JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Heteroaromatic π-Stacking Energy Landscapes

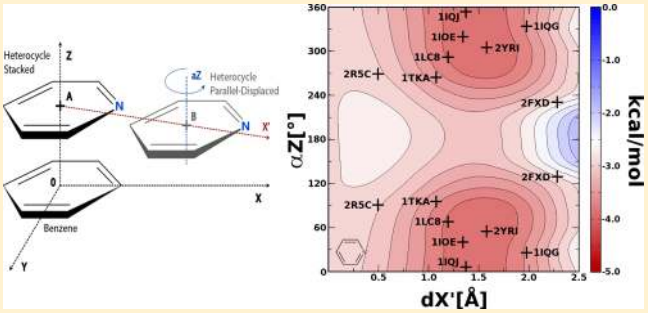
Roland G. Huber,[+] Michael A. Margreiter,[+] Julian E. Fuchs,[+] Susanne von Grafenstein,[+] Christofer S. Tautermann,[#] Klaus R. Liedl,*[,+] and Thomas Fox[#]

[+]Institute of General, Inorganic and Theoretical Chemistry, Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, CCB − Center for Chemistry and Biomedicine, Innrain 80/82, A-6020 Innsbruck, Austria

[#]Computational Chemistry, Lead Identification and Optimization Support, Boehringer Ingelheim Pharma GmbH & Co. KG, Birkendorfer Straße 65, D-88397 Biberach an der Riss, Germany

Supporting Information

ABSTRACT: In this study we investigate π-stacking interactions of a variety of aromatic heterocycles with benzene using dispersion corrected density functional theory. We calculate extensive potential energy surfaces for parallel-displaced interaction geometries. We find that dispersion contributes significantly to the interaction energy and is complemented by a varying degree of electrostatic interactions. We identify geometric preferences and minimum interaction energies for a set of 13 5- and 6-membered aromatic heterocycles frequently encountered in small drug-like molecules. We demonstrate that the electrostatic properties of these systems are a key determinant for their orientational preferences. The results of this study can be applied in lead optimization for the improvement of stacking interactions, as it provides detailed energy landscapes for a wide range of coplanar heteroaromatic geometries. These energy landscapes can serve as a guide for ring replacement in structure-based drug design.



INTRODUCTION

Heteroaromatic rings are a key component in most known drug molecules.[1] They provide rigid building blocks for anchoring substituents in defined geometric positions, thereby determining the overall molecular shape. They also play a crucial role in molecular recognition by proteins.[2] Their unique electronic structure with a distinct π-cloud located parallel above and below the ring plane allows for a variety of interaction patterns.[3] The distribution as well as the overall density of electrons within the conjugated π-system of a prototypical benzene molecule can either be modulated by substituents in various geometric arrangements[4,5] or by exchanging ring carbon atoms by nitrogen, oxygen, or sulfur creating heterocyclic systems.[6] Their unique interactions are a reason for their frequent occurrence in natural as well as synthetic bioactive molecules.

Interaction patterns exhibited by aromatic heterocycles comprise hydrophobic, polar, hydrogen bonding, cation-π,[7−9] amid-π,[10] halogen-π,[11] and π-stacking interactions.[12] π-Stacking interactions have been investigated in detail by a wide range of experimental and computational methods: for benzene dimers,[13−16] in DNA[17] or in proteins.[18,19] However, it should be noted that the study of noncovalent interactions remains challenging for computational as well as experimental approaches.[20]

Extensive *ab initio* calculations have been performed on the benzene dimer.[21,22] In addition to the intrinsic relevance of this specific system, these calculations are often used to evaluate the performance of quantum chemical and force field methods in describing aromatic systems.[23,24] Previous studies used coupled cluster with singles, doubles, and perturbative triples (CCSD(T))[25] up to complete basis set (CBS) extrapolation[26] to evaluate accurate interaction energies.[27,28] Møller−Plesset perturbation theory was found inadequate to describe the interaction energies of π-stacking complexes, as it tends to significantly overestimate interaction energies and underestimate equilibrium distances at the CBS limit.[29] Compared to the benzene dimer, studies of interactions of aromatic heterocycles are far less numerous. Some attention has been paid to the interaction of benzene with pyridine and the pyridine dimer.[30] As the main focus of this study is to characterize the potential energy surfaces (PES) for a diverse set of 13 aromatic heterocycles, the coupled cluster approach would be prohibitively expensive. Nevertheless, we employed CCSD(T) calculations using CBS extrapolation and the cc-pVTZ basis set[31] as a reference to identify a suitable density functional method for our calculations.

Density functional theory (DFT)[32] has been the most widely used approach in electronic structure calculations, providing a reasonable trade-off between accuracy and computational efficiency. Since long-range dispersion effects are known to

Received: March 21, 2014

   dx.doi.org/10.1021/ci500183u | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

be an important contribution to the interaction energy of the benzene dimer, computationally cheap Hartree−Fock calculations and most DFT methods are unsuitable to study π-stacking.[33] Recently, several density functionals including a correction term for long-range dispersion interactions have been developed.[34] Various functional forms for including such a term have been proposed. The most common functional form for a correction term is an r^{-6} potential using a nucleus-specific dispersion contribution for each atom.[35] These parameters are combined for each pair of atoms to give a dispersion parameter for the respective interaction. Furthermore, an exponential damping function is usually applied to avoid unphysical attraction at short-range. Some studies suggest an overestimation of the stability of large $\pi-\pi$ complexes in pairwise long-range corrected DFT methods.[36,37] They propose the inclusion of an empirical 3-body dispersion term.[38,39] However, our reference calculations at the CCSD(T)/CBS level of theory suggest that this effect is not relevant for the size of the systems investigated in this study.

Interestingly, only a limited set of aromatic heterocycles frequently appears in drug molecules. We performed extensive DFT calculations to characterize π-stacking interactions between benzene and a variety of aromatic heterocycles relevant in drug design as identified by Ertl and co-workers.[40] Our central goal is to understand the effect of heteroaromatic stacking in protein−ligand recognition processes and provide guidelines for ligand optimization along the lines of Bissantz et al.[41] We present PES for a variety of aromatic heterocycles interacting with benzene. Benzene is used as a simplified model of a phenylalanine residue. A variety of interaction geometries deviating from the optimal relative orientation for heterocyclic-aromatic interactions can be observed in crystal structures (Figure 1).

