ORGANIC LETTERS

2006
Vol. 8, No. 17
3861−3864

Attempts to Assemble a Universal NMR Database without Synthesis of NMR Database Compounds

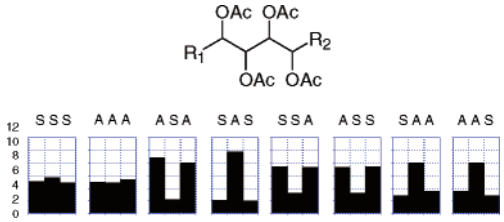
Hirofumi Seike, Indranath Ghosh, and Yoshito Kishi*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

kishi@chemistry.harvard.edu

Received June 27, 2006

ABSTRACT



The feasibility of assembling $^3J_{H,H}$ profiles from NMR data collected on relevant, but not necessarily specific, NMR database compounds representing a given stereocluster was demonstrated. By this approach, the $^3J_{H,H}$ profile was created for the contiguous tetraol peracetate stereocluster. The reliability and applicability of the database thus assembled were demonstrated for known peracetates derived from two heptoses.

Through the work on palytoxins, AAL toxin/fumonisin, and maitotoxin,[1] we experimentally demonstrated that: (1) the spectroscopic profiles of the stereoclusters present in these natural products are inherent to the specific stereochemical arrangement of the small substituents on the carbon backbone and are independent from the rest of the molecules and (2) steric and stereoelectronic interactions between structural clusters connected either directly or with a one-methylene bridge are significant, but interactions between stereoclusters connected with a two- or more-methylene bridge are negligible. On the basis of these experimental observations, we advanced the logic of a universal NMR database approach for stereochemical assignment of (acyclic) compounds. With the use of the ^{13}C chemical shift profiles as the primary means, the feasibility and reliability of this approach were first demonstrated, and then its applicability and usefulness were shown with the stereochemical assignment of the desertmycin/oasomycin class of antibiotics,[2] the mycolactones,[3] tetrafibricin,[4] and glisoprenin A.[5]

The universal NMR database approach of assigning stereochemistry in any given molecule consists of the following steps: (1) identify a stereocluster present in the molecule, (2) compare its NMR profile with the NMR profile of each diastereomer possible for an NMR database compound that adequately represents this stereocluster, and (3) predict its stereochemistry on the basis of the profile fitness of the unknown stereocluster against each diastereomer. In

(1) For the references on palytoxin, AAL toxin/fumonisin, and maitotoxin, see refs 1 and 2, refs 3 and 4, and refs 5 and 6, respectively, cited in: *J. Am. Chem. Soc.* **2003**, *125*, 14379.

(2) (a) Kobayashi, Y.; Lee, J.; Tezuka, K.; Kishi, Y. *Org. Lett.* **1999**, *1*, 2177. (b) Lee, J.; Kobayashi, Y.; Tezuka, K.; Kishi, Y. *Org. Lett.* **1999**, *1*, 2181. (c) Kobayashi, Y.; Tan, C.-H.; Kishi, Y. *Helv. Chim. Acta* **2000**, *83*, 2562. (d) Kobayashi, Y.; Tan, C.-H.; Kishi, Y. *Angew. Chem., Int. Ed.* **2000**, *39*, 4279. (e) Tan, C.-H.; Kobayashi, Y.; Kishi, Y. *Angew. Chem., Int. Ed.* **2000**, *39*, 4282. (f) Kobayashi, Y.; Hayashi, N.; Tan, C.-H.; Kishi, Y. *Org. Lett.* **2001**, *3*, 2245. (g) Hayashi, N.; Kobayashi, Y.; Kishi, Y. *Org. Lett.* **2001**, *3*, 2249. (h) Kobayashi, Y.; Hayashi, N.; Kishi, Y. *Org. Lett.* **2001**, *3*, 2253. (i) Kobayashi, Y.; Tan, C.-H.; Kishi, Y. *J. Am. Chem. Soc.* **2001**, *123*, 2076. (j) Kobayashi, Y.; Hayashi, N.; Kishi, Y. *Org. Lett.* **2002**, *4*, 411.

(3) (a) Benowitz, A. B.; Fidanze, S.; Small, P. L. C.; Kishi, Y. *J. Am. Chem. Soc.* **2001**, *123*, 5128. (b) Fidanze, S.; Song, F.; Szlosek-Pinaud, M.; Small, P. L. C.; Kishi, Y. *J. Am. Chem. Soc.* **2001**, *123*, 10117.

