



HAL
open science

Non-Covalent Interactions using Local Correlation Methods: Energy Partitioning, Geometry Optimisation and Harmonic Frequency Calculations

Jamie Platts, J. Grant Hill

► **To cite this version:**

Jamie Platts, J. Grant Hill. Non-Covalent Interactions using Local Correlation Methods: Energy Partitioning, Geometry Optimisation and Harmonic Frequency Calculations. *Molecular Physics*, 2010, 108 (11), pp.1497-1504. 10.1080/00268971003757977 . hal-00602607

HAL Id: hal-00602607

<https://hal.science/hal-00602607>

Submitted on 23 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Non-Covalent Interactions using Local Correlation Methods:
Energy Partitioning, Geometry Optimisation and Harmonic
Frequency Calculations**

Journal:	<i>Molecular Physics</i>
Manuscript ID:	TMPH-2010-0047.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	04-Mar-2010
Complete List of Authors:	Platts, Jamie; Cardiff University, Chemistry Hill, J.; Washington State University, Chemistry
Keywords:	Local correlation, Non-covalent interactions, Hydrogen bonding, Pi-stacking, Zero point energy



Non-Covalent Interactions using Local Correlation Methods: Energy Partitioning, Geometry Optimisation and Harmonic Frequency Calculations

James A. Platts* and J. Grant Hill[†]

School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, UK

[†] Present address: Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA.

* Author for correspondence:

Phone: +44-2920-874950

FAX: +44-2920-874030

E-mail: platts@cardiff.ac.uk

Abstract

The performance of density fitting, local correlation methods (DF-LMP2 and DF-LCCSD) in studies of non-covalent interactions is tested against literature data for a standard set of 22 intermolecular complexes. Partitioning of interaction energy in the local correlation approach, based on the classes of occupied and virtual orbital involved in the interaction, clearly distinguishes the three types of interaction present in the set of complexes, in agreement with previous classifications. Geometry optimisation is found to be straightforward with DF-LMP2 without the need for counterpoise correction, resulting in geometries very close to previous, counterpoise-corrected structures. Spin-component scaling of gradients to correct for the known shortcomings of conventional MP2 has only a small effect on geometries in most cases, but significantly alters the distance between aromatic rings in stacked complexes. Harmonic frequency calculation is made possible by efficient use of parallel computing resources, and confirms all structures to be true minima, unlike previous estimates using density functional theory. Corrections for the change in zero-point vibrational energy are determined from this data, and typically constitute between 10 and 50% of the overall binding energy of the complex.

Introduction

Non-covalent interactions play an important part in a wide range of chemical and biological phenomena, for instance folding of proteins and nucleic acids into their correct 3-dimensional shapes.¹ Intermolecular forces acting within molecular clusters are widely used as models of such non-covalent interactions, since their effects can be more easily separated from intra-molecular forces. *Ab initio* theoretical methods are increasingly used to complement experimental study of such clusters. The majority of theoretical treatments of non-covalent interactions employ the supermolecular approach, in which interaction energy is calculated as the difference in energy between the cluster and its constituent monomers. However, two significant drawbacks of this approach are the need for large orbital basis sets, including polarisation and diffuse functions, for proper description of all interactions between molecules, and the related problem of basis set superposition error (BSSE).²

These problems can be ameliorated, to a large extent, by use of local correlation methods, which effectively eliminate BSSE by restricting excitations from occupied orbitals to those virtual orbitals that are in close spatial proximity.³ Moreover, in combination with the density fitting (DF) approximation, methods such as DF-LMP2⁴ and DF-LCCSD^{5,6,7,8} make significant efficiency savings over conventional calculations, approaching linear scaling in favourable cases, hence allowing application of suitable orbital basis sets for description of non-covalent interactions. A further advantage of the local correlation approach is that intermolecular correlation energy can be partitioned into physically meaningful contributions, such as dispersion and ionic terms, on the basis of the occupied and virtual orbitals involved.³ This approach is complementary to decomposition schemes, and is based solely on the local correlation treatment of the cluster and so requires no further calculations or significant computational resources.

BSSE can also be accounted for by use of the counterpoise method,⁹ but for a cluster of n molecules, this procedure requires $(n+1)$ energy calculations in the full basis set of the cluster, substantially increasing the computational resources required. The situation becomes still less satisfactory if geometry optimisation of the cluster is a goal, as separate gradient calculations for the cluster and all fragments are necessary. Without such corrections, geometry optimisation is likely to result in underestimation of intermolecular separations and overestimation of stabilities. Alternative approaches, such as symmetry adapted perturbation theory (SAPT) eliminate BSSE by construction, but are currently limited to pairs of molecules.¹⁰

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Molecular clusters are dynamic objects, and understanding of their vibrational properties is key to understanding their thermodynamic and spectroscopic properties. The near-universal starting point for theoretical description of vibrational behaviour is the harmonic approximation, which, despite its known shortcomings, is widely used to interpret thermodynamic and spectroscopic properties of interest. The efficiency and lack of BSSE in local correlation methods is again a significant advantage in calculating the necessary data. Recently, Hrenar *et al.* described the use of DF-LMP2 methods for numerical evaluation of the Hessian using finite differences of analytical gradients, made possible by using a highly parallelised algorithm.¹¹ In doing so, they demonstrated the utility of this approach, and identified scaling factors that correct for at least some of the deficiencies of the harmonic approach.

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
Grimme's spin-component scaling (SCS) method,¹² and variants thereof, have been shown to correct for deficiencies in MP2 for stacked complexes, but may actually degrade performance for hydrogen bonded complexes.¹³ Recently, Marchetti *et al.* proposed a dispersion-weighted SCS scheme that combines the accuracy of conventional MP2 for electrostatic complexes and of SCS-MP2 for dispersion bound ones.¹⁴ In the current investigation, density fitted, local correlation methods are applied to a range of intermolecular interactions, encompassing complexes bound mainly by electrostatic effects, by dispersion forces, or by a combination of these. These complexes are taken from the recent review of Hobza and co-workers,¹⁵ who suggest that they are representative of most important classes of non-covalent interaction. DF-LMP2, and DF-LCCSD where possible, methods are used to partition the interaction energy, to optimise geometries, and to calculate harmonic frequencies. In doing so, the aim is to show that DF-LMP2 and its SCS variant are capable of accurately and efficiently predicting the geometry and harmonic vibrational motion of complexes bound by non-covalent interactions, paving the way for studies of larger systems for which benchmark data is not available.

50 51 52 53 54 55 56 57 58 59 60 **Computational Methods**

All calculations were performed with the MOLPRO suite of programs,¹⁶ employing the correlation consistent aug-cc-pVTZ orbital basis set throughout.^{17,18} Density fitting of MP2 energies employed the corresponding MP2FIT basis set,¹⁹ while the underlying Hartree-Fock calculation also employed density fitting²⁰ with the corresponding JKFIT basis.²¹ DF in LCCSD for integrals over three and four external orbitals used the larger aug-cc-pVQZ/MP2FIT auxiliary basis sets to minimize fitting errors. The localised orbitals required for local correlation methods were generated via the Pipek-Mezey method,²² while the orbital domain selection followed the

1
2
3 procedure of Boughton and Pulay,²³ with merging of rotationally invariant π -domains where
4 appropriate. Initial test calculations indicated that using domains calculated at large intermolecular
5 separation and then frozen for calculations on the interacting system produced almost negligible
6 differences in geometry and harmonic frequencies when compared with domains computed solely
7 for the interacting system (so-called floating domains). For example, with the small methane dimer
8 system a maximum difference of 6 cm^{-1} between normal modes was observed, while for the larger
9 parallel-displaced benzene dimer system the maximum difference was 0.2 cm^{-1} . As deviations of
10 this size are likely to be small compared with both the basis set and method errors, floating
11 domains were used for all geometry optimisations and calculation of numerical harmonic
12 frequencies.
13
14
15
16
17
18
19
20
21

22
23 Analytical gradients for DF-LMP2 methods²⁴ are available in MOLPRO, enabling efficient
24 geometry optimisation. All geometry optimisations reported commenced from the reported
25 literature structure, and employed default criteria for convergence. Calculation of root mean square
26 (RMS) deviation between initial and final structures used Chemcraft.²⁵ Following geometry
27 optimisation, harmonic frequencies were calculated by numerical differentiation of analytic
28 gradients using central differences, employing the default step size of 0.01 a.u. This is a sizeable
29 task for the larger molecules considered, but was made feasible by use of a parallel algorithm,²⁶ as
30 implemented in MOLPRO, in which individual energy calculations on different geometry
31 displacements used in the formation of gradients and Hessians are performed independently on
32 separate processors. Individual DF-LMP2 calculations did not take account of molecular symmetry
33 (as this usually leads to symmetry equivalent localised orbitals), but this numerical algorithm can
34 use symmetry to reduce the number of displacements required. The local character of the orbitals
35 allows for the partitioning of the intermolecular interaction energy into intramolecular, dispersive,
36 and ionic components of the correlation energy.³ This is achieved by dividing the system into
37 monomers and separating the excitations from localised orbitals into different classes based on the
38 domains the pair of electrons are excited into.
39
40
41
42
43
44
45
46
47
48
49
50
51

