

# Conformational Analysis of Enantiomerization Coupled to Internal Rotation in Triptycyl-n-helicenes

A. Carreras,<sup>a</sup> L. Fuligni,<sup>b</sup> P. Alemany,<sup>b</sup> M. Llunell,<sup>b</sup> J. M. Bofill,<sup>c</sup> and W. Quapp<sup>d</sup>

<sup>a</sup> Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal, 4. 20018 Donostia, Euskadi, Spain.

<sup>b</sup> Departament de Ciència de Materials i Química Física & Institut de Química Teòrica i Computacional (IQTCUB). Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain.

<sup>c</sup> Departament de Química Inorgànica i Orgànica & Institut de Química Teòrica i Computacional (IQTCUB). Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain.

<sup>d</sup> Mathematisches Institut, Universität Leipzig, PF 100920, D-04009 Leipzig, Germany.

## Index

### A. Reduced PES computation

#### A.1. Local search

#### A.2. Full reduced PES

### B. Characteristic points

#### Notes and References

Minimum energy geometries (Three TnH.xyz files containing the geometry in cartesian coordinates are provided as supplementary material)

### A. Reduced PES computation

For the computation of the reduced PES for each TnH molecule one must obtain at least a regular grid of points on this surface, although this is not a task that can be simply performed using any of the algorithms related to the computation of molecular optimizations available in standard quantum chemistry packages such as Gaussian<sup>1</sup> or Gamess<sup>2</sup>. Since these programs just allow to freeze specific internal coordinates during energy optimizations, it is not possible to apply those procedures to the coordinates chosen in this work to build the PES since, as described in the main text, they are defined as non-fixed combinations of individual internal coordinates that may be different for each point in the grid.

Since there are many different geometries compatible for each pair of  $\alpha_R$  and  $\alpha_H$  values, to calculate a single point on the reduced PES of a TnH molecule we must obtain the optimized geometry with the minimum energy for this pair of  $\alpha_R$  and  $\alpha_H$  values. As explained in the main text, these two angles are, however, not actual dihedral angles involving four atoms in the molecule, but rather a lineal combination of four dihedral angles  $\beta$ . So, in the end, it is necessary to determine the set of values for these four dihedral angles that minimize the energy. For this reason, for each  $(\alpha_R, \alpha_H)$  pair, an infinite set of  $\beta$  values must be considered, although, by definition, two of these four angles (from now on  $\beta''$ ) depend also from the value of the other two (from now on  $\beta'$ ) and from the values of  $n_L$  and  $n_S$  that are used to indicate the position of the helicene ends with respect to the different triptycyl blades.

Taking these considerations into account it is easy to see that to find the geometry for a single point on the reduced PES it will be necessary to perform many restricted geometry optimizations, changing each time the values of the two  $\beta'$  angles, recalculate the corresponding values of the two  $\beta''$  angles, and perform a geometry optimization freezing the four considered dihedral angles  $\beta$  (see Fig. S1). Since the software that is employed for optimizations with frozen coordinates is Gaussian v.09, the dihedral angles in the input geometry must be close to the calculated values for these angles, otherwise the program is not able to properly fix the chosen internal coordinates. Because of this restriction, each point of the PES must be computed starting from an already calculated point in the neighbourhood, which also implies that not all the points of the PES can be computed simultaneously. Since this procedure must be repeated until the minimum value for the potential energy is reached, each single PES point is obtained from many different constrained optimizations and it has been necessary to develop an algorithm to efficiently perform all these optimizations to find minimum energy geometries to construct the reduced PES.

Given the symmetry of the PES as a consequence of the existence of a reflection plane as explained in the main text, it is, however, not necessary to compute points for all the variable ranges. Considering the rotation of the molecule, the range of  $\alpha_R$  can be limited between  $0^\circ$  and  $120^\circ$ . The range for the other variable,  $\alpha_H$ , can be reduced thanks to the inversion symmetry of the PES around the achiral points at  $\alpha_H = 0^\circ$ , allowing us to obtain points with negative  $\alpha_H$  values from those with positive ones.

