Taylor & Francis
Taylor & Francis Group

Check for updates

Computational drug discovery of an inhibitor of APOBEC3B as a treatment for epithelial cancers

Dominic A. Caputa[a], Quin P. Blankenship[b], Zachary D. Smith[b,c], Molly M. Huebner[b], Zoe A. Vetter[a,b], Richard W. Parks[b,c], Saul Armendariz Lobera[c], Emmett M. Leddin[b], Cooper A. Taylor[d], Carol A. Parish[d] and Bill R. Miller III[b]

[a]Physics Department, Truman State University, Kirksville, MO, USA; [b]Chemistry Department, Truman State University, Kirksville, MO, USA; [c]Biology Department, Truman State University, Kirksville, MO, USA; [d]Department of Chemistry, University of Richmond, Richmond, VA, USA

Communicated by Ramaswamy H. Sarma

ABSTRACT

Cancer is one of the leading causes of death in the U.S., and tumorous cancers such as cervical, lung, breast, and ovarian cancers are the most common types. APOBEC3B is a nonessential cytidine deaminase found in humans and theorized to defend against viral infection. However, overexpression of APOBEC3B is linked to cancer in humans, which makes APOBEC3B a potential cancer treatment target through competitive inhibition for several tumorous cancers. Computational studies can help reveal a small molecule inhibitor using high-throughput virtual screening of millions of candidates with relatively little cost. This study aims to narrow the field of potential APOBEC3B inhibition candidates for future *in vitro* assays and provide an effective scaffold for drug design studies. Another goal of this project is to provide critical amino acid targets in the active site for future drug design studies. This study simulated 7.8 million drug candidates using high-throughput virtual screening and further processed the top scoring 241 molecules from AutoDock Vina, DOCK 6, and *de novo* design. Using virtual screening, *de novo* design, and molecular dynamics simulations, a competitive inhibitor candidate was discovered with the average binding free energy score of -46.03 kcal/mol, more than 10 kcal/mol better than the substrate control (dCMP). These results indicate that this molecule (or a structural derivative) may be an effective inhibitor of APOBEC3B and prevent host genome mutagenesis resulting from protein overexpression. Another important finding is the confirmation of essential amino acid targets, such as Tyr250 and Gln213 within the active site of APOBEC3B. Therefore, study used novel computational methods to provide a theoretical scaffold for future drug design studies that may prove useful as a treatment for epithelial cancers.

Abbreviations: APOBEC3B: apolipoprotein B mRNA editing enzyme; dCMP: deoxycytidylate; kcal/mol: kilocalories/mole; MD: molecular dynamics; MM-GBSA: molecular mechanics-generalized born surface area; ns: nanosecond; PCA: principal component analysis; ps: picosecond; RMSD: root-mean-squared deviation; RMSF: root-mean-squared fluctuation; ssDNA: single stranded DNA

Introduction

Cancer is the second leading cause of death in the United States, and tumorous cancers make up a significant portion of those deaths per year (CDCBreastCancer, 2023). Ovarian, cervical, breast, and lung cancer are the most common types of tumorous cancers. Cancers are inherently difficult to treat because there are many independent factors. A common pharmaceutical route is to identify the location of the cancer and treat using general chemotherapy, which has several short-term and chronic side effects.

Apolipoprotein B mRNA editing enzyme (abbreviated as APOBEC3B) is a human cytidine deaminase that has a role in innate immunity and has been linked to the mutagenesis of cancers (Hou et al., 2021). APOBEC3B is a nonessential human globular protein that is part of the larger APOBEC3 family and is responsible for the deamination of cytosine to uracil in single-stranded DNA (Harris, 2015). The catalytic site

on APOBEC3B, characterized primarily by the presence of a zinc ion, binds cytosine noncovalently, catalyzes its conversion to uracil, and then releases the product. Human APOBEC3B natively mutates single stranded DNA (ssDNA), which prevents viral DNA from being used for viral replication and infecting the host (Kouno et al., 2017). Overexpression of APOBEC3B in humans causes damage to host DNA, leading to a variety of random mutations (Harris, 2015). Cancer can result from the accumulation of mutations caused by overexpression of APOBEC3B (NCI, 2023). APOBEC3B is often overexpressed near cancerous tumors and in patients who are experiencing a relapse (Zhang et al., 2022). APOBEC3B has been linked to common cancers, such as breast cancer, ovarian cancer, lung cancer, and bladder cancer (Asaoka et al., 2019), and has been proposed to be an important drug target to treat cancers through competitive inhibition (Olson et al., 2018). Inhibition of APOBEC3B may severely limit tumor growth in patients without other
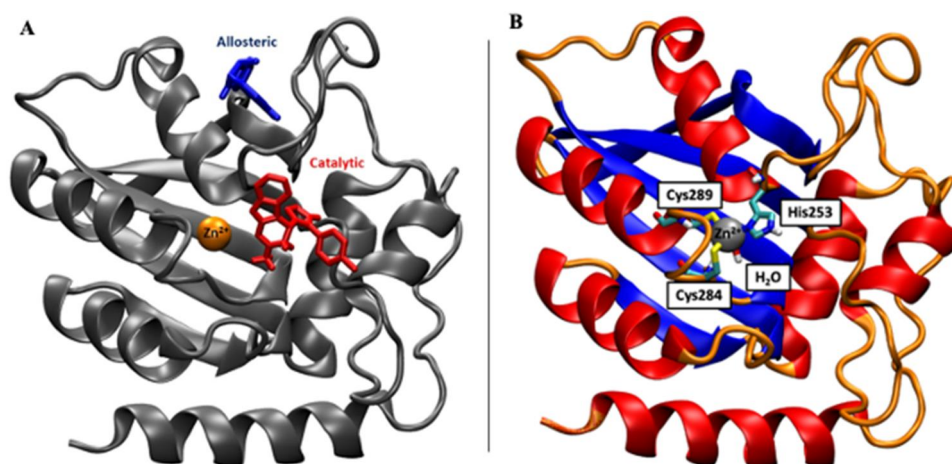
Figure 1. (A) Depiction of the representative ligands binding to the catalytic (red) and allosteric (blue) binding sites on the surface of APOBEC3B (gray) as observed from docking ZINC12 ligands. For perspective, the Zn^{2+} (orange) is shown. (B) The structure of APOBEC3B from the 5CQH PDB (Shi et al., 2015). The secondary structure of APOBEC3B (PDB: 5CQH) is depicted with α-helices (red), β-sheets (blue), and coils (orange) shown. The catalytic Zn^{2+} ion (gray sphere) is displayed coordinated to His253, Cys284, Cys289, and the catalytic water molecule.

underlying conditions and is therefore another treatment option to prevent the pre-cancerous state caused by APOBEC3B.

The APOBEC3B protein (Figure 1B) contains 382 amino acids with two globular Zn-coordinating domains, each containing a hydrophobic core consisting of five β-strands surrounded by six α-helices. There is a pseudo active site at the N-terminal domain and a catalytic active site that is at the C-terminal domain. There has been experimental success in the crystallization of APOBEC3B in isolation and when bound to ligands. In 2016, Shi et al. crystallized APOBEC3B (PDB: 5TD5) with the catalytic site accessible to ssDNA (4-mer) (Shi et al., 2017). In our simulations, we used the crystal structure (PDB: 5CQH) of the catalytic domain (amino acids 189–379) solved by Shi et al. in 2015 through X-ray diffraction (1.73 Å resolution) bound by a dCMP nucleotide and four ethylene glycol molecules (Shi et al., 2015). The 5CQH crystal structure solved by Shi et. Al was used as the receptor structure for all docking calculations and provided the starting structure for all molecular dynamics (MD) simulations performed in this study. This closed conformation, which is stabilized by Arg211 (loop 1) and Tyr315 (loop 7) stacking with Tyr313 positioned over the active site, represents an APOBEC3B conformation similar to what would be expected in the presence of a small molecule inhibitor and suggests that conformational changes occur to allow substrate binding. The Zn-coordinating motif includes His253, Cys284, Cys289, and the catalytic water molecule (substituted with ethylene glycol in the 5CQH crystal structure), which is hydrogen bonding with the activating Glu255. This structure represents a promising drug target for cancer treatment. Having a robust understanding of the structure and dynamics of APOBEC3B is key to designing an efficient inhibition candidate.

Historically, small molecule inhibitors of APOBEC3B have not proved to be effective, so other methods have been explored. Previous research used a kinase to phosphorylate APOBEC3B to inhibit its function (Matsumoto et al., 2019) and ssDNA was used to inhibit APOBEC3B in vivo (Barzak et al., 2019). There are other studies using macro-scale molecules, such as antibodies that bind with a K_d ∼70 nM and show promise with APOBEC3B inhibition (Tang et al., 2021). Additionally, the Epstein-Barr virus binds to APOBEC3B preferentially to prevent APOBEC3B-ssDNA binding, which provides safety from the viral immune response by APOBEC3B (Cryo-EM structure of the EBV ribonucleotide reductase BORF2 and mechanism of APOBEC3B inhibition., 2023). Most previous studies have dealt with large molecules, while few have sought to experimentally determine a small molecule inhibitor for APOBEC3B. Herein a small molecular inhibitor for APOBEC3B would be classified as a compound that competes for binding with the native ssDNA substrate. One such recent study discovered a compound (MN23) with relatively low efficacy ($IC_{50} = 0.15 \mu M$) (Olson, n.d.). A pair of potential small molecules, 2-pyrimidone and sodium iodide, can force APOBEC3B into an inactive state and prevent its cytidine deaminase activity were discovered in 2019 (Shi et al., 2019). However, neither of these small molecules were considered potent inhibitors in the study performed by Shi et al. (Shi et al., 2019) Experimental results have not yet provided a potent small molecule inhibitor for APOBEC3B. One previous study utilized computational methods by docking small molecules to narrow down ZINC12 ligands for in vitro assays (King et al., 2021). Previous computational studies have been limited to only using docking methods, while our study greatly increases the computational rigor used to identify possible drug candidates by combining docking with molecular dynamics simulations and free energy calculations to yield small molecules with better predicted efficacy. Traditionally, drug discovery relied solely on the chemical and medical intuition of medicinal chemists and the isolation of natural products to make the process more efficient. This slowed drug development since researchers had to devise a structure for the drug, perform a synthesis, purify and characterize the drug, and then finally test the drug in vitro. Advances in structural biochemical techniques have also resolved many important protein structures that have allowed in silico methods to aid the drug discovery process. With the advent of computational chemistry methods, the

