



HAL
open science

BenzAI: A Program to Design Benzenoids With Defined Properties Using Constraint Programming

Adrien Varet, Nicolas Prcovic, Cyril Terrioux, Denis Hagebaum-Reignier,
Yannick Carissan

► **To cite this version:**

Adrien Varet, Nicolas Prcovic, Cyril Terrioux, Denis Hagebaum-Reignier, Yannick Carissan. BenzAI: A Program to Design Benzenoids With Defined Properties Using Constraint Programming. *Journal of Chemical Information and Modeling*, 2022, 62 (11), pp.2811–2820. 10.1021/acs.jcim.2c00353 . hal-03691364

HAL Id: hal-03691364

<https://hal.science/hal-03691364v1>

Submitted on 9 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BenzAI: A Program to Design Benzenoids With Defined Properties Using Constraint Programming

Adrien Varet,[†] Nicolas Prcovic,[†] Cyril Terrioux,^{*,†} Denis Hagebaum-Reignier,[‡]
and Yannick Carissan^{*,‡}

[†]*Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, 13013 France*

[‡]*Aix Marseille Univ, CNRS, Centrale Marseille, ISM2, Marseille, 13013 France*

E-mail: cyril.terrioux@univ-amu.fr; yannick.carissan@univ-amu.fr

Abstract

The BenzAI program can automatically generate benzenoids with defined structural constraints, like the number of hexagons, the number of carbon and/or hydrogen atoms, the existence of symmetries, the number of Kekulé structures, their diameter, or more elaborate criteria like their irregularity. BenzAI allows both automatic generation and manual building of benzenoids including or excluding defined patterns. For all benzenoids with less than 10 rings, it can extract IR spectra from a database that we generated. It computes local aromaticity for closed-shell and mono radical species using circuit-based algorithms. We present practical examples of what can be done with BenzAI for a diversity of cases.

Introduction

This article is devoted to the presentation of the BenzAI software that can generate any benzenoid, perform an analysis of its electronic properties in terms of Kekulé structures, and provide infrared spectra stored in an external database. A great feature of BenzAI is the fact that the generation of the structure can be done with various constraints especially with the constraint to contain or not one or multiple patterns. We believe BenzAI provides data of in-

terest for a broad spectrum of fields in physics and chemistry, e.g., astrochemistry, material design, synthetic chemistry, among others.

Benzenoids are polycyclic aromatic hydrocarbons (PAHs) made of fused six-member rings only. They are well-studied in various domains because of their electronic and optical properties.¹ They are also studied in detail in interstellar chemistry because they are suspected to be present in interstellar clouds and are believed to act as catalysts for chemical reactions taking place in space.² The aromatic character of benzenoids can be studied using the Clar sextet rule for ground states.³⁻⁵ For large benzenoids, the shape of edges has a large influence on their properties.^{6,7} They are also a challenge for synthetic chemists, which lead to the synthesis of remarkable molecules on surfaces.⁸ All these points can be addressed with BenzAI. Aromaticity can be easily evaluated even for radical species, for which BenzAI provides Clar covering and ring bond order analysis. Infrared spectra of multiple species can be determined through the automated consultation of our database, and building of benzenoids with a defined irregularity parameter. As for properties defined by the shape of the edges or the presence of a defined pattern, we have integrated such features, for which practical examples are given in the manuscript.

In the literature, there some software and

databases available to generate or study PAH structures. Among them, the CaGe software is based on graph theory and allows the construction of diverse 2D or 3D PAHs.⁹ It allows to generate 3D structures with fused 5-member rings and a variety of specific PAHs. Tománek and Frederick have setup a database, from which one can recover the geometries of all isomers of fullerenes with less than 720 carbon atoms.¹⁰ This work intends to study the curvature of these objects.¹¹ Another software to analyse geometry deformation in PAHs is available as a website and computes POAV (π Orbital Axis Vector) values and further developments to study geometrical deformation in PAHs.^{12–14} The NASA database is also available online and contains a large number of PAHs.¹⁵ This work is intended for infrared spectra studies. Recently, a large set of PAH and their local aromaticity was computed.¹⁶ Some software are also available to study aromaticity in general such as the Aroma package, which scans NICS values on a grid above a molecular plane.^{17–19} The newly developed `ims3d.py` suite does the same kind of work in a general fashion, which allows for the study of aromaticity in distorted molecules, for which the determination of a plane is ambiguous.²⁰ In this context, BenzAI aims at being an easy to use software flexible and expandable. It handles fragment combination for the generation of benzenoids, which is not covered by the software previously cited and deals with the electronic structure analysis based on graph theory. Furthermore, it is connected to a rich database dedicated to aromaticity. The full feature of this database will be published soon.

IUPAC defines aromaticity as a gain in energy of systems that provide for their enhanced thermodynamic stability (relative to acyclic structural analog).²¹ Thus, strictly speaking, it is a global property. However, as chemists, we know that rings in molecules are not equivalent and some are more likely to react (less stabilized by aromaticity) than others. Thus, BenzAI can compute local indices of aromaticity. This is not new and the literature gives a plethora of such indices. The interested reader can look at the article by Solà, who deciphers the reasons

for such a large number of criteria.²²

In the first part, we present the methods used to solve the benzenoid generation problem. The second part deals with the computation of local aromaticity and the third shows practical examples of what can be achieved with BenzAI. BenzAI is also connected to a database of benzenoids with less than ten cycles. For these molecules, the infrared spectra were calculated at the 6-31G/B3LYP level of theory. Whenever a molecule in this database is generated and if the database is reachable, the results are automatically retrieved and can be explored within BenzAI.

Underlying Concepts And Methods

Generation With Constraints

We briefly recall here the strategy used to generate all the benzenoids having a given number n of hexagons. For further details, please refer to our previously published work on the generation of benzenoids.^{23,24} Constraint programming is a branch of artificial intelligence, which focuses on automated reasoning. A generic solver is used whatever the nature of the problem. The problem is expressed in terms of constraints among variables in a given formalism sent to the solver, which outputs the solutions. We use constraint programming, as it easily allows for the extension of the generation of such molecules under combined specific constraints as opposed to bespoke approaches, which could be faster but do not guarantee that additional constraints could be used without a complete rewriting of the code.^{25,26}

We define the graph of a benzenoid B as the undirected graph B_h in which each hexagon is a vertex and an edge links two hexagons that are adjacent in the molecule. This is often referred to as the inner dual graph. Any benzenoid of n hexagons can be embedded in a coronenoid of size at most $k(n) = \lfloor \frac{n}{2} + 1 \rfloor$. The model relies on the fact that the generation of the n hexagon benzenoids can be translated in graph theory as finding all the connected sub-graphs of the

hexagon graph of a coronenoid of size $k(n)$. Let $B_h^{c(k(n))}$ be the graph of the coronenoid of size $k(n)$ and respectively n_c and m_c its number of vertices and edges. We use a Boolean variable x_i per vertex and a Boolean variable $y_{i,j}$ per edge of $B_h^{c(k(n))}$. x_i is **true** when the i^{th} hexagon of $B_h^{c(k(n))}$ is part of B_h , while $y_{i,j}$ is **true** when the i^{th} and j^{th} hexagons of $B_h^{c(k(n))}$ are fused in B_h . The search of all valid B_h contained inside $B_h^{c(k(n))}$ is done under the following constraints:

- If two hexagons are fused in the benzenoid graph B_h , they must be part of B_h and conversely: $y_{i,j} = \text{true} \Leftrightarrow x_i = x_j = \text{true}$;
- B_h must have n hexagons: $\sum x_i = n$;
- B_h must be connected: any pair of hexagons can be linked by a path.
- If six hexagons form a ring, the ring in the middle must be part of the graph (coronene case).

