Methods for computing nuclear magnetic resonance chemical shielding in large systems. Multiple cluster and charge field approaches ☆

Angel C. de Dios and Eric Oldfield

Department of Chemistry, University of Illinois at Urbana-Champaign, 505 South Mathews Avenue, Urbana, IL 61801, USA

Received 23 November 1992; in final form 26 January 1993

Ab initio calculations show that additivity of the intermolecular shielding exists in a model system consisting of fluorobenzene interacting with hydrogen fluoride molecules, $C_6H_5F-(HF)_n$, where $n=1-5$. These results indicate that it should be possible to perform chemical shielding calculations on a large system by dividing it into a series of smaller clusters. For M atoms divided into M/N clusters of N atoms, the time savings for large M is on the order of $M^3/16N^3$, a time savings of ≈ 60 for $M=100$, $N=10$. We demonstrate the feasibility of using point charges to model long-range electrostatic field effects on shielding by comparing the results of full ab initio calculations with those obtained by using point charges to represent the HF molecules in the $C_6H_5F-(HF)_n$ clusters. This comparison shows generally good agreement between the two approaches so long as the point charges are >2.5 Å from all the atoms in the molecule to which the nucleus belongs, a situation which should pertain for many macromolecules. Addition of 1000 point charges to the C_6H_5F system increased computational time by only 50% and appears to offer promise for investigations of chemical shielding in proteins and nucleic acids, where both short-range (electronic) and longer-range (electrostatic field) effects may be important.

1. Introduction

The nuclear magnetic resonance (NMR) chemical shielding nonequivalences in proteins and nucleic acids, caused primarily by folding into their native conformation, have been widely known for a number of years [1–3], and without such nonequivalences modern multi-dimensional NMR studies of protein structure [4,5] would not be possible. It is thus somewhat surprising that relatively little work aimed at interpreting chemical shifts in such systems has been reported. For example, ^{13}C chemical shifts (for a given type of atomic site) often have a 5–10 ppm range [2,6,7], ^{15}N shift ranges are up to about 35 ppm [8,9], ^{17}O shifts are about 15 ppm [10] and ^{19}F shifts are up to ≈ 15–20 ppm [11–13]. With this increasing body of data on chemical shifts in proteins (and, to a lesser extent, in nucleic acids), a sound interpretation of these chemical shift observations becomes even more desirable, since it should be possible to use such information to help refine protein structures in solution, as well as explore topics such as protein electrostatics [14]. Previous efforts in this area have focused primarily on the ^1H nucleus [15,16], where differences in chemical shielding have been attributed to factors such as ring currents, magnetic susceptibility anisotropies, and electric field effects. Since there are many contributions that need to be accounted for, but only a single observable, multi-parameter optimization methods have generally been used to fit large data sets, and the optimized parameters then employed to further refine solution NMR structures [17]. In other cases, even more empirical methods [18] have found utility in structural analysis [19].

While these semi-empirical approaches for ^1H NMR are useful for analyzing e.g. helical or sheet conformations, they have not yet been successful in interpreting the NMR spectra of the heavier elements – ^{13}C, ^{15}N, ^{17}O and ^{19}F, which all have much larger shielding ranges. For these nuclei, the ring cur-

☆ This work was supported by the United States National Institutes of Health (grant HL-19481).

rent and other susceptibility effects have the same magnitude as in ^{1}H NMR – about 1–2 ppm – and thus play only a minor role in shielding. What is left, then, are the major "electrical" and "electronic" interactions, which we can conveniently categorize as short-range (σ_s) or long-range (σ_ℓ), where the total shielding (σ_t) is given by

$$\sigma_t = \sigma_s + \sigma_\ell + \sigma_o , \quad (1)$$

in which σ_o represents the other, smaller or classical interactions mentioned above. We use the breakdown shown in eq. (1) since we believe it helps clarify what needs to be calculated. For example, for ^{15}N in a helical or sheet segment, it can be anticipated that there will be a highly localized helix-sheet-side-chain contribution to shielding, σ_s, due to the different torsion angles ψ, ϕ, χ, etc., as well as a much longer-range contribution, σ_ℓ, due to the charge field of the protein. For say a ^{19}F atom in a 5-fluorotryptophan, we expect that the much longer-range weak electrical interactions, σ_ℓ, will dominate, as shown elsewhere [14,20], since σ_s will not change from residue to residue. For C^{17}O bound to Fe in a number of similar heme proteins, again, σ_ℓ can be expected to dominate [21]. For some of the heavier nuclei, it is thus possible to concentrate on the long-range electric field effects which appear to be dominant in these systems and, hopefully, to develop more rigorous approaches to studying the dependence of the shielding properties on electrostatic fields, while in other cases, e.g. C$^\alpha$, ^{15}NH, it seems likely that methods must be developed to calculate *both* σ_s and σ_ℓ.