(4) Kobayashi, Y.; Czechtizky, W.; Kishi, Y. *Org. Lett.* **2003**, *5*, 93.

(5) (a) Ghosh, I.; Zeng, H.; Kishi, Y. *Org. Lett.* **2004**, *6*, 4715. (b) Ghosh, I.; Kishi, Y.; Tomoda, H.; Omura, S. *Org. Lett.* **2004**, *6*, 4719. (c) Adams, C. M.; Ghosh, I.; Kishi, Y. *Org. Lett.* **2004**, *6*, 4723.

10.1021/ol061580t CCC: $33.50
Published on Web 07/27/2006

connection with the above steps, we also advanced: (1) the logic and guidelines to identify a stereocluster in a given molecule[2a,6] and (2) the procedure to assemble an NMR database for a given stereocluster.[2a,6] Using the contiguous dipropionate stereocluster[2a,6] and the contiguous polyol stereocluster[7] as examples, Figure 1 outlines the standard

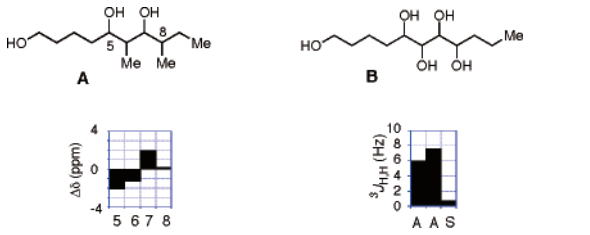


Figure 1. Standard procedure for creating the universal NMR database. **Left panel:** ^{13}C chemical shift database for the contiguous dipropionate. Step 1: select the NMR database compound **A** representing the contiguous dipropionate stereocluster. Step 2: synthesize all the eight diastereomers possible for **A**. Step 3: establish the ^{13}C chemical shift for the relevant ^{13}C nuclei of each diastereomer. Step 4: create the ^{13}C chemical shift profile for each diastereomer as a deviation ($\Delta\delta$) in chemical shift from the average chemical shift for each nucleus. **Right panel:** Vicinal ^{1}H/^{1}H spin-coupling database for the contiguous tetraol. Step 1: select the NMR database compound **B** representing the contiguous tetraol stereocluster. Step 2: synthesize all the eight diastereomers possible for **B**. Step 3: establish the ^{1}H/^{1}H spin-coupling constants for each diastereomer. Step 4: create the $^3J_{H,H}$ profile for each diastereomer with direct use of the observed spin-coupling constants. Abbreviations: A = anti and S = syn.

steps required for creation of a ^{13}C chemical shift database and a vicinal spin-coupling database, respectively.

Among these steps, step 2, which involves "synthesis of all the diastereomers possible for **A** and **B**", requires the most significant laboratory effort, and consequently, it would be beneficial to explore the possibility of eliminating this step from our approach. In this respect, we should specifically point out the difference in procedure between assembling a chemical shift profile and a spin-coupling constant profile. A chemical shift profile is created by using the deviation of chemical shift from the average value of all the diastereomers possible for the NMR database compound **A**. In contrast, a ^{1}H/^{1}H spin-coupling ($^3J_{H,H}$) profile is assembled by directly using experimentally observed values, thereby suggesting the possibility of creating a $^3J_{H,H}$ profile, with the use of the NMR data collected on relevant, but not necessarily specific, NMR database compounds representing a given stereocluster. An appealing implication of this analysis is that, theoretically, a $^3J_{H,H}$ profile can be created simply from spin-coupling data taken directly from relevant literature examples, thus eliminating the most time- and labor-intensive step from the universal database approach.