drug screening and development pace has increased exponentially. Molecular docking is a quick and simplistic computational chemistry method of virtually screening thousands or millions of compounds to determine potential binding modes of small molecules to proteins. Docking generates possible conformations of each ligand and uses empirical formulas based on experiment to calculate rudimentary binding scores for each pose. Computational drug screening is a fast, safe and inexpensive way to screen millions of drug candidates in a timely manner and may hold the key to finding an effective inhibitor to APOBEC3B (Cui et al., 2020; Kant et al., 2021; Navyashree et al., 2021; Prada-Gracia et al., 2016; Sharma et al., 2020). Furthermore, molecular dynamics (MD) simulations utilize classical Newtonian physics to model the interactions of molecules over time, and can be used by refine protein-ligand docking poses and provide more rigorous energy calculations to quantify intermolecular interactions between a protein and a drug molecule.

In this study, we utilized molecular docking, MD, binding free energy calculations, and *de novo* ligand growth in the pursuit of potential novel competitive inhibitors of APOBEC3B. Approximately 7.8 million molecules from the ZINC12 clean leads database were virtually screened and 211 top-scoring molecules from the database were selected for further analysis using MD. The drugs we have simulated show potential for inhibition and are strong candidates for future *in vitro* assays. Our results are encouraging as they demonstrate substantial progress compared to previously published potential APOBEC3B inhibitors. Continued testing of these potential inhibitors *in vitro* may yield effective inhibition of APOBEC3B and prevent the mutagenic activity exhibited by APOBEC3B.

Methods

A potent inhibitor of APOBEC3B has yet to be discovered, thus a high-throughput *in silico* approach is critical to narrow the chemical space of potential inhibitors. The computational methods utilized in this study included docking 7.8 million molecules obtained from a similarity search on the ZINC15 database using cytosine (which binds to the active site natively) as the target of similarity, then performing atomic-level molecular dynamics simulations and binding free energy calculations on 211 molecules that showed the most promise from docking. Finally, we performed novel *de novo* ligand growth to generate new potential inhibitors and processed those new ligands through the same workflow. We drew from the ZINC12 database (Irwin et al., 2012) for the majority of our ligands and then used those as a starting point for testing against the deoxycytidylate (dCMP) control. The 5CQH PDB[10] was obtained from the Protein Data Bank (Berman et al., 2003) and used as the receptor for the remainder of the project with slight modifications as detailed below.

Protein model preparation

Using MODELLER (Eswar et al., 2006; Fiser et al., 2000; Sali and Blundell, 1993) online loop refinement within Chimera

(Pettersen et al., 2004), we modeled in the missing random coil region spanning amino acids Ala242–Tyr250 (loop 3) in the 5CQH PDB. We used AMBER's *MCPB.py* toolkit (Li and Merz, 2016) to generate parameters for the metal-centered active site, including forming bonds between the catalytic Zn^{2+} ion and Cys284@S, Cys289@S, His353@ND, and the oxygen of a crystallographic water. These modifications allowed for simulations to maintain a reasonable conformation of the protein using the metal center of APOBEC3B.

Virtual screening

PyRx (Small-Molecule Library, 2023) was used to define a grid box and generate pdbqt files for modeling the binding site. The QuickVina algorithm (Alhossary et al., 2015) of the AutoDock Vina software (Eberhardt et al., 2021; Trott and Olson, 2010) was used to perform flexible molecular docking with a docking cubic box of width 25 Å length centered on (X = −1.500, Y = −7.700, Z = 15.000) on the 5CQH PDB after adding the zinc and loop as detailed above. The grid box location was chosen based on its proximity to the catalytic zinc ion. 7.8 million molecules were acquired from the ZINC12 (Irwin et al., 2012) clean leads database and docked to this structure using an exhaustiveness of 8 (default) and scored by the AutoDock Vina scoring function. The default mode of docking was utilized in AutoDock Vina, which outputs the best nine poses for each docked molecule (defaults). For each molecule, we selected the top scoring pose for further analysis. A binding free energy cutoff of −7.8 kcal/mol from QuickVina (Li and Merz, 2016) was utilized to filter the number of candidates proceeding to molecular dynamics and further analysis, resulting in 211 top-ranked ligands simulated (see below).

Molecular dynamics

All protein-ligand complexes were prepared for MD and simulated using the AmberTools (Trott and Olson, 2010) and Amber 20 software package (Case et al., 2022). The 5CQH PDB structure was used as the starting protein conformation for APOBEC3B in all simulations herein. Additionally, the ligand coordinates for each potential inhibitor were taken from the top scoring docking mode from PyRx (see above) in reference to the original 5CQH coordinates. The Amber *ff14SB* force field (Maier et al., 2015) parameters were applied to the protein and GAFF (Wang et al., 2004) parameters were used for the ligands. The *antechamber* module (Maier et al., 2015) in AmberTools was used to generate partial atomic charges for all ligands. We used AmberTools' *tleap* to neutralize the system for each ligand with ∼3 Na^+ ions (to counter the negative overall charge of the protein) and solvated a truncated octahedron periodic box with TIP3P water models (Jorgensen et al., 1983), resulting in a unit cell consisting of ∼23,000 atoms. We used the GPU-accelerated *pmemd* code (Salomon-Ferrer et al., 2013) for Amber 20 to perform MD (Salomon-Ferrer et al., 2013; Weinbach & Elber, 2005). Minimization of the solvated system consisted of seven stages, each comprising a maximum of 5000 steps. The first 1000 steps contained the steepest descent in potential, with the remaining 4000 being of conjugate gradient minimization. The first of the seven stages was given a restraining

force of 10.0 kcal/mol/Å^2 on the heavy atoms of the solute and this was methodically lowered to 0.0 kcal/mol/Å^2 by stage seven. After all seven minimization steps, the solvated box was heated linearly from 10 K to 300 K at 1 atm pressure over 2.0 ns while the solute atoms were again restrained with a restraint weight of 10 kcal/mol/Å^2. During equilibration the restraining force was lowered from 10 to 0.0 kcal/mol/Å^2 every 500 ps over seven stages until fully unrestrained. After equilibration, unrestrained MD was performed at a constant pressure of 1 atm with isotropic scaling, and constant temperature of 300 K was maintained using a Langevin thermostat. Each protein-ligand complex was simulated using a 2-fs timestep with the SHAKE algorithm (Weinbach and Elber, 2005) for at least 250 ns, which was enough to reach a conformational equilibrium (Figures S3–S6). During unrestrained MD, each atom in the ligand was given a random initial velocity. After running unrestrained MD on each of the ligands that passed the docking cutoff score (211 ligands with a docking score of −7.8 kcal/mol or better), we selected the 30 top performing ligands from the ZINC database for further analysis, representing the best ∼15% of ligands simulated for 250 ns. These top 30 ligands were simulated in triplicate (each replicate initiated with different starting atomic velocities) for 250 ns (750 ns total) to determine an a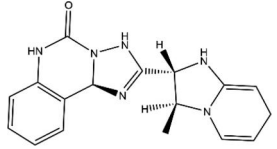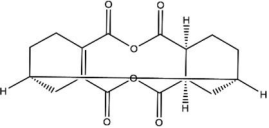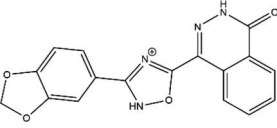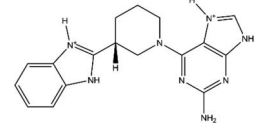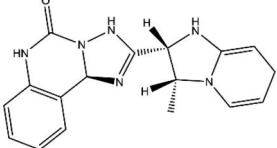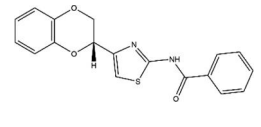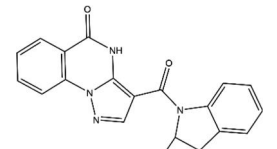verage binding free energy that represents enhanced sampling of the potential energy surface. All initial simulations were performed using identical constraints; longer simulations were used on the best performing ligands to further analyze their behaviors.
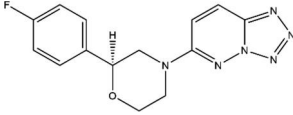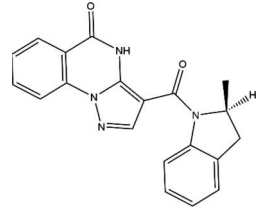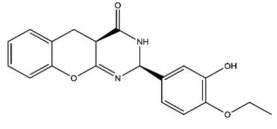
Molecular dynamics analysis

Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) (Sun et al., 2014) calculations were performed using Amber's *MMPBSA.py* (Miller et al., 2012) to determine the average binding free energy scores from the MD trajectories for all ligands. Binding free energies (kcal/mol) were calculated every 0.1 ns (all recorded frames) using the GB model developed by A. Onufriev, D. Bashford and D.A. Case (Onufriev et al., 2000; 2004) in conjunction with *bondi* radii parameters. The binding free energies were further decomposed on a per-residue basis to quantify each residue's contribution to protein-ligand binding for all simulations. The analysis also parses out the energies based on contribution type, such as electrostatic, van der Waals, polar solvation, and non-polar solvation energies. All other parameters for *MMPBSA.py* were set to default values.

