Acta Crystallographica Section B Structural Science

ISSN 0108-7681

Andreas Hofmann, Jeanette Netzel and Sander van Smaalen*

Laboratory of Crystallography, University of Bayreuth, D-95440 Bayreuth, Germany

Correspondence e-mail: smash@uni-bayreuth.de

Accurate charge density of trialanine: a comparison of the multipole formalism and the maximum entropy method (MEM)

Received 2 August 2006 Accepted 2 December 2006

An accurate charge density study of trialanine is presented with the maximum entropy method (MEM), on the basis of the same reflection data as was used for a multipole refinement [Rödel et al. (2006). Org. Biomol. Chem. 4, 475-481]. With the MEM, the optimum fit to the data is found to correspond to a final value of χ^2 which is less than its statistical expectation value N_{Ref} , where N_{Ref} is the number of reflections. A refinement strategy is presented that determines the optimal goal for χ^2 . It is shown that the MEM and the multipole method are on a par with regard to the reproduction of atomic charges and volumes, general topological features and trends in the charge density in the bond critical points (BCPs). Regarding the values of the charge densities in the BCPs, agreement between quantum chemical calculations, the multipole method and MEM is good, but not perfect. In the case of the Laplacians, the coincidence is not as good and especially the Laplacians of the C-O bonds differ strongly. One of the reasons for the observed differences in the topological parameters in the BCPs is the fact that MEM densities still include the effects of thermal motion, whereas multipole densities are free from the effects of thermal motion. Hydrogen bonds are more convincingly reproduced by the MEM than by multipole models.

1. Introduction

Measured data of any kind are usually afflicted by statistical noise. For a reasonable interpretation of an experiment, it is necessary to extract as much information as possible from the data within the limits imposed by the statistical noise. One way of doing this is the maximum entropy approach which is used to find the most probable values that correspond to the measured data, by maximizing the informational entropy (Buck & Macauley, 1994; Jaynes, 1957, 1979, 1986).

Although the MEM has been successfully used to tackle various crystallographic problems in the fields of data processing, powder diffraction and solving the phase problem (Gilmore, 1996), there is still some dispute about its usefulness in the determination of accurate charge densities. This is mainly due to the fact that the reconstructed electron density $[\rho^{\text{MEM}}(r)]$ is affected by artefacts such as spurious maxima or 'ripples' in the charge-density distribution which are specific to the MEM (Jauch & Palmer, 1993; Jauch, 1994; de Vries *et al.*, 1996; Roversi *et al.*, 1998; Palatinus & van Smaalen, 2002; Roversi *et al.*, 2002). In recent years, several improvements have been introduced into the MEM that should solve these problems (*e.g.* applying a non-uniform prior and prior-derived *F* constraints; de Vries *et al.*, 1996; Iversen *et al.*, 1997; Palatinus & van Smaalen, 2005).

© 2007 International Union of Crystallography Printed in Singapore – all rights reserved

Table 1	
Crystallographic data	a (Rödel, 2003).

Formula	$C_9H_{17}N_3O_4 \cdot H_20$
M_r	243.3
Space group	Monoclinic, C2
Ź	8
<i>T</i> (K)	20
a (Å)	18.441 (2)
b (Å)	5.215 (1)
c (Å)	24.854 (3)
β (°)	98.765 (2)
$V(Å^3)$	2362.4 (5)
$\sin \theta / \lambda \max (\text{\AA}^{-1})$	1.15
Unique data (measured / prior derived)	14 895/281 077
Completeness (%)	93.3
R _{int}	0.0295
R (ISAM)	0.0314

The eventual aim of these improvements is to achieve a $\rho^{\text{MEM}}(r)$ map which is at least comparable in quality to electron-density maps $[\rho^{\text{multipole}}(r)]$ that are obtained by refinements based on the multipole formalism (Hansen & Coppens, 1978). We are particularly interested in charge-density studies of polypeptides and small proteins - a field where the multipole refinement of each individual atom becomes more and more difficult to perform owing to the increasing number of parameters and the occurrence of correlations between them. Before using the MEM on yet unstudied peptides we want to show that the MEM can produce reliable charge-density maps for these kinds of systems. Therefore, we performed an MEM refinement on the same 20 K X-ray diffraction data of the tripeptide L-alanyl-L-alanine (trialanine) which has already been used for a multipole refinement (Rödel et al., 2006).

2. The maximum entropy method

The basis for the application of the MEM is a discretized electron density on an $N_1 \times N_2 \times N_3 = N_p$ grid over the unit cell, with $\rho^k = \rho(\mathbf{x}_k)$ and \mathbf{x}_k being the position of pixel k. In this work, the entropy S of a discrete electron density is defined as

$$S = -\sum_{k=1}^{N_p} \rho^k \log\left(\frac{\rho^k}{\rho_{\text{prior}}^k}\right),\tag{1}$$

where the values of ρ_{prior} define the prior or reference electron density. The basic principle of the MEM is that the optimal electron density is defined to be the electron density { ρ^k } that maximizes the entropy *S*, while one or more constraints are fulfilled. Besides the normalization of { ρ^k },

$$C_0 = -1 + \frac{1}{\rho_{\text{total}}} \cdot \sum_{k=1}^{N_P} \rho^k, \qquad (2)$$

the most important constraint is the so-called F constraint which incorporates the measured structure factors in the maximum entropy calculation

$$C_F = -\chi^2 + \sum_{hkl}^{N_{\text{ref}}} \left(w_{hkl} \frac{\left| F_{hkl}^{\text{obs}} - F_{hkl}^{\text{MEM}} \right|^2}{\sigma^2} \right).$$
(3)

Here F_{hkl}^{obs} and F_{hkl}^{MEM} denote the measured and MEM-calculated phased structure factors of the (hkl) reflection. The w_{hkl} factor allows for weighting; its value is 1.0 if no weights are applied.