To test this possibility, the $^3J_{H,H}$ profiles for contiguous tetraol appeared to be a good starting point. Previously, we selected stereocluster **B** (Figure 1), synthesized all the diastereomers possible for **B**, and then used the NMR characteristics observed for each diastereomer to assemble the $^3J_{H,H}$ profiles, as well as the ^{13}C and ^{1}H chemical shift profiles, for the contiguous tetraol stereocluster. Through this work, we demonstrated that spin−spin coupling profiles composed of three contiguous $^3J_{H,H}$'s are, at least to the first order of analysis, sufficient for stereochemical analysis of an unknown polyol.[7]

As a case study, we assembled $^3J_{H,H}$ profiles for the contiguous tetraol stereocluster, using only $^3J_{H,H}$ values obtained from relevant but scattered literature examples (Figure 2). Gratifyingly, we found that the profiles thus

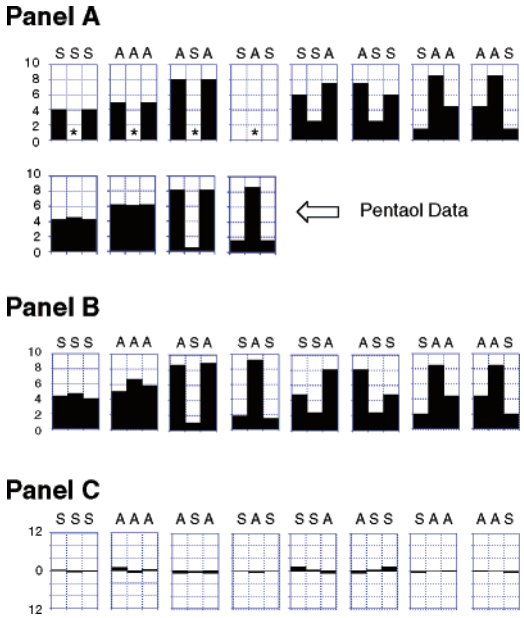


Figure 2. Comparison of the contiguous tetraol $^3J_{H,H}$ profiles assembled via two different approaches. **Panel A.** First line: $^3J_{H,H}$ profiles assembled with the use of the NMR database compound **B**. The $^3J_{H,H}$'s indicated by an asterisk could not be determined from 1D ^{1}H NMR spectra. Second line: $^3J_{H,H}$ profiles assembled with the use of the NMR database pentaol corresponding to **B**. **Panel B:** $^3J_{H,H}$ profiles assembled with the use of spin-coupling values reported in the literature. 3, 3, 11, 6, 14, 14, 5, and 5 literature examples were used to assemble the profile for the SSS, AAA, ASA, SAS, SSA, ASS, SAA, and AAS subgroups, respectively. A complete quotation of the literature is given in the Supporting Information. **Panel C:** The difference between the $^3J_{H,H}$ profiles created via two different approaches. For the SSS, AAA, ASA, and SAS subgroups, the $^3J_{H,H}$ profile of the pentaol was used to estimate the difference between the two $^3J_{H,H}$ profiles. Abbreviations: A = anti and S = syn.

obtained match exceptionally well with the profiles constructed in the previous work.[7] This exercise demonstrates that NMR profiles, at least $^3J_{H,H}$ profiles, can be assembled directly from the NMR characteristics found on relevant, but not necessarily specific, NMR database compounds such as **B** shown in Figure 1. In our view, this should have a significant impact on the universal NMR database approach, as the most time- and labor-intensive part in this approach

(6) Kishi, Y. *Tetrahedron* **2002**, *58*, 6239.
(7) Higashibayashi, S.; Czechtizky, W.; Kobayashi, Y.; Kishi, Y. *J. Am. Chem. Soc*. **2003**, *125*, 14379.

could be eliminated, as long as enough examples are reported in the literature for the stereocluster in question.

Being encouraged by this case study, we chose polyol peracetates as examples, to further study the scope of this approach. Our choice was obviously dictated by the fact that $^3J_{H,H}$ data were readily available in the literature for a variety of different polyacetates. Using all the relevant data reported in the literature, we assembled the $^3J_{H,H}$ profiles consisting of three contiguous $^3J_{H,H}$ constants for all of the eight subgroups (Figure 3). Through this exercise, several important aspects have emerged. First, the deviation in the $^3J_{H,H}$ profile within the same subgroup is negligibly small. Second, if the difference in the two termini was ignored, the SSS, AAA, SAS, and ASA subgroups contain a plane of symmetry. Therefore, we expect that they should exhibit a $^3J_{H,H}$ profile with a symmetrical nature. Indeed, a symmetrical pattern is recognized in their profiles. Third, if the difference in the two termini was ignored, the relative stereochemistry in the AAS and SAA subgroups, as well as in the SSA and ASS subgroups, is the same but arranged in the opposite direction. Therefore, we expect that these pairs should have profiles that are mirror images of each other, and indeed, this is the case. These characteristics support the notion that the $^3J_{H,H}$ profiles thus assembled indeed represent the structural property inherent to each subgroup and, therefore, argue for the possibility of using these $^3J_{H,H}$ profiles to predict the relative stereochemistry of unknown compounds.

