

Experimental, Hartree–Fock, and Density Functional Theory Investigations of the Charge Density, Dipole Moment, Electrostatic Potential, and Electric Field Gradients in L-Asparagine Monohydrate

William D. Arnold,[†] Lori K. Sanders,[†] Michael T. McMahon,[†] Anatoliy V. Volkov,[‡] Guang Wu,[‡] Philip Coppens,[‡] Scott R. Wilson,[†] Nathalie Godbout,[†] and Eric Oldfield*[†]

Contribution from the Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, and Chemistry Department, State University of New York at Buffalo, Buffalo, New York 14260-3000

Received February 1, 2000

Abstract: We have investigated the charge density, $\rho(\mathbf{r})$, its curvature, $\partial^2\rho/\partial r_{ij}$, the dipole moment, μ , and the electrostatic potential, $\Phi(\mathbf{r})$, in L-asparagine monohydrate by using high-resolution single-crystal X-ray crystallography and quantum chemistry. In addition, we have compared electric field gradient, $\nabla\mathbf{E}$, results obtained from crystallography and quantum chemistry with those obtained from single-crystal ^{14}N nuclear magnetic resonance spectroscopy. A multipole model of the X-ray $\rho(\mathbf{r})$ is compared to Hartree–Fock and density functional theory predictions, using two different large basis sets. The quality of the calculated charge densities is evaluated from a simultaneous comparison of eight Hessian-of- $\rho(\mathbf{r})$ tensors at bond critical points between non-hydrogen atoms. These tensors are expressed in an icosahedral representation, which includes information on both tensor magnitude and orientation. The best theory-versus-experiment correlation is found at the B3LYP/6-311++G(2d,2p) level, which yields a slope of 1.09 and an R^2 value of 0.96. Both DFT and HF results give molecular dipole moments in good accord with the value extracted from the X-ray diffraction data, 14.3(3) D, and both sets of calculations are found to correctly reproduce the experimental molecular electrostatic potential, $\Phi(\mathbf{r})$. The intermolecular hydrogen bond $\rho(\mathbf{r})$ is also subjected to a detailed theoretical and experimental topological analysis, and again good agreement is found between theory and experiment. For the comparison of the $\nabla\mathbf{E}$ tensors, the icosahedral representation is again used. There is found to be moderate accord between theory and experiment when using results obtained from diffraction data, but much better accord when using results obtained from NMR data (slope = 1.14, R^2 = 0.94, for the 12 icosahedral tensor elements for N1 and N2). Overall, these results strongly support the idea that both HF and DFT methods give excellent representations of the electrostatic properties $\rho(\mathbf{r})$, $\partial^2\rho/\partial r_{ij}$, μ , $\Phi(\mathbf{r})$, and $\nabla\mathbf{E}$, for crystalline L-asparagine monohydrate, encouraging their future use in situations where experimental results are lacking, such as in peptides and in enzyme active sites.

Introduction

There is currently considerable interest in using quantum chemical methods to investigate structure and bonding in molecules of ever increasing size and to help predict and refine the structures of molecules using spectroscopic observables.¹ In our group at the University of Illinois, we have been using quantum chemical methods to help interpret both isotropic chemical shifts and chemical shift tensors in proteins and model systems, to provide new approaches to protein structure refinement.^{1–4} In the case of $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, and $^{13}\text{C}^\gamma$ shift determinations, we have generally used Hartree–Fock (HF) methods,^{5,6} while in the case of metalloporphyrins, we have used

density functional theory (DFT) methods with hybrid functionals to investigate both metal and ligand shieldings,^{7–10} since these give the best agreement between theory and experiment. And, as a bonus from the SCF part of these calculations, we have access to a large base of electrostatic properties which can be derived at little extra computational cost. The general question then arises: How accurate might these computed electrostatic properties, such as the charge density, $\rho(\mathbf{r})$, its curvature, $\partial^2\rho/\partial r_{ij}$, the dipole moment, μ , the electrostatic potential, $\Phi(\mathbf{r})$, and the electric field gradient, $\nabla\mathbf{E}$, be?

We report here high-resolution single-crystal X-ray diffraction data (obtained by using synchrotron radiation with an area detector) on L-asparagine·H₂O, which contains a hydrogen-bonded amide group, and we investigate the $\rho(\mathbf{r})$, $\partial^2\rho/\partial r_{ij}$, μ , $\Phi(\mathbf{r})$, and $\nabla\mathbf{E}$ values determined both experimentally (from

[†] University of Illinois.

[‡] State University of New York.

(1) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217–225.

(2) deDios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491–1496.

(3) Pearson, J. G.; Wang, J.; Markley, J. L.; Le, H.; Oldfield, E. *J. Am. Chem. Soc.* **1995**, *117*, 8823–8829.

(4) McMahon, M. T.; deDios, A. C.; Godbout, N.; Salzmann, R.; Laws, D. D.; Le, H.; Havlin, R. H.; Oldfield, E. *J. Am. Chem. Soc.* **1998**, *120*, 4784–4797.

(5) Havlin, R. H.; Le, H.; Laws, D. D.; deDios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 11951–11958.

(6) Pearson, J. G.; Le, H.; Sanders, L. K.; Godbout, N.; Havlin, R. H.; Oldfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 11941–11950.

(7) Salzmann, R.; Kaupp, M.; McMahon, M.; Oldfield, E. *J. Am. Chem. Soc.* **1998**, *120*, 4771–4783.

(8) Godbout, N.; Oldfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 8065–8069.

(9) Havlin, R. H.; McMahon, M.; Srinivasan, R.; Le, H.; Oldfield, E. *J. Phys. Chem.* **1997**, *101*, 8908–8913.

(10) Godbout, N.; Havlin, R.; Salzmann, R.; Wojdelski, M.; Debrunner, P. G.; Oldfield, E. *J. Phys. Chem.* **1998**, *102*, 2342–2350.

diffraction data and from NMR, ref 11) and theoretically (as a function of type of theory and basis set size). Luger and co-workers have previously established the transferability of X-ray charge densities and Laplacians among the conserved heavy-atom bonds in asparagine, glutamic acid, serine, threonine, proline, aspartic acid, alanine, and glycine,¹² and theoretical HF results have been compared with X-ray determinations of $\rho(\mathbf{r})$, $\nabla^2\rho(\mathbf{r})$, μ , and $\Phi(\mathbf{r})$ in L-alanine,^{13,14} DL-proline,¹⁵ and DL-aspartic acid.¹⁶ More recently, hybrid functional DFT methods have been applied to the study of DL-histidine, but using a relatively modest basis set.¹⁷ Here, we compare the results of both HF and DFT investigations of these properties, using more extended basis sets in both cases. In addition, we investigate the curvature in $\rho(\mathbf{r})$ at bond critical points, as expressed by the Hessian-of- $\rho(\mathbf{r})$ -tensor, \mathbf{H} . By analogy with previous investigations of chemical shielding tensors, this $\partial^2\rho/\partial r_{ij}$ comparison is expected to be a more rigorous test of the quality of a calculation than $\nabla^2\rho(\mathbf{r})$ alone, given that fortuitous error cancellation is less likely to occur. We also consider both the magnitude and the orientation of the Hessian tensors and the electric field gradient (EFG) tensors using an icosahedral representation, introduced previously for chemical shielding tensors by Grant et al.¹⁸ Moreover, we consider both X-ray and NMR determinations of the EFG tensor to provide the most thorough overall test of the quality of the calculations.

Both X-ray diffraction and quantum chemical determinations of the $\rho(\mathbf{r})$ topology in biomolecules are size-limited, and despite the growing utility of synchrotron radiation in structural biology, it seems that accurate topological charge density analysis in proteins is still intractable. An accurate $\rho(\mathbf{r})$ requires a highly overdetermined, aspherical, anharmonic model, and while focused synchrotron X-ray beams have allowed atomic resolution for several protein structures,^{19,20} radiation damage may complicate the collection of a sufficient number of reflections for a full topological $\rho(\mathbf{r})$ analysis. Consequently, quantum chemical methods may begin to play an important role, since although quantum chemical calculations also suffer severe size restrictions, they have been shown in many cases to accurately reproduce measured, spectroscopic observables in quite large systems.² By analogy then, it seems reasonable to believe that important local electrostatic structural information can be deduced from large cluster calculations, so that topics such as ligand stabilization via hydrogen bonding in heme proteins, or active-site transition-state stabilization by local electrostatic field effects in enzymes, can be investigated.²¹ But first, we need to find the best way to compute and measure electrostatic properties such as $\rho(\mathbf{r})$, $\nabla^2\rho(\mathbf{r})$, $\partial^2\rho/\partial r_{ij}$, μ , $\Phi(\mathbf{r})$, and $\nabla\mathbf{E}$ in model systems.

