

Determination of the Relative Configuration of Organic Compounds Using NMR and DG: A Systematic Approach for a Model System[§]

Matthias Köck* and Jochen Junker

Institut für Organische Chemie, Johann Wolfgang Goethe-Universität, Marie-Curie-Str. 11, 60439 Frankfurt, Germany;
Phone: +49-69-798-29143; Fax: +49-69-798-29128; (km@org.chemie.uni-frankfurt.de)

Received: 16 June 1997 / Accepted: 29 July 1997 / Published: 9 September 1997

Abstract

The assignment of the configuration of natural products or organic compounds in general is often carried out by X-ray crystallography or chemical synthesis. NMR spectroscopy is often used only in a qualitative way in the context of configuration assignment. This paper will illuminate a systematic investigation of the combined use of NMR spectroscopy and distance geometry (DG) calculations for the determination of the relative configuration of an organic model compound. The NMR/DG approach allows to quantify configurational assignments.

Keywords: NMR spectroscopy, Distance geometry, Floating chirality, Configurational assignment

Introduction

The determination of the relative and absolute configuration of natural products is essential for an understanding of their biological activity or ecological relevance. In synthetic organic chemistry the knowledge of the stereochemistry of the intermediates is necessary to understand stereoselective reactions. Presently, the standard methods for determining the relative or absolute configuration of organic molecules are X-ray crystallography, chemical synthesis and NMR spectroscopy. The first two methods, although powerful, have significant drawbacks. X-ray crystallography requires crystalline products whereas the chemical synthesis is usually very time consuming and by no means reliable. Here we dis-

cuss how effective NOE (nuclear Overhauser enhancement) derived distance restraints in combination with computational methods can be used for configurational assignments.

Methods

NOE derived distance restraints may be used in a qualitative way or as restraints in EM or MD simulations. Both approaches are problematic for compounds with a large number of unknown stereogenic centers. Therefore, a method is required which allows the determination of all unknown centers simultaneously and without the necessity of crystalline products. Distance geometry (DG) [1] in combination with distance bounds driven dynamics (DDD) [2] calculations using interproton distances (NOE restraints) and floating chirality (fc, means no chiral restraints for stereogenic centers) [3] are discussed here as a novel method for the determination of the relative configuration (fc-rDG/DDD).

[§] Presented at the 11. Molecular Modeling Workshop, 6 -7 May 1997, Darmstadt, Germany

* To whom correspondence should be addressed

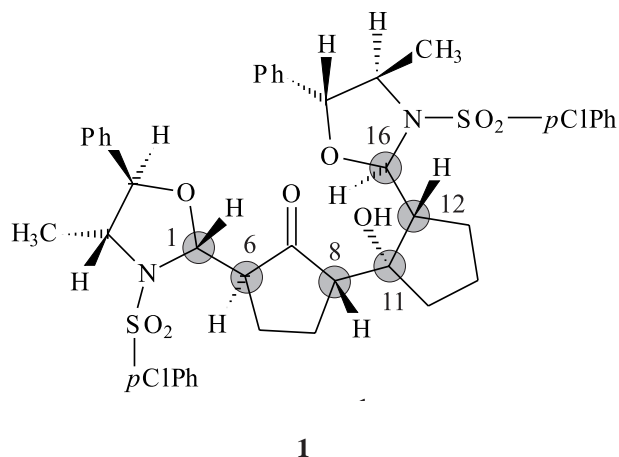


Figure 1. Structural formula of the bisoxazolidine **1**, the six unknown stereogenic centers in **1** are indicated by grey circles. The configurations of the six unknown centers are: 1-R, 6-R, 8-R, 11-R, 12-S and 16-R, abbr. as RRRRSR.

