Topological Properties of the Peptide Bond in Glycyl-L-threonine Dihydrate Based on a Fast Synchrotron/CCD-Diffraction Experiment at 100 K

Birger Dittrich,[a] Ralf Flaig,[a] Tibor Koritsánszky,[b] Hans-Georg Krane,[c] Wolfgang Morgenroth,[d] and Peter Luger*[a]

Abstract: The charge density of glycyl-L-threonine dihydrate is extracted from a synchrotron data set of 98405 reflections collected at 100 K with a Bruker CCD area detector up to a resolution of $d = 0.38$ Å ($\sin\theta/\lambda = 1.32$ Å^{-1}). The data are interpreted in terms of the "rigid pseudoatom" model. The topology of the experimental density is analyzed and compared with the topology obtained experimentally for the constituting amino acids and to that derived from Hartree–Fock calculations for the isolated molecule. All critical points of the electron density at the covalent and hydrogen bonds, as well as those of the Laplacian, were located, thereby deriving quantitative topological data for the peptide and side chain bonds. Bond topological indices in the dipeptide compare well with those of the corresponding bonds in the building amino acids, thus suggesting transferability of electronic properties of atoms and functional groups when these are derived by Bader's partitioning. Discrepancies between theoretical and experimental results could be attributed to crystal field effects.

Keywords: ab initio calculations · amino acids · electrostatic interactions · synchroton radiation · topological analysis

Introduction

Although functional complementarity and similarity are key concepts in understanding molecular recognition in biochemical processes they are hard to access with available quantum chemical methods, basically because of the size of the systems to be considered. A practical approach to this problem is to study a relevant part of the system, the so-called active site, which is believed to be characteristic of the entire process. Considerable efforts have been made recently to model biopolymers, either with the aid of semiempirical methods, or by constructing their properties from those of their constituting fragments.[1] In this respect the distribution of electronic charge ($\rho(\boldsymbol{r})$) is a property of utmost importance. The necessary information for the physically correct partitioning of an electronic system into transferable functional groups is embedded in the topology of $\rho(\boldsymbol{r})$. A chemical group is defined as a bounded region of real space based on the interatomic surface of zero flux in the gradient vector field $\nabla\rho(\boldsymbol{r})$.[2] Local features of the Laplacian distribution ($\nabla^2\rho(\boldsymbol{r})$) are recognized as plausible descriptors and predictors of molecular complementarity. In this sense, charge concentrations/depletions of one molecule can be considered to be complementary to depletions/concentrations of the other.

Chang and Bader[3] have recently demonstrated how this method works in the construction of polypeptides and of their properties by linking peptide units together. The basic point in the theoretical synthesis is how the local topology of each fragment's density changes upon the alteration of the chemical environment or the molecular conformation.

The approach outlined above requires the knowledge of the electron density which can be obtained not only by theoretical methods but also from precise, highly resolved X-ray diffraction data.[4] Experimental density, derived in such a way, refers to the crystalline state and does not correspond to a single quantum state. In spite of this, recent studies have shown, that it exhibits topological equivalence to the density obtained by quantum chemical calculations for the ground state isolated system.[5–10] The experimental procedure is being revolutionized due to technical advances—especially the application of area detectors combined with either synchrotron or conventional radiation[11]—which have rationalized

[a] Prof. Dr. P. Luger, Dipl.-Chem. B. Dittrich, Dipl.-Chem. R. Flaig
Institut für Kristallographie, Freie Universität Berlin
Takustrasse 6, 14195 Berlin (Germany)
Fax: (+49) 30 83853464
E-mail: luger@chemie.fu-berlin.de

[b] Prof. Dr. T. Koritsánszky
Department of Chemistry, University of the Witwatersrand
Private Bag 3, WITS 2050, Johannesburg (South Africa)

[c] Dr. H.-G. Krane
Mineralogisch-Petrologisches Institut der Universität Bonn
Poppelsdorfer Schloß, 53115 Bonn (Germany)

[d] Dr. W. Morgenroth
Institut für Mineralogie und Petrographie
Universität Hamburg, Grindelallee 48, 20146 Hamburg (Germany)

  0947-6539/00/0614-2582 $ 17.50+.50/0

the use of experimental densities parallel or alternative to those obtained from wave functions.

In the course of our ongoing studies we have been analyzing the experimental densities of amino acids to explore the extent to which the topology obtained by interpreting X-ray data is reproducible for chemically analogous systems and to see how it compares with theoretical predictions.[11–13]

We have been particularly interested in the comparison of the local topological indices of $\rho(\boldsymbol{r})$, which are characteristic for relevant atomic interactions: The location ($\boldsymbol{r}_b$) of the bond critical point (CP, defined by $\nabla\rho(\boldsymbol{r}_b)=0$), the value of $\rho(\boldsymbol{r}_b)$ and its Laplacian at this point, the bond path length, the bond ellipticity ε (the ratio of the negative principal curvatures at the bond CP) and finally the location and extent of bonded and non-bonded valence shell charge concentrations (VSCC).

It has been demonstrated that under similar refinement conditions these parameters are statistically equal for chemically analogous atomic interactions in different amino acid molecules and when significant differences arise, especially in polar bonds, then these can be attributed to different crystal field effects.[13]

The plausible next step towards the experimental verification of Bader's method to construct peptides from peptide residues is to examine how these bond topological indices are reproducible in an oligopeptide compared with its building block amino acids, which were previously found.

This paper reports on the topological analysis of the experimental charge density of the dipeptide glycyl-L-threonine extracted from CCD/synchrotron low temperature data and compares the result with those obtained for the individual amino acids. First results of this work were presented in [14].

Abstract in German: *Die Ladungsdichteverteilung von Glycyl-L-threonin Dihydrat wurde aus einem Synchrotrondatensatz von 98405 Reflexen bestimmt, der bei 100 K mit einem Bruker CCD Flächendetektor bis zu einer Auflösung von $d=0.38$ Å (bzw. $\sin\theta/\lambda=1.32$ Å^{-1}) gemessen wurde. Der Datensatz wurde mit Hilfe des "Pseudoatom"-Formalismus interpretiert. Eine topologische Analyse der experimentellen Ladungsdichte wurde durchgeführt und die Ergebnisse wurden einerseits mit der Topologie der entsprechenden Monoaminosäuren Gly und L-Thr verglichen, andererseits mit Befunden von Hartree–Fock-Rechnungen an dem isolierten Molekül des Dipeptids. Alle kritischen Punkte auf den kovalenten Bindungen und den Wasserstoffbrücken wurden sowohl für die Ladungsdichte als auch für die Laplacefunktion bestimmt, so daß quantitative topologische Daten für die Peptidbindung und die Bindungen in den Seitengruppen zur Verfügung stehen. Letztere stimmen gut mit den entsprechenden Bindungen in den beitragenden Aminosäuren Gly und L-Thr überein, was für die Übertragbarkeit elektronischer Eigenschaften von Atomen oder funktionellen Gruppen spricht, wenn diese nach dem Bader'schen Partitionierungskonzept behandelt werden. Abweichungen zwischen theoretischen und experimentellen Ergebnissen konnten intermolekularen Wechselwirkungen im Kristall zugeordnet werden.*

A different but strongly related approach in the description of density similarities between analogous peptide residues in different chemical environments makes use of the fact that the multipole model, applied to interpret X-ray data, relies on atomic partitioning. Chemically equivalent pseudo-atoms are expected to give the same contribution to the total density. This assumption, which is usually imposed as constraint in the refinement, has gained considerable support from experimental charge density studies of small peptides and related compounds[15–18] indicating that the multipole populations of chemically equivalent atoms are, indeed, statistically equal, despite second neighbor and crystal field effects. The transferable properties of the density parameters can be used to construct "average" aspherical densities for atoms occuring in peptides and thus, to improve the scattering model[19] used to interpret their X-ray diffraction data.

