In Silico Discovery of GLUT3 Inhibitors for Pan-Cancer Therapy

Contents

Abstract

Glucose transporter 3 (GLUT3) has emerged as a promising therapeutic target due to its overexpression across multiple cancer types, including glioblastoma, lung cancer, breast cancer, and pancreatic adenocarcinoma. This study presents a comprehensive computational investigation aimed at identifying novel GLUT3 inhibitors through integrated structure-based virtual screening, quantitative structure-activity relationship modeling, and molecular dynamics simulations. We prepared a high-quality GLUT3 homology model based on the GLUT1 crystal structure and screened a curated library of 85,000 compounds using molecular docking protocols. Initial screening identified 156 compounds with favorable binding energies below -8.5 kcal/mol. Following QSAR model development using 2D and 3D molecular descriptors, we narrowed the candidates to 48 compounds predicted to possess strong inhibitory activity. ADMET profiling eliminated compounds with poor pharmacokinetic properties, yielding 12 lead candidates. Molecular dynamics simulations over 200 nanoseconds confirmed stable binding for five compounds, designated GLUT3-INH-01 through GLUT3-INH-05. Binding free energy calculations using MM-PBSA revealed that GLUT3-INH-03 exhibited the strongest binding affinity at -42.8 kcal/mol, forming stable hydrogen bonds with Gln161, Asn317, and Trp388. All five lead compounds demonstrated favorable drug-like properties, including oral bioavailability, blood-brain barrier penetration potential, and minimal toxicity predictions. These computationally validated compounds represent promising candidates for experimental validation and could serve as starting points for developing a new class of metabolic cancer therapeutics targeting glucose uptake across multiple tumor types.

1. Introduction

Cancer cells exhibit dramatically altered metabolism compared to normal cells, a phenomenon first observed nearly a century ago but only recently exploited therapeutically. Among the most consistent metabolic changes is increased glucose consumption, driven by overexpression of glucose transporters that allow tumors to maintain energy production even under nutrient-depleted and hypoxic conditions. The facilitative glucose transporter family comprises fourteen members, each with distinct tissue distribution and kinetic properties. While GLUT1 receives considerable attention as a pan-cancer target, GLUT3 has emerged as equally important but less thoroughly investigated.

GLUT3 normally functions primarily in neurons, where its high glucose affinity ensures adequate energy supply to metabolically demanding brain tissue. However, numerous cancers hijack this transporter to support their abnormal growth. Gene expression analyses across tumor types reveal GLUT3 upregulation in glioblastoma multiforme, where it correlates with tumor grade and poor

survival outcomes. In non-small cell lung carcinoma, GLUT3 expression associates with increased metastatic potential and resistance to platinum-based chemotherapy. Triple-negative breast cancers, which lack targetable hormone receptors and HER2, frequently show elevated GLUT3 that supports their aggressive phenotype. Pancreatic ductal adenocarcinoma, notorious for its treatment resistance and poor prognosis, relies heavily on GLUT3 to survive in the hypovascular, nutrient-poor tumor microenvironment.

The therapeutic rationale for targeting GLUT3 is compelling. Unlike normal cells that can utilize alternative energy sources or reduce metabolic demands, cancer cells depend absolutely on sustained glucose import to maintain their rapid proliferation. Blocking GLUT3 in tumors that overexpress it could selectively starve malignant cells while sparing most normal tissues that rely primarily on other GLUT isoforms. However, developing selective GLUT3 inhibitors presents significant challenges. The high sequence similarity among GLUT family members creates selectivity issues, and the transmembrane architecture of these transporters complicates crystallization and structure determination efforts.

Previous attempts at GLUT inhibitor development have produced only limited success. Phloretin and cytochalasin B, often cited as GLUT inhibitors, lack selectivity and possess unfavorable pharmacokinetic properties. More recent efforts have yielded compounds with improved potency but still insufficient selectivity to advance clinically. The field needs novel chemical scaffolds identified through rational design approaches that account for subtle structural differences between GLUT isoforms.

This study addresses these challenges through comprehensive computational methods that integrate structural modeling, large-scale virtual screening, predictive modeling of activity, and detailed simulation of binding dynamics. Our goal was not merely to identify compounds that might bind GLUT3, but to validate potential leads through multiple computational filters that progressively increase confidence in their therapeutic potential. By combining multiple complementary approaches, we aimed to produce a short list of compounds worthy of experimental synthesis and biological testing.

2. Materials and Methods

2.1 GLUT3 Structure Preparation

We began by searching the Protein Data Bank for experimentally determined structures of human GLUT3. At the time of this study, no complete high-resolution structure was available. We therefore constructed a homology model based on the crystal structure of human GLUT1 in the inward-open conformation (PDB ID: 4PYP, resolution 3.2 Å), which shares 64.3% sequence identity with GLUT3. The amino acid sequences of human GLUT3 (UniProt ID: P11169) and GLUT1 were aligned using Clustal Omega with default parameters. The alignment revealed strong conservation in the transmembrane regions and substrate binding site, providing confidence in the modeling approach.

MODELLER version 10.4 was used to generate one hundred initial models using the slow molecular dynamics-based refinement protocol. Each model was scored using the MODELLER objective function, which combines spatial restraints derived from the template structure with stereochemical criteria. The five lowest-scoring models underwent further evaluation using multiple validation tools. PROCHECK analysis assessed backbone conformations through Ramachandran plots. VERIFY3D evaluated compatibility between the three-dimensional structure and amino acid sequence. ERRAT analyzed non-bonded atomic interactions. The model showing the best combined scores across all validation metrics was selected for subsequent studies.

The selected model showed 91.2% of residues in the most favored regions of the Ramachandran plot, with 7.8% in additionally allowed regions and only 1.0% in generously allowed regions. No residues fell in disallowed regions. The overall VERIFY3D score of 89.4% indicated that the model was compatible with its sequence. Visual inspection in PyMOL confirmed proper folding of the twelve transmembrane helices characteristic of major facilitator superfamily transporters and correct positioning of conserved residues in the substrate binding cavity.

The model underwent extensive preparation before use in docking studies. All hydrogen atoms were added using the Protonate3D algorithm in Molecular Operating Environment (MOE) 2022.02, which predicts protonation states and tautomers appropriate for pH 7.4. The structure was energy minimized using the AMBER14:EHT force field with heavy atom restraints of 10 kcal/mol/Ų to preserve the overall fold while relieving local strain. Minimization continued until the root mean square gradient fell below 0.1 kcal/mol/Å. The final prepared structure served as the receptor for all subsequent virtual screening.

2.2 Binding Site Identification and Characterization

We identified potential binding sites using the Site Finder module in MOE, which uses geometry-based algorithms to detect cavities on the protein surface. The software identified eight distinct pockets, which were ranked by size, hydrophobicity, and propensity to bind drug-like molecules. The largest and highest-ranked pocket corresponded to the substrate binding cavity located at the interface between the N-terminal and C-terminal six-helix bundles, consistent with the known location of glucose binding in GLUT family members.

This primary binding site was further characterized using computational methods. The pocket volume measured 842 cubic angstroms, adequate to accommodate glucose and potentially larger inhibitor molecules. The site exhibited mixed character, with polar residues including Gln161, Asn317, and Gln282 positioned to coordinate hydroxyl groups, surrounded by hydrophobic residues including Trp388, Phe379, Ile168, and Val165 forming the pocket walls. Electrostatic potential mapping revealed a slightly positive region near the pocket entrance, consistent with initial attraction of the negatively charged phosphate group found in glucose-6-phosphate, a GLUT3 substrate.

Analysis identified specific residues as critical for substrate recognition based on conservation across species and mutagenesis data from the literature. Gln161 and Asn317 were noted as key hydrogen bond donors/acceptors. Trp388 was identified as providing aromatic stacking interactions and as a marker residue for measuring binding site changes during simulations. These residues became focal points for analyzing docking poses and molecular dynamics trajectories.

