Density functional theory study of the interaction of vinyl radical, ethyne, and ethene with benzene, aimed to define an affordable computational level to investigate stability trends in large van der Waals complexes

UNIVERSITÀ DEGLI STUDI DI TORINO

Density Functional Theory study of the interaction of vinyl radical, ethyne, and ethene with benzene, aimed to define an affordable computational level to investigate stability trends in large van der Waals complexes.

Andrea Maranzana, **Anna Giordana**, **Antonius Indarto**,† and **Glauco Tonachini**[a]

Dipartimento di Chimica, Università di Torino, Corso Massimo D'Azeglio 48, I-10125 Torino, Italy

Vincenzo Barone

Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126, Pisa, Italy

Mauro Causà

Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale, Università di Napoli "Federico II", Via Cintia, 80126 Napoli, Italy

Michele Pavone

Dipartimento di Scienze Chimiche, Università di Napoli "Federico II", Complesso Universitario di Monte Sant'Angelo, Via Cintia, I-80126 Napoli, Italy

Running title: *DFT on van der Waals complexes*

Keywords: van der Waals / complexes / aromatic hydrocarbons / PAH / adsorption / soot /

†present address: Department of Chemical Engineering, Bandung Institute of Technology
Benny Subianto (Labtek V) Building, Jl. Ganesa 10 Bandung 40132, Indonesia

[a] corresponding author: e-mail address glauco.tonachini@unito.it; fax ++39-011-2367648; group web site: http://www.thecream.unito.it/

e-mail addresses of all Authors:
andrea.maranzana@unito.it
anna.giordana@hotmail.com
antonius.indarto@che.itb.ac.id
vincenzo.barone@sns.it
mauro.causa@unina.it
mipavone@unina.it

Abstract. Our purpose is to identify a computational level sufficiently dependable and affordable to assess trends in the interaction of a variety of radical or closed shell unsaturated hydrocarbons A adsorbed on soot platelet models B. These systems, of environmental interest, would unavoidably have rather large sizes, thus prompting to explore in this paper the performances of relatively low-level computational methods and compare them with higher-level reference results. To this end, the interaction of three complexes between non-polar species, vinyl radical, ethyne, or ethene (A) with benzene (B) is studied, since these species, involved themselves in growth processes of Polycyclic Aromatic Hydrocarbons (PAHs) and soot particles, are small enough to allow high-level reference calculations of the interaction energy ΔE_{AB}. Counterpoise-corrected interaction energies ΔE_{AB} are used at all stages. (1) Density Functional Theory (DFT) unconstrained optimizations of the A–B complexes are carried out, using the B3LYP-D, ωB97X-D, and M06-2X functionals, with six basis sets: 6-31G(d), 6-311 (2d,p), and 6-311++G(3df,3pd); aug-cc-pV*DZ* and aug-cc-pV*TZ*; N07T. (2) Then, unconstrained optimizations by Møller-Plesset second order Perturbation Theory (MP2), with each basis set, allow subsequent single point Coupled Cluster Singles Doubles and perturbative estimate of the Triples energy computations with the same basis sets [CCSD(T)//MP2]. (3) Based on an additivity assumption of (i) the estimated MP2 energy at the Complete Basis Set limit [$E_{\mathbf{MP2/CBS}}$] and (ii) the higher-order correlation energy effects in passing from MP2 to CCSD(T) at the aug-cc-pV*TZ* basis set, $\Delta E_{\mathbf{CC\text{-}MP}}$, a CCSD(T)/CBS estimate is obtained and taken as a computational energy reference.

At DFT, variations in ΔE_{AB} with basis set are not large for the title molecules, and the three functionals perform rather satisfactorily even with rather small basis sets [6-31G(d) and N07T], exhibiting deviation from the computational reference of less than 1 kcal mol^{-1}. The zero-point vibrational energy corrected estimates $\Delta(E_{AB}$+ZPE), obtained with the three functionals and the 6-31G(d) and N07T basis sets, are compared with experimental D_0 measures, when available. In particular, this comparison is finally extended to the naphthalene and coronene dimers and to three π–π associations of different PAHs (R, made by 10, 16, or 24 C atoms) and P (80 C atoms).

I. INTRODUCTION

Soot aerosol contributes in a significant way to the total mass of atmospheric aerosol. Since the relatively disordered growing of the graphenic platelets of soot is generated in the same combustion processes (at relatively low O_2 concentrations) which bring about the formation of polycyclic aromatic hydrocarbons (PAHs) and derivatives, these species share the same nature of soot and can be found in association with it.[1,2] The purpose of the present paper is an attempt to set the scene to study the formation mechanism of PAHs (or soot platelets) in particular when the growing molecular system **A** is adsorbed on a soot platelet **B** ("particle phase"). This situation will involve in general adsorption of a growing PAH-like unsaturated system, and their interaction with small species which have the function of PAH growth building blocks. The overall size of these composite systems, of environmental interest, would unavoidably be rather large.[3]

Therefore, it seems advisable to identify an affordable and sufficiently reliable computational level to describe the possible adsorption of both small molecules or PAH-like species on a locally graphenic, "PAH-like", surface. This necessity prompts to explore the performances of relatively low-level computational methods and their comparison with some reference higher-level results, as those provided by a coupled cluster approach. It is obvious that this comparison limits the size of the test complexes severely. Hence the three small-size title systems have been selected. In the following, we will obviously focus our attention on the performances offered by basis sets of moderate size, used in conjunction with three DFT functionals.

In the end, we will test the performances of these low-level calculations for more extended systems, in which both π–π and CH–π interactions can be present.[4] In particular, (1) the naphthalene dimer offers a test for the computationally less expensive basis sets, since an experimental datum is available. Then, (2) the coronene dimer will be studied, where a comparison with benchmark calculations by Janowski and Pulay is possible.[5,6,7] Finally, (3) the association between naphthalene, pyrene, and coronene (10, 16, and 24 carbon atoms, respectively) and a larger PAH with 80 carbon atoms will allow us to compare the energetics with the exfoliation energy of graphite.

Typical test calculations found in the literature are run on the benzene dimer,[8,9,10,11] but other associations of aromatics have also been studied.[5,6,7,12,13,14,15] Several papers have already appeared in the literature, reporting experimental[16,17,18,19] or computational results,[20,21] (or both)[22] for the benzene-ethene, benzene-ethyne, or the closely related benzene-butadiyne (diacetylene) complexes.[23,24] In particular, we recall the extensive study of the benzene-ethyne and benzene-ethene systems, carried out both experimentally and computationally by Shibasaki, Fujii, Mikami, and Tsuzuki.[25] An energy analysis for a large variety of differently oriented model

systems with benzene, among which systems presenting CH-π, π-π, π^+-π, metal cation-π interactions (the ethyne–benzene complex included) has also been presented by Kim and coworkers.[26] They also discussed in recent years the differences between aromatic π interactions, aliphatic π interactions, and non-π stacking interactions by studying the dimers of ethane, ethene, ethyne, cyclohexane, and benzene,[27] then edge-to-face aromatic interactions,[28] and also the interactions of DNA bases with π systems (benzene, naphthalene, graphene).[29] In fact, many theoretical studies on van der Waals (vdW) complexes between molecules have appeared so far. We can mention in particular a review by Müller-Dethlefs and Hobza,[30] and two feature articles by Sherrill and coworkers.[9,10] In more recent years, an issue of *PCCP* appeared, dedicated to the subject "Stacking Interactions",[31] and a Perspective issue of *J. Chem. Phys.* on "Advances and Challenges in treating van der Waals Dispersion Forces in Density Functional Theory".[32] A very recent review by Riley and Hobza addresses the importance of aromatic interactions in chemistry and biochemistry, concluding that "accurate binding energies for aromatic complexes should be based on computations made at the (estimated) CCSD(T)/ complete basis set limit (CBS) level of theory".[33] This point is widely accepted. However, in the past years, theoretical investigations, focused on π–π and CH–π interactions, were carried out not only at the Coupled Cluster,[34] but also at the Møller-Plesset Perturbation Theory to the second order (MP2)[35] or Density Functional (DFT)[36] theory levels, in an effort to establish a suitable reference and the reliability of the results obtained.[37] In particular this last approach is pertinent and germane to the spirit of the present study. The problem of non-covalent interactions has been addressed by a variety of DFT approaches (DFT calculations are obviously attractive because of their practicability). Local and semilocal (gradient corrected) implementations of DFT have been proved to fail in describing dispersion interactions correctly, mostly giving repulsive potential energy curves.[38,39] However, some functionals offered seemingly a better performance.[40,41] Moreover, various improvements of the DFT scheme have been put forth, of which some implementations introduce an *ad hoc* term to take into account dispersion forces, while other follow different approaches.[42,43,44,45,46] The inclusion of empirical dispersion terms in DFT functionals (and in semiempirical methods)[47] has been reviewed recently.[48]

II. METHODS

DFT functionals. The structures of the complexes are determined initially within the Density Functional Theory (DFT),[36] by fully unconstrained gradient-based energy minimizations,[49] focusing on the performances of the B3LYP–D[11,50,51] and ωB97X–D[52] functionals, and comparing with the M06-2X functional.[42] The first two density functional models have been corrected with the semiempirical pairwise term for dispersion, as formulated by Grimme.[44b] The M06-2X, instead, has been purposely parametrized with a benchmark set of molecules that included vdW complexes. Our choice of DFT models and semi-empirical dispersion terms was motivated by two considerations. The first was related to computational feasibility. Thus, we discarded the vdW density functional proposed by Langreth et al.,[53] because, despite its accuracy and theoretical appeal, vdw-DFT numerical performances are still too low to allow for large-scale simulations. The second one was related to transferability of our results to other related systems. Thus, we recalled to Grimme's DFT-D approach that has proven to be very effective, expecially for carbonaceous species.[54] The effect of using a finer integration grid was explored.[55]

Basis sets. We used, with the mentioned functionals, the Pople basis sets 6-31G(d), 6-311G (2d,p), and 6-311++G(3df,3pd);[56] then Dunning's aug–cc–pV*DZ* and aug–cc–pV*TZ* basis sets,[57] and finally the more recently proposed N07T basis set.[11] Pople's basis sets range from double to triple–ζ in the valence shell only, augmented by diffuse functions on both carbon and hydrogen (++), and carry a different number of polarization functions. Dunning's valence polar-ized double and triple–ζ basis sets carry 3s,2p,1d or 4s,3p,2d,1f functions on C, respectively, and 2s,1p or 3s,2p,1d on H. They are augmented by diffuse functions (aug-), and one for every valence and polarization functions in use for each atom is supplemented (e.g. the aug–cc–pv*TZ* basis has one s, one p, one d, and one f diffuse functions on C, and one s, one d, and one p diffuse functions on H). N07T is instead a valence triple-ζ quality basis set that carries single diffuse p and s functions on C and H, respectively, and only one set of d and p polarization functions on C and H, respectively. It was derived from the primitive set of the aug–cc–pV*QZ* basis set,[57] by a new contraction scheme defined on the basis of atomic DFT(B3LYP) calculations. Given the good radial flexibility accompanied by a minimal presence of polarization functions, its performance is expected to allow an interesting comparison between the roles of radial and angular flexibility. The DFT(B3LYP-D)/N07T computational level has been proved to give satisfactory results for the classical benzene dimer test case.[11] The contraction scheme for each basis set is reported in the headings of Tables 1-3, for the reader's convenience. Counterpoise-corrected interaction energies ΔE_{AB} are used at all stages (see below).

Reference calculations. Frozen core coupled cluster single point energy calculations[34] were also carried out, at the CCSD(T) level and using all basis sets, in correspondence of optimum geometries obtained at the Møller-Plesset second order Perturbation Theory (MP2),[35] without constraints, with the same basis sets. In particular, CCSD(T) and MP2 energies have been used, both with the aug–cc–pV*TZ* basis set, to estimate the higher order correlation energy effects ($\Delta E_{CC\text{-}MP}$). Then, the energy differences between MP2/aug–cc–pV*TZ* and MP2/aug–cc–pV*QZ* have been used to obtain a Complete Basis Set (CBS) limit estimate, by using the two-term extrapolation formula by Halkier et al.,[58] which we label $E_{MP2/CBS}$. Assuming additivity between the two quantities so defined, as oftentimes accepted (see for instance refs. 10 or 26), we estimate CCSD(T)/CBS result (details in the Supplementary Material, Section 1, Table 1S). Hence, this estimate at CCSD(T)/CBS//MP2/aug-ccPV*TZ* is assumed here to represent a dependable coupled cluster reference.[59,60]

Counterpoise corrections. At all stages (DFT, MP2, and CCSD(T) calculations), a counterpoise (CP) correction was applied to get the final ΔE_{AB} values,[61] to better assess the stability of the complexes by correcting for the basis set superposition error (BSSE). In the case of the stablest minimum of the vinyl radical-benzene complex, DFT test were carried out to assess if the CP correction was inclined to affect the inter-moiety distance significantly, and, as a consequence, the CP correction itself. On the basis of the data obtained (see details in the Supplementary Material, Section 2, Figures 2S-4S), it was decided to proceed by single-point CP energy corrections on the geometry optimized without CP. These CP-corrected ΔE_{AB} data are represented graphically in the **Figures 2**, **4**, and **6**, while the numerical values are compared in **Tables 1-3** with and without CP correction.

Thermochemistry. A vibrational analysis was carried out in all cases to define the nature of the critical points, and the thermochemistry assessed. In particular, the zero-point vibrational correction to the energy (ZPE), was used to draw, when possible, a comparison between our computed "$\Delta[E_{AB}$+ZPE]" values with available experimental D_0 values. When the experimental datum is not available, the CP-corrected interaction energy ΔE_{AB} (= $-D_e$) is used to draw comparisons within computational data.

All calculations were carried out using the GAUSSIAN09 system of programs.[62] The MOLDEN program has been exploited for the graphics.[63]

III. RESULTS AND DISCUSSION

1. The molecular models.

The most widely accepted mechanism for PAH and soot growth is the "*H*ydrogen *A*bstraction – C_2H_2 *A*ddition" mechanism (dubbed *HACA* by Frenklach and Wang),[64] which was put forward independently by Bockhorn's[65] and Frenklach's[66] groups in the years 1983-84. Variants were also proposed, as the Bittner-Howard mechanism.[67] These mechanistic hypotheses, which involve a growing unsaturated hydrocarbon (**X**–CH=CH•) radical and ethyne (acetylene), were considered in theoretical studies of gas-phase reactions.[68,69] A simple molecular model could thus involve ethyne (the building unit) and styrenyl (**X** = Ph), as representative of the growing unsaturated radical, though others can naturally be conceived. But here the necessary comparison with some reference high-level result [e.g. CCSD(T)] mandatorily limits the size of the complexes that can be considered. To this end, the interaction of ethyne, ethene, or the vinyl radical (representative of the moiety **A**) with benzene (representative of the moiety **B**) is studied at DFT, MP2, and coupled cluster. The stable geometries of their vdW complexes are investigated, mainly to compare the performances of several computational levels, with particular attention to the less demanding ones (interaction energies ΔE_{AB} in Tables 1-3). These species (involved themselves in growth processes of PAHs and soot particles) are small enough to allow high-level reference calculations of ΔE_{AB}. The models chosen here are indeed – forcedly – severely limited as regards the size: on one side, benzene would stay for larger aromatics (or even soot platelets!); on the other, the vinyl radical is the smallest representative of the growing **X**–CH=CH• radical, where **X** (here hydrogen!) stays for the aromatic part of the growing radical.

The molecular systems chosen are all non-polar. In a series of papers, in which two of the functionals chosen here (ωB97x-D and M06-2X) were also used, Dannenberg and coworkers showed that, if polar interactions are present, dispersion interactions become relatively less important.[70] When induction plays an important role, as in peptides, these functional tend to overestimate the interaction energy.[71] In studying the interaction of pyridine and p-benzo-quinone with a variety of geometry arrangements and functionals (among which ωB97x-D and M06-2X, used here), they also found that functionals designed to treat dispersion behave erratically as the predictions of the most stable structure vary considerably.[72] Similarly, the interaction of phenylalanines has been very recently investigated.[73] All these systems differ from those studied in that they are polar molecules. Our present results are to be intended as pertinent to situations where induction plays a limited role (with the possible exception of ethyne).[74]

The final choice, hopefully coming from the present study, is thus expected to come from a trade-off between dependability and feasibility, keeping in mind that larger and more realistic systems will be too demanding for high-level approaches. We will begin by presenting the results obtained here, at a variety of computational levels, for benzene-ethyne (§ 2) and benzene-ethene (§ 3), for which a comparison can be drawn with previous studies. Then we will similarly discuss the features of the benzene-vinyl radical complex (not studied so far to our knowledge), which can be compared with those obtained for the previous systems (§ 4).

