Experimental and Theoretical Investigation of the ^{13}C and ^{15}N Chemical Shift Tensors in Melanostatin—Exploring the Chemical Shift Tensor as a Structural Probe

Mark Strohmeier and David M. Grant*

Contribution from the Department of Chemistry, University of Utah, 315 South. 1400 East, Salt Lake City, Utah 84112-0850

Received July 16, 2003; E-mail: grant@chem.utah.edu

Abstract: The determination of backbone conformations in powdered peptides using ^{13}C and ^{15}N shift tensor information is explored. The ^{13}C and ^{15}N principal shift values in natural abundance ^{13}C and ^{15}N melanostatin (L-Pro-L-Leu-Gly amide) are measured using the FIREMAT technique. Furthermore, the orientation of the C−N bond in the ^{13}C shift principal axis system for the backbone carbons is obtained from the presence of the ^{13}C−^{14}N dipolar coupling. The Ramachandran angles for the title compound are obtained from solid-state NMR data by comparing the experimentally determined shift tensor information to systematic theoretical shielding calculations on N-formyl-L-amino acid-amide models. The effects of geometry optimization and neglect of intermolecular interactions on the theoretical shielding values in the model compounds are investigated. The sets of NMR derived Ramachandran angles are assembled in a set of test structures that are compared to the available single-crystal X-ray structure. Shift tensor calculations on the test structures and the X-ray structure are used to further assess the importance of intermolecular interactions when the shift tensor is used as a structural probe in powdered peptides.

Introduction

In recent years, there has been increased interest in obtaining structural information from chemical shift principal values.[1−4] It has been shown for several nonpolar terpenes that reasonable structural prediction can be made by combining density functional theory (DFT) calculations and accurate principal shift value measurements when intermolecular effects are negligible or can be avoided by appropriately choosing the ^{13}C positions that probe the structure.[5−7] In addition, progress has been made toward determination of the Ramachandran angles Φ and Ψ (see Figure 1 for definitions) in peptides by relating chemical shift measurements of C$^\alpha$ and C$^\beta$ to quantum mechanical calculations on small model systems.[2,8] This includes several studies on invertebrate calmodulin and Staphylococcal nuclease where calculations on known X-ray structures and optimized structures reproduce the known solution isotropic chemical shifts of the valine residues.[9,10] In addition, the sensitivity of the
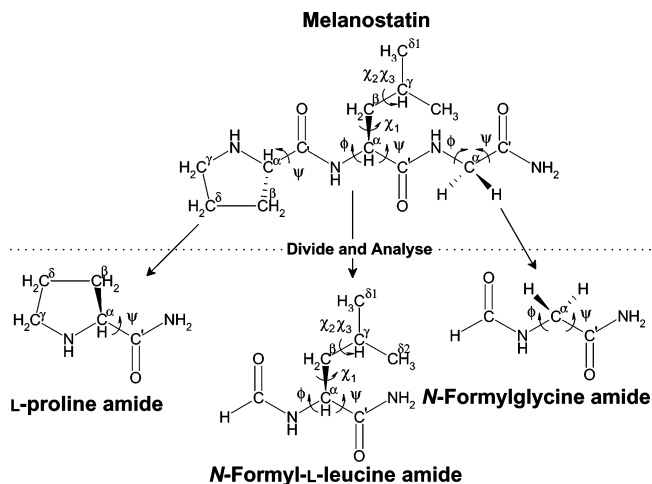
Figure 1. Melanostatin and the three molecular models that were utilized for shift tensor surface calculations are shown. The nomenclature and the Ramachandran and the side chain angles are indicated.

"CSA" (Δσ*), defined as the difference between the chemical shift component along the C$^\alpha$−H and the average shift perpendicular to the C$^\alpha$−H bond, to structural differences has also been investigated. It was found that calculations on a relatively small N-formylamino acid-amide fragment of the peptide backbone are sufficient to account for experimental observations in several proteins.[11,12] The sensitivity of Δσ* results mainly from the change of shift tensor orientation

(1) Barich, D. H.; Pugmire, R. J.; Grant, D. M.; Iuliucci, R. J. *J. Phys. Chem. A* **2001**, *105*, 6780−6784.
(2) Heller, J.; Laws, D. D.; Tomaselli, M.; King, D. S.; Wemmer, D. E.; Pines, A.; Havlin, R. H.; Oldfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 7827−7831.
(3) Sakurai, M.; Wada, M.; Inoue, Y.; Tamura, Y.; Watanabe, Y. *J. Phys. Chem.* **1996**, *100*, 1957−1964.
(4) Harper, J. K.; Facelli, J. C.; Barich, D. H.; McGeorge, G.; Mulgrew, A. E.; Grant, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 10 589−10 595.
(5) Harper, J. K.; Mulgrew, A. E.; Li, J. Y.; Barich, D. H.; Strobel, G. A.; Grant, D. M. *J. Am. Chem. Soc.* **2001**, *123*, 9837−9842.
(6) Harper, J. K.; Grant, D. M. *J. Am. Chem. Soc.* **2000**, *122*, 3708−3714.
(7) Harper, J. K.; Barich, D. H.; Hu, J. Z.; Strobel, G. A.; Grant, D. M. *J. Org. Chem* **2003**, *68*, 4609−4614.
(8) Havlin, R. H.; Laws, D. D.; Bitter, H.-M. L.; Sanders, L. K.; Sun, H.; Grimley, J. S.; Wemmer, D. E.; Pines, A.; Oldfield, E. *J. Am. Chem. Soc.* **2001**, *123*, 10 362−10 369.
(9) Laws, D. D.; Le, H.; de Dios, A.; Havlin, R. H.; Oldfield, E. *J. Am. Chem. Soc.* **1995**, *117*, 9542−9546.
(10) Pearson, J. G.; Le, H.; Sanders, L. K.; Godbout, N.; Havlin, R. H.; Oldfield, E. O. *J. Am. Chem. Soc.* **1997**, *119*, 11 941−11 950.

resulting from different Ramachandran angles and emphasize the importance of orientational information when shift tensors are experimentally accessible.[8]

Several refinement protocols utilizing shift tensor information have been proposed recently. Cross and co-workers[13,14] utilized shift tensor and NH dipolar coupling measurements in oriented samples to refine the structure of membrane proteins. Witter et al.[15] implemented a procedure using fast semiempirical shift tensor calculations and measured solid-state isotropic shifts and molecular mechanics energies to refine the proton positions in mannitol. Despite the differences of these two approaches, both minimize a constructed energy function that is defined by the sum of the molecular mechanics energy and a function quantizing the agreement between theoretical and experimentally observed shifts, similar to the restrained optimization procedures used for structure elucidation based on solution state NMR data.

In this paper, the feasibility of structural analysis of peptide samples utilizing the ^{13}C and ^{15}N shift tensor principal components obtained with FIREMAT experiments on a natural abundance sample is investigated. In contrast to rather nonpolar hydrocarbons such as terpenes, the chemical shift principal values in solid peptides can be affected considerably by intermolecular interactions, which complicates the analysis as simple structural models cannot include these effects. Melanostatin (L-Pro-L-Leu-Gly amide, Figure 1) also known as melanin-release inhibiting factor or MIF, is a neurohormone that is responsible for modulating dopamine receptors in the central nervous system.[16,17] The high-resolution of the ^{13}C FIREMAT spectra reveal that the ^{13}C$-^{14}$N dipolar coupling cannot be simplified, and the quadrupolar interaction of the ^{14}N has to be included in the spectral analysis even at 9.4 T. Including the ^{13}C$-^{14}$N dipolar coupling into the spectral analysis of the FIREMAT spectra allows for the extraction of five out of the six components of the coupled C$^\alpha$ and C$'$ shift tensors, yielding valuable orientational information on the shift tensor.[18]

The determination of the Φ/Ψ dihedral angles from the chemical shift tensor data is done by comparison of the experimental values with theoretical Φ/Ψ shift tensor surfaces for *N*-formyl-glycine amide, *N*-formyl-L-leucine amide, and L-proline amide model systems given in Figure 1. Extended geometry optimizations of the structural models and subsequent shift tensor calculations are done using DFT. For studies on larger systems and for crystal structure prediction only computationally less demanding low levels of theory are currently feasible for geometry optimization. Thus, the effect of different geometry optimization methods on the theoretical shielding tensors of the *N*-formylamino acid amide model systems is investigated for several points in the Φ/Ψ space to validate

systematic errors in the resulting shielding surfaces. The effect of long and short-range electrostatic intermolecular interactions and truncation of the molecular model on the resulting chemical shift tensor components calculations is investigated utilizing the embedded ion method (EIM) with the known X-ray structure.[19−21]

The EIM has been shown to increase the agreement between experimental principal shift values and theoretical principal shielding values dramatically when strong lattice effects are present such as in several carbonates and thiocarbonates as well as in nucleosides and amino acids.[22−24] The conformation obtained from the analysis of the experimental shift tensor data is compared with the X-ray structure for melanostatin. Discrepancies between predicted and X-ray structure are discussed in light of neglect of intermolecular interactions, geometry optimization, and truncation of the molecular model.

Experimental Section

Sample Preparation. Natural abundance Melanostatin (L-Pro-L-Leu-Gly amide) was obtained from Sigma-Aldrich and used without any further purification. Crystalline samples for X-ray and NMR analysis were obtained by recrystallization from a water/methanol mixture by evaporation of methanol at room temperature. The crystalline material was filtered off and washed with water, air-dried, and ground to a fine white powder for NMR analysis. A single crystal for X-ray structure determination was directly obtained from the mother solution and air-dried.[21] The samples are stable toward water loss and the CPMAS spectra remain unchanged even after placing the crystalline material in high vacuum for several days.