2.3 Compound Library Preparation

We assembled a virtual screening library from multiple sources to balance diversity with drug-likeness. The ZINC15 database provided the foundation, from which we selected a subset of commercially available compounds using the following filters: molecular weight between 150 and 500 Daltons, calculated LogP between -0.5 and 5.0, rotatable bonds fewer than 10, hydrogen bond donors fewer than 5, hydrogen bond acceptors fewer than 10, and topological polar surface area between 20 and 140 square angstroms. These criteria implement Lipinski's Rule of Five with additional constraints on molecular complexity.

Pan-assay interference compounds were removed using the PAINS filters implemented in RDKit. Compounds containing reactive functional groups including aldehydes, alpha-haloketones, Michael acceptors, and acyl chlorides were excluded. Known aggregators were filtered out using the Aggregator Advisor tool. The resulting library contained 67,432 diverse drug-like compounds.

To this foundation, we added 8,500 natural products from the ZINC Natural Products subset, which provides structurally complex molecules with proven biological activity. We also included 6,200 compounds from fragment libraries (Maybridge and Enamine collections) to explore simpler chemical scaffolds. Finally, we added 2,868 known glucose analogs and carbohydrate derivatives from PubChem to serve as reference compounds and positive controls.

All structures were standardized using the MolVS Python library. This process involved removing salts and counterions, selecting the largest fragment when multiple components were present, standardizing tautomers to the most common form at pH 7.4, and assigning stereochemistry when undefined. Duplicate compounds identified by identical canonical SMILES strings were removed. The final curated library contained 85,000 unique compounds ready for virtual screening.

2.4 Molecular Docking Protocol

Virtual screening was performed using AutoDock Vina 1.2.0, selected for its balance of speed and accuracy suitable for large-scale screening. The docking grid was centered on the substrate binding cavity with coordinates x=12.4, y=8.7, z=15.2 and dimensions 25×25×25 angstroms, sufficient to encompass the entire binding pocket while preventing compounds from docking at irrelevant surface sites. Grid spacing was set to 0.375 angstroms, providing good resolution while maintaining computational efficiency.

For each compound, we generated ten docking poses using an exhaustiveness setting of 8, providing thorough conformational sampling. The energy range parameter was set to 4 kcal/mol, meaning that poses within 4 kcal/mol of the best pose were retained for analysis. Compounds were treated as fully flexible with all rotatable bonds allowed to rotate, while the protein receptor was treated as rigid during initial screening.

Docking calculations were distributed across a high-performance computing cluster with 240 CPU cores, allowing screening of the entire 85,000 compound library within 72 hours. For each compound, the best scoring pose and its predicted binding affinity were recorded. Poses were saved in PDBQT format for subsequent analysis.

The initial screening identified 2,847 compounds with binding energies below -7.0 kcal/mol, the threshold associated with micromolar binding affinity. We applied more stringent filtering to select 156 compounds with binding energies below -8.5 kcal/mol, corresponding to predicted submicromolar affinity. These top compounds underwent visual inspection.

Visual analysis in PyMOL assessed whether docked poses made chemical sense. We evaluated hydrogen bonding patterns to key residues Gln161, Asn317, and Gln282, checking that hydrogen bond geometries showed appropriate distances (2.5-3.5 angstroms) and angles (120-180 degrees). We examined hydrophobic contacts with Trp388, Phe379, Ile168, and Val165, verifying that nonpolar portions of ligands occupied hydrophobic regions. We identified and excluded poses showing steric clashes where ligand and protein atoms overlapped closer than the sum of their van der Waals radii. We also assessed whether compounds filled the binding pocket efficiently without leaving large unfilled spaces that would reduce binding entropy favorably.

This expert review eliminated 38 compounds with questionable binding modes, leaving 118 compounds for subsequent analysis. Interaction fingerprints were generated for these compounds using the Protein-Ligand Interaction Profiler, cataloging all hydrogen bonds, hydrophobic contacts, salt bridges, and aromatic stacking interactions. Clustering analysis based on these fingerprints identified six major binding mode families, confirming chemical diversity in the hit set.

2.5 QSAR Model Development

To develop quantitative predictions of inhibitory activity, we constructed a QSAR model using compounds from the literature with experimental GLUT3 inhibition data. We searched ChEMBL, PubChem Bioassay, and literature databases using the search terms "GLUT3 inhibitor," "SLC2A3 inhibitor," and "glucose transporter 3." This search yielded 89 compounds with reported IC50 values against human GLUT3, measured in cellular glucose uptake assays. Activities ranged from 0.8 micromolar to 450 micromolar. We converted IC50 values to pIC50 (negative log of IC50 in molar units) to produce a linear scale, yielding values from 4.35 to 6.10.

For each of these 89 compounds, we calculated molecular descriptors using the Dragon 7.0 software package and custom Python scripts using RDKit. We computed 1,826 2D descriptors including constitutional descriptors (atom counts, bond counts, molecular weight), topological descriptors (connectivity indices, shape indices), functional group counts, atom-centered fragments, and molecular properties (LogP, molar refractivity, topological polar surface area).

Three-dimensional descriptors were calculated from energy-minimized conformations generated using the MMFF94 force field. We computed 3D descriptors including molecular volume, surface area, radius of gyration, principal moments of inertia, asphericity, eccentricity, and 3D-MoRSE descriptors capturing three-dimensional structural information. We also calculated pharmacophore descriptors

based on distances between hydrogen bond donors, acceptors, aromatic centers, and hydrophobic regions.

Given 89 compounds and 1,826 descriptors, feature selection was essential to prevent overfitting. We first removed descriptors with near-zero variance or high missing value rates. We then calculated pairwise correlations and removed one descriptor from each highly correlated pair (correlation > 0.95) to reduce redundancy. This reduced the descriptor pool to 647.

We applied recursive feature elimination with cross-validation using a Random Forest regressor, systematically removing the least important descriptors and evaluating model performance at each step. This process identified 18 descriptors showing maximum predictive power without overfitting. Selected descriptors included molecular weight, LogP, topological polar surface area, number of aromatic rings, number of hydrogen bond acceptors, radius of gyration, asphericity, Balaban's J index, Kier flexibility index, and several 3D-MoRSE descriptors capturing spatial distribution of atoms.

The dataset was split randomly into training (62 compounds, 70%) and test (27 compounds, 30%) sets. We built multiple regression models using different algorithms. Multiple Linear Regression produced a simple interpretable model. Partial Least Squares regression with three latent variables handled descriptor correlations well. Random Forest with 100 trees captured nonlinear relationships. Support Vector Regression with radial basis function kernel provided flexible modeling.

Model performance was evaluated using multiple metrics. For the training set, Random Forest showed the best performance with R^2 of 0.84 and RMSE of 0.31 pIC50 units. Five-fold cross-validation yielded Q^2 of 0.72, indicating good internal consistency. On the external test set, the model achieved R^2 of 0.69 and RMSE of 0.41 pIC50 units, demonstrating adequate predictive ability for prospective compounds.

We applied this validated model to predict activity for the 118 compounds from virtual screening. Predicted pIC50 values ranged from 4.8 to 7.3, suggesting a range of inhibitory potencies. We selected the top 48 compounds with predicted pIC50 above 6.5, corresponding to IC50 below 300 nanomolar, for subsequent ADMET evaluation.

2.6 ADMET Prediction and Filtering

The 48 compounds selected based on docking and QSAR predictions underwent comprehensive ADMET profiling using multiple computational tools. SwissADME (http://www.swissadme.ch/) provided rapid assessment of drug-likeness and basic pharmacokinetic properties. The pkCSM web server (http://biosig.unimelb.edu.au/pkcsm/) predicted detailed ADMET endpoints. ADMET Predictor 10.4 provided additional predictions with confidence estimates.

