VIP

The Calculation of Indirect Nuclear Spin–Spin Coupling Constants in Large Molecules

Mark A. Watson,[b] Paweł Sałek,[c] Peter Macak,[d] Michał Jaszuński,[e] and Trygve Helgaker*[a]

Abstract: We present calculations of indirect nuclear spin–spin coupling constants in large molecular systems, performed using density functional theory. Such calculations, which have become possible because of the use of linear-scaling techniques in the evaluation of the Coulomb and exchange-correlation contributions to the electronic energy, allow us to study indirect spin–spin couplings in molecules of biological interest, without having to construct artificial model systems. In addition to presenting a statistical analysis of the large number of short-range coupling constants in large molecular systems, we analyse the asymptotic dependence of the indirect nuclear spin–spin coupling constants on the internuclear separation. In particular, we demonstrate that, in a sufficiently large one-electron basis set, the indirect spin–spin coupling constants become proportional to the inverse cube of the internuclear separation, even though the diamagnetic and paramagnetic spin-orbit contributions to the spin–spin coupling constants separately decay as the inverse square of this separation. By contrast, the triplet Fermi contact and spin-dipole contributions to the indirect spin–spin coupling constants decay exponentially and as the inverse cube of the internuclear separation, respectively. Thus, whereas short-range indirect spin–spin coupling constants are usually dominated by the Fermi contact contribution, long-range coupling constants are always dominated by the negative diamagnetic spin-orbit contribution and by the positive paramagnetic spin-orbit contribution, with small spin-dipole and negligible Fermi contact contributions.

Keywords: density functional calculations · NMR spectroscopy · response theory · spin–spin coupling constants

Introduction

Apart from the nuclear shielding constants, the indirect nuclear spin–spin coupling constants represent the most important source of information in high-resolution nuclear magnetic resonance (NMR) spectroscopy. Their importance is twofold. First, since the magnitude of the indirect spin–spin coupling constants depends critically on the nature of the molecular electron distribution, their measurement provides invaluable information on the chemical bonding of the coupled nuclei. Second, since the indirect spin–spin coupling constants are sensitive to the molecular geometry, they are extensively used for structural elucidation. In either case, the independent quantum-chemical calculation of indirect spin–spin coupling constants can play an important role in the interpretation of the measured coupling constants in terms of the electronic and geometric structure. However, until now, the rigorous calculation of spin–spin coupling constants has been restricted to small systems, containing much less than hundred atoms.[1–4] This restriction, which arises from the severe demands that such calculations place on the description of the electronic system, is particularly unfortunate since, nowadays, perhaps the greatest need for the reliable theoretical prediction of spin–spin coupling constants stems from a central application area of high-resolution NMR spectroscopy—namely, the elucidation of the native structures of biological molecules. In such studies, the indirect spin–spin coupling constants play a crucial role; the vicinal (three-bond) coupling constants, in particular, have for

[a] Prof. Dr. T. Helgaker
Department of Chemistry, University of Oslo
Box 1033 Blindern, 0315 Oslo (Norway)
Fax: (+47) 22855441
E-mail: trygve.helgaker@kjemi.uio.no

[b] Dr. M. A. Watson
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge, CB2 1EW (UK)

[c] Dr. P. Sałek
Laboratory of Theoretical Chemistry
The Royal Institute of Technology
Roslagstullsbacken 15, Stockholm 10691 (Sweden)

[d] Dr. P. Macak
The AlbaNova University Center, Institute of Physics
Stockholm 10691 (Sweden)

[e] Prof. Dr. M. Jaszuński
Institute of Organic Chemistry
Polish Academy of Sciences
01 224 Warszawa, Kasprzaka 44 (Poland)

 DOI: 10.1002/chem.200306065

a long time been used to determine dihedral angles. One-bond coupling constants, geminal (two-bond) coupling constants, and even long-range proton–proton coupling constants have also found widespread use in structure elucidation. Finally, the recent discovery of coupling constants transmitted through hydrogen bonds[5] (see also reference [3]) has made the spin–spin constants even more useful for the determination of the high-order structures of biopolymers.

Clearly, for quantum chemistry to make a substantial contribution to modern NMR structural chemistry, there is a need for the development of computational techniques that are not only accurate, but also applicable to large molecular systems. We here present such an approach. High accuracy and wide applicability is achieved by combining the techniques recently developed for the accurate calculation of spin–spin coupling constants in small molecules by Kohn–Sham density functional theory (DFT),[6–10] with linear-scaling techniques for the calculation of Coulomb interactions in large systems.[11–16] In this manner, we are now in a position to calculate the indirect spin–spin coupling constants in molecules that contain several hundred atoms in a routine fashion, with an accuracy and reliability similar to that in small molecules.

In small molecules, the accurate calculation of indirect nuclear spin–spin coupling constants has proved to be significantly more difficult than that of most other properties, in particular, that of the other NMR parameters, such as the nuclear magnetic shielding constants and the nuclear quadrupole coupling constants. First, in Ramsey's nonrelativistic theory, there are several distinct mechanisms contributing to the spin–spin coupling constants: the diamagnetic spin-orbit (DSO) term, the paramagnetic spin-orbit (PSO) term, the spin-dipole (SD) term, and the Fermi contact (FC) term.[17] Although the FC term often dominates the short-range coupling constants, none of these contributions can be neglected, increasing the programming and computational efforts over that of most other molecular properties, such as polarisabilities and shielding constants. Second, the FC and SD terms involve triplet perturbations, requiring a flexible description of the molecular electronic structure, not attained by restricted Hartree–Fock theory, for example. Third, for an accurate calculation of the FC contribution, it is necessary to provide a good description of the electron density at the nuclei, necessitating the use of large Gaussian basis sets. Nevertheless, a variety of techniques have over the years been developed and routinely used for the calculation of indirect nuclear spin–spin coupling constants in small to medium-sized molecules, the most popular ab initio methods being the second-order polarisation propagator approach (SOPPA) from 1987,[18–23] the multiconfigurational self-consistent field (MCSCF) approach from 1992[24–31] and the coupled-cluster approach from 1994[32–40] (for each of the methods, there are numerous other applications). In spite of these developments, DFT remains the only reliable method currently capable of treating systems that contain more than ten to twenty atoms and therefore applicable to systems of biological interest. Although the quality is not quite as high as that of the coupled-cluster singles-and-doubles (CCSD) approximation and in particular the coupled-cluster single-doubles-and-triples (CCSDT) approximation, comparisons indicate that the results obtained with DFT are usually quite accurate, in particular, for the coupling constants studied in this paper.[41,42]

For small and medium-sized molecular systems, the calculation of indirect spin–spin coupling constants by DFT is dominated by the evaluation of the Coulomb and exchange-correlation contributions to the total energy and to the linear transformations required for the iterative solution of the response equations. Such calculations cannot be extended to large systems unless the Coulomb and exchange-correlation contributions to the Kohn–Sham matrices, constructed in the course of the optimisation of the electronic energy, and the solution of the linear equations (to generate the perturbed densities) is carried out in a manner that does not scale rapidly with system size. In our implementation, the calculation of these contributions has been implemented in a linear-scaling manner such that, for sufficiently large systems, they no longer dominate the calculation of the coupling constants.

Results and Discussion

The Results and Discussion section is divided into three parts. We begin, by presenting the theory underlying our DFT calculations of indirect spin–spin coupling constants in large molecular systems. Then, in another section, we analyse the dependence of the various Ramsey contributions to the coupling constants on the separation between the coupled nuclei. In the final section we present calculations on the valinomycin and hexapeptide molecules. Since the number of indirect spin–spin coupling constants increases quadratically with the size of the system, even a moderately large molecule contains a very large number of spin–spin coupling constants, most of which are inaccessible to experiment. In addition to studying, in a statistical manner, those short-range coupling constants that can be measured experimentally, in the last section we also examine the behaviour of the indirect spin–spin coupling constants at large internuclear separations, in particular, the asymptotic behaviour of the DSO, PSO, FC and SD contributions. In this section, we also take the opportunity to examine and illustrate some well-known relationships of spin–spin coupling theory, such as the Dirac vector model[43] and the Karplus relation.[44,45] However, no attempts are made at comparisons with experiment, as this would require a careful consideration of molecular conformations and environmental effects, which is beyond the scope of the present article.

The theory of spin–spin coupling-constant calculations: In Ramsey's nonrelativistic theory, the elements of the reduced indirect nuclear spin–spin coupling tensor $\mathbf{K}_{KL}$ of nuclei K and L are given by the following sum-over-states expression of second-order perturbation theory [Eq. (1)]:[17,46]

$$\begin{aligned}\mathbf{K}_{KL} = &\langle 0|\mathbf{h}_{KL}^{DSO}|0\rangle - 2\sum_{n_S\neq 0}\frac{\langle 0|\mathbf{h}_K^{PSO}|n_S\rangle\langle n_S|(\mathbf{h}_L^{PSO})^T|0\rangle}{E_{n_S}-E_0}\\ &-2\sum_{n_T}\frac{\langle 0|\mathbf{h}_K^{FC}+\mathbf{h}_K^{SD}|n_T\rangle\langle n_T|(\mathbf{h}_L^{FC})^T+(\mathbf{h}_L^{SD})^T|0\rangle}{E_{n_T}-E_0}\end{aligned} \quad (1)$$

In Equation (1) the first summation is over all excited singlet states $|n_S\rangle$ of energy E_{n_S} and the second summation is over all triplet states $|n_T\rangle$ of energy E_{n_T}. In atomic units, the DSO, PSO, FC and SD operators are given by Equations (2)–(5).

$$\mathbf{h}_{KL}^{DSO} = \alpha^4 \sum_i \frac{\mathbf{r}_{iK}^T \mathbf{r}_{iL} \mathbf{I}_3 - \mathbf{r}_{iK} \mathbf{r}_{iL}^T}{r_{iK}^3 r_{iL}^3} \tag{2}$$

$$\mathbf{h}_{K}^{PSO} = -i\alpha^2 \sum_i \frac{\mathbf{r}_{iK} \times \boldsymbol{\nabla}_i}{r_{iK}^3} \tag{3}$$

$$\mathbf{h}_{K}^{FC} = \frac{8\pi\alpha^2}{3} \sum_i \delta(\mathbf{r}_{iK}) \mathbf{s}_i \tag{4}$$

$$\mathbf{h}_{K}^{SD} = \alpha^2 \sum_i \frac{3\mathbf{r}_{iK}^T \mathbf{s}_i \mathbf{r}_{iK} - r_{iK}^2 \mathbf{s}_i}{r_{iK}^5} \tag{5}$$