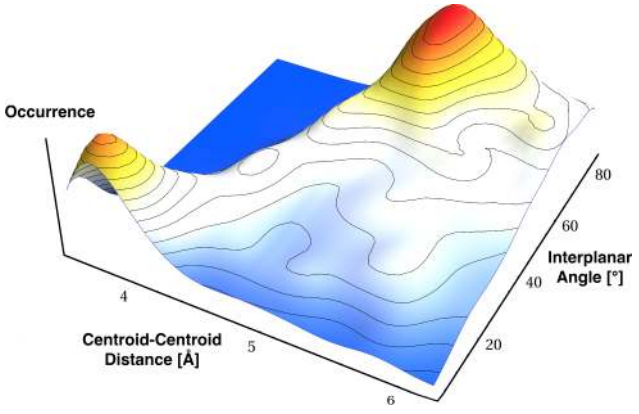


Figure 1. Observed interactions of aromatic heterocycles in ligands with aromatic protein side chains in the PDB extracted with Relibase+. The two distinct maxima represent a parallel-displaced (low interplanar angle, left peak) as well as a T-shaped interaction geometry (high interplanar angle, right peak). For the purposes of this study, we focus on parallel-displaced geometries usually associated with π-stacking.

Therefore, we calculated PES for all investigated systems. Furthermore, the geometric requirements of a certain binding mode do not allow for arbitrary placement of interaction partners. We provide fundamental data useful in structure-based lead optimization by suggesting possible ring substitutions that provide a gain in affinity through strengthened π-stacking interactions.

METHODS

A coordinate system was defined for all complexes as outlined in Figure 2. We set the origin to the centroid of the benzene

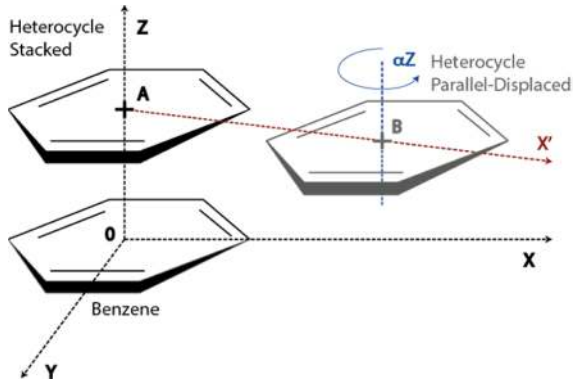


Figure 2. Procedure for scanning parallel-displaced PES: After initial identification of the stacked (point A) and parallel-displaced (point B) minima (ring centers denoted by +), a path X′ (red) for lateral displacement is chosen through these points. Starting from point A, the aromatic heterocycle is moved along X′ in increments of 1/10 the distance A−B. While moving the aromatic heterocycle through 20 increments of X′ displacement beyond point B, it is rotated through 360° in 30° intervals around a ring-plane normal axis in its center of geometry αZ (blue) at each point.

molecule, the *Z*-axis as a normal on the ring plane of the benzene molecule and the *X*-axis from the origin through one of the C−H bonds of the benzene molecule. A second benzene molecule was placed in the exact same orientation at an interplanar distance of 3.0 Å. The interplanar distance of this complex was subsequently varied in steps of 0.1 Å. The most favorable interaction energy along this path was defined as the "stacked" geometry (Point A, Figure 2) and used as a starting point for all further geometric manipulations.

Interaction energies were calculated by subtracting the energy of the two monomers, as suggested by the supermolecular approach:

$$E_{Interaction} = E_{Complex} - E_{Monomer\,A} - E_{Monomer\,B}$$

In order to identify a suitable DFT functional for this study, the stacked PES of the benzene dimer was scanned by performing *ab initio* calculations using Gaussian09 Rev. A.02[42] at the CCSD(T)/CBS level of theory by using aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ to obtain the interaction energy at the MP2/CBS limit through three-point extrapolation as described by Helgaker et al.[43] The difference between MP2/aug-cc-pVDZ and CCSD(T)/aug-cc-pVDZ was added to the extrapolated MP2/CBS energy to obtain the CCSD(T)/CBS value. The monomer geometry of benzene was obtained by using the default geometry in GaussView 4.1. For further validation, lateral displacement along dX was varied in order to sample interaction energies of parallel-displaced geometries at the CCSD(T)/cc-pVTZ level of theory. At each interplanar distance dZ and lateral displacement dX single point energy calculations were performed.

This calculation was repeated with the Hartree−Fock method and several density functional methods comprising B3LYP[44] and the dispersion corrected methods M06-2X[45] and ωB97XD[46] using the cc-pVDZ, aug-cc-pVDZ, cc-pVTZ and aug-cc-pVTZ basis sets from Dunning and co-workers.[31]

dx.doi.org/10.1021/ci500183u | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Results obtained using ωB97XD/cc-pVTZ were consistently within less than 0.5 kcal/mol compared to CCSD(T)/CBS calculations for stacked geometries (see Supporting Information Figure 1) and CCSD(T)/cc-pVTZ calculations of parallel-displaced geometries (see Supporting Information Figure 2). Therefore, all subsequent interaction energy and dipole moment calculations were performed at the ωB97XD/cc-pVTZ level of theory.

A set of seven 5-membered and six 6-membered heteroaromatic systems (see Figure 3) was prepared by fully

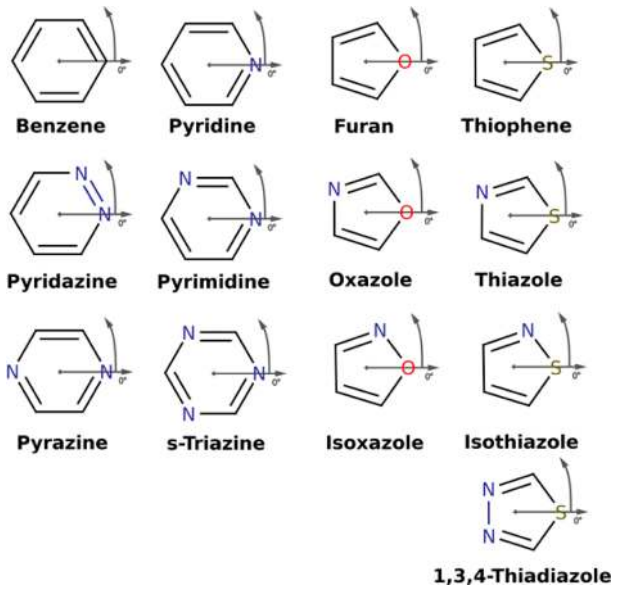


Figure 3. Initial ring orientations: All investigated aromatic heterocycles, in their αZ = 0° orientation and the direction of rotation along αZ are indicated in this plot.

optimizing each aromatic heterocycle as a monomer at the ωB97XD/cc-pVTZ level of theory. A complex with benzene was prepared from each resulting geometry in the relative orientation outlined in Figures 2 and 3. To scan the PES of these complexes, rotational and translational degrees of freedom were explored by performing single point energy calculations.

