Journal of Molecular Graphics and Modelling 29 (2010) 178–187

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Evaluation of dispersion-corrected density functional theory (B3LYP-DCP) for compounds of biochemical interest

Sten O. Nilsson Lill*

Department of Chemistry, University of Gothenburg, SE-412 96 Göteborg, Sweden

ARTICLE INFO

Article history:
Received 21 April 2010
Received in revised form 1 June 2010
Accepted 2 June 2010
Available online 15 June 2010

Keywords:
Density functional theory
Dispersion correction
Aromatic interactions
Peptide
Conformational analysis

ABSTRACT

An evaluation of a dispersion-corrected density functional theory method (B3LYP-DCP) [I.D. Mackie, G.A. DiLabio, Interactions in large, polyaromatic hydrocarbon dimers: application of density functional theory with dispersion corrections, J. Phys. Chem. A 112 (2008) 10968–10976] for three systems of biochemical interest is presented. Firstly, structures and energies of isomers of the tripeptide Phe-Gly-Phe have been compared with CCSD(T)/CBS//RI-MP2/cc-pVTZ literature values. In the system aromatic interactions compete with XH–π (X = C, N) interactions and hydrogen bonds which makes it a reliable model for proteins. The resulting mean absolute deviation between B3LYP-DCP and CCSD(T)/CBS relative energies is found to be 0.50 kcal mol^{-1}. Secondly, a phenylalanine derivative featuring a CH–π interaction has been investigated. A comparison between the optimized geometry and X-ray crystal data shows that B3LYP-DCP accurately predicts the interaction between the two aromatic rings. Thirdly, the dipeptide Ac-Phe-Phe-NH_2 which contains an edge-to-face interaction between two aromatic rings has been studied. The study demonstrates the general applicability of the B3LYP-DCP method on systems which features interactions typically present in biochemical compounds.

1. Introduction

In the last decades it has become more and more evident that weak interactions play significant roles in biological recognition. Examples of this are interactions between aromatic residues in enzymes in general, interactions between drugs and aromatic residues in the active site of an enzyme, interactions between carbohydrates and the protein matrix, XH–π (X = C, N, O, S) interactions, and hydrogen bonds [1–11]. A very illustrative example of this is the blend of interactions found between the anti-Alzheimer drug Aricept (2-((1-benzylpiperidin-4-yl)methyl)-5,6-dimethoxy-2,3-dihydroinden-1-one) and the enzyme acetylcholinesterase (Fig. 1) [2].

One of the major challenges in computational chemistry is to be able to handle weak interactions such as hydrogen bonds and dispersive interactions in an accurate way. Although a high-level wave function method such as CCSD(T) is capable of accurately describing these types of interactions it is generally not applicable to systems of large size due to its inherent demand of computational resources. Therefore, for the less time-consuming density functional theory (DFT) methods, considerable effort has been put into the development of correction terms to treat dispersive interactions between non-covalently bonded molecules. For example, the group of DiLabio has developed dispersion-correcting atom-centred potentials (DCPs) for use with a variety of common DFT functionals and basis sets [12,13]. The DCP is added in the form of two Gaussian functions, one long-range attractive and a second short-ranged repulsive. For larger systems this was found necessary to avoid over-binding [12,13]. The coefficients for the two functions were optimized to reproduce the dispersion binding in a training set consisting of non-covalently bound dimers of hydrocarbons and of heteroatom containing molecules. It was found that binding energies could be predicted within 15%, and monomer separations within 0.1 Å, of high-level *ab initio* data [12,13]. DCPs have successfully been applied to study dispersive binding in a diverse set of systems, including cyclic aromatic hydrocarbons [12–14], benzothiophenes [15], small hydrocarbons and pentacene on hydrogen-terminated silicon surfaces [16,17], and recently also for hydrocarbon isomerisation reactions and olefin monomer insertion reactions [18]. The DCP approach is attractive in three ways: it gives the possibility of switching the correction term "on or off" and therefore directly gives the answer to how important dispersive interactions are for the system under investigation. DCPs are also applicable in a number of general computational chemistry programs, and one does not need specialised functionals or large basis sets [12,13]. An alternative way of handling dispersive interactions has been described by Grimme and co-workers. They have developed DFT-D in which a damped empirical atom pair specific

* Tel.: +46 31 7869103; fax: +46 31 7723840.
E-mail address: stenil@chem.gu.se.

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.06.001

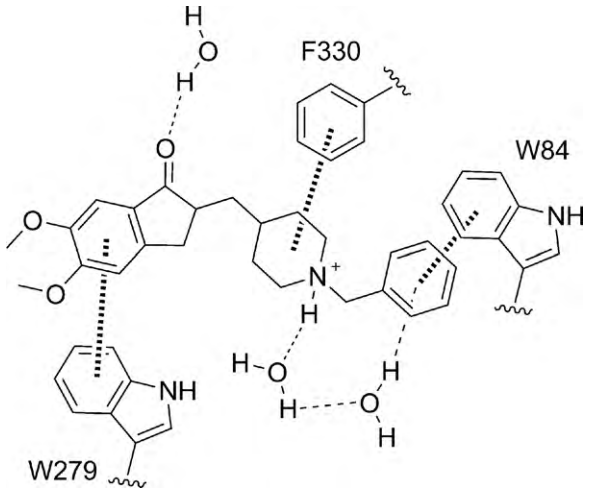

Fig. 1. Interactions (dashed lines) between Aricept and the active site of acetylcholinesterase.

potential of C_6/R^6 type is added on top of existing DFT functionals [19–21]. As a comparison, with DFT-D the dispersion attraction stems from attraction between nuclei, while in the DCP-approach the attraction comes from the added potential that acts between electrons and nuclei. Using the successful DFT-D approach one can also switch the correction term "on or off" and directly estimate its influence on energies and geometries. Besides these techniques, new DFT functionals in which weak interactions are well described compared to how standard DFT functionals perform have recently been released, for example M06 and M06-2X [22]. In addition, truly non-local DFT-methods, such as the vdW-DF approach based on the Andersson–Langreth–Lundqvist (ALL)-functional has shown great promise [23,24].