Since we are particularly interested in the performances of the three DFT functionals chosen when used with rather small and inexpensive basis sets [6-31G(d) and N07T], we will compare them with experimental D_0 data, when available (**Table 4**). Finally, we are also interested in π–π interactions (adsorption of the growing aromatic system), but the very small complexes mentioned above are kept together, even when the geometry maximizes π–π interactions, at best by a tiny interaction energy. Then, in almost all the relevant complexes, C–H···π interactions are present. Then, having defined some suitable computational level, to assess its performances when π–π interactions are dominant, we have also carried out optimizations for the stacked naphthalene dimer. For this dimer, in which mainly dispersion forces operate, we have compared computationally estimated and measured D_0 values (**Table 5**). In the end, a further test is carried out to determine the association energies between naphthalene, pyrene, or coronene (10, 16, or 24 carbon atoms, respectively) and a larger PAH, circumcircumpyrene, with 80 carbon atoms. These are aimed to assess a trend against another experimental datum, the exfoliation energy of graphite.

2. The benzene-ethyne complex.

Here, similarly to what was found in previous studies, two critical points on a very flat energy surface were found upon energy minimization (**Figure 1**).[16] In one, labeled as "⊥", ethyne is perpendicular to benzene, pointing one hydrogen toward the center of the ring (structure **1a**). The other critical point, labeled as "**||**", sees ethyne parallel to the benzene molecular plane, and one of the two π systems of ethyne can be thought of as pointing toward the π system of benzene (structure **1b**). In the ⊥ form **1a**, which is clearly more stable, the distance between the ethyne C closest to benzene and the benzene plane is ca. 3.5 Å. In the case of the **||** form **1b**, the two moieties are separated by a distance of ca. 3.5 Å, taken as the distance between the midpoint of the ethyne carbon-carbon bond and the benzene plane.

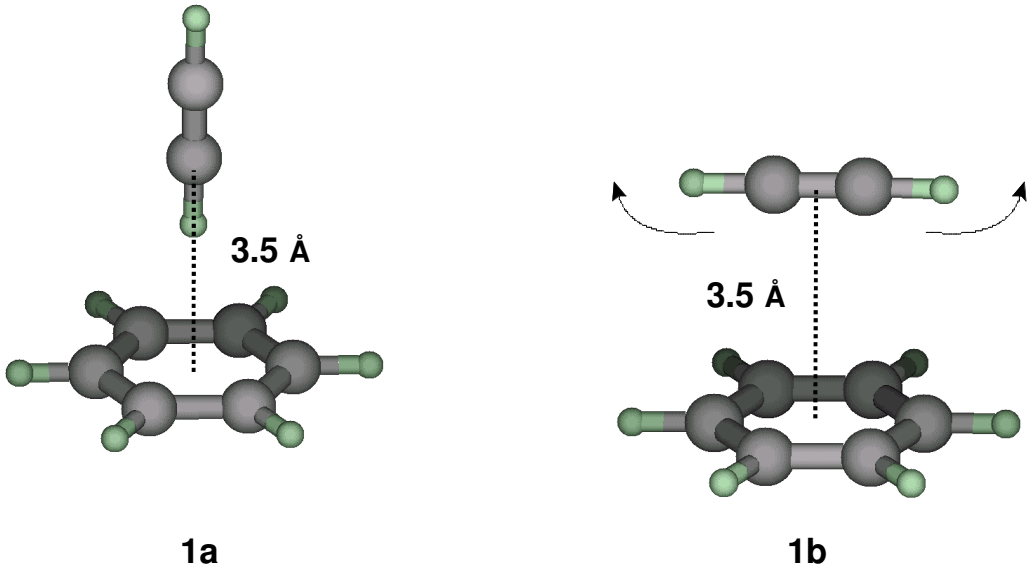
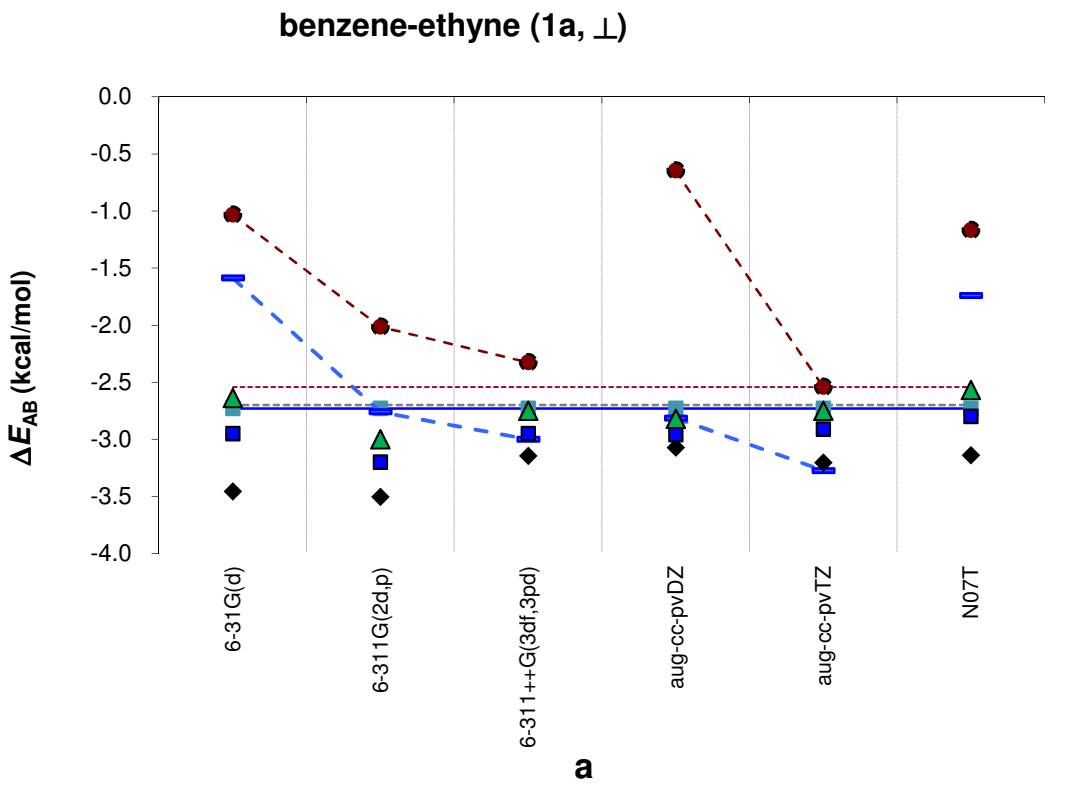


Figure 1

Van der Waals complexes of ethyne and benzene **1a** and **1b**. The **||** arrangement **1b** is classified at some computational level as a transition structure between two equivalent ⊥ stable structures **1a** (as indicated by the curly arrows). At other computational levels, it is classified instead as an energy minimum. In any case, just to provide some information on the inter-moiety distance, the optimum M06-2X/aug-cc-pV*TZ* values for the inter-moiety distances are reported.

Figure 2 summarizes the results obtained at the various computational levels for the energy minima **1a** and **1b**. Data for both structures (raw ΔE_{AB} values together with the CP-corrected ones) are reported in **Table 1**: inspection of these data shows that variations in ΔE_{AB} estimates with the basis set are (as expected) more moderate if CP correction is introduced. Therefore, given that the purpose of the present study is to assess if the use of modest basis sets for large systems can be acceptable, CP corrections appears to be mandatory. The ⊥ arrangement of the two molecules, **1a**, is described as lower in energy than the separate molecules (zero reference level) in all cases. The DFT results are apparently less basis set dependent than CCSD(T), or MP2. These theory levels, though, show a more regular downwards trend upon basis set extension, with MP2 inclined to overbind (dashed lines in **Figure 2**). DFT data show less regular trends, as already observed by Dannenberg and coworkers.[70] In this arrangement, given the sp hybridation of the carbon atom and the consequent significant polarity of the C–H bond,[75] electrostatic interactions are more effective than in the **||** arrangement. Therefore, this association has been considered as an example of π–hydrogen bond based cluster, rather than a van der Waals complex[23,76] (see the analysis carried out in ref. 25). It is also interesting to examine, in ref. 25, Table 2, with relevant text, where Tsuzuki and coworkers point out that dispersion is the main contributor to the complex stability, while an electrostatic term plays a non-negligible role (in the case of ethyne, but much less for ethene).



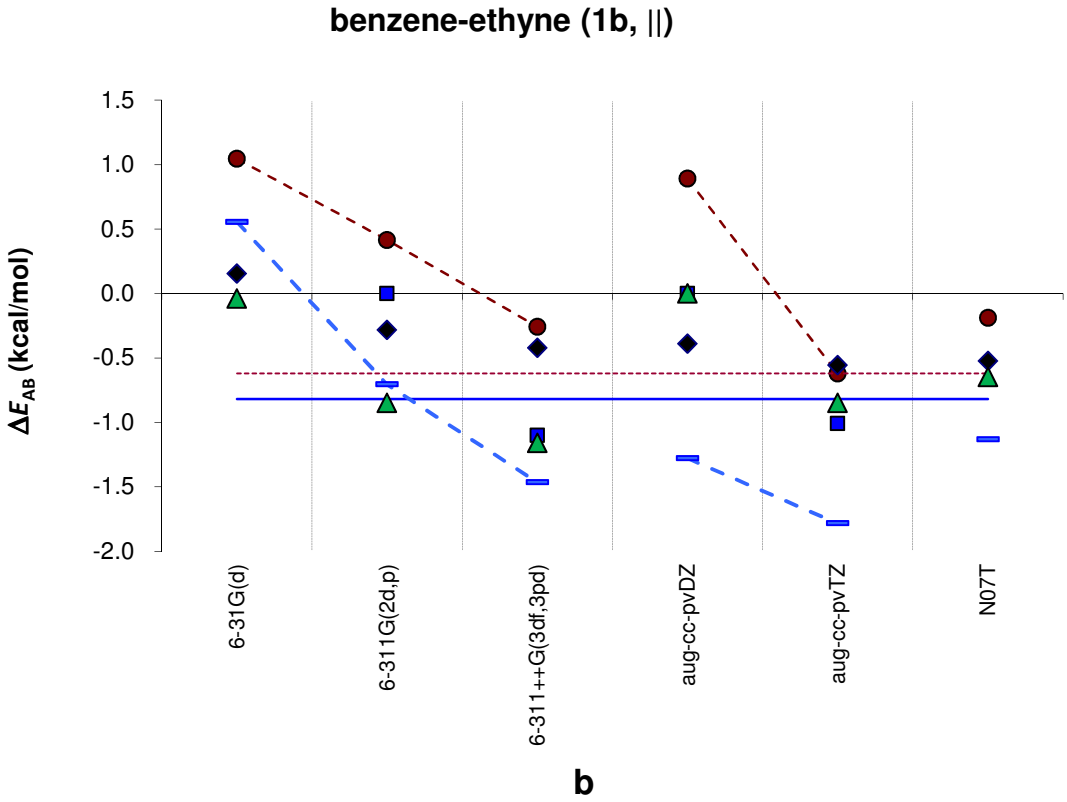


Figure 2

CP-corrected stabilization energy ($\Delta E_{AB} = -D_e$) of the complexes of ethyne and benzene for the **1a** (⊥) and **1b** (**||**) arrangements (plots **a** and **b**, respectively). Black diamonds: B3LYP-D; blue squares: ωB97X-D; green triangles: M06-2X. Blue dashes and dashed lines: MP2. Dark red filled circles and dashed lines: CCSD(T). Missing points in **b**: the parallel structure corresponds to a TS. Though the ΔE_{AB} values displayed in **a** are not corrected for zero-point vibrational energy (available only at DFT), the experimental D_0 = 2.7 kcal mol^{-1} is reported, for sake of comparison, as a dashed gray line (see however **Table 4**). Our computational reference, [CCSD(T)/CBS], is indicated by a blue continuous line located at: **1a**, -2.73; **1b**, -0.82. [CCSD(T)/aug-cc-pV*TZ*] is indicated by a dashed dark red horizontal line .

A similar analysis was carried out by Singh et al.,[26] who found, for an inter-moiety distance of 3.54 Å, an interaction energy of -2.85 kcal mol^{-1} at CCSD(T)/CBS. The dispersion and electrostatic contributions were evaluated as -3.10 and -2.23 kcal mol^{-1}, respectively.

The CCSD(T) reference was already deemed to be quite good by Shibasaki et al.,[25] whose best estimate was 2.4 kcal mol^{-1}, in correspondence of MP2/cc-pV*TZ* geometries. Later papers by Tsuzuki and others[77] report an estimate of -2.8 kcal mol^{-1} (compare with Table 1 of ref. **80**) The experimental estimate of the interaction energy is D_0 = 2.7 ± 0.2 kcal mol^{-1}).[25] We approach to this value with our D_e estimates at the CCSD(T)/aug-cc-pV*TZ* and CCSD(T)/6-311++G(3df,3pd) computational levels (-2.5 and -2.3 kcal mol^{-1}, respectively). If we take CCSD(T)/CBS as a reference for D_e, (for **1a**, ΔE_{AB} = -2.7), **Figure 2a** indicates that some DFT calculations for the ⊥ **1a** structure tend to overbind to some extent, in the order B3LYP-D > ωB97X–D > M06-2X.

In view of our goal, it is noteworthy that the two more modest basis sets, 6-31G(d) and N07T, give a good result, especially when they are associated with the M06-2X and ωB97X-D functionals: the observed discrepancy can be deemed acceptable. By contrast with the ⊥ arrangement, the **||** structure **1b** is found either less stable, or more stable, than the separate moieties, depending on the computational level. At the DFT level, B3LYP-D sets it slightly above the zero when used with 6-31G(d). This is not the case in the other calculations, where DFT describes it as slightly bound, by ca. 1 kcal mol^{-1} at most.

The **||** arrangement was always looked for and optimized as an energy minimum. Yet, the harmonic vibrational analysis carried out on it (1) indicates in any case a direction in which the *E* surface is exceedingly flat, and (2) produces, with some methods, one normal mode with a very low imaginary vibrational frequency (this is indicated in **Table 1** by TS, transition structure). Just to give an example, while ν_1 = 40.7 cm^{-1} at B3LYP-D/aug-cc-pV*DZ*, ν_1 = 7.6*i* cm^{-1} at MP2/aug-cc-pV*DZ* level.[78] Since structures of this kind, in which dispersion forces play a major role, acquire a larger stability as the model system increases its size, structure **1b** will be not discussed any further. Instead, the naphthalene and coronene dimers, and the π–π association between two rather large PAHs (possibly interesting examples in which dispersion dominates) will be discussed to some extent in the last part of this section.

