Density Functional Calculations of Proton Chemical Shifts in Model Peptides

Doree Sitkoff and David A. Case*

Contribution from the Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037

Received June 27, 1997. Revised Manuscript Received September 26, 1997[⊗]

Abstract: Density-functional chemical shielding calculations are reported for the alanine dipeptide with a variety of backbone torsion angles and for methane and *N*-methylacetamide complexes with rare gases, monatomic ions, water, and other amides. These fragment systems model electrostatic, nonbonded, and hydrogen bonding interactions in proteins and have been investigated at a variety of geometries. The results are compared to empirical formulas that relate intermolecular shielding effects to peptide group magnetic anisotropies, electrostatic polarization of the C−H and N−H bonds, magnetic contributions from C−C and C−H bonds, and close contact effects. Close contacts are found to deshield protons involved in close nonbonded contacts that typically occur in hydrogen bonds. "Lone pair" charges improve the model for electrostatic effects and are important for understanding the angular dependence of shifts for protons involved in hydrogen bonds. C−C and C−H bond anisotropy contributions help to explain the torsional dependence of amide proton shifts in alanine dipeptide. Good agreement is found between the empirical formulas and the quantum chemistry results, allowing a reassessment of empirical formulas that are used in the analysis of chemical shift dispersion in proteins.

1. Introduction

It has been known for some time that NMR chemical shift values reflect molecular structure. Local variations in angles and bond lengths can affect chemical shieldings through changes in electronic structure, and more distant groups in the environment can contribute to chemical shifts as well, by affecting the magnetic field or electron density at nuclei of interest via anisotropic magnetic susceptibility (often from aromatic rings), electrostatic, and close contact interactions. In most NMR studies of macromolecular structure, however, little knowledge has been gleaned from chemical shift information other than by using them as distinctive labels for nuclei under study.

This has begun to change in recent years, which have seen an increase in the use of chemical shift values themselves in deriving structural information on macromolecules.[1−10] Much of the progress is due to empirical analyses of experimentally measured shifts and their relationships to structure. In some cases, simple correlations between shifts and local conformation have been found, such as between experimental protein C^{α} and C^{β} shifts and backbone torsion angles.[11,12] More detailed empirical methods have been developed for calculating proton shifts, which are generally dominated by environmental effects rather than local structure; these approaches include equations based on classical models of ring current,[13] magnetic anisotropy,[14] and electrostatic[15] effects on shifts, parameterized to experimental shift data.[16−18] The empirical approaches generally afford only a limited understanding of conformational shift effects, however, because the multitude of contributions and conditions present in the experimental systems (solvent molecules, conformational variations, in addition to combinations of electrostatic, close contact, and magnetic anisotropy contributions) make it difficult to fully test the empirical shift models and develop reliable parameters.

Recent advances in quantum chemical studies[5,6,19−23] are beginning to yield significant contributions to chemical shift calculation methods by enabling us to strengthen our knowledge and understanding of the relationship between structure and chemical shifts. This has been true in recent years for the case of ^{13}C shifts, where quantum calculations on small molecules in the gas phase have been used to thoroughly trace the dependence of carbon peptide shieldings on ϕ and ψ backbone angles.[24,25] Another modern use of quantum shift calculations

[⊗] Abstract published in *Advance ACS Abstracts,* December 1, 1997.

(1) Ösapay, K.; Theriault, Y.; Wright, P. E.; Case, D. A. *J. Mol. Biol.* **1994**, *244*, 183−197.
(2) Szilgyi, L. *Prog. NMR Spectr.* **1995**, *27*, 325−443.
(3) Kuszewski, J.; Gronenborn, A. M.; Clore, G. M. *J. Magn Reson. Ser. B* **1995**, *107*, 293−297.
(4) Kuszewski, J.; Qin, J.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson. Ser. B* **1995**, *106*, 92−96.
(5) Le, H.; Pearson, J. G.; de Dios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1995**, *117*, 3800−3807.
(6) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217−225.
(7) Williamson, M. P.; Kikuchi, J.; Asakura, T. *J. Mol. Biol.* **1995**, *247*, 541−546.
(8) Celda, B.; Biamonti, C.; Arnau, M. J.; Tejero, R.; Montelione, G. T. *J. Biomol. NMR* **1995**, *5*, 161−172.
(9) Copie, V.; Battles, J. A.; Schwab, J. M.; Torchia, D. A. *J. Biomol. NMR* **1996**, *7*, 335−340.
(10) MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. *J. Biomol. NMR* **1996**, *7*, 256−260.
(11) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490−5492.
(12) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311−333.
(13) Haigh, C. W.; Mallion, R. B. *Prog. NMR Spectr.* **1980**, *13*, 303−344.
(14) McConnell, H. M. *J. Chem. Phys.* **1957**, *27*, 226−229.
(15) Buckingham, A. D. *Can. J. Chem.* **1960**, *38*, 300−307.
(16) Ösapay, K.; Case, D. A. *J. Am. Chem. Soc.* **1991**, *113*, 9436−9444.
(17) Williamson, M. P.; Asakura, T. *J. Magn. Reson. Ser. B* **1993**, *101*, 63−71.
(18) Asakura, T.; Taoka, K.; Demura, M.; Williamson, M. P. *J. Biomol. NMR* **1995**, *6*, 227−236.
(19) Chesnut, D. B.; Phung, C. G. In *Nuclear Magnetic Shieldings and Molecular Structure*; Tossell, J. A., Ed.; Kluwer Academic Publishers: Dordrecht, 1993; pp 221−241.
(20) Jiao, D.; Barfield, M.; Hruby, V. J. *J. Am. Chem. Soc.* **1993**, *115*, 10883−10887.
(21) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer, H. F., III *J. Am. Chem. Soc.* **1994**, *116*, 3967−3972.
(22) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer, H. F., III *J. Am. Chem. Soc.* **1995**, *117*, 2632−2637.
(23) Malkin, V. G.; Malkina, O. L.; Casida, M. E.; Salahub, D. R. *J. Am. Chem. Soc.* **1994**, *116*, 5898−5908.

S0002-7863(97)02143-4 CCC: $14.00

Table 1. Molecules for Which DFT and Empirical Shifts Were Calculated[a]

system	protons	δ_{emp} contributions				
		anis	*pol*	*ccch*	*cc*	*S1*[b]
methane-He	HC	•				
methane-Ar	HC	•				
methane-Ne	HC	•				
methane-F^-	HC	•				
methane-water	HC	•	•	•[c]		
methane-NMA	HC	•	•	•	•	•
NMA-He	HN	•				
NMA-Ar	HN	•				
NMA-Ne	HN	•				
NMA-F^-	HN	•	•			
NMA−water	HN	•	•	•[c]		
NMA−NMA	HN	•	•	•	•	•[c]
alanine dipeptide[d]	HN_1,Ha,HB3,HN_2	•	•	•	•	•[c]

[a] Molecules whose proton shifts were investigated are in all capital letters. See Figure 1 for molecular structures and atom names. See text for complete descriptions of molecular geometries. Empirical contributions are abbreviated as follows: *anis* = peptide magnetic anisotropy; *pol* = electrostatic polarization; *ccch* = C−C and C−H bond magnetic anisotropy; *cc* = close contact. [b] Protons having negligible or constant close contact interactions with nearby atoms. These DFT shifts were used to fit peptide anisotropy and electrostatic empirical shift parameters. The full set of protons in this table, denoted *S2*, were subsequently used to fit the close contact empirical shift parameters. [c] Subset of protons with negligible or constant close contact interactions. [d] Shifts from low energy alanine dipeptide structures having a calculated DFT shift of less than 7 ppm, excluding HN_1 shifts for $(\phi,\psi) = (-30,120)$, $(\phi,\psi) = (-30,150)$, $(\phi,\psi) = (-60,180)$ (see Methods section.) HB3 denotes the average of shifts from the three methyl protons on the CB carbon.

is in the development and parameterization of physical models of environmental shift effects; this can be done by studying small molecules in simple geometries and combinations, so that different contributions to shifts and their behavior with conformational change can be isolated and quantitated. Empirical models of ring current and electrostatic effects on proton shifts were recently studied in this manner.[26]

In this work, quantum chemical shift calculations are used to investigate the structural dependence of magnetic anisotropy, electrostatic and close contact empirical contributions to shifts in proteins, by examining shifts for small molecules in geometries that probe these interactions. The shifts are used to develop, test, and parameterize empirical models of these physical effects. The results yield an empirical approach that calculates gas phase proton shifts for alanine dipeptide with improved accuracy over previous empirical models, which were developed and parameterized to experimental shift data in proteins. In combination with recent similar work on ring current theory,[26] the new empirical model will afford a fresh approach to interpreting and predicting proton shifts in macromolecules.

2. Methods

2.1. Structures Examined. The model systems for which Density Functional Theory (DFT) quantum calculations were performed are given in Table 1 and Figure 1. All of the structures were built and minimized in monomer form in CHARMM using the CHARMM22 all-hydrogen force field. Alanine dipeptide conformers were minimized with a force of 1000.0 kcal/mol constraining the dihedral angles for a total of 1000 steps. The systems can be divided into three major groups: the first group has methane as a probe molecule, as in an earlier paper on ring current effects;[26] the second uses N-methylacetamide (NMA) as a probe molecule, primarily to examine hydrogen-bonding and other nonbonded contributions to the amide proton shift; and the third set consists of 57 low-energy conformers of the alanine dipeptide, to probe a variety of effects in a large peptide fragment. Details of the geometries are given in the following paragraphs.

