ACS APPLIED BIO MATERIALS

www.acsabm.org

Forum Article

Identifying Potential Ligand Binding Sites on Glycogen Synthase Kinase 3 Using Atomistic Cosolvent Simulations

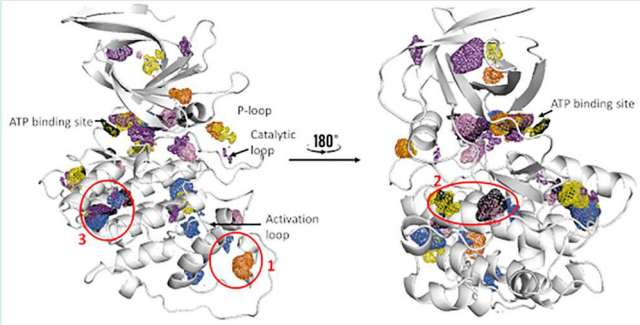
Debarati DasGupta,* Ramin Mehrani, Heather A. Carlson, and Sumit Sharma*

Cite This: https://doi.org/10.1021/acsabm.2c01079

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Glycogen synthase kinase 3 β (GSK3β) is a serine/threonine kinase that phosphorylates several protein substrates in crucial cell signaling pathways. Owing to its therapeutic importance, there is a need to develop GSK3β inhibitors with high specificity and potency. One approach is to find small molecules that can allosterically bind to the GSK3β protein surface. We have employed fully atomistic mixed-solvent molecular dynamics (MixMD) simulations to identify three plausible allosteric sites on GSK3β that can facilitate the search for allosteric inhibitors. Our MixMD simulations narrow down the allosteric sites to precise regions on the GSK3β surface, thereby improving upon the previous predictions of the locations of these sites.



KEYWORDS: *MixMD, cosolvent simulations, GSK3 allosteric site mapping, structure-based drug design*

1. INTRODUCTION

Protein phosphorylation is an important post-translational modification of proteins. This reversible process involves transfer of a phosphate group of purine nucleotide triphosphates, that is, adenosine (ATP) and guanosine triphosphate (GTP) to the serine, threonine, and tyrosine residues of proteins. Protein phosphorylation is catalyzed by enzymes called protein kinases. Protein kinases mediate numerous signal transduction pathways in eukaryotic cells and control a large number of cellular processes, including transcription, apoptosis, homeostasis, cell cycle progression, intercellular communication, etc. Dysregulation of protein kinases is linked to various human ailments, thus affording the development of their agonists and antagonists as a promising route for disease therapy. There are more than 500 kinases in the human genome.[1] Glycogen synthase kinase 3 (GSK3) is a conserved serine/threonine kinase present in all eukaryotes that regulates numerous cellular functions. GSK3 exists in two isoforms, namely, GSK3α and GSK3β, that share 68% amino acid similarity, and their ATP binding site is conserved with 95% sequence similarity. Both these isoforms are expressed in humans. While first identified as a kinase for glycogen synthase, at present there are over 40 known GSK3 substrates. Not surprisingly, abnormal activity of GSK3 is associated with a host of diseases, including cancer, non-insulin-dependent diabetes mellitus, pathological inflammation, asthma, myeloid leukemia, cardiac hypertrophy, and neurological and neuro-developmental disorders such as bipolar disorder and Alzheimer's disease.[2] GSK3's multifarious roles in many cellular events make it an important therapeutic target for the treatment of these disorders.[2−8]

GSK3 inhibitors may be classified as ATP-competitive, non-ATP-competitive, and allosteric. The majority of the known small molecule GSK3 inhibitors are ATP-competitive. These inhibitors are not highly selective toward GSK3 because of similarities of the ATP binding pockets between different kinases. The compound 9-ING-41 is an ATP-competitive inhibitor that has entered phase-II clinical trials.[9] Wagner et al.[10,11] recognized that the ATP binding sites of GSK3β and GSK3α differ by a single residue: Arg-113 (GSK3β)/Glu-196 (GSK3α) and so developed GSK3β selective inhibitors exploiting the different interactions and steric requirements of the ATP binding sites in the two isoforms. In another work, researchers identified 4-hydroxy-4-phenyl-3-(pyridine-4-ylmethyl)thiazolidine-2-thione (in short, COB-187) as a specific and potent ATP competitive inhibitor of GSK3α and GSK3β.[10,11] Tideglusib is a non-ATP competitive inhibitor that is now under phase-II clinical trials. However, recent studies have shown that tideglusib binds to several different kinases and is only moderately selective toward GSK3.[10,11]

An important concern in developing GSK3 inhibitors is that they should not upregulate oncogenes in the process of

Special Issue: Computational Advances in Biomaterials

Received: December 29, 2022
Accepted: April 20, 2023

suppressing GSK3 activity. For instance, the inhibitors should not alter GSK3's role in β-catenin signaling. One approach to ensure this is to develop GSK3 inhibitors that target the allosteric binding sites.[12] Palinurin, a natural product extracted from marine sponge, has been identified as a non-ATP/substrate GSK3 inhibitor.[13] So far, the search for allosteric inhibitors of GSK3 has remained little explored.[14,15] While ATP binding sites, with their well-defined cleft, are clearly delineated on the protein surface, allosteric binding sites are more challenging to find. A pioneering work in this field is from Palomo et al.,[16] wherein the researchers studied 25 PDB structures of GSK3β using Fpocket and Hpocket binding site identification algorithms to map the potential binding sites. Along with the ATP-, the substrate-, and the binding sites of peptides axin/fratide, Palomo et al. identified four potential allosteric binding sites. In a subsequent study, Palomo et al.[17] identified a family of allosteric modulators associated with the binding site between the residues Arg209 and Thr235 of GSK3β (labeled as pocket 7 in their work[16]). Brogi et al.[18] developed a dual inhibitor of human adenosine kinase and GSK3β and then docked the inhibitor on the pocket 7 of GSK3β using Autodock to determine its relevant interactions with the protein.[18] Silva et al.[19] employed a number of binding site identification algorithms (Fpocket, Superstar, metaPocket, Sitemap, and PARS) to determine which residues are included to be a part of the pocket 7 by these algorithms. Then, they docked a previously identified allosteric modulator on the four allosteric sites found by Palomo et al.[16] Carullo et al.[20] synthesized a family of square-amide based compounds and evaluated their potencies as ATP noncompetitive inhibitors of GSK3β. While these are illuminating studies, they are based on a static structure of a protein to identify the binding sites and ligand binding affinities. Fpocket uses a geometry-based criterion for finding the binding sites. Recent work has shown that protein geometry may not be correlated to the behavior of interfacial water and ligand binding propensity.[21] Therefore, searching for ligand binding sites solely based on protein geometry may be misleading. The present work stands out as the first dynamical investigation of allosteric binding sites of GSK3β, which will be immensely useful to the practitioners who invoke docking and molecular simulation methods to identify allosteric inhibitors of GSK3β.

