Binding Affinities of Disheveled PDZ Protein-Ligand Complexes

Aarav Singh[1], Andrew Jubintoro[1], Harsha Kanchalara[1], Patrick Blankenberg[1], Dr. Jie Zheng[1]

[1]Stein Eye Institute, Department of Ophthalmology, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA

Disclosures: All authors have no potential conflict of interest.

ABSTRACT:

Wnt/ß-catenin signaling is critical for cell growth and development, with its hyperactive dysregulation implicated in the development of cancer. Current therapeutic research on inhibition of Wnt/ß-catenin signaling is impeded by the high cost of experimentally determining binding affinities. Consequently, interest has risen in screening potential inhibitors binding affinities with computational tools to reduce costs. Here, we test the validity of a computational molecular dynamics simulator, Binding Free Energy Estimator 2 (BFEE2), in determining peptide and small molecule ligand affinity for Wnt/ß-catenin signaling. We focus on the Dishevelled (DVL) PDZ domain, a key mediator in WNT signaling through its ability to bind to various peptide ligands. We analyze the binding affinities of several DVL PDZ domain-peptide complexes against previously established results to determine the validity of computational analysis. We conclude that computational molecular dynamics simulations are moderately effective in determining DVL peptide ligand complex affinity with an R^2 of 0.44; however, the simulation was not effective with small molecule drug inhibitors with an R^2 of 0.06.

1. INTRODUCTION:

The Wnt signalling pathway is crucial in cell growth, differentiation, and development. In the canonical cascade, Wnt ligands bind Frizzled receptors, which recruit Dishevelled-1 (DL1); DV1 subsequently promotes B-catenin stabilization. However, overactivation of DVL1 can drive excessive B-catenin accumulation and uncontrolled cell proliferation. In the eye specifically, this hyperproliferation contributes to tumorigenesis, such as retinoblastoma. Thus, identifying potent small-molecule inhibitors that target the DV1 PDZ domain is critical, as they can interfere with Wnt signaling and restore normal control of cell growth [8]. In this context, a convenient pre-screening tool would enable more efficient evaluation of new ligands, helping to prioritize those with stronger binding affinities to the Dishevelled protein for more detailed study.

The 22 DVL protein ligand molecules gathered for analysis were a combination of peptide ligand complexes and small molecule inhibitors downloaded from RCSB PDB [12]. They spanned from the DVL1, 2, and 3 domain isoforms. Protein-ligand complexes 2F0A, 3FY5, 2REY, 6LCA, 1MC7, 6LCB, 8WWR, and 8YR7 didn't have any ligand complexes associated with the protein so they were ineligible to use. 3CBZ protein ligand complex wasn't fully generated in RCSB and the ligand would fly off during an MD simulation, so that complex was ineligible. 6LCB and 6ZC8 generated too high a penalty to render an accurate BFEE2 simulation so they were ineligible. The complex's RCSB codes for the peptide ligands are 3CBX, 3CBY, 3CCO0, and 1L6O. The RCSB codes of the complexes with small molecule ligands are 6ZC6, 6ZBZ, 6ZBQ, 6ZC3, 6ZC5, 6ZC7, 2MX6, and 2KAW. Molecular dynamics simulations, specifically the Binding Free Energy Estimator 2 (BFEE2) software, and CGenFF, a small molecule parameterization software, were utilized to calculate the standard binding free energies of these complexes and compare them to experimental data already published.