(a) Methane as Probe Molecule. This follows our earlier work using methane as a probe molecule to study intermolecular effects on aliphatic C−H bonds. Rare gases He, Ne, and Ar (probing van der Waals dispersion effects) and a fluoride ion (probing electrostatic effects) were placed near one of the protons in methane, along the direction colinear with the C−H bond vector. Complexes started from an H···X distance of 1.725 Å (X = He), 1.825 Å (X = Ne), 2.225 Å (X = Ar), and 3.225 Å (X = F^-), and distances were increased in 0.1 Å steps for a total of 1.5 Å. Methane molecules were also placed at various distances and orientations from an NMA molecule both in and out of the NMA plane. Distances between the center of NMA and the methane molecule were approximately 5 Å. These calculations primarily probe the magnetic anisotropy of the NMA peptide group. Finally, a water molecule was placed near methane in a geometry similar to the rare gas complexes, such that the methane C and one of its H atoms were colinear with the water O, and the C−H methane bond vector bisected the H−O−H water angle. Distances sampled ranged from 2.125 to 4.625 Å in 0.1 Å increments. Similar geometries were generated for NMA near methane, with the C−H methane bond vector colinear with the NMA carbonyl bond, and the same distances as in the methane−water case were sampled.

(b) NMA as Probe Molecule. Similar calculations were used with NMA as a probe, looking at intermolecular effects on the amide proton shift. First, noble gas atoms He, Ne, and Ar were placed near the amide proton, along the direction colinear with the N−H bond vector. Similar calculations were carried out for a fluoride ion in place of the rare gas, to model simple electrostatic effects. Starting distances were 1.525 Å (He), 1.725 Å (Ne), 2.125 Å (Ar), and 3.025 Å (F^-). Next, hydrogen bonding effects were investigated by placing water molecules near NMA. In the "linear" series of geometries, shown in Figure 1a, the complexes were similar to those generated for methane near water: the water was oriented so that the water oxygen was hydrogen bonded to the NMA amide proton, with both NMA and water occupying the same plane and the N−H NMA vector bisecting the H−O−H water angle. The H···O distance was varied from 1.825 to 4.625 Å in 0.1 Å increments. In the "angled" series, a constant H···O distance of 2.225 Å was used. Starting from the linear NMA−water structure, two sets of angled complexes were generated. The first, shown on the left side of Figure 1b, was created by translating the water molecule in the NMA plane so that the N−H···O angle ranged from 0 to 60°. To create the second set shown on the right of Figure 1b, starting again from the linear NMA−water structure at 2.225 Å distance, the water was translated out of the plane, at 5−10° increments through a total of 45°. The water molecule was not rotated relative to the NMA molecule in any of these structures. Finally, linear amide−amide hydrogen bonds, depicted in Figure 1c, were constructed from two NMA molecules by translating one molecule relative to the position of the other such that both occupied the same plane and the amide proton of one NMA formed a linear hydrogen bond with the carbonyl O of the second NMA. Then, starting from a linear NMA−NMA structure at a distance of 2.225 Å, two sets of

(24) de Dios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1994**, *116*, 5307−5314.
(25) Pearson, J. G.; Wang, J.-F.; Markley, J. L.; Le, H.; Oldfield, E. *J. Am. Chem. Soc.* **1995**, *117*, 8823−8829.
(26) Case, D. A. *J. Biomol. NMR* **1995**, *6*, 341−346.

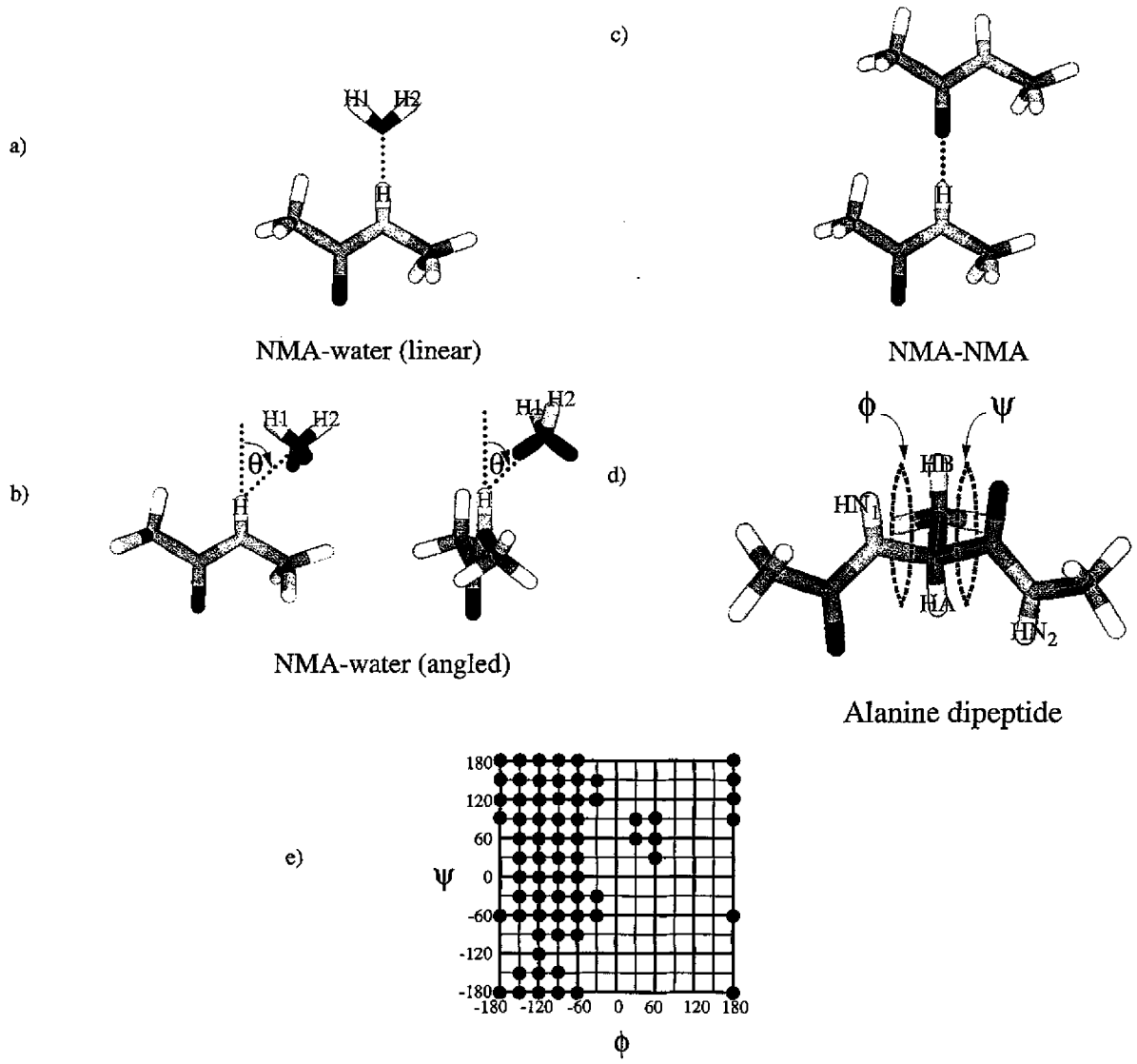

Figure 1. A subset of the molecules whose calculated shifts compose the shift pool. (a) NMA−water linear complex. (b) The two types of angled NMA−water complexes. Lone pairs on O are shown in black. Structure on the left was created by translating the water molecule within the NMA plane and has both water lone pair lobes pointing out of the NMA plane. Structure on the right was created by translating the water molecule out of the NMA plane and has one water lone pair lobe pointing toward the the NMA amide proton. (c) NMA−NMA dimer. (d) Alanine dipeptide. Torsions were sampled every 30° in ϕ and ψ space. Only conformations that occupied the two lowest energy levels according to a recent study by Thornton and co-workers[28] were retained; these are shown in (e).

angled hydrogen-bonded complexes were generated exactly as in the angled NMA−water case.

(c) Alanine Dipeptide Calculations. The ϕ and ψ backbone torsions shown in Figure 1d were sampled every 30° over the entire ϕ/ψ space. High energy structures were removed from the set; these were defined as torsions in the excluded or highest allowed energy sections of Ramachandran space as determined recently by statistical sampling of alanine torsions angles in the Brookhaven Protein Structure Databank[27] by Thornton and co-workers.[28] This criterion kept 57 out of 144 possible alanine dipeptide torsions. The resultant torsions are shown in Figure 1e. We decided to concentrate on "low-energy" structures to minimize (but not eliminate) the contributions to shifts from changes in local bond lengths and angles, which become increasingly important for more highly strained structures. For a similar reason we also excluded the HN_1 amide proton shift of the structures (ϕ, ψ) = (−30,120), (−30,150), (−60,180) (see Figure 1d for the names of the alanine dipeptide protons). In these structures the N-terminal peptide group (to which HN_1 belongs) was significantly nonplanar; the CH3−C−N−Ca torsion angle values were between 165° and 168°, while the remaining low-energy alanine dipeptide structures were clustered between 171° and 180°. Finally, we removed the two largest shifts from the pool, HN_2 from (ϕ, ψ) = (−60,30), (−60,60), to prevent their having a disproportionate effect on the fitted parameters.

2.2. Quantum Chemistry Calculations. Shielding tensors were computed using the *deMon* program,[23] which combines density functional theory with a sum-over-states perturbation approach. In this method, Kohn-Sham orbitals are inserted into a standard formula for chemical shielding,[29] and energy denominators are approximated by differences in Kohn-Sham orbital energies, corrected for changes in the exchange correlation potential that occur upon excitation. The gauge invariance requirement is treated using the individual gauge for localized

(27) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.

