Accurate computations of the structures and binding energies of the imidazole···benzene and pyrrole···benzene complexes

Sandra Ahnen [a], Anna-Sophia Hehn [a], Konstantinos D. Vogiatzis [a,b,1], Maria A. Trachsel [c], Samuel Leutwyler [c], Wim Klopper [a,b,*]

[a] Institute of Physical Chemistry, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 2, D-76131 Karlsruhe, Germany
[b] Center for Functional Nanostructures, Karlsruhe Institute of Technology (KIT), Wolfgang-Gaede-Straße 1a, D-76131 Karlsruhe, Germany
[c] Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

ARTICLE INFO

ABSTRACT

Using explicitly-correlated coupled-cluster theory with single and double excitations, the intermolecular distances and interaction energies of the T-shaped imidazole···benzene and pyrrole···benzene complexes have been computed in a large augmented correlation-consistent quadruple-zeta basis set, adding also corrections for connected triple excitations and remaining basis-set-superposition errors. The results of these computations are used to assess other methods such as Møller–Plesset perturbation theory (MP2), spin-component-scaled MP2 theory, dispersion-weighted MP2 theory, interference-corrected explicitly-correlated MP2 theory, dispersion-corrected double-hybrid density-functional theory (DFT), DFT-based symmetry-adapted perturbation theory, the random-phase approximation, explicitly-correlated ring-coupled-cluster-doubles theory, and double-hybrid DFT with a correlation energy computed in the random-phase approximation.

1. Introduction

1.1. Motivation

In view of the importance of the NH···π interactions in biochemistry and structural biology, the purpose of the present work is to provide benchmark-quality equilibrium structures and binding energies of the imidazole···benzene (Im···Bz) and pyrrole···benzene (Py···Bz) complexes by means of quantum-chemical computations at the explicitly-correlated CCSD(T) coupled-cluster level (coupled-cluster theory with single and double excitations plus a correction for connected triples), which approaches closely the basis-set limit of that level. The results of these computations will be used to assess other computational methods, at lower levels of theory, which could be used to provide theoretical support of our experimental vibronic-spectroscopy studies of these and other, related systems [1–4].

In terms of cost effectiveness, recent method development in the area of density-functional theory (DFT) has given us a wide variety of promising approaches. Accurate low-cost methods are particularly useful to study not only dimers but also larger systems such as tri- or tetramers, which we are also interested in. Dispersion-corrected density-functional theory (DFT-D) seems particularly promising [5,6], and we shall compare our basis-set-limit CCSD(T) values for the imidazole- and pyrrole···benzene complexes to results from various DFT-D approaches. We shall also investigate DFT methods that compute the electron-correlation energy in the random-phase approximation (RPA) [7] as well as DFT-based symmetry-adapted perturbation theory (DFT-SAPT) [8,9]. Where applicable, basis-set corrections shall be computed using explicitly-correlated wavefunctions, for example for approaches using double-hybrid functionals. In this respect, the present work contains a few uncommon approaches, for example computing the RPA correlation energy from the *explicitly-correlated* ring-coupled-cluster-doubles method, or using (explicitly-correlated) RPA correlation energies in double-hybrid functionals.

The present article is organised as follows: computational details of the present work are given in Section 2. Here, it will be described how the initial equilibrium geometries have been optimised (at the SCS-MP2/aug-cc-pVTZ level, *vide infra*) and how they have been refined at our composite scheme used for method

---

* Corresponding author at: Institute of Physical Chemistry, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 2, D-76131 Karlsruhe, Germany. Tel.: +49 721 60847263.
E-mail address: klopper@kit.edu (W. Klopper).
[1] Current address: Department of Chemistry, University of Minnesota, 207 Pleasant Street S.E., Minneapolis, MN 55455-0431, USA.

assessment. Our interaction-energy benchmarks are compared with several other approaches, whose computational details are given as supplementary data. Results are discussed in Section 3 and the article ends with some conclusions collected in Section 4.

2. Computational details

2.1. Interaction and binding energies

Counterpoise-corrected energies $E_{AB}^{cp}(AB)$ are defined as [10]

$$E_{AB}^{cp}(AB) = E_{AB}(AB) - E_{AB}(A) - E_{AB}(B) + E_A(A) + E_B(B), \qquad (1)$$

where $E_N(M)$ is the energy of the molecular system M in the basis N and where the geometries of A and B are those in the complex AB. The notation N = AB means that the respective molecular system is computed using all of the basis functions of the whole complex AB. For example, $E_{AB}(A)$ is the energy of fragment A obtained from a computation in the basis of complex AB, with the geometry of A in the complex AB.

Notice that in the present article, we distinguish between the quantities "interaction energy", a purely theoretical quantity which is not an observable, and "binding energy", which after correction for zero-point vibrational energy *is* an observable.

The *interaction energy* is computed either *without* (w/o CP; ΔE_2) or *with* applying a counterpoise correction (w/CP; ΔE_2^{cp}). These interaction energies as defined as follows:

$$\text{w/o CP}: \quad \Delta E_2 = E_{AB}(AB) - E_A(A) - E_B(B); \qquad (2)$$
$$\text{w/CP}: \quad \Delta E_2^{cp} = E_{AB}^{cp}(AB) - E_A(A) - E_B(B). \qquad (3)$$

Notice that the interaction energies ΔE_2 and ΔE_2^{cp} are two-body terms that are defined by dissociating the complex into individual fragments A and B while keeping the geometries of these fragments fixed. The one-body term

$$\Delta E_1 = E_A(A) - E_A^{eq}(A) + E_B(B) - E_B^{eq}(B) \qquad (4)$$

accounts for the deformation of the fragment geometries in the complex AB, where $E_N^{eq}(M)$ is the energy of the molecular system M in its own equilibrium geometry. Finally, by adding the one- and two-body terms, the counterpoise-corrected and not-counterpoise-corrected *binding energies* are computed as

$$\text{w/o CP}: \quad D_e = -\Delta E_2 - \Delta E_1; \qquad (5)$$
$$\text{w/CP}: \quad D_e^{cp} = -\Delta E_2^{cp} - \Delta E_1, \qquad (6)$$

respectively.

