

Article

Screening and Analysis of Potential Inhibitors of SHMT2

Bojin Chen ^{1,2} and John Z. H. Zhang ^{1,2,3,4,5,*} 

- ¹ Shanghai Engineering Research Center of Molecular Therapeutics & New Drug Development, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062, China; chenbojin135@163.com
- ² NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China
- ³ Faculty of Synthetic Biology and Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
- ⁴ Department of Chemistry, New York University, New York, NY 10003, USA
- ⁵ Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, China
- * Correspondence: john.zhang@nyu.edu

Abstract: Serine hydroxymethyltransferase 2 (SHMT2) has garnered significant attention as a critical catalytic regulator of the serine/glycine pathway in the one-carbon metabolism of cancer cells. Despite its potential as an anti-cancer target, only a limited number of inhibitors have been identified so far. In this study, we employed seven different scoring functions and skeleton clustering to screen the ChemDiv database for 38 compounds, 20 of which originate from the same skeleton structure. The most significant residues from SHMT2 and chemical groups from the inhibitors were identified using ASGBIE (Alanine Scanning with Generalized Born model and Interaction Entropy), and the binding energy of each residue was quantitatively determined, revealing the essential features of the protein–inhibitor interaction. The two most important contributing residues are TYR105 and TYR106 of the B chain followed by LEU166 and ARG425 of the A chain. The findings will be greatly helpful in developing a thorough comprehension of the binding mechanisms involved in drug–SHMT2 interactions and offer valuable direction for designing more potent inhibitors.

Keywords: virtual screening; binding free energy; molecular dynamics simulation; alanine scan; interaction entropy



Citation: Chen, B.; Zhang, J.Z.H. Screening and Analysis of Potential Inhibitors of SHMT2. *Biophysica* **2023**, *3*, 651–667. <https://doi.org/10.3390/biophysica3040044>

Academic Editors: Javier Sancho and Paulino Gómez-Puertas

Received: 7 October 2023
Revised: 27 November 2023
Accepted: 29 November 2023
Published: 3 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The metabolic process of cancer cells is very different compared to that of normal cells, and this abnormal cellular metabolic process is termed reprogrammed metabolism, in which one-carbon unit (1C) metabolism [1–3] is necessary to maintain the proliferation and growth of cancer cells. One-carbon unit (1C) is mainly derived from the conversion of serine and glycine in mitochondria, and is transported intracellularly as tetrahydrofolate (THF), which is a carrier of the substances that are essential for cell proliferation [1,3–13]. Hence, predicting tumor behavior and inhibiting the essential pathways involved in one-carbon metabolism presents an opportunity to restrain tumor progression.

Serine hydroxymethyltransferase is a critical enzyme involved in the synthesis of serine and glycine. In human cells, there are two subtypes of this enzyme present in the cytoplasm and mitochondria, named SHMT1 (cytoplasmic) and SHMT2 (mitochondrial) according to their location. Among them, SHMT2 is related to the reprogramming metabolic process of many cancers, such as liver cancer, breast cancer, B-cell lymphoma, lung cancer, tongue squamous cell carcinoma, and glioma, and other cancers and tumors have an obvious high expression [14–16]. Given its role in cancer metabolism, SHMT2 has garnered as an attractive therapeutic target, receiving widespread attention in recent years.

Research [17] indicates that both antibodies and small molecules can exert inhibitory effects on proteins. Currently, there are two major classes of molecules known for their inhibitory effects on SHMT2. These include pyrano- and pyrazole-like molecules as well as structural analogues of folic acid. The former class has been shown to inhibit the activity of plant SHMT2 proteins and has even been published as an herbicide [18]. Researchers have also tested compounds with this backbone on human-derived SHMT2 and observed the same inhibitory effect [19]. The binding strength of pyrano- and pyrazole-like molecules to SHMTs has been determined with IC50 values, ranging from 10 nM to 5 μ M [14]. In contrast, folic acid structural analogue molecules such as pemetrexed and lomitrexed exhibit weak binding to SHMT2, with IC50 values ranging from 100 μ M [20,21].

The drug development process is lengthy and costly, and in order to address these two major issues, computer-aided drug design (CADD) [22] has emerged. Currently, there are two types of computer-aided drug design methods: ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS) [23,24]. After obtaining the selected potential inhibitor molecules, it is necessary to quantitatively analyze the protein–ligand interactions and accurately predict the binding affinity between the protein and the molecules. There are several computational methods available for predicting binding energy, among which MM/PB(GB)SA [25] is widely used. However, this method does not account for the influence of entropy on binding free energy. In order to predict binding energy more accurately between the protein and the ligand molecules, our research group has developed a novel method named ASGBIE [26–28]. It involves using alanine scanning [29], MM/GBSA, and interaction entropy [30].

He et al. calculated the binding affinities of 28 ligands to SHMT2 [28] and also performed structure-based virtual screening (SBVS) on the Specs database, resulting in the identification of three potential inhibitors that are verified by experiment [31]. The scaffold of the inhibitor molecules identified by Liping He et al. consists of the pyranopyrazole scaffold. In contrast, among the small molecules that have screened, only a fraction of compounds exhibits the pyranopyrazole scaffold, while the remaining small molecules possess distinct scaffolds.

In this study, a screening was performed on 1.6 million molecules from the ChemDiv database, resulting in the identification of 38 potential SHMT2 inhibitor molecules. To quantitatively predict the binding energy between the protein SHMT2 and its inhibitor molecules, the ASGBIE method was employed to analyze the protein–ligand interactions. A comparison was made with the binding mode of the protein–ligand molecule in the crystal structure 5V7I, revealing the involvement of the same key residues in the binding energy, consistent with the crystal structure.

2. Results and Discussion

2.1. Virtual Screening Results

The results of skeleton clustering are shown in Figure 1. During the skeleton clustering process, a classification criterion cutoff of 0.4 was chosen. As a result, 5000 molecules were divided into 435 clusters based on their skeletons. The largest cluster consisted of 41 compounds, while 264 clusters contained only a single compound each. Moreover, the distribution of compounds within the clusters provided valuable insights into the structural diversity present in the dataset.

Twenty molecules were selected based on the backbone clustering for their potential inhibitory activity. All 20 of these compounds share a common backbone structure of 6-amino-1,4-dihydropyrano[2,3-c] pyrazole-5-carbonitrile, as illustrated in Figure 2, which is the molecular scaffold structure of the existing inhibitor. In addition, the other 18 compounds with the highest GlideScore in the clusters containing molecules with different structures were selected. Furthermore, 38 molecules were selected that might be associated with molecules with SHMT2 inhibitory activity. Table S1 displays the structures for each of these 38 molecules.

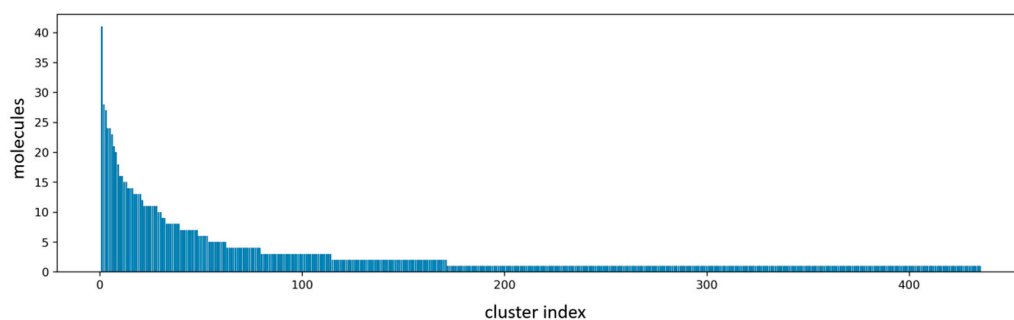


Figure 1. The result of clustering (similarity cutoff is 0.4). The horizontal axis represents the number of clusters obtained, while the vertical axis represents the number of molecules contained in each cluster.

