

2024 S.T. Yau High School Science Award (Asia)

Research Report

The Team

Registration Number: Chem-126

Name of team member: Cyrus Ng

School: Island School

Name of team member: Dorottya Papp

School: Island School

Name of supervising teacher: Gareth Cawson

Job Title: Chemistry Teacher

School: Island School

Title

Investigating Binding Affinities of the HPV E6-Associated Protein LXXLL Motif and E7 LXCXE Motif with Quinolines via Molecular Docking

Date

14 August 2024

Investigating Binding Affinities of the HPV E6-Associated Protein LXXLL Motif and E7 LXCXE Motif with Quinolines via Molecular Docking

Cyrus Ng, Dorottya Papp

Abstract

The HPV oncoproteins E6 and E7 bind to and induce ubiquitination in p53 tumour suppressor proteins and retinoblastoma proteins, causing malignant HPV-associated tumours by slowing apoptosis and promoting rapid cell division. The primary binding sites for drug targeting are the LXXLL motif on E6-associated proteins and the LXCXE motif on E7. Quinolines are a class of drugs that have shown the potential to bind to these motifs due to their pi-stacking aromatic rings, hypercyclicality, and hydrophobic areas. Molecular docking by Autodock Vina was performed on the two motifs with 64 FDA-approved quinolines. Benchmark-normalised Z-score analysis was then used to compare binding affinities on the two motifs and rank the quinolines relative to a benchmark drug. Nelfinavir, which has shown the ability to inhibit both oncoproteins in *in vitro* studies, was used as the benchmark drug. The overall data exhibited a positively skewed normal distribution, with the highest average Z-scores obtained by simeprevir, capmatinib, saquinavir, and dihydroergotamine. Simeprevir exhibited the highest binding affinities relative to the benchmark, with an average Z-score of 4.55. Investigating the combined binding affinities of quinolines with oncoproteins *in silico* identified new drugs for further analysis with molecular dynamics and *in vitro* studies, with possible scope to enhance HPV treatment through targeted therapeutic techniques.

Keywords: HPV, Quinoline, E6, E6AP, E7, LXXLL Motif, LXCXE Motif, Apoptosis, Oncoprotein, In-silico, Binding Affinity, Targeted Therapeutics

Acknowledgement

We thank our supervising Chemistry teacher, Mr. Gareth Cawson, for his help in reviewing and guiding this research project. He read deeply into the project and ensured our science was accurate. We also thank Mr. Howard Chow, a biomedical engineer, for commenting on our initial literature review. Finally, we thank our families and friends for their support and patience during the long project process.




2024 S.-T. Yau High School Science Award
仅用于2024丘成桐中学科学奖公示

Commitments on Academic Honesty and Integrity

We hereby declare that we

1. are fully committed to the principle of honesty, integrity and fair play throughout the competition.
2. actually perform the research work ourselves and thus truly understand the content of the work.
3. observe the common standard of academic integrity adopted by most journals and degree theses.
4. have declared all the assistance and contribution we have received from any personnel, agency, institution, etc. for the research work.
5. undertake to avoid getting in touch with assessment panel members in a way that may lead to direct or indirect conflict of interest.
6. undertake to avoid any interaction with assessment panel members that would undermine the neutrality of the panel member and fairness of the assessment process.
7. observe the safety regulations of the laboratory(ies) where we conduct the experiment(s), if applicable.
8. observe all rules and regulations of the competition.
9. agree that the decision of YHSA(Asia) is final in all matters related to the competition.

We understand and agree that failure to honour the above commitments may lead to disqualification from the competition and/or removal of reward, if applicable; that any unethical deeds if found, will be disclosed to the school principal of team member(s) and relevant parties if deemed necessary; and that the decision of YHSA(Asia) is final and no appeal will be accepted.


Name of team member: Cyrus Ng

Name of team member: Dorottya Papp

Name of supervising teacher: Gareth Cawson

Noted and Endorsed by:

Principal Stephen Loggie


Table of Contents

Abstract	2
Acknowledgement	3
Commitments on Academic Honesty and Integrity	4
Literature Review	6
Methodology	10
Results and Discussion	12
Limitations & Future Research	19
References	20
Appendix Contents	23

Literature Review

In the world of targeted therapeutics, HPV is a budding focus. With 12 high-risk strains and 200 additional variants, no single vaccine currently in circulation offers universal protection. The 12 high-risk strains, specifically HPV 16 and HPV 18, are the leading cause of 5 types of cancer, including cervical cancer [1]. The virus attacks the stratified epithelium in the basal layer, creating the infamous warts on the skin. Aided by E6 and E7 oncoproteins, some lesions may become malignant and cause associated cancers [2]. Developing targeted therapeutics for the oncoproteins associated with HPV would aid in treating HPV variants not protected by vaccines and provide an alternative to those unable to receive the vaccine regimen.

E6 and E7 are viral oncoproteins that induce malignant HPV cancers by disrupting cell division and apoptosis rates to promote tumour growth [3]. E7 inhibits the retinoblastoma protein (pRb). This protein prevents excessive cell division by inducing cellular senescence, inhibiting transcription factors that push cells into the S (synthesis) phase of their cell cycle [4]. On the other hand, E6 inhibits p53, a tumour suppressor protein that initiates apoptosis and DNA repair. Thus, when these oncoproteins successfully bind to their targets, they induce cancerous behaviour. The two oncoproteins are required for malignant cancer's expansion and survival.

The E7 oncoprotein's function is to degrade pRb, which leads to uncontrolled division of cancerous cells. The component which carries out this function through binding to pRb and inducing ubiquitination is the LXCXE motif, consisting of leucine (L), any amino acid (X), cysteine (C), any amino acid (X) and glutamic acid (E). The LXCXE motif is located in E7's conserved region 2 (CR2), a sequence which remains unchanged with different strains of HPV. Like E6, E7 contains zinc finger domains that provide structural stability to carry out its function. Experimental evidence shows the E7 contains one significant zinc finger domain located in the conserved region 3 (CR3), towards the carboxy-terminal end of the protein. This domain contains two CXXC motifs separated by 29 amino acid residues [5].

The E6 oncoprotein's function is to facilitate the degradation of the p53 tumour suppressor gene, which regulates the cell cycle and promotes apoptosis. The E6 oncoprotein is supported by enzymes called E6-associated proteins (E6APs) that catalyse ubiquitination, a

process which marks the p53 tumour suppressor gene for proteasomal degradation by transferring ubiquitin from E2 ubiquitin-conjugating enzymes via thioester bonds [6]. For ubiquitination, the E6AP must form a functional complex with an E6 oncoprotein for greater stability. This formation of the E6-E6AP complex occurs through binding between the LXXLL motif on E6AP and zinc finger domains on E6 oncoproteins located near its N-terminal and C-terminal [7]. The LXXLL motif is characterised by leucine (L) residues at positions 1, 4, and 5, with any two amino acids (X) in between, and is often recognised as a critical part of protein-protein interactions. On the other hand, zinc finger domains are small protein structural motifs that coordinate zinc ions through cysteine molecules to stabilise their folding⁴. Crystallographic analysis shows that E6 zinc finger domains form a deep hydrophobic pocket which stably binds to the LXXLL motifs due to the similarly hydrophobic nature of leucine side chains and the exclusion of water molecules. Furthermore, lysine and arginine residues surrounding the zinc finger domains contain positive charges, which interact with negative charges on the LXXLL motifs' glutamic residues, providing further stability for the E6-E6AP complex [8].