The parallel-displaced minimum (Point B, Figure 2) was identified by a sequence of coarse manual optimization followed by free geometry optimization. Manual optimization was performed by displacing the aromatic heterocycle from the stacked minimum along the X direction in steps of 0.1 Å while keeping the benzene at the origin stationary. At each point, the aromatic heterocycle was rotated to sequentially align each ring atom with the *X*-axis in order to identify the preferred orientation. Subsequently, a free geometry optimization was performed starting from the energetically most favorable orientation identified manually. The aromatic heterocycle centroid of the resulting geometry was defined as Point B.

After identifying both the respective stacked and parallel-displaced minimum geometries, a straight transition path X′ through these points was defined (Figure 2). Each aromatic heterocycle was moved along this path in increments of 1/10 of the distance A−B for a total of 20 increments while keeping the ring planes parallel. At each increment of X′ displacement, the aromatic heterocycle was rotated in 12 steps of 30° in the counterclockwise direction. Single point energies were calculated for each geometry in order to sample the PES starting from the stacked configuration in point A through the parallel-displaced minimum in point B and beyond. Location of points A and B as well as the slope for X′ for all aromatic heterocycles are given in Supporting Information Table 2. dZ denotes the interplanar distance at the respective point. As dX denotes lateral displacement, it is 0 for the stacked geometry and is given only for the parallel-displaced minimum at point B. The slope denotes the change in dZ per 1 Å of lateral displacement along dX when following the transition path X′. Using this approach, we can scan the most relevant region of the respective PES of parallel-displaced arrangements using just 2 parameters.

All calculated single points for each system were combined to form a PES for parallel-displaced geometries of the respective system. Smoothed contour plots were created by cubic interpolation using MATLAB R2012a.[47] Contour plots of the PES for all systems near their respective parallel-displaced minima are depicted in Figure 4.

Experimental data was extracted from the PDB[48] using Relibase+ (version 3.1).[49] For the analysis of geometries, all aromatic 5- and 6-rings within a distance of less than 6.5 Å of any aromatic side chain atom of a receptor were selected. Heme and the purine and pyrimidine DNA and RNA bases were excluded from this selection in order to focus the analysis on drug-like molecules. Furthermore, all NMR structures as well as all X-ray structures with a resolution worse than 2.5 Å were not considered in this analysis. In order to validate our theoretical findings we specifically identified interactions of nonannelated pyridine substructures interacting with phenylalanine. We identified 8 protein−ligand complexes and mapped them to our PES of the benzene−pyridine interaction.

RESULTS AND DISCUSSION

The tested density functionals show significant differences in their ability to accurately describe π-stacking interactions: B3LYP did not show any discernible favorable interaction between two benzene molecules. The M06-2X functional exhibited oscillations in interaction energy in areas away from the minimum. This was deemed undesirable, since the study aims at characterizing large areas of the PES in order to estimate preferred as well as disfavored interaction geometries for the respective aromatic heterocycles.

Whereas ωB97XD with both double-ζ basis sets gave unsatisfactory results with energy deviations exceeding 1 kcal/mol compared to CCSD(T)/CBS, the triple-ζ basis sets were able to yield consistent interaction energies to CCSD(T)/CBS with deviations smaller than 0.5 kcal/mol. Usage of aug-cc-pVTZ did not yield significant improvements in accuracy over the cc-pVTZ basis set. Therefore, cc-pVTZ was used for all subsequent calculations in order to optimize computational efficiency. PES were calculated as outlined in the Methods section for all 12 heteroaromatic complexes and the benzene dimer. The resulting energy landscapes at a distance of up to 2.5 Å in the X′ direction are given in Figure 4. 3D depictions of the parallel-displaced minima as identified through the PES scans are given in Figure 5.

Placing an electronegative atom over the benzene ring surface is energetically disfavored in all cases. This also holds true for multiple heteroatoms in a single ring, e.g. in the case of pyrazine, where two distinct saddle points in the rotational dimension can be observed at αZ angles of 0° and 180°, respectively (Figure 4e). The preferred rotational orientation is

dx.doi.org/10.1021/ci500183u | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

