HIGH SCHOOL EDITION
Journal of Student Research
Volume 10 Issue 4 (2021)

Identifying Promising Anticancer Sulforaphane Derivatives Using QSAR, Docking, and ADME Studies

Prachi Heda[1], Sahana Ravishankar[2], Aditi Shankar[3], Shrimayi Chaganti[4], Dishita Rajan[5], Riya Parekh[6] and Gayathri Renganathan[#]

[1]Dougherty Valley High School, San Ramon, CA, USA
[2]Amador Valley High School, Pleasanton, CA, USA
[3]Mission San Jose High School, Fremont, CA, USA
[4]Monta Vista High School, Cupertino, CA, USA
[5]Westmont High School, Campbell, CA, USA
[6]Irvington High School, Fremont, CA, USA
[#]Advisor

ABSTRACT

In cruciferous vegetables such as broccoli, sulfur-rich isothiocyanates, most notably sulforaphane (SFN), has shown anti-cancer properties, including cell cycle regulation, inducing apoptosis, and metastasis. Sulforaphane is a natural antioxidant that regulates several signal transduction pathways controlling oxidative stress, cellular defense, and cardiovascular disease. This paper attempts to elucidate the most drug-like SFN derivative through computational methods, including molecular docking, Quantitative Structure-Activity Relationship (QSAR), and absorption, distribution, metabolism, and excretion (ADME) analysis. From our studies, we can conclude that phenylethyl isothiocyanate (PEITC) has the most therapeutic potential out of a small set of 7 SFN derivatives. It is a confirmed lead-like compound by testing QSAR descriptors, notably the Dragon consensus drug-like score and lead-like score 2. PEITC also proves to be the most bioavailable derivative, as it is predicted to have high gastrointestinal absorption (GIA) and blood-brain barrier (BBB) permeability. In addition, it is in the optimal range for 5 out of 6 bioavailability properties proposed by the Abbot Bioavailability Score.[1] Lastly, from docking studies, PEITC had the highest average binding affinity overall, meaning that it holds a vital role in cancer prevention through molecular mechanisms.

Introduction

In the pursuit of finding novel cancer treatments, researchers have explored different methods such as radiation therapy, immunotherapy, hormone therapy, and the most widely researched: chemotherapy. The use of natural products in cancer therapy has increased over the past few decades as they are considered inexpensive, applicable, and accessible compared to conventional chemotherapy agents. Natural products and antioxidants are also known to be more environmentally friendly and have fewer toxic side effects.

Sulforaphane (SFN) is a sulfur-rich isothiocyanate found in cruciferous vegetables. It is a highly researched isothiocyanate derived from the hydrolysis of glucosinolate glucoraphanin by an enzyme called myrosinase. SFN is known for its high chemical reactivity caused by the electrophilicity of the central carbon of the isothiocyanate (—NCS) group. All glucosinolates have basic structures consisting of a β-D-thioglucose group, a sulfonated oxime group, and an amino acid-derived side chain.[2]

Multiple in-vitro and in-vivo studies have shown sulforaphane's chemopreventive activity against several types of cancer through the suppression of cellular proliferation, metastasis, angiogenesis, and other hallmarks of cancer.[3] Additionally, SFN has been found to possess antioxidant, antimicrobial, anti-aging, neuroprotective,

antidiabetic, and anti-inflammatory properties.[4] Its wide range of therapeutic properties has led it to become one of the most widely studied isothiocyanates.

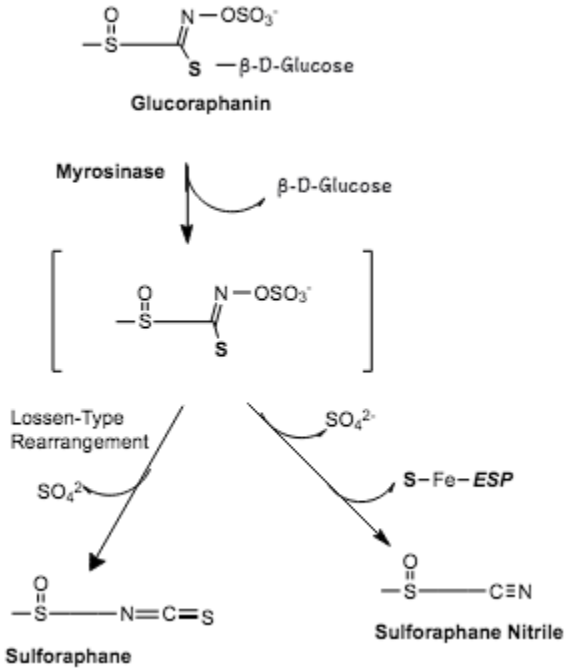


Figure 1. Conversion of Glucoraphanin to Sulforaphane. Image adapted from Matusheski et al.[5].

Despite extensive research concerning SFN and its properties, limited information is available regarding SFN's role in signaling pathways. It is widely accepted that SFN plays a role in KEAP1-NRF2 signaling.[6] The transcription factor NRF2 is a master regulator of cell survival responses to endogenous and exogenous stressors. This master regulator controls many essential pathways, such as cellular antioxidant defense. It induces NRF2/ARE/Prdx6 activity during aging, oxidative stress, and while protecting against cardiovascular disease.[7]

SFN also regulates several other signal transduction pathways. Many other prospective targets have been noted in literature, such as nuclear factor kappa B (NF-κB), phosphatidylinositol-3-kinase (PI3K), protein kinase B (Akt), and others have been noted.[8] Other targets include MMPs 2 and 9, DNMT1, KEAP1, and HDACS. MMPs 2 and 9 are involved in the invasion and metastasis of malignant cancer cells and are known to be regulated by MAPKs, including ERK1 and 2, P-38, and SAPK/JNK, suggesting that MAPKs can be a potential target for SFN.[9] Another potential target for SFN is the DNMT1 enzyme, which is often overexpressed in many cancers, resulting in abnormal patterns of DNA methylation.[10] The KEAP1 protein is a target of SFN as well. When KEAP1 is inhibited, it will prevent NRF2 degradation by the ubiquitin-proteasome system, increasing NRF2 expression levels. This results in the accumulation of NRF2 in the nucleus, where it induces the transcription of antioxidative and cytoprotective genes, leading to activation of the cell defense system.[11] Finally, SFN has been shown to inhibit HDACS 1-4 and 6, a common overactivated and expressed enzyme in cancers. The inhibition of HDACS can reactivate NRF2 expression, which prevents the growth of cancer cells.[12]

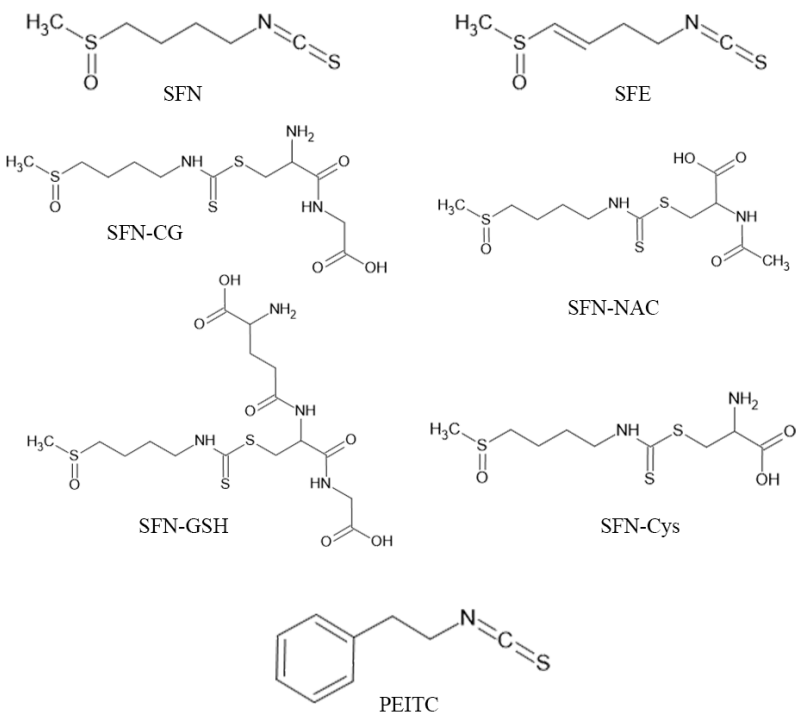


Figure 2. Structure of Sulforaphane and derivatives.

