Abstract: A new method for refinement of 3D molecular structures by geometry optimization is presented. Prerequisites are a force field and a very fast procedure for the calculation of chemical shifts in every step of optimization. To the energy, provided by the force field (COSMOS force field), a pseudoenergy, depending on the difference between experimental and calculated chemical shifts, is added. In addition to the energy gradients, pseudoforces are computed. This requires the derivatives of the chemical shifts with respect to the coordinates. The pseudoforces are analytically derived from the integral expressions of the bond polarization theory. Single chemical shift values attributed to corresponding atoms are considered for structural correction. As a first example, this method is applied for proton position refinement of the D-mannitol X-ray structure. A crystal structure refinement with 13C chemical shift pseudoforces is carried out.


Key words: chemical shift; chemical shift restriction; structure refinement; pseudoforces; bond polarization theory; crystal simulation

Introduction

The 13C NMR chemical shift in liquids, liquid crystals, and solids has become a powerful and versatile tool for the study of structure, morphology, and dynamics of these molecular systems. During the last few years, considerable progress has been achieved by the structure investigation, especially of proteins and peptides. It was recognized that there is a strong correlation between protein backbone conformation and Cα and Cβ secondary chemical shift (the difference between observed shift and random coil shift).

It is also well known that the helical Cβ sites are deshielded by about 5 ppm compared to those of sheet geometry. Additionally, DFT-IGLO calculations show that the Cα—Hα chemical shift bond contribution is the most sensitive and could have conformational differences as large as 22 ppm. There is also a change in anisotropy, about 12 ppm, and in tensor orientation. For carbonyl sites it was found that they are strongly affected by hydrogen bonding and charge field perturbations. However, the chemical shift also provides information about backbone dynamics, ring flip rates, proximity and orientation of aromatic rings, ionization states, oxidation states, and internuclear O—H distances.

Simple and effective techniques for identifying protein secondary structure are either based on the chemical shift index, database searches for chemical shift and sequence homology, or application of statistical methods. Other indirect 13C chemical shift structure refinement methods utilize chemical shielding surfaces, which are functions of the peptide backbone dihedral angles (φ, ψ) and side chain torsion angles (χ). These shielding surfaces are either determined experimentally from protein crystal structures or calculated from single amino acids. All of these methods are restricted to special classes of compounds. By applying Bayesian probabilities on chemical shift surfaces that depend on dihedral angles, probability surfaces (Z surfaces) are derived. These are used to obtain the most probable torsion angles given due to the NMR data.

Direct force field refinement methods introduce additional chemical shift pseudoenergy terms. They are functions of the differences between experimental and estimated data. These expectation values are obtained, for instance, from the empirical surface of Spera and Bax and depend on the dihedral backbone and side chain torsion angles. For targeting the expectation values on the observed Cα and Cβ chemical shifts, different potentials, such as the square well function or the harmonic potential, are used. The latter, for instance, is given by

\[ E_{CS} = k_{CS} ([\delta(\phi, \psi)] - \delta^{obs})^2 \]  

where the force constant \( k_{CS} \) is about 1.0 kcal mol\(^{-1}\) ppm\(^{-2}\).

Recently, a refinement technique of orientational restraints for uniform aligned membrane proteins with respect to the external magnetic field was developed.

All methods mentioned so far are rather restricted to singular classes of compounds and even to special residues such as Cα, Cβ, and Cγ sites of proteins and peptides. It seems difficult to apply them to complexes and ionic systems. So far no method exists for...
explicit $^{13}$C chemical shift structure determination. In this work, we derive a chemical shift pseudopotential from the bond polarization approach and provide a direct method of 3D structure refinement with $^{13}$C chemical shift data.

Ab initio MD simulations$^{12}$ and CS calculations$^{13}$ are highly recommended but are still very time consuming, even when using parallel architectures and QM/MM methods.$^{14,15}$ Therefore, it is very important to introduce a simplification for the Hamiltonian, keeping as much as possible of the electronic information. Within the COSMOS$^{16}$ force field approach the charges of the Coulomb energy are calculated in every step of geometry optimization using the semiempirical bond polarization theory (BPT).$^{17}$ These atomic charges depend on the 3D arrangement of all atoms. For the calculation of the chemical shifts these charges are also necessary. For the other types of interactions classical force field potentials are used.

