Modeling ${}^{2h}J_{iso}(N, N)$ in nucleic acid base pairs: Ab initio characterization of the ${}^{2h}J(N, N)$ tensor in the methyleneimine dimer as a function of hydrogen bond geometry

David L. Bryce^a & Roderick E. Wasylishen^{b,*}

^aDepartment of Chemistry, Dalhousie University, Halifax, NS, Canada B3H 4J3 ^bDepartment of Chemistry, University of Alberta, Edmonton, AB, Canada T6G 2G2

Received 22 November 2000; Accepted 16 February 2001

Key words: ab initio calculations, hydrogen bonds, indirect spin-spin coupling tensors, *J*-coupling, ${}^{2h}J(N, N)$, MCSCF, nitrogen couplings, nucleic acid base pairs, scalar couplings

Abstract

The observation of ${}^{2h}J_{iso}(N, N)$ coupling has prompted considerable interest in this phenomenon from experimentalists and theoreticians due to the potential these couplings hold for the determination of secondary and tertiary structure in biologically important molecules. Here, we present an ab initio (MCSCF) study of the complete ${}^{2h}J(N, N)$ tensor for a model methyleneimine dimer system as a function of (*i*) the N-N separation, r_{NN} , and (*ii*) the hydrogen bond angle, θ . This simple system models the ${}^{2h}J(N, N)$ tensor of nucleic acid base pairs. Results indicate that although the Fermi-contact mechanism dominates ${}^{2h}J_{iso}(N, N)$, the coupling tensor is anisotropic due to contributions from the Fermi-contact spin-dipolar cross term. The variation in ${}^{2h}J_{iso}(N, N)$ as a function of r_{NN} is fit to an exponential decay. The influence of θ on the coupling constant is less pronounced but must be considered if experimental coupling constants are to be used for quantitative structure determination. Our results for this simple model system demonstrate that ${}^{2h}J_{iso}(N, N)$ is a valuable probe of hydrogen bonding in nucleic acid base pairs.

In solution, indirect nuclear spin-spin coupling constants ('scalar' or J coupling constants) are important in providing information about the secondary and tertiary structure of biological macromolecules (Marino et al., 1999; Case, 2000; Mollova and Pardi, 2000). One of the most spectacular developments in recent years has been the observation of unexpectedly large J couplings (\sim 7–10 Hz) between heavy nuclei across hydrogen bonds. The first such reports were concerned with ¹⁵N-¹⁵N J coupling in RNA (Dingley and Grzesiek, 1998) and DNA (Pervushin et al., 1998). Subsequently, an abundance of these couplings have been detected (e.g., Dingley et al., 1999, 2000; Hennig and Geierstanger, 1999; Majumdar et al. 1999; Hennig and Williamson, 2000; Pervushin et al., 2000). In addition to ${}^{2h}J(N, N)$, many reports of J couplings across hydrogen bonds involving other nuclei have been reported, e.g., ${}^{h}J(N, C)$ (Cordier and Grzesiek, 1999a; Cornilescu et al., 1999a, b; Wang et al., 1999), ${}^{2h}J(H, C)$ (Cordier et al., 1999b), ${}^{2h}J(H, H)$ (Fierman et al., 2000), ${}^{2h}J(P, H)$ and ${}^{3h}J(P, N)$ (Löhr et al., 2000; Mishima et al., 2000).

The theory of indirect nuclear spin–spin coupling was originally developed by Ramsey (1953). The isotropic coupling constant, J_{iso} , is equal to one-third the trace of **J**, a second rank tensor. In its principal axis system (PAS), the symmetric part of the **J** tensor is diagonal and has up to three independent elements, J_{11} , J_{22} , and J_{33} , ordered according to the convention $|J_{33} - J_{iso}| \ge |J_{11} - J_{iso}| \ge |J_{22} - J_{iso}|$. The anisotropy, ΔJ , of the tensor is defined as $J_{33} - (J_{11} + J_{22})/2$, and the asymmetry parameter, η , is given by $(J_{22} - J_{11})/(J_{33} - J_{iso})$. The present contribution is focused on elucidating the properties of the complete ^{2h}**J**(N, N) tensor. How does the ^{2h} J_{iso} (N, N) coupling constant depend on structure? What is