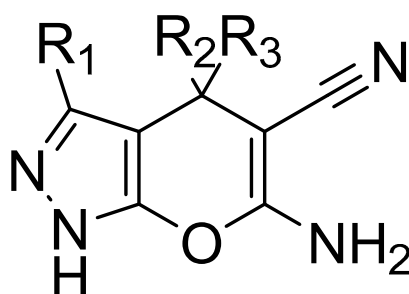
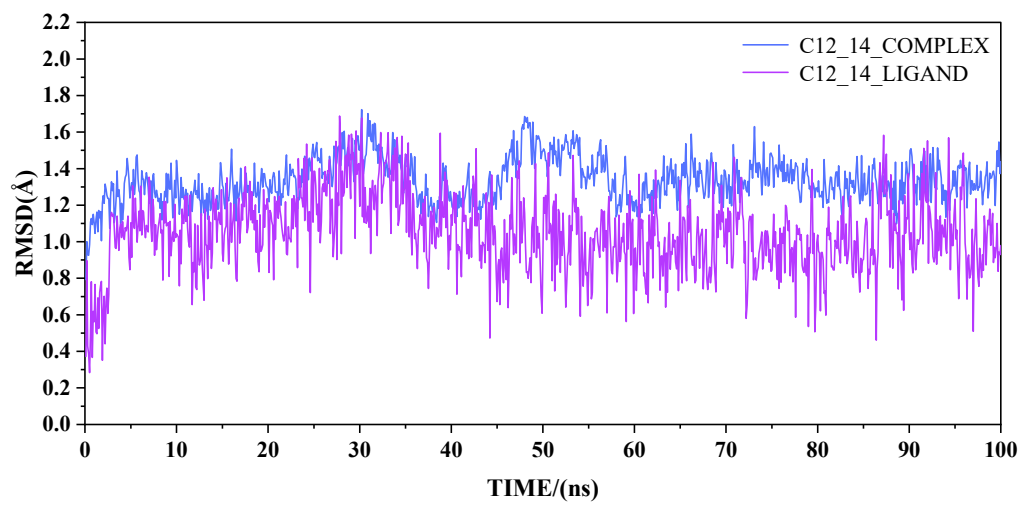


Figure 2. The common scaffold of all compounds in the selected 12th cluster, the pyranopyrazole structure.

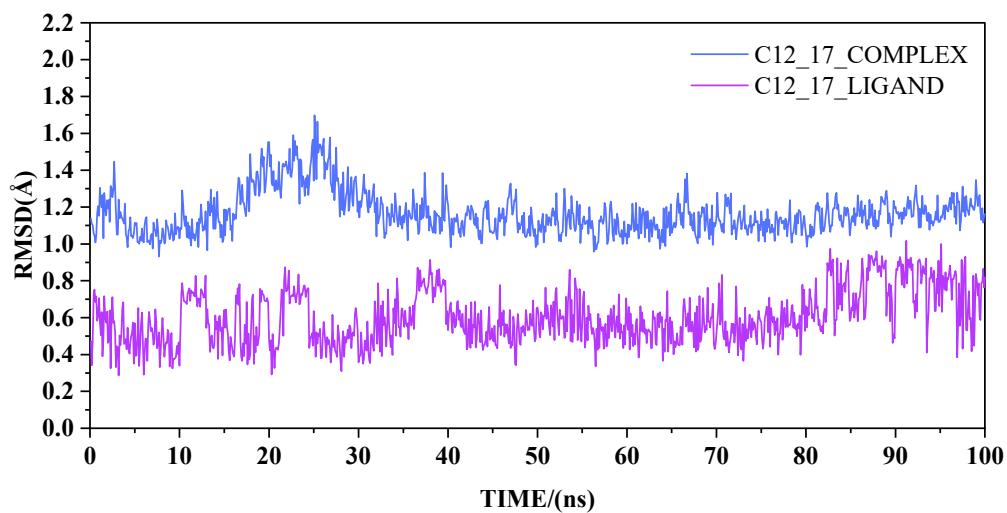
The ranking of all molecules in the seven scoring functions is presented in Table S2. Each molecule number starts with the letter C, where the first number denotes the order of the 19 selected clusters and the second number represents the Glide scoring ranking of the molecule in the cluster.

2.2. Result from Molecular Dynamics Simulation

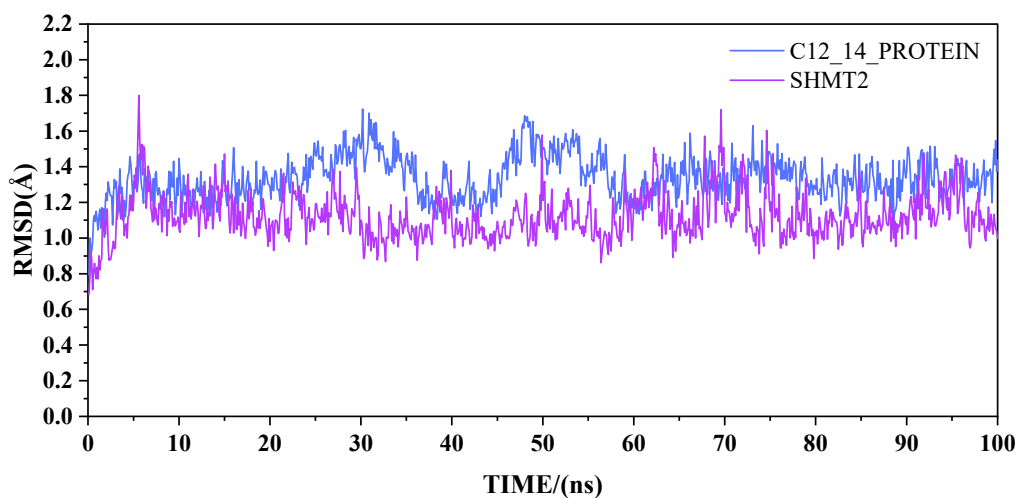
The compounds C12_14 and C12_17 possess the same molecular backbone but exhibit a significant difference in their comprehensive ranking. To assess the influence of simulation time on the simulation results, molecular dynamics simulations were performed, using a 100 ns MD simulation, on the complexes formed by these two compounds with SHMT2. Figure 3a,b show the RMSD values of these two compounds over a 100 ns simulation. Specifically, Figure 3a represents the simulation trajectory of the C12_14–SHMT2 complex system, while Figure 3b represents the C12_17–SHMT2 complex system. As shown in the figure, the RMSD of both complex systems remains below 2 Å throughout the entire simulation process, and the fluctuation of these two systems is quite small in scale (less than 1 angstrom). We also plotted RMSD values of both ligand-bound and apo proteins in Figure 3c. Also, the RMSD of the ligand-bound protein is slightly larger than that of the apo-protein, but the difference is quite small and is largely believed to be the result of random fluctuations of individual trajectories. Thus, the conformational change in the protein upon ligand binding should be quite small.



(a)



(b)



(c)

Figure 3. RMSD of two SHMT2–ligand complex systems and that of the apo-SHMT2 system: (a) the C12_14–SHMT2 complex system, (b) the C12_17–SHMT2 complex system, and (c) the apo-SHMT2 system in comparison with the C12_14–SHMT2 system.

2.3. Binding Free Energies

Data for the binding free energy calculations for the 38 ligands and 8Z1 in the 5V7I complex are provided in Figure 4 and Table S3. From Figure 4 it is evident that among the 38 compounds studied, C14_1 and C12_14 are two compounds with different skeletons, and they exhibit superior binding ability towards the receptor protein, with binding free energies of -32.27 kcal/mol and -25.22 kcal/mol, respectively. It is worth noting that the binding free energies of the top two compounds with SHMT2 exceed -25 kcal/mol. The binding free energy of 8Z1 is -24.61 kcal/mol. These findings indicate that C14_1 and C12_14 compounds exhibit exceptional binding affinity towards SHMT2 and hold significant promise as drug candidates for inhibiting the biological activity of the SHMT2 target protein. Our focus will be on investigating the binding mechanism of 8Z1, C14_1, and C12_14 with the receptor proteins.

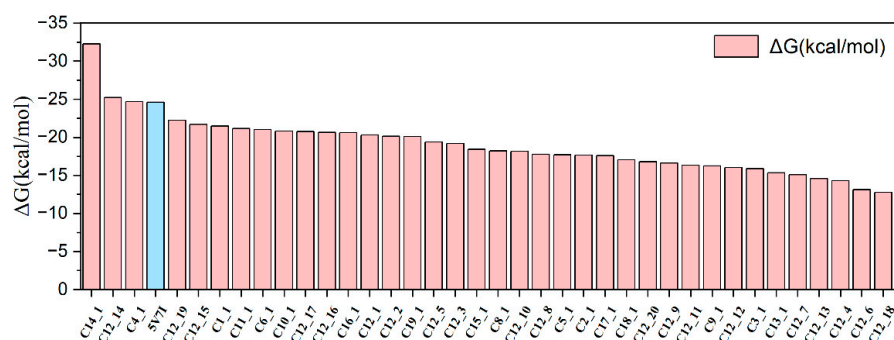


Figure 4. Binding free energy ranking of 38 complex systems and that of 5V7I (in blue).

2.4. Identification of Hot Spots

The ASGBIE method was used to calculate the detailed results of each part for the complex system C14_1–SHMT2, as shown in Table 1. Only the energy of each part of the amino acid residue with $\Delta\Delta G < -1.0$ kcal/mol is displayed in the table. The binding free energy of the composite system is composed of enthalpy and entropy, with enthalpy being composed of four components: van der Waals energy, electrostatic energy, polar solvation energy, and non-polar solvation energy. The data in the table indicate that the van der Waals energies are significantly higher than those of the other components, and thus dominate the binding mechanism in the complex system.