Further constraints were added to avoid redundant solutions, or can be added to constrain the solutions to be of a certain symmetry. Many libraries are available to represent and solve efficiently such problems. In this work, we exploit the open-source Java library *Choco Solver*²⁷ that natively proposes graph variables and the more usual graph-related constraints (notably **connected** constraint). An example of such a generation is given as an example in Figure 1.

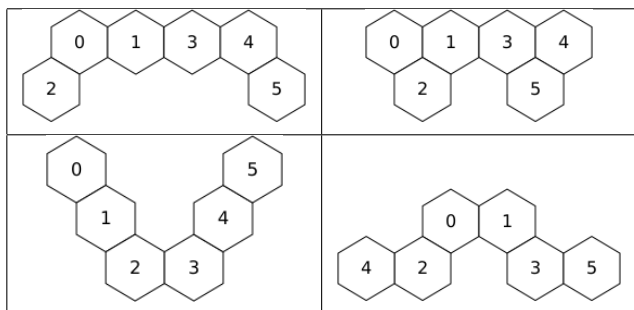


Figure 1: Generation of the only four benzenoids with six cycles and a mirror symmetry.

As for the generation of benzenoids with a constraint on fragments, it is based on the concept of extended benzenoid graph. The extension consists in assigning the hexagons of the

graph a value: plus (+ colored in green), minus (− colored in red), or neutral (0 colored in orange). When a hexagon is assigned a plus value, it must be in the graph of the solution. On the contrary, a minus value imposes the absence of the hexagon in the solution. Neutral hexagons are neither required nor excluded. This allows for the generation of benzenoid graphs, which represent fragments. The generation of the benzenoid is thus done with the constraint that the extended graph of one or more fragments must be in the solution or excluded from the solution. All details and models used for this are published elsewhere.²⁴ As an example, we show in Figure 2 the generation of all benzenoids with seven hexagons made of the combination of two fragments.

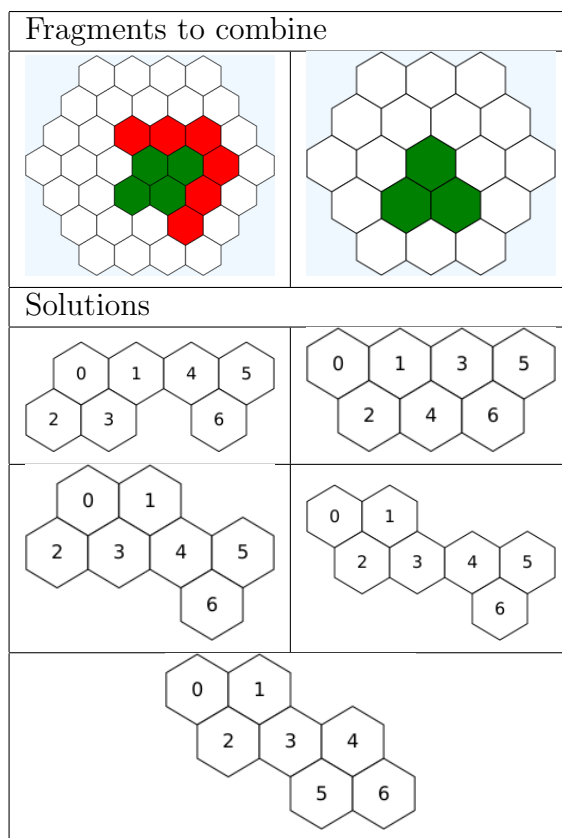


Figure 2: Example of the generation of benzenoids with seven hexagons, made of all possible combinations of two fragments. For fragments, green hexagons indicate that they must be present in the solution, whereas red hexagons must not be in the solution. More detailed information is available in the supplementary information (Figure SI 6).

Analysis of the Electronic Structure of a Benzenoid

The values obtained in the framework of this work are subject to limitations intrinsic to the method. The geometries used for the topological aromaticity calculations are not optimized (all hexagons are considered regular) and the aromaticity indices as well as the bond orders are quantities which differ from the ones extracted from quantum chemical calculations. However, the trends are usually good as shown in our recent work.²⁸

Local Aromaticity

In our program, local aromaticity is computed following the work of Randić²⁹ based on minimal conjugated circuits.

Definition 1 *Let B be a benzenoid and K one of its Kekulé structure. C is a minimal conjugated circuit of the hexagon h of B (also called h -MCC) with respect to K if C is one of the covering circuits of h having the smallest size.*

Such a circuit is a loop made of alternating single and double bonds in a Kekulé structure. In benzenoids, these circuits can only be of size $4n + 2$ (with n a positive integer). The algorithm suggested by Randić is described in Algorithm 1. The difficulty lies in obtaining all possible Kekulé structures and their set of conjugated circuits. The detailed implementation of this model using the approaches by the Lin and Fan method³⁰ or the improved method by Lin³¹ with constraint programming is thoroughly described in a previous work.³² In short, this approach consists in enumerating all possible h-MCCs in all Kekulé structures of a benzenoid and attributing to each of them a "resonance energy", R_n , which value is tabulated according to their size ($4n + 2$). The sum of all these energy contributions normalized to the number of Kekulé structures gives the total resonance energy - or global aromaticity - of the benzenoid. The contribution of each ring of the benzenoid, called ring resonance energy (RRE), then corresponds to local aromaticity.