These constraints are chosen in a way that requires them to become zero when the conditions they represent are fulfilled. However, since only the derivative of the constraints occur in the iterations, the absolute value of χ^2 is irrelevant for the minimization procedure. On the other hand, its value is important as a stopping criterion. Convergence is tested by comparison of the constraint value [see (3)] computed with $w_{hkl} = 1.0$ with the stopping criterion. In the historical MEM the stopping criterion corresponds to the classical leastsquares refinement (Gull & Skilling, 1999). The constraint is fulfilled if $\chi^2 = N_{\text{Ref.}}$ According to Gull & Skilling (1999), the historic MEM is not Bayesian and therefore imperfect. The constraint $\chi^2 = N_{\text{Ref}}$ is only an approximation to the maximization of the true likelihood $Pr(F^{obs}|\rho^{MEM})$; no single selected ρ^{MEM} can fully represent the posterior probability $\Pr(\rho^{\text{MEM}}|F^{\text{obs}})$ which theory demands, and it is difficult to define the number N_{Ref} of fully independent data in a suitable invariant manner. It is well known that the constraint χ^2 = N_{Ref} gives systematically under-fitted reconstructions. The reason is that the χ^2 statistics between F^{obs} and F^{MEM} will indeed average to N_{Ref} if ρ^{MEM} is the real electron-density distribution. However, this is unattainable and the computed ρ^{MEM} will necessarily be biased towards the data, so that the misfit is reduced. Accordingly, $\chi^2 = N_{\text{Ref}}$ is too pessimistic. Therefore, Gull & Skilling (1999) recommend the use of the classical MEM which is truly Bayesian and does not rely on the χ^2 statistics as a convergence criterion. Unfortunately, this classic MEM is incompatible with the necessary MEM enhancements such as prior-derived F constraints (Palatinus & van Smaalen, 2005) and ad hoc weighting (de Vries et al., 1996). In order to retain compatibility of these MEM enhancements, the historical MEM for the calculation of electron density maps should be kept. However, since χ^2 = $N_{\rm Ref}$ is too pessimistic, a value of χ^2 smaller than $N_{\rm Ref}$ is desirable and, as pointed out above, mathematically justified. Since the optimal value of χ^2 also depends on the number of reflections, we define $\chi^2_{aim} = \chi^2 / N_{Ref}$, with the expectation value of $\langle \chi^2_{aim} \rangle = 1.0$ in the case of classical least-squares refinement.

It has been shown by Jauch & Palmer (1993) that the distribution of the normalized residuals for ρ^{MEM} is not Gaussian as desired, but that a few strong low-angle reflections account for the main part of χ^2 . The remaining reflections, however, are over-fitted to satisfy the requirement $\chi^2/N_{\text{Ref}} = \chi^2_{\text{aim}}$. Several methods have been suggested to counterbalance this effect (de Vries *et al.* 1994; Iversen *et al.*, 1997; Palatinus & van Smaalen, 2005). In this work the *ad hoc* weighting scheme as suggested by de Vries *et al.* (1994) is

applied. The resulting equation for the F constraint is (3), where the weights w_{hkl} are defined as

$$w_{hkl} = \frac{1}{|H_{hkl}|^{n}} \cdot \left(\frac{1}{N_{\text{ref}}} \sum_{hkl}^{N_{\text{ref}}} \frac{1}{|H_{hkl}|^{n}}\right)^{-1}$$
(4)

and $\mathbf{H} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$. These weights will be denoted as Hn (*n* is the power of the inverse reciprocal lattice vector). *H*0 means no *ad hoc* weighting ($w_{hkl} = 1$ for all *hkl*). A weighting scheme Hn (n > 0) results in a more Gaussian-like distribution of the residuals. Based on an empirical investigation, de Vries *et al.* (1994) found that n = 4 (*H*4) gives the best results.

3. Experimental

3.1. Refinement

Data collection (Mo $K\alpha$ radiation at 20 K) and data reduction have been described in Rödel *et al.* (2006), who generously gave us a copy of the reflection data file. The most important crystallographic data are summarized in Table 1. Refinements with the independent spherical atom model (ISAM) were performed with the computer program *JANA*2000 (Petricek *et al.*, 2000), using the coordinates from the multipole refinement as starting positions for all non-H atoms. C–H bond lengths were fixed to the values known from neutron scattering experiments at low temperatures (Steiner & Saenger, 1993). This choice was motivated by the fact that H atoms at neutron distances provide the desired reference point for the comparison of ISAM and final densities. Furthermore, initial MEM calculations with H atoms either at neutron positions or at positions known from free



Figure 1

Perspective representation of both crystallographic independent trialanine molecules, together with the water molecules and with all hydrogen bonds (dashed lines). Bond critical points are given for all hydrogen bonds (green dots) and all C-C, C-N and C-O bonds (cyan dots). All 12 hydrogen bonds are shown. Symmetry related atoms refer to the following symmetry operations: (i) $-\frac{1}{2} + x$, $-\frac{1}{2} + y$, z; (ii) x, 1 + y, z; (iii) 1 - x, y, -z; (iv) 1 - x, y, 1 - z.

refinements against X-ray data have shown a more smooth convergence of the MEM in the case of neutron positions for H atoms, despite the slightly worse fit of the ISAM refinement with neutron positions ($R_F = 0.031$) compared with the ISAM refinement with X-ray positions (0.029; see Rödel *et al.*, 2006). An instability factor of 0.005 was used. The crystal structure is shown in Fig. 1. The input file for *BayMEM* (phased reflection file) was then created with *JANA*2000 and the pro-crystal prior [based on the final positions and the displacement parameters (anisotropic for C, N and O; isotropic for H) of the spherical refinement] was created with the module PRIOR of *BayMEM* (van Smaalen *et al.*, 2003).

3.2. MEM

All calculations were performed on a Compaq-DEC ES40 Workstation. The prior density file was computed with the module PRIOR; the electron-density map analysis based on Bader's AIM approach was performed with the module EDMA of *BayMEM* (van Smaalen *et al.*, 2003). The MEM calculations were performed with the latest version of *BayMEM* (van Smaalen *et al.*, 2003), using an adapted version of the commercially available *MEMsys5* algorithm package



Figure 2

Distribution of residuals $(|F^{obs}(\mathbf{H})| - |F^{MEM}(\mathbf{H})|)/\sigma$ for weights H0 through H5 and for (a) $\chi^2_{aim} = 1.0$ and (b) $\chi^2_{aim} = 0.425$. H0 means no ad hoc weighting. The Gaussian curve is shown in red. The insets show a magnification of the outer regions. The number of reflections in intervals of 0.2 wide $\Delta F/\sigma$ are given.

research papers



Figure 3

Difference maps $[\rho^{\text{MEM}} - \rho^{\text{prior}}; \text{ images } (a), (c), (e)]$ and residual maps [inverse Fourier transformation of $F^{\text{obs}} - F^{\text{MEM}}; \text{ images } (b), (d), (f)]$ of the peptide bond plane (N4-C3-O3a) for $\chi^2_{\text{aim}} = 1.0 (a), (b), \chi^2_{\text{aim}} = 0.425 (c), (d)$ and $\chi^2_{\text{aim}} = 0.2 (e), (f)$. Contour lines at 0.05 e Å⁻³, red dotted lines denote negative, blue lines denote positive values.

