An Ab Initio Quantum Chemical Investigation of ^{43}Ca NMR Interaction Parameters for the Ca^{2+} Sites in Organic Complexes and in Metalloproteins

Alan Wong,[†] Danielle Laurencin,[†] Gang Wu,[‡] Ray Dupree,[†] and Mark E. Smith*[,†]

Department of Physics, University of Warwick, Coventry, CV4 7AL, U.K., and Department of Chemistry, Queen's University, 90 Bader Lane, Kingston, Ontario, K7L 3N6, Canada

Received: February 3, 2008; Revised Manuscript Received: July 6, 2008

We have carried out an extensive ab initio quantum chemical (QC) ^{43}Ca NMR study on a series of Ca−O organic compounds and three different Ca-bound proteins and found that the HF/6-31G* level of function can reliably predict ^{43}Ca NMR interaction parameters (δ_{iso} and χ_q), especially for organic solids. This QC study finds correlations between Ca−O bond environment (mean distance and coordination number) and $\delta_{iso}(^{43}Ca)$. Although relatively small values of $\chi_q(^{43}Ca)$ are found for Ca−O organic compounds with a coordination number between 6 and 10, the QC shows that $\chi_q(^{43}Ca)$ is sensitive to the Ca−O coordination geometry of the Ca^{2+} sites in metalloproteins—a potentially important observation. An application of such ab initio QC ^{43}Ca NMR studies is in characterizing the Ca−O bonding environment around target Ca^{2+} sites. As an example, we propose a new potential analytical approach using the absolute ^{43}Ca chemical shielding constant to investigate the hydration shell of Ca^{2+} in a dilute $CaCl_2$ aqueous solution. Furthermore, by adopting a NMR methodology similar to that reported in Wong et al. *Chem. Phys. Lett*. **2006**, *427*, 201, natural abundance ^{43}Ca MAS NMR spectra of Ca(L-glutamate)$_2$•4H$_2$O were recorded, and $\delta_{iso}(^{43}Ca)$ and the quadrupolar parameter (P_q) were estimated to be 6.6 ppm and 0.8 MHz, respectively.

Introduction

Approximately one-third of all known proteins require a metal ion for their structural stability and/or physiological function. Thus, it is important to understand the roles of metal ions in metalloproteins. Among the biologically important divalent cations, Ca^{2+} is one of the most abundant ions in biological systems. It is a constituent element of a majority of the inorganic phase of the bone materials in animals.[1] Although Ca^{2+} is an important cation in life sciences, there are relatively few spectroscopic techniques able to detect Ca^{2+}. It is invisible to UV and EPR spectroscopic investigation because of its closed electron shells. Despite the difficulties of characterizing Ca^{2+}, solution ^{43}Ca NMR has been successfully applied to probe high-affinity Ca^{2+} binding sites in specific cell types and enabled titration experiments to investigate the Ca^{2+} binding properties.[2−5] Most of these biological systems are dilute in Ca^{2+} ranging from 10^{-7} to 10^{-5} M. On the other hand, there are special cases where a high level of Ca^{2+} can be found in biological systems such as specialized cells involved in bone formation and the inorganic phase of bone materials. These, however, cannot be readily studied by solution ^{43}Ca NMR, because in some respects these materials are more like solids. For this reason, solid-state ^{43}Ca NMR is a potentially attractive method to characterize Ca^{2+} in such materials. However, the natural isotopic abundance of the NMR-active ^{43}Ca nucleus is rather low (0.14%), and ^{43}Ca also has a small magnetic moment, such that it is termed a low-γ nucleus.[6] These two intrinsic properties hinder ^{43}Ca NMR studies of solid materials, because they lead to a weak NMR signal. Early solid-state ^{43}Ca NMR studies relied on the use of large sample-volume rotors (9.5 and 14 mm diameter) to increase the signal by maximizing the amount of sample.[7,8] These studies clearly illustrated the potential of ^{43}Ca NMR spectroscopy for accurately characterizing Ca−O environments in simple inorganic materials. Since then, there has been a growing interest in solid-state ^{43}Ca NMR studies. For example, ^{43}Ca NMR has very recently been used successfully to characterize Ca^{2+} in glasses,[9−12] in geopolymers,[13] in hydroxapatite,[14] and in the polymorphs of $CaCO_3$ including vaterite.[15] To date, there is only one systematic solid-state ^{43}Ca NMR study on a series of Ca−O environments in organic compounds[16] that mimic the Ca^{2+} binding sites in biomaterials. The lack of ^{43}Ca NMR studies on Ca-containing bio/organic compounds is mainly due to the larger molecular size in these materials which result in lower Ca content compared with the results of previous studies of inorganic materials and hence a weaker signal.

In general, especially for NMR studies of insensitive nuclei, ab initio quantum chemical (QC) NMR calculations are often carried out in parallel with NMR experiments to provide more background information about the local environment prior to detailed experimental investigations. To date, QC NMR calculations have proven to be a good complementary tool for characterizing metal−oxygen interactions, either from the metal ions (^{23}Na,[17] ^{25}Mg,[18] and ^{67}Zn[19]) or from the oxygen[20] perspective. Because of the extreme difficulty of solid-state ^{43}Ca NMR studies for large biomolecular systems, it is even more important to establish a computational approach for ^{43}Ca NMR parameters as a complementary tool to study the Ca^{2+} binding sites in macromolecules. Kwan et al.[3] have applied QC ^{43}Ca NMR calculations along with ^{43}Ca NMR experiments to predict the structural environments around the Ca^{2+} site in a Ca-bound guanosine complex. However, to our knowledge, there is no systematic QC study of ^{43}Ca NMR parameters to establish correlations of the local environment around Ca^{2+} sites and the corresponding ^{43}Ca NMR interaction parameters for organic Ca^{2+} complexes. Therefore, the primary objectives of this study

* Author for correspondence. Tel: 44-(0)24-7652-2380. Fax: 44-(0)24-7669-2016. Email: M.E.Smith.1@warwick.ac.uk.
† University of Warwick.
‡ Queen's University.

10.1021/jp801015y CCC: $40.75
Published on Web 09/06/2008

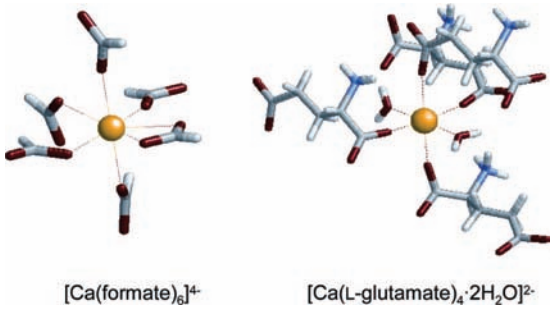

Figure 1. Molecular clusters of Ca(formate)$_2$ and Ca(L-glutamate)$_2$•4H$_2$O used for the QC calculations of ^{43}Ca NMR parameters.

are (i) to systematically evaluate the accuracy of QC ^{43}Ca NMR calculations and (ii) to further examine the influence of the Ca−O distance and coordination on the ^{43}Ca NMR interaction parameters (the isotropic chemical shift (δ_{iso}) and the quadrupole coupling constant (χ_q)) in organic Ca−O systems.

