Supporting Information

PyRod – Tracing Water Molecules in Molecular Dynamics Simulations

*David Schaller†, Szymon Pach† and Gerhard Wolber†,**

† Pharmaceutical and Medicinal Chemistry, Freie Universität Berlin, Königin-Luise-Str. 2+4, 14195 Berlin, Germany

**Corresponding Author, e-mail: gerhard.wolber@fu-berlin.de*

Table of contents

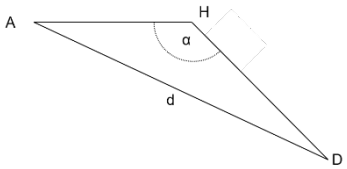
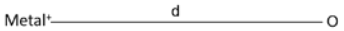
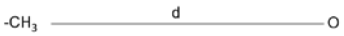
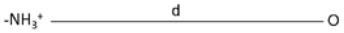
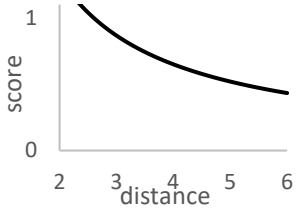
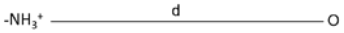
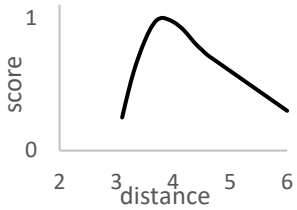
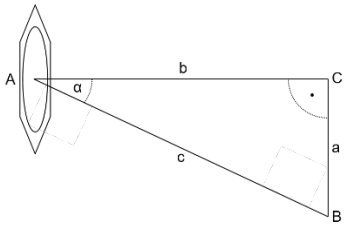
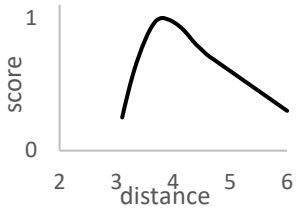
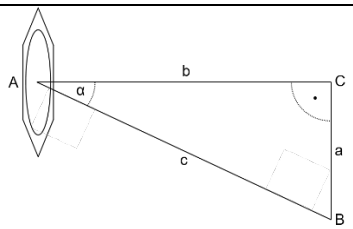
1. Feature Definition

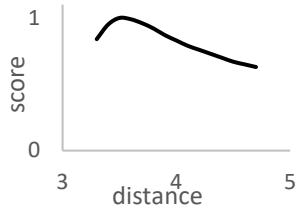
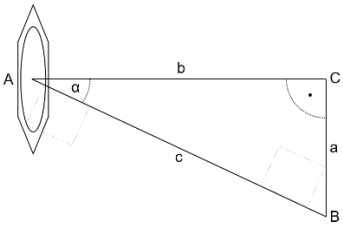
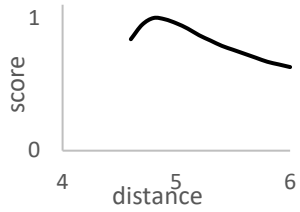
Table S1: Chemical feature definition and functional protein groups.

protein chemical feature	**selection criteria**	**corresponding ligand chemical features for interaction**
hydrogen bond acceptor	• nitrogen, oxygen or sulfur with acceptor capabilities • imidazole nitrogen checked for protonation state	• hydrogen bond donor
hydrogen bond donor	• nitrogen, oxygen or sulfur with donor capabilities • checked for protonation state	• hydrogen bond acceptor
hydrophobic interaction	• uncharged carbon or sulfur atoms not bonded to oxygen or nitrogen	• hydrophobic interaction
positive ionizable	• N-terminus, lysine amine, arginine guanidine, histidine imidazole • checked for protonation state	• negative ionizable • aromatic interaction
negative ionizable	• C-terminus, aspartate carboxylate, glutamate carboxylate, cysteine thiolate, serine hydroxylate, threonine hydroxylate, tyrosine hydroxylate • checked for protonation state	• positive ionizable
aromatic interaction	• phenylalanine benzene, tyrosine benzene, tryptophan indole (benzene and pyrrole), histidine imidazole (checked for protonation)	• aromatic interaction • positive ionizable
metal	• name specified via config-file	• hydrogen bond acceptor

2. Scoring Functions

Table S2: Scoring functions for chemical features in PyRod. HB – hydrogen bond, HI – hydrophobic interaction, II – ionizable interaction, AI – aromatic interaction. * scored per grid point.

