

Article

An Assessment of Dispersion-Corrected DFT Methods for Modeling Nonbonded Interactions in Protein Kinase Inhibitor Complexes

Yan Zhu ¹, Saad Alqahtani ^{1,2} and Xiche Hu ^{1,*}

¹ Department of Chemistry and Biochemistry, University of Toledo, Toledo, OH 43606, USA; yan.zhu2@rockets.utoledo.edu (Y.Z.); salqahtani2@ksu.edu.sa (S.A.)

² Department of Chemistry, King Saud University, Riyadh 12372, Saudi Arabia

* Correspondence: xiche.hu@utoledo.edu; Tel.: +1-4195301513

Abstract: Accurate modeling of nonbonded interactions between protein kinases and their small molecule inhibitors is essential for structure-based drug design. Quantum chemical methods such as density functional theory (DFT) hold significant promise for quantifying the strengths of these key protein–ligand interactions. However, the accuracy of DFT methods can vary substantially depending on the choice of exchange–correlation functionals and associated basis sets. In this study, a comprehensive benchmarking of nine widely used DFT methods was carried out to identify an optimal approach for quantitative modeling of nonbonded interactions, balancing both accuracy and computational efficiency. From a database of 2139 kinase-inhibitor crystal structures, a diverse library of 49 nonbonded interaction motifs was extracted, encompassing CH– π , π – π stacking, cation– π , hydrogen bonding, and salt bridge interactions. The strengths of nonbonded interaction energies for all 49 motifs were calculated at the advanced CCSD(T)/CBS level of theory, which serve as references for a systematic benchmarking of BLYP, TPSS, B97, ω B97X, B3LYP, M062X, PW6B95, B2PLYP, and PWPB95 functionals with D3BJ dispersion correction alongside def2-SVP, def2-TZVP, and def2-QZVP basis sets. The RI, RIJK, and RIJCOSX approximations were used for selected functionals. It was found that the B3LYP/def2-TZVP and RIJK RI-B2PLYP/def2-QZVP methods delivered the best combination of accuracy and computational efficiency, making them well-suited for efficient modeling of nonbonded interactions responsible for molecular recognition of protein kinase inhibitors in their targets.

Keywords: dispersion-corrected DFTs; benchmarking; protein kinase inhibitors; motifs of nonbonded interactions; CCSD(T)/CBS reference; molecular recognition



Citation: Zhu, Y.; Alqahtani, S.; Hu, X. An Assessment of Dispersion-Corrected DFT Methods for Modeling Nonbonded Interactions in Protein Kinase Inhibitor Complexes. *Molecules* **2024**, *29*, 304. <https://doi.org/10.3390/molecules29020304>

Academic Editors: Filipe Menezes and Grzegorz Popowicz

Received: 31 October 2023

Revised: 30 December 2023

Accepted: 3 January 2024

Published: 6 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Protein kinases are a large family of enzymes that play a central role in cellular regulation by catalyzing the transfer of the γ -phosphate group from ATP to their target substrates. This process, known as protein phosphorylation, is a fundamental regulatory mechanism for numerous cellular functions, including metabolism, cell growth and division, differentiation, apoptosis, and gene expression [1]. Dysregulation of protein kinase activity has been implicated in many human diseases, such as cancer, inflammation, diabetes, nervous system disorders, and cardiovascular disease [2]. Cancer cells, in particular, rely on protein phosphorylation by specific kinases to promote abnormal proliferation, metastasis, angiogenesis, and survival. Over 400 diseases are thought to be linked to protein kinases, either directly or indirectly [3]. This has spurred considerable interest in developing small molecule protein kinase inhibitors (PKIs) capable of modulating kinase function for therapeutic benefit [4]. The majority of PKIs are ATP binding site competitive inhibitors, which are the primary targets of this research.

Ever since the first approval of the PKI drug imatinib by the FDA for the treatment of chronic myeloid leukemia in 2001, ATP binding site competitive PKIs have emerged as one of the most significant therapeutic agents in the 21st century for treating a variety of diseases, including cancer, heart disease, and diabetes [4]. However, many challenges remain in PKI drug discovery, such as drug resistance and off-target-mediated toxicity [2,5,6]. The key to overcoming these challenges lies in the targeted molecular design of more potent and selective PKIs. To do so, an atomic level understanding of the molecular recognition between PKIs and their specific kinase targets is required [6].

It is nonbonded interactions that mediate molecular recognition between PKIs and their target protein kinases, as in all ligand–protein complexes [7,8]. Traditionally, the consideration of nonbonded interactions mainly included hydrogen bonding and salt bridge interactions. However, in recent years, more and more evidence suggests that π -moiety involved interactions, such as π – π stacking interactions [9], CH– π interactions [10], cation– π interactions [11], are just as important as hydrogen bonding and salt bridges [6,12,13]. As is established in Ref. [12], hereinafter, all these π -moiety involved interactions will be collectively termed “nonbonded π -interactions”.

To better understand molecular recognition of PKIs in the binding pockets of protein kinases, we are establishing a library of 3D binding motifs between PKIs and their interacting residues inside protein kinases. In a previous study [6], we carried out a large-scale data mining of the Protein Data Bank (PDB), which resulted in the establishment of a database of 2139 non-redundant high-resolution X-ray crystal structures of PKIs bound to protein kinases. As you will see in the Results and Discussion section, care is taken to ensure that the library of 3D binding motifs samples are representative of all major modes of nonbonded interactions occurring in the database of 2139 PKI bound protein kinases, including hydrogen bonding, salt bridge, CH– π interaction, π – π stacking interactions, and cation– π interaction.

Subsequently, intermolecular interaction energies for all motifs in the library were calculated at the highest level of electronic structure theory currently available, i.e., the coupled cluster methods with single, double, and perturbative triple excitations CCSD(T) at complete basis set (CBS) [14]. Besides gaining mechanistic insights on molecular recognition of PKIs in protein kinases, this work has two practical implications: (1) The library of 3D motifs based on real life PKI-bound complexes will provide medicinal chemists with working structure motifs that can guide future design of next generation PKIs; and (2) The magnitudes of intermolecular interaction energies for all motifs of nonbonded intermolecular interactions calculated at the advanced CCSD(T)/CBS level of theory will serve as references for theoreticians to benchmark existing or newly developed methods of intermolecular interaction energy calculations. The latter is exactly what we carried out in the second half of this article: benchmarking of the widely applied density functional theory (DFT) methods against CCSD(T)/CBS.

Nonbonded interactions are essentially a juxtaposition of several elements, including electrostatic interactions, exchange repulsion interactions, induction, and dispersion forces. Of these, dispersion forces constitute the dominant attractive forces between neutral molecules [15,16]. Dispersion forces arise from the mutual correlation of electrons that belong to interacting monomers (intermolecular correlation effects). For a proper treatment of dispersions, post-Hartree–Fock methods that include electron correlation corrections are needed. Therefore, for the study of biological macromolecular systems, applicable wavefunction-based post-Hartree–Fock methods are quite limited due to their large system size. Over the years, second-order Møller–Plesset perturbation (MP2) has been the working horse for calculations of nonbonded interactions because of its efficiency and convenience [17]. The highly accurate CCSD(T) method has been widely applied to small molecular systems. However, its applicability to large biological systems is severely limited due to the extremely high demand of the CCSD(T) method for both CPU time and core memory until recently.

The DFT method is an alternative approach that provides an unrivaled practical balance of accuracy and computational cost. Although the DFT method has been widely applied to study covalently bonded molecules, its application to nonbonded intermolecular interactions is limited due to its inability to capture long-range correlation effects [18]. As a result, typical pure DFT functionals do not take dispersion interactions into consideration. Fortunately, a new generation of DFT methods have been developed with dispersion correction, making it possible to apply DFT methods for analyzing biological systems that are interacting mainly through dispersion interactions [16]. However, the accuracies of dispersion corrected DFTs varied widely. In this work, we performed a systematic benchmarking of nine widely used exchange–correlation functionals, BLYP [19,20], TPSS [21], B97 [22], ω B97X [23], B3LYP [24,25], M062X [26], PW6B95 [27], B2PLYP [28] and PWPB95 [29]. Those functionals span the entire spectrum of the Jacob’s ladder, including two (meta-)GGAs, five (meta-)hybrid functionals, and two double-hybrid functionals. With the inclusion of the D3 level dispersion correction developed by Grimme [30], their performance is calibrated against the ‘gold standard’ CCSD(T) method for a diverse set of 49 motifs of nonbonded interactions extracted from X-ray structures of kinase–inhibitor complexes. Here, we aim at determining the best performing dispersion-corrected DFT method in terms of not only the accuracy but also the computational efficiency. Our objective is to arrive at a recommendation on the best performing DFT protocol that is applicable to routine modeling of protein–ligand binding in kinases.

Another important consideration for the theoretical treatment of nonbonded interactions is the judicious selection of a basis set that effectively balances accuracy and efficiency. Two widely used series of basis sets in electronic structure methods, which take into account electron correlations, are the def2 series developed by Ahlrichs and Weigend [31] and Dunning’s correlation-consistent family of basis sets (cc-pVXZ) [32]. The CC-pVXZ basis sets, known for their correlation consistent design, were selected for the CCSD(T) calculations. The def2 series of basis sets were employed in the DFT methods for their balanced accuracy and computational efficiency. The latter is achieved through density fitting, also known as resolution of the identity (RI) or fitting auxiliary basis sets. Density fitting has been seamlessly incorporated into calculations that employ the def2 series of basis sets in the ORCA program [33]. In this work, def2-SVP, def2-TZVP, and def2-QZVP basis sets were tested in all DFT methods for a systematical study of basis set dependence.

Historically, benchmarking of theoretical approaches for nonbonded interaction energy calculations has been based on prototypical model systems of nonbonded interactions [34,35]. The latter are typically at equilibrium geometries or artificially scaled distances between the nonbonded interacting pairs [36]. This is problematic because, for a given theoretical method, the accuracy of the calculated interaction energy is known to be sensitively geometry (distance) dependent. In contrast, the 3D motifs of nonbonded interactions in our study were obtained directly from actual real life PKI-bound protein complexes. Thus, it is our expectation that the winner of this benchmarking study (i.e., the top performing DFT approach) will be better suited for molecular modeling of PKI binding with protein kinases in any future structure-based drug design of potent and selective PKIs.

As noted earlier, the CCSD(T) interaction energies not only serve as a benchmark for this study but also offer a valuable resource for future development and validation of methods for calculating intermolecular interactions. Moreover, the established library of 3D nonbonded interaction motifs represents a significant contribution in its own right, providing a valuable dataset for future studies exploring protein–ligand recognition and beyond.