We also used the *cpptraj* package (Roe and Cheatham, 2013) included in AmberTools to calculate and compare the

Table 1. Top 10 docked ZINC12 molecules.

ZINC ID	Docking Score (kcal/mol)	2D Structure	Allosteric or Catalytic	ZINC ID	Docking Score (kcal/mol)	2D Structure	Allosteric or Catalytic
29528216	−8.5		A	39741670	−8.2		A
40448341	−8.5		C	73637923	−8.2		A
47194834	−8.4		A	20901837	−8.2		A
95529176	−8.3		A	39741667	−8.2		A
06703437	−8.2		C	47194833	−8.1		A

The molecules depicted were docked and discovered to have the lowest binding free energy according to the AutoDock Vina scoring function, as well as an allosteric or catalytic designation based on the location of the final docked pose.

root-mean-squared deviations (RMSD), root-mean-squared residue fluctuations (RMSF), hydrogen bonding analysis, and principal component analysis (PCA) for the simulations. The VMD software (Humphrey et al., 1996) was used for visualization and generating structural images (Table 1).
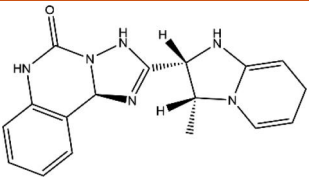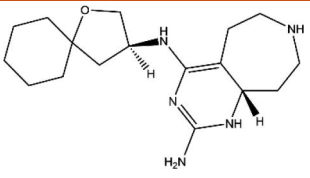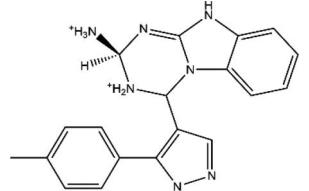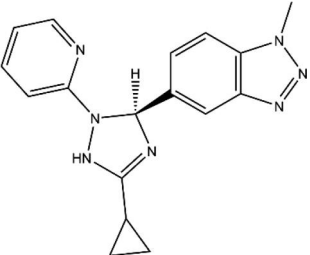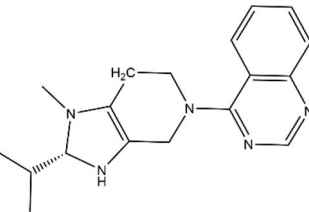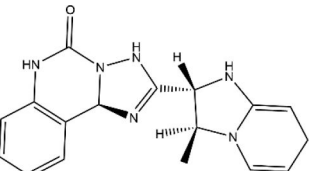
Pharmokinetics and drug-likeness calculations were performed on the top scoring inhibitors from MD used the SwissADME web server (Daina et al., 2017).

2.6. DOCK 6 and de novo design

DOCK 6 (Allen et al., 2017) was used to create a model of APOBEC3B that would act as a receptor for de novo ligand growth, which creates a series of novel ligands specific to the receptor structure that could serve as potential inhibitor candidates. The best-performing catalytic lig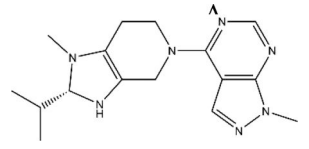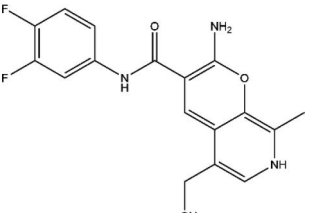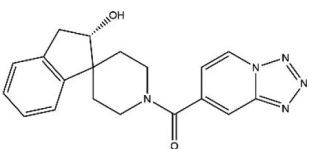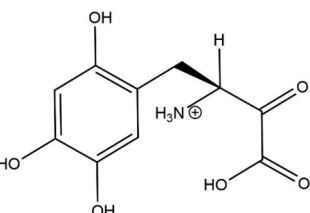and was determined through MM-GBSA post-processing of MD simulations (above). The ligand, pre-screened catalytic ligand ZINC09338246 (Table 2) from the ZINC12 database, was used as a model to help DOCK 6 establish a proper docking site. The sphere selection was constrained to 1.4 Å within the active site on the receptor to prevent over saturation of ligand growth. The fragment library used for de novo ligand growth was propagated from a PDB compilation of 25 of the top-scoring ligands from the screened ZINC12 ligands that were simulated in triplicate for 250 ns. After de novo ligand growth was performed, the top 30 newly generated, grown ligands were each simulated for 250 ns once bound to

Table 2. Top 10 ZINC12 ligands from MD.

Name or ZINC ID	MMGBSA Binding Energy (kcal/mol)	2D Structure	Allosteric or Catalytic (A/C)	Name or ZINC ID	MMGBSA Binding Energy (kcal/mol)	2D Structure	Allosteric or Catalytic (A/C)
39741667	−33.7		A	72478471	−28.1		A
9338246	−33.3		C	65494139	−24.9		A
69493336	−30.8		A	39741670	−23.2		A
69512841	−29.2		A	03911065	−21.9		A
72152602	−28.2		A	MN23	−22.9		C

A list of the top 10 candidates from the ZINC12 database ordered by average MM-GBSA binding free energy, as well as their locations in the allosteric or catalytic binding sites. MN23 originates from an *in vitro* study by Olson et al. (n.d.) that we used as a control to compare with our proposed drug compounds.

APOBEC3B (following the same procedure above) for comparison to the ZINC12 ligands. All other parameters were used as default based on available DOCK 6 and *de novo* protocols.

Results and discussion

In this study, we performed the first high-throughput virtual screening of the APOBEC3B enzyme in search of a novel inhibitor that could develop into a future cancer treatment. After docking 7.8 million clean lead molecules from the ZINC12 database, we simulated the top-ranking 211 compounds bound to the APOEBC3B active site using molecular dynamics and simulated the top 30 compounds in triplicate for 250-ns each. These compounds were then further narrowed using average binding free energy calculations to the top 25 potential inhibitors, which were used by DOCK 6 to generate a fragment library for *de novo* structure-based drug design. The best 30 de novo drugs from DOCK 6 were also simulated and analyzed using molecular dynamics and free energy calculations. Overall, we report several novel potential inhibitors (both purchasable and/or synthesizable) specific to APOBEC3B that may lead to the development of new cancer treatments.

Docking results

The ZINC clean leads database was used to access structures for 7.8 million molecules, which were virtually screened using QVina. Table 1 depicts the top 10 drug candidates along with their corresponding docking scores (Table S1 provides structures and docking scores for all 211 docked compounds that met the −7.8 kcal/mol cutoff criteria). The allosteric binding site described in Table 1 is located primarily in the space between Asp260 and Leu238, ∼12.5 Å from the Zn^{2+} in the catalytic site (Figure 1A). There are clear similarities in functional groups present among the ligand structures, with exceptions such as ZINC40448341. One of these similarities is nitrogen-containing π systems, which is unsurprising considering the natural substrate is a pyrimidine (cytosine). The active site was decided based on proximity to a zinc ion parameterized for the system, however some ligands docked to a pocket adjacent to the active site and thus would not be considered competitive inhibitors.

Upon inspection, ZINC29528216 has key interactions with Arg257 and Arg252 through hydrogen bonding interactions. ZINC40448341 shows strong interactions with Tyr313, His253, Arg211, and Asn240 inside the catalytic site. Tyr313 shows *in vitro* binding relevance by having important π-stacking interactions with the cytosine base in the active site as well as helping create a hydrophobic pocket for the −1 base in the ssDNA chain (Shi et al., 2017). His253 shows binding relevance as well through similar π-stacking interactions with the cytosine in the active site, proving crucial to binding in addition to Tyr313 (Shi et al., 2015). Asn240 also proves to be important by hydrogen bonding with the backbone of the cytosine base *in vitro* (Shi et al., 2017). However, it is important to note that neither of these

ligands performed well enough to score in the top 10 ZINC12 ligands after modeling them with molecular dynamics simulations (see below).

Molecular dynamics simulations

After docking, molecular dynamics simulations were used to provide a more accurate and dynamic view of the interactions involved in these protein-drug complexes. As noted in the Methods, a Qvina docking score cutoff of −7.8 kcal/mol was established to filter out less promising candidates and allow for further analysis with the most promising candidates. This resulted in 211 docked molecules being analyzed using MD. Table S1 depicts the structures and binding free energy results for all simulated molecules. A table of the top 10 performing candidates from our MD analysis is shown below (Table 2), and it is important to note that not all of these candidates match with those of the top 10 docked candidates. That is to be expected due to the rigor and increased sampling of MD simulations when compared to docking. However, there are some conserved candidates between these tables such as ZINC39741667 and ZINC39741670. These two ligands are particularly similar in structure, and thus the consistency in their results across docking and MD helps give credibility to the consistency of our methodology. The structures of the most promising ligands from docking similarly resemble those from MD, such as the nitrogen ring systems and the carbonyls that resemble the natural substrate (cytosine) of APOBEC3B. Many inhibitors seek to emulate the natural substrate due to the specificity of the binding site, so this similarity is expected.

A control for this project was established by simulating a single dCMP nucleotide. This control dCMP was removed from a larger crystal structure of ssDNA (PDB 5TD5), docked to APOBEC3B, and simulated in triplicate using the same procedure detailed below for 250 ns. This control scored an average of −34.09 kcal/mol, though we recognize that an extended ssDNA sequence would likely have a more potent binding affinity. One molecule (MN23) listed in the table is a drug previously described by Olson et al. who tested compounds *in vitro*, which resulted in an IC_{50} value of 0.15 μM (Olson, n.d.) This drug was also used as a comparison with these newly tested ZINC and *de novo* ligands and the top ZINC candidates from this project scored up to 10 kcal/mol more favorably than this comparison drug ($\Delta G_{bind} = -22.9$ kcal/mol).

