Conformational Analysis of Cinchonine and Cinchonidine by Tensor Decomposition of Molecular Dynamics Trajectories

Karlo Sović,1 Tea Ostojić,1 Sara Cepić,1 Alma Ramić,1 Renata Odžak,2 Mirjana Skočibušić,3 Tomica Hrenar,1,* Ines Primožič1,#

1 Department of Chemistry, Faculty of Science, University of Zagreb, Horvatovac 102a, HR-10000 Zagreb, Croatia
2 Department of Chemistry, Faculty of Science, University of Split, R. Boškovića 33, HR-21000 Split, Croatia
3 Department of Biology, Faculty of Science, University of Split, R. Boškovića 33, HR-21000 Split, Croatia

* Corresponding author’s e-mail address: hrenar@chem.pmf.hr
# Corresponding author’s e-mail address: ines@chem.pmf.hr

RECEIVED: July 8, 2019 ★ REVISED: September 8, 2019 ★ ACCEPTED: September 9, 2019

THIS PAPER IS DEDICATED TO PROF. KATA MLINARIĆ-MAJERSKI ON THE OCCASION OF HER 70TH BIRTHDAY

Abstract: Full conformational space of cinchonine and cinchonidine has been investigated by means of statistical analysis of quantum chemical molecular dynamics simulations. Recently developed procedure comprising principal component analysis of molecular dynamics trajectories was applied on cinchonine and cinchonidine as well as on their protonated and methylated quaternary derivatives. The method for full conformational analysis includes Cartesian coordinates sampling through quantum chemical molecular dynamics simulations, reduction of dimensionality by principal component analysis, determination of probability distributions in a reduced space of Cartesian coordinates and search for all the strict extrema points in probability distribution functions. In order to gain crucial insight in the understanding of chirality induction of these alkaloids, comparison of the determined conformational spaces of pseudo-enantiomers has been made. It was shown that protonation of the quinuclidine nitrogen atom stabilizes the conformers with the intramolecular 1N–H∙∙∙9O hydrogen bond whereas methylation on the same position results in the reduction of the domain of internal coordinates responsible for the conformational space.

Keywords: cinchonine, cinchonidine, full conformational analysis, quantum chemical molecular dynamics, principal component analysis.

INTRODUCTION

Cinchona alkaloids are well-known natural products isolated from bark of the Cinchona tree with a wide range of applications in chemistry and biology. They are used in stereoselective synthesis both in homogenous and heterogenous catalysis, as a chiral resolving agents or chiral auxiliaries in enantioselective chromatography. Furthermore, they have interesting biological properties which include antimalarial, anti-inflammatory and antiarrhythmic activity.

Their structure consists of a bicyclic aliphatic quinuclidine ring and aromatic quinoline ring which are connected through a chiral carbon atom bearing hydroxyl group. Stereochemistry of these alkaloids is complex since they contain five stereocenters, at C-(3), C-(4), N-(1), C-(8) and C-(9) atoms but differ in the configuration only at two carbon atoms, C-(8) and C-(9), Scheme 1. Naturally occurring Cinchona alkaloids occurs in diastereomeric pairs, but they behave like pseudo-enantiomers when they are used as catalysts in asymmetric organic reactions. Understanding of their structural behaviour can provide crucial information on chiral inducing and discriminating ability of these alkaloids.

Because of their structure and stereochemistry, Cinchona alkaloids have complex conformational behaviour and a number of conformers are possible. Knowledge about the dominant conformers can be crucial for understanding reaction pathways and stereochemical outcome of the catalyzed reactions. Conformational behaviour of a few Cinchona alkaloids has been previously investigated with experimental techniques - NMR spectroscopy, NOESY experiments in D6-acetone suggested the presence of 3 conformers but due to the small population of other
conformers, their presence in solution is difficult to confirm.\(^{[15]}\)

Conformational investigations of this class of alkaloids, based on computational methods, have been undertaken by Dijkstra et al.\(^{[16]}\) in 1989 by performing molecular mechanics (MM) calculations in vacuo with MM2P and MMX force fields. Later, the interest shifted towards the understanding of the conformational space of these alkaloids on a Pt(111) surface and such investigation was conducted by Angelo Vargas and Alfons Baiker\(^{[17]}\) using fully relativistic Hamiltonian, including spin-orbit coupling. In 2008, conformational behaviour of cinchonidine was revisited by Baiker et al.\(^{[10]}\) with DFT metadynamics by using Car-Parrinello scheme.