Density models and refinement strategy

The generalized scattering factor model based on the Hansen–Coppens formalism was applied.[20] The starting atomic parameters were taken from the spherical atom refinement (SHELXL)[21] starting with the isotropic room temperature structure.[22] The multipole refinements were carried out with the full-matrix LSQ program (XDLSM) of the XD program package.[23] In all cases the quantity $\Sigma_{\mathbf{H}} w_{\mathbf{H}} (|F_{obs}(\mathbf{H})-k|F_{cal}(\mathbf{H})|)^2$ was minimized using the statistical weight $w_{\mathbf{H}}=\sigma^{-2}(F_{obs}(\mathbf{H}))$. The core and the spherical valence density of the heavy atoms were composed of Hartree–Fock wave functions expanded over Slater type basis functions, while the scattering factors of the hydrogen atoms were calculated from the exact radial density functions using $\kappa=1.38$. This contraction allows a simple analytical approximation to the scattering factor introduced by Stewart et al.[24] For the deformation terms single-zeta orbitals with energy-optimized Slater exponents were taken and kept fixed.[25]

The hexadecapolar level of the multipolar expansion was used for carbon, nitrogen, and oxygen, while dipoles were used for the hydrogen atoms. A local mirror symmetry was applied to the carbon atoms in the carboxylate (C(4)) and methylene (C(1)) groups (atomic numbering scheme see Figure 1), while C_{3v} symmetry was imposed for the methyl (C(6)) group. No symmetry restriction was imposed on the N(1) atom of the ammonium group as it is involved in hydrogen bonding but the densities of the hydrogen atoms attached to it were constrained to be the same.

The unit cell contains two water molecules, one of them disordered only at the oxygen site, that is the disordered oxygen atoms share the same hydrogen atoms. For these sites the conventional refinement led to well resolved peaks of a one to two ratio which was maintained during the multipole refinement. The densities of the disordered oxygen atoms (O(61) and O(62)) were maintained equivalent to that of the well resolved water oxygen O(5). The molecule was kept neutral during the refinement and no charge transfer between the dipeptide molecule and the water molecules was allowed.

In the multipole model used, a scale factor, the atomic positions (93), anisotropic (90), and isotropic (16) temperature parameters of the heavy and of the hydrogen atoms, were refined together with their charge density parameters

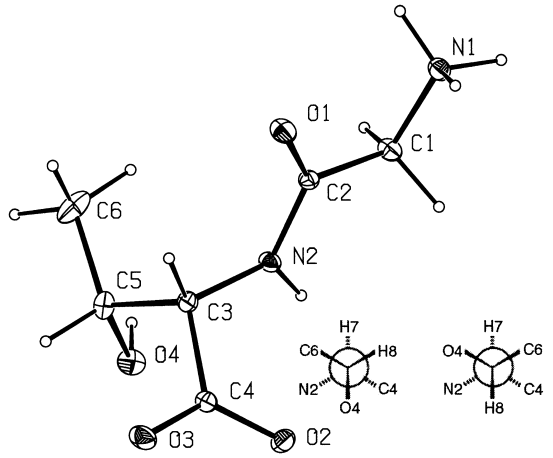

Figure 1. ORTEP representation of the experimental geometry of Gly-L-Thr. Thermal ellipsoids are shown at a probability of 50%. The left Newman projection illustrates the conformation along C5–C3, the right one shows the different conformation resulting from the geometry optimization.

(317). 9005 out of 10028 symmetry independent reflections, which met the criterion $[F_{obs}(\mathbf{H}) > 2.5\sigma(F_{obs}(\mathbf{H})]$, were included in the refinement resulting in a reflection to parameter ratio of 21. It yielded an agreement factor of 2.52% and a goodness of fit of 1.64. No significant residues of electron density were found. Figure 2 shows exemplarily the residual density map of $\rho(\boldsymbol{r})$ for the peptide group after multipole refinement. This map is almost noiseless, which suggests that the multipoles did adequately fit the experimental data. Additionally the Hirshfeld test,[26] which indicates the proper deconvolution of thermal and bonding effects, was satisfied.

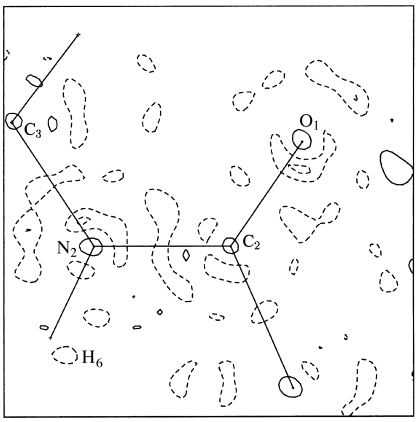

Figure 2. Residual map of the peptide group in Gly-L-Thr with contour intervals of 0.1 e Å^{-3}.

Theoretical calculations

ab initio Calculations were performed with the GAUSSIAN98[27] program package at the HF level of theory utilizing the 6-311++G(3df,3pd) as the highest standard basis set for single point energy calculations. The default options were used for SCF convergence and threshold limits applied for the final changes in the maximum forces and displacements in the geometry optimization. Because the 6-311++G(d,p) optimization, followed by an evaluation of the harmonic vibrational frequencies, resulted in a conformation different from that obtained by X-ray diffraction, single point calculations with different basis sets with the experimental geometry were performed. The wave functions resulting from these calculations were evaluated with the program package AIMPAC.[28]

