Comparison of the computational NMR chemical shifts of choline with the experimental data

C. Alcorn\textsuperscript{1}, M. Cuperlovic-Culf\textsuperscript{1,2} and K. Ghandi\textsuperscript{1}\textsuperscript{*}

1. Department of Chemistry and Biochemistry, Mount Allison University
2. Institute for Information Technology, National Research Council of Canada

kghandi@mta.ca; Miroslava.Cuperlovic-Culf@nrc-cnrc.gc.ca

Abstract. One of the main biological markers of the presence of cancer in living patients is an over-expression of total choline (tCho), which is the sum of free choline and its derivatives. \textsuperscript{1}H Magnetic Resonance Spectroscopy, or H-MRS, enables the quantification of tCho via its proton spectra, and thus has the potential to be a diagnostic tool for the presence of cancer and an accurate early indicator of the response of cancer to treatment. However, it remains difficult to quantify individual choline derivatives, since they share a large structural similarity ((CH\textsubscript{3})\textsubscript{3}-N\textsuperscript{+}-CH\textsubscript{2}-CH\textsubscript{2}-O\textsuperscript{-}), of which the strongest signal detectable by MRS is that of the choline “head group”: the three methyl groups bonded to the nitrogen. This work used ACENet, a high performance computing system, to attempt to model the NMR parameters of choline derivatives, with the focus of this report being free choline. Optimized structures were determined using Density Functional Theory and the B3LYP electron correlation functional. The Polarizable Continuum Model was used to evaluate solvent effects. The Gauge-Invariant Atomic Orbital method was found to be the superior method for calculating the NMR parameters of cholines.

1. Introduction
A wide variety of cancers radically alter the metabolism of free choline (Cho) and its derivatives in patients [1-5]. The increase in total choline (tCho) concentration observed in tumours is universal for many types of cancers, and in fact the quantity of tCho present \textit{in vivo} is used as a diagnostic biomarker of the presence and aggressiveness of cancer [1-2]. Quantification of tCho is also useful in differentiating between malignant and benign tumours [4], the effectiveness of cancer treatments, and the targeting of radiation therapy [2].

While tCho is a useful biomarker of cancer, knowledge of the specific choline derivative concentrations present in tumours would give more information about the altered metabolic pathways, and thus clues on tumour sub-types, prognosis and treatment planning. For example, in the case of breast and ovarian cancer cells, it has been shown that phosphocholine (PCho) concentration is 10-fold higher than that of glycerophosphocholine (GPCho) [1-2,4]. In fact, the ratio of PCho to GPCho is altered in most types of cancers [1]. Change in the PCho/GPCho ratio can be correlated with increased activity of choline kinase. Introduction of choline kinase inhibitors have been shown to exhibit anticancer activity [4].

Although \textit{in vitro} high field \textsuperscript{1}H and \textsuperscript{31}P NMR spectroscopy can be used for measurement of concentrations of individual choline derivatives, it remains difficult to determine their concentrations accurately due to the overlapping peaks of different choline derivatives. Therefore, computational methods are necessary to model the NMR parameters of choline derivatives. This work used ACENet, a high performance computing system, to attempt to model the NMR parameters of choline derivatives, with the focus of this report being free choline. Optimized structures were determined using Density Functional Theory and the B3LYP electron correlation functional. The Polarizable Continuum Model was used to evaluate solvent effects. The Gauge-Invariant Atomic Orbital method was found to be the superior method for calculating the NMR parameters of cholines.
in vivo. Cholines are detected in a clinical setting by Magnetic Resonance Spectroscopy (MRS) techniques [1-3,5-7]. The main proton peak of interest with respect to total choline detection are the nine chemically equivalent protons found on the three methyl groups bonded to the quaternary nitrogen (see Figures 1–3).

![Figure 1. Chemical structure of acetylcholine](image1)

![Figure 2. Chemical structure of choline](image2)

![Figure 3. Chemical structure of phosphocholine](image3)

All choline derivatives share this structural similarity, making the quantification of specific choline derivatives difficult due to the resolution of current clinical MRS instruments, which stems from technical difficulties for in vivo detection such as low field regulations, magnetic field inhomogeneities, sample size, and a sample embedded in vivo, as well as inability to run experiments for extended periods of time [6]. Thus, knowledge of subtle differences in the NMR spectra of choline derivatives may help us in developing a model spectrum for a mixed sample of choline derivatives, for direct comparison to an MRS spectrum. This model can then be utilized to decipher the individual choline derivative concentrations present in a mixed sample.
The only structural differences in the three choline derivatives of interest to us (acetylcholine (AChO), Cho, and PCho) are the groups bonded to the oxygen in the CH₂O group, as shown in Figures 1–3 respectively. These differences will affect the NMR parameters in two ways. First, these groups will donate or withdraw unequal amounts of electron density from the remainder of the molecule by simple electronegativity, and secondly, these groups provide several sites for possible hydrogen bonding, which in turn would cause polarization of the electron density by the solvent.