2. Results and Discussion

2.1. Library of 3D Motifs of Nonbonded Interactions

In order to construct a comprehensive library of three-dimensional (3D) binding motifs to model the interactions between PKIs and protein kinases, we meticulously selected representative motifs of pairwise nonbonded interactions between PKIs and their interacting residues inside the ATP binding pocket of protein kinases. In a previous study, we

carried out a large-scale data mining of the PDB, which resulted in the establishment of a database of 2139 non-redundant high-resolution X-ray crystal structures of kinase–inhibitor complexes [6]. Based on this database, 49 unique motifs of nonbonded interaction pairs between PKIs and their interacting residues were extracted. These represent key modes of nonbonded interactions, including 13 CH– π interactions, 12 π – π stacking interactions, 8 cation– π interactions, 8 hydrogen bonding, and 8 salt bridge interactions. The selection of these motifs was governed by two main criteria. The first criterion prioritizes protein kinases that are the targets of disease treatment. For example, multiple cyclin-dependent kinases (CDKs) were included, owing to their crucial role in cell cycle regulation and their status as high-priority targets for kinase inhibitor development. The second criterion ensures that these 49 motifs comprehensively cover the diverse spectrum of possible intermolecular interactions, such as hydrogen bonding, salt bridge, and all non-bonded π -interactions, in proportion to their occurrence frequencies in the database of 2139 PKI-Protein complexes [6].

Table 1 provides a full list of relevant details on all 49 motifs. For easy reference, a motif ID is assigned in column 1. Column 2 defines the mode of nonbonded interactions. Column 3 lists the PDB IDs for the PKI-bound complexes from which the binding motifs are extracted. Column 4 lists the interacting pairs, with PKIs labeled by the three letter ligand IDs of the corresponding PDB file and their interacting residues labeled by the residue IDs. Geometrical features for the interacting pairs are given in Columns 5 and 6. For hydrogen bonding pairs, two sets of H-bond angles and distances are listed since there are dual hydrogen bonds in all motifs except motif 40. For all other modes of nonbonded interactions, the closest atom to atom distance between the PKI and its interacting residue is tabulated. For π – π stacking interactions, the angle is measured between the two π planes of the interacting partners. The last column lists the reference in which the X-ray crystal structure was reported.

Table 1. List of motifs of nonbonded interaction pairs between PKIs and their interacting residues ^a.

No.	Mode of Interaction	PDB ID	Intermolecular Pair	Angle	Distance (Å) ^b	Ref.
1	CH– π	1H1Q	2A6...A31		3.32	[37]
2		4BDB	ODO...L354	-	3.93	[38]
3		2BTS	U32...I10	-	3.55	[39]
4		2XIY	XIY...L174	-	3.57	[40]
5		1UU9	BI3...L88	-	3.81	[41]
6		2IO6	330...I374	-	4.41	[42]
7		4GUE	QCT...L147	-	3.61	[43]
8		3OOG	YTP...A70	-	3.34	-
9		4FKO	20K...I10	-	3.72	-
10		3QL8	X01...L134	-	3.23	-
11		4RC2	3O8...L174	-	3.43	-
12		4RC3	3OA...A56	-	3.27	-
13		3QX4	X4B...L134	-	3.61	-
14	π – π stacking	2V7O	DRN...F90	89.64 ^c	3.77	[44]
15		3I4B	Z48...Y134	6.8	4.31	[45]
16		4DGO	0JW...F113	43.85	3.55	-
17		3QQF	X07...F82	70.14	3.76	-
18		3B2T	M33...Y566	10.09	3.83	[46]
19		3R6X	X84...F82	74.54	3.71	-
20		4WG4	UWB...Y131	5.57	4.08	-
21		1PXN	CK6...F82	16.8	3.62	[47]
22		3ROY	22Z...F82	68.31	3.75	-
23		2C6O	4SP...F82	20.3	4.01	[48]
24		3RAK	03Z...F82	41.29	3.81	[49]
25		4EL9	AFE...F79	19.17	3.48	[50]

Table 1. Cont.

No.	Mode of Interaction	PDB ID	Intermolecular Pair	Angle	Distance (Å) ^b			Ref.
26	Cation- π	1PXL	CK4...K33	-	4.62			[51]
27		1FGI	SU1...K514	-	5.15			[52]
28		1M2Q	MNX...K68	-	4.21			[53]
29		2VGP	AD6...K180	-	5.06			[54]
30		3OWP	2SB...K72	-	3.59			-
31		3SQQ	99Z...K33	-	5.12			[49]
32		3RJC	06Z...K33	-	4.77			[49]
33		3QTW	X3A...K89	-	3.51			[49]
34	H-Bond ^d	3SW4	18K...L83	158.59	166.51	2.88	2.94	-
35		3ZLY	YSO...M146	153.12	164.48	2.97	3.26	[55]
36		3RZB	02Z...L83	143.48	162.20	2.96	3.06	[49]
37		3RPY	27Z...L83	141.82	167.13	2.89	3.24	[49]
38		3R28	XA0...L83	155.87	134.34	2.88	2.98	-
39		3QZI	X72...L83	132.02	148.51	2.98	2.98	-
40		3SXF	BK5...Y131	161.45		3.14		-
41		2R3P	3SC...L83	141.41	168.79	2.82	3.35	[56]
42	Salt bridge	3MA3	01I...K67	-	3.38			[57]
43		4L9I	8PR...D271	-	3.96			[58]
44		1NVQ	UCN...E91	-	3.48			[59]
45		4JIK	1KO...D148	-	3.22			[60]
46		1BX6	BA1...E127	-	5.05			[61]
47		4WG4	UWB...E135	-	4.16			-
48		4MTA	2D2...K67	-	3.67			[62]
49		1E9H	INR...K33	-	3.65			[63]

^a: All ligands (designated by the three-letter ligand IDs) and their interacting residues are extracted from chain A of the corresponding PDB files. ^b: The closest distance between nonhydrogen atoms (C, O, or N) of the PKI and those of its interacting residue. ^c: Angle (in degree) for π - π stacking motifs is measured between the two π planes. ^d: All motifs of H-bonding interactions contain dual H-bonding except motif 40. Thus, two sets of parameters are given.

Three dimensional structures for all 49 motifs of nonbonded interaction pairs between PKIs and their interacting residues are depicted in Figure 1 in the same motif count order as in Table 1. The four letter codes after the motif count represent the PDB IDs for the PKI bound complexes from which the binding motifs are extracted. For each binding motif, the PKI and its interacting residue are labeled by the three-letter ligand ID and the residue ID number of the corresponding PDB file, respectively. Given the large size of PKI molecules, only the functional group of a given PKI that directly interacts with the protein was kept as part of the 3D motif. The coordinates of all the non-hydrogen atoms were extracted from their respective PDB files. Hydrogen atoms were patched to satisfy the valency by means of a geometry optimization calculation at the HF/6-31+G levels using the Gaussian 09 program [64]. It should be cautioned that the X-ray crystal structures from which the motifs were extracted may potentially have been distorted by crystal packing. However, the extent of such distortion is a matter of hot debate, with some seeing it as minor adjustments while others view it as significant.

In a previous study, we established a pharmacophore model for ATP-competitive kinase inhibitors [6]. According to this pharmacophore model, PKIs are a special class of small molecules that feature a scaffold of one or more aromatic rings that are linked with one or more hydrophilic functional groups. The former has the structural role of acting as a scaffold and the functional role of participating in nonbonded π -interactions (i.e., π - π stacking interactions, CH- π interactions, cation- π interactions, and XH- π interactions (XH = NH, SH, OH)) with hydrophobic regions I, II, and III (see Ref. [6] for details), and the adenine region. The latter ensure water solubility and form hydrogen bonds with the hinge region and other hydrophilic residues of the ATP binding pocket. This pharmacophore model, to a large extent, functions as a blueprint to guide the se-

lection of the 49 binding motifs sampled here. As shown in Figure 1, in all 13 motifs of CH- π interactions, aromatic rings of PKIs interact with the CH groups of the aliphatic residues Ala, Leu, and Ile. In the case of π - π stacking interactions, aromatic rings of PKIs interact with aromatic residues Phe, Tyr, and Trp. As can be seen from Table 1, a wide spectrum of π - π stacking angles are sampled, ranging from parallel displaced to T-shaped configuration, and everything in between. In this study, the coordinates for all motifs of nonbonded interactions were directly extracted from the experimentally determined X-ray crystal structures. Inside the protein, the conformation of the interacting pairs is constrained by a multitude of forces. As a results, the extracted motifs of non-bonded interactions are generally not in energetically optimal conformations that are achievable in the gas phase, such as the perfect T-shaped or parallel-displaced configurations for π - π stacking interaction pairs [13]. Motifs of cation- π interaction feature aromatic rings of PKIs interacting with the positively charged side chains of the basic residue Lys. In motifs of hydrogen bonding pairs, the backbone N and O atoms of residues in the hinge region are involved in dual hydrogen bonding with H bond donors and acceptors of PKIs. This mirrors a major hydrogen bonding pattern for molecular recognition of FDA approved small molecule protein kinase inhibitor drugs in protein kinases [12]. For motifs of salt bridges, positively charged basic residue Lys and negatively charged acidic residues (Asp and Glu) in protein kinases interact with oppositely charged functional groups of PKIs, respectively.

CH- π Interactions

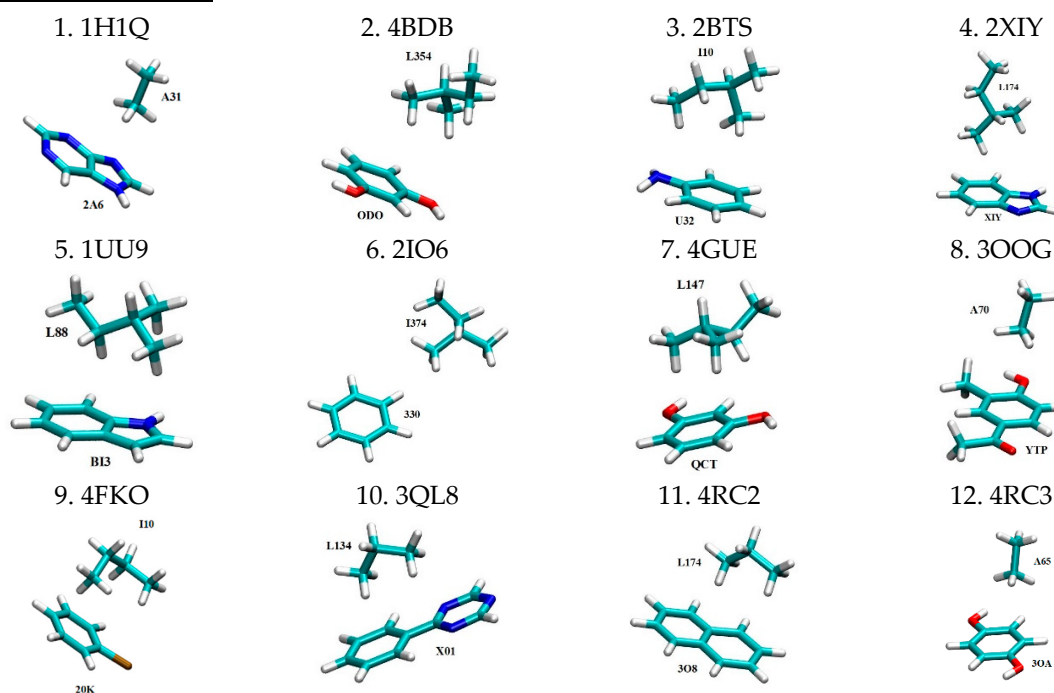


Figure 1. Cont.