2. METHODS:

PDB ID	Domain-Isoform	Ligand	Description
1L6O	DVL-2	Dapper peptide ligand	XENOPUS laevis DVL-2 PDZ domain bound to the Dapper peptide
6ZC6	DVL-3	QEN small molecule ligand	Human DVL-3 PDZ domain in complex with a small-molecule inhibitor "QEN"
3CC0	DVL-2	N3 inhibitory peptide ligand	Human DVL-2 PDZ domain complexed with the N3 inhibitory peptide

3CBY	DVL-2	N1 inhibitory peptide ligand	Human DVL-2 PDZ domain bound to inhibitory N1 peptide
3CBX	DVL-2	C1 inhibitory peptide	Human DVL-2 PDZ domain bound to inhibitory peptide C1
6ZBZ	DVL-3	QE5 small molecule ligand	Human DVL-3 PDZ domain with small-molecule inhibitor QE5
6ZBQ	DVL-3	QE2 small molecule ligand	Human DVL-3 PDZ domain with small-molecule inhibitor QE2
6ZC3	DVL-3	JOR small molecule ligand	Human DVL-3 PDZ domain with small-molecule inhibitor JOR
6ZC4	DVL-3	QEZ small molecule ligand	Human DVL-3 PDZ domain with small-molecule inhibitor QEZ
6ZC7	DVL-3	V31 small molecule ligand	Human DVL-3 PDZ domain with small-molecule inhibitor V31
2MX6	DVL-1	PHQ(WV) peptide ligand	Mouse DVL-1 PDZ domain bound to a short peptide PHQWV
2KAW	DVL-1	SUZ small molecule ligand	Mouse DVL-1 PDZ domain in complex with a small-molecule inhibitor (sulindac-derived)

Table 1. The 12 PDZ domains consisting of peptide or small molecule ligands are summarized in the table above.

2.1 Pre-Processing

Pre-processing of peptide molecules involved downloading the necessary .pdb files from RCSB and deleting any unnecessary water or solvent ions left over from the original X-ray or NMR technique. If the complex came as dimers, that is there were two proteins and two ligands, then the protein of one complex is paired with the ligand of the second complex, and the other protein and ligand were deleted so just a single protein-ligand complex existed. Additionally, for peptide ligands the COOH chain was de-protonated to COO- because that O- ion is well known to make hydrogen bonds within the PDZ domain.

For the peptide complexes, the pre-processing stage was straightforward. Visual Molecular Dynamics (VMD)'s AutoPSF tool was used to build a .psf file (a system connectivity/identity file that enumerates all atoms and how they're connected so the MD engine can build the bonded interaction lists). The topology file (general information containing composition and connectivity of the system) used to build the AutoPSF is "top_all_36_prot.rtf" [1].

The pdf and psf were then padded with 10 A around the molecular system to prevent interactions with periodic images and allow for realistic solvation. The entire system was solvated to 0.15 M NaCl to mimic biological conditions [5].

The small molecules similarly followed these pre-processing steps, but the ligand, not being made up of amino acids and instead an organic compound, needed to have the necessary topology and parameter files (information defining bond strengths and force constraints i.e how atoms interact) generated. Thus, CGenFF, a new parameterization software developed by

SilcsBio, was used. The .mol2 ligand file downloaded from RCSB was inputted into CGenff and generated the necessary topology files and parameterization with relatively low penalty error (sub 50). [2,3,4].

2.2 BFEE2

Binding Free Energy Estimator 2 is the protocol for the computation of standard binding free energies of protein-ligand complexes. Once the pre-processing stage was finished, the final complex.pdb and complex.psf files were input into BFEE2 along with the necessary CGenFF parameter files. After the protein and ligand files were inputted, the system was selected to pin down the protein in the middle of the system box, use reflecting boundaries as needed, and run at 300 K which is standard for NAMD simulations. Given the computational demand on an average workstation laptop, the UCLA Hoffman2 supercluster was used with NAMD GPU CUDA acceleration.