The first complete application of the fc-rDG/DDD method to determine the relative configuration of an organic compound (bisoxazolidine **1** [4], see Figure 1) was described in 1994 [5]. In this investigation all unknown stereogenic centers were allowed to change their sense of chirality during the simulation. The six unknown stereogenic centers of **1** (grey circles, see Figure 1) were determined simultaneously with a high degree of accuracy [5a]. In this investigation 57 NOE

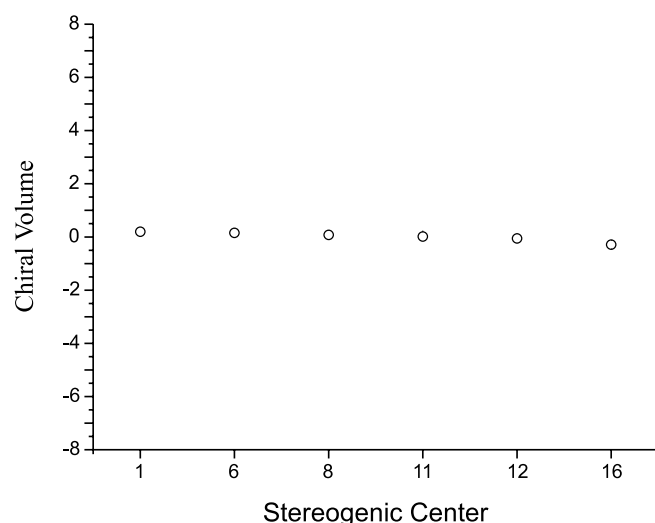


Figure 2. Plot of the average chiral volume of the six unknown stereogenic centers (C-1, C-6, C-8, C-11, C-12 and C-16) of **1** in fc-DG/DDD simulations without distance restraints. The expected chiral volume for a carbon atom is ± 8.5 , the average value found in these simulations is around zero. The standard deviation is too small to be seen in the graph.

cross peaks of **1** were analysed and used as restraints (r) in the simulations [5a]. The first results with the fc-rDG/DDD method indicated that this method could to be a very useful tool for the determination of the relative configuration of organic compounds.

Results

Here we investigate the fc-rDG/DDD method in a systematic way. Four combinatorial possibilities for DG/DDD simulations with floating chirality and NOE restraints (a to d) exist as listed below. Only the fc-rDG/DDD method (a) is of interest for the determination of the relative configuration. The other methods are needed to investigate if the DG/DDD method meets the basic demands of configurational assignments.

- (a) fc-DG/DDD with experimental restraints (fc-rDG/DDD)
- (b) fc-DG/DDD without experimental restraints (fc-DG/DDD)
- (c) DG/DDD without experimental restraints (DG/DDD)
- (d) DG/DDD with experimental restraints (rDG/DDD)

There are two major aspects which have to be investigated for the fc-DG/DDD calculations without experimental distance restraints (b). In these calculations an equal population of the *R* and *S* configuration at every stereogenic center is expected and therefore the averaged chiral volume for each stereogenic center should be zero over all generated structures. Furthermore, all possible absolute configurations (in case of molecule **1** there are 64) should be equally popu-

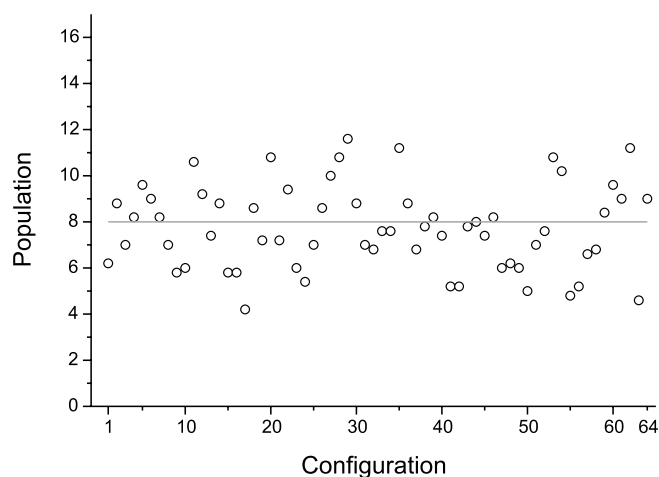


Figure 3. Plot of the population of all 64 configurations obtained from the fc-DG/DDD simulations. The mean value is indicated by a circle. As expected all possible configurations show a similar population.

