Received: 27 May 2017 | Revised: 29 June 2017 | Accepted: 11 August 2017

DOI: 10.1111/cbdd.13084

RESEARCH ARTICLE

# Aromatic interactions at the ligand–protein interface: Implications for the development of docking scoring functions

**Michal Brylinski**[1,2]

[1]Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA

[2]Center for Computation & Technology, Louisiana State University, Baton Rouge, LA, USA

**Correspondence**
Michal Brylinski, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA.
Email: michal@brylinski.org

**Funding information**
National Institute of General Medical Sciences, Grant/Award Number: R35GM119524

The ability to design and fine-tune non-covalent interactions between organic ligands and proteins is indispensable to rational drug development. Aromatic stacking has long been recognized as one of the key constituents of ligand–protein interfaces. In this communication, we employ a two-parameter geometric model to conduct a large-scale statistical analysis of aromatic contacts in the experimental and computer-generated structures of ligand–protein complexes, considering various combinations of aromatic amino acid residues and ligand rings. The geometry of interfacial π–π stacking in crystal structures accords with experimental and theoretical data collected for simple systems, such as the benzene dimer. Many contemporary ligand docking programs implicitly treat aromatic stacking with van der Waals and Coulombic potentials. Although this approach generally provides a sufficient specificity to model aromatic interactions, the geometry of π–π contacts in high-scoring docking conformations could still be improved. The comprehensive analysis of aromatic geometries at ligand–protein interfaces lies the foundation for the development of type-specific statistical potentials to more accurately describe aromatic interactions in molecular docking. A Perl script to detect and calculate the geometric parameters of aromatic interactions in ligand–protein complexes is available at https://github.com/michal-brylinski/earomatic. The dataset comprising experimental complex structures and computer-generated models is available at https://osf.io/rztha/.

**KEYWORDS**
π–π interactions, aromatic interactions, ligand binding, ligand docking, molecular docking, molecular modeling, non-covalent interactions, parallel stacking, perpendicular stacking, protein–ligand complexes

## 1 | INTRODUCTION

Low molecular weight ligands, such as endogenous compounds and synthetic drugs, reversibly bind to proteins by forming multiple non-covalent interactions predominantly with the side chains of binding pocket residues. In contrast to strong covalent bonds, these rather weak intermolecular contacts comprise a variety of interactions that do not involve sharing electrons. Key interactions between ligands and macromolecules include hydrogen bonds,[1,2] π–π aromatic stacking,[3,4] cation–π interactions,[5,6] hydrophobic effects,[7,8] halogen bonds,[9,10] and salt bridges.[11,12] A significant effort is directed to study the geometrical properties and energetics of these non-covalent bonds because of their paramount importance in molecular recognition and practical applications in drug discovery.[8,13,14] Pharmacology exploits the fact that bioactive compounds have a sufficient specificity and potency to bind and modulate the function of macromolecular targets. At the outset of drug development, the selection of a molecular scaffold requisite for binding is often followed by the optimization of the adjoining chemical moieties non-covalently interacting with pocket residues. Here, the goal is to maximize the affinity of a drug candidate toward the target macromolecule and to reduce its dissociation from the functional site. On that account, the ability to fine-tune a network of non-covalent interactions promoting high-affinity binding of small molecules to their targets is critical for the success of rational drug discovery.

Two distinct, yet complementary computational techniques are used to gain insights into the structure and energy landscape of non-covalent interactions in biological systems. The first approach employs quantum mechanical (QM) calculations, sometimes in combination with molecular mechanics (MM). For instance, the geometrical preferences of various types of hydrogen bonds frequently present at ligand–protein interfaces have been investigated with QM.[15] The derived distance and angle values describing the geometry of hydrogen bonds can subsequently be utilized by empirical methods to effectively model hydrogen bonds upon ligand binding. Another study employed a hybrid QM/MM method to analyze polarization effects playing a significant role in the determination of ligand–protein complex structures.[16] In this approach, fixed charges on ligand atoms obtained from force field parameterization are replaced by those calculated by QM/MM to improve the accuracy of the modeling of molecular assemblies. Finally, the relative strength of π–π and cation–π interactions was investigated as a function of the geometry and protonation state in histidine-aromatic complexes with quantum chemistry methods.[17] In addition to important differences in the stability of aromatic interactions in the gas phase, water, and protein-like environments, it was found that π–π stacking is essential for the favorable electron correlation, whereas cation–π contacts produce further electrostatic contributions. The advantage of QM methods is that these calculations can be applied to a variety of systems and the results obtained for idealized molecules are usually straightforward to interpret. Nonetheless, because of limits on the system size as well as the fact that only simple molecules in vacuum are subject to QM calculations, the derived geometric and energy parameters may not be suitable to reliably model a biological system with its highly complex and heterogeneous environment.

On that account, another computational approach to explore the geometry and energy landscape of various non-covalent interactions at ligand–protein interfaces builds on the accumulated knowledge of the atomic structures of molecular assemblies. For example, accurate potentials of mean force (PMF) can be derived from a large number of complexes deposited in the Protein Data Bank (PDB).[18] The Biomolecular Ligand Energy Evaluation Protocol (BLEEP) was developed to estimate the affinity of ligand binding from the complex structure.[19] BLEEP considers 40 different atom types and employs a reverse Boltzmann methodology to convert the distribution of interaction distances into energy-like pair potential functions. As it was anticipated, these potentials promote short-range polar contacts and hydrogen bonding, whereas the range of hydrophobic interactions is distinctively longer.

Another example of an atomic PMF derived from a database of ligand–protein complexes is the Astex Statistical Potential (ASP).[20] A unique feature of this new potential is that it accounts for differences in the exposure of various types of protein atoms toward ligand-binding sites. Employing ASP in molecular docking considerably improves the accuracy of pose prediction not only across a large validation set of ligand–protein complexes, but also for a small testing set of pharmaceutically relevant targets. Finally, two distance-dependent statistical scoring functions were developed using probability theory, PoseScore to identify native ligand-binding geometries and RankScore to distinguish between binding ligands and non-binding molecules.[21] Both potentials were derived from a set of 8,885 crystallographic structures of ligand–protein complexes employing optimized atomic distance thresholds and including non-native ligand geometries. In addition to experimental structures, the performance of PoseScore and RankScore was evaluated against computer-generated protein models with encouraging results. Because of their remarkable accuracy, these tools can support drug development by predicting ligand–protein complex structures and helping identify potentially bioactive compounds.

In addition to parameters for contact-based and distance-dependent pair potentials routinely derived from interaction statistics in the PDB, this large collection of molecular structures can also be used to parameterize other types of potential functions. For instance, a sophisticated descriptor-based scoring function integrating evolutionary constraints with physics-based energy terms implemented in the GEAUXDOCK docking program[22,23] was parameterized against a representative snapshot of ligand–protein complexes extracted from the PDB. Scoring terms in GEAUXDOCK include electrostatic and van der Waals interactions, hydrogen bonds, hydrophobic interactions, generic and pocket-specific contact potentials, a pseudo-pharmacophore potential, and position restraints on family conserved anchor substructures and the binding site center.

