Predicting the Strength of Stacking Interactions between Heterocycles and Aromatic Amino Acid Side Chains

Andrea N. Bootsma,[†§] Analise C. Doney,[†] and Steven E. Wheeler[†§*]

[†]Department of Chemistry, Texas A&M University, College Station, TX 77842

[§]Center for Computational Quantum Chemistry, Department of Chemistry, University of Georgia, Athens, GA 30602

ABSTRACT: Despite the ubiquity of stacking interactions between heterocycles and aromatic amino acids in biological systems, our ability to predict their strength, even qualitatively, is limited. Based on rigorous *ab initio* data, we have devised a simple predictive model of the strength of stacking interactions between heterocycles commonly found in biologically active molecules and the amino acid side chains Phe, Tyr, and Trp. This model provides reliable predictions of the stacking ability of a given heterocycle based on readily-computed heterocycle descriptors, obviating the need for quantum chemical computations of stacked dimers. We show that the values of these descriptors, and therefore the strength of stacking interactions with aromatic amino acid side chains, follow simple predictable trends and can be modulated by changing the number and distribution of heteroatoms within the heterocycle. This provides a simple conceptual model for understanding stacking interactions in protein binding sites and tuning the strength of stacking interactions in drug design.

Introduction

Arene stacking interactions, broadly defined as approximately parallel face-to-face interactions between aromatic rings,[1-3] impact everything from the binding of ligands by proteins to the solid-state packing of organic electronic materials and the performance of asymmetric catalysts.[4-10] By understanding the factors that impact the strength of these interactions, we can explain their role in chemical and biochemical systems and exploit their power in the context of design.[11] Stacking interactions between heterocycles and the aromatic amino acid side chains Phe, Tyr, and Trp are particularly important in biological systems,[12-14] contributing to the binding of artificial inhibitors, natural substrates, co-factors, and nucleic acids by proteins (see Fig 1).[15-20] Changes in the number and distribution of heteroatoms within such heterocyclic systems can modulate their ability to stack with binding site aromatic residues, tuning their overall binding over a broad range.[15, 21-22]

The favorable stacking between aromatic rings has long been recognized, and Hunter and Sanders[23] provided a foundational conceptual model of the factors that impact the strength and geometry of these interactions. Over the last 15 years, high-accuracy quantum chemical studies and new experiments have provided a more nuanced view of these interactions, particularly with regard to the impact of substituents.[24-34] However, despite progress in the area of heterocycle stacking,[35-44] the factors that impact how strongly a given heterocycle will stack are still not completely understood. This limits our ability to understand stacking interactions in many biological contexts and hamstrings efforts at structure-based drug design. For example, even though tools exist for enumerating potential heterocyclic fragments during exploratory chemistry and lead optimization,[45-46] there is currently no means of ranking the ability of such fragments to stack with binding site aromatic residues short of expensive quantum chemical computations.

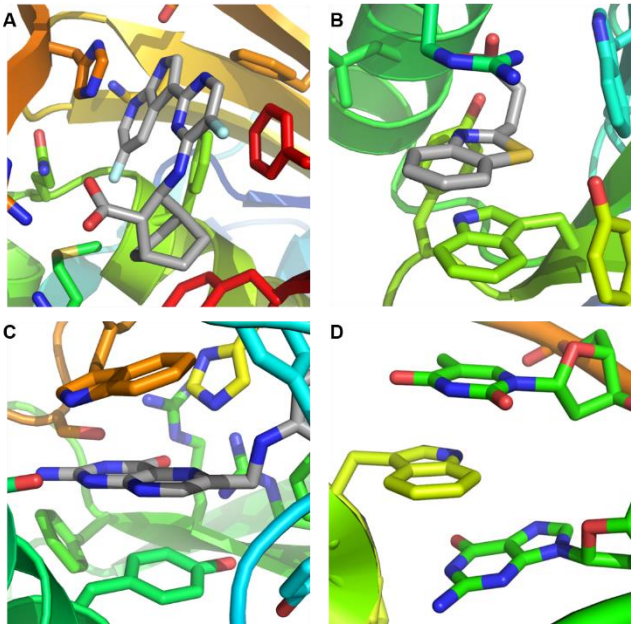


Figure 1. Examples of stacking interactions between heterocycles and aromatic amino acid side chains in biological systems. A. Inhibitor VX-787 (grey) bound to the influenza A viral polymerase (PDB: 4P1U); B. A quinazolinone-based inhibitor (grey) bound to a zinc-finger ubiquitin binding domain (PDB: 6CEF); C. Folate (grey) bound to the human folate receptor (PDB: 4KMZ). D. Single stranded DNA (green) bound to human TDP-41 (PDB: 4IUF).

In 2009, Hohenstein and Sherrill[39] published high-accuracy gas-phase computations of the pyridine-benzene dimer, showing that electrostatic interactions result in a 0.5 kcal mol^{-1} enhancement in the stacking interaction compared to the benzene dimer. Subsequently, Stahl and co-workers[47] published a guide to molecular interactions for medicinal chemists based on analyses of crystal structure data and representative drug binding sites. This included

the general advice that stacking interactions of electron-deficient rings are generally preferred over stacking of electron-rich rings and noted that the preferred orientation of stacked arenes can often be rationalized by the alignment of molecular dipoles or partial atomic charges.[4] In a different context, Corminboeuf and co-workers[48] showed that π-electron depletion correlates with the strength of stacking interactions, lending further support to the notion that electron-deficient arenes stack more effectively than their electron-rich counterparts.

There have been several efforts to distill our understanding of heterocycle stacking into predictive models based on simple molecular descriptors. Such models could potentially be used to predict the strength of stacking interactions without the need for expensive quantum chemistry computations. For instance, Huber and co-workers[49] demonstrated that molecular dipole moments are correlated with the strength of stacking interactions of monocyclic heterocycles with benzene. However, subsequent work by An, *et al.*[50] on stacking interactions between a broader set of biologically relevant heterocycles and 9-methyladenine showed that dipole moments alone do not capture trends in stacking interactions for bicyclic and larger heterocycles. Thus, it appears that the development of a predictive model of the strength of stacking interactions applicable to the diverse heterocycles appearing in biologically active molecules will require more advanced molecular descriptors.

We have demonstrated the utility of recently developed heterocycle descriptors[51-52] in predictive models of non-covalent interactions relevant to biological systems. For instance, we were able to obtain robust predictions of stacking interactions between Asp-Arg salt-bridges and heterocycles commonly found in pharmaceuticals[52] using heterocycle descriptors derived from the electrostatic potential (ESP) and electric field of the heterocycle. Here, we use these descriptors to develop a predictive model of the strength of stacking interactions between heterocycles found in biologically active molecules and Phe, Tyr, and Trp side chains, which informs the development of an understanding of these interactions.