Algorithm 1: Algorithm used to determine local aromaticity of all benzenic cycles of a benzenoid

Data: \mathcal{K} the complete set Kekulé structures the benzenoid under study
 $\{h_i\}_{i=1}^{n_h}$ the benzenic cycles (h stands for hexagon) of the benzenoid under study
 $\{R_i\}$ the resonance energies associated to $6, 10, \dots, 4i + 2$ conjugated electrons
Result: Local aromaticity of each cycle

```
1 foreach  $K \in \mathcal{K}$  do
2   determine  $\mathcal{C}$ , the complete set of
   conjugated circuit of  $K_i$ 
3   foreach  $C \in \mathcal{C}$  do
4     for  $k \leftarrow 1$  to  $n_h$  do
5       if  $h_k$  is involved in  $C$ 
6         then
7            $e \leftarrow$  number of electrons
           conjugated in  $C$ 
8            $E[k] \leftarrow E[k] + R_e$ 
9         end
10    end
11  end
12  for  $k \leftarrow 1$  to  $n_h$  do
13     $E[k] \leftarrow \frac{E[k]}{|\mathcal{K}|}$ 
14  end
15 end
```

Clar cover

From Solà 2013:³³ *Clar's rule states that the Kekulé resonance structure with the largest number of disjoint aromatic π -sextets, i.e., benzene-like moieties, is the most important for the characterization of properties of polycyclic aromatic hydrocarbons (PAHs). Aromatic π -sextets are defined as six π -electrons localized in a single benzene-like ring separated from adjacent rings by formal CC single bonds.*³⁴

For a given benzenoid, a Clar cover is obtained by maximizing the number of Clar sextets of its electronic structure. In our approach, we can generate radical Clar covers by imposing the constraint that the molecule must carry one, or more radicals.

Ring bond order

Clar formulas can be numerically represented by **ring bond orders** (RBO).²⁹ The bond order is a well-known concept among the chemist community and was first introduced by L. Pauling.³⁵ In the framework of π conjugated systems, a "pure" π bond as in the ethylene molecule, has a bond order of 1. When one sums up all bond orders of the six C–C bonds forming each ring, one obtains an RBO. The archetypal RBO of a π system is the RBO of the benzene molecule and is equal to 3. Thus, the closest to 3 an RBO is, the more aromatic the ring is. In BenzAI, the bond order b of a benzenoid B is calculated as the ratio, in percent, of the number of Kekulé structures in which b is a double bond and the total number of Kekulé structures of B .

Interface

Molecules are grouped into collections. These collections can be filled by the user via drawing or importing a molecule or automatic generation using constraints. Within a collection, one can select a set of molecules to create a new collection or for particular calculations, for instance, calculation of local aromaticity. Molecules can be appended to an existing collection by unit or by batches of selected molecules. A filter can be applied to a collection to generate a new collection. Figure 3 shows screenshots of BenzAI for a typical input and the corresponding output.

Drawing and Import/Export

The drawing module allows for the easy drawing of a benzenoid on a hexagonal grid. The default grid is made of three concentric crowns of hexagons: the smallest crown is just a single ring, i.e. the benzene molecule, two crowns correspond to the coronene molecule, and so on. Such concentric crowns of hexagons are thus called coronenoids, as they contain a coronene substructure. A simple click on a hexagon of the grid adds it to the desired molecule. There can be holes in the benzenoid but all cycles

must be fused. Once drawn, the molecule is added to the desired collection of molecules. Hence many molecules can be drawn by hand and analyzed.

Benzenoids can also be exported from the application using the export menu, either individually or by collection. Benzenoid structures can be exported from the application using the export menu, either individually or by collection, to the following formats: graph format,³⁶ CML (Chemical Markup Language), XYZ and PNG format (Portable Network Graphics). They can be imported later on in BenzAI only using the graph format. Importing a collection consists simply in selecting several files (these files must be located in the same directory).

Generation of Benzenoid Structures at Will

The BenzAI program can generate benzenoid structures with the following constraints, which can be combined:²³

Based on the chemical composition:

- Total number of hexagons
- Total number of carbon atoms
- Total number of hydrogen atoms

Based on the structure:

- Irregularity parameter ξ and/or number of solo, duo, trio, and quarto hydrogen atoms (see definition below)
- Diameter: it is defined as the maximal length of the shortest path connecting any hexagon to another. This length is equal to the number of edges that are crossed along the path. For instance, the molecule shown in Figure 4 has a diameter of 4, the maximal length of the shortest path is either 0-1-2-5-7 or 3-1-2-5-7.
- Coronenoid, rectangle, rhombus, catacondensed
- With/without holes (Coronoids)

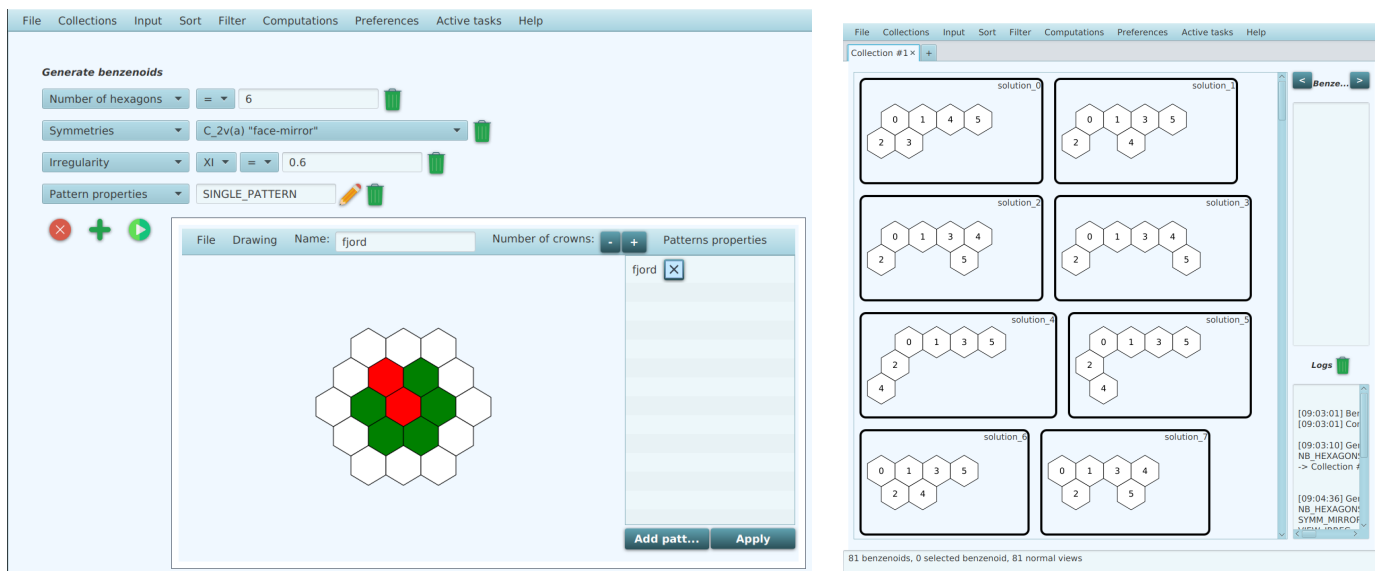


Figure 3: Screenshots of the BenzAI graphical interface: generator interface (left) result of a generation (right).

