

# Permeability-Driven Virtual Screening Identifies Drug-Like Allosteric Inhibitors of MTHFD2 for Targeted Therapy in Non-Small Cell Lung Cancer

Abhay Chathuruthy

*Received April 17, 2025*

*Accepted August 22, 2025*

*Electronic access September 30, 2025*

Non-small cell lung cancer is the leading cause of cancer-related deaths in the U.S. Current chemotherapies lack selectivity, making enzymes over-expressed in cancer cells such as Methylenetetrahydrofolate Dehydrogenase 2 (MTHFD2) promising therapeutic targets. While existing inhibitors selectively bind to MTHFD2 with high potency, their poor permeability and selectivity prevents clinical implementation. This study aims to identify permeable inhibitors that target MTHFD2 through its allosteric site using computational simulations in Schrodinger's Maestro. Computational screening techniques in Maestro were performed on a library of compounds, prioritizing predicted Caco-2 permeability and strongest binding affinity. The top 10 compounds identified in this study demonstrated high permeability while the control indicated low permeability. While the top compounds permeability was higher than the controls (Caco-2 684 nm/sec vs. 0.864 nm/sec, respectively), a two-tailed Mann Whitney U-Test could not detect a statistically significant difference in the binding score between them. However, numerical values suggest that this study's top compound may be able to enter cancer cells more effectively than current MTHFD2 inhibitors with a slightly stronger level of binding potency ( $\Delta G$  binding score  $-72.39$  kcal/mol for top compound vs.  $-71.53$  kcal/mol for control, respectively). Drug likeness of these top 11 compounds was evaluated for clinical application, with 2 offering the most promising results. These findings identify computational compounds with predicted properties that could address the permeability barrier of current inhibitors. Further validation of the compounds' in-vitro potency may contribute to the development of a more effective, targeted non-small cell lung cancer treatment.

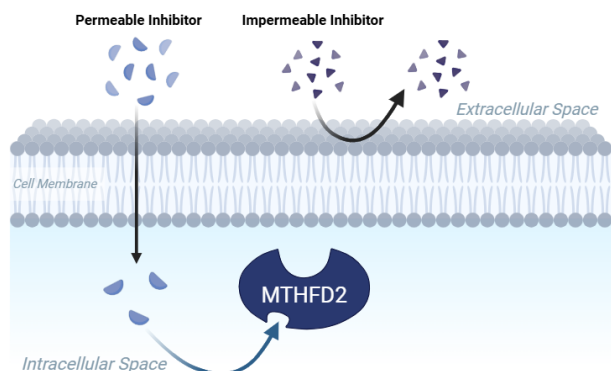
## Introduction

### Lung Cancer Statistics

Lung Cancer is the leading cause of cancer-related deaths in the United States and it accounts for 20% of all cancer deaths. The American Cancer Society estimates 226,650 new cases (110,680 in men and 115,970 in women) and 124,730 deaths (64,190 in men and 60,540 in women) in 2025 alone, with non-small cell lung cancer (NSCLC) accounting for 87% of these cases<sup>1</sup>. Despite advances in therapies and treatments, Surveillance, Epidemiology and End Results (SEER) data indicates a 5-year survival rate of 37.1% for regional and 9.7% for distant NSCLC<sup>2</sup>. Current therapies like chemotherapy are non-specific and impact both the cancerous cells and the normal cells. Chemotherapy has a number of potentially lethal side effects due to its chemotoxicity, and some life-threatening toxicities include myelosuppression, febrile neutropenia, severe oral mucositis, and sepsis<sup>3</sup>. Such statistics highlight the need for more selective therapies for NSCLC. One strategy for achieving such selectivity is targeting metabolic enzymes that are overexpressed in cancer cells but minimally expressed in healthy cells. Among these, MTHFD2 has attracted significant interest.

