8614 J. Org. Chem. **1997,** *62,* 8614−8615

How Many NOE Derived Restraints Are Necessary for a Reliable Determination of the Relative Configuration of an Organic Compound? Application to a Model System

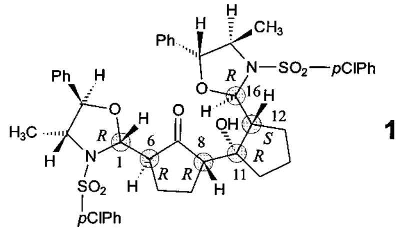
Matthias Köck* and Jochen Junker

Institut für Organische Chemie, Johann Wolfgang Goethe-Universität, Marie-Curie-Str. 11, D-60439 Frankfurt, Germany

Received May 28, 1997 (Revised Manuscript Received September 5, 1997)

The determination of the relative and absolute configuration of natural products is essential to understand their biological activity. In synthetic organic chemistry, the knowledge of the stereochemistry of intermediates is necessary to understand stereoselective reactions. Presently, the standard methods for determining the relative or absolute configuration of organic compounds are X-ray crystallography, chemical synthesis, and NMR spectroscopy. In the past, NMR was usually applied in a qualitative way for the configurational assignment, which is susceptible to errors. Interproton distance information in NMR spectroscopy is obtained from NOE (nuclear Overhauser enhancement) effects. Here, we want to demonstrate that the relative configuration of an organic compound can be determined on the basis of a limited number of NOEs.

The model compound **1** was originally obtained by Hoppe et al.[1] as a pure diastereomer with 10 stereogenic centers.[2] The stereochemistry of the oxazolidines was known prior to the NMR investigations. The remaining six unknown stereogenic centers (gray circles) of **1** have been determined simultaneously by NMR and distance geometry (DG)[4] with a high degree of reliability.[3] In this investigation, 57 NOE cross peaks of **1** were analyzed and used as restraints (r) in the simulations.[3] The simulations were carried out with a combination of DG and distance bounds driven dynamics (DDD)[5] without chiral constraints for the unknown stereogenic centers (floating chirality, fc).[3,6] The floating chirality approach[7] combined with distance geometry (fc-rDG/DDD) allows the stereogenic centers to adopt the configuration according to the experimental data during the simulation.[8,9]



1

To answer the question “How many NOEs are sufficient to define the configuration with a reasonable degree of accuracy?”, several fc-rDG/DDD simulations with a reduced number of NOE derived restraints for **1** were carried out.[9] In all DG simulations 50 structures of **1** were generated. In the following the term *all generated structures* denotes these 50 structures. The structures are ranked by their pseudoenergies,[10] and in usual DG/DDD investigations only the structures that best fit the experimental data (lowest pseudoenergy) are considered. Therefore, the 20 best structures with respect to the experimental data are also discussed.

Out of the 57 experimental NOE restraints, sets of 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, and 55 NOEs were randomly generated. Each of the fc-rDG/DDD simulations for the 11 subdata sets were run 50 times. To obtain a data set without any experimental error, sets of distance restraints for the 57 interproton interactions were also generated from the X-ray structure of model **1**[11] (“synthetic” or “theoretical” NOEs[9,12]). The same procedure was applied as described before.

The results for the 11 experimental subdata sets (5−55 NOEs) are shown in a three-dimensional plot in Figure 1. The simulations with the synthetic NOEs gave a very similar result, and therefore, these are not shown. The three-dimensional representation makes it possible to see all 64 configurations and how they evolve by the number of NOEs. The *z*-axis in Figure 1 gives the ratio of structures with the correct configuration (1*R*,6*R*,8*R*,11*R*,12*S*,16*R*, abbreviated as *RRRRSR*) to all generated structures (RCONF). All generated structures were taken into account in Figure 1a, whereas only those 20 structures with the lowest pseudoenergy were used for Figure 1b. All obtained values of RCONF for the synthetic NOEs are a little bit larger in comparison to the experimental simulation. This was expected since these distances do not have any experimental error.

The calculations with only five NOEs resulted, as expected, in an almost equal distribution over all configurations. But with an increasing number of NOEs the percentage of the correct configuration (*RRRRSR*) compared to the wrong configurations grows (see Figure 1). For 15 and more NOEs the population of the correct configuration is significantly higher than that of any other configuration. When using the best 20 structures (with the lowest pseudoenergy) the correct configuration (*RRRRSR*) is almost twice as populated as the second most frequent configuration already with 10 NOEs. With 20 NOEs about 50% of the generated structures have the correct configuration out of 64 possible configurations, increasing to about 60% for 25 NOEs (see Figure 1b). Using 35 or more NOEs about 90% of the 20 structures have the correct configuration.