Computational Methods

Ab initio QC calculations of the NMR parameters were performed on a SunFire 25000 system (with 72 x dual-core UltraSPARC-IV+1.5 GHZ processors and 576 GB of memory) by using the Gaussian03 package.[21] Simplified molecular clusters were used for all Ca−O compounds to reduce computational time. To ensure the consistency of the calculations for all clusters, a general three-step protocol for constructing the cluster models was adopted. The steps were (i) experimental X-ray structures used directly without any geometry optimization; (ii) each cluster consisted of all molecules (ligands and water) in the first-coordination sphere around the target Ca^{2+}; and (iii) the hydrogen atoms are added in the idealized positions for molecules without crystallographic hydrogen positions. The bond distances for C−H, N−H, and O−H are taken as 0.96, 0.90, and 1.0 Å, respectively. Figure 1 shows the molecular cluster, used in the QC calculations, of the target Ca^{2+} site in Ca(formate)$_2$ and in Ca(L-glutamate)$_2$•4H$_2$O as examples. The same protocol was also applied to the Ca^{2+} binding sites in Ca-bound proteins (calmodulin, parvalbumin,and α-lactalbumin).

The total number of atoms in the molecular clusters varies from 20 to 180, depending on the local environment around the target Ca^{2+} ion. The computational time varied from 5 min to more than 10 h. We also found that inclusion of molecules from the second-coordination sphere usually produces only very small differences in the calculated NMR results. For example, when using HF/6-31G* to calculate the ^{43}Ca NMR data of the Ca^{2+} site in Ca(formate)$_2$, the difference was 0.3 ppm and 0.2 MHz for δ_{iso}(^{43}Ca) and χ_q(^{43}Ca), respectively, between a large cluster of 80 atoms (first- and second-coordination molecules) and a small cluster of 25 atoms (first-coordination molecules alone).

Ab initio QC calculations were carried out by using HF and B3LYP methods with various basis sets of STO-6G, 6-31G*, 6-311++G**, and cc-pVTZ on the Ca atom and 6-31G* on other atoms (C, N, O, and H). 6-311++G** was also tested on the O atoms in Ca(formate)$_2$, and little difference was found in ^{43}Ca NMR parameters calculated compared to those obtained with 6-31G*. The principal components of the ^{43}Ca electric-field-gradient (EFG) tensor components are defined as $|q_{zz}| > |q_{yy}| > |q_{xx}|$ where $q_{zz} + q_{yy} + q_{xx} = 0$.[22] The principal absolute chemical-shielding (CS) tensor components were computed by the ab initio self-consistent field coupled with the gauge-invariant-atomic-orbital method, where $\sigma_{iso} = (\sigma_{11} + \sigma_{22} + \sigma_{33})/3$ with $\sigma_{33} > \sigma_{22} > \sigma_{11}$. To make direct comparison between the calculated ^{43}Ca NMR parameters (χ_q, η_q, and δ_{iso}) and the experimental results, the following equations were used,

$$\chi_q[\mathrm{MHz}] = e^2 q_{zz} Q/h = -243.96 Q[\mathrm{barn}] q_{zz}[\mathrm{au}] \quad (1)$$

$$\eta_q = (q_{xx} - q_{yy})/q_{zz} \quad (2)$$

$$\delta_{iso}[\mathrm{ppm}] = \sigma_{ref} - \sigma_{iso} \quad (3)$$

where Q is the nuclear quadrupole moment, and a value of −0.048 barns was used for ^{43}Ca.[23] σ_{ref} is an absolute shielding constant for the primary chemical-shift reference sample.

^{43}Ca NMR Experimental Methods

Calcium(L-glutamate)$_2$•4H$_2$O was purchased from Aldrich (U.K.) and used directly for NMR experiments. ^{43}Ca NMR spectra were recorded on Chemagnetics Infinity 360 and 600 Bruker Advance II$^+$ spectrometers at 8.45 and 14.1 T, respectively. At these two fields, the operating frequencies for ^{43}Ca were 24.32 and 40.36 MHz, respectively. A large-volume 9.5 mm T3 Varian MAS probe was used with sample spinning at 3500−4000 Hz. For spectra obtained at 8.45 T, approximately 200 000 transients were averaged with a delay of 0.5 s. At 14.1 T, approximately 65 000 transients were collected with a delay of 1.0 s. To test the effects of concentration on the shift referencing, a saturated aqueous $CaCl_2$ solution was successively diluted, and various specific concentrations were also made up. A shift difference of 8.0 ppm was observed between the saturated solution and infinite dilution, with no change of shift observed for further dilution below 2.0 M. Hence, spectra were referenced to 1.0 M $CaCl_2$ solution at 0 ppm, which also effectively corresponds to an infinite dilute $CaCl_2$ solution.[24]

Results and Discussion

As mentioned earlier, one difficulty in carrying out ^{43}Ca NMR is the low natural abundance for the NMR-active ^{43}Ca isotope. The two common approaches to overcome such a problem are ^{43}Ca-enrichment and increasing the sample volume. We have chosen to use a large volume 9.5 mm diameter rotor for the organic solids. It should be pointed out that the ^{43}Ca-enrichment approach would be the more appropriate choice for the limited sample sizes that are likely to be available for biological materials. Figure 2 shows solid-state ^{43}Ca magic angle spinning (MAS) spectra of Ca(L-glutamate)$_2$•4H$_2$O at 8.45 and 14.1 T. Because of the lack of well-defined features (i.e., discontinuities) in both spectra, the δ_{iso}(^{43}Ca) and quadrupolar parameter P_q ($P_q = \chi_q(1 + \eta_q^2/3)^{1/2}$) were determined by using the field dependence of the center of gravity of the resonance[25] (δ_{cg}, $\delta_{cg} = \delta_{iso} - (1/392)(P_q^2/\upsilon_o^2) \times 10^6$) and found to be approximately $\delta_{iso} = 6.6 \pm 1.0$ ppm and $P_q = 0.8 \pm 0.5$ MHz.