^{*}To whom correspondence should be addressed. E-mail: Roderick.Wasylishen@UAlberta.Ca



Figure 1. Model methyleneimine dimer system used for calculations of ^{2h}J(N, N) tensors. The nitrogen-nitrogen separation is denoted by r_{NN} and the hydrogen bond angle is indicated by θ .

the relative importance of the four J coupling mechanisms? Is the **J** tensor anisotropic? The latter question is of importance because one cannot easily separate direct dipolar interactions from anisotropic indirect nuclear spin–spin coupling interactions (Bryce and Wasylishen, 2000a; Wasylishen, 1996). In typical solution NMR experiments, one measures an *effective dipolar coupling constant*,

$$R_{\rm eff} = R_{\rm DD} - \Delta J/3 \tag{1}$$

where R_{DD} is the direct dipolar coupling constant which varies inversely with the motionally averaged cube of the internuclear distance. To determine highly accurate internuclear distances from experimental dipolar coupling constants (i.e., R_{eff}), the influence of ΔJ must be considered. Also, in principle, ΔJ will contribute to the efficiency of dipole–dipole relaxation (Blicharksi, 1972; Spiess, 1978; Wasylishen, 1996).

In this communication, we present the results of high-level ab initio multiconfigurational selfconsistent field (MCSCF) calculations of the complete 2h **J**(15 N, 15 N) tensor¹ for a simple model system, the dimer of methyleneimine (Figure 1). Use of this system models the ${}^{2h}\mathbf{J}(N, N)$ tensor of nucleic acid base pairs, the building blocks of DNA and RNA. It must be emphasized that accurate high-level ab initio calculations of complete J tensors for much larger molecular fragments are impractical at this time. Also, from a more fundamental point of view, it is advantageous to study relatively simple model systems where the effects of specific variables such as hydrogen bond length and bond angle may be systematically investigated. There have been few previous computational studies of couplings between nitrogen atoms across hydrogen bonds (Dingley et al., 1999; Scheurer and

Brüschweiler, 1999; Benedict et al., 2000; Del Bene et al., 2000a, b; Pecul et al., 2000). Of these studies, only Pecul and co-workers have employed the MCSCF approach to calculating ${}^{2h}J_{iso}(N, N)$.

The MCSCF method for calculating J tensors has been implemented in the Dalton Quantum Chemistry Program (Helgaker et al., 1997, 1999). In this implementation, all contributions to the coupling tensor are computed: spin-orbital (diamagnetic, DSO, and paramagnetic, PSO), spin-dipolar (SD), Fermi-contact (FC). There is also a FC \times SD cross term which contributes to ΔJ but not to J_{iso} . Recent studies have established the reliability of the MCSCF technique for calculating complete J tensors (Vahtras et al., 1992; Barszczewicz et al., 1994, 1995; Helgaker et al., 1998; Kaski et al., 1998, 1999; Bryce and Wasylishen, 2000a, b). For the present calculations, we have employed the experimental gas-phase geometry of the methyleneimine monomer (Johnson and Lovas, 1972; Pearson and Lovas, 1977) and systematically varied the N-N distance and N-H...N bond angle in the planar dimer in order to gain a thorough understanding of the effects of the hydrogen bond geometry on the properties of the ^{2h}J(N, N) tensor. All MCSCF calculations were carried out using Dalton running on an IBM RS/6000 workstation with dual 200 MHz processors. The calculations make use of the C_s point group symmetry of the planar dimer. The wave function was constructed using the following restricted active space (RAS): (40/00/10, 2/32), where the orbitals are of a' and a" symmetries respectively, and the orbital spaces are listed in the order (inactive/RAS1/RAS2/RAS3). Two electrons were allowed to be excited into RAS3. The cc-pCVDZ basis set has been used in all calculations; this corresponds to a total of 102 orbitals. Due to computational limitations inherent to the MC-SCF technique and inherent to Dalton, the use of larger basis sets was not feasible for our model system. Calculation of all of the indirect spin-spin coupling tensors in the model system required approximately 45 h of CPU time for each geometry; this time was reduced to 11 h for each geometry if only the 2h **J**(N, N) tensor was calculated.