However, some subgroups exhibit similar $^3J_{H,H}$ profiles, thereby raising the question of their usefulness for assigning stereochemistry in unknown compounds. To address this issue, we have estimated the difference in profiles projected for all the combinations of subgroups (Figure 3). As described previously, in the universal NMR database approach, we predict the stereochemistry of a given unknown compound on the basis of the degree of fitness of the NMR profile found for the unknown compound with that of the relevant NMR database. For the present case, the $^3J_{H,H}$ profile of an unknown polyol peracetate is experimentally established and then compared with the $^3J_{H,H}$ profile of the eight subgroups. Thus, the magnitude of the profile difference ($\sum|\Delta Hz|$) among the eight subgroups becomes critical to predict the stereochemistry of an unknown compound. To estimate the minimum profile difference required, we calculated the deviation of the observed $^3J_{H,H}$ values from the relevant subgroup profile for each literature example and found that in no case was the total deviation greater than 3.3 Hz. Thus, we assume that a profile difference of 3.3 Hz or greater ($\sum|\Delta Hz| \geq 3.3$ Hz) should be sufficient to differentiate between all the possible combinations of the eight subgroups. With this assumption, the $^3J_{H,H}$ profiles thus assembled should be able to predict the relative stereochemistry of 23 out of the 28 possible combinations of eight subgroups (Figure 4).[8−10]

As demonstrated in the case of contiguous polyols, the $^3J_{H,H}$ profile composed of three $^3J_{H,H}$'s is a powerful tool to analyze the stereochemistry of higher polyols.[7] Applying the reported procedure to the present case, we should be able to predict the relative stereochemistry for all, except five, of the 120 possible combinations for the contiguous pentaol