The increasing applicability of SFN has subsequently led to a correlating interest in the metabolites of SFN and related isothiocyanate phytochemicals. Metabolites of SFN include sulforaphane glutathione (SFN-GSH), sulforaphane cysteine-glycine (SFN-CG), sulforaphane cysteine (SFN-Cys), and sulforaphane N-acetylcysteine (SFN-NAC). Related compounds include sulforaphene (SFE) and phenylethyl isothiocyanate (PEITC).

Background

Molecular Docking

Molecular Docking is the in-silico study of how molecular structures interact with other molecular structures. This includes protein-protein interactions, protein-nucleic acid interactions, and most commonly, protein-small molecule interactions. Docking is a frequently used process in drug design screening, as it can predict the binding affinity of any compound with a protein readily available in databases, such as the RCSB protein data bank.[13] There are many web servers and software that can conduct molecular docking. A particularly efficient open-source software is AutoDockVina (ADV).[14] ADV has a user-friendly interface and can swiftly predict binding affinities. In this study, molecular docking by AutoDockVina was used to predict which proteins SFN and related compounds might interact with in-vitro. This information helps determine potential molecular pathways. Additionally, molecular docking was used to establish which small molecules, or ligands, had a higher chance of binding to a specified protein. From this, the biological activity of each compound was compared.

SwissADME

SwissADME offers various helpful computational and molecular tools that allow users to gain a deeper insight into drugs or molecules. ADME stands for Absorption, Distribution, Metabolism, and Excretion. Since our study targets sulforaphane for improving bioavailability, the most valuable tools on SwissADME were the chemical structure, bioavailability radar, and drug-likeness features. The Bioavailability Radar considers the lipophilicity, size, polarity, solubility, flexibility, and saturation of the drug to determine drug-likeness. Other descriptors like molecular weight (MW), molecular refractivity (MR), count of specific atom types, and polar surface area (PSA) can help determine if the drug can cross the blood-brain barrier to reach the brain and to evaluate other ADME properties. Furthermore, the drug-likeness feature allows us to determine if SFN could be an oral drug with respect to bioavailability. To determine if SFN can be an oral drug candidate, its drug-likeness would be determined from its structural and physicochemical inspections.[15]

Quantitative Structure Activity Relationship (QSAR)

QSAR computational modeling methods reveal relationships between chemical structure and biological activity. With QSAR models, mathematical equations can investigate properties that may be relevant to drug design. Additionally, these models screen large libraries of compounds for optimal candidates for ADME properties, as stated above. Because of these characteristics, QSAR modeling is essential in the pharmaceutical industry.[16] ADME and other properties calculated by QSAR software are called descriptors, and thousands have been reported. To create an adequate model, the correct descriptors must be chosen. In the software AlvaDesc, descriptors are grouped into blocks. These include constitutional indices, functional group counts, pharmacophore descriptors, and many more. In this study, drug-like indices were of particular interest. Two different types of scores can be calculated within the drug-like indices block: drug-likeness and lead-like scores. Druglikeness scores are a qualitative concept that evaluates if a substance is potentially bioavailable. Lead-like scores use filters to determine if a compound has pharmacological activity.[17]

Methods

Avogadro and Orca

First, each derivative was drawn in Avogadro V 1.2.0,[18] a free molecule editor and visualizer. Then the geometry of each structure was optimized using an MMFF94 force field with 10,000 steps, steepest descent, and 10-e7 convergence. After geometry optimization, Orca,[19] an organic chemistry program, was used to conduct Density Functional Theory (DFT) optimizations. An orca input was generated using a B3LYP functional set, and the orca command was implemented using a Mac terminal. The output PDB files were used for molecular docking.

AutoDockTools and AutoDockVina

AutoDockTools[20] was used to prepare the ligand and protein files. Ligand PDB files obtained from Avogadro were imported into AutoDockTools, where the ligand function was used to convert the file into a PDBQT format. To prepare each protein for docking, the PDB file was first obtained from RCSB Protein Data Bank[13]. After opening each protein in AutoDockTools, water molecules were removed while non-polar hydrogens and Gasteiger charges were added. Then each protein was converted to a PDBQT file using the macromolecule function. Each PDBQT file was opened using a text editor, and then charges were added to ions that required them. To complete docking by AutoDockVina,[21] a configuration file was made. Each configuration file consisted of a ligand file name, protein file name, exhaustiveness of 16, and grid box parameters that included the coordinates of the entire protein to conduct blind docking. To run AutoDockVina, the vina command line was entered into an Ubuntu terminal along with the configuration file for each docking run. The output was a PDBQT file that contained nine ligand conformations and a TXT file that contained the binding scores for each ligand conformation. One last additional step was taken here to ease the visualization process. A vina split command was run to split the output PDBQT file into nine separate files, each with a single conformer.

UCSF Chimera

UCSF Chimera[22] was used to view the protein-ligand complex and visualize potential hydrogen bonds between ligands and proteins.

Discovery Studio Visualizer

Discovery Studio Visualizer[23] (DSV) was also used for non-covalent visualization. A protein-ligand complex was imported into DSV, and the 2-D diagram feature was used to view interactions such as Pi-Alkyl, Amide Pi-Stacked, Pi Cation, Pi-Pi Stacked, Pi-Sulfur, Van Der Waals, and others.

alvaDesc

alvaDesc[24] was used to obtain QSAR data for each derivative. First, the free web server cheminfo[25] was used to generate a SMILES code for each derivative since they were not readily available on mainstream chemistry databases. Then an online SMILES translator provided by NIH[26] was used to convert each SMILES code into a Kekule MOL file. This process ensured the standardization of each compound. Next, each compound was loaded into alvaDesc, and basic drug-like indices were selected as the molecular descriptor of choice.

SwissADME

SwissADME[15] is an accessible web server that analyzes the absorption, distribution, metabolism, and excretion (ADME) of chemical compounds to assess their drug-likeness. The same list of SMILES that were created earlier was entered into SwissADME.

Results

QSAR

Two descriptors of interest were chosen: DLS_cons and LLS_02. DLS_cons (Dragon consensus drug-like score) are defined as the average of the seven drug-like scores (DLS_1 - DLS_7) provided by alvaDesc. These seven drug-like scores have been defined as the ratio between the number of satisfied rules and the total number of rules, represented by the equation LS=nRules/tRules, where *tRules* is the total number of rules and *nRules* is the number of satisfied rules. Drug-like indices evaluate qualitative properties to determine whether a compound is potentially bioavailable. Below is each DLS index.

Drug Like Scores

Score/ Rule	H-bond donors	H-bond acceptors	Molecular weight	Moriguchi's logP	Rotatable bond #	# of rings	Formal charge	Descriptor *C3p*	Descriptor *h-p*	Descriptor *Unsat-p*	Descriptor *NO_C3*	Polar surface area
DLS_1	≤ 5	≤ 10	≤ 500	≤ 4.15	-	-	-	-	-	-	-	-
DLS_2	≤ 5	1 - 8	200 - 450	-2.0 - 4.5	1 - 9	≤ 5	-	-	-	-	-	-
DLS_3	≤ 5	≤ 10	200 - 500	-5.0 - 5.0	≤ 8	-	-2 - 2	-	-	-	-	-
DLS_4	≤ 5	2 - 10	78 - 500	-0.5 - 5.0	-	-	-	0.15 - 0.8	0.6 - 1.6	0.10 - 0.45	-	-

DLS_5	-	-	-	-	-	-	-	-	-	≤ 0.4	0.10 - 1.80	-
DLS_6	≤ 5	≤ 10	≤ 500	≤ 5	≤ 10	-	-	-	-	-	-	≤ 140
DLS_7	-	-	-	-	≤ 10	-	-	-	-	-	-	≤ 140

Table 1. Drug-like scores. Molecular descriptor provided by alvaDesc.

Since DLS_cons is an average of the above indices, this index accurately portrays whether each sulforaphane derivative can become a drug.