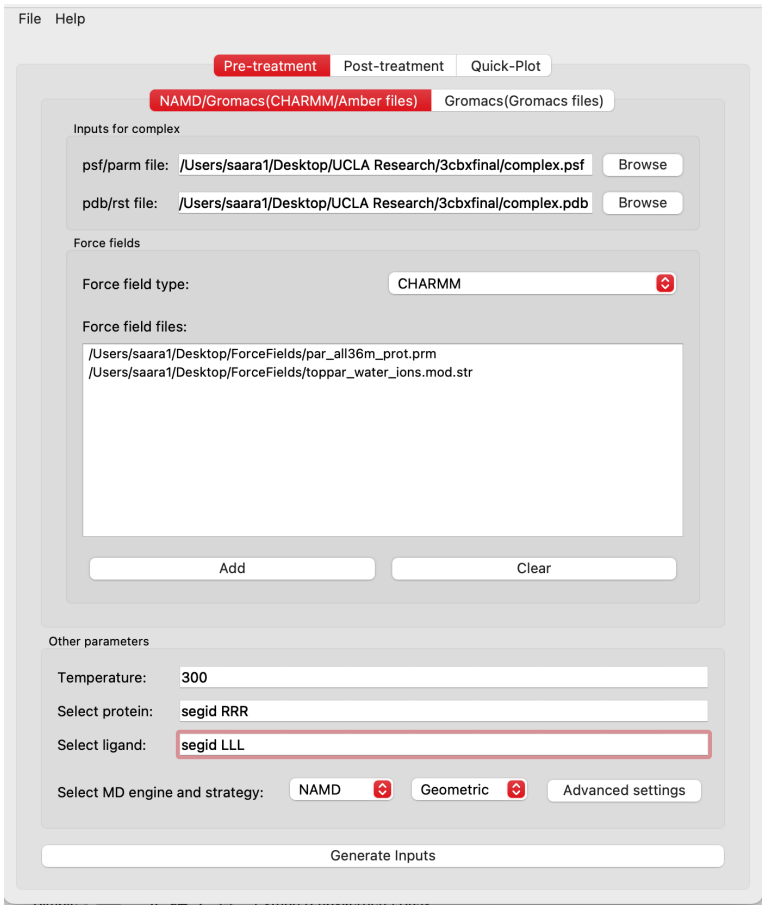

Figure 1. BFEE2 setup of 3CBX protein-ligand complex.

Each step of the BFEE2 protocol targets a specific component of the ligand's binding affinity by applying Colvars (software library used in MD simulations to bias runs) restraints to sample the energy landscape. Three different forms of restraints were used: harmonic, angular, and RMSD. A harmonic distance restraint keeps the ligand near the binding pocket by constraining the center-of-mass distance between the protein and ligand. Angular restraints limit the ligand's rotational freedom relative to the protein, enabling systematic sampling of orientational degrees of freedom such as eulerTheta, eulerPhi, eulerPsi, polarTheta, and polarPhi. The RMSD restraint ensures the ligand remains within a defined structural envelope, distinguishing bound from unbound states [6,7].

As part of Colvars, BFEE2 used Adaptive Biasing Force (ABF) method to provide enhanced sampling. ABF identifies which degrees of freedom, like ligand rotation or separation, are energetically unfavorable then applies a small, adaptive biasing force to help the molecule overcome these energy barriers and explore the configurational space better.

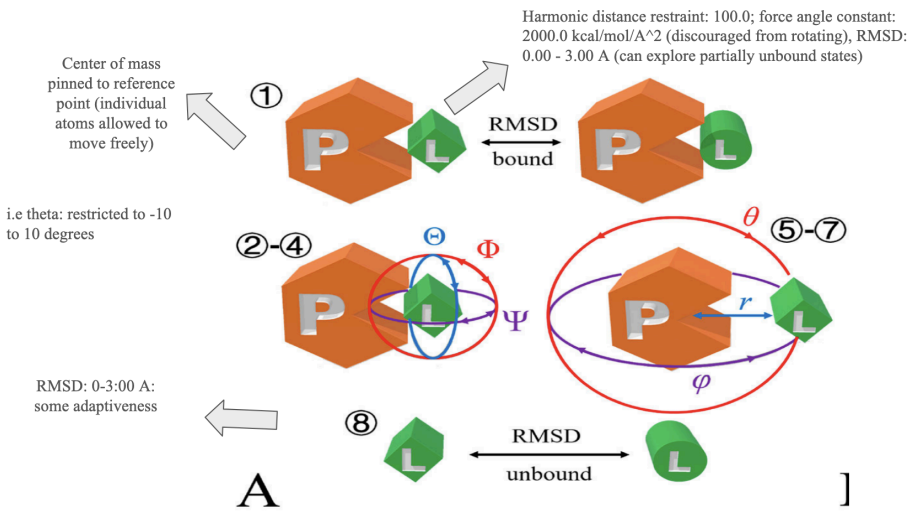

Figure 2. BFEE2 geometric route setup taken from BFEE2 paper; steps 1-8 along with their respective restraints [7].

