

A Test of the Single-Conformation Hypothesis in the Analysis of NMR Data for Small Polar Molecules: A Force Field Comparison

Neysa Nevins,[†] Daniel Cicero,[‡] and James P. Snyder^{*†}

Department of Chemistry, Emory University, 1515 Pierce Drive, Atlanta, Georgia 30322, and Istituto di Ricerche di Biologia Molecolare P. Angeletti, 00040 Pomezia, Italy

Received December 15, 1998

A tricyclic ketone with seven stereogenic centers produced in >95:5 diastereomeric excess by an asymmetric Diels–Alder reaction has been subjected to a careful 2-D NMR(CDCl₃)/MD analysis by Reggelin and co-workers and interpreted in terms of a single conformation in CDCl₃ solution. The present work examines the validity of this interpretation. Specifically, the conformational profile of the endo isomer of the tricyclic ketone was examined by Monte Carlo searches with both the MM2* and MMFF force fields in MacroModel using the GB/SA CHCl₃ solvent continuum model. The two sets of conformations were then combined and optimized with the MM3(96) force field. The structures of the resulting conformations and the NMR-derived geometric variables were together subjected to a NAMFIS analysis (NMR analysis of molecular flexibility in solution). Ideally, the procedure deconvolutes the thermally averaged NMR data into a small family of conformations that optimally represents the *J*-derived torsions and the NOE-derived distances. It is concluded that the tricyclic ketone under study is not well-characterized by a single conformation in CDCl₃ solution, but by a set of rapidly equilibrating conformers that produces a deceptively averaged NMR spectrum. It would appear that the previously suggested structure is a virtual conformation obligated to compress multiconformer features into a single 3-D construct. In addition, we find that current force fields do not uniformly represent the conformational profile of polar molecules such as **1**. The cause is traced primarily to the variable treatment of electrostatic effects.

Introduction

Multidimensional NMR has become a powerful complement to X-ray crystallography for the determination of molecular structure in solution. Among others, the three-dimensional (3-D) architectures of proteins, oligonucleotides, and macromolecular complexes have been convincingly established in aqueous media.¹ The methodology generally involves the measurement of a large number of geometrically determined variables such as nuclear Overhauser effects (NOEs) and coupling constants (*J*s) followed by a search of the target structure in 3-D space constrained by the same variables. For example, starting from a number of points, restrained molecular dynamics (rMD) routinely leads to a common, internally consistent 3-D structure.^{2–4} The tactics are successful, in part, because the macromolecules in question are relatively conformationally immobile. Apart from small subsets of disordered loops, the molecules present essentially a single target conformation.

Peptides or small flexible organic molecules show a fundamentally different conformational profile in solution across a range of temperatures. The compounds are usually characterized by families of rapidly equilibrating conformations characterized by an averaged NMR spectrum. Nonetheless, it has become popular to apply the

macromolecular techniques to small molecules in solvent under the assumption of a single important conformation. Structures reasonably consistent with NOESY, ROESY, and coupling constant data are thereby derived and offered as a meaningful solution to the conformational problem.^{5–10} Ensemble-averaging, in which multiple starting geometries converge to a single conformation or to a small set of viable conformations, helps to prevent biasing.^{11–13} However, multiconformer refinement can lead to overfitting the data (i.e., too many parameters and too little data). To overcome this problem, cross-validation methods can offer a means for evaluating the quality of a fit.¹⁴ For example, a cross-validation procedure implemented with distance geometry/distance-restrained simulated annealing has shown that multiple conformations best fit the NMR data for small proteins,

(5) Cachau, R. E.; Gussio, R.; Beutler, J. A.; Chmurny, G. N.; Hilton, B. D.; Muschik, G. M.; Erickson, J. W. *Int. J. Supercomputer Applic. High Performance Comput.* **1994**, *8*, 24–34.

(6) Paloma, L. G.; Guy, R. K.; Wrasidlo, W.; Nicholaou, K. C. *Chem. Biol.* **1994**, *1*, 107.

(7) Walker, S.; Valentine, K. G.; Kahne, D. *J. Am. Chem. Soc.* **1990**, *112*, 6428.

(8) Walker, S.; Yang, D.; Kahne, D. *J. Am. Chem. Soc.* **1991**, *113*, 4716.

(9) Walker, S.; Murnick, J.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 7954.

(10) Walse, B.; Ullner, M.; Lindbladh, C.; Bulow, L.; Drakenberg, T.; Teleman, O. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 11–22.

(11) Bonvin, A. M.; Boelens, R.; Kaptein, R. *J. Biomol. NMR* **1994**, *4*, 143–149.

(12) Scheek, R. M.; Torda, A. E.; Kemmink, J.; van Gunsteren, W. F. In *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy*; Hoch, J. C., Poulsen, F. M., Redfield, C., Eds.; Plenum Press: New York, 1991; pp 209–217.

(13) Mierke, D. F.; Kurz, M.; Kessler, H. *J. Am. Chem. Soc.* **1994**, *116*, 1042–1049.

(14) Bonvin, A. M. J. J.; Brünger, A. T. *J. Mol. Biol.* **1995**, *250*, 80–93.

[†] Emory University.

[‡] Istituto di Ricerche di Biologia Molecolare P. Angeletti.

(1) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.

(2) Brünger, A. T.; Clore, G. M.; Gronenborn, A. M.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3801.

(3) Kuszewski, J.; Nilges, M.; Brünger, A. T. *J. Biomol. NMR* **1992**, *2*, 33–56.

(4) Powers, R.; Garrett, D. S.; March, C. J.; Frieden, E. A.; Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1993**, *32*, 6744–6762.

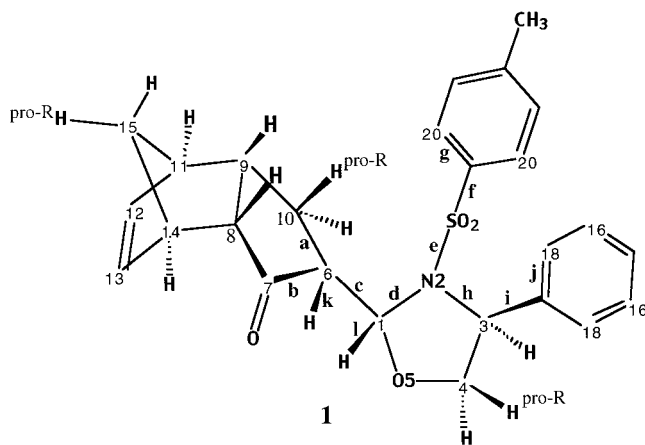


Figure 1. Tricyclic ketone **1** with numeric (for interproton distances) and alphabetical (for dihedral angles) labeling.

including the IgG binding domain of protein G and interleukin 8.^{14,15}

Other approaches for deconvoluting dynamic NMR averages include genetic algorithms,^{16,17} MEDUSA,^{18,19} CPA,²⁰ PEPFLEX-II,²¹ and NAMFIS.²² Genetic algorithms have been used to generate conformations that are consistent with solution NOE distance information.¹⁶ MEDUSA and CPA produce clusters of conformations that likewise represent conformational space by utilizing NOE distance constraints. PEPFLEX-II and NAMFIS operate on structures from a conformational search generated independently of NMR data (e.g., from a force field only). They subsequently seek to extract conformers that best match a combination of NOEs and *J*s.

