

Accuracy in determining interproton distances using Nuclear Overhauser Effect data from a flexible molecule

Catharine R. Jones, Craig P. Butts* and Jeremy N. Harvey

Full Research Paper

Open Access

Address:
Department of Chemistry, University of Bristol, Cantock's Close,
Bristol, BS8 1TS, United Kingdom

Email:
Catharine R. Jones - chcrj@bris.ac.uk; Craig P. Butts* -
craig.butts@bris.ac.uk; Jeremy N. Harvey -
Jeremy.Harvey@bristol.ac.uk

* Corresponding author

Keywords:
conformation; internuclear distances; NMR spectroscopy; NOE

Beilstein J. Org. Chem. **2011**, *7*, 145–150.
doi:10.3762/bjoc.7.20

Received: 03 November 2010
Accepted: 10 December 2010
Published: 01 February 2011

Guest Editor: J. Murphy

© 2011 Jones et al; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

The determination of accurate NOE-derived interproton distances and confirmation/prediction of relative populations in multi-conformer, flexible small molecules was investigated with the model compound 4-propylaniline. The low accuracy assumed for semi-quantitative NOE distance restraints is typically taken to suggest that large numbers of constraints need to be used in the dynamical analysis of flexible molecules, and this requires, for example, the measurement and Karplus-type analysis of scalar coupling constants ($^3J_{CH}$ and $^3J_{HH}$). Herein we demonstrate that, contrary to this common perception, NOE measurements alone are accurate enough to establish interproton distances, and hence conformational detail, in flexible molecules to within a few percent of their ensemble-averaged values, hence reducing the demand for additional restraints in such dynamic analyses.

Introduction

Information obtained from Nuclear Overhauser Effect (NOE) experiments in NMR spectroscopy is widely employed in the determination of stereochemical and conformational details [1]. It is traditionally used in a qualitative or semi-quantitative manner to establish gross differences between conformers. However, the quantitative use of NOEs is often discounted or at least considered to be only approximately accurate. This perceived inaccuracy in NOE-distance relationships arises because many factors may perturb NOE intensities, including spin diffusion, selective polarisation transfer, variation in τ_c

between spins, accuracy of signal integration and conformational flexibility [1]. Despite this, we have recently shown that surprisingly accurate NOE-derived distances can be obtained in small organic molecules, and that many of the perturbing factors do not contribute significantly to NOE intensities when the molecule of interest is in the fast tumbling regime and measurements are made within the Initial Rate Approximation limits [2]. In this previous report we observed mean errors of ~3% in distances obtained from 1D NOESY experiments on the rigid molecule strychnine in *d*₆-benzene (as

compared to their computationally-derived values [3]). With accurate distances obtained in a rigid organic molecule, it seems sensible to examine whether this approach can be extended more generally to multi-conformational systems with similar accuracy. Further, it is likely that the accurate interproton distance-assessments from NOE will allow accurate modelling of conformer populations in solution with fewer NOE constraints, or indeed improvements in the accuracy of modelling with the same number of NOE constraints.

If flexible systems exhibiting multiple conformations in solution are interconverting rapidly on the NMR time-scale, then conformational exchange will lead to ensemble-averaging of the observed NOEs for each corresponding interproton distance in each contributing conformer. One approach to analysing such ensemble-averaged NOEs is to assume the molecule will occupy a number of distinct low-energy conformations with particular populations in solution. The ensemble-averaged NOE-determined distances can thus be used, along with computations of conformer geometries, to confirm the structures and energies/populations of contributing conformers. An excellent description of the advantages and disadvantages of such an approach was made by Kozerski et al., using conformational ensemble fitting to NOE data from 2,3-dihydrobenzofuran derivatives to determine stereochemical and conformational information [4]. Critically, they highlight the challenge in fitting multi-conformer, multi-isomeric models, which require large numbers of NOE contacts in order to extract the best-fit to these loose data. On the other hand, with more accurate distances available, we suggest it will be possible to identify not only geometry, but also fit conformation populations to within reasonable errors.

