# Application of the DP4 Probability Method to Flexible Cyclic Peptides with Multiple Independent Stereocenters: The True Structure of Cyclocinamide A

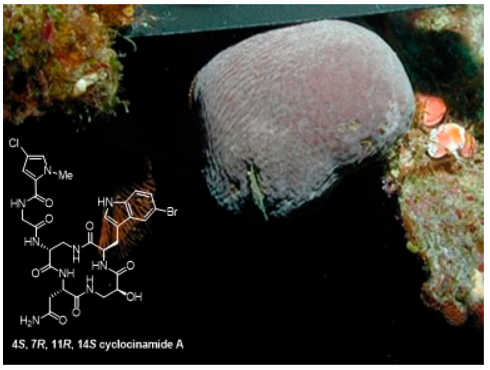
Jason K. Cooper,[†] Kelin Li,[‡] Jeffrey Aubé,[‡] David A. Coppage,[§] and Joseph P. Konopelski*[,§]

[†]Joint Center for Artificial Photosynthesis and Chemical Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[‡]Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

[§]Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, United States

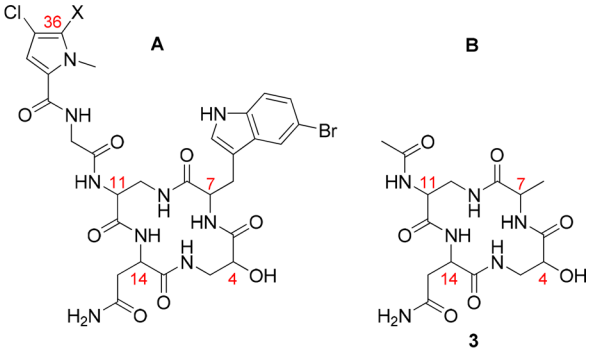
**ABSTRACT:** A DP4 protocol has been successfully utilized to establish the true structure of the natural product cyclocinamide A, a flexible cyclic peptide with four isolated stereocenters. Benchmarking the necessary level of theory required to successfully predict the NMR spectra of three previously synthesized isomers of cyclocinamide A led to the prediction of the natural stereochemistry as 4*S*, 7*R*, 11*R*, 14*S*, which has been confirmed by total synthesis.




In 1997, Crews and co-workers described the isolation, structure, and biological activity (solid tumor activity in a disk diffusion assay) of the cyclic peptide cyclocinamide A (CC-A).[1] This initial work, together with a second publication,[2] proposed the all-*S* stereoisomer (**1a**, Figure 1A) for this natural product constructed of $\alpha$- and $\beta$-amino acids. Initial synthetic work, which did not benefit from the stereochemical claims made in the second CC-A publication, produced the 4*R*, 7*S*, 11*R*, 14*S* (**1b**)[3] and 4*R*, 7*S*, 11*S*, 14*S* (**1c**)[4] isomers. However, neither was found to be the natural product. Ireland and co-workers isolated a similar compound that they dubbed cyclocinamide B (CC-B, **2**), which was claimed to differ from its predecessor in two ways: the addition of a second chlorine atom at C36 and defined stereochemistry of 4*S*, 7*R*, 11*S*, 14*R*.[5]

One of us (UCSC) has prepared **1a**,[6] the 4*S*, 7*S*, 11*R*, 14*S* isomer **1d**,[6] **2**,[7] and ***ent*-1b**[7] (i.e., the compound with the core of **2** and the glycine-pyrrole side chain of compounds **1**). None of these compounds corresponded to either natural product as determined by comparison of the $^{1}$H and $^{13}$C NMR data. However, it was determined that CC-A and CC-B were either enantiomers or identical with respect to their stereochemistry by comparison of their respective NMR data.[7]

Herein, we report the assignment of the stereochemistry of cyclocinamide A as **1e** (4*S*, 7*R*, 11*R*, 14*S*) through the extension of the DP4 analysis to this 14-membered,




**1a** (4*S*, 7*S*, 11*S*, 14*S*), X = H), "all-*S*", UCSC synthesis
**1b** (4*R*, 7*S*, 11*R*, 14*S*, X = H), "4*R*, 11*R*", Montana State synthesis
**1c** (4*R*, 7*S*, 11*S*, 14*S*, X = H), "4*R*", Wayne State synthesis
**1d** (4*S*, 7*S*, 11*R*, 14*S*, X = H), "11*R*", UCSC synthesis
***ent*-1b** (4*S*, 7*R*, 11*S*, 14*R*, X = H), "7*R*, 14*R*", UCSC synthesis
**1e** (4*S*, 7*R*, 11*R*, 14*S*, X = H), "7*R*, 11*R*", *present synthesis*
**2** (4*S*, 7*R*, 11*S*, 14*R*, X = Cl), UCSC synthesis

**Figure 1.** (A) Cyclocinamide structures **1** and **2**, stereochemistry and shorthand designation; (B) truncated structure of CC-A **3** for computational studies.

