MNRAS **000**, 1–4 (2019) Preprint 9 March 2019 Compiled using MNRAS LATEX style file v3.0

Supplementary Material for Identification of Stable Configurations in the Superhydrogenation Sequence of Polycyclic Aromatic Hydrocarbon Molecules

Pernille A. Jensen,[1] Mirko Leccese,[2] Frederik D. S. Simonsen,[1] Anders W. Skov,[1] Matteo Bonfanti,[2] John D. Thrower,[1] Rocco Martinazzo[2]★ and Liv Hornekær[1,3]†
[1]*Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark*
[2]*Department of Chemistry, UniversitÃą degli Studi di Milano, Milan, Italy*
[3]*Interdisciplinatry Nano-Science Centre (iNano), Aarhus University, Denmark*

Accepted XXX. Received YYY; in original form ZZZ

1 CHOICE OF THE DENSITY FUNCTIONAL

Despite the continuous theoretical progress, the choice of the functional most appropriate for an application remains a key step to obtain accurate results from Density Functional Theory (DFT) electronic-structure calculations. For the problem considered in the main text, hydrogenation of a PAH, we decided to use the so-called M06-2X functional because of the indications we gained from a preparatory investigation of the system energetics, that is described in some detail in the following.

We considered the adsorption of the first H atom on coronene, testing several exchange and correlation (xc) functionals against the few data available for this process, in particular the results of accurate coupled-cluster (CCSD(T)) calculations extrapolated to the complete basis-set (CBS) limit (Wang et al. 2012). However, since the system lacks of extensive data, further analysis was necessary to assess the quality of the functionals and make our choice robust. To this end, we considered the thermochemistry of selected literature databases that seemed most appropriate to describe relevant aspects of the PAHs chemistry, and used them to compare the performance of the functionals.

We employed the reasonably accurate Pople's 6-31+G(d,p) atom-centered basis set and the xc-functionals listed in Table 1, along with B3LYP that represents a rather popular but often unsatisfactory choice. The functionals of Table 1 are considered to be among the best available nowadays and show outstanding perfomances for the application fields suggested by the developers (Thrular and coworkers) and given in the table ((Zhao & Truhlar 2006b; Peverati & Truhlar 2011)). They are all of the *meta-hybrid* type and thus include the spin-labeled kinetic energy density in the set of fundamental variables, and a variable fraction of exact Hartree-Fock exchange (X%), in some cases dependent on the interelectronic separation (range-separated functionals).

Benchmarking of the functionals is usually performed w.r.t. several thermochemical and kinetic databases that gather key chemical-physical parameters of selected chemical processes (obtained either from experiments or high-quality electronic structure calculations). However, the results are typically application-dependent and thus a careful choice of the databases that fit best to one's problem seems to be the best strategy for choosing the appropriate functional. For our problem we considered the following databases:

Table 1. Functional form (M=meta, HM = hybrid-meta , MHF = meta (full) Hartree-Fock, RSHM = range-separated hybrid, CT=charge-transfer,TDDFT=time-dependent DFT), Hartree-Fock exchange percentage (X%) and application area of the main Minnesota functionals (in paranthesis the year of development). (Zhao & Thrular 2007; Peverati & Truhlar 2011)

Functional	Type	X%	Suggested use
MPW1B95 (2004)	HM	31	Covalent and non-covalent thermochemistry, hydrogen bonding, weak interactions, CT
MPWB1K (2004)	HM	44	Thermochemical kinetics, hydrogen bonding, weak interactions, CT
M06-L (2006)	M	0	Main-group thermochemistry, kinetics, non-covalent interactions
M06 (2008)	HM	26	Main-group thermochemistry, kinetics, metallochemical interactions non-covalent interactions
M06-2X (2008)	HM	52	Main-group thermochemistry, kinetics, non-covalent interactions
M06-HF (2006)	MHF	100	Long-range CT via TDDFT spectroscopic properties non-covalent interactions
M11 (2012)	RSHM	42.8-100	Replaces M06-2X, M06-HF, M06 best performance for CT via TDDFT