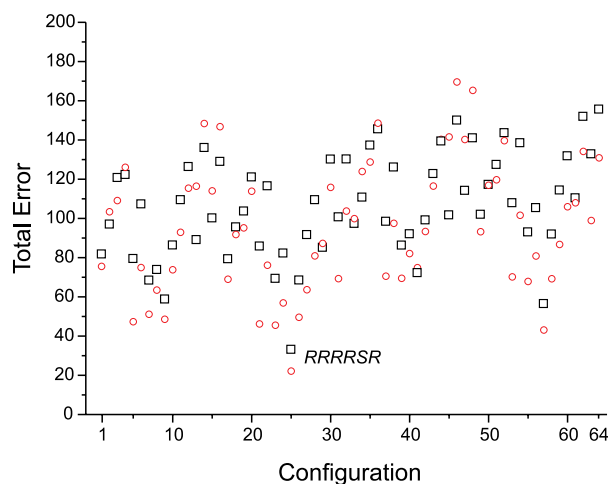
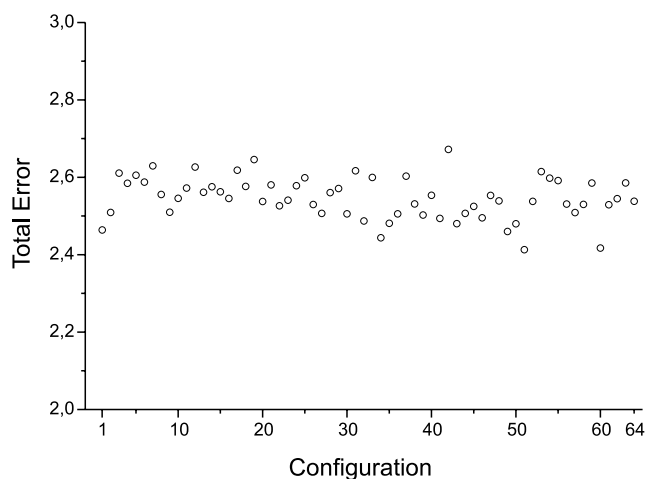


Figure 4. a) Result of the DG/DDD simulations, the average total errors for each configuration is plotted. The values for all configurations are very similar, indicating that DG/DDD was able to fulfill very well the holonomic restraints; calculations do not favor a certain stereoisomer. b) Results of rDG/DDD simulations with experimental (black squares) and theoretical (red circles) distance restraints. All 64 possible configurations of **1** were investigated (on the x-axis). The y-axis shows the total error, averaged over all generated structures. The correct configuration (RRRRSR, number 25) shows by far the smallest total error for the experimental and theoretical restraints.

lated. The fc-DG/DDD calculations were carried out five times with 500 structures. As expected the sum over all chiral volumes for each of the six floating stereogenic centers were close to zero in the fc-DG/DDD calculations (see Figure 2). The 64 different possible absolute configurations were nearly equally populated (see Figure 3). This result was expected since no experimental restraints were applied. It indicates that the results are not biased and that the resulting configurations in fc-rDG/DDD calculations are due to the experimental restraints.

The DG/DDD calculations with chiral restraints (c) were carried out to verify that none of the 64 absolute configurations of **1** is favored by their total errors if no experimental restraints are applied. It is expected that all configurations show similar total errors. The simulations were run with 50 structures for each possible configuration (64) and resulted for all configurations in structures which fulfilled, both the holonomic restraints and the given chiral restraints very well (total errors between 2.4 and 2.7, see Figure 4a). Therefore no configuration is favored by geometrical reasons.