The reliability of ab initio QC ^{43}Ca NMR calculations is investigated in this study, especially for Ca−O environments in organic compounds where solid-state ^{43}Ca NMR data were acquired previously.[16] The QC results are summarized in Table 1. Figure 3 shows good agreement, with a slope of 0.97, between the solid-state NMR results for δ_{iso}(^{43}Ca) and the QC values of σ_{iso}(^{43}Ca) obtained by using HF/6-31G*. The σ_{ref} can be extracted from the y-intercept value and is found to be 1271 ppm with a standard error of 5 ppm. The σ_{ref}(^{43}Ca) value corresponds to a fully hydrated Ca^{2+} ion in an infinite dilute $CaCl_2$ solution. It should be noted that the quality of the agreement could probably be further improved by readjusting (or optimizing) the hydrogen-atom positions in the cluster models. It is interesting to find that HF/6-31G* has also been

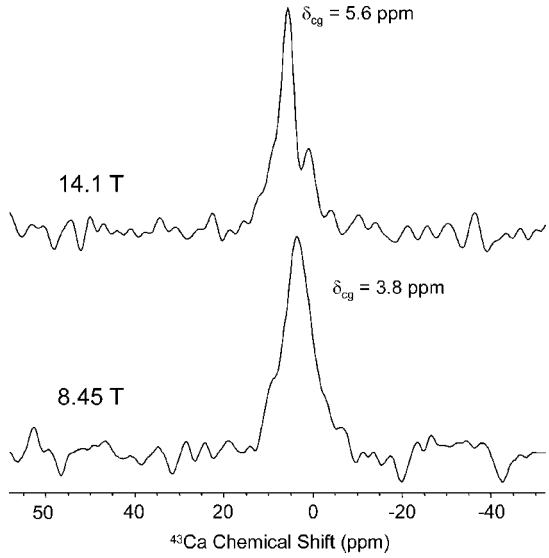

Figure 2. ^{43}Ca MAS spectra of Ca(L-glutamate)$_2$·4H$_2$O at 8.45 and 14.1 T.

used for other light atoms (^{23}Na[17] and ^{25}Mg[18]) and resulted in reliable NMR parameters when compared with the corresponding experimental results. Although we have shown that HF/6-31G* can produce reliable values of $\delta_{iso}(^{43}Ca)$ for Ca−O compounds, it is important to compare the results with those obtained by using other available basis functions. The results are summarized in Table 2. Figure 4 shows the variation of the calculated results of four different basis sets (STO-6G, 6-31G*, 6-311++G**, and cc-pVTZ) with HF (black line) and B3LYP (red line) functional methods. In general, HF/6-31G*, HF/6-311++G**, HF/cc-pVTZ, and B3LYP/cc-pVTZ give a slope close to 1.0; however, HF/6-31G* yields the most accurate results with the smallest standard error, ±5.4 ppm. On the other hand, HF and B3LYP methods with a small basis function, STO-6G, produce results with large standard errors, ±30 ppm. In addition, from the somewhat limited available experimental quadrupolar parameters, HF/6-31G* also produces the closest agreement with the experimental ^{43}Ca χ_q and η_q data. Because of the good agreement with the experimental data from HF/6-31G* calculations, the discussion below is therefore based on the results computed from this basis.

To extend the application of QC ^{43}Ca NMR calculations, the possibility of predicting the ^{43}Ca NMR parameters for high-affinity Ca-binding sites in Ca-bound proteins (EF-hand and α-lactalbumin) were explored. The EF-hand protein is defined by its helix−loop−helix secondary structure, where the carboxylic functional groups in the loop region play key roles for selectively binding to the Ca^{2+} ions. Calmodulin (PDB:1EXR) and parvalbumin (PDB:4CPV) are well-studied EF-hand proteins that consist of high-affinity Ca-binding sites (four and two Ca^{2+} sites, respectively).[26] As shown in Figures 5 and 6, all Ca^{2+} ions are coordinated to seven oxygen atoms forming a bipentagon coordination geometry. These Ca^{2+} ions are tightly bound to their corresponding sites in aqueous solution with association constants (K_a) in the order of 10^7-10^9 M^{-1}. Andersson et al.[4] have successfully probed these strongly bound Ca^{2+} ions by using solution ^{43}Ca NMR and reported $\delta_{iso}(^{43}Ca)$ to be 10 ± 1 ppm relative to a Ca^{2+} signal from 1 M $CaCl_2$ solution. The QC $\delta_{iso}(^{43}Ca)$ values from the cluster models shown in Figures 5 and 6 are slightly overestimated but are in relatively good agreement with the experimental values, being in the range of 11−18 ppm. The variation in QC $\delta_{iso}(^{43}Ca)$ can be attributed to the range in Ca−O distances (2.38−2.43 Å) found in the different Ca^{2+} sites.

A different class of calcium binding protein, α-lactalbumin (PDB:1ALC), is also investigated here. The protein is composed of two helices linked by a tight loop. The lone Ca^{2+} site (with $K_a \approx 2 \times 10^7$ M^{-1})[5] is found to have a coordination environment very similar to those in calmodulin and parvalbumin. It is formed by oxygens from four aspartates and one lysine residues, and the distorted bipentagon coordination sphere is completed by two water molecules with a mean Ca−O distance of 2.36 Å. The QC value 3.5 ± 5 ppm from the cluster model is found to be slightly too large compared to the observed shift −5.8 ppm.[5] On the basis of the comparison of the three different Ca-bound proteins, the $\delta_{iso}(^{43}Ca)$ shift difference between the QC values from crystal clusters and the observed values in solution may be attributed to the fact that the Ca^{2+} coordination environments are slightly different between the solid phase and the solution phase. For example, the upfield shift observed from the solution phase for α-lactalbumin may suggest that the Ca−O coordination sphere is larger (or expanded) in the aqueous phase compared to that in the crystal phase.

Although all Ca^{2+} sites in these proteins consist of a bipentagon coordination geometry, the bipentagon coordination of Ca3 in calmodulin and of Ca1 and Ca2 in paravalbumin are more distorted compared to that of other Ca^{2+} sites, including α-lactalbumin, where the angle between the axial oxygen atoms (O···Ca···O) is much less than 180°. The QC $\chi_q(^{43}Ca)$ values for these latter sites are consistently larger (2−3 MHz) compared with those in the less distorted sites (1−2 MHz). The QC results suggest that $\chi_q(^{43}Ca)$ is sensitive to the local Ca−O coordination geometry environment in metalloproteins and illustrate the potential of characterizing the Ca−O coordination environments in Ca-bound proteins by solid-state ^{43}Ca NMR spectroscopy.