Table 1: ΔE_{AB} Interaction Energies[a] of the vdW Ethyne-Benzene Complexes.

	basis set:	**6-31G(d)**	**6-311G(2d,p)**	**6-311++G(3df,3pd)**	**aug-cc-pV*DZ***	**aug-cc-pV*TZ***	**N07T**
C/H contraction scheme:[b]		(10s4p1d / 4s)	(11s5p2d / 5s1p)	(12s6p3d1f / 6s3p1d)	(16s5p2d / 5s2p)	(17s6p3d1f / 6s3p2d)	(12s7p1d / 7s1p)
Theory level	arrangement[c]	→ [3s2p1d / 2s]	→ [4s3p2d / 3s1p]	→ [5s4p3d1f / 4s3p1d]	→ [4s3p2d / 3s2p]	→ [5s4p3d1f / 3s3p2d]	→ [6s4p1d / 4s1p]
***B3LYP-D*:**	⊥ **1a**	(-4.3) **-3.5**	(-4.0) **-3.5**	(-3.6) **-3.1**	(-3.9) **-3.1**	(-3.4) **-3.2**	(-3.3) **-3.1**
	|| **1b**	(-1.3) **0.2**	(-0.9) **-0.3**	(-0.7) **-0.4**	(-0.7) **-0.4**	(-0.6) **-0.6**	(-0.6) **-0.5**
ω*B97X-D*:	⊥ **1a**	(-3.6) **-2.9**	(-3.6) **-3.2**	(-3.4) **-3.0**	(-3.7) **-3.0**	(-3.1) **-2.9**	(-2.9) **-2.8**
	|| **1b**	**TS**	**TS**	(-1.4) **-1.1**	**TS**	(-1.2) **-1.0**	**TS**
***M06-2X*:**	⊥ **1a**	(-3.4) **-2.6**	(-3.4) **-3.0**	(-3.7) **-2.8**	(-3.7) **-2.8**	(-3.0) **-2.7**	(-2.7) -2.6
	|| **1b**	(-1.3) **-0.04**	(-1.4) **-0.9**	(-1.5) **-1.2**	**TS**	(-1.0) **-0.9**	(-0.7) **-0.6**
***MP2*:**	⊥ **1a**	(-3.6) **-1.6**	(-4.0) **-2.8**	(-4.5) **-3.0**	(-5.9) **-2.8**	(-4.4) **-3.3**	(-4.6) **-1.7**
	|| **1b**[c]	(-0.8) **0.6**	(-2.0) **-0.7**	(-2.5) **-1.5**	(-3.0) **-1.3**	(-2.5) **-1.8**	(-3.1) **-1.1**
***CCSD(T)*:**[d]	⊥ **1a**	(-3.0) **-1.0**	(-3.2) **-2.0**	(-3.9) **-2.3**	(-3.9) **-0.6**	(-3.6) **-2.5**	(-4.0) **-1.2**
	|| **1b**[c]	(-0.3) **1.0**	(-0.8) **0.4**	(-1.3) **-0.3**	(-0.5) **0.9**	(-1.2) **-0.6**	(-2.1) **-0.2**

[a] CP-corrected energy differences (in bold) relative to the separate moieties; uncorrected values in parenthesis. Units: kcal mol^{-1}.
[b] (number of primitive gaussians of each angular momentum) → [number of contracted gaussians].
[c] When the || arrangement corresponds to a 1st order saddle point it is indicated by **TS**. It is so also at the MP2 level.
[d] The CCSD(T)/CBS reference value is, for **1a**: -2.73 kcal mol^{-1}.

3. The benzene-ethene complex.

Two critical points were found also in this case. The relevant structures could be dubbed "side ⊥" (**2a**) and "top ⊥" (**2b**) (**Figure 3**). They correspond to structures G and F-1, respectively, reported by Ōki, Takano, and Toyota (see their Figure 1),[79] who studied this system at HF, MP2, and DFT(B3LYP) levels of theory, all with the 6-31G(d,p) basis set. In structures **2a** and **2b**, the hydrogen atoms are significantly involved in the interaction. The role of the C–H···π interaction is discussed within a comparison of the ethyne-benzene and ethene-benzene complexes in ref. 25 (compare their Figure 5). Here we can just mention again that dispersion is indicated in papers by Tsuzuki and coworkers as the main contributor to the stability of these complexes, accompanied by an electrostatic term, which plays for ethene a lesser role than for ethyne.[20,25,74,80] The closest C–C distance, in the case of the side ⊥ form **2a**, is ca. 3.7 Å. In the top ⊥ form **2b**, in which a C–H bond points toward a benzene carbon, the distance between the hydrogen involved in the interaction and the carbon is ca. 2.9 Å. In Table 2 of ref. **79**, C-H···H-C distances of 2.65 (B3LYP) and 2.51 (MP2) Å are indicated for G, and "ethene-H···closest benzene C" distance of 2.88 Å (MP2) is reported for F-1 (our H-H distance in **2a** is 2.42 at M06-2X/aug-cc-pV*TZ;* compare Table 2S in Section 3 of the Supplementary Material, and Section 6, for a more complete set of data).

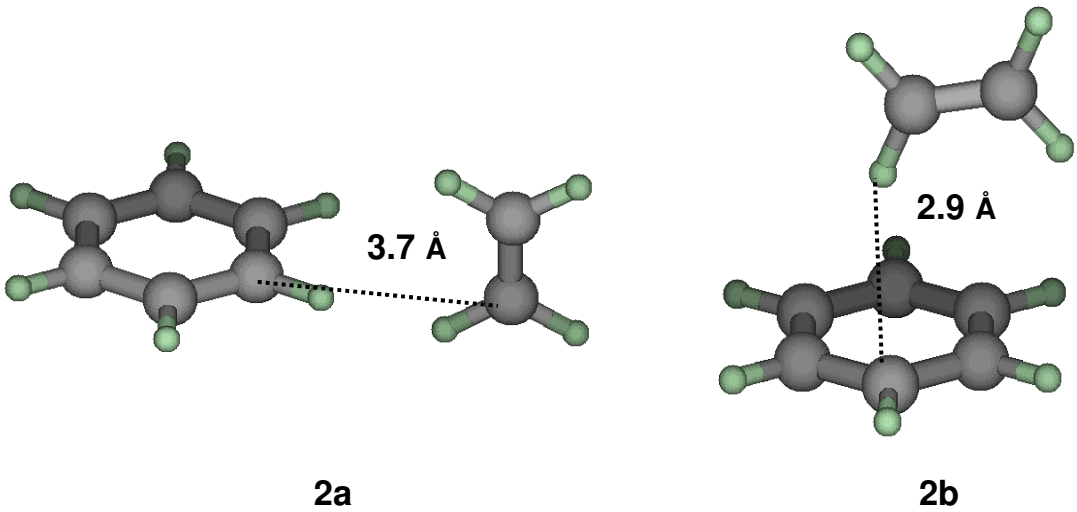


Figure 3

Complexes of ethene and benzene. The **2a** (side ⊥) and **2b** (top ⊥) arrangements. In both cases, to provide some information on the inter-moiety distance, the optimum M06-2X/aug-cc-pV*TZ* values for the inter-moiety distances are reported.

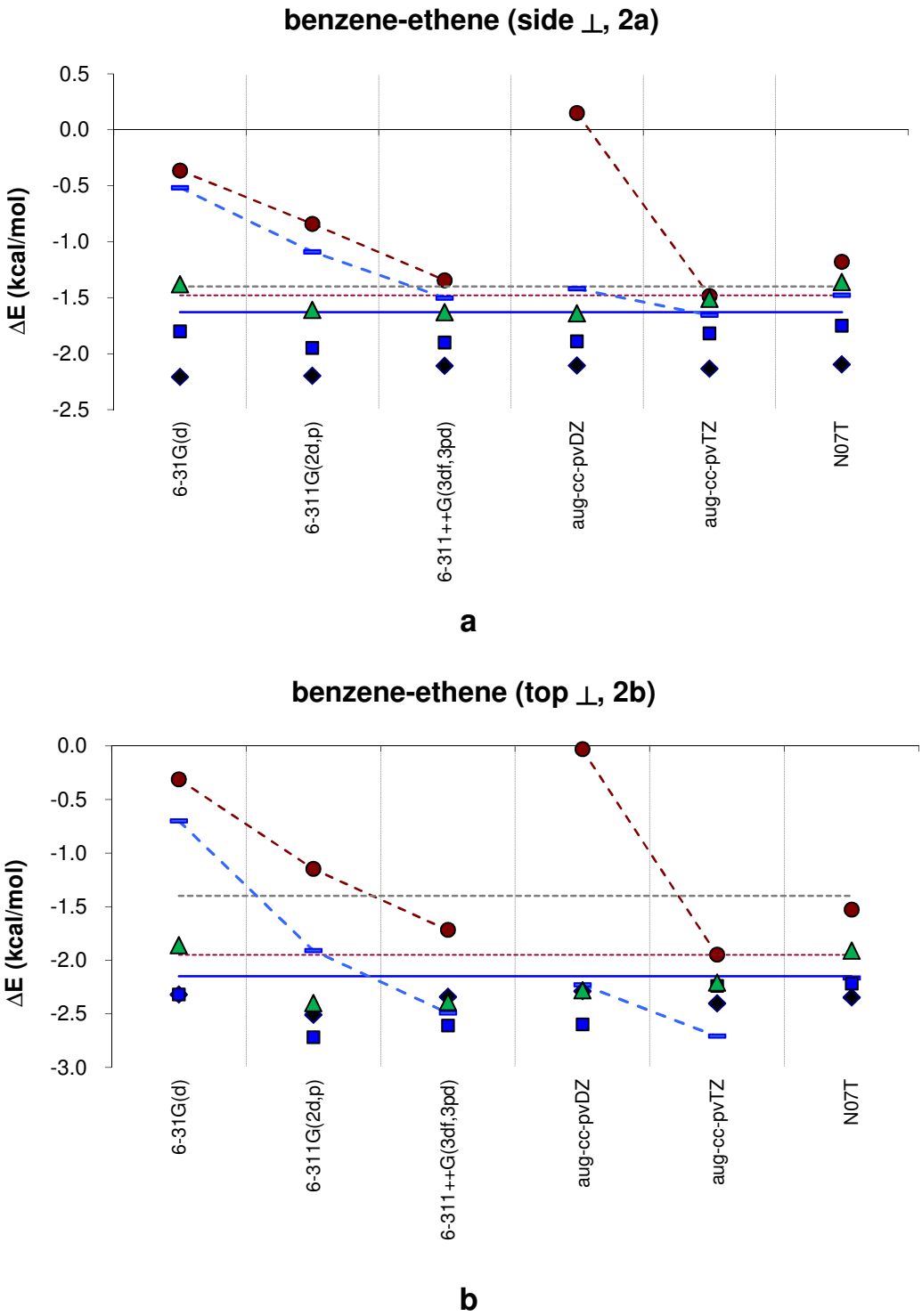


Figure 4

CP-corrected stabilization energy ($\Delta E_{AB} = -D_e$) of the complexes of ethene and benzene for the **2a** (side ⊥) arrangement and the **2b** (top ⊥) arrangement (plots **a** and **b**, respectively). Black diamonds: B3LYP-D; blue squares: ωB97X-D; green triangles: M06-2X. Blue dashes and dashed lines: MP2. Dark red filled circles and dashed lines: CCSD(T). Though the ΔE_{AB} values displayed here are not corrected for zero-point vibrational energy (available only at DFT), the experimental D_0 = 1.4 kcal mol^{-1} is reported for sake of comparison as a dashed gray line (see however **Table 4**). Our computational reference, [CCSD(T)/CBS], is indicated by a blue continuous line, at -1.63 for **2a**, and -2.15 for **2b**; [CCSD(T)/aug-cc-pV*TZ*] is indicated by a dashed dark horizontal red line.

Table 2 collects the original, raw, ΔE_{AB} data together with the CP-corrected values. It shows again that variations in ΔE_{AB} values upon change of basis set are, not unexpectedly, smaller if CP correction is applied. Given the purpose of this study, CP correction is again indicated as compulsory. In **Figure 4** the stability of the complexes **2a** and **2b** with respect to the separate ethene and benzene moieties (zero level) is reported. The **2b** form is more stable by ca. 0.4 (or 0.5) kcal mol^{-1}, at CCSD(T)/aug-cc-pV*TZ* (or CBS). Also all DFT functionals describe it as somewhat more stable. DFT is in some cases overbinding to some extent with respect to our best D_e reference, and we observe for **2a** the same trend seen for the preceding case: B3LYP-D > ωB97X–D > M06-2X. The trend is not as regular for **2b**. We also notice a modest basis set dependence of the DFT data (but not a regular trend upon basis set enlargement; compare ref. **70**). By contrast, ΔE_{AB} changes with the basis set are significantly larger at coupled cluster and MP2,[81] though these theory levels apparently exhibit some convergent behavior with basis set extension. MP2 seems again inclined to overbind. In particular, the CCSD(T)/aug-cc-pV*DZ* datum underestimates ΔE_{AB} remarkably and departs more from the D_e reference than any DFT datum. In **2a** and **2b**, with the more modest basis sets, especially M06-2X and ωB97X-D provide results close to the CCSD(T)/CBS reference (-1.63 for **2a**, and -2.15 for **2b**) one approaching from above the other from below. It could be compared with those from Tsuzuki's group (benzene-ethene interaction energy at -2.06 – -2.17 kcal mol^{-1}). Given that the experimental estimate of the interaction energy is D_0 = 1.4 ± 0.2 kcal mol^{-1}, the chosen reference appears to be rather good. Our estimates at the CCSD(T)/aug-cc-pV*TZ* and CCSD(T)/6-311++G(3df,3pd) computational levels are: for **2a**, D_e = -1.5 and -1.3 kcal mol^{-1}, respectively. For **2b** they fall not as close: D_e = -1.9 and-1.7 kcal mol^{-1}, respectively. We can further notice that, at coupled cluster, N07T falls not far from our chosen reference (at variance with the ethyne result for **1a**). The CCSD(T) performance had already been discussed by Shibasaki et al., whose best computational estimate was 1.7 kcal mol^{-1}.[25]

Table 2: ΔE_{AB} Interaction Energies[a] of the vdW Ethene-Benzene Complexes.

	basis set:	**6-31G(d)**	**6-311G(2d,p)**	**6-311++G(3df,3pd)**	**aug-cc-pv*DZ***	**aug-cc-pv*TZ***	**N07T**
C/H contraction scheme:[b]		(10s4p1d / 4s)	(11s5p2d / 5s1p)	(12s6p3d1f / 6s3p1d)	(16s5p2d / 5s2p)	(17s6p3d1f / 6s3p2d)	(12s7p1d / 7s1p)
Theory level	arrangement	→ [3s2p1d / 2s]	→ [4s3p2d / 3s1p]	→ [5s4p3d1f / 4s3p1d]	→ [4s3p2d / 3s2p]	→ [5s4p3d1f / 3s3p2d]	→ [6s4p1d / 4s1p]
***B3LYP-D*:**	**side ⊥ 2a**	(-3.4) **-2.2**	(-2.7) **-2.2**	(-2.3) **-2.1**	(-2.7) **-2.1**	(-2.2) **-2.1**	(-2.2) **-2.1**
	top ⊥ 2b	(-3.4) **-2.3**	(-3.1) **-2.5**	(-2.8) **-2.3**	(-3.0) **-2.3**	(-2.5) **-2.4**	(-2.5) **-2.3**
ω*B97X-D*:	**side ⊥ 2a**	(-2.6) **-1.8**	(-2.4) **-1.9**	(-2.1) **-1.9**	(-2.4) **-1.9**	(-1.9) **-1.8**	(-1.8) **-1.8**
	top ⊥ 2b	(-3.1) **-2.3**	(-3.2) **-2.7**	(-3.0) **-2.6**	(-3.2) **-2.6**	(-2.7) **-2.2**	(-2.5) **-2.2**
***M06-2X*:**	**side ⊥ 2a**	(-2.3) **-1.4**	(-2.1) **-1.6**	(-1.9) **-1.6**	(-2.2) **-1.6**	(-1.7) **-1.5**	(-1.5) **-1.4**
	top ⊥ 2b	(-2. 8) **-1.9**	(-2.9) **-2.4**	(-2.9) **-2.4**	(-3.1) **-2.3**	(-2.4) **-2.2**	(-2.1) **-1.9**
MP2:	**side ⊥ 2a**	(-1.6) **-0.5**	(-2.1) **-1.1**	(-2.2) **-1.5**	(-2.8) **-1.4**	(-2.4) **-1.7**	(-2.8) **-1.5**
	top ⊥ 2b	(-2.4) **-0.7**	(-3.3) **-1.9**	(-4.0) **-2.5**	(-5.0) **-2.2**	(-3.7) **-2.7**	(-3.8) **-2.2**
***CCSD(T):*[c]**	**side ⊥ 2a**	(-1.4) **-0.4**	(-1.9) **-0.8**	(-2.0) **-1.3**	(-1.3) **0.1**	(-2.1) **-1.5**	(-2.5) **-1.2**
	top ⊥ 2b	(-1.9) **-0.3**	(-2.5) **-1.1**	(-3.2) **-1.7**	(-2.9) **0.0**	(-3.0) **-1.9**	(-3.2) **-1.5**

[a] CP-corrected energy differences (in bold) relative to the separate moieties; uncorrected values in parenthesis. Units: kcal mol^{-1}.
[b] (number of primitive gaussians of each angular momentum) → [number of contracted gaussians].
[c] The CCSD(T)/CBS reference values are, for **2a**, -1.63, and **2b**, -2.15 kcal mol^{-1}.