#### A.1. Local search

For any pair of  $\alpha_R$  and  $\alpha_H$  values to be computed, a minimum energy geometry search respect to the  $\beta'$  angles must be conducted (see Fig. S2). The space for this search is, however, too large to be fully explored, but considering that each search starts from a close minimum geometry, the set of dihedral angles corresponding to the minimum geometry for the new values of  $\alpha_R$  and  $\alpha_H$  should be close to those corresponding to the starting geometry since the PES is a continuous surface. The global search of the minimum energy geometry can be thus converted into a local search starting from the values of the  $\beta'$  angles obtained from a nearby already computed geometry.

The algorithm that we have developed explores a grid of  $\beta'$  values close to the starting values, recalculating for each point in the scanning area the corresponding values of  $\beta''$ . Once the neighbour angles have been scanned, the search restarts taking the minimum energy geometry as the center of a new neighbourhood as shown in Fig. S3. The local search ends when the central point is also the minimum energy geometry.

Although the characteristics of the neighbourhood of a point such as its size and the distance between points are parameterized in the final algorithm, these values have been fixed after some previous tests to optimize convergence and computational cost. The displacement used to scan the neighbours of the  $\beta'$  angles is always  $1^\circ$ . The neighbourhood's size is defined by a parameter  $n$  indicating the number of

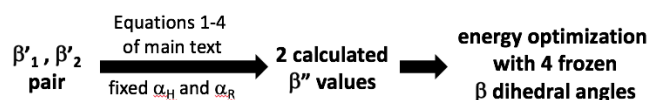


Fig. S1. Schematic procedure corresponding to the single energy optimization for a given pair of  $\beta'$  angles.

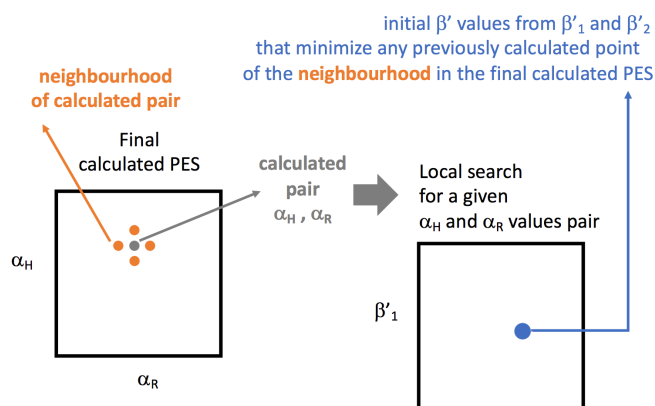


Fig. S2. Schematic representation of the initial procedure of the local search of  $\beta$  dihedral angles that minimize the energy for a given  $\alpha_H$  and  $\alpha_R$  values.

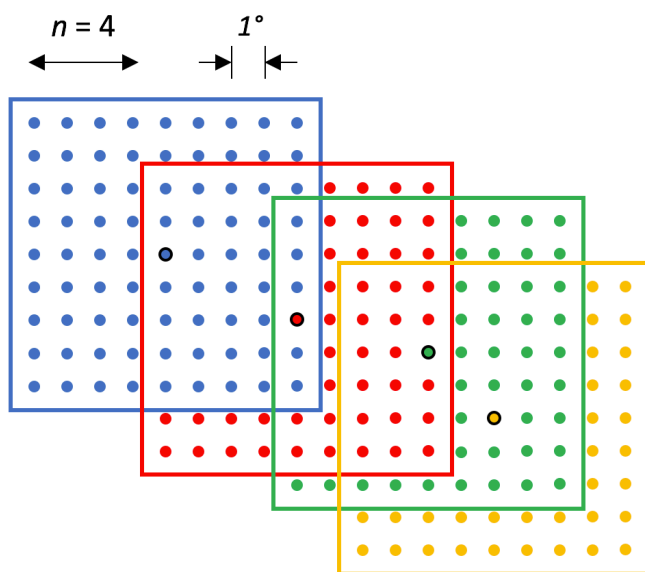


Fig. S3. Representation of a local scan performed by the algorithm with a scanning window of size  $2n+1$  with  $n=4$ . The squares are the scanning windows, the dots the points scanned. The central points are marked by a black circle and represent new minima. The points scanned on each neighbourhood are of the same color of the corresponding window. The yellow window makes the last search because the central point is a local minimum and total minimum in the actual scanned area.

neighbours in each direction. So, the scanning area contains  $(2n+1)^2$  points. The absolute maximum allowed variation of the  $\beta'$  angles from the central values is always set to  $4^\circ$  ( $n = 4$ ). This value is small enough to achieve a good performance of the algorithm because of the reduced number of optimizations performed for each scanned neighbourhood, while, at the same time, it is large enough to obtain the global minimum geometry for each pair of  $\alpha$  angles with a reduced number of consecutive calculations.

