frontiers in Pharmacology

ORIGINAL RESEARCH
published: 24 April 2018
doi: 10.3389/fphar.2018.00395

Aromatic Rings Commonly Used in Medicinal Chemistry: Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities

*Marcelo D. Polêto[1], Victor H. Rusu[2], Bruno I. Grisci[3], Marcio Dorn[3], Roberto D. Lins[4] and Hugo Verli[1]**

[1] Grupo de Bioinformática Estrutural, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [2] Swiss National Supercomputing Centre, Lugano, Switzerland, [3] Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [4] Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, Brazil

OPEN ACCESS

Edited by:
Adriano D. Andricopulo, University of São Paulo, Brazil

Reviewed by:
Antonio Monari, Université de Lorraine, France
Gustavo Trossini, Universidade de São Paulo, Brazil

***Correspondence:**
Hugo Verli
hverli@cbiot.ufrgs.br

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: *09 November 2017*
Accepted: *05 April 2018*
Published: *24 April 2018*

Citation:
Polêto MD, Rusu VH, Grisci BI, Dorn M, Lins RD and Verli H (2018) Aromatic Rings Commonly Used in Medicinal Chemistry: Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities. Front. Pharmacol. 9:395. doi: 10.3389/fphar.2018.00395

The identification of lead compounds usually includes a step of chemical diversity generation. Its rationale may be supported by both qualitative (SAR) and quantitative (QSAR) approaches, offering models of the putative ligand-receptor interactions. In both scenarios, our understanding of which interactions functional groups can perform is mostly based on their chemical nature (such as electronegativity, volume, melting point, lipophilicity etc.) instead of their dynamics in aqueous, biological solutions (solvent accessibility, lifetime of hydrogen bonds, solvent structure etc.). As a consequence, it is challenging to predict from 2D structures which functional groups will be able to perform interactions with the target receptor, at which intensity and relative abundance in the biological environment, all of which will contribute to ligand potency and intrinsic activity. With this in mind, the aim of this work is to assess properties of aromatic rings, commonly used for drug design, in aqueous solution through molecular dynamics simulations in order to characterize their chemical features and infer their impact in complexation dynamics. For this, common aromatic and heteroaromatic rings were selected and received new atomic charge set based on the direction and module of the dipole moment from MP2/6-31G* calculations, while other topological terms were taken from GROMOS53A6 force field. Afterwards, liquid physicochemical properties were simulated for a calibration set composed by nearly 40 molecules and compared to their respective experimental data, in order to validate each topology. Based on the reliance of the employed strategy, we expanded the dataset to more than 100 aromatic rings. Properties in aqueous solution such as solvent accessible surface area, H-bonds availability, H-bonds residence time, and water structure around heteroatoms were calculated for each ring, creating a database of potential interactions, shedding light on features of drugs in biological solutions, on the structural basis for bioisosterism and on the enthalpic/entropic costs for ligand-receptor complexation dynamics.

Keywords: drug design, GROMOS, aromatic rings, functional groups, interactions

Ascletis Ex. 2056
Page 1 of 20

1. INTRODUCTION

The development of a drug is a multi step process, usually starting with the identification of hit compounds. The challenging task of optimizing these compounds into leads and finally into drugs is commonly facilitated by computer aided drug design (CADD) techniques (Anderson, 2003; Sliwoski et al., 2013; Bajorath, 2015). With the growing information on protein structure on the last years, structure based drug design (SBDD) has become a significant tool for hit discovery (Anderson, 2003; Lounnas et al., 2013; Lionta et al., 2014). When structural information of the receptor is absent, molecular fingerprints of approved drugs are also used to search for new ligands in a process also known as ligand based drug design (LBDD) (Lee et al., 2011). Nevertheless, there are still considerable challenges associated to the predictiveness of ligand potency and affinity via computational methods (Paul et al., 2010; Csermely et al., 2012).

In general, optimization of lead compounds is based in qualitative or quantitative structure-activity relationships (SAR or QSAR, respectively) (Shahlaei, 2013). These relationships are usually based in molecular descriptors to predict ligand pharmacodynamics and pharmacokinetics, such as log*P* to access lipophilicity, log*S* to access solubility or p*K*a to access the ionic state of a compound, along with other topological, geometrical and physicochemical descriptors (Danishuddin and Khan, 2016). While some correlations have reasonable power of predictiveness, many descriptors have no biological meaning and can mislead the optimization process. As highlighted by Hopkins et al. (2014), high-throughput screening methods have been linked to the rise of hits with inflated physicochemical properties during the optimization process (Keserü and Makara, 2009). Also, recent reviews have shown an increase of molar mass in the recent medicinal chemistry efforts (Leeson and Springthorpe, 2007) and many authors correlate this strategy with the likelihood of poor results of such compounds (Gleeson, 2008; Waring, 2009, 2010; Gleeson et al., 2011).

Many chemical moieties are regularly used in medicinal chemistry to produce chemical diversity (Bemis and Murcko, 1996; Welsch et al., 2010; Taylor et al., 2014), a practice well-known as fragment based drug design (FBDD), and its use for pharmacophore modeling and to prevent high toxicity is not recent (Gao et al., 2010). Particularly, aromatic rings are extensively used in drugs due to their well known synthetic and modification paths (Aldeghi et al., 2014). For example, at least, one aromatic ring can be found in 99% of a database containing more than 3,500 evaluated by the medicinal chemistry department of Pfizer, AstraZeneca (AZ) and GlaxoSmithKlin (GSK) (Roughley and Jordan, 2011). Still, little is known about their chemical features in biological solution, such as H-bonds availability, lifetime of H-bonds, solvent accessibility, and conformational ensemble. In this sense, molecular dynamics (MD) simulations can provide useful information with atomistic resolution and access the aforementioned features of chemical groups in water, providing fundamental data to drive medicinal chemistry approaches.

Still, dynamical properties of chemical moieties in biological solution are usually neglected in drug design and very difficult to access (Ferenczy and Keseru, 2010; Reynolds and Holloway, 2011; Hopkins et al., 2014). Even though MD simulations have been used in medicinal chemistry to generate different receptor conformers and to validate binding poses predicted by docking (Zhao and Caflisch, 2015; Ganesan et al., 2017), simulations of free ligand in solution is rarely used to access the conformational ensemble and energies associated with solvation due to the challenge on solving conformational flexibility and internal energies (Butler et al., 2009; Blundell et al., 2016). When solvated, the enthalpic and entropic costs of disrupting a H-bond or dismantling the entire solvation shell of a ligand can be the determinant step to provide the proper energy of binding (Biela et al., 2012; Blundell et al., 2013; Mondal et al., 2014). Yet, free-energy of binding is often predicted via geometrical or alchemical transformations (Zwanzig, 1954; Aqvist et al., 1994; Woo and Roux, 2005; Gumbart et al., 2013), alongside with recent developments in funnel metadynamics (Limongelli et al., 2013). More recently, thermodynamical features of ligands have been experimentally investigated in order to enhance binding and efficiency (Freire, 2009; Ferenczy and Keseru, 2010; Reynolds and Holloway, 2011). Ligand features such as H-bonds lifetime, effects of vicinity in H-bonds availability and strength, accessible surface area and water structure around binding sites can provide substantial information for designing new molecular entities (Blundell et al., 2016).

Different force fields have been used for drug design purposes, such as MMF94 (Halgren, 1996), OPLS-AA (Jorgensen et al., 1996), and GAFF (Wang et al., 2004). While these force fields parameterized their electrostatic terms using *ab initio* calculations, the GROMOS force fields (derived from the Groningen Molecular Simulation package) used free-energy of solvation as target (Daura et al., 1998; Oostenbrink et al., 2004) to empirically assign atomic partial charges. Thus, in this work, we have chosen the GROMOS force field to simulate the dynamical behavior of 103 aromatic rings (including a calibration subset of 42 molecules) mostly commonly used in drug design and their interactions with solvent in order to access thermodynamical properties in solution. These interactions, in turn, offer a reference for future rational drug design studies, as describe in details how several functional groups interact with their surroundings.

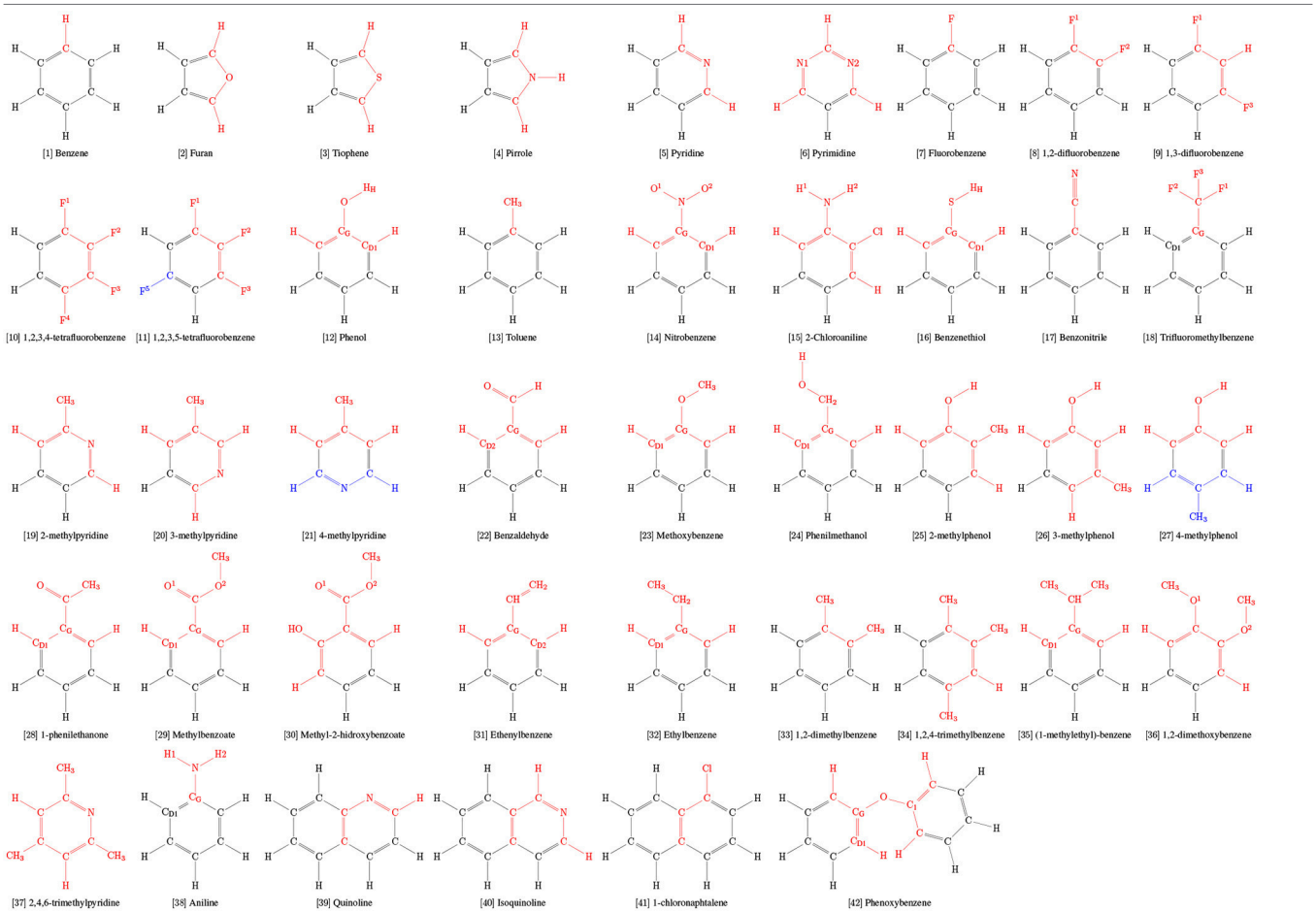
2. METHODS

2.1. Selection of Rings

A series of 103 aromatic rings commonly used in drug design were selected for this study (Broughton and Watson, 2004; Jordan and Roughley, 2009; Welsch et al., 2010; Taylor et al., 2014, 2017). Among them, a calibration set of 42 molecules (**Table 1**), for which physical-chemical properties are known, were selected from the benchmark developed by Caleman et al. (2012). Briefly, both works of Taylor et al. (2014, 2017) employed a detailed search of substructure frequencies from FDA Orange Book and cross referenced with ChEMBL, DrugBank, Nature, Drug Reviews, the FDA Web site, and the Annual Reports in Medicinal Chemistry; the work of Broughton and Watson (2004) employed search of substructure frequencies in MDL Drug Data

Ascletis Ex. 2056
Page 2 of 20

TABLE 1 | Charge groups (colored) and aromatic rings used as calibration set in this work.



Report database by using a "Phase II" keyword; and the work of Welsch et al. (2010) have pinpointed privileged scaffolds from natural-products works throughout literature.

2.2. Topology Construction

Structures for these aromatic rings were built using Avogadro (Hanwell et al., 2012). Molecular mechanical (MM) topological parameters as bonds, angles, and Lennard-Jones parameters were taken from GROMOS53A6 (Oostenbrink et al., 2004). Due to the well–known good performance of MP2 methods for small aromatic rings (Li et al., 2015; Matczak and Wojtulewski, 2015), atomic partial charges were based on quantum mechanical (QM) calculations using MP2 theory (Møller and Plesset, 1934), 6-31G* (Petersson et al., 1988) basis set and implicit solvent *Polarizable Continuum Model* (PCM) (Mennucci and Tomasi, 1997) followed by a RESP fitting (Bayly et al., 1993). The so obtained partial charges were adjusted in the MM to reproduce the QM dipole moment of the ring. The angle θ formed between the QM and MM model dipole moment vectors was monitored through an in house script to make sure the angle had the lowest value possible, guaranteeing the conservation of the QM dipole moment direction. For our calibration set, the module of the MM partial charges were adjusted to better reproduce the physicochemical properties of the organic liquids. Following the philosophy of charge group assignment, groups were limited, at maximum, to the atoms at the *ortho* position on each ring. In more complex substitution patterns, a superimposition of two charge groups was required to correctly describe the chemical group. In such cases, the Coulombic terms of the overlapping atoms were adjusted to correctly describe the direction of the total dipole moment of the ring. For molecules containing linear constraints (benzonitrile), *virtual sites* were added in order to preserve the total moment of inertia and mass, thus preserving the linearity of these groups (Feenstra et al., 1999).

2.3. New Torsional Potentials

The quantum mechanical torsional profile of every dihedral angle was calculated using Gaussian (Frisch et al., 2016) (RRID:SCR_014897). Molecular structures were built using Avogadro (Hanwell et al., 2012) and their geometry were optimized using Hartree-Fock method (Fock, 1930; Hartree and Hartree, 1935) and basis set 3-21G* (Dobbs and Hehre, 1986). Afterwards, the *Scan* routine was used to calculate the total energy of the molecule conformation for each dihedral

Ascletis Ex. 2056
Page 3 of 20

orientation, adopting a *tight* convergence criteria, with geometric optimization, MP2/6-31G* and steps of 30°. In order to calculate the torsional profile for molecular mechanics model, dihedral orientations were kept fixed during minimization using restraint forces for the same angles evaluated by quantum calculations. Both profiles were submitted to the Rotational Profiler server (Rusu et al., 2014) to obtain appropriate sets of classical mechanics parameters that provided a better fitting to the QM-obtained torsional profile.

2.4. General Simulation Settings

All simulations were carried out using the GROMACS 5.0.7 package (Abraham et al., 2015) (RRID:SCR_014565). In order to create parameters compatible with the GROMOS family, we have followed previous literature (Daura et al., 1998; Schuler et al., 2001; Oostenbrink et al., 2004) settings: twin-range scheme was used with short- and long-range cutoff distances of 0.8 and 1.4 nm, respectively. Also, the reaction-field method was applied to correct the effects of electrostatic interactions beyond the long-range cutoff distance (Barker and Watts, 1973; Tironi et al., 1995), using the dielectric constant as ε_{RF} for organic liquid simulations and $\varepsilon_{RF} = 62$ for simulations in water (Heinz et al., 2001; Oostenbrink et al., 2004). The LINCS algorithm (Hess et al., 1997; Hess, 2008) was used to constrain all covalent bonds, using a cubic interpolation, a Fourier grid of 0.12 nm and timestep of 2 fs. Configurations were saved at every 2 ps for analysis.