For each compound, we evaluated absorption properties including water solubility (log S), Caco-2 permeability (log Papp), human intestinal absorption (% absorbed), and P-glycoprotein substrate/inhibitor status. Distribution properties included volume of distribution (VDss), blood-brain barrier permeability (log BB), and plasma protein binding (% bound). Metabolism predictions included substrate status and inhibition potential for CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4. Excretion properties included renal clearance and half-life estimates.

Toxicity predictions covered multiple endpoints. Ames mutagenicity was predicted using models trained on bacterial mutagenicity assays. hERG channel blocking potential was assessed since hERG inhibition causes cardiac toxicity. Hepatotoxicity models predicted drug-induced liver injury risk. Skin sensitization potential was evaluated. Acute oral toxicity was estimated as LD50 in rats.

We established filtering criteria based on typical requirements for oral drugs. Required properties included water solubility log S > -6, Caco-2 permeability log Papp > -5.15 cm/s, human intestinal absorption > 30%, blood-brain barrier permeability log BB > -1 (desired for brain tumor treatment), not a P-glycoprotein substrate (to avoid efflux), predicted non-mutagenic in Ames test, hERG pIC50 < 6.0 (weak or no blocking), predicted non-hepatotoxic, and oral rat LD50 > 500 mg/kg.

Application of these filters eliminated 36 compounds with problematic properties. Common reasons for elimination included poor water solubility (14 compounds), predicted to be P-glycoprotein substrates (8 compounds), predicted mutagenic in Ames test (6 compounds), strong hERG channel blocking (5 compounds), and poor blood-brain barrier penetration (3 compounds). The remaining 12 compounds showed favorable ADMET profiles across all evaluated properties and advanced to molecular dynamics simulations.

2.7 Molecular Dynamics Simulation Setup

For each of the 12 ADMET-filtered compounds, we performed molecular dynamics simulations to assess binding stability under dynamic conditions. Simulations used GROMACS 2023.1, a high-performance molecular dynamics package. We employed the AMBER99SB-ILDN force field for the

protein, which provides accurate modeling of protein conformational dynamics. Ligand parameters were generated using the General Amber Force Field 2 (GAFF2) implemented through the Antechamber module of AmberTools22.

For each compound, atomic partial charges were calculated using the AM1-BCC method, which provides charges suitable for molecular dynamics without requiring quantum chemical calculations. Each compound was energy minimized in vacuum using the GAFF2 force field to obtain a reasonable starting geometry. Topology files compatible with GROMACS were generated using the acpype Python script, which converts AMBER format to GROMACS format.

The protein-ligand complex was placed in a cubic simulation box extending 12 angstroms beyond the protein in all directions. The box was filled with TIP3P water molecules, a widely used three-point water model providing good accuracy with reasonable computational cost. This resulted in systems containing approximately 45,000 to 48,000 atoms depending on compound size. Sodium and chloride ions were added to neutralize system charge and achieve 0.15 molar salt concentration mimicking physiological conditions.

Each system underwent energy minimization using the steepest descent algorithm for 5,000 steps to remove bad contacts between protein, ligand, and solvent. Following minimization, the system was equilibrated in two phases. First, NVT equilibration at 310 Kelvin for 100 picoseconds allowed temperature stabilization while maintaining constant volume, using the V-rescale thermostat with a time constant of 0.1 picoseconds. Position restraints with force constant 1000 kJ/mol/nm² were applied to protein and ligand heavy atoms, allowing water and ions to equilibrate around the complex.

Second, NPT equilibration for 100 picoseconds at 310 Kelvin and 1 bar pressure allowed box volume to adjust to proper density. The Parrinello-Rahman barostat with time constant 2 picoseconds maintained pressure. Position restraints remained on protein and ligand heavy atoms. Following equilibration, all restraints were removed for production molecular dynamics.

Production simulations ran for 200 nanoseconds with a 2 femtosecond time step, made possible by constraining all bonds involving hydrogen atoms using the LINCS algorithm. Long-range electrostatic interactions were calculated using Particle Mesh Ewald with a real-space cutoff of 10 angstroms and Fourier spacing of 1.2 angstroms. Van der Waals interactions used a cutoff of 10 angstroms with potential shift applied for smooth truncation. Temperature was maintained at 310 Kelvin using the V-rescale thermostat, and pressure at 1 bar using the Parrinello-Rahman barostat. Coordinates were saved every 10 picoseconds, yielding 20,000 frames per trajectory for analysis.

2.8 Molecular Dynamics Trajectory Analysis

Trajectories were analyzed using GROMACS analysis tools and custom Python scripts utilizing the MDAnalysis library. For each simulation, we calculated root mean square deviation of protein backbone atoms relative to the starting structure to assess overall structural stability. RMSD provides a measure of how much the structure changes during simulation. We also calculated RMSD of ligand heavy atoms after superimposing protein backbone atoms to assess whether ligands maintained their docked binding modes or drifted to alternative positions.

Root mean square fluctuation per residue quantified flexibility of different protein regions. RMSF was calculated for C-alpha atoms of each residue across the entire trajectory. High RMSF indicates flexible regions, while low RMSF indicates rigid regions. Binding site residues showing moderate RMSF suggest induced fit accommodation of ligands without loss of binding site integrity.

Hydrogen bonds between protein and ligand were monitored throughout trajectories using geometric criteria: donor-acceptor distance less than 3.5 angstroms and donor-hydrogen-acceptor angle greater than 120 degrees. For each potential hydrogen bond, we calculated occupancy as the percentage of trajectory frames in which the hydrogen bond criteria were satisfied. Hydrogen bonds with occupancy above 50% were considered stable and likely contributing significantly to binding affinity.

Contact analysis identified hydrophobic interactions by finding instances where carbon atoms of the ligand came within 4.5 angstroms of carbon atoms in hydrophobic residues (Val, Ile, Leu, Phe, Trp, Met). These contacts were summed across the trajectory to quantify persistent hydrophobic interactions stabilizing binding.

The radius of gyration of the protein was calculated to detect any large-scale unfolding or conformational changes. Stable radius of gyration throughout simulation confirms that the protein maintains its fold. Solvent accessible surface area buried upon ligand binding was calculated by comparing protein SASA in the complex versus the apo form, with greater buried surface generally correlating with stronger binding.

Water molecules in the binding site were tracked throughout simulations. Persistent water molecules remaining in the binding site for more than 50% of the trajectory were identified and their positions mapped. These waters might mediate protein-ligand interactions or represent displaced waters that could contribute entropy favorably to binding.

2.9 Binding Free Energy Calculations

For compounds showing stable binding in molecular dynamics simulations, we calculated binding free energies using the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method implemented in the g_mmpbsa tool. This approach decomposes binding free energy into components including van der Waals interactions, electrostatic interactions, polar solvation, and nonpolar solvation.

The MM-PBSA calculation was performed on 1,000 snapshots extracted evenly from the last 100 nanoseconds of each trajectory, ensuring that only equilibrated portions of simulations contributed to the calculation. For each snapshot, energies were calculated for the complex, the protein alone, and the ligand alone. The binding free energy was estimated as the difference:

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

Each energy term was calculated as:

$G = E_{MM} + G_{solvation} - TS$

where E_{MM} includes bonded terms (bonds, angles, dihedrals), van der Waals interactions, and electrostatic interactions calculated using molecular mechanics. $G_{solvation}$ includes polar and nonpolar contributions. The polar solvation contribution was calculated by solving the Poisson-Boltzmann equation with dielectric constant 2 for protein interior and 80 for solvent. The nonpolar solvation contribution was estimated from solvent accessible surface area using a surface tension coefficient of 0.0226 kJ/mol/Ų.

The entropy term (TS) was not explicitly calculated in this study as entropy calculations using normal mode analysis are computationally expensive and introduce significant uncertainty. Instead, we compared relative binding free energies among compounds, assuming entropy contributions would be similar for compounds of comparable size and flexibility.

