Selection of an Appropriate Basis Set for Accurate Description of Binding Energy: A First Principles Study

Salih AKBUDAK

1 Adiyaman University, Arts and Sciences Faculty, Physics Department, 02040, Adiyaman

(Alınış / Received: 06.06.2017, Kabul / Accepted: 26.09.2017, Online Yayınlanma / Published Online: 01.11.2017)

Abstract: Binding energies of Ar₂ and H₂ dimers have been investigated using correlation consistent cc-pVXZ and aug-cc-pVXZ basis sets together with Coupled Cluster with Singles and Doubles with Perturbative Triples (CCSD(T)) method. Two point extrapolations to complete basis set limit has been applied to reduce basis set incompleteness (BSIE) error. Discrepancy of our theoretical binding energy values from current experimental binding energy values in literature both for Ar₂ and H₂ dimers observed to be less than 1 kcal/mol.

1. Introduction

Over the years there has been significant improvements in computer technology which has boosted up theoretical calculations carried out in physics, chemistry and material science and engineering [1-5]. Electronic structure calculations which try to understand the characteristics of system by solving Schrödinger equation comprehend most of the theoretical calculations in physics and chemistry [6-12]. However, exact solution of Schrödinger equation is only possible for a limited number of problems like particle in a box, hydrogen atom and harmonic oscillator. If the system of interest is composed of two and more electrons, we need to take into account electron correlations in the system which increase computation time by N³ where N is the number of electrons. Schrödinger equation can only be solved approximately. Density Functional Theory which is one of the distinguished electronic structure theory is implemented under different softwares like VASP, Quantum Espresso, Orca, Gaussian, Siesta i.e., [13-16]. These different softwares use different basis sets. For instance VASP and Quantum Espresso use plane-wave based basis sets [17-22], ORCA and Gaussian softwares use gaussian basis sets [23-26]. Siesta uses numeric-atom centered orbital basis sets [27-28].

In all electronic structure methods accuracy of Schrödinger equation is determined by the exhange correlation method and basis set. So we can say that there are two types of errors in electronic structure calculations. One is due to exhange correlation method and other is due to basis set. If the exhange correlation method selected for calculations is from Hartree Fock (HF) to CCSD(T) computation time increases significantly with respect to number of basis functions. So we have to find a balance between computation time and accuracy. However we should keep in mind that there are many basis sets in the literature. Among the most common basis sets are Dunning's and Huzinaga's correlation consistent basis sets [29-36], John A. Pople's basis sets [37-43], Peterson's basis sets [44-46] and Jan Erik Almlöf's basis sets [47-49]. In this paper we have used Dunning's correlation consistent basis sets together with CCSD(T) method. Since CCSD(T) is an advanced
correlated method which reduces the error to less than 1 kcal/mol, it is accepted as gold standard method among all advanced methods. So, using this method we observed the effect of cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ basis sets on binding energies of Ar₂ and H₂ dimers. Due to the fact that Ar and H are one of the most abundant elements in nature, we have made CCSD(T) calculations to reveal the effects of different correlation consistent basis sets on binding energies of these dimers. Accurate description of binding energies of dimers is a challenging case due to overlapping between basis function of one atom with the basis function of the other atom which causes a basis set superposition error [50, 51].

Binding energies of dimers and molecules have been intensively studied both experimentally and theoretically [52-68]. To mention some of these studies, Deible et al. studied the binding energy of the beryllium dimer using Quantum Monte Carlo method [64], Ian S.O Pimienta studied Ionic Dimers [52], E. Miliodoros et al. carried out a benchmark study of the pi-pi binding energy in the benzene dimer [53], A.A Zadoroshnya et al. studied ionization-induced structural changes in uracil dimers and their spectroscopic signatures [57], P. A Pienizek et al. studied Electronic Structure of the benzene dimer [50, 51].

So we have investigated the effect of different basis sets on binding energies and reported our results at CCSD(T) level. CCSD(T) is the gold standard method that gives the most accurate results so by doing this we tried to choose the best performing basis set that will have a low computation cost. Although computation time for dimers is not a big problem if we go beyond dimers and use more complex molecules computation time increases drastically. Thus selection of an appropriate basis set for calculations that will enable the desired accuracy can reduce computation time significantly.

2. Material and Method

For all calculations we used ORCA software [16]. As in all electronic structure methods, the first task is to solve electronic Schrödinger equation exactly. Time-independent Schrödinger equation is given as:

\[
\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right) \Psi = E \Psi \tag{1}
\]

To have detailed information about the system wavefunction should be exactly known. Wave function in Schrödinger equation can be written as a linear combination of basis functions;

\[
\Psi = \sum_\mu c_\mu \Phi_\mu \tag{2}
\]

here \(\Psi\) is wavefunction, \(c\) is a constant and \(\Phi\) is basis function. Wavefunction can be obtained self-consistently from Equation 2. However to obtain exact value of wavefunction ‘b’ in Equation 2 should go to infinity \(\infty\) which means that we should include infinitely many basis functions in our basis set. Unfortunately this is computationally impractical because computation time scale as \(N^3\) where \(N\) is number of basis functions. Since we can not have infinitely basis functions in our basis set we always encounter a basis set incompleteness error (BSIE). One popular way to alleviate this BSIE is to use complete basis set extrapolation techniques. In this paper we used Halkier’s two point extrapolation technique [70]. Two point extrapolation scheme is given as;

\[
E[X] = E[\infty] + \frac{A}{X^3} \tag{3}
\]

\[
E[Y] = E[\infty] + \frac{A}{Y^3} \tag{4}
\]

using Equations 3 and 4 real equation of CBS can be derived as;

\[
CBS = E_{XY}[\infty] = \frac{E[X]X^3 - E[Y]Y^3}{X^3 - Y^3} \tag{5}
\]

where \(X\) and \(Y\) are maximum angular momentum numbers in basis set. For example if we do extrapolations from DZ to TZ, we take \(X\) as 3 and \(Y\) as 2.