The second fact which had to be investigated for these simulations without floating chirality is if the DG/DDD method favors the correct configuration when NOE restraints are applied. In these calculations (d), the correct configuration should show the lowest total error. The simulations were carried out with the 57 experimental NOEs for all 64 absolute configurations and repeated with the corresponding dis-

tance restraints derived from the X-ray structure (later called synthetic or theoretical NOEs) [a]. The rDG/DDD calculations (d) were carried out with 50 structures for each of the 64 configurations (see Figure 4b). With experimental and "synthetic" data sets the correct configuration (1-*R*, 6-*R*, 8-*R*, 11-*R*, 12-*S* and 16-*R*, abbr. as RRRRSR) shows the smallest violations of the distance restraints and the lowest total error [b] (total error of about 20 for the synthetic data and about 30 for the experimental data). The total errors of the next best configurations were double the lowest. This result demonstrates that the experimental restraints strongly favor the correct configuration.

New simulations with the fc-rDG/DDD method (a) were also carried out for **1**. In the first investigation of **1** from 1994, floating chirality was only applied to the six unknown chiral centers of **1** (see grey circles in Figure 1). There are two options to reduce the degree of determination of the system. One is to increase the number of floating centers during the simulation and the other is to reduce the number of NOEs.

Investigations of the fc-rDG/DDD calculations were carried out with an increased number of floating centers. The application of floating chirality to more centers increases the configurational and conformational freedom and therefore the "noise" of wrong configurations. When floating chirality was applied to four methylene groups (the protons of the fifth are isochronous) beside the six unknown centers it still gave very reliable results. The relative population of the correct configuration (RRRRSR) was about 55%. The next most populated configuration was 12%. The calculations with floating chirality for all ten stereogenic centers (the six unknown stereogenic centers and the four centers with known configuration in the oxazolidines) resulted in a population of the

[a] In order to get the artificial restraints all interproton distances below 8 Å were generated from the X-ray structure. The generated data was compared with the experimental NOEs, and the corresponding distances were used as "synthetic" NOEs.

[b] The pseudoenergy is not a real energy, it is a quality factor which describes the degree of satisfaction of the distances and the chiral volumes. The distance term consists of experimental and holonomic restraints. The latter ones are given by the constitution of the molecule.