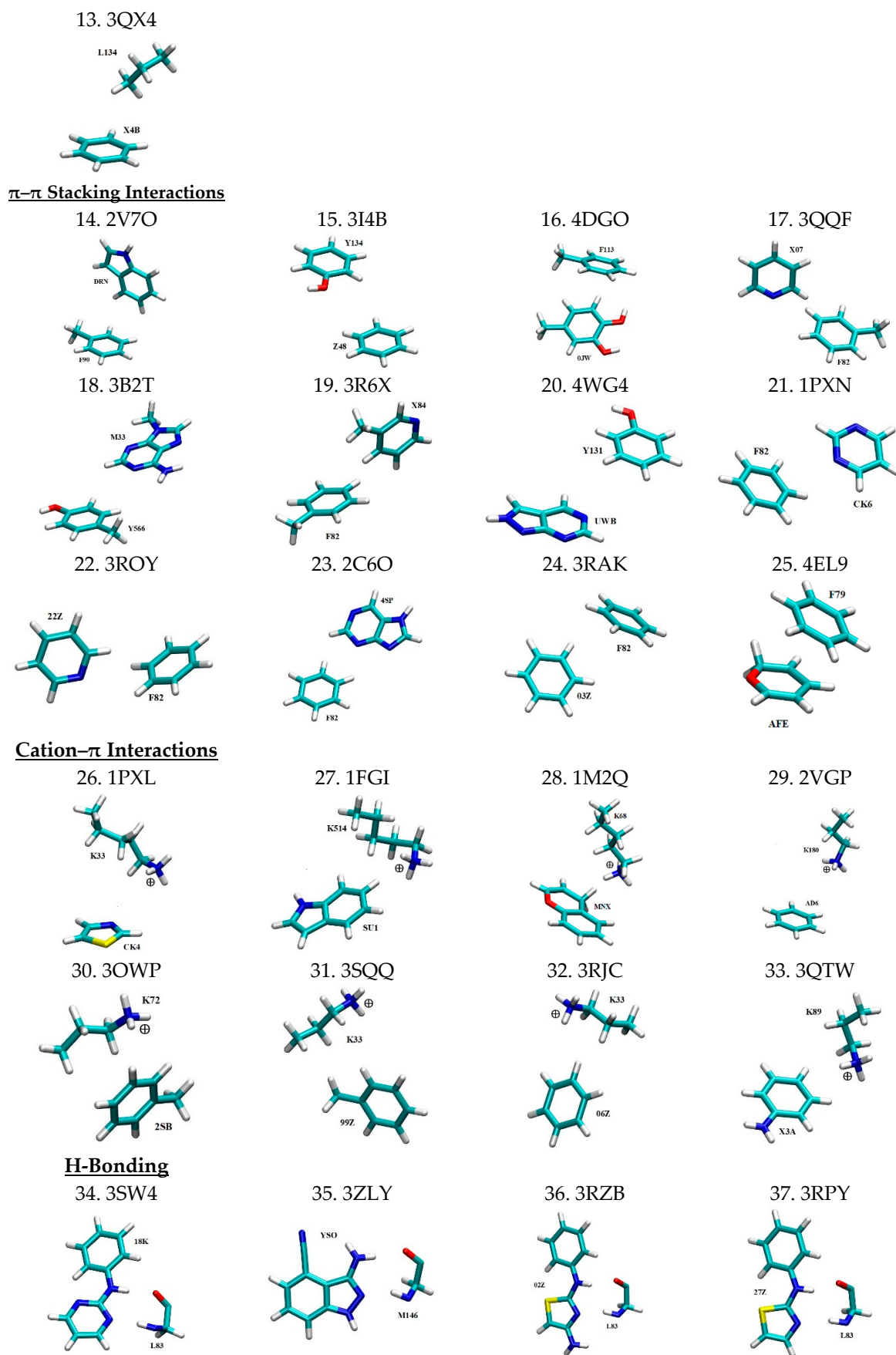


Figure 1. Cont.

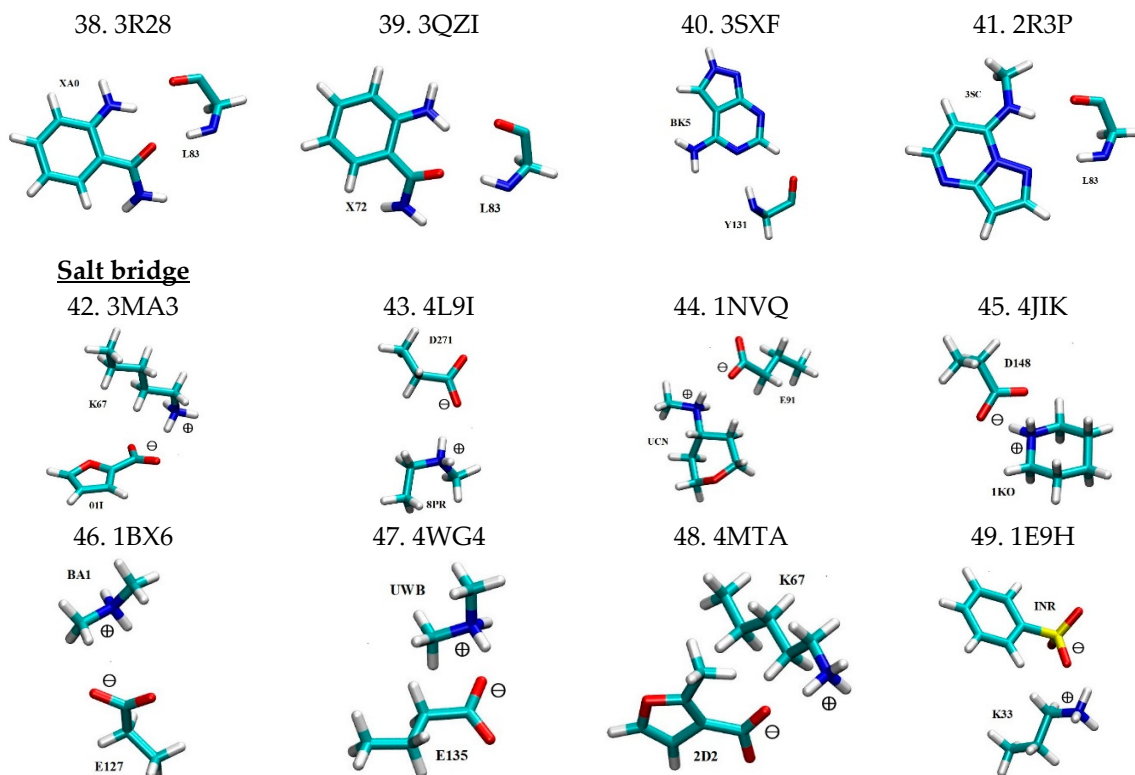


Figure 1. Motifs of nonbonded interactions between PKIs (labeled by the three-letter ligand IDs) and their interacting residues (labeled by the residue IDs) in a licorice representation with carbon, oxygen, nitrogen, sulfur, hydrogen and chlorine atoms colored in cyan, red, blue, yellow, white and brown, respectively. The electric charges of constituting monomers are labeled for motifs of cation- π and salt bridge interactions. All motifs have a spin multiplicity of 1. This figure is produced with the program VMD 1.9.3 [65].

2.2. Energies of Intermolecular Interactions Calculated at the CCSD(T)/CBS Level

The strengths of the intermolecular interaction energies for all 49 motifs were calculated at both the gas phase and the solution phase. The former serves as the reference for benchmarking the DFT methods in the next subsection. The latter aims at a realistic evaluation of the strengths of the intermolecular interactions in the aqueous media where the actual biological interactions occur. Table 2 lists both the gas phase ($\Delta E_{\text{CCSD(T)}}^{\text{g}}$) and the solution phase ($\Delta E_{\text{int}}^{\text{aq}}$) intermolecular interaction energies for all 49 motifs. The gas phase interaction energies were calculated at the CCSD(T) level and extrapolated to the complete basis limit (CBS) via the two-point basis set extrapolation scheme [66,67], with the basis set superposition error (BSSE) corrections (see Section 3 for details). Following Equation (3), the CCSD(T)/CBS level interaction energy was obtained by $\Delta E_{\text{CBS}}^{\text{CCSD(T)}} = \Delta E_{\text{CBS}}^{\text{MP2}} + (\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}})|_{\text{aug-cc-pVDZ}}$. The first term $\Delta E_{\text{CBS}}^{\text{MP2}}$ defines the MP2 level interaction energy with complete basis set, and was evaluated using Equation (4) with the cc-pVTZ and cc-pVQZ basis sets. The second term, commonly referred to as “coupled cluster correction” in the literature [68], was calculated using the aug-cc-pVDZ basis set. Given the demanding scaling of the CCSD(T) method with the system size, the CCSD(T)/aug-cc-pVDZ protocol employed here represents what is computationally feasible currently for the large size motifs studied. It is important to note that achieving convergence for the “coupled cluster correction” term might necessitate a basis set more extensive than aug-cc-pVDZ [68]. In the case of benzene dimer, for example, the “coupled cluster correction” term for π - π stacking interactions evaluated at the CCSD(T)/aug-cc-pVDZ level deviates from that of “best estimate” CCSD(T)/CBS by approximately 0.1 kcal/mol [68]. The solution phase interaction energies were obtained

indirectly by means of a thermodynamic cycle according to Equation (5) (see Theory and Methods): $\Delta E_{\text{int}}^{\text{aq}} = \Delta E_{\text{CCSD(T)}}^{\text{g}} + \Delta E_{\text{Deh}}$. The dehydration energy ΔE_{Deh} itself was calculated utilizing the SM5.42R solvation model of Cramer and Truhlar [69].

Table 2. Intermolecular interaction energies (in kcal/mol) calculated at the CCSD(T)/CBS level for the 3D binding motifs of protein kinase inhibitors in protein kinases.

