

# Interaction energy landscapes of aromatic heterocycles through a reliable yet affordable computational approach.

Matheus Jacobs,<sup>a</sup> Leandro Greff Da Silveira,<sup>a,b</sup> Giacomo Prampolini,<sup>c,\*</sup>  
Paolo Roberto Livotto,<sup>a</sup> and Ivo Cacelli<sup>c,d</sup>

<sup>a</sup>*Instituto de Química, Universidade Federal do Rio Grande do Sul,  
Avenida Bento Gonçalves 9500, CEP 91501-970 Porto Alegre, Brazil*

<sup>b</sup>*Departamento de Ciências Exatas e da Terra,  
Universidade Regional Integrada do Alto Uruguay e da Missões (URI),  
Avenida Assis Brasil 709, CEP 98400-00 Frederico Westphalen, Brazil*

<sup>c</sup>*Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR),  
Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy*

<sup>d</sup>*Dipartimento di Chimica e Chimica Industriale, Università di Pisa,  
Via G. Moruzzi 3, I-56124 Pisa, Italy*

August 29, 2017

---

\*Corresponding author

## Abstract

Noncovalent interactions between homo-dimers of several aromatic heterocycles (pyrrole, furan, thiophene, pyridine, pyridazine, pyrimidine and pyrazine) are investigated at *ab initio* level, employing the Möller-Plesset second order perturbation theory, coupled with small Gaussian basis sets (6-31G\* and 6-31G\*\*) with specifically tuned polarization exponents. The latter are modified using a systematic and automated procedure, the MP2<sup>mod</sup> approach, based on a comparison with high level CCSD(T) calculations extrapolated to complete basis set. The MP2<sup>mod</sup> results achieved with the modified 6-31G\*\* basis set show an excellent agreement with CCSD(T)/CBS reference energies, with a standard deviation less than 0.3 kcal/mol. Exploiting its low computational cost, the MP2<sup>mod</sup> approach is then used to explore sections of the intermolecular energy of the considered homo-dimers, with the aim of rationalizing the results. It is found that the direct electrostatic interaction between the monomers electron clouds is at the origin of some observed features, and in many cases multipoles higher than dipole play a relevant role, although often the interplay with other contributions to the noncovalent forces (as for instance induction,  $\pi$ - $\pi$  or XH- $\pi$  interactions) makes rather difficult a simple rationalization.

# 1 Introduction

Noncovalent interactions between molecular species bearing aromatic moieties play a key role in many fields of chemistry, materials science and biochemistry,<sup>1</sup> and are thought to be crucial to different cutting-edge applications, as molecular recognition,<sup>2</sup> base stacking in DNA,<sup>3</sup> advanced organic photovoltaic,<sup>4,5</sup> novel drugs design,<sup>6,7</sup> just to cite a few examples. Yet, from a more fundamental point of view, a widely applicable rationale on their origin and peculiar features is still argument of debate.<sup>1,8-17</sup> In particular, the role of the contribution of the the  $\pi$  system to the total interaction energy, through the so called<sup>1,2,16</sup>  $\pi$ -stacking,  $\pi$ -ion or XH- $\pi$  interactions, is far from being completely understood. The well known Hunter-Sanders electrostatic model<sup>18</sup> has succeeded in predicting the relative stability and the effects of electron withdrawing/donating substituents in many aromatic pairs in the face-to-face (FF), T-shaped (TS) and parallel displaced (PD) arrangements. In a subsequent paper,<sup>19</sup> however, Hunter and coworkers noticed that, beside electrostatic effects discussed in the original paper, the direct interaction with substituents linked to a ring also give a relevant contribution. On the same foot, several failures of the predictions based on the Hunter-Sanders model can be corrected resorting to the work of Wheeler and coworkers,<sup>13,16</sup> who suggest that substituent effects in aromatic moieties can be explained in term of direct interaction of the functional group with the aromatic ring, rather than mediated by the aromatic  $\pi$  systems. Notwithstanding some differences, both Hunter and Sanders and Wheeler’s models agree in indicating the delicate interplay among electrostatics, dispersion, exchange/repulsion and direct interactions as responsible for most of the peculiar features of aromatic interactions, including their dependence on the relative spatial disposition of the interacting molecules and/or on the chemical nature of their substituents.

Despite most of these important conclusions have been deduced essentially from accurate calculations, carried out on a small number of selected dimer geometries,<sup>9,10,13,14,16,17,20-22</sup> it is evident that extensive explorations of the interaction energy surface (IPES) could provide new insights on aromatic non-covalent interactions. Unfortunately, a reliable IPES representation still remains a challenge, even for small aromatic dimers. In fact, on the one hand the complex dependence of the aromatic interactions upon orientation

requires sampling a rather large number of geometries, on the other hand the aforementioned interplay between the different contributions to the total interaction energy calls for an accurate and well balanced calculation method. Due to the important role expected from the dispersion contribution, a reliable energy estimate can be obtained only resorting to postHF methods,<sup>23-25</sup> in a wavefunction (WF) based framework, or accounting for dispersion through some correction schemes,<sup>26-28</sup> if density functional theory (DFT) is exploited. Although DFT methods are in general computationally more convenient, it has been shown<sup>16,24,29-40</sup> that the choice of the functional, the basis set and the correction scheme for dispersion have a remarkable impact on the accuracy of the final estimates. This would require a careful benchmark of several dispersion corrected (DFT-D) functionals against some high-level data, possibly considering a representative sample of dimer arrangements.<sup>39</sup> In this work we chose to follow a different strategy, based on WF based approach, with the aim of obtaining accurate results at a reasonable computational cost, at least comparable with that required by standard DFT methods.

Within a WF framework, one of the most reliable methods is the Coupled Cluster with Single, Double and perturbatively included connected Triple excitations, CCSD(T). Its accuracy can be further increased by extrapolating the results to the Complete Basis Set limit (CBS) according to slightly different schemes, separately proposed by the groups of Tsuzuki,<sup>21</sup> Hobza<sup>8</sup> and Sherrill.<sup>22</sup> For noncovalent interactions, CCSD(T)/CBS is nowadays considered the gold standard of quantum chemistry.<sup>24</sup> Yet, its computational demand limits the number of calculations to few geometries and does not allow for the exploration of large portions of the IPES, even for systems with less than 100 electrons as for those here considered. The second order Möller-Plesset perturbation theory (MP2) is computationally more convenient, but it is well-known<sup>22,23,31,36</sup> that severely overestimates the aromatic interaction energies, in particular when large basis sets are used. In order to circumvent this problem, much attention has been recently devoted<sup>23,37,41-43</sup> to the search for more effective basis sets, capable to deliver accurate results without renouncing to MP2 computational convenience. Within this framework, resorting to the pioneering work of Kroon-Batenburg and Van Duijneveldt,<sup>44</sup> Hobza and coworkers<sup>8,20,45,46</sup> noticed that the interaction energies of the benzene dimer were generally overestimated by MP2,

but a significant improvement with respect to reference CCSD(T)/CBS data<sup>21, 22, 24, 25</sup> could be obtained by using suitably modified small basis sets. The modified set for benzene was labeled 6-31G\*(0.25), because the modification consisted in decreasing to 0.25 the value of the exponent of the polarization  $d$  function of the Carbon atoms in the 6-31G\* basis set. Following this idea, the procedure was extended to other basis sets (as 6-31G\*\* or cc-pvDz) and polarization exponents other than those on Carbon were modified, obtaining rather accurate results for a relatively wide range of compounds, as for instance DNA bases,<sup>45, 47–52</sup> liquid crystal forming molecules,<sup>53–56</sup> Nitrogen containing hetero-cycles<sup>37, 50, 57</sup> and, most recently, the quinhydrone dimer and eumelanin building blocks,<sup>42, 43</sup> where the name MP2<sup>mod</sup> was first proposed.

In this work the MP2<sup>mod</sup> approach will be adopted in the calculation of the interaction energy landscapes as well as in the identification of some relevant local minima of aromatic homo-dimers. The first aim is to set up and validate an automated procedure to tune the MP2<sup>mod</sup> method. To this end, specific exponents of the polarization functions in small basis sets (6-31G\* and 6-31G\*\*) are optimized against reference CCSD(T)/CBS calculations, purposely carried out for the selected targets. Once the best-fit modified basis sets are obtained and carefully validated, the MP2<sup>mod</sup> computational convenience can be exploited to compute and analyze the molecular IPESs for each investigated species. Indeed, the second goal of this work is to investigate and possibly rationalize the changes in the interaction patterns triggered by different heteroatoms (N, O and S), and how these could drive the most favorable disposition and orientation of two interacting rings.

Although many computational studies have been reported for the smallest prototype of aromatic interactions (i.e. the benzene dimer)<sup>20–22, 58–61</sup> and an increasing number of reports has been recently dedicated to the effect of substituents on the aromatic ring,<sup>1, 8–10, 12–17</sup> surprisingly much less attention has been dedicated to the effect of the substitution within the aromatic skeleton. To the best of our knowledge only few heterocyclic molecules have been studied such as for instance pyrrole,<sup>40, 62–64</sup> furan,<sup>65</sup> thiophene,<sup>66–68</sup> pyridine,<sup>14, 16, 30, 69–73</sup> pyrazine<sup>71, 74, 75</sup> and pyrimidine.<sup>76</sup> Nonetheless, a systematic comparison of large portions of dimer IPESs, carried out at the same level of accuracy among a selected sample of different homo-dimers, seems to be missing. Only very recently,<sup>38</sup> a

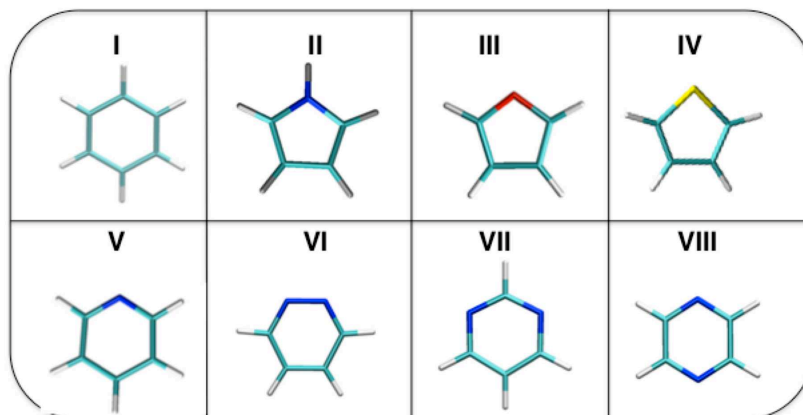


Figure 1: Target molecules. I) benzene; II) pyrrole; III) furan; IV) thiophene; V) pyridine; VI) pyridazine; VII) pyrimidine; VIII) pyrazine. Standard atom colors (white for H, cyan for C, blue for N, red for O and yellow for S) have been used in all figures throughout the text.

systematic study involving heteroaromatic landscapes has been reported, where a number of hetero-dimers, made up by stacking the benzene molecule with a five- or six-membered heteroaromatic ring, was investigated at CCSD(T)/CBS and at DFT-D level. Despite the wealth of information retrieved, two important features still call for further investigation. First, the previous studies were often limited to stacked structures, and the possible competition with other arrangements (such as T-shaped ones) has been seldom considered. Furthermore, the systematic discussion proposed in Ref.,<sup>38</sup> should be extended to homo-dimers, that is to pairs of aromatic rings, each bearing at least one heteroatom. To fill this lack, in this work the stacked and T-shaped conformational landscapes of eight aromatic homo-dimers (namely, benzene, pyrrole, furan, thiophene, pyridine, pyridazine, pyrimidine and pyrazine), displayed in Figure 1, are explored and compared.