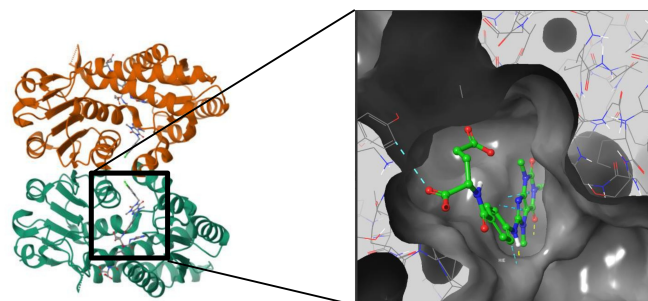
### MTHFD2 as a Cancer Target

One promising metabolic target is methylenetetrahydrofolate dehydrogenase 2 (MTHFD2), a mitochondrial enzyme that supports the folate-mediated one-carbon metabolism. It is a bi-functional enzyme that converts  $\text{CH}_2\text{-THF}$  to 5,10-methylenetetrahydrofolate through dehydrogenase activity. Then, it converts  $\text{CH}^+\text{-THF}$  to 10-CHO-THF through a cyclohydrolase activity. Ultimately, it catalyzes the production of 10-CHO-THF, which is then sent to the next enzyme in the metabolism<sup>4</sup>. This enzyme has been shown to be overexpressed in a variety of cancers while having low or undetectable expression in normal adult tissues<sup>5</sup>. Additionally, it has been shown to be involved in the proliferation, migration, and apoptosis of lung cancer cells by mediating MCM4, MCM7, ZEB1, Vimentin, and SNAI1, which contributed to tumor cell growth<sup>6</sup>. Data confirmed that the MTHFD2 gene and protein was upregulated in lung cancer patients and was correlated with poor survival in lung cancer<sup>6</sup>. Additional testing on NSCLC cancer cell lines found that the inhibition of MTHFD2 suppressed the progression of the cell lines and induced apoptosis of NSCLC cells<sup>7</sup>. Although MTHFD2's role in cancer is known, certain drug design challenges have hindered the effectiveness of current inhibitors.



**Fig. 1** Schematic that illustrates the difference between permeable and impermeable inhibitors for MTHFD2. Permeable inhibitors (blue, left) can pass through the cell membrane to bind to the protein while the impermeable inhibitors (purple, right) cannot permeate through the membrane and cannot interact with the target. (Student Created)

bioisosteric modifications to improve the potency and permeability. Addressing these challenges could identify inhibitors that are biologically and clinically effective.



**Fig. 2** Structure of MTHFD2 that highlights the allosteric site. (Left) Structure of MTHFD2 (PDB: 7EHM) with the black box showing the allosteric site. (Right) Zoomed-in representation of the allosteric pocket with an inhibitor binding to it.

## Problem Statement and Rationale

Currently, MTHFD2 inhibitors have been identified, but their lack of permeability limits their cellular efficacy. Additionally, the enzyme lies within two membrane-bound organelles which makes the transport of drug into both the mitochondria and the nucleus important for clinical implementation<sup>8</sup>. TH9619, an identified inhibitor for MTHFD2, was found to be highly potent ( $IC_{50} = 47$  nM), but had poor passive permeability due to its low lipophilicity (cLogP: -2.25) and large polar surface area (239 Å<sup>2</sup>)<sup>9</sup>. The same inhibitor targeted MTHFD1, a similarly structured protein that does not have the unique expression pattern, indicating low selectivity (and potential off-target effects)<sup>10</sup>. Because of this, TH9619 inhibits both MTHFD2 and the DC domain of MTHFD1, which results in thymidine depletion and subsequently the suppression of leukemic tumor growth<sup>10</sup>. A 2024 field review by Ramos et al. noted that TH9619 is not being transported into the mitochondria and it prevents the inhibitor from binding to mitochondrial MTHFD2<sup>10</sup>. This characteristic creates a cellular uptake challenge that prevents TH9619 from being implemented clinically as it is incapable of permeating into the mitochondria. To overcome selectivity barriers, Ma et al. crystallized the structure of MTHFD2 with xanthine derivative 15. This compound was found to bind to MTHFD2 through an allosteric site (a unique site on the protein) which inhibited MTHFD2 and exhibited no activity against MTHFD1<sup>11</sup>. Previous studies have performed pharmacophore-based virtual screenings of natural products that target MTHFD2's allosteric site, but they overlook cell permeability and optimization of the compounds<sup>12</sup>. This study also targets MTHFD2's allosteric site but performs an early-stage permeability screening, uses a chemically diverse library beyond natural products, and applies

## Significance and Purpose

An inhibitor that is selective to MTHFD2 and permeable addresses two problems associated with current MTHFD2 inhibitors: off-target effects and limited clinical viability. If confirmed by in-vitro and in-vivo studies in the future, the compounds identified could become starting points for a selective therapy for non-small cell lung cancer.

## Objectives

This study aims to identify cell-permeable, allosteric MTHFD2 inhibitors through permeability focused virtual screening. The investigation intends to generate a suitable list of top compounds for biochemical validation.