Method

The determination of interproton distances from NOE data previously described by us is based on comparison of relative NOE intensities for pairs of spins in 1D transient NOESY experiments [2]. Assuming that the molecule of interest is in the fast tumbling regime and that the Initial Rate Approximation holds true, the normalised NOE intensity between two spins I and S , η_{IS} , is proportional to the cross-relaxation rate, σ_{IS} , between these spins and the mixing time, τ_m , of the experiment (Equation 1). In turn, the cross-relaxation rate, σ_{IS} , between spins I and S is proportional to the internuclear distance between spins I and S (r_{IS}^{-6}) as described in Equation 2. A more complete description of these equations and their use in determining interproton distances can be found in references [1] and [2].

$$\eta_{IS} = \sigma_{IS} \tau_m \quad (1)$$

$$\sigma_{IS} = k r_{IS}^{-6} \quad (2)$$

where

$$k = \left(\frac{\mu_0}{4\pi} \right) \frac{\hbar^2 \gamma^4}{10} \left(\frac{6\tau_c}{1 + 4\omega^2 \tau_c^2} - \tau_c \right)$$

Assuming that the values defining k (ω -Larmor frequency, τ_c -rotational correlation time, γ -magnetogyric ratio) remain constant for *each spin pair in a given selective inversion experiment*, the ratio of intensities of a pair of NOE signals, $\eta_{I1S}:\eta_{I2S}$, *within that experiment* can thus be assumed to be proportional to the ratio of their internuclear distances (Equation 3). Thus, by comparing η_{I1S} and η_{I2S} *within the same selective inversion experiment*, we only need to know one distance, e.g., r_{I1S} , in order to calculate the second distance, r_{I2S} .

$$\frac{\eta_{I1S}}{\eta_{I2S}} = \frac{r_{I1S}^{-6}}{r_{I2S}^{-6}} \quad (3)$$

As outlined above, when there are multiple conformations describing a flexible molecular system, the matter of internuclear distance determination becomes more challenging. A general outline of the treatment of flexible small molecules using NOE experiments for both conformational and population analysis can be found in reference [1]. Larger flexible systems with multiple conformations such as peptides and proteins have been investigated using NOE-derived data, where techniques such as ensemble-averaged full relaxation matrix approaches are used to simulate NOEs [5-7]. The relative proportions of the various conformations contributing to the ensemble are then iteratively adjusted to derive the best fit to the experimentally measured NOEs. Accurate determination of populations in these analyses has always been limited by the inherently low accuracy of the NOE-derived restraints used, resulting in a broad range of conformer populations fitting the observed NOE restraints.

On the other hand, with the high accuracy provided by the NOE-distance analysis we employ, we have recently identified and quantified a previously unrecognised conformer of strychnine through a relatively minor deviation in a single measured NOE-derived interproton distance across the seven-membered ring of strychnine [8]. This intra-ring distance was observed to be ~ 0.6 Å (15%) shorter than expected by X-ray crystallography [9] and DFT [3], which would traditionally be considered an acceptable ‘experimental error’ for NOE-derived interproton distances. However, given that the average errors for NOE-

derived distances in our earlier study [2] were ~3%, this was identified as significantly anomalous. In the event, a second low-energy conformer was identified with a population of ~2.2% compared to the major conformer. By incorporating this second conformer, the error for the problematic distance was reduced from 15% to 3%, while no other significant changes in the NOE-derived distances arose.

We now sought to investigate the conformer populations of a flexible molecule where conformational exchange causes the perturbation of more than one distance, and hence more than one NOE intensity. Herein, we report that the high accuracy reported for relatively rigid molecules is maintained for multiple NOE-derived distances arising from conformationally flexible molecules when compared to their time-averaged computationally-derived distances in the alkyl chain of a small molecule – 4-propylaniline. These NOE-derived distances can thus be applied to modelling the populations (and hence energy differences) of the conformers, again with good accuracy when compared to their calculated values.