The most important term in electronic Schrödinger equation is exchange correlation term since it can not be determined analytically for systems having two and more electrons. So, one has to apply an approximation method. In this paper, we have chosen Coupled Cluster Singles and Doubles with Perturbative Triples method which is a post-Hartee-Fock method. When applying this method to medium sized molecules results are satisfying. After obtaining CCSD(T) total energies for the system we have obtained binding energies of H₂ and Ar₂ dimers using conventional binding energy formula;

\[
\Delta E_{\text{binding}} = E_{\text{molecule}} - 2E_{\text{atom}} \tag{6}
\]

where \(E\) is the total energy \([69]\).

3. Results

Figure 1 (a)-(e) shows the absolute binding energy values with respect to cc-pVXZ (X=D, T, Q and 5) and aug-cc-pVXZ (X= D, T and Q) basis sets for Ar₂ dimer. All electronic structure calculations are carried out at CCSD(T) level. Extrapolated binding energies for Figures 1(a)-(e) are 0.0049, 0.0054, 0.005, 0.0067 and 0.0064 eV respectively.
Figure 2 shows the binding energies of \( \text{Ar}_2 \) with respect to correlation consistent cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ basis sets. Binding energies obtained with these basis sets are 0.0024, 0.0035, 0.0043, 0.0047, 0.0072, 0.0069 and 0.0066 eV respectively. From these observations the best available basis sets for description of binding energy of \( \text{Ar}_2 \) are cc-pVTZ and cc-pVQZ since energy discrepancy from aug-cc-pVQZ is less than 2 meV. Computationally the most demanding basis set is aug-cc-pVQZ. If cc-pVTZ and cc-pVQZ basis sets are used, and 3Z to QZ two point extrapolation is applied, computation time decreases significantly compared to basis sets used in figures 1 (c)-(d). Results obtained with this basis sets are in chemical accuracy limit (1 kcal/mol or 43 meV) since experimental binding energy of \( \text{Ar}_2 \) dimer is 12 meV [65,71].

Figures 3 (a)-(e) show the absolute binding energy values of \( \text{H}_2 \) dimer with respect to different correlation consistent basis sets of Dunning [29]. Figures 3 (a) to (e) CBS extrapolated values are obtained by applying DZ to 3Z, 3Z to QZ, QZ to 5Z two point extrapolations for cc-pVXZ (X=D, T, Q, 5) and DZ, TZ and QZ extrapolations for aug-cc-pVXZ(X=D, T, Q) basis sets. Extrapolated binding energy values shown by CBS in Figures 3 (a)-(e) are 4.39, 4.01, 3.41, 3.29 and 3.26 eV respectively.

Figure 4 shows binding energy of \( \text{H}_2 \) dimer for cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ basis sets at CCSD(T) level. Binding energies calculated are 6.09, 4.89, 4.38, 3.91, 3.17 eV for cc-pVXZ (X=D, T, Q, 5) basis sets and 3.25 eV for aug-cc-pVXZ(X=D, T, Q) basis sets. Experimental binding energy of \( \text{H}_2 \) dimer is 4.39 eV [71] so if we choose cc-pVDZ and cc-pVTZ basis sets and make extrapolations to complete basis set limit (CBS) from DZ to TZ we obtain a binding energy of 4.39 eV for \( \text{H}_2 \). So theoretical value we obtained at CCSD(T) level is 0.01 eV less than experimental value. And we can also use cc-pVTZ basis sets which has a binding energy of 4.38 eV. If we use this basis set we do not need to make extrapolations.
In summary, we have investigated the effect of different basis sets on binding energies and tried to find the best basis set for each dimer in terms of computation time and accuracy. In principle what we expect for a specific calculation is that the larger is the basis set we use the higher the accurate results we have. But, especially for large molecules, not for dimers computation time is also very important, because when deciding to use a basis set, computation time is an important constraint, so selection of a basis set of modest size that would give the desired accuracy for calculation is the main objective in electronic structure calculations. The need for these types of studies is indispensable since there is a huge number of basis sets in literature and they are growing exponentially. So, it is always a need to compare the performance of existing basis sets for a system and find the best performing basis set in terms of computation time and accuracy. And we believe that this paper will initiate more studies about different molecules and dimers that will shed light to computational chemists and computational physicists. The investigation of large molecules with different basis sets will be carried out in the future.

References


benchmark set of biologically relevant molecules. Physical Chemistry Chemical Physics, 8(2006), 5287-5293.