NPC Natural Product Communications

2016
Vol. 11
No. 10
1499 - 1504

# Pharmacophore Models Derived From Molecular Dynamics Simulations of Protein-Ligand Complexes: A Case Study

**Marcus Wieder[a,b], Ugo Perricone[c], Thomas Seidel[a,*] and Thierry Langer[a]**

[a]*Department of Pharmaceutical Chemistry*, *Faculty of Life Sciences*, *University of Vienna*, *Vienna*, *Austria*
[b]*Department of Computational Biological Chemistry*, *Faculty of Chemistry*, *University of Vienna*, *Vienna*, *Austria*
[c]*Dipartimento di Scienze e Tecnologie Biologiche Chimiche e Farmaceutiche "STEBICEF"*, *Università di Palermo*, *Palermo*, *Italy*

*thomas.seidel@univie.ac.at*

**Received: February 12th, 2016; Accepted: May 3rd, 2016**

A single, merged pharmacophore hypothesis is derived combining 2000 pharmacophore models obtained during a 20 ns molecular dynamics simulation of a protein-ligand complex with one pharmacophore model derived from the initial PDB structure. This merged pharmacophore model contains all features that are present during the simulation and statistical information about the dynamics of the pharmacophore features. Based on the dynamics of the pharmacophore features we derive two distinctive feature patterns resulting in two different pharmacophore models for the analyzed system – the first model consists of features that are obtained from the PDB structure and the second uses two features that can only be derived from the molecular dynamics simulation. Both models can distinguish between active and decoy molecules in virtual screening. Our approach represents an objective way to add/remove features in pharmacophore models and can be of interest for the investigation of any naturally occurring system that relies on ligand-receptor interactions for its biological activity.

**Keywords:** Molecular Dynamics Simulation, Pharmacophore Modeling, Protein-Ligand Flexibility, Structure-based Modeling.

Pharmacophore models are defined as the ensemble of steric and electronic features that are necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or block) its biological response [1]. These features include H-bond acceptors (HBA), H-bond donors (HBD), positive and negative ionizable groups (PI/NI), as well as hydrophobic regions (H) and aromatic rings (AR). Structure-based pharmacophore models use the 3D structure of a protein-ligand complex to extract a set of essential pharmacophore features that describe the binding mode of the ligand. Such pharmacophore models can then be used to identify novel active compounds using virtual screening protocols [2-4].

Structure-based pharmacophore models face two difficulties: on one hand, they are sensitive to the atomic coordinates of the protein-ligand complex from which they are derived and on the other hand it is challenging to assess the derived pharmacophore features in a conscious and objective way. The first issue is inherently linked to concerns related to crystallography [5] since more than 80% of protein-ligand complexes in the Protein Data Bank (PDB) are determined using X-ray crystallography [6]. These concerns include notably the fidelity of ligand coordinates in protein-ligand complexes [7,8] and the influence of crystal contacts and solvent on protein structures [9,10]. Proteins and small organic molecules, such as natural products, or from pure synthetic origins are inherently dynamic and display a wide range of motions, which range from the vibrations of individual bonds to collective, large structural movements. The crystal structure of the protein-ligand complex represents only a single snapshot of a dynamic system and thus provides neither information about the conformational flexibility of the ligand, nor about motions of the residues in and near the binding pocket [11-13]. As a result, pharmacophore models derived from such structures might include features that are artifacts, caused either by the crystal environment or the single set of coordinates of the structure [14].

One very general way to avoid dependence on a single set of coordinates is the use of molecular dynamics (MD) simulations to generate multiple sets of coordinates and use these as the basis for pharmacophore model generation. MD simulations have proven to be invaluable for understanding dynamics of biomolecules [15,16], solvent effects [17] etc., and form the basis for advanced techniques, such as the calculation of free energies of protein-ligand binding [18].

In a recent article [19] we investigated the possibility of improving pharmacophore models using molecular dynamic simulations. In this case study we present preliminary results that extend the analysis conducted [19] for one protein-ligand system. We analyze the variability of the interaction partners of the pharmacophore model and analyze the occurrence of features as a function of time. From this analysis two pharmacophore models are derived based on the frequency of interactions and the time resolved dynamics of the pharmacophore features.