The goal of this work is to investigate potential energy surfaces of cinchonidine derivatives (cinchonidine, CD-1; cinchonidinium CD-2; 1-methyl cinchonidinium, CD-3) and cinchonine derivatives (cinchonine, CN-1; cinchoninium CN-2; 1-methyl cinchoninium, CN-3) (Scheme 2), in vacuo to make a comparison of the determined conformational spaces on a semiempirical level of theory, using PM7 Hamiltonian.\(^{[18]}\) We used an adapted approach with the statistical analysis of the molecular dynamics trajectory using a dimensionality reduction and search through the probability distribution of molecular geometries in reduced space.

**EXPERIMENTAL**

Molecular dynamics simulations were performed using on-the-fly calculations of forces in each point of the simulation and velocity Verlet algorithm for integration.\(^{[21]}\) The forces were calculated using the PM7 Hamiltonian\(^{[17]}\) implemented in MOPAC2016.\(^{[20,21]}\) All molecular dynamics simulations were performed using our own program qcc.\(^{[22–26]}\) To ensure that most of the phase space relevant for the conformational space of the investigated compounds was adequately sampled, the initial temperature for Maxwell distribution of velocities was set at 798.15 K and this temperature was kept constant throughout the simulations using a velocity scaling algorithm. The step size was 0.5 fs and total length of simulation was 2.5 ns (total of 5 000 000 steps).

Each molecular dynamics run for a molecule with \(N\) atoms produced a 2\(^{\text{nd}}\) order data tensor with dimensions \(\text{number-of-steps} \times 3N\). Principal component analysis was used as a dimensionality reduction tool and performed using a NIPALS algorithm implemented in our own program moonee.\(^{[28–30]}\) Probability distributions were generated in a reduced space using the \(n\)-dimensional parallelepiped and counting all the points that belong in it. These probability distributions were then subjected to a procedure for finding all local maxima that correspond to local minima on the corresponding potential energy surface. A parallelized combinatorial optimization algorithm for the arbitrary number of ways (dimensions) implemented in moonee was used. A set of structures determined from probability distributions was used as an initial guess for the full conformational analysis.\(^{[31]}\) All structures from the initial guess were optimized and subsequently clustered. Optimizations of these structures were performed using the B3LYP-D3 functional and 6-311++G(d,p) basis set. Harmonic frequency analyses were performed, and standard Gibbs energies of formation were calculated at \(T = 298.15\) K and \(p = 101325\) Pa. All quantum-chemical calculations were performed using the Gaussian 16 program.\(^{[32]}\)

**RESULTS AND DISCUSSION**

Statistical Analysis of Molecular Dynamics Trajectories

For all investigated compounds trajectory of 5 000 000 steps was sampled by conducting molecular dynamics simulation using on-the-fly calculations of forces. Extraction

### Scheme 1
Structure of cinchonidine (CD) and cinchonine (CN) with the chiral centres highlighted.

### Scheme 2
Investigated Cinchonidine (CD) derivatives (cinchonidine, CD-1; cinchonidinium CD-2; 1-methyl cinchonidinium, CD-3) and Cinchonine (CN) derivatives (cinchonine, CN-1; cinchoninium CN-2; 1-methyl cinchoninium, CN-3).
of initial guess structures for full conformational analysis was performed by determining all strict local maxima in the probability distribution function of molecules geometries. These points represented a sampling space from which the initial guess structures for conformational analysis were extracted. This extraction was performed by finding all strict local maxima in the probability distribution of the molecular geometry coordinates. Since the molecule spends more time around the minima points on the PES, probability distribution function in these parts of phase space would have strict local maximum that correspond to strict local minima on the corresponding potential energy surface.

Dimensionality of this search for strict extrema problem is of the order $3N$, where $N$ is the number of atoms in the molecule. To reduce the problem to fewer dimensions, 2nd order tensor decomposition tool principal component analysis (PCA) was used. In the PCA data matrix, $X$ of rank $r$ is decomposed as a sum of $r$ matrices $t_i p_i$ of rank 1:

$$X = \sum_{i=1}^{r} t_i p_i$$  \hspace{1cm} (1)

PCA enables one to find the best linear projections for a high dimensional set of data in the least-squares sense. Scores $t_i$ represent projections of the original points on the principal component (PC) direction and can be used for classification or building of probability distributions, whereas loadings $p_i$ represent the eigenvectors of data covariance (or correlation) matrix and can be used for the identification of variability among the data. More details on PCA can be found in the literature."}\textsuperscript{[33]}

To determine the final number of principal components and total number of steps necessary for conducting the full conformational analysis, we calculated the strict local maxima plateau during the course of simulation\textsuperscript{[26]} by following the procedure presented on Scheme 3.