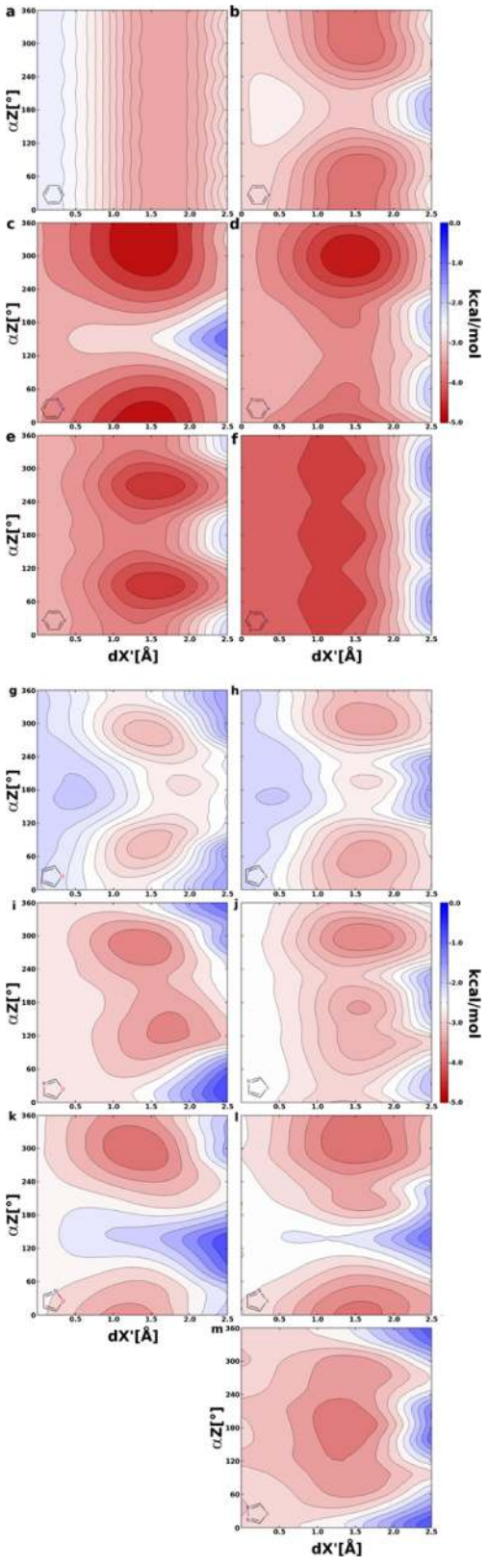


Figure 4. PES for all investigated aromatic heterocycles. Energy surfaces of parallel-displaced geometries for 6-membered (a benzene, b pyridine, c pyridazine, d pyrimidine, e pyrazine, f s-triazine) and 5-membered (g furan, h thiophene, i oxazole, j thiazole, k isoxazole, l isothiazole, m 1,3,4-thiadiazole) rings investigated in this study are given here. Lateral displacement is denoted by dX′, rotation angle by αZ. A common coloring scheme across all plots gives an indication of the relative energetics of the interactions.

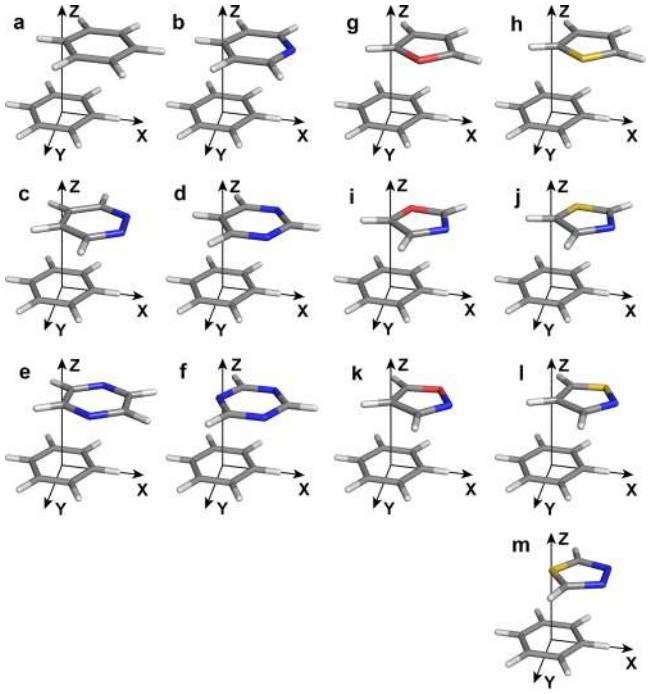


Figure 5. 3D depictions of the minimum geometries identified on the PES for all investigated systems. (a benzene, b pyridine, c pyridazine, d pyrimidine, e pyrazine, f s-triazine, g furan, h thiophene, i oxazole, j thiazole, k isoxazole, l isothiazole, m 1,3,4-thiadiazole).

highly dependent on the type of aromatic heterocycle as shown by the PES in Figure 4. A consistent interaction pattern can be observed for all investigated aromatic heterocycles: All systems exhibit a pronounced minimum at a parallel-displaced geometry at dX′ of approximately 1.5 Å as outlined in Figure 4. The interplanar angle resulting from unrestricted optimizations is below 10° in all cases. Interaction energies and minimum geometries resulting from unrestricted optimization are given in Figure 6.

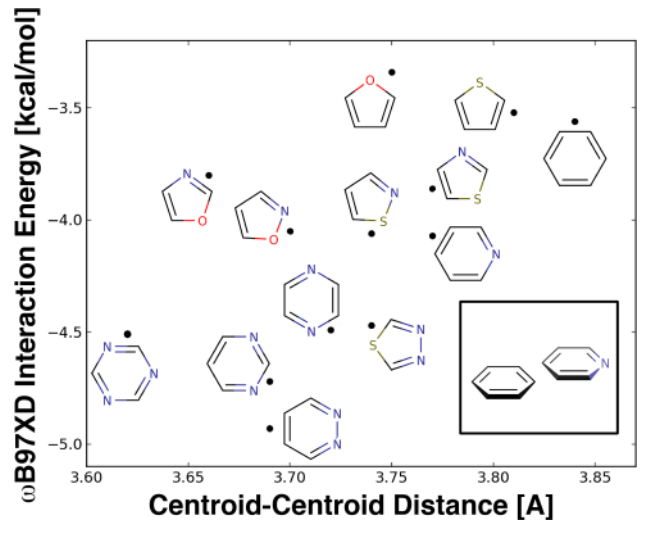


Figure 6. Interaction energies for all investigated heterocycles at the parallel-displaced minimum: The *x*-axis corresponds to the centroid distance at the energy minimum. Rotation of heterocycles is consistent with the minimum geometry analogous to the benzene−pyridine example.

dx.doi.org/10.1021/ci500183u | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

The interaction minima for aromatic heterocycles in parallel-displaced geometries are up to 1.5 kcal/mol more favorable than for the benzene dimer. Unfavorable orientations carry an energy penalty of up to 2 kcal/mol over benzene−benzene even in close proximity (displacement below 0.5 Å) to the minimum geometry. Nitrogen is generally oriented away from the ring center of the interaction partner in the respective minimum geometry.

A closer look at the series of 1,2-, 1,3-, and 1,4-diazabenzenes (pyridazine, pyrimidine, and pyrazine) reveals an interesting trend: In the case of pyrazine, it is geometrically impossible for both nitrogens to point away from the ring center of the interacting benzene due to the 1,4 positioning of the nitrogen atoms. It also has the weakest interaction energy at the minimum geometry of this series. Pyrimidine has a strong preferential orientation with both nitrogen atoms pointing away from the benzene ring center. It also shows a more favorable interaction energy compared to pyrazine. Pyridazine has the strongest interaction energy of all molecules in this series. Both nitrogen atoms are oriented away from the benzene ring center. Furthermore, pyridazine disfavors an orientation with the nitrogen atoms located directly above the benzene ring most strongly of all molecules in this series.