2.4.1. Organic Liquids Simulations

In order to build the organic liquid systems, cubic boxes of 2×2×2 nm were created, each with a single organic molecule. A total of 125 of these boxes were stacked, forming an unique box with conventional periodic boundary conditions treatment of 10×10×10 nm which was simulated under high pressure (100 bar) to induce liquid phase. The systems were then simulated and equilibrated at 1 bar. Afterwards, the boxes were staggered to obtain systems with 1000 molecules in liquid phase and simulated at 1 bar until the total energy drift converged to values below 0.5 J/(mol×ns×Degrees of Freedom). Such criterion is necessary to make sure that the fluctuating properties could be accurately calculated (Caleman et al., 2012). All simulations were carried out with Berendsen pressure and temperature coupling algorithm due to their efficiency in molecular relaxations (Berendsen et al., 1984), using $\tau_T = 0.2$ ps and $\tau_P = 0.5$ ps. When available, experimental values of isothermal compressibility and dielectric constant were used as an additional parameter for liquid simulations. Otherwise, the compressibility of the most chemically similar molecule was used. The experimental dielectric constants from each liquid were also used as parameters in the simulations (Oostenbrink et al., 2004).

In order to calculate the densities of liquids (ρ), simulations at constant pressure were carried out for 10 ns and ρ were calculated using block averages of 5 blocks. Enthalpy of vaporization (ΔH_{vap}) were calculated by block averaging the same 10 ns of liquid simulation to obtain $E_{pot}(l)$ and another 100 ns of gas phase simulation using a stochastic dynamics integrator (SD) (Van Gunsteren and Berendsen, 1988) with a single molecule in vacuum, to obtain $E_{pot}(g)$ as the equation:

$$\Delta H_{vap} = (E_{pot}(g) + k_B T) - E_{pot}(l) \tag{1}$$

Aiming to calculate the dielectric constant (ε), the simulation of the liquid boxes from which ρ were obtained were extended up to 60 ns. Convergence calculations of ε were done using running averages and ε were evaluated only after convergence. In order to calculate thermal expansion coefficients (α_P) and classic isobaric heat capacities ($C_{P\text{cla}}$), three constant pressure simulations were carried out for 5 ns each, with temperatures T, T+10K, and T-10K, for each liquid. The calculations of α_P and $C_{P\text{cla}}$ were done using the finite difference method (Kunz and van Gunsteren, 2009):

$$\alpha_P \approx \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P \approx -\frac{\ln\langle\rho\rangle_{T_2} - \ln\langle\rho\rangle_{T_1}}{T_2 - T_1} \tag{2}$$

and:

$$C_P \approx \left(\frac{\partial U}{\partial T}\right)_P \approx \frac{\langle U\rangle_{T_2} - \langle U\rangle_{T_1}}{T_2 - T_1} \tag{3}$$

In order to calculate isothermal compressibilities (κ_T), three constant volume simulations were carried out for 5 ns each, with pressures 1, 0.9, and 1.1 bar. The calculations of κ_T was also done using the finite difference method:

$$\kappa_T \approx \frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \approx -\frac{\ln\rho_2 - \ln\rho_1}{\langle P\rangle_{\rho_2} - \langle P\rangle_{\rho_1}} \tag{4}$$

2.4.2. Solvation Free Energy Simulations

Simulations in water were carried out to evaluate the solvation free energies (ΔG_{hyd}) of 30 molecules at 1 bar and 298 K. Each aromatic ring (solute) was centered into a cubic box with appropriate dimensions to reproduce the density of SPC water models (0.997 g/cm^3). In free-energy calculations using thermodynamic integration (TI) method, a coupling parameter λ is used to perturb solute-solvent interactions.

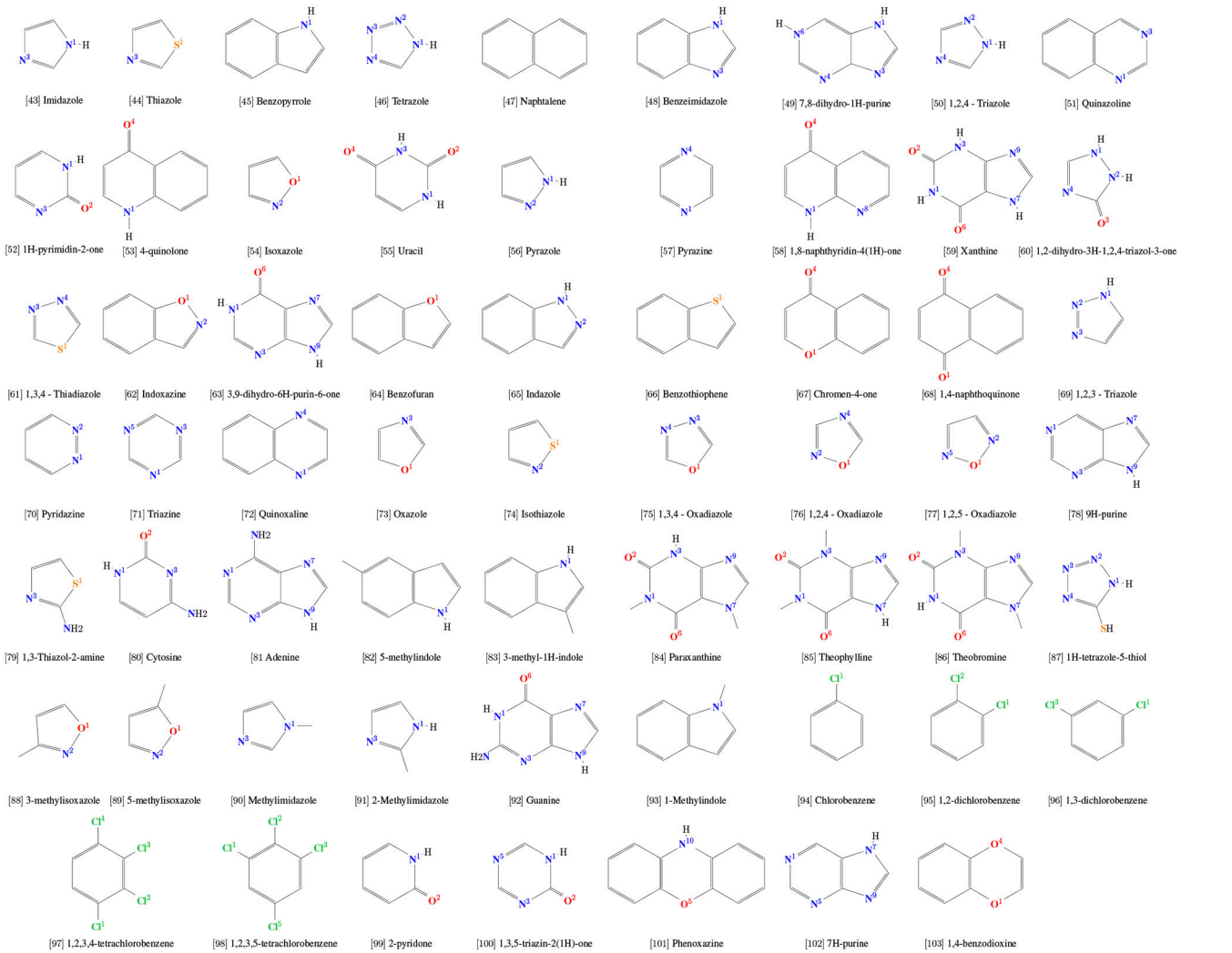
$$\Delta G_{\text{sim}} = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda \tag{5}$$

in which H is the Halmiltonian, $\lambda = 0$ refers to the state in which the solute fully interacts with the solvent and $\lambda = 1$ refers to the state in which the solute-solvent interactions do not exist. In our setup, Coulombic interactions were decoupled first, and the Lennard-Jones interactions after, using a soft-core potential to avoid issues related to strong Lennard-Jones interactions (Beutler et al., 1994). A soft-core power was set to 1 and α_{LJ} set to 0.5, following recommendations of Shirts and Pande (2005). Both interactions were decoupled using λ values: 0, 0.02, 0.04, 0.07, 0.1, 0.15, 0.2, ..., 0.8, 0.85, 0.9, 0.93, 0.96, 0.98, 1, totalizing 50 λ simulations.

Our simulation protocol consisted of an initial steepest-descent minimization, followed by a L-BFGS minimization until a maximum force of 10 kJ/(mol^{-1} nm^{-1}) was reached.

Ascletis Ex. 2056
Page 4 of 20

TABLE 2 | Dataset of aromatic rings evaluated in this work. Heteroatoms are highlighted in colors.



After, initial velocities were assigned and the systems were equilibrated for 100 ps using a NVT ensemble at each λ. The systems were subjected to another 100 ps of equilibration on a NPT ensemble, using the Parrinello-Rahman pressure coupling algorithm (Parrinello and Rahman, 1981), a $\tau_t = 5$ ps time constant for coupling and a compressibility of 4.5×10^5 bar^{-1}. Finally, production simulations were done using the Langevin integrator (Van Gunsteren and Berendsen, 1988) to sample the $\langle \partial H/\partial \lambda \rangle_\lambda$ until convergence. Therefore, simulations time varied between 1 and 5 ns. In addition, the last frame of the production phase of each λ was used as input for the next subsequent λ.

2.4.3. Simulation of Rings in Water

After an extensive comparison of simulated and experimental physicochemical properties of our calibration set and consequent validation, the same strategy of topological construction was applied to other 61 rings commonly used in drug design (**Table 2**) for which experimental properties are not available, totalizing 103 aromatic rings in this study. Hence, in order to evaluate chemical features and interactions of aromatic rings with their surroundings, a total set of 103 aromatic was simulated in water, including all 42 molecules present in the calibration set (**Table 1**). Each solute was placed in a cubic box with a distance of 1.0 nm to its edges. The boxes were then filled with SPC water model and minimized long enough eliminate any possible clashes until convergence at a maximum force of 0.1 kJ/mol×nm. After, the system was equilibrated in a NVT ensemble at 298.15 K using the Nosé-Hoover algorithm (Nosé, 1984) for temperature coupling. Production runs of 250 ns were carried out with temperature and pressure coupling handled by V-rescale (Bussi et al., 2007) and Parrinelo-Rahman (Parrinello and Rahman, 1981) algorithms, using $\tau_T = 0.1$ ps and $\tau_P = 2.0$ ps. The GROMACS tools *hbond*, *rdf*, and *sorient* were used to calculate H-bonds related properties and solvation structure around the heteroatom using a block-averaging approach over 5 box of 50 ns.

Ascletis Ex. 2056
Page 5 of 20

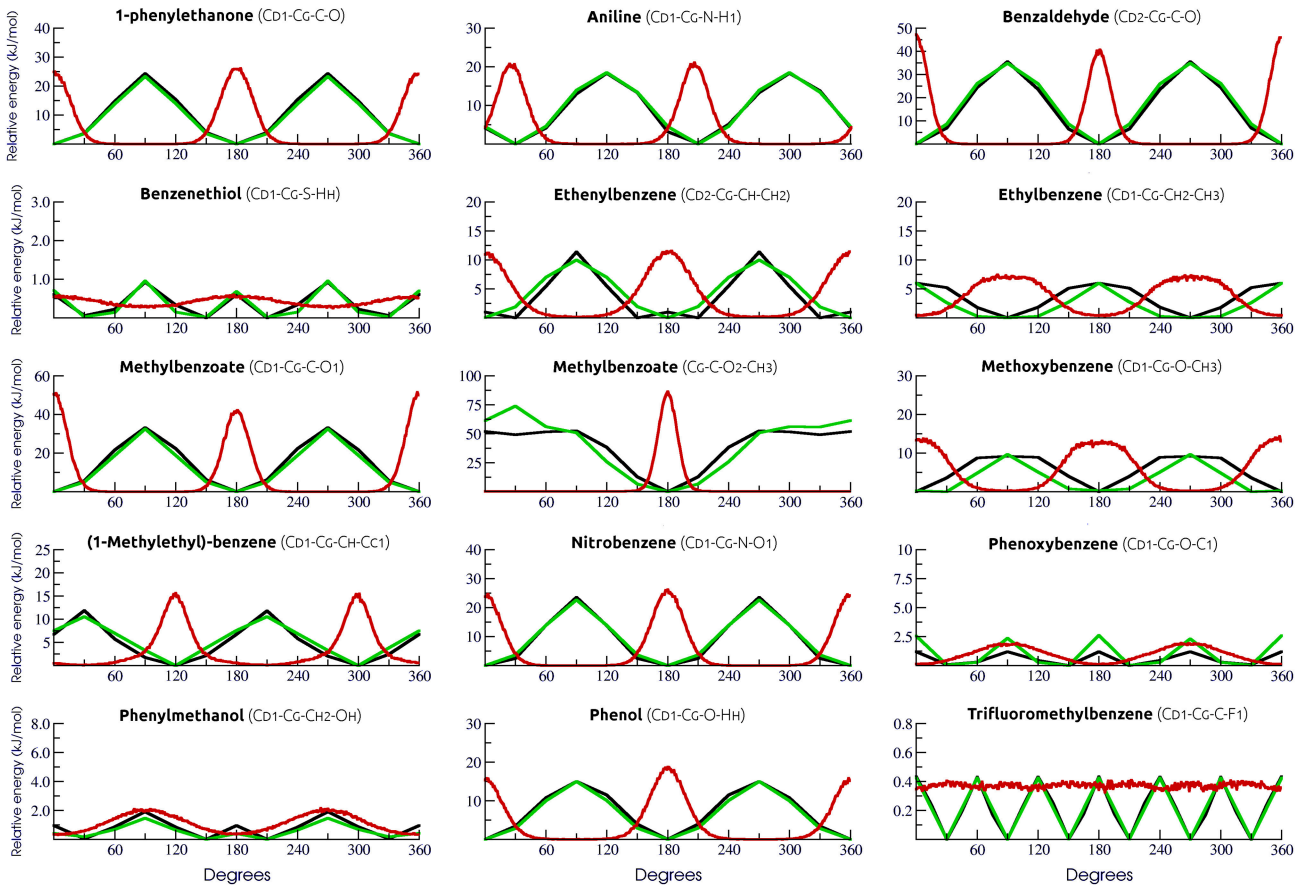


FIGURE 1 | Evaluation of torsional parameters and dihedral distribution. QM and adjusted MM torsional profiles are shown in black and green, respectively. In red, the dihedral distribution during simulations.

3. RESULTS

3.1. New Torsional Profiles

In order to accurately describe the torsional angles of the selected aromatic rings, a total of 15 new dihedral potentials were derived by fitting the MM profiles to the corresponding QM-calculated ones (Table S1). Fittings were conducted using the Rotational Profiler server (Rusu et al., 2014). For all cases, the use of new parameters yield almost identical values of minimum and barrier amplitudes to those calculated by QM (**Figure 1**). Dihedral distribution throughout simulations was also evaluated.

3.2. Physical-Chemical Properties

In order to validate our strategy of topology building, boxes of organic liquids were simulated to obtain physical-chemical properties for each compound. Reference experimental values (Table S2) were used to calculate the absolute error of each property and to guide adjustments on the coulombic terms in order to mitigate deviations. We have calculated the θ angle between QM and MM dipole moments and the final version of our calibration set (**Table 1**) yielded an average θ angle of $2.5^\circ \pm 6.1^\circ$, suggesting that our MM models conserve the direction of the QM dipole moment, preserving the electrostatic potential of each molecule.

Following the GROMOS philosophy (Oostenbrink et al., 2004; Horta et al., 2016), density (ρ), enthalpy of vaporization (ΔH_{vap}), and free energy of solvation (ΔG_{hyd}) were used as targets for the parametrization, while isothermal coefficient (α_P), isothermal compressibility (κ_T), dielectric constant (ε), and classic isobaric heat capacity ($C_{P\text{cla}}$) were calculated as benchmarks for GROMOS performance and compared with the results obtained in Caleman et al. (2012) and Horta et al. (2016) (**Table 3**). Linear regression between experimental and simulated values were calculated in order to access the prediction power of the employed strategy (**Figure 2**). The equations further reported were calculated excluding outliers (values higher than 2 standard deviations).