We have employed the NAMFIS method to analyze the solution conformer populations of tricyclic ketone **1** (see Figure 1), an enantiomerically pure diastereomer synthesized by Diels–Alder cycloaddition in a study of asymmetric induction.²³ The molecule was previously subjected by Reggelin and co-workers to a careful NOESY analysis in CDCl₃ and a subsequent molecular dynamics (MD) search for a *single strongly preferred solution conformation* using the resulting intramolecular distances and coupling constants as constraints.^{24,25} A structure with a diastereomeric configuration and conformation similar to that observed by X-ray crystallography for a closely related structure was derived.²⁶

A molecule such as **1** with four single and easily rotated bonds and a flexible saturated five-membered ring would normally be expected to exist in solution in many

conformational states. Using the NAMFIS methodology, we show that this is indeed the case by applying three popular molecular mechanics procedures to the problem. We also demonstrate that an unambiguous determination of conformer population for **1** cannot, unfortunately, be achieved with the current generation of force fields.

Methods

In the present work, conformational profiles of the endo and exo isomers of **1** were obtained by 1000 step Monte Carlo (MCMM) searches²⁷ with the MM2* and MMFF²⁸ force fields and the GB/SA CHCl₃ solvent continuum model²⁹ in MacroModel.³⁰ For the endo isomer, these searches resulted in 59 and 36 conformations, respectively. The global minima were found 35 and 51 times, respectively, for these force fields. Do these searches represent an essentially complete coverage of the conformational surface available to **1**? Two criteria argue in favor. First, Still and co-workers have defined D_{\min} , the smallest number of times a conformer is found, as an indicator of search efficiency.²⁷ The very effective usage-directed Monte Carlo (MCMM) search methods show satisfactorily high D_{\min} in the 7–15 range. Applications of the related SUMM procedure^{31a} suggest that a D_{\min} of at least 5 for each found conformer is a reasonable indicator of search completeness.^{31b} Second, Saunders has suggested a simple means of estimating the probability that a conformational search is complete: $1 - (1 - (1/N))^M$ where N is the total number of conformers and M is the number of search steps.³² Let us assume that the actual number of conformers for **1** is 100, somewhat higher than we have found. Then, the odds of finding one of the 100 at least once in 1000 search steps is estimated to be $1 - (1 - (1/100))^{1000} = 0.99996$. Alternatively stated, there is a 99.6% chance ($(0.99996)^{100}$) of not missing any of the 100 conformers in a 1000 step search. For a total of only 59 conformers, the odds of not missing a single conformer is essentially 100%. Thus, both the number of conformers located and the number of times the global minima were found provide a high measure of confidence that a thorough search of the low energy regions of conformation space has been achieved.

The MM2* and MMFF sets of conformations were then combined and optimized with MM3(96)³³ to give 84 conformations. For all three force fields, only conformers within 10 kcal/mol of the global minimum were included in the final data set. The adequacy of this energy cutoff is supported by two sets of observations. First, it is a long-standing rule-of-thumb that 3 kcal/mol from the global minimum is the cutoff for empirical isomer observation ($\Delta\Delta G = 3.0$ kcal/mol translates to a 99.4/0.6 isomer ratio at 298 K) and a cap on the conformational energy ligands are willing to pay when bound to the active site of enzymes.³⁴ Thus, a well-parametrized force field should be able to locate experimental conformers within a 3 kcal/mol window. Conformational searching studies with full energy optimization have employed just this cutoff for hydrocarbons.³⁵ Second, for molecules with richer functionality a more conservative strategy is generally adopted. In practice,

(15) Brünger, A. T.; Clore, G. M.; Gronenborn, A. M.; Saffrich, R.; Nilges, M. *Science* **1993**, *261*, 328–331.

(16) Beckers, M. L.; Buydens, L. M.; Pikkemaat, J. A.; Altona, C. J. *Biomol. NMR* **1997**, *9*, 25–34.

(17) Pearlman, D. A. *J. Biomol. NMR* **1996**, *8*, 49–66.

(18) Brüscheiler, R.; Blackledge, M. J.; Ernst, R. R. *J. Biomol. NMR* **1991**, *1*, 3–11.

(19) Blackledge, M. J.; Brüscheiler, R.; Griesinger, C.; Schmidt, J. M.; Xu, P.; Ernst, R. R. *Biochemistry* **1993**, *32*, 10960–10974.

(20) Landis, C.; Allured, V. S. *J. Am. Chem. Soc.* **1991**, *113*, 9493–9499.

(21) Wang, J. J.; Hodges, R. S.; Sykes, B. D. *J. Am. Chem. Soc.* **1995**, *117*, 8627–8634.

(22) Cicero, D. O.; Barbato, G.; Bazzo, R. *J. Am. Chem. Soc.* **1995**, *117*, 1027–1033.

(23) Hoffmann, H.; Bolte, M.; Berger, B.; Hoppe, D. *Tetrahedron Lett.* **1993**, *34*, 6537–6540.

(24) Reggelin, M.; Hoffmann, H.; Kock, M.; Mierke, D. F. *J. Am. Chem. Soc.* **1992**, *114*, 3272–3277.

(25) Mierke, D. F.; Reggelin, M. *J. Org. Chem.* **1992**, *57*, 6365–6367.

(26) Citation 18 in ref 24.

(27) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.

(28) (a) Halgren, T. *J. Comput. Chem.* **1996**, *17*, 490–519. (b) Halgren, T. *J. Comput. Chem.* **1996**, *17*, 520–552. (c) Halgren, T. *J. Comput. Chem.* **1996**, *17*, 553–586. (d) Halgren, T.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587–615. (e) Halgren, T. *J. Comput. Chem.* **1996**, *17*, 616–641.

(29) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

(30) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Kiskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

(31) (a) Goodman, J. M.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110–1117. (b) Cf. footnote 5 of ref 31a.

(32) Saunders, M. *J. Comput. Chem.* **1989**, *10*, 203–208.

(33) Allinger, N. L.; Yuh, Y. H.; Li, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

(34) (a) Boström, J.; Norrby, P.-O.; Liljefors, T. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 383–396. (b) Vieth, M.; Hirst, J. D.; Brooks, C. L., III. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 563–572.

