33232037>
- Banerjee, P., Kemmler, E., Dunkel, M., & Preissner, R. (2024). ProTox 3.0: a webserver for the prediction of toxicity of chemicals. *Nucleic acids research*, – 52(W1). – W513–W520. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/38647086>
- Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, – 74(3). – P. 229–263. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/38572751>
- Bugnon, M., Röhrig, U. F., Goullieux, M., Perez, M. A. S., Daina, A., Michielin, O., & Zoete, V. (2024). SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic acids research*, – 52(W1). – W324–W332. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/38686803>
- Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, – 7. – 42717 p. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/28256516>
- Ehrt, C., Schulze, T., Graef, J., Diedrich, K., Pletzer-Zelgert, J., & Rarey, M. (2025). Pro-teinsPlus: a publicly available resource for protein structure mining. *Nucleic acids research*, – 53(W1). – W478–W484. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/40326518>
- Gu, M., Sun, S., You, Q., & Wang, L. (2023). Forward or Backward: Lessons Learned from Small Molecule Drugs Approved by FDA from 2012 to 2022. *Molecules (Basel, Switzerland)*, – 28(24). – 7941 p. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/38138431>
- Hu, X., Wu, Y., Yao, M., Chen, Z., & Li, Q. (2025). The other side of the coin: protein deubiquitination by Ubiquitin-Specific Protease 1 in cancer progression and therapy. *Future medicinal chemistry*, – 17(3). – P. 329–345. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/39819213>
- Huang, M. L., Shen, G. T., & Li, N. L. (2022). Emerging potential of ubiquitin-specific proteases and ubiquitin-specific proteases inhibitors in breast cancer treatment. *World journal of clinical cases*, – 10(32). – P. 11690–11701. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/36405275>
- Huang, P., Wang, Y., Zhang, P., & Li, Q. (2023). Ubiquitin-specific peptidase 1: assessing its role in cancer therapy. *Clinical and experimental medicine*, – 23(7). – P. 2953–2966. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/37093451>

- Jendele, L., Krivak, R., Skoda, P., Novotny, M., & Hoksza, D. (2019). PrankWeb: a web server for ligand binding site prediction and visualization. *Nucleic acids research*, – 47(W1). – W345–W349. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/31114880>
- Kashyap, V. K., Sharma, B. P., Pandey, D., Singh, A. K., Peasah-Darkwah, G., Singh, B., Roy, K. K., Yallapu, M. M., & Chauhan, S. C. (2024). Small Molecule with Big Impact: Metarrestin Targets the Perinucleolar Compartment in Cancer Metastasis. *Cells*, – 13(24). – 2053 p. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/39768145>
- Kitamura H. (2023). Ubiquitin-Specific Proteases (USPs) and Metabolic Disorders. *International journal of molecular sciences*, – 24(4). – 3219 p. Retrieved from URL: <https://pubmed.ncbi.nlm.nih.gov/36834633>
- Lemonidis, K., Arkinson, C., Rennie, M. L., & Walden, H. (2022). Mechanism, specificity, and function of FANCD2-FANCI ubiquitination and deubiquitination. *The FEBS journal*, 289(16), 4811–4829. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/34137174>
- Mazloumi Aboukheili, A. M., & Walden, H. (2025). USP1 in regulation of DNA repair pathways. *DNA repair*, 146, 103807. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/39848025>
- Pedreñe, A., Mosquera-Sulbarán, J., Muñoz, N., Tene, D., & Robalino, J. (2021). Nanoantibodies: small molecules, big possibilities. *Biotechnology*, 102(3), 321–336. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/36606147>
- Röhrig, U. F., Goullieux, M., Bugnon, M., & Zoete, V. (2023). Attracting Cavities 2.0: Improving the Flexibility and Robustness for Small-Molecule Docking. *Journal of chemical information and modeling*, 63(12), 3925–3940. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/37285197>
- Sahu M, Suryawanshi H. (2021). Immunotherapy: The future of cancer treatment. *J Oral Maxillofac Pathol*, 25(2), 371. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/34703141>
- Siegel RL, Kratzer TB, Wagle NS, Sung H, Jemal A. (2026). Cancer statistics, 2026. *CA Cancer J Clin*, 76(1), e7004. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC12798275>
- Sonkin, D., Thomas, A., & Teicher, B. A. (2024). Cancer treatments: Past, present, and future. *Cancer genetics*, 286–287, 18–24. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/38909530>
- Sunseri, J., & Koes, D. R. (2016). Pharmit: interactive exploration of chemical space. *Nucleic acids research*, 44(W1), W442–W448. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/27095195>
- Torrado, C., Ashton, N. W., D’Andrea, A. D., & Yap, T. A. (2025). USP1 inhibition: A journey from target discovery to clinical translation. *Pharmacology & therapeutics*, 271, 108865. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/40274197>

submitted 27.07.2026;

accepted for publication 11.08.2026;

published 30.09.2026

© Bethany Ma

Contact: bxyma2016@gmail.com