Accurate charge densities of amino acids and peptides by the Maximum Entropy Method

Von der Universität Bayreuth
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

von

Jeanette Netzel

aus Bergen

1. Gutachter: Prof. dr. S. van Smaalen
2. Gutachter: Prof. Dr. P. Luger

Tag der Einreichung: 15. August 2008
Tag des Kolloquiums: 15. Januar 2009

Contents

Publications

Parts of the present thesis have been published in international scientific literature or have been submitted for publication:

Chapter 3:
A. Hofmann, J. Netzel and S. van Smaalen.
Accurate charge density of Tri-alanine: A comparison of the multipole formalism and the maximum entropy method (MEM).
Acta Crystallogr. B, **63**, 285–295 (2007).

Chapter 4:
J. Netzel, A. Hofmann and S. van Smaalen.
Accurate charge density of α-glycine by the maximum entropy method.
CrystEngComm, **10**, 335–343 (2008).

Chapter 5:
J. Netzel and S. van Smaalen.
Joined analysis of topological properties of hydrogen bonds and covalent bonds from accurate charge density studies by the maximum entropy method.
Submitted to *CrystEngComm* (2008).

Chapter 1

Introduction

The ubiquity of the crystalline state makes crystallography an interdisciplinary science of importance for material science, synthetic chemistry and biology. An understanding of the properties of a crystal can only be developed, if its structure, that is the spatial arrangement of the atoms, is known. X-ray diffraction is the method of choice for the determination of crystal structures.

A simple description of the crystal structure is provided by the independent spherical atom model (ISAM), which is obtained by refinement against X-ray diffraction data. The ISAM describes the positions of the atoms with their spherical electron densities in the unit cell and the anisotropic atomic displacements of these atoms (atomic displacement parameters) about their positions due to thermal motion. However, the reorganization of valence electrons due to chemical bonding in molecules is not considered within the ISAM. Thus, it does not describe the true electron density with respect to the experimental data.

The multipole model allows to recover bonding effects on the density by refinement of additional parameters against the diffraction data.[1,2] Besides the coordinates and atomic displacement parameters (ADP) employed for the ISAM, the multipole model additionally employs multipolar expansions of the atomic electron density. These expansions constitute a spherical core, a spherical valence electron density and an aspherical valence density, and can be refined with respect to their population coefficients, radial functions and parameters for expansion or contraction of the radial functions. Thus, the aspherical-atom density obtained by the multipole model deviates from the density based on ISAM.

However, the summation of all refinable multipole parameters of one atom yields

a large number of parameters to be refined for the total molecule. This effect becomes more severe for increasing size of the unit cell and leads to correlated parameters in most cases. The increasing number of correlated multipole parameters is usually counteracted by introduction of constraints on the parameters or to refine only selected parameters that are considered physically important. Thus, the multipole model imposes restrictions on the density, thus leading to artifacts in the density or to models describing the density incompletely. Another problem may arise from the employment of radial functions, because they may be inflexible[3,4,5,6,7] at distances remote from the respective nuclei and restrict the distribution of density within the applied multipole parameters.

In contrast to the multipole model, the Maximum Entropy Method (MEM) provides a model-independent electron density and it does not suffer from correlated parameters by its very principle. The MEM reconstructs the density on a grid over the unit cell and provides a stable solution, even if incomplete data sets of integrated diffraction intensities are employed as experimental data. However, the electron densities reconstructed with the MEM may contain artifacts[8,9,10,11,12] of different nature than artifacts in the multipole densities. Several extensions to the MEM have overcome such deficiencies and established this approach as serious alternative for the multipole method. While the multipole model provides static electron densities, the MEM produces dynamic densities.

For the purpose of accurate charge-density studies, the MEM requires data sets of high quality, that are datasets containing all reflections up to a high resolution of $(\sin(\theta)/\lambda)_{max} > 1.00$ Å^{-1}, and data measured at low temperatures of about 20 K, to reduce thermal motion which is included in dynamic MEM densities.

The objective of the present thesis is Accurate Charge Density studies of biological molecules, *i.e.* amino acids and tripeptides, by the MEM and to provide a description of the reconstruction of these densities. For that purpose, data of several compounds, of which the reconstruction of accurate charge densities is described, were obtained from the literature. The employment of such high-quality data allows a comparison of the MEM densities with the densities from the literature obtained from the multipole method on the same data. It will be shown that the MEM, along with its extensions, describes the electron density in a more realistic way than the multipole method does and that the MEM allows a good characterization of chemical bonding. Because the MEM does not suffer from dependent parameters,

it is the expectation that this method will be more easily applied to diffraction data of protein crystals than multipole refinements.

Applications and the principle of the MEM are described in Chapter 2. The procedure of the determination of the charge density along with extensions to the MEM to enhance the quality of the MEM densities are given. Eventually, the Quantum Theory of Atoms in Molecules,[13] on which the analysis of the MEM densities is based, is described.

Chapter 3 reports on the computational details of the application of the MEM on the tripeptide L-alanyl–L-alanyl–L-alanine (trialanine) and presents the results of the analysis of the MEM density and its bond critical points with respect to the Atoms in Molecules (AIM) theory,[13] in comparison with the corresponding results of the multipole refinement and from quantum chemical calculations.[14] Similarities and differences between the densities obtained by the MEM and multipole refinements are discussed.

An application of the MEM to the amino acid α-glycine is discussed in Chapter 4. It gives a description of the computational details and presents the resulting MEM densities in comparison with the densities obtained from the multipole refinement on the same data set, along with the AIM analysis of the MEM density and its bond critical points. The centrosymmetric structure of α-glycine allowed a comparison of the phases of the reflections with the corresponding phases obtained from the multipole model. The result of that comparison is discussed.

Chapter 5 focusses on the study of hydrogen bonds from several amino acids and tripeptides, and reports on the results of the comparison of the electron densities from the MEM and the multipole model. Via the AIM[13] analysis, the topological and energetic properties at bond critical points of the hydrogen bonds and of covalent bonds are obtained and systematic dependencies on the distances between atoms involved are reported. The contribution of the prior density to the properties at bond critical points is discussed.

A summary and conclusions are given in Chapter 6, which is present as German equivalent in Chapter 7.

Appendices provide supplementary information to the studied compounds regarding perspective views of the structures, details to the MEM calculations and the analysis of the MEM densities such as the values of properties at the bond critical points.

Chapter 2

The Maximum Entropy Method in crystallography

2.1 Applications of the MEM

The MEM is a general method for data analysis, which is employed to extract the maximum amount of information from the data, without the introduction of artifacts or assumptions concerning a model.[15,16]

In X-ray diffraction studies, the MEM can be used for deconvolution of powder diffraction data,[17] to extract phases from intensities of Bragg reflections,[18,19] or to determine the most probable electron density distribution in the unit cell.[20,21,17] The reconstruction of the density in the unit cell can be employed to address two kinds of problems. One application is the determination of atomic positions in the unit cell. A prominent example is the determination of the position of the yttrium atom within the carbon cage of the metallofullerene $Y@C_{82}$.[22] The other application is the reconstruction of so-called Accurate Charge Densities in the unit cell. In view of Chapters 3, 4 and 5, it is mentioned, that Accurate Charge Densities by the MEM are topic of the present thesis. Charge densities of several amino acids and tripeptides were reconstructed by the MEM and analyzed to study chemical bonds.

Accurate Charge Density studies require accurate X-ray diffraction data and the lowest possible temperature for data collection.[23] Despite the latter requirement for data collection, some groups have reported Accurate Charge Density studies of MoO_3,[24] $PbTiO_3$[25] and $BaTiO_3$[25] by the combination of Rietveld refinement of

powder diffraction data, collected at high temperatures up to 800K, and the MEM. They claimed to yield information about the nature of the bonding by inspection of the MEM electron densities and the ascertainment of the existence of overlapping densities between atoms. However, a comparison between ISAM and MEM densities has not been made in those studies–as I will present in this thesis–and it can be questioned that high-temperature studies will have obtained information beyond the ISAM.

In other publications, it has been claimed to identify hydrogen bonds by inspection of charge density maps produced from data that has been collected at temperatures higher than 90 K; overlapping densities between the atoms have been used to ascertain the existence of these relatively weak bonds.[26,27]

Besides the inspection of charge density maps to characterize the nature of bonds, other groups have claimed to yield insight into the strength of bonds by determination of the density at the middle of the bond between two atoms.[28,29] In view of Section 2.2.3, it is mentioned that the sophisticated analysis of charge densities according to the Atoms in Molecules theory[13] provides a possibility to interpret electron densities quantitatively, to identify and characterize chemical bonds.

In order to extract information about the electron configuration from charge density distribution maps, the MEM has been used. It has been claimed to extract 4f electrons of cerium phosphide[30] or to identify the spin configuration of the pyroborate $Mn_2B_2O_5$.[31]

The MEM can be employed to reveal disorder such as partially occupied atomic sites, as in the case of partially occupied zinc atoms in Zn_4Sb_3,[32] or multiple orientations of molecules or functional groups. Orientational disorders of the AlO_4 tetrahedra in strontium europium aluminate, $Sr_{0.864}Eu_{0.136}Al_2O_4$,[33] or of pentamethyl-cyclopentadienyl-lithium[34] were found. Multiple conformations of the cage structure of a nano-porous lime-alumina compound, $12CaO{\cdot}7Al_2O_3$, were revealed by the MEM.[35]

Aperiodic crystals can also be studied with the aid of the MEM. Modulation functions have been extracted from the (3+d)-dimensional superspace MEM densities of the misfit layer compound $(LaS)_{1.14}NbS_2$, the high-pressure phase III of bismuth and ammonium tetrafluoroberyllate $(NH_4)_2BeF_4$.[36,37] For the latter a charge density study has been performed to investigate the deformation of the density of that compound due to chemical bonding;[36,37] modulations of the distances between

hydrogen atoms and acceptor atoms of hydrogen bonds have been determined.[37]

In neutron diffraction studies, the MEM can be applied to determine the scattering length density distribution.[38,39] Three-dimensional images of the scattering lengths densities created by the MEM allow the analysis of atomic distribution. Nuclear densities can be obtained by the MEM from neutron diffraction and can reveal disorders as well.[40]

2.2 Accurate charge densities by the MEM

2.2.1 Principle of the MEM - BayMEM

For the reconstruction of the accurate charge density by the MEM, a grid of $N_p = N_1 \times N_2 \times N_3$ points over the unit cell is defined. The density $\rho_k = \rho(\mathbf{x}_k)$ is discretized on this grid and described by its values at the grid points $\mathbf{x}_k$ where $k = 1, \ldots, N_p$.

The MEM is based on the principle that the most probable density $\{\rho_k\}$ is that one, which simultaneously maximizes the informational entropy S, fulfills the constraint of the normalization C_0 of the density and fits the diffraction data.[15,41] The informational entropy S of $\{\rho_k\}$ is defined as

$$S = -\sum_{k=1}^{N_p} \rho_k \log\left(\frac{\rho_k}{\rho_k^{prior}}\right) \tag{2.1}$$

where $\rho_k^{prior} = \rho^{prior}(\mathbf{x}_k)$ is used as reference density or prior density [see example in Figure 2.1(a)] in the MEM. The maximum of the entropy $S = 0$ is fulfilled for $\rho_k = \rho_k^{prior}$. Without incorporation of further information such as diffraction data, the prior density is obtained as the most probable density.

The constraint of the normalization of the density is given by

$$C_0 = \frac{V_{cell}}{N_p} \sum_{k=1}^{N_p} (\rho_k) - N_e \tag{2.2}$$

where V_{cell} is the volume of the unit cell and N_e is the number of electrons in the unit cell. Under the condition that the total density which is dicretized on the grid over the unit cell equals the number of electrons in the unit cell, the normalization constraint is fulfilled, when $C_0 = 0$.

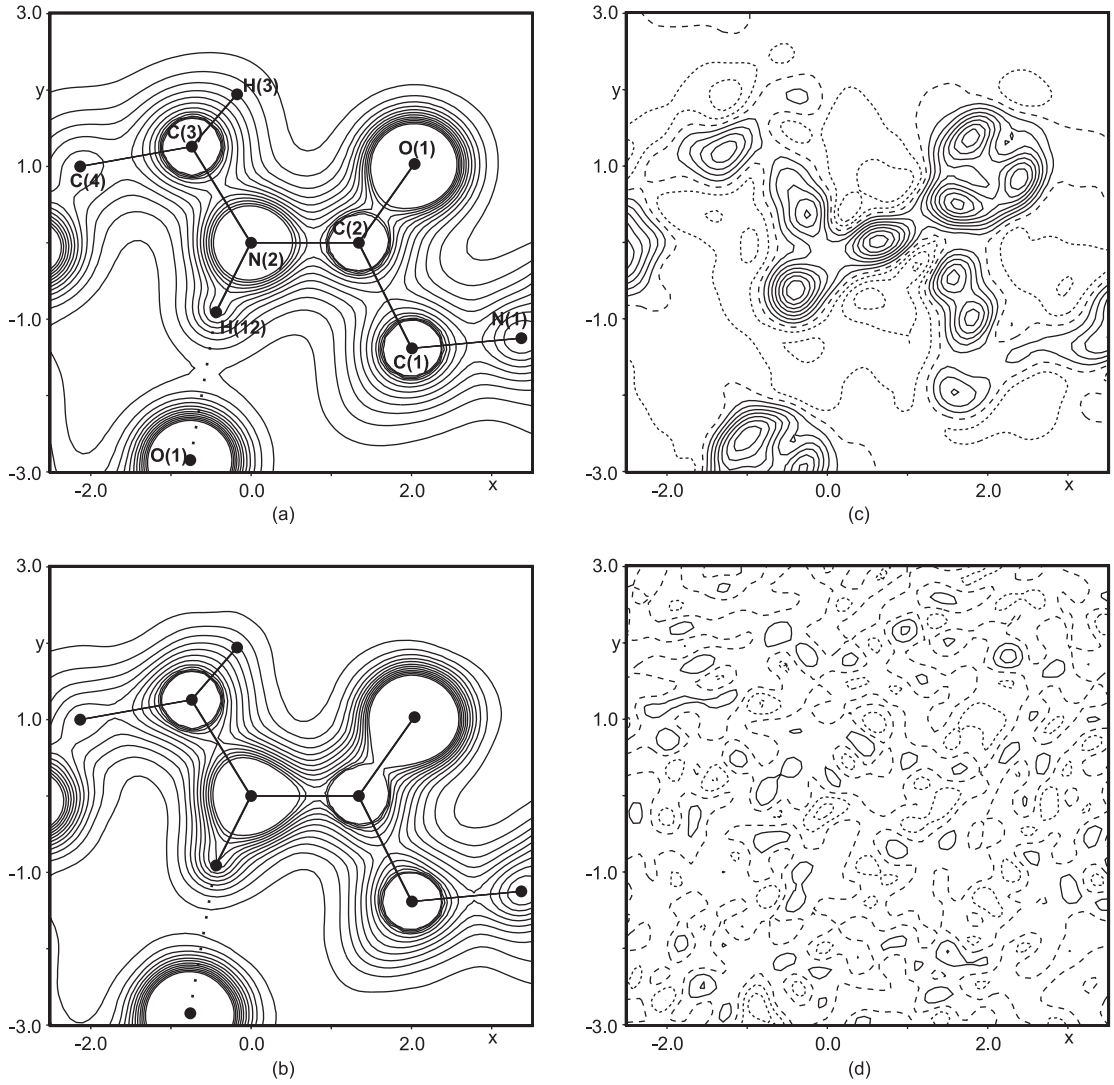

Figure 2.1: Examples of maps that are produced with the MEM. Sections of 6×6 Å^2 through the N(2)–C(2)–O(1) peptide bond plane of L-alanyl-L-tyrosyl-alanine with water as solvent. (a) ρ^{prior}, used as reference density in the MEM, here non-uniform prior density. (b) ρ^{MEM}, density optimized by the MEM. (c) Dynamic difference map [$\rho^{MEM} - \rho^{prior}$], visualizes the deformation of the density due to chemical bonding. (d) Difference Fourier map, visualizes the amount of unfitted density. Dots denote atomic maxima/atomic positions in the N(2)–C(2)–O(1) plane and out of that plane, respectively. Solid lines are contours of positive value, dotted lines are negative contours, and dashed lines represent the contour of zero value. (a) and (b): Contour of equal density are given from 0.2 to 2.5 electrons/Å^3 in steps of 0.2 electrons/Å^3. (c): Contour lines are at intervals of 0.05 electrons/Å^3. (d): Contour lines are at intervals of 0.10 electrons/Å^3.

Diffraction data are taken into account by the F-constraint[21,42] $C_{F^2} = 0$ with

$$C_{F^2} = -\chi^2_{aim} + \frac{1}{N_F}\sum_{i=1}^{N_F} w_i \left(\frac{|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|}{\sigma(\mathbf{H}_i)} \right)^2 \tag{2.3}$$

where $F_{obs}(\mathbf{H}_i)$ is the phased observed structure factor of the Bragg reflection with scattering vector $\mathbf{H}_i$ and standard uncertainty (s.u.) $\sigma(\mathbf{H}_i)$. $F_{obs}(\mathbf{H}_i)$ are obtained from the ISAM refinement, for which details for the studied compounds are given in Appendices A,C,D,F,G. $F_{MEM}(\mathbf{H}_i)$ are obtained from $\{\rho_k\}$ by discrete Fourier transform. The summation includes all N_F observed structure factors. The standard MEM employs weights $w_i = 1$ and $\chi^2_{aim} = 1$.[43,21,44] The difference $|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|$ is considered to be consistent with the experimental error $\sigma(\mathbf{H}_i)$. By that assumption, the summation over the measured structure factors would equal N_F and thus, for the standard version of the MEM with $\chi^2_{aim} = 1$, the F-constraint would be fulfilled, $C_{F^2} = 0$.

To determine the most probable density, the problem of the maximum entropy has to be solved. The method of undetermined Lagrange multipliers is employed for that purpose. The maximum of the Lagrangian

$$Q = S - \lambda C_{F^2} \tag{2.4}$$

has to be determined for variations of λ and $\{\rho_k\}$. Q is maximal when $C_{F^2} = 0$ and $\{\rho_k\}$ fulfill a set of N_p non-linear equations

$$\rho_j = \rho_j^{prior} \exp\left[-\lambda \frac{\partial C_{F^2}}{\partial \rho_j} \right] \tag{2.5}$$

which cannot be solved analytically.[43,21] Thus, this set of N_p equations has to be iteratively solved.

In the computer program BayMEM,[45] two iterative procedures, *i.e.* the Cambridge algorithm[43,46,47] and the Sakato-Sato algorithm,[21,48] are available to determine the values $\{\rho_k^{MEM}\}$ of the density $\rho^{MEM}(\mathbf{x})$ optimized by the MEM [see example in Figure 2.1(b)]. The Cambridge algorithm is incorporated into the computer program BayMEM[45] via the MemSys5 package.[49]

The MEM equations $C_{F^2} = 0$ (Equation 2.3) and Equation 2.5 are iteratively solved. The Cambridge algorithm[43,46,47] simultaneously optimizes the Lagrange parameter λ and $\{\rho_k\}$ using

$$\rho_j^{(n)} = \rho_j^{prior} \exp\left[-\lambda \left(\frac{\partial C_{F^2}}{\partial \rho_j} \right)^{(n)} \right] \quad . \tag{2.6}$$

The iteration procedure starts with a sufficiently small value of λ at the global maximum of $S = 0$ (Equation 2.1), where $\{\rho_j^{(n)}\} = \{\rho_j^{(0)}\} = \{\rho_j^{(prior)}\}$. For the next iterative step, λ is marginally increased and the optimized $\{\rho_j^{(n+1)}\}$ determined, maximizing Q (Equation 2.4). The value of λ is increased again, which can be done along with the iteration of $\{\rho_j^{(n)}\}$ itself, until convergence of the iterations and the global maximum of Q is reached. Convergence of iterations is considered to be reached when the F-constraint drops below its expectation value zero, $C_{F^2} = 0$.

The Sakato-Sato algorithm[21] chooses a sufficiently small value of λ and then determines $\{\rho_k^{MEM}\}$ iteratively, employing the modified MEM equation[48]

$$\rho_j^{(n+1)} = \rho_j^{(n)} \exp\left[-\lambda \left(\frac{\partial C_{F^2}}{\partial \rho_j}\right)^{(n)}\right] \quad . \tag{2.7}$$

Both the Cambridge and the Sakato-Sato algorithm employ initially $\{\rho_k^{(0)}\} = \{\rho_k^{prior}\}$ for the iterative procedure. Unlike the Cambridge algorithm, the Sakato-Sato algorithm additionally replaces $\{\rho_k^{prior}\}$ by the values $\{\rho_k^{(n)}\}$ of the previous iteration. The principle of the MEM does not allow an updating of the prior. A direct comparison of both algorithms with the computer program BayMEM[45] shows that the Cambridge algorithm leads to a density which is marginally better than from the Sakato-Sato algorithm. Thus, the MEM calculations of the present work have been performed with the computer program BayMEM[45] and the Cambridge algorithm, incorporating several enhancements to the MEM, which are discussed in Section 2.2.2.

