Data Article

In silico geometric and energetic data of all possible simple rotamers made of non-metal elements

Taweetham Limpanuparb\textsuperscript{a,*}, Sopanant Datta\textsuperscript{a}, Kridtin Chinsukserm\textsuperscript{a}, Peerayar Teeraniramitr\textsuperscript{b,c}

\textsuperscript{a} Mahidol University International College, Mahidol University, Salaya, Nakhon Pathom 73170, Thailand
\textsuperscript{b} Mahidol University International Demonstration School, Mahidol University, Salaya, Nakhon Pathom 73170, Thailand
\textsuperscript{c} Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand

\begin{abstract}

This article presents theoretical data on geometric and energetic features of halogenated rotamers of the following backbone structures, C–C, N–N, P–P, O–O, S–S, N–P, C–N, C–P, C–O, C–S, N–O, N–S, P–O and P–S. The data is considered to be comprehensive combinations of non-metal elements in the form abcx–ydef whereby a,b,c,d,e,f are halogen (fluorine to iodine), hydrogen or a lone pair and x,y are carbon, nitrogen, phosphorus, oxygen and sulfur. Data were obtained from \textit{ab initio} geometry optimization and frequency calculations at HF, B3LYP, MP2 and CCSD levels of theory on 6-311++G(d,p) basis set. In total, 8535 non-enantiomeric structures were produced by custom-made codes in Mathematica and Q-Chem quantum chemical package. Extracted geometric and energetic data as well as raw output files, codes and scripts associated with the data production are presented in the data repository.

© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license. (\url{http://creativecommons.org/licenses/by/4.0/})

\end{abstract}

\begin{keywords}

Rotamers
Haloethanes
Relative stability
Steric effects
\end{keywords}

\begin{articleinfo}

Article history:
Received 30 December 2019
Revised 24 February 2020
Accepted 10 March 2020
Available online 19 March 2020

Keywords:
Rotamers
Haloethanes
Relative stability
Steric effects
\end{articleinfo}

\* Corresponding author.
E-mail address: taweetham.lim@mahidol.edu (T. Limpanuparb).
Specifications table

| Subject            | Chemistry         |
|--------------------|-------------------|
| Specific subject area | Physical and theoretical chemistry/spectroscopy |
| Type of data       | Tables and Q-Chem output files |
| How data were acquired | Quantum chemical computation on Q-Chem 5.2.1, developer version |
| Data format        | Raw and analyzed |
| Parameters for data collection | Hartree-Fock (HF)/6-311++G(d,p), Becke, 3-parameter, Lee–Yang–Parr (B3LYP)/6-311++G(d,p), Second order Møller–Plesset perturbation theory (MP2)/6-311++G(d,p), Coupled Cluster Singles and Doubles (CCSD)/6-311++G(d,p) |
| Description of data collection | Data were obtained from ab initio geometry optimization and frequency calculations. In total, 8535 non-enantiomeric structures were produced and processed by custom-made codes. |
| Data source location | Mahidol University, Salaya, Thailand |
| Latitude and longitude: 13.792790, 100.325707 |
| Data accessibility | Repository name: mendeley.com |
| Data identification number: DOI: 10.17632/m2h6yg9nzp |
| Direct URL to data: https://data.mendeley.com/datasets/m2h6yg9nzp |

Value of the data

- The origin of energetic preference for staggered structure in ethane [1-5] and gauche structure in 1,2-difluoroethane [6-11] has long been debated and sometimes controversial. The comprehensive data set presented in this article fills the gap in the literature and can be used for further analysis and discussion in relevant topics such as gauche effect [7,8,12] and bent bond [7,13].
- Similar to cis effect where the cis or (Z) isomer is more stable than trans or (E) isomer [14] and relative stability of positional isomers of substituted benzenes [15], gauche effect is demonstrated in this data set by many examples where steric hindrance alone fails to account for the observed relative stability trend.
- For reference purpose, 15665 rotamers are identified with internal numbering, SMILES and PubChem CID. (Out of 15665 rotamers, 1713 rotamers (11%) are identified with CID, of which only 631 are unique.) These can be used in future theoretical or experimental work involving two-center non-metal rotamers.
- Source codes and raw data are available for reproduction of the work and further analysis. For example, molecular dipole moment and vibrational spectrum can be extracted from the raw output. Source codes can be used to generate molecules of related classes for further calculation.

