

Diabolus in Chemistry?

Julien Pilmé* and Riccardo Spezia*

Sorbonne Université, CNRS, Laboratoire de Chimie Théorique CC 137 – 4, place Jussieu

F. 75252 PARIS CEDEX 05 – France

E-mail: julien.pilme@sorbonne-universite.fr; riccardo.spezia@sorbonne-universite.fr

Abstract

This article explores the possible presence of a pentacle valence bond structure in C_5 cyclic molecules. At this end, we have used quantum chemistry tools to elucidate the possible arrangement and the nature of chemical bonds within linear, cyclic, and three-dimensional structures only formed by five carbon atoms. While the linear structure is clearly the most stable one, local minima were obtained for both bi- and three-dimensional structures. Beyond the historical satanic symbol, the pentacle arrangement corresponds to an unusual formal structure with five crossing between C-C bonds. Our calculations show that this diabolic cyclic C_5 scheme is a relevant resonant structure and, furthermore, it is also present in the more known cyclo-pentadienyl molecule.

1 Introduction

The ideas of regular geometries and symmetries had historically a crucial role in imaging and determining chemical structures. From a long time, the concept of valency, the nature of the chemical bond, is often related to structural considerations. One well-known example is given by the Kekule's structure of benzene.¹ The concept of resonant structures also comes

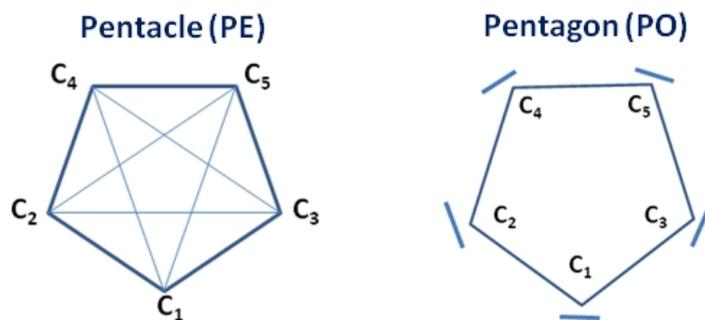
from possible geometrical patterns in which chemical bonds can be disposed fulfilling the valence of atoms constituting the molecules.² Another example of structure discovery motivated by geometrical considerations is the bicyclo[2.2.0]hexa-2,5-diene structure (an isomer of benzene) proposed theoretically by Dewar³ and synthesised by van Tamelen and Pappas almost one hundred year after.⁴ This alternative arrangement can be seen as coming from conceptual possible arrangement of bonds which are considered as "linkages between Daltonian atoms".²

We can call this a geometry-driven discovery which we can find in many other examples: from the simple water molecule,⁵ to transition metal complexes, in which group theory, leaded by symmetry, is the common way of describing them, to solid state chemistry (and physics) in which the symmetry is extended to periodicity and spatial groups are used.^{6,7}

In organic chemistry, triangles, squares, rectangles, pentagons, hexagons etc ... made by carbon atoms are ubiquitous and they are used to simply rationalize bond making and breaking, thanks to the ability of carbon atoms to make multiple bonds. However, in organic chemistry in solution, carbon based structures are composed also by other atoms (at least H atoms), with the relevant exception of fullerenes. Interestingly, their structures were first suggested and observed experimentally only many years after.⁸

Amongst the possible geometrical figures, one seems missing to us: the pentacle. This is an old geometrical figure composed by five vertices all linked together and, when the tip downwards, it constitutes the well-known satanic symbol. If, as usual, each segment is a chemical bond, this will generate a (diabolic) Lewis structure corresponding to a molecule composed by five carbon atoms (and no hydrogen or other atoms). This particular cyclic C₅ molecule can have, in principle, the bonds arranged in a way they form a pentacle motif (see Scheme 1).