2.2.2 Extensions to the MEM

A flat prior density with

$$\rho_k^{prior} = \frac{N_e}{V_{cell}} \tag{2.8}$$

does not incorporate other information than the number of electrons in the unit cell, which are uniformly distributed over the volume of the unit cell. Whereas a non-uniform prior contains information about atoms and their positions in the unit cell. The use of a flat prior would result in MEM densities that contain noise and artifacts, of which the effect on the density is larger than effects due to chemical bonding.[50] The preference of a non-uniform prior above a flat prior as reference electron density has shown to be favorable, because it enhances the quality of the MEM density.[10]

The non-uniform prior is obtained by employment of the independent spherical atom model (ISAM), which provides coordinates and atomic displacement parameters for the computation of the non-uniform prior electron density [see example in Figure 2.1(a)].

Spurious maxima of the MEM densities are eliminated by the method of prior-derived F-constraint (PDC)[51] with

$$C_{F^2}^{PDC} = -\chi^2_{aim} + \frac{1}{N_{all}} \sum_{i=1}^{N_F} w_i \left(\frac{|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|}{\sigma(\mathbf{H}_i)} \right)^2 + \frac{1}{N_{all}} \sum_{l=1}^{N_{PDC}} w_l \left(\frac{|F_{prior}(\mathbf{H}_l) - F_{MEM}(\mathbf{H}_l)|}{\sigma(\mathbf{H}_l)} \right)^2 \tag{2.9}$$

where $N_{all} = N_F + N_{PDC}$. The iterations are performed with the summation of Equation 2.9, which includes an extra term (compare to Equation 2.3) incorporating reflections that were not measured. $F_{prior}(\mathbf{H}_l)$ up to a high resolution, *e.g.* up to $\sin(\theta)/\lambda = 2.5$ Å^{-1}, are obtained from $\{\rho_k^{prior}\}$ by discrete Fourier transform. The smallest standard uncertainty of the experimental reflections is selected for $\sigma(\mathbf{H}_l)$.

Due to decreasing scattered intensities with increasing scattering angle, structure factors may be measured as weak or unobserved. By the method of PDC, they are obtained by Fourier transform of the prior density. These calculated structure factors are good estimates for structure factors of high-angle reflections, because mainly core electrons, which are considered to be well described by the ISAM, contribute to high-order reflections. The employment of the PDC enhances the quality of the densities produced by the MEM. However, the method of PDC can only be successful if a certain minimum of resolution of the experimental data is available, *e.g.* up to $\sin(\theta)/\lambda > 0.9$ Å^{-1},[51] which is a requirement on the data that is generally necessary for the purpose of charge density studies.

Within the F-constraint (Equation 2.3) and the PDC (Equation 2.9), respectively, static weights

$$w_i = \frac{1}{|\mathbf{H_i}|^n} \left(\frac{1}{N_F} \sum_{i=1}^{N_F} \frac{1}{|\mathbf{H_i}|^n} \right)^{-1} \tag{2.10}$$

have been chosen according to de Vries *et al.*,[52] where $\mathbf{H_i}$ is the scattering vector of the Bragg reflection and n a small positive integer. The standard MEM employs $w_i = 1$,[43,21,44] which would lead to a non-Gaussian distribution of the residuals

$\Delta F(\mathbf{H}_i)/\sigma_i$[46,52] with

$$\Delta F(\mathbf{H}_i)/\sigma_i = \frac{1}{\sigma_i}\,[F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)] \qquad . \tag{2.11}$$

The studies of trialanine (Chapter 3)[42] and α-glycine (Chapter 4)[53] have shown, that weighting of $\mathbf{H_i}^n$ with $n = 3$, 4 and 5 produce Gaussian distribution of the residuals $\Delta F(\mathbf{H}_i)/\sigma_i$ (Equation 2.11). Large residuals for low-order reflections, that would occur with $w_i = 1$, are suppressed and a larger weight is given to reflections with short scattering vectors. In accordance with de Vries *et al.*,[52] a weighting with $n = 4$ has been chosen for all MEM calculations of the present work. The Gaussian distribution of $\Delta F(\mathbf{H}_i)/\sigma_i$ of the studied compounds is displayed in Figure H.1 of Appendix H.

The choice of an optimal value of χ^2_{aim} as stopping criterion for the MEM calculation is of high importance, since it determines the point of convergence of the MEM iterations through the criterion $C_{F^2} = 0$ (Equation 2.3). For employment of $C^{PDC}_{F^2}$ (Equation 2.9) by the method of PDC instead of C_{F^2}, the convergence of the iterations is still tested by C_{F^2} (Equation 2.3).

The stopping criterion for the standard versions of the MEM corresponds to $\chi^2_{aim} = 1$ (Equation 2.3),[43,21,44] denoted as historical MEM.[49] χ^2_{aim} equals one when the difference $|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|$ is consistent with the experimental error $\sigma(\mathbf{H}_i)$. Theoretically, convergence would be reached with values of $\chi^2_{aim} < 1$, because the average difference $|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|/\sigma(\mathbf{H}_i)$ is expected to be smaller than one.[47,46] The behavior, that χ^2_{aim} at the point of convergence is smaller than one, is taken into account by the so-called classical MEM.[47,46] If standard uncertainties of observed reflections have been estimated smaller than their true values, values of $\chi^2_{aim} > 1$ can be determined.

One important tool for the evaluation of the quality of the MEM densities is the difference Fourier map, which visualizes the amount of unfitted density [see example in Figure 2.1(d)]. By implementation of that tool into the computer program BayMEM,[45] it is possible to produce hard evidence for the determination of the optimal value of χ^2_{aim}.

Inspection of dynamic difference maps $[\rho^{MEM} - \rho^{prior}]$ [see example in Figure 2.1(c)], also denoted as dynamic deformation maps, and difference Fourier maps obtained by the MEM, is employed to determine the optimal value of χ^2_{aim}. It has shown that, if χ^2_{aim} is too small, it would result in overfitted data, leading to

the incorporation of experimental errors and noise into the densities. The dynamic deformation map in the case of a too low χ^2_{aim} would show ripples of the contour lines, while the difference Fourier map would be flat and featureless. Whereas, a too large value of χ^2_{aim} results in a large amount of unfitted data not taken account into the density map because the MEM calculation did not converge yet. Thus, remaining density would be visible as structure larger than the noise in the difference Fourier map, while the $[\rho^{MEM} - \rho^{prior}]$ map would exhibit smooth contour lines.

Extensive calculations on trialanine and α-glycine (Chapters 3 and 4)[42,53] have yielded two different approaches for the purpose of the determination of the optimal value of χ^2_{aim}. One approach employs the classical MEM[47,46] initially to determine the optimal value of χ^2_{aim}. The classical MEM cannot be combined with the method of PDC,[51] because it is properly defined only for an F-constraint based on experimental data. Thus, the classical MEM is performed without the PDC.[51] The value of χ^2 at the point of convergence of the classical MEM divided by N_F yields the effective value of χ^2_{aim}, which is subsequently set as χ^2_{aim} for the historical MEM[49] (Chapter 4).[53] Once the value χ^2_{aim} is determined, the historical MEM, employing the PDC[51] and static weighting of $\mathbf{H_i}^n$ with $n = 4$ according to de Vries *et al.*,[52] is performed to reconstruct the optimized MEM density.

It has been shown, that the χ^2_{aim} obtained in that way, may lead in some cases (Appendices F and G) to overfitted data, which would be visible as ripples in the difference maps. In such cases the value of χ^2_{aim} from the classical MEM is too small and thus used as benchmark for a manual search for the optimal value. For that procedure, the value of χ^2_{aim} from the classical MEM is gradually increased and used subsequently in the historical MEM. By inspection of the deformation maps and the difference Fourier maps, the optimal χ^2_{aim} is pinpointed (Appendices F and G). Eventually, by this procedure that MEM calculation with the optimal χ^2_{aim} is chosen.

The other method for the determination of the optimal χ^2_{aim} starts with the historical MEM without preceding classical MEM. For that approach, series of historical MEM calculations, employing the PDC[51] and static weighting of $\mathbf{H_i}^n$ with $n = 4$ according to de Vries *et al.*,[52] with arbitrarily chosen values of χ^2_{aim} around one, are performed (Chapter 3).[42] By inspection of the resulting $[\rho^{MEM} - \rho^{prior}]$ maps and difference Fourier maps, the point of convergence is estimated by a small amount of unfitted density, visible as structure in the difference Fourier maps, and

incorporated noise, visible as ripples in the difference maps. With respect to the inspection of these maps, that MEM calculation with the optimal χ^2_{aim} is chosen.

2.2.3 The Atoms in Molecules Theory

The Quantum Theory of Atoms in Molecules[13] allows to interpret electron densities quantitatively and attain information about chemical bonding. Analysis of the topological properties of the charge density $\rho(\mathbf{x})$ in the three-dimensional space yields the number and kind of its critical points, which constitute maxima, minima and saddle points of the charge density, where the first derivatives of $\rho(\mathbf{x})$ vanish, $\nabla\rho = 0$. Characterization of critical points of the charge density provides insight into the conventional chemical structure of the respective molecule.

Critical points are denoted by (ω, σ). ω is equal to the number of non-zero eigenvalues, λ_1, λ_2 and λ_3, of the Hessian matrix of the density. Energetically stable critical points have the value $\omega = 3$. The algebraic signs of the eigenvalues λ_1, λ_2 and λ_3 determine σ. A (3,-3) critical point possesses three negative curvatures and corresponds to a local maximum of ρ at that point, and describes an atomic maximum. Two negative curvatures, with a maximum of ρ in the plane defined by these two axes, and one positive curvature with a minimum of ρ along this axis, define a (3,-1) bond critical point (BCP), which describes a chemical bond. A ring critical point (3,+1), describing a chemical ring structure, is defined by two positive curvatures, with a minimum of ρ in the plane defined by these two axes, and one negative curvature with a maximum of ρ along this axis. The point (3,+3) defines a cage critical point, describing a chemical cage structure, with all three curvatures positive and ρ as local minimum at that point.

Due to the properties of the density at a critical point, whether it has a maximum or a minimum on the three axes of its principal curvature, the space of the charge distribution is partitioned into atomic basins, each containing one nucleus. Thus, a single atom is defined as an entity of a nucleus and its associated atomic basin. The sum of all atoms with their assigned basins form the molecule. Additionally, the existence of (3,-1) bond critical points of the density provides a boundary (interatomic surface) between basins of connected or neighboring atoms. Integration of the charge over the volume of an atomic basin yields the atomic charges.

The existence of local maxima of the density, *i.e.* (3,-3) critical points, indicate positions of atomic maxima in the unit cell [see examples in Figures 2.1(a) and (b)]

and their associated atomic basins contain the charges of the respective atoms. Ring or cage structures are revealed by (3,+1) ring and (3,+3) cage critical points, respectively. The presence of BCPs with their properties such as the density $\rho(BCP)$, the eigenvalues λ_1, λ_2 and λ_3, and Laplacians $\nabla^2\rho(BCP) = \lambda_1 + \lambda_2 + \lambda_3$, allow characterization of chemical bonds. A negative Laplacian, with a large magnitude, and a large $\rho(BCP)$ indicate a charge concentration at the BCP, resulting in a sharing of electronic charge by the nuclei involved. Such interactions are shared-shell interactions, which are typical for covalent bonds. $\nabla^2\rho(BCP) > 0$, together with a relatively small $\rho(BCP)$, indicate charge a depletion at the BCP. Such properties are regarded as closed-shell interactions and are found with non-covalent interactions such as van der Waals, ionic, dipolar interactions or hydrogen bonds. Thus, with the aid of the AIM theory,[13] the chemical structure is recovered and chemical bonds of various types are characterized.

Electron density maps have been analyzed according to Bader's AIM theory[13] with the module EDMA of the program BayMEM.[45] Local maxima of the density, atomic basins, atomic charges and bond critical points (BCPs) with their densities $\rho(BCP)$, their principal curvatures (eigenvalues) and their Laplacians $\nabla^2\rho(BCP)$ were determined.

2.2.4 Summary

The MEM can be used in crystallography for deconvolution of powder diffraction data, to solve the phase problem or to determine the electron density in the unit cell. Electron densities can be employed to locate atoms in the unit cell or to extract, in addition to atomic positions, information about chemical bonding. Disorder such as partially occupied atomic sites, multiple orientations or multiple conformations of molecules or functional groups can be revealed by the MEM. For the purpose of the determination of Accurate Charge Densities via the computer program BayMEM,[45] the density is defined by its values on a grid over the unit cell. By an iterative search, the most probable density is determined, which simultaneously maximizes the informational entropy, fulfills the normalization constraint and is subject to the diffraction data via the F-constraint.[50] The quality of the optimized electron densities is enhanced by several extension to the MEM. Artifacts in the MEM densities are eliminated by usage of a non-uniform prior[10] and the method of PDC.[51] Residuals $\Delta F(\mathbf{H}_i)/\sigma_i$ show Gaussian distribution by appropriate choice of static weighting.[52]

The point of convergence of the MEM calculations is determined by the choice of an optimal χ^2_{aim} with the aid of inspection of the dynamic deformation maps and difference Fourier maps (Chapters 3 and 4).[42,53] The latter has been implemented into the computer program BayMEM[45] to provide a tool for evaluating the quality of the MEM densities. Analysis of the MEM densities according to the AIM theory[13] yields the positions of atoms in the unit cell and information about chemical bonds such as hydrogen bonds and covalent bonds.

Chapter 3

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

Abstract

An accurate charge density study of trialanine is presented with the maximum entropy method (MEM), on 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 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.

3.1 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.[54,15,16,55]

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,[17] 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 $[\rho^{MEM}(r)]$ is affected by artefacts such as spurious maxima or 'ripples' in the charge-density distribution which are specific to the MEM.[8,9,10,12,50,56] 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[10,57,51]).

The eventual aim of these improvements is to achieve a $\rho^{MEM}(r)$ which is at least comparable in quality to electron-density maps $[\rho^{multipole}(r)]$ that are obtained by refinements based on the multipole formalism.[1] 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 20K X-ray diffraction data of the tripeptide L-alanyl-L-alanyl-L-alanine (trialanine) which has already been used for a multipole refinement.[14]

3.2 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

$$S = -\sum_{k=1}^{N_p} \rho^k \log\left(\frac{\rho^k}{\rho^k_{prior}}\right) , \tag{3.1}$$

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 $\{\rho^k\}$ that maximizes the entropy S, while one or more constraints are fulfilled. Besides the normalization of $\{\rho^k\}$,

$$C_0 = -1 + \frac{1}{\rho_{total}} \cdot \sum_{k=1}^{N_p} \rho^k \tag{3.2}$$

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

$$C_F = -\chi^2 + \sum_{hkl}^{N_{ref}} \left(w_{hkl} \frac{|F^{obs}_{hkl} - F^{MEM}_{hkl}|^2}{\sigma^2} \right) \quad . \tag{3.3}$$

Here F^{obs}_{hkl} and F^{MEM}_{hkl} 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 3.3) computed with $w_{hkl} = 1.0$ with the stopping criterion. In the historical MEM the stopping criterion corresponds to the classical least-squares refinement.[58] The constraint is fulfilled if $\chi^2 = N_{Ref}$. According to Gull & Skilling,[58] the historic MEM is not Bayesian and therefore imperfect. The constraint $\chi^2 = N_{Ref}$ 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 $\Pr(F^{MEM}|\rho^{obs})$ 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 $\chi^2 = N_{Ref}$ 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, $\chi^2 = N_{Ref}$ is too pessimistic. Therefore, Gull & Skilling[58] recommended 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[51] and *ad hoc* weighting.[10] In order to retain compatibility of these MEM enhancements, the historical MEM for the calculation of electron density maps should be kept. However, since $\chi^2 = N_{Ref}$ is too pessimistic, a value of χ^2 smaller than N_{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[8] 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 $\chi^2/N_{Ref} = \chi^2_{aim}$. Several methods have been suggested to counterbalance this effect.[52,57,51] In this work the *ad hoc* weighting scheme as suggested by de Vries *et al.*[52] is applied. The resulting equation for the F constraint is (3.3), where the weights w_{hkl} are defined as

$$w_{hkl} = \frac{1}{|\mathbf{H}_{hkl}|^n} \cdot \left(\frac{1}{N_{ref}} \sum_{hkl}^{N_{ref}} \frac{1}{|\mathbf{H}_{hkl}|^n} \right)^{-1} \tag{3.4}$$

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). $H0$ 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.*[52] found that $n = 4$ ($H4$) gives the best results.

Table 3.1: Crystallographic data.[59]

Formula	$C_9H_{17}N_3O_4 \cdot H_2O$
M_r	243.3
Space group	Monoclinic, $C2$
Z	8
T (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
$\sin(\theta)/\lambda$ max (Å^{-1})	1.15
Unique data (measured/prior-derived)	14895/281077
Completeness (%)	93.3
R_{int}	0.0295
R(ISAM)	0.0314

3.3 Experimental

3.3.1 Refinement

Data collection (Mo $K\alpha$ radiation at 20 K) and data reduction have been described in Rödel *et al.*,[14] who generously gave us a copy of the reflection data file. The most important crystallographic data are summarized in Table 3.1.

Refinements with the independent spherical atom model (ISAM) were performed with the computer program *JANA*2000,[60] 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.[61] 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 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.*[14]). An instability factor of 0.005 was used. The crystal structure is shown in Fig. 3.1. 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, O; isotropic for H) of the spherical refinement] was created with the module PRIOR of *BayMEM*.[45]

3.3.2 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*.[45] The MEM calculations were performed with the latest version of *BayMEM*,[45] using an adapted version of the commercially available *MEMsys5* algorithm package.[58] For the grid-based MEM the unit cell was divided into 216×64×324 voxels, corresponding to voxel edge lengths of 0.085 × 0.081 × 0.077 Å^3. 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 $\sin\theta/\lambda$ region 0.9 - 2.5 Å^{-1}, as suggested by Palatinus and van Smaalen.[51]

Choice of parameters: weighting

As already mentioned, the unmodified MEM has a tendency to dramatically underfit some strong low-angle reflections, whereas a number of the remaining reflections are over-fitted to satisfy the requirement $\chi^2/N_{Ref} = \chi^2_{aim}$. In order to counterbalance this effect the *ad hoc* weighting scheme proposed by de Vries *et al.*[52] was utilized. The usage of this weighting scheme results in a more Gaussian-like distribution of the residuals. Fig. 3.2 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.

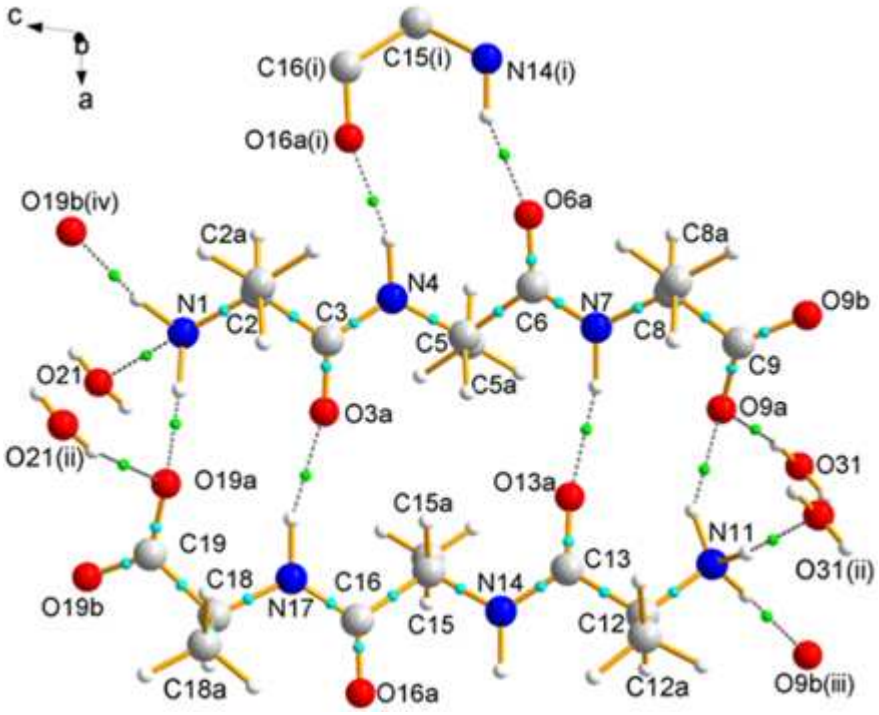

Figure 3.1: Perspective representation of both crystallographic independent trialanine molecules, together with the water molecules and with all hydrogen bonds (dashed lines). Bond critical points are given for all hydrogen bonds (green dots) and all C–C, C–N and C–O bonds (cyan dots). All 12 hydrogen 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) $1-x, y, 1-z$.