52 53 **Results and Discussion**

54 Table 1 presents the results of partitioning the DF-LMP2 correlation energy into its intermolecular
55 components, along with the interaction energy at the SCF level, for the S22 set of complexes of
56 Hobza and co-workers.¹⁵ Following reference 3, dispersion refers to simultaneous excitations on
57 different monomers, ionic to excitation from an occupied orbital on one monomer to a virtual
58
59
60

orbital on the other along with a single excitation on the acceptor monomer, and dispersion-exchange to cross-excitations. Such definitions stem naturally from the local correlation approach, but may not be directly comparable to assignments from methods such as SAPT.¹⁰ Intramolecular correlation can also be calculated in this fashion, but does not show significant differences from the sum of monomer correlation energies. The three intermolecular correlation terms are taken directly from the DF-LMP2 cluster calculation, while the SCF interaction energy was calculated using the supermolecular approach. In all cases, dispersion-exchange energies are essentially negligible, and are included only for completeness.

Table 1 SCF interaction energies and partitioned DF-LMP2 intermolecular energies (kcal mol⁻¹)

	SCF	Dispersion	Ionic	Dispersion-Exchange
Ammonia dimer	-1.27	-1.23	-0.77	-0.01
Water dimer	-3.72	-1.06	-1.23	+0.01
Formic acid dimer	-15.49	-4.20	-5.63	+0.01
Formamide dimer	-12.32	-3.73	-4.09	+0.08
Uracil dimer HB	-9.31	-4.82	-4.79	+0.07
2-Pyridoxine...2-aminopyridine	-5.81	-5.52	-4.76	-0.03
Adenine...Thymine HB	-6.40	-5.68	-4.79	+0.01
Methane dimer	+0.35	-0.73	-0.14	0.00
Ethene dimer	+0.79	-1.88	-0.60	+0.01
Benzene...methane	+1.60	-2.19	-0.86	+0.02
Benzene dimer stack	+6.33	-7.22	-2.78	+0.02
Pyrazine dimer	+7.60	-7.77	-3.11	+0.03
Uracil dimer stack	+5.20	-9.15	-3.13	+0.01
Indole...benzene stack	+8.62	-10.60	-4.46	+0.11
Adenine...Thymine stack	+5.72	-14.05	-5.42	+0.05
Ethene...ethyne	-0.52	-1.01	-0.58	0.01
Benzene...water	-0.35	-2.01	-1.05	+0.01
Benzene...ammonia	+0.74	-2.10	-0.95	+0.01
Benzene...cyanide	-1.55	-2.75	-1.63	+0.04
Benzene dimer T	+2.54	-3.84	-1.57	+0.05
Indole...benzene T	+2.00	-5.13	-2.55	+0.05

Phenol dimer	+0.43	-4.72	-2.14	+0.02
--------------	-------	-------	-------	-------

The first seven clusters, denoted as hydrogen bonded in the original work, all display significant stabilisation at the SCF level, reflecting the importance of electrostatics in such interactions. Nonetheless, dispersion and ionic correlation effects contribute significantly to the overall stabilisation, with approximately equal importance. In all seven cases, the SCF stabilisation is the largest single term, but in three (ammonia dimer, 2-pyridoxine...2-aminopyridine and adenine...thymine) the combined correlation terms outweigh the SCF contribution. For example, in Watson-Crick hydrogen bonded adenine...thymine the SCF, dispersion and ionic contributions are 39%, 35% and 29% of the overall binding energy, respectively.

The next eight complexes were denoted as predominantly bound through dispersion,¹⁵ an assignment supported by the present analysis. In all eight clusters, SCF energies are repulsive, with all stabilisation stemming from correlation effects. Unlike the hydrogen bonded complexes, dispersion effects are much greater (typically by a factor of two to five) than ionic ones in these complexes, more than counteracting the repulsive SCF term. This is exemplified by the archetypal stacked, parallel-displaced benzene dimer, for which dispersion effects are 2.5 times greater than ionic ones. The final seven complexes are of so-called “mixed” type, with no single contribution dominating. This designation is again supported by the current analysis; SCF energies are attractive in three cases and repulsive in the remaining four. Dispersion contributions are generally larger than ionic ones, but not to the same extent as for the dispersion-bound clusters. Complexes that could be designated as containing X—H... π (X = C, N, O) non-covalent bonds, such as benzene...water, benzene...ammonia and benzene...HCN, show significant differences from the conventional hydrogen bonds. Most notably, dispersion contributions are significantly greater than SCF ones. The clear contrast in the origin of stabilisation between the three classes of complex is illustrated in Figure 1.

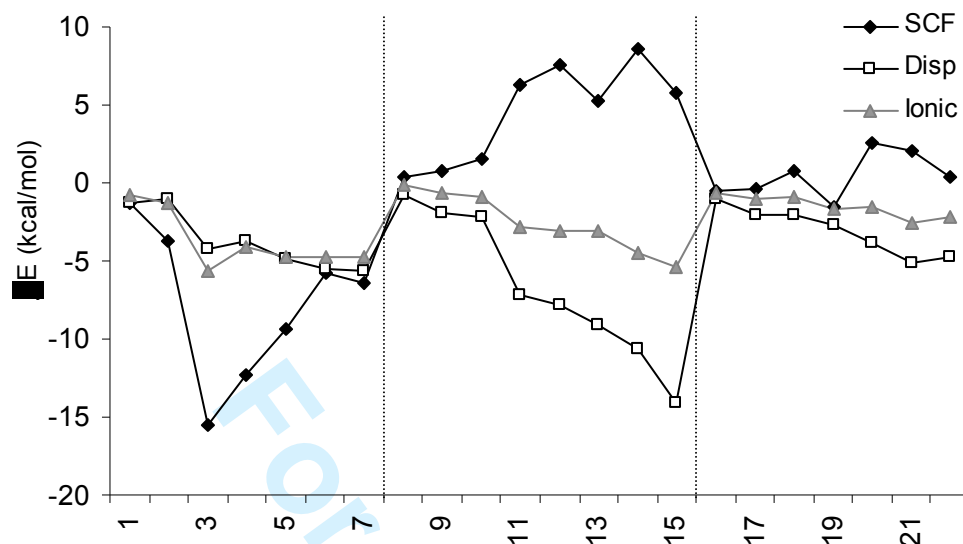


Figure 1 Illustration of energy partitioning data from Table 1. Vertical dashed lines show separation of classes of complex used in ref. 15. The numbering on the abscissa corresponds to the ordering of the complexes in Table 1.

The analysis of Table 1 is based on MP2 methods, known to perform adequately for H-bonding but to substantially over-estimate interaction energies of stacked complexes. Analogous partitioning of correlation energy with the DF-LCCSD method (energy partitioning of perturbative triples contributions is not currently possible) has also been performed. For inclusion of all relevant electron pairs in the LCCSD treatment, modification of the default distance based domain selection criteria is required: after some trial and error, all orbital pairs within a distance of 8 a.u. are treated as strong pairs, *i.e.*, at the CCSD level. Unfortunately, this renders DF-LCCSD calculations of the larger complexes unfeasible with a moderately large basis set such as aug-cc-pVTZ.

The partitioning of the DF-LCCSD correlation energy is presented in Table 2 for selected, smaller complexes, and shows that the trends observed using DF-LMP2 methods are largely preserved at this improved level of theory. For dispersion bound and mixed complexes, the tendency of MP2 to overestimate interaction energies is evident in both dispersion and ionic terms, with rather larger reductions in the former than the latter. This is exemplified in data for the ethene dimer, for which dispersion contributions are reduced by $0.1 \text{ kcal mol}^{-1}$ and ionic by $0.02 \text{ kcal mol}^{-1}$. This compares well with the $\Delta\text{CCSD(T)}$ correction to the overall interaction of $+0.09$ reported in ref 15.

