

Claremont Colleges

## Scholarship @ Claremont

---

CMC Senior Theses

CMC Student Scholarship


---

2021

# Using NMR Spectroscopy and Computational Chemistry to Confirm the Structure of Novel Antibiotic Nocamycin O

Stephanie Lewis

Follow this and additional works at: [https://scholarship.claremont.edu/cmc\\_theses](https://scholarship.claremont.edu/cmc_theses)

 Part of the [Analytical Chemistry Commons](#), [Biochemistry Commons](#), [Medicinal-Pharmaceutical Chemistry Commons](#), and the [Molecular Biology Commons](#)

---

### Recommended Citation

Lewis, Stephanie, "Using NMR Spectroscopy and Computational Chemistry to Confirm the Structure of Novel Antibiotic Nocamycin O" (2021). *CMC Senior Theses*. 2592.  
[https://scholarship.claremont.edu/cmc\\_theses/2592](https://scholarship.claremont.edu/cmc_theses/2592)

This Open Access Senior Thesis is brought to you by Scholarship@Claremont. It has been accepted for inclusion in this collection by an authorized administrator. For more information, please contact [scholarship@cuc.claremont.edu](mailto:scholarship@cuc.claremont.edu).

Using NMR Spectroscopy and Computational Chemistry to Confirm the Structure of  
Novel Antibiotic Nocamycin O

A Thesis Presented

By

Stephanie Lewis

To the W. M. Keck Science Department  
Of Claremont McKenna, Pitzer, and Scripps Colleges

In partial fulfillment of  
The degree of Bachelor of Arts

Senior Thesis in Molecular Biology

November 23<sup>rd</sup>, 2020

## Table of Contents

<b>Abstract</b>	<b>3</b>
<b>Introduction</b>	<b>4</b>
Novel Antibiotic Discovery: Methods	4
Nocamycin O as a Potential Antibiotic for Commercial Use	5
NMR Spectroscopy and its Theoretical Supplements	12
NMR Theory and the Importance of Chemical Shift Calculations in Structure Determination	13
NMR in Practice	18
<b>Methods</b>	<b>22</b>
NMR Data Acquisition in MeOD-d4	22
NMR Data Acquisition in DMSO-d6	23
Chemical Shift Calculations	24
<b>Results</b>	<b>25</b>
NMR in MeOD and dDMSO - Experimental Results	25
Confirmation of Experimental Chemical Shift Assignments - DFT in Spartan '18	30
Residual Analysis of Density Functional Calculations	32
Confirmation of Novelty - Comparison of Nocamycin I and Nocamycin O	33
<b>Discussion</b>	<b>37</b>
Structure Confirmation of Nocamycin O	37
DFT Calculations	38
Nocamycin O as a Novel Member of the Nocamycin Family	40
<b>Conclusion</b>	<b>42</b>
<b>Acknowledgements</b>	<b>43</b>
<b>References</b>	<b>44</b>
<b>Supplementary Figures</b>	<b>48</b>
Section 1: MeOD NMR Data	48
Section 2: dDMSO NMR Data	54

## Abstract

In recent years, many medically promising antibiotics have been discovered in nature, especially in insect-microbe symbioses. One of the better-studied examples of this kind of defensive relationship is that of fungus-growing ants and the antibiotic-producing Actinobacteria. These bacteria produce several defensive chemicals with myriad uses, including one antibiotic that inhibits the growth of several bacterial strains, including other Actinobacteria. This antibiotic (known as nocamycin O) is a promising candidate for medicinal use due to its similarities to bacterial RNA polymerase inhibitors tirandamycin and streptolydigin, which inhibit several human pathogens. The determination of the structure of nocamycin O will be an important first step toward determining its function and its potential utility in the medical field. This can be done efficiently and accurately using nuclear magnetic resonance spectroscopy (NMR). NMR can be used on its own to attempt to solve the structure of a compound, or in tandem with virtual chemical shift calculations that act as a check to correct the experimentally-derived structure. Overall, NMR and chemical shift calculations have become integral components to biochemical and biomedical research because they make structure elucidation much easier. My research sought to confirm the structure of nocamycin O using prior NMR data for the compound, as well as novel 2D NMR data collected in MeOD and DMSO with complementary  $^{13}\text{C}$ -NMR spectrum calculations performed using DFT in Spartan '18. Comparative analysis of NMR spectra for nocamycin O and nocamycin I revealed key differences in chemical shift values; the carbon with the additional -OH in nocamycin O experienced a shift change of almost 40 ppm, while other carbons in the molecule showed a change of 5-10 ppm. These changes were likely due to a difference in nuclear environment at these positions, which was confirmed via the DFT calculations and ROESY spectrum.

## Introduction

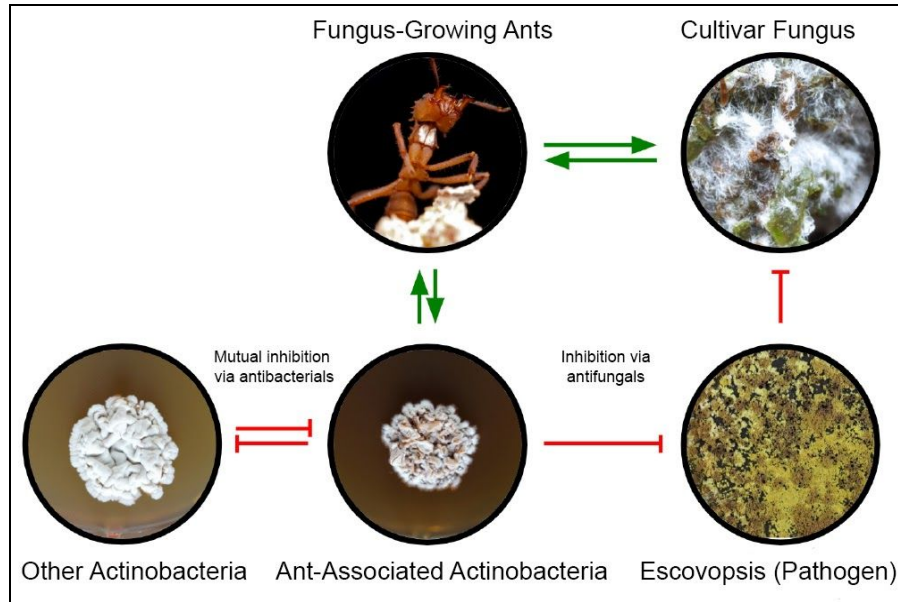
### *Novel Antibiotic Discovery: Methods*

Over the last few decades, more and more highly antibiotic-resistant pathogenic bacteria have been discovered. This necessitates the introduction and use in medicine of new antibiotics that bacteria have not yet evolved to resist. However, the rate of approval of new antibiotics has not kept pace with need and has decreased dramatically in the last twenty years.<sup>1</sup> The most common traditional methods of antibiotic discovery are discovery-based chemical screening (screening small molecules for a certain desired activity without prior knowledge of their functions), and target-oriented screening (focused on identifying compounds that attack a known target), but these methods have grown ineffective.<sup>2</sup> The slowdown may be due to the rediscovery of many already-known substances, as it appears that certain classes of molecules are more likely to be discovered in these types of screening campaigns. To combat that, combinatorial chemistry was created, which involves *in silico* screening of fragments and experimental screening of fragment libraries followed by linkage of promising fragments in hope of creating a functional and useful drug.<sup>3</sup> However, though many new chemical compounds have been discovered, few have proven at all useful in terms of antibiotic or antifungal activity. This is because the combinatorial chemistry methods appear to only cover a limited amount of the chemical diversity represented in most chemical libraries.<sup>2</sup> Most drugs currently in use, as well as those produced in nature, show a much more diverse representation of chemical types that have more diverse and potent functions. The discovery of more of these diverse and well-functioning compounds could be achieved through natural products discovery, but not in the traditional sense. Instead of focusing on high-throughput screening or chimerically stitching two functional parts together, the attention should be on finding more naturally-occurring compounds, in

particular from plants and microbes and especially from defensive symbioses. In recent years, this type of natural products research has been the most productive in terms of new drug discoveries to be tested. There are currently over 100 new naturally-derived drugs in clinical development, mainly as cancer treatments and antibiotics.<sup>1</sup>

#### *Nocamycin O as a Potential Antibiotic for Commercial Use*

In terms of novel antibiotic discovery, insect-microbe symbioses have shown much promise in recent years. These types of symbioses are common in many different types of animals and can come in two forms. Nutritional symbioses imply that the bacteria either produces nutrients the host needs or aids with digestion, while defensive symbiosis usually means that the bacteria produces a chemical that the host weaponizes for its own defense.<sup>4-7</sup> One of the more promising and better-studied examples of this kind of defensive relationship is that of fungus-growing ants and the antibiotic-producing Actinobacteria. The bacteria live off of secretions produced by the ant, and in return, they manufacture a wide variety of molecules with demonstrated antibacterial properties.<sup>7,8</sup> The presence of *Actinobacteria* has been shown to correlate with the efficacy of the ants' efforts to suppress attacks by the pathogen *Escovopsis*, which feeds upon the cultivar fungus that is the ants' main food source (Figure 1). Antifungal molecules were also isolated from strains of *Streptomyces* found in similar fungus gardens on the same types of ants; their general function is suppression of nest pathogens, and they represent an exciting new area of study.<sup>7,8</sup>



**Figure 1:** Graphic representation of the symbiotic relationship between fungus-growing ants and some strains of Actinobacteria, adapted from Kim et. al. (2019).

A particularly interesting facet of this symbiosis is that most strains of Actinobacteria show strong antagonistic effects against other Actinobacteria strains isolated from different ants. This is likely because when they coexist in the same environment, Actinobacteria are in direct competition with each other to maintain their favorable position in the symbiosis. This is well-illustrated in the response of *Trachymyrmex septentrionalis*-associated Actinobacteria to other types of Actinobacteria.<sup>9</sup> Most of the bacterial isolates coming from *T. septentrionalis* show weak or no inhibition of other types of Actinobacteria also associated with other *T. septentrionalis* ants and extremely strong inhibition of the other Actinobacteria 17SM-1 and 18AZ-4, which come from *T. smithi* and *T. arizonensis* ants respectively.<sup>9</sup> All of the strains that showed any inhibition at all were able to efficiently inhibit the growth of other types of Actinobacteria and many Gram-positive bacteria including the human pathogens *Staphylococcus aureus* and *Enterococcus faecalis*.<sup>9</sup> The fact that the paradigm of this particular defensive symbiosis parallels human medicine (both use antagonistic molecules like antibiotics to suppress

pathogens) and that molecules associated with this paradigm have been shown to inhibit growth of human pathogens is exciting because it represents an ecologically-guided pathway to discovering new antibiotics with a higher probability of success in humans, as they are already compatible with animal hosts.<sup>7</sup>

One such antibiotic is nocamycin O, an analog of nocamycin I (also known as BU-2313B). It was isolated from a bacterium in the genus *Amycolatopsis*, a symbiont of *Trachymyrmex smithi* that is another type of Actinobacteria. Four different isolates of *Amycolatopsis* were tested by Rose Kim (a previous researcher in the Van Arnam Lab) for potential antibiotic activity; the most inhibitory isolate, 17SM-2, was selected for further study. This strain was subjected to resident-intruder assays to determine whether the compound it produced had antifungal or antibacterial activity (Figure 2):

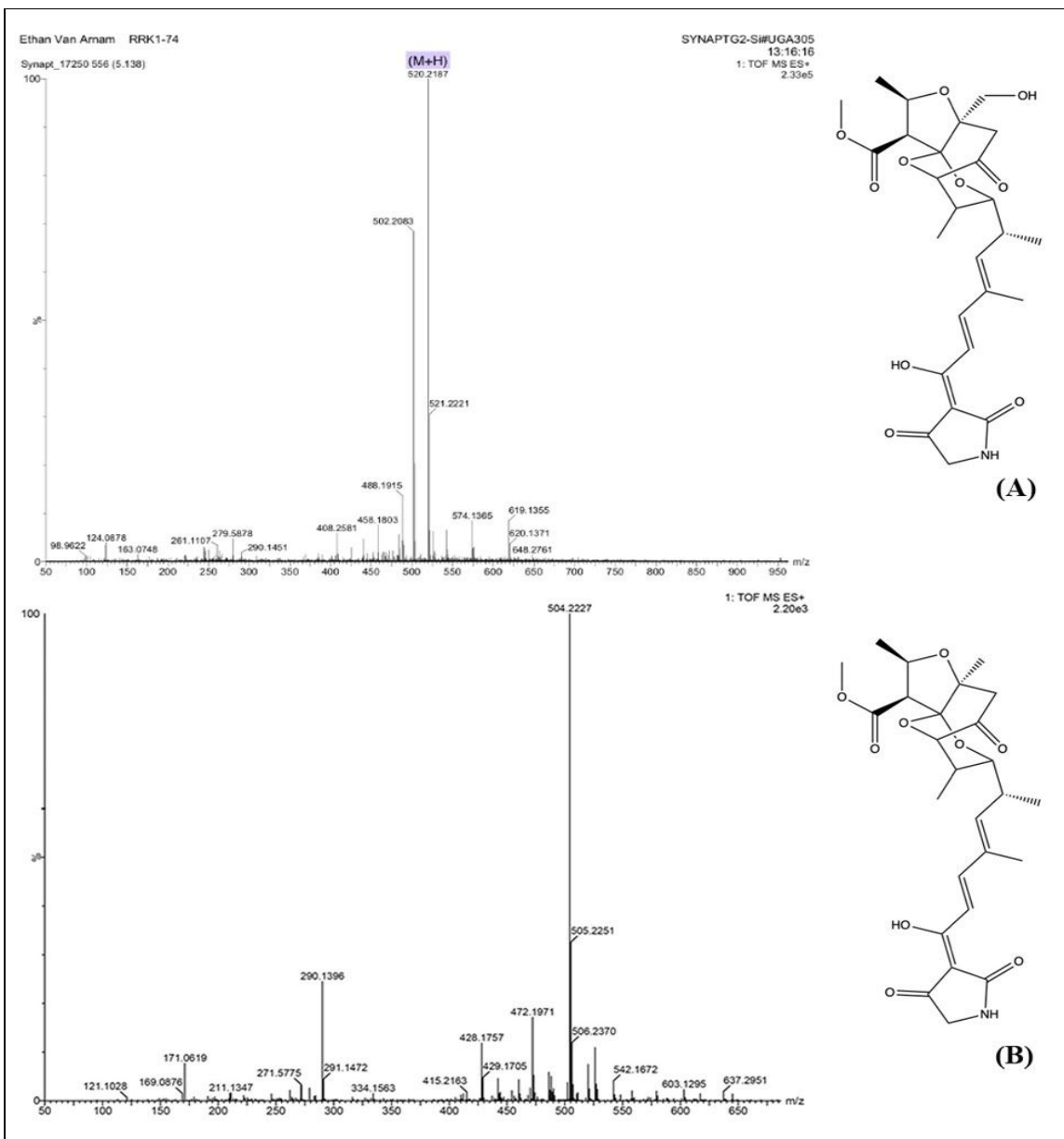


**Figure 2:** Resident-intruder assays to evaluate antifungal (left) and antibacterial (right) activity in 17SM-2, from Kim et. al. (2019).

The antibacterial compound was extracted from the bacteria and then fractionated, first using C18 chromatography and then high-performance liquid chromatography (HPLC). The most active fraction was then subjected to liquid chromatography-high resolution mass spectrometry (LC-HRMS). This analysis showed a (M+H)<sup>+</sup> peak of 520.2187, which gives a molecular mass of 519.2133 +/- 0.0051. Additionally, no matches were found in the Dictionary of Natural Products, indicating that this antibiotic is possibly a novel find.<sup>11</sup> Similarities between nocamycin

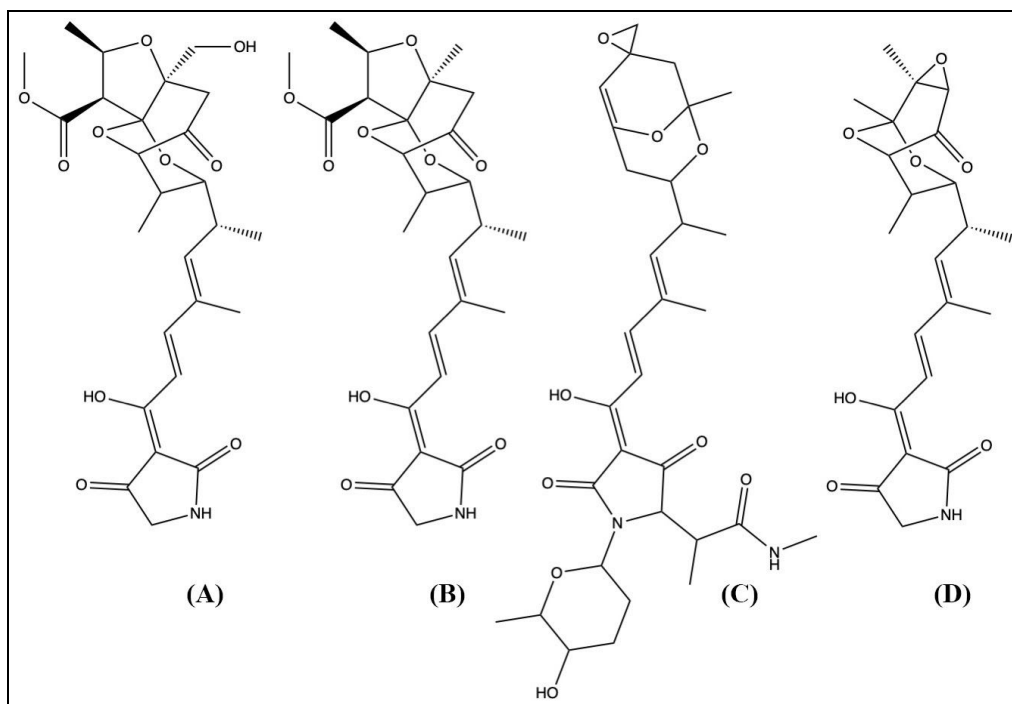


I data and preliminary MS data for nocamycin O support the idea that it belongs to the nocamycin family (Figure 3). The strongest support for this idea, however, comes from biosynthetic gene cluster analysis; >75% of the genes encoding nocamycin I are very similar to genes for nocamycin O, and all genes except one have an analog in the other genome. One- and two-dimensional NMR (nuclear magnetic resonance) experiments were used to elucidate a tentative structure.



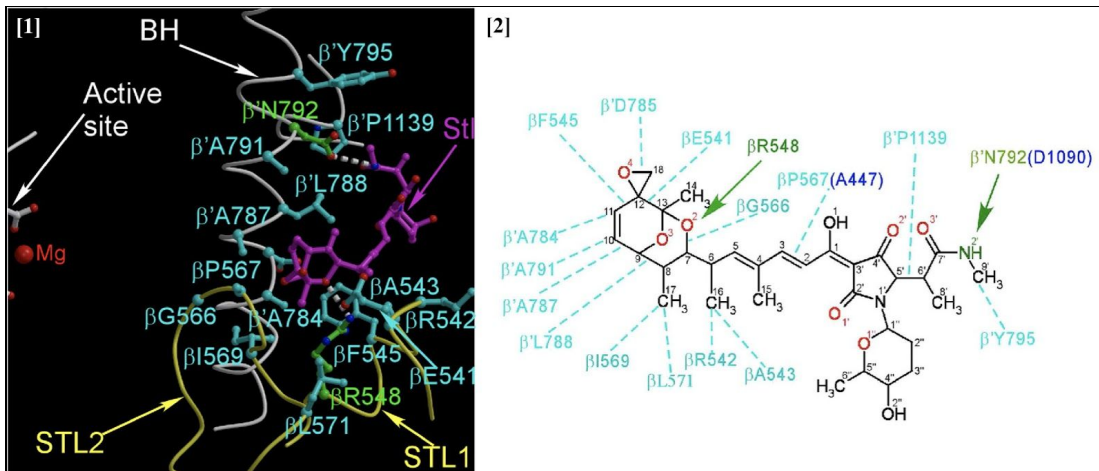
**Figure 3:** Mass spectrometry results with tentative structure for (A) nocamycin O and (B) nocamycin I. nocamycin I data from Cogan et. al. (2020).