Table 1. Average energy values for each residue ($\Delta\Delta G < -1.0$ kcal/mol) in the C14_1–SHMT2 complex system. ($\Delta\Delta E_{vdw}$ is van der Waals energy, $\Delta\Delta E_{ele}$ is electrostatic energy, $\Delta\Delta GB$ is polar solvation energy, and $\Delta\Delta NP$ is binding enthalpy, IE is binding entropy, and $\Delta\Delta G_{bind}^{x \rightarrow a}$ is binding free energy).

Residue	$\Delta\Delta E_{vdw}$	$\Delta\Delta E_{ele}$	$\Delta\Delta GB$	$\Delta\Delta NP$	$\Delta\Delta H$	IE	$\Delta\Delta G_{bind}^{x \rightarrow a}$
TYR105_B	-5.05	0.05	0.71	-0.28	-4.57	0.39	-4.18
LEU166_A	-4.50	-0.07	0.55	-0.35	-4.36	0.44	-3.92
ARG425_A	-3.18	-0.08	0.02	-0.22	-3.46	0.24	-3.22
TYR106_B	-2.40	0.13	-0.17	-0.17	-2.61	0.15	-2.46
LYS409_A	-1.82	-0.01	0.06	-0.17	-1.94	0.14	-1.80
GLU98_B	-1.75	-0.22	0.23	-0.14	-1.88	0.17	-1.71
TYR176_A	-1.93	0.00	0.34	-0.19	-1.78	0.16	-1.62
PHE320_B	-2.11	-0.55	1.14	-0.12	-1.64	0.21	-1.43
ASN408_A	-1.37	-0.31	0.15	-0.11	-1.63	0.23	-1.40
LEU172_A	-1.74	-0.30	0.59	-0.13	-1.59	0.21	-1.38
HIP171_A	-1.21	-0.17	-0.01	-0.09	-1.48	0.23	-1.25
ASN410_A	-1.38	-0.38	0.52	-0.10	-1.33	0.25	-1.08

In addition, to verify the reliability of the calculation results, a comparison was made between the residue contributions of C14_1-SHMT2 and 5V7I, as shown in Figure 5. It illustrates that the residue TYR105, located on the B chain near the binding pocket, is the residue that contributes most to the binding of the ligand and receptor proteins. The residues LEU166 on the A chain and TYR106 on the B chain also significantly contribute to the binding energy.

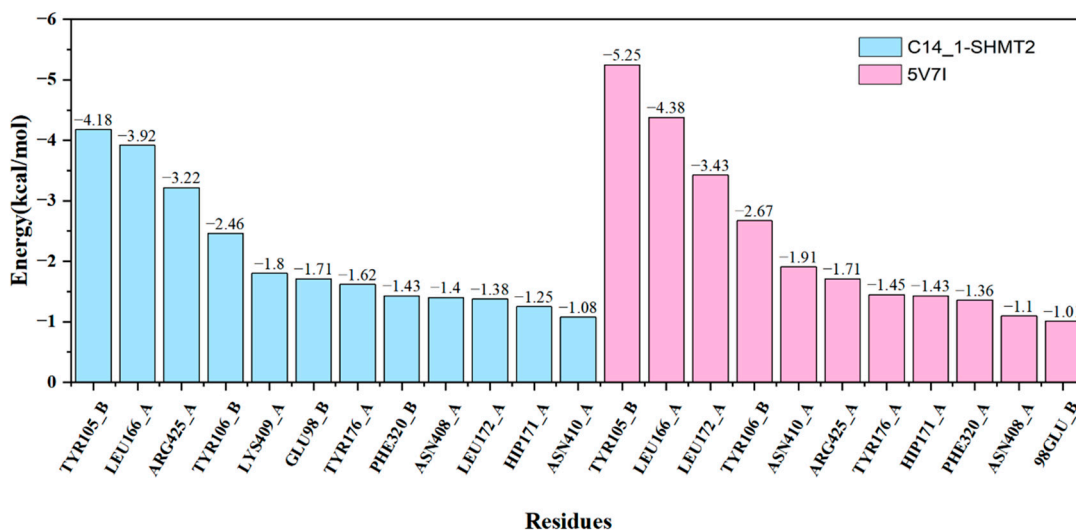


Figure 5. The contribution of residues to the complex C14_1-SHMT2 and 5V7I within 5 Å of the pocket (only residues with $\Delta\Delta G < -1.0$ kcal/mol are listed).

Not all amino acid residues in proteins significantly contribute to the binding of small drug molecules to target proteins. Only a few key residues located near the binding pocket play a crucial role in binding molecules to receptor proteins. These amino acids are commonly referred to as hot spots. To uncover the binding mechanism of receptor proteins and molecules and provide ideas and schemes for designing drugs that target SHMT2, it is essential to analyze the binding mechanism of hot spots to molecules and study their contribution to the binding free energy. In this study, hot spots were defined as those that contributed less than -2 kcal/mol to the binding free energy.

Compounds C14_1 and C12_14 have different scaffolds, with C12_14 featuring a furan-pyrazole scaffold structure. These two compounds exhibit the highest binding free energies when interacting with the protein SHMT2; thus, we selected these two complex systems and 5V7I for binding mechanism analysis. Figure 6 ranks the hot spot residues in each complex system according to their energy contribution.

Figure 6a shows that among the complex systems of compound C14_1 and SHMT2, residue TYR105_B contributes the most to the binding free energy, with an energy of -4.18 kcal/mol, followed by LEU166_A residue with -3.92 kcal/mol. Furthermore, residue LYS409 contributes -1.80 kcal/mol of binding energy, which is very close to the energy contribution of hot spot residues.

The total binding energy of the complex system C12_14-SHMT2 is second only to that of C14_1-SHMT2. From Figure 6b, it can be found that six residues in this system are predicted to be hot spots, with the highest incidence. The A-chain residue LEU166 contributes the most energy of -4.39 kcal/mol, followed by the B-chain residue TYR105 and the A-chain residue ARG425, which generate energy equal to -3.67 kcal/mol and -3.27 kcal/mol, respectively. These three hot spot residues with the highest binding energy contributions were the same as those in the C14_1-SHMT2 system.

As can be observed from Figure 6c, the hot spot residues in 5V7I are identical to those in the two aforementioned systems, and the binding energy contribution of residue TYR105 in chain B is the highest.

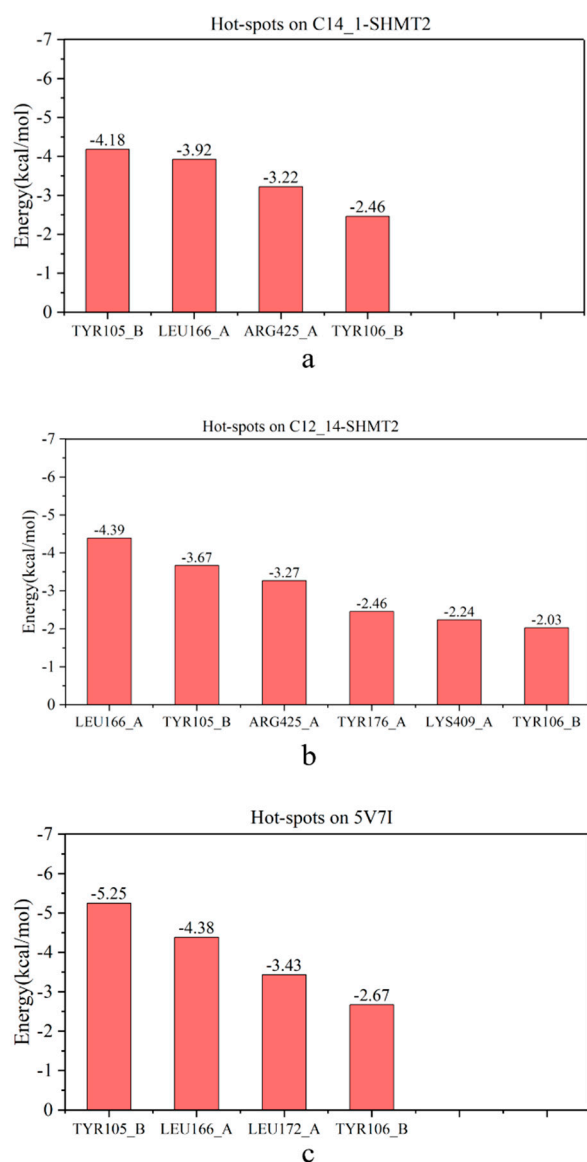


Figure 6. The total energy contribution of hot spot residues: (a) the hot spot residue of C14_1-SHMT2 are TYR105_B, LEU166_A, ARG425_A, and TYR106_B; (b) the hot spot residue of C12_14-SHMT2 include LEU166_A, TYR105_B, ARG425_A, TYR176_A, and LYS409_A; and (c) the hot spot residue of 5V7I are TYR105_B, LEU166_A, LEU172_A, and TYR106_B.