Because of the problem expressed above regarding the use of constrained angles in Gaussian, during the local search it is better to perform each new optimization starting from a previously calculated geometry, as closer as possible, in order to avoid convergence problems in Gaussian that easily emerge when there is too much difference between the initial and the desired dihedral angles. This means that our algorithm must also store all the computed converged geometries in the current neighbourhood.

The local search cannot be executed in a straightforward loop because each calculation must always start from the central point and the calculation for other points cannot start until a closest configuration has been finished and stored. Since each restricted optimization has a different computational cost and convergence particularities one cannot know beforehand how distant the closest calculated geometry will be. To handle these requirements in an easier manner, the virtual search window is separated in 3 parts that are handled distinctly: the upper region, the bottom region and the middle line passing through the central point. Although it is not possible to do a full check for all the results of the optimizations, each output geometry is partially checked measuring a selected set of C-H and C-C distances to detect the presence of broken or newly created bonds. Once an optimization converges and it passes this distance check, it is not necessary to calculate it again. During the local search, any result is stored in a file that allows the use of this information, decreasing the overall computational cost, since there are a lot of common points in the consecutive scanned neighbourhoods.

One further advantage in the efficiency of the algorithm is related to the early termination of the local search with respect to a total analysis of the neighbourhood because each time a local minimum in the border of the neighbourhood is found the search restarts (that point becomes the new central point). In this way, it is possible to skip some optimizations that may fail due to the impossibility to apply the constraints to the input geometry. In fact, during the initial testing of the algorithm it appeared that the calculations performed with dihedral angles far from the central value fail more easily. Avoiding as many such optimizations as possible further improves the performance of the algorithm.

## A.2. Full reduced PES

Given the considerations of the previous sections, the amount of computation necessary to obtain the reduced PES of a single TnH molecule is quite large. Evaluating optimistically that the mean number of optimizations per each surface point

is about 100 (81 neighbours) for a regular grid of the PES with a displacement of  $5^\circ$  between points that are calculated on the rhomboidal area indicated above, the total number of optimizations that must be performed to obtain a single PES is

$$100 \cdot ((120^\circ/5^\circ) \cdot (210^\circ/5^\circ))/2 \approx 50000$$

Due to the high number of optimizations required, to obtain the points of the surface in a reasonable time, it is necessary to employ a semi-empirical method that is, computationally speaking, much faster than Density Functional Theory based methods. Among the semi-empirical methods available in Gaussian, the Austin Model 1 (AM1) has been chosen because the same method was employed by Kelly in his theoretical calculations.<sup>3</sup>

The algorithm starts from one of the 6 isomeric minimum energy configurations of the molecule and recursively tries to compute new points on the PES that are close to already calculated points until all the points on the grid have been computed. Since it is assumed that each point is computed from an optimized geometry of one of the closest already computed points on the PES, it is not possible to compute all the points or even a large number of them in parallel. In fact, there is only a small number of points that can be calculated at any given moment, depending on the points of the PES that have already been computed. To parallelize the execution, the code we have developed is managed so that one process calculates a different point of the PES handling the concurrence with other processes for the computations that can be performed according to the existing points. This is done using the file system as a shared memory space between the nodes where the processes can be executed. All these calculations have been performed on a SUN cluster provided by the Institute of Theoretical and Computational Chemistry of the University of Barcelona (IQTC-UB). This machine contains hundreds of CPUs located in different nodes with access to the same physical memory. This kind of architecture allows the storage of the computed information about each geometry so that every running task can access it.