In these equations α ($\approx 1/137$) is the fine-structure constant and the summations are over all electrons. We note that whereas the DSO and PSO operators in Equations (2) and (3) are singlet operators, coupling the nuclear magnetic moments to the orbital motion of the electrons, the FC and SD operators in Equations (4) and (5) are triplet operators, coupling the nuclear moments to the spin of the electrons. Usually, the indirect nuclear spin–spin interactions are expressed not in terms of the reduced coupling tensor $\mathbf{K}_{KL}$ of Equation (1), but in terms of the indirect nuclear spin–spin coupling tensor [Eq. (6)] in which γ_K and γ_L are the magnetogyric ratios of the two nuclei, and h is Planck's constant.

$$\mathbf{J}_{KL} = h \frac{\gamma_K}{2\pi} \frac{\gamma_L}{2\pi} \mathbf{K}_{KL} \tag{6}$$

Moreover, for freely tumbling molecules as studied in high-resolution NMR spectroscopy, only the isotropic part of the coupling tensor is observed [Eq. (7)].

$$J_{KL} = {}^1/_3 \mathrm{Tr} \mathbf{J}_{KL} \tag{7}$$

Since the FC operator is isotropic and the SD operator anisotropic, there are no mixed FC/SD and SD/FC contributions to J_{KL}, which may therefore be uniquely decomposed into DSO, PSO, FC and SD contributions.

In the present paper, we are concerned with the evaluation of the indirect spin–spin coupling constants in large molecules by means of DFT. The evaluation of the spin–spin coupling constants does not follow Ramsey's sum-over-states expression, but rather the general scheme of response theory,[47] identifying the reduced coupling tensor $\mathbf{K}_{KL}$ of Equation (1) with the second derivative of the total electronic energy with respect to magnetic moments $\mathbf{M}_K$ and $\mathbf{M}_L$ of nuclei K and L.[1,4] We will outline this theory in the following two subsections: in the first we will discuss the DFT energy and in the second we comment on the calculation of spin–spin coupling constants, with emphasis on computational aspects related to large molecules. Our implementation of Kohn–Sham DFT for the evaluation of indirect nuclear spin–spin coupling constants is a modification of the DFT spin–spin implementation by Helgaker, Watson and Handy[9] in a development version of Dalton 1.2,[48] based on the original multiconfigurational self-consistent-field (MCSCF) spin–spin implementation by Vahtras and coworkers.[49]

The molecular electronic energy: In Kohn–Sham DFT, the total electronic energy of a closed-shell molecular system is given by Equation (8):

$$E = 2 \sum_i h_{ii} + 2 \sum_{ij} (g_{iijj} - \gamma g_{ijji}) + E_{XC}[\rho] + h_{nuc} \tag{8}$$

in which, in addition to the one- and two-electron Hamiltonian integrals [Eqs. (9) and (10)]

$$h_{pq} = -{}^1/_2 \int \phi_p^*(\mathbf{r}) \nabla^2 \phi_q(\mathbf{r}) \mathrm{d}\mathbf{r} - \sum_K Z_K \int \phi_p^*(\mathbf{r}) r_K^{-1} \phi_q(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{9}$$

$$g_{pqrs} = \int \int \phi_p^*(\mathbf{r}_1) \phi_q(\mathbf{r}_1) r_{12}^{-1} \phi_r^*(\mathbf{r}_2) \phi_s(\mathbf{r}_2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{10}$$

we have introduced the proportion of exact exchange γ, the exchange-correlation energy $E_{XC}[\rho]$ as a functional of the electron density ρ, and the nuclear-nuclear repulsion energy h_{nuc}. In Equation (8), the summations are over the doubly occupied molecular orbitals $\phi_i(\mathbf{r})$. Here and elsewhere, we use the convention that the indices i and j denote occupied orbitals, the indices a and b denote virtual orbitals and the indices p, q, r and s are used to represent unspecified (occupied or virtual) orbitals. In Equation (9), Z_K is the charge of nucleus K and r_K its distance to the electron.

In the following, we shall assume that the exchange-correlation functional can be written in the form given in Equation (11):

$$E_{XC}[\rho] = \int f(\rho_s(\mathbf{r}), \rho_t(\mathbf{r}), \varsigma_{ss}(\mathbf{r}), \varsigma_{st}(\mathbf{r}), \varsigma_{tt}(\mathbf{r})) \mathrm{d}\mathbf{r} \tag{11}$$

in which $\rho_s(\mathbf{r})$ and $\rho_t(\mathbf{r})$ are the density and spin density, respectively, which may be expressed in terms of the alpha and beta spin densities $\rho_\alpha(\mathbf{r})$ and $\rho_\beta(\mathbf{r})$ as Equation (12):

$$\begin{aligned} \rho_s(\mathbf{r}) &= \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r}) \\ \rho_t(\mathbf{r}) &= \rho_\alpha(\mathbf{r}) - \rho_\beta(\mathbf{r}) \end{aligned} \tag{12}$$

and where we have introduced the density-gradient scalar products [Eq. (13)]:

$$\begin{aligned} \varsigma_{ss}(\mathbf{r}) &= \boldsymbol{\nabla}\rho_s(\mathbf{r}) \cdot \boldsymbol{\nabla}\rho_s(\mathbf{r}) \\ \varsigma_{st}(\mathbf{r}) &= \boldsymbol{\nabla}\rho_s(\mathbf{r}) \cdot \boldsymbol{\nabla}\rho_t(\mathbf{r}) \\ \varsigma_{tt}(\mathbf{r}) &= \boldsymbol{\nabla}\rho_t(\mathbf{r}) \cdot \boldsymbol{\nabla}\rho_t(\mathbf{r}) \end{aligned} \tag{13}$$

This form of the exchange-correlation functional is appropriate for the generalised gradient approximation (GGA). In the simpler local-density approximation (LDA), there is no dependence on the density gradient in $E_{XC}[\rho]$, only on $\rho_s(\mathbf{r})$ and $\rho_t(\mathbf{r})$. The detailed form of the exchange-correlation functional depends on the particular DFT approximation that is made in the calculation.[50,51]

For the optimised unperturbed closed-shell system, the molecular orbitals satisfy the condition that Kohn–Sham matrix elements [Eq. (14)]

$$F_{pq} = h_{pq} + \sum_j (2g_{pqjj} - \gamma g_{pjjq}) + \int \left(\frac{\partial f}{\partial \rho_s} \phi_p \phi_q + 2 \frac{\partial f}{\partial \varsigma_{ss}} \nabla \rho_s \cdot \nabla \phi_p \phi_q \right) d\mathbf{r} \quad (14)$$

between occupied orbitals and virtual orbitals vanish: $F_{ai} = 0$. For such systems, the alpha and beta densities are identical and the spin density vanishes $\rho_t(\mathbf{r}) = 0$; when the system is perturbed, the spin density no longer vanishes.

In our implementation, the evaluation of the Coulomb part of the Kohn–Sham matrix [Eq. (14)] is carried out in a mixed manner. Whereas all contributions that involve overlapping charge distributions are evaluated by explicit integration as described in reference [52], using the McMurchie–Davidson scheme,[53] the contributions that involve non-overlapping charge distributions are evaluated by fast multipole techniques so that, for sufficiently large systems, the evaluation of the Coulomb part of the Kohn–Sham matrix scales linearly with system size. The new code is based on the original Dalton code, rewritten for the efficient evaluation of the Coulomb part of the Kohn–Sham matrix as described in reference [54].

The exchange-correlation code of Dalton has furthermore been rewritten and adapted for large systems, as described in reference [54]. In particular, the DFT integration of the exchange-correlation contribution to the Kohn–Sham matrix [Eq. (14)] is now carried out in an efficient, linear-scaling manner. Although only LDA calculations are presented here, the DFT integrator handles gradient-corrected functionals at little extra cost. However, no linear-scaling techniques have been implemented for the exact Hartree–Fock exchange with $\gamma \neq 0$. For large systems, therefore, we are at present restricted to pure DFT calculations.

The optimisation of the electronic energy is carried out in the standard self-consistent manner, but has been extended to include a second-order trust-region-based Newton method, based on the original MCSCF implementation of Jensen, Jørgensen and Ågren.[55] Since the optimisation of the energy is often more difficult for large than for small systems, the extension of the DFT optimisation to include the second-order method is important. The electronic Hessian needed for the second-order optimisation of the energy is given in the next section.

The calculation of spin–spin coupling tensors: From general considerations of response theory,[4] the indirect reduced spin–spin coupling tensors may be written in the form given in Equation (15):

$$\mathbf{K}_{KL} = \mathbf{K}_{KL}^{DSO} + \sum_{ai} \boldsymbol{\lambda}_{K,ai}^{PSO} (\mathbf{h}_{L,ai}^{PSO})^T + \sum_{ai} (\lambda_{K,ai}^{FC} \mathbf{I}_3 + \boldsymbol{\lambda}_{K,ai}^{SD})(h_{L,ai}^{FC} \mathbf{I}_3 + \mathbf{h}_{L,ai}^{SD})^T \quad (15)$$

which may be compared with Ramsey's expression [Eq. (1)]. In the following we shall discuss each of the terms in Equation (15), beginning with the DSO term.

The DSO contribution to the coupling tensor is a three-by-three matrix containing the expectation values given in Equation (16).

$$\mathbf{K}_{KL}^{DSO} = 2\alpha^4 \sum_i \int \phi_i(\mathbf{r}) r_K^{-3} r_L^{-3} (\mathbf{r}_K^T \mathbf{r}_L \mathbf{I}_3 - \mathbf{r}_K \mathbf{r}_L^T) \phi_i(\mathbf{r}) d\mathbf{r} \quad (16)$$

These values are calculated numerically by using the same integrator as for the exchange-correlation energy. Previously, this contribution was evaluated in a mixed analytical–numerical fashion that was very expensive for large systems;[49] in the present implementation, the DSO contribution requires only a fraction of the time needed for the calculation of indirect spin–spin coupling constants.