As an example, much deeper potential wells representing strong interactions are assigned by GEAUXDOCK to salt bridges between guanidinium groups in arginine residues and ligand carboxyl moieties, compared to those less favorable, for example, between arginine and amide groups. Moreover, different types of hydrogen bounds have distinct geometries and strengths. Although the mean interaction distance of hydrogen bond between the hydroxyl group on threonine and ligand primary amine is shorter than that for the tyrosine hydroxyl and an amide group, the latter is slightly stronger at the optimal distance. Also, force field parameters to model hydrophobic interactions in GEAUXDOCK are in line with the physicochemical properties of ligand atoms, for example, aromatic carbons and halogens tend toward non-polar residues, whereas amine nitrogen and carboxylate oxygen atoms clearly prefer a polar microenvironment. Indubitably, these carefully derived parameters are pivotal for the accuracy of subsequent molecular modeling simulations. Other studies systematically explore structural data in the PDB to calculate distance- and angle-dependent statistical potentials for hydrogen bonds,[15] investigate the geometrical and energetic features of halogen bonds in biological molecules,[9,24] as well as characterize hydrophobic and aromatic interactions at the ligand–protein interface.[8,25]

Although π–π stacking, defined as an attractive, non-covalent interaction between aromatic rings, is not as widespread as hydrogen bonds and hydrophobic contacts, it plays a vital role in biological recognition and the organization of biomolecular structures. The benzene dimer is a prototypical system to study π–π aromatic stacking. However, investigating this simple system poses significant practical challenges due to its relatively small binding energy of about 2–3 kcal/mol and the fact that the dimer is stable only at low temperatures. An experimental evidence of the perpendicular conformation in crystal and liquid benzene was obtained by molecular beam electric deflection study.[26] A subsequent theoretical research on π–π interactions indicated that favorable perpendicular and offset-parallel configurations correspond to energy minima of comparable depth, whereas the less stable eclipsed geometry represents an energetic saddle point.[27,28] Indeed, perpendicular and offset-parallel configurations are dominant in the crystal structures of simple aromatic compounds[29] and proteins,[30] in contrast to infrequently observed eclipsed stacking.

As the major energetic contributors to π–π interactions are London dispersion forces and electrostatics, many molecular force fields and scoring functions simulate aromatic stacking implicitly with van der Waals and Coulomb potentials rather than employing explicit terms. On that account, a systematic evaluation of the geometry of aromatic interactions in computer-generated models compared to that in the experimental structures of organic ligand–protein complexes can cast light on the accuracy of the modeling of π–π stacking in pharmaceutical design. In this communication, we first conduct a large-scale statistical analysis or interfacial aromatic contacts in the crystal structures of organic ligand–protein complexes employing a two-parameter geometric model. Our study considers various combinations of aromatic protein residues and ring structures in ligand molecules. Subsequently, complex models constructed by contemporary docking software are carefully assessed in terms of the predicted geometry of aromatic contacts. The results have important ramifications for the development of molecular force fields and scoring functions for structure-based drug discovery.

# 2 | METHODS AND MATERIALS

## 2.1 | Geometry of aromatic interactions

Aromatic rings in ligand molecules are identified with the Chemistry::Ring::Find module available in PerlMol.[1] This module implements a breadth-first ring finding algorithm to identify the Smallest Set of Smallest Rings.[31,32] Based on atomic contacts detected with ligand–protein contact (LPC) software,[33] we subsequently select those interactions involving aromatic residues, phenylalanine (F), tyrosine (Y), tryptophan (W) and histidine (H), and ligand aromatic atoms. This procedure produces a complete list of interacting aromatic rings in a given structure of a ligand–protein complex. Next, we employ a two-parameter model, presented in Figure 1, to describe the geometry of each pair of interacting rings. The first parameter in this model is the Cartesian distance between the geometric centers of two rings, referred to as the distance. The second parameter is an angle between normal vectors of two aromatic rings, referred to as the angle. Further, the interaction type is specified using a notation $A^P$:L, where *A* is an amino acid, *P* is the number of amino acid ring atoms, and *L* is the number of ligand ring atoms. For instance, $W^6$:5 denotes an interaction between a 6-member benzene ring of tryptophan and a 5-member aromatic ring of the ligand.

## 2.2 | Dataset of ligand–protein complexes

The protocol employed in this work to compile the dataset of ligand–protein complexes is similar to that previously developed to generate representative and non-redundant sets of ligand–protein complexes for benchmarking of *e*FINDSITE[34] and other binding site prediction algorithms. First, we identified in the PDB protein chains composed of 50–999 amino acids that non-covalently bind small organic molecules. Next,

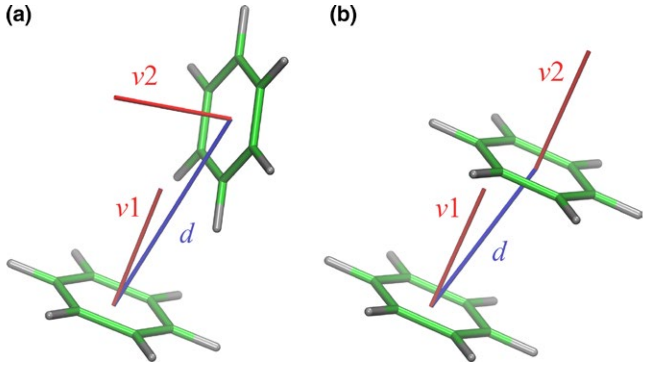

**FIGURE 1** Two-parameter model describing the geometry of aromatic interactions. Parameters are the Cartesian distance *d* between the geometric centers of aromatic rings (blue) and an angle between the normal vectors *v*1 and *v*2 of ring planes (red). Two distinct low-energy configurations of the benzene dimer are shown, (a) perpendicular and (b) parallel [Colour figure can be viewed at wileyonlinelibrary.com]

we retained those proteins binding a single ligand whose Tanimoto coefficient (TC) to at least one FDA-approved drug is ≥0.5. The TC is calculated for 1,024-bit molecular fingerprints with OPENBABEL[35] against FDA-approved drugs in the DrugBank database.[36] Subsequently, protein sequences were clustered with CD-HIT[37] at 40% sequence similarity, and a representative set of proteins binding chemically dissimilar ligands (a pairwise TC of <0.5) at different locations (at least 8 Å apart) were selected from each homologous cluster. Finally, we kept only those complexes stabilized by at least one aromatic interaction between the ligand and the receptor protein identified with LPC.[33] This procedure resulted in a non-redundant and representative dataset of 3,079 proteins bound to ligands containing aromatic moieties, referred to as the daTaset tO evalUate alGoritHms for ligand Docking (TOUGH-D1) dataset.

## 2.3 | Molecular docking

In addition to experimental structures acquired from the PDB,[18] a series of docking models were constructed for TOUGH-D1 complexes with two academic programs, AUTODOCK VINA[38] and RDOCK.[39] VINA has an excellent scoring power according to a recent benchmarking study,[40] and it is the most widely used molecular docking tool.[41] RDOCK is a newly released program that is more accurate and faster than other academic codes. For VINA, MGL tools 1.5.6[42] and OPEN BABEL 2.4.1[43] were used to add polar hydrogens and partial charges, as well as to convert target proteins and library compounds to the PDBQT format. The optimal search space centered on the binding site was defined for each docking ligand from its radius of gyration as described previously.[44] Molecular docking was carried out by AUTODOCK VINA 1.1.2 with the exhaustiveness parameter set to 1,000. For RDOCK, OPEN BABEL 2.4.1[43] was used to convert receptor proteins and ligands to the required Tripos MOL2 and SDFile formats. The docking box was defined by the rcavity program employing the reference ligand method. Simulations were conducted by RDOCK 2013.1 with the default scoring function and 100 docking runs per ligand. Finally, we executed VINA for each ligand–protein system with the –local_only option to generate near-native conformations and the –randomize_only option to generate 100 random configurations avoiding atomic clashes.