In the series oxazole to thiazole or isoxazole to isothiazole no significant changes of the interaction energies of these systems with benzene are observed. However, due to the larger size of the sulfur atom, the interplanar distance at the minimum is increased by approximately 0.1 Å in the case of isothiazole and by 0.05 Å in the case of thiazole compared to isoxazole and oxazole, respectively. Still, replacing furan by thiophene increases the interaction energy slightly and also increases the equilibrium distance by approximately 0.05 Å.

If we consider only systems with nonzero dipole moment (excluding benzene, pyrazine, and s-triazine), the resulting correlation of the magnitude of the dipole moment with the minimum interaction energy in parallel-displaced geometry is striking. The squared correlation coefficient is 0.78 as outlined in Figure 7. In general, an increasing dipole moment causes increasing interaction energy for the respective aromatic heterocycle. A significant portion of the observed variance in interaction energy at the parallel-displaced minimum geometry can be explained by the magnitude of the dipole moment. All systems with a dipole moment of zero do have nonzero quadrupole moments by which they can show electrostatic interactions as well. The observed deviation of pyrimidine from this trend is likely caused by the partial compensation of dipole moment vectors induced by the 1,3-positioning of the nitrogen atoms. This will lead to higher electrostatic moments being more pronounced, explaining the high interaction energy of pyrimidine.

As expected, the benzene dimer (Figure 4a) shows no preferred orientation. While this is also true for s-triazine (Figure 4f), s-triazine is the only investigated system featuring a low-energy stacked geometry. Pyridine (Figure 4b) prefers any parallel-displaced orientation locating the nitrogen atom away from the benzene π-cloud. This is also true for pyridazine (Figure 4c), showing a stronger gain in interaction energy than pyridine for favored configurations. Whereas pyrimidine (Figure 4d) also prefers to orient its nitrogen atoms away from the benzene π-cloud, pyrazine (Figure 4e) does not allow this configuration due to the 1,4 positioning of its nitrogen atoms. Pyrazine prefers to have its nitrogen atoms oriented perpendicular to the axis of displacement over one nitrogen atom placed above and the other outside the π-cloud of benzene.

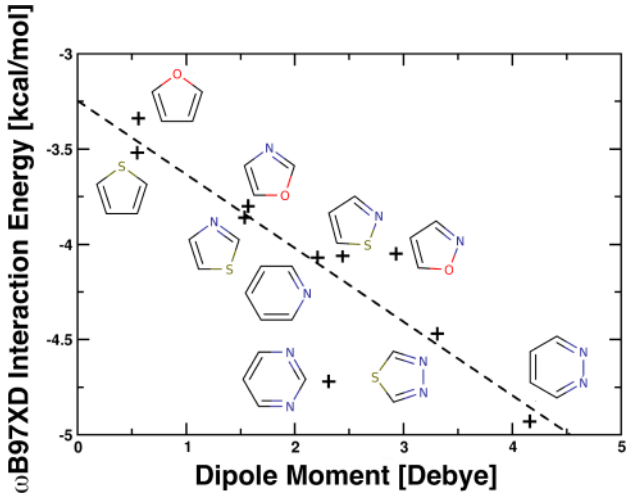


Figure 7. Correlation of interaction energy of 10 aromatic heterocycles and benzene at the minimum geometry with dipole moment: All systems without dipole moment were excluded from this graph (benzene, pyrazine and s-triazine). (left to right: thiophene, furan, thiazole, oxazole, pyridine, pyrimidine, isothiazole, isoxazole, 1,3,4-thiadiazole) Almost 80% of the variance in interaction energy can be explained by the magnitude of the dipole moment alone. A higher dipole moment (μ, Debye) yields a more favorable interaction energy (IE, kcal/mol) at the minimum. (μ = −0.37*IE-3.25 (R^2 = 0.78)).

Furan (Figure 4g) and thiophene (Figure 4h) show very similar interaction patterns in terms of energetics as well as orientational preference. Both systems favor placing the heteroatom above the hydrogen atoms of benzene located at a 60−90° angle of the axis of displacement. Due to the larger size of sulfur, the location of the minimum in thiophene is at a slightly (0.2 Å dX′) larger distance than for furan. The similarity of oxo- and thio-substituted aromatic heterocycles holds also true for oxazole (Figure 4i) and thiazole (Figure 4j) as well as isoxazole (Figure 4k) and isothiazole (Figure 4l). In all cases, the positioning of the nitrogen atom is again the crucial determinant of preferred orientation. Nitrogen is again preferably oriented away from the π-cloud of benzene. Placing either oxygen or sulfur directly above the center of the benzene molecule does not lower the interaction energy. The case of 1,3,4-thiadiazole (Figure 4m) is especially interesting when compared to isothiazole or isoxazole, as 1,3,4-thiadiazole exhibits the most favorable interaction energy at an orientation that is an energetic maximum for both isoxazole and isothiazole.

From the presented data we deduce two main contributing factors to the energy and geometric preference of hetero-aromatic stacking interactions. A significant part of the interaction energy is explained by dispersion interactions. This is exemplified by the benzene dimer, which has no dipole moment, exhibiting only weakly polarized C−H bonds, yet showing a favorable interaction of larger than 2 kcal/mol at the parallel-displaced minimum. Dispersion interactions do not show a significant directional preference and almost exclusively depend on the approximately constant contact area of the involved species. In addition to the dispersion component, we find an equally important electrostatic contribution. We presume that the geometric preference of specific systems is mainly driven by these electrostatic interactions. This is consistent with the findings of Sherrill in his recent study of

dx.doi.org/10.1021/ci500183u | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

stacking interactions of benzene and pyridine homo- and heterodimers.[30]

We find that the minimum interaction energy is strongly dependent on the magnitude of the dipole moment. Correlating the minimum interaction energy with the magnitude of the dipole moment of a heterocyclic monomer revealed an unexpectedly strong relationship between these factors. Around 80% of the observed variance in minimum interaction energy is explained by this simple descriptor. This finding is especially relevant, since the most commonly occurring natural aromatic heterocycles, i.e. pyrimidine and purine in the nucleobases, have a nonzero dipole moment.[50] Exploiting electrostatic interaction patterns might facilitate the design of environment-specific stacking interactions by rational choice of specific aromatic heterocycles.