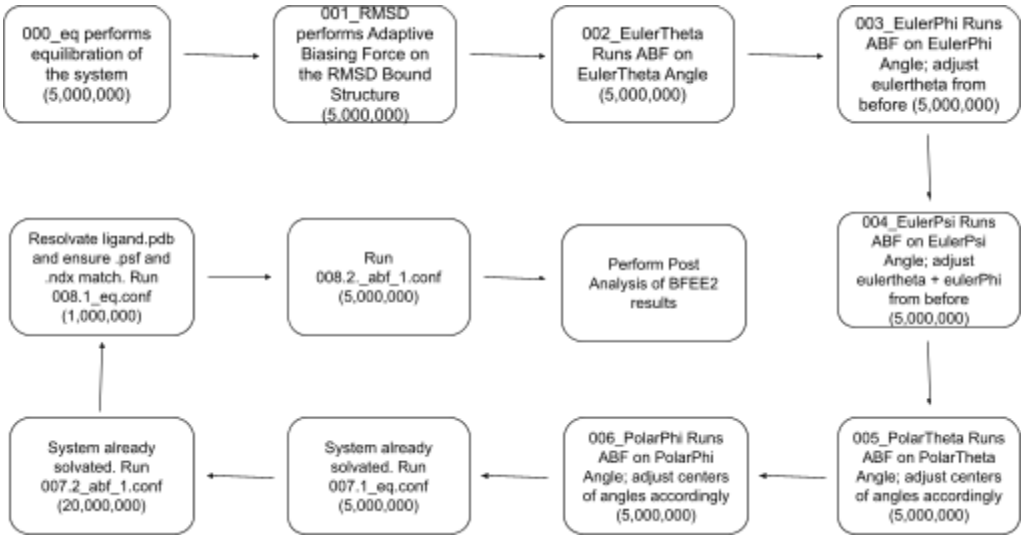

Figure 3. The default parameters of each step of the BFEE2 simulation are outputted with their corresponding steps. For small protein ligand complexes, the default parameterization approximately reaches RMSD convergence.

The default parameters of each step of the BFEE2 simulation are outlined. After running 5,000,000 steps in a BFEE2, a potential of mean force (PMF) is generated, displaying how the free energy changes along a selected collective variable (e.g. distance, angle, or orientation between protein and ligand). The binding free energy is then calculated using Boltzmann-weighted integration, which evaluates the probability of the ligand being bound vs unbound to yield the absolute binding free energy. Each step of the BFEE2 simulation corresponds to the free energy associated with a specific restraint, and summing up these values gives the total binding free energy of the ligand to the protein. To ensure the simulation was running properly, two key checkpoints were used.. First, in the 000_eq (equilibration) folder, the DCD trajectory file was loaded into VMD alongside the corresponding PSF structure file to visually confirm that the ligand remained in the binding pocket and did not drift or unbind prematurely, which is an essential sign of a properly equilibrated system. Second, for each simulation window, the RMSD plot was inspected to verify convergence, which is typically indicated by the RMSD values flattening out gradually toward the end of the run.

