This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proofreading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This Accepted Manuscript will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Structural revision of glabramycins B and C, antibiotics from the fungus Neosartorya glabra by DFT calculations of NMR chemical shifts and coupling constants

Yang Li*

Department of Chemistry, Yale University, 225 Prospect Street, P.O. Box 208107, New Haven, CT 06520.

1 Author’s present address: Georgetown University Law Center, 600 New Jersey Ave NW, Washington, DC 20001. Tel: (202) 643-2539. Email: yl625@georgetown.edu
The $^{13}$C NMR spectra and vicinal proton-proton coupling constants of two tricyclic macrolactone natural products were analyzed using computational methods, which resulted in their structural revisions.
Abstract
Glabramycins B and C are antibacterial natural products produced by the fungal organism *Neosartorya glabra*. Their stereochemical structures have not been completely defined to date. In this work, DFT calculations are employed to predict the expected Carbon-13 NMR chemical shifts and key vicinal proton-proton coupling constants for all of the candidate stereoisomers. By comparison with experimentally measured values, the complete relative stereochemical configurations for glabramycins B and C are established.

Keywords: *Neosartorya glabra*, Glabramycin, structural revision, macrolactone

1. Introduction

Glabramycins B and C are two macrolactone natural products first isolated from the fungus *Neosartorya glabra*, in an effort to systematically discover clinically useful antibiotics. Both glabramycins exhibit antibacterial activity against *S. pneumoniae*; additionally, glabramycin C possesses further antibacterial activities against *S. aureus* and *B. subtilis*. It has come to our attention that the glabramycins bear striking structural similarities to another natural product Sch-642305. In addition to the structural similarity, the two organisms that produce these natural products are also closely related in taxonomy, *Penicillium verrucosum*, which produces Sch-642305, and *Neosartorya glabra*, which produces glabramycins, are both members of the *Trichocomaceae* family of fungi. Yet when comparing the structures of Sch-642305 and glabramycin C, it is noticeable that the two molecules have different relative configurations at the critical ring junction C11 (Figure 1). Furthermore, the relative configuration of the methyl group at C20 in both glabramycins B and C was not determined in the original isolation report. Therefore the aim of this study is to elucidate the stereochemical configurations at both C11 and C20 in the two compounds glabramycin B and C.

![Figure 1](image-url)

Figure 1. Structures of Sch-642305, glabramycin B, and glabramycin C with relative configurations and atom numbering indicated. Structural differences are highlighted.

Traditionally total synthesis is one of the most important methods (and sometimes the only one) for resolving such stereochemical ambiguities. However, while the structure of Sch-642305 has been confirmed beyond any doubt by several total syntheses, no such efforts have been reported for the glabramycins. Undoubtedly, synthesizing the compounds in question requires significant efforts. Thus while total synthesis is the golden standard for structural proof and is still reserved as a method of last resort, almost all molecular structures are currently determined by spectroscopic methods, the most important of which is NMR. The measured NMR parameters such as chemical shifts and coupling constants can yield a wealth of information about atom connectivity, geometry, and functionalities present in the molecule in question. Although organic chemists tend to conceive these parameters as semi-empirical quantities, rigorous physical theories defining them have been proposed since the earliest days of NMR. In the decades following the 1950s, along with theoretical developments, calculations of NMR parameters from first principles were attempted. These calculations were mostly carried out by specialists on very simple molecules up to approximately the last decade. In recent years, the exponential increase of computer speed and the wide availability of relatively easy-to-use software packages have made highly accurate computational chemistry methods accessible to the practicing bench chemist. The *ab initio* calculation of NMR chemical shifts has proven in numerous cases to be a useful tool in the structural elucidation of natural products, such as in the cases of viridiol, samoquasine A, spiroleucettadine, hassananes, and obtusallenes. By comparing the predicted spectroscopic parameters of candidate structures to the experimentally observed values, the correct structure can often be identified. Statistical techniques have also been applied in the comparison of calculated and experimental values, so that the “matchness” can be quantified in probabilistic terms.