Shown in Figure 2 is a plot of the isotropic and anisotropic portions of ^{2h}**J**(N, N) as a function of the nitrogen-nitrogen separation ($r_{\rm NN}$), with the N-H^{...}N angle fixed at 180°. The magnitude of the isotropic portion decays exponentially as a function of $r_{\rm NN}$ (Equation 2, $R^2 = 0.9959$); however, the fit tends to get worse for very small values of $r_{\rm NN}$. The datum

¹All results reported herein are for ¹⁵N-¹⁵N couplings.



Figure 2. Plot of the isotropic (J_{iso}) and anisotropic (ΔJ) portions of ${}^{2h}\mathbf{J}(\mathbf{N}, \mathbf{N})$ calculated for the model system shown in Figure 1 as a function of nitrogen-nitrogen separation, with a N-H…N bond angle of 180° and a N-H bond length of 1.02 Å. The data for the isotropic portion are fit to an exponential decay (the datum at $r_{\rm NN} = 2.02$ Å is not included in the fit). The smoothed line connecting the data points for the anisotropic portion is simply to illustrate the trend.

for $r_{\rm NN} = 2.02$ Å has not been included in the fits (Equations 2 and 3).

$$^{2n}J_{iso}(N, N) / Hz = 9640 \exp(-2.73r_{NN})$$
 (2)

$$r_{\rm NN}/{\rm \AA} = -0.37 \ln({}^{2\rm h}J_{\rm iso}({\rm N},{\rm N})) + 3.36$$
 (3)

An analogous correlation between hydrogen bond length in proteins and ${}^{3h}J_{iso}(N, C)$ has been described by Cornilescu et al. (1999b), based on experimental data. Typical N-N separations across nucleic acid base pairs are approximately 2.8–2.9 Å, although there are cases where the separation can be anomalously small, e.g., 2.2-2.6 Å (Jursa and Kypr, 1993). The correlation obtained in Equations 2 and 3 is reasonable when compared to existing experimental data. For example, the experimentally observed value of ${}^{2h}J_{iso}(N, N)$ in Watson-Crick base pairs ranges from approximately 6 to 8 Hz and in Hoogsteen base pairs from about 7 to 10 Hz (Dingley et al., 1999). Employing Equation 2 and standard N-N separations (2.8–2.9 Å) gives values of 3.5–4.6 Hz for ${}^{2h}J_{iso}(N, N)$. Using a typical value of 7 Hz for ${}^{2h}J_{iso}(N, N)$ in Equation 3 yields a short but certainly reasonable N-N distance of 2.64 Å. Although the reliability of calculated **J** tensors has increased greatly in the past several years (Helgaker et al., 1999), it should be emphasized that the computational results (Equations 2 and 3) should not be regarded as being accurate to better than 10%. Regardless of the computational technique and model system employed, if



Figure 3. Plot of the isotropic (J_{iso}) and anisotropic (ΔJ) portions of ^{2h}J(N, N) calculated for the model system shown in Figure 1 as a function of N-H^{...}N bond angle, with the N_{donor}-H separation fixed at 1.02 Å and the N_{acceptor}^{...}H separation fixed at 1.75 Å. The smoothed lines serve as guides to the trends.

such a level of accuracy were desired, rovibrational corrections would be required to account for the fact that the calculations are performed at the equilibrium geometry while experimental data are inevitably for molecules in particular rovibrational states. Additionally, quantitative comparison between experiment and theory necessitates a consideration of medium effects on the coupling tensor (Contreras et al., 2000). Such corrections are beyond the scope of this communication. Our main intent here is to study *trends* in the J tensor as a function of hydrogen bond geometry. Independent of the nitrogen-nitrogen distance, the FC mechanism dominates (97-99%) the isotropic coupling. The dominance of the FC term was also noted by Del Bene et al. (2000a) and Pecul et al. (2000). It is important to note, however, that despite this dominance, the anisotropic portion of ${}^{2h}J(N, N)$ is not insignificant compared to ${}^{2h}J_{iso}(N, N)$ (Figure 2). In fact, for very long hydrogen bonds, the magnitude of ΔJ equals and even surpasses the magnitude of the isotropic portion. Contributions to ΔJ are dominated by the FC \times SD cross term. The asymmetry of the J tensor, η , decreases smoothly from 0.47 to 0.20 as the nitrogen-nitrogen distance is increased over the range shown in Figure 2.