NMR Spectroscopy. High-speed MAS and FIREMAT[25] experiments were performed on a CMX200 (4.7 T) and a CMX400 (9.4 T) spectrometer operating at 50.305 and 100.621 MHz carbon frequencies, respectively. Both spectrometers are equipped with 7.5 mm PENCIL probes and a feedback circuit to synchronize the pulse sequence to the rotor position. In all experiments, transverse magnetization was produced by cross polarization from protons. The π pulses for protons and carbons were approximately 8 μs on both spectrometer systems. TPPM[26,27] ^{1}H decoupling was used with a phase shift angle of 36° for the ^{13}C FIREMAT experiments at 4.7 and 9.4 T. The spinning speed for the FIREMAT experiments was $f_r = 500$ Hz at both fields. At 50.305 and 100.621 MHz, 16 and 32 evolution increments were taken, respectively. The corresponding spectral widths in the acquisition dimension were 48 kHz and 96 kHz, and in the evolution dimension, they were 8 kHz and 16 kHz, respectively. All ^{13}C spectra were referenced externally to the TMS shift scale by the downfield resonance of adamantane at 38.56 ppm. The ^{15}N FIREMAT spectrum was recorded at 4.7 T field. The spinning speed was $f_r = 400$ Hz with 8 evolution increments taken. The spectral widths were in the acquisition dimension and evolution dimension were 19.2 and 3.2 kHz, respectively. CW ^{1}H decoupling was used with a field of 62 kHz. The π pulse length for ^{15}N was 16 μs. The ^{15}N spectrum was referenced externally to the nitromethane scale using the α glycine resonance at −346.43 ppm.

(11) Szabo, C. M.; Sanders, L. K.; Arnold, W.; Grimley, J. S.; Godbout, N.; McMahon, M. T.; Moreno, B.; Oldfield, E. In *Modeling NMR Chemical Shifts*; Facelli, J., deDios, A. C., Eds.; ACS Symposium Series 732; 1999; p 40.
(12) Sun, H.; Sanders, L.; Oldfield, E. *J. Am. Chem. Soc.* **2002**, *124*, 5486−5495.
(13) Bertram, R.; Quine, J. R.; Chapman, M. S.; Cross, T. A. *J. Magn. Reson.* **2000**, *147*, 9−16.
(14) Kim, S.; Quine, J. R.; Cross, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 7292−7298.
(15) Witter, R.; Priess, W.; Sternberg, U. *J. Comput. Chem.* **2002**, *23*, 298−305.
(16) Mishra, R. K.; Chiu, S.; Chiu, P.; Mishra, C. P *Methods. Find. Exp. Clin. Pharmacol.* **1983**, *5*, 203−233.
(17) Celis, M. E.; Taleisnik, S.; Walter, R. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1428−1433.
(18) Strohmeier, M.; Alderman, D. W.; Grant, D. M. *J. Magn. Reson.* **2002**, *155*, 263−277.

(19) Stueber, D.; Guenneau, F. N.; Grant, D. M. *J. Chem. Phys.* **2001**, *114*, 9236−9243.
(20) Reed, L. L.; Johnson, P. L. *J. Am. Chem. Soc.* **1973**, *95*, 7523−7524.
(21) CCDC 216376 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/conts/retrieving.html (or from the CCDC, 12 Union Road, Cambridge CB2 1EZ, UK.; fax: +44 1223 336033; e-mail: deposit@ccdc.cam.ac.uk).
(22) Stueber, D.; Orendt, A. M.; Facelli, J. C.; Parry, R. W.; Grant, D. M. *Solid State NMR* **2002**, *22*, 29.
(23) Stueber, D.; Grant, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 10 539.
(24) Strohmeier, M.; Stueber, D.; Grant, D. M. *J. Chem. Phys. A* **2003**, *107*, 7629−7642.
(25) Alderman, D. W.; McGeorge, G.; Hu, J. Z.; Pugmire, R. J.; Grant, D. M. *Mol. Phys.* **1998**, *95* (6), 1113−1126.
(26) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951−6958.
(27) McGeorge, G.; Alderman, D. W.; Grant, D. M. *J. Magn. Reson.* **1999**, *137*, 138−143.
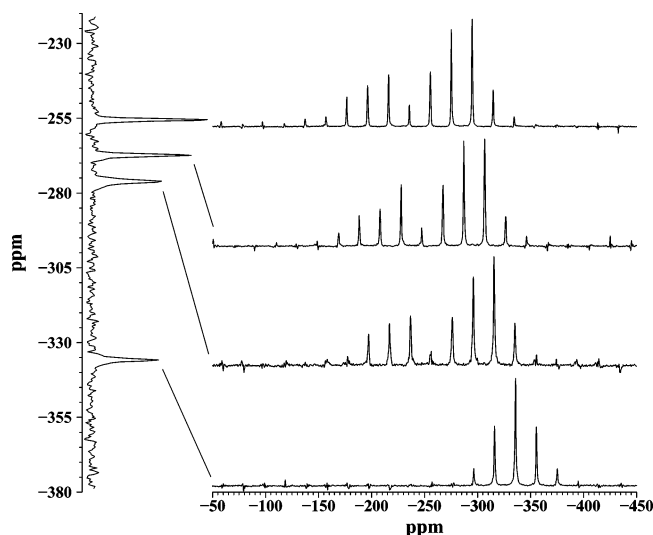
Figure 2. Melanostatin 4.7 T ^{15}N FIREMAT spectrum. The one-dimensional P2DSS spectrum is shown on the left-hand side, TIGER extracted sideband patterns for each of the isotropic peaks are given on the right-hand array of spectra.
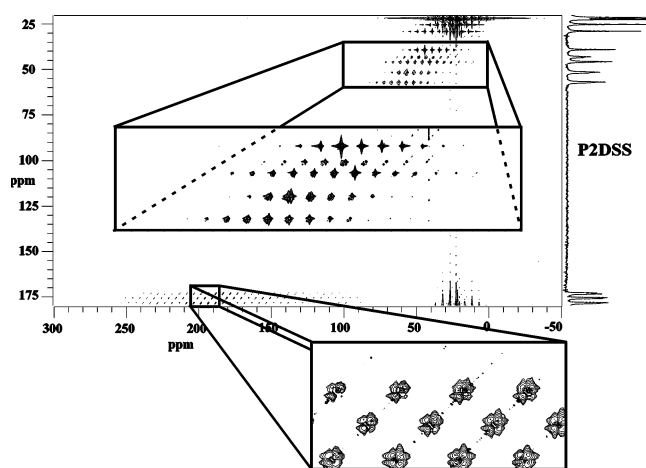
Figure 3. Melanostatin 9.4 T ^{13}C FIREMAT data shown as 2D absolute value mode spectrum. Sideband patterns in the acquisition dimension are resolved for each ^{13}C isotropic position in the evolution dimension (P2DSS spectrum). The top expansion box shows the complete C$^\alpha$ and the bottom expansion box a portion of the C′ sideband patterns.

The ^{15}N FIREMAT spectrum was TIGER[28] processed, and the resulting sideband patterns were fit to a shift tensor model using the banded matrix[29] approach. The P2DSS[25,30] spectrum and the extracted sideband pattern are shown in Figure 2.

Spectral analysis of the ^{13}C data utilized the approach described previously.[18] The high-resolution of the two-dimensional ^{13}C spectra in Figure 3 clearly demonstrates that the dipolar coupling to the quadrupolar nucleus ^{14}N has an profound effect on the line shape of coupled shifts even at 9.4 T. Thus, the spectral analysis must account for this inhomogeneous broadening mechanism by including the quadrupolar coupling of the ^{14}N nucleus in addition to the ^{13}C$-^{14}$N dipole−dipole coupling. The quadrupolar coupling constants χ, asymmetries η and geometrical parameters applied in the fitting procedure, given in Table 1 are estimated from experimental values for peptides and amides reported in the literature.[31,32] The quadrupolar coupling tensors in peptides are rather insensitive to conformational changes and

Table 1. QCC Tensors Utilized for Fitting of ^{13}C FIREMAT Spectra

	χ/MHz	η	θ^a (deg)	ϕ^a (deg)
Gly N-1	−2.65	0.33	90	90
Gly N^b	−3.10	0.40	90	70
Leu N^b	−3.10	0.40	90	70
Pro N^b	−4.00	0.20	104	64

^{a} The angles θ and ϕ are the polar angles that define the CN bond vector in the EFG principal axis system. ^{b} The orientation of the C$^\alpha$N bond vector is reported. The C′N and Pro C$^\delta$N orientations follow when standard geometry is assumed.

therefore estimates, based on experimental values in similar compounds, are quite reasonable.[33] The situation is different, however, for the amine nitrogen in the proline residue, which is hydrogen bonded via its lone pair; a reliable estimate based on values from solid-state experiments is presently unavailable for similar compounds.

Quantum Mechanical Computations. All theoretical computations are performed with the quantum chemistry suite *Gaussian 98* revision A.11.[34] Chemical shielding tensors were calculated at the B3LYP/D95** level[35−39] using GIAO.[40,41] All computations were performed on a Beowulf cluster of personal computers with AMD, Intel Pentium II and III processors operating from 0.35 to up to 1.333 GHz, as well as on a SGI Origin 2000. Complete Ramachandran Φ/Ψ shielding surfaces for the L-leucine and glycine residues were generated using *N*-formyl-L-leucine amide and *N*-formyl-glycine amide fragments (Figure 1) that were used in previous studies of Φ/Ψ shift surfaces.[8] All structural parameters except the dihedral angles Φ and Ψ and the side chain angles χ_1, χ_2, and χ_3 in *N*-formyl-L-leucine amide were geometry optimized using B3LYP with the D95** basis set.[42] The side chain angles in *N*-formyl-L-leucine amide were set to the most populated conformation: $\chi_1 = -60°$, $\chi_2 = -60°$, and $\chi_3 = 180°$.[8] The molecular symmetry of *N*-formylglycine amide surface allows for the calculation of only half the surface to cover the complete Φ/Ψ space. Therefore, 324 points at 15° intervals in Φ/Ψ space from 0° < Φ < 180° and −165° < Ψ < 180° were calculated, allowing the generation of a full symmetric surface with 576 points for *N*-formylglycine amide. These computations including geometry optimization and shift tensor calculations took on average 2.8 h on 2 processors in parallel on the Beowulf cluster.[43] On the *N*-formyl-L-leucine amide surface 288 points on a rectilinear grid with 21.2° intervals in Φ/Ψ space were calculated. These computations including geometry optimization and shift tensor calculations took on average 22.9 h on 2 processors in parallel on the Origin

(28) McGeorge, G.; Hu, J. Z.; Mayne, C. L.; Alderman, D. W.; Pugmire, R. J.; Grant, D. M. *J. Magn. Reson.* **1997**, *129*, 134−144.
(29) Sethi, N. K.; Alderman, D. W.; Grant, D. M. *Mol. Phys.* **1990**, *71*, 217−238.
(30) Gan, Z. *J. Am. Chem. Soc.* **1992**, *114*, 8307−8309.

