Ab Initio Study of the Conformational Dependence of the Nonplanarity of the Peptide Group

Michael Ramek

Institut für Physikalische und Theoretische Chemie, Technische Universität Graz, A-8010 Graz, Austria

Ching-Hsing Yu, Joshua Sakon, and Lothar Schäfer*

Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701

Received: July 12, 2000; In Final Form: August 28, 2000

To study the nonplanarity of peptide bonds, the conformationally dependent variations of the N−C torsional angle, ω_2 (Figure 1), of the central peptide group in *N*-formyl L-alanyl L-alanine amide (ALA-ALA) was investigated using a database of 11 664 RHF/4-21G ab initio gradient optimized structures. The database was generated by optimizing the geometries of ALA-ALA at grid points in its four-dimensional (ϕ_1,ψ_1,ϕ_2,ψ_2) conformational space (Figure 1) defined by 40° increments along the outer torsions ϕ_1 and ψ_2, and by 30° increments along the inner torsions ψ_1 and ϕ_2. Using cubic spline functions, the grid structures were then used to construct analytical representations of complete surfaces of ω_2 in (ϕ_1,ψ_1,ϕ_2,ψ_2)-space. Analyses of the conformational surfaces of ω_2 reveal that the peptide N−C torsion is a smoothly varying function of associated ϕ and ψ angles and that, for many conformational regions, deviations from planarity are the rule rather than the exception. Comparisons with protein crystallographic data show that, in contrast to peptide torsional angles calculated for an entire protein, the ω_2 angles of smaller model peptides, such as ALA-ALA, cannot be used to model peptide groups in proteins, because of long-range effects present in the latter but not the former. This finding indicates the general difficulty of predicting the exact positions of backbone torsional angles in proteins from smaller model peptides. Furthermore, the results confirm the directional nature of polypeptide chains. That is, conformation transmission effects from neighboring groups differ, depending on whether they are transmitted from right to left or from left to right in the peptide chain.

Introduction

In the recent past there has been renewed interest in the nonplanarity of peptide bonds in peptides and proteins. On the basis of a survey of crystallographic data, MacArthur and Thornton[1] found that "substantial deviations from planarity can be tolerated with a standard deviation in the angle of up to 6° about a mean value for the trans peptide that is less than 180°". Furthermore, these authors[1] found evidence for a systematic dependence of the N−C peptide torsional angle, ω, on the ϕ-(N−C(α)) and ψ(C(α)-C′) backbone torsional angles in proteins as well as in small peptides. Rick and Cachau[2] derived similar conclusions from an analysis of protein structures in the protein data bank, demonstrating that the torsional rotation of the peptide bond is environmentally dependent in that different secondary structure elements in proteins are characterized by different degrees of the nonplanarity of ω. All of these findings are in contrast with the historic concept of the essential planarity of the trans peptide group, originally postulated by Corey and Pauling[3] and generally adopted in peptide conformational analyses until Ramachandran[4] recognized the need for nonplanar peptide units in polypeptide chains.

The variation of ω with ϕ and ψ is a special example, illustrating the concept of *local geometry*.[5] The geometries of peptides are local in the sense that the backbone structural parameters depend acutely on where a given molecule is in its ϕ,ψ-space. The concept is in direct contrast to attempts of defining peptides and proteins in terms of ideal and rigid geometries, such as those proposed by Engh and Huber[6] for use in protein X-ray crystallography. Because of the incompleteness of experimental data, ab initio calculations were often used[7−9] as an auxiliary source of information on local geometry trends in peptides and proteins, with particular interest directed toward important backbone parameters, such as the N−C(α) and C(α)-C′ bond lengths and the N−C(α)-C′ bond angle.[7−9]

The large number of ab initio calculations of peptides published during the past decade[10−113] demonstrates the general utility of quantum chemical procedures in studies of such molecules. In the current paper we will use the results of ab initio geometry optimizations of *N*-formyl L-alanyl L-alanine amide (ALA-ALA) for a systematic analysis of the dependence of the ω(N−C) torsional angle on the ϕ and ψ angles of adjacent residues. Apart from the context of this paper, the quantitative structural information on ALA-ALA given below is useful in parameter refinements for empirical molecular modeling procedures. For example, we are currently employing the database described below in attempts to further develop molecular dynamics simulation procedures of the adsorption of organic materials on the clay mineral/aqueous solution interface.[114−116]

Computational Procedures

For the present study we have generated[117] a structural database consisting of 11 664 ab initio gradient optimized structures of ALA-ALA. This database was obtained by optimizing the geometries of this compound at grid points in

* Corresponding author. E-mail: schafer@protein.uark.edu.

10.1021/jp002498x CCC: $19.00
Published on Web 10/03/2000

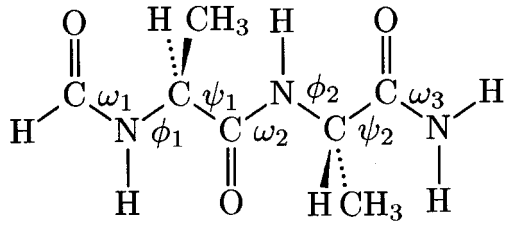

Figure 1. Schematic representation of *N*-formyl L-alanine L-alanyl amide.

its four-dimensional ($\phi_1,\psi_1,\phi_2,\psi_2$) conformational space (Figure 1) defined by 40° increments along the outer torsions ϕ_1 and ψ_2, and by 30° increments along the inner torsions ψ_1 and ϕ_2.[117] The smaller step size for the latter was chosen because the properties of the central peptide bond are of particular interest. At each of the $9 \times 12 \times 12 \times 9 = 11\,664$ grid points geometry optimizations were performed via RHF/4-21G[118] calculations, in which the torsions ϕ_1, ψ_1, ϕ_2, and ψ_2 were kept constant while all other structural parameters were relaxed without any constraints. The orientation of the torsions ω_1 and ω_2 (Figure 1) was trans in all cases. The specific torsion of interest for this study is ω_2 because it is the link of two complete residues.

To evaluate the results of the ab initio calculations, an auxiliary program was written using natural cubic spline functions to generate an analytical representation of the potential energy surface (PES) of ALA-ALA from the 11 664 grid points. By the help of this program it was possible to locate the local minima, which were subsequently optimized in RHF/4-21G[118] and MP2/6-311G**[119,120] calculations. Additional computational details are given in ref 117.

Results and Discussion

(a) Notation. For the purposes of this paper we adopt the following notation. The spline function generated analytical representation of the PES of ALA-ALA can be used for calculating complete surfaces of the torsional dependence of ω_2 on ϕ_1, ψ_1, ϕ_2, and ψ_2. To present the properties of four-dimensional surfaces in a two-dimensional medium, the conformational dependence of ω_2 on the two amino acid residues of ALA-ALA can be explored when one of them is kept in a fixed orientation while the other one is allowed to move freely. Throughout this paper we will refer to the former as the *constrained residue*, and to the latter, as the *moving residue*.

The terms *residue 1* (with ω_1, ϕ_1 and ψ_1) and *residue 2* (with ω_2, ϕ_2, and ψ_2) are used in agreement with Figure 1. The shorthand notation "Phi$_i$_psi$_i$_xx", where $i = 1$ or 2, is used to label graphs of surfaces of ω_2 that are obtained when ϕ_i and ψ_i vary, i.e., residue *i* is the moving residue, while residue *j* (with $i \neq j$) is the constrained residue, with torsional angles, ϕ_j and ψ_j, fixed in the conformational regions xx = ar, al, br, or bt, respectively. Because of software limitations, the latinicized symbols 'al', 'ar', 'bt', and 'br' are used in the graphs for the right- and left-handed α-helical regions, α_R (=ar), α_L (=al), for the extended forms, β_S (=bt), and the bridge region, δ_R (=br), respectively. In agreement with a convention generally accepted in protein crystallography,[121] we have selected $\phi = \psi = 55°$ for al, $\phi = -75°$ and $\psi = -45°$ for ar, $\phi = -165°$, and $\psi = 165°$ for bt, and $\phi = -90°$ and $\psi = 0°$ for br.

By the term *calculated values* we refer to all parameter values (such as N−C(α), C(α)−C′, N−C(α)−C′, or ω_2) which were calculated at specific ϕ_1, ψ_1, ϕ_2, and ψ_2 torsions by spline function interpolation of the RHF/4-21G database of ALA-ALA. When calculated values are compared with experimental parameter values, as in Table 1, the latter were always taken from the crystallographic database established by Karplus.[121]

TABLE 1: Root Mean Square Deviations between the Ab Initio Bond Lengths (Å) and Angles (deg) of ALA-ALA and Values Found in Crystallographic Data of Proteins[a]

	total			Bmc		
	1.75	1.56	1.40	1.75	1.56	1.40
C(α)−C′(1)	0.020	0.019	0.019	0.020	0.019	0.019
C(α)−C′(2)	0.021	0.019	0.019	0.021	0.019	0.019
N−C(α)(1)	0.021	0.020	0.019	0.021	0.020	0.019
N−C(α)(2)	0.021	0.021	0.020	0.021	0.021	0.020
N−C(α)−C′(1)	3.09	2.93	2.89	3.01	2.90	2.85
N−C(α)−C′(2)	3.21	3.05	2.96	3.14	3.03	2.93

[a] To obtain the rms deviations of this table, sets of backbone torsional angles ($\phi_i,\psi_i,\phi_{i+1},\psi_{i+1}$) and their associated backbone bond lengths and angles were taken from the "total" and "Bmc" protein crystallographic data files reported by Karplus.[121] The bond lengths C(α)−C′, N−C(α), and the angle N−C(α)−C′ were then calculated by spline-function interpolation at each point ($\phi_i,\psi_i,\phi_{i+1},\psi_{i+1}$) in the four-dimensional torsional space of ALA-ALA, using the HF/4-21G ab initio values of the latter to obtain parameter values for both residues 1 and 2. Residue numbers are indicated in parentheses behind each parameter type. The rms deviations were then calculated between the ab initio values and the crystallographic values of the bond lengths and angles associated with each set of ($\phi_i,\psi_i,\phi_{i+1},\psi_{i+1}$). The crystallographic values were taken from three selections of proteins, characterized by resolutions of ≤ 1.75 Å, ≤ 1.56 Å, and ≤ 1.40 Å.