In an aqueous environment, a ligand will bind to a protein only if it has a stronger affinity for a site as compared to water and ions. Binding of a ligand on a protein surface is associated with the loss in translational and conformational entropy of the ligand and, in some cases, a gain in the entropy of interfacial water.[22] Geometry-based algorithms and/or algorithms that only consider a static structure of a protein to identify binding sites are unable to capture these entropic effects as well as the competition between ligand and water molecules for the binding sites. The dynamics of the protein itself plays a role in ligand binding, which is also missed in these coarse approaches. In this work, we have employed cosolvent molecular dynamics simulations (MixMD) to identify plausible allosteric sites on GSK3β. MixMD is a powerful, yet computationally tractable approach for mapping hotspots on proteins and has been used to map active sites, allosteric sites, protein−protein interfaces, and cryptic pockets.[23−27] Through the MixMD simulation methodology, we have identified three different potential allosteric sites on GSK3β. The sites identified through our methodology align well with the findings of Palomo et al.[16] Since Fpocket is a geometry-based algorithm that works on static protein structures, the Palomo et al. work identified large regions on the protein that can serve as binding sites. Our MixMD approach, on the other hand, is based on molecular dynamics and, therefore, is able to identify precise regions on the GSK3β protein surface as potential binding sites. Hence, this work presents an improvement over the previously identified allosteric sites.[16]

2. SIMULATION SYSTEM AND METHODS

In MixMD simulations, a single protein chain is simulated in an aqueous medium with a fixed concentration of a small molecule, which is termed as a molecular probe/cosolvent. Through a series of MD simulations performed by choosing different polar, charged, and hydrophobic molecular probes, the locations on the protein surface that register a large concentration of these probes are identified as putative binding sites. For the MixMD simulations, crystal structure of GSK3β, PDB ID 6V6L, was obtained from the protein data bank (PDB).[28] We chose the PDB structure 6V6L because it has 2.1 Å resolution and no missing residues at the catalytic site. This crystal structure has 11 missing residues, mostly in the N and C termini ends of the sequence, which are not essential for the kinase activity. The structure was prepped by adding missing residues and hydrogen atoms and setting all protonation states in accordance with a pH of 7.2. Histidine, Glutamine, and Asparagine residue tautomers were manually corrected, and the structure was stripped of all nonprotein atoms. These protein structure preparation steps were performed using Molecular Operating Environment (MOE).[29] To neutralize the system, seven chloride ions were added using the tleap module of AMBER20.[30] The GSK3β structure so obtained when compared to the structure predicted by α-fold has a backbone RMSD of 0.89 Å, which confirms that our initial structure aligns closely with the α-fold model (Figure S1 in Supporting Information). Six different molecular probes were used in the MixMD simulations: acetonitrile (C3N), isopropanol (IPA), pyrimidine (1P3), *N*-methylacetamide (NMA), imidazole (IMI), and ethanol (EOH). The selected molecular probes are drug-like fragments so that the identified binding sites have high druggability. We used the same probes that have been used in previously published MixMD simulation-based works.[31−36] The C3N, EOH, and IPA have the ability to form hydrogen bonds with protein pockets and are considered polar moieties, although the IPA group can act as a hydrophobic marker and has the potential to probe deep pockets. The IMI group is hydrophobic, participates in $\pi-\pi$ interactions, and therefore maps the hinge binding regions of kinases which are mostly heteroaromatic cores. Force field parameters of IPA, 1P3, C3N, IMI, and EOH were taken from Lexa et al.[37] NMA parameters were developed by Caldwell and Kollman.[38] "The force field parameters of the probes that were used in this study have been well-validated and are the ones that have been used in the previous MixMD simulation studies."[32,39−41] Implementation of the MixMD method was as per the protocol delineated in previous works.[23,24,42] In the initial configuration, the apo protein was layered with an 8-Å-thick shell of molecular probes. Then the system was solvated with water represented by the TIP3P[43] water model to achieve 5%v/v concentration of the probe. Simulations were performed in Amber20 using the ff119SB force field.[44] All hydrogen atoms were restrained using the SHAKE[45] algorithm. MD simulations were performed with a time step of 2 fs using Pmemd.cuda[46] on GPU architecture for equilibration and production runs and Sander was used for energy minimization and temperature ramping steps. Long-range electrostatics were treated using Particle Mesh Ewald.[46] The real space parts of Coulombic interactions and nonbonded interactions were cut off at 10 Å. The system was energy minimized using a 20 kcal/mol/A^2 harmonic restraint on the protein via steepest descent for 15 000 steps. The system was then gradually heated from 0 to 300 K in 300 ps in the canonical ensemble. During the heating step, the protein backbone was restrained using a harmonic restrain of 3 kcal/mol/A^2. After reaching the target temperature of 300 K, the system was subjected to equilibration in the isothermal−isobaric ensemble at 300 K and 1 bar pressure. Three equilibration steps, each 1 ns long, were