Table 1. Crystallographic Data for L-Asparagine Monohydrate

| | |
|---|--|
| formula | C ₄ H ₁₀ N ₂ O ₄ |
| formula weight | 150.14 |
| crystal system | orthorhombic |
| space group | P2 ₁ 2 ₁ 2 ₁ |
| crystal size, color | 0.06 × 0.08 × 0.1 mm ³ , colorless |
| <i>a</i> | 5.583(1) Å |
| <i>b</i> | 9.738(1) Å |
| <i>c</i> | 11.693(1) Å |
| $\alpha = \beta = \gamma$ | 90° |
| <i>Z</i> | 4 |
| volume | 635.72 Å ³ |
| density | 1.569 g cm ⁻³ |
| <i>T</i> | 20(1) K |
| λ | 0.643 Å |
| ($\sin \theta/\lambda$) _{max} | 1.07 Å ⁻¹ |
| reflections collected | 41 609 |
| unique reflections | 6363 |
| reflections included in refinement | 4911 |
| <i>R</i> _{int} | 0.037 |
| <i>R</i> [<i>F</i>] | 0.0204 |
| <i>R</i> _w [<i>F</i>] | 0.0262 |
| GOF | 0.9 |
| <i>N</i> _{var} | 201 |
| <i>N</i> _{ref} / <i>N</i> _{var} | 24.4 |

Experimental Section

A sample of L-asparagine monohydrate (Sigma, St. Louis, MO) was recrystallized from hot water, and single-crystal X-ray diffraction data were collected at the SUNY X3A1 beamline at the National Synchrotron Light Source, Brookhaven National Laboratory. X-ray crystallographic data for L-asparagine monohydrate are summarized in Table 1. Using a sideways-reflecting curved Si(111) monochromator, a wavelength of 0.643 Å was obtained. A well-faceted asparagine monohydrate crystal with dimensions 0.06 × 0.08 × 0.1 mm³ was glued onto the tip of an amorphous carbon fiber, which was then mounted on the coldfinger of a two-stage close-cycle helium DISPLEX CT211 cryostat. The cryostat was mounted on the φ -table of a HUBER D-511.1 four-circle diffractometer,²² equipped with a Bruker SMART 1000 CCD area detector. A special "antiscattering" device²³ was mounted inside the chamber of the cryostat in order to reduce the scattering of the direct beam by the graphite walls of the vacuum chamber. All measurements were obtained at 20(1) K. The detector surface was located 4.7 cm from the crystal, which is the shortest distance allowed by the vacuum chamber.

Six data sets were collected at two different orientations of the crystal using three different settings (0°, 35°, and 55°) of the detector 2 θ arm. The data were collected by rotation of the φ axis from 0 to 390° (orientation 1) and to 420° (orientation 2) with an interval of 0.3° per frame. The exposure times were 5, 7, and 20 s for 0°, 35°, and 55° 2 θ settings, respectively. The unit cell parameters (*a* = 5.583(1) Å, *b* = 9.738(1) Å, *c* = 11.693(1) Å, $\alpha = \beta = \gamma = 90^\circ$) were determined using 692 reflections. The intensities were integrated with the SAINT software package.²⁴ During the integration, the orientation matrix was optimized after every 50 frames. The incident beam intensity was normalized on the basis of the counts from a beam monitor placed behind the beam-defining slits. No absorption correction was applied because of the small size of the specimen and the low absorption coefficient.

The 41 609 symmetry-equivalent and redundant measurements (0.099 < $\sin \theta/\lambda$ < 1.073 Å⁻¹) were averaged in the program SORTAV.²⁵ The sharp vertical profile of the synchrotron beam together with a slight instability of a cryostat necessitated a special φ -correction procedure.²⁶ After correction, averaging in point group 222 resulted in 6363 unique

(22) Graafsma, H.; Sagerman, G.; Coppens, P. *J. Appl. Crystallogr.* **1991**, *24*, 961–962.

(23) Darovsky, A.; Bolotovskiy, R.; Coppens, P. *J. Appl. Crystallogr.* **1994**, *27*, 1039–1040.

(24) Siemens. *SAINT. Program to Integrate and Reduce Raw Crystallographic Area Detector Data*; Siemens Analytical X-ray Instruments Inc.: Madison, WI, 1996.

(25) Blessing, R. H. *J. Appl. Crystallogr.* **1997**, *30*, 421–426.

(11) Naito, A.; McDowell, C. A. *J. Chem. Phys.* **1984**, *81*, 4795–4803.

(12) Flaig, R.; Koritsanszky, T.; Janczak, J.; Krane, H.-G.; Morgenroth, W.; Luger, P. *Angew. Chem., Int. Ed.* **1999**, *38*, 1397–1400.

(13) Gatti, C.; Bianchi, R.; Destro, R.; Merati, F. *J. Mol. Struct.* **1992**, *255*, 409–433.

(14) Destro, R.; Marsh, R.; Bianchi, R. *J. Phys. Chem.* **1988**, *92*, 966–973.

(15) Koritsanszky, T.; Flaig, R.; Zobel, D.; Krane, H.-G.; Morgenroth, W.; Luger, P. *Science* **1998**, *279*, 356–358.

(16) Flaig, R.; Koritsanszky, T.; Zobel, D.; Luger, P. *J. Am. Chem. Soc.* **1998**, *120*, 2227–2238.

(17) Coppens, P.; Abramov, Y.; Carducci, M.; Korjov, B.; Novozhilova, I.; Alhambra, C.; Pressprich, M. *J. Am. Chem. Soc.* **1999**, *121*, 2585–2593.

(18) Alderman, D. W.; Sherwood, M. H.; Grant, D. M. *J. Magn. Reson.* **1993**, *A101*, 188–197.

(19) Wilson, K. S. *Nature Struct. Biol.* **1998**, *5*, 627–630.

(20) Cusack, S.; Helrhali, H.; Bram, A.; Burghammer, M.; Perrakes, A.; Riek, C. *Nature Struct. Biol.* **1998**, *5*, 634–637.

(21) Warshel, A.; Creighton, S. In *Computer Simulation of Biomolecular Systems*; van Gunsteren, W. F.; Weiner, P. K., Eds.; ESCOM: Leiden, 1989; pp 120–138.

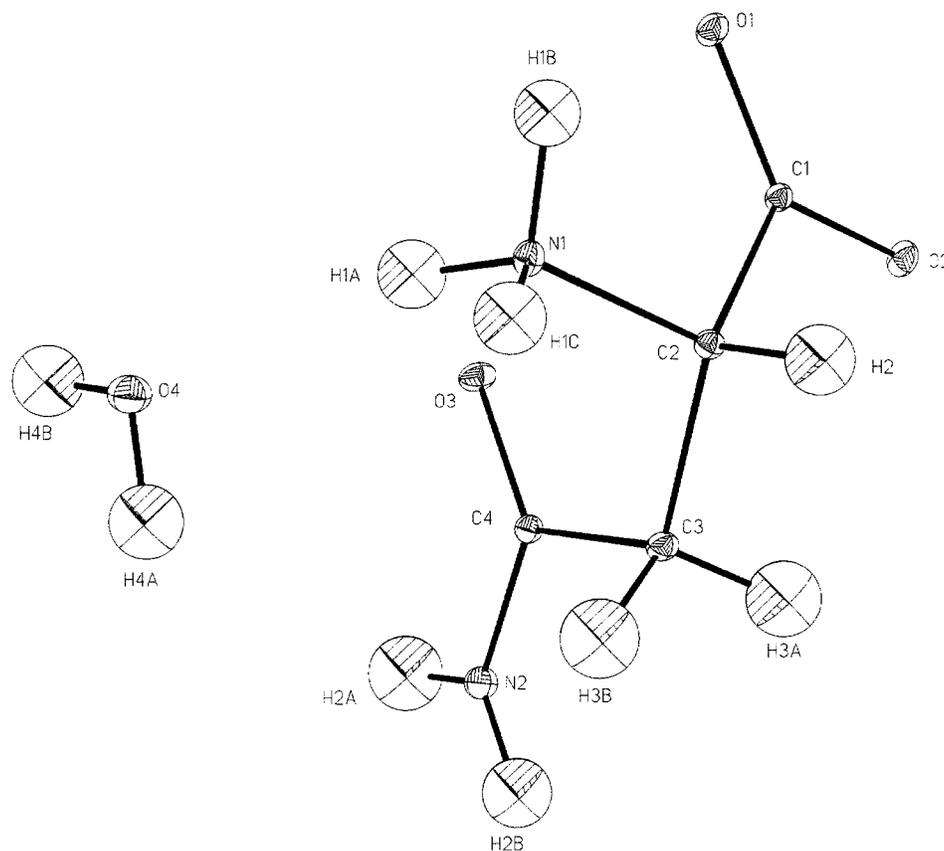


Figure 1. SHELXL drawing of L-asparagine monohydrate showing the numbering scheme used in this work. Thermal ellipsoids are drawn at 50% probability.

reflections with an average multiplicity of 6.5 per unique reflection and a final *R*-merge factor of 3.7%. Only 5231 unique reflections measured three or more times were used in subsequent analysis.

Conventional IAM refinement was performed in SHELXL-97²⁷ based on F^2 using all unique reflections. Comparison of F_{obs} and F_{calc} for strong reflections showed extinction not to be a factor, as was expected given the small crystal size and the relatively short wavelength used.