For CD-1 this plateau was reached at approximately 2,000,000 points in the simulation and with 4 principal components, Fig. 1. After this point in the probability density distribution, the number of strict local maxima remains constant. These 4 principal components were describing 84.58% of the total variance (Table 1) and the total number of initial guess structures obtained for CD-1 was 79. Total number of initial guess structures for CD-2 and CD-3 were 84 and 48, respectively.

For CN-1 this plateau was again reached at approx. 2,000,000 points in the simulation and with 4 principal components, Figure 2. These 4 principal components were...
describing 87.60 % of the total variance (Table 2.) and the
total number of initial guess structures for CN-1 was 59.
Total number of initial guess structures for CN-2 and CN-3
were 77 and 48, respectively.

**Full Conformational Analysis of CD-1, CD-2, and CD-3**

In cases of CD-1, CD-2, and CD-3, search for strict local
maxima in the probability distribution was performed in
4-dimensional space. Since the total variance explained in
these first four PCs was more than 80 % (Table 1.), majority
of the phase space was investigated thus insuring the
completeness of conformational analyses. Each of these
strict local maxima corresponds to a structure that is in
vicinity of the extrema point on PES. Geometry
optimization for these points were conducted at B3LYP-
D3/6-311++G(d,p) level of theory. Optimized geometries
for each compounds were subjected to the clustering
procedure and for CD-1 five conformers with the
abundance larger than 5 % were obtained (Table 3, Figure
3a). Conformational space is mainly determined by the
position of quinoline ring and orientation of O–H group
(Figure 4 a ). It is confirmed that on this high level of the
theory, lowest energy conformer is the one labelled as

<table>
<thead>
<tr>
<th>Principal component</th>
<th>CD-1</th>
<th>CD-2</th>
<th>CD-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC01</td>
<td>35.51</td>
<td>35.51</td>
<td>32.72</td>
</tr>
<tr>
<td>PC02</td>
<td>27.13</td>
<td>62.64</td>
<td>30.61</td>
</tr>
<tr>
<td>PC03</td>
<td>11.94</td>
<td>74.59</td>
<td>10.99</td>
</tr>
<tr>
<td>PC04</td>
<td>10.00</td>
<td>84.58</td>
<td>11.41</td>
</tr>
<tr>
<td>PC05</td>
<td>6.69</td>
<td>91.28</td>
<td>90.66</td>
</tr>
<tr>
<td>PC06</td>
<td>3.03</td>
<td>94.30</td>
<td>92.77</td>
</tr>
<tr>
<td>PC07</td>
<td>1.54</td>
<td>95.84</td>
<td>94.46</td>
</tr>
<tr>
<td>PC08</td>
<td>1.22</td>
<td>97.07</td>
<td>95.88</td>
</tr>
<tr>
<td>PC09</td>
<td>1.12</td>
<td>98.19</td>
<td>97.07</td>
</tr>
<tr>
<td>PC10</td>
<td>0.29</td>
<td>98.48</td>
<td>97.74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Principal component</th>
<th>CD-1</th>
<th>CD-2</th>
<th>CD-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC01</td>
<td>36.35</td>
<td>36.35</td>
<td>33.14</td>
</tr>
<tr>
<td>PC02</td>
<td>28.92</td>
<td>65.27</td>
<td>30.00</td>
</tr>
<tr>
<td>PC03</td>
<td>11.36</td>
<td>76.63</td>
<td>11.58</td>
</tr>
<tr>
<td>PC04</td>
<td>10.97</td>
<td>87.60</td>
<td>10.57</td>
</tr>
<tr>
<td>PC05</td>
<td>5.17</td>
<td>92.77</td>
<td>90.96</td>
</tr>
<tr>
<td>PC06</td>
<td>1.78</td>
<td>94.55</td>
<td>92.86</td>
</tr>
<tr>
<td>PC07</td>
<td>1.68</td>
<td>96.24</td>
<td>94.38</td>
</tr>
<tr>
<td>PC08</td>
<td>1.12</td>
<td>97.36</td>
<td>95.74</td>
</tr>
<tr>
<td>PC09</td>
<td>0.81</td>
<td>98.17</td>
<td>96.96</td>
</tr>
<tr>
<td>PC10</td>
<td>0.42</td>
<td>98.60</td>
<td>97.67</td>
</tr>
</tbody>
</table>