https://doi.org/10.1021/acsabm.2c01079
ACS Appl. Bio Mater. XXXX, XXX, XXX−XXX

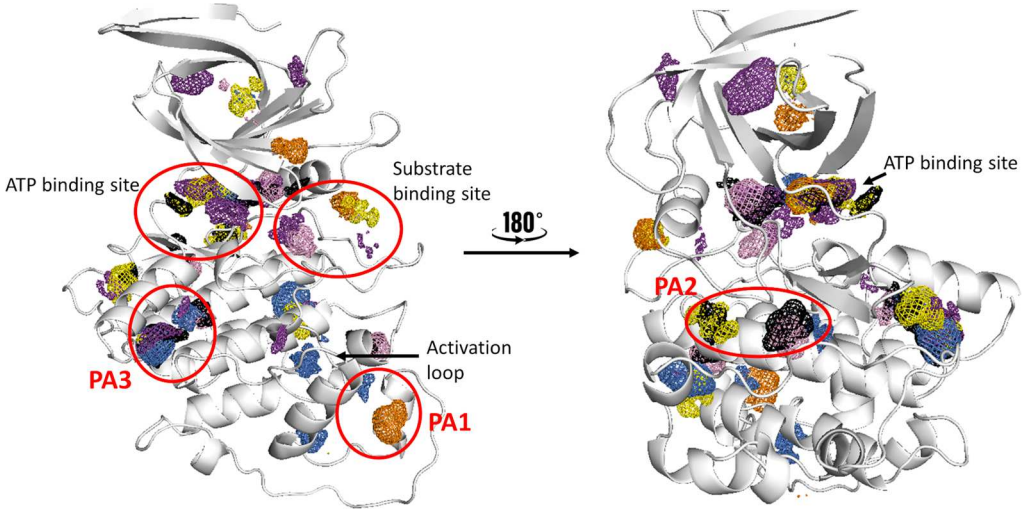


Figure 1. Hotspots mapped on GSK3β using 6 small water miscible probes (C3N—orange, IPA—blue, 1P3—purple, IMI—black, NMA—yellow, and EOH—light pink). The solvent maps are contoured at 35σ. GSK3β has several important sites mapped including the P-loop, the catalytic site, the C lobe area, the catalytic and the activation loops, and three possible allosteric sites (PA1, PA2, and PA3) that are circled in the figure.

performed wherein the protein backbone restrain was reduced from 5 to 2 to 1 kcal/mol/A^2. A final unconstrained equilibration step was performed for 75 ns in the isothermal−isobaric ensemble, which was followed by a 30 ns production run. For each probe, ten independent runs were performed to ensure adequate sampling and the MD configurations were written after every 250 ps. To confirm that our simulations are sufficiently long, for each probe we extended five independent runs to 100 ns and ensured that the probe mapping is consistent for each 20 ns window (30−50 ns, 70−90 ns, and 80−100 ns) and with the first 30 ns. Therefore, for each probe, we have 650 ns of simulation data.

3. RESULTS AND DISCUSSION

To generate a local density map of the probes/cosolvent molecules, the simulation system was divided into grids of 0.5 Å spacing.[47] In each voxel of the grid, average density of the probe molecules was determined. A mesh of cosolvent density, called a normalized occupancy map, was created by considering deviation from the mean density, given by $\xi = \frac{x-\mu}{\sigma}$, where x is the ensemble averaged density in the voxel of the grid, μ is the mean density of the probe in the simulation system, and σ is the standard deviation of the density of the probe in the voxel of the grids. The normalized occupancy maps were visualized in Pymol.[48] The probe/cosolvent densities were contoured at $\xi > 35$. The regions with large ξ associated with three or more probes were considered the potential binding sites. Large-valued contours indicate high concentration of the probes. Figure 1 shows the hotspots on GSK3β ($\xi > 35$) identified in the MixMD simulations.

A total of 6 sites were mapped by three or more probes. The ATP binding site was found to be the dominant one and was mapped by 5 probes (C3N, IPA, 1P3, IMI, and NMA) at $\xi >$ 70. Other sites were mapped with $\xi \in (35,50)$. Several conserved sites on GSK3β, including the P-loop, the C lobe area, and the catalytic and activation loops registered large ξ values in the MixMD simulations (Figure S2, Supporting Information). In addition, three possible allosteric sites were also identified, which are labeled as sites PA1, PA2, and PA3 in the Figure 1.

In the MixMD simulations, the GSK3β protein did not undergo any conformational change. Figure 2 shows the backbone RMSD plots of the MD trajectories associated with each probe. The backbone RMSD fluctuations were within 2−3 Å in the last 80 ns of the simulation time, except in the case of the IPA probe, where for 15 ns (from 35 to 50 ns), the RMSD jumped to 3.8 Å but then reverted back to 2−3 Å. The backbone fluctuations arise mainly from the motion of the P loop and the catalytic loop.

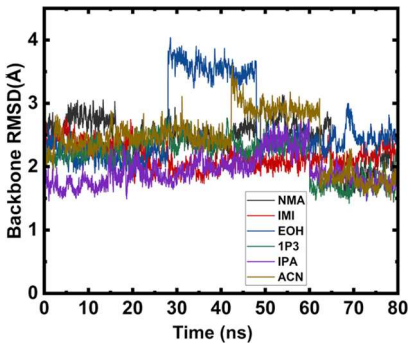


Figure 2. Backbone RMSD fluctuations of representative MD trajectories for each probe for the last 80 ns of the simulation. The RMSD remained within 2−3 Å as compared to the crystal structure, showing that the GSK3β protein did not undergo any conformational change in the presence of the probes.

The allosteric site 1 (PA1), mapped by the probes 1P3, NMA, IMI, and IPA, is located at the G-H helix intersection with several aromatic-rich residues lining the binding pocket. This site is hydrophobic with the following residues within 4 Å of the pocket: Phe-206, Ile-258, Pro-263, Pro-253, Asn-263, and Tyr-265. The 1P3 and IMI probes are involved in $\pi-\pi$ interactions with the aromatic groups of Tyr-265 and Phe-206 residues. The NMA probe forms H-bonds with the Asn-263. IPA and IMI have hydrophobic interactions with the Pro-263, Pro-253, and Ile-258 groups. The allosteric site 2 (PA2) is a negatively charged, slightly shallow cleft, composed of acidic and charged side chains: Asp-82, His-83, Arg-144, Tyr-148, and Thr-340. Not surprisingly, this site is mapped by the polar probes: NMA, 1P3, and IMI. NMA and 1P3 have Coulombic interactions with Asp-82, His-83, and Arg-144. IMI has $\pi-\pi$ interactions with Tyr-148. The allosteric site 3 (PA3) is mainly

https://doi.org/10.1021/acsabm.2c01079
ACS Appl. Bio Mater. XXXX, XXX, XXX−XXX

Table 1. Free Energy of Binding, ΔG_i, of Different Molecular Probes at the Three Potential Allosteric Sites on GSK3β[a]

Site ID	C3N	IPA	1P3	NMA	IMI	EOH
PA1	−1.5 (0.2)	−2.5 (0.4)	−3.9 (0.2)	−2.2 (0.2)	−2.9 (0.2)	−1.5 (0.3)
PA2	−1.8 (0.2)	−2.3 (0.3)	−3.7 (0.4)	−2.6 (0.3)	−2.7 (0.3)	−1.7 (0.3)
PA3	−1.8 (0.2)	−2.8 (0.3)	−4.2 (0.3)	−2.9 (0.3)	−3.0 (0.2)	−2.2 (0.5)

[a]The ΔG_i are listed in kcal/mol and are the mean values calculated from the five 100 ns MD runs. The errors are standard deviations of ΔG_i from the five runs.