The aspherical atom refinement was carried out using the XD package²⁸ and was based on F . Only 4911 reflections with $F_o > 3\sigma(F_o)$ were used. The aspherical atom model implemented in XD is based on the Hansen–Coppens multipole formalism.²⁹ The formalism describes the static electron density in the crystal by a superposition of aspherical *pseudoatoms*, the charge density of which is modeled by a nucleus-centered multipole expansion,

$$\rho_{\mathbf{k}}(\mathbf{r}) = P_c \rho_c(r) + P_v \kappa^3 \rho_v(\kappa r) + \kappa' \sum_{l=1}^4 R_l(\kappa' r) \sum_{m=1}^l P_{lm\pm} d_{lm\pm}(\mathbf{r}/r)$$

where ρ_c and ρ_v are spherically averaged free-atom Hartree–Fock core and valence densities normalized to one electron, $d_{lm\pm}$ are real spherical harmonic angular functions, R_l are normalized Slater-type radial functions, and κ and κ' are dimensionless expansion–contraction parameters, which can be refined in the fitting procedure along with the populations P_v and $P_{lm\pm}$. HF densities are used for the spherically averaged core (ρ_c) and valence (ρ_v) shells.³⁰ The default conventional

sets^{28,31} of R_l functions were used for all pseudoatoms. κ and κ' parameters of the hydrogen atoms were fixed at a value of 1.2 throughout all refinements. To reduce the number of parameters, the multipole coefficients of the non-hydrogen atoms were constrained to obey local *mm2* symmetry for C(1) and N(2) atoms, 3-fold symmetry (3) for the N(1) atom, and mirror-plane symmetry (*m*) for C(3), C(4), O(4), and O(3) (Figure 1). No local symmetry constraints were applied to the C(2) atom (α carbon). All hydrogen atoms were given cylindrical symmetry. Chemical equivalency constraints were applied to atoms O(1) and O(2); hydrogens H(1A), H(1B), and H(1C); hydrogens H(3A) and H(3B); hydrogens H(2A) and H(2B); and hydrogens H(4A) and H(4B). Finally, a molecular electroneutrality constraint was applied separately to the asparagine molecule and to the water molecule.

In the multipole refinement, the procedure recently described by Abramov et al. was followed.³² In the first stage, high-order refinements ($\sin \theta/\lambda > 0.7 \text{ \AA}^{-1}$) were performed to determine unbiased positional and thermal parameters for the non-hydrogen atoms. Hydrogen positions were obtained by extending X–H distances to their standard neutron diffraction values ($C_{\text{primary}}\text{–H} = 1.092 \text{ \AA}$, $C_{\text{secondary}}\text{–H} = 1.099 \text{ \AA}$, $N_{\text{NH}_3}\text{–H} = 1.035 \text{ \AA}$, $N_{\text{NH}_2}\text{–H} = 1.010 \text{ \AA}$, and $O_{\text{water}}\text{–H} = 0.960 \text{ \AA}$). These distances were kept constant in subsequent refinements. In the next stage a κ -refinement ($P_{lm\pm} = 0$) was performed with all structural parameters, except the isotropic thermal parameters of the hydrogens, being fixed at the previously refined values. Finally, a κ' -restricted multipole model (KRMM) refinement³² was carried out. In this type of refinement the κ' parameters are fixed at values derived from multipole refinements of theoretical structure factors, obtained from periodic Hartree–Fock (PHF) calculations at the 6-31G** level on a series of organic compounds.³³ The multipole expansion was truncated at the octupole level ($l_{\text{max}} = 3$) for the non-hydrogen atoms and at the quadrupole level ($l_{\text{max}} = 2$) for the hydrogens. The resulting multipole

(26) Volkov, A.; Wu, G.; Coppens, P. *J. Synchrotron Rad.* **1999**, *6*, 1007–1015.

(27) Sheldrick, G. M. *SHELXL97. Program for Crystal Structure Refinement*; University of Göttingen: Göttingen, Germany, 1997.

(28) Koritsanzky, T.; Howard, S.; Su, Z.; Mallinson, P. R.; Richter, T.; Hansen, N. K. *XD. Computer Program Package for Multipole Refinement and Analysis of Electron Densities from Diffraction Data*; Free University of Berlin: Berlin, Germany, 1997.

(29) Hansen, N. K.; Coppens, P. *Acta Crystallogr.* **1978**, *A34*, 909–921.

(30) Clementi, E.; Roetti, C. *At. Data Nucl. Data Tables* **1974**, *14*, 177–478.

(31) Coppens, P. *X-ray Charge Densities and Chemical Bonding*; Oxford University Press: New York, 1997.

(32) Abramov, Y.; Volkov, A.; Coppens, P. *Chem. Phys. Lett.* **1999**, *311*, 81–86.

(33) Volkov, A.; Abramov, Y.; Coppens, P. Manuscript in preparation.

Table 2. Topological Properties at Heavy-Atom (3,−1) Bond Critical Points

| bonded pair | $\rho(\mathbf{r}_b)$ (e \AA^{-3}) | $\nabla^2\rho(\mathbf{r}_b)$ (e \AA^{-5}) | λ_1 (e \AA^{-5}) | λ_2 (e \AA^{-5}) | λ_3 (e \AA^{-5}) | method |
|-------------|---|---|------------------------------------|------------------------------------|------------------------------------|-------------------------|
| C1–O1 | 2.69(2) | −26.6(1) | −23.72 | −21.92 | 19.00 | X-ray |
| | 2.61 | −17.74 | −27.34 | −25.21 | 34.81 | HF/6-311++G(2d,2p) |
| | 2.64 | −15.62 | −28.83 | −26.63 | 39.85 | HF/6-311++G(3df,3pd) |
| | 2.56 | −18.71 | −24.82 | −22.29 | 28.39 | B3LYP/6-311++G(2d,2p) |
| | 2.59 | −17.15 | −26.20 | −23.49 | 32.54 | B3LYP/6-311++G(3df,3pd) |
| C1–O2 | 2.63(2) | −29.6(1) | −22.65 | −20.79 | 13.82 | X-ray |
| | 2.69 | −22.10 | −25.69 | −23.48 | 27.07 | HF/6-311++G(2d,2p) |
| | 2.72 | −19.70 | −27.00 | −24.75 | 32.05 | HF/6-311++G(3df,3pd) |
| | 2.64 | −21.94 | −23.24 | −20.66 | 21.96 | B3LYP/6-311++G(2d,2p) |
| | 2.68 | −20.25 | −24.46 | −21.71 | 25.92 | B3LYP/6-311++G(3df,3pd) |
| C1–C2 | 1.67(2) | −10.55(4) | −11.25 | −10.12 | 10.82 | X-ray |
| | 1.73 | −16.61 | −13.16 | −12.35 | 8.91 | HF/6-311++G(2d,2p) |
| | 1.77 | −18.53 | −13.54 | −12.72 | 7.73 | HF/6-311++G(3df,3pd) |
| | 1.64 | −12.70 | −12.06 | −11.24 | 10.59 | B3LYP/6-311++G(2d,2p) |
| | 1.67 | −14.04 | −12.35 | −11.48 | 9.79 | B3LYP/6-311++G(3df,3pd) |
| C2–C3 | 1.62(3) | −10.01(5) | −10.45 | −10.13 | 10.57 | X-ray |
| | 1.73 | −15.67 | −12.33 | −11.96 | 8.61 | HF/6-311++G(2d,2p) |
| | 1.77 | −17.34 | −12.57 | −12.23 | 7.46 | HF/6-311++G(3df,3pd) |
| | 1.66 | −12.87 | −11.60 | −11.24 | 9.97 | B3LYP/6-311++G(2d,2p) |
| | 1.69 | −14.11 | −11.77 | −11.45 | 9.11 | B3LYP/6-311++G(3df,3pd) |
| C3–C4 | 1.71(3) | −12.27(5) | −10.87 | −12.09 | 10.69 | X-ray |
| | 1.79 | −17.10 | −13.36 | −12.63 | 8.89 | HF/6-311++G(2d,2p) |
| | 1.83 | −18.92 | −13.70 | −12.89 | 7.68 | HF/6-311++G(3df,3pd) |
| | 1.70 | −13.73 | −12.43 | −11.70 | 10.40 | B3LYP/6-311++G(2d,2p) |
| | 1.74 | −15.08 | −12.72 | −11.89 | 9.53 | B3LYP/6-311++G(3df,3pd) |
| C2–N1 | 1.67(3) | −8.64(8) | −9.40 | −10.71 | 11.47 | X-ray |
| | 1.55 | −10.60 | −9.43 | −8.65 | 7.48 | HF/6-311++G(2d,2p) |
| | 1.57 | −8.74 | −9.85 | −8.98 | 10.09 | HF/6-311++G(3df,3pd) |
| | 1.57 | −13.13 | −10.19 | −9.71 | 6.78 | B3LYP/6-311++G(2d,2p) |
| | 1.59 | −13.16 | −10.35 | −9.78 | 6.97 | B3LYP/6-311++G(3df,3pd) |
| C4–N2 | 2.27(3) | −27.2(1) | −18.88 | −16.35 | 8.04 | X-ray |
| | 2.34 | −28.86 | −21.09 | −19.58 | 11.81 | HF/6-311++G(2d,2p) |
| | 2.37 | −26.80 | −22.01 | −20.46 | 15.66 | HF/6-311++G(3df,3pd) |
| | 2.30 | −27.31 | −19.66 | −17.16 | 9.51 | B3LYP/6-311++G(2d,2p) |
| | 2.32 | −26.90 | −20.26 | −17.66 | 11.02 | B3LYP/6-311++G(3df,3pd) |
| C4–O3 | 2.69(3) | −25.4(2) | −21.77 | −23.09 | 19.32 | X-ray |
| | 2.73 | −19.53 | −27.05 | −24.84 | 32.37 | HF/6-311++G(2d,2p) |
| | 2.77 | −17.09 | −28.53 | −26.25 | 37.69 | HF/6-311++G(3df,3pd) |
| | 2.69 | −20.43 | −24.69 | −22.30 | 26.53 | B3LYP/6-311++G(2d,2p) |
| | 2.72 | −18.68 | −26.08 | −23.53 | 30.93 | B3LYP/6-311++G(3df,3pd) |

population and radial screening parameters are summarized in Table S1 (Supporting Information). Results of the Hirshfeld rigid bond test³⁴ show excellent agreement between mean-square displacement amplitudes for bonded non-hydrogen atoms, with a maximum discrepancy of 0.0007 \AA found for the conventional IAM refinement and 0.0003 \AA for the final KRMM refinement (Table S2, Supporting Information).

