

Interaction energies in non-covalently bound intermolecular complexes derived using the subsystem formulation of density functional theory

Marcin Dułak · Tomasz A. Wesolowski

Received: 1 November 2006 / Accepted: 1 February 2007 / Published online: 13 March 2007
© Springer-Verlag 2007

Abstract Interaction energies for a representative sample of 39 intermolecular complexes are calculated using two computational approaches based on the subsystem formulation of density functional theory introduced by Cortona (*Phys. Rev. B* **44**:8454, 1991), adopted for studies of intermolecular complexes (Wesolowski and Weber in *Chem. Phys. Lett.* **248**:71, 1996). The energy components (exchange-correlation and non-additive kinetic) expressed as explicit density functionals are approximated by means of gradient-free- (local density approximation) of gradient-dependent- (generalized gradient approximation) approximations. The sample of the considered intermolecular complexes was used previously by Zhao and Truhlar to compare the interaction energies derived using various methods based on the Kohn-Sham equations with high-level quantum chemistry results considered as the reference. It stretches from rare gas dimers up to strong hydrogen bonds. Our results indicate that the subsystem-based methods provide an interesting alternative to that based on the Kohn-Sham equations. Local density approximation, which is the simplest approximation for the relevant density functionals and which does not rely on any empirical data, leads to a computational approach comparing favorably with more than twenty methods based on the Kohn-Sham equations including the ones, which use extensively empirical parameterizations. For various types of non-bonding interactions, the strengths and weaknesses of gradient-free and gradient-dependent approximations to exchange-correlation and non-additive kinetic energy density functionals are discussed in detail.

Keywords Non-bonding interactions · Stacking · Hydrogen bonding · Density functional theory · Subsystems

Introduction

The difficulties of the Kohn-Sham-based methods in describing the energetics of weakly bound intermolecular complexes are well documented. Semi-local approximations to the exchange-correlation energy functional, in which the density of the exchange-correlation energy is expressed as an analytic function of such local quantities as: electron density (LDA) [1], electron density and its gradients (GGA), or electron density, its gradients, and higher derivatives and/or density of the kinetic energy (metaGGA), are incapable to describe the correct asymptotics of the potential energy surface at the limit of non-overlapping electron densities of the molecules forming the complex. Even close to the equilibrium geometry, where the electron densities of interacting molecules do overlap to some extent, LDA leads to a systematic overestimation of the interaction energy and GGA to erratic results depending strongly on the form of the gradient-dependency in the approximate exchange-correlation functional [2].

In the original applications of the subsystem formulation of density functional theory (DFT) to ionic solids, the subsystems corresponded to atoms, and the LDA functionals were used together with additional approximations on the symmetry and localization of orbitals for each subsystem [3]. In our adaptation of this formalism to molecular systems, LDA and GGA functionals can be used and no restrictions are made on symmetry or localization of orbitals in each subsystem [4]. Therefore, this implementation is ideally suited for testing the overall accuracy of the used approximate density functionals against adequate reference data. At the LDA and

M. Dułak · T. A. Wesolowski (✉)
Département de Chimie Physique, Université de Genève,
30, quai Ernest-Ansermet,
CH-1211 Genève 4, Switzerland
e-mail: Tomasz.Wesolowski@chiphys.unige.ch

GGA levels, the subsystem formulation of DFT leads usually to better interaction energies than their Kohn-Sham counterparts, as indicated in our previously reported studies concerning hydrogen-bonded complexes [5], complexes formed by non-polar molecules [6] or complexes involving aromatic rings [2, 7, 8].

This work aims at systematization of the picture emerging from our previous studies. To this end, we apply two versions of the formalism (LDA and GGA) to evaluate interaction energies in a set of 39 representative intermolecular complexes at their equilibrium geometries. For the chosen set, accurate reference interaction energies derived from wavefunction-based methods are available. Zhao and Truhlar used these data in their systematic evaluation of performance of various approximations to the exchange-correlation energy functional applied within the Kohn-Sham framework [9–11]. Therefore, using the same systems makes it possible to compare directly the accuracy of the two considered computational methods based on the subsystem formulation of DFT with that of Kohn-Sham-based methods. It is also worthwhile to note that the strength of intermolecular interactions varies in a wide range (up to -28.80 kcal/mol) and includes complexes more strongly bound than complexes for which the subsystem formulation of DFT was applied so far.

The complexes in the test set are divided into six groups [10, 11]:

- *hydrogen-bonded* (HB6/04) $\text{NH}_3\text{-NH}_3$, HF-HF , $\text{H}_2\text{O-H}_2\text{O}$, $\text{NH}_3\text{-H}_2\text{O}$, $\text{HCONH}_2\text{-HCONH}_2$, and HCOOH-HCOOH ,
- *charge transfer type* (CT7/04): $\text{C}_2\text{H}_4\text{-F}_2$, $\text{NH}_3\text{-F}_2$, $\text{C}_2\text{H}_2\text{-ClF}$, HCN-ClF , $\text{NH}_3\text{-Cl}_2$, $\text{H}_2\text{O-ClF}$, and $\text{NH}_3\text{-ClF}$,
- *dominated by dipolar interactions* (DI6/04): $\text{H}_2\text{S-H}_2\text{S}$, HCl-HCl , $\text{H}_2\text{S-HCl}$, $\text{CH}_3\text{Cl-HCl}$, $\text{HCN-CH}_3\text{SH}$, and $\text{CH}_3\text{SH-HCl}$,
- *Weakly bonded* (WI7/05): He-Ne , He-Ar , Ne-Ne , Ne-Ar , $\text{CH}_4\text{-Ne}$, $\text{C}_6\text{H}_6\text{-Ne}$, and $\text{CH}_4\text{-CH}_4$,
- *π -stacked* (PPS5/05): $(\text{C}_2\text{H}_2)_2$, $(\text{C}_2\text{H}_4)_2$, sandwich (S $(\text{C}_6\text{H}_6)_2$), T-shaped (T $(\text{C}_6\text{H}_6)_2$), and parallel-displaced (PD $(\text{C}_6\text{H}_6)_2$),
- *nucleobase pairs* (BP8/05): adenine-thymine (A-T), guanine-cytosine (G-C), antiparallel cytosine dimer (anti C-C), displaced cytosine dimer (displ C-C), parallel cytosine dimer (par C-C), uracil dimer (U-U), Watson-Crick adenine-thymine (WC A-T), and Watson-Crick guanine-cytosine (WC G-C).

The quantity of primary interest for this study is the interaction energy which is calculated at exactly the same geometries as the ones used in the analyses by Zhao and Truhlar [9–11]. This type of calculations assure that the numerical differences between our results and those in the

compared database can be attributed to three factors: *i*) the used basis sets, *ii*) numerical procedures, and *iii*) the approximations to the relevant density functionals. The first two factors are fully controllable in our implementation of the formalism. This work focuses, therefore, on the issue of the performance of LDA and GGA functionals in the subsystem formulation of DFT.

Computational details

Practical applications of the subsystem formulation of DFT involve solving two sets of one-electron equations for embedded orbitals referred to here as Kohn-Sham equations with constrained electron density (KSCED), where each set of embedded orbitals is used to construct the electron density of the corresponding subsystem. In these equations, two types of quantities used also in the Kohn-Sham formulation of DFT are approximated by means of explicit functionals of electron density: the exchange-correlation functional ($E_{xc}[\rho]$) and the non-additive kinetic energy bi-functional ($T_s^{nad}[\rho_A, \rho_B] = T_s[\rho_A + \rho_B] - T_s[\rho_A] - T_s[\rho_B]$, where $T_s[\rho]$ denotes the kinetic energy in the reference system of non-interacting electrons). In this work, LDA and GGA density functionals are considered. We will use the labels KSCED LDA and KSCED GGA for the corresponding computational methods. In the KSCED LDA calculations, the exchange functional is approximated using the expression for the uniform gas of non-interacting electrons by Dirac [12], the correlation energy is approximated using the Vosko et al. [13] parametrization (equation [4.4] in Ref. [13] referred frequently as “VWN V”) of the Ceperley-Alder [14] reference data for correlation energy in the uniform electron gas, and the non-additive kinetic energy is approximated using the Thomas-Fermi formula for the kinetic energy [15, 16]. Note that the above approximate functionals do not rely on any empirical data. In the KSCED GGA calculations, the Perdew-Wang (PW91) [17, 18] exchange-correlation functional is used whereas the non-additive kinetic energy bi-functional $T_s^{nad}[\rho_A, \rho_B]$ is approximated according to the formula: $T_s^{nad}[\rho_A, \rho_B] \approx T_s^{nad(GGA97)}[\rho_A, \rho_B] = T_s^{LC94}[\rho_A + \rho_B] - T_s^{LC94}[\rho_A] - T_s^{LC94}[\rho_B]$, where $T_s^{LC94}[\rho]$ denotes the Lembarki-Chermette [19] functional of the kinetic energy. The $T_s^{nad(GGA97)}[\rho_A, \rho_B]$ was shown to provide a good approximation to the non-additive kinetic energy functional and potential in the case of weakly overlapping densities [20].