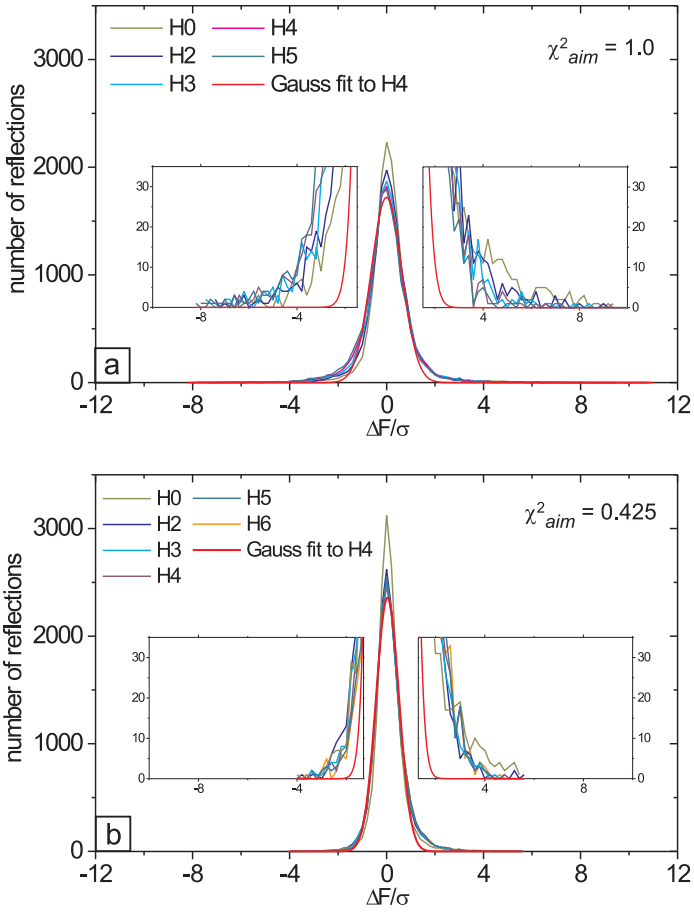

Figure 3.2: Distribution of residuals $(|F^{obs}\mathbf{H}| - |F^{MEM}\mathbf{H}|)/\sigma$ for weights $H0$ through $H5$ and for (a) $\chi^2_{aim} = 1.0$ and (b) $\chi^2_{aim} = 0.425$. $H0$ means no *ad hoc* weighting. The Gaussian curve is shown in red. The insets show a magnification of the outer regions. The number of reflections in intervals of 0.2 wide $\Delta F/\sigma$ are given.

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^{obs}, which means that the FWHM is smaller and consequently the Gaussian curve is higher than in the case where $\chi^2_{aim} = 1.0$. Nevertheless, the observed outliers deviate less from zero ($\Delta F/\sigma$ = -3.4 to 4.8 *versus* -8 to 9.4 for $\chi^2_{aim} = 1.0$, $H4$) than would be expected solely from the reduced width of the histogram (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 corresponding 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.*[52] Close inspection of the tails of the Gauss curve (magnified region in the insets of Fig. 3.2) 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 under-fitted structure factors, it cannot completely suppress the deviations from the true Gaussian shape of the final distribution of residuals.

The choice of parameters: χ^2_{aim}

The historical MEM uses $\chi^2_{aim} = N_{Ref}$ ($\chi^2_{aim} = 1.0$) 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 3.2). 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. 3.3b) clearly demonstrate that a significant amount of electron density is not fitted by the MEM calculation. This means that the MEM

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

Table 3.2: Averaged ρ (first line) and $\nabla^2\rho$ (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[14] were achieved with the smaller value of χ^2_{aim} (0.425).

	Prior (ISAM)	MEM ($\chi^2_{aim} = 1.0$)	MEM ($\chi^2_{aim} = 0.425$)	Multipole	B3LYP/6-311 ++G(d,p) Calc.
$C_{peptide}$–$O_{peptide}$	2.11(1)	2.33(1)	2.55(2)	2.87(4)	2.65(2)
	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(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_α–$N_{ammonium}$	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_α–$N_{peptide}$	1.45(1)	1.58(1)	1.74(1)	1.76(4)	1.69(1)
	2(1)	-3(2)	-8(3)	-11(2)	-14.0(1)
C_α–C_β	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_α–$C_{peptide}$	1.19(1)	1.32(3)	1.52(5)	1.71(5)	1.71(1)
	0(1)	-4(2)	-7(3)	-11(2)	-14.6(2)
C_α–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)

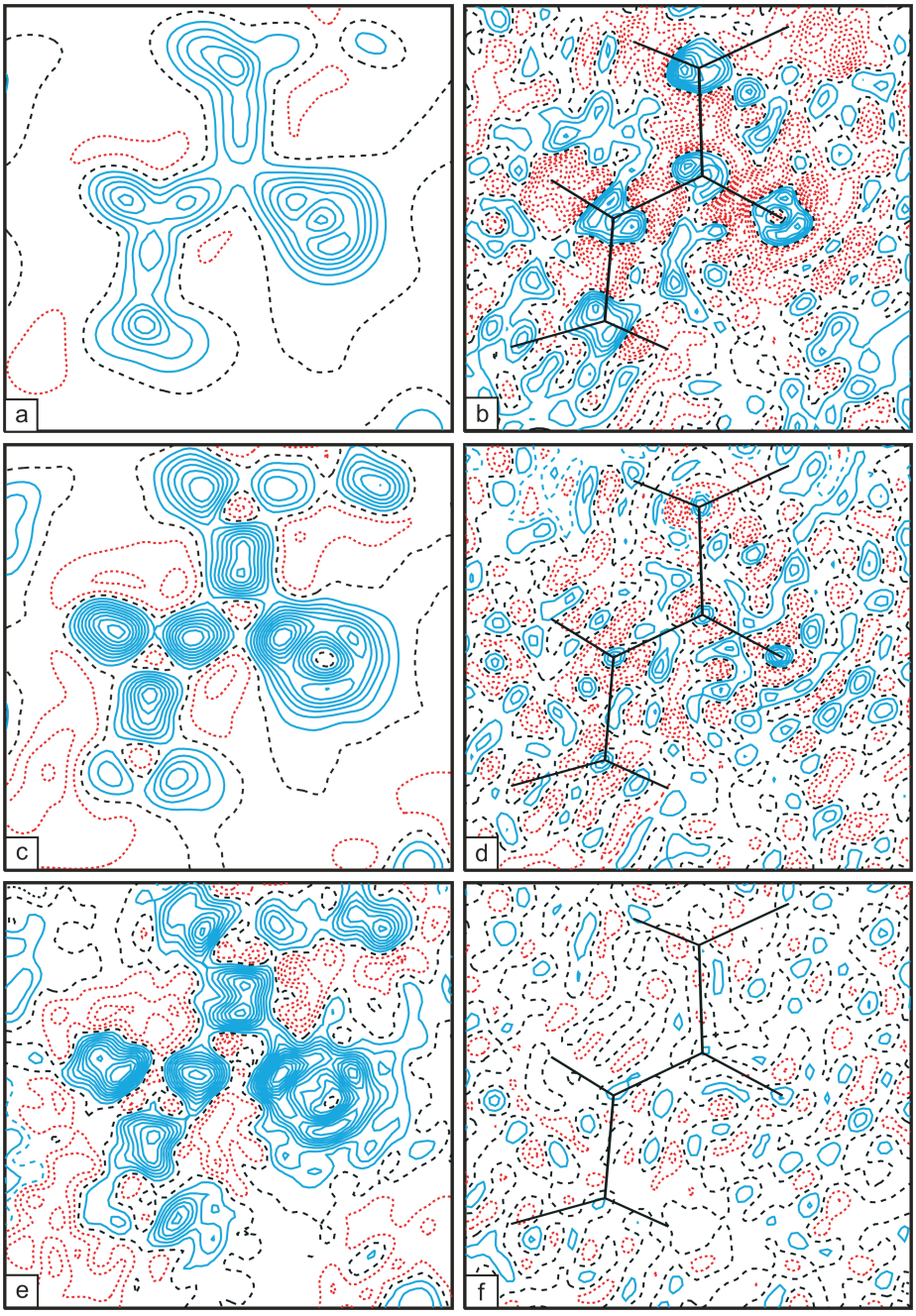

Figure 3.3: Difference maps [$\rho^{MEM} - \rho^{prior}$; images (a), (c), (e)] and residual maps [inverse Fourier transformation of $F^{obs} - F^{MEM}$; images (b), (d), (f)] of the peptide bond plane (N4–C3–O3a) for $\chi^2_{aim} = 1.0$ (a), (b), $\chi^2_{aim} = 0.425$ (c), (d) and $\chi^2_{aim} = 0.2$ (e), (f). Contour lines at 0.05 e Å^{-3}, red dotted lines denote negative, blue lines denote positive values.

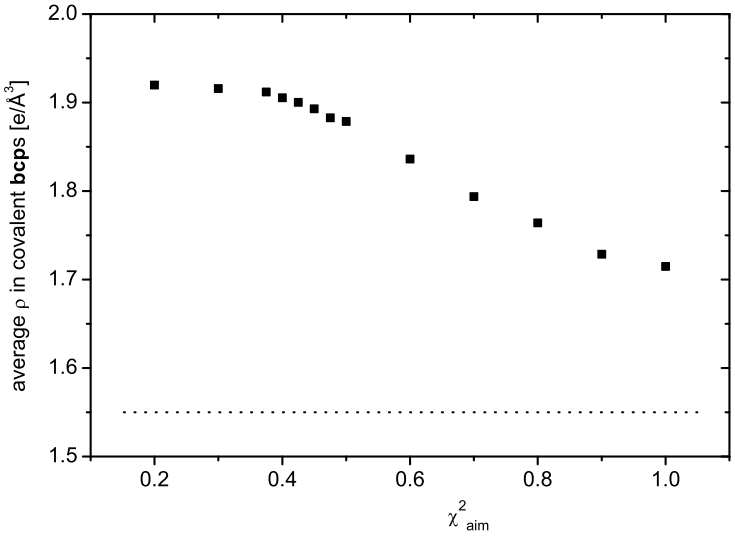

Figure 3.4: Electron density in the BCPs averaged over all C–C, C–N and C–O covalent bonds for different MEM densities depending on χ^2_{aim}. The averaged electron density in the BCPs of ρ^{prior} is 1.55 e Å^{-3} (dotted line).

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. 3.4 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. 3.3e). 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. 3.3c). However, the corresponding residual map is not entirely feature-free. Especially the areas close to the atom centres show deviations of up to ± 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. 3.3e, 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[53] 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.

3.4 Results and discussion

3.4.1 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.[8,9,10,12,50] 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 AIM theory.[13] Since these quantities are additive, 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 only 0.04 Å^3 higher than the unit-cell volume of 2362.43 Å^3 (for comparison: $\sum V_{multipole} = 2345.9$ Å^3). By integrating the electron

density within one atomic basin and subtracting the corresponding proton charges, the overall charge of a fragment can be calculated. The sum over all charges in the unit cell results in exactly in $0.00e$, as it is required by the electroneutrality of the unit cell (for comparison: $\sum Q_{multipole} = 0.04e$).

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 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. 3.5). 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. 3.5). Analysis of the model densities shows that the 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 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 to $+0.2(1)e$.

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 3.3 shows a comparison of atom charges and volumes between ISAM, MEM

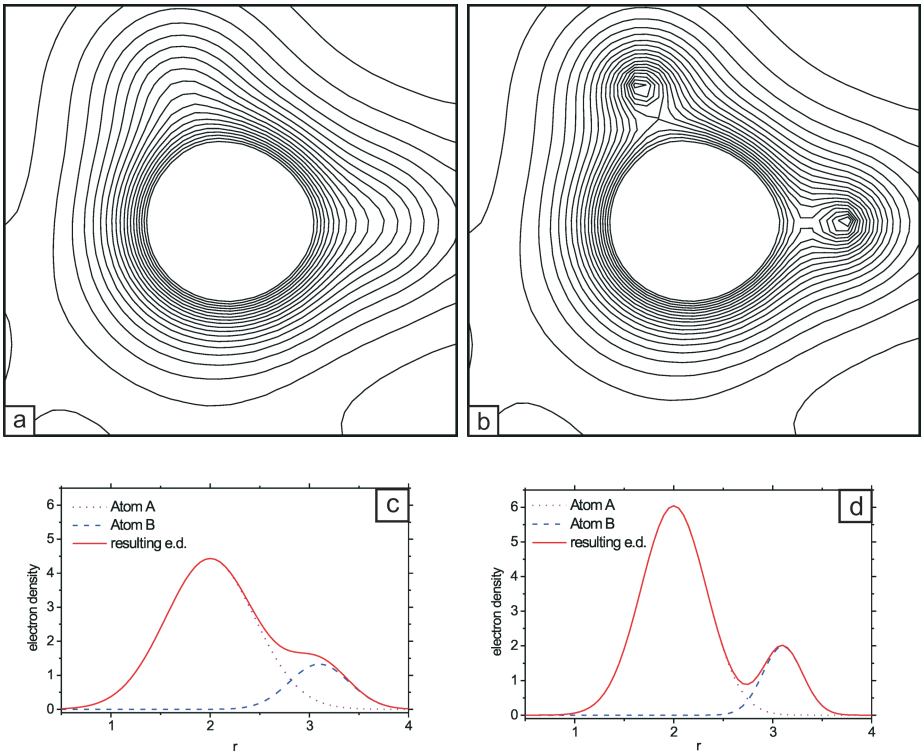

Figure 3.5: H1*a*–N1–H1*b* sections of electron densities (contour lines at 0.1 e Å^{-3}). (a) Prior electron density. (b) Model electron density similar to the prior, but with ADPs set to zero. Owing to the thermal motion, the individual electron-density peaks are broader and the resulting electron density does not show individual maxima for the H atoms (a), (c). Graphs (c) and (d) schematically illustrate this effect as one-dimensional sections.

Table 3.3: Comparison of averaged atomic charges (e) and volumes (Å^3).

Note that for $N_{peptide}$ and $N_{ammonnium}$ 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)
$O_{peptide}$	-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)
$O_{short\ carboxy}$	-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)
$NH_3/N_{ammonium}$	*+0.01(1)*	*25.9(9)*	*+0.40(5)*	*24(1)*	-1.2(1)	15.5(4)

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 [$Q_{MEM} = +1.40(1)e$, $Q_{multipole} = +1.17(5)e$]. It seems noteworthy that the charge of the carbon of the peptide bond averages almost $0.2e$ higher for the MEM than for the multipole refinement [$Q_{MEM} = 1.29(5)$; $Q_{multipole} = 1.1(4)$], 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 $0.19e$ 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.*[62,63]). 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

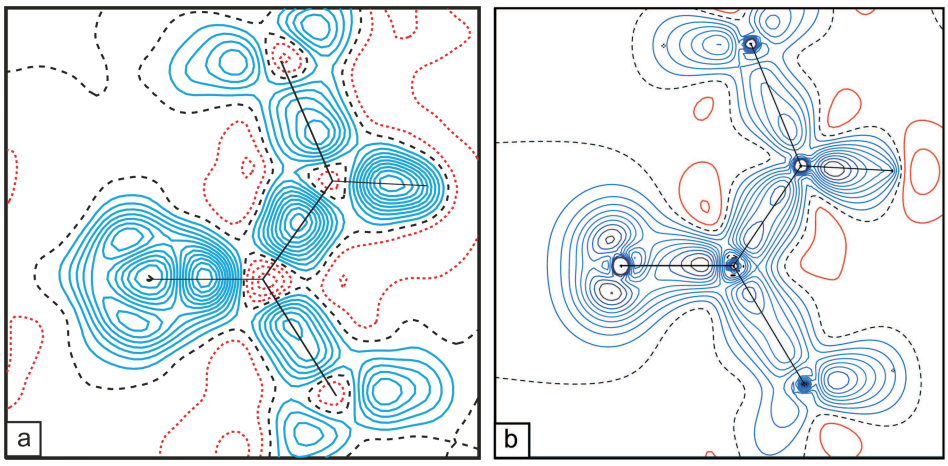

Figure 3.6: (a) MEM difference map ($\rho^{MEM} - \rho^{prior}$) and (b) multipole deformation map[59] of the peptide bond plane (O6*a*–C6–N7). Contour lines at 0.05 e Å^{-3} are for MEM and 0.10 e Å^{-3} for multipole maps; red dotted lines denote negative values, blue lines denote positive values.

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_{MEM} = 16.7(4) Å^3, $V_{multipole}$ = 20.4(2) Å^3]. Since the MEM volume is in perfect agreement with the average values from the literature [V = 16.5(9) Å^3][64] 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 that 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.

3.4.2 Covalent bonds

Although MEM difference maps ($\rho^{MEM} - \rho^{prior}$) 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. 3.6 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 AIM theory[13] 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 BCPs, 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 3.2. The positions of the BCPs are shown in Fig. 3.1. 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 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.*[65,3,66]). According to Volkov *et al.*[3] 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.*[14] 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. 3.4) 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. 3.3e). 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 replacement 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 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 [ρ(C–O$_{peptid}$) > ρ(C–O$_{long\ carboxy}$) > ρ(C–O$_{short\ carboxy}$) > ρ(C$_{peptide}$–N$_{peptide}$) > ρ(C$_\alpha$–N$_{peptide}$) $\geq$ ρ(C$_\alpha$–N$_{ammonium}$) > ρ(C$_\alpha$–N$_{carboxy}$) $\simeq$ ρ(C$_\alpha$–N$_{peptide}$) > ρ(C$_\alpha$–N$_\beta$)]. 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$_\alpha$–N$_{peptide}$ bond results from a significant 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$_\alpha$–N$_{peptide}$ 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$_{peptide}$ = 1.75 e Å^{-3}; N$_{peptide}$–C$_\alpha$–C$_{peptide}$ = 1.68 e Å^{-3}). For the other bonds the influence of the next nearest neighbour is less distinct.

According to Bader's AIM theory,[13] the Laplacian (sum of the eigenvalues of the

Hessian matrix) in the BCP gives valuable information about the type of the 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[62] 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 calculated with experimental coordinates and ADPs with a corresponding prorystal 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 $\pm$ 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.*[3]), easily explains the observed discrepancies between the multipole- and MEM-derived Laplacians of the C–O bonds.