## 2 Methods

Unless otherwise stated, all the calculations were performed with the GAUSSIAN09 suite of programs.<sup>77</sup> The equilibrium geometry of each isolated monomer was obtained by geometry optimization at DFT level, using the B3LYP functional with the Dunning's

correlation consistent cc-pVTZ basis set.

The reference dimer interaction energies, employed for both the tuning of the MP2<sup>mod</sup> exponents and the validation of the MP2<sup>mod</sup> local minima, were computed at CCSD(T)/CBS level. In most of its current implementations,<sup>25</sup> the CCSD(T)/CBS scheme exploits the empirical observation that the difference ( $\Delta^{CCSD(T)-MP2}$ ) between CCSD(T) and MP2 interaction energies ( $\Delta E$ ), evaluated for medium to large basis sets ( $X$ ), is almost constant:

$$\Delta^{CCSD(T)-MP2} = \Delta E_X^{CCSD(T)} - \Delta_X^{MP2} \quad (1)$$

The CCSD(T) interaction energy at the complete basis set (CBS) limit ( $\Delta E_{CBS}^{CCSD(T)}$ ) was then retrieved as

$$\Delta E_{CBS}^{CCSD(T)} = \Delta E_{CBS}^{MP2} + \Delta^{CCSD(T)-MP2} \quad (2)$$

where the MP2 interaction energy at CBS ( $\Delta E_{CBS}^{MP2}$ ) was estimated through an extrapolation scheme. In this work, the Halkier extrapolation scheme<sup>78</sup> was employed to estimate  $\Delta E_{CBS}^{MP2}$ , making use of the augmented *aug-cc-pvDz* and *aug-cc-pvTz* basis sets, whereas the  $\Delta^{CCSD(T)-MP2}$  difference was estimated at *aug-cc-pvDz* level. In all CCSD(T) and MP2 calculations the standard Counterpoise (CP) correction<sup>79</sup> was applied to take care of the basis set superposition error (BSSE). The modified basis sets employed in the MP2<sup>mod</sup> calculations of the conformational landscapes were tuned specifically for each target dimer through the following procedure:

1. Based on Hunter-Sanders predictions and/or on the available literature results, four representative dimer arrangements (*e.g.* FF, TS, PD, etc.) are identified.
2. Interaction energy curves are preliminary computed, at MP2<sup>mod</sup>/6-31G\*(0.25) level, by displacing the monomers from the formerly identified arrangements along a chosen coordinate (see the insets of Figures 2, 3 or 4). During the scan, the internal geometry of each monomer is kept fixed.
3. Three different dimer geometries are selected along each resulting curve: one in the short-range, the second around the energy minimum, and the last in the long-

range. For each of the twelve sampled configurations, the dimer interaction energy is computed at CCSD(T)/CBS level and included in the reference database.

4. The optimization of the exponent of the polarization functions for MP2<sup>mod</sup> is carried out with the EXOPT code<sup>42,43,55,80</sup> using the aforementioned CCSD(T)/CBS database.

Two different MP2<sup>mod</sup> basis sets were tuned for each target molecule. In the first case the exponents of the *d* functions on (N, O, S) in the 6-31G\* basis set were optimized, while keeping the analogous exponent for the Carbon fixed at 0.25. This modified basis set was called 6-31G\*(0.25,  $\alpha_X$ ),  $X=N,O,S$ . In the larger 6-31G\*\* basis set, which also includes *p* polarization functions on hydrogen atoms, all three  $\alpha_C$ ,  $\alpha_X$  and  $\alpha_H$  exponents were modified, yielding the 6-31G\*\*( $\alpha_C$ ,  $\alpha_X$ ,  $\alpha_H$ ) basis sets.

Finally, to evaluate the performances of the MP2<sup>mod</sup> strategy with respect to reference CCSD(T)/CBS interaction energies, two different indicators have been employed, as recently suggested by Řezáč and Hobza:<sup>25</sup> the root mean squared error,

$$RMSE = \left[ \frac{1}{N_{geom}} \sum_{k=1}^{N_{geom}} \left( \Delta E^{MP2_k^{mod}} - \Delta E^{CCSD(T)_k} \right)^2 \right]^{\frac{1}{2}} \quad (3)$$

and the largest unsigned error (LUE).

## 3 Results

### 3.1 MP2<sup>mod</sup> tuning

#### *Benzene*

Despite it does not contain any heteroatom, the benzene dimer, the simplest prototype of aromatic interaction, has been included in the present work for two reasons. First, it might be employed as a useful reference to evaluate the effect of the heteroatom insertion within the monomer rings. Next, the large number of available literature data concerning benzene dimers can be exploited to evaluate the accuracy of CCSD(T)/CBS protocol here employed for building reference data. In particular, one can resort to the CCSD(T)/CBS curves, reported by Sherrill and coworkers,<sup>61,81</sup> for four selected dimer arrangements,



namely FF, TS and two PD geometries (see Figure 2 insets). With respect to the present work, Sherrill’s group adopted<sup>61,81</sup> a slightly different and more accurate procedure to carry out CCSD(T)/CBS extrapolation. Indeed, as detailed in the original paper,<sup>61</sup> larger basis sets have been employed for both the MP2 extrapolation at CBS,  $\Delta E_{CBS}^{MP2}$  in equation (2), and the  $\Delta^{CCSD(T)-MP2}$  correction (1): the former was obtained through the Halkier scheme applied to the *aug-cc-pvTZ* and *aug-cc-pvQZ*, while  $\Delta E_{CBS}^{MP2}$  was computed at *aug-cc-pvTZ* level.

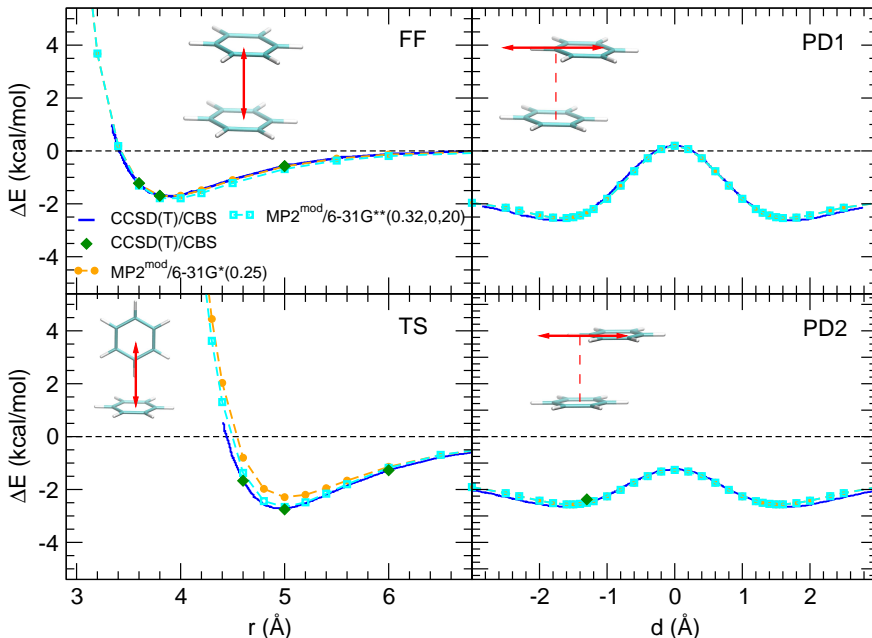


Figure 2: Results of  $MP2^{mod}$  calculations, carried out with the 6-31G\*(0.25) and 6-31G\*\*(0.32,0.20) basis sets (orange and cyan curves, respectively), on the four representative benzene dimer arrangements displayed in the insets. CCSD(T)/CBS reference data, either reported in literature<sup>61</sup> (blue curves) or computed in this work (green diamonds), are also shown for comparison. In PD1 and PD2 arrangements, the separation between the planes containing the two rings (dashed lines in the insets) were set to 3.4 Å and 3.6 Å, respectively.

To evaluate the accuracy of the CBS extrapolation procedure adopted herein, the interaction energy calculations for seven geometries selected along the curves, were per-

formed at CCSD(T)/CBS with the protocol described in the Methods section and compared to literature values reported by Sherrill and coworkers for the same conformers.<sup>61,81</sup> By looking at Figure 2, it appears that the differences in the CCSD(T) extrapolation schemes to CBS do not cause significant alterations in the expected values: indeed, the RMSE and LUE values reported in Table 1, indicate that the two procedures differ by few hundredths of kcal/mol. These results prompted us to confidently employ our CBS extrapolation scheme for all compounds here considered.

Next, MP2<sup>mod</sup> curves were computed for a large number of inter-ring separations for each of the considered arrangements. For benzene, the original 6-31G\*(0.25) basis set was first employed. The results are shown in Figure 2, whereas the computed interaction energies for selected conformations are reported in detail in Table 1. It is apparent

Type	R (Å)	$\Delta E^{CCSD(T)}$		$\Delta E^{MP2^{mod}}$	
		Ref. [61]	this work	6-31G*(0.25)	6-31G**(0.32,0.20)
FF	3.6	-1.25	-1.22	-1.20	-1.31
	3.8	-1.68	-1.70	-1.70	-1.79
	5.0	-0.55	-0.58	-0.50	-0.68
TS	4.6	-1.69	-1.67	-0.80	-1.37
	5.0	-2.70	-2.75	-2.29	-2.65
	6.0	-1.27	-1.27	-1.14	-1.18
PD	-1.3	-2.30	-2.37	-2.34	-2.37
RMSE		0.04	0.00	0.38	0.14
LUE		0.05	0.00	0.87	0.30

Table 1: Results of CCSD(T)/CBS and MP2<sup>mod</sup> calculations performed on the representative benzene dimer arrangements displayed in the insets of Figure 2 with the 6-31G\*(0.25) and 6-31G\*\*(0.32,0.20) basis sets. CCSD(T)/CBS values were either reported in literature<sup>61</sup> (third column) or computed in this work (fourth column). All interaction energies, RMSE and LUE are reported in kcal/mol.

that MP2<sup>mod</sup>/6-31G\*(0.25) calculations yields rather accurate predictions with respect to reference values, in particular for stacked conformations. The agreement is somewhat worse for TS arrangements, where the TS minimum is almost isoenergetic with the PD one. Indeed, most of the contribution to the RMSE (0.38 kcal/mol) comes from two TS arrangements, namely at short (0.87 kcal/mol) and intermediate (0.44 kcal/mol) distances. At difference with the stacked configurations, where the main contribution resides in  $\pi$ - $\pi$  interactions, in the TS arrangements the CH- $\pi$  interaction between the Hydrogen pointing

toward the neighboring aromatic plane should play a predominant role. Therefore it is interesting to verify if such inaccuracies could be reduced if polarization effects are also considered on the Hydrogen atoms. In this case, of course, it is possible that the best value for the exponent of the Carbon polarization  $d$  functions deviates from the 0.25 value often employed in literature.