The calculations are carried out using our numerical implementation of the subsystem formulation of DFT (the program deMon2K-KSCED [21]) based on the program deMon2K [22]. In all calculations, the following program options and parameters are applied: 10^{-6} a.u. self-consistent field (SCF) energy convergence criterion, pruned “MEDI-

UM” (75,302)p [23] grid. The aug-cc-pVTZ [24, 25] and the MG3S [26] basis sets were used. Two types of basis set expansions are considered for orbitals corresponding to each subsystem: centered on the monomer or centered on the dimer. The corresponding calculations are labeled as KSCED(m) and KSCED(s), respectively, following the convention of Ref. [27]. Classical electron-electron interactions (Coulomb) are evaluated using auxiliary fitting functions referred to as GEN-A2*, which are automatically generated for any given orbital basis sets [28]. Further details concerning the formal framework of the applied computational methods and the numerical implementation can be found in Refs. [3–5].

The interaction energies discussed in the comparisons with literature data were corrected for the effects associated with the basis set superposition error (BSSE) and the grid superposition error (GSE). The origin and treatment of these two types of errors is similar. In the KSCED(s) case, BSSE originates from the fact that for each subsystem, the basis sets includes also atomic orbitals belonging to the other subsystem. Such orbitals are missing in the Kohn-Sham calculations to derive the total energy of the isolated subsystem. The Boys-Bernardi counterpoise technique [29] is used to correct for BSSE. In this case, the counterpoise technique eliminates also the grid superposition error because the same grid is used to calculate the energy of each isolated monomer as well as for the dimer. The situation is different in the case of KSCED(m) calculations. The set of centers of the atomic basis sets is the same for both the isolated- and embedded monomer. The grids differ however. A larger one - corresponding to the whole dimer - is used for the embedded monomer, whereas a smaller one is used for isolated monomer. This imbalance of the grid sizes gives rise to a small numerical effect, for which is nevertheless also accounted.

Because of these particularities in correcting both types of errors in the interaction energy, the procedure to evaluate the corrected interaction energy denoted as $\Delta E(BSSE)$ throughout the text is given below. We use the notation from Ref. [30] and start with the KSCED(s) case.

- the electronic energy of a molecular system M at geometry G computed with basis set σ is denoted by $E_G^\sigma(M)$,
- $\Delta E = E_{AB}^{\alpha\cup\beta}(AB) - [E_A^\alpha(A) + E_B^\beta(B)]$ is the interaction energy without BSSE correction,
- $\Delta E(fCP) = E_{AB}^{\alpha\cup\beta}(AB) - [E_{AB}^{\alpha\cup\beta}(A) + E_{AB}^{\alpha\cup\beta}(B)]$ is the interaction energy corrected by means of the function counterpoise (fCP) technique
- $\Delta E(BSSE) = \Delta E(fCP) + E_{rel}^\alpha(A) + E_{rel}^\beta(B) = \Delta E(fCP) + E_{rel}$ is the fCP corrected interaction energy including the relaxation effects of the geometry of each monomer,

$$\begin{aligned} - E_{rel}^\alpha(A) &= E_{AB}^\alpha(A) - E_A^\alpha(A), \\ - E_{rel}^\beta(B) &= E_{AB}^\beta(B) - E_B^\beta(B). \end{aligned}$$

Note that, in the KSCED(m) case all the above formulas and definitions apply provided that the symbol $\alpha\cup\beta$ means that the α basis set is used for the subsystem A and β for the subsystem B. $\Delta E(fCP)$ defined above takes into account the correction for GSE in the KSCED(m) case.

The reported results are given as binding energies (the negative of the above-mentioned interaction energies) in kcal/mol. In the statistical analysis the errors have the following definitions ($E_i^{Ref.}$ is the reference binding energy taken from Refs. [9–11], E_i is the calculated binding energy, and N is the number of the complexes in the given set):

- MD (Mean difference) = $\frac{1}{N} \sum_{i=1}^N (E_i^{Ref.} - E_i)$,
- MAD (Mean absolute difference) = $\frac{1}{N} \sum_{i=1}^N |E_i^{Ref.} - E_i|$,
- MRAD (Mean relative absolute difference) = $\frac{1}{N} \sum_{i=1}^N |E_i^{Ref.} - E_i| / E_i^{Ref.}$,
- MAX (Maximal error) = $\max_{i=1}^N |E_i^{Ref.} - E_i|$,
- MAXR (Maximal relative error) = $\max_{i=1}^N |E_i^{Ref.} - E_i| / E_i^{Ref.}$.

Each of the above statistical measurements of differences between the results of a given approximate method and the reference data has a particular bias. If MAD is used, the statistical weight of systems with small binding energies such as that in the WI7/05 database, is reduced. Therefore, MRAD is used in most of the discussions.

Results

Performance of the LDA and GGA functionals in KSCED calculations

In this section, we consider the BSSE corrected numerical values of the interaction energies ($\Delta E(BSSE)$) obtained using the KSCED(s) type of expansion as the basis for comparisons with results of other methods, in which this correction is also taken into account. This choice is motivated by our intention to discuss the results approaching to the basis set limit as close as possible. As far as supermolecular and monomer expansions are concerned, the former one is obviously more appropriate because it includes all the functions present in the monomer expansion scheme and the functions localized on a partner molecule as well. Since the magnitudes of the counterpoise correction ($\Delta E(BSSE) - \Delta E$) are systematically smaller for aug-cc-pVTZ than for MG3S (compare Tables 1 and 3, 2 and 4), we start with the discussion of the results obtained using the former one collected in Tables 1 and 2.

Table 1 Binding energies (in kcal/mol) calculated using the LDA and aug-cc-pVTZ basis set