Docking models constructed for TOUGH-D1 complexes were evaluated with the Contact Mode Score (CMS)[45] against experimental binding poses. The CMS is a new metric assessing the conformational similarity based on intermolecular ligand–protein contacts, which is less dependent on the ligand size compared to the widely used root-mean-square deviation. It ranges from about 0 for random binding poses to 1 for identical configurations. Two sets of conformations were compiled from docking models generated by VINA, a native-like set comprising those configurations having a CMS of ≥0.5 and a random set of models whose CMS to the experimental structure is <0.3. In addition, we prepared a set of high-scoring models for each docking program based on the predicted binding affinity.

## 2.4 | Data analysis and visualization

Two-dimensional histograms of the distribution of geometric parameters describing aromatic interactions, the distance and the angle, were smoothed with the kernel density estimation (KDE) technique.[46,47] KDE is a nonparametric method estimating the probability density function of a set of variables based on a finite data sample. The parameter space was first discretized to a $100 \times 100$ matrix and then populated with observations, that is, distance and angle values computed for each aromatic interaction in a given dataset. To smooth the data with KDE, each observation was represented by a Gaussian kernel, which is a non-negative function integrating to 1 and with a mean of 0. MATRIX2PNG 1.2.2[48] was then used to generate heat maps showing the correlative distribution of geometric parameters for aromatic interactions. In addition to the visual analysis of heat maps, the overlap between two probability distributions is measured with the $G$ test[49,50]:

$$G = 2\sum_{i} \mathrm{O}_i \times \ln\left(\frac{\mathrm{O}_i}{\mathrm{E}_i}\right) \tag{1}$$

where $\mathrm{O}_i$ and $\mathrm{E}_i$ are the observed and expected counts in the $i$th cell and the sum is taken over all nonempty cells in the $100 \times 100$ matrix. Finally, p$K_\mathrm{a}$ values are assigned to histidine residues with PROPKA 3.1[51,52] and the molecular structures of ligand–protein complexes are visualized with VISUAL MOLECULAR DYNAMICS 1.9.3.[53]

# 3 | RESULTS

## 3.1 | Geometry of aromatic contacts in experimental complex structures

The TOUGH-D1 dataset of experimental structures compiled in this study comprises 3,079 ligand–protein complexes forming a total number of 8,148 interactions between 4,967 aromatic rings of organic ligands and 5,961 protein residues. The amino acid composition of these interactions is 37.4% $\mathrm{F}^6$, 23.5% $\mathrm{Y}^6$, 24.7% $\mathrm{H}^5$, 5.5% $\mathrm{W}^5$, and 8.9% $\mathrm{W}^6$. Further, 56.0% and 44.0% of aromatic interactions involve 6- and 5-member rings in ligand molecules, respectively, whose atomic makeup is 75.2% carbon, 24.7% nitrogen, 0.1% sulfur, and 0.03% oxygen. Table 1 shows the composition of aromatic interactions in TOUGH-D1 complexes. Phenylalanine, tyrosine, and tryptophan residues form more interactions with 6-member ligand rings, in contrast to histidine residues that prefer to interact with 5-member ligand rings. Aromatic

**TABLE 1** Composition of aromatic interactions across the TOUGH-D1 dataset

| | Ligand ring | |
|---|---|---|
| Amino acid ring | 6-member (%) | 5-member (%) |
| $F^6$ | 21.8 | 15.6 |
| $Y^6$ | 15.4 | 8.1 |
| $W^5$ | 3.4 | 2.1 |
| $W^6$ | 5.8 | 3.1 |
| $H^5$ | 9.6 | 15.1 |

Interactions between phenylalanine ($F^6$), tyrosine ($Y^6$), tryptophan ($W^5$ and $W^6$) and histidine ($H^5$) residues, and 6- and 5-member ligand aromatic rings are considered.

interactions involving a bicyclic side chain of tryptophan more often are formed through its 6-member benzene ring than a 5-member nitrogen-containing pyrrole ring.

Figure 2 shows the correlative distribution of distance and angle values for all interaction types calculated for the experimental structures of ligand–protein complexes included in the TOUGH-D1 dataset. The distance ranges from 3 to 8 Å, and the angle is within the acute range of 0°–90°. In general, most interactions have two distinct regions of a high probability density. The first recognizable geometry is described by a relatively close distance of about 3.5–4.5 Å between the centers of aromatic rings, and a small angle of 0°–15° between the normal vectors of ring planes. This range of parameters corresponds to a parallel aromatic stacking, shown for the benzene dimer as a prototypical system in Figure 1b. The second densely populated region described by longer distances of 5–6 Å and near-right angles corresponds to a perpendicular aromatic interaction, another typical geometry illustrated for the benzene dimer in Figure 1a.

Distance and angle values obviously mutually depend on one another, for example, rotating one aromatic ring by 90° in order to change the parallel stacking to a perpendicular orientation pushes rings farther apart because of the van der Waals repulsion between the clouds of ring atoms. Further, there are certain differences between interactions involving various amino acids and ligand rings. For instance, the presence of a hydroxyl group in tyrosine perceptibly changes the geometry of aromatic interactions with respect to phenylalanine. Notably more intermediate configurations between parallel and perpendicular regions can be observed for $F^6$:6 (Figure 2a) and $F^6$:5 (Figure 2b) compared to $Y^6$:6 (Figure 2c) and $Y^6$:5 (Figure 2d). The 5-member aromatic ring in tryptophan favors a parallel stacking with 6-member ligand rings (Figure 2e), but forms both parallel and perpendicular interactions with 5-member ligand rings (Figure 2f). Because of the fused ring structure of tryptophan side chains, $W^6$:6 has a dual peak within the perpendicular region (Figure 2g), and the majority of $W^6$:5 interactions are parallel with a somewhat broader range of distances between rings of 3–5.5 Å (Figure 2h).

Finally, histidine residues form both parallel and perpendicular interactions with 6-member ligand rings (Figure 2i), whereas perpendicular geometries are the most common for those interactions involving 5-member ligand rings (Figure 2j). Depending on the pH, a histidine switches between the double-protonated form with both δ and ε nitrogen atoms protonated, and the neutral state with either δ or ε nitrogen protonated. On that account, in Figure 3, we show the distribution of geometrical parameters for $H^5$:6 and $H^5$:5 interactions in two groups identified based on the predicted p$K_a$ shift from the model value

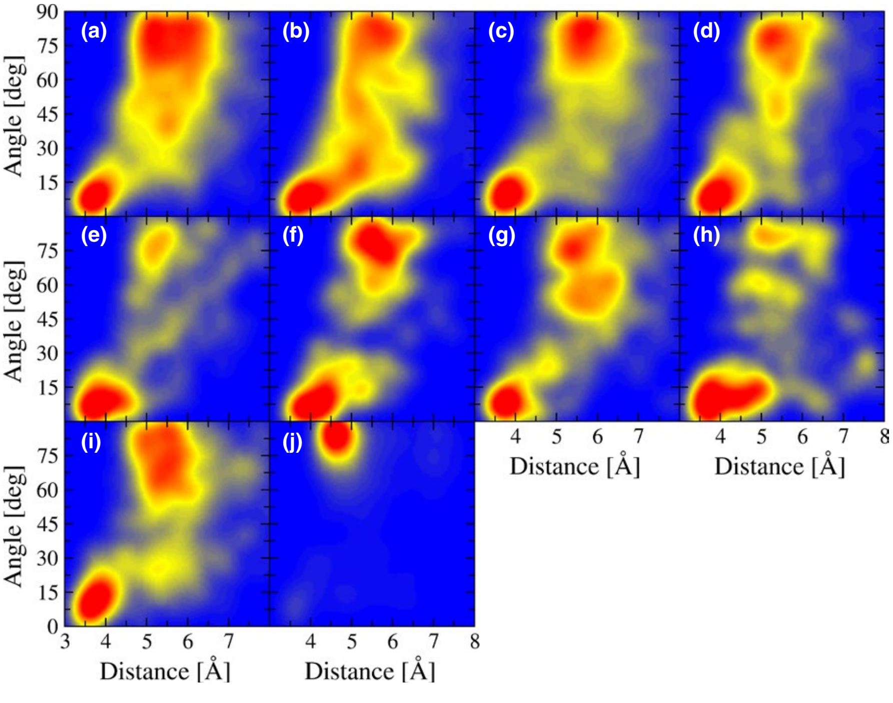

**FIGURE 2** Heat maps showing the distribution of the geometrical properties of aromatic interactions across the experimental structures of ligand–protein complexes included in the TOUGH-D1 dataset. The interaction geometry is described by two parameters, an angle between the normal vectors of aromatic rings and a distance between ring centers. The following interaction types are presented: (a) $F^6$:6, (b) $F^6$:5, (c) $Y^6$:6, (d) $Y^6$:5, (e) $W^5$:6, (f) $W^5$:5, (g) $W^6$:6, (h) $W^6$:5, (i) $H^5$:6, and (j) $H^5$:5 [Colour figure can be viewed at wileyonlinelibrary.com]