The remaining terms to the indirect nuclear spin–spin coupling tensor [Eq. (15)] correspond to the sum-over-states terms in Ramsey's expression [Eq. (1)] and contain, for each occupied-virtual orbital pair *ai*, the three-dimensional column vector $\mathbf{h}_{L,ai}^{PSO}$ [Eq. (17)], the scalar $h_{L,ai}^{FC}$ [Eq. (18)] and the three-by-three traceless symmetric matrix $\mathbf{h}_{L,ai}^{SD}$ [Eq. (19)]:

$$\mathbf{h}_{L,ai}^{PSO} = 2\alpha^2 \int \phi_a(\mathbf{r}) r_L^{-3} \mathbf{r}_L \times \nabla \phi_i(\mathbf{r}) d\mathbf{r} \quad (17)$$

$$h_{L,ai}^{FC} = \frac{8\pi\alpha^2}{3} \phi_a(\mathbf{R}_L) \phi_i(\mathbf{R}_L) \quad (18)$$

$$\mathbf{h}_{L,ai}^{SD} = \alpha^2 \int \phi_a(\mathbf{r}) r_L^{-5} (3\mathbf{r}_L \mathbf{r}_L^T - r_L^2 \mathbf{I}_3) \phi_i(\mathbf{r}) d\mathbf{r} \quad (19)$$

in which $\mathbf{R}_L$ is the position of nucleus L. The elements of $\lambda_{L,ai}^{PSO}$, $\lambda_{L,ai}^{FC}$, and $\lambda_{L,ai}^{SD}$ in Equation (15) represent the responses of the system to the nuclear magnetic moments and are obtained from the solution of ten sets of linear equations [Eq. (20)–(22)]:

$$\sum_{bj} {}^{II}G_{ai,bj}^{ss} \boldsymbol{\lambda}_{L,bj}^{PSO} = -\mathbf{h}_{L,ai}^{PSO} \quad (20)$$

$$\sum_{bj} {}^{RR}G_{ai,bj}^{tt} \lambda_{L,bj}^{FC} = -h_{L,ai}^{FC} \quad (21)$$

$$\sum_{bj} {}^{RR}G_{ai,bj}^{tt} \boldsymbol{\lambda}_{L,bj}^{SD} = -\mathbf{h}_{L,ai}^{SD} \quad (22)$$

The matrices on the left-hand sides of these equations are diagonal subblocks of the total electronic Hessian, reflecting the different symmetries of the perturbations (imaginary singlet for PSO and real triplet for FC and SD). If we let u represent either s (singlet) or t (triplet), the nonzero elements of the electronic Hessian are given by Equations (23) and (24):

$${}^{II}G_{ai,bj}^{uu} = \delta_{ij} F_{ab} - \delta_{ab} F_{ij} - \gamma(g_{abij} - g_{ajib}) \quad (23)$$

$$
\begin{aligned}
{}^{RR}G^{uu}_{ai,bj} = {} & \delta_{ij}F_{ab} - \delta_{ab}F_{ij} + 4\,\delta_{su}g_{aibj} - \gamma(g_{abij} + g_{ajib}) + \\
& \int \frac{\partial^2 f}{\partial \rho_u^2}\phi_a\phi_i\phi_b\phi_j \mathrm{d}\mathbf{r} + 2\int \frac{\partial f}{\partial \varsigma_{uu}}(\nabla\phi_a\phi_i)\cdot(\nabla\phi_b\phi_j)\mathrm{d}\mathbf{r} + \\
& \int \frac{\partial^2 f}{\partial \varsigma_{su}^2}[(\nabla\rho_s\cdot\nabla\phi_a\phi_i)(\nabla\rho_s\cdot\nabla\phi_b\phi_j)]\mathrm{d}\mathbf{r} + \\
& \int \frac{\partial^2 f}{\partial \rho_u \partial \varsigma_{su}}[(\nabla\rho_s\cdot\nabla\phi_a\phi_i)\phi_b\phi_j + \phi_a\phi_i(\nabla\rho_s\cdot\nabla\phi_b\phi_j)]\mathrm{d}\mathbf{r}
\end{aligned}
\tag{24}
$$

Whereas the imaginary blocks of the electronic Hessian [Eq. (23)] are the same in the singlet and triplet cases, the real blocks [Eq. (24)] are different in the two cases. We note that the real singlet Hessian—that is, Equation (24) with u = s—is needed for the second-order optimisation of the electronic energy. In a slightly different form, the expressions for the singlet and triplet electronic Hessians were first given by Bauernschmitt and Ahlrichs.[56]

The solution of the linear equations Equations (20)–(22) is carried out in a direct iterative manner, as described in reference [57]. The electronic Hessian is never constructed explicitly, neither for the solution of the response equations Equations (20)–(22), nor for the second-order optimisation of the electronic energy. Instead, we directly construct products of the electronic Hessian with trial vectors. The time-consuming steps of these transformations are the construction of the Kohn–Sham matrix (with a modified density) and the direct linear transformation of the trial vector with the DFT part of the electronic Hessian. We also note that, in canonical Kohn–Sham theory, the Kohn–Sham matrix is diagonal ($F_{pq} = \delta_{pq}\varepsilon_p$), implying that the leading (and usually dominant) contribution to the electronic Hessian, $\delta_{ij}F_{ab} - \delta_{ab}F_{ij}$, is diagonal as well. In pure DFT ($\gamma = 0$), the imaginary Hessian then becomes diagonal, making the solution of the PSO response equations [Eq. (20)] trivial.

To obtain all couplings to a given nucleus L, we need to solve the response equations [Eqs. (20)–(22)] only for that particular nucleus; for the remaining nuclei K ≠ L, it is sufficient to calculate the right-hand sides of Equations (17)–(19). This observation follows from Equation (15), recalling that the coupling tensors are symmetric in K and L. This approach is particularly important in large molecules, where we frequently are interested in the couplings to only a subset of all nuclei. However, if all couplings among all nuclei are required, then we must solve a total of $10n$ linear equations, where n is the number of magnetic nuclei in the molecule.

In our present implementation, the main difficulty with calculations of indirect nuclear spin–spin coupling constants in large molecules is the optimisation of the electronic energy rather than the solution of the linear equations. Because of the small HOMO–LUMO gap typical of large systems, the energy optimisation is frequently more difficult than in small systems and much work remains to be done to improve this part of the calculation. By contrast, the solution of the linear Equations (20)–(22) is rather straightforward and not more difficult than for small systems. Since the PSO, FC and SD operators are local, screening techniques are very effective at reducing the cost of the evaluation and convergence is usually achieved in about five iterations, with no iterations needed for the diagonal PSO equations in pure DFT.

The spatial dependence of spin–spin coupling constants: Among the four different mechanisms that contribute to the isotropic indirect spin–spin coupling, by far the most important is the FC mechanism. Formally, its importance stems from the large prefactor of $8\pi/3$ in Equation (4), which ensures that the FC contribution to the coupling constants often dominates over the DSO, PSO and SD contributions, at least for one-bond couplings. Nevertheless, none of the contributions can be a priori neglected, since each may become important in special cases, as we shall see in the following sections.

The spatial dependence of the FC, SD, PSO and DSO integrals: All four operators [Eqs. (2)–(5)] that contribute to the indirect nuclear spin–spin coupling constants are local with respect to the positions of the nuclei, making these constants small for large internuclear separations. The FC operator [Eq. (4)], in particular, is extreme in this respect, vanishing everywhere except at the nuclei. Indeed, we shall later see that, for sufficiently large separations, the FC contribution to the spin–spin coupling constants becomes negligible. Instead, long-range spin–spin coupling constants are usually dominated by the DSO and PSO contributions, which are often negligible at short distances.

To examine the distance dependence of the contributions to the spin–spin coupling constants in more detail, consider two spherical Gaussian functions with exponents a and b and centred at positions **A** and **B**, respectively [Eqs. (25) and (26)].

$$G_a(\mathbf{r}_A) = \exp(-ar_A^2), \quad \mathbf{r}_A = \mathbf{r} - \mathbf{A} \tag{25}$$

$$G_b(\mathbf{r}_B) = \exp(-br_B^2), \quad \mathbf{r}_B = \mathbf{r} - \mathbf{B} \tag{26}$$

Introducing the centre **P** of the Gaussian orbital product $G_a(\mathbf{r}_A)G_b(\mathbf{r}_B)$ and the reduced exponent μ as Equation (27)

$$
\begin{aligned}
\mathbf{P} &= \frac{a}{a+b}\mathbf{A} + \frac{b}{a+b}\mathbf{B} \\
\mu &= \frac{ab}{a+b}
\end{aligned}
\tag{27}
$$

we obtain for the operators in Equations (2)–(5) the following asymptotic dependence [Eqs. (28)–(31)] of the one-electron integral on the distances R_{PK} and R_{PL} from the orbital centre **P** to the nuclear centres K and L, respectively.

$$\langle G_a | h_K^{FC} | G_b \rangle \propto \exp(-\mu R_{PK}^2) \tag{28}$$

$$\langle G_a | h_K^{SD} | G_b \rangle \propto R_{PK}^{-3} \tag{29}$$

$$\langle G_a | h_{KL}^{DSO} | G_b \rangle \propto R_{PK}^{-2} R_{PL}^{-2} \tag{30}$$

$$\langle G_a | h_K^{PSO} | G_b \rangle \propto R_{PK}^{-2} \tag{31}$$

Placing the origin of the coordinate system at one of the nuclei K or L, we then find that the numerators in Ramsey's expression [Eq. (1)] or equivalently the right-hand sides [Eqs. (17)–(19)] of the corresponding response-theory expression [Eq. (15)] decay at least exponentially for the FC contribution, as R_{KL}^{-3} for the SD contribution, and as R_{KL}^{-2} for the DSO and PSO contributions.

If Slater functions had been used in place of Gaussians in Equations (28)–(31), the results would have been more complicated but essentially the same except that the FC integrals would then decay in an exponential manner. Indeed, in the important special case in which $\mathbf{A}=\mathbf{B}$, the only difference is that, for Slater functions, μR_{PK}^2 is replaced by $(a+b)R_{PK}$ in the FC expression [Eq. (28)]. Since, in the remainder of this section, we are only concerned with the slower decay of the DSO and PSO contributions, this difference does not affect our conclusions.