Since the mechanisms of nocamycin I and O inhibition have not been well studied, research on similar molecules was conducted to gain a better idea of ways in which these molecules could potentially inhibit bacterial growth. Chemically similar (and better characterized) relatives include tirandamycin and streptolydigin, also part of the tetramic acid family. It has been well-documented that streptolydigin inhibits bacterial growth by interfering with bacterial RNA polymerase by binding to its St1 pocket, bridge helix, and trigger-loop regions; part of its structure is close enough to contact the backbone of the DNA nontemplate strand, allowing streptolydigin to interfere with RNAP function through direct interactions with the DNA.<sup>12</sup> This is relevant because while streptolydigin has a similar structure to the nocamycin family, it is different enough that the binding affinity and properties may be much different. Nocamycins I and II are unique from the rest of the family in that they have a fused tricyclic ring system, while others such as streptolydigin and tirandamycin have a bicyclic system (Figure 4).



**Figure 4:** Structure comparison of nocamycin O (A), nocamycin I (B), streptolydigin (C), and tirandamycin (D).

Streptolydigin is a known inhibitor of bacterial RNA polymerase in a number of Gram-positive bacteria. Nocamycin I (the closest chemical relative to nocamycin O) also demonstrates broad antimicrobial activity against many Gram-positive and a few Gram-negative bacteria.<sup>13-15</sup> These bacteria include the two anaerobes *Bacteroides fragilis* and *Propionibacterium acnes* and the aerobe *Streptococcus pyogenes*, a known cause of infections such as tonsillitis, scarlet fever, cellulitis, and necrotizing fasciitis in humans.<sup>13,14</sup> Additionally, it was noted that increasing the size or chain length of the 5' substituent (a synthetic addition to the the CH<sub>2</sub> carbon in the tetramic acid ring) led to a decrease in efficacy of the antibiotic, with the most effective synthetic nocamycin I analog having a methyl substituent in the R<sub>1</sub> position; it was still not as effective as the original compound.<sup>14</sup> This has implications for nocamycin activity in comparison to other tetramic acid antibiotics like streptolydigin; most of those implications have to do with structure and how the antibiotic associates with its target, bacterial RNA polymerase. Though it is currently unknown how or whether nocamycin interacts with RNA polymerase to inhibit bacterial growth, based upon its structural similarities to streptolydigin, our knowledge of how streptolydigin interacts with RNA polymerase, and the similar activity profiles between the two (both inhibit bacteria in the *Streptococcus* family and the *Bacillus* family, as well as *Neisseria meningitidis*), it is likely that the nocamycin family associates with RNA polymerase in a similar fashion.<sup>12,16</sup>



**Figure 5:** [1] streptolydigin (in purple) associating with bacterial RNAP and [2] all of its van der Waals interactions (blue dashes) and hydrogen bonds (green arrows) with RNAP, from Temiakov et. al. (2005).

It has been suggested that the streptolol portion (the portion containing the bicyclic ring system) of streptolydigin is the main determinant of its affinity for RNA polymerase; the tetramic acid portion of the molecule has no specific binding sites on the protein surface and thus cannot contribute to its interactions with RNA polymerase, and may even weaken them.<sup>15</sup> The sugar moiety (the six-membered ring containing carbons 1''-5'' in Figure 4) also has no contact with RNA polymerase; it may even hinder streptolydigin binding by competing for positions with other parts of RNA polymerase.<sup>15</sup> Nocamycin O lacks both a bulky tetramic acid portion and a sugar moiety, as it contains only a five-membered nitrogenous ring in that region (Figure 3). Assuming that nocamycin O maintains a similar conformation/orientation toward RNA polymerase as streptolydigin upon binding with it, nocamycin O could potentially inhibit RNA polymerase more efficiently than streptolydigin. The hydrogen bond with  $\beta$ R548 can be maintained, as that part of nocamycin O is structurally identical to streptolydigin, and the hydrogen bond lost by the replacement of the acetamide could easily be filled by the nitrogen in the tetramic acid ring of nocamycin O. In addition, nocamycin O also contains an -OH group on

one methyl substituent of the tricyclic system, which opens up the possibility of additional/previously unobserved noncovalent interactions with the hydrophilic/polar amino acids in that region of RNA polymerase (Figure 3). The combination of the less bulky tetramic acid group and the additional hydrogen bonding opportunity implies that nocamycin O could be an even stronger inhibitor of RNA polymerase than either nocamycin I or streptolydigin, and that it has potential as an antibiotic against human pathogens as well. Additionally, there is already a precedent for the addition of an -OH group at this position, as tirandamycin B, another tetramic acid antibiotic, has this addition.<sup>17</sup>

As shown by the importance of chemical structure in the functioning of streptolydigin, proper elucidation of chemical structure is an important first step in determining how a molecule functions *in vivo*. Thus, proper determination of the structure of nocamycin O will be an integral and enlightening step toward the discovery of its relationship with RNA polymerase and its binding properties, as well as an integral component to its establishment as a novel molecule. One method that has been briefly mentioned was selected for this task: NMR spectroscopy. The utility of NMR spectroscopy and related theoretical models in terms of structure determination will be discussed below.

### *NMR Spectroscopy and its Theoretical Supplements*

NMR spectroscopy is useful for several reasons; it is a non-destructive technique (the molecule does not decay after analysis), it has a wide range of applications in almost every branch of science, and it has hundreds of different types of experiments meant to reveal different aspects of the molecular structure.<sup>18</sup> In the most basic sense, NMR involves exciting the nuclei of the molecule in a specific way by generating a series of radiofrequency and/or magnetic gradient

pulses followed by signal acquisition to gather the desired information about the molecule. This information can then be compiled into a readable format from which a molecular structure can be obtained.

In an ideal situation, issues caused by suboptimal experimental environments and background noise in NMR acquisition would be minimal, but in practice, the spectra obtained depend a great deal on external factors, many of which are mechanical- or solvent-related. Common causes of inaccuracy in chemical shift assignments include distorted or hidden peaks caused by misplacement of the magnets (incorrect shimming) or poor signal-to-noise ratio, insufficient acquisition time, and most importantly, broad peaks caused by decreased sensitivity of the experiment due to insufficient sample volume.<sup>18</sup> It is clear that spectral complexity impedes proper structural determination in experimental NMR acquisition, and with larger molecules, this issue is compounded due to the possibility of peaks appearing on top of each other, a phenomenon known as spectral crowding. However, it should be noted that many of these issues can be addressed by simply changing the modality or phase state of the experiment. Many problems associated with 1-dimensional solution-state NMR can be fixed either by transitioning to solid-state NMR or by performing multiple 2-dimensional experiments, as these will show where couplings occur on the crowded <sup>1</sup>H-NMR spectrum. In addition, theoretical calculations can be used to help confirm shift assignments by showing what the shift values for each nucleus in a structure should be under various solvent conditions.

### *NMR Theory and the Importance of Chemical Shift Calculations in Structure Determination*

Presently, computational chemistry software packages are becoming more widely known as valid alternatives and complements to the experimental methods used to perform certain

experiments and calculations. These include energetics calculations, molecular structure optimizations, and calculation of the values of various molecular characteristics, including their IR (infrared) and NMR spectra. With theoretical calculations, the issues of spectral complexity and suboptimal experimental conditions are eliminated. Additionally, NMR chemical shift calculations are an excellent way to confirm the structure proposed by the experimental data. NMR spectra are created based on the specific behavior of each NMR-active nucleus in the molecule in response to an external magnetic field, in this case, a large magnet.<sup>19</sup> The central question that NMR helps answer is the determination of chemical structure; assigning shift values to each atom in the molecule can help confirm that structure. The basic principle underlying NMR and all theoretical models based upon it is that the frequency at which each atom's spin aligns with the external magnetic field is directly related to its chemical environment. The frequency at which each atom resonates can be altered by a change to its environment, and depending on what changes, can be positive or negative.<sup>19,20</sup> Shielding is caused by the electron cloud surrounding each atom; when the external magnetic field is applied, the field it induces directly opposes it, which reduces the strength of the external magnetic field in that location.<sup>20</sup> Being bonded to more electronegative atoms causes electron density to be siphoned away from that specific nucleus, resulting in less of a shielding effect and an increase in the shift value.<sup>21</sup> The shift value of each nucleus is a value in ppm (a standardized unit, parts per million, that allows for values taken on different spectrometers to be the same) which is usually between 0 and 14 for protons, and directly depends on the frequency (in hertz) that causes the atom to align with the external field; more shielded atoms will respond to lower frequencies and thus have a lower ppm value.<sup>20</sup>

An even smaller deviation in the magnetic field can be observed due to the orientation of other nearby atoms of the same type. Essentially, while the external magnetic field affects the resonant frequency of each nucleus, so do the induced magnetic fields of the nearby atoms.<sup>19,20</sup> Each of these factors depend on the electron density around each atom as well as the way these electrons affect the local magnetic environment, two things that empirical methods (which assign chemical shifts based on prior experimentally-determined values for chemically-similar nuclei) fail to account for because they generally ignore electronic contributions. Thus, quantum chemical methods such as density functional theory, discussed later in this section, become necessary in order to account for these small but important indirect interactions in the overall energy of the molecule.

Empirical methods for calculating shifts involve extrapolation of experimental knowledge; data are usually taken from chemical shift tables for each unique molecular environment, adjusted based on electronegativity trends, and averaged over the different conformers of the molecule to give each atom a unique value in the NMR spectrum.<sup>19</sup> Other methods involve a mass search of molecule databases to find molecular environments which are chemically similar to the atom currently being analyzed; the advantage of these methods is that each can be performed and completed quickly. However, these empirically-based models have drawbacks; when the differences in chemical shift are too small, the program cannot reliably distinguish between individual magnetic signals, making this type of calculation less useful for large molecules like biopolymers and antibiotics. Quantum-mechanical calculations, the primary calculation format found in this thesis, use an entirely different approach to predict the chemical shifts of the target molecule than empirical methods; this source is primarily the electronic structure of the molecule.<sup>19</sup> While empirical methods use previously acquired data and



chemically similar matches to construct an NMR spectrum, quantum mechanical calculations use the electronic structure to determine the magnetic shielding and spin-spin coupling constants, interactions which are determined solely by the electronic structure of the molecule (Figure 6).

$$\begin{aligned}
 E(\mathbf{B}, \boldsymbol{\mu}) &= E_0 + \left\{ E^{(10)}\mathbf{B} + \sum_i E^{(01)}\boldsymbol{\mu}_i \right\} \\
 &+ \frac{1}{2}(\mathbf{B}^T E^{(20)}\mathbf{B}) + \sum_i \frac{1}{2}(\mathbf{B}^T E_i^{(11)}\boldsymbol{\mu}_i) \\
 &+ \frac{1}{2} \sum_i \sum_{i \neq j} \boldsymbol{\mu}_i^T E_{ij}^{(02)}\boldsymbol{\mu}_j + \dots, \\
 E_i^{(11)} &= d^2 E(\mathbf{B}, \boldsymbol{\mu}) / d\mathbf{B} d\boldsymbol{\mu}_i \quad \text{for } \mathbf{B} = 0, \boldsymbol{\mu} = 0 \quad [11] \\
 E_{ij}^{(02)} &= d^2 E(\mathbf{B}, \boldsymbol{\mu}) / d\boldsymbol{\mu}_i d\boldsymbol{\mu}_j \quad \text{for } \mathbf{B} = 0, \boldsymbol{\mu} = 0 \\
 & \hspace{15em} [12]
 \end{aligned}$$

**Figure 6:** Taylor series describing the energy of the molecule in terms of the electronic structure including magnetic shielding ( $E_i^{(11)}$ ) and spin-spin coupling ( $E_{ij}^{(02)}$ ), from Gryff-Keller (2011).

In Figure 6,  $\mathbf{B}$  and  $\boldsymbol{\mu}$  are quantities describing the magnetic field and nuclear magnetic moments, two quantities which are necessary for calculating the magnetic shielding and coupling constants. These are integral for the accurate calculation of the shift value for each nucleus due to the utilization of the electronic structure as outlined above. Therefore, these quantities cannot be measured using simple analytical techniques; due to the complexity of the connection between these parameters and the electronic structure, connections cannot be made between spin-spin coupling (small deviations in the field due to signals produced by neighboring atoms), magnetic shielding, and other well-defined NMR parameters using empirical methods.<sup>19</sup> Though quantum mechanical calculations come with their own set of issues (first and foremost an

extremely slow processing rate), a quantum mechanical perspective is helpful for linking NMR parameters to molecular structure.

The major principles used in NMR shift calculations are based upon principles of quantum mechanics as applied to molecules, a subset of quantum mechanics known as quantum chemistry, whose theories underlie the programs used in computational chemistry. This involves the application of mathematical and theoretical principles to the solution of chemical problems and can be used to calculate values for chemical properties virtually.<sup>22</sup> Two broader areas of quantum chemistry are statistical mechanics and electronic structure theory; of these, the more relevant in terms of NMR spectra is electronic structure theory. Electronic structure theory utilizes the laws of quantum mechanics rather than classical mechanics to mathematically calculate the energy and related properties of molecules while taking into account the atoms' electrons, something that molecular mechanics and empirical methods do not do. The only form of electronic structure calculations that are entirely quantum theory-based are known as ab initio calculations; one of the most commonly used methods in general and the one most often used in terms of this thesis is density functional theory, or DFT.

Density functional theory (hereafter abbreviated as DFT) appears here to offer a compelling alternative to more computationally-taxing methods like Hartree-Fock, which involves the approximation of the Schrödinger equation via a complex multivariable wave function  $\Psi(x_1, x_2, \dots, x_N)$ .<sup>19,23</sup> It simplifies the solution to the Schrödinger equation by employing a simpler method to calculate the energies; in DFT, the  $N$ -electron wave function and the associated Schrodinger equation are replaced with the simpler electron density equation  $\rho(r)$  and its associated calculational scheme (Figure 7):

$$E_{\text{TF}}[\rho(r)] = C_{\text{F}} \int \rho^{5/3} dr - Z \int \frac{\rho(r)}{r} dr + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 \quad [1]$$

$$E_v[\rho] = T[\rho] + V_{ne}[\rho] + V_{ee}[\rho] = \int \rho(r)v(r) dr + F_{\text{HK}}[\rho] \quad [2]$$

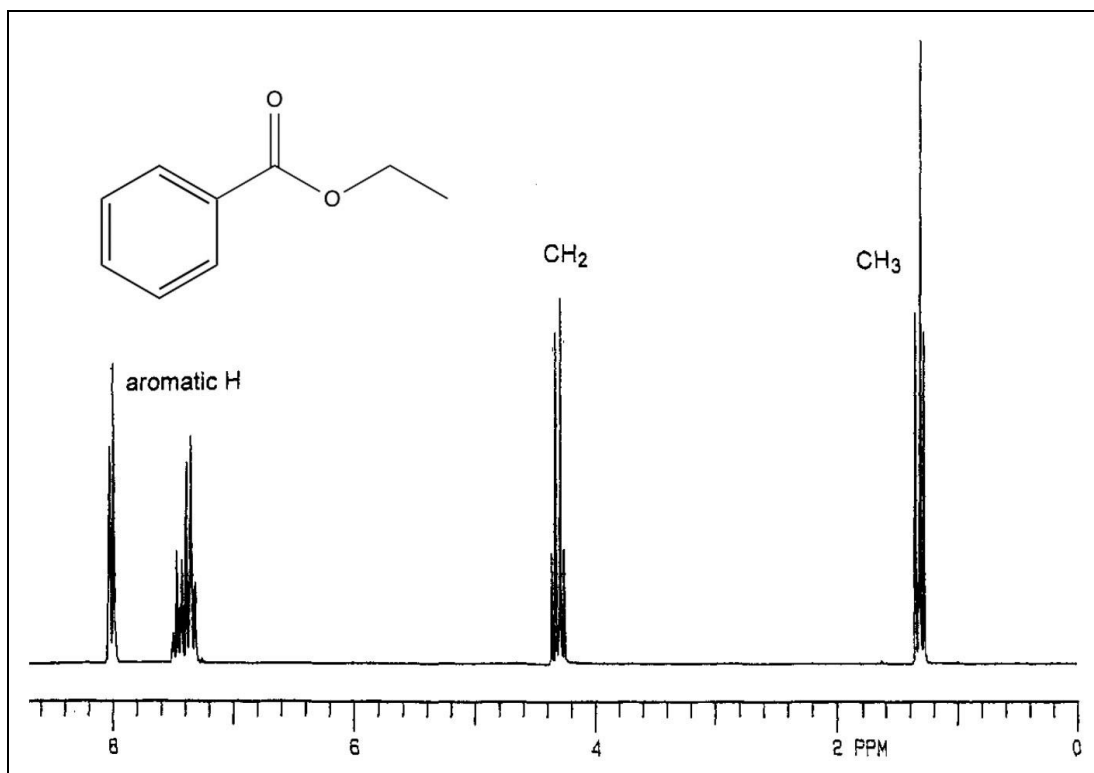
**Figure 7:** [1] representing the energy of an atom in terms of its electron density, and [2] the ground-state energy with respect to  $\rho$ , from Parr and Yang (1995).

Figure 7 shows an equation for the energy of an atom that depends only on the electron density; the Hohenberg-Kohn theorems modified that further to say that it determines all properties of the ground state, including the kinetic ( $T[\rho]$ ), potential ( $V[\rho]$ ), and total ( $E[\rho]$ ) energies.<sup>23</sup> It is important to note that it is impossible to have a true value of  $F_{\text{HK}}[\rho]$ , the universal functional (a function that takes another function as its input) of  $\rho(r)$ , which is notoriously difficult to explicitly calculate. However, approximations are fairly easy to do if one assumes that the value of  $F_{\text{HK}}[\rho]$  in a uniform electron gas translates well to real systems.<sup>24</sup> Following the calculation of  $E[\rho]$ , the energies with respect to the nuclear magnetic moment  $\mu$  and the external magnetic field  $\mathbf{B}$  (as seen in Figure 6) can be compared to the ground-state energy in the absence of these variables,  $E_0$ , and the same theoretical model (in this case, DFT) is then used to generate a simulated NMR spectrum. The calculated chemical shift values for the tentative structure can then be plotted against the experimentally determined shift values to make certain that the structure and shift value assignments are correct.

### *NMR in Practice*

In practice, NMR spectra are much easier to understand than the theoretical calculations required to simulate them. The principle behind NMR spectroscopy is very similar to that of the fluorescence of light; when an external magnetic field is applied, each NMR-active nucleus can align their spins either with or against the magnetic field. When the nuclei are in the xy-plane,

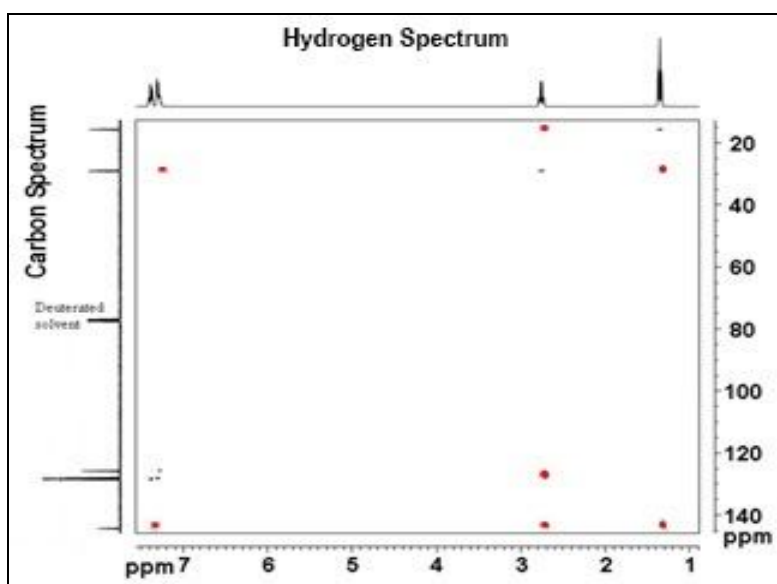
the signal will be visible. As the nucleus relaxes back to align with the z-axis, its signal decays; the different rates at which the nuclei relax and its immediate environment account for differences in signal intensity and chemical shift value, respectively. The surroundings for each atom play a role as well. For both  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR (the most often-analyzed nuclei in NMR), as more electron density is shifted away from the nucleus in question, the higher the chemical shift gets (Figure 8). Certain functional groups will consistently cause higher shift values, making NMR useful for identifying functional groups and determining molecular structure.