3.4.3 Hydrogen bonds

Rödel *et al.*[14] found 12 N–H$\cdots$O and O–H$\cdots$O hydrogen bonds in the asymmetric unit (see Fig. 3.1). Their topological properties were also computed with the MEM and the values are compared with the multipole results in Table 3.4. In 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

Table 3.4: ρ and $\nabla^2\rho$ 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.[14]

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

	Prior (ISAM)		MEM		Multipole	
	ρ	$\nabla^2\rho$	ρ	$\nabla^2\rho$	ρ	$\nabla^2\rho$
N1–H1$a\cdots$O19b	0.39	1.75	0.43	0.77	0.28	4.60
N1–H1$b\cdots$O21	0.22	2.19	0.21	1.88	0.13	2.22
N1–H1$c\cdots$O19a	0.35	2.72	0.33	1.78	0.25	4.05
N4–H4$a\cdots$O16a	0.24	2.02	0.21	2.45	0.15	2.94
N7–H7$a\cdots$O13a	0.21	2.26	0.20	1.43	0.13	2.42
N11–H11$c\cdots$O31	0.34	2.52	0.37	1.41	0.20	4.02
N11–H11$b\cdots$O9a	0.22	2.39	0.19	1.60	0.13	2.52
N11–H11$a\cdots$O9b	0.39	1.40	0.45	-1.57	0.28	4.97
N14–H14$a\cdots$O6a	0.23	2.02	0.19	2.19	0.14	2.89
N17–H17$a\cdots$O3a	0.17	1.92	0.15	2.11	0.09	1.86
N21–H21$a\cdots$O19a	0.22	2.11	0.27	1.54	0.20	4.02
N31–H31$a\cdots$O9a	0.27	2.44	0.33	1.58	0.24	4.94

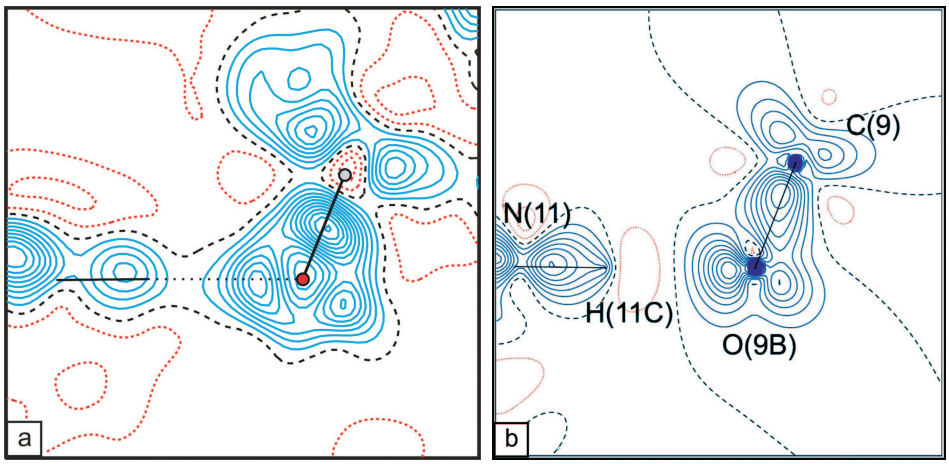

Figure 3.7: (a) MEM difference map ($\rho^{MEM} - \rho^{prior}$) and (b) multipole deformation map[59] of the hydrogen-bond plane (H11*c*–O9*b*–C9). Contour lines at 0.05 e Å^{-3} are for MEM and 0.10 e Å^{-3} for multipole maps; red dotted lines denote negative values, blue lines denote positive values.

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 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. 3.7) 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 oxygen 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

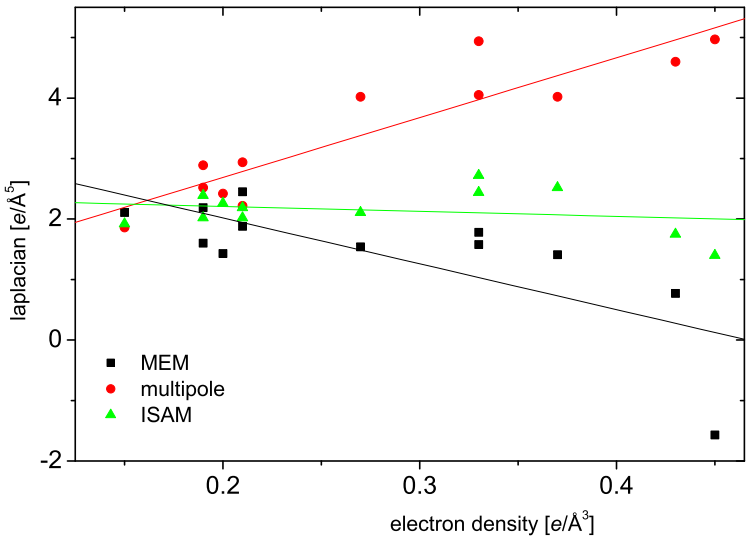

Figure 3.8: Correlation between the value of the Laplacian and electron density in BCPs of hydrogen bonds. In the case of ISAM (triangles) there is no significant slope, whereas in the case of the multipole method (circles) the slope is clearly positive, and in the case of the MEM (squares) it is significantly negative.

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.*[3]).

In general, Bader's AIM theory[13] 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. 3.8) 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$\cdots$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 electron density 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.

3.5 Conclusions

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 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 or between the MEM 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.[67,64]

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

Chapter 4

Accurate charge density of α-glycine by the maximum entropy method

Abstract

Accurate electron densities of α-glycine have been obtained by the maximum entropy method (MEM) applied to low-temperature X-ray diffraction data by Destro *et al.* (*J. Phys. Chem. A*, 2000, **104**, 1047–1054). Difference Fourier maps have been found to provide a good stopping criterion for the iterations in the MEM, in agreement with our previous findings for trialanine. Properties according to Bader's atoms-in-molecules theory are reported for the MEM electron density. These properties are found to be in agreement with the properties of the MEM electron density as they have been obtained previously for the tripeptide trialanine, thus showing the consistency of the MEM approach when applied to centrosymmetric and acentric organic compounds. The dynamic MEM electron density compares favourably with the static electron density obtained from the multipole model by Destro *et al.* (2000), with differences being attributed to the specific nature of each method. The independent spherical atom model (ISAM) and the multipole model provide different phases for 17 reflections of which only two are of the type 'observed'. A MEM calculation with reflection phases from the multipole model leads to an electron density that is only marginally different from the MEM electron density with phases from the ISAM refinement. This suggests, at least for centrosymmetric structures, that the ISAM is sufficiently good to be used as basis for the MEM

approach to accurate electron density studies.

4.1 Introduction

Accurate electron density studies provide information on chemical bonding that can be used to develop models for the stability and chemical reactivity of molecules, molecular aggregates and inorganic compounds.[13] The established experimental method of accurate electron density studies is structure refinement of the multipole model against single-crystal X-ray diffraction data.[68] Despite its success the multipole method suffers from the problem of dependent parameters that becomes more severe for an increasing size of the problem. This problem is usually solved by the introduction of constraints on the parameters that go beyond the constraints imposed by the symmetry of the crystal structure.

The maximum entropy method (MEM) is a method that can be used for model-independent reconstruction of a variety of image types.[54] One crystallographic application is the reconstruction of the electron density in the unit cell.[17] Inherently, the MEM does not suffer from interdependencies between parameters, but it may produce artifacts in the reconstructed images that are specific to the method. Accordingly, the feasibility of the MEM as a method for accurate electron density studies has been controversial.[8,9,10,11,12,50] Several procedures have emerged as being necessary for accurate electron density studies by the MEM. These include the use of special weighting schemes for the reflection data in the χ^2 constraint,[52] the use of a procrystal prior density as opposed to a flat prior,[10] and the employment of prior-derived F-constraints.[51] Furthermore it was found that the successful application of the MEM requires extensive data sets of accurate X-ray diffraction data.[10] The determination of the MEM electron density requires an iterative procedure. Recently we have shown that the point where the iterations are stopped has a large influence on the final electron density, and a procedure was developed for the determination of the optimal stopping point of the iterations.[42]

Accurate electron density studies by the multipole method have been performed for small peptides with the goal of developing databases of transferable multipole parameters for applications of the multipole model to proteins.[69,70,71] Accurate electron densities have also been obtained for most amino acids.[72] In two recent studies we have demonstrated the feasibility of the MEM for accurate electron den-

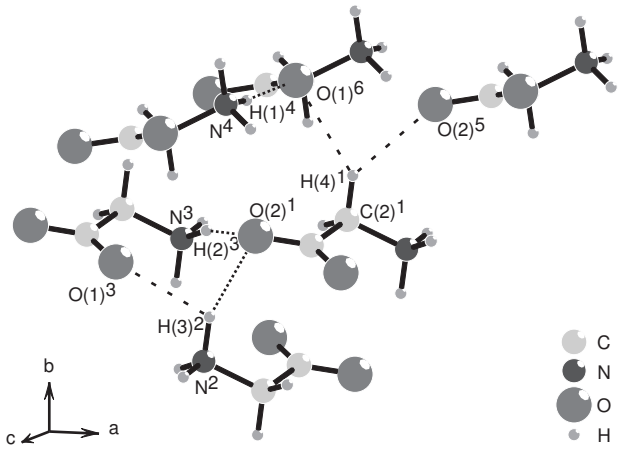

Figure 4.1: Perspective view of α-glycine obtained with atomic coordinates from the ISAM. Hydrogen bonds are indicated by dotted lines, weak hydrogen bonds are indicated by dashed lines. Symmetry operators are denoted by superscripts: $1:(x, y, z)$; $2:(-x, -y, -z)$; $3:(-1+x, y, z)$; $4:(-0.5+x, 0.5-y, 0.5+z)$; $5:(0.5+x, 0.5-y, -0.5+z)$; $6:(-0.5+x, 0.5-y, -0.5+z)$.

sity studies of tripeptides by comparing MEM densities and multipole densities of the trialanine and ala-pro-ala.[42,73] Both compounds crystallize in acentric space groups. In order to further establish the MEM as an alternative to multipole refinements for accurate electron density studies we presently report the results of the MEM applied to α-glycine (Fig. 4.1). This amino acid has been selected because it has a centrosymmetric crystal structure and because excellent diffraction data at very low temperatures as well as a very good multipole refinement are available for this compound.[74] Employing the same diffraction data as have previously been used for the multipole refinement we show that a good agreement between MEM and multipole densities is achieved, while typical differences can be attributed to the characteristic features of both methods.

4.2 Computational details

Single-crystal X-ray diffraction data were obtained from Destro *et al.*,[74] who also reported a multipole refinement with these data (Table 4.1). The MEM has been applied following procedures described in detail elsewhere.[10,52,51,42,75] As a first step

Table 4.1: Crystallographic data from Destro *et al.*[74] and summary of the ISAM refinement (present work). Reflections with $I/\sigma(I) > 3$ are classified as observed

Chemical formula	$C_2O_2NH_5$
Space group	$P2_1/n$
Z	4
a/Å	5.0866 (20)
b/Å	11.7731 (30)
c/Å	5.4595 (20)
β/°	111.99 (15)
V/Å^3	303.2
F(000)	160
Temperature/K	23
Wavelength, λ/Å	0.71073
$[\sin(\theta)/\lambda]_{max}$/ Å^{-1}	1.15
Number of unique reflections (obs/all)	3483/3822
Multipole refinement[74]	
R_F(obs)	0.0129
wR_{F^2}(obs)	0.0172
GoF(obs)	1.041
ISAM refinement	
R_F(obs/all)	0.0233/0.0260
wR_{F^2}(obs/all)	0.0525/0.0535
GoF(obs/all)	2.02/1.96

we have performed a refinement of the independent spherical atom model (ISAM) against all data (Table 4.1). An anisotropic extinction correction was incorporated into the refinement, because it was found to be important by Destro *et al.*[74] Hydrogen atoms were included at distances from the carbon and nitrogen atoms, corresponding to the distances known from neutron diffraction (1.09 Å for C-H and 1.03 Å for N-H).[76] All refinements were performed with the computer program JANA2000.[60]

The ISAM refinement was used to obtain scaled and phased observed structure factors corrected for anomalous scattering, according to a procedure given by Bagautdinov *et al.*.[75] These data were used in the F-constraint of the MEM.[51] The ISAM provides coordinates and temperature parameters for all atoms, which were used to compute the procrystal electron density with the computer program PRIOR.[45] This density, $\rho^{prior}(\mathbf{x})$, is used as reference electron density or prior in the MEM.

The entropy S of an electron density $\rho(\mathbf{x})$ is defined as[46]

$$S = -\sum_{k=1}^{N_p} \rho_k \log\left(\rho_k/\rho_k^{prior}\right) , \tag{4.1}$$

where $\rho_k = \rho(\mathbf{x}_k)$ are the values of the electron density on a grid of $N_p = N_1 \times N_2 \times N_3$ points over the unit cell. $\rho_k^{prior} = \rho^{prior}(\mathbf{x}_k)$ are the corresponding values of the prior.

The principle of maximum entropy states that the most probable electron density $\{\rho_k\}$ is the density that maximizes S and simultaneously fits the data. The latter are taken into account through the so-called F-constraint[21]

$$C_{F^2} = -\chi^2_{aim} + \frac{1}{N_F}\sum_{i=1}^{N_F} w_i \left(\frac{|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|}{\sigma(\mathbf{H}_i)}\right)^2 \tag{4.2}$$

where $F_{obs}(\mathbf{H}_i)$ is the phased observed structure factor of the Bragg reflection with scattering vector $\mathbf{H}_i$ and $\sigma(\mathbf{H}_i)$ is its standard uncertainty (s.u.). $F_{MEM}(\mathbf{H}_i)$ is obtained by discrete Fourier transform of the electron density $\{\rho_k\}$. The summation extends over all observed reflections N_F. Static weights

$$w_i = \frac{1}{|\mathbf{H_i}|^n}\left(\frac{1}{N_F}\sum_{i=1}^{N_F}\frac{1}{|\mathbf{H_i}|^n}\right)^{-1} \tag{4.3}$$

have been chosen according to de Vries *et al.*,[52] where n is a small positive integer. Iterations are considered to be converged once C_{F^2} drops below zero for the first

time. The point of convergence thus depends on the value of χ^2_{aim}. The choice of an optimal χ^2_{aim} as well as the choice of appropriate weights [value of n in eqn (4.3)] are discussed in Section 4.3.1. While eqn (4.2) was used to define the stopping criterion of the MEM, the iterative procedure itself has been performed with the summation of eqn (4.2) being extended towards all reflections up to $\sin(\theta)/\lambda = 2.5$ Å^{-1} by the method of prior-derived F-constraints (PDC).[51]

Calculations according to the MEM have been performed with the computer program BayMEM,[45] employing the Cambridge maximum entropy algorithm.[46,58] The electron density and the prior have been defined on a grid of $64 \times 144 \times 72$ pixels over the unit cell, which corresponds to a pixel size of $0.080 \times 0.082 \times 0.076$ Å^3. Electron density maps were analyzed according to Bader's atoms in molecules (AIM) theory[13] with the module EDMA of BayMEM, thus extracting the local maxima of the density, atomic basins, atomic charges and bond critical points (BCPs) with their densities ρ_{bcp} and eigenvalues λ_i^{bcp} $(i = 1, 2, 3)$ of the Hessian matrix.

4.3 Results

4.3.1 Determination of parameters

MEM calculations according to the procedures described in section 4.2 result in MEM-optimized electron densities, $\rho^{MEM}(\mathbf{x})$, that still depend on two parameters: the value of χ^2_{aim} [eqn (4.2)] and the value of n [eqn (4.3)]. The historical MEM employs $\chi^2_{aim} = 1$ and $n = 0$, reflecting the expectation value of χ^2 being equal to N_F for weights proportional to the inverse square of the s.u.'s [eqn (4.2)]. It has been shown that this estimate of χ^2 is too pessimistic and leads to underfitted data, while $n = 0$ usually leads to a distribution of values of

$$\Delta F(\mathbf{H}_i)/\sigma_i = \frac{1}{\sigma_i} \left[F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)\right] \tag{4.4}$$

that is far from the required Gaussian distribution.[52,46] The classical MEM is based on an alternative stopping criterion for the iterative procedure that is based on both the constraint and the entropy.[46] It is properly defined only for the case of $n = 0$ and an F-constraint based on the data, *i.e.* it cannot be combined with PDC.

In a first approach we have performed several runs of the MEM based on the F-constraint without PDC and with values of n equal to three (H3 weights), four

(H4) and five (H5), respectively. Iterations were stopped according to the criterion of the classical MEM approach.[46] The value of χ^2 at convergence can be computed for each run. Divided by N_F it provides an effective value for χ^2_{aim} that would have led to the same density as the classical MEM in a MEM procedure with χ^2_{aim} and eqn (4.2) as stopping criterion. Effective values of 0.1762, 0.3131 and 0.5504 for χ^2_{aim} have been obtained by the classical MEM with weights H3, H4 and H5, respectively. These values of χ^2_{aim} were then used in MEM calculations with χ^2_{aim} and eqn (4.2) as stopping criterion, now with an F-constraint including PDC. The quality of the resulting maps and the fit to the data were analysed by inspection of the difference maps $\rho^{MEM}(\mathbf{x}) - \rho^{prior}(\mathbf{x})$ and difference Fourier maps of $\Delta F(\mathbf{H}_i)$ (Fig. 4.2).

It appears that the calculation with H5 and $\chi^2_{aim} = 0.5504$ has stopped too early, because the difference Fourier map is structured in the region of the atoms (Fig. 4.2e). The calculation with H3 and $\chi^2_{aim} = 0.1762$ results in a difference Fourier map without any structure, but the contours in the difference density are not perfectly smooth anymore, thus suggesting that some noise of the data has been fitted. Finally, the calculation with H4 and $\chi^2_{aim} = 0.3131$ provides both a featureless difference Fourier map and a difference density that is smooth. For the present compound and data it is thus found that the combination of weights H4 and a stopping criterion defined by $\chi^2_{aim} = 0.3131$ represents the optimal fit to the data.

This result does not allow the conclusion that weights other than H4 are inappropriate, because other weights might lead to the correct $\rho^{MEM}(\mathbf{x})$ for values of χ^2_{aim} different from the values determined by the classical MEM. Therefore, we have performed a series of MEM calculations for all combinations of χ^2_{aim} equal to 0.2, 0.3, 0.4, 0.5 and 1.0 and weights H0, H3, H4 and H5.

A first evaluation of the quality of the fits to the data is provided by the histograms of the residuals [eqn (4.4)]. For $\chi^2_{aim} = 1$ (represents the historical MEM) and weights H0 an unfavorable distribution of residuals is obtained with a few outliers of large values of $|\Delta F/\sigma|$, in accordance with the discussion in the literature (Fig. 4.3a).[52] MEM runs with weights H3, H4 or H5 result in distributions of residuals much closer to the Gaussian distribution than MEM runs with H0 do, but notable differences are still present. The Gaussian distribution is much better approximated for MEM calculations with $\chi^2_{aim} = 0.3$ (Fig. 4.3b). Weights H0 still have outliers but they are less severe than for the calculation with $\chi^2_{aim} = 1$. Outliers are not present in MEM calculations with weights H3, H4 and H5, while only minor

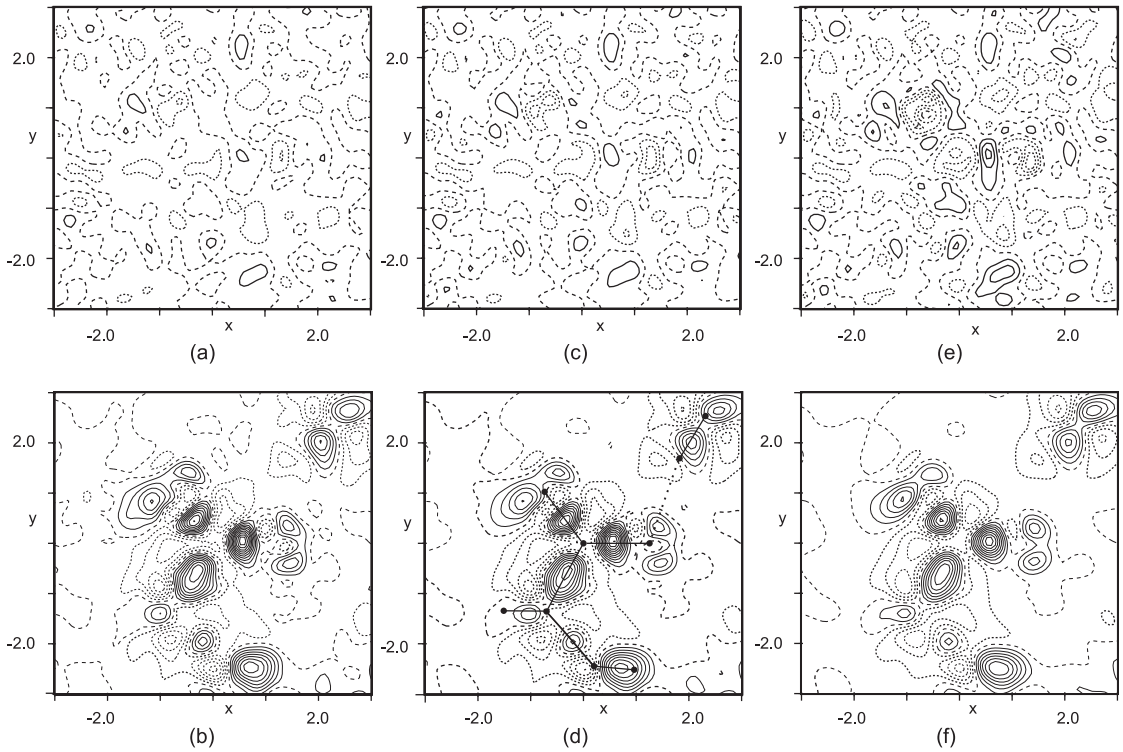

Figure 4.2: Sections of area 6×6 Å^2 through the C(1)–O(1)–O(2) plane of difference Fourier maps and difference maps $\rho^{MEM} - \rho^{prior}$ for selected MEM calculations. (a) Difference Fourier map for weights H3 and $\chi^2_{aim} = 0.1762$. (b) Difference map with $\Delta\rho(min/max) = -0.37/0.56$ electrons/Å^3. (c) Difference Fourier map for weights H4 and $\chi^2_{aim} = 0.3131$. (d) Difference map with $\Delta\rho(min/max) = -0.36/0.56$ electrons/Å^3. (e) Difference Fourier map for weights H5 and $\chi^2_{aim} = 0.5504$. (f) Difference map with $\Delta\rho(min/max) = -0.33/0.54$ electrons/Å^3. Contour lines are at intervals of 0.05 electrons/Å^3. Solid lines are contours of positive value, dotted lines are negative contours, and dashed lines represent the contour of zero value.