Results and Discussion

We systematically searched for low-lying stacked dimers of toluene, *p*-methylphenol, and 3-methylindole (as models of Phe, Tyr, and Trp side chains, respectively; see Chart 1) with a training set of **46** heterocycles that are representative of those in biologically active molecules (**1**-**46**, Chart 1).[53-54] The gas-phase interaction energies for all unique stacked dimers were then computed using robust *ab initio* methods (See Computational Methods). Interaction energies for the global minimum energy stacked dimers of each heterocycle with Phe, Tyr, and Trp are listed in Table 1. Figure 2 shows geometries of selected dimers. The computed interaction energies for the heterocycles in this training set span more than 10 kcal mol^{-1}. This includes a span of 7.2 kcal mol^{-1} across the monocycles, from –3.6 kcal mol^{-1} for pyrazole (**8**) stacked with Phe to –10.8 kcal mol^{-1} for 1,2,3,4-tetrazine (**25**) stacked with Trp. As for the bicyclic systems, computed stacking energies range from –7.1 kcal mol^{-1} for isoindole (**28**) stacked with Phe to –15.8 kcal mol^{-1} for 3,9-dihydro-purine-2,6-dione (**37**) stacked with Trp. Stacking interactions of heterocycles with Tyr are, on average, ~10% stronger than those with Phe, while stacking interactions with Trp are ~40% stronger. Apart from this broad trend, there is considerable spread in predicted interaction energies for each amino acid side chain. This highlights the wide range over which heterocycle stacking interactions with a given aromatic amino acid side chain can vary while also revealing a powerful means of tuning stacking interactions in the context of drug design.[15, 21-22]

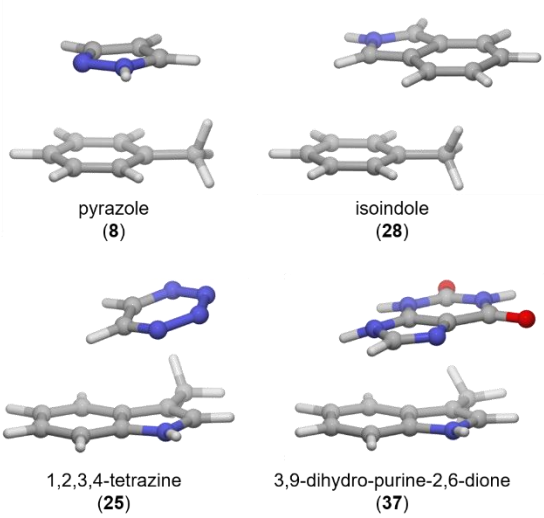


Figure 2. Stacked dimers of selected heterocycles with Phe (top) and Trp (bottom) side chains.

In order to understand the trends in stacking interaction energies across this set of heterocycles, the stacked dimers were analyzed using symmetry-adapted perturbation theory (SAPT).[55-58] SAPT enables the decomposition of interaction energies into electrostatic, dispersion, exchange-repulsion, and induction effects, and can be used to identify the effects that drive trends in interaction energies.[59] Overall, these data show that the stacking abilities of the heterocycles in Chart 1 are driven by dispersion and electrostatic effects (see SI Figure S1). This is in accord with the previous results of Hohenstein and Sherrill[39] for the stacked benzene-pyridine dimer.

A. Predictive Models of Stacking Interactions

Based on this insight, we sought to develop predictive multivariate models of the stacking energy between each of these heterocycles and the aromatic amino acids Phe, Tyr, and Trp based on readily computed descriptors. We considered our recently developed heterocycle descriptors[52] along with other physically motivated quantities such as dipole moment, quadrupole moment, volume, and polarizability (see SI Table S7 for full list of descriptors), focusing on descriptors that capture electrostatic effects and dispersion interactions. Initial exploratory fitting of the data for Phe, Tyr, and Trp separately revealed that the coefficients in each fit were proportional to the size of the amino acid side chain. This motivated consideration of functional forms in which the total interaction energy is scaled by the number of heavy atoms in the amino acid side chain (N_{HA}^{AA}).

Ultimately, we were able to develop four predictive models of the stacking interaction of heterocycles with Phe, Tyr, and Trp containing four adjustable parameters and one with three parameters. These five models were fit to the global minimum energy stacking interaction energies of **1**-**46** and provide nearly equally robust predictions for the training set data. Two are presented as equations 1 and 2 below (see SI Figures S2-S5 for other models):

Chart 1. Heterocycle training set along with the model amino acid side chains Phe, Tyr, and Trp.

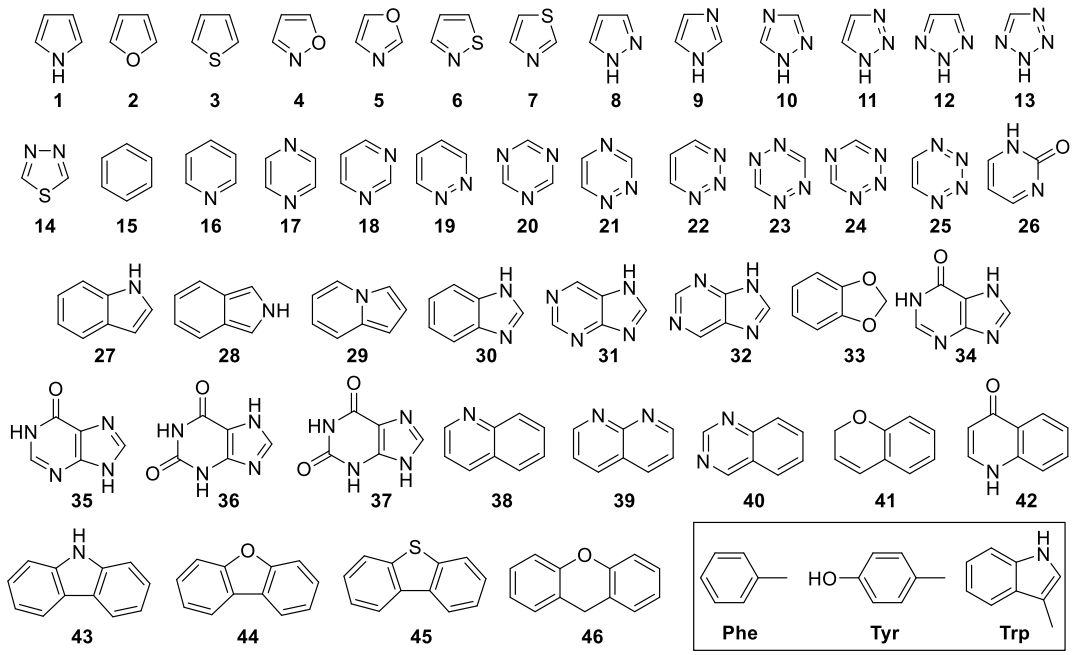


Table 1. Interaction energies for global minimum energy stacked dimers of heterocycles **1**-**58** with Phe, Tyr, and Trp side chains along with heterocycle descriptors, all in kcal mol^{-1}.[a]