Table 2 Partitioned DF-LCCSD interaction energies for selected complexes (kcal mol⁻¹)

Name	Dispersion	Ionic	Dispersion-Exchange
Ammonia dimer	-1.29	-0.79	-0.01
Water dimer	-1.20	-1.24	-0.03
Methane dimer	-0.72	-0.16	0.00
Ethene dimer	-1.78	-0.58	0.00
Benzene...methane	-2.03	-0.82	+0.01
Ethene...ethyne	-0.95	-0.54	0.00
Benzene...water	-1.90	-0.99	-0.01
Benzene...ammonia	-1.96	-0.85	-0.01

This analysis was performed at the literature geometries, which were obtained with either MP2 or CCSD(T) optimisation with triple- or quadruple- ζ basis sets and including counterpoise correction. This can be a laborious process, and might be significantly eased by the advantages of the DF-LMP2 method discussed above, particularly the removal of BSSE. Table 3 reports the results of full, unconstrained geometry optimisation of the same S22 complexes using DF-LMP2/aug-cc-pVTZ, starting from the literature geometry. Proper application of local correlation methods requires symmetry to be disabled, but the optimisation algorithm retains any elements of symmetry present in the initial geometry.

Optimisation generally proceeded smoothly, in most cases reaching the default convergence criteria in MOLPRO (maximum gradient below 3×10^{-4} a.u.) in less than 10 steps. Table 3 demonstrates that the resulting optimised geometries are very similar to their literature equivalents in all but one case, that being the phenol dimer. Performance for hydrogen bonded and dispersion bound complexes is similar, whereas the slightly larger RMS deviation for the mixed complexes is almost entirely due to the phenol dimer (omitting this point gives a mean RMS of 0.018 Å for the remaining six complexes). No significant difference in the quality of agreement is observed between MP2 and CCSD(T) optimised literature geometries. Figure 2 shows the literature and optimised geometries of the phenol dimer, which by inspection are highly similar. The hydrogen bond O...H distances in the two cases are almost identical (1.937 vs. 1.935 Å). However, the dihedral angle between the mean planes of each molecule are different, with values of 60.5° in the literature geometry and 56.3° in the DF-LMP2 optimised structure. Thus, DF-LMP2 brings the

aromatic rings into slightly closer contact (distance between centroids is 5.12 Å vs. 5.39 Å in the literature structure). This subtle change notwithstanding, the overall conclusion is that this efficient and counterpoise-free approach to geometry optimisation satisfactorily and consistently reproduces conventional, corrected geometries.

Table 3 RMS deviations from literature geometry (Å)

Name	DF-LMP2	SCSN
Ammonia dimer ^a	0.025	0.027
Water dimer ^a	0.030	0.037
Formic acid dimer ^a	0.017	0.016
Formamide dimer ^a	0.025	0.028
Uracil dimer HB	0.008	0.026
2-Pyridoxine...2-aminopyridine	0.009	0.034
Adenine...Thymine HB	0.011	0.028
Methane dimer ^a	0.005	0.050
Ethene dimer ^a	0.017	0.049
Benzene...methane	0.005	0.039
Benzene dimer stack	0.029	0.068
Pyrazine dimer	0.021	0.048
Uracil dimer stack	0.023	0.053
Indole...benzene stack	0.017	0.039
Adenine...Thymine stack	0.018	0.035
Ethene...ethyne ^a	0.014	0.017
Benzene...water	0.026	0.012
Benzene...ammonia	0.019	0.021
Benzene...cyanide	0.019	0.016
Benzene dimer T	0.015	0.027
Indole...benzene T	0.012	0.023
Phenol dimer	0.183	0.173
Mean	0.025	0.039
Mean H-bond	0.018	0.028
Mean dispersion	0.017	0.048

Mean Mixed	0.041	0.041
Mean CCSD(T) ^b	0.019	0.032

^a Literature geometry from CP-corrected CCSD(T)/AVQZ optimisation; all others from CP-corrected MP2/AVTZ. ^b For literature geometries optimised with CCSD(T) only.

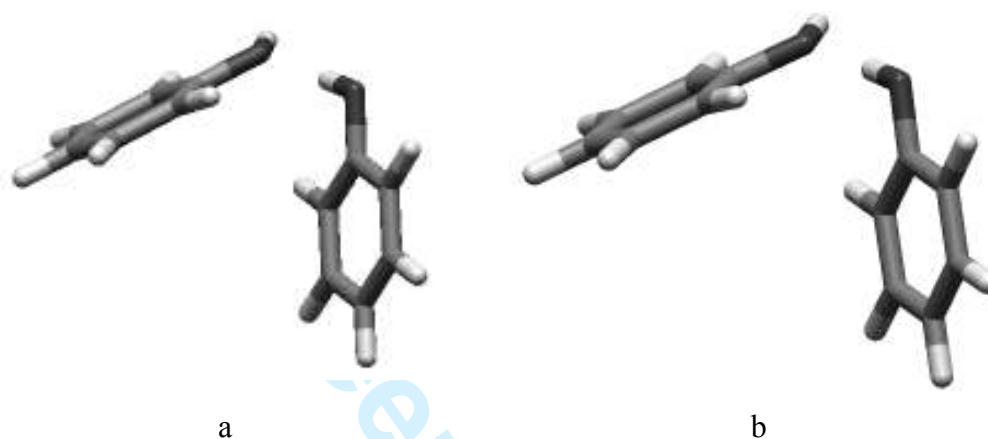


Figure 2 Phenol dimer a) literature geometry; b) DF-LMP2 optimised.

Both conventional MP2 and DF-LMP2 are well known to overestimate the interaction energy of stacked complexes, and might therefore be expected to underestimate intermolecular separation in such complexes. This is apparent in potential energy curves for the stacked benzene dimer, which reach a minimum at vertical separation of *ca.* 3.5 Å at CBS(T) level and *ca.* 3.35 Å at MP2.²⁷ Recently, it was demonstrated that a new parameterisation of SCS designed specifically for non-covalent interactions and denoted SCSN,^{13a} overcomes this deficiency, resulting in high-accuracy interaction energies for all classes of complex with errors of less than 0.5 kcal mol⁻¹. In this approach, correlation energy from same- and opposite-spin electron pairs are weighted by scaling factors in order to correct for systematic errors in standard MP2. The same scaling factors can be applied to calculation of analytic gradients, allowing optimisation to minima on this scaled potential energy surface. Table 3 also reports the RMS deviations from literature geometry using this approach.

In general, SCSN optimised structures are slightly further from literature values than unscaled DF-LMP2, both for MP2 and CCSD(T) reference data. Overall performance is similar to unscaled optimisation for hydrogen bonded and mixed complexes, but differences are much larger for dispersion bound complexes. The relatively small RMS deviations reported in Table 3 mask some

rather more significant changes; for instance, in the stacked benzene dimer, the separation of mean planes is 3.361 Å in the literature (MP2/TZ-CP optimised) geometry, whereas our SCSN optimisation results in a separation of 3.487 Å, much closer to the minimum of the CBS(T) curve.²⁷ A similar trend is apparent in other stacked complexes, albeit to a lesser extent: in the pyrazine dimer, the molecular planes are separated by 3.261 Å in literature geometry and 3.270 Å in the SCSN geometry. In stacked adenine-thymine, molecular planes are not quite parallel, so for the current investigation the perpendicular distance is defined as that from the mean plane of adenine to the centroid of the 6-membered ring of thymine, which is 3.174 Å in literature and 3.205 Å in SCSN optimised geometry. Thus, it is evident that the excess stabilisation of stacked complexes at MP2 level carries through into potential energy surfaces, and can be simply and efficiently corrected by the empirical SCSN scheme.

A second SCSN geometry optimisation was performed for the T-shaped benzene dimer, following the recent study by Wang *et al.* that showed the lowest energy geometry of this complex to deviate from the ideal C_{2v} geometry.²⁸ Starting from a slightly offset geometry, optimisation proceeded smoothly to a point just 0.06 kcal mol⁻¹ lower in energy than that reported in Table 3. Shown in Figure 3, the SCSN geometry is similar to that reported in ref.28, with the donor C—H directed towards a C—C bond rather than the centre of the acceptor ring.