**Figure 8:** H-NMR spectrum of ethyl benzoate to illustrate chemical shift, from Balci (2005).

The splitting of the peaks seen in Figure 8 has to do with the spin orientation of neighboring nuclei, also known as spin-spin coupling. The pattern has to do with the number of different nuclei in the immediate vicinity of the nucleus; the  $\text{CH}_2$  will be split three times (into a quartet) because its only immediate hydrogen neighbor is the methyl group.<sup>25</sup> Once the spectral pattern of one type of nucleus is known, it can be used to assign another type of nucleus and

figure out structure and attachments through two-dimensional NMR, which is the correlation of shift values for one type of nucleus with the shift values of another type of nucleus. For example, certain two-dimensional experiments will align proton signals with the carbon signals in their surrounding environments, giving information about which proton is attached to which carbon. The most common and useful of these heteronuclear two-dimensional experiments are Heteronuclear Single Quantum Coherence (HSQC) and Heteronuclear Multiple Bond Coherence (HMBC).



**Figure 9:** Overlay of HSQC and HMBC spectrum of ethylbenzene, from Hoffman (2020). Signals present on the HMBC that are absent on the HSQC are highlighted in red.

HSQC is one of many short-range heteronuclear correlation methods. It is especially useful because it illustrates exactly which hydrogen is attached to which carbon. All correlation signals not highlighted in Figure 9 make up the HSQC spectrum, where each hydrogen signal on the horizontal axis has only one correlation to the carbon spectrum on the vertical axis because hydrogen can only make one bond. Using this, one can assign hydrogens to their respective carbons. HMBC is, as its name suggests, indicative of the carbon associations up to 4 bonds away from each hydrogen atom, and can be used to illustrate the chemical environment of each

hydrogen.<sup>30,31</sup> As seen in Figure 9, the hydrogens of the methyl group in ethylbenzene are coupled to both the carbon in the benzylic position and the first carbon in the aromatic ring, and this is especially useful because that carbon lacks hydrogens and so is invisible on the HSQC. Using the HSQC and HMBC spectra in conjunction with the <sup>1</sup>H spectrum and its splitting is often enough to determine the connectivity of the atoms.

After the connectivity of the atoms is determined, additional spectra can be taken that are useful for determining aspects of the molecule's conformation and orientation in space, such as cis/trans alkene structure. The experiment used to detect these changes is known as a ROESY (rotating-frame Overhauser Effect spectroscopy) spectrum, which uses Nuclear Overhauser Effects to detect small changes in inter- and intramolecular interactions.<sup>28</sup> The Nuclear Overhauser Effect can be defined as the enhancement of a spin signal due to the dipole-dipole relaxation effect, or more simply, the spin of the excited nucleus relaxes because the magnetization is transferred over to the other coupled nucleus in the pair.<sup>28</sup> However, there is an amount of artifact that can appear on a ROESY spectrum from protons that are coupled to each other and are fairly similar in chemical shift due to transfer of the signal by proximity or proton exchange.<sup>29</sup> This can be remedied by also performing a TOCSY (Total Correlated Spectroscopy) experiment, which is somewhat of a long-range COSY coupling experiment. Any protons that show up there are then known to be artifacts on the ROESY.<sup>29</sup> This technique is desirable for discovering the intramolecular reactions as well as the conformation of medium-sized molecules such as nocamycin O.

In cases where spectra are difficult to read due to spectral crowding or a poor signal to noise ratio, more spectra may need to be taken in different solvents to fully understand the connections between atoms. In some cases, peaks may even disappear due to proton exchange

with the solvent, so it becomes quite challenging to know if the preliminary structure is correct. This is where the utility of chemical shift calculations using models like HF and DFT becomes apparent. If the structure put into the model is inaccurate, the calculated shifts will reflect that, and changes to the proposed structure can be made (and additional NMR experiments run to confirm that the changes are correct). As a molecule gets larger, NMR structures become harder to solve, and thus chemical shift calculations become more important. Given that most biomolecules are somewhat larger and more complex than synthetically generated compounds, both NMR and its theoretical counterpart are integral components in chemical and biochemical research.

In order to fully understand the relationship between RNA polymerase and nocamycin O, the molecular structure must be comprehensively elucidated. The definitive determination of the structure of nocamycin O will be integral in a) determining its novelty, b) allowing us to determine how this novelty impacts its interaction with its cellular target, and c) eventually conducting trials against a number of pathogens to determine its value in the medical and research fields. This study seeks to fully determine the nocamycin O structure using prior NMR data collected by students in the Van Arnam lab along with novel 2D NMR data collected between January and March of 2020, with complementary  $^{13}\text{C}$ -NMR chemical shift calculations performed using Hartree-Fock and density functional theory.

## **Methods**

### *NMR Data Acquisition in MeOD- $d_4$*

Prior to NMR acquisition, the purified nocamycin O was supplied by Elisabeth Lawton. The nocamycin O was evaporated to dryness in a SpeedVac, held under an additional vacuum of

approximately 100 millitorr overnight, and redissolved in 200 mL of deuterated methanol (MeOD-d<sub>4</sub>). All spectra were collected between the dates of February 13 and 28, 2020 on a 500 MHz Bruker BioSpin GmbH NMR instrument. Additionally, all spectra were collected between 294.6 and 298.1 K with a 5 mm TXI 1H/ D-13C/ 15N Z-GRD Z8161/ 0208 probe setting other than the initial 1D carbon spectrum, which was collected by Elisabeth Lawton on October 7, 2019 with a 5 mm PABBO BB/ 19F-1H/ D Z-GRD Z109128/ 0109 probe setting and zgig30 pulse sequence. The second spectrum collected was a 1D proton spectrum, gathered using pulse sequence zg30 with 32 scans. Following the initial proton spectrum, 5 other experiments were performed: 2D COSY (pulse sequence cosygpppqf, 8 scans), 2D HSQC (pulse sequence hsqcedetgppsp.3, 32 scans), 2D H2BC (pulse sequence h2bcetgpl3, 22 scans), 2D HMBC (pulse sequence hmbcetgpl3nd, 62 scans), and 2D ROESY (pulse sequence roesyphpp.2, 14 scans, relaxation delay 2.0 seconds). Following acquisition, all spectra were imported to and analyzed using MestReNova v12.0.3-21384 software. All heteronuclear 2D experiments were analyzed using the previously collected <sup>13</sup>C spectrum as a reference (which was referenced to the solvent, MeOD-d<sub>4</sub>), and the proton dimensions of all experiments were also referenced to the MeOD hydrogen peak (a quintet centered at 3.31 ppm).

#### *NMR Data Acquisition in DMSO-d<sub>6</sub>*

The previously used sample (MeOD-d<sub>4</sub>) was evaporated to dryness in a SpeedVac, redissolved in 1 mL of regular methanol (MeOH), and allowed to sit at room temperature for 16 hours so that proton exchange could occur. The sample was then re-evaporated to dryness on the



SpeedVac, further held under vacuum at approximately 100 millitorr for 15 hours, and finally redissolved in 200 mL of DMSO. Additional NMR data was collected on the 500 MHz Bruker BioSpin GmbH with probe setting 5 mm TXI 1H/ D-13C/ 15N Z-GRD Z8161/ 0208. Temperature was not controlled for these experiments. The first spectrum collected was a 1D proton spectrum (pulse sequence zg30, 32 scans), followed by a 2D HSQC (pulse sequence hsqcedetgppsp.3, 32 scans), 2D HMBC (pulse sequence hmbcetgpl3nd, 64 scans), 2D COSY (pulse sequence cosygpmfppqf, 32 scans), 2D H2BC (pulse sequence h2bcetgpl3, 96 scans), 2D ROESY (pulse sequence roesyphpr.2, 32 scans, relaxation delay 2.0 seconds), and a 2D TOCSY (pulse sequence mlevphpr.2, 32 scans). An attempt to capture a 1D carbon spectrum was unsuccessful due to low probe sensitivity, so all spectra were referenced using the successful 1D proton spectrum (which was referenced to the DMSO peak at 2.50 ppm) and the solvent peak present in the carbon spectrum (the only definitive peak present at 39.50 ppm). All spectra were analyzed in MestReNova v12.0.3-21384 and shift values were assigned.

### *Chemical Shift Calculations*

Following initial NMR data acquisition, chemical shift calculations for the proposed structure of nocamycin O as well as for the chemically similar compound nocamycin I were generated. This was done using Spartan '18 molecular modeling software and precisely followed a procedure outlined by Hehre et. al. as summarized below.<sup>30</sup>

The first step was performed using a systematic or sparse-systematic conformational search using the MMFF (Merck Molecular Force Field) molecular mechanics model to remove

any extremely high-energy conformers present in the initial search; all conformers with energies more than 40 kJ/mol above the lowest-energy conformer were discarded. In step 2, the Hartree-Fock method was utilized along with the 3-21G basis set to calculate equilibrium geometries for each conformer (between 20 and 25 are typically left at this stage). All conformers outside of a 40 kJ/mol energy threshold were again eliminated, as well as any duplicate conformers. In step 3, the  $\omega$ B97X-D density functional model was used with the 6-31G\* basis set to recalculate energies with a more accurate molecular model, and all conformers with a relative energy greater than 15 kJ/mol were discarded.

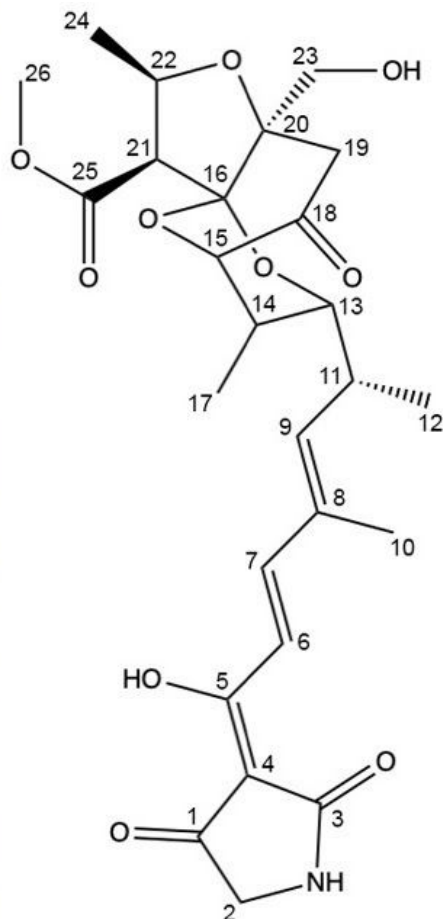
Steps 4 and 5 endeavored to repeat steps 2 and 3 under stricter molecular models and with more accurate basis sets. In step 4, the  $\omega$ B97X-D density functional model was used with the 6-31G\* basis set to recalculate equilibrium geometries for the remaining conformers, eliminating those with a relative energy greater than 10 kJ/mol. Step 5 recalculated conformer energies (usually around 8 or 9 conformers are left at this stage) with the  $\omega$ B97X-V density functional model and the 6-311+G(2df,2p) dual basis set. After step 5, both  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR spectra were generated for all conformers with a relative energy less than 10 kJ/mol and each shift value was averaged according to Boltzmann weight.

## Results

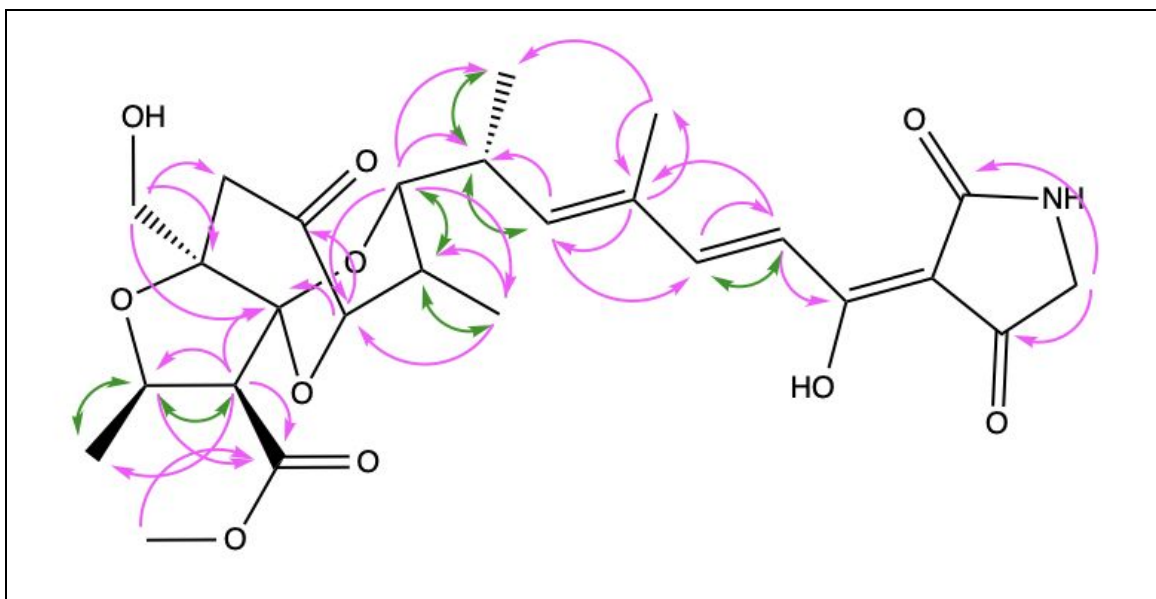
### *NMR in MeOD and dDMSO - Experimental Results*

The preliminary structure of nocamycin O was largely confirmed and supported by the data gathered from the initial NMR structure analysis in MeOD (Figure 10):

Position	$\delta$ C	$\delta$ H	Mult (J in Hz)
1	197.23		
2	51.07	3.61 d (12.1)	
3	179.29		
4	103		
5	184.36		
6	125.65	7.58 d (15.6)	
7	145.22	7.29 d (15.6)	
8	136.5		
9	139.13	5.87 d (9.9)	
10	12.66	1.88 d (1.2)	
11	35.51	2.87 ddd (9.2, 6.8, 2.1)	
12	17.05	1.01 d 6.9)	
13	81.18	3.58 dd (9.2, 2.4)	
14	36.19	2.03 dt (12.2, 6.6)	
15	80.2	4.11 m	
16	107.38		
17	11.75	0.84 d (7.0)	
18	209		
19	46	Not observed	N/A
		Not observed	N/A
20	85.36		
21	64.21	3.00 d (8.0)	
22	74.16	4.56 dq (8.3, 6.0)	
23	61.1	3.44 d (12.3)	
		4.09 d (12.3)	
24	20.41	1.29 d (6.0)	
25	170.94		
26	52.68	3.81 s	



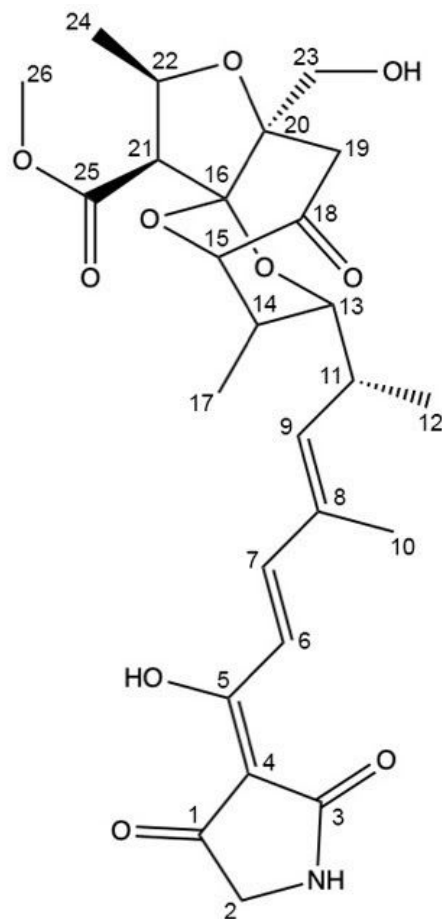
**Figure 10:** Spectral data for nocamycin O in MeOD (values obtained via 1D  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR). The initial data was mostly sufficient to prove most of the structure previously proposed as correct; however, two issues were identified through the NMR data collection. The first problem was that it was suspected that several  $\alpha$ -hydrogens had been replaced with deuteriums. This is most clearly seen in the signal for C-19, where no proton signals were observed (even though it should show up as a  $\text{CH}_2$ ). This necessitated another round of data collection in an aprotic solvent, which will be discussed later in this section. The second problem was the low confidence of the chemical shift assignment at C-4, likely because of the scarcity of protons in that region of the molecule. There were no HMBC, H2BC, or COSY signals coupled to that carbon (Figure 11).



**Figure 11:** Key NMR correlations (COSY - green, HMBC - pink) in MeOD.

The HMBC and H2BC coupling signals were extensive and useful for confirming the tricyclic structure of the upper part of the molecule. However, there is less evidence supporting the -OH group at C-23, as C-19, which is adjacent, was not coupling to anything. This was due to proton exchange at that location; however, the fact that the shift value of C-23 is greater than 60 ppm is indicative of its likely attachment to an oxygen atom. The inability to confirm the existence of the -OH group or the assignment at C-4 necessitated more data collection. dDMSO was chosen as the aprotic solvent (as it easily dissolves water-insoluble compounds), and the compound was redissolved, first in MeOH to reverse the proton exchange, and then in dDMSO. The same spectra were obtained for the compound in dDMSO as in MeOD, with the addition of a TOCSY spectrum. Spectral assignments as well as splitting and coupling constants for the proton spectrum were determined (Figure 12).