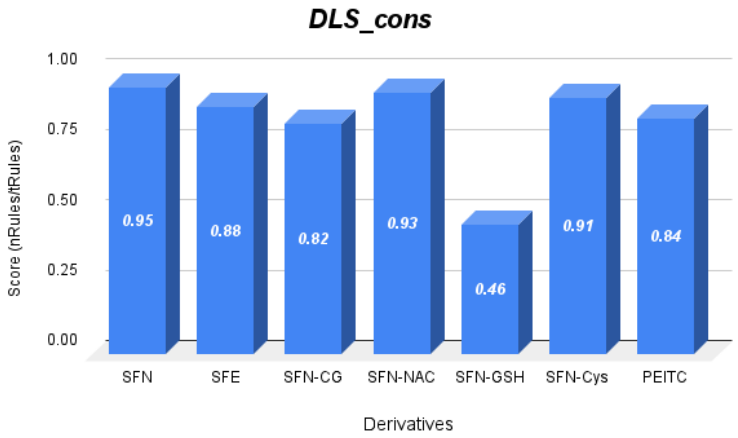


Figure 3. DLS_cons score for SFN and its derivatives and metabolites. A score of 1 indicates that a compound is a good candidate to be a drug, whereas a score of 0 indicates that a compound will likely not be a drug.

While drug-likeness evaluates bioavailability for compounds resembling existing drugs, the term "lead-like" is used for compounds possessing the structural and physicochemical profile of a quality lead. LLS indices are filters used to select those compounds qualified to be a lead in drug discovery. Typically lead compounds, compared to drugs, have lower molecular complexity, a smaller number of rings (nCIC), a smaller number of rotatable bonds (RBN), a lower MW and are more polar.[17] The index LLS_02 is a lead-like score derived from the rules proposed by Monge *et al*[27]. and is based on the following eight rules.

- H-bond donors ≤ 5
- H-bond acceptors in ≤ 9
- molecular weight MW ≤ 460
- Moriguchi's logP MLOGP in the range from –4.0 to 4.2
- rotatable bond number RBN ≤ 10
- number of rings nCIC ≤ 4
- number of halogens nX ≤ 7

- total number of nN + nO ≥ 1

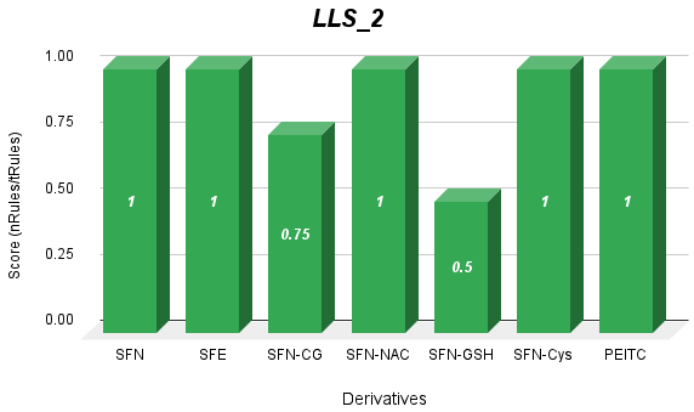


Figure 4. LLS_2 score for SFN and its derivatives and metabolites. A score of 1 indicates that a compound has lead-like properties, whereas a score of 0 indicates that a compound does not.

SwissADME]

SwissADME provides data on blood-brain-barrier (BBB) permeation and passive human gastrointestinal absorption (GIA), both of which are very important as these compounds must have sound absorption, have a low level of toxicity, and be orally bioavailable. Additionally, the website has other characteristics, such as Pgp substrate and bioavailability scores.

ADME Properties

Compound	GI absorption	BBB permeant	Pgp Substrate	Bioavailability Score
SFN	High	No	No	0.55
SFE	High	No	No	0.55
SFN-CG	Low	No	No	0.55
SFN-NAC	Low	No	No	0.11
SFN-GSH	Low	No	Yes	0.11
SFN-Cys	Low	No	No	0.55
PEITC	High	Yes	No	0.55

Table 2. Predicted ADME properties. Data retrieved from SwissADME

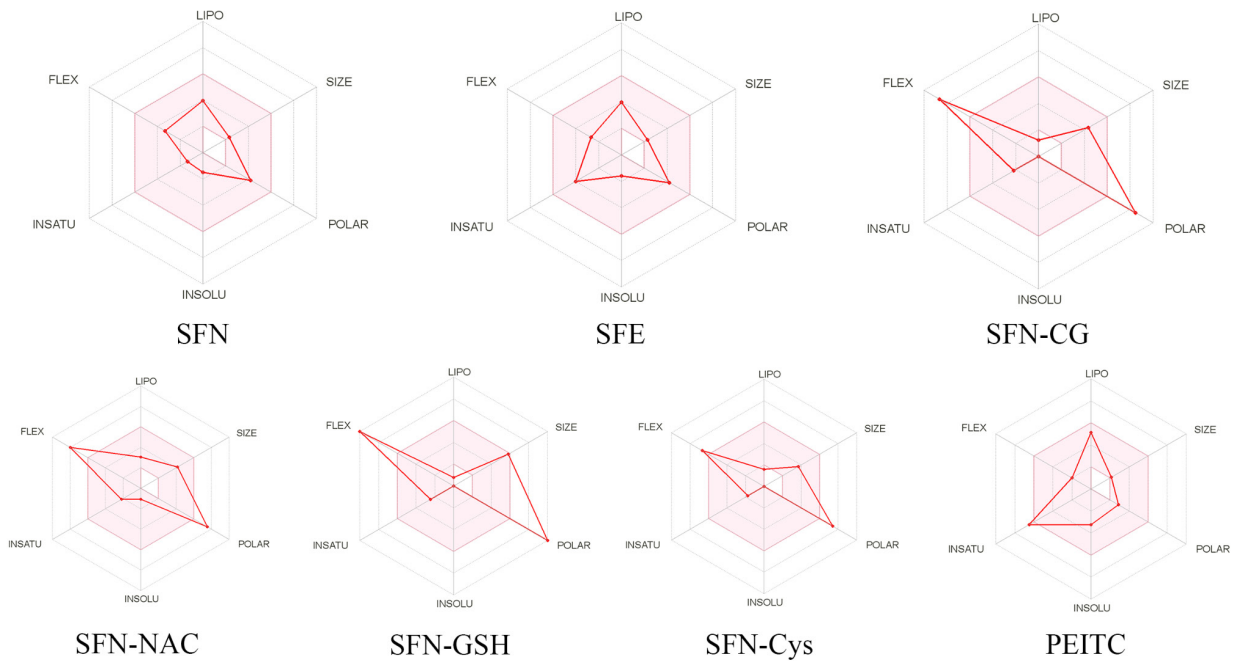


Figure 5. Bioavailability Radars of SFN, SFE, SFN-CG, SFN-NAC, SFN-GSH, SFN-Cys, and PEITC. Retrieved from SwissADME.

Docking

With a thorough investigation of literature, 13 protein targets were selected.

Protein Targets

PDB Code	Description	Class
3WNS[28]	Macrophage Migration Inhibitory Factor (MIF)	Isomerase Inhibitor
2J5E[29]	Epidermal growth factor receptor (EGFR)	Transferase
4LY1[30]	Histone deacetylase 2 (HDAC2)	Hydrolase Inhibitor
4CQ8[31]	Dihydroorotate dehydrogenase (DHODH)	Oxidoreductase
6DWN[32]	Human Cytochrome P450 1A1 (CYP1A1)	Oxidoreductase
5OMG[33]	p38alpha	Transferase
6RQ4[34]	Mitogen-activated kinase 1 (ERK2)	Signaling Protein
6PYE[35]	Histone deacetylase 6 (HDAC6)	Hydrolase
4WXX[36]	DNA Methyltransferase 1 (DNMT1)	Transferase
4QHE[37]	Human apurinic/apyrimidinic endonuclease 1 (APE1)	Lyase

1JL0[38]	Human S-Adenosylmethionine Decarboxylase (AdoMetDC)	Lyase
5FNU[39]	Kelch-like ECH associated protein 1 (KEAP1)	Transcription
1WE1[40]	Heme oxygenase-1 (HO-1)	Oxidoreductase

Table 3. Proteins selected for docking, with PDB code and class.

Each SFN derivative was docked to each protein, and the docking score in kcal/mole was recorded. Note, the lower the docking score, the higher the binding affinity.