πTC13 is a π-system database (Zhao & Truhlar 2006a), comprising various reference values

i) proton affinities of conjugated polyenes
ii) proton affinities of conjugated Schiff bases
iii) energy separations between cumulenes and polyines
iv) torsional potentials of butadiene and styrene
v) bond length alternation in butadiene and octatetraene

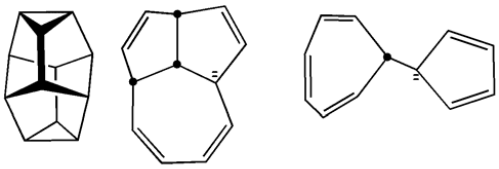

Figure 1. Some hydrocarbons considered in HC7.

PAH5 is a database of PAH isomerization energies (Karton 2016), using CCSD(T) results with a CBS extrapolation as a reference. The isomerization reactions taken into account are the following:

i) phenanthrene → anthracene
ii) triphenylene → chrysene
iii) triphenylene → benzo[a]anthracene
iv) triphenylene → benzo[c]phenanthrene
v) triphenylene →naphthacene

HC7 is a database of hydrocarbon data that are sensitive to medium-range correlations energies (Zhao & Thrular 2007), namely

i) isomerization of $(CH)_{12}$ into the three isomers of Fig. 1
ii) $(CH_3)_3CC(CH_3)_3 \rightarrow$ n-C_8H_{18}
iii) n-C_6H_{14} + 4 $CH_4 \rightarrow 5C_2H_6$
iv) n-C_8H_{18} + 6 $CH_4 \rightarrow 7C_2H_6$
v) adamantane + $3C_2H_4 \rightarrow 2C_2H_2$
vi) bicyclo[2.2.2]octane → $3C_2H_4 + C_2H_2$

S22A is a database representing non-covalent interactions for some small to relatively large (30 atoms) complexes of common molecules with C,N,O and H, and single, double and triple bonds (Takatani et al. 2010). Here, the reference values are CCSD(T)/CBS results. The database is divided into three subsets

i) hydrogen bonded complexes: $(NH_3)_2$; $(H_2O)_2$; Formic acid dimer; Formamide dimer; Uracil dimer (C_{2h}); 2-pyridoxine·2-aminopyridine; Adenine·Tymine
ii) complexes with predominant dispersion stabilization: $(CH_4)_2$; $(C_2H_4)_2$; Benzene·CH_4; Benzene dimer; Pyrazine dimer; Uracil dimer (C_2); Indole·Benzene; Adenine·Tymine stack
iii) mixed complexes in which electrostatic and dispersion contributions are similar in magnitude: Ethene·ethine; Benzene·H_2O; Benzene·NH_3; Benzene·HCN; Benzene dimer; Indole·Benzene T-shape; Phenol dimer

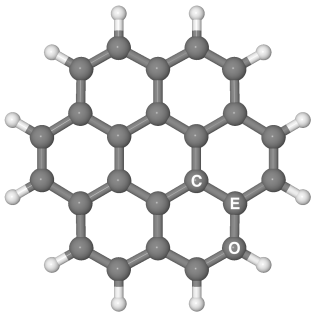

Figure 2. The three different sites on coronene molecule.

The performance of the functionals of Table 1 on the above datasets are given in Table 2 as mean unsigned error (MUE, in kcal/mol) (Zhao & Truhlar 2006a; Karton 2016; Zhao & Thrular 2007; Takatani et al. 2010), alongside the results of our benchmark calculations. For the latter we considered adsorption at the three physically distinct sites of the coronene molecule, namely the outer edge (o), the inner edge (e) and the central carbon (c) indicated schematically in Fig. 2. We computed the corresponding binding and the barrier energies, using *quasi* second-order optimization algorithms (without constraints) and frequency calculation checks, as described in the main text.