In this contribution, we used theoretical chemistry to scrutiny the possibility that such particular diabolic resonant structure is present in C₅ molecule. At this end, we studied the different possibilities for C₅ to be arranged in space: as chain, as a ring and as a branched



Scheme 1: Two Lewis structures of cyclic C_5 .

structure. Furthermore, we have also quantified the contribution of pentacle resonant structure in the well-known cyclopentadienyl, $C_5H_5^-$, structure.

2 Theory

2.1 Quantum Chemical topology

In this paper, we assume that the reader is familiar with the quantum chemical topology (QCT) of scalar fields because numerous presentations of the methodology have already been published in the literature.^{9,10} Briefly, QCT is devoted to answer general questions about the chemical bonding in molecules and solids, and predict or explain chemical reactivity trends. This approach relies on the theory of gradient dynamical systems that enables a partitioning of the molecular space into basins. The most used one is the electron density, giving rise to the quantum theory of atoms-in-molecules (*QTAIM*).^{11,12} These basins being only atomic ones, a topological atom can be defined as the union of a nucleus with its associated electron density basin. Another widely used function is the Electron Localization Function (ELF) usually interpreted as a signature of the distribution of electron-pairs in the molecular space.¹³ The ELF topology depicts some non-atomic valence basins in addition to valence and core basins surrounding nuclei with atomic number $Z > 2$. In all the cases, the basins are delimited by zero-flux surfaces and the integration of the electron density over each basin directly provides its corresponding population. For example, the *QTAIM*

atomic charge, q , is calculated by subtracting the electron population of the topological atom.. Introduced by Bader and Stephens¹⁴ and later recovered by Fradera et al.¹⁵ the delocalization index (δ) is a measure of the electron-sharing between two atoms and can be compared to other bond order indices.

2.2 Molecular Similarity

Molecular similarity is a fundamental concept in chemistry. It is crucial to many aspects of chemical reasoning and for example, it stands as a cornerstone assumption in medicinal chemistry.¹⁶ Conversely, dissimilarity plays a significant role in an expanding array of applications particularly in combinatorial chemistry or for the the virtual screening.¹⁷ Several types of mathematical methods can be used to represent the molecular dissimilarity. Among these various approaches, the electron localization-delocalization real and symmetric matrix (LDM), is a representative graph of a molecule where all atoms (vertices) are interconnected by the QTAIM delocalization indices.¹⁸ The summation of off-diagonal elements of a LDM corresponds to the total delocalized population of the molecule. The trace of a LDM ($n \times n$) is the total localized electron population of the molecule:

$$LDM = \begin{pmatrix} \lambda_A & \delta_{AB}/2 & \dots & \delta_{AN}/2 \\ \delta_{BA}/2 & \lambda_B & \dots & \delta_{BN}/2 \\ \dots & \dots & \dots & \dots \\ \delta_{NA}/2 & \delta_{NB}/2 & \dots & \lambda_N \end{pmatrix}$$

where λ_i are the localization indices and δ_{ij} the delocalization indices (positive when atoms interact). Although several limitations have been identified, the LDM matrix is a powerful tool to measure the similarity/dissimilarity of different molecules themselves or as a predictor in QSAR methods.¹⁹ There does not exist a unique way to compare matrices. However, the Euclidian-type norm (termed as Frobenius distance) is commonly used to evaluate the dissimilarities between matrices. The dissimilarities between two molecules A

and B can be then evaluated using the Frobenius distance $d(A, B)$ applied to the LDM, :

$$d(A, B) = \sqrt{\sum_{i,j} |a_{ij} - b_{ij}|^2}$$

Where a_{ij} and b_{ij} are the corresponding matrix elements of LDMs of A and B molecules respectively. Smaller the value of the distance, more similar are the two molecules. With this point, the Frobenius distance can be used for measuring the similarity between structures.^{20,21}