Charge density and electric field gradient tensor calculations were performed by using the Gaussian 94 and Gaussian 98 programs^{35,36} on a cluster of Silicon Graphics/Cray Origin 200 computers in our

laboratory at the University of Illinois, and on Origin-2000 and Exemplar clusters at the National Center for Supercomputing Applications, located in Urbana, IL. Unless otherwise noted, the experimental X-ray geometry of L-asparagine monohydrate determined in this work was employed. Both Hartree–Fock (HF) and density functional theory (DFT) methods were used, with the uniform basis sets 6-311++G(2d,2p) and 6-311++G(3df,3pd). DFT calculations employed the Becke 3 parameter hybrid exchange functional³⁷ and the LYP correlation functional,³⁸ B3LYP. Topological analyses of the calculated $\rho(\mathbf{r})$ were carried out by using Bader's AIMPAC program suite,³⁹ while the corresponding experimental (X-ray) properties were extracted from the diffraction data by using the XD program package.²⁸ Electric field gradient tensors were obtained from the X-ray diffraction data using the Molprop93 program.⁴⁰

Results and Discussion

Associated with every chemical bond is a point, \mathbf{r}_b , at which the first derivative of the charge density, $\rho(\mathbf{r})$, is zero.⁴¹ At this so-called bond critical point (BCP) there are three nonzero principal curvatures in $\rho(\mathbf{r}_b)$: two negative and one positive. Such a BCP is classified as (3,−1): 3 for the number of nonzero

(34) Hirshfeld, F. L. *Acta Crystallogr.* **1976**, *A32*, 239–244.

(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; 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 C.2; Gaussian, Inc.: Pittsburgh, PA, 1995.

(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(37) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785–789.

(39) Bader, R. F. W. *AIMPAC Program Package*; McMaster University: Hamilton, ON, Canada, 1998.

(40) Su, Z.; Coppens, P. *Acta Crystallogr.* **1992**, *A48*, 188–197.

(41) Bader, R. F. W. *Atoms in Molecules—A Quantum Theory*; Clarendon Press: Oxford, 1990.

Table 3. Least-Squares Fit Statistics for Theory-versus-Experiment Property Correlations

| | slope | intercept | R^2 value | method |
|---|-------|-----------|-------------|-------------------------|
| heavy-atom bond | 0.98 | 0.08 | 0.98 | HF/6-311++G(2d,2p) |
| charge density | 0.97 | 0.12 | 0.97 | HF/6-311++G(3df,3pd) |
| | 0.99 | 0.00 | 0.99 | B3LYP/6-311++G(2d,2p) |
| | 0.99 | 0.02 | 0.99 | B3LYP/6-311++G(3df,3pd) |
| heavy-atom bond | 1.26 | 1.71 | 0.95 | HF/6-311++G(2d,2p) |
| principal curvatures | 1.37 | 2.65 | 0.93 | HF/6-311++G(3df,3pd) |
| | 1.12 | 1.16 | 0.97 | B3LYP/6-311++G(2d,2p) |
| | 1.20 | 1.74 | 0.96 | B3LYP/6-311++G(3df,3pd) |
| heavy-atom bond | 1.21 | 1.43 | 0.93 | HF/6-311++G(2d,2p) |
| Hessian tensors ^a | 1.31 | 2.27 | 0.89 | HF/6-311++G(3df,3pd) |
| | 1.09 | 0.96 | 0.96 | B3LYP/6-311++G(2d,2p) |
| | 1.16 | 1.46 | 0.93 | B3LYP/6-311++G(3df,3pd) |
| hydrogen bond | 0.73 | 0.07 | 0.69 | HF/6-311++G(2d,2p) |
| charge density | 0.85 | 0.06 | 0.77 | HF/6-311++G(3df,3pd) |
| | 0.86 | 0.07 | 0.78 | B3LYP/6-311++G(2d,2p) |
| | 0.91 | 0.07 | 0.76 | B3LYP/6-311++G(3df,3pd) |
| hydrogen bond | 0.92 | -0.16 | 1.00 | HF/6-311++G(2d,2p) |
| principal curvatures | 0.98 | -0.22 | 1.00 | HF/6-311++G(3df,3pd) |
| | 0.94 | -0.22 | 1.00 | B3LYP/6-311++G(2d,2p) |
| | 0.99 | -0.27 | 1.00 | B3LYP/6-311++G(3df,3pd) |
| hydrogen bond | 0.92 | -0.16 | 0.99 | HF/6-311++G(2d,2p) |
| Hessian tensors ^a | 0.97 | -0.21 | 0.99 | HF/6-311++G(3df,3pd) |
| | 0.93 | -0.22 | 0.99 | B3LYP/6-311++G(2d,2p) |
| | 0.98 | -0.27 | 0.99 | B3LYP/6-311++G(3df,3pd) |
| EFG tensor | 2.04 | 0.00 | 0.85 | HF/6-311++G(2d,2p) |
| principal values | 1.99 | 0.00 | 0.84 | HF/6-311++G(3df,3pd) |
| | 1.85 | 0.00 | 0.84 | B3LYP/6-311++G(2d,2p) |
| | 1.81 | 0.00 | 0.83 | B3LYP/6-311++G(3df,3pd) |
| EFG tensor ^a | 1.88 | 0.00 | 0.72 | HF/6-311++G(2d,2p) |
| | 1.85 | 0.00 | 0.73 | HF/6-311++G(3df,3pd) |
| | 1.73 | 0.00 | 0.74 | B3LYP/6-311++G(2d,2p) |
| | 1.71 | 0.00 | 0.74 | B3LYP/6-311++G(3df,3pd) |
| EFG tensor ^a | 1.35 | 0.00 | 0.79 | HF/6-311++G(2d,2p) |
| (theory vs ¹⁴ N NMR ^b) | 1.32 | 0.00 | 0.77 | HF/6-311++G(3df,3pd) |
| | 1.28 | 0.00 | 0.81 | B3LYP/6-311++G(2d,2p) |
| | 1.27 | 0.00 | 0.79 | B3LYP/6-311++G(3df,3pd) |
| EFG tensor ^a | 1.22 | 0.00 | 0.93 | HF/6-311++G(2d,2p) |
| (theory vs ¹⁴ N NMR, ^b | 1.19 | 0.00 | 0.91 | HF/6-311++G(3df,3pd) |
| point charges) | 1.14 | 0.00 | 0.94 | B3LYP/6-311++G(2d,2p) |
| | 1.13 | 0.00 | 0.93 | B3LYP/6-311++G(3df,3pd) |

^a Icosahedral representation.¹⁸ ^b Reference 11.

principal curvatures in $\rho(\mathbf{r}_b)$, and -1 for the algebraic sum of curvature signs. Experimental and calculated electron densities at the $(3,-1)$ bond critical points between heavy atoms in L-asparagine are shown in Table 2. The X-ray charge density topologies are subject to both the experimental conditions and the data analysis methodology.¹⁶ Nevertheless, we find good overall accord between our experimental $\rho(\mathbf{r}_b)$ and $\nabla^2\rho(\mathbf{r}_b)$ and those previously reported for five bonds at 100 K.¹² A direct comparison of both $\rho(\mathbf{r}_b)$ and $\nabla^2\rho(\mathbf{r}_b)$ together produces a correlation coefficient of 98% between the two data sets.

The two basis sets we used for the calculations differ only in the polarization functions, with the more extended basis adding f functions to heavy atoms and d functions to hydrogens. Both basis sets have been used previously in Hartree-Fock calculations to predict the topology of $\rho(\mathbf{r})$ in amino acids.¹⁵⁻¹⁷ Our results show that the correlation between theoretical and experimental $\rho(\mathbf{r}_b)$ is largely insensitive to both the theoretical method used and the basis sets employed, with both HF and DFT methods producing approximately 1:1 correlations, as summarized in Table 3. For the HF comparison, we find a slope of 0.98 and an R^2 value of 0.98 for the smaller basis calculation and a slope of 0.97 and an R^2 value of 0.97 for the larger basis calculation, while with the DFT method we find a slope of 0.99 and an R^2 value of 0.99 for both basis sets. This can be judged to be very good accord between theory and experiment,

consistent with previous work on DL-aspartic acid and DL-proline,^{15,16} where a root-mean-square deviation (rmsd) between theoretical and experimental $\rho(\mathbf{r})$ values of $0.12 \text{ e } \text{\AA}^{-3}$ was found for DL-aspartic acid (HF, 6-311++G(3df, 3pd)) and $0.08 \text{ e } \text{\AA}^{-3}$ for DL-proline (HF, 6-311++G(3df,3pd)), to be compared with our rmsds of $0.09 \text{ e } \text{\AA}^{-3}$ for the 6-311++G(3df,3pd) HF calculation and $0.04 \text{ e } \text{\AA}^{-3}$ for the same basis DFT calculation. For the smaller basis calculations, 6-311++G(2d,2p), the rmsd was $0.07 \text{ e } \text{\AA}^{-3}$ for the HF calculation and $0.05 \text{ e } \text{\AA}^{-3}$ for the DFT calculation. There is, therefore, very good agreement between theory and experiment for both HF and DFT methods, and both basis sets. This may suggest, however, that $\rho(\mathbf{r})$ is simply not a particularly sensitive property with which to test the quality of a calculation. We therefore next investigated the curvatures in $\rho(\mathbf{r})$ at bond critical points, $\partial^2\rho/\partial r_{ij}$, to see to what extent the experimental values can be evaluated theoretically.