**Research Question:** How do computationally identified MTHFD2 inhibitors with high cell permeability compare to current inhibitors in permeability and potency?

**Hypothesis:** MTHFD2 inhibitors identified with a permeability-driven virtual screening will have higher Caco-2 permeability values and more favorable  $\Delta G$  binding scores compared to current inhibitors. To ground these objectives in context, it is important to define the scope of the work and the limitations of the study.

## Scope and Limitations

This study is entirely computational and limited to in-silico discovery. It covers protein preparation of a crystal structure of MTHFD2, structure based virtual screening of 140,000 compounds, SwissADME, and human colorectal adenocarcinoma cell (Caco-2) permeability predictions, and MM-GBSA calculations. It excludes all wet-lab validation. All information is

limited to the 140,000 compounds in the Blumberg Ligand Library. Because the values were calculated using simulations and estimations, they will require biochemical assays, cell-based testing, and in-vivo studies to validate the results. With these boundaries established, the methodology was designed to identify and evaluate promising compounds.

## Methodology Overview

140,000 compounds from the Blumberg ligand library were run through a virtual screening. The compounds were screened in several phases to ensure they were permeable, potent, and selective. First, the hits with optimal Caco-2 values and fit in the Lipinski Rule of 5 were prioritized. Later, the permeable compounds were screened based on their binding affinity to MTHFD2's allosteric site. Finally, the compound with the strongest binding interaction was optimized using Biososteric enumeration and the ADME properties of the top compounds were measured. Detailed protocols can be found in the methods section.

## Results

Among the compounds that were run through the virtual screening, 11 candidates with the strongest binding affinity and high permeability were selected for presentation. These compounds were labeled as the top compounds.

The Caco-2 values and  $\Delta G$  Binding Scores of the top 11 compounds as well as the MTHFD2 control are shown in Table 2. The MTHFD2 control showed low permeability with a Caco-2 value below 25 nm/sec, whereas all the top 11 compounds exceeded the 500 nm/sec cutoff for strong membrane transport. Additionally, three compounds from the top 11 compounds (BAS 2500328, ChemBridge.5608, BAS 1927109) exhibited Caco-2 values higher than the FDA-approved drug. These compounds were labeled as the "Permeability-Benchmarked Hits." However, their docking scores were substantially weaker than the MTHFD2 control by at least 10.81 kcal/mol (less negative  $\Delta G$ ). The top-binding compound (K838-0044) was the only compound in the top 11 to surpass the MTHFD2 control in binding affinity, but it was only slightly stronger (0.86 kcal/mol). Additionally, a two-tailed Mann-Whitney U Test could not detect a statistically significant difference between the  $\Delta G$  binding score of the control and K838-0044 ( $U = 16$ ,  $n_1 = n_2 = 6$ ).

To further assess the drug-like properties of the top 11 compounds, the ADME properties of the compounds were analyzed. Figure 3 shows the ADME properties of the top 11 compounds. Cells are color-coded based on whether the metric falls within the defined ideal range (green) and outside it (red), detailed in the Variables and Measurements section. Only one permeability-benchmarked hit (BAS 25000328) fulfilled the oral

drug-likeness criteria while Y050-00557, not a permeability-benchmarked hit, also fulfilled it. These properties are consistent with oral drug-likeness. Of the candidates that did not fulfill the drug-likeness criteria, eight failed the "Fraction Csp<sup>3</sup>" metric, which indicates a low proportion of sp<sup>3</sup> hybridized carbon atoms.

Additionally, Table 3 displays the structural formula of the strongest binding compound, permeability-benchmarked hits, and the identified drug-like compounds (candidates of interest). Structurally, K838-0044 contains a moderate molecular weight (463.6 g/mol) and a balanced total polar surface area (104.0 Å<sup>2</sup>), supporting cellular uptake, along with various functional groups (specifically benzenes and amides) which may contribute to stronger binding interactions. In contrast, the MTHFD2 control has a high molecular weight and a large total polar surface area from the multiple carboxylic acids, which limits its permeability. Additionally, BAS2500328 has the lowest total polar surface area (30.49 Å<sup>2</sup>), and this is likely the reason it is significantly more permeable than the other compounds.

The compound with the strongest binding affinity (K838-0044) was optimized using biososteric enumeration to improve both permeability and binding strength. Figure 4 shows the aromatic ring modifications made to the compound. Figure 5 presents the  $\Delta G$  Binding Scores associated with those modifications, while Figure 6 displays the corresponding Caco-2 values.