C, S	Compound	KSCED(m)			KSCED(s)			-E _{rel}	Ref.
		-ΔE	-ΔE(fCP)	-ΔE(BSSE)	-ΔE	-ΔE(fCP)	-ΔE(BSSE)		
01, 01	NH ₃ -NH ₃	3.63	3.60	3.63	3.74	3.69	3.72	0.04	3.15
02, 02	HF-HF	3.83	3.73	3.83	4.08	3.88	3.98	0.10	4.57
03, 03	H ₂ O-H ₂ O	4.79	4.71	4.79	4.92	4.80	4.88	0.08	4.97
04, 04	NH ₃ -H ₂ O	6.44	6.36	6.44	6.70	6.62	6.69	0.07	6.41
05, 05	HCONH ₂ -HCONH ₂	14.98	15.39	14.98	15.59	15.87	15.45	-0.41	14.94
06, 06	HCOOH-HCOOH	13.91	14.87	13.91	15.28	16.00	15.04	-0.96	16.15
	MD (MAD)	0.43 (0.62)	0.26 (0.56)	0.43 (0.62)	-0.02 (0.49)	-0.11 (0.45)	0.07 (0.53)		
	MAX (NMAX)	2.24 (6)	1.28 (6)	2.24 (6)	0.87 (6)	0.93 (5)	1.11 (6)		
	MRAD (MAXR, NMAXR)	0.08 (0.16, 2)	0.08 (0.18, 2)	0.08 (0.16, 2)	0.07 (0.19, 1)	0.08 (0.17, 1)	0.08 (0.18, 1)		
07, 01	C ₂ H ₄ -F ₂	-0.04	0.11	-0.04	0.24	0.22	0.08	-0.15	1.06
08, 02	NH ₃ -F ₂	0.15	0.47	0.15	0.37	0.64	0.32	-0.32	1.81
09, 03	C ₂ H ₂ -ClF	0.89	1.08	0.89	1.87	1.93	1.74	-0.19	3.81
10, 04	HCN-ClF	2.44	2.56	2.44	3.18	3.13	3.01	-0.11	4.86
11, 05	NH ₃ -Cl ₂	2.79	3.08	2.79	4.05	4.20	3.92	-0.29	4.88
12, 06	H ₂ O-ClF	2.69	2.81	2.69	3.61	3.64	3.52	-0.12	5.36
13, 07	NH ₃ -ClF	0.98	2.32	0.97	5.90	7.06	5.71	-1.35	10.62
	MD (MAD)	3.21 (3.21)	2.85 (2.85)	3.22 (3.22)	1.88 (1.88)	1.65 (1.65)	2.01 (2.01)		
	MAX (NMAX)	9.64 (7)	8.30 (7)	9.65 (7)	4.72 (7)	3.56 (7)	4.91 (7)		
	MRAD (MAXR, NMAXR)	0.72 (1.04, 1)	0.64 (0.90, 1)	0.72 (1.04, 1)	0.48 (0.80, 2)	0.44 (0.79, 1)	0.52 (0.92, 1)		
14, 01	H ₂ S-H ₂ S	2.02	1.98	2.02	2.18	2.05	2.09	0.04	1.66
15, 02	HCl-HCl	1.82	1.77	1.82	2.00	1.88	1.93	0.05	2.01
16, 03	H ₂ S-HCl	3.33	3.23	3.33	3.68	3.49	3.58	0.09	3.35
17, 04	CH ₃ Cl-HCl	2.65	2.64	2.65	2.91	2.83	2.84	0.01	3.55
18, 05	HCN-CH ₃ SH	4.05	4.04	4.05	4.27	4.15	4.17	0.01	3.59
19, 06	CH ₃ SH-HCl	3.96	3.91	3.97	4.48	4.32	4.39	0.06	4.16
	MD (MAD)	0.08 (0.35)	0.12 (0.38)	0.08 (0.35)	-0.20 (0.42)	-0.07 (0.35)	-0.11 (0.38)		
	MAX (NMAX)	0.90 (4)	0.91 (4)	0.90 (4)	0.68 (5)	0.72 (4)	0.71 (4)		
	MRAD (MAXR, NMAXR)	0.12 (0.25, 4)	0.13 (0.26, 4)	0.12 (0.25, 4)	0.14 (0.31, 1)	0.12 (0.23, 1)	0.13 (0.26, 1)		
20, 01	He-Ne	0.08	0.08	0.08	0.09	0.08	0.08	0.00	0.04
21, 02	He-Ar	0.11	0.11	0.11	0.12	0.11	0.11	0.00	0.06
22, 03	Ne-Ne	0.08	0.08	0.08	0.12	0.09	0.09	0.00	0.08
23, 04	Ne-Ar	0.13	0.13	0.13	0.16	0.13	0.13	0.00	0.13
24, 05	CH ₄ -Ne	0.15	0.15	0.15	0.25	0.17	0.17	0.00	0.22
25, 06	C ₆ H ₆ -Ne	0.20	0.23	0.20	0.43	0.34	0.31	-0.02	0.47
26, 07	CH ₄ -CH ₄	0.43	0.43	0.43	0.43	0.45	0.44	0.00	0.51
	MD (MAD)	0.05 (0.07)	0.04 (0.07)	0.05 (0.07)	-0.01 (0.05)	0.02 (0.05)	0.03 (0.05)		
	MAX (NMAX)	0.27 (6)	0.24 (6)	0.27 (6)	0.08 (7)	0.13 (6)	0.16 (6)		
	MRAD (MAXR, NMAXR)	0.41 (1.00, 1)	0.40 (1.00, 1)	0.41 (1.00, 1)	0.48 (1.25, 1)	0.37 (1.00, 1)	0.38 (1.00, 1)		
27, 01	C ₂ H ₂ -C ₂ H ₂	1.59	1.59	1.59	1.66	1.64	1.64	0.00	1.34
28, 02	C ₂ H ₄ -C ₂ H ₄	1.06	1.06	1.06	1.12	1.10	1.10	0.00	1.42
29, 03	S C ₆ H ₆ -C ₆ H ₆	0.29	0.29	0.29	0.49	0.28	0.28	0.00	1.81
30, 04	T C ₆ H ₆ -C ₆ H ₆	1.02	1.02	1.03	1.23	1.08	1.09	0.01	2.74
31, 05	PD C ₆ H ₆ -C ₆ H ₆	0.53	0.55	0.55	0.81	0.60	0.60	0.00	2.78
	MD (MAD)	1.12 (1.22)	1.12 (1.22)	1.11 (1.21)	0.96 (1.08)	1.08 (1.20)	1.08 (1.20)		
	MAX (NMAX)	2.25 (5)	2.23 (5)	2.23 (5)	1.97 (5)	2.18 (5)	2.18 (5)		
	MRAD (MAXR, NMAXR)	0.54 (0.84, 3)	0.54 (0.84, 3)	0.54 (0.84, 3)	0.49 (0.73, 3)	0.54 (0.85, 3)	0.54 (0.85, 3)		

C denotes the number of compound and S its number within the given set, respectively. KSCED(m), KSCED(s), and Ref. denote monomolecular KSCED, supermolecular KSCED, and reference results respectively. NMAX and NMAXR are the numbers of the complexes with the given maximal (MAX and MAXR, respectively) errors. For the definitions of errors and interaction energies see [Computational details](#).

Table 2 Binding energies (in kcal/mol) calculated using the GGA and aug-cc-pVTZ basis set

C, S	Compound	KSCED(m)			KSCED(s)			$-E_{rel}$	Ref.
		$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$	$-\Delta E$	$-\Delta E(fCP)$	$-\Delta EE(BSSE)$		
01, 01	NH ₃ -NH ₃	4.16	4.14	4.16	4.26	4.22	4.24	0.02	3.15
02, 02	HF-HF	4.68	4.61	4.68	4.88	4.72	4.80	0.08	4.57
03, 03	H ₂ O-H ₂ O	5.61	5.56	5.60	5.70	5.62	5.66	0.04	4.97
04, 04	NH ₃ -H ₂ O	7.40	7.37	7.40	7.62	7.58	7.61	0.03	6.41
05, 05	HCONH ₂ -HCONH ₂	17.02	17.64	17.02	17.56	18.00	17.38	-0.61	14.94
06, 06	HCOOH-HCOOH	16.83	18.42	16.84	18.06	19.37	17.79	-1.58	16.15
	MD (MAD)	-0.92 (0.92)	-1.26 (1.26)	-0.92 (0.92)	-1.31 (1.31)	-1.55 (1.55)	-1.22 (1.22)		
	MAX (NMAX)	2.08 (5)	2.70 (5)	2.08 (5)	2.62 (5)	3.22 (6)	2.44 (5)		
	MRAD (MAXR, NMAXR)	0.13 (0.32, 1)	0.15 (0.31, 1)	0.13 (0.32, 1)	0.17 (0.35, 1)	0.18 (0.34, 1)	0.16 (0.35, 1)		
07, 01	C ₂ H ₄ -F ₂	0.76	0.76	0.76	1.02	0.89	0.88	0.00	1.06
08, 02	NH ₃ -F ₂	1.26	1.31	1.26	1.48	1.52	1.46	-0.06	1.81
09, 03	C ₂ H ₂ -ClF	3.07	3.00	3.07	4.18	4.03	4.09	0.07	3.81
10, 04	HCN-ClF	4.54	4.44	4.54	5.32	5.07	5.17	0.09	4.86
11, 05	NH ₃ -Cl ₂	4.99	5.01	4.99	6.38	6.30	6.27	-0.03	4.88
12, 06	H ₂ O-ClF	4.95	4.84	4.95	5.94	5.76	5.87	0.11	5.36
13, 07	NH ₃ -ClF	7.37	7.83	7.37	13.94	14.25	13.79	-0.46	10.62
	MD (MAD)	0.78 (0.81)	0.74 (0.78)	0.78 (0.81)	-0.84 (0.94)	-0.77 (0.91)	-0.73 (0.88)		
	MAX (NMAX)	3.25 (7)	2.79 (7)	3.25 (7)	3.32 (7)	3.63 (7)	3.17 (7)		
	MRAD (MAXR, NMAXR)	0.18 (0.31, 7)	0.18 (0.28, 1)	0.18 (0.31, 7)	0.16 (0.31, 7)	0.16 (0.34, 7)	0.17 (0.30, 7)		
14, 01	H ₂ S-H ₂ S	2.36	2.32	2.36	2.48	2.39	2.43	0.04	1.66
15, 02	HCl-HCl	2.34	2.31	2.34	2.47	2.41	2.44	0.04	2.01
16, 03	H ₂ S-HCl	3.98	3.92	3.98	4.24	4.14	4.20	0.06	3.35
17, 04	CH ₃ Cl-HCl	3.68	3.62	3.68	3.87	3.79	3.85	0.06	3.55
18, 05	HCN-CH ₃ SH	4.64	4.63	4.64	4.83	4.74	4.75	0.00	3.59
19, 06	CH ₃ SH-HCl	5.10	5.07	5.11	5.51	5.42	5.46	0.04	4.16
	MD (MAD)	-0.63 (0.63)	-0.59 (0.59)	-0.63 (0.63)	-0.85 (0.85)	-0.76 (0.76)	-0.80 (0.80)		
	MAX (NMAX)	1.05 (5)	1.04 (5)	1.05 (5)	1.35 (6)	1.26 (6)	1.30 (6)		
	MRAD (MAXR, NMAXR)	0.22 (0.42, 1)	0.21 (0.40, 1)	0.22 (0.42, 1)	0.29 (0.49, 1)	0.26 (0.44, 1)	0.28 (0.46, 1)		
20, 01	He-Ne	0.31	0.30	0.30	0.32	0.31	0.31	0.00	0.04
21, 02	He-Ar	0.32	0.32	0.32	0.32	0.32	0.32	0.00	0.06
22, 03	Ne-Ne	0.40	0.40	0.40	0.43	0.41	0.41	0.00	0.08
23, 04	Ne-Ar	0.43	0.43	0.43	0.46	0.44	0.44	0.00	0.13
24, 05	CH ₄ -Ne	0.50	0.50	0.50	0.59	0.52	0.52	0.00	0.22
25, 06	C ₆ H ₆ -Ne	0.79	0.82	0.80	1.00	0.93	0.91	-0.02	0.47
26, 07	CH ₄ -CH ₄	0.84	0.84	0.84	0.85	0.86	0.86	0.00	0.51
	MD (MAD)	-0.30 (0.30)	-0.30 (0.30)	-0.30 (0.30)	-0.35 (0.35)	-0.33 (0.33)	-0.32 (0.32)		
	MAX (NMAX)	0.33 (7)	0.35 (6)	0.33 (6)	0.53 (6)	0.46 (6)	0.44 (6)		
	MRAD (MAXR, NMAXR)	2.86 (6.75, 1)	2.83 (6.50, 1)	2.82 (6.50, 1)	3.10 (7.00, 1)	2.95 (6.75, 1)	2.94 (6.75, 1)		
27, 01	C ₂ H ₂ -C ₂ H ₂	1.93	1.93	1.93	1.99	1.98	1.97	0.00	1.34
28, 02	C ₂ H ₄ -C ₂ H ₄	1.65	1.64	1.65	1.71	1.68	1.69	0.01	1.42
29, 03	S C ₆ H ₆ -C ₆ H ₆	0.94	0.94	0.94	1.15	0.93	0.93	0.00	1.81
30, 04	T C ₆ H ₆ -C ₆ H ₆	1.78	1.78	1.79	2.02	1.85	1.87	0.02	2.74
31, 05	PD C ₆ H ₆ -C ₆ H ₆	1.41	1.44	1.44	1.69	1.48	1.48	0.00	2.78
	MD (MAD)	0.48 (0.80)	0.47 (0.80)	0.47 (0.80)	0.31 (0.68)	0.43 (0.79)	0.43 (0.79)		
	MAX (NMAX)	1.37 (5)	1.34 (5)	1.34 (5)	1.09 (5)	1.30 (5)	1.30 (5)		
	MRAD (MAXR, NMAXR)	0.39 (0.49, 5)	0.38 (0.48, 5)	0.38 (0.48, 5)	0.34 (0.49, 1)	0.39 (0.49, 3)	0.39 (0.49, 3)		