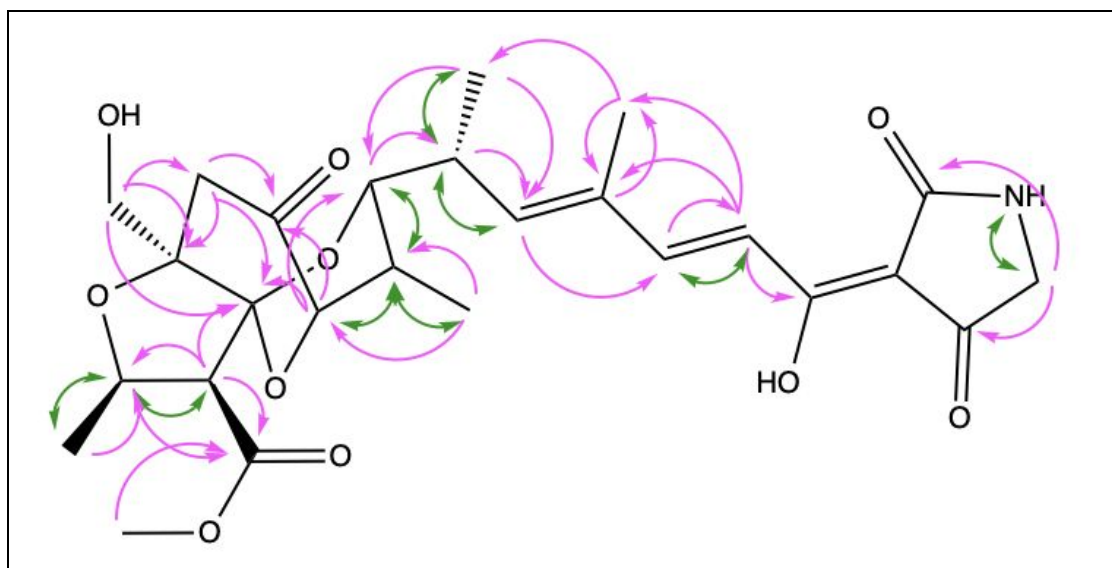
Position	$\delta$ C	Type	$\delta$ H	Mult (J in Hz)
1	194.31	C		
1*		NH	6.52 s	
2	49.82	CH <sub>2</sub>	3.32 s	
3	177.34	C		
4	Not observed	C		
5	181.82	C		
5*		OH	Not observed	N/A
6	127.2	CH	7.65 d (15.56)	
7	141.19	CH	7.04 d (15.56)	
8	134.83	C		
9	136.26	CH	5.65 d (9.98)	
10	12.9	CH <sub>3</sub>	1.76 s	
11	33.92	CH	2.75 t (8.44, 8.44)	
12	16.83	CH <sub>3</sub>	0.92 d (6.83)	
13	79.08	CH	3.64 d (11.6)	
14	34.54	CH	1.87 dt (6.56, 6.56, 12.6)	
15	78.59	CH	4.14 d (5.66)	
16	106.13	C		
17	11.5	CH <sub>3</sub>	0.74 d (7.04)	
18	208.21	C		
19	46.14	CH <sub>2</sub>	2.57 s (obscured) 2.97 s (obscured)	
20	84.49	C		
21	62.75	CH	2.88 d (8.32)	
22	72.34	CH	4.42 q (6.78, 6.78, 6.42)	
23	59.58	CH <sub>2</sub>	3.38 d (12.3) 3.81 d (12.25)	
23*		OH	not observed	N/A
24	20.31	CH <sub>3</sub>	1.21 d (6.05)	
25	169.36	C		
26	52.12	CH <sub>3</sub>	3.7 s	



**Figure 12:** Spectral data for nocamycin O in dDMSO (values obtained via 1D <sup>1</sup>H-NMR and 2D HSQC).

Following the second round of data collection, the issues relating to insufficient data at C-19 and C-23 were resolved. Both protons attached to the carbon at C-19 became visible (though their splitting is somewhat unclear due to the dDMSO solvent peak present at 2.50 ppm and the CH peak from H-21 at 2.88 ppm). The new visibility at C-19 means that all three CH<sub>2</sub> carbons originally proposed to be a part of the structure are now fully confirmed, a finding that can be corroborated by the phase-edited HSQC (in which all CH<sub>2</sub> peaks show up as blue instead of red). This increases the likelihood that nocamycin O is a novel member of the nocamycin family (the spectrum for nocamycin I, its nearest relative, has only two CH<sub>2</sub> groups).<sup>31</sup> However, this experiment was not without its issues. Neither -OH proton appeared coupled to anything;

there were some extremely wide signals around the 2.90-3.0 ppm region, but due to the presence of the C-19 proton, it was difficult to tell if that was the signal of only one proton or of two protons crowded together (Figure S12). Additionally, insufficient signal-to-noise prevented us from obtaining an interpretable  $^{13}\text{C}$  spectrum, so while any carbons that had HSQC or HMBC couplings could be identified, anything that had no homonuclear or heteronuclear proton associations became invisible. Thus, the carbon at C-4 had no confirmed chemical shift value because it had no HMBC or H2BC correlations (Figure 13).

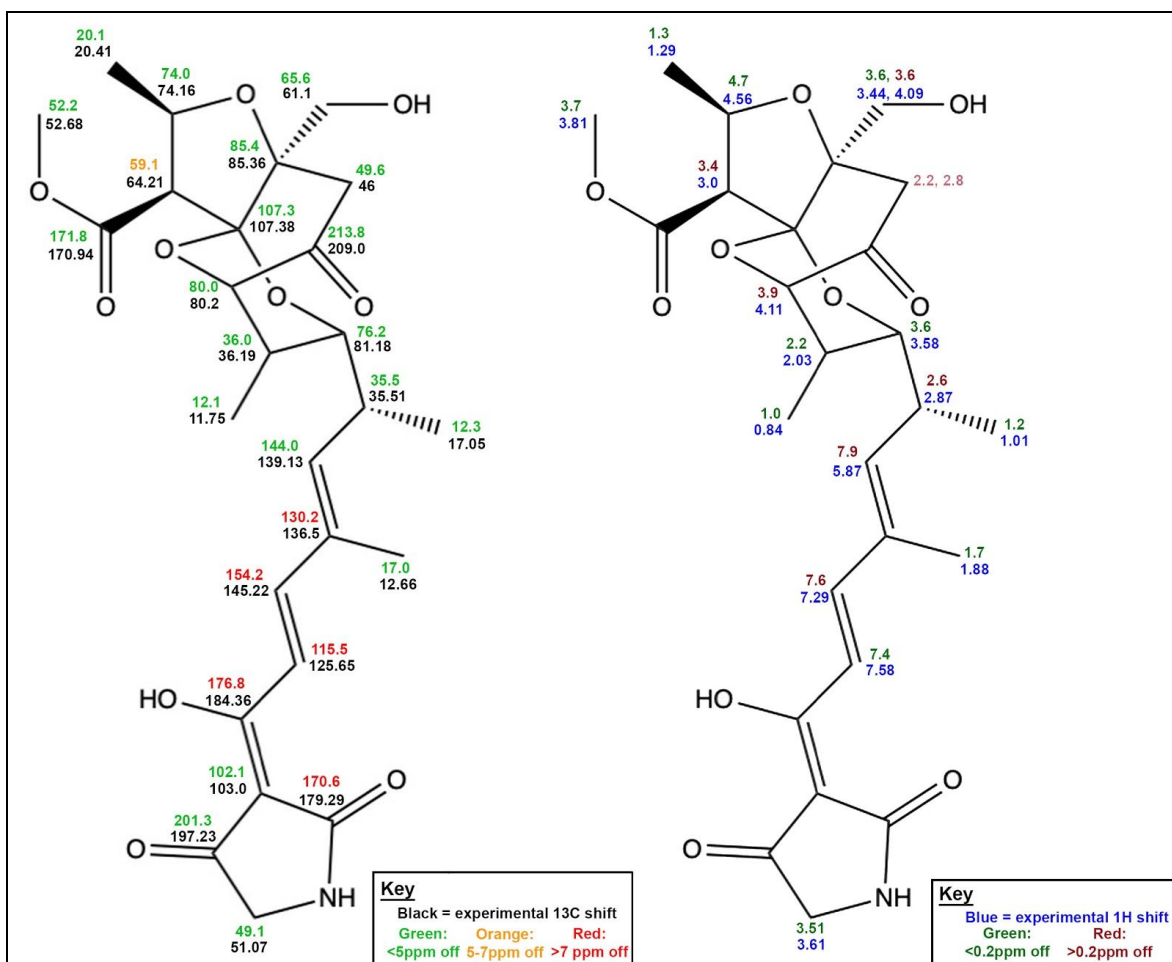


**Figure 13:** Key NMR correlations (COSY - green, HMBC - pink) in dDMSO.

Given the HMBC signals between C-1, C-2, and C-3, the COSY correlation between C-2 and the NH, the prior data in MeOD, and the similarity of the MS/MS and NMR data to that of nocamycin I, the chemical identity of the carbon at C-4 is likely that of an alkene/ $\alpha$ -carbon, though its actual identity and shift value cannot be confirmed based on the experimentally-obtained data.

## Confirmation of Experimental Chemical Shift Assignments - DFT in Spartan '18

To confirm that the proposed structure was correct, chemical shift calculations utilizing density functional theory were performed in Spartan '18 in order to compare the experimentally-determined value with the calculated value of each nucleus. The results of those calculations as compared to the experimental shifts in MeOD are shown below (Figure 14):

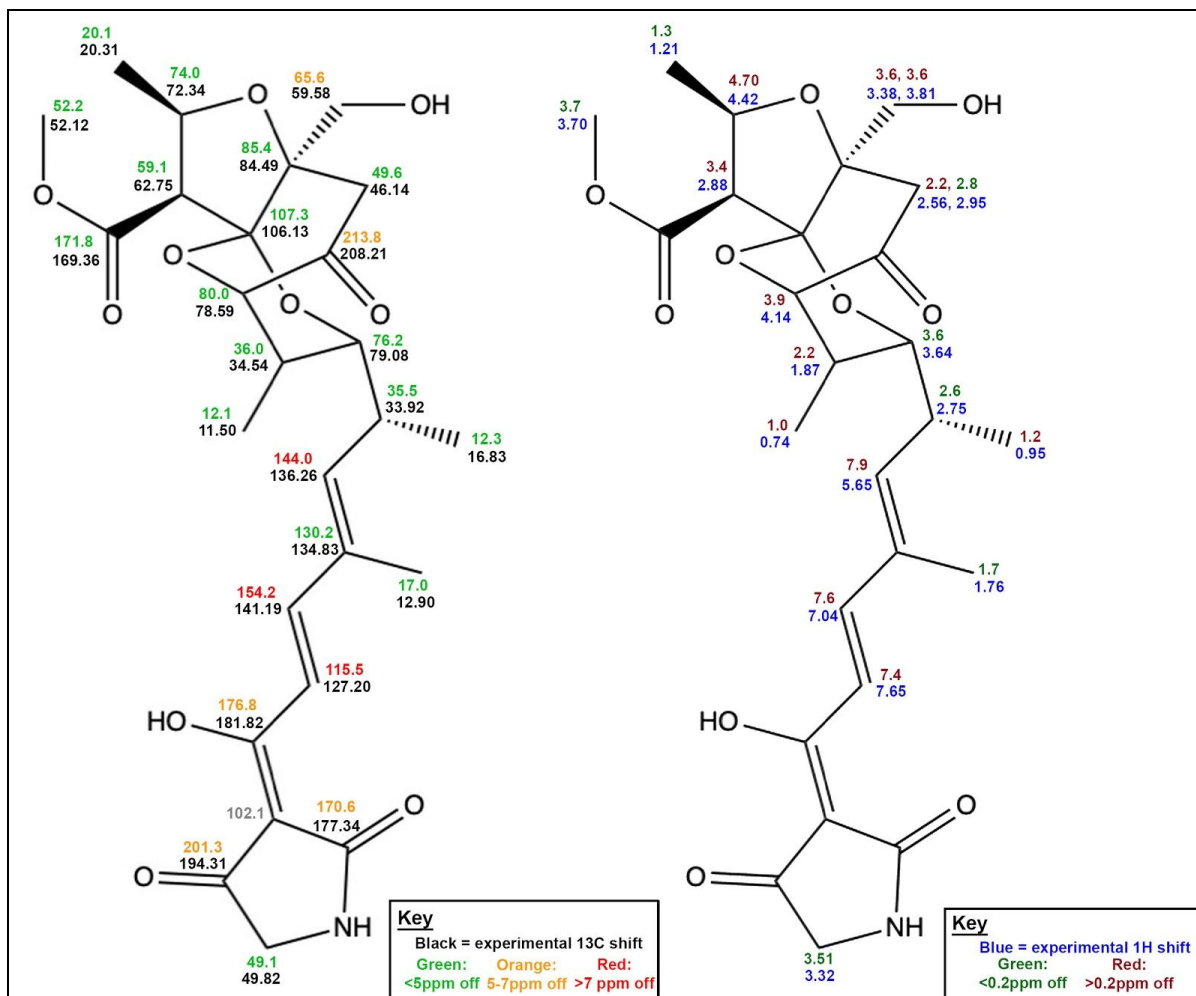


**Figure 14:** Comparison of experimental chemical shift values (MeOD) to calculated values. The proton values for C-19 (2.2, 2.8) are a paler color because they had no experimental equivalent.

The calculated and experimental carbon shift values were overall in agreement for the tricyclic portion of the molecule and slightly less in agreement for the olefinic middle portion of the molecule. The calculated proton shifts did not appear to follow a pattern in terms of which parts of the molecule had more accurate values. Overall, the calculated values did agree with the



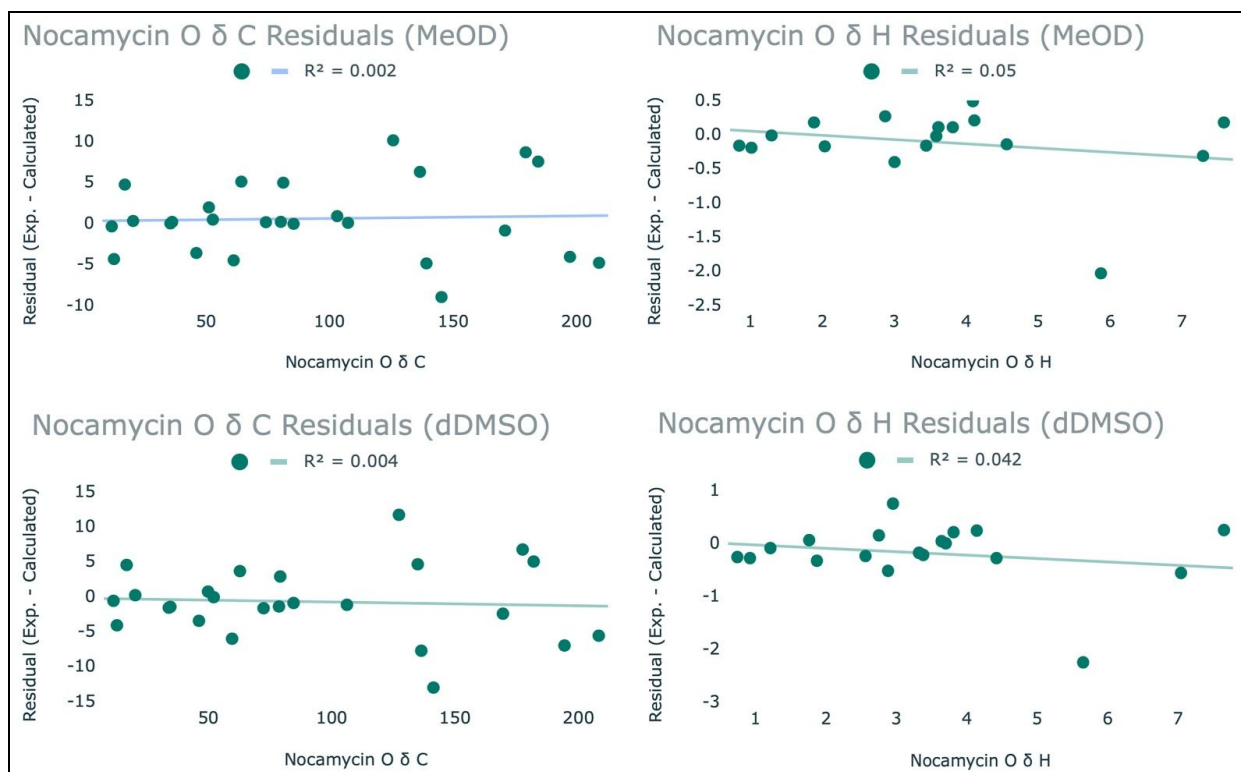
experimentally determined values, especially for the carbon spectrum. The dDMSO comparisons followed a largely similar trend (Figure 15):





## Residual Analysis of Density Functional Calculations

To ensure that the results of the density functional calculations could be considered accurate and that there was not an unknown source of systematic error present in the protocol, residual analysis was performed comparing the calculated chemical shift values to the experimental values in both MeOD and dDMSO (Figure 16):



**Figure 16:** Residual analysis comparing the calculated chemical shifts to the experimental values. All trendlines follow the linear least squares fit model and  $R^2$  values closer to zero are indicative of less correlation between chemical shift value and magnitude of error.

The residual analysis of the calculated and experimental values reveals no significant or worrying trends in the data; while the residual values for the carbon shifts do get larger as the chemical shift values increase, the graph, the trendline and the  $R^2$  value show that this error is not systematic (i.e. all positive errors or all negative errors). Thus, the protocol can be considered reliable and the calculated chemical shift values that it generates can be considered accurate.

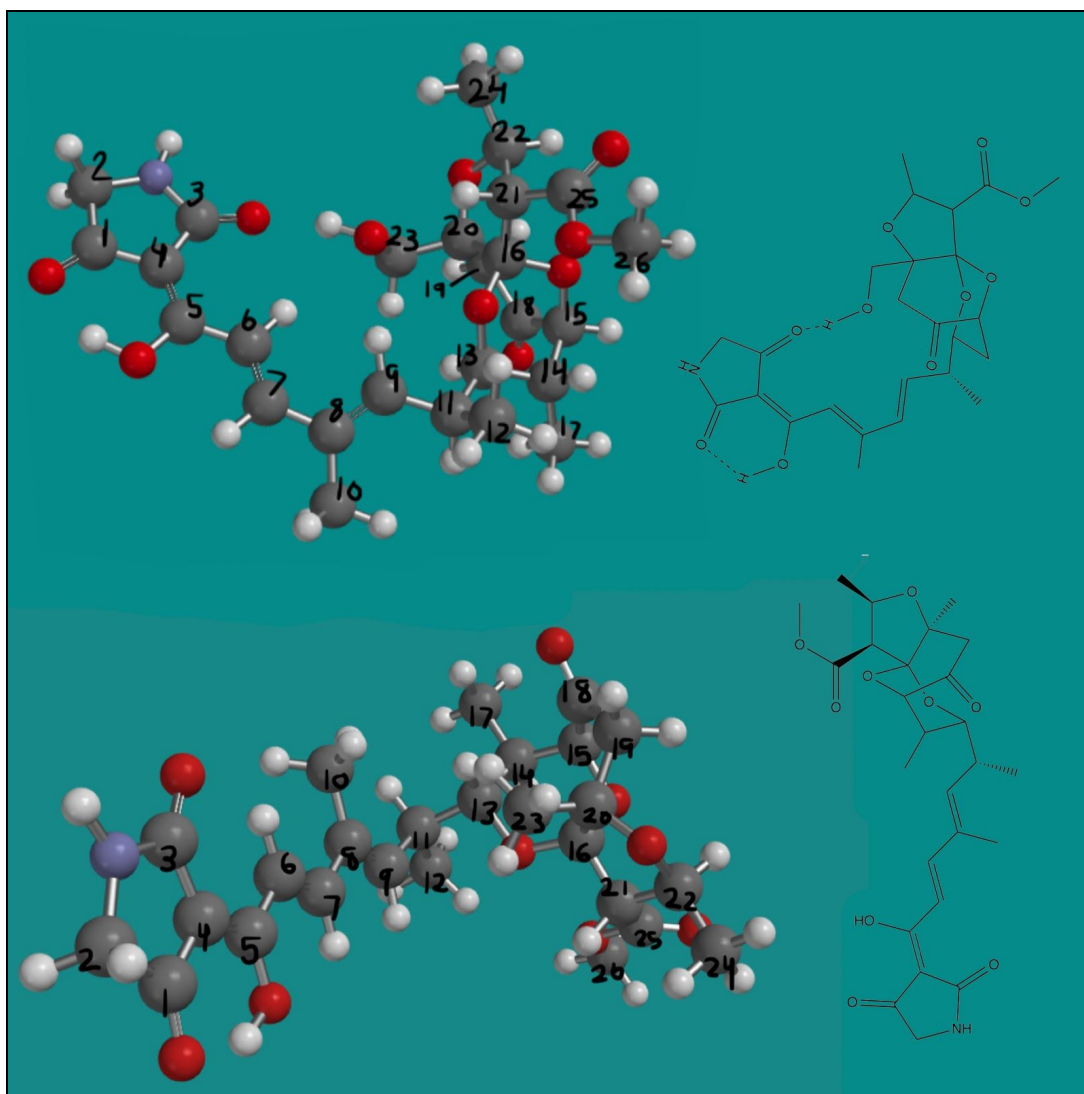
### Confirmation of Novelty - Comparison of Nocamycin I and Nocamycin O