There has been much progress during the past few years with regard to ab initio methods of calculating chemical shielding. In fact, computed values for shielding of the first and second row elements have approached experiment so closely that rovibrational corrections are now necessary [22]. Much of the progress in this area can be attributed to the introduction of gauge-including (GIAO, ref. [23]), individual gauge (IGLO, ref. [24]) and local origin (LORG, ref. [25]) methods for calculation of shielding [23–25]. These methods have greatly reduced basis set requirements, and have made study of larger systems possible with use of considerably less computer resources.

However, even with these recent developments, the calculation of chemical shifts in proteins – which typically contain at least a thousand atoms, has remained insurmountable. Thus, it is necessary to make a number of simplifications in the calculations when dealing with these kinds of systems. In the molecular clusters we consider in this Letter, which we maintain are useful models for macromolecular shielding, we divide the chemical shift at a particular site into the same contributions outlined in eq. (1): σ_s, the short-range contribution, which arises from the electrons belonging to the atom which contains the nucleus of interest, and its interaction with the other atoms in the molecule, and σ_ℓ, the long-range or intermolecular contribution, which comes from the electrical perturbations caused by the presence of other molecules in the vicinity of the nucleus of interest. In a protein, of course, most interactions are "intramolecular", but it is logical to break down this total interaction into intra-residue ($\equiv$intramolecular) and inter-residue ($\equiv$intermolecular) interactions, as we will show elsewhere [26].

So, in a protein, to a first approximation, the atoms belonging to the residue that contains the nucleus of interest can be regarded as contributing to the shielding intramolecularly, while all the other amino-acid residues contribute "intermolecularly". In this fashion, the intramolecular shielding can be computed using conventional ab initio methods, since only a relatively small number of atoms need to be considered. The next question is – how to obtain the intermolecular shift? Recently, ab initio calculations of the intermolecular contributions to shielding of the rare gases, either in the gas phase, or as adsorbed species, have been shown to be successful [27]. This work indicated that the intermolecular contributions to shielding were additive, and it satisfactorily reproduced the experimental shieldings. In particular, it was found that the change in the shielding of an Ar atom caused by the presence of two neighboring Ar atoms was equivalent to twice the change in shielding brought about by a single neighboring Ar atom. While these rare-gas results may naturally be unique, they nevertheless stimulated us to investigate the possibility that a similar behavior might be exhibited in more complex molecular clusters, and by inference, in macromolecules. We are thus interested in developing tractable ways of handling these weak or long-range electro-

static contributions to shielding, and thus first give a brief review of the topic, as follows.

2. Theoretical background

The effect of a uniform electric field on a molecular electronic property, P, such as chemical shielding, was first suggested by Stephen [28] and Buckingham [29,30] as being expressible as a power series in the uniform field F,

$$P_{\alpha\beta}=P^{(0)}_{\alpha\beta}+P^{(1)}_{\alpha\beta,\gamma}F_\gamma+P^{(2)}_{\alpha\beta,\gamma,\delta}F_\gamma F_\delta+\dots . \quad (2)$$

The intermolecular shielding observed at a particular site can thus be regarded as a result of the electric fields arising from its neighbors. In proteins, these fields would be far from uniform, hence eq. (2) should include additional terms that take into account this non-uniformity, as suggested by Buckingham and Lawley (for the H atom, ref. [31]) and Batchelor [32],

$$P_{\alpha\beta}=P^{(0)}_{\alpha\beta}+P^{(1)}_{\alpha\beta,\gamma}F_\gamma+P^{(1)}_{\alpha\beta,\gamma\delta}F_{\gamma\delta}+P^{(2)}_{\alpha\beta,\gamma,\delta}F_\gamma F_\delta+P^{(2)}_{\alpha\beta,\gamma\delta,\epsilon\zeta}F_{\gamma\delta}F_{\epsilon\zeta}+\dots , \quad (3)$$

where $F_{\gamma\delta}$ is the $\gamma\delta$ element of the field gradient tensor. If this expansion converges, then with a knowledge of the coefficients listed in eq. (3) and the field and the field gradient tensor at the nucleus in question, the shielding can be calculated. The coefficients in eq. (3) can be obtained by calculating corrections to the energy of the system arising from several perturbations: the external magnetic field, the external electric field, and the nuclear magnetic moment; and over the past 30 years, there has been considerable work aimed at determining these coefficients [20,29–33]. Most recently, Dykstra and co-workers have developed an open-ended way of differentiating electronic wavefunctions and energies, derivative Hartree–Fock theory [34,35], and have reported a series of so-called (multipole) shielding polarizability tensors,

$$A_{\alpha\beta,\gamma}=\frac{\partial\sigma_{\alpha\beta}}{\partial F_\gamma}, \quad \text{etc.}, \quad (4)$$

for the uniform fields, field gradients, and so forth.