The curvature in the charge density at a BCP is described by a real symmetric tensor known as the Hessian-of- $\rho(\mathbf{r})$, \mathbf{H} , which has nine elements of the form $\partial^2\rho/r_i r_j$.⁴¹ When diagonalized, \mathbf{H} is expressed by three elements, λ_{1-3} , which correspond to curvatures in $\rho(\mathbf{r}_b)$ along three principal axes. Table 2 lists these principal $\rho(\mathbf{r}_b)$ curvatures at the $(3,-1)$ bond critical points between heavy atoms in L-asparagine·H₂O. The Laplacian of the electron density at a BCP, $\nabla^2\rho(\mathbf{r}_b)$, is simply the sum of these three principal curvatures, and for a rigorous comparison

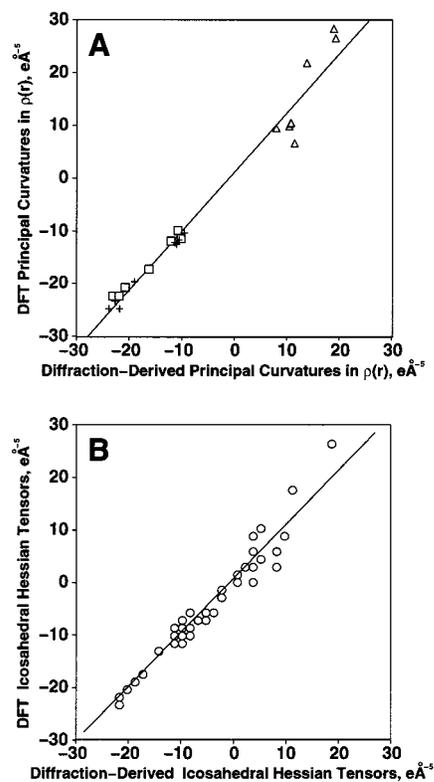


Figure 2. Curvatures in $\rho(\mathbf{r}_b)$ at (3,-1) BCPs along non-hydrogen atom bonds (A) Principal curvatures in $\rho(\mathbf{r}_b)$: B3LYP 6-311++G-(2d,2p) versus X-ray diffraction. + = λ_1 , \square = λ_2 , \triangle = λ_3 . (B) Icosahedral Hessian-of- $\rho(\mathbf{r}_b)$ tensors: B3LYP 6-311++G(2d,2p) versus X-ray diffraction.

between theory and experiment, it is clearly desirable to be able to predict not only $\nabla^2\rho(\mathbf{r}_b)$ but each component curvature as well. This situation is quite analogous to that found in calculations of chemical shielding or chemical shifts, in which the isotropic shielding is simply one-third the sum of the three principal values of the shielding tensor. It is well known that theory-versus-experiment comparisons between individual shielding tensor elements, rather than their sum, tends to preclude fortuitous cancellations of error. The most reliable theoretical method, therefore, is likely to be that which most accurately reproduces all three principal curvatures, rather than simply the Laplacian.

We therefore investigated the slope and R^2 values between the experimentally determined principal curvatures in $\rho(\mathbf{r})$ and those predicted by quantum chemistry, using HF and DFT methods, with both the 6-311G++(2d,2p) and 6-311++G-(3df,3dp) basis sets (Table 2 and Figure 2A). Using the results presented in Table 2, we obtain (Table 3) the following slope and R^2 values: 1.26/0.95 (HF, small basis); 1.37/0.93 (HF large basis); 1.12/0.97 (DFT small basis); and 1.20/0.96 (DFT, large basis). While all four sets of calculations are very respectable, marginally better agreement between theory and the diffraction experiment is found for the DFT results, with the B3LYP/6-311++G(2d,2p) basis set calculation, shown graphically in Figure 2A, giving the best result.

These principal curvatures in $\rho(\mathbf{r})$ represent only the magnitudes of the Hessian tensors, \mathbf{H} , however. Using the icosahedral representation described by Grant et al.,¹⁸ we therefore also compared \mathbf{H} tensor orientations, which correspond to the directions of principal $\rho(\mathbf{r})$ curvature. Here, the components of the icosahedral tensor, χ_{1-6} , can be defined from a tensor in

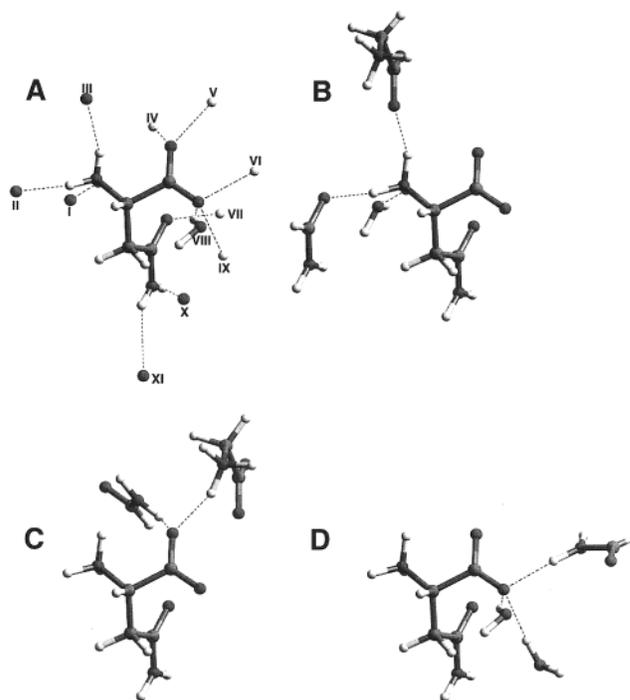


Figure 3. Hydrogen bonds in L-asparagine monohydrate. (A) The complete hydrogen-bonding environment of a single asparagine molecule: I, H1A-O4; II, H1C-O3; III, H1B-O1; IV, O1-H2B; V, O1-H1B; VI, O2-H2A; VII, O3-H1C; VIII, O2-H4A; IX, O2-H4B; X, H2A-O2; XI, H2B-O1. (B) The quantum chemical model used for hydrogen bonds involving H1A-C. (C) The model used for hydrogen bonds involving O1. (D) The model used for hydrogen bonds involving O2.

either the Cartesian, \mathbf{C} , or the principal value/principal axis representation, \mathbf{P} :

$$\chi_1 = a^2C_{xx} + b^2C_{yy} - 2abC_{xy} = (a_1 - bm_1)^2P_{11} + (a_2 - bm_2)^2P_{22} + (a_3 - bm_3)^2P_{33}$$

$$\chi_2 = a^2C_{xx} + b^2C_{yy} + 2abC_{xy} = (a_1 + bm_1)^2P_{11} + (a_2 + bm_2)^2P_{22} + (a_3 + bm_3)^2P_{33}$$

$$\chi_3 = a^2C_{yy} + b^2C_{zz} - 2abC_{yz} = (am_1 - bn_1)^2P_{11} + (am_2 - bn_2)^2P_{22} + (am_3 - bn_3)^2P_{33}$$

$$\chi_4 = a^2C_{yy} + b^2C_{zz} + 2abC_{yz} = (am_1 + bn_1)^2P_{11} + (am_2 + bn_2)^2P_{22} + (am_3 + bn_3)^2P_{33}$$

$$\chi_5 = a^2C_{zz} + b^2C_{xx} - 2abC_{zx} = (an_1 - bl_1)^2P_{11} + (an_2 - bl_2)^2P_{22} + (an_3 - bl_3)^2P_{33}$$

$$\chi_6 = a^2C_{zz} + b^2C_{xx} + 2abC_{zx} = (an_1 + bl_1)^2P_{11} + (an_2 + bl_2)^2P_{22} + (an_3 + bl_3)^2P_{33}$$

where $a = 0.8507$, $b = 0.5257$, and l_i , m_i , and n_i are elements of the $(l, m, n)^T$ direction cosine for the i th principal axis.¹⁸

Tensors are commonly expressed in their principal axis systems (PAS) as three principal values and corresponding direction cosines with respect to a molecular reference frame. Thus, to compare the orientations of such tensors one must compare the angles between each PAS and the molecular