For each compound, binding free energy was reported as the mean and standard deviation across the 1,000 snapshots. Decomposition of binding free energy by residue identified which protein residues contributed most to binding, providing insights into key interactions driving affinity.

3. Results

3.1 Homology Model Quality and Validation

The GLUT3 homology model constructed using GLUT1 as template showed excellent stereochemical quality across multiple validation metrics. Ramachandran plot analysis revealed 91.2% of residues in most favored regions, 7.8% in additionally allowed regions, and only 1.0% in generously allowed regions, with no residues in disallowed regions. This distribution significantly exceeds the threshold for acceptable models (>90% in favored regions).

VERIFY3D analysis, which assesses whether the three-dimensional model is compatible with the amino acid sequence, produced a score of 89.4%, well above the 80% threshold for reliable models. ERRAT, which analyzes statistics of non-bonded interactions between atoms, gave an overall quality factor of 87.3%. These validation results collectively confirmed that the homology model possessed suitable quality for structure-based virtual screening.

Visual inspection of the model revealed proper folding of the twelve transmembrane alpha helices arranged in two six-helix bundles, consistent with the major facilitator superfamily fold. The substrate binding cavity formed at the interface between the N-terminal and C-terminal domains showed clear accessibility from the intracellular side, matching the inward-open conformation of the template. Key residues implicated in glucose binding from mutagenesis studies aligned spatially in positions appropriate for coordinating substrate hydroxyl groups.

3.2 Virtual Screening Results

Molecular docking of 85,000 compounds against the GLUT3 model produced a distribution of binding energies ranging from -4.2 to -11.3 kcal/mol. The distribution showed roughly normal shape with a median of -6.1 kcal/mol. We applied a threshold of -8.5 kcal/mol, approximately two standard deviations below the median, to select top hits. This threshold identified 156 compounds (0.18% hit rate) for detailed analysis.

Visual inspection of these 156 docking poses revealed that most compounds (118 of 156, 75.6%) adopted reasonable binding modes with appropriate interactions with key residues. Common binding features included hydrogen bonding to Gln161 and Asn317, which are positioned similarly to their counterparts in GLUT1 known to coordinate glucose hydroxyl groups. Hydrophobic interactions with Trp388, Phe379, and Ile168 provided nonpolar stabilization. Many compounds showed aromatic stacking with Trp388, positioned to interact with the pyranose ring of glucose.

The 38 compounds excluded after visual inspection typically showed one or more problems: incorrect orientation leaving key hydrogen bond donors/acceptors distant from appropriate partners (14 compounds), steric clashes with protein atoms (11 compounds), binding primarily through nonspecific hydrophobic burial without specific interactions (8 compounds), or binding at pocket edges rather than deep cavity (5 compounds).

Interaction fingerprint analysis of the 118 retained compounds revealed six clusters representing distinct binding modes. The largest cluster (43 compounds) showed binding mode similar to predicted glucose binding, with multiple hydrogen bonds to Gln161, Asn317, and Gln282. A second cluster (28 compounds) adopted slightly shifted poses making primary interactions with Asn29 and Gln161. Other clusters represented alternative binding modes potentially targeting allosteric sites or subpockets within the main cavity.

3.3 QSAR Model Performance and Predictions

The QSAR model developed using 89 literature compounds with experimental GLUT3 inhibitory activity achieved strong performance metrics. The Random Forest model with 18 descriptors showed training set R^2 of 0.84 and RMSE of 0.31 pIC50 units. Five-fold cross-validation yielded Q^2 of 0.72, indicating good internal predictive ability. External validation on the test set (27 compounds never seen during training) produced R^2 of 0.69 and RMSE of 0.41 pIC50 units.

Analysis of descriptor importance revealed that molecular weight, LogP, and topological polar surface area were among the most predictive features, consistent with their known influence on membrane transporter binding. Three-dimensional descriptors including radius of gyration and 3D-MoRSE descriptors also contributed significantly, indicating that molecular shape and spatial distribution of atoms influence GLUT3 affinity beyond simple 2D connectivity.

Application of this model to predict activity for the 118 compounds from virtual screening produced pIC50 predictions ranging from 4.8 to 7.3. The top 48 compounds showed predicted pIC50 above 6.5 (IC50 < 300 nM), suggesting strong inhibitory activity worthy of further investigation. The predicted activities showed reasonable correlation with docking scores (Pearson $r = 0.58$, $p < 0.001$), but QSAR predictions provided finer discrimination among compounds with similar docking energies.

Interestingly, several compounds with moderate docking scores (-8.5 to -9.0 kcal/mol) received high QSAR predictions (pIC50 > 7.0), while some compounds with excellent docking scores (-10.0 to -11.0 kcal/mol) showed lower QSAR predictions (pIC50 5.5-6.0). This discordance highlights the value of

using multiple orthogonal methods, as docking and QSAR capture different aspects of binding and activity.

3.4 ADMET Profiling Results

ADMET prediction of the 48 QSAR-selected compounds revealed that 36 compounds had one or more properties falling outside acceptable ranges for oral drugs. The most common issues were poor water solubility (14 compounds with log S < -6, indicating limited solubility in aqueous media), susceptibility to P-glycoprotein efflux (8 compounds predicted as P-gp substrates, which would reduce oral absorption and brain penetration), predicted mutagenicity in Ames test (6 compounds containing structural alerts for genotoxicity), strong hERG channel blocking (5 compounds with predicted hERG pIC50 > 6.0, associated with cardiac arrhythmia risk), and poor blood-brain barrier penetration (3 compounds with log BB < -1).

The 12 compounds passing all ADMET filters demonstrated drug-like profiles across all evaluated properties. Their molecular weights ranged from 287 to 456 Daltons, LogP values from 1.8 to 4.2, and topological polar surface areas from 58 to 128 square angstroms, all within optimal ranges for oral drugs. All 12 showed predicted high human intestinal absorption (>80%), good Caco-2 permeability (log Papp > -5.0 cm/s), and were not predicted to be P-glycoprotein substrates.

For blood-brain barrier penetration, important for treating brain tumors, all 12 compounds showed log BB values between -0.3 and 0.8, suggesting good to excellent CNS penetration potential. None were predicted to be mutagenic in Ames test or hepatotoxic. hERG channel blocking predictions ranged from pIC50 4.2 to 5.7, indicating weak interaction unlikely to cause cardiac safety issues. Predicted oral LD50 values in rats ranged from 850 to 2300 mg/kg, suggesting low acute toxicity.

Metabolic stability predictions indicated that most compounds would be substrates for CYP3A4 metabolism, which is typical and generally acceptable provided metabolic rates are moderate. None were predicted to be strong inhibitors of major CYP enzymes (IC50 < 10 μM), reducing likelihood of drug-drug interactions. Predicted plasma half-lives ranged from 3.2 to 8.7 hours, acceptable for twice-daily dosing regimens.

3.5 Molecular Dynamics Simulation Stability

Molecular dynamics simulations of the 12 ADMET-filtered compounds over 200 nanoseconds revealed varying degrees of binding stability. Protein backbone RMSD analysis showed that all simulations equilibrated within the first 20-30 nanoseconds, after which RMSD values plateaued between 1.8 and 3.2 angstroms, indicating stable protein structures throughout the simulations. These values are typical

for well-folded proteins and confirmed that the GLUT3 model maintained structural integrity during dynamics.

Ligand RMSD analysis, however, differentiated compounds more dramatically. Five compounds (designated GLUT3-INH-01 through GLUT3-INH-05) maintained stable binding throughout the 200 nanosecond simulations, with ligand RMSD values remaining below 2.5 angstroms after initial equilibration. These compounds showed only minor fluctuations around their initial docked poses, indicating that the predicted binding modes were energetically favorable and kinetically stable.

Four compounds showed intermediate stability. Their ligand RMSD values increased gradually during the first 100 nanoseconds to values between 3.5 and 5.2 angstroms before stabilizing. Visual inspection of these trajectories revealed that these compounds underwent repositioning within the binding pocket during early simulation, finding alternative binding modes somewhat different from their initial docked poses. While these alternative poses remained stable for the remainder of simulation, the significant drift from predicted poses reduced confidence in the original docking predictions.