Results and Discussion

The molecular structure with the chosen atom numbering scheme is shown in an ORTEP[29] representation in Figure 1. The peptide chain has the extended conformation as already stated by Yadava and Padmanabhan[22] (see corresponding torsional angles ϕ, ψ, χ, and ω in Table 1). The side chain conformation in the threonine residue expressed by the torsion χ around C^{α}–C^{β} is shown in the left insert in Figure 1. In the crystal the conformations of C^{γ} and O^{γ} are *–gauche* ($\chi_{2'} = -64.2(1)^{\circ}$) and *+gauche* ($\chi_{2''} = 59.1(1)^{\circ}$) with respect to N(2).[30] In the optimized structure C^{γ} and O^{γ} are in *trans* ($\chi_{2'} = -168.4^{\circ}$) and *–gauche* positions ($\chi_{2''} = -43.6^{\circ}$) as illustrated in the right insert in Figure 1. This arrangement was already mentioned by Yadava and Padmanabhan to be realized in the crystal of L-threonine itself and in L-threonine-L-phenylalanine-*p*-nitrobenzyl ester hydrobromide (Mallikarjunan, 1969).[31]

In contrast to the findings of Yadava and Padmanabhan it is now clear that the molecule is in the expected zwitterionic form with C–O bond lengths in the carboxylate group rather close to each other (C(4)–O(2) = 1.2650(3), C(4)–O(3) = 1.2504(4) Å). As already found in the charge density study of D,L-aspartic acid,[12] the theoretical distances of the two C–O bonds in the carboxylate group deviate from each other to a larger extent than the corresponding experimental values do. The C^{α}–C^{β} and C^{γ}–O^{γ} distances, reported by Yadava and Padmanabhan as unusually long, are in normal ranges, 1.5434(4) and 1.4324(3) Å, respectively. However, the theoretical C^{α}–C′ distance (1.5720 Å) is extremely long, which was also found in the aspartic acid study.[12]

The static deformation density in the plane of the peptide bond is shown in Figure 3, where it is compared with the theoretical map obtained from the wave function at HF/6-311++(d,p) level. Both maps are qualitatively comparable and show the expected details in the bonding and nonbonding regions. A relatively high density on the peptide bond can already be seen, which suggests a certain double bond character as will be discussed later quantitatively. A topological analysis of the experimental $\rho(\boldsymbol{r})$ was performed using the property program XDPROP of the XD system,[23] while the theoretical density was interpreted with the program AIMPAC.[28] Figure 4 shows relief plots of the negative Laplacian function. The Laplacian distributions obtained by the two methods are very similar and show the expected density accumulation in the bonds, although the bonded valence shell charge concentrations (VSCC) are more pronounced in the experimental than in the theoretical map. A quantitative comparison of the results of different theoretical calculations and the experiment is given in Table 2. Theoretical bond densities, expressed by the values of $\rho(\boldsymbol{r}_b)$, show

  0947-6539/00/0614-2584 $ 17.50+.50/0 *Chem. Eur. J.* **2000**, *6*, No. 14

Table 1. Selected experimental and optimized geometrical parameters.

Bond distances[a][Å]	exptl	HF/6–311G++(d,p)
N(1)–C(1)	1.4728(3)	1.4966
N(1)–H(1)	1.03	1.0368
N(1)–H(2)	1.00	1.0053
N(1)–H(3)	1.00	1.0048
C(1)–C(2)	1.5214(3)	1.5413
C(1)–H(4)	1.08	1.0795
C(1)–H(5)	1.08	1.0801
N(2)–C(2)	1.3403(3)	1.2913
O(1)–C(2)	1.2343(3)	1.2199
N(2)–H(6)	1.00	1.0094
N(2)–C(3)	1.4493(3)	1.4614
C(3)–C(4)	1.5390(3)	1.5720
C(3)–H(7)	1.08	1.0836
O(2)–C(4)	1.2650(3)	1.2362
O(3)–C(4)	1.2504(4)	1.2128
C(3)–C(5)	1.5434(4)	1.5409
O(4)–C(5)	1.4324(3)	1.4137
O(4)–H(9)	0.94	0.9421
C(5)–C(6)	1.5176(4)	1.5209
C(5)–H(8)	1.08	1.0842
C(6)–H(10)	1.08	1.0807
C(6)–H(11)	1.08	1.0854
C(6)–H(12)	1.08	1.0893
Bond angles [°]	**exptl**	**HF/6–311G++(d,p)**
N(1)-C(1)-C(2)	109.9(1)	104.66
C(1)-C(2)-O(1)	121.4(1)	115.48
C(1)-C(2)-N(2)	113.8(1)	114.81
O(1)-C(2)-N(2)	124.8(1)	129.71
C(2)-N(2)-C(3)	123.5(1)	127.71
N(2)-C(3)-C(4)	111.9(1)	105.15
N(2)-C(3)-C(5)	110.5(1)	110.37
C(3)-C(4)-O(2)	118.2(1)	113.64
C(3)-C(4)-O(3)	116.6(1)	115.25
O(2)-C(4)-O(3)	125.2(1)	131.10
C(4)-C(3)-C(5)	109.4(1)	112.07
C(3)-C(5)-C(6)	111.9(1)	112.52
C(3)-C(5)-O(4)	108.0(1)	110.93
O(4)-C(5)-C(6)	111.7(1)	110.83
Torsional angles [°]	**exptl**	**HF/6–311G++(d,p)**
N(1)-C(1)-C(2)-O(1)	10.7(1)	8.31
N(1)-C(1)-C(2)-N(2) [ψ_1]	–169.4(1)	–171.87
C(3)-N(2)-C(2)-O(1)	4.0(1)	–1.44
C(3)-N(2)-C(2)-C(1) [ω_1]	–175.9(1)	–178.78
C(2)-N(2)-C(3)-C(4) [ϕ_2]	–123.4(1)	–152.87
C(2)-N(2)-C(3)-C(5)	114.4(1)	86.09
N(2)-C(3)-C(4)-O(2) [$\psi_{2'}$]	–8.1(1)	–13.88
N(2)-C(3)-C(4)-O(3) [$\psi_{2'}$]	172.4(1)	166.96
N(2)-C(3)-C(5)-C(6) [$\chi_{2'}$]	–64.2(1)	–168.42
N(2)-C(3)-C(5)-O(4) [$\chi_{2''}$]	59.1(1)	–43.61
C(5)-C(3)-C(4)-O(2)	114.7(1)	106.05
C(5)-C(3)-C(4)-O(3)	–64.8(1)	–73.12
C(4)-C(3)-C(5)-O(4)	–64.6(1)	–160.45
C(4)-C(3)-C(5)-C(6)	172.1(1)	74.75

[a] The oxygen, nitrogen and carbon–hydrogen distances were set to the given values.