Het	ESP^{Het}_{mean}	ESP^{Het}_{range}	ESP^{Het}_{max}	E^{Phe}_{int}	E^{Tyr}_{int}	E^{Trp}_{int}	Het	ESP^{Het}_{mean}	ESP^{Het}_{range}	ESP^{Het}_{max}	E^{Phe}_{int}	E^{Tyr}_{int}	E^{Trp}_{int}
1	–5.9	14.7	3.0	–	–	–4.5	**31**	–0.7	21.8	11.3	–9.4	–11.2	–11.0
2	–3.9	6.3	–0.7	–4.2	–4.3	–5.3	**32**	–1.1	14.9	8.2	–7.9	–9.1	–11.8
3	–4.5	7.0	–0.9	–4.5	–5.2	–5.8	**33**	–3.0	12.4	4.7	–7.6	–8.3	–13.6
4	–1.8	11.5	3.4	–5.1	–5.1	–7.2	**34**	–0.3	15.3	6.0	–9.4	–11.7	–12.4
5	–1.8	9.0	2.3	–5.1	–5.1	–6.7	**35**	–1.0	20.4	7.2	–9.2	–10.5	–10.2
6	–2.4	10.1	2.7	–5.5	–5.8	–7.2	**36**	1.1	16.4	8.9	–9.4	–11.5	–14.0
7	–2.5	8.8	1.9	–5.5	–5.5	–7.3	**37**	0.9	27.1	14.8	–8.8	–11.2	–14.4
8	–4.0	12.2	3.2	–3.6	–4.5	–5.1	**38**	–3.9	10.2	1.6	–8.1	–8.5	–14.2
9	–3.7	17.7	6.2	–	–	–8.4	**39**	–2.3	16.1	4.2	–8.9	–9.2	–15.8
10	–1.7	13.8	6.7	–4.3	–4.3	–	**40**	–1.8	9.1	3.0	–8.6	–9.3	–10.5
11	–1.8	17.8	6.9	–	–	–6.0	**41**	–4.2	12.8	3.3	–7.7	–9.1	–11.8
12	–2.1	7.0	2.7	–4.8	–5.1	–6.2	**42**	–3.2	22.8	8.6	–9.1	–10.1	–10.7
13	0.0	13.0	6.6	–5.2	–4.8	–	**43**	–5.8	13.2	3.0	–10.0	–11.2	–10.9
14	–0.5	12.7	4.1	–6.1	–6.3	–9.1	**44**	–4.6	6.9	–0.6	–8.8	–9.9	–13.3
15	–4.7	7.4	–1.4	–4.5	–4.9	–6.0	**45**	–5.0	7.6	–0.5	–9.2	–10.2	–14.2
16	–2.8	10.0	2.2	–5.6	–5.8	–7.7	**46**	–4.4	10.6	2.5	–9.8	–11.4	–11.8
17	–0.9	5.4	1.6	–6.0	–6.4	–8.1		**Mean**			–7.0	–7.8	–9.9
18	–0.8	8.8	4.1	–6.1	–6.4	–8.4	**47**	0.2	9.3	4.6	–	–5.6	–7.2
19	–0.7	15.4	5.5	–6.5	–6.7	–9.5	**48**	0.3	12.9	4.9	–5.7	–5.7	–8.0
20	1.1	3.8	2.5	–5.9	–6.6	–4.5	**49**	–4.8	11.0	2.1	–7.3	–7.6	–9.0
21	1.1	12.3	6.2	–7.0	–7.6	–5.3	**50**	–3.5	15.0	3.8	–8.3	–8.9	–11.3
22	1.2	17.6	8.5	–7.3	–8.6	–8.4	**51**	–4.2	10.2	2.3	–7.4	–7.9	–10.6
23	2.9	7.0	5.6	–6.8	–7.2	–9.9	**52**	–3.7	7.8	0.8	–7.8	–8.1	–9.9
24	3.0	11.2	6.4	–7.1	–7.8	–10.3	**53**	–5.3	8.0	–1.3	–7.3	–8.0	–10.0
25	3.1	16.8	10.6	–8.0	–8.7	–9.8	**54**	–2.0	10.5	3.8	–8.5	–9.8	–11.3
26	0.1	22.4	9.3	–7.3	–7.3	–9.9	**55**	0.0	9.7	5.2	–9.1	–10.3	–11.5
27	–6.2	15.5	3.5	–7.3	–8.6	–10.8	**56**	–1.9	14.1	2.9	–8.7	–11.0	–11.9
28	–6.7	17.7	4.4	–7.1	–7.8	–10.4	**57**	–2.1	14.5	3.9	–8.6	–9.3	–11.9
29	–5.1	11.4	1.1	–7.6	–8.1	–11.2	**58**	–2.0	18.3	6.4	–8.9	–9.6	–13.2
30	-4.5	17.0	6.4	-6.4	-9.3	-10.4							

[a] Dashes indicate that no stacked local energy minima were located. The tautomers of **10** and **13** were also considered but no stacked minima were located. Mean stacking interaction energies are for the training set (**1**-**46**).

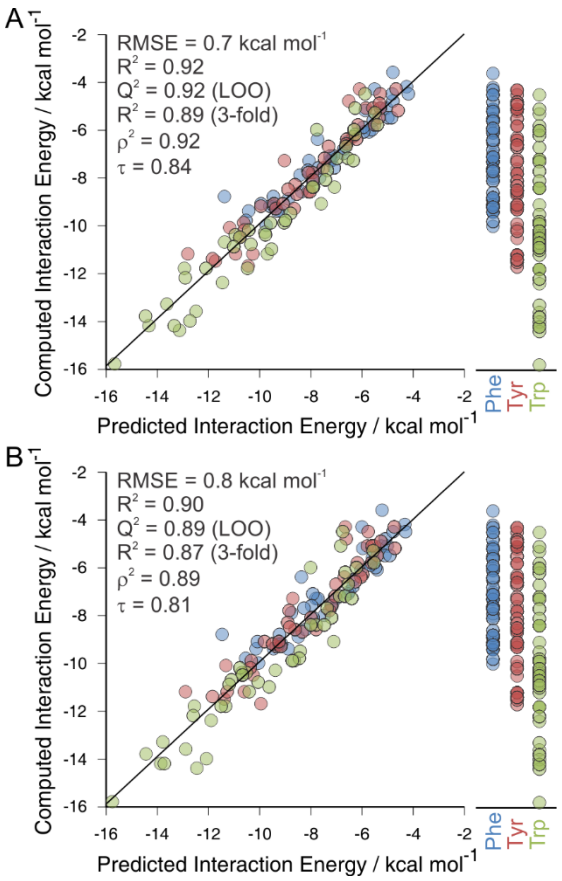


Figure 3. DLPNO-CCSD(T) interaction energies vs predicted interaction energies from equations 1 (A) and 2 (B) for global minimum energy stacked dimers of Phe, Tyr, and Trp side chains with the training set (**1**-**46**). The strip plots on the right show the same computed interaction energies grouped by amino acid.

$$\Delta E_{pred} = N_{HA}^{AA}(-0.036ESP_{mean}^{Het} - 0.013ESP_{range}^{Het} - 0.095N_{HA}^{Het}) - 1.36 \quad (1)$$

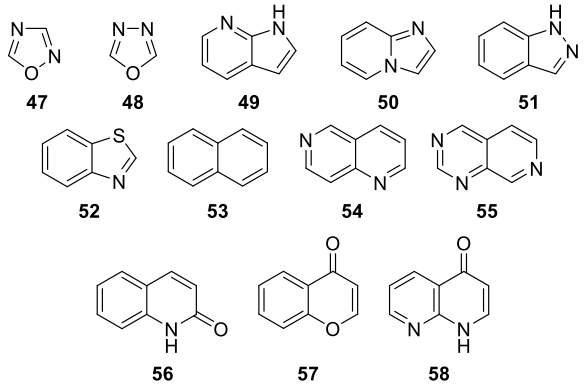
$$\Delta E_{pred} = N_{HA}^{AA}(-0.032ESP_{max}^{Het} - 0.087N_{HA}^{Het}) - 1.46 \quad (2)$$

In these equations, N_{HA} refers to the heavy atom count[60] of either the amino acid (N_{HA}^{AA}) or heterocycle (N_{HA}^{Het}), while ESP_{mean}^{Het}, ESP_{range}^{Het}, and ESP_{max}^{Het} are heterocycle descriptors[52] based on the mean, range, and maximum values of the ESP within the projection of the van der Waals volume of the heterocycle in a plane 3.25 Å away. The three other fits differ from equation 1 by the replacement of ESP_{range}^{Het} with ESP_{max}^{Het}, ESP_{σ}^{Het} (the standard deviation of the ESP in the plane 3.25 Å from the heterocycle), or F_{mean}^{Het} (the mean value of the electric field in this plane).