This is clearly a very elegant way of obtaining the weak or long-range electrical contributions to shielding, σ_ℓ, and based on other work [14,20,21] there are good reasons to believe that eq. (3) can give a good description of some protein chemical shifts – e.g. for ^{19}F or C^{17}O, while in other cases, such as C$^\alpha$, ^{15}NH, it seems likely that the A and B tensors themselves will be functions of ψ, ϕ and χ, the backbone and sidechain torsion angles. We are thus exploring other routes, in which $\sigma_s+\sigma_\ell$ are computed together by using ab initio methods. A full ab initio calculation does not, of course, rely on the convergence implicit in eq. (3), but as mentioned before, the large number of atoms in a protein impedes such calculations. However, if the intermolecular contributions to shielding are additive, then there should be a significant decrease in computational time by dividing the molecule up into much smaller clusters, taking advantage of the $\approx N^4$ basis set size dependence. Alternatively, there may be even simpler ways of incorporating electrical polarization effects, e.g. by introducing the charge field at the SCF level.

In this theoretical study we focus our attention on the molecule fluorobenzene, C_6H_5F. Fluorobenzene is small enough that a basis set of triple-zeta quality can be employed without taking too much computer time. Hydrogen fluoride is chosen to act as the perturbing molecule. The additivity of the chemical shielding is examined by introducing successive HF (or FH) molecules into the system. Simple additivity of shielding is unlikely to hold over the whole range of possible separations between C_6H_5F and HF, so we investigate where additivity is actually demonstrable.

The second method we describe models σ_ℓ by including the purely classical electrostatic influence of the intermolecular partners in the ab initio calculation of the chemical shielding in the representative molecular fragment. Such an approach has been previously applied to understanding the influence on quadrupole coupling constants (qcc) by partner molecules in van der Waals molecules. In a study of the ^{14}N qccs of HCN in dimer complexes (HCN–X, X=HF, HCN, HCCH), Jaman et al. [36] represented the electrical potential of the partner molecule by that of their low-order electrical multipoles, and this electrical potential was incorporated into the Hamiltonian of the ab initio calculation of the qcc. This model approach, termed charge field perturbation (CFP), reproduced the experimentally ob-

served trends. This approach has also been used for the H_2O–HCN system [37], and Cummins et al. [38] have applied a similar procedure to examine the ^{17}O qcc of water in ice, representing the other water molecules by point charges.

This method can therefore be regarded as an extreme case of the locally-dense basis set scheme proposed by Chesnut and Moore [39]. If the intermolecular shift is dominated by weak electrical interactions, then it seems reasonable to believe that it should be possible to describe the HF molecule by point charges (or else the interaction is not weak). The partial charges used for HF can be obtained either by a Mulliken population analysis after an SCF calculation, or can be extracted from the experimentally measured dipole moment and bond length. The atoms of perturbing molecules are thus replaced with partial charges, which do not add to the number of electrons and basis functions of the system, and as described below we compare results of full ab initio calculations on C_6H_5F–$(HF)_n$ with those obtained by using the point-charge model. Related work aimed at modeling a crystallographic charge field for saccharide chemical shielding calculations is also being pursued by Grant [40].

3. Computational

SCF and shielding calculations were carried out using the TX90 program of Pulay and co-workers [41,42], which features an efficient implementation of the GIAO (gauge-including atomic orbital) chemical shielding method proposed by Ditchfield [23]. The basis set used in all calculations presented is the 6-311G basis set of Pople and co-workers [43], augmented with a set of d-type polarization functions on the heavy atoms, and a set of p-type functions on the hydrogens (6-311G**). All computations were performed on an IBM RISC/6000, model 350 computer (IBM Corporation, Austin, TX, USA). The partial charges used in representing the HF molecule were obtained by a Mulliken population analysis with the basis set 6-311G++(3d, 3p) obtained by using the GAUSSIAN 88 program [44]. The geometries of C_6H_5F and HF used in the calculations were both obtained from experiment [45,46].

4. Results and discussion

The isotropic shielding obtained for ^{19}F in C_6H_5F at its experimental equilibrium geometry is 343.80 ppm, with the following traceless principal components: $\sigma_{11} = -66.0$ ppm, $\sigma_{22} = 18.6$ ppm, and $\sigma_{33} = 47.3$ ppm. These calculated traceless components agree favorably with the experimental values $\sigma_{11} = -58.0$ ppm, $\sigma_{22} = 7.0$ ppm, and $\sigma_{33} = 51.0$ ppm obtained using solid state NMR [20,47]. σ_{11} is along y, σ_{22} along x and σ_{33} is oriented along z, using the axis system given in fig. 1. Absolute experimental shieldings in the limit of zero density for C_6H_5F have not been reported, and it would be inappropriate to compare the calculated values with condensed phase data.

For the intermolecular shift studies, there are three sites on which an HF molecule can be placed, as shown in fig. 1, which we designate HF_x, HF_y and HF_z. For our investigation of the additivity of chemical shielding, we have taken a fluorobenzene molecule perturbed by either an axial (HF_x), equatorial (HF_y) or apical (HF_z) HF molecule, as shown in fig. 1, and computed the change in shielding on interaction with HF. Table 1 gives results for the three dimers, obtained using a 3.0 Å separation between the fluorobenzene F and HF hydrogen. We then carried out full ab initio calculations of the shieldings for the five C_6H_5F–$(HF)_2$ trimers, the five tetramers, the three pentamers, and the hexamer, and then deduced the corresponding shielding changes due to simple additivity of the dimer results, as shown in table 1.