**Received:** June 5, 2018
**Published:** July 9, 2018

conformationally flexible, cyclic peptide with four unrelated stereocenters. In addition, we present the total synthesis of this isomer, thus verifying the DP4 prediction.

The failure to find correspondence between the claimed and true structures of natural products by synthetic means has been the experience of many other researchers[8] and has led us to a new strategy. The prediction of $^{1}H$ and $^{13}C$ chemical shifts by purely computational means has advanced greatly in recent years.[9] In addition, the classic problem of natural product identification, that of having one set of experimental NMR data that could be assigned to one of several stereochemically possible structures, has been explored by Smith and Goodman[10] with the development of the DP4 application. Use of this procedure results in the assignment of a probability of identity between a given experimental set of $^{1}H$ and/or $^{13}C$ NMR spectra and each of the corresponding computationally derived data sets for the stereoisomers.[11]

Implementation of this methodology to the present problem, however, was not without its challenges. The stereogenic centers in CC-A are isolated from one another; no useful NOE data beyond the nearest neighbor relationships were obtained, and the four amide bonds effectively isolate each spin system.[1] Furthermore, there was no detectable interaction between the glycine-pyrrole fragment at C11, the tryptophan residue at C7, and the asparagine side chain at C14. In addition, the flexibility of CC-A would require the incorporation of a number of low-energy conformations that contribute to the final calculated spectra.

Structure **3** was employed in the calculations. Such an approach was deemed reasonable due to the lack of interactions between side chains on the macrocycle (*vide supra*). In addition, we had previously argued that evaluation of the chemical shift data rested most reasonably on the ring $sp^3$ carbons and hydrogens, where all the stereochemistry resides, and not on the more conformationally mobile side chains.

The work initially focused on computing the $^{1}H$ and $^{13}C$ chemical shifts of the three isomers most recently prepared: **1a** (all-*S*), **1d** (11*R*), and ***ent*-1b** (7*R*, 14*R*). The computational protocols would be continually refined until the DP4 application could assign a high probability of identity upon comparison of a given experimental set of $^{1}H$ and $^{13}C$ chemical shifts with the corresponding calculated spectra for that isomer when challenged with the calculated spectra of all three isomers. Once this benchmarking task was complete, the same methodology would be applied to all possible stereoisomers of cyclocinamide A for comparison with the experimental shifts of the natural product.[12]

Table 1 gives an overview of the initial conformational search results for each isomer using molecular mechanics calculations (Spartan 10[13]). However, the initial levels of theory and basis sets employed within Gaussian 09[14] used to refine the energy levels and provide the calculated spectral data did not produce adequate results. Specifically, we were unable to obtain correspondence of the experimental vs calculated values for the 7*R*, 14*R* isomer with the DP4 method from our initial calculations. Successful implementation was finally achieved by an additional geometry optimization from the conformations with energies <4 kcal $mol^{-1}$ from the minimum using B3LYP/6-311+G(d,2p), which also included a frequency calculation on the final structures to obtain the Gibbs free energies. As shown in

**Table 1. Computational Data and DP4 Predictions**

| | | minimum energy conformations | | |
|---|---|---|---|---|
| isomer | stereochemistry | molecular mechanics[a] | Gaussian[b] | DP4 protocol predictions[c] (%) |
| **1a** | all-*S* | 105 | 13 | 0.0 |
| ***ent*-1b** | 7*R*, 14*R* | 39 | 8 | 26.3 |
| **1c** | 4*R* | 64 | 16 | 0.0 |
| **1d** | 11*R* | 102 | 19 | 0.3 |
| **1e** | 7*R*, 11*R* | 121 | 15 | **73.2** |
| **1f** | 7*R* | 71 | 8 | 0.2 |
| **1g** | 11*R*, 14*R* | 61 | 9 | 0.0 |
| **1h** | 14*R* | 50 | 7 | 0.0 |