Table 1. Error Estimates Assigned to NOE Distance Categories

| error | NOE distances (x in Å) |
|-----------|---------------------------|
| ± 0.1 | $x < 2.5$ |
| 0.2 | $2.5 \leq x < 3.0$ |
| 0.3 | $3.0 \leq x < 3.5$ |
| 0.4 | $3.5 \leq x < 6.0$ |

conformational searches and filters are performed with energy windows from 3 to 12 kcal/mol above the global minimum.^{35,36} Nonetheless, a number of studies have located the X-ray crystal conformation within 2 kcal/mol of the force field global minimum.³⁷

A NAMFIS²² analysis was carried out for each of these data sets with 50 NOE-determined distances and the seven key $^3J_{\text{HH}}$ coupling constants $J_{(\text{H1-H6})}$, $J_{(\text{H3-H4}^{\text{pro-R}})}$, $J_{(\text{H3-H4}^{\text{pro-S}})}$, $J_{(\text{H6-H10}^{\text{pro-R}})}$, $J_{(\text{H6-H10}^{\text{pro-S}})}$, $J_{(\text{H9-H10}^{\text{pro-R}})}$, $J_{(\text{H9-H10}^{\text{pro-S}})}$ (see Figure 1 for numbering). The remaining 13 reported J s were excluded from the analysis since they are either geminal couplings or fall between protons on the conformationally rigid norbornene bicycle. One distance, that between the tolyl methyl and *o*-phenyl protons, was likewise excluded from the calculation. It is important to recall that a large majority of the included values are averages representing the rapidly equilibrated and unequally populated conformations of the compound under study, structure **1**, in the present case. In practice, NAMFIS operates on a "complete" set of molecular conformations and deconvolutes the thermally averaged NMR variables into a small family of structures that optimally represents the J -derived torsions and the NOE-derived distances. Thus, the method requires the NMR distances, J s, and a complete set of optimized conformations for **1** as input. For each computer-generated conformation, the distances between protons corresponding to experimental NOEs (d_{exp}) are calculated (d_{calc}). Likewise, each conformer is sampled for torsion angles corresponding to the J_{exp} 's. NAMFIS estimates the corresponding J_{calc} 's by means of Haasnoot, Leeuw, and Altona's parameter set *B* for the Karplus equation.³⁸

At this point, one has assembled for each of the conformations in the dataset a set of proton-proton distances and J s that map onto the averaged experimental data. They are subjected to a nonlinear least-squares procedure to yield the best match of the data in terms of a selected set of conformations assigned populations depending on the contribution to the overall fit. Goodness-of-fit is expressed as the sum of square differences (SSD) between measured and modeled variables as shown by eq 1. For m NOE distances, d_{calc} and d_{exp} are the calculated and experimental values, respectively, and d_{err} is the NOE error estimate (see Table 1). For n proton-proton coupling constants ($^3J_{\text{HH}}$), $^3J_{\text{calc}}$ and $^3J_{\text{exp}}$ are the calculated and experimental values, respectively, and w is a weighting factor. For the present NAMFIS analysis, $m = 50$, $n = 7$, and $w = 3$ for all $^3J_{\text{HH}}$'s.

$$\text{SSD} = \frac{1}{2} \left(\sum_{i=1}^m [(d_{\text{calc}}(i) - d_{\text{exp}}(i))/d_{\text{err}}(i)]^2 + \frac{1}{2} \left(\sum_{j=1}^n [w(j) \times (^3J_{\text{calc}}(j) - ^3J_{\text{exp}}(j))/1.5]^2 \right) \right) \quad (1)$$

The E04UPF NAG FORTRAN Library Routine was employed for the minimization of the SSDs according to eq 1

(35) Saunders, M.; Houk, K. N.; Wu, Y.-D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 1419–1427.

(36) (a) Subramanian, G.; Paterlini, M. G.; Larson, D. L.; Portoghese, P. S.; Ferguson, D. M. *J. Med. Chem.* **1998**, *41*, 4777–4789. (b) Alcaro, S.; Coleman, R. S. *J. Org. Chem.* **1998**, *63*, 4620–4625. (c) Kurebayashi, H.; Mine, T.; Harada, K.; Usui, S.; Okajima, T.; Fukazawa, Y. *Tetrahedron* **1998**, 13495–13504. (d) Filizola, M.; Llorens, O.; Carteni-Farina, M.; Perez, J. *Bioorg. Med. Chem.* **1998**, *6*, 1491–1500.

(37) Linden, A.; Beckhaus, H.-D.; Verevkin, S. P.; Ruchardt, C.; Ganguly, B.; Fuchs, B. *J. Org. Chem.* **1988**, *53*, 8205–8211.

(38) Haasnoot, C. G. G.; De Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783–2792.

Table 2. Contributions of NOEs and $^3J_{\text{HH}}$'s to the Sum of Square Difference (SSD) Equation^a for the Endo Isomer of **1**

| | MD | MMFF | MM2* | MM3(96) | combined ^b |
|-----------------------------------|-----|------|------|---------|-----------------------|
| NOEs ^c | 923 | 107 | 171 | 87 | 84 |
| $^3J_{\text{HH}}$'s ^d | 70 | 21 | 18 | 17 | 17 |
| total | 993 | 128 | 189 | 104 | 101 |

^a See eq 1. ^b Combined set of conformations from MMFF, MM2*, and MM3(96). ^c See Table 3. ^d See Table 4.

subject to various constraints.^{39,40} Details are provided in ref 22. Although the constraints obviously include the d_{exp} and $^3J_{\text{exp}}$ values, error limits can also be placed on the NMR data. These additional constraints can be generalized effective weighting factors or specific error estimates for each NMR measurement. The population of conformers (set of mole fractions) that optimally fits both the experimental NMR data and the associated specific errors has been defined a "feasible solution".²² In the present case, a separate error analysis has not been performed for each NMR distance and torsion. Instead, relative error estimates were assigned to the NOE distance categories as given in Table 1. Fitting the experimental and force field generated variables under these conditions has been termed a "best fit solution".²² It represents the level of NAMFIS treatment afforded compound **1** in this work. Roughly speaking, the outcome is 80% probable at the 95% confidence level.

Results and Discussion

NAMFIS Analysis. One primary question is central to our study: Is there more than one predominant conformation for **1** in CDCl₃ solution? If so, what is the nature of the conformational profile? The study proceeded in two steps. First, conformational searches were performed separately for both the endo and exo Diels–Alder adducts. Second, intramolecular proton-proton distances and selected H–C–C–H dihedral angles for collections of conformers obtained from the different force fields were NAMFIS fitted to 50 NOE-derived distances and seven $^3J_{\text{HH}}$'s of **1**.

