A Stable *cis*-Azobenzene in Aqueous Solution

Eric J. Chambers and Ian S. Haworth*

Dept. of Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, Los Angeles, CA 90033, USA

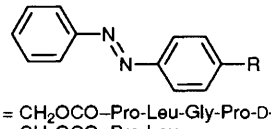
In aqueous solution the peptides phenylazobenzyloxycarbonyl(pz)–Pro-Leu-Gly-Pro-D-Arg and pz–Pro-Leu have at least one third of their azobenzene moieties present in the *cis* configuration.

The peptide phenylazobenzyloxycarbonyl(pz)–Pro-Leu-Gly-Pro-D-Arg (**1**, Fig. 1) enhances absorption of drugs into the gastrointestinal tract.[1] We report here the results of some preliminary NMR investigations into the conformation of this molecule and its derivative pz–Pro-Leu **2**. These show that at least one third of the azobenzene moieties exist in the *cis* conformation in aqueous solution. To our knowledge this is the first observation of such a stable *cis*-azobenzene. The stability is due to interactions with the peptide chain since in 4-hydroxyazobenzene **3** little of the *cis* isomer could be detected. The *cis*/*trans* equilibration of the azobenzene moiety in **1** and **2** occurs very slowly and the ^{1}H NMR spectra are further complicated by a second equilibrium, that between *cis* and *trans* proline residues.†

The ground state energies in the *cis*/*trans* equilibrium of free azobenzene **4** strongly favour the *trans* conformation,[2,3] but *cis*-azobenzene can be formed at ambient temperature by ultraviolet irradiation of the *trans* isomer.[4] However, even under these conditions a ^{13}C enriched sample was required to observe the ^{13}C NMR spectrum of the *cis* form in chloroform.[5] The aromatic region of the ^{1}H NMR spectrum of **1** in D_2O is shown 10 min [Fig. 2(*a*)] and 15 days [Fig. 2(*b*)] after dissolution. The difference between the two spectra is striking and the appearance of the four upfield peaks was unexpected. These peaks were assigned to the *cis* azobenzene isomer (vide infra). It is emphasized that the solution was not exposed to direct UV light throughout this period. Furthermore, direct UV irradiation of **1** in D_2O had no apparent effect. Similar irradiation of **1** in DMSO resulted in a minor enhancement of the upfield signals, consistent with their assignment to the *cis* form. However, only 15% of the *cis* form was observed under these conditions, considerably less than in the 15 day-unirradiated aqueous solution of **1**. The UV spectrum of the 15-day sample confirmed the presence of significant amounts of the *cis* azobenzene moiety in aqueous solution for **1** and **2**.

The appearance (excepting the relative intensities) of the aromatic signals of the ^{1}H and ^{13}C spectra was similar in D_2O and DMSO (D_2O shake). The low solubility of **1** in D_2O prevented the use of ^{13}C and 2D NMR techniques in this solvent. Hence assignment of the aromatic region of **1** in D_2O was based upon ^{1}H/^{1}H DQFCOSY and ^{1}H–^{13}C HETCOR experiments on both **1** and **4** in DMSO. The ^{13}C assignments for azobenzene in DMSO agree well with those previously obtained in $CDCl_3$.[5] In the COSY experiment no cross-peaks were observed between the group of aromatic signals assigned to the *trans* azobenzene and those assigned to the *cis* conformation.

To determine if the stability of the *cis* isomer is a consequence of the peptide substituent or an inherent property of the azobenzene moiety in aqueous solution we recorded NMR spectra of 4-hydroxyazobenzene **3**. This derivative was used because of the low solubility of **4** in D_2O. We did not observe a significant proportion of the *cis* conformation of **3**.



1 R = CH_2OCO–Pro-Leu-Gly-Pro-D-Arg
2 R = CH_2OCO–Pro-Leu
3 R = OH
4 R = H
5 Pro-Leu-Gly-Pro-D-Arg

Fig. 1

How is the *cis*-azobenzene conformation of **1** and **2** stabilized by the Pro-Leu chain? We believe that hydrophobic interactions between the peptide side chains and the aromatic rings provide the additional free energy which stabilizes the otherwise unfavourable *cis* conformation. The methyl region of the initial ^{1}H NMR spectrum of **1** and **2** in D_2O and the changes observed for these signals with time provide indirect evidence of different leucine/azobenzene orientations in different conformations.

An NOE difference experiment has confirmed that in at least one of the *trans* azobenzene conformations of **2** in DMSO the leucine sidechain is within 5 Å of the aromatic protons. Nonselective irradiation of the methyl signals resulted in enhancement of the *o*, *m* and *p* signals of the *trans* azobenzene.