In some of the comparisons described below, region average values are involved. To determine region average values, regions in ϕ_1,ψ_1-space were defined by a 15° grid, yielding the region boundaries shown in Tables 2 to 4. For each region, experimental average parameter values were calculated from crystallographic parameters[121] in (ϕ_i, ψ_i, ϕ_{i+1}, ψ_{i+1}) protein subunits which were ordered by region in accordance with their torsional angles, ϕ_i and ψ_i. C*alculated region averages* are for ab initio parameter values, which were calculated at the crystallographic (ϕ_i, ψ_i, ϕ_{i+1}, ψ_{i+1}) torsional angles by spline function interpolation of the RHF/4-21G database of ALA-ALA. To determine averages, the calculated parameters were ordered by their ϕ_1 and ψ_1 torsional angles and averaged for each region.

(b) Comparison of Experimental and Calculated Structural Parameters. To evaluate the accuracy of the ab initio structures, several comparisons with experimental structures were executed. In Table 1, root-mean-square (rms) deviations are presented for some calculated and experimental structural parameters. It is seen that, for bond lengths such as N−C(α) and C(α)−C′, the rms deviations are on the order of 0.02 Å; and 3° for the bond angle, N−C(α)−C′.

Calculated and experimental region average values are compared in Tables 2−4 for regions with a population of >20 residues, as found in the protein database by Karplus.[121] For N−C(α), C(α)−C′, and N−C(α)−C′, the rms deviations between calculated and experimental region average values are 0.008 Å, 0.006 Å, and 1.10°, respectively.

In Figures 2−4, the calculated and experimental region averages are compared. As found before,[58] there is close correlation for N−C(α)−C′, with well-known systematic deviations in the α-helical and β-extended regions (previously[8,58] termed helix-contraction and β-expansion), whereas calculated and experimental trends in bond lengths, such as N−C(α), C(α)−C′, typically do not agree.

It is possible to expect that ω-angles obtained from RHF/4-21G geometry optimizations are not sufficiently accurate to be meaningfully compared with experimental values. To explore this matter, the 48 RHF/6-31G* optimized energy minima of ALA-ALA, described before,[117] were re-optimized using MP2/6-311G** gradient geometry optimization. Two of the 48 HF/

TABLE 2: Average Values of Crystallographic and Calculated N−C(α) Bond Lengths(Å) of Residue 1 of ALA-ALA in Various Regions of Peptide ϕ/ψ-space[a]

region	X-ray	calc	diff	pop
1 [−165,−150] [165, 180]	1.474	1.455	0.019	30
2 [−150,−135] [165, 180]	1.471	1.455	0.016	28
3 [−135,−120] [165, 180]	1.468	1.456	0.013	21
4 [−165,−150] [150, 165]	1.474	1.456	0.018	46
5 [−150,−135] [150, 165]	1.470	1.456	0.014	67
6 [−135,−120] [150, 165]	1.468	1.456	0.012	67
7 [−120,−105] [150, 165]	1.463	1.456	0.007	45
8 [−150,−135] [135, 150]	1.471	1.458	0.012	44
9 [−135,−120] [135, 150]	1.470	1.458	0.012	85
10 [−120,−105] [135, 150]	1.469	1.458	0.011	69
11 [−105, −90] [135, 150]	1.468	1.458	0.010	55
12 [−150,−135] [120, 135]	1.466	1.462	0.005	27
13 [−135,−120] [120, 135]	1.470	1.462	0.008	78
14 [−120,−105] [120, 135]	1.465	1.462	0.003	81
15 [−105, −90] [120, 135]	1.470	1.461	0.008	79
16 [−90, −75] [165, 180]	1.473	1.453	0.020	29
17 [−75, −60] [165, 180]	1.471	1.454	0.017	21
18 [−105, −90] [150, 165]	1.466	1.455	0.011	23
19 [−90, −75] [150, 165]	1.468	1.454	0.013	51
20 [−75, −60] [150, 165]	1.471	1.455	0.016	65
21 [−90, −75] [135, 150]	1.463	1.457	0.006	54
22 [−75, −60] [135, 150]	1.472	1.458	0.014	114
23 [−60, −45] [135, 150]	1.470	1.459	0.010	48
24 [−90, −75] [120, 135]	1.475	1.461	0.014	48
25 [−75, −60] [120, 135]	1.471	1.461	0.009	58
26 [−60, −45] [120, 135]	1.470	1.463	0.008	31
27 [−135,−120] [105, 120]	1.471	1.465	0.006	35
28 [−120,−105] [105, 120]	1.470	1.465	0.005	43
29 [−105, −90] [105, 120]	1.477	1.465	0.012	49
30 [−120,−105] [15, 30]	1.467	1.464	0.003	27
31 [−105, −90] [15, 30]	1.466	1.463	0.002	21
32 [−105, −90] [0, 15]	1.469	1.462	0.007	43
33 [−90, −75] [0, 15]	1.459	1.463	−0.004	28
34 [−120,−105] [−15, 0]	1.463	1.462	0.002	22
35 [−105, −90] [−15, 0]	1.468	1.462	0.006	52
36 [−90, −75] [−15, 0]	1.466	1.463	0.003	80
37 [−75, −60] [−15, 0]	1.471	1.468	0.003	58
38 [−90, −75] [−30, −15]	1.470	1.463	0.007	48
39 [−75, −60] [−30, −15]	1.472	1.469	0.003	223
40 [−60, −45] [−30, −15]	1.468	1.469	−0.002	41
41 [−90, −75] [−45, −30]	1.468	1.468	0.000	43
42 [−75, −60] [−45, −30]	1.473	1.469	0.003	603
43 [−60, −45] [−45, −30]	1.471	1.469	0.002	248
44 [−75, −60] [−60, −45]	1.472	1.470	0.003	200
45 [−60, −45] [−60, −45]	1.469	1.470	−0.001	149
46 [45, 60] [30, 45]	1.465	1.473	−0.008	22
47 [60, 75] [15, 30]	1.464	1.471	−0.007	23

[a] Region numbers and boundaries of regions in ϕ/ψ-space of residue 1 are given in column "region". The boundaries of each region are those of a 15° grid and are defined by the lower and upper values, ϕ_{lower} and ϕ_{upper}, respectively, of ϕ_1, given in the first bracket of each line of column "region", and the lower and upper values, ψ_{lower} and ψ_{upper}, respectively, of ψ_1, given in the second bracket. That is, each region defines a range, $\phi_{upper} > \phi_1 \geq \phi_{lower}$, and $\psi_{upper} > \psi_1 \geq \psi_{lower}$, respectively. Region numbering is arbitrary, starting with the β-region, proceeding to α_R and ending at α_L. Note that only regions with a population of >20, as found in the Bmc database of ref 121, were included in the analysis; all others were omitted. Region populations are given in column "pop". For each region, average experimental parameter values were calculated (col. "X-ray) from N−C(α) bond lengths, which were taken from crystal structures and ordered by region in accordance with their associated torsional angles ϕ_1 and ψ_1. The corresponding region averages for ab initio bond lengths were obtained for values calculated at the crystallographic torsional angles (ϕ_1, ψ_1, ϕ_2, ψ_2) and are given in the column "calc". Differences between experimental minus ab initio values are given in the column "diff". The rms deviation for all regions listed in this table (population weighted) is 0.008 Å. All crystallographic values were taken from the Bmc set of molecules given by Karplus,[121] with resolution ≤ 1.56 Å.