(28) Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. *Proteins: Str. Func. Gen.* **1992**, *12*, 345−364.

(29) Ramsey, N. F. *Phys. Rev.* **1950**, *78*, 699.

orbitals (IGLO) approach.[30] Full details of the method are given elsewhere.[23] The calculations used the (11s7p2d/6s2p)[7s6p2d/4s2p] "IGLO-III" orbital basis set of Kutzelnigg and co-workers;[30] this is a relatively large basis set, with 11 s-type and 7 p-type Gaussians on first row atoms (contracted to 7s/6p) along with two uncontracted polarization functions. For the rare gas atoms this basis set was not available; for these atoms we used the *deMon* orbital basis sets: DZV (He), DZVP (Ne), and TZVP (Ar). All calculations used the Perdew-Wang-91 (PW91) exchange-correlation potential[31] and the "Loc.1" correction for energy denominators.[23] For all of the bimolecular complexes, proton shifts are reported relative to the computed value for the monomer, *i.e.*, these are the "secondary" shifts due to the presence of the second molecule. For the alanine dipeptide, only the variation in shifts due to conformational changes is analyzed here; for convenience, shifts are reported relative to the computed proton shielding for methane, 31.19 ppm.

2.3. Empirical Shift Calculations. Details of the magnetic anisotropy, electrostatic, and close contact empirical shielding computational models are described in the sections below.

2.3.1. Magnetic Anisotropy Contributions. An asymmetry in the magnetic susceptibility of a chemical group leads, in the presence of an external magnetic field, to electrostatic currents in the group. These currents are the source of additional magnetic fields—and thus chemical shifts—at atoms surrounding the asymmetric group. As shown by McConnell,[14] when the asymmetric "source" group and the atom whose shift is being evaluated are far apart, the shift is given by

$$\delta_{anis} = (3L_0R^3)^{-1} \sum_{i=x,y,z} \chi_{ii}(3\cos^2\theta_i - 1) \qquad (1)$$

where L_0 is Avogadro's constant, R is the distance between the shifted atom and the asymmetric group, χ_{ii} is a component of the magnetic susceptibility tensor, and θ_i is the angle between the i axis and the vector **R**. For the case of an axially symmetric magnetic anisotropy, this becomes

$$\delta_{anis} = (3L_0R^3)^{-1}\Delta\chi(3\cos^2\theta - 1) \qquad (2)$$

where θ is the angle between the vector **R** and the normal to the plane of axial symmetry and $\Delta\chi$ is the difference between magnetic susceptibilities along the axis of symmetry and within the plane of symmetry.

Outside of aromatic ring systems which are known to be highly magnetically asymmetric, many chemical groups have relatively small magnetic anisotropies, as determined from susceptibility measurements on molecules in the gas phase primarily by Flygare and co-workers.[32] Due in part to its aromatic character, an exception is the peptide group. Data for formamide suggest that the susceptibility tensor for this group is roughly axially symmetric about the normal to the amide plane:

$$2\chi_{xx} - \chi_{yy} - \chi_{zz} = 2.2$$

$$2\chi_{yy} - \chi_{xx} - \chi_{zz} = 8.0 \qquad (3)$$

$$\Delta\chi = -\frac{1}{2}(\chi_{xx} + \chi_{yy}) + \chi_{zz} = -5.1 \qquad (4)$$

where z is the unit vector perpendicular to the peptide plane, y is the unit vector bisecting the peptide angle NCO, x is the unit vector perpendicular to y and z, and the susceptibilities are in units of 10^{-6} cm^3/mol. For comparison, magnetic anisotropy values for water are an order of magnitude smaller.

In the study below we consider two different magnetic anisotropy models for the peptide group. In the first model we assume the peptide group anisotropy has axial symmetry. In the second, we relax this requirement, allowing the group to be asymmetric. In both cases, we take the center of the peptide group magnetic anisotropy to be 0.7 Å from the carbon along the bisector of the NCO angle, as was done in previous work.[16] The DFT data is used to empirically fit the susceptibility values in eqs 1 or 2. The resultant values are compared to measured anisotropies and to other estimates in the literature.

In addition to "group" susceptibility parameters, commonly used for aromatic rings or for the peptide group as discussed above, there is a strong tradition of analysis of susceptibility anisotropies in terms of "bond" contributions.[33−35] It is possible to create models of this sort for the peptide group,[36] but this involves several adjustable parameters, and the predictions of nearby proton shifts are not markedly improved over simpler group models.[37] Bond-based models, though, have long been used for saturated systems, where electron pair bonds dominate the electronic structure and there is little delocalization.[34,38,39] These contributions are generally smaller than those from delocalized groups but can nevertheless lead to systematic shifts. A frequently cited example is the difference in chemical shift of 0.48 ppm for axial and equatorial protons in cyclohexane, which is thought to be due to anisotropy contributions from C−C and C−H bonds.[40,41] Below, we study the ability of a model incorporating C−C and C−H bond anisotropies to model part of the torsional dependence of amide proton shifts in peptides. The calculations use the axially symmetric magnetic susceptibility values determined by Flygare[38]

$$\Delta\chi_{C-C} = -7.7$$

$$\Delta\chi_{C-H} = -2.5 \qquad (5)$$

for C−C and C−H bonds, respectively, where the vector **z** in eq 4 is parallel to the C−X bond vector, and units are 10^{-6} cm^3/mol. The center of the C−C magnetic anisotropy is assumed to be midway between the two carbon atoms. For the C−H bond, the magnetic anisotropy center is taken as 0.77 Å away from the carbon along the C−H bond, in accordance with previous theoretical analyses.[34]

2.3.2. Polarization Effects. A significant contribution to chemical shifts can also arise from distant polar groups, which can polarize the electron cloud around the chosen nucleus and thereby increase or decrease the local shielding by electrons. The most significant term for a proton is expected to be proportional to the projection of the local electric field onto the X−H bond vector, where X is the atom connected to H.[15] Higher order terms can contribute as well although they are

(30) Kutzelnigg, W.; Fleischer, U.; Schindler, M. *NMR, Basic Principles Prog.* **1990**, *23*, 165.
(31) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(32) Flygare, W. H. *Chem. Rev.* **1974**, *74*, 653−687.
(33) Zrcher, R. F. *Prog. NMR Spectr.* **1967**, *2*, 205−257.
(34) Jackman, L. M.; Sternhell, S. *Applications of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry*; Pergamon: Oxford, 1969.
(35) ApSimon, J. W.; Beierbeck, H. *Can. J. Chem.* **1971**, *49*, 1328−1334.
(36) Asakura, T.; Niizawa, Y.; Williamson, M. P. *J. Magn. Reson.* **1992**, *98*, 646−653.
(37) Ösapay, K.; Case, D. A. *J. Biomol. NMR* **1994**, *4*, 215−230.
(38) Schmalz, T. G.; Norris, C. L.; Flygare, W. H. *J. Am. Chem. Soc.* **1973**, *95*, 7961−7967.
(39) Schneider, H.-J.; Schmidt, G. *J. Chem. Soc., Perkin Trans. II* **1985**, 2027−2031.
(40) Harris, R. K. *Nuclear Magnetic Resonance Spectroscopy, A Physicochemical View*; Longman Scientific & Technical: Essex, England, 1986.
(41) Bovey, F. A. *Nuclear Magnetic Resonance Spectroscopy*; Academic Press: San Diego, California, 1988.

Table 2. Lone Pair Charge Parameters[a]

moiety	LP−O	LP−O−LP	q_{LP}	q_O	q_X
water[b]				−0.720	0.360
ST2 water[c]	0.8	109.47	−0.2357	0.000	0.2357
water	0.8	109.47	−0.200	0.000	0.200
sp^3 O−X_2	0.8	109.47	−0.200	0.000	$(0.400 + q_O^{orig})/2 + q_X^{orig}$
sp^2 O−X	0.8	120.00	−0.200	0.000	$0.400 + q_O^{orig} + q_X^{orig}$
sp^2 N−X_3	0.8	120.00	−0.200	0.000	$(0.200 + q_N^{orig})/3 + q_X^{orig}$

[a] LP−O denotes the distance between each lone pair and oxygen (or nitrogen in the case of sp^2 N−X_3), in Å; LP−O−LP is the angle between lone pairs, in degrees; q_{LP}, q_O, q_X are charges on the lone pair, oxygen (or nitrogen) and atom X where X is an atom bonded to O (or N), respectively, in units of *e*. q^{orig} is the partial charge on the atom prior to making lone pair charge adjustments. [b] Atom-centered charge model, fit to yield a dipole moment of 2.0 D. [c] From Stillinger and Rahman.[50]

expected to be small for the systems considered here.[42] The shift due to polarization effects is thus given by

$$\delta_{pol} = -A\mathbf{E}(\mathrm{X} - \mathrm{H}) - B\mathbf{E}^2 - \cdots \qquad (6)$$

where **E** is the electric field and *A* and *B* are proportionality constants specific to the X−H bond.