- With/without a combination of patterns. Patterns can be joint, disjoint, or share hexagons, they can be saved and loaded in a user-friendly manner.
- Using symmetries: 13 types of symmetry are possible for benzenoids, using either reflection or rotation symmetry (or any combination of both of types) referred to with their Schoenflies symmetry point group symbol.³⁷

Based on generation itself:

- Number of solutions desired
- Total time for the generation

Please note that the number of solutions can be extremely high. So we advise the user to make extensive use of the total time generation criterion as unwise requests can exhaust his patience or RAM, whichever comes first.

Filtering

Any collection can be filtered with the same criteria used for the generation, except for the number of solutions or total time. In addition to the generation criteria, the number of Kekulé

structures can be used as a new filtering criterion. When set to zero, this allows a simple way to filter non-kekulean benzenoids from the solutions of a collection. The filtering option is thus a convenient way to generate numerous sub-collections according to some criteria from a common parent collection having given composition or structural characteristics. Hence, one can explore a collection through many aspects.

Sorting

Any generated collection can be sorted out in increasing or decreasing order with respect to structural parameters (number of carbon atoms, hydrogen atoms, rings, or irregularity) or with respect to the number of Kekulé structures. As an example, let us generate the set of all benzenoids with 32 carbon atoms. This leads to a collection of 330 molecules. Sorting by decreasing order of ξ shows that only on such molecule has a ξ value of 1. This molecule is shown in Figure 4.

Computations on a collection

Using the "Computations" tab of BenzAI, one can perform two types of analysis on either an

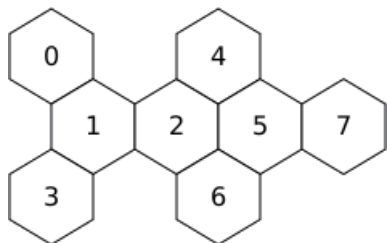


Figure 4: Only benzenoid with 32 carbon atoms with a ξ value of 1 obtained by filtering the collection of 330 benzenoids with 32 carbon atoms by decreasing ξ value.

entire collection or on manually selected solutions of a collection:

- an electronic structure analysis according to the aromaticity criteria previously mentioned (RRE, Clar cover, RBO).
- a histogram analysis with respect to the irregularity parameter

These analyses will be illustrated below in the "Use Cases" section.

Database

If possible, BenzAI can retrieve data from the database we computed, which contains all possible benzenoids of less than 10 cycles and their vibrational spectra computed at the 6-31g/B3LYP level of theory. All geometries were optimized and do not present any negative frequency. These spectra can be seen individually or they are merged using Boltzmann statistics based on the DFT energies of each molecule. For example, let us generate the set of benzenoids with four hexagons and retrieve their vibrational spectra from the database, Figure 5. BenzAI allows for the visualization of 3 spectra, each of them being a Boltzmann's weighted spectrum of all molecules with the same number of carbon and hydrogen atoms.

Use Cases

Aromaticity of a Set of Molecules

In order to show the capabilities of BenzAI for the computation of local aromaticity and re-

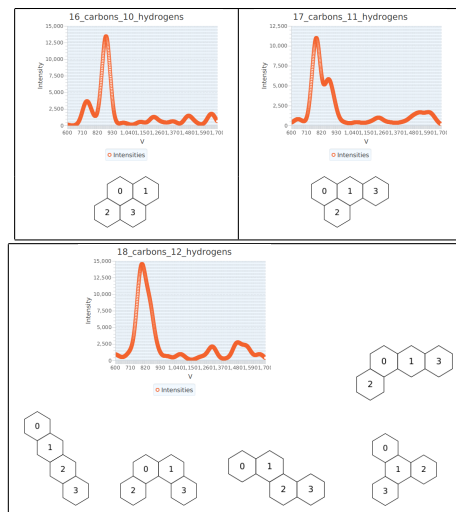


Figure 5: Vibrational spectra extracted from the database for the benzenoids with four hexagons. Three spectra are obtained as these benzenoids can have either 16, 17 or 18 carbon atoms. The spectrum obtained for 18 carbon atoms is a Boltzmann weighted spectrum of the spectra of the five benzenoids.

lated properties, we show the results obtained from the set of all the benzenoids with five rings. All results are shown in the supplementary information (Figures SI 1-5). In this section, we shall consider a subset of these molecules (Figure 6). Benzo[m]tetraphene and Benzo[e]pyrene have only one Clar structure. The local aromaticity calculation and the ring bond order representation fully agree. Pentacene has five Clar structures, which differ by the cycle carrying the Clar sextet. Thus, the ring bond order representation gives an equal value for each ring. However, in acenes, the local aromaticity is larger for the central cycles. This is well captured by the local aromaticity calculation based on electronic circuits. The Benzo[a]tetraphene is similar to pentacene with two sliding sextets (Figure 7). The local aromaticity of the molecule is well described and fully in line with the Clar cover. The ring bond order description lacks accuracy for the right-most ring (by construction) but the description of aromaticity with electronic circuits captures the fact that this ring is less aromatic than its neighbour. The perylene case is of high interest as it shows a central ring with no local aro-

Name	Local aromaticity	Clar cover	Ring bond order
Benzo[m]tetraphene			
Benzo[e]pyrene			
Pentacene			
Benzo[a]tetraphene			
Perylene			
C ₁₇ H ₁₃			

Figure 6: From left to right: local aromaticity using Lin’s method, Clar cover and ring bond order for a subset of 5-member ring benzenoids. Darker blue color indicates higher local aromaticity or higher ring bond order. For the Clar cover representation, only one among all possible Clar structures is given. Note that the circle around one atom indicates a radical.

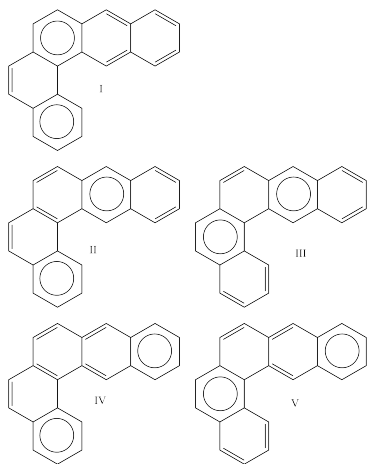


Figure 7: Clar structures with two sextets of benzo[a]tetraphene.

maticity, in agreement with the Clar cover, in which the central ring cannot carry a sextet. It is well-known that perylene can be seen as two naphthalene moieties linked by two bonds of essentially single character. However, the ring

bond order of the central ring is not null, because it takes into account the Kekulé structures for which its bonds can have a double character. This is never the case for the two central bonds but the other ones are double in the Kekulé structures of naphthalene. Finally, the C₁₇H₁₃ molecule is not kekuleable, because it has an odd number of π electrons. The local aromaticity cannot be evaluated by methods based on electronic circuits made of alternating single and double bonds because of the radical delocalization. The Clar cover picture shows a carbon, which carries the radical and all its second neighbors can have a radical character as well, Figure 8. Then the ring bond order picture allows for an evaluation of local aromaticity as well as the localization of the radical. In the case of a radical species, the ring bond order is obtained by summing on all possible electronic structures, which have one radical and maximize the number of Clar sextets.