For the  $\Delta G$  Binding Scores, the second modification (V2) improved the binding affinity the most by -6.59 kcal/mol (from V0 to V2). This modification demonstrated a stronger binding interaction than both the MTHFD2 control and the unmodified K838-0044.

Regarding Caco-2 values, the first modification (V1) improved permeability the most by 128.2 nm/sec (V0 to V1). Although the permeability of the modified compound was higher than the unmodified K838-0044, it remained lower than the FDA-approved control.

## Discussion

### Restatement of Key Findings

A virtual screening of more than 140,000 compounds resulted in 11 candidates that had high permeability while the MTHFD2 control exhibited low permeability. Three of those compounds (BAS 2500328, ChemBridge.5608, BAS 1927109) displayed Caco-2 values higher than the FDA-approved control (Permeability-Benchmarked Hits). Further testing using SwissADME predictions identified two ligands with oral-drug likeness. Eight out of the 11 compounds failed the "Fraction Csp<sup>3</sup> metric, indicating a low proportion of sp<sup>3</sup> hybridized carbon atoms and a flat molecular structure. This is likely due to the presence of multiple aromatic rings, specifically benzene,

**Table 1** Lipinski Rule of 5 Values and Estimated Percent Human Oral Absorption of Top Compounds and Control (Student Created)

Compound Name	Molecular Weight (Da)	LogP o/w	HBD	HBA	Oral Absorption (%)
MTHFD2 Control	60	4.5	3.3	11.3	26.2
FDA Approved Control	555	7.7	1.0	3.8	100
K838-0044	464	4.4	1.0	8.5	100
BAS 2500328	283	4.7	1.0	2.5	100
ChemBridge.12766	363	3.4	1.0	5.5	95.9
ChemBridge.9851	410	3.8	2.0	7.5	100
ChemBridge.1065	329	4.1	1.0	4.0	100
L273-0378	483	3.4	2.0	9.0	96.2
K781-9015	403	3.1	2.0	7.0	93.8
ChemBridge.5608	378	4.6	2.0	4.8	100
BAS 1927109	340	2.5	2.0	7.0	92.8
Y050-0557	437	4.2	1.0	7.8	100
G350-0016	358	4.0	1.0	4.5	100

**Table 2** Caco-2,  $\Delta G$  Binding Scores, and Binding Scores of Top 11 Compounds and MTHFD2 Control (Student Created)

Compound Name	Caco-2 (nm/sec)	$\Delta G$ Binding Score (kcal/mol)
MTHFD2 Control	0.8640	-71.53
FDA Approved Drug	1556	X
K838-0044	684.3	-72.39
BAS 2500328	8192	-55.45
ChemBridge.12766	553.6	-68.52
ChemBridge.9851	512.6	-58.44
ChemBridge.1065	1194	-60.86
L273-0378	592.8	-58.32
K781-9015	530.5	-65.45
ChemBridge.5608	1562	-58.21
BAS 1927109	710.1	-35.19
Y050-0557	1326	-56.29
G350-0016	2940	-60.72

in the identified compounds. It suggests that the compounds may have unfavorable absorption/distribution and could be more metabolically unstable in clinical settings.

A two-tailed Mann-Whitney U Test could not detect a statistically significant difference between the  $\Delta G$  Binding Scores of K838-0044, the compound with the strongest binding affinity, and the MTHFD2 control ( $U = 16$ ,  $n_1 = n_2 = 6$ ). However, numerical values suggest that K838-0044 has a slightly stronger binding score ( $\Delta G$  binding score  $-72.39$  kcal/mol for K838-0044 vs.  $-71.53$  kcal/mol for control) while being significantly more permeable (Caco-2 684 nm/sec vs. 0.864 nm/sec, respectively). Bioisosteric enumeration of K838-0044 improved interaction strength by  $-6.59$  kcal/mol (from V0 to V2) and Caco-2 permeability by 128.2 nm/sec (from V0 to V1) compared to the unmodified K838-0044.

## Implications and Significance

These results demonstrate that a permeability-focused virtual screening can overcome a major obstacle in MTHFD2 inhibitors: poor permeability. By integrating Caco-2 calculations with docking and MM-GBSA scoring, this workflow identified candidates with strong binding affinity and oral drug-likeness through computational estimations. This study also identifies new compounds capable of binding to MTHFD2's allosteric site, providing a foundation for further structure-based optimization.