(31) Rabbani, S. R.; Edmonds, D. T.; Gosling, P.; Palmer, M. H. *J. Magn. Reson.* **1987**, *72*, 230−237.
(32) Edmonds, D. T.; Speight, P. A. *Phys. Lett.* **1971**, *34A* (6), 325−326.
(33) Torrent, M.; Mansour, D.; Day, E. P.; Morokuma, K. *J. Phys. Chem. A* **2001**, *105*, 4546−4557.
(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, G. E.; 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.; Salvador, P.; Dannenberg, J. J.; 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.; 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.11; Gaussian, Inc.: Pittsburgh, PA, 2001.
(35) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(36) Becke, A. D. *Phys. Rev.* **1988**, *A 38*, 3098−3100.
(37) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(38) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.
(39) Dunning, T. H., Jr.; Hay, P. J. In *Modern Theoretical Chemistry*; Schaefer, H. F., III, Ed; Plenum: New York, 1976; pp 1−28.
(40) London, F. *J. Phys. Radium* **1937**, *8*, 397−409.
(41) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789−807.
(42) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2000; p 125.
(43) Times for individual computations varied substantially from the average due to the heterogeneous nature of the Beowulf cluster.

2000. For the L-proline residue model GIAO shift calculations were performed on a geometry optimized L-proline amide fragment at the B3LYP/D95** level of theory. Due to the rigid structure of proline and since it constitutes the N-terminus of the tripeptide only variations in Ψ were sampled every 30°.

To test the effect of different levels of theory on the geometry optimization and the theoretical shift tensors, the geometry of the *N*-formyl-L-leucine amide model fragment was optimized using four different methods: B3LYP/D95**, HF/3-21G,[44–47] AM1,[48–52] and AMBER.[53] The effect of truncation of the model was tested on the *N*-formylglycine amide model fragment.

Isolated molecule (IM) and EIM[19] shielding calculations were performed on the melanostatin X-ray geometry. Prior to shielding calculations all hydrogen positions as well as the proline C^δ and C^γ positions were optimized at the B3LYP/D95** level of theory, as their position were insufficiently determined by the X-ray analysis.[20,21] The EIM utilized the Ewald program described previously.[54,55] To accommodate the large unit cell dimensions of melanostatin the Ewald summation parameters had to be adjusted from previously reported values to obtain convergence of the Ewald sum that calculates the lattice potential.[55] A convergence parameter of $\alpha = 0.1$ and summation limits of ± 16 and ± 10 unit cells in all crystallographic directions for the real and reciprocal sums, respectively, were found sufficient. Convergence of the Ewald sum was assumed when the calculated lattice potential was within 10 nV upon changes of the convergence parameter α and increase of the summation limits.

Parameters of the Ewald calculations were as follows: $N_t = 76\,624$, $N_1 = 300$, $N_{uc} = 535$. The Ewald potential inside Zone 1 was reproduced during the iterations by the N_t point charges with an rms of less than 10 μV. Partial atomic charges for the molecule of interest were obtained using the natural population analysis (NPA) as implemented in *Gaussian 98*.[56–62] Self-consistency of the point charge array in the EIM was assumed when the change in NPA partial atomic charges of the molecule of interest was $< 0.001\,e$ between iterations. This procedure was shown to give equivalent shielding tensors in amino acids as the EIM utilizing B3LYP with 6-311G**[63,64] basis set and NPA charge calculations as well as B3LYP with D95** basis set utilizing partial atomic charges fitted to reproduce the electrostatic of the molecule of interest (ChelpG[65]).[24]

All calculated ^{13}C shieldings were converted to the TMS shift scale assuming a shift of the bare nucleus of 187 ppm and a linear scaling

factor of -1. This conversion is taken from shielding shift correlations reported for several amino acids.[24] All calculated ^{15}N shieldings were converted to the nitromethane shift scale using the reported shift of the bare nucleus of -135.8 ppm and a linear scaling factor of -1.[66,67] All comparisons between theoretical shift tensor surfaces and experimental shift tensors are made in the icosahedral representation.[68]

Comparing Chemical Shift Tensors. A faithful comparison between two chemical shift tensors is achieved by calculating the scalar distance between them, which is defined as the norm of the quadratic difference

$$d^2 = \frac{\int |f(x) - g(x)|^2 \mathrm{d}x}{\int \mathrm{d}x} \tag{1}$$

where the integrals are over the functions domain. Alderman et al.[68] have shown that the second rank chemical shift tensor may be expressed in an icosahedral representation so that the distance squared between two symmetric tensors δ and ϵ is easily obtained by 1/6 of the sum of square differences of their icosahedral components (a–f)[68,69]

$$d^2 = \frac{1}{6}[(\delta_a - \epsilon_a)^2 + (\delta_b - \epsilon_b)^2 + (\delta_c - \epsilon_c)^2 + (\delta_d - \epsilon_d)^2 + (\delta_e - \epsilon_e)^2 + (\delta_f - \epsilon_f)^2] \tag{2}$$

The distance squared may be used to define a figure-of-merit function for comparison of experimental shift tensors with theoretical shift tensors of a molecular model with adjustable structural parameters. For a molecular system, shift tensors can be obtained for all atomic positions and hence a shift function for a molecule may be defined as a vector with $6 \cdot N$ elements, where N is the number of atomic positions of a given nucleus, for example ^{13}C

$$\vec{S}_C = (\delta_a{}^1 \cdots \delta_f{}^1 \delta_a{}^2 \cdots \delta_f{}^2 \cdots \delta_a{}^{Nc} \cdots \delta_f{}^{Nc}) \tag{3}$$

Similarly the agreement between two molecular shift tensor functions S and T may be quantized by the distance squared between the two corresponding vectors

$$d^2 = \frac{1}{6N_C}|\vec{S}_C - \vec{T}_C|^2 = \frac{1}{N_C}\sum_{i=1}^{N_C}\frac{1}{6}\sum_{k=a}^{f}(\delta_k^i - \epsilon_k^i)^2 \tag{4}$$

Similar expressions are obtained for other nuclei encountered in a molecule such as ^{15}N, ^{1}H, etc. Although it is not necessary to separate the molecular shift function for the various nuclei, it is beneficial for normalization. In a fitting procedure that fits the experimentally determined shift tensors the model shift tensor functions are calculated from a theoretical molecular model that depends on certain molecular structural parameters. A commonly used figure-of-merit function for this type of linear and nonlinear fitting problems is chi squared (χ^2),[70] which is defined for the most general case as follows[71]

$$\chi^2 = \sum_{i=1}^{K}\frac{[y_i - y_i(a_1;a_2;...;a_M)]}{\sigma_i^2} \tag{5}$$

Here, y_i is the observable from the experiment, $\{a_1;a_2;...;a_M\}$ are the adjustable parameters of the model, σ_i is the standard deviation of

(44) Roothan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
(45) McWeeny, R.; Dierksen, G. *J. Chem. Phys.* **1968**, *49*, 4852−4856.
(46) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797−2803.
(47) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939−947.
(48) Dewar, M.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.
(49) Davis L. P.; Guidry, R. M.; Williams, J. R.; Dewar, M. J. S.; Rzepa, H. S. *J. Comput. Chem.* **1981**, *2*, 433−445.
(50) Dewar, M. J. S.; McKee, M. L.; Rzepa, H. S. *J. Am. Chem. Soc.* **1978**, *100*, 3606−3607.
(51) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(52) Dewar, M. J. S.; Reynolds, C. H. *J. Comput. Chem.* **1986**, *2*, 140−143.
(53) Cornell, W. D.; Cieplak, P.; Baylay, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5193.
(54) Derenzo, S. E.; Klintenberg, M. K.; Weber, M. J. *J. Chem. Phys.* **2000**, *112* (5), 2074−2081.
(55) Klintenberg, M.; Derenzo, S. E.; Weber, M. J. *J. Comput. Phys. Commun.* **2000**, *131*, 120−128.
(56) Carpenter, J. E.; Weinhold, F. *J. Mol. Struct.* **1988**, *169*, 41−62.
(57) Carpenter, J. E. Ph D Thesis, University of Wisconsin, 1987.
(58) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(59) Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066−4073.
(60) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 1736−1740.
(61) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.
(62) Weinhold, F.; Carpenter, J. E. *The Structure of Small Molecules and Ions*; Plenum: New York, 1988; p 227.
(63) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639−5648.
(64) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.
(65) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.

(66) Jameson, C. J.; Mason, J. In *Multinuclear NMR*; Mason, J., Ed.; Plenum Press: New York, 1987; p 336.
(67) Jameson, C. J.; Jameson, A. K.; Oppusunggu, D.; Wille, S.; Burrell, M.; Mason, J. *J. Chem. Phys.* **1981**, *74* (1), 81−88.
(68) Alderman, D. W.; Sherwood, M. H.; Grant, D. M. *J. Magn. Reson. A* **1993**, *101*, 188−197.
(69) In the original paper, the icosahedral components are numbered 1 through 6, here these components are designated from a to f to avoid confusion with the Cartesian representation of principal shift values that are defined from 11 to 33.