Perhaps the most important part of the study of this compound is its confirmation as a novel compound. Thus, a comparison study between nocamycin O and nocamycin I (its most chemically similar relative) was conducted to determine if there were any significant and/or obvious differences in data collected for the two compounds. This study was done using data analysis of previously collected NMR data for nocamycin I and density functional calculations to generate calculated chemical shift values for nocamycin I.<sup>31</sup> A table comparing the experimentally derived chemical shift values of nocamycin I and nocamycin O in MeOD is shown below (Figure 17):

position	$\delta$ C (Noc O)	$\delta$ C (Noc I)	Delta (I - O)	$\delta$ H (Noc O)	$\delta$ H (Noc I)	Delta (I - O)
1	197.23	Not observed	N/A			
2	51.07	52.4	1.33	3.61	3.82	0.21
3	179.29	Not observed	N/A			
4	103	Not observed	N/A			
5	184.36	176.2	-8.16			
6	125.65	117.2	-8.45	7.58	7.22	-0.36
7	145.22	150.6	5.38	7.29	7.54	0.25
8	136.5	135.9	-0.6			
9	139.13	145.3	6.17	5.87	6.12	0.25
10	12.66	12.3	-0.36	1.88	1.92	0.04
11	35.51	36.2	0.69	2.87	2.94	0.07
12	17.05	17.1	0.05	1.01	1.06	0.05
13	81.18	80.1	-1.08	3.58	3.58	0
14	36.19	36.5	0.31	2.03	2.02	-0.01
15	80.2	80.2	0	4.11	4.09	-0.02
16	107.38	107.8	0.42			
17	11.75	11.9	0.15	0.84	0.85	0.01
18	209	209.7	0.7			
19	46	52.1	6.1	Not observed	2.63	N/A
				Not observed	2.63	N/A
20	85.36	83.2	-2.16			
21	64.21	64.4	0.19	3.00	2.9	-0.10
22	74.16	74.4	0.24	4.56	4.53	-0.03
23	61.1	21.8	-39.3	3.44	1.42	-2.02
				4.09		N/A
24	20.41	20.5	0.09	1.29	1.27	-0.02
25	170.94	172.1	1.16			
26	52.68	52.5	-0.18	3.81	3.81	0

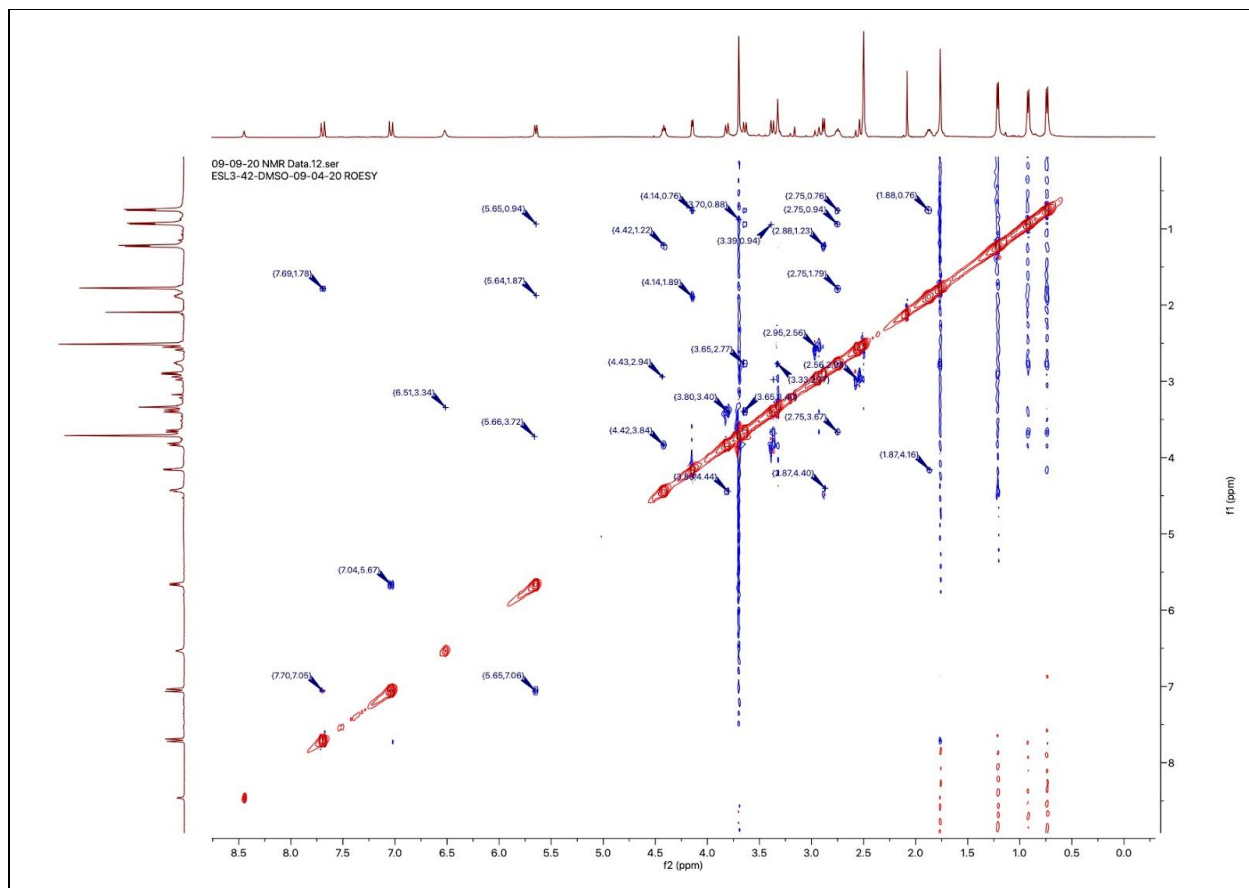
**Figure 17:** Comparison of experimental chemical shift values for nocamycin I (from Mo et. al. 2017) and nocamycin O (both in dDMSO). Proton shift differences of greater than 0.4 ppm and carbon shift differences of greater than 3 ppm have been highlighted.

As shown in Figure 17, C-2, C-5, C-6, C-7, C-9, C-19, and C-23 have marked differences in either their carbon or proton chemical shifts. The differences in chemical shift values seen in the olefinic section (C-5 to C-9) of the molecule likely have nothing to do with structural changes to the compound but rather conformational changes and a subsequent difference in the environment of these carbons in space. This theory was supported by the differences seen in the conformations best supported by the density functional calculations performed on nocamycin I and nocamycin O, as those conformations appeared to be quite different (Figure 18).



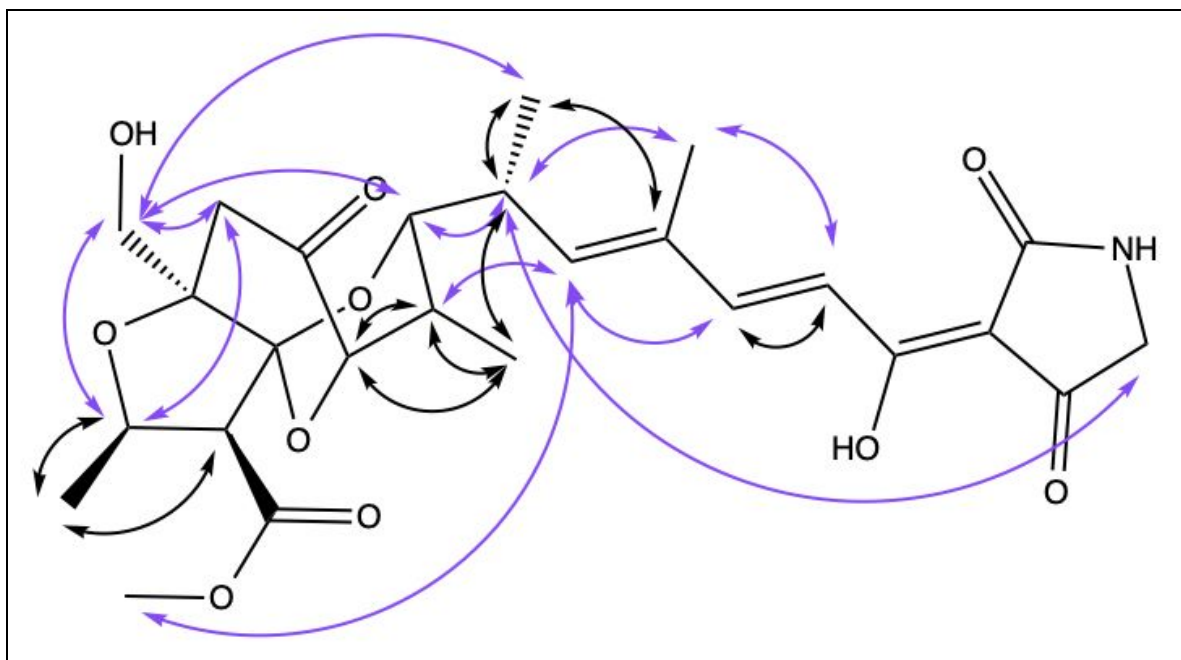
**Figure 18:** Conformational differences between best conformer of nocamycin O (top) and nocamycin I (bottom). Skeletal structures included for clarity.

The theory that nocamycin O adopts a novel conformation as compared to nocamycin I, and is therefore a novel molecule, is further supported by the additional ROESY spectrum gathered for the compound (Figure 19):



**Figure 19:** ROESY (2D proton correlation) spectrum for nocamycin O.

Though there is a considerable amount of t2 and t1 noise confounding the readings closer to the diagonal, quite a few interesting observations can be made based on the data further away from it. Notably, both 5.65 ppm (C-9) and 0.92 ppm (C-12) appear to be correlated with protons that are structurally very far away from them (Figure 20):



**Figure 20:** All ROESY correlations present on the spectrum in dDMSO. Purple arrows represent true NOE, while black arrows represent artifact that is also seen on the TOCSY (Figure S14).

C-9 (5.65 ppm) appears to be correlated with C-26 (3.70 ppm), C-10 (2.75 ppm) appears to be correlated with C-2 (3.32 ppm), and C-12 (0.92 ppm) appears to be correlated with C-23 (3.38 ppm), none of which are structurally near each other at all. None of these three key correlations appear on the TOCSY either, so it is unlikely that the correlations are artifacts or errors. Therefore, they must be spatially close to each other, as that is the only explanation for the signals correlations. These observations point to a unique conformation of nocamycin O similar to that shown in Figure 18 (though not exactly like that, as that conformation would make the correlation of C-26 and C-9 very unlikely). This is overwhelmingly due to the addition of the -OH group, which allows different and/or previously unobserved hydrogen bonding opportunities, increasing the likelihood that nocamycin O is a novel antibiotic.

## Discussion

### *Structure Confirmation of Nocamycin O*

The two rounds of experimental NMR data acquisition (in MeOD and then dDMSO) each individually confirmed certain aspects of the compound's structure. The MeOD run contained 26 different carbon signals (plus an additional one that was eventually discovered to be residual formic acid from the HPLC extraction), which supported the structure based on the significant genetic similarity between nocamycin O and nocamycin I (discovered through the analysis of the biosynthetic gene cluster) as well as the MS results (Figure 3). However, the lack of proton signals at C-19 made it difficult to confirm the presence of the -OH group at C-23, which was also suggested by the 16-amu mass difference between the MS results for nocamycin I and nocamycin O. Due to this lack of coupling and resulting ambiguity of which shift value belonged to C-19 and which belonged to C-23 (the MeOD shift values at those positions were 46.0 and 61.1 ppm, and 61.1 was first incorrectly assigned to C-19), more data collection was deemed necessary. The second data collection was done in dDMSO after a brief solvation in MeOH to exchange the protons back, and this time there was enough information to confirm the identities of C-19 and C-23, but no shift at all was seen for C-4 due to insufficient reach of the HMBC couplings and lack of a decent  $^{13}\text{C}$  spectrum due to poor signal-to-noise ratio (Figure 13). Though the inability to confirm C-4 in dDMSO was undesirable, the restoration of the  $\text{CH}_2$  at C-19 means that the tricyclic structure of the upper portion of the molecule as well as everything from C-5 onward can be fairly confirmed via the HMBC and COSY correlations in dDMSO, which are more extensive than those seen in MeOD (Figures 11 and 13). However, other than the singular NH coupling seen in the ROESY spectrum (Figure 19), no other heteroatomic couplings appear to be observed on either of the HMBC or H2BC spectra. It is somewhat likely that some

additional proton exchange may still be present in the sample via hydrogen bonding (especially at the -OH attached to C-5), thus rendering some or all of the -OH groups invisible on said spectra. Even though the spectrum mostly confirms the predicted structure (and agrees with the MS data) further NMR spectrometry analysis should be conducted on the compound (specifically a new HMBC of the compound that contains the missing -OH groups, as this is the only spectrum they are likely to appear on) to ensure that this conclusion is correct and to visualize all of the peaks seen in the MeOD and dDMSO experiments together in one comprehensive set of spectra.

#### *DFT Calculations*

The density functional theory calculations were additionally helpful for solving the issue of whether the shift at C-23 was 46 or 61 ppm. However, the amount of error present for other parts of the molecule did begin to raise questions about the validity of the protocol. While most of the carbon shift calculations were within the acceptable error range, the approximations were noticeably bad for C-5 through C-8 (the olefinic section of the molecule), as well as for many of the proton shifts. An acceptable margin of error for this protocol (as defined by the authors) is approximately 4-5 ppm for carbon shifts and less than 0.2 ppm for proton shifts, and there were a few carbon shift values that were off by 10 ppm or more.<sup>32</sup> While there does not appear to be any systematic error associated with the protocol, the approximations for the carbon shifts also appear to get less accurate as the experimental shift value increases, a trend that is more noticeable in the dDMSO residual plot but is present in both (Figure 16). There are several possible reasons for these inaccuracies; those which are most relevant here are conformational

effects on the experimental results, accuracy of the calculations and its impact on the Boltzmann weight assigned by the program, and the complex nature of the calculation of proton spectra.

Possibly the factor having the largest impact on the accuracy of the DFT calculations is conformational change in the experimental sample, which may be solvent-related or not. As a molecule with a fair amount of free rotation and cis/trans alkene options, nocamycin O has thousands of possible conformations. Its possibilities are somewhat restricted due to the ring-locked conformation of the upper region of the molecule, but the olefinic section of the molecule has numerous possible conformations and shapes. Density functional theory is also known to be inaccurate in measuring the shifts of alkenes and alkynes specifically, likely because it doesn't account for these possible conformational differences.<sup>33</sup> Additionally, differences in shift values for cis and trans forms of the same alkene can be between 0.2 and 0.4 ppm, while carbon shifts have an even larger margin of error.<sup>34</sup> The DFT calculations were also performed in the gas phase, which ignores any effects the solvent could possibly have on the conformation of the molecule. Thus, it is likely that the conformational distribution for nocamycin O is different in the gas phase and the solvated phase, resulting in particularly large errors in the most conformationally flexible region.

Inaccuracies created by the program also include errors stemming from the accuracy of the model that the calculations favor as well as the Boltzmann weight determined by the program in the final step. If the Boltzmann weight in the gas phase is different from the actual conformational distribution, this could result in an inaccurate approximation of the chemical shift values for those conformationally diverse areas, on the order of 5-8 ppm.<sup>32</sup> The Boltzmann weight seen *in silico* can also have a great effect on the calculated J values (coupling constants) which are helpful for determining the splitting and chemical identity of peaks in proton



spectra.<sup>32,35</sup> This is important for the accuracy of the proton spectrum because the coupling constants are the second derivatives of the energy with respect to the nuclear magnetic moment  $\mu$ . This quantity is determined by the program solely using the electronic structure of the molecule, which is extremely conformation-dependent.<sup>19,36</sup> Thus, the proton spectra are more likely to be incorrect if solvent effects on molecular conformation are not taken into account. The carbon spectrum does not appear to be as sensitive to this, though the increase in residual error as the  $^{13}\text{C}$  shift values increase does represent a possibly problematic issue with the protocol. Overall, the DFT calculations represent a decent confirmation of the structure of nocamycin O, but the values only appear trustworthy for the non-olefinic portions of the  $^{13}\text{C}$  spectrum. To ensure accuracy of the calculations and determine whether solvent interaction really has a large effect on conformation and on the calculated shift values, the calculations should be performed again in Spartan '18 under polar solvent conditions instead of in the gas phase.

#### *Nocamycin O as a Novel Member of the Nocamycin Family*

Perhaps the most successful portion of this thesis was the confirmation of nocamycin O as a new compound and a member of the nocamycin family. The analysis of the experimental NMR shift values for nocamycin O and nocamycin I revealed no large differences in a considerable portion of the molecule, indicating that the molecules are definitely chemically similar (Figure 17). However, certain areas of the molecule exhibited extremely large differences in shift values, namely C-2, C-5, C-6, C-7, C-9, C-19, and C-23. Most notably, the carbon shift value at C-23 increased nearly 38 ppm from nocamycin I to nocamycin O, while the proton shift value increased by nearly 2 ppm. The standard shift for a methyl group (C-23 identity in nocamycin I) is around 10-15 ppm, though C-23 is understandably a bit higher due to its

surrounding oxygen atoms. The carbon shift value for an RCH<sub>2</sub>OH methylene carbon is between 50 and 90 ppm; the C-23 value for nocamycin O falls right into this range at 59.58 ppm.<sup>37</sup> This supports a change in the chemical identity of C-23, and it also supports the idea that the addition at C-23 is an -OH group. The 16-amu difference in the MS data for the two compounds basically confirms this idea, because that would correspond to an identical molecule with one additional oxygen atom (Figure 3).

Other marked changes include the drastic differences seen in the olefinic section of the molecule (C-5, C-6, C-7, and C-9). Each of these carbons has a shift differential between 5 and 10 ppm, a significant amount of difference to be sure, but not enough to indicate an additional atom in the region.<sup>37</sup> Instead, the difference in chemical shifts at these positions is more likely to be a result of a change in the chemical environment directly adjacent to the nuclei. The experiment used to detect these changes was a ROESY spectrum (Figure 19). Our ROESY experiment detected several interesting interactions between structurally distant parts of the nocamycin O molecule. C-12 (0.92 ppm) appears to be correlated with at least one of the protons in the CH<sub>2</sub> group at C-23 (3.38 ppm), and C-9 (5.65 ppm) and C-26 (3.70 ppm) are also correlated, as well as C-10 (2.75 ppm) and C-2 (3.32 ppm). As these nuclei are too far apart to couple on anything dependent on bond interaction (i.e. an HMBC or H2BC), the only explanation for this coupling is that magnetic transfer is occurring through spatial association between the nuclei. The DFT calculations support this association as correct; the collapsed conformation of nocamycin O appears to be more stable than the linear, stretched-out conformation that nocamycin I favors. The linear conformation also would not allow the association of any of these three pairs, because they are extremely far apart in the linear conformation (Figure 18). This in conjunction with the drastic differences seen between the

experimental spectra for nocamycin O and nocamycin I point to nocamycin O being a closely related but novel member of the nocamycin antibiotic family.

## **Conclusion**

This thesis sought to fully elucidate the structure of nocamycin O, a novel antibiotic, using experimental and theoretical methods to confirm its place in the nocamycin family and also its distinct differences from all known members of said family. The experimental NMR data confirmed its novelty (especially when compared to its structurally similar analog nocamycin I), but there were several holes in the structure data that will have to be addressed in order to definitively confirm the correct structure. Chemical shift calculations served to confirm that most of the  $^{13}\text{C}$  shift assignments made experimentally were correct; the calculated proton shift values were less accurate, and calculations should be re-run in a polar solvent environment to determine if that has any effect on the accuracy of the proton shift approximations. Finally, a combination of both of these elements helped to further prove the novelty through spatially-dependent proton couplings (ROESY) present on the spectrum for nocamycin O that were not present for nocamycin I, as well as conformer analysis using DFT that showed extremely different preferred geometries for the two molecules.

Through this structural analysis, nocamycin O is shown to be a novel member of the nocamycin family. Given that other members of this family, as well as other tetramic acid-derived compounds like streptolydigin, have potent antibacterial effects due to their inhibition of RNA polymerase, this points to nocamycin O as an exciting area for future study as well as a potentially important antibiotic for use in the field of medicine.<sup>15,16</sup>

## **Acknowledgements**

Thank you to Claremont McKenna's SIE fund as well as Dr. Ethan Van Arnam's lab at W. M. Keck Science Center for funding this research. Additionally, thank you to Ethan Van Arnam and Bethany Caulkins for agreeing to be readers for this (massive) undertaking. Thank you to Janet Tran for acting as my primary fact-checker, idea-giver, and fellow thesis-writer; I would not have wanted to do this without your help. Thank you to Elisabeth Lawton, Rose Kim, and Georgia Scherer of the Van Arnam lab for their important foundational work upon which I built this thesis and thank you to Juan Santos, Arjun Deol, and Ananya Koneti for their invaluable support and insights during this trying semester.

