

# How different is pyrimidine as a core component of DNA base from its diazine isomers: A DFT study?

Subhojyoti Chatterjee | Feng Wang

Molecular Model Discovery Laboratory,  
Department of Chemistry and  
Biotechnology, Faculty of Science,  
Engineering and Technology, Swinburne  
University of Technology, Hawthorn,  
Melbourne, Victoria 3122, Australia

## Correspondence

Feng Wang, Molecular Model Discovery  
Laboratory, Department of Chemistry and  
Biotechnology, Faculty of Science,  
Engineering and Technology, Swinburne  
University of Technology, Hawthorn,  
Melbourne, Victoria, 3122, Australia.  
Email: fwang@swin.edu.au

## Abstract

Recent photoemission spectroscopic (X-ray photoemission spectra) study revealed less dramatic chemical changes for pyrimidine (PyM, 1, 3-diazine) with in its ionization potential. Present systematic study using density functional theory calculations shows that PyM is indeed quite different from its diazine isomers (PyD, 1, 2-diazine and PyA, 1, 4-diazine). It is discovered that the most stable isomer PyM is relaxed from  $C_{2v}$  to  $C_1$  point symmetry with a total electronic energy deduction of  $-15.86 \text{ kcal.mol}^{-1}$ . Although not substantial, PyM has the smallest molecule shape (electronic spatial extent) and the largest HOMO-LUMO energy gap of 5.65 eV; only one absorption band in the region of 200–300 nm of the UV-Vis spectrum but three clusters of chemical shift in the carbon and hydrogen NMR spectra. The energy decomposition analyses revealed that the interaction energy ( $\Delta E_{\text{int}}$ ) of PyM is preferred over PyA by  $4.08 \text{ kcal.mol}^{-1}$  and over PyD by  $22.32 \text{ kcal.mol}^{-1}$ , with the preferred N—C—N bond revealed by graph theory.

## KEYWORDS

chemical graph, DNA base thymine, energy decomposition analysis, nuclear magnetic resonance, pyrimidine, UV-Vis absorption, X-ray photoemission spectroscopy

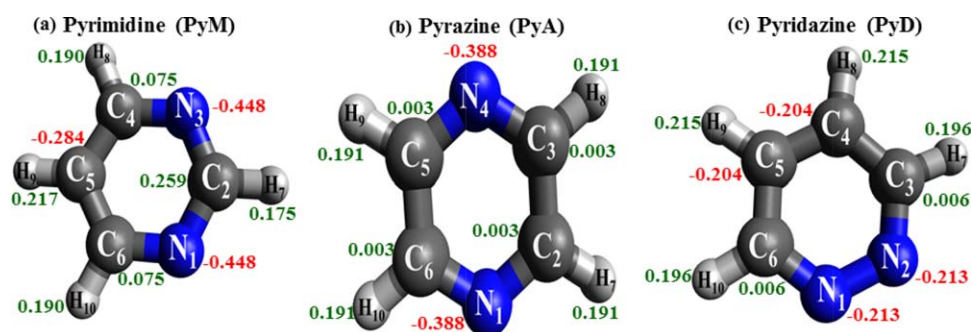
## 1 | INTRODUCTION

Deoxyribonucleic acid (DNA) bases, as the building blocks of many important natural and synthetic compounds, have been studied for many years.<sup>[1–3]</sup> The pyrimidine bases, cytosine (C), thymine (T), and uracil (U) (U for RNA) share a pyrimidine pentagon ring fusion<sup>[4,5]</sup> and thus, pyrimidine (PyM) forms the core component of all DNA bases.<sup>[6–8]</sup> In addition to DNA bases, PyM moiety is also found in other natural derivatives like vitamin B1 and alloxan as well as in synthetic compounds such as the anti-HIV drug, zidovudine.<sup>[9–11]</sup>

Previous photoemission spectroscopic (X-ray photoemission spectra [XPS]) study of DNA bases<sup>[12–14]</sup> revealed that compared to purine DNA bases, pyrimidine DNA bases are engaged with less dramatic chemical changes in its ionization potentials (IP). The Pyrimidine ring is subject to diamine conformers, that is, Pyridazine (PyD) and Pyrazine (PyA), that is, aromatic heterocyclic isomers with nitrogens at positions 1 and 3 for PyM, 1 and 2 for PyD, and 1 and 4 for PyA.<sup>[6]</sup> Normally, the diazines are described as  $\pi$ -deficient and any kind of electronegative group substitution significantly increases the  $\pi$ -deficiency, making a decrease in its basicity. This facilitates reduction in resonance stability; leading to ring cleavage reaction than electrophilic aromatic substitution,

although nucleophilic aromatic substitution is preferred.<sup>[15]</sup> Conversely, unlike PyM, PyA is a symmetrical molecule which favors the electrophilic aromatic substitution.<sup>[16]</sup> Like PyM, PyA also has diverse pharmacological activities such as anticancer or anti-tuberculosis, along with forming an important structural flavoring unit for food products.<sup>[17]</sup> In contrast to its other aromatic ring isomers, the most unstable isomer PyD has a rare occurrence in nature. Due to its application as a pharmacophore unit, PyD are usually industrial synthesised via 1,4-diketones or 4-ketoacids with hydrazine.<sup>[18]</sup> In addition, PyD also makes an essential core structure of several drugs such as cefozopran, cadralazine, cilazapril, and minaprine.<sup>[19–21]</sup>

It is well-known that structure determines properties of molecules in chemistry. Although the entire heterocyclic diazine ring is the structural positional isomers of each other,<sup>[22]</sup> the challenge of the decade lies on the structural preference of pyrimidine (PyM) as a core component of DNA base over its other isomeric form, Pyrazine (PyA) and Pyridazine (PyD) and thus a systematic study on its molecular properties is crucial for better chemistries and building blocks of life.<sup>[23]</sup> A number of previous studies of the diazine isomers concentrated on the properties such as outermost  $n$  or  $\pi$  orbital (Rydberg) valence transition energies,<sup>[24–27]</sup> the oxidative potential or discrete energy levels and to



**FIGURE 1** Optimized structures of diazine isomers in their ground electronic states using B3LYP/aug-cc-pVTZ. The natural bond orbital (NBO) charges are present in parentheses

energy bands,<sup>[28]</sup> single-hole states,<sup>[29]</sup> and vertical electron affinities, etc.<sup>[30,31]</sup> Thus, a systematic study of Pyrimidine,<sup>[12–14,32,33]</sup> and its isomers, Pyrazine and Pyridazine, is needed to help in the understanding of why nature makes PyM the key component of DNA base.

## 2 | COMPUTATIONAL METHODS AND DETAILS

The density functional theory (DFT) model employed is the same as in the previous study for pyrimidine and purine bases.<sup>[12–14,32]</sup> That is, B3LYP/aug-cc-pVTZ is employed for geometric calculations in the present study for the diazines. This model has been proven a good model to study DNA bases and produce results which agree well with experiments, although they were developed and optimized for post-HF method. Additionally, cc-pVTZ is better than 6-311G (d, p) or similar and being a trimmed basis set they are smaller than augmented basis set, making them computationally faster and suitable for the present study.<sup>[12,32,34–39]</sup> Based on the obtained geometries, the NMR chemical shift was calculated using the M06-2X/6-31 + G\* model, after it's comparison with the internal reference standard, tetra-methyl silane (TMS). The Gauge invariant atomic orbital (GAIO) method was employed to calculate the chemical shift of <sup>13</sup>C, <sup>15</sup>N, and <sup>1</sup>H NMR for the diazine isomers.

