SPIN
JURNAL KIMIA & PENDIDIKAN KIMIA
https://journal.uinmataram.ac.id/index.php/spin

COMPUTATIONAL STUDY: THE EFFECT OF BASIS SET ON THE ENERGY STABILITY OF HYDROCARBONS AND HALOGENIC ACIDS

Nilawati✉[1], Yogi Surya Priadana[1], Yoga Surya Priadana[1], Zainina Azani[1], Mardia Solin[1], Hamidah[1], Ika Bella Pratiwi[1]

[1]Pendidikan Kimia, FKIP Universitas Samudra, Jln. Kampus Meurandeh, Langsa 24416, Indonesia.
DOI: 10.20414/spin.v6i2.10289

History Article
Accepted:
June 02, 2024
reviewed:
November 20, 2024
Published:
December 20, 2024

Keywords:
Alkanes; base sets; computing; energy; halogen acids

ABSTRACT

This research is motivated by the many base sets contained in computational calculations. The selection and use of base sets play an important role in determining the accuracy of the calculation of theoretical molecular properties as well as in lowering the BSSE (base set superposition error). This study aims to analyze the influence of base set variations on the calculation of hydrocarbon and halogen acid energy stability using computational methods. The research method used is an experimental method. Molecular modeling using Avogadro software with the molecules used are halogen acid molecules and alkanes C1 - C5. The calculation uses NwChem software with the DFT method, functional BL3YP with 4 different base sets, namely STO-3G, 321-G, cc-pVDZ, and cc-pVTZ. The resulting data is then analyzed and compared with database calculations. The results of the study show that the higher the base set used, the more accurate the calculation results obtained. The base set that shows the most accurate results in this study is cc-pVTZ. However, it should be noted that the most accurate set bases can be different for each molecule.

ABSTRAK

Penelitian ini dilatarbelakangi oleh banyaknya *basis set* yang terdapat dalam perhitungan komputasi. Pemilihan dan penggunaan *basis set* memainkan peran penting dalam menentukan akurasi perhitungan sifat – sifat molekul secara teoritis serta untuk menurunkan BSSE (*basis set superposition error)*. Penelitian ini bertujuan untuk menganalisis pengaruh variasi *basis set* terhadap perhitungan stabilitas energi hidrokarbon dan asam halogen menggunakan metode komputasi. Metode penelitian yang digunakan adalah metode eksperimental. Pemodelan molekul menggunakan software Avogadro dengan molekul yang digunakan adalah molekul asam halogen dan alkana C1 - C5. Perhitungan menggunakan *software* NwChem dengan metode DFT, fungsional BL3YP dengan 4 *basis set* yang berbeda yakni STO-3G, 321-G, cc-pVDZ, dan cc-pVTZ. Data yang dihasilkan kemudian dianalisis dan dibandingkan dengan perhitungan *database*. Hasil penelitian menunjukkan bahwa semakin tinggi *basis set* yang digunakan maka semakin akurat hasil perhitungan yang didapatkan. Adapun *basis set* yang menunjukkan hasil yang paling akurat dalam penelitian ini adalah cc-pVTZ. Namun, perlu dicatat bahwa *basis set* paling akurat dapat berbeda untuk setiap molekul.

How to Cite

Nilawati., Priadana, Y. S., Priadana, Y. S., ..., & Pratiwi, I. B. (2024). Computational Study: The Effect of Basis Set on The Energy Stability of Hydrocarbons and Halogenic Acids. *SPIN-Jurnal Kimia & Pendidikan Kimia*. 6(2). 194-201.