Results and Discussion

A B3LYP/6-31G* conformational search of 4-propylaniline leads not surprisingly to four non-degenerate low-energy conformers, with the propyl chain either in an anti, **1a** and **1b**, or a gauche, **2a** and **2b**, conformation (Figure 1). The anti conformers, **1a** and **1b**, differ from each other only in the opposite pyramidalisation of the amino group: The gauche conformers **2a** and **2b** are similarly related, with two further degenerate gauche conformers of **2a** and **2b** arising by rotation of 120° around the C5–C6 bond. In each case, the isomer ‘a’ is the one in which the aniline H atoms are on the same side of the benzene ring as the carbon chain. The optimised H–H distances in **1a** are very similar to those found in **1b**, and likewise for **2a** and **2b**.

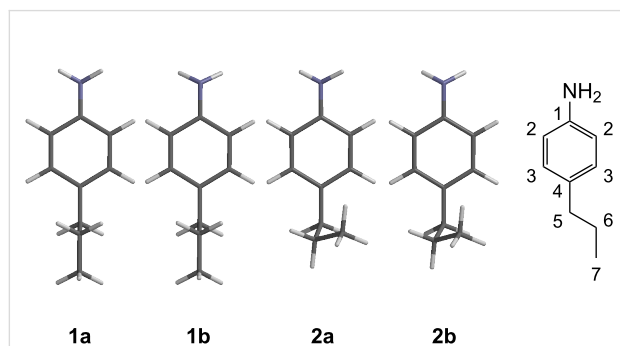


Figure 1: Four low energy conformations of 4-propylaniline obtained from B3LYP/6-31G* conformational search and labelled 2D structure (**1a/2a** differ from **1b/2b** in the pyramidalization at the amine nitrogen, see main text for details).

The calculated energies of conformers **1a** and **1b** differ by less than 0.01 kJ/mol from each other, and lie only 1.99 and 1.94 kJ/mol lower than **2a** and **2b**. After correction for zero-point energy and thermal and entropic corrections, the relative calculated ΔG at 298 K for conformers **1a**, **1b**, **2a** and **2b** were 0.09, 0.00, 3.20 and 3.14 kJ/mol, respectively. Finally, on including the solvation free energies (in CDCl_3) for all species, this leads to corresponding predicted free energies of 0.10, 0.00, 3.55 and 3.39 kJ/mol, respectively. As the B3LYP relative electronic energies are not expected to be highly accurate, single-point calculations using two very accurate local correlated methods were used to refine the gas phase relative energies. At the LCCSD(T0) level of theory, relative electronic energies of 0.07, 0.03, 0.09 and 0.00 kJ/mol, respectively, are obtained. The better description of dispersion interactions stabilises the more compact gauche conformers as expected. LPNO-CEPA-1/cc-pVTZ calculations, yield relative energies of 0.03, 0.00, 0.20 and 0.14 kJ/mol, very close to the coupled-cluster values, suggesting that these quantities are reliable to ± 0.5 kJ/mol and perhaps better.

Including the corrections for zero-point energy, thermal and entropic corrections, and solvation, these correlated ab initio calculations lead to predicted relative free energies in solution for the four conformers of 0.15, 0.00, 1.63 and 1.43 kJ/mol based on the LCCSD(T0) calculations, and very similar values of 0.14, 0.00, 1.76 and 1.60 kJ/mol based on the CEPA calculations. Taking into account the two-fold degeneracy of each of the gauche conformers **2a** and **2b**, this gives expected equilibrium populations of 23, 24, 25 and 27% or 24, 25, 25 and 26%, respectively. As the distances in each pair of conformers are so similar, this can be described more concisely as being predicted populations of **1 (a + b)** and **2 (a + b)** of 47% and 53% or 49% and 51%.