## Connection to Objectives

The primary objective was achieved: eleven compounds with high permeability were identified. Of these, three surpassed the FDA-approved control and two met the full ADME criteria, making them promising candidates for clinical implementation. A secondary aim was to analyze binding strength, and one com-

Compound Name	XLOGP3	MW	TPSA	Log S (ESOL)	Fraction Csp3	Num. Rotatable bonds
Control	2.95	603.4	177.6	-5.09	0.23	11
K838-0044	4.14	463.6	104.0	-5.20	0.20	8
BAS 2500328	4.18	283.4	30.49	-4.32	0.33	5
ChemBridge.12766	2.79	363.2	66.48	-3.96	0.12	4
ChemBridge.9851	3.54	410.4	84.50	-4.53	0.05	8
ChemBridge.1065	3.94	328.8	49.41	-4.48	0.22	4
L273-9015	2.77	482.6	135.4	-4.17	0.43	10
K781-9015	3.72	403.3	83.65	-4.58	0.06	7
ChemBridge.5608	4.65	377.8	67.01	-5.29	0.05	6
BAS 1927109	2.87	340.4	83.65	-3.92	0.06	5
Y050-00557	3.73	437.0	84.09	-4.81	0.38	6
G350-0016	4.30	358.4	60.68	-5.08	0.14	4

### Key

- Outside Ideal Range
- Within Ideal Range
- (rows with all green cells) the compound meets all major criteria for oral drug-likeness
- Permeability-Benchmarked Hits (surpassed FDA approved control in Caco-2)

**Fig. 3** Heat Map evaluating ADME metrics for Top 11 compounds. Columns represent lipophilicity (XLOGP3), molecular weight (MW), topological polar surface area (TPSA), solubility (LOG S (ESOL)), saturation (Fraction Csp3), and flexibility (Num. Rotatable bonds). Green cells indicate that the value falls within the drug-like range and red cells indicate that the value falls outside of the drug-like range (scoring criteria can be found in the Variables and Measurements section of the methodology). (Student Created)

pound was found to have a comparable  $\Delta G$  Binding Score to the MTHFD2 control.

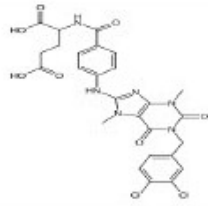
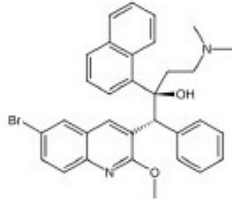
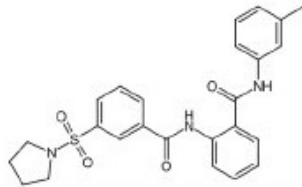
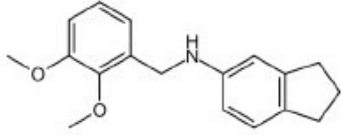
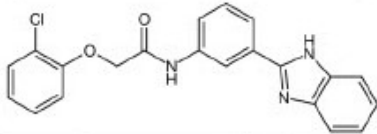
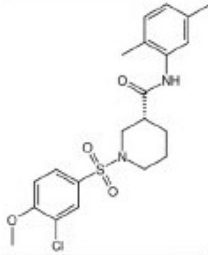
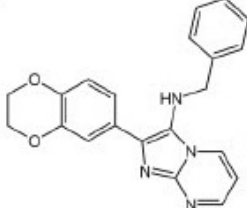
### Recommendations for Future Work

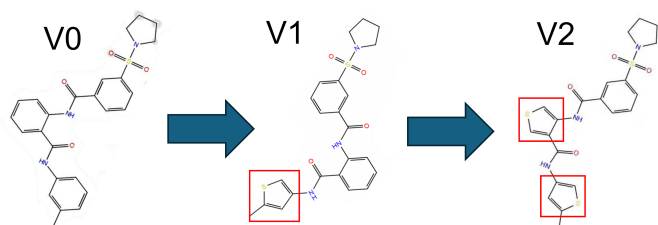
Future work will assess the metabolic stability of the top compounds, including CYP enzyme interactions, to ensure thera-

peutic effectiveness. Potential resistance mechanisms, such as MTHFD2 mutations, will also be investigated. Wet lab validation will be performed to confirm in-silico results, including Caco-2 assays for permeability and IC<sub>50</sub> assays for potency. Testing on lung cancer cell lines may reveal whether the identified compounds are viable treatments. Further bioisosteric enumeration could improve both binding affinity and permeabil-