As a preliminary test of the quantitative abilities of NAMFIS in the present context, we examined the exo isomer. The corresponding MMFF and MM2* datasets yielded a relatively poor fit to the NMR data with overall SSD's of 820 and 879, respectively. The outcome is consistent with the previous NOE/rMD examination, which ruled out the exo isomer as the product of the Diels–Alder reaction.²⁴ By contrast, the endo data set could be satisfied with seven conformations corresponding to overall SSDs of 128, 189, and 104 for the MMFF, MM2*, and MM3(96) force fields, respectively (Tables 2 and 6).

Tables 3 and 4 compare NOE-based proton-proton distances and experimental $^3J_{\text{HH}}$'s with those for the rMD structure and the NAMFIS force field results. The major source of difference between the calculated and measured NOE distances involves the separation of protons on the phenyl subunit and those in the saturated rings, as well as the H-4^{pro-R} - H-8 distance. In all cases, the calculated distances are larger than those derived by NMR (Table

(39) (a) Gill, P. E.; Murray, W.; Wright, M. H. *Practical Optimization*; Academic Press: London and New York, 1981. (b) Hock, W.; Schittkowski, K. *Test Examples for Nonlinear Programming Codes, Lecture Notes in Economics and Mathematical Systems 187*; Springer-Verlag: Berlin and New York, 1981.

(40) For details on the NAG FORTRAN Library see: <http://www.nag.co.uk/numeric/FLOLCH/mk18.html>.

Table 3. Proton-Proton Separations from NMR (NOESY) Analysis, a Single Restrained Molecular Dynamics (rMD) Structure, and Weighted Averages of Families of Conformers Calculated with NAMFIS for the MM2*, MMFF, and MM3(96) Force Fields

| proton | proton | NMR | MD | Δ | MMFF | Δ | MM2* | Δ | MM3(96) | Δ |
|-------------------------|-----------|------|------|----------|------|----------|------|----------|---------|----------|
| H-15pro-S | H-15pro-R | 1.85 | 1.78 | -0.07 | 1.78 | -0.07 | 1.82 | -0.03 | 1.82 | -0.03 |
| H-14 | H-15pro-R | 2.62 | 2.65 | 0.03 | 2.63 | 0.01 | 2.66 | 0.04 | 2.66 | 0.04 |
| H-14 | H-15pro-S | 2.70 | 2.68 | -0.02 | 2.67 | -0.03 | 2.71 | 0.01 | 2.69 | -0.01 |
| H-13 | H-15pro-R | 3.58 | 3.49 | -0.09 | 3.41 | -0.17 | 3.45 | -0.13 | 3.45 | -0.13 |
| H-13 | H-14 | 2.73 | 2.69 | -0.04 | 2.63 | -0.10 | 2.66 | -0.07 | 2.67 | -0.06 |
| H-12 | H-15pro-R | 3.60 | 3.49 | -0.11 | 3.42 | -0.18 | 3.46 | -0.14 | 3.46 | -0.14 |
| H-12 | H-14 | 4.61 | 4.29 | -0.32 | 4.27 | -0.34 | 4.31 | -0.30 | 4.33 | -0.28 |
| H-12 | H-15pro-S | 2.54 | 2.70 | 0.16 | 2.66 | 0.12 | 2.70 | 0.16 | 2.69 | 0.15 |
| H-8 | H-14 | 2.50 | 2.62 | 0.12 | 2.55 | 0.05 | 2.56 | 0.06 | 2.57 | 0.07 |
| H-9 | H-15pro-S | 2.42 | 2.51 | 0.09 | 2.49 | 0.07 | 2.54 | 0.12 | 2.56 | 0.14 |
| H-9 | H-12 | 4.47 | 4.29 | -0.18 | 4.23 | -0.24 | 4.23 | -0.24 | 4.25 | -0.22 |
| H-6 | H-13 | 3.06 | 3.63 | 0.57 | 3.33 | 0.27 | 3.17 | 0.11 | 3.16 | 0.10 |
| H-6 | H-12 | 3.12 | 3.58 | 0.46 | 3.12 | 0.00 | 3.02 | -0.10 | 2.97 | -0.15 |
| H-10pro-R | H-9 | 2.27 | 2.25 | -0.02 | 2.27 | 0.00 | 2.31 | 0.04 | 2.31 | 0.04 |
| H-10pro-R | H-6 | 2.66 | 2.97 | 0.31 | 2.98 | 0.32 | 3.00 | 0.34 | 3.00 | 0.34 |
| H-10pro-S | H-13 | 4.69 | 4.44 | -0.25 | 4.46 | -0.23 | 4.42 | -0.27 | 4.42 | -0.27 |
| H-10pro-S | H-12 | 2.84 | 3.04 | 0.20 | 3.03 | 0.19 | 2.91 | 0.07 | 2.92 | 0.08 |
| H-10pro-S | H-9 | 2.83 | 2.87 | 0.04 | 2.80 | -0.03 | 2.84 | 0.01 | 2.84 | 0.01 |
| H-10pro-S | H-6 | 2.32 | 2.30 | -0.02 | 2.35 | 0.03 | 2.38 | 0.06 | 2.36 | 0.04 |
| H-10pro-S | H-10pro-R | 1.82 | 1.74 | -0.08 | 1.75 | -0.07 | 1.79 | -0.03 | 1.79 | -0.03 |
| H-1 | H-6 | 2.34 | 2.53 | 0.19 | 2.54 | 0.20 | 2.48 | 0.14 | 2.55 | 0.21 |
| H-1 | H-10pro-R | 3.48 | 3.72 | 0.24 | 3.32 | -0.16 | 3.04 | -0.44 | 3.41 | -0.07 |
| H-1 | H-10pro-S | 4.42 | 4.14 | -0.28 | 3.57 | -0.85 | 3.36 | -1.06 | 4.05 | -0.37 |
| H-11 | H-15pro-R | 2.62 | 2.65 | 0.03 | 2.62 | 0.00 | 2.66 | 0.04 | 2.65 | 0.03 |
| H-11 | H-15pro-S | 2.71 | 2.68 | -0.03 | 2.67 | -0.04 | 2.71 | 0.00 | 2.70 | -0.01 |
| H-11 | H-12 | 2.70 | 2.68 | -0.02 | 2.64 | -0.06 | 2.65 | -0.05 | 2.67 | -0.03 |
| H-11 | H-9 | 2.53 | 2.52 | -0.01 | 2.50 | -0.03 | 2.53 | 0.00 | 2.55 | 0.02 |
| H-11 | H-10pro-R | 3.79 | 3.71 | -0.08 | 3.75 | -0.04 | 3.78 | -0.01 | 3.81 | 0.02 |
| H-11 | H-10pro-S | 2.81 | 3.00 | 0.19 | 2.98 | 0.17 | 2.95 | 0.14 | 2.98 | 0.17 |
| H-4pro-R | H-8 | 3.59 | 4.38 | 0.79 | 5.09 | 1.50 | 5.07 | 1.48 | 5.09 | 1.50 |
| H-4pro-R | H-10pro-R | 3.21 | 3.38 | 0.17 | 3.17 | -0.04 | 3.37 | 0.16 | 3.30 | 0.09 |
| H-4pro-R | H-1 | 3.67 | 3.86 | 0.19 | 3.79 | 0.12 | 3.80 | 0.13 | 3.77 | 0.10 |
| H-4pro-S | H-1 | 2.60 | 3.08 | 0.48 | 2.82 | 0.22 | 2.80 | 0.20 | 2.73 | 0.13 |
| H-4pro-S | H-4pro-R | 1.78 | 1.79 | 0.01 | 1.77 | -0.01 | 1.81 | 0.03 | 1.79 | 0.01 |
| H-3 | H-10pro-R | 4.55 | 4.11 | -0.44 | 4.56 | 0.01 | 4.69 | 0.14 | 4.55 | 0.00 |
| H-3 | H-1 | 3.19 | 3.55 | 0.36 | 3.58 | 0.39 | 3.46 | 0.27 | 3.44 | 0.25 |
| H-3 | H-4pro-R | 2.88 | 2.86 | -0.02 | 2.79 | -0.09 | 2.84 | -0.04 | 2.80 | -0.08 |
| H-3 | H-4pro-S | 2.31 | 2.23 | -0.08 | 2.34 | 0.03 | 2.37 | 0.06 | 2.37 | 0.06 |
| H16,16' | H-10pro-R | 3.90 | 4.70 | 0.80 | 4.77 | 0.87 | 4.85 | 0.95 | 4.68 | 0.78 |
| H16,16' | H-10pro-S | 4.18 | 5.55 | 1.37 | 4.78 | 0.60 | 5.30 | 1.12 | 4.59 | 0.41 |
| H16,16' | H-4pro-R | 3.79 | 5.28 | 1.49 | 4.92 | 1.13 | 4.96 | 1.17 | 4.94 | 1.15 |
| H16,16' | H-3 | 3.47 | 5.09 | 1.62 | 5.03 | 1.56 | 5.04 | 1.57 | 5.05 | 1.58 |
| H18,18' | H-10pro-R | 2.89 | 3.32 | 0.43 | 3.11 | 0.22 | 2.84 | -0.05 | 2.95 | 0.06 |
| H18,18' | H-10pro-S | 3.08 | 4.33 | 1.25 | 3.40 | 0.32 | 3.12 | 0.04 | 3.17 | 0.09 |
| H18,18' | H-4pro-R | 2.79 | 3.45 | 0.66 | 2.58 | -0.21 | 2.81 | 0.02 | 2.63 | -0.16 |
| H18,18' | H-3 | 2.82 | 3.04 | 0.22 | 2.92 | 0.10 | 2.83 | 0.01 | 2.85 | 0.03 |
| H20,20' | H-6 | 3.38 | 3.70 | 0.32 | 4.28 | 0.90 | 3.49 | 0.11 | 3.69 | 0.31 |
| H20,20' | H-1 | 2.48 | 3.82 | 1.34 | 2.76 | 0.28 | 3.12 | 0.64 | 2.72 | 0.24 |
| H20,20' | H-4pro-S | 3.86 | 5.88 | 2.02 | 4.01 | 0.15 | 4.09 | 0.23 | 3.85 | -0.01 |
| H20,20' | H-3 | 2.49 | 4.92 | 2.43 | 2.79 | 0.30 | 3.15 | 0.66 | 2.77 | 0.28 |
| raw SSD ^a | | | | 23.9 | | 9.9 | | 11.1 | | 7.9 |
| NAMFIS SSD ^b | | | | 923 | | 107 | | 171 | | 87 |