Time dependent density functional theory (TD-DFT) has been used to calculate the UV-Vis spectrum in methyl cyclohexane solution as well as in gas phase using B3LYP/aug-cc-pVTZ model.<sup>[40]</sup> The lowest 60 excited states of singlet-singlet transitions were employed for the absorption spectra calculation. Absorption profiles were simulated using a half-width half maximum (HWHM) of .11eV, 1.5 eV, and .06 eV for pyrimidine, pyrazine, and pyridazine, respectively. The conductor-like polarisable continuum solvent model (CPCM) was employed.

The valence spectra of the diazines were calculated using the outer valence Green's function (OVGF) theory.<sup>[41–45]</sup> The complete valence binding energy (vertical ionization potential) spectra of diazine isomers were calculated using the DFT based SAOP/et-pVQZ model.<sup>[32,46,47]</sup> The asymptotic correct LB94/et-pVQZ<sup>[12,48]</sup> model was used to calculate the vertical core ionization energies of diazine isomers. Further, the orbital momentum profiles of the isomers as orbital

momentum distributions (MDs)<sup>[49–53]</sup> were calculated with the momentum cut-off at 10 a.u.<sup>[54]</sup>

Energy decomposition analysis (EDA) based on Morokuma<sup>[55]</sup> is employed to reveal the isomer differences<sup>[56–67]</sup> and the extended transition state (ETS) of Ziegler and Rauk.<sup>[68]</sup>

$$\Delta E_{\text{Int}} = \Delta E_{\text{elect}} + \Delta E_{\text{Pauli}} + \Delta E_{\text{orb}} \quad (1)$$

Where,  $\Delta E_{\text{Int}}$  is the stabilizing term called interaction energy that includes the instantaneous interaction between electrostatic energy ( $\Delta E_{\text{elect}}$ ), quantum mechanical Pauli energy ( $\Delta E_{\text{Pauli}}$ ) and orbital interaction energy ( $\Delta E_{\text{orb}}$ ) based on atomic "fragments."<sup>[56,69–71]</sup> The calculation are produced using the B3LYP/TZ2P model.<sup>[68,72]</sup> All quantum mechanical calculations are performed using the Gaussian 09 (G09) computational chemistry package<sup>[73]</sup> and the Amsterdam density functional (ADF) computational chemistry program.<sup>[74–76]</sup>

The diazine isomers are also studied using graph theory which has been applied to study biomolecules.<sup>[35,51]</sup> The LDU decomposition (Equation 2), that is, lower, diagonal, and upper triangular matrix<sup>[77,78]</sup> was constructed based on the Doolittle algorithm<sup>[79]</sup> for numerical analysis of these diazine isomers,

$$PAQ = LU. \quad (2)$$

Where A is a square matrix; P is a permutation matrix; Q is a permutation matrix that reorders the columns of matrix A and LU are lower and upper triangular matrices, and A = LDU, where D is a diagonal matrix.

## 3 | RESULTS AND DISCUSSION

### 3.1 | Geometrical and other molecular properties

The diazine molecule isomers (C<sub>4</sub>H<sub>4</sub>N<sub>2</sub>), that is, Pyrimidine (PyM), Pyrazine (PyA), and Pyridazine (PyD) are hexagon aromatic compounds, consisting of four carbon atoms and two nitrogen atoms. Figure 1 gives the chemical structures of the isomers, where the positions of the N atoms determine the isomers. The total electronic energies of the three isomers indicate that the central component of DNA base PyM is the most stable isomer among the three structures. Table 1 compares the selected properties of the isomers and validates with available theoretical and experiment results.<sup>[1,80–82]</sup> Table S1 of Supporting Information

TABLE 1 Comparison of selected molecular properties of pyrimidine, pyrazine, and pyridazine<sup>a</sup>

Properties	Pyrimidine (PyM)	Pyrazine (PyA)	Pyridazine (PyD)
Sym. (State)	C <sub>1</sub> (X1A)	C <sub>2v</sub> (X1A1)	C <sub>s</sub> (X1A')
$\Delta E$ (E <sub>h</sub> )	0.0 <sup>b</sup>	4.390	22.590
R6 (Å)	8.106	8.109	8.152
RA (GHz)	6.318	6.440	6.290
RB (GHz)	6.126	5.990	6.024
RC (GHz)	3.110	3.103	3.077
<R <sup>2</sup> > (a.u.)	409.399	409.980	412.984
$\mu$ (Debye)	2.360 (2.330) <sup>c</sup>	.000	4.240
ZPE (kcal.mol <sup>-1</sup> )	48.219	48.091	47.699
$\Delta\epsilon$ (HOMO – LUMO (eV))	5.650	5.280	4.870

<sup>a</sup>Using the B3LYP/aug-cc-pVTZ model.<sup>b</sup>E<sub>h</sub> = −165,926.05 (kcal.mol<sup>-1</sup>); R6 denotes the ring perimeter of the isomer; <R<sup>2</sup>> is the electronic spatial extent;  $\mu$  the dipole moment.<sup>c</sup>Expt. dipole moment<sup>32</sup>; ZPE is zero-point energy; {RA, RB, RC} are the Rotational Constants; HOMO is highest occupied molecular orbital and LUMO is lowest un-occupied molecular orbital.

provide the fully geometric parameters for the isomers. A good agreement between the experimentally measured and theoretically calculated bond lengths (pm) and angle (°) is achieved with the mean unsigned error being 1.04 pm, .75 pm, and .17 pm as well as .19°, .08°, and .04° for PyM, PyA, and PyD, respectively. The major differences between the isomer pair are produced from the positions of the N–C bonds locally. For example, the ring perimeter<sup>[83]</sup> of PyM, PyA, and PyD are given by 8.106 Å, 8.109 Å, and 8.152 Å, respectively, varying only as small as .04 Å. PyM exhibits the smallest ring perimeter among the isomers.

The electronic spatial extent, <R<sup>2</sup>>, is a predictive size of a molecule and reflects the compact nature of the structure. The <R<sup>2</sup>> values exhibit that PyM has the smallest size among its diazine isomers (409.399 a.u.) followed by PyA of 409.980 a.u. and PyD has the largest size of 412.984 a.u. The size changes of the isomers are also revealed by their rotational constants A, B, and C. For example, the rotational constants A, B, and C of PyM are given by 6.318 GHz, 6.126 GHz, and 3.110 GHz, respectively, indicate that the diazine is a flat molecule. A noticeable difference of the diazine isomers is the dipole moment, depending on the symmetry. For example, PyD has the largest dipole moment of 4.24 Debye, but the symmetric PyA does not have permanent dipole moment. The natural bond orbital (NBO)<sup>[84]</sup> charges are also provided on the structures in Figure 1. The NBO of the N atoms of PyM (−.44e) possess the most negative charges among the isomers, whereas the NBO on the nitrogen atoms of PyA and PyD are less negative with −.38e and −.21e, respectively. More interestingly, PyM has the least balance of the NBO negative charges. For example, the negative NBO charges of nitrogen atoms of PyM (−.44e) are balanced by

C<sub>(5)</sub> on the other side of the ring with NBO of −.284e; whereas the negative NBO charges of the nitrogens of PyD (−.213e) are better balanced by C<sub>(4)</sub> and C<sub>(5)</sub> of −.204e on the other side of the ring, and the NBOs of the nitrogens (−.388e) of PyA are well balanced by each other as they are located on the opposite sides of the ring.