The methyl region of the ^{1}H NMR spectrum of **1** in D_2O taken immediately following dissolution is shown in Fig. 3(*a*). Since in this solution we have very little of the *cis* azobenzene moiety, we can assign the two double doublet signals centred at δ 0.9 (l) and 0.7 (k) on the basis of isomerism around the pz–Pro (designated Pro 1 in **1**) amide bond. Signal k is probably due to the *trans* Pro 1 conformer, assuming this to be the preferred isomer,[6] whilst the *cis* Pro 1 conformer yields signal l. In the COSY spectrum no cross-peaks are observed between signals l and k. We obtained very similar spectra for **2** following dissolution in both D_2O and DMSO. Interestingly, for **1** in DMSO several of the doublet components are further

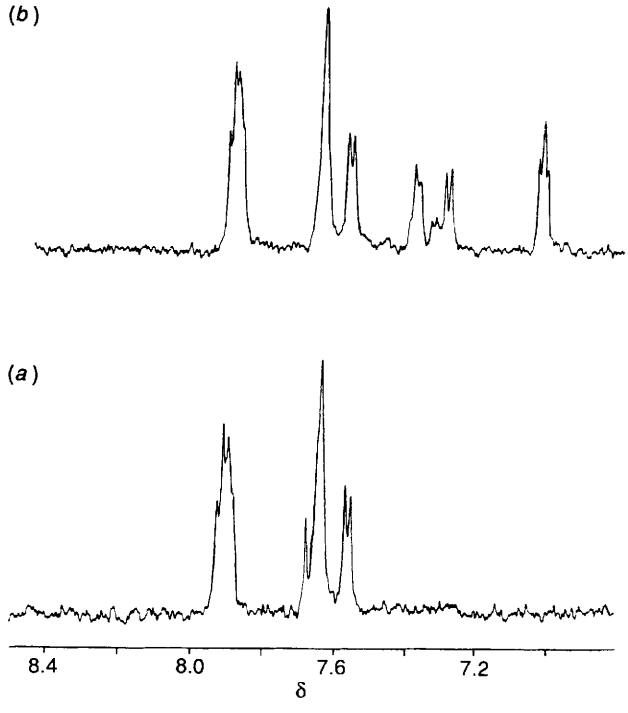


Fig. 2 Aromatic region of the 500 MHz ^{1}H NMR spectrum of **1** after (*a*) 10 min. and (*b*) 15 days. Experimental conditions: saturated aqueous solution, *ca* 300 K, TSP reference.

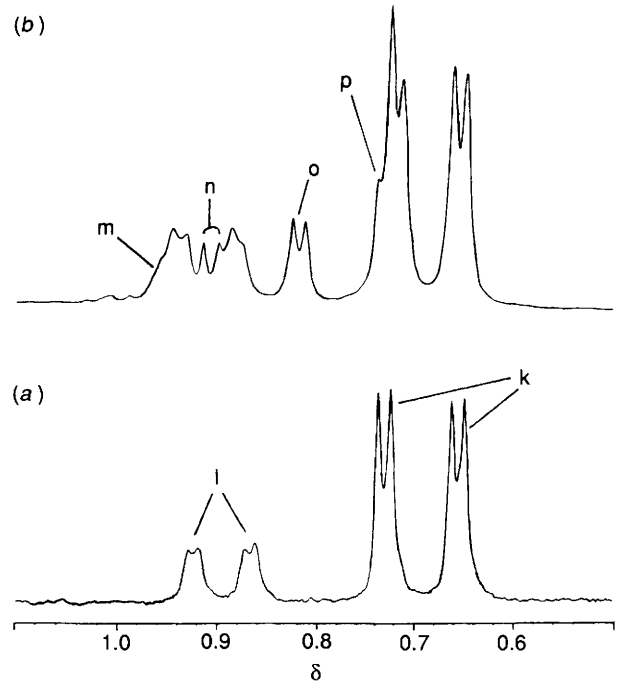


Fig. 3 Methyl region of the 500 MHz ^{1}H NMR spectrum of an aqueous solution of **1** after (*a*) 10 min and (*b*) 15 days

split, possibly reflecting the influence of a third equilibrium, *cis*/*trans* about Pro 2 (we note that we observed the *cis* Pro 2 conformation in the free peptide Pro 1-Leu-Gly-Pro 2-D-Arg **5**, but this had no apparent effect on the leucine methyl groups, which appeared as a double doublet at δ 0.95). Heating of **1** or **2** in DMSO to 360 K resulted in near-coalescence of signals l and k, showing them to be due to a now rapidly exchanging equilibrium. In the same spectra the aromatic signals (*trans* and *cis*) were unaffected. The solvent dependence of the conformational equilibria of **1** is further illustrated by the appearance of the methyl region in $CDCl_3$, in which the signals arising from the *cis* and *trans* isomers are unresolved at 298 K, suggesting the Pro 1 equilibrium to be in rapid exchange in this solvent even at room temperature.