Table 2. Topological parameters of bonds formed by non-hydrogen atoms.[a]

Bond	ρ	$\nabla^2\rho$	ε	d	Method
N(1)–C(1)	1.58	–1.7	0.06	1.007	HF/6–311++G(d,p)
	1.62	–9.2	0.06	0.993	HF/6–311++G(2d,2p)
	1.62	–6.9	0.06	1.002	HF/6–311++G(3df,3pd)
	1.76(3)	–12.7(1)	0.05	0.825	experimental
C(1)–C(2)	1.81	–18.9	0.07	0.798	(d,p)
	1.79	–17.3	0.06	0.786	(2d,2p)
	1.83	–19.5	0.06	0.791	(3df,3pd)
	1.73(2)	–12.5(1)	0.07	0.757	experimental
O(1)–C(2)	2.71	–6.9	0.04	0.822	(d,p)
	2.78	–15.8	0.06	0.818	(2d,2p)
	2.81	–15.4	0.06	0.822	(3df,3pd)
	2.77(3)	–26.3(2)	0.07	0.734	experimental
N(2)–C(2)	2.34	–25.9	0.16	0.869	(d,p)
	2.39	–32.9	0.06	0.843	(2d,2p)
	2.41	–31.1	0.06	0.862	(3df,3pd)
	2.33(3)	–19.3(1)	0.44	0.779	experimental
N(2)–C(3)	1.64	–2.1	0.05	0.989	(d,p)
	1.68	–9.7	0.06	0.978	(2d,2p)
	1.69	–7.6	0.06	0.984	(3df,3pd)
	1.84(1)	–11.2(1)	0.31	0.828	experimental
C(3)–C(4)	1.76	–18.0	0.07	0.818	(d,p)
	1.74	–16.2	0.06	0.802	(2d,2p)
	1.78	–18.5	0.06	0.810	(3df,3pd)
	1.69(2)	–11.8(1)	0.08	0.783	experimental
O(2)–C(4)	2.52	–9.1	0.02	0.844	(d,p)
	2.59	–17.9	0.06	0.841	(2d,2p)
	2.63	–17.0	0.06	0.844	(3df,3pd)
	2.66(3)	–30.2(2)	0.07	0.800	experimental
O(3)–C(4)	2.61	–7.8	0.03	0.834	(d,p)
	2.68	–16.6	0.06	0.831	(2d,2p)
	2.70	–16.3	0.06	0.835	(3df,3pd)
	2.73(3)	–31.4(2)	0.04	0.778	experimental
O(4)–C(5)	1.63	–1.9	0.12	0.974	(d,p)
	1.67	–8.2	0.06	0.967	(2d,2p)
	1.67	–6.0	0.06	0.973	(3df,3pd)
	1.74(2)	–10.7(1)	0.03	0.837	experimental
C(3)–C(5)	1.73	–16.7	0.05	0.780	(d,p)
	1.71	–15.1	0.06	0.777	(2d,2p)
	1.75	–17.1	0.06	0.765	(3df,3pd)
	1.70(2)	–12.4(1)	0.03	0.777	experimental
C(5)–C(6)	1.77	–17.4	0.04	0.766	(d,p)
	1.75	–16.0	0.06	0.766	(2d,2p)
	1.79	–17.9	0.06	0.767	(3df,3pd)
	1.67(2)	–10.9(1)	0.20	0.748	experimental

[a] ρ and ∇^2 denote the electron density and the Laplacian at the bond critical point. ε is the bond ellipticity and d is the distance from the first atom defining the bond to the CP. Units are in e and Å.

moderate basis set dependence and agree with each other within 4% for the theoretical values and within 4–5% between experiment and theory (except for N(2)–C(3), where the difference between experiment and theory is around 10%).

The trend in the strength of the covalent bonds, revealed by this topological parameter, is of chemical relevance. Among the C–C bonds, C(1)–C(2) located next to the peptide bond has the highest bond density, however, the difference to other C–C bonds is small. The C–O bond order decreases in the following sequence: C=O, C–OO$^-$, and C–OH. The peptide bond, N(2)–C(2), is the strongest among the N–C bonds with a $\rho(\boldsymbol{r}_b)$ value of 2.33(3) e Å^{-3}, this indicates a certain double bond character. The experiment yields extra density also on the neighboring bond N(2)–C(3) with $\rho(\boldsymbol{r}_b) = 1.84(1)$ e Å^{-3} while N(1)–C(1) ($\rho(\boldsymbol{r}_b) = 1.76(3)$ e Å^{-3}) is weaker in comparison. As far as we know, these quantities are the first topological data obtained experimentally for the peptide bond. Considerable differences between theory and experiment, as well as between the different basis sets, are found for the $\nabla^2\rho(\boldsymbol{r}_b)$ values. Here the effect of introducing high angular momentum basis functions is especially pronounced. For the polar bonds (C–O and C–N) the experiment gives considerably lower Laplacian values at the bond critical point (CP)

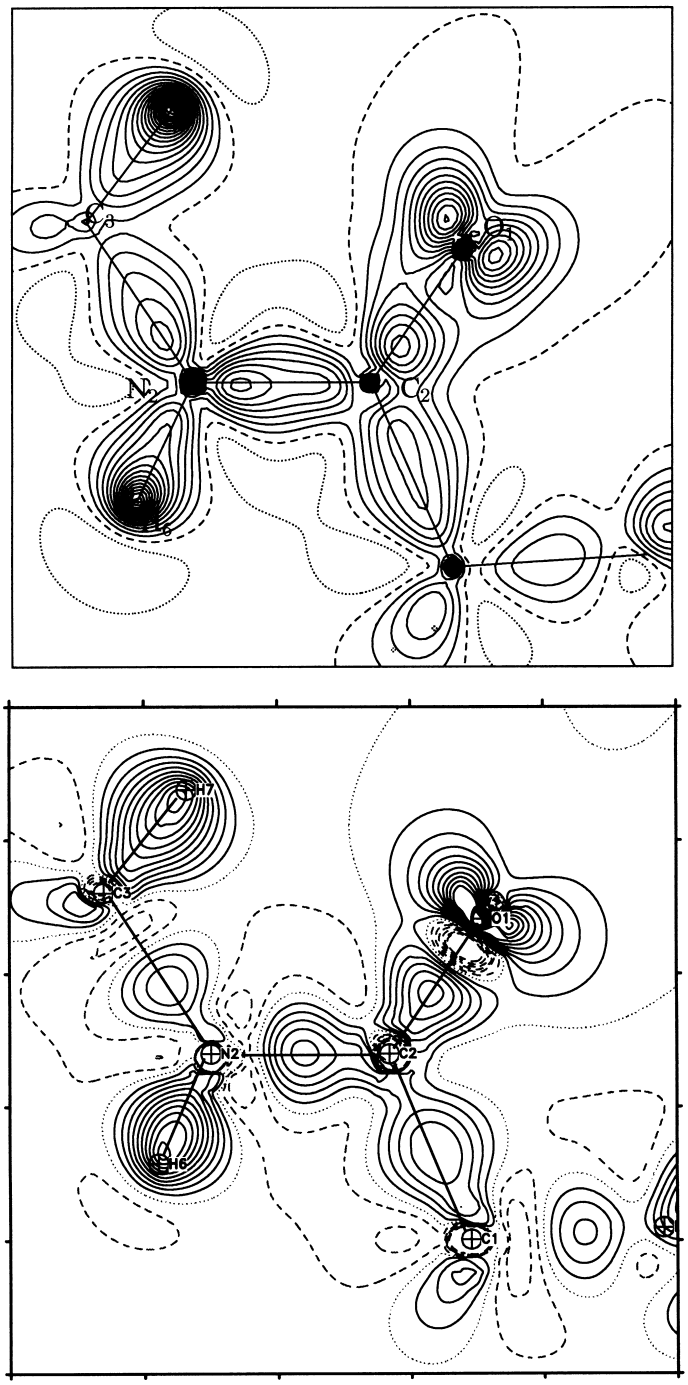

Figure 3. Static experimental (above) and theoretical (below) deformation density map in the plane of the peptide bond in Gly-L-Thr; contour intervals are 0.1 e Å^{-3}.