^a "Raw SSD" is the sum of square differences between calculated and NMR. derived NOE distances, i.e., no error (d_{err}) weighting.

^b "NAMFIS SSD" is the sum of square differences calculated from the first part of eq 1.

3). We have no completely satisfactory explanation for these discrepancies.

Table 5 reports the mole fractions (MF) predicted by NAMFIS for the three most populated conformations of **1** within each of the three force field sets. Each trio of conformers accounts for 69–80% of the total solution population. The remaining 20–31% is found in four additional conformations projected to contribute 0.03–0.10 to the overall mole fractions (Table 6). Interestingly, although one conformer appears in all three NAMFIS solutions for the three force fields, most of the seven conformers arising from one force field are not found in the NAMFIS set by the other two. The most important common conformer for all three solutions is the highest

populated conformer for MM2* (39%) and MM3(96) (47%) and a lower populated conformer for MMFF (2%) (Table 6).

Although all three force fields present a different group of conformers as "best fits" to the NMR quantities, the separate sets are nonetheless responsible for the relatively low SSD's recorded in Tables 2–4. Pooling the MM2*, MMFF and MM3(96) conformers provides a combined set of 179 structures. NAMFIS treatment of the latter furnishes an SSD slightly lower but comparable to that obtained from the individual force field samples (Table 2). Of the seven structures obtained from the combined set, five are from the MM3(96) set and comprise 84% of the total derived conformer population. In addi-

Table 4. Selected Proton Coupling Constants ($^3J_{\text{HH}}$'s) Determined from NMR and NAMFIS for the Endo Conformer of **1**

| | NMR | "rMD" ^a | MMFF | MM2* | MM3(96) | combined |
|------------------------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|
| coupling constant | $^3J_{\text{HH}}$ | $^3J_{\text{HH}}$ | $^3J_{\text{HH}}$ | $^3J_{\text{HH}}$ | $^3J_{\text{HH}}$ | $^3J_{\text{HH}}$ |
| $J_{\text{H1-H6}}$ | 2.9 | 3.6 | 2.0 | 2.8 | 2.4 | 2.3 |
| $J_{\text{H3-H4}^{\text{pro-R}}}$ | 2.4 | 1.7 | 2.4 | 2.8 | 2.5 | 2.6 |
| $J_{\text{H3-H4}^{\text{pro-S}}}$ | 6.7 | 7.0 | 6.4 | 6.3 | 6.2 | 6.3 |
| $J_{\text{H6-H10}^{\text{pro-R}}}$ | 10.0 | 6.0 | 8.8 | 8.9 | 8.9 | 8.9 |
| $J_{\text{H6-H10}^{\text{pro-S}}}$ | 10.4 | 9.9 | 8.8 | 8.8 | 8.9 | 9 |
| $J_{\text{H9-H10}^{\text{pro-R}}}$ | 10.0 | 10.3 | 10.0 | 10.1 | 10.1 | 10.2 |
| $J_{\text{H9-H10}^{\text{pro-S}}}$ | 2.4 | 2.5 | 1.9 | 1.7 | 1.8 | 1.7 |
| raw SSD ^b | | 17 | 5.2 | 4.6 | 4.3 | 4.3 |
| NAMFIS SSD ^c | | 70 | 21 | 18 | 17 | 17 |