Following these considerations, the computed CCSD(T)/CBS data were employed in the EXOPT procedure, and the exponents of the  $d$  and  $p$  polarization functions on C and H atoms of the 6-31G\*\* basis sets were tuned. The best-fit exponents were  $\alpha_C = 0.32$  and  $\alpha_H = 0.20$ . Both a visual inspection of the MP2<sup>mod</sup>/6-31G\*\*(0.32, 0.20) curves displayed in Figure 2 and the RMSE and LUE reported in Table 1 indicate that a sensibly better agreement with the reference data is achieved: the addition of tuned  $p$  polarization functions leads to more accurate results in the TS arrangements, whereas does not significantly disturb those obtained for stacked conformations.

### ***Five-membered rings***

Considering the promising results achieved for benzene, the MP2<sup>mod</sup> exponent parameterization was extended to the whole investigated sample, displayed in Figure 1. In order to explore the possibility that such exponents could have a certain degree of transferability to the same atom within different molecules, the small 6-31G\* basis set was again first employed, taking the polarization exponent of Carbon fixed to 0.25. Conversely, in the larger 6-31G\*\* basis set this constraint was released and all exponents were optimized.

The same procedure adopted for benzene was first applied to the dimers of five-membered rings pyrrole, furan and thiophene. For each molecule, a set of four representative arrangements was either retrieved from the literature (for pyrrole<sup>43</sup> and thiophene<sup>66</sup>) or constructed exploiting the Hunter-Sanders rules (for furan, see Supporting Information). Preliminary energy curves were thereafter obtained, for each dimer arrangement, by displacing one monomer along a selected coordinate and computing the interaction energy of at MP2<sup>mod</sup>/6-31G\*(0.25) level, *i.e.* making use of a  $d$  exponent of 0.25 for all heavy atoms. From the resulting energy profiles twelve geometries (three for each curve, see Figure 3 and Figures A and B in the Supporting Information) were selected. Their CCSD(T)/CBS interaction energies were computed and stored in a reference database

to be used in the exponent tuning of both 6-31G\* and 6-31G\*\* basis sets, using the the EXOPT code.

As an example, the results of the tuning procedure carried out for the thiophene dimer are displayed in Figure 3. Both CCSD(T)/CBS interaction energies reported in literature

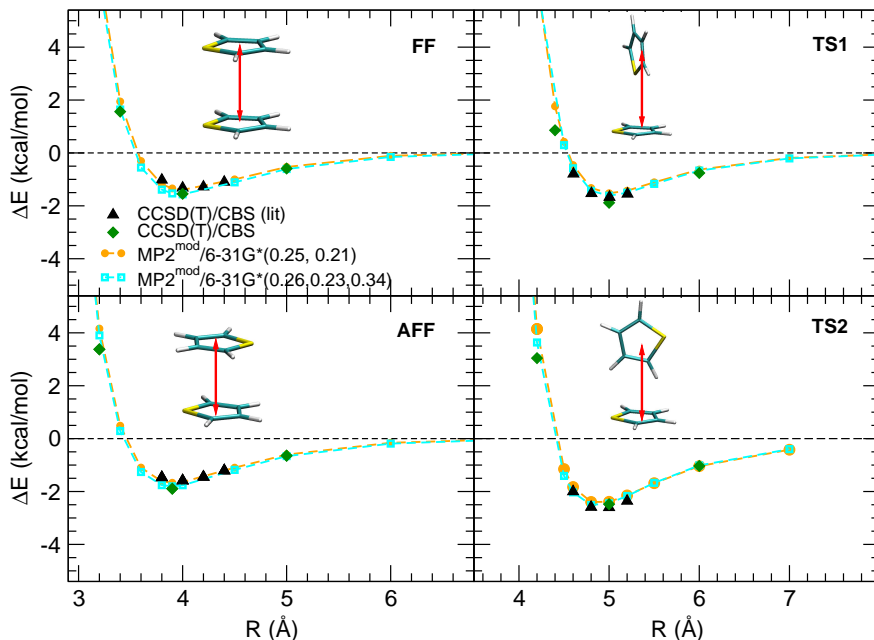


Figure 3: Results of EXOPT exponent optimization on the representative thiophene dimer arrangements chosen for the fit. CCSD(T)/CBS reference data, either found in literature<sup>66</sup> (black triangles) or computed in this work (green diamonds) are compared with the MP2<sup>mod</sup> ones, obtained with the 6-31G\*(0.25, 0.21) and 6-31G\*\*(0.26, 0.23, 0.34) modified basis sets (orange and cyan curves, respectively).

by Tsuzuki and coworkers<sup>66</sup> and those computed in this work are displayed, together with the pertinent arrangements shown in the insets. Notwithstanding the different CBS extrapolation schemes (see Refs. [66] and [21] for further details), it appears, as found for the benzene dimer, that the agreement between the two sets of CCSD(T) data is very good. The twelve reference interaction energies computed in this work were those used for the EXOPT tuning. The best exponents are  $\alpha_S=0.21$  for the 6-31G\* basis set and

$\alpha_C=0.26$ ,  $\alpha_S=0.23$ ,  $\alpha_H=0.34$ , for the 6-31G\*\* basis sets. From a visual inspection of the resulting MP2<sup>mod</sup> curves, a very good agreement with the reference data is apparent at all geometries. Similar results were obtained for pyrrole and furan, whose details are given in the Supporting Information (Figures A and B).

The results obtained on the benchmarked five-membered rings are summarized in Table 2, where the MP2<sup>mod</sup> RMSE with respect to reference CCSD(T)/CBS data is reported for both the basis sets. It is evident that the current MP2<sup>mod</sup> calculations

Dimer	Basis set	Exponents			Number of geometries	RMSE	LUE
		$\alpha_C$	$\alpha_X$	$\alpha_H$			
benzene	6-31G*	0.25	-	-	7	0.38	0.87
	6-31G**	0.32	-	0.20		0.14	0.30
pyrrole	6-31G*	0.25	0.37	-	12	0.76	2.13
	6-31G**	0.26	0.37	0.60		0.42	0.85
furan	6-31G*	0.25	0.32	-	13	0.34	0.69
	6-31G**	0.21	0.36	0.27		0.17	0.36
thiophene	6-31G*	0.25	0.21	-	14	0.47	1.08
	6-31G**	0.26	0.23	0.34		0.31	0.77

Table 2: Summary of MP2<sup>mod</sup> basis set tuning with respect to CCSD(T)/CBS reference values for the investigated five-membered aromatic heterocycles. The tuned exponents are explicitly reported for both modified basis sets (6-31G\* and 6-31G\*\*), RMSE and LUE are reported in kcal/mol. In the first row, benzene is also reported for comparison.

take advantage of the largest basis sets, which allow to halve the standard deviation for all systems. The comparable accuracy obtained for all dimers indicates that the improvement is almost system independent and the larger flexibility of the 6-31G\*\* basis set is capable to correctly describe the polarization effects arising from the presence of the second molecule. Finally, the results achieved with both basis sets are in line with those reported for the benzene molecule, except for the RMSE of the pyrrole dimer with the smaller basis set.

### ***Six-membered rings***

The exponent tuning was extended to the four six-membered aromatic heterocycles displayed in Figure 1, namely pyridine, pyrazine, pyrimidine and pyridazine. It might be worth noticing that in the previous set the difference among the investigated compounds consisted in the heteroatom embedded in the ring (N,S or O), whereas this set includes

molecules differing only by the number of Nitrogen atoms and by their relative position within the backbone. Considering the more similar chemical features of the selected six-membered rings, a more subtle and complex situation should be expected. Nonetheless, the tuning procedure applied for the benzene and the five-membered ring molecules was straightforwardly extended to the six-member heterocycles. By considering that in this new sample very similar  $MP2^{mod}$  performances were registered, to not overwhelm the reader with unnecessary information, only the outcomes observed for pyridazine will be discussed in the following, whereas a detailed description of the whole results is again given in the Supporting Information (Figures C, D and E).

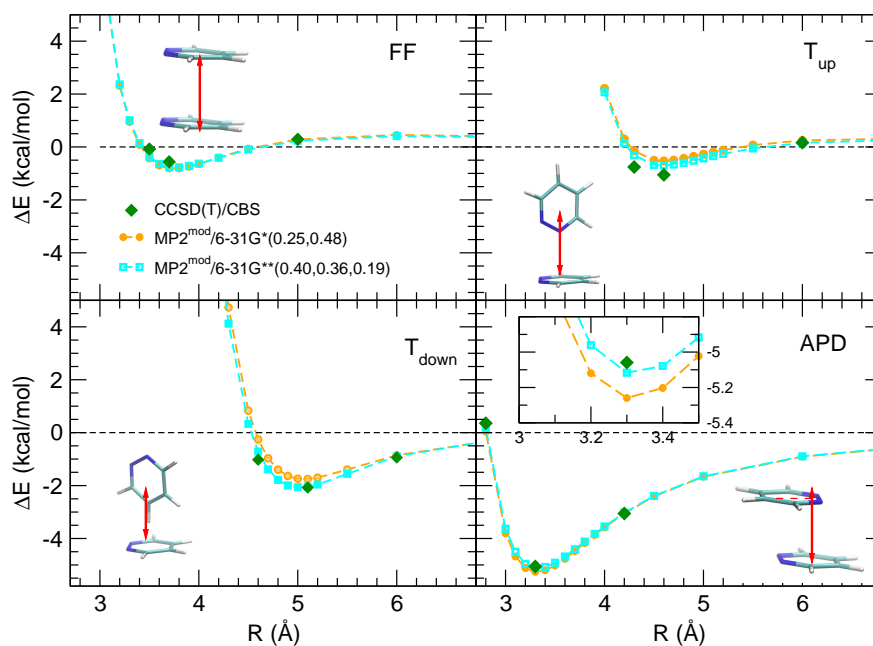


Figure 4: Results of EXOPT optimization on the representative pyridazine dimer arrangements chosen for the tuning. CCSD(T)/CBS reference data computed in this work (green diamonds) are compared with the  $MP2^{mod}$  ones, obtained with the two different basis sets. In the right bottom panel, the minimum APD region is evidenced in the inset.

The reference sample includes five representative arrangements of the pyridazine dimer,

namely FF, antiparallel FF (AFF), antiparallel displaced (APD), and two TS configurations, with either the Hydrogen ( $T_{down}$ ) or the N-N bond ( $T_{up}$ ) of one monomer is pointing toward the ring center of the other one (see the insets in Figure 4). The CCSD(T)/CBS energies were again computed in the three significant points (repulsive branch, minimum and long-range) along each curve except the AFF one, since it seemed redundant with the APD one. Based on the resulting CCSD(T) data, the two basis sets were tuned, obtaining  $\alpha_N = 0.48$  for the 6-31G\* basis set and  $\alpha_C = 0.40$ ,  $\alpha_N = 0.36$  and  $\alpha_H = 0.19$  for the 6-31G\*\* basis set. The considered interaction curves are reported in Figure 4 together with the reference data. From a first visual inspection it is clear as most of the features encountered in the previous cases still hold for pyridazine. The addition of Hydrogen polarization functions allows for a more accurate representation of the CH- $\pi$  interactions taking place in the T-down conformer, and the increased flexibility of the larger basis set reflects in a better performance, evident, for instance, in the APD stacked arrangements.