It is worth noting that this *ab initio* method is different from the complementary approach of predicting NMR shifts from existing data points. The *ab initio* method, while orders of magnitude slower, can handle unusual molecules and give results with more reliability and accuracy.
2. Results and discussion

2.1 Structural revision of glabramycin C

2.1.1 Calculation of chemical shifts in $^{13}$C NMR

Given the availability of a wealth of NMR parameters disclosed in the original isolation report,1 we chose to pursue a computational NMR investigation on the structures of glabramycins B and C. We started the investigation with glabramycin C and identified four possibilities for its true relative configuration. They arise from the diastereomeric combination of the stereochemical configurations at the ring junction C11 and the methyl C20. These four cases were labeled (11R,20R), (11R,20S), (11S,20R), and (11S,20S) (structures shown in Figure 2). Due to the polycyclic nature of these structures, there was only one low-energy conformer for each of the candidate isomers, as we found out in conformational searches using molecular mechanics. The geometries of the four structures were minimized at the B3LYP/6-31G** level, and their $^{13}$C NMR chemical shifts were then calculated at the GIAO/mPW1PW91/6-31G** level of theory. While B3LYP is widely recognized for giving excellent molecular geometry, its performance in predicting $^{13}$C NMR chemical shifts is questionable. In a benchmarking comparison of the accuracy of density functionals for the prediction of $^{13}$C NMR chemical shifts of 15 natural products, Cimino and coworkers found that the performance of B3LYP was inferior to that of mPW1PW91.33 Wu and coworkers similarly concluded that the LYP correlation functional was less accurate than PW91 for NMR calculations in a more extensive comparison of 21 exchange-correlation functionals.34 We therefore chose mPW1PW91 as the density functional to carry out the NMR calculations. The particular basis set was chosen as a good compromise between speed and accuracy. The expected mean absolute error is less than 1.5 ppm from the experimental values for this particular combination of density functional and basis set.33 The IEF-PCM solvation model was employed in these steps to enhance the accuracy of the calculation.32, 35

The calculations produced a set of predicted $^{13}$C chemical shifts for each one of the four candidate stereoisomers of glabramycin C. The results are plotted in Figure 2. The differences (in ppm) between the calculated and observed $^{13}$C chemical shift are shown as vertical bars for each of the carbon atoms. Surprisingly, the predicted $^{13}$C NMR shifts for all four of the candidate stereoisomers deviated significantly from the experimental measurements, far beyond the range of expected calculation errors. We eventually traced the problem to two sources. Firstly, the predictions for the carboxylate C1 were off by as much as 6–8 ppm. This was most likely due to the neglect of carboxylic acid dimerism in the calculation, as documented in previous computational36, 37 and experimental studies.38 However, modeling the dimeric form of glabramycin C would be prohibitively expensive because the computational cost of DFT calculations increases as $N^3$ for a system of $N$ electrons.39 We therefore chose to simply exclude the C1 chemical shift as a data point in all subsequent statistical comparisons.35 Secondly, we came to the conclusion that the C18 and C19 $^{13}$C NMR peaks were most likely misassigned in the original isolation report and had to be exchanged. Because their corresponding $^1$H NMR peaks were very close, the original assignments of C18 and C19 by HMBC could have easily been confused. Other researchers have also encountered similar issues of spectrum misassignment when comparing calculated and experimental $^{13}$C NMR shifts.40

[Figure 2]

Figure 2. Comparing calculated and observed $^{13}$C NMR chemical shifts for the four proposed stereochemical models of glabramycin C. Chemical shift differences in ppm are plotted for each C atom in the molecule. The chemical shift values used to plot the charts here were scaled by linear regression against the observed values (See experimental section 4.2). Note that in all four plots the predicted chemical shifts for the carboxylate C1 atoms were all underestimated; and the predicted chemical shifts for C18 and C19 atoms are consistently under- and overestimated, respectively, indicative of misassignment.

With these two issues rectified, the calculation data were reprocessed, and the resulting predictions were again compared against the experimental chemical shifts (Figure 3). The 11R isomers had nearly twice the mean absolute errors as the 11S isomers. The errors were mostly concentrated in the region of C10 to C15. This was the region of the six-membered ring of glabramycin C, and therefore was expected to show the most deviations had the C11 ring junction stereochemistry been misassigned. This suggested that 11S was most likely to be the correct relative configuration. It was also noticed that the two 11S isomers had relatively small errors, even though the C20 NMR shift was predicted to be 7 ppm more downfield than the experimental value in the (11S,20R) model. However, this error was
absent in the (11S,20S) isomer, suggesting that 20S was most likely to be the true relative configuration of glabramycin C. Overall, based on the mean absolute error and the maximum error, the (11S,20S) isomer was overwhelmingly more favorable than the other three possibilities. The calculations were also repeated with several other combinations of density functionals, basis sets and solvation models (Table 1), and the same trend was observed. In Table 1, entry 3 employed a different solvation model (CPCM), while the calculations in entries 1 and 2 were done in a vacuum. Entry 4 used a different functional and basis set (B3LYP/cc-pVDZ) for the calculation of chemical shifts. Different methods for geometry optimization were also explored. Entries 2 and 3 used mPW1PW91/6-31G**, while entry 1 was done using a non-DFT method (HF/3-21G). In all of the cases, the (11S,20S) isomer always had the smallest mean absolute error (MAE) and maximum error (ME) when compared with the other candidate isomers.