Computational chemistry is a very effective tool in calculating the NMR parameters of molecules [8-10], such as complicated organic molecules [11-12] and molecules of biological importance [13-15]. Accurate geometries are paramount when using computational methods to calculate NMR parameters. In this work we calculated the geometry of Cho in both the gas and liquid phase, and compared our results to experimental crystal structures. We then calculated the NMR parameters and compared our results with the experimental data to gauge the accuracy of NMR calculations. From calculation, it is possible to obtain chemical shifts, coupling constants as well as relaxation parameters; however this work deals only with the chemical shifts.

Computational chemistry is a very effective tool in calculating NMR parameters of ionic molecules [8-10]. In order to use computational methods to calculate NMR parameters, we must have accurate geometries. In this work we calculated the geometry of Cho in the gas and liquid phase, and compared our method to experimental crystal structures. We then calculated the NMR parameters and compared our results with the experimental data to gauge the accuracy of NMR calculations. From calculations it is possible to obtain chemical shifts, coupling constants as well as relaxation parameters; however this work deals only with the chemical shifts.

2. Computational methods

As mentioned above, the first type of calculation performed in this work was gas phase quantum calculations. Moving from the gas to the liquid phase, the simplest method for accounting for solvent effects is through what is known as continuum solvent models. The PCM model functions by creating a cavity in the solvent for the solute, and surrounds the cavity with a continuous medium with dielectric constant ε. PCM defines the shape of the cavity as the overlap of several atom centered spheres, whose radii are 1.2 times the van-der-Wall radius of the atom it is centered on [16]. The electrostatic potential and surface point charge, \( \phi(\mathbf{r}) \) and \( Q_s \), are initially unknown; they are solved by iteration [17]. Iterations are performed until the electron density is stabilized to within a certain threshold value.

In this work, the Integral Equation Formalism variant of the Polarizable Continuum Model (IEF-PCM) [18] was used exclusively, with water as the solvent medium, having a dielectric constant of \( \varepsilon_r = 78.3553 \).

It must be mentioned that solvent effects in this study were limited to liquid water to simulate solvation in the cytoplasm of blood-free cells, though the types of environments found in vivo are much more complicated than simple water. Other dissolved constituents of the cytoplasm are not considered at this time, nor is the binding of cholines to lipids or cell membranes.

All calculations in this work were performed using the Gaussian 03 package [19]. Structures were optimized using Density Functional Theory (DFT) [20] with the B3LYP correlation function [21-22], at a large Pople’s basis set [23] of 6-311++G(2d,2p). DFT [20] defines all of the electrons as a density function with the contributions of the total energy as a sum of kinetic energy (\( E_T \)), electron-nuclear attraction (\( E_{ea} \)), and the electron-electron repulsion which is a sum of the Coulomb (\( E_C \)) integral, and an exchange/correlation energy (\( E_{XC} \)) [24].

Numerous exchange/correlation energies exist. One of the successful ones, used in this work, is the hybrid B3LYP [21-22] functional derived by Lee, Yang and Parr, which is a linear combination of exchange energies found by other methods.

Another exchange/correlation relation, the PBE0 functional [25], was also used in this work. PBE0 is a variant of B3LYP, which only differs from B3LYP by weighting the exchange and correlation energies by 25% and 75% respectively.
The geometry of Cho was optimized in both the gas phase and by using the IEF-PCM model, and was compared with experimental crystal structures [26] for structural confirmation. From the optimized structures, calculations of the nuclear shielding were done using both the Gauge-Invariant Atomic Orbital (GIAO) [27-28] and Continuous Set of Gauge Transformations (CSGT) methods [29], at the same method and basis set.

A full description of the GIAO and CSGT methods for the calculation of NMR parameters is beyond the scope of this paper. In short, the GIAO method is derived by considering the application of a magnetic field as an energetic perturbation to the original Self-Consistent Field (SCF) equations of the Hartree-Fock (or DFT) method [18]; the first term of the Taylor series expansion of these equations with respect to the magnetic field describes the magnetic shielding tensor $\sigma_B$ [27]. However, the GIAO method introduces a gauge-dependence by multiplying the wavefunctions $\psi_s$ by a gauge-dependent term; the CSGT method avoids this problem by calculating the current density induced by the perturbing magnetic field for all points, and taking those points as gauge origin [29].