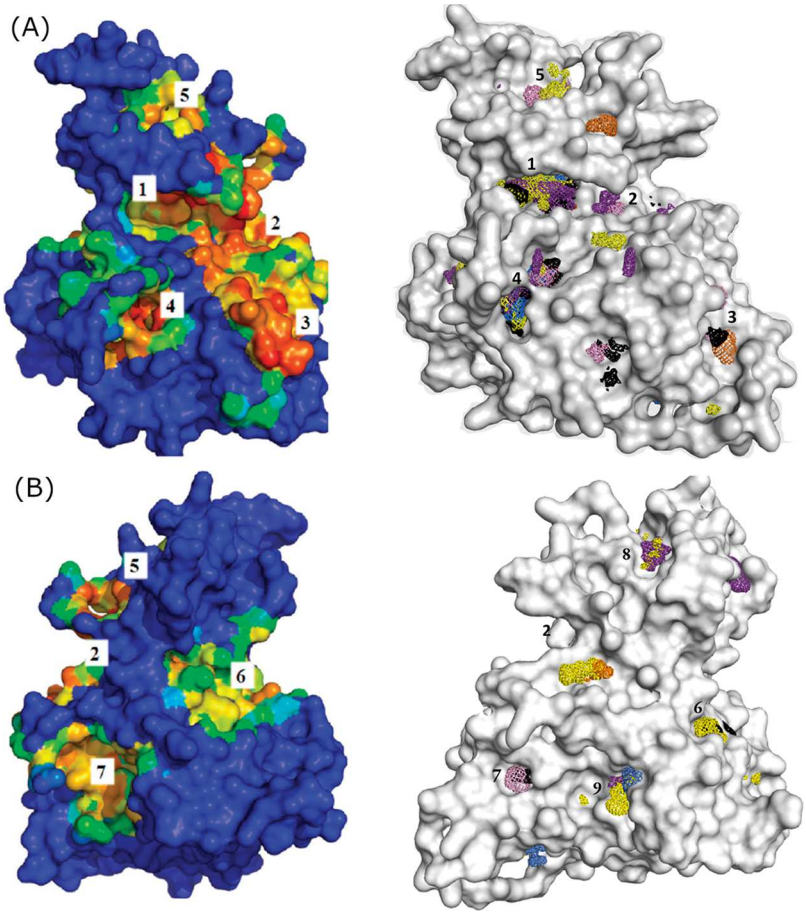


Figure 3. Comparison of the binding sites identified by Palomo et al.[16] (left column) and in this work (right column). (A) shows sites 1−5. Site 1 is the ATP binding site. Site 2 is the substrate binding site. (B) shows the site 2, and sites 5−7. The sites 3, 4, and 7 in Palomo et al. correspond to PA1, PA3, and PA2, respectively. Site 5 in Palomo et al. was not found to be conserved in all crystal structures of GSK3β. It was only mapped by two probes in the MixMD simulations. The MixMD contours appear different than Figure 1 because the surface plot representation of the protein hides some contours. The left column figures in (A) and (B) are reproduced from ref 16. Copyright 2011 American Chemical Society.

lined up with amino acid residues: Arg-125, Tyr-117, Ala-120, Pro-232, Gln-231, Tyr-199, and Ser-124. This site is mapped by the probes: NMA, 1P3, and IMI. The 1P3 and IMI probes are involved in $\pi-\pi$ interactions with the aromatic group of Tyr-199. The NMA and 1P3 probes interact with Arg-125 and Gln-231 residues via Coulombic interactions. IMI has hydrophobic interactions with Tyr-117, Ala-120, and Pro-232.

We estimated free energy of binding of the molecular probes on the three potential allosteric sites based on the probability of occupancy of the sites, $\Delta G_i = -RT\ln {}^{N_i}/_{N_o}$. Here N_i is the probability of a molecular probe being bound to the site and N_o is the probability of occupying a bulk region of the same volume as the site. This methodology has been outlined in a recent work.[12] The estimated free energies of binding of all the probes are listed in Table 1. For the sites PA1, PA2, and PA3, 1P3 shows the highest binding affinity followed closely by IMI and NMA. Since C3N is composed of 3 heavy atoms, it has the weakest affinity among all the probes simulated.

Figure 3 compares the binding sites identified by Palomo et al.[16] and in this work. The MixMD contours appear different in this figure as compared to Figure 1 because the surface plot representation of the protein hides some contours. Site 1 is the ATP binding site. Site two is the substrate binding site. The sites 3, 4, and 7 in Palomo et al. correspond to the potential allosteric sites PA1, PA3, and PA2, respectively. The PA2 is essentially a smaller hotspot that overlaps with site 7 of Palomo et al. as well as another hotspot close to it (Figure 1). MixMD simulations provide precise locations of the binding probes and therefore narrow down the locations of the putative binding sites. Site 5 in Palomo et al. was not found to be conserved in all crystal structures of GSK3β. It was only mapped by two probes in the MixMD simulations and therefore was not identified as a potential allosteric site.