In this work, a general chemical shift pseudopotential, utilizing a structure dependent chemical shift calculation, is introduced into the force field. This provides a general and fast tool to estimate the chemical shift pseudoforces during every step of geometry optimization as well as molecular dynamics simulation. These pseudoforces correct the classical force field interactions. As soon as the geometry does not give the proper theoretical chemicals shift values, strong error minimizing forces drive the structure into more reliable conformations. Therefore, this method is applicable to a wide range of liquid and solid state compounds containing single and double bonds such as C—C, C=C, C—H, C—N, C=N, C—O, and C=O. Additionally, it can be used for structure investigation of metalloproteins.

Theory

Introduction

When nuclear magnetic dipoles, such as $^{13}$C nuclei, interact with external magnetic fields, the surrounding electrons shield or deshield the nuclei. This property is identified as the magnetic shielding tensor and can be represented as an expectation value $\sigma_N = \langle \Psi_0 | \sigma_N | \Psi_0 \rangle$. The wave function $|\Psi_0\rangle$, belonging to the ground-state energy $E_0$, is the solution of the Hamilton operator or Fock operator without external perturbations. The shielding operator $\sigma_N$ yields within the Coulomb gauge restriction the general form (Ramsey expression$^{18}$)

$$\hat{\sigma}_N = \hat{\sigma}_N^{\text{dia}} - 2 \sum_{k \neq 0} \frac{\hat{\sigma}_N^{\text{orb}} |\Psi_k(0)\rangle \langle \Psi_k(0)| (\hat{\sigma}_N^{\text{pso}})^T}{E_k - E_0}$$

The shielding operator contains the excited states $|\Psi_k(0)\rangle$ belonging to the energy $E_k$. Involved electron interactions$^{19}$ are described by the operators $\hat{\sigma}_N^{\text{orb}}$, $\hat{\sigma}_N^{\text{pso}}$, and $\hat{\sigma}_N^{\text{dia}}$. The coupling of the external field to the orbital motion is expressed by orbital angular-momentum operators:

$$\hat{J}^{\text{orb}}_N = \frac{1}{2} \sum_i \hat{J}_i$$

The nuclear spin-orbit operator or orbital hyperfine operator couples the nuclear magnetic moments to the orbital motion of the electrons:

$$\hat{J}_N^{\text{pso}} = \alpha^2 \sum_i \frac{\hat{J}_N}{r_i^3}$$

The diamagnetic electron operator has the form

$$\hat{J}_N^{\text{dia}} = \frac{\alpha^2}{2} \sum_i \frac{(r_i O r_i N) - (r_i N r_i O)}{r_i^3}$$

The shielding operator can be written as a sum of one electron operators. It is clear that the second order perturbation theory is included into the shielding operator itself [eq. (2)]. In principle the calculation of this NMR parameter is independent of the exact wave function model.

For most calculations the Ramsey expression is not useful because it requires an explicit representation of the excited states. In practice variational methods are used because all computational methods of ab initio theory can be cast in a variational form, such as HF, FCI, MCSCF, CC, and MP.$^{20}$ The quantum mechanical computational procedures and computer hardware have made it possible to perform such calculations for molecular fragments large enough to reflect the essential features of the local environment.$^{21}$ So, for systems of second row atoms (H–Ne) a minimum 3-21G$^*$ basis set is sufficient. For more accurate chemical shifts the carbon sites are treated with a 6-311+G*2$^*$ basis set. For N-formyl-L-alanine amide residues (10 atoms) 59 contracted Gaussian functions are needed, if only the C* and C# sites are of interest. At the HF level the CPU time goes with 59$^4$ and it takes 24.2 min for a single chemical shift calculation on a 100 Mflops/s machine.$^{24}$ Current standard GHz PCs run about 450 to 500 Mflops/s peak speed. However, the computational effort for molecular dynamics simulations or geometry optimizations is still too large for a routine task. Only parallel computers are capable of performing such ab initio calculations. Therefore, good work has to be done to establish other fast computational methods. One possible approach is a semiempirical quantum chemical method. In this work, the BPT is used to compute the chemical shifts and their pseudoforces.

