Rapid and precise thermochemical calculations by quantum chemical methods

Ph.D. thesis

By: Adrienn Ruzsinszky

Supervisor: Dr. Gábor Csonka

Budapest University of Technology and Economics

Department of Inorganic Chemistry

2003

Introduction

Estimation of correlation energy plays a central role in theoretical chemistry. The electron correlation is the error of Hartree-Fock method and equals to the difference of the non-relativistic complete CI and numerically calculated Hartree-Fock energy or extrapolated to infinite basis set. This error is only one percent of the total Hartree-Fock energy, but it is critical in the point of view of chemistry. This error leads to 220 kcal/mol average absolute error of the standard enthalpies of formation for the 222 test molecules of the G3/99 database, thus the non-corrected Hartree-Fock method is not applicable for thermochemical calculations.

Analysis of exact correlation energy on free atoms shows that the correlation energy depends on the nuclear charge, on the number of electron and on their spin. Knowledge of the correlation energy of atoms cannot be easily transported to the molecular environment because the spin system of the molecules differs from those of atoms. Our goal is to develop a new method that is more rapid than the previous methods and it is able to yield enthalpies of formation within the chemical accuracy (1-2 kcal/mol) for classical molecules. The Rapid Estimation of the Basis set Error and Correlation Energy from Partial charges (REBECEP) method can be summarized as follows:

1. We partition the molecular correlation energy among atoms.

2. The correlation energy of an atom is calculated with the use of the partial charges by linear interpolation.

3. We add the correlation energy estimated from sum of atomic parts to the Hartree-Fock energy calculated with a given basis.

4. The total energy corrected by the zero-point vibration energy, the thermal correction and the atomic energies according to a known procedure yields the non-relativistic enthalpy of formation of the molecule.

For developing and testing the method we constructed a database of 161 molecules that contains the molecular geometry, the atomic partial charges of the atoms of the molecules, the Hartree-Fock energy, the zero-point energies, and the experimental data. With the use of a multilinear equation system we determine the atomic correction factors reproducing best the experimental enthalpies of formation of the selected molecules. The largest disadvantage of this method is that it requires the expensive frequency analysis. For this reason we have developed a parameter system that can treat implicitly the zero-point energy and thermal

corrections: Rapid Estimation of the Enthalpies of Formation from Hartree-Fock results (REEF-HF).

In the assessment of approximate functionals for the exchange-correlation energy, great weight is given to the accuracy of enthalpies of formation calculated from the atomization energy. The error of the previously developed non empirical functionals for the molecules studied by us comes from the inconsistency between the energy of the free atoms and molecules. This error cancels out in the chemical reactions that do not involve free atoms. Relying only on the standard enthalpies of formation derived from atomization energies one can obtain misleading results about the quality of the given method for reaction enthalpies. Smaller errors arise due to the imperfection of the basis set and due to neglection of the relativistic effects. The calculated DFT enthalpies of formation can be improved occasionally by introducing an atomic energy parameter that corrects the energy of the free atom. We have tested a new, non empirical functional and we have given the expected precision of that functional for the calculation of standard enthalpies of formation.

Thesis

1. We have successfully developed a parameter set that is able to correct the Hartree-Fock total energy calculated with 6-31G(d) and 6-311+G(2d,p) basis sets for H, C, N, O, F, Si, S, and Cl containing molecules within the REEF-HF (Rapid Estimation of Enthalpies of Formation from Hartree-Fock results) method that treats implicitly the zero-point correction and thermal corrections.

2. REEF-HF calculations on 161 molecules show that experimental enthalpy of formation can be approximated with 1.5 kcal/mol average absolute error even in the case of larger molecules, whereas this error is about 1 kcal/mol in the many thousand times more expensive G3 method. We have compared our results for large molecules to the results obtained by the widely used B3LYP method. This latter, considerably more expensive method can reproduce the standard enthalpies of formation with 7 kcal/mol average absolute error.

3. Using NPA partial charges as correction to the HF/6-31G(d) total energy we can obtain better results than with Mulliken charges. The average absolute error is 1.97 kcal/mol with NPA charges, and 2.14 kcal/mol with Mulliken charges.

4. The results calculated with 6-311+G(2d,p) basis and NPA charges are considerably better than the 6-31G(d) results (e.g. the average error decreases from 1.97 to 1.55

kcal/mol). For Cl containing molecules considerable improvement can be achieved due to the increase of the size of the basis set. Using Mulliken charges we couldn't notice such improvement. Considering this we don't recommend the use of Mulliken charges with 6-311+G(2d,p) basis set in REEF-HF calculations.

5. Independent on partial charge calculation method the largest negative error (subtracting the calculated value from the experimental) was obtained for azulene (about -9 kcal/mol). The other extreme deviation was shown by the butanedinitrile with the largest positive error (+6 kcal/mol). Considerable error can be observed for the carbonic-difluoride, cyclopropene and bicyclobutane. Our analysis yielded that the REEF-HF method gives systematic error for strained rings and t-butyl group containing molecules.

6. REEF-HF results with 6-311+G(2d,p) basis and NPA charges are available for 159 molecules. If we leave out the 5 most problematic molecules the average absolute error decreases to 1.38 kcal/mol for the remaining 154 molecules. REEF-HF method can be used well for molecular families.