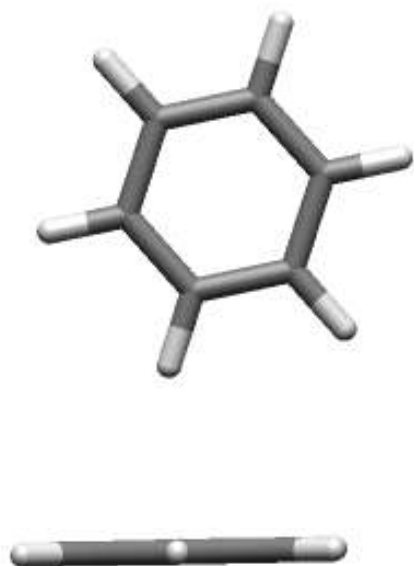


Figure 3 SCSN optimised geometry of T-shaped benzene dimer

The fully optimised DF-LMP2 geometries were then used to calculate harmonic frequencies using numerical differentiation. Firstly, and most importantly, this procedure confirms that all but one

1
2
3
4 optimised structures are true minima at this theoretical level, as confirmed by the lack of imaginary
5 frequencies. As noted above, any molecular symmetry present in the literature geometry was
6 retained in these calculations, such that our data lends support to the symmetries employed in the
7 original study of Hobza and co-workers. This is in marked contrast to a recent DFT-based study of
8 the same set of molecules, in which one or more imaginary frequencies were observed for several
9 complexes.²⁹ The only exception is the T-shaped benzene dimer, for which the high-symmetry
10 (C_{2v}) form is found to be a transition state, with a single imaginary frequency of -10.8 cm^{-1} . Re-
11 optimisation with lower (C_s) symmetry results in a geometry similar to that shown in Figure 3,
12 which is found to be a true minimum on the potential energy surface.
13
14
15
16
17
18
19
20

21 A similar procedure was followed for the SCSN-optimised geometries. In this case, both the C_{2v}
22 and C_s optimised forms of the T-shaped benzene dimer are found to be true minima, echoing the
23 findings of Wang *et al.*,²⁸ who used dispersion-corrected DFT to show that both forms are stable
24 and are connected by a transition state with a barrier of just $0.01 \text{ kcal mol}^{-1}$. In this regard, SCSN
25 again seems to be slightly more accurate than unscaled DF-LMP2 in finding both forms to be
26 stable, but the inherent errors associated with this approach are much larger than these energy
27 differences, such that any conclusion based on this data must be tentative at best. All remaining
28 structures are found to be minima using this approach. Two geometries of the stacked (parallel-
29 displaced) benzene dimer were tested, one with a C—H bond situated over the partner benzene and
30 a second with a C—C bond over the other benzene. Both are found to be minima using SCSN,
31 with effectively identical energies within the expected accuracy of the method (less than 0.1 kcal
32 mol^{-1} difference).
33
34
35
36
37
38
39
40
41
42

43
44 As well as confirming structures as minima, harmonic frequency calculations can be used to
45 estimate the change in zero-point vibrational energy (ZPVE) that accompanies formation of a
46 complex from constituent monomers. Such changes stem largely from conversion of translational
47 and rotational degrees of freedom of individual molecules into intermolecular vibrations. Such
48 vibrations are likely to be strongly anharmonic in nature, such that harmonic data can only be a
49 first-order approximation to the true ZPVE correction. Nonetheless, obtaining even harmonic
50 vibrational data for the larger complexes considered here is a considerable challenge, and the data
51 presented in Table 4 is a step towards establishing accurate ZPVE corrections for systems of this
52 size.
53
54
55
56
57
58
59
60

Table 4 Harmonic zero point energy corrections for binding (kcal mol⁻¹)

Name	DF-LMP2	Scaled ^a	SCSN
Ammonia dimer	1.70	1.62	1.66
Water dimer	2.37	2.27	2.33
Formic acid dimer	2.08	1.99	2.06
Formamide dimer	2.64	2.53	2.52
Uracil dimer HB	1.21	1.16	1.15
2-Pyridoxine...2-aminopyridine	1.01	0.97	0.98
Adenine...thymine HB	0.95	0.90	0.89
Methane dimer	1.16	1.11	1.14
Ethene dimer	0.83	0.80	0.82
Benzene...methane	0.64	0.62	0.70
Benzene dimer stack	0.53	0.51	0.56
Pyrazine dimer	0.61	0.58	0.51
Uracil dimer stack	0.58	0.56	0.56
Indole...benzene stack	0.25	0.24	0.43
Adenine...thymine stack	0.72	0.68	0.61
Ethene...ethyne	0.71	0.68	0.73
Benzene...water	1.06	1.01	1.04
Benzene...ammonia	0.70	0.67	0.79
Benzene...cyanide	0.60	0.57	0.71
Benzene dimer T	0.52	0.49	0.67
Indole...benzene T	0.62	0.60	0.59
Phenol dimer	0.95	0.91	0.99

^a Scaled from raw DF-LMP2 data using the recommended scaling factor of 0.956

Δ ZPVE values on formation of complexes from monomers are largest for the hydrogen bonded complexes, and are in general rather smaller for dispersion bound and mixed complexes. The single largest change is found for the formamide dimer, closely followed by water and formic acid dimers. The value for the water dimer is similar to previous estimates using MP2 methods with the harmonic approximation (2.1 – 2.3 kcal mol⁻¹), and as expected rather larger than values that take anharmonicity into account (1.8 – 2.1 kcal mol⁻¹).³⁰ The larger aromatic systems containing two strong hydrogen bonds show much smaller Δ ZPVE, falling to *ca.* 1 kcal mol⁻¹ for 2-pyridoxine...2-

1
2
3 aminopyridine and adenine...thymine. Values for dispersion bound and mixed complexes are in
4 general rather smaller than for hydrogen bonded ones. For example, $\Delta ZPVE$ for the stacked uracil
5 dimer is less than half the value for the hydrogen bonded isomer, with a similar pattern for the
6 adenine-thymine complexes. A slightly surprising exception to this pattern is the methane dimer,
7 whose $\Delta ZPVE$ value of $1.16 \text{ kcal mol}^{-1}$ is more than twice the binding energy. This may represent
8 the breakdown of the harmonic approximation, or be a result of noise introduced by the use of
9 numerical differences in such a weakly bound case.
10
11
12
13
14
15
16
17

18 In addition to raw DF-LMP2 data, Table 4 also contains $\Delta ZPVE$ values scaled by the value of
19 0.956 recommended by Hrenar *et al.*,¹¹ and values from numerical frequency calculation based on
20 SCSN energies. Changes using the single scaling factor are generally very small, and do not
21 significantly alter the conclusions drawn from unscaled data. A similar pattern emerges from most
22 of the SCSN data, with one exception: in the stacked indole-benzene complex, the SCSN value is
23 almost double that from unscaled DF-LMP2. In the absence of benchmark data, it is not possible to
24 decide which of these values is the more reliable, but in the context of other stacked complexes the
25 raw DF-LMP2 value seems rather small.
26
27
28
29
30
31
32
33

34 Conclusions

35 The utility of local correlation methods, when combined with density fitting, has been
36 demonstrated in the study of non-covalent interactions. By construction, these methods allow for
37 decomposition of correlation energy into contributions from dispersion and ionic excitations, thus
38 giving insight into the origin of stabilisation at no additional computational cost to the correlation
39 calculation. This analysis, at both DF-LMP2 and DF-LCCSD levels, clearly distinguishes
40 complexes bound by hydrogen bonding, dispersion, or a mixture of these effects. The lack of basis
41 set superposition error in local methods, coupled with their computational efficiency, makes them
42 ideally suited for simple geometry optimisation and harmonic frequency calculation. Geometry
43 optimisation leads to structures in close agreement with published, counterpoise-corrected
44 geometries, demonstrating the utility of this approach. Harmonic frequency analysis, using both
45 unscaled and SCSN-scaled MP2 data, indicates that most complexes within the S22 set are at least
46 local minima on their respective potential energy surfaces. The only exception to this conclusion is
47 the C_{2v} T-shaped benzene dimer, which is a minimum with SCSN scaling but a saddle point with
48 unscaled MP2. Zero-point vibrational energy corrections to binding energies have also been
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

estimated, using raw data and two distinct scaling methods. In general, all three methods are in reasonable agreement, with corrections of between 20% and 50% of uncorrected binding energies.

Acknowledgements

This work was funded by the UK EPSRC, grant ref. EP/C013328/1.