TABLE 3: Average Values of Crystallographic and Calculated C(α)−C Bond Lengths(Å) of Residue 1 of ALA-ALA in Various Regions of Peptide ϕ/ψ-space[a]

region	X-ray	calc	diff	pop
1 [−165,−150] [165, 180]	1.523	1.527	−0.004	30
2 [−150,−135] [165, 180]	1.526	1.526	0.000	28
3 [−135,−120] [165, 180]	1.527	1.529	−0.001	21
4 [−165,−150] [150, 165]	1.528	1.525	0.003	46
5 [−150,−135] [150, 165]	1.524	1.525	−0.002	67
6 [−135,−120] [150, 165]	1.528	1.526	0.002	67
7 [−120,−105] [150, 165]	1.529	1.529	0.000	45
8 [−150,−135] [135, 150]	1.528	1.525	0.003	44
9 [−135,−120] [135, 150]	1.525	1.526	−0.002	85
10 [−120,−105] [135, 150]	1.527	1.528	−0.001	69
11 [−105, −90] [135, 150]	1.532	1.529	0.002	55
12 [−150,−135] [120, 135]	1.524	1.525	−0.001	27
13 [−135,−120] [120, 135]	1.528	1.527	0.001	78
14 [−120,−105] [120, 135]	1.529	1.529	0.000	81
15 [−105, −90] [120, 135]	1.529	1.530	−0.001	79
16 [−90, −75] [165, 180]	1.530	1.534	−0.004	29
17 [−75, −60] [165, 180]	1.527	1.536	−0.009	21
18 [−105, −90] [150, 165]	1.539	1.530	0.009	23
19 [−90, −75] [150, 165]	1.534	1.532	0.002	51
20 [−75, −60] [150, 165]	1.529	1.533	−0.004	65
21 [−90, −75] [135, 150]	1.529	1.531	−0.002	54
22 [−75, −60] [135, 150]	1.528	1.534	−0.006	114
23 [−60, −45] [135, 150]	1.527	1.536	−0.009	48
24 [−90, −75] [120, 135]	1.527	1.532	−0.004	48
25 [−75, −60] [120, 135]	1.534	1.535	−0.002	58
26 [−60, −45] [120, 135]	1.528	1.538	−0.010	31
27 [−135,−120] [105, 120]	1.533	1.528	0.005	35
28 [−120,−105] [105, 120]	1.531	1.530	0.002	43
29 [−105, −90] [105, 120]	1.531	1.531	0.000	49
30 [−120,−105] [15, 30]	1.526	1.536	−0.010	27
31 [−105, −90] [15, 30]	1.522	1.537	−0.015	21
32 [−105, −90] [0, 15]	1.525	1.535	−0.010	43
33 [−90, −75] [0, 15]	1.526	1.537	−0.011	28
34 [−120,−105] [−15, 0]	1.522	1.531	−0.009	22
35 [−105, −90] [−15, 0]	1.522	1.532	−0.011	52
36 [−90, −75] [−15, 0]	1.524	1.534	−0.010	80
37 [−75, −60] [−15, 0]	1.521	1.534	−0.014	58
38 [−90, −75] [−30, −15]	1.522	1.532	−0.009	48
39 [−75, −60] [−30, −15]	1.523	1.532	−0.009	223
40 [−60, −45] [−30, −15]	1.523	1.533	−0.011	41
41 [−90, −75] [−45, −30]	1.525	1.529	−0.004	43
42 [−75, −60] [−45, −30]	1.526	1.530	−0.003	603
43 [−60, −45] [−45, −30]	1.525	1.531	−0.006	248
44 [−75, −60] [−60, −45]	1.526	1.530	−0.004	200
45 [−60, −45] [−60, −45]	1.529	1.531	−0.002	149
46 [45, 60] [30, 45]	1.523	1.539	−0.016	22
47 [60, 75] [15, 30]	1.517	1.537	−0.020	23

[a] The symbols of Table 2 were used, referring to C(α)−C. The population-weighted rms deviation is 0.006 Å for all regions listed in this table.

6-31G* minima, $\epsilon\epsilon$ and $\epsilon\delta_L$[117] were found not stable in MP2-space. During the MP2/6-311G** geometry optimization they converted to the single minimum listed in Table 5 at $\phi_1 = 73.8°$, $\psi_1 = -58.5°$, $\phi_2 = 74.5°$, and $\psi_2 = -67.8°$. For most other structures, the ϕ,ψ torsional angles of the RHF/6-31G* energy minima are within a few degrees of the corresponding angles in the MP2/6-311G** geometries. Nevertheless, in some cases relatively large ϕ,ψ shifts of 20° to 30° are found.

Table 5 presents a comparison of ω_2-angles taken from the MP2/6-311G** and RHF/4-21G structures of ALA-ALA. The graphic representation in Figure 5 documents the similarity of the two sets.

(c) Torsional Dependence of ω_2 Angles. When calculated and experimental backbone bond lengths are compared with each other, the overall rms deviations are relatively small, i.e., 0.008 and 0.006 Å for the N−C(α) and C(α)−C′ bond lengths, respectively, but individual region average bond lengths (Tables 2 and 3, and Figures 2 and 3) disagree on the order of 0.02 Å. Closer inspection of Figure 2 seems to suggest that the agreement for N−C(α) is good in helical regions (#30−40) but

TABLE 4: Average Values of Crystallographic and Calculated N−C(α)−C Bond Angles (deg) of Residue 1 of ALA-ALA in Various Regions of Peptide ϕ/ψ-space[a]

region	X-ray	calc	diff	pop
1 [−165,−150] [165, 180]	108.56	106.75	1.80	30
2 [−150,−135] [165, 180]	109.53	107.25	2.28	28
3 [−135,−120] [165, 180]	109.10	107.80	1.30	21
4 [−165,−150] [150, 165]	108.02	106.63	1.39	46
5 [−150,−135] [150, 165]	108.76	107.12	1.64	67
6 [−135,−120] [150, 165]	109.46	107.64	1.82	67
7 [−120,−105] [150, 165]	109.53	108.34	1.19	45
8 [−150,−135] [135, 150]	108.49	106.63	1.86	44
9 [−135,−120] [135, 150]	109.30	107.11	2.19	85
10 [−120,−105] [135, 150]	108.81	107.85	0.96	69
11 [−105, −90] [135, 150]	109.57	108.71	0.86	55
12 [−150,−135] [120, 135]	107.73	106.08	1.65	27
13 [−135,−120] [120, 135]	107.63	106.66	0.97	78
14 [−120,−105] [120, 135]	108.29	107.30	0.99	81
15 [−105, −90] [120, 135]	108.53	108.08	0.45	79
16 [−90, −75] [165, 180]	109.80	109.41	0.38	29
17 [−75, −60] [165, 180]	110.14	109.67	0.47	21
18 [−105, −90] [150, 165]	109.22	108.96	0.26	23
19 [−90, −75] [150, 165]	110.25	109.76	0.50	51
20 [−75, −60] [150, 165]	110.28	109.96	0.32	65
21 [−90, −75] [135, 150]	110.36	109.49	0.86	54
22 [−75, −60] [135, 150]	110.04	109.64	0.41	114
23 [−60, −45] [135, 150]	110.51	109.65	0.86	48
24 [−90, −75] [120, 135]	109.04	108.96	0.09	48
25 [−75, −60] [120, 135]	109.24	109.20	0.04	58
26 [−60, −45] [120, 135]	110.94	109.29	1.65	31
27 [−135,−120] [105, 120]	106.30	106.32	−0.03	35
28 [−120,−105] [105, 120]	106.41	106.90	−0.49	43
29 [−105, −90] [105, 120]	106.67	107.53	−0.86	49
30 [−120,−105] [15, 30]	112.14	112.76	−0.62	27
31 [−105, −90] [15, 30]	113.75	113.51	0.24	21
32 [−105, −90] [0, 15]	113.09	113.60	−0.52	43
33 [−90, −75] [0, 15]	113.66	114.59	−0.93	28
34 [−120,−105] [−15, 0]	114.89	113.26	1.62	22
35 [−105, −90] [−15, 0]	113.55	113.32	0.23	52
36 [−90, −75] [−15, 0]	113.32	114.05	−0.73	80
37 [−75, −60] [−15, 0]	113.34	114.80	−1.46	58
38 [−90, −75] [−30, −15]	113.17	113.26	−0.10	48
39 [−75, −60] [−30, −15]	112.39	113.87	−1.49	223
40 [−60, −45] [−30, −15]	113.65	114.70	−1.05	41
41 [−90, −75] [−45, −30]	111.27	112.18	−0.90	43
42 [−75, −60] [−45, −30]	111.04	112.24	−1.20	603
43 [−60, −45] [−45, −30]	112.15	112.87	−0.72	248
44 [−75, −60] [−60, −45]	110.73	111.68	−0.95	200
45 [−60, −45] [−60, −45]	111.53	112.02	−0.50	149
46 [45, 60] [30, 45]	112.02	111.76	0.26	22
47 [60, 75] [15, 30]	115.47	113.27	2.20	23

[a] The symbols of Table 2 were used, referring to N−C(α)−C. The population-weighted rms deviation is 1.10° for all regions listed in this table.

poor elsewhere, whereas just the opposite is true (Figure 3) for C(α)−C′. If these trends are not random but systematic, we have currently no explanation to offer. We prefer to think that the trends described above are not systematic, because similar discrepancies in bond lengths were observed before[122] and were rationalized in terms of solid-state effects; that is, it is well-known among crystallographers[122] that effects of temperature, crystal structure, and packing and molecular volume effects are readily detected in bond lengths but not bond angles, which depend mainly on intramolecular properties. Indeed, it is seen from Table 4 and Figure 4 that, in the case of the N−C(α)−C′ bond angles, calculated and experimental trends in region average values are in good agreement *and* the two sets of values are close, as indicated by the rms deviation of 1.10°.

The case of the peptide torsion, ω_2, deserves special consideration. When the calculated and experimental region average values of ω_2 are compared, *both* the relative trends and the rms values indicate considerable disagreement. The considerable disagreement between individual angles is documented in Figure 6; the rms deviation for ω_2 in (ϕ_i, ψ_i, ϕ_{i+1}, ψ_{i+1}) subunits is 4.8° when the residues are ordered in regions in accordance with the values of ϕ_i and ψ_i; and 6.4° when they are ordered in accordance with the values of ϕ_{i+1} and ψ_{i+1}. Note that the overall range of parameter variations both in N−C(α)−

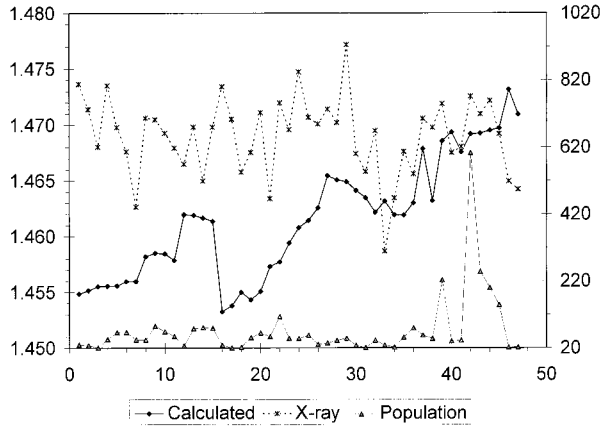

Figure 2. Region average values of N−C(α) bond lengths in proteins. The experimental and calculated values (left-hand ordinate), region populations (right-hand ordinate), and region numbers (abscissa) of Table 2 are shown.