### 3.2 | Chemical environment of the diazines

The chemical environments for the nitrogens of the diazines are identical in the same isomer, whereas the carbon environments of PyM and PyD split into three and twofold. As all carbons in PyA are equivalent so that they do not split in the C1s spectra of PyA. As a result, the chemical shift of the carbon and nitrogens of the diazines in their core ionization spectra and NMR spectra are more or less predictable. Figure 2 compares the calculated C-NMR, H-NMR, and N-NMR of the three diazines with available experimental result. The agreement is excellent as shown in the parentheses in blue next to the shift, with the mean sign error being only 1.56 ppm.

In the isomers, each carbon atom only bonds one hydrogen atom, so that the chemical shift of the C-NMR and H-NMR exhibit similar patterns. However, for the C-NMR spectra, the carbons directly bond with nitrogens, that is, C<sub>(2)</sub> (159.08 ppm) and C<sub>(4)</sub> (156.92 ppm) of PyM has the largest chemical shift among all the C-NMR of the isomers, followed by the pair of carbons directly connecting to the nitrogens in PyD, C<sub>(3)</sub> and C<sub>(6)</sub> at 152.25 ppm, the four equivalent carbons, C<sub>(2)</sub>, C<sub>(3)</sub>, C<sub>(5)</sub>, and C<sub>(6)</sub> in PyA at 144.93 ppm, then the other pair of carbons, C<sub>(4)</sub> and C<sub>(5)</sub> in PyD at 126.66 ppm, and finally C<sub>(5)</sub> in PyM at 117.93 ppm which does not directly bond with any nitrogens. The C-NMR

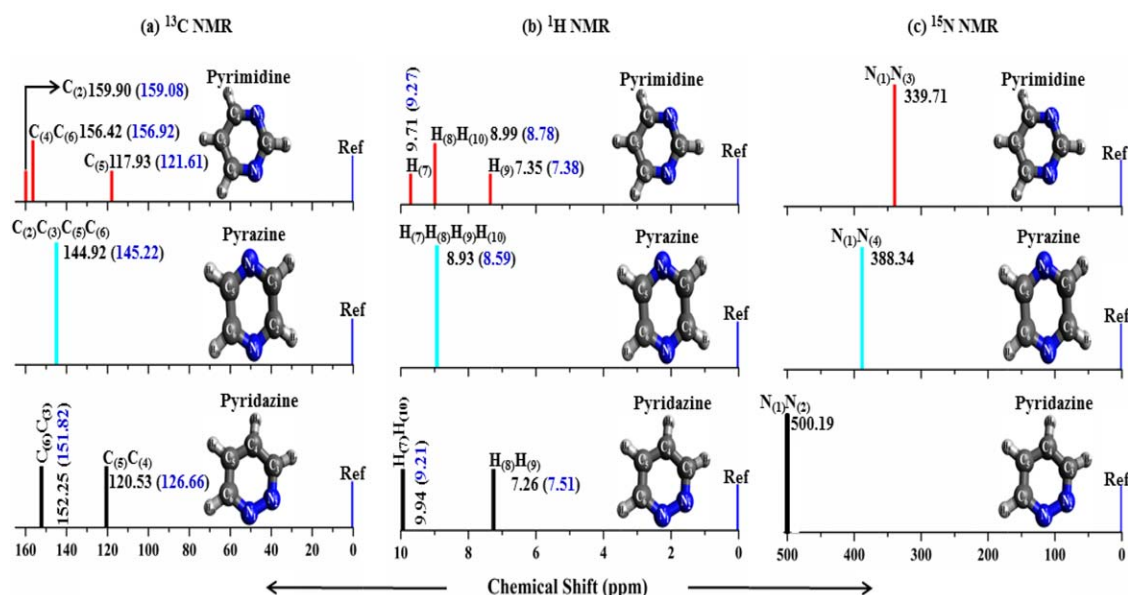


FIGURE 2 Comparison of the experimental and calculated carbon, proton and nitrogen NMR spectra of diazine isomers using the M06-2X/6-31 + G\* model. The values in the bracket represent the experimental shift

( $\Delta\delta = 41.97$  ppm) and H-NMR ( $\Delta\delta = 1.98$  ppm) of PyM exhibits the largest chemical shift among the diazines, which indicates that the chemical environment of the C—H groups in PyM are more diverse than its other diazine isomers. Conversely, the calculated chemical shift of the N-NMR of PyM is the smallest ( $\delta = 339.71$  ppm), in comparison to its other diazine isomers and hence the chemical environment of the nitrogens in PyM is less strong. The different chemical environment of the C and N atoms in the diazine isomers can be studied using the XPS<sup>[12,32]</sup> (Table S2 Supporting Information), where the carbons bonded directly with nitrogens in PyM possess the C1s ionization energy larger than 291 eV, whereas the C1s ionization energies of all other carbon atoms in the diazine isomers are less than 291 eV.

### 3.3 | Valence binding energy spectra and HOMO momentum distribution

Chemical information and bonding of the molecule are dominated by information from valence space. Because of its delocalized character with small energy intervals, it is more complex than the core space. Figure 3 compares the valence IP spectra of the diazines using the same model as previous study.<sup>[12]</sup> The electronic structural associations of the three valence spectra are apparent in the figure. For example, the first IP of the diazine isomers is highest with 9.88 eV for PyM, followed by 9.71 eV, and 9.34 eV for PyA and PyD, respectively. As seen from the spectra, the valence space electronic structures of the diazines are quite different: the three inner valence orbitals (12a, 13a, and 14a) of PyM in the region of IP > 16 eV are well separated; but the three orbitals of PyM in the mid-valence region of 16–14 eV are closely located to produce a single peak, whereas the outer valence region of IP < 14 eV splits into a closed located pair of two orbitals (18a–19a) followed by the well separated HOMO-1 (20a) and HOMO (21a), as shown in Figure 3. Table S3 Supporting Information reports the valence shell

binding energy of diazine isomers,<sup>[23,26]</sup> which are compared with available literature and measurement results. The agreement with experiment is in general good. Although in the inner valence regions, the single particle approximation in the OVGF model breaks in case of PyA due to the spectroscopic pole strength being below the cut-off of .85,<sup>[85]</sup> the statistically averaged orbital potential (SAOP) model

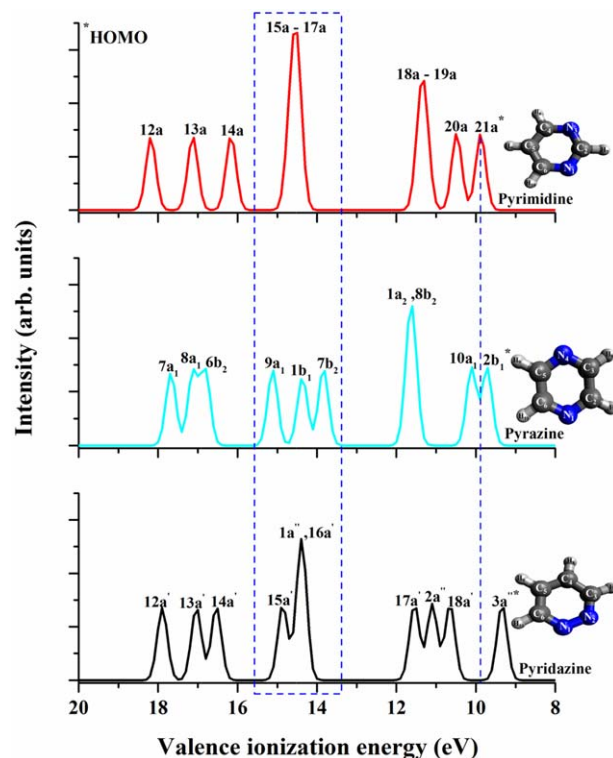
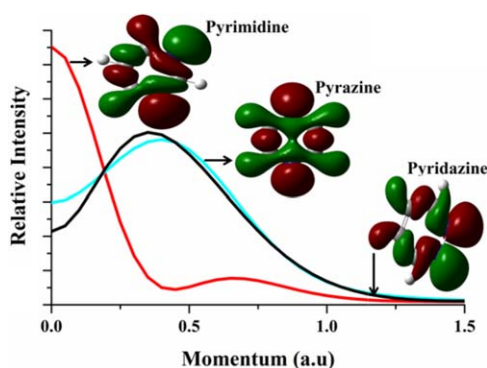