In essence, ESP_{max}^{Het}, ESP_{range}^{Het}, ESP_{σ}^{Het}, and F_{mean}^{Het} all capture the variation in the electrostatic potential due to the heterocycle in the region of the stacked amino acid side chain. Any of these descriptors, when paired with ESP_{mean}^{Het}, and N_{HA}^{Het} result in robust predictions of stacking interactions. Molecular dipole moments (μ) provide similar information, and a functional form analogous to equation 1 featuring μ instead of ESP_{range}^{Het} performs only slightly less well (see SI figure S2). However, the ESP and electric-field derived descriptors are arguably more versatile than dipole moments since they capture differences in electrostatic character among rings with zero net dipole (*e.g.* benzene, pyrazine, and 1,3,5-triazine).[51] In equation 2, ESP_{mean}^{Het} and ESP_{range}^{Het} are replaced by ESP_{max}^{Het}. This single electrostatic descriptor captures the fact that changing the maximum ESP can be accomplished independently by either shifting all ESP values while keeping the range constant or 'tilting' the ESP surface while keeping the mean ESP value constant. While we will focus on equations 1 and 2 below, the performance and interpretation of the other fits is similar.

Predicted interaction energies from equations 1 and 2 for the training set are plotted against the DLPNO-CCSD(T) interaction energies in Figures 3A and 3B, respectively. These fits are both accurate (RMSE of 0.7-0.8 kcal mol^{-1}) and robust to multiple levels of cross-validation, including Leave-One-Out (LOO), 10-fold, 5-fold and 3-fold (see Figure 3; full details of cross-validation are available in SI Table S1).

Chart 2. Heterocycle test set.



To probe the generality of equations 1 and 2, we considered an external test set of 12 additional heterocycles (see Chart 2). Stacking interaction energies of all unique stacked dimers of these heterocycles with model Phe, Tyr, and Trp side chains were evaluated as done for the training set (see SI Tables S10-S12). Figure 4 shows the interaction energies for the global minimum energy stacked dimers of the test set predicted using equations 1 and 2 versus DLPNO-CCSD(T) data. The predictions from equations 1 and 2 are highly correlated with the computed interaction energies for these 12 heterocycles and the RMSE values (0.5-0.6 kcal mol^{-1}) are slightly smaller than for the training set.

For both the training set and test set, equations 1 and 2 provide equally sound predictions across different classes of heterocycles. For instance, predicted interaction energies are reliable for both monocyclic and bicyclic heterocycles, as well as for heterocycles containing two or more heteroatoms compared to those with one or fewer heteroatoms. Plots of the data split across these categories can be seen in SI Figures S6 and S7.

Often, quantitative predictions of heterocycle stacking are secondary to simply ranking stacking ability. To assess the ability of equations 1 and 2 to provide reliable rankings, we considered Spearman's rank correlation coefficient (ρ^2) and Kendall τ. Both equations 1 and 2 accurately predict the ranked order of heterocycle stacking with Phe, Tyr, and Trp (see Figures 3 and 4); their performance is similarly strong when ranking the stacking energies of heterocycles with a single amino acid.

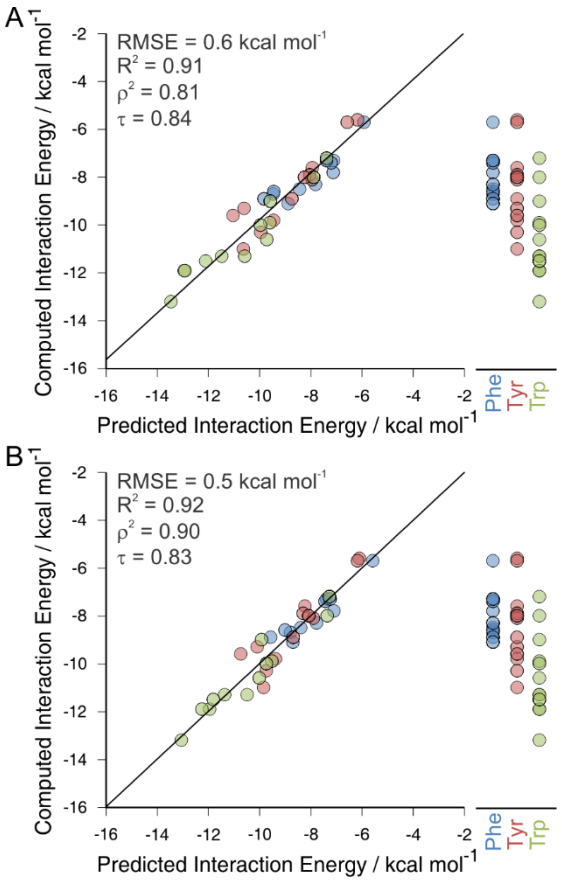


Figure 4. DLPNO-CCSD(T) interaction energies vs predicted interaction energies from equations 1 (A) and 2 (B) for global minimum energy stacked dimers of Phe, Tyr, and Trp side chains with the external test set (**47**-**58**). The strip plots on the right show the same computed interaction energies grouped by amino acid.

The robustness and generality of equations 1 and 2 indicates that the potential stacking interactions of a given heterocycle with Phe, Tyr, or Trp can be reliably predicted without resorting to expensive QM computations of stacked dimers. Instead, one needs to simply evaluate ESP_{mean}^{Het} and ESP_{mean}^{Het}or ESP_{max}^{Het}, which only require simple computations on the isolated heterocycles. The result is a ~10^5 reduction in computational cost but minimal loss in accuracy.

B. Conceptual Model of Stacking Interactions

The simplicity of equations 1 and 2 lends itself to interpretation of the factors that impact heterocycle stacking and the development of qualitative guidelines for maximizing these interactions in the context of design. The only descriptor of the amino acid side chain in these two models is the heavy atom count, N_{HA}^{AA}.[60] This simple descriptor (7 for Phe, 8 for Tyr, and 10 For Trp) scales the strength of stacking interactions to reflect the general observation that stacking interactions with Tyr and Trp are, on average, ~10% and ~40% stronger than with Phe, respectively. This enhanced stacking for the larger amino acid side chains could be a result of dispersion-driven interactions alone or could also be due in part to charge penetration effects.[61]

The heterocycle descriptors appearing in these equations (ESP_{mean}^{Het}, ESP_{range}^{Het}, ESP_{max}^{Het}, and N_{HA}^{Het}), on the other hand, confirm a prominent role of electrostatic effects (possibly including charge penetration effects)[61] and dispersion interactions, in accord with the SAPT data. Overall, each additional heavy atom in the heterocycle leads to an approximate 0.7, 0.8 and 1.0 kcal mol^{-1} increase in the stacking interaction with Phe, Tyr, and Trp, respectively. Regarding electrostatics, stacking interactions are enhanced in heterocycles with both overall positive ESP values and a large range of ESP values in the plane 3.25 Å from the heterocycle as well as those with large maximum ESP values in this plane. More precisely, equation 1 indicates that a 1 kcal mol^{-1} increase in ESP_{mean}^{Het} and ESP_{range}^{Het} provide ~0.3 and ~0.1 kcal mol^{-1} enhancement in stacking energies, whereas increasing ESP_{max}^{Het} by 1 kcal mol^{-1} should enhance stacking interactions by ~0.3 kcal mol^{-1}. It should be noted that ESP_{range}^{Het} and ESP_{max}^{Het} are more sensitive to the introduction of heteroatoms and other functionality than is ESP_{mean}^{Het} (*vide infra*), and modulating the values of any of these descriptors provide opportunities to tune the strength of stacking interactions over a broad range.

We can consider specific examples of heterocycles in order to gain a more intuitive understanding of these heterocycle descriptors and the impact of changes in the number and distribution of heteroatoms on stacking interactions (see Table 1). For instance, the impact of heterocycle size (as captured by N_{HA}^{Het}) is demonstrated by stacking interactions of pyrrole (**1**), indole (**27**), and carbazole (**43**). These three heterocycles have similar values for ESP_{mean}^{Het}, ESP_{range}^{Het}, and ESP_{max}^{Het}, but vary in size. Computed interaction energies for these three heterocycles stacked with Trp, for example, follow the trend in size, going from –4.5 kcal mol^{-1} for **1**, to –11.2 kcal mol^{-1} for **27**, finally to –14.2 kcal mol^{-1} for **43**.