The used protein-ligand complex has the PDB code 2OJ9 and represents the crystal structure of the IGF-1R (insulin-like growth factor-1 receptor) kinase domain in complex with a benzimidazole inhibitor. Overexpression of IGF-1R has been demonstrated in a variety of tumors, including glioma, lung, ovary, breast, carcinomas, sarcomas, and melanoma [20]. This protein-ligand complex was chosen from the analyzed complexes in [19] because the pharmacophore model contains a balanced number of the most common features (3 hydrophobic features, 3 hydrogen bond donor features, 2 hydrogen bond acceptor features, 3 aromatic features).

Typically, most ligands of protein kinases bind in the hinge region at the folding cleft of the N- and C-lobes. Common scaffolds that bind this region contain two hydrogen bond features, usually a donor-acceptor pair that interact with the hinge backbone [21]. The PDB pharmacophore model displays the typical hydrogen bond

interaction pattern with MET103 and GLU101 (shown in Figure 2), as described in the literature [20].

A 20 ns MD simulation of the solvated protein-ligand structure was computed, for every 10 ps the coordinates were saved resulting in 2000 coordinate sets. In addition, we also considered the coordinates of the starting structure. A pharmacophore model was derived from each structure that was obtained during the MD simulation. For further analysis a consensus pharmacophore model (a merged pharmacophore model) was generated which consists of all features that are present either in the experimental structure or in any snapshot generated during the multiple MD simulations, thus it incorporates information about the dynamics of the protein-ligand complex. The frequency with which individual features are present permits the ranking/prioritization of the features if needed and to detect outliers, i.e., features seen only rarely. Additionally, the interaction partners for each pharmacophore feature were analyzed and an interaction map (interaction matrix) was constructed. An interaction map allows quantitative analysis of the interaction partners of the pharmacophore features. As a final step, the frequency of the pharmacophore features was analyzed as a function of time. Combining these different analysis methods for the dynamics of the pharmacophore features allows the conscious derivation of pharmacophore models that are different from the corresponding model obtained with the PDB structure.

The trajectory of the protein-ligand complex was visually inspected to ensure that no large scale movements took place and that the ligand remained within the binding site at all times. The root mean square deviation (RMSD) values of the Cα-atoms of the protein were in an acceptable range (the RMSD plot for the ligand and protein backbone is shown in Figure 2). In contrast, the RMSD values of the ligand are rather high (ranging to a maximum of 9 Angstrom). In Figure 2, representative structures for the first (Figure 3.1) to the fourth quarter (Figure 3.4) of the MD simulation are shown. As can be seen, the pyridine moieties of the ligand rotate freely during the MD simulation, but also the translation of the imidazole contributes to the elevated RMSD values. Nevertheless, the protein-ligand complex was stable during the simulation.

The frequency of the specific pharmacophore features and the interaction map for 2OJ9 are shown in Figure 1.