Following dissolution of **1** in D_2O, we estimate a *trans* : *cis* Pro 1 ratio of about 80 : 20 with an essentially 100 : 0 *trans* : *cis* azobenzene ratio [Fig. 3(*a*) and 2(*a*), respectively]. The methyl region of the same sample at isomeric equilibrium of the azobenzene moiety [Fig. 3(*b*)] has several additional signals compared to the initial spectrum. These are doublets at δ 0.9 (n) and 0.8 (o), a third signal (p) at *ca.* δ 0.74 partially obscured by the signal k and a signal (m) obscured by signal l. Consistent with the *trans* : *cis* azobenzene equilibrium [Fig. 2(*b*)], these represent about 30% of the total methyl signal.

Molecular dynamics simulations on **2** in a periodic box of water using the AMBER 4.0 forcefield[7] suggested the configuration of the pz–Pro amide bond in **2** is important in facilitating side chain/aromatic-ring interactions. Calculations were performed starting from (gas phase) minimized conformations for each of the four azo/Pro isomers. The leucine sidechain remained close to the phenyl rings in the MD simulation of the *trans*-azo, *cis*-Pro isomer. The *cis*-Azo, *cis*-Pro isomer adopted a conformation in which the hydrophobic azobenzene, the proline ring, and the leucine sidechain were aligned on one side of the molecule. The other side of the molecule is more hydrophilic, the leucine carboxylic acid being aligned with the ester of the benzyloxycarbonyl group. This is of interest considering the membrane penetrating properties of the molecule. For the *trans*-azo, *trans*-Pro and *cis*-azo, *trans*-Pro isomers the *trans* proline configuration prevents significant association of the leucine sidechain with the azobenzene.

The central observation in this study is the existence of a stable *cis*-azobenzene in aqueous solution. Further experiments and simulations using **1**, **2** and mutated peptides are in progress and these should allow us to make a definitive assignment of the NMR spectra and provide a complete explanation for the *cis*-azobenzene stabilization.

We would like to thank Alan Kershaw and Dr T. Flood for time on the U.S.C. Chemistry Department Bruker AMX 500 NMR spectrometer and Dr Vince Lee for samples of **1**, **2** and **5** and for many helpful discussions.

Received, 21st March 1994; Com. 4/01679C

Footnote

† *Chemical shift assignments* for azobenzene moieties: **1** in DMSO, *trans*, ^{13}C δ 151.91(C-1), 122.38(C-2), 129.34(C-3), 131.37(C-4), 140.54(C-4′); ^{1}H δ 7.89(H-2), 7.63(H-3), 7.52(H-4); *cis*, ^{1}H δ 6.812(H-2′), 6.847(H-2), 7.196(H-3′), 7.320(H-3), 7.297(H-4); **1** in D_2O, *trans*, ^{1}H δ 7.90(H-2′), 7.82(H-2), 7.641(H-3), 7.552(H-3′), 7.86(H-4); *cis*, ^{1}H δ 6.979(H-2, H-2′), 7.371(H-3), 7.263(H-3′), 7.30(H-4). **3** in D_2O, *trans*, ^{1}H δ 7.853(H-2), 7.620(H-3), 7.573(H-4), 7.815(H-2′), 7.037(H-3′). **4** in DMSO, *trans*, ^{13}C δ 152.37(C-1), 122.98(C-2), 129.99(C-3), 132.08(C-4); *cis*, ^{13}C, δ 153.99(C-1), 120.32(C-2), 129.36(C-3), 127.78(C-4); *trans*, ^{1}H δ 7.921(H-2), 7.648(H-3), 7.603(H-4); *cis*, 6.855(H-2), 7.326(H-3), 7.204(H-4). Numbers labelled with a prime denote the disubstituted ring. ^{13}C chemical shifts are referenced to the DMSO solvent peak (δ 39.5).

References

1 V. H. L. Lee, *BioPharm.*, 1992, **5**, 39.
2 S. Monti, G. Orlandi and P. Palmieri, *Chem. Phys.*, 1982, **71**, 87.
3 A. W. Adamson, A. Vogler, H. Kunkely and R. Wachter, *J. Am. Chem. Soc.*, 1978, **100**, 1298.
4 A. Vogel, *Textbook of Practical Organic Chemistry*, 4th edn. Longman, London, 1978, p. 208.
5 H. J. Shine and W. Subotkowski, *Magn. Reson. Chem.*, 1991, **29**, 969.
6 C. Grathwohl and K. Wuthrich, *Biopolymers*, 1976, **15**, 2025.
7 D. A. Pearlman, D. A. Case, J. C. Caldwell, G. L. Seibel, U. C. Singh, P. Weiner and P. A. Kollman, AMBER, University of California, San Francisco, CA, 1991.