In previous studies, a general correlation between the experimental value of $\delta_{iso}(^{43}Ca)$ and the mean Ca−O distance was found,[7,8,16] where δ_{iso} increases as the mean Ca−O distance decreases. With the reliable NMR results from HF/6-31G*, a better empirical correlation can now be established between the Ca−O environment and ^{43}Ca NMR parameters by also considering other different functional groups of Ca−O organic compounds, such as ether oxygens (C−O−C) from crown-ethers, hydroxyl oxygens (C−OH) from carbohydrates, and carbonyl oxygens (C═O) from acetonate. This provides a larger range of Ca−O distances and Ca−O coordination numbers (CN). Figure 7a shows that there is a clear linear dependence with little ambiguity for all different Ca−O environments: the QC $\delta_{iso}(^{43}Ca)$ decreases as the Ca−O bond weakens (δ_{iso} [ppm] $= -154(d_{Ca-O}) + 380$). This observed trend can be qualitatively rationalized by Ramsey's chemical-shielding theory.[27] As the Ca−O interaction weakens, less donor electron density is contributed to the bonding interaction, reducing the paramagnetic contribution, producing a decrease in chemical shift. Because the mean metal−oxygen distance increases as the metal−oxygen CN increases, an approximately linear decrease in ^{43}Ca shift is also observed for larger Ca−O CN (Figure 7b). Lower-quality crystal structures will decrease to accuracy of the QC values, but on the basis of variations of the NMR parameters observed, indications of the local calcium environment would still be forthcoming.

The QC $\chi_q(^{43}Ca)$ values are typically 1−3 MHz for Ca−O in organic compounds with CN of 6−10, although for CN of 4 (tetrahedral) and 5 (square pyramidal), χ_q is found to be much larger, 5−7 MHz. However, the majority of Ca^{2+} sites found in biomolecules have CN of 6, 7, or 8 with the neighboring

TABLE 1: Summary of HF/6-31G* ^{43}Ca NMR Parameters for Ca−O Compounds and Ca Hydrated Models

Ca compound	CN	d(Ca−O)/Å	cal. σ_{iso}/ppm	cal. δ_{iso}[a]/ ±5 ppm	cal. χ_q/MHz	cal. η_q	expt. δ_{iso}[b]/ppm	expt. χ_q/MHz	CSD[c] refcodes
With NMR Data									
Ca(L-glutamate)$_2$·4H$_2$O	6	2.317	1258.6	17.2	0.78	0.96	6.6[d]	∼0.8[d]	LGLUCA
Ca(formate)$_2$	7	2.421	1263.3	7.9	1.85	0.55	13.5[e]	1.24	CAFORM04
Ca(D-gluconate)$_2$·5H$_2$O	7	2.382	1260.5	10.7	1.42	0.80	16.5[e]	1.25	ZUZVOF
Ca(acetate)$_2$·H$_2$O Ca1	7	2.404	1250.4	20.7	1.67	0.75	15.8[e]		CEJLIM
Ca2	8	2.502	1275.6	−4.4	1.74	0.60	−3.8[e]		CEJLIM
Ca(ascorbate)$_2$·2H$_2$O	8	2.462	1280.5	−9.4	0.58	0.50	−6.3[e]		CAASCO01
Ca(oxalate)·3H$_2$O	8	2.453	1264.3	6.8	2.95	0.37	3.8[e]	1.55	ZZZUOQ01
Ca calmodulin[f] Ca1	7	2.40	1254.8	16.3	1.15	0.44	10	1.15	PDB:1EXR
Ca2	7	2.38	1252.9	18.2	1.91	0.81	10	1.15	PDB:1EXR
Ca3	7	2.43	1260.1	11.0	2.26	0.36	10	1.15	PDB:1EXR
Ca4	7	2.39	1257.0	14.2	1.10	0.76	10	1.15	PDB:1EXR
Ca parvalbumin[f] Ca1	7	2.41	1257.9	13.3	3.16	0.73	10	1.3	PDB:4CPV
Ca2	7	2.42	1255.9	15.2	3.10	0.64	10	1.3	PDB:4CPV
Ca α-lactalbumin[f]	7	2.36	1267.6	3.5	1.46	0.49	−5.8	0.7	PDB:1ALC
Without NMR data[g]									
diphenylethanolate **1**	4	2.268	1237.2	33.9	6.35	0.13			YALJEA
methylphenolate **2**	5	2.325	1265.3	5.9	5.66	0.51			SICGUG
bis(acetylacetonate) **3**	6	2.338	1242.8	28.4	2.27	0.43			BOLTIF
hexakis(urea) **4**	6	2.325	1249.5	21.6	0.46	0.19			CABRUR10
monosaccharide **5**	8	2.478	1259.5	11.6	1.42	0.26			ALARCA
monosaccharide **6**	9	2.552	1276.9	−5.8	1.94	0.10			CANAGL01
disaccharide **7**	9	2.507	1275.5	−4.4	2.40	0.28			ALTRCA
crown-ether **8**	10	2.603	1303.2	−32.1	1.94	0.31			DUJJOH
crown-ether **9**	10	2.573	1297.8	−26.7	2.14	0.44			FEWCAP
Hydration Model[h]									
6-water	6	2.442	1263.9	7.2	0.00	0.00			
7-water	7	2.488	1274.3	−3.2	0.63	0.35			
8-water	8	2.529	1285.3	−14.2	0.91	0.00			
9-water	9	2.591	1303.3	−32.3	0.89	0.89			

[a] δ_{iso} [ppm] = 1271.1 − σ_{iso}. [b] Experimental errors for δ_{iso} ±0.5−1.0 ppm and χ_q ±0.1 MHz.[4,5,16] [c] From Cambridge Structural Database. [d] This work. [e] Re-referenced the original solid δ_{iso}[16] to an appropriate primary reference, 1.0 M CaCl(aq),[24] by adding 8 ppm. [f] From solution NMR data;[4,5] X-ray structural data were obtained from PDB. [g] **1** bis((m∼2∼-2-(4-chlorophenyl)-1,1-diphenyl-ethanolato-O,O)-(2-(4-chlorophenyl)-1,1-diphenyl-ethanolato)-tetrahydrofuran-calcium) toluene solvate; **2** bis(2,6-di-t-butyl-4-methylphenoxy)-tris(tetrahydrofuran-O)-calcium tetrahydrofuran solvate; **3** cis-bis(acetylacetonato)-diaqua-calcium monohydrate; **4** hexakis(urea) calcium bromide; **5** α-L-arabinose calcium chloride tetrahydrate; **6** calcium sodium α-D-galacturonate hexahydrate; **7** α-D-allopyranosyl-α-D-allopyranoside calcium chloride pentahydrate; **8** bis(nitrato-O,O')-(2R,3R,11R,12R)-2,3,11,12-tetramethyl-1,4,7,10,13,16-hexaoxa cyclo-octadecane-O,O,O,O,O,O-calcium; **9** (18-crown-6)-dinitrato-calcium. [h] From ab initio molecular orbital calculations:[29] models were structural-optimized with RHF/HUZSP*(p)//RHF/HUZSP*(p) level.