Table 2. Experimental and Theoretical Icosahedral Shift Components of Melanostatin/ppm

			experimental			isolated molecule				EIM			
		δ_a	δ_c	δ_e	δ_{iso}	δ_a	δ_c	δ_e	δ_{iso}	δ_a	δ_c	δ_e	δ_{iso}
Gly	C′	228.5	161.6	133.5	174.5	226.5	137.8	132.1	165.5	224.4	161.2	126.5	170.7
	Cα	62.6	36.9	35.7	45.1	65.9	35.0	35.1	45.3	66.8	35.0	35.5	45.8
Leu	C′	234.0	161.8	135.3	177.0	233.4	148.7	133.7	171.9	231.1	161.4	130.5	174.3
	Cα	63.2	51.8	45.5	53.5	65.9	51.4	47.9	55.1	66.3	52.9	49.0	56.1
	Cβ	52.8	38.0	32.1	41.0	57.7	40.8	23.7	40.7	57.9	40.4	31.4	43.3
	Cγ	28.4	22.2	22.8	24.5	32.9	24.8	18.6	25.4	28.8	22.3	22.0	24.4
	Cδ1	37.8	21.4	11.8	23.7	37.5	19.2	5.1	20.6	38.6	21.5	7.7	22.6
	Cδ2	34.6	22.8	13.2	23.5	44.4	21.5	7.1	24.3	33.5	21.3	9.8	21.5
Pro	C′	234.5	169.2	135.5	179.7	235.9	163.0	134.9	177.9	231.0	171.1	132.8	178.3
	Cα	74.8	56.6	45.2	58.9	77.9	56.7	39.3	57.9	76.0	58.8	42.3	59.0
	Cβ	47.2	20.3	25.2	30.9	53.0	15.9	7.4	25.4	50.3	19.6	24.4	31.5
	Cγ	38.1	20.8	22.6	27.2	41.2	12.4	21.0	24.9	41.2	18.2	23.8	27.7
	Cδ	66.7	41.0	35.4	47.7	73.8	38.1	32.2	48.0	67.9	42.3	33.8	48.0
Gly	N-1	−210.9	−325.7	−292.1	−276.2	−249.2	−355.6	−342.3	−315.7	−218.3	−332.3	−303.9	−284.9
	N	−205.8	−313.3	−283.1	−267.4	−222.0	−338.6	−322.2	−294.3	−203.9	−313.6	−283.9	−267.1
Leu	N	−191.2	−301.3	−274.1	−255.5	−201.3	−316.0	−310.2	−275.8	−186.7	−302.5	−285.1	−258.1
Pro	N	−305.4	−347.6	−354.6	−335.9	−315.3	−378.7	−372.4	−355.5	−312.5	−363.5	−359.8	−345.3

observation i, and K is the number of observations. For fitting molecular shift tensor functions, χ^2 can be expressed in terms of the d^2 for the different nuclei and their respective standard deviations, σ

$$\chi^2 = \frac{1}{\sigma_c^2}\sum_{i=1}^{N_c}\sum_{k=a}^{f}(\delta_k^i - \epsilon_k^i(a_1;a_2;...;a_M))^2 +$$

$$\frac{1}{\sigma_N^2}\sum_{i=1}^{N_N}\sum_{k=a}^{f}(\delta_k^i - \epsilon_k^i(a_1;a_2;...;a_M))^2$$

$$= \frac{|\vec{S}_C - \vec{T}_C|^2}{\sigma_C^2} + \frac{|\vec{S}_C - \vec{T}_C|^2}{\sigma_N^2}$$

$$= \frac{6N_C}{\sigma_C^2}d_C^2(a_1;a_2;...;a_M) + \frac{6N_N}{\sigma_N^2}d_C^2(a_1;a_2;...;a_M) \quad (6)$$

In addition to the adjustable structural parameters such as the Ramachandran angles Φ and Ψ, the rotation that relates the coordinate systems of the two compared tensors is often only partially known or completely unknown from an experiment on a powdered solids. When one or two parameters of the rotation are unknown then the rotational parameters are included in the fitting parameters that minimize the function of merit. In the case where the rotation is completely unknown, it has been shown that a distance based figure-of-merit function is minimized when the principal axes of the two tensors are collinear and the principal components associated with each axis are ordered by magnitude.[68]

Once the problem is cast in terms of the known χ^2 statistic, goodness of fit criteria and confidence limits are easily derived if the distributed errors are assumed to be normal. The goodness of fit can be estimated from χ^2_{min} whereas the confidence limits may be obtained from constant $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ boundaries.[71] When the errors are governed by systematic errors (when an insufficient model is assumed) then estimates of confidence limits become inaccurate and outliers become more probable.

In this study, the resulting discretely sampled χ^2 Ramachandran surfaces are expanded in a complex 2D Fourier series of the form

$$\chi^2 = \sum_{n=0}^{N}\sum_{m=0}^{M}A_{nm}e^{i(n\Phi+m\Psi)} \text{ where } A_{nm} = A_{-n-m}^* \quad (7)$$

Table 3. Orientation of the CN Bond Vector in the Chemical Shift Principal Axis System (deg)

		experimental			isolated molecule				EIM			
		<11	<22	<33	<11	<22	<33	<exp[a]	<11	<22	<33	<exp[a]
Gly	C′	43.2	46.9	90.0[b]	34.3	55.7	89.5	8.9	40.7	49.3	89.9	2.5
	Cα	82.1	43.0	48.1	82.0	55.0	36.2	12.0	84.6	50.2	40.3	7.9
Leu	C′	33.4	66.6	90.0[b]	27.0	63.0	90.4	1.7	26.8	63.2	90.1	1.3
	Cα	84.2	39.8	51.8	80.5	43.1	48.5	5.0	82.0	41.2	50.0	2.7
Pro	C′	31.1	68.9	90.0[b]	23.0	67.1	88.7	1.3	23.9	66.1	89.0	1.5
	Cα	24.8	66.3	88.9	82.4	79.4	23.1	77.0	86.5	27.5	72.9	63.4
	Cδ	75.5	52.1	41.6	87.9	85.9	4.6	37.0	82.1	89.0	8.0	38.2

[a] Angle between experimental and theoretical CN bond vectors assuming collinear chemical shift principal axis system. [b] This angle was set to the reported value in the fitting procedure because of local symmetry considerations.

The minimum is determined from the continuous expansion. The relevant $\Delta\chi^2$ surfaces are generated by subtracting the minimum χ^2_{min}.

Results and Discussion

¹³C and ¹⁵N Chemical Shift Principal Values. The experimental chemical shift principal values are given in Table 2. The full FIREMAT spectrum recorded at 9.4 T on natural abundance ¹³C in Figure 3 reveals that inhomogeneously broadened lines in the form of 2:1 doublets are observed for the ¹⁴N dipolar coupled Cα and C′ resonance lines. This demonstrates that the ¹⁴N quadrupolar coupling must be included in the spectral analysis to extract optimal principal shift values.[18] The orientations of the CN bond vector in the chemical shift principal axis system obtained from including the ¹³C−¹⁴N dipolar coupling in the spectral analysis are given in Table 3; the experimental dipolar coupling constants are shown in Table 4. The narrow lines, due to the crystallinity of the sample, are sufficient to resolve the two Cδ methyl groups in the leucine residue with an isotropic shift difference of only 0.2 ppm. Hence, all ¹³C positions are well resolved and allow for the determination of the principal shift values. Spectra recorded at 4.7 T have more pronounced 2:1 doublets in the Cα and C′ isotropic shift regions

(70) This figure-of-merit function, χ^2, should not be confused with the quadrupolar coupling constant, χ, which only occurs in first power, whereas the figure-of-merit function χ^2 appears always squared.

(71) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling W. T. *Numerical Recipies in FORTRAN*; Cambridge University Press: Cambridge 1989; p 510.
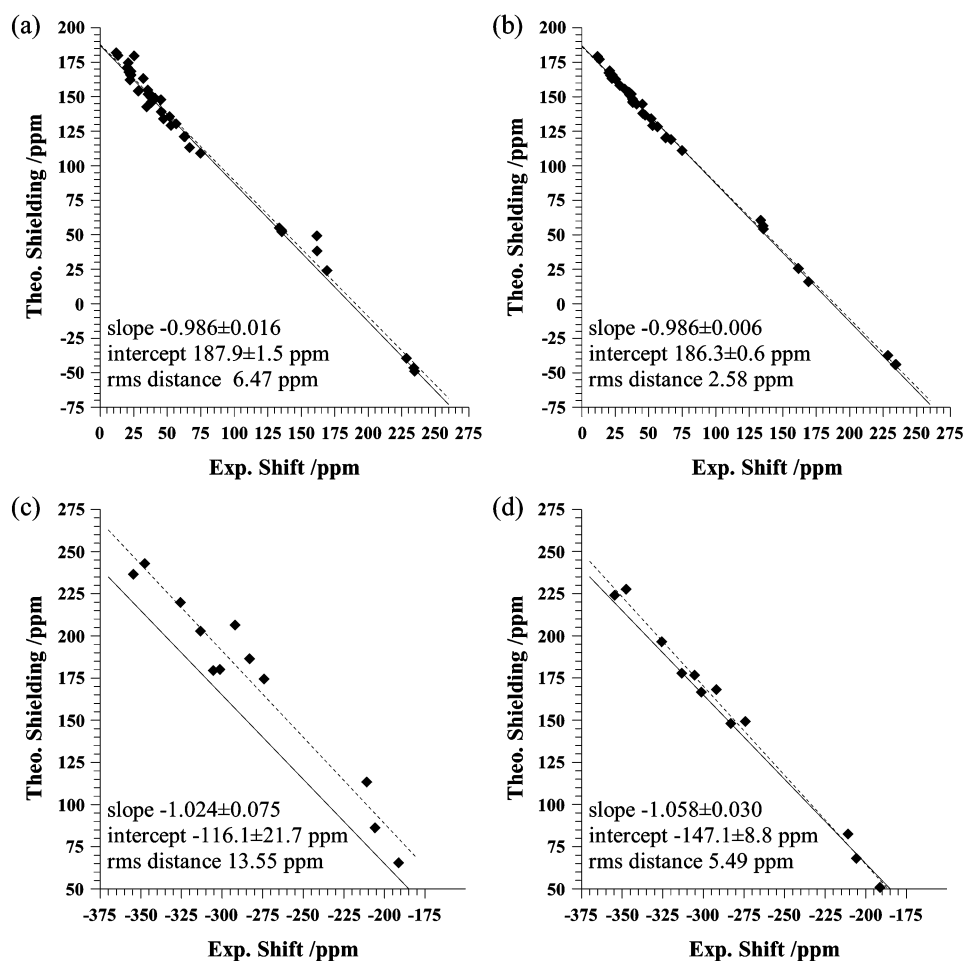
Figure 4. Correlation between icosahedral theoretical shieldings and experimental shifts. The best fit lines to the data are shown as dashed lines with slopes, intercepts and rms distances reported in the plot, solid lines represent the expected lines using the shielding shift conversion discussed in the text. (a) ^{13}C IM calculation. (b) ^{13}C EIM calculation. (c) ^{15}N IM calculation. (d) ^{15}N EIM calculation.