C denotes the number of compound and S its number within the given set, respectively. KSCED(m), KSCED(s), and Ref. denote monomolecular KSCED, supermolecular KSCED, and reference results, respectively. NMAX and NMAXR are the numbers of the complexes with the given maximal (MAX and MAXR, respectively) errors. For the definitions of errors and interaction energies, see [Computational details](#).

Table 3 Binding energies (in kcal/mol) calculated using the LDA and MG3S basis set

C, S	Compound	KSCED(m)			KSCED(s)			$-E_{rel}$	Ref.
		$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$	$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$		
01, 01	NH ₃ -NH ₃	3.77	3.74	3.77	3.98	3.79	3.82	0.04	3.15
02, 02	HF-HF	3.93	3.85	3.93	4.44	3.92	4.00	0.08	4.57
03, 03	H ₂ O-H ₂ O	5.08	5.01	5.08	5.47	4.92	4.99	0.07	4.97
04, 04	NH ₃ -H ₂ O	6.55	6.49	6.55	7.05	6.68	6.74	0.05	6.41
05, 05	HCONH ₂ -HCONH ₂	14.42	14.96	14.42	15.63	15.75	15.21	-0.54	14.94
06, 06	HCOOH-HCOOH	13.24	14.20	13.24	15.36	15.69	14.73	-0.96	16.15
	MD (MAD)	0.53 (0.82)	0.32 (0.57)	0.53 (0.82)	-0.29 (0.60)	-0.09 (0.48)	0.12 (0.55)		
	MAX (NMAX)	2.91 (6)	1.95 (6)	2.91 (6)	0.83 (1)	0.81 (5)	1.42 (6)		
	MRAD (MAXR, NMAXR)	0.10 (0.20, 1)	0.08 (0.19, 1)	0.10 (0.20, 1)	0.10 (0.26, 1)	0.08 (0.20, 1)	0.08 (0.21, 1)		
07, 01	C ₂ H ₄ -F ₂	-0.12	0.04	-0.12	0.33	0.18	0.02	-0.15	1.06
08, 02	NH ₃ -F ₂	-0.15	0.18	-0.15	0.45	0.50	0.16	-0.33	1.81
09, 03	C ₂ H ₂ -ClF	0.72	0.97	0.72	1.78	1.80	1.55	-0.25	3.81
10, 04	HCN-ClF	2.34	2.52	2.34	3.39	3.17	2.99	-0.18	4.86
11, 05	NH ₃ -Cl ₂	2.76	3.12	2.76	4.04	4.12	3.77	-0.35	4.88
12, 06	H ₂ O-ClF	2.69	2.87	2.69	3.93	3.58	3.40	-0.18	5.36
13, 07	NH ₃ -ClF	0.69	2.11	0.68	5.88	6.83	5.41	-1.42	10.62
	MD (MAD)	3.35 (3.35)	2.94 (2.94)	3.35 (3.35)	1.80 (1.80)	1.75 (1.75)	2.16 (2.16)		
	MAX (NMAX)	9.93 (7)	8.51 (7)	9.94 (7)	4.74 (7)	3.79 (7)	5.21 (7)		
	MRAD (MAXR, NMAXR)	0.77 (1.11, 1)	0.67 (0.96, 1)	0.77 (1.11, 1)	0.45 (0.75, 2)	0.47 (0.83, 1)	0.56 (0.98, 1)		
14, 01	H ₂ S-H ₂ S	1.93	1.90	1.93	2.20	2.00	2.03	0.03	1.66
15, 02	HCl-HCl	1.73	1.69	1.73	2.07	1.83	1.87	0.04	2.01
16, 03	H ₂ S-HCl	3.14	3.07	3.14	3.71	3.39	3.46	0.07	3.35
17, 04	CH ₃ Cl-HCl	2.51	2.54	2.51	2.96	2.78	2.75	-0.02	3.55
18, 05	HCN-CH ₃ SH	4.04	4.03	4.03	4.33	4.19	4.19	0.01	3.59
19, 06	CH ₃ SH-HCl	3.75	3.74	3.76	4.51	4.24	4.26	0.02	4.16
	MD (MAD)	0.20 (0.44)	0.23 (0.45)	0.20 (0.44)	-0.24 (0.44)	-0.02 (0.34)	-0.04 (0.35)		
	MAX (NMAX)	1.04 (4)	1.01 (4)	1.04 (4)	0.74 (5)	0.77 (4)	0.80 (4)		
	MRAD (MAXR, NMAXR)	0.15 (0.29, 4)	0.15 (0.28, 4)	0.15 (0.29, 4)	0.15 (0.33, 1)	0.12 (0.22, 4)	0.12 (0.23, 4)		
20, 01	He-Ne	0.07	0.07	0.07	0.10	0.07	0.07	0.00	0.04
21, 02	He-Ar	0.11	0.11	0.11	0.13	0.11	0.11	0.00	0.06
22, 03	Ne-Ne	0.07	0.07	0.07	0.19	0.08	0.08	0.00	0.08
23, 04	Ne-Ar	0.11	0.11	0.11	0.22	0.12	0.12	0.00	0.13
24, 05	CH ₄ -Ne	0.12	0.12	0.12	0.28	0.15	0.15	0.00	0.22
25, 06	C ₆ H ₆ -Ne	0.16	0.19	0.17	0.43	0.20	0.18	-0.03	0.47
26, 07	CH ₄ -CH ₄	0.41	0.41	0.41	0.45	0.43	0.43	0.00	0.51
	MD (MAD)	0.07 (0.09)	0.06 (0.08)	0.06 (0.09)	-0.04 (0.07)	0.05 (0.07)	0.05 (0.08)		
	MAX (NMAX)	0.31 (6)	0.28 (6)	0.30 (6)	0.11 (3)	0.27 (6)	0.29 (6)		
	MRAD (MAXR, NMAXR)	0.45 (0.83, 2)	0.44 (0.83, 2)	0.45 (0.83, 2)	0.74 (1.50, 1)	0.39 (0.83, 2)	0.39 (0.83, 2)		
27, 01	C ₂ H ₂ -C ₂ H ₂	1.59	1.60	1.59	1.68	1.65	1.64	-0.01	1.34
28, 02	C ₂ H ₄ -C ₂ H ₄	1.06	1.05	1.05	1.12	1.08	1.08	0.00	1.42
29, 03	S C ₆ H ₆ -C ₆ H ₆	0.18	0.18	0.18	0.52	0.16	0.16	0.00	1.81
30, 04	T C ₆ H ₆ -C ₆ H ₆	1.08	1.07	1.09	1.41	1.06	1.07	0.01	2.74
31, 05	PD C ₆ H ₆ -C ₆ H ₆	0.46	0.49	0.49	1.01	0.57	0.57	0.00	2.78
	MD (MAD)	1.14 (1.24)	1.14 (1.24)	1.14 (1.24)	0.87 (1.01)	1.11 (1.24)	1.11 (1.23)		
	MAX (NMAX)	2.32 (5)	2.29 (5)	2.29 (5)	1.77 (5)	2.21 (5)	2.21 (5)		
	MRAD (MAXR, NMAXR)	0.56 (0.90, 3)	0.56 (0.90, 3)	0.55 (0.90, 3)	0.46 (0.71, 3)	0.56 (0.91, 3)	0.56 (0.91, 3)		
32, 01	A-T	3.21	4.23	3.23	4.80	4.44	3.44	-0.99	11.60
33, 02	G-C	9.24	11.27	9.23	10.89	11.61	9.57	-2.04	16.19
34, 03	anti C-C	4.71	6.18	4.64	6.07	6.37	4.83	-1.54	9.90