Table 2 presents the calculated shieldings when the hydrogen of HF is placed 5.0 Å from the aromatic fluorine. The calculated shieldings reported in tables 1 and 2 both have counterpoise corrections, which were obtained by placing ghost orbitals at the hydrogen fluoride positions. The results of tables 1 and

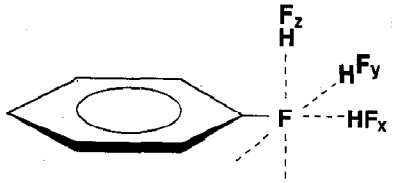

Fig. 1. Fluorobenzene–hydrogen fluoride cluster axis system.

Table 1
^{19}F NMR chemical shielding of fluorobenzene in 3 Å $C_6H_5F-(HF)_n$ clusters

System	$\sigma-\sigma(C_6H_5F)$ [a] (ppm)	Cluster sum (ppm)	Difference (%)
$C_6H_5F-(HF_x)$	5.984 (4.026) {*A*}		
$C_6H_5F-(HF_y)$	2.807 (2.190) {*B*}		
$C_6H_5F-(HF_z)$	4.941 (1.429) {*C*}		
$C_6H_5F-(HF_x)(HF_y)$	8.527 (5.891)	8.791 {*A*+*B*}	3.1
$C_6H_5F-(HF_x)(HF_z)$	10.625 (5.218)	10.925 {*A*+*C*}	2.8
$C_6H_5F-(HF_y)(HF_z)$	7.553 (3.580)	7.748 {*B*+*C*}	2.6
$C_6H_5F-(HF_y)_2$	5.526 (4.209)	5.614 {2*B*}	1.6
$C_6H_5F-(HF_z)_2$	9.750 (2.829)	9.882 {2*C*}	1.4
$C_6H_5F-(HF_x)(HF_y)(HF_z)$	12.986 (7.057)	13.732 {*A*+*B*+*C*}	5.7
$C_6H_5F-(HF_x)(HF_y)_2$	10.981 (7.611)	11.598 {*A*+2*B*}	5.6
$C_6H_5F-(HF_x)(HF_z)_2$	15.154 (6.386)	15.866 {*A*+2*C*}	4.7
$C_6H_5F-(HF_y)(HF_z)_2$	12.179 (4.925)	12.689 {*B*+2*C*}	4.2
$C_6H_5F-(HF_z)(HF_y)_2$	10.095 (5.555)	10.555 {2*B*+*C*}	4.6
$C_6H_5F-(HF_x)(HF_y)_2(HF_z)$	15.276 (8.748)	16.539 {*A*+2*B*+*C*}	8.3
$C_6H_5F-(HF_x)(HF_y)(HF_z)_2$	17.347 (8.185)	18.673 {*A*+*B*+2*C*}	7.6
$C_6H_5F-(HF_y)_2(HF_z)_2$	14.563 (6.838)	15.496 {2*B*+2*C*}	6.4
$C_6H_5F-(HF_x)(HF_y)_2(HF_z)_2$	19.488 (9.829)	21.480 {*A*+2*B*+2*C*}	10.2

[a] Counterpoise corrections used are enclosed in parentheses. The shielding values given are those computed, but are probably only accurate to a few percent in absolute terms.

Table 2
^{19}F NMR chemical shielding on fluorobenzene in 5 Å $C_6H_5F-(HF)_n$ clusters

System	$\sigma-\sigma(C_6H_5F)$ [a] (ppm)	Cluster sum (ppm)	Difference (%)
$C_6H_5F-(HF_x)$	1.704 {*A*}		
$C_6H_5F-(HF_y)$	0.036 {*B*}		
$C_6H_5F-(HF_z)$	0.984 [*C*}		
$C_6H_5F-(HF_x)(HF_y)$	1.730	1.740 {*A*+*B*}	0.6
$C_6H_5F-(HF_x)(HF_y)(HF_z)$	2.690	2.724 {*A*+*B*+*C*}	1.3
$C_6H_5F-(HF_x)(HF_y)_2(HF_z)$	2.710	2.760 {*A*+2*B*+*C*}	1.8
$C_6H_5F-(HF_x)(HF_y)_2(HF_z)_2$	3.658	3.744 {*A*+2*B*+2*C*}	2.4

[a] Counterpoise corrections are negligible at this distance.