Table 4. Dipolar Coupling Constants and CN Bond Lengths Obtained from X-ray Crystallography and from FIREMAT Data

		X-ray		FIREMAT	
		r C–N/Å	D/Hz	r C–N/Å	D/Hz
Gly	**C′–N-1**	1.333	923	1.36	862
	C$^\alpha$–N	1.442	728	1.45	712
Leu	**C′–Gly N**	1.342	903	1.31	978
	C$^\alpha$–N	1.458	705	1.46	707
Pro	**C′–Leu N**	1.341	905	1.34	915
	C$^\alpha$–N	1.465	695	1.44	739
	C$^\delta$–N	1.476	679	1.52	618

and unambiguously reveal the nature of the splittings at 9.4 T. Spectral assignment is based on the EIM shielding calculations and can be made with >94% confidence for the C$^\alpha$ carbons and with >83% confidence for the C$^\delta$ positions in the leucine side chain using *F*-test statistics. Close inspection of the spectra reported by Garbow and Werther confirms the assignment for the C$^\alpha$ and C′ positions.[72] Using *F*-test statistics the ^{15}N assignment based on the EIM shielding calculations can be made with a confidence of >97%.

Intermolecular Effects–EIM Calculations. In the EIM calculations intermolecular interactions in the crystal are approximated with an electrostatic lattice potential that is calculated from Ewald lattice sums. This lattice potential is reproduced in

the EWALD program by an array of point charges at crystallographic positions. When an isolated molecule is placed inside the array of point charges all intermolecular hydrogen bonds are described electrostatically. This procedure was shown to give accurate ^{13}C shift tensors in hydrogen bonded zwitterionic amino acid crystals with an rms distance of 4.1 ppm at the B3LYP/D95** level of theory.[24] The correlation of the ^{13}C theoretical icosahedral shieldings for an IM melanostatin with the experimental icosahedral shifts is shown in Figure 4a, whereas the correlation for the EIM calculated shieldings is shown in Figure 4b. The resulting theoretical shifts are reported in Table 2. The agreement between theoretical and experimental shift tensors is dramatically improved when intermolecular interactions are incorporated in the calculations, with the rms distance decreasing from 6.47 to 2.58 ppm. As expected, the δ_c components of the C′ icosahedral shifts are severely affected by intermolecular hydrogen bonding. Again, the icosahedral representation is used to increase the statistical sensitivity and changes are observed only in δ_c as it incorporates δ_{11} and δ_{22}. Opposite shifts upon H-bonding were observed in δ_{11} and δ_{22}. Hence, the effect in the δ_c component of the icosahedral representation is less pronounced than in δ_{22} in the Cartesian representation.[82] Nevertheless, a considerable improvement for the EIM calculated shieldings over the isolated molecule

(72) Garbow, J. R.; McWerther, C. A. *J. Am. Chem. Soc.* **1993**, *115*, 238−244.

(73) Stueber, D.; Grant, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 10 539−10 551.

Table 5. Effects of Molecular Truncation and Geometry Optimization on Structure and C^α Theoretical Chemical Shifts in *N*-Formyl-glycine Amide

	$\Phi/°$	$\Psi/°$	CN/Å	CC/Å	CCN/°	CH/Å	δ_a/ppm	δ_c/ppm	δ_e/ppm
experimental	69.5	18.3	1.442	1.514	115.2	-	62.6	36.9	35.7
MIF/H-opt/EIM[a]	69.5	18.3	1.442	1.514	115.2	1.094	66.8	35.0	35.5
MIF/H-opt/IM[b]	69.5	18.3	1.442	1.514	115.2	1.094	65.9	35.0	35.1
model/X-ray[c]	69.5	18.3	1.442	1.514	115.2	1.094	65.4	34.4	33.9
model/X-ray/opt[d]	69.5	18.3	1.459	1.541	117.4	1.094	67.9	35.4	37.1
model/75/15[e]	75	15	1.458	1.541	117.3	1.094	67.8	35.2	36.3
model/75/30[f]	75	30	1.458	1.541	115.9	1.094	67.6	35.4	38.9

[a] MIF/H-opt/EIM: MIF, hydrogen optimized at B3LYP/D95**, EIM calculated shifts. [b] MIF/H-opt/IM: MIF, hydrogen optimized at B3LYP/D95**, IM calculated shifts. [c] Model/X-ray: *N*-formylglycine amide with MIF geometry for all common bonds. Open valences are capped with hydrogens. All hydrogen positions are optimized at the B3LYP/D95** level. [d] Model/X-ray/opt: *N*-formylglycine amide model with MIF geometry for all common bonds. Open valences are capped with hydrogens. Full geometry optimization with fixed Φ and Ψ dihedral angles at the B3LYP/D95** level. [e] Model/75/15: *N*-formyl-glycine amide model with $\Phi = 75°$ and $\Psi = 15°$. Full geometry optimization with fixed Φ and Ψ dihedral angles at the B3LYP/D95** level. [f] Model/75/30: *N*-formylglycine amide model with $\Phi = 75°$ and $\Psi = 30°$. Full geometry optimization with fixed Φ and Ψ dihedral angles at the B3LYP/D95** level.

calculations at this position is obtained. Even though sp^3 carbon shift tensors are considered to be less sensitive to intermolecular interactions, they also are improved significantly when the lattice effects are included in the calculations. For example, the IM and EIM calculated icosahedral shift components of Leu C^β, Pro C^β, and Pro C^γ differ by up to 17 ppm in Table 2. This has important implications for the structure determination using sp^3 shift tensors, because intermolecular interactions are difficult to measure and are generally neglected for simplification.

The CN bond vector orientations in the chemical shift principal axis system are compared between theory and experiment in Table 3. The agreement between theory and experiment is excellent except for the C^α and C^δ positions in the proline residue. This is most likely due to a less accurate estimate for the quadrupolar coupling tensor for the N position in the proline since quadrupolar coupling constants for similar compounds are unavailable in the literature. Fortunately, the shift tensor components are less sensitive to errors in the quadrupolar coupling constants and reliable shift values are determined from the fitting.[18] The effect of intermolecular interactions on the shift tensor orientation is less than 7°, with the exception of the Pro C^α shift tensor, where a 52° difference is observed between the IM and EIM calculations. This is presumably due to the fact that the Pro N lone pair is involved in hydrogen bonding.

The results of the EIM and IM ^{15}N shielding calculations are also summarized in Figure 5c,d and Table 2. In accordance with previous results on shielding calculations in several amino acids, large improvements in the agreement between theory and experiment are observed when intermolecular interactions are included in the calculations using point charge arrays that faithfully reproduce the electrostatic lattice potential.[73,24] Quan-

tum mechanical calculations of intermolecular hydrogen bonds have been shown to produce even better agreement between experiment and theory,[24] but these are not feasible, because of the size of the system and the number of intermolecular hydrogen bonds that exist. Nevertheless, the EIM shielding calculations provide sufficient agreement to aid the discussion of intermolecular effects on the ^{15}N shielding calculations on test structures. The magnitude of the improvements seen for nitrogen shift tensors, when intermolecular interactions are included in the calculations, reveals the high sensitivity of the nitrogen shifts on electrostatic forces. The largest difference between EIM and isolated molecule shielding calculations is 38 ppm found for the δ_c component in Gly N, the second smallest difference, however, is already more than 12 ppm. This sensitivity renders the nitrogen shifts a superb probe of molecular conformation as well as intermolecular arrangement in the crystal, provided that the molecular model is detailed enough to faithfully reproduce the shift tensor structure relationship.

Simplification of the Molecular Model. Before a structure elucidation based on the chemical shift tensor can be attempted, the effect of truncation of the molecular model and optimization of bond lengths and angles on the calculated shift tensors has to be investigated. The alternative of sampling the full geometrical space using a complete molecule and including neighboring molecules in the molecular model becomes increasingly difficult and prohibitively time-consuming. Omitting the geometry optimization in shift tensor surface calculations by fixing the local geometry to average X-ray bond lengths and angles might result in unrealistic steric interaction as the system is not allowed to relax. Thus, geometry optimization has to be an essential step in building a molecular model to obtain a reduced set of geometric parameters.