BPT

If a quantum chemical method is combined with a force field, a suitable form for the representation of the molecular wave function is localized bond orbitals. For every bond defined within the force field a bond orbital can be constructed. In addition to the bond orbitals, antibonds are introduced to account for polarizations. Delocalizations can be treated in the same manner. The molecular wave function $|\Psi_0\rangle$ is the solution of the energy equation for the Hamilton operator $\hat{H}$ and can be constructed from a perturbation series

$$|\Psi_0\rangle = |\Phi_0\rangle + \sum_{i \neq 0} \frac{\langle \Phi_i | \hat{H} | \Phi_0 \rangle}{E_0 - E_i} |\Phi_i\rangle + \cdots,$$

where the molecular ground state wave function $|\Phi_0\rangle$ of the energy $E_0$ is chosen to be a Slater determinant constructed from
The inner shell contributions of the shielding operator can be considered as a sum of one electron operators [see eq. (2)]. In general expression for any local one electron expectation value with overlap approximation. The basis functions of the hybrids are Slater atomic orbitals (SAOs). The exponentials are taken from Burns. It is common to identify the Hamiltonian $\hat{H}$ of eq. (6) with the Fock operator $\hat{F}$ \(^0\). The most important term of the CI is clearly the polarization. This is indicated by rewriting the exited wave function as a polarization state $|\Phi_i\rangle = |\Phi(\vec{r}_i^{*})\rangle$. If we restrict the perturbation series to this first order polarization term it is straight forward to get an expression for the expectation value of one electron operators $\hat{O}$ using eq. (6):

$$
\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = \langle \Phi_0 | \hat{O} | \Phi_0 \rangle + \sum_{i \neq 0} \left( \frac{\langle \Phi(\vec{r}_i^{*}) | \hat{O} | \Phi_0 \rangle}{E_0 - E_i} \langle \Phi_0 | \Phi(\vec{r}_i^{*}) \rangle \right)
$$  

(7)

The Fock operator $\hat{F}$ \(^0\) is thought to be separated into two parts, A and B ($\hat{F} = \hat{F}_A + \hat{F}_B$). Part A is designated to the bond in consideration and B is designated to the rest of the system. This splitting can be performed because delocalizations from A to B are neglected. By using the same ideas as in the PCILO method of Malrieu,\(^28\) it is assumed that the polarity parameters \(d\) are adjusted in a way that the requirement $\langle i_A | \hat{F}_A | i_A^{*} \rangle = 0$ is fulfilled. From eq. (7) it follows

$$
\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = 2 \sum_A \left( \langle i_A | \hat{O} | i_A \rangle + \frac{\langle i_A^{*} | \hat{O} | i_A \rangle}{E_0 - E_i} \langle i_A | \hat{F}_B | i_A^{*} \rangle \right)
$$  

(8)

We write the chemical shift operator $\hat{O} = \hat{\sigma}_{\text{ref}} - \hat{\sigma}_N = \hat{\delta}_N$. It can be considered as a sum of one electron operators [see eq. (2)]. The inner shell contributions of the shielding operator $\hat{\sigma}_N$ and the reference $\hat{\sigma}_{\text{ref}}$ are the same and cancel each other out. The shift operator acts on all electrons of bond orbitals and lone pairs. Hence, the sum of eq. (8) runs over all bonds of subsystem A. However, the expectation value of the shift is roughly proportional to the inverse distance of the electrons from the nucleus N. Only bonds directly connected to the nucleus N are taken into account in first approximation and the sum over A breaks down.

Within the BPT approach the Fock operator $\hat{F}_B$ is approximated by a point charge distribution $\tilde{V}_B$ of system B [eq. (17)]. These are the net atomic charges located at the atomic nuclei positions. Double occupied bond orbitals are only suitable for the description of $\sigma$-bonds and ideal $\pi$-bonds. In order to describe molecular systems with delocalized $\pi$-bonds the occupation number becomes lower than two. Double bonds are treated as superimposed resonance structures. There is no easy way to obtain the occupation number theoretically. Within the framework of the semiempirical BPT the occupation number of conjugated $\pi$-bonds is estimated from an empirical valence formula $\nu = \exp(R - r)/0.37$. The occupation number of a double bond is defined as $n = 2\nu$, and the $\pi$-bond contribution occupation number becomes $n = 2(\nu - 1)$. It depends on the equilibrium value of an ideal single bond $R$ and on the actual bond length $r$. Taking eq. (8) and introducing bond specific constants $\delta_i$ as well as $A_i = 4\langle i | \delta_i^{*} | i \rangle / \Delta E$, one obtains for the chemical shift of atom N

$$
\langle \Psi_0 | \hat{\delta}_N | \Psi_0 \rangle = \sum_{i \in N} \delta_i n_i + A_i n_i^2 \langle i | \tilde{V}_B | i^{*} \rangle
$$  