Experimentally, a selection of interproton distances: H3–H5, H3–H6, H3–H7, H5–H7, was determined using NOE intensities from 1D NOESY experiments. The intensities of the measured NOE signals were first corrected for the chemical equivalence/symmetry in each group by dividing the NOE intensity between signals *I* and *S*, η_{IS} , by $n_I n_S$, where n_I and n_S are the number of chemically equivalent spins in the groups giving rise to signals *I* and *S* respectively [1]. In order to convert the corrected NOE intensities into physically realistic interproton distances, *vide supra*, they need to be scaled against a single fixed reference distance for which an NOE has also been measured [1,2]. The obvious candidate distance, between the fixed aromatic H2 and H3 protons, could not be used as the experimental NOE signal between these two protons showed strong coupling artefacts that obscured the NOE intensity. Instead, the NOE data were internally calibrated from the

conformationally averaged H3–H5 NOE, assuming a reference distance for this proton pair of 2.77 Å (which is the calculated ensemble-average using the populations for **1a/b** and **2a/b** as determined above). The relative NOE intensities were then converted to ensemble-averaged internuclear distances, r_{NOE} , by applying the r^{-6} analysis *vide supra*. Where the same ensemble-averaged interproton distance was measured by two NOE experiments, e.g., H3–H6, H6–H3 (where the labels in italics are the inverted proton respectively) the experimental distance, H3–H6, was taken as the average of the pair and the results are presented in Table 1.

Table 1: NOE-derived and ensemble-averaged interproton distances.

	r_{NOE} (Å)	r_{calc} (Å)	% error
H3–H5 (ref)	2.77	2.77	-
H3–H6	3.21	3.34	3.83
H3–H7	4.16	4.05	2.54
H5–H7	3.13	3.08	1.67

To calculate the effective interproton distances, r_{calc} , for the conformationally averaged molecule, i.e., those distances which theoretically should be determined by the NOE experiments, each interproton distance, r_{IS} , as determined in the B3LYP/6-31G* structural optimisation for the conformers **1a**, **1b**, **2a** and **2b** was converted into a corresponding NOE intensity for each conformer using the r^{-6} relationship ($\eta_{\text{IS}} = r_{\text{IS}}^{-6}$). The measured NOEs all involve protons in chemically equivalent groups, e.g., H3 and H3' or H6 and H6', so the effective distances were obtained from the calculated structures (and thus relative NOE values determined) for *all* the contributing chemically equivalent pairs in each case, e.g., the H3–H6 NOE derived from $r_{\text{H3–H6}}$, $r_{\text{H3–H6}'}$, $r_{\text{H3}'–H6}$ and $r_{\text{H3}'–H6}'$, which were combined with $\langle r_{\text{IS}}^{-6} \rangle = (\sum r_{\text{IS}}^{-6}) / (n_I n_S)$, where n_I and n_S are the number of equivalent spins in the groups *I* and *S*, respectively. These individual NOE intensities for each of the six conformers were then weighted using the calculated Boltzmann populations from above, to convert them into effective distances.

A comparison of the resulting NOE-derived distances, r_{NOE} , and ensemble-averaged distances, r_{calc} , from calculations for H3–H6, H3–H7 and H5–H7 are shown in Table 1. The observed errors for these distances derived from a free-rotating group are in the range of 1–4% (mean 2.68%) and this compares very well with the observations in our previous work on the rigid strychnine system, which gave mean errors of ~3% in *d*₆-benzene and ~4% in CDCl₃ [2,3].

An alternative method of data analysis might assume that the NOE measurements and the DFT conformer structures are

accurate, but that the computed free energies (and hence the populations) are not. Thus, one might use the NOE and geometry data to model the conformer populations and hence the relative energies. In theory, there are too many conformers (four non-degenerate) to model their populations against the measured NOEs. However, if one assumes the degeneracy of **1a/b** and also **2a/b**, on near-symmetry grounds, then the relative populations of conformers **1a** and **1b** would be equal, (*x*), as would **2a** and **2b**, (1–*x*), reducing the problem to a single unknown with three NOE restraints. Figure 2 shows the plot of the average errors, compared to experiment, arising from weighting the calculated NOE intensities (*vide supra*) for each conformer by a range of populations (*x*) of **1a/b**. The best fit to experiment is obtained at ~55% **1a/b**, (and thus ~45% **2a/b**), corresponding to a free energy difference of ca. 2.1–2.2 kJ/mol between conformers **1a/b** and **2a/b**. This compares extremely well with the highest level calculations, which suggest a corresponding free energy difference of ca. 1.5 kJ/mol. This remarkable match in populations and energies supports the proposition that fitting populations of conformers from the NOE-derived distances is an inherently accurate approach when the geometries of the contributing conformers can be accurately described.