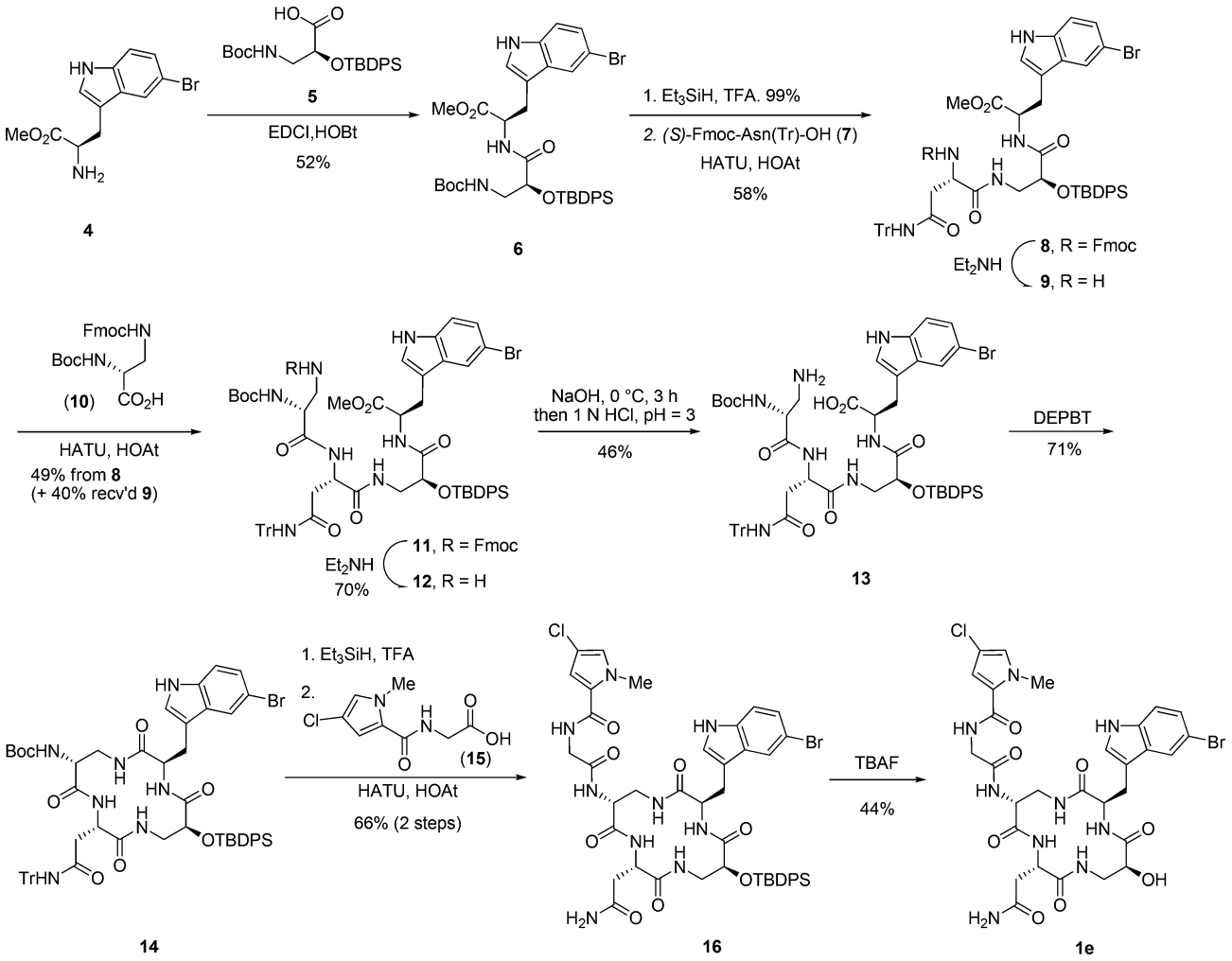
[a]Number of low energy conformations from molecular mechanics calculations. [b]Number of low energy (<4 kcal/$mol^{-1}$) conformations from single-point energy calculations. [c]Probabilities using DP4 applet from comparison of cyclocinamide A (CC-A) experimental NMR data with computational NMR data from all isomers.