^a Structure obtained by MMFF optimization with all of the rMD constraints (ref 24) in place. ^b "Raw SSD" is the sum of square differences between calculated and NMR $^3J_{\text{HH}}$'s. ^c "NAMFIS SSD" is the sum of square differences calculated from the second part of eq 1.

tion, there is one each from MM2* (14%) and MMFF (1%). As illustrated by Table 5, NAMFIS estimation of conformer mole fraction does not generally identify low energy conformations as dominant in the derived populations. In fact, the 10 lowest energy conformations from any of the force field searches lead to less satisfying correlations of the NMR data with overall SSD's of 155, 248, and 165 for MMFF, MM2*, and MM3(96), respectively (compare Table 2).

Force Field Dependent Conformer Energies. The NAMFIS mole fractions and corresponding force field energies relative to individual global minima are also given in Table 5. Each force field delivers a different global minimum. It is noteworthy that conformers that contribute significantly to the NAMFIS best solution are often calculated to be much higher in energy than the lowest from the different force fields. We attribute this largely to the fact that compound **1** contains eight polar bonds. Hence, a considerable number of intramolecular dipole-dipole interactions between atoms separated by more than three bonds are present. Since polarization effects are absent from the force fields and long-range dipole effects are not incorporated in the underlying molecular mechanics parameters, the relative energies of conformations of **1** in the force field framework are inconsistent and most likely inaccurate. The inconsistency arises from the fact that the main energy contribution to the total steric energies is the electrostatic term in the force field equation, a term constituted differently for each of the three force fields used in this study. MMFF employs bond charge increment parameters derived mainly from HF/6-31G* dipole moments. MM3(96) uses a bond dipole approach in which experimental dipole moments were used for parametrization. MM2*, a modified version of the earlier MM2 force field, substitutes a partial charge scheme for the original bond dipole equations. In unfavorable cases, the latter substitution can lead to large errors due to parameter inconsistency.

Thus, across a family of conformations for a given polar molecule, different force fields can be expected to provide not only a variable estimate of relative energies for the same set of conformations but also a distinct set of geometries. The point is reinforced by inspection of the global minima cited in Table 6. MM3(96) and MM2* deliver essentially the same global minima, but differ qualitatively from MMFF in torsions def and efg. Fur-

thermore, each of the force fields used in this study deliver a different set of three dominant conformations as the best NAMFIS fit to the empirical data (Tables 5 and 6). The differential long-range electrostatic interactions from force field to force field serve to selectively stabilize certain conformations and destabilize others. In the same spirit, certain conformations that exist as local minima in one framework are metastable in another. Consequently, the existence of local minima, the distribution of conformers within the energy manifold, and the identity of the global minimum are unfortunately force field dependent for the highly polar species **1**.

It is possible that low quality or missing force field parameters contribute to the uneven performance of the force fields compared. For example, with respect to **1**, MM2* operates with four tosyl and seven oxazolidine-related torsions of low quality. MMFF in Macromodel carries 12 low quality and 2 medium quality parameters in the same region of the molecule. MM3(96) fares best with only six similar low quality torsions. Nonetheless, similar observations of poor performance have been made for the abilities of MM2* and MM3* to represent the relative conformational energies of chair and twist-boat forms of substituted dioxanes.⁴¹ In this case, electrostatic energy is likewise the single most important component of the conformational energies. Reparametrization of AMBER with restrained electrostatic-potential-based atomic charges and MP2/6-31G* energies corrects the discrepancies and provides a general molecular mechanical model for 1,3-dioxoalkanes.⁴²

In the context of the present problem, it would appear that MMFF is the superior method. From Table 6 (2nd and 11th columns), it can be seen that MMFF explains 68% (MF = 0.68) of the data via four conformations with energies ≤ 3.3 kcal/mol. The low energies obtained with Becke3LYP density functional theory (see below) in Table 5 are confirmatory. MM2* by comparison accounts for 55 and 65% of the MF at energies of ≤ 2.2 and < 3.9 kcal/mol, respectively. The Becke3LYP energies are, nonetheless, considerably higher. MM3(96) falls in third place with 32% of the MF accommodated at ≤ 4.1 kcal/mol. The moderate success with which MMFF matches NAMFIS mole fractions against low energy conformers may be due to its representation of molecular electrostatics with uniformly applied electrostatic potential atomic charges^{28b} similar to AMBER⁴² and OPLS.

In molecules with only a few long-range dipolar interactions, inclusion of solvent effects can sometimes help to ameliorate excessive electrostatic interactions. For the MM2* and MMFF conformational searches of adduct **1**, we have employed the GB/SA CHCl₃ continuum solvation model using charges generated by Macromodel. As indicated by the relative energies in Table 5, this expedient does not improve matters by comparison to the empirically based NAMFIS solution.

DFT Calculations. To obtain an independent energy assessment, single-point density functional theory (DFT) calculations using the Becke3LYP/6-31G* formalism were performed in Gaussian 94⁴³ for the three dominant conformers of each force field (Table 5). Energy evaluation of the MMFF optimized structures results in an

(41) Rychovsky, S. D.; Yang, G.; Powers, J. P. *J. Org. Chem.* **1993**, *58*, 5251.

(42) Howard, A. E.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 243-261.

Table 5. Mole Fractions (MF)^a and Energies (ΔE)^b of Major Conformers That Comprise the NAMFIS “Best Solution” Fit to NMR-Determined NOEs and ³J_{HHS}

| no. ^d | MMFF ^c | | | no. ^e | MM2* ^c | | | no. ^f | MM3(96) | | |
|------------------|-------------------|------------------------|---------------------------|------------------|-------------------|------------------------|---------------------------|------------------|---------|------------------------|---------------------------|
| | MF | ΔE^{FF} | ΔE^{B3LYP} | | MF | ΔE^{FF} | ΔE^{B3LYP} | | MF | ΔE^{FF} | ΔE^{B3LYP} |
| 1 | 0.23 | 0.0 | 0.0 | 18 | 0.39 | 1.2 | 12.6 | 29 | 0.13 | 3.7 | +2.3 |
| 14 | 0.37 | 0.8 | -0.5 | 40 | 0.10 | 5.1 | 5.0 | 35 | 0.19 | 4.1 | -3.3 |
| 35 | 0.20 | 9.9 | 3.7 | 58 | 0.20 | 8.1 | 8.2 | 45 | 0.47 | 4.8 | -3.5 |

^a MF is the mole fraction of each conformer within the family that comprises the “best solution”. ^b ΔE^{FF} (force field) and ΔE^{B3LYP} (single-point Becke3LYP/6-31G*) are the energies (kcal/mol) with respect to the conformer in each force field that is the global minimum. ^c GB/SA CHCl₃ solvent continuum model employed in conformational search. ^d Conformer number in increasing energy order for the MMFF force field (36 total). ^e Conformer number in increasing energy order for the MM2* force field (59 total). ^f Conformer number in increasing energy order for the MM3(96) force field (84 total).