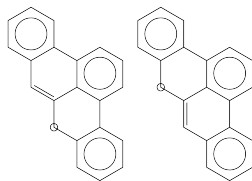


Figure 8: Clar structures with three sextets of $C_{17}H_{13}$. The atoms surrounded by a circle carry the radical.

Irregularity

The irregularity parameter ξ , is defined as follows:²

$$\xi = \frac{N_{trio} + N_{quarto}}{N_{solo} + N_{duo} + N_{trio} + N_{quarto}} \quad (1)$$

It is used in astrochemistry to characterize the regularity of polycyclic aromatic hydrocarbons. BenzAI allows for the calculation of the distribution of ξ in a collection of molecule and the generation of molecules with a constraint ξ value or constraint values of N_{solo} , N_{duo} , N_{trio} or N_{quarto} . Recently, Bouwman suggested that the NASA PAH database lacks large molecules (with 34 carbon atoms or more) with high irregularity values.² In Figure 9, after generating the full set of benzenoids with nine hexagons or less and filtering them by the number of carbon atoms (less than or equal to 34 and more than 34 carbon atoms), we compare the distribution of ξ in both sets. As suggested, the NASA database over-represents molecules with a low value of ξ . This is of the highest importance because the kind of hydrogen of a PAH (solo, duo, etc.) was shown to have an influence on its IR absorption band.^{15,38,39}

Molecular Design

Generation of Large Symmetric Molecules

BenzAI allows for the generation of a determined number of benzenoids with a high number of benzene rings and specific properties. In the following, we shall generate five benzenoids with 24 cycles, a mirror symmetry and constrained ξ values ($\xi < 0.2$, $\xi = 0.2, 0.4, 0.6, 0.8$ and $\xi > 0.8$). Results of this generation are

shown in Figure 10. It is interesting to see that higher ξ values lead to more branched structures (as opposed to more compact for low ξ values), which have non-aromatic cycles. Of course, such the generation of only five benzenoids per ξ value cannot be generalized, but if one combines these results with the distribution of ξ shown in Figure 9, it is of interest to extend the available databases by a tool such as BenzAI, to better represent the variety of benzenoids.

Prediction of Magnetic Properties

Some of us have recently studied chiral non-planar PAHs based on triphenylene. These molecules were obtained by a trimerization of helicenes (formally a [2+2+2] reaction, Figure 11).⁴⁰ The trimerization leads to the formation of a new C_6 cycle, which is referred to as the central cycle in the following discussion. Such molecules could exhibit magnetic properties if their π extensions would stabilize unpaired electrons. Thus, we search monomeric units with stable unpaired electrons and an even number of carbons, *i.e.* concealed non-kekulean benzenoids.⁴¹ We know that the $C_{13}H_9$ triangulene carries a radical localized on the second neighbor of the central carbon atom, Figure 12.^{42,43} There must be two such units by monomer for the total molecule to have an even number of electrons and these two units must not share a ring. The combination of these constraints can be handled by BenzAI. In the generator, we define the $C_{13}H_9$ as a pattern as shown in Figure 13 and ask for two such disjoint units. In the framework of this article, we decided to use only the two solutions shown in Figure 14: one solution **A** which should exhibit a diradical character (hence a trimer of septet ground state) and **B**, which should lead to a closed-shell wave-function. By construction, 3-**A** (Figure 15) should carry six unpaired electrons localized on the triangulene moieties whereas 3-**B** should be a closed-shell molecule. In order to verify these hypotheses, we computed the energy of the spin states of both molecules, Table 1.

Hence, the use of BenzAI allowed us to find a

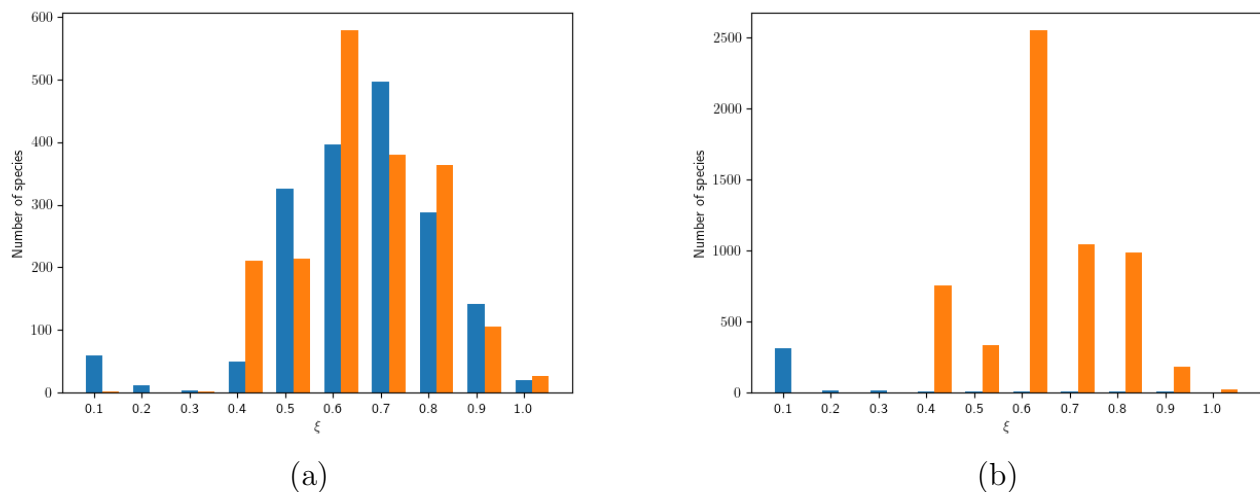


Figure 9: Comparison of the distributions of the irregularity parameter (less or equal than 34 carbon atoms (a) and more than 34 carbon atoms (b)) as shown by Bouwman *et al.*² (blue) and this work (orange) for the full set of benzenoids with nine or less hexagons.

Table 1: Energy of the spin states of 3-**A** and 3-**B** in kJ/mol obtained at the PBE0/def2-SV(P) level of theory.

	3-A	3-B
singlet	241	0
triplet	83	13
quintet	80	34
septet	0	53

concealed non-kekulean PAH with a high spin ground state.