Consistent series of molecules can be identified wherein variation of the placement of heteroatoms causes a systematic shift in the preferred interaction geometry and strength: For one, this is exemplified by a series of nitrogen-containing aromatic heterocycles: pyridine-pyridazine-pyrimidine-pyrazine. Throughout the series, interaction energy increases steadily. Second, substitution of ring atoms with atoms from the same group of the periodic table cause differences in centroid-centroid distances consistent with the size differences of these atoms while not affecting interaction energies significantly. Such substitutions are also demonstrated for the cases of oxazole/thiazole and isoxazole/isothiazole, respectively.

Ideal relative orientations and distances can be identified from our energy scans. Most investigated systems show broad energy minima that enable a certain degree of flexibility in relative orientation. However, certain orientations should be avoided, as improper positioning of electronegative atoms is associated with a significant energy penalty. Pyridine for example allows a rotation of around 240° without any significant deviation form the minimum interaction energy. The remaining 120° of possible relative orientations with nitrogen directly above the benzene ring can reduce the interaction energy by more than 1 kcal/mol, thereby making the interaction less favorable than an interaction with benzene. This effect is even more pronounced in pyridazine with two adjacent nitrogen atoms, where the energy penalty for improper relative rotation is more than 2 kcal/mol.

Interaction geometries of pyridyl substrucures with phenylalanine were identified using Relibase+ as outlined in the Methods section and subsequently mapped to our coordinate system by calculating a best-fit plane through the centroid of the phenyl substructure of phenylalanine, calculating the projection of the ring centroid of the pyridyl substructure of the ligand to this plane and the respective ring rotation αZ. It can be seen, that the majority of interactions occur in regions identified as energetically highly favorable (complexes 1TKA, 1LC8, 1IOE, 1IQJ, 2YRI, 1IQG). However, certain complexes present the pyridyl substructure to the receptor in a nonoptimal way (complexes 2R5C, 2FXD). These examples illustrate the necessity for a more detailed picture of interaction energies in nonminimum regions. Our calculations suggest that reorientation or replacement of the pyridyl moiety in 2FXD would provide a gain in affinity.

1IQJ is a co-crystal of coagulation Factor Xa[51] with the inhibitor M55124 (Figure 8). It is well-known that the hydrophobic S4 pocket of Factor Xa interacts with aromatic substructures.[52] The depicted example shows excellent π-stacking with pyridine. According to our analysis, the

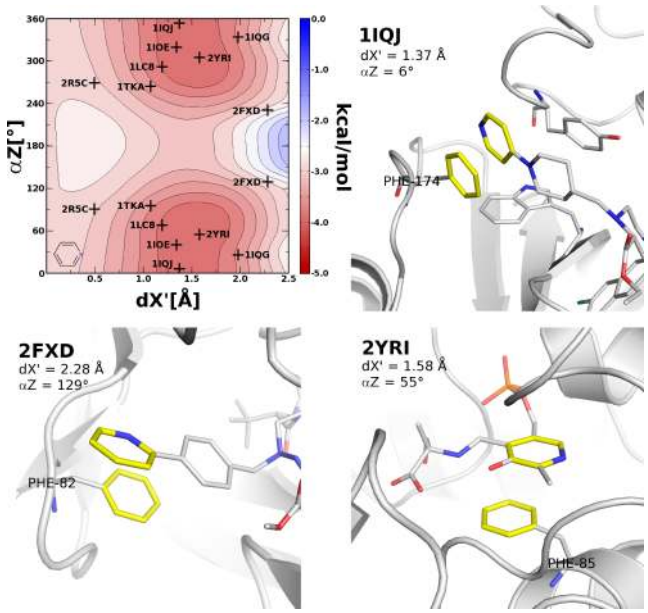


Figure 8. Occurrences of parallel-displaced Phe-pyridine interactions in protein−ligand complexes identified in the PDB are overlaid with the corresponding benzene−pyridine PES. Most interaction geometries occur in a relative orientation identified as favorable, as exemplified by the depicted interactions in 1IQJ and 2YRI. However, some interaction geometries are clearly in a nonpreferred region as exemplified by 2FXD.

interaction geometry of the pyridyl moiety in M55124 is in the optimal orientation relative to the side chain Phe-174 in terms of its stacking interaction (IE = −3.92 kcal/mol). However, we cannot exclude the possibility for the pyridine to be protonated, which would add a cation-π component to this interaction.

2YRI is an X-ray structure of alanine-pyruvate aminotransferase in complex with an inhibitor that exhibits a substituted pyridine as its central structural feature (Figure 8). Based on our analysis of interaction energies, the core pyridine is ideally positioned relative to Phe-85 in terms of π-stacking energy. The broad interaction energy minimum allows for this moiety to be rotated at around 60° along αZ relative to the Factor Xa-M55124 interaction described previously and still be optimal in terms of energy (IE = −3.97 kcal/mol).

2FXD[53] is a crystal structure of HIV-1 protease in complex with the inhibitor atazanavir (Figure 8). Atazanavir contains a benzene moiety, which is directly linked to the analyzed pyridine ring. Observed stacking interactions of the pyridyl moiety with Phe-82 show a distinct deviation from identified optimal relative geometries. According to our calculations, this interaction geometry is more than 1 kcal/mol unfavorable over optimal Phe-pyridine stacking (IE = −2.84 kcal/mol). However, the neighboring benzene in atazanavir is also involved in stacking with Phe-82. We argue that rotating the pyridine by 120° for the nitrogen to be in para-position would improve stacking interactions. However, this might further reduce coplanarity of the benzene and pyridine rings in atazanavir, as the added hydrogen in the ortho-position would induce additional strain energy in the ligand. This reduction in coplanarity would reduce stacking interactions for both the pyridine and benzene rings. Other marketed drugs that target HIV-1 protease frequently contain a single benzene ring in an analogous position (e.g., ritonavir, lopinavir).[54]

dx.doi.org/10.1021/ci500183u | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Two important classes of effects were neglected in this study. We expect substituent effects to have a significant impact, both in terms of geometric preferences and interaction energies. Substituent effects are known to change electrostatic properties of aromatic ring systems (e.g., the inversion of electrostatics for hexafluorobenzene at the extreme). Furthermore, different substituents introduce additional interaction centers as well as steric requirements. Although a detailed investigation of substituent effects is beyond the scope of this study, we assume the general workflow presented here to be applicable to the investigation of substituent effects as well.

