

### Abstract

Protein-ligand reactions commonly exhibit halogen bonds of the form C-X... $\pi$ . The polarizability of halogen atoms forms a positive region called a  $\sigma$ -hole that can interact with electronegative molecules. This study investigated the influence of dispersion between the thirty-one independent bromine, chlorine, and hydrogen structures with ammonia, formaldehyde, and benzene. The Hartree-Fock calculation utilized Dispersion Functional Theory along with three functionals and two basis sets. The calculated interaction energies were compared to benchmark CCSD(T)/aug-cc-pVQZ values. Results showed that dispersion correction significantly improved molecular modeling outcomes. The study also indicated that running calculations with the functional BLYP and dispersion correction D3 produces a reasonably accurate result given the computational cost.

### Background

Halogen bonds are prevalent in protein-ligand interactions. These biological interactions commonly have a C-X... $\pi$  form. The polarizability of halogen atoms leads to the formation of a positive region on the halogen atom known as a  $\sigma$ -hole (sigma-hole). As demonstrated in Fig. 1, the positive region of the halogen complex (C-X) will interact with the electronegative region of the benzene ring ( $\pi$ ).

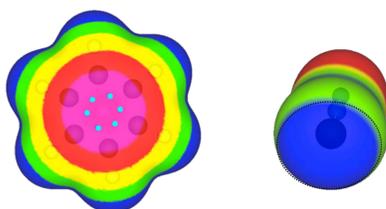


Figure 1: Electron density map for BrCN...C<sub>6</sub>H<sub>6</sub>

Furthermore, the size of the halogen atoms causes dispersion to significantly influence molecular interaction. Consequently, when modeling the C-X... $\pi$  interaction with the Hartree-Fock method in accordance to Dispersion Functional Theory (DFT), dispersion correction is required.

Many studies have investigated halogen interaction with common electronegative atoms such as oxygen and nitrogen. Although C-X... $\pi$  formations are common in nature, halogen interaction with structures like benzene remain understudied.

In an effort to model these structures and compare the accuracy of various functionals using the DFT calculation method is computationally efficient and accurate. A variety of functionals are compatible with the DFT calculation method. Furthermore, the computational cost of the basis set is reasonably balanced with the accuracy of the results.

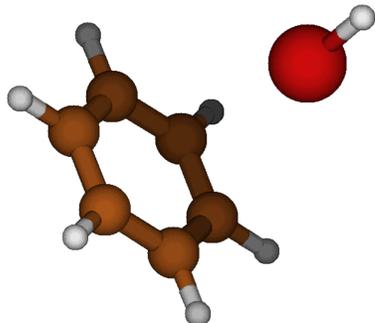


Figure 2: HBr...C<sub>6</sub>H<sub>6</sub> complex

### Methods

The study investigated the independent interaction of thirty-one bromine, chlorine, and hydrogen structures with ammonia, formaldehyde, and benzene ( $\pi$  groups). All bimolecular interactions were prepared at the MP2/TZVP level of theory. Instances of Hartree-Fock calculations used either TZVP basis sets or QZVP basis sets compatible with DFT. The study also investigated the versatility of the common functionals Becke-Lee-Yang-Parr (BLYP), Becke-3-Parameter-Lee-Yang-Parr (B3LYP), and Becke-07 (B97). Calculations also integrated dispersion correction with D2 and D3 methods. The interaction energy was determined from the subsequent dimer energy values (Eq. 1).

$$\Delta E_{int} = E_{AB} - E_A - E_B$$

Eq. 1: Interaction energy for bimolecular interactions.

The resulting interaction energy values for the pairs of functionals and dispersion corrections were compared to benchmark values determined from CCSD(T)/aug-cc-pVQZ methods.

### Conclusions

The Hartree-Fock calculations reveal intriguing traits about the various functionals and dispersion correction with regard to  $\pi$  groups. Across all molecule interactions, lack of dispersion correction caused egregious error (Table 1). The significant reduction in error introduced after dispersion correction supports the  $\sigma$ -hole phenomenon. Additionally, the BLYP functional had the most accurate calculation when coupled with dispersion correction. While the TZVP basis set is less computationally demanding, it may reduce accuracy as compared to the QZVP basis set. Interestingly, side-by-side comparison of TZVP calculations with QZVP demonstrates a reasonable average percent error as low as 15% as compared to the QZVP's 5%.

From the tested  $\pi$  groups, the average percent error increases from ammonia to formaldehyde to benzene (Table 2, Fig. 3). After introducing dispersion correction, the error significantly decreased across all  $\pi$  groups. This trend shows the significance of dispersion in the C-X... $\pi$  interactions.