Examination of the values for the electrostatic descriptors ESP_{mean}^{Het} ESP_{range}^{Het}, and ESP_{max}^{Het} across the heterocycles reveals clear trends that can be distilled into simple rules for their estimation, and thus the strength of stacking interactions. For example, ESP_{mean}^{Het} depends primarily on the number of S, O, imino N (-N=), amino N (-NH-), and carbonyl groups present in the heterocycle. Overall, each S, O, imino N, and carbonyl increases ESP_{mean}^{Het} by 0.5, 1.0, 2.0, and 3.5 kcal mol^{-1}, respectively. In contrast, each amino nitrogen decreases ESP_{mean}^{Het} by 1 kcal mol^{-1}. These effects are additive, as seen through the consecutive addition of imino nitrogens to benzene (**15**) to give pyridine (**16**), pyrimidine (**18**), and 1,2,4-triazine (**21**). The ESP_{mean}^{Het} values for these rings systematically increase from -4.7 to 1.1 kcal mol^{-1}, accompanied by an improvement in interaction energy with Trp from –6.0 kcal mol^{-1} for benzene to –9.9 kcal mol^{-1} for 1,2,4-triazine. This dependence of ESP_{mean}^{Het} on the number of these different heteroatoms is the origin of the general observation that electron deficient heterocycles (*e.g.* those containing many imino nitrogens) tend to engage in stronger stacking interactions,[4, 47-48] and ESP_{mean}^{Het} provides a readily-computed means of quantifying the degree of electron deficient character for a given heterocycle.

However, changes in ESP_{mean}^{Het} alone do not fully capture variations in the strength of stacking interactions with heterocycles. For example, despite the same number of heteroatoms in pyrazine (**17**), pyrimidine (**18**) and pyridazine (**19**), and their correspondingly similar ESP_{mean}^{Het} values (–0.8 ± 0.1 kcal mol^{-1}), the stacking interactions of these three rings with Trp vary from –8.1 to –9.5 kcal mol^{-1}. This can be attributed to the variation in ESP_{range}^{Het} across this series.

Unlike ESP_{mean}^{Het}, which depends only on the number of heteroatoms, the values of ESP_{range}^{Het} and ESP_{max}^{Het} depend on both the number of heteroatoms as well as their relative position within the heterocycle. Qualitatively, when a pair of 'like' heteroatoms (S, O, imino N, and C=O vs amino N) are on nearby positions, there is a cooperative impact on these two descriptors. For instance, while the addition of a single imino nitrogen to form pyridine (**16**) from benzene (**15**) shifts ESP_{range}^{Het} and ESP_{max}^{Het} by 2.6 and 3.6 kcal mol^{-1}, respectively, the introduction of a second, adjacent imino *N* to form pyridazine (**19**) has 2-3 times the effect. On the other hand, when like heteroatoms are on opposite sides of a ring, their impact on ESP_{range}^{Het} cancels to a large degree. This can be seen by comparing pyrimidine (**18**) and pyrazine (**17**), which both have smaller values of ESP_{range}^{Het} and ESP_{max}^{Het} than pyridine. In fact, ESP_{range}^{Het} for pyrazine is even smaller than for benzene. Similar trends hold for heterocycles containing any combination of S, O, amino N, and C=O groups.

For 'unlike' heteroatoms, the opposite trend occurs, with the impact of adjacent heteroatoms canceling and opposing heteroatoms exhibit positive cooperativity. This can be understood by recalling that amino and imino nitrogens have opposite impacts on the ESP, as demonstrated by the opposite sign of their impact on ESP_{mean}^{Het}. For example, for pyrazole (**8**), which has amino and imino nitrogens at adjacent positions, the value of ESP_{range}^{Het} (12.2 kcal mol^{-1}) is lower than that of pyrrole (**1**, ESP_{range}^{Het} = 14.7 kcal mol^{-1}). At the same time, the 1,3 amino-imino nitrogen pair in imidazole (**9**) leads to an ESP_{range}^{Het} value of 17.7 kcal mol^{-1}.

Thus, ESP_{range}^{Het} and ESP_{max}^{Het} are maximized when like heteroatoms and functional groups are grouped together on one side of the ring and the 'unlike' types on the opposite side of the ring. For example, the strongest stacking interaction predicted here is for 3,9-dihydro-purine-2,6-dione (**37**) stacked with Trp (see Figure 2). In this case, the presence of two C=O groups, including one adjacent to an imino nitrogen, overwhelms the modest effects of the three amino nitrogens, leading to a substantial values of ESP_{range}^{Het} and ESP_{max}^{Het}, a high ESP_{mean}^{Het}, and very strong stacking interaction. This tendency for heterocycles with certain heteroatom types clustered on one side to stack more strongly is exhibited by the antifolate drugs pemetrexed, aminopterin, and methotrexate, which bind to the human folate receptor α (hFRα) more weakly than folate.[18] Binding of these ligands is primarily modulated by the stacking of a terminal heterocyclic fragment with Trp and Tyr residues (See Figure 1C). The pteridin-4(3H)-one in folate features more imino nitrogens and carbonyl groups clustered together than either the pteridine in methotrexate and aminopterin or the 3,7-dihydro-4H-pyrrolo[2,3-d]pyrimidin-4-one in pemetrexed, which is expected to result in stronger stacking with both Trp and Tyr. While this is just a single example, it portends the use of this conceptual model to guide the judicious choice of heterocycle fragments to maximize stacking interactions with binding side aromatic groups in drug design.

Conclusions

We have developed simple predictive models (equations 1 and 2; see also SI Figures S2-S5) of the maximum stacking interactions between heterocycles found in biologically active molecules and the amino acid side chains Phe, Tyr, and Trp based on readily computed heterocycle descriptors. These models depend on the size of the heterocycle (as captured by the heavy-atom count),[60] as well as the electrostatic descriptors ESP_{mean}^{Het}, ESP_{range}^{Het}, and ESP_{max}^{Het}.[52] These latter three descriptors follow simple, predictable trends and can be modulated by varying the number and distribution of heteroatoms (including carbonyl groups) within the heterocycle. This provides a clear conceptual framework for understanding trends in stacking interactions between heterocycles and aromatic amino acid side chains and designing heterocycles that maximize such interactions.

Overall, stacking interactions with Phe, Tyr, and Trp are enhanced by:[62]

1. Increasing the numbers of S, O, imino N, and carbonyl groups, with the size of this effect increasing across this series;
2. Grouping 'like' heteroatoms (S, O, imino N, and carbonyl vs. amino N) on opposing sides of the heterocycle.

The former increases the mean ESP in the plane near the heterocycle while the latter enhances the variation in the ESP in this region, as captured by ESP_{range}^{Het} and ESP_{max}^{Het}. Using these guidelines, heterocycles can be qualitatively ranked in their ability to stack with Phe, Tyr, and Trp. For more quantitative predictions or rankings, the descriptors appearing in equations 1 and 2 can be rapidly evaluated using widely available computational tools.