References

- ¹ A. Pavlov, P.M. Mitrasinovic, *Curr. Org. Chem.*, 2010, **14**, 129.
- ² P. Hobza, H.L. Selzle, E.W. Schlag, *J. Phys. Chem.*, 1996, **100**, 18790.
- ³ M. Schütz, G. Rauhut, H.-J. Werner, *J. Phys. Chem. A*, 1998, **102**, 5997.
- ⁴ H.-J. Werner, F. R. Manby, P. J. Knowles, *J. Chem. Phys.*, 2003, **118**, 8149.
- ⁵ C. Hampel, H.-J. Werner, *J. Chem. Phys.*, 1996, **104**, 6286.
- ⁶ M. Schütz, H.-J. Werner, *J. Chem. Phys.*, 2001, **114**, 661.
- ⁷ M. Schütz, *Phys. Chem. Chem. Phys.*, 2002, **4**, 3941.
- ⁸ M. Schütz, F. R. Manby, *Phys. Chem. Chem. Phys.*, 2003, **5**, 3349.
- ⁹ S. F. Boys, F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
- ¹⁰ B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887.
- ¹¹ T. Hrenar, G. Rauhut, H.-J. Werner, *J. Phys. Chem. A*, 2006, **110**, 2060.
- ¹² S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095.
- ¹³ a) J.G. Hill, J.A. Platts, *J. Chem. Theor. Comput.*, 2007, **3**, 80. b) R.A. Bachorz, F. A. Bischoff, S. Hofener, W. Klopper, P. Ottiger, R. Leist, J. A. Frey, S. Leutwyler, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2758.
- ¹⁴ O. Marchetti and H.-J. Werner, *J. Phys. Chem. A*, 2009, **113**, 11580.
- ¹⁵ P. Jurečka, J. Šponer, J. Černý, P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985.
- ¹⁶ MOLPRO, version 2008.1, a package of *ab initio* programs, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz and others, see <http://www.molpro.net>.
- ¹⁷ T. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007.
- ¹⁸ R. A. Kendall, T. Dunning Jr., R. J. Harrison, *J. Chem. Phys.*, 1991, **96**, 6797.
- ¹⁹ F. Weigend, A. Köhn, C. Hättig, *J. Chem. Phys.*, 2002, **116**, 3175.
- ²⁰ R. Polly, H.-J. Werner, F. R. Manby, P. J. Knowles, *Mol. Phys.*, 2004, **102**, 2311.
- ²¹ F. Weigend, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4285.
- ²² J. Pipek, P. G. Mezey, *J. Chem. Phys.*, 1989, **90**, 4916.
- ²³ J. W. Boughton, P. Pulay, *J. Comput. Chem.*, 1993, **14**, 736.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
-
- ²⁴ M. Schütz, H.-J. Werner, R. Lindh, F. R. Manby, *J. Chem. Phys.*, 2004, **121**, 737.
- ²⁵ Chemcraft v1.6: www.chemcraftprog.com
- ²⁶ J. Neugebauer, M. Reiher, C. Kind, B.A. Hess, *J. Comput. Chem.*, 2002, **23**, 895.
- ²⁷ a) S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J. Am. Chem. Soc.*, 2000, **122**, 3746. b) M. O. Sinnokrot, C. D. Sherrill, *J. Am. Chem. Soc.*, 2004, **126**, 7690. c) J.G. Hill, J.A. Platts, H.-J. Werner, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4072.
- ²⁸ W. Wang, M. Pitoňák, P. Hobza, *ChemPhysChem*, 2007, **8**, 2107.
- ²⁹ K. Gkionis, J.G. Hill, S.P. Oldfield, J.A. Platts, *J. Mol. Model.*, 2009, **15**, 1051.
- ³⁰ M. W. Feyereisen, D. Feller, D.A. Dixon, *J. Phys. Chem.*, 1996, **100**, 2993.

Non-Covalent Interactions using Local Correlation Methods: Energy Partitioning, Geometry Optimisation and Harmonic Frequency Calculations

James A. Platts* and J. Grant Hill[†]

School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, UK

[†] Present address: Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA.

* Author for correspondence:

Phone: +44-2920-874950

FAX: +44-2920-874030

E-mail: platts@cardiff.ac.uk

Abstract

The performance of density fitting, local correlation methods (DF-LMP2 and DF-LCCSD) in studies of non-covalent interactions is tested against literature data for a standard set of 22 intermolecular complexes. Partitioning of interaction energy in the local correlation approach, based on the classes of occupied and virtual orbital involved in the interaction, clearly distinguishes the three types of interaction present in the set of complexes, in agreement with previous classifications. Geometry optimisation is found to be straightforward with DF-LMP2 without the need for counterpoise correction, resulting in geometries very close to previous, counterpoise-corrected structures. Spin-component scaling of gradients to correct for the known shortcomings of conventional MP2 has only a small effect on geometries in most cases, but significantly alters the distance between aromatic rings in stacked complexes. Harmonic frequency calculation is made possible by efficient use of parallel computing resources, and confirms all structures to be true minima, unlike previous estimates using density functional theory. Corrections for the change in zero-point vibrational energy are determined from this data, and typically constitute between 10 and 50% of the overall binding energy of the complex.

Introduction

Non-covalent interactions play an important part in a wide range of chemical and biological phenomena, for instance folding of proteins and nucleic acids into their correct 3-dimensional shapes.¹ Intermolecular forces acting within molecular clusters are widely used as models of such non-covalent interactions, since their effects can be more easily separated from intra-molecular forces. *Ab initio* theoretical methods are increasingly used to complement experimental study of such clusters. The majority of theoretical treatments of non-covalent interactions employ the supermolecular approach, in which interaction energy is calculated as the difference in energy between the cluster and its constituent monomers. However, two significant drawbacks of this approach are the need for large orbital basis sets, including polarisation and diffuse functions, for proper description of all interactions between molecules, and the related problem of basis set superposition error (BSSE).²

These problems can be ameliorated, to a large extent, by use of local correlation methods, which effectively eliminate BSSE by restricting excitations from occupied orbitals to those virtual orbitals that are in close spatial proximity.³ Moreover, in combination with the density fitting (DF) approximation, methods such as DF-LMP2⁴ and DF-LCCSD^{5,6,7,8} make significant efficiency savings over conventional calculations, approaching linear scaling in favourable cases, hence allowing application of suitable orbital basis sets for description of non-covalent interactions. A further advantage of the local correlation approach is that intermolecular correlation energy can be partitioned into physically meaningful contributions, such as dispersion and ionic terms, on the basis of the occupied and virtual orbitals involved.³ This approach is complementary to decomposition schemes, and is based solely on the local correlation treatment of the cluster and so requires no further calculations or significant computational resources.

BSSE can also be accounted for by use of the counterpoise method,⁹ but for a cluster of n molecules, this procedure requires $(n+1)$ energy calculations in the full basis set of the cluster, substantially increasing the computational resources required. The situation becomes still less satisfactory if geometry optimisation of the cluster is a goal, as separate gradient calculations for the cluster and all fragments are necessary. Without such corrections, geometry optimisation is likely to result in underestimation of intermolecular separations and overestimation of stabilities. Alternative approaches, such as symmetry adapted perturbation theory (SAPT) eliminate BSSE by construction, but are currently limited to pairs of molecules.¹⁰

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Molecular clusters are dynamic objects, and understanding of their vibrational properties is key to understanding their thermodynamic and spectroscopic properties. The near-universal starting point for theoretical description of vibrational behaviour is the harmonic approximation, which, despite its known shortcomings, is widely used to interpret thermodynamic and spectroscopic properties of interest. The efficiency and lack of BSSE in local correlation methods is again a significant advantage in calculating the necessary data. Recently, Hrenar *et al.* described the use of DF-LMP2 methods for numerical evaluation of the Hessian using finite differences of analytical gradients, made possible by using a highly parallelised algorithm.¹¹ In doing so, they demonstrated the utility of this approach, and identified scaling factors that correct for at least some of the deficiencies of the harmonic approach.

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
Grimme's spin-component scaling (SCS) method,¹² and variants thereof, have been shown to correct for deficiencies in MP2 for stacked complexes, but may actually degrade performance for hydrogen bonded complexes.¹³ Recently, Marchetti *et al.* proposed a dispersion-weighted SCS scheme that combines the accuracy of conventional MP2 for electrostatic complexes and of SCS-MP2 for dispersion bound ones.¹⁴ In the current investigation, density fitted, local correlation methods are applied to a range of intermolecular interactions, encompassing complexes bound mainly by electrostatic effects, by dispersion forces, or by a combination of these. These complexes are taken from the recent review of Hobza and co-workers,¹⁵ who suggest that they are representative of most important classes of non-covalent interaction. DF-LMP2, and DF-LCCSD where possible, methods are used to partition the interaction energy, to optimise geometries, and to calculate harmonic frequencies. In doing so, the aim is to show that DF-LMP2 and its SCS variant are capable of accurately and efficiently predicting the geometry and harmonic vibrational motion of complexes bound by non-covalent interactions, paving the way for studies of larger systems for which benchmark data is not available.