No.	Mode of Interaction	PDB ID	Intermolecular Pair	$\Delta E_{\text{CCSD(T)}}^{\text{g}}$ (kcal/mol)	ΔE_{Deh} (kcal/mol)	$\Delta E_{\text{int}}^{\text{aq}}$ (kcal/mol) ^a
1	CH- π	1H1Q	2A6...A31	−1.9	0.5	−1.4
2		4BDB	ODO...L354	−2.2	0.4	−1.8
3		2BTS	U32...I10	−2.4	0.8	−1.6
4		2XIY	XIY...L174	−2.3	0.3	−2.0
5		1UU9	BI3...L88	−3.6	0.3	−3.3
6		2IO6	330...I374	−0.7	0.0	−0.7
7		4GUE	QCT...L147	−3.4	0.7	−2.7
8		3OOG	YTP...A70	−1.9	0.7	−1.2
9		4FKO	20K...I10	−2.3	0.0	−2.3
10		3QL8	X01...L134	−2.5	0.0	−2.5
11		4RC2	3O8...L174	−2.1	0.2	−1.9
12		4RC3	3OA...A56	−1.0	−1.0	−2.0
13		3QX4	X4B...L134	−1.0	0.1	−0.9
14	π - π Stacking	2V7O	DRN...F90	−1.3	0.0	−1.3
15		3I4B	Z48...Y134	−1.0	0.3	−0.7
16		4DGO	0JW...F113	−4.4	1.6	−2.8
17		3QQF	X07...F82	−2.4	0.7	−1.7
18		3B2T	M33...Y566	−2.0	0.5	−1.5
19		3R6X	X84...F82	−1.9	0.4	−1.5
20		4WG4	UWB...Y131	−0.6	0.3	−0.3
21		1PXN	CK6...F82	−0.7	0.6	−0.1
22		3ROY	22Z...F82	−2.4	0.6	−1.8
23		2C6O	4SP...F82	−1.1	0.8	−0.3
24		3RAK	03Z...F82	−1.7	0.1	−1.6
25		4EL9	AFE...F79	−1.9	0.4	−1.5
26	Cation- π	1PXL	CK4...K33	−6.5	5.3	−1.2
27		1FGI	SU1...K514	−4.4	2.5	−1.9
28		1M2Q	MNX...K68	−2.0	−0.3	−2.3
29		2VGP	AD6...K180	−1.5	1.0	−0.5
30		3OWP	2SB...K72	−9.2	4.5	−4.7
31		3SQQ	99Z...K33	−2.1	0.8	−1.3
32		3RJC	06Z...K33	−2.1	0.5	−1.6
33		3QTW	X3A...K89	−8.4	8.6	0.2
34	H-Bond	3SW4	18K...L83	−6.0	4.8	−1.2
35		3ZLY	YSO...M146	−4.9	3.7	−1.2
36		3RZB	02Z...L83	−7.0	5.2	−1.8
37		3RPY	27Z...L83	−7.0	4.4	−2.6
38		3R28	XA0...L83	−4.8	4.0	−0.8
39		3QZI	X72...L83	−4.6	0.6	−4.0
40		3SXF	BK5...Y131	−3.8	3.1	−0.7
41		2R3P	3SC...L83	−6.1	3.3	−2.8
42	Salt bridge	3MA3	01I...K67	−111.0	102.6	−8.4
43		4L9I	8PR...D271	−100.3	94.5	−5.8
44		1NVQ	UCN...E91	−100.0	98.6	−1.4
45		4JIK	1KO...D148	−115.9	112.2	−3.7
46		1BX6	BA1...E127	−68.2	67.5	−0.7
47		4WG4	UWB...E135	−92.6	87.5	−5.1
48		4MTA	2D2...K67	−106.6	99.4	−7.2
49		1E9H	INR...K33	−101.5	98.5	−3.0

^a The solution phase interaction energy was calculated via $\Delta E_{\text{int}}^{\text{aq}} = \Delta E_{\text{CCSD(T)}}^{\text{g}} + \Delta E_{\text{Deh}}$ according to Equation (5).

As shown in the fifth column of Table 2, the pairwise interactions between PKIs and their targets are energetically favorable in the gas phase; all the calculated interaction energies are negative (attractive). For example, hydrogen bonding interactions exhibit energies ranging from -7.0 to -3.8 kcal/mol, with an average of -5.5 kcal/mol. The energies for salt bridge interactions span from an extraordinary -115.9 to -68.2 kcal/mol, averaging at -99.5 kcal/mol. Non-bonded π -interactions have a smaller magnitude, with an average gas-phase interaction energies of -2.1 kcal/mol, -1.8 kcal/mol, and -4.5 kcal/mol for CH- π interaction, π - π stacking interactions, and cation- π interactions, respectively.

The last column of Table 2 presents the solution phase interaction energies for all 49 motifs, which sheds some light on the molecular recognition of PKIs in protein kinases. All modes of nonbonded interactions in the 49 motifs display a favorable (negative) interaction energy, making an energetically stabilizing contribution to the binding of PKIs with protein kinases. Although the dispersion forces dominated CH- π interactions and π - π stacking interactions are relatively weak in the gas phase in comparison with polar and charged interactions (hydrogen bonding, salt bridge, and cation- π interactions), there is a much lower (or near zero) energetic cost of dehydration involved. As a result, significant interaction strengths remain in the aqueous phase for both CH- π interactions and π - π stacking interactions. This agrees with our published findings that, in addition to hydrogen bonding, aromatic rings of PKIs function as important molecular determinants for the binding of PKIs with protein kinases. From the perspective of aromatic rings, hydrophobic regions I, II, and III (see Ref. [6] for details), plus the adenine region, are the most relevant. Those four regions are loaded with hydrophobic residues that can participate in nonbonded π -interactions (i.e., π - π stacking interactions, CH- π interactions). The aliphatic residues Ala, Val, Leu, and Ile in the hydrophobic regions of the ATP binding site can form CH- π interactions with aromatic rings of PKIs. Additionally, aromatic residues Phe, Tyr, and Trp can form π - π stacking interactions with aromatic rings of PKIs.

It is worth noting, from the perspective of methodology, the importance of including dehydration energy in any theoretical treatment of nonbonded interactions. For example, one particularly noteworthy observation involves the strength of salt bridge interactions, which feature very strong gas-phase interaction energies of around -100 kcal/mol. While these energies might appear extremely strong, it becomes evident that the dehydration energy ΔE_{Deh} is also substantial, roughly around 100 kcal/mol. Consequently, the average intermolecular interaction energies reduce from -99.5 kcal/mol in the gas phase to -4.4 kcal/mol in the solution phase for salt bridge interactions. The same can be said about other motifs of nonbonded interactions involving polar or charged species. The average dehydration energies are 2.9 and 3.6 kcal/mol for cation- π interactions and hydrogen bonding, respectively. This underlines the necessity of incorporating desolvation energy corrections in any theoretical calculations aiming for a realistic modeling of ligand binding in proteins.

2.3. Benchmarking of DFT Methods

Based on the gas phase interaction energies of 49 motifs of nonbonded interactions calculated at the CCSD(T)/CBS level above, the performance of BLYP, TPSS, B97, ω B97X, B3LYP, M062X, PW6B95, B2PLYP and PWPB95 functionals was benchmarked. As shown in Table 3, these nine exchange-correction functionals span the entire spectrum of the Jacob's ladder, featuring two (meta-)GGAs, five (meta-)hybrid functionals, and two double-hybrid functionals. Given the dual objectives of our performance benchmarking, i.e., accuracy and efficiency (see the Introduction section), different algorithms of implementation for several functionals are also evaluated. For functionals BLYP, TPSS, B97, B2PLYP, and PWPB95, the Resolution of Identity (RI) approximation (also called Density Fitting) [70] is implemented. In addition, the PW6B95 functional is implemented with the RIJCOSX approximation that uses density fitting for the Coulomb (J) integrals and numerical chain-of-sphere integration for the HF Exchange integrals (COSX) [71]. For efficiency, B2PLYP and PWPB95 functionals were also implemented with the resolution of identity (RI) approximation for

the perturbation step and RIJK (RI approximation is applied to both Coulomb and HF Exchange integrals) [72] for the HF-SCF step. Basis set dependence was also studied systematically. For this purpose, the Ahlrichs def2 basis set family [31,73] was employed, including def2-SVP (split valence), def2-TZVP (triple zeta valence), and def2-QZVP (quadruple zeta valence).

Table 3. List of DFT methods.

Functional	Implementation ^a	HF Exchange ^b (%)	Type ^c	Ex. Functional Corr Functional	Ref.
BLYP	RI BLYP	0	GGA	Becke 1988 Lee-Yang-Parr 1988	[19,20]
TPSS	RI TPSS	0	Meta-GGA	TPSS TPSS	[21,74]
B97	RI B97	26.93	hybrid GGA	B97-2 B97-3	[29,75]
ω B97X	ω B97X	15.77	range-separated hybrid-GGA	LRC hybrid functionals	[23]
B3LYP	B3LYP	20	hybrid GGA	Becke 1988 Lee-Yang-Parr 1988	[19,20]
M062X	M062X	54	Meta-hybrid GGA	M06-2X M06-2X	[26]
PW6B95	PW6B95 RIJCOSX PW6B95	28	Meta-hybrid GGA	PW6B95 PW6B95	[27]
B2PLYP	RI B2PLYP RIJK RI B2PLYP	54	Double-Hybrid-GGA	Becke 1988 Lee-Yang-Parr 1988	[19,20] [28]
PWPB95	RI PWPB95 RIJK RI PWPB95	50	Double-Hybrid- Meta GGA	Perdew–Wang32 Becke9533	[76,77] [29,78]

^a: RI denotes the Resolution of Identity (RI) approximation [70]. RIJCOSX is an approximation that uses density fitting for the Coulomb (J) integrals and numerical chain-of-sphere integration for the HF Exchange integrals (COSX) [71]. “RI RIJK” stands for resolution of identity (RI) approximation for the perturbation step and RIJK (RI approximation is applied to both Coulomb and HF Exchange integrals) [72] for the HF-SCF step. ^b: the percentage of HF exchange in the functional. ^c: GGA denotes generalized gradient approximation.

Gas phase interaction energies for all 49 motifs of nonbonded interactions were calculated according to Equation (1) for all nine exchange–correlation functionals listed in Table 3. In order to compensate for the lack of adequate account of dispersion interactions, the D3 level dispersion correction [30] is added for all DFT calculations. BSSE was corrected by the Boys and Bernardi Counter Poise Method [79]. The resulting gas phase interaction energies for each of the nine dispersion-corrected DFT methods, in combination with basis sets def2-SVP, def2-TZVP, and def2-QZVP are tabulated in Tables S1–S3 of the Supplementary Material section, respectively. It is worth noting that, for functionals with various numerical algorithms of implementation, i.e., RI, RIJCOSX, and RI RIJK, there was no noticeable difference in the calculated interaction energies from those calculated with their corresponding original functionals. Hence, interaction energies are presented only using the original functionals in Tables S1–S3. The same applies to all the subsequent data analyses presented in all the tables and figures hereafter. As can be seen from the Supplementary Tables S1–S3, a majority of the DFT methods, with dispersion corrections, yielded interaction energies that are either identical or closely aligned with those obtained from CCSD(T)/CBS calculations in the gas phase. This close agreement emphasizes the potential utility of specific dispersion-corrected DFT methods as accurate yet computationally efficient tools for modeling these critical biological interactions.

To conduct an in-depth assessment of the performance of the various DFT methods and basis sets under consideration, we employed a series of statistical metrics for a comparative

analysis against the benchmark CCSD(T)/CBS calculations. Specifically, we calculated the Root Mean Square Deviation (RMSD), Mean Unsigned Error (MUE), Average Signed Error (AVG), and the Maximum Deviation (MAX). RMSD serves as a widely accepted metric in computational chemistry for performance evaluation, facilitating a meaningful comparison between DFT-derived interaction energies and those ascertained via CCSD(T) calculations. The MUE metrics offer additional insights into the primary discrepancies associated with the DFT methods relative to the benchmark CCSD(T) values.