Table 4. Topological Properties at Hydrogen Bond (3,−1) Critical Points

| | bond pair | $\rho(\mathbf{r}_b)$ ($e \text{ \AA}^{-3}$) | $\nabla^2\rho(\mathbf{r}_b)$ ($e \text{ \AA}^{-5}$) | λ_1 ($e \text{ \AA}^{-5}$) | λ_2 ($e \text{ \AA}^{-5}$) | λ_3 ($e \text{ \AA}^{-5}$) | method | |
|-------------------|------------------|---|---|--------------------------------------|--------------------------------------|--------------------------------------|-------------------------|-------------------------|
| NH ₃ | N1—H1A...O4 | 0.15(2) | 2.89(1) | −0.74 | −0.68 | 4.31 | X-ray diffraction | |
| | | 0.16 | 2.32 | −0.78 | −0.73 | 3.83 | HF/6-311++G(2d,2p) | |
| | | 0.16 | 2.41 | −0.82 | −0.77 | 4.00 | HF/6-311++G(3df,3pd) | |
| | | 0.17 | 2.19 | −0.83 | −0.78 | 3.80 | B3LYP/6-311++G(2d,2p) | |
| | | 0.17 | 2.27 | −0.87 | −0.82 | 3.96 | B3LYP/6-311++G(3df,3pd) | |
| | N1—H1B...O1 | 0.173(0) | 3.433(0) | −0.96 | −0.91 | 5.30 | X-ray diffraction | |
| | | 0.22 | 2.63 | −1.00 | −0.99 | 4.62 | HF/6-311++G(2d,2p) | |
| | | 0.20 | 2.57 | −1.13 | −1.11 | 4.81 | HF/6-311++G(3df,3pd) | |
| | | 0.21 | 2.47 | −1.09 | −1.08 | 4.64 | B3LYP/6-311++G(2d,2p) | |
| | | 0.22 | 2.44 | −1.21 | −1.19 | 4.84 | B3LYP/6-311++G(3df,3pd) | |
| | N1—H1C...O3 | 0.193(1) | 3.79(2) | −1.10 | −1.09 | 5.98 | X-ray diffraction | |
| | | 0.19 | 3.05 | −1.20 | −1.17 | 5.42 | HF/6-311++G(2d,2p) | |
| | | 0.23 | 2.92 | −1.38 | −1.33 | 5.63 | HF/6-311++G(3df,3pd) | |
| | | 0.24 | 2.84 | −1.32 | −1.27 | 5.43 | B3LYP/6-311++G(2d,2p) | |
| | | 0.25 | 2.74 | −1.48 | −1.42 | 5.64 | B3LYP/6-311++G(3df,3pd) | |
| CONH ₂ | N2—H2A...O2 | 0.11(2) | 2.65(1) | −0.70 | −0.68 | 4.03 | X-ray diffraction | |
| | | 0.17 | 2.28 | −0.86 | −0.82 | 3.96 | HF/6-311++G(2d,2p) | |
| | | 0.18 | 2.27 | −0.97 | −0.93 | 4.17 | HF/6-311++G(3df,3pd) | |
| | | 0.18 | 2.13 | −0.93 | −0.90 | 3.96 | B3LYP/6-311++G(2d,2p) | |
| | | 0.19 | 2.12 | −1.04 | −1.00 | 4.16 | B3LYP/6-311++G(3df,3pd) | |
| | N2—H2B...O1 | 0.064(0) | 1.716(0) | −0.35 | −0.29 | 2.36 | X-ray diffraction | |
| | | 0.10 | 1.36 | −0.37 | −0.34 | 2.07 | HF/6-311++G(2d,2p) | |
| | | 0.10 | 1.41 | −0.38 | −0.33 | 2.12 | HF/6-311++G(3df,3pd) | |
| | | 0.10 | 1.29 | −0.39 | −0.35 | 2.03 | B3LYP/6-311++G(2d,2p) | |
| | | 0.10 | 1.33 | −0.40 | −0.35 | 2.08 | B3LYP/6-311++G(3df,3pd) | |
| | H ₂ O | O4—H4A...O2 | 0.12(2) | 3.34(1) | −0.65 | −0.59 | 4.58 | X-ray diffraction |
| | | | 0.16 | 2.41 | −0.82 | −0.79 | 4.02 | HF/6-311++G(2d,2p) |
| | | | 0.17 | 2.44 | −0.93 | −0.90 | 4.27 | HF/6-311++G(3df,3pd) |
| | | | 0.18 | 2.27 | −0.89 | −0.87 | 4.03 | B3LYP/6-311++G(2d,2p) |
| | | | 0.18 | 2.31 | −0.99 | −0.96 | 4.26 | B3LYP/6-311++G(3df,3pd) |
| O4—H4B...O2 | | 0.110(0) | 3.58(1) | −0.56 | −0.55 | 4.69 | X-ray diffraction | |
| | | 0.17 | 2.32 | −0.80 | −0.75 | 3.87 | HF/6-311++G(2d,2p) | |
| | | 0.18 | 2.37 | −0.90 | −0.86 | 4.13 | HF/6-311++G(3df,3pd) | |
| | | 0.19 | 2.19 | −0.88 | −0.83 | 3.90 | B3LYP/6-311++G(2d,2p) | |
| | | 0.19 | 2.24 | −0.97 | −0.93 | 4.14 | B3LYP/6-311++G(3df,3pd) | |

reference frame, and experimental errors in individual Cartesian tensor elements must be indirectly translated into errors in degrees or radians. A comparison of tensors in the icosahedral representation, however, is much more convenient, since both the magnitudes and the orientations of the tensors are quantitatively evaluated at the same time by comparing the six icosahedral tensor elements, χ_{1-6} . Also, the six icosahedral elements are equally weighted in any coordinate frame, making this representation ideal for fitting a least-squares line through theoretical-versus-experimental tensor data.¹⁸ Using this approach, we find very good agreement between theory and experiment for both the magnitudes and the orientations of the Hessian tensors. We again find that the DFT method with a 6-311++G(2d,2p) basis set provides the best theory-versus-experiment correlation (slope = 1.09, $R^2 = 0.96$, Table 3, Figure 2B).

The combination of X-ray diffraction and quantum chemistry also provides a direct means to investigate hydrogen bonding via topological analysis of the associated electron density.⁴² A (3,−1) BCP was therefore located for each of the seven distinct hydrogen bonds shown in the structure of L-asparagine·H₂O (Figure 3A), and the topology of $\rho(\mathbf{r})$ was analyzed at these points. To provide hydrogen bond partners, we incorporated additional small molecules into the calculations, basically as reported previously in the context of chemical shielding calculations in proteins^{1,2} and in L-asparagine·H₂O.⁴³ For example, to describe the N1—H···O hydrogen bonds (Figure 3B), we utilized a central asparagine molecule, two appropriately positioned asparagine fragments, and one water molecule. The

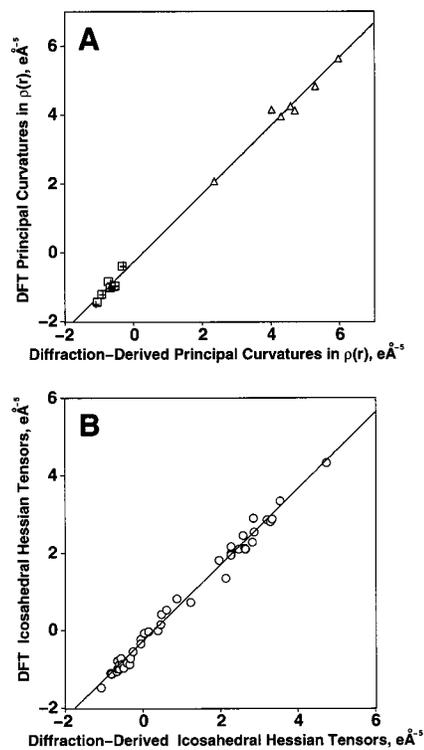


Figure 4. Curvatures in $\rho(\mathbf{r}_b)$ at (3,−1) BCPs along hydrogen bonds. (A) Principal curvatures in $\rho(\mathbf{r}_b)$: B3LYP 6-311++G(3df,3pd) versus X-ray diffraction. + = λ_1 , \square = λ_2 , \triangle = λ_3 . (B) Icosahedral Hessian-of- $\rho(\mathbf{r}_b)$ tensors: B3LYP 6-311++G(3df,3pd) versus X-ray diffraction.

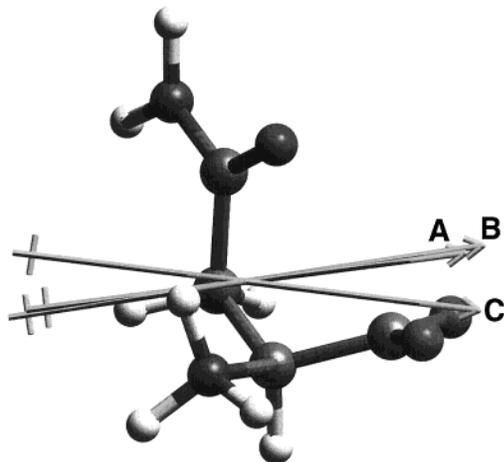
(42) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1994.

(43) Scheurer, C.; Skrynnikov, N. R.; Lienin, S. F.; Straus, S. K.; Brühweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 4242–4251.