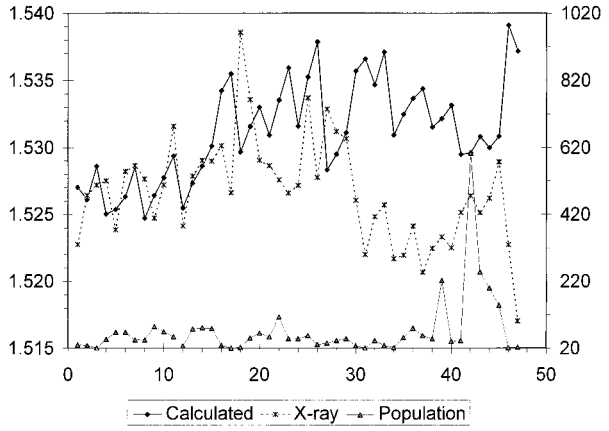

Figure 3. Region average values of C(α)−C′ bond lengths in proteins. The experimental and calculated values (left-hand ordinate), region populations (right-hand ordinate), and region numbers (abscissa) of Table 3 are shown.

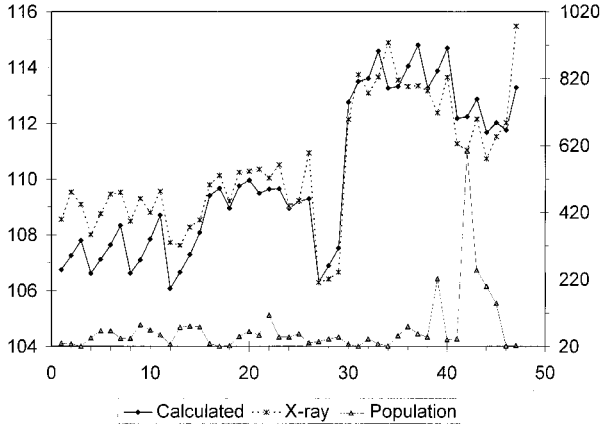

Figure 4. Region average values of N−C(α)−C′ bond angles in proteins. The experimental and calculated values (left-hand ordinate), region populations (right-hand ordinate), and region numbers (abscissa) of Table 4 are shown.

TABLE 5: Relative Energies (kcal/mol) and Torsional Angles (ϕ_1, ψ_1, ϕ_2, ψ_2, and $\omega 2$ in deg) for 47 MP2/6-311G Optimized Structures of *N*-Formyl L-alanyl L-alanine Amide[a]**

energy	ϕ_1	ψ_1	ϕ_2	ψ_2	ω2/MP2	ω2/HF
0.000	−82.19	76.37	−86.12	78.63	189.19	185.31
1.730	−160.18	153.92	−81.51	82.07	175.46	179.14
1.995	−70.31	−22.50	−105.28	12.68	179.58	174.62
2.098	73.60	−61.83	−80.86	81.33	182.22	184.15
2.217	−81.84	81.93	−146.18	17.73	195.94	189.76
2.399	−58.72	130.69	58.24	32.85	180.05	185.23
2.509	−131.73	14.70	−81.66	80.74	177.50	179.52
2.559	−160.11	162.65	−158.35	169.17	173.43	174.92
2.618	−82.65	77.42	71.42	−46.79	181.49	174.74
2.841	−78.84	80.28	49.63	−148.68	202.20	198.55
2.957	−81.09	87.18	−162.15	155.05	188.19	180.66
3.108	52.47	−133.74	−106.89	17.76	176.57	175.09
3.109	−52.50	133.72	107.00	−17.92	183.44	185.85
3.224	63.12	30.38	−88.79	79.33	187.13	184.04
3.605	61.11	31.49	58.02	33.70	175.50	180.42
3.686	−87.49	65.33	58.44	36.73	176.94	181.66
3.714	71.06	−72.22	−54.28	142.06	161.15	167.59
3.933	−122.99	11.52	−159.05	167.53	174.40	174.89
4.034	−77.56	82.22	−181.24	−36.79	190.88	186.63
4.076	50.92	−140.36	−82.09	81.37	176.57	182.11
4.317	−160.37	155.89	75.01	−69.05	171.77	173.10
4.493	60.27	34.95	−170.15	153.94	186.80	182.57
4.494	−158.53	170.08	−144.77	23.64	187.24	185.63
4.628[b]	73.82	−58.51	74.47	−67.77	177.01	177.71
4.705	73.69	−58.07	74.77	−46.07	176.56	174.71
4.818	−163.48	−61.94	−58.79	155.81	153.88	152.02
4.850	−158.53	54.83	51.21	−157.31	207.92	206.90
4.901	−116.77	12.85	74.13	−69.69	173.81	175.71
5.038	75.98	−53.19	−74.97	−23.25	187.72	181.76
5.228	−160.61	155.32	62.42	38.34	164.55	173.73
5.389	72.88	−68.21	61.78	39.03	169.28	177.36
5.433	−172.24	−40.73	−84.28	82.50	186.82	182.74
5.685	58.99	42.32	70.90	−71.96	185.28	177.74
5.692	60.90	38.59	−139.27	15.98	193.83	189.58
5.821	75.37	−57.50	−159.05	160.13	182.48	182.52
6.097	52.82	−137.26	−153.67	171.76	173.68	175.83
6.220	−159.18	169.38	54.02	−141.58	187.51	184.06
6.404	−168.97	−38.93	−164.06	163.58	185.34	181.05
6.663	73.03	−50.69	51.91	−141.22	194.52	191.17
6.665	−123.76	13.01	62.55	37.85	164.51	171.79
6.976	−159.60	157.66	−164.38	−40.56	167.79	171.00
7.099	64.40	17.66	158.83	−36.95	186.47	182.84
8.078	−171.75	−37.44	71.47	−70.63	184.41	178.81
8.243	−135.13	15.79	−164.36	−37.18	167.27	170.78
9.184	−168.39	−34.33	52.80	−143.59	196.43	188.09
9.610	−173.82	−46.12	59.71	37.86	175.27	179.25
11.049	−172.81	−43.99	−163.14	−48.85	178.82	175.12

[a] Energies (kcal/mol) are listed in column "energy" and are relative to −663.17553265889 Hartree. Values listed in column "ω2/MP2" are the ω2 angles (degree) obtained from the MP2/6-311G** geometry optimizations. In column "ω2/HF", the HF/4-21G ω2 angles are listed which were calculated at the torsional angles given in columns "ϕ_1", "ψ_1", "ϕ_2", and, "ψ_2.

C′ and ω_2 are ∼10°, but the rms deviation is 1.1° for the former and nearly 5° for the latter.

From this result, one might derive the impression that RHF/4-21G calculations cannot provide any useful information on the conformational properties of peptide bonds. However, this result really represents an interesting and somewhat unexpected finding. That is, the actual values of peptide torsions in proteins cannot be simulated by those obtained from smaller model systems, such as ALA-ALA, because of long-range interaction, which are present in the former but not the latter. This thesis is supported by the fact that RHF/4-21G geometry optimizations performed on the entire protein crambin yielded peptide torsions[123] in significantly better agreement than Figure 6 with

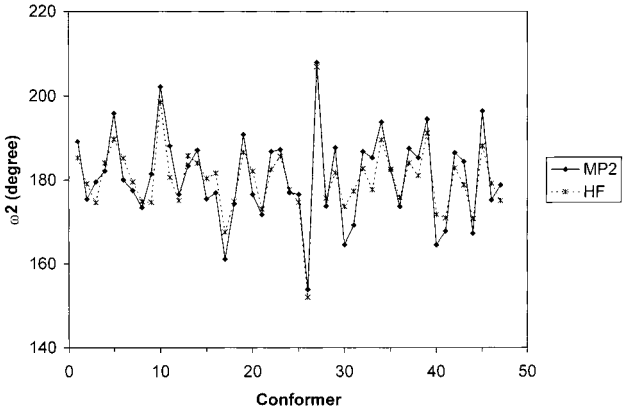

Figure 5. Comparison of RHF/4-21G and MP2/6-311G** calculated ω_2 angles (Table 5) in *N*-formyl L-alanine L-alanyl amide. The abscissa is for conformation numbers following the sequence of Table 5.