A quick look at the table reveals that, with the exception of M06-L, all the functionals considered underestimate the reaction barrier of the CCSD(T) reference and generally (now excluding M06-HF) overestimate the binding energies. The trends among the functionals are similar irrespective of the adsorption site, and indicate a general improvement of the results when using *meta* xc-functionals and a reasonable amount of HF exchange. This conclusion is supported by the MUE values on the above mentioned reference datasets, that show a considerable improvement when going from B3LYP to a M* functional. Particularly impressive is the performance of M06-2X which shows rather small MUE values in "critical" databases. In particular, SS2A is a databases usually considered for dispersion forces, and M06-2X behaves rather well for these applications. That is, even though the functional (and the others in the table) was not developed with the dispersive forces in mind, the use of the kinetic energy density and the ensuing more extensive parametrization made it adequate to describe the non-covalent interactions, at least at binding geometries.

Overall, our results clearly indicate that M06-2X is the most appropriate functional for investigating coronene hydrogenation.

2 CHOICE OF ADSORPTION SITES

The choice of the appropriate sites to check for hydrogenation, at each hydrogenation step, is one further issue that deserves some comments. Indeed, a brute-force approach that considers all possible adsorption sites is yet out of reach nowadays even for a small-sized molecule like coronene with its only 24 hydrogenation sites (disregarding the possible difference between adsorption on one face or the other). The total number of sites N is several milions, as a simple calculation shows. For, if the sites were distinguishable (*i.e.* in the absence of symmetry) this number would be given by $N = \sum_{n=0}^{24} \binom{24}{n} \equiv 2^{24}$, a number that should be divided by a symmetry factor of the order $\sim 6 - 12$ to account for the hexago-

Table 2. Binding energies, D (eV) and sticking barriers, E_s (meV), for the hydrogen addition to three different sites of coronene, MUE (Mean Unsigned Error) (kcal/mol) for three different database, πTC13 (π-system thermochemistry), SS2A (Non covalent binding energies), HC7/11 (Medium-large correlation energies for hydrocarbons). All the values reported (this work and literature data) are not CP-corrected. "I" and "II" stands refer to two different basis set, respectively cc-pVDZ and cc-pVTZ.

					MUE			
		X%	D	E_s	πTC13	SS2A	HC7/11	PAH5
Outer Edge	B3LYP	20	1.46	73	6.06	3.77	16.8	6.2
	MPW1B95	31	1.44	137				
	MPWB1K	44	1.49	154				
	M06	26	1.53	168	4.08	1.06	2.78	2.2
	M06-2X	52	1.35	187	1.51	0.40	2.15	2.1
	M06-HF	100	1.24	209	2.29	0.62	2.29	5.4
	M06-L	0	1.43	219	6.52	0.80	3.35	3.1
	M11	42.8-100	1.54	135	2.12	0.44	3.74	8.3
	PW91[a]		1.45	60				
Center	B3LYP	20	0.66	262	6.06	3.77	16.8	6.2
	MPW1B95	31	0.63	302				
	MPWB1K	44	0.65	321				
	M06	26	0.74	321	4.08	1.06	2.78	2.2
	M06-2X	52	0.57	328	1.51	0.40	2.15	2.1
	M06-HF	100	0.51	305	2.29	0.62	2.29	5.4
	M06-L	0	0.62	407	6.52	0.80	3.35	3.1
	M11	42.8-100	0.73	281	2.12	0.44	3.74	8.3
	PW91[a]		~ 0.725	~ 180				
	B3LYP[b]	20	0.66	250				
	ROMP2/II [c]		0.29	400				
	ROCCSD(T)/I [c]		0.58	370				
Edge	B3LYP	20	0.71	279	6.06	3.77	16.8	6.2
	MPW1B95	31	0.67	342				
	MPWB1K	44	0.70	363				
	PW91[a]		~ 0.725	~ 180				