3 Results and discussion

3.1 Structural parameters

We consider three arrangements of C_5 : linear 1D (or more in general a segment), 2D (a cyclic structure) and 3D (for which we considered different spatial arrangements). The most stable geometry corresponds to a linear, 1D, structure with $D_{\infty h}$ symmetry, which was previously detected in the circumstellar shell of IRC+10216.^{22,23} However, we found stable structures also for 2D and 3D arrangements: all the geometries are shown in Figure 1 while structural and energetic results are gathered in Table 1. One 2D structure is a minimum (termed as Diabolus-Irregular) and it corresponds to a distorted pentagon where the bonds are not all the same and with a C_{2v} symmetry. We also searched for a regular pentagonal structure (termed as Diabolus-Regular) with a D_{5h} symmetry where all bonds and angles are equal. However, we can find only one geometry and it turned out to be a saddle point. Moving to 3D arrangement, we obtained a stable structure which corresponds to a bipyramid trigonal one (D_{3h} symmetry) while the tetrahedral structure (T_d symmetry) is a saddle point. From a connection with simple concepts that are familiar to chemists, the valence-shell electron pair repulsion model (VSEPR) model²⁴ proven its indisputable utility for helping chemists to somehow with hands rationalize molecular architectures. Interestingly, the C_5 arrangement

deviates from the standard VSEPR prediction. Indeed, while the predicted VSEPR-type suggests a tetrahedral three-dimensional arrangement ($AX_4 \equiv CC_4$), it is worth noting that the tetrahedral structure is not a minimum; in fact, it is found to be more than 250 kcal/mol above the linear structure.