[Figure 3]
Figure 3. Comparison between calculated and observed $^{13}$C NMR chemical shifts for the four proposed stereochemical models of glabramycin C, after excluding C1 atom and swapping the spectral assignments for C18 and C19 atoms. The chemical shift values used to plot the charts here were scaled by linear regression against the observed values (See experimental section 4.2).

[Table 1]
Table 1. Evaluating the calculated $^{13}$C chemical shifts for the four candidate isomers of glabramycin C against experimental values. MAE (mean absolute error) and ME (maximum error) values were calculated after linear regression corrections.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Geometry</th>
<th>NMR</th>
<th>vacuum</th>
<th>CHCl$_3$</th>
<th>CHCl$_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HF/3-21G</td>
<td>GIAO</td>
<td>mPW1PW91/6-31G**</td>
<td>mPW1PW91/6-31G**</td>
<td>mPW1PW91/6-31G**</td>
</tr>
<tr>
<td>2</td>
<td>mPW1PW91/6-31G**, vacuum</td>
<td>GIAO</td>
<td>mPW1PW91/6-31G**, vacuum</td>
<td>mPW1PW91/6-31G**, vacuum</td>
<td>mPW1PW91/6-31G**, vacuum</td>
</tr>
<tr>
<td>3</td>
<td>mPW1PW91/6-31G**, vacuum</td>
<td>GIAO</td>
<td>mPW1PW91/6-31G**, vacuum</td>
<td>mPW1PW91/6-31G**, CPCM CHCl$_3$</td>
<td>mPW1PW91/6-31G**, CPCM CHCl$_3$</td>
</tr>
<tr>
<td>4</td>
<td>B3LYP/6-31G**, IEF-PCM CHCl$_3$</td>
<td>GIAO</td>
<td>B3LYP/cc-pVDZ, IEF-PCM CHCl$_3$</td>
<td>GIAO mPW1PW91/6-31G**, IEF-PCM CHCl$_3$</td>
<td>GIAO mPW1PW91/6-31G**, IEF-PCM CHCl$_3$</td>
</tr>
<tr>
<td>5</td>
<td>B3LYP/6-31G**, IEF-PCM CHCl$_3$</td>
<td>GIAO</td>
<td>mPW1PW91/6-31G**, IEF-PCM CHCl$_3$</td>
<td>mPW1PW91/6-31G**, IEF-PCM CHCl$_3$</td>
<td>mPW1PW91/6-31G**, IEF-PCM CHCl$_3$</td>
</tr>
</tbody>
</table>

Mean absolute error in ppm.

Maximum error in ppm.

Not calculated.

2.1.2 Calculation of $^{3}J_{HH}$ coupling constants in $^{1}$H NMR

While $^{13}$C NMR chemical shifts are certainly influenced by stereochemical isomerism, the proton-proton vicinal coupling constants ($^{3}J_{HH}$) are much more direct indicators of stereochemistry. They are more amenable to interpretation by organic chemists and are used extensively in the determination of dihedral angles. However, calculation of indirect spin-spin coupling constants is significantly more challenging than that of chemical shifts. There are several distinct mechanisms that contribute to the coupling constants. The interactions of the nuclear magnetic fields with the orbital movement of electrons are represented by the diamagnetic spin-orbit term and the paramagnetic spin-orbit term; the interactions between the electronic spins and the nuclear magnetic fields are represented by the spin-dipole term and the Fermi-contact term. The relative magnitudes of these terms vary depending on the types of nuclei involved and the chemical bonds between them, among many other factors. Therefore, all of the terms may have to be accounted for in order to achieve accurate results, and consequently can require much more computational resources than calculating NMR shifts. Furthermore, it is often observed that very large basis sets are needed to obtain accurate results. For chemists with only modest computational resources, these calculations are rarely undertaken. Therefore, the calculation of NMR coupling constants for the structural assignment of natural products has only been employed in a few occasions.

Bally and Rablen investigated this challenging problem in great detail. They systematically benchmarked the performance of a large number of available methods and procedures for the calculation of proton-proton coupling constants. Several simplifications were identified that greatly improved the computational efficiency. They found that for calculating the $^{3}J_{HH}$ coupling constant in the context of the usual organic molecules, it was sufficient to only consider the Fermi-contact term among the many contributing terms, and then scale the resulting number empirically to make up for the other contributions. They also advocated using a “mixed basis set” for the calculation. In their approach, the 1s part of the basis set for H atoms is decontracted and augmented to improve the description of
electron density in the inner region of the atom, while all other atoms are treated by the standard 6-31G** Pople basis set to reduce the computational cost. They reported that accurate results (within 0.5 Hz) can be obtained in very reasonable lengths of time using their method.