of 6.5 for the imidazole side chain. According to PROPKA3,[51,52] as many as 84.6% of histidines interacting with ligand aromatic rings are acidic with negative p$K_a$ shifts from the model value, most likely because the protonation of histidine residues has a stabilizing effect of about 1–3 kcal/mol.[17,54] Consequently, interaction geometries shown in Figure 2i,j are biased toward protonated histidine residues, for which the distributions of distance and angle values are also independently presented in Figure 3a,b. Interestingly, the protonation of the imidazole side chain has a notable effect on the geometry of aromatic interactions. The vast majority of $H^5$:5 interactions involving protonated histidine residues are perpendicular, likely due to the additional favorable contribution from the cation–π electrostatic energy.[17] In contrast, the distributions of geometrical properties of π–π interactions involving neutral histidine residues shown in Figure 3c,d are qualitatively similar to those computed for other aromatic amino acids.

## 3.2 | Examples of aromatic interactions stabilizing ligand–protein complexes

We selected a series of representative examples to examine the molecular structures of aromatic interactions in the context of high probability density regions in heat maps presented in Figure 2. Figure 4 shows four ligand–protein complexes stabilized by various aromatic contacts. Predominant interactions are exemplified by those formed by pyridoxal-5′-phosphate and phenylalanine residues. The offset-parallel aromatic stacking in glutamine aminotransferase from *Thermus thermophilus* HB8 (Figure 4a, PDB-ID: 1v2d, chain A)[55] has a distance of 3.8 Å and an angle of 8.5°, whereas the perpendicular stacking in maize serine racemase (Figure 4b, PDB-ID: 5cvc, chain B)[56] has a distance of 4.9 Å and an angle of 79.5°. Further, complexes shown in Figure 4c,d typify interactions between nucleotides and tyrosine residues. The offset-parallel aromatic stacking between ADP and Y474 in isocitrate dehydrogenase kinase/phosphatase from *Escherichia coli* (Figure 4c, PDB-ID: 3lc6, chain A)[57] consists of two interactions, $Y^6$:5 and $Y^6$:6. The former involves a 5-member ring a1 and has a distance (angle) of 4.0 Å (7.4°), and the latter involves a 6-member ring a2 and has a distance (angle) of 4.0 Å (6.3°). ATP bound to Hmd co-occurring protein HcgE from *Methanothermobacter marburgensis* forms two perpendicular interactions with Y91 (Figure 4d, PDB-ID: 3wv8, chain B).[58] In this complex structure, $Y^6$:5 (ring a1) has a distance of 5.1 Å and an angle of 76.9°, whereas $Y^6$:6 (ring a2) has a distance of 5.1 Å and an angle of 77.2°.

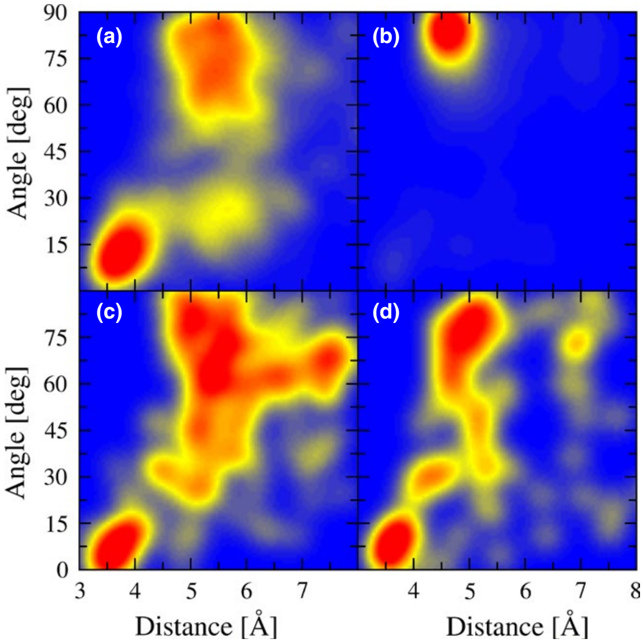

**FIGURE 3** Heat maps showing the distribution of the geometrical properties of aromatic interactions involving histidine residues across the experimental structures of TOUGH-D1 complexes. The interaction geometry is described by two parameters, an angle between the normal vectors of aromatic rings and a distance between ring centers. Interactions are divided into two groups with (a, b) negative and (c, d) positive shifts from the model p$K_a$ value for the imidazole side chain. The following interaction types are presented: (a, c) $H^5$:6 and (b, d) $H^5$:5 [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 5 presents three examples of ligand–protein complexes stabilized by multiple aromatic interactions involving tryptophan and histidine residues. Proflavin forms a 3-layer parallel stacking with two tryptophan residues, W95 and W126, when bound to multidrug binding protein EbrR from *Streptomyces lividans* (Figure 5a, PDB-ID: 3hth, chain A). For example, the middle ring a2 in proflavin forms two offset-parallel interactions with W95, $W^5$:6 with a distance (angle) of 4.6 Å (4.2°) and $W^6$:6 with a distance (angle) of 3.7 Å (4.1°), as well as a near-parallel interaction with W126, $W^5$:6 whose distance (angle) is 5.9 Å (36.5°). In addition, the EbrR/proflavin complex is stabilized by aromatic interactions with F67 and Y107; for instance, proflavin interacts with F67 through a perpendicular stacking $F^6$:6 with a distance (angle) of 5.2 Å (83.9°). A series of perpendicular aromatic interactions are formed between tryptophan residues W47 and W52 of the human catalytic elimination antibody 13G5 and a hapten molecule (Figure 5b, PDB-ID: 3fo2, chain B).[59] Distance (angle) values for selected contacts involving W47, $W^5$:5, $W^6$:6, and $W^5$:6 are 5.7 Å (85.1°), 6.3 Å (85.6°), and 6.1 Å (85.6°), respectively. Also, histidine H98 perpendicularly interacts with the ring a1 of the ligand with a distance of 4.8 Å and an angle of 75.1°. The last example is metallo-β-lactamase from *Bacteroides fragilis* bound to a potent inhibitor (Figure 5c, PDB-ID: 1a8t, chain A).[60] This complex is stabilized by a variety of aromatic interactions involving three histidine residues, H84, H145, and H206, as well as tryptophan W32. For instance, H206 forms a parallel stacking $H^5$:6 against ring a3 with a distance (angle) of 5.6 Å