2 show that the intermolecular shift, to a good first approximation, can be regarded as additive. Calculations were also performed in which the hydrogen of HF is 2.5 Å from the aromatic fluorine. The differences are larger, up to 20%, indicating the loss of additivity at very short distances. Table 3 shows the results of calculations, also at 3.0 Å, but with the HF molecule having a reversed orientation (i.e. $C_6H_5F...FH$). Here, as in tables 1 and 2, a relatively good additivity relationship is exhibited, and at 5 Å separation, table 4, there is even better agreement. All of the results of tables 1–4 are presented in graphical form in fig. 2, where the excellent overall agreement using the cluster approach is apparent.

A calculation involving a C_6H_5F molecule interacting with an HF molecule takes about two hours of CPU time (at 20 Mflops). On the other hand, a C_6H_5F molecule with five HF molecules takes about twelve hours. Using additivity, one reduces this time to six hours. The savings in time will be even more pronounced with larger numbers of atoms. For example, if we consider a large species having M atoms

Table 3
^{19}F NMR chemical shielding of fluorobenzene in 3 Å $C_6H_5F-(FH)_n$ clusters

System	$\sigma-\sigma(C_6H_5F)$ [a] (ppm)	Cluster sum (ppm)	Difference (%)
$C_6H_5F-(FH_x)$	−4.853 (0.910) {A}		
$C_6H_5F-(FH_y)$	2.832 (−0.577) {B}		
$C_6H_5F-(FH_z)$	−4.509 (0.434) {C}		
$C_6H_5F-(FH_x)(FH_y)$	−1.892 (0.340)	−2.021 {A+B}	6.8
$C_6H_5F-(FH_x)(FH_z)$	−8.067 (1.346)	−9.362 {A+C}	16.1
$C_6H_5F-(FH_y)(FH_z)$	−1.730 (−0.133)	−1.677 {B+C}	3.1
$C_6H_5F-(FH_y)_2$	5.733 (−1.114)	5.664 {2B}	1.2
$C_6H_5F-(FH_z)_2$	−9.073 (0.846)	−9.018 {2C}	0.6
$C_6H_5F-(FH_x)(FH_y)(FH_z)$	−6.372 (0.787)	−6.530 {A+B+C}	2.5
$C_6H_5F-(FH_x)(FH_y)_2$	1.175 (−0.189)	0.811 {A+2B}	31.0
$C_6H_5F-(FH_x)(FH_z)_2$	−14.026 (1.760)	−13.871 {A+2C}	1.1
$C_6H_5F-(FH_y)(FH_z)_2$	−6.082 (0.289)	−6.186 {B+2C}	1.7
$C_6H_5F-(FH_z)(FH_y)_2$	1.419 (−0.661)	1.155 {2B+C}	18.6
$C_6H_5F-(FH_x)(FH_y)_2(FH_z)$	−3.228 (0.267)	−3.698 {A+2B+C}	14.5
$C_6H_5F-(FH_x)(FH_y)(FH_z)_2$	−10.906 (1.211)	−11.039 {A+B+2C}	1.2
$C_6H_5F-(FH_y)_2(FH_z)_2$	−4.130 (0.910)	−3.354 {2B+2C}	18.8
$C_6H_5F-(FH_x)(FH_y)_2(FH_z)_2$	−7.689 (0.699)	−8.207 {A+2B+2C}	6.7

[a] Counterpoise corrections used are enclosed in parentheses.

Table 4
^{19}F NMR chemical shielding of fluorobenzene in 5 Å $C_6H_5F-(FH)_n$ clusters

System	$\sigma-\sigma(C_6H_5F)$ [a] (ppm)	Cluster sum (ppm)	Difference (%)
$C_6H_5F-(FH_x)$	−1.473 {A}		
$C_6H_5F-(FH_y)$	−0.166 {B}		
$C_6H_5F-(FH_z)$	−0.806 {C}		
$C_6H_5F-(FH_x)(FH_y)$	−1.638	−1.639 {A+B}	0.06
$C_6H_5F-(FH_x)(FH_y)(FH_z)$	−2.446	−2.445 {A+B+C}	0.04
$C_6H_5F-(FH_x)(FH_y)_2(FH_z)$	−2.605	−2.611 {A+2B+C}	0.23
$C_6H_5F-(FH_x)(FH_y)_2(FH_z)_2$	−3.415	−3.417 {A+2B+2C}	0.06

[a] Counterpoise corrections are negligible at this distance.

(or basis functions), division into M/N clusters of N atoms and assuming that the overall length of the calculation goes *approximately* as the fourth power of the number of basis functions, then an overall time savings of $\approx M^3/16N^3$ is obtained using the cluster approach. For say 100 atoms, a 20000 hour calculation is thus reduced to ≈125 hours (at 20 Mflops), or overnight using a relatively small cluster of RISC machines. While the actual time savings will vary from program to program, considerable time savings can be anticipated in most cases, as we find experimentally, due to additivity.