7. For 27 selected hydrocarbons from our database the NPA or Mulliken charges corrected HF/6-31G(d) total energy show only 0.77 kcal/mol or 0.64 kcal/mol average absolute errors, respectively. This is a very good result compared to atomic equivalent corrected B3LYP/6-311G(d,p) results, since the latter method yields 0.81 kcal/mol average error of for 23 more simple hydrocarbons. Among the selected 27 molecules there were some more problematic as naphthalene or spiropentane.

8. Comparing 69 experimental, HF/6-31G(d) (scaled by 0.8929), B3LYP/6-31G(2df,p) (scaled by 0.9854) and estimated zero-point energies based on the molecular stoichiometry or the atomic charges we obtained that HF and B3LYP methods yield the best results (~0.4 kcal/mol average absolute error). The average error of the estimated zero-point energies based on the molecular stoichiometry or the atomic charges was about 1 kcal/mol.

9. Using the explicit zero-point energy estimation based on partial charges in the REBECEP method the results worsen by 0.2 kcal/mol on average. This signals that for the estimation of the standard enthalpies of formation it is more appropriate to calculate zero-point energies implicitly like in the REEF-HF method, because the quality of the results will not deteriorate.

10. We have tested a new non-published functional developed by Tao, Perdew Staroverov and Scuseria (TPSS). The small and systematic error of this functional can be reduced

with the right selection of the basis [6−311G(d,p)], so we can achieve extraordinary accuracy for the standard enthalpy of formation of the 50 organic molecules of the G3-3 subset containing H, C, N, O and F atoms. The 6−311G (d,p) basis works well with the TPSS functional, achieving a remarkable cancellation of moderate errors. The average error of TPSS/6−311G(d,p) is close to zero, the standard deviation is about 2.7 kcal/mol. I we change from the G2/97 data base to the G3-3 data base the average error of TPSS/6−311G(d,p) model decreases from 3.2 kcal/mol to 2.4 kcal/mol.

11. We have concluded that SWVN5 and PBE functionals systematically overbinds and PKZB functional underbinds. For the molecules studied most of the error of the earlier fully or nearly non-empirical functionals resides in the energies of the free atoms.

12. We proposed to add small energy corrections to the calculated atomic energies in order to obtain the best possible agreement in a least square sense with experimental enthalpies of formation. Remarkable increase in performance was observed for the SVWN5 functional in the case of 50 larger molecules, the average error decreased from 217.7 kcal/mol to 4.4 kcal/mol this latter value is almost the half of the error of non-corrected B3LYP method for the same molecules. In the case of PBE functional the average absolute error decreases from 32,2 to 2.2 kcal/mol. The error of TPSS/6-311G(d,p) model doesn't change, indicating the consistent atomic and molecular energies of the non corrected TPSS method.

Publications

1. Csonka, G. I.; **Ruzsinszky, A.;** Tao, J; Perdew J.P.
   Energies of Organic Molecules and Atoms in Density Functional Theory
   ACS, submitted

2. **Ruzsinszky, A.;** Csonka, G. I.
   Rapid Estimation of Enthalpies of Formation from Hartree-Fock Total Energy and
   Partial Charges for Compounds Containing Si, S, and Cl Atoms
   J. Phys. Chem. A. **2003**, accepted

3. **Ruzsinszky, A**.; Csonka, G. I.; Margitfalvi, J. L.; Kristyán, S.
   Performance of Rapid Estimation of Zero Point Energies, Part 2
   J. Phys. Chem. A. **submitted**.

4. **Ruzsinszky, A**.; Kristyán, S.; Margitfalvi, J. L.; Csonka, G. I.
   Rapid Estimation of Zero Point Energies of Molecules using Hartree-Fock Atomic
   Partial Charges
   J. Phys. Chem. A. **2003**, 107, 1833-1839.

5. **Ruzsinszky, A.**; Van Alsenoy, C.; Csonka, G. I.
   Implicit Zero-Point Vibration Energy and Thermal Corrections in Rapid Estimation of
   Enthalpies of Formation from Hartree-Fock Total Energy and Partial Charges
   J. Phys. Chem. A. **2003**, 107, 736-744.

6. **Ruzsinszky, A**., Van Alsenoy, C., Csonka, G. I., Optimal Selection of Partial Charge
   Calculation Method for Rapid Estimation of Enthalpies of Formation from Hartree-
   Fock Total Energy
   J. Phys. Chem. A **2002**, 106, 12139-12150

7. Csonka, G. I., **Ruzsinszky, A**., Oláh, J., Van Alsenoy, C., Organizing Atomic Partial
   Charges into a Database
   J. Mol. Struct. (THEOCHEM), **2002,** 589-591, 1-5.

8. Kristyán, S., **Ruzsinszky, A**., Csonka, G. I. Reproducing Gaussian-3 Total Energy
   Using Fitted Atomic Correlation Parameters for RECEP (Rapid Estimation of
   Correlation Energy from Partial Charges) Method and Hartree Fock Results,
   J. Phys. Chem. A **2001**, 105, 1926.

9. Kristyán, S., **Ruzsinszky, A.**, Csonka, G. I**.**, Accurate Thermochemistry from
   Corrected Hartree Fock Results: Rapid Estimation of Nearly Experimental Quality
   Total Energy Using the Small 6-31G(d) Basis Set,
   Theor. Chem. Acc. **2001** 106, 319-328.

10. Kristyán, S., **Ruzsinszky, A.**, Csonka, G. I., The Performance of Rapid Estimation of
    Basis Set Error and Correlation Energy from Partial Charges (REBECEP) Method on
    New Molecules of G3/99 Test Set,
    Theor. Chem. Acc. **2001** 106, 404-411.