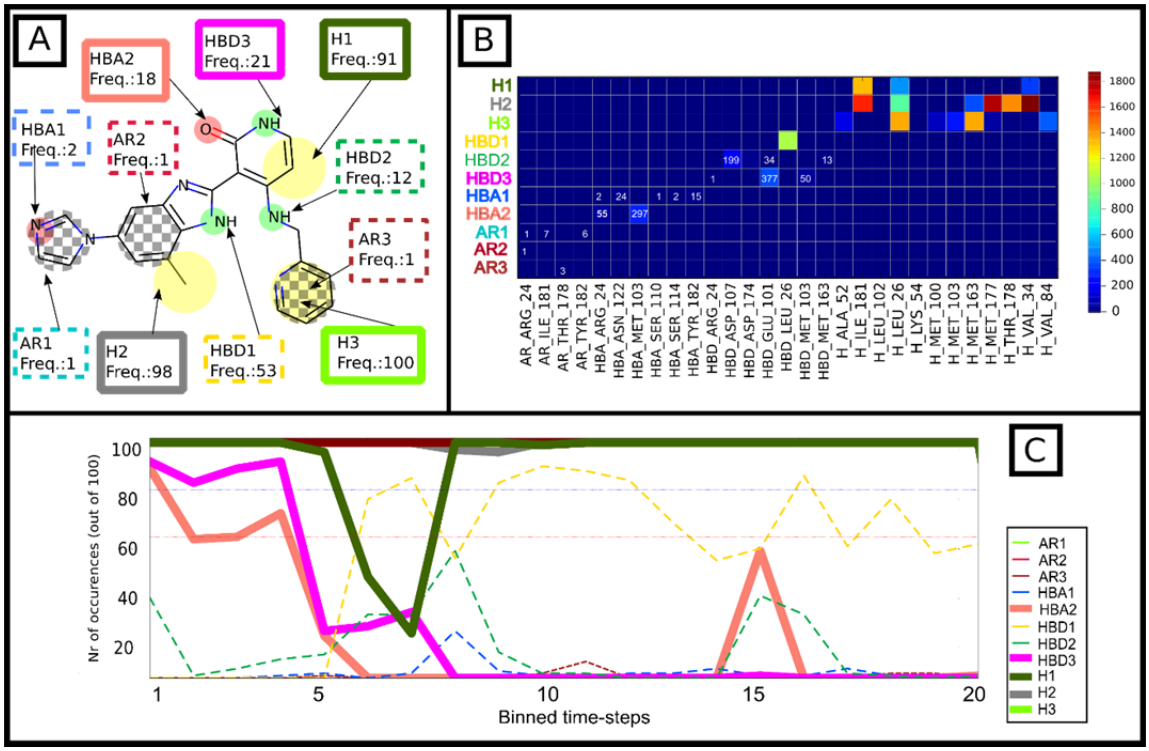


**Figure 1:** Analysis of the dynamics of the pharmacophore features of 2OJ9. (A) Shows a 2D representation of the ligand with the pharmacophore features mapped on the structure. Yellow spheres represent hydrophobic (H) interactions, small green circles indicate hydrogen bond donor (HBD) features and small red circles hydrogen bond acceptor (HBA) features. Black and white chess-fields represent aromatic features (AR). For every feature a box is shown, providing the feature name that is used in part (B) and (C) of the figure and information about the statistical frequency (given in percent and rounded to integers) of the specific feature. Dashed outlined boxes indicate features that are not present in the PDB pharmacophore; continuous lined boxes indicate features that are present in the PDB pharmacophore. The color of the boxes are consistent with the colored row labels in part (B) and the colored lines in part (C) of the figure. (B) Shows the interaction matrix as a heat-map. The row names indicate the pharmacophore features and the column names the interaction partners of the pharmacophore features. The column names consist of three parts, separated by underscores: the first part indicates the feature type, the second part the 3-letter amino acid code and the third part the residue number of the amino acid. The entries in the interaction map are color coded, ranging from dark blue to dark red (as shown in the legend of Figure 1B.). The absolute values of the cells in the interaction map are written as numbers for all feature types other than hydrophobic features, if the number of interaction is below 400. (C) Shows the statistical frequency of the features as a function of time. Thick enclosing lines indicate pharmacophore features that were present in the pharmacophore model obtained with the PDB structure, whereas thin, dashed lines indicate features that are not present in the PDB structure. The Y-axis corresponds to the number of occurrences of the specific feature per binned time-step and the X-axis corresponds to the binned time-steps. For a detailed description of this plot see the Methods Section, e.g. 'Frequency Plot'.