Encouraged by their report, and also noting that the conformational rigidity of the polycyclic glabramycins was advantageous for the calculation of \( J_{10,11} \) values, we computed the theoretical values of \( J_{10,11} \) and \( J_{11,12} \) for all four possible candidate structures of glabramycin C (Table 2). Upon comparison with experimental measurements, it was clear that the 11S models in general agreed with the experimental values. One of the four models (11S,20S) had a particularly close match, with a difference of no more than 0.3 Hz from the experimental values. The coupling constant calculation, taken together with the \( ^{13}C \) NMR chemical shift calculation, strongly suggests that the (11S,20S) model is the correct stereochemical structure of glabramycin C.

Table 2. Experimental and calculated \( J_{10,11} \) values (Hz) for the four stereochemical models of Glabramycin C.

<table>
<thead>
<tr>
<th>Isomer</th>
<th>Experimental</th>
<th>Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>11R,20R</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td>11R,20S</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>11S,20R</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>11S,20S</td>
<td>1.4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

\(^{a}\)Data extracted from isolation report.\(^{b}\)Dihedral angle of H10-C10-C11-H11 in geometry optimized models.\(^{c}\)Dihedral angle of H11-C11-C12-H12 in geometry optimized models.

2.2 Structural revision of glabramycin B

We next turned our attention to glabramycin B, which is almost identical to glabramycin C except for a saturated C12-C13 bond (Figure 1). We repeated the \( ^{13}C \) NMR shift calculation on this compound, and found that the NMR peaks for C12 and C13 were again misassigned and should be swapped, as in the case of glabramycin C (see above §2.1.1). The results of the calculations, after making the adjustment, are shown in Figure 4. The (11S,20S) isomer is much closer to the experimental values than the other three candidates.

Figure 4. Comparison between the calculated and the observed \( ^{13}C \) NMR chemical shifts for the four proposed stereochemical models of glabramycin B. The chemical shift values used to plot the charts here were scaled by linear regression against the observed values (See experimental section 4.2).

3. Conclusion

In conclusion, through the computational analysis of \( ^{13}C \) chemical shifts and \( J_{10,11} \) coupling constants in the NMR spectra of glabramycins B and C, we have revised the relative stereochemical assignment at the critical C11 ring junction of the tricyclic core of the molecules. Furthermore, the originally undefined methyl stereocenter at C20 has now been established.

Despite all of the developments in chromatographic and spectroscopic methodologies, the isolation and structural determination of complex natural products are still universally recognized as challenging tasks, and especially so in the determination of stereochemistry. The studies presented here serve as an additional example in a growing repertoire of works that demonstrate the utility of computational tools in aiding the efforts of structural determination.

4. Experimental section

4.1 Conformational search

A large number of conformers with random torsional angles were generated and then optimized using MMFF as implemented in MacroModel. A number of the lowest energy conformers were then optimized again with DFT to verify their relative energies. The second lowest conformers for the four diastereomers were >1 kcal/mol above the ground conformers and were therefore not included in subsequent calculations in view of their small Boltzmann weights.
4.2 $^{13}$C NMR chemical shift

The molecular structures were first minimized by MMFF as implemented in MacroModel.$^{50}$ All torsional angles were systematically varied to search for the global minimum. The molecular mechanics minimized structures were then imported to Gaussian$^{51}$ for all subsequent steps. Geometry optimization was conducted at the B3LYP/6-31G(d,p) level of theory. IEF-PCM solvation in chloroform was used for the glabramycin C stereoisomers; methanol was used for the glabramycin B stereoisomers. Chloroform and methanol were the solvents in which experimental NMR spectra were acquired. $^{13}$C NMR shielding constants $\sigma$ were calculated at the GIAO/mPW1PW91/6-31G(d,p) level, with IEF-PCM solvation in the appropriate solvent. Chemical shifts were calculated as $\delta_{\text{calc}} = \sigma_{\text{ref}} - \sigma$, where $\sigma_{\text{ref}} = 196.594$ was the chemical shielding constant of tetramethylsilane calculated at the same level of theory. The values of $\delta_{\text{calc}}$ were then scaled linearly: $\delta_{\text{scaled}} = a + b\delta_{\text{calc}}$. The scaling constants $a$ and $b$ were determined by least squares linear regression between the calculated chemical shifts $\delta_{\text{calc}}$ and the experimental ones $\delta_{\exp}$. The differences between the theoretical and the experimental chemical shifts, $\Delta\delta = \delta_{\text{scaled}} - \delta_{\exp}$, were calculated for each carbon atom and plotted in Figures 2 to 5. Mean absolute error was calculated as $\text{MAE} = \Sigma|\Delta\delta|/n$; maximum error was defined as the largest $|\Delta\delta|$. The carboxylate C1 atom was excluded from the calculations, unless otherwise noted.