## References

- [1]: Harvey, A. Natural Products in Drug Discovery. *Drug Discovery Today* 2008, 13(19-20), 894–901.
- [2]: Wohlleben, W.; Mast, Y.; Stegmann, E.; Ziemert, N. Antibiotic Drug Discovery. *Microbial Biotechnology* 2016, 9 (5), 541–548.
- [3]: Liu, R.; Li, X.; Lam, K. Combinatorial Chemistry in Drug Discovery. *Current Opinion in Chemical Biology* 2017, 38, 117–126.
- [4]: Kellner, R. L. L. Stadium-Specific Transmission of Endosymbionts Needed for Pederin Biosynthesis in Three Species of Paederus Rove Beetles. *Entomologia Experimentalis et Applicata* 2003, 107(2), 115–124.
- [5]: Clay, K. Defensive Symbiosis: a Microbial Perspective. *Functional Ecology* 2014, 28(2), 293–298.
- [6]: Kaltenpoth, M.; Engl, T. Defensive Microbial Symbionts in Hymenoptera. *Functional Ecology* 2013, 28(2), 315–327.
- [7]: Van Arnem, E. B.; Currie, C. R.; Clardy, J. Defense Contracts: Molecular Protection in Insect-Microbe Symbioses. *Chemical Society Reviews* 2018, 47(5), 1638–1651.
- [8]: Cafaro, M. J.; Poulsen, M.; Little, A. E. F.; Price, S. L.; Gerardo, N. M.; Wong, B.; Stuart, A. E.; Larget, B.; Abbot, P.; Currie, C. R. Specificity in the Symbiotic Association between Fungus-Growing Ants and Protective Pseudonocardia Bacteria. *Proceedings of the Royal Society B: Biological Sciences* 2010, 278(1713), 1814–1822.
- [9]: Chang, P.; Rao, K.; Longo, L.; Lawton, E.; Scherer, G.; Van Arnem, E. Thiopeptide Defense by an Ant's Bacterial Symbiont. *Journal of Natural Products* 2020, 83, 725–729.
- [10]: Kim, R.; Lawton, E.; Van Arnem, E. B. Identification of an Antibiotic Compound Produced by *Amycolatopsis*, a bacterial symbiont of *Trachymyrmex smithi*. Southern California Undergraduate Research Conference, 2019.
- [11]: Lawton, E.; Kim, R.; Van Arnem, E. B. Structural Characterization of an Antibiotic Compound Produced by an *Amycolatopsis* Bacterial Symbiont of *Trachymyrmex smithi*. Keck Undergraduate Summer Research Symposium, 2019.

- [12]: Tuske, S.; Sarafianos, S.; Wang, X.; Hudson, B.; Sineva, E.; Mukhopadhyay, J.; Birktoft, J.; Leroy, O.; Ismail, S.; Clark, A.; Dharia, C.; Napoli, A.; Laptenko, O.; Lee, J.; Borukhov, S.; Ebricht, R.; Arnold, E. Inhibition of Bacterial RNA Polymerase by Streptolydigin: Stabilization of a Straight-Bridge-Helix Active-Center Conformation. *Cell* 2005, 122(4), 541–552.
- [13]: Mo, X.; Shi, C.; Gui, C.; Zhang, Y.; Ju, J.; Wang, Q. Identification of Nocamycin Biosynthetic Gene Cluster from *Saccharothrix Syringae* NRRL B-16468 and Generation of New Nocamycin Derivatives by Manipulating Gene Cluster. *Microbial Cell Factories* 2017, 16(1).
- [14]: Toda, S.; Nakagawa, S.; Naito, T.; Kawaguchi, H. Bu-2313, a New Antibiotic Complex Active against Anaerobes. III. Semi-Synthesis of Bu-2313 A and B, and Their Analogs. *The Journal of Antibiotics* 1980, 33(2), 173–181.
- [15]: Temiakov, D.; Zenkin, N.; Vassilyeva, M. N.; Perederina, A.; Tahirov, T. H.; Kashkina, E.; Savkina, M.; Zorov, S.; Nikiforov, V.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Severinov, K.; Vassilyev, D. G. Structural Basis of Transcription Inhibition by Antibiotic Streptolydigin. *Molecular Cell* 2005, 19(5), 655–666.
- [16]: Tsukiura, H.; Tomita, K.; Hanada, M.; Kobaru, S.; Tsunakawa, M.; Fujisawa, K.-I.; Kawaguchi, H. Bu-2313, a New Antibiotic Complex Active against Anaerobes. I. Production, Isolation and Properties of Bu-2313 A and B. *The Journal of Antibiotics* 1980, 33(2), 157–165.
- [17]: Carlson, J. C.; Fortman, J. L.; Anzai, Y.; Li, S.; Burr, D. A.; Sherman, D. H. Identification of the Tirandamycin Biosynthetic Gene Cluster from *Streptomyces* Sp. 307-9. *ChemBioChem* 2010, 11(4), 564–572.
- [18]: Torres, A. M.; Price, W. S. Common Problems and Artifacts Encountered in Solution-State NMR Experiments. *Concepts in Magnetic Resonance Part A* 2016, 45A (2).
- [19]: Gryff-Keller, A. Theoretical Modeling of <sup>13</sup>C NMR Chemical Shifts - How to Use the Calculation Results. *ChemInform* 2012, 43 (26).
- [20]: Balci, M. *Basic 1H- and 13C-NMR spectroscopy*; Elsevier: Amsterdam, 2005.
- [21]: Jackowski, K.; Kubiszewski, M.; Wilczek, M. <sup>13</sup>C and <sup>1</sup>H nuclear magnetic shielding and spin–spin coupling constants of <sup>13</sup>C-enriched bromomethane in the gas phase. *Chemical Physics Letters* 2007, 440, 176-179.
- [22]: Hehre, W. J.; Radom, L.; Schleyer, P. V. R.; Pople, J. A. *Ab initio molecular orbital theory*; John Wiley & Sons, Inc.: New York, NY, 1986.

- [23]: Parr, R. G.; Weitao, R. G. Y. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, NY, 1995.
- [24]: Eriksson, O.; Wills, J. First Principles Theory of Magneto-Crystalline Anisotropy. *Electronic Structure and Physical Properties of Solids Lecture Notes in Physics* 2000, 247–285.
- [25]: Hoffman, R. <http://chem.ch.huji.ac.il/nmr/whatisnmr/whatisnmr.html> (accessed Sep 4, 2020).
- [26]: Hoffman, R. <http://chem.ch.huji.ac.il/nmr/techniques/2d/hetcor/hetcor.html> (accessed Sep 4, 2020).
- [27]: Hoffman, R. <http://chem.ch.huji.ac.il/nmr/techniques/2d/hmbc/hmbc.html> (accessed Sep 4, 2020).
- [28]: ur-Rahman, A.; Choudhary, M. I. *Applications of NMR Spectroscopy*; Bentham Science Publishers Ltd.: Shāriqah, United Arab Emirates, 2015; Vol. 3.
- [29]: Decatur, J. NOESY and ROESY. [https://nmr.chem.columbia.edu/sites/default/files/content/NOESY and ROESY experiments.pdf](https://nmr.chem.columbia.edu/sites/default/files/content/NOESY%20and%20ROESY%20experiments.pdf).
- [30]: Hehre, W.; Klunzinger, P.; Deppmeier, B.; Driessen, A.; Uchida, N.; Hashimoto, M.; Fukushi, E.; Takata, Y. Efficient Protocol for Accurately Calculating <sup>13</sup>C Chemical Shifts of Conformationally Flexible Natural Products: Scope, Assessment, and Limitations. *Journal of Natural Products* 2019, 82, 2299–2306.
- [31]: Mo, X.; Gui, C.; Wang, Q. Elucidation of a Carboxylate O-Methyltransferase NcmP in Nocamycin Biosynthetic Pathway. *Bioorganic & Medicinal Chemistry Letters* 2017, 27(18), 4431–4435.
- [32]: Tähtinen, P.; Bagno, A.; Klika, K. D.; Pihlaja, K. Modeling NMR Parameters by DFT Methods as an Aid to the Conformational Analysis of Cis-Fused 7a(8a)-Methyl Octa(Hexa)Hydrocyclopenta[d][1,3]Oxazines and [3,1]Benzoxazines. *Journal of the American Chemical Society* 2003, 125(15), 4609–4618.
- [33]: Iron, M. A. Evaluation of the Factors Impacting the Accuracy of <sup>13</sup>C NMR Chemical Shift Predictions Using Density Functional Theory—The Advantage of Long-Range Corrected Functionals. *Journal of Chemical Theory and Computation* 2017, 13 (11), 5798–5819.

[34]: Curtin, D. Y.; Dayagi, S. Effects On The N.m.r. Spectra Of The Stereochemistry Of 3,4-Diaryladipic Esters And 1,2-Diarylcyclopentanes. *Canadian Journal of Chemistry* 1964, 42(4), 867–877.

[35]: Bally, T.; Rablen, P. R. Quantum-Chemical Simulation of <sup>1</sup>H NMR Spectra. 2.† Comparison of DFT-Based Procedures for Computing Proton–Proton Coupling Constants in Organic Molecules. *The Journal of Organic Chemistry* 2011, 76(12), 4818–4830.

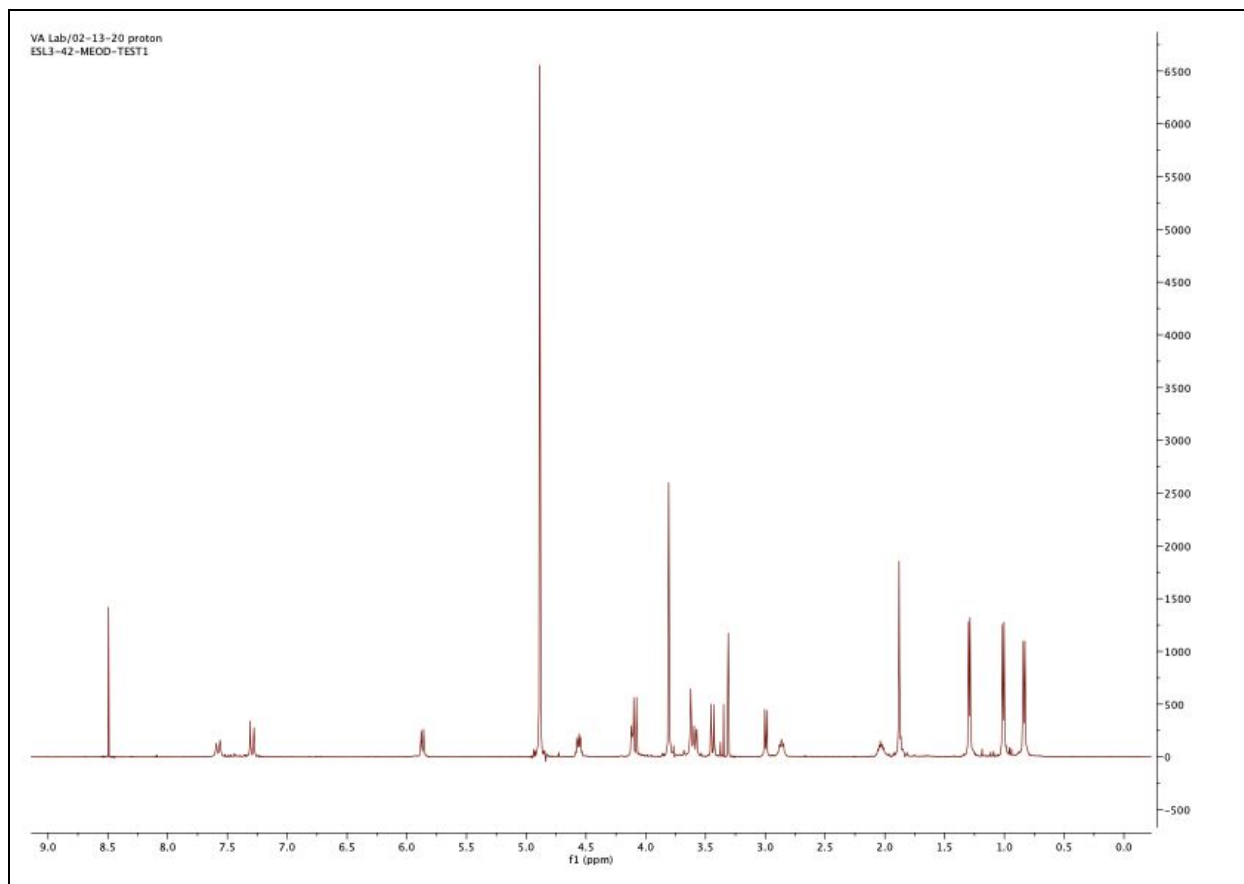
[36]: Brena, B.; Zhuang, G. V.; Augustsson, A.; Liu, G.; Nordgren, J.; Guo, J.-H.; Ross, P. N.; Luo, Y. Conformation Dependence of Electronic Structures of Poly(Ethylene Oxide). *The Journal of Physical Chemistry B* 2005, 109(16), 7907–7914.

[37]: Clark, J. Interpreting C-13 NMR Spectra.  
<https://www.chemguide.co.uk/analysis/nmr/interpretc13.html> (accessed Oct 17, 2020).

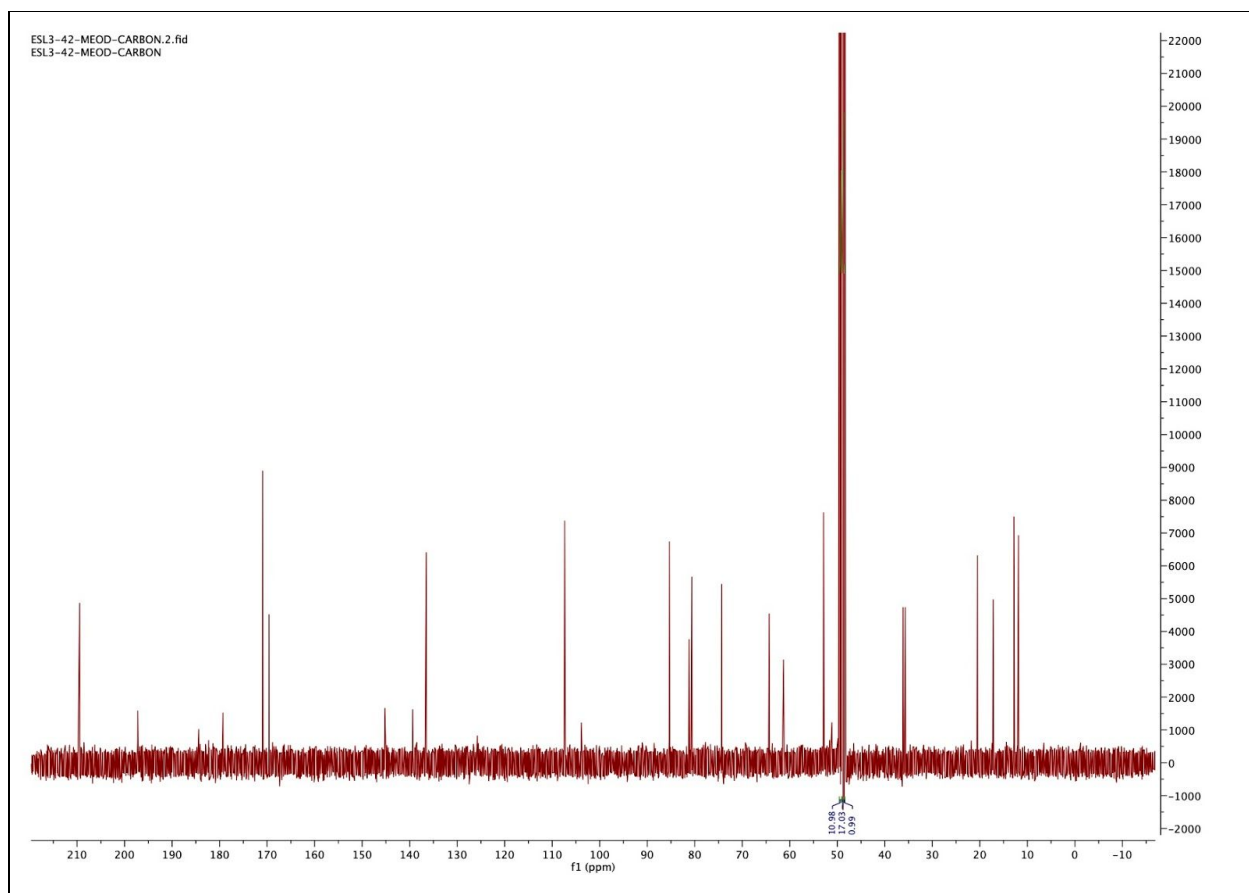
[38]: Cogan, D. P.; Ly, J.; Nair, S. K. Structural Basis for Enzymatic Off-Loading of Hybrid Polyketides by Dieckmann Condensation. *ACS Chemical Biology* 2020.



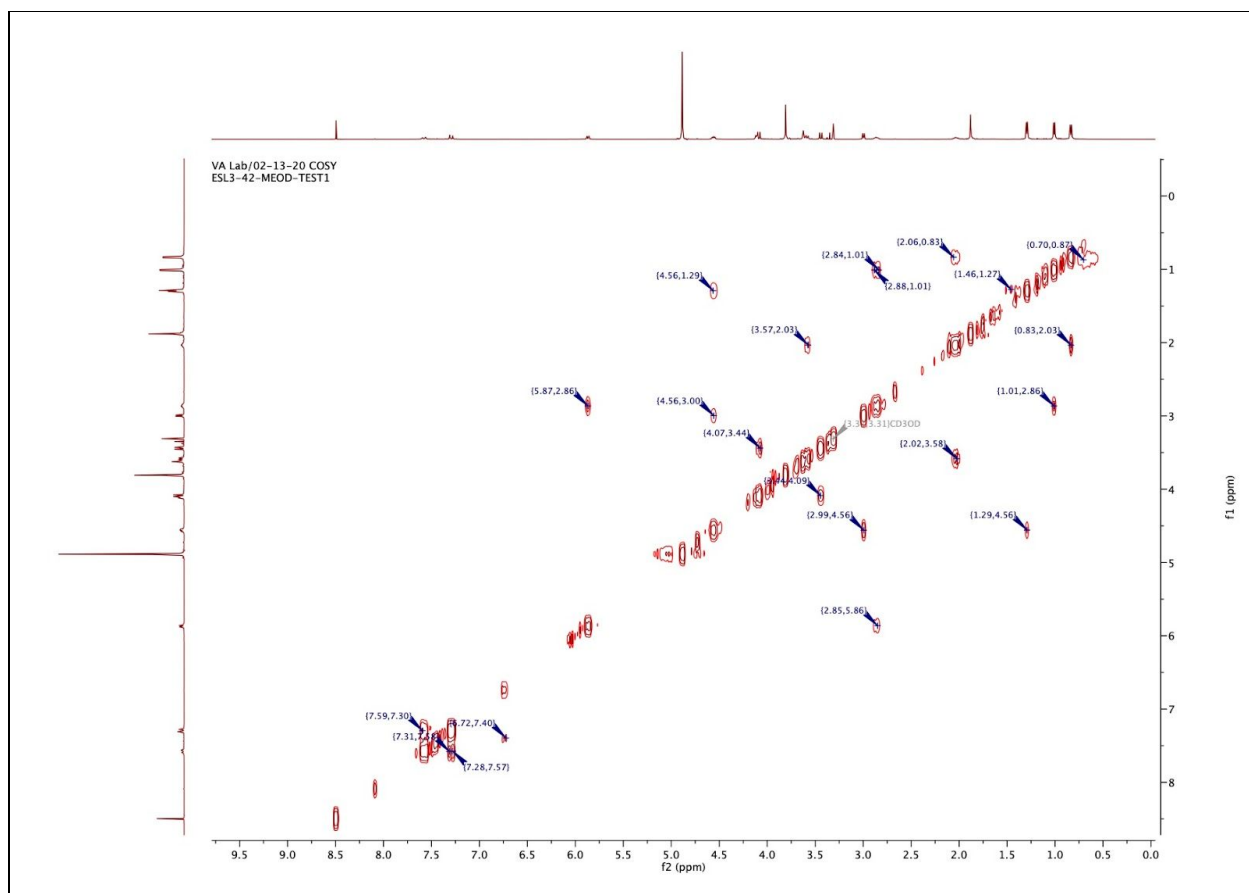
**Supplementary Figures**  
*Section 1: MeOD NMR Data*