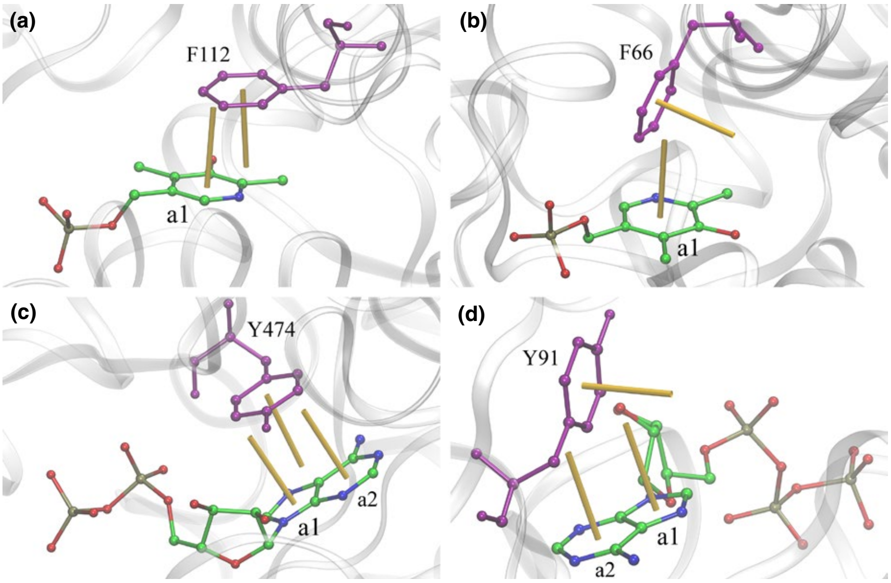

**FIGURE 4** Examples of aromatic stacking in ligand–protein complexes involving phenylalanine and tyrosine residues. (a, b) pyridoxal-5′-phosphate and phenylalanine residues, and (c, d) ADP/ATP and tyrosine residues. Two distinct stacking geometries are presented: (a, c) offset-parallel and (b, d) perpendicular. Ligands are colored by atom type (C—green, N—blue, O—red, P—ocher), whereas aromatic binding residues are purple. Normal vectors of aromatic rings are shown as yellow sticks [Colour figure can be viewed at wileyonlinelibrary.com]

(22.0°), and a perpendicular interaction $H^5$:5 against ring a1 with a distance (angle) of 4.0 Å (72.0°). Furthermore, both rings of W32 make parallel contacts with ring a3 of the inhibitor whose distances (angles) are 5.7 Å (11.8°) for $W^5$:6 and 5.5 Å (11.4°) for $W^6$:6.

## 3.3 | Aromatic contacts in computer-generated complex structures

Heat maps presented in Figure 2 clearly demonstrate that aromatic interactions in the experimental structures of ligand–protein complexes tend to adopt certain geometries. An important question from a modeling point of view is whether similar stacking configurations are observed in the theoretical models of complex structures generated by molecular docking. To look into this facet of ligand docking, we analyze aromatic interactions in four sets of theoretical models of ligand–protein assemblies constructed for the TOUGH-D1 dataset. In addition to heat maps presented in Figure 6, the deviation from the reference probability distribution obtained for experimental structures is quantified by the *G* test. *G*-values are reported in Table 2 for all interaction types with smaller values indicating a better agreement with the reference distribution. Heat maps generated for native-like configurations (1st column in Figure 6) are very similar to those shown in Figure 2 for experimental structures. Further, the average *G*-value for this set is as low as 0.184. This can be expected because employing a CMS threshold of ≥0.5 ensures that the modeled ligand–protein contacts are highly correlated with those in experimental structures.

VINA and RDOCK construct ligand conformations in which the geometry of aromatic interaction does not deviate far away from that in experimental complexes (2nd and 3rd columns in Figure 6). Although the average *G*-values for both docking tools are comparable, there are notable differences with respect to individual interaction types. In general, binding poses generated by VINA have better geometries of aromatic interactions involving 6-member rings in ligand molecules than RDOCK. For instance, *G*-values for $F^6$:6, $Y^6$:6, $W^5$:6, and $H^5$:6 modeled by VINA are 0.122, 0.118, 0.359, and 0.312, respectively, compared to 0.179, 0.165, 0.753, and 0.340 for RDOCK. However, RDOCK seems to model certain interactions involving 5-member rings more accurately than VINA; for example, *G*-value for $W^5$:5 ($H^5$:5) is 0.581 (0.278) for RDOCK and 1.065 (0.601) for VINA. The quality of other interaction types, such as $F^6$:5, $Y^6$:5, $W^6$:6, and $H^5$:6, in docking models is to a large extent independent of the docking algorithm. For comparison, the last column in Figure 6 shows the distribution of distance and angle values across a dataset of random ligand–protein configurations. The correlative distribution of geometric parameters for aromatic interactions is totally lost when ligands are arbitrarily positioned within their target binding sites only to avoid steric clashes. In addition, *G*-values for random conformations reported in Table 2 are much higher than those calculated for high-scoring models by VINA and RDOCK; for example, the average *G*-value across the random dataset is as high as 2.185, compared to only 0.401 for VINA and 0.399 for RDOCK.

## 4 | DISCUSSION

In this communication, we report the results of a statistical analysis of aromatic contacts at the ligand–protein interface. A two-parameter model considering a distance between the geometric centers of aromatic rings and an angle between normal vectors of ring planes is employed to investigate interfacial π–π interactions in experimental as well as computer-generated complex structures. Our analysis of X-ray crystallography data is consistent