A part of Figure 1a is shown in a more detailed way in Figure 2. For 15 NOEs using all generated structures the correct configuration is populated twice (2.2) as much

---

(1) Conde-Friboes, K.; Hoppe, D. *Tetrahedron* **1992**, *48*, 6011.

(2) The sum formula of **1** is $C_{42}H_{44}Cl_2N_2O_8S_2$, resulting in a molecular weight of 839.8.

(3) Reggelin, M.; Köck, M.; Conde-Frieboes, K.; Mierke, D. F. *Angew. Chem.* **1994**, *106*, 822; *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 753.

(4) Crippen, G. M.; Havel, T. F. *Distance Geometry and Molecular Conformation*; Research Studies Press Ltd.: Somerset, England, 1988.

(5) (a) Kaptein, R.; Boelens, R.; Scheek, R. M.; van Gunsteren, W. F. *Biochemistry* **1988**, *27*, 5389. (b) Scheek, R. M.; van Gunsteren, W. F.; Kaptein, R. *Methods Enzymol.* **1989**, *177*, 204.

(6) Mierke, D. F.; Reggelin, M. *J. Org. Chem.* **1992**, *57*, 6365.

(7) (a) Weber, P. L.; Morrison, R.; Hare, D. *J. Mol. Biol.* **1988**, *204*, 483. (b) Holak, T. A.; Gondol, D.; Otlewski, J.; Wilusz, T. *J. Mol. Biol.* **1989**, *210*, 635.

(8) Köck, M.; Junker, J. *J. Mol. Model.* **1997**, *3*, 403.

(9) Since NOEs cannot distinguish between enantiomers, it is only possible to assign the relative configuration of an organic compound. In case of the bisoxazolidine **1** where the configuration of four stereogenic centers (the centers in the oxazolidines) were known prior the NMR analysis, the correct absolute configuration is obtained.

(10) (a) Havel, T. F; Wüthrich, K. *J. Mol. Biol.* **1985**, *182*, 281. (b) Oshiro, C. M.; Thomason, J.; Kuntz, I. D. *Biopolymers* **1991**, *31*, 1049.

(11) The pseudoenergy is not a real energy, it is a quality factor that describes the degree of satisfaction of the given distances and the given chiral volumes with the actual values. The distance term consists of experimental and holonomic restraints. The latter ones are defined by the constitution of the molecule.

(12) Hirschler, J.; Berger, B.; Bolte, M. *Acta Crystallogr.* **1994**, *C50*, 1279.

(13) Havel, T. F.; Wüthrich, K. *Bull. Math. Biol.* **1984**, *46*, 673.

(14) Köck, M.; Junker, J. Unpublished results.