Table 1, the number of low energy conformations ranged from 7 for the 14*R* structure to 19 for **1d** (11*R*). These structures were then used to calculate the $^{1}H$ and $^{13}C$ NMR shielding constants with the mPW1PW91/6-311+G(3df,2p) level of theory. To transform the shielding constants to chemical shift data, *N*-methylacetamide in DMSO was employed as the reference, calculated at the same level of theory.[9,15,16] The Gibbs free energy of each conformer was used to assess its contribution to the Boltzmann distribution of structures contributing to the final NMR shielding constants. From these results, 92−99% identity was obtained from the DP4 analysis for each of the three structures; only the core $sp^3$ centers ($^{1}H$ and $^{13}C$ spectra) were used in the comparison (Supporting Information, Tables S1 and S2).[17]

With a confirmed computational method in place, the final DP4 analysis was performed on all eight possible diastereomers of CC-A. Table 1 gives the final results, which identified the 7*R*, 11*R* isomer **1e** as the most probable stereochemical match to the natural product. The only other isomer with an appreciable, but distinctly lower, probability of identity with CC-A was ***ent*-1b**, which the previous synthetic work had shown was not the desired product.

Based on this analysis, we prepared isomer **1e** using an analogous route to that employed in the production of **2** and ***ent*-1b** (Scheme 1). Thus, methyl *R*-5-bromotryptophan **4** was coupled with **5** (prepared as previously reported[6]) to afford dipeptide **6**, which itself was subjected to TFA deprotection and coupled with commercially available (*S*)-Fmoc-Asn(Tr)-OH **7** to provide the 4*S*, 7*S*, 11*R*-tripeptide **8** (CC-A numbering). Additional deprotection and coupling, this time with (*R*)-**10**, provided tetrapeptide **11** in a moderate 49% yield from **8**, but with 40% of recovered **9**, which could be recycled. Carrying out this reaction for an extended time did not lead to improved yields. The conversion of **11** to seco-acid **13** required deprotection of the Fmoc amine and carboxylic acid saponification. The latter reaction proved sensitive, with the best results being obtained from carefully monitored treatment with a base at 0 °C for 3 h; extended exposure of **13** to the base at room temperature led to extensive decomposition. Cyclization was accomplished by treating **13** with DEPBT, affording compound **14** in 71% yield. Reductive deprotection of **14** removed both the Boc and the terminal amide trityl group, setting up amine coupling with carboxylic acid **15**[6] to give hexapeptide **16** in

**Scheme 1. Synthesis of 7*R*, 11*R*-Cyclocinamide A**




66% yield (2 steps). Finally, TBDPS removal by treatment with TBAF gave the target alkaloid **1e** in 44% yield.

The $^{1}H$ and $^{13}C$ NMR spectra of **1e** were compared to the values originally reported by Crews[1] (Table S3). Figures S1 and S2 provide a graphical comparison of the ring $sp^3$ carbons and hydrogens of all four of the recently prepared CC-A isomers (**1a**, ***ent*-1b**, **1d**, and **1e**) to the natural material. Table S4 provides the corresponding $^{13}C$ data in tabular form color-coded for the magnitude of chemical shift variation from that of the natural material; the same comparison of natural CC-B with synthetic **2** is also provided in this table. The close correspondence of the 7*R*, 11*R* isomer spectral data with corresponding natural product data strongly support the assignment of compound **1e** as cyclocinamide A. The most notable difference in the synthetic sample is the doublet appearing at $\delta$ 6.00 ppm, which we assign to the hydroxyl OH group (a likely candidate for exchange in the natural sample).[18] Otherwise, the largest disparity between our synthetic sample and the originally isolated compound is the magnitude of the dextrorotatory specific rotation: $[\alpha]^{21}_{D}$ +102.5 for the synthetic material vs $[\alpha]^{21}_{D}$ +29 for the original isolated sample (*c* 0.1, MeOH for both).

In conclusion, we have succeeded in establishing the absolute stereochemistry and structure of the marine natural product cyclocinamide A. In addition, we have demonstrated that the DP4 protocol can be effectively utilized to guide synthetic target identification on flexible molecules, provided the necessary level of theory for the problem can be ascertained.[19]

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.8b01756.