4.3 Calculation of $^3J_{HH}$ NMR coupling constant

On the same geometry-optimized structures, proton-proton coupling constants were calculated using literature recommended procedures.$^{48}$ The B3LYP/6-31G** level of theory was used. The core basis functions for hydrogen atoms were uncontracted and augmented. Only the contribution of Fermi-contact term to the nuclear spin couplings was calculated. The resulting coupling constants were empirically scaled by 0.9155.

Acknowledgments

The author is grateful to the Chemistry Department of Yale University for its graduate training program.

Supplementary Material

Additional data tables, figures, coordinates for optimized structures, and sample Gaussian input file are available.
References

### Table 1. Evaluating the calculated $^{13}$C chemical shifts for the four candidate isomers of glabramycin C against experimental values. MAE (mean absolute error) and ME (maximum error) values were calculated after linear regression corrections.

*Results are tabulated for the following computational methodologies:
1. Geometry: HF/ 3-21G, vacuum; NMR: GIAO mPW1PW91/6-31G**, vacuum
2. Geometry: mPW1PW91/6-31G**, vacuum; NMR: GIAO mPW1PW91/6-31G**, vacuum
3. Geometry: mPW1PW91/6-31G**, vacuum; NMR: GIAO mPW1PW91/6-31G**, CPCM CHCl$_3$
5. Geometry: B3LYP/ 6-31G**, IEF-PCM CHCl$_3$; NMR: GIAO mPW1PW91/6-31G**, IEF-PCM CHCl$_3$

*Mean absolute error in ppm.
"Maximum error in ppm.
*Not calculated.

<table>
<thead>
<tr>
<th>Calc. Method</th>
<th>Deviations from experimental values</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>ME</td>
<td>MAE</td>
<td>ME</td>
<td>MAE</td>
</tr>
<tr>
<td>1</td>
<td>4.0</td>
<td>12.3</td>
<td>3.6</td>
<td>10.9</td>
</tr>
<tr>
<td>2</td>
<td>3.6</td>
<td>9.3</td>
<td>3.1</td>
<td>8.4</td>
</tr>
<tr>
<td>3</td>
<td>3.6</td>
<td>10.4</td>
<td>3.1</td>
<td>10.8</td>
</tr>
<tr>
<td>4</td>
<td>4.1</td>
<td>10.2</td>
<td>3.6</td>
<td>9.5</td>
</tr>
<tr>
<td>5</td>
<td>3.6</td>
<td>9.2</td>
<td>3.1</td>
<td>9.6</td>
</tr>
</tbody>
</table>

### Table 2. Experimental and calculated $^1$H values (Hz) for the four stereochemical models of Glabramycin C.

*Data extracted from isolation report.

\[ J_{\text{HH}} \] \(^{13}C\) values (Hz) for the four stereochemical models of Glabramycin C.

*Calculated values for candidate models

<table>
<thead>
<tr>
<th></th>
<th>Exp. values</th>
<th>Calculated values for candidate models</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_{10,11}$</td>
<td>8.5</td>
<td>14.9</td>
</tr>
<tr>
<td>$\phi_{10,11}$</td>
<td>-</td>
<td>170°</td>
</tr>
<tr>
<td>$J_{11,12}$</td>
<td>4.5</td>
<td>2.4</td>
</tr>
<tr>
<td>$\phi_{11,12}$</td>
<td>-</td>
<td>94°</td>
</tr>
</tbody>
</table>
Figure 1
143x53mm (300 x 300 DPI)
(11R,20R)
Mean Abs Err = 4.0 ppm
Max Err = 11.8 ppm

(11R,20S)
Mean Abs Err = 3.4 ppm
Max Err = 10.3 ppm

(11S,20R)
Mean Abs Err = 2.1 ppm
Max Err = 7.1 ppm

(11S,20S)
Mean Abs Err = 2.0 ppm
Max Err = 7.3 ppm

R = \text{carbon atom}

254x190mm (300 x 300 DPI)