The performances of all MP2<sup>mod</sup> calculations carried out for the nitrogen containing heterocycles are summarized in Table 3. The quality of the MP2<sup>mod</sup> predictions is again

Dimer	Basis set	Exponents			Number of geometries	RMSE	LUE
		$\alpha_C$	$\alpha_X$	$\alpha_H$			
benzene	6-31G*	0.25	-	-	7	0.38	0.87
	6-31G**	0.32	-	0.20		0.14	0.30
pyridine	6-31G*	0.25	0.35	-	12	0.39	0.84
	6-31G**	0.36	0.18	0.20		0.29	0.67
pyridazine	6-31G*	0.25	0.48	-	12	0.40	0.77
	6-31G**	0.40	0.36	0.19		0.21	0.44
pyrimidine	6-31G*	0.25	0.26	-	12	0.45	0.81
	6-31G**	0.46	0.33	0.20		0.24	0.55
pyrazine	6-31G*	0.25	0.42	-	12	0.41	0.61
	6-31G**	0.57	0.31	0.20		0.28	0.53

Table 3: Summary of MP2<sup>mod</sup> basis set tuning with respect to CCSD(T)/CBS reference values for the investigated six-membered aromatic heterocycles. The tuned exponents are explicitly shown in the third, fourth and fifth column, whereas in the last one RMSE is reported in kcal/mol. In the first row, benzene is also reported for comparison.

very good, as the RMSE with respect to CCSD(T)/CBS values is always less than 0.3 kcal/mol, when MP2<sup>mod</sup> is coupled with the larger 6-31G\*\* modified basis sets. It might

be also interesting to note that the LUE registered for both basis sets corresponds, for all investigated samples, to short range repulsive arrangements, where the interaction energy curve is very steep.

Besides its accuracy, another strength of the  $\text{MP2}^{mod}$  procedure stands in its computational feasibility. As shown in the Supporting Information (Figure H), the small dimensions of the modified basis sets allow for a significant decrease of the computational cost not only with respect to  $\text{CCSD(T)}/\text{CBS}$  but also if compared to standard DFT-D approaches. Yet, it should be noticed that the latter DFT calculations were performed strictly following the indication of the original papers (see Supporting Information for details), *i.e.* with rather large basis sets. Despite a more fair comparison could be set up using smaller and specifically tuned basis sets for each investigated DFT-D functional, such computational burden is beyond the aim of the present work.

## 3.2 Interaction energy landscapes

The computational convenience of the proposed  $\text{MP2}^{mod}$  method can be exploited to investigate relevant portions of the IPES, with the aim of rationalizing the different behavior of the considered molecules. To set up a significant comparison among the investigated species, the investigated IPES sections were created at the same inter-ring distance, which, based on the previous one-dimensional scans, appears to be one of the most relevant coordinate for the interaction energy. As illustrated in Figure 5, two separate 2D scans were prepared for each molecular dimer, fixing the distance between the centers of mass at 3.5 Å and 4.6 Å, which correspond (see Figures 2, 3, 4 and A-E in the Supporting Information) to stacked and TS minima, respectively. For the stacked case, as shown in the left panels of Figure 5, the two coordinates subjected to the scan were the displacement along one ring plane ( $d$ ) and the angle around the axis perpendicular to the cycle ( $\beta$ ). For all molecules,  $d$  was varied between 0 and 1.8 Å in steps of 0.3, whereas  $\beta$  was scanned in the  $[0^\circ - 180^\circ]$  range in steps of  $30^\circ$ . Conversely, in the 2D scan carried out for the starting TS conformer, the  $\alpha$  and  $\beta$  angles were varied, again in the  $[0^\circ - 180^\circ]$  in steps of  $30^\circ$  (see right panels of Figure 5).



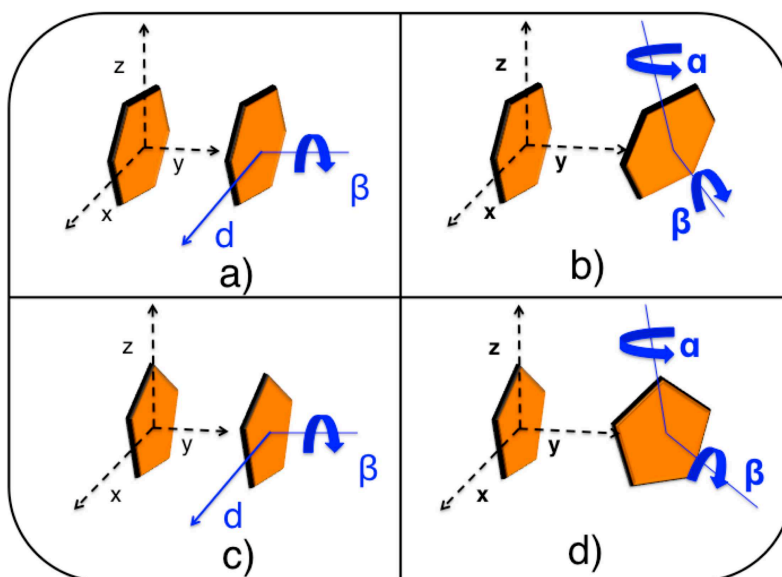


Figure 5: Procedure for scanning IPES sections of five-membered (bottom) and six-membered (top) rings. Left: two-dimensional scans of stacked-like arrangements. The inter-ring distance was kept fixed at  $3.5 \text{ \AA}$  for all benchmarked molecules, and the displacement  $d$  and the angle  $\beta$  were scanned in steps of  $0.3 \text{ \AA}$  and  $30^\circ$ , respectively. Right: two-dimensional scans of stacked ( $\alpha=0^\circ$ ) and TS ( $\alpha=90^\circ$ ) arrangements. In this case, the inter-ring distance was kept fixed at  $4.6 \text{ \AA}$  for all benchmarked molecules, and the  $\alpha$  and  $\beta$  angles were both scanned in steps of  $30^\circ$ .

All energy scans were performed at  $\text{MP2}^{mod}$  level, with the larger 6-31G\*\* modified basis sets. The computational convenience of the current  $\text{MP2}^{mod}$  method is confirmed by the fact that each scan required less than 30 minutes on a 24 processors (Xeon 2.60 GHz) computer. For comparison, on the same machine, a single CCSD(T)/CBS point required around 500 minutes. The results are shown in Figures 6 and 7, before discussing them into some detail, a final validation was further performed on the  $\text{MP2}^{mod}$  accuracy.

### 3.3 Final validation

Beside its computational convenience, for a balanced description of the IPES, it is of the foremost importance that the  $\text{MP2}^{mod}$  method does not alter the delicate interplay existing between the different components that concur to aromatic interactions, as for instance the stacked  $\pi$ - $\pi$  interactions or the CH- $\pi$  ones, occurring in the TS like arrangements. In this section we report some tests on the quality of  $\text{MP2}^{mod}$  energy predictions, for arrangements different from the ones employed in the basis set tuning. To this aim,

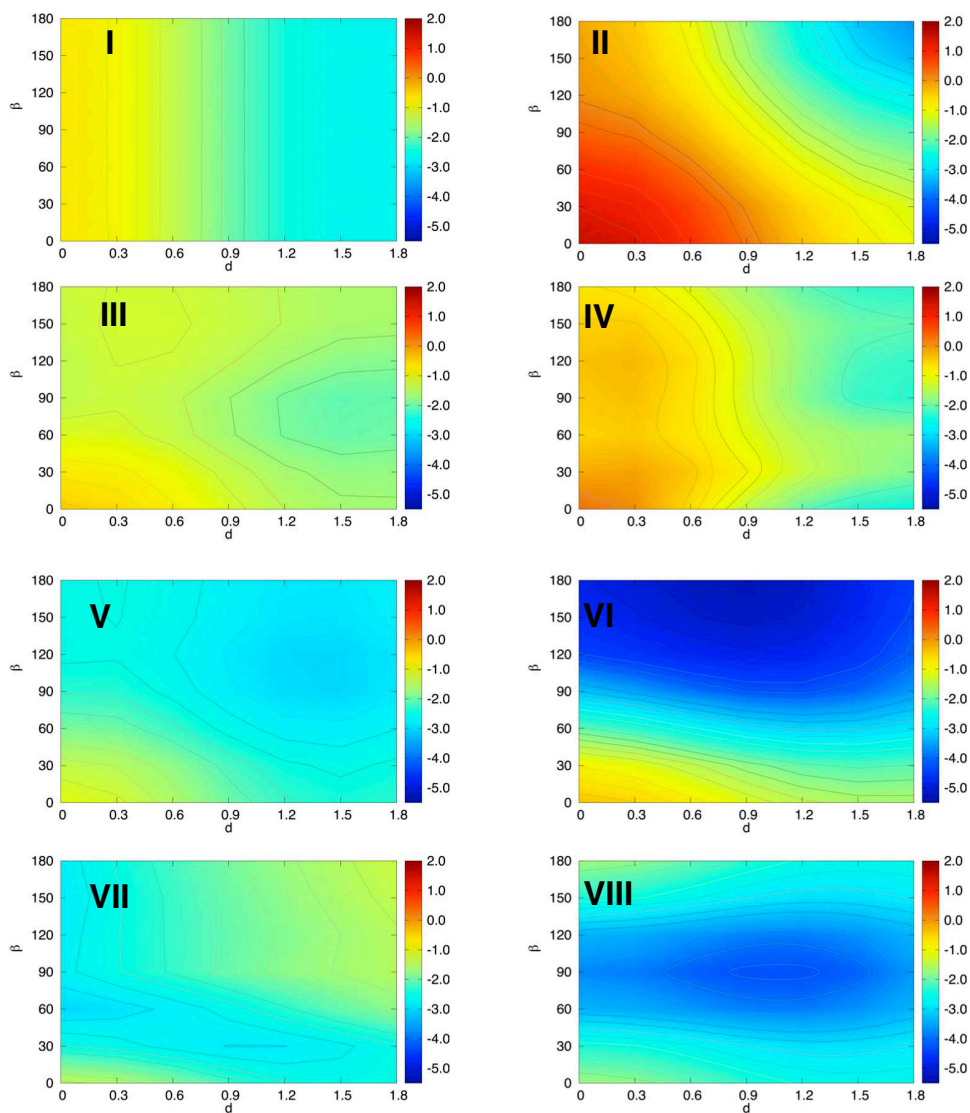


Figure 6:  $MP2^{mod}$  two-dimensional interaction energy scans, computed on dimers of the compounds shown in Figure 1 arranged in stacked-like geometries.  $d$  and  $\beta$  displacements refer to the a) and c) panels of Figure 5. I) benzene; II) pyrrole; III) furan; IV) thiophene; V) pyridine; VI) pyridazine; VII) pyrimidine; VIII) pyrazine.

the most favorable geometric arrangements for each scan (*i.e.* the IPES local minima within the step size employed in the 2D scans) were sorted from the scan and used for an additional CCSD(T)/CBS calculation. Such geometries are displayed in Figures F and G in the Supporting Information and the results summarized in Table 4.