3. RESULTS:

After running all the MD simulations based on the BFEE2 setup designated above, BFEE2 computational binding affinities were compared to experimental binding affinities.

```
Results:
ΔG(site,c)          = -10.70 kcal/mol
ΔG(site,eulerTheta)  = -0.29 kcal/mol
ΔG(site,eulerPhi)    = -0.30 kcal/mol
ΔG(site,eulerPsi)    = -0.20 kcal/mol
ΔG(site,polarTheta)  = -0.07 kcal/mol
ΔG(site,polarPhi)    = -0.11 kcal/mol
(1/beta)*ln(S*I*C0)  = -17.56 kcal/mol
ΔG(bulk,c)          = 13.51 kcal/mol
ΔG(bulk,o)          = 6.61 kcal/mol

Standard Binding Free Energy:
ΔG(total)           = -9.10 kcal/mol
```

Figure 4. BFEE2 sample output from 3CBY protein-ligand complex.

PDB ID	**ΔGbind(calc) (kcal/mol)**	**ΔGbind(exp) (kcal/mol)**
1L6O	-8.06	-7.08
6ZC6	0.50	-6.67
3CCO	-7.86	-7.33
3CBY	-9.10	-8.13
3CBX	-8.37	-8.45
6ZBZ	3.42	-5.29
6ZBQ	-3.12	-5.62
6ZC3	-0.64	Negligible affinity (>1000[ITC])
6ZC4	8.44	-6.43
6ZC7	-5.93	-6.42
2MX6	-3.93	-7.07(pending data from PI)
2KAW	6.61	-6.97

Table 2. Comparison of computational vs experimental binding affinities for all protein-ligand complexes [8,9,10,11].

After performing all BFEE2 experiments, an R^2 test was performed to evaluate how well the computed binding free energies correlated with the experimental values. Graphs were split up based on if a protein-ligand complex had a peptide or small molecule ligand.

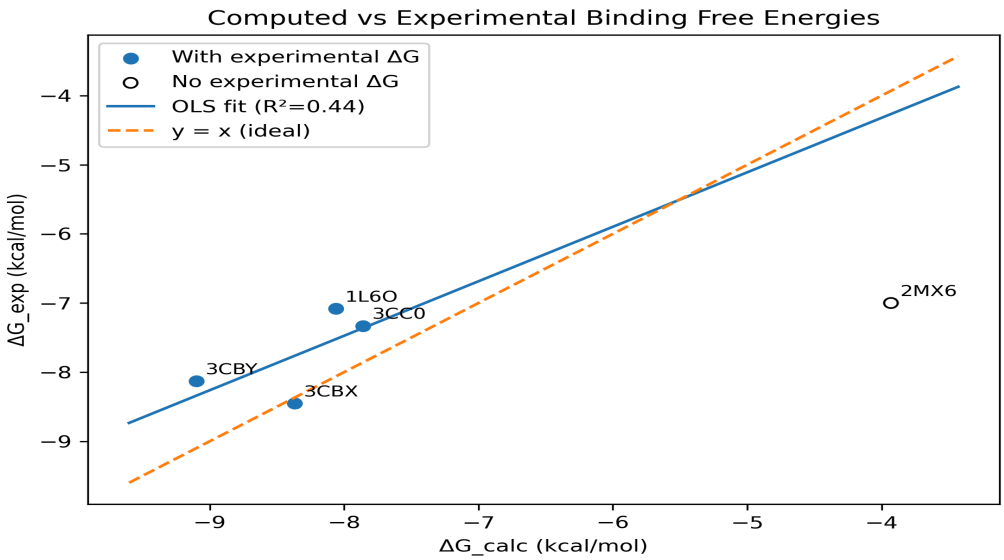

Figure 5. Scatter plot with regression line of peptide ligands. 2MX6 currently omitted as experimental data is pending.