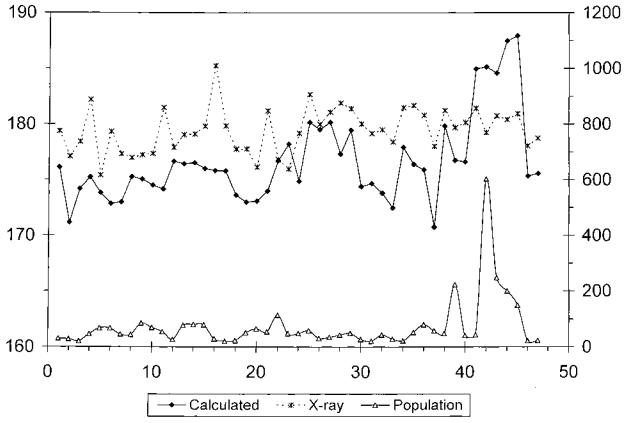

Figure 6. Comparison of calculated and experimental region average values of peptide torsional angles, ω. The experimental and calculated values of ω_2 (left-hand ordinate), region populations (right-hand ordinate), and region numbers (abscissa) have the same meaning as in Figure 2. Parameter values were determined as described in the text.

the experimental values.[124] Excluding the torsion of the terminal residue, which is usually subject to external effects, the rms deviation for ω in the case of crambin is 4.6°. This value is of the same magnitude as the rms deviation for ω_2 presented above, but the overall parameter range in the case of crambin is more than 20°, or twice as large as in ALA-ALA, and individual parameters are compared for which the rms deviations are always larger than for comparisons of average values. Furthermore, as seen from Figure 7, the relative trends are in significantly better agreement than those of Figure 6. Thus, the discrepancy between calculated region average values for ω_2, taken from ALA-ALA, and corresponding averages from protein crystal structures indicates the incompatibility of the two datasets but does not rule out the ab initio calculations of ω_2 in ALA-ALA as a source of reasonable estimates of the conformational effects of *neighboring* residues on the nonplanarity of the peptide link.

For this purpose, then, to make use of the RHF/4-21G ab initio database for ALA-ALA, a large number of different functional dependences of ω_2 can be selected for presentation, but only a limited documentation can be considered here.

In Figure 8, the surface phi1_psi1_ar is presented. The changes in ω_2 seen in Figure 8 are those that are obtained when residue 1 is the moving residue and residue 2 is constrained to the α-helical position. Comparing with Figure 9, it is seen that the functional dependence of ω_2 on ϕ_1 and ψ_1 is very different

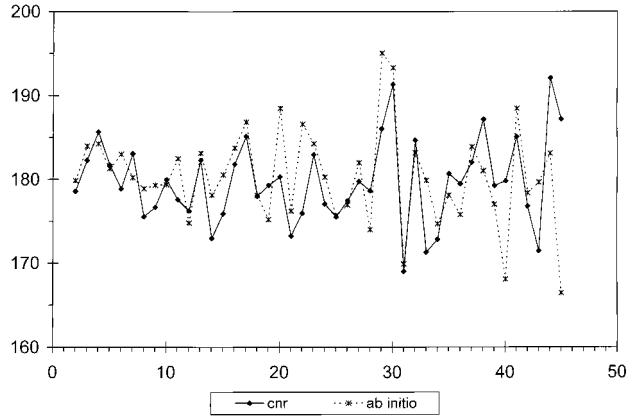

Figure 7. Comparison of calculated and experimental values of peptide torsional angles, ω, in the protein crambin. The calculated values of ω were taken from RHF/4-21G gradient geometry optimizations[123] executed on the entire molecule. The experimental values are from the crystal structure.[124] The abscissa is for residue numbers in agreement with the crystallographic study.[124]

when residue 2 is locked into a different region; such as the β region. Such results are consistently found: the local values of ω_2 are acutely dependent on the conformational states of the two residues that it connects. Similar trends are established by the experimental data. For example, it is seen from Figure 10 that different distributions of ω-angles are found in crystal structures for (ϕ_i, ψ_i, ϕ_{i+1}, ψ_{i+1}) pairs of residues, when residue i is α_R, while residue j is either α_R or δ_R; or when residue i is β_S, while residue j is either α_R or β_S. In all cases, different distributions result when a conformational change occurs in one of the two residues but not in the other.

In Figure 11, the surface phi2_psi2_ar is presented. Compared with that of phi1_psi1_ar (Figure 8), Figure 11 shows that the conformational effects, by residue 1 on ω_2, differ from those by residue 2. That is, conformation transmission effects in peptide chains are different when transmitted from right to left, than when transmitted from left to right.

In Figure 12, the surface for psi1_phi2_bt is shown. To produce that surface, ψ_1 and ϕ_2 were varied throughout their respective spaces, while the remaining two torsions were constrained at $\phi_1 = -165°$ and $\psi_2 = +165°$. Again, ω_2 is obtained as a smoothly varying function of associated torsional angles. The large deviations of ω_2 from 180° in this case (Figure 12) are particularly noteworthy.

(d) Library of Ab Initio Molecular Structural Data. The geometries calculated for this study have been added to a growing library[125] of molecular structures, which we obtained in the recent past by ab initio geometry optimizations of basic organic functional groups. As in the current study, most of the structures deposited in these database were obtained by RHF/4-21G geometry optimizations,[118] which previous experience has shown to be rather accurate for molecules of the kind considered here.

In addition to the more than 11 600 structures of N-formyl L-alanyl L-alanine amide described above, the library currently contains some 468 structures of *n*-hexane, spanning the entire three-dimensional conformational space of that molecule. Furthermore, it contains the RHF/4-21G optimized structures and conformational geometry functions of some fifty X−C−C−Y systems determined at 30° intervals of their respective one-dimensional torsional space; the complete set of several thousand structures determined for the two-dimensional X−C−C−C and C−C−C−Y torsional space of twenty eight X−C−C−C−Y systems, where X,Y= CH3, F, C, OH, NH2, COH, and COOH, for which optimizations were performed at 30° grid points; and several hundred RHF/4-21G structures determined at 30° grid points in the two-dimensional (ϕ,ψ)-space of the model dipeptides *N*-acetyl *N*′-methyl glycine amide and the alanine homologue.

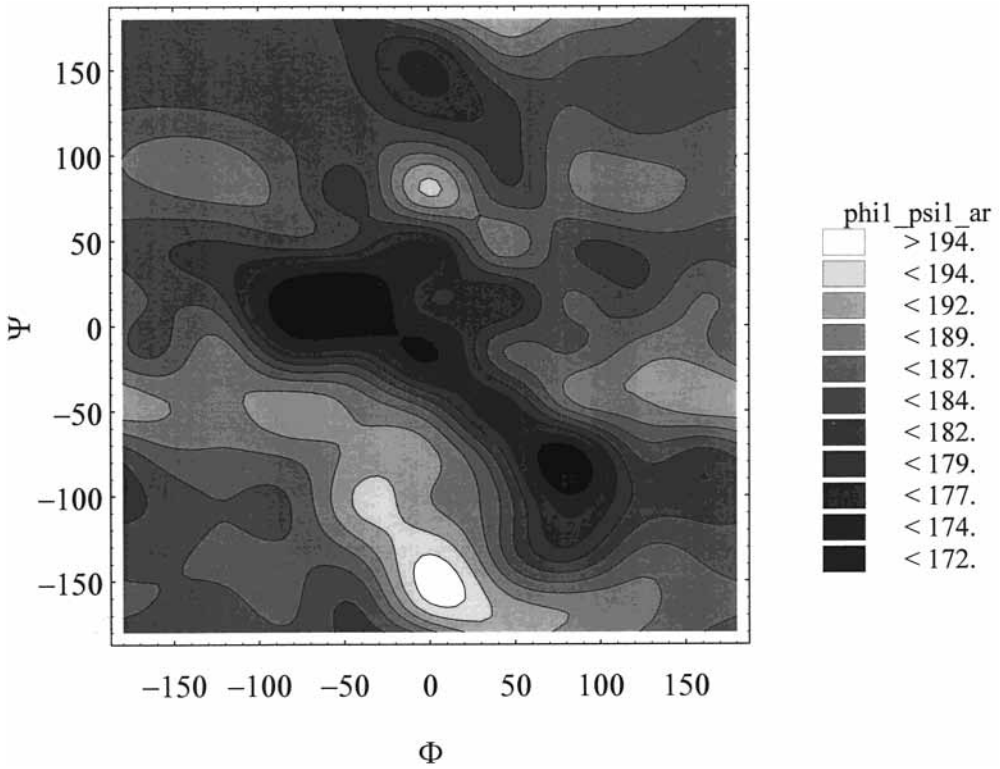

Figure 8. The functional dependence of ω_2 in *N*-formyl L-alanyl L-alanine amide on ϕ_1 and ψ_1. The parameter surface shown is obtained when residue 1 is the moving residue and residue 2 is constrained in the α-helical region.

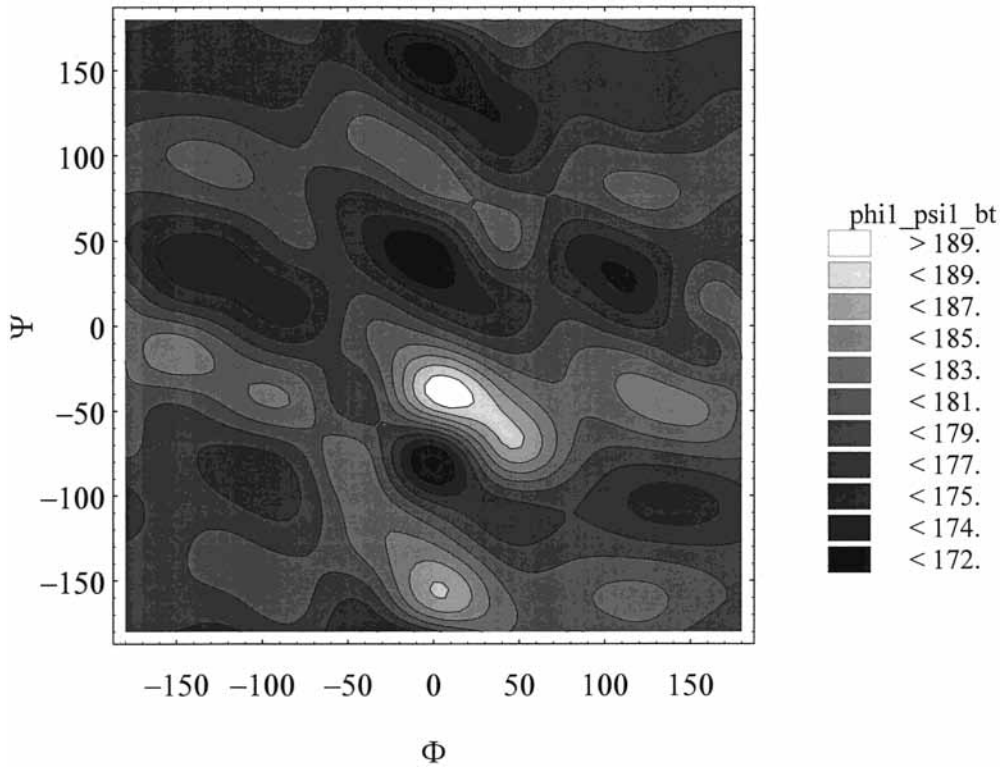

Figure 9. The functional dependence of ω_2 in *N*-formyl L-alanyl L-alanine amide on ϕ_1 and ψ_1. The parameter surface shown is obtained when residue 1 is the moving residue and residue 2 is constrained in the β region.