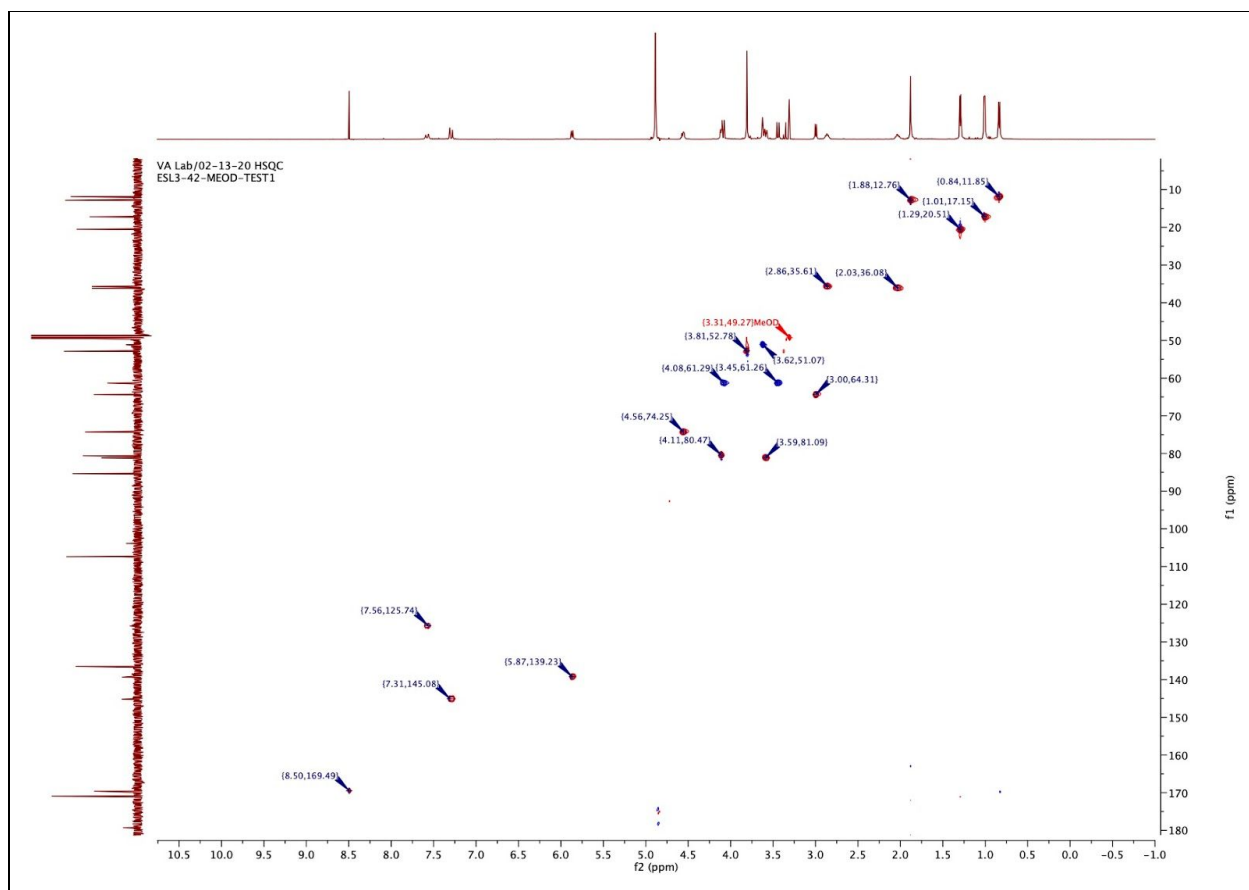
**Figure S1:** <sup>1</sup>H-NMR spectrum of nocamycin O in MeOD.



**Figure S2:**  $^{13}\text{C}$ -NMR spectrum of nocamycin O in MeOD.



**Figure S3:** 2D COSY spectrum of nocamycin O in MeOD.



**Figure S4:** 2D HSQC spectrum of nocamycin O in MeOD.

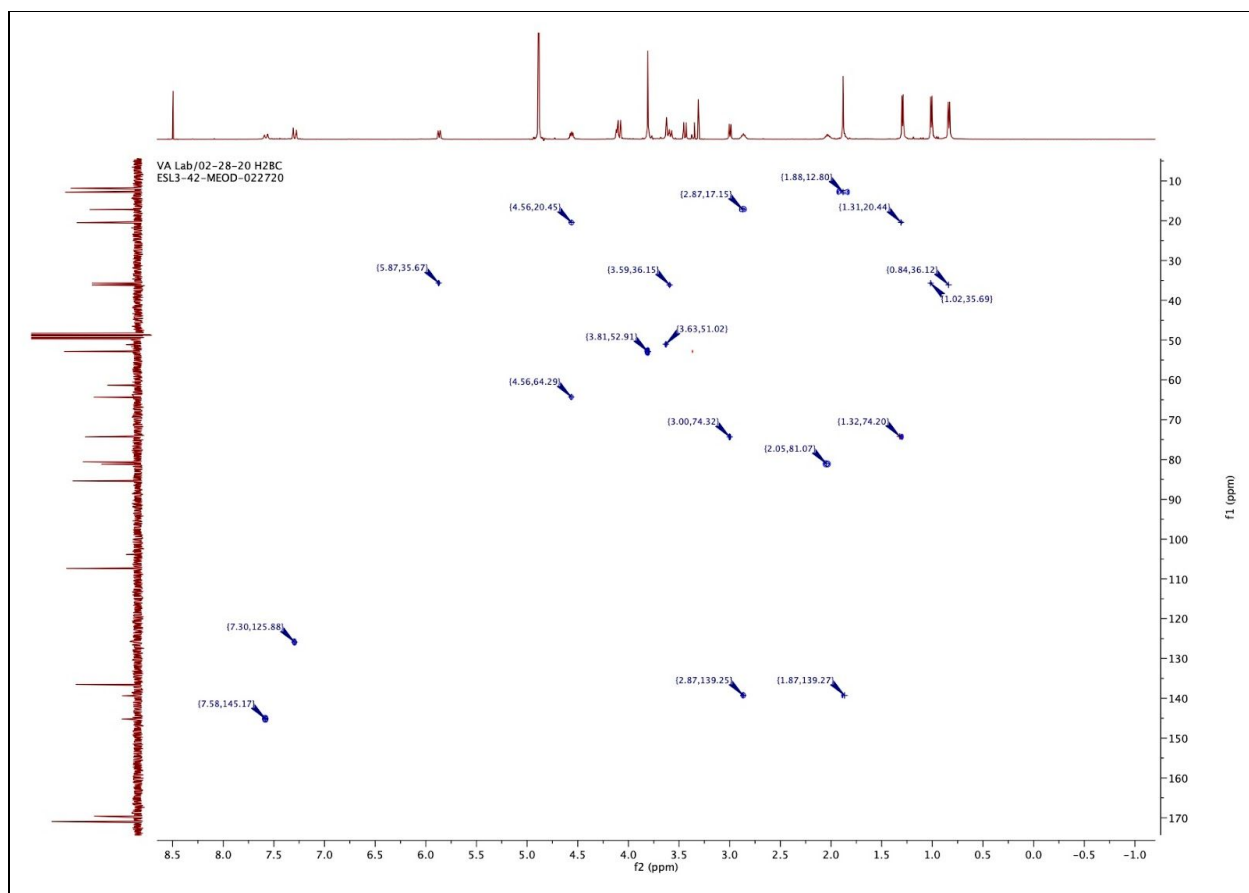


Figure S5: 2D H2BC spectrum of nocamycin O in MeOD.

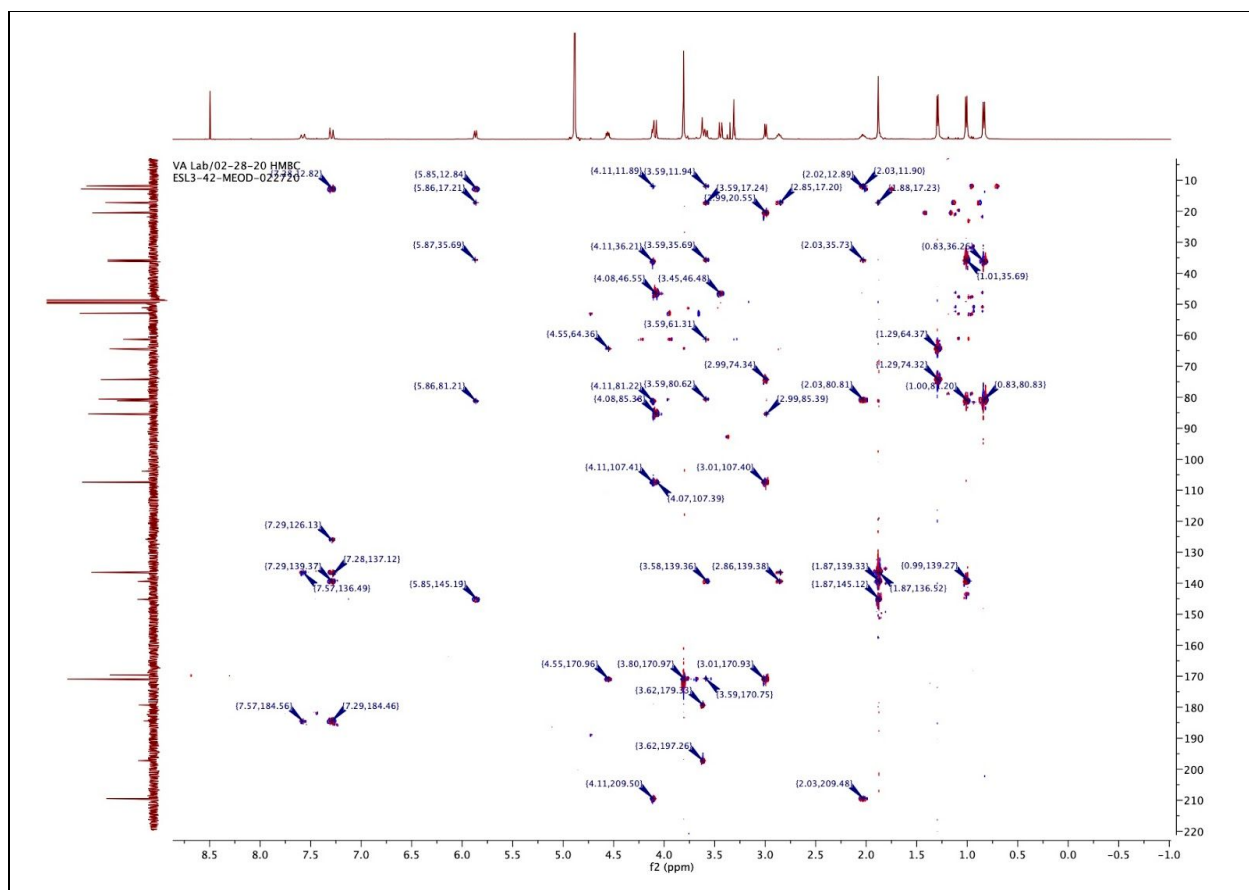
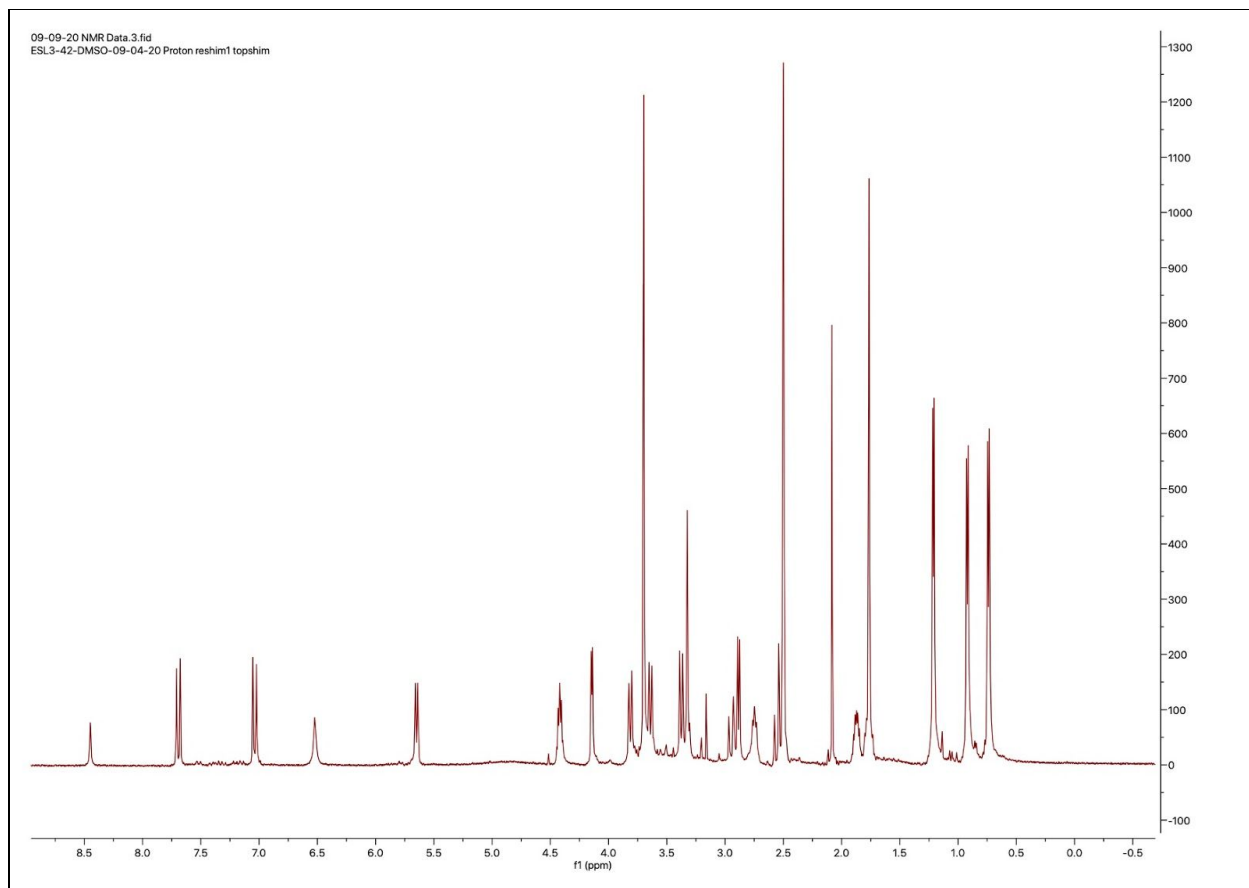


Figure S6: 2D HMBC spectrum of nocamycin O in MeOD.

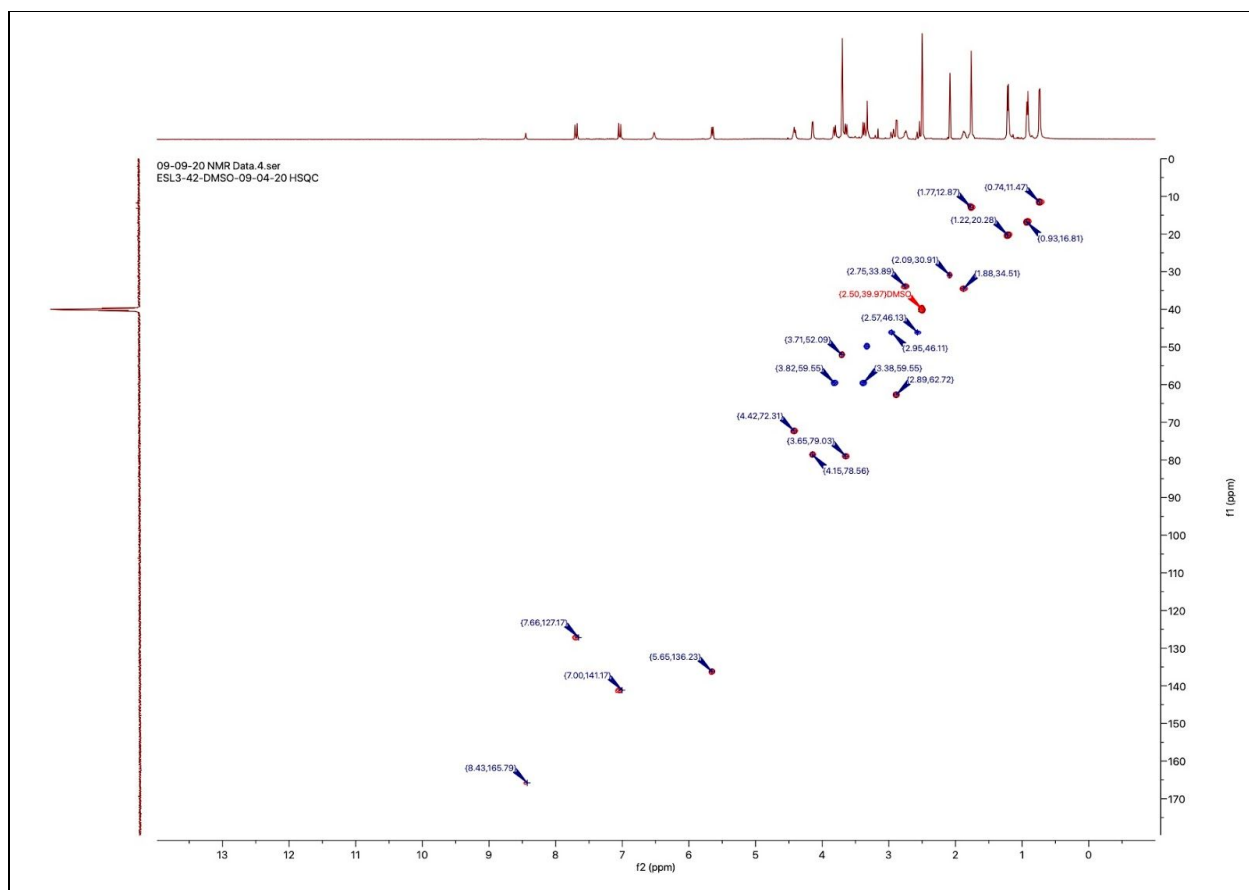
Hydrogens	Carbon (HSQC)	COSY	HMBC	H2BC
0.84	11.75	2.03	36.19, 80.2	36.19
1.01	17.05	2.87	35.51, 81.18, 139.13	35.51
1.29	20.41	4.56	64.21, 74.16	20.41, 74.16
1.88	12.66	none	17.05, 136.5, 139.13, 145.22	12.66, 139.13
2.03	36.19	0.84, 3.58	11.75, 12.66, 35.51, 81.18, 209.0	81.18
2.87	35.51	2.03	17.05, 136.5, 139.13	17.05, 139.13
3.0	64.21	4.56	20.41, 74.16, 85.36, 107.38, 170.94	74.16
3.44	61.1	4.09	46.0	none
3.58	81.2	2.03	11.75, 17.05, 35.51, 61.1, 80.2, 139.13, 170.94	35.51
3.61	51.07	none	179.29, 197.23	51.07
3.81	52.68	none	170.94	52.68
4.09	61.1	3.44	46.0, 85.36, 107.38	none
4.11	80.2	none	11.75, 36.19, 81.18, 107.38, 209.0	none
4.56	74.16	1.29, 3.0	64.21, 170.94	20.41, 64.21
5.87	139.13	2.87	12.66, 17.05, 35.51, 81.18, 145.22	35.51
7.29	145.22	7.58	12.66, 125.65, 136.5, 139.13, 184.36	125.65
7.58	125.65	7.29	136.5, 184.36	145.22

Figure S7: Comprehensive correlation table of all correlations/couplings seen in MeOD.