Many years ago, Buckingham suggested that an appropriate value for *A* for a C−H bond would be -2×10^{-12} esu^{-1}.[15] Modern quantum mechanical methods can now be used to estimate the derivative of the proton shielding with respect to an external electric field. These calculations suggest larger values of *A*, close to -4×10^{-12} esu^{-1};[43,44] recent empirical estimates from fluorine-substituted hydrocarbons yielded a value of -3.7×10^{-12}esu^{-1}.[45] A similar value, $A = -3.1\ (\pm 0.24) \times 10^{-12}$ esu^{-1}, was obtained recently for the C−H bond in methane by performing density functional shift calculations on a methane probe near small ring molecules.[26] For protons in methane, *B* has been estimated at -0.3 to -0.4×10^{-18} esu^{-1} via shielding hyperpolarizability quantum chemistry calculations.[42,46] In most circumstances, estimated contributions from the second term in eq 6 are quite small compared to the first term. For example, the largest electrostatic field encountered in the structures studied here, ~0.1 *e*/Å^2, leads to linear and quadratic field-induced proton chemical shifts of 1.49 and 0.07 ppm, respectively, using $A = -3.1 \times 10^{-12}$ esu^{-1} and $B = -0.35 \times 10^{-18}$ esu^{-1}. We have thus retained only the first term in eq 6 below. As has been observed in quantum calculations, X−H bonds can have different *A* values depending on the identity of X. In this work we have fitted separate *A* values for C−H and N−H bonds.

Electrostatic fields were estimated using Coulomb's law and partial charges taken from the Amber 94 force field,[47] which were chosen to facilitate use of these formulas in macromolecular calculations. Partial charges for NMA were obtained by combining the AMBER charges for the *N*-methyl and acetyl peptide protecting groups. We did not use the default AMBER water charges, however, which are derived from the TIP3P water model,[48] as TIP3P water has the aqueous (~2.4 D) rather than gas phase (1.85 D) dipole moment value. The inflated charges would result in a smaller fitted *A* value in eq 6 for protons near water relative to the other molecules in the study. Instead, partial charges for water were derived by placing equal and opposite charges on the O and two H atoms so that the ratio of calculated to experimental water gas phase dipole moments[49] matched the ratio for NMA:

$$\left(\frac{\mu_{calc}^{AMBER}}{\mu_{\exp}}\right)_{\mathrm{NMA}} = \left(\frac{\mu_{calc}}{\mu_{\exp}}\right)_{\mathrm{water}}$$

Using the values $\mu_{calc}^{AMBER} = 4.0$ D and $\mu_{\exp} = 3.7$ D for NMA, and $\mu_{\exp} = 1.85$ D for water, this leads to a target dipole moment for water of $\mu_{calc} = 2.0$ D.

For O and N atoms, lone pair charge models, shown in Table 2, were also investigated as a possible alternative to the atom-centered charges assigned in AMBER. The lone pair charge assignments for water were based on the ST2 water model developed by Stillinger and co-workers,[50] in which, charges of -0.2357 *e* are placed at the lone pair sites, 0.8 Å away from the water's O, at LP−O−LP angles of 109.47°, equivalent to the H−O−H angle. Identical and opposite charges are assigned to the H atom centers, and the O atom center has a charge of 0.0. In the current work, the ST2 model was modified by reducing the magnitude of the charges assigned to the H's and lone pairs in water to -0.200 *e* in order to obtain a water dipole moment of 2.0 D, as discussed above. Lone pair charges on molecules other than water were derived by analogy to the ST2 model: -0.200 *e* were placed at the lone pair atom sites and the sum 0.400 *e* + original charge on O was divided among atoms bonded to the O. The charge on the O was 0.0 *e*. Further details on lone pair placement and charge assignments are given in Table 2. For NMA, the lone pair and AMBER charges produced identical dipole moments to within 0.2 D. We note that this approach is just one of many ways to assign non-atom-centered charges to atoms; further investigation into lone pair charge distributions and their effects on chemical shifts is underway in our laboratory.

For complexes, only the charges on the molecule opposite to the one whose shift was being considered were used to calculate the electric fields. For the alanine dipeptide, charges on atomic groups directly bonded to the proton in question were ignored, since the positions of these atoms with respect to the proton should change very little with ϕ and ψ torsion angle variation. The atoms in the alanine dipeptide whose charges were skipped were as follows: For HN_1, the first peptide group O, C, N, H, and the bonded CA; for HA and HB: CA, HA, CB, HB1, HB2, HB3; for HN_2, the second peptide group O, C, N, H, and the C-terminal methyl group. See Figure 1d for the names of the alanine dipeptide protons.

(42) Augspurger, J. D.; Dykstra, C. E. *J. Phys. Chem.* **1991**, *95*, 9230−9238.
(43) Grayson, M.; Raynes, W. T. *Magn. Reson. Chem* **1995**, *33*, 138−143.
(44) Coriani, S.; Rizzo, A.; Ruud, K.; Helgaker, T. *Mol. Phys.* **1996**, *88*, 931−947.
(45) Abraham, R. J.; Warne, M. A.; Griffiths, L. *J. Chem. Soc., Perkin Trans. 2* **1997**, 203−207.
(46) Raynes, W. T. In *Nuclear Magnetic Shieldings and Molecular Structure*; Tossell, J. A., Ed.; Kluwer Academic Publishers: Dordrcht, 1993; pp 401−420.
(47) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(49) *CRC Handbook of Chemistry and Physics*, 73rd ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, 1992−1993.
(50) Stillinger, F. H.; Rahman, A. *J. Chem. Phys.* **1974**, *60*, 1545−1557.

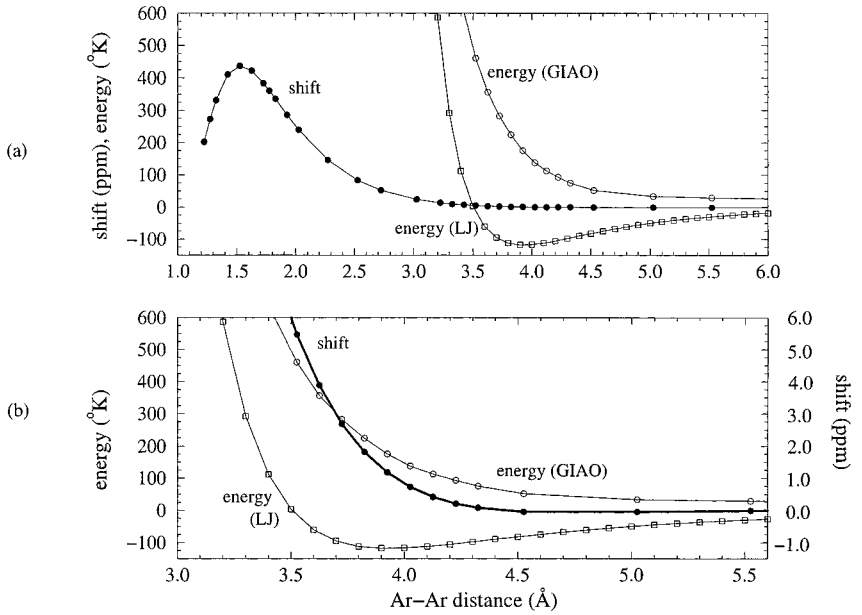

Figure 2. (a) GIAO shifts (filled circles), DFT single point energies (empty circles), and Lennard-Jones energies (open squares) for an Ar−Ar dimer. GIAO calculations were performed using the DFT approach with the Becke 1988 exchange function and the Perdew/Wang 1991 correlation function and the cc-pVTZ basis set.[69] Lennard-Jones energies used the values $\epsilon k = 117.7$°K and $\sigma = 3.504$.[60] (b) Same as (a), but expanded along the *x*-axis and with respect to shift scale.

2.3.3. Close-Contact Contributions. The study of solvent effects and nonbonded contributions to proton chemical shifts has a long and somewhat confusing history. The notion that dispersion contributes to chemical shifts was first proposed by Stephen[51] and Buckingham *et al.*[52] At close proximity, London forces due to correlations of fluctuating dipoles can induce a buildup of electron density between molecules. The resultant loss of electron density near the nuclei is expected to decrease chemical shielding according to the average square of the electric field, which can be further described[46,53] in accordance with the Drude model[54] for atoms, to give a "close contact" shift:

$$\delta_{cc} = -B\langle \mathbf{E}^2 \rangle = -B\frac{3}{2}\frac{U_1 U_2}{U_1 + U_2}\frac{\alpha_2}{r^6} \qquad (7)$$

Here $\langle \mathbf{E}^2 \rangle$ is the average square fluctuating field at atom 1 induced by atom 2, U_i is the ionization energy of atom i, α_2 is the polarizability of atom 2 and r is the distance between the atoms, and B comes from eq 6.