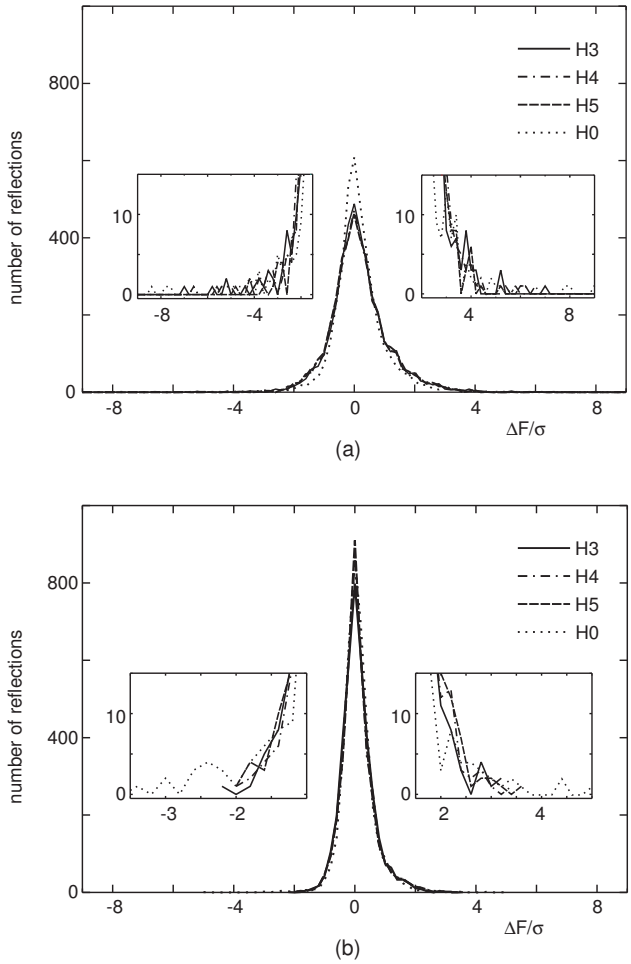

Figure 4.3: Distribution of residuals $\Delta F(\mathbf{H}_i)/\sigma_i$ [eqn (4.4)] for weights H3, H4, H5 and H0. (a) $\chi^2_{aim} = 1$. (b) $\chi^2_{aim} = 0.3$. The insets show magnifications of the outer regions of the curvatures.

differences are found between histograms obtained for MEM calculations with the latter three weights. We have chosen H4 as weights for the final calculations.

Difference Fourier maps and $\rho^{MEM}(\mathbf{x})$ have been analyzed for MEM calculations with weights H4 and stopping criteria provided by $\chi^2_{aim} = 0.2$, 0.3131, 0.5 and 1.0 (Figs. 4.2c and 4.4). Maps of similar appearances have been obtained for H3 and H5. The maps for different values of χ^2_{aim} confirm the observations made in Fig. 4.2. Too high values of χ^2_{aim} lead to underfitted data as represented by structured difference Fourier maps. Too low values of χ^2_{aim} lead to fitting of the noise in the data as represented by very flat difference Fourier maps and a noisy appearance of the contour lines in $\rho^{MEM}(\mathbf{x})$ and $\Delta\rho^{MEM}(\mathbf{x})$. The optimal value of χ^2_{aim} is approximately 0.3. We have chosen the MEM electron density obtained with $\chi^2_{aim} = 0.3131$ and weights H4 for a more detailed analysis by the AIM theory.

A good fit to the data thus requires weights Hn different from 1 ($n > 0$), while the precise value of n is not of large influence. H4 seems to be a good choice. The optimal value of χ^2_{aim} can be obtained by inspection of difference Fourier maps and $\rho^{MEM}(\mathbf{x})$ or $\Delta\rho^{MEM}(\mathbf{x})$. These results are in accordance with the results obtained for trialanine.[42] Weights H4 are also the weights recommended by de Vries *et al.*[52]

4.3.2 Phases of the Bragg reflections

The F-constraint incorporates the observed structure factor amplitudes combined with reflection phases from the calculated structure factors of the ISAM. The MEM determines an electron density that fits both the amplitudes and phases and the MEM thus implicitly assumes that the difference between structure factors of a good structure model ($R_F = 0.023$ for the ISAM) and the true structure factors is mainly in their amplitudes and not in their phases. This assumption can be tested by comparing $F_{MEM}(\mathbf{H}_i)$ with $F^{ISAM}_{cal}(\mathbf{H}_i)$ and $F^{mult}_{cal}(\mathbf{H}_i)$. The latter indicate the structure factors of the multipole model[74] that are supposed to represent the true structure factors. Since α-glycine is centrosymmetric the structure factors corrected for anomalous scattering have phases 0 and π and the implicit assumption sketched above can be tested by counting the number of reflections for which the phases are different between any pair of models.

It turns out that the ISAM and the multipole model lead to different phases for 17 reflections of which only two reflections are of the type observed (Table 4.2). The MEM is able to flip the phases of five reflections of the type less than, which

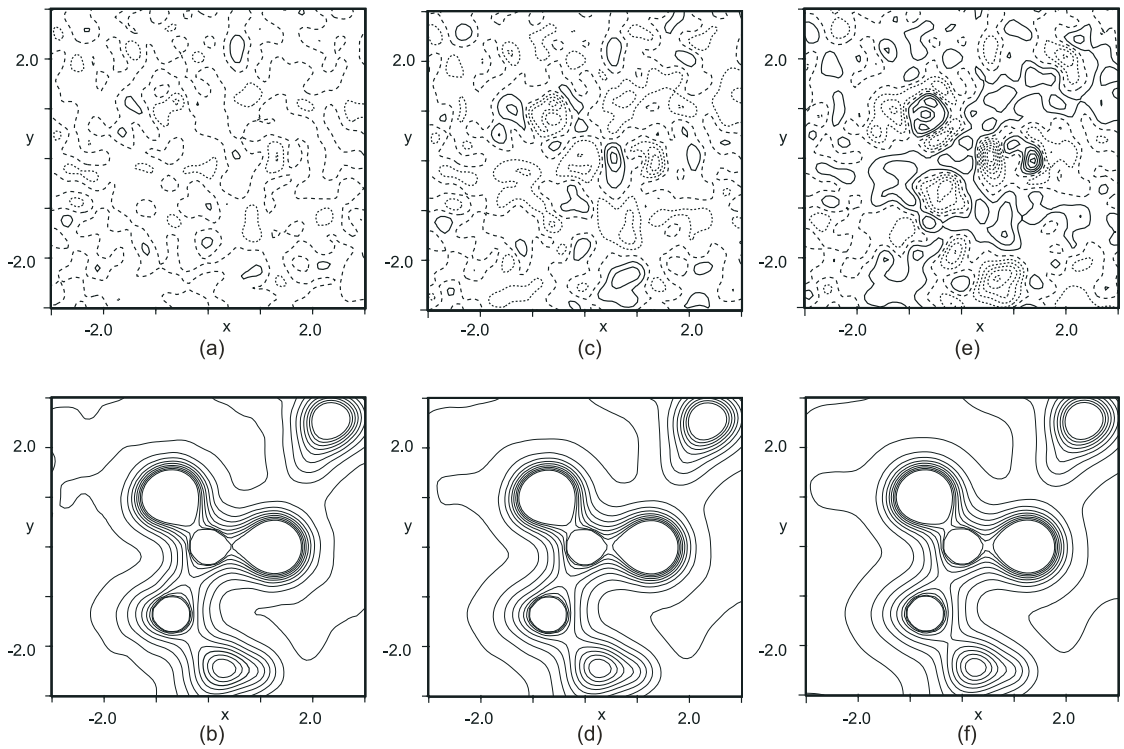

Figure 4.4: Sections of area 6×6 Å^2 through the C(1)–O(1)–O(2) plane of difference Fourier maps and ρ^{MEM} for selected MEM calculations with weights H4. (a) Difference Fourier map for $\chi^2_{aim} = 0.2$. (b) $\rho^{MEM}(\mathbf{x})$. (c) Difference Fourier map for $\chi^2_{aim} = 0.5$. (d) $\rho^{MEM}(\mathbf{x})$. (e) Difference Fourier map for $\chi^2_{aim} = 1.0$. (f) $\rho^{MEM}(\mathbf{x})$. Contour lines are at intervals of 0.05 electrons/Å^3 for the difference Fourier maps and at intervals of 0.3 electrons/Å^3 from 0.1 to 2.5 electrons/Å^3 for $\rho^{MEM}(\mathbf{x})$. Solid lines are contours of positive value, dotted lines are negative contours, and dashed lines represent the contour of zero value.

Table 4.2: Numbers of structure factors with different phases for several pairs of models. The total number of reflections is 3822. The number in square brackets indicates the number of observed reflections ($I/\sigma_I > 3$) with different phases. ISAM: independent spherical atom model; MP: multipole model; MEM: calculation with weights H4 and $\chi^2_{aim} = 0.3131$; MP-MEM: MEM calculation with phases from the MP

	ISAM	MEM	MP	MP-MEM
ISAM	0	5 [0]	17 [2]	17 [2]
MEM		0	20 [2]	12 [2]
MP			0	10 [0]
MP-MEM				0

allows the conclusion that the MEM basically keeps the reflections phases as they are present in the F-constraint. For further tests we have performed an additional MEM calculation with H4 and $\chi^2_{aim} = 0.3131$, in which the phases of the multipole refinement have been used in the F-constraint (MP-MEM model). In this case the MEM is able to flip the phase of 10 reflections of type less than (Table 4.2). This larger number of reflections with flipped phases is probably due to the fact that the PRIOR was based on the ISAM, while the reflection phases correspond to the multipole model. At present we do not have software to compute a dynamic density map for a multipole model, which would have been the appropriate choice in the latter case. The small influence of the choice of phases on the results is confirmed by inspection of the MEM and MP-MEM electron densities. The difference $\rho^{MP-MEM}(\mathbf{x})$-$\rho^{MEM}(\mathbf{x})$ has maximum and minimum values of 0.056 and -0.070 electrons/Å^3, respectively (Fig. 4.5). Differences between the MEM and MP-MEM densities at the BCPs are even smaller (Table 4.3). In view of the accuracy of the methods these differences seem to be immaterial. It can be concluded that the ISAM describes the reflection phases with sufficient accuracy, while it misses part of the amplitudes of the reflections.

4.4 Discussion

The electron density $\rho^{MEM}(\mathbf{x})$ as obtained with H4 and $\chi^2_{aim} = 0.3131$ has been analyzed according to Bader's AIM theory.[13] Local maxima are identified with

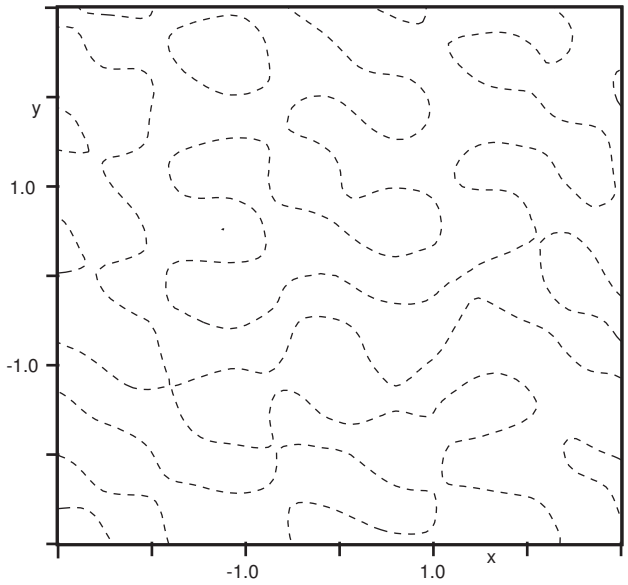

Figure 4.5: Section of area 6×6 Å^2 through the C(1)–O(1)–O(2) plane of the difference ($\rho^{MP-MEM} - \rho^{MEM}$). Contour intervals at 0.05 electrons/Å^3. $\Delta\rho(min)/\Delta\rho(max)$= $-0.070/0.056$ electrons/Å^3. Dashed lines represent the contour of zero value.

Table 4.3: Electron densities and Laplacians at the BCPs of covalent bonds. Values are given for ρ_{bcp} (electrons/Å^3; first line) and $\nabla^2\rho_{bcp}$ (electrons/Å^5; second line) derived from $\rho^{prior}(\mathbf{x})$, $\rho^{MEM}(\mathbf{x})$, the static electron density of the multipole model[74] and $\rho^{MP-MEM}(\mathbf{x})$ as well as $\rho^{MEM}(\mathbf{x})$ of trialanine[42]

Bond	Prior	MEM	Multipole	MP-MEM	Trialanine
Short C_{carbox}–O	2.05	2.49	2.77	2.52	2.47
	9.52	9.09	-32.8	6.59	23
Long C_{carbox}–O	2.02	2.34	2.67	2.37	2.44
	2.61	11.19	-30.5	9.30	10
C_α–C_{carbox}	1.19	1.55	1.78	1.57	1.48
	-0.24	-12.87	-15.6	-13.16	-9
C_α–N_{amino}	1.40	1.50	1.69	1.51	1.67
	2.14	-6.27	-11.9	-6.40	-8
C_α–H	1.21	1.70	1.99	1.69	–
	-6.01	-27.42	-22.7	-26.66	–
C_α–H	1.20	1.61	1.91	1.62	–
	-4.40	-19.10	-21.2	-19.66	–

atomic positions (Table 4.4). For non-hydrogen atoms the positions obtained from $\rho^{MEM}(\mathbf{x})$ lie within 0.005 Å from the positions in the multipole model, while atomic positions in the multipole model and in the ISAM differ by up to 0.001 Å. Both differences are small, whereby the differences between positions in the structure model and positions derived from $\rho^{MEM}(\mathbf{x})$ are fully explained by the accuracy of the positions derived from $\rho^{MEM}(\mathbf{x})$. It was previously shown that the accuracy by which maxima in the electron density can be identified is better than 10% of the pixel size.[45] The presently observed differences of atomic positions lie well within this limit of ∼0.008 Å. The positions of hydrogen atoms are less accurate (Table 4.4). Most noteworthy is that $\rho^{MEM}(\mathbf{x})$ does not exhibit local maxima corresponding to the three hydrogen atoms bonded to the nitrogen atom (atoms H(1), H(2) and H(3)). It has previously been shown that this feature is a property of dynamic density maps as opposed to static density maps that always show local maxima for all atoms.[42]

Volumes of the atomic basins add up to 302.72 Å^3 that is only 0.15% smaller than the volume of the unit cell (Table 4.1). This indicates that all density has been captured by the atomic basins as listed in Table 4.5. Since hydrogen atoms H(1), H(2) and H(3) do not constitute local maxima in $\rho^{MEM}(\mathbf{x})$ their volumes and electrons are incorporated into the basin of the nitrogen atom which they are covalently bonded to. We thus obtain only the net charge for the NH_3 group and not for the individual atoms N, H(1), H(2) and H(3). Net atomic charges differ substantially from the charges derived from the monopole populations in the multipole model,[74] in accordance with the expectation that the AIM analysis of the electron density and monopole charges of the multipole model lead to different atomic charges. The MEM electron density and the static electron density calculated from the multipole model have similar appearances (Fig. 4.6). Differences between these two densities at the bond critical points (BCPs) are of similar magnitude as has previously been determined for trialanine (Table 4.3).[42] For covalent bonds between non-hydrogen atoms the MEM density has values of ρ_{bcp} that are smaller than the values derived from the multipole model by, on the average, 12.8%. This discrepancy can be explained partly by the fact that the MEM produces dynamic densities, while the multipole model is usually analyzed through its static density, and partly because MEM densities will be slightly too low because not all aspects of the data have been fitted at convergence, and partly because the multipole model has been found to produce values of electron densities that are larger than the corresponding

Table 4.4: Atomic positions derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms. $|\Delta\mathbf{x}^{MEM-MP}|$ is the distance (Å) between atomic positions derived from the MEM and those in the multipole model (MP). $|\Delta\mathbf{x}^{MEM-ISAM}|$ is the distance (Å) between atomic positions from the MEM and in the ISAM. $|\Delta\mathbf{x}^{ISAM-MP}|$ is the distance (Å) between corresponding positions in the ISAM and the MP

Atom	x	y	z	$\lvert\Delta\mathbf{x}^{MEM-MP}\rvert$	$\lvert\Delta\mathbf{x}^{MEM-ISAM}\rvert$	$\lvert\Delta\mathbf{x}^{ISAM-MP}\rvert$
C(1)	0.069585	0.125102	0.066349	2.95×10^{-3}	2.86×10^{-3}	0.44×10^{-3}
C(2)	0.058968	0.145918	0.786756	5.49×10^{-3}	5.08×10^{-3}	0.51×10^{-3}
N	0.295218	0.088595	0.740114	3.67×10^{-3}	3.78×10^{-3}	0.31×10^{-3}
O(1)	0.301733	0.093705	0.236264	2.31×10^{-3}	3.22×10^{-3}	0.93×10^{-3}
O(2)	0.844486	0.142321	0.107005	3.01×10^{-3}	2.05×10^{-3}	0.98×10^{-3}
H(4)	0.067969	0.229047	0.758047	8.85×10^{-2}	9.53×10^{-3}	3.92×10^{-2}
H(5)	0.871272	0.116960	0.651328	7.29×10^{-2}	7.36×10^{-3}	1.91×10^{-2}

Table 4.5: Volumes of the atomic basins and the number of electrons in the atomic basins and the net atomic charge (Q) for $\rho^{MEM}(\mathbf{x})$. The basin of nitrogen encompasses the hydrogen atoms H(1), H(2) and H(3) because the latter atoms do not constitute local maxima in the density

Atom	Volume	Electrons	Q
C(1)	4.17	6.25	−0.25
C(2)	10.69	4.37	+1.63
NH_3	20.68	9.72	+0.28
O(1)	15.68	9.16	−1.16
O(2)	15.46	9.09	−1.09
H(4)	2.00	0.40	+0.60
H(5)	7.00	1.01	−0.01
Total	75.68	40.00	0.00

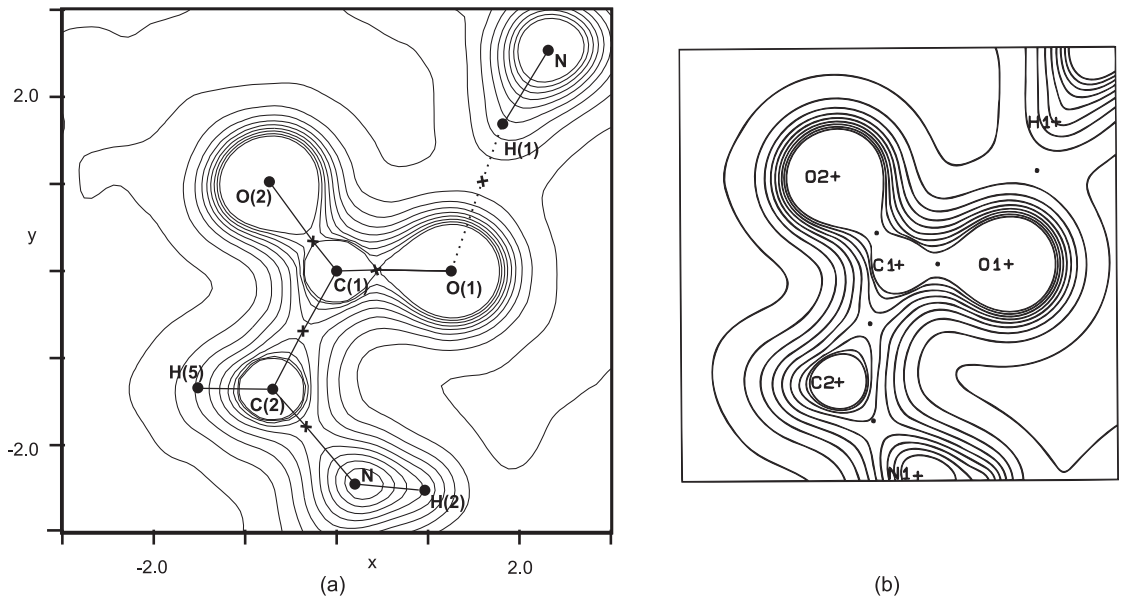

Figure 4.6: Section through the C(1)–O(1)–O(2) plane of the electron density of α-glycine. (a) Section of area 6×6 Å^2 through ρ^{MEM}; crosses indicate bond critical points. (b) Section of area 5×5 Å^2 through the static electron density corresponding to the multipole model (reprinted with permission from Destro *et al.*[74]); dots indicate bond critical points. Contours of equal density are given from 0.1 to 2.5 electrons/Å^3 in steps of 0.3 electrons/Å^3.