A somewhat broader question is how the asymptotic behaviour of the Gaussians (which are used in this paper) may affect the FC results presented in later sections. What matters, however, is not the behaviour of individual orbital products as reflected in Equation (28), but rather the behaviour of linear combinations of such products. Thus, to the extent that the basis set is complete, the long-range behaviour of the FC interaction is correctly represented by Gaussians. This condition is relatively easy to satisfy (in an approximate manner) in the interior of the molecule, where basis functions distributed on different nuclei may contribute to the description of the long-range FC interaction. Far away from the nuclei, on the other hand, the transmission of the FC interaction may be poorly described unless the basis set has been augmented by diffuse functions. In a later section, we shall see that the FC interaction decays exponentially in the internuclear separation, even when the electron density is expanded in Gaussians.

The cancellation of DSO and PSO contributions at large separations: Although the simple analysis given above indicates that the spin–spin couplings between two nuclei should decay as R_{KL}^{-2}, the decay may be even faster. First, we have not taken into account the long-range behaviour of the wave function, nor have we considered the possibility of cancellation between the different terms. Indeed, we shall now show that, in the limit of large separations R_{KL}, the DSO and PSO terms very nearly cancel, leading to an overall decay rate of at least R_{KL}^{-3}. To the best of our knowledge, the results presented in this section are new and not previously discussed in the literature.

Consider the isotropic parts of the DSO and PSO contributions to the indirect spin–spin coupling tensor [Eqs. (32) and (33), respectively], both of which decay as R_{KL}^{-2}:

$$K_{KL}^{DSO}=\frac{2\alpha^4}{3}\langle 0|\frac{\mathbf{r}_K}{r_K^3}\cdot\frac{\mathbf{r}_L}{r_L^3}|0\rangle=\mathcal{O}(R_{KL}^{-2}) \tag{32}$$

$$K_{KL}^{PSO}=-\frac{2\alpha^4}{3}\sum_{n\neq 0}\frac{\langle 0|\frac{\mathbf{l}_K}{r_K^3}|n\rangle\cdot\langle n|\frac{\mathbf{l}_L}{r_L^3}|0\rangle}{E_n-E_0}=\mathcal{O}(R_{KL}^{-2}) \tag{33}$$

Here, we have suppressed the summation over the electrons, and the angular momentum $\mathbf{l}_K$ with respect to centre K is related to the linear momentum $\mathbf{p}=-i\nabla$ as in Equation (34):

$$\mathbf{l}_K=\mathbf{r}_K\times\mathbf{p} \tag{34}$$

The same applies to $\mathbf{l}_L$. We shall now show that, for sufficiently large R_{KL}, the sum of the DSO and PSO contributions [Eqs. (32) and (33)] decays as R_{KL}^{-3} rather than as R_{KL}^{-2}. First, introducing the Taylor expansion [Eq. (35)]

$$\frac{\mathbf{r}_L}{r_L^3}=\frac{\mathbf{R}_{KL}}{R_{KL}^3}+\frac{\mathbf{I}_3R_{KL}^2-3\mathbf{R}_{KL}\mathbf{R}_{KL}^T}{R_{KL}^5}\mathbf{r}_K+\;\cdots\cdots \tag{35}$$

we note that the second factor in the numerator of Equation (33) can be written in the form given in Equation (36):

$$\langle n|\frac{\mathbf{l}_L}{r_L^3}|0\rangle=\langle n|\frac{\mathbf{R}_{KL}\times\mathbf{p}}{R_{KL}^3}|0\rangle+\mathcal{O}(R_{KL}^{-3}) \tag{36}$$

Next, using the identity $\mathbf{p}=i[H,\mathbf{r}]$, we obtain Equation (37) for this factor:

$$\langle n|\frac{\mathbf{l}_L}{r_L^3}|0\rangle=(E_n-E_0)\langle n|\frac{\mathbf{R}_{KL}\times i\mathbf{r}}{R_{KL}^3}|0\rangle+\mathcal{O}(R_{KL}^{-3}) \tag{37}$$

Inserting this result into Equation (33) and invoking the resolution of the identity, we find that, for sufficiently large internuclear separations, the PSO term may be written as the following expectation value [Eq. (38)]:

$$K_{KL}^{PSO}=-\frac{2\alpha^4}{3}\langle 0|\frac{\mathbf{l}_K}{r_K^3}\cdot\frac{\mathbf{R}_{KL}\times i\mathbf{r}}{R_{KL}^3}|0\rangle+\mathcal{O}(R_{KL}^{-3}) \tag{38}$$

which may be further rewritten as Equation (39):

$$K_{KL}^{PSO}=-\frac{2\alpha^4}{6}\langle 0|\frac{\mathbf{l}_K}{r_K^3}\cdot\frac{\mathbf{R}_{KL}\times i\mathbf{r}}{R_{KL}^3}-\frac{\mathbf{R}_{KL}\times i\mathbf{r}}{R_{KL}^3}\cdot\frac{\mathbf{l}_K}{r_K^3}|0\rangle+\mathcal{O}(R_{KL}^{-3}) \tag{39}$$

Noting the vector identity [Eq. (40)], which is valid provided that $\mathbf{b}$ is independent of $\mathbf{r}$, we find that Equation (39) may be rewritten in the form of Equation (41) in which we have used Equation (35) to establish the final relationship between the DSO and PSO contributions:

$$[\mathbf{a}(\mathbf{r})\times\nabla]\cdot(\mathbf{b}\times\mathbf{r})-(\mathbf{b}\times\mathbf{r})\cdot[\mathbf{a}(\mathbf{r})\times\nabla]=2\mathbf{a}(\mathbf{r})\cdot\mathbf{b} \tag{40}$$

$$K_{KL}^{PSO}=-\frac{2\alpha^4}{3}\langle 0|\frac{\mathbf{r}_K}{r_K^3}\cdot\frac{\mathbf{R}_{KL}}{R_{KL}^3}|0\rangle+\mathcal{O}(R_{KL}^{-3})=-K_{KL}^{DSO}+\mathcal{O}(R_{KL}^{-3}) \tag{41}$$

Adding together the DSO and PSO contributions to the spin–spin coupling constant, we thus obtain Equation (42).

$$K_{KL}^{DSO}+K_{KL}^{PSO}=\mathcal{O}(R_{KL}^{-3}) \tag{42}$$

This demonstrates that the total orbital contribution decays as R_{KL}^{-3}. In particular, we note that this is the same decay rate as for the spin-mediated contributions [Eq. (43)]:

$$K_{KL}^{FC} + K_{KL}^{SD} = \mathcal{O}(R_{KL}^{-3}) \quad (43)$$

We conclude that, for sufficiently large internuclear separations, the total indirect spin–spin contributions decay at least as fast as R_{KL}^{-3}. We note, however, that, for approximate wave functions and computational models, this result holds only in the limit of a complete one-electron basis, since we have assumed the validity of the relation $\mathbf{p} = i[H,\mathbf{r}]$.

Let us now consider the signs of the DSO and PSO contributions for large separations. From the expression for the DSO contribution to the indirect spin–spin coupling constants in Equation (32), we note that, for large separations R_{KL}, most of the contributions to the integral comes from points between the nuclei K and L. At such points, $\mathbf{r}_K \cdot \mathbf{r}_L < 0$, producing an overall negative DSO contribution to the couplings. Conversely, for small separations R_{KL}, most points give $\mathbf{r}_K \cdot \mathbf{r}_L > 0$, yielding an overall positive DSO contribution. From similar arguments, we find that the PSO contribution to the couplings will be mostly negative for small internuclear separations and positive for very large separations. We conclude that, for large separations Equations (44) and (45) apply:

$$K_{KL}^{DSO} \propto -R_{KL}^{-2} < 0, \; R_{KL} \gg 0 \quad (44)$$

$$K_{KL}^{PSO} \propto R_{KL}^{-2} > 0, \; R_{KL} \gg 0 \quad (45)$$

No such relationships are valid for the FC and SD contributions, which may both either be positive or negative for large distances. In the next section, we shall see examples of this behaviour.

The calculation of spin–spin coupling constants: To illustrate the behaviour of indirect spin–spin coupling constants in large molecules, we present in this section model calculations on the valinomycin and hexapeptide molecules. No attempts are here made at comparisons with experiment, which would require a careful consideration of molecular conformations and environmental effects. Rather, our purpose is to examine some general features of spin–spin coupling constants in large systems, not previously studied theoretically or experimentally. In addition, we shall take advantage of the vast number of spin–spin coupling constants produced by calculations on large systems to study, in a statistical manner, some features of spin–spin coupling constants in a more convincing manner than is possible in smaller molecules, for example, the dependence of the spin–spin coupling constants on the number of intervening bonds and on the internuclear separation, the Dirac vector model,[43] and the Karplus relation.[44,45]

In Figure 1, we have plotted, on a logarithmic scale and as a function of the internuclear separation, all reduced indirect nuclear spin–spin coupling constants that involve at least one carbon atom in valinomycin $C_{54}H_{90}N_6O_{18}$ depicted in Figure 2. All four contributions (FC, SD, DSO and PSO) discussed here and later have been calculated at the LDA

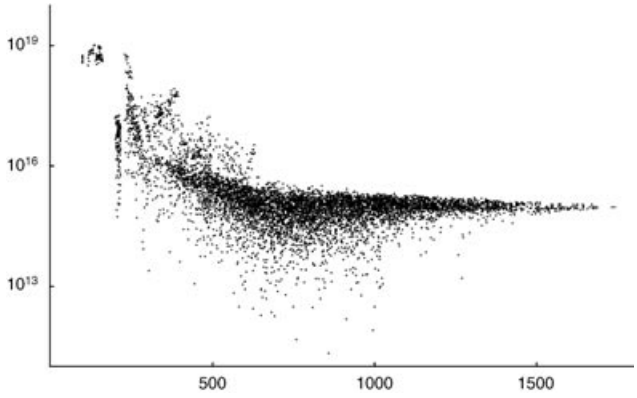

Figure 1. The absolute value of the reduced indirect spin–spin coupling constants ($NA^{-2}m^{-3}$) to all carbon atoms in valinomycin, plotted on a logarithmic scale as a function of the internuclear separation (pm). The spin–spin coupling constants have been calculated at the LDA/6-31G level of theory.