(Gull & Skilling, 1999). For the grid-based MEM the unit cell was divided into $216 \times 64 \times 324$ voxels, corresponding to voxel edge lengths of $0.085 \times 0.081 \times 0.077$ Å³. In order to minimize magnitudes of artefacts in ρ^{MEM} due to series-termination effects, the missing high-angle reflections were calculated based on the procrystal prior electron density in the

(FWHM 2.2 versus 3.0 for $\chi^2_{aim} = 1.0$). It is noteworthy that a smaller χ^2_{aim} also reduces the impact of the weighting. Although there is still a significant difference between using no weights (H0) and, for example, weights H2, the differences in the residual distributions become minute for H3, H4 and H5. In accordance with coincident histograms the corre-

sin θ/λ region 0.9–2.5 Å⁻¹, as suggested by Palatinus & van Smaalen (2005).

3.2.1. Choice of parameters: weighting. As already mentioned, the unmodified MEM has a tendency to dramatically under-fit some strong low-angle reflections, whereas a number of the remaining reflections are overfitted to satisfy the requirement $\chi^2/N_{\rm Ref} = \chi^2_{\rm aim}$. In order to counterbalance this effect the ad hoc weighting scheme proposed by de Vries et al. (1994) was utilized. The usage of this weighting scheme results in a more Gaussian-like distribution of the residuals. Fig. 2 clearly shows that with increasing power of the weighting (from H0 to H5), the number and the magnitude of the corresponding deviation of the under-fitted reflections is reduced. Consequently, fewer reflections are over-fitted, leading to a flattening of the peak of the histogram. For our dataset it seems that higher Hn provide better results. Unfortunately, with the current algorithm in use, higher Hn also mean considerably longer computation times, so that a value higher than H5 cannot be calculated within a reasonable time.

If a smaller value for χ^2_{aim} (0.425) is used, the resulting deviations from the optimal Gaussian distribution are less pronounced. Of course, using a smaller χ^2_{aim} fits F^{MEM} closer to $F^{\rm obs}$, which means that the FWHM is smaller and consequently the Gaussian curve is higher than in the case where χ^2_{aim} = 1.0. Nevertheless, the observed outliers deviate less from zero $(\Delta F/\sigma = -3.4 \text{ to } 4.8 \text{ versus } -8 \text{ to}$ 9.4 for $\chi^2_{aim} = 1.0$, H4) than would be expected solely from the reduced width of the histogram sponding electron densities are in perfect agreement (*e.g.* deviations of the electron density ρ^{MEM} in the BCPs less than 1%). Therefore, there is no evidence to indicate that one of these weights should be preferred above the others. We have chosen H4 since this is also the value which is recommended by de Vries *et al.* (1994).

Close inspection of the tails of the Gaussian curve (magnified region in the insets of Fig. 2) shows that there is still some bias in the result. This can be seen in the slightly asymmetrical distribution of residuals where the negative (left) part of the curve declines more steeply than the positive part. This indicates that although the *ad hoc* weighting scheme H4 successfully reduces the problem with over- and underfitted structure factors, it cannot completely suppress the deviations from the true Gaussian shape of the final distribution of residuals.

3.2.2. Choice of parameters: χ^2_{aim} . The historical MEM uses $\chi^2 = N_{\text{Ref}} (\chi^2_{\text{aim}} = 1.0)$ as the stopping criterion. This is in agreement with classic χ^2 refinements. However, the quality of the resulting charge-density map is poor. Firstly, the electron densities in the covalent BCPs are considerably lower than expected from the theoretical calculations or multipole refinements (Table 2). Secondly, small features, such as the lone pairs of O atoms, are not reproduced at all. The origin of this failure is found in the difference-Fourier maps.¹ These maps (Fig. 3b) clearly demonstrate that a significant amount of electron density is not fitted by the MEM calculation. This means that the MEM algorithm stops too early, before the optimal electron density is reached. Therefore, the stopping criterion has to be modified. However, since we could not find a reasonable theoretical way to predict the best value for χ^2_{aim} in advance, we decided to pursue a more empirical approach. Several MEM calculations with different values for χ^2_{aim} (0.2– 1.0) were performed and analyzed. Fig. 4 shows a plot of the average electron density in the BCPs of all non-hydrogen containing covalent bonds *versus* the used χ^2_{aim} . It can be seen that lowering the value of χ^2_{aim} increases the electron density in the BCP until all the significant electron density is represented in the MEM densities. At this point, the residual map shows only statistical noise. A further reduction of the value of χ^2_{aim} (below 0.375) only forces the calculation to include more noise to achieve a better fit to the data, which results in distorted electron-density maps (e.g. $\chi^2_{aim} = 0.2$, see Fig. 3e). The ultimate goal is to find exactly that value for χ^2_{aim} where significant features of the residual maps are suppressed below the noise level, whereas the corresponding electron-density map is not distorted. We have chosen $\chi^2_{aim} = 0.425$ as the optimal value for this system. At this value of χ^2_{aim} the average electron density in the BCPs is only 1% lower than in the case of $\chi^2_{aim} = 0.2$, but the features in the corresponding deformation maps are still nice and smooth (see Fig. 3c). However, the corresponding residual map is not entirely feature-free. Especially the areas close to the atom centres show deviations of up to \pm 0.2 e Å⁻³. Unfortunately, the statistical noise is of the same

magnitude. If the MEM is forced to include the remaining electron density in the fit (by reducing the value of χ^2_{aim}) this noise will also be included in the resulting electron density maps. For example, for $\chi^2_{aim} = 0.2$ (Figs. 3e and f) the residual map is free of features, whereas the corresponding difference map now contains many deformations which have their origin in the statistical noise. In a study on glycine (manuscript in preparation) we could show that a better dataset ($R_{int} = 0.015$ versus 0.030 in the present case) allows a closer fit to the data, minimizing significant features in the residual maps. However, it should not be forgotten that an absolute deviation in the electron density of up to ± 0.2 e Å⁻³, close to the position of the C, N or O atoms, is only a small relative deviation of a maximum of 2%, and even less (0.4%) directly at the atomic position.