50 51 52 53 54 55 56 57 58 59 60 **Computational Methods**

All calculations were performed with the MOLPRO suite of programs,¹⁶ employing the correlation consistent aug-cc-pVTZ orbital basis set throughout.^{17,18} Density fitting of MP2 energies employed the corresponding MP2FIT basis set,¹⁹ while the underlying Hartree-Fock calculation also employed density fitting²⁰ with the corresponding JKFIT basis.²¹ DF in LCCSD for integrals over three and four external orbitals used the larger aug-cc-pVQZ/MP2FIT auxiliary basis sets to minimize fitting errors. The localised orbitals required for local correlation methods were generated via the Pipek-Mezey method,²² while the orbital domain selection followed the

1
2
3 procedure of Boughton and Pulay,²³ with merging of rotationally invariant π -domains where
4 appropriate. Initial test calculations indicated that using domains calculated at large intermolecular
5 separation and then frozen for calculations on the interacting system produced almost negligible
6 differences in geometry and harmonic frequencies when compared with domains computed solely
7 for the interacting system (so-called floating domains). For example, with the small methane dimer
8 system a maximum difference of 6 cm^{-1} between normal modes was observed, while for the larger
9 parallel-displaced benzene dimer system the maximum difference was 0.2 cm^{-1} . As deviations of
10 this size are likely to be small compared with both the basis set and method errors, floating
11 domains were used for all geometry optimisations and calculation of numerical harmonic
12 frequencies.
13
14
15
16
17
18
19
20
21

22
23 Analytical gradients for DF-LMP2 methods²⁴ are available in MOLPRO, enabling efficient
24 geometry optimisation. All geometry optimisations reported commenced from the reported
25 literature structure, and employed default criteria for convergence. Calculation of root mean square
26 (RMS) deviation between initial and final structures used Chemcraft.²⁵ Following geometry
27 optimisation, harmonic frequencies were calculated by numerical differentiation of analytic
28 gradients using central differences, employing the default step size of 0.01 a.u. This is a sizeable
29 task for the larger molecules considered, but was made feasible by use of a parallel algorithm,²⁶ as
30 implemented in MOLPRO, in which individual energy calculations on different geometry
31 displacements used in the formation of gradients and Hessians are performed independently on
32 separate processors. Individual DF-LMP2 calculations did not take account of molecular symmetry
33 (as this usually leads to symmetry equivalent localised orbitals), but this numerical algorithm can
34 use symmetry to reduce the number of displacements required. The local character of the orbitals
35 allows for the partitioning of the intermolecular interaction energy into intramolecular, dispersive,
36 and ionic components of the correlation energy.³ This is achieved by dividing the system into
37 monomers and separating the excitations from localised orbitals into different classes based on the
38 domains the pair of electrons are excited into.
39
40
41
42
43
44
45
46
47
48
49
50
51

52 53 **Results and Discussion**

54 Table 1 presents the results of partitioning the DF-LMP2 correlation energy into its intermolecular
55 components, along with the interaction energy at the SCF level, for the S22 set of complexes of
56 Hobza and co-workers.¹⁵ Following reference 3, dispersion refers to simultaneous excitations on
57 different monomers, ionic to excitation from an occupied orbital on one monomer to a virtual
58
59
60

orbital on the other along with a single excitation on the acceptor monomer, and dispersion-exchange to cross-excitations. Such definitions stem naturally from the local correlation approach, but may not be directly comparable to assignments from methods such as SAPT.¹⁰ Intramolecular correlation can also be calculated in this fashion, but does not show significant differences from the sum of monomer correlation energies. The three intermolecular correlation terms are taken directly from the DF-LMP2 cluster calculation, while the SCF interaction energy was calculated using the supermolecular approach. In all cases, dispersion-exchange energies are essentially negligible, and are included only for completeness.

Table 1 SCF interaction energies and partitioned DF-LMP2 intermolecular energies (kcal mol⁻¹)

	SCF	Dispersion	Ionic	Dispersion-Exchange
Ammonia dimer	-1.27	-1.23	-0.77	-0.01
Water dimer	-3.72	-1.06	-1.23	+0.01
Formic acid dimer	-15.49	-4.20	-5.63	+0.01
Formamide dimer	-12.32	-3.73	-4.09	+0.08
Uracil dimer HB	-9.31	-4.82	-4.79	+0.07
2-Pyridoxine...2-aminopyridine	-5.81	-5.52	-4.76	-0.03
Adenine...Thymine HB	-6.40	-5.68	-4.79	+0.01
Methane dimer	+0.35	-0.73	-0.14	0.00
Ethene dimer	+0.79	-1.88	-0.60	+0.01
Benzene...methane	+1.60	-2.19	-0.86	+0.02
Benzene dimer stack	+6.33	-7.22	-2.78	+0.02
Pyrazine dimer	+7.60	-7.77	-3.11	+0.03
Uracil dimer stack	+5.20	-9.15	-3.13	+0.01
Indole...benzene stack	+8.62	-10.60	-4.46	+0.11
Adenine...Thymine stack	+5.72	-14.05	-5.42	+0.05
Ethene...ethyne	-0.52	-1.01	-0.58	0.01
Benzene...water	-0.35	-2.01	-1.05	+0.01
Benzene...ammonia	+0.74	-2.10	-0.95	+0.01
Benzene...cyanide	-1.55	-2.75	-1.63	+0.04
Benzene dimer T	+2.54	-3.84	-1.57	+0.05
Indole...benzene T	+2.00	-5.13	-2.55	+0.05

Phenol dimer	+0.43	-4.72	-2.14	+0.02
--------------	-------	-------	-------	-------

The first seven clusters, denoted as hydrogen bonded in the original work, all display significant stabilisation at the SCF level, reflecting the importance of electrostatics in such interactions. Nonetheless, dispersion and ionic correlation effects contribute significantly to the overall stabilisation, with approximately equal importance. In all seven cases, the SCF stabilisation is the largest single term, but in three (ammonia dimer, 2-pyridoxine...2-aminopyridine and adenine...thymine) the combined correlation terms outweigh the SCF contribution. For example, in Watson-Crick hydrogen bonded adenine...thymine the SCF, dispersion and ionic contributions are 39%, 35% and 29% of the overall binding energy, respectively.

The next eight complexes were denoted as predominantly bound through dispersion,¹⁵ an assignment supported by the present analysis. In all eight clusters, SCF energies are repulsive, with all stabilisation stemming from correlation effects. Unlike the hydrogen bonded complexes, dispersion effects are much greater (typically by a factor of two to five) than ionic ones in these complexes, more than counteracting the repulsive SCF term. This is exemplified by the archetypal stacked, parallel-displaced benzene dimer, for which dispersion effects are 2.5 times greater than ionic ones. The final seven complexes are of so-called “mixed” type, with no single contribution dominating. This designation is again supported by the current analysis; SCF energies are attractive in three cases and repulsive in the remaining four. Dispersion contributions are generally larger than ionic ones, but not to the same extent as for the dispersion-bound clusters. Complexes that could be designated as containing X—H... π (X = C, N, O) non-covalent bonds, such as benzene...water, benzene...ammonia and benzene...HCN, show significant differences from the conventional hydrogen bonds. Most notably, dispersion contributions are significantly greater than SCF ones. The clear contrast in the origin of stabilisation between the three classes of complex is illustrated in Figure 1.

Figure 1 Illustration of energy partitioning data from Table 1. Vertical dashed lines show separation of classes of complex used in ref. 15. The numbering on the abscissa corresponds to the ordering of the complexes in Table 1.

The analysis of Table 1 is based on MP2 methods, known to perform adequately for H-bonding but to substantially over-estimate interaction energies of stacked complexes. Analogous partitioning of correlation energy with the DF-LCCSD method (energy partitioning of perturbative triples contributions is not currently possible) has also been performed. For inclusion of all relevant electron pairs in the LCCSD treatment, modification of the default distance based domain selection criteria is required: after some trial and error, all orbital pairs within a distance of 8 a.u. are treated as strong pairs, *i.e.*, at the CCSD level. Unfortunately, this renders DF-LCCSD calculations of the larger complexes unfeasible with a moderately large basis set such as aug-cc-pVTZ.

The partitioning of the DF-LCCSD correlation energy is presented in Table 2 for selected, smaller complexes, and shows that the trends observed using DF-LMP2 methods are largely preserved at this improved level of theory. For dispersion bound and mixed complexes, the tendency of MP2 to overestimate interaction energies is evident in both dispersion and ionic terms, with rather larger reductions in the former than the latter. This is exemplified in data for the ethene dimer, for which dispersion contributions are reduced by 0.1 kcal mol⁻¹ and ionic by 0.02 kcal mol⁻¹. This compares well with the Δ CCSD(T) correction to the overall interaction of +0.09 reported in ref 15.

Table 2 Partitioned DF-LCCSD interaction energies for selected complexes (kcal mol⁻¹)

Name	Dispersion	Ionic	Dispersion-Exchange
Ammonia dimer	-1.29	-0.79	-0.01
Water dimer	-1.20	-1.24	-0.03
Methane dimer	-0.72	-0.16	0.00
Ethene dimer	-1.78	-0.58	0.00
Benzene...methane	-2.03	-0.82	+0.01
Ethene...ethyne	-0.95	-0.54	0.00
Benzene...water	-1.90	-0.99	-0.01
Benzene...ammonia	-1.96	-0.85	-0.01

This analysis was performed at the literature geometries, which were obtained with either MP2 or CCSD(T) optimisation with triple- or quadruple- ζ basis sets and including counterpoise correction.

This can be a laborious process, and might be significantly eased by the advantages of the DF-LMP2 method discussed above, particularly the removal of BSSE. Table 3 reports the results of full, unconstrained geometry optimisation of the same S22 complexes using DF-LMP2/aug-cc-pVTZ, starting from the literature geometry. Proper application of local correlation methods requires symmetry to be disabled, but the optimisation algorithm retains any elements of symmetry present in the initial geometry.