One class of compounds that have the potential to bind to the LXXLL motif are quinolines. Quinolines are organic compounds with one benzene ring, and one pyridine ring joined at two adjacent carbons. The basic structure of quinoline is shown in Figure 1, though there are many quinoline derivatives with various side groups.

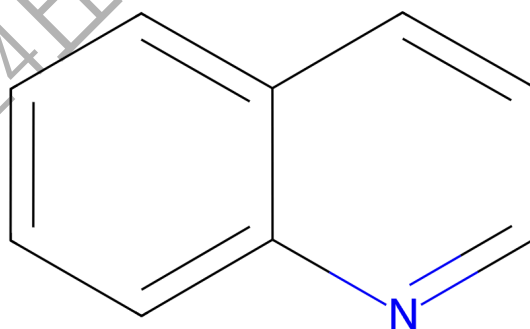


Fig. 1: Chemical structure of quinoline

For medicinal purposes, quinolines are contained in many anti-malarial drugs and are used to manufacture pellagra-preventing nicotinic acid [5]. Currently, 65 quinoline-containing drugs are FDA-approved. By identifying quinoline-based drugs that could potentially inhibit E6AP, drug repurposing could take place. Drug repurposing reduces risk and cost as a

clinically tested and safe chemical is used. If successful, it reduces the time between an identified issue and viable treatment being distributed to patients.

Quinolines have several properties that make them a possible candidate for binding to the LXXLL and LXCXE motifs. Firstly, the heterocyclic nature of quinoline allows it to mimic the helical structure of the two motifs. The drug can fit better into the pocket and create a stronger bond by being cyclical. This would also increase the drug's efficacy as an inhibitor as it would compete with the E6 protein/pRb and thus disrupt the pathway [9]. Secondly, aromatic rings can form non-covalent attractive forces via interactions called pi-stacking. An 'offset stack' forms due to the alignment of positive and negative electrostatic potential [10], and thus, a stronger force of attraction is created between the drug and E6AP/E7. Lastly, the motifs contain hydrophobic pockets. Quinoline compounds have been evidenced to bind to various hydrophobic spaces, such as a plasma protein called human serum albumin [11]. An isoquinoline (isoquinoline-3-carboxamide derivative), an isomer of quinoline with the same aromatic rings, was found to be a possible anti-tumour lead, which further prompts this investigation into whether quinolines could potentially do the same [12]. Drug repurposing investigations typically start from the *in silico* stage, specifically molecular docking, to discover binding affinity.

Binding free energy (ΔG_{bind}), which in this case equates to binding affinity, quantifies the strength and stability of interaction between molecules such as drugs, which are ligands, and proteins. It is the difference in free energy, or ability to do work, between the molecules' bound state (complex) and unbound states. The equation of binding free energy is as follows:

$$\Delta G_{bind} = \Delta G_{complex} - (\Delta G_{protein} + \Delta G_{ligand})$$

$\Delta G_{complex}$ is the free energy of the protein-ligand complex, while $\Delta G_{protein}$ and ΔG_{ligand} are the free energies of the unbound proteins and ligands, respectively [13].

A negative binding affinity suggests a spontaneous and favourable binding interaction, whereas a positive value indicates that the binding is not favourable. By understanding the binding affinities between quinoline-based drugs and binding sites of associated oncoproteins, researchers can gain insight into the strength of binding and, therefore, the

potential effectiveness of the drug as a competitive inhibitor at specific binding sites, providing clues for therapeutic targets. In one experiment by Wang, Baddock, Mafi et al., the binding affinity between E6AP and 16E6 at the site of the LXXLL motif was recorded at -7.78 kcal/mol [14].

This experiment aims to find a potentially inhibiting drug, which would be evident by matching the binding affinity of a benchmark known to previously inhibit the oncoproteins or have a more negative binding affinity than the literature value. A benchmark drug identified was nelfinavir, a drug found to target E6 and E7 in an investigation from 2021. Nelfinavir successfully inhibited E6 and E7 and increased p53 levels, indicating an interruption in the cancer-promoting metabolic activities of the oncoproteins. This is due to binding to the LXXLL and LXCXE motifs on the oncoproteins [15].

The literature reveals the significance of effective binding between LXXLL motifs on E6APs and zinc finger domains on E6 oncoproteins. If quinolines have sufficient binding free energy with either component, they could act as a competitive inhibitor for forming the E6-E6AP complex and, therefore, ubiquitination and eventual degradation of the p53 tumour suppressor gene. More importantly, they may also bind to the LXCXE motif on E7, inhibiting both tumour-promoting processes associated with HPV. Hence, this study will investigate the binding free energy between quinolines and the binding regions of E6AP and E7 to discover the most promising drug to oppose and inhibit HPV-associated cancers.

Methodology

To discover which quinoline is the most effective against HPV oncogenesis, docking simulations *in silico* with 64 of the 65 quinolines (saquinavir mesylate could not be docked due to errors in the downloadable files) on the E6AP LXXLL motif and E7 LXCXE motif were performed. This study will then use benchmark-normalised z-score analysis to combine data from drug interactions with E6 and E7 and suggest quinolines that may effectively inhibit both proteins. If successful, further medical research into these drugs could be undertaken.

Performing static molecular docking simulations between quinolines and LXXLL motif on E6AP and LXCXE motif on E7 will produce binding affinity values. The quinolines were determined by a search on the ChEMBL database, where the parameters were set to only the approved drugs rather than any in stages of clinical trials. The E6AP and E7 files were

downloaded as .pdb from the Protein Data Bank and UniProt. Once obtained, the files were uploaded to pyMOL to reveal their amino acid sequence, where the coordinates of the LXXLL and LXCXE motifs (found as LQELL and LYCYE) were identified, as seen in Figures 2 and 3.