✉Correspondence Author:
nila526262@gmail.com
p-ISSN: 2580-2623
e-ISSN: 2745-685

INTRODUCTION

In computational chemistry, the choice of basis set plays an important role in determining the accuracy of theoretical calculations of molecular properties (Daga et al., 2020; Jensen Frank, 2017). A basis set is defined as a set of mathematical functions used to describe or explain the behavior of electrons in atoms (Jensen Frank, 2017). A basis set is a wave function for an individual atom that is very important in forming the basis of quantum chemical calculations. The wave function of a molecule is obtained from a linear combination of atomic orbitals (LCAO) mainly based on density and ab initio calculations. The accuracy of the wave function depends on the quality of the basis set (Coşkun & Ertürk, 2022; Lehtola, 2021). The more accurate the basis set used, the better the wave function and the more accurate the calculation results obtained (Morgante & Peverati, 2020). Based on this statement, computational calculations usually use the largest basis set (Kirschner et al., 2020). However, using the largest basis set does not guarantee accurate results. Better results require calculations with a larger basis set so that the results we get now still have uncertainty (Nabil et al., 2023)

One of the important calculations theoretically is to find the strength of the interaction between components used to check the stability of the system. The interaction between the constituent molecules of a system has several names. Interaction energy (also known as adsorption energy) and binding energy have been used to indicate the strength of the interaction between the constituents. Using a supermolecular approach, the interaction energy (E_{int}) and binding energy (E_{bin}) can be calculated by:

$$E_{int} \text{ or } E_{bin} = E_{complex\ AB} - E_{monomer\ A} - E_{monomer\ B}$$

The difference between E_{int} and E_{bin} is the way the monomer energy is obtained. For E_{int}, the monomer energy is "calculated at the same core position as in the total system", while for E_{bin}, the monomer core is at its optimal position (Grabowski, 2017).

In computational chemistry, when performing calculations using a limited basis set, one will encounter "basis set superposition errors" or BSSE. When calculating the bond energy between two molecules, a smaller basis set is used for each molecule, which causes the fragment energies (separate molecules) to tend to be higher than they should be. This results in an overestimation of the bond energy. As monomer A approaches monomer B, the dimer can be artificially stabilized because monomer A uses the additional basis functions of monomer B to describe its electron distribution, and vice versa. The error arises from inconsistent treatment of monomers — they can access the additional functions of the other monomer at shorter intermolecular distances, but at larger intermolecular distances, the other monomer is too far away (overlap integrals are too small) for its functions to provide stabilization. Inconsistent treatment of the basis sets for each monomer because of varying intermolecular distances is the source of the basis set superposition error (BSSE). The smaller the basis set, the more

severe the effect (Nagy & Jensen, 2017; Sherrill, 2018).

Hydrocarbons and halogenated acids are two classes of compounds that have many uses in various fields, including industry, pharmaceuticals, energy and fuels, as solvents and reagents. Hydrocarbons include organic compounds composed of carbon and hydrogen atoms. Halogen acids are compounds containing halogen atoms (fluorine, chlorine, bromine, iodine) bound to hydrogen atoms. These two classes of compounds have different electronic structures so consideration is needed in choosing a basis set.

Molecular energy is one of the components or parts calculated in computational chemistry. Molecules with lower energy tend to be more stable than molecules with higher energy. The lower the energy, the more difficult it is for a molecule to react or decompose (Jensen Frank, 2017). In addition to determining stability, molecular energy also affects reactivity. Chemical reactions tend to occur if the total energy of the product is lower than that of the reactants (Asmuruf et al., 2017).

Several studies have been conducted, stating that the basis set affects the calculation of molecular energy (Myllys et al., 2016; Nabil et al., 2023; Pérez-Barcia et al., 2023; Susanti et al., 2020). However, specific studies on the effect of the basis set on the energy of hydrocarbon and halogen acid molecules are still limited. Therefore, this study highlights the effect of the basis set on the energy of hydrocarbon and halogen acid molecules and investigates the basis set that has the most stable molecular energy. Where using the most accurate basis set, will increase the accuracy of the calculation so as to provide information on the calculation of molecular properties that are closer to experimental values.

METHOD

This research begins by drawing a stable molecular structure that will be calculated for energy using Avogadro software. The molecules that will be used are shown in Table 1:

Table 1. Molecules whose energy will be calculated based on several basis sets.