Table 4 elucidates the errors associated with the nine DFT methods in relation to the benchmark CCSD(T)/CBS calculations. Overall, the MAE values are slightly less than the RMSD values. This is not unexpected since RMSD penalizes large discrepancies more severely while offering an excellent measure of overall accuracy. Conversely, MAE, though less prone to outliers, overlooks the direction of errors (over/underestimation) and does not penalize large errors as heavily as RMSD. The AVG values are even smaller. Although AVG provides a simple measure of the error's central tendency, it is easily skewed by outliers and lacks information about the error distribution. The MAX% values are quite large, which should not be a cause of alarm. That is because %MAX highlights the worst-case scenario for individual predictions. To a large extent, the small magnitude of interaction energies for several modes of nonbonded interactions (see Tables 2 and S1–S3) are responsible for the large discrepancy in MAX%. So, we will focus on the RMSD values for our error analysis. As can be seen from Table 4, the def2-QZVP basis set consistently outperformed its def2-TZVP and def2-SVP counterparts. Notably, the most accurate RMSD value recorded was 0.46 kcal/mol, achieved using the B3LYP functional in combination with the def2-QZVP basis set. The B3LYP/def2-TZVP combination yielded second best RMSD of 0.50 kcal/mol. The third best RMSD value of 0.51 kcal/mol belongs to the double hybrid functional B2PLYP with the def2-QZVP basis set. Interestingly, for a subset of DFT methods—namely ω B97X, TPSS, B97, and PW6B95—the def2-QZVP and def2-TZVP basis sets yielded nearly identical RMSD values. However, it is crucial to point out that these RMSD values were considerably higher than those produced by the B3LYP and B2PLYP methods.

Table 4. Errors of the studied methods for 49 motifs with respect to the benchmark (CCSD(T)/CBS) calculations.

DFT Method	RMSD (kcal/mol)			MAE (kcal/mol)			AVG (kcal/mol)			MAX%		
	def2-QZVP	def2-TZVP	def2-SVP	def2-QZVP	def2-TZVP	def2-SVP	def2-QZVP	def2-TZVP	def2-SVP	def2-QZVP	def2-TZVP	def2-SVP
BLYP	0.70	0.73	1.25	0.48	0.48	0.76	0.15	0.09	−0.28	33.2	34.0	38.1
TPSS	0.76	0.77	0.97	0.52	0.51	0.64	0.23	0.21	−0.09	39.7	41.2	38.5
B97	0.96	0.95	1.11	0.70	0.66	0.77	0.14	0.11	−0.18	50.4	51.3	51.0
ω B97X	0.75	0.77	0.88	0.43	0.43	0.55	0.23	0.20	0.01	43.1	43.5	39.8
B3LYP	0.46	0.50	0.96	0.29	0.32	0.55	0.08	0.05	−0.24	25.9	26.6	26.6
M062X	1.13	1.17	1.14	0.93	0.94	1.00	0.93	0.93	0.82	103.1	106.4	118.6
PW6B95	0.93	0.92	0.90	0.57	0.54	0.62	0.50	0.47	0.23	46.1	47.6	54.4
B2PLYP	0.51	0.66	0.93	0.35	0.45	0.75	0.32	0.45	0.50	26.9	43.4	76.3
PWBP95	0.98	1.07	1.05	0.68	0.74	0.79	0.67	0.74	0.65	51.7	61.9	86.5

RMSD: root mean square deviation. MAE: mean absolute error. AVG: average signed error. MAX%: the largest percentage error in the set relative to the CCSD(T)/CBS interaction energy.

The basis set dependence is further illustrated in Figure 2. As depicted, the performance of most DFT methods is closely tied to the sizes of the basis sets employed. Generally speaking, the larger def2-QZVP basis set delivers superior results when compared to the def2-TZVP basis set, which in turn outperforms the more diminutive def2-SVP basis set across a majority of the DFT methods tested. While there are a few anomalies to this trend, it is worth noting that the disparities in RMSD values calculated at the def2-QZVP and def2-TZVP levels are negligible. This analysis accentuates the importance of carefully selecting the appropriate basis set for electronic structure calculations, especially those focused on delineating complex biological interactions. The demonstrated improvement in performance with increasing basis set size suggests that for studies requiring a high degree

of accuracy, the use of larger basis sets such as def2-QZVP is recommended, although the computational cost must also be considered in the selection process.

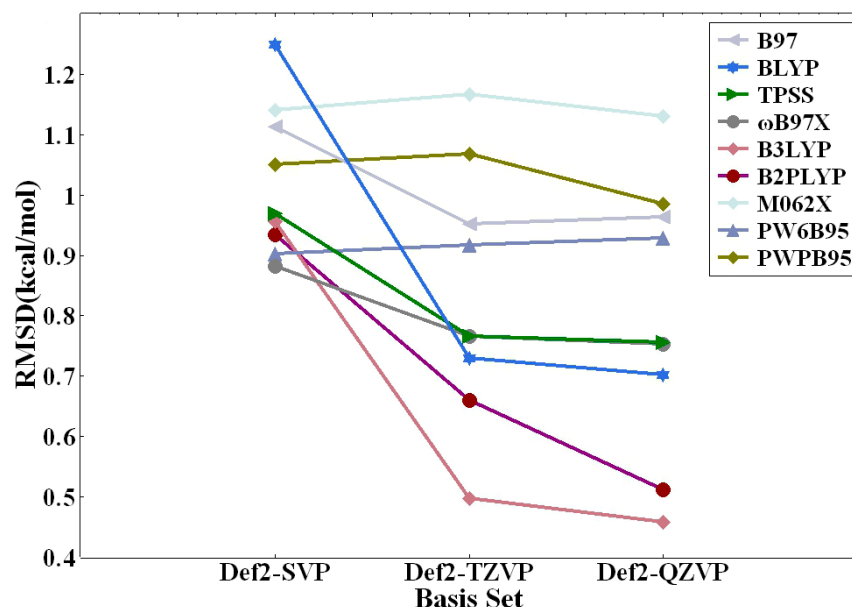


Figure 2. Basis set dependence of RMSD of calculated intermolecular interaction energies (in kcal/mol) for the nine DFT methods against those of CCSD(T)/CBS.

Figure 3 provides a more detailed and distributed analysis of the root-mean-square deviation (RMSD) in calculated interaction energies as a function of modes of nonbonded interactions across nine different DFT methods, each assessed at three distinct basis set levels: def2-SVP, def2-TZVP, and def2-QZVP. Overall, the performance of all nine DFT methods, as judged from the RMSD values, are generally much better for nonbonded π -interactions (CH- π , π - π stacking, and cation- π interactions) than the polar and charged interactions (hydrogen bonding and salt bridge interactions). Notably, Figure 3a unveils an unexpected observation: a remarkably high RMSD for salt bridge interactions when using methods like B3LYP, BLYP, PW6B95, PW6B95 RIJCOSX, RI B97, RI TPSS, and ω B97X, coupled with the def2-SVP basis set. Intriguingly, for the def2-TZVP and def2-QZVP basis sets, the highest RMSD is associated with hydrogen bond interactions. A careful examination of detailed interaction energies as listed in Tables 2 and S1–S3 for hydrogen bonded motifs uncovers two important characteristics of the errors. First, the computed nonbonded interaction energy from all DFT approaches is significantly less negative than the very attractive (negative) values obtained using the CCSD(T)/CBS method. Second, there appears to be a consistent bias (positive difference between DFT and CCSD(T)) across all DFT techniques, as indicated by the strong resemblance between the magnitudes of the mean unsigned error (MUE) and average error (AVG) and the root-mean-square deviation (RMSD).

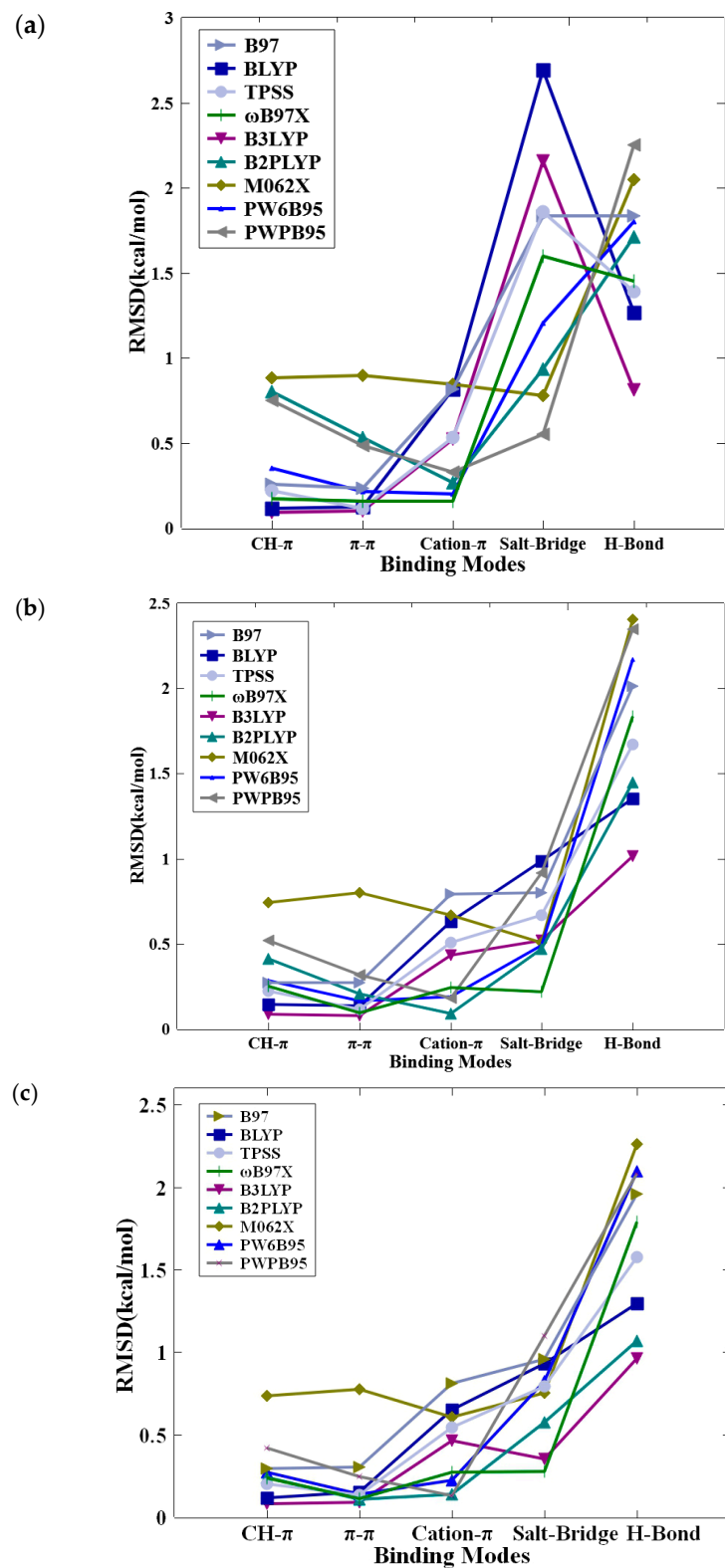


Figure 3. RMSD of calculated intermolecular interaction energies (in kcal/mol) for the nine DFT methods against those of CCSD(T)/CBS as a function of modes of nonbonded interactions. (a) DFT/def2-SVP, (b) DFT/def2-TZVP, (c) DFT/def2-QZVP.