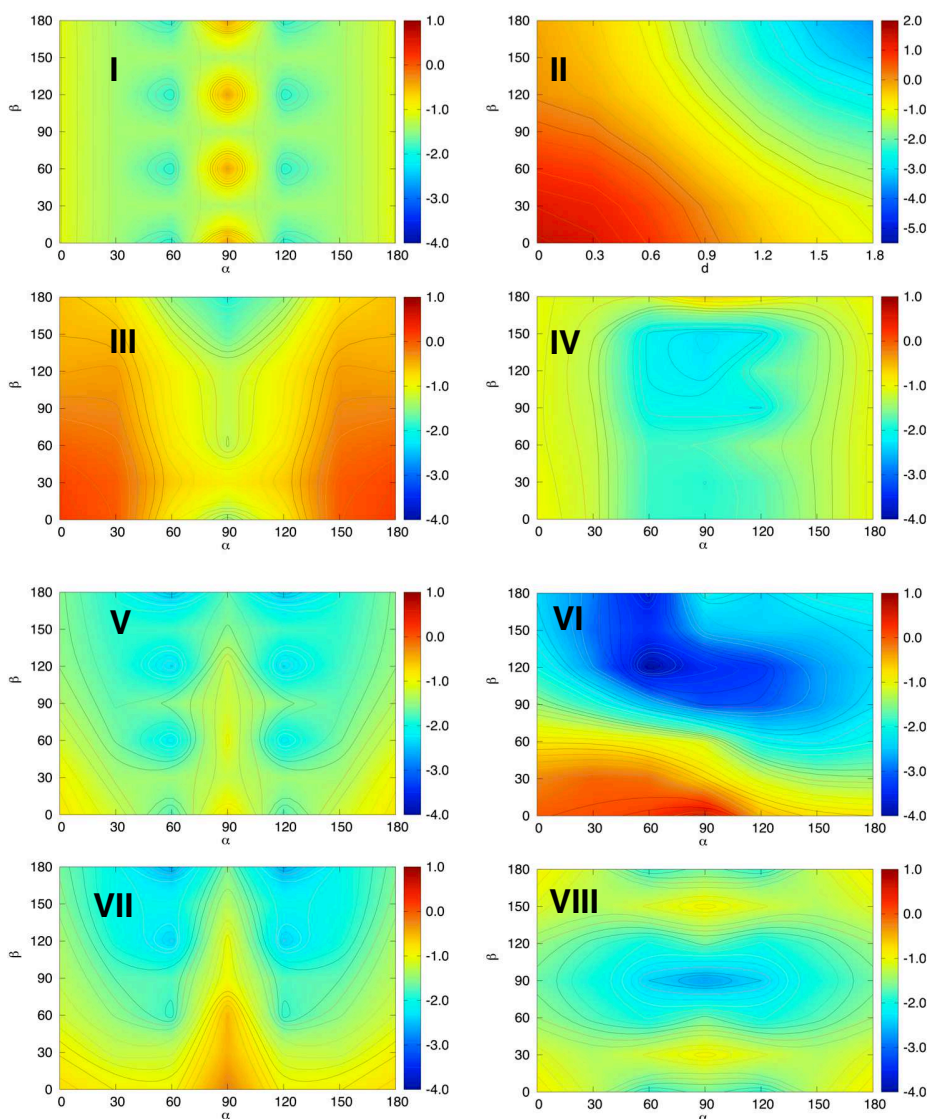


Figure 7:  $\text{MP2}^{\text{mod}}$  two-dimensional interaction energy scans, computed on dimers of the compounds shown in Figure 1 arranged in TS-like geometries.  $\alpha$  and  $\beta$  displacements refer to the b) and d) panels of Figure 5. I) benzene; II) pyrrole; III) furan; IV) thiophene; V) pyridine; VI) pyridazine; VII) pyrimidine; VIII) pyrazine.

From the comparison of  $\text{MP2}^{\text{mod}}$  vs.  $\text{CCSD(T)}/\text{CBS}$  results, it is evident that  $\text{MP2}^{\text{mod}}$  is capable of maintaining a similar accuracy even outside the training set. Moreover, the ratio between the interaction energy in a stacked and a TS conformation found at  $\text{CCSD(T)}/\text{CBS}$  level is kept almost unaltered by  $\text{MP2}^{\text{mod}}$ . In fact, this latter feature al-

Molecule	$R_{ring-ring}$	$d/\alpha$	$\alpha/\beta$	$\Delta E$		
				HF	MP2 <sup>mod</sup> /6-31G**	CCSD(T)/CBS
Benzene	3.5	1.8	30	3.97	-2.64	-2.75
	4.6	60	60	2.98	-1.89	-1.99
Pyrrole	3.5	1.8	180	0.38	-3.72	-4.15
	4.6	90	90	1.41	-1.93	-2.31
Furan	3.5	1.5	90	2.18	-2.06	-2.07
	4.6	90	90	0.62	-2.27	-2.36
Thiophene	3.5	1.8	90	4.31	-2.35	-2.62
	4.6	90	150	2.66	-2.35	-2.52
Pyridine	3.5	1.5	120	3.05	-3.11	-3.07
	4.6	60	180	1.65	-2.48	-2.60
Pyridazine	3.5	0.9	180	1.19	-5.30	-5.24
	4.6	60	120	-1.03	-3.98	-4.16
Pyrimidine	3.5	0	60	2.83	-3.10	-3.14
	4.6	60	180	0.50	-2.63	-2.91
Pyrazine	3.5	1.2	90	1.74	-4.18	-4.10
	4.6	90	90	1.04	-2.75	-3.16

Table 4: Interaction energies (kcal/mol) computed for the minimum energy geometries found in the 2D energy scans (shown in Figures F and G in the Supporting Information) at MP2<sup>mod</sup>/6-31G\*\* and CCSD(T)/CBS level. For the former calculation the uncorrelated Hartree-Fock term (HF) is also reported. Inter ring distance ( $R_{ring-ring}$ ) and displacement ( $d$ ) are displayed in Å, whereas  $\alpha$  and  $\beta$  angles are in degrees.

allows us to confidently employ MP2<sup>mod</sup> to exhaustively sample larger regions of the IPESs, as for instance required for force-field parameterization.<sup>37,53,55,56,80,82-91</sup> This will be the subject of a forthcoming work, currently in progress in our group.

## 4 Discussion

The validation of the reliability of the MP2<sup>mod</sup> predictions, achieved in the previous section, allows us to discuss the results reported in Table 4 and Figures 6 and 7 with increased confidence. As far as the five membered ring are concerned, it appears from Table 4 that the stacked interaction energy involving pyrrole (-3.72 kcal/mol) is significantly larger than those computed for furan and thiophene (-2.06 and -2.35 kcal/mol). This result agrees with the recent findings of Wheeler’s group,<sup>17</sup> who computed at DFT-D level the interaction of methyl-adenine stacked with several hetero-cycles of pharmaceutical inter-

est, finding the same trends in going from pyrrole to furan. Furthermore, the significant contribution of the NH- $\pi$  interactions to the total energy, already discussed by some of us previously<sup>43</sup> and also confirmed by Ref. [17], clearly appears by looking at the minimum pyrrole stacked geometry (see Figure F in the Supporting Information). By looking at both the minima considered in Table 4 and at the IPES sections displayed in Figures 6 and 7, it can be observed that for furan and thiophene the TS arrangements show similar or even lower energies with respect to the stacked ones. This is in agreement with the accurate CCSD(T)/CBS data published by Tsuzuki<sup>66</sup> (and here displayed in Figure 3), who first hypothesized the TS stability as a probable cause of the herringbone structure of thiophene crystals.

Turning to the six-membered, Nitrogen containing rings, both Wheeler’s group<sup>17</sup> and Huber and coworkers<sup>38</sup> respectively considered the interaction of azines with methyl-adenine and benzene. It was found not only that the presence of the heteroatom increased the interaction energy with respect to the benzene dimer, but also that the number of Nitrogen atoms and their relative position affect the interaction strength, resulting in the series pyridazine > pyridine > pyrimidine  $\simeq$  pyrazine. By looking at the data reported in Table 4, all the above mentioned observations are confirmed, apparently with the exception of pyrazine, whose stacked interaction energy (-4.18 kcal/mol) is lower than in pyridazine (-5.30 kcal/mol), but more favorable than both pyridine (-3.11 kcal/mol) and pyrimidine (-3.10 kcal/mol). However, Figure 6 reveals that when the  $\beta$  extrema (*i.e.* FF with  $\beta=0^\circ$  or AFF, with  $\beta=180^\circ$ ) are considered, the ordering found in References [38] and [17] is recovered, as the pyrazine dimer is significantly stabilized by the  $\beta$ -rotation to a ”cross-like” conformer ( $\beta=0^\circ$ ), which could not be formed in the pyrazine-methyl-adenine or pyrazine-benzene hetero-dimers. This latter finding is also in agreement with the results of Mishra *et al.*,<sup>71</sup> who found a similar ”cross-like” ( $\beta \simeq 0^\circ$ ) minimum energy arrangement, with a CCSD(T)/CBS computed energy of -4.14 kcal/mol. A further confirmation can be found in the more recent work by Tekin and coworkers,<sup>75</sup> who report a very similar cross displaced stacked conformation, found after a dimer geometry optimization carried out at SCS-MP2 level, that yielded a final binding energy of 4.3 kcal/mol, again in agreement with the MP2<sup>mod</sup> value.

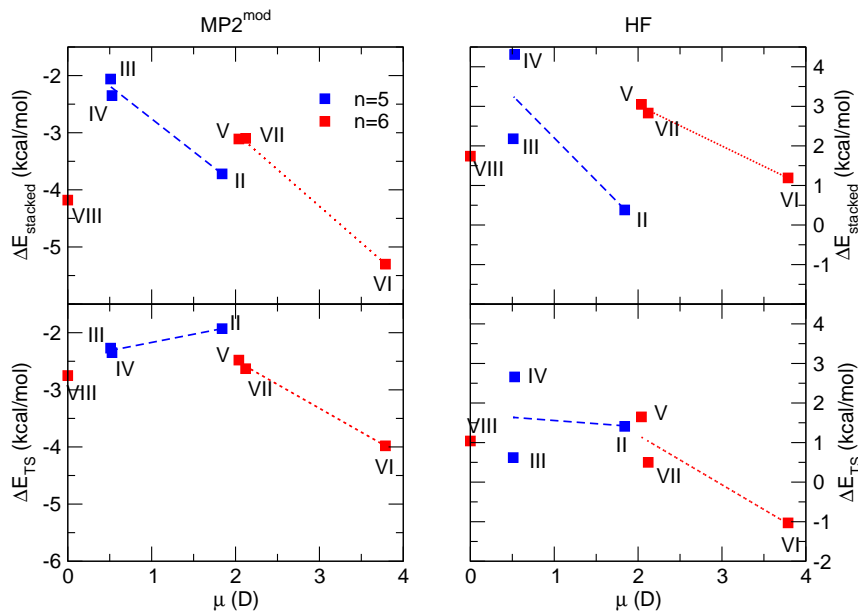


Figure 8: The intermolecular energies, computed at  $MP2^{mod}$  (left panels) and Hartree-Fock (HF, right panels) level for the minima reported in Table 4, are plotted vs. the molecular dipole moments of the involved species. Top panels: stacked conformations; bottom panels: T-shaped conformations. Roman numbers refer to the species shown in Figure 1: II: pyrrole; III: furan; IV: thiophene; V: pyridine; VI: pyridazine; VII: pyrimidine; VIII: pyrazine.