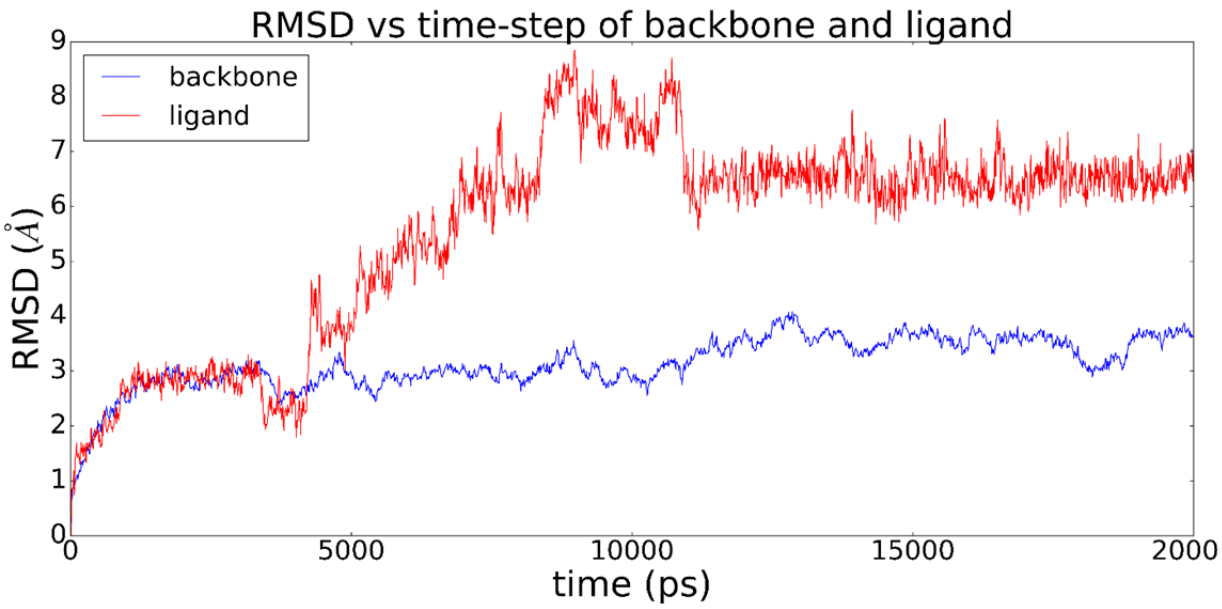


**Figure 2:** The root mean square deviation (RMSD) in Angstrom (Å) of the protein backbone (in blue) and the ligand (in red) as a function of time for the analyzed protein-ligand.

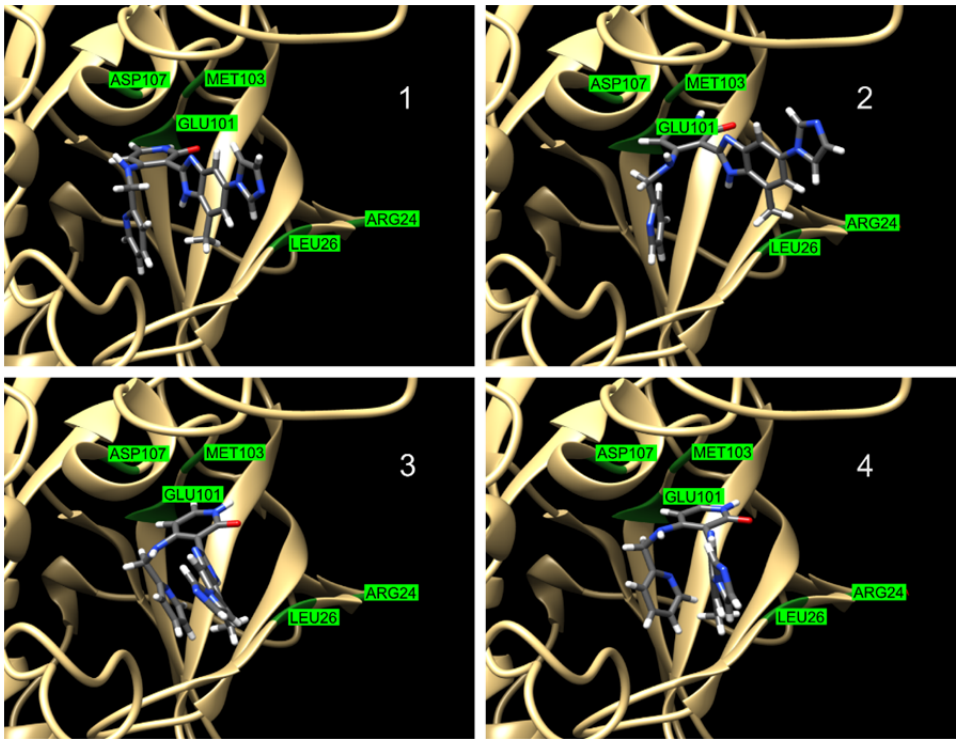


**Figure 3:** The ligand inside the binding pocket is shown at 4 different timesteps. The length of the MD simulation is divided into 4 equally long parts and clustering is performed based on the RMSD of the ligand. A representative ligand structure is extracted from the most populated cluster and shown from Figure 3.1 (representative structure of most populated cluster from 0 to 5 ns) to 3.4 (representative structure of most populated cluster from 15 to 20 ns). The amino acids that are most common in hydrogen bond interactions are explicitly labeled.