(9)

The sum runs over all bond contributions belonging to atom N. The polarization parameters $\delta_i$ and $A_i$ can be once obtained from a calibration.\(^31\) A collection of crystal structures and single crystal chemical shift measurements\(^32\) is used to establish a set of linear equations in the form of eq. (9). To some extent, \textit{ab initio} results are also exploited. If the parameters $\delta_i$ and $A_i$ are determined, only the matrix elements of the bond polarization energy $\langle i | \tilde{V}_B | i^{*} \rangle$ [eqs. (22) and (23)] and the occupation numbers have to be calculated. They are all expressed analytically and hence the calculation efficiency of the BPT formula is obvious.

Formula (9) holds for the chemical shift and is also a general expression for any local one electron expectation value with strong polarization dependency. Beside the $^{13}$C chemical shifts,\(^31\) the atomic charges\(^33\) are also calibrated within this approximation. These atomic coordinate dependent charges enter the matrix elements of the bond polarization energy [eqs. (22) and (23)]. By investigating the charge equations, which look similar to eq. (9), it is obvious that charges are estimated from charges. Therefore, calculating charges means solving a set of linear equations. Most CPU time is spent to build up and solve this set of linear charge equations rather than computing the chemical shifts or their related pseudoforces.

### Molecular Force Field with Pseudoforces

The driving force leading to molecular equilibrium is the minimization of the Gibbs free enthalpy:

$$
G(p, T, N) = E_{\text{kin}} + E_{\text{pot}} - TS + pV
$$  

(10)

It is a function of temperature and pressure and represents the energy of a molecular system at constant temperature $T$ and pressure $p$ excluding volume work $pV$ and heat exchange $-TS$. It includes entropic driving forces as, for instance, hydrophobic and hydrophilic effects. In the harmonic approximation,\(^34\) for instance, the gradient of the Gibbs potential can be calculated from the molecular force field and the gradient of the Hessian determinant. The potential energy ($E_{\text{pot}}$) describes the interactions between residues, atoms or atom groups, and molecules. In the calculations the COSMO-S force field\(^19\) is used. Force contributions to the free energy beyond the potential energy are entropic in nature and involve in the harmonic approximation the second and third derivatives of the potential energy and scale linearly with the temperature. Unfortunately, they cannot be calculated efficiently and are not used in practice. However, it follows that in equilibrium state at $T > 0$ the molecular forces do not vanish. Therefore, additional contributions
are introduced to the force field, targeting calculated expectation values to experimental values, in our case $^{13}$C chemical shifts. In order to obtain energetic corrections the polarization energy contribution of the bonds around nucleus N has to be calculated. Using eq. (7) with $\hat{O} = \hat{F}_N$, the part of the Fock operator containing the bond electrons belonging to atom position N, it can be deduced

$$E_N \propto \frac{\langle \Phi_0 | \hat{V}_B | \Phi_i^+ \rangle^2}{E_0 - E_i}$$

(11)

From this, the energetic change due to chemical shift variation can be calculated and gives

$$E_{P}^{N}[\delta] = E_{P}^{N}[\delta_0] + \frac{\partial E_{P}^{N}}{\partial \delta} |_{\delta_0} \Delta \delta + o(\Delta \delta^2)$$

(12)

In this approach the chemical shift is represented as functional of the polarization energies $E_{P}^{N}$ belonging to the nucleus in consideration and vice versa. It can be expanded in a Taylor series around a desired value $\delta_0$:

$$\Delta E_{P}^{N} = -2(\delta - \delta_0) \sum_{i \neq 0}^{N} \frac{\partial E_{P}^{N}}{\partial \delta}(\Phi_0 | \hat{V}_B | \Phi_i^+) \Delta \delta + o(\Delta \delta^2)$$