3. Results and discussion
The computational work was done stepwise in the following order. 1) First, structures were optimized to an energy minimum (with and without IEF-PCM for solvent effects). 2) A frequency calculation was performed on the optimized structure for confirmation of a local energy minimum configuration. 3) Calculations of the nuclear shielding were done using both the GIAO and CSGT methods.

The structures of the gas phase and IEF-PCM optimized choline cation is compared to that of the choline chloride crystal structure found by Senko and Templeton [26], with bond lengths, angles and dihedral angles reported in Table 1. Atom labels are referenced from the Cho geometry shown in Figure 4.

Only minor differences are seen in bond lengths for both gas and liquid phase optimized geometries. However, the gas phase geometry reports significant differences in bond angles, and a very large difference in the dihedral angle of the gas phase. For the IEF-PCM optimized geometry, the only significant variation from the crystal structure is observed in one of the dihedral angles. However, since we are comparing the solid crystal structure, which has paired cations/anions and very closely packed, to a solvated cation alone, an 18° difference in the dihedral angle is reasonable. The dihedral angle itself is only a “twist” in the backbone of the molecule, which could be expected to be required to form a rigid solid. However, the 84° difference reported in our gas phase calculations is much too large and is a reflection of the fact that the positive charge on the nitrogen will attract the electronegative oxygen atom, which is easy to achieve through torsion. This attraction seems to be adequately dampened by the solvent field.
Table 1. Comparison of the experimental crystal structure of Cho [26] to our calculated geometries

| Parameter | Crystal Structure of Cho [26] | Gas Phase | IEF-PCM |
|-----------|-------------------------------|-----------|---------|
| r(N1-C3)  | 1.52                          | 1.51      | 1.50    |
| r(N1-C5)  | 1.51                          | 1.51      | 1.50    |
| r(N1-C13) | 1.54                          | 1.51      | 1.51    |
| r(N1-C6)  | 1.60                          | 1.52      | 1.53    |
| r(C2-C17) | 1.56                          | 1.53      | 1.51    |
| r(C17-O20)| 1.45                          | 1.41      | 1.42    |
| ∠(O20-C17-C2) | 112                      | 103      | 111    |
| ∠(C17-C2-N1) | 111                      | 117      | 117    |
| ∠(C5-N1-C2-C17) | 172                      | 180      | 168    |
| ∠(O20-C17-C2-N1) | 84                      | 0        | 66     |

Bond lengths shown in angstroms, angles in degrees.

Figure 4. The optimized structure of Cho (with IEF-PCM, using B3LYP/6-311++G(2d,2p)) with atom labels for reference to the geometry in Table 1.
Figure 5. Gaussian NMR spectrum of the choline cation. Label assignments are in Figure 4.

Considering that our IEF-PCM method gives an acceptable geometry, we feel confident to use it to calculate the NMR parameters for free choline. However, the NMR parameters were also calculated from our gas phase geometry for a comparison of the difference between calculated NMR parameters method by method.

Calculated chemical shifts are referenced to a separate calculation of TMS, at the same method and basis set. The rotational averaging of chemically equivalent protons is not represented in these types of calculations, since a geometry optimization leads to a rigid structure. Thus, a GIAO NMR output can have a different chemical shift for chemically equivalent protons, as can be seen in Figure 5. For this reason, the values of chemically equivalent protons reported in this work have been averaged.

Table 2 lists the experimental $^1$H NMR signals of Cho, obtained from the spectrum of choline chloride from the Human Metabolome Data Base (HMDB) [30]. The spectrum is shown in Figure 6. The calculated NMR chemical shifts of gas phase optimized Cho are listed in Table 3, which are compared to experimental results. As can be seen in Table 3, the chemical shifts calculated by the GIAO method have a lower mean absolute deviation than those calculated by the CSGT method for both the PBE0 and B3LYP correlation functions. For this reason, only the GIAO method is considered further in this work. Lastly, when incorporating the IEF-PCM model, we can see from Table 3 that the B3LYP functional gives the most accurate chemical shifts, though the mean absolute deviation is in very close in agreement to that of the PBE0 functional.
Figure 6. $^1$H NMR spectrum of 115 mM Cho [11]. See Table 2 for assignment of peaks.
### Table 2. Experimental chemical shifts of Cho in aqueous solution

| H Signal          | Cho (ppm) |
|-------------------|-----------|
| $(\text{CH}_3)_3\text{N}$ | 3.19 (9)  |
| NCH$_2$           | 3.51 (1.93) |
| CH$_2$O           | 4.06 (1.88) |

Relative areas of integration are shown in parenthesis, with normalized values underlined.