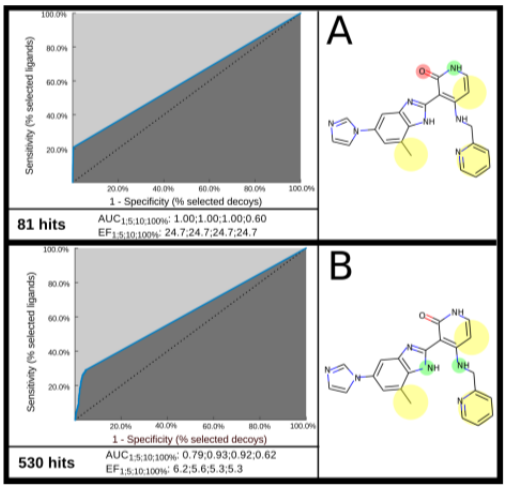


**Figure 4:** The receiver operating characteristic (ROC) curves and the two pharmacophore models are shown. In the ROC curve the true positive rate on the Y axis is plotted against the false positive rate on the X axis. The number of total hits and the enrichment factor (EF) are shown at 1%, 5%, 10% and 100%, respectively. In (A) the pharmacophore model obtained with the PDB structure and the virtual screening results are shown. In (B) the pharmacophore models with two MD derived hydrogen bond donor features and the virtual screening results are shown. For the description of the graphical 2D representation of the pharmacophore features see legend of Figure 1.

The initial pharmacophore hypothesis (shown in Figure 1A) includes 5 pharmacophore features (which will be called PDB features) and during the MD simulation 6 additional pharmacophore features (3 aromatic features, 1 hydrogen bond acceptor and 2 hydrogen bond donor features) are revealed (which will be called MD derived features).

As can be seen in Figure 1A, most of the MD derived pharmacophore features have a lower statistical frequency than the PDB features – only the MD derived feature HBD1 occurs more often than either HBD3 or HBA2 (both are PDB features). A further observation that can be drawn from Figure 1A is that hydrophobic features are far more stable during the MD simulation than hydrogen bond features and that aromatic features are the most unstable feature type. This is in accordance with our previous findings [19].

A closer look at Figure 1B reveals why most hydrophobic features have such stable and high frequencies – they interact with multiple interaction partners at the same time, thus preserving the interaction even in the case when one interaction partner leaves the range of influence of the ligand. It should be noted that the presented hydrogen bond features also have multiple interaction partners, but, indicated by the numbers for the different interaction partners for the hydrogen bond features, this feature type changes rarely between them, and if so, the change is slow and infrequent.

Figure 1C shows additional time resolved information about the frequencies of the pharmacophore features. As can be seen, the three hydrophobic features occur steadily above 95% for all binned time-steps – with the exception of H1 between the binned time-step 5 and 8. Around the same binned-time steps the frequency of HBA2 and HBD3 (both PDB features) drops and HBD1 (MD derived feature) appears with a subsequent frequency of around 70%. The high frequency of HBD1 is only partly represented in the total frequency (as seen in Figure 1A), since the feature was not present for the first quarter of the simulation. Our analysis provides an explanation for what happens at these binned time-steps. The RMSD value of the ligand starts to rise (shown in Figure 2) and the movement of the ligand results in a change of the presented interaction partner, and, therefore, we observe the drop in the frequencies for HBD3 and HBA2.

The data presented in Figure 1, especially in Figure 1C, suggest that the pharmacophore model which is appropriate for the first quarter of the simulation (based on the frequencies of the features) does not represent the second half of the simulation. Considering the different frequencies of the pharmacophore features, two pharmacophore models are proposed: The first model contains the three hydrophobic features and HBD3 and HBA2 – this is the pharmacophore model derived from the PDB structure (and will be called subsequently PDB pharmacophore model). The second pharmacophore model contains the three hydrophobic features, but HBA2 and HBD3 are exchanged in favor of HBD1 and HBD2 (and this model will be called MD derived pharmacophore model). These two pharmacophore models represent the pharmacophore features with different frequencies in the beginning and at the end of the MD simulation. Especially in the light of the work in [20], the reported findings are interesting. Although in the presented study a different tautomer was used than in [20], the typical hydrogen pattern with MET103 and GLU101 is present. However, it appears as if the interaction with LEU26 and ASP107 (as shown in Figure 1B) can also play an important role. In the following section the virtual screening results with these two pharmacophore models against a library of known active and calculated decoys will be shown and discussed in detail.