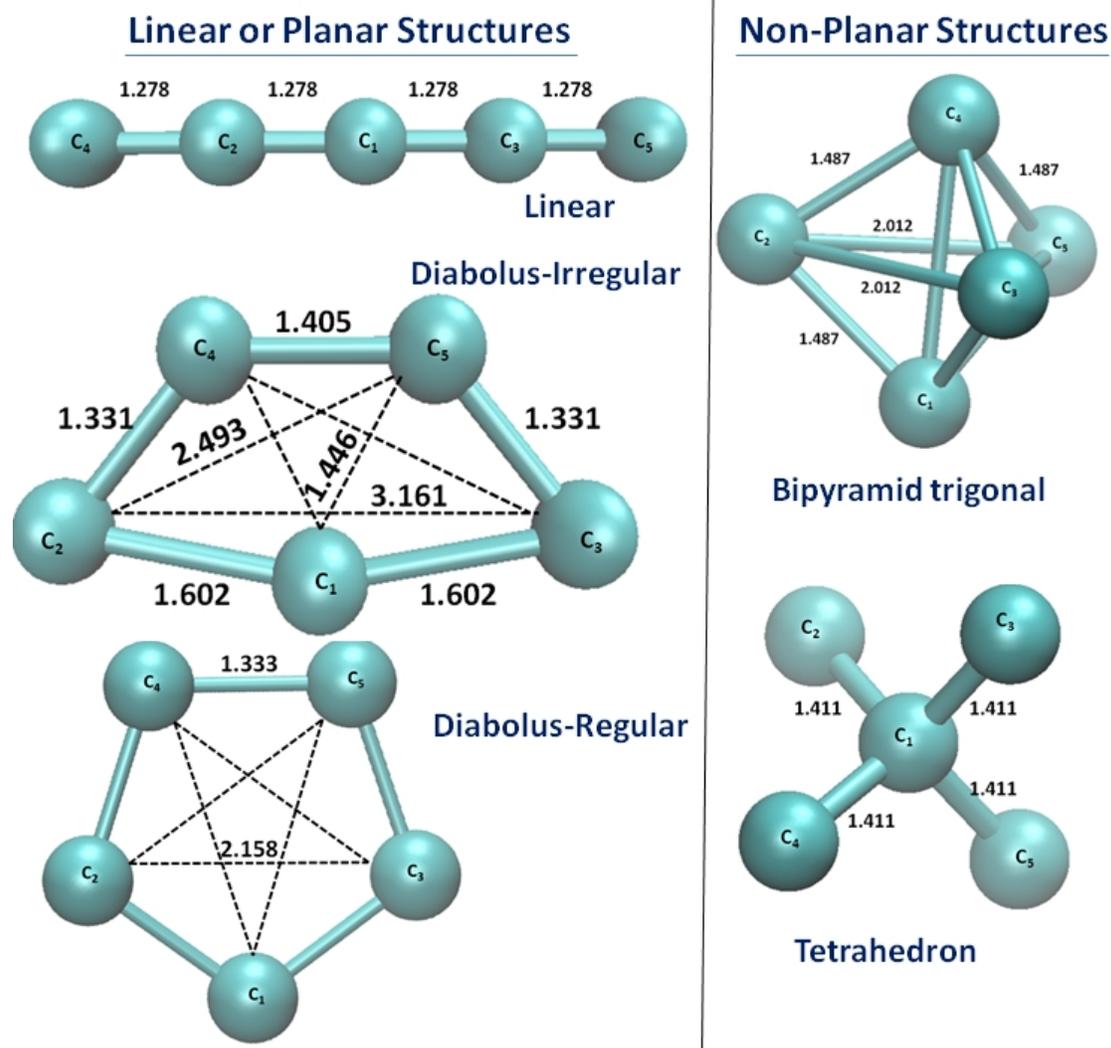


Figure 1: Optimized C_5 Structures. Geometries have been optimized at the M06-2X/aug-cc-pVTZ level of theory. Distances are given in Å.

All of these structures have significantly higher energies, ranging from 40 to 260 kcal/mol above the linear structure. The two planar structures closely look like the formal pentacle and pentagon structures (termed Diabolus-Irregular and -Regular). Note that Diabolus-Irregular (C_{2v}) being a minimum, it could potentially be observed in the interstellar medium. We can

Table 1: Structural and energy parameters of C_5 structures optimized at the M06-2X/aug-cc-pVTZ level of theory. The carbons numbering is given in Figure 1.

Molecule C_5	Symmetry	ΔE (kcal/mol)	Type	R_{12} (Å)	R_{23} (Å)	R_{14} (Å)
Linear	$D_{\infty h}$	0.0	minimum	1.278	2.556	3.834
Diabolus-Irregular	planar, C_{2v}	43.8	minimum	1.602	3.161	1.446
Bipyramid trigonal	D_{3h}	80.6	minimum	1.487	2.012	1.857
Diabolus-Regular	planar, D_{5h}	174.3	saddle point	1.333	2.158	2.158
Tetrahedron	T_d	259.5	saddle point	1.411	2.305	1.411
$C_5H_5^-$	planar, D_{5h}		minimum	1.408	2.279	2.279

also observe that the C_5 cycle in the well-known cyclopentadienyl $C_5H_5^-$ (D_{5h}) exhibits structural parameters similar enough to the single Diabolus-Regular system (D_{5h}). For example, all the inter-atomic distances in Diabolus-Regular have been found with a same value of 2.158 Å close to the corresponding distance of 2.279 Å calculated in $C_5H_5^-$ at the M06-2X/aug-cc-pVTZ level of theory (see Table 1).

3.2 The linear structure

Another important concern assessed in this work, is related to the validity of expected Lewis structure of the ground state of the C_5 molecule regarding the geometry and its electron pair domains. With ten electron pairs, the expected Lewis structure gathers eight bonding domains (four σ bonds and four π bonds) together with two lone pairs belonging to the σ system and located around the terminal carbon atoms (carbene like structure). As shown in Figure 2, the occurrence of ELF domains in the linear geometry perfectly matches with the Lewis prediction on the location of domains.

In addition to core basins, the ELF topology yields several valence basins. Among them, two non-bonding basins accounting for the carbon lone-pairs are located on the right and on the left of the linear structure. Three similar bonding basins, accounting for the covalent electron-pairing are identified in this structure. However, as shown in Table 2, the computed QTAIM charges and delocalization indices do not match with a full symmetric structure where all the C-C bonds are identical.