In the recent computational study performed on dimers of benzene and small heterocycles, Huber and coworkers<sup>38</sup> found a correlation between the computed stacking interactions and the molecular dipole moments of the target heterocycle. More recently, Wheeler and coworkers<sup>17</sup> confirmed that the molecular dipoles could be used as predictors of the strength of the stacking interaction energy involving small heterocycles, but showed that the same does not hold for larger cycles. To get a deeper insight into this feature, the molecular dipoles of the heterocycles investigated herein were plotted in Figure 8 against  $MP2^{mod}$  interaction energies computed for the local minima reported in Table 4. From a first inspection, it appears that all targets with a non-null dipole moment present some correlation between the dipole strength and the total stacked interaction energy (top left panel of Figure 8). Interestingly, the same holds even for TS conformations, as evident

in the top left panel. From a closer look, however, some discrepancies with the previous findings emerge. First, a good correlation is obtained only if five- and six-membered rings are considered separately. Next, the slope of the correlation regression for the interaction energy of five-membered rings in TS arrangements has an opposite sign with respect to the one obtained for six-membered cycles. Finally, the HF component of the total interaction energy shows a much weaker (or null) correlation, even if it should account for most of the first order electrostatic contribution. It should be though noted that the whole comparison with the correlation reported in Reference [38] and [17] could be undermined by an important difference among the sample compounds considered in these works and in the present case. In fact, in References [38] and [17], the investigated heterocycles are interacting respectively with a molecule (benzene) with null dipole moment or with a significantly polar molecule as methyl adenine. Conversely, here homo-dimers were considered, either with null (benzene, pyrazine) or not null dipole moments. Furthermore, the whole above analysis is limited to few selected geometries and only concerns with a rather simple indicator as the molecular dipole moment of each monomer.

The MP2<sup>mod</sup> computational convenience allows for the calculation of a much larger number of dimer arrangements, as done for instance in the extended sampling of the IPES landscapes displayed in Figures 6 and 7. In the attempt to find a rationale of such landscapes and the relationships between the electrostatic contribution and the total interaction energy, a more detailed analysis is displayed, for stacked dimers, in Figure 9. Each set of points of the same color in this Figure, connected by the least square line, refers to the arrangements at a given displacement ( $d$  in Figure 6) with the second molecule undergone to a rotation ( $\alpha$ ) around the axis perpendicular to the molecular plane and passing in the center of the polygon, as shown in Figure 5. The abscissa data are an estimate of the electrostatic interaction energy  $E(\text{ESP})$  obtained by the electrostatic potential (ESP) charges, as computed for the monomer at B3LYP/cc-pVDz level of theory. Therefore Figure 9 can furnish some insights on the contribution of electrostatic energy to the whole interaction energy.

For the five-membered atom ring dimers pyrrole shows by far the most significant  $\Delta E/E(\text{ESP})$  correlation between the total and electrostatic energies, clearly indicating

that the minimum of each subset with the dipoles directed in opposite directions is essentially determined by the electrostatic energy. As the slope of the curves is slightly larger than unity we can deduce that there are other cooperative energy components that reinforce the electrostatic profile. Furan dimer presents a much lower (and always positive) electrostatic energy and shows weakly correlated curves with scattered data, although some features similar to those of pyrrole seem also to be present. Thiophene does not show any significant correlation. The different behavior of these systems may find a rationale in the dipole moments of the monomers (1.98, 0.51 and 0.44 D, respectively for Pyrrole, Furan and Thiophene) allowing to speculate that different energy contributions concur to stabilize the antiparallel arrangements at all plane-to-plane distances so that in the case of large electrostatic energies (corresponding to large dipole moments) the  $\Delta E/E(\text{ESP})$  correlation can be evidenced.

For the six-membered atom rings dimers the correlation between the total energy and its electrostatic (ESP derived) contribution is rather evident with the exception of pyrimidine. In this case it appears that the slope of the curves is close to the unity, indicating that the electrostatic energy is by far the main responsible of the energy profile along the rotation around the axis orthogonal to the molecular plane. A similar anomalous behavior

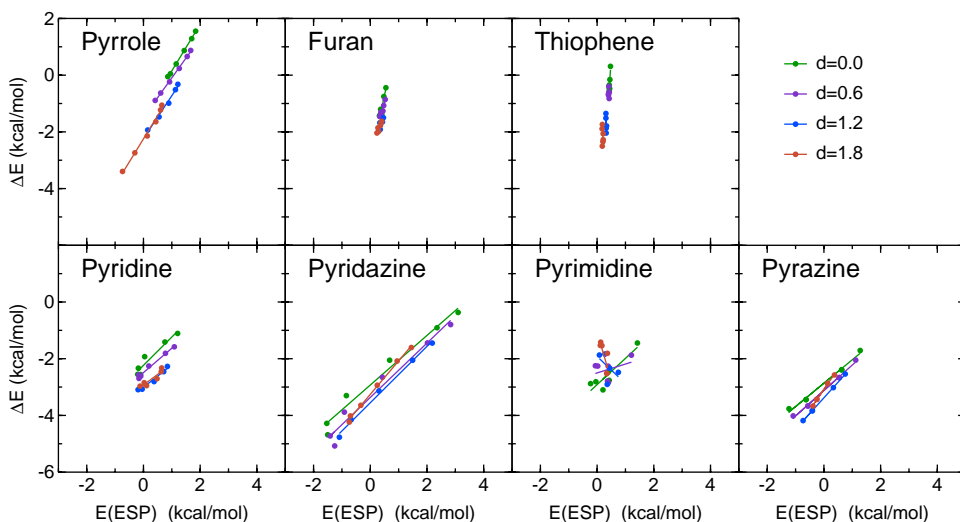


Figure 9:  $\text{MP2}^{\text{mod}}$  interaction energies vs ESP energy, computed along different  $d$  displacements in the conformational landscape of Figure 6.



of pyrimidine was found by Huber and coworkers,<sup>38</sup> when benzene-pyrimidine correlations were considered. This was ascribed to a partial compensation of the dipole moments induced by the two nitrogen atoms, which could enhance the effects of higher electrostatic moments. It might be worth noticing that, at variance with the five-membered atom ring dimers, even the pyrazine dimer (null dipole moment) shows a significant electrostatic energy which strongly correlates with the interaction energy, whose minima are quite pronounced and resemble those of pyridazine. This resemblance with pyridazine, which has the highest dipole moment among the considered six-membered atom rings dimers, may appear rather unexpected. Notice that, besides the interaction energy, the resemblance include also the electrostatic energy, which is rather similar for the two systems. A possible qualitative explanation could arise from the electrostatic interaction among the several C-N bond local dipoles, or, equivalently, from the higher order electrostatic moments, whose interaction energy is expected to show a marked dependence on the relative orientation.

## 5 Conclusions

The first aim of this work concerned with the automation and validation of the  $\text{MP2}^{mod}$  procedure, which essentially consists in tuning some exponents of small dimensioned basis sets to be coupled with the standard MP2 technique. In this regard the  $\text{MP2}^{mod}$  approach has been shown to be capable to deliver very accurate predictions of the interaction energy at a very low computational cost, even when hetero-aromatic interaction are involved. The  $\text{MP2}^{mod}$  performances maintain a similar high level of accuracy for all tested samples, even outside the training set employed in the basis set tuning. At least for the tested sample,  $\text{MP2}^{mod}$  outperforms most of the benchmarked DFT-D functionals, both in accuracy and in computational convenience, at least when large basis sets are employed. The high accuracy and the low computational cost of the  $\text{MP2}^{mod}$  protocol represents an important step toward the possibility of sampling large portions of IPESs, paving the way for building reference databases. These could be in turn exploited for developing accurate QM derived FFs, allowing for *in silico* investigations on larger length and time scale or on more complex systems. Furthermore, the achieved automation of the  $\text{MP2}^{mod}$  proce-

ture can be easily extended to other similar or even larger hetero-cycles, as for instance those commonly employed in drug design. In this framework the similarity of some of the retrieved exponents (as for instance those for the N atoms in the investigated diazines), seems to suggest that a certain degree of transferability might hold for the exponents, so that the expensive CCSD(T) calculations could be avoided when the target molecule is composed of building blocks already tackled by MP2<sup>mod</sup>, as for instance biphenyl or poly-thiophene.

Next, MP2<sup>mod</sup> computational feasibility and reliability have been exploited to fulfill the second goal of this work, that is to make a further step in the comprehension of the intricacies of the noncovalent interactions among aromatic hetero-cycles. To this end selected portions of the heterocycles IPESs were computed, allowing for an extended analysis of the resulting interaction patterns. On the one hand our findings confirm the overall picture available in literature regarding heteroaromatic stacked interaction: the complex orientational dependence, the peculiar effects of substitution and the correlation with simple molecular descriptors. On the other hand, the large amount of geometries considered, has allowed us to not only consolidate the statistical validity of previous findings, but also to extend the discussion to other possible competing arrangements, as for instance the T-shaped ones, which have been seldom considered in the literature but can play an important role in condensed phases.

### Supporting Information

Additional data and several details about the reported calculations not included in this paper.

## References

- [1] Fagnani, D.; Sotuyo, A.; Castellano, R. In  *$\pi\pi$  Interactions. Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Elsevier, 2017.
- [2] 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.