We propose that the difference-Fourier map provides a good stopping criterion for the MEM. A value of χ^2_{aim} should be chosen, for which apparent features in the difference Fourier map are of similar magnitudes as the noise. Our analysis has shown that the corresponding electron density then is close to a limiting density corresponding to a fit to noise-free data.

4. Results and discussion

4.1. Atom charges and volumes

One of the known artefacts of the MEM is the possible presence of spurious local maxima in the electron-density maps (Jauch & Palmer, 1993; Jauch, 1994; de Vries *et al.*, 1996; Roversi *et al.*, 1998; Palatinus & van Smaalen, 2002). Close inspection of the present MEM electron density ($\chi^2_{aim} = 0.425$) shows that all maxima in the map are atom based, indicating that the strategies which were introduced to avoid these artefacts were indeed successful. The volume (V) and charge (Q) of each atomic basin were calculated on the basis of Bader's (1990) AIM theory. Since these quantities are addi-



Figure 4

Electron density in the BCPs averaged over all C–C, C–N and C–O covalent bonds for different MEM densities depending on χ^2_{aim} . The averaged electron density in the BCPs of ρ^{prior} is 1.55 e Å⁻³ (dotted line).

¹ In order to inspect the corresponding difference-Fourier maps, the *BayMEM* program was extended by the option to calculate residual maps.

Table 2

Averaged ρ (first line) and $\nabla^2 \rho$ values (second line) at the BCPs (in e Å⁻³ and e Å⁻⁵, respectively) for the different bond types in trialanine.

The final values which are discussed in comparison to the multipole/quantum chemical results (Rödel et al., 2006) were achieved with the smaller value of χ^2_{aim} (0.425).

	Prior (ISAM)	$\frac{\text{MEM}}{(\chi^2_{\text{aim}} = 1.0)}$	MEM $(\chi^2_{aim} = 0.425)$	Multipole	B3LYP/6-311 + + G(d,p) Calc.
C _{pentide} -O _{pentide}	2.11 (1)	2.33 (1)	2.55 (2)	2.87 (4)	2.65 (2)
peptide peptide	14 (5)	8 (7)	23 (9)	-29(3)	-10.2(1)
C _{peptide} -N _{peptide}	1.77 (1)	1.95 (3)	2.18 (5)	2.43 (3)	2.29 (1)
1 1 1 1 1 1	-1(2)	-8(2)	-17(4)	-23(1)	-23.7(2)
Long C _{carbox} -O	2.03 (1)	2.22 (1)	2.44 (1)	2.72 (6)	2.49 (1)
-	16 (3)	10 (4)	10 (5)	-27(4)	-12.1(2)
Short C _{carbox} -O	2.06 (1)	2.29(1)	2.47 (3)	2.82(1)	2.57 (1)
	15 (1)	13 (5)	23 (5)	-33.1(7)	-11.2(3)
$C_{\alpha} - N_{ammonium}$	1.39 (1)	1.50 (3)	1.67 (4)	1.76 (7)	1.59 (1)
	3.8 (2)	0 (1)	-8(1)	-11(4)	-12.2(3)
$C_{\alpha} - N_{peptide}$	1.45 (1)	1.58 (1)	1.74 (1)	1.76 (4)	1.69 (1)
1 1	2 (1)	-3(2)	-8(3)	-11(2)	-14.0(1)
$C_{\alpha} - C_{\beta}$	1.19 (1)	1.33 (2)	1.48 (2)	1.59 (2)	1.61 (1)
	0.1 (1)	-4.2(5)	-9 (1)	-9(1)	-12.8(1)
$C_{\alpha} - C_{peptide}$	1.19(1)	1.32 (3)	1.52 (5)	1.71 (5)	1.71 (1)
1 1	0 (1)	-4(2)	-7(3)	-11(2)	-14.6(2)
$C_{\alpha} - C_{carbox}$	1.18 (1)	1.35 (2)	1.53 (1)	1.78 (1)	1.69 (1)
	-0.8(1)	-5.35 (5)	-5.4(3)	-11.2(4)	-11.2(3)

tive, one expects that the sum of volumes in the unit cell equals the volume of the unit cell. Indeed, the sum of the volumes of the atomic basins in the unit cell is 2362.47 $Å^3$ and therefore



A significant difference between the MEM and the multipole method lies in the handling of thermal motion. The multipole method combines a sophisticated atom model with the classical anisotropic displacement parameters. Therefore, the electron density map can be regarded without the effects of thermal motion. Since the MEM approach is model-independent, MEM electron density maps always represent the distribution of the electron density

in the unit cell at the temperature of the measurement, *i.e.* static structure and the effects of thermal motion are not easily separated. Features are broader in MEM densities than they

(c) Atom A (d)Atom A Atom B Atom B esulting e.d resulting e.d

Figure 5

H1a - N1 - H1b sections of electron densities (contour lines at 0.1 e Å⁻³). (a) Prior electron density. (b) Model electron density similar to the prior, but with ADPs set to zero. Owing to the thermal motion, the individual electron-density peaks are broader and the resulting electron density does not show individual maxima for the H atoms (a), (c). Graphs (c) and (d) schematically illustrate this effect as one-dimensional sections.

are in static electron-density maps. We have found that one property of dynamic density maps is that H atoms do not necessarily give rise to local maxima. This effect is demonstrated by a comparison of the dynamic and static electron densities of trialanine, as they were computed from the ISAM (Fig. 5). The static density exhibits local maxima for all atoms. However, the dynamic density reveals H atoms only as a shoulder on the local maximum corresponding to the non-H atom to which they are covalently bonded (Fig. 5). Analysis of the model densities shows that failure to observe local maxima in the ρ^{MEM} for some H atoms is the result of thermal smearing and not a feature of the MEM. Since Bader's AIM analysis requires such a local maximum to calculate atomic basins with all their properties, the corresponding analysis cannot be carried out for most H atoms. The analysis of model densities shows that all the charge of the

Table 3 Comparison of averaged atomic charges (e) and volumes (\mathring{A}^3) .