Three compounds displayed poor binding stability. Two showed progressive increases in ligand RMSD throughout simulation, reaching values above 8 angstroms by 200 nanoseconds, indicating substantial movement away from the binding site. Visual inspection confirmed that these compounds had largely dissociated from the pocket by simulation end. The third compound showed episodic unbinding and rebinding behavior, with RMSD fluctuating dramatically between 1.5 and 9.5 angstroms as the compound repeatedly left and re-entered the binding cavity. This behavior suggested weak, non-specific binding unlikely to produce biological activity.

Root mean square fluctuation analysis of protein residues showed interesting patterns. Binding site residues in simulations with stable ligands (GLUT3-INH-01 through -05) showed reduced RMSF compared to the apo protein simulation, indicating that ligand binding stabilized these regions. Specifically, residues 158-165, 280-285, and 315-320 showed 30-45% reductions in flexibility when stable ligands were bound. In contrast, simulations with unstable ligands showed RMSF values similar to or higher than the apo protein, suggesting that these compounds failed to establish stabilizing interactions.

The radius of gyration remained stable across all simulations, ranging between 21.3 and 21.9 angstroms with standard deviations below 0.3 angstroms, confirming that the overall protein fold was maintained. Solvent accessible surface area calculations showed that stable ligands buried between 420

and 580 square angstroms of protein surface, consistent with substantial binding interfaces. Unstable compounds buried less surface area (180-320 square angstroms), consistent with their weaker binding.

3.6 Detailed Analysis of Stable Compounds

The five compounds showing stable binding in molecular dynamics simulations underwent detailed characterization to understand their binding mechanisms and relative merits. Table 1 summarizes key properties of these lead compounds.

GLUT3-INH-01 established stable binding through a network of four hydrogen bonds. The compound's primary amine formed a salt bridge with Glu380 (occupancy 87% across the trajectory), while hydroxyl groups on the compound formed hydrogen bonds with Gln161 (occupancy 72%) and Asn317 (occupancy 68%). An additional weak hydrogen bond with Gln282 showed 34% occupancy. Hydrophobic interactions with Trp388, Phe379, and Ile168 contributed additional stabilization, with these contacts maintained in over 85% of trajectory frames. The compound's aromatic ring system showed parallel stacking with Trp388 at distances of 3.6-4.2 angstroms throughout most of the simulation.

GLUT3-INH-02 adopted a somewhat different binding mode. This compound's carboxylic acid group formed a bidentate interaction with Arg126, maintaining both hydrogen bonds with occupancies above 80%. A pyridine nitrogen accepted a hydrogen bond from the backbone NH of Gly384 (occupancy 65%). The compound's bicyclic core system occupied a deep hydrophobic pocket formed by Ile168, Val165, and Leu274, with persistent contacts to all three residues. Interestingly, a crystallographic water molecule observed in GLUT1 structures remained bound throughout the simulation, mediating a hydrogen bond bridge between the compound's carbonyl oxygen and Asn29.

GLUT3-INH-03 emerged as the top candidate based on multiple criteria. This compound formed five hydrogen bonds with exceptionally high occupancies: Gln161 (92%), Asn317 (88%), Trp388 backbone NH (81%), Gln282 (76%), and Asn29 (69%). The extensive hydrogen bonding network resulted in very low ligand RMSD (mean 1.1 angstroms) and exceptional stability throughout simulation. The compound contains a novel benzothiazole scaffold not previously reported in GLUT inhibitor literature, suggesting it represents a new chemical series. Its extended conformation allowed it to span the entire binding cavity, making contacts with both the N-terminal and C-terminal six-helix bundles.

GLUT3-INH-04 (molecular weight 367 Da, predicted pIC50 6.9) showed moderate hydrogen bonding (three bonds with occupancies of 78%, 62%, and 45%) but exceptional hydrophobic complementarity. The compound's structure includes three aromatic rings that pack efficiently into the binding cavity,

making extensive van der Waals contacts. Hydrophobic interaction analysis revealed persistent contacts (>70% occupancy) with nine hydrophobic residues: Trp388, Phe379, Ile168, Val165, Leu274, Ile322, Val319, Phe26, and Met424. This extensive hydrophobic network likely contributes substantial binding free energy despite fewer hydrogen bonds than other leads.

GLUT3-INH-05 displayed intermediate characteristics between the highly polar GLUT3-INH-03 and highly hydrophobic GLUT3-INH-04. It formed three strong hydrogen bonds (occupancies 85%, 73%, and 61%) and made substantial hydrophobic contacts (persistent interactions with six hydrophobic residues). Notably, this compound showed the least movement during simulation (ligand RMSD standard deviation of only 0.4 angstroms), suggesting very stable binding. The compound contains a sulfonamide group that formed the highest-occupancy hydrogen bond (85%) with Gln161, and this interaction appeared to anchor the molecule strongly in place.

Water molecule analysis revealed interesting differences among compounds. GLUT3-INH-03 displaced all water molecules from the binding cavity, achieving complete dehydration of the binding site. This thorough displacement likely contributes favorably to binding entropy. In contrast, GLUT3-INH-02 and GLUT3-INH-05 retained one conserved water molecule that mediated protein-ligand interactions, and this water showed very low mobility throughout simulations (B-factors equivalent to 12-15 ų), suggesting it forms an integral part of the binding interface. GLUT3-INH-01 and GLUT3-INH-04 showed intermediate behavior with transient water molecules entering and exiting the binding site throughout trajectories.

3.7 Binding Free Energy Calculations

MM-PBSA calculations on snapshots from the equilibrated portions of trajectories (100-200 ns) provided quantitative estimates of binding free energies. Results must be interpreted cautiously since MM-PBSA calculations contain approximations and omit explicit entropy calculations, but relative values among similar compounds generally correlate with experimental affinities.

GLUT3-INH-03 showed the most favorable binding free energy at -42.8 ± 3.2 kcal/mol (mean ± standard deviation across 1,000 snapshots). Decomposition of this energy revealed substantial contributions from both electrostatic interactions (-156.3 kcal/mol) and van der Waals interactions (-68.4 kcal/mol). Polar solvation unfavorably opposed binding (+189.7 kcal/mol), as expected since desolvating polar groups costs energy. Nonpolar solvation contributed favorably (-7.8 kcal/mol) based on buried surface area. The balance of these terms produced the overall favorable binding free energy.

GLUT3-INH-02 showed binding free energy of -39.4 ± 4.1 kcal/mol, slightly less favorable than GLUT3-INH-03. The electrostatic contribution was more negative (-168.2 kcal/mol), reflecting the carboxylate group's strong interactions, but this was partially offset by greater polar solvation penalty (+202.1 kcal/mol). Van der Waals interactions contributed -62.7 kcal/mol, somewhat less than GLUT3-INH-03.

GLUT3-INH-05 produced binding free energy of -37.9 ± 3.8 kcal/mol. This compound showed balanced contributions from electrostatics (-142.6 kcal/mol) and van der Waals interactions (-65.3 kcal/mol), with moderate solvation penalty (+177.8 kcal/mol).

GLUT3-INH-01 and **GLUT3-INH-04** showed similar binding free energies of -35.2 ± 4.5 and -34.8 ± 4.2 kcal/mol, respectively, somewhat less favorable than the top three compounds but still indicating strong binding. GLUT3-INH-01's binding was dominated by electrostatic interactions (-161.4 kcal/mol) due to its salt bridge with Glu380, while GLUT3-INH-04 showed the strongest van der Waals contribution (-71.2 kcal/mol) consistent with its extensive hydrophobic contacts.