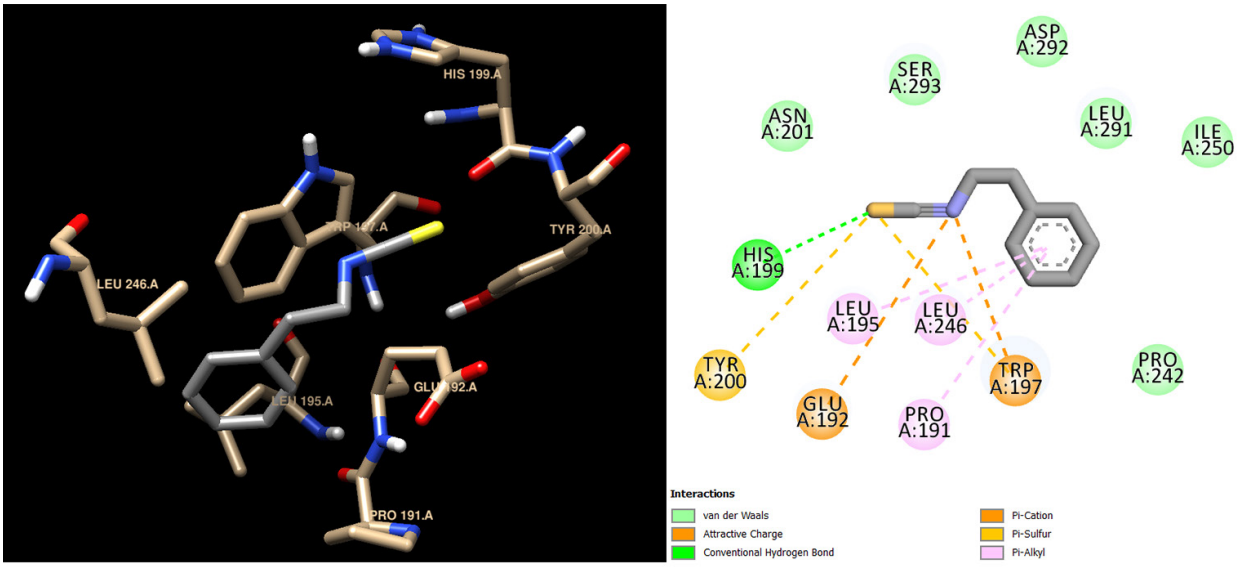
Docking score (kcal/mol)

Derivative/Protein	3WNS	2J5E	4LY1	4WXX	5OMG	6RQ4	6PYE	4CQ8	4QHE	1JL0	6DWN	5FNU	1WE1	Average Derivative score
SFN	-4.7	-3.5	-3.7	-4.2	-4.1	-3.8	-3.8	-4.5	-4.0	-4.7	-3.8	-4.4	-4.4	-4.12
SFE	-5.0	-3.7	-3.6	-4.3	-4.8	-4.0	-4.3	-4.7	-4.3	-5.0	-4.8	-4.5	-4.7	-4.44
SFN-CG	-6.5	-4.4	-4.1	-5.2	-4.7	-4.6	-4.8	-5.5	-5.5	-6.0	-6.1	-5.9	-5.3	-5.28
SFN-NAC	-6.4	-4.3	-3.7	-4.9	-5.2	-4.9	-4.2	-4.3	-5.3	-5.9	-5.9	-6.0	-4.8	-5.06
SFN-GSH	-8.0	-4.9	-4.1	-5.6	-4.9	-6.0	-5.4	-4.9	-5.3	-6.5	-4.7	-6.0	-5.5	-5.52
SFN-Cys	-5.9	-4.3	-3.7	-4.7	-4.7	-4.4	-4.3	-3.9	-4.7	-5.5	-3.8	-5.5	-5.2	-4.66
PEITC	-6.1	-5.1	-4.8	-5.6	-6.6	-5.0	-6.5	-6.1	-5.4	-6.9	-6.8	-5.2	-6.6	-5.9
Average protein score	-6.09	-4.31	-3.96	-4.93	-5.0	-4.67	-4.76	-4.84	-4.93	-5.79	-5.13	-5.36	-5.21	

> -4.0	-4.0 ≤ score < -5.0	-5.0 ≤ score < -6.0	-6.0 ≤ score ≤ -6.5	> -6.5

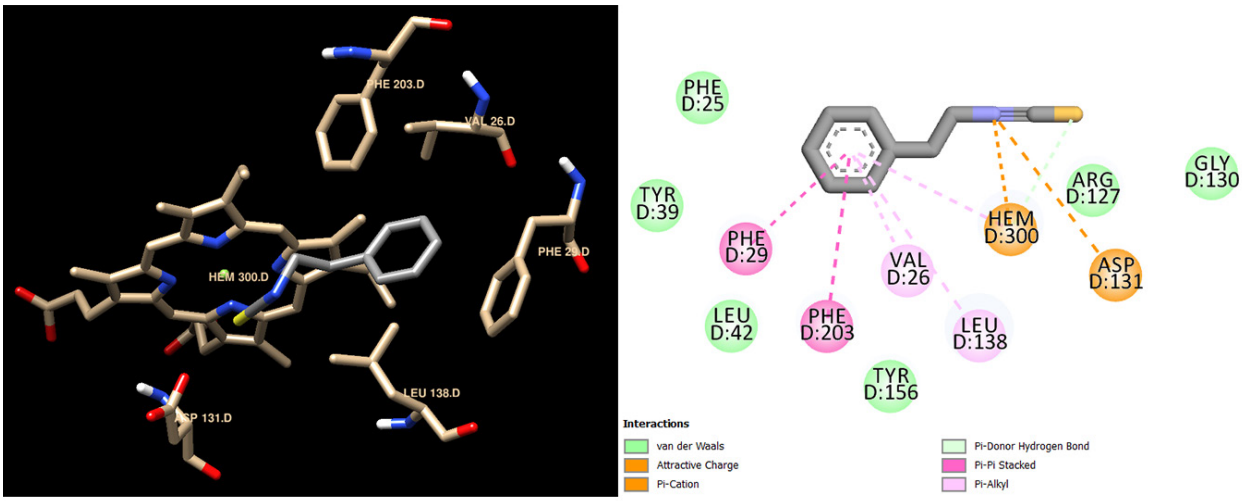
Table 4. Docking scores. Averages of each protein and derivative are shown.

Interactions between the protein and ligand with the lowest docking scores were visualized.



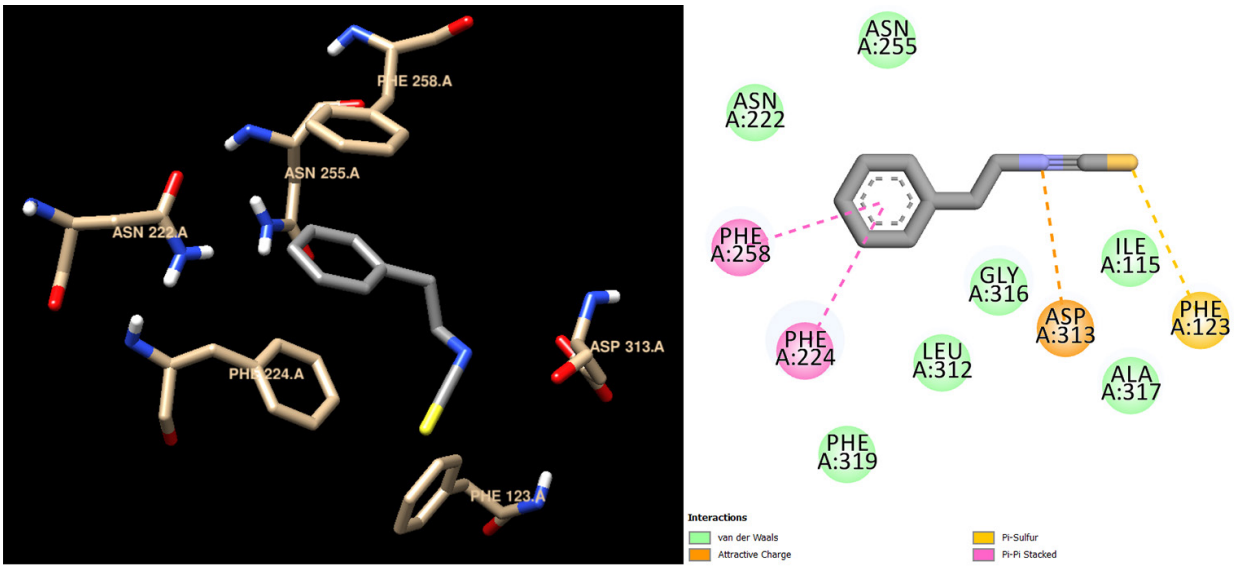
a. b.

Figure 6. a. 3D and b. 2D view of PEITC docked to p38alpha (PDB 5OMG).