https://doi.org/10.1021/acsabm.2c01079
ACS Appl. Bio Mater. XXXX, XXX, XXX−XXX

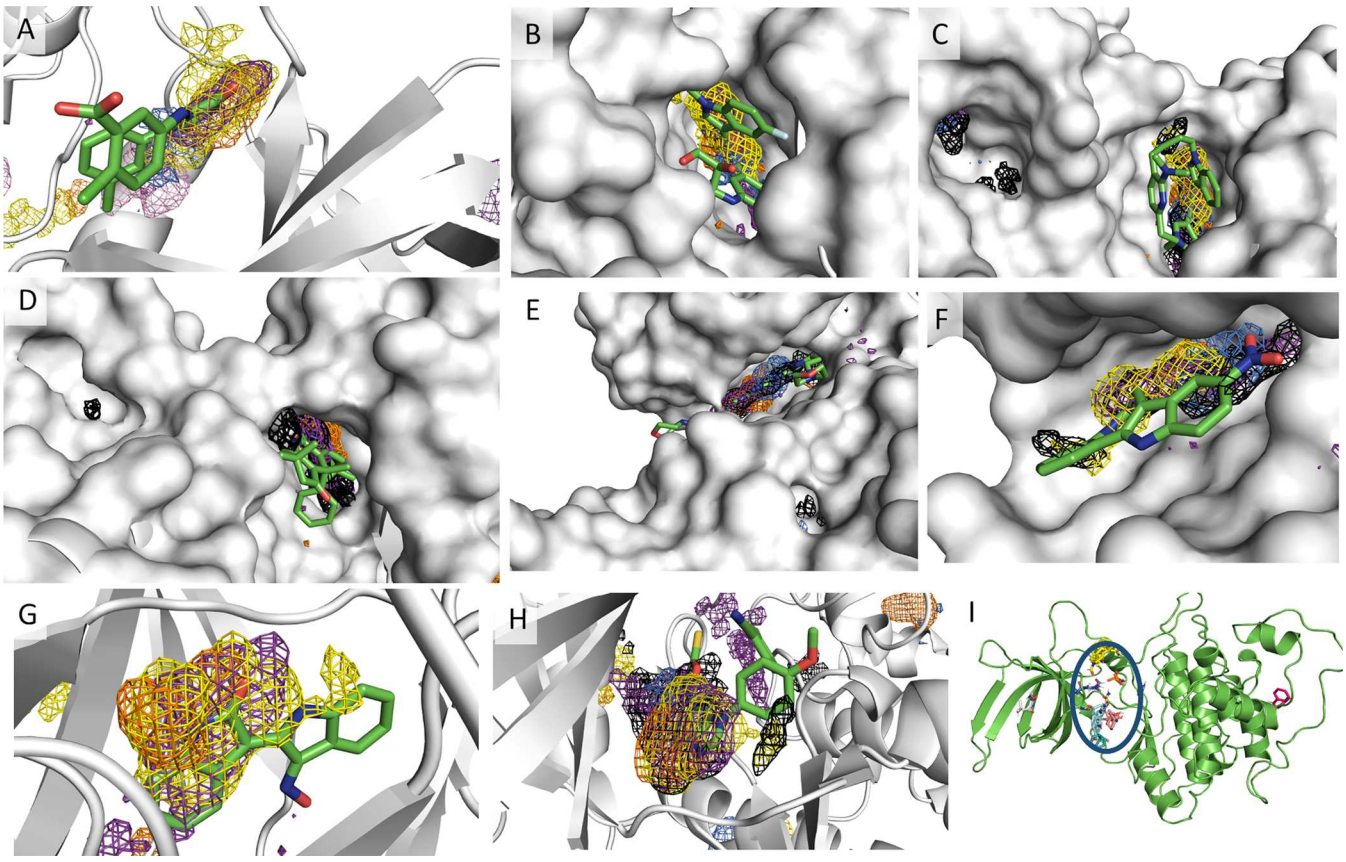


Figure 4. ATP binding site of GSK3 was found to be the dominant binding site in our MixMD simulations with molecular probes contoured at 70σ levels. The contours obtained for the molecular probes coincide with the ligands that have been identified to the ATP binding site. PDB structures of these ligands in the bound configuration are shown in the panels A to H with PDB IDs: A. 5KPM; B. 6Q8K; C.1H8F; D. 1Q41; E. 3F88; F. 2OW3; G. 1R0E; H. 1Q4L. Panel I shows result from the FTMap technique. FTMap was able identify only the ATP binding site on the GSK3 protein.

A search of the protein data bank shows that there are several ligands that are known to bind to the ATP binding site of GSK3β (PDB IDs: 5KPM, 6Q8K, 1H8F, 1Q41, 3F88, 2OW3, 1R0E, and 1Q4L). In Figure 4, we compare the configurations of these ligands when bound to the ATP binding site with the contours of the molecular probes obtained from the MixMD simulations. We find that the bound locations of the ligands coincide quite well with the molecular probe contours highlighting the effectiveness of the MixMD methodology. As our MixMD simulations revealed, the ATP binding site is the most favorable one for ligand binding, and that is why ligands in nanomolar concentrations can bind to this site and many such ligands have been found in nature. The potential allosteric sites have smaller affinities for ligands. As has been highlighted in previous work, ligands targeting these sites are expected to mildly modulate GSK3β activity and will be needed in higher concentrations.[14,15,17−19,49] Therefore, it is not surprising that there are no ligands found in nature that bind to these potential allosteric sites in nanomolar concentrations. We used Fourier Transform Map (FTMap)[50] as an independent technique to identify binding sites on GSK3β to compare against our findings. FTMap is a computational technique in which 16 different molecular probes are docked on to a protein surface in a large number of conformations to identify the strongest binding sites. This methodology is directly analogous to the X-ray technique called multiple solvent crystal structures wherein probe molecules are used to search for binding sites by performing X-ray crystallography. While FTMap has proven to be an effective technique, its drawback is that it does not consider protein flexibility and competition between water and molecular probes for the binding sites. We find that the FTMap is only able to identify the ATP binding site on GSK3β (Figure 4 Panel I and Figure S3).

We also employed another recently developed druggability prediction tool, PockDrug Server[51] to identify potential bindable sites for GSK3β. PockDrug uses a consensus algorithm based on the binding site predictions made by various binding site prediction algorithms: Fpocket,[52] DrugPred,[53] SiteMap,[54] and DoGSiteScorer.[55] The results from the PockDrug server are given in Table S1 (Supporting Information). Interestingly, PockDrug can identify the allosteric sites 1 and 3 with a high druggability score. The site 1 is labeled as the site P6 in the Table S1 and has a volume of 727 Å^3. The site 3 corresponds to the site P4 in the Table S1 and has a cavity volume of 908 Å^3. Since PockDrug is a geometry-based algorithm that operates on a static structure of protein and does not include the dynamics of water, it is not as accurate as the MixMD methodology that we have implemented. Nevertheless, it corroborates our assessment that the allosteric site 1 is mainly hydrophobic. Similarly, FTMap fails to identify the potential allosteric sites identified by the MixMD methodology as has been noted before for other kinases.[33]