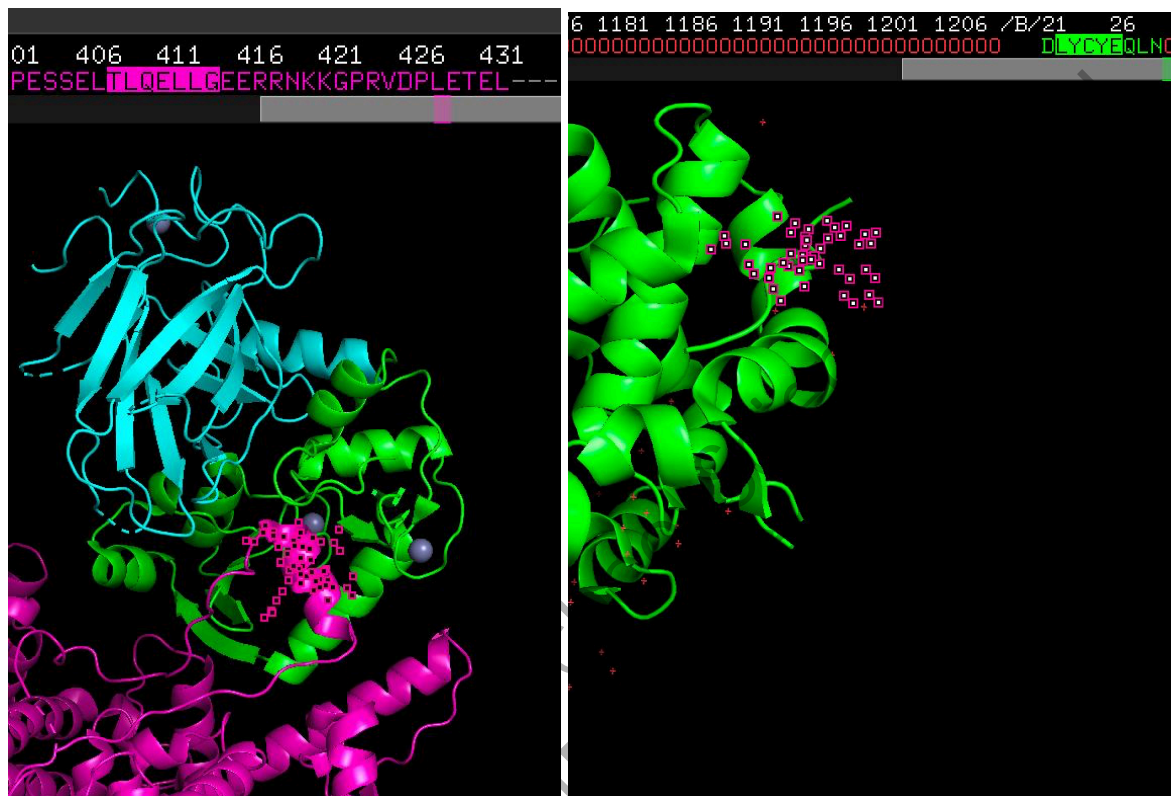


Fig. 2 (left): Structure of E6AP as displayed on PyMOL. The LXXLL motif is highlighted in the sequence.

Fig. 3 (right): Structure of E7 as displayed on PyMOL. The LXCXE motif is highlighted in the sequence.

The 3D conformational structure of each quinoline was downloaded as a .sdf file from PubChem4, which was converted to .pdbqt through OpenBabel5. All files were uploaded to Autodock Tools6, where water was removed to prevent interference in the complexes, hydrogens were added as crystal structures often lack them, Gastigier charges were computed to standardise the baseline of electrostatic interactions and the outputs were downloaded as .pdbqt files. Finally, Autodock Vina produced an output containing binding affinities of different orientations. As the orientation with the lowest (most negative) binding affinity is the most thermodynamically favourable, it was the only one logged. Each docking interaction between drug and protein was repeated and averaged 20 times via a Python computer programme, as shown in Figure 4.

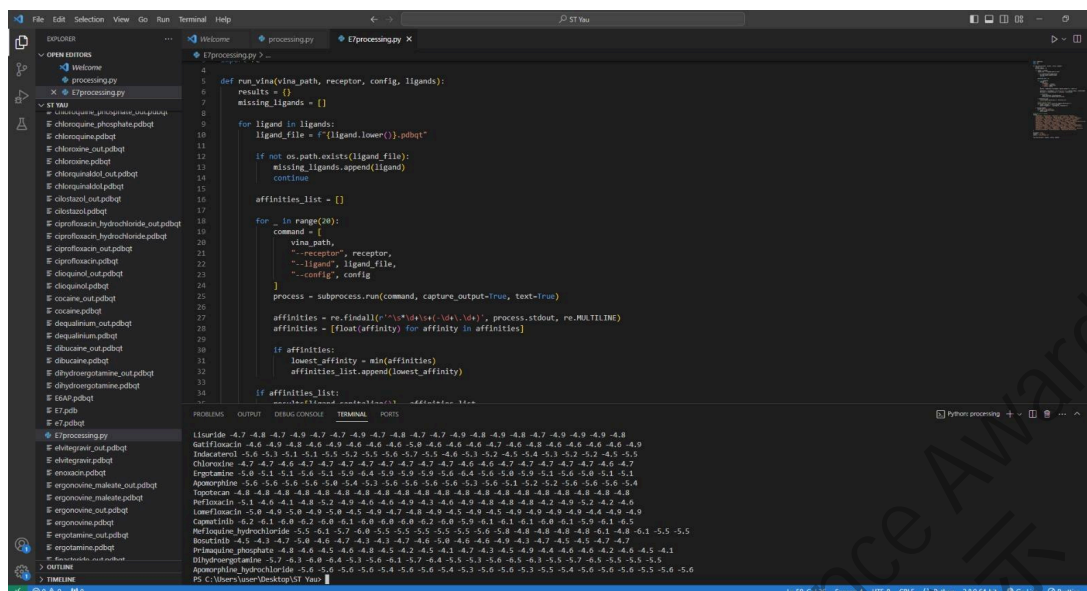


Fig. 4 Python computer programme used to run Autodock Vina.

The novel part of this study used benchmark-normalised Z-score analysis to compare the combined binding affinities on E6AP LXXLL and E7 LXCXE motifs of different quinolines relative to nelfinavir, the benchmark drug shown to inhibit both oncoproteins. The Z-score indicates the distance of data points from the mean in standard deviations, providing a relative comparison metric. This method was a modified statistical approach from a machine learning study on breast cancer detection, which similarly uses Z-scores to normalise the impact of more than one biological factor.

First, Z-scores were calculated for the binding affinities between drug and protein with the formulae $Z = (v - \mu) / \sigma$, representing the Z-score, v the binding affinity value for that interaction, μ the mean, and σ the standard deviation of all binding affinities. The mean and standard deviation were calculated separately for binding interactions with E6 and E7. The Z-score of nelfinavir for E6 and E7 was then subtracted from each drug's Z-score to yield each quinoline's binding affinity relative to nelfinavir. Because the Z-score conversions automatically normalise the two datasets, the Z-score of each drug with E6 and E7 were simply averaged and plotted on a bar chart. Although E6 and E7 data can be differently weighted in the final calculations, they will be considered equally in this study because the view of which is more essential to the process is contested [16].