Experimental measurements of changes in chemical shifts of nonpolar, magnetically isotropic molecules as a function of density, temperature, and solvent support the existence of a close contact or van der Waals contribution to chemical (de)shielding, that can be modeled with a $1/r^6$ dependence.[53,55−59] Although this has the same distance dependence as eq 7, two arguments suggest that dispersion effects are not the dominant interaction. First, Hartree−Fock or DFT calculations on rare gas dimers (and other nonpolar interactions) show a deshielding in rough accord with observation,[53] even though dispersion energetics are not correctly modeled at this level of theory. This is depicted in Figure 2, which shows the energy of an argon dimer as a function of interatomic distance, as computed via density functional theory and via the Lennard-Jones equation, using typical literature parameters.[60] Chemical shifts, computed by

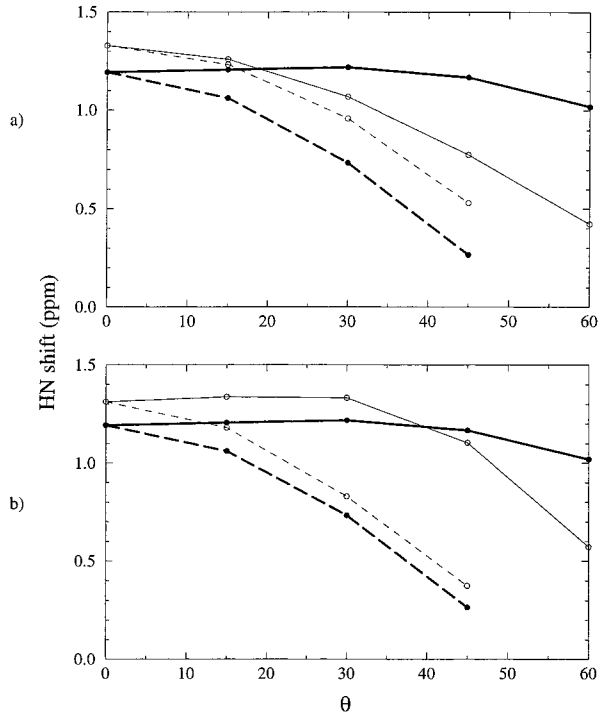

Figure 3. DFT and empirically calculated complex-induced shifts (complex shift − monomer shift) for the NMA amide proton in the NMA−water complex as a function of the NMA−water hydrogen bond angle, θ, defined as 180° − the N−H···O angle. Results are shown using either an atom-centered (a) or ST-2 based lone pair (b) charge model for O. Solid and dashed lines denote shifts for geometries in which lone pair points into and away from the NMA amide proton, respectively (geometries shown in right and left sides of Figure 1b, respectively.) Heavy and fine lines denote DFT and empirically calculated shifts, respectively.

(51) Stephen, M. J. *Mol. Phys.* **1958**, *1*, 223.
(52) Buckingham, A. D.; Schaefer, T.; Schneider, W. G. *J. Chem Phys.* **1960**, *32*, 1227−1233.
(53) Jameson, C. J.; de Dios, A. C. *J. Chem. Phys.* **1992**, *97*, 417−434.
(54) Hirschfelder, J. O.; Curtiss, C. F.; Bird, R. B. In *Molecular theory of gases and liquids*; Wiley: New York, 1954.
(55) Luhmer, M.; Dejaegere, A.; Reisse, J. *Magn. Reson. Chem.* **1989**, *27*, 950−952.
(56) Bennett, B.; Raynes, W. T. *Mag. Reson. Chem.* **1991**, *29*, 946−954.
(57) Giessner-Prettre, C.; Gresh, N.; Maddaluno, J. *J. Magn. Reson.* **1992**, *99*, 605−610.
(58) Lau, E. Y.; Gerig, J. T. *J. Chem. Phys.* **1995**, *103*, 3341−3349.
(59) Lau, E. Y.; Gerig, J. T. *J. Am. Chem. Soc.* **1996**, *118*, 1194−1200.

the Gauge-Independent Atomic Orbital (GIAO) approach,[61] are also shown. The dispersion-induced energy minimum is completely absent from the DFT energy curve. Moreover, the calculated chemical shift rises continuously as the argon atoms approach each other, well past the interval where the repulsive forces overcome dispersive effects according to Lennard-Jones theory. Second, mean square fields much larger than those arising from eq 7 would be required to explain the observed (and quantum-mechanically calculated) close-contact shifts.[46,58,62] Nevertheless, some of the general ideas of eq 7 appear to have rough validity: the $1/r^6$ distance dependence is approximately correct (although other functional forms would probably also fit the available data), and the predicted dependence on ionization potentials and polarizabilities also seems to hold.[53,58]

In this work, we explore using eq 7 to calculate the close contact shift contribution, treating B as an adjustable parameter. Ionization constants and polarizability values were taken from the CRC Handbook of Chemistry and Physics.[49] Only atoms four or more bonds away were considered in calculating the close contact contribution to chemical shifts. The results below show that a single value for B appears to account for close-contact shifts for both N−H and C−H protons interacting with rare gas atoms and with first-row atoms N and O. A somewhat larger B value is required for protons interacting with C atoms. Further investigations into the nature of close-contact deshieldings will be reported elsewhere.

2.4. Parameter Fitting. The total empirical shift is given by

$$\delta_{empirical} = \delta_{anis} + \delta_{pol} + \delta_{cc} + \delta_{const} \tag{8}$$

where the anisotropy (δ_{anis}), electrostatic polarization (δ_{pol}), and close contact (δ_{cc}) contributions are computed via the models and equations described in the sections above, and the constant (δ_{const}) comprises the local shielding differences between protons in the alanine dipeptide and protons in methane, which was used as a reference. Constants were not necessary for the protons in complexes as their shifts were referenced to the monomer values.

In order to fit the peptide anisotropies, the electrostatic A and and close contact B coefficients, and constants in eq 8, we used a nonlinear optimization program to minimize

$$\rho(z) = \sum_{shifts} \left(1 + \frac{1}{2}z^2\right) \tag{9}$$

where $z = (\delta_{DFT} - \delta_{empirical})/c$ with $c = 0.5$ ppm. In comparison with conventional least-squares optimization, this has the effect of reducing the importance of shifts whose errors are much larger than the rms error over the data set. In order to determine the predictive ability of the correlations and the uncertainty in the parameters, we repeated the optimizations ten times, each time removing 1/10 of the points to be fitted. This "jackknife" procedure can then be used to develop an estimate of the uncertainties in the final parameters:[63] use the ten sets of parameter estimates to define *pseudovalues* for each parameter as

$$y^*_j = 10y_{all} - 9y_j, \quad j = 1, 2, \cdots 10 \tag{10}$$

where y_{all} is the parameter estimate when all of the data is considered, and y_j the estimate when the subset j is omitted. Then the jackknife estimate for the parameter is the mean of the y^*_j values, and the estimate of its uncertainty is determined by standard formulas for the uncertainty of a mean.[63]

3. Results

In the initial sections below, we consider the efficacy of the various empirical models in computing contributions to proton shifts for the structures described in the Methods above. Subsequently, shift contributions as a function of backbone torsion angles (in the alanine dipeptide) and as a function of distance (in complexes) are discussed.

3.1. Shifts Due to Magnetic Anisotropy and Electrostatic Polarization Shift Contributions. For our initial calculations, protons involved in close-contact interactions were removed from the DFT shift pool, and empirical magnetic anisotropy and electrostatic parameters were fit to the remaining data. Close contact between a proton and atom X was defined as a calculated close contact shift contribution (given by eq 7 with $B = -0.4 \times 10^{-18}$ esu^{-1}) of 0.010 ppm or greater. For X = O, N, C, H, F, He, Ne, and Ar this translates into a H−X cutoff distance of 2.9, 3.0, 3.2, 2.8, 2.9, 2.4, 2.6, and 3.3 Å, respectively. The remaining set of DFT shifts, which we refer to as *S1*, was used to fit parameters for peptide magnetic anisotropy, electrostatic and constant empirical contributions to proton chemical shifts via the jackknife procedure described in the Methods section. The results are shown in Tables 3 and 4 and Figure 4 for the symmetric planar peptide group magnetic anisotropy model, the ST2-based lone pair charge model for O, and magnetic anisotropy contributions from C−C and C−H bonds (see Methods for detailed descriptions of empirical models). These and results for other empirical models are described in the text below.

Overall, the *S1* shifts are well fit by the empirical shift equations, as demonstrated in the fit statistics in Table 4. Variations in the empirical models, including an asymmetric model for peptide magnetic anisotropy, removal of magnetic anisotropy contributions from C−C and C−H bonds, and lone pair charges for N or atom-centered charges for both O and N produced fits of comparable quality to that shown in Table 4 for the *S1* shift set. The latter result is as expected, since at the medium to long range distances explored in shift set *S1*, detailed descriptions of the contributing groups, such as lone pair electron distribution, can be approximated by more macroscopic, in this case atom-centered, charge models. In addition, the protons most affected by C−C and C−H magnetic anisotropies are absent from *S1* as they are involved in close contact interactions.

The parameter values for the electrostatics shift contribution were well-determined; both the H−C and H−N parameters had jackknife uncertainties (in parentheses in the table) on the order of 1%, and their variations with changes in empirical models were of the same order. Only slightly larger variations were seen when the shift set was increased to include protons in close contact with other atoms (see Table 3, Fit 3). The electrostatics parameter A of -3.8×10^{-12} esu^{-1} for a C−H bond is very close to other recent estimates cited above. Some previous empirical shift fitting studies have determined smaller electrostatic shift parameter values,[16,17] but those were based on fits to shifts collected in aqueous solution, where electrostatic interactions are dampened by dielectric screening effects. Interestingly, as seen from Table 3, the fitted parameter for a C−H bond is about 10% greater than that for a N−H bond,

(60) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976.
(61) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8324−8328.
(62) Abraham, R. J.; Warne, M. A.; Griffiths, L. *J. Chem. Soc., Perkin Trans. 2* **1997**, 881−886.
(63) Mosteller, F.; Tukey, J. W. *Data Analysis and Regression. A Second Course in Statistics*; Addison-Wesley: Reading, MA, 1977.