Furthermore, to further analyze the binding mechanism of the ligand–protein complex systems and the role of hot spot residues, we calculated the average energy contribution of all hot spot residues in the 38 systems, as shown in Figure 7.

Figure 7 shows that the B-chain residue TYR105 and the A-chain residue LEU166 were consistently predicted to be hot spot residues in all 38 complex systems, with an average energy contribution of -3.56 kcal/mol and -3.10 kcal/mol, respectively. This suggests that these two residues have the most significant impact on binding affinity and are the most critical amino acids in the binding process. Additionally, the A-chain residue ARG425 and the B-chain residue TYR106 also play important roles in binding energy, with an average energy contribution of -3.09 kcal/mol and -2.72 kcal/mol. Moreover, there was a slightly less predictable, but still significant hot spot on the A-chain residue TYR176, with an average energy contribution of -2.46 kcal/mol. Additionally, residue 98GLU_B was identified as a hot spot only once and contributed a binding energy of -2.06 kcal/mol.

Based on the above study, of the complex systems formed by the binding of 38 compounds to SHMT2 proteins, four hot spot residues made the most significant contribution

to the binding mechanism. These residues were LEU166 and ARG425 from the A chain and TYR105 and TYR106 from the B chain. These residues were found to have the highest frequency of occurrence and the greatest influence on the binding mechanism.

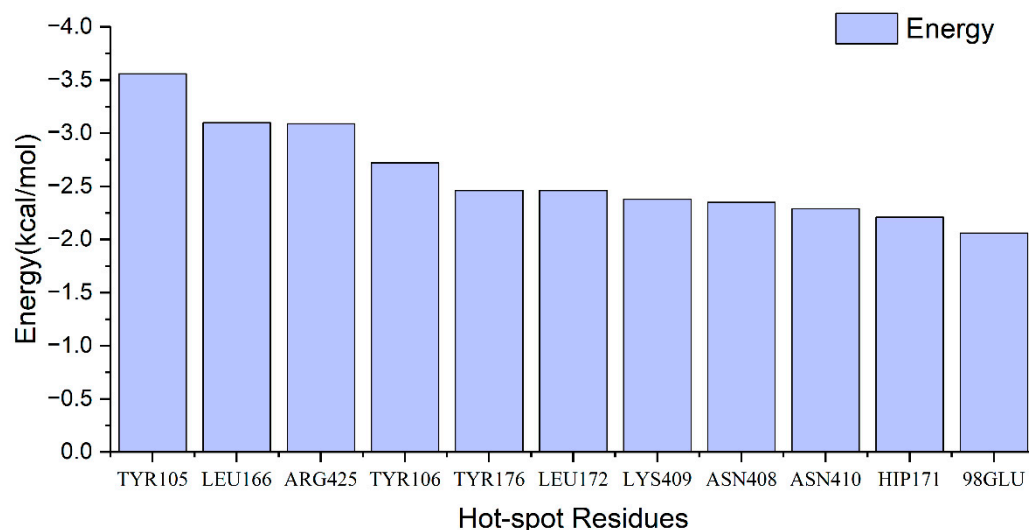


Figure 7. Average energy contribution of hot spot residues across the 38 complex systems.

2.5. Analysis of Binding Mechanism

Exploring the way in which hot spot residues bind to molecules and their contribution to the binding of both is essential to understand the binding mechanism between molecules and receptor proteins. This knowledge can even provide useful guidelines and suggestions for the design of targeted SHMT2 inhibitor molecules. To achieve this goal, we decomposed the binding free energy on the hot spot residues of each system into the contributions of van der Waals energy, nonpolar solvation energy, electrostatic energy, and polar solvation energy. Figure 8 displays the different types of energy contributions for all hot spot residues predicted in the three systems.

A comprehensive analysis of the energy terms for each residue in the binding energy contribution clearly shows that van der Waals interactions of hot spot residues dominate the entire binding process. For example, in the C14_1-SHMT2 complex system, the van der Waals energies of the four hot spot residues were -5.05 kcal/mol, -4.50 kcal/mol, -3.18 kcal/mol, and -2.40 kcal/mol, while the electrostatic energies were only 0.05 kcal/mol, -0.07 kcal/mol, -0.08 kcal/mol, and 0.13 kcal/mol, which were much less than the energy contribution of the van der Waals part. The figure also indicates that the main contribution to the binding energy in other systems is also the van der Waals energy part, while nonpolar solvation energy, electrostatic energy, and polar solvation energy have little effect on the overall binding effect.

Through an analysis of the binding free energy of the three complex systems and the specific energy contribution of individual residues near the binding pocket, we were able to identify the four hot spot residues, TYR105_B, LEU166_A, ARG425_A, and TYR106_B, which play a major role in the binding mechanism. To deeply investigate the interaction pattern between the compound molecules and the hot spot residues, we used the spatial conformation of the C14_1-SHMT2 system as a representative, and compared it with the interaction pattern between the protein and ligand in 5V7I, which is named 8Z1. The interaction pattern of the compound molecules with the hot spot residues was investigated in depth, as shown in Figure 9.

In the C14_1-SHMT2 system, the main contribution of the four hot spot residues is from the van der Waals energy part, with the B-chain residue TYR105 having the largest contribution. The figure shows that TYR105_B has a strong π - π interaction with the benzene ring of the ligand molecule, while the A-chain residue LEU166 has a significant π -alkyl interaction with the benzene ring of the ligand molecule, but the distance to the ligand

molecule is greater compared to residue TYR105_B, and the π -alkyl interaction is also weaker than the π - π interaction. These findings are consistent with the calculated energy results. Additionally, van der Waals interactions exist between the hot spot residues ARG425_A and TYR106_B and the ligand molecule, with the A-chain residue ARG425 being closer to the ligand and having a stronger effect. This is why the contribution of ARG425_A in the calculated results is larger than that of TYR106_B, but less than the energy contribution of TYR105_B and LEU166_A.

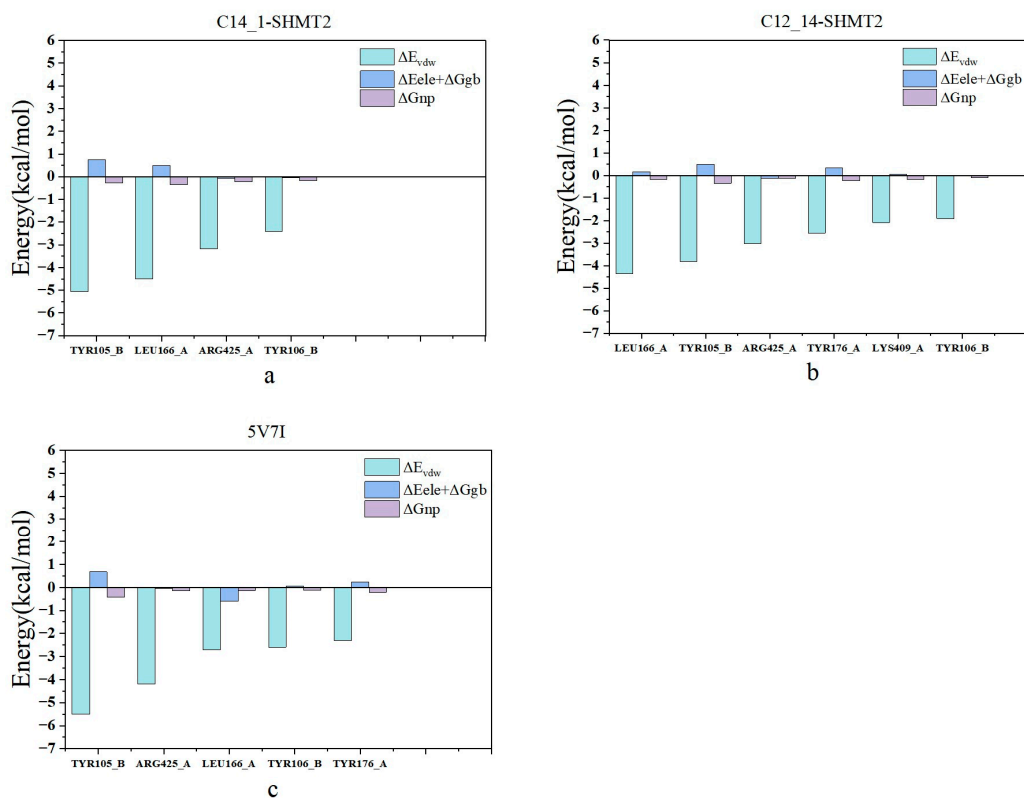


Figure 8. Different types of energy contributions from hot spot residues: (a) C14_1-SHMT2, (b) C12_14-SHMT2, and (c) 5V7I.