We now consider briefly the question: where does this additivity come from? First, the additivity of the intermolecular shielding suggests that eq. (3) is indeed a good representation of the underlying physics. The field and field gradient terms can be regarded as sums of contributions coming from each molecule in the cluster. For example, in the case of the $C_6H_5F-(HF)_3$ tetramer (an axial HF_x, an equatorial HF_y and an apical HF_z), the following relationships hold:

$$\text{total } F_y = F_y(\text{axial HF}) + F_y(\text{equatorial HF}) + F_y(\text{apical HF}) , \quad (5)$$

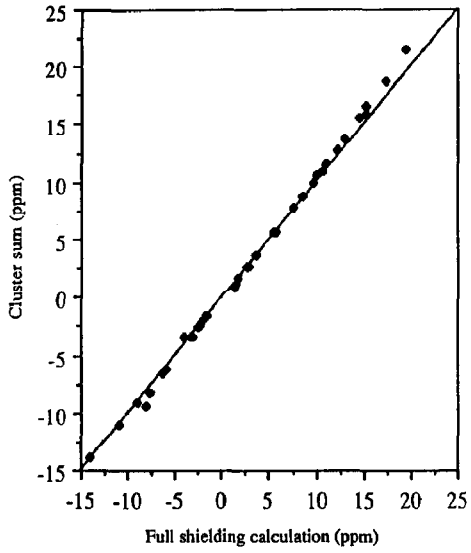

Fig. 2. Graph showing relationship between ^{19}F NMR chemical shielding of F in fluorobenzene in a series of $C_6H_5F-(HF)_n$ clusters calculated using a full ab initio method versus shielding computed by using a multiple cluster additivity approach. Positive values indicate increased shielding over that observed in free C_6H_5F.

and

$$\text{total } F_{\gamma\delta} = F_{\gamma\delta}(\text{axial HF}) + F_{\gamma\delta}(\text{equatorial HF}) + F_{\gamma\delta}(\text{apical HF}) . \quad (6)$$

The additivity of the intermolecular shift is a consequence of the additivity of these purely electrostatic terms. At closer distances, additivity breaks down, which is to be expected because at close range one needs to take into account dispersion, and mutual polarization amongst the molecules. It is thus fortunate that in the systems of greatest interest (to us) the relevant range of distances is typically $\gg 3.0$ Å, below that where additivity breaks down.

The reduction in computational time (and disk space) achieved for proteins and nucleic acids, using additivity, although significant, may still not be adequate to enable calculations for very large systems, on a small machine. The applicability of eq. (3), however, points out that an alternative and even faster approach may be possible, if indeed the change in chemical shielding is taken to be a result of the electric field and field gradients the nucleus in question is experiencing. In particular, it may be possible to describe this highly non-uniform electrostatic field by representing the perturbing molecules with partial atomic charges. The partial charges we have used for representing the HF molecule are 0.5644 and -0.5644 au for H and F, respectively. The shielding results using partial charges are shown in table 5. The agreement between the full ab initio and point-charge calculations is very satisfactory, except for the one case where the HF molecule is lying in the plane containing the aromatic ring, but perpendicular to the C–F bond. The discrepancy here may be due to the fact that although this hydrogen position is 3.0 Å away from the aromatic fluorine, it is only 1.7 Å from one of the ortho-hydrogens of the phenyl ring. At this very close approach, the partial charges may no longer be good representations of the entire perturbation.

Fig. 3 shows in graphical form the results of another set of calculations which exhibit excellent agreement between the full ab initio and point-charge results. Here, only an axial HF molecule is considered (as point charges), and the different shieldings are obtained by using different separations between the C_6H_5F and the point charges. The range of chem-

Table 5
^{19}F NMR chemical shielding of fluorobenzene in 3 Å $C_6H_5F-(HF)_n$ clusters using full ab initio and point-charge models

System	Full ab initio (ppm)	Point charges (ppm)	Difference (%)
$C_6H_5F-(HF_x)$	5.984	6.272	4.8
$C_6H_5F-(HF_y)$	2.807	1.646	41.3
$C_6H_5F-(HF_z)$	4.941	5.470	10.7
$C_6H_5F-(HF_x)(HF_y)$	8.527	7.798	8.5
$C_6H_5F-(HF_x)(HF_y)(HF_z)$	12.986	12.921	0.5
$C_6H_5F-(HF_x)(HF_y)_2(HF_z)$	15.276	14.371	5.9
$C_6H_5F-(HF_x)(HF_y)_2(HF_z)_2$	19.488	19.265	1.1

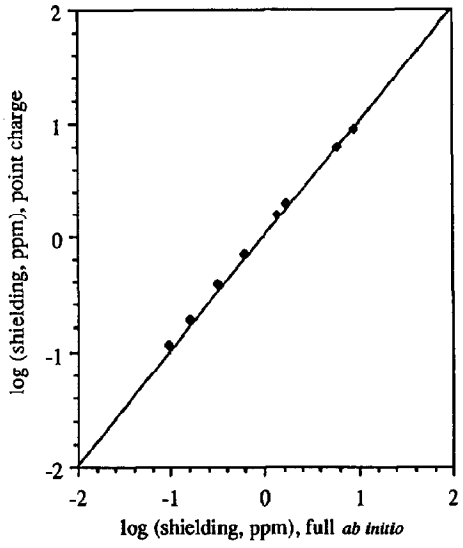

Fig. 3. Graph showing relationship between ^{19}F NMR chemical shielding of F in fluorobenzene in a series of C_6H_5F–HF axial dimers calculated using a full ab initio method versus shielding computed by using partial atomic charges to represent the HF molecules, as a function of separation distance (r=2.5→ 20 Å).