Conclusion

This article presents the BenzAI program, an open-source software available at the following URL: <https://benzai-team.github.io/BenzAI/>. This program allows for the generation of benzenoids under structural constraints. Then, it allows for the determination of the local aromaticity of the hexagons and the calculation of IR spectra provided that the generated benzenoids have less than 10 hexagons. In this article, we show concrete applications of this software in diverse aspects of chemistry. The beauty of constraint programming lies in the ease to add new constraints in the future without having to modify a bespoke program. Thus,

we are already working on new features like allowing for cycles with more or less than six carbon atoms, which could extend the possibilities of the program.

Supporting Information Available

Complementary to Figure 2, the generation of benzenoids with seven hexagons, made of all possible combinations of two fragments with the fragments colored for easier recognition. Complementary to Figure 6, The full set of all benzenoids with five rings is available together with their local aromaticity computed with the Lin, Lin and Fan and ring bond order methods as well as their Clar cover.

Data and Software Availability

The BenzAI is available free of charge at <https://benzai-team.github.io/BenzAI/>.

Acknowledgement The authors thank the ANR project DEMOGRAPH (ANR-16-CE40-0028) and the ministère de l'éducation nationale for funding the PhD of Adrien Varet.

References

- (1) Rakhi, R.; Suresh, C. H. Optoelectronic Properties of Polycyclic Benzenoid Hydrocarbons of Various Sizes and Shapes for Donor- π -Acceptor Systems: A DFT Study. *ChemistrySelect* **2021**, *6*, 2760–2769.
- (2) Bouwman, J.; Castellanos, P.; Bulak, M.; Terwisscha van Scheltinga, J.; Cami, J.; Linnartz, H.; Tielens, A. G. G. M. Effect of Molecular Structure on the Infrared Signatures of Astronomically Relevant PAHs. *A&A* **2019**, *621*, A80.
- (3) Wang, Y. Quantitative Resonance Theory Based on the Clar Sextet Model. *JOURNAL OF PHYSICAL CHEMISTRY A* **2022**, *126*, 164–176.
- (4) Anjalikrishna, P. K.; Gadre, S. R.; Suresh, C. H. Antiaromaticity-Aromaticity Interplay in Fused Benzenoid Systems Using Molecular Electrostatic Potential Topology. *JPCA* **2021**, *125*, 5999–6012.
- (5) Jousselein-Oba, T.; Mamada, M.; Wright, K.; Marrot, J.; Adachi, C.; Yassar, A.; Frigoli, M. Synthesis, Aromaticity, and Application of peri-Pentacenopentacene: Localized Representation of Benzenoid Aromatic Compounds. *ACIE* **2022**, *61*, e202112794.
- (6) Lin, C.-K. Theoretical study of nitrogen-doped graphene nanoflakes: Stability and spectroscopy depending on dopant types and flake sizes. *JCC* **2018**, *39*, 1387–1397.
- (7) Pelloni, S.; Lazzeretti, P. Polygonal current models for polycyclic aromatic hydrocarbons and graphene sheets of various shapes. *JCC* **2018**, *39*, 21–34.
- (8) Arikawa, S.; Shimizu, A.; Shiomi, D.; Sato, K.; Shintani, R. Synthesis and Isolation of a Kinetically Stabilized Crystalline Triangulene. *JACS* **2021**, *143*, 19599–19605.
- (9) Brinkmann, G.; Friedrichs, O. D.; Liskens, S.; Peeters, A.; Van Cleemput, N. CaGe - a Virtual Environment for Studying Some Special Classes of Plane Graphs - an Update. *MATCH-COMMUNICATIONS IN MATHEMATICAL AND IN COMPUTER CHEMISTRY* **2010**, *63*, 533–552.
- (10) Tomanek, D.; Frederick, N. Fullerene Database <http://nanoten.com/science/fullerene/>.
- (11) Guan, J.; Jin, Z.; Zhu, Z.; Chuang, C.; Jin, B.-Y.; Tománek, D. Local Curvature and Stability of Two-Dimensional Systems. *Phys. Rev. B* **2014**, *90*, 245403.
- (12) Salvato-Vallverdu, G.; Sabalot-Cuzzubbo, J.; Bégué, D.; Cresson, J. Pychemcurv <https://pychemcurv.readthedocs.io/en/latest/>.
- (13) Haddon, R. C. Hybridization and the Orientation and Alignment of π -Orbitals in Nonplanar Conjugated Organic Molecules: π -Orbital Axis Vector Analysis (POAV2). *J. Am. Chem. Soc.* **1986**, *108*, 2837–2842.
- (14) Sabalot-Cuzzubbo, J.; Salvato-Vallverdu, G.; Begue, D.; Cresson, J. Relating the Molecular Topology and Local Geometry: Haddon’s Pyramidalization Angle and the Gaussian Curvature. *J. Chem. Phys.* **2020**, *152*.
- (15) Bauschlicher, J.; Peeters, E.; Allamandola, L. J. The Infrared Spectra of Very Large Irregular Polycyclic Aromatic Hydrocarbons (PAHs): Observational Probes of Astronomical PAH Geometry, Size and Charge. *ApJ* **2009**, *697*, 311–327.
- (16) Alvarez-Ramirez, F.; Ruiz-Morales, Y. Database of Nuclear Independent Chemical Shifts (NICS) versus NICSZZ of Polycyclic Aromatic Hydrocarbons (PAHs). *J. Chem. Inf. Model.* **2020**, *60*, 611–620.