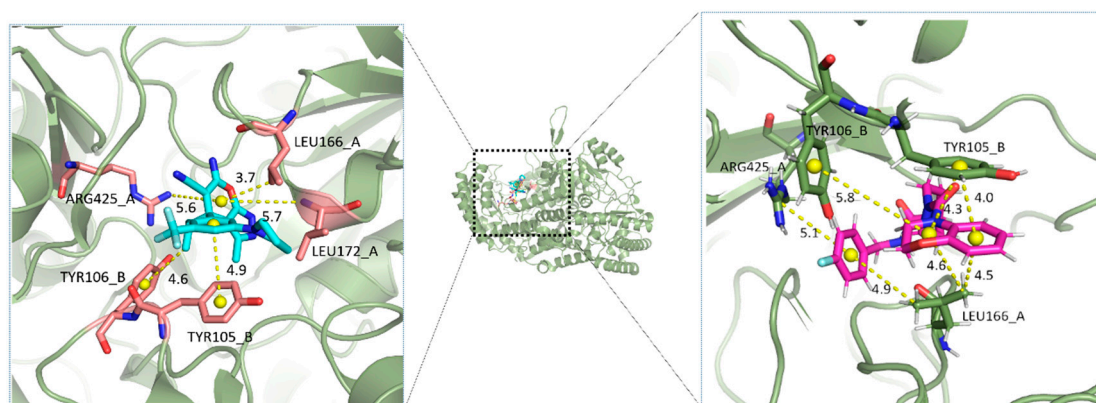


Figure 9. Ligand-residues interactions. The left panel represents the interactions between the ligand and surrounding residues in the 5V7I. The right panel represents the interactions between compound C14_1 and SHMT2.

To verify the accuracy and reliability of the binding mechanism analysis, the interaction between the ligand molecule and the protein in 5V7I was analyzed. In the crystal structure of SHMT2, there is a strong π - π interaction between the B-chain residue TYR105 and the

ligand molecule, and the other three residues near the binding pocket form a strong π -alkyl interaction with the ligand molecule, which shows that the van der Waals energy is the part that makes the main contribution to the binding interaction, which is consistent with the conclusion based on the calculation results.

Based on the predictions of hot spot residues and the analysis of spatial conformation, we found that four hot spot residues, namely, TYR105_B, LEU166_A, ARG425_A, and TYR106_B, are crucial for the formation of a stable complex between the compound and SHMT2 protein. During the ligand–acceptor binding process, the dominant contribution to the binding comes from the van der Waals energy fraction, primarily due to the formation of a large number of π - π interactions and π -alkyl interactions between the benzene ring of the molecule and the surrounding residues.

3. Method

The flow chart is shown in Figure 10. In this study, we aimed to screen a large number of molecules in the ChemDiv database, which initially contained 1.6 million molecules. To reduce the number of molecules and increase the likelihood of finding drug-like compounds, we conducted a preliminary screening based on the “Five Principles of Drug-like Drugs”. In order to refine the selection of potential inhibitor molecules, we utilized seven different scoring functions, including GlideScore [32–37], ad4_scoring, dkoes_scoring, vina, vinardo (these four functions are in the Smina molecular docking program [38], Delta-Vina [39,40], and AA-Score [41], to perform consensus scoring on the top 10,000 compounds obtained using GlideScore. From these results, we selected 1254 molecules that were ranked in top 5000 across all seven scoring methods.

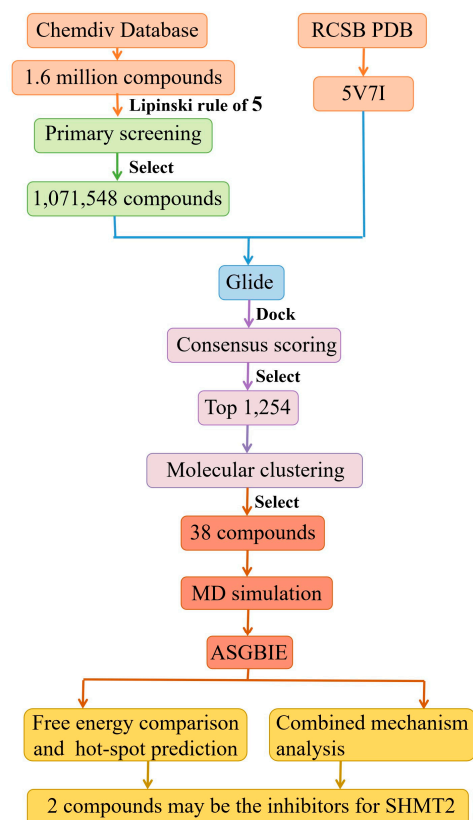


Figure 10. Research flow chart.

The selection of potential drugs is a critical step in drug discovery. Given that the properties of molecules are closely related to their structures, we utilized the Butina algorithm to cluster 1254 molecules based on their structural similarities. By doing so, we were able to efficiently and accurately screen and ultimately identify 38 potential drug candidates. To

accurately determine the binding strength of 38 potential inhibitor molecules to SHMT2, we conducted molecular dynamics simulations [42,43] of 38 protein–ligand complex systems. Using the ASGBIE method, we calculated the binding free energy of each system and analyzed the binding mechanism of each molecule.

3.1. Molecular Docking

We obtained 1.6 million compounds from ChemDiv (<http://www.chemdiv.com> (accessed on 24 May 2022)), which were then screened using the Lipinski's rule of five, which is a set of guidelines used in drug discovery and designed to evaluate the drug-likeness of chemical compounds. For the primary screening, we established the following criteria 1. molecular weight greater than 250 and less than 500; 2. hydrogen bond donor count less than 5; 3. hydrogen bond acceptor count less than 10; 4. lipophilicity coefficient less than 5; and 5. number of rotatable bonds not exceeding 10. As a result of this primary screening, we obtained 1,071,548 molecules that met the necessary criteria. Next, we optimized the compounds according to various pH environments using the LigPrep function of the Glide module. For the docking process, we utilized the crystal structure of the SHMT2 obtained from the PDB database (PDB ID: 5V7I [14]). The SHMT2 comprises two chains, A and B, as illustrated in Figure 11.

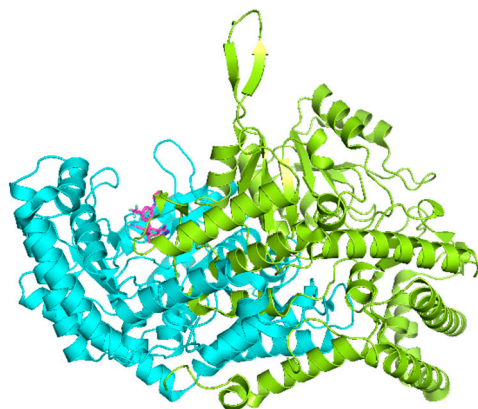


Figure 11. Cocrystal structure of 5V7I. The SHMT2 consists of two chains, A (green) and B (blue), and the ligand molecule (pink) is bound to SHMT2.

To complete the missing hydrogen atoms of the protein, we employed Schrödinger. The force field of OPLS-2005 was set to a pH parameter of 7.4 to optimize the structure and minimize the energy of the protein. A protein docking box was constructed with a central coordinate based on the location of the small molecule near the 5V7I binding pocket, and the box radius was set to 30 Å. Next, molecular docking was performed using the ligand docking operation of the Glide module, and the docking of compounds from the ChemDiv database to SHMT2 was performed with SP precision.

3.2. Consensus Scoring

The Glide module was used to dock proteins with the ligands of the compound library, and the GlideScore [44] ranking was obtained. The top 10,000 scores were used to select the best conformations of the molecules, and the structures of the complexes after Glide docking were scored by consensus using other six different scoring functions, namely, ad4_scoring [45], dkoes_scoring [38], vina [46], vinardo [47], Δ vinaXGB [39], and AA-Score [41] (developed by our research group), and the molecules were ranked according to their scores.

3.3. Molecular Clustering

To begin with, the top 5000 molecules are selected from the consensus scoring ranking. These 1254 compounds are then skeleton-clustered using the Butina algorithm [48], which

employs molecular fingerprints and Tanimoto coefficients to rapidly group the dataset into clusters based on structural similarities among the compound molecules. When the number of molecules in the compound library is less than 100,000, the Butina clustering method can be implemented with the aid of Rdkit.