Results and discussion

Drug	E6AP LXXLL	Z-Score	Z-Score - Benchmark
Simeprevir	-7.4	4.5	4.41
Capmatinib	-6.1	2.1	1.99
Saquinavir	-5.7	1.3	1.18
Dihydroergotamine	-5.8	1.7	1.55
Risperidone	-5.4	0.8	0.70
Montelukast	-5.7	1.4	1.23
Ergotamine	-5.5	1.1	0.93
Indacaterol Maleate	-5.3	0.7	0.55
Bromocriptine	-5.3	0.6	0.50
Indacaterol	-5.2	0.5	0.41
Nelfinavir (Benchmark)	-5.0	0.1	0.00

Table 1: Z-scores in relation to E6AP LXXLL for the top 10 ranked drugs and the benchmark

Drug	E7 LXCXE	Z-Score	Z-Score - Benchmark
Simeprevir	-10.1	4.9	4.68
Capmatinib	-7.5	1.7	1.53
Saquinavir	-7.6	1.8	1.61
Dihydroergotamine	-7.0	1.1	0.88
Risperidone	-7.1	1.3	1.08
Montelukast	-6.5	0.5	0.32
Ergotamine	-6.7	0.8	0.59
Indacaterol Maleate	-7.0	1.1	0.95
Bromocriptine	-7.0	1.1	0.91
Indacaterol	-7.0	1.1	0.97
Nelfinavir (Benchmark)	-6.3	0.2	0.00

Table 2: Z-scores in relation to E7 LXCXE for the top 10 ranked drugs and the benchmark

Drug	Average Z-Score Relative to Benchmark
------	---------------------------------------

Simeprevir	4.55
Capmatinib	1.76
Saquinavir	1.40
Dihydroergotamine	1.22
Risperidone	0.89
Montelukast	0.78
Ergotamine	0.76
Indacaterol Maleate	0.75
Bromocriptine	0.70
Indacaterol	0.69
Nelfinavir (Benchmark)	0.00

Table 3: Average Z-scores for the top 10 ranked drugs and the benchmark

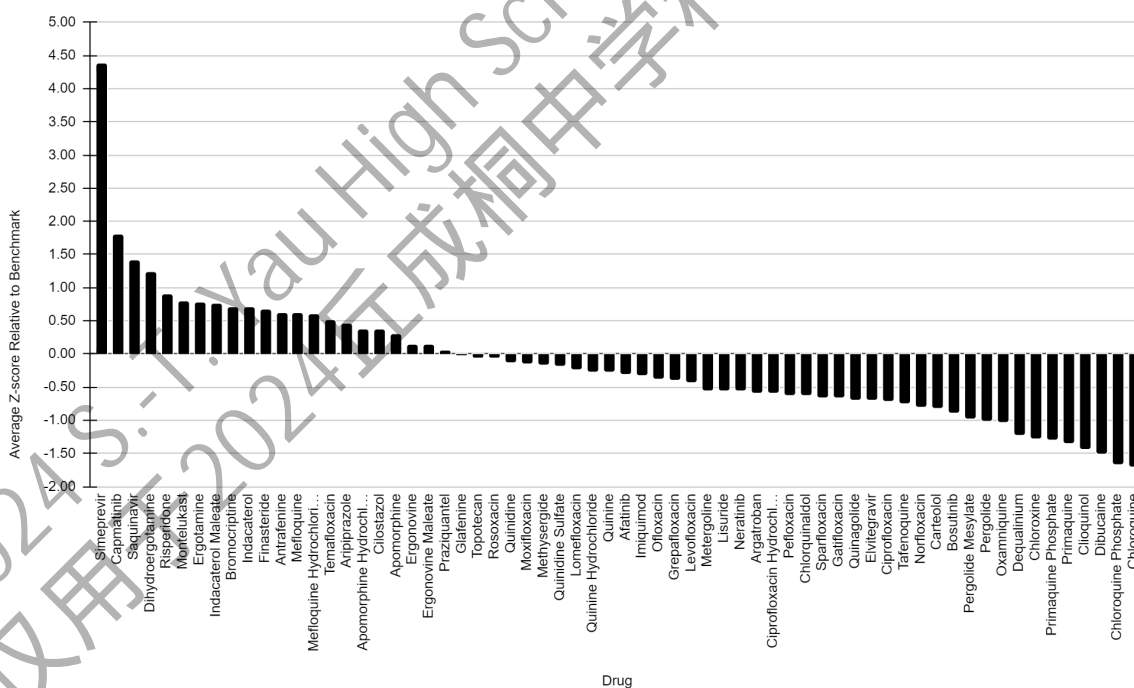


Fig. 5: Bar chart showing average z-score relative to benchmark for tested quinoline drugs. The y-axis denotes the number of standard deviations from nelfinavir's binding affinities.

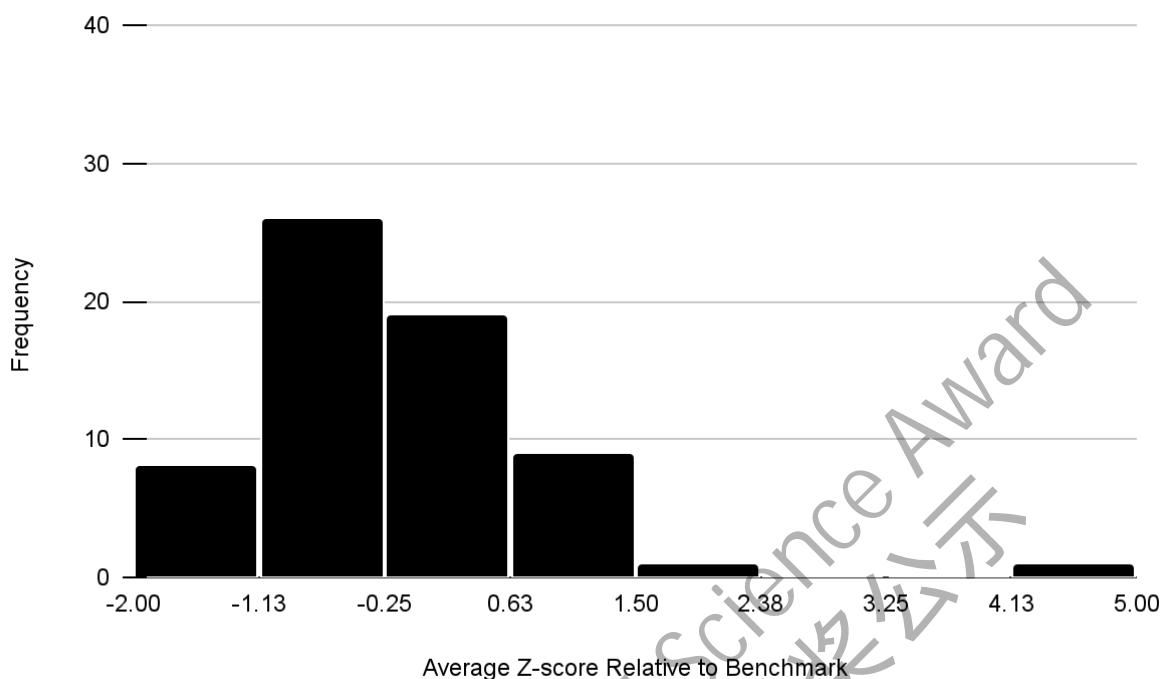


Fig. 6: Histogram of average z-score relative to benchmark for tested quinoline drugs