Regarding the targeted properties, our calibration set yielded the equations $y = 0.9118x + 0.1001$ for density, $y = 1.0699x - 1.6491$ for enthalpy of vaporization and $y = 0.8676x + 0.8929$ for free energy of solvation, with correlation coefficients of $R = 0.92$, $R = 0.96$, and $R = 0.89$, respectively. In terms of average deviation (AVED), our calibration set overestimates ρ in 0.008 g/cm^3, ΔH_{vap} in 1.51 kJ/mol and underestimates ΔG_{hyd} in 3.35

Ascletis Ex. 2056
Page 6 of 20

TABLE 3 | Average deviation between experimental and simulated physicochemical properties of aromatic rings evaluated in our calibration set. Simulated GAFF and OPLS-AA values were obtained from Caleman et al. (2012) and 2016H66 values from Horta et al. (2016). Density (ρ) in g/cm^3, enthalpy of vaporization (ΔH_{vap}) in kJ/mol, thermal expansion coefficient (α_P) in 10^{-3}/K, isothermal compressibility (κ_T) in 1/GPa, dielectric constant (ε), classic isobaric heat capacity (Cp_{cla}) in J/mol×K, and free-energy of solvation (ΔG_{hyd}) in kJ/mol.

Properties	Force field	Statistical N	Average Dev.	St. Dev.	R coefficient
ρ	This work	42	0.008	0.051	0.92
	2016H66	6	0.016	0.019	0.99
	GAFF	40	−0.008	0.045	0.93
	OPLS-AA	40	0.001	0.025	0.98
ΔH_{vap}	This work	42	1.514	4.457	0.96
	2016H66	6	2.257	6.758	0.96
	GAFF	40	2.298	5.419	0.88
	OPLS-AA	40	3.243	5.216	0.90
Cp_{cla}	This work	42	88.201	33.440	0.77
	2016H66	6	98.712	35.232	0.63
	GAFF	37	133.884	40.225	0.84
	OPLS-AA	37	129.397	35.330	0.91
α_P	This work	42	0.146	0.210	0.82
	2016H66	6	0.171	0.148	0.91
	GAFF	40	0.224	0.220	0.58
	OPLS-AA	40	0.155	0.210	0.64
κ_T	This work	42	0.046	0.500	0.70
	2016H66	6	0.276	0.279	0.71
	GAFF	40	0.054	0.150	0.77
	OPLS-AA	40	−0.016	0.130	0.78
ε	This work	42	−4.523	5.650	0.65
	2016H66	6	−2.217	2.515	0.89
	GAFF	29	−4.254	2.740	0.97
	OPLS-AA	33	−4.564	5.600	0.72

kJ/mol. Without the outliers, the AVED for ΔG_{hyd} improves to 2.83 kJ/mol.

Non-targeted properties were calculated to evaluate how they behaved in our simulations. Linear regressions yielded equations $y = 0.93825x + 0.1406$ for α_P (R = 0.82), $y = +0.90079x - 0.0140$ for κ_T ($R = 0.70$), $y = 0.2581x + 1.8961$ for ε ($R = 0.65$), and $y = 0.8989x + 100.5$ for Cp_{cla} ($R = 0.77$). In terms of AVED, α_P is overestimated in 0.14 10^{-3}/K and κ_T is overestimated in 0.0465 1/GPa. As expected (Caleman et al., 2012; Horta et al., 2016), ε is poorly described due to the lack of polarization effects, resulting in a underestimation of −4.52 in the dielectric constant. On other hand, Cp_{cla} was overestimated by 88.2 J/mol×K, a behavior aligned with recent works in literature (Caleman et al., 2012; Horta et al., 2016). Individual AVED and absolute errors can be found in Tables S4, S5 in Supplementary Material, along with experimental properties in Table S3.

3.3. Interactions in Water

In order to quantitatively evaluate the behavior of heteroaromatic rings in water and their interactions with the aqueous surrounding, some properties were calculated throughout 250 ns of simulation. From these calculations, we were capable to assess the average H-bond ($Aver_{HB}$) of each heteroatom along with its residence time (τ_{HB}), lifetime ($lifetime_{HB}$), the free-energy of breakage of a H-bond (ΔG_{HB}), and the percentage of simulation time that a given heteroatom was involved in, at least, one H-bond (*Percent*). We were also capable to obtain the optimal binding distance between an heteratom and water (OBD_{HB}), along with the coordination number (CN_{HB}) at the OBD_{HB} and the average orientation of water molecules surrounding the heteroatom. These data are compiled in **Tables 4**, **5**.

4. DISCUSSION

4.1. Topology Building Strategy

The accurate description of organic compounds' chemical diversity, mainly in the context of drugs and medicinal chemistry, is a challenging task in molecular mechanics since it must be described as broadly as possible by the force field fragments. However, the most common sets of MM parameters

Ascletis Ex. 2056
Page 7 of 20

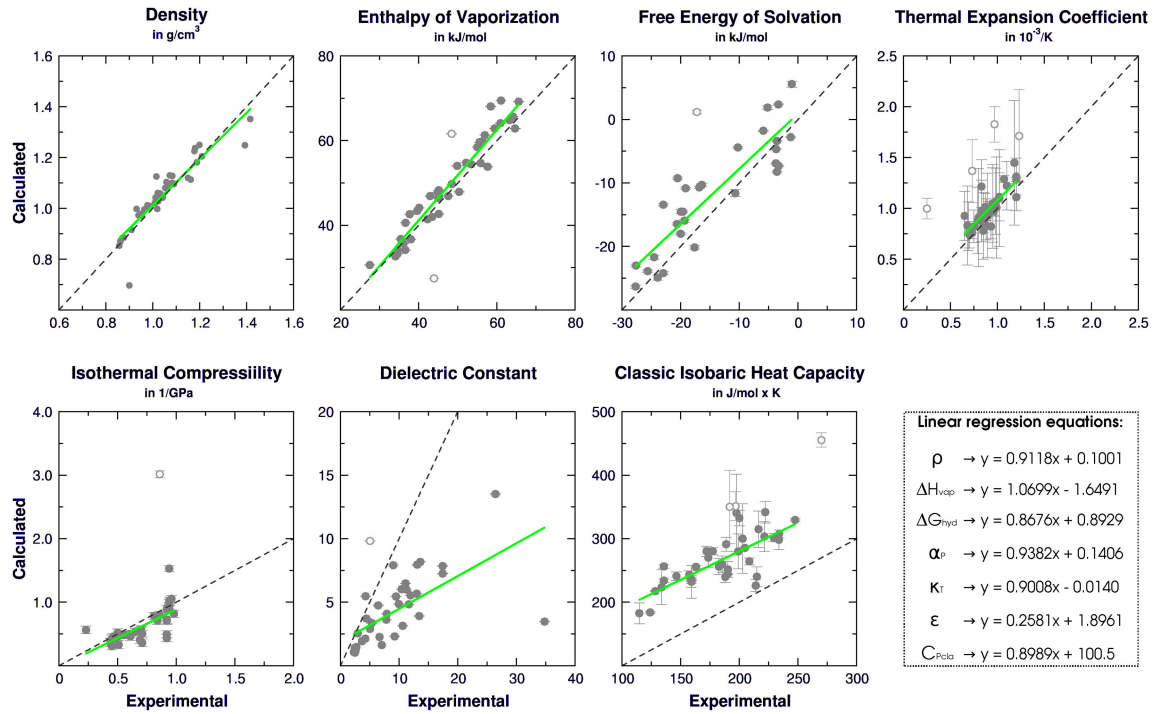


FIGURE 2 | Correlation between experimental and calculated physical-chemical properties of organic liquids for 42 aromatic compounds on the calibration set. Standard deviations are shown as bars, linear regressions are shown as green and empty dots represent outliers.

employed in biomolecules simulations are usually centered on the monomeric constituents of biopolymers and lipids, while parameters for synthetic compounds, as well as other common non-polymeric biological molecules (e.g., natural products), must be included from specific calculations or external sets of parameters.

In this sense, a proper description of torsional terms will impact directly the dynamical behavior of these small molecules, even considering that, when evaluating ligand-receptor complexes, the influence of these terms might be mitigated due to the ligand movement restriction inside the binding pocket. Still, accommodation of flexible docking derived poses, fine tunning of induced fit, and characterization of ligands conformational induction vs. selection (with potential inferences of the entropic costs of binding) require dihedrals potentials specifically adjusted to organic compounds. Hence, new parameters were generated in this work exclusively for 15 dihedrals in aromatic rings in our calibration set (**Figure 1**). In general, our results revealed that our MM parameters yielded a good description of the QM torsional profile, with the exceptions of [16] tiophenol, [42] phenoxybenzene, [24] phenylmethanol, and [18] trifluoromethylbenzene. For these molecules, the distribution profile was almost evenly spread, most likely due to the low energy barrier (below 2.5 kJ/mol), indicating that transient states are commonly achieved during our simulations in SPC water model. Simulations of these particular molecules in vacuum revealed little influence of water solvation in the dihedral profile (data not shown).

In another sense, the choice of an atomic charge set for ligands can drastically impact thermodynamical binding properties such as complexation free-energy and desolvation. Therefore, we employed in this work a dipole moment based strategy to describe the Coulombic contribution using physicochemical properties of organic liquids as target. The prediction power of our strategy was compared to recent comparisons of aromatic compounds in liquid phase (Caleman et al., 2012; Horta et al., 2016) and summarized in **Table 3**. In general, our calibration set yielded similar or lower average deviations than benchmarks made with OPLS-AA, GAFF, and 2016H66 sets for all physicochemical properties evaluated in this work. The main difference was in terms of Cp_{cla}, for which GAFF and OPLS-AA overestimate nearly 40 J/mol×K more than our parameters. Still, all four parameters sets overestimates Cp_{cla}. In addition, the GROMOS53A5 force field was designed to reproduce physicochemical properties, and later on adjusted to reproduce free energy of solvation and hydration (GROMOS53A6) (Oostenbrink et al., 2004). The average deviation on density, enthalpy of vaporization and free-energy of solvation of GROMOS53A5 were 0.0389 g/cm^3, −0.4 and 3.8 kJ/mol, respectively. These values are very similar to our results, as shown in **Table 3**, reiterating the quality of our parameters.

It is important to mention that the employed benchmark set was built using the same Lennard-Jones parameters used in the benzene ring of phenylalanine in GROMOS53A6. While GROMOS53A6 produces a $\Delta G_{hyd} = 0.0$ kJ/mol for benzene

Ascletis Ex. 2056
Page 8 of 20

TABLE 4 | Properties of heteroaromatic rings in water. Average H-bonds ($Aver_{HB}$), H-bond residence time (τ_{HB}) is ps, H-bond lifetime ($lifetime_{HB}$) in 1/ps, free-energy of H-bond breakage (ΔG_{HB}) in kJ/mol, percentage of simulation with at least one formed H-bond (*Percent*.), coordination number of water (*CN*), optimal binding distance with water (OBD_{HB}) in nm, and overall water orientation around the heteroatom (*Orientation*).

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	*Percent*	CN	OBD_{HB}	*Orientation*
Water	Ow	1.73 ± 0.62	2.11 ± 0.02	0.47 ± 0.00	6.38 ± 0.03	98.58	4.11 ± 2.83	0.18 ± 0.00	Undefined
	OH^1	0.87 ± 0.35	1.80 ± 0.03	0.55 ± 0.01	5.98 ± 0.05	86.25	4.11 ± 2.83	0.18 ± 0.00	O-oriented
	OH^2	0.86 ± 0.35	1.83 ± 0.03	0.54 ± 0.01	6.03 ± 0.04	86.07	4.11 ± 2.83	0.18 ± 0.00	O-oriented
Phenol	O	1.10 ± 0.62	1.61 ± 0.03	0.62 ± 0.01	5.70 ± 0.04	85.96	1.46 ± 1.03	0.18 ± 0.00	Undefined
	OH	0.96 ± 0.20	9.49 ± 0.18	0.11 ± 0.00	10.11 ± 0.05	96.04	0.90 ± 0.01	0.17 ± 0.00	O-oriented
Phenylmethanol	O	1.42 ± 0.58	2.58 ± 0.03	0.39 ± 0.00	6.88 ± 0.02	96.51	2.68 ± 1.59	0.18 ± 0.00	Undefined
	OH	0.95 ± 0.24	5.37 ± 0.06	0.19 ± 0.00	8.70 ± 0.03	94.25	1.13 ± 0.01	0.17 ± 0.00	O-oriented
2-methylphenol	O	1.04 ± 0.59	1.88 ± 0.04	0.53 ± 0.01	6.09 ± 0.05	84.80	1.05 ± 0.00	0.18 ± 0.00	Undefined
	OH	0.95 ± 0.23	9.46 ± 0.17	0.11 ± 0.00	10.10 ± 0.04	94.53	0.87 ± 0.01	0.17 ± 0.00	O-oriented
3-methylphenol	O	1.08 ± 0.61	1.74 ± 0.02	0.58 ± 0.01	5.90 ± 0.03	85.83	1.43 ± 1.00	0.18 ± 0.00	Undefined
	OH	0.96 ± 0.19	10.12 ± 0.19	0.10 ± 0.00	10.27 ± 0.05	96.30	0.90 ± 0.01	0.17 ± 0.00	O-oriented
4-methylphenol	O	1.08 ± 0.61	1.73 ± 0.02	0.58 ± 0.01	5.89 ± 0.03	85.70	1.10 ± 0.01	0.18 ± 0.00	Undefined
	OH	0.96 ± 0.20	10.00 ± 0.21	0.10 ± 0.00	10.24 ± 0.05	96.21	0.90 ± 0.01	0.17 ± 0.00	O-oriented
Benzenethiol	S	0.67 ± 0.65	0.38 ± 0.01	2.63 ± 0.05	2.13 ± 0.04	57.29	0.81 ± 0.17	0.23 ± 0.00	Undefined
	SH	0.77 ± 0.43	1.00 ± 0.02	1.00 ± 0.02	4.52 ± 0.05	76.38	2.08 ± 0.02	0.23 ± 0.00	O-oriented
Aniline	N	0.93 ± 0.58	1.64 ± 0.02	0.61 ± 0.01	5.75 ± 0.03	79.89	1.01 ± 0.01	0.19 ± 0.00	Undefined
	NH^1	0.63 ± 0.49	1.15 ± 0.03	0.87 ± 0.02	4.87 ± 0.06	62.48	1.25 ± 0.38	0.22 ± 0.00	O-oriented
	NH^2	0.63 ± 0.50	0.99 ± 0.02	1.01 ± 0.02	4.51 ± 0.04	62.05	1.39 ± 0.25	0.23 ± 0.00	O-oriented
2-chloroaniline	N	0.86 ± 0.50	2.29 ± 0.04	0.44 ± 0.01	6.59 ± 0.05	79.39	0.92 ± 0.00	0.19 ± 0.00	Undefined
	NH^1	0.51 ± 0.51	1.00 ± 0.03	1.00 ± 0.03	4.53 ± 0.06	50.60	1.33 ± 0.14	0.23 ± 0.00	O-oriented
	NH^2	0.56 ± 0.51	0.87 ± 0.03	1.15 ± 0.04	4.18 ± 0.09	55.20	3.82 ± 5.29	0.23 ± 0.01	O-oriented
	Cl	0.24 ± 0.45	0.32 ± 0.08	3.26 ± 0.68	1.66 ± 0.56	22.67	18.94 ± 10.05	0.36 ± 0.00	Undefined
Pyridine	N	1.41 ± 0.71	1.33 ± 0.02	0.75 ± 0.01	5.24 ± 0.03	91.46	1.59 ± 0.01	0.20 ± 0.00	Undefined
Pyrimidine	N^1	1.06 ± 0.68	0.91 ± 0.02	1.10 ± 0.02	4.30 ± 0.05	80.71	1.23 ± 0.01	0.20 ± 0.00	Undefined
	N^2	0.98 ± 0.68	0.81 ± 0.02	1.24 ± 0.03	4.00 ± 0.06	76.96	1.17 ± 0.01	0.20 ± 0.00	Undefined
2-methylpyridine	N	1.52 ± 0.70	1.74 ± 0.04	0.57 ± 0.01	5.90 ± 0.06	93.96	1.68 ± 0.00	0.20 ± 0.00	Undefined
3-methylpyridine	N	1.43 ± 0.71	1.34 ± 0.04	0.74 ± 0.02	5.26 ± 0.07	91.65	1.61 ± 0.01	0.20 ± 0.00	Undefined
4-methylpyridine	N	1.46 ± 0.71	1.44 ± 0.04	0.69 ± 0.02	5.44 ± 0.07	92.57	1.62 ± 0.01	0.20 ± 0.00	Undefined
2,4,6-trimethylpyridine	N	0.36 ± 0.53	0.36 ± 0.04	2.79 ± 0.31	2.00 ± 0.28	33.67	24.48 ± 3.47	0.42 ± 0.09	Undefined
Quinoline	N	1.64 ± 0.68	2.00 ± 0.05	0.50 ± 0.01	6.25 ± 0.06	96.10	1.78 ± 0.01	0.19 ± 0.00	Undefined
Isoquinoline	N	1.26 ± 0.68	1.22 ± 0.04	0.82 ± 0.02	5.02 ± 0.07	88.67	1.43 ± 0.01	0.20 ± 0.00	Undefined
Benzonitrile	N	1.63 ± 0.72	1.30 ± 0.01	0.77 ± 0.01	5.17 ± 0.02	95.50	1.88 ± 0.01	0.19 ± 0.00	Undefined
Furan	O	0.42 ± 0.57	0.29 ± 0.01	3.41 ± 0.07	1.49 ± 0.05	37.99	31.54 ± 2.60	0.46 ± 0.01	Undefined
Tiophene	S	0.15 ± 0.37	0.25 ± 0.03	4.07 ± 0.49	1.07 ± 0.32	14.03	18.25 ± 5.37	0.37 ± 0.00	Undefined
Pyrrole	NH	0.92 ± 0.29	3.80 ± 0.06	0.26 ± 0.00	7.84 ± 0.04	91.73	0.38 ± 0.00	0.18 ± 0.00	O-oriented
	N	0.74 ± 0.67	1.33 ± 0.03	0.75 ± 0.02	5.23 ± 0.06	60.90	0.62 ± 0.01	0.23 ± 0.00	Undefined
Fluorobenzene	F^1	0.30 ± 0.49	0.30 ± 0.03	3.35 ± 0.32	1.54 ± 0.24	27.84	13.84 ± 5.27	0.36 ± 0.01	Undefined