Our MD results from the molecules screened from the ZINC database suggest ZINC09338246 is the second most promising inhibitor of APOBEC3B. PCA results on the top four normal modes (Figure S11) from MD displayed a stable enzyme core near the binding pocket, while the solvent exposed loops were more dynamics. Figure 2 shows the binding and efficiency of ZINC09338246 from the 250 ns MD simulation. Figure 2A shows a representative conformation of the APOBEC3B-ZINC09338246 complex from MD, depicting how deeply the molecule binds within the active site. A free energy decomposition analysis (Figure 2D) of ZINC09338246 bound to APOBEC3B for 250 ns shows that the amino acids
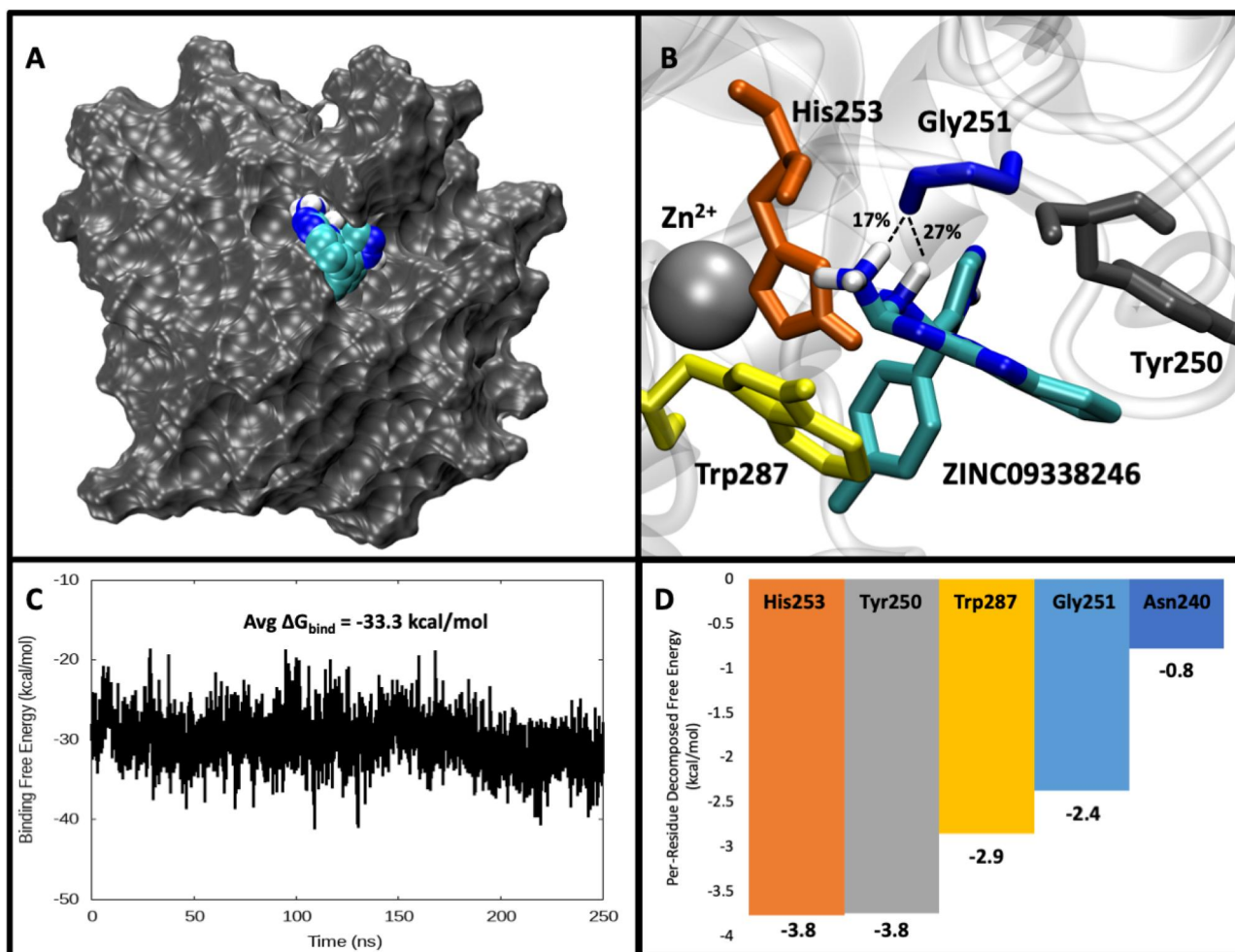
Figure 2. Comprehensive analysis of ZINC09338246 MD. A) 3D surface representation of the protein-ligand complex with the surface of APOBEC3B in gray and ZINC09338246 shown in VDW representation. B) Representation of the ZINC09338246 binding pocket showing the catalytic Zn^{2+} ion and top amino acid residues contributing to binding: His253 (orange), Tyr250 (gray), Trp287 (yellow), and Gly251 (blue). the dotted lines represent the hydrogen bonds observed during the simulation between Gly251 and ZINC09338246, along with the percent occupancy of those hydrogen bonds during the simulation. C) a graph of the binding free energy (kcal/mol) during the 250 ns APOBEC3B-ZINC09338246 MD simulation with the average binding free energy shown. D) Depiction of the decomposed free energy (kcal/mol) of each of the top five amino acids from APOBEC3B that contribute to binding. Additionally, Figure S4 depicts the root-mean-squared deviation and fluctuation of the ZINC09338246, which indicates the stability of the ligand. Figure S8 depicts a table of values relevant to the hydrogen bonding as well as an accompanying 2D structure to assist with atom naming conventions. Finally, Table S7 features an ADMET prediction which reports 0 violations in the Lipinski rules for druglikeness and a log[P] = −0.06 which indicates a slight difficulty in crossing the cell membrane once administered. It should also be noted that while Asn240 is mentioned in panel D, it was removed from the image in panel B due to obstruction of view.

that contribute most to binding are the His253 and Tyr250 primarily through π-stacking interactions. His253 is important for *in vitro* π-stacking with the cytosine base (Shi et al., 2015; 2017), thus ZINC09338246 is replacing a key interaction with the native ssDNA. Trp287 interacts much the same way, while Gly251 and Asn240 contribute more through hydrogen bonding and dipole-dipole interactions (Figure 2B). Trp287 is structurally important to APOBEC3B by stabilizing the closed loop conformation of APOBEC3B (Shi et al., 2015), making Trp287 a logical amino acid to target for inhibition. Along with Trp287, Asn240 is another important residue hypothesized to be interacting well with the C2 carbonyl group of the cytosine base *in vitro* (Shi et al., 2015), again demonstrating the potential potency of ZINC09338246 by directly competing with the natural ligand in the catalytic site. Based on the binding free energy throughout the simulation (Figure 2C), the ligand appears to change conformation slightly, resulting in a marginally more potent bound conformation at

~190 ns (Figure 2C). Through visual inspection, the ligand was shown to begin forming a water bridge with Gly248 through hydrogen bonding with the ligand for ~25% of the simulation, which correlates with the decrease in binding free energy observed in Figure 2C at ~190 ns.