FIGURE 3 Comparison of simulated valence binding energy spectra of diazine isomers using the OVGF/aug-cc-pVTZ model





**FIGURE 4** Theoretical momentum distribution (TMD) of diazine isomers of the highest occupied molecular orbital (HOMO)

continues to produce good agreement to the experiment, as observed in the valence structure of diazine isomers PyM, PyA, and PyD. In general, the inner valence region of IP > 16 eV of PyM, PyA, and PyD are significantly similar and are dominated by the C2s orbitals. Here, it should be also noted that, the IP energies of the inner valence space are only estimations.

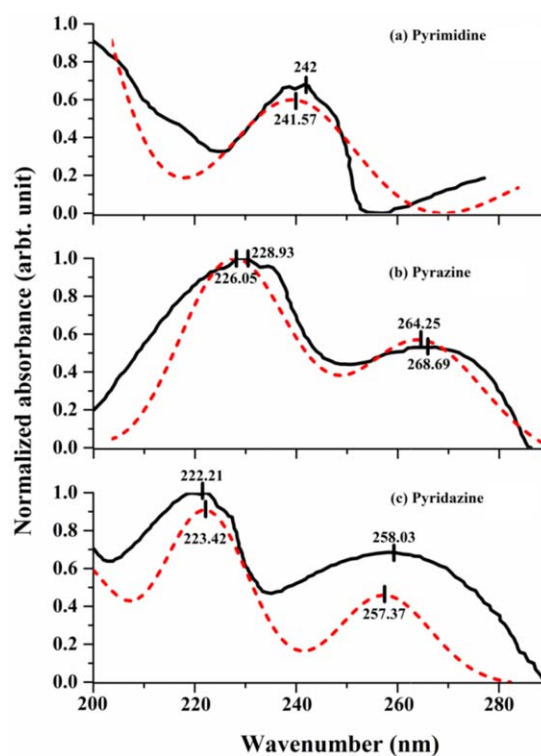
The IP of the highest occupied molecular orbital (HOMO, 21a) of PyM is the highest ionization energy (9.88 eV) among the isomers. To further probe the information of HOMO, the theoretical momentum distributions (TMD, i.e., the cross sections) of the HOMO diazine isomers are compared in Figure 4 using dual space analysis (DSA).<sup>[86]</sup> All theoretical cross-section have been scaled to have identical maximum vertices with reference to PyM, thus permitting comparison for the shape rather than for absolute cross-section. It was interesting that the momentum distribution of the HOMO's of PyM exhibit half-bell shaped cross-sections, indicating strong s-electron dominance with hybridized orbital nature and exhibiting more anti-bonding character with the maximum cross section for the major peak located at  $\sim 0.5$  a.u. Conversely, the TMD differences in the HOMOs for PyA and PyD exhibit bell shaped curved, dominated by p-like orbitals.<sup>[86]</sup> This distinctive different distributions of HOMOs for PyM versus PyA and PyD can be seen by the molecular electron density distributions, also presented in the same figure. Although the valence orbitals are delocalized in the entire molecule, the HOMO of PyM (21a) exhibits fewer nodal planes in comparison to its diazine isomers PyA (2b<sub>1</sub>) and PyD (3a''), hereby maximizing the orbital density of the orbital. In the momentum region of  $P < 0.5$  a.u., orbital TMD of the HOMOs of PyM shows opposite features to its other isomers, PyA and PyD. For example, the TMD is maximum at  $P = 0$  a.u. for PyM but minima for PyA and PyD; whereas at approximately  $P = 0.4$  a.u., the TMD of PyM exhibits a minimum but a maximum for PyA and PyD.

### 3.4 | UV-Vis spectra and electronic structure of diazines

The valence electronic information of the isomers revealed in the previous section indicated that the different valence electronic

structures of the diazines will lead to quite different UV-Vis spectra. But UV-Vis absorption spectrum provides result of transitions among the occupied molecular orbitals and the virtual orbitals.<sup>[45]</sup> Figure 5 compares the calculated UV-Vis absorption spectra of PyM, PyA, and PyD in the region of 200–300 nm in methyl cyclohexane solution with the measurements.<sup>[87,88]</sup> Note that the TD-DFT calculated spectra has been blue shifted by 30 nm as the relative peak positions are important. The relative UV-Vis spectral peak positions are in excellent agreement with the measurements in methyl cyclohexane solution. For example, the two spectral peaks UV-Vis spectra of PyA are 39.66 nm apart in measurement, whereas this separation is calculated by 38.20 nm; similarly, the measurement is 35.82 nm and the calculated is 33.95 nm for PyD.

The UV-Vis spectrum of PyM, Figure 5(a) shows only a single band at 241.57 (14) nm in the calculations, with the measured spectrum (black) being 242 nm.<sup>[87]</sup> The less stable isomers PyA and PyD both show two bands at approximately 226 (.45) nm and 264 (.15) nm for PyA; 223 (.23) nm and 257 (.49) nm for PyD. Table S4 Supporting Information reports the TD-DFT simulated UV-Vis spectra of PyM, PyA, and PyD *in silico* in a broader UV-Vis region of 190–300 nm. The transition bands in the region under 300 nm are contributions from many transitions among the low energy frontier orbitals, such as HOMO-*N* and LUMO + *M*, where *N* and *M* being small integers (1, 2, 3



**FIGURE 5** Comparison of the observed 87,88 and calculated UV-Vis spectra in methyl cyclohexane of diazine isomers (a) Pyrimidine (HWHH = .11eV), (b) Pyrazine (HWHH = 1.5 eV), and (c) Pyridazine (HWHH = .06eV) using B3LYP/aug-cc-pVTZ model. The simulated spectra were shifted by 30nm

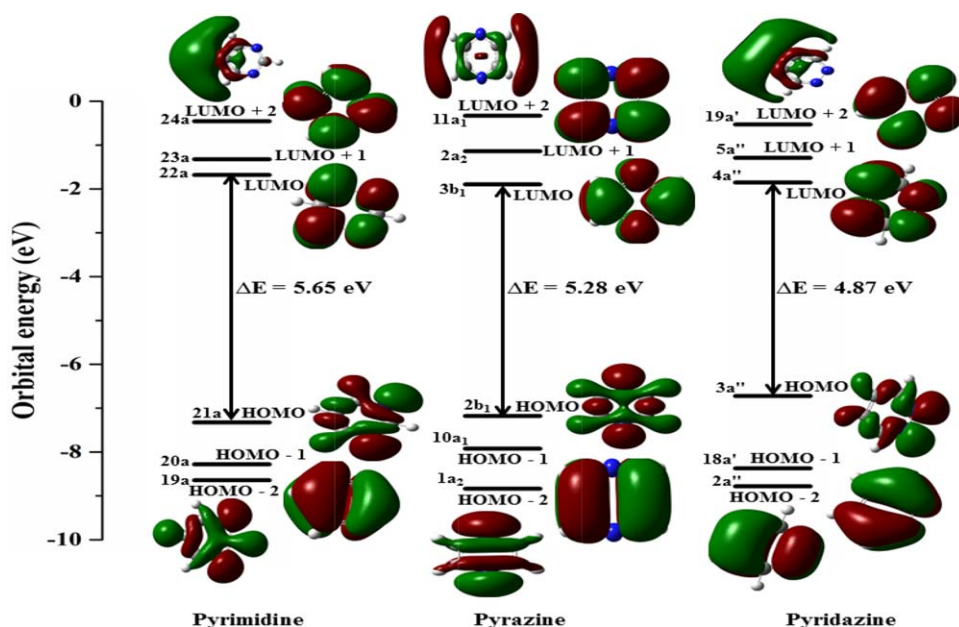


FIGURE 6 Frontier orbitals energy diagrams of diazine isomers at the B3LYP/aug-cc-pVTZ level of theory. Orbital numbering and symmetry are indicated