The components of the icosahedral shift are compared in Table 5 for different models for the glycine residue in melanostatin. By comparing the ^{13}C shifts for MIF/H-opt/IM and Model/X-ray it is apparent that truncation of the molecular model results only in differences of less than 1.2 ppm. The difference between MIF/H-opt/EIM and MIF/H-opt/IM, which could be attributed to intermolecular interactions, are less than 0.9 ppm. Optimization of all heavy atom bond lengths and angles, while fixing the Ramachandran angles Φ and Ψ, results in a large effect on the shift tensor of up to 3.2 ppm in the components (by comparison of Model/X-ray and Model/X-ray/opt). Geometry optimization yields a gas-phase geometry which

(74) *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, 1996; pp 9-2 and pp 9-24.
(75) Jeffrey, G. A.; Ruble, J. R.; McMullan, R. K.; DeFrees, D. J.; Binkley, J. S.; Pople, J. A. *Acta Crystallogr.* **1980**, *B36*, 2292–2299.
(76) Jeffrey, G. A.; Ruble, J. R.; McMullan, R. K.; DeFrees, D. J.; Pople, J. A. *Acta Crystallogr.* **1981**, *B37*, 1381–1387.
(77) Jeffrey, G. A.; Ruble, J. R.; McMullan, R. K.; DeFrees, D. J.; Pople, J. A. *Acta Crystallogr.* **1981**, *B37*, 1885–1890.
(78) Iuliucci, R. J.; Facelli, J. C.; Alderman, D. W.; Grant, D. M. *J. Am. Chem. Soc.* **1995**, *117*, 2336–2343.
(79) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908–5920.
(80) Chekmenev, E. Y.; Xu, R. Z.; Maeshuta, M. S.; Wittebort, R. J. *J. Am. Chem. Soc.* **2002**, *124*, 11 894–11 899.
(81) Grant, D. M.; Gunneau, F.; Hu, J. Z.; Alderman, D. W.; Pugmire, R. J. Experimental NMR Conference, Asilomar 2000.
(82) Gu, Z.; Zambrano, R.; McDermott, A. *J. Am. Chem. Soc.* **1994**, *116*, 6368–6372.
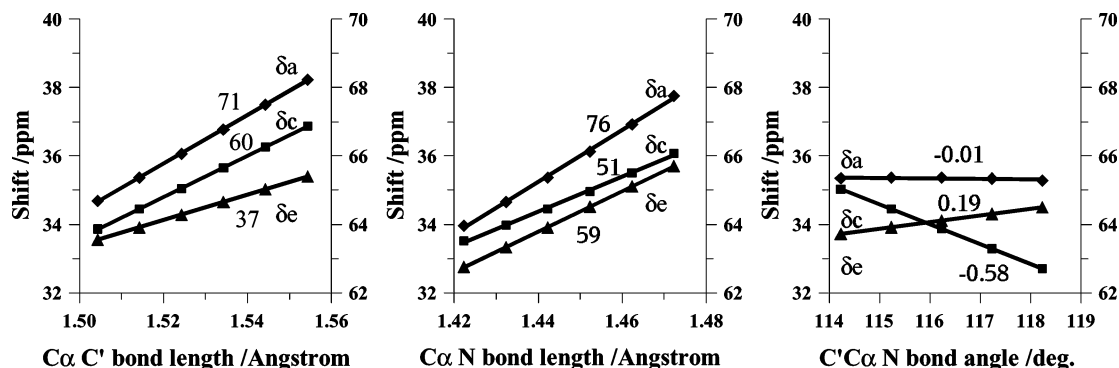
Figure 5. Shielding gradients of the icosahedral C^α shift components calculated for the *N*-formylglycine amide model. Shift gradients are given in ppm/Å and ppm/°. The δ_a components are given relative to the right axis, δ_c and δ_e relative to the left axis.

Table 6. Effects of Geometry Optimization on Structure and C^α Theoretical Chemical Shift in *N*-Formyl-L-leucine Amide

|	Φ/°	Ψ/°	CN/Å	CC/Å	CCN/°	CH/Å	δ_a/ppm	δ_c/ppm	δ_e/ppm
experimental	−59.7	126.9	1.458	1.520	108.2		63.2	51.8	45.5
MIF/H-opt/EIM			1.458	1.520	108.2	1.097	66.3	52.9	49.0
MIF/H-opt/IM			1.458	1.520	108.2	1.097	65.9	51.4	47.9
B3LYP/D95**	−60	120	1.465	1.548	110.2	1.095	67.1	52.9	47.8
HF/3-21G			1.463	1.534	109.1	1.078	64.4	51.9	46.9
AM1			1.445	1.552	110.9	1.132	67.6	54.9	49.9
AMBER			1.466	1.543	112.1	1.090	66.7	51.1	50.3
B3LYP/D95**	105	−135	1.471	1.552	106.7	1.099	70.7	55.4	53.7
HF/3-21G			1.478	1.536	106.1	1.074	71.9	51.1	52.3
AM1			1.443	1.554	109.3	1.140	69.6	56.7	55.9
AMBER			1.470	1.542	108.3	1.089	72.1	50.6	52.6
B3LYP/D95**	−60	−90	1.470	1.554	112.7	1.094	68.1	56.8	44.5
HF/3-21G			1.468	1.539	112.1	1.079	64.5	54.1	42.9
AM1			1.440	1.550	113.3	1.139	70.5	55.6	45.6
AMBER			1.467	1.546	112.9	1.090	66.1	56.1	49.0
B3LYP/D95**	30	150	1.476	1.549	116.2	1.095	78.7	54.5	50.4
HF/3-21G			1.469	1.531	115.9	1.078	74.9	51.8	47.5
AM1			1.444	1.549	114.7	1.133	74.8	53.1	52.7
AMBER			1.479	1.554	119.6	1.089	76.3	54.9	48.6

differs considerably from average heavy atom bond lengths that are found by X-ray in the solid state.[74−77] The bond lengths and bond angles in the vicinity of the important C^α are changed from the X-ray values by 0.027 Å for the CC bond length and 2.2° for the CCN bond angle. Effects in the shift components of uncertainties in bond lengths and angles are substantial; in acenaphthene, they have been calculated to be as large as 300 ppm/Å and 8.4 ppm/°.[78] In the *N*-formylglycine amide model fixed at X-ray geometry with optimized hydrogen positions gradients for the C^α icosahedral shift, components were calculated at the B3LYP/D95** level to be up to 75 ppm/Å for the CN and CC bonds and −0.54 ppm/° for the CCN angle as shown in Figure 5. These gradients, although smaller than for the sp^2 carbons in acenaphthene, reproduce the shift tensor differences between the truncated model using X-ray bond lengths and angles and the geometry optimized model. The magnitude of the changes in the chemical shift observed for truncation and optimization is of the order of changes observed for changes in Ψ of at least 15°. Because of fortuitous cancellation of errors, the truncated optimized model gives similar shifts as the EIM, an effect that will be discussed in more detail for ^{15}N below.

In this work, the highest feasible level of theory, B3LYP/D95**, was used for geometry optimizations as the computationally less demanding semiempirical and molecular mechanics approaches may not perform well in finding good geometries off the global energy minimum.[79] However, in previous

literature, studies of Ramachandran shift surfaces various approaches with different lower levels of theory have been used for the geometry optimizations.[2,8−12] Despite the deficiencies of these lower levels of theory, they are currently the only practical methods for large systems in simulated annealing and crystal structure prediction procedures. Although there have been studies dealing with the differences in optimized geometries at different levels of theory, there are no studies on the resulting effect on the calculated chemical shift tensor components. Therefore, to assess this effect, several points in the Φ/Ψ space of *N*-formyl-L-leucine amide were optimized with AMBER, AM1, and HF/3-21G and then these geometries were used for shift tensor calculations at the B3LYP/D95**level. The results are summarized in Table 6.

The effect caused by truncation and optimization using B3LYP/D95** with the Ramachandran angles found by X-ray is approximately 1.2 ppm, similar to the results found for *N*-formylglycine amide. Bond distances found for the different optimization methods differ by considerable amounts from bond lengths found by X-ray diffraction and B3LYP/D95**. These differences in the local geometry in the vicinity of C^α result in significant differences in the shift tensor components of up to 6.1 ppm for AM1 and AMBER optimized geometries at Φ = 105° and Ψ = −135°. These differences in shift tensor components caused by differences in local geometry can result in substantial differences in the resulting shielding surfaces, especially when shallow shift surfaces are encountered. Thus,
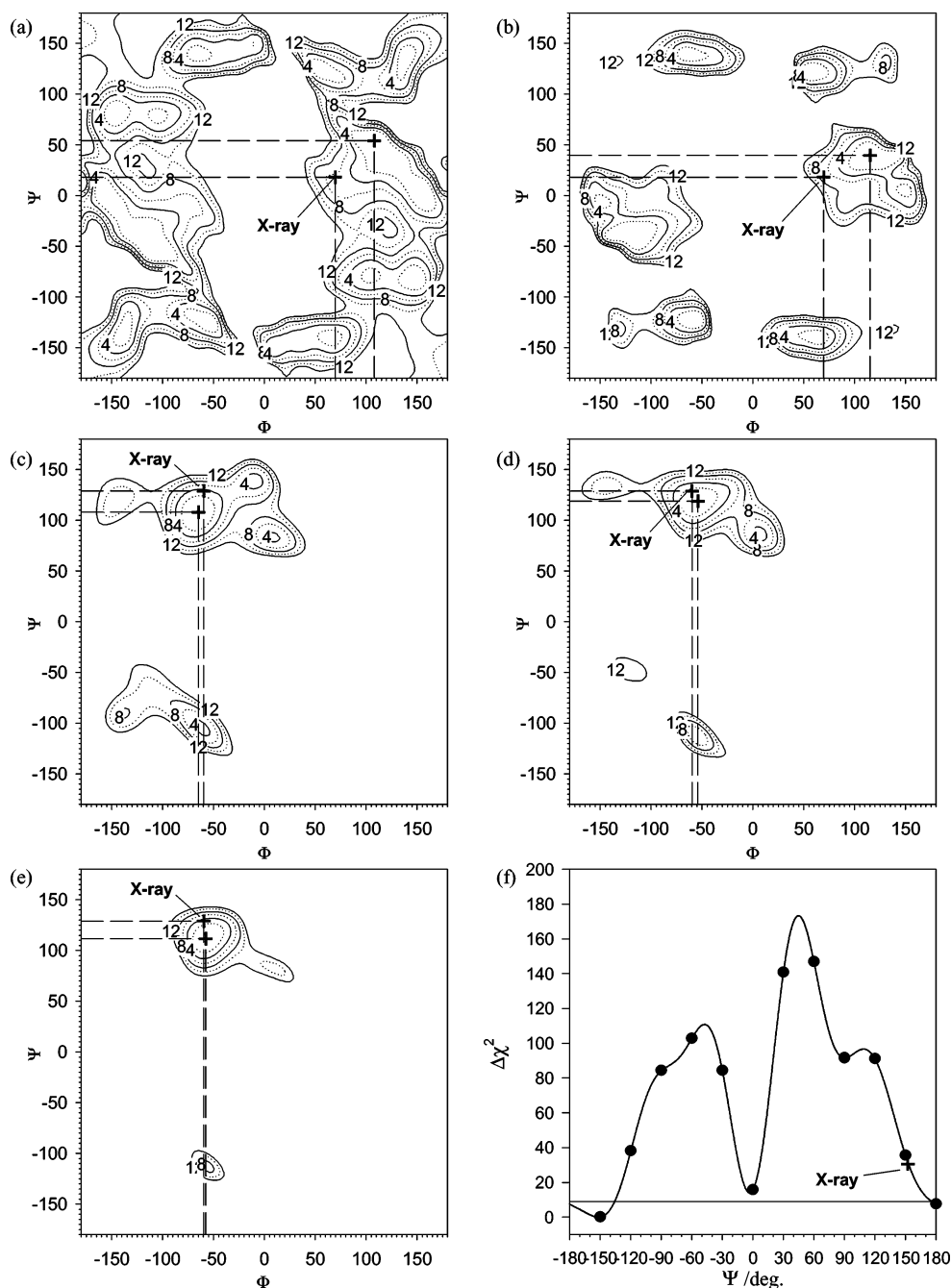
Figure 6. $\Delta\chi^2$ Ramachandran surfaces (a) Gly C^α shift principal values. (b) Gly C^α principal shift values and CN bond vector orientation. (c) Leu C^α principal shift values. (d) Leu C^α principal values and CN bond vector orientation. (e) As in (d) including C^β principal shift values. (e) Proline residue all aliphatic ^{13}C principal shift values. X-ray and global minima are indicated by + in (a−f).