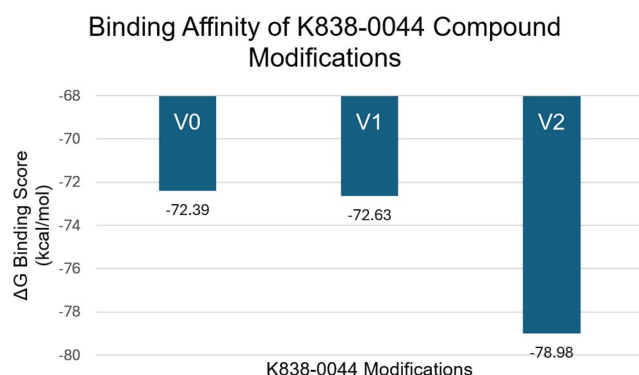


**Table 3:** Classification, name, and structural formula of the candidates of interest and controls (images student created using Leskoff)

Classification	Compound Name	Structural Formula
Control	MTHFD2 Control	
Control	FDA Approved Drug	
Top Compound (Strongest Binding Affinity)	K838-0044	
Permeability-Benchmarked Hit /Drug-Like	BAS 2500328	
Permeability-Benchmarked Hit	ChemBridge.5608	
Drug-Like	Y050-00557	
Permeability-Benchmarked Hit	G350-0016	



**Fig. 4** Structure of MTHFD2 that highlights the allosteric site. (Left) Structure of MTHFD2 (PDB: 7EHM) with the black box showing the allosteric site. (Right) Zoomed-in representation of the allosteric pocket with an inhibitor binding to it.



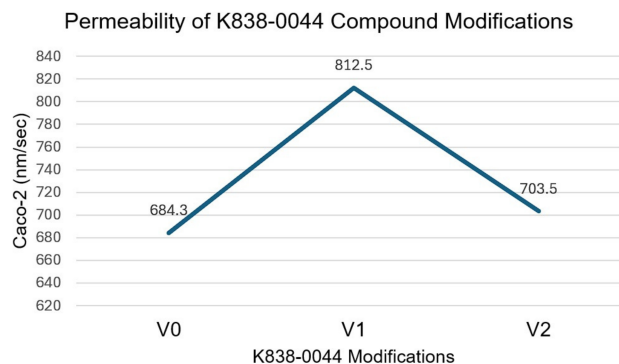
**Fig. 5** Bar graph of the  $\Delta G$  Binding Scores of the aromatic ring modifications made to K838-0044. The x-axis shows the different modifications made to K838-0044: V0 (original K838-0044), V1 (K838-0044 with one thiophene substitution), and V2 (K838-0044 with 2 thiophene substitutions). The y-axis shows the  $\Delta G$  binding score that corresponds to each modification with more negative values indicating stronger binding affinity. (Student Created)

ity. Long molecular dynamics simulations may refine docking score rankings.

## Limitations

All findings are predictive and rely on a single crystal structure, which may limit reliability. Permeability and binding affinity values are computational predictions not yet validated in vitro. The Caco-2 metric was predicted using QikProp, a regression model trained on experimental data from 150 compounds. In one study, QikProp's log PCaco-2 predictions had an  $r^2$  of 0.71 and RMS error of 0.56 log(nm/sec)<sup>13</sup>. Other ADME metrics were also predicted using validated QikProp models.

Binding scores were calculated using Prime MM-GBSA. Chandna et al. (2015) reported a correlation coefficient up to 0.77 between computational predictions and experimental results for N-myristoyltransferase inhibitors<sup>14</sup>. Although MTHFD2 benchmarking has not been performed, these find-



**Fig. 6** Line graph of the Caco-2 values of the aromatic ring modifications made to K838-0044. The x-axis shows the different modifications made to K838-0044: V0 (original K838-0044), V1 (K838-0044 with one thiophene substitution), and V2 (K838-0044 with 2 thiophene substitutions). The y-axis shows the predicted Caco-2 permeability values that corresponds to each modification with higher values indicating greater permeability. (Student Created)

ings support Prime MM-GBSA's use for identifying promising targets.