Table 4.6: Electron densities and Laplacians at the BCPs of hydrogen bonds. Values are given for ρ_{bcp} (electrons/Å^3; first line) and $\nabla^2\rho_{bcp}$ (electrons/Å^5; second line) derived from $\rho^{prior}(\mathbf{x})$, $\rho^{MEM}(\mathbf{x})$, the static electron density of the multipole model,[74] and $\rho^{MP-MEM}(\mathbf{x})$. BCPs of O(1)$^3\cdots$H(3)^2–N^2 and O(1)$^6\cdots$H(4)^1–C(2)^1 were not found in Prior density map. See Fig. 4.1 for the definition of the symmetry operators

Bond	Prior	MEM	Multipole	MP-MEM
O(1)$^6\cdots$H(1)^4–N^4	0.334	0.303	0.286	0.291
	2.69	2.98	2.47	3.37
O(2)$^1\cdots$H(2)^3–N^3	0.283	0.249	0.215	0.257
	3.12	3.94	2.25	2.70
O(2)$^1\cdots$H(3)^2–N^2	0.191	0.201	0.125	0.198
	1.97	0.90	1.59	1.04
O(1)$^3\cdots$H(3)^2–N^2	–	0.108	0.057	0.103
	–	-1.67	1.18	0.42
O(1)$^6\cdots$H(4)^1–C(2)^1	–	0.129	0.059	0.112
	–	-0.19	0.96	0.52
O(2)$^5\cdots$H(4)^1–C(2)^1	0.098	0.124	0.067	0.120
	1.12	-1.31	1.07	-0.94

values in theoretical densities.[42] Unfortunately, an appropriate quantum mechanical calculation of zwitterionic α-glycine is not available to us, so that a comparison to theoretical densities cannot be made.

Bonds of similar character in α-glycine and trialanine, respectively, have similar values of ρ_{bcp}, thus showing the consistency of the MEM approach (Table 4.3).

Three hydrogen bonds have been identified by their BCPs in $\rho^{MEM}(\mathbf{x})$, in accordance with the results of the multipole refinements (Fig. 4.1).[74] Three more O$\cdots$H contacts lead to BCPs in $\rho^{MEM}(\mathbf{x})$ (Table 4.6). They have been classified by Destro *et al.*[74] as short intermolecular contacts, but they can also be considered as very weak hydrogen bonds. Electron densities at BCPs of hydrogen bonds are on average 62.43 % larger than the corresponding density values obtained from the multipole model. This result suggests—as noticed before for the case of trialanine[42]—that the MEM provides a more accurate description of the electron density in hydrogen bonds than it is provided by the multipole model. This interpretation follows from known problems of the multipole model in describing electron densities around

hydrogen atoms due to inflexibility of the model parameters,[3,4,5,6,7] while the MEM can reproduce any density distribution without restrictions imposed by a model.

At the BCPs the Hessian matrix possesses two negative eigenvalues, corresponding to principal directions in the plane perpendicular to the bond, and one positive eigenvalue, corresponding to the direction approximately parallel to the bond (Table 4.7). Both features are in agreement with the general properties of covalent bonds.[13] A fair agreement is found between the numerical values of the negative eigenvalues of $\nabla^2\rho_{bcp}^{MEM}$ and the corresponding values for the electron density obtained from the multipole model, while differences can be attributed to the effects of thermal motion as they are present in MEM densities but absent in static densities of the multipole model.[42] Large discrepancies between the two densities are found for the positive eigenvalues of the Hessian matrix at the BCPs of the C–O bonds. In turn this leads to large discrepancies of the values of the Laplacian between the two densities (Table 4.3). These differences between MEM and multipole model densities can only be explained by differences between the two types of densities that go beyond the effects of thermal smearing. In this respect it is interesting to note that large discrepancies have been found between the Laplacians at the BCPs of polar bonds (like C–O bonds) as obtained by the multipole model and by quantum theoretical calculations.[68] While there is an established way of interpreting the Laplacian of static densities, this is not yet available for dynamic densities—even if the dynamic character pertains to zero-point vibrations only—therefore we refrain from a detailed interpretation of the Laplacian.

4.5 Conclusions

The accurate electron density of α-glycine has been successfully determined by the maximum entropy method (MEM). The stopping criterion of the MEM iterations is provided by inspection of the difference Fourier maps, in accordance with the analysis on trialanine.[42] At convergence they are free of features—thus ensuring that all information in the data has been captured by $\rho^{MEM}(\mathbf{x})$—while they will retain a noisy appearance, so that $\rho^{MEM}(\mathbf{x})$ does not incorporate the effects of noise in the data. The latter property is also visible in difference maps, $\Delta\rho(\mathbf{x})$, that will adopt noisy contour lines as soon as noise of the data have been fitted. Numerically the stopping criterion is summarized by the value of the F-constraint at convergence,

Table 4.7: Equivalent eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the Hessian matrix at BCPs and the bond ellipticities ϵ of multipole refinement and the MEM. No BCPs of N–H bonds were located, therefore no eigenvalues available

	Multipole				MEM			
Bond	λ_1	λ_2	λ_3	ϵ	λ_1	λ_2	λ_3	ϵ
C(1)–O(1)	-26.93	-25.09	19.18	0.07	-22.48	-21.29	52.86	0.06
C(1)–O(2)	-25.51	-22.60	17.66	0.13	-21.08	-17.22	49.50	0.22
C(1)–C(2)	-13.55	-12.17	10.10	0.11	-12.83	-10.61	10.58	0.21
C(2)–N	-12.27	-11.77	12.18	0.04	-11.01	-8.34	13.08	0.32
C(2)–H(4)	-19.78	-19.37	16.50	0.02	-18.12	-15.71	6.41	0.15
C(2)–H(5)	-18.68	-17.98	15.43	0.04	-15.42	-11.34	7.66	0.36
Hydrogen Bonds								
$O(1)^6 \cdots H(1)^4$–N^4	-1.91	-1.83	6.21	0.04	-1.87	-1.32	6.17	0.42
$O(2)^1 \cdots H(2)^3$–N^3	-1.49	-1.18	4.92	0.26	-1.93	-0.31	6.18	5.15
$O(2)^1 \cdots H(3)^2$–N^2	-0.66	-0.60	2.85	0.10	-1.39	-1.02	3.30	0.37
$O(1)^3 \cdots H(3)^2$–N^2	-0.24	-0.21	1.63	0.09	-1.02	-1.01	0.36	0.01
$O(1)^6 \cdots H(4)^1$–$C(2)^1$	-0.24	-0.18	1.37	0.33	-0.87	-0.55	1.23	0.58
$O(2)^5 \cdots H(4)^1$–$C(2)^1$	-0.27	-0.23	1.57	0.17	-1.49	-0.95	1.14	0.57

χ^2_{aim}.

The MEM electron density is determined by the weights in the F-constraint that in turn depend on the s.u.'s of the data (eqn (4.2)). Weights proportional to $1/\sigma_F^2$ (weights of type H0) lead to an unfavorable point of convergence of the MEM and de Vries *et al.*[52] have proposed a weighting scheme H4 ($n = 4$ in eqn (4.3)) as optimal choice. In accordance with the study on trialanine[42] we have presently found that weights H4 are a good choice indeed, but that the precise choice of weights (H3, H4 or H5) does not influence the result very much, if convergence is defined according to the appropriate value of χ^2_{aim}. A favorable property of the stopping criterion based on inspection of difference Fourier maps is that it automatically corrects for an error of scale in the estimated s.u.'s of the data.[42]

For the present data we have found that the optimal value of χ^2_{aim} coincides with the value of the F-constraint that is obtained by the stopping criterion of the classical MEM[46] in case of weights H4. We consider this a fortuitous coincidence of stopping criteria, however this result might also be taken as an argument in favor of H4 as opposed to other choices of the weights. Reflection phases obtained by the ISAM have been found to be a sufficiently good approximation to the true reflection phases, with a wrong assignment for only two reflections of the type 'observed'.

The MEM electron density of α-glycine compares well with the MEM electron density of trialanine.[42] Discrepancies between ρ_{bcp} of similar bonds in the two compounds can be explained by the different environments of the ammonium groups and carboxylic oxygen atoms. In trialanine they form hydrogen bonds with water molecules,[14] while in water-free α-glycine hydrogen bonds occur between oxygen atoms and the NH_3 group or the C–H group (Table 4.6). Discrepancies between $\rho^{MEM}(\mathbf{x})$ and the static multipole densities are similar for α-glycine and trialanine,[42] with the specific features of both methods as origin for these differences.

Substantial differences between MEM electron densities and static multipole densities have been found for the values of the second derivatives at the BCPs of the electron densities, again in accordance with the results for trialanine.[42] The most important difference concerns the principal value of the Hessian matrix for the directions along chemical bonds, which in turn leads to large variations of the Laplacian. These discrepancies are in first approximation due to the effect of thermal smearing as it is present in the dynamic densities produced by the MEM and that is absent in static densities corresponding to the multipole model. At a temperature of

23 K most of the thermal motion will be zero point vibrations that cannot be avoided at any temperature. It is therefore an open question whether correlations between physical and chemical properties on the one hand and properties of electron densities on the other hand can better be obtained from static electron densities or from dynamic electron densities that at least incorporate the zero point vibrational effects. Unfortunately, lack of reference values prevent the interpretation of the Laplacians of dynamic densities at present.

Chapter 5

Joined analysis of topological properties of hydrogen bonds and covalent bonds from accurate charge density studies by the maximum entropy method

Abstract

Accurate Charge Densities have been determined by the Maximum Entropy Method (MEM) from low-temperature X-ray diffraction data of six different crystals of amino acids and peptides. A comparison of dynamic deformation densities of the MEM with static deformation densities of multipole models shows that the MEM leads to a better description of the electron density in hydrogen bonds. Topological properties at bond critical points (BCPs) have been analyzed according to Bader's Atoms in Molecules (AIM) theory. They are found to depend systematically on the bond length, but with different functions for covalent C–C, C–N and C–O bonds and for hydrogen bonds together with covalent C–H and N–H bonds. Similar dependencies are known for AIM properties derived from static multipole densities, albeit with different functions than presently found for dynamic MEM densities. For covalent bonds they indicate that, in general, the electron density at the BCP is

largest for multipole densities and smallest for procrystal densities, while the reverse is true for hydrogen bonds. The ratio of potential and kinetic energy densities $|V(BCP)|/G(BCP)$ is successfully used for a classification of hydrogen bonds according to their distance $d(\mathrm{H}\cdots\mathrm{O})$ between the hydrogen atom and the acceptor atom. The classification based on MEM densities coincides with the usual classification of hydrogen bonds as strong, intermediate and weak [G.A. Jeffrey: *An introduction to hydrogen bonding*, Oxford University Press (1997)]. Short ($d(\mathrm{H}\cdots\mathrm{O})$ < 1.5 Å), strong hydrogen bonds are found to possess mainly covalent character, hydrogen bonds of intermediate length and intermediate strength have mixed covalent/ionic character, and long ($d(\mathrm{H}\cdots\mathrm{O}) > 2.2$ Å), weak hydrogen bonds are mainly stabilized by ionic interactions. MEM and procrystal densities lead to similar values of the density at BCPs of hydrogen bonds, but differences are shown to prevail, such that it is found that only the Accurate Charge Density, as represented by MEM densities in the present work, can lead towards a correct characterisation of chemical bonding.

5.1 Introduction

Inter- and intramolecular hydrogen bonds are important in both molecular and biological chemistry, because they contribute a large part of the interactions responsible for the conformations and functions of many compounds in those fields. Different approaches and methods have been employed to determine geometrical, topological, energetic and functional properties of hydrogen bonds. Besides spectroscopic methods, like Infrared spectroscopy, X-ray diffraction is an important tool for providing answers to structural questions regarding hydrogen bonds. Koch and Popelier[77] proposed eight criteria that establish the existence of hydrogen bonds. Geometrical, energetic and IR spectroscopic properties were suggested by Jeffrey,[78] thus allowing a classification of strong, medium and weak hydrogen bonds.

A sophisticated approach to analyze topological properties of electron densities is provided by Bader's Atoms in Molecules (AIM) theory.[13] The AIM theory allows the determination of bond critical points (BCPs) and their properties, such as the electron density and its Laplacian, leading to the detection of hydrogen bonds in crystal structures as well as providing a quantitative characterization of the type and strengths of these bonds. Analyses according to the AIM theory[13] of electron

densities of amino acids and peptides have been performed by Destro *et al.*,[74,79] Benabicha *et al.*,[62] Pichon-Pesme *et al.*,[63] Wagner *et al.*,[80] Flaig *et al.*,[81] Scheins *et al.*,[82] Mebs *et al.*,[72] Checinska *et al.*,[83] Rödel *et al.*[14] and Kalinowski *et al.*.[84]

Abramov introduced a method which uses the densities at BCPs and their Laplacians for calculating the kinetic energy densities at BCPs.[85] Employment of the local virial theorem[13] allows the calculation of the potential energy densities at BCPs. These energy densities provide information about the character of the bond analyzed.[85,86,87] Extensive studies of energy densities and topological properties at BCPs of hydrogen bonds have been performed by Espinosa *et al.*.[88,89,90,91]

Experimental charge densities are usually based on the multipole model.[1] Alternatively, they can be determined by the maximum entropy method (MEM).[21,42,73,53,26] MEM electron densities (ρ^{MEM}) have successfully been used to study disorder in crystal structures. Most prominent application has been the determination of the location of the metal atom in endohedral fullerenes.[22] Earlier studies have stressed artifacts in MEM densities, which have magnitudes equal to deformation densities of chemical bonds, and thus would prohibit the use of the MEM in accurate charge density studies.[8,9,10,11,12] These problems have been overcome by a combination of extensions to the MEM, including the use of a procrystal prior density,[10] the use of static weights in the F-constraint,[52] the use of prior-derived F-constraints[51] and the definition of a criterion of convergence for the MEM iterations, which is based on difference Fourier maps.[42] The MEM has the potential to become the method of choice in accurate charge density studies on proteins,[73,26] because the MEM—unlike multipole refinements—does not suffer from correlations between parameters.

The present work reports the analysis of MEM electron densities of several amino acids and peptides. The study includes the analysis of geometrical, topological and energetic properties of all 52 hydrogen bonds that have been identified in these compounds. The quantitative analysis is supplemented by a descriptive analysis of electron densities in the regions of the hydrogen bonds. Since the role of a promolecule (procrystal) has been discussed as being important for the extraction of information of bonding,[92,93] the contribution of the prior density to properties of chemical bonds is discussed. The systematic dependence of properties of hydrogen bonds on the distance between hydrogen atom and acceptor atom is supplemented by an analysis of the properties of covalent bonds with respect to the bond distance.

5.2 Computational details

5.2.1 MEM calculations

Single-crystal x-ray diffraction data of L-alanine (Ala),[79] an L-phenylalanine formic acid complex (Phe),[72] L-alanyl-L-tyrosyl-L-alanine (Ala-Tyr-Ala) with water as solvent[83] and Ala-Tyr-Ala with ethanol as solvent[83] were kindly provided by Destro or by Luger and co-workers, who have already reported multipole refinements against these data (Table 5.1).[79,72,83] We have used these data to perform refinements of the independent spherical atom model (ISAM) with the computer program JANA2000.[60] The coordinates and atomic displacement parameters (ADP) obtained by the multipole refinements were used as starting model for the ISAM refinement. Hydrogen atoms were fixed at distances known from neutron diffraction.[76,72,94,95] A riding model of U_{iso}(H)= 1.2 U_{eq}(N,C) and U_{iso}(H)= 1.5 U_{eq}(O) was employed to calculate the ADPs of hydrogen atoms. For all three water molecules of the asymmetric unit of Ala-Tyr-Ala with water, the angle H–O–H was restrained to 104.5°. The coordinates of H(2b) of Phe, of H(15), H(61), H(62), H(71), H(72), H(81) and H(82) of Ala-Tyr-Ala with water, and of H(15) and H(16) of Ala-Tyr-Ala with ethanol were independently refined subject to constraints on the distances X–H towards values known from neutron diffraction, because a reasonable geometrical restraint was not available for these hydrogen atoms. Coordinates of all other hydrogen atoms were obtained by attachment to their neighbor atoms with tetrahedral or trigonal angle restraints according to chemical meaning. Agreement indices for ISAM refinements are given in Table 5.1.

According to a procedure by Bagautdinov *et al.*,[75] the ISAM refinement was employed to obtain phased and scaled observed structure factors corrected for anomalous scattering, which were used for the MEM calculations. The coordinates and ADPs of the ISAM were used to compute the procrystal electron density [prior density or PRIOR, $\rho^{prior}(\mathbf{x})$] with the computer program PRIOR.[45] The PRIOR was calculated on a grid over the unit cell and used as reference density in the MEM calculations. Equal grids were chosen for the PRIOR and MEM densities, such that the pixel size did not exceed 0.1 × 0.1 × 0.1 Å^3 (Table 5.1).

The MEM is based on the principle that the most probable density $\{\rho_k\}$ is that one, which simultaneously fits the diffraction data and maximizes the informational

entropy S, with

$$S = -\sum_{k=1}^{N_p} \rho_k \log\left(\rho_k/\rho_k^{prior}\right) , \tag{5.1}$$

where $\rho_k = \rho(\mathbf{x}_k)$ are the values of the electron density on a grid over the unit cell of $N_p = N_1 \times N_2 \times N_3$ points. $\rho_k^{prior} = \rho^{prior}(\mathbf{x}_k)$ are the corresponding values of the PRIOR.

Diffraction data are taken into account by the method of undetermined Lagrange multipliers employing the F-constraint $C_{F^2} = 0$ with[21,42]

$$C_{F^2} = -\chi^2_{aim} + \frac{1}{N_F}\sum_{i=1}^{N_F} w_i \left(\frac{|F_{obs}(\mathbf{H}_i) - F_{MEM}(\mathbf{H}_i)|}{\sigma(\mathbf{H}_i)}\right)^2 \tag{5.2}$$

$F_{obs}(\mathbf{H}_i)$ is the phased observed structure factor of the Bragg reflection with scattering vector $\mathbf{H}_i$ and standard uncertainty (s.u.) $\sigma(\mathbf{H}_i)$. $F_{MEM}(\mathbf{H}_i)$ is obtained by discrete Fourier transform of the electron density $\{\rho_k\}$. The summation extends over all observed reflections N_F. Static weights

$$w_i = \frac{1}{|\mathbf{H_i}|^n}\left(\frac{1}{N_F}\sum_{i=1}^{N_F}\frac{1}{|\mathbf{H_i}|^n}\right)^{-1} \tag{5.3}$$

with $n = 4$ have been chosen according to de Vries *et al.*[52] The summation of Eq. 5.2 has been extended towards all reflections up to $\sin(\theta)/\lambda = 2.5$ Å^{-1} according to the method of prior-derived F-constraints (PDC).[51] Because the PDC extends the F-constraint by terms involving calculated structure factors of the PRIOR, its use is recommended only if experimental data are available up to a certain minimum resolution, *e.g.* up to at least $\sin(\theta)/\lambda > 0.9$ Å^{-1}.[51] This condition is fulfilled for all data sets considered in this article.