Note that for $N_{peptide}$ and $N_{ammonium}$ the contribution of the H atoms could not be separated from that of the N atom. Therefore, in the case of MEM and prior, the total charge and volume for the whole fragment $(-NH_3, -NH-)$ is given (*in italics*).

	Prior		MEM		Multipole	
	Q	V	Q	V	Q	V
NH/N _{peptide}	-0.47 (2)	16.0 (2)	-0.5(1)	15 (1)	-1.03(3)	13.8 (5)
C _{peptide}	+0.69(9)	8.5 (9)	+1.29(5)	6.0 (4)	+1.1(4)	6.0 (1)
Opeptide	-0.49(9)	13.2 (4)	-1.2(1)	17 (1)	-1.13(3)	18 (1)
O _{long carboxy}	-0.52(7)	14 (1)	-1.1(1)	15.8 (2)	-1.02(4)	16.9 (3)
Oahort carboxy	-0.63(5)	16(1)	-0.91(5)	16.7 (4)	-1.00(4)	20.4 (2)
C _{carboxy}	+0.89(4)	8.1 (2)	+1.40(1)	6.1 (3)	+1.17(5)	6.1 (3)
$NH_3/\dot{N}_{ammonium}$	+0.01 (1)	25.9 (9)	+0.40 (5)	24 (Ì)	-1.2 (1)	15.5 (4)

affected H atoms will be added to the non-H atom to which it is covalently bonded. Exceptions are the carbon-bonded H atoms, which exhibited an electron-density maximum of their own for all C_{α} -H and almost half of the C_{β} -H atoms. The average charge for H atoms which are bonded to the α -C atoms is +0.22 (9)*e*, in the case of the C_{β} -bonded H atoms which can be analysed it averages +0.2 (1)*e*.

Failure to observe local maxima for some H atoms does not imply that these atoms are 'not found'. Their densities can be revealed by subtracting from the densities ρ^{prior} or ρ^{MEM} a model electron density that has been computed in a way similar to the prior, but with non-H atoms only.

Table 3 shows a comparison of atom charges and volumes between ISAM, MEM and multipole models. It is obvious that both multipole and MEM charges differ significantly from the charges calculated on the basis of the ISAM. In general, the charges calculated with the MEM are the same as the results of the multipole refinement within the standard deviation. Only the carbon of the carboxylate shows a difference from this trend, being significantly more positive [Q_{MEM} =

+1.40 (1)*e*, $Q_{\text{multipole}} = +1.17 (5)$ *e*]. It seems noteworthy that the charge of the carbon of the peptide bond averages almost 0.2*e* higher for the MEM than for the multipole refinement [$Q_{\text{MEM}} = 1.29 (5)$; $Q_{\text{multipole}} = 1.1 (4)$], but the variance is rather high in the latter case. Future investigations will be required to show whether this is a general difference between the multipole method and the MEM.

It is remarkable that in contrast to the multipole refinement, the MEM indicates a difference of 0.19*e* between the charges of the O atoms of the long [Q = 1.1 (1)e] and the short [Q = 0.91 (5)e]

C-O bond of the carboxylic group. This finding is in agreement with general chemical knowledge and this charge difference between the two carboxylic O atoms has already been observed for multipole refinements in several cases (*e.g.* Benabicha *et al.*, 2000; Pichon-Pesme *et al.*, 2000). It is interesting that this is not the case for the multipole refinement of trialanine where these two charges are practically identical, while the MEM recovers this difference for the same data.

Like the charges, the atomic volumes resulting from the MEM and multipole refinements are equal to each other within standard deviations. The only exception is the volume of the O atom which is bound by the shorter carboxylate bond $[V_{\text{MEM}} = 16.7 \text{ (4) } \text{Å}^3, V_{\text{multipole}} = 20.4 \text{ (2) } \text{Å}^3]$. Since the MEM volume is in perfect agreement with the average values from the literature $[V = 16.5 \text{ (9) } \text{Å}^3; \text{Dittrich}, 2002]$ we believe that the MEM results are closer to the true values in this case.

The atomic volumes of the MEM do not differ much from the corresponding ISAM values. However, there is a tendency for the atomic volumes, which were estimated by the MEM, to be slightly larger in the case of O atoms and slightly smaller in

the case of C atoms.



Figure 6

(a) MEM difference map ($\rho^{\text{MEM}} - \rho^{\text{prior}}$) and (b) multipole deformation map (Rödel, 2003) of the peptide bond plane (O6*a*-C6-N7). Contour lines at 0.05 e Å⁻³ are for MEM and 0.10 e Å⁻³ for multipole maps; red dotted lines denote negative values, blue lines denote positive values.

4.2. Covalent bonds

Although MEM difference maps $(\rho^{\text{MEM}} - \rho^{\text{prior}})$ do not represent the same quantity as the static deformation maps from the multipole method, both maps visualize the differences in electron density between the MEM or multipole densities and the density based on the ISAM. In Fig. 6 a comparison of these maps is shown for the O6-C6-N7 peptide bond. Features in the multipole deformation maps are more smooth than features in MEM difference maps. Smooth features are inherent to the multipole method, since this method uses smooth functions (multipoles) for modelling the electron densities, whereas the MEM refines electron densities on a grid. Nevertheless, both maps show the same features. They exhibit an accumulation of electron density in the areas between the atoms, indicating the formation of a covalent bond. Furthermore, the two lone pairs of O atoms are clearly visible in both maps.