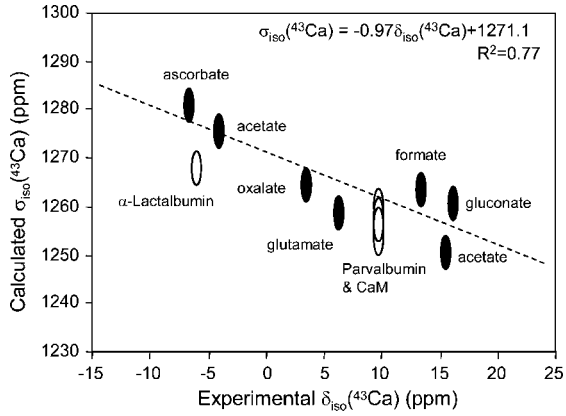

Figure 3. Correlation between the HF/6-31G* σ_{iso}(^{43}Ca) values and the NMR δ_{iso}(^{43}Ca) values of organic compounds. The least-squares linear fit is indicated by the dashed line. For comparison, the solution ^{43}Ca NMR shifts of Ca-bound proteins are also included and represented by open circles, whereas the Ca−O organic compounds are represented by solid circle. The QC (±5.4 ppm) and experimental (±1 ppm) errors are indicated by the circle size.

donor oxygen atoms, and therefore, a relatively small χ_q (1−3 MHz) should be expected for the Ca^{2+} binding sites in biological systems. The small quadrupolar contribution is advantageous for ^{43}Ca NMR studies, because the residual second-order quadrupolar line broadening contribution ($\Delta\upsilon_q^{(2)}$) under MAS would then be small (for I = 7/2, $\Delta\upsilon_q^{(2)}$= (5/147)[$\chi_q(6 + \eta_q)$]^2/(224υ_o)).[22] For ^{43}Ca in a site with χ_q(^{43}Ca) of 1.0−3.0 MHz and η_q of 1.0, the quadrupolar broadening is expected to be only 180−1700 Hz at 14.1 T. On the basis of the range found for δ_{iso}(^{43}Ca) in this study, ∼100 ppm, a typical ^{43}Ca MAS signal (e.g., χ_q = 2.0 MHz) at 14.1 T will have a typical width

TABLE 2: Summary of the Least-Square Linear Fit between the Solid-State NMR δ_{iso}(^{43}Ca) Results and the QC σ_{iso}(^{43}Ca) Values with Different Levels of Calculation

method/basis set	σ_{ref}[a] /ppm	slope of the least-square fit	standard error[b] /ppm
HF/STO-6G	1396.9	1.93	30.8
HF/6-31G*	1271.1	0.97	5.4
HF/6-311++G**	1260.1	0.83	11.6
HF/cc-pVTZ	1279.2	0.92	14.8
B3LYP/STO-6G	1376.3	1.97	30.6
B3LYP/6-31G*	1232.5	1.22	9.4
B3LYP/6-311++G**	1194.7	1.37	11.2
B3LYP/cc-pVTZ	1219.8	0.88	13.0

[a] Reference for 1.0 M $CaCl_2$ solution. [b] Standard error for predicting the calculated shielding constant $S_{y,x} = \{(1/n(n-2))[n\Sigma y^2 - (\Sigma y)^2 - [n\Sigma xy - (\Sigma x)(\Sigma y)]^2/(n\Sigma x^2 - (\Sigma x)^2)]\}^{1/2}$.

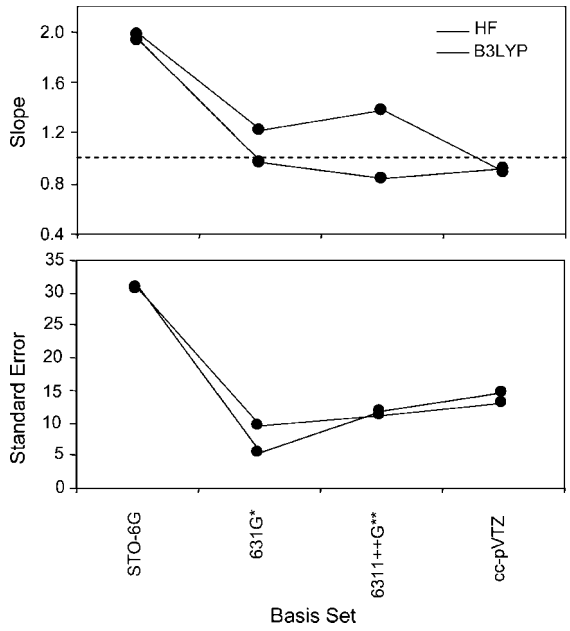

Figure 4. Comparison of $\delta_{iso}(^{43}Ca)$ between different functional methods and basis sets: HF (bottom line); B3LYP (top line). The comparison was done for all values of organic compounds determined from solid-state NMR,[15] see Table 1.

of ~18% of the entire chemical-shift span, which reduces to ~8% at 21.15 T. The relatively small line broadening compared with the potential chemical-shift differences is expected to improve the chances of spectrally distinguishing Ca^{2+} sites with different environments without undertaking difficult and time-consuming experiments such as 2D multiple-quantum MAS. Indeed, recently, two ^{43}Ca NMR signals were resolved for the distinct Ca^{2+} sites in $Ca(acetate)_2 \cdot H_2O$ at 14.1 T[16] and in Ca hydroxyapatite at 18.8 T.[14] An advantage of the small $\chi_q(^{43}Ca)$ is that it is not necessary to use fast MAS to produce the narrowing effect, so that large-volume sample rotors can be used to characterize Ca^{2+} ions at natural isotopic abundance when sufficient sample is available. Furthermore, because ^{43}Ca is a quadrupole nucleus ($I = 7/2$), robust signal enhancement techniques[28] which manipulate the nuclear spin-population between the satellite and central transitions have the potential to enhance the signal up to 3-fold for MAS spectra.

Similarly to ^{23}Na, ^{39}K, and ^{25}Mg, the ^{43}Ca chemical-shift anisotropy (CSA) is expected to be small for the light Ca^{2+} ion. The QC ^{43}Ca span ($\Omega = \sigma_{33} - \sigma_{11}$) for the organic solids is found to be 20−40 ppm. This can be compared to the static second-order quadrupolar line width for the range of typical $^{43}Ca(\chi_q)$ values from such environments which will be approximately 50−400 ppm. Hence, the CSA contribution will be small but large enough to complicate the spectral simulation, especially if these two tensors are not aligned, adding more unknown variables to the simulation of the static spectrum. For this reason, excellent signal-to-noise (S/N) would be necessary at several applied magnetic fields to ensure that accurate simulation is possible. Such measurements are very time-consuming, especially for ^{43}Ca. Moreover, the variation in span $\Omega(^{43}Ca)$ is small, suggesting that it is not a sensitive probe to characterize the local environment. Thus, $\delta_{iso}(^{43}Ca)$ which covers a large range, as demonstrated in the $\delta_{iso}(^{43}Ca)$ versus Ca−O correlation plot shown in Figure 7, is the more valuable NMR parameter to study the Ca^{2+} environment and should be the principal NMR parameter for characterizing the Ca^{2+} binding sites.