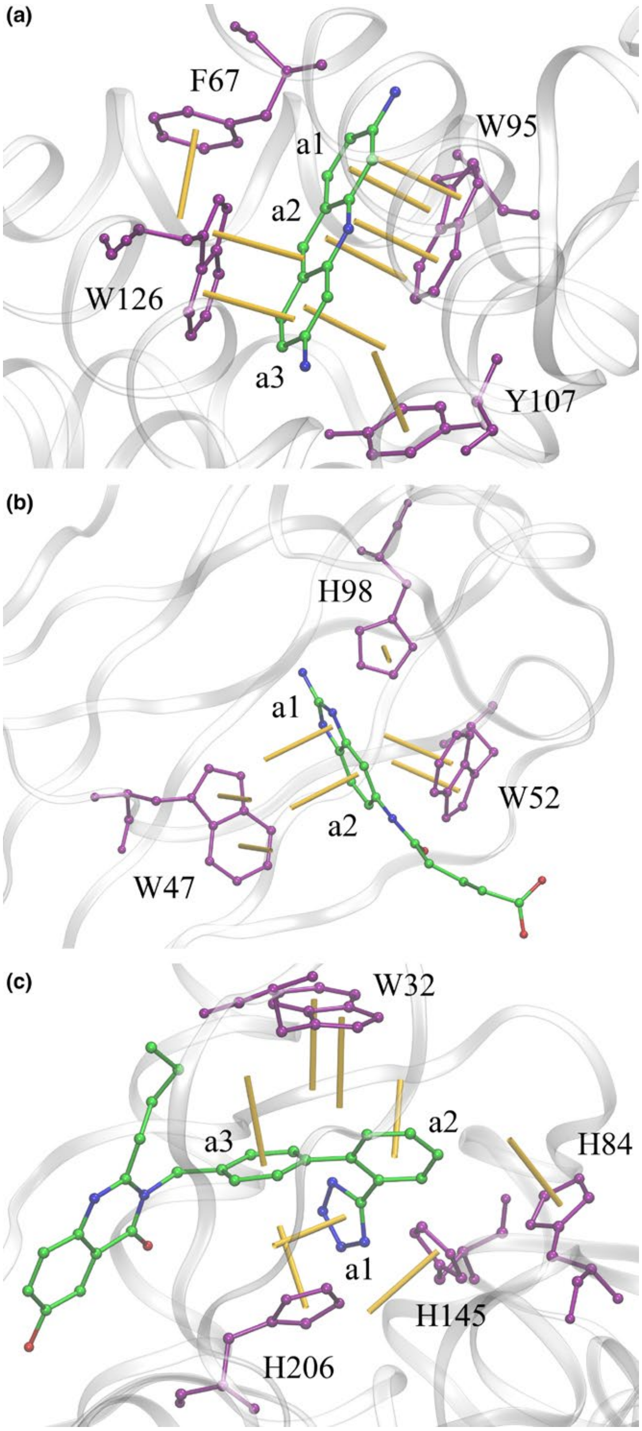

**FIGURE 5** Examples of aromatic stacking in ligand–protein complexes involving tryptophan and histidine residues. (a) EbrR complexed with proflavin, (b) catalytic antibody 13G5 complexed with a hapten, and (c) metallo-β-lactamase complexed with a biphenyl tetrazole inhibitor. Ligands are colored by atom type (C—green, N—blue, O—red), whereas aromatic binding residues are purple. Normal vectors of aromatic rings are shown as yellow sticks [Colour figure can be viewed at wileyonlinelibrary.com]

with the geometrical and energetic properties of aromatic stacking reported previously. Exploring the potential energy surface for the benzene dimer with ab initio methods revealed two nearly isoenergetic structures, perpendicular and offset-parallel configurations.[61] The ideal offset-parallel conformation with an angle between ring planes of 0° has the lowest energy of −3.33 kcal/mol at a horizontal displacement of 1.54 Å, whereas perfectly perpendicular structures with an angle of 90° have interaction energies of −2.84 kcal/mol (point-face) and −2.51 kcal/mol (edge-face).[27] Another QM study predicted that perpendicular and offset-parallel configurations are nearly isoenergetic with binding energies of 2.7 and 2.8 kcal/mol, respectively. These two low-energy arrangements of aromatic rings appear as distinct areas of high probability density on the distance-angle maps constructed in this study for different combinations of protein aromatic residues and ligand ring structures present in the TOUGH-D1 dataset.

With respect to the geometry of π–π interactions, a distance of 4.96 Å between the centers of mass of individual rings in the perpendicular benzene dimer in the gas phase was measured by rotational experiments.[62] Similar distances of 5.0–5.1 Å were obtained with QM calculations.[27,28,61,63] These values match a lower bound of the intermonomer distance range in ligand–protein complexes, which in our analysis extends to about 6 Å. This broader distance range is likely caused by divers chemical moieties attached to aromatic rings in ligand molecules. It has been reported that adding a substituent to one of the rings in the benzene dimer impacts the π–π interaction energy and geometry compared to unsubstituted systems.[64] Specifically, Monte Carlo simulations demonstrated that the presence of a substituent tends to increase the intermonomer separation in perpendicular conformations. For instance, the lowest energy structure of the benzene/phenol dimer has a distance of 5.6 Å, which increases to 6.2 Å as the dimer adopts an edge-to-edge configuration. Furthermore, ab initio calculations together with the analysis of X-ray crystallography data showed that aromatic molecules form stacks with the vertical separation of 3.3–4.1 Å between ring planes in various parallel orientations,[27–29] closely matching the distance range for the parallel aromatic stacking in our analysis. Interestingly, a quantum chemistry study of the benzene dimer revealed the presence of a shallow minimum on the path interconverting offset-parallel structures through a perpendicular saddle point.[27] This configuration corresponds to a tilt angle between phenyl ring planes of about 45°, which can also be observed as a medium probability density area at a distance of approximately 5.5 Å in our distance-angle maps generated for phenylalanine and tyrosine residues.

Previous studies investigating aromatic stacking at the ligand–protein interface are often focused only on nucleotide binding; for instance, it was reported that 65% of adenylate-protein complexes form π–π interactions between adenine bases and aromatic side chains with two predominant orientations, offset-parallel and perpendicular.[65] A similar model

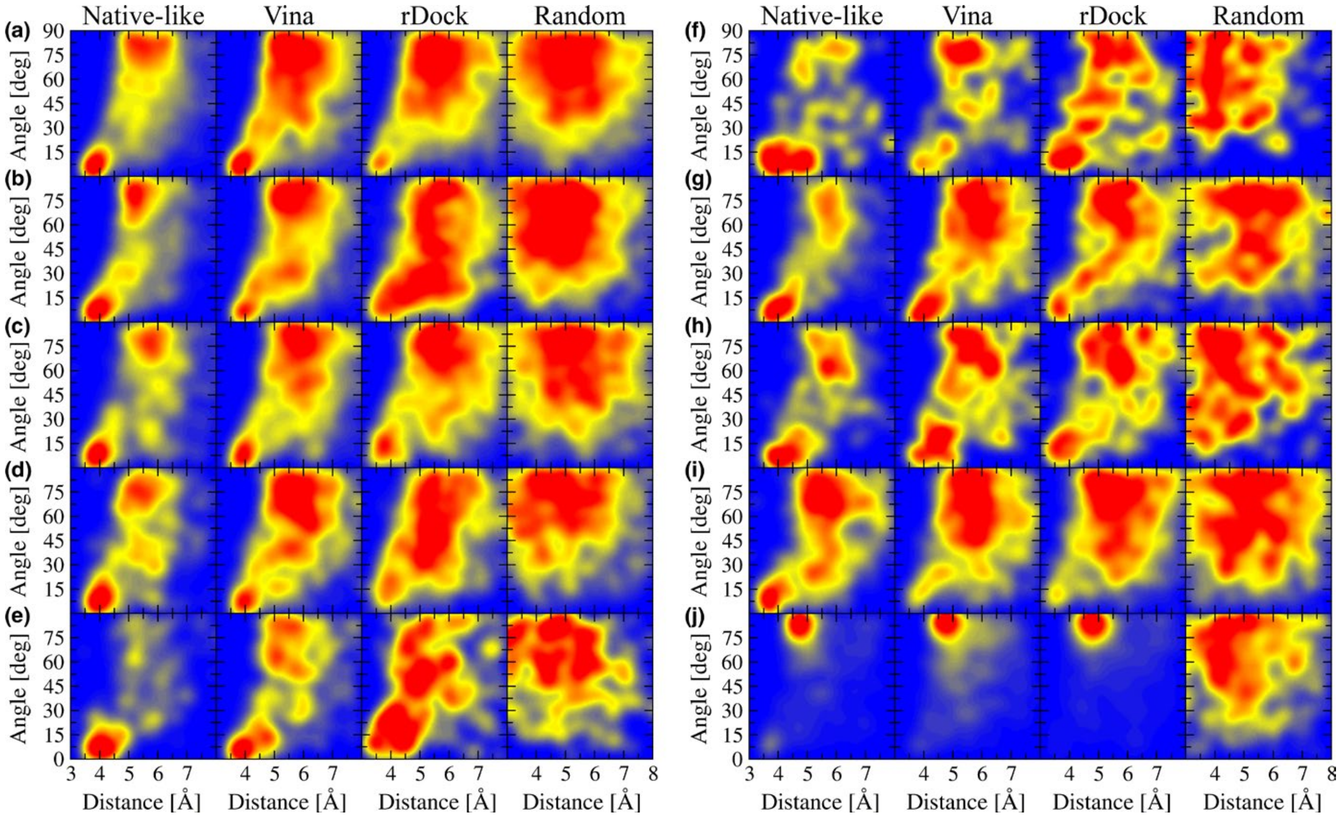

**FIGURE 6** Heat maps showing the distribution of the geometrical properties of aromatic interactions across the computer-generated models of ligand–protein complexes included in the TOUGH-D1 dataset. The interaction geometry is described by two parameters, an angle between the normal vectors of aromatic rings and a distance between ring centers. The following interaction types are presented: (a) $F^6$:6, (b) $F^6$:5, (c) $Y^6$:6, (d) $Y^6$:5, (e) $W^5$:6, (f) $W^5$:5, (g) $W^6$:6, (h) $W^6$:5, (i) $H^5$:6, and (j) $H^5$:5. Four models are considered for each interaction type: native-like conformations, high-scoring models reported by VINA and RDOCK, and ligands randomly positioned within binding sites [Colour figure can be viewed at wileyonlinelibrary.com]