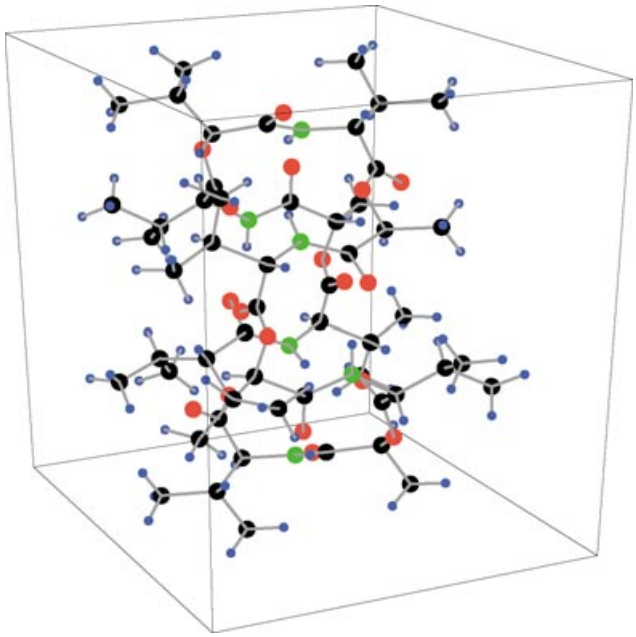

Figure 2. The valinomycin molecule with hydrogen atoms in blue, carbon atoms in black, oxygen atoms in red, and nitrogen atoms in green.

level of theory[58,59] in the 6-31G basis,[60] at an experimentally derived geometry and for the isotopic species ^1H, ^{13}C, ^{14}N and ^{17}O. While the use of a more accurate DFT model and a larger basis is certainly feasible for valinomycin and even larger systems, the biggest problem with calculations of spin–spin coupling constants in such molecules is the potentially large number of conformational structures that contribute. The valinomycin calculation presented here should therefore be viewed as a model calculation. Nevertheless, as we shall see, such calculations can teach us much about the behaviour of indirect nuclear spin–spin coupling constants in large systems.

There are a total of 7587 unique spin–spin coupling constants to carbon in valinomycin (4860 CH couplings, 1431 CC couplings, 972 CO couplings, and 324 CN cou-

plings). However, most of these spin–spin couplings can neither be observed nor assigned experimentally; for example, at the LDA/6-31G level, only 497 couplings are larger than 1 Hz and only 1163 are larger than 0.1 Hz in magnitude. In Figure 1, the observable, large one-bond, geminal and vicinal coupling constants are located to the left, whereas the unobservable, small long-range coupling constants are located to the right.

In the following, we shall first consider the short-range spin–spin coupling constants of valinomycin in the next subsection, followed by the consideration of the long-range couplings of valinomycin and the smaller hexapeptide molecule. We will conclude by considering the Dirac vector model and the Karplus relation in the valinomycin molecule.

Table 1. Statistical measures and relative contributions of LDA/6-31G indirect spin–spin coupling constants in valinomycin. The mean absolute (mabs), mean, minimum and maximum values are given in Hz, the relative contributions are percentages of the total spin–spin coupling constants; n is the number of intervening bonds and N the number of occurrences of this n-bond coupling.

atoms	n	N	mabs	mean	min	max	FC	SD	DSO	PSO
CH	1	84	131.0	131.0	111.5	172.4	98.1	0.4	1.0	0.5
	2	126	1.8	−0.7	−4.1	5.5	81.7	3.1	30.4	−15.2
	3	174	3.2	3.1	−0.3	10.3	152.9	0.7	−73.5	19.9
	4	114	0.3	0.0	−0.5	1.5	22.4	7.1	92.6	−22.1
	5	144	0.1	0.1	−0.2	0.7	31.3	2.3	63.1	3.4
CC	1	42	47.1	47.1	30.4	73.3	97.9	1.8	0.8	−0.6
	2	51	1.1	0.4	−1.7	4.2	100.3	1.5	3.2	−4.9
	3	54	1.8	1.8	−0.2	5.3	385.3	118.6	−991.4	587.6
	4	54	0.2	0.0	−0.5	0.5	103.9	−1.6	3.4	−5.7
	5	72	0.1	0.1	0.0	0.4	212.6	46.1	−234.8	76.1
CN	1	12	11.6	11.6	7.4	16.2	107.7	2.1	1.2	−11.1
	2	18	2.0	1.7	−0.7	6.4	86.0	5.2	−3.1	11.9
	3	18	0.6	0.6	−0.1	1.8	39.9	43.0	−8.8	26.0
	4	18	0.1	0.0	−0.2	0.2	91.9	−2.7	26.5	−15.8
	5	18	0.0	0.0	−0.1	0.0	80.4	0.0	25.5	−5.8
CO	1	24	27.5	27.5	13.1	43.0	77.3	−3.8	−1.0	27.4
	2	30	7.0	−7.0	−24.7	−0.7	117.0	2.7	2.5	−22.2
	3	36	0.7	−0.7	−2.5	0.2	159.7	−1.1	18.4	−76.9
	4	78	0.2	−0.2	−1.5	0.1	75.0	10.9	19.5	−5.4
	5	48	0.1	0.0	−0.4	0.0	83.2	−3.9	18.6	2.2

Short-range coupling constants in valinomycin: In Figure 3, we have plotted, on a logarithmic scale as a function of the internuclear separation, all calculated indirect spin–spin coupling constants that are greater than 0.1 Hz in magnitude. Since there are many such coupling constants and since they cannot individually be compared with experiment, they have not been tabulated here. Instead, we have listed some statistical data on the spin–spin coupling constants in valinomycin for up to five intervening bonds in Table 1. This table also contains information on the relative importance of the FC, SD, DSO and PSO contributions to the total spin–spin coupling constants, calculated by averaging over all calculated spin–spin coupling constants. Note that these statistical data are based on large numbers of individual coupling constants.

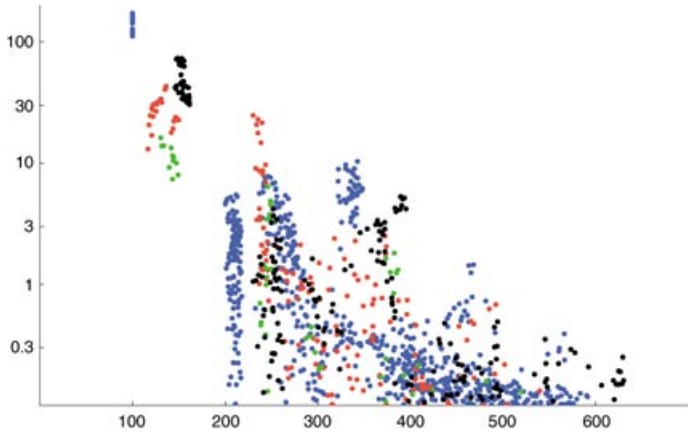

Figure 3. The absolute value of the indirect spin–spin coupling constants (Hz) in valinomycin greater than 0.1 Hz, on a logarithmic scale as a function of the internuclear distance (pm). We have used blue, black, red and green for the CH, CC, CO and CN coupling constants, respectively. The spin–spin coupling constants have been calculated at the LDA/6-31G level of theory.

The largest coupling constants in valinomycin are the one-bond coupling constants, with mean absolute values of 131, 47, 12 and 28 Hz for $^1J_{CH}$, $^1J_{CC}$, $^1J_{CN}$ and $^1J_{CO}$, respectively. However, as seen from the upper left corner of Figure 3 up to 162 pm, there are large variations among these couplings—the largest one-bond coupling constants are 60% larger and the smallest 65% smaller than the mean values. All one-bond coupling constants are positive. We also note that, whereas the $^1J_{CH}$ and $^1J_{CC}$ couplings are completely dominated by the FC contribution, the $^1J_{CN}$ and $^1J_{CO}$ couplings have sizable PSO contributions.

As seen from Figure 3 between 200 and 262 pm as well as from Table 1, the geminal coupling constants in valinomycin are much smaller than the one-bond coupling constants. Specifically, for the CH, CC, CN, and CO interactions, the mean absolute values of the geminal couplings constitute 1.4%, 2.3%, 17%, and 25%, respectively, of the corresponding one-bond couplings. As expected from the Dirac vector model,[43] the geminal couplings are predominantly negative (vide infra). However, the spans of the coupling constants are large: −4.1 to 5.5 Hz for $^2J_{CH}$, −1.7 to 4.2 Hz for $^2J_{CC}$, −0.7 to 6.4 for $^2J_{CN}$ and −24.7 to −0.7 for $^2J_{CO}$. Like the one-bond coupling constants, the geminal coupling constants are dominated by the FC contribution, although we note that the $^2J_{CH}$ constants have large DSO and PSO contributions of opposite signs.

The vicinal coupling constants, which are located between 237 to 395 pm in Figure 3, are about as large as the geminal couplings, in particular, for CH and CC. Thus, the mean absolute values of $^3J_{CC}$ and $^3J_{CH}$ are 3.2 and 1.8 Hz, respectively, whereas the corresponding values for $^2J_{CC}$ and $^2J_{CH}$ are 1.8 and 1.1 Hz, respectively; also, the spans of the CC and CH vicinal couplings are larger than those of the corresponding geminal couplings. By contrast, for CN and CO,

the vicinal couplings are much smaller than the geminal ones. Following the Dirac vector model, most of the vicinal coupling constants are positive. Finally, unlike the one- and two-bond coupling constants, the vicinal coupling constants are not dominated by the FC contribution. Indeed, except for the small SD contributions to the $^3J_{CH}$ and $^3J_{CO}$ couplings, all four contributions are large. For the vicinal $^3J_{CC}$ couplings, in particular, there are large contributions of opposite signs from the DSO and PSO operators.

From Table 1, we note that the long-range four- and five-bond couplings are mostly smaller than 1 Hz, the largest $^4J_{CH}$ coupling being 1.5 Hz. However, we shall see later that some of the couplings larger than 1 Hz in Figure 3 are very long-range in the sense that the two coupled nuclei are only remotely bonded, through more than ten bonds. Finally, except for the $^4J_{CC}$ couplings, the long-range spin–spin coupling constants are not dominated by the FC contribution. In the next subsection, we shall study the long-range behaviour of the spin–spin coupling constants in more detail.

The dependence of the coupling constants on the internuclear separations: From Figure 1, we see that the indirect nuclear spin–spin coupling constants initially decay fast (exponentially) with increasing separation between the coupled nuclei. However, beyond about 500 pm, the decay is much slower. Indeed, for large separations, the decay of the spin–spin coupling constants is surprisingly slow.