- [3] Mak, C. H. Unraveling Base Stacking Driving Forces in DNA. *J. Phys. Chem. B* **2016**, *120*, 6010–6020.
- [4] Brédas, J.-L.; Norton, J. E.; Cornil, J.; Coropceanu, V. Molecular Understanding of Organic Solar Cells: The Challenges *Acc. Chem. Res.* **2009**, *42*, 1691–1699.
- [5] Shewmon, N. T.; Watkins, D. L.; Galindo, J. F.; Zerdan, R. B.; Chen, J.; Keum, J.; Roitberg, A. E.; Xue, J.; Castellano, R. K. Enhancement in Organic Photovoltaic Efficiency through the Synergistic Interplay of Molecular Donor Hydrogen Bonding and  $\pi$ -Stacking *Advanced Functional Materials* **2015**, *25*, 5166–5177.
- [6] Cragg, P. J. *Supramolecular Chemistry - From Biological Inspiration to Biomedical Applications*; Springer Netherlands, 2010.
- [7] Meanwell, N.; Scriven, E. F.; Ramsden, C. A., Eds.; *Advances in Heterocyclic Chemistry*, Vol. 123; Academic Press, 2017; pp 245 – 361.
- [8] Hobza, P.; Zahradník, R.; Müller-Dethlefs, K. The World of Non-Covalent Interactions: 2006 *Collect. Czech. Chem. Comm.* **2006**, *71*, 443–531.
- [9] Grimme, S. Do Special Noncovalent  $\pi\pi$  Stacking Interactions Really Exist? *Angew. Chem. Int. Ed.* **2008**, *47*, 3430–3434.
- [10] Hohenstein, E. G.; Duan, J.; Sherrill, C. D. Origin of the Surprising Enhancement of Electrostatic Energies by Electron-Donating Substituents in Substituted Sandwich Benzene Dimers. *J. Am. Chem. Soc.* **2011**, *133*, 13244–13247.
- [11] Watt, M.; Hardebeck, L. K. E.; Kirkpatrick, C. C.; Lewis, M. Face-to-Face Arene - Arene Binding Energies : Dominated by Dispersion but Predicted by Electrostatic and Dispersion / Polarizability Substituent Constants *J. Am. Chem. Soc.* **2011**, 3854–3862.
- [12] Martinez, C. R.; Iverson, B. L.; Shimizu, K. D.; Swager, T.; Kouk, K. N.; Persson, M.; Yamaguchi, K.; Kim, K.; Iii, W. G.; Stoddart, J.; III, W. A. G.; Stoddart, J. Rethinking the Term  $\pi$ -Stacking *Chemical Science* **2012**, *3*, 2191–2201.

- [13] Raju, R. K.; Bloom, J. W. G.; Wheeler, S. E. Broad Transferability of Substituent Effects in  $\pi$ -Stacking Interactions Provides New Insights into Their Origin *J. Chem. Theory Comput.* **2013**, *9*, 3479–3490.
- [14] Sherrill, C. D. Energy Component Analysis of  $\pi$  Interactions *Acc. Chem. Res.* **2013**, *46*, 1020–1028.
- [15] 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*, 14060–14067.
- [16] Wheeler, S. E.; Bloom, J. W. G. Toward a More Complete Understanding of Noncovalent Interactions Involving Aromatic Rings *J. Phys. Chem. A* **2014**, *118*, 6133–6147.
- [17] 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. Mod.* **2016**, *56*, 906–914.
- [18] Hunter, C. A.; Sanders, J. K. M. The nature of  $\pi$ - $\pi$  interactions *J. Am. Chem. Soc.* **1990**, *112*, 5525–5534.
- [19] Cockroft, S. L.; Perkins, J.; Zonta, C.; Adams, H.; Spey, S. E.; Low, C. M. R.; Vinter, J. G.; Lawson, K. R.; Urch, C. J.; Hunter, C. A. Substituent Effects on Aromatic Stacking Interactions *Organic & Biomolecular Chemistry* **2007**, *5*, 1062–1080.
- [20] 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. Chem. Phys.* **1996**, *100*, 18790–18794.
- [21] Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the  $\pi/\pi$  Interaction: Model Chemistry Calculations of Benzene Dimer Interaction *J. Am. Chem. Soc.* **2002**, *124*, 104–112.

- [22] 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.
- [23] Riley, K. E.; Platts, J. A.; Řezáč, J.; Hobza, P.; Hill, J. G. Assessment of the Performance of MP2 and MP2 Variants for the Treatment of Noncovalent Interactions. *J. Phys. Chem. A* **2012**, *116*, 4159–69.
- [24] Řezáč, J.; Hobza, P. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the Gold Standard, CCSD(T) at the Complete Basis Set Limit? *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.
- [25] Řezáč, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071.
- [26] Grimme, S. Density Functional Theory with London Dispersion Corrections *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 211.
- [27] Ehrlich, S.; Moellmann, J.; Grimme, S. Dispersion-Corrected Density Functional Theory for Aromatic Interactions in Complex Systems. *Acc. Chem. Res.* **2013**, *46*, 916–926.
- [28] Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods *Chem. Rev.* **2016**, *116*, 5105–5154.
- [29] Burns, L.; Vázquez-Mayagoitia, A.; Sumpter, B. G.; Sherrill, C. D. Density-Functional Approaches to Noncovalent Interactions: a Comparison of Dispersion Corrections (DFT-D), Exchange-Hole Dipole Moment (XDM) Theory, and Specialized Functionals. *J. Chem. Phys.* **2011**, *134*, 084107.
- [30] Vazquez-Mayagoitia, A.; Sherrill, C. D.; Apra, E.; Sumpter, B. G.; Apra, E. An Assessment of Density Functional Methods for Potential Energy Curves of Nonbonded Interactions: The XYG3 and B97-D Approximations *J. Chem. Theory Comput.* **2010**, *6*, 727–734.

- [31] Riley, K. E.; Pitoňák, M.; Cerný, J.; Hobza, P. On the Structure and Geometry of Biomolecular Binding Motifs (Hydrogen-Bonding, Stacking,  $XH \cdots \pi$ ): WFT and DFT Calculations *J. Chem. Theory Comput.* **2010**, *6*, 66–80.
- [32] Paytakov, G.; Dinadayalane, T.; Leszczynski, J. Toward Selection of Efficient Density Functionals for van der Waals Molecular Complexes: Comparative Study of  $CH \cdots \pi$  and  $NH \cdots \pi$  Interactions. *J. Phys. Chem. A* **2015**, *119*, 1190–1200.
- [33] Forni, A.; Pieraccini, S.; Rendine, S.; Sironi, M. Halogen Bonds with Benzene: an Assessment of DFT Functionals. *J. Comp. Chem.* **2014**, *35*, 386–94.
- [34] Morgado, C.; Vincent, M. A.; Hillier, I. H.; Shan, X. Can the DFT-D Method Describe the Full Range of Noncovalent Interactions Found in Large Biomolecules? *Phys. Chem. Chem. Phys.* **2007**, *9*, 448–51.
- [35] Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. Assessment of the Performance of the M052X and M062X Exchange-Correlation Functionals for Noncovalent Interactions in Biomolecules *J. Chem. Theory Comput.* **2008**, *4*, 1996–2000.
- [36] Marianski, M.; Oliva, A.; Dannenberg, J. J. A Reinvestigation of the Dimer of para-Benzoquinone and Pyrimidine with MP2, CCSD(T), and DFT Using Functionals Including Those Designed to Describe Dispersion. *J. Phys. Chem. A* **2012**, *116*, 8100–5.
- [37] Prampolini, G.; Livotto, P. R.; Cacelli, I. Accuracy of Quantum Mechanically Derived Force-Fields Parameterized from Dispersion-Corrected DFT Data: The Benzene Dimer as a Prototype for Aromatic Interactions. *J. Chem. Theory Comput.* **2015**, *11*, 5182–5196.
- [38] 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. Mod.* **2014**, *54*, 1371–1379.
- [39] Prampolini, G.; Campetella, M.; De Mitri, N.; Livotto, P. R.; Cacelli, I. Systematic and Automated Development of Quantum Mechanically Derived Force Fields: The

- Challenging Case of Halogenated Hydrocarbons *J. Chem. Theory Comput.* **2016**, *12*, 5525–5540.
- [40] Sierański, T. The Intricacies of the Stacking Interaction in a Pyrrole-Pyrrole System *Struct. Chem.* **2016**, *27*, 1107–1120.
- [41] Goldey, M. B.; Belzunces, B.; Head-Gordon, M. Attenuated MP2 with a Long-Range Dispersion Correction for Treating Nonbonded Interactions *J. Chem. Theory Comput.* **2015**, *11*, 4159–4168.
- [42] Barone, V.; Cacelli, I.; Crescenzi, O.; D’Ischia, M.; Ferretti, A.; Prampolini, G.; Villani, G. Unraveling the Interplay of Different Contributions to the Stability of the Quinhydrone Dimer *RSC Adv.* **2014**, *4*, 876.
- [43] Prampolini, G.; Cacelli, I.; Ferretti, A. Intermolecular Interactions in Eumelanins: a Computational Bottom-up Approach. I. Small Building Blocks *RSC Adv.* **2015**, *5*, 38513–38526.
- [44] Kroon-Batenburg, L.; Van Duijneveldt, F. The Use of a Moment-Optimized DZP Basis Set for Describing the Interaction in the Water Dimer *J. Mol. Struct. THEOCHEM* **1985**, *22*, 185–199.
- [45] Šponer, J.; Leszczynski, J.; Hobza, P. Base Stacking in Cytosine Dimer. A Comparison of Correlated Ab Initio Calculations with Three Empirical Potential Models and Density Functional Theory Calculations *J. Comput. Chem.* **1996**, *17*, 841–850.
- [46] Riley, K.; Hobza, P. Assessment of the MP2 Method, along with Several Basis Sets, for the Computation of Interaction Energies of Biologically Relevant Hydrogen Bonded and Dispersion Bound Complexes *J. Phys. Chem. A* **2007**, *111*, 8257–8263.
- [47] Šponer, J.; Gabb, H. A.; Leszczynski, J.; Hobza, P. Base-base and Deoxyribose-Base Stacking Interactions in B-DNA and Z-DNA: a Quantum-Chemical Study. *Biophys. J.* **1997**, *73*, 76–87.
- [48] Hobza, P.; Šponer, J. Toward True DNA Base-Stacking Energies: MP2, CCSD(T), and Complete Basis Set Calculations *J. Am. Chem. Soc.* **2002**, *124*, 11802–11808.

- [49] Řeha, D.; Kabeláč, M.; Ryjáček, F.; Šponer, J.; Šponer, J.; Elstner, M.; Suhai, S.; Hobza, P. Intercalators. 1. Nature of Stacking Interactions between Intercalators (Ethidium, Daunomycin, Ellipticine, and 4,6-Diaminide-2-phenylindole) and DNA Base Pairs. Ab Initio Quantum Chemical, Density Functional Theory, and Empirical Potential Study *J. Am. Chem. Soc.* **2002**, *124*, 3366–3376.
- [50] Waller, M. P.; Robertazzi, A.; Platts, J. A.; Hibbs, D. E.; Williams, P. A. Hybrid Density Functional Theory for  $\pi$ -stacking Interactions: Application to Benzenes, Pyridines, and DNA Bases *J. Comp. Chem.* **2006**, *27*, 491–504.
- [51] Jurecka, P.; Hobza, P. True Stabilization Energies for the Optimal Planar Hydrogen-Bonded and Stacked Structures of Guanine···Cytosine, Adenine···Thymine, and Their 9- and 1-Methyl Derivatives: Complete Basis Set Calculations at the MP2 and CCSD(T) Levels and Comparison with Ex *J. Am. Chem. Soc.* **2003**, *125*, 15608–15613.
- [52] Rutledge, L. R.; Durst, H. F.; Wetmore, S. D. Evidence for Stabilization of DNA/RNAProtein Complexes Arising from Nucleobase-Amino Acid Stacking and T-Shaped Interactions *J. Chem. Theory Comput.* **2009**, *5*, 1400–1410.
- [53] Cacelli, I.; Cinacchi, G.; Prampolini, G.; Tani, A. Computer Simulation of Solid and Liquid Benzene with an Atomistic Interaction Potential Derived from Ab Initio Calculations. *J. Am. Chem. Soc.* **2004**, *126*, 14278–14286.
- [54] Amovilli, C.; Cacelli, I.; Cinacchi, G.; Gaetani, L.; Prampolini, G.; Tani, A. Structure and Dynamics of Mesogens Using Intermolecular Potentials Derived from Ab Initio Calculations *Theor. Chem. Accounts* **2007**, *117*, 885–901.
- [55] Cacelli, I.; Lami, C. F.; Prampolini, G. Force-field Modeling through Quantum Mechanical Calculations: Molecular Dynamics Simulations of a Nematogenic Molecule in its Condensed Phases. *J. Comp. Chem.* **2009**, *30*, 366–378.
- [56] Cacelli, I.; Cimoli, A.; De Gaetani, L.; Prampolini, G.; Tani, A. Chemical Detail Force Fields for Mesogenic Molecules *J. Chem. Theory Comput.* **2009**, *5*, 1865–1876.