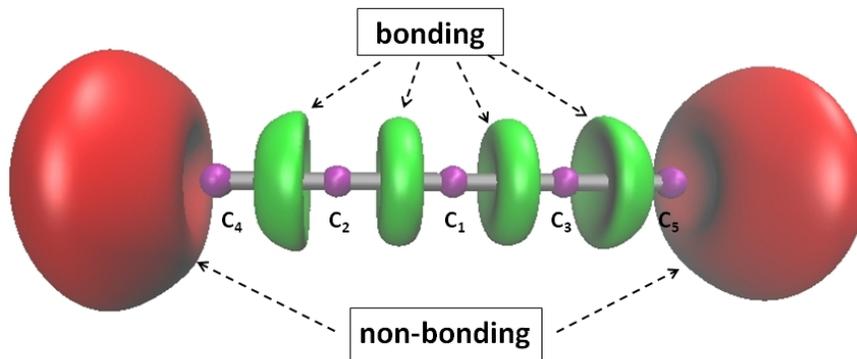


Figure 2: ELF localization domains (isosurface = 0.8) for the C_5 linear structure computed at the M06-2X/aug-cc-pVTZ level of theory. Color code: magenta for core basins, red for valence non-bonding basins and green for bonding basins.

Table 2: QTAIM analysis. Atomic populations \bar{N} and delocalization index (δ) in electrons. The carbons numbering is given in Figure 1.

Linear C_5	$\bar{N}(e)$	$\delta(e)$
C_1	6.53	$C_1 - C_2 : 1.20$
C_2	6.17	$C_2 - C_4 : 1.94$
C_3	6.17	$C_3 - C_1 : 1.20$
C_4	5.57	$C_4 - C_2 : 1.94$
C_5	5.57	$C_5 - C_3 : 1.94$

Indeed, the delocalization index is close to 2 for the C_2-C_4 and C_3-C_5 bonds while it is close to 1 for the C_1-C_2 and C_1-C_3 bonds. These results align with negative QTAIM charges for the central atoms (C_1 , C_2 , C_3) and positive for the terminal atoms (C_4 and C_5). Comparing the ELF and QTAIM data leads us to propose a resonant Lewis scheme swinging between covalent and ionic structures where localized lone pairs (4 σ electrons) are considered at the terminal carbons. In addition, all 8 π electrons can be delocalized in each structure. The resulting resonant scheme is given in Figure 3 and takes into account all of these results.

To give a flavor of the potential weights of these selected Lewis structures, we can use the QTAIM populations, as done by Silvi and others.²⁵ For symmetry reasons, we can only get three atomic charges and thus, we can access to three different weights as shown in

pentacle Lewis structure. As shown in Figure 4, the ELF topology of both Diabolus-Irregular (minima) and Diabolus-Regular (saddle point) yields several valence basins.