The clustering process involves the following steps: 1. defining the function for calculating the molecular fingerprint, 2. reading the compound molecules from the library, 3. computing the molecular fingerprint and clustering based on the cutoff value, and 4. obtaining the resulting clusters. The cutoff represents the threshold for Tanimoto similarity, where compounds with a Tanimoto distance greater than the cutoff are classified into different clusters. Tanimoto similarity is expressed using the equation:

$$T_C(A, B) = \frac{N_C}{N_A + N_B - N_C} \quad (1)$$

In the formula, N_A denotes the number of features inherent to molecule A, N_B the number of features inherent to molecule B, and N_C represents the tally of features shared between molecules A and B. The selected cutoff should result in a reduced number of clusters containing individual elements after clustering, with a cluster composition characterized by a modest distribution of molecular quantities and a relatively smooth trend in variation. In this study, a cutoff value of 0.4 was set during molecular clustering. As a result, the clustering analysis yielded clusters where molecules within each cluster shared the same skeletal structure. However, there were significant differences in the skeletal structures between molecules that were assigned to different clusters.

3.4. Molecular Dynamics Simulation

The GAFF force field [49] was utilized to obtain force field parameters for the 22 compounds. Gaussian16 and Antechamber in Amber18 were used to calculate RESP charges [50–54] for 22 ligands. The protonation of the protein–ligand complex was conducted at pH 7.4 to maintain consistency with experimental conditions. The nonstandard protein residue LLP parameters were prepared in accordance with the Amber tutorial (<http://ambermd.org/tutorials/basic/tutorial5/> (accessed on 24 May 2022)). Finally, molecular dynamics (MD) simulations were performed using the ff14SB force field [55] and Amber18 [56,57]. To solvate the protein–ligand complex, TIP3P water was used to fill a cubic box extending 10 Å away from the protein, and the cutoff for nonbonded interactions was set to 10 Å. To neutralize the systems, counterions were added. Each system underwent a minimization step of 4000 steps and was heated from 0 to 300 K in 1 ns with Langevin dynamics, followed by a 10 ns molecular dynamics (MD) simulation in the NPT ensemble using a time step of 2 fs. To constrain bonds involving hydrogen atoms, the SHAKE algorithm [58] was employed. To compare the impact of time duration on simulation results, we chose two complex systems and performed molecular dynamics simulations of 100 ns duration on each.

To determine whether the trajectory has achieved stability, it is necessary to calculate the RMSD. The calculation of the RMSD values for the complex is performed by aligning the protein's backbone atoms and all atoms of the small molecule, with the initial docked conformation as the reference. A total of 10,000 frames from the trajectory are sampled, with every tenth frame selected for comparison against the initial conformation. For calculating the ligand RMSD values, all atoms of the ligand are aligned with the initial structure, and the difference between the two structures is quantified.

3.5. Binding Free Energy Calculation

The computational alanine scanning approach assumes that the mutated alanine residue contributes insignificantly to the binding free energy. Therefore, the difference in free energy before and after alanine mutation provides a quantitative measure of the contribution of the specific residue to the total binding free energy. Thus, the difference in

binding free energy resulting from the mutation of residue x to a (alanine) can be defined as follows:

$$\Delta\Delta G_{bind}^{x\rightarrow a} = \Delta G_{bind}^a - \Delta G_{bind}^x = \Delta\Delta G_{gas}^{x\rightarrow a} + \Delta\Delta G_{sol}^{x\rightarrow a} \quad (2)$$

The binding free energy for both the gas phase and solvation components was evaluated using the following scheme:

$$\Delta\Delta G_{gas}^{x\rightarrow a} = \Delta G_{gas}^a - \Delta G_{gas}^x \quad (3)$$

$$\Delta\Delta G_{sol}^{x\rightarrow a} = \Delta G_{sol}^a - \Delta G_{sol}^x \quad (4)$$

To determine the gas phase components of the binding free energy between and residues a and x , we used the notation ΔG_{gas}^a and ΔG_{gas}^x . We applied the alanine scanning approach to all pocket residues located within 6 Å of any ligand atoms to obtain residue-specific contributions to the total binding energy.

We evaluated the gas phase free energy between the residue and the ligand by summing the enthalpy and entropy. The enthalpy was calculated using the standard molecular mechanics method, while the entropy was calculated using the IE (interaction energy) method [30,59–62].

$$\Delta G_{gas}^x = \langle E_{int}^x \rangle - T\Delta S_{int}^x = \langle E_{int}^x \rangle + K T \ln \langle e^{\beta\Delta E_{int}^x} \rangle \quad (5)$$

and similarly, for the mutant,

$$\Delta G_{gas}^a = \langle E_{int}^a \rangle + K T \ln \langle e^{\beta\Delta E_{int}^a} \rangle \quad (6)$$

Here, E_{int}^x and E_{int}^a represent the interaction energies, including van der Waals and electrostatic energies, between the ligand and residue x and a , respectively. The exponential argument β is equal to $1/KT$, and $\Delta E_{int}^{x/a}$ is the deviation from the average $E_{int}^{x/a}$. The exponential average was evaluated using discrete time averaging:

$$\langle e^{\beta\Delta E_{int}^x} \rangle = \frac{1}{N} \sum_{i=1}^N e^{\beta\Delta E_{int}^x(t_i)} \quad (7)$$

where N is the number of MD snapshots. Finally, Equation (3) becomes the following:

$$\Delta\Delta G_{gas}^{x\rightarrow a} = \Delta\Delta E_{gas}^{x\rightarrow a} - T\Delta\Delta S_{gas}^{x\rightarrow a} = \langle E_{int}^a \rangle - \langle E_{int}^x \rangle + K T \left[\ln \langle e^{\beta\Delta E_{int}^a} \rangle - \ln \langle e^{\beta\Delta E_{int}^x} \rangle \right] \quad (8)$$

In Equation (4), the solvation energy was calculated using the OBC GBSA model (igb = 2 in Amber) [63–65], with a dielectric constant of 1, 3, and 5 for nonpolar, polar, and charged residues, respectively [65,66]. By combining the alanine scanning and IE methods described above, the total protein–ligand binding free energy can be approximated by the summation [26,67]:

$$\Delta G_{bind} = - \sum_x \Delta\Delta G_{bind}^{x\rightarrow a} \quad (9)$$

where the summation is performed with all contributing residues (6 Å around the ligand). Although Equation (9) is not theoretically rigorous, it provides a good estimate of the total binding free energy, especially when relative binding free energies are important for a series of compounds targeting the same receptor. In this study, a total of 1000 snapshots from the last 5 ns of the MD trajectory were used to calculate the binding energy using the ASGBIE method [26–28,60].

For comparison, we also calculated the binding free energy using the conventional MM/GBSA method with the MMPBSA.py script [68], and the igb parameter was set to twp. We set the dielectric constant to one in the MM/GBSA calculation, as this is typical

in standard MM/GBSA calculations. A total of 100 snapshots were used to calculate the MM/GBSA energy. Considering the computational cost, we used 10 snapshots to perform the normal mode calculation for each complex.

4. Conclusions and Discussion

To identify the potential inhibitor molecules of SHMT2, we employed seven scoring functions for consensus scoring in the ChemDiv database. Subsequently, we performed skeleton clustering on the compounds obtained by consensus scoring, resulting in a list of 38 drugs with potential inhibitory activity. Following this, we conducted molecular dynamics simulations on these 38 complex systems and calculated the binding free energy using the alanine scanning binding interaction entropy method. Based on the strength of their binding affinity, we ranked the systems and selected the top two systems for the further analysis of their hot spots.

After identifying the hot spots of each system, we studied the similarities and differences between them, as well as their mechanisms of action with molecules. Based on our research, we have arrived at the following conclusions.

We used the ASGBIE method to calculate the binding free energy of the 38 screened compounds. The results showed that the binding strength of C14_1 and C12_14 to SHMT2 ranked among the top two, and the complex system formed by these two compound molecules and SHMT2 remained stable during kinetic simulation. By analyzing and comparing the binding affinity of each system, we found that these two compounds exhibited the highest binding affinity to SHMT2.

Our analysis revealed that the amino acid residues TYR105_B, LEU166_A, ARG425_A, and TYR106_B, located near the SHMT2 binding pocket, were consistently identified as hot spots in both complex systems. These hot spots were found to contribute less than -2 kcal/mol to the binding free energy in each of the different systems.

We conducted an analysis of the binding mechanism of three complex systems and found a large number of van der Waals interactions between the compounds and hot spot residues, including π - π interactions and π -alkyl interactions. These findings suggest that the presence of benzene rings is favorable for the binding of molecules to proteins, with the van der Waals energy contribution having the most significant impact on the binding free energy.