- [57] 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*, 6038–6044.
- [58] Bludský, O.; Rubeš, M.; Soldán, P.; Nachtigall, P. Investigation of the Benzene-Dimer Potential Energy Surface: DFT/CCSD(T) Correction Scheme *J. Phys. Chem.* **2008**, *128*, 1–8.
- [59] Janowski, T.; Pulay, P. High Accuracy Benchmark Calculations on the Benzene Dimer Potential Energy Surface *Chem. Phys. Lett.* **2007**, *447*, 27–32.
- [60] Podeszwa, R.; Bukowski, R.; Szalewicz, K. Potential Energy Surface for the Benzene Dimer and Perturbational Analysis of  $\pi$ - $\pi$  Interactions *J. Phys. Chem. A* **2006**, *110*, 10345–10354.
- [61] Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. An Assessment of Theoretical Methods for Nonbonded Interactions: Comparison to Complete Basis Set Limit Coupled-Cluster Potential Energy Curves for the Benzene Dimer, the Methane Dimer, Benzene-Methane, and Benzene-H<sub>2</sub>S *J. Phys. Chem. A* **2009**, *113*, 10146–10159.
- [62] Park, H.; Lee, S. Ab Initio Investigations of the Pyrrole Dimer: a Direct Observation of the  $\pi$ -facial Hydrogen Bond *Chem. Phys. Lett.* **1999**, *301*, 487–492.
- [63] Lukeš, V.; Breza, M.; Biskupič, S. Interaction Energy Anisotropy of the Pyrrole Dimer: Ab Initio Theoretical Study *Theor. Chem. Accounts* **1999**, *101*, 319–324.
- [64] Kabelác, M.; Hobza, P.; Spirko, V. The Structure and Vibrational Dynamics of the Pyrrole Dimer. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3885–91.
- [65] Pei, K.; Li, H. Structure and Property of the Furan Dimer (C<sub>4</sub>H<sub>4</sub>O)<sub>2</sub>: a Theoretical Study *J. Mol. Struct.* **2004**, *693*, 141–144.
- [66] Tsuzuki, S.; Honda, K.; Azumi, R. Model Chemistry Calculations of Thiophene Dimer Interactions: Origin of  $\pi$ -Stacking *J. Am. Chem. Soc.* **2002**, *124*, 12200–12209.

- [67] Scherlis, D. A.; Marzari, N.  $\pi$ -Stacking in Thiophene Oligomers as the Driving Force for Electroactive Materials and Devices. *J. Am. Chem. Soc.* **2005**, *127*, 3207–3212.
- [68] Mackie, I. D.; McClure, S. A.; DiLabio, G. A. Binding in Thiophene and Benzothio-  
phene Dimers Investigated by Density Functional Theory with Dispersion-Correcting  
Potentials. *J. Phys. Chem. A* **2009**, *113*, 5476–5484.
- [69] Mishra, B. K.; Sathyamurthy, N.  $\pi\pi$  Interaction in Pyridine *J. Phys. Chem. A* **2005**,  
*109*, 6–8.
- [70] Hohenstein, E. G.; Sherrill, C. D. Effects of Heteroatoms on Aromatic  $\pi$ - $\pi$  Interac-  
tions: Benzene-Pyridine and Pyridine Dimer *J. Phys. Chem. A* **2009**, *113*, 878–886.
- [71] Mishra, B. K.; Arey, J. S.; Sathyamurthy, N. Stacking and Spreading Interaction in  
N-Heteroaromatic Systems *J. Phys. Chem. A* **2010**, *114*, 9606–9616.
- [72] Mottishaw, J. D.; Sun, H. Effects of Aromatic Trifluoromethylation, Fluorination,  
and Methylation on Intermolecular  $\pi$ - $\pi$  interactions. *J. Phys. Chem. A* **2013**, *117*,  
7970–9.
- [73] Ninković, D. B.; Andrić, J. M.; Zarić, S. D. Parallel Interactions at Large Horizontal  
Displacement in Pyridine-Pyridine and Benzene-Pyridine Dimers *ChemPhysChem*.  
**2013**, *14*, 237–243.
- [74] Mishra, B. K.; Sathyamurthy, N. Stacking Interaction in Pyrazine Dimer *J. Theor.*  
*Comp. Chem.* **2006**, *5*, 609–619.
- [75] Sütay, B.; Tekin, A.; Yurtsever, M. Intermolecular Interactions in Nitrogen-  
Containing Aromatic Systems *Theor. Chem. Accounts* **2012**, *131*, 1120.
- [76] Reyes, A.; Tlenkopatchev, M. a.; Fomina, L.; Guadarrama, P.; Fomine, S. Local MP2-  
Based Method for Estimation of Intermolecular Interactions in Aromatic Molecules.  
Benzene, Naphthalene, and Pyrimidine Dimers. A Comparison with Canonical MP2  
Method *J. Phys. Chem. A* **2003**, *107*, 7027–7031.

- [77] 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.; 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.; 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.; Iyengar, S. S.; Tomasi, J.; ; Cossi, M.; Rega, N.; Millam, J. M.; 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.; Parandekar, P. V.; Mayhall, N. J.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fo, D. J.; Gaussian09, Revision C.01; Gaussian, Inc.; Wallingford CT; 2009.
- [78] Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. Basis-set Convergence in Correlated Calculations on Ne, N<sub>2</sub> and H<sub>2</sub>O *Chem. Phys. Lett.* **1998**, *286*, 243.
- [79] Boys, S.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors *Mol. Phys.* **1970**, *19*, 553–566.
- [80] Cacelli, I.; Cimoli, A.; Livotto, P. R.; Prampolini, G. An Automated Approach for the Parameterization of Accurate Intermolecular Force-Fields: Pyridine as a Case Study. *J. Comp. Chem.* **2012**, *33*, 1055–1067.
- [81] Sherrill, C. D.; Sumpter, B. G.; Sinnokrot, M. O.; Marshall, M. S.; Hohenstein, E. G.; Walker, R. C.; Gould, I. A. N. R. Assessment of standard force field models against high-quality ab initio potential curves for prototypes of pi-pi, CH/pi, and SH/pi interactions. *J. Comp. Chem.* **2009**, *30*, 2187–93.

- [82] Cacelli, I.; Prampolini, G. Parametrization and Validation of Intramolecular Force Fields Derived from DFT Calculations *J. Chem. Theory Comput.* **2007**, *3*, 1803–1817.
- [83] Prampolini, G. Parametrization and Validation of Coarse Grained Force-Fields Derived from ab Initio Calculations *J. Chem. Theory Comput.* **2006**, *2*, 556–567.
- [84] Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham, T. E. Quantum Mechanically Derived AMBER-Compatible Heme Parameters for Various States of the Cytochrome P450 Catalytic Cycle. *J. Comp. Chem.* **2012**, *33*, 119–33.
- [85] Yakovenko, O. Y.; Li, Y.; Oliferenko, A.; Vashchenko, G. M.; Bdzholo, V. G.; Jones, S. J. M. Ab Initio Parameterization of YFF1, a Universal Force Field for Drug-Design Applications. *J. Mol. Model.* **2012**, *18*, 663–73.
- [86] Piquemal, J.-P.; Jordan, K. D. From Quantum Mechanics to Force Fields: New Methodologies for the Classical Simulation of Complex Systems *Theor. Chem. Accounts* **2012**, *131*, 1207.
- [87] Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G. Joyce and Ulysses: Integrated and User-Friendly Tools for the Parameterization of Intramolecular Force Fields from Quantum Mechanical Data. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3736–51.
- [88] Grimme, S. A General Quantum Mechanically Derived Force Field (QMDFFF) for Molecules and Condensed Phase Simulations *J. Chem. Theory Comput.* **2014**, 4497–4514.
- [89] Schmidt, J. R.; Yu, K.; McDaniel, J. G. Transferable Next-Generation Force Fields from Simple Liquids to Complex Materials *Acc. Chem. Res.* **2015**, *48*, 548–556.
- [90] Vanduyfhuys, L.; Vandenbrande, S.; Verstraelen, T.; Schmid, R.; Waroquier, M.; Van Speybroeck, V. QuickFF: A Program for a Quick and Easy Derivation of Force Fields for Metal-Organic Frameworks from Ab Initio Input. *J. Chem. Theory Comput.* **2015**, *36*, 1015–27.

- [91] Cole, D. J.; Vilseck, J. Z.; Tirado-Rives, J.; Payne, M. C.; Jorgensen, W. L. Biomolecular Force Field Parameterization via Atoms-in-Molecule Electron Density Partitioning *J. Chem. Theory Comput.* **2016**, *12*, 2312–2323.

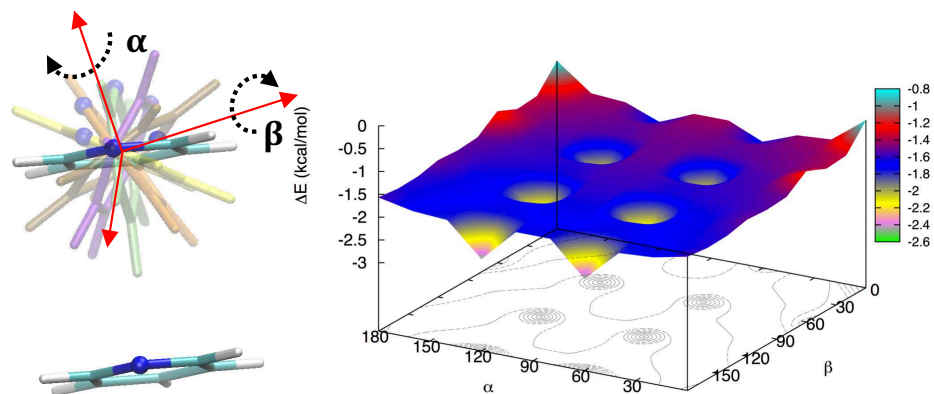


Figure 10: Table of Contents graphic