Halogen Acid	Hydrocarbons
Flurry Acid (HF)	Methane (CH_4)
Hydrochloric Acid (HCl)	Ethane ($_{C2}H_6$)
Hydrobromic Acid (HBr)	Propane (C_3H_8)
Iodide Acid (HI)	Butane (C_4H_{10})
	Pentane (C_5H_{12})

Furthermore, the initial structure that has been created is optimized using DFT theory with the B3LYP functional. By using the B3LYP function, changes in interaction energy can be calculated. B3LYP is ubiquitous in molecular studies, making this level of theory a standard in many fields of study (Nabil et al., 2023). NwChem software is used for structure calculations. The basis sets used in this study are the 3-21G, STO-3G, cc-pVDZ, and cc-pVTZ basis sets. Computational calculations use the metapad and NwRun applications. The calculation results that have been obtained are then processed using excel, analyzed and the results validated by comparing the calculation results with the database contained in the Computational Chemistry Comparison and Benchmark DataBase (JOHNSON III & RUSSELL, 2020).:

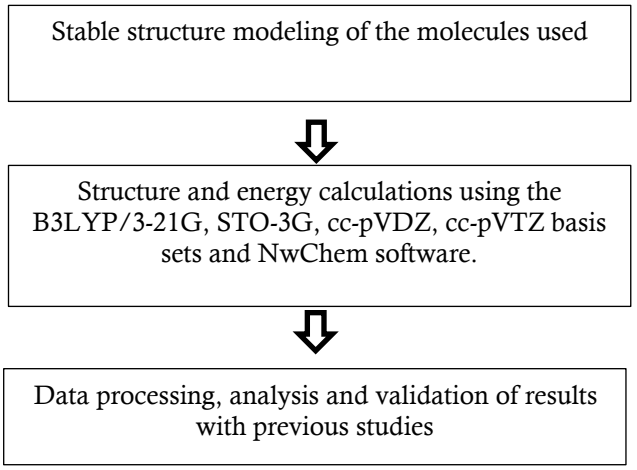

Figure 1. Flowchart of structural modeling (adapted from Hasby et al., 2016)

RESULT AND DISCUSSION

Physical and solid state structural data for many compounds provide evidence for the formation of intermolecular hydrogen bonds. Hydrogen bonds are formed between H atoms attached to an electronegative atom and an electronegative atom that has a lone pair of electrons. Halogen acids are formed from hydrogen (H) and halogen atoms (F, Cl, Br, I). All halogen acids (HX) are gases at 298K with a sharp sour odor. The bond length is defined by the optimal distance between the atomic nuclei that leads to maximum stability. Each covalent bond has a characteristic bond strength and bond length (Housecroft & Sharpe, 2012)

In the context of chemistry, molecular energy is very important to determine the stability and geometric structure of a molecule. Based on the calculations that have been done, the energy value of hydrofluoric acid is the highest followed by hydrochloric acid, hydrobromic acid, and hydroiodic acid. The results of the calculations can be seen in table 2.

Table 2. Results of calculating the energy prices of halogen acids based on the basis set used in kJ/mol.

Name / Basis Set	Hydrofluoric Acid	Hydrochloric Acid	Hydrobromic	Iodide Acid
STO-3G	-259.660,6517	-1.196.551,229	-6.688.313,933	-18.003.697,37
STO-3G database	-259.716,7389	-1.196.599,657	-6.686.027,603	-18.092.716,5
3-21G	-263.693,5392	-1.203.986,498	-6.727.513,787	-18.092.684,12
3-21G database	-262.190,2934	-1.204.025,75	-6.727.547,924	-18.092.910,88
cc-pVDZ database	-263.693,5392	-1.209.888,599	-6.760.028,061	-
cc-pVTZ	-263.766,6591	-1.209.903,381	-6.760.199,884	-
cc-pVTZ database	-263.819,6157	-1.209.943,399	-6.760.232,412	-
H – Halogen Bond Length	0,93	1,30	1,45	1,65