Of course, in real biological systems stacking interactions rarely occur in isolation, and simple dimer models such as those presented above will not always provide a complete picture.[63] However, the conceptual and predictive models presented should provide a sound starting point for understanding and optimizing stacking interactions in complex environments. Similarly, the heterocycles considered here do not contain substituents, which is atypical for biologically active heterocycles. However, previous work[30, 33, 64-66] has demonstrated that substituent effects on stacking interactions are independent of the presence of heteroatoms on the substituted ring. Consequently, the vast body of previous work on substituent effects in stacking interactions[24-34] should be directly transferrable to the qualitative and quantitative models presented above. Finally, equations 1 and 2 concern the maximal possible stacking interactions exhibited by a given heterocycle, which will not necessarily be achievable in a given context. With these caveats in mind, this new view of stacking interactions between heterocycles and aromatic amino acid side chains should facilitate both a greater understanding of stacking in biological systems and provide updated principles[47] for maximizing stacking interactions in protein binding sites.

Computational Methods

Local energy minima were identified by systematically sampling six (or 12, for non-symmetric heterocycles) orientations of the heterocycle at each of nine initial points located in a 3 x 3 grid 3.6 Å above the heterocycle (fewer for Phe due to symmetry). Each of these starting configurations was optimized in the gas phase at the B97D/def2-TZVPP level of theory.[67-69] This dispersion-corrected DFT functional paired with a triple-ζ basis set should provide reliable geometries for these stacked dimers.[67] The resulting energy minima were then classified as either stacked or edge-to-face (T-shaped) geometries. Stacked geometries were those for which the centroid-centroid distance between the nearest two rings was less than 4.1 Å and the tilt angle between heterocycle planes was less

than 20°. For each heterocycle/amino acid combination, there were up to 24 unique stacked energy minima based on an RMSD cutoff of 0.4 Å. DLPNO-CCSD(T)/cc-pVQZ single point energies[70-74] were computed using 'Normal' PNO cutoffs for each unique stacked energy minimum to obtain accurate interaction energies (*i.e.* the energy difference between the optimized dimer and corresponding optimized monomers). Benchmark computations reveal that these DLPNO-CCSD(T) interaction energies are systematically 1.0 ± 0.3 kcal mol^{-1} from the corresponding canonical CCSD(T)/CBS values (see SI Table S2). Solution-phase interaction energies (in diethyl ether) were computed by combining the gas-phase DLPNO-CCSD(T) energies with solvation corrections computed at the ωB97X-D/def2-TZVPP level of theory.[75] While the above focus is on the global minimum energy stacked dimer for each heterocycle with the three amino acid side chain models, data for all local stacked energy minima are available in the SI Tables S4-S6. All DFT optimizations were performed using Gaussian09,[76] while Orca 4.0[77] was used for the DLPNO-CCSD(T) single point energies. SAPT computations were performed at the SAPT0/jun-cc-pVDZ level of theory[55-58, 78] using Psi4.[79]

ASSOCIATED CONTENT

Additional computational details and data, energies, Cartesian coordinates. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Email: swheele2@uga.edu.

Author Contributions

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

Funding Sources

National Science Foundation Grant CHE-1254897.

ACKNOWLEDGMENT

Portions of this research were conducted with resources provided by Texas A&M University (http://hprc.tamu.edu) and the Georgia Advanced Computing Resource Center (http://gacrc.uga.edu). We dedicate this paper to L. J. Kingsley and A. M. Westerman.

ABBREVIATIONS

DLPNO-CCSD(T), domain pair natural orbital coupled cluster theory with single and double excitations and perturbative triple excitations; cc-pVQZ, correlation consistent polarized valence quadruple-zeta; SAPT, symmetry-adapted perturbation theory; B97-D, Becke's 1997 exchange-correlation functional with an empirical dispersion correction.

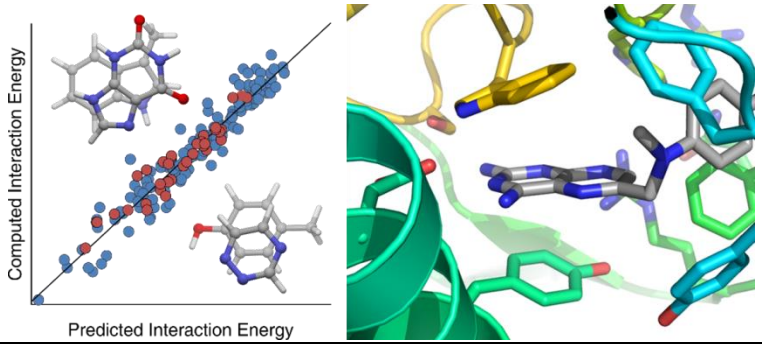
REFERENCES

1. Martinez, C. R.; Iverson, B. L., Rethinking the Term "π-Stacking". *Chem. Sci.* **2012,** *3*, 2191-2201.
2. Grimme, S., Do Special Noncovalent π-π Stacking Interactions Really Exist? *Angew. Chem. Int. Ed.* **2008,** *47*, 3430-3434.
3. Bloom, J. W.; Wheeler, S. E., Taking the aromaticity out of aromatic interactions. *Angew Chem Int Ed Engl* **2011,** *50* (34), 7847-9.
4. 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.
5. Schneider, H. J., Binding mechanisms in supramolecular complexes. *Angew Chem Int Ed Engl* **2009,** *48* (22), 3924-77.
6. Salonen, L. M.; Ellermann, M.; Diederich, F., Aromatic rings in chemical and biological recognition: energetics and structures. *Angew Chem Int Ed Engl* **2011,** *50* (21), 4808-42.
7. Wheeler, S. E.; Seguin, T. J.; Guan, Y.; Doney, A. C., Noncovalent Interactions in Organocatalysis and the Prospect of Computational Catalyst Design. *Acc. Chem. Res.* **2016,** *49* (5), 1061-9.
8. Neel, A. J.; Hilton, M. J.; Sigman, M. S.; Toste, F. D., Exploiting non-covalent pi interactions for catalyst design. *Nature* **2017,** *543* (7647), 637-646.
9. Thakuria, R.; Nath, N. K.; Saha, B. K., The Nature and Applications of π–π Interactions: A Perspective. *Cryst. Growth Des.* **2019,** *19* (2), 523-528.
10. Sutton, C.; Risko, C.; Brédas, J.-L., Noncovalent Intermolecular Interactions in Organic Electronic Materials: Implications for the Molecular Packing vs Electronic Properties of Acenes. *Chem. Mater.* **2015,** *28* (1), 3-16.
11. Schenkel, L. B.; Olivieri, P. R.; Boezio, A. A.; Deak, H. L.; Emkey, R.; Graceffa, R. F.; Gunaydin, H.; Guzman-Perez, A.; Lee, J. H.; Teffera, Y.; Wang, W.; Youngblood, B. D.; Yu, V. L.; Zhang, M.; Gavva, N. R.; Lehto, S. G.; Geuns-Meyer, S., Optimization of a Novel Quinazolinone-Based Series of Transient Receptor Potential A1 (TRPA1) Antagonists Demonstrating Potent in Vivo Activity. *J. Med. Chem.* **2016,** *59* (6), 2794-809.
12. Burley, S. K.; Petsko, G. A., Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **1985,** *229* (4708), 23-8.
13. Chelli, R.; Gervasio, F. L.; Procacci, P.; Schettino, V., Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *J. Am. Chem. Soc.* **2002,** *124* (21), 6133-43.
14. McGaughey, G. B.; Gagne, M.; Rappe, A. K., pi-Stacking interactions. Alive and well in proteins. *J. Biol. Chem.* **1998,** *273* (25), 15458-63.
15. Clark, M. P.; Ledeboer, M. W.; Davies, I.; Byrn, R. A.; Jones, S. M.; Perola, E.; Tsai, A.; Jacobs, M.; Nti-Addae, K.; Bandarage, U. K.; Boyd, M. J.; Bethiel, R. S.; Court, J. J.; Deng, H.; Duffy, J. P.; Dorsch, W. A.; Farmer, L. J.; Gao, H.; Gu, W.; Jackson, K.; Jacobs, D. H.; Kennedy, J. M.; Ledford, B.; Liang, J.; Maltais, F.; Murcko, M.; Wang, T.; Wannamaker, M. W.; Bennett, H. B.; Leeman, J. R.; McNeil, C.; Taylor, W. P.; Memmott, C.; Jiang, M.; Rijnbrand, R.; Bral, C.; Germann, U.; Nezami, A.; Zhang, Y.; Salituro, F. G.; Bennani, Y. L.; Charifson, P. S., Discovery of a Novel, First-in-Class, Orally Bioavailable Azaindole Inhibitor (VX-787) of Influenza PB2. *J. Med. Chem.* **2014,** *57* (15), 6668-6678.
16. Severin, C.; Rocha de Moura, T.; Liu, Y.; Li, K.; Zheng, X.; Luo, M., The cap-binding site of influenza virus protein PB2 as a drug target. *Acta Crystallographica Section D* **2016,** *72* (2), 245-253.
17. Ferreira de Freitas, R.; Harding, R. J.; Franzoni, I.; Ravichandran, M.; Mann, M. K.; Ouyang, H.; Lautens, M.; Santhakumar, V.; Arrowsmith, C. H.; Schapira, M., Identification and Structure–Activity Relationship of HDAC6 Zinc-Finger Ubiquitin Binding Domain Inhibitors. *J. Med. Chem.* **2018,** *61* (10), 4517-4527.
18. Wibowo, A. S.; Singh, M.; Reeder, K. M.; Carter, J. J.; Kovach, A. R.; Meng, W.; Ratnam, M.; Zhang, F.; Dann, C. E., 3rd, Structures of human folate receptors reveal biological trafficking states and diversity in folate and antifolate recognition. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110* (38), 15180-8.
19. Hudson, W. H.; Ortlund, E. A., The structure, function and evolution of proteins that bind DNA and RNA. *Nat. Rev. Mol. Cell Biol.* **2014,** *15*, 749.
20. Kuo, P.-H.; Chiang, C.-H.; Wang, Y.-T.; Doudeva, L. G.; Yuan, H. S., The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Res.* **2014,** *42* (7), 4712-4722.
21. Lee, W. G.; Gallardo-Macias, R.; Frey, K. M.; Spasov, K. A.; Bollini, M.; Anderson, K. S.; Jorgensen, W. L., Picomolar inhibitors of HIV reverse transcriptase featuring bicyclic replacement of a cyanovinylphenyl group. *J. Am. Chem. Soc.* **2013,** *135* (44), 16705-13.
22. Kim, J. T.; Hamilton, A. D.; Bailey, C. M.; Domaoal, R. A.; Wang, L.; Anderson, K. S.; Jorgensen, W. L., FEP-guided selection of bicyclic