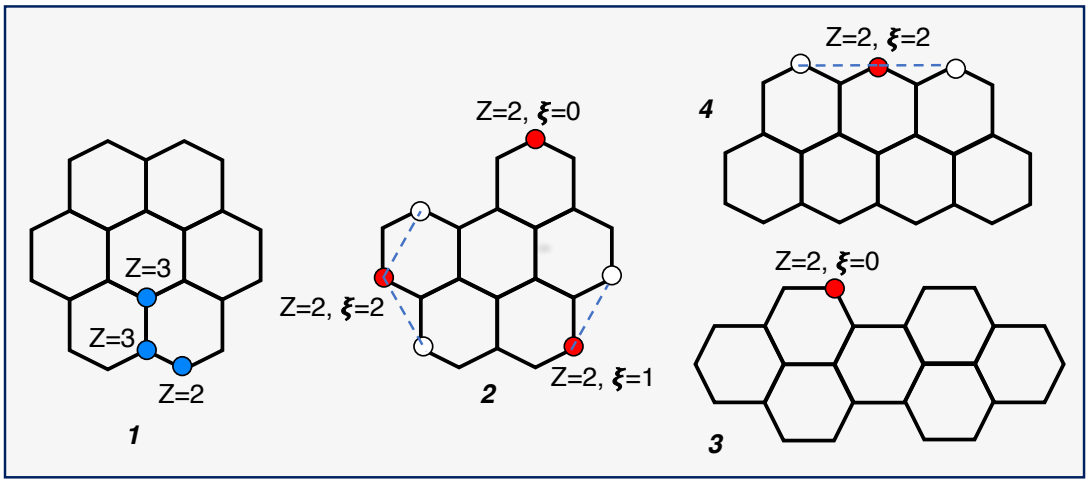

Figure 3. Exemplificatory structures showing the concepts of π coordination and π hyperconjugation. **1** The coordination numbers Z of the first adsorption sites in coronene. **2** Edge sites ($Z = 2$) with different hyperconjugation number ξ as indicated, with their second (edge) neighbors highlighted. **3** and **4** show the prototypical cases of armchair ($\xi = 0$) and zig-zag ($\xi = 2$) edges.

nal symmetry of the substrate[1]. Hence, one needs some "guiding principles" to identify at the outset, for a given structure $C_{24}H_{12+n}$, where the next H atom would preferably absorb. Fortunately, this is possible for typical C sp^2 structures since, at a crude level, they form bipartite systems with special symmetries in their electronic energy spectra (Bonfanti et al. 2011). Briefly, as mentioned in the main text, one distinguishes two cases, depending on whether the systems presents or not an unpaired electron. In the latter case (which

[1] To be precise, one should count the number of equivalent classes of configurations, *i.e.* only those configurations that cannot be superimposed by applying a symmetry operation of the pristine substrate. Thus, for instance, the symmetry number would be 8 for monomers, since out of the 24 possible configurations only 3 are distinct.

arises for $C_{24}H_{12+n}$ with n even) the main factors driving the addition of a H atom are 'π-coordination' and 'π-hyperconjugation' (in decreasing order of importance). The former determines the localization of frontier orbitals at the edges of the π electronic system, enhancing the reactivity of those sp^2 sites that have the smallest number of sp^2 neighbors. This is a rather reasonable rule that states that the π undercoordinated sites are the most reactive ones, as if they presented some sort of dangling bonds to be saturated[2]. In this respect, true edge sites (E sites) are all characterized by $Z = 2$, while inner sites ($Z = 3$) can be conveniently distinguished in graphitic sites (G) (with three neighbors with $Z = 3$) and inner-edge sites (with only two or one three-coordinated neighbors, F_2 and F_1 sites, respectively).

Hyperconjugation, on the other hand, discriminates between sites with the same π-coordination but different number of next-to-nearest neighbors in the π lattice, as exemplified by the structures shown in Fig. 3. Hydrogen affinity increases with this number and makes zig-zag edges markedly more reactive than armchair ones. In using these rules it is worth noticing that they apply to the π sites, and thus the concepts of "coordination number" refers to the number of C sp^2 neighbors and *not* to the total coordination number.