O1 acceptor molecule was replaced by a glycine zwitterion, while the O3 acceptor was replaced by an acetamide molecule. Construction of such fragments required only truncation of the

Table 5. Molecular Dipole Moment of L-Asparagine^a

| | μ_x | μ_y | μ_z | $ \mu $ |
|-------------------------|---------|---------|---------|---------|
| X-ray diffraction | 5.0(3) | -3.0(1) | 13.1(3) | 14.3(3) |
| HF/6-311++G(2d,2p) | 8.8 | -0.9 | 11.3 | 14.3 |
| HF/6-311++G(3df,3pd) | 8.7 | -0.9 | 11.2 | 14.3 |
| B3LYP/6-311++G(2d,2p) | 7.9 | -0.8 | 10.5 | 13.2 |
| B3LYP/6-311++G(3df,3pd) | 7.8 | -0.8 | 10.5 | 13.1 |

^a All values in debyes.**Figure 5.** Orientation of the molecular dipole moment vector of L-asparagine: (A) B3LYP/6-311++G(3df,3pd), (B) HF/6-311++G(3df,3pd), and (C) X-ray diffraction.

original asparagine lattice, leaving the original asparagine geometry intact. Hydrogen atom geometries were fixed to the neutron diffraction values.⁴⁴ Three such cluster calculations (Figure 3B–D) were sufficient to describe all seven hydrogen bonds, and the calculated topological properties are shown in Table 4, while Figure 4A shows a comparison between the 21 theoretical and experimental principal curvatures, λ_{1-3} , for the seven hydrogen bonds. As with the covalent, heavy-atom bonds, the correlation is very good, with a slope of 0.99 and an R^2 value of 1.00 for the B3LYP/6-311++G(3df,3pd) calculation (Table 3). The predicted directions of principal curvature in $\rho(\mathbf{r}_b)$ at hydrogen bond BCPs, expressed by icosahedral Hessian tensors, are also in excellent agreement with the diffraction values (Figure 4B), with the B3LYP/6-311++G(3df,3pd)

calculation producing a theory-versus-experiment correlation of slope = 0.98 and $R^2 = 0.99$ (Table 3).

We next investigated three other electrical/electrostatic properties: the dipole moment, μ ; the electrostatic potential, $\Phi(\mathbf{r})$; and the electric field gradient, $\nabla\mathbf{E}$. The dipole moment vector results from the molecular charge distribution, and both its magnitude and its direction can be extracted from X-ray diffraction data.^{31,45} Table 5 compares the molecular dipole moment of L-asparagine as determined by diffraction with the values calculated by using both HF and DFT methods, using both basis sets. The water molecule was removed during the evaluation of the X-ray dipole moment, as permitted by XD.²⁸ Within experimental error, the *ab initio* HF method predicts dipole moment magnitudes in very good agreement with the X-ray value. The B3LYP DFT method produces a dipole moment which is slightly smaller (1 D) than the X-ray value, an effect which has been noted previously for DL-histidine.¹⁷ The molecular orientations of the experimental and calculated μ vectors are shown in Figure 5.

In addition to a dipole moment, the molecular charge distribution also induces a more complex, three-dimensional electrostatic potential, $\Phi(\mathbf{r})$. An experimental $\Phi(\mathbf{r})$ was drawn from the X-ray data in the manner described by Su and Coppens.⁴⁰ Figure 6A shows the X-ray potential, mapped onto a $0.11 \text{ e } \text{\AA}^{-3}$ isosurface of the experimentally determined electron density, $\rho(\mathbf{r})$. All values of $\Phi(\mathbf{r})$ which spatially intersect this isodensity are depicted. The corresponding quantum chemical potential and density are depicted in Figure 6B (HF, large basis) and C (DFT, large basis) and clearly reproduce the main experimental charge distribution seen in L-asparagine (Figure 6A). Both basis sets produced essentially identical pictures (data not shown). The small differences which are seen between Figure 6A (experiment) and Figure 6B,C (calculation) are consistent with those previously reported with alanine, proline, and aspartic acid, which were attributed to crystal-field polarization effects.^{13–16} That is, the calculations evaluate the molecular electrostatic potential of an isolated, gas-phase L-asparagine molecule, while in its crystalline state each L-asparagine molecule is hydrogen bonded to three water molecules and eight other asparagines. Thus, it is completely reasonable that the $\Phi(\mathbf{r})$ of a single asparagine molecule extracted from the crystal will be subject to additional polarizations. We did not attempt such large supermolecule calculations,

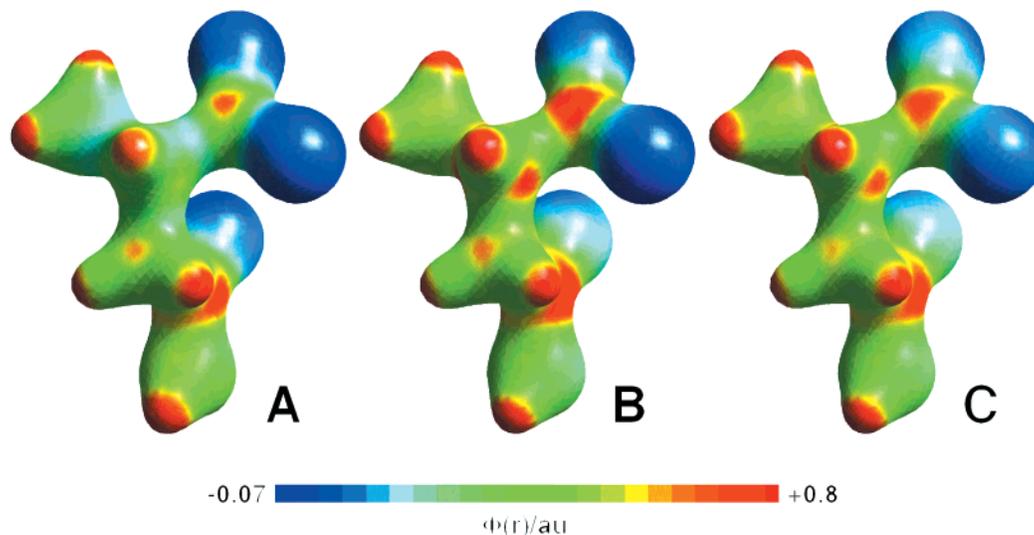
**Figure 6.** Molecular electrostatic potential, $\Phi(\mathbf{r})$, mapped from -0.07 to $+0.8 \text{ e}/4\pi\epsilon_0 a_0$, onto a $0.11 \text{ e } \text{\AA}^{-3}$ isosurface of the electron density: (A) X-ray diffraction, (B) HF/6-311++G(3df,3pd) calculation, and (C) B3LYP/6-311++G(3df,3pd) calculation.

Table 6. Principal Elements^a of Oxygen and Nitrogen Electric Field Gradient Tensors in L-Asparagine Monohydrate

| atom | V ₃₃ | V ₂₂ | V ₁₁ | method |
|------|-----------------|-----------------|-----------------|-------------------------|
| O1 | 0.17(4) | 0.5(1) | -0.7(1) | X-ray |
| | 0.33 | 1.18 | -1.51 | HF/6-311G++(2d,2p) |
| | 0.35 | 1.13 | -1.48 | HF/6-311G++(3df,3pd) |
| | 0.36 | 1.02 | -1.39 | B3LYP/6-311G++(2d,2p) |
| | 0.37 | 0.99 | -1.37 | B3LYP/6-311G++(3df,3pd) |
| O2 | 0.17(4) | 0.5(1) | -0.7(1) | X-ray |
| | 0.38 | 1.04 | -1.42 | HF/6-311G++(2d,2p) |
| | 0.40 | 0.99 | -1.39 | HF/6-311G++(3df,3pd) |
| | 0.36 | 0.89 | -1.25 | B3LYP/6-311G++(2d,2p) |
| | 0.37 | 0.85 | -1.23 | B3LYP/6-311G++(3df,3pd) |
| O3 | 0.27(5) | 0.6(1) | -0.9(2) | X-ray |
| | 0.77 | 1.03 | -1.80 | HF/6-311G++(2d,2p) |
| | 0.78 | 0.98 | -1.77 | HF/6-311G++(3df,3pd) |
| | 0.75 | 0.92 | -1.67 | B3LYP/6-311G++(2d,2p) |
| | 0.76 | 0.89 | -1.65 | B3LYP/6-311G++(3df,3pd) |
| N1 | 0.02(3) | 0.02(3) | -0.04(5) | X-ray |
| | -0.01 | -0.11 | 0.12 | HF/6-311G++(2d,2p) |
| | -0.02 | -0.11 | 0.13 | HF/6-311G++(3df,3pd) |
| | 0.00 | -0.11 | 0.11 | B3LYP/6-311G++(2d,2p) |
| | 0.00 | -0.11 | 0.11 | B3LYP/6-311G++(3df,3pd) |
| | -0.0955(6) | -0.1496(6) | 0.2451(6) | ¹⁴ N NMR |
| N2 | -0.02(9) | -0.05(3) | 0.07(6) | X-ray |
| | 0.44 | 0.50 | -0.94 | HF/6-311G++(2d,2p) |
| | 0.45 | 0.48 | -0.93 | HF/6-311G++(3df,3pd) |
| | 0.42 | 0.47 | -0.89 | B3LYP/6-311G++(2d,2p) |
| | 0.42 | 0.46 | -0.89 | B3LYP/6-311G++(3df,3pd) |
| | 0.191(2) | 0.379(2) | -0.570(2) | ¹⁴ N NMR |

$$^a V_{ii} = -\nabla E_{ii}, |V_{11}| \geq |V_{22}| \geq |V_{33}|.$$

which are probably best performed by using periodic boundary conditions, since already at the single-molecule level there is very good agreement between theory and experiment for $\Phi(\mathbf{r})$.