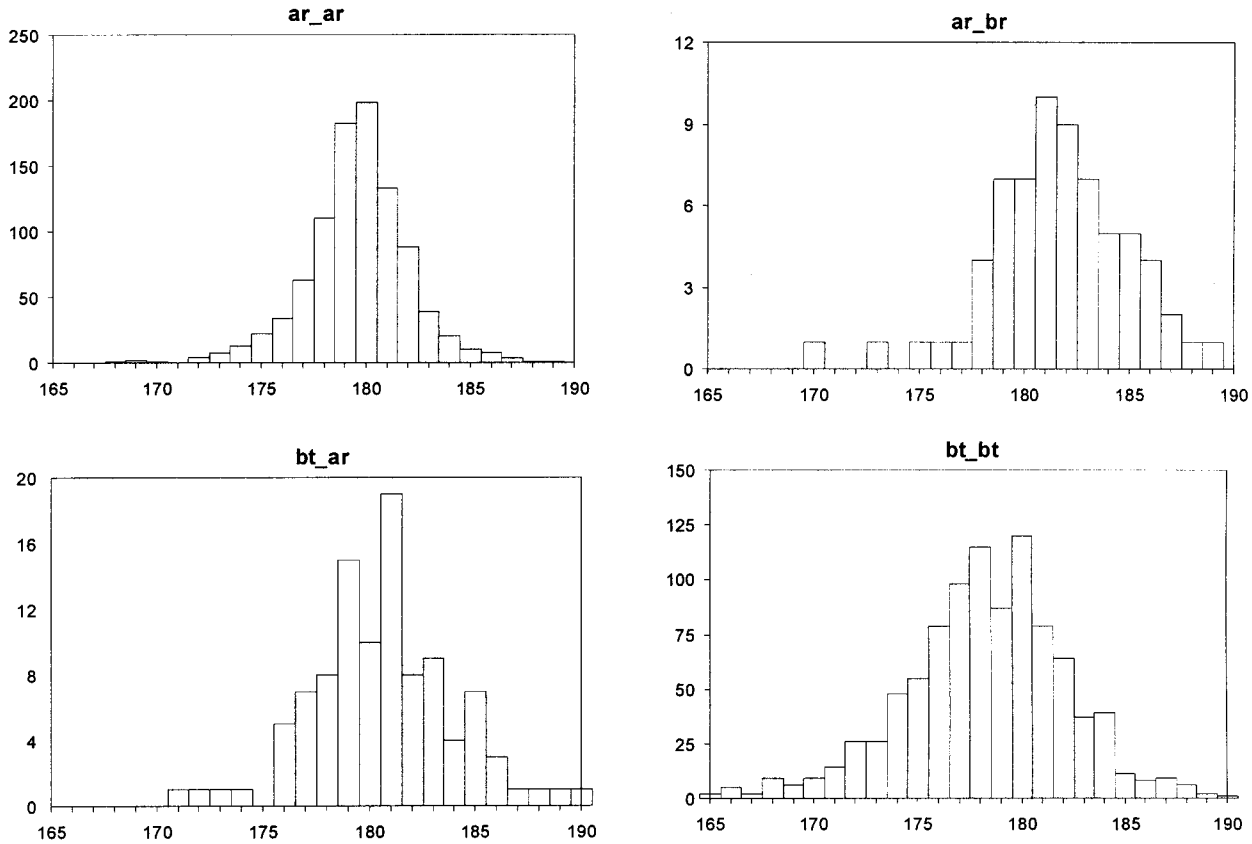

Figure 10. Distributions of ω-angles found in protein crystal structures[121] for (ϕ_i, ψ_i, ϕ_{i+1}, ψ_{i+1}) pairs of residues, when residue i is α_R while residue j is either α_R or δ_R (top; ar_ar, ar_br, respectively); or when residue i is β_S, while residue j is either α_R or β_S. (bottom; bt_ar, bt_bt, respectively).

The molecular library further contains the requisite programs for spline function interpolation, standard geometry function additivity, and gradient surface construction. By selecting a set of applicable torsional angles as input, the software included with the database will readily calculate the values of various backbone bond distances, bond angles of a given system, and

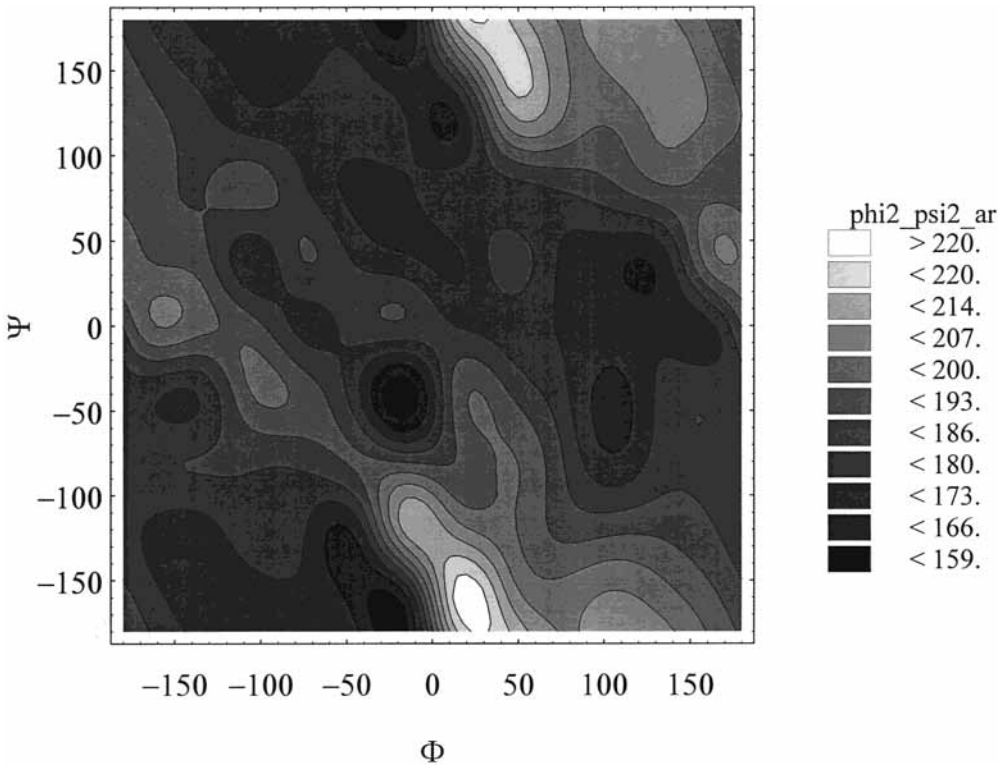

Figure 11. The functional dependence of ω_2 in *N*-formyl L-alanyl L-alanine amide on ϕ_2 and ψ_2. The parameter surface shown is obtained when residue 2 is the moving residue and residue 1 is constrained in the α-helical region.

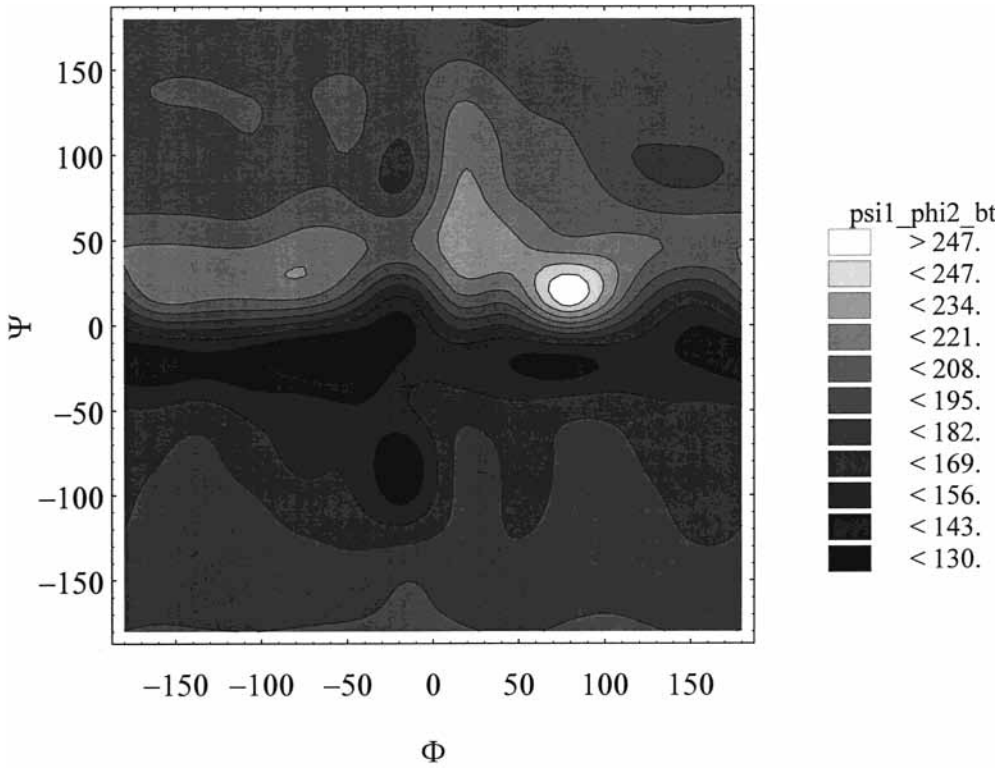

Figure 12. The functional dependence of ω_2 in *N*-formyl L-alanyl L-alanine amide on ϕ_2 and ψ_1. The parameter surface is obtained when ψ_1 and ϕ_2 are varied throughout their respective spaces, while the remaining two torsions are constrained at $\phi_1 = -165°$ and $\psi_2 = +165°$.