(Continued)

Ascletis Ex. 2056
Page 9 of 20

TABLE 4 | Continued

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	Percent	CN	OBD_{HB}	Orientation
1,2-difluorobenzene	F^1	0.24 ± 0.45	0.33 ± 0.07	3.14 ± 0.61	1.74 ± 0.49	22.91	12.15 ± 2.45	0.37 ± 0.01	Undefined
	F^2	0.24 ± 0.45	0.34 ± 0.07	3.08 ± 0.64	1.79 ± 0.51	22.90	13.31 ± 3.02	0.37 ± 0.01	Undefined
1,3-difluorobenzene	F^1	0.23 ± 0.45	0.36 ± 0.10	2.91 ± 0.58	1.94 ± 0.59	22.23	14.99 ± 5.60	0.36 ± 0.01	Undefined
	F^3	0.23 ± 0.45	0.32 ± 0.04	3.20 ± 0.36	1.66 ± 0.29	22.22	11.70 ± 1.33	0.36 ± 0.00	Undefined
1,2,3,4-tetrafluorobenzene	F^1	0.17 ± 0.38	0.36 ± 0.08	2.88 ± 0.61	1.97 ± 0.53	16.08	12.99 ± 3.94	0.37 ± 0.01	Undefined
	F^2	0.18 ± 0.40	0.42 ± 0.15	2.66 ± 0.78	2.22 ± 0.79	17.44	12.72 ± 2.45	0.37 ± 0.01	Undefined
	F^3	0.18 ± 0.40	0.33 ± 0.04	3.11 ± 0.45	1.74 ± 0.34	17.31	13.89 ± 3.36	0.37 ± 0.01	Undefined
	F^4	0.16 ± 0.38	0.43 ± 0.21	2.72 ± 0.83	2.20 ± 0.98	16.05	11.25 ± 1.36	0.36 ± 0.00	Undefined
1,2,3,5-tetrafluorobenzene	F^1	0.17 ± 0.39	0.37 ± 0.10	2.84 ± 0.64	2.01 ± 0.61	16.51	12.34 ± 3.05	0.36 ± 0.01	Undefined
	F^2	0.16 ± 0.37	0.48 ± 0.25	2.47 ± 0.82	2.47 ± 1.03	15.19	12.79 ± 3.60	0.37 ± 0.01	Undefined
	F^3	0.17 ± 0.39	0.44 ± 0.10	2.40 ± 0.53	2.42 ± 0.56	16.59	14.91 ± 5.13	0.37 ± 0.01	Undefined
	F^5	0.21 ± 0.43	0.33 ± 0.06	3.17 ± 0.54	1.70 ± 0.44	20.59	11.51 ± 1.36	0.36 ± 0.01	Undefined
Trifluoromethylbenzene	F^1	0.10 ± 0.30	1.64 ± 1.99	1.26 ± 0.91	4.66 ± 2.10	9.56	14.84 ± 3.37	0.39 ± 0.02	Undefined
	F^2	0.10 ± 0.30	0.64 ± 0.37	1.95 ± 0.73	3.10 ± 1.16	9.66	15.37 ± 3.44	0.40 ± 0.02	Undefined
	F^3	0.10 ± 0.30	2.82 ± 5.74	1.05 ± 0.41	5.00 ± 2.38	9.54	14.94 ± 3.23	0.40 ± 0.02	Undefined
1-chloronaphthalene	Cl	0.37 ± 0.55	0.28 ± 0.02	3.55 ± 0.21	1.39 ± 0.15	33.96	15.53 ± 7.71	0.36 ± 0.00	Undefined
1-phenylethanone	O	0.87 ± 0.66	0.56 ± 0.01	1.79 ± 0.04	3.09 ± 0.05	71.14	1.08 ± 0.01	0.19 ± 0.00	Undefined
Benzaldehyde	O	1.03 ± 0.66	0.78 ± 0.01	1.28 ± 0.02	3.91 ± 0.04	80.79	1.22 ± 0.01	0.18 ± 0.00	Undefined
Nitrobenzene	O^1	0.14 ± 0.36	0.43 ± 0.13	2.51 ± 0.67	2.34 ± 0.72	13.79	13.60 ± 4.08	0.38 ± 0.01	Undefined
	O^2	0.14 ± 0.36	0.46 ± 0.15	2.37 ± 0.62	2.48 ± 0.73	13.81	16.48 ± 5.81	0.38 ± 0.01	Undefined
Methylbenzoate	O^1	0.95 ± 0.67	0.69 ± 0.01	1.45 ± 0.02	3.62 ± 0.04	75.85	1.13 ± 0.01	0.19 ± 0.00	Undefined
	O^2	0.17 ± 0.38	0.28 ± 0.05	3.63 ± 0.58	1.37 ± 0.43	16.40	25.27 ± 2.00	0.45 ± 0.00	Undefined
2-hydroxy-methylbenzoate	O	0.96 ± 0.58	1.48 ± 0.02	0.67 ± 0.01	5.51 ± 0.03	81.17	1.07 ± 0.00	0.18 ± 0.00	Undefined
	O^1	0.94 ± 0.64	1.07 ± 0.17	0.96 ± 0.15	4.65 ± 0.39	76.43	1.07 ± 0.13	0.18 ± 0.00	Undefined
	O^2	0.12 ± 0.33	0.36 ± 0.23	3.43 ± 1.03	1.65 ± 1.10	12.11	23.62 ± 8.28	0.33 ± 0.11	Undefined
	OH	0.05 ± 0.22	0.21 ± 0.04	4.83 ± 0.80	0.66 ± 0.46	5.25	0.40 ± 0.01	0.18 ± 0.01	O-oriented
Methoxybenzene	O	0.36 ± 0.51	0.33 ± 0.02	3.03 ± 0.14	1.78 ± 0.12	34.78	0.42 ± 0.01	0.20 ± 0.00	H-oriented
1,2-dimethoxybenzene	O^1	0.39 ± 0.54	0.38 ± 0.02	2.62 ± 0.15	2.14 ± 0.14	36.51	20.52 ± 10.11	0.28 ± 0.08	H-oriented
	O^1	0.39 ± 0.54	0.38 ± 0.05	2.64 ± 0.32	2.14 ± 0.32	36.49	12.48 ± 11.98	0.24 ± 0.00	H-oriented
Phenoxybenzene	O	0.32 ± 0.49	0.29 ± 0.02	3.40 ± 0.18	1.50 ± 0.13	31.12	5.69 ± 10.11	0.23 ± 0.01	Undefined

Colors represent different functional groups: red for oxygen, blue for nitrogen, orange for sulfur and green for halogen containing groups.

(phenylalanine side-chain), our benzene parameters yield a $\Delta G_{hyd} = -3.4$ kJ/mol, a much closer value to the experimental data ($\Delta G_{hyd} = -3.6$ kJ/mol). Nevertheless, the AVED value reveals a underestimation for free energy of hydration in our parameter set. A possible reason is that chemical functions such as nitro, fluorine, chlorine, and aldehydic carbonyls are not commonly found in biomolecules and, therefore, the LJ parameters used in GROMOS53A6 may not be properly extrapolated to synthetic compounds. Moreover, we have tested ether oxygens LJ parameters reported in Horta et al. (2011) in our pure liquid simulations of [2]furan and [23]methoxybenzene, leading to approximately the same behavior in their respective physical-chemical properties (data not shown).

4.2. Properties in Solution: Influence of Nearby Substitutions in H-Bonds

In order to access quantitative informations regarding how aromatic rings interact with their surroundings, we performed molecular dynamics simulations for 103 aromatic rings most commonly used in drug design, including our 42 molecules

Ascletis Ex. 2056
Page 10 of 20

TABLE 5 | Properties of heteroaromatic rings in water. Average H-bonds ($Aver_{HB}$), H-bond residence time (τ_{HB}) is ps, H-bond lifetime ($lifetime_{HB}$) in 1/ps, free-energy of H-bond breakage (ΔG_{HB}) in kJ/mol, percentage of simulation with at least one formed H-bond (*Percent*.), coordination number of water (*CN*), optimal binding distance with water (OBD_{HB}) in nm, and overall water orientation around the heteroatom (*Orientation*).

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	*Percent*	*CN*	OBD_{HB}	*Orientation*
Water	Ow	1.73 ± 0.62	2.11 ± 0.02	0.47 ± 0.00	6.38 ± 0.03	98.58	4.11 ± 2.83	0.18 ± 0.00	Undefined
	OH1	0.87 ± 0.35	1.80 ± 0.03	0.55 ± 0.01	5.98 ± 0.05	86.25	4.11 ± 2.83	0.18 ± 0.00	O-oriented
	OH2	0.86 ± 0.35	1.83 ± 0.03	0.54 ± 0.01	6.03 ± 0.04	86.07	4.11 ± 2.83	0.18 ± 0.00	O-oriented
Imidazole	N^1	0.08 ± 0.27	0.33 ± 0.09	3.27 ± 0.86	1.68 ± 0.69	7.58	26.01 ± 7.18	0.41 ± 0.01	Undefined
	N^1H	0.56 ± 0.52	0.35 ± 0.00	2.89 ± 0.04	1.90 ± 0.03	55.10	34.04 ± 1.08	0.45 ± 0.02	Undefined
	N^3	1.30 ± 0.72	1.01 ± 0.01	0.99 ± 0.00	4.56 ± 0.01	87.72	7.56 ± 12.01	0.20 ± 0.00	Undefined
Thiazole	S^1	0.04 ± 0.20	–	–	–	4.21	16.56 ± 5.11	0.38 ± 0.01	Undefined
	N^3	0.53 ± 0.60	0.37 ± 0.01	2.73 ± 0.08	2.04 ± 0.07	47.16	0.62 ± 0.11	0.22 ± 0.00	Undefined
Benzopyrrole	N^1	0.09 ± 0.29	0.28 ± 0.05	3.68 ± 0.49	1.32 ± 0.37	8.60	15.88 ± 2.99	0.38 ± 0.00	Undefined
	N^1H	0.63 ± 0.51	0.66 ± 0.01	1.50 ± 0.03	3.51 ± 0.04	61.63	21.05 ± 1.45	0.41 ± 0.01	Undefined
Tetrazole	N^4	0.87 ± 0.69	0.54 ± 0.01	1.85 ± 0.03	3.00 ± 0.04	69.85	1.17 ± 0.02	0.21 ± 0.00	Undefined
	N^3	0.89 ± 0.74	0.53 ± 0.02	1.87 ± 0.06	2.97 ± 0.08	68.12	1.09 ± 0.32	0.23 ± 0.00	Undefined
	N^2	0.31 ± 0.51	0.27 ± 0.02	3.75 ± 0.22	1.26 ± 0.15	29.37	26.38 ± 7.22	0.41 ± 0.00	Undefined
	N^1	0.02 ± 0.14	–	–	–	1.95	21.88 ± 2.06	0.41 ± 0.01	Undefined
	N^1H	0.00 ± 0.00	–	–	–	0.00	0.50 ± 0.03	0.24 ± 0.00	O-oriented
Benzeimidazole	N^1	0.06 ± 0.23	0.21 ± 0.03	4.87 ± 0.67	0.62 ± 0.34	5.53	16.49 ± 2.30	0.40 ± 0.01	Undefined
	N^1H	0.78 ± 0.44	1.21 ± 0.02	0.82 ± 0.01	5.01 ± 0.03	77.23	32.26 ± 2.16	0.46 ± 0.01	Undefined
	N^3	1.08 ± 0.71	0.86 ± 0.01	1.16 ± 0.02	4.16 ± 0.04	80.28	1.32 ± 0.01	0.20 ± 0.00	Undefined
7,8-dihydro-1H-purine	N^6H	0.42 ± 0.51	0.34 ± 0.01	2.91 ± 0.09	1.88 ± 0.08	41.30	30.55 ± 3.92	0.46 ± 0.01	Undefined
	N^1	0.04 ± 0.19	0.20 ± 0.02	5.21 ± 0.66	0.46 ± 0.31	3.73	16.89 ± 12.64	0.28 ± 0.00	O-oriented
	N^6	0.04 ± 0.19	–	–	–	3.70	28.25 ± 2.59	0.46 ± 0.01	Undefined
	N^4	0.33 ± 0.51	0.37 ± 0.04	2.76 ± 0.31	2.02 ± 0.29	31.14	24.55 ± 2.24	0.45 ± 0.01	Undefined
	N^1H	0.96 ± 0.21	6.49 ± 0.11	0.15 ± 0.00	9.16 ± 0.04	95.83	0.79 ± 0.65	0.18 ± 0.00	O-oriented
	N^3	1.83 ± 0.70	1.97 ± 0.09	0.51 ± 0.02	6.20 ± 0.11	97.65	3.38 ± 1.14	0.20 ± 0.00	Undefined
1,2,4 - Triazole	N^2	1.50 ± 0.72	1.53 ± 0.04	0.65 ± 0.02	5.59 ± 0.06	92.83	1.70 ± 0.00	0.20 ± 0.00	Undefined
	N^1	0.66 ± 0.66	0.83 ± 0.02	1.20 ± 0.04	4.07 ± 0.07	55.78	2.10 ± 3.51	0.22 ± 0.00	Undefined
	N^1H	0.98 ± 0.15	11.47 ± 0.25	0.09 ± 0.00	10.58 ± 0.05	97.94	3.76 ± 0.00	0.25 ± 0.03	O-oriented
	N^4	0.83 ± 0.68	0.61 ± 0.01	1.63 ± 0.02	3.32 ± 0.04	67.99	0.97 ± 0.01	0.20 ± 0.00	Undefined
Quinazoline	N^1	0.64 ± 0.63	0.49 ± 0.02	2.04 ± 0.07	2.76 ± 0.09	56.17	0.71 ± 0.21	0.21 ± 0.00	Undefined
	N^3	0.43 ± 0.56	0.31 ± 0.01	3.19 ± 0.05	1.65 ± 0.04	39.34	24.86 ± 3.16	0.28 ± 0.08	H-oriented
1H-pyrimidin-2-one	O^2	1.20 ± 0.72	0.79 ± 0.01	1.27 ± 0.02	3.93 ± 0.03	84.99	1.42 ± 0.02	0.20 ± 0.00	Undefined
	N^1H	0.41 ± 0.51	0.33 ± 0.01	3.07 ± 0.13	1.75 ± 0.10	39.74	24.01 ± 4.58	0.44 ± 0.02	Undefined
	N^1	0.01 ± 0.09	–	–	–	0.73	19.00 ± 2.14	0.41 ± 0.00	Undefined
	N^3	0.89 ± 0.62	0.84 ± 0.00	1.19 ± 0.01	4.11 ± 0.01	75.00	1.05 ± 0.00	0.20 ± 0.00	Undefined
4-quinolone	O^4	1.74 ± 0.74	1.39 ± 0.02	0.72 ± 0.01	5.35 ± 0.04	96.17	3.85 ± 1.52	0.19 ± 0.00	Undefined
	N^1	0.01 ± 0.10	–	–	–	1.03	23.69 ± 2.01	0.47 ± 0.01	Undefined
	N^1H	0.66 ± 0.49	0.80 ± 0.02	1.26 ± 0.04	3.96 ± 0.07	65.61	26.90 ± 2.65	0.47 ± 0.01	Undefined
Isoxazole	O^1	0.59 ± 0.62	0.36 ± 0.00	2.82 ± 0.02	1.96 ± 0.02	52.43	7.34 ± 13.27	0.22 ± 0.01	Undefined
	N^2	0.60 ± 0.63	0.35 ± 0.01	2.86 ± 0.05	1.92 ± 0.04	52.00	0.77 ± 0.04	0.24 ± 0.00	H-oriented
Uracil	N^3	0.02 ± 0.16	1.99 ± 1.45	1.04 ± 0.90	5.38 ± 2.21	2.48	17.32 ± 1.01	0.42 ± 0.01	Undefined
	N^3H	0.33 ± 0.49	0.29 ± 0.01	3.47 ± 0.10	1.45 ± 0.07	31.98	25.64 ± 8.27	0.43 ± 0.02	Undefined
	O^2	0.39 ± 0.54	0.28 ± 0.01	3.61 ± 0.13	1.35 ± 0.09	36.17	19.18 ± 8.50	0.37 ± 0.01	Undefined
	O^4	1.24 ± 0.71	0.87 ± 0.01	1.15 ± 0.01	4.19 ± 0.02	86.66	4.08 ± 5.14	0.20 ± 0.00	Undefined
	N^1	0.01 ± 0.07	–	–	–	0.54	29.81 ± 2.46	0.46 ± 0.01	Undefined
	N^1H	0.45 ± 0.52	0.37 ± 0.01	2.67 ± 0.07	2.09 ± 0.06	44.10	30.23 ± 1.97	0.47 ± 0.01	Undefined