ZINC39741667 was determined to be the strongest binding ZINC molecule from this study based on binding free energy calculations from MD (Figure 3). ZINC39741667 is observed as binding in an allosteric site (Figure 3A), as reported in Table 2. Thus, ZINC39741667 should not be considered a competitive inhibitor, but still demonstrates promise as a potentially effective drug candidate. Furthermore, the amino acids interacting with ZINC39741667 (Figure 3B) help depict the difference in binding pockets as none of these amino acids are present in the decompositions of the other top molecules from ZINC or *de novo*, which are binding in the catalytic active site. Arg257 stands out as the most stabilizing amino acid in this simulation of this protein-ligand
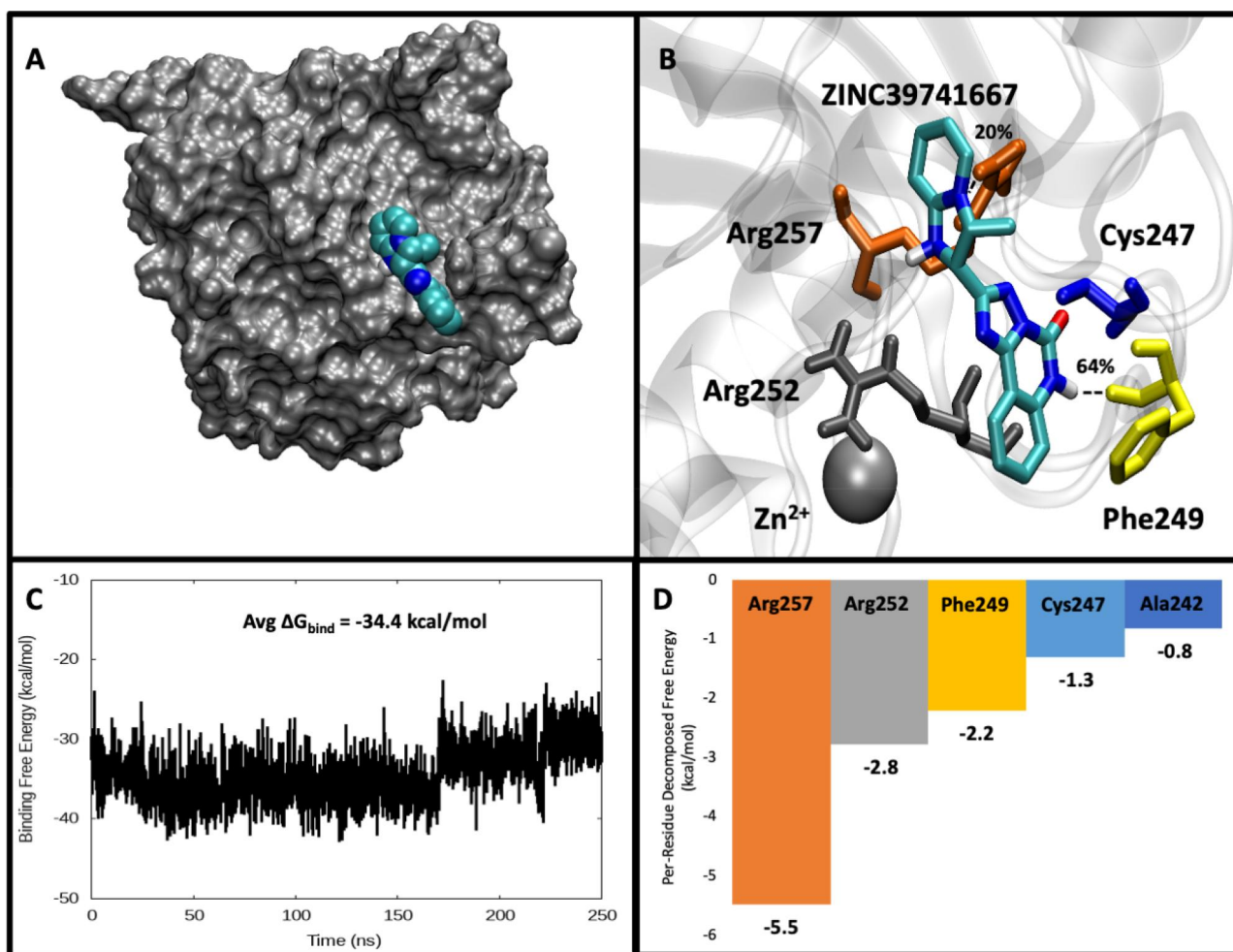
Figure 3. Comprehensive analysis of ZINC39741667 MD. A) 3D surface representation of the protein-ligand complex with the surface of APOBEC3B in gray and ZINC39741667 shown in VDW representation. B) Representation of the ZINC09338246 binding pocket showing the catalytic Zn^{2+} ion and top amino acid residues contributing to binding: Arg257 (orange), Arg252 (gray), Phe249 (yellow), and Cys247 (blue). C) a graph of the binding free energy (kcal/mol) during the 250 ns APOBEC3B-ZINC39741667 MD simulation with the average binding free energy shown. D) Depiction of the decomposed free energy (kcal/mol) of each of the top five amino acids from APOBEC3B that contribute to binding. Additionally, Figure S4 depicts the root-mean-squared deviation and fluctuation of the ZINC39741667, which indicates the stability of the ligand. Figure S8 depicts a table of values relevant to the hydrogen bonding as well as an accompanying 2D structure to assist with atom naming conventions. Finally, Table S7 features an ADMET prediction which reports 0 violations in the Lipinski rules for druglikeness and a log[P] = 1.80 which indicates a drug that absorbs well orally and intestinally without being sequestered into adipocytes.
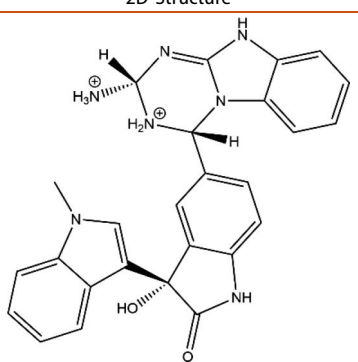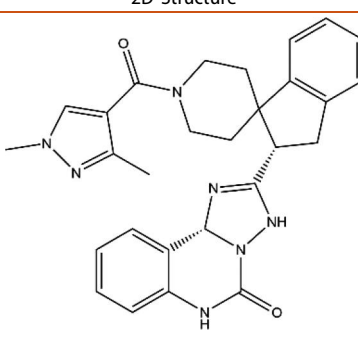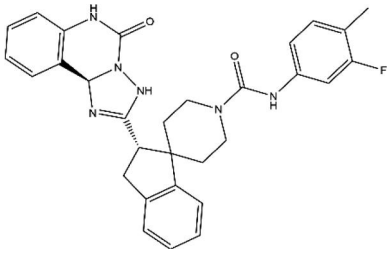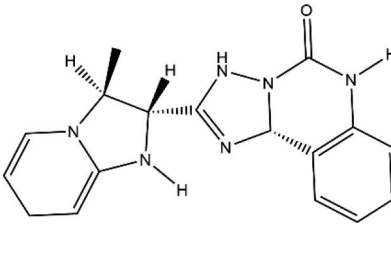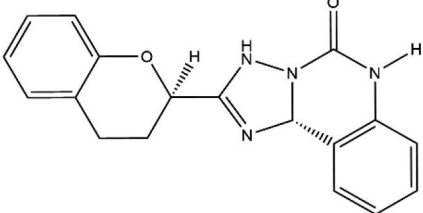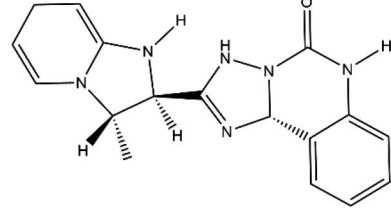
complex; this is primarily due to the dipole interactions between the sidechain and the tertiary amide atoms present in the molecule. Arg252 has similar interactions between the sidechain and the oxygen-nitrogen fragment of the molecule. Phe249 contributes in two considerable ways that influence ligand binding, one being a π-stacking interaction and the other is a hydrogen bond between the backbone and the ZINC39741667 (Figure 3B). Shi et al. also observed important protein-substrate interactions in this region (amino acids 242–255), such as direct or water-mediated hydrogen bonds with the backbone phosphate of the +1 nucleotide (Shi et al., 2017). The binding free energy begins to worsen near the end of the simulation (Figure 3C), which is indicative of the molecule becoming less stable in the binding pocket. This change in stability can be accounted for by the π-stacking interaction breaking around 170 ns and by making the molecule less stable for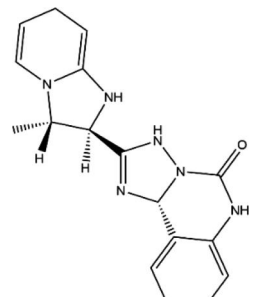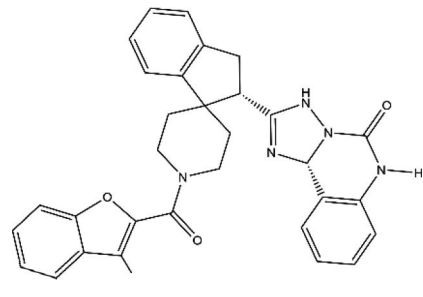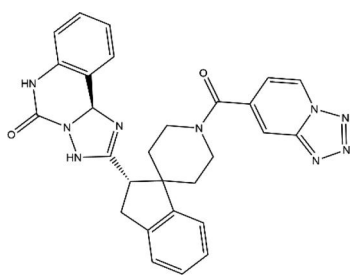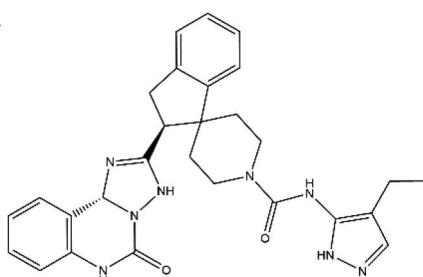 the rest of the simulation. PCA results on the top four normal modes (Figure S12) from MD suggest the structure of APOBEC3B is most stable when

bound to ZINC39741667 compared with the other potential inhibitors analyzed (Figures S11–S14). While this ligand has a better average binding affinity than ZINC09338246, it should be noted that it is an allosteric inhibitor, and it does become less energetically stable toward the end of the simulation, suggesting potentially greater instability than the previous ligand at larger time frames.

De novo design

Using *de novo* ligand design with DOCK 6, 919 ligands were generated from the fragment library created using the top performing 25 ZINC12 clean leads. The DOCK 6 software was used to dock these novel ligands to the APOBEC3B active site, and the top 30 were then simulated with MD (Table S2). However, while the protocol for MD and MM-GBSA calculations were identical to the ZINC12 clean leads, the DOCK 6 docking scores are not directly comparable to the QVina

Table 3. Top 10 *de novo* ligands.

Name	2D Structure	MMGBSA Binding Energy (kcal/mol)	Name	2D Structure	MMGBSA Binding Energy (kcal/mol)
DN29		−46.0	DN06		−23.4
DN08		−29.4	DN24		−22.3
DN28		−26.2	DN26		−21.6
DN19		−25.8	DN02		−22.0
DN04		−25.1	DN14		−21.6

A list of the top 10 performing *de novo* ligands found in this study ordered by most favorable average MM-GBSA binding free energy to least favorable. All *de novo* ligands docked to the same binding site (catalytic) as ZINC09338246.