The screening results (the receiver operator curve, enrichment factor and number of total hits) for the different pharmacophore models are shown in Figure 4. Both pharmacophore models are able to discriminate between actives and decoys, and thus both provide good early enrichment.

The PDB pharmacophore model gives rise to 81 hits and the enrichment factor at 1.5% is 24.7. The pharmacophore hypothesis is able to retrieve 45 of the 226 active compounds.

The MD derived pharmacophore model leads to 530 hits, whereas the early enrichment factor at 1.5% is 6.2. The pharmacophore model retrieves 66 of the 226 active compounds in the library.

A closer look at the hit-list obtained with both pharmacophore models reveals that the PDB pharmacophore model retrieves 33 active molecules that are not present in the hit-list obtained with the MD derived pharmacophore model. The MD derived model retrieves 54 unique hits – the hit-list of both pharmacophore models share only 12 active molecules. This is not surprising since the pharmacophore models are different and represent two distinct interaction modes.

These two pharmacophore models can be used together – the PDB pharmacophore model is more likely able to distinguish between active and decoy models, but the MD derived pharmacophore model correctly identifies a higher number of active molecules. Since both models share only 12 active molecules in the resulting hit list, combining the results of these models results in a higher number of active candidates than only using the PDB pharmacophore.

In conclusion, MD simulations can reveal otherwise hidden pharmacophore features that are not present in the pharmacophore model derived from the experimental crystal structure. Using

additional information obtained from MD simulation, i.e. time resolved frequency information and interaction plots, it is possible to construct pharmacophore models that integrate the dynamic of the ligand inside of the binding pocket. Furthermore, this approach provides an objective way to add MD derived pharmacophore features to PDB derived pharmacophore models or, on the other side, remove PDB features that are less important based on the observed frequencies.

## Experimental

***PDB structure preparation:*** The quality and correctness of the PDB structure was audited using the WHAT IF web interface to WHAT CHECK [22,23]. The structure was analyzed with the software PropKa 3.1, and the protonation state of the protein and the ligand was assigned under the assumption of a solution pH of 7 [24,25]. Furthermore, in any calculations the ligand-tautomer suggested by the PDB was used.

***Retrieving representative structures:*** Chimera was used to cluster the MD simulation based on the RMSD deviation of the ligand [26]. The MD simulation was divided in 4 parts (from 1-5 ns, 5-10 ns, 10-15 ns and 15-20 ns) and the parts were individually analyzed. Representative structures for the most populated cluster in the 4 quarters were retrieved.

***MD simulation:*** We used the CHARMM-GUI web interface [27] to solvate the protein and set up the simulations. All MD simulations were carried out with CHARMM [28], utilizing the CHARMM/OpenMM coupling [29]. Parameters and molecular topologies for the ligand was generated based on the CGenFF force field [30]. The protein-ligand complex was solvated in cubic boxes of TIP3P water. The water box for 2OJ9 consists of 20344 water molecules and a total of 66062 atoms. Electrostatic interactions were computed using the particle-mesh-Ewald method [31]. SHAKE was used to keep all bonds involving hydrogen atoms rigid. After initial equilibration for 500 ps with a 1 fs time step, the system was simulated at 303.15 K for 20 ns using Langevin dynamics; the pressure was kept around 1 atm by a Monte Carlo barostat. The time step during the production calculation was 2 fs; coordinates were saved every 10 ps, resulting in 2000 coordinate sets for one simulation.

***RMSD analysis:*** The stability of the simulations was monitored by computing the root mean square deviations for the protein and ligand, using the MDAnalysis package [32]. The RMSDs were calculated as follows: all coordinates saved during the MD were fitted against the starting structure based on the coordinates of the Cα-atoms of the protein. Using the starting structure as reference, we computed for these reoriented coordinates the RMSD of the Cα-atoms for the protein and the RMSD of the heavy atoms of the ligand.

***Pharmacophore model analysis:*** For analysis of the pharmacophore models all water molecules were discarded. LigandScout 4.09.1 was used to generate a structure based pharmacophore model for each frame saved during the MD simulation (2000 pharmacophore models in total) and for the PDB structure [33].