heterocycles in lead optimization for non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J. Am. Chem. Soc.* **2006,** *128* (48), 15372-3.
23. Hunter, C. A.; Sanders, J. K. M., The Nature of π-π Interactions. *J. Am. Chem. Soc.* **1990,** *112*, 5525-5534.
24. Sinnokrot, M. O.; Sherrill, C. D., Substituent Effects in π-π Interactions: Sandwich and T-Shaped Configurations. *J. Am. Chem. Soc.* **2004,** *126*, 7690-7697.
25. Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J., Electrostatic control of aromatic stacking interactions. *J. Am. Chem. Soc.* **2005,** *127* (24), 8594-5.
26. Cockroft, S. L.; Perkins, J.; Zonta, C.; Adams, H.; Spey, S. E.; Low, C. M.; Vinter, J. G.; Lawson, K. R.; Urch, C. J.; Hunter, C. A., Substituent effects on aromatic stacking interactions. *Org. Biomol. Chem.* **2007,** *5* (7), 1062-80.
27. Wheeler, S. E.; Houk, K. N., Substituent effects in the benzene dimer are due to direct interactions of the substituents with the unsubstituted benzene. *J. Am. Chem. Soc.* **2008,** *130* (33), 10854-5.
28. Ringer, A. L.; Sherrill, C. D., Substituent Effects in Sandwich Configurations of Multiply Substituted Benzene Dimers Are Not Solely Governed By Electrostatic Control. *J. Am. Chem. Soc.* **2009,** *131*, 4574-4575.
29. Cockroft, S. L.; Hunter, C. A., Desolvation and substituent effects in edge-to-face aromatic interactions. *Chem Commun (Camb)* **2009,** (26), 3961-3.
30. Wheeler, S. E., Local nature of substituent effects in stacking interactions. *J. Am. Chem. Soc.* **2011,** *133* (26), 10262-74.
31. Watt, M.; Hardebeck, L. K. E.; Kirkpatrick, C. C.; Lewis, M., Face-to-Face Arene-Arene Bidning Energies: Dominated by Dispersion but Predicted by Electrostatic and Dispersion/Polarizability Substituent Constants. *J. Am. Chem. Soc.* **2011,** *133*, 3854-3862.
32. Muchowska, K. B.; Adam, C.; Mati, I. K.; Cockroft, S. L., Electrostatic modulation of aromatic rings via explicit solvation of substituents. *J. Am. Chem. Soc.* **2013,** *135* (27), 9976-9.
33. Wheeler, S. E.; Bloom, J. W., Toward a more complete understanding of noncovalent interactions involving aromatic rings. *J. Phys. Chem. A* **2014,** *118* (32), 6133-47.
34. Hwang, J.; Li, P.; Carroll, W. R.; Smith, M. D.; Pellechia, P. J.; Shimizu, K. D., Additivity of substituent effects in aromatic stacking interactions. *J. Am. Chem. Soc.* **2014,** *136* (40), 14060-7.
35. Mignon, P.; Loverix, S.; De Proft, F.; Geerlings, P., Influence of Stacking on Hydrogen Bonding: Quantum Chemical Study on Pyridine−Benzene Model Complexes. *J. Phys. Chem. A* **2004,** *108* (28), 6038-6044.
36. Piacenza, M.; Grimme, S., Van der Waals interactions in aromatic systems: structure and energetics of dimers and trimers of pyridine. *ChemPhysChem* **2005,** *6* (8), 1554-8.
37. Swart, M.; van der Wijst, T.; Fonseca Guerra, C.; Bickelhaupt, F. M., Pi-pi stacking tackled with density functional theory. *J. Mol. Model.* **2007,** *13* (12), 1245-57.
38. Itahara, T.; Imaizumi, K., Role of nitrogen atom in aromatic stacking. *J. Phys. Chem. B* **2007,** *111* (8), 2025-32.
39. Hohenstein, E. G.; Sherrill, C. D., Effects of heteroatoms on aromatic pi-pi interactions: benzene-pyridine and pyridine dimer. *J. Phys. Chem. A* **2009,** *113* (5), 878-86.
40. Geng, Y.; Takatani, T.; Hohenstein, E. G.; Sherrill, C. D., Accurately characterizing the pi-pi interaction energies of indole-benzene complexes. *J. Phys. Chem. A* **2010,** *114* (10), 3576-82.
41. Karthikeyan, S.; Nagase, S., Origins of the stability of imidazole-imidazole, benzene-imidazole, and benzene-indole dimers: CCSD(T)/CBS and SAPT calculations. *J. Phys. Chem. A* **2012,** *116* (7), 1694-700.
42. Li, P.; Zhao, C.; Smith, M. D.; Shimizu, K. D., Comprehensive experimental study of N-heterocyclic pi-stacking interactions of neutral and cationic pyridines. *J. Org. Chem.* **2013,** *78* (11), 5303-13.
43. Jacobs, M.; Greff Da Silveira, L.; Prampolini, G.; Livotto, P. R.; Cacelli, I., Interaction Energy Landscapes of Aromatic Heterocycles through a Reliable yet Affordable Computational Approach. *J. Chem. Theory Comput.* **2018,** *14* (2), 543-556.
44. Gung, B. W.; Wekesa, F.; Barnes, C. L., Stacking interactions between nitrogen-containing six-membered heterocyclic aromatic rings and substituted benzene: studies in solution and in the solid state. *J. Org. Chem.* **2008,** *73* (5), 1803-8.
45. Tyagarajan, S.; Lowden, C. T.; Peng, Z.; Dykstra, K. D.; Sherer, E. C.; Krska, S. W., Heterocyclic Regioisomer Enumeration (HREMS): A Cheminformatics Design Tool. *J. Chem. Inf. Model.* **2015,** *55* (6), 1130-5.
46. Mok, N. Y.; Brown, N., Applications of Systematic Molecular Scaffold Enumeration to Enrich Structure-Activity Relationship Information. *J. Chem. Inf. Model.* **2017,** *57* (1), 27-35.
47. Bissantz, C.; Kuhn, B.; Stahl, M., A medicinal chemist's guide to molecular interactions. *J. Med. Chem.* **2010,** *53* (14), 5061-84.
48. Gonthier, J. F.; Steinmann, S. N.; Roch, L.; Ruggi, A.; Luisier, N.; Severin, K.; Corminboeuf, C., pi-Depletion as a criterion to predict pi-stacking ability. *Chem Commun (Camb)* **2012,** *48* (74), 9239-41.
49. Huber, R. G.; Margreiter, M. A.; Fuchs, J. E.; von Grafenstein, S.; Tautermann, C. S.; Liedl, K. R.; Fox, T., Heteroaromatic pi-stacking energy landscapes. *J. Chem. Inf. Model.* **2014,** *54* (5), 1371-9.
50. An, Y.; Doney, A. C.; Andrade, R. B.; Wheeler, S. E., Stacking Interactions between 9-Methyladenine and Heterocycles Commonly Found in Pharmaceuticals. *J. Chem. Inf. Model.* **2016,** *56* (5), 906-914.
51. Bootsma, A. N.; Wheeler, S. E., Stacking Interactions of Heterocyclic Drug Fragments with Protein Amide Backbones. *ChemMedChem* **2018,** *13* (8), 835-841.
52. Bootsma, A. N.; Wheeler, S. E., Tuning Stacking Interactions between Asp-Arg Salt Bridges and Heterocyclic Drug Fragments. *J. Chem. Inf. Model.* **2019,** *59* (1), 149-158.
53. Vitaku, E.; Smith, D. T.; Njardarson, J. T., Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *J. Med. Chem.* **2014,** *57* (24), 10257-10274.
54. Broughton, H. B.; Watson, I. A., Selection of heterocycles for drug design. *J. Mol. Graph. Model.* **2004,** *23* (1), 51-58.
55. Jeziorski, B.; Moszynski, R.; Szalewicz, K., Perturbation-Theory Approach to Intermolecular Potential-Energy Surfaces of Van-Der-Waals Complexes. *Chem. Rev.* **1994,** *94* (7), 1887-1930.
56. Szalewicz, K., Symmetry-adapted perturbation theory of intermolecular forces. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2012,** *2* (2), 254-272.
57. Hohenstein, E. G.; Sherrill, C. D., Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2010,** *133* (1), 014101.
58. Hohenstein, E. G.; Sherrill, C. D., Density fitting and Cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of pi-pi interactions in linear acenes. *J. Chem. Phys.* **2010,** *132* (18), 184111.
59. Bloom, J. W.; Raju, R. K.; Wheeler, S. E., Physical Nature of Substituent Effects in XH/pi Interactions. *J. Chem. Theory Comput.* **2012,** *8* (9), 3167-74.
60. Sanders, J. M., Optimal pi-stacking interaction energies in parallel-displaced aryl/aryl dimers are predicted by the dimer heavy atom count. *J. Phys. Chem. A* **2010,** *114* (34), 9205-11.
61. Gryn'ova, G.; Corminboeuf, C., Implications of Charge Penetration for Heteroatom-Containing Organic Semiconductors. *J. Phys. Chem. Lett.* **2016,** *7* (24), 5198-5204.
62. The above discussion is based on gas-phase stacking interactions. However, we have also computed stacking interaction energies for all unique dimers in diethyl ether (ε = 4.24) as a model of a typical protein environment and refit equation 1 to these data. While the correlation are not quite as strong as seen for the gas-phase data (R^2 = 0.82), the overall trends remain unchanged (see SI Figure S8 and S9) and the guidelines presented should apply to stacking interactions in both the gas-phase and dielectric environments.
63. Gryn'ova, G.; Nicolai, A.; Prlj, A.; Ollitrault, P.; Andrienko, D.; Corminboeuf, C., Charge transport in highly ordered organic nanofibrils: lessons from modelling. *Journal of Materials Chemistry C* **2017,** *5* (2), 350-361.
64. Raju, R. K.; Bloom, J. W.; An, Y.; Wheeler, S. E., Substituent effects on non-covalent interactions with aromatic rings: insights from computational chemistry. *ChemPhysChem* **2011,** *12* (17), 3116-30.
65. Raju, R. K.; Bloom, J. W.; Wheeler, S. E., Broad Transferability of Substituent Effects in pi-Stacking Interactions Provides New Insights into Their Origin. *J. Chem. Theory Comput.* **2013,** *9* (8), 3479-90.