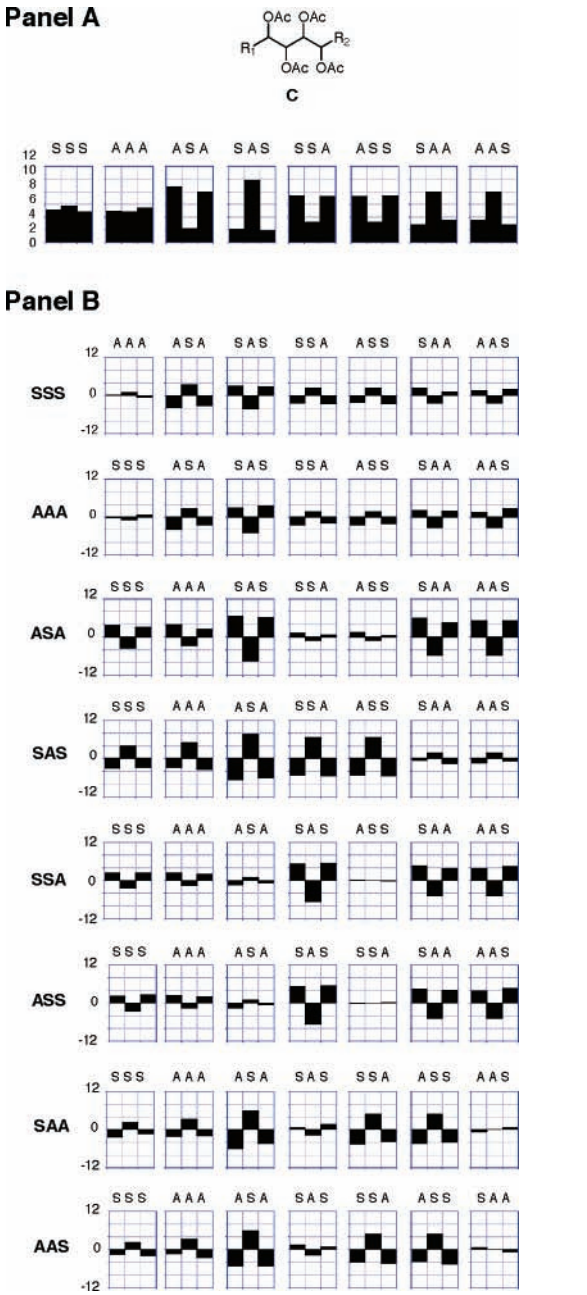


Figure 3. $^3J_{H,H}$ profiles of contiguous tetraol peracetates. **Panel A:** $^3J_{H,H}$ profiles of the contiguous tetraol peracetate stereocluster **C** assembled from the spin-coupling constants of relevant compounds reported in the literature. 3, 3, 5, 22, 7, 7, 4, and 4 literature examples were used to assemble the profile for the SSS, AAA, ASA, SAS, SSA, ASS, SAA, and AAS subgroups, respectively. A complete quotation of the literature is given in the Supporting Information. **Panel B:** The difference in $^3J_{H,H}$ profiles projected for all the possible combinations of the eight subgroups. Note that there are 28 unique combinations, but each unique combination is duplicated in this presentation, for example, SSS/AAA and AAA/SSS. Abbreviations: A = anti and S = syn.

(8) Out of the five indistinguishable combinations, two are in the borderline region, i.e., SSA vs ASA ($\sum|\Delta Hz| = 3.1$ Hz) and ASS vs ASA ($\sum|\Delta Hz| = 3.1$ Hz).[9]

(9) For complete analysis, see Supporting Information.

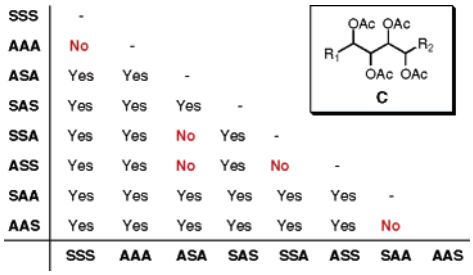


Figure 4. Predicted capacity of the $^3J_{H,H}$ profiles of the contiguous tetraol peracetate **C** to differentiate all the possible combinations of the eight subgroups. With the assumption that the magnitude of the profile difference, $\sum|\Delta Hz| \geq 3.3$ Hz, is sufficient to achieve the task, all, except the five highlighted in red, of the combinations of eight subgroups can be differentiated.[8] Abbreviations: A = anti and S = syn.

peracetates.[11] Interestingly, four out of the five combinations are in the borderline region,[12] but the profile difference for the remaining combination, i.e., the AAAA vs the SSSS subgroup, is too small to predict their stereochemistry. In this connection, however, we should note that the profile difference between the AAA and SSS subgroups is sufficiently large ($\sum|\Delta Hz|$ = ca. 4.4 Hz) in the corresponding contiguous tetraol series. Thus, even for the case of AAA(A) vs SSS(S), the stereochemical differentiation should be possible via $^3J_{H,H}$ profile analysis of the corresponding polyols.

To demonstrate the applicability and usefulness of the $^3J_{H,H}$ profiles reported in this paper for predicting the relative stereochemistry of unknown compounds, we again conducted a case study. On searching the literature, we found the $^3J_{H,H}$'s reported for the peracetates derived from two heptoses.[13] Following the procedure established for the contiguous polyols,[7] we predicted the stereochemistry for **1** and **2** (Figure 5). With the use of the reported data, the overall $^3J_{H,H}$ profiles were prepared and then imaginarily dissected into two $^3J_{H,H}$ profiles, each composed of three $^3J_{H,H}$'s. Each of the resultant profiles was then compared with the $^3J_{H,H}$ profiles shown in Figure 3. This comparison immediately allowed us to conclude the relative stereochemistry of the two peracetates as shown in Figure 5. The stereochemistry thus derived matches the stereochemistry reported for **1** and **2**.