Figure 2 also demonstrates the requirement for the relatively accurate determination of internuclear distances in order to fit conformationally flexible systems. Assuming a mean 3% error in the interproton distances determined by NOEs (black line in Figure 2), the experimental results are consistent with a population of **1a/b** lying in the range 40 to 65%. However, typical semi-quantitative NOE studies are assumed to lead to much larger errors in distances, in the order of ± 10% (red line in Figure 2), or even assign distance-ranges to NOE intensities such as strong (2–3 Å), medium (3–4 Å) and weak (> 4 Å). The entire population range of conformers **1a/b** falls below the red line in Figure 2, showing that such loose assumptions of NOE accuracy would effectively allow *any* population of **1a/b** and **2a/b** to be fitted acceptably to the experimental data, requiring other constraints to be invoked in order to model even this simple system.

In summary, these results suggest that the methods we have previously described for extracting accurate interproton distances from NOE data in rigid systems can be extended with little, or no, loss of accuracy to relatively flexible small molecules. The high accuracy of this distance data allows it to be applied to assessing/confirming the relative populations of contributing conformers in small, flexible molecules with reasonable certainty, even where very few restraints are available – succeeding where traditional semi-quantitative NOE analysis would not.

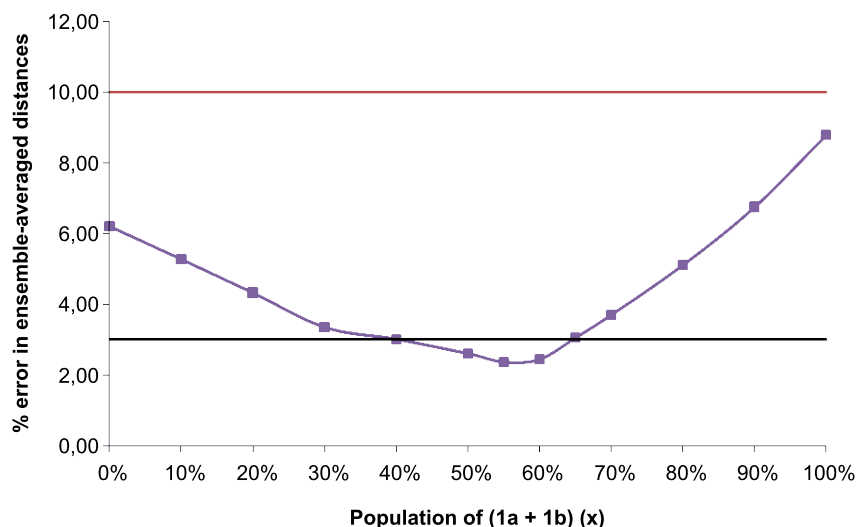


Figure 2: Plot of the mean calculated error arising from fitting the population distribution of conformers **1a/b** (x) and **2a/b** ($1-x$) to the experimental NOE-derived interproton distances. Exemplar error thresholds are represented at 3% (black line) and 10% (red line).

Experimental

NMR samples were prepared in 5 mm tubes with 0.7 ml CDCl_3 and ~10 mg 4-propylaniline, in air without degassing. NMR data were all collected on a 500 MHz Varian VNMRS Direct-Drive spectrometer equipped with an indirect observe probe. 1D selective transient NOESY spectra (64 k data points, 8 kHz spectral width, 500 ms mixing time, 4.096 sec. acquisition time, 1 s relaxation delay, 512 scans (45 minutes/irradiation)) were obtained using the Varian Chempack NOESY1D sequence which is based on the DPFGENOE (double-pulse field gradient spin-echo NOE) excitation sculpted selective sequence reported by Stott et al. [10] and incorporates a zero-quantum filter element [11]. NOE build-up curves were obtained with mixing times up to 900 ms and the critical constancy of relative NOE intensities within each irradiation was confirmed (as well as the linearity of absolute NOE intensities).