The default feature constraints for the generation of the pharmacophore models were used as described in the manual [34]. The resulting 2001 pharmacophore models were analyzed as follows. Each pharmacophore feature has 2 properties: the ligand atoms that are part of the feature, and the feature type. If both properties of a pharmacophore feature were present in two models, then this feature was considered identical and the frequency count of this specific feature was incremented. In this manner we obtained statistics for how often a certain feature was present during the course of the MD simulation. Separate statistics were made for features not present in the PDB pharmacophore model, i.e., features only seen during the MD simulation. Using this frequency information, the merged pharmacophore model encompassing all features seen during the simulations was constructed by mapping the features on a representative 2D structure of the ligand.

Virtual screening was performed using known active and calculated decoy molecules obtained from the DUD-E database [35]. The database provided 226 active and 9395 decoys. All molecules were prepared as libraries for the screening using the command line tool idbgen provided by LigandScout. Conformers were generated using the icon best option in idbgen; this option produces a maximum number of 200 conformations for each molecule processed.

***Interaction matrix:*** The columns of the interaction matrix indicate all amino acid residues that are involved in a pharmacophore feature at some point during the MD simulations; the rows designate all pharmacophore features and the values in the matrix indicate how often a specific amino acid was involved in a specific pharmacophore feature. In this way it is possible to analyze the number of interaction partners and also their statistical frequency. The numeric values were coded as a heat map – the colors range from blue (zero interaction) to dark red (interaction at every time step). The numeric values for hydrogen bond and aromatic interactions are given explicitly in the heat map for values below 400. The interaction map was generated using the python package matplotlib [36].

***Frequency plot:*** For the MD simulations a frequency plot was calculated. This plot shows the occurrence of the features as a function of time. This is calculated as follows: The pharmacophore models are chronologically sorted and for every pharmacophore feature an occurrence list is calculated. Every time a pharmacophore model at a specific time step displays a specific feature 1 is inserted at the time step defined position in the list, otherwise 0 gets inserted. This results in a list with 2001 entries for every pharmacophore feature, which contains zeros and ones. In the end these lists are reduced by summing over chunks of 100 entries – resulting in a new list with 20 entries containing numbers between 100 and 0. To obtain a graphical representation, these lists are subsequently plotted using the python package matplotlib.

## References

[1] Ganellin C, Lindberg P, Mitscher L. (**1998**) Glossary of terms used in medicinal chemistry. *Pure and Applied Chemistry*, ***70***, 1129-1143.
[2] Yang S-Y. (**2010**) Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discovery Today*, ***15***, 444-450.
[3] Sanders MPA, McGuire R, Roumen L, de Esch IJP, de Vlieg J, Klomp JPG. (**2012**) From the protein's perspective: the benefits and challenges of protein structure-based pharmacophore modeling. *Medicinal Chemistry Communications*, ***3***, 28-38.
[4] Mannhold R, Kubinyi H, Folkers G. (**2006**) *Pharmacophores and Pharmacophore Searches.* Wiley-VCH, Weinheim, 1-395.