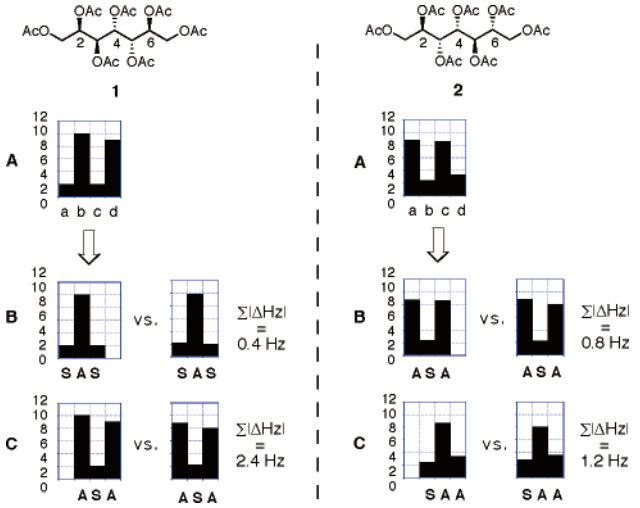


Figure 5. Profile analysis of the $^3J_{H,H}$ coupling constants reported for the peracetate derived from two heptoses. **A**: Overall profiles reported for the peracetates, where a, b, c, and d represent the vicinal spin-coupling constants (Hz) observed for H2/H3, H3/H4, H4/H5, and H5/H6, respectively. **B**: $^3J_{H,H}$ profile composed of the three $^3J_{H,H}$'s of H2/H3−H3/H4−H4/H5; the profile of **1** matches that of the SAS subgroup ($\sum|\Delta Hz| = 0.4$ Hz), whereas the profile of **2** matches that of the ASA subgroup ($\sum|\Delta Hz| = 0.8$ Hz). **C**: $^3J_{H,H}$ profile composed of the three $^3J_{H,H}$'s of H3/H4−H4/H5−H5/H6; the profile of **1** matches that of the ASA subgroup ($\sum|\Delta Hz| = 2.4$ Hz), whereas the profile of **2** matches that of the SAA subgroup ($\sum|\Delta Hz| = 1.2$ Hz).[14] On the basis of this analysis, the relative stereochemistry of **1** and **2** was predicted as indicated. Abbreviations: A = anti and S = syn.

In summary, we have demonstrated the feasibility of assembling $^3J_{H,H}$ profiles from NMR data collected on relevant, but not necessarily specific, NMR database compounds representing a given stereocluster. By this approach, we created the $^3J_{H,H}$ profile for the contiguous tetraol peracetate stereocluster and showed the reliability and applicability for the peracetates derived from two heptoses. Using sagittamide A as an example, we will further demonstrate the usefulness and reliability of this approach in the accompanying paper.[15] In addition, we are studying the possibility of extending this approach to create a chemical shift NMR database.

Acknowledgment. We gratefully acknowledge Eisai Research Institute for financial support.

Supporting Information Available: Literature examples, including structure and observed spin−spin coupling constants, used to create the $^3J_{H,H}$ profiles for contiguous tetraols, tetraol peracetates, and contiguous triol peracetates; analysis to differentiate between various subgroups for the contiguous tetraols and tetraol peracetates and contiguous triol peracetates; and analysis to differentiate between various subgroups for the contiguous pentaol peracetates. This material is available free of charge via the Internet at http://pubs.acs.org.

OL061580T

(10) With the use of the same procedure, the $^3J_{H,H}$ profile was also assembled for the contiguous triol peracetates, and its capacity to discriminate one subgroup from the remaining subgroups was also assessed. Overall, the $^3J_{H,H}$ profile composed with only two $^3J_{H,H}$'s should be effective to predict the relative stereochemistry of five out of the six possible combinations of four subgroups. The details of this analysis are given in Supporting Information.

(11) The details of this analysis are given in Supporting Information.

(12) These are SASS vs SASA ($\sum|\Delta Hz| = 3.1$ Hz), SSAS vs ASAS ($\sum|\Delta Hz| = 3.1$ Hz), SSAA vs ASAA ($\sum|\Delta Hz| = 3.1$ Hz), and AASS vs AASA ($\sum|\Delta Hz| = 3.1$ Hz).[11]

(13) (a) Moore, R. E.; Barchi, J. J., Jr.; Bartolini, G. *J. Org. Chem.* **1985**, *50*, 374. (b) Angyal, S. J.; Saunders, J. K.; Grainger, C. T.; Le Fur, R.; Williams, P. G. *Carbohydr. Res.* **1986**, *150*, 7. (c) Angyal, S. J.; Le Fur, R. *J. Org. Chem.* **1989**, *54*, 1927.

(14) For profile comparison **C** in the heptaacetate **2** series, there was a second candidate (SAS subgroup) noticed. However, it was discarded because its profile fitness ($\sum|\Delta Hz| = 2.9$ Hz) was much worse than the candidate shown in Figure 5.

(15) Seike, H.; Ghosh, I.; Kishi, Y. *Org. Lett.* **2006**, *8*, 3865.