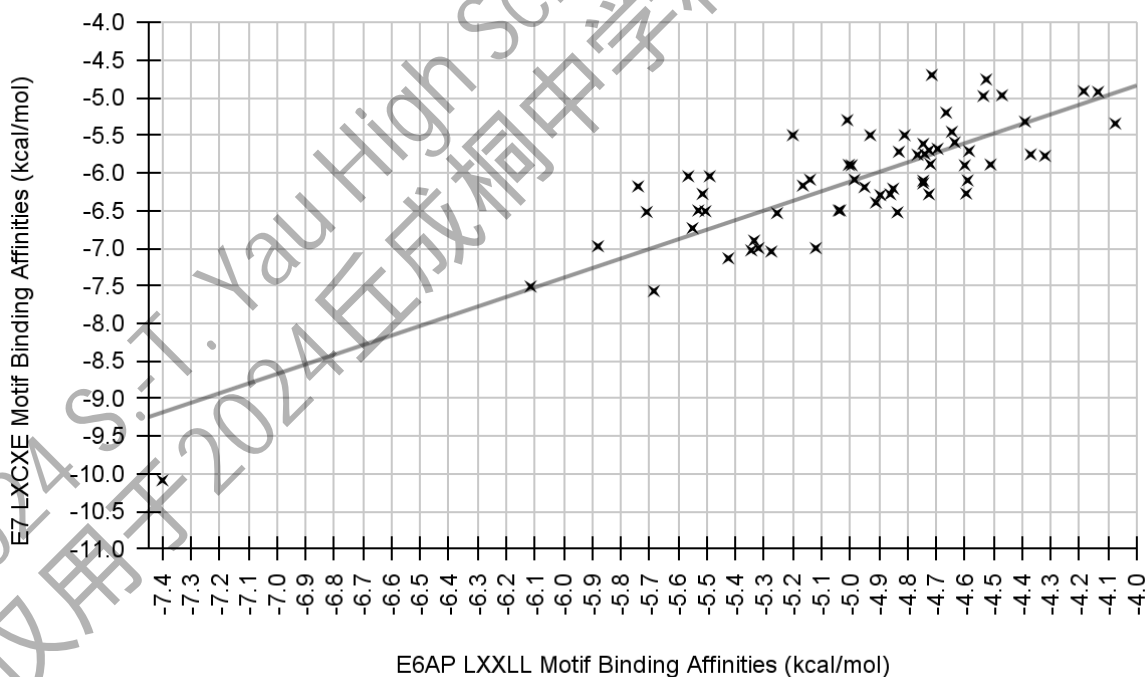


Fig.7: Scatter plot of binding affinities for E6AP LXXLL motif and E7 LXCXE motif interactions. The trend line has an R^2 value of 0.69.

The bar chart in Figure 5 was obtained by applying the methodology to the 64 FDA-approved quinolines. The histogram shown in Figure 6 exhibits a normal distribution with a positive skew, suggesting most extreme values are on the right; specifically, there is a

noticeable rise in binding affinity past the fourth-highest ranking drug, meaning the quinolines with outstanding binding affinities are simeprevir, capmatinib, saquinavir and dihydroergotamine. As seen in Table 3, simeprevir obtained an average Z-score relative to the benchmark of 4.55, meaning it has the lowest combined binding affinity and is the only outlier in the dataset, nearly doubling the binding affinity of the following drug, capmatinib.

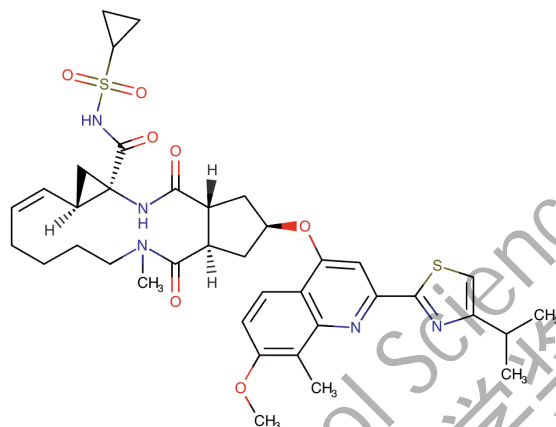


Fig. 8: Chemical structure of simeprevir

Simeprevir (C₃₈H₄₇N₅O₇S₂) is a macrocyclic molecule; its structure is shown in Figure 8. The ringed portion allows for it to fit into the two helical motifs. As with other drugs, the aromatic rings could engage in pi-stacking with the amino acid residues. The methoxy group, an electron-donating group, has the potential to form hydrogen bonds with the motifs. Hydrogen bonds are the strongest intermolecular forces, and a large amount would contribute to a high binding affinity. Other areas that could be involved in hydrogen bonding are the thiazole ring, which contains nitrogen, and the sulfonamide group. As a large molecule, it provides a large interaction interface between itself and the motifs, creating multiple opportunities for chemical interactions. The abundant opportunity for hydrogen bonds and other chemical interactions is most likely the reason for simeprevir's outlier status. Simeprevir also shares many chemical properties with the next-ranked drug, capmatinib.

Capmatinib, available under the brand name Taveclo, is an oral anti-cancer medication first approved in 2020 [17]. It is only used to treat non-small-cell lung cancer in cases where the cells have a particular genetic mutation.

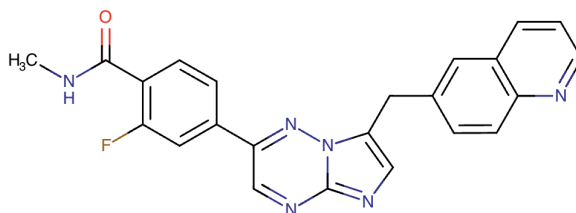


Fig 9: Chemical structure of capmatinib

The chemical formula of capmatinib is $C_{23}H_{17}FN_6O$, and its structure is shown in Figure 9. It contains a quinoline, an imidazole and a triazine ring (imidazo[1,2-B][1,2,4]triazine) as well as a benzamide (2-fluoro-N-methylbenzamide). The tricyclic structure leads it to be rigid. It does have some flexibility provided by the piperazine ring, which could explain the success in binding to both motifs investigated in this paper. There is a methoxy group, which is rich in electrons and thus allows for strong hydrogen bonding. The multiple nitrogens within the structure can act as hydrogen bond acceptors, forming even stronger bonds. Like the other quinoline compounds, aromatic rings bind to hydrophobic pockets and have a cyclic shape to fit into the helical LXXLL motif. However, these are not unique to capmatinib. The quinoline compounds were chosen due to the last two traits, but there is a wide range of binding affinities. This suggests that the quinoline compound is not the deciding factor in whether or not a drug has strong potential.

Capmatinib is a tyrosine kinase inhibitor (TKI), indicative of its anti-cancer properties and uses, like other quinoline-containing drugs on the list afatinib, neratinib and bosutinib. However, they did not score similar values in the simulation, suggesting that capmatinib has unique properties that are not common to TKIs. One is the tricyclic cycle, which creates a larger, more rigid interaction surface, leading to increased interactions between the LXXLL motif and the drug. The structure and components of capmatinib likely create more favourable interactions, such as hydrogen bonding. However, much is unknown about the specifics, so the reasons behind capmatinib's differentiation from other TKIs require more research.