The final program is written in the C language and employs Gaussian and bash scripts, respectively, to perform the constrained optimizations and to analyze and extract the results of successfully terminated optimizations without the need for external libraries, aside from the standard C libraries. The algorithm starts from a minimum configuration of the TnH molecule that is analyzed to compute the first point in the PES, because the energy minimum is not among the points of the grid at which the PES is computed. To avoid concurrency problems, this step is performed only by the process that starts the computation first and the others wait for the first point to be found. This generates another bottleneck in the program but it does not greatly affect the total computational time.

After this first step, the processes can start to recursively calculate and explore neighbour points of the grid departing from one of the already calculated ones. This procedure is performed by two subroutines that check if the current scanned point is already calculated or if it cannot be calculated, that is, if the optimization in Gaussian does not converge or if there are any problems with the constraining of

the angles. These subroutines are employed to scan the rhomboid region in different directions, up and down, to avoid close loops that may appear when exploring the boundaries of this region. All values in the code are parametrized so that the program can obtain the full reduced PES for each one of the TnH molecules just by specifying the length of the helicene fragment and the displacement between the points on the PES as input parameters.

### B. Characteristic points

Gamess is employed to obtain the SPs of the molecules and to compute the IRC from those geometries. Points of the reduced PES where the extreme hydrogen atom of one of the helicene ends is over one of the triptycyl blades are employed as input geometries for the calculation of the transition states. The computed trajectories of the IRC are then analyzed to calculate the relative combinations of dihedral angles  $\alpha_R$  and  $\alpha_H$  corresponding to this RP and replicated according to the symmetry of the PES to allow its projection on the whole surface.

All computed energy data are represented on a scale relative to the energy of the minimum. It is usually done in this type of calculations and it is especially necessary when calculations are performed with different software. In both calculations (PES with Gaussian and IRC with Gamess) energy barriers should be, and are, approximately of the same value, considering that there is a small computational error. All the values of energy are expressed in kcal mol<sup>-1</sup>.

Due to the huge computational time needed to compute each PES, together with some bugs of the first versions of the code and some sporadic problems with the IQTC-UB cluster, even employing a large number of processes for each reduced PES calculation it was not possible to obtain the complete set of points of the surface of the three TnH molecules. However, the number of points obtained is close to the complete set, so that the computed points are already sufficient for a good graphical representation of the PES of these molecules using standard interpolation techniques.

As stated above, the search for a VRI is not usually easy, but since the PES is known as well as the TS involved in the rotation and enantiomerization processes, the position of VRI points can be expected to lie between these two TSs because the presence of a VRI indicates that the RP connects the two SPs. According to the definition of a VRI, the Hessian at this point has an eigenvector with a zero eigenvalue that is orthogonal to the gradient. This means that projecting the Hessian in the direction orthogonal to the gradient must give vibrational frequencies very close to 0. Given this consideration, in the approach to estimate the VRI, the first step is to find the intermediate point between the geometries of the TSs. This can be achieved obtaining the geometry for which each dihedral angle  $\beta$  equals the mean value of the corresponding values for the two TSs. In this way, the corresponding values of  $\alpha_R$  and  $\alpha_H$  in the resulting geometry are very close to the mean values of the  $\alpha_R$  and  $\alpha_H$  of the TSs. Once this geometry is found, an estimation of the VRI point can be obtained by slightly changing the dihedral angles in the molecule to obtain other geometries that are in the

neighbourhood of the middle point. This is achieved moving the two ends of the helicene fragment towards one or the other of the two TSs (4 possible combinations). Geometries are obtained through constrained optimizations with the Gaussian program, while the projection of the Hessian is performed using Gamess. The vibrational frequencies resulting from the projection are a measure of how much the corresponding molecular structure is close to the VRI. In particular, the seven frequencies corresponding, respectively, to projections of the Hessian with respect to translations, global rotations and gradient, must be very close to zero, while the frequency corresponding to the eighth vibrational mode must be as close to zero as possible.

### Notes and references

- 1 M. J. Frisch et al., "Gaussian 09, Revision B. 01," Gaussian Inc., Wallingford, CT, 2004.
- 2 M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347].
- 3 T. R. Kelly, I. Tellitu and J. P. Sestelo, *Angew. Chem. Int. Ed. Engl.*, 1997, **36**, 1866.