ical shifts in fig. 3 covers two orders of magnitude, and a range of separations which probably encompasses the relevant range of separations for intermolecular interactions in proteins and nucleic acids, 3–20 Å. When taken together with the results of table 5, these results strongly indicate that the point-charge model appears to be quite accurate. In addition, the increase in computational time due to the introduction of point charges is very small. In fact, a calculation involving C_6H_5F and one thousand point charges takes only about an hour more than a similar computation on C_6H_5F alone.

5. Conclusions

The results we have shown above indicate that the intermolecular contributions to NMR chemical shielding, as modeled by C_6H_5F interacting with HF molecules, are mostly additive. This additivity relation supports the notion that an electrostatic picture is a viable representation of intermolecular chemical shielding, at least in this particular system. The apparent adequacy of the electrostatic treatment suggests that the atoms of perturbing residues in macromolecules can be represented by simple point charges, at least in regions of relatively uniform dielectric, such as many buried peptide backbone sites. In this manner, chemical shifts involving a large number of atoms can be calculated. For residues which have very little change in short-range (or electronic) shielding from site to site, e.g. 5-^{19}F Trp residues, the shielding polarizability approach discussed elsewhere [20,21,33] appears to be the method of choice. However, when both long-range (σ_ℓ) *and* short-range (σ_s) contributions to shielding need to be included, then either the cluster method or the point charge (or a more sophisticated but nevertheless SCF level) model may offer certain computational advantages.

Based on this work, and that reported elsewhere, we thus conclude: (1) intermolecular (inter-residue) contributions to shielding may often be additive; (2) additivity supports the idea that shielding can be expressed as a convergent power series expansion in the potential; (3) point-charge models work because of (2); (4) shielding polarizability models work for the same reasons, and as noted elsewhere, mutual polarization corrections [48] appear relatively small. Thus, a number of new methods are now available for calculating intermolecular (inter-residue) contributions to shielding, and these can be used to investigate both static structure, dynamic structure, and electrostatics, in complex systems. "Brute force" hardware improvements will still be essential, nevertheless, for acceptable progress with macromolecules.

Acknowledgement

We thank Professor P. Pulay, Professor J.F. Hinton and Dr. K. Wolinski for providing us with a copy of their TX90 program, Professor Cynthia J. Jameson and Professor Clifford E. Dykstra for helpful discussions, and a referee for constructive comments.

References

[1] C.C. McDonald and W.D. Phillips, J. Am. Chem. Soc. 91 (1969) 1513.