Table 3 (continued)

C, S	Compound	KSCED(m)			KSCED(s)			$-E_{rel}$	Ref.
		$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$	$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$		
35, 04	displ C-C	5.35	5.85	5.31	6.42	5.92	5.38	-0.54	9.43
36, 05	par C-C	-4.97	-5.45	-4.94	-3.92	-5.28	-4.77	0.51	-2.45
37, 06	U-U	4.05	4.58	4.00	5.35	4.77	4.19	-0.58	10.30
38, 07	WC A-T	15.06	15.12	15.06	16.68	16.09	16.03	-0.07	15.40
39, 08	WC G-C	28.62	29.18	28.62	30.81	30.55	29.98	-0.56	28.80
	MD (MAD)	4.24 (4.24)	3.53 (3.62)	4.25 (4.25)	2.76 (3.58)	3.09 (3.70)	3.81 (4.27)		
	MAX (NMAX)	8.39 (1)	7.37 (1)	8.37 (1)	6.80 (1)	7.16 (1)	8.16 (1)		
	MRAD (MAXR, NMAXR)	0.47 (1.03, 5)	0.44 (1.22, 5)	0.47 (1.02, 5)	0.36 (0.60, 5)	0.43 (1.16, 5)	0.46 (0.95, 5)		

C denotes the number of compound and S its number within the given set, respectively. KSCED(m), KSCED(s), and Ref. denote monomolecular KSCED, supermolecular KSCED, and reference results respectively. NMAX and NMAXR are the numbers of the complexes with the given maximal (MAX and MAXR, respectively) errors. For the definitions of errors and interaction energies see [Computational details](#).

KSCED LDA performs very well for the complexes dominated by electrostatic interactions (HB6/04 and DI6/04). For hydrogen bonded complexes, MRAD equals 8% whereas it reaches 13% for that dominated by dipolar interactions. For weakly bounded complexes (WI7/05), MRAD reaches 38%. Complexes with strong “charge transfer” character (CT7/04, MRAD=52%) and π -stacked systems (PPS5/05, MRAD=54%) are not satisfactorily described at the KSCED LDA level. Introduction of gradients into approximate functionals (KSCED GGA) worsens the description of hydrogen bounded complexes (MRAD increases to from 8% to 16%) and dipole-dipole-bounded ones (MRAD increases from 13% to 28%). For most weakly bound ones, KSCED GGA leads to qualitatively wrong results (MRAD amounts to almost 300%). Compared to KSCED LDA, KSCED GGA improves, however, the interaction energies in two classes of systems. In the CT7/04 group, MRAD is reduced from 52% to 17%, whereas it is reduced from 54% to 39% in the case of π -stacked complexes. These trends are in line with our observations made reported in Ref. [6]. The role of gradient-dependent terms on the accuracy of the interaction energies in one group of systems (hydrogen-, dipolar-, and weakly bonded) is qualitatively different than in the other (π -stacking and “charge transfer” complexes).

Tables 3 and 4 collect the results obtained using the MG3S basis set. Additionally, data for the nucleic acid base pairs (BP8/05) are also included. The statistical errors of the binding energies calculated using the aug-cc-pVTZ and MG3S are similar reflecting the same trends as the ones discussed previously for the compounds in the CT7/04, WI7/05, HB6/04, DI6/04, and PPS5/05 databases. In the case of nucleic acids base pairs, introducing gradient dependency into the approximate functionals decreases MRAD from 46% to 37%. Inspection of individual base-pairs indicates that KSCED LDA is a better approximation

for the two most strongly bound pairs (hydrogen bonded) whereas KSCED GGA performs better for the weaker ones (π -stacked).

Table 5 shows MRADs for both methods discussed in this work (KSCED LDA and KSCED GGA) together with MRADs calculated based on the supplementary material in Ref. [9] for the same reference data. We bring here to the reader’s attention the fact that the WI9/04 [9] database is used here instead of the discussed previously WI7/05 in order to make possible direct comparisons with the results reported in Ref. [9]. Compared to Table 8 of Ref. [9], where a different statistical measure was used, the order of functionals does not change significantly. The overall MRAD equals 30% and 88% for KSCED LDA and KSCED GGA, respectively. The corresponding MAD amount to 0.73 and 0.77 kcal/mol. Using these overall measures of the accuracy of a given method places KSCED LDA among the best ones. For comparison, MRAD corresponding to MP2 is larger (31%). It is important to note that all Kohn-Sham-based methods for which MRAD is smaller than that of KSCED LDA use empirical parameters, which is not the case of the KSCED LDA method which is based entirely on first principles.

The adequacy of KSCED LDA methods is not uniform in all classes of intermolecular systems. In the case of the WI9/04 database, it is *the best among all considered methods*. For the complexes with the strong “charge transfer” character, it is, however, less good but as indicated in the previous discussions KSCED GGA is more adequate as *one of the best among all considered methods*.

Monomolecular vs. supermolecular basis set

Comparing the KSCED(m) and the KSCED(s) interaction energies $\Delta E(BSSE)$ (see Tables 1, 2, 3 and 4) shows that using supermolecular basis sets to construct orbitals for each

Table 4 Binding energies (in kcal/mol) calculated using the GGA and MG3S basis set