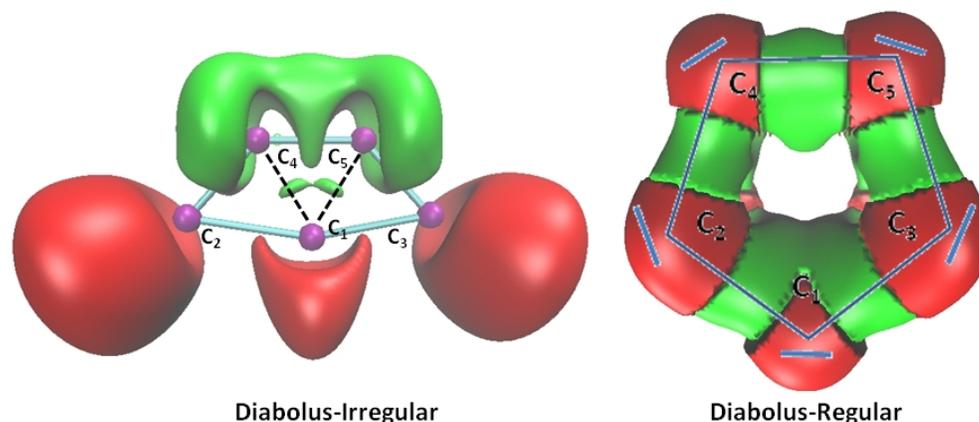


Figure 4: ELF localization domains (isosurface = 0.8) for the C_5 cyclic structures calculated at the M06-2X/aug-cc-pVTZ level of theory. Color code: magenta for core basins, red for valence non-bonding basins and green for bonding basins.

Among them, non-bonding basins accounting for the carbon lone-pairs are located on only three carbons for Diabolus-Irregular while all carbons are involved for Diabolus-Regular. For the latter, this aligns perfectly with the formal Lewis structure Pentagon (see figure 4) where no inner C-C bonds are found. Interestingly, two similar bonding basins accounting for the covalent electron-pairing are identified inside this Diabolus-Irregular which exhibits rather striking resemblances to the pentacle formal structure (see Figure 4). In order to quantitatively evaluate the similarity between the DFT-optimized structures and the formal Lewis structures (pentacle and pentagon), we propose to use the LDM formalism previously detailed in the theory section. Note that the way used in the previous section to build a resonant system of equations from QTAIM populations cannot be applied to the cyclic forms because all the formal atomic populations in both structures (pentacle and pentagon) are equal to 6 electrons (see Figure 1). However, we can use the LDM matrices to the formal C_5 Pentacle (PE) and Pentagon (PO) structures, which are defined as follows :

Given the two reference matrices (PE and PO), we can use the delocalization indices computed from the different DFT cyclic structures and measure the Frobenius distances as

$$\text{PE} = \begin{pmatrix} 4 & 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 4 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 4 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & 4 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 & 4 \end{pmatrix} \quad \text{PO} = \begin{pmatrix} 5 & 1/2 & 1/2 & 0 & 0 \\ 1/2 & 5 & 0 & 1/2 & 0 \\ 1/2 & 0 & 5 & 0 & 1/2 \\ 0 & 1/2 & 0 & 5 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 5 \end{pmatrix}$$

explained in the Theory section. For example, the following LDM matrix corresponds to the Diabolus-Regular molecule :

$$\begin{pmatrix} 4.008 & 0.853 & 0.853 & 0.142 & 0.142 \\ 0.853 & 4.008 & 0.142 & 0.853 & 0.142 \\ 0.853 & 0.142 & 4.008 & 0.142 & 0.853 \\ 0.142 & 0.853 & 0.142 & 4.008 & 0.853 \\ 0.142 & 0.142 & 0.853 & 0.853 & 4.008 \end{pmatrix}$$

All the results are reported in Table 3 where we also report the distance for PE and PO as reference as well as what obtained from C_5H_5^- .

Table 3: Weights of diabolus structures **(a)** Frobenius distance calculated with respect to the Pentacle formal structure **(b)** Frobenius distance calculated with respect to the Pentagon formal structure **(c)** Calculated from Eq. 1 **(d)** Calculated from Eq. 2.

Molecule C_5	$d_{PE}^{(a)}$	$d_{PO}^{(b)}$	$\omega_{PE}^{(c)}$	$\omega_{PO}^{(d)}$
Pentacle (PE)	0.0	2.738	1.0	0.0
Diabolus-Irregular	1.441	2.017	0.58	0.42
Diabolus-Regular	1.588	2.524	0.61	0.39
C_5H_5^-	1.677	2.018	0.54	0.46
Pentagon (PO)	2.738	0.0	0.0	1.0