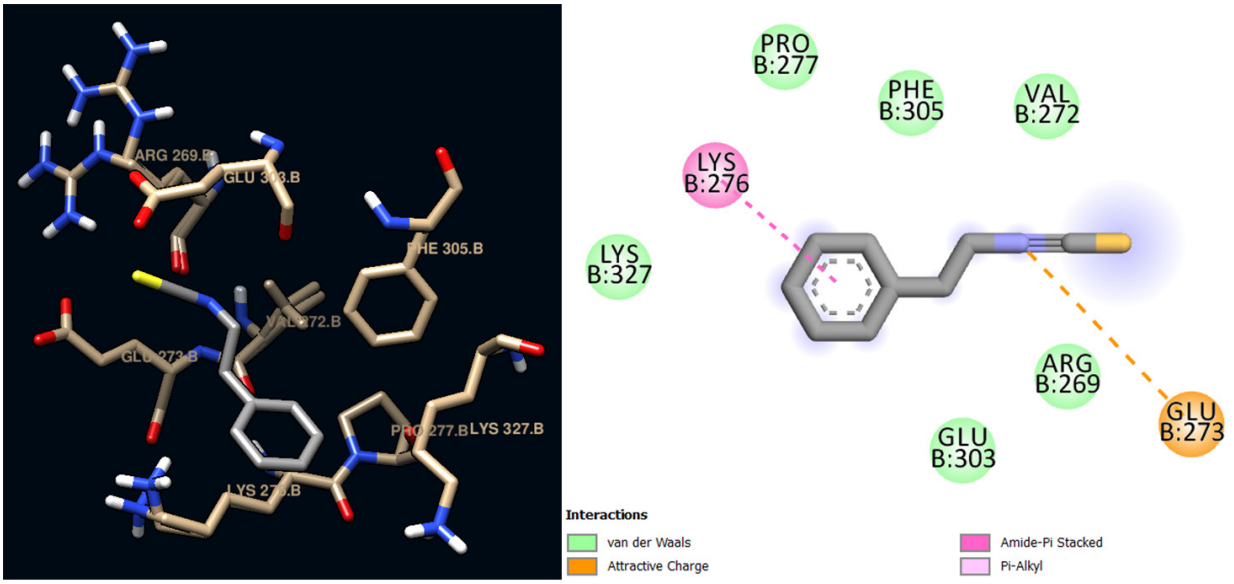
a. b.

Figure 7. a. 3D and b. 2D view of PEITC docked to Heme Oxygenase-1 (PDB 1WE1).



a. b.

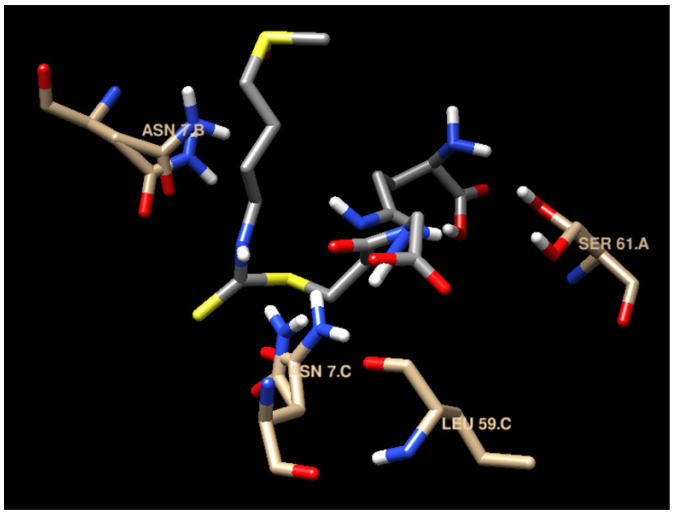
Figure 8. a. 3D and b. 2D view of PEITC docked to Human Cytochrome P450 1A1 (PDB 6DWN).



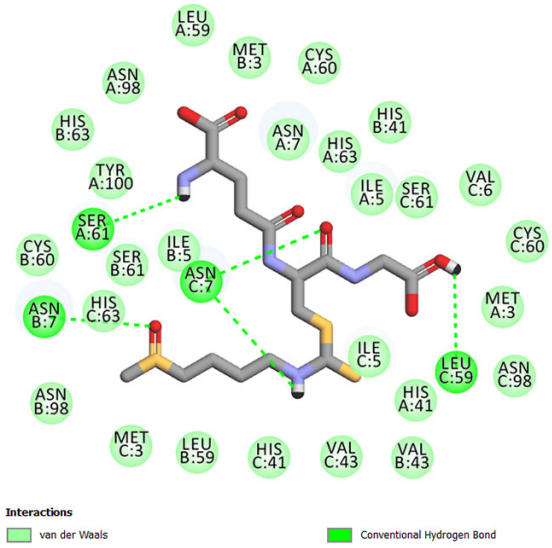
a. b.

Figure 9. a. 3D and b. 2D view of PEITC docked to Human S-Adenosylmethionine Decarboxylase (PDB 1JL0).

Figure 10. a. 3D and b. 2D view of SFN-GSH docked to Macrophage Migration Inhibitory Factor (PDB 3WNS).



a.



b.

Discussion

QSAR

All derivatives had a DLS_con score of above 0.4, and six out of seven derivatives had a favorable DLS_con score of above 0.8. The highest score achieved was 0.95 by SFN itself, reinforcing the promise of SFN as a pharmaceutical. SFN-NAC had a DLS_con score of 0.93, with the only violations being the total number of rotatable bonds and polar surface area (TPSA) parameters. Next in line was SFN-Cys with a score of 0.91, then SFE with a surprisingly high score of 0.88. After SFE was PEITC with a score of 0.84. This score is lower than expected, as PEITC is the only derivative in the set that contains an aromatic ring. With aromatic rings being abundant in approved pharmaceuticals due to their rigidity and functionality, our team expected PEITC to have a higher drug-like score. However, in the DLS_cons parameters, the only specification regarding rings was that there must be less than or equal to four rings in a given compound. Therefore, aromatic rings do not play a significant role in the DLS_cons score, which explains why PEITC had a lower score than derivatives without aromatic rings. After PEITC, SFN-CG had a score of 0.82, and SFN-GSH had a score of 0.46. It was expected that SFN-GSH would have the lowest score because of its large structure and thus would violate MW, TPSA, RBN, and more parameters.

For LLS_2, five out of seven derivatives had a perfect score of 1. These derivatives include PEITC, SFY-Cys, SFN-NAC, SFE, and SFN. These results correlate with the results of DLS_cons, as these five derivatives had a DLS_con score greater than or equal to 0.84. SFN-CG had an LLS_2 score of 0.75, as it violated RBN and nN+nO greater than or equal to 1 parameter. Lastly, SFN-GSH had a poor LLS_2 score of 0.5. SFN-GSH violated the H-bond donors, molecular weight, MLOGP, and RBN parameters. Once again, this is not surprising due to the large structure of SFN-GSH.

SwissADME

According to Table 1, out of the seven derivatives, SFN, SFE, and PEITC had high GI absorption. Of these three, only PEITC had BBB permeability. Additionally, information regarding each derivative acting as a substrate of the Pgp (P-glycoprotein) implies whether the derivative has sufficient drug absorption and distribution. As an efflux transporter, Pgp limits the bioavailability of orally administered drugs by pumping them back into the lumen.

Therefore, a substrate of Pgp can result in increased bioavailability of the susceptible drug.[41] SFN-GSH is the only Pgp substrate, so this indicates that SFN-GSH has the potential to bind to Pgp. However, it is also shown to have low GI absorption and being BBB impermeable. This shows that even potential Pgp substrates are not necessarily bioavailable. Lastly, in the table is the Bioavailability Score, also known as the Abbot Bioavailability Score, proposed by Martin YC, which predicts a compound's probability of having at least 10% bioavailability in a rat or measurable Caco-2 permeability.[42] SFN, SFE, PEITC, SFN-Cys, and SFN-CG have a bioavailability score of 0.55. SFN-NAC and SFN-GSH both have a bioavailability score of 0.11.