(Continued)

Ascletis Ex. 2056
Page 11 of 20

TABLE 5 | Continued

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	Percent	CN	OBD_{HB}	Orientation
Pyrazole	N^1H	0.00 ± 0.00	–	–	–	0.00	18.32 ± 2.00	0.40 ± 0.01	Undefined
	N^1	0.00 ± 0.00	–	–	–	0.00	17.76 ± 1.63	0.40 ± 0.00	Undefined
	N^2	0.72 ± 0.66	0.44 ± 0.01	2.29 ± 0.07	2.48 ± 0.08	60.69	0.96 ± 0.03	0.21 ± 0.00	Undefined
Pyrazine	N^1	1.15 ± 0.66	1.15 ± 0.03	0.87 ± 0.02	4.88 ± 0.07	85.83	6.55 ± 10.50	0.19 ± 0.00	Undefined
	N^4	1.15 ± 0.65	1.15 ± 0.02	0.87 ± 0.02	4.86 ± 0.05	85.87	6.66 ± 10.74	0.19 ± 0.00	Undefined
1,8-naphthyridin-4(1H)-one	O^4	1.13 ± 0.73	0.71 ± 0.02	1.41 ± 0.04	3.68 ± 0.07	81.51	1.44 ± 0.02	0.20 ± 0.00	Undefined
	N^8	0.28 ± 0.48	0.28 ± 0.02	3.57 ± 0.28	1.38 ± 0.19	26.76	25.87 ± 2.34	0.45 ± 0.01	Undefined
	N^1	0.05 ± 0.23	0.40 ± 0.14	2.81 ± 0.93	2.10 ± 0.84	5.34	22.03 ± 2.09	0.44 ± 0.01	Undefined
	N^1H	0.40 ± 0.51	0.34 ± 0.02	2.99 ± 0.20	1.82 ± 0.17	39.32	26.16 ± 2.29	0.45 ± 0.01	Undefined
Xanthine	N^1H	0.40 ± 0.51	0.38 ± 0.01	2.66 ± 0.10	2.10 ± 0.09	39.37	28.47 ± 4.09	0.46 ± 0.02	Undefined
	O^2	0.52 ± 0.60	0.32 ± 0.03	3.13 ± 0.23	1.71 ± 0.19	46.33	17.09 ± 8.14	0.37 ± 0.01	Undefined
	N^7	0.02 ± 0.12	–	–	–	1.51	25.71 ± 2.69	0.45 ± 0.02	Undefined
	N^7H	0.47 ± 0.52	0.43 ± 0.01	2.33 ± 0.08	2.43 ± 0.08	46.50	26.50 ± 2.03	0.46 ± 0.01	Undefined
	N^1	0.02 ± 0.15	–	–	–	2.33	21.48 ± 4.05	0.44 ± 0.01	Undefined
	N^3	0.03 ± 0.17	–	–	–	3.09	19.25 ± 5.54	0.41 ± 0.01	Undefined
	O^6	0.46 ± 0.57	0.31 ± 0.01	3.25 ± 0.06	1.61 ± 0.05	42.64	8.82 ± 4.23	0.33 ± 0.06	Undefined
	N^9	0.28 ± 0.47	0.28 ± 0.03	3.61 ± 0.36	1.36 ± 0.25	27.40	26.67 ± 2.40	0.46 ± 0.01	Undefined
1,2-dihydro-3H-1,2,4-triazol-3-one	N^1H	0.95 ± 0.24	4.50 ± 0.09	0.22 ± 0.00	8.25 ± 0.05	94.28	1.32 ± 1.05	0.17 ± 0.00	O-oriented
	N^2H	0.48 ± 0.52	0.37 ± 0.01	2.72 ± 0.05	2.04 ± 0.05	46.54	21.76 ± 2.73	0.38 ± 0.00	Undefined
	N^4	1.21 ± 0.67	1.11 ± 0.01	0.90 ± 0.01	4.80 ± 0.03	87.18	1.39 ± 0.00	0.20 ± 0.00	Undefined
	O^3	1.26 ± 0.76	0.79 ± 0.01	1.27 ± 0.01	3.93 ± 0.02	85.27	1.54 ± 0.02	0.20 ± 0.00	Undefined
	N^1	0.02 ± 0.13	–	–	–	1.75	17.03 ± 9.51	0.28 ± 0.00	O-oriented
	N^2	0.03 ± 0.16	–	–	–	2.52	27.49 ± 5.89	0.38 ± 0.00	Undefined
1,3,4 - Thiadiazole	S^1	0.02 ± 0.15	–	–	–	2.32	19.74 ± 5.86	0.39 ± 0.01	Undefined
	N^3	1.33 ± 0.73	1.17 ± 0.03	0.86 ± 0.02	4.91 ± 0.05	88.26	18.35 ± 13.63	0.21 ± 0.00	Undefined
	N^4	1.34 ± 0.73	1.16 ± 0.01	0.86 ± 0.01	4.89 ± 0.02	88.35	1.70 ± 0.02	0.21 ± 0.00	Undefined
Indoxazine	N^2	0.69 ± 0.65	0.45 ± 0.00	2.20 ± 0.02	2.57 ± 0.02	59.11	0.94 ± 0.02	0.22 ± 0.00	Undefined
	O^1	0.74 ± 0.66	0.48 ± 0.01	2.08 ± 0.05	2.72 ± 0.06	62.08	0.84 ± 0.21	0.21 ± 0.00	Undefined
3,9-dihydro-6H-purin-6-one	N^1H	0.46 ± 0.52	0.40 ± 0.01	2.48 ± 0.09	2.28 ± 0.09	45.20	26.74 ± 2.00	0.45 ± 0.00	Undefined
	N^1	0.02 ± 0.14	–	–	–	1.93	22.39 ± 5.57	0.42 ± 0.01	Undefined
	N^9	0.03 ± 0.17	0.71 ± 0.45	1.99 ± 1.06	3.23 ± 1.48	3.05	20.26 ± 3.47	0.42 ± 0.01	Undefined
	N^9H	0.47 ± 0.52	0.36 ± 0.00	2.80 ± 0.03	1.97 ± 0.03	45.60	26.51 ± 3.63	0.44 ± 0.02	Undefined
	N^3	0.11 ± 0.32	0.47 ± 0.19	2.40 ± 0.71	2.49 ± 0.88	11.22	27.05 ± 1.99	0.46 ± 0.01	Undefined
	O^6	1.35 ± 0.77	0.82 ± 0.03	1.21 ± 0.04	4.05 ± 0.08	87.77	1.68 ± 0.02	0.20 ± 0.00	Undefined
	N^7	0.57 ± 0.61	0.44 ± 0.02	2.28 ± 0.11	2.49 ± 0.12	50.12	0.65 ± 0.15	0.23 ± 0.00	Undefined
Benzofuran	O^1	0.50 ± 0.59	0.35 ± 0.01	2.90 ± 0.10	1.89 ± 0.09	44.67	23.40 ± 11.48	0.32 ± 0.11	Undefined
Indazole	N^2	0.40 ± 0.55	0.29 ± 0.02	3.45 ± 0.20	1.46 ± 0.14	36.29	22.45 ± 4.12	0.42 ± 0.01	Undefined
	N^1	0.17 ± 0.39	0.22 ± 0.02	4.62 ± 0.40	0.75 ± 0.22	16.72	16.16 ± 2.36	0.39 ± 0.01	Undefined
	N^1H	0.55 ± 0.52	0.45 ± 0.00	2.22 ± 0.02	2.55 ± 0.02	53.47	18.47 ± 4.54	0.40 ± 0.01	Undefined
Benzothiophene	S^1	0.14 ± 0.35	0.36 ± 0.09	2.99 ± 0.83	1.91 ± 0.67	13.13	17.61 ± 6.64	0.37 ± 0.00	Undefined
Chromone	O^4	1.17 ± 0.73	0.74 ± 0.01	1.35 ± 0.02	3.78 ± 0.04	83.02	1.46 ± 0.01	0.20 ± 0.00	Undefined
	O^1	0.14 ± 0.35	0.39 ± 0.12	2.81 ± 0.70	2.06 ± 0.71	13.72	25.00 ± 1.49	0.46 ± 0.01	Undefined
1,4-naphthoquinone	O^4	0.64 ± 0.61	0.44 ± 0.01	2.27 ± 0.03	2.49 ± 0.03	56.82	0.83 ± 0.01	0.20 ± 0.00	Undefined
	O^1	0.64 ± 0.61	0.44 ± 0.01	2.25 ± 0.08	2.52 ± 0.08	57.12	0.82 ± 0.02	0.20 ± 0.00	Undefined
1,2,3 - Triazole	N^1H	0.82 ± 0.41	1.16 ± 0.02	0.86 ± 0.01	4.89 ± 0.04	80.97	14.14 ± 17.25	0.24 ± 0.10	O-oriented
	N^3	1.11 ± 0.73	0.78 ± 0.01	1.29 ± 0.02	3.90 ± 0.03	80.38	1.47 ± 0.02	0.21 ± 0.00	Undefined
	N^2	1.09 ± 0.74	0.75 ± 0.01	1.34 ± 0.02	3.81 ± 0.03	78.90	1.47 ± 0.00	0.21 ± 0.00	Undefined

(Continued)

Ascletis Ex. 2056
Page 12 of 20

TABLE 5 | Continued

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	*Percent*	*CN*	OBD_{HB}	*Orientation*
	N^1	0.11 ± 0.32	0.21 ± 0.00	4.73 ± 0.09	0.67 ± 0.05	11.02	16.72 ± 8.74	0.39 ± 0.03	Undefined
Pyridazine	N^1	1.42 ± 0.76	1.25 ± 0.01	0.80 ± 0.01	5.09 ± 0.02	89.58	1.83 ± 0.00	0.21 ± 0.00	Undefined
	N^2	1.41 ± 0.76	1.24 ± 0.03	0.81 ± 0.02	5.05 ± 0.07	89.42	1.83 ± 0.00	0.21 ± 0.00	Undefined
Triazine	N^1	0.28 ± 0.48	0.29 ± 0.03	3.46 ± 0.38	1.47 ± 0.27	26.10	29.60 ± 0.96	0.45 ± 0.01	Undefined
	N^3	0.27 ± 0.48	0.29 ± 0.03	3.48 ± 0.35	1.45 ± 0.25	25.94	29.63 ± 1.97	0.46 ± 0.00	Undefined
	N^5	0.27 ± 0.48	0.28 ± 0.02	3.62 ± 0.24	1.34 ± 0.16	25.78	33.15 ± 1.59	0.45 ± 0.01	Undefined
Quinoxaline	N^4	0.34 ± 0.51	0.30 ± 0.02	3.40 ± 0.19	1.50 ± 0.14	32.01	24.66 ± 0.64	0.37 ± 0.10	Undefined
	N^1	0.34 ± 0.51	0.30 ± 0.02	3.34 ± 0.19	1.54 ± 0.14	31.83	25.96 ± 2.58	0.33 ± 0.11	Undefined
Oxazole	O^1	0.34 ± 0.51	0.29 ± 0.01	3.51 ± 0.16	1.42 ± 0.12	31.80	31.43 ± 3.12	0.41 ± 0.10	Undefined
	N^3	0.42 ± 0.56	0.29 ± 0.01	3.48 ± 0.14	1.44 ± 0.10	38.27	33.46 ± 2.32	0.43 ± 0.09	Undefined
Isothiazole	S^1	0.05 ± 0.22	0.70 ± 0.11	1.45 ± 0.20	3.63 ± 0.36	5.11	13.50 ± 1.67	0.37 ± 0.00	Undefined
	N^2	0.30 ± 0.49	0.29 ± 0.02	3.42 ± 0.21	1.48 ± 0.15	28.32	29.65 ± 2.44	0.46 ± 0.01	Undefined
1,3,4 - Oxadiazole	N^3	0.96 ± 0.70	0.65 ± 0.02	1.53 ± 0.04	3.47 ± 0.07	74.33	1.32 ± 0.00	0.22 ± 0.00	Undefined
	N^4	0.96 ± 0.70	0.64 ± 0.01	1.55 ± 0.02	3.44 ± 0.04	73.98	1.31 ± 0.01	0.21 ± 0.00	Undefined
	O^1	0.09 ± 0.29	0.71 ± 0.42	1.88 ± 0.84	3.29 ± 1.36	8.69	29.27 ± 5.26	0.44 ± 0.02	Undefined
1,2,5 - Oxadiazole	O^1	0.64 ± 0.66	0.36 ± 0.01	2.76 ± 0.06	2.01 ± 0.06	53.93	12.92 ± 14.89	0.28 ± 0.09	H-oriented
	N^2	0.35 ± 0.53	0.28 ± 0.02	3.59 ± 0.20	1.37 ± 0.14	32.97	27.66 ± 1.31	0.43 ± 0.01	Undefined
	N^5	0.36 ± 0.53	0.28 ± 0.01	3.54 ± 0.14	1.40 ± 0.10	33.18	29.19 ± 3.89	0.44 ± 0.01	Undefined
1,2,4 - Oxadiazole	N^2	0.57 ± 0.61	0.35 ± 0.01	2.83 ± 0.04	1.95 ± 0.04	50.92	7.40 ± 13.39	0.23 ± 0.01	Undefined
	N^4	0.57 ± 0.57	0.43 ± 0.01	2.30 ± 0.07	2.46 ± 0.07	52.94	0.69 ± 0.01	0.20 ± 0.00	Undefined
	O^1	0.53 ± 0.59	0.34 ± 0.01	2.98 ± 0.05	1.82 ± 0.04	47.71	15.71 ± 18.47	0.33 ± 0.12	Undefined
9H-purine	N^3	0.14 ± 0.35	0.62 ± 0.38	2.12 ± 0.88	2.95 ± 1.29	13.39	22.11 ± 1.98	0.42 ± 0.01	Undefined
	N^9H	0.93 ± 0.28	3.72 ± 0.10	0.27 ± 0.01	7.79 ± 0.07	92.10	0.44 ± 0.00	0.18 ± 0.00	O-oriented
	N^1	0.89 ± 0.64	0.82 ± 0.01	1.22 ± 0.02	4.03 ± 0.04	73.61	1.05 ± 0.01	0.20 ± 0.00	Undefined
	N^9	0.06 ± 0.25	0.23 ± 0.03	4.47 ± 0.62	0.84 ± 0.35	6.42	3.15 ± 3.45	0.29 ± 0.00	O-oriented
	N^7	0.43 ± 0.56	0.34 ± 0.03	2.93 ± 0.23	1.87 ± 0.20	39.57	33.93 ± 0.52	0.29 ± 0.10	H-oriented
1,3-Thiazol-2-amine	N^3	0.33 ± 0.51	0.32 ± 0.03	3.12 ± 0.28	1.72 ± 0.22	30.55	*nan* ± *nan*	0.48 ± 0.01	Undefined
	S^1	0.04 ± 0.19	–	–	–	3.62	27.86 ± 3.96	0.45 ± 0.02	Undefined
	N	0.77 ± 0.53	1.33 ± 0.02	0.75 ± 0.01	5.24 ± 0.04	72	0.81 ± 0.00	0.19 ± 0.00	Undefined
	NH^1	0.74 ± 0.46	1.29 ± 0.01	0.78 ± 0.01	5.16 ± 0.03	72.89	1.35 ± 0.38	0.21 ± 0.01	O-oriented
	NH^2	0.73 ± 0.46	1.15 ± 0.01	0.87 ± 0.01	4.88 ± 0.03	72.43	1.20 ± 0.44	0.21 ± 0.01	O-oriented
Cytosine	N^1	0.09 ± 0.29	0.36 ± 0.14	3.07 ± 0.78	1.86 ± 0.81	8.45	25.84 ± 1.36	0.44 ± 0.01	Undefined
	N^1H	0.32 ± 0.48	0.29 ± 0.01	3.41 ± 0.06	1.48 ± 0.04	31.74	29.19 ± 2.52	0.45 ± 0.00	Undefined
	N	0.88 ± 0.53	2.19 ± 0.03	0.46 ± 0.01	6.47 ± 0.04	79.3	8.19 ± 5.54	0.19 ± 0.00	Undefined
	NH^1	0.73 ± 0.46	1.81 ± 0.02	0.55 ± 0.01	6.00 ± 0.02	72.76	1.06 ± 0.48	0.21 ± 0.01	O-oriented
	NH^2	0.70 ± 0.47	1.37 ± 0.02	0.73 ± 0.01	5.30 ± 0.04	69.18	0.87 ± 0.39	0.20 ± 0.00	O-oriented
	O^1	1.20 ± 0.79	0.69 ± 0.02	1.44 ± 0.04	3.62 ± 0.07	81.29	1.40 ± 0.39	0.21 ± 0.00	Undefined
	N^3	0.93 ± 0.73	0.70 ± 0.05	1.44 ± 0.10	3.63 ± 0.18	71.11	1.34 ± 0.32	0.23 ± 0.00	Undefined
Adenine	N^3	1.98 ± 0.68	3.01 ± 0.10	0.33 ± 0.01	7.26 ± 0.08	98.70	3.96 ± 1.48	0.19 ± 0.00	Undefined
	N^9	0.18 ± 0.40	0.31 ± 0.02	3.26 ± 0.27	1.60 ± 0.20	16.99	25.32 ± 2.77	0.35 ± 0.00	Undefined
	N^9H	0.30 ± 0.47	0.35 ± 0.01	2.89 ± 0.11	1.90 ± 0.09	29.54	23.64 ± 1.66	0.34 ± 0.00	Undefined
	N^7	0.17 ± 0.39	0.30 ± 0.02	3.33 ± 0.23	1.55 ± 0.18	16.86	29.43 ± 2.90	0.45 ± 0.01	Undefined
	N^1	0.14 ± 0.36	0.80 ± 0.49	1.96 ± 1.26	3.43 ± 1.72	13.30	21.59 ± 5.25	0.39 ± 0.00	Undefined
	N	0.92 ± 0.48	3.15 ± 0.08	0.32 ± 0.01	7.37 ± 0.06	84.23	4.04 ± 6.16	0.19 ± 0.00	Undefined
	NH^1	0.68 ± 0.48	1.85 ± 0.04	0.54 ± 0.01	6.05 ± 0.05	67.41	0.66 ± 0.06	0.20 ± 0.00	O-oriented