S0022-3263(97)00930-4 CCC: $14.00

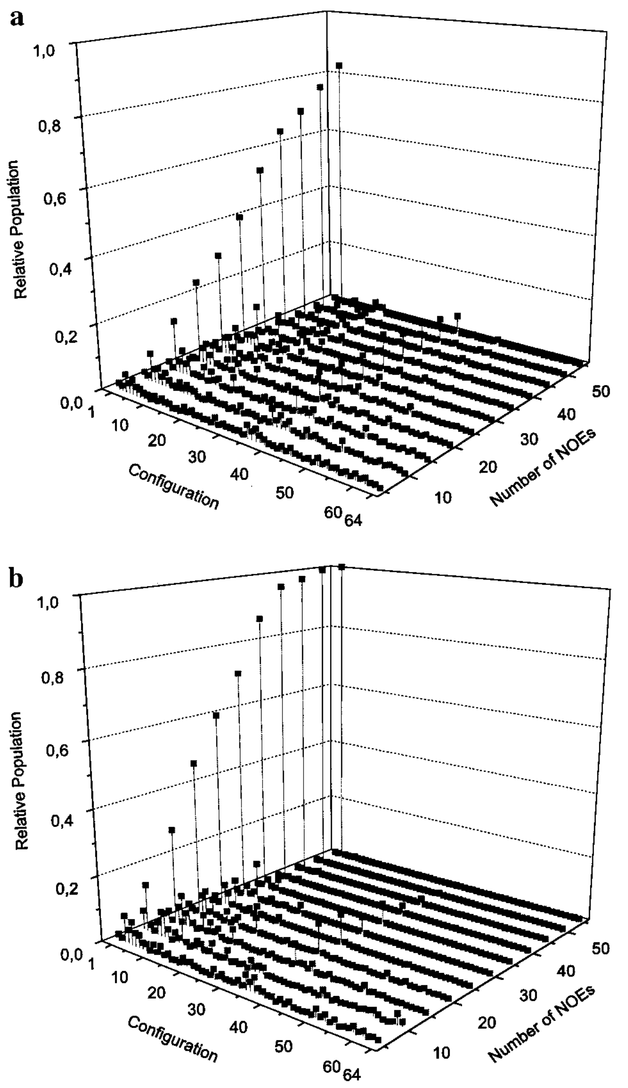


Figure 1. Plot of the fc-rDG/DDD simulations with a random selection of NOEs (5 to 55, in steps of 5) from the 57 experimental NOE restraints (a) for all generated structures and (b) for the 20 structures with the lowest pseudoenergy. The *z*-axis shows the ratio of the correct absolute configuration to all generated structures (relative population; in the text RCONF), the *y*-axis the number of NOEs of the subdata set, and the *x*-axis shows all 64 configurations. The configurations 1−64 are given in a binary order. The *R* configuration denotes zero in the binary code and the *S* configuration denotes 1. The digit to the right counts one, the one to the left 32. To the result the number 1 has to be added. Configuration 1 is *RRRRRR* (0 + 1 = 1), configuration 3 is *RRRRSR* (2 + 1 = 3) and configuration 64 is *SSSSSS* (63 + 1 = 64). The filled squares represent the average value over all 50 simulations.

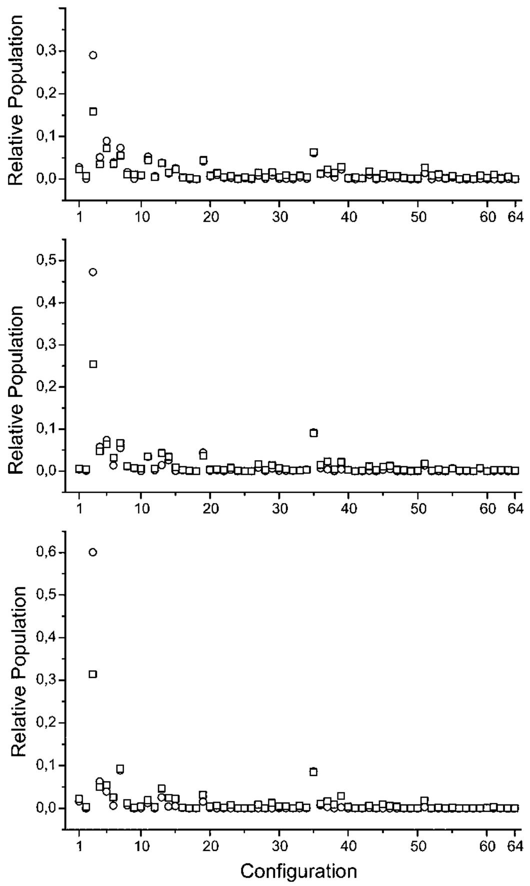


Figure 2. Part of Figure 1 that shows the results with 15 (top), 20 (middle), and 25 (bottom) experimental distances for all generated structures (squares) and for the best 20 structures (circles).

as the second most frequent population. Using 20 NOEs the ratio of RCONF increases to 2.8 times and with 25 NOEs to 3.4 times (see Figure 2). For 25 NOEs about 31% of structures have the correct absolute configuration. The next most populated structures have the configuration *RRRSSR* and *SRRRSR*, which are populated by 9% and 8%, respectively. Both structures differ only by the configuration of one stereogenic center (centers 11 and 1, respectively) in comparison to the correct configuration (*RRRRSR*). These results indicate that the number of NOEs for a reliable determination of the configuration can be about 30−40% of the 57 experimental ones. From a spectroscopic point of view usually the longer distances would disappear earlier in NOE measurements due to signal-to-noise restrictions. The shorter distances have been shown to have a larger influence on the assignment of the configuration.[13] Since in the demonstrated approach all distances have the same probability to disappear, the results are expected to be even better in a real case with a limited number of NOEs.

In summary, the fc-rDG/DDD simulations are a reliable method for the determination of the relative configuration of the bisoxazolidine **1**. Even with only 20 NOE derived restraints the relative configuration could be determined with a high degree of reliability. This result is confirmed with the theoretical NOE data set, which allows us to verify the fc-rDG/DDD method without experimental errors. But it has to be mentioned that the results may be different for other organic compounds. To allow a more general comment the importance of the NOEs has to be studied. One possibility would be to analyze a sphere around every proton and judge the different contributions of interproton distances. It will be of major importance that the NOE contacts are distributed over the sphere. Talking in two dimensions, this mean that there are NOEs in both directions (“push” and “pull” NOEs).

Acknowledgment. This work was supported by the Fonds der Chemischen Industrie (Li 149/4 and Sachkostenzuschup) and the Deutsche Forschungsgemeinschaft (Ko 1314/3-1 and 3-2). We acknowledge the support of Prof. Dr. Christian Griesinger. We thank Dr. Ruud M. Scheek (University of Groningen, The Netherlands) for providing us his DG program package and Dr. John P. Marino for careful reading of the manuscript.

Supporting Information Available: Experimental details and Figure 3 (2 pages).

JO970930+