The bioavailability radar of each compound was evaluated using the SwissADME tool. The bioavailability radar enables a first glance at the drug-likeness of a molecule. The pink area in each figure represents the ideal range of each property. Six properties were taken into account: lipophilicity, size, polarity, solubility, flexibility, and saturation. The optimal range for lipophilicity would be between -0.7 and +0.5. Lipophilicity is a critical physicochemical parameter that contributes to the absorption, distribution, metabolism, excretion, and toxicity of a drug.[43] The optimal size or molecular weight range would be between 150 and 500 g/mol. The optimal polarity is between 20 and 130 Å 2. The best solubility should not be higher than six as having a soluble molecule greatly aids many drug development activities, mainly the ease of handling and formation. For saturation, the fraction of carbons in the sp3 hybridization should not be less than 0.25. Finally, the flexibility should be no more than nine rotatable bonds.[15] According to Figure 3, SFN, SFE, and PEITC all relatively stay in the optimal ranges for each property, indicating that they have the highest drug-likeness. The other four compounds were not predicted to be orally bioavailable because they exceeded the ranges for flexibility and polarity.

Docking

After docking studies, five protein-ligand complexes with the highest binding affinity were further analyzed. While the docking scores of all protein-ligand complexes can be found in Table 3, the ones selected include PEITC docked to Human S-Adenosylmethionine Decarboxylase (PDB 1JL0), p38alpha (PDB 5OMG), Heme Oxygenase-1 (PDB 1WE1), Human Cytochrome P450 1A1 (PDB 6DWN), and SFN-GSH docked to Macrophage Migration Inhibitory Factor (PDB 3WNS). These complexes were chosen for their highest binding affinity and because the average derivative/protein docking score was the highest in their respective groups. The average derivative docking score for PEITC was -5.9 kcal/mol, and the average docking score for the Macrophage Migration Inhibitory Factor was -6.09 kcal/mol.

PEITC docked to p38alpha had a docking score of -6.6 kcal/mol. As shown in Figure 6, the active site residues include ASN201, SER293, ASP292, LEU291, ILE250, PRO242, as PEITC has van der Waals interactions with them. Other interactions include a hydrogen bond between the sulfur present in PEITC and the positively charged HIS199. A hydrogen bond is a chemical bond between a strongly electronegative atom such as oxygen, fluorine, or nitrogen and a partially positive hydrogen molecule. These bonds have an intermolecular attraction stronger than van der Waals forces but weaker than ionic and covalent bonds. The same sulfur forms pi-sulfur interactions with the aromatic ring in TYR200. Pi-Sulfur interactions are non-covalent interactions in which the lone pair of valence electrons in the electron cloud of sulfur interacts with the electron cloud of the aromatic ring of the aromatic amino acid, tyrosine. The nitrogen in PEITC interacts with GLU192 and TRP197 in the form of pi-cation interactions. Pi-Cation interactions are strong non-covalent interactions between the negatively charged pi face of an aromatic ring and positively charged cations.[44] These interactions contribute to protein stability, molecular recognition, and drug-receptor interactions. Lastly, PEITC's aromatic rings form pi-alkyl interactions with LEU195, LEU246, and PRO191. Pi-Alkyl interactions are a form of non-covalent interactions that are vital in protein-ligand recognition. The interaction occurs between the electron group of an element from the alkyl group and the aromatic rings of a protein.[45] Although p38alpha is a tumor suppressor, it can also function as a tumor promoter and can acquire cancerous roles in cell metabolism and angiogenesis and has been linked to low p38alpha activations.[46] PEITC can inhibit the p38alpha activations in cancers, thereby reducing angiogenesis and other tumor-promoting mechanisms. Since PEITC has a high docking score to p38alpha, it could play a role in inhibiting p38alpha pathways.

PEITC docked to Heme Oxygenase-1 (HO-1) had a docking score of -6.6 kcal/mol. As shown in Figure 7, Van der Waals interactions occur with residues PHE25, TYR39, LEU42, TYR156, ARG127, and GLY130. The aromatic ring present in PEITC formed pi-pi stacking interactions with PHE29 and PHE203 and pi-alkyl interactions with VAL26, LEU138, and the nearby heme molecule. Pi-pi stacking interactions are non-covalent attractive forces between aromatic rings that have pi bonds which are covalent chemical bonds. Pi-pi stacking contributes to drug delivery and molecular recognition.[47] Lastly, the electronegative sulfur on PEITC formed a hydrogen bond with the heme molecule. HO-1 activation is known to be susceptible to proteolytic cleavage, and in addition, the nuclear localization of HO-1 facilitates tumor growth and proliferation. Inhibiting the activation of HO-1 can reduce tumor formation and its proliferation.[48] PEITC could play a role in inhibiting HO-1 activation since PEITC has a high binding affinity to HO-1.

PEITC docked to Human Cytochrome P450 1A1 had a docking score of -6.8 kcal/mol. As shown in Figure 8, PETIC formed van der Waals interactions with ASN222, ASN225, PHE319, LEU312, GLY316, ALA317, and ILE115. Additionally, the aromatic ring formed pi-pi stacking interactions with PHE258 and PHE224. The nitrogen atom on PEITC had attractive charges with the negatively charged ASP313. Lastly, the sulfur atom had pi-sulfur interactions with PHE123. Cytochromes P450 1A1 is an extrahepatic monooxygenase involved in the metabolism of endogenous substrates and drugs.[49] It has been implicated in the bioactivation of carcinogens. Therefore, compounds that could regulate the activity of CYPs are essential for the prevention of chemical-induced carcinogenesis.[50] Since PEITC had a high binding affinity to CYP 1A1, it could inhibit the enzyme and prevent tumor initiation.

PEITC docked to Human S-Adenosylmethionine Decarboxylase (AdoMetDC) had a docking score of -6.9 kcal/mol. As shown in Figure 9, the active site residues include LYS327. PRO277, PHE305, VAL272, GLU303, and ARG269, with which PEITC has van der Waals interactions. Van der Waals interactions occur between atoms in close proximity that generate electrical interactions. Individually, these forces are the weakest compared to the strength (kJ/mol) of hydrogen bonds and ionic interactions. However, a collective of Van der Waals forces between atoms or molecules has considerably strong interactions. Additionally, an Amide-Pi stacked interaction is formed between the surface of the amide bond on LYS276 and the surface of the aromatic ring present in PEITC. The last interaction shown in Figure 4.b is an attractive charge between the nitrogen in PEITC and the negatively charged GLU273. Attractive charges result from the force generated between two particles of opposite charges, as seen by the bond between PEITC and GLU273. AdoMetDC is an enzyme in the production of polyamines, which are ubiquitous organic cations needed for normal cell differentiation and proliferation. Tumor tissues often have elevated levels of polyamines, therefore increasing cellular proliferation. Inhibiting AdoMetDC in cancers with high expression of polyamines can reduce irregular cell proliferation.[51] Since PETIC has a high docking score to AdoMetDC, it may play a role in the inhibition of AdoMetDC.

Lastly, SFN-GSH docked to Macrophage Migration Inhibitory Factor (MIF) had the highest binding affinity with a docking score of -8.0 kcal/mol. Seen in Figure 10, the most prevalent interaction in this protein-ligand complex was van der Waals interactions. While van der Waals forces are considered weak, in large numbers they can become much stronger. The number of van der Waals interactions can contribute to the binding affinity of this complex, which was the highest out of the 91 docking runs completed. The only other interactions present were hydrogen bonds with residues SER61, ASN7 in chain B, ASN7 in chain C, and LEU59. MIF is a pleiotropic, proinflammatory cytokine that plays a great role in the initiation of immune and inflammatory response. However, it also has a unique tautomerase enzymatic activity, causing an elevation of MIF. Elevated levels of MIF promote proliferation, angiogenesis, and metastasis of almost all cancer cells. The inhibition of the tautomerase activity interferes with the interaction of MIF to other proteins, which thereby inhibits cancer properties.[28]

Conclusion

In this research, QSAR, ADME, and docking studies were conducted to determine the most promising anticancer therapeutic in a small set of SFN derivatives and metabolites. Consistent with our hypothesis, PEITC dominated as the most promising small molecule therapy overall.

QSAR studies conducted with the alvaDesc software determined the drug-likeness and lead likeness of each compound. PEITC had a DLS_cons score of 0.84, which was the fifth-highest score. Even though this score is lower than other derivatives scores, 0.84 is an adequate score while evaluating 12 overall parameters within all the drug-like descriptors utilized in the dragoon consensus score. Additionally, PEITC had a perfect LLS_2 score of 1, confirming that it is a lead-like compound.

ADME studies with SwissADME predicted properties regarding bioavailability. PEITC was in the optimal range of 5 out of 6 bioavailability properties, as shown by Figure 3. While both SFN and SFE were in the optimal ranges for all bioavailability properties, it is crucial to consider each compound's permeation characteristics. PEITC has high GI absorption and is the only compound that is blood-brain barrier permeable. This characteristic is vital since therapeutics must be easily absorbed by the body to have any predicted effect.