The comparison of deviation reveals that without dispersion correction, the calculation tends to underbind the molecular interaction. Introducing dispersion correction leads to a tendency to overbind. (Table 3)

Concluding from various functional and dispersion correction methods in molecular modeling, a researcher using limited computational resources can confidently use the TZVP BLYP-D3 method to produce results with an acceptable average percent error of 15%. For researchers with access to better computer resources, the QZVP BLYP-D3 method will produce results with an average percent error of only 5%.

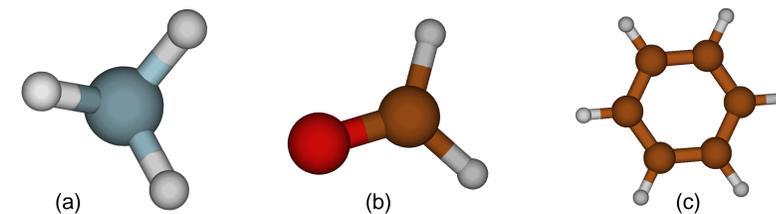


Figure 3:  $\pi$  groups in increasing order of dispersion influence: (a) ammonia, (b) formaldehyde, and (c) benzene.

### Results

Average Percent Error

Functional-Dispersion correction	def2-TZVP	def2-QZVP
BLYP	87.17 %	96.99 %
BLYP-D3	14.39 %	5.08 %
B3LYP	69.48 %	76.67 %
B3LYP-D3	16.24 %	7.40 %
B97	53.48 %	59.29 %
B97-D2	35.33 %	27.33 %
B97-D3	22.44 %	18.63 %

Table 1: Average percent errors for functional and dispersion correction pairs. Values are separated into TZVP and QZVP basis set categories. Error determined by comparing computed values and benchmark CCSD(T)/aug-cc-pVQZ values.

Table 2: Average percent errors for the  $\pi$  groups ammonia, formaldehyde, and benzene. Values are separated into TZVP and QZVP basis set categories. Error determined by comparing computed values and benchmark CCSD(T)/aug-cc-pVQZ values.

Functional-Dispersion correction	Average percent error					
	def2-TZVP			def2-QZVP		
	Ammonia	Formaldehyde	Benzene	Ammonia	Formaldehyde	Benzene
BLYP	17.10 %	48.97 %	136.92 %	32.24 %	58.06 %	144.78 %
BLYP-D3	25.17 %	8.47 %	12.63 %	8.48 %	1.08 %	5.60 %
B3LYP	14.96 %	34.73 %	110.72 %	22.72 %	42.37 %	117.43 %
B3LYP-D3	26.10 %	12.73 %	13.68 %	11.03 %	5.09 %	6.97 %
B97	10.77 %	27.05 %	85.38 %	14.79 %	33.88 %	91.46 %
B97-D2	31.86 %	20.92 %	44.05 %	18.16 %	14.09 %	37.97 %
B97-D3	35.94 %	24.08 %	15.71 %	22.25 %	17.25 %	17.73 %

Table 3: Average computed value deviation from benchmark CCSD(T)/aug-cc-pVQZ values. Negative values result from overbinding, while positive corresponds to underbinding.

Functional-Dispersion correction	Average deviation					
	def2-TZVP			def2-QZVP		
	Ammonia	Formaldehyde	Benzene	Ammonia	Formaldehyde	Benzene
BLYP	0.26915	1.23293	3.57758	0.92148	1.46976	3.79067
BLYP-D3	-0.98888	-0.22984	-0.34266	-0.33655	0.00700	-0.12957
B3LYP	0.00880	0.83902	2.86147	0.59497	1.03736	3.04613
B3LYP-D3	-1.03249	-0.37027	-0.39902	-0.44633	-0.17193	-0.21435
B97	-0.17185	0.67105	2.20981	0.37221	0.85077	2.38305
B97-D2	-1.14847	-0.53200	-1.26501	-0.60442	-0.35229	-1.09177
B97-D3	-1.41857	-0.70453	-0.21385	-0.87451	-0.52481	-0.04061

### Future work

The investigation of the significance of dispersion influence for C-X...  $\pi$  interactions can be further improved by testing additional  $\pi$  groups. Further analysis of functional ability can factor running time into the efficiency and accuracy of its use in molecular modeling. Furthermore, new types of functionals should be assessed in the Hartree-Fock calculations.

### References

Riley, K. E., Ford, C. L., & Demouchet, K. (2015). Comparison of hydrogen bonds, halogen bonds, C-H...  $\pi$  interactions, and C-X...  $\pi$  interactions using high-level ab initio methods. *Chemical Physics Letters*, 621, 165-170.

### Acknowledgements

This material is based upon work supported by the National Science Foundation under the NSF EPSCoR Cooperative Agreement No. EPS-1003897 with additional support from the Louisiana Board of Regents.