On the other hand, when the system presents an unpaired electron (a case which arises for $C_{24}H_{12+n}$ when n is odd) the reactivity is baised towards the lattice sites where such an unpaired electron mostly resides since, obviously, spin pairing of the extra H atom (*i.e.*, bond formation) is easier at those positions. It has long been realized that in bipartite lattices the "unpaired" electron density localizes on the majority sublattice, a result that can also be obtained from (chemical) resonance theory (Bonfanti et al. 2018). Such density, however, is unevenly distributed on such sublattice, since it localizes close to the "defect" that generated it (in the so-called *ortho* and *para* positions) and possibly shaped by the presence of edges (*i.e.*, with a preference towards low coordination and high hyperconjugation sites of the majority sublattice).

The simple rules described above have been used to generate stepwise the guess structures with extra hydrogen atoms. Fig. 5, for instance, displays the complete set of hydrogenation sites that have been considered during the investigation of the main hydrogenation pathway (checking, when meaningful, adsorption on either sides of the molecule). Each hydrogenated level (nH) is pictorially represented as a molecular graph where black circles denote carbons that have already been hydrogenated in previous steps, and blue circles mark those positions that have been inspected at that specific step.

Inspection of Fig. 5 shows that the rules above prove to be rather useful in predicting the most reactive sites, at each hydrogenation level. On the contrary, we found that a reactivity analysis based on the condensed (radical) Fukui indices (as obtained from a Hirschfeld partitioning scheme of the electron density) behaves rather poorly and fails, from the 4th step onward, in identifying the site with the largest binding energy. The good correlation between binding energies and coordination / hyperconjugation numbers is shown in Fig. 3, where the main results have been re-organized according to these numbers. The figure clearly shows the increase of the binding energy when decreasing Z and, for $Z = 2$, when increasing ξ (the correlation is even better if one limits the comparison to the energies for binding H to the different sites of the *same* structure, which is precisely the situation one addresses during the hydrogenation process).

[2] The argument can be given a firmer ground by exploiting the bipartite nature of the system and squaring the Hamiltonian (Bonfanti et al. 2011).

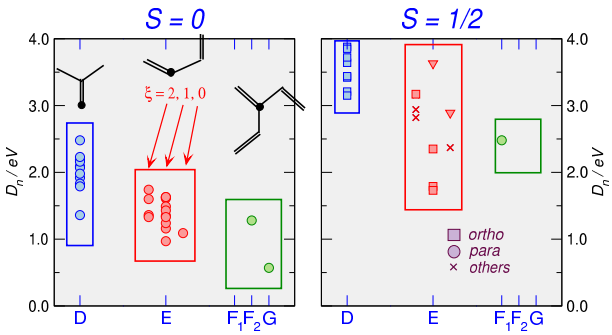

Figure 4. Correlation between hydrogen binding energy and coordination numbers. Left: H sticking on even numbered species (*i.e.*, singlets), for different sites, $Z = 1, 2$ and 3 for D, E and F/G sites. Also indicated the hyperconjugation numbers. See text for details. Right: same as in the left panel for odd-numbered species (*i.e.*, doublets).

REFERENCES

Bonfanti M., Casolo S., Ponti A., Tantardini G., Martinazzo R., 2011, J. Chem. Phys., 135, 164701
Bonfanti M., Achilli S., Martinazzo R., 2018, J. Phys. Condens. Matter, 30, 283002
Karton A., 2016, J. Comp. Chem.
Peverati R., Truhlar D. G., 2011, Phys. Chem Lett., 2, 2810
Takatani T., Hohenstein E. G., Malagoli M., Marshall M. S., Sherril C. D., 2010, J. Chem. Phys., 8
Wang Y., Qian H., Morokuma K., Irle S., 2012, J. Phys. Chem. A, 116, 7154
Zhao Y., Thrular D. G., 2007, Theor. Chem. Account, 120, 215
Zhao Y., Truhlar D. G., 2006a, Org. Lett., 8, 5753âĂŞ5755
Zhao Y., Truhlar D., 2006b, J. Phys. Chem. A, 110, 13126âĂŞ13130

This paper has been typeset from a TEX/LATEX file prepared by the author.