1. Data description

There are 15 folders for C–C, N–N, P–P, O–O, S–S, N–P, O–S, C–N, C–P, C–O, C–S, N–O, N–S, P–O and P–S. In each folder, there are four subfolders for four different methodologies, HF, B3LYP, MP2 and CCSD. In addition to raw output files (.out) and geometry in Z-matrix and Cartesian coordinate format (.xyz), the following summary table files (.csv) are provided in each subfolder:

- A single csv file in xyz subfolder containing geometric data of 7 bond lengths in Å, 12 bond angles and 9 torsional angles in degree (if lone pair(s) are involved, there will be less numbers of geometric parameters and ‘de’ is shown in place of a numerical value.)
- Energetic data, in separate csv files, include electronic energy \(E_{\text{elec}}\) in a.u. (Hartree), thermal correction to enthalpy \(H_{\text{corr}}\) in kcal mol\(^{-1}\), zero-point vibrational energy \(E_{\text{ZPE}}\) in kcal mol\(^{-1}\) and entropy \(S\) in cal mol\(^{-1}\) K\(^{-1}\).

An example of these data is shown in Fig. 1. Names for compounds exist in two different formats and due to symmetry, there are up to six ways to write these out regardless of the format. Therefore, a rotamer name may not exactly match a file name in many instances. Source codes, scripts and examples are provided in a separate folder.
2. Experimental design, materials, and methods

Exhaustive listing of all rotamers can be done in many different approaches. We completed our comprehensive lists of all rotamers by extending the approach we have used for substituted benzenes [15]. Rotamers as viewed by Newman projection can be equivalent to substituted benzenes with two additional conditions. First, the list of substituent elements must include a lone pair of electrons. Second, rotamers are less symmetric compared to benzenes with regards to rotation and flipping. Q-Chem 5.2.1 [16], IQmol 2.13 [17] and Wolfram Mathematica 12.0 [18] were used in the same way as described previously [15]. In addition to compounds in Tables 1–8, Preliminary calculations were also completed for all combination of single atom of C, N, P, O, S and hydrogen/halogen atoms from F to I.

Tables 1–8 provide a comprehensive listing of all rotamers considered in this work. The listing is first organized by the number of substituent elements and pattern of empirical formulas. An explanation on how to calculate the number of chemical empirical formulas in each table is given in Table 9. Rotamer structures are also listed for each pattern. Each rotamer structure can be rotated three times unless it is symmetric (cannot rotate) or has chiral center(s) (∗2 for each center). Asterisks (∗) shown in Tables 1, 2, 4, 6–8 indicate chiral centers. There are two special cases of meso compounds in Tables 1 and 2 which have a reduced number of rotamer structures. Similar symmetrical cases were also found in previous study of substituted benzenes [15]. Since enantiomeric structures are identical in energy, only one of the two enantiomeric structures is considered for each pair. Table 10 provides an overview of all computational jobs described in this paper.

All optimization jobs converged in Q-Chem before geometry information were extracted. The converged results are not necessarily the same conformer as the input. For example, some anti forms are turned into gauche forms during the geometry optimization. Almost all of the converged rotamers were confirmed to be local minima by frequency calculation (16,729 out of 16,840 jobs). However, there were 111 frequency jobs with exactly one imaginary frequency
### Table 1
List of 2675 possible C–C rotamers in 210 formulas (1505 non-enantiomeric rotamers).