with the exception of N(2)–C(2) compared with the theoretical calculations, whereas for the C–C bonds the opposite trend is seen. These findings are in line with the results of a number of earlier studies on small organic molecules.[5, 6, 12, 13, 32]

For the contributing amino acids, L-threonine[13] and glycine,[33] full topological analyses have been executed so that a quantitative comparison between the monomers and the dipeptide can be made, see Table 3. The agreement is exceptionally good for $\rho(\boldsymbol{r}_b)$, except for the two bonds of the

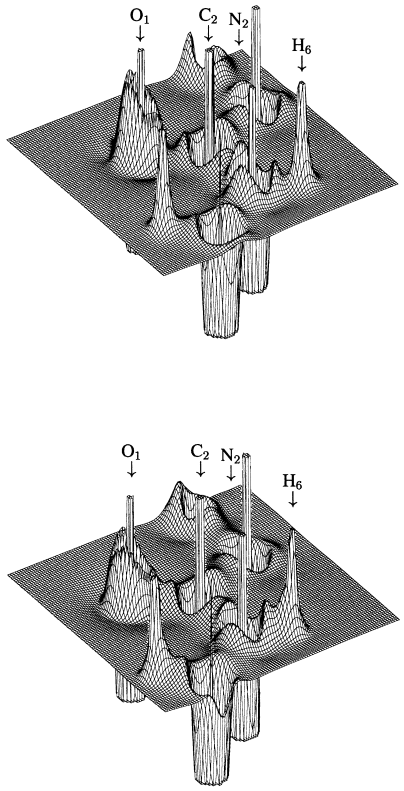

Figure 4. Experimental (above) and theoretical (below) Laplacian map in the plane of the peptide bond in Gly-L-Thr.

dipeptide with C(3) as contributor. This atom is next to the peptide bond, so that it is no longer equivalent to a C$^{\alpha}$ atom of a monomeric amino acid. As already mentioned above, the bond C(3)–N(2) has some extra charge (1.84 e Å^{-3} compared with 1.71 e Å^{-3} found as average of nine amino acids for

Table 3. Comparison of bond topological parameters of Gly-L-Thr, L-Thr and Gly.[a]

Bond	Compound	ρ	$\nabla^2\rho$	ε	d
C–O(1)	Gly-L-Thr	2.66(3)	– 30.2(2)	0.07	0.800
	L-Thr	2.64(4)	– 30.7(3)	0.04	0.814
	Gly	2.67	– 30.5	–	–
C–O(2)	Gly-L-Thr	2.73(3)	– 31.4(2)	0.07	0.778
	L-Thr	2.78(4)	– 38.1(2)	0.08	0.796
	Gly	2.77	– 32.8	–	–
N–C$_{\alpha}$	Gly-L-Thr	1.76(3)	– 12.7(1)	0.05	0.825
	L-Thr	1.72(3)	– 13.2(1)	0.18	0.851
	Gly	1.69	– 11.9	–	–
C(3)–C(4)	Gly-L-Thr	1.69(2)	– 11.8(1)	0.07	0.783
C–Cα	L-Thr	1.80(3)	– 15.4(1)	0.25	0.725
	Gly	1.78	– 15.6	–	–
C(3)–C(5)	Gly-L-Thr	1.70(2)	– 12.4(1)	0.03	0.777
C(1)–C(3)	L-Thr	1.80(3)	– 13.8(1)	0.05	0.756
C(5)–O(4)	Gly-L-Thr	1.71(2)	– 10.7(1)	0.03	0.837
C(3)–O(3)	L-Thr	1.85(3)	– 11.4(1)	0.16	0.799
C(5)–C(6)	Gly-L-Thr	1.67(2)	– 10.9(1)	0.20	0.748
C(3)–C(4)	L-Thr	1.66(3)	– 12.8(1)	0.05	0.833

[a] ρ and $\nabla^2\rho$ denote the electron density and the Laplacian at the bond critical point. ε is the bond ellipticity and d is the distance from the first atom defining the bond to the CP. Units are in e and Å.

0947-6539/00/0614-2586 $ 17.50+.50/0 *Chem. Eur. J.* **2000**, *6*, No. 14

Table 4. Hydrogen bonds and their bond topological parameters.

D–H···A[a]	Symm. op.	D–H	H···A	D···A	D–H···A	ρ	$\nabla^2\rho$	ε
O(4)–H(9)···O(3)	$1-x, -\frac{1}{2}+y, 1\frac{1}{2}-z$	0.94	1.8463	2.7638(4)	164.5	0.17(2)	2.9(1)	0.12
O(5)–H(13)···O(2)	$1-x, \frac{1}{2}+y, -\frac{1}{2}-z$	0.95	1.8130	2.7449(4)	166.2	0.27(1)	3.4(1)	0.06
O(5)–H(14)···O(3)	$1\frac{1}{2}-x, 1-y, \frac{1}{2}+z$	0.95	1.7462	2.6827(4)	168.1	0.31(1)	4.2(1)	0.05
O(61)[b]–H(15)···O(2)	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$	0.95	1.9761	2.8664(10)	155.3	0.09(1)	2.5(1)	0.14
O(61)[b]–H(16)···O(5)	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$	0.97	1.7789	2.7500(8)	177.8	0.20(1)	4.1(1)	0.03
N(1)–H(1)···O(4)	$\frac{1}{2}+x, 1\frac{1}{2}-y, 1-z$	1.03	2.4843	3.0922(4)	117.1	0.04(1)	0.6(1)	0.03
N(1)–H(1)···O(5)	x, y, z	1.03	1.9401	2.8233(4)	141.9	0.15(1)	2.5(1)	0.03
N(1)–H(2)···O(2)	$-\frac{1}{2}+x, 1\frac{1}{2}-y, 1-z$	1.00	1.8222	2.8221(3)	179.6	0.17(1)	2.8(1)	0.09
N(1)–H(3)···O(61)[b]	x, y, z	1.00	1.7343	2.7069(9)	163.1	0.36(1)	5.5(1)	0.09
N(2)–H(6)···O(1)	$-\frac{1}{2}+x, 1\frac{1}{2}-y, 1-z$	1.00	2.0197	2.9894(3)	162.7	0.14(2)	1.8(1)	0.05
C(1)–H(4)···O(1)	$\frac{1}{2}+x, 1\frac{1}{2}-y, 1-z$	1.08	2.5076	3.2187(3)	122.5	0.05(1)	0.7(1)	0.13
C(1)–H(5)···O(4)	$1-x, -\frac{1}{2}+y, 1\frac{1}{2}-z$	1.08	2.3058	3.3802(4)	172.9	0.07(1)	1.1(1)	0.05

[a] Units are in e and degree and Å. D denotes the donating, A the accepting atom. ρ und $\nabla^2\rho$ denote the electron density and the Laplacian at the bond critical point. ε is the bond ellipticity. [b] Data are given for the disorderd atom with more probability.