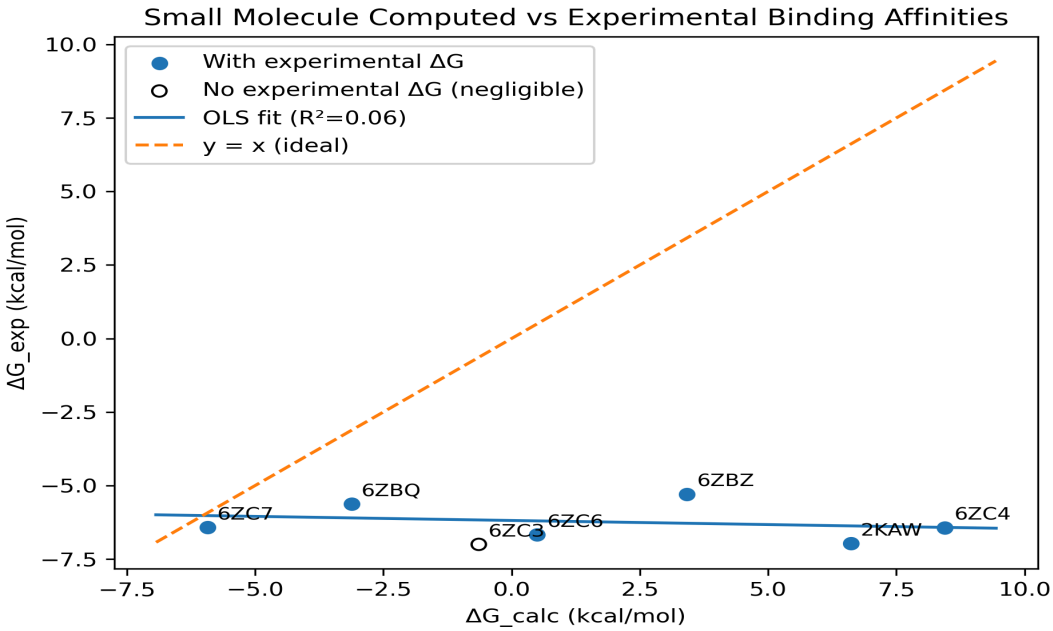

Figure 6. Scatter plot with regression line of small molecule ligands. 6ZC3 omitted as experimental delta G is negligible (>1000[ITC]).

4. DISCUSSION:

BFEE2 simulations indicate that the peptide-PDZ complexes produced a markedly higher R^2 (0.44 in the peptide subset) compared with the small molecule inhibitors that indicated a very low coefficient of determination ($R^2 = 0.06$). This difference stems from how peptide ligands have much better force-field coverage. All amino acid residues are native to CHARMM'S force field and have thus been extensively validated for its properties. This accurate parameterization allows simulations to generate experimental measurements more accurately. Small molecule inhibitors displayed almost no correlation with experimental delta G binding affinity ($R^2 = 0.06$). It can be reasoned that this is due to the CGenFF-derived topologies and parameters generated. Even though the program can conveniently generate a parameter file from a .mol2 ligand, the parameterization still has low but nonzero penalty scores for its approximations in dihedrals, charges, and nonbonded interactions. This is because the small molecules have more diverse chemical groups (halogens, rings, and heteroatoms) that still have not been extensively validated. These inaccuracies can accumulate during simulations as long as BFEE2's and shift the free energies by multiple kcal/mol [2,3,4,7].

The results of the BFEE2 give good quantitative predictions for peptide binding; however, small-molecule inhibitors will most likely require manual parameterization following CGenFF with programs like VMD FFTK(force field toolkit) that require quantum-mechanical calculations to obtain parameters.

PDB ID	Pitfalls and Caveats
1L6O	Peptide complex ran with no issues
6ZC6	Lone pair caused 008_RMSDUnbound to produce a "rattle shake algorithm" and crash. LP kept for 001-007 and taken out for 008_RMSDUnbound with just ligand
3CCO	Peptide complex ran with no issues
3CBY	Peptide complex ran with no issues
3CBX	Peptide complex ran with no issues
6ZBZ	CGenFF penalty errors low but still existent
6ZBQ	CGenFF penalty errors low but still existent
6ZC3	CGenFF penalty errors low but still existent
6ZC4	Lone pair caused 008_RMSDUnbound to produce a "rattle shake algorithm" and crash. LP kept for 001-007 and taken out for 008_RMSDUnbound with just ligand
6ZC7	Lone pair caused 008_RMSDUnbound to produce a "rattle shake algorithm" and crash. LP kept for 001-007 and taken out for 008_RMSDUnbound with just ligand
2MX6	2MX6 ligand not showing full connectivity. Ligand .str from CGenFF couldn't work with
2KAW	Sulindac generated higher penalty errors.