differences in extracted Ramachandran angles may be observed when experimental data are compared to shielding surfaces that utilize different methods for geometry optimization.

Comparing Experimental Melanostatin and Model C^α Chemical Shift Surfaces. The comparison of the experimental C^α shift tensor components of the glycine and leucine residues and the theoretical model C^α shift tensors is presented in Figure 6a−e, where the $\Delta\chi^2$ Ramachandran surfaces are shown. The Ramachandran angles that give the minimum $\Delta\chi^2$ values are given in Table 7 along with the X-ray determined values. Assuming normal distributed errors the 68% confidence boundary is given for a two parameter fit by a 2.30 $\Delta\chi^2$ contour, whereas the 90% and 99% confidence boundary are given by

Table 7. Melanostatin Ramachandran Angles as Determined by X-ray and SSNMR

	X-ray		SSNMR	
	Φ (deg)	Ψ (deg)	ϕ (deg)	Ψ (deg)
Pro		152.4		−151.2
Leu	−59.7	128.9	−57.6	111.6
Gly	69.5	18.3	115.2	39.6

The X-ray leucine side chain angles are $\chi_1 = -61.4°$, $\chi_2 = -66.6°$, $\chi_3 = 170.2°$.

4.61 and 9.21 $\Delta\chi^2$ contours, respectively. These confidence boundaries are only an estimate as systematic errors due to the neglect of any lattice effects, truncation of the molecular model

and geometry optimization as discussed in the previous section may be quite significant. The only random error that is distributed in a normal manner is encountered in the original FID, from which the principal shift values are extracted and is less than 1 ppm for high signal-to-noise spectra and is small compared to the systematic differences.[80,81] The standard deviation needed in the χ^2 statistics (eq 6) is assumed to be 3.5 ppm. This value is obtained from the standard deviation for the correlation of theoretical shieldings and experimental sp^3 carbon shifts at the B3LYP/D95** level of theory using the EIM for a set of several amino acids and melanostatin.

Comparing only the C$^\alpha$ principal values of the shift tensor for glycine and leucine in Figure 6a,c results in several local minima in Φ/Ψ space that cannot be distinguished from the global minimum with high confidence. The area that gives good agreement for the measured C$^\alpha$ chemical shift components for the glycine residue is large due to the symmetry of the *N*-formylglycine amide model. On the *N*-formyl-L-leucine amide surface, Figure 6c only four distinct minima are found. When the orientational information on the CN bond vector in the chemical shift principal axis system is included in Figure 6b,d several of the local minima can be distinguished and the area that can be excluded with confidence is greatly enlarged in both surfaces. This results in six symmetry related minima on the glycine surface. The leucine surface is reduced to two low-lying minima when the orientational information is included in the comparison. In the case of leucine, the C$^\beta$ chemical shift tensor may also be included to further distinguish the Ramachandran surface as shown in Figure 6e. This effectively removes the ambiguities so that one unique set of Ramachandran angles is obtained. To find the Ψ angle of the proline residue, all proline sp^3 carbon shift tensors are compared to the model shifts in Figure 6f. This comparison results in one global minimum that can be distinguished from local minima with high confidence.

The X-ray geometries generally are within a reasonable $\Delta\chi^2$ boundary considering the simplification of the theoretical model and the resulting systematic errors. Determination of the glycine conformation is complicated by the fact that the C$^\alpha$ has local symmetry, resulting in a less distinct shift tensor surface. In addition the lack of a C$^\beta$ position prevents the comparison of additional data. Nevertheless, three of the six structures have reasonable glycine Φ angles. The other three conformations are related by the local symmetry around the C$^\alpha$.

The leucine Φ/Ψ angles are close to the X-ray values and the surface obtained including the orientational information Figure 6d and additionally the C$^\beta$ shifts in Figure 6e are well-defined with fewer probable local minima. The good agreement between X-ray and shift derived conformation for the leucine residue is a result of the relative insensitivity of the leucine C$^\alpha$ shifts to intermolecular effects and the lack of local symmetry at the C$^\alpha$ position. The insensitivity to intermolecular effects of the Leu C$^\alpha$ position is presumably due to its buried position inside the tripeptide.

The proline Ψ angle that is determined from the shift tensor components differs from the X-ray value by 56°. This discrepancy is mainly due to quite large intermolecular effects that are observed for the C$^\beta$ and C$^\gamma$ shifts (see Table 2). Fitting the experimental shift tensor components with a model that excludes these intermolecular interactions results in an excellent best fit.

The obtained Ψ angle, however, is considerably offset from the X-ray value.

Assessing Test Structures with ^{13}C and ^{15}N Shift Principal Values and Total Energy. Combining all dihedral angles determined from the Ramachandran surfaces yields six possible conformations of the glycine residue in the melanostatin tripeptide. Subjecting each of these melanostatin test structures to a geometry optimization where the backbone dihedral angles and the leucine side chain dihedral angles are fixed yields the six structures in Figure 7. Comparing all sp^3 ^{13}C shift tensors reveals similar rms distances ranging from 2.97 to 3.29 ppm and χ^2 values ranging from 23.2 to 26.6 (i.e., $\Delta\chi^2 < 3.4$) for these test structures and consequently they are indistinguishable with a high confidence. This is due to the fact that the glycine residue constitutes the C-terminus and lacks a side chain so that steric restrains imposed by change of the Φ and Ψ dihedral angles is small.