Optimisation generally proceeded smoothly, in most cases reaching the default convergence criteria in MOLPRO (maximum gradient below 3×10^{-4} a.u.) in less than 10 steps. Table 3 demonstrates that the resulting optimised geometries are very similar to their literature equivalents in all but one case, that being the phenol dimer. Performance for hydrogen bonded and dispersion bound complexes is similar, whereas the slightly larger RMS deviation for the mixed complexes is almost entirely due to the phenol dimer (omitting this point gives a mean RMS of 0.018 Å for the remaining six complexes). No significant difference in the quality of agreement is observed between MP2 and CCSD(T) optimised literature geometries. Figure 2 shows the literature and optimised geometries of the phenol dimer, which by inspection are highly similar. The hydrogen bond O...H distances in the two cases are almost identical (1.937 vs. 1.935 Å). However, the dihedral angle between the mean planes of each molecule are different, with values of 60.5° in the literature geometry and 56.3° in the DF-LMP2 optimised structure. Thus, DF-LMP2 brings the aromatic rings into slightly closer contact (distance between centroids is 5.12 Å vs. 5.39 Å in the literature structure). This subtle change notwithstanding, the overall conclusion is that this efficient and counterpoise-free approach to geometry optimisation satisfactorily and consistently reproduces conventional, corrected geometries.

Table 3 RMS deviations from literature geometry (Å)

Name	DF-LMP2	SCSN
Ammonia dimer ^a	0.025	0.027
Water dimer ^a	0.030	0.037
Formic acid dimer ^a	0.017	0.016
Formamide dimer ^a	0.025	0.028
Uracil dimer HB	0.008	0.026
2-Pyridoxine...2-aminopyridine	0.009	0.034
Adenine...Thymine HB	0.011	0.028

Methane dimer ^a	0.005	0.050
Ethene dimer ^a	0.017	0.049
Benzene...methane	0.005	0.039
Benzene dimer stack	0.029	0.068
Pyrazine dimer	0.021	0.048
Uracil dimer stack	0.023	0.053
Indole...benzene stack	0.017	0.039
Adenine...Thymine stack	0.018	0.035
Ethene...ethyne ^a	0.014	0.017
Benzene...water	0.026	0.012
Benzene...ammonia	0.019	0.021
Benzene...cyanide	0.019	0.016
Benzene dimer T	0.015	0.027
Indole...benzene T	0.012	0.023
Phenol dimer	0.183	0.173
Mean	0.025	0.039
Mean H-bond	0.018	0.028
Mean dispersion	0.017	0.048
Mean Mixed	0.041	0.041
Mean CCSD(T) ^b	0.019	0.032

^a Literature geometry from CP-corrected CCSD(T)/AVQZ optimisation; all others from CP-corrected MP2/AVTZ. ^b For literature geometries optimised with CCSD(T) only.

Figure 2 Phenol dimer a) literature geometry; b) DF-LMP2 optimised.

Both conventional MP2 and DF-LMP2 are well known to overestimate the interaction energy of stacked complexes, and might therefore be expected to underestimate intermolecular separation in such complexes. This is apparent in potential energy curves for the stacked benzene dimer, which reach a minimum at vertical separation of *ca.* 3.5 Å at CBS(T) level and *ca.* 3.35 Å at MP2.²⁷ Recently, it was demonstrated that a new parameterisation of SCS designed specifically for non-covalent interactions and denoted SCSN,^{13a} overcomes this deficiency, resulting in high-accuracy

1
2
3 interaction energies for all classes of complex with errors of less than 0.5 kcal mol⁻¹. In this
4 approach, correlation energy from same- and opposite-spin electron pairs are weighted by scaling
5 factors in order to correct for systematic errors in standard MP2. The same scaling factors can be
6 applied to calculation of analytic gradients, allowing optimisation to minima on this scaled
7 potential energy surface. Table 3 also reports the RMS deviations from literature geometry using
8 this approach.
9
10
11
12
13

14
15 In general, SCSN optimised structures are slightly further from literature values than unscaled DF-
16 LMP2, both for MP2 and CCSD(T) reference data. Overall performance is similar to unscaled
17 optimisation for hydrogen bonded and mixed complexes, but differences are much larger for
18 dispersion bound complexes. The relatively small RMS deviations reported in Table 3 mask some
19 rather more significant changes; for instance, in the stacked benzene dimer, the separation of mean
20 planes is 3.361 Å in the literature (MP2/TZ-CP optimised) geometry, whereas our SCSN
21 optimisation results in a separation of 3.487 Å, much closer to the minimum of the CBS(T)
22 curve.²⁷ A similar trend is apparent in other stacked complexes, albeit to a lesser extent: in the
23 pyrazine dimer, the molecular planes are separated by 3.261 Å in literature geometry and 3.270 Å
24 in the SCSN geometry. In stacked adenine-thymine, molecular planes are not quite parallel, so for
25 the current investigation the perpendicular distance is defined as that from the mean plane of
26 adenine to the centroid of the 6-membered ring of thymine, which is 3.174 Å in literature and
27 3.205 Å in SCSN optimised geometry. Thus, it is evident that the excess stabilisation of stacked
28 complexes at MP2 level carries through into potential energy surfaces, and can be simply and
29 efficiently corrected by the empirical SCSN scheme.
30
31
32
33
34
35
36
37
38
39
40
41
42

43
44 A second SCSN geometry optimisation was performed for the T-shaped benzene dimer, following
45 the recent study by Wang *et al.* that showed the lowest energy geometry of this complex to deviate
46 from the ideal C_{2v} geometry.²⁸ Starting from a slightly offset geometry, optimisation proceeded
47 smoothly to a point just 0.06 kcal mol⁻¹ lower in energy than that reported in Table 3. Shown in
48 Figure 3, the SCSN geometry is similar to that reported in ref.28, with the donor C—H directed
49 towards a C—C bond rather than the centre of the acceptor ring.
50
51
52
53
54
55
56
57

58 **Figure 3** SCSN optimised geometry of T-shaped benzene dimer
59
60

1
2
3 The fully optimised DF-LMP2 geometries were then used to calculate harmonic frequencies using
4 numerical differentiation. Firstly, and most importantly, this procedure confirms that all but one
5 optimised structures are true minima at this theoretical level, as confirmed by the lack of imaginary
6 frequencies. As noted above, any molecular symmetry present in the literature geometry was
7 retained in these calculations, such that our data lends support to the symmetries employed in the
8 original study of Hobza and co-workers. This is in marked contrast to a recent DFT-based study of
9 the same set of molecules, in which one or more imaginary frequencies were observed for several
10 complexes.²⁹ The only exception is the T-shaped benzene dimer, for which the high-symmetry
11 (C_{2v}) form is found to be a transition state, with a single imaginary frequency of -10.8 cm^{-1} . Re-
12 optimisation with lower (C_s) symmetry results in a geometry similar to that shown in Figure 3,
13 which is found to be a true minimum on the potential energy surface.
14
15
16
17
18
19
20
21
22
23

24 A similar procedure was followed for the SCSN-optimised geometries. In this case, both the C_{2v}
25 and C_s optimised forms of the T-shaped benzene dimer are found to be true minima, echoing the
26 findings of Wang *et al.*,²⁸ who used dispersion-corrected DFT to show that both forms are stable
27 and are connected by a transition state with a barrier of just $0.01 \text{ kcal mol}^{-1}$. In this regard, SCSN
28 again seems to be slightly more accurate than unscaled DF-LMP2 in finding both forms to be
29 stable, but the inherent errors associated with this approach are much larger than these energy
30 differences, such that any conclusion based on this data must be tentative at best. All remaining
31 structures are found to be minima using this approach. Two geometries of the stacked (parallel-
32 displaced) benzene dimer were tested, one with a C—H bond situated over the partner benzene and
33 a second with a C—C bond over the other benzene. Both are found to be minima using SCSN,
34 with effectively identical energies within the expected accuracy of the method (less than 0.1 kcal
35 mol^{-1} difference).
36
37
38
39
40
41
42
43
44
45
46
47

48 As well as confirming structures as minima, harmonic frequency calculations can be used to
49 estimate the change in zero-point vibrational energy (ZPVE) that accompanies formation of a
50 complex from constituent monomers. Such changes stem largely from conversion of translational
51 and rotational degrees of freedom of individual molecules into intermolecular vibrations. Such
52 vibrations are likely to be strongly anharmonic in nature, such that harmonic data can only be a
53 first-order approximation to the true ZPVE correction. Nonetheless, obtaining even harmonic
54 vibrational data for the larger complexes considered here is a considerable challenge, and the data
55 presented in Table 4 is a step towards establishing accurate ZPVE corrections for systems of this
56 size.
57
58
59
60