https://doi.org/10.1021/acsabm.2c01079
ACS Appl. Bio Mater. XXXX, XXX, XXX−XXX

The probe molecules used in MixMD simulation are like fragments of actual drug molecules, and so how do we know that the sites identified by these probes will also serve as putative binding sites for drug molecules? An important point to remember is that in the MixMD approach, a site is considered as a binding site only when three or more different probes show affinity for the site and have extensive site mapping at $\xi > 35$. This rule has been optimized to obtain a clean signal-to-noise ratio and to filter out sites that may not be amenable for small molecule drug design. Different probes are chosen to capture different kinds of interactions (hydrophobic, polar, hydrogen bonding, $\pi-\pi$ interactions). So, these criteria ensure that the identified sites have affinity for a multitude of functional groups and a drug molecule with similar groups will experience a strong binding affinity for these sites. Therefore, the binding sites identified by us using MixMD, albeit smaller than those identified by geometry-based algorithms, are still large enough to harbor drug-like molecules. In a mini perspective, Ghanakota and Carlson[33] have discussed the need for cosolvent based MD simulations to determine potential hotspots on protein targets for structure based drug discovery. They compared the performance of binding site mapping algorithms like FTMap with MixMD simulations. While static protein structure-based algorithms are computationally fast, they lack the accuracy offered by MixMD simulations. FTMap samples probe molecules on a densely spaced grid. The 16 probes used in FTMap are all similarly sized to the fragments used in MixMD. For the 5 kinases studied by Ghanakota and Carlson, FTMap was successful in characterizing their ATP binding site but was unable to map the potential allosteric sites in any of them. A similar failure was observed with the SiteMap tool of Schrodinger. Often, distinct conformational changes like closing or opening of a binding site (allosteric/cryptic pockets) and side-chain movements are important, which are only captured by molecular dynamics-based methods. Along with the identification of hotspots, MixMD simulation results are useful for fragment-based drug design, wherein a set of small fragments are grown to lead molecules that eventually become drug candidates.[35] MixMD simulations have also helped in identifying cryptic binding sites, which are defined as the sites that become available upon structural reorganization of a protein.[31]

■ CONCLUSIONS

We employed mixed-solvent molecular dynamics (MixMD) simulations to identify potential allosteric sites on GSK3β protein. In these simulations, we find that small water miscible probes employed in the MixMD simulations not only efficiently map the ATP binding site but also map three other potential allosteric sites. These sites can be targeted for designing GSK3β specific inhibitors. It should be noted that we refer to the identified sites as "potential" allosteric sites because we do not provide evidence that binding of a ligand at any of these sites will result in the modulation of GSK3β activity. Allosteric inhibition of GSK3β has garnered significant interest as it is expected to help to achieve specificity without impacting other kinases. We argue through this work that to find allosteric inhibitors of GSK3β, the researchers should target the potential allosteric sites identified by us for their search. The potential allosteric sites identified through MixMD simulations align well with those from Palomo et al.[16] but our work narrows down the regions on the protein surface associated with these sites. These sites have been the basis of previous studies to develop allosteric inhibitors for GSK3β.[17−19]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsabm.2c01079.

Figure S1: Backbone alignment of crystal structure of GSK3β (PDB ID 6V6L) after generating missing residues with the α-fold full length model and Md averaged structure. Figure S2: Various conserved sites on GSK3β. Figure S3: FTMAP results of GSK3β showing all sites mapped by 16 molecular probes. Figure S4: Contours at different values of ξ showing high density regions of the probes used in MixMD simulations. Table S1. Potential binding sites on GSK3β identified using Pocket Druggability Prediction (PockDrug) server. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Sumit Sharma − *Department of Chemical and Biomolecular Engineering, Ohio University, Athens, Ohio 45701, United States*; orcid.org/0000-0003-3138-5487; Email: sharmas@ohio.edu

Debarati DasGupta − *College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109, United States*; Email: debarati_dasgupta@hotmail.com

Authors

Ramin Mehrani − *Department of Mechanical Engineering, Ohio University, Athens, Ohio 45701, United States*

Heather A. Carlson − *College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109, United States*; orcid.org/0000-0002-7495-1699

Complete contact information is available at: https://pubs.acs.org/10.1021/acsabm.2c01079

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation CDS&E grant 1953311. Computational resources for this work were provided by the National Science Foundation XSEDE grant DMR190005. H.C. would like to acknowledge funding from the National Institutes of General Medical Sciences (R01 GM065372).

■ REFERENCES

(1) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298* (5600), 1912−1934.

(2) Wagner, F. F.; Benajiba, L.; Campbell, A. J.; Weïwer, M.; Sacher, J. R.; Gale, J. P.; Ross, L.; Puissant, A.; Alexe, G.; Conway, A.; Back, M.; Pikman, Y.; Galinsky, I.; DeAngelo, D. J.; Stone, R. M.; Kaya, T.; Shi, X.; Robers, M. B.; Machleidt, T.; Wilkinson, J.; Hermine, O.; Kung, A.; Stein, A. J.; Lakshminarasimhan, D.; Hemann, M. T.; Scolnick, E.; Zhang, Y. L.; Pan, J. Q.; Stegmaier, K.; Holson, E. B. Exploiting an Asp-Glu "switch" in glycogen synthase kinase 3 to design paralog-selective inhibitors for use in acute myeloid leukemia. *Sci. Transl Med.* **2018**, *10*, 431.

(3) Beurel, E.; Grieco, S. F.; Jope, R. S. Glycogen synthase kinase-3 (GSK3): regulation, actions, and diseases. *Pharmacol Ther* **2015**, *148*, 114−31.

(4) Wu, X.; Stenson, M.; Abeykoon, J.; Nowakowski, K.; Zhang, L.; Lawson, J.; Wellik, L.; Li, Y.; Krull, J.; Wenzl, K.; Novak, A. J.; Ansell, S. M.; Bishop, G. A.; Billadeau, D. D.; Peng, K. W.; Giles, F.; Schmitt, D. M.; Witzig, T. E. Targeting glycogen synthase kinase 3 for therapeutic benefit in lymphoma. *Blood* **2019**, *134* (4), 363−373.

(5) Balasubramaniam, M.; Mainali, N.; Bowroju, S. K.; Atluri, P.; Penthala, N. R.; Ayyadevera, S.; Crooks, P. A.; Shmookler Reis, R. J. Structural modeling of GSK3β implicates the inactive (DFG-out) conformation as the target bound by TDZD analogs. *Sci. Rep* **2020**, *10* (1), 18326.

(6) Giglio, J.; Fernandez, S.; Martinez, A.; Zeni, M.; Reyes, L.; Rey, A.; Cerecetto, H. Glycogen Synthase Kinase-3 Maleimide Inhibitors As Potential PET-Tracers for Imaging Alzheimer's Disease: (11)C-Synthesis and In Vivo Proof of Concept. *J. Med. Chem.* **2022**, *65* (2), 1342−1351.