(Continued)

Ascletis Ex. 2056
Page 13 of 20

TABLE 5 | Continued

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	Percent	CN	OBD_{HB}	Orientation
5-methylindole	NH^2	0.70 ± 0.47	1.50 ± 0.02	0.67 ± 0.01	5.53 ± 0.04	68.92	0.67 ± 0.03	0.20 ± 0.00	O-oriented
	N^1	0.26 ± 0.47	0.39 ± 0.02	2.56 ± 0.11	2.20 ± 0.11	24.48	12.49 ± 0.43	0.34 ± 0.05	Undefined
	N^1H	0.89 ± 0.33	3.45 ± 0.04	0.29 ± 0.00	7.60 ± 0.03	88.55	0.37 ± 0.00	0.19 ± 0.00	O-oriented
3-methyl-1H-indole	N^1	0.18 ± 0.40	0.35 ± 0.05	2.94 ± 0.40	1.88 ± 0.34	17.01	17.42 ± 2.94	0.38 ± 0.00	Undefined
	N^1H	0.71 ± 0.48	1.07 ± 0.01	0.94 ± 0.01	4.69 ± 0.03	70.01	17.37 ± 2.17	0.39 ± 0.01	Undefined
Paraxanthine	O^6	0.58 ± 0.58	0.48 ± 0.01	2.08 ± 0.05	2.72 ± 0.06	52.88	0.76 ± 0.02	0.21 ± 0.00	Undefined
	N^3	0.01 ± 0.10	–	–	–	1.02	18.20 ± 4.10	0.39 ± 0.01	Undefined
	N^3H	0.54 ± 0.52	0.65 ± 0.02	1.55 ± 0.06	3.45 ± 0.09	52.58	18.43 ± 2.43	0.40 ± 0.01	Undefined
	O^2	0.61 ± 0.61	0.45 ± 0.01	2.22 ± 0.05	2.55 ± 0.06	54.13	0.78 ± 0.02	0.21 ± 0.00	Undefined
	N^9	0.79 ± 0.60	0.86 ± 0.01	1.17 ± 0.02	4.15 ± 0.04	69.54	0.95 ± 0.01	0.20 ± 0.00	Undefined
	N^7	0.00 ± 0.05	–	–	–	0.21	28.11 ± 1.53	0.47 ± 0.00	Undefined
	N^1	0.03 ± 0.16	–	–	–	2.66	24.93 ± 0.00	0.48 ± 0.00	Undefined
Theophylline	N^7H	0.33 ± 0.49	0.33 ± 0.01	3.00 ± 0.11	1.80 ± 0.09	32.06	23.41 ± 1.68	0.44 ± 0.01	Undefined
	O^6	0.30 ± 0.48	0.27 ± 0.01	3.73 ± 0.16	1.27 ± 0.11	28.37	13.20 ± 1.45	0.38 ± 0.01	Undefined
	N^3	0.02 ± 0.15	–	–	–	2.41	20.89 ± 2.00	0.45 ± 0.01	Undefined
	O^2	0.60 ± 0.62	0.40 ± 0.01	2.53 ± 0.09	2.23 ± 0.09	53.03	0.63 ± 0.21	0.22 ± 0.00	Undefined
	N^9	0.16 ± 0.38	0.29 ± 0.03	3.44 ± 0.33	1.48 ± 0.24	15.73	27.76 ± 1.85	0.46 ± 0.01	Undefined
	N^7	0.02 ± 0.13	–	–	–	1.78	23.16 ± 3.82	0.43 ± 0.01	Undefined
	N^1	0.01 ± 0.12	–	–	–	1.48	25.52 ± 2.34	0.48 ± 0.01	Undefined
Theobromine	O^6	0.26 ± 0.46	0.26 ± 0.01	3.89 ± 0.10	1.16 ± 0.06	25.11	12.25 ± 1.29	0.37 ± 0.01	Undefined
	N^3	0.00 ± 0.06	–	–	–	0.33	27.56 ± 1.00	0.48 ± 0.01	Undefined
	O^2	0.97 ± 0.68	0.69 ± 0.01	1.46 ± 0.01	3.59 ± 0.02	76.39	1.22 ± 0.00	0.20 ± 0.00	Undefined
	N^9	0.10 ± 0.30	0.28 ± 0.02	3.54 ± 0.27	1.40 ± 0.19	9.65	17.11 ± 2.33	0.42 ± 0.01	Undefined
	N^7	0.01 ± 0.10	–	–	–	1.01	25.82 ± 1.01	0.47 ± 0.01	Undefined
	N^1H	0.00 ± 0.00	–	–	–	0.00	20.86 ± 1.70	0.41 ± 0.02	Undefined
	N^1	0.03 ± 0.18	2.28 ± 1.56	0.99 ± 0.92	5.67 ± 2.35	3.16	18.46 ± 3.28	0.40 ± 0.00	Undefined
2H-tetrazol-5-thiol	N^1H	0.66 ± 0.50	0.56 ± 0.01	1.79 ± 0.04	3.08 ± 0.06	65.29	24.63 ± 6.12	0.43 ± 0.01	Undefined
	S	0.08 ± 0.28	1.98 ± 2.59	1.55 ± 1.03	4.47 ± 2.73	8.34	24.11 ± 12.41	0.35 ± 0.00	Undefined
	SH	0.65 ± 0.59	0.36 ± 0.01	2.75 ± 0.08	2.02 ± 0.07	59.18	15.22 ± 7.54	0.36 ± 0.00	Undefined
	N^3	1.05 ± 0.75	0.67 ± 0.01	1.50 ± 0.03	3.52 ± 0.06	76.69	1.44 ± 0.02	0.21 ± 0.00	Undefined
	N^2	0.47 ± 0.58	0.34 ± 0.00	2.93 ± 0.03	1.86 ± 0.03	42.55	21.28 ± 4.89	0.40 ± 0.01	Undefined
	N^1	0.01 ± 0.09	–	–	–	0.90	16.82 ± 4.27	0.39 ± 0.02	Undefined
	N^4	0.54 ± 0.61	0.37 ± 0.01	2.68 ± 0.10	2.09 ± 0.09	47.48	31.80 ± 0.00	0.33 ± 0.11	Undefined
3-methylisoxazole	O^1	0.87 ± 0.70	0.59 ± 0.01	1.71 ± 0.02	3.20 ± 0.03	69.34	0.99 ± 0.25	0.21 ± 0.00	Undefined
	N^2	0.94 ± 0.72	0.62 ± 0.01	1.61 ± 0.02	3.35 ± 0.03	72.36	1.32 ± 0.02	0.22 ± 0.00	Undefined
5-methylisoxazole	O^1	1.06 ± 0.71	0.79 ± 0.02	1.27 ± 0.03	3.95 ± 0.06	78.70	1.31 ± 0.02	0.20 ± 0.00	Undefined
	N^2	1.03 ± 0.73	0.73 ± 0.01	1.37 ± 0.02	3.75 ± 0.03	76.61	1.42 ± 0.01	0.22 ± 0.00	Undefined
Methylimidazole	N^3	1.51 ± 0.68	1.51 ± 0.03	0.66 ± 0.01	5.55 ± 0.05	94.20	11.94 ± 12.59	0.19 ± 0.00	Undefined
	N^1	0.03 ± 0.18	0.45 ± 0.28	2.86 ± 1.15	2.20 ± 1.28	3.26	29.76 ± 1.84	0.46 ± 0.01	Undefined
2-Methylimidazole	N^3	1.76 ± 0.68	2.28 ± 0.04	0.44 ± 0.01	6.57 ± 0.05	97.30	3.63 ± 0.94	0.19 ± 0.00	Undefined
	N^1	0.11 ± 0.32	0.23 ± 0.02	4.45 ± 0.34	0.83 ± 0.19	10.86	15.18 ± 1.88	0.40 ± 0.01	Undefined
	N^1H	0.87 ± 0.36	1.86 ± 0.02	0.54 ± 0.01	6.06 ± 0.03	86.05	0.35 ± 0.01	0.19 ± 0.00	O-oriented
Guanine	N^1	0.00 ± 0.06	–	–	–	0.40	6.42 ± 4.70	0.27 ± 0.00	O-oriented
	N^1H	0.98 ± 0.15	11.66 ± 0.29	0.09 ± 0.00	10.62 ± 0.06	97.86	2.00 ± 0.29	0.17 ± 0.00	O-oriented
	N^7	0.98 ± 0.65	0.81 ± 0.01	1.24 ± 0.02	3.99 ± 0.03	78.40	1.19 ± 0.01	0.20 ± 0.00	Undefined
	N^3	1.51 ± 0.64	2.33 ± 0.08	0.43 ± 0.02	6.62 ± 0.09	95.37	2.98 ± 1.13	0.19 ± 0.00	Undefined
	N	0.58 ± 0.57	1.16 ± 0.03	0.86 ± 0.02	4.90 ± 0.06	54.33	0.27 ± 0.02	0.19 ± 0.00	Undefined
	NH^1	0.71 ± 0.47	2.24 ± 0.03	0.45 ± 0.01	6.53 ± 0.03	70.16	29.15 ± 1.28	0.35 ± 0.00	Undefined
	NH^2	0.67 ± 0.48	1.62 ± 0.02	0.62 ± 0.01	5.72 ± 0.03	66.50	24.74 ± 7.45	0.34 ± 0.00	Undefined

(Continued)