Lastly, docking studies were utilized to predict the binding affinity of each compound to a small set of proteins. Overall, PEITC had the highest average binding affinity, meaning that it had a high binding affinity to multiple proteins in the set. These notable proteins include AdoMetDC, p38alpha, HO-1, and Human Cytochrome P450 1A1. The binding affinity of PEITC to these proteins indicates that PETIC has a role in cancer prevention through reducing proliferation, carcinogenesis, and angiogenesis.

Acknowledgments

We gratefully acknowledge the Olive Children's Foundation and its community of corporate sponsors and supporters for funding our research. Finally, we thank the Aspiring Scholars Directed Research Program for providing us with the knowledge and direction to conduct research. Author contributions are as follows. GR conceptualized this project, PH conducted and obtained experimental results, and subsequently wrote the methods, results, discussion, and conclusion sections. SR, SC, DR, and AS worked on the discussion section. RP, SC, AS, and PH collaborated on the introduction.

References

1. Martin Y. C. (2005). A bioavailability score. *Journal of medicinal chemistry*, *48*(9), 3164–3170. https://doi.org/10.1021/jm0492002
2. *Sulforaphane Molecule-- Anti-Aging and role in Cancer*. Sulforaphane Molecule -- Anti-Aging and Senolytic Properties. (n.d.). https://www.worldofmolecules.com/anti-aging-and-senolytics/sulforaphane-molecule.html.
3. Fofaria, N. M., Ranjan, A., Kim, S. H., & Srivastava, S. K. (2015). Mechanisms of the Anticancer Effects of Isothiocyanates. *The Enzymes*, *37*, 111–137. https://doi.org/10.1016/bs.enz.2015.06.001
4. Kim, J. K., & Park, S. U. (2016). Current potential health benefits of sulforaphane. *EXCLI journal*, *15*, 571–577. https://doi.org/10.17179/excli2016-485
5. Matusheski, N. V., Juvik, J. A., & Jeffery, E. H. (2004). Heating decreases epithiospecifier protein activity and increases sulforaphane formation in broccoli. *Phytochemistry*, *65*(9), 1273–1281. https://doi.org/10.1016/j.phytochem.2004.04.013
6. Matusheski, N. V., Juvik, J. A., & Jeffery, E. H. (2004). Heating decreases epithiospecifier protein activity and increases sulforaphane formation in broccoli. *Phytochemistry*, *65*(9), 1273–1281. https://doi.org/10.1016/j.phytochem.2004.04.013

7. Kensler, T. W., Egner, P. A., Agyeman, A. S., Visvanathan, K., Groopman, J. D., Chen, J. G., Chen, T. Y., Fahey, J. W., & Talalay, P. (2013). Keap1-nrf2 signaling: a target for cancer prevention by sulforaphane. *Topics in current chemistry*, *329*, 163–177. https://doi.org/10.1007/128_2012_339
8. Kubo, E., Chhunchha, B., Singh, P., Sasaki, H., & Singh, D. P. (2017). Sulforaphane reactivates cellular antioxidant defense by inducing Nrf2/ARE/Prdx6 activity during aging and oxidative stress. *Scientific reports*, *7*(1), 14130. https://doi.org/10.1038/s41598-017-14520-8
9. Jabbarzadeh Kaboli, Parham & Afzalipoor, Masoomeh & Mohammadi, Mahsa & Abiri, Ardavan & Mokhtarian, Roya & Vazifehmand, Reza & Amanollahi, Shima & Sani, Shaghayegh & Li, Mingxing & Zhao, Yueshui & Wu, Xu & Shen, Jing & Cho, Chi & Xiao, Zhangang. (2020). Targets and mechanisms of sulforaphane derivatives obtained from cruciferous plants with special focus on breast cancer – contradictory effects and future perspectives. Biomedicine & Pharmacotherapy. 121. 109635. 10.1016/j.biopha.2019.109635.
10. Keum, Y. S., Yu, S., Chang, P. P., Yuan, X., Kim, J. H., Xu, C., Han, J., Agarwal, A., & Kong, A. N. (2006). Mechanism of action of sulforaphane: inhibition of p38 mitogen-activated protein kinase isoforms contributing to the induction of antioxidant response element-mediated heme oxygenase-1 in human hepatoma HepG2 cells. *Cancer research*, *66*(17), 8804–8813. https://doi.org/10.1158/0008-5472.CAN-05-3513
11. Kaufman-Szymczyk, A., Majewski, G., Lubecka-Pietruszewska, K., & Fabianowska-Majewska, K. (2015). The Role of Sulforaphane in Epigenetic Mechanisms, Including Interdependence between Histone Modification and DNA Methylation. *International journal of molecular sciences*, *16*(12), 29732–29743. https://doi.org/10.3390/ijms161226195
12. Wang, H., Khor, T. O., Yang, Q., Huang, Y., Wu, T. Y., Saw, C. L., Lin, W., Androulakis, I. P., & Kong, A. N. (2012). Pharmacokinetics and pharmacodynamics of phase II drug metabolizing/antioxidant enzymes gene response by anticancer agent sulforaphane in rat lymphocytes. *Molecular pharmaceutics*, *9*(10), 2819–2827. https://doi.org/10.1021/mp300130k
13. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank *Nucleic Acids Research*, 28: 235-242.
14. Santos-Martins, D., Eberhardt, J., Tillack, A.F., Forli, S. (2020). AutoDock Vina 1.2.0. Improved software with new docking methods, expanded force field, and Python bindings.
15. Daina, Antoine & Michielin, Olivier & Zoete, Vincent. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports. 7. 42717. 10.1038/srep42717.
16. Kwon, S., Bae, H., Jo, J., & Yoon, S, (2019). Comprehensive ensemble in QSAR prediction for drug discovery. *BMC bioinformatics, 20*(1), 521. https://doi.org/10.1186/s12859-019-3135-4
17. AlvaDesc molecular descriptors (n.d.). https://www.alvascience.com/alvadesc-descriptors
18. Hanwell, M.D., Curtis, D.E., Lonie, D.C. et al. (2012, August 13) *Avogadro: an advanced semantic chemical editor, visualization, and analysis platform.* J Cheminform 4, 17 (2012). https://doi.org/10.1186/1758-2946-4-17.
19. Neese, Frank, (2017, July 17) *Software update: the ORCA program system, version 4.0.* Wiley Interdiscip. Rev.: Comput. Mol. Sci., 8, e1327, https://doi.org/10.1002/wcms.1327
20. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, *30*(16), 2785–2791. https://doi.org/10.1002/jcc.21256
21. Trott, O., & Olson, A. J. (2010). 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

22. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084
23. Dassault Systèmes BIOVIA, (2017) *Discovery Studio Modeling Environment*, San Diego: Dassault Systèmes, 2016. https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-discovery-studio/visualization/
24. Mauri, A. (2020). *alvaDesc: A tool to calculate and analyze molecular descriptors and fingerprints.* In K. Roy (Ed.), Ecotoxicological QSARs (pp. 801–820). Humana Press Inc. https://lnkd.in/dHmPQS5
25. Create a molfile. (n.d.). *Cheminfo*. http://www.cheminfo.org/Chemistry/Generate_molfiles/index.html.
26. U.S. Department of Health and Human Services. (n.d.). *Online SMILES Translator*. National Institutes of Health. https://cactus.nci.nih.gov/translate/.
27. Monge, A., Arrault, A., Marot, C., & Morin-Allory, L. (2006). Managing, profiling and analyzing a library of 2.6 million compounds gathered from 32 chemical providers. *Molecular diversity*, *10*(3), 389–403. https://doi.org/10.1007/s11030-006-9033-5
28. Spencer, E. S., Dale, E. J., Gommans, A. L., Rutledge, M. T., Vo, C. T., Nakatani, Y., Gamble, A. B., Smith, R. A., Wilbanks, S. M., Hampton, M. B., & Tyndall, J. D. (2015). Multiple binding modes of isothiocyanates that inhibit macrophage migration inhibitory factor. *European journal of medicinal chemistry*, *93*, 501–510. https://doi.org/10.1016/j.ejmech.2015.02.012
29. Blair, J. A., Rauh, D., Kung, C., Yun, C. H., Fan, Q. W., Rode, H., Zhang, C., Eck, M. J., Weiss, W. A., & Shokat, K. M. (2007). Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nature chemical biology*, *3*(4), 229–238. https://doi.org/10.1038/nchembio866
30. Lauffer, B. E., Mintzer, R., Fong, R., Mukund, S., Tam, C., Zilberleyb, I., Flicke, B., Ritscher, A., Fedorowicz, G., Vallero, R., Ortwine, D. F., Gunzner, J., Modrusan, Z., Neumann, L., Koth, C. M., Lupardus, P. J., Kaminker, J. S., Heise, C. E., & Steiner, P. (2013). Histone deacetylase (HDAC) inhibitor kinetic rate constants correlate with cellular histone acetylation but not transcription and cell viability. *The Journal of biological chemistry*, *288*(37), 26926–26943. https://doi.org/10.1074/jbc.M113.490706
31. Ross, L. S., Gamo, F. J., Lafuente-Monasterio, M. J., Singh, O. M., Rowland, P., Wiegand, R. C., & Wirth, D. F. (2014). In vitro resistance selections for Plasmodium falciparum dihydroorotate dehydrogenase inhibitors give mutants with multiple point mutations in the drug-binding site and altered growth. *The Journal of biological chemistry*, *289*(26), 17980–17995. https://doi.org/10.1074/jbc.M114.558353
32. Bart, A. G., & Scott, E. E. (2018). Structures of human cytochrome P450 1A1 with bergamottin and erlotinib reveal active-site modifications for binding of diverse ligands. *The Journal of biological chemistry*, *293*(50), 19201–19210. https://doi.org/10.1074/jbc.RA118.005588
33. Bartolini, D., Bührmann, M., Barreca, M. L., Manfroni, G., Cecchetti, V., Rauh, D., & Galli, F. (2019). Co-crystal structure determination and cellular evaluation of 1,4-dihydropyrazolo[4,3-c] [1,2] benzothiazine 5,5-dioxide p38α MAPK inhibitors. *Biochemical and biophysical research communications*, *511*(3), 579–586. https://doi.org/10.1016/j.bbrc.2019.02.063
34. Kidger, A. M., Munck, J. M., Saini, H. K., Balmanno, K., Minihane, E., Courtin, A., Graham, B., O'Reilly, M., Odle, R., & Cook, S. J. (2020). Dual-Mechanism ERK1/2 Inhibitors Exploit a Distinct Binding Mode to Block Phosphorylation and Nuclear Accumulation of ERK1/2. *Molecular cancer therapeutics*, *19*(2), 525–539. https://doi.org/10.1158/1535-7163.MCT-19-0505
35. Reßing, N., Sönnichsen, M., Osko, J. D., Schöler, A., Schliehe-Diecks, J., Skerhut, A., Borkhardt, A., Hauer, J., Kassack, M. U., Christianson, D. W., Bhatia, S., & Hansen, F. K. (2020). Multicomponent Synthesis, Binding Mode, and Structure-Activity Relationship of Selective Histone Deacetylase 6 (HDAC6) Inhibitors with Bifurcated Capping Groups. *Journal of medicinal chemistry*, *63*(18), 10339–10351. https://doi.org/10.1021/acs.jmedchem.9b01888

36. Zhang, Z. M., Liu, S., Lin, K., Luo, Y., Perry, J. J., Wang, Y., & Song, J. (2015). Crystal Structure of Human DNA Methyltransferase 1. *Journal of molecular biology*, *427*(15), 2520–2531. https://doi.org/10.1016/j.jmb.2015.06.001
37. He, H., Chen, Q., & Georgiadis, M. M. (2014). High-resolution crystal structures reveal plasticity in the metal binding site of apurinic/apyrimidinic endonuclease I. *Biochemistry*, *53*(41), 6520–6529. https://doi.org/10.1021/bi500676p
38. Ekstrom, J. L., Tolbert, W. D., Xiong, H., Pegg, A. E., & Ealick, S. E. (2001). Structure of a human S-adenosylmethionine decarboxylase self-processing ester intermediate and mechanism of putrescine stimulation of processing as revealed by the H243A mutant. *Biochemistry*, *40*(32), 9495–9504. https://doi.org/10.1021/bi010736o
39. Davies, T. G., Wixted, W. E., Coyle, J. E., Griffiths-Jones, C., Hearn, K., McMenamin, R., Norton, D., Rich, S. J., Richardson, C., Saxty, G., Willems, H. M., Woolford, A. J., Cottom, J. E., Kou, J. P., Yonchuk, J. G., Feldser, H. G., Sanchez, Y., Foley, J. P., Bolognese, B. J., Logan, G., … Kerns, J. K. (2016). Monoacidic Inhibitors of the Kelch-like ECH-Associated Protein 1: Nuclear Factor Erythroid 2-Related Factor 2 (KEAP1:NRF2) Protein-Protein Interaction with High Cell Potency Identified by Fragment-Based Discovery. *Journal of medicinal chemistry*, *59*(8), 3991–4006. https://doi.org/10.1021/acs.jmedchem.6b00228
40. Sugishima, M., Migita, C. T., Zhang, X., Yoshida, T., & Fukuyama, K. (2004). Crystal structure of heme oxygenase-1 from cyanobacterium Synechocystis sp. PCC 6803 in complex with heme. *European journal of biochemistry*, *271*(22), 4517–4525. https://doi.org/10.1111/j.1432-1033.2004.04411.x
41. Finch, A., & Pillans, P. (2014). P-glycoprotein and its role in drug-drug interactions. *Australian Prescriber*, *37*(4), 137–139. https://doi.org/10.18773/austprescr.2014.050
42. Martin Y. C. (2005). A bioavailability score. *Journal of medicinal chemistry*, *48*(9), 3164–3170. https://doi.org/10.1021/jm0492002
43. P, Surat. (2021, February 01). Importance of Solubility and Lipophilicity in Drug Development. AZoLifeSciences. Retrieved on July 02, 2021 from https://www.azolifesciences.com/article/Importance-of-Solubility-and-Lipophilicity-in-Drug-Development.aspx.
44. Dougherty D. A. (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science (New York, N.Y.)*, *271*(5246), 163–168. https://doi.org/10.1126/science.271.5246.163
45. Meyer, E. A., Castellano, R. K., & Diederich, F. (2003). Interactions with aromatic rings in chemical and biological recognition. *Angewandte Chemie (International ed. in English)*, *42*(11), 1210–1250. https://doi.org/10.1002/anie.200390319
46. Grossi, V., Peserico, A., Tezil, T., & Simone, C. (2014). p38α MAPK pathway: a key factor in colorectal cancer therapy and chemoresistance. *World journal of gastroenterology*, *20*(29), 9744–9758. https://doi.org/10.3748/wjg.v20.i29.9744
47. Zhuang, W. R., Wang, Y., Cui, P. F., Xing, L., Lee, J., Kim, D., Jiang, H. L., & Oh, Y. K. (2019). Applications of π-π stacking interactions in the design of drug-delivery systems. *Journal of controlled release : official journal of the Controlled Release Society*, *294*, 311–326. https://doi.org/10.1016/j.jconrel.2018.12.014
48. Jozkowicz, A., Was, H., & Dulak, J. (2007). Heme oxygenase-1 in tumors: is it a false friend?. *Antioxidants & redox signaling*, *9*(12), 2099–2117. https://doi.org/10.1089/ars.2007.1659
49. Walsh, A. A., Szklarz, G. D., & Scott, E. E. (2013). Human cytochrome P450 1A1 structure and utility in understanding drug and xenobiotic metabolism. *The Journal of biological chemistry*, *288*(18), 12932–12943. https://doi.org/10.1074/jbc.M113.452953
50. Lněničková, K., Dymáková, A., Szotáková, B., & Boušová, I. (2017). Sulforaphane Alters β-Naphthoflavone-Induced Changes in Activity and Expression of Drug-Metabolizing Enzymes in Rat Hepatocytes. *Molecules (Basel, Switzerland)*, *22*(11), 1983. https://doi.org/10.3390/molecules22111983

51. Koomoa, D. L., Borsics, T., Feith, D. J., Coleman, C. C., Wallick, C. J., Gamper, I., Pegg, A. E., & Bachmann, A. S. (2009). Inhibition of S-adenosylmethionine decarboxylase by inhibitor SAM486A connects polyamine metabolism with p53-Mdm2-Akt/protein kinase B regulation and apoptosis in neuroblastoma. *Molecular cancer therapeutics*, *8*(7), 2067–2075. https://doi.org/10.1158/1535-7163.MCT-08-1217