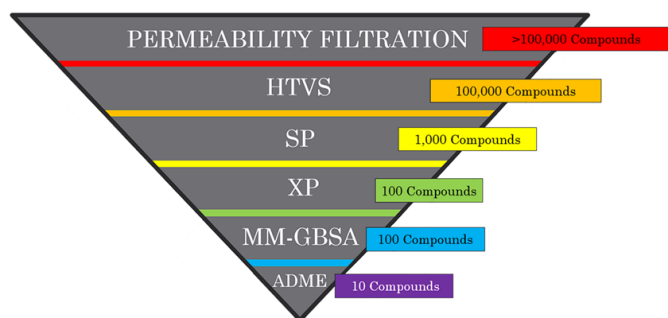
The V2 modification to K838-0044 improved binding score by 6.59 kcal/mola large change for a single ring modification. This result should be interpreted cautiously, as the calculation could not be independently repeated due to loss of software access after an internship. The Mann-Whitney U Test was conducted with a small sample size ( $n = 6$ ), which limits statistical power and does not generate a p-value. No additional controls were used to validate binding site boundaries beyond Glide's default ligand-based grid generation. Although a known potent MTHFD2 inhibitor was used as a positive control for MM-GBSA predictions, no benchmarking studies were performed to validate docking, MM-GBSA, or Caco-2 predictions.

Caco-2 permeability estimates a compound's ability to pass through the cell membrane but has limitations for predicting mitochondrial uptake, which is critical for MTHFD2 targeting. While computational screening identifies promising candidates, in vitro testing is essential to confirm functional inhibition and clinical relevance.

## Closing Thought

By demonstrating how permeability-focused virtual screening can identify drug-like compounds against MTHFD2, this study illustrates how computational methods can accelerate the development of tumor-selective therapies. Pairing these predictions with experimental validation offers a tangible path toward the first clinically viable MTHFD2 inhibitor for non-small cell lung cancer.

## Methods



**Fig. 7** Maestro screening workflow implemented to identify permeable compounds with the strongest binding affinity to MTHFD2's allosteric site. Each stage represents a screening step with increasingly stricter criteria (top to bottom), and the number of compounds processed at each step is shown. (Student Created)

## Research Design

### Experimental Design

This study used an in-silico, screening-based experimental design with two main phases. The first phase involved high-throughput virtual screening focused on permeability and docking of a chemical library on MTHFD2's allosteric site. The second phase optimized the ligand with the strongest binding affinity to improve permeability and interaction strength, and predicted the ADME properties of the top 11 compounds using SwissADME. This approach prioritized early identification of drug-like and permeable allosteric inhibitors for MTHFD2.

### Computational Sample

The dataset used was the Blumberg Ligand Library from the PA Biotechnology Center, containing 139,588 crystallized compound structures. A crystal structure of MTHFD2 (PDB: 7EHM) was prepared using Maestro's Protein Preparation Wizard. Glide was used to generate a receptor grid centered on the co-crystallized allosteric ligand from PDB: 7EHM. Grid dimensions were automatically set by Glide based on the reference ligand.

### Variables and Measurements

Software tools used:

- **Maestro:** Simulates interactions between small molecules and proteins.
- **SwissADME:** Predicts pharmacokinetic properties.

- **Protein Data Bank:** Provides crystal structures of proteins.

### Key Variables:

- **Caco-2 Permeability** (nm/sec): Predicts membrane permeability. High: > 500, Low: < 25 (QikProp manual). RMS error: 0.56 log(nm/sec).
- **$\Delta G$  Binding Score** (kcal/mol): Estimates binding free energy to MTHFD2.

### ADME Properties:

- **XLOGP3:** Lipophilicity. Ideal:  $-0.7$  to  $+5.0$
- **MW:** Molecular weight. Ideal: 150500 g/mol
- **TPSA:** Topological Polar Surface Area. Ideal: 20130 Å<sup>2</sup>
- **Log S (ESOL):** Solubility. Ideal:  $-6$  to  $0$
- **Fraction Csp3:** Saturation. Ideal: 0.251
- **Num. Rotatable Bonds:** Flexibility. Ideal: 09

### Controls

- **FDA Approved Control:** Bedaquiline, a mitochondrial-targeting drug, used to benchmark permeability.
- **MTHFD2 Control:** A known MTHFD2 inhibitor from PDB 7EHM, used to benchmark binding affinity.

### Data Collection and Procedure

The compound library was filtered using Lipinski's Rule of Five and logP o/w limits (06). Screening steps:

1. **HTVS:** Rapid docking to identify top 1000 compounds by interaction strength.
2. **SP Docking:** More accurate scoring to identify top 100 compounds.
3. **XP Docking:** Extra precision scoring and pose generation.
4. **MM-GBSA:** Free energy estimates ( $\Delta G$ ) accounting for protein flexibility.