C, S	Compound	KSCED(m)			KSCED(s)			$-E_{rel}$	Ref.
		$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$	$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$		
01, 01	NH ₃ -NH ₃	4.34	4.32	4.34	4.55	4.35	4.37	0.02	3.15
02, 02	HF-HF	4.85	4.79	4.85	5.33	4.81	4.86	0.06	4.57
03, 03	H ₂ O-H ₂ O	5.93	5.89	5.93	6.30	5.77	5.80	0.04	4.97
04, 04	NH ₃ -H ₂ O	7.51	7.49	7.51	8.01	7.66	7.68	0.02	6.41
05, 05	HCONH ₂ -HCONH ₂	16.49	17.23	16.49	17.62	17.94	17.20	-0.74	14.94
06, 06	HCOOH-HCOOH	16.11	17.69	16.11	18.10	19.06	17.49	-1.58	16.15
	MD (MAD)	-0.84 (0.85)	-1.20 (1.20)	-0.84 (0.85)	-1.62 (1.62)	-1.57 (1.57)	-1.20 (1.20)		
	MAX (NMAX)	1.55 (5)	2.29 (5)	1.55 (5)	2.68 (5)	3.00 (5)	2.26 (5)		
	MRAD (MAXR, NMAXR)	0.15 (0.38, 1)	0.17 (0.37, 1)	0.15 (0.38, 1)	0.24 (0.44, 1)	0.20 (0.38, 1)	0.17 (0.39, 1)		
07, 01	C ₂ H ₄ -F ₂	0.77	0.79	0.78	1.23	0.94	0.93	-0.01	1.06
08, 02	NH ₃ -F ₂	1.11	1.18	1.11	1.75	1.53	1.46	-0.07	1.81
09, 03	C ₂ H ₂ -ClF	2.99	2.98	2.99	4.12	3.91	3.92	0.01	3.81
10, 04	HCN-ClF	4.59	4.56	4.58	5.59	5.19	5.22	0.03	4.86
11, 05	NH ₃ -Cl ₂	5.08	5.16	5.07	6.43	6.22	6.14	-0.08	4.88
12, 06	H ₂ O-ClF	5.14	5.08	5.13	6.36	5.79	5.84	0.05	5.36
13, 07	NH ₃ -ClF	7.06	7.57	7.05	13.54	13.57	13.05	-0.52	10.62
	MD (MAD)	0.81 (0.87)	0.73 (0.81)	0.81 (0.87)	-0.95 (0.96)	-0.68 (0.79)	-0.59 (0.73)		
	MAX (NMAX)	3.56 (7)	3.05 (7)	3.57 (7)	2.92 (7)	2.95 (7)	2.43 (7)		
	MRAD (MAXR, NMAXR)	0.19 (0.39, 2)	0.18 (0.35, 2)	0.19 (0.39, 2)	0.17 (0.32, 5)	0.14 (0.28, 7)	0.14 (0.26, 5)		
14, 01	H ₂ S-H ₂ S	2.30	2.27	2.29	2.53	2.35	2.38	0.02	1.66
15, 02	HCl-HCl	2.28	2.25	2.28	2.57	2.36	2.39	0.02	2.01
16, 03	H ₂ S-HCl	3.84	3.80	3.84	4.33	4.06	4.10	0.04	3.35
17, 04	CH ₃ Cl-HCl	3.59	3.56	3.59	3.97	3.76	3.78	0.02	3.55
18, 05	HCN-CH ₃ SH	4.66	4.65	4.65	4.90	4.78	4.78	0.00	3.59
19, 06	CH ₃ SH-HCl	4.95	4.97	4.96	5.60	5.37	5.37	-0.01	4.16
	MD (MAD)	-0.55 (0.55)	-0.53 (0.53)	-0.55 (0.55)	-0.93 (0.93)	-0.73 (0.73)	-0.75 (0.75)		
	MAX (NMAX)	1.07 (5)	1.06 (5)	1.06 (5)	1.44 (6)	1.21 (6)	1.21 (6)		
	MRAD (MAXR, NMAXR)	0.19 (0.39, 1)	0.19 (0.37, 1)	0.19 (0.38, 1)	0.32 (0.52, 1)	0.25 (0.42, 1)	0.26 (0.43, 1)		
20, 01	He-Ne	0.30	0.29	0.29	0.33	0.30	0.30	0.00	0.04
21, 02	He-Ar	0.31	0.31	0.31	0.33	0.31	0.31	0.00	0.06
22, 03	Ne-Ne	0.41	0.40	0.40	0.53	0.42	0.42	0.00	0.08
23, 04	Ne-Ar	0.44	0.44	0.44	0.55	0.45	0.45	0.00	0.13
24, 05	CH ₄ -Ne	0.50	0.50	0.50	0.65	0.53	0.53	0.00	0.22
25, 06	C ₆ H ₆ -Ne	0.86	0.89	0.87	1.13	0.90	0.88	-0.02	0.47
26, 07	CH ₄ -CH ₄	0.84	0.85	0.84	0.87	0.87	0.86	0.00	0.51
	MD (MAD)	-0.31 (0.31)	-0.31 (0.31)	-0.31 (0.31)	-0.41 (0.41)	-0.32 (0.32)	-0.32 (0.32)		
	MAX (NMAX)	0.39 (6)	0.42 (6)	0.40 (6)	0.66 (6)	0.43 (6)	0.41 (6)		
	MRAD (MAXR, NMAXR)	2.85 (6.50, 1)	2.80 (6.25, 1)	2.80 (6.25, 1)	3.52 (7.25, 1)	2.92 (6.50, 1)	2.91 (6.50, 1)		
27, 01	C ₂ H ₂ -C ₂ H ₂	1.98	1.99	1.98	2.05	2.02	2.01	0.00	1.34
28, 02	C ₂ H ₄ -C ₂ H ₄	1.73	1.72	1.73	1.77	1.72	1.73	0.01	1.42
29, 03	S C ₆ H ₆ -C ₆ H ₆	1.02	1.02	1.02	1.30	0.98	0.98	0.00	1.81
30, 04	T C ₆ H ₆ -C ₆ H ₆	1.98	1.97	1.98	2.24	1.90	1.92	0.02	2.74
31, 05	PD C ₆ H ₆ -C ₆ H ₆	1.60	1.63	1.63	2.02	1.62	1.62	0.00	2.78
	MD (MAD)	0.36 (0.74)	0.35 (0.73)	0.35 (0.73)	0.14 (0.57)	0.37 (0.76)	0.37 (0.76)		
	MAX (NMAX)	1.18 (5)	1.15 (5)	1.15 (5)	0.76 (5)	1.16 (5)	1.16 (5)		
	MRAD (MAXR, NMAXR)	0.37 (0.48, 1)	0.37 (0.49, 1)	0.36 (0.48, 1)	0.30 (0.53, 1)	0.38 (0.51, 1)	0.38 (0.50, 1)		
32, 01	A-T	7.15	7.93	7.18	8.55	7.93	7.18	-0.75	11.60
33, 02	G-C	13.16	14.92	13.14	14.65	15.10	13.32	-1.78	16.19
34, 03	anti C-C	7.68	8.77	7.60	8.91	8.82	7.66	-1.16	9.90

Table 4 (continued)

<i>C</i> , <i>S</i>	Compound	KSCED(m)			KSCED(s)			$-E_{rel}$	Ref.
		$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$	$-\Delta E$	$-\Delta E(fCP)$	$-\Delta E(BSSE)$		
35, 04	displ C-C	7.39	7.62	7.34	8.36	7.59	7.31	-0.28	9.43
36, 05	par C-C	0.58	-3.17	0.61	1.53	-3.10	0.68	3.78	-2.45
37, 06	U-U	6.58	7.16	6.52	7.76	7.22	6.58	-0.64	10.30
38, 07	WC A-T	17.57	18.36	17.57	19.05	19.20	18.41	-0.79	15.40
39, 08	WC G-C	31.49	33.44	31.49	33.52	34.66	32.70	-1.95	28.80
	MD (MAD)	0.95 (2.92)	0.52 (2.42)	0.96 (2.94)	-0.40 (2.69)	0.22 (2.63)	0.67 (3.18)		
	MAX (NMAX)	4.45 (1)	4.64 (8)	4.42 (1)	4.72 (8)	5.86 (8)	4.42 (1)		
	MRAD (MAXR, NMAXR)	0.36 (1.24, 5)	0.21 (0.32, 1)	0.36 (1.25, 5)	0.36 (1.62, 5)	0.21 (0.32, 1)	0.37 (1.28, 5)		

C denotes the number of compound and *S* its number within the given set, respectively. KSCED(m), KSCED(s), and Ref. denote monomolecular KSCED, supermolecular KSCED, and reference results respectively. NMAX and NMAXR are the numbers of the complexes with the given maximal (MAX and MAXR, respectively) errors. For the definitions of errors and interaction energies see [Computational details](#).

subsystem is not necessary for all systems except for the complexes with a strong “charge transfer” character (CT7/04 database). For such systems, KSCED(s) and KSCED(m) interaction energies differ qualitatively reaching the order of magnitude of the interaction energy itself for both KSCED LDA and KSCED GGA. For instance, the largest difference between GGA KSCED(s) and KSCED(m) equals to 6.42 kcal/mol for NH₃-ClF whereas the reference binding energy amounts to 10.62 kcal/mol in this complex.

Potential energy surface scan: stacked C-C

The previous sections concerned the accuracy of the energy for a particular geometry of the dimer. Here, we analyze the potential energy curves corresponding to changes of intermolecular distance in a stacked cytosine dimer. Such analysis is made in view of possible applications of the subsystem-based DFT formalism in practical simulations and complements a similar analysis for hydrogen bound complexes reported elsewhere [5]. The stacked cytosine dimer was chosen because of the notorious difficulties the Kohn-Sham-based methods face in describing potential energy surface in such systems. Figure 1 shows several potential energy curves obtained using gradient-free (LDA) and gradient-dependent (GGA) approximations to the relevant density functionals applied in Kohn-Sham and KSCED formal frameworks. Each point of the curve corresponds to the same internal geometry of each monomer and the relative orientation denoted as structure 14 in Ref. [31] but a different intermolecular distance.

Equilibrium intermolecular distance and the well depth, the key parameters of the potential energy curve, depend strongly on the calculation method used. The Kohn-Sham LDA calculations lead to surprisingly good well depth, which agrees with the reference data within few percent.

The equilibrium intermolecular distance is, however, too short by about 0.1 Å. Introducing gradient dependency into the exchange-correlation functional by means of using the PW91 approximation worsens significantly both the intermolecular distance and the well depth. Turning back to the results of KSCED calculations, Fig. 1 shows clearly that KSCED GGA performs better than KSCED LDA. The two key parameters of the potential energy surface are slightly worse than the benchmark data (intermolecular distance being about 0.2 Å too short and the well depth underestimated by about 25%). These shifts are in line with the trends in the benzene dimer [7].

Surprisingly, the KS LDA results are closer to the reference data than probably any existing functional (see comparison of different DFT methods in Ref. [11]). The good performance of LDA in Kohn-Sham calculations is probably accidental in this case. The PW91 approximation to the exchange-correlation functional in Kohn-Sham calculations leads to significantly worse parameters of the potential energy surface (see Fig. 1) in agreement with Ref. [31]. This example shows also that formulating trends based on an appraisal of approximations to the exchange-correlation functional using a particular testing set should be made carefully. The same concerns the performance of approximations to the density functionals in calculations based on the subsystem formulation of DFT.

Conclusions

Interaction energies derived using two types of approximations (LDA and GGA) for exchange-correlation and non-additive kinetic energy applied within the framework of the subsystem formulation of density functional theory were obtained for 39 weakly bound complexes at their equilibrium geometry.