2.2. Composite scheme for the CCSD(T) basis-set limit

The composite scheme that we have used to refine the equilibrium geometries and to compute the interaction energies of the two complexes at the basis-set limit begins with explicitly-correlated coupled-cluster computations with single and double excitations in the CCSD(F12) approximation of Ref. [11]. (For recent reviews on explicitly-correlated coupled-cluster theory, see Refs. [12,13].) These computations were performed with the RICC2 module [12,14–16] of the program package TURBOMOLE [17,18] using the aug-cc-pVQZ basis [19,20] for all atoms. As complementary auxiliary basis set (CABS [21]) for the CCSD(F12) computations, we have used the aug-cc-pVQZ-OptRI basis of Yousaf and Peterson [22]. CCSD(F12) computations require two more auxiliary basis sets, that is, a basis set for the density-fitting approximation to certain two-electron integrals (MP2 fitting basis; CBAS) and a basis set for the density-fitting approximation to the matrix representation of the exchange operator of Hartree–Fock theory (exchange fitting basis; JKBAS [23]). As CBAS, we used the aug-cc-pwCV5Z CBAS basis (aug-cc-pV5Z for H) [24,25]. As JKBAS, we used the

cc-pV5Z JKBAS basis for H, C, and N, all taken from the TURBOMOLE repository [18].

The CCSD(F12) computations were performed using a restricted Hartree–Fock (RHF) reference, using ansatz 2B [F+K] (see Ref. [16] for details), using the fixed-amplitudes ansatz of Ten-no [26], and applying the approximations of Ref. [27] to the integrals over the commutator between the geminal and the operator of kinetic energy. The Slater-type geminal [28] with exponent $\gamma = 1.4\ a_0^{-1}$ [29] was represented by six Gaussian-type geminals, taking the exponents and contraction coefficients from Ref. [30].

To the CCSD(F12) energy, we have added a correction for connected triple excitations. First, we have computed the energy contribution $E_{(T)}$ in the aug-cc-pVTZ basis [19,20] using the program package MOLPRO [31,32]. Then, following Ref. [33], we have computed the energy contribution $E_{(T*)}$ by scaling the connected-triples correction $E_{(T)}$ obtained in the aug-cc-pVTZ basis by the ratio of the MP2-F12 and MP2 correlation energies,

$$E_{(T*)} = x_{F12} E_{(T)}; \qquad (7)$$
$$x_{F12} = E_{MP2\text{-}F12}/E_{MP2}. \qquad (8)$$

In Eq. (8), the $E_{MP2\text{-}F12}$ correlation energy has been computed in the aug-cc-pVQZ basis, whereas the E_{MP2} correlation energy has been computed in the aug-cc-pVTZ basis. In previous work, to ensure size extensivity of the correlation approach, Marchetti and Werner [34] as well as Marshall and Sherrill [35] have computed x_{F12} once and for all for the complex AB and have kept this ratio fixed for the computation of the (T*) contributions to the correlation energies of the fragments. In contrast to this, we decided to take the supermolecule approach: x_{F12} is determined individually for each interfragment separation R. For the dissociation limit, this means that the energies $E_{MP2\text{-}F12}$ and E_{MP2} are computed as the *sum* of the respective energies of the fragments A and B,

$$x_{F12}(R = \infty) = \frac{E_{MP2\text{-}F12}(A) + E_{MP2\text{-}F12}(B)}{E_{MP2}(A) + E_{MP2}(B)}. \qquad (9)$$

This supermolecule approach provides a consistent energy behaviour when the intermolecular distance is increased to infinity, but our dissociation limit is not equal to the sum of the two energies obtained in *two individual* x_{F12} computations of A and B.

In the framework of the counterpoise approach, a corresponding ratio x_{F12}^{cp} is computed from the counterpoise-corrected energies $E_{MP2\text{-}F12}^{cp}$ and E_{MP2}^{cp} and multiplied with the counterpoise-corrected triples contribution $E_{(T)}^{cp}$. The *counterpoise-corrected* CCSD(F12)+(T*) level of computation was chosen as our composite scheme to approach the CCSD(T) basis-set limit.

All explicitly-correlated computations include a correction (E_{CABS}) for single excitations into the complementary auxiliary basis set [36], and all electron-correlation computations of the present work were performed in the frozen-core approximation, unless otherwise noted.

For comparison, we have also computed the CCSD(F12)+(T*) energies in the aug-cc-pVDZ and aug-cc-pVTZ basis sets. Except for the exponents of the Slater-type geminals ($\gamma = 1.1$ or $1.2\ a_0^{-1}$ in conjunction with aug-cc-pVDZ or aug-cc-pVTZ, respectively [29]), these computations were performed in full analogy to the computations in the aug-cc-pVQZ basis, that is, with the auxiliary basis sets aug-cc-pwCV(X+1)Z CBAS and cc-pV(X+1)Z JKBAS for X = D, T.

2.3. Geometry optimisations

Initial equilibrium geometries of the imidazole···benzene and pyrrole···benzene complexes were optimised at the not-counterpoise-corrected SCS-MP2/aug-cc-pVTZ level (spin-component-scaled MP2 theory, see Ref. [37]) using the RICC2 module [14,15]