(7) Saraswati, A. P.; Ali Hussaini, S. M.; Krishna, N. H.; Babu, B. N.; Kamal, A. Glycogen synthase kinase-3 and its inhibitors: Potential target for various therapeutic conditions. *Eur. J. Med. Chem.* **2018**, *144*, 843−858.

(8) Rana, A. K.; Singh, D. Targeting glycogen synthase kinase-3 for oxidative stress and neuroinflammation: Opportunities, challenges and future directions for cerebral stroke management. *Neuro-pharmacology* **2018**, *139*, 124−136.

(9) Gaisina, I. N.; Gallier, F.; Ougolkov, A. V.; Kim, K. H.; Kurome, T.; Guo, S.; Holzle, D.; Luchini, D. N.; Blond, S. Y.; Billadeau, D. D. From a Natural Product Lead to the Identification of Potent and Selective Benzofuran-3-yl-(Indol-3-yl) Maleimides as Glycogen Synthase Kinase 3β Inhibitors that Suppress Proliferation and Survival of Pancreatic Cancer Cells. *J. Med. Chem.* **2009**, *52* (7), 1853−1863.

(10) Noori, M. S.; Bhatt, P. M.; Courreges, M. C.; Ghazanfari, D.; Cuckler, C.; Orac, C. M.; McMills, M. C.; Schwartz, F. L.; Deosarkar, S. P.; Bergmeier, S. C.; McCall, K. D.; Goetz, D. J. Identification of a novel selective and potent inhibitor of glycogen synthase kinase-3. *Am. J. Physiol Cell Physiol* **2019**, *317* (6), C1289−c1303.

(11) Ghazanfari, D.; Noori, M. S.; Bergmeier, S. C.; Hines, J. V.; McCall, K. D.; Goetz, D. J. A novel GSK-3 inhibitor binds to GSK-3β via a reversible, time and Cys-199-dependent mechanism. *Bioorg. Med. Chem.* **2021**, *40*, 116179.

(12) Buch, I.; Fishelovitch, D.; London, N.; Raveh, B.; Wolfson, H. J.; Nussinov, R. Allosteric regulation of glycogen synthase kinase 3β: a theoretical study. *Biochemistry* **2010**, *49* (51), 10890−901.

(13) Bidon-Chanal, A.; Fuertes, A.; Alonso, D.; Pérez, D. I.; Martínez, A.; Luque, F. J.; Medina, M. Evidence for a new binding mode to GSK-3: Allosteric regulation by the marine compound palinurin. *Eur. J. Med. Chem.* **2013**, *60*, 479−489.

(14) Palomo, V.; Martinez, A. Glycogen Synthase Kinase 3 (GSK-3) Inhibitors: a Patent Update (2014−2015). *Expert Opin. Ther. Pat.* **2017**, *27* (6), 657−666.

(15) Roca, C.; Campillo, N. E. Glycogen Synthase Kinase 3 (GSK-3) Inhibitors: A Patent Update (2016−2019). *Expert Opin. Ther. Pat.* **2020**, *30* (11), 863−872.

(16) Palomo, V.; Soteras, I.; Perez, D. I.; Perez, C.; Gil, C.; Campillo, N. E.; Martinez, A. Exploring the Binding Sites of Glycogen Synthase Kinase 3. Identification and Characterization of Allosteric Modulation Cavities. *J. Med. Chem.* **2011**, *54* (24), 8461−8470.

(17) Palomo, V.; Perez, D. I.; Roca, C.; Anderson, C.; Rodríguez-Muela, N.; Perez, C.; Morales-Garcia, J. A.; Reyes, J. A.; Campillo, N. E.; Perez-Castillo, A. M. Subtly Modulating Glycogen Synthase Kinase 3 β: Allosteric Inhibitor Development and Their Potential for the Treatment of Chronic Diseases. *J. Med. Chem.* **2017**, *60* (12), 4983−5001.

(18) Brogi, S.; Ramunno, A.; Savi, L.; Chemi, G.; Alfano, G.; Pecorelli, A.; Pambianchi, E.; Galatello, P.; Compagnoni, G.; Focher, F. First Dual AK/GSK-3β Inhibitors Endowed With Antioxidant Properties as Multifunctional, Potential Neuroprotective Agents. *Eur. J. Med. Chem.* **2017**, *138*, 438−457.

(19) Silva, G. M.; Borges, R. S.; Santos, K. L. B.; Federico, L. B.; Francischini, I. A. G.; Gomes, S. Q.; Barcelos, M. P.; Silva, R. C.; Santos, C. B. R.; Silva, C. H. T. P. Revisiting the Proposition of Binding Pockets and Bioactive Poses for GSK-3β Allosteric Modulators Addressed to Neurodegenerative Diseases. *Int. J. Mol. Sci.* **2021**, *22* (15), 8252.

(20) Carullo, G.; Bottoni, L.; Pasquini, S.; Papa, A.; Contri, C.; Brogi, S.; Calderone, V.; Orlandini, M.; Gemma, S.; Varani, K. Synthesis of Unsymmetrical Squaramides as Allosteric GSK-3β Inhibitors Promoting β-Catenin-Mediated Transcription of TCF/LEF in Retinal Pigment Epithelial Cells. *ChemMedChem.* **2022**, *17*, No. e202200456.

(21) Rego, N. B.; Xi, E.; Patel, A. J. Identifying Hydrophobic Protein Patches to Inform Protein Interaction Interfaces. *Proc. Nat. Acad. Sci. U. S. A.* **2021**, *118* (6), No. e2018234118.

(22) Setny, P.; Baron, R.; McCammon, J. A. How can Hydrophobic Association be Enthalpy Driven? *J. Chem. Theory Comput.* **2010**, *6* (9), 2866−2871.

(23) Ghanakota, P.; Carlson, H. A. Moving Beyond Active-Site Detection: MixMD Applied to Allosteric Systems. *J. Phys. Chem. B* **2016**, *120* (33), 8685−95.

(24) Ghanakota, P.; DasGupta, D.; Carlson, H. A. Free Energies and Entropies of Binding Sites Identified by MixMD Cosolvent Simulations. *J. Chem. Inf Model* **2019**, *59* (5), 2035−2045.