| Number of elements | Empirical formula (Distribution of elements) | Number of empirical formulas | Rotamer structure | Rotamers per formula | Number of rotamers |
|--------------------|---------------------------------------------|------------------------------|-------------------|----------------------|-------------------|
| 1                  | $C_2\alpha_6$ (6)                          | 5                            | $\alpha_3C-C\alpha_3$ | 1                    | 1                 |
|                    |                                             |                              |                   | 5                    | 5                 |
| 2                  | $C_2\alpha_5\beta$ (1–5)                    | 20                           | $\alpha_2\beta C-C\alpha_3$ | 1                    | 1                 |
|                    | $C_2\alpha_2\beta_4$ (2–4)                  | 20                           | $\alpha_2\beta C-C\beta_2$ | 3                    | 2                 |
|                    | $C_2\alpha_3\beta_3$ (3–3)                  | 10                           | $\alpha_2\beta C-C\beta_3$ | 1                    | 1                 |
| 3                  | $C_2\alpha_4\beta_3\gamma_4$ (1–1–4)       | 30                           | $\alpha_2\beta\gamma C-C\gamma_3$ | 2                    | 1                 |
|                    |                                             |                              |                   | 60                   | 30                |
|                    | $C_2\alpha_2\beta_3\gamma_3$ (1–2–3)       | 60                           | $\alpha_2\beta\gamma C-C\beta_2\gamma_2$ | 3                    | 2                 |
|                    |                                             |                              |                   | 360                  | 180               |
|                    | $C_2\alpha_2\beta_2\gamma_2$ (2–2–2)       | 10                           | $\alpha_2\beta\gamma C-C\alpha\beta\gamma_2$ | 9                    | 5                 |
|                    |                                             |                              |                   | 90                   | 60                |
| 4                  | $C_2\alpha_5\beta\gamma_3\delta_3$ (1–1–1–3) | 20                           | $\alpha_2\beta\delta C-C\gamma_3\delta_2$ | 18                   | 9                 |
|                    |                                             |                              |                   | 360                  | 180               |
|                    | $C_2\alpha_2\beta_2\gamma_2\delta_2$ (1–1–2–2) | 30                           | $\alpha_2\beta\gamma C-C\beta_2\gamma_2\delta_2$ | 12                   | 6                 |
|                    |                                             |                              |                   | 360                  | 180               |
|                    |                                             |                              |                   | 360                  | 180               |
| 5                  | $C_2\alpha_4\beta\gamma_3\delta\varepsilon_3$ (1–1–1–2) | 5                           | $\alpha_2\beta\gamma C-C\gamma_3\delta\varepsilon_3$ | 36                   | 18                |
|                    |                                             |                              |                   | 180                  | 90                |

### Table 2
List of 975 possible N–N (or P–P) rotamers in 70 formulas (500 non-enantiomeric rotamers).

| Number of elements | Empirical formula (Distribution of elements) | Number of empirical formulas | Rotamer structure | Rotamers per formula | Number of rotamers |
|--------------------|---------------------------------------------|------------------------------|-------------------|----------------------|-------------------|
| 1                  | $N_2\alpha_4$ (4)                          | 5                            | $\alpha_2N-N\alpha_2$ | 3                    | 2                 |
|                    |                                             |                              |                   | 15                   | 10                |
| 2                  | $N_2\alpha_3\beta$ (1–3)                    | 20                           | $\alpha_2\beta N-N\beta_2$ | 6                    | 3                 |
|                    | $N_2\alpha_2\beta_2$ (2–2)                  | 10                           | $\alpha_2\beta N-N\alpha\beta$ | 9                    | 5                 |
|                    |                                             |                              |                   | 120                  | 60                |
|                    |                                             |                              |                   | 90                   | 50                |
| 3                  | $N_2\alpha_4\beta_3\gamma_3$ (1–1–2)       | 30                           | $\alpha_2\gamma N-N\gamma_2$ | 3                    | 2                 |
|                    |                                             |                              |                   | 30                   | 20                |
| 4                  | $N_2\alpha_4\beta\gamma_3\delta$ (1–1–1–1) | 5                            | $\alpha_2\beta\gamma N-N\gamma_3\delta$ | 36                   | 18                |
|                    |                                             |                              |                   | 180                  | 90                |
ranging from 4.9i cm⁻¹ to 513.1i cm⁻¹. The list of files is provided in data folder. Most of them (74) are HF calculations and the rest (37) are B3LYP. Upon closer inspection, all rotamers of O–O, S–S and O–S with an imaginary frequency (55) are in an anti form. Similarly, 28 of 34 rotamers of N–N, P–P and N–P with an imaginary frequency are in an anti form from the perspective of the two lone pairs. As these observations suggest that the anti form is not stable, gauche effect is evident in these classes of compounds.