C^{α}–$N(H_3)$).[13] This is at the expense of charge depletion in the bonds C(3)–C(4) and C(3)–C(5).

A summary of the geometrical and topological parameters of the hydrogen bonds is listed in Table 4. All N–H and O–H groups act as donors, N(1)–H(1) is donor in a bifurcated hydrogen bond to O(4) and O(5), where, however, the rather long N···O and H···O distances indicate N(1)–H(1)···O(4) as the weakest of all hydrogen bonds. Data for the contacts involving the disordered water molecule are considered only for the site O(61) with the higher occupation. O(1) and O(4) are also bonded to C–H donors with O···H distances shorter than the sum of the van der Waals radii (2.72 Å). For all cases low density and positive Laplacians were found at the CPs, as it is expected for closed shell interactions. There is a satisfactory correlation between the hydrogen bond strength as expressed by the hydrogen-acceptor distance and the density found at the hydrogen bond CP. This relation was first mentioned by Espinosa et al.[34]

The distribution of $\nabla^2\rho(\boldsymbol{r})$ allows an exact assignment also of non-bonding electron pairs. Table 5 lists the experimental non-bonded VSCCs of the oxygen atoms in terms of the $\rho(\boldsymbol{r})$ and the $\nabla^2\rho(\boldsymbol{r})$ values of the $(3,+3)$ CPs of the Laplacian. The geometrical arrangements of the lone pairs at each oxygen atom are given by the O–CP distances and the C-O-CP and the CP_1-O-CP_2 angles. The relative locations of the non-bonded VSSCs seem to be correlated with the strengths and directions of the hydrogen bonds. For example, the larger CP_1-O-CP_2 angle at O(2), as compared with that at O(3), can be attributed to the former being an acceptor in three rather than two strong hydrogen bonds. The extra interaction, in which the O(2) atom is involved, disturbs the expected symmetric arrangement of the VSSCs (as suggested by the VSEPR model) compared with that found for the O(3) atom. The lone pairs at the carbonyl oxygen O(1) exhibit a symmetric geometry, that can be attributed to an sp^2 hybridization in spite of the the different strength of the hydrogen bonds this atom participates in.

Table 5. Nonbonded valence shell charge concentration of oxygen atoms.[a]

Atom	ρ	R	$\nabla^2\rho$	C-O-CP	CP_1-O-CP_2
O(1)	6.44	0.338	–147.1	117.77	–
	6.60	0.337	–161.0	123.38	117.79
O(2)	6.49	0.339	–150.9	108.49	–
	6.20	0.342	–125.9	93.39	157.39
O(3)	6.49	0.339	–145.3	118.16	–
	6.20	0.341	–137.6	109.57	129.21
O(4)	6.33	0.340	–125.0	93.90	–
	6.05	0.343	–121.8	155.51	94.10

[a] Units are e, Å and degrees. R is the radial distance of the $(3,+3)$ CP of the Laplacian from the oxygen atom, C-O-CP is the angle formed by the C–O bond and the O–CP vector, CP_1-O-CP_2 is the angle formed by the CP_1-O and O-CP_2 vectors.

The three dimensional representation of the Laplacian distribution displayed in Figure 5 can be used as a tool to map fine charge reorganizations due to weak intermolecular interactions such as hydrogen bonding. The holes in the green isosurface ($\nabla^2\rho(\boldsymbol{r}) = +12.5$ e Å^{-5}) indicate regions of smaller charge density depletion caused by the local hydrogen bond charge concentration. Each vector from a hydrogen donor atom to an acceptor oxygen atom passes these regions. Therefore the Laplacian distribution makes directions visible, where chemical interactions and intermolecular recognition processes are favored.

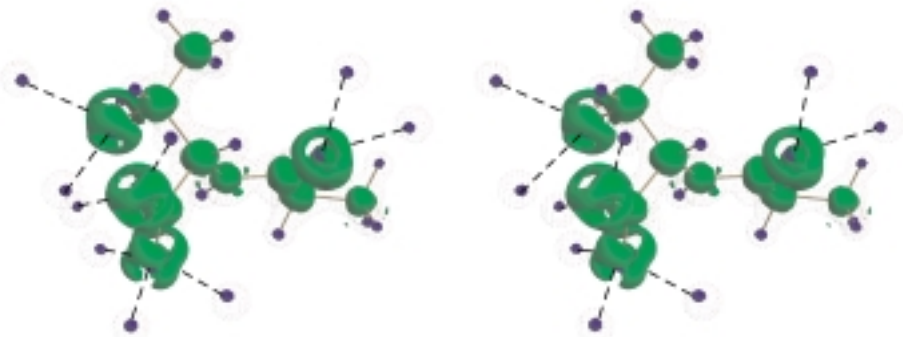

Figure 5. Stereo representation of three isosurfaces of the Laplacian function of the dipeptide Gly-L-Thr. Blue = -100 e Å^{-5}, red dots = zero surface, green = $+12$ e Å^{-5}. In addition to the molecule all hydrogen atoms of symmetry related molecules that are involved in hydrogen bonds are drawn. All connecting vectors from the oxygen atoms to these hydrogen atoms pass through holes in the green isosurface.