4. The benzene-vinyl radical complex.

We will label as "perpendicular" (⊥), collectively, those arrangements in which the two π systems are "perpendicular" in an approximate way, i.e. where vinyl interacts with benzene mostly with one hydrogen (**Figure 5**, structures **3a-3d**). The parallel minimum (**||**) of the vinyl radical-benzene association is defined as that presenting the direct interaction of the two π systems, which point one towards the other (structure **3e** in **Figure 5**). A measure of the closeness between the two moieties can be indicated, in **3a-d**, by the distance between the vinyl hydrogen more involved in the interaction and the closest carbon in benzene (2.8 – 2.9 Å).[82] Structure **3e** is found only with the B3LYP-D functional, since the optimization ends up in one ⊥ minimum when using the other two functionals; in this case we could look at the closest C–C distance, ca. 3.6 Å. Structure **3e** is described as bound by ca. 1 kcal mol^{-1} with the largest basis sets. **Table 3** collects the CP-corrected ΔE_{AB} data displayed in **Figure 6** together with the raw data. They confirm again that variations in ΔE_{AB}'s with basis sets are less important if the CP correction is applied, thus strongly suggesting this procedure within the intention of the present study. The different ⊥ complexes **3a-d** are lower in energy than the separate systems with all functionals, and are endowed with comparable stability (**Table 3**), with complex **3a** slightly more stable in most cases. At CCSD(T)/aug-cc-pV*TZ* the **3a** form is the most stable, by ca. -2.8 kcal mol^{-1}. Also B3LYP-D and ωB97X–D describe it as the most stable (at -2.6 and -3.0 kcal mol^{-1}), but M06-2X, though setting it at -3.0 kcal mol^{-1}, puts **3d** slightly lower, at -3.1 kcal mol^{-1}.

All DFT functionals show a modest basis set dependence. As noted before, the coupled cluster data are by contrast more basis set dependent, though they appear to converge clearly to an asymptote from above. DFT is in most cases somewhat overbinding with respect to our best D_e reference, CCSD(T)/CBS, with both overbinding and basis set dependence more pronounced for **3d**. The values obtained by the different functionals are rather close one to the other for **3a-c**, while in **3d** we observe the opposite trend with respect to the preceding case: B3LYP-D < ωB97X–D < M06-2X, with a larger dispersion of data. As in the preceding case, ΔE_{AB} changes with the basis set are significantly larger at coupled cluster and MP2, though they show some convergent behavior with basis set extension, and a MP2 tendency to give too negative values, as in the preceding cases. In particular, we note that the CCSD(T)/aug–cc–pV*DZ* datum underestimates ΔE_{AB} significantly and in most cases moves off from the D_e reference more than any DFT result. In the Supplementary Material, Section 4, the atomic spin densities are reported together with the expectation values of the total spin operator. It can be seen there that only

tiny spin polarization effects (or none) are induced by the vinyl radical into the interacting benzene.

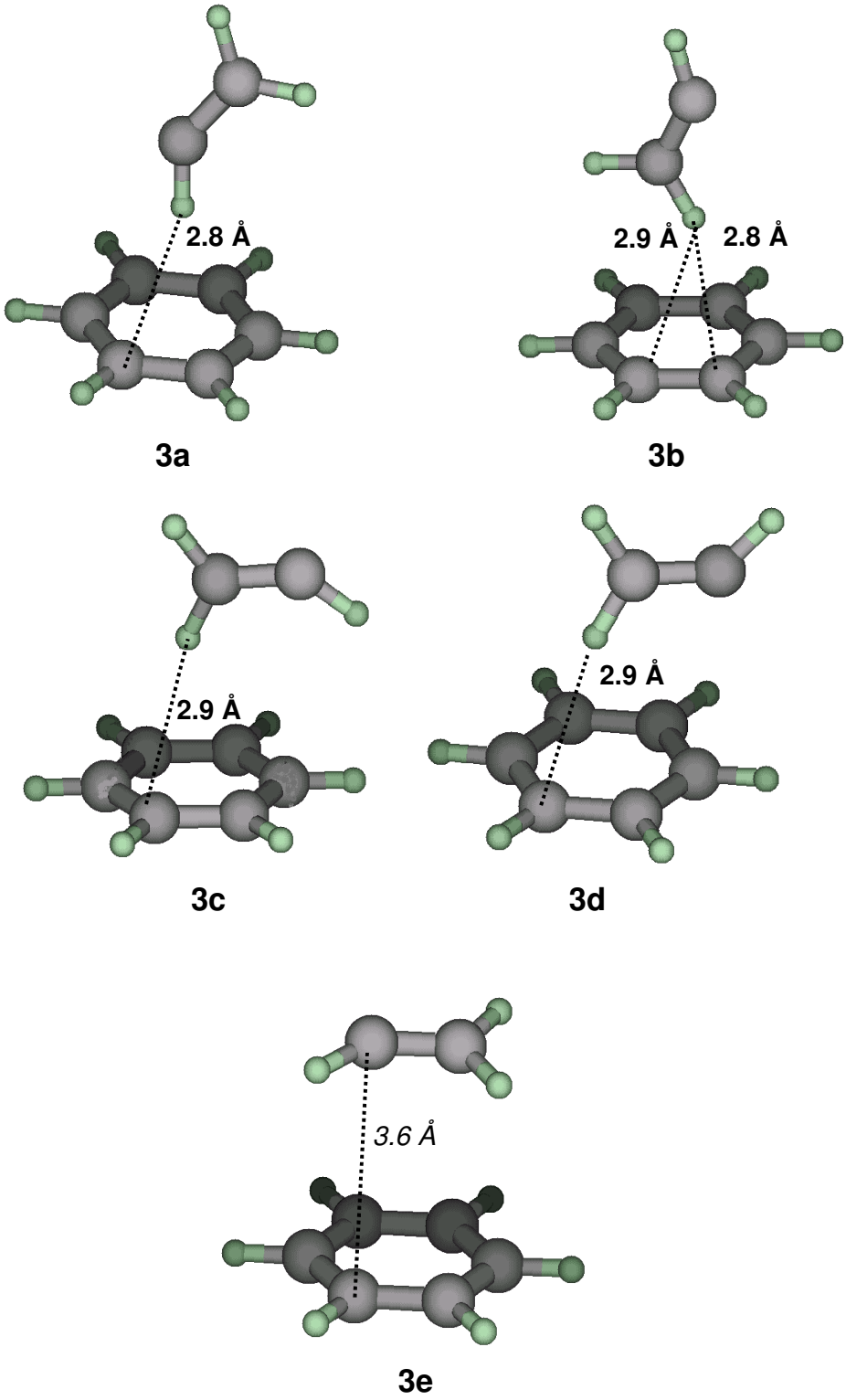
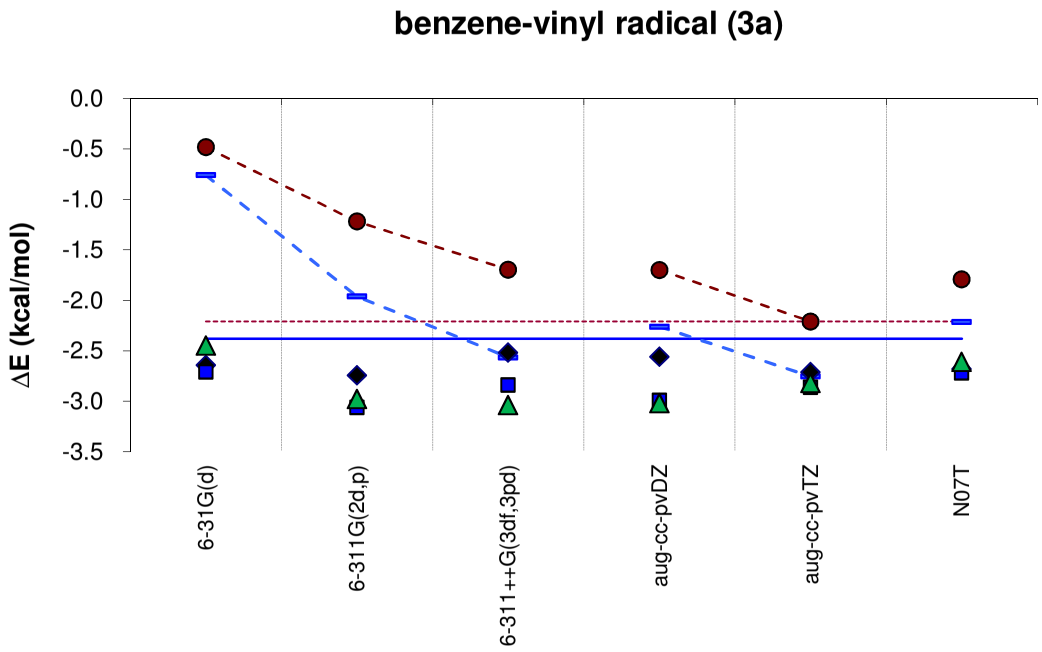
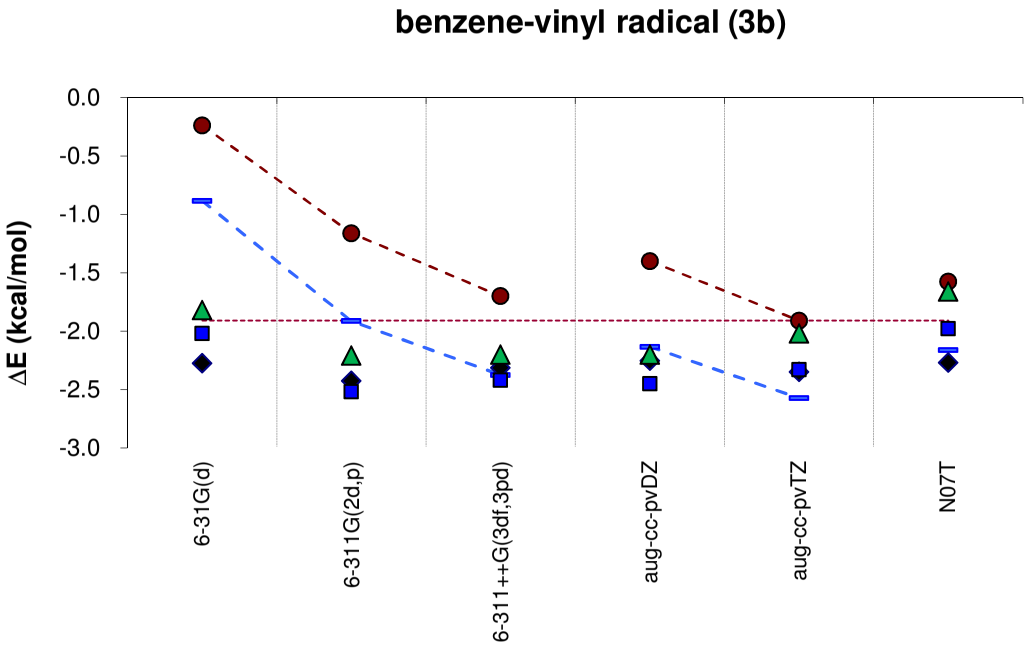


Figure 5

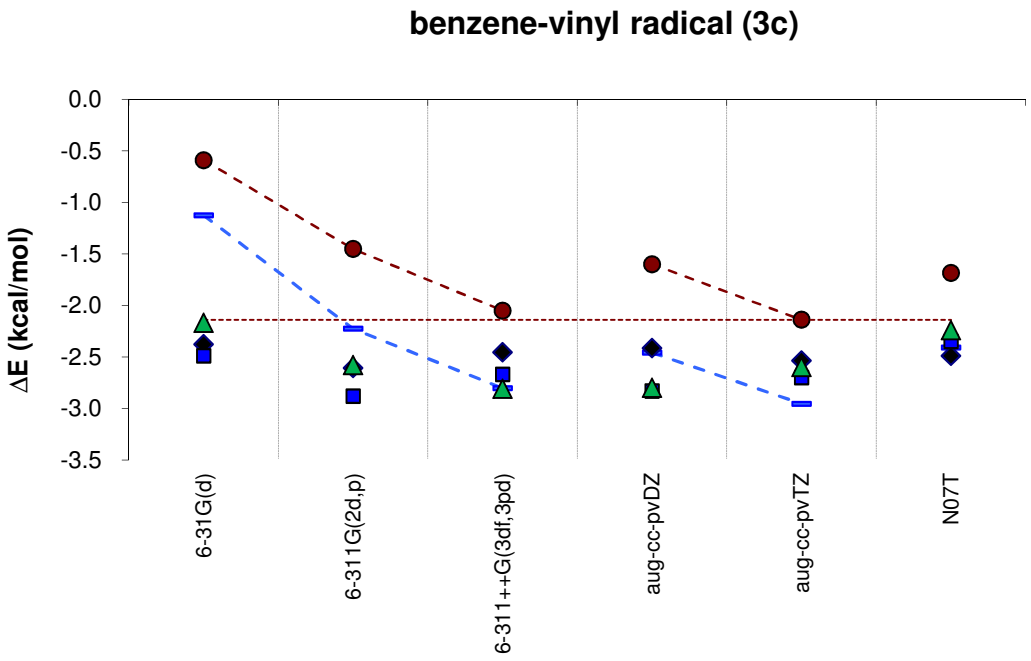
Complexes of the vinyl radical and benzene, labeled as structures **3a-e**. For **3a-d**, to provide some information on the inter-moiety distance, the optimum M06-2X /aug-cc-pV*TZ* values for the distance between one vinyl atom and one benzene carbon are reported. For **3e**, not found at M06-2X and ωB97X-D, the B3LYP-D value is reported (italic).


benzene-vinyl radical (3a)
ΔE (kcal/mol)
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
6-31G(d)
6-311G(2d,p)
6-311++G(3df,3pd)
aug-cc-pvDZ
aug-cc-pvTZ
N07T

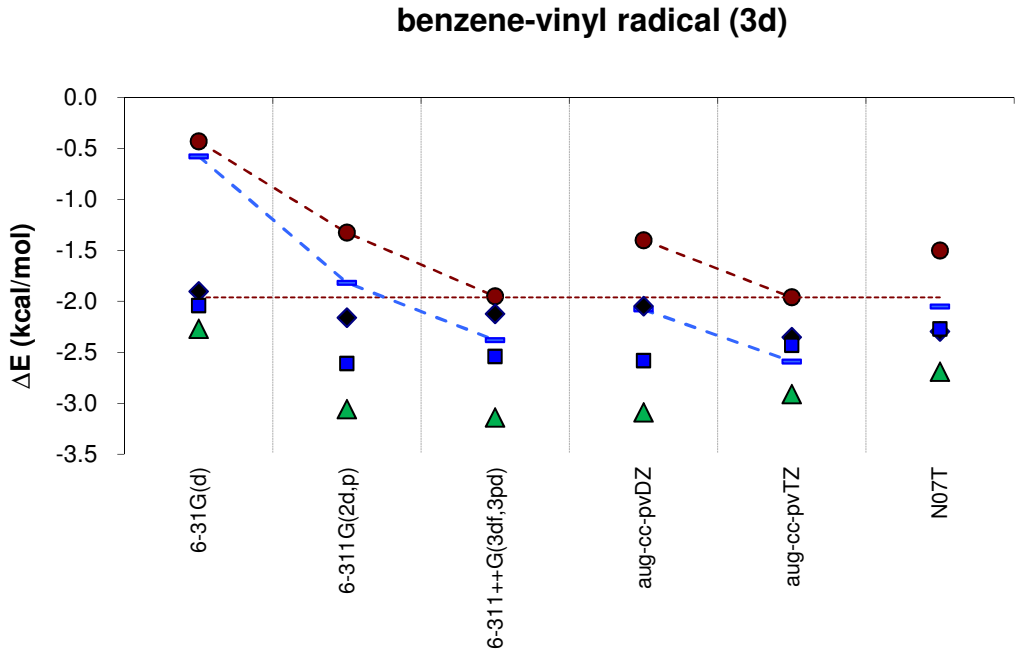
a


benzene-vinyl radical (3b)
ΔE (kcal/mol)
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
6-31G(d)
6-311G(2d,p)
6-311++G(3df,3pd)
aug-cc-pvDZ
aug-cc-pvTZ
N07T

b



c



d

Figure 6

CP-corrected stabilization energy (ΔE_{AB}) of the **3a-d** complexes of vinyl and benzene. Black diamonds: B3LYP-D; blue triangles: M06-2X; blue squares: ωB97X-D. Blue dashes: MP2. Dark red filled circles: CCSD(T). Our computational reference, [CCSD(T)/CBS], is indicated by a blue continuous line, located at: **3a**, -2.38; **3b**, -2.04;**3c**, -2.30; **3d**,-2.11. [CCSD(T)/aug-cc-pV*TZ*] is indicated by a dashed dark red horizontal line.