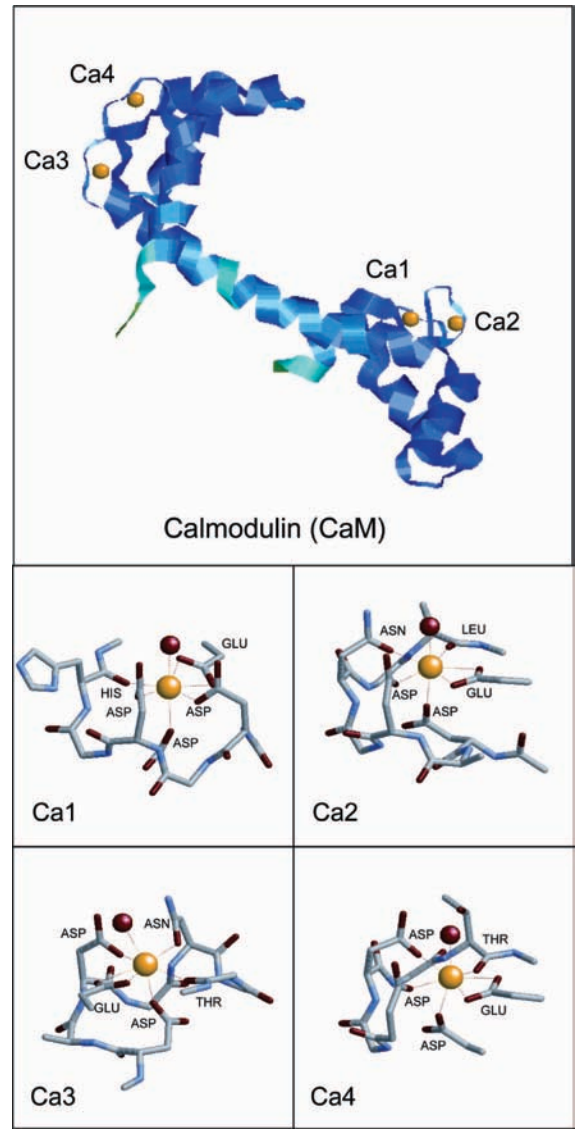

Figure 5. Crystal structure of calmodulin (PDB:1EXR). The four molecular clusters used for QC ^{43}Ca NMR calculation are depicted without hydrogen atoms.

There are ongoing investigations on the solvation structure of Ca^{2+} in aqueous solution by using computational and experimental approaches. However, studies to date have yielded a wide range of different results and conclusions. To our knowledge, there are still no definite conclusions regarding the exact number of water molecules in the Ca^{2+} first hydration shell. For example, Katz et al.[29] used DFT molecular orbital calculations to study a variety of hydrated Ca^{2+} clusters and found that clusters with 6, 7, and 8 water molecules (or CN of 6, 7, and 8) in the first hydration shell of Ca^{2+} are energetically similar. Pinna and co-workers[30] used X-ray diffraction and demonstrated that in $CaCl_2$ aqueous solutions, the ions can be treated as independently hydrated and found that the Ca^{2+} ions in 1.0, 2.0, and 4.0 M $CaCl_2$ solutions have CN 6. In a recent study, Dellago and co-workers,[31] by using Car−Parrinello molecular dynamics simulations to characterize the properties of the Ca^{2+} hydration shell, concluded that the CN fluctuates between 7 and 8, with a CN of 7 being the preferred hydration shell around Ca^{2+}. Fulton and co-workers[32] have carried out EXAFS and XANES experiments on both concentrated and dilute Ca^{2+} aqueous solution and found that the water coordination in both consist of a mean CN of ~7.

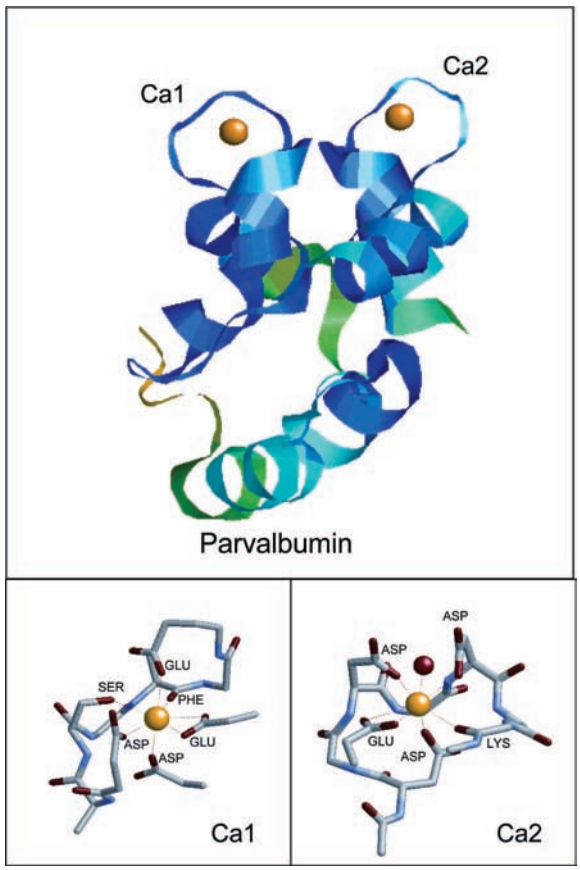

Figure 6. Crystal structure of parvalbumin (PDB:4CPV). The two molecular clusters used for QC ^{43}Ca NMR calculation are depicted without hydrogen atoms.