Table 6. Tosyl/Phenyl (def, egf, hij) and Five-Membered Rings (kcl, dck, acd, abd) Torsions for the rMD Structure of Reggelin et al.,^a an X-ray Structure,^b the MMFF rMD Structure,^c and NAMFIS-Fitted Conformations (see Figure 1)

| | MF ^d | torsions | | | | | | | conf no. ^e | rel energy ^e |
|--------------------|-------------------|----------|------|------|------|------|------|------|-----------------------|-------------------------|
| | | def | efg | hij | kcl | dck | acd | bcd | | |
| rMD ^a | | | | | -44 | 76 | -47 | -164 | | |
| X-ray ^b | | | | | -65 | 53 | -65 | 175 | | |
| “rMD” ^c | | 67 | 59 | -67 | -48 | 78 | -48 | -164 | | |
| MMFF | 0.37 | -69 | 92 | -19 | -71 | 54 | -71 | 172 | 5 | 0.8 |
| | 0.23 ^f | -68 | 93 | -79 | -72 | 52 | -74 | 169 | 1 | 0.0 |
| | 0.20 | -68 | 90 | -28 | 54 | 176 | 55 | -66 | 35 | 9.9 |
| | 0.08 | 67 | -105 | -152 | -71 | 54 | -73 | 172 | 17 | 5.2 |
| | 0.06 | -66 | 93 | 96 | -69 | 56 | -71 | 175 | 7 | 1.3 |
| | 0.03 | -68 | 87 | -10 | -170 | -44 | -166 | 78 | 31 | 9.0 |
| | 0.02 | -69 | -92 | -151 | -69 | 55 | -71 | 174 | 15 | 3.3 |
| MM2* | 0.39 | -47 | -63 | -172 | -58 | 65 | -60 | -180 | 18 | 1.2 |
| | 0.20 | 45 | 97 | 160 | 50 | 176 | 56 | -68 | 58 | 8.1 |
| | 0.10 | -53 | -46 | 177 | 75 | -159 | 82 | -44 | 40 | 5.1 |
| | 0.10 | -46 | 131 | -149 | 79 | -156 | 85 | -39 | 20 | 2.2 |
| | 0.10 | -39 | -65 | -141 | -57 | 67 | -59 | -177 | 28 | 3.9 |
| | 0.07 | 136 | 77 | -14 | 52 | 178 | 58 | -67 | 50 | 6.4 |
| | 0.06 | -135 | 55 | 85 | -53 | 71 | -54 | -172 | 11 | 0.9 |
| | 0.00 ^f | 86 | -116 | -109 | -57 | 66 | -58 | -179 | 1 | 0.0 |
| MM3(96) | 0.47 | -67 | -83 | 166 | -68 | 57 | -65 | 175 | 45 | 4.8 |
| | 0.19 | -66 | 83 | 99 | -69 | 55 | -67 | 172 | 35 | 4.1 |
| | 0.13 | 143 | -72 | 48 | -54 | 70 | -51 | -171 | 29 | 3.7 |
| | 0.08 | -95 | 46 | 82 | -50 | 76 | -45 | -166 | 80 | 8.8 |
| | 0.07 | -66 | -86 | 174 | -172 | -47 | -169 | 73 | 53 | 5.4 |
| | 0.03 | 98 | 69 | -108 | -52 | 74 | -48 | -168 | 46 | 4.9 |
| | 0.03 | -37 | -76 | 0 | -58 | 67 | -54 | -175 | 79 | 8.7 |
| | 0.00 ^f | 102 | -97 | -110 | -56 | 71 | -51 | -171 | 1 | 0.0 |

^a Reference 24. ^b Reference 26. ^c Structure obtained by MMFF optimization with all of the rMD constraints (see ref 24) in place. ^d MF is the conformer mole fraction. ^e Conformer number (no.) and relative energies (E_{rel}) for a given FF family ranked from the global minimum (1) to highest energy (36, 59, and 84, respectively; see footnotes d–f in Table 5). ^f The global minimum for the designated force field.

apparent correlation between low calculated conformer energies (both force field and DFT) and the highest NAMFIS mole fractions for two of the three key conformations. Neither MM2* nor MM3(96) yields such a three-way correspondence. The simplest interpretation of the energy results in Table 5 is that neither the force fields nor DFT are able to predict the experimentally based NAMFIS outcome.

We recognize that the single-point DFT furnish an energy scaling that is clearly dependent on the nature of the internal coordinates predicted by an individual force field. While full DFT optimization would remove the energy discrepancies between force fields, single point energies within a given conformational family (e.g.,

MMFF) can be regarded as semiquantitatively reliable. The situation is parallel to the practice of optimizing molecular geometries with one *ab initio* basis set and calculating the energy at another.⁴⁴

NAMFIS Bypasses the Energy Catastrophe. The conformer energy comparisons emphasize an important element of a NAMFIS-like deconvolution analysis. Namely, explicit force-field energies are ignored during the fitting exercise. In this way, artifacts arising from energy miscalculation introduced by long-range interactions among polar functionalities are sidestepped. Accordingly, a large energy window is defined (e.g., 10 kcal/mol) within which all conformations generated by the search are retained. Assuming that geometries across the conformational potential energy surface are well represented, the NAMFIS treatment then identifies structures consistent with the experimentally determined NMR data. The approach thus relies on the strength of a well-

(43) Gaussian 94, Revision B.1: Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 1995.

(44) That is, Becke3LYP/6-31G**/MMFF is equivalent to Becke3LYP/6-31G**/HF/3-21G; cf. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.