Table 3: ΔE_{AB} Interaction Energies[a] of the vdW Vinyl Radical-Benzene Complexes.

	basis set:	6-31G(d)	6-311G(2d,p)	6-311++G(3df,3pd)	aug-cc-pv*DZ*	aug-cc-pv*TZ*	N07T
	C/H contraction scheme:[b]	(10s4p1d / 4s)	(11s5p2d / 5s1p)	(12s6p3d1f / 6s3p1d)	(16s5p2d / 5s2p)	(17s6p3d1f / 6s3p2d)	(12s7p1d / 7s1p)
Theory level	arrangement	→ [3s2p1d / 2s]	→ [4s3p2d / 3s1p]	→ [5s4p3d1f / 4s3p1d]	→ [4s3p2d / 3s2p]	→ [5s4p3d1f / 3s3p2d]	→ [6s4p1d / 4s1p]
B3LYP-D:	⊥ **3a**	(-3.5) **-2.6**	(-3.3) **-2.7**	(-2.9) **-2.5**	(-3.1) **-2.6**	(-2.8) **-2.7**	(-2.8) **-2.7**
	⊥ **3b**	(-3.2) **-2.3**	(-2.9) **-2.4**	(-2.7) **-2.3**	(-2.9) **-2.3**	(-2.5) **-2.4**	(-2.4) **-2.3**
	⊥ **3c**	(-3.6) **-2.4**	(-3.3) **-2.6**	(-2.9) **-2.5**	(-3.0) **-2.4**	(-2.7) **-2.5**	(-2.6) **-2.5**
	⊥ **3d**	(-3.3) **-1.9**	(-2.8) **-2.2**	(-2.5) **-2.1**	(-2.7) **-2.0**	(-2.5) **-2.4**	(-2.4) **-2.3**
	‖ **3e**	(-1.4) **-0.2**	(-1.6) **-0.9**	(-1.3) **-1.4**	(-1.3) **-1.0**	(-1.7) **-1.5**	(-1.6) **-1.5**
ωB97X-D:	⊥ **3a**	(-3.6) **-2.7**	(-3.6) **-3.1**	(-3.4) **-2.8**	(-3.5) **-3.0**	(-3.0) **-2.9**	(-2.8) **-2.7**
	⊥ **3b**	(-2.9) **-2.0**	(-2.9) **-2.5**	(-2.8) **-2.4**	(-3.0) **-2.4**	(-2.5) **-2.3**	(-2.3) **-2.0**
	⊥ **3c**	(-3.4) **-2.5**	(-3.4) **-2.9**	(-3.2) **-2.7**	(-3.4) **-2.8**	(-2.9) **-2.7**	(-2.7) **-2.3**
	⊥ **3d**	(-3.0) **-2.0**	(-3.1) **-2.6**	(-3.0) **-2.5**	(-3.2) **-2.6**	(-2.6) **-2.4**	(-2.3) **-2.3**
***M06-2X*:**	⊥ **3a**	(-3.5) **-2.4**	(-3.6 **-3.0**	(-3.5) **-3.0**	(-3.6) **-3.0**	(-3.0) **-2.8**	(-2.7) **-2.6**
	⊥ **3b**	(-2.6) **-1.8**	(-2.6) **-2.2**	(-2.6) **-2.2**	(-2.9) **-2.2**	(-2.2) **-2.0**	(-2.0) **-1.7**
	⊥ **3c**	(-3.2) **-2.2**	(-3.3) **-2.6**	(-3.3) **-2.8**	(-3.4) **-2.8**	(-2.8) **-2.6**	(-2.5) **-2.2**
	⊥ **3d**	(-3.5) **-2.3**	(-3.7) **-3.1**	(-3.6) **-3.1**	(-3.8) **-3.1**	(-3.1) **-2.9**	(-2.8) **-2.7**
MP2:	⊥ **3a**	(-2.7) **-1.1**	(-3.6) **-2.2**	(-4.2) **-2.8**	(-5.1) **-2.5**	(-4.0) **-3.0**	(-4.1) **-2.4**
	⊥ **3b**	(-2.3) **-0.9**	(-3.1) **-1.9**	(-3.7) **-2.4**	(-4.8) **-2.1**	(-3.6) **-2.6**	(-3.6) **-2.2**
	⊥ **3c**	(-2.4) **-0.8**	(-3.4) **-2.0**	(-3.9) **-2.6**	(-4.9) **-2.3**	(-3.8) **-2.8**	(-3.8) **-2.2**
	⊥ **3d**	(-2.2) **-0.6**	(-3.2) **-1.8**	(-3.8) **-2.4**	(-4.8) **-2.1**	(-3.6) **-2.6**	(-3.6) **-2.1**
CCSD(T):[c]	⊥ **3a**	(-2.2) **-0.6**	(-2.8) **-1.5**	(-3.5) **-2.1**	(-4.3) **-1.7**	(-3.2) **-2.2**	(-3.5) **-1.8**
	⊥ **3b**	(-1.9) **-0.5**	(-2.4) **-1.2**	(-3.0) **-1.7**	(-4.1) **-1.4**	(-2.9) **-1.9**	(-3.0) **-1.6**
	⊥ **3c**	(-2.0) **-0.4**	(-2.7) **-1.3**	(-3.3) **-2.0**	(-4.3) **-1.6**	(-3.1) **-2.1**	(-3.3) **-1.7**
	⊥ **3d**	(-1.9) **-0.3**	(-2.5) **-1.2**	(-3.2) **-1.7**	(-4.2) **-1.4**	(-2.9) **-2.0**	(-3.1) **-1.5**

[a] CP-corrected energy differences (in bold) relative to the separate moieties; uncorrected values in parenthesis. Units: kcal mol^{-1}.

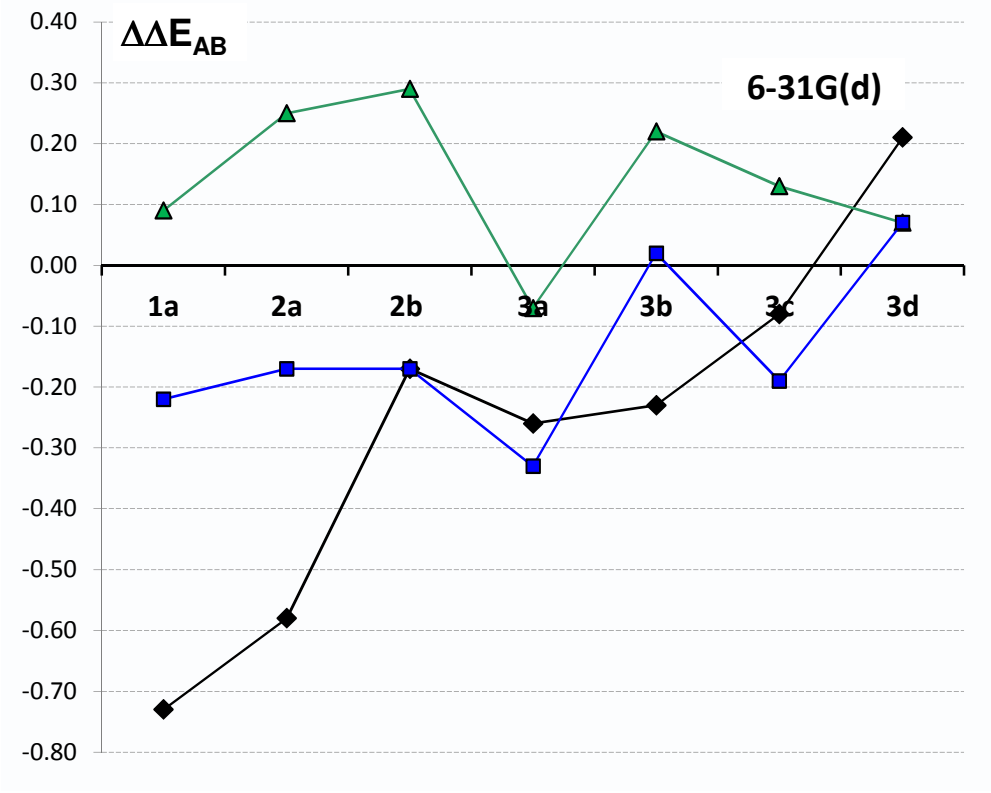
[b] (number of primitive gaussians of each angular momentum) → [number of contracted gaussians]

[c] The CCSD(T)/CBS reference values are **3a**, -2.38; **3b**, -2.04; **3c**, -2.30; **3d**, -2.11 kcal mol^{-1}.

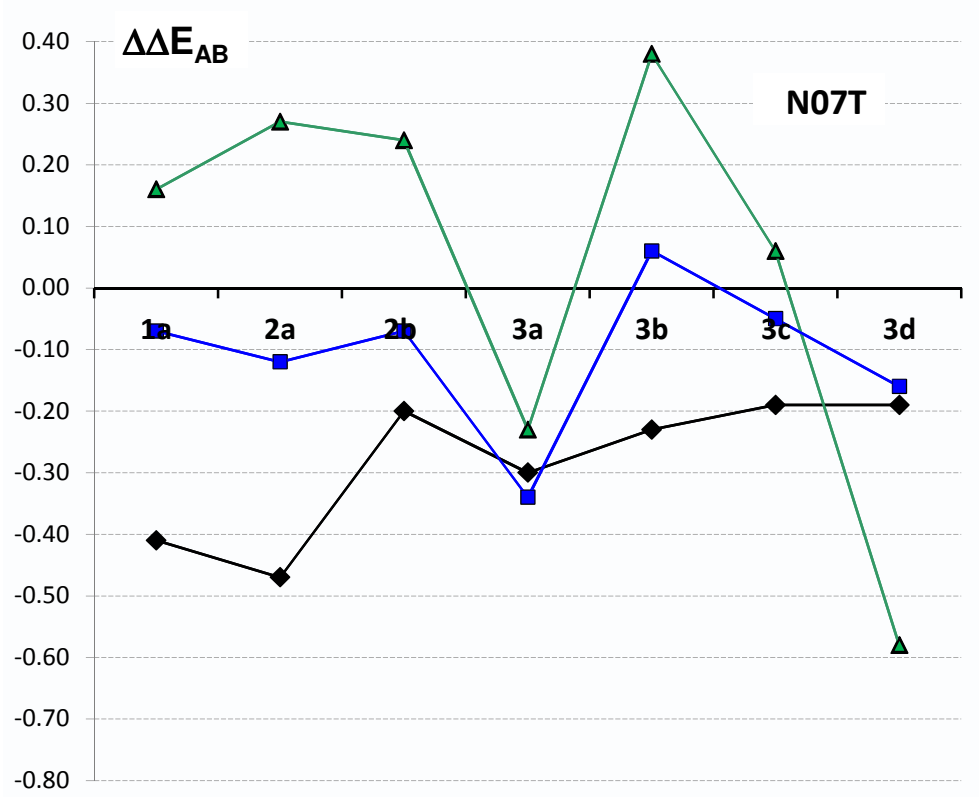
5. Performances of the smaller basis sets.

5.1 D_e comparison with coupled cluster reference.

As stated, our main focus is on the DFT performances with those basis sets, which are most affordable when large systems are studied. Hence we will presently comment on the (CP corrected) performances of the less flexible but also less demanding basis set, 6-31G(d), in **Figure 7a**, and those of the sp-richer but still not too expensive basis, N07T, in **Figure 7b**. These are displayed as deviations ($\Delta\Delta E_{AB}$) from the best computational reference here available, CCSD(T)/CBS. The CCSD(T)/CBS values used to draw the plots are: **1a**, -2.73; **2a**, -1.63; **2b**, -2.15; **3a**, -2.38; **3b**, -2.04; **3c**, -2.30; **3d**, -2.11. The following features can be observed (compare lines of the same color). The variations are less than 1 kcal mol^{-1} for any DFT method. B3LYP-D (black) is mostly overbinding, with one exception (**3d**). With 6-31G(d) B3LYP-D spans the range -0.73 to 0.21 (**3d**); with N07T it is narrower, -0.47 to -0.19 kcal mol^{-1}. The same, but to a lesser degree, can be stated for ωB97X-D (blue), which spans with 6-31G(d) the range -0.33 to 0.07 kcal mol^{-1}, thus exhibiting more moderate variations, and with N07T from -0.34 to 0.06 kcal mol^{-1}). M06-2X (green) shows more often variations in excess, ranging from -0.07 to 0.29 kcal mol^{-1}, with 6-31G(d), and from -0.58 to 0.38 with N07T. If the goal, in dealing with very large systems, is assessing trends rather than the accuracy of the single result, even small basis sets could be acceptable, and the results provided by this approach could be deemed as overall rather encouraging. Though we are aware that we have not explored a wide set of cases, the indication emerging from this study are somewhat in favor, when choosing 6-31G(d), of the M06-2X and ωB97X-D functionals; with N07T, of ωB97X-D.



a



b

Figure 7

Deviation in complex stabilization energy ($\Delta\Delta E_{AB}$ at DFT, in kcal mol^{-1}) from the chosen computational reference, CCSD(T)/CBS (abscissa axis) reported for each complex, as indicated by the labels. (**a**): 6-31G(d) basis set. (**b**): N07T basis set. Black diamonds and lines: B3LYP-D. Green triangles and lines: M06-2X**.** Blue squares and lines: ωB97X-D.

5.2 Comparison with experimental D_0 data.

Benzene-ethyne and benzene-ethene. The experimental results (D_0) available so far (to our knowledge) are reported in **Table 4** and compared with the D_0 estimates for the most stable complex structures of the ethyne and ethene complexes with benzene. These data are displayed in **Figure 8**. Their computed D_0 estimates come from the CP- and ZPE-corrected DFT energy differences, $\Delta[E_{AB}+ZPE]$, provided by the 6-31G(d) and N07T basis sets. The comparison is drawn, for each system, only for the most stable structures. For **1a** (ethyne-benzene) and **2b** (ethene-benzene), B3LYP-D provides estimates which are within the experimental error bars if 6-31G(d) is used (bold figures), while they are less in accord if N07T is chosen. With ωB97X–D, the requirement of falling within the experimental error bars is saytisfied only for **2b**, with N07T (bold figure). When using M06-2X, it is met with both basis sets, again, only for **2b** (bold figure) In **Table 4** we have reported also CCSD(T) ΔE_{AB} data. CCSD(T) had been already validated as a reliable computational reference in different studies (see for instance ref. **25**), and this fact may be particularly beneficial in cases, as the vinyl radical-benzene complex, for which no experimental data are available.