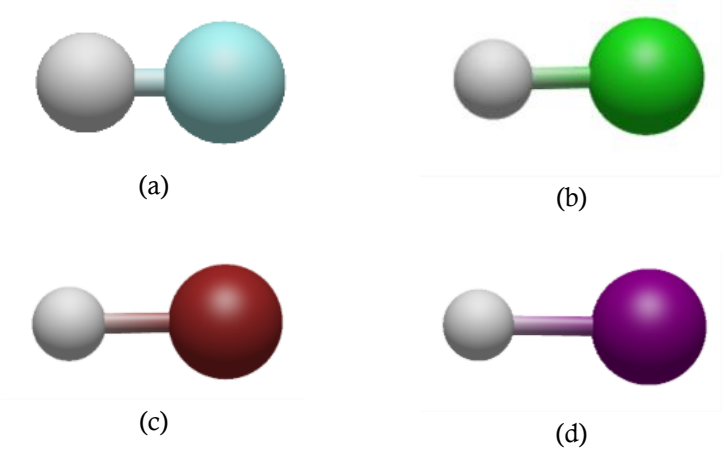

Figure 1 (a) molecular structure of hydrofluoric acid, (b) molecular structure of hydrochloric acid, (c) molecular structure of hydrobromic acid, (d) molecular structure of hydrofluoric acid.

Based on Table 2, the calculation results using 4 different basis sets, namely STO-3G, 3-21G, cc-pVDZ, and cc-pVTZ, show that hydrofluoric acid has the highest energy, followed by hydrochloric acid, hydrobromic acid, and the lowest is iodic acid energy. Judging from the difference in the results of the database energy and the basis set of the calculation results, there is no significant difference indicating that the calculation results are not far off. In table 2, there is data on the bond length, the bond length of HI is the furthest. This is in line with the nature of the periodic table of elements where from top to bottom the atomic radius will increase due to the increasing number of atomic shells. Atomic radius affects the bond length because it is largely determined by the number of atomic radii involved between the bonded atoms. The greater the energy of a molecule, the more unstable it will be. Conversely, the smaller the energy produced, the more stable the geometric structure of the compound will be (Yuniar et al., 2019). From Table 2, the most stable compounds are iodic acid, hydrobromic acid, hydrochloric acid, and hydrofluoric acid.

Simple hydrocarbons that only contain single bonds are called alkanes. The length of the carbon chain in alkanes varies, starting from methane (CH_4) with one carbon atom, to alkanes with longer carbon chains, such as ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), and pentane (C_5H_{12}). Each subsequent member in the alkane series is formed by the addition of one C atom and two H atoms to the previous member (Petrucci et al., 2013).

Differences in the length of the carbon chain in alkanes have a significant effect on the properties of their molecules, including molecular energy. Based on the results of calculations and comparisons with the database, the more carbon chains in alkane compounds, the higher the energy stability will be, as indicated by the order of molecular energy. Table 3 shows that the energy of the pentane compound is the lowest compared to butane, propane, ethane, and methane compounds.

Table 3. The calculation of the molecular energy of alkane hydrocarbons based on the basis set used in kJ/mol units.

Name / Basis Set	Methane	Ethane	Propane	Butane	Pentane
STO-3G	-104.030,7701	-206.663,8043	-308.488,0673	-410.310,6072	-512.130,1194
STO-3G database	-105.831,0695	-207.130.2791	-309.110.7273	-411.131,3062	-513.135.4179
3-21G	-104.982,2067	-208.054,8734	-310.559,4882	-413.056,7833	-515.557,8188
3-21G database	-105.811,8324	-208.481.9758	-311.127.3056	-413.813,3883	-516.482.5104
cc-pVDZ database	-106.375,6743	-209.607,3123	-312.812.7636	-416.057,8416	-519.287,6655
cc-pVTZ	-105.822,0567	-209.282,9821	-312.385,5064	-415.479,8922	-512.130,1194
cc-pVTZ database	-106.433,1439	-209.699,3335	-312.939.4676	-4.162.1919,44	-493.228,5462