their gradients, at any given point in the conformational space. A copy of the library on CD is available from the authors on request.[125]

Apart from heuristic purposes illustrating the local nature[5] of molecular geometries, the database is useful in parameter developments for empirical molecular modeling procedures, such as in the clay/organic matter parameter development program[114−116] currently pursued in our laboratory.

Conclusions

The extent of disagreement (Figure 6) between calculated and experimental region average values of ω_2 is an interesting and

perhaps unexpected finding. In our opinion, this result shows to what extent torsional angles in proteins can be affected by long-range interactions, which are not present in small model peptides, like ALA-ALA, but in proteins. Apart from ω_2, such effects must be expected for all backbone torsional angles in proteins, indicating the essential difficulty of predicting precisely the positions of torsional minima of proteins, from the minima of smaller peptides.

The discrepancies in calculated and experimental values of ω_2 are in contrast to the generally good agreement found for primary structural parameters (Tables 1−4, and Figures 2−4), and for RHF/4-21G ω angles calculated for crambin as a whole and their crystal counterparts (Figure 7). In view of Table 5, we conclude that inclusion of electron correlation and polarization functions in the MP2/6-311G** calculations does not produce peptide torsions that are very different from calculations, such as RHF/4-21G, which are devoid of correlation and polarization functions. This finding is perhaps somewhat amazing for a parameter of the kind considered here, and it may indicate that, in short range, peptide nonplanarity is largely a matter of steric effects.

On the basis of the considerations above, we conclude that ab initio calculations of ALA-ALA can provide a good estimate of the flexibility of ω_2 as a function of interactions between its immediately neighboring residues. In this context, no matter what kind of functional dependence is considered, the peptide torsional angle is generally found to be a smoothly varying function of the associated ϕ and ψ angles. For many regions of ϕ,ψ-space, significant deviations from the planarity of the peptide group are the rule rather than the exception. Apart from the general variability, there are also constant patterns. For example, it seems that deviations from peptide planarity are always large when one of the associated residues is in the bridge region ($\phi = -90°$ and $\psi = 0°$).

Furthermore, from the results presented above, the directional properties of protein chains can be inferred. That is, conformation transmission effects directed from residue 1 to residue 2 are different than effects directed in the opposite direction.

Acknowledgment. The authors gratefully acknowledge support by USDA CSREES grant 99-35107-7782 and by the IBM Shared University Research Program. Special thanks are due to Prof. Collis Geren, Vice Chancellor for Research, University of Arkansas, and Dr. Jamie Coffin, IBM.

References and Notes

(1) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1996**, *264*, 1180.
(2) Rick, S. W.; Cachau, R. E. *J. Chem. Phys.* **2000**, *112*, 5230.
(3) Corey, R. B.; Pauling, L. *Proc. R. Soc. Ser. B* **1953**, *141*, 10.
(4) Ramachandran, G. N. *Biopolymers* **1968**, *7*, 1494.
(5) Schäfer, L.; van Alsenoy, C.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439.
(6) Engh, R.; Huber, R. *Acta Crystallogr. Sect. A* **1991**, *47*, 392.
(7) Scarsdale, J. N.; van Alsenoy, C.; Klimkowski, V. J.; Schäfer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438.
(8) Jiang, X.; Cao, M.; Teppen, B.; Newton, S. Q.; Schäfer, L. *J. Phys. Chem.* **1995**, *99*, 10521.
(9) Yu, C.-H.; Schäfer, L.; Ramek, M. *J. Phys. Chem. A* **1999**, *103*, 8337.
(10) Aida, M. *Bull. Chem. Soc. Jpn.* **1993**, *66*, 6, 3423.
(11) Adamo, C.; Dillet, V.; Barone, V. *Chem. Phys. Lett.* **1996**, *263*, 113.
(12) Aleman, C.; Puiggali, J. *J. Phys. Chem. B* **1997**, *101*, 3441.
(13) Amodeo, P.; Barone, V. *J. Am. Chem. Soc.* **1992**, *114*, 9085.
(14) Antohi, O.; Naider, F.; Sapse, A.-M. *J. Mol. Struct.* **1996**, *360*, 99.
(15) Bakhshi, A. K.; Otto, P.; Liegener, C.-M.; Rehm, E.; Ladik, J. *Int. J. Quantum Chem.* **1990**, *38*, 573.
(16) Balazs, A. *J. Mol. Struct.* **1991**, *245*, 111.
(17) Barone, V.; Fraternali, F.; Cristinziano, P. L. *Macromolecules* **1990**, *23*, 2038.
(18) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908.
(19) Böhm, H.-J. *J. Am. Chem. Soc.* **1993**, *115*, 6152.
(20) Böhm, H.-J.; Brode, S. *J. Comput. Chem.* **1995**, *16*, 146.
(21) Böhm, H.-J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129.
(22) Boon, G.; De Proft, F.; Langenaeker, W.; Geerlings, P. *Chem. Phys. Lett.* **1998**, *295*, 122.
(23) Bour, P.; Keiderling, T. A. *J. Am. Chem. Soc.* **1993**, *115*, 9602.
(24) Broda, M. A.; Rzeszotarska, B.; Smelka, L.; Rospenk, M. *J. Peptide Research* **1997**, *50*, 342.
(25) Cerda, B.; Hoyau, S.; Ohanessian, G.; Wesdemiotis, C. *J. Am. Chem. Soc.* **1998**, *120*, 2437.
(26) Chakraborty, D.; Manogaran, S. *J. Phys. Chem. A* **1997**, *101*, 6964.
(27) Chang, C.; Bader, R. F. W. *J. Phys. Chem.* **1992**, *96*, 1654.
(28) Cheam, T. C. *J. Mol. Struct.* **1993**, *295*, 259.
(29) Cheam, T. C. *J. Mol. Struct.* **1992**, *274*, 289.
(30) Cheam, T. C.; Krimm, S. *J. Mol. Struct. (THEOCHEM)* **1990**, *206*, 173.
(31) Chesnut, D. B.; Phung, C. G. *Chem. Phys. Lett.* **1991**, *183*, 505.
(32) Cornell, W. D.; Gould, I. R.; Kollman, P. A. *J. Mol. Struct.* **1997**, *392*, 101.
(33) Crisma, M.; Valle, G.; Formaggio, F.; Toniolo, C.; Bagno, A. *J. Am. Chem. Soc.* **1997**, *119*, 4136.
(34) Cui, C.; Kim, K. S. *J. Phys. Chem. A* **1999**, *103*, 2751.
(35) Dive, G.; Dehareng, D.; Ghuysen, J. M. *J. Am. Chem. Soc.* **1994**, *116*, 2548.
(36) Endredi, G.; Liegener, C.-M.; McAllister, M. A.; Perczel, A.; Ladik, J.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1994**, *306*, 1.
(37) Endredi, G.; Perczel, A.; McAllister, M. A.; Csonka, G. I.; Ladik, J.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1997**, *391*, 15.
(38) Faerman, C. H.; Price, S. L. *J. Am. Chem. Soc.* **1990**, *112*, 4915.
(39) Fang, D.-C.; Fu, X.-Y.; Tang, T.-H.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1998**, *427*, 243.
(40) Fernández, B.; Ríos, M. A.; Carballeira, L. *J. Comput. Chem.* **1991**, *12*, 78.
(41) Fischer, S.; Dunbrack, R. L., Jr.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931.
(42) Frau, J.; Donsos, J.; Munoz, F.; Blanco, F. G. *Biopolymers* **1998**, *45*, 119.
(43) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369.
(44) Garmer, D. R. *J. Phys. Chem. B* **1997**, *101*, 2945.
(45) Gervasio, F. L.; Guarna, A.; Giolitti, A.; Schettino, V. *J. Phys. Chem. B* **2000**, *104*, 1108.
(46) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250.
(47) Gould, I. R.; Hillier, I. H. *J. Chem. Soc., Chem. Commun.* **1993**, 951.
(48) Gould, I. R.; Kollman, P. A. *J. Phys. Chem.* **1992**, *96*, 9255.
(49) Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104.
(50) Grant, J. A.; Williams, R. L.; Scheraga, H. A. *Biopolymers*, **1990**, *30*, 929.
(51) Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104.
(52) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.
(53) 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.
(54) Hetzer, G.; Pulay, P.; Werner, H.-J. *Chem. Phys. Lett.* **1998**, *290*, 143.
(55) Jensen, J. H.; Baldridge, K. K.; Gordon, M. S. *J. Phys. Chem.* **1992**, *96*, 8340.
(56) Jewsbury, P.; Yamamoto, S.; Minato, T.; Saito, M.; Kitagawa, T. *J. Am. Chem. Soc.* **1994**, *116*, 11586.
(57) Jiao, D.; Barfield, M.; Hruby, V. J. *J. Am. Chem. Soc.* **1993**, *115*, 10883.
(58) Jiang, X.; Yu, C.-H.; Cao, M.; Newton, S. Q.; Paulus, E. F.; Schäfer L. *J. Mol. Struct.* **1997**, *403*, 83.
(59) Jonnson, M.; Wayner, D. D. M.; Armstrong, D. A.; Yu, D.; Rauk, A. *J. Chem. Soc., Perkin Trans.* **1998**, *2*, 1967.
(60) Kaschner, R.; Hohl, D. *J. Phys. Chem. A* **1998**, *102*, 5111.
(61) Kim, B. H.; Cho, S. G.; Ha, T. K. *J. Org. Chem.* **1999**, *64*, 5036.
(62) Kim, K. S.; Cui, C.; Cho, S. J. *J. Phys. Chem.* **1998**, *102*, 461.
(63) Klassen, J. S.; Kebarle, P. *J. Am. Chem. Soc.* **1997**, *119*, 6552.
(64) Ladik, J.; Sutjianto, A.; Otto, P. *J. Mol. Struct. (THEOCHEM)* **1991**, *228*, 271.
(65) Lewis, J. P.; Pawley, N. H.; Sankey, O. F. *J. Phys. Chem. B* **1997**, *101*, 10576.
(66) Liegener, C.-M.; Sutjianto, A.; Ladik, J. *Chem. Phys.* **1990**, *145*, 385.