Bader's (1990) AIM theory provides an excellent tool to study and compare the topology of electron-density maps on a quantitative basis. According to his theory the character of a bond is determined by the values of the electron density and the values of the second derivatives of the electron density in the BCP, which can be described approximately as the 'saddle point' of the electron density between two atoms. The corresponding average values in the BCPs for all non-hydrogen containing covalent bonds of the same type have been calculated and are listed in Table 2. The positions of the BCPs are shown in Fig. 1. The coincidence between the ρ_{bcp} derived from MEM densities and multipole densities is not as good as it is for atom charges and atomic volumes. In most cases the electron densities in the BCPs are lower for the MEM than for the multipole method; the average MEM density for a particular bond type is 84-99% of its corresponding multipole counterpart. It is interesting to compare the MEM and multipole results with the results based on quantum chemical calculations, especially in the case of the heterogeneous bonds. Here the MEM values are a maximum of 5% smaller than the theoretical values, whereas at the same time the multipole values are up to 10% higher. These topological discrepancies between multipole and theoretical charge densities, particularly at the BCP of polar bonds, have been previously observed (e.g. Bach et al., 2001; Volkov et al., 2000; Gatti et al., 1992). According to Volkov et al. (2000) the main origin for this lies in the nature of the radial functions of the multipole model. However, in the case of the homogenous, non-polarized carbon-carbon bonds, the multipole values match the theoretical electron densities to within 5%, whereas the MEM values are up to 13% smaller. Since the MEM just fits electron densities without any knowledge of atom types, it is hard to believe that this different behaviour for homogenous versus heterogeneous bonds is actually inherent to the MEM. It could be possible that the quantum chemical calculations are less reliable for the carbon-carbon bond. This is also indicated by a HF calculation which was performed by Rödel et al. (2006) for the trialanine. Here the computed densities for the HF and the density functional theory calculations are exactly the same (within the standard uncertainties), except in the case of carbon-carbon bonds where the electron densities differ by 0.1 e $Å^{-3}$ in the BCPs.

These considerations aside, the MEM results are still slightly too low with respect to the quantum chemical results. There are two explanations for this behaviour. Firstly, the electron density in the BCP could only be fitted to 99% of the possible value (see Fig. 4) because 100% would have required such a low value for χ^2_{aim} that a considerable amount of noise would have been incorporated in the resulting electron-density maps (as in Fig. 3*e*). Secondly, the MEM densities

incorporate the effect of thermal motion. Since the position of the BCP marks two maxima and only one minima of the local electron density, thermal motion will reduce the electron density in the BCP in comparison to the thermal-motion-free calculational/multipole method. The magnitude of this effect can only be roughly estimated. An analysis of an electron-density map generated on the basis of the ISAM with the coordinates of the trialanine atoms from the ISAM refinement and by 90% reduced experimental ADPs (atomic displacement parameters) revealed that the ρ values in the BCPs which are less affected by thermal motion are up to 3% higher than in the case where unmodified experimental ADPs have been used.

More important than an exact replication of the literature values which were generated by different methods (multipole, quantum chemical calculations) is the fact that independent of the actual values the general trend of the electron density of the different bond types is the same for the MEM, the multipole method and the theoretical calculations $\left[\rho(C-)\right]$ $O_{peptide}) > \rho(C - O_{long carboxy}) > \rho(C - O_{short carboxy}) >$ $\rho(C_{peptide} - N_{peptide}) > \rho(C_{\alpha} - N_{peptide}) \ge \rho(C_{\alpha} - N_{ammonium}) >$ $\rho(C_{\alpha} - C_{carboxy}) \simeq \rho(C_{\alpha} - C_{peptide}) > \rho(C_{\alpha} - C_{\beta})]$. Furthermore, the reproducibility of the electron densities in the BCP of a specific bond is very high, as can be seen by the small variance of the corresponding average values (MEM: maximum deviation: ± 0.05 and average variance ± 0.03 e Å⁻³; multipole: maximum deviation ± 0.07 and average variance: ± 0.04 e Å⁻³). In this respect it is interesting to note that the variance of ± 0.05 e Å⁻³ for the average C_{α}-C_{peptide} bond results from a significantly higher electron density in the BCP of the atom pair adjacent to NH_3^+ [1.55 (2) e Å⁻³] in comparison to the corresponding value for the $C_{\alpha}-C_{peptide}$ bond $[1.48 (2) e Å^{-3}]$ in the central amino acid. Exactly the same trend is found by the multipole method ($N_{ammonium}$) $C_{\alpha} - C_{\text{peptide}} = 1.75 \text{ e} \text{ Å}^{-3}; N_{\text{peptide}} - C_{\alpha} - C_{\text{peptide}} = 1.68 \text{ e} \text{ Å}^{-3}).$ For the other bonds the influence of the next nearest neighbour is less distinct.

According to Bader's AIM theory (Bader, 2000), the Laplacian (sum of the eigenvalues of the Hessian matrix) in the BCP gives valuable information about the type of bonding interaction (open or closed shell). When the Laplacians are compared, one has to keep in mind that the reproducibility of these values is lower than in the case of the densities, even for the same kind of bond. In general, the average values for MEM Laplacians in the BCPs are slightly more positive than their multipole counterparts, but due to their high standard deviations this is not discussed further. However, the Laplacians of the C–O bonds are explicitly positive for the MEM and strongly negative for the multipole method. Benabicha's (2000) comparison of the topological parameters of a number of C-O bonds which were investigated by the multipole method shows that although the variance in the Laplacians is rather high, all the Laplacians are explicitly negative. The reason for this difference of the MEM and the multipole method lies within the fact that thermal motion effects are included in the MEM and excluded in the multipole method. This is proven by a comparison of a procrystal electron density

Table 4

 ρ and $\nabla^2 \rho$ values at the BCPs (in e Å⁻³ and e Å⁻⁵, respectively) for the different hydrogen bonds in trialanine in comparison to the multipole results (Rödel *et al.*, 2006).

The corresponding values c	calculated on	the basis of the	ISAM are also given	ί.
----------------------------	---------------	------------------	---------------------	----