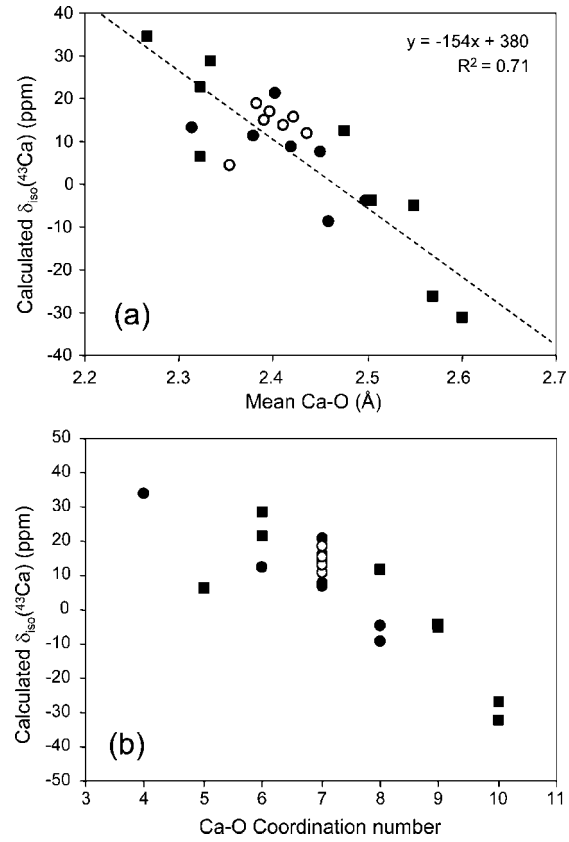

Figure 7. Correlations of the HF/6-31G* $\delta_{iso}(^{43}Ca)$ shifts: (a) mean Ca−O distance and (b) Ca−O coordination number. Solid circles: Ca−O compounds with solid-state NMR data. Solid squares: Ca−O compounds without experimental NMR data. Open circles: Ca-bound proteins.

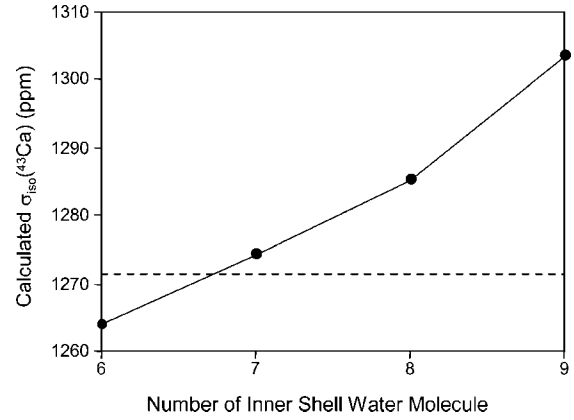

Figure 8. Correlation between HF/6-31G* shielding constant and number of inner-shell water molecules. The dashed line indicates the experimental shielding constant found for the dilute $CaCl_2$ solution.

In this study, the $\sigma(^{43}Ca)$ value for the infinite dilute $CaCl_2$ solution was determined to be 1271 ppm. Because the HF/6-31G* basis function gives reliable $\sigma(^{43}Ca)$, the effect of hydration environments around the Ca^{2+} in dilute $CaCl_2$ solution can be examined by calculating $\sigma(^{43}Ca)$ for the possible Ca^{2+} hydration models. The QC $\sigma(^{43}Ca)$ values for a series of Ca^{2+} hydration clusters (CN = 6, 7, 8, and 9), predicted from ab initio molecular orbital calculations at RHF/HUZSP*(p)//RHF/HUZSP*(p) level,[29] are summarized in Table 1. It is important to note that all calculations were done with the same basis function, HF/6-31G*, as the one used in the determination of σ_{ref}. This should give minimum systematic errors from the functional method itself. Figure 8 shows a decrease in chemical shift as the CN increases, which is consistent with the Ca−O distance and CN dependences observed for Ca−O compounds above. More importantly, σ_{ref} is comparable to the value computed for CN of 7, where seven water molecules are coordinated to the Ca^{2+} ion, suggesting that seven water molecules are preferred in the first hydration shell of the Ca^{2+} ions in a dilute $CaCl_2$ solution. This finding is consistent with a recent molecular dynamics simulation study by Dellago and co-workers[31] and also with the EXAFS and XANES measurements by Fulton and co-workers.[32] To establish an unambiguous conclusion for the hydration environments around Ca^{2+}, one can improve the quality of the data analysis by (1) collecting more experimental NMR shifts on Ca−O organic compounds to improve the reliability of σ_{ref} for $CaCl_2$ solution from a comparison between the experimental $\delta(^{43}Ca)$ and the corresponding $\sigma(^{43}Ca)$ (see Figure 3); (2) performing molecular dynamics simulations for $CaCl_2$ solution to give more reliable Ca^{2+} hydration cluster models for the NMR calculation; (3) considering the second hydration shell (water or/and counterion) around the Ca^{2+} site for the QC NMR. These three suggestions should provide more definite information regarding Ca^{2+} hydration environments; however, this is beyond the scope of the present study. Nonetheless, the above-described methodology, by combining ^{43}Ca NMR experimental data and QC calculations, demonstrates the potential for studying the Ca−O environment in Ca-containing complexes, such as organic molecules and metalloproteins.

As a final note, it should be emphasized that the observed ^{43}Ca MAS NMR signal of Ca(L-glutamate)$_2$•4H$_2$O (MW = 404 g/mol) with a total mass of ~500 mg at a low magnetic field,

8.45 T, is very encouraging. On the basis of the ^{43}Ca MAS NMR signal, a similar S/N can be acquired for 10 mg of a 60% ^{43}Ca-enriched sample with a molecular weight of ~7.5 kDa at 21.1 T. This strongly suggests that solid-state ^{43}Ca NMR spectroscopy could be used for probing the Ca^{2+} ions in small metalloproteins with limited sample available. In addition, according to the QC ^{43}Ca NMR parameters, the quadrupolar broadening effects for Ca−O environments (CN = 6, 7, and 8) most commonly found in biomolecules are expected to be relatively small, 490 Hz (8 ppm) at 21.15 T for $\chi_q(^{43}Ca) = 2.0$ MHz; therefore, the loss of S/N from the quadrupolar broadening will be small.

Conclusion

The present study has shown that ab initio QC calculations of the ^{43}Ca NMR parameters can produce reliable results for Ca−O environments in organic molecules and in Ca-bound proteins by using a small basis set, HF/6-31G*. It has also been shown that $\delta_{iso}(^{43}Ca)$ is a sensitive probe of the Ca−O bonding environment, with $\delta_{iso}(^{43}Ca)$ varying by about 80 ppm and being correlated with the mean Ca−O distance. Such a correlation may be useful in identifying the Ca^{2+} ion binding environments for unknown systems based on solid-state ^{43}Ca NMR results alone. In addition, the QC results also suggest that $\chi_q(^{43}Ca)$ is sensitive to the Ca−O geometry of the coordination environment in Ca-bound proteins. A potential application has been outlined to investigate the Ca−O environment by combining the use of QC and solid-state ^{43}Ca NMR. With current advances in NMR instrumentation and the small $\chi_q(^{43}Ca)$ values anticipated for Ca−O environments in most organic compounds, there is clearly considerable potential for the use of solid-state ^{43}Ca NMR spectroscopy in the characterization of Ca^{2+} binding sites in biomolecules.