### Table 3. Calculated chemical shifts of gas phase optimized Cho by method

| Proton Group | CSGT, PB0 | CSGT, B3LYP | GIAO, PB0 | GIAO, B3LYP | GIAO, PB0 | GIAO, B3LYP | Exp. [30] |
|--------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|
| NCH$_2$      | 3.57      | 3.59        | 3.37      | 3.40        | 3.45      | 3.51        | 3.51      |
| CH$_2$O      | 4.44      | 4.48        | 4.21      | 4.26        | 4.26      | 4.33        | 4.06      |
| N$(\text{CH}_3)_3$ | 3.22      | 3.22        | 3.04      | 3.05        | 3.06      | 3.09        | 3.19      |
| O-H          | 1.12      | 1.28        | 1.10      | 1.26        | 2.34      | 3.21        | N/A       |

| Mean Absolute Deviation | 0.163 | 0.184 | 0.145 | 0.152 | 0.130 | 0.126 | -         |

The chemical shifts listed in Table 3 are calculated from our gas phase geometry, which could be less accurate than our geometry optimized with IEF-PCM. However, the values listed in Table 3 are helpful in gauging the mean absolute deviation in chemical shifts from method to method. The largest difference in mean absolute deviation reported in Table 3 is 0.06 ppm, which is on par with the absolute errors between calculated and experimental chemical shifts. Thus, each method is reasonably accurate. However, continuing our analysis using all the methods reported in Table 3 would be computationally expensive, and thus we rely on the GIAO method, using both IEF-PCM and the B3LYP functional.

Table 4 lists the IEF-PCM/GIAO/B3LYP calculated chemical shifts of the IEF-PCM optimized structure of Cho listed in Table 2. As can be seen in Table 4, the mean absolute deviation is higher than the lowest value in Table 3. However, the mean absolute deviation is increased mainly by the larger error on the CH$_2$O peak, which is also overestimated by all methods reported in Table 2. Furthermore, the remarkable accuracy of the calculated chemical shift of the $(\text{CH}_3)_3\text{N}$ peak, which is the main peak of interest in H-MRS, is quite promising.

As mentioned in our introduction, explicit hydrogen bonding can also alter the chemical shifts of molecules in solution. This type of solvation has been neglected in this work. Thus, the future work in this area remains to see if the values reported in Table 4 can be further refined by the addition of explicit water molecules to the optimized structures of cholines.
Table 4. Calculated chemical shifts of IEF-PCM optimized Cho by IEF-PCM/GIAO/B3LYP

| Signal   | Cho (ppm)  |
|----------|------------|
| NCH\textsubscript{2} | 3.56 (3.51) |
| CH\textsubscript{2}O | 4.45 (4.06) |
| N(CH\textsubscript{3})\textsubscript{3} | 3.19 (3.19) |
| OH      | 3.18 (N/A) |

Mean Absolute Deviation: 0.153

Experimental chemical shifts shown in parenthesis.

4. Conclusion

Optimized structures using DFT, with both the B3LYP functional and IEF-PCM for evaluation of solvent effects gives a reasonable geometry compared to the crystal structure of choline. The small differences between our computational results and the crystal structure (mostly the dihedral angles) are most probably due to the very different nature of the crystalline and aqueous environments. The GIAO method was shown to lead to reasonable agreement with the experimental chemical shifts when combined with IEF-PCM and the B3LYP functional. Slight differences in chemical shifts could be due to explicit solvation effects (hydrogen bonding). Our method is a first step towards future calculation of NMR parameters and line shapes for choline and choline derivatives under different conditions present in vivo.

5. Acknowledgements

Computational work was performed on the Atlantic Computational Excellence Network (ACENet) within the consortium of Compute Canada, who has our thanks for their support. This research was financially supported by the Natural Sciences and Engineering Research Council of Canada and the National Research Council Institute for Information Technology (NRC-IIT). Their support is gratefully acknowledged. Special thanks go to Joey Bernard from ACENet for extensive technical support in the area of parallel computing.