....7), representing frontier orbital energy levels (Transition shown in Table S4 Supporting Information are those  $> 5\%$ ). The band positions in this region are not due to any single transitions. The major transitions with oscillator strength  $f > .12$  are also given in the Table S4 Supporting Information. As seen in the table, the simulated major two bands  $\lambda_1$  and  $\lambda_2$  in PyA and PyD were observed with approximately 35 nm splitting for the diazine isomers. Interestingly, the HOMO-LUMO transitions for PyM, PyA, and PyD locate in the region above 300 nm which outside of the measurement,<sup>[87]</sup> that is, the transitions at 310.69 nm of PyM, at 340.89 nm of PyA, and 356.05 nm of PyD, are dominated by the HOMO-LUMO transition but with a small oscillatory strength ( $f$ ) of  $\sim .006$  and thus was not considered in the present study.

To further inspect such transitions, frontier orbitals diagrams from HOMO-2 (occupied) to LUMO + 2 (unoccupied) of PyM, PyA, and PyD, respectively, together with their corresponding orbital distributions were represented in Figure 6. As shown in the figure, the energy gaps between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), is the largest for PyM (5.65 eV) among the diazine isomers, whereas the HOMO-LUMO energy gap for PyD is the smallest (4.87 eV) in the diazine isomers (for a more comprehensive information, refer to Table S4 Supporting Information). While in general the smaller the allowed transition energy the larger is the probability of the transitions, not all transitions are allowed by symmetry. Relaxation of the structure of PyM from a higher symmetry  $C_{2v}$  to a lower symmetry  $C_1$  can make the unoccupied orbitals more accessible, as the HOMO and LUMO of PyM in higher symmetry structure may be not allowed. A couple of competitive transitions, that is, an  $n \rightarrow n^*$  transition of HOMO - 1 (20a) to LUMO (22a), and an  $n \rightarrow \sigma^*$  transition of HOMO (21a) to LUMO + 2 (24a) are more accessible

for PyM. For example, the first energy gap ((HOMO - 1) - LUMO) is given by 6.60 eV, whereas the next energy gap (HOMO - (LUMO + 2)) for PyM is given by 6.83 eV. Thus, these transitions show the dominance of p-orbitals and help in conjugation. Further, as indicated in Table S4 Supporting Information, large contributions from transitions in this region are also due to a number of other transitions among its low lying frontier orbitals, such as the band at 210.44 (.38) nm of PyM exhibits three major transitions, HOMO - 3  $\rightarrow$  LUMO (70%), HOMO - 1  $\rightarrow$  LUMO + 1 (12%) and HOMO  $\rightarrow$  LUMO + 4 (7%) etc. These closely located energies for the HOMO (21a) and HOMO - 1 (20a) of PyM contribute to the first two peaks in the valence binding energy spectrum with a small energy difference of only .48 eV, as shown in Figure 3.

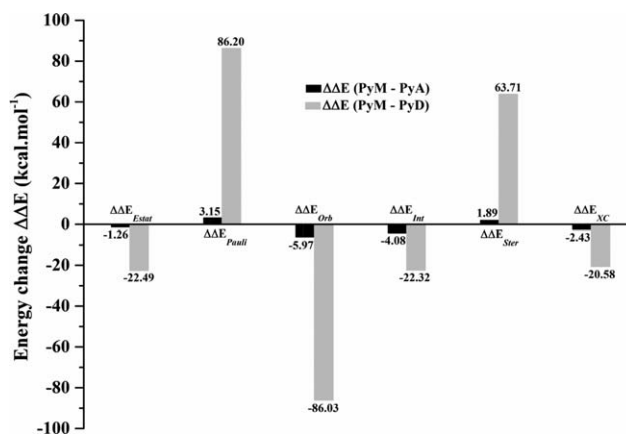





FIGURE 7 Comparison of energy terms for the diazine isomers with energy decomposition

TABLE 2 Comparison of largest eigenvalues and their vector components (absolute values) in case of Pyrimidine, Pyridazine, and Pyrazine

Graph			
	Pyrimidine (PyM)	Pyridazine (PyA)	Pyrazine (PyD)
CASE A: Considering only the C-atoms as nodes			
LEV <sup>a</sup>	8.841	8.590	8.196
Node	VC(LEV) <sup>b</sup>	Node	VC(LEV)
$\lambda_2$	.517	$\lambda_3$	.343
$\lambda_4$	.418	$\lambda_4$	.409
$\lambda_5$	.477	$\lambda_5$	.534
$\lambda_6$	.576 <sup>c</sup>	$\lambda_6$	.563
CASE B: Considering only the N-atoms as nodes			
LEV	2.828	2.236	2.013
Node	VC(LEV)	Node	VC(LEV)
$\lambda_1$	.577	$\lambda_1$	.408
$\lambda_3$	.961@	$\lambda_2$	.831
		$\lambda_4$	.707

<sup>a</sup>Largest eigenvalue.<sup>b</sup>Vector components corresponding to largest eigenvalue.<sup>c</sup>The largest vector component of the largest eigenvalue [LVC (LEV)] is shown in bold for C constructed graph and @bold italic for N constructed graph.