Saquinavir is one of the many drugs used in the regimen to prevent and treat HIV. Due to low oral bioavailability, it is administered with ritonavir, another protease inhibitor, to improve efficacy.

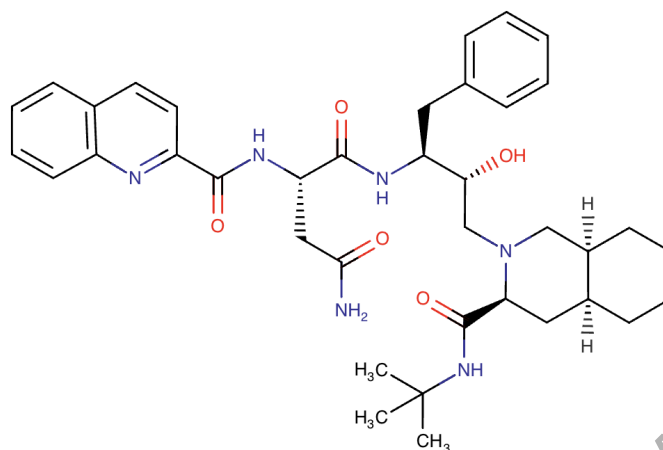


Fig 10: Chemical structure of saquinavir

Saquinavir (C₃₈H₅₀N₆O₅) contains a hydrophobic decahydroisoquinoline group; its structure is shown in Figure 10. This allows it to take up much space in the hydrophobic areas of the LXXLL and LXCXE motifs. This may also contribute to intermolecular Van der Waals forces, which accumulate a strong attraction between the motifs and the drug. The typical quinoline-induced aromatic rings, pi-stacking and helical form all likely contribute to the binding affinity. The amide group also acts as a hydrogen bond donor. Like dihydroergotamine and unlike the previous two drugs, it is highly flexible.

The fourth drug with potential is dihydroergotamine (DHE). DHE is an analgesic medicine administered either through injection or nasal spray. It is not preventative and thus has no reason to be administered unless there is an ongoing headache, as overuse negatively affects many organs.

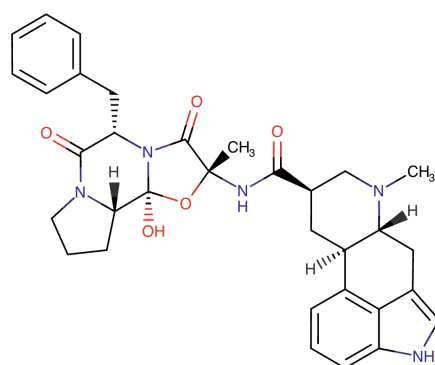


Fig 11: Chemical structure of dihydroergotamine

DHE, like other ergot alkaloids, contains an ergoline core with an indole ring. It has the chemical formula $C_{33}H_{37}N_5O_5$, its structure shown in Figure 11. Although already established as a non-crucial factor, quinolines still possess an aromatic ring that may have been involved in bonding to the helical LXXLL. The shape of the ergoline core is tetracyclic and 3D. It is a large compound with multiple rings and some rotatable bonds, which result in a highly flexible compound. Its large interface and great flexibility could be the reason for its ability to bind very strongly to both E7 and E6AP. Similarly to capmatinib, DHE has many potential spots on its peripheral that could accept hydrogen bonds.

Several conclusions can be drawn from these results by looking at the structures of the two most successful candidates. First, and as mentioned before, quinoline is not the necessary component. However, the hydrophobic regions, such as aromatic rings, are more pronounced in the higher-ranked drugs, indicating it is a potentially significant factor. Secondly, the larger, more complex, multi-ring-containing compounds tend to attain lower binding affinities. The four drugs analysed meet those criteria. This is likely due to the more extensive interaction interface available and the increase in opportunity for favourable chemical interactions. As E6 is another large protein, larger compounds are more likely to work as inhibitors, which seems more promising. pRb, which binds to E7, has an amino acid chain roughly 3 times the size of the average protein, further encouraging this conclusion [18]. Another factor that seems to affect results is the drugs' flexibility. The multiple conformations of the drug can form a better 'fit' for the motifs, much like the induced fit model for enzymes. However, the first two ranked drugs are more rigid than the flexible dihydroergotamine and subsequent drugs, so it is still unclear. Furthermore, simeprevir and capmatinib share several chemical properties beyond quinoline. Both have a methoxy group and rings with ample potential for hydrogen bond accepting, suggesting hydrogen bonds are an abundant force in the binding to the motifs. Lastly, the salt forms of drugs indicate that the overall charge of a drug is not relevant to the binding. If it were, quinine hydrochloride would have a much different average Z-score than quinine, but in the graph here, they are almost indistinguishable. However, this study concludes that simeprevir should be a central focus of future research due to its outstanding binding affinities relative to the benchmark drug and other quinolines.

Finally, a positive correlation exists between binding affinities with E6AP and E7, as shown by the linear trendline with a positive gradient in Figure 7. There are no major anomalies, but

the trend is moderate, with an R^2 value of 0.69. This trend offers two insights: 1) Drugs effective against one of the oncoproteins are also more likely to be effective against the other, and 2) The chemical attributes of the LXXLL and LXCXE motif are likely similar considering each drug has similar relative efficacy in binding to both. These insights are helpful for future drug discovery surrounding HPV E6AP LXXLL and E7 LXCXE motifs.

Limitations & Future Research

This study aimed to discover which quinolines are worthy of further investigation as inhibitors of E6AP and E7, achieved by judging drugs based on relative comparisons with a benchmark. However, there are three major limitations to static docking in the context of HPV E6 and E7 *in silico* drug discovery: 1) the protein surface is kept rigid, meaning it does not simulate the more accurate, induced fit model for protein-ligand interactions, and 2) it is not done in a realistic environment which simulates the human body, such as with regards to solvent and temperature. Molecular dynamics simulations on software like GROMACS, which addresses both concerns, should hence be the next step for the top four quinolines highlighted in this study. Nonetheless, *in vitro* studies, such as the one which confirmed the efficacy of the benchmark drug nelfinavir, are vital in verifying *in silico* data. Furthermore, although this study discussed two of the most promising binding sites on E6AP and E7, other binding sites may still be playing a role in inducing ubiquitination, such as the Taz2 binding site on E7, left untested. These untested binding sites also include allosteric sites offering methods for non-competitive inhibition, where the drug would cause changes in the structure of the oncoproteins. In addition, further docking is needed to discover the potential for quinoline drugs to disrupt the zinc finger domains on E6 and E7, which are responsible for structural stability in binding interactions.