Another dimension of consideration, particularly relevant for large biological systems, is the computational cost associated with DFT calculations. A delicate trade-off exists between methodological accuracy and computational efficiency. The accuracy can be quantified by RMSD values relative to the CCSD(T)/CBS benchmarks. The computational cost is measured by CPU time. For equal configuration of computing hardware, all DFT calculations were carried out on Dell workstations equipped with 2.7 GHz Intel Core i7 6820HQ CPU and 32 GB core memory. The ORCA program was executed in parallel on four processors using the PAL4 command. CPU times are calculated as the average over the entire set of 49 motifs of nonbonded interactions. Figure 4 presents a comparative analysis of various DFT methods, evaluating them on the dual criteria of accuracy and computational cost. The plot maps accuracy against CPU time (in min) for various implementations of the nine exchange-correlation functionals (see Table 2) coupled to def2-TZVP and def2-QZVP basis sets. Lower RMSD values signify greater accuracy and short CPU times indicate less computational expense. Consequently, the optimal DFT method would ideally be situated in the lower-left quadrant of the plot, representing both high accuracy and low computational cost. Based on these considerations, the B3LYP/def2-TZVP combination, with an RMSD of 0.50 kcal/mol and an average CPU time of 27.7 min, aligns closest with the criteria of achieving the lowest RMSD while also requiring the least CPU time, making it the top performing choice under the metrics considered. For applications that require the highest accuracy, B3LYP/def2-QZVP offers the best RMSD value of 0.46 kcal/mol, but with an extremely high average CPU cost of 479.8 min. The next best performing combination is RIJK RI-B2PLYP/def2-QZVP with an RMSD of 0.51 kcal/mol and an averaged CPU time of 33.2 min. For applications that can tolerate a slight loss of accuracy, the RIJK RI-B2PLYP/def2-TZVP combination (with a RMSD 0.66 kcal/mol) has the advantage of consuming nearly one fourth of the CPU time (7.7/33.2) required by RIJK RI-B2PLYP/def2-QZVP.

As pointed out earlier, for functionals with various numerical algorithms of implementation, i.e., RI, RIJCOSX, and RI RIJK, there was no noticeable difference in accuracy. Remarkably, large differences in CPU time were observed. For example, the averaged CPU time for RIJK RI B2PLYP/def2-QZVP and RI B2PLYP/def2-QZVP are 33.2 min and 444.2 min, respectively. This represents a CPU time reduction of 13.4 fold, which can be a favorable factor of consideration for large biological systems like the PKI-bound protein kinases.

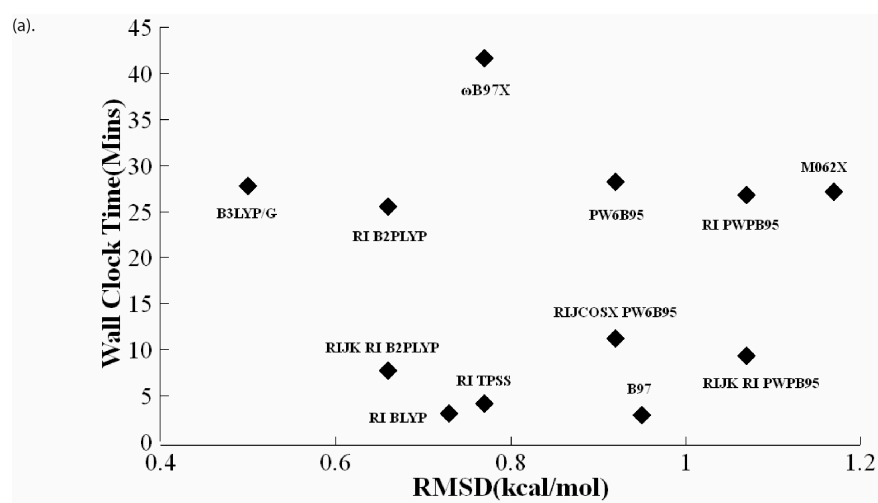


Figure 4. Cont.

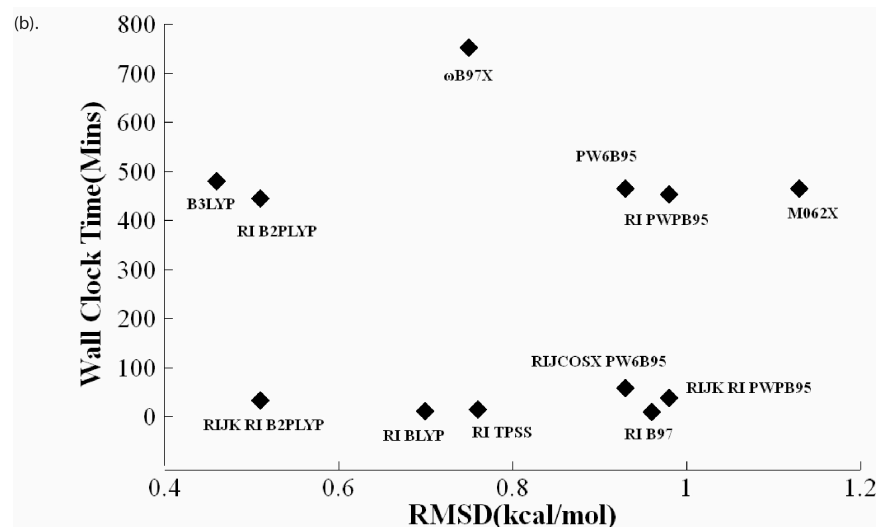


Figure 4. Performance of nine DFT methods and their variants as applied to the 49 binding motifs of PKIs in protein kinases. The average RMSD in interaction energies is calculated with respect to the CCSD(T)/CBS method. For each DFT method, the calculation times represent the averaged time over all 49 motifs. (a) DFT/def2-TZVP and (b) DFT/def2-QZVP.

3. Theory and Methods

Quantum Chemical Calculation of Intermolecular Interaction Energies

Gas phase intermolecular interaction energy. The energies of intermolecular interactions in the gas phase were calculated by means of the supramolecular approach. In the supramolecular approach, the energy of interaction between molecules *A* and *B* is defined as the difference between the energy of the interacting dimer and the energies of the monomers:

$$\Delta E_{int}^g = E_{AB} - E_A - E_B \quad (1)$$

Nine DFT methods are employed to calculate the gas phase intermolecular interaction energies according to Equation (1) for benchmarking purpose: BLYP [19,20], TPSS [21], B97 [22], ωB97X [23], B3LYP [24,25], M062X [26], PW6B95 [27], B2PLYP [28] and PWPB95 [29]. For each of the DFT methods, three basis sets are applied: def2-QZVP, def2-TZVP, and def2-SVP [31,73]. It is well known that traditional pure DFT functionals lack an account of dispersion interactions, and thus are inadequate for the treatment of dispersion force dominated nonbonded interactions such as CH–π interactions and π–π stacking interactions. To overcome this problem, the D3 level dispersion correction developed by Grimme [30] is applied. Specifically, the atom-pairwise dispersion correction with the Becke–Johnson damping scheme (D3BJ) [80] is adopted. Here, the dispersion correction is calculated as follows by adding over all atom pairs in the systems:

$$E_{disp}^{D3(BJ)} = -\frac{1}{2} \sum_{A \neq B} s_6 \frac{C_6^{AB}}{R_{AB}^6 + [f(R_{AB}^0)]^6} + s_8 \frac{C_8^{AB}}{R_{AB}^8 + [f(R_{AB}^0)]^8} \quad (2)$$

where C_6^{AB} and C_8^{AB} are the averaged dispersion coefficients for the interacting atoms *A* and *B* at sixth and eighth order. R_{AB} represents internuclear distance between atom *A* and atom *B* and S_n ($n = 6, 8$) are global scaling factors that are used to adjust the correction to the repulsive behavior of the used DFT method [81].

In practical calculations, the coordinates of all atoms are taken directly from X-ray crystal structures of the PKI-bound protein kinases. The positions of hydrogen atoms are patched through geometry optimization calculations at the HF/6-31+G level using the Gaussian 09 program [64]. All the DFT calculations are carried out using the ORCA 04 [33] program. Last but not least, the basis set superposition error (BSSE) was corrected by the Boys and Bernardi Counter Poise Method [79].

The CCSD(T)/CBS Method. To establish a benchmarking reference, the CCSD(T) method at the complete basis set is employed to calculate the gas phase intermolecular interaction energy according to Equation (1). As mentioned in Section 2.1, the CCSD(T)/CBS Method is widely regarded as the gold standard for treatment of intermolecular interaction energies. Due to the high computational cost of CCSD(T) calculation, the complete basis set (CBS) limit is achieved by means of an extrapolation scheme [14]. In that scheme, the CCSD(T)/CBS level interaction energy is given by:

$$\Delta E_{\text{CBS}}^{\text{CCSD(T)}} = \Delta E_{\text{CBS}}^{\text{MP2}} + (\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}}) \Big|_{\text{small basis set}} \quad (3)$$

where $\Delta E_{\text{CBS}}^{\text{MP2}}$ is the MP2 level interaction energy with complete basis set. It is calculated using the two-point basis set extrapolation scheme developed by Halkier and co-workers [66,67]:

$$E_{\text{corr,lim}} = \frac{X^3}{X^3 - (X-1)^3} E_{\text{corr,X}} - \frac{(X-1)^3}{X^3 - (X-1)^3} E_{\text{corr,X-1}} \quad (4)$$

where $E_{\text{corr,X}}$ is the correlation energy obtained with the correlation consistent basis set (cc-pVXZ) with cardinal number ($X = \text{T and Q}$) and $E_{\text{corr,lim}}$ is the basis set limit value of the correlation energy. The post-MP2 correction ($\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}}$), also known as “coupled cluster correction”, is calculated using the augmented correlation consistent basis set aug-cc-pVDZ.

Solution phase intermolecular interaction energy. The energy of intermolecular interaction in the solution phase was evaluated indirectly by means of a thermodynamic scheme. For a detailed description of the scheme, interested readers can refer to Reference [6]. According to the scheme, the binding energy for complex formation in the solution phase can be evaluated indirectly by calculating intermolecular interaction energies in the gas phase, ΔE_{int}^g , followed by a correction for the dehydration energy ΔE_{Deh} :

$$\Delta E_{\text{int}}^{\text{aq}} = \Delta E_{\text{int}}^g + \Delta E_{\text{Deh}} \quad (5)$$

The dehydration energy for the complex formation is defined by:

$$\Delta E_{\text{Deh}} = \Delta G_{\text{AB}}^{\text{sol}} - \Delta G_{\text{A}}^{\text{sol}} - \Delta G_{\text{B}}^{\text{sol}} \quad (6)$$

where ΔG_i^{sol} , $i = \text{AB, A, B}$ represents the free energies of solvation for the complex AB, and the monomers A, B, respectively. The SM5.42R Solvation Model of Cramer and Truhlar [69] as implemented in the 2008 R1 version of GAMESS [82] was adopted for the evaluation of those free energies of solvation. The SM5.42R model was chosen for its reported improvement of accuracy in predicting solvation energies [69]. In comparison to PCM and COSMO, it takes into account a wider range of solvent parameters, enabling a more complex and in-depth description of solvent–solute interactions.