**Competing Interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
## Table 6
List of 3125 possible C-N (or C-P) rotamers in 126 formulas (1625 non-enantiomeric rotamers).

| Number of elements | Empirical formula (Distribution of elements) | Number of empirical formulas | Rotamer structure | Rotamers per formula | Number of rotamers |
|--------------------|---------------------------------------------|----------------------------|-------------------|---------------------|-------------------|
|                    |                                             |                            |                   | Inclusive of        | Excessive of      | Inclusive of      | Excessive of      |
|                    |                                             |                            |                   | enantiomers         | enantiomers       | enantiomers      | enantiomers      |
| 1                  | CNa₃ (5)                                    | 5                          | α₃C–Na₂           | 1                   | 1                 | 5               | 5               |
| 2                  | CNaβ₄ (1–4)                                 | 20                         | αβ₄C–Nβ₂          | 3                   | 2                 | 60              | 40              |
|                    |                                             |                            | β₃C–Nαβ           | 2                   | 1                 | 40              | 20              |
|                    |                                             |                            | αβ₃C–Nαβ          | 6                   | 3                 | 120             | 60              |
|                    |                                             |                            | α₂βC–Nβ₂          | 3                   | 2                 | 60              | 40              |
|                    |                                             |                            | β₃C–Na₂           | 1                   | 1                 | 20              | 20              |
| 3                  | CNaβγᵢ₃ (1–1–3)                             | 30                         | αβγᵢC–Nγᵢ₂       | 6                   | 3                 | 180             | 90              |
|                    |                                             |                            | βᵢC–Nαγ           | 12                  | 6                 | 360             | 180             |
|                    |                                             |                            | γᵢC–Nαβ           | 2                   | 1                 | 60              | 30              |
|                    |                                             |                            | αβγᵢC–Nᵢβγ       | 12                  | 6                 | 360             | 180             |
|                    |                                             |                            | βᵢγC–Nᵢαγ         | 12                  | 6                 | 360             | 180             |
|                    |                                             |                            | αβᵢC–Nγᵢ₂         | 6                   | 4                 | 180             | 120             |
|                    |                                             |                            | αγᵢC–Nβᵢ₂         |                      |                   |                 |                 |
| 4                  | CNaβγδ₂ (1–1–1–2)                           | 30                         | αβδC–N¹γδ         | 36                  | 18                | 720             | 360             |
|                    |                                             |                            | βγδC–N¹αδ          | 36                  | 18                | 720             | 360             |
|                    |                                             |                            | αβγδC–Nœδ₂         | 6                   | 3                 | 120             | 60              |
|                    |                                             |                            | βγδC–N¹αγ          | 18                  | 9                 | 360             | 180             |
|                    |                                             |                            | γδC–Nαβ           |                      |                   |                 |                 |
| 5                  | CNaβγδε (1–1–1–1)                           | 1                          | αβγδεC–Nεδε       | (10)                |                   | 120             | 60              |
|                    |                                             |                            | (10)              |                     |                   | 120             | 60              |