(67) Luna, A.; Amekraz, B.; Tortajada, J.; Morizur, J. P.; Mo, O.; Yanez, M. *J. Am. Chem. Soc.* **1998**, *120*, 5411.
(68) Mapel, J. R.; Hwang, M.-J.; Jalkanen, K. J.; Stockfisch, T. P.; Hagler, A. T. *J. Comput. Chem.* **1998**, *19*, 430
(69) Maxwell, D. S.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 984.
(70) McAllister, M. A.; Perczel, A.; Császár, P.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1993**, *288*, 181.
(71) McAllister, M. A.; Perczel, A.; Császár, P.; Viviani, W.; Rivail, J.-L.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1993**, *288*, 161.
(72) Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, *359*, 653.
(73) Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1990**, *112*, 9016.
(74) Mohle, K.; Hofmann, H.-J. *Biopolymers* **1996**, *38*, 781.
(75) Mohle, K.; Hofmann, H.-J. *J. Peptide Res.* **1998**, *51*, 19.
(76) Mohle, K.; Gubmann, M.; Hofmann, H.-J. *J. Comput. Chem.* **1997**, *18*, 1415.
(77) No, K. T.; Grant, J. A.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1990**, *94*, 4740.
(78) Okimoto, N.; Tsukui, T.; Hata, M.; Hoshino, T.; Tsuda, M. *J. Am. Chem. Soc.* **1999**, *121*, 7349.
(79) Otto, P.; Sutjianto, A. *J. Mol. Struct. (THEOCHEM)* **1991**, *231*, 277.
(80) Perczel, A.; Ángyán, J. G.; Kajtár, M.; Viviani, W.; Rivail, J.-L.; Marcoccia, J.-F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256.
(81) Perczel, A.; Farkas, M. A. O.; Császár, A. G.; Csizmadia, I. G. *Can. J. Chem.* **1997**, *75*, 1120.
(82) Perczel, A.; Farkas, M. A. O.; Marcoccia, J. M.; Csizmadia, I. G. *Int. J. Quantum Chem.* **1997**, *61*, 797.
(83) Perczel, A.; Farkas, M. A. O.; Csizmadia, I. G. *J. Comput. Chem.* **1996**, *17*, 821.
(84) Perczel, A.; Kajtár, M.; Marcoccia, J.-F.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)*, **1991**, *232*, 291.
(85) Perczel, A.; McAllister, M. A.; Császár, P.; Csizmadia, I. G. *Can. J. Chem.* **1994**, *72*, 2050.
(86) Perczel, A.; McAllister, M. A.; Császár, P.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 4849.
(87) Perczel, A.; Viviani, W.; Csizmadia, I. G. In *Molecular Aspects of Biotechnology: Computational Models and Theories*; Bertran, J., Ed.; Kluwer Academic Publishers: Dordrecht, 1992; p 39.
(88) Price, S. L.; Faerman, C. H.; Murray, C. W. *J. Comput. Chem.* **1991**, *12*, 1187.
(89) Price, S. L.; Stone, A. J. *J. Chem. Soc., Faraday Trans.* **1992**, *88*, 1755.
(90) Ramek, M.; Kelterer, A.-M.; Teppen, B. J.; Schäfer, L. *J. Mol. Struct.* **1995**, *352/353*, 59.
(91) Rao, B. G.; Tilton, R. F.; Singh, U. C. *J. Am. Chem. Soc.* **1992**, *114*, 4447.
(92) Rauk, A.; Yu, D.; Armstrong, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 8848.
(93) Resat, H.; Maye, P. V.; Mezei, M. *Biopolymers* **1997**, *41*, 73.
(94) Roux, B. *Chem. Phys. Lett.* **1993**, *212*, 231.
(95) Sapse, A.-M.; Jain, D. C.; de Gale, D.; Wu, T. C. *J. Comput. Chem.* **1990**, *11*, 573.
(96) Schäfer, L.; Newton, S. Q.; Cao, M.; Peeters, A.; Van Alsenoy, C.; Wolinski, K.; Momany, F. A. *J. Am. Chem. Soc.* **1993**, *115*, 272.
(97) Scheiner, S.; Wang, L. *J. Am. Chem. Soc.* **1993**, *115*, 1958.
(98) Shang, H. S.; Head-Gordon, T. *J. Am. Chem. Soc.* **1994**, *116*, 1528.
(99) Shirazian, S.; Gronert, S. *J. Mol. Struct. (THEOCHEM)* **1997**, *397*, 107.
(100) Sokalski, W. A.; Keller, D. A.; Ornstein, R. L.; Rein, R. *J. Comput. Chem.* **1993**, *14*, 970.
(101) Sternberg, U.; Koch, F.-T.; Möllhoff, M. *J. Comput. Chem.* **1994**, *15*, 524.
(102) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer, H. F., III *J. Am. Chem. Soc.* **1995**, *117*, 2632.
(103) Torii, H.; Tasumi, M. *J. Mol. Struct.* **1993**, *300*, 171.
(104) Van Alsenoy, C.; Cao, M.; Newton, S. Q.; Teppen, B.; Perczel, A.; Csizmadia, I. G.; Momany, F. A.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1993**, *286*, 149.
(105) Vank, J. C.; Sosa, C. P.; Perczel, A.; Csizmadia, I. G. *Can. J. Chem.* **2000**, *78*, 395.
(106) Viviani, W.; Rivail, J.-L.; Csizmadia, I. G. *Theor. Chim. Acta* **1993**, *85*, 189.
(107) Viviani, W.; Rivail, J.-L.; Perczel, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 8321.
(108) Voisin, C.; Cartier, A. *J. Mol. Struct. (THEOCHEM)*, **1993**, *286*, 35.
(109) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1994**, *116*, 12022.
(110) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 12423.
(111) Williams, D. E. *Biopolymers* **1990**, *29*, 1367.
(112) Yang, W. *J. Mol. Struct. (THEOCHEM)* **1992**, *255*, 461.
(113) Zhang, K.; Zimmerman, D. M.; Chung-Phillips, A.; Cassady, C. J. *J. Am. Chem. Soc.* **1993**, *115*, 10812.
(114) Teppen, B. J.; Rasmussen, K.; Bertsch, P. M.; Miller, D. M.; Schäfer, L. J. *Phys. Chem.* **1997**, *101*, 1579.
(115) Teppen, B. J.; Yu, C.-H.; Miller, D. M.; Schäfer, L. *J. Comput. Chem.* **1998**, *19*, 144.
(116) Yu, C.-H.; Newton, Q. S.; Norman, M. A.; Miller, D. M.; Schäfer, L.; Teppen, B. J. *Clay and Clay Minerals*, in press.
(117) Ramek, M.; Yu, C.-H.; Schäfer, L. *Can. J. Chem.* **1998**, *76*, 566.
(118) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.
(119) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225.
(120) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(121) Karplus, P. A. *Protein Science* **1996**, *5*, 1406.
(122) Yu, C.-H.; Paulus, E. F.; Ramek, M.; Schäfer, L. *J. Mol. Struct.*, **2000**, *500*, 259.
(123) Van Alsenoy, C.; Yu, C.-H.; Peeters, A.; Martin, J. M. L.;. Schäfer, L. J. *Phys. Chem.* **1998**, *102*, 2246.
(124) Teeter, M. M.; Roe, S. M.; Heo, N. H. J. *Biol. Chem.* **1993**, *239*, 292.
(125) Ramek, M.; Yu, C.-H.; Schäfer, L. "*MOLL − A Molecular Library of Ab Initio Structures of Basic Organic Functional Groups*", University of Arkansas, Fayetteville, AR, USA.; and Technical University of Graz, Austria, 1999.