	Prior (ISAM)		MEM		Multipole	
	ρ	$ abla^2 ho$	ρ	$ abla^2 ho$	ρ	$ abla^2 ho$
N1-H1 a ···O19 b	0.39	1.75	0.43	0.77	0.28	4.60
$N1-H1b\cdots O21$	0.22	2.19	0.21	1.88	0.13	2.22
$N1-H1c\cdots O19a$	0.35	2.72	0.33	1.78	0.25	4.05
N4-H4 a ···O16 a	0.24	2.02	0.21	2.45	0.15	2.94
N7-H7a···O13a	0.21	2.26	0.20	1.43	0.13	2.42
$N11 - H11c \cdot \cdot \cdot O31$	0.34	2.52	0.37	1.41	0.20	4.02
N11 $-$ H11 b ···O9 a	0.22	2.39	0.19	1.60	0.13	2.52
N11 $-$ H11 a ···O9 b	0.39	1.40	0.45	-1.57	0.28	4.97
N14 $-$ H14 a ···O6 a	0.23	2.02	0.19	2.19	0.14	2.89
N17−H17a···O3a	0.17	1.92	0.15	2.11	0.09	1.86
$O21 - H21a \cdot \cdot \cdot O19a$	0.22	2.11	0.27	1.54	0.20	4.02
O31−H31a···O9a	0.27	2.44	0.33	1.58	0.24	4.94

calculated with experimental coordinates and ADPs with a corresponding procyrstal electron density where the ADPs have been set to a tenth of their experimental value. The surprising result is that the average Laplacians of the C–O bonds are dramatically more negative (14–39 e Å⁻⁵) in the latter case, whereas the Laplacians of the other bonds are mainly unaffected (maximum change ± 2 e Å⁻⁵). It can be assumed that the topology of the MEM electron density map is also prone to this effect. This effect, in combination with the fact that the multipole-derived Laplacians are usually too negative compared with the theoretical values owing to the nature of the radial functions of the multipole model (Volkov *et al.*, 2000), easily explains the observed discrepancies between the multipole- and MEM-derived Laplacians of the C–O bonds.



Figure 7

(a) MEM difference map $(\rho^{\text{MEM}} - \rho^{\text{prior}})$ and (b) multipole deformation map (Rödel *et al.*, 2006) of the hydrogen-bond plane (H11*c*-O9*b*-C9). Contour lines at 0.05 e Å⁻³ are for MEM and 0.10 e Å⁻³ for multipole maps; red dotted lines denote negative, blue lines denote positive values.

4.3. Hydrogen bonds

Rödel et al. (2006) found 12 N- $H \cdots O$ and $O - H \cdots O$ hydrogen bonds in the asymmetric unit (see Fig. 1). Their topological properties were also computed with the MEM and the values are compared with the multipole results in Table 4. In the case of the charge densities, the general trend is the same for the MEM and the multipole method: long hydrogen bond distances exhibit very small charge densities in the BCPs and the charge density increases when the distance decreases. However, in contrast to the BCPs of covalent bonds, the electron densities are on average 49% higher for MEM densities than for the multipole method. Five of the

strongest hydrogen bonds (interactions involving the charged carboxylate and ammonium groups) even exhibit a charge density which is larger than in the case of the ISAM. This is remarkable because in the case of the multipole method all BCPs of hydrogen bonds are smaller than the corresponding ISAM values. The reason for this feature is that the single electron of the H atom is mostly located in the covalent bond. Therefore, the probability of finding it on the opposite site of the bond is reduced in comparison to the ISAM which assumes an even distribution of the electron density around the nucleus. The higher value of the electron density in the case of the MEM therefore indicates how exceptionally strong the corresponding hydrogen-bonding interactions are. So why do the MEM and the multipole results differ so significantly in this respect? The difference-/deformation-density maps of the

strongest hydrogen bond (Fig. 7) show two significant differences. First, the MEM shows a clear increase in electron density between the H and O atoms, whereas the multipole method exhibits a decrease in electron density at the same position. Secondly, the polarization of the O atom is much more directed and significantly narrower for the MEM than for the multipole method. This might also be the reason for the observed different behaviour. The applied multipoles for the O atoms cannot be used to satisfactorily model such a fine and distinct interaction over a longer distance, as is the case in hydrogen bonding. This also explains the difference in the Laplacians (the average values for the multipole method are more than two times higher than for the MEM), which are known to depend on the radial functions of the multipole model (Volkov *et al.*, 2000).

In general, Bader's (1990) AIM theory postulates that covalent interactions are characterized by high electron densities and negative Laplacians in the BCPs, whereas closed-shell (ionic) interactions are characterized by small electron densities and positive Laplacians. Consequently, hydrogen bonds are described as closed-shell interactions. However, a plot of the Laplacians *versus* the electron density for the BCPs of the hydrogen bonds (Fig. 8) reveals some interesting trends. In the case of the multipole method the Laplacians increase with electron density, whereas in the case of the MEM the Laplacians decrease with increasing electron density. In the case of the bond with the highest electron density, even a negative Laplacian was observed. Owing to the discussed insufficiencies of the multipole method in the modelling of hydrogen bonds, we believe that the trend observed in the MEM density is closer to reality. These results suggest that with increasing strength of the hydrogen bonds, their ionic nature is more and more mixed with covalent interactions. A close analysis reveals that the most affected hydrogen bonds are between the oppositely charged NH_3^+ and COO^{-} groups, either *via* a direct hydrogen bond (N-H···O) or via a water molecule $(N-H\cdots O-H\cdots O)$. A second explanation for the high electron densities in the hydrogen bonds in MEM maps is provided by the possibility that in a tiny fraction of these groups, the proton is not located near the N atom, but near the O atom (N \cdots H-O; N \cdots H-O \cdots H-O), with a very small probability, leading to a different electron-density distribution. Owing to the averaging character of the X-ray diffraction experiment, the resulting electrondensity distribution would be a mixture of both states according to their frequency of occurrence. This could very well simulate a more covalent type of bond with an increased electron density on the bond path in comparison to the ISAM.



Figure 8

Correlation between the value of the Laplacian and electron density in BCPs of hydrogen bonds. In the case of the ISAM (triangles) there is no significant slope, whereas in the case of the multipole method (circles) the slope is clearly positive, and in the case of the MEM (squares) it is significantly negative.

The computer program BayMEM was extended by the option to calculate difference-Fourier maps. Analysis of these maps revealed that the electron-density maps which are generated with the historical MEM ($\chi^2_{aim} = 1.0$) do not model the actual electron-density distribution sufficiently. A closer fit to the data (smaller χ^2_{aim}) gives better results, but if the value of χ^2_{aim} chosen is too low, noise becomes incorporated in the maps. The ideal value of χ^2_{aim} , where the fit to the data is optimal without introducing significant amounts of noise, has to be established empirically for each system by close analysis of the corresponding difference and residual maps. The best value for χ^2_{aim} was found to be 0.425 in the present case. A useful property of this empirical method is that the final MEM density will become independent from possible errors in the scale of the measured standard uncertainties of the diffracted intensities.