When comparing the interaction of molecules and residues near the binding pocket in the crystal structure of SHMT2, we found that the hot spots that played the main role were TYR105_B, TYR106_B, LEU166_A, and ARG425_A. This indicates that these four hot spots played a crucial role in the binding process and were the key residues that enabled the compound to bind to SHMT2.

In conclusion, this study provides valuable insights into the binding mechanism of SHMT2 and its inhibitor molecules. By analyzing the hot spots and binding free energy contributions of key residues, we have identified the crucial role of TYR105_B, TYR106_B, LEU166_A, and ARG425_A in the binding process. The study also highlights the importance of van der Waals interactions in ligand-receptor protein binding and suggests that clustering and consensus scoring methods can be employed in screening SHMT2 inhibitor drugs. Overall, these findings can guide the modification and design of SHMT2 inhibitor molecules and potentially aid in the development of cancer treatments.

However, the present study does lack the experimental validation of the computational predictions and has not addressed the mutational effect as well as the selectivity issue of the compounds between SHMT1 and SHMT2. We hope to address these issues in our future study.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biophysica3040044/s1>, Table S1. The structures of 38 compounds screened. Table S2. Compound rankings among the seven scoring functions. Table S3. Specific values of binding free energy for 38 compounds and 8Z1 to SHMT2.

Author Contributions: Formal analysis & investigation, J.Z.H.Z.; writing—original draft preparation, B.C.; writing—review and editing, B.C. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the National Natural Science Foundation of China (21933010, 22250710136).

Data Availability Statement: Data are contained within the article.

Acknowledgments: This work was supported by the Shanghai Engineering Research Center of Molecular Therapeutics and New Drug Development R&D Program. This work was also supported in part through the NYU IT High Performance Computing resources, services, and staff expertise (22333006). We thank the NYU Shanghai HPC facility for providing computational resources.

Conflicts of Interest: The authors declare no competing financial interests.

References

1. Guertin, D.A.; Wellen, K.E. Acetyl-CoA metabolism in cancer. *Nat. Rev. Cancer* **2023**, *23*, 156–172. [[CrossRef](#)] [[PubMed](#)]
2. Locasale, J.W. Serine, glycine and one-carbon units: Cancer metabolism in full circle. *Nat. Rev. Cancer* **2013**, *13*, 572–583. [[CrossRef](#)] [[PubMed](#)]
3. Peng, H.; Wang, Y.; Luo, W. Multifaceted role of branched-chain amino acid metabolism in cancer. *Oncogene* **2020**, *39*, 6747–6756. [[CrossRef](#)] [[PubMed](#)]
4. Asai, A.; Konno, M.; Koseki, J.; Taniguchi, M.; Vecchione, A.; Ishii, H. One-carbon metabolism for cancer diagnostic and therapeutic approaches. *Cancer Lett.* **2020**, *470*, 141–148. [[CrossRef](#)] [[PubMed](#)]
5. Clare, C.E.; Brassington, A.H.; Kwong, W.Y.; Sinclair, K.D. One-Carbon Metabolism: Linking Nutritional Biochemistry to Epigenetic Programming of Long-Term Development. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 263–287. [[CrossRef](#)]
6. Cuthbertson, C.R.; Arabzada, Z.; Bankhead, A., 3rd; Kyani, A.; Neamati, N. A Review of Small-Molecule Inhibitors of One-Carbon Enzymes: SHMT2 and MTHFD2 in the Spotlight. *ACS Pharmacol. Transl. Sci.* **2021**, *4*, 624–646. [[CrossRef](#)] [[PubMed](#)]
7. Ducker, G.S.; Rabinowitz, J.D. One-Carbon Metabolism in Health and Disease. *Cell Metab.* **2017**, *25*, 27–42. [[CrossRef](#)]
8. Kato, Y.; Maeda, T.; Suzuki, A.; Baba, Y. Cancer metabolism: New insights into classic characteristics. *Jpn. Dent. Sci. Rev.* **2018**, *54*, 8–21. [[CrossRef](#)]
9. Newman, A.C.; Maddocks, O.D.K. One-carbon metabolism in cancer. *Br. J. Cancer* **2017**, *116*, 1499–1504. [[CrossRef](#)]
10. Tong, J.; Krieger, J.R.; Taylor, P.; Bagshaw, R.; Kang, J.; Jeedigunta, S.; Wybenga-Groot, L.E.; Zhang, W.; Badr, H.; Mirhadi, S.; et al. Cancer proteome and metabolite changes linked to SHMT2. *PLoS ONE* **2020**, *15*, e0237981. [[CrossRef](#)]
11. Ji, L.; Tang, Y.; Pang, X.; Zhang, Y. Increased Expression of Serine Hydroxymethyltransferase 2 (SHMT2) is a Negative Prognostic Marker in Patients with Hepatocellular Carcinoma and is Associated with Proliferation of HepG2 Cells. *Med. Sci. Monit.* **2019**, *25*, 5823–5832. [[CrossRef](#)] [[PubMed](#)]
12. Wilke, A.C.; Doebele, C.; Zindel, A.; Lee, K.S.; Rieke, S.A.; Ceribelli, M.; Comoglio, F.; Phelan, J.D.; Wang, J.Q.; Pikman, Y.; et al. SHMT2 inhibition disrupts the TCF3 transcriptional survival program in Burkitt lymphoma. *Blood* **2022**, *139*, 538–553. [[CrossRef](#)] [[PubMed](#)]
13. Zhang, P.; Yang, Q. Overexpression of SHMT2 Predicts a Poor Prognosis and Promotes Tumor Cell Growth in Bladder Cancer. *Front. Genet.* **2021**, *12*, 682856. [[CrossRef](#)] [[PubMed](#)]
14. Ducker, G.S.; Ghergurovich, J.M.; Mainolfi, N.; Suri, V.; Jeong, S.K.; Hsin-Jung Li, S.; Friedman, A.; Manfredi, M.G.; Gitai, Z.; Kim, H.; et al. Human SHMT inhibitors reveal defective glycine import as a targetable metabolic vulnerability of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11404–11409. [[CrossRef](#)]
15. Wang, B.; Wang, W.; Zhu, Z.; Zhang, X.; Tang, F.; Wang, D.; Liu, X.; Yan, X.; Zhuang, H. Mitochondrial serine hydroxymethyltransferase 2 is a potential diagnostic and prognostic biomarker for human glioma. *Clin. Neurol. Neurosurg.* **2017**, *154*, 28–33. [[CrossRef](#)] [[PubMed](#)]
16. Woo, C.C.; Chen, W.C.; Teo, X.Q.; Radda, G.K.; Lee, P.T. Downregulating serine hydroxymethyltransferase 2 (SHMT2) suppresses tumorigenesis in human hepatocellular carcinoma. *Oncotarget* **2016**, *7*, 53005–53017. [[CrossRef](#)] [[PubMed](#)]
17. Tennenhouse, A.; Khmelnitsky, L.; Khalaila, R.; Yeshaya, N.; Noronha, A.; Lindzen, M.; Makowski, E.; Zaretsky, I.; Sirkis, Y.F.; Galon-Wolfenson, Y. Reliable energy-based antibody humanization and stabilization. *bioRxiv* **2022**. [[CrossRef](#)]
18. Witschel, M.; Stelzer, F.; Hutzler, J.; Qu, T.; Mietzner, T.; Kreuz, K.; Grossmann, K.; Aponte, R.; Hoeffken, H.W.; Calo, F. Pyrazolopyrans Having Herbicidal and Pharmaceutical Properties. U.S. Patent US9480259B2, 1 November 2016.
19. Marani, M.; Paone, A.; Fiascarelli, A.; Macone, A.; Gargano, M.; Rinaldo, S.; Giardina, G.; Pontecorvi, V.; Koes, D.; McDermott, L. A pyrazolopyran derivative preferentially inhibits the activity of human cytosolic serine hydroxymethyltransferase and induces cell death in lung cancer cells. *Oncotarget* **2016**, *7*, 4570. [[CrossRef](#)]
20. Zhao, R.; Goldman, I.D. Resistance to antifolates. *Oncogene* **2003**, *22*, 7431–7457. [[CrossRef](#)]
21. Stover, P.; Schirch, V. 5-Formyltetrahydrofolate polyglutamates are slow tight binding inhibitors of serine hydroxymethyltransferase. *J. Biol. Chem.* **1991**, *266*, 1543–1550. [[CrossRef](#)]
22. Marshall, G.R. Computer-aided drug design. *Annu. Rev. Pharmacol. Toxicol.* **1987**, *27*, 193–213. [[CrossRef](#)] [[PubMed](#)]
23. Lyne, P.D. Structure-based virtual screening: An overview. *Drug Discov. Today* **2002**, *7*, 1047–1055. [[CrossRef](#)] [[PubMed](#)]