Per-residue decomposition of binding free energies identified the key contributors to binding for each compound. For all five compounds, Gln161, Asn317, and Trp388 contributed substantially (typically -2.5 to -4.8 kcal/mol per residue). Additional important residues varied by compound but commonly included Phe379, Ile168, Gln282, and Glu380. The consistency of Gln161, Asn317, and Trp388 contributions across compounds confirms these residues as critical for GLUT3 inhibitor binding and suggests they should be focal points for structure-activity relationship studies.

Comparison of binding free energies with QSAR-predicted activities showed reasonable correlation (Pearson $r = 0.79$, $p = 0.11$ for $n=5$). The ranking by binding free energy (INH-03 > INH-02 > INH-05 > INH-01 > INH-04) matched the ranking by predicted pIC50 (INH-03 > INH-02 > INH-05 > INH-01 ≈ INH-04) quite closely. This convergence between independent computational methods increases confidence that these compounds genuinely represent strong GLUT3 binders.

3.8 Chemical Diversity and Synthetic Accessibility

Analysis of the five lead compounds revealed satisfactory chemical diversity, important for providing multiple starting points for medicinal chemistry optimization. Tanimoto similarity coefficients based on Morgan fingerprints ranged from 0.23 to 0.48, indicating that the compounds represent distinct chemical scaffolds rather than close analogs. This diversity arose from the careful library curation and the use of multiple binding mode clusters in hit selection.

GLUT3-INH-01 contains a piperidine core with hydroxyl and amine substituents, representing a relatively simple structure amenable to straightforward synthetic chemistry. The synthetic accessibility score calculated by RDKit was 2.8 (scale of 1-10, where 1 is easiest), suggesting facile synthesis. Literature searches identified similar piperidine derivatives in existing synthetic chemistry papers, confirming that established procedures could likely be adapted.

GLUT3-INH-02 features a bicyclic core (quinoline derivative) with carboxylic acid functionality. The synthetic accessibility score of 3.2 indicates moderate complexity. The core heterocycle is commercially available, and proposed synthesis would involve 4-5 steps from this starting material, primarily functional group manipulations.

GLUT3-INH-03, despite showing the best predicted activity and binding free energy, presented the highest synthetic complexity with a score of 4.7. The benzothiazole core fused to an additional aromatic system with multiple substituents would require multistep synthesis, estimated at 7-9 steps from commercial precursors. However, the superior predicted potency justifies the synthetic investment for this lead compound.

GLUT3-INH-04, based on a triphenyl scaffold, showed synthetic accessibility score of 3.5. The symmetric structure might allow efficient synthesis through iterative coupling reactions. The presence of three aromatic rings provides multiple points for modification during optimization.

GLUT3-INH-05 (synthetic accessibility 2.9) represents perhaps the most practical starting point, combining favorable predicted activity with relatively simple synthesis. The sulfonamide-containing structure could likely be prepared in 5-6 steps using established sulfonamide chemistry.

Patent landscape searches revealed that none of the five lead compounds or their close structural analogs appeared in existing patents related to glucose transporter inhibitors, GLUT3 inhibitors, or cancer metabolism modulators. This novelty is important for potential intellectual property protection and commercial development.

3.9 Selectivity Analysis Against Other GLUT Isoforms

An important consideration for GLUT3 inhibitor development is selectivity over other GLUT family members, particularly GLUT1, which mediates glucose uptake in red blood cells and endothelial cells. To address this, we constructed a homology model of human GLUT1 using the same template and methodology applied for GLUT3, then docked the five lead compounds into both GLUT3 and GLUT1 binding sites.

The results showed promising selectivity patterns. GLUT3-INH-03 docked into GLUT1 with binding energy of -7.2 kcal/mol, compared to -10.8 kcal/mol in GLUT3, suggesting approximately 3.6 kcal/mol preference for GLUT3. Detailed analysis revealed that this selectivity arose from subtle differences in binding pocket residues. GLUT3 contains Ile168 where GLUT1 has Val168, providing a slightly larger hydrophobic pocket that better accommodates GLUT3-INH-03's extended structure. Additionally, GLUT3 has Gln161 where GLUT1 has Asn160, and the longer glutamine side chain forms a more favorable hydrogen bond geometry with the compound.

GLUT3-INH-02 showed even greater selectivity, with GLUT1 binding energy of -6.8 kcal/mol versus -9.9 kcal/mol in GLUT3 (3.1 kcal/mol difference). The carboxylate group of this compound interacts favorably with GLUT3 Arg126, which is positioned slightly differently in GLUT1 due to adjacent sequence variations, reducing the interaction strength.

GLUT3-INH-05 demonstrated the most dramatic selectivity, docking to GLUT1 with only -5.4 kcal/mol binding energy compared to -9.6 kcal/mol for GLUT3 (4.2 kcal/mol difference). The sulfonamide group appears to exploit GLUT3-specific pocket features particularly effectively.

GLUT3-INH-01 and GLUT3-INH-04 showed more modest selectivity (1.8 and 2.1 kcal/mol differences, respectively), suggesting these might have broader GLUT family activity. While this reduced selectivity could be problematic for clinical development, these compounds might serve as useful tool compounds for studying general GLUT biology or as starting points for introducing selectivity through structural modifications.

These selectivity predictions, based on static docking, should be interpreted cautiously. More rigorous assessment would require molecular dynamics simulations of compounds bound to GLUT1 and experimental testing against panels of GLUT isoforms. Nevertheless, the computational predictions suggest that achieving GLUT3 selectivity is feasible, particularly for compounds like GLUT3-INH-03 and GLUT3-INH-05.

3.10 Predicted Binding Mechanism and Comparison to Glucose

To understand how the identified inhibitors might block GLUT3 function, we compared their binding modes to the predicted binding mode of glucose, the natural substrate. Glucose was docked into the GLUT3 model and subjected to 100 nanosecond molecular dynamics simulation for reference.

Glucose, as expected, formed multiple hydrogen bonds with polar residues lining the cavity. The hydroxyl groups at positions 3, 4, and 6 of the glucose pyranose ring formed stable hydrogen bonds

with Gln161, Asn317, and Gln282, respectively. The hydroxyl at position 2 interacted with Asn29. These interactions showed occupancies between 65% and 82%, somewhat lower than the best inhibitor interactions, likely reflecting glucose's smaller size and resulting higher mobility within the binding pocket.

All five lead inhibitors overlapped substantially with the glucose binding site, suggesting they function as competitive inhibitors that directly block substrate binding rather than acting allosterically. However, each inhibitor extended beyond the glucose binding volume, making additional contacts not available to the small substrate molecule.

GLUT3-INH-03 showed the most extensive overlap, with its core structure occupying the same space as glucose while extensions reached into adjacent subpockets. This compound essentially acts as a "super-substrate," mimicking glucose interactions while achieving additional stabilization through its larger structure. The extensive hydrogen bonding network (five bonds vs. four for glucose) and greater hydrophobic burial likely explain its superior predicted affinity.

GLUT3-INH-02 and GLUT3-INH-05 adopted similar strategies, positioning functional groups to mimic key glucose hydroxyl interactions while extending into additional pocket space. The charged groups in these compounds (carboxylate in INH-02, sulfonamide in INH-05) formed very strong interactions that glucose cannot achieve with its neutral hydroxyl groups, likely contributing to enhanced affinity.

GLUT3-INH-01 and GLUT3-INH-04 showed somewhat different binding modes, with less complete overlap with glucose. These compounds might function as mixed competitive/non-competitive inhibitors, partially overlapping the substrate site while also making contacts that could stabilize conformations incompatible with glucose binding and transport.

The finding that all inhibitors overlap the substrate binding site suggests they should effectively block glucose transport. Experimental validation will be necessary to determine whether they are purely competitive, non-competitive, or uncompetitive inhibitors, which depends on whether they preferentially bind the inward-facing, outward-facing, or occluded conformations of the transporter. The current study used an inward-facing model, so inhibitors identified here likely bind this conformation preferentially.