(13)

This polarization energy change due to chemical shift variation implies a new force contribution and is regarded as the chemical shift pseudoforce. The experimental values are identified as $\delta_0 = \delta^E$ and the theoretical chemical shift as $\delta = \delta^T$. Taking the gradient, one obtains the chemical shift pseudoforce in terms of bond and antibond wave functions:

$$F_{CS}^{\alpha} = 2(\delta^T - \delta^E) \sum_{i \neq 0} \frac{\partial}{\partial \sigma} \left( \frac{\langle i | \hat{V}_B | i^+ \rangle}{\langle i | \hat{V}_B | j^+ \rangle} \right)$$

(14)

It is clear that the direction of any force contribution is the gradient of the bond polarization matrix element. The chemical shift matrix element $\langle i | \hat{S} | j^+ \rangle$ can be expressed by the ratio of the corresponding bond polarization parameter of the charge parametrization to shift parametrization, $\tilde{A}_\alpha / \tilde{A}_\sigma$. Because of simplicity we differentiate only between $\sigma$- and $\pi$-bond contributions and approximate $A_\alpha$ by $\tilde{A}_\sigma / \delta$ and $\tilde{A}_\pi / \delta$:

$$F_{CS}^{\alpha} = 2(\delta^T - \delta^E) \sum_{i \neq 0} \tilde{A}_\sigma \frac{\partial}{\partial \sigma} \left( \frac{\langle i | \hat{V}_B | i^+ \rangle}{\langle i | \hat{V}_B | j^+ \rangle} \right) + \sum_{j \neq 0} \tilde{A}_\pi \frac{\partial}{\partial \pi} \left( j | \hat{V}_B | j^+ \right)$$

(15)

The two weight factors $\tilde{A}_\sigma$ and $\tilde{A}_\pi$ are the mean chemical shift bond polarization parameters. The shift parameter $\delta$ can be approximated by $\delta = \tilde{A} \Delta E / 4$. In order to compute the potential $\hat{V}_B$, as it can be seen from eq. (15), most CPU power is used to derive the charges. Also, the calculation of the chemical shift [eq. (9)] needs the charge distribution. For instance, the estimation of the charge distribution and the chemical shift tensors of the 14 center carbon atoms of a D-mannitol crystal with 442 atoms in total, using crystallographic boundary conditions additionally, takes 5 s on a 450 MHz Pentium II (156 Mflops/s peak speed).

The pseudoforce scales with the relative difference between theoretical and experimental chemical shift $(\delta^T - \delta^E) / \delta$. This is the result of the Taylor expansion of the polarization energy [eq. (12)]. The scaling is multiplied with a scaling constant $S$ to control the magnitude of the pseudoforce with respect to the other interactions of the force field. In order to limit the pseudoforce at large differences $(\delta^T - \delta^E)$ a modified force scaling factor

$$C(\delta^T - \delta^E) = S \frac{\exp((\delta^T - \delta^E) / \delta) - \exp((\delta^E - \delta^T) / \delta)}{\exp((\delta^T - \delta^E) / \delta) + \exp((\delta^E - \delta^T) / \delta)}$$

(16)

is introduced. The functional behavior is shown in Figure 1.

**Application**

**Chemical Shift Geometry Optimization of Crystalline D-Mannitol**

Our method can be used to refine the structures of large molecular systems, for instance, proton positions in crystals. Because of the lack of core electron interaction with the radiation field, proton positions are not always readily derived from X-ray investigations, especially for large molecules. Although for small molecules high resolution structures are common we improve the proton positions of $\beta$-D-mannitol. This acyclic sugar alcohol $\beta$-D-mannitol, $C_6H_{14}O_6$, is of medical and biological interest. As the result of the chiral nature it is optically active in aqueous solution and crystallizes from it in the noncentrosymmetric $\beta$ form. It has the orthorhombic space group P2$_1$2$_1$2$_1$ and was X-ray refined and investigated by its basic physical properties. Additionally, cross-polarization and magic-angle spinning $^{13}$C-NMR spectra were recorded and the isotopic chemical shifts were listed by Wasylishein.