MEM calculations have been performed with the computer program BayMEM,[45] employing the Cambridge maximum entropy algorithm.[46,49] Convergence of the iterations is considered to be reached once C_{F^2} has dropped below zero (Eq. 5.2), and it thus depends on the value of χ^2_{aim}. Following procedures introduced earlier,[53,42] an optimal value for χ^2_{aim} was determined for each data set by inspection of difference Fourier maps and dynamic deformation maps

$$\rho^{def}_{MEM}(\mathbf{x}) = \rho^{MEM}(\mathbf{x}) - \rho^{prior}(\mathbf{x}) \tag{5.4}$$

for several values of χ^2_{aim} (Table 5.1). Details of the MEM calculations of α-glycine and trialanine have been described elsewhere.[53,42]

An optimal value for χ^2_{aim} is necessary to obtain accurate and reliable electron density maps by the MEM.[73] In theory,[47,46] one would only expect values $\chi^2_{aim} < 1$, but values $\chi^2_{aim} > 1$ can appear if standard uncertainties of measured reflection intensities have been estimated smaller than their true values. The standard uncertainties in turn are related to the Goodness of Fit (GoF) of the refinements, with

$$\mathrm{GoF} = \left(\frac{1}{N_F - P_{refined}} \sum_{i=1}^{N_F} \left[\frac{(|F_{obs}(\mathbf{H}_i)| - |F_{calc}(\mathbf{H}_i)|)}{\sigma(\mathbf{H}_i)} \right]^2 \right)^{\frac{1}{2}} \tag{5.5}$$

where $P_{refined}$ is the number of refined parameters and $F_{calc}(\mathbf{H}_i)$ are the calculated structure factors of the model. For underestimated standard uncertainties, the value of the GoF would be larger than one at convergence of the refinement.

For data sets with multipole refinements resulting in a GoF close to one, we obtained $\chi^2_{aim} < 1$, and for data sets with refinements resulting in a GoF close to two, values of $\chi^2_{aim} > 1$ were obtained (Table 5.1). This indicates that the stopping criterion for the MEM calculation depends on the scale of the standard uncertainties of the intensities. Thus, the accuracy of the standard uncertainties can be estimated from consideration of the value of χ^2_{aim} as determined in the MEM procedure. Values of χ^2_{aim} smaller than one indicate that the standard uncertainties are estimated close to their true values, whereas values of $\chi^2_{aim} > 1$ indicate that the standard uncertainties have been underestimated.

5.2.2 Analysis of the MEM density

The difference between the ISAM and the aspherical electron distribution obtained by the MEM has been analysed by dynamic deformation maps (Eq. 5.4). In particular, sections containing selected atoms allow the visualisation of the difference densities in hydrogen bonds (Fig. 5.1).

Electron density maps have been analysed according to Bader's AIM theory[13] with the module EDMA of the program BayMEM.[45] For each density map, EDMA provides the positions and values of local maxima of the density, the atomic basins, the atomic charges and the positions of bond critical points (BCPs) together with their densities $\rho(BCP)$, their principal curvatures λ_1, λ_2 and λ_3 (eigenvalues of the Hessian matrix) and their Laplacians $\nabla^2\rho(BCP)$. Both the PRIOR and the MEM densities have been analysed in the same way. Covalent bonds have been identified

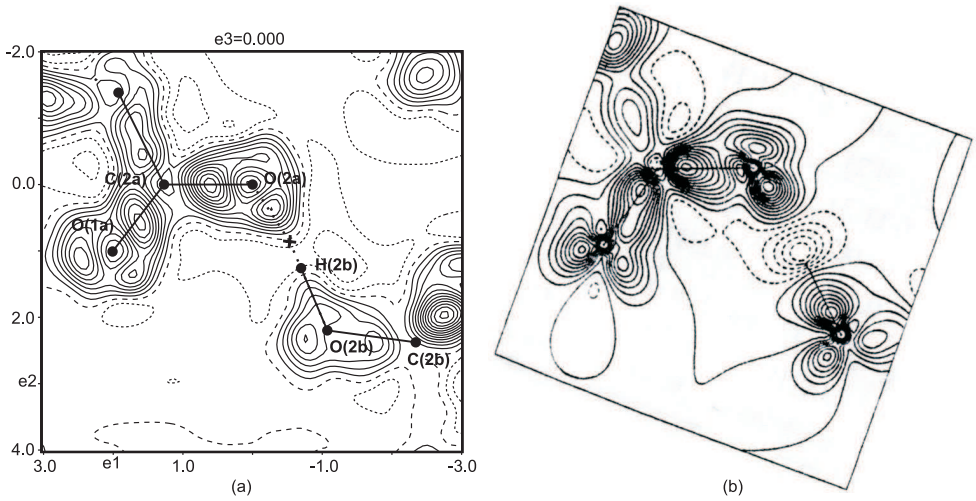

Figure 5.1: Sections of deformation maps containing the atoms O(2a)–C(2a)–O(1a) and showing the hydrogen bond O(2b)–H(2b)$\cdots$O(2a) of the L-phenylalanine formic acid complex. (a) Section of area 6 $\times$ 6 Å^2 of the dynamic deformation density of the MEM (Eq. 5.4). Contour intervals: 0.05 electrons/Å^3. Solid lines indicate positive contours, dotted lines negative contours and dashed lines represent the zero contour. $\Delta\rho(min/max)$ = -0.15/0.61 electrons/Å^3. The cross indicates the BCP of the hydrogen bond, with d(H$\cdots$O)= 1.45 Å, $\rho(BCP) = 0.599$ electrons/Å^3 and $\nabla^2\rho(BCP) = 0.32$ electrons/Å^5. (b) Static deformation density of the multipole model (reprinted with permission from Mebs *et al.*[72]). Contour intervals: 0.10 electrons/Å^3.

by BCPs with values $\rho(BCP)$ larger than $\sim$1.0 electrons/Å^3. A BCP with a smaller value of $\rho(BCP)$ in the region of a potential donor-acceptor pair for hydrogen bonds was used to establish the existence of a hydrogen bond. The positions of BCPs of covalent and hydrogen bonds in MEM densities match BCPs in electrons densities obtained from the multipole model.

The kinetic, potential and total energy densities at BCPs have been obtained from PRIOR and MEM densities according to a procedure proposed by Abramov.[85] The kinetic energy density $G(BCP)$ at a BCP is given by[85]

$$G(BCP) = \frac{3}{10}(3\pi^2)^{\frac{2}{3}}\,\rho(BCP)^{\frac{5}{3}} + \frac{1}{6}\nabla^2\rho(BCP) \tag{5.6}$$

with $\rho(BCP)$ and $\nabla^2\rho(BCP)$ in atomic units. Employing the local virial theorem,[13] the potential energy density $V(BCP)$ at a BCP is

$$V(BCP) = \frac{1}{4}\nabla^2\rho(BCP) - 2G(BCP)\,. \tag{5.7}$$

The total energy density $H(BCP)$ at a BCP then is defined as

$$H(BCP) = G(BCP) + V(BCP)\,. \tag{5.8}$$

It is noticed that Eqs. 5.6 and 5.7 have been derived for static electron densities, *i.e.* within the Born-Oppenheimer approximation. We apply these relations to dynamic densities as obtained by the PRIOR and the MEM. While systematic dependencies of, for example, $G(BCP)$ on $d(\mathrm{H}\cdots\mathrm{O})$ are found (Section 5.3.3), the interpretation of these quantities as kinetic and potential energy densities needs to be established or correction factors need to be found. This is beyond the scope of the present manuscript.

Systematic dependencies on bond lengths have been established for various topological and energetic properties at BCPs of covalent C–C, C–N, C–O, C–H and N–H bonds and of hydrogen bonds. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance $d(\mathrm{H}\cdots\mathrm{O})$.

5.3 Results and discussion

5.3.1 Electron densities in hydrogen bonds

The dynamic deformation map of the MEM (Eq. 5.4) and the static deformation map of the multipole model have similar appearances for the L-phenylalanine formic acid complex (Fig. 5.1). Distinct features, like lone pairs of oxygen atoms and an accumulation of electron density in regions of covalent bonding, are uncovered by both the MEM and the multipole model. However, the hydrogen bond appears differently in these two densities: along the bond path of the hydrogen bond, the MEM leads to a positive difference density [Fig. 5.1(a)], whereas the multipole method exhibits a negative deformation density in this region [Fig. 5.1(b)]. These observations are consistent with those on trialanine and α-glycine.[42,53] They can be interpreted as being due to (i) the differences between dynamic and static densities, and (ii) the known inflexibility of the multipole model in describing densities at remote positions from atomic maxima.[3,4,5,6,7]

Further support for this interpretation comes from the comparison of the dynamic deformation density of the MEM (Eq. 5.4) with an experimental dynamic deformation map that has been computed as the difference Fourier map of $[F_{obs}(\mathbf{H}) - F_{ISAM}(\mathbf{H})]$ whereby phases for $F_{obs}(\mathbf{H})$ have been obtained from a multipole model.[79] The deformation density $\rho^{def}_{MEM}(\mathbf{x})$ (Eq. 5.4) along the bond path of the N–H$\cdots$O hydrogen bond in L-alanine exhibits similar features as $\rho^{def}_{MEM}(\mathbf{x})$ of the O–H$\cdots$O hydrogen bond in L-phenylalanine [Figs. 5.1(a) and 5.2(a)]. The experimental difference Fourier map with phases from a standard multipole model exhibits a density of N–H that is less protruded towards oxygen than $\rho^{def}_{MEM}(\mathbf{x})$ does, while the minimum density along the H$\cdots$O bond path is approx. 0.1 electrons/Å^3 lower than in $\rho^{def}_{MEM}(\mathbf{x})$ [Fig. 5.2(b)]. Phases of an extended multipole model (anisotropic ADPs and multipole parameters up to quadrupole terms for hydrogen) then lead to a dynamic difference Fourier map that is even closer to $\rho^{def}_{MEM}(\mathbf{x})$ [Fig. 5.2(c)].[96] Unfortunately, a dynamic deformation density of the multipole model [difference Fourier map of $F_{calc}(\mathbf{H})$ of the multipole model with $F_{ISAM}(\mathbf{H})$] is not available to us, so that the performance of the extended multipole model cannot be evaluated in this case.

It has been noticed before,[6,98,99] that anisotropic ADPs and higher-order multipole terms of hydrogen atoms are important for a proper description of the electron

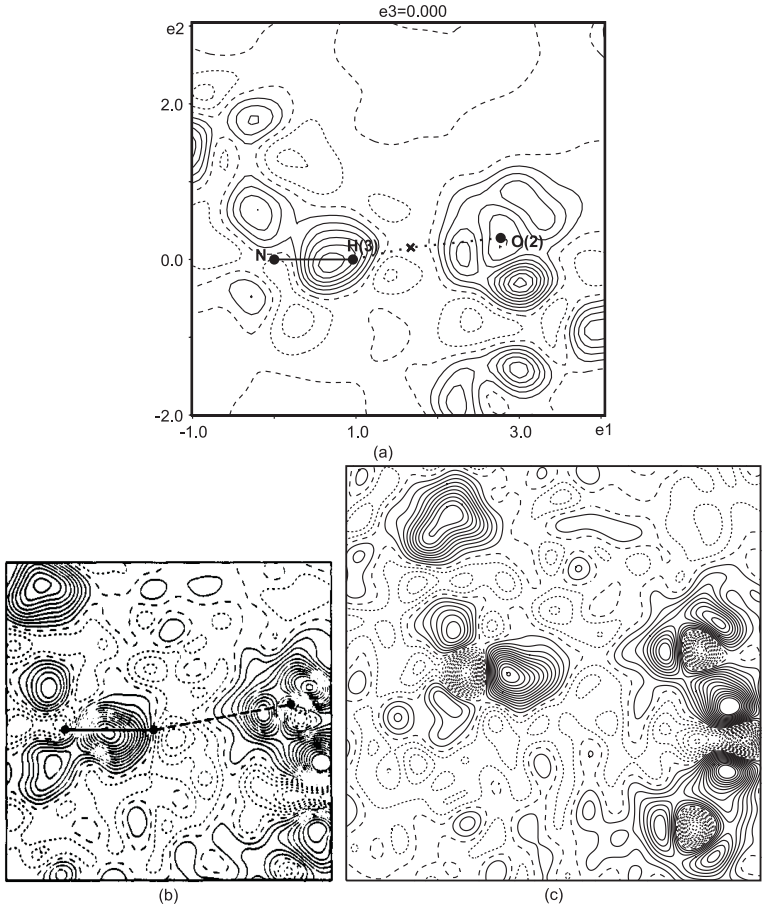

Figure 5.2: Sections of deformation densities containing the atoms N–H(3)–O(2) and showing the hydrogen bond N–H(3)$\cdots$O(2) of L-alanine. (a) Section of area 5×5 Å^2 through the dynamic deformation density of the MEM (Eq. 5.4). $\Delta\rho(min/max)$ = -0.13/0.42 electrons/Å^3. The cross indicates the BCP of the hydrogen bond with d(H$\cdots$O)= 1.76 Å, $\rho(BCP)$ = 0.326 electrons/Å^3 and $\nabla^2\rho(BCP)$ = -1.46 electrons/Å^5. (b) Section of 4×4 Å^2 through the experimental dynamic deformation density (reprinted with permission from Destro *et al.*[79]). (c) Section of 5×5 Å^2 through the experimental dynamic deformation density with phases from the extended multipole model (Destro *et al.*[96,97]). Contour intervals: 0.05 electrons/Å^3; contour lines as defined in Fig. 5.1.

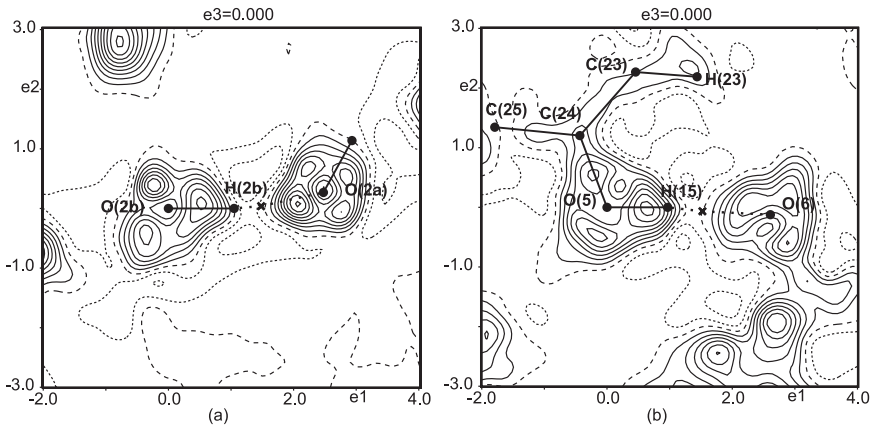

Figure 5.3: Sections of area 6×6 Å^2 of dynamic deformation densities of the MEM (Eq. 5.4), showing hydrogen bonds of the type O–H$\cdots$O. (a) The plane containing the atoms O(2b)–H(2b)–O(2a) of the L-phenylalanine formic acid complex. $\Delta\rho(min/max)$ = -0.15/0.46 electrons/Å^3. This is a different plane containing the same hydrogen bond as displayed in Fig. 5.1(a). (b) The plane containing the atoms O(5)–H(15)–O(6) of Ala-Tyr-Ala with ethanol. $\Delta\rho(min/max)$ = -0.12/0.37 electrons/Å^3. Properties of the hydrogen bond O(5)–H(15)$\cdots$O(6): d(H$\cdots$O)= 1.67 Å, $\rho(BCP)$ = 0.435 electrons/Å^3 and $\nabla^2\rho(BCP)$ = 2.05 electrons/Å^5. Crosses indicate BCPs; contour interval: 0.05 electrons/Å^3; contour lines as defined in Fig. 5.1.

density around hydrogen atoms. However, a refinement of these parameters is not possible for systems substantially larger than simple amino acids, due to the problem of dependent parameters in the multipole model. Usually, the treatment of hydrogen atoms does not go beyond isotropic ADPs[62,63,84,80,100] and dipolar terms within the multipole model.[101,102,83,103] The extended multipole model will thus remain a special feature for crystals of small molecules. On the other hand, the MEM is applicable to both small and large systems, and it leads to a proper description of the deformation density with phases from the ISAM.

Positive dynamic difference densities around the BCPs between the hydrogen atom and the acceptor atom turn out to be a feature of all the three types of hydrogen bonds studied in the present work. For hydrogen bonds of the type O–H$\cdots$O and N–H$\cdots$O, this feature is very pronounced (Figs. 5.1–5.4), whereas this behaviour is less pronounced in hydrogen bonds of the type C–H$\cdots$O (Fig. 5.5). Since hydrogen bonds of the latter type can be considered as very weak or even as non-conventional

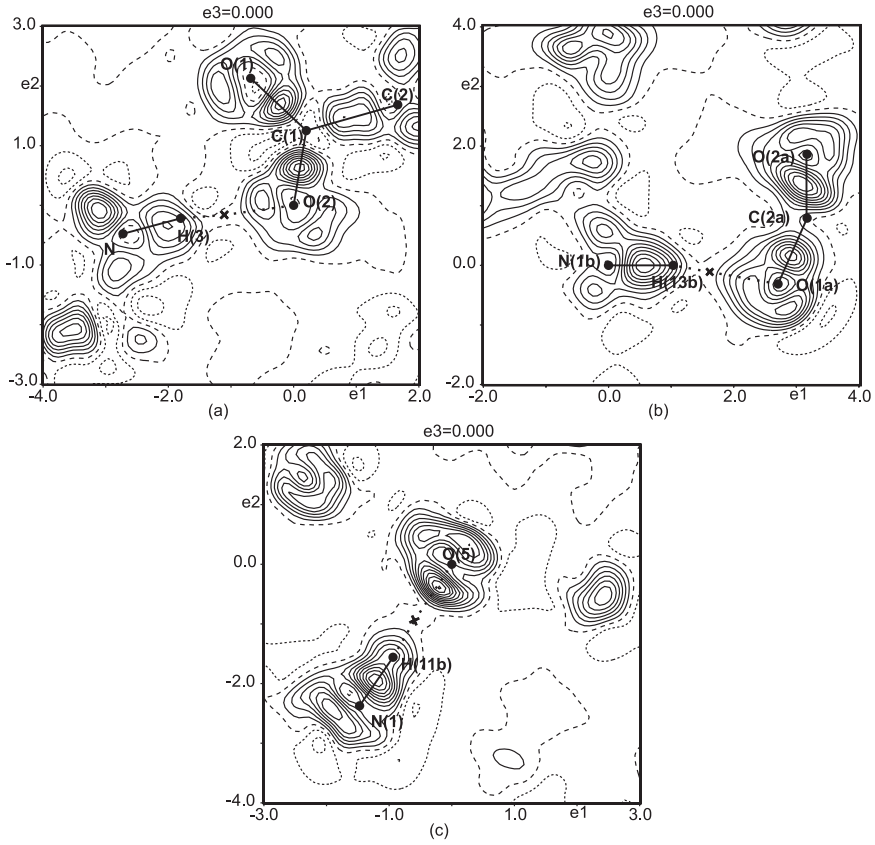

Figure 5.4: Sections of area 6×6 Å^2 of dynamic deformation densities of the MEM (Eq. 5.4), showing hydrogen bonds of the type N–H$\cdots$O. (a) The plane containing the points O(2)–BCP–H(3) of L-alanine. $\Delta\rho(min/max)$ = -0.13/0.43 electrons/Å^3. This is a different plane containing the same hydrogen bond as displayed in Fig. 5.2(a). (b) The plane containing the atoms N(1b)–H(13b)–O(1a) of the L-phenylalanine formic acid complex. $\Delta\rho(min/max)$ = -0.11/0.50 electrons/Å^3. Properties of the hydrogen bond N(1b)–H(13b)$\cdots$O(1a) are: d(H$\cdots$O)= 1.71 Å, $\rho(BCP)$ = 0.402 electrons/Å^3 and $\nabla^2\rho(BCP)$ = 2.82 electrons/Å^5. (c) The plane containing the atoms O(5)–H(11b)–N(1) of Ala-Tyr-Ala with water. $\Delta\rho(min/max)$ = -0.12/0.56 electrons/Å^3. Properties of the hydrogen bond N(1)–H(11b)$\cdots$O(5) are: d(H$\cdots$O)= 1.76 Å, $\rho(BCP)$ = 0.364 electrons/Å^3 and $\nabla^2\rho(BCP)$ = 0.57 electrons/Å^5. Crosses indicate BCPs, contour intervals: 0.05 electrons/Å^3, contour lines as defined in Fig. 5.1.