66. Wheeler, S. E., Understanding substituent effects in noncovalent interactions involving aromatic rings. *Acc. Chem. Res.* **2013,** *46* (4), 1029-38.
67. Grimme, S., Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006,** *27* (15), 1787-99.
68. Becke, A., Density-Functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *J. Chem. Phys.* **1997,** *107*, 8554-8560.
69. 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.
70. Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F., Sparse maps-A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015,** *143* (3), 034108.
71. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F., Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013,** *139* (13), 134101.
72. Riplinger, C.; Neese, F., An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013,** *138* (3), 034106.
73. Neese, F.; Hansen, A.; Liakos, D. G., Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009,** *131* (6), 064103.
74. Dunning, T. H., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989,** *90* (2), 1007-1023.
75. Chai, J.-D.; Head-Gordon, M., Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008,** *10* (44), 6615-6620.
76. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.; Izmaylov, A.; Bloino, J.; Zheng, G.; Sonnenberg, J.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., JA; Peralta, J.; Ogliaro, F.; Bearpark, M.; Heyd, J.; Brothers, E.; Kudin, K.; Staroverov, V.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.; Iyengar, S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.; Klene, M.; Knox, J.; Cross, J.; Bakken, V.; Adamo, C.; ramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; Martin, R.; Morokuma, K.; Zakrzewski, V.; Voth, G.; Salvador, P.; Dannenberg, J.; Dapprich, S.; Daniels, A.; Farkas, O.; Foresman, J.; Ortiz, J.; Cioslowski, J.; Fox, D. *Gaussian 09, Revision D.01*, Gaussian, Inc.: 2009.
77. Neese, F., Software update: the ORCA program system, version 4.0. *WIREs Comput. Mol. Sci.* **2018,** *8* (1).
78. Papajak, E.; Truhlar, D. G., Convergent Partially Augmented Basis Sets for Post-Hartree–Fock Calculations of Molecular Properties and Reaction Barrier Heights. *J. Chem. Theory Comput.* **2011,** *7*, 10-18.
79. Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D., PSI4: an open-source ab initio electronic structure program. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2012,** *2* (4), 556-565.

TOC Graphic


Computed Interaction Energy
Predicted Interaction Energy