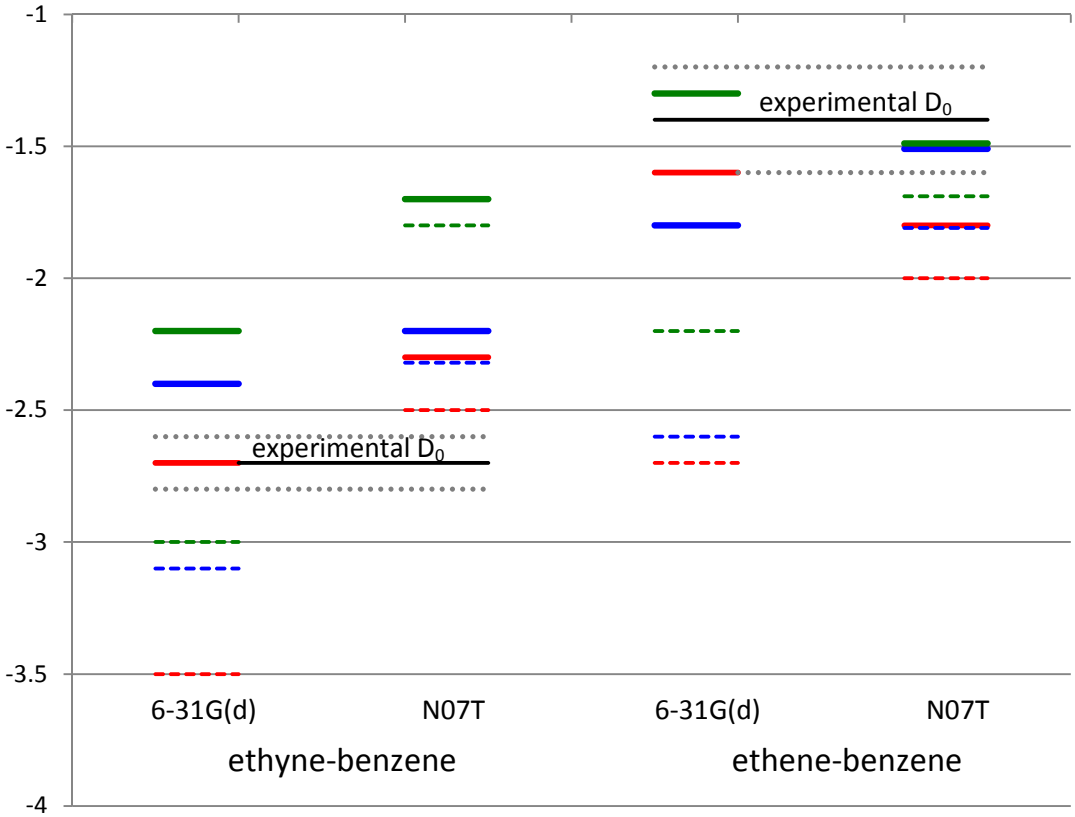


Figure 8

Stabilization energy $\Delta[E_{AB}+ZPE]$, in kcal mol^{-1}, for the the **ethyne-benzene 1a** (left two columns) and **ethene-benzene 2b** (right two columns) complexes. **Black lines**: experimental D_0 data; dashed lines hint to the relevant error bars. Dashed segments: raw ΔE_{AB} results; continuous segments: CP-corrected ΔE_{AB} results. **Red**: B3LYP-D results. **Green**: M06-2X results. **Blue**: ωB97X-D results. For all functionals: first and third columns of segments: 6-31G(d)//6-31G(d); second and fourth columns: N07T//N07T.

It has been sometimes suggested that the CP correction can in some cases prove excessive (though the argument has been debated),[10,61b,83] especially with moderate-size basis sets.[84] Kim has drawn attention toward the procedure of not applying it fully (e.g. by recommending a 50% correction, though at the same time suggesting caution).[85] We have tested this empirical damping of the correction, which we expect to work only if the raw and CP-corrected values happen to take in between the experimental measure in an approximately symmetrical way (see details in the Supplementary Material, Section 7).

The naphthalene dimer. It has been often stated that, if the basis set is not flexible enough, the dispersion forces cannot be represented accurately, and the computation will be presumably biased in favour of perpendicular arrangements, with C–H···π interactions, in which other forces can operate significantly. However, for very small systems, as those just discussed, a parallel arrangement has in any case very little stability, with any basis set. Therefore we have also addressed the case of π–π interactions.[4] To begin with, we have carried out, at the same computational levels, some optimizations for the stacked naphthalene dimer (distorted crossed geometry), in which mainly dispersion forces operate. The relevant data are collected in **Table 5**. In line **A**, our $\Delta[E_{AB}$+ZPE] estimates can be compared with the experimental D_0 value.[86] In particular, B3LYP-D and M06-2X give, with the less extended basis 6-31G(d), values which fall within the experimental error bar (bold figures). In addition, since a computational "benchmark" datum by Janowski and Pulay is available,[5] we have also drawn a comparison (line **B**) between D_e values. Here, is the ωB97X-D/6-31G(d) combination to fall close, with ΔE_{AB} = -6.0 kcal mol^{-1}, to their best estimate, which is -6.1 kcal mol^{-1}. Also other computational estimates can be compared with our results.[87,88,89]

Table 4: Comparison of computed and experimental D_0 interaction energies.[a]

Functional:	*B3LYP-D*[a]		*ωB97X-D*[a]		*M06-2X*[a]		*CCSD(T)*	DFT ZPE corrections[c]			experimental
basis set:	6-31G(d)	N07T	6-31G(d)	N07T	6-31G(d)	N07T	CBS[c]	*B3LYP-D*	*ωB97X-D*	*M06-2X*	**D_0**[b]
								aug-cc-pV*TZ*			
Ethyne-benzene **1a**	**-2.7**	-2.3	-2.4	-2.2	-2.2	-1.7	*-2.7*	*+ 0.59*	*+ 0.47*	*+ 0.69*	**2.7 ±0.1**
Ethene-benzene **2b**	**-1.6**	-1.8	-1.8	**-1.5**	**-1.3**	**-1.5**	*-2.2*	*+ 0.99*	*+ 0.79*	*+ 0.78*	**1.4 ±0.2**
Vinyl-benzene **3a**	-1.6	-1.8	-1.9	-2.1	-1.7	-1.9	*-2.4*	*+ 0.86*	*+ 0.83*	*+ 0.97*	–

[a] $\Delta[E_{AB}$+ZPE] for the most stable structures: they are CP-corrected ΔE_{AB} values (= $-D_e$), to which the ZPE correction is applied . They are relative to the separate moieties, and can be compared to the experimental D_0. Units: kcal mol^{-1}.

[b] D_0 value: see ref. **25**.

[c] ΔE_{AB} (= $-D_e$) values. To allow a loose comparison of coupled cluster results with experimental data for ethyne end ethene, D_0 might be estimated by adding the reported DFT ZPE differences obtained by the three functionals to the CCSD(T)/CBS ΔE_{AB} (figures in italic).

Table 5: Naphthalene dimer. Comparison of interaction energies with: (A) experimental data; (B) benchmark computed data.

	Functional:	*B3LYP-D*		*ωB97X-D*		*M06-2X*		experimental **D_0**[c]
	basis set:	6-31G(d)	N07T	6-31G(d)	N07T	6-31G(d)	N07T	
A	$\Delta[E_{AB}$+ZPE]^b	**-3.8**	-5.6	-5.5	-6.4	**-3.7**	-5.0	**2.9 ±0.9**
B	(ΔE_{AB}) and ΔE_{AB}[a]	(-8.3) –4.2	(-5.9) –5.2	(-8.9) -6.0	(-7.4) -7.0	(-7.3) –4.0	(-5.8) –5.2	Janowski & Pulay(ref.4) -6.1

[a] Line **A**: $\Delta[E_{AB}$+ZPE] are CP-corrected ΔE_{AB} values, to which the ZPE correction is applied, to compare with the experimental D_0 value.

[b] Line **B**: raw ΔE_{AB} values in parenthesis, followed on the right by CP-corrected values, which provide an estimate to $-D_e$, to be compared with the benchmark datum from ref. **4**.

[c] D_0 value: see ref. **86**.

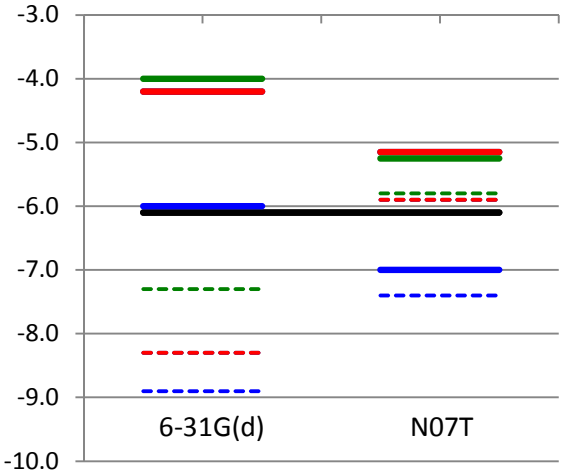


Figure 9

Complex stabilization energy (ΔE_{AB}, in kcal mol^{-1}) for the naphthalene dimer. **Black line**: Janowski and Pulay best estimate (ref. 7), taken as reference. Dashed segments: raw results; continuous segments: CP-corrected results. **Red**: our B3LYP-D results. **Green**: our M06-2X results. **Blue**: our ωB97X-D results. For all functionals: left 6-31G(d)// 6-31G(d); right: N07T//N07T.

We can recall here that the DFT(B3LYP-D) approach was proposed[11] and its performances discussed for the benzene dimer (a classical test case). As regards the estimate of D_0 obtained at this level in conjunction with the N07T basis set, the obtained values, ranging from 2.1 to 2.6 kcal mol^{-1}, compared satisfactorily with the experimental estimate (2.4 ±0.4 kcal mol^{-1}).[86] The benzene dimer was also studied with the M06-2X functional, giving values of 2.8 and 3.7 kcal mol^{-1} (with and without CP, respectively).[42]

We can finally note that, on the basis of the present results, multiple polarization functions of higher angular quantum number seem to be not too effective**,** whilst enhancement of the radial polarization through enrichment of the sp part of the basis set (or its less severe contraction) is helpful in providing a smaller BSSE in the case of π–π interactions (see **Figure 9** and compare also the full set of results reported in the **Tables 1**, **2**, and **3**). This effect is flanked by a lowering of the CP-corrected estimate of ΔE_{AB} (arrows). A greater stabilization does not translate to a more correct estimate, though, as can be seen in the case of the ωB97X-D functional (continuous blue segments).

5.3 A further D_e comparison with a computational benchmark: *The coronene dimer.*

We have also carried out unconstrained optimizations for the coronene dimer (in its parallel displaced arrangement), in which, again, mainly dispersion forces operate. In **Figure 10** we display some calculations, commented hereafter, together with similar results by Zhao and Truhlar at DFT(M06-2X) (red segments).[13] High-quality calculations by Janowski and Pulay[5,6,7] are again available, which allow a comparison with our "inexpensive" results (black line, taken as a reference): their best ΔE_{AB} value, obtained at the QCISD(T)/aDZ level (inter-moiety distance optimized by using DFT(B3LYP)/aug-cc-pV*TZ* monomer geometry), is -20.0 kcal mol^{-1}. We report in the following raw values in parenthesis, followed by CP-corrected values. The functional ωB97X–D (blue segments)with N07T gives (-24.1) -22.4 kcal mol^{-1}, and with 6-31G(d) (-26.7) -20.8, which could be seen as rather close to their value. Rather surprisingly, when using M06-2X (green segments), a worse result is obtained: (-17.8) -15.5 with N07T, and (-19.7) -13.2 kcal mol^{-1} with 6-31G(d).

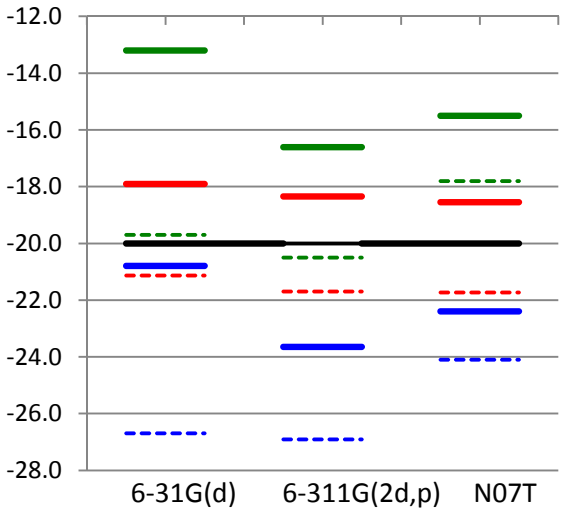


Figure 10

Complex stabilization energy (ΔE_{AB}, in kcal mol^{-1}) for the coronene dimer. **Black line**: Janowski and Pulay best estimate (ref. 7), taken as reference. Dashed segments: raw results; continuous segments: CP-corrected results. **Red**: Zhao and Truhlar at DFT(M06-2X) results (ref. 13, Table 3). **Green**: our M06-2X results. **Blue**: our ωB97X-D results. For all functionals: left 6-31G(d)//6-31G(d); center: 6-311G(2d,p)//6-31G(d); right: N07T//N07T.

To probe the basis set dependence of the two sets of results, we have also carried out single point 6-311G(2d,p) energy evaluations at the 6-31G(d) optimum geometries (center, blue and green segments).[90] The arrows emphasize that a significant basis set dependence of the results is found for this larger system (the sequence is [3s2p1d / 2s] → [4s3p2d / 3s1p] → [6s4p1d / 4s1p] as regards the contracted gaussians). By changing in a parallel fashion, one of the functionals (M06-2X) improves its performance, while the other (ωB97X–D) gets worse: M06-2X/6-311G(2d,p)//6-31G(d) gives (-20.47) -16.61 and ωB97X–D/6-311G(2d,p)//6-31G(d) gives (-20.47) -16.61 kcal mol^{-1}.

5.4 The association of two different PAHs.

We can futher test the performances of the most "inexpensive" computational levels [M06-2X/6-31G(d) and ωB97X-D/6-31G(d)] by studying the π–π association of a first smaller PAH, coronene, made by 24 carbon atoms and labeled R24, to another larger PAH, circumcircumpyrene, made by 80 carbon atoms (P80), large enough to allow a bordered projection of R24 onto it. The R24:P80 association is sketched below (**Figure 11**). If we aim to get in the complex a projection of the first PAH onto the second one well within its perimeter (bordered projection), an ample enough bordering is suggested by the tendency of two PAHs to slide to some extent one onto the other with respect to a fully eclipsed arrangement. This slipping (accompanied possibly by some reorientation, see for instance ref. **88**) can take place in different manners, and we report here the most stable situation. The CP- and ZPE-corrected energy change associated to cluster formation, $\Delta E_{R:P}$, divided by the number of carbons contained in R, can provide a value "per C atom" that can be considered typical of R, $\Delta E_C(R)$. It is probably the most important contribution to the interaction, since it is related to those parts of the two molecules that are directly superimposed. This presumably major component of interaction is seen as "perpendicular" to the (approximate) molecular planes. Bordering can be expected to produce oblique interactions and a further contribution to the interaction energy. In fact, though probably smaller, they induce some curvature in P80 (in correspondence of which, a deformation energy in P80 is estimated to be only +0.7 kcal mol^{-1}). In **Figure 11**, perpendicular interactions in coronene:circumcircumpyrene are symbolically indicated by a few double-headed arrows. Within this setting, $\Delta E_C(R)$ values are approximately defined by dividing the interaction energy by the number of carbons contained in the smaller PAH. Similarly, $\Delta H_{R:P}$ and $\Delta H_C(R24)$ (enthalpies) can be assessed.

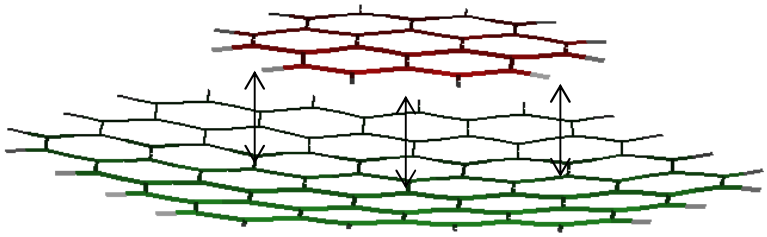


Figure 11

The R24:P80 complex

By the same approach, two smaller systems have been studied too, to see some trend in the energy and enthalpy computed values. These systems are R10:P80 (naphthalene:circum-circumpyrene) and R16:P80 (pyrene:circumcircumpyrene).

By proceeding along this line, we find (full set of data in the **Supplementary material**, **Section 5**) when using M06-2X values in the range ΔE_C = -0.94 to -0.97 (ΔH_C = -0.90 to -0.94) with full CP correction and taking into account the ZPE.