From the two Frobenius distances, it is also possible to define the weight of Pentacle (ω_{PE}) and Pentagon (ω_{PO}) as follows:

$$\omega_{PE} = 1 - \frac{d_{PE}}{d_{PE} + d_{PO}} \quad (1)$$

$$\omega_{PO} = 1 - \frac{d_{PO}}{d_{PE} + d_{PO}} \quad (2)$$

Results are reported in Table 3. Surprisingly, the pentacle Lewis structure makes a dominant contribution in both cyclic structures Diabolus-Regular and Diabolus-Irregular, reaching a weight exceeding fifty percent, even for $C_5H_5^-$. Note that for Diabolus-Regular, it was rather difficult to characterize the contribution of the pentacle Lewis structure from only ELF localization domains (see Figure 4) since none inner C-C bonding basin has been found inside the pentagon. Recalling the hidden diabolic aspect lurking in the shadow, these results illustrate that the rationalization of the electronic structure of cyclic structures needed a thorough analysis of a set of delocalization indices revealing the role of inner delocalized density in the pentacle cycle.

4 Conclusions and Outlooks

It is remarkable how small and simple systems, such as the C_5 molecule, still offer a rich playground for a better understanding of usual concepts, such as the resonance. In particular, when the five carbon atoms are arranged in a ring, we obtained that a formal structure with bonds passing inside the ring is an important one: this intriguingly corresponds to a pentacle, which was never considered before. This "diabolic" bonding pattern, is particularly intriguing because the bonds formally "cross" each other. Notably, this pattern is not negligible also in the hydrogenated form, the less exotic $C_5H_5^-$ molecule. While a pentacle is possible only in C_5 systems, it will be interesting to know if other "crossing bonds" structures are possible resonant structure and how this is related (or not) to their reactivity.

Furthermore, we have found that not only the linear 1D C_5 structure is a minimum (at it is the most stable structure), but also 2D and 3D structures exist. Thus, It would be interesting to see how increasing the carbon length, 2D and 3D geometries become more and more stable and if such diabolic crossing bonds are responsible to their stabilization of other structures with growing number of carbon atoms. Is the devil hidden in other structures? This question will surely deserve more investigations.

Computational details

The M06-2X hybrid functional²⁷ level with the Gaussian 16 software was used for all calculations.²⁸ The standard all-electron aug-cc-pVTZ basis set was used for all atoms.²⁹ All the geometries were fully optimized without symmetry constraint. Each minima displays only positive eigenvalues. The quantum chemical topology analyses have been performed using the TopChem2 program package.³⁰

Acknowledgement

We thank an anonymous twitter user who brought us to the interest of a possible pentacle bonding structure.

References

- (1) Kekulé, F. A. Untersuchungen über aromatische Verbindungen. *Liebigs Ann. Chem.* **1866**, *137*, 129–136.
- (2) Lewis, G. N. The Atom and the Molecule. *J. Am. Chem. Soc.* **1916**, *38*, 762–785.
- (3) Dewar, J. On the Oxidation of Phenyl Alcohol, and a Mechanical Arrangement adapted to illustrate Structure in the Non-saturated Hydrocarbons. *Proc. R. Soc. Edinb.* **1869**, *6*, 82–86.
- (4) van Tamelen, E. E.; Pappas, S. P. Bicyclo [2.2.0]hexa-2,5-diene. *J. Am. Chem. Soc.* **1963**, *85*, 3297–3298.
- (5) Armstrong, H. E. On the Constitution and Properties of Water and Hydrogen Chloride. *Philosophical Transactions of the Royal Society of London* **1873**, *163*, 1–56.