B3LYP, which is a standard DFT-approach, in general shows great performance for thermochemistry [25,26], but due to its repulsive long-range behavior [12] B3LYP is not recommended for weakly interacting systems [26]. Recently [14], the dispersion-corrected B3LYP/6-31+G(d,p)-DCP method was evaluated against the S22 benchmark set [27] which contains a range of complexes interacting via dispersive interactions, hydrogen bonds, or combinations of these. With a mean deviation of only 0.8 kcal mol^{-1} to benchmark wave function calculations for dimerization energies, it was shown that B3LYP/6-31+G(d,p)-DCP is a useful method for weakly interacting systems [14]. The 6-31+G(d,p) basis set used was the one recommended by the DCP developers to give a good balance between computing time and good quality results [13], and if a method with such a fairly small basis set gives reliable results it would make it very useful for studies of systems of larger size. However, it is of interest to continue to evaluate B3LYP/6-31+G(d,p)-DCP also for systems of biochemical interest to demonstrate the general applicability of the method. In this study, the following systems have been investigated which separates the paper in three sections:

(a) The tripeptide Phe-Gly-Phe for which high level *ab initio* energies and structures are available [28]. This allows for an adequate evaluation of different isomers calculated using B3LYP-DCP. In the isomers a competition between aromatic ring-ring interactions, XH–π (X = C, N) interactions, and hydrogen bonds occurs.
(b) A phenylalanine derivative with available X-ray structural data demonstrating an intramolecular CH–π interaction [29]. This allows for direct structural validation of B3LYP-DCP on this type of weak interaction. In addition, ^{1}H NMR data, sensitive to structural changes and therefore useful chemical probes, are evaluated.
(c) The dipeptide Ac-Phe-Phe-NH_2, which has been characterized in the gas-phase by IR-spectroscopy and computational studies [30]. This system shows an edge-to-face interaction between the two aromatic rings. A comparison of calculated and experimental vibrational frequencies, also sensitive to different interactions, is made.

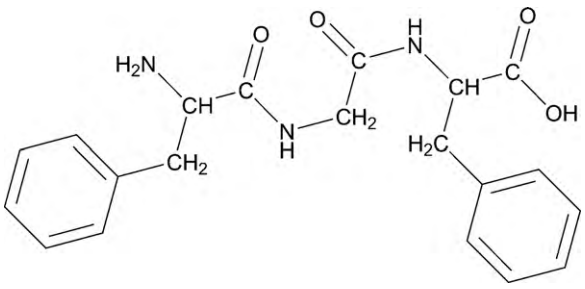

Fig. 2. The tripeptide system Phe-Gly-Phe.

2. Computational methods

All compounds were optimized using B3LYP[31–33]/6-31+G(d,p)[34–36]-DCP [13] in Gaussian03 [37]. Examples of input for performing this type of calculation using Gaussian can be found in the Supporting Information. Since the dispersion correcting potential as explained in the original references explicitly was parameterized to cover the basis set superposition error, no such correction was employed here [12,13]. All geometries were characterized as minima or saddle points on the potential energy surface by using the sign of the eigenvalues of the force constant matrix obtained from a frequency calculation. Single point calculations for the tripeptide system were also performed using B3LYP/6-31+G(d,p)-DCP//RI-MP2/cc-pVTZ for a comparison. The MP2 geometries were taken from Ref. [28]. Chemical shieldings, *i.e.* the second derivative of the energy with respect to an external magnetic field and the nuclear magnetic moment, were calculated with the standard Gauge Independent Atomic Orbital (GIAO) [38] method using B3LYP/6-31+G(d,p), both with and without the DCP included, or with the larger triple zeta basis set 6-311+G(d,p)-DCP, relative to TMS optimized in T_d-symmetry. Figures and plots of calculated structures were generated using Chemcraft [39].

3. Results and discussion

3.1. The tripeptide system Phe-Gly-Phe

In a recent study by Valdes et al. the tripeptide Phe-Gly-Phe was investigated using benchmark CCSD(T)/CBS//RI-MP2/cc-pVTZ calculations (Fig. 2) [28].

The tripeptide system includes aromatic side-chains which are capable of forming stacked or T-shaped complexes due to stabilizing interactions between aromatic rings. In addition it may form CH–π and NH–π interactions and hydrogen bonds of NH···O=C, OH···O=C, or NH···NH_2 type, so it appears to be a realistic model system for competitive interactions typically observed in proteins. In this study, we take advantage of the recent data from Valdes et al.'s study [28] on different isomers of the tripeptide Phe-Gly-Phe and compare with results generated using B3LYP-DCP.

In the study by Valdes et al., different computational approaches were evaluated against CCSD(T)/CBS//RI-MP2/cc-pVTZ calculations for fifteen different conformers (Fig. 3) [28]. These were ranked in terms of relative potential energies and RMS structural devi-