Table 5 Relative mean absolute errors (MRAD) of binding energies (1=100%) calculated using different methods and the MG3S basis set, (including the correction for BSSE), sorted accordingly to the average MRAD

method	HB6/04	CT7/04	DI6/04	WI9/04	average
B971	0.04 (0.10, 5)	0.25 (0.52, 2)	0.09 (0.15, 4)	0.41 (0.75 , 1)	0.22 (0.75, 20)
B98	0.07 (0.12, 5)	0.16 (0.35, 2)	0.14 (0.25, 4)	0.49 (0.98, 7)	0.24 (0.98, 26)
PBE1KCIS	0.05 (0.12, 5)	0.18 (0.46, 2)	0.14 (0.25, 4)	0.51 (0.90, 7)	0.25 (0.90, 26)
PBE1PBE	0.03 (0.07, 5)	0.16 (0.28, 2)	0.13 (0.19, 4)	0.56 (1.08, 7)	0.25 (1.08, 26)
MPW1B95	0.10 (0.17, 1)	0.06 (0.12, 4)	0.23 (0.37, 1)	0.55 (0.89, 6)	0.26 (0.89, 25)
MPWB1K	0.09 (0.16, 1)	0.17 (0.41, 1)	0.24 (0.37, 1)	0.56 (0.89, 6)	0.29 (0.89, 25)
MP2	0.10 (0.14, 2)	0.09 (0.27, 2)	0.08 (0.12, 6)	0.78 (1.50, 3)	0.31 (1.50, 22)
TPSS1KCIS	0.10 (0.16, 1)	0.25 (0.61, 2)	0.20 (0.33, 4)	0.65 (1.37, 7)	0.33 (1.37, 26)
X3LYP	0.05 (0.11, 1)	0.20 (0.54, 2)	0.22 (0.38, 1)	0.70 (1.47, 7)	0.33 (1.47, 26)
TPSSh	0.10 (0.17, 1)	0.29 (0.66, 2)	0.22 (0.33, 1)	0.82 (1.57, 7)	0.40 (1.57, 26)
mPW1PW91	0.10 (0.18, 1)	0.10 (0.20, 1)	0.24 (0.38, 1)	0.97 (1.76, 7)	0.40 (1.76, 26)
MPWKIS1K	0.12 (0.20, 1)	0.28 (0.62, 1)	0.32 (0.46, 1)	0.79 (1.49, 7)	0.41 (1.49, 26)
MPW1KCIS	0.15 (0.22, 1)	0.19 (0.53, 2)	0.30 (0.43, 1)	0.85 (1.71, 7)	0.41 (1.71, 26)
PBE	0.03 (0.07, 5)	0.81 (1.82, 2)	0.11 (0.29, 6)	0.55 (1.00, 1)	0.41 (1.82, 08)
TPSSKCIS	0.10 (0.15, 1)	0.58 (1.41, 2)	0.19 (0.30, 4)	0.71 (1.41, 7)	0.44 (1.41, 08)
BHandHLYP	0.06 (0.14, 1)	0.24 (0.55, 1)	0.32 (0.51, 1)	0.94 (1.47, 7)	0.44 (1.47, 26)
TPSS	0.10 (0.17, 1)	0.55 (1.28, 2)	0.20 (0.31, 4)	0.81 (1.59, 7)	0.46 (1.59, 26)
MPW1K	0.08 (0.16, 1)	0.25 (0.71, 1)	0.25 (0.40, 1)	1.02 (1.63, 7)	0.46 (1.63, 26)
mPWB95	0.13 (0.18, 1)	0.70 (1.77, 2)	0.21 (0.30, 1)	0.66 (1.25, 1)	0.46 (1.77, 08)
MPW3LYP	0.05 (0.08, 2)	0.34 (0.79, 2)	0.13 (0.23, 1)	1.04 (2.75, 1)	0.46 (2.75, 20)
PBEKCIS	0.05 (0.12, 5)	0.78 (1.88, 2)	0.11 (0.17, 4)	0.74 (1.75, 1)	0.47 (1.88, 08)
mPWKCIS	0.16 (0.22, 1)	0.56 (1.54, 2)	0.26 (0.41, 4)	0.88 (1.82, 7)	0.51 (1.82, 26)
mPWPW91	0.12 (0.18, 1)	0.59 (1.52, 2)	0.21 (0.33, 4)	1.01 (1.96, 7)	0.55 (1.96, 26)
mPWPBE	0.13 (0.19, 1)	0.58 (1.48, 2)	0.22 (0.34, 4)	1.04 (2.00, 7)	0.55 (2.00, 26)
B972	0.21 (0.29, 1)	0.16 (0.31, 1)	0.36 (0.50, 4)	1.27 (1.88, 7)	0.56 (1.88, 26)
XLYP	0.11 (0.19, 1)	0.66 (1.78, 2)	0.28 (0.45, 1)	1.10 (1.98, 7)	0.60 (1.98, 26)
B3LYP	0.11 (0.21, 1)	0.14 (0.42, 2)	0.34 (0.54, 1)	1.68 (2.50, 2)	0.66 (2.50, 21)
mPWLYP	0.06 (0.12, 5)	0.84 (2.14, 1)	0.14 (0.23, 1)	1.39 (3.75, 1)	0.70 (3.75, 20)
O3LYP	0.40 (0.55, 1)	0.33 (0.60, 3)	0.74 (0.97, 1)	1.29 (2.65, 7)	0.72 (2.65, 26)
PW91PW91	0.07 (0.12, 4)	0.95 (2.08, 1)	0.14 (0.36, 6)	1.64 (5.00, 1)	0.81 (5.00, 20)
HCTH	0.22 (0.30, 5)	0.52 (1.55, 1)	0.23 (0.45, 4)	2.00 (5.25, 1)	0.87 (5.25, 20)
OLYP	0.51 (0.68, 1)	0.43 (0.67, 2)	0.87 (1.13, 1)	1.55 (3.12, 7)	0.90 (3.12, 26)
BB1K	0.18 (0.30, 1)	0.35 (0.85, 1)	0.42 (0.61, 1)	2.14 (3.50, 1)	0.90 (3.50, 20)
B3P86	0.07 (0.16, 1)	0.16 (0.30, 7)	0.21 (0.36, 1)	2.72 (5.00, 1)	0.97 (5.00, 20)
BLYP	0.19 (0.31, 1)	0.52 (1.55, 2)	0.42 (0.65, 1)	2.37 (3.67, 2)	1.02 (3.67, 21)
B1B95	0.22 (0.36, 1)	0.23 (0.58, 1)	0.46 (0.67, 1)	2.54 (4.25, 1)	1.02 (4.25, 20)
B3PW91	0.20 (0.33, 1)	0.17 (0.50, 1)	0.42 (0.63, 1)	2.94 (4.75, 1)	1.09 (4.75, 20)
SVWN5	0.54 (0.66, 1)	1.78 (3.20, 1)	0.89 (1.19, 6)	1.18 (2.25, 1)	1.13 (3.20, 07)
SPWL	0.54 (0.66, 1)	1.78 (3.20, 1)	0.89 (1.19, 6)	1.18 (2.25, 1)	1.13 (3.20, 07)
SVWN3	0.57 (0.70, 1)	1.84 (3.30, 1)	0.95 (1.24, 6)	1.31 (2.50, 1)	1.21 (3.30, 07)
VSXC	0.09 (0.13, 4)	1.02 (3.01, 1)	0.28 (0.54, 4)	3.15 (5.68, 6)	1.35 (5.68, 25)
BB95	0.29 (0.43, 1)	0.41 (1.25, 2)	0.49 (0.72, 1)	3.38 (6.00, 1)	1.36 (6.00, 20)
BP86	0.15 (0.27, 1)	0.49 (1.29, 2)	0.29 (0.47, 1)	3.81 (7.25, 1)	1.44 (7.25, 20)
BPW91	0.28 (0.43, 1)	0.33 (1.01, 2)	0.49 (0.73, 1)	3.88 (6.50, 1)	1.50 (6.50, 20)
BPBE	0.28 (0.44, 1)	0.32 (0.96, 2)	0.50 (0.75, 1)	3.95 (6.75, 1)	1.52 (6.75, 20)
G96LYP	0.46 (0.77, 1)	0.33 (0.59, 2)	0.97 (1.43, 1)	8.10 (16.0, 1)	2.99 (16.0, 20)
KSCED(LDA)	0.08 (0.21, 1)	0.56 (0.98, 1)	0.12 (0.23, 4)	0.36 (0.83, 2)	0.30 (0.98, 07)
KSCED(GGA)	0.17 (0.39, 1)	0.14 (0.26, 5)	0.26 (0.43, 1)	2.34 (6.50, 1)	0.88 (6.50, 20)

The maximal relative absolute error (MAXR) and the number of the complex with this error in the set are given in parentheses. For the definitions of errors see [Computational details](#).