### ADME Predictions

SwissADME predicted ADME values to assess oral drug-likeness, including: XLOGP3, MW, TPSA, Log S, Fraction Csp3, and Num. Rotatable Bonds.



---

## Bioisosteric Enumeration

K838-0044 was modified to improve binding and permeability:

- **V1:** One benzene replaced with thiophene.
- **V2:** V1 plus an additional benzene replaced with thiophene.

These modifications yielded the greatest improvements.

## Mann-Whitney U-Test

Six MM-GBSA jobs were run for both K838-0044 and the MTHFD2 control. A two-tailed Mann-Whitney U-Test compared  $\Delta G$  scores:

- **Null Hypothesis:** No difference in ranks.
- **Alternative Hypothesis:** Difference in ranks.
- **U-statistic:** Calculated manually; values  $< 5$  considered significant.

## Data Analysis

$\Delta G$  Binding Scores and Caco-2 values were compared using Maestro. Heatmaps and visualizations were created in Excel.

## Ethical Considerations

Only publicly available molecular and structural data were used. No human, animal, or clinical data were involved; no ethics approval required.

## Acknowledgements

I would like to thank my mentors Dr. Scott Willett and Ryan Ford for their guidance throughout this project. I am also grateful to my adult sponsor Mark Hayden and to my parents for their continued support.

## References

- 1 A. C. Society, *Lung Cancer Early Detection, Diagnosis, and Staging*, <https://www.cancer.org/cancer/types/lung-cancer/detection-diagnosis-staging/survival-rates.html>, Retrieved May 16, 2025, from.
- 2 E. Surveillance and E. Results, *Cancer Stat Facts: Lung and Bronchus Cancer. National Cancer Institute*, <https://seer.cancer.gov/statfacts/html/lungb.html>, Retrieved May 16, 2025, from.
- 3 R. Juthani, S. Punatar and I. Mittra, *New light on chemotherapy toxicity and its prevention*.
- 4 Z. Z and L. GKK, *More Than a Metabolic Enzyme: MTHFD2 as a Novel Target for Anticancer Therapy?*
- 5 R. Nilsson, M. Jain and N. Madhusudhan, *Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer*.
- 6 J. Mo, Z. Gao and L. Zheng, *Targeting mitochondrial one-carbon enzyme MTHFD2 together with pemetrexed confers therapeutic advantages in lung adenocarcinoma*.
- 7 F. Zhou, Z. Yuan, Y. Gong, L. Li, Y. Wang, X. Wang, C. Ma, L. Yang, Z. Liu, L. Wang, H. Zhao, C. Zhao and X. Huang, *Pharmacological targeting of MTHFD2 suppresses NSCLC via the regulation of ILK signaling pathway*.
- 8 J. Huang, Y. Qin, C. Lin, X. Huang and F. Zhang, *MTHFD2 facilitates breast cancer cell proliferation via the AKT signaling pathway*.
- 9 N. Bonagas, N. Gustafsson, M. Henriksson, P. Marttila, R. Gustafsson, E. Wiita, S. Borhade, A. Green, K. Vallin, A. Sarno, R. Svensson, C. Gk-trk, T. Pham, A. Jemth, O. Loseva, V. Cookson, N. Kiweler, L. Sandberg, A. Rasti, J. Unterlass and T. Helleday, *Pharmacological targeting of MTHFD2 suppresses acute myeloid leukemia by inducing thymidine depletion and replication stress*.
- 10 L. Ramos, M. Henriksson, T. Helleday and A. C. Green, *Targeting MTHFD2 to Exploit Cancer-Specific Metabolism and the DNA Damage Response*.
- 11 L. Lee, Y. Peng, H. Chang, T. Hsu, C. Lu, C. Huang, C. Hsueh, F. Kung, C. Kuo, W. Jiaang and S. Wu.
- 12 N. Rana, D. Patel and M. Parmar, *Targeting allosteric binding site in methylenetetrahydrofolate dehydrogenase 2 (MTHFD2) to identify natural product inhibitors via structure-based computational approach*.
- 13 W. Jorgensen and Schrödinger, *QikProp Technical Manual*, <https://qikprop.molssi.org/static/pdf/QP30.manual.pdf>, Retrieved August 9, 2025, from.
- 14 N. Chandna, M. Kumari, C. Sharma, M. Vijjulatha and J. Kapoor, *QM/MM Docking Strategy and Prime/MM-GBSA Calculation of Celecoxib Analogues as N-myristoyltransferase Inhibitors*.