- (17) Stanger, A. Nucleus-Independent Chemical Shifts (NICS): Distance Dependence and Revised Criteria for Aromaticity and Antiaromaticity. *J. Org. Chem.* **2006**, *71*, 883–893.
- (18) Stanger, A. Obtaining Relative Induced Ring Currents Quantitatively from NICS. *J. Org. Chem.* **2010**, *75*, 2281–2288.
- (19) Gershoni-Poranne, R.; Stanger, A. The NICS-XY-Scan: Identification of Local and Global Ring Currents in Multi-Ring Systems. *Chem.-Eur. J.* **2014**, *20*, 5673–5688.
- (20) Artigas, A.; Hagebaum-Reignier, D.; Carissan, Y.; Coquerel, Y. Visualizing Electron Delocalization in Contorted Polycyclic Aromatic Hydrocarbons. *Chem. Sci.* **2021**, *12*, 13092–13100.
- (21) McNaught, A. D.; Wilkinson, A. *IUPAC Compendium of Chemical Terminology, 2nd Ed. (the "Gold Book")*. Compiled by A. D. McNaught and A. Wilkinson.; Blackwell Scientific Publications, Oxford (1997) Online version (2019-) created by S. J. Chalk.
- (22) Solà, M. Why Aromaticity Is a Suspicious Concept? Why? *Frontiers in Chemistry* **2017**, *5*.
- (23) Carissan, Y.; Hagebaum-Reignier, D.; Prcovic, N.; Terrioux, C.; Varet, A. Using Constraint Programming to Generate Benzenoid Structures in Theoretical Chemistry. Principles and Practice of Constraint Programming. Cham, 2020; pp 690–706.
- (24) Carissan, Y.; Hagebaum-Reignier, D.; Prcovic, N.; Terrioux, C.; Varet, A. Exhaustive Generation of Benzenoid Structures Sharing Common Patterns. 27th International Conference on Principles and Practice of Constraint Programming (CP 2021). Dagstuhl, Germany, 2021; pp 19:1–19:18.
- (25) Dias, J. R. Benzenoid Series Having a Constant Number of Isomers. *Journal of Chemical Information and Computer Sciences* **1990**, *30*, 61–64.
- (26) Dias, J. R. The Most Stable Class of Benzenoid Hydrocarbons and Their Topological Characteristics - Total Resonant Sextet Benzenoids Revisited. *Journal of Chemical Information and Computer Sciences* **1999**, *39*, 144–150.
- (27) Fages, J.-G.; Lorca, X.; Prud'homme, C. Choco Solver User Guide Documentation. <https://choco-solver.readthedocs.io/en/latest/>.
- (28) Carissan, Y.; Dim, C.-A.; Hagebaum-Reignier, D.; Prcovic, N.; Terrioux, C.; Varet, A. Computing the Local Aromaticity of Benzenoids Thanks to Constraint Programming. Principles and Practice of Constraint Programming. Cham, 2020; pp 673–689.
- (29) Randić, M.; Balaban, A. T. Local Aromaticity and Aromatic Sextet Theory beyond Clar. *Int. J. Quantum Chem.* **2018**, *118*, e25657.
- (30) Lin, C.-d.; Fan, G.-q. Algorithms for the Count of Linearly Independent and Minimal Conjugated Circuits in Benzenoid Hydrocarbons. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 782–787.
- (31) Lin, C.-d. Efficient Method for Calculating the Resonance Energy Expression of Benzenoid Hydrocarbons Based on the Enumeration of Conjugated Circuits. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 778–783.
- (32) Carissan, Y.; Dim, C.-A.; Hagebaum-Reignier, D.; Prcovic, N.; Terrioux, C.; Varet, A. Computing the Local Aromaticity of Benzenoids Thanks to Constraint Programming. Principles and Practice of Constraint Programming. Cham, 2020; pp 673–689.
- (33) Solà, M. Forty Years of Clar's Aromatic π -Sextet Rule. *Front. Chem.* **2013**, *1*.

- (34) Hauptmann, S. *The Aromatic Sextet: Von E. Clar*; John Wiley & Sons LTD, London, New York, Sydney, Toronto 1972; 128 Seiten mit zahlreichen Formelbildern; Format 13 × 20 cm; broschiert £ 1, 50. *Zeitschrift für Chemie* **1973**, *13*, 200–200.
- (35) Pauling, L.; Brockway, L. O.; Beach, J. Y. The Dependence of Interatomic Distance on Single Bond-Double Bond Resonance 1. *Journal of the American Chemical Society* **1935**, *57*, 2705–2709.
- (36) https://benzai-team.github.io/BenzAI/graph_format.
- (37) Kovic, J.; Pisanski, T.; Balaban, A. T.; Fowler, P. W. On Symmetries of Benzenoid Systems. *Match-Communications in Mathematical and in Computer Chemistry* **2014**, *72*, 3–26.
- (38) Honny, S.; Van Kerckhoven, C.; Peeters, E.; Tielens, A. G. G. M.; Hudgins, D. M.; Allamandola, L. J. The CH Out-of-Plane Bending Modes of PAH Molecules in Astrophysical Environments. *A&A* **2001**, *370*, 1030–1043.
- (39) Ricca, A.; Roser, J. E.; Peeters, E.; Boersma, C. Polycyclic Aromatic Hydrocarbons with Armchair Edges: Potential Emitters in Class B Sources. *ApJ* **2019**, *882*, 56.
- (40) Artigas, A.; Hagebaum-Reignier, D.; Carissan, Y.; Coquerel, Y. Visualizing Electron Delocalization in Contorted Polycyclic Aromatic Hydrocarbons. *Chem. Sci.* **2021**, 10.1039.D1SC03368A.
- (41) Cyvin, S. J.; Brunvoll, J.; Cyvin, B. N. Search for Concealed Non-Kekulean Benzenoids and Coronoids. *J. Chem. Inf. Model.* **1989**, *29*, 236–244.
- (42) Rico-Sotomayor, E. M.; Barquera-Lozada, J. E. Triangulenes and their ions: reaching the limits of Clar’s rule. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24704–24711.
- (43) Misra, A.; Schmalz, T. G.; Klein, D. J. Clar Theory for Radical Benzenoids. *Journal of Chemical Information and Modeling* **2009**, *49*, 2670–2676, PMID: 19916509.