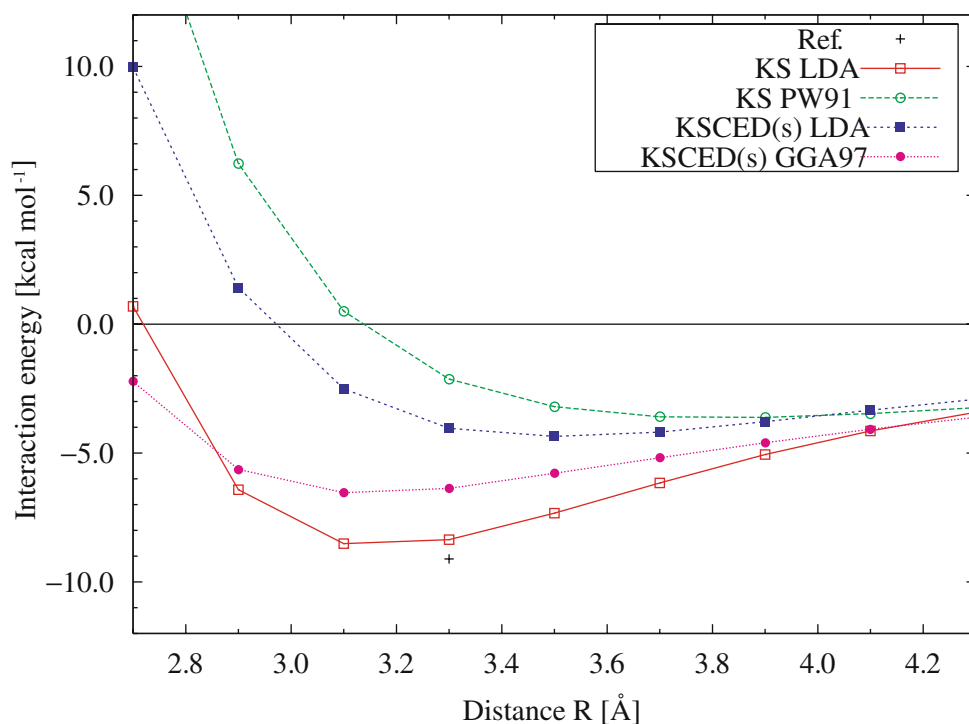


Fig. 1 Interaction energy of the “face to face” stacked cytosine dimer (structure 14 in the Ref. [31]) as a function of the distance R between the planes of the cytosine monomers (geometries taken from [31]).

The benchmark data (well depth and the equilibrium interplanar distance) taken from Ref. [31] is indicated as Ref

For several classes of systems, such as hydrogen-bonded complexes, complexes with a strong dipole-dipole character, the weakest complexes of van der Waals type, the local density approximation performs very well. Compared to the results obtained using parameter-free methods based on many-body perturbation theory (MP2) and the Kohn-Sham formalism (LDA, PBE, TPSS, TPSSh), KSCED LDA, being entirely parameter-free, is the most accurate for weakly bound complexes, second best (after KS PBE) for hydrogen bonded complexes, and third best for complexes with strong dipole-dipole character (mean relative absolute deviations from the reference data are only slightly larger than that of either MP2 or KS PBE). KSCED LDA, compares also favorably with Kohn-Sham-based methods using empirical approximations to the exchange-correlation functional, remaining in the group of methods performing the best. In these types of systems, introducing gradient-dependency into the relevant density functionals worsens noticeably the accuracy of the interaction energies. KSCED LDA fails, however, to describe properly π -stacking interactions and complexes of strong “charge-transfer” character. In this type of systems, introducing gradient-dependency into the relevant density functionals results in a significant improvement in the accuracy. The origin of this opposite effect of gradient-dependency remains to be clarified.

From the practical point of view, the calculations are computationally more costly than the corresponding LDA or GGA Kohn-Sham calculations because they were performed using supermolecular expansion for each electron density ρ_A and ρ_B . A much less computationally demanding type of calculations, in which the monomer-centered basis set is used for each subsystem (KSCED(m)) does not lead to noticeable deterioration of the interaction energies in most cases. This simplification is, however, not adequate in the case of complexes with a “charge transfer” character (CT7/04 database).

Acknowledgements This work has been supported by Swiss National Scientific Foundation.

References

1. Kohn W, Sham LJ (1965) *Phys Rev* 140:A1133–A1138, DOI [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133)
2. Wesolowski TA, Ellinger Y, Weber J (1998) *J Chem Phys* 108:6078–6083
3. Cortona P (1991) *Phys Rev B* 44:8454–8458, DOI [10.1103/PhysRevB.44.8454](https://doi.org/10.1103/PhysRevB.44.8454)

4. Wesolowski TA, Weber J (1996) *Chem Phys Lett* 248:71–76, DOI [10.1016/0009-2614\(95\)01281-8](https://doi.org/10.1016/0009-2614(95)01281-8)
5. Kevorkyants R, Dulak M, Wesolowski TA (2006) *J Chem Phys* 124:024104, DOI [10.1063/1.2150820](https://doi.org/10.1063/1.2150820)
6. Wesolowski TA, Tran F (2003) *J Chem Phys* 118:2072–2080, DOI [10.1063/1.1534090](https://doi.org/10.1063/1.1534090)
7. Tran F, Weber J, Wesolowski TA (2001) *Helv Chim Acta* 84:1489–1503, DOI [10.1002/1522-2675\(20010613\)84:6;1489::AID-HLCA1489;3.0.CO;2-D](https://doi.org/10.1002/1522-2675(20010613)84:6;1489::AID-HLCA1489;3.0.CO;2-D)
8. Wesolowski TA, Morgantini PY, Weber J (2002) *J Chem Phys* 116:6411–6421
9. Zhao Y, Truhlar DG (2005) *J Chem Theory Comput* 1:415–432, DOI [10.1021/ct049851d](https://doi.org/10.1021/ct049851d)
10. Zhao Y, Truhlar DG (2005) *J Phys Chem A* 109:5656–5667, DOI [10.1021/jp050536c](https://doi.org/10.1021/jp050536c)
11. Zhao Y, Truhlar DG (2005) *Phys Chem Chem Phys* 7:2701–2705, DOI [10.1039/b507036h](https://doi.org/10.1039/b507036h)
12. Dirac PAM (1930) *Proc Camb Philos Soc* 26:376
13. Vosko SH, Wilk L, Nusair M (1980) *Can J Phys* 58:1200–1211
14. Ceperley DM, Alder BJ (1980) *Phys Rev Lett* 45:566–569, DOI [10.1103/PhysRevLett.45.566](https://doi.org/10.1103/PhysRevLett.45.566)
15. Thomas LH (1927) *Proc Camb Philos Soc* 23:542
16. Fermi E (1928) *Z Phys* 48:73
17. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C (1992) *Phys Rev B* 46:6671–6687, DOI [10.1103/PhysRevB.46.6671](https://doi.org/10.1103/PhysRevB.46.6671)
18. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C (1993) *Phys Rev B* 48:4978–4978, DOI [10.1103/PhysRevB.48.4978.2](https://doi.org/10.1103/PhysRevB.48.4978.2)
19. Lembarki A, Chermette H (1994) *Phys Rev A* 50:5328–5331, DOI [10.1103/PhysRevA.50.5328](https://doi.org/10.1103/PhysRevA.50.5328)
20. Wesolowski TA, Chermette H, Weber J (1996) *J Chem Phys* 105:9182–9190, DOI [10.1063/1.472823](https://doi.org/10.1063/1.472823)
21. Dulak M, Wesolowski TA (2005) *Int J Quantum Chem* 101:543–549, DOI [10.1002/qua.20309](https://doi.org/10.1002/qua.20309)
22. Köster AM, Flores-Moreno R, Geudtner G, Goursot A, Heine T, Reveles JU, Vela A, Salahub DR (2003) deMon 2003, NRC, Canada. <http://www.deMon-software.com/>
23. Gill PMW, Johnson BG, Pople JA (1993) *Chem Phys Lett* 209:506–512, DOI [10.1016/0009-2614\(93\)80125-9](https://doi.org/10.1016/0009-2614(93)80125-9)
24. Kendall RA, Dunning TH, Harrison RJ (1992) *J Chem Phys* 96:6796–6806, DOI [10.1063/1.462569](https://doi.org/10.1063/1.462569)
25. Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 02/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller or Karen Schuchardt for further information. <http://www.emsl.pnl.gov/forms/basisform.html>
26. Lynch BJ, Zhao Y, Truhlar DG (2003) *J Phys Chem A* 107:1384–1388, DOI [10.1021/jp021590l](https://doi.org/10.1021/jp021590l)
27. Wesolowski TA (1997) *J Chem Phys* 106:8516–8526, DOI [10.1063/1.473907](https://doi.org/10.1063/1.473907)
28. Köster AM, Calaminici P, Escalante S, Flores-Moreno R, Goursot A, Patchkovskii S, Reveles JU, Salahub DR, Vela A, Heine T (2004) The deMon user's guide, version 1.0.3, 2003–2004. <http://www.deMon-software.com/>
29. Boys SF, Bernardi F (1970) *Mol Phys* 19:553
30. Xantheas SS (1996) *J Chem Phys* 104:8821–8824, DOI [10.1063/1.471605](https://doi.org/10.1063/1.471605)
31. Jurecka P, Sponer J, Hobza P (2004) *J Phys Chem B* 108:5466–5471, DOI [10.1021/jp049956c](https://doi.org/10.1021/jp049956c)