It has been suggested that the dominant factor in determining the value of ${}^{2h}J_{iso}(N, N)$ is the nitrogennitrogen distance (Del Bene et al., 2000a; Pecul et al. 2000). To test this hypothesis, we have also carried out calculations of ${}^{2h}J(N, N)$ as a function of the N-H^{...}N bond angle, while holding the N_{donor}-H and N_{acceptor}...H separations fixed at 1.02 Å and 1.75 Å, respectively. The results are presented in Figure 3. When the angle is reduced from 180° to 170°, there is a slight decrease in the value of ${}^{2h}J_{iso}(N, N)$. However, as the angle is diminished further, to 140° , ${}^{2h}J_{iso}(N, N)$ increases to 8.3 Hz. The FC mechanism dominates the J tensor regardless of the N-H...N bond angle. The value of ΔJ is less sensitive to the bond angle than is J_{iso} , and remains constant at about 3 Hz. Majumdar and coworkers (1999) have reported a comparatively small value for ${}^{2h}J_{iso}(N, N)$ in an A-A mismatch segment of DNA (2.45 Hz); it has been proposed (Majumdar et al., 1999; Dingley et al., 2000) that the reason for the reduced value of the coupling constant is due to an unusual hydrogen bond geometry. In addition to the possibility that the N-N distance in the A-A mismatch is simply longer than in typical nucleic acid base pairs, consideration of the data presented in Figure 3 indicates that a slight deviation from linearity of the hydrogen bond ($\sim 170^{\circ}$) could also contribute to the reduced value of $^{2h}J_{iso}(N, N)$.

Although the anisotropy of ^{2h}**J**(N, N) is small for all of the geometries considered here, it is nevertheless important to specify the orientation of the tensor in the molecular framework. For a nitrogen-nitrogen separation of 3.02 Å in a linear hydrogen bond, we find the principal components of the symmetric part of the ^{2h}**J**(N, N) tensor to be: $J_{11} = 0.93$ Hz, $J_{22} = 1.07$ Hz, $J_{33} = 2.00$ Hz. In its PAS, the tensor is oriented such that J_{22} is perpendicular to the dimer plane and J_{33} is approximately along the vector connecting the two nitrogen nuclei.

As a testament to the remarkable magnitude of $^{2h}J_{iso}(N, N)$ observed in nucleic acid base pairs, we consider that prior to the observation of these couplings, one of the largest known two-bond nitrogennitrogen couplings was through covalent bonds in urea, ${}^{2}J_{iso}(N, N) = \pm 5.1$ Hz (Stilbs and Forsén, 1976). To our knowledge, the largest ${}^{2h}J_{iso}(N, N)$ reported to date is 11.0 ± 1.0 Hz between the imidazole rings of histidine bases in sperm whale apomyoglobin, pH = 4.9 (Hennig and Geierstanger, 1999). Given that the largest known two-bond proton-proton coupling is found in formaldehyde (${}^{2}J_{iso}(H, H) = 41 \text{ Hz}$) (Shapiro et al., 1963; Pople and Bothner-By, 1965), it is reasonable to assume that ${}^{2}J_{iso}(N, N)$ is relatively large in urea since these two compounds have similar geometries. MCSCF calculations of ${}^{2}\mathbf{J}(N, N)$ in urea (C_{2v} ; RAS:3100/0000/5421/4321; cc-pCVTZ) indicate that $^{2}J_{iso}(N, N)$ is +5.9 Hz, thus providing the sign of this coupling and showing good agreement with experiment. The calculations show that the coupling is dominated by the FC mechanism; however, given the N-N distance of 2.3 Å (Swaminathan et al., 1984), it is clear that the simple correlation between ${}^{2h}J_{iso}(N, N)$

and $r_{\rm NN}$ shown in Figure 2 and Equations 2 and 3 cannot be extended to include two-bond N-N couplings in σ -bonded systems.