scores due to the difference in scoring functions, and thus the DOCK 6 docking scores have been omitted from Table 3 (and Table S2) for clarity. As shown in Table 3, several of the *de novo* ligands (DN08, DN29, etc.) have a conserved fragment involving a purine-like ring that is a direct product of

the *de novo* ligand growth process using ZINC09338246 as one of the model ligands in the fragment library generation. The 30 most promising ligands from DOCK 6 and *de novo* ligand design were simulated using Amber (see results below). These top 30 ligands docked in the same vicinity of
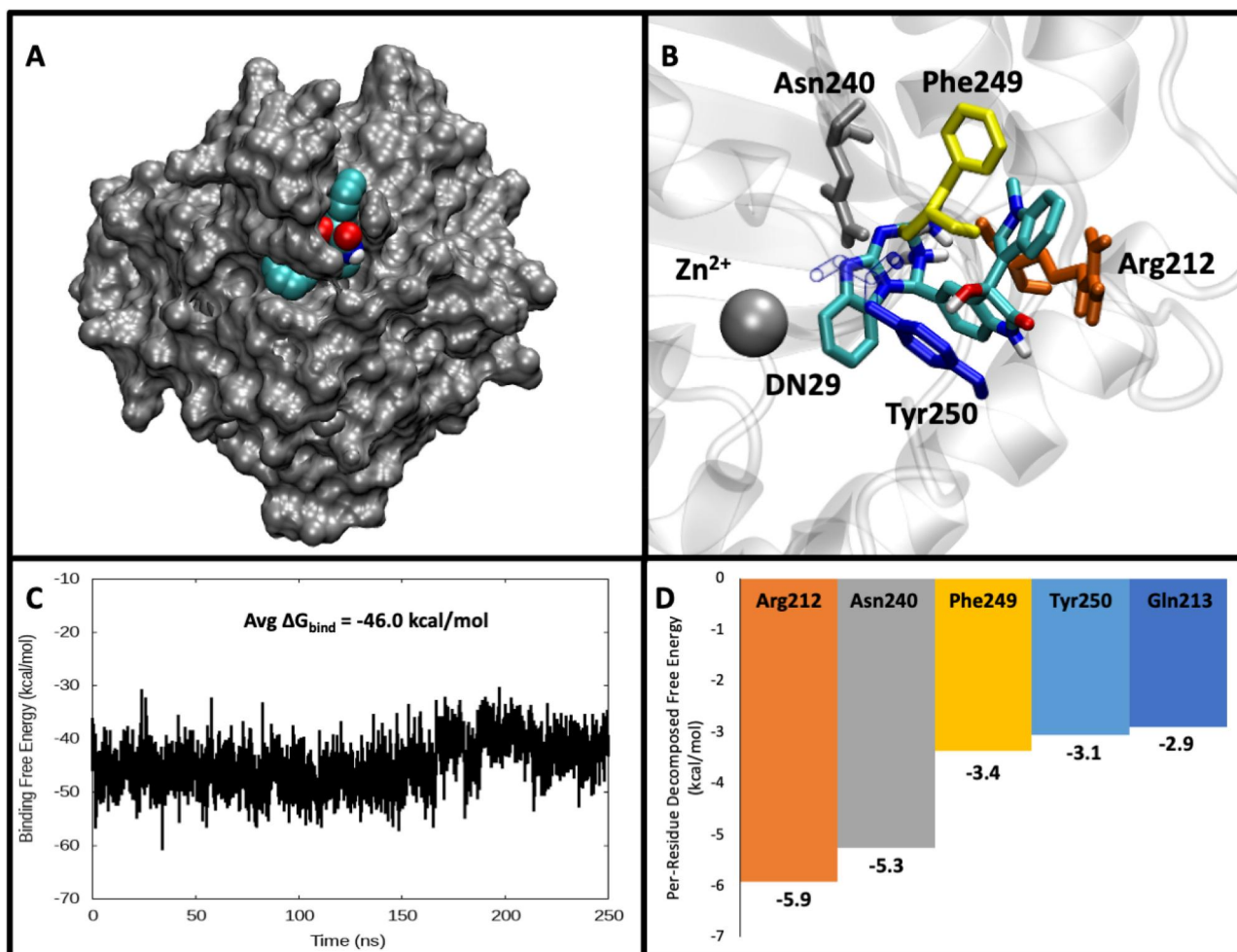
Figure 4. Comprehensive analysis of DN29 MD. A) 3D surface representation of the protein-ligand complex with the surface of APOBEC3B in gray and DN29 shown in VDW representation. B) Representation of the DN29 binding pocket showing the catalytic Zn^{2+} ion and top amino acid residues contributing to binding: Arg212 (orange), Asn240 (gray), Phe249 (yellow), and Tyr250 (blue). C) a graph of the binding free energy (kcal/mol) during the 250 ns APOBEC3B-DN29 MD simulation with the average binding free energy shown. D) Depiction of the decomposed free energy (kcal/mol) of each of the top five amino acids from APOBEC3B that contribute to binding. Additionally, Figure S4 depicts the root-mean-squared deviation and fluctuation of the DN29, which indicates the stability of the ligand. Figure S8 depicts a table of values relevant to the hydrogen bonding as well as an accompanying 2D structure to assist with atom naming conventions. Finally, Table S7 features a table with ADMET predictions with 0 violations of the Lipinski rules for druglikeness and a log[P] = −0.37 which indicates a slight difficulty in crossing the cell membrane once administered.

the catalytic binding site as the conserved fragment we see with ZINC09338246 and are typically interacting with Arg212 through hydrogen bonding interactions.

DN29 is clearly the best scoring molecule produced from both the ZINC database screening and *de novo* design methodology. DN29 shares a similar structure and chemical features with ZINC09338246, especially regarding the imidazole ring that sits deep in the catalytic site. DN29 scored on average −12 kcal/mol lower than the other top performing molecules from either methodology and that is due to several compounding factors. An important structural part of this molecule is the guanidinium fragment that interacts well with the aromatic structures around it, such as Trp287 and Tyr250. These residues show high-affinity π-stacking interactions with this fragment throughout the simulation and both contribute heavily to the binding free energy. Another important part of this molecule is the presence of polar atoms that interact most with Arg212 and Asn240. These polar interactions heavily

contribute to the binding free energy as seen in Figure 4D. These fragments were all pulled from the library created from the ZINC12 database molecules, which helps explain the structural similarity to molecules such as ZINC09338246.

Molecular dynamics on de novo molecules

The ZINC12 clean leads yielded no candidates that outperformed the control dCMP, however the results from *de novo* ligand growth were more promising. The control dCMP scored an MM-GBSA average binding free energy of −34.09 kcal/mol over the course of running 250 ns MD in triplicate, and the *de novo* candidate, DN29, scored an average of −46.03 kcal/mol. Figure 4 shows the binding efficiency of DN29 from the 250 ns MD simulation. Figure 4A shows a representative conformation of the APOBEC3B-DN29 complex from MD, depicting how deeply the molecule binds within
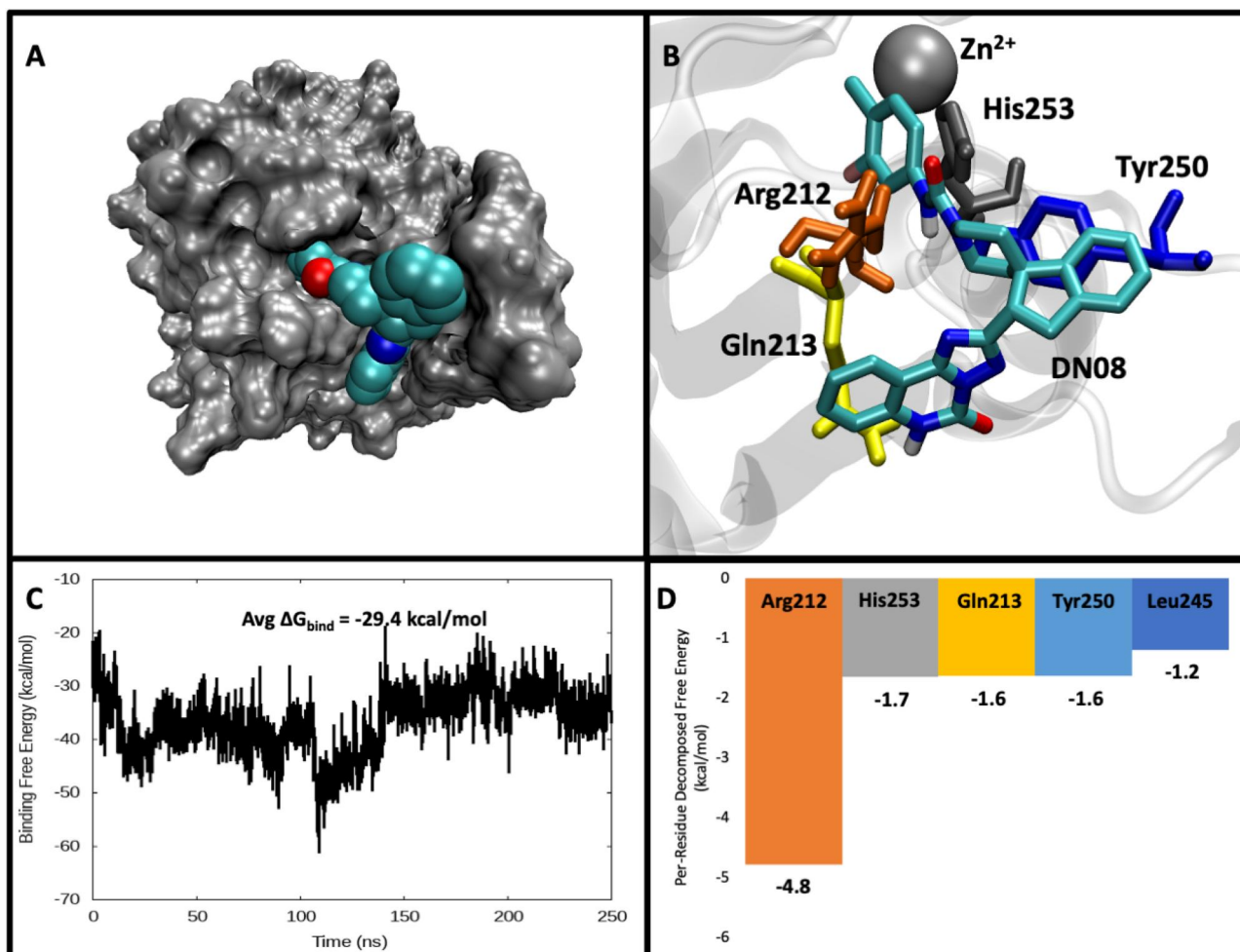
Figure 5. Comprehensive analysis of DN08 MD. A) 3D surface representation of the protein-ligand complex with the surface of APOBEC3B in gray and DN08 shown in VDW representation. B) Representation of the DN08 binding pocket showing the catalytic Zn^{2+} ion and top amino acid residues contributing to binding: Arg212 (orange), His253 (gray), Gln213 (yellow), and Tyr250 (blue). C) a graph of the binding free energy (kcal/mol) during the 250-ns APOBEC3B-DN08 MD simulation with the average binding free energy shown. D) Depiction of the decomposed free energy (kcal/mol) of each of the top five amino acids from APOBEC3B that contribute to binding. Additionally, Figure S4 depicts the root-mean-squared deviation and fluctuation of the DN08, which indicates the stability of the ligand. Figure S8 depicts a table of values relevant to the hydrogen bonding as well as an accompanying 2D structure to assist with atom naming conventions. Finally, Table S7 features ADMET predictions with 2 violations of the Lipinski rules for druglikeness (MW > 500, MLOGP > 4.15) and a log[P] = 3.85 which indicates a preference for being sequestered into adipocytes without reaching its target cell.