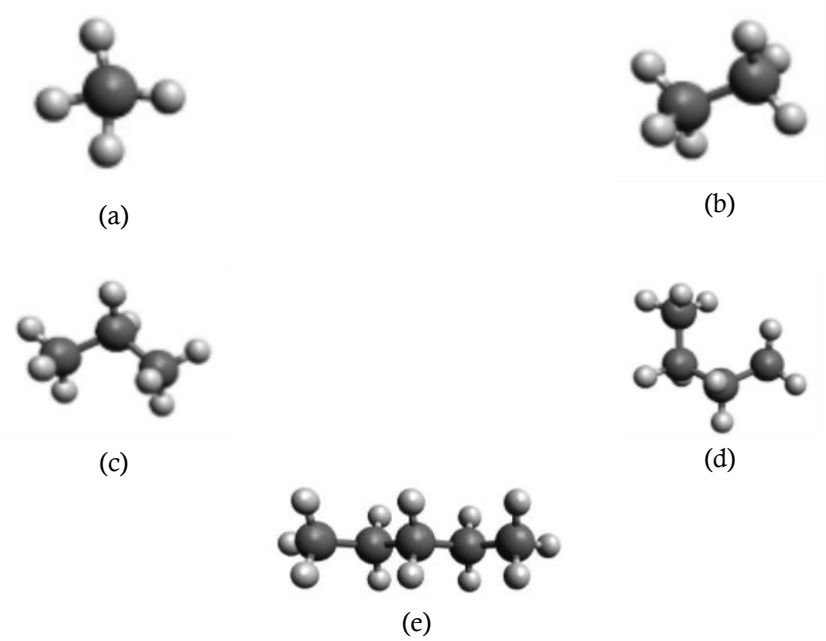

Figure 2 (a) molecular structure of methane, (b) molecular structure of ethane, (c) molecular structure of propane, (d) molecular structure of butane, (e) molecular structure of pentane.

Based on the basis set used and compared with the database used, the BL3YP/cc-PVTZ basis set shows the smallest results compared to other databases. The more accurate the basis set used, the better the results. In this case, the smaller the energy produced, the more accurate the basis set used. The selection of a basis set in computational calculations is very important because it determines the accuracy of the calculation results within the limits of the related theoretical method.

The selection of the right basis set requires a good understanding of the system being studied, the problem being faced, the level of precision required, the acceptable computation time, and the program used. The influence of the basis set is very important in molecular modeling. Differences in the use of basis sets for an element or molecule can affect the calculation results obtained. For this reason, the use of the most appropriate basis set will provide the most optimal results.

For each compound to be calculated, the accurate basis set can vary, such as research conducted by (Myllys et al., 2016) using sulfuric acid, ammonia, dimethylamine, formic acid and water compounds with the DFT method concluded that the cc-pV5Z and cc-pV6Z basis sets provide the most accurate results and are very close to experimental values, especially for the ωB97X-D function. Furthermore, research conducted by (Pérez-Barcia et al., 2023) using a larger cc-pVTZ

basis set did increase the convergence of the total interaction energy but did not increase the convergence of the different interaction energy components. The addition of the basis set resulted in a general increase in electrostatic energy, while the polarization energy decreased and Pauli changed slightly. These results are in accordance with the importance of using a large basis set to accurately represent molecular electric polarization. Further research conducted by (Nabil et al., 2023), concluded that the most accurate basis set for calculating acetamide (AA) and N-methyl acetamide (NMA) adsorbed on the kaolinite surface is (1) B-311++G(2df,2pd), providing interaction energy close to the complete basis set (CBS) extrapolation value with a maximum time estimate of around 4%. (2) This Dunning's correlation-consistent basis set also has good performance producing interaction energy that matches the CBS extrapolation value.

CONCLUSION

The selection of basis sets in computational calculations plays a very important role in the accuracy of theoretical calculations of molecular properties. The higher the basis set used, the more accurate the calculation results will be. In this study, halogen acid and alkane molecules (c1-c5) were used. Using the DFT method and the BL3YP functional, the calculation results and comparisons with the database, the cc-pVTZ basis set is the basis set that produces the most stable calculations compared to the STO-3G, 3-21G, and cc-pVDZ basis sets. The lower the energy of a compound, the more stable its geometric structure will be. The order of stability in halogen acids is HI> HBr> HCl> HI. In alkanes is Pentane> Butane> Propane> Ethane> Methane.

REFERENCES

Asmuruf, F., Futwembun, A., & Kopeuw, A. J. (2017). Karakteristik Struktur Elektronik Pada Senyawa Goniotalamin, Infektokaryon Dan Kurzicalkolakton Menggunakan Metode Density Functional Theory (DFT). *AVOGADRO Jurnal Kimia*, *1*(1).