References

1. National Cancer Institute. (2023, April 4). HPV and Cancer. Retrieved July 28, 2024, from National Cancer Institute website:
<https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-and-cancer>
2. McMurray, H. R., Nguyen, D., Westbrook, T. F., & Mcance, D. J. (2001). Biology of human papillomaviruses. *International Journal of Experimental Pathology*, 82(1), 15–33. <https://doi.org/10.1046/j.1365-2613.2001.00177.x>
3. Castel, P., Rauen, K. A., & McCormick, F. (2020). The duality of human oncoproteins: drivers of cancer and congenital disorders. *Nature Reviews Cancer*, 20(7), 383–397. <https://doi.org/10.1038/s41568-020-0256-z>
4. Foley, M. (1998). Quinoline Antimalarials Mechanisms of Action and Resistance and Prospects for New Agents. *Pharmacology & Therapeutics*, 79(1), 55–87. [https://doi.org/10.1016/s0163-7258\(98\)00012-6](https://doi.org/10.1016/s0163-7258(98)00012-6)
5. McLaughlin-Drubin, M. E., & Münger, K. (2009). The human papillomavirus E7 oncoprotein. *Virology*, 384(2), 335–344. <https://doi.org/10.1016/j.virol.2008.10.006>
6. Greenfield, I., Nickerson, J., Penman, S., & Stanley, M. (1991). Human papillomavirus 16 E7 protein is associated with the nuclear matrix. *Proceedings of the National Academy of Sciences*, 88(24), 11217–11221. <https://doi.org/10.1073/pnas.88.24.11217>
7. Vande Pol, S. B., & Klingelhutz, A. J. (2013). Papillomavirus E6 oncoproteins. *Virology*, 445(1-2), 115–137. <https://doi.org/10.1016/j.virol.2013.04.026>
8. Manzo-Merino, J., Thomas, M., Fuentes-Gonzalez, A. M., Lizano, M., & Banks, L. (2013). HPV E6 oncoprotein as a potential therapeutic target in HPV related cancers. *Expert Opinion on Therapeutic Targets*, 17(11), 1357–1368. <https://doi.org/10.1517/14728222.2013.832204>
9. Rogers, K. (2023, June 8). Competitive inhibition | Biochemistry. Retrieved from Encyclopedia Britannica website:
<https://www.britannica.com/science/competitive-inhibition>
10. Illustrated Glossary of Organic Chemistry. (2018, August 14). Illustrated Glossary of Organic Chemistry - Aromatic interaction (aromatic stacking; pi stacking). Retrieved from www.chem.ucla.edu website:
https://www.chem.ucla.edu/~harding/IGOC/A/aromatic_aromatic_interaction.html

11. Wang, R., Hu, X., Pan, J., Gong, D., & Zhang, G. (2018). Interaction between quinoline yellow and human serum albumin: spectroscopic, chemometric and molecular docking studies. *Journal of the Science of Food and Agriculture*, 99(1), 73–82. <https://doi.org/10.1002/jsfa.9144>
12. Gao, F., Liu, H., Li, L., Guo, J., Wang, Y., Zhao, M., & Peng, S. (2015). Design, synthesis, and testing of an isoquinoline-3-carboxylic-based novel anti-tumor lead. *Bioorganic & Medicinal Chemistry Letters*, 25(20), 4434–4436. <https://doi.org/10.1016/j.bmcl.2015.09.014>
13. Raha, K., & Merz, K. M. (2005). Chapter 9 Calculating Binding Free Energy in Protein–Ligand Interaction. *Annual Reports in Computational Chemistry*, 1(9), 113–130. [https://doi.org/10.1016/s1574-1400\(05\)01009-1](https://doi.org/10.1016/s1574-1400(05)01009-1)
14. John, Baddock, H. T., Mafi, A., Foe, I. T., Bratkowski, M., Lin, T.-Y., ... Nile, A. H. (2024). Structure of the p53 degradation complex from HPV16. *Nature Communications*, 15(1). <https://doi.org/10.1038/s41467-024-45920-w>
15. Park, S., Auyeung, A., Lee, D. L., Lambert, P. F., Carchman, E. H., & Sherer, N. M. (2021). HIV-1 Protease Inhibitors Slow HPV16-Driven Cell Proliferation through Targeted Depletion of Viral E6 and E7 Oncoproteins. *Cancers*, 13(5), 949–949. <https://doi.org/10.3390/cancers13050949>
16. Arushi Vats³, Trejo-Cerro, O., Thomas, M., & Banks, L. (2021). Human Papillomavirus E6 and E7: What remains? *Tumour Virus Research*, 11(3), 200213–200213. <https://doi.org/10.1016/j.tvr.2021.200213>
17. Center. (2024). FDA approves capmatinib for metastatic non-small cell lung cancer. Retrieved August 14, 2024, from U.S. Food and Drug Administration website: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-capmatinib-metastatic-non-small-cell-lung-cancer>
18. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2024). The Shape and Structure of Proteins. Retrieved August 14, 2024, from Nih.gov website: <https://www.ncbi.nlm.nih.gov/books/NBK26830/>
19. Serge Pérez, & Igor Tvaroška. (2014). Carbohydrate–Protein Interactions. *Advances in Carbohydrate Chemistry and Biochemistry*, 71(1), 9–136. <https://doi.org/10.1016/b978-0-12-800128-8.00001-7>
20. Jansma, A. L., Martinez-Yamout, M. A., Liao, R., Sun, P., H. Jane Dyson, & Wright, P. E. (2014). The High-Risk HPV16 E7 Oncoprotein Mediates Interaction between the Transcriptional Coactivator CBP and the Retinoblastoma Protein pRb. *Journal of*

- Molecular Biology*, 426(24), 4030–4048. <https://doi.org/10.1016/j.jmb.2014.10.021>
21. Yim, E.-K., & Park, J.-S. (2005). The Role of HPV E6 and E7 Oncoproteins in HPV-associated Cervical Carcinogenesis. *Cancer Research and Treatment*, 37(6), 319. <https://doi.org/10.4143/crt.2005.37.6.319>
22. ChEMBL Database. (2023). Retrieved August 14, 2024, from Ebi.ac.uk website: <https://www.ebi.ac.uk/chembl/>
23. Data, P. (2023). RCSB PDB - 8GCR: HPV16 E6-E6AP-p53 complex. Retrieved August 14, 2024, from Rcsb.org website: <https://www.rcsb.org/structure/8GCR>
24. UniProt. (2024). Retrieved August 14, 2024, from Uniprot.org website: <https://www.uniprot.org/uniprotkb/P03129/entry>
25. OPENBABEL - Chemical file format converter. (2024). Retrieved August 14, 2024, from Cheminfo.org website: <https://www.cheminfo.org/Chemistry/Cheminformatics/FormatConverter/index.html>
26. Trott, O., & Olson, A. J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>
27. Henderi Henderi. (2021). Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer. *IJIS International Journal of Informatics and Information Systems*, 4(1), 13–20. <https://doi.org/10.47738/ijis.v4i1.73>

Appendix Contents

The zip file contains 3 supplementary data tables that were not able to be presented in this document due to lack of clarity and coherence when formatted.