## Table 7
List of 625 possible C-O (or C-S) rotamers in 70 formulas (375 non-enantiomeric rotamers).

| Number of elements | Empirical formula (Distribution of elements) | Number of empirical formulas | Rotamer structure | Rotamers per formula | Number of rotamers |
|--------------------|---------------------------------------------|----------------------------|-------------------|---------------------|-------------------|
|                    |                                             |                            |                   | Inclusive of        | Excessive of      | Inclusive of      | Excessive of      |
|                    |                                             |                            |                   | enantiomers         | enantiomers       | enantiomers      | enantiomers      |
| 1                  | COα₄ (4)                                    | 5                          | α₃C–Oα            | 1                   | 1                 | 5               | 5               |
| 2                  | COαβ₃ (1–3)                                 | 20                         | αβ₃C–Oβ           | 3                   | 2                 | 60              | 40              |
|                    |                                             |                            | β₃C–Oα            | 1                   | 1                 | 20              | 20              |
|                    |                                             |                            | αβ₃C–Oα           | 6                   | 4                 | 60              | 40              |
| 3                  | COαβγᵢ (1–1–2)                              | 30                         | αβγᵢC–Oγᵢ        | 6                   | 3                 | 180             | 90              |
|                    |                                             |                            | αγᵢC–Oβᵢ         | 6                   | 4                 | 180             | 120             |
|                    |                                             |                            | βγᵢC–Oαᵢ         |                      |                   |                 |                 |
| 4                  | COαβγδ (1–1–1–1)                            | 5                          | αβγδC–Oδ         | 24                  | 12                | 120             | 60              |
### Table 8
List of 375 possible N–O (or N–S, P–O, P–S) rotamers in 35 formulas (200 non-enantiomeric rotamers).

| Number of elements | Empirical formula (Distribution of elements) | Number of empirical formulas | Rotamer structure | Inclusive of enantiomers | Exclusive of enantiomers | Inclusive of enantiomers | Exclusive of enantiomers |
|--------------------|---------------------------------------------|-----------------------------|------------------|------------------------|------------------------|------------------------|------------------------|
| 1                  | NOαβγ (1–2–3)                              | 5                           | αβN–Oγ           | 3                      | 2                      | 15                     | 10                     |
| 2                  | NOαβγ (1–1–1)                              | 5                           | αβN–Oγ           | 3                      | 2                      | 120                    | 60                     |
| 3                  | NOαβγ (1–1–1)                              | 10                          | αβN–Oγ           | 18                     | 9                      | 180                    | 90                     |

### Table 9
Examples for number of empirical formula calculation.

| Number of elements | Empirical formula (Distribution of elements) | Number of empirical formulas$^a$ |
|--------------------|---------------------------------------------|---------------------------------|
| 3                  | C$_2$αβγδε (1–2–3)                          | k = 3, n$_1$ = 1, n$_2$ = 1, n$_3$ = 1 therefore $\frac{5!}{3!} = 60$ |
| 4                  | C$_2$αβγδε (1–1–1–1–2)                     | k = 5, n$_1$ = 4, n$_2$ = 2 therefore $\frac{5!}{4!} = 30$ |

$^a$ The number of empirical formulas is calculated by using the expression $(\frac{\prod}{\prod})$, where
- $\prod$ is the number of possible substituent elements (H, F, Cl, Br, I),
- $\prod$ is the actual number of substituent elements and
- $!$ is the product of the factorial of the number of substituent elements with the same subscript.

### Table 10
Summary of 43450 computational jobs included in this paper (opt for geometry optimization and freq for frequency calculation).

| Class of compound | Number of rotamers | HF opt | B3LYP opt | MP2 opt | CCSD opt |
|-------------------|--------------------|--------|-----------|---------|----------|
| Preliminary CNPOS| 175                | 170    | All       | All     | All      | All      | All      |
| C–C               | 2675               | 1505   | All       | All     | All      | All      | 23       |
| N–N               | 975                | 500    | All       | All     | All      | All      | 30       |
| P–P               | 975                | 500    | All       | All     | All      | All      | 30       |
| O–O               | 45                 | 30     | All       | All     | All      | All      | All      |
| S–S               | 45                 | 30     | All       | All     | All      | All      | All      |
| C–N               | 3125               | 1625   | All       | All     | All      | All      | 33       |
| C–O               | 3125               | 1625   | All       | All     | All      | All      | 33       |
| C–S               | 625                | 375    | All       | All     | All      | All      | 33       |
| N–P               | 1875               | 950    | All       | All     | All      | All      | 34       |
| N–O               | 375                | 200    | All       | All     | All      | All      | 34       |
| N–S               | 375                | 200    | All       | All     | All      | All      | 34       |
| P–O               | 375                | 200    | All       | All     | All      | All      | 34       |
| P–S               | 375                | 200    | All       | All     | All      | All      | 34       |
| O–S               | 75                 | 50     | All       | All     | All      | All      | All      |
| Total             | 15840              | 8535   | 8535      | 8535    | 8535     | 110      | 665      |
Acknowledgments

We are grateful to MUIC for financial support. We thank Wanutcha Lorpaiboon for her initial input in the early stage of this work.