We could show that at optimal χ^2_{aim} the MEM and the multipole method are on a par regarding the reproduction of atomic charges, volumes, general topological features and the trends in the charge density in the BCPs. Quantum chemical calculations, multipole method and MEM give the same charge densities in the BCPs within $\pm 8\%$ of the corresponding average. Whether the agreement between the multipole method and the theory is better depends on the type of bond. In the case of the Laplacians, the agreement is not as good and the Laplacians of the C–O bonds differ especially strongly. The peculiarities of both methods are responsible for these differences.

In general, the distinct and fine features of hydrogen bonds are more convincingly reproduced by the model-independent MEM since the restrictions on radial functions limit the reproduction of such features in the case of the multipole method. Consequently, the hydrogen bonds are more pronounced in the MEM determination and the corresponding topological parameters differ from the multipole derived values.

The disadvantage of the model-independent nature of the MEM is that the best achievable charge-density map depends strongly on the completeness and accuracy of the data. An extensive data set of high-quality data collected with the crystal cooled to a very low temperature is therefore necessary for an accurate charge-density study with the MEM. The MEM cannot repair systematic errors in the data, and MEM electron densities will be adversily affected if such errors are present.

The present results for trialanine show that electron densities obtained with the MEM are not better than those obtained with multipole refinements. However, an improvement upon multipole refinements has not been the goal of our study. The MEM electron density is shown to be of comparable quality to the multipole density for trialanine, thus indicating that the MEM can be an alternative for multipole refinements in accurate charge-density studies. The MEM has the potential to improve upon multipole refinements in cases where the latter have known problems. Specifically this applies to cases where some kind of disorder can be present, like the positions of the H atoms in hydrogen bonds and to compounds with very large unit cells, for which multipole parameters are severely correlated, while the MEM does not suffer from correlated parameters by principle. The application of the MEM to diffraction data of protein crystals is part of our future research program. Further improvements to the MEM would be the use of a prior based on a multipole model with multipole parameters obtained from a database of transferable multipole parameters or from quantum chemical calculations on small model compounds (Pichon-Pesme *et al.*, 2004; Dittrich *et al.*, 2005).

We are indebted to P. Luger and E. Rödel for making the Xray diffraction data available to us. Financial support was obtained from the German Science Foundation (DFG) within the framework of SPP1178.

References

- Bach, A., Lentz, D. & Luger, P. (2001). J. Phys. Chem. A, 105, 7405–7412.
- Bader, R. F. W. (1990). *Atoms in Molecules: a Quantum Theory*. Oxford Science Publications. Oxford: Clarendon Press.
- Benabicha, F., Pichon-Pesme, V., Jelsch, C., Lecomte, C. & Khmou, A. (2000). Acta Cryst. B56, 155–165.
- Buck, B. & Macaulay, V. A. (1994). Editors. *Maximum Entropy in Action*. Oxford University Press.
- Dittrich, B. (2002). PhD thesis, Freie Universität Berlin.
- Dittrich, B., Hubschle, C. B., Messerschmidt, M., Kalinowksi, R., Girnt, D. & Luger, P. (2005). *Acta Cryst.* A61, 314–320.
- Gatti, C., Bianchi, R., Destro, R. & Merati, F. (1992). J. Mol. Struct. (Theochem.) 255, 409–433.
- Gilmore, C. J. (1996). Acta Cryst. A52, 561-589.
- Gull, S. F. & Skilling, J. (1999). Quantified Maximum Entropy. MemSys5. User's Manual. Bury St Edmunds, Suffolk, UK.
- Hansen, N. K. & Coppens, P. (1978). Acta Cryst. A34, 909-921.

- Iversen, B. B., Jensen, J. L. & Danielsen, J. (1997). Acta Cryst. A53, 376–387.
- Jauch, W. (1994). Acta Cryst. A50, 650-652.
- Jauch, W. & Palmer, A. (1993). Acta Cryst. A49, 590-591.
- Jaynes, E. T. (1957). Phys. Rev. 106, 620-630.
- Jaynes, E. T. (1979). In The Maximum Entropy Formalism, edited by R. D. Levine & M. Tribus, pp. 15–118. Cambridge: MIT Press.
- Jaynes, E. T. (1986). Maximum Entropy and Bayesian Methods in Applied Statistics, edited by J. H. Justice, pp. 26–58. Cambridge University Press.
- Palatinus, L. & van Smaalen, S. (2002). Acta Cryst. A58, 559-567.
- Palatinus, L. & van Smaalen, S. (2005). Acta Cryst. A61, 363–372.
- Petricek, V., Dusek, M. & Palatinus, L.(2000). *JANA*2000. Institute of Physics, Praha, Czech Republic.
- Pichon-Pesme, V., Jelsch, C., Guillot, B. & Lecomte, C. (2004). Acta Cryst. A60, 204–208.
- Pichon-Pesme, V., Lachekar, H., Souhassou, M. & Lecomte, C. (2000). Acta Cryst. B56, 728–737.
- Rödel, E. (2003). Experimentelle Untersuchung der Ladungsdichte des Tripeptids L-Alanyl-L-alanyl-L-alanin mit Hilfe von hochaufgelöster Röntgenbeugung bei 20 Kelvin. Diplomarbeit, Freie Universität Berlin.
- Rödel, E., Messerschmidt, M., Dittrich, B. & Luger, P. (2006). Org. Biomol. Chem. 4, 475–481.
- Roversi, P., Irwin, J. J. & Bricogne, G. (1998). Acta Cryst. A54, 971– 996.
- Roversi, P., Irwin, J. J. & Bricogne, G. (2002). Electron, Spin and Momentum Densities and Chemical Reactivity, edited by P. G. Mezey & B. E. Roberston. Dordrecht: Kluwer Academic Publishers.
- Smaalen, S. van, Palatinus, L. & Schneider, M. (2003). Acta Cryst. A**59**, 459–469.
- Steiner, T. & Saenger, W. (1993). Acta Cryst. A49, 379-384.
- Vries, R. Y. de, Briels, W. J. & Feil, D. (1994). Acta Cryst. A50, 383– 391.
- Vries, R. Y. de, Briels, W. J. & Feil, D. (1996). Phys. Rev. Lett. 77, 1719–1722.
- Volkov, A., Abramov, Y., Coppens, P. & Gatti, C. (2000). Acta Cryst. A56, 332–339.