References

[1] Glunde K and Bhujwalla Z 2011 *Semin. Oncol.* **38**, 26-41
[2] Glunde K, Jacobs M and Bhujwalla Z 2006 *Expert Rev. Mol. Diagn.* **6**, 821-29
[3] Morita N, Harada M, Otsuka H, Melhem E and Nishitani, H 2010 *Magn. Reson. Med. Sci.* **9**, 167-75
[4] Katz-Brull R, Seger D, Rivenson-Segal D, Rushkin E and Degani H 2002 *Cancer Research* **62**, 1966-70
[5] Haddadin I, McIntosh A, Meisamy S, Corum C, Styczynski-Snyder A, Powell N, Nelson M, Yee D, Garwood M and Bolan P 2009 *NMR Biomed.* **22**, 65-76
[6] Soares D and Law M 2009 *Clin. Radiol.* **64**, 12-21
[7] Ramirez de Molina A, Rodriguez-González A, Gutiérrez R, Martinez-Piñeiro L, Sánchez J, Bonilla F, Rosell R and Lacal J 2002 *Biochem. Biophys. Res. Commun.* **296**, 580-83
[8] Kaupp M, Buhl M and Malkin V 2004 *Calculation of NMR and EPR Parameters: Theory and Applications*, Wiley-VCH, Verlag GmbH & Co. KGaA, Weinheim, Germany
[9] Helgaker T, Jaszunski M and Ruud K 1999 *Chem. Rev.* **99**, 293-352
[10] Bagno A and Saielli G 2007 *Theor. Chem. Acc.* **117**, 603-19
[11] Dwan J, Durant D and Ghandi K 2008 *Cent. Eur. J. Chem.* **6**(3), 3447-58
[12] Rychnovsky S 2006 *Org. Lett.* **8**, 2895-8.
[13] Clore M and Schwieters C 2002 \textit{Curr. Opin. Struc. Biol.} 12, 146-53.

[14] Gervais, C, Dupree, R, Pike, K, Bonhomme, C, Profeta, M, Pickard, C and Mauri, F 2005 \textit{J. Phys. Chem. A}, 109, 6960-9

[15] Allouche A, Aubert-Frecon Mand Raveron-Demilly D 2007 \textit{Phys. Chem. Chem. Phys.} 9, 3098-103

[16] Tomasi J and Persico M 1994 \textit{Chem. Rev.} 94, 2027-94

[17] Levine I 2008 \textit{Quantum Chemistry 6th Ed.}, Prentice-Hall, Englewood Cliffs, NJ

[18] Tomasi J, Mennucci B and Cammi R 2005 \textit{Chem. Rev.} 105, 2999-3093

[19] Gaussian 03, Revision C.02, Frisch M, Trucks G, Schlegel H, Scuseria G, Robb M, Cheeseman J, Scalmani G, Barone V, Mennucci B, Petersson G, Nakatsuji H, Caricato M, Li X, Hratchian H, Izmaylov A, Bloino J, Zheng G, Sonnenberg J, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery, Jr J, Peralta J, Ogliarom F, Bearpark M, Heyd J, Brothers E, Kudin, K, Staroverov V, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant J, Iyengar S, Tomasi J, Cossi M, Rega N, Millam N, Klene M, Knox J, Cross J, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann R, Yazyev O, Austin A, Cammi R, Pomelli C, Ochterski J, Martin R, Morokuma K, Zakrzewski V, Voth G, Salvador P, Dannenberg J, Dapprich S, Daniels A, Farkas Ö, Foresman J, Ortiz J, Cioslowski J and Fox D 2003 \textit{Gaussian Inc.}, Wallingford, Connecticut

[20] Parr R and Yang W 1989 \textit{Density Functional Theory of Atoms and Molecules}, Oxford University Press, Oxford, United Kingdom

[21] Becke A 1993 \textit{J. Chem. Phys.} 98, 5648-52

[22] Lee C, Yang W and Parr R 1988 \textit{Phys. Rev. B} 37, 785-89

[23] Foresman J and Frisch A 1996 \textit{Exploring Chemistry with Electronic Structure Methods, 2nd Ed.} Gaussian Inc, Pittsburgh, PA

[24] Engel T and Reid P 2006 \textit{Physical Chemistry 2nd Ed.}, Pearson Education, San Francisco, CA

[25] Adamo C and Barone V 1999 \textit{J. Chem. Phys.} 110, 6158-69

[26] Senko M and Templeton D 1960 \textit{Acta Cryst.} 13, 281-85

[27] Ditchfield R 1974 \textit{Mol. Phys.} 27, 789-807

[28] Schreckenbach G and Ziegler T 1995 \textit{J. Phys. Chem.} 99, 606-11

[29] Keith T and Bader R 1993 \textit{Chem. Rev. Lett.} 210, 223-31

[30] Wishart D, Tzur D, Knox C, Eisner R, Chi Gio A, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L., Lewis M, Coutouly M.-A, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau D, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan G, MacInnis G, Meljie A, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes B, Vogel H and Querengesser L 2007 \textit{Nucleic Acids Res.} 35, D521-6