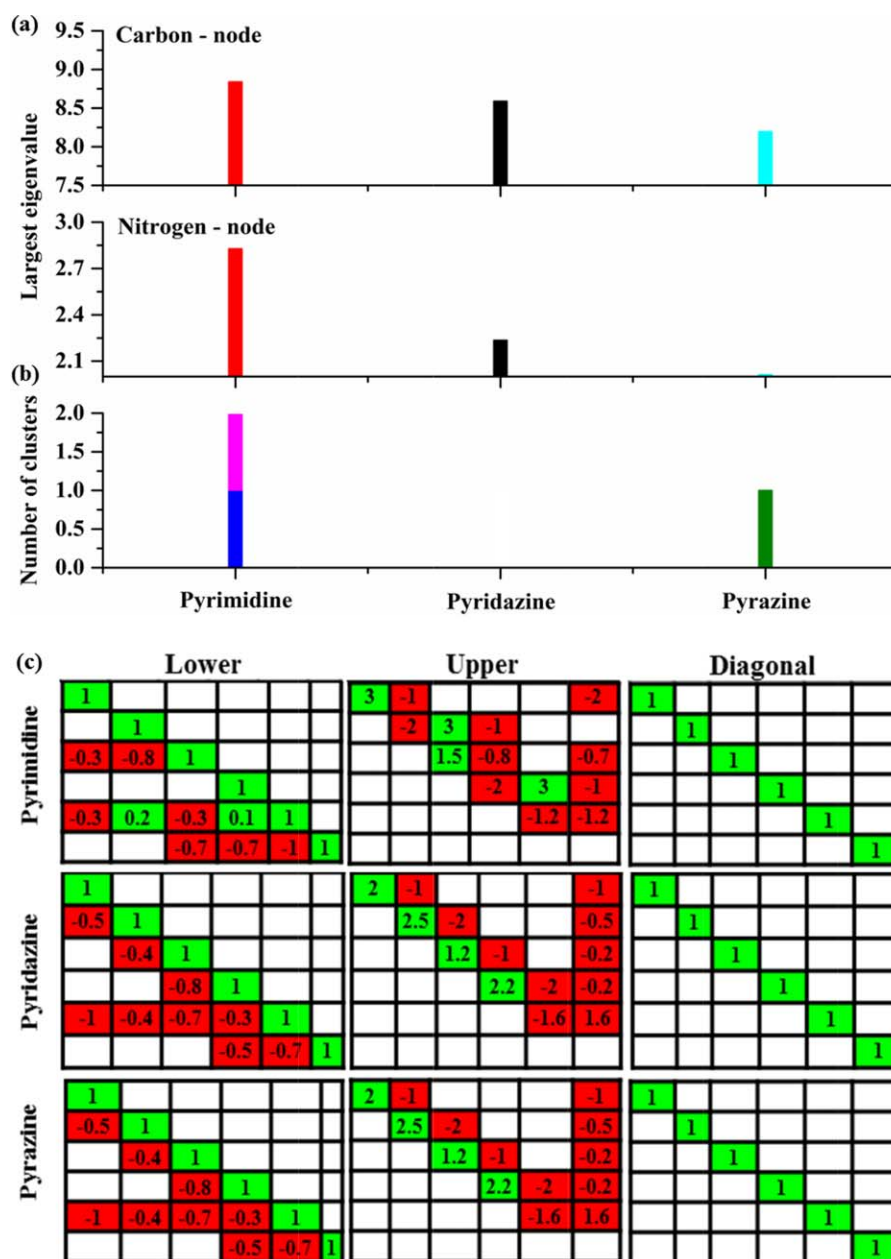
### 3.5 | Energy decomposition analysis

Energy decomposition analysis (EDA)<sup>[89]</sup> is able to reveal the energy differences between the isomers based on the physical origin of the energies. It may be able to provide more information why PyM is selected over its diazine bases PyA and PyD as a DNA base by nature. The total interaction energy ( $\Delta E_{\text{int}}$ ) is the sum of three independent energy components, that is, electrostatic attractive energies ( $\Delta E_{\text{Estat}}$ ), the quantum mechanical Pauli repulsive energy ( $\Delta E_{\text{Pauli}}$ ) and the orbital attractive interaction energy ( $\Delta E_{\text{Orb}}$ ). These energies, together with steric repulsion energy ( $\Delta E_{\text{ster}}$ ), which is the sum of  $\Delta E_{\text{Pauli}}$  and  $\Delta E_{\text{Estat}}$ , and steric exchange component ( $\Delta E_{\text{XC}}$ ), which is the sum of  $\Delta E_{\text{Orb}}$  and  $\Delta E_{\text{ster}}$ , provide further insight of the isomer differences. Figure 7 compares the access energies ( $\Delta \Delta E$ ) between the PyA and PyD with respect to PyM. As seen in this figure, the access energy components between PyA and PyM (the black bars) are significantly smaller than those of PyD and PyM energy components (the grey bars). The total interaction energy change,  $\Delta \Delta E_{\text{int}}$  results in a small negative energy value, indicating that the PyM isomer is the most stable isomer, in agreement with the previous total electronic energy calculations, however, due to different reasons. PyM is more stable than PyA due to large access orbital energy ( $\Delta E_{\text{Orb}}$ ) which experiences the largest energy change of  $-5.97 \text{ kcal.mol}^{-1}$  between the isomers; whereas in comparison to PyD, PyM is more stable due to the large access electrostatic energy  $\Delta E_{\text{Estat}}$  of  $-22.49 \text{ kcal.mol}^{-1}$ , due to the cancelation of large access repulsive Pauli energy  $86.2 \text{ kcal.mol}^{-1}$  and equally large access orbital energy

$\Delta E_{\text{Orb}}$  of  $-86.03 \text{ kcal.mol}^{-1}$ . As a result, PyM and its PyA isomers are more physically similar to each other than PyM and PyD.

### 3.6 | Cluster and LUD decomposition matrix

Finally, chemical graph theory which has been previously applied to biomolecules,<sup>[35,51]</sup> was applied to explore the isomers by constructing distance matrix (DM), to investigate the influence of the nitrogen position over the cyclic aromatic ring. As the information regarding the contribution of each node in the chemical graph is best captured by the vector component (VC) of the largest eigenvalue (LEV),<sup>[90,91]</sup> the C<sub>(6)</sub> of the diazine isomers connected to N<sub>(1)</sub> and C<sub>(5)</sub> influences the largest vector component (LVC) of the LEV (Table 2), where PyM exhibits the highest LEV of .576 followed by PyD (.563) and PyA (.540). If only the N-atoms are considered as nodes (case B), the highest VC (LEV) was observed for the second nitrogen present in the cyclic ring, with N<sub>(3)</sub> for PyM being the highest, that is, .961 (PyA for .707 and PyD for .831), which indicated that the positions of N-atoms of the isomers are in favor of PyM, that is, principal component of the DNA base selection. The LEV for the cyclic aromatic ring, PyM, PyA and PyD for C and N – node DM are distinct and thus can serve as a molecular topological index (MTI), where the score follows the trend of PyM being the highest [see Figure 8(a)]. For example, the quantified MTI for PyM (8.841/2.828), PyD (8.590/2.236), and PyA (8.196/2.013) in accordance with C/N – atom node, shows influence of the isomer PyM over PyA and PyD diazine ring.



**FIGURE 8** Representation of molecular topological indices based on largest eigenvalue for graphs with (a) C-nodes (hydrogen and nitrogen atom suppressed) and N-nodes (hydrogen and carbon atom suppressed), (b) Number of clusters formed based on the VC[SEV], and (c) LUD decomposition matrix

Recognition of cluster and pattern of similar connectivity,<sup>[92–94]</sup> can be studied from the spectral analysis of connected graphs.<sup>[95,96]</sup> According to Vishveshwara et al.,<sup>[97,98]</sup> the vector components of the second smallest eigenvalue carry the information about the cluster forming residues. Figure 8(b) shows the number of clusters for PyM is 2 which is larger than that of PyA (1 cluster) and of PyD (0 cluster).

Chemical graph spectra can also be treated as an essential tool for elucidation of structural differentiation. With the molecules under study, an important arena of exploration can be for designing a signature pattern that would help to differentiate structures that differ only in their nitrogen group positioning and the backbone topology. This phenomenon is captured by the numerical analysis of the LDU decom-

position [Figure 8(c)], where the complete lower and upper triangular matrix carrying the same scores for PyA and PyD. The major difference is that for the PyM, it arises due to the difference in the positions and bonding patterns of the two N atoms w.r.t. to the four C atoms. For example, in case of PyM its N—C—N bond, whereas it is C—N—C and N—N—C bond for PyA and PyD, respectively.