To understand the reason for this behaviour, we have in Figure 4 plotted separately the negative and positive FC, SD, DSO and PSO contributions to the reduced spin–spin coupling constants. (Reduced coupling constants are used to ensure that the signs of the contributions are not arbitrary.) From these plots we see that whereas the FC contribution dominates the couplings for small distances, it decays much faster than the other contributions with respect to the separation of the coupled nuclei. Indeed, as predicted by our simple analysis above, the FC contribution decays exponentially with increasing separation. Note that all FC couplings occupy a rather broad diagonal band, which extends over approximately three orders of magnitude. For sufficiently large separations, there appears to be a random distribution of positive and negative FC contributions.

The SD contribution is smaller than the FC contribution for small internuclear separations, but decays more slowly (as R_{KL}^{-3}) and eventually dominates over the FC contribution for large separations, see Figure 4. However, beyond about 500 pm, all spin–spin coupling constants are completely dominated by the DSO and PSO contributions, as expected from the R_{KL}^{-3} decay of the SD contribution and the R_{KL}^{-2} decay of the DSO and PSO contributions.

The plots in Figure 4 confirm the predicted signs of Equations (44) and (45): for small separations, the DSO contributions are predominantly positive and the PSO contributions negative; conversely, for large separations, the DSO contributions are negative and the PSO contributions positive. However, from a careful inspection of Figure 4 it follows that, for large separations, the DSO contributions are slightly larger than the PSO contributions in magnitude, indicating that the cancellation between the DSO and PSO contributions is incomplete in the small 6-31G basis. As a result, the sum of the DSO and PSO contributions and hence the total spin–spin coupling constants become negative at large distances. The cancellation of these terms at large separations is discussed in greater detail below.

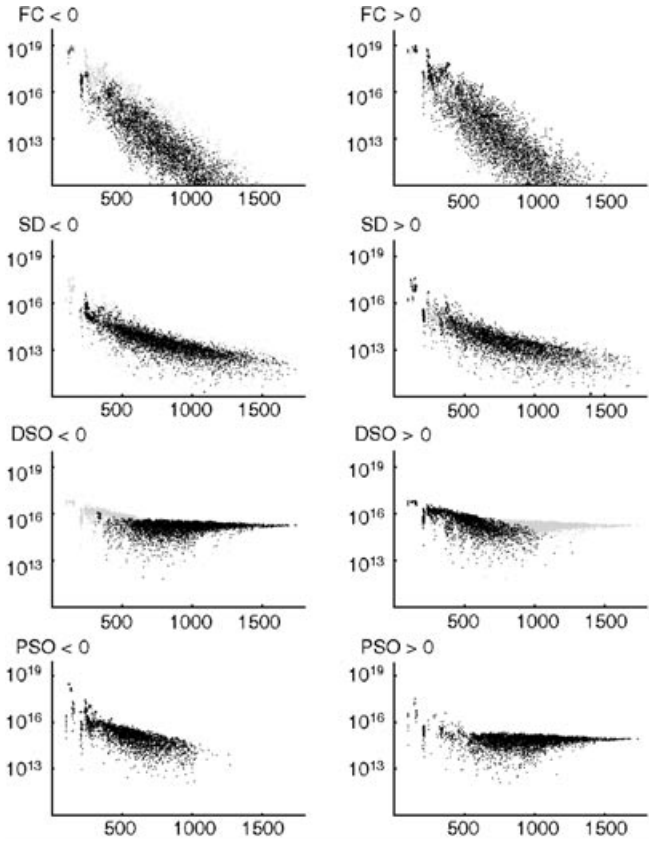

Figure 4. The absolute values of the FC, SD, DSO and PSO contributions to the reduced LDA/6-31G spin–spin coupling constants ($NA^{-2}\,m^{-3}$) to all carbon atoms in valinomycin, plotted on a logarithmic scale as a function of the internuclear separation (pm). Negative contributions are plotted on the left, positive contributions on the right (overlaid on the contributions of the opposite sign).

The dependence of the spin–spin coupling constants on the number of intervening bonds: Having discussed the decay of the coupling constants with increasing internuclear separation, let us now consider how the coupling constants depend on the number of intervening bonds. In Figure 5, we have plotted the maximum, mean, and minimum absolute spin–spin couplings in valinomycin as functions of the number of intervening bonds and on the internuclear separation.

Although the two plots of the mean spin–spin coupling constants in Figure 5 are quite similar, there are some striking differences between the plots of maximum couplings constants. Whereas the plotted maximum coupling constant decreases monotonically with increasing separation in the bottom view, it exhibits several peaks as a function of the number of intervening bonds in the top view. In particular, the peaks at 11, 13, and 15 bonds correspond to the long-range couplings $^{11}J_{CO} = -2.4$, $^{13}J_{CO} = -1.1$ and $^{15}J_{CO} = -2.3$ Hz. However, these large long-range interactions do not arise as a result of special, unusually large couplings

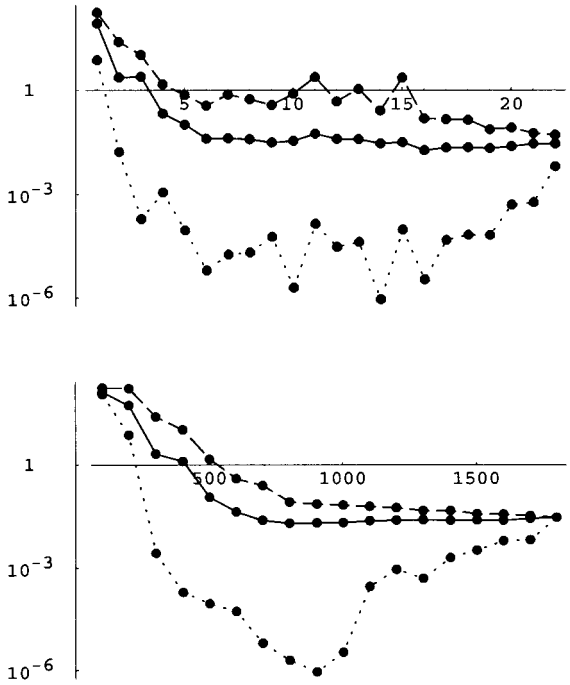

Figure 5. The maximum (dashed line), mean (full line) and minimum (dotted line) absolute spin–spin coupling constants in valinomycin (Hz) as a function of the number of intervening bonds (top) and the internuclear separation (pm; bottom). The spin–spin coupling constants have been calculated at the LDA/6-31G level of theory.

through a large number of covalent bonds. Rather, they arise as a result of short-range through-space interactions, the respective internuclear separations being only 318, 343 and 340 pm. Finally, we note that all long-range couplings larger than 1 Hz are dominated by the FC interaction (−2.3, −1.0 and −2.2 Hz, respectively). In short, large long-range couplings arise only when the coupled nuclei are sufficiently close for the FC interaction to contribute significantly in a through-space manner.

Basis-set convergence of the long-range spin–spin coupling constants: To illustrate the basis-set convergence of the long-range spin–spin coupling constants, we have carried out calculations on the hexapeptide $C_{16}O_6N_6H_{28}$ model system shown in Figure 6, using different basis functions. The hexapeptide model is much smaller than valinomycin but just as long, making it suitable for studies of long-range coupling constants.

In Figure 7, we have plotted the different contributions to the LDA spin–spin couplings involving at least one of the end H, C, N and O atoms in the hexapeptide model, illustrating the different asymptotic behaviour of these contributions to the spin–spin coupling constants in the large HII basis.[61,62] We note in particular the near coincidence of the points representing the DSO and PSO contributions, indicating near convergence of these contributions with respect to the basis set.

To exhibit the basis-set convergence in a clearer manner, we have in Figure 8 plotted the maximum absolute values of the different contributions to the spin–spin coupling constants in the hexapeptide model for different internuclear separations. Each maximum value has been obtained from a comparison of all coupling constants in intervals of 200 pm. We note the extreme sensitivity of the PSO contribution to the quality of the basis set. As the basis set increases in the

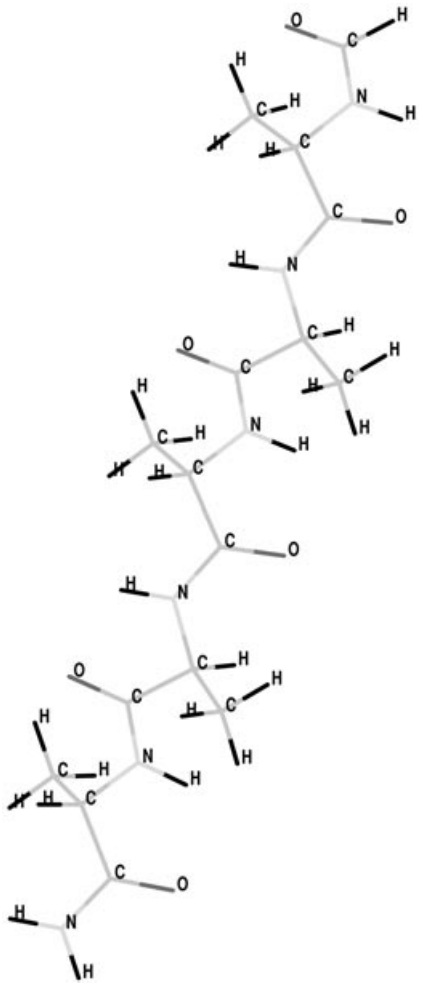

Figure 6. The model hexapeptide molecule. The indirect spin–spin coupling constants have been calculated for the top H, C, N and O atoms.

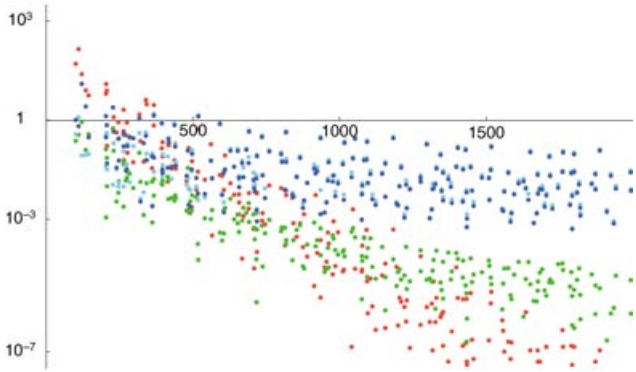

Figure 7. The FC (red), SD (green), DSO (light blue), and PSO (dark blue) contributions to all indirect spin–spin coupling constants involving at least one of the H, C, N and O end atoms in the hexapeptide model, calculated at the LDA/HII level of theory. The absolute values of the contributions (Hz) have been plotted on a logarithmic scale, as a function of the internuclear separation (pm).