The comparison between the experimental and theoretical chemical shifts, derived by the BPT method, of the D-mannitol

![Figure 1. Functional dependence of the relative force scaling factor C/S with respect to the difference between theoretical and experimental chemical shift: line (––) with shift parameter $\delta = 1.0 \text{ ppm}$; (—) with $\delta = 3.9 \text{ ppm}$, the standard deviation of the BPT chemical shifts; and (–) with $\delta = 10.0 \text{ ppm}$.](image-url)
Table 1. Experimental Data on the D-Mannitol Carbon Atoms (Column 1), Theoretical Results of the X-Ray Structure (Column 2), and of the Chemical Shift Force Field Refined Structure at Various Pseudoforce Scaling Constants (Columns 3 to 10).

<table>
<thead>
<tr>
<th>Chemical shift [ppm]</th>
<th>Experimental Structure</th>
<th>X-Ray Structure</th>
<th>Theoretical 13C Calculations Using the COSMOS Program Applying the Bond Polarization Theory (BPT)</th>
<th>Scaling Constant of the Chemical Shift Pseudoforce</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C1</td>
<td>64.3</td>
<td>61.64</td>
<td>64.47</td>
<td>64.45</td>
</tr>
<tr>
<td>C2</td>
<td>71.7</td>
<td>70.46</td>
<td>73.18</td>
<td>73.17</td>
</tr>
<tr>
<td>C3</td>
<td>69.3</td>
<td>71.23</td>
<td>73.23</td>
<td>73.22</td>
</tr>
<tr>
<td>C4</td>
<td>67.4</td>
<td>68.85</td>
<td>71.47</td>
<td>71.47</td>
</tr>
<tr>
<td>C5</td>
<td>70.5</td>
<td>70.3</td>
<td>72.54</td>
<td>72.54</td>
</tr>
<tr>
<td>C6</td>
<td>62.8</td>
<td>61.23</td>
<td>63.25</td>
<td>63.24</td>
</tr>
<tr>
<td>Deviation</td>
<td>0.168</td>
<td>2.54</td>
<td>2.54</td>
<td>2.49</td>
</tr>
<tr>
<td>Displacement [Å]</td>
<td>—</td>
<td>0.195</td>
<td>0.197</td>
<td>0.199</td>
</tr>
</tbody>
</table>

Rows 1 to 6 show the experimental and theoretical 13C chemical shift values of the X-ray structure before and after the proton position refinement at various scalings. Row 7 compares the average chemical shift deviations from experimental values. Row 8 compares the displacement with respect to the crystal structure and rows 9, 10, 11 list the pseudo-, bonded, and total energies.

An investigation of the average coordinate displacement of the protons with respect to the crystal structure in Table 1 shows a decreasing from 0.197 to 0.128 Å with respect to the scaling. At scaling constants higher than 100 no further change is obtained. There is an outlier at the value of 10. It can be understood by analyzing the total energy plot (Fig. 3).

At small scaling factors the electrostatic contributions dominate the total force field energy. However, as the scaling increases, the pseudoforces enlarge. They are overthrown by the harmonic bond length, bond angle, and dihedral angle forces of the protons. These contributions dominate up to a scaling constant $S$ of 10. At higher scalings, the bonded energy does not increase drastically, but the negative pseudoforces start to dominate all other contribu-

![Figure 2. Theoretical chemical shift deviation from experiment with respect to the logarithm of the pseudoforce scaling constant.](image-url)
Chemical Shift Driven Geometry Optimization

Figure 3. Total energy with respect to the logarithm of the pseudoforce scaling constant.

Figure 4. Superposition of the X-ray and $^{13}$C refined structure (scaling constant $S = 1000$). At the proton positions the 50% probability spheres are drawn.

Figure 5. Refined $\beta$-D-mannitol crystal structure indicating hydrogen bridges. The intramolecular hydrogen bond (*) appears only after refinement.

Nevertheless, the structure stays stable. This means that the proton positions are controlled exclusively by the chemical shift pseudoforces. The errors of the NMR experiment and the chemical shift calculations limit the precision of the structure determination. However, the refinement deviation is smaller than the uncertainties of the X-ray diffraction. The given average proton displacement parameter derived from the temperature factor is about 0.2 Å. The standard deviation of the refined structure ($S = 100$, $S = 1000$) with respect to the X-ray structure is about 0.13 Å.