One can than compare, with due prudence, the estimates (by changing the sign) with the exfoliation energy per carbon atom (ΔE_C^{exf}) typical of graphite.[91] ΔE_C^{exf} was estimated experimentally by Zacharia, Ulbricht, and Hertel as 1.20(±0.12) kcal mol^{-1}.[92] Hence, M06-2X/6-31G(d) approaches this 1.08-1.32 range from below, while ωB97X-D estimates seem to be too large.

However, if we extend the comparison to Janowski and Pulay's best computational estimate for the interaction energy between coronene and circumcoronene (-32.1 kcal mol^{-1}), we find that their estimate for coronene on a graphite surface was -37.4 kcal mol^{-1}, which translates to a value estimated by those Authors higher than the experimental one, i.e. -1.56 kcal mol^{-1} per C atom.[6] ωB97X-D would be put under a more favourable light, in this case.

IV. CONCLUSIONS

The purpose of our exploration admittedly stands in some contrast with the usual recommendation that a large basis set with multiple polarization functions is necessary to get accurate results for vdW association energies. But we have to consider that, in dealing with large systems, rather than aspiring to get a significant accuracy on the single datum, we will rather aim to define dependable trends.

Therefore, we have examined the performances of three DFT functionals (B3LYP-D, ωB97X-D, and M06-2X) in association with a variety of basis sets, in providing estimates of the CP-corrected complexation energy, ΔE_{AB} (= $-D_e$) for fully optimized complexes. The D_e estimates have been compared with CCSD(T)/CBS values (extrapolation to the "complete basis set", CBS) results, taken as our computational reference (see **Figure 7**). We could define such a reference only for small complexes, namely those of ethyne, ethene, or the vinyl radical with benzene. All DFT functionals show a modest basis set dependence of the ΔE_{AB} estimates, while the coupled cluster data are, by contrast, more basis set dependent, though they appear to converge towards a limit. The same can be said for the MP2 results; however MP2 tends, upon basis set extension, to overbind (see **Figures 2**, **4**, **6**).

We have been interested especially to the performance of DFT functionals with basis sets of limited extension, as 6-31G(d) and N07T. They produce, for the title complexes, ΔE_{AB} estimates which span a range δ around the CCSD(T)/CBS reference of somewhat less than 1 kcal mol^{-1}, at most. In particular, B3LYP-D is mostly overbinding with both basis sets, with δ = 0.94 [6-31G(d)], δ = 0.28 [N07T]; M06-2X is mostly overbinding with 6-31G(d), with δ = 0.36, while with N07T δ = 0.96; ωB97X-D is mostly overbinding with both basis sets, with δ = 0.40 for both (see again **Figure 7**).

Then, we focused on the performances of the three functionals, with the 6-31G(d) and N07T basis sets only, comparing our computed D_0 estimates for the most stable structures (obtained by adding the zero-point vibrational correction to the CP-corrected ΔE_{AB} values) with available experimental D_0 values (**Table 4**, first two lines). The results for the complexes benzene-ethyne and benzene-ethene are rather sparse, in the sense that, depending on the system, one or the other computational level performs best. The computed D_0 values fall within the experimental error bar for the experimental D_0 in 5 cases out of 12.

A similar test was then carried out on the naphthalene dimer, in an approximately C_2 π–π arrangement (**Table 5**). Here we found our best estimates at B3LYP-D and M06-2X, both with

6-31G(d): at those levels, our computational naphthalene dimer D_0 result falls within the experimental error bar. Also a high-quality computation was available in this case, and the ωB97X-D/6-31G(d) combination is the one which falls the closest. Similarly, we could compare a high-quality calculation for the (π–π) coronene dimer, and again is the ωB97X-D/6-31G(d) combination which gives the best result, very close to the benchmark.

As a final test, we studied three different π–π associations between two PAHs of different size at M06-2X/6-31G(d) and ωB97X-D/6-31G(d), with the purpose of drawing a comparison with the exfoliation energy per C atom (ΔE_C^{exf}) determined experimentally for graphite. These are R10:P80 (made by naphthalene interacting with circumcircumpyrene), R16:P80 (pyrene-circumcircumpyrene), and R24:P80 (coronene-circumcircumpyrene). The R:P interaction energies (sign inverted), divided by the number of C atoms in the smaller component, are expected to approximate in some way ΔE_C^{exf}. M06-2X/6-31G(d) approaches the experimental 1.08-1.32 kcal mol^{-1} per C atom range from below, whilst ωB97X-D estimates seem to be too large.

The results collected in this study encourage to deal with large systems at DFT, further exploring the use of either the B3LYP-D, the ωB97X-D, or the M06-2X functionals, in conjunction with a modestly-sized basis set. From the far from complete comparison reported above, DFT(M06-2X)/6-31G(d) or DFT(ωB97X-D)/6-31G(d) might emerge as possible functional/basis set combinations of choice to assess trends within large systems.

Acknowledgements. This work was conducted in the frame of EC FP6 NoE ACCENT+ (Atmospheric Composition Change, the European NeTwork of Excellence). Funding through the project PRIN 2009 is acknowledged (Programmi di ricerca scientifica di Rilevante Interesse Nazionale - D.M. 19 marzo 2010 n. 51 - prot. 2009SLKFEX_005).

Supplementary Material file. This material includes: a Table with the data used to estimate the CCSD(T)/CBS energy differences (Section 1); tests on the position of the minimum-energy intermoiety distance in the complexes when CP correction is taken into account (Section 2); a Table with intermoiety distances at different combinations of DFT functional and basis set (Section 3); spin densities associated to atoms, and total spin expectation values, for the structures involving the vinyl radical and benzene (Section 4); CP-corrected energy and enthalpy differences for the association of two different PAHs (Section 5); the geometries and energies of all optimized structures (Section 6).

REFERENCES AND NOTES

[1] W. F. Cooke, J. J. Wilson, J. Geophys. Res. **101**, 19395 (1996). C. Liousse, J. E. Penner, C. Chuang, J. J. Walton, H. Eddleman, H. Cachier, J. Geophys. Res. **101**, 19411 (1996). B. J. Finlayson-Pitts, J. N. Jr. Pitts, *Chemistry of the Upper and Lower Atmosphere* (Academic Press: 2000) Chapter 10, pp 453-460.

[2] K.-H. Homann, Angew. Chem. Int. Ed. **37**, 2435 (1998).

[3] Apart from adsorption onto the surface of a carbonaceous particle, one could consider stacking between PAH units. There are indications that, at combustion temperatures, stacking interactions alone between modest size PAH units cannot start the nucleation of a carbon nanoparticle [H. Sabbah, L. Biennier, S. J. Klippenstein, I. R. Sims, B. R. Rowe J. Phys. Chem. Lett. **1**, 2962 (2010)], and that this might occur only if larger PAHs can be present in concentrations significantly higher than currently known [T. S. Totton, A. J. Misquitta, M. Kraft Phys. Chem. Chem. Phys. **14**, 4081 (2012)]. However, stacking could also accompany a previous binding of PAH-like units through σ bond formations [A. Giordana, A. Maranzana, G. Tonachini J. Phys. Chem. C **115**, 1732 (2011). A. Giordana, A. Maranzana, G. Tonachini J. Phys. Chem. C **115**, 17237 (2011)].

[4] We can mention that the π interactions are of interest also because they have been utilized for various types of molecular assembly [see for instance: J. Y. Lee, B. H. Hong, W. Y. Kim, S. K. Min, Y. Kim, M. V. Jouravlev, R. Bose, K. S. Kim, I.-C. Hwang, L. J. Kaufman, C. W. Wong, P. Kim, K. S. Kim, Nature 460, 498-501 (2009)] and functionalizations of graphene for nanodevices (see for instance: S. K. Min, W. Y. Kim, Y. Cho, K. S. Kim Nature Nanotech. 6, 162-165 (2011); V. Georgakilas, M. Otyepka, A. B. Bourlinos, V. Chandra, N. Kim, K. C. Kemp, P. Hobza, R. Zboril, K. S. Kim Chem. Rev. 112, 6156 (2012); E. C. Lee, D. Kim, P. Jurečka, P. Tarakeshwar, P. Hobza, K. S. Kim, J. Phys. Chem. A**, 111***,* 3446 (2007).

[5] T. Janowski, P. Pulay, J. Am. Chem. Soc., **134**, 17520 (2012).

[6] T. Janowski, P. Pulay, Theor. Chem. Acc., **130**, 419 (2011).

[7] T. Janowski , A. R. Ford, P. Pulay, Mol. Phys., **108**, 249 (2010).

[8] For studies dealing with the benzene dimer, see for instance: P. C. Jha, Z. Rinkevicius, H. Ågren, P. Seal, S. Chakrabarti, Phys. Chem. Chem. Phys. **10**, 2715 (2008). R. A. DiStasio, Jr., G. von Helden, R. P. Steele, M. Head-Gordon, Chem. Phys. Lett. **437**, 277 (2007). T. Janowski, P. Pulay, Chem. Phys. Lett. **447,** 27 (2007). R. Podeszwa, R. Bukowski, K. Szalewicz, J. Phys. Chem. A **110**, 10345 (2006). J. Grant Hill, J. A. Platts, H.-J. Werner, Phys. Chem. Chem. Phys.

8, 4072 (2006). N. K. Lee, S. Park, K. S. Kim J. Chem. Phys. **116**, 7910 (2002).

[9] M. O. Sinnokrot, C. D. Sherrill, J. Phys. Chem. A **110**, 10656-10668 (2006).

[10] C. D. Sherrill, T. Takatani, E. G. Hohenstein, J. Phys. Chem. A **113**, 10146 (2009).

[11] M. Pavone, N. Rega, N.; V. Barone, Chem. Phys. Lett. **452**, 333 (2008). M. Biczysko, G.Piani, M. Pasquini,N. Schiccheri, G. Pietraperzia, M. Becucci, M. Pavone, V. Barone J. Chem. Phys. **127**, 144303 (2007). M. Pasquini, N. Schiccheri, G. Piani, G. Pietraperzia, M. Becucci, M. Biczysko, M. Pavone, V. Barone J. Phys. Chem. A **111**, 12363 (2007).

[12] R. Podeszwa, K. Szalewicz, Phys. Chem. Chem. Phys. **10**, 2735 (2008).

[13] Y. Zhao, D. G. Truhlar, J. Phys. Chem. C **112,** 4061 (2008).

[14] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, J. Chem. Phys. **120**, 647 (2004).

[15] L. F. Molnar, X. He, B. Wang, K. M. Merz, J. Chem. Phys. **131**, 065102 (2009).

[16] E. Carrasquillo, T. S. Zwier, D. H. Levy, J. Chem. Phys. **83**, 4990 (1985). In this experimental study, two different isomers of the C_6H_6-C_2H_2 association were identified, of which one is described similar to our **1a**.

[17] M. Y. Shelley, H.-L. Dai, T. Troxler, J. Chem. Phys. **111**, 9081 (1999).

[18] R. K. Sampson, S. M. Bellm, J. R. Gaskooke, W. D. Lawrance, Chem. Phys. Lett. **372**, 307 (2003).

[19] A. Fujii, S. Morita, M. Miyazaki, T. Ebata, N. Mikami, J. Phys. Chem. A **108**, 2652 (2004).

[20] S. Tsuzuki, K.Honda, T. Uchimaru, M. Mikami, K. Tanabe, J. Am. Chem. Soc. **122**, 3746 (2000). The computed interaction energies were: -2.83 and -2.06 kcal mol^{-1}, for the ethyne-benzene and ethene-benzene complexes, respectively. Dispersion is indicated as the main source of attraction, while the electrostatic contribution, -2.01 and -0.65 kcal mol^{-1}, respectively, has also a role.

[21] J. J. Novoa, F. Mota, Chem. Phys. Lett. **318**, 318, 345 (2000). Here, the ethyne-benzene and ethene-benzene complexes CP-corrected association energies at MP2/6-31+G(2d,p) are: 2.59 and 0.80 kcal mol^{-1}, respectively. These values cannot be compared directly with our MP2 estimates, but are close to those reported in Tables 1 and 2 for the 6-311G(2d,p) basis set.

[22] K. Sundararajan, K. S. Viswanathan, A. D. Kulkarni, S. R. Gadre, J. Mol. Struct. **613**, 209 (2002). In this investigation, two structures were found computationally for the C_6H_6-C_2H_2

complex, one, 1A, in which ethyne points with one hydrogen toward the center of benzene, as in our **1a**, the other, 1B, exchanges the roles, and has two benzene hydrogens pointing toward the midpoint of ethyne's triple bond. Experimentally, only 1A was observed.

[23] C. Ramos, P. R. Winter, J. A. Stearns, T. S. Zwier, J. Phys. Chem. A **107**, 10280 (2003).

[24] A. Fujii, S. Morita, M. Miyazaki, T. Ebata, N. Mikami, J. Phys. Chem. A **108** , 2652 (2004).

[25] K. Shibasaki, A. Fujii, N. Mikami, S. Tsuzuki, J. Phys. Chem. A **111**, 753 (2007). Here, MP2/cc-pV*TZ* optimized geometries were used for the estimation of the $E_{CCSD(T)}$(limit). The error of $E_{CCSD(T)}$(limit) associated with an inaccuracy of the optimized geometries was judged very small (probably <0.04 kcal mol^{-1}). See also note 87 in ref. **80**.

[26] N. J. Singh, S. K. Min, D. Y. Kim, K. S. Kim, J. Chem. Theory Comput. **5**, 515 (2009).

[27] K. S. Kim, S. Karthikeyan, N. J. Singh, J. Chem. Theory Comput. **7**, 3471 (2011).

[28] E. C. Lee, B. H. Hong, J. Y. Lee, J. C. Kim, D. Kim, Y. Kim, P. Tarakeshwar, K. S. Kim, J. Am. Chem. Soc., **127**, 4530 (2005).

[29] Y. Cho, Seung Kyu Min, J. Yun, W. Y. Kim, A. Tkatchenko, K. S. Kim, J. Chem. Theory Comput. **9**, 2090 (2013).

[30] K. Müller-Dethlefs, P. Hobza, Chem. Rev. **100**, 143 (2000).

[31] P. Hobza, Phys. Chem. Chem. Phys. **10**, 2581 (2008).

[32] A. Michaelides, J. Chem. Phys. **137**, 120901 (2012).

[33] K. E. Riley, P. Hobza, Acc. Chem. Res., **46**, 927 (2013).

[34] F. Coester, H. Kümmel, Nucl. Phys. **17**, 477 (1960). J. Cízek, J. Chem. Phys. **45**, 650 (1966). J. Paldus, J. Cízek, I. Shavitt, Phys. Rev. A **5**, 50 (1972). For a review, see: T. D. Crawford, H. F. Schaefer III, *An Introduction to Coupled Cluster Theory for Computational Chemists , in Reviews in Computational Chemistry* (Wiley-VCH, 2000) vol. 14, pp 33-136.

[35] C. Møller, M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[36] R. G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989) Chapter 3; pp 47-69.

[37] We can forcedly mention here only a selection of papers, each followed by the indication of the methods used. P. Hobza, H.-L. Selzle, E. W. Schlag, J. Am. Chem. Soc. **116**, 3500 (1994), HF and MP2. E. J. Meijer, M. Sprik, J. Chem. Phys. **105**, 8684 (1996), DFT. S. Tsuzuki, T. Uchimaru, M. Mikami, K. Tanabe, J. Phys. Chem. A **102**, 2091 (1998), MP2 and CCSD(T).

S. Tsuzuki, T. Uchimaru, K. Matsumura, M. Mikami, K. Tanabe, Chem. Phys. Lett. **319**, 547-554 (2000), MP2 and CCSD(T). H. Ruuska, T. A. Pakkanen, J. Phys. Chem. B **105**, 9541 (2001), HF and MP2. M. O. Sinnokrot, E. F. Valeev, C. D. Sherrill, J. Am. Chem. Soc. **124**, 10887 (2002), MP2 and CCSD(T). S. D. Chakarova, E. Schröder, Mat. Sci. Eng. **C25**, 787 (2005), DFT. J. D. Kubicki, Env. Sci. Technol. **40**, 2298 (2006), MM+MP2.