In summary, this first-principles investigation of $^{2h}\mathbf{J}(N, N)$ in the methyleneimine dimer has provided insight into the relative importance of the various coupling mechanisms and the dependence of both the isotropic and anisotropic portions on local molecular structure. In agreement with previous studies, we find that the nitrogen-nitrogen distance is dominant in determining ${}^{2h}J_{iso}(N, N)$ across hydrogen bonds; however, if highly accurate correlations are to be made between ${}^{2h}J_{iso}(N, N)$ and local molecular structure, consideration of the hydrogen bond angle is also necessary. ${}^{2h}J_{iso}(N, N)$ provides a measure of hydrogen bond strength insofar as it is related to hydrogen bond length. The general conclusions reached regarding the properties of ${}^{2h}\mathbf{J}(N, N)$ and the dependence of $^{2h}J_{iso}(N, N)$ on local molecular structure should be applicable to similar systems of biological relevance, specifically nucleic acid base pairs.

Acknowledgements

The authors thank all members of the solid-state NMR group at the University of Alberta for valuable comments. R.E.W. thanks NSERC for operating grants and for a grant to purchase the workstation used in the present work. D.L.B. thanks NSERC, the Izaak Walton Killam Trust, the Walter C. Sumner Foundation, and Dalhousie University for postgraduate scholarships.

References

- Barszczewicz, A., Helgaker, T., Jaszuński, M., Jørgensen, P. and Ruud, K. (1994) J. Chem. Phys., 101, 6822–6828.
- Barszczewicz, A., Helgaker, T., Jaszuński, M., Jørgensen, P. and Ruud, K. (1995) J. Magn. Reson., A114, 212–218.
- Benedict, H., Shenderovich, I.G., Malkina, O.L., Malkin, V.G., Denisov, G.S., Golubev, N.S. and Limbach, H.-H. (2000) J. Am. Chem. Soc., 122, 1979–1988.
- Blicharski, J.S. (1972) Z. Naturforsch., A27, 1355-1357.
- Bryce, D.L. and Wasylishen, R.E. (2000a) J. Am. Chem. Soc., 122, 3197–3205.
- Bryce, D.L. and Wasylishen, R.E. (2000b) J. Am. Chem. Soc., 122, 11236–11237.
- Case, D.A. (2000) Curr. Opin. Struct. Biol., 10, 197-203.
- Contreras, R.H., Peralta, J.E., Giribet, C.G., Ruiz de Azúa, M.C. and Facelli, J.C. (2000) Annu. Rep. NMR Spectrosc., 41, 55–184.
- Cordier, F. and Grzesiek, S. (1999a) J. Am. Chem. Soc., **121**, 1601– 1602.

- Cordier, F., Rogowski, M., Grzesiek, S. and Bax, A. (1999b) J. Magn. Reson., 140, 510–512.
- Cornilescu, G., Hu, J.-S. and Bax, A. (1999a) J. Am. Chem. Soc., 121, 2949–2950.
- Cornilescu, G., Ramirez, B.E., Frank, M.K., Clore, G.M., Gronenborn, A.M. and Bax, A. (1999b) J. Am. Chem. Soc., 121, 6275–6279.
- Del Bene, J.E., Ajith Perera, S. and Bartlett, R.J. (2000a) J. Am. Chem. Soc., **122**, 3560–3561.
- Del Bene, J.E. and Bartlett, R.J. (2000b) J. Am. Chem. Soc., 122, 10480–10481.
- Dingley, A.J. and Grzesiek, S. (1998) J. Am. Chem. Soc., **120**, 8293–8297.
- Dingley, A.J., Masse, J.E., Peterson, R.D., Barfield, M., Feigon, J. and Grzesiek, S. (1999) J. Am. Chem. Soc., 121, 6019–6027.
- Dingley, A.J., Masse, J.E., Feigon, J. and Grzesiek, S. (2000) J. Biomol. NMR, 16, 279–289.
- Fierman, M., Nelson, A., Khan, S.I., Barfield, M. and O'Leary, D.J. (2000) Org. Lett., 2, 2077–2080.
- Helgaker, T., Jensen, H.J.Aa., Jørgensen, P., Olsen, J., Ruud, K., Ågren, H., Andersen, T., Bak, K.L., Bakken, V., Christiansen, O., Dahle, P., Dalskov, E.K., Enevoldsen, T., Fernandez, B., Heiberg, H., Hettema, H., Jonsson, D., Kirpekar, S., Kobayashi, R., Koch, H., Mikkelsen, K.V., Norman, P., Packer, M.J., Saue, T., Taylor, P.R. and Vahtras, O. (1997) *Dalton: An electronic* structure program, *Release 1.0.*
- Helgaker, T., Jaszuński, M., Ruud, K. and Górska, A. (1998) *Theor. Chem. Acc.*, **99**, 175–182.
- Helgaker, T., Jaszuński, M. and Ruud, K. (1999) Chem. Rev., 99, 293–352.
- Hennig, M. and Geierstanger, B.H. (1999) J. Am. Chem. Soc., 121, 5123–5126.
- Hennig, M. and Williamson, J.R. (2000) Nucleic Acids Res., 28, 1585–1593.
- Johnson, D.R. and Lovas, F.J. (1972) Chem. Phys. Lett., 15, 65-68.
- Jursa, J. and Kypr, J. (1993) Gen. Physiol. Biophys., 12, 401-419.
- Kaski, J., Lantto, P., Vaara, J. and Jokisaari, J. (1998) J. Am. Chem. Soc., 120, 3993–4005.
- Kaski, J., Lantto, P., Rantala, T.T., Schroderus, J., Vaara, J. and Jokisaari, J. (1999) J. Phys. Chem., A103, 9669–9677.
- Löhr, F., Mayhew, S.G. and Rüterjans, H. (2000) J. Am. Chem. Soc., 122, 9289–9295.