Acknowledgment. This work was supported by BBSRC through Grant BB/C000471/1. EPSRC and the University of Warwick are thanked for the partial funding of NMR instrumentation at Warwick. Quantum chemical calculations were performed at HPCVL at Queen's University. A.W. would like to thank the NSERC (Canada) for a Post-Doctoral Fellowship Award, D.L. thanks the 6th European Community Framework Program for a Marie Curie Intra-European Fellowship, and R.D. thanks the Leverhulme Trust. We would also like to thank the referees for their valuable comments.

References and Notes

(1) Dorozhkin, S. V. *J. Mater. Sci.* **2007**, *42*, 1061.
(2) (a) Aramini, J. M.; Vogel, H. J. *Biochem. Cell Biol.* **1998**, *76*, 210, and references therein. (b) Drakenberg, T.; Johansson, C.; Forsén, S. *Method Mol. Biol.* **1997**, *60*, 299, and references therein.
(3) Kwan, I. C. M.; Wong, A.; She, Yi.-M.; Smith, M. E.; Wu, G. *Chem. Commun.* **2008**, 682.
(4) Andersson, T.; Drakenberg, T.; Forsén, S.; Thulin, E.; Swärd, M. *J. Am. Chem. Soc.* **1982**, *104*, 576.
(5) Aramini, J. M.; Drakenberg, T.; Hiraoki, T.; Ke, Y.; Nitta, K.; Vogel, H. J. *Biochemistry* **1992**, *31*, 6761.
(6) Smith, M. E. *Annu. Rep. NMR Spectrosc.* **2000**, *43*, 121.
(7) Dupree, R.; Howes, A. P.; Kohn, S. C. *Chem. Phys. Lett.* **1997**, *276*, 399.
(8) Lin, Z.; Smith, M. E.; Sowrey, F. E.; Newport, R. J. *Phys. Rev. B* **2004**, *69*, 224107.
(9) Shimoda, K.; Tobu, Y.; Kanehashi, K.; Nemoto, T.; Saito, K. *Chem. Lett.* **2005**, *34*, 1588.
(10) Shimoda, K.; Tobu, Y.; Kanehashi, K.; Nemoto, T.; Saito, K. *Solid State Nucl. Magn. Reson.* **2006**, *30*, 198.
(11) Shimoda, K.; Tobu, Y.; Shimoikeda, Y.; Nemoto, T.; Saito, K. *J. Magn. Reson.* **2007**, *186*, 114.
(12) Angeli, F.; Gaillard, M.; Jollivet, P.; Charpentier, T. *Chem. Phys. Lett.* **2007**, *440*, 324.
(13) MacKenzie, K. J. D.; Smith, M. E.; Wong, A. *J. Mater. Chem.* **2007**, *17*, 5090.
(14) (a) Laurencin, D.; Wong, A.; Hanna, J. V.; Dupree, R.; Smith, M. E. *J. Am. Chem. Soc.* **2008**, *130*, 2412. (b) Laurencin, D.; Wong, A.; Dupree, R.; Smith, M. E. *Magn. Reson. Chem.* **2008**, *46*, 347.
(15) Bryce, D. L.; Bultz, E. B.; Aebi, D. *J. Am. Chem. Soc.* **2008**, *130*, 9282.
(16) Wong, A.; Howes, A. P.; Dupree, R.; Smith, M. E. *Chem. Phys. Lett.* **2006**, *427*, 201.
(17) (a) Wong, A.; Wu, G. *J. Phys. Chem. A* **2003**, *107*, 579. (b) Tossell, J. A. *J. Phys. Chem. B* **2001**, *105*, 11060.
(18) Wong, A.; Ida, R.; Mo, X.; Gan, Z.; Poh, J.; Wu, G. *J. Phys. Chem. A* **2006**, *110*, 10084.
(19) (a) Lipton, A. S.; Ellis, P. D. *J. Am. Chem. Soc.* **2007**, *129*, 9192. (b) Zhang, Y.; Mukherjee, S.; Oldfield, E. *J. Am. Chem. Soc.* **2005**, *127*, 2370. (c) Ida, R.; Wu, G. *J. Phys. Chem. A* **2002**, *106*, 11234.
(20) (a) Wong, A.; Thurgood, G.; Dupree, R.; Smith, M. E. *Chem. Phys.* **2007**, *337*, 144. (b) Kwan, I. C. M.; Mo, X.; Wu, G. *J. Am. Chem. Soc.* **2007**, *129*, 2398. (c) Chekmenev, E. Y.; Waddel, K. W.; Hu, J.; Gan, Z.; Wittebort, R. J.; Cross, T. A. *J. Am. Chem. Soc.* **2006**, *128*, 9849.
(21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(22) (a) MacKenzie, K. J. D.; Smith. M. E. Multinuclear Solid-State NMR of Inorganic Materials. *Pergamon Materials Series*; Pergamon/Elsevier: Oxford, 2002; Vol. 6. (b) Smith, M. E.; van Eck, E. R. H. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 159.
(23) Pyykkö, P. *Mol. Phys.* **2001**, *99*, 1617.
(24) Farmer, R. M.; Popov, A. I. *Inorg. Nucl. Chem. Lett.* **1981**, *17*, 51.
(25) Kohn, S. C.; Smith, M. E.; Dirken, P. J.; van Eck, E. R. H.; Kentgens, A. P. M.; Dupree, R. *Geochim. Cosmochim. Acta* **1998**, *62*, 79.
(26) Gifford, J. L.; Walsh, M. P.; Vogel, H. J. *Biochem. J.* **2007**, *405*, 199.
(27) (a) Ramsey, N. F. *Phys. Rev.* **1952**, *86*, 243. (b) Ramsey, N. F. *Phys. Rev.* **1951**, *83*, 540. (c) Ramsey, N. F. *Phys. Rev.* **1950**, *78*, 699.
(28) Siegel, R.; Nakashima, T. T.; Wasylishen, R. E. *Conc. Magn. Reson. A* **2005**, *26*, 47.
(29) Katz, A. K.; Glusker, J. P.; Beebe, S. A.; Bock, C. W. *J. Am. Chem. Soc.* **1996**, *118*, 5752.
(30) Licheri, G.; Piccaluga, G.; Pinna, G. *J. Chem. Phys.* **1976**, *64*, 2437.
(31) Naor, M. M.; Nostrand, K. V.; Dellago, C. *Chem. Phys. Lett.* **2003**, *369*, 159.
(32) Fulton, J. L.; Heald, S. M.; Badyal, Y. S.; Simonson, J. M. *J. Phys. Chem. A* **2003**, *107*, 4688.

JP801015Y