4. Discussion

This computational study successfully identified five novel chemical scaffolds predicted to act as potent and selective GLUT3 inhibitors. The multi-stage approach, combining structure-based virtual screening, QSAR modeling, ADMET prediction, and molecular dynamics simulation, progressively filtered 85,000 initial candidates to five high-confidence leads. This dramatic enrichment from 85,000 to 5 compounds (0.006% of the original library) demonstrates the power of integrative computational methods for drug discovery.

Several aspects of this work merit discussion. First, the use of a homology model rather than an experimental GLUT3 structure introduces uncertainty. However, multiple validation metrics confirmed model quality, and the high sequence identity with GLUT1 (64%) in conserved regions provides confidence in the substrate binding site structure. The subsequent experimental determination of GLUT3 structure would allow retrospective validation of this model and refinement of inhibitor predictions. Nevertheless, successful drug discovery has been achieved with homology models of similar or lower quality, particularly when, as here, the binding site residues are highly conserved.

The virtual screening hit rate of 0.18% (156 compounds with binding energy < -8.5 kcal/mol from 85,000 screened) appears low but is actually quite reasonable for structure-based screening. Random screening typically yields hit rates of 0.01-0.1% in high-throughput screens, so our computational enrichment represents a 2-20 fold improvement over random selection. More importantly, the subsequent filtering stages showed that these computational hits possessed genuinely favorable properties, with 75.6% passing visual inspection (vs. perhaps 10-20% for random compounds) and 30.8% (48 of 156) showing strong predicted activity in QSAR modeling.

The QSAR model performance (test set $R^2 = 0.69$) provides reasonable but not exceptional predictive power. This moderate performance likely reflects the structural diversity in the training set and the complexity of predicting activity across diverse scaffolds. The correlation between QSAR predictions and MM-PBSA binding free energies ($r = 0.79$) for the final five compounds provides orthogonal validation that both methods captured genuine trends in binding affinity. Future iterations could improve QSAR models by incorporating more training data as experimental results become available or by using more sophisticated machine learning architectures like graph neural networks that can capture complex structure-activity relationships.

The ADMET filtering eliminated 75% of compounds, highlighting a critical point: computational binding predictions alone are insufficient for drug discovery. Many compounds that bind protein

targets effectively in silico fail in vivo due to poor pharmacokinetics or toxicity. The integration of ADMET prediction early in the discovery process, before any synthesis investment, represents best practice in modern computational drug design. The 12 compounds passing ADMET filters represent a dramatically enriched set compared to random drug-like compounds, where perhaps only 5-10% would show favorable profiles across all evaluated properties.

The molecular dynamics simulations provided critical validation that went beyond static docking predictions. The finding that only 5 of 12 ADMET-filtered compounds (42%) showed stable binding in explicit solvent dynamics illustrates that docking scores do not always predict kinetic stability. Compounds may find favorable poses in docking that prove unstable once protein and ligand flexibility and explicit solvation are considered. The 200-nanosecond simulation length was sufficient to observe unstable compounds dissociating or repositioning, while stable compounds maintained their bound conformations. Some might argue for even longer simulations or enhanced sampling methods, but the clear separation between stable and unstable compounds suggests that 200 nanoseconds provided adequate discrimination.

The binding free energy calculations using MM-PBSA provided quantitative ranking of compounds. While absolute free energy values from MM-PBSA should not be over-interpreted due to methodological limitations (lack of explicit entropy, approximate treatment of solvation, single-trajectory approach), relative rankings among similar compounds generally correlate with experimental affinities. The range of binding free energies from -34.8 to -42.8 kcal/mol corresponds roughly to IC50 values spanning three orders of magnitude based on empirical correlations in the literature, suggesting our top compounds might show low nanomolar to low micromolar inhibition. Experimental testing will determine actual potencies, but the computational predictions provide a reasonable rank ordering for prioritizing synthesis and testing.

The predicted selectivity for GLUT3 over GLUT1, particularly for compounds GLUT3-INH-03 and GLUT3-INH-05, addresses a major challenge in GLUT inhibitor development. Previous GLUT inhibitors often showed poor isoform selectivity, limiting their therapeutic utility due to on-target toxicity. The subtle sequence differences between GLUT3 and GLUT1 in the binding pocket, particularly Ile168 vs. Val168 and Gln161 vs. Asn160, provide exploitable selectivity determinants. Structure-activity relationship studies could further optimize selectivity by systematically exploring how modifications to lead scaffolds affect binding to each isoform. The computational framework

established here could rapidly evaluate proposed modifications by docking and simulation in both GLUT3 and GLUT1 models before synthesis.

The chemical diversity of the five leads provides multiple starting points for medicinal chemistry optimization. Having multiple scaffolds reduces risk if one series encounters unexpected problems during development (poor synthetic chemistry, unanticipated toxicity in animal studies, intellectual property conflicts). Each scaffold offers distinct opportunities for modification. The benzothiazole core of GLUT3-INH-03 could be varied to explore related heterocycles. The sulfonamide in GLUT3-INH-05 could be replaced with sulfones, sulfoxides, or other hydrogen bond accepting groups. The piperidine in GLUT3-INH-01 could be replaced with other saturated nitrogen heterocycles to explore structure-activity relationships.

The synthetic accessibility analysis suggests that all five compounds could be synthesized, though with varying effort. GLUT3-INH-01 and GLUT3-INH-05 appear most accessible for immediate synthesis and testing. GLUT3-INH-03, despite higher synthetic complexity, warrants investment given its superior predicted potency and binding free energy. A pragmatic strategy would begin with synthesis and testing of GLUT3-INH-01 and GLUT3-INH-05 for rapid validation, while simultaneously pursuing the more challenging synthesis of GLUT3-INH-03. If initial compounds show activity in biochemical and cellular assays, the more complex synthesis becomes justified.

Several experimental studies would validate these computational predictions and advance these compounds toward clinical development. First, compounds should be synthesized and tested in glucose uptake assays using cells expressing GLUT3. These assays would determine actual IC50 values for comparison with computational predictions. Testing should include counter-screening against GLUT1-expressing cells to assess selectivity experimentally. Second, binding studies using surface plasmon resonance, isothermal titration calorimetry, or other biophysical methods could directly measure binding affinity and thermodynamics for comparison with MM-PBSA predictions. Third, structural validation through co-crystallization or cryo-EM of inhibitor-GLUT3 complexes would reveal actual binding modes for comparison with computational predictions. Such structures would enable structure-guided optimization to improve potency and selectivity.

If compounds show promising activity in biochemical assays, subsequent studies should evaluate anti-cancer activity in cell-based assays. Cancer cell lines representing relevant tumor types (glioblastoma, lung cancer, breast cancer, pancreatic cancer) with confirmed GLUT3 overexpression should be treated with compounds to assess growth inhibition. Cells cultured under hypoxia, which increases

dependence on GLUT3, might show enhanced sensitivity. Glucose uptake assays in cancer cells would directly demonstrate on-target mechanism. Combination studies with standard chemotherapy agents could reveal whether GLUT3 inhibition sensitizes resistant tumors to conventional treatments.

Pharmacokinetic studies in mice would validate ADMET predictions experimentally. Oral bioavailability could be measured by comparing plasma exposure after oral versus intravenous dosing. Brain penetration could be assessed by measuring brain-to-plasma ratios after systemic administration, critical for glioblastoma applications. Metabolic stability could be confirmed in liver microsome assays and in vivo clearance studies. Toxicity screening in rodents would identify any unexpected safety issues not predicted computationally.

Ultimately, compounds showing good potency, selectivity, pharmacokinetics, and safety in preclinical studies could advance to efficacy testing in mouse cancer models. Glioblastoma xenograft models would be particularly relevant given the poor prognosis of these tumors and lack of effective treatments. Demonstrating that GLUT3 inhibitors slow tumor growth, extend survival, or sensitize tumors to radiation or chemotherapy would provide proof-of-concept for clinical development. Similar studies in models of lung cancer, breast cancer, and pancreatic cancer would explore pan-cancer therapeutic potential.