**Figure 2.** Total number of determined strict local maxima in the probability distribution of CN-1 geometries in reduced space calculated by principal component analysis of molecular dynamics trajectory.
Open(3) in previous work.[10] For CD-2, and CD-3, five and three conformers were obtained (Table 4 and Table 5, Figure 3b and Figure 3c). Energy distribution of conformers shows that an addition of proton to quinuclidine moiety causes stabilization of the Gibbs energies of formation for conformers relative to the most stable one. The most stable conformer for CD-2 is the one with intramolecular hydrogen bond N–H⋯O. Addition of the methyl group does not allow creation of this hydrogen bond and the lowest energy conformers are defined by the position of the vinyl group (Figure 3c and Figure 4a).

### Table 3. Relative standard Gibbs energies of formation and relative abundance for conformers of compound CD-1 calculated by using B3LYP functional and 6-311++G(d,p) basis set with D3 version of Grimme’s dispersion. (at \( T = 298.15 \) K and \( p = 101325 \) Pa).

<table>
<thead>
<tr>
<th>Conformer</th>
<th>( \Delta G^\circ / \text{kJ mol}^{-1} )</th>
<th>Abundance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.5932</td>
</tr>
<tr>
<td>2</td>
<td>4.923</td>
<td>0.0814</td>
</tr>
<tr>
<td>3</td>
<td>4.985</td>
<td>0.0794</td>
</tr>
<tr>
<td>4</td>
<td>5.474</td>
<td>0.0652</td>
</tr>
</tbody>
</table>

* Only conformers with abundance ≥ 5% are presented.

### Table 4. Relative standard Gibbs energies of formation and relative abundance for conformers of compound CD-2 calculated by using B3LYP functional and 6-311++G(d,p) basis set with D3 version of Grimme’s dispersion. (at \( T = 298.15 \) K and \( p = 101325 \) Pa).

<table>
<thead>
<tr>
<th>Conformer</th>
<th>( \Delta G^\circ / \text{kJ mol}^{-1} )</th>
<th>Abundance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.2454</td>
</tr>
<tr>
<td>2</td>
<td>0.901</td>
<td>0.1706</td>
</tr>
<tr>
<td>3</td>
<td>1.107</td>
<td>0.1570</td>
</tr>
<tr>
<td>4</td>
<td>1.606</td>
<td>0.1284</td>
</tr>
<tr>
<td>5</td>
<td>2.630</td>
<td>0.0849</td>
</tr>
</tbody>
</table>

* Only conformers with abundance ≥ 5% are presented.

Full Conformational Analysis of CN-1, CN-2, and CN-3

Similar to the previous case, the search for strict local maxima in the probability distribution of CN-1, CN-2, and CN-3 was performed in 4-dimensional space. Total variance explained in these first four PCs was again more than 80% (Table 2.) that insure that the most important parts of the phase space were investigated. Optimized geometries at the B3LYP-D3/6-311++G(d,p) level of theory for each compounds were subjected to the clustering procedure and for CN-1 three conformers with the abundance larger than 5% were obtained (Table 6, Figure 5a). As in the case of CD-1 conformational space is determined by the position of quinoline ring and the orientation of O–H group (Figure 4b). For CN-2, and CN-3, five and two conformers were obtained respectively (Table 7 and Table 8, Figure 5b and Figure 5c). Energy distribution of conformers confirms previous findings that an addition of proton to quinuclidine moiety causes stabilization of the Gibbs energies of formation for conformer with intramolecular hydrogen bond whereas addition of the methyl group leads to greater energy separation for conformers. Addition of the methyl group does not allow creation of this hydrogen bond and the lowest energy conformers are defined by the position of the vinyl group (Figure 4b).

### Table 6. Relative standard Gibbs energies of formation and relative abundance for conformers of compound CN-1 calculated by using B3LYP functional and 6-311++G(d,p) basis set with D3 version of Grimme’s dispersion. (at \( T = 298.15 \) K and \( p = 101325 \) Pa).