[38] S. Kristyán, P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).

[39] S. Tsuzuki, T. Uchimaru, K. Tanabe, Chem. Phys. Lett. **287**, 202 (1998).

[40] E. R. Johnson, R. A. Wolkow, G. A. DiLabio, Chem. Phys. Lett. **394**, 334 (2004) .

[41] S. Tsuzuki, H. P. Lüthi, J. Chem. Phys. **114**, 3949 (2001).

[42] Noteworthy are in particular the new functionals put forth by Zhao and Truhlar (based on the improvement of the exchange-correlation functional). Their performances are discussed in ref. **10** and in: Y. Zhao, D. G. Truhlar, J. Chem. Theory Comput. **2**, 1009 (2006). Y. Zhao, D. G. Truhlar, J. Chem. Theory Comput. **3**, 289 (2007). Y. Zhao, D. G. Truhlar, Phys. Chem. Chem. Phys. **10**, 2813 (2008). E. G. Hohenstein, S. T. Chill, C. D. Sherrill, J. Chem. Theory Comput. **4**, 1996 (2008). Y. Zhao, D. G. Truhlar, J. Chem. Phys. **125**, 194101 (2006). Y. Zhao, D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).

[43] S. A. Kafafi, J. Phys. Chem. A **102**, 10404 (1998). F. Tran, J. Weber, T. A. Wesołowski, Helv. Chim. Acta **84**, 1489 (2001). X. Xu, W. A. Goddard, III, W. A. PNAS **101**, 2673 (2004). M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004). O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, D. Sebastiani, Rev. Lett. **93**, 153004 (2004). O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani Phys. Rev. B **71**, 195119 (2005). A. J. Misquitta, R. Podeszwa, B. Jeziorski, K. Szalewicz, J. Chem. Phys. **123** , 214103 (2005). S. D. Chakarova, E. Schröder, J. Chem. Phys. **122**, 054102 (2005). A. Heßelmann, G. Jansen, M. Schütz J. Chem. Phys. **122**, 014103 (2005). P. Jurečka, J. Černý, P. Hobza, D. R. Salahub, J. Comput. Chem. **28**, 28, 555 (2006). O. I. Obolensky, V. V. Semenikhina, A. V. Solov'yov, W. Greiner, Int. J. Quantum Chem. **107**, 1335(2007). E. Tapavicza, I.-C. Lin, O. A. von Lilienfeld, I. Tavernelli, M. D. Coutinho-Neto, U. Rothlisberger, J. Chem. Theory Comp. **3**, 1673 (2007). L. Zhechkov, T. Heine, S. Patchkovskii, G. Seifert, H. A. Duarte, J. Chem. Theory Comput. **1**, 841 (2005).

[44] S. Grimme, J. Comp. Chem. **25**, 1463 (2004). S. Grimme, J. Comp. Chem. **27**, 1787 (2006).

[45] M. E. Foster, K. Sohlberg, Phys. Chem. Chem. Phys. **10**, 2813 (2008).

[46] S. D. Chakarova-Käck, A. Vojvodic, J. Kleis, P. H., E. Schröder New Journal of Physics **12**, 013017 (2010).

[47] See for instance: J. Ez, J. Fanfrlk, D. Salahub, P. Hobza, J. Chem. Theory Comput. **5**, 1749 (2009).

[48] M. E. Foster, K. Sohlberg, Phys. Chem. Chem. Phys. **12**, 307 (2010).

[49] A. Pople, P. M. W. Gill, B. G. Johnson, Chem. Phys. Lett. **199**, 557 (1992). H. B. Schlegel, Computational Theoretical Organic Chemistry (Csizsmadia, I. G., Daudel, R., Eds.; Reidel Publ. Co.: 1981) pp.129-159. H. B. Schlegel, J. Chem. Phys. **77**, 3676 (1982). H. B. Schlegel, J. S. Binkley, J. A. Pople, J. Chem. Phys. **80**, 1976 (1984). H. B. Schlegel, J. Comput. Chem. **3**, 214 (1982).

[50] V. Barone, M. Biczysko, M. Pavone Chem. Phys. **346**, 247 (2008).

[51] B3: A. D. Becke, Phys. Rev. A **38**, 3098 (1998). A. D. Becke, ACS Symp. Ser. **394**, 165 (1989). J. A. Pople, P. M. W. Gill, B. G. Johnson, Chem. Phys. Lett. **199**, 557 (1992). A. D. Becke, J. Chem. Phys. **98**, 5648 (1993). LYP: C. Lee, W. Yang, R. G. Parr Phys. Rev. B **37**, 785 (1988).

[52] D. Chai, M. Head-Gordon, J. Chem. Phys. **128**, 084106 (2008).

[53] D. C. Langreth, M. Dion, H. Rydberg, E. Schröder, P. Hyldgaard, B. I. Lundqvist, Int. J. Quantum Chem., **101**, 599 (2005).

[54] J. Antony, S. Grimme, Phys. Chem. Chem. Phys. **8**, 5287 (2006).

[55] F. Jensen, In *Introduction to Computational Chemistry* (John Wiley & Sons, 1999) Chapter 5. The keywords "scf=tight int=ultrafine opt=tight" were used in the GAUSSIAN optimizations. However, no significant change were observed in a test run on the vinyl radical-benzene complex **3e**. The ΔE results are reported below. Compare refs. **38** and **40** for discussions on the role of the integration scheme in connection with BSSE and PES definition.

basis set:	6-31G(d)	6-311G (2d,p)	6-311++G(3df,3pd)	aug-cc-pv*DZ*
integration grid				
standard	-1.43	-1.59	-1.72	-1.34
"ultrafine"	-1.44	-1.59	-1.69	-1.34

[56] 6-31G: W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys. **56**, 2257 (1972). 6-31G(d): P. C. Hariharan, J. A. Pople, Theor. Chim. Acta **28**, 213 (1973). 6-311G(d,p): K. Raghavachari, J. S. Binkley, R. Seeger, J. A. Pople, Chem. Phys. **72**, 650 (1980). Multiple polarization functions: M. J. Frisch, J. A. Pople, J. S. Binkley, J. Chem. Phys. **80**, 3265 (1984).

[57] D. E. Woon, T. H. Jr. Dunning, J. Chem. Phys. **98**, 1358 (1993). T. H. Jr. Dunning, J. Chem. Phys. **90**, 1007 (1989). K. A. Peterson, D. E. Woon, T. H. Jr. Dunning, J. Chem. Phys. **100**, 7410 (1994).

[58] Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.

[59] In some cases, a divergent behavior in Møller–Plesset perturbative series has been reported: J. Olsen, O. Christiansen, H. Koch, P. Jørgensen, J. Chem. Phys. **105**, 5082 (1996).

[60] It can be commented that though the relatively "inexpensive" MP2 has been often said to provide a rather satisfactory assessment of binding in vdW complexes, it has also been indicated, as inclined to overbinding [P. Hobza, H. L. Selzle, E. W. Schlag, J. Phys. Chem. **100**, 18790-18794 (1996), and ref. **41**]. By contrast, the CCSD(T) implementation of coupled cluster theory gives accurate binding energies, but is computationally very demanding. Finally, there are cases where MP2 and CCSD(T) interaction energies have been reported to show a significant basis set dependence (refs. **25** and **41**; for CCSD(T), see, as an example, Figures 2, 4, 6 in the present study).

[61] (a) S. F. Boys, F. Bernardi, Mol. Phys. **19**, 553 (1970).
(b) S. Simon, M. Duran, J. J. Dannenberg, J. Chem. Phys. **105**, 11024 (1996).
Recently, a new counterpoise method, different from the one applied in this work, has been put forward: F. Jensen, J. Chem. Theory Comput. **6**, 100 (2010).

[62] Gaussian 09, Revision A.2. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

[63] MOLDEN: G. Schaftenaar, J. H. Noordik, J. Comput.-Aided Mol. Design **14**, 123 (2000) (http://www.cmbi.ru.nl/molden/molden.html).

[64] M. Frenklach, H. Wang, Proc. Combust. Inst. **23**, 1559 (1991).

[65] H. Bockhorn, F. Fetting, H. W. Wenz, Ber. Bunsen-Ges. Phys. Chem. **97**, 1067 (1983).

[66] M. Frenklach, D. W. Clary, W. C. Gardiner, S. E. Stein, Proc. Combust. Inst. **20**, 887 (1984).

[67] J. D. Bittner, J. B. Howard, Symp. Int. Combust. Inst.**18**, 1105 (1981).

[68] C. W. Jr. Bauschlicher, A. Ricca, Chem. Phys. Lett. **326**, 283 (2000). H. Richter, O. A. Mazyar, R. Sumathi, W. H. Green, J. B. Howard, J. W. Bozzelli, J. Phys. Chem. A **105**, 1561 (2001). B. V. Unterreiner, M. Sierka, R. Ahlrichs, Phys. Chem. Chem. Phys. **20046**, 4377 (2004). V. V. Kislov, N. I. Islamova, A. M. Kolker, S. H. Lin, A. M. Mebel, J. Chem. Theory. Comput. **1**, 908 (2005).

[69] We have explored the growth mechanism in a study connected to the present one, for species adsorbed on a soot platelet, represented by a PAH-like model: A. Indarto, A. Giordana, G. Ghigo, A. Maranzana, G. Tonachini, Phys. Chem. Chem. Phys. **12**, 9429 (2010).

[70] D. Roy, M. Marianski, N. T. Maitra, J. J. Dannenberg, J. Chem. Phys. **137**, 134109 (2012).

[71] M. Marianski, A. Asensio, J. J. Dannenberg, J. Chem. Phys. **137**, 044109 (2012). These Authors find, when studying α-helical and antiparallel β-sheet polyalanines by functionals designed to treat dispersion (B97-D, ωB97x-D, M06, M06L, and M06-2X), that they significantly overesti-mate interaction enthalpies of folding for the α-helix, thus predicting unreasonable structures. Though their performance was better when the basis set is extended up to aug-cc-pVTZ , they comment that such computations would hardly be feasible for systems as large as the peptides.

[72] M. Marianski, X. Oliva, J. J. Dannenberg, J. Phys.Chem. A, **116**, 8100 (2012). Here, when studying the interaction between pyrimidine and p-benzoquinone, MP2 calculations with large basis sets were found to possibly overestimate dispersion (also observed here, see Figures 2, 3, and 4). It was in agreement with published results on the π-stacked benzene dimer, though does not appear to be true for all dispersion stabilized complexes. The Authors offered as a possible reason the fact that the interactions in these structures are not purely due to dispersion but to a combination of induction and dispersion.

[73] G. Pohl, J. A. Plumley, J. J. Dannenberg, J. Chem. Phys. **138**, 245102 (2013);

[74] The electrostatic contribution depends on the electronegativity χ of carbon, which depends in turn on its hybridation. Since C(sp) in ethyne has more s character than C(sp^2) in ethene, this empirical quantity is consequently estimated to be larger for carbon in the former: χ=

3.29 for the sp hybridation, and χ= 2.75 for the sp^2 (for sp^3, χ= 2.5). Thus, the interaction involving one ethyne C–H bond more closely resembles a hydrogen bond. For a complete discussion, with analysis of the different (dispersion, electrostatic,…) contributions to the stability of the complex, the reader is referred to the papers by Tsuzuki and Fujii (refs. **25** and **80**).The χ values parallel the estimated pK_a values for alkynes and alkenes which provide a measure of the stability of the corresponding carbanion (ethyne: pK_a =25, ethene pK_a =44; from M. B. Smith, J. March In *March's Advanced Organic Chemistry* (Wiley, 2001) Chapter 8; Table 8.1.

[75] Carbon hybridation reflects on different properties: chemical shifts (see, C. S. Foote, B. L. Iverson, E. V. Anslyne *Organic Chemistry* - ch. 13); C-H bond strengths (ethane: $D_0 \approx 99$, ethene: $D_0 \approx 105$; ethyne: $D_0 \approx 111$ kcal mol^{-1}); approximate acidities of C-H bonds (ethane: $pK_a \approx 50$, ethene: $pK_a \approx 44$; ethyne: $pK_a \approx 25$ kcal mol^{-1}), due to the different s character of the hybrid orbital to which the carbanion lone pair is associated. M. B. Smith, J. March *March's Advanced Organic Chemistry. Reactions, Mechanisms, and Structure*, 6th edition, (John Wiley & Sons 2007) Chapter 8, p 388.

[76] Y. Kodama, K. Nishihata, M. Nishio, Y. Iitaka J. Chem. Soc. Perkin **2**, 1490 (1976).

[77] A. Tekin, G. Jansen, Phys. Chem. Chem. Phys., 2007, 9, 1680.

[78] This structure corresponds to a first order saddle point on the energy hypersurface, and the surface itself is exceedingly flat in the direction of the corresponding transition vector. The relevant vibration (a swaying motion, hinted to by the curled arrows in Figure 1) clearly connects this **||** transition structure (TS) with two equivalent ⊥ energy minima, which differ by the ethyne hydrogen involved in the interaction with the benzene ring.

[79] M. Ōki ; S. Takano, S. Toyota, Bull. Chem. Soc. Jpn. **73**, 2221-2230 (2000).

[80] S. Tsuzuki, A. Fujii, Phys. Chem. Chem. Phys. **10**, 2584 (2008) and references therein.

[81] Our MP2/6-31G(d) or 6-311G(2d,p) data can be compared with those reported in Table 7 of ref. **79**, namely -0.4 kcal mol^{-1} for structure G and -0.8 for F-1, both CP-corrected.

[82] The ⊥ minima **3a** and **3b** bear some resemblance with the geometry J of the benzene-ethene complex displayed in Figure 4 of ref. **20a** (rather than with I).

[83] F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de RijdtJ, and J. H. van Lenthe, Chem. Rev. 94, 1873 (1994).

[84] K. R. Liedl J. Chem. Phys. 108, 3199 (1998). T. H. Dunning J. Phys. Chem. A 104, 9062 (2000).

[85] K. S. Kim, P. Tarakeshwar, and J. Y. Lee, Chem. Rev. 100, 4145-4185 (2000).

[86] T. Fujiwara, E. C. Lim, J. Phys. Chem. A **107**, 4381 (2003).

[87] C. Gonzalez, E. C. Lim, J. Phys. Chem. A **107**, 10105 (2003). Here the BSSE-uncorrected interaction energy of the crossed (C_2) dimer is reported (Table 2) as 18.6 kcal mol^{-1}, at the MP2/6-31G(d) optimized geometry, with d exponent=0.25.

[88] M. Rubes, O. Bludsky, P. Nachtigall, Chem Phys Chem **9**, 1702 (2008). Here a whole series of parallel displaced arrangements is examined (see Figure 1). The closest to our computation is PD-1 (C_2), whose interaction energy is estimated -6.23 kcal/mol^{-1}.

[89] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikamia J. Chem. Phys. **120**, 647 (2004), in which two almost isoenergetic structures are found with the lowest energy, the slipped-parallel (C_i) structure -5.73 kcal/mol^{-1} and the cross (D_{2d}) structure -5.28 kcal/mol^{-1}.

[90] The significant difference in ΔE_{AB} values does not appear to originate from differences of the geometries obtained with the two functionals. A single-point energy evaluation test was carried out in which the M06-2X energies were obtained for the optimum ωB97X–D geometry, and vice versa. The CP-corrected ΔE_{AB} values did not vary significantly. M06-2X/6-31G(d) //ωB97X-D/6-31G(d): ΔE_{AB} = -13.3 kcal mol^{-1} (while, optimized at M06-2X/6-31G(d), is -13.2) and ωB97X-D/6-31G(d) //M06-2X/6-31G(d): ΔE_{AB} = -20.0 kcal mol^{-1} (optimized at ωB97X-D/6-31G(d): -20.8). Hence differences originate from the functional used to calculate the energies.

[91] See for instance their Figure 4 in: J. D. Herdman, J. H. Miller J. Phys. Chem. A **112**, 6249 (2008). It displays the binding energy per carbon atom for a series of homo-molecular dimer complexes.

[92] R. Zacharia, H. Ulbricht, T. Hertel, Phys Rev B **69**, 155406 (2004).