The electrostatic potential (EP), calculated from the multipole model using the method of Su and Coppens,[35] is displayed on the left side of Figure 6. It represents an isolated molecule which is extracted from the crystal, but still bears the polarization effects induced by intermolecular interactions. Thus, a three dimensional distribution of the EP in a chemical environment is obtained which surely simulates, for example, physiological conditions better than an EP extracted from an isolated molecule, being shown on the right side of Figure 6 as result of the quantum chemical calculations. The EP not only plays an important role in chemical reactivity, but also in the evaluation of molecular recognition (for example, drug receptor interaction). Isosurface representations in Figure 6

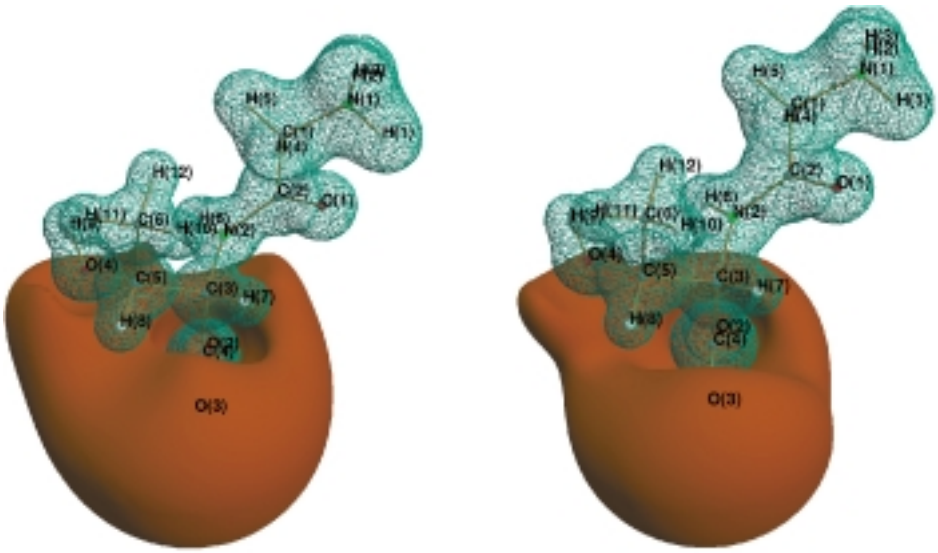

Figure 6. Experimental (left) and theoretical (right) electrostatic potential in a three dimensional isosurface representation of Gly-L-Thr. The surfaces correspond to potentials of $-0.18\,e\,Å^{-1}$ (red) and $0.6\,e\,Å^{-1}$ (blue).

both show a large basin of electronegative region around the carboxylate group that is extended continuously to the lone pair region of O(4) and partly to the carbonyl oxygen O(1). This region is more extended in the experimental than in the theoretical distribution. We attribute this to the polarization effects arising from intermolecular interactions in the crystal.

Conclusion

This detailed investigation of the dipeptide glycyl-L-threonine includes a topological analysis of the charge density and the related properties which is based on a fast synchrotron experiment. HF wavefunctions were also evaluated. On the experimental side, we found an excellent accordance of electronic properties between the constituting amino acids and the dipeptide. We provide quantitative topological data for the electronic properties of the peptide bond, which have not appeared in the literature until now.

Peptides and proteins can be regarded as consisting of peptide bonds forming a backbone to which different side chain residues are attached. The investigation of glycyl-L-threonine, and ongoing work on another dipeptide and a hexapeptide suggest, that one has to differentiate between side chains, for which a high degree of transferability is found and the backbone, where the situation is different. The COO^- and NH_3^+ termini of the amino acid building units change chemically upon forming a peptide bond. Thus there is no straightforward way to deduce transferability of bond topological parameters from amino acids to peptides. There is ongoing work in our group to address this problem by using more suitable model compounds. Further studies of this kind should reveal, how different side chain residues influence the electronic properties of the peptide bond.

The influence of intermolecular interactions on multipole-refined densities has been the subject of many recent studies.[34, 36, 37] The latest model study, based on theoretical noise-free static data, concluded that these effects can be measurable.[38] Hydrogen bonding manifests itself, as seen in terms of interaction density, in charge rearrangements at both the donor and acceptor atoms resulting in extra charge accumulations in the donor-hydrogen bond, as well as in the valence region of the acceptor. The latter charge concentration changes the topology of the density in the carbonyl bond, giving rise to a shift in the location of the bond CP and to an enhancement of the bonded VSSC. In a polar bond the bond CP is located closer to the less electronegative atom, at a place where the Laplacian changes rapidly. For all C–O bonds in glycyl-L-threonine the experimental bond CPs were found to be shifted away from the carbon atoms relative to their theoretical positions. This is in accord with our findings in the charge density study on aspartic acid.[12] The sensitivity of the topology of the density to crystal field effects, though not as apparent as that of the electrostatic potential, suggests that topological analysis is certainly becoming an efficient tool for detailed examination of weak interactions in molecular crystals.

Experimental Section

A colorless crystal of glycyl-L-threonine with the dimensions $0.60 \times 0.40 \times 0.38\ mm^3$ was grown by applying vapor diffusion (water/ethanol) methods with the commercially available product. X-ray data were measured on a Huber four-circle diffractometer at the beamline D3 of the storage ring DORIS III at the HASYLAB/Deutsches Elektronen Synchrotron, Hamburg. A wavelength of 0.5 Å for the primary beam was used. The temperature was maintained at 100 K during the measurement with an Oxford Cryosystems N_2 gas stream cooling device. The CCD area detector allowed to measure 98405 reflections in 2 d up to a resolution of $\sin\theta/\lambda = 1.32\ Å^{-1}$ (or $d = 0.38$ Å). Data were collected for two different positions of the detector. For the 2θ positions 30° and 50° a total number of 2420 frames were collected with a scan width of 0.1° in ω and an exposure time of 3 and 5 s, respectively.

The measurement strategy was planned with ASTRO,[39] the data collection was monitored with SMART[39] and the frames were integrated and

0947-6539/00/0614-2588 $ 17.50+.50/0 *Chem. Eur. J.* **2000**, *6*, No. 14

Table 6. Crystal data and structure refinement for glycyl-L-threonine.

empirical formula	$C_6H_{12}N_2O_4 \cdot 2H_2O$
formula weight [g mol^{-1}]	212.2
crystal system	orthorhombic
space group	$P2_12_12_1$ (No. 19)
Z	4
temperature [°K]	100
unit cell dimensions:	
a [Å]	9.5731(7)
b [Å]	10.0496(5)
c [Å]	10.5731(5)
$\alpha=\beta=\gamma$ [°]	90.0
V [Å^3]	1017.19(6)
calculated density [g cm^{-3}]	1.386
$F(000)$	456.0
absorption coefficient μ [mm^{-1}]	0.07
crystal size [mm^3]	$0.6 \times 0.4 \times 0.38$
λ [Å]	0.5
max. 2θ [°]	82.67
$(\sin\theta/\lambda)_{max}$ [Å^{-1}]	1.32
limiting indices	$-12 \le h \le 24$, $-26 \le k \le 24$, $-24 \le l \le 27$
number of collected reflections	98405
symmetry independent reflections	10028
reflections with $F_o > 2.5\sigma(F_o)$	9005
completeness	98%
mean/max redundancy from ASTRO	13/72
redundancy after integration	9.8
R_{int}	0.0423
R_w (spherical/aspherical)	0.034, 0.019
R_1 (spherical/aspherical)	0.033, 0.025
$R_{all}(F)$	0.0337
Gof (spherical/aspherical)	2.89, 1.64

corrected with the SAINT[39] and SADABS[40] programs. Further details on the crystal data and the experimental conditions are given in Table 6.