Table 3. Empirical Chemical Shift Parameters[a]

								alanine dipeptide constants			
fit	shift set	$\Delta\chi_1$	$\Delta\chi_2$	*A(HN)*	*A(HC)*	*B*	*B(C)*	HN_1	HN_2	*HA*	*HB3*
1	*S1*	−6.25(0.97)		−3.46(0.02)	−3.81(0.03)					3.22(0.08)	1.54(0.04)
2	*S2*	−6.25		−3.46	−3.81	2.82(0.08)	4.39(0.50)	5.01(0.02)	5.21(0.04)	3.01(0.02)	1.47(0.02)
3	*S2*	−5.93(0.62)		−3.45(0.06)	−3.73(0.05)	2.86(0.61)	4.31(0.61)	5.01(0.02)	5.21(0.04)	3.05(0.06)	1.47(0.02)
4[b]	*S2*	7.95(1.62)	4.91(0.67)	−3.42(0.06)	−3.78(0.03)	2.86(0.09)	4.35(0.59)	5.02(0.02)	5.21(0.04)	3.11(0.07)	1.48(0.02)

[a] Parameters are the result of applying the jackknife procedure (see Methods) to the indicated set of shift data. See Table 1 and text for shift set definitions. All results are for a planar symmetric magnetic anisotropy model for the peptide group and for a lone pair charge model on O unless otherwise noted; results for other models are discussed in the text. $\Delta\chi_1$ and $\Delta\chi_2$ are the magnetic susceptibilities for the peptide group; for the planar symmetric model $\Delta\chi_1 = \Delta\chi$ in eqs 2 and 4; for the asymmetric model $\Delta\chi_1$ and $\Delta\chi_2$ correspond to $2\chi_{xx} - \chi_{yy} - \chi_{zz}$ and $2\chi_{yy} - \chi_{xx} - \chi_{zz}$, respectively, in eq 3; axes are defined in the text. *A*(*HN*) and *A*(*HC*) are the electrostatics parameters in eq 6 for protons attached to a nitrogen and carbon atom, respectively. *B* and *B*(*C*) correspond to $-B$ in eq 7 for protons near O, N and rare gas atoms, and near C, respectively. The anisotropy parameters are in units of 10^{-6} cm^3/mol; electrostatics parameters are in units of 10^{-12} esu^{-1}; close contact parameters are in units of 10^{-18} esu^{-2}; and constants are in ppm. Jackknife estimate of standard deviations in parameter values are in parentheses. Missing parentheses indicate the parameter value was held constant in the fit. [b] Asymmetric peptide magnetic anisotropy model was used.

Table 4. Statistics on the Fits[a]

		no. of shifts[b]							
fit	shift set	HN	HC	DFT range[c]	Pearson *r*	rms error	slope	int	max err(±)
1	*S1*	72	127	2.132	0.982	0.064	0.962	0.026	−0.166/0.166
2	*S2*	254	289	4.473	0.959	0.136	0.940	0.047	−0.487/0.398
3	*S2*	254	289	4.473	0.959	0.135	0.934	0.047	−0.498/0.397
4	*S2*	254	289	4.473	0.959	0.135	0.932	0.048	−0.489/0.398

[a] Results using the four sets of jackknife parameters from the fits listed in Table 3. See Table 1 and text for shift set definitions. Final five columns give the linear correlation coefficient, the root-mean-square difference between the DFT and empirical values in ppm, the slope and intercept of the best fit line to the empirical versus DFT shifts, and the maximum signed errors in empirically fitted shifts in ppm. [b] HN: protons attached to a nitrogen; HC: protons attached to a carbon. [c] Span of DFT shifts in the set in ppm.

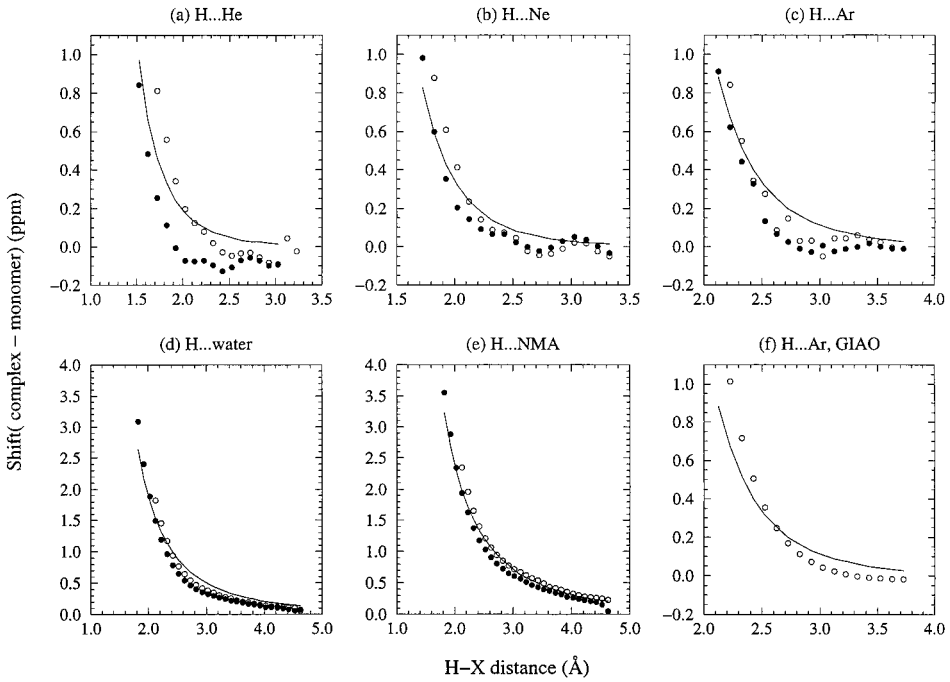

Figure 4. Quantum chemically calculated complex-induced shifts for the amide proton of NMA (filled circles) and the most closely approached methyl proton of methane (empty circles) near rare gas atoms {(a) through (c) and (f)} and near water (d) and NMA (e). The total empirical shift, computed using parameters from Fit 2 in Table 3, is shown as a solid line. Quantum shifts for (a)−(e) were calculated via the *deMon* program,[23] and (f) was calculated via the Hartree−Fock GIAO approach with the 6-31++G** basis set.

while experimentally determined bond polarizabilities find a 10% difference in H−N vs H−C in the opposite direction.[64] It is not clear, however, that chemical shielding at a nucleus and electron density movement along the center of a bond should respond in an identical manner to electric fields at their respective sites.

Also shown in Table 3 are the fitted magnetic anisotropy parameters. The best fit value using the symmetric planar peptide anisotropy model was $\Delta\chi = -6.25 \pm 0.97 \times 10^{-6}$ cm^3/mol. Previous estimates of peptide magnetic anisotropy include -7.9×10^{-6} cm^3/mol from fits of experimental protein shifts to empirical shift equations similar to the ones discussed here,[16] -5.51×10^{-6} cm^3/mol from measured Zeeman effects and bulk magnetic susceptibilities of formamide,[64,65] and -5.36×10^{-6} cm^3/mol based on a modification of ring current theory as applied to the peptide group.[66] Parameters were also fit to the

(64) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533−8542.
(65) Tigelaar, H. L.; Flygare, W. H. *J. Am. Chem. Soc.* **1972**, *94*, 343−346.
(66) Pauling, L. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 2293−2294.

shift set using the asymmetric planar magnetic anisotropy model; however, the resultant values were highly uncertain due to the lack of short-range-induced shifts in set *S1*, and the shift statistics were nearly identical to those obtained with the symmetric peptide model. These results suggest that for the geometries tested here, the effects of the magnetic asymmetry of the peptide group are not significant, thus a symmetric planar model is sufficient.

Overall, the *S1* results suggest that any of the electrostatic and magnetic anisotropy empirical shift models under consideration here could be used to compute relatively accurate DFT shifts, using physically reasonable parameters, for protons involved in medium to long range interactions. The following two sections consider short-range shift contributions.

3.2. Results for NMA−Water Complexes: Tests of the Lone Pair versus Atom-Centered Charge Models. In order to determine whether the lone pair and atom-centered charge models both gave acceptable shifts at close contact distances, empirical shifts were calculated and compared with DFT values for eight NMA−water conformations with hydrogen bond angles ranging from 0 to 60° (hydrogen bond angle is defined as 180 − angle(N−H···O)). Two overall geometries were investigated: In the first geometry, the water was translated in the plane of the NMA, such that, at each angle, the O−LP distance was equivalent for both lone pairs; neither lone pair lobe pointed directly into the NMA amide proton (see left side of Figure 1b). In the second geometry, the water was translated out of the NMA plane such that one of the lone pair lobes pointed toward the amide proton (see right side of Figure 1b). For all of the complexes, the hydrogen bond distance from HN to O was 2.225 Å, and the H−O−H and NMA planes were parallel. In the following discussion, the first and second sets of geometries are referred to as having their lone pair lobes pointing either "away from" or "toward" the amide proton, respectively. The total calculated empirical and DFT shifts are shown in Figure 3. The empirical shifts were computed using the jackknife parameters from fits to the shift pool *S1*. Close contact shifts, which will be discussed in the following section, were also added in; however, this contribution is identical for all angles because the HN···O distance does not change. Variations in the empirical shift as a function of angle are thus entirely due to the electrostatic component. The empirical electrostatic shifts were calculated using either an atom-centered charge distribution for water or a charge distribution based on the ST250 lone pair water model of Stillinger and co-workers (see Methods section).