24. Walters, W.P.; Stahl, M.T.; Murcko, M.A. Virtual screening—An overview. *Drug Discov. Today* **1998**, *3*, 160–178. [[CrossRef](#)]
25. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z.; Hou, T. End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem. Rev.* **2019**, *119*, 9478–9508. [[CrossRef](#)] [[PubMed](#)]
26. Liu, X.; Peng, L.; Zhou, Y.; Zhang, Y.; Zhang, J.Z. Computational alanine scanning with interaction entropy for protein–ligand binding free energies. *J. Chem. Theory Comput.* **2018**, *14*, 1772–1780. [[CrossRef](#)] [[PubMed](#)]
27. Liu, X.; Peng, L.; Zhang, J.Z. Accurate and efficient calculation of protein–protein binding free energy–interaction entropy with residue type-specific dielectric constants. *J. Chem. Inf. Model.* **2018**, *59*, 272–281. [[CrossRef](#)] [[PubMed](#)]
28. He, L.; Bao, J.; Yang, Y.; Dong, S.; Zhang, L.; Qi, Y.; Zhang, J.Z. Study of SHMT2 inhibitors and their binding mechanism by computational alanine scanning. *J. Chem. Inf. Model.* **2019**, *59*, 3871–3878. [[CrossRef](#)]
29. Moreira, I.S.; Fernandes, P.A.; Ramos, M.J. Computational alanine scanning mutagenesis—An improved methodological approach. *J. Comput. Chem.* **2007**, *28*, 644–654. [[CrossRef](#)]
30. Duan, L.; Liu, X.; Zhang, J.Z. Interaction entropy: A new paradigm for highly efficient and reliable computation of protein–ligand binding free energy. *J. Am. Chem. Soc.* **2016**, *138*, 5722–5728. [[CrossRef](#)]
31. Han, Y.; He, L.; Qi, Y.; Zhao, Y.; Pan, Y.; Fang, B.; Li, S.; Zhang, J.Z.; Zhang, L. Identification of three new compounds that directly target human serine hydroxymethyltransferase 2. *Chem. Biol. Drug Des.* **2021**, *97*, 221–230. [[CrossRef](#)]
32. Chaput, L.; Mouawad, L. Efficient conformational sampling and weak scoring in docking programs? Strategy of the wisdom of crowds. *J. Cheminform.* **2017**, *9*, 37. [[CrossRef](#)] [[PubMed](#)]
33. Rao, S.; Sanschagrin, P.C.; Greenwood, J.R.; Repasky, M.P.; Sherman, W.; Farid, R. Improving database enrichment through ensemble docking. *J. Comput. Aided Mol. Des.* **2008**, *22*, 621–627. [[CrossRef](#)]
34. Vass, M.; Tarcsay, A.; Keseru, G.M. Multiple ligand docking by Glide: Implications for virtual second-site screening. *J. Comput. Aided Mol. Des.* **2012**, *26*, 821–834. [[CrossRef](#)] [[PubMed](#)]
35. Xu, D.; Meroueh, S.O. Effect of Binding Pose and Modeled Structures on SVMGen and GlideScore Enrichment of Chemical Libraries. *J. Chem. Inf. Model.* **2016**, *56*, 1139–1151. [[CrossRef](#)] [[PubMed](#)]
36. Repasky, M.P.; Shelley, M.; Friesner, R.A. Flexible ligand docking with Glide. *Curr. Protoc. Bioinform.* **2007**, *8*, 8–12. [[CrossRef](#)]
37. Corbeil, C.R.; Englebienne, P.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 1. Development and validation of FITTED 1.0. *J. Chem. Inf. Model.* **2007**, *47*, 435–449. [[CrossRef](#)]
38. Koes, D.R.; Baumgartner, M.P.; Camacho, C.J. Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1893–1904. [[CrossRef](#)]
39. Lu, J.; Hou, X.; Wang, C.; Zhang, Y. Incorporating Explicit Water Molecules and Ligand Conformation Stability in Machine-Learning Scoring Functions. *J. Chem. Inf. Model.* **2019**, *59*, 4540–4549. [[CrossRef](#)]
40. Wang, C.; Zhang, Y. Improving scoring–docking–screening powers of protein–ligand scoring functions using random forest. *J. Comput. Chem.* **2017**, *38*, 169–177. [[CrossRef](#)]
41. Pan, X.; Wang, H.; Zhang, Y.; Wang, X.; Li, C.; Ji, C.; Zhang, J.Z.H. AA-Score: A New Scoring Function Based on Amino Acid-Specific Interaction for Molecular Docking. *J. Chem. Inf. Model.* **2022**, *62*, 2499–2509. [[CrossRef](#)]
42. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19. [[CrossRef](#)]
43. Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169. [[CrossRef](#)] [[PubMed](#)]
44. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749. [[CrossRef](#)]
45. Huey, R.; Morris, G.M.; Olson, A.J.; Goodsell, D.S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152. [[CrossRef](#)] [[PubMed](#)]
46. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)]
47. Quiroga, R.; Villarreal, M.A. Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PLoS ONE* **2016**, *11*, e0155183. [[CrossRef](#)] [[PubMed](#)]
48. Butina, D. Unsupervised data base clustering based on daylight’s fingerprint and Tanimoto similarity: A fast and automated way to cluster small and large data sets. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 747–750. [[CrossRef](#)]
49. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)]
50. Dupradeau, F.Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. The R.E.D. tools: Advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821–7839. [[CrossRef](#)]
51. Janeczek, M.; Kuhrova, P.; Mlynsky, V.; Otyepka, M.; Sponer, J.; Banas, P. W-RESP: Well-Restrained Electrostatic Potential-Derived Charges. Revisiting the Charge Derivation Model. *J. Chem. Theory Comput.* **2021**, *17*, 3495–3509. [[CrossRef](#)]
52. Okiyama, Y.; Watanabe, H.; Fukuzawa, K.; Nakano, T.; Mochizuki, Y.; Ishikawa, T.; Ebina, K.; Tanaka, S. Application of the fragment molecular orbital method for determination of atomic charges on polypeptides. II. Towards an improvement of force fields used for classical molecular dynamics simulations. *Chem. Phys. Lett.* **2009**, *467*, 417–423. [[CrossRef](#)]

53. Shul'ga, D.A.; Oliferenko, A.A.; Pisarev, S.A.; Palyulin, V.A.; Zefirov, N.S. Modeling of atomic RESP charges with the use of topological calculation schemes. *Dokl. Chem.* **2006**, *408*, 76–79. [[CrossRef](#)]
54. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Kollman, P.A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **2002**, *115*, 9620–9631. [[CrossRef](#)]
55. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [[CrossRef](#)] [[PubMed](#)]
56. Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, T.E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41. [[CrossRef](#)]
57. Case, D.A.; Cheatham, T.E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [[CrossRef](#)] [[PubMed](#)]
58. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34. [[CrossRef](#)]
59. Qiu, L.; Yan, Y.; Sun, Z.; Song, J.; Zhang, J.Z. Interaction entropy for computational alanine scanning in protein–protein binding. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1342. [[CrossRef](#)]
60. Yan, Y.; Yang, M.; Ji, C.G.; Zhang, J.Z. Interaction entropy for computational alanine scanning. *J. Chem. Inf. Model.* **2017**, *57*, 1112–1122. [[CrossRef](#)]
61. Song, J.; Qiu, L.; Zhang, J.Z. An efficient method for computing excess free energy of liquid. *Sci. China Chem.* **2018**, *61*, 135–140. [[CrossRef](#)]
62. Ben-Shalom, I.Y.; Pfeiffer-Marek, S.; Baringhaus, K.-H.; Gohlke, H. Efficient approximation of ligand rotational and translational entropy changes upon binding for use in MM-PBSA calculations. *J. Chem. Inf. Model.* **2017**, *57*, 170–189. [[CrossRef](#)]
63. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [[CrossRef](#)]
64. Massova, I.; Kollman, P.A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discov. Des.* **2000**, *18*, 113–135. [[CrossRef](#)]
65. Onufriev, A.; Bashford, D.; Case, D.A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins Struct. Funct. Bioinform.* **2004**, *55*, 383–394. [[CrossRef](#)]
66. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82. [[CrossRef](#)] [[PubMed](#)]
67. Zhou, Y.; Liu, X.; Zhang, Y.; Peng, L.; Zhang, J.Z. Residue-specific free energy analysis in ligand bindings to JAK2. *Mol. Phys.* **2018**, *116*, 2633–2641. [[CrossRef](#)]
68. Miller, B.R., III; McGee, T.D., Jr.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA. py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.