Details and tables of computational data, comparison of NMR data between final synthetic compound and natural product, full experimental, and NMR spectra (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: joek@ucsc.edu.

### ORCID

Jason K. Cooper: 0000-0002-7953-4229
Jeffrey Aubé: 0000-0003-1049-5767
Joseph P. Konopelski: 0000-0002-7910-8596

## Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Partial support for the synthetic chemistry in the form of start-up funding awarded from the University of North Carolina at Chapel Hill to J.A. is gratefully acknowledged. We also acknowledge the UNC Department of Chemistry Mass Spectrometry Core Laboratory for HR-MS analysis, and Dr. Sarah M. Scarry and Dr. Greg Young (UNC) for assistance with NMR. The work of D.A.S. was supported by grants from the NIH (R01 CA47135 and R25 GM104552). One of us (J.P.K.) would like to offer his sincere thanks to his UC colleagues, Prof. D. Tantillo and Dr. M. Lodewyk (UC Davis), for their many conversations regarding the computational work, and Prof. P. Crews (UCSC) for many fruitful discussions on the subject of this paper and for sharing the underwater image of *Psammocinia* aff. *bulbosa* used in the graphical abstract.

## ■ REFERENCES

(1) Clark, W. D.; Corbett, T. F.; Valeriote, F.; Crews, P. *J. Am. Chem. Soc.* **1997**, *119*, 9285−9286.
(2) Rubio, B. K.; Robinson, S. J.; Avalos, C. E.; Valeriote, F. A.; de Voogd, N. J.; Crews, P. *J. Nat. Prod.* **2008**, *71*, 1475−1478.
(3) Grieco, P. A.; Reilly, M. *Tetrahedron Lett.* **1998**, *39*, 8925−8928.
(4) Liu, L. *Ph.D. Dissertation*; Wayne State University, Detroit, MI, 2002.
(5) Laird, D. W.; LaBarbera, D. V.; Feng, X.; Bugni, T. S.; Harper, M. K.; Ireland, C. M. *J. Nat. Prod.* **2007**, *70*, 741−746.
(6) Garcia, J. M.; Curzon, S. S.; Watts, K. R.; Konopelski, J. P. *Org. Lett.* **2012**, *14*, 2054−2057.
(7) Curzon, S. S.; Garcia, J. M.; Konopelski, J. P. *Tetrahedron Lett.* **2015**, *56*, 2991−2994.
(8) See: Nicolaou, K. C.; Snyder, S. A. *Angew. Chem., Int. Ed.* **2005**, *44*, 1012−1044.
(9) Lodewyk, M. W.; Siebert, M. R.; Tantillo, D. J. *Chem. Rev.* **2012**, *112*, 1839−1862.
(10) Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12946−12959.
(11) For a recent reference to a modified DP4 analysis, denoted DP4+, see: Zanardi, M. M.; Suarez, A. G.; Sarotti, A. M. *J. Org. Chem.* **2017**, *82*, 1873−1879 and references therein.
(12) With four stereogenic centers, 16 possible stereoisomers (8 pairs of enantiomers) are possible. However, since the analysis was by NMR, which is transparent to enantiomers, the determination would be among the 8 unique diastereomers. See Supporting Information for full details.
(13) *Spartan '10*; Wavefunction: Irvine, CA.
(14) *Gaussian 09 v.*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2009.
(15) Fritz, H.; Kristinsson, H.; Mollenkopf, M.; Winkler, T. *Magn. Reson. Chem.* **1990**, *28*, 331−336.
(16) See Supporting Information for all calculation tables.
(17) It should be noted that the synthesis of the all-*S* isomer **1a** was established through the use of Marfey's method and subsequent HPLC analysis, establishing the stereochemical purity of the product. The convergence of experimental and calculated NMR data not only validated this computational approach but also verified that the final cyclizations leading to **1d** and ***ent*-1b** occurred without racemization, thus preserving stereochemical integrity.
(18) The free hydroxyl group in cyclocinamide B is observed at $\delta$ 6.0; see ref 5.
(19) For a different approach to the stereochemical identification of a cyclic peptide, see: Andersson, H.; Demaegdt, H.; Vauquelin, G.; Lindeberg, G.; Karlén, A.; Hallberg, M.; Erdélyi, M.; Hallberg, A. *J. Med. Chem.* **2010**, *53*, 8059−8071.