[5] Davis AM, St-Gallay SA, Kleywegt GJ. (**2008**) Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discovery Today*, ***13***, 831-841.
[6] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. (**2000**) The Protein Data Bank. *Nucleic Acids Research*, ***28***, 235-242.
[7] Liebeschuetz J, Hennemann J, Olsson T., Groom CR. (**2012**) The good, the bad and the twisted: a survey of ligand geometry in protein crystal structures. *Journal of Computer-Aided Molecular Design*, ***26***, 169-183.
[8] Reynolds C. (**2014**) Protein-ligand cocrystal structures: We can do better. *ACS Medicinal Chemistry Letters*, ***5***, 727-729.
[9] Davis AM, Teague SJ, Kleywegt GJ. (**2003**) Application and limitations of x-ray crystallographic data in structure-based ligand and drug design. *Angewandte Chemie International Edition*, ***42***, 2718-2736.
[10] Terada T, Kidera A. (**2012**) Comparative molecular dynamics simulation study of crystal environment effect on protein structure. *Journal of Physical Chemistry B*, ***116***, 6810-6818.
[11] Mirjalili V, Feig M. (**2013**) Protein structure refinement through structure selection and averaging from molecular dynamics ensembles. *Journal of Chemical Theory and Computation*, ***9***, 1294-1303.
[12] Mirjalili V, Noyes K, Feig M. (**2014**) Physics-based protein structure refinement through multiple molecular dynamics trajectories and structure averaging. *Proteins: Structure, Function, and Bioinformatics*, ***82***, 196-207.
[13] Whitesides GM, Krishnamurthy VM. (**2005**) Designing ligands to bind proteins. *Quarterly Reviews of Biophysics*, ***38***, 385-395.
[14] Barillari C, Marcou G, Rognan D. (**2008**) Hot-spots-guided receptor-based pharmacophores (HS-Pharm): a knowledge-based approach to identify ligand-anchoring atoms in protein cavities and prioritize structure-based pharmacophores. *Journal of Chemical Information and Modeling*, ***48***, 1396-410.
[15] Adcock SA, McCammon JA. (**2006**) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical Reviews*, ***106***, 1589-615.
[16] Karplus M, McCammon JA. (**2002**) Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, ***9***, 646-652.
[17] Soares CM, Teixeira VH, Baptista AM. (**2003**) Protein structure and dynamics in nonaqueous solvents: insights from molecular dynamics simulation studies. *Biophysical Journal*, ***84***, 1628-1641.
[18] Deng Y, Roux B. (**2004**) Computations of standard binding free energies with molecular dynamics simulations. *Journal of Physical Chemistry B*, ***108***, 16567-16576.
[19] Wieder M, Perricone U, Boresch S, Seidel T, Langer T. (**2016**) Evaluating the stability of pharmacophore features using molecular dynamics simulations. *Biochemical and Biophysical Research Communications*, ***470***, 685-689.
[20] Velaparthi U, Wittman M, Liu P, Stoffan K, Zimmermann K, Sang X, Carboni J, Li A, Attar R, Gottardis M, Greer A, Chang CY, Jacobsen BL, Sack JS, Sun Y, Langley DR, Balasubramanian B, Vyas D. (**2007**) Discovery and initial SAR of 3-(1H-benzo[d]imidazol-2-yl)pyridin-2(1H)-ones as inhibitors of insulin-like growth factor 1-receptor (IGF-1R). *Bioorganic & Medicinal Chemistry Letters*, ***17***, 2317-2321.
[21] Xing L, Klug-Mcleod J, Rai B, Lunney EA. (**2015**) Kinase hinge binding scaffolds and their hydrogen bond patterns. *Bioorganic & Medicinal Chemistry*, ***23***, 6520-6527.
[22] Hooft RWW, Vriend G, Sander C, Abola EE. (**1996**) Errors in protein structures. *Nature*, ***381***, 272-272.
[23] Vriend G. (**1990**) WHAT IF: a molecular modeling and drug design program. *Journal of Molecular Graphics*, ***8***, 52-56.
[24] Olsson MHM, Søndergaard CR, Rostkowski M, Jensen JH. (**2011**) PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation*, ***7***, 525-537.
[25] Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. (**2011**) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *Journal of Chemical Theory and Computation*, ***7***, 2284-2295.
[26] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. (**2004**) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, ***25***, 1605-1612.
[27] Steffen C, Thomas K, Huniar U, Hellweg A, Rubner O, Schroer A. (**2010**) TmoleX--a graphical user interface for TURBOMOLE. *Journal of Computational Chemistry*, ***31***, 2967-2970.
[28] Brooks B, Brooks C. (**2009**) CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry*, ***30***, 1545-1614.
[29] Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi Y, Jo S, Pande VS, Case DA, Brooks CL, MacKerell AD, Klauda JB, Im W. (**2016**) CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of Chemical Theory and Computation*, ***12***, 405-413.
[30] Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD. (**2010**) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, ***31***, 671-690.
[31] Darden T, York D, Pedersen L. (**1993**) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics*, ***98***, 10089-10092.
[32] Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O. (**2011**) MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, ***32***, 2319-2327.
[33] Wolber G, Langer T. (**2005**) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling*, ***45***, 160-169.
[34] LigandScout3 Manual, http://www.inteligand.com/ligandscout3/manual/ph__pharmacophore_feature_definitions.html (accessed December 15, 2015).
[35] Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. (**2012**) Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *Journal of Medicinal Chemistry*, ***55***, 6582-6594.
[36] Hunter JD. (**2007**) Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, ***9***, 90-95.