[2] A. Allerhand, R.F. Childers and E. Oldfield, Biochemistry 12 (1973) 1335.
[3] W.C. Chu and J. Horowitz, Nucleic Acids Res. 17 (1989) 7241;
W.J. Metzler, C. Wang, D.B. Kitchen, R.M. Levy and A. Pardi, J. Mol. Biol. 214 (1990) 711.
[4] G.M. Clore and A.M. Groenenborn, Progr. NMR Spectry. 23 (1991) 43.
[5] A. Bax, M. Ikura, L. Kay, D.A. Torchia and R. Tschudin, J. Magn. Reson. 86 (1990) 304.
[6] L.K. Nicholson, L.E. Kay, D.M. Baldisseri, J. Arango, P.E. Young, A. Bax and D.A. Torchia, Biochemistry 31 (1992) 5253.
[7] S. Spera and A. Bax, J. Am. Chem. Soc. 113 (1991) 5490.
[8] J. Glushka, M. Lee, S. Coffin and D. Cowburn, J. Am. Chem. Soc. 111 (1989) 7716.
[9] D.A. Torchia, S.W. Sparks and A. Bax, Biochemistry 28 (1989) 5509;
H.B.R. Cole, S.W. Sparks and D.A. Torchia, Proc. Natl. Acad. Sci. US 85 (1988) 6362.
[10] K.D. Park, K. Guo, F. Adebodun, M.L. Chiu, S.G. Sligar and E. Oldfield, Biochemistry 30 (1991) 2333.
[11] M.P. Gamcsik, J.T. Gerig and R.B. Swenson, Biochim. Biophys. Acta 874 (1986) 372.
[12] L.A. Luck and J.J. Falke, Biochemistry 30 (1991) 4248.
[13] E. Oldfield, B. Montez, J. Patterson, S. Harrell, H. Le and S. Lian, unpublished results.
[14] E. Oldfield, J. Pearson, F.S. Lee and A. Warshel, unpublished results.
[15] K. Ösapay and D.A. Case, J. Am. Chem. Soc. 113 (1991) 9436.
[16] M.P. Williamson, T. Asakura, E. Nakamura and M. Demura, J. Biomol. NMR 2 (1992) 83.
[17] D.A. Case, NATO Advanced Study Institute, University of Maryland, College Park, MD, poster presentation, July (1992).
[18] D.S. Wishart, B.D. Sykes and F.M. Richards, Biochemistry 31 (1992) 1647.
[19] H.R. Mott, P.C. Driscoll, J. Boyd, R.M. Cooke, M.P. Weir and I.D. Campbell, Biochemistry 31 (1992) 7741.
[20] C.E. Dykstra, J.D. Augspurger, J. Pearson and E. Oldfield, NATO Advanced Study Institute, College Park, MD, in press.
[21] J.D. Augspurger, C.E. Dykstra and E. Oldfield, J. Am. Chem. Soc. 113 (1991) 2447.
[22] P.W. Fowler, G. Riley and W.T. Raynes, Mol. Phys. 42 (1981) 1463;
P. Lazzeretti, R. Zanasi, A.J. Sadlej and W.T. Raynes, Mol. Phys. 62 (1987) 605;
C.J. Jameson, C. de Dios and A.K. Jameson, J. Chem. Phys. 95 (1991) 1069, 9042.
[23] R. Ditchfield, J. Chem. Phys. 56 (1972) 5688.
[24] M. Schindler and W. Kutzelnigg, J. Am. Chem. Soc. 105 (1983) 1360.
[25] A.E. Hansen and T.D. Bouman, J. Chem. Phys. 82 (1985) 5035.
[26] A.C. de Dios and E. Oldfield, unpublished results.
[27] C.J. Jameson and A.C. de Dios, J. Chem. Phys. 97 (1992) 417.
[28] M.J. Stephen, Mol. Phys. 1 (1957) 223.
[29] A.D. Buckingham, Can. J. Chem. 38 (1960) 300.
[30] B. Day and A.D. Buckingham, Mol. Phys. 32 (1976) 343.
[31] A.D. Buckingham and K.P. Lawley, Mol. Phys. 3 (1960) 219.
[32] J.G. Batchelor, J. Am. Chem. Soc. 97 (1975) 3410.
[33] J.Gruninger and H.F. Hameka, J. Chem. Phys. 48 (1968) 4878;
J.P. Riley and W.T. Raynes, Mol. Phys. 32 (1976) 569;
A.J. Sadlej and W.T. Raynes, Mol. Phys. 35 (1978) 101;
M. Žaucer and A. Ažman, Z. Naturforsch. 34a (1979) 1279;
M.J. Packer and W.T. Raynes, Mol. Phys. 69 (1990) 391;
J. Augspurger, J. Pearson, E. Oldfield, C.E. Dykstra, K.D. Park and D. Schwartz, J. Magn. Reson. 100 (1992) 342.
[34] C.E. Dykstra and P.G. Jasien, Chem. Phys. Letters 109 (1984) 388.
[35] J.D. Augspurger and C.E. Dykstra, J. Phys. Chem. 95 (1991) 9230.
[36] A.I. Jaman, T.C. Germann, H.S. Gutowsky, J.D. Augspurger and C.E. Dykstra, Chem. Phys. 154 (1991) 281.
[37] H.S. Gutowsky, T.C. Germann, J.D. Augspurger and C.E. Dykstra, J. Chem. Phys. 96 (1992) 5808.
[38] P.L. Cummins, G.B. Bacskay and N.S. Hush, Mol. Phys. 61 (1987) 795.
[39] D.B. Chesnut and K.D. Moore, J. Comput. Chem. 10 (1989) 648.
[40] D. Grant, private communication.
[41] P. Pulay, K. Wolinksi and J.F. Hinton, The Texas Program, University of Arkansas, Fayetteville, AR (1991).
[42] K. Wolinski, J.F. Hinton and P. Pulay, J. Am. Chem. Soc. 112 (1990) 8251.
[43] R. Krishnan, J.S. Binkley, R. Seeger and J.A. Pople, J. Chem. Phys. 72 (1980) 650.
[44] M.J. Frisch, M. Head-Gordon, H.B. Schlegel, K. Raghavachari, J.S. Binkley, C. Gonzalez, D.J. DeFrees, D.J. Fox, R.A. Whiteside, R. Seeger, C.F. Melius, J. Baker, R.L. Martin, L.R. Kahn, J.J.P. Stewart, E.M. Fluder, S. Topiol and J.A. Pople, GAUSSIAN 88 (Gaussian, Inc., Pittsburgh, 1988).
[45] L. Hygaard, I. Bojensen, T. Pedersen and J. Rastrup-Andersen, J. Mol. Struct. 2 (1968) 209.
[46] G. Herzberg, Spectra of diatomic molecules (Van Nostrand, Princeton, 1950).
[47] M. Mehring, NMR Basic principles and progress (Springer, Berlin, 1976).
[48] J. Augspurger and C. E. Dykstra, private communications.