Finally we consider the determination and prediction of ∇E , the electric field gradient (EFG) tensor. An X-ray ∇E at each nuclear position can be extracted from the Hansen–Coppens multipole population parameters,²⁹ so we initially compare the principal values of these EFGs for the three oxygen and two nitrogen nuclei in L-asparagine with those predicted by using HF and DFT methods (Table 6). Although there is little correlation between theory and experiment for the nitrogen nuclei, the calculated EFGs for the three oxygen nuclei are in good agreement with the X-ray values (Table 3). All four methods produce R^2 values of 0.98 for the theory-versus-experiment correlation, with the DFT calculations giving slightly better slopes (Figure 7A).

To compare the orientations of these tensors, we also expressed the EFGs in the icosahedral representation.¹⁸ Here, the predicted oxygen ∇E tensors are in moderate accord with the X-ray diffraction values (Table 3), with the best theory-versus-experiment correlation (slope = 1.70 and $R^2 = 0.82$) produced by the B3LYP/6-311++G(3df,3pd) calculation (Figure 7B and Table 3). There is again no correlation between theory and diffraction experiment EFGs at the two nitrogen nuclei. While one possibility is that the calculations simply break down for this particular electrostatic property, it seemed more reasonable, given the discussion of this topic in the literature, that the EFG at the nucleus is simply more difficult to deduce from the experimental diffraction data. In addition, in recent work Ernst and co-workers have shown excellent agreement between experimental and theoretical (DFT) values for the amide ¹⁵N chemical shielding tensor in L-asparagine·H₂O.⁴³ If the quantum chemical EFG calculations are in fact accurate, then

(44) Verbist, J. J.; Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *Acta Crystallogr.* **1972**, B28, 3006–3013.

(45) Spackman, M. A.; Byrom, P. G. *Acta Crystallogr.* **1996**, B52, 1023–1035.

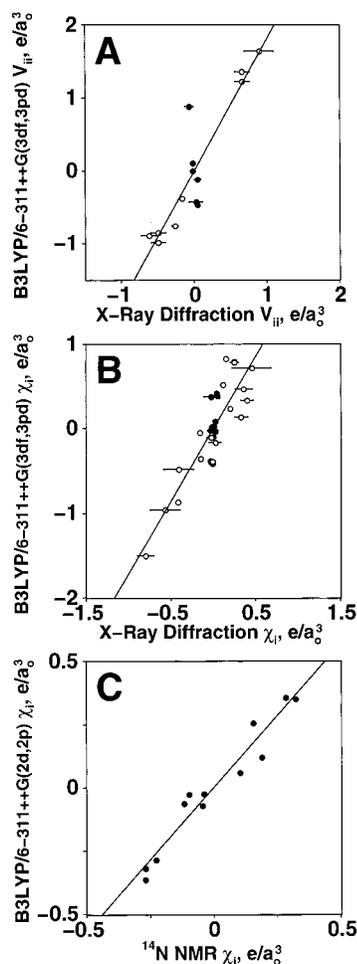


Figure 7. Electric field gradient tensors at the oxygen and nitrogen nuclei in L-asparagine. (A) Principal values of the O1, O2, and O3 EFG tensors (○): B3LYP/6-311++G(3df,3pd) versus X-ray diffraction (slope = 1.86, $R^2 = 0.98$). Nitrogen EFG principal values (N1, N2, ●) are not included in the correlation. (B) Icosahedral EFG tensors at O1, O2, O3 (○) and N1, N2 (●): B3LYP/6-311++G(3df,3pd) versus X-ray diffraction (slope = 1.70, $R^2 = 0.82$). Nitrogen values are not included in the correlation. (C) Icosahedral EFG tensors at nitrogen nuclei: point charge B3LYP/6-311++G(2d,2p) versus single-crystal ¹⁴N NMR (slope = 1.14, $R^2 = 0.94$).

it should be possible to use highly accurate, single-crystal ¹⁴N and ¹⁷O NMR determinations of the EFG to validate the calculations. While ¹⁷O ∇E tensors have not yet been reported for L-asparagine·H₂O, the ¹⁴N tensors have been reported in a single-crystal study by McDowell and co-workers.¹¹ The calculated nitrogen ∇E tensors are indeed correlated with these NMR tensors, though the agreement is somewhat less than ideal (slope = 1.28 and $R^2 = 0.81$ for the B3LYP/6-311++G(2d,2p) theory-versus-experiment correlation) (Table 3). To investigate whether this correlation could be improved by considering crystal lattice effects, a charge field was introduced into a second series of nitrogen EFG calculations. Mulliken atomic charges were determined *ab initio* at the 6-311++G(2d,2p) level for each atom in the L-asparagine·H₂O asymmetric unit. Crystal symmetry operations were then used to construct a large lattice of charged atoms, from which a single asparagine molecule, to be treated explicitly with basis functions, was chosen. Atoms which fell within a 15-Å radius from the center-of-mass of this molecule were included as point charges (ca. 2500). These large charge-field calculations produced EFG tensors in excellent agreement with those measured by ¹⁴N NMR (Table 3). The

best theory-versus-experiment correlation is again found with the B3LYP/6-311++G(2d,2p) method (Figure 7C), which produced a slope of 1.14 and an R^2 value of 0.94. In contrast, the X-ray result shows a systematic underestimation of the magnitude of O1, O2, and O3 EFGs (Table 6) and poorer R^2 values, when comparing $\nabla\mathbf{E}$ tensors in the icosahedral representation. An antishielding factor, like the Sternheimer factor routinely applied in the analysis of ^{57}Fe Mössbauer data, may also be appropriate for the X-ray data but is clearly irrelevant for the NMR determination, which probes the EFG at the nucleus in a direct manner.

Clearly then, the differences found with the diffraction-derived correlation appear not to originate primarily either in the structure or in the quantum chemical calculations, since both the magnitudes and the orientations of the EFG tensor are accurately evaluated theoretically from the asparagine structure, as shown in Figure 7C, when using NMR data. Thus, while $\rho(\mathbf{r})$, $\partial^2\rho/\partial r_{ij}$, μ , and $\Phi(\mathbf{r})$ are all well described from diffraction data, $\nabla\mathbf{E}$ is simply more difficult to extract, and for the moment it appears to be most readily deduced by using NMR methods. Nevertheless, the fact that all properties— $\rho(\mathbf{r})$, $\partial^2\rho/\partial r_{ij}$, μ , $\Phi(\mathbf{r})$, and $\nabla\mathbf{E}$ —can be accurately predicted from moderate-sized HF or DFT calculations using, as appropriate, either diffraction or NMR data strongly supports the use of such methods in providing accurate representations of the electrical, electrostatic, and bonding properties in amino acids, and by inference in peptides and larger systems as well.

Conclusion

The results we have shown above are of interest for several reasons. First, we have shown that the charge density, $\rho(\mathbf{r})$, the dipole moment, μ , the molecular electrostatic potential, $\Phi(\mathbf{r})$, and the curvatures of the charge density, $\partial^2\rho/\partial r_{ij}$, at (3, -1) bond critical points and hydrogen bond (3, -1) critical points in L-asparagine monohydrate extracted from X-ray diffraction data are all in good accord with values computed theoretically by using either HF or DFT methods. Second, we find that the oxygen EFG tensors, $\nabla\mathbf{E}$, obtained by X-ray diffraction are in somewhat less good accord with the calculations, while for the nitrogen nuclei little or no correlation exists. However, these

(46) de Dios, A. C.; Laws, D. D.; Oldfield, E. *J. Am. Chem. Soc.* **1994**, *116*, 7784–7786.

same theoretical nitrogen EFGs are in moderate accord with those obtained by using single-crystal NMR, which is arguably the most accurate and amenable probe of the EFG at the nucleus in such systems. When point charges are added to the calculations, as reported previously for chemical shift anisotropy calculations,⁴⁶ excellent agreement between theoretical and experimental nitrogen NMR EFGs is found. Third, we have extended the use of the icosahedral representation¹⁸ to the Hessian-of- $\rho(\mathbf{r})$ and EFG tensors, in which both tensor magnitude and orientation are conveniently represented by six equally weighted icosahedral tensor elements. This enables a simple and general means of comparing tensor orientation information and facilitates a more rigorous evaluation of experimental-versus-theoretical property comparisons. Overall, these results validate the idea that quantum chemical calculations on amino acids enable the accurate prediction of many electrostatic properties— $\rho(\mathbf{r})$, $\partial^2\rho/\partial r_{ij}$, μ , $\Phi(\mathbf{r})$, and $\nabla\mathbf{E}$ —which gives additional confidence in using both HF and DFT methods to probe the local electrostatic properties of even more complex systems.

Acknowledgment. This work was supported by the United States Public Health Service (Grant GM-50694), and was carried out in part by use of the SGI/Cray Origin 2000 and Power Challenge clusters at the National Center for Supercomputing Applications (funded in part by the U.S. National Science Foundation, Grant CHE-97002ON). Support of the work at SUNY/Buffalo by the National Science Foundation (CHE9615586) is gratefully acknowledged. The SUNY X3 beamline at NSLS is supported by the Division of Basic Energy Sciences of the U.S. Department of Energy (DE-FG02-86ER45231). Research was carried out in part at the National Synchrotron Light Source at Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Division of Materials Sciences and Division of Chemical Sciences.

Supporting Information Available: Tables of multipole population parameters, radial screening parameters, and mean-square displacement amplitudes for non-hydrogen atoms (PDF). X-ray crystallographic files, in CIF format, are also available. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA000386D