type	geometry	cutoffs	scoring function
HB		$d_{oxygen} \leq 3.2$ $d_{nitrogen} \leq 3.3$ $d_{sulfur} \leq 3.9$ $\alpha \leq 130°$	occupancy
Metal		$d \leq 3$	occupancy, part of hydrogen bond acceptor and negative ionizable
HI		$d \leq 5$	number of hydrophobic atoms scaled by buriedness
II		$d \leq 6$	occupancy scaled by distance 
Cation-π*		$3.1 \leq d \leq 6$	occupancy scaled by distance 
π-Cation*		$\alpha < 30°$ $3.1 < b \leq 6.0$	occupancy scaled by distance 
$AI_{\pi\text{-stacked}}$*		$\alpha < 45°$ $3.3 \leq b \leq 4.7$ $a \leq 2.0$	occupancy scaled by distance

			
$AI_{t\text{-}stacked}$*		$\alpha < 45°$ $4.7 < b \leq 6.0$ $a \leq 0.5$ or $\alpha \geq 45°$ $4.6 \leq a \leq 6.0$ $b \leq 0.5$	occupancy scaled by distance 

3. PyRod combinatorial library parameters

Table S3: Parameters used for combinatorial library generation.

		CDK2	D3R	HIV1P
minimal	independent chemical features	3	3	3
	hydrogen bonding features	1	1	2
	ionizable features	0	1	0
	aromatic features	0	0	0
	hydrophobic features	1	1	2
maximal	independent chemical features	5	5	5
	hydrogen bond features	4	4	4
	ionizable features	1	1	1
	aromatic features	2	2	2
	hydrophobic features	3	3	3
number of features in super pharmacophore		15	16	15
number of generated combinations		816	2441	588

4. Early enrichment factors

Table S4: Early enrichment factors for best performing pharmacophore model generated by PyRod and for docking results from DOCK 3.6[1] as benchmarked in the DUD-E publication[2].