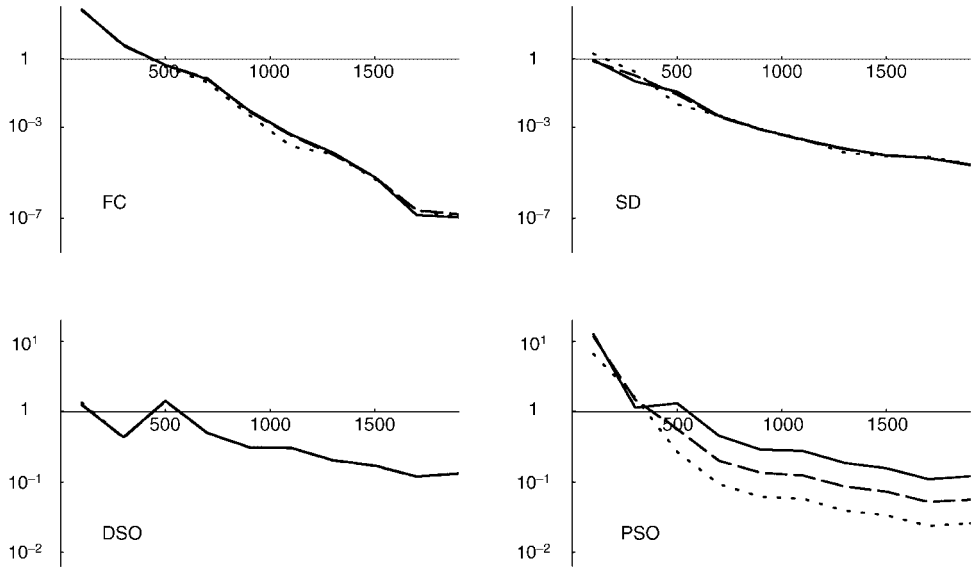

Figure 8. The convergence of the maximum absolute values (over 200 pm intervals) of the FC, SD, DSO and PSO contributions to the indirect spin–spin coupling constants (Hz) involving at least one of the H, C, N and O end atoms in the hexapeptide model, plotted as functions of the internuclear separation (pm). The calculations have been carried out at the LDA level in three basis sets: STO-3G (dotted line), 6-31G (dashed line), and HII (full line).

sequence STO-3G,[63] 6-31G[60] and HII,[61,62] the (positive) PSO contributions increase by almost an order of magnitude. By comparison, the basis-set dependence of the remaining contributions is insignificant at long separations.

As already established, the long-range spin–spin coupling constants are dominated by negative DSO and positive PSO contributions. Moreover, whereas each of these contributions separately decays as R_{KL}^{-2}, their sum converges as R_{KL}^{-3}. However, because of the extreme sensitivity of the long-range PSO contribution to the basis set, this faster convergence of the total spin–spin coupling constants is difficult to achieve in a finite basis.

In Figure 9, we have plotted, in a manner similar to that in Figure 8, the maximum total spin–spin coupling constants in the hexapeptide model in intervals of 200 pm. As the basis set increases from STO-3G to HII, the (negative) long-range total spin–spin coupling constants decrease by about an order of magnitude as the description of the positive PSO contribution becomes more accurate. Clearly, in our present LDA/6-31G calculations of the long-range coupling constants of valinomycin, we cannot claim to have reached basis-set saturation with respect to the PSO contribution to the spin–spin coupling constants.

We note that for both molecules studied here (valinomycin and hexapeptide), we observe the theoretically predicted asymptotic long-range behaviour of the spin–spin coupling constants discussed earlier. These predictions were made without any consideration of molecular structure, analysing only the dependence of the one-electron integrals on the internuclear distance. Thus, at smaller distances we can expect significant deviations from these simple predictions. In particular, coupling constants across a few bonds—say, five to seven bonds—may sometimes be observed because they are much larger than expected from the distance between the coupled nuclei. Such large long-range couplings arise because of special bonding situations not encountered here.

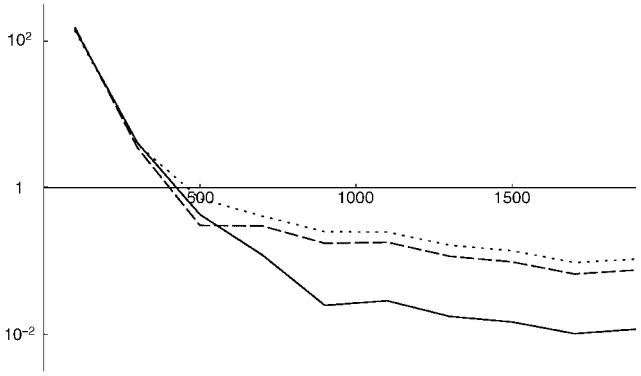

Figure 9. The convergence of the maximum absolute values (averaged over 200 pm intervals) of the total indirect spin–spin coupling constants (Hz) involving at least one of the H, C, N and O end atoms in the hexapeptide model, plotted as functions of the internuclear separation (pm). The calculations have been carried out at the LDA level of theory in three basis sets: STO-3G (dotted line), 6-31G (dashed line), and HII (full line).

The Dirac vector model: Let us now consider the dependence of the sign of the coupling constants on the number of intervening bonds. In Figure 10, we have plotted the proportion of positive reduced spin–spin coupling constants in valinomycin as a function of the number of intervening bonds and as a function of the internuclear separation. For small separations, we observe the well-known tendency of alternating signs for one-bond, geminal, and vicinal coupling constants. Beyond five- or six-bond couplings, this alternation disappears as the FC contribution becomes less important until, at large distances, all couplings become negative, as expected from our discussion above. No such alternation is observed when the proportion of positive reduced coupling constants is plotted as a function of the separation between the nuclei. Indeed, this behaviour is expected from the Dirac vector model,[43] which relates the relative signs of

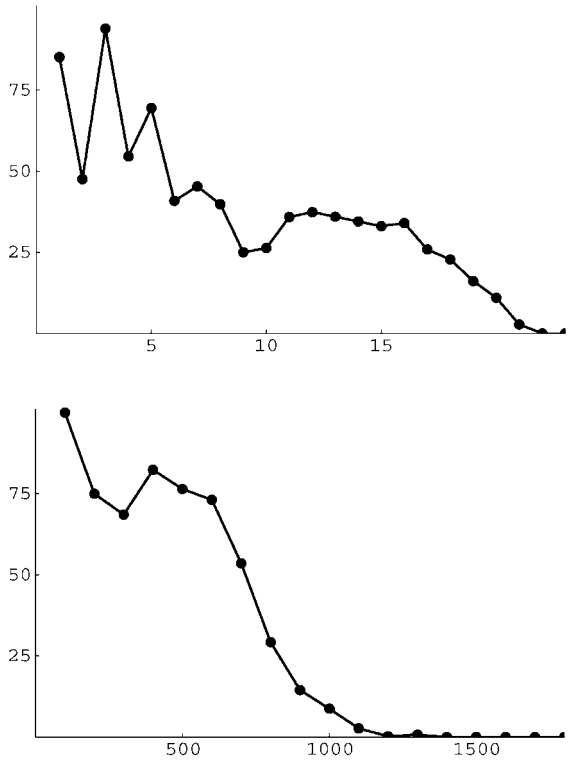

Figure 10. The percentage of positive reduced indirect spin–spin couplings in valinomycin as a function of the number of intervening bonds (top) and the internuclear separation (pm; bottom). The spin–spin coupling constants have been calculated at the LDA/6-31G level of theory.

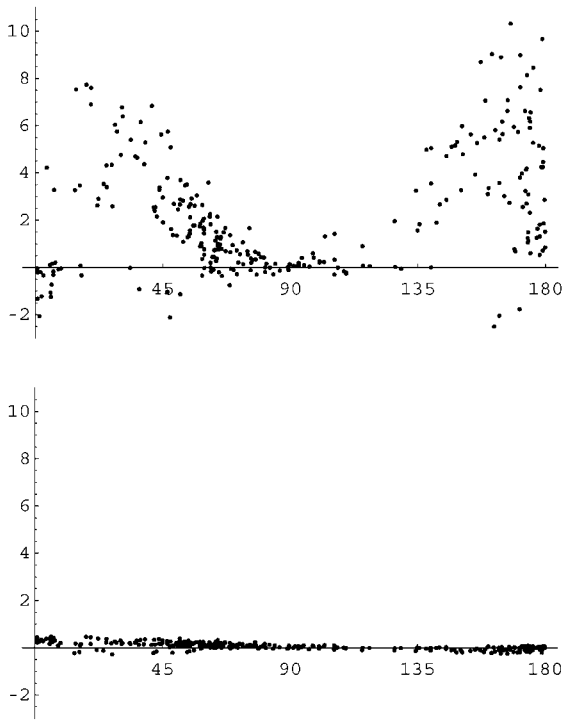

Figure 11. The vicinal coupling constants in valinomycin (Hz) involving at least one carbon atom as a function of the dihedral angle of the coupled nuclei. Top: the total spin–spin couplings; bottom: the non-FC contributions. The spin–spin coupling constants have been calculated at the LDA/6-31G level of theory.

the coupling constants to the number of intervening bonds rather than to the internuclear separation. We note here that this alternation is restricted to the FC term, which dominates short-range couplings. No such alternation is observed for the other contributions to the coupling constants.

The predominance of negative couplings for large separations in Figure 10 follows from the dominance of the DSO contribution at large separations as observed in Figure 4. It is possible that, in a large basis, when a more complete cancellation of the DSO and PSO terms is achieved, there will be an even distribution of positive and negative coupling constants at large separations.

Vicinal coupling constants and the Karplus relation: Among the 7587 coupling constants that we have calculated in valinomycin, there are a total of 282 vicinal couplings that involve one or two carbon atoms. In Figure 11, we have plotted all these couplings as functions of the dihedral angle associated with the coupled atoms. This figure clearly demonstrates that such couplings always vanish for dihedral angles close to 90°. For angles different from 90°, there is a large variation in the magnitude of the coupling constants. In part, this variation arises from the fact that, in the figure, we have combined all types of vicinal couplings, regardless of the type of coupled nuclei and the type of intervening nuclei. As can be seen from the separate plot for the non-FC contributions to these constants, the Karplus curve arises solely from the FC contribution to the vicinal coupling constants.