In Figure 4, the superposition of the X-ray and the proton refined structure ($S = 1000$) is shown. The spheres at the proton positions are the isotropic 50% probability ellipsoids. The concept of the probability plots was taken from ORTEP-III.38

It is well known that hydrogen bond systems stabilize the crystal structure. In the COSMOS force field,16 hydrogen bonds are searched according to the geometric rule39 that the total bond length contraction caused by H bond formation should be equal to or greater than twice the van der Waals radius of the hydrogen atom (1.6 Å). If the O···H distance gets lower than the mean covalent bond radii of H and O plus 1.6 Å, then the van der Waals interaction between these two atoms is switched off. After the refinement we obtained one additional intramolecular hydrogen bond. The X-ray data provided two. Figure 5 shows that every mannitol molecule in the lattice has three intramolecular hydrogen bonds.

Results and Conclusion

The bond polarization model is an effective method for calculating $^{13}$C chemical shifts and their gradients. From the BPT chemical shift, pseudoforces for geometry optimization and molecular dynamics are derived. It is shown that the chemical shift calculation due to the BPT can predict the structure of molecular systems. The coordinate dependency of the BPT chemical shift is due to the polarizing charges and $\pi$ bond length. Inner coordinates, bond angles and bond length, have indirect influence.

However, it should be noted that the standard deviation of the BPT chemical shifts is about 3.9 ppm and there are absolute errors up to 10 ppm. Empirical investigations on the reproducibility of $^{13}$C NMR shift values40 also show a deviation up to 2.5 ppm and value ranges up to 7 ppm. Thus, in general, properly parameterized and measured chemical shift values should be used for structure refinement.

Appendix

Integrals

The coordinate dependent part of the chemical shift, eq. (9), and of the pseudopotentials, are the occupation numbers and the matrix elements of the Fock operator $\langle \uparrow | \hat{F}_B | \uparrow \rangle$. In the BPT approximation the Fock operator is a sum of all point charges polarizing the bonds $| \uparrow \rangle$:

$$\hat{F}_B \approx \hat{V}_B = \sum_{\mathbf{r}} \frac{q_{\mathbf{r}}}{| \mathbf{r} - \mathbf{z} |}$$

(17)
Table 2. Bond Orbital Matrix Elements of the Fock Operator (a) and the Derivatives of These Matrix Elements with Respect to $\mathbf{R}$ (b), $\cos(\theta)$ (c), and $\phi$ (d).

<table>
<thead>
<tr>
<th>Atomic orbitals</th>
<th>Integrals and Integral Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(ns)$</td>
<td>$\langle \phi^{nlm}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(ns)$</th>
<th>$(ns)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(np_z)$</td>
<td>$\frac{1}{R^2} + \frac{3}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} (3 \cos^2(\theta) - 1) \right)$</td>
</tr>
<tr>
<td>$(ns)$</td>
<td>$(np_z)$</td>
</tr>
<tr>
<td>$\frac{1}{R^2} - \frac{3}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} (3 \cos^2(\theta) - 1) \right)$</td>
<td></td>
</tr>
<tr>
<td>$(np_x)$</td>
<td>$(np_x)$</td>
</tr>
<tr>
<td>$\frac{1}{R^2} - \frac{3}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} (3 \cos^2(\theta) - 1) \right)$</td>
<td></td>
</tr>
<tr>
<td>$(np_y)$</td>
<td>$(np_y)$</td>
</tr>
<tr>
<td>$\frac{1}{R^2} - \frac{3}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} (3 \cos^2(\theta) - 1) \right)$</td>
<td></td>
</tr>
</tbody>
</table>