to that used in the present study was employed to survey the structures of complexes between proteins and adenine- and guanine-containing ligands in the PDB.[4] Reported values of the vertical distance between ring planes of less than 4.5 Å for parallel and less than 5.5 Å for perpendicular orientations are in line with our analysis of the X-ray crystallography data. Nonetheless, a somewhat broader range of distances observed for the perpendicular configurations in our distance-angle maps arise from the fact that geometric parameters describing aromatic stacking in ligand–protein complexes are derived in the present study from a more diverse collection of compounds containing aromatic groups.

In addition to the experimental structures of ligand–protein complexes, we examine the geometry of aromatic interactions in theoretical models constructed by molecular docking. To the best of our knowledge, scoring functions implemented in modern docking programs treat π–π stacking as van der Waals and Coulombic interactions. This description of aromatic interactions is sufficient in some cases. For instance, docking of a series of agonists to the binding pocket of the homology model of the serotonin 5-$HT_{2C}$ G protein-coupled

**TABLE 2** Deviation of the geometry of aromatic interactions from experimental structures for computer-generated models of TOUGH-D1 complexes

| | Protein–ligand conformations | | | |
|---|---|---|---|---|
| Interaction | Native-like | VINA | RDOCK | Random |
| $F^6$:6 | 0.045 | 0.122 | 0.179 | 1.622 |
| $F^6$:5 | 0.113 | 0.199 | 0.203 | 2.198 |
| $Y^6$:6 | 0.082 | 0.118 | 0.165 | 1.733 |
| $Y^6$:5 | 0.132 | 0.333 | 0.354 | 2.232 |
| $W^5$:6 | 0.274 | 0.359 | 0.753 | 2.325 |
| $W^5$:5 | 0.296 | 1.065 | 0.581 | 3.479 |
| $W^6$:6 | 0.131 | 0.274 | 0.271 | 1.585 |
| $W^6$:5 | 0.418 | 0.628 | 0.862 | 3.049 |
| $H^5$:6 | 0.130 | 0.312 | 0.340 | 1.530 |
| $H^5$:5 | 0.215 | 0.601 | 0.278 | 2.093 |
| Average | 0.184 | 0.401 | 0.399 | 2.185 |

The deviation is measured with the *G* test for native-like conformations, high-scoring models constructed by VINA and RDOCK, and ligands randomly positioned within binding sites.

receptor with the PatchDock server[66] produced ligand conformations interacting with tryptophan and phenylalanine residues through parallel and perpendicular aromatic stacking.[67] Another example is a molecular modeling study conducted with Glide[68] for subnanomolar affinity antagonists of the cannabinoid receptor $CB_2$.[69] Docking calculations revealed a number of parallel and perpendicular aromatic contacts between chloromethylphenyl and methylbenzyl rings of the compounds and a cluster of aromatic residues in $CB_2$, suggesting that π–π interactions are critical to the efficacy of these antagonists.

Postprocessing of docking models is a common practice to improve the prediction accuracy of molecular docking. For instance, molecular mechanics with continuum solvation offers a rigorous approach to decompose free energy into individual contributions from various interaction groups enhancing the screening and ranking power of autodock.[70,71] In addition, it has been reported that re-ranking guanosine triphosphate (GTP) docking poses generated by the gold[72] program by explicitly accounting for the π–π stacking yields a higher accuracy of the modeling of GTP–protein complexes compared to the default goldscore scoring function.[4] Our assessment of docking models constructed by autodock vina[38] and rdock[39] for a diverse collection of ligands and proteins included in the TOUGH-D1 dataset indicates that although state-of-the-art molecular docking force fields provide sufficient specificity to reliably model aromatic interactions, the geometry of π–π contacts in high-scoring conformations could still be improved. The comprehensive analysis of aromatic geometries at ligand–protein interfaces presented in this study lies the foundation for the development of type-specific statistical potentials to more accurately treat π–π interactions in molecular docking. A Perl script to detect and calculate the geometric parameters of aromatic interactions in ligand–protein complexes is available at https://github.com/michal-brylinski/earomatic. The TOUGH-D1 dataset comprising experimental structures and computer-generated models of ligand–protein complexes is available at https://osf.io/rztha/.

## ACKNOWLEDGEMENTS

The author is grateful to Louisiana State University for providing computing resources. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM119524. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## CONFLICT OF INTEREST

None.

## ENDNOTE

[1] (http://www.perlmol.org/) PerlMol. *PerlMol*.

## ORCID

*Michal Brylinski* http://orcid.org/0000-0002-6204-2869

## REFERENCES

[1] S. K. Panigrahi, G. R. Desiraju, *Proteins* **2007**, *67*, 128.

[2] M. A. Williams, J. E. Ladbury, in Hydrogen Bonds in Protein–Ligand Complexes (Eds: H.-J. Bohm, G. Schneider), FRG: Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim **2003**, pp. 137–161.

[3] D. D. Boehr, A. R. Farley, G. D. Wright, J. R. Cox, *Chem. Biol.* **2002**, *9*, 1209.

[4] T. V. Pyrkov, D. V. Pyrkova, E. D. Balitskaya, R. G. Efremov, *Acta Nat.* **2009**, *1*, 124.

[5] N. S. Scrutton, A. R. Raine, *Biochem. J.* **1996**, *319*, 1.

[6] R. Wu, T. B. McMahon, *J. Am. Chem. Soc.* **2008**, *130*, 12554.

[7] R. G. Efremov, A. O. Chugunov, T. V. Pyrkov, J. P. Priestle, A. S. Arseniev, E. Jacoby, *Curr. Med. Chem.* **2007**, *14*, 393.

[8] A. M. Gallina, P. Bork, D. Bordo, *J. Mol. Recognit.* **2014**, *27*, 65.

[9] Y. Lu, Y. Wang, W. Zhu, *Phys. Chem. Chem. Phys.* **2010**, *12*, 4543.

[10] P. Zhou, F. Tian, J. Zou, Z. Shang, *Mini Rev. Med. Chem.* **2010**, *10*, 309.

[11] P. Kukic, J. E. Nielsen, *Future Med. Chem.* **2010**, *2*, 647.

[12] M. T. Neves-Petersen, S. B. Petersen, *Biotechnol. Annu. Rev.* **2003**, *9*, 315.

[13] P. Zhou, J. Huang, F. Tian, *Curr. Med. Chem.* **2012**, *19*, 226.

[14] K. Chen, L. Kurgan, *PLoS ONE* **2009**, *4*, e4473.

[15] Z. Liu, G. Wang, Z. Li, R. Wang, *J. Chem. Theory Comput.* **2008**, *4*, 1959.

[16] A. E. Cho, V. Guallar, B. J. Berne, R. Friesner, *J. Comput. Chem.* **2005**, *26*, 915.

[17] E. Cauet, M. Rooman, R. Wintjens, J. Lievin, C. Biot, *J. Chem. Theory Comput.* **2005**, *1*, 472.

[18] H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook, C. Zardecki, *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 899.