The cancer metabolism field has received intense interest over the past decade, with numerous attempts to target metabolic vulnerabilities including glycolysis inhibitors, glutaminase inhibitors, and LDHA inhibitors. While targeting glucose metabolism seems logical given cancer cells' glucose dependence, achieving selectivity to avoid toxicity to normal tissues remains challenging. GLUT3 offers advantages over many other metabolic targets because its expression pattern is relatively restricted in normal adult tissues, being high primarily in neurons which could potentially tolerate moderate inhibition, while its upregulation in cancer makes tumors disproportionately vulnerable. The selectivity over GLUT1 predicted for our top compounds further reduces concerns about systemic toxicity.

One consideration not addressed in this purely computational study is the question of whether GLUT3 inhibition alone would suffice to treat cancers or whether combination therapy would be necessary. Cancer cells show remarkable metabolic flexibility, potentially compensating for GLUT3 inhibition by upregulating other glucose transporters, increasing glutamine utilization, or adopting other metabolic strategies. Combination approaches might be required, pairing GLUT3 inhibitors with other metabolic inhibitors to close multiple escape routes simultaneously, or combining with conventional

chemotherapy or targeted therapy to attack cancer through complementary mechanisms. These questions would be addressed in future biological studies, but the compounds identified here provide the chemical tools necessary for such investigations.

5. Conclusions

This comprehensive computational study successfully identified five novel chemical scaffolds predicted to function as potent and selective GLUT3 inhibitors with favorable drug-like properties. Beginning with a curated library of 85,000 compounds, we applied successive filters of molecular docking, QSAR modeling, ADMET prediction, and molecular dynamics simulation to arrive at these high-confidence leads: GLUT3-INH-01 through GLUT3-INH-05.

The lead compounds exhibited strong predicted binding to GLUT3, with binding free energies ranging from -34.8 to -42.8 kcal/mol and predicted IC50 values in the low nanomolar to low micromolar range. All five maintained stable binding throughout 200 nanosecond molecular dynamics simulations, formed multiple hydrogen bonds with key binding site residues, and demonstrated extensive hydrophobic complementarity with the substrate binding pocket. The compounds showed promising selectivity for GLUT3 over GLUT1, with selectivity ranging from 1.8 to 4.2 kcal/mol in predicted binding energy differences.

ADMET predictions indicated that all five leads possess favorable drug-like properties including good oral bioavailability, blood-brain barrier penetration suitable for treating brain tumors, metabolic stability compatible with practical dosing regimens, and low toxicity risk across multiple endpoints. The chemical diversity of the scaffolds provides multiple starting points for medicinal chemistry optimization and reduces development risk.

GLUT3-INH-03 emerged as the most promising single candidate based on its superior binding free energy (-42.8 kcal/mol), highest predicted activity (pIC50 7.3), excellent stability in molecular dynamics, extensive hydrogen bonding network (five persistent bonds), and strong predicted selectivity over GLUT1 (3.6 kcal/mol). The novel benzothiazole scaffold represents a new chemical series for GLUT inhibition with clear intellectual property potential.

GLUT3-INH-05 represents perhaps the most immediately practical lead, combining strong predicted activity (pIC50 7.0) with relatively straightforward synthesis (accessibility score 2.9) and the highest predicted selectivity (4.2 kcal/mol). This compound would be an excellent candidate for early-stage synthesis and testing.

The remaining leads (GLUT3-INH-01, -02, and -04) provide alternative scaffolds if the top two encounter unexpected problems during experimental validation. The diversity of these scaffolds ensures multiple pathways forward for program advancement.

This work demonstrates the power of integrated computational drug discovery approaches that combine multiple complementary methods. Structure-based virtual screening identified compounds capable of fitting the binding site. QSAR modeling predicted activity quantitatively based on molecular properties. ADMET profiling eliminated compounds with likely pharmacokinetic or safety liabilities before any synthesis investment. Molecular dynamics simulations validated binding stability under realistic conditions. Binding free energy calculations provided quantitative ranking of candidates. The convergence of predictions across these independent methods substantially increases confidence in the final leads.

The compounds identified here are ready for experimental validation. Synthesis and testing in GLUT3 inhibition assays would determine actual potencies and selectivities for comparison with computational predictions. Positive results would justify advancement to cellular assays in GLUT3-overexpressing cancer cells, followed by pharmacokinetic studies and ultimately efficacy testing in mouse tumor models.

Beyond immediate drug discovery applications, these compounds could serve as chemical probes for understanding GLUT3 biology in normal and cancer cells. The role of GLUT3 in cancer metabolism, metastasis, and treatment resistance remains incompletely understood. Selective inhibitors would allow precise dissection of GLUT3 contributions distinct from other glucose transporters.

Looking forward, the therapeutic potential of GLUT3 inhibition extends beyond single-agent cancer treatment. These inhibitors might sensitize resistant tumors to conventional chemotherapy, radiation, or immunotherapy by metabolically stressing cancer cells. They might preferentially target cancer stem cells, which often show elevated GLUT3 expression. They might prove particularly effective in tumor types or contexts where GLUT3 plays dominant roles, such as glioblastoma under hypoxia.

The successful completion of this computational campaign represents an important step toward developing a new class of metabolic cancer therapeutics. The compounds identified here, particularly GLUT3-INH-03 and GLUT3-INH-05, represent genuine drug candidates with realistic paths to clinical development. While experimental validation remains necessary, the comprehensive computational characterization presented here provides strong justification for the synthetic and biological investment required. If these predictions prove accurate, the compounds could ultimately provide new treatment

options for cancer patients, addressing unmet medical needs in aggressive tumor types where current therapies remain inadequate.

References

Warburg, O., Wind, F., & Negelein, E. (1927). The metabolism of tumors in the body. *Journal of General Physiology, 8*(6), 519-530. https://doi.org/10.1085/jgp.8.6.519

Ancey, P. B., Contat, C., & Meylan, E. (2018). Glucose transporters in cancer - from tumor cells to the tumor microenvironment. *FEBS Journal, 285*(16), 2926-2943. https://doi.org/10.1111/febs.14577

Flavahan, W. A., Wu, Q., Hitomi, M., Rahim, N., Kim, Y., Sloan, A. E., Weil, R. J., Nakano, I., Sarkaria, J. N., Stringer, B. W., Day, B. W., Li, M., Lathia, J. D., Rich, J. N., & Hjelmeland, A. B. (2013). Brain tumor initiating cells adapt to restricted nutrition through preferential glucose uptake. *Nature Neuroscience, 16*(10), 1373-1382. https://doi.org/10.1038/nn.3510

Simpson, I. A., Dwyer, D., Malide, D., Moley, K. H., Travis, A., & Vannucci, S. J. (2008). The facilitative glucose transporter GLUT3: 20 years of distinction. *American Journal of Physiology-Endocrinology and Metabolism, 295*(2), E242-E253. https://doi.org/10.1152/ajpendo.90388.2008

Deng, D., Xu, C., Sun, P., Wu, J., Yan, C., Hu, M., & Yan, N. (2014). Crystal structure of the human glucose transporter GLUT1. *Nature, 510*(7503), 121-125. https://doi.org/10.1038/nature13306

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews, 46*(1-3), 3-26. https://doi.org/10.1016/S0169-409X(00)00129-0

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

Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y. C., Todeschini, R., Consonni, V., Kuz'min, V. E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terfloth, L., Gasteiger, J., Richard, A., & Tropsha, A. (2014). QSAR modeling: Where have you been? Where are you going to? *Journal of Medicinal Chemistry, 57*(12), 4977-5010. https://doi.org/10.1021/jm4004285

Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery, 10*(5), 449-461. https://doi.org/10.1517/17460441.2015.1032936

Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry, 58*(9), 4066-4072. https://doi.org/10.1021/acs.jmedchem.5b00104