(c) Integrals and Integral Derivatives

<table>
<thead>
<tr>
<th>Atomic orbitals</th>
<th>Integrals and Integral Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(ns)$</td>
<td>$\frac{\partial}{\partial \mathbf{R}} \langle \phi^{nlm}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(ns)$</th>
<th>$(ns)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(np_z)$</td>
<td>$\frac{6}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} \cos(\theta) \right)$</td>
</tr>
<tr>
<td>$(ns)$</td>
<td>$(np_z)$</td>
</tr>
<tr>
<td>$\frac{1}{\sqrt{R^2}} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} \right)$</td>
<td></td>
</tr>
<tr>
<td>$(np_x)$</td>
<td>$(np_x)$</td>
</tr>
<tr>
<td>$\frac{6}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} \cos(\theta) \cos(2\phi) \right)$</td>
<td></td>
</tr>
<tr>
<td>$(np_y)$</td>
<td>$(np_y)$</td>
</tr>
<tr>
<td>$\frac{6}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} \cos(\theta) \sin(2\phi) \right)$</td>
<td></td>
</tr>
</tbody>
</table>

(d) Integrals and Integral Derivatives

<table>
<thead>
<tr>
<th>Atomic orbitals</th>
<th>Integrals and Integral Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(ns)$</td>
<td>$\frac{\partial}{\partial \phi} \langle \phi^{nlm}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(ns)$</th>
<th>$(ns)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(np_z)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$(np_z)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$(np_z)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$(np_x)$</td>
<td>$(np_x)$</td>
</tr>
<tr>
<td>$\frac{6}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} (1 - \cos^2(\theta))(2 \sin(\phi)) \right)$</td>
<td></td>
</tr>
<tr>
<td>$(np_y)$</td>
<td>$(np_y)$</td>
</tr>
<tr>
<td>$\frac{6}{R^4} \left( \frac{(2n+1)(2n+2)}{5(2n)^2} (1 - \cos^2(\theta))(2 \cos(\phi)) \right)$</td>
<td></td>
</tr>
</tbody>
</table>
The charges are computed before the chemical shift evaluation takes place.

In order to calculate the matrix elements \( \langle i | \hat{\mathbf{V}}_B | i^* \rangle \), bond and antibond wave functions have to be explicitly given:

\[
\langle i | \hat{\mathbf{V}}_B | i^* \rangle = \sqrt{\frac{1 - d^2}{2}} \langle \chi_a | \hat{\mathbf{V}}_B | \chi_b \rangle + \sqrt{\frac{1 + d}{2}} \langle \chi_b | \hat{\mathbf{V}}_B | \chi_a \rangle
\]

Bonds \( |i\rangle \) and antibonds \( |i^*\rangle \) are created from bond orbitals and hybrid functions:

\[
\langle \mathcal{L} | \chi \rangle = \chi(\mathcal{L}) = \sum_k h_k \phi_k(\zeta, \mathcal{L})
\]

The hybrids themselves are formed from Slater atomic orbitals:

\[
\langle \mathcal{L} | \phi \rangle = \phi_{\text{Slater}}(\zeta, \mathcal{L}) = \sqrt{\frac{(2\zeta)^{n+1}}{(2n)!}} \zeta^{-n-1} \exp(-\zeta r) Y_{nm}(\theta, \phi)
\]

The exponents \( \zeta \) are calculated according to the rules of Burns.

If the polarity parameter \( d \) is considered much lower than 1 and overlap elements such as \( \langle \chi_a | \hat{\mathbf{V}}_B | \chi_b \rangle \) are regarded as small, the matrix elements of the bond polarization energy have the form

\[
\langle i | \hat{\mathbf{V}}_B | i^* \rangle = \frac{1 - d^2}{2} \left[ \langle \chi_a | \hat{\mathbf{V}}_B | \chi_a \rangle - \langle \chi_b | \hat{\mathbf{V}}_B | \chi_b \rangle \right]
\]

Therefore, calculating bond orbital matrix elements \( V_B = \langle \chi | \hat{\mathbf{V}}_B | \chi \rangle \) means solving a sum of integrals:

\[
V_B(R_x) = \sum_x \sum_k h_k^2 \int d^3 R \frac{q_x}{R_x - d} \phi_k(\mathcal{L})
\]

The factor \( 0.5 \sqrt{1 - d^2} \) becomes part of the polarization parameter \( A_1 \). To evaluate the pseudoforces the gradients also have to be taken. In Table 2 the integrals and their derivatives, which are necessary for sp\(^2\) and sp\(^3\) carbon bond contributions, are listed.

References

32. (a) Veeman, W. S. Prog NMR Spectrosc 1984, 20, 193; (b) Sherwood, M. H.; Alderman, D. W.; Grant, M. G. J Magn Reson 1989, 466.