Table 4 Harmonic zero point energy corrections for binding (kcal mol⁻¹)

Name	DF-LMP2	Scaled ^a	SCSN
Ammonia dimer	1.70	1.62	1.66
Water dimer	2.37	2.27	2.33
Formic acid dimer	2.08	1.99	2.06
Formamide dimer	2.64	2.53	2.52
Uracil dimer HB	1.21	1.16	1.15
2-Pyridoxine...2-aminopyridine	1.01	0.97	0.98
Adenine...thymine HB	0.95	0.90	0.89
Methane dimer	1.16	1.11	1.14
Ethene dimer	0.83	0.80	0.82
Benzene...methane	0.64	0.62	0.70
Benzene dimer stack	0.53	0.51	0.56
Pyrazine dimer	0.61	0.58	0.51
Uracil dimer stack	0.58	0.56	0.56
Indole...benzene stack	0.25	0.24	0.43
Adenine...thymine stack	0.72	0.68	0.61
Ethene...ethyne	0.71	0.68	0.73
Benzene...water	1.06	1.01	1.04
Benzene...ammonia	0.70	0.67	0.79
Benzene...cyanide	0.60	0.57	0.71
Benzene dimer T	0.52	0.49	0.67
Indole...benzene T	0.62	0.60	0.59
Phenol dimer	0.95	0.91	0.99

^a Scaled from raw DF-LMP2 data using the recommended scaling factor of 0.956

$\Delta ZPVE$ values on formation of complexes from monomers are largest for the hydrogen bonded complexes, and are in general rather smaller for dispersion bound and mixed complexes. The single largest change is found for the formamide dimer, closely followed by water and formic acid dimers. The value for the water dimer is similar to previous estimates using MP2 methods with the harmonic approximation (2.1 – 2.3 kcal mol⁻¹), and as expected rather larger than values that take anharmonicity into account (1.8 – 2.1 kcal mol⁻¹).³⁰ The larger aromatic systems containing two

1
2
3 strong hydrogen bonds show much smaller $\Delta ZPVE$, falling to *ca.* 1 kcal mol⁻¹ for 2-pyridoxine...2-
4 aminopyridine and adenine...thymine. Values for dispersion bound and mixed complexes are in
5 general rather smaller than for hydrogen bonded ones. For example, $\Delta ZPVE$ for the stacked uracil
6 dimer is less than half the value for the hydrogen bonded isomer, with a similar pattern for the
7 adenine-thymine complexes. A slightly surprising exception to this pattern is the methane dimer,
8 whose $\Delta ZPVE$ value of 1.16 kcal mol⁻¹ is more than twice the binding energy. This may represent
9 the breakdown of the harmonic approximation, or be a result of noise introduced by the use of
10 numerical differences in such a weakly bound case.
11
12
13
14
15
16
17
18
19

20 In addition to raw DF-LMP2 data, Table 4 also contains $\Delta ZPVE$ values scaled by the value of
21 0.956 recommended by Hrenar *et al.*,¹¹ and values from numerical frequency calculation based on
22 SCSN energies. Changes using the single scaling factor are generally very small, and do not
23 significantly alter the conclusions drawn from unscaled data. A similar pattern emerges from most
24 of the SCSN data, with one exception: in the stacked indole-benzene complex, the SCSN value is
25 almost double that from unscaled DF-LMP2. In the absence of benchmark data, it is not possible to
26 decide which of these values is the more reliable, but in the context of other stacked complexes the
27 raw DF-LMP2 value seems rather small.
28
29
30
31
32
33
34
35

36 Conclusions

37 The utility of local correlation methods, when combined with density fitting, has been
38 demonstrated in the study of non-covalent interactions. By construction, these methods allow for
39 decomposition of correlation energy into contributions from dispersion and ionic excitations, thus
40 giving insight into the origin of stabilisation at no additional computational cost to the correlation
41 calculation. This analysis, at both DF-LMP2 and DF-LCCSD levels, clearly distinguishes
42 complexes bound by hydrogen bonding, dispersion, or a mixture of these effects. The lack of basis
43 set superposition error in local methods, coupled with their computational efficiency, makes them
44 ideally suited for simple geometry optimisation and harmonic frequency calculation. Geometry
45 optimisation leads to structures in close agreement with published, counterpoise-corrected
46 geometries, demonstrating the utility of this approach. Harmonic frequency analysis, using both
47 unscaled and SCSN-scaled MP2 data, indicates that most complexes within the S22 set are at least
48 local minima on their respective potential energy surfaces. The only exception to this conclusion is
49 the C_{2v} T-shaped benzene dimer, which is a minimum with SCSN scaling but a saddle point with
50 unscaled MP2. Zero-point vibrational energy corrections to binding energies have also been
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

estimated, using raw data and two distinct scaling methods. In general, all three methods are in reasonable agreement, with corrections of between 20% and 50% of uncorrected binding energies.

Acknowledgements

This work was funded by the UK EPSRC, grant ref. EP/C013328/1.

References

- ¹ A. Pavlov, P.M. Mitrasinovic, *Curr. Org. Chem.*, 2010, **14**, 129.
- ² P. Hobza, H.L. Selzle, E.W. Schlag, *J. Phys. Chem.*, 1996, **100**, 18790.
- ³ M. Schütz, G. Rauhut, H.-J. Werner, *J. Phys. Chem. A*, 1998, **102**, 5997.
- ⁴ H.-J. Werner, F. R. Manby, P. J. Knowles, *J. Chem. Phys.*, 2003, **118**, 8149.
- ⁵ C. Hampel, H.-J. Werner, *J. Chem. Phys.*, 1996, **104**, 6286.
- ⁶ M. Schütz, H.-J. Werner, *J. Chem. Phys.*, 2001, **114**, 661.
- ⁷ M. Schütz, *Phys. Chem. Chem. Phys.*, 2002, **4**, 3941.
- ⁸ M. Schütz, F. R. Manby, *Phys. Chem. Chem. Phys.*, 2003, **5**, 3349.
- ⁹ S. F. Boys, F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
- ¹⁰ B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887.
- ¹¹ T. Hrenar, G. Rauhut, H.-J. Werner, *J. Phys. Chem. A*, 2006, **110**, 2060.
- ¹² S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095.
- ¹³ a) J.G. Hill, J.A. Platts, *J. Chem. Theor. Comput.*, 2007, **3**, 80. b) R.A. Bachorz, F. A. Bischoff, S. Hofener, W. Klopper, P. Ottiger, R. Leist, J. A. Frey, S. Leutwyler, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2758.
- ¹⁴ O. Marchetti and H.-J. Werner, *J. Phys. Chem. A*, 2009, **113**, 11580.
- ¹⁵ P. Jurečka, J. Šponer, J. Černý, P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985.
- ¹⁶ MOLPRO, version 2008.1, a package of *ab initio* programs, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz and others, see <http://www.molpro.net>.
- ¹⁷ T. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007.
- ¹⁸ R. A. Kendall, T. Dunning Jr., R. J. Harrison, *J. Chem. Phys.*, 1991, **96**, 6797.
- ¹⁹ F. Weigend, A. Köhn, C. Hättig, *J. Chem. Phys.*, 2002, **116**, 3175.
- ²⁰ R. Polly, H.-J. Werner, F. R. Manby, P. J. Knowles, *Mol. Phys.*, 2004, **102**, 2311.
- ²¹ F. Weigend, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4285.
- ²² J. Pipek, P. G. Mezey, *J. Chem. Phys.*, 1989, **90**, 4916.
- ²³ J. W. Boughton, P. Pulay, *J. Comput. Chem.*, 1993, **14**, 736.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
-
- ²⁴ M. Schütz, H.-J. Werner, R. Lindh, F. R. Manby, *J. Chem. Phys.*, 2004, **121**, 737.
- ²⁵ Chemcraft v1.6: www.chemcraftprog.com
- ²⁶ J. Neugebauer, M. Reiher, C. Kind, B.A. Hess, *J. Comput. Chem.*, 2002, **23**, 895.
- ²⁷ a) S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J. Am. Chem. Soc.*, 2000, **122**, 3746. b) M. O. Sinnokrot, C. D. Sherrill, *J. Am. Chem. Soc.*, 2004, **126**, 7690. c) J.G. Hill, J.A. Platts, H.-J. Werner, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4072.
- ²⁸ W. Wang, M. Pitoňák, P. Hobza, *ChemPhysChem*, 2007, **8**, 2107.
- ²⁹ K. Gkionis, J.G. Hill, S.P. Oldfield, J.A. Platts, *J. Mol. Model.*, 2009, **15**, 1051.
- ³⁰ M. W. Feyereisen, D. Feller, D.A. Dixon, *J. Phys. Chem.*, 1996, **100**, 2993.