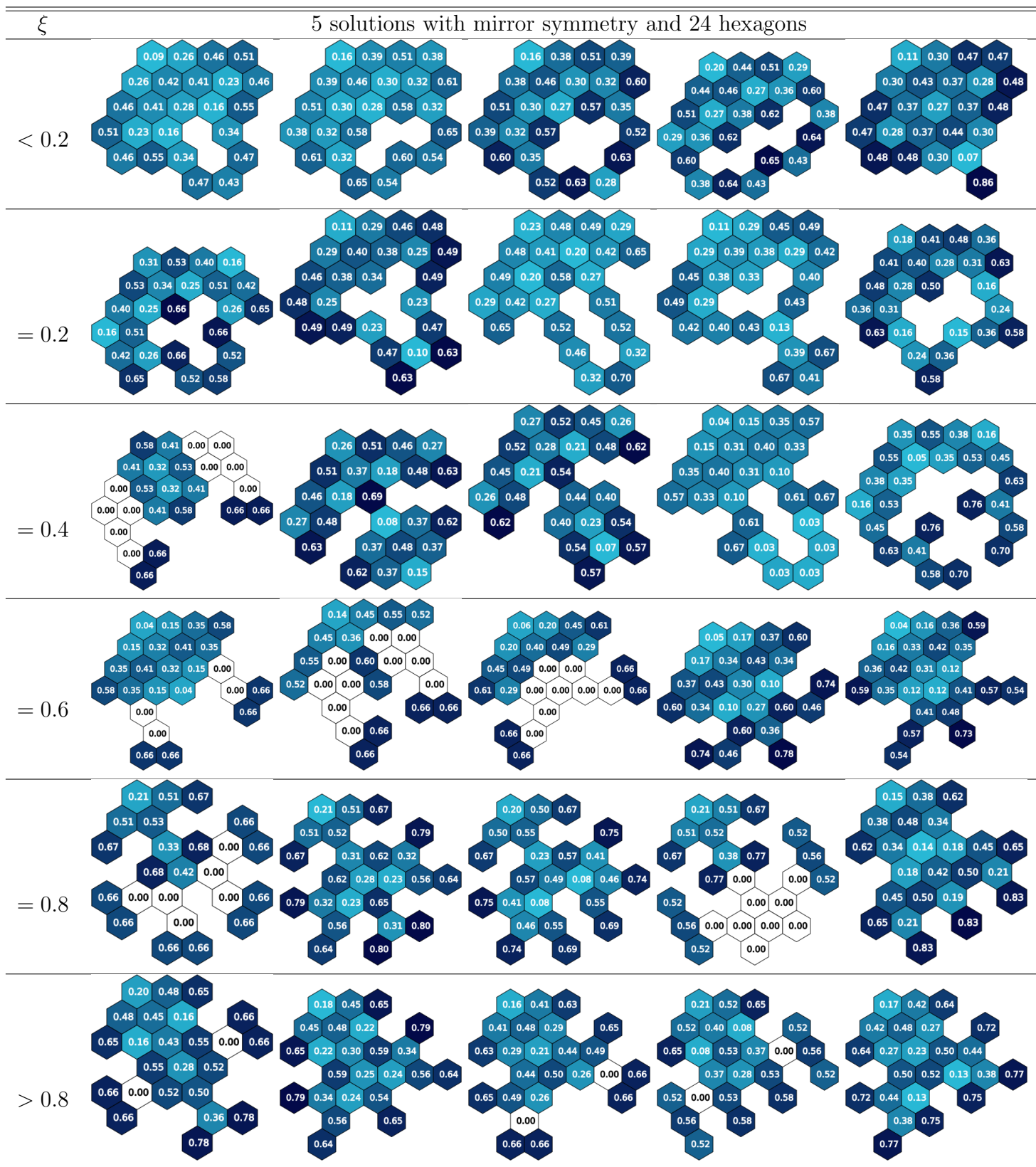


Figure 10: Five first solutions obtained by six successive benzenoid generation done with BenzAI with the following constraints: 24 benzene rings, mirror symmetry, and an irregularity parameter $\xi < 0.2$, $\xi = \{0.2, 0.4, 0.6, 0.8\}$ and $\xi > 0.8$. Blue color scale indicates Lin aromaticity. White color indicates no local aromaticity. All these solutions were obtained within a few seconds on an Intel core i7 laptop.

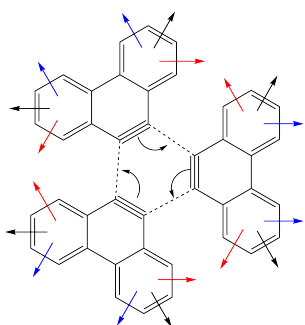


Figure 11: Formal [2+2+2] trimerization of three monomeric units. The arrows indicate the possible π extensions of these monomers.

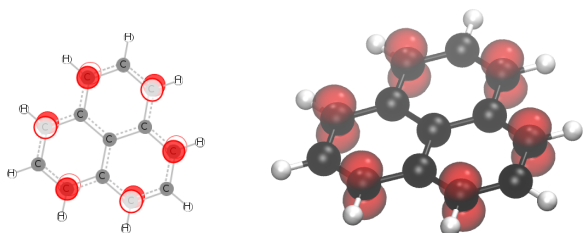


Figure 12: Singly occupied molecular orbital at the Hückel level of theory (left) and spin density of the $C_{13}H_9$ radical computer at the PBE0/def2-SV(P) level of theory (right).

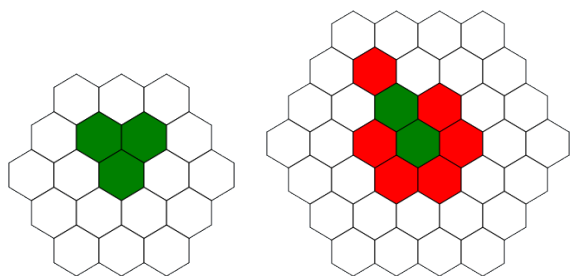


Figure 13: Pattern definition for the $C_{13}H_9$ radical carrier fragment (left) and of the central cycle fragment (right). Green cycles must be in the solution whereas red cycles are excluded from the solution.

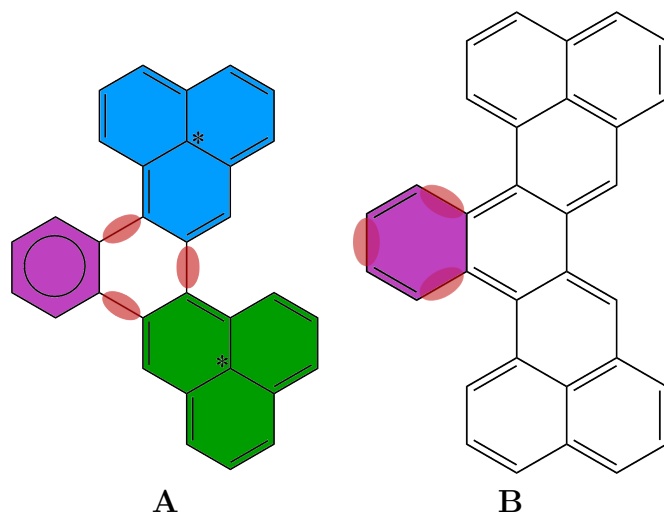


Figure 14: Compound **A** (left) and **B** (right). The stard positions and equivalent carry a radical. The purple cycle will be the central cycle of the trimer, the blue and green triangulene units each carry a radical. These two electrons cannot be paired: it is a concealed non-kekulean benzenoid. Bonds surrounded in red are single in all triplet Kekulé structures.

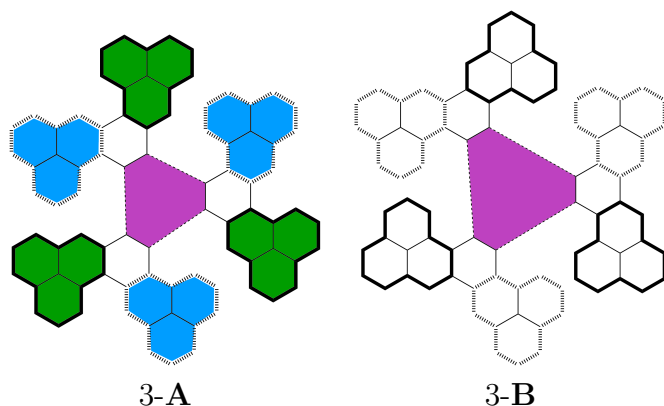


Figure 15: Trimerization of compound **A** (left) and of compound **B** (right). The dashed lines indicate the bonds which will lead to the formation of the central cycle. The bold lines indicate that the part of the molecule comes towards the reader, whereas the hollow lines go away from the reader. Blue and green triangulene moieties carry one radical.

Graphical TOC Entry