(25) Ghanakota, P.; van Vlijmen, H.; Sherman, W.; Beuming, T. Large-Scale Validation of Mixed-Solvent Simulations to Assess Hotspots at Protein-Protein Interaction Interfaces. *J. Chem. Inf Model* **2018**, *58* (4), 784−793.

(26) Smith, R. D.; Carlson, H. A. Identification of Cryptic Binding Sites Using MixMD with Standard and Accelerated Molecular Dynamics. *J. Chem. Inf Model* **2021**, *61* (3), 1287−1299.

(27) Makley, L. N.; Johnson, O. T.; Ghanakota, P.; Rauch, J. N.; Osborn, D.; Wu, T. S.; Cierpicki, T.; Carlson, H. A.; Gestwicki, J. E. Chemical validation of a druggable site on Hsp27/HSPB1 using in silico solvent mapping and biophysical methods. *Bioorg. Med. Chem.* **2021**, *34*, 115990.

(28) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235−42.

(29) Vilar, S.; Cozza, G.; Moro, S. Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. *Curr. Top. Med. Chem.* **2008**, *8* (18), 1555−1572.

(30) Case, D. A.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S.R.; Cerutti, D. S.; Cheatham, T. E.; Cruzeiro, I. V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kovalenko, A.; Krasny, R.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Man, V.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wilson, L.; Wolf, R. M.; Wu, X.; Xiong, Y.; Xue, Y.; York, D. M.; Kollman, P. A. *AMBER2020*; 2020.

(31) Smith, R. D.; Carlson, H. A. Identification of Cryptic Binding Sites Using MixMD With Standard and Accelerated Molecular Dynamics. *J. Chem. Inf. Model.* **2021**, *61* (3), 1287−1299.

(32) Ghanakota, P.; Carlson, H. A. Moving Beyond Active-Site Detection: MixMD Applied to Allosteric Systems. *J. Phys. Chem. B* **2016**, *120* (33), 8685−8695.

(33) Ghanakota, P.; Carlson, H. A. Driving Structure-Based Drug Discovery Through Cosolvent Molecular Dynamics: Miniperspective. *J. Med. Chem.* **2016**, *59* (23), 10383−10399.

(34) Ghanakota, P.; DasGupta, D.; Carlson, H. A. Free energies and entropies of binding sites identified by MixMD cosolvent simulations. *J. Chem. Inf. Model.* **2019**, *59*, 2035.

(35) Lal Gupta, P.; Carlson, H. A. Cosolvent Simulations with Fragment-Bound Proteins Identify Hot Spots to Direct Lead Growth. *J. Chem. Theory Comput.* **2022**, *18* (6), 3829−3844.

https://doi.org/10.1021/acsabm.2c01079
ACS Appl. Bio Mater. XXXX, XXX, XXX−XXX

(36) Chan, W.; Carlson, H.; Traynor, J. Cosolvent Molecular Dynamics Simulation-Based Discovery of Potential Allosteric Sites on Regulator of G Protein Signaling 4. *FASEB J.* **2020**, *34* (S1), 1−1.
(37) Lexa, K. W.; Goh, G. B.; Carlson, H. A. Parameter choice matters: validating probe parameters for use in mixed-solvent simulations. *J. Chem. Inf Model* **2014**, *54* (8), 2190−9.
(38) Caldwell, J. W.; Kollman, P. A. Structure and Properties of Neat Liquids Using Nonadditive Molecular Dynamics: Water, Methanol, and N-Methylacetamide. *J. Phys. Chem.* **1995**, *99*, 6208−6219.
(39) Lexa, K. W.; Goh, G. B.; Carlson, H. A. Parameter Choice Matters: Validating Probe Parameters for Use in Mixed-Solvent Simulations. *J. Chem. Inf. Model.* **2014**, *54* (8), 2190−2199.
(40) Grabuleda, X.; Jaime, C.; Kollman, P. A. Molecular Dynamics Simulation Studies of Liquid Acetonitrile: New Six-Site Model. *J. Comput. Chem.* **2000**, *21* (10), 901−908.
(41) Caldwell, J. W.; Kollman, P. A. Structure and Properties of Neat Liquids Using Nonadditive Molecular Dynamics: Water, Methanol, and N-Methylacetamide. *J. Phys. Chem.* **1995**, *99* (16), 6208−6219.
(42) Ghanakota, P.; Carlson, H. A. Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *J. Med. Chem.* **2016**, *59* (23), 10383−10399.
(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(44) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput* **2015**, *11* (8), 3696−713.
(45) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J.C Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *Journal Of Computational Physics* **1977**, *23* (3), 327−341.
(46) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput* **2013**, *9* (9), 3878−88.
(47) Roe, D. R.; Cheatham, T. E., 3rd PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput* **2013**, *9* (7), 3084−95.
(48) Schrödinger *The PyMol Molecular Graphics System, 1.7.4.1*; 2020.
(49) Avrahami, L.; Licht-Murava, A.; Eisenstein, M.; Eldar-Finkelman, H. GSK-3 Inhibition: Achieving Moderate Efficacy With High Selectivity. *BBA-Proteins Proteom.* **2013**, *1834* (7), 1410−1414.
(50) Ngan, C. H.; Bohnuud, T.; Mottarella, S. E.; Beglov, D.; Villar, E. A.; Hall, D. R.; Kozakov, D.; Vajda, S. FTMAP: extended protein mapping with user-selected probe molecules. *Nucleic Acids Res.* **2012**, *40*, W271−5.
(51) Hussein, H. A.; Borrel, A.; Geneix, C.; Petitjean, M.; Regad, L.; Camproux, A.-C. PockDrug-Server: a new web server for predicting pocket druggability on holo and apo proteins. *Nucleic Acids Res.* **2015**, *43* (W1), W436−W442.
(52) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* **2009**, *10* (1), 1−11.
(53) Krasowski, A.; Muthas, D.; Sarkar, A.; Schmitt, S.; Brenk, R. DrugPred: a structure-based approach to predict protein druggability developed using an extensive nonredundant data set. *J. Chem. Inf Model* **2011**, *51* (11), 2829−2842.
(54) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf Model* **2009**, *49* (2), 377−389.
(55) Volkamer, A.; Kuhn, D.; Rippmann, F.; Rarey, M. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* **2012**, *28* (15), 2074−2075.

https://doi.org/10.1021/acsabm.2c01079
ACS Appl. Bio Mater. XXXX, XXX, XXX−XXX