Table 3. Limitations encountered with each peptide ligand complex.

5. CONCLUSION:

BFEE2 proved useful as a pre-screening tool for dishevelled peptide ligands, with R^2 value 0f 0.44 showing moderate correlation to experimental data. In contrast, small molecule inhibitors showed almost no correlation ($R^2 = 0.06$), highlighting how CGenFF small ligand parameterization generates decent but not accurate enough estimates. Researchers investigating disheveled peptide proteins or other peptide ligands shouldn't need extensive manual intervention with parameterization or BFEE2, but should have access to a good GPU in order to conveniently run the NAMD simulations. Future research would involve perfecting BFEE2 runs to ensure the best RMSD convergence and doing manual parameterization of the small molecule files through tools like FFTK in order to obtain more accurate computational binding affinities.

REFERENCES:

1. Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Henin J, Jiang W, McGreevy R, Melo MCR, Radak BK, Skeel RD, Singharoy A, Wang Y, Roux B, Aksimentiev A, Luthey-Schulten Z, Kale LV, Schulten K, Chipot C, Tajkhorshid E. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys.* 2020;153:044130. doi:10.1063/5.0014475
2. Vanommeslaeghe K, Acharya C, Kundu S, Zhong S, Shim JE, Darian E, Guvench O, Lopes P, Vorobyov I, MacKerell AD Jr. CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem.* 2010;31:671-690. doi:10.1002/jcc.21367
3. Vanommeslaeghe K, MacKerell AD Jr. Automation of the CHARMM General Force Field (CGenFF) I: Bond perception and atom typing. *J Chem Inf Model.* 2012;52:3144-3154. doi:10.1021/ci300363c
4. Vanommeslaeghe K, Raman EP, MacKerell AD Jr. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *J Chem Inf Model.* 2012;52:3155-3168. doi:10.1021/ci3003649
5. Humphrey W, Dalke A, Schulten K. VMD—Visual molecular dynamics. *J Mol Graph.* 1996;14:33-38. doi:10.1016/0263-7855(96)00018-5
6. Fiorin G, Klein ML, Henin J. Using collective variables to drive molecular dynamics simulations. *J Chem Inf Model.* 2021;61(7):3104-3111. doi:10.1021/acs.jcim.1c00269 (BFEE2 overview)
7. Yan Y, Liu X, Wang F, Liu W, Sun Y, Zhao X, et al. BFEE2: A robust and versatile free energy estimation toolkit for molecular simulations. *Front Mol Biosci.* 2023;10:10082674. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10082674/ (BFEE2 in depth)
8. Zhang Y, Appleton BA, Wiesmann C, Lau T, Costa M, Hannoush RN, Sidhu SS. Inhibition of Wnt signaling by Dishevelled PDZ peptides. *Nat Chem Biol.* 2009 Apr;5(4):217-219. doi:10.1038/nchembio.152. Epub 2009 Mar 1. PMID: 19252499.
9. van der Wel PCA, Brogi M, Pusey M. Structure and dynamics of small-molecule inhibitors binding to Dvl1 PDZ: Insights from NMR methods. *Magn Reson.* 2021;2(2):355-371. doi:10.5194/mr-2-355-2021
10. Zhang J, Li J, Wang J, Li T, Chen D, Yang Z, Song X. Sulindac selectively inhibits Wnt signaling: Insight into an oncogene target. *PLoS One.* 2010;5(12):e14286. doi:10.1371/journal.pone.0014286
11. Tao Q, Yokota C, Puzon-Michel K, Pierfelice T, Lu H. Structural and functional analyses of Dishevelled PDZ interactions with peptides. *Dev Cell.* 2002;3(5):659-671. doi:10.1016/S1534-5807(02)00140-5
12. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* 2000;28:235-242. doi:10.1093/nar/28.1.235