<table>
<thead>
<tr>
<th>Conformer</th>
<th>( \Delta G^\circ / \text{kJ mol}^{-1} )</th>
<th>Abundance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.5251</td>
</tr>
<tr>
<td>2</td>
<td>1.316</td>
<td>0.3088</td>
</tr>
<tr>
<td>3</td>
<td>3.990</td>
<td>0.1060</td>
</tr>
</tbody>
</table>

* Only conformers with abundance ≥ 5% are presented.

### Table 7. Relative standard Gibbs energies of formation and relative abundance for conformers of compound CN-2 calculated by using B3LYP functional and 6-311++G(d,p) basis set with D3 version of Grimme’s dispersion. (at \( T = 298.15 \) K and \( p = 101325 \) Pa).

<table>
<thead>
<tr>
<th>Conformer</th>
<th>( \Delta G^\circ / \text{kJ mol}^{-1} )</th>
<th>Abundance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.4680</td>
</tr>
<tr>
<td>2</td>
<td>4.021</td>
<td>0.0924</td>
</tr>
<tr>
<td>3</td>
<td>4.073</td>
<td>0.0905</td>
</tr>
<tr>
<td>4</td>
<td>4.270</td>
<td>0.0836</td>
</tr>
<tr>
<td>5</td>
<td>4.299</td>
<td>0.0826</td>
</tr>
</tbody>
</table>

* Only conformers with abundance ≥ 5% are presented.
Figure 3. Conformational space of compound a) CD-1, b) CD-2 and c) CD-3 obtained from 4-dimensional probability distribution space calculated by principal component analysis of the molecular dynamics trajectory. Standard Gibbs energies of formation were computed at $T = 298.15$ K and $p = 101325$ Pa (B3LYP-D3/6-311++G(d,p)).

Figure 4. Total number of determined strict local maxima in the probability distribution of CN-1 geometries in reduced space calculated by principal component analysis of molecular dynamics trajectory.

Table 8. Relative standard Gibbs energies of formation and relative abundance for conformers of compound CN-3 calculated by using B3LYP functional and 6-311++G(d,p) basis set with D3 version of Grimme’s dispersion. (at $T = 298.15$ K and $p = 101325$ Pa).

<table>
<thead>
<tr>
<th>Conformer</th>
<th>$\Delta G^\circ$ / kJ mol$^{-1}$</th>
<th>Abundance$^{(a)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.7343</td>
</tr>
<tr>
<td>2</td>
<td>3.052</td>
<td>0.2144</td>
</tr>
</tbody>
</table>

$^{(a)}$ Only conformers with abundance ≥ 5% are presented.
CONCLUSION

Recently developed method for the full conformational analysis was applied to determine conformational space of pseudo-enantiomers of Cinchona alkaloids as well as their quaternary N-protonated and N-methyl derivatives. Probability distributions of molecular conformations were generated in the 4-dimensional reduced space using the tensor decomposition of Cartesian coordinates space sampled by quantum chemical molecular dynamics simulations. Dimension of the reduced space and length of simulations were estimated by constructing strict local maxima plateau.

For calculated probability distributions extrema points were determined. Each particular strict local maximum point in probability distribution was mapped to the corresponding potential energy surface and used as an initial guess structure for the full conformational analysis. Since the described method of full conformational analysis can be applied to any cyclic/noncyclic molecule, conformational spaces for investigated molecules were determined in dependence on free single bonds torsional movement and concerted torsional bond movement (quinuclidine ring) present in the investigated compounds.

Figure 5. Conformational space of compound a) CN-1, b) CN-2 and c) CN-3 obtained from 4-dimensional probability distribution space calculated by principal component analysis of the molecular dynamics trajectory. Standard Gibbs energies of formation were computed at $T = 298.15$ K and $p = 101325$ Pa (B3LYP-D3/6-311++G(d,p)).
Protonation of quinuclidine nitrogen atom caused separation of the conformer energy levels and stabilization of the conformers with intramolecular hydrogen bond. Addition of the methyl group at the same position revealed that domain of the conformational space was mainly determined by the position of the vinyl group. While the conformational spaces of pseudo-enantiomers were mainly determined by the position of the quinoline rings, quaternization of each pseudo-enantiomer did not result in a different position of that moiety in the lowest energy conformers.

**Acknowledgment.** This work was supported by the Croatian Science Foundation, Project No: IP-2016-06-3775 ADESIRE.

**REFERENCES**