As can be seen from Figure 3, for increasingly nonlinear hydrogen bonded geometries the DFT shift decreased when the lone pair vector pointed away from the HN, but it remained nearly constant when the lone pair vector pointed toward the HN. Using the atom-centered charge model (Figure 3a), the empirical shifts decreased for all nonlinear hydrogen bond geometries, regardless of the direction of the lone pair vector, in contrast to the the DFT results. The ST2-based lone pair charge model, however (Figure 3b), reproduced the qualitative behavior of the DFT shifts for structures with hydrogen bonds deviating up to approximately 35° from linearity, and agreement improved further when a recently modified version of the ST2 model[67] was used (results not shown). The errors encountered for highly nonlinear structures probably involve contributions from close contacts between the amide N and the water O that are not included in the empirical shift model; for instance, at an angle of 60°, the distance between the water O and the NMA N is only 2.8 Å.

Similar results were obtained for the geometric shift dependence of the hydrogen bonded amide proton in an NMA−NMA complex (results not shown). For nonlinear hydrogen-bonded structures, the DFT calculated shifts depended upon the direction of the lone pair vector, with larger shifts occurring when the vector pointed directly toward the amide proton. This behavior was reproduced up to 45° from the linear hydrogen bond geometry when a lone pair rather than atom-centered charge model was used for the NMA carbonyl group. Nonelectrostatic empirical contributions (magnetic anisotropy, close contacts) cannot account for the angled NMA dimer DFT shift behavior without using nonphysical parameter values.

Assuming that the DFT shifts are accurate, these results suggest that not only proton−oxygen distances and angles but also the positions of the oxygen lone pairs can significantly affect chemical shifts of hydrogen bonded protons. These effects may be considered to derive from the directionality of the lone pair orbitals of the hydrogen bond acceptor and are well approximated, at least at small hydrogen bond angles, using the ST2-based lone pair electrostatic model for charges on O.

3.3. Close Contact Contributions. DFT and empirically calculated proton shifts as a function of intermolecular distance are shown in Figure 4 for methane and NMA complexed with either rare gas atoms, water, or NMA. It is clear that in all cases, as the molecules approach each other, the proton shifts increase dramatically. For methane and NMA near the rare gas atoms, there is some apparent noise in the chemical shift data; we attribute this to the poorer quality basis sets that were used for the rare gases. Figure 4f depicts the shift for methane near argon as calculated via the Hartree−Fock GIAO approach with the 6-31++G** basis set. A comparison of part c with f of Figure 4 demonstrates that despite the noise in the *deMon* shifts, their overall magnitude is consistent with shifts computed via the GIAO approach, except at the closest distances, where the *deMon* shift is several tenths of a ppm below the GIAO value. In all of the complexes, at a constant H−X distance, the methyl shifts are larger than the amide values. However, the difference was not significant enough to warrant separate *B* values for H−C versus H−N protons when fit to the full set of shift data, as described in further detail below.

To parameterize the close contact empirical shift term in eq 7, we used the full *S2* shift set, which contains data for both methyl and amide protons involved in short- as well as medium- and long-range interactions. The *S2* set is described in further detail in Table 1 and in the Methods section. *B* values in eq 7 were obtained by fitting to *S2* while holding the electrostatic and anisotropy parameters constant at their *S1*-fit values. In some cases we tried fitting multiple *B* values depending on the identity of the atoms either in close contact with or bonded to the proton whose shift was being fit. In some fits, the electrostatic and anisotropy parameters were permitted to vary as well. Results of the fits are shown in Tables 3 and 4 for the symmetric peptide anisotropy model, the ST2-based lone pair charge model for O and magnetic anisotropy contributions from C−C and C−H bonds. Results for other models are also in the table or are discussed below.

Irrespective of which anisotropy or charge model was adopted, using separate *B* values for H−C vs H−N protons had little significant effect on the fit statistics, most likely because extremely close contacts (where the methyl vs amide differences are largest) are rare for methyl protons. The statistics were significantly improved, however, by setting $B = 0.0$ for contributions from H atoms. (Recent empirical fits by Abraham *et al.*[68] to alkanes also suggested that H−H interactions were

(67) Head-Gordon, T.; Stillinger, F. H. *J. Chem. Phys.* **1993**, *98*, 3313−3327.

less deshielding than other close contacts.) The maximum errors were further reduced by approximately 0.11 ppm by fitting a separate, larger *B* value for close contact contributions from C versus all other heavy atoms. This need for separate *B* values depending on the identity of the interacting atoms may be due to inaccuracies in using atomic polarizability and ionization energy values in eq 7, rather than values reflective of the atom's molecular environment. The larger *B* value for C appears to result primarily from methyl group contributions, which are otherwise underestimated when *B* for close contacts between protons is set to zero.

The best fit *B* parameter values from Fit 2 in Table 3 are fairly well-determined, though not as certain as the electrostatic parameters, as can be seen from the jackknife uncertainties in Fits 1 and 3. The parameters change by up to 5% when different anisotropy and charge models are used or when all of the parameters are allowed to vary in the fit. For comparison, estimated values for the shielding hyperpolarizability value *B* in eq 6 from the literature (for protons in molecules comparable to those in *S2*)[44,46] are on the order of -0.4×10^{-18} esu^{-1} or about one order of magnitude smaller than the fitted *B* parameters for eq 7 (Table 3). Similar discrepancies have also been observed in studies of the temperature effects on gaseous methane shifts[46] and solvent induced shielding of neon atoms.[58] These results support the notion that observed close-contact shifts do not arise primarily from London dispersion interactions but are more likely to arise from exchange−repulsion overlap effects that have a similar distance dependence.

3.4. C−C and C−H Bond Anisotropies. DFT shifts in the *S2* pool were also compared with empirical shifts calculated either with or without the C−C and C−H bond anisotropy contribution. Axially symmetric bond magnetic susceptibility anisotropy values of -7.7×10^{-6} cm^3/mol for C−C and -2.5×10^{-6} cm^3/mol for C−H, as determined by Flygare,[32] were used in eq 2 to calculate the magnetic anisotropy shift contributions, with the axis of symmetry defined to be along the C−C and C−H bonds. These anisotropy values produce good agreement between the experimental, DFT, and empirically calculated equatorial versus axial proton shift in cyclohexane and between DFT and empirically calculated shifts in other small hydrocarbon molecules (Sitkoff and Case, unpublished results). Excluding contributions from C−C and C−H protons led to a large error (∼0.6−0.7 ppm) for HN_1 protons in alanine dipeptide structures with backbone torsions of $\phi = 60$. These errors were reduced to approximately 0.3−0.4 ppm when C−C and C−H bond anisotropies were included, depending on which models were used for the peptide anisotropy and O and N charges.

3.5. Overall Fit of Empirical to DFT Data. The statistics on empirical versus DFT calculated shifts for the data set *S2* are shown in Table 4 for ST2-based lone pair charges on O, C−C, and C−H magnetic anisotropy contributions and either a symmetric or asymmetric peptide magnetic anisotropy. The statistics for *S2* are worse overall than those for *S1*; for example, using results from Fits 1 and 2 in Table 3, rms errors increased from 0.06 ppm for *S1* to 0.14 ppm for *S2*, and the maximum unsigned error increased from 0.17 to 0.49 ppm. For many of these protons, however, the close contact and/or electrostatics contributions are quite large, and at short-range distances these are extremely sensitive to positions, thus small changes in the fitted sensitivity to distance can lead to large errors in calculated shifts.

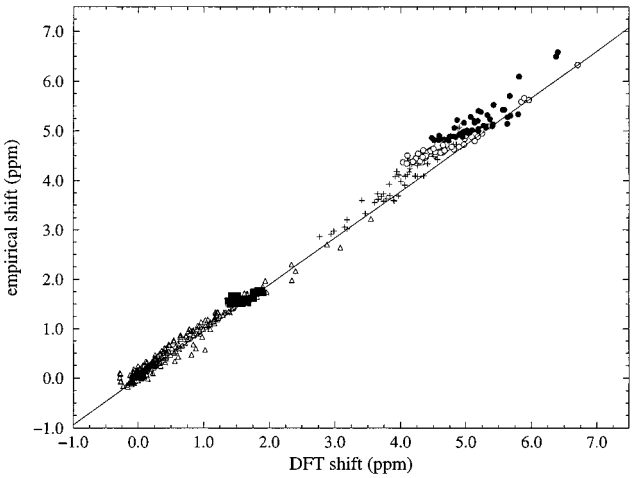

Figure 5. Empirical versus DFT shifts for the protons in set *S2* using parameters from Fit 2 in Table 3. Symbols for the alanine dipeptide protons are HA: +; HB3: filled squares; HN1: unfilled circles; HN2: filled circles. All remaining shifts are shown as unfilled triangles.

Using an atom-centered rather than ST2-based lone pair charge model for O worsened the fit of empirical to DFT shifts in some of the alanine dipeptide amide protons as well as in the angled NMA−water and NMA−NMA complexes. Adding a lone pair rather than atom-centered charge model for N had no significant effect on the fit statistics, nor did using the asymmetric anisotropy model, refitting the anisotropy parameters to the whole data set. The asymmetric magnetic susceptibilities resulting from the latter fit were considerably better defined than in the fit to set *S1*; the values corresponding to the anisotropies in eq 3 were 8.0 ± 1.6 and $4.9 \pm 0.7 \times 10^{-6}$ cm^3/mol; for comparison, the experimental anisotropies for formamide are 2.2 and 8.0×10^{-6} cm^3/mol.[65] The fact that the parameters are close in value in the in- and out-of-plane directions and that the fit statistics do not change significantly between models suggests that even at short-range distances, an axially symmetric peptide anisotropy model is sufficient to describe chemical shifts in peptides and small molecule complexes. This result is in agreement with a previous empirical shift study that considered proton shifts in proteins.[16]

Empirical versus DFT calculated shifts for the entire set *S2* are shown in Figure 5, using the models and parameters from Fit 2 in Table 3. The majority of the DFT shifts are well fit by the parameterized empirical shift equations. One outlier is the NMA−water complex that has the largest tested nonlinear hydrogen bond in the NMA plane (60° from linearity). As was already described, this structure probably involves close contact interactions between the water and the NMA N.