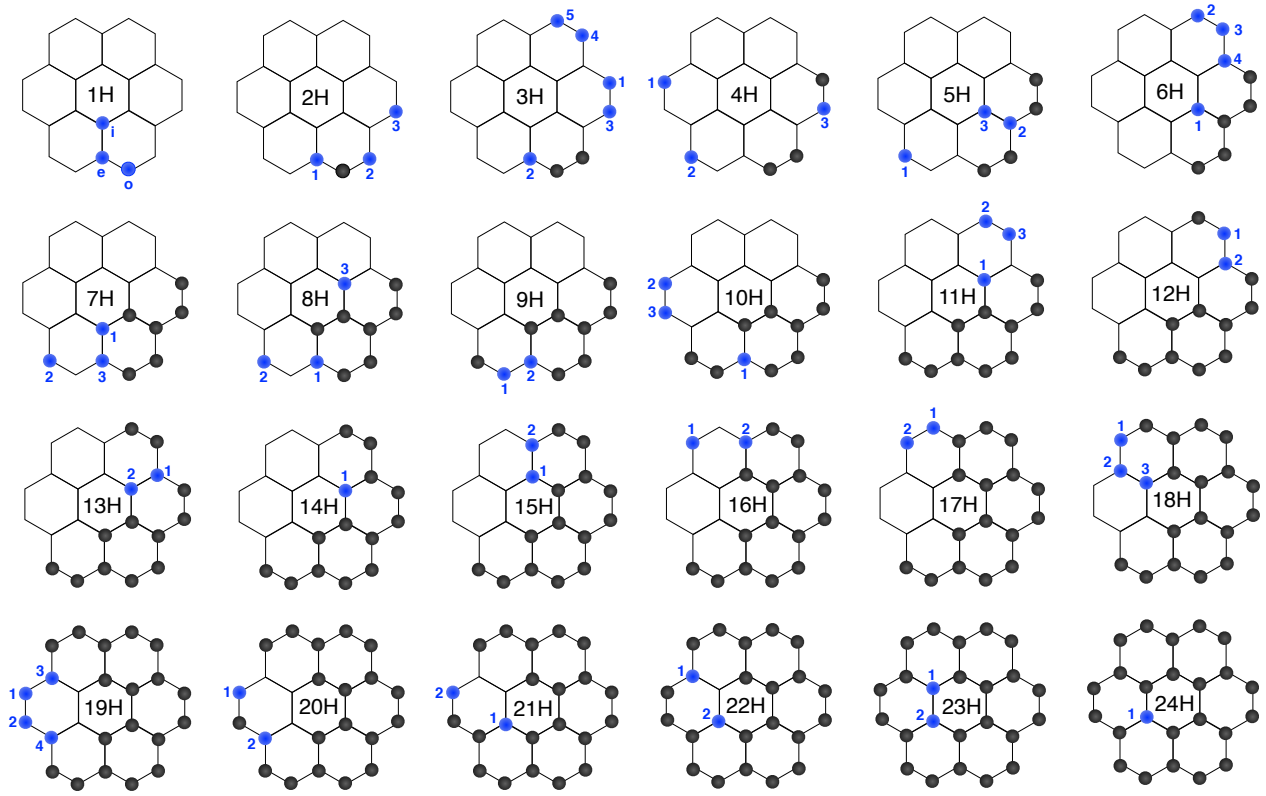

Figure 5. Black circles mark hydrogenated sp^3 carbons while blue circles mark sites whose binding energy has been computed at each hydrogenation level (nH).