Moreover, effects of solvation play an important role in protein−ligand affinity. As the ligand and the protein are solvated in aqueous solution prior to complex formation, (partial) desolvation of both interaction partners occurs during binding. We assume that different solvation energies of the investigated aromatic heterocycles have a significant impact on the overall binding energetics. However, we do not expect solvation to be crucial for the identified geometric preferences.

CONCLUSION

In the course of this study we evaluated dispersion corrected density functionals for their capacity to reproduce CCSD(T) energy surfaces. We considered the ωB97XD and M06-2X functionals developed by the Head-Gordon[45] and Truhlar[44] groups, respectively. We found that both functionals are well suited to reproduce energies and positions of the minima of high-level post-HF calculations. We found ωB97XD to be better suited for our purposes, as M06-2X exhibited oscillating energies in geometries away from the local energy minima.

Energetic differences between different interacting aromatic heterocycles are not immediately obvious even to the trained chemist. Therefore, our study can provide a reference for structure-based drug design by comparing ring orientations and their associated energies in a comprehensive manner. The calculated PES allow to immediately identify which ring replacements might produce a gain in affinity in cases where π-stacking interactions are relevant.

The ωB97XD functional employing high-ζ basis sets is well suited to investigate aromatic stacking interactions. We performed extensive PES scans for 13 complexes of 5- and 6-membered aromatic heterocycles with benzene. We identified optimal geometries and interaction energies at the parallel-displaced minimum. Furthermore, we present PES of parallel-displaced heteroaromatic systems as a guide to identify especially favorable or unfavorable interaction geometries. Moreover, we identified a strong dependence on electrostatic effects for the preferred interaction geometry as well as the minimum energy.

We conclude that dispersion is a significant contributor to the overall interaction energy, which is augmented to a varying degree by electrostatic effects. The presented PES as well as the identified optimal interaction geometries can provide guidance for structure-based lead optimization. This information can assist a medicinal chemist in identifying an appropriate aromatic heterocycle for a given interaction geometry (compare Bissantz et al.[41]). An analogous methodology will be applied to study T-shaped aromatic interactions.

ASSOCIATED CONTENT

Supporting Information

Figure S1: Comparison of CCSD(T)/CBS with MP2/CBS, HF/CBS, and ωB97XD/cc-pVTZ for stacked geometries; Table S1: Values for Figure S1; Figure S2: Comparison of CCSD(T)/cc-pVTZ and ωB97XD/cc-pVTZ PES for the benzene dimer; Table S2: Interaction energies and dipole moments calculated at ωB97XD/cc-pVTZ; Table S2: Positions of points A and B, slope of X′ axis for all systems. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Klaus.Liedl@uibk.ac.at.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

R.G.H. and J.E.F. are recipients of a DOC Fellowship of the Austrian Academy of Sciences at the Institute of General, Inorganic and Theoretical Chemistry. This work was supported by the Austrian Ministry of Science BMWF as part of the Konjunkturpaket II of the Focal Point Scientific Computing at the University of Innsbruck. The authors would like to thank Hannes G. Wallnoefer for his involvement in establishing this project. This work was funded in part by Boehringer Ingelheim Pharma GmbH & Co. KG. This work was supported in part by the Austrian Science Fund FWF project P23051 "Targeting Influenza Neuraminidase".

ABBREVIATIONS

DFT: denstiy functional theory; CCSD(T): coupled cluster singles doubles perturbative triples; CBS: complete basis set; PES: potential energy surface(s); IE: interaction energy

REFERENCES

(1) Feher, M.; Schmidt, J. M. Property distributions: Differences between drugs, natural products and molecules from combinatorial chemistry. *J. Chem. Inf. Model.* **2003**, *43*, 218−227.

(2) Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem., Int. Ed.* **2003**, *92*, 1210−1250.

(3) Schneider, H.-J. Binding Mechanisms in Supramolecular Complexes. *Angew. Chem., Int. Ed.* **2009**, *48*, 3924−3977.

(4) Fukui, K.; Yonezawa, T.; Nagata, C.; Shingu, H. Molecular Orbital Theory of Orientation in Aromatic, Heteroaromatic and Conjugated Molecules. *J. Chem. Phys.* **1954**, *22*, 1433−1442.

(5) Brown, H. C.; Okamaoto, Y. Directive Effects in Aromatic Substitution 30. Electrophilic Substituent Constants. *J. Am. Chem. Soc.* **1958**, *80*, 4979−4987.

(6) Pople, J. A.; Gordon, M. Molecular Orbital Theory of Electronic Structure of Organic Compunds 1. Substituent Effects and Dipole Momenzs. *J. Am. Chem. Soc.* **1967**, *89*, 4253−4261.

(7) Ma, J. C.; Dougherty, D. A. The cation-pi interaction. *Chem. Rev.* **1997**, *97*, 1303−1324.

(8) Dougherty, D. A. Cation-pi interactions in chemistry and biology: A new view of benzene, Phe, Tyr and Trp. *Science* **1996**, *271*, 163−168.

(9) Gallivan, J. P.; Dougherty, D. A. Cation-pi interactions in structural biology. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9459−9464.

(10) Harder, M.; Kuhn, B.; Diederich, F. Efficient Stacking on Protein Amide Fragments. *ChemMedChem.* **2013**, *8*, 397−404.

(11) Wallnoefer, H. G.; Fox, T.; Liedl, K. R.; Tautermann, C. S. Dispersion dominated halogen-pi interactions: energies and locations of minima. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14941−14949.