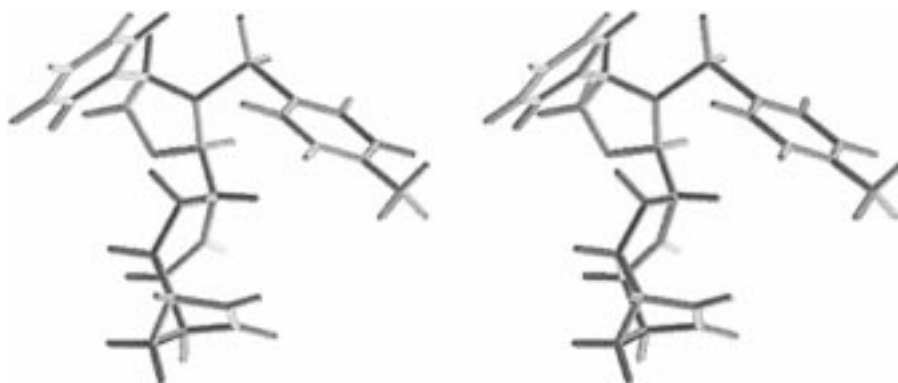


Figure 2. Reconstituted rMD conformation of **1** in stereo projection; “rMD”.

parametrized force field (accurate geometries) while avoiding a weakness for highly polar molecules (inaccurate relative energies).

The rMD Single Conformation. The restrained molecular dynamics (rMD) structure for **1** was compared with the “best fit” family of endo conformers. Table 6 gives dihedral values for torsions about the N–S, S–C_{tol}, and C–C_{ph} bonds (first three entries labeled def, efg, and hij) and for the central C1–C6 bond between the oxazolidine and cyclopentanone (last four entries labeled kcl, dck, acd, and bcd; see Figure 1 for labeling). Table 6 also includes torsions from the rMD structure of Reggelin and co-workers,²⁴ an X-ray structure reported by Reggelin et al.,²⁶ a reconstituted rMD structure (“rMD”, see below), and the structures selected by NAMFIS (“best solution”) from conformational manifolds produced by MMFF, MM2*, and MM3(96).

As mentioned above, all three force fields give different sets of conformers as best fits to the NMR data. Nevertheless, while the rMD conformer is distinct from the top one or two dominant conformers in the NAMFIS population with respect to torsion about the tosyl and phenyl exocyclic bonds, it is similar to them about the C1–C6 bond. For the former (first three entries in Table 6) there is a great deal of variation about these dihedral angles, in part due to low rotational barriers. None of the conformers from any of the NAMFIS analyses have torsions about the tosyl and phenyl bonds that are similar to the rMD structure. Clearly, the single rMD conformation is not consistent with NOEs between tosyl/phenyl protons and those of the five-membered rings (Table 3). This outcome confirms a later distance geometry (DG) analysis by Mierke and Reggelin that reports standard deviations of 100° and greater about these torsional angles.²⁵ An additional discrepancy between calculation and experiment is found for the NOE distance between H-4^{pro-R} and H-8 (see Table 3). Both NAMFIS (with all three force fields) and rMD methods determine this distance to be about 1.5 and 0.8 Å, respectively, longer than experiment. It is likely that the NMR-determined distance is in error.

For the dihedral angle between oxazolidine and cyclopentanone (i.e., C1–C6), a large percentage of conformers in all three endo datasets have torsions consistent with the rMD structure. In the present study, 76%, 53%, and 93% of conformers comprising the NAMFIS best solutions for the MMFF, MM2*, and MM3(96) force fields, respectively, meet this criterion.

The rMD Structure is High Energy. In the present work, the rMD conformation of **1** was reconstituted from



Figure 3. MD structure (darker) constrained to interproton distances and torsions reported by Reggelin and co-workers (ref 24) overlaid on MMFF optimized structure (lighter).

the data reported by Reggelin and colleagues by reoptimizing the structure with MMFF, MM2*, and MM3*⁴⁵ under the restraint that both the torsional angles around C1–C6 and the intramolecular distances conform to the reported rMD values (i.e., “rMD”, Figure 2).²⁴ NAMFIS matching of the rMD structure and the NMR data led to relatively large SSDs for the NOEs and the ³J_{HH}'s (SSD = 1000, Tables 2–4).

The rMD structure was subsequently reoptimized without constraints with MMFF, MM2*, and MM3* leading to energy drops of 19, 52, and 73 kcal/mol, respectively. The structure is clearly distant from the nearest local minimum on the potential energy surface for all three force fields. The rMD structures before and after MMFF unconstrained optimization, respectively, are overlaid in Figure 3.

The above comparison between the energies of MM energy minima and rMD has been criticized by a reviewer from the viewpoint that molecular dynamics simulations partition the total energy between kinetic and potential

(45) MM3(96) is not equipped to optimize a structure with more than ten internal constraints. As a result, we employed MM3*/Macromodel as a surrogate for this calculation. All missing parameters were supplemented with MM3(96) values.

energy contributions. Consequently, an MD structure is expected to fall at higher energy than a corresponding energy minimum. On the face of it, this is correct. In the case of **1**, however, the MD search was conducted with *ground-state geometric constraints* derived from NMR measurements. Were **1** a single conformation in CDCl₃, this strategy should have defined the solution conformer such that neither geometry nor energy vary when the structure is optimized with and without reference to rMD. The fact that fully optimized rMD (in bond distance and bond angle space) falls 19–73 kcal/mol when torsional fixation is lifted implies clearly that its instability is due to conformational misfitting and not to residual MD kinetic energy. The latter is effectively eliminated by the geometry optimization process.

Conclusions

The NAMFIS analysis supports two features of the previous NMR analysis for **1**. Namely, the latter is the endo isomer, and the dominant conformations sustain an inter-ring torsion $\phi_{bcd} \cong 180^\circ$ (see Figure 1). Importantly, we conclude that tricyclic ketone **1** is not well-characterized by a single conformation in CDCl₃ solution, but by a set of rapidly equilibrating conformers that produces a deceptively averaged NMR spectrum. It would appear that the previously suggested structure is a virtual

conformation that combines multiconformer features in a single geometry. As such, the probability of its existence on the conformational surface of the endo Diels–Alder adduct is negligible. In general, the presumption of single or strongly preferred conformation for a small molecule characterized by easily rotated bonds involves both a risk and an inappropriate assumption. The assertion is best tested by independent means prior to mapping NMR distances and torsions onto a single aggregate of atoms.

Quite apart from the dynamic complexity of adduct **1** in CDCl₃ solution, we demonstrate that current molecular mechanics schemes are unable to uniformly reproduce the conformational profile of the molecule. This includes a common collection of conformations, relative energies, and a reproducible global minimum. Our limited single-point DFT calculations (Table 5) reveal a similar trend. Traced primarily to the differential treatment of electrostatics from force field to force field, the effect is quite likely to be general for molecules with three or more polar functionalities separated by four or more bonds. While further investigation remains to be done, it suggests that the current generation of force fields and charge dependent quantum mechanical procedures may be incapable of mapping experimental conformations for highly polar molecules with well-separated functional groups.

JO9824450