Ascletis Ex. 2056
Page 14 of 20

TABLE 5 | Continued

Molecule	Atom	$Aver_{HB}$	τ_{HB}	$lifetime_{HB}$	ΔG_{HB}	*Percent*	*CN*	OBD_{HB}	*Orientation*
	O^6	1.91 ± 0.71	2.01 ± 0.02	0.50 ± 0.01	6.26 ± 0.03	98.32	3.73 ± 1.35	0.23 ± 0.05	Undefined
	N^9	0.01 ± 0.09	–	–	–	0.76	3.81 ± 3.08	0.28 ± 0.00	O-oriented
	N^9H	0.97 ± 0.19	11.16 ± 0.09	0.09 ± 0.00	10.51 ± 0.02	96.44	4.90 ± 5.94	0.17 ± 0.00	O-oriented
1-Methylindole	N^1	0.22 ± 0.44	0.65 ± 0.20	1.80 ± 0.84	3.29 ± 0.96	20.67	26.89 ± 0.73	0.47 ± 0.00	Undefined
Chlorobenzene	Cl^1	0.22 ± 0.44	0.34 ± 0.08	3.08 ± 0.70	1.80 ± 0.58	20.83	31.47 ± 1.09	0.36 ± 0.00	Undefined
1,2-dichlorobenzene	Cl^1	0.17 ± 0.39	0.38 ± 0.05	2.70 ± 0.33	2.09 ± 0.32	16.31	19.59 ± 8.61	0.36 ± 0.00	Undefined
	Cl^2	0.17 ± 0.39	0.39 ± 0.05	2.64 ± 0.38	2.14 ± 0.34	16.29	18.60 ± 8.25	0.36 ± 0.00	Undefined
1,3-dichlorobenzene	Cl^1	0.17 ± 0.40	0.38 ± 0.08	2.78 ± 0.63	2.06 ± 0.56	16.74	30.86 ± 3.91	0.36 ± 0.00	Undefined
	Cl^3	0.17 ± 0.39	0.37 ± 0.06	2.79 ± 0.40	2.01 ± 0.37	16.55	26.58 ± 8.67	0.36 ± 0.00	Undefined
1,2,3,4-tetrachlorobenzene	Cl^4	0.13 ± 0.35	0.39 ± 0.12	2.75 ± 0.67	2.11 ± 0.69	12.90	25.88 ± 5.26	0.36 ± 0.00	Undefined
	Cl^1	0.13 ± 0.35	0.43 ± 0.14	2.52 ± 0.58	2.32 ± 0.69	13.06	23.23 ± 6.96	0.37 ± 0.00	Undefined
	Cl^2	0.11 ± 0.32	0.43 ± 0.12	2.47 ± 0.63	2.37 ± 0.66	10.45	22.80 ± 7.24	0.37 ± 0.00	Undefined
	Cl^3	0.11 ± 0.31	0.64 ± 0.32	1.93 ± 0.80	3.14 ± 1.13	10.32	22.24 ± 7.06	0.36 ± 0.00	Undefined
1,2,3,5-tetrachlorobenzene	Cl^5	0.16 ± 0.38	0.29 ± 0.06	3.55 ± 0.69	1.44 ± 0.49	15.90	27.66 ± 8.74	0.36 ± 0.00	Undefined
	Cl^1	0.14 ± 0.36	0.40 ± 0.11	2.68 ± 0.57	2.15 ± 0.60	14.00	23.15 ± 9.67	0.36 ± 0.00	Undefined
	Cl^2	0.11 ± 0.32	0.48 ± 0.25	2.51 ± 0.92	2.46 ± 1.09	10.97	22.95 ± 7.34	0.37 ± 0.00	Undefined
	Cl^3	0.14 ± 0.36	0.79 ± 0.71	1.98 ± 0.84	3.24 ± 1.66	13.90	23.20 ± 8.70	0.36 ± 0.00	Undefined
2-pyridone	O^2	1.55 ± 0.75	1.11 ± 0.02	0.90 ± 0.01	4.79 ± 0.04	93.28	1.82 ± 0.00	0.19 ± 0.00	Undefined
	N^1	0.07 ± 0.27	0.24 ± 0.02	4.21 ± 0.40	0.98 ± 0.24	7.37	19.48 ± 4.86	0.43 ± 0.02	Undefined
	N^1H	0.78 ± 0.43	1.40 ± 0.02	0.71 ± 0.01	5.36 ± 0.03	77.75	26.08 ± 3.66	0.44 ± 0.01	Undefined
1,3,5-triazin-2(1H)-one	N^3	1.09 ± 0.70	1.00 ± 0.03	1.01 ± 0.03	4.52 ± 0.08	80.87	1.35 ± 0.00	0.20 ± 0.00	Undefined
	N^5	0.11 ± 0.32	0.38 ± 0.20	3.10 ± 0.99	1.91 ± 1.04	10.86	26.48 ± 3.24	0.45 ± 0.01	Undefined
	N^1	0.03 ± 0.17	6.76 ± 12.20	1.62 ± 1.41	5.18 ± 4.11	3.06	25.39 ± 7.62	0.43 ± 0.02	Undefined
	N^1H	0.61 ± 0.51	0.55 ± 0.02	1.81 ± 0.06	3.06 ± 0.08	59.92	30.73 ± 1.51	0.46 ± 0.02	Undefined
	O^2	0.61 ± 0.66	0.41 ± 0.02	2.45 ± 0.09	2.31 ± 0.09	51.10	28.96 ± 4.10	0.35 ± 0.00	Undefined
Phenoxazine	O^5	0.68 ± 0.65	0.45 ± 0.01	2.23 ± 0.04	2.54 ± 0.04	58.43	0.83 ± 0.01	0.21 ± 0.00	Undefined
	$N^{10}H$	0.64 ± 0.50	1.10 ± 0.03	0.91 ± 0.02	4.76 ± 0.06	62.98	23.69 ± 5.31	0.45 ± 0.02	Undefined
	N^{10}	0.14 ± 0.36	0.20 ± 0.01	4.98 ± 0.25	0.55 ± 0.13	13.62	14.89 ± 2.51	0.40 ± 0.01	Undefined
7H-purine	N^1	0.40 ± 0.55	0.31 ± 0.01	3.18 ± 0.09	1.66 ± 0.07	37.45	30.48 ± 3.09	0.32 ± 0.10	Undefined
	N^7H	0.48 ± 0.52	0.35 ± 0.01	2.82 ± 0.05	1.96 ± 0.05	47.21	30.09 ± 1.79	0.46 ± 0.01	Undefined
	N^3	0.53 ± 0.61	0.37 ± 0.02	2.68 ± 0.12	2.09 ± 0.11	46.52	28.45 ± 4.00	0.45 ± 0.02	Undefined
	N^9	0.42 ± 0.56	0.32 ± 0.01	3.13 ± 0.08	1.70 ± 0.06	38.40	29.46 ± 1.47	0.41 ± 0.08	Undefined
	N^7	0.02 ± 0.15	1.28 ± 0.78	1.34 ± 1.03	4.52 ± 1.88	2.34	22.55 ± 4.08	0.43 ± 0.01	Undefined
1,4-benzodioxine	O^4	0.49 ± 0.57	0.39 ± 0.01	2.58 ± 0.07	2.18 ± 0.07	45.03	0.50 ± 0.13	0.21 ± 0.00	Undefined
	O^1	0.49 ± 0.57	0.39 ± 0.01	2.58 ± 0.06	2.18 ± 0.06	44.98	0.57 ± 0.02	0.21 ± 0.00	Undefined

Colors represent different functional groups: red for oxygen, blue for nitrogen, orange for sulfur and green for halogen containing groups.

calibration set. These information are condensed in the **Tables 4**, **5**. Simulations were carried for 250 ns to properly sample multiple events of H-bond breakages and solvation shell rearrangements.

Our results reveal non-obvious information about the H-bond availability and strength, as in the case of [5]pyridine/[6]pyrimidine/[56]pyrazine/[70]pyridazine/[71]triazine series (**Figure 3**). While exchanging a pyridine by a pyrimidine ring might lead to apparent gain of a H-bond acceptor, nitrogens of pyrimidine present a ΔG_{HB} of nearly 1 kJ/mol lower than pyridine. Moreover, the *Percent* of time with at least one formed H-bond between water and pyridine nitrogen is higher than the ones in pyrimidine. When comparing pyridine with pyrazine (an addition of another N in *para*), H-bonds are very similar, so as the second and third solvation layers. Also, acceptance capacity in pyrimidine ring is very similar to triazine, where all three nitrogens are located in *meta*. Intriguingly, values for pyridine are very similar to the ones calculated for pyridazine, with a slight increase in OBD_{HB} and a more compact second layer of solvation, as shown in **Figure 3A**. These results suggest that

Ascletis Ex. 2056
Page 15 of 20

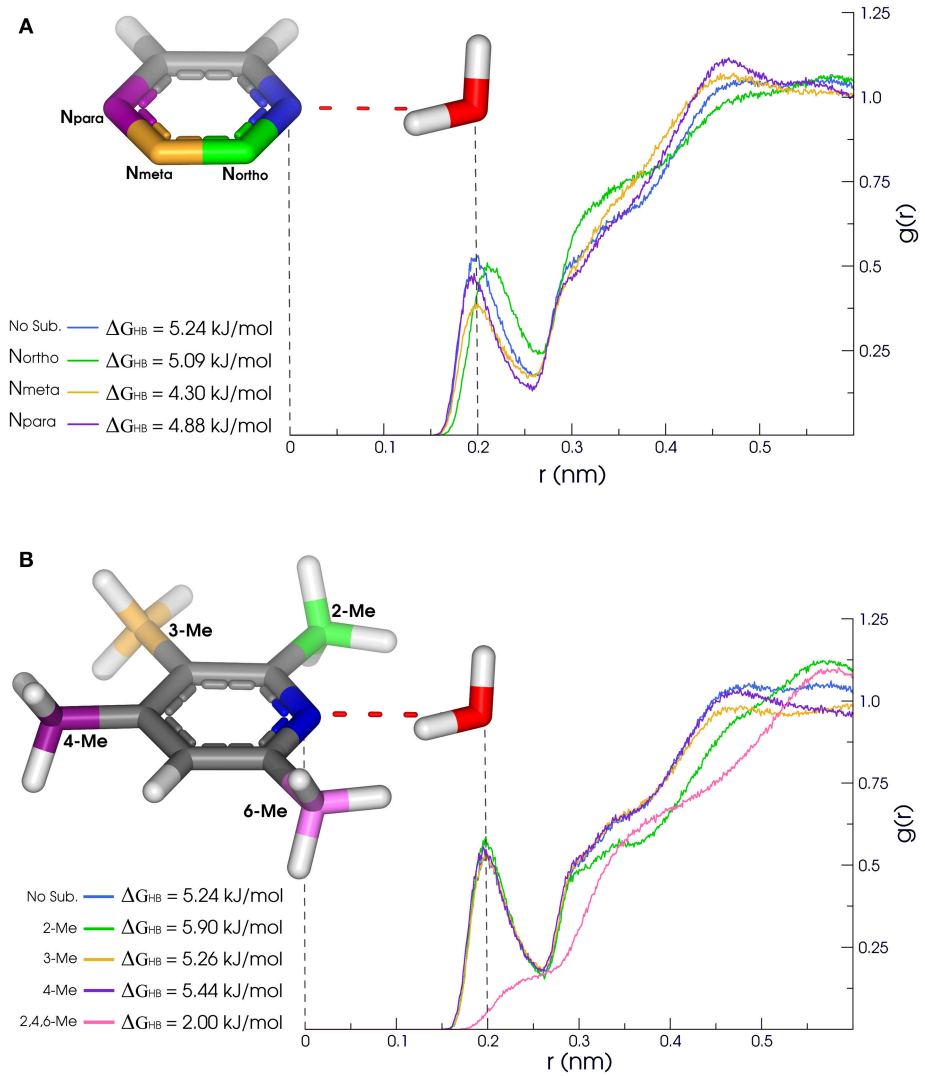


FIGURE 3 | (A) Methyl substituitions: 2-Me (green), 3-Me (yellow), **(B)** Nearby N substitution: N_{ortho} (green), N_{meta} (yellow), 4-Me (purple) and 2,4,6-Me (pink). N_{para} (purple). Solvation properties of aromatic rings in pyridine family. Radial distribution functions (RDFs) and H-bonding strength of N^1 (blue) are affected by substitutions in *ortho*, *meta*, and *para*.

another nitrogen acceptor in *meta* decreases nitrogen acceptance capacity, while another nitrogen acceptor in *ortho* has low effect in H-bond capacity, but a considerable effect in the solvation layers structures. In this sense, these features can impact the binding inside receptors. Pyridazine, for example, has a larger OBD_{HB} than pyridine, suggesting that these molecules can occupy the binding pocket in a different manner, impacting the entropic cost of binding.

Other cases have been equally surprising, like the [39]quinoline/[40]isoquinoline. The main difference between them is the location of the acceptor nitrogen (closer to C^8 in the quinoline fused ring). Counterintuitively, the $Aver_{HB}$ of isoquinoline is slightly lower than for quinoline, such as the τ_{HB}, and the ΔG_{HB} is almost 1.25 kJ/mol lower. The same properties for pyridine ring are somewhat between these values of quinoline and isoquinoline. In addition, ΔG_{HB} for [51]quinazoline and [72]quinoxaline rings are almost 3 kJ/mol lower than quinoline and isoquinoline. In this sense, quinazoline and quinoxaline would be better candidates in fragment-based drug design due to the lower energetic cost of desolvation, while maintaining the H-bond capacity inside the receptor. Another case in terms of aromatic nitrogen hydrogen bond acceptor is the [37]2,4,6-trimethylpyridine (**Figure 3B**). The presence of methyl groups in both *ortho* positions drastically reduces the availability of H-bonds, as shown in **Figure 3**, and diminish the residence time of the accepted H-bond. But the presence of only one methyl group in *ortho* appears to have a modest effect, slightly favoring the presence of H-bond in

Ascletis Ex. 2056
Page 16 of 20

nitrogen of [19]2-methylpyridine. Moreover, the second and third solvation layers of 2- and 2,4,6-trimethylpyridine are dismantled, while the same behavior is not observed for [20]3- and [21]4-methylpyridine.

Other non-obvious events can be observed regarding H-bond donation in hydroxyls groups. In case of [12]phenol, the necessary energy to break a donated H-bond (∼10 kJ/mol) is almost the double to break an accepted one (∼5.70 kJ/mol), in alignment with the QM data reported by Parthasarath et al. (2005) in HF, MP2, and DFT level. And while phenol and [24]phenylmethanol might appear interchangeable during the lead optimization process, the ΔG_{HB} of accepted and donated H-bonds in the hydroxyl group is almost 1 kJ/mol higher for phenylmethanol. While targeting thermodynamics of binding during drug design, these energy costs of desolvation can play a crucial role. As expected, benzenethiol was revealed to be a poor acceptor of hydrogen bonds in our simulations, but a reasonable H-bond donator. In terms of vicinity effects, methylation in *ortho* seems to have little effect on hydroxyl groups, since the properties evaluated for the series [12]phenol/[25]2-methylphenol/[26]3-methylphenol/[27]4-methylphenol have very similar behavior.

It is well– know that halogens are widely used for drug design, and the role of halogen bonds (X-bonds) and H-bonds role have been investigated thoroughly (Rendine et al., 2011; Ford and Ho, 2016; Lin and Mackerell, 2017). In general, the H-bonding strength decreases with the halogen radius (F > Cl > Br > I), while the halogen bond strength increases (Rendine et al., 2011). In this work, we investigated how fluorine and chlorine behave as H-bond acceptors in water. In the case of [7]fluorobenzene, the $\Delta G_{HB} = 1.54 \pm 0.24$ is in accordance with a weak H-bond (Domagała et al., 2017). The other fluorinated rings in the series (1,2-, 1,3-, 1,2,3,4-, and 1,2,3,5-tetrafluorobenzene [8-11]) have similar values, varying from 1.5 to 2.2 kJ/mol. Regarding the chlorinated rings series (chlorobenzene, 1,2-, 1,3-, 1,2,3,4-, and 1,2,3,5-tetrachlorobenzene [94–98]), ΔG_{HB} ranged from 1.80 to 3.24 kJ/mol, contradicting the expected behavior. X-bonding are often poorly described in MM, since it treats atoms as a sphere with isoelectric surface and thus not describing the necessary positive potential required for such interaction. In fact, we have visually evaluated that waters surrounding fluorine and chlorine have their hydrogens oriented toward the halogens, confirming our measure of H-bonds and not X-bonds.

Regarding oxygen atoms within the aromatic ring, $Aver_{HB}$ are generally lower than expected. It is well known that oxygens in heterocycles act as H-bond acceptor (Kaur and Khanna, 2011), but our model does not reproduce this tendency. It is important to notice that GROMOS53A6 does not have specific parameters for oxygens within aromatic rings, and LJ parameters from ethers were employed. Not surprisingly, the calculated properties for the oxygen atom in furan and benzofuran are very similar to methoxybenzene and phenoxybenzene. This result suggests that the description of the properties in aqueous solutions of aromatic rings containing oxygen might be improved by specific LJ parameters. Moreover, we have tested ether LJ parameters reported in Horta et al. (2011) for our simulations of furan and methoxybenzene in water, yielding lower $Aver_{HB}$ and ΔG_{HB} (data not shown). The new force field parameters developed in this work can be obtained upon request.

4.3. Impacts in Drug Design

Recently, several authors have questioned the LE approach as optimization tool and its actual power to lead to high affinity compounds (Abad-Zapatero, 2007; Morgan et al., 2011; Cavalluzzi et al., 2017). Another recent review (DeGoey et al., 2017) has pointed out the emergence of approved drugs that violate Lipinski's rules of 5 and correlated them to properties such as number of aromatic rings and rotatable bonds. Freire (2009) have proposed an experimental thermodynamic approach to guide the drug design process and these results led to believe that tweaking ligand enthalpy and entropy of binding is not only experimentally possible, but also possible to predict. Therefore, the GROMOS series of force fields present an extra advantage here due to their calibration to reproduce free-energy of solvation and other thermodynamical properties.

In this sense, we have parameterized and validated a calibration set of 42 aromatic rings commonly used in drug design using thermodynamical properties in condensed phase. After, we performed a study with a larger dataset of 103 heteroaromatic rings in order to understand how these molecules interact with water and to prospect and map potential interactions with target-receptors. The water molecules probe the occurrence of hydrogen bonds, and the absence of these interactions, as well as the distance from the first solvation sphere, may probe sites for hydrophobic interactions. With these information at hand, medicinal chemists and pharmacologists may employ quantitative estimations on how each functional group may or may not interact with its target protein, as well as identify the potential influence of close chemical modifications. These properties (and a handful of others) are compiled in **Tables 4**, **5**, and can be used as reference during lead optimization process.

The strategy employed here could be used to amplify the spectrum of drug fragments with accurate description of chemical events simulated by molecular dynamics. In addition, it can improve the description of drug-receptor complexation dynamics of other molecules of interest, molecular recognition of drugs and signal transduction mediated by conformational changes of ligands. In fact, by assessing the strength and availability of interactions between aromatic rings and water solvent, the results presented here not only offer detailed quantitative information about potential interactions that each individual aromatic ring can make with its surrounding, but also shed light upon the energetics of biological events, such as dismantling solvation shells — an important step in the ligand binding process.

5. CONCLUSIONS

In this work, we have successfully produced topologies for a calibration set of 42 aromatic rings using as target physicochemical properties of respective organic liquids. Our

Ascletis Ex. 2056
Page 17 of 20

strategy revealed a very competitive prediction power when compared alongside with other force fields, while presenting a simple approach to describe aromatic rings through molecular dynamics simulations that can be easily extrapolated to other rings. In addition to that, H-bond availability and solvent accessibility are difficult and non-obvious informations to predict from bidimensional data, but still essential for medicinal chemistry purposes. Here, we have simulated in aqueous solvent more than 100 aromatic rings commonly used in drug design in order to assess dynamical chemical properties, such as average H-bonds, their lifetime, residence time and free energy of breakage. Thus, we have described a low cost approach based on molecular dynamics simulations to access valuable information that could be useful both to predict the enthalpic cost of desolvation and for interpretation of pharmacological data by a medicinal chemist or pharmacologist. Our results provide a large database of quantitative information for a total of 103 aromatic rings most commonly used in drug design that can guide medicinal chemists in future drug design efforts.