(12) Janiak, C. A critical account of pi-pi stacking in metal complexes with aromatic nitrogen-containing ligands. *J. Chem. Soc., Dalton Trans.* **2000**, *21*, 3885−3896.
(13) Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. Extracting Hydrophobic Free Energies from Experimental Data − Relationship to Protein Folding and Theoretical Models. *Biochemistry* **1991**, *30*, 9686−9697.
(14) Shetty, A. S.; Zhang, J. S.; Moore, J. S. Aromatic pi-stacking in solution as revealed through the aggregation of phenylacetylene macrocycles. *J. Am. Chem. Soc.* **1996**, *118*, 1019−1027.
(15) Börnsen, K. O.; Selzle, H. L.; Schlag, E. W. Spectra of Isotopically Mixed Benzene Dimers − Details on the Interaction in the VdW Bond. *J. Chem. Phys.* **1986**, *85*, 1726−1732.
(16) Grover, J. R.; Walters, E. A.; Hui, E. A. Dissociation Energies of the Benzene Dimer and Dimer Cation. *J. Phys. Chem.* **1987**, *91*, 3233−3237.
(17) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T. Factors contributing to aromatic stacking in water: Evaluation in the context of DNA. *J. Am. Chem. Soc.* **2000**, *122*, 2213−2222.
(18) McGaughey, G. B.; Gagne, M.; Rappe, A. K. Pi-stacking interactions − Alive and well in proteins. *J. Biol. Chem.* **1998**, *273*, 15458−15463.
(19) Burley, S. K.; Petsko, G. A. Aromatic-Aromatic Interaction − A Mechanism of Protein Structure Stabilization. *Science* **1985**, *229*, 23−28.
(20) Müller-Dethlefs, K.; Hobza, P. Noncovalent Interactions: A Challenge for Experiment and Theory. *Chem. Rev.* **2000**, *100*, 143−167.
(21) Hobza, P.; Selzle, H. L.; Schlag, E. W. Potential Energy Surface of the Benzene Dimer − Ab Initio Theoretical Study. *J. Am. Chem. Soc.* **1994**, *116*, 3500−3506.
(22) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(23) Dion, M.; Rydberg, H.; Schroder, E.; Langreth, D. C.; Lundqvist, B. I. Van der Waals density functional for general geometries. *Phys. Rev. Lett.* **2004**, *92*, 246401.
(24) 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 elemets H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(25) Purvis, G. D.; Bartlett, R. J. A Full Coupled Cluster Singles and Doubles Model − The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910−1918.
(26) Feller, D. Application of Systematic Sequences of Wave Functions to the Water Dimer. *J. Chem. Phys.* **1992**, *96*, 6104−6114.
(27) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the ab initio limit for pi-pi interactions: The benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.
(28) Hobza, P.; Selzle, H. L.; Schlag, E. W. Potential Energy Surface for the benzene dimer. Results of ab initio CCSD(T) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced. *J. Phys. Chem.* **1996**, *100*, 18790−18794.
(29) Sinnokrot, M. O.; Sherrill, C. D. Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped and parallel-displaced configurations. *J. Phys. Chem. A* **2004**, *108*, 10200−10207.
(30) Sherrill, C. D. Energy Component Analysis of π Interactions. *Acc. Chem. Res.* **2012**, *46*, 1020−1028.
(31) 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.
(32) Kohn, W.; Sham, L. J. Self-consistent Equations including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, 1133−1138.
(33) Andersson, Y.; Langreth, D. C.; Lundqvist, B. I. Van der Waals interactions in density functional theory. *Phys. Rev. Lett.* **1996**, *76*, 102−105.
(34) Becke, A. D.; Johnson, E. R. A density functional model of the dispersion interaction. *J. Chem. Phys.* **2005**, *123*, 154101.
(35) Grimme, S. Density Functional Theory with London Dispersion Corrections. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211−228.
(36) DiStasio, R. A., Jr.; von Lilienfeld, A.; Tkatchenko, A. Collective many-body van der Waals interactions in molecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14791−14795.
(37) Tkachenko, A.; Alfè, D.; Kim, K. S. First-Principles Modeling of Non-Covalent Interactions in Supramolecular Systems: The Role of Many-Body Effects. *J. Chem. Theory Comput.* **2012**, *8*, 4317−4322.
(38) Reilly, A. M.; Tkatchenko, A. Seamless and Accurate Modeling of Organic Molecular Materials. *J. Phys. Chem. Lett.* **2013**, *4*, 1028−1033.
(39) Risthaus, T.; Grimme, S. Benchmarking of London Dispersion-Accounting Density Functional Theory Methods on Very Large Molecular Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 1580−1591.
(40) Ertl, P.; Jelfs, S.; Muehlbacher, J.; Schuffenhauer, A.; Selzer, P. Quest for the rings. In silico exploration of ring universe to identify novel bioactive heteroaromatic scaffolds. *J. Med. Chem.* **2006**, *49*, 4568−4573.
(41) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061−5084.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.02*; Gaussian, Inc.: Wallingford, CT, 2009.
(43) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639.
(44) 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.
(45) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transistion elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(46) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(47) MATLAB R2012a. The MathWorks Inc.: Natick, MA, 2012.
(48) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(49) Hendlich, M.; Bergner, A.; Günther, J.; Klebe, G. Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. *J. Mol. Biol.* **2003**, *326*, 607−620.
(50) Yakovchuck, P.; Protozanova, E.; Frank-Kamenetskii, M. D. Base Stacking and Base Pairing Contributions into Thermal Stability of the DNA Double Helix. *Nucleic Acid Res.* **2006**, *34*, 564−574.
(51) Borensztajn, K.; Peppelenbosch, M. P.; Spek, C. A. Factor Xa: at the crossroads between coagulation and signaling in physiology and disease. *Trends Mol. Med.* **2008**, *14*, 429−440.
(52) Chou, Y. L.; Davey, D. D.; Eagen, K. A.; Griedel, B. D.; Karanjawala, R.; Phillips, G. B.; Sacchi, K. L.; Shaw, K. J.; Wu, S. C.; Lentz, D.; Liang, A. M.; Trinh, L.; Morrissey, M. M.; Kochanny, M. J.

Structure-activity relationships of substituted benzothiophene-anthranilamide factor Xa inhibitors. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 507−511.

(53) Klei, H. E.; Kish, K.; Lin, P. F.; Guo, Q.; Friborg, J.; Rose, R. E.; Zhang, Y.; Goldfarb, V.; Langley, D. R.; Wittekind, M.; Sheriff, S. X-Ray Crystal Structures of Human Immunodeficiency Virus Type 1 Protease Mutants Complexed with Atazanavir. *J. Virol.* **2007**, *81*, 9525−9535.

(54) Kiran Kumar Reddy, G. S.; Ali, A.; Nalam, M. N. L.; Anjum, G. S.; Cao, H.; Nathans, R. S.; Schiffer, C. A.; Rana, T. M. Design and Synthesis of HIV-1 Protease Inhibitors Incorporating Oxazolidinones as P2/P2′ Ligands in Pseudosymmetric Dipeptide Isosters. *J. Med. Chem.* **2007**, *50*, 4316−4328.