References

[1] V. Pophristic, L. Goodman, Hyperconjugation not steric repulsion leads to the staggered structure of Ethane, Nature 411 (2001) 565–568.
[2] F. Weinhold, A new twist on molecular shape, Nature 411 (2001) 539–541.
[3] P.R. Schreiner, Teaching the right reasons: lessons from the mistaken origin of the rotational barrier in ethane, Angew. Chem. 41 (2002) 3579–3582.
[4] F.M. Bickelhaupt, E.J. Baerends, The case for steric repulsion causing the staggered conformation of ethane, Angew. Chem. 42 (2003) 4183–4188.
[5] Y. Mo, W. Wu, L. Song, M. Lin, Q. Zhang, J. Gao, The magnitude of hyperconjugation in ethane: a perspective from ab initio valence bond theory, Angew. Chem. 43 (2004) 1986–1990.
[6] D. Wu, A. Tian, H. Sun, Conformational properties of 1,3-difluoropropene, J. Phys. Chem. A. 102 (1998) 9901–9905.
[7] L. Goodman, H. Gu, V. Pophristic, Gauche effect in 1,2-difluoroethane. hyperconjugation, bent bonds, steric repulsion, J. Phys. Chem. A. 109 (2005) 1223–1229.
[8] N.C. Craig, A. Chen, K.H. Suh, S. Klee, G.C. Mellau, B.P. Winnewisser, M. Winnewisser, Contribution to the study of the gauche effect. The complete structure of the anti rotamer of 1,2-difluoroethane, J. Am. Chem. Soc. 119 (1997) 4789–4790.
[9] D. O’Hagan, Understanding organofluorine chemistry. An introduction to the C–F bond, Chem. Soc. Rev. 37 (2008) 308–319.
[10] M.A.X. Muir, A systematic density functional study of fluorination in methane, ethane and ethylene, Mol. Phys. 89 (1996) 211–237.
[11] Á. Ganyecz, M. Kállay, J. Csontos, High accuracy quantum chemical and thermochemical network data for the heats of formation of fluorinated and chlorinated methanes and ethanes, J. Phys. Chem. A. 122 (2018) 5993–6006.
[12] E.L. Eliel, S.H. Wilen, L.N. Mander, Stereochemistry of Organic Compounds 1994, Wiley & Sons, Inc., New York, 1994.
[13] K.B. Wiberg, Bent bonds in organic compounds, Acc. Chem. Res. 29 (1996) 229–234.
[14] K. Chinsukserm, W. Lorpaiboon, P. Teeraniramitr, T. Limpanuparb, Geometric and energetic data from ab initio calculations of haloethene, haloline, halomethyleneephosphine, haloinophosphine, halodiazene, halodiphosphene and halocyclopropane, Data Brief 27 (2019) 104738.
[15] S. Datta, T. Limpanuparb, Geometric and energetic data from quantum chemical calculations of halobenzenes and xylenes, Data Brief. (2020) 105386, doi:10.1016/j.dib.2020.105386.
[16] Y. Shao, Z. Gan, E. Epifanovsky, A.T. Gilbert, M. Wormit, J. Kussmann, A.W. Lange, A. Behn, J. Deng, X. Feng, Advances in molecular quantum chemistry contained in the Q-Chem 4 program package, Mol. Phys. 113 (2015) 184–215.
[17] A.T.B. Gilbert, IQmol, http://iqmol.org, 2019.
[18] Wolfram Research Inc, Mathematica, Champaign, Illinois, 2019.