the catalytic site as well as how loop 3 of the protein closes over the molecule. A free energy decomposition analysis (Figure 4D) of DN29 bound to APOBEC3B for 250 ns shows that the amino acids that contribute most to binding are Arg212 and Asn240. The interaction with Arg212 is primarily due to the interaction between the carbonyl oxygen of DN29 and the sidechain polar hydrogens of Arg212. The interaction between Asn240 and the molecule primarily relies on the dipole-dipole interaction between the DN29 and Asn240. Asn240 is important for *in vitro* as it has been shown to interact with the C2 carbon of the backbone of the cytosine in ssDNA (Shi et al., 2017). Gln213, is also one of the top contributing amino acids (Figure 4D), and has been theorized to be an important amino acid residue *in vitro* due to its preference for adenosine at the +1 position (Shi et al., 2017). Each of these residues are critical and would serve as rational targets for future drug design. Based on the binding free energy throughout the simulation (Figure 4C), the ligand appears to be energetically stable, as well as structurally stable within the binding pocket, which indicates a stable

binding conformation throughout the simulation. In fact, the binding free energy appears to be trending downwards as time progresses, indicating a stabilizing binding affinity. Through visual inspection, the increase in binding free energy beginning at ~190 ns can be attributed to loop 3 detaching from DN29, while the ligand remains stable in the binding site. However, the loop does return to its initial position interacting with DN29 later in the simulation. PCA results (Figure S13) from MD suggest the structure of APOBEC3B shares dynamic similarities (though greater magnitude) when bound to DN29 compared with ZINC39741667 (Figure S12) for all four of the top normal modes.

DN08 is the second most potent inhibitor found using *de novo* design. PCA results on the top four normal modes (Figure S14) from MD suggest the binding pocket of APOBEC3B near the Zn^{2+} ion is most dynamic when bound to DN08 compared with our other analyzed ligands (Figure S11–S13). Figure 5A shows the protein-ligand complex and the most notable difference from DN29 is the lack of protein enclosure of the ligand. This gives an initial indicator of the

lower binding free energy observed compared to DN29. DN08 did not perform as well as DN29, and that is clear based on the much more erratic binding free energy during the simulation (Figure 5C) relative to DN29 (Figure 4C). Figure 5B shows the relevant interactions between the protein and DN08. DN08 does share a strong π-cation interaction between itself and Arg212 with DN29; this interaction is consistent among the best performing ligands in this study. We hypothesize Arg212 could become one of the most important amino acids to target in future drug trials. DN08 also has interactions with the Tyr250, similar to DN29 and ZINC09338246, which makes it another important target residue for drug design. His253 and Tyr250 are two other reoccurring residues shown in Figure 5D which, according to Shi et al., contribute via important direct or water-mediated hydrogen bonds with the backbone phosphate of the +1 nucleotide between the nucleotide and amino acids 242–255 (Shi et al., 2017). One amino acid that does not appear in Figure 5D is Tyr313, which forms a T-shaped π-stacking interaction with the cytosine in the active site in vitro, as well as interacting with the −1 thymine in the ssDNA chain (Shi et al., 2017). Tyr313 actually has very similar contribution to the binding free energy (–1.0 kcal/mol) as Leu245, but falls just short of the top five interacting residues shown (Figure 5D). The amide oxygen on DN08 hydrogen bonds with the side chain hydroxyl of Tyr313 for ∼80% of the simulation, and a halogen-π interaction also stabilizes the DN08 fluorine near Tyr313 for much of the simulation, similar to the interaction Tyr313 has with the natural substrate in vitro. DN08 is among the most promising molecules in our study, and it helps to illuminate the importance of Arg212 for inhibitor binding.

Overall, our MD results from the molecules created in the de novo design process suggest that DN29 is the most promising inhibitor of APOBEC3B. DN29 is a de novo grown ligand that has not yet been synthesized to our knowledge. However, future work will investigate the most feasible synthetic pathway for DN29. Currently, a synthetic pathway for DN29 is underway through the Olson lab at Truman State University. Another important note to make is that we are currently unaware of any catalytic activity the enzyme may have on DN29 itself in the active site, since it, too, has an amine group deep in the binding site near the zinc atom, which coordinates the cytidine deaminase reaction in vivo. These studies will be investigated more in future work as well to determine an IC$_{50}$ value for DN29 and to discover the impact of DN29 on the catalytic efficiency of APOBEC3B in vitro. However, this drug remains a promising scaffold for future APOBEC3B drug development. DN08 is another promising inhibitor for APOBEC3B and focuses on other related interactions between the protein and bound ssDNA that could be important for inhibitor design. De novo ligand design has yielded the most promising drugs in this study and has provided successful results in only one iteration. With successive iterations of de novo ligand design and more expansive libraries, de novo ligand design has the promise to generate more potent drug candidates in the future for APOBEC3B.

Conclusion

To help treat cancer proliferation, we have searched for and designed a small molecule inhibitor for APOBEC3B. This small molecule inhibitor could be another tool in the fight against tumorous cancers and provides us with a straightforward method for treating these cancers. Previous small molecule studies have proved largely ineffective when testing in vitro with IC$_{50}$ values in the μM range, which is not comparable to marketable drug candidates. Therefore, our work in creating a well-performing drug candidate is a huge step in the fight against APOBEC3B mutagenesis. The ultimate goal of this project is to provide a scaffold for further drug development and reveal the most important amino acids to bind to when designing an APOBEC3B-specific inhibitor.

We are the first to conduct high-throughput virtual screening and MD on this scale as well as the first to utilize de novo ligand design for APOBEC3B. These analyses have shown that a small molecule inhibitor for APOBEC3B is plausible and could allow for further testing on their efficacy against APOBEC3B. The ZINC database has provided a functional scaffold in the way of ZINC09338246 that could compete with the natural substrate in some capacity and was used to generate promising novel potential inhibitors using de novo ligand design. Our use of de novo ligand growth has yielded a potentially successful, and synthesizable, drug candidate to test in vitro. Certain amino acid targets seem to be essential as the top de novo candidates interact with similar residues as well as the top performing ZINC ligand. Through binding free energy decompositions using MMPBSA.py these residues appear to be Arg212, Gln213, and Tyr250, which makes them the most rational targets for future drug trials and design. We believe the small molecules described herein provide the best-known lead molecules and scaffolds for future drug design efforts for APOBEC3B, and by utilizing computational chemistry methods we were able to reasonably screen millions of compounds in chemical space to narrow down the search for potent APOBEC3B inhibitors.

Future work would include performing syntheses and in vitro inhibition assays (Nair and Rein, 2014) on the top ZINC and de novo candidates. In these studies, we would test the efficacy of these drugs on APOBEC3B along with enantiomeric studies of the most potent candidates and determine how these drugs affect APOBEC3B catalysis using fluorescence methods as described by Nair (Nair and Rein, 2014). This project has provided future drug designers with a scaffold from which to continue to build ligands and has provided a list of important residues to consider when designing an inhibitor of APOBEC3B.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alhossary, A., Handoko, S. D., Mu, Y., & Kwoh, C.-K. (2015). Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*, *31*(13), 2214–2216. https://doi.org/10.1093/bioinformatics/btv082

Allen, W. J., Fochtman, B. C., Balius, T. E., & Rizzo, R. C. (2017). Customizable de Novo design strategies for DOCK: Application to HIVgp41 and other therapeutic targets. *Journal of Computational Chemistry*, *38*(30), 2641–2663. https://doi.org/10.1002/jcc.25052

Asaoka, M., Ishikawa, T., Takabe, K., & Patnaik, S. K. (2019). APOBEC3-mediated RNA editing in breast cancer is associated with heightened immune activity and improved survival. *International Journal of Molecular Sciences*, *20*(22), 5621. https://doi.org/10.3390/ijms20225621

Barzak, F. M., Harjes, S., Kvach, M. V., Kurup, H. M., Jameson, G. B., Filichev, V. V., & Harjes, E. (2019). Selective inhibition of APOBEC3 enzymes by single-stranded DNAs containing 2′-Deoxyzebularine. *Organic & Biomolecular Chemistry*, *17*(43), 9435–9441. https://doi.org/10.1039/c9ob01781j

Berman, H., Henrick, K., & Nakamura, H. (2003). Announcing the Worldwide Protein Data Bank. *Nature Structural & Molecular Biology*, *10*(12), 980–980. https://doi.org/10.1038/nsb1203-980

Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I. Y., Berryman, J. T., Brozell, S. R., Cerutti, D. S., Cheatham, I. I. I., Cisneros, T. E., Cruzeiro, G. A., Darden, V. W. D., Duke, T. A., Giambasu, R. E., Gilson, G., Gohlke, M. K., Goetz, H., Harris, A. W., Izadi, R., Izmailov, S., … Amber, P. A. (2022). Jr 2022,

CDCBreastCancer. (2023). *An Update on Cancer Deaths in the United States*. Centers for Disease Control and Prevention. Retrieved March 5, 2023, from https://www.cdc.gov/cancer/dcpc/research/update-on-cancer-deaths/index.htm.