Geometry optimisation at the B3LYP/6-31G* level was carried out using the Gaussian 03 package and frequencies were computed to characterise the minima and derive statistical mechanical corrections to the electronic energies. Gas-phase single point energies were then calculated at the four minima in Gaussian, with B3LYP/6-31G* and a polarisable continuum model (IEF-PCM, parameters for chloroform solvent, $\epsilon = 4.9$). The LCCSD(T0)/cc-pVTZ calculations [12,13] were performed as implemented in the MOLPRO2008 package [14]. Inspection of the orbital domains for the different conformers show that a consistent set is obtained for all, hence the domain error on relative energies should be small. The LPNO CEPA-1/cc-pVTZ calculations [15] were performed using the implementation in the ORCA 2.8 package [16].

Acknowledgements

We thank the University of Bristol for funding for CRJ and the EPSRC (EP/F013515/1) and BBSRC (BB/F011539/1) for instrumentation funding. We thank Frank Neese for providing the ORCA program package.

References

- Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd ed.; John Wiley & Sons, Inc.: New York, 2000.
- Butts, C. P.; Jones, C. R.; Towers, E. C.; Flynn, J. L.; Appleby, L.; Barron, N. J. *Org. Biomol. Chem.* **2011**, *9*, 177–184. doi:10.1039/C0OB00479K
- Bagno, A.; Rastrelli, F.; Saielli, G. *Chem.–Eur. J.* **2006**, *12*, 5514–5525. doi:10.1002/chem.200501583
- Kozerski, L.; Krajewski, P.; Pupek, K.; Blackwell, P. G.; Williamson, M. P. *J. Chem. Soc., Perkin Trans. 2* **1997**, 1811–1818. doi:10.1039/a700149e
- Bonvin, A. M. J. J.; Brünger, A. T. *J. Biomol. NMR* **1996**, *7*, 72–76.
- Xu, Q.; Gitti, R.; Bush, C. A. *Glycobiology* **1996**, *6*, 281–288. doi:10.1093/glycob/6.3.281
- Wang, J.; Hodges, R. S.; Sykes, B. D. *J. Am. Chem. Soc.* **1995**, *117*, 8627–8634. doi:10.1021/ja00138a019
- Butts, C. P.; Jones, C. R.; Harvey, J. N. *Chem. Commun.* **2011**, *47*, 1193–1195. doi:10.1039/C0CC04114A
- Messerschmidt, M.; Scheins, S.; Luger, P. *Acta Crystallogr., Sect. B: Struct. Sci.* **2005**, *61*, 115–121. doi:10.1107/S0108768104032781
- Stott, K.; Keeler, J.; Van, Q. N.; Shaka, A. J. *J. Magn. Reson.* **1997**, *125*, 302–324. doi:10.1006/jmre.1997.1110
- Thrippleton, M. J.; Keeler, J. *Angew. Chem.* **2003**, *42*, 3938–3941. doi:10.1002/anie.200351947
- Schütz, M. *J. Chem. Phys.* **2000**, *113*, 9986–10001. doi:10.1063/1.1323265

13. Schütz, M.; Manby, F. R. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349–3358. doi:10.1039/b304550a
14. MOLPRO, Version 2010.1, a package of *ab initio* programs, Werner, H. J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M. and others, see <http://www.molpro.net>.
15. Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009**, *130*, No. 114108. doi:10.1063/1.3086717
16. ORCA, Version 2.8; Neese, F.; Wennmohs, F.: Bonn, 2010.

License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (<http://www.beilstein-journals.org/bjoc>)

The definitive version of this article is the electronic one which can be found at:
[doi:10.3762/bjoc.7.20](https://doi.org/10.3762/bjoc.7.20)