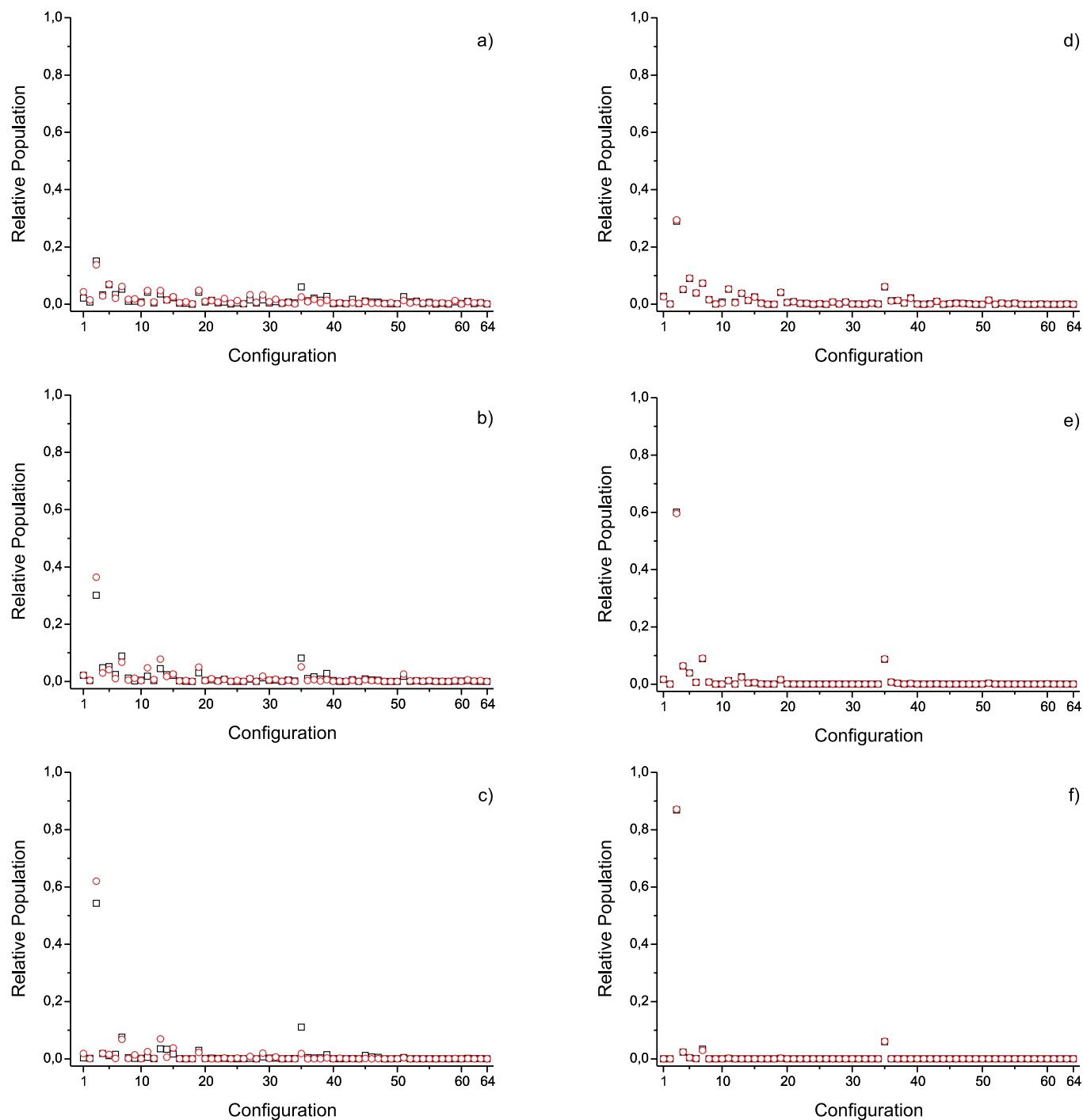


Figure 5. Graphical representation of three of the *fc-rDG/DDD* simulations with a random selection of NOEs out of 57 NOE restraints (experimental and synthetic, respectively). In each graph the results for the experimental (black squares) and the synthetic (red circles) NOE set are shown. The simulations run with 15 (a and d), 25 (b and e) and 35 (c and f) NOEs are exemplary shown. On the x-axis the possible configurations are indicated, on the y-axis the respective relative population is shown. The left graphs show the population of the different configurations in the context of the complete simulation data while on the right side the context is reduced to the best 20 structures.

correct relative configuration (*RRRRSR* and *SSSSRS*) of about 30% (the next configuration was populated by 10%). Even when floating chirality was applied to the four methylene groups and all ten stereogenic centers, the correct relative configuration (*RRRRSR* and *SSSSRS*) was still the most populated. About one-third of the generated structures have the correct relative configuration and both enantiomers were almost equally populated.

Further investigations were carried out to determine the minimal number of restraints necessary to obtain reliable results with the *DG/DDD* method [6, 7]. Out of the 57 ex-