AUTHOR CONTRIBUTIONS

MP carried out quantum calculations, molecular dynamics simulations, data analyses, and drafted the manuscript. VR contributed in the simulations protocols and manuscript draft. BG wrote in house scripts for dipole-based charge assignment and data analyses. MD contributed to manuscript draft. RL contributed to simulations protocols and manuscript draft. HV contributed to data analyses and manuscript draft.

FUNDING

The authors thank the funding agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS). This work was partially supported by grants from FAPERGS/PRONUPEQ (16/2551-0000520-6).

ACKNOWLEDGMENTS

Research developed with support of the Centro Nacional de Supercomputação (CESUP), from Universidade Federal do Rio Grande do Sul (UFRGS). We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2018.00395/full#supplementary-material

REFERENCES

Abad-Zapatero, C. (2007). Ligand efficiency indices for effective drug discovery. *Expert Opin. Drug Dis.* 2, 469–488. doi: 10.1517/17460441.2.4.469

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1-2, 19–25. doi: 10.1016/j.softx.2015.06.001

Aldeghi, M., Malhotra, S., Selwood, D. L., and Chan, A. W. (2014). Two- and three-dimensional rings in drugs. *Chem. Biol. Drug Des.* 83, 450–461. doi: 10.1111/cbdd.12260

Anderson, A. C. (2003). The process of structure-based drug design. *Chem. Biol.* 10, 787–797. doi: 10.1016/j.chembiol.2003.09.002

Aqvist, J., Medina, C., and Samuelsson, J. E. (1994). A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* 7, 385–391.

Bajorath, J. (2015). Computer-aided drug discovery. *F1000 Res.* 4:630. doi: 10.12688/f1000research.6653.1

Barker, J. A. and Watts, R. O. (1973). Monte carlo studies of the dielectric properties of water-like models. *Mol. Phys.* 26, 789–792.

Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* 97, 10269–10280.

Bemis, G. W. and Murcko, M. A. (1996). The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 39, 2887–2893.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81, 3684–3690.

Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R., and van Gunsteren, W. F. (1994). Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* 222, 529–539.

Biela, A., Khayat, M., Tan, H., Kong, J., Heine, A., Hangauer, D., et al. (2012). Impact of ligand and protein desolvation on ligand binding to the S1 pocket of thrombin. *J. Mol. Biol.* 418, 350–366. doi: 10.1016/j.jmb.2012.01.054

Blundell, C. D., Nowak, T., and Watson, M. J. (2016). Measurement, interpretation and use of free ligand solution conformations in drug discovery. *Prog. Med. Chem.* 55, 45–147. doi: 10.1016/bs.pmch.2015.10.003

Blundell, C. D., Packer, M. J., and Almond, A. (2013). Quantification of free ligand conformational preferences by NMR and their relationship to the bioactive conformation. *Bioorg. Med. Chem.* 21, 4976–4987.doi: 10.1016/j.bmc.2013.06.056

Broughton, H. B. and Watson, I. A. (2004). Selection of heterocycles for drug design. *J. Mol. Graph Model.* 23, 51–58.doi: 10.1016/j.jmgm.2004.03.016

Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101. doi: 10.1063/1.2408420

Butler, K. T., Luque, F. J., and Barril, X. (2009). Toward accurate relative energy predictions of the bioactive conformation of drugs. *J. Comput. Chem.* 30, 601–610. doi: 10.1002/jcc.21087

Caleman, C., van Maaren, P. J., Hong, M., Hub, J. S., Costa, L. T., and van der Spoel, D. (2012). Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J. Chem. Theor. Comput.* 8, 61–74. doi: 10.1021/ct200731v

Cavalluzzi, M. M., Mangiatordi, G. F., Nicolotti, O., and Lentini, G. (2017). Ligand efficiency metrics in drug discovery: the pros and cons from a practical perspective. *Expert Opin. Drug Dis.* 12, 1087–1104. doi: 10.1080/17460441.2017

Csermely, P., Korcsmáros, T., Kiss, H. J., London, G., and Nussinov, R. (2012). Structure and dynamics of molecular networks: A novel paradigm of drug discovery. A comprehensive review. *Pharmacol. Ther.* 138, 333–408. doi: 10.1016/j.pharmthera.2013.01.016

Danishuddin and Khan, A. U. (2016). Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug Discov. Today* 21, 1291–1302. doi: 10.1016/j.drudis.2016.06.013

Daura, X., Mark, A. E., and Van Gunsteren, W. F. (1998). Parametrization of aliphatic CHn united atoms of GROMOS96 force field. *J. Comput. Chem.* 19, 535–547.

Ascletis Ex. 2056
Page 18 of 20

DeGoey, D. A., Chen, H. J., Cox, P. B., and Wendt, M. D. (2017). Beyond the rule of 5: lessons learned from AbbVie's drugs and compound collection. *J. Med. Chem.* 61, 2636–2651. doi: 10.1021/acs.jmedchem.7b00717

Dobbs, K. D. and Hehre, W. J. (1986). Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third-and fourth-row, main-group elements. *J. Comput. Chem.* 7, 359–378.

Domagała, M., Lutyńska, A., and Palusiak, M. (2017). Halogen bond versus hydrogen bond: The many-body interactions approach. *Int. J. Quantum Chem.* 117:e25348. doi: 10.1002/qua.25348

Feenstra, K. A., Hess, B., and Berendsen, H. J. C. (1999). Improving efficiency of large time-scale molecular dynamics simulations of hydrogen rich systems. *J. Comput. Chem.* 20, 786–798.

Ferenczy, G. G. and Keseru, G. M. (2010). Thermodynamics guided lead discovery and optimization. *Drug Discov. Today* 15, 919-932. doi: 10.1016/j.drudis.2010.08.013

Fock, V. (1930). Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Phys.* 61, 126–148.

Ford, M. C. and Ho, P. S. (2016). Computational tools to model halogen bonds in medicinal chemistry. *J. Med. Chem.* 59, 1655–1670. doi: 10.1021/acs.jmedchem.5b00997

Freire, E. (2009). A thermodynamic approach to the affinity optimization of drug candidates. *Chem. Biol. Drug Des.* 74, 468–472. doi: 10.1111/j.1747-0285.2009.00880.x

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 09, Revision A.02*. Wallingford, CT.

Ganesan, A., Coote, M. L., and Barakat, K. (2017). Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discov. Today* 22, 249–269. doi: 10.1016/j.drudis.2016.11.001

Gao, Q., Yang, L., and Zhu, Y. (2010). Pharmacophore based drug design approach as a practical process in drug discovery. *Curr. Comput. Aid Drug* 6, 37–49. doi: 10.2174/157340910790980151

Gleeson, M. P. (2008). Generation of a set of simple, interpretable ADMET rules of thumb. *J. Med. Chem.* 51, 817–834. doi: 10.1021/jm701122q

Gleeson, M. P., Hersey, A., Montanari, D., and Overington, J. (2011). Probing the links between *in vitro* potency, ADMET and physicochemical parameters. *Nat. Rev. Drug Discov.* 10, 197–208. doi: 10.1038/nrd3367

Gumbart, J. C., Roux, B., and Chipot, C. (2013). Standard binding free energies from computer simulations: what is the best strategy? *J. Chem. Theor. Comput.* 9, 794–802. doi: 10.1021/ct3008099

Halgren, T. A. (1996). Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* 17, 520–552.

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* 4:17. doi: 10.1186/1758-2946-4-17

Hartree, D. R. and Hartree, W. (1935). Self-consistent field, with exchange, for beryllium. *Proc. R. Soc. A Math. Phys.* 150, 9–33.

Heinz, T. N., van Gunsteren, W. F., and Hünenberger, P. H. (2001). Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* 115, 1125–1136. doi: 10.1063/1.1379764

Hess, B. (2008). P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theor. Comput.* 4, 116–122. doi: 10.1021/ct700200b

Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18, 1463–1472.

Hopkins, A. L., Keserü, G. M., Leeson, P. D., Rees, D. C., and Reynolds, C. H. (2014). The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* 13, 105–121. doi: 10.1038/nrd4163

Horta, B. A., Merz, P. T., Fuchs, P. F., Dolenc, J., Riniker, S., and Hünenberger, P. H. (2016). A GROMOS-compatible force field for small organic molecules in the condensed phase: the 2016H66 parameter set. *J. Chem. Theor. Comput.* 12, 3825–3850. doi: 10.1021/acs.jctc.6b00187

Horta, B. A., Fuchs, P. F., van Gunsteren, W. F., and Hünenberger, P. H. (2011). New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters. *J. Chem. Theor. Comput.* 7, 1016–1031. doi: 10.1021/ct1006407

Jordan, A. M. and Roughley, S. D. (2009). Drug discovery chemistry: a primer for the non-specialist. *Drug Discov Today* 14, 731–744. doi: 10.1016/j.drudis.2009.04.005

Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118, 11225–11236.

Kaur, D. and Khanna, S. (2011). Intermolecular hydrogen bonding interactions of furan, isoxazole and oxazole with water. *Comput. Theor. Chem.* 963, 71–75. doi: 10.1016/j.comptc.2010.09.011

Keserü, G. M. and Makara, G. M. (2009). The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discov.* 8, 203–212. doi: 10.1038/nrd2796

Kunz, A. P., and van Gunsteren, W. F. (2009). Development of a nonlinear classical polarization model for liquid water and aqueous solutions: COS/D. *J. Phys. Chem. A* 113, 11570–11579. doi: 10.1021/jp903164s

Lee, C. H., Huang, H. C., and Juan, H. F. (2011). Reviewing ligand-based rational drug design: the search for an ATP synthase inhibitor. *Int. J. Mol. Sci.* 12, 5304–5318. doi: 10.3390/ijms12085304

Leeson, P. D. and Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 6, 881–890. doi: 10.1038/nrd2445

Li, S., Smith, D. G., and Patkowski, K. (2015). An accurate benchmark description of the interactions between carbon dioxide and polyheterocyclic aromatic compounds containing nitrogen. *Phys. Chem. Chem. Phys.* 17, 16560–16574. doi: 10.1039/c5cp02365c

Limongelli, V., Bonomi, M., and Parrinello, M. (2013). Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6358–6363. doi: 10.1073/pnas.1303186110

Lin, F. Y. and Mackerell, A. D. (2017). Do halogen-hydrogen bond donor interactions dominate the favorable contribution of halogens to ligand-protein binding? *J. Phys. Chem. B* 121, 6813–6821. doi: 10.1021/acs.jpcb.7b04198

Lionta, E., Spyrou, G., Vassilatis, D. K., and Cournia, Z. (2014). Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* 14, 1923–1938. doi: 10.2174/1568026614666140929124445

Lounnas, V., Ritschel, T., Kelder, J., McGuire, R., Bywater, R. P., and Foloppe, N. (2013). Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Comput. Struct. Biotechnol. J.* 5:e201302011. doi: 10.5936/csbj.201302011

Matczak, P. and Wojtulewski, S. (2015). Performance of Møller-Plesset second-order perturbation theory and density functional theory in predicting the interaction between stannylenes and aromatic molecules. *J. Mol. Model.* 21, 41. doi: 10.1007/s00894-015-2589-1

Mennucci, B. and Tomasi, J. (1997). Continuum solvation models: a new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* 106, 5151–5158.

Møller, C. and Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Phys. Rev.* 46, 618–622.

Mondal, J., Friesner, R. A., and Berne, B. J. (2014). Role of desolvation in thermodynamics and kinetics of ligand binding to a kinase. *J. Chem. Theor. Comput.* 10, 5696–5705. doi: 10.1021/ct500584n

Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., and Greyson, D. (2011). The cost of drug development: a systematic review. *Health Policy* 100, 4–17. doi: 10.1016/j.healthpol.2010.12.002

Nosé, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52, 255–268.

Oostenbrink, C., Villa, A., Mark, A. E., and van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* 25, 1656–1676. doi: 10.1002/jcc.20090

Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* 52, 7182–7190.

Parthasarath, R., Subramanian, V., and Sathyamurthy, N. (2005). Hydrogen bonding in phenol, water, and phenol-water clusters. *J. Phys. Chem. A* 109, 843–850. doi: 10.1021/jp046499r

Ascletis Ex. 2056
Page 19 of 20

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., et al. (2010). How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* 9, 203. doi: 10.1038/nrd3078

Petersson, G. A., Bennett, A., Tensfeldt, T. G., Al-Laham, M. A., Shirley, W. A., and Mantzaris, J. (1988). A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *J. Chem. Phys.* 89, 2193–2218.

Rendine, S., Pieraccini, S., Forni, A., and Sironi, M. (2011). Halogen bonding in ligand—receptor systems in the framework of classical force fields. *Phys. Chem. Chem. Phys.* 13:19508. doi: 10.1039/c1cp22436k

Reynolds, C. H. and Holloway, M. K. (2011). Thermodynamics of ligand binding and efficiency. *ACS Med. Chem. Lett.* 2, 433–437. doi: 10.1021/ml200010k

Roughley, S. D. and Jordan, A. M. (2011). The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J. Med. Chem.* 54, 3451–3479. doi: 10.1021/jm200187y

Rusu, V. H., Baron, R., and Lins, R. D. (2014). PITOMBA: Parameter Interface for Oligosaccharide Molecules Based on Atoms. *J. Chem. Theor. Comput.* 10, 5068–5080. doi: 10.1021/ct500455u

Schuler, L. D., Daura, X., and van Gunsteren, W. F. (2001). An improved FROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* 22, 1205–1218. doi: 10.1002/jcc.1078

Shahlaei, M. (2013). Descriptor selection methods in quantitative structure-activity relationship studies: a review study. *Chem. Rev.* 113, 8093–8103. doi: 10.1021/cr3004339

Shirts, M. R. and Pande, V. S. (2005). Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys.* 122:144107. doi: 10.1063/1.1873592

Sliwoski, G., Kothiwale, S., Meiler, J., and Lowe, E. W. (2013). Computational methods in drug discovery. *Pharmacol. Rev.* 66, 334–395. doi: 10.1124/pr.112.007336

Taylor, R. D., MacCoss, M., and Lawson, A. D. (2017). Combining molecular scaffolds from FDA approved drugs: application to drug discovery. *J. Med. Chem.* 60, 1638–1647. doi: 10.1021/acs.jmedchem.6b01367

Taylor, R. D., MacCoss, M., and Lawson, A. D. (2014). Rings in drugs. *J. Med. Chem.* 57, 5845–5859. doi: 10.1021/jm4017625

Tironi, I. G., Sperb, R., Smith, P. E., and van Gunsteren, W. F. (1995). A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* 102, 5451–5459.

Van Gunsteren, W. F. and Berendsen, H. J. C. (1988). A leap-frog algorithm for stochastic dynamics. *Mol. Simulat.* 1, 173–185.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general Amber force field. *J. Comput. Chem.* 25, 1157–1174. doi: 10.1002/jcc.20035

Waring, M. J. (2009). Defining optimum lipophilicity and molecular weight ranges for drug candidates-Molecular weight dependent lower log D limits based on permeability. *Bioorg. Med. Chem. Lett.* 19, 2844–2851. doi: 10.1016/j.bmcl.2009.03.109

Waring, M. J. (2010). Lipophilicity in drug discovery. *Expert Opin. Drug Dis.* 5, 235–248. doi: 10.1517/17460441003605098

Welsch, M. E., Snyder, S. A., and Stockwell, B. R. (2010). Privileged scaffolds for library design and drug discovery. *Curr. Opin. Chem. Biol.* 14, 347–361. doi: 10.1016/j.cbpa.2010.02.018

Woo, H. J. and Roux, B. (2005). Calculation of absolute protein-ligand binding free energy from computer simulations. *Proc. Natl. Acad. Sci. U.S.A.* 102, 6825–6830. doi: 10.1073/pnas.0409005102

Zhao, H. and Caflisch, A. (2015). Molecular dynamics in drug design. *Eur. J. Med. Chem.* 91, 4–14. doi: 10.1002/ijch.201400009

Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I. nonpolar gases. *J. Chem. Phys.* 22, 1420–1426.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer GT and handling Editor declared their shared affiliation.

Copyright © 2018 Polêto, Rusu, Grisci, Dorn, Lins and Verli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Ascletis Ex. 2056
Page 20 of 20