Conclusions

We have demonstrated that indirect nuclear spin–spin coupling constants now can be calculated for large molecular systems, containing more than 100 atoms. Such calculations have been carried out for the valinomycin and hexapeptide molecules and used to study the behaviour of long-range coupling constants. Whereas short-range coupling constants are often dominated by the FC term, long-range coupling constants are dominated by the DSO and PSO terms, which separately decay in a manner inversely proportional to the square of the distance between the coupled nuclei, but taken together decay in a manner inversely proportional to the cube of this distance. At large separations, the DSO and PSO contributions to the reduced coupling constants become negative and positive, respectively. However, because of a very slow basis-set convergence of the PSO term at large internuclear separations, this faster decay is only observed with large basis sets. The FC and SD contributions to the indirect spin–spin coupling constants decay exponentially and inversely proportionally to the cube of the internuclear distance, respectively. In large molecular systems, indirect spin–spin couplings larger than 1 Hz may occur for atoms separated by more than ten bonds if these atoms are

close together, because of strong through-space FC interactions.

Acknowledgments

This work was supported by the EU training and research network MOLPROP ("Molecular Properties and Molecular Materials", Contract No. HPRN-CT-2000-00013), by the Norwegian Research Council through a travel grant to M.J. (Grant No. 141932), and by the Norwegian Program for Supercomputing (nn1118k). We are indebted to Dr. K. Wolinski, who provided the geometry of the hexapeptide molecule (from the Parallel Quantum Solutions data base).

[1] T. Helgaker, M. Jaszuński, K. Ruud, *Chem. Rev.* **1999**, *99*, 293.
[2] J. Vaara, J. Jokisaari, R. E. Wasylishen, D. L. Bryce, *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 233.
[3] R. H. Contreras, V. Barone, J. C. Facelli, J. E. Peralta, *Annu. Rep. NMR Spectrosc.* **2003**, *51*, 167.
[4] T. Helgaker, M. Pecul, in *Calculation of NMR and EPR parameters, Theory and Applications* (Eds.: M. Kaupp, M. Bühl, V. Malkin), Wiley-VCH, Weinheim, **2004**, p. 101.
[5] A. J. Dingley, F. Cordier, S. Grzesiek, *Concepts Magn. Reson.* **2001**, *13*, 103.
[6] V. G. Malkin, O. L. Malkina, D. R. Salahub, *Chem. Phys. Lett.* **1994**, *221*, 91.
[7] R. M. Dickson, T. Ziegler, *J. Phys. Chem.* **1996**, *100*, 5286.
[8] V. Sychrovský, J. Gräfenstein, D. Cremer, *J. Chem. Phys.* **2000**, *113*, 3530.
[9] T. Helgaker, M. Watson, N. C. Handy, *J. Chem. Phys.* **2000**, *113*, 9402.
[10] J. Autschbach, T. Ziegler, *J. Chem. Phys.* **2000**, *113*, 936.
[11] M. C. Strain, G. E. Scuseria, M. J. Frisch, *Science* **1996**, *271*, 51.
[12] E. Schwegler, M. Challacombe, *J. Chem. Phys.* **1996**, *105*, 2726.
[13] C. A. White, B. G. Johnson, P. M. W. Gill, M. Head-Gordon, *Chem. Phys. Lett.* **1996**, *253*, 268.
[14] M. Challacombe, E. Schwegler, *J. Chem. Phys.* **1997**, *106*, 5526.
[15] E. Schwegler, M. Challacombe, M. Head-Gordon, *J. Chem. Phys.* **1997**, *106*, 9708.
[16] C. Ochsenfeld, C. A. White, M. Head-Gordon, *J. Chem. Phys.* **1998**, *109*, 1669.
[17] N. F. Ramsey, *Phys. Rev.* **1953**, *91*, 303.
[18] J. Geertsen, J. Oddershede, G. E. Scuseria, *J. Chem. Phys.* **1987**, *87*, 2138.
[19] R. D. Wigglesworth, W. T. Raynes, S. P. A. Sauer, J. Oddershede, *Mol. Phys.* **1998**, *94*, 851.
[20] S. P. A. Sauer, C. K. Møller, H. Koch, I. Paidarová, V. Špirko, *Chem. Phys.* **1998**, *238*, 385.
[21] R. D. Wigglesworth, W. T. Raynes, S. Kirpekar, J. Oddershede, S. P. A. Sauer, *J. Chem. Phys.* **2000**, *112*, 3735.
[22] P. F. Provasi, G. A. Aucar, S. P. A. Sauer, *J. Chem. Phys.* **2001**, *115*, 1324.
[23] L. B. Krivdin, S. P. A. Sauer, J. E. Peralta, R. H. Contreras, *Magn. Reson. Chem.* **2002**, *40*, 187.
[24] O. Vahtras, H. Ågren, P. Jørgensen, H. J. Aa. Jensen, S. B. Padkjær, T. Helgaker, *J. Chem. Phys.* **1992**, *96*, 6120.
[25] A. Barszczewicz, M. Jaszuński, K. Kamieńska-Trela, T. Helgaker, P. Jørgensen, O. Vahtras, *Theor. Chim. Acta* **1993**, *87*, 19.
[26] T. Helgaker, M. Jaszuński, K. Ruud, A. Górska, *Theor. Chem. Acc.* **1998**, *99*, 175.
[27] M. Pecul, J. Sadlej, *Chem. Phys.* **1998**, *234*, 111.
[28] P.-O. Åstrand, K. Ruud, K. V. Mikkelsen, T. Helgaker, *J. Chem. Phys.* **1999**, *110*, 9463.
[29] P. Lantto, J. Vaara, *J. Chem. Phys.* **2001**, *114*, 5482.
[30] M. Jaszuński, K. Ruud, *Chem. Phys. Lett.* **2001**, *336*, 473.
[31] J. Casanueva, J. San Fabián, E. Diéz, A. L. Esteban, *Chem. Phys. Lett.* **2002**, *361*, 159.
[32] S. A. Perera, H. Sekino, R. J. Bartlett, *J. Chem. Phys.* **1994**, *101*, 2186.
[33] H. Sekino, R. J. Bartlett, *Chem. Phys. Lett.* **1994**, *225*, 486.
[34] S. A. Perera, M. Nooijen, R. J. Bartlett, *J. Chem. Phys.* **1996**, *104*, 3290.
[35] S. A. Perera, R. J. Bartlett, *J. Am. Chem. Soc.* **1996**, *118*, 7849.
[36] J. E. Del Bene, R. J. Bartlett, *J. Am. Chem. Soc.* **2000**, *122*, 10480.
[37] J. E. Del Bene, M. J. T. Jordan, S. A. Perera, R. J. Bartlett, *J. Phys. Chem. A* **2001**, *105*, 8399.
[38] A. A. Auer, J. Gauss, *J. Chem. Phys.* **2001**, *115*, 1619.
[39] A. Wu, D. Cremer, A. A. Auer, J. Gauss, *J. Phys. Chem. A* **2002**, *106*, 657.
[40] A. A. Auer, J. Gauss, M. Pecul, *Chem. Phys. Lett.* **2003**, *368*, 172.
[41] M. Pecul, T. Helgaker, *Int. J. Mol. Sci.* **2003**, *4*, 143.
[42] T. A. Ruden, T. Helgaker, M. Jaszuński, *Chem. Phys.* **2004**, *296*, 53.
[43] R. M. Lynden-Bell, R. K. Harris, *Nuclear Magnetic Resonance Spectroscopy.* Nelson, London, **1969**.
[44] M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11.
[45] M. Karplus, *J. Am. Chem. Soc.* **1963**, *85*, 2870.
[46] P. Pyykkö, in *Calculation of NMR and EPR parameters, Theory and Applications* (Eds.: M. Kaupp, M. Bühl, V. Malkin), Wiley-VCH, Weinheim, **2004**, p. 7.
[47] J. Olsen, P. Jørgensen, *J. Chem. Phys.* **1985**, *82*, 3235.
[48] T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen, K. Ruud, H. Ågren, A. A. Auer, K. L. Bak, V. Bakken, O. Christiansen, S. Coriani, P. Dahle, E. K. Dalskov, T. Enevoldsen, B. Fernandez, C. Hättig, K. Hald, A. Halkier, H. Heiberg, H. Hettema, D. Jonsson, S. Kirpekar, R. Kobayashi, H. Koch, K. V. Mikkelsen, P. Norman, M. J. Packer, T. B. Pedersen, T. A. Ruden, A. Sanchez, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. O. Sylvester-Hvid, P. R. Taylor, O. Vahtras, DALTON, an ab initio electronic structure program, Release 1.2 (2001) (see http://www.kjemi.uio.no/software/dalton/dalton.html).
[49] O. Vahtras, H. Ågren, P. Jørgensen, H. J. Aa. Jensen, S. B. Padkjær, T. Helgaker, *J. Chem. Phys.* **1992**, *96*, 6120.
[50] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed., Wiley-VCH, Weinheim, **2001**.
[51] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, **1989**.
[52] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, Chichester, **2000**.
[53] L. E. McMurchie, E. R. Davidson, *J. Comput. Phys.* **1978**, *26*, 218.
[54] M. A. Watson, P. Sałek, P. Macak, T. Helgaker, *J. Chem. Phys.* in press.
[55] H. J. Aa. Jensen, P. Jørgensen, H. Ågren, *J. Chem. Phys.* **1987**, *87*, 451.
[56] R. Bauernschmitt, R. Ahlrichs, *Chem. Phys. Lett.* **1996**, *256*, 454.
[57] P. Jørgensen, H. J. Aa. Jensen, J. Olsen, *J. Chem. Phys.* **1988**, *89*, 3654.
[58] P. A. M. Dirac, *Proc. Cambridge Philos. Soc.* **1930**, *26*, 376.
[59] S. J. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200.
[60] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257.
[61] S. Huzinaga, *Approximate Atomic Functions. Technical Report*, University of Alberta, Edmonton, **1971**.
[62] W. Kutzelnigg, U. Fleischer, M. Schindler, in *NMR Basic Principles and Progress*, *Vol. 23*, (Eds.: P. Diehl, E. Fluck, H. Günther, R. Kosfeld, J. Seelig), Springer, Berlin, **1990**.
[63] W. J. Hehre, R. F. Stewart, J. A. Pople, *J. Chem. Phys.* **1969**, *51*, 2657.

Received: December 24, 2003
Published online: August 5, 2004