Section 2: dDMSO NMR Data

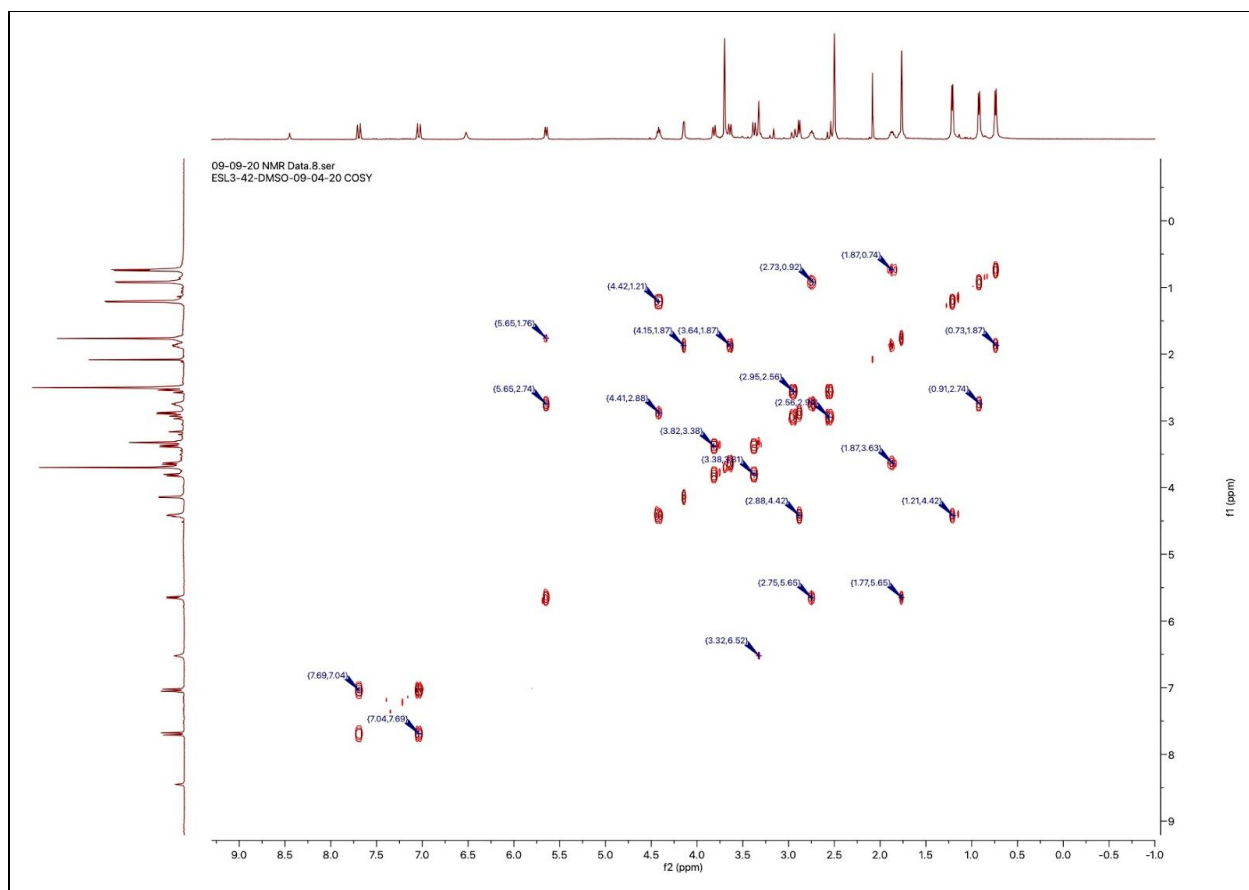


**Figure S8:** 1D <sup>1</sup>H-NMR spectrum for nocamycin O in dDMSO.

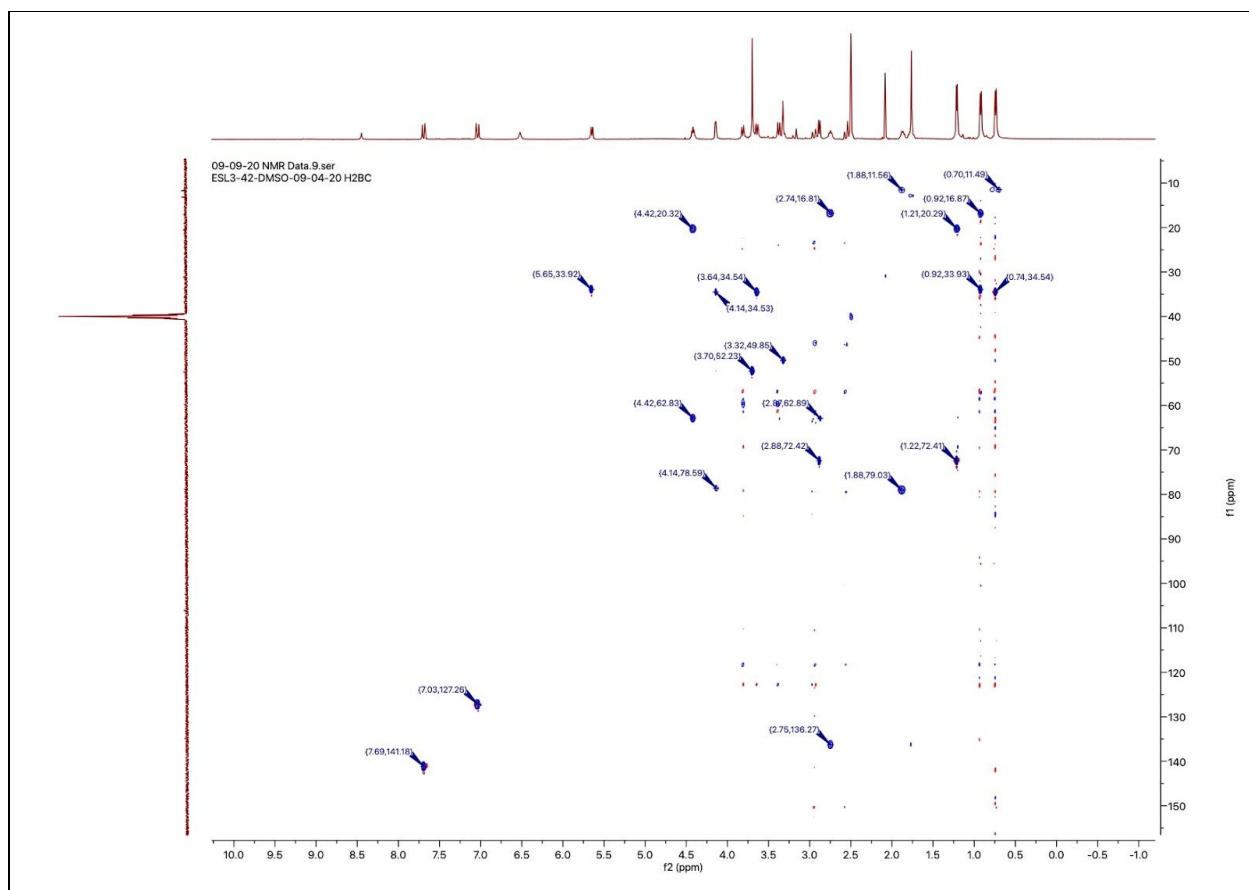


**Figure S9:** 2D HSQC spectrum for nocamycin O in dDMSO.





**Figure S10:** 2D COSY spectrum for nocamycin O in dDMSO.



**Figure S11:** 2D H2BC spectrum for nocamycin O in dDMSO.

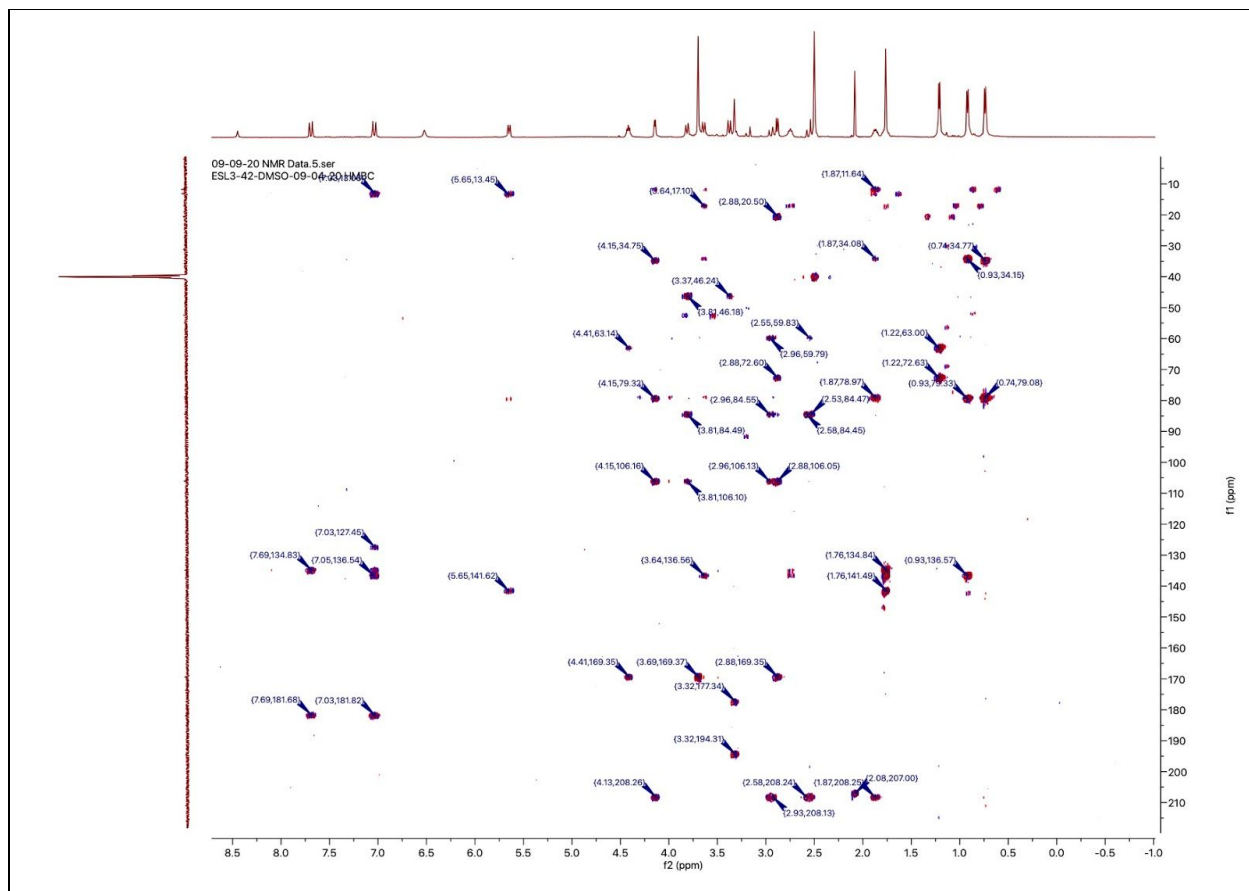
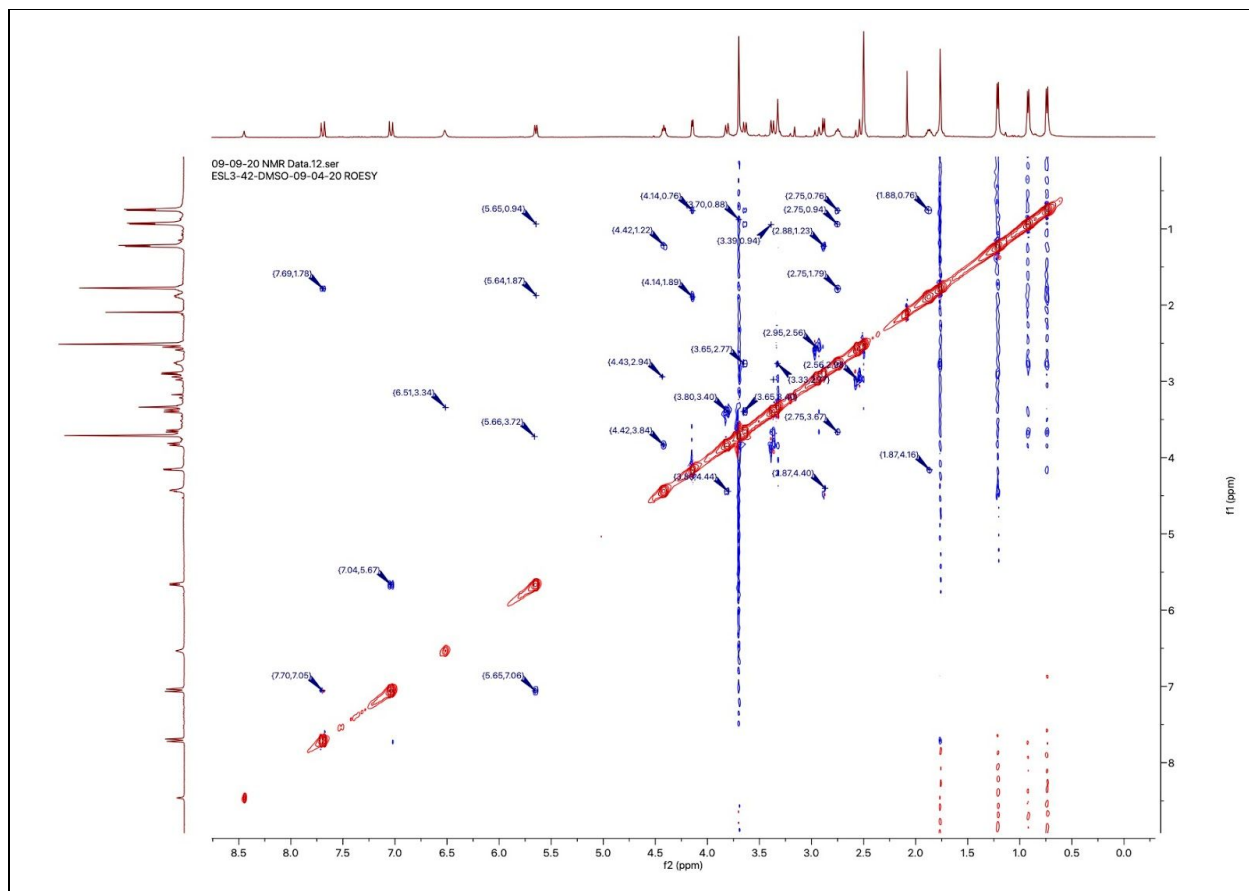
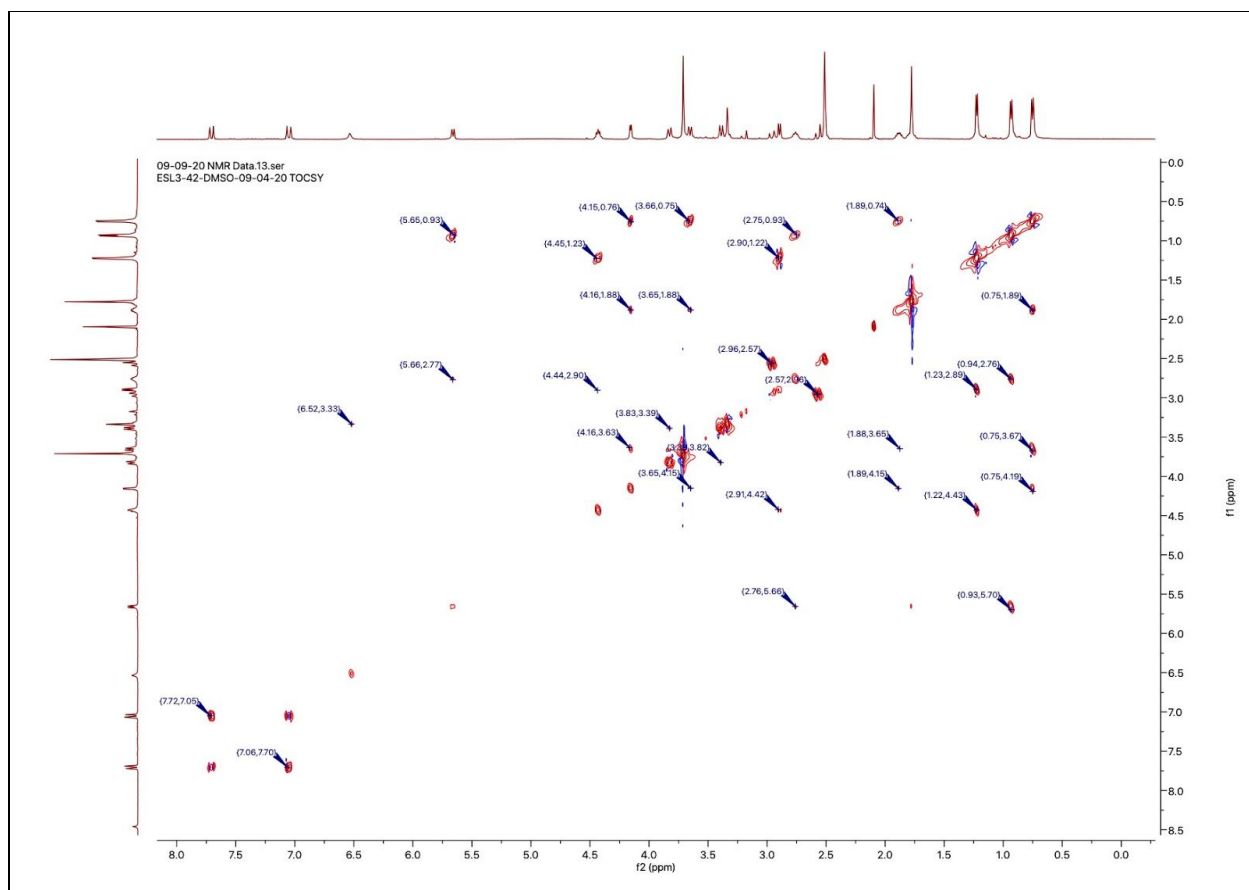


Figure S12: 2D HMBC spectrum for nocamycin O in dDMSO.



**Figure S13:** 2D ROESY spectrum for nocamycin O in dDMSO.



**Figure S14:** 2D TOCSY spectrum for nocamycin O in dDMSO.

Hydrogens	Carbon (HSQC)	COSY	HMBC	H2BC	ROESY	TOCSY
0.74	11.5	1.87	34.54, 78.59	34.54	1.87, 2.75, 4.14	1.87, 2.75, 4.14
0.92	16.83	2.75	33.92, 79.06, 136.26	33.92	2.75, 3.38, 5.65	2.75, 5.65
1.21	20.31	4.42	62.75, 72.34	72.34	2.88, 4.42	2.88, 4.42
1.76	12.9	5.65	134.83, 141.19	none	2.75, 7.65	none
1.87	34.54	0.74, 3.64	11.5, 78.59, 208.21	33.92	0.74, 4.14, 5.65	0.74, 1.87, 4.14
2.56	46.14	2.95	84.49, 208.21	none	2.95	2.95
2.75	33.92	0.92, 5.65	46.14, 84.49, 106.13	16.83, 136.26	0.74, 0.92, 1.76, 3.32, 3.64	0.92, 5.65
2.88	62.75	4.42	20.31, 72.34, 106.13, 169.36	72.34	1.21	0.92, 1.21, 4.42, 5.65
2.95	46.14	2.56	59.58, 84.49, 106.13, 208.21	none	2.56, 3.38, 4.42	2.56
3.32	49.82	6.52	177.34, 194.31	none	2.75, 6.52	6.52
3.38	59.58	3.81	46.14	none	0.92, 2.95, 3.64, 3.81	3.81
3.64	79.06	1.87	16.83, 136.26	34.54	2.75, 3.38	0.74, 1.87, 4.14
3.7	52.12	none	169.36	none	5.65	none
3.81	59.58	3.38	46.14, 84.49, 106.13	none	3.38, 4.42	3.38
4.14	78.59	1.87	34.54, 79.06, 106.13, 208.21	34.54	0.74, 1.87	0.74, 1.87, 3.64
4.42	72.34	1.21, 2.88	62.75, 169.36	20.31, 62.75	1.21, 2.88, 3.81	1.21, 2.88
5.65	136.26	1.76, 2.75	12.9, 141.19	33.92	0.92, 1.87, 3.7, 7.04	0.92, 2.75
6.52	(nitrogen)	3.32	none	none	3.32	3.32
7.04	141.19	7.65	134.83, 181.82	127.2	5.65, 7.65	7.65
7.65	127.2	7.04	134.83, 181	141.19	1.76, 7.04	7.04

Figure S15: Comprehensive correlation table of all correlations/couplings seen in MeOD.