Coşkun, M., & Ertürk, M. (2022). Double hyperbolic cosine *basis set*s for LCAO calculations. *Molecular Physics*, *120*(17). https://doi.org/10.1080/00268976.2022.2109527

Daga, L. E., Civalleri, B., & Maschio, L. (2020). Gaussian *Basis set*s for Crystalline Solids: All-Purpose *Basis set* Libraries vs System-Specific Optimizations. *Journal of Chemical Theory and Computation*, *16*(4). https://doi.org/10.1021/acs.jctc.9b01004

Grabowski, S. J. (2017). New type of halogen bond: Multivalent halogen interacting with π- and σ-electrons. *Molecules*, *22*(12). https://doi.org/10.3390/molecules22122150

Housecroft, C., & Sharpe, A. (2012). Inorganic Chemistry (Catherine E. Housecroft and Alan G. Sharpe). In *Pearson*.

Jensen Frank. (2017). Introduction to Computational Chemistry, 3rd Edition | Wiley. In *Wiley*.

JOHNSON III, & RUSSELL, D. (2020). *NIST Computational Chemistry Comparison and Benchmark Database*. NIST Standard Reference Database Number 101. https://cccbdb.nist.gov/

Kirschner, K. N., Reith, D., & Heiden, W. (2020). The performance of Dunning, Jensen, and Karlsruhe *basis set*s on computing relative energies and geometries. *Soft Materials*, *18*(2–3). https://doi.org/10.1080/1539445X.2020.1714656

Lehtola, S. (2021). Straightforward and Accurate Automatic Auxiliary *Basis set* Generation for Molecular

Calculations with Atomic Orbital *Basis set*s. *Journal of Chemical Theory and Computation*, *17*(11). https://doi.org/10.1021/acs.jctc.1c00607

Morgante, P., & Peverati, R. (2020). The devil in the details: A tutorial review on some undervalued aspects of density functional theory calculations. In *International Journal of Quantum Chemistry* (Vol. 120, Issue 18). https://doi.org/10.1002/qua.26332

Myllys, N., Elm, J., & Kurtén, T. (2016). Density functional theory *basis set* convergence of sulfuric acid-containing molecular clusters. *Computational and Theoretical Chemistry*, *1098*. https://doi.org/10.1016/j.comptc.2016.10.015

Nabil, N. A. M., Ang, L. S., & Sulaiman, S. (2023). *Basis set* Effects on the Stabilities and Interaction Energies of Small Amide Molecules Adsorbed on Kaolinite Surface. *Indonesian Journal of Chemistry*, *23*(4). https://doi.org/10.22146/ijc.79795

Nagy, B., & Jensen, F. (2017). *Basis sets in Quantum Chemistry*. https://doi.org/10.1002/9781119356059.ch3

Pérez-Barcia, Á., Cárdenas, G., Nogueira, J. J., & Mandado, M. (2023). Effect of the QM Size, *Basis set*, and Polarization on QM/MM Interaction Energy Decomposition Analysis. *Journal of Chemical Information and Modeling*, *63*(3). https://doi.org/10.1021/acs.jcim.2c01184

Petrucci, R. H., Harwood, W. S., Herring, G. E., & Madura, J. (2013). General Chemistry: Principles & Modern Applications. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).

Piela, L. (2006). Ideas of Quantum Chemistry. In *Ideas of Quantum Chemistry*. https://doi.org/10.1016/B978-0-444-52227-6.X5000-5

Susanti, E. D., Junia, S. R., Fadilla, R. N., & Supardi, A. (2020). The Effect of *Basis set* on Quantum Tunneling Probability with the Case of trans-HCOH Isomerization. *Journal of Physics: Conference Series*, *1445*(1). https://doi.org/10.1088/1742-6596/1445/1/012010

Yuniar, Simatupang, E., Putri, A., & Marwati, Y. (2019). PEMODELAN ISOMERISASI STRUKTUR MOLEKUL C6H14 MELALUI STUDI KOMPUTASI. *CHEMICA : Jurnal Pendidikan Kimia Dan Ilmu Kimia*, *2*(1).