perimental NOE restraints 50 data subsets of 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 and 55 NOEs were extracted randomly. Therefore, all 50 NOE restraint files for each of the 11 data sets are different. Similar sets of distance restraints were also generated from the X-ray structure [8] ("synthetic NOEs") [6, 9]. The upper and lower bounds for the restraints were set to $\pm 10\%$ of the actual distance. The fc-rDG/DDD simulations were run with 50 structures for each sub-data set. The results for three data sets (15, 25 and 35 NOEs) of the fc-rDG/DDD calculations with reduced number of NOEs are shown in Figure 5. For the analysis of this graphical representation first all generated structures (45 to 50) were used (Figure 5a to c) and then only the best 20 (Figure 5d to f). About 42% of structures with the correct absolute configuration are obtained for 25 NOEs, when inspecting all structures. These results indicate that even with a relatively small number of NOEs the configuration of **1** can be determined with a high degree of accuracy. The corresponding results for the "synthetic NOEs" are included in Figure 5 (red circles). The graphical representation shows in principle the same result as for the experimental NOEs (black squares), proving the quality of the experimental data set as well as the conclusion which were drawn from it. When the best 20 generated structures (Figure 5b) are inspected, only a small difference can be observed between the calculations with the experimental and the synthetic NOEs.

Conclusion

The systematic investigation using bisoxazolidine **1** has shown that the fc-rDG/DDD simulations are a reliable method for the determination of the relative configuration. Even with ten floating centers and without the diastereotopic assignment of the methylene groups the calculations favor the correct relative configuration. Furthermore a relatively small number of NOE restraints (20 to 25) was sufficient to determine the correct configuration. This result is confirmed using the theoretical NOE data set which allows the verification of the fc-rDG/DDD method without experimental errors. In the future, it has to be demonstrated that this method can be successfully applied to a wide variety of organic compounds.

Acknowledgment: This work was supported by the Fonds der Chemischen Industrie (Li 149/4) 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 Prof. Dale F. Mierke (University of Worcester, USA) for modified versions of some of the DG programs. We further thank Dr. John P. Marino for careful reading of the manuscript.

References

1. a) Havel, T. F.; Kuntz, I. D.; Crippen, G. M. *Bull. Math. Biol.* **1983**, *45*, 665-720; b) Havel, T. F. *DISGEO*, Quantum Chemistry Exchange Program Exchange No. 507, 1986, Indiana University; c) Crippen, G. M.; Havel, T. F. *Distance Geometry and Molecular Conformation*, Research Studies Press LTD., Somerset, England, **1988**; d) Kuntz, I. D.; Thomason, J. F.; Oshiro, C. M. *Methods Enzymol.* **1989**, *177*, 159-204. e) Havel, T. F. *Prog. Biophys. Molec. Biol.* **1991**, *56*, 43-78.
2. a) Kaptein, R.; Boelens, R.; Scheek, R. M.; van Gunsteren, W. F. *Biochemistry* **1988**, *27*, 5389-5395; b) Scheek, R. M.; van Gunsteren, W. F.; Kaptein, R. *Methods Enzymol.* **1989**, *177*, 204-218.
3. a) Weber, P. L.; Morrison, R.; Hare, D. *J. Mol. Biol.* **1988**, *204*, 483-487; b) Holak, T. A.; Gondol, D.; Otlewski, J.; Wilusz, T. *J. Mol. Biol.* **1989**, *210*, 635-648.
4. Conde-Frieboes, K.; Hoppe, D. *Tetrahedron* **1992**, *48*, 6011-6020.
5. a) Reggelin, M.; Köck, M.; Conde-Frieboes, K.; Mierke, D. F. *Angew. Chem.* **1994**, *106*, 822-824; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 753-755; b) Mierke, D. F.; Reggelin, M. *J. Org. Chem.* **1992**, *57*, 6365-6367.
6. a) Havel, T. F.; Wüthrich, K. *J. Mol. Biol.* **1985**, *182*, 281-294; b) Oshiro, C. M.; Thomason, J.; Kuntz, I. D. *Biopolymers* **1991**, *31*, 1049-1064.
7. Köck, M.; Junker, J., *J. Org. Chem.* accepted.
8. Hirschler, J.; Berger, B.; Bolte, M. *Acta Crystallogr.* **1994**, *C50*, 1279-1281.
9. Havel, T. F.; Wüthrich, K. *Bull. Math. Biol.* **1984**, *46*, 673-698.