The protons that are least well fit overall are the amide protons in the alanine dipeptide. As is evident from Figure 5, many of these DFT shifts correlate with the empirical shifts but with a slope of less than one; that is, not all of the DFT shift variation can be explained via the empirical equations. A likely reason for the poorer fits encountered for these protons is that the shift calculations assume that contributions from closely bonded atoms are constant for all alanine dipeptide structures. In fact, there are many small changes in the local structure of the entire peptide group. These changes probably affect the electron density, and thus shift values, at the amide proton due to the highly electron dense and resonant character of the peptide bond.

(68) Abraham, R. J.; Edgar, M.; Glover, R. P.; Warne, M. A.; Griffiths, L. *J. Chem. Soc., Perkin Trans. 2* **1997**, 333−341.

(69) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, J. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94 (Revision B.2)*; Gaussian, Inc.: Pittsburgh, PA, 1995.

Table 5. Alanine Dipeptide HA Shifts[a]

quantity	*deMon*	S&C	O&C	W&A
$\alpha - \beta$	−0.44	−0.54	−0.42	−0.59
(60,60) − β	−1.28	−1.56	−0.75	−0.52
min	(60,90)	(60,90)	(60,180)	(−30,−60)
max	(−120,−60)	(−120,−60)	(−120,−60)	(−120,120)
range	2.18	2.29	0.94	0.74

[a] DFT and empirically calculated shifts for HA in alanine dipeptide using the DFT *deMon* method, and the following empirical approaches: S&C: eq 8 using parameters from Fit 2 in Table 3; O&C: Osapay and Case;[37] and W&A: Williamson and Asakura.[17] All shifts were calculated for the alanine dipeptide structures described in the Methods section. α shift values are the average of shifts for $(\phi,\psi) = (-60,-30)$ and $(-60,-60)$; β shift values are for $(\phi,\psi) = (-120,120)$. Maximum and minimum shifts and range apply to the "allowed" region of Ramachandran space only, as defined in the Methods section. All shifts are in ppm.

3.6. Alanine Dipeptide Chemical Shifts and Chemical Shift Contributions. An overall description of the *deMon* calculated shifts for HA in alanine dipeptide is given in Table 5, in parallel with results calculated via the empirical methods discussed here and via two previous empirical approaches.[16,17] In considering the shifts calculated by these earlier approaches, it is important to recall that those methods were parameterized against experimentally measured shifts in proteins, where electrostatic effects are screened by aqueous solvent. It is thus unrealistic to expect them to yield accurate electrostatic and therefore total shifts in the gas phase. Despite this caveat, all of the methods find that, as was described previously,[16] most of the variation in the HA shift in the allowed region of (ϕ,ψ) space is along the ϕ coordinate. Likewise, in all of the methods, α proton shifts are 0.5 ± 0.09 ppm smaller than β proton shifts; it turns out that peptide magnetic anisotropy contributions are primarily responsible for the α-β shift difference, and this contribution is approximately equal in all of the empirical methods. The main distinction between shifts computed here and in the previous empirical approaches is that the range of calculated shifts is larger here by more than a factor of 2. A similar difference in the calculated shift range is found for HB3 and HN_2 protons, while for HN_1 the ranges differ by more than a factor of 4 (or a difference of 2 ppm). The close contact term and C−C and C−H bond magnetic anisotropy terms, which were not present in the previous empirical shift implementations, and the electrostatic term *A*, which has a larger value than in previous methods, appear to be about equally responsible for the improved agreement between shifts computed by the current empirical approach and the *deMon* quantum mechanical method.

In Figure 6 the peptide magnetic anisotropy, electrostatic, C−C and C−H bond anisotropy and close contact shift contributions, and the error in DFT versus empirical shift are plotted for some of the protons in the shift pool. Data are shown in order along the *x*-axis for the hydrogen bonded HN proton in the NMA dimer, methane protons near NMA, and the HN_1, HA, HB, and HN_2 protons in the alanine dipeptide. The structures are sorted either from smallest to largest in terms of distance (NMA dimer) or by error in calculated versus empirical shifts (NMA-methane, alanine dipeptide). Variations in the electrostatic term dominate the HN_1 proton shifts, with the exception of torsions in which $\phi = 60$, for which there is a large variation in C−C and C−H bond anistropy. In contrast the HN_2 proton generally has large contributions from all four empirical sources, although the peptide anisotropy and close contact terms largely cancel each other. This difference between alanine dipeptide amide groups appears to be due to the geometry of the backbone chain, which limits the proximity between HN_1 and the C-terminal peptide group. The HA protons have large electrostatic and peptide anisotropy shifts,

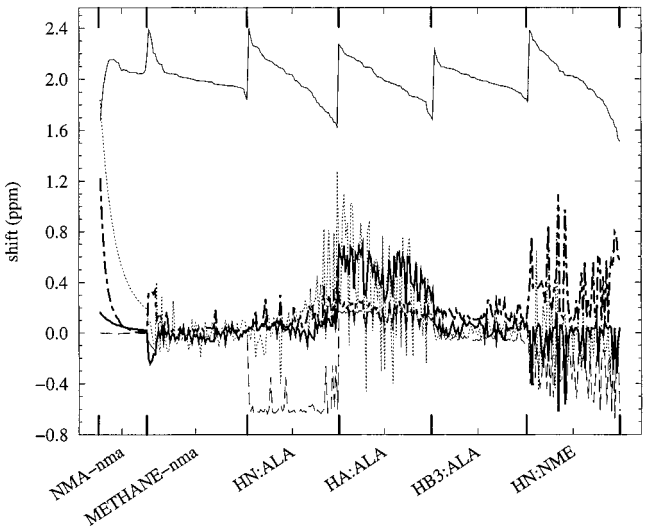

Figure 6. Electrostatic (···), peptide magnetic anisotropy (—), close contact (— · —), and C−C and C−H bond anisotropy (- - -) empirical contributions to proton shifts for the hydrogen bonded proton in the NMA-dimers, the methane protons in the nonlinear methane−NMA complexes, and the HN1, HA, HB, and HN2 protons in the alanine dipeptide conformations. The NMA dimer shifts are in order according to increasing dimer distance; the other protons are sorted within each proton type according to the error in empirically calculated shifts. Errors (empirical − DFT shift + 2.0 ppm) are shown in the figure as a fine solid line.

resulting from the two peptide groups they are sandwiched between. The HB protons are far enough away from the peptide backbone atoms that the electrostatic, anisotropy and close contact contributions are all close to zero. Overall, no major systematic errors in the shift calculations are observed.

4. Discussion

Quantum chemical shielding calculations have recently begun to significantly advance our understanding of the relationship between structure and chemical shifts.[24−26] The quantum chemical approaches offer the opportunity to explore in detail not only local but also environmental shielding effects in a controlled manner and to develop and parameterize empirical models of chemical shifts accordingly. By varying simple geometries of momoners and complexes in the work presented here, we have isolated and explored contributions to proton shifts from close contacts, electrostatics, and peptide and bond magnetic anisotropies. A consistent, physically reasonable set of empirical parameters has been established which describes both the simplest, isolated shifts as well as shifts in more complicated systems including hydrogen-bonded complexes and the alanine dipeptide. The in-depth probing made possible by the quantum chemical calculations should increase the accuracy not only of total shift magnitudes calculated by the empirical method but also of the breakdown into empirical shift contributions as well.

The empirical shift equation and associated parameters developed here yield improved agreement between calculated empirical and DFT proton shifts relative to two previous empirical shift calculation methods[16,17] for the most complicated system explored in this work, the alanine dipeptide molecule. As was mentioned above, this improvement is due in part to the value of the electrostatic parameter *A*, which was smaller in the previous studies, since they were developed to calculate shifts in aqueous protein systems in which electrostatic interactions are screened. A significant improvement in agreement was also gained from the addition of the close contact term

and bond anisotropy terms, which were absent from the previous empirical models. The necessity of including these terms is clear from an examination of the quantum chemically calculated shifts for protons near rare gas atoms and both the experimental and the quantum mechanically computed delta shifts for equatorial versus axial protons in cyclohexane.

The systems studied in this work are small and limited to the gas phase; it is of critical importance to determine whether the models developed here will yield improved shift results for larger biomolecules in aqueous solvents. Preliminary results of applying this new empirical model to shift calculations in proteins, using ring current parameters developed previously in this lab from fits to quantum chemical shifts of protons near aromatic molecules,[26] and a modified electrostatics parameter to account for the screening properties of water suggest that agreement between calculated and experimental backbone shifts is improved over previous empirical models. The explicit treatment of close contact shielding effects and the lone pair charge distribution on O atoms have a significant effect on shift calculations for amide protons, which are often involved in hydrogen bonds, and whose shifts have proven difficult to reproduce in the past.[16,18] We plan to incorporate these methods into a structural refinement program in order to improve derivations of protein structure by more accurately extracting structural information from measured chemical shift data.

Acknowledgment. This work was supported by NIH Grant GM45811. D.S. was supported by NIH fellowship GM18427. The *deMon* density functional programs were kindly supplied by D. R. Salahub, V. G. Malkin, and O. L. Malkina. We thank Annick Dejaegere for helpful discussions throughout the course of this work.

JA9721430