Crystallographic data (excluding structure factors) for the structure reported in this paper have been deposited with the Cambridge Crystallographic Data Centre as supplementary publication no. CCDC-136044. Copies of the data can be obtained free of charge on application to CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: (+44) 1223-336-033; e-mail: deposit@ccdc.cam.ac.uk).

Acknowledgement

We thank the Deutsche Forschungsgemeinschaft DFG (Project No. Lu 222/22-1), the Bundesminister für Bildung, Wissenschaft, Forschung und Technologie, BMBF (Grant no. 05 647 KEA1) and the Fonds der Chemischen Industrie for financial support. Moreover we are grateful to Prof. R. Destro, Milan/Italy, for having provided us with his glycine topological data.

[1] P. L. A. Popelier, R. F. W. Bader, *J. Phys. Chem.* **1994**, *98*, 4473.
[2] R. F. W. Bader, *Atoms in Molecules—A Quantum Theory*, Clarendon Press, Oxford, **1990**.
[3] C. Chang, R. F. W. Bader, *J. Phys. Chem.* **1992**, *96*, 1654.
[4] P. Coppens, *X-Ray Charge Densities and Chemical Bonding*, Oxford University Press, New York, **1997**.
[5] C. Gatti, R. Bianchi, R. Destro, F. Merati, *J. Mol. Struct.* (*THEOCHEM*) **1992**, *255*, 409.
[6] C. Flensburg, S. Larsen, R. F. Stewart, *J. Phys. Chem.* **1995**, *99*, 10130.
[7] T. Koritsánszky, J. Buschmann, P. Luger, *J. Phys. Chem.* **1996**, *100*, 10547.
[8] P. Fuhrmann, T. Koritsánszky, P. Luger, *Z. Kristallogr.* **1997**, *212*, 213.
[9] T. Koritsánszky, J. Buschmann, D. Lentz, P. Luger, G. Perpetuo, M. Röttger, *Chem. Eur. J.* **1999**, *5*, 3413.
[10] T. Koritsánszky, D. Zobel, P. Luger, *J. Phys. Chem. A* **2000**, *104*, 1549.
[11] T. Koritsánszky, R. Flaig, D. Zobel, H. G. Krane, W. Morgenroth, P. Luger, *Science* **1998**, *279*, 356.
[12] R. Flaig, T. Koritsánszky, D. Zobel, P. Luger, *J. Am. Chem. Soc.* **1998**, *120*, 2227.
[13] R. Flaig, T. Koritsánszky, J. Janczak, H. G. Krane, W. Morgenroth, P. Luger, *Angew. Chem.* **1999**, *111*, 1494; *Angew. Chem. Int. Ed.* **1999**, *38*, 1397.
[14] B. Dittrich, A. Wagner, R. Flaig, P. Luger, W. Morgenroth, H. G. Krane, *Z. Kristallogr.* **1999**, Supp. No. 16, 74.
[15] M. Souhassou, C. Lecomte, R. H. Blessing, A. Aubry, M. M. Rohmer, R. Wiest, M. Bénard, M. Marraud, *Acta Crystallogr.* **1991**, *B47*, 253.
[16] M. Souhassou, C. Lecomte, N. E. Ghermani, M. M. Rohmer, R. Wiest, M. Bénard, R. H. Blessing, *J. Am. Chem. Soc.* **1992**, *114*, 2371.
[17] V. Pichon-Pesme, C. Lecomte, R. Wiest, M. Bénard, *J. Am. Chem. Soc.* **1992**, *114*, 2713.
[18] R. Wiest, V. Pichon-Pesme, M. Bénard, C. Lecomte, *J. Phys. Chem.* **1994**, *98*, 1351.
[19] C. Jelsch, V. Pichon-Pesme, C. Lecomte, A. Aubry, *Acta Crystallogr.* **1998**, *D54*, 1306.
[20] N. K. Hansen, P. Coppens, *Acta Crystallogr.* **1978**, *A34*, 909.
[21] G. M. Sheldrick, *SHELX 97*, Universität Göttingen, **1997**.
[22] V. S. Yadava, V. M. Padmanabhan, *Acta Crystallogr.* **1973**, *B29*, 854.
[23] T. Koritsánszky, S. Howard, P. R. Mallinson, Z. Su, T. Richter, N. K. Hansen, *XD—A computer program package for multipole refinement and analysis of electron densities from diffraction data. User manual*, Freie Universität Berlin, **1997**.
[24] R. F. Stewart, E. R. Davidson, W. T. Simpson, *J. Chem. Phys.* **1965**, *42*, 3175.
[25] E. Clementi, C. Roetti, *Atomic Data and Nuclear Data Tables* **1974**, *14*, 177.
[26] F. L. Hirshfeld, *Acta Crystallgr.* **1976**, *A32*, 239.
[27] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, GAUSSIAN 98, Revision A.6, Gaussian, Inc., Pittsburgh PA, **1998**.
[28] J. Cheeseman, T. A. Keith, R. F. W. Bader, *AIMPAC program package*, McMaster University, Hamilton, Ontario, **1992**.
[29] M. N. Burnett, C. K. Johnson, ORTEP-III, Oak Ridge Thermal Ellipsoid Plot Program for Crystal Structure Illustrations, Oak Ridge National Laboratory Report ORNL-6895, Oak Ridge, Tennessee, **1996**.
[30] E. Benedetti, G. Morelli, G. Némethy, H. A. Scheraga, *Int. J. Peptide Protein Res.* **1983**, *22*, 1.
[31] M. Mallikarjunan, S. T. Rao, K. Venkatesan, V. R. Sarma, *Acta Crystallogr.* **1969**, *B25*, 220.
[32] S. T. Howard, M. B. Hursthouse, C. W. Lehmann, E. A. Poyner, *Acta Crystallogr.* **1995**, *B51*, 328.
[33] R. Destro, P. Roversi, M. Barzaghi, R. E. Marsh, *J. Phys. Chem. A* **2000**, *104*, 1047.
[34] E. Espinosa, M. Souhassou, H. Lachekar, C. Lecomte, *Acta Crystallogr.* **1999**, *B55*, 563.
[35] Z. W. Su, P. Coppens, *Acta Crystallogr.* **1992**, *A48*, 188.
[36] M. P. C. M. Krijn, D. Feil, *J. Chem. Phys.* **1988**, *89*, 5787.
[37] M. P. C. M. Krijn, H. Graafsma, D. Feil, *Acta Crystallogr.* **1988**, *B44*, 609.
[38] M. A. Spackman, P. G. Byrom, M. Alfredsson, K. Hermansson, *Acta Crystallogr.* **1999**, *A55*, 30.
[39] Programs ASTRO 1995–1996, SMART 1996, SAINT 1994–1996, Bruker-AXS Inc., Madison, WI (USA).
[40] R. H. Blessing, *Acta Crystallogr.* **1995**, *A51*, 33.

Received: November 11, 1999 [F2133]