## 4 | CONCLUSIONS

Nitrogen atom positions in the diazine aromatic ring result in the DNA base crucial component PyM and its diazine isomers PyA and PyD. C1s



XPS along with  $^{13}\text{C}$  NMR is sensitive to the changes in the chemical environment of the carbon atoms caused by their interactions with the N atoms and is therefore ideal for isomer study. PyM as the most stable isomer exhibits a relaxation to a  $C_1$  point symmetry, leading to a total electronic energy deduction of  $-15.86 \text{ kcal.mol}^{-1}$  (B3LYP/aug-cc-TZVP). Evidences from the present study indeed indicate that PyM possesses properties which are very different from its isomers. PyM has the smallest molecule shape  $\langle R^2 \rangle$  and the largest HOMO-LUMO energy gap of 5.65 eV. The valence IP spectrum of PyM shows well separated inner valence orbitals (12a–14a), closed positioned mid-valence orbitals (15a–17a) and blue shifted (larger energy) HOMO (21a). The TMD of the HOMO of PyM reveals significant difference between PyM and its other isomers: the HOMO of PyM is more sp-electron dominance whereas the HOMOs of the isomers are dominated by p-electrons. Such very different valence electronic structure also contributes to the UV-Vis spectrum of PyM, which only shows a single band in the region of 200–300nm. The individual chemical environment of PyM and its isomers are also seen in its NMR spectra by their chemical shifts. The present energy decomposition analyses (EDA) revealed that the interaction energy ( $\Delta E_{\text{int}}$ ) of the pentagon ring of PyM is preferred over PyA ( $\Delta \Delta E_i = -4.08 \text{ kcal.mol}^{-1}$ ) and PyD ( $\Delta \Delta E_j = -22.32 \text{ kcal.mol}^{-1}$ ) isomers, being dominated by orbital interaction (PyM-PyA) but electrostatic interaction (PyM-PyD). Finally, the chemical graph, LUD decomposition matrix, has been employed for the first time to study diazine isomers which indicated the preferential N—C—N bond for the DNA interaction, thus making PyM evident as a core component of DNA base.

## ACKNOWLEDGMENTS

FW acknowledges Australia Research Council (ARC) through the Discovery Project (DP) for funding and PhD scholarship for SC. The authors also acknowledge the merit based supercomputing time from National Computational Infrastructure (NCI) and Swinburne University Supercomputing facility. SC thanks Qudsia Arooj and Alex Winnett for assistance.

## REFERENCES

- [1] S. Breda, I. D. Reva, L. Lapinski, M. J. Nowak, R. Fausto, *J. Mol. Struct.* **2006**, 786, 193.
- [2] O. O. Brovarets, D. M. Hovorun, *J. Biomol. Struct. Dyn.* **2015**, 33, 28.
- [3] O. O. Brovarets, D. M. Hovorun, *J. Biomol. Struct. Dyn.* **2015**, 33, 925.
- [4] E. D. Raczynska, M. Makowski, M. Hallmann, B. Kaminska, *RSC Adv.* **2015**, 5, 36587.
- [5] M. E. Hamdi, W. Tiznado, J. Poater, M. Sola, *J. Org. Chem.* **2011**, 76, 8913.
- [6] H. C. Brown, Determination of organic structures by physical methods, E. A. Baude, F. C. Nachod, Eds.; New York, NY: Academic Press, 1955, 6, 5.
- [7] T. L. Gilchrist, *Heterocyclic Chemistry*, New York: Longman, **1997**, 3, 432.
- [8] J. A. Joule, *Heterocyclic compound* (5th ed.), K. Mills, Eds.; Oxford: Wiley, **2010**, 5, 250.
- [9] K. Wright, *Nature* **1986**, 323, 283.
- [10] D. J. Jeffries, *BMJ (Clinical Research Ed.)* **1989**, 298, 1132.
- [11] D. E. Clercq, *Biochem. Pharmacol.* **1994**, 47, 155.
- [12] F. Wang, Q. Zhu, E. Ivanova, *J. Synchrotron Radiat.* **2008**, 15, 624.
- [13] O. O. Brovarets, R. O. Zhurakivsky, D. M. Hovorun, *Phys. Chem. Chem. Phys.* **2014**, 16, 3715.
- [14] O. O. Brovarets, D. M. Hovorun, *Phys. Chem. Chem. Phys.* **2013**, 15, 20091.
- [15] M. Movassaghi, M. D. Hill, *J. Am. Chem. Soc.* **2006**, 128, 14254.
- [16] R. A. Abramovitch, E. M. Smith, Vol. 14, Wiley, New York **1974**, 1.
- [17] M. Asif, *Int. J. Adv. Sci. Res.* **2015**, 1, 05.
- [18] M. Tisler, B. Stanovnik, *Adv. Heterocycl. Chem.* **1968**, 9, 211.
- [19] M. Asif, *Int. J. Pharm. Chem.* **2015**, 05, 397.
- [20] M. Asif, *Chron. Young Sci.* **2010**, 1, 3.
- [21] M. Asif, *Curr. Med. Chem.* **2012**, 19, 2984.
- [22] S. G. Turk, *J. Chem.* **2011**, 35, 803.
- [23] A. W. Potts, D. M. P. Holland, A. B. Trofimov, J. Schirmer, L. Karlsson, K. Siegbahn, *J. Phys. B: At. Mol. Opt. Phys.* **2003**, 36, 3129.
- [24] Y. Li, J. Wan, X. Xu, *J. Comput. Chem.* **2007**, 28, 1658.
- [25] R. So, S. Alavi, *J. Comput. Chem.* **2007**, 28, 1776.
- [26] M. Stener, P. Decleva, D. M. P. Holland, D. A. Shaw, *J. Phys. B: At. Mol. Opt. Phys.* **2011**, 44, 075203.
- [27] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. D. Simone, K. C. Prince, A. B. Trofimov, E. V. Gromov, I. L. Zaytseva, J. Schirmer, *Chem. Phys.* **2008**, 347, 360.
- [28] F. Wang, *J. Phys.: Conf. Ser.* **2008**, 141, 012019.
- [29] D. M. P. Holland, A. W. Potts, L. Karlsson, M. Stener, P. Decleva, *Chem. Phys.* **2011**, 390, 25.
- [30] Y. Takahata, A. Okamoto, D. P. Chong, *Int. J. Quantum Chem.* **2006**, 106, 2581.
- [31] X. Qian, P. Umari, N. Marzari, *Phys. Rev. B* **2011**, 84, 075103.
- [32] Z. Yang, P. Duffy, F. Wang, *Int. J. Quantum Chem.* **2013**, 113, 2312.
- [33] J. S. Tan, S. X. M. Boerrigter, R. P. Scaringe, K. R. Morris, *J. Comput. Chem.* **2012**, 33, 950.
- [34] E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, D. G. Truhlar, *J. Chem. Theory Comput.* **2011**, 7, 3027.
- [35] S. Chatterjee, M. Ahmed, F. Wang, *Radiat. Phys. Chem.* **2016**, 119, 1.
- [36] A. Ganesan, F. Wang, *J. Chem. Phys.* **2009**, 131, 044321.
- [37] C. T. Falzon, F. Wang, *J. Chem. Phys.* **2005**, 123, 214307.
- [38] C. T. Falzon, F. Wang, W. Pang, *J. Phys. Chem. B* **2006**, 110, 9713.
- [39] Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.* **2008**, 41, 157.
- [40] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* **2003**, 24, 669.
- [41] D. Danovich, *WIREs Comput. Mol. Sci.* **2011**, 1, 377.
- [42] A. P. W. Arachchilage, F. Wang, V. Feyer, O. Plekan, K. C. Prince, *J. Chem. Phys.* **2010**, 133, 174319.
- [43] A. Ganesan, F. Wang, M. Brunger, K. C. Prince, *J. Synchrotron. Radiat.* **2011**, 18, 733.
- [44] L. Selvam, V. Vasilyev, F. Wang, *J. Phys. Chem. B* **2009**, 113, 11496.
- [45] F. Chen, F. Wang, *Molecules* **2009**, 14, 2656.
- [46] P. R. T. Schipper, O. V. Gritsenko, S. J. A. V. Gisbergen, E. J. Baerends, *J. Chem. Phys.* **2000**, 112, 1344.
- [47] D. P. Chong, Y. Takahata, *Chem. Phys. Lett.* **2006**, 418, 286.
- [48] R. V. Leeuwen, E. J. Baerends, *Phys. Rev. A* **1994**, 49, 2421.