- Majumdar, A., Kettani, A. and Skripkin, E. (1999) *J. Biomol. NMR*, **14**, 67–70.
- Marino, J.P., Schwalbe, H. and Griesinger, C. (1999) Acc. Chem. Res., 32, 614–623.
- Mishima, M., Hatanaka, M., Yokoyama, S., Ikegami, T., Wälchli, M., Ito, Y. and Shirakawa, M. (2000) J. Am. Chem. Soc., 122, 5883–5884.
- Mollova, E.T. and Pardi, A. (2000) Curr. Opin. Struct. Biol., 10, 298–302.
- Pearson Jr., R. and Lovas, F.J. (1977) J. Chem. Phys., 66, 4149– 4156.
- Pecul, M., Leszczynski, J. and Sadlej, J. (2000) J. Phys. Chem., A104, 8105–8113.
- Pervushin, K., Ono, A., Fernández, C., Szyperski, T., Kainosho, M. and Wüthrich, K. (1998) Proc. Natl. Acad. Sci. USA, 95, 14147– 14151.
- Pervushin, K., Fernández, C., Riek, R., Ono, A., Kainosho, M. and Wüthrich, K. (2000) J. Biomol. NMR, 16, 39–46.
- Pople, J.A. and Bothner-By, A.A. (1965) *J. Chem. Phys.*, **42**, 1339–1349.
- Ramsey, N.F. (1953) Phys. Rev., 91, 303-307.
- Scheurer, C. and Brüschweiler, R. (1999) J. Am. Chem. Soc., 121, 8661–8662.
- Shapiro, B.L., Kopchik, R.M. and Ebersole, S.J. (1963) *J. Chem. Phys.*, **39**, 3154–3155.
- Spiess, H.W. (1978) In NMR Basic Principles and Progress, Vol. 15 (Eds, Diehl, P., Fluck, E. and Kosfeld, R.), Springer-Verlag, Berlin, pp. 55–214.
- Stilbs, P. and Forsén, S. (1976) Org. Magn. Reson., 8, 384-386.
- Swaminathan, S., Craven, B.M. and McMullan, R.K. (1984) Acta Crystallogr., B40, 300–306.
- Vahtras, O., Ågren, H., Jørgensen, P., Jensen, H.J.A., Padkjær, S.B. and Helgaker, T. (1992) J. Chem. Phys., 96, 6120–6125.
- Wang, Y.-X., Jacob, J., Cordier, F., Wingfield, P., Stahl, S.J., Lee-Huang, S., Torchia, D., Grzesiek, S. and Bax, A. (1999) J. Biomol. NMR, 14, 181–184.
- Wasylishen, R.E. (1996) In Encyclopedia of Nuclear Magnetic Resonance (Eds, Grant, D.M. and Harris, R.K.), Wiley Inc., Chichester, pp. 1685–1695.