- (6) Hofmann, A. On the Combining Power of Atoms. *Philosophical Transactions of the Royal Society of London* **1862**, *152*, 83–224.
- (7) Singh, V. Applications of Group Theory in Chemistry. *J. Chem. Ed.* **2007**, *84*, 1919–1923.
- (8) Kroto, H.; Heath, J.; O'Brien, S.; Curl, R.; Smalley, R. C60: Buckminsterfullerene. *Nature* **1985**, *318*, 162–163.
- (9) Popelier, P. L. A. In *The Chemical Bond II: 100 Years Old and Getting Stronger*; Mingos, D. M. P., Ed.; Springer International Publishing: Cham, 2016; pp 71–117.
- (10) Silvi, B.; Gillespie, R. J.; Gatti, C. In *Comprehensive Inorganic Chemistry II (Second Edition)*, second edition ed.; Reedijk, J., Poeppelemeier, K., Eds.; Elsevier: Amsterdam, 2013; pp 187–226.
- (11) Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893–928.
- (12) Popelier, P. L. A. *The Chemical Bond*; John Wiley Sons, Ltd, 2014; Chapter 8, pp 271–308.
- (13) Silvi, B.; Savin, A. Classification of chemical bonds based on topological analysis of electron localization functions. *Nature* **1994**, *371*, 683–686.
- (14) Bader, R. F. W.; Stephens, M. E. Spatial localization of the electronic pair and number distributions in molecules. *J. Am. Chem. Soc.* **1975**, *97*, 7391–7399.
- (15) Fradera, X.; Austen, M. A.; Bader, R. F. W. The Lewis Model and Beyond. *J. Phys. Chem. A* **1999**, *103*, 304–314.
- (16) Maggiora, G.; Vogt, M.; Stumpfe, D.; Bajorath, J. Molecular Similarity in Medicinal Chemistry. *J. Med. Chem.* **2014**, *57*, 3186–3204.

- (17) Bender, A.; Glen, R. C. Molecular similarity: a key technique in molecular informatics. *Org. Biomol. Chem.* **2004**, *2*, 3204–3218.
- (18) Matta, C. F.; Sumar, I.; Cook, R.; Ayers, P. W. In *Applications of Topological Methods in Molecular Chemistry*; Chauvin, R., Lepetit, C., Silvi, B., Alikhani, E., Eds.; Springer International Publishing: Cham, 2016; pp 53–88.
- (19) Cook, R. Principal components of localization-delocalization matrices: new descriptors for modeling biological activities of organic compounds. Part I: mosquito insecticides and repellents. *Struct Chem* **2017**, *28*, 1525–1535.
- (20) Matta, C. F. Molecules as networks: A localization-delocalization matrices approach. *Comput. Theor. Chem.* **2018**, *1124*, 1–14.
- (21) Timm, M. J.; Matta, C. F.; Massa, L.; Huang, L. The Localization-Delocalization Matrix and the Electron-Density-Weighted Connectivity Matrix of a Finite Graphene Nanoribbon Reconstructed from Kernel Fragments. *J. Phys. Chem. A* **2014**, *118*, 11304–11316.
- (22) Bernath, P. F.; Hinkle, K. H.; Keady, J. J. Detection of C₅ in the Circumstellar Shell of IRC+10216. *Science* **1989**, *244*, 562–564.
- (23) Galazutdinov, G.; Musaev, F.; Krelowski, J. On the detection of the linear C₅ molecule in the interstellar medium. *MNRAS* **2001**, *325*, 1332–1334.
- (24) Gillespie, R. Fifty years of the VSEPR model. *Coord. Chem. Rev.* **2008**, *252*, 1315–1327.
- (25) Silvi, B. How topological partitions of the electron distributions reveal delocalization. *Phys. Chem. Chem. Phys.* **2004**, *6*, 256–260.
- (26) Pichierri, F. Molecular Structure of the Octatetranlyl Anion, C₈H⁻: A Computational Study. *J. Phys. Chem. A* **2008**, *112*, 7717–7722, PMID: 18666765.

- (27) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (28) Frisch, M. J. et al. Gaussian~16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.
- (29) Dunning, J., Thom H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (30) Kozłowski, D.; Pilmé, J. New insights in quantum chemical topology studies using numerical grid-based analyses. *J. Comput. Chem.* **2011**, *32*, 3207–3217.