Cryo-EM structure of the EBV ribonucleotide reductase BORF2 and mechanism of APOBEC3B inhibition. Science Advances. Retrieved March 5, 2023, from https://www.science.org/doi/10.1126/sciadv.abm2827.

Cui, W., Aouidate, A., Wang, S., Yu, Q., Li, Y., & Yuan, S. (2020). Discovering anti-cancer drugs via computational methods. *Frontiers in Pharmacology*, *11*, 733. https://doi.org/10.3389/fphar.2020.00733

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*(1), 42717. https://doi.org/10.1038/srep42717

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. *Journal of Chemical Information and Modeling*, *61*(8), 3891–3898. https://doi.org/10.1021/acs.jcim.1c00203

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M.-Y., Pieper, U., & Sali, A. (2006). Comparative protein structure modeling using modeller. *Current Protocols in Bioinformatics*, *Chapter 5*. Unit-5.6. https://doi.org/10.1002/0471250953.bi0506s15

Fiser, A., Do, R. K., & Sali, A. (2000). Modeling of loops in protein structures. *Protein Science*, *9*(9), 1753–1773. https://doi.org/10.1110/ps.9.9.1753

Harris, R. S. (2015). Molecular mechanism and clinical impact of APOBEC3B-catalyzed mutagenesis in breast cancer. *Breast Cancer Research*, *17*(1), 8. https://doi.org/10.1186/s13058-014-0498-3

Hou, S., Lee, J. M., Myint, W., Matsuo, H., Kurt Yilmaz, N., & Schiffer, C. A. (2021). Structural basis of substrate specificity in human cytidine deaminase family APOBEC3s. *The Journal of Biological Chemistry*, *297*(2), 100909. https://doi.org/10.1016/j.jbc.2021.100909

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., & Coleman, R. G. (2012). ZINC: A free tool to discover chemistry for biology. *Journal of Chemical Information and Modeling*, *52*(7), 1757–1768. https://doi.org/10.1021/ci3001277

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935. https://doi.org/10.1063/1.445869

Kant, K., Rawat, R., Bhati, V., Bhosale, S., Sharma, D., Banerjee, S., & Kumar, A. (2021). Computational identification of natural product leads that inhibit Mast Cell Chymase: An exclusive plausible treatment for Japanese Encephalitis. *Journal of Biomolecular Structure & Dynamics*, *39*(4), 1203–1212. https://doi.org/10.1080/07391102.2020.1726820

King, J. J., Borzooee, F., Im, J., Asgharpour, M., Ghorbani, A., Diamond, C. P., Fifield, H., Berghuis, L., & Larijani, M. (2021). Structure-based design of first-generation small molecule inhibitors targeting the catalytic pockets of AID, APOBEC3A, and APOBEC3B. *ACS Pharmacology & Translational Science*, *4*(4), 1390–1407. https://doi.org/10.1021/acsptsci.1c00091

Kouno, T., Silvas, T. V., Hilbert, B. J., Shandilya, S. M. D., Bohn, M. F., Kelch, B. A., Royer, W. E., Somasundaran, M., Kurt Yilmaz, N., Matsuo, H., & Schiffer, C. A. (2017). Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nature Communications*, *8*(1), 15024. https://doi.org/10.1038/ncomms15024

Li, P., & Merz, K. M. Jr. (2016). MCPB.Py: A python based metal center parameter builder. *Journal of Chemical Information and Modeling*, *56*(4), 599–604. https://doi.org/10.1021/acs.jcim.5b00674

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). Ff14SB: Improving the accuracy of protein side chain and backbone parameters from Ff99SB. *Journal of Chemical Theory and Computation*, *11*(8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255

Matsumoto, T., Shirakawa, K., Yokoyama, M., Fukuda, H., Sarca, A. D., Koyabu, S., Yamazaki, H., Kazuma, Y., Matsui, H., Maruyama, W., Nagata, K., Tanabe, F., Kobayashi, M., Shindo, K., Morishita, R., Sato, H., & Takaori-Kondo, A. (2019). Protein kinase a inhibits tumor mutator APOBEC3B through phosphorylation. *Scientific Reports*, *9*(1), 8307. https://doi.org/10.1038/s41598-019-44407-9

Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.Py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation*, *8*(9), 3314–3321. https://doi.org/10.1021/ct300418h

Nair, S., & Rein, A. (2014). In vitro assay for cytidine deaminase activity of APOBEC3 protein. *Bio Protoc*, *4*(20), e1266.

Navyashree, V., Kant, K., & Kumar, A. (2021). Natural chemical entities from Arisaema Genus might be a promising break-through against japanese encephalitis virus infection: A molecular docking and dynamics approach. *Journal of Biomolecular Structure & Dynamics*, *39*(4), 1404–1416. https://doi.org/10.1080/07391102.2020.1731603

NCI. (2023). What is cancer? Retrieved March 5, 2023, from https://www.cancer.gov/about-cancer/understanding/what-is-cancer.

Olson, M. E. A. (n.d.). Thesis submitted to the Faculty Of University Of Minnesota.

Olson, M. E., Harris, R. S., & Harki, D. A. (2018). APOBEC enzymes as targets for virus and cancer therapy. *Cell Chemical Biology*, *25*(1), 36–49. https://doi.org/10.1016/j.chembiol.2017.10.007

Onufriev, A., Bashford, D., & Case, D. A. (2000). Modification of the generalized born model suitable for macromolecules. *The Journal of Physical Chemistry B*, *104*(15), 3712–3720. https://doi.org/10.1021/jp994072s

Onufriev, A., Bashford, D., & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified

generalized Born Model. *Proteins*, 55(2), 383–394. https://doi.org/10.1002/prot.20033

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Prada-Gracia, D., Huerta-Yépez, S., & Moreno-Vargas, L. M. (2016). Application of computational methods for anticancer drug discovery, design, and optimization. *Boletín Médico Del Hospital Infantil de México (English Edition)*, 73(6), 411–423. https://doi.org/10.1016/j.bmhime.2017.11.040

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, 9(7), 3084–3095. https://doi.org/10.1021/ct400341p

Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779–815. https://doi.org/10.1006/jmbi.1993.1626

Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation*, 9(9), 3878–3888. https://doi.org/10.1021/ct400314y

Sharma, A., Vora, J., Patel, D., Sinha, S., Jha, P. C., & Shrivastava, N. (2020). Identification of Natural Inhibitors against Prime Targets of SARS-CoV-2 Using Molecular Docking, Molecular Dynamics Simulation and MM-PBSA Approaches. *Journal of Biomolecular Structure & Dynamics*, 40(7), 3296–3311. https://doi.org/10.1080/07391102.2020.1846624

Shi, K., Carpenter, M. A., Banerjee, S., Shaban, N. M., Kurahashi, K., Salamango, D. J., McCann, J. L., Starrett, G. J., Duffy, J. V., Demir, Ö., Amaro, R. E., Harki, D. A., Harris, R. S., & Aihara, H. (2017). Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nature Structural & Molecular Biology*, 24(2), 131–139. https://doi.org/10.1038/nsmb.3344

Shi, K., Carpenter, M. A., Kurahashi, K., Harris, R. S., & Aihara, H. (2015). Crystal structure of the DNA deaminase APOBEC3B catalytic domain. *The Journal of Biological Chemistry*, 290(47), 28120–28130. https://doi.org/10.1074/jbc.M115.679951

Shi, K., Demir, Ö., Carpenter, M. A., Banerjee, S., Harki, D. A., Amaro, R. E., Harris, R. S., & Aihara, H. (2019). Active site plasticity and possible modes of chemical inhibition of the human DNA Deaminase APOBEC3B. *FASEB bioAdvances*, 2(1), 49–58. https://doi.org/10.1096/fba.2019-00068

Small-Molecule Library. (2023). Screening by Docking with PyRx. | SpringerLink. Retrieved March 5, 2023, from https://link.springer.com/protocol/10.1007/978-1-4939-2269-7_19.

Sun, H., Li, Y., Shen, M., Tian, S., Xu, L., Pan, P., Guan, Y., & Hou, T. (2014). Assessing the performance of MM/PBSA and MM/GBSA Methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. *Physical Chemistry Chemical Physics*, 16(40), 22035–22045. https://doi.org/10.1039/C4CP03179B

Tang, H., Demir, Ö., Kurniawan, F., Brown, W. L., Shi, K., Moeller, N. H., Carpenter, M. A., Belica, C., Orellana, K., Du, G., LeBeau, A. M., Amaro, R. E., Harris, R. S., & Aihara, H. (2021). Structural characterization of a minimal antibody against human APOBEC3B. *Viruses*, 13(4), 663. https://doi.org/10.3390/v13040663

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. https://doi.org/10.1002/jcc.21334

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9), 1157–1174. https://doi.org/10.1002/jcc.20035

Weinbach, Y., & Elber, R. (2005). Revisiting and parallelizing SHAKE. *Journal of Computational Physics*, 209(1), 193–206. https://doi.org/10.1016/j.jcp.2005.03.015

Zhang, X., Wu, Z., Hao, Y., Yu, T., Li, X., Liang, Y., Li, J., Huang, L., Xu, Y., Li, X., Xu, X., Wang, W., Xu, G., Zhang, X., Lv, Q., Fang, Y., Xu, R., & Qian, W. (2022). Aberrantly activated APOBEC3B is associated with mutant P53-driven refractory/relapsed diffuse large B-Cell Lymphoma. *Frontiers in Immunology*, 13, 888250. https://doi.org/10.3389/fimmu.2022.888250