4. Conclusions

Accurate quantification of the strengths of nonbonded interactions in large biological systems, like PKIs in their targeted proteins, is crucial for studying molecular recognition of ligands in proteins. However, high-accuracy computational methods often come at a substantial computational cost. This study navigates this trade-off by comprehensively assessing nine density functional theory (DFT) methods for their ability to balance accuracy and efficiency.

From a database of 2139 kinase–inhibitor crystal structures, we extracted a library of 49 diverse nonbonded interaction motifs, encompassing CH– π , π – π stacking, cation– π , hydrogen bonding, and salt bridge interactions. Using CCSD(T)/CBS as reference, we systematically benchmarked BLYP, TPSS, B97, ω B97X, B3LYP, M062X, PW6B95, B2PLYP, and PWPB95 functionals with D3BJ dispersion correction alongside def2-SVP, def2-TZVP,

and def2-QZVP basis sets. RI, RI-JK, and RIJCOSX approximations were employed for specific functionals. The B3LYP/def2-TZVP combination emerges as the optimal choice, demonstrating a balanced performance with an RMSD of 0.50 kcal/mol and an average CPU time of 27.7 min. For applications demanding utmost accuracy, the B3LYP/def2-QZVP combination excels in RMSD (0.46 kcal/mol) but incurs a significantly higher average CPU cost of 479.8 min. The RIJK RI-B2PLYP/def2-QZVP combination provides a viable alternative, achieving a balance with an RMSD of 0.51 kcal/mol and an averaged CPU time of 33.2 min. Notably, the RIJK RI-B2PLYP/def2-TZVP combination, while slightly less accurate (RMSD 0.66 kcal/mol), presents a substantial reduction in CPU time (7.7/33.2) against RIJK RI-B2PLYP/def2-QZVP, making it advantageous for applications that can tolerate a moderate loss of accuracy. Remarkably, different numerical implementations of the same functional can drastically impact CPU time, with RIJK RI B2PLYP/def2-QZVP requiring 13.4 times less time than RI B2PLYP/def2-QZVP.

These findings provide valuable insights for choosing the optimal DFT method for studying nonbonded interactions in large biological systems, particularly when computational expense is a major concern. This work paves the way for robust, high-throughput structure-based drug design targeting protein kinases through accurate modeling of critical nonbonded interactions.

Moreover, the calculated CCSD(T) energies extend beyond this study, offering a valuable benchmark for future testing of methods for calculating intermolecular interactions. Furthermore, the established library of 3D motifs stands as a significant contribution itself, providing a valuable resource for realistic molecular modeling of ligand binding with protein in general.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules29020304/s1>, Table S1: Gas phase intermolecular interaction energies at the DFT/def2-SVP level for the 49 binding motifs of PKIs in protein kinases. Table S2: Gas phase intermolecular interaction energies at the DFT/def2-TZVP level for the 49 binding motifs. Table S3: Gas phase intermolecular interaction energies (in kcal/mol) at the DFT/def2-QZVP level for the 49 binding motifs.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, X.H. and Y.Z.; methodology, Y.Z., S.A. and X.H.; software, Y.Z. and S.A.; validation Y.Z. and S.A.; formal analysis, Y.Z.; investigation, Y.Z.; resources, X.H.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, X.H.; visualization, Y.Z.; supervision, X.H.; project administration, X.H.; funding acquisition, X.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Coordinates of 49 motifs of nonbonded interactions between PKIs and their respective binding protein kinases are available from https://rocketsutoledo-my.sharepoint.com/:f/g/personal/xhu_rockets_utoledo_edu/EjY6HV9di7BBmZVh_MVoyUUBLJIKBHyDskH3GjBAzd-80A?e=aeKLVn. Access to the data set needs to be authorized by the authors upon request. The data are not publicly available due to internal campus network security protocol.

Acknowledgments: We are pleased to acknowledge the Ohio Supercomputer Center for a generous allocation of supercomputer time.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Manning, G.; Plowman, G.D.; Hunter, T.; Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* **2002**, *27*, 514–520. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Cohen, P.; Cross, D.; Janne, P.A. Kinase drug discovery 20 years after imatinib: Progress and future directions. *Nat. Rev. Drug Discov.* **2021**, *20*, 551–569. [\[CrossRef\]](#)
3. Melnikova, I.; Golden, J. *Targeting Protein Kinases*; Nature Publishing Group: New York, NY, USA, 2004.
4. Ayala-Aguilera, C.C.; Valero, T.; Lorente-Macias, A.; Baillache, D.J.; Croke, S.; Unciti-Broceta, A. Small Molecule Kinase Inhibitor Drugs (1995–2021): Medical Indication, Pharmacology, and Synthesis. *J. Med. Chem.* **2022**, *65*, 1047–1131. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Badrinarayan, P.; Sastry, G.N. Rational Approaches Towards Lead Optimization of Kinase Inhibitors: The Issue of Specificity. *Curr. Pharm. Des.* **2013**, *19*, 4714–4738. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Zhu, Y.; Alqahtani, S.; Hu, X.C. Aromatic Rings as Molecular Determinants for the Molecular Recognition of Protein Kinase Inhibitors. *Molecules* **2021**, *26*, 1776. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Persch, E.; Dumele, O.; Diederich, F. Molecular Recognition in Chemical and Biological Systems. *Angew. Chem.-Int. Ed.* **2015**, *54*, 3290–3327. [\[CrossRef\]](#)
8. Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061–5084. [\[CrossRef\]](#)
9. Hunter, C.A.; Singh, J.; Thornton, J.M. π - π interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol.* **1991**, *218*, 837–846. [\[CrossRef\]](#)
10. Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/ π Interaction: Evidence, Nature, and Consequences*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
11. Gallivan, J.P.; Dougherty, D.A. Cation- π interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459–9464. [\[CrossRef\]](#)
12. Zhu, Y.; Hu, X.C. Molecular Recognition of FDA-Approved Small Molecule Protein Kinase Drugs in Protein Kinases. *Molecules* **2022**, *27*, 7124. [\[CrossRef\]](#)
13. Meyer, E.A.; Castellano, R.K.; Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem.-Int. Ed.* **2003**, *42*, 1210–1250. [\[CrossRef\]](#)
14. Hobza, P. 2 Theoretical studies of hydrogen bonding. *Annu. Rep. Sect. "C" (Phys. Chem.)* **2004**, *100*, 3–27. [\[CrossRef\]](#)
15. Chalasinski, G.; Gutowski, M. Weak Interactions Between Small Systems—Models for Studying the Nature of Intermolecular Forces and Challenging Problems for Ab Initio Calculations. *Chem. Rev.* **1988**, *88*, 943–962. [\[CrossRef\]](#)
16. Grimme, S.; Antony, J.; Schwabe, T.; Mück-Lichtenfeld, C. Density functional theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (bio)organic molecules. *Org. Biomol. Chem.* **2007**, *5*, 741–758. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Mao, L.; Wang, Y.; Liu, Y.; Hu, X. Molecular determinants for ATP-binding in proteins: A data mining and quantum chemical analysis. *J. Mol. Biol.* **2004**, *336*, 787–807. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Kristyan, S.; Pulay, P. Can (Semi)Local Density-Functional Theory Account for the London Dispersion Forces. *Chem. Phys. Lett.* **1994**, *229*, 175–180. [\[CrossRef\]](#)
19. Becke, A.D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Lee, C.T.; Yang, W.T.; Parr, R.G. Development of the colle-salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988**, *37*, 785–789. [\[CrossRef\]](#)
21. Tao, J.M.; Perdew, J.P.; Staroverov, V.N.; Scuseria, G.E. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401. [\[CrossRef\]](#)
22. Becke, A.D. Density-functional thermochemistry.5. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560. [\[CrossRef\]](#)
23. Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Becke, A.D. Density-Functional Thermochemistry.3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [\[CrossRef\]](#)
25. Stephens, P.J.; Devlin, F.J.; Chabalowski, C.F.; Frisch, M.J. Ab-initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627. [\[CrossRef\]](#)
26. Zhao, Y.; Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
27. Zhao, Y.; Truhlar, D.G. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A* **2005**, *109*, 5656–5667. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals-Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309. [\[CrossRef\]](#) [\[PubMed\]](#)

30. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)]
31. Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305. [[CrossRef](#)]
32. Dunning, T.H. Gaussian-Basis Sets for Use in Correlated Molecular Calculations: 1. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023. [[CrossRef](#)]
33. Neese, F. The ORCA program system. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2012**, *2*, 73–78. [[CrossRef](#)]
34. Karton, A.; Martin, J.M.L. Prototypical π - π dimers re-examined by means of high-level CCSDT(Q) composite *ab initio* methods. *J. Chem. Phys.* **2021**, *154*, 124117. [[CrossRef](#)] [[PubMed](#)]
35. Rezac, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071. [[CrossRef](#)] [[PubMed](#)]
36. Rezac, J.; Riley, K.E.; Hobza, P. S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438. [[CrossRef](#)] [[PubMed](#)]
37. Davies, T.G.; Bentley, J.; Arris, C.E.; Boyle, F.T.; Curtin, N.J.; Endicott, J.A.; Gibson, A.E.; Golding, B.T.; Griffin, R.J.; Hardcastle, I.R. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nat. Struct. Mol. Biol.* **2002**, *9*, 745. [[CrossRef](#)]
38. Silva-Santisteban, M.C.; Westwood, I.M.; Boxall, K.; Brown, N.; Peacock, S.; McAndrew, C.; Barrie, E.; Richards, M.; Mirza, A.; Oliver, A.W. Fragment-based screening maps inhibitor interactions in the ATP-binding site of checkpoint kinase 2. *PLoS ONE* **2013**, *8*, e65689. [[CrossRef](#)]
39. Vulpetti, A.; Casale, E.; Roletto, F.; Amici, R.; Villa, M.; Pevarello, P. Structure-based drug design to the discovery of new 2-aminothiazole CDK2 inhibitors. *J. Mol. Graph. Model.* **2006**, *24*, 341–348. [[CrossRef](#)]
40. Schulz, M.N.; Fanghänel, J.; Schäfer, M.; Badock, V.; Briem, H.; Boemer, U.; Nguyen, D.; Husemann, M.; Hillig, R.C. A crystallographic fragment screen identifies cinnamic acid derivatives as starting points for potent Pim-1 inhibitors. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 156–166. [[CrossRef](#)]
41. Komander, D.; Kular, G.S.; Schüttelkopf, A.W.; Deak, M.; Prakash, K.; Bain, J.; Elliott, M.; Garrido-Franco, M.; Kozikowski, A.P.; Alessi, D.R. Interactions of LY333531 and other bisindolyl maleimide inhibitors with PDK1. *Structure* **2004**, *12*, 215–226. [[CrossRef](#)]
42. Smaill, J.B.; Baker, E.N.; Booth, R.J.; Bridges, A.J.; Dickson, J.M.; Dobrusin, E.M.; Ivanovic, I.; Kraker, A.J.; Lee, H.H.; Lunney, E.A. Synthesis and structure–activity relationships of N-6 substituted analogues of 9-hydroxy-4-phenylpyrrolo [3, 4-c] carbazole-1, 3 (2H, 6H)-diones as inhibitors of Wee1 and Chk1 checkpoint kinases. *Eur. J. Med. Chem.* **2008**, *43*, 1276–1296. [[CrossRef](#)]
43. Derewenda, U.; Artamonov, M.; Szukalska, G.; Utepbergenov, D.; Olekhovich, N.; Parikh, H.I.; Kellogg, G.E.; Somlyo, A.V.; Derewenda, Z.S. Identification of quercitrin as an inhibitor of the p90 S6 ribosomal kinase (RSK): Structure of its complex with the N-terminal domain of RSK2 at 1.8 Å resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 266–275. [[CrossRef](#)] [[PubMed](#)]
44. Rellos, P.; Pike, A.C.; Niesen, F.H.; Salah, E.; Lee, W.H.; Von Delft, F.; Knapp, S. Structure of the CaMKII δ /calmodulin complex reveals the molecular mechanism of CaMKII kinase activation. *PLoS Biol.* **2010**, *8*, e1000426. [[CrossRef](#)]
45. Aronov, A.M.; Tang, Q.; Martinez-Botella, G.; Bemis, G.W.; Cao, J.; Chen, G.; Ewing, N.P.; Ford, P.J.; Germann, U.A.; Green, J. Structure-guided design of potent and selective pyrimidylpyrrole inhibitors of extracellular signal-regulated kinase (ERK) using conformational control. *J. Med. Chem.* **2009**, *52*, 6362–6368. [[CrossRef](#)] [[PubMed](#)]
46. Lew, E.D.; Bae, J.H.; Rohmann, E.; Wollnik, B.; Schlessinger, J. Structural basis for reduced FGFR2 activity in LADD syndrome: Implications for FGFR autoinhibition and activation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19802–19807. [[CrossRef](#)] [[PubMed](#)]
47. Wang, S.; Meades, C.; Wood, G.; Osnowski, A.; Anderson, S.; Yuill, R.; Thomas, M.; Mezna, M.; Jackson, W.; Midgley, C. 2-Anilino-4-(thiazol-5-yl) pyrimidine CDK inhibitors: Synthesis, SAR analysis, X-ray crystallography, and biological activity. *J. Med. Chem.* **2004**, *47*, 1662–1675. [[CrossRef](#)]
48. Richardson, C.M.; Williamson, D.S.; Parratt, M.J.; Borgognoni, J.; Cansfield, A.D.; Dokurno, P.; Francis, G.L.; Howes, R.; Moore, J.D.; Murray, J.B. Triazolo [1, 5-a] pyrimidines as novel CDK2 inhibitors: Protein structure-guided design and SAR. *Bioorganic Med. Chem. Lett.* **2006**, *16*, 1353–1357. [[CrossRef](#)]
49. Schonbrunn, E.; Betzi, S.; Alam, R.; Martin, M.P.; Becker, A.; Han, H.; Francis, R.; Chakrasali, R.; Jakkaraj, S.; Kazi, A. Development of highly potent and selective diaminothiazole inhibitors of cyclin-dependent kinases. *J. Med. Chem.* **2013**, *56*, 3768–3782. [[CrossRef](#)]
50. Utepbergenov, D.; Derewenda, U.; Olekhovich, N.; Szukalska, G.; Banerjee, B.; Hilinski, M.K.; Lannigan, D.A.; Stukenberg, P.T.; Derewenda, Z.S. Insights into the inhibition of the p90 ribosomal S6 kinase (RSK) by the flavonol glycoside SL0101 from the 1.5 Å crystal structure of the N-terminal domain of RSK2 with bound inhibitor. *Biochemistry* **2012**, *51*, 6499–6510. [[CrossRef](#)]
51. Wu, S.Y.; McNae, I.; Kontopidis, G.; McClue, S.J.; McInnes, C.; Stewart, K.J.; Wang, S.; Zheleva, D.I.; Marriage, H.; Lane, D.P. Discovery of a novel family of CDK inhibitors with the program LIDAEUS: Structural basis for ligand-induced disordering of the activation loop. *Structure* **2003**, *11*, 399–410. [[CrossRef](#)]
52. Mohammadi, M.; McMahon, G.; Sun, L.; Tang, C.; Hirth, P.; Yeh, B.K.; Hubbard, S.R.; Schlessinger, J. Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science* **1997**, *276*, 955–960. [[CrossRef](#)]
53. De Moliner, E.; Moro, S.; Sarno, S.; Zagotto, G.; Zanotti, G.; Pinna, L.A.; Battistutta, R. Inhibition of protein kinase CK2 by anthraquinone-related compounds a structural insight. *J. Biol. Chem.* **2003**, *278*, 1831–1836. [[CrossRef](#)] [[PubMed](#)]

54. Andersen, C.B.; Wan, Y.; Chang, J.W.; Riggs, B.; Lee, C.; Liu, Y.; Sessa, F.; Villa, F.; Kwiatkowski, N.; Suzuki, M. Discovery of selective aminothiazole aurora kinase inhibitors. *ACS Chem. Biol.* **2008**, *3*, 180–192. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Amaning, K.; Lowinski, M.; Vallee, F.; Steier, V.; Marcireau, C.; Ugolini, A.; Delorme, C.; Foucalt, F.; McCort, G.; Derimay, N. The use of virtual screening and differential scanning fluorimetry for the rapid identification of fragments active against MEK1. *Bioorganic Med. Chem. Lett.* **2013**, *23*, 3620–3626. [\[CrossRef\]](#)
56. Fischmann, T.O.; Hruza, A.; Duca, J.S.; Ramanathan, L.; Mayhood, T.; Windsor, W.T.; Le, H.V.; Guzi, T.J.; Dwyer, M.P.; Paruch, K. Structure-guided discovery of cyclin-dependent kinase inhibitors. *Biopolym. Orig. Res. Biomol.* **2008**, *89*, 372–379. [\[CrossRef\]](#) [\[PubMed\]](#)
57. López-Ramos, M.; Prudent, R.; Moucadel, V.; Sautel, C.F.; Barette, C.; Lafanechère, L.; Mouawad, L.; Grierson, D.; Schmidt, F.; Florent, J.-C. New potent dual inhibitors of CK2 and Pim kinases: Discovery and structural insights. *FASEB J.* **2010**, *24*, 3171–3185. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Homan, K.T.; Wu, E.; Wilson, M.W.; Singh, P.; Larsen, S.D.; Tesmer, J.J. Structural and functional analysis of G protein-coupled receptor kinase inhibition by paroxetine and a rationally designed analog. *Mol. Pharmacol.* **2014**, *85*, 237–248. [\[CrossRef\]](#)
59. Zhao, B.; Bower, M.J.; McDevitt, P.J.; Zhao, H.; Davis, S.T.; Johanson, K.O.; Green, S.M.; Concha, N.O.; Zhou, B.-B.S. Structural basis for Chk1 inhibition by UCN-01. *J. Biol. Chem.* **2002**, *277*, 46609–46615. [\[CrossRef\]](#)
60. Meng, Z.; Ciavarrì, J.P.; McRiner, A.; Zhao, Y.; Zhao, L.; Reddy, P.A.; Zhang, X.; Fischmann, T.O.; Whitehurst, C.; Siddiqui, M.A. Potency switch between CHK1 and MK2: Discovery of imidazo [1,2-a] pyrazine-and imidazo [1,2-c] pyrimidine-based kinase inhibitors. *Bioorganic Med. Chem. Lett.* **2013**, *23*, 2863–2867. [\[CrossRef\]](#)
61. Narayana, N.; Diller, T.C.; Koide, K.; Bunnage, M.E.; Nicolaou, K.; Brunton, L.L.; Xuong, N.-H.; Ten Eyck, L.F.; Taylor, S.S. Crystal structure of the potent natural product inhibitor balanol in complex with the catalytic subunit of cAMP-dependent protein kinase. *Biochemistry* **1999**, *38*, 2367–2376. [\[CrossRef\]](#)
62. Kutchukian, P.S.; Wassermann, A.M.; Lindvall, M.K.; Wright, S.K.; Ottl, J.; Jacob, J.; Scheufler, C.; Marzinzik, A.; Brooijmans, N.; Glick, M. Large scale meta-analysis of fragment-based screening campaigns: Privileged fragments and complementary technologies. *J. Biomol. Screen.* **2015**, *20*, 588–596. [\[CrossRef\]](#)
63. Davies, T.G.; Tunnah, P.; Meijer, L.; Marko, D.; Eisenbrand, G.; Endicott, J.A.; Noble, M.E. Inhibitor binding to active and inactive CDK2: The crystal structure of CDK2-cyclin A/indirubin-5-sulphonate. *Structure* **2001**, *9*, 389–397. [\[CrossRef\]](#)
64. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; et al. *Gaussian 09, Revision A. 02*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
65. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Halkier, A.; Klopper, W.; Helgaker, T.; Jorgensen, P.; Taylor, P.R. Basis set convergence of the interaction energy of hydrogen-bonded complexes. *J. Chem. Phys.* **1999**, *111*, 9157–9167. [\[CrossRef\]](#)
67. Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A.K. Basis-set convergence in correlated calculations on Ne, N₂, and H₂O. *Chem. Phys. Lett.* **1998**, *286*, 243–252. [\[CrossRef\]](#)
68. Sherrill, C.D.; Takatani, T.; Hohenstein, E.G. An Assessment of Theoretical Methods for Nonbonded Interactions: Comparison to Complete Basis Set Limit Coupled-Cluster Potential Energy Curves for the Benzene Dimer, the Methane Dimer, Benzene-Methane, and Benzene-H₂S. *J. Phys. Chem. A* **2009**, *113*, 10146–10159. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Li, J.; Zhu, T.; Hawkins, G.D.; Winget, P.; Liotard, D.A.; Cramer, C.J.; Truhlar, D.G. Extension of the platform of applicability of the SM5. 42R universal solvation model. *Theor. Chem. Acc.* **1999**, *103*, 9–63. [\[CrossRef\]](#)
70. Vahtras, O.; Almlöf, J.; Feyereisen, M.W. Integral approximations for lcao-scf calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518. [\[CrossRef\]](#)
71. Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A ‘chain-of-spheres’ algorithm for the Hartree-Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109. [\[CrossRef\]](#)
72. Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291. [\[CrossRef\]](#)
73. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065. [\[CrossRef\]](#)
74. Staroverov, V.N.; Scuseria, G.E.; Tao, J.; Perdew, J.P. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137. [\[CrossRef\]](#)
75. Hamprecht, F.A.; Cohen, A.J.; Tozer, D.J.; Handy, N.C. Development and assessment of new exchange-correlation functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271. [\[CrossRef\]](#)
76. Becke, A.D. Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, *104*, 1040–1046. [\[CrossRef\]](#)
77. Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel-and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102. [\[CrossRef\]](#)
78. Jung, Y.; Lochan, R.C.; Dutoi, A.D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802. [\[CrossRef\]](#)
79. Boys, S.F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553–566. [\[CrossRef\]](#)

80. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [[CrossRef](#)]
81. Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473. [[CrossRef](#)]
82. Schmidt, M.W.; Baldridge, K.K.; Boatz, J.A.; Elbert, S.T.; Gordon, M.S.; Jensen, J.H.; Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.J.; et al. General Atomic and Molecular Electronic-Structure System. *J. Comput. Chem.* **1993**, *14*, 1347–1363. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.