1. [A1] E6AP Binding Affinity Raw Data.pdf
2. [A2] E7 Binding Affinity Raw Data.pdf
3. [A3] Full Processed Data, Z-Scores.pdf

2024 S.-T. Yau High School Science Award
仅用于2024丘成桐中学科学奖公示

Drug	E6AP LXXLL	Z-Score	Z-Score - Benchmark	E7 LXCXE	Z-Score	Z-Score - Benchmark	Average Z-Score Relative to Benchmark
Simeprevir	-7.4	4.5	4.41	-10.1	4.9	4.68	4.55
Capmatinib	-6.1	2.1	1.99	-7.5	1.7	1.53	1.76
Saquinavir	-5.7	1.3	1.18	-7.6	1.8	1.61	1.40
Dihydroergotamine	-5.8	1.7	1.55	-7.0	1.1	0.88	1.22
Risperidone	-5.4	0.8	0.70	-7.1	1.3	1.08	0.89
Montelukast	-5.7	1.4	1.23	-6.5	0.5	0.32	0.78
Ergotamine	-5.5	1.1	0.93	-6.7	0.8	0.59	0.76
Indacaterol Maleate	-5.3	0.7	0.55	-7.0	1.1	0.95	0.75
Bromocriptine	-5.3	0.6	0.50	-7.0	1.1	0.91	0.70
Indacaterol	-5.2	0.5	0.41	-7.0	1.1	0.97	0.69
Finasteride	-5.3	0.6	0.53	-6.9	1.0	0.79	0.66
Antrafenine	-5.7	1.4	1.29	-6.2	0.1	-0.09	0.60
Mefloquine	-5.5	1.0	0.89	-6.5	0.5	0.30	0.60
Mefloquine Hydrochloride	-5.5	1.0	0.85	-6.5	0.5	0.31	0.58
Temafloxacin	-5.1	0.2	0.12	-7.0	1.1	0.91	0.52
Aripiprazole	-5.5	1.0	0.87	-6.3	0.2	0.03	0.45
Apomorphine Hydrochloride	-5.5	1.1	0.96	-6.0	-0.1	-0.26	0.35
Cilostazol	-5.2	0.5	0.38	-6.5	0.5	0.34	0.36
Apomorphine	-5.5	0.9	0.82	-6.0	-0.1	-0.26	0.28
Ergonovine	-5.0	0.1	-0.03	-6.5	0.5	0.30	0.14
Ergonovine Maleate	-5.0	0.1	-0.04	-6.5	0.5	0.30	0.13
Praziquantel	-5.1	0.3	0.21	-6.2	0.1	-0.10	0.05
Glafenine	-5.1	0.3	0.16	-6.1	0.0	-0.20	-0.02
Topotecan	-4.8	-0.3	-0.41	-6.5	0.5	0.53	-0.04
Rosoxacin	-4.9	-0.2	-0.27	-6.4	0.4	0.17	-0.05
Quinidine	-4.9	-0.2	-0.30	-6.3	0.2	0.05	-0.13
Moxifloxacin	-4.9	-0.1	-0.20	-6.2	0.1	-0.08	-0.14
Methysergide	-5.0	0.0	-0.13	-6.1	0.0	-0.20	-0.17
Quinidine Sulfate	-4.8	-0.2	-0.37	-6.3	0.2	0.04	-0.16
Lomefloxacin	-4.8	-0.3	-0.39	-6.2	0.1	-0.06	-0.22
Quinine Hydrochloride	-5.0	0.0	-0.09	-5.9	-0.3	-0.44	-0.27
Quinine	-5.0	0.0	-0.11	-5.9	-0.3	-0.43	-0.27
Afatinib	-4.7	-0.5	-0.62	-6.3	0.2	0.04	-0.29
Imiquimod	-5.2	0.4	0.27	-5.6	-0.7	-0.92	-0.33
Otofloxacin	-4.7	-0.5	-0.58	-6.1	0.0	-0.14	-0.36
Grepafoxacin	-4.7	-0.5	-0.58	-6.1	0.0	-0.18	-0.38
Levofloxacin	-4.6	-0.7	-0.87	-6.3	0.2	0.02	-0.42
Metergoline	-4.6	-0.8	-0.87	-6.1	0.0	-0.19	-0.53
Lisuride	-4.8	-0.3	-0.42	-5.7	-0.5	-0.65	-0.54
Neratinib	-4.7	-0.5	-0.63	-5.9	-0.3	-0.45	-0.54
Argatroban	-4.9	-0.1	-0.24	-5.5	-0.7	-0.92	-0.58
Ciprofloxacin Hydrochloride	-4.7	-0.4	-0.55	-5.8	-0.4	-0.61	-0.58
Pefloxacin	-4.7	-0.5	-0.59	-5.8	-0.4	-0.62	-0.60
Chlorquinaldol	-5.0	0.0	-0.08	-5.3	-1.0	-1.17	-0.63
Sparfloxacin	-4.6	-0.7	-0.86	-5.9	-0.3	-0.43	-0.64
Gatifloxacin	-4.7	-0.5	-0.62	-5.7	-0.5	-0.68	-0.65
Quinagolide	-4.7	-0.5	-0.58	-5.6	-0.6	-0.78	-0.68
Elvitegravir	-4.8	-0.3	-0.46	-5.5	-0.7	-0.92	-0.69
Ciprofloxacin	-4.7	-0.6	-0.68	-5.7	-0.5	-0.70	-0.69
Tafenoquine	-4.5	-0.9	-1.02	-5.9	-0.3	-0.45	-0.74
Norfloxacin	-4.6	-0.8	-0.88	-5.7	-0.5	-0.67	-0.78
Carteolol	-4.6	-0.7	-0.79	-5.6	-0.6	-0.81	-0.80
Bosutinib	-4.6	-0.7	-0.77	-5.5	-0.8	-0.98	-0.87
Pergolide Mesylate	-4.3	-1.2	-1.29	-5.8	-0.4	-0.61	-0.95
Pergolide	-4.3	-1.3	-1.38	-5.8	-0.4	-0.59	-0.98
Oxamniquine	-4.6	-0.6	-0.73	-5.2	-1.1	-1.29	-1.01
Dequalinium	-4.4	-1.1	-1.25	-5.3	-1.0	-1.14	-1.20
Chloroxine	-4.7	-0.5	-0.64	-4.7	-1.7	-1.90	-1.27
Primaquine Phosphate	-4.5	-0.9	-0.98	-5.0	-1.4	-1.56	-1.27
Primaquine	-4.4	-1.0	-1.10	-5.0	-1.4	-1.57	-1.34
Cloquinoxol	-4.5	-0.9	-1.00	-4.8	-1.6	-1.83	-1.41
Dibucaine	-4.0	-1.7	-1.84	-5.3	-0.9	-1.11	-1.48
Chloroquine Phosphate	-4.2	-1.5	-1.64	-4.9	-1.5	-1.64	-1.64
Chloroquine	-4.1	-1.6	-1.73	-4.9	-1.4	-1.63	-1.68
Nelfinavir (Benchmark)	-5.0	0.1	0.00	-6.3	0.2	0.00	0.00
	Average	S.D.					
E6AP LXXLL	-5.0	0.53					
E7 LXCXE	-6.1	0.82					