	PyRod	Dock 3.6
CDK2	30.3	14
D3R	7.7	4
HIV1P	54.6	5
ERα	-	15
$A_{2A}R$	-	22

5. Equilibration

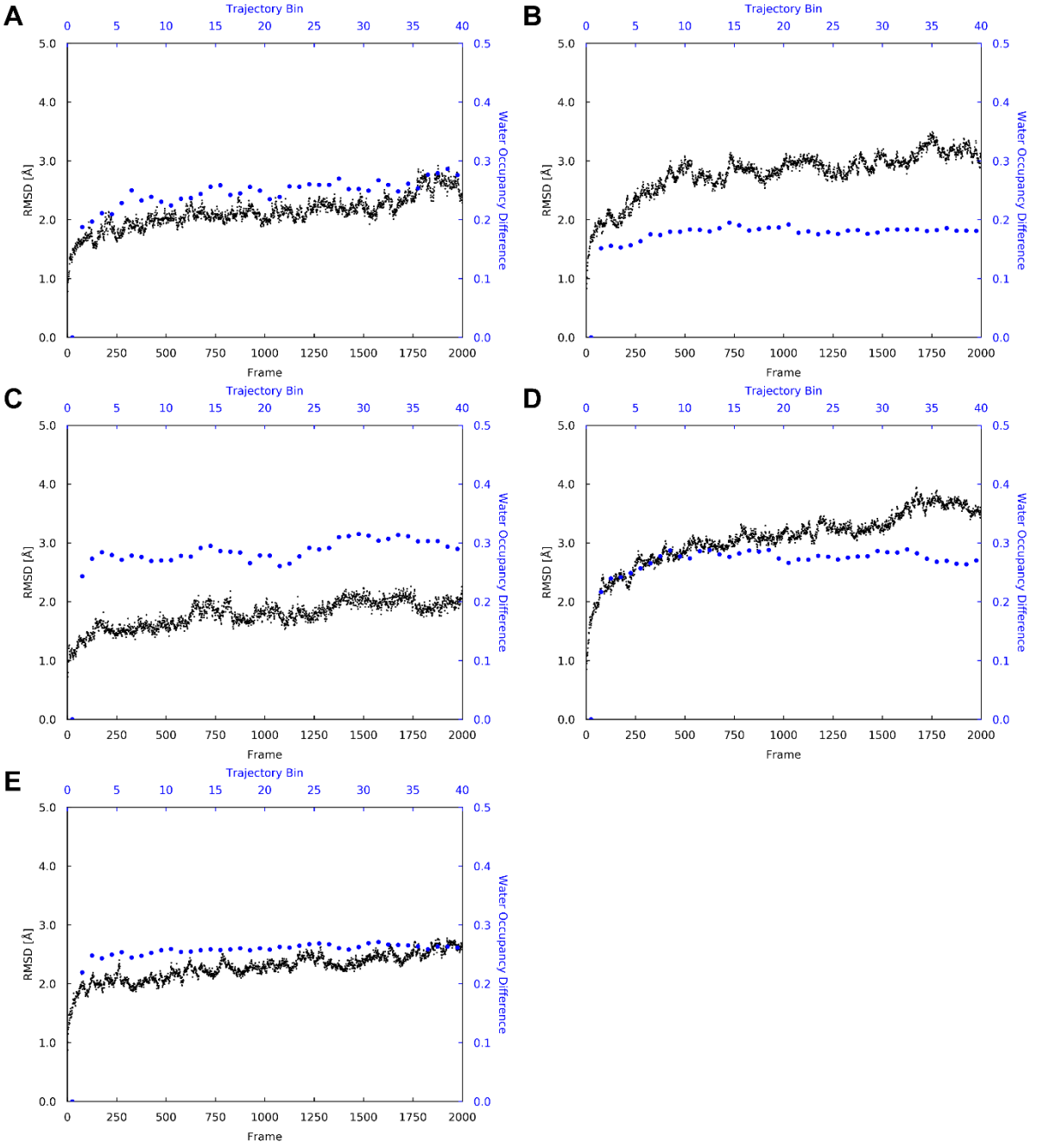


Figure S1: RMSD and water occupancy difference plots for one replication of each system tested, i.e. cyclin-dependent kinase 2 (5if1[3], **A**), dopamine D3 receptor (3pbl[4], **B**), HIV-1 protease (1nh0[5], **C**), estrogen receptor α (1xpc[6], **D**) and adenosine A_{2A} receptor (5iu4[7], **E**). RMSD of protein heavy atoms is shown in black and was calculated using VMD 1.9.2[37]. The water occupancy difference is shown in blue and was calculated as follows. The trajectory of each system was separated into trajectory bins of 50 frames. The first bin contains frames 1-50 and the last bin contains frames 1951-2000. PyRod was used for each trajectory bin to calculate water occupancies. The water occupancies of each bin were then compared grid-point wise to the water occupancies of the first bin, whereas a grid point pair was classified as different if the water occupancy values differ by more than 5. A water occupancy difference of 0.3 corresponds to a situation where 30 % of the grid points of that bin show a different water occupancy compared to the first bin.

6. References

(1) Coleman, R. G.; Carchia, M.; Sterling, T.; Irwin, J. J.; Shoichet, B. K. Ligand Pose and Orientational Sampling in Molecular Docking. *PLoS One* **2013**, *8* (10), e75992.

(2) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55* (14), 6582–6594.

(3) Ayaz, P.; Andres, D.; Kwiatkowski, D. A.; Kolbe, C.-C.; Lienau, P.; Siemeister, G.; Lücking, U.; Stegmann, C. M. Conformational Adaption May Explain the Slow Dissociation Kinetics of Roniciclib (BAY 1000394), a Type I CDK Inhibitor with Kinetic Selectivity for CDK2 and CDK9. *ACS Chem. Biol.* **2016**, *11* (6), 1710–1719.

(4) Chien, E. Y. T.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. *Science* **2010**, *330* (6007), 1091–1095.

(5) Brynda, J.; Rezacova, P.; Fabry, M.; Horejsi, M.; Stouracova, R.; Sedlacek, J.; Soucek, M.; Hradilek, M.; Lepsik, M.; Konvalinka, J. A Phenylnorstatine Inhibitor Binding to HIV-1 Protease: Geometry, Protonation, and Subsite-Pocket Interactions Analyzed at Atomic Resolution. *J. Med. Chem.* **2004**, *47* (8), 2030–2036.

(6) Blizzard, T. A.; Dininno, F.; Morgan, J. D.; Chen, H. Y.; Wu, J. Y.; Kim, S.; Chan, W.; Birzin, E. T.; Yang, Y. T.; Pai, L.-Y.; Fitzgerald, P. M. D.; Sharma, N.; Li, Y.; Zhang, Z.; Hayes, E. C.; Dasilva, C. A.; Tang, W.; Rohrer, S. P.; Schaeffer, J. M.; Hammond, M. L. Estrogen Receptor Ligands. Part 9: Dihydrobenzoxathiin SERAMs with Alkyl Substituted Pyrrolidine Side Chains and Linkers. *Bioorg. Med. Chem. Lett.* **2005**, *15* (1), 107–113.

(7) Segala, E.; Guo, D.; Cheng, R. K. Y.; Bortolato, A.; Deflorian, F.; Doré, A. S.; Errey, J. C.; Heitman, L. H.; IJzerman, A. P.; Marshall, F. H.; Cooke, R. M. Controlling the Dissociation of Ligands from the Adenosine A2A Receptor through Modulation of Salt Bridge Strength. *J. Med. Chem.* **2016**, *59* (13), 6470–6479.