[19] J. B. O. Mitchell, R. A. Laskowski, A. Alex, J. M. Thornton, *J. Comput. Chem.* **1999**, *20*, 1165.

[20] W. T. Mooij, M. L. Verdonk, *Proteins* **2005**, *61*, 272.

[21] H. Fan, D. Schneidman-Duhovny, J. J. Irwin, G. Dong, B. K. Shoichet, A. Sali, *J. Chem. Inf. Model.* **2011**, *51*, 3078.

[22] Y. Fang, Y. Ding, W. P. Feinstein, D. M. Koppelman, J. Moreno, M. Jarrell, J. Ramanujam, M. Brylinski, *PLoS ONE* **2016**, *11*, e0158898.

[23] Y. Ding, Y. Fang, W. P. Feinstein, J. Ramanujam, D. M. Koppelman, J. Moreno, M. Brylinski, M. Jarrell, *J. Comput. Chem.* **2015**, *36*, 2013.

[24] P. Auffinger, F. A. Hays, E. Westhof, P. S. Ho, *Proc. Natl Acad. Sci. USA* **2004**, *101*, 16789.

[25] S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme, M. Schroeder, *Nucleic Acids Res.* **2015**, *43*, W443.

[26] K. C. Janda, J. C. Hemminger, J. S. Winn, S. E. Novick, S. J. Harris, W. Klemperer, *J. Chem. Phys.* **1975**, *63*, 1419.

[27] R. L. Jaffe, G. D. Smith, *J. Chem. Phys.* **1996**, *105*, 2780.

[28] M. O. Sinnokrot, E. F. Valeev, C. D. Sherrill, *J. Am. Chem. Soc.* **2002**, *124*, 10887.

[29] T. Dahl, *Acta Chem. Scand.* **1994**, *48*, 95.

[30] C. A. Hunter, J. Singh, J. M. Thornton, *J. Mol. Biol.* **1991**, *218*, 837.

[31] A. R. Leach, D. P. Dolata, K. Prout, *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 316.

[32] J. Figueras, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 986.

[33] V. Sobolev, A. Sorokine, J. Prilusky, E. E. Abola, M. Edelman, *Bioinformatics* **1999**, *15*, 327.

[34] M. Brylinski, W. P. Feinstein, *J. Comput. Aided Mol. Des.* **2013**, *27*, 551.

[35] R. Guha, M. T. Howard, G. R. Hutchison, P. Murray-Rust, H. Rzepa, C. Steinbeck, J. Wegner, E. L. Willighagen, *J. Chem. Inf. Model.* **2006**, *46*, 991.

[36] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, *Nucleic Acids Res.* **2006**, *34*, D668.

[37] L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, *Bioinformatics* **2012**, *28*, 3150.

[38] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455.

[39] S. Ruiz-Carmona, D. Alvarez-Garcia, N. Foloppe, A. B. Garmendia-Doval, S. Juhos, P. Schmidtke, X. Barril, R. E. Hubbard, S. D. Morley, *PLoS Comput. Biol.* **2014**, *10*, e1003571.

[40] Z. Wang, H. Sun, X. Yao, D. Li, L. Xu, Y. Li, S. Tian, T. Hou, *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964.

[41] Y. C. Chen, *Trends Pharmacol. Sci.* **2015**, *36*, 78.

[42] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785.

[43] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, *J. Cheminform.* **2011**, *3*, 33.

[44] W. P. Feinstein, M. Brylinski, *J. Cheminform.* **2015**, *7*, 18.

[45] Y. Ding, Y. Fang, J. Moreno, J. Ramanujam, M. Jarrell, M. Brylinski, *Comput. Biol. Chem.* **2016**, *64*, 403.

[46] E. Parzen, *Ann. Math. Stat.* **1962**, *33*, 1065.

[47] M. Rosenblatt, *Ann. Math. Stat.* **1956**, *27*, 832.

[48] P. Pavlidis, W. S. Noble, *Bioinformatics* **2003**, *19*, 295.

[49] J. H. McDonald, G-Test of Goodness-of-Fit, Sparky House Publishing, Baltimore, MD **2014**, pp. 53–58.

[50] R. R. Sokal, F. J. Rohlf, Biometry: The Principles and Practice of Statistics in Biological Research, 2nd ed., Freeman, New York **1981**.

[51] M. H. Olsson, C. R. Sondergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 525.

[52] C. R. Sondergaard, M. H. Olsson, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 2284.

[53] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 27.

[54] R. Loewenthal, J. Sancho, A. R. Fersht, *Biochemistry* **1991**, *30*, 6775.

[55] M. Goto, R. Omi, I. Miyahara, A. Hosono, H. Mizuguchi, H. Hayashi, H. Kagamiyama, K. Hirotsu, *J. Biol. Chem.* **2004**, *279*, 16518.

[56] L. Zou, Y. Song, C. Wang, J. Sun, L. Wang, B. Cheng, J. Fan, *Acta Crystallogr. F Struct. Biol. Commun.* **2016**, *72*, 165.

[57] J. Zheng, Z. Jia, *Nature* **2010**, *465*, 961.

[58] T. Fujishiro, J. Kahnt, U. Ermler, S. Shima, *Nat. Commun.* **2015**, *6*, 6895.

[59] E. W. Debler, R. Muller, D. Hilvert, I. A. Wilson, *Proc. Natl Acad. Sci. USA* **2009**, *106*, 18539.

[60] J. H. Toney, P. M. Fitzgerald, N. Grover-Sharma, S. H. Olson, W. J. May, J. G. Sundelof, D. E. Vanderwall, K. A. Cleary, S. K. Grant, J. K. Wu, J. W. Kozarich, D. L. Pompliano, G. G. Hammond, *Chem. Biol.* **1998**, *5*, 185.

[61] P. Hobza, H. L. Selzle, E. W. Schlag, *J. Phys. Chem.* **1996**, *100*, 18790.

[62] E. Arunan, H. S. Gutowsky, *J. Chem. Phys.* **1993**, *98*, 4294.

[63] P. Hobza, H. L. Selzle, E. W. Schlag, *J. Am. Chem. Soc.* **1994**, *116*, 3500.

[64] T. Smith, L. V. Slipchenko, M. S. Gordon, *J. Phys. Chem. A* **2008**, *112*, 5286.

[65] L. Mao, Y. Wang, Y. Liu, X. Hu, *J. Mol. Biol.* **2004**, *336*, 787.

[66] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H. J. Wolfson, *Nucleic Acids Res.* **2005**, *33*, W363.

[67] T. Cordova-Sintjago, N. Villa, L. Fang, R. G. Booth, *Mol. Phys.* **2014**, *112*, 398.

[68] M. P. Repasky, M. Shelley, R. A. Friesner, *Curr. Protoc. Bioinform.* **2007**, Chapter 8: Unit 8 12.

[69] E. Kotsikorou, F. Navas, 3rd, M. J. Roche, A. F. Gilliam, B. F. Thomas, H. H. Seltzman, P. Kumar, Z. H. Song, D. P. Hurst, D. L. Lynch, P. H. Reggio, *J. Med. Chem.* **2013**, *56*, 6593.

[70] H. Sun, Y. Li, M. Shen, S. Tian, L. Xu, P. Pan, Y. Guan, T. Hou, *Phys. Chem. Chem. Phys.* **2014**, *16*, 22035.

[71] H. Sun, Y. Li, S. Tian, L. Xu, T. Hou, *Phys. Chem. Chem. Phys.* **2014**, *16*, 16719.

[72] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727.

**How to cite this article:** Brylinski M. Aromatic interactions at the ligand–protein interface: Implications for the development of docking scoring functions. *Chem Biol Drug Des*. 2018;91:380–390. https://doi.org/10.1111/cbdd.13084