- [49] M. V. Roux, P. Jimenez, J. Z. Davalos, O. Castano, M. T. Molina, R. Notario, M. Herreros, J. L. M. Abboud, *J. Am. Chem. Soc.* **1996**, 118, 12735.
- [50] Z. J. Yang, F. Wang, *New. J. Chem.* **2014**, 38, 1031.
- [51] S. Chatterjee, F. Wang, *J. Theor. Comput. Chem.* **2015**, 14, 1550014.
- [52] F. Wang, *J. Phys. Chem. A* **2003**, 107, 10199.
- [53] C. G. Ning, Y. R. Huang, S. F. Zhang, J. K. Deng, K. Liu, Z. H. Luo, F. Wang, *J. Phys. Chem. A* **2008**, 112, 11078.
- [54] F. Wang, L. Selvan, G. F. Gribakin, C. M. Surko, *J. Phys. B* **2010**, 43, 165207.
- [55] K. Morokuma, *J. Chem. Phys.* **1971**, 55, 1236.
- [56] E. D. Glendening, A. Streitwieser, *J. Chem. Phys.* **1994**, 100, 2900.
- [57] M. V. Hopffgarten, G. Frenking, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, 2, 43.
- [58] H. Hirao, *Chem. Phys. Lett.* **2007**, 443, 141.
- [59] T. Ziegler, A. Rauk, *Inorg. Chem.* **1979**, 18, 1558.
- [60] T. Ziegler, A. Rauk, *Inorg. Chem.* **1979**, 18, 1755.
- [61] P. Su, H. Li, *J. Chem. Phys.* **2009**, 131, 014102.
- [62] W. J. Stevens, W. H. Fink, *Chem. Phys. Lett.* **1987**, 139, 15.
- [63] Y. Mo, J. Gao, S. D. Peyerimhoff, *J. Chem. Phys.* **2000**, 112, 5530.
- [64] Y. Mo, P. Bao, J. Gao, *Phys. Chem. Chem. Phys.* **2011**, 13, 6760.
- [65] A. J. Misquitta, R. Podeszwa, B. Jeziorski, K. Szalewicz, *J. Chem. Phys.* **2005**, 123, 214103.
- [66] K. Kitaura, K. Morokuma, *Int. J. Quantum Chem.* **1976**, 10, 325.
- [67] R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, M. Head-Gordon, *J. Phys. Chem. A* **2007**, 111, 8753.
- [68] T. Ziegler, A. Rauk, *Theor. Chim. Acta* **1977**, 46, 1.
- [69] A. E. Reed, L. A. Curtiss, F. Weinhold, *Chem. Rev.* **1988**, 88, 899.
- [70] K. Szalewicz, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, 2, 254.
- [71] F. Wang, S. Islam, V. Vasilyev, *Mater* **2015**, 8, 7723.
- [72] Z. Gomez-Sandoval, E. Pena, C. F. Guerra, F. M. Bickelhaupt, M. A. Mendez-Rojas, G. A. Merino, *Inorg. Chem.* **2009**, 48, 2714.
- [73] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, Jr., J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Revision a.02.*, Gaussian, Inc., Wallingford, CT **2009**.
- [74] G. T. Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. V. Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, 22, 931.
- [75] C. F. Guerra, J. G. Snijders, G. T. Velde, E. J. Baerends, *Theor. Chem. Acc.* **1998**, 99, 391.
- [76] E. J. Baerends, T. Ziegler, J. Autschbach, D. Bashford, A. Bérces, F. M. Bickelhaupt, C. Bo, P. M. Boerrigter, L. Cavallo, D. P. Chong, L. Deng, R. M. Dickson, D. E. Ellis, M. van Faassen, L. Fan, T. H. Fischer, C. F. Guerra, A. Ghysels, A. Giammona, S. J. A. V. Gisbergen, A. L. Yakovlev, *ADF2012, SCM, Theoretical Chemistry*, Vrije Universiteit, Amsterdam, The Netherlands. Available at: <http://www.scm.com>
- [77] A. M. Turing, *Q. J. Mech. Appl. Math.* **1948**, 1, 287.
- [78] H. A. Roger, J. R. Charles, Vol. 3, Cambridge University Press, **1985**, Ch. 9, 159.
- [79] J. Kyte, R. F. Doolittle, *J. Mol. Biol.* **1982**, 157, 105.
- [80] L. Fernholt, C. Romming, *Acta. Chem. Scand. A* **1978**, 32, 271.
- [81] B. J. M. Bormans, G. D. With, F. C. Mijlhoff, *J. Mol. Struct.* **1977**, 42, 121.
- [82] S. Craddock, C. Purves, D. W. H. Rankin, *J. Mol. Struct.* **1990**, 220, 193.
- [83] F. Wang, M. T. Downton, N. Kidwani, *J. Theor. Comput. Chem.* **2005**, 4, 247.
- [84] F. Weinhold, *J. Comput. Chem.* **2012**, 33, 2363.
- [85] M. Ahmed, A. Ganesan, F. Wang, V. Feyer, O. Plekan, K. C. Prince, *J. Phys. Chem. A* **2012**, 116, 8653.
- [86] F. Wang, *J. Phys. Chem. A* **2003**, 107, 4999.
- [87] L. Asbrink, C. Fridh, B. O. Jonsson, E. Lindholm, *Int. J. Mass. Spectrom. Ion. Phys.* **1972**, 8, 215.
- [88] L. Asbrink, C. Fridh, B. O. Jonsson, E. Lindholm, *Int. J. Mass. Spectrom. Ion. Phys.* **1972**, 8, 229.
- [89] F. M. Bickelhaupt, E. J. Baerends, *Angew. Chem. Int. Ed.* **2003**, 42, 4183.
- [90] I. Gutman, V. Gineityte, M. Lepovi, M. Petrovi, *J. Serb. Chem. Soc.* **1999**, 64, 673.
- [91] S. Vishveshwara, K. V. Brinda, N. Kannan, *J. Theor. Comput. Chem.* **2002**, 1, 1.
- [92] L. Hagen, A. B. Kahng, *IEEE Trans. Comput. Aided Des.* **1992**, 11, 1074.
- [93] E. R. Barnes, J. Siam, *Algebra. Discrete. Methods* **1982**, 7, 541.
- [94] W. E. Donath, A. J. Homan, *IBM J. Res. Dev.* **1973**, 17, 420.
- [95] P. Gould, *Inst. Br. Geog. Trans.* **1967**, 42, 53.
- [96] K. Tinkler, *Inst. Br. Geog. Trans.* **1972**, 55, 17.
- [97] S. M. Patra, S. Vishveshwara, *Biophys. Chem.* **2000**, 84, 13.
- [98] B. S. Sanjeev, S. M. Patra, S. Vishveshwara, *J. Chem. Phys.* **2001**, 114, 1906.

## SUPPORTING INFORMATION

Additional Supporting Information can be found in the online version of this article at the publisher's website.

**How to cite this article:** S., Chatterjee, F., Wang *Int. J. Quantum Chem.* 2016, DOI: 10.1002/qua.25229