It is apparent that the agreement of experimental and theoretical shifts of these hypothetical structures is better than the theoretical calculations of the sp^3 carbons on an isolated molecule with X-ray conformation and optimized hydrogen positions. However, the accurate shift tensor calculation utilizing the EIM method on the X-ray structure results in an excellent agreement between experimental and theoretical shifts. This clearly indicates that most of the structural differences observed for the test structures result from the systematic errors due to neglect of any lattice effects and the slight discrepancies of bond lengths and angles due to geometry optimization. The sp^2 ^{13}C shift tensor components are omitted in this analysis, as they are known to be very sensitive to intermolecular effects such as hydrogen bonding and long-range electrostatic interactions.[43,82−84]
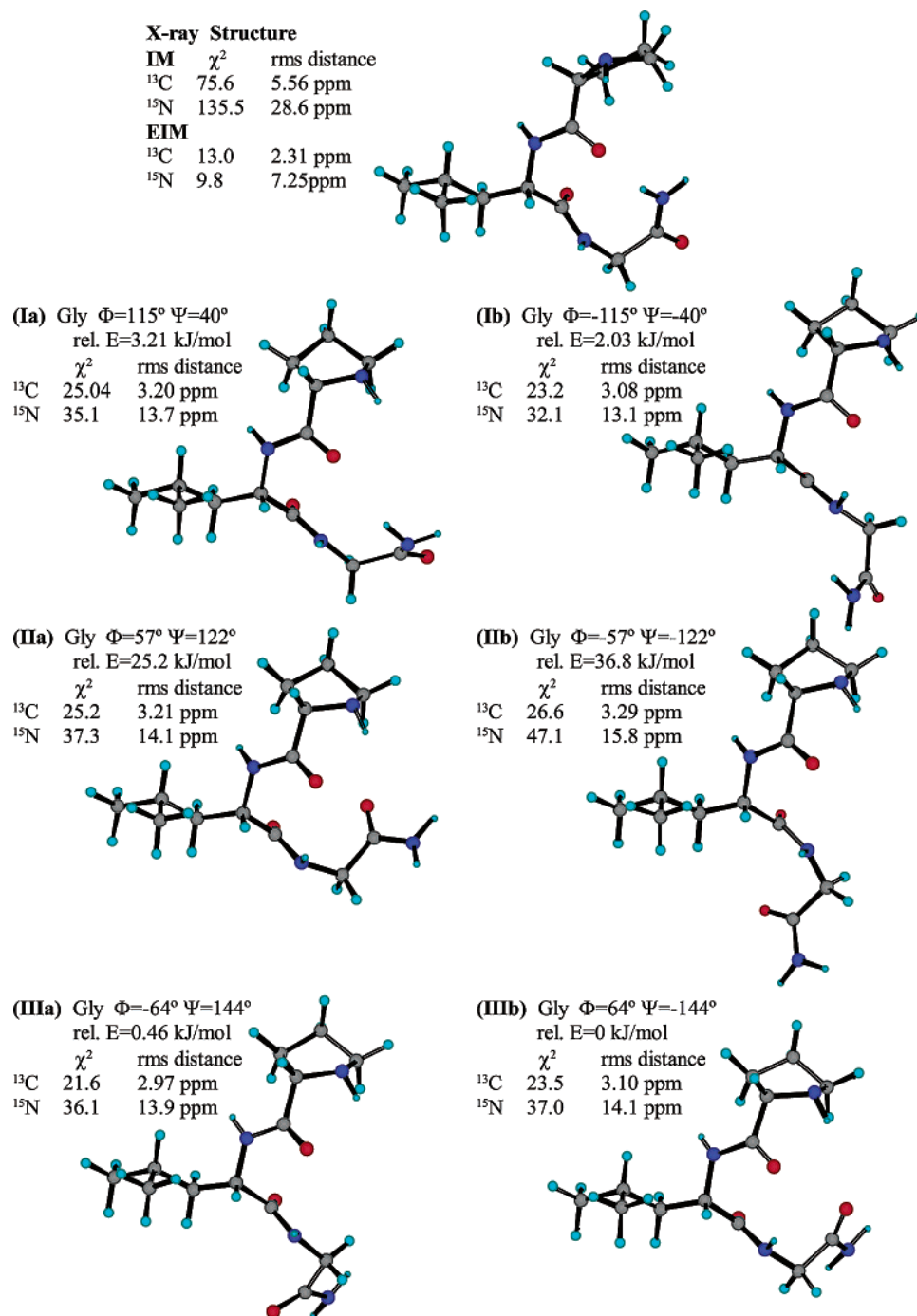
Possible molecular conformations may also be distinguished energetically, because only low energy conformations will be populated in solution and crystallize. Comparing the calculated energies of the test structures reveals energy differences between these conformations of nearly 36 kJ/mol with the lowest four conformations within 3.2 kJ/mol. Intermolecular interactions such as hydrogen bonding and crystal packing effects in the solid state, nevertheless, contribute to the overall conformational energy. These effects are neglected in the calculation and may alter the relative energies. The energies of moderate hydrogen bonds encountered in biological samples, are on the order of 16−50 kJ/mol.[85] Hence, a distinction of these test structures based upon energy considerations is impossible with any confidence.

The agreement between the experimental ^{15}N chemical shifts and the theoretical shifts calculated for the six test structures in Figure 7 is better than for the isolated molecule calculations on the X-ray structure. This artifact is presumably due to the slight change in local geometry, as the shift tensor gradients for the ^{15}N shift tensor components was calculated to be up to 301 ppm/Å for the Gly N−Formyl C′ bond in the *N*-formylglycine amide model. This large gradient has a profound effect on ^{15}N shielding calculations on test structures even when geometry optimized CN bond lengths resemble X-ray bond lengths relatively well. For example, the X-ray determined Gly N−Leu

(83) De Dios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1994**, *116*, 11 485−11 488.
(84) Facelli, J. C.; Gu, Z.; McDermott, A. *Mol. Phys.* **1995**, *86*, 865−872.
(85) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48−76.

X-ray Structure

IM	χ^2	rms distance
^{13}C	75.6	5.56 ppm
^{15}N	135.5	28.6 ppm
EIM		
^{13}C	13.0	2.31 ppm
^{15}N	9.8	7.25ppm

(Ia) Gly $\Phi=115°$ $\Psi=40°$
rel. E=3.21 kJ/mol

	χ^2	rms distance
^{13}C	25.04	3.20 ppm
^{15}N	35.1	13.7 ppm

(Ib) Gly $\Phi=-115°$ $\Psi=-40°$
rel. E=2.03 kJ/mol

	χ^2	rms distance
^{13}C	23.2	3.08 ppm
^{15}N	32.1	13.1 ppm

(IIa) Gly $\Phi=57°$ $\Psi=122°$
rel. E=25.2 kJ/mol

	χ^2	rms distance
^{13}C	25.2	3.21 ppm
^{15}N	37.3	14.1 ppm

(IIb) Gly $\Phi=-57°$ $\Psi=-122°$
rel. E=36.8 kJ/mol

	χ^2	rms distance
^{13}C	26.6	3.29 ppm
^{15}N	47.1	15.8 ppm

(IIIa) Gly $\Phi=-64°$ $\Psi=144°$
rel. E=0.46 kJ/mol

	χ^2	rms distance
^{13}C	21.6	2.97 ppm
^{15}N	36.1	13.9 ppm

(IIIb) Gly $\Phi=64°$ $\Psi=-144°$
rel. E=0 kJ/mol

	χ^2	rms distance
^{13}C	23.5	3.10 ppm
^{15}N	37.0	14.1 ppm

Figure 7. Comparison of X-ray and solid-state NMR derived conformations.

C′ bond length is 1.342 Å; in the six test structures, this bond length ranges from 1.363 to 1.386 Å. This 0.021 to 0.044 Å difference alone leads to a 6.3 to 13.2 ppm shift difference in the affected component between test structure and X-ray structure, respectively. Including intermolecular effects utilizing the EIM on the X-ray structure results in an excellent agreement between experiment and theory, reassuring the quality of the X-ray geometry.

This high sensitivity to the geometry makes the ^{15}N shifts a superb probe of molecular and electronic structure as well as intermolecular arrangement; the interpretation of the experimental shifts, however, becomes more challenging as intermolecular effects must also be included and accurate local geometries are necessary. Simplified molecular models are

insufficient to reproduce shifts for positions involved in intermolecular interactions so that the geometrical space cannot be readily explored. Therefore, the ^{15}N principal shift values and sp^2 ^{13}C principal shift values cannot be used to further refine the melanostatin conformation until more refined test structures that include intermolecular interactions and possibly crystallographic information are available. These test structures may in the future be provided by ab initio crystallographic prediction algorithms or from other experimental methods such as X-ray powder diffraction when single crystals are not available.[86-88]

Conclusions

Accurate ^{13}C and ^{15}N chemical shift tensor components are determined in melanostatin from natural abundance ^{13}C and ^{15}N

FIREMAT experiments. The high-resolution of the experimental ^{13}C FIREMAT spectrum requires the inclusion of the ^{13}C$-^{14}$N dipolar and the ^{14}N quadrupolar interactions in the spectral analysis. This allows for the determination of the CN bond vector in the chemical shift principal axis system of the coupled resonances. The structure of the sample was confirmed by single-crystal X-ray diffraction. A comparison of theoretical shift tensor calculations of an isolated molecule and a molecule embedded in a point charge array to mimic intermolecular interactions reveals sensitivity of even the sp^3 carbon shift tensor components to lattice effects. The effect observed for the C′ carbon shifts can be attributed to hydrogen bonding as reported before. The excellent agreement between EIM calculated ^{13}C and ^{15}N shieldings and experimental shifts confirm the high quality of the X-ray geometry.

N-formylamino acid amide model fragments are used to calculate Ramachandran shielding surfaces. Truncation of the model was shown to have only a minor effect on the shift tensors, whereas geometry optimization was shown to have a substantial effect on the shift tensor, which leads to a systematic error in the determination of the dihedral backbone angles Φ and Ψ. The backbone dihedral angles of melanostatin are determined from comparing C$^\alpha$ shift tensor components to the calculated surfaces from these model fragments using χ^2 statistics. Including the orientational information of the shift tensors in the comparison resolves ambiguities and increases the areas on the surface that may be rejected. The determination of the proline Ψ angle is impaired by the large intermolecular effects observed for the side chain carbon shift tensors resulting in a 56° discrepancy to the X-ray value as intermolecular effects are neglected in the model calculations. The determination of the glycine angles is complicated by the local symmetry around C$^\alpha$ so that the resulting surface contains several minima and is rather featureless. The conformation of the L-leucine residue determined from comparing C$^\alpha$ and C$^\beta$ shift tensors to the calculated model surfaces yields good agreement with X-ray values. Calculations performed on the test structures with different probable glycine angles all show good agreement for calculated and experimental sp^3 carbon shift tensor components. The ambiguities, nevertheless, in the glycine Φ and Ψ angles cannot be resolved.

In summary, the current technique of comparing experimental and theoretical C$^\alpha$ shift tensors to extract Ramachandran angles in solids may yield conformations that are reliable and consistent with X-ray diffraction, when impairments due to geometry optimization and intermolecular effects are small compared to the surfaces shift gradients. More refined test structures that include intermolecular interactions are needed to fully exploit the sensitivity of the chemical shift tensor for structure determination and the accuracy with which the ^{13}C and ^{15}N shift tensor components can be calculated with the EIM. Future methods may include crystal structure prediction in conjunction with X-ray powder diffraction or other approaches that are able to provide intermolecular arrangements so that intermolecular interactions are faithfully incorporated in shift tensor calculations of test structures.

(86) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzybschenko, A.; Erk, P.; Gavezotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooji, W. J. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Wiliams, D. E. *Acta Crystallogr.* **2002**, *B58*, 647−661.

(87) Smith, E. D. L.; Hammond, R. B.; Jones, M. J.; Roberts, K. J.; Mitchell, J. B. O.; Price, S. L.; Harris, R. K.; Apperley, D. C.; Cherryman, J. C.; Docherty, R. *J. Phys. Chem. B* **2001**, *105*, 5818−5826.

(88) Rajeswaran, M.; Blanton, T. N.; Zumbulyadis, N.; Giesen, D. J.; Conesa-Moratilla, C.; Misture, S. T.; Stephens, P. W.; Huq, A. *J. Am. Chem. Soc.* **2002**, *124*, 14 450−14 459.