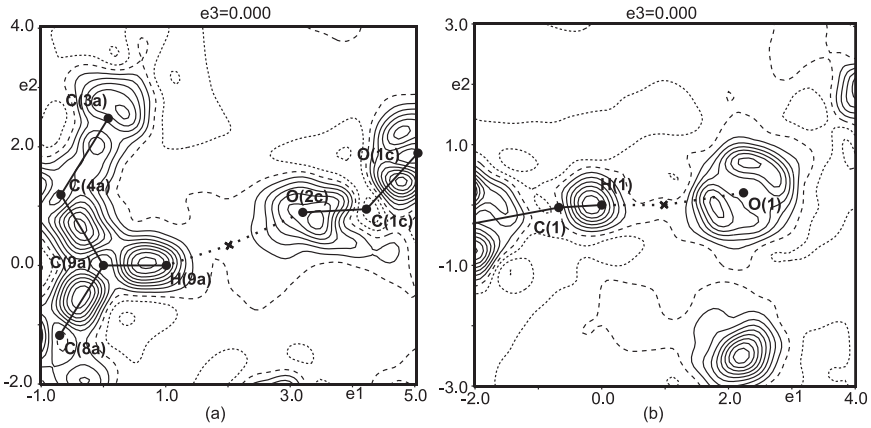

Figure 5.5: Sections of area 6×6 Å^2 of dynamic deformation densities of the MEM (Eq. 5.4), showing hydrogen bonds of the type C–H$\cdots$O. (a) The plane containing the atoms C(9a)–H(9a)–O(2c) of the L-phenylalanine formic acid complex. $\Delta\rho(min/max)$ = -0.13/0.42 electrons/Å^3. Properties of the hydrogen bond C(9a)–H(9a)$\cdots$O(2c) are: d(H$\cdots$O)= 2.36 Å, $\rho(BCP)$ = 0.111 electrons/Å^3 and $\nabla^2\rho(BCP)$ = 0.91 electrons/Å^5. (b) The plane containing the points H(1)–BCP–O(1) of Ala-Tyr-Ala with ethanol. $\Delta\rho(min/max)$ = -0.15/0.43 electrons/Å^3. Properties of the hydrogen bond C(1)–H(1)$\cdots$O(1) are: d(H$\cdots$O) = 2.19 Å, $\rho(BCP)$ = 0.150 electrons/Å^3 and $\nabla^2\rho(BCP)$ = -0.25 electrons/Å^5. Crosses indicate BCPs, contour intervals at 0.05 electrons/Å^3, contour lines as defined in Fig. 5.1.

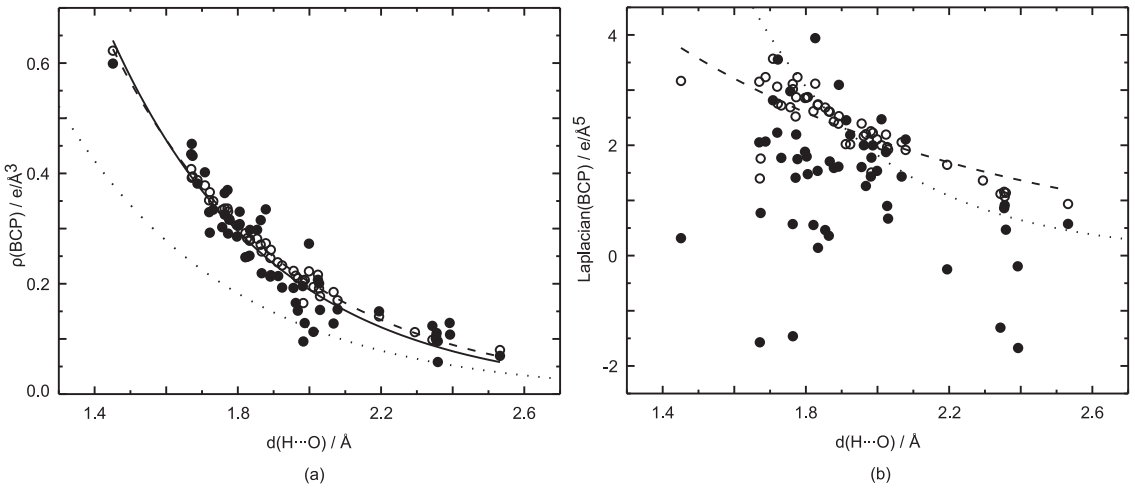

Figure 5.6: Topological properties of hydrogen bonds from MEM (filled circles) and PRIOR (open circles) densities in dependence on the distance $d(\mathrm{H}\cdots\mathrm{O})$. (a) $\rho^{MEM}(BCP)$ and (b) $\nabla^2\rho^{MEM}(BCP)$. The solid line represents the fit to the MEM values given by $\rho(BCP) = 16.14\,(3.76)\,\exp[-2.22\,(13)d(\mathrm{H}\cdots\mathrm{O})]$. Dashed lines are fits to PRIOR values, with $\rho(BCP) = 12.27\,(68)\,\exp[-2.05\,(3)\,d(\mathrm{H}\cdots\mathrm{O})]$ and $\nabla^2\rho(BCP) = 17.77\,(4.35)\,\exp[-1.07\,(13)\,d(\mathrm{H}\cdots\mathrm{O})]$. Dotted lines are functions derived by Espinosa[88] from fits to topological properties of multipole densities, with $\rho(BCP) = 8\,(4)\,\exp[-2.1\,(3)\,d(\mathrm{H}\cdots\mathrm{O})]$ and $\nabla^2\rho(BCP) = 330\,(180)\,\exp[-2.6\,(3)\,d(\mathrm{H}\cdots\mathrm{O})]$.

hydrogen bonds,[104] we have established that only stronger hydrogen bonds have a large potential to draw electrons into the BCP, resulting in an accumulation of charge between the hydrogen atom and the acceptor atom. The potential to draw electrons into the regions of BCPs is confirmed by the energy densities at the BCPs, which show increasing magnitudes with decreasing distance $d(\mathrm{H}\cdots\mathrm{O})$ (Section 5.3.3).

5.3.2 Topological properties of hydrogen bonds

Densities at the BCPs of hydrogen bonds of both MEM and PRIOR densities depend exponentially on the distance $d(\mathrm{H}\cdots\mathrm{O})$ [Fig. 5.6(a)]. For the PRIOR this dependence is almost perfect, while values of $\rho(BCP)$ derived from MEM densities exhibit a larger scatter about the exponential curve. Contributions to this scatter will come from the properties of the MEM that it will have fitted part of the noise in the data and that it suffers from series termination effects due to incompleteness

of the data. Furthermore, part of the scatter of values will be a real property of the compounds studied, which is caused by variations of the thermal motion between different atoms, thus leading to differences in dynamic densities even if the static density would be similar.

A quantum mechanical theory does not exist, that would demand an exponential relationship between $\rho(BCP)$ and $d(\mathrm{H}\cdots\mathrm{O})$. Deviations from an average exponential relation can thus be caused by variations of the properties of the bonds, *e.g.* as caused by variations of their environments. It is noted that Espinosa *et al.*[88] have established an exponential relationship between $\rho^{MP}(BCP)$ of hydrogen bonds and $d(\mathrm{H}\cdots\mathrm{O})$ for static multipole densities, albeit with different values of the parameters in the exponential function than have presently been determined for MEM and PRIOR densities (Fig. 5.6). The values of $\rho^{MP}(BCP)$ also show a substantial scatter about the proposed exponential dependence on the distance $d(\mathrm{H}\cdots\mathrm{O})$.[88]

The dynamic densities at the BCPs of hydrogen bonds in the PRIOR and the MEM densities are in general larger than the corresponding values of the static multipole density. In view of the discussion in Section 5.3.1, we believe this to be a real effect that is caused by the dynamic *vs* static character of the densities and by the inflexibility of the multipole model in the region of hydrogen bonds.[3,4,5,6,7]

An exponential dependence on $d(\mathrm{H}\cdots\mathrm{O})$ is also observed for the values of the second derivatives of $\rho(\mathbf{x})$ at the BCPs of hydrogen bonds, as they are provided by the three eigenvalues $\lambda_1(BCP)$, $\lambda_2(BCP)$ and $\lambda_3(BCP)$ of the Hessian matrix (Fig. 5.7). As for the values of the densities themselves, the exponential relationship is almost perfectly fulfilled for the PRIOR, while some scatter of values about the exponential curve can be observed for the eigenvalues derived from the MEM densities. A similar behavior of $\lambda_3(BCP)$—the curvature at the BCP in the direction of the bond path—has been reported for static multipole densities by Espinosa *et al.*.[88] Following the proposal by Espinosa *et al.*,[88] $\lambda_3(BCP)$ as derived from dynamic MEM densities, might thus form a suitable parameter for the classification of hydrogen bonds.

Values of the Laplacian $\nabla^2\rho(BCP) = \lambda_1(BCP) + \lambda_2(BCP) + \lambda_3(BCP)$ exhibit an exponential dependence on $d(\mathrm{H}\cdots\mathrm{O})$ for the PRIOR, while they show a considerable scatter for the MEM densities [Fig. 5.6(b)]. First of all, these variations $\nabla^2\rho(BCP)$ can be explained by the fact, that $\lambda_1(BCP) + \lambda_2(BCP) \approx -\lambda_3(BCP)$ [Fig. 5.7 (d)], such that minor variations of the values of the individual

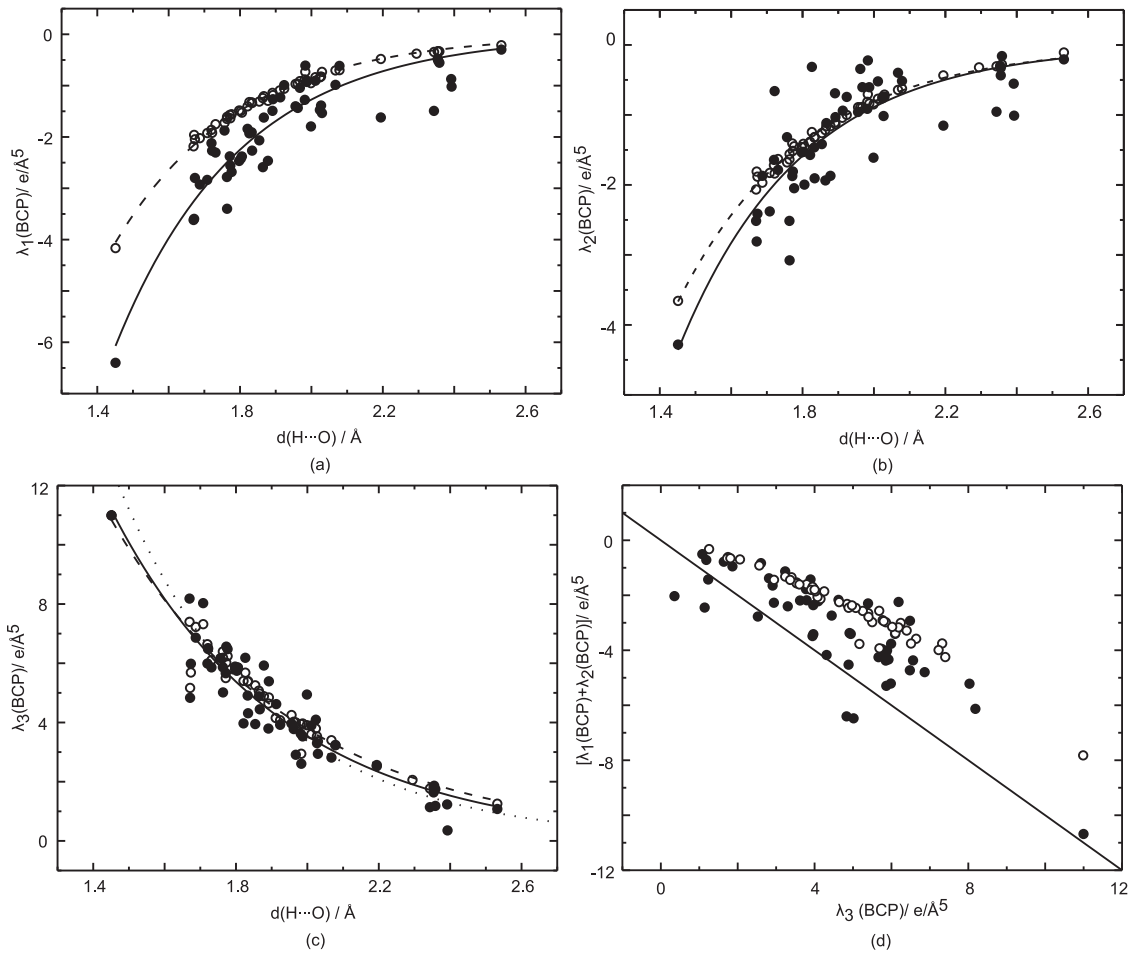

Figure 5.7: Curvatures at the BCPs of hydrogen bonds for MEM (filled circles) and PRIOR (open circles) densities. (a) $\lambda_1(BCP)$, (b) $\lambda_2(BCP)$ and (c) $\lambda_3(BCP)$ as a function of the distance $d(\mathrm{H}\cdots\mathrm{O})$. (d) $\lambda_1(BCP) + \lambda_2(BCP)$ *vs* $\lambda_3(BCP)$. Solid lines represent fits to MEM values with $\lambda_1(BCP) = -378.51\,(117.17)$ $\exp[-2.85\,(18)\,d(\mathrm{H}\cdots\mathrm{O})]$, $\lambda_2(BCP) = -292.51\,(144.41)$ $\exp[-2.90\,(29)\,d(\mathrm{H}\cdots\mathrm{O})]$ and $\lambda_3(BCP) = 233.39\,(53.80)$ $\exp[-2.09\,(13)\,d(\mathrm{H}\cdots\mathrm{O})]$. Dashed lines represent fits to PRIOR values with $\lambda_1(BCP) = -261.92\,(15.83)$ $\exp[-2.87(4)d(\mathrm{H}\cdots\mathrm{O})]$, $\lambda_2(BCP) = -200.00\,(13.58)$ $\exp[-2.76(4)d(\mathrm{H}\cdots\mathrm{O})]$ and $\lambda_3(BCP) = 176.54\,(23.44)$ $\exp[-1.92(7)d(\mathrm{H}\cdots\mathrm{O})]$. The dotted line is the function determined by Espinosa *et al.*[88] for multipole densities, with $\lambda_3(BCP) = 410(80)$ $\exp[-2.4(1)d(\mathrm{H}\cdots\mathrm{O})]$. The solid line in (d) is the function $\lambda_1(BCP) + \lambda_2(BCP) = -\lambda_3(BCP)$.

eigenvalues are magnified towards large variations of $\nabla^2\rho(BCP)$. We believe the source of these variations to be, on the one hand, artifacts of the MEM and noise in the data, and, on the other hand, variations of the thermal motion between different structures (see the discussion on $\rho(BCP)$ above). Especially the latter property has previously been shown to be an important effect, where relatively small variations of thermal parameters lead to large variations of $\nabla^2\rho(BCP)$, while they hardly effect $\rho(BCP)$.[42]

Because thermal motion depends on the crystal packing, part of the observed variations of $\nabla^2\rho(BCP)$ will reflect true variations of the dynamic MEM electron densities, as they are the result from true variations of the thermal motion. Nevertheless, the observed negative values of $\nabla^2\rho(BCP)$ most probably are caused by noise in the diffraction data. The outliers of $\nabla^2\rho(BCP)$ do not belong to a particular data set, which excludes the explanation that one of the data sets might be particularly affected by noise or systematic errors.

5.3.3 Energetic properties of hydrogen bonds

The kinetic, potential and total energy densities at the BCPs of hydrogen bonds show a nearly perfect exponential dependence on $d(\mathrm{H}\cdots\mathrm{O})$ for PRIOR densities (Fig. 5.8). Corresponding values from MEM densities scatter around an average exponential dependence. A larger or smaller scatter is obtained, depending on the relative importance of $\rho(BCP)$ and $\nabla^2\rho(BCP)$ in determining each property (Eqs. 5.6–5.8).

Similar exponential relationships have been obtained for static multipole densities by Espinosa *et al.*.[89] Although the functions for energy densities derived from dynamic MEM densities are different from those for static multipole densities, the differences are smaller than in the case of $\rho(BCP)$ and $\nabla^2\rho(BCP)$. This indicates a compensating effect on going from static to dynamic densities, where, on the average, a larger value of $\rho(BCP)$ is compensated by a smaller value of $\nabla^2\rho(BCP)$ (Fig. 5.6).

The kinetic and potential energy densities provide information about the nature of the bond. The potential energy density $|V(BCP)|$ describes the ability of the system to concentrate electrons at the BCPs, while the kinetic energy density $G(BCP)$ describes the tendency of the electrons to spread out.[89] Accordingly, values $|V(BCP)|/G(BCP) < 1$ are considered to indicate a depletion of electrons at the

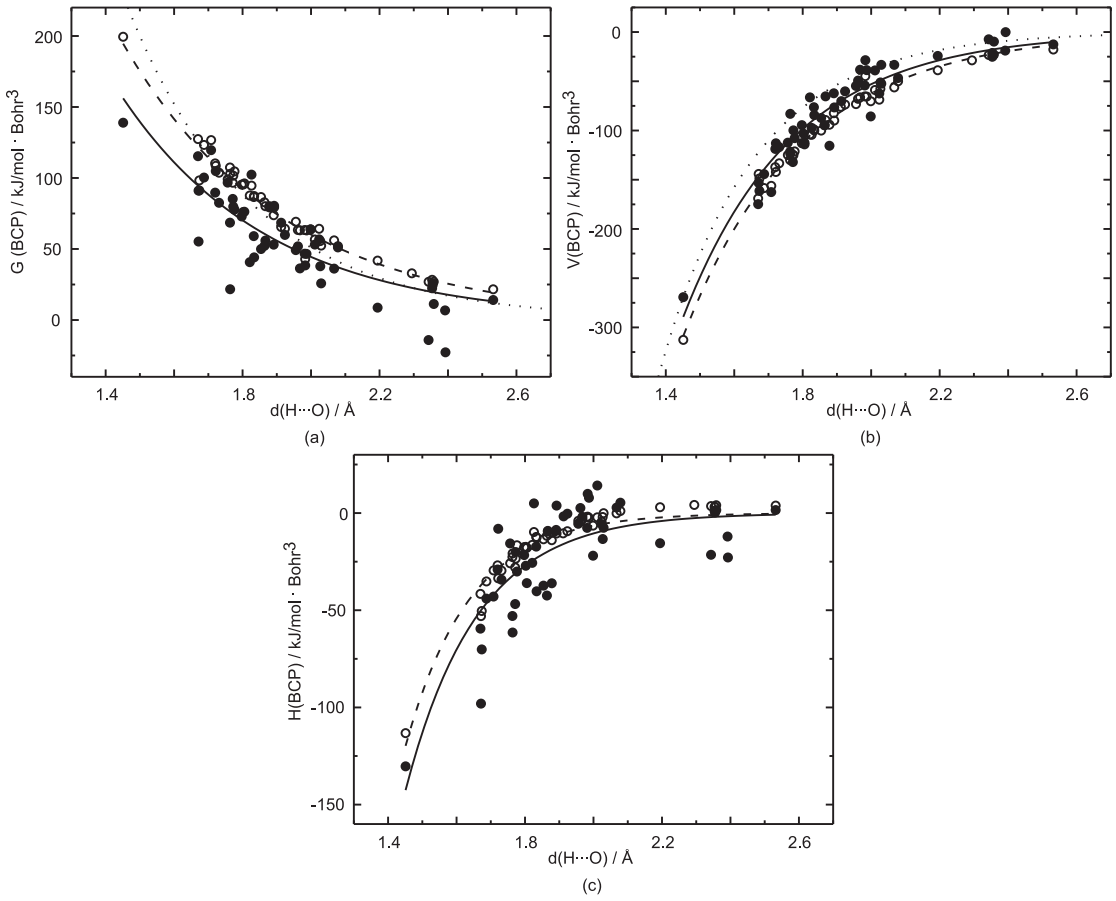

Figure 5.8: Energetic properties of hydrogen bonds of MEM (filled circles) and PRIOR (open circles) densities in dependence on the distance $d(\mathrm{H}\cdots\mathrm{O})$. (a) Kinetic energy density $G(BCP)$, (b) Potential energy density $V(BCP)$, and (c) Total energy density $H(BCP)$. Solid lines represent fits to the MEM values with $G(BCP) = 4.331\,(1.879)\,10^3$ $\exp[-2.29\,(25)\,d(\mathrm{H}\cdots\mathrm{O})]$, $V(BCP) = -2.572\,(627)\,10^4$ $\exp[-3.09\,(14)\,d(\mathrm{H}\cdots\mathrm{O})]$ and $H(BCP) = -1.419\,(1.040)\,10^5$ $\exp[-4.76\,(46)\,d(\mathrm{H}\cdots\mathrm{O})]$. Dashed lines correspond to fits to PRIOR values with $G(BCP) = 4.335\,(590)\,10^3$ $\exp[-2.14\,(8)\,d(\mathrm{H}\cdots\mathrm{O})]$, $V(BCP) = -2.113\,(178)\,10^4$ $\exp[-2.91\,(5)\,d(\mathrm{H}\cdots\mathrm{O})]$ and $H(BCP) = -2.721\,(782)\,10^5$ $\exp[-5.33\,(18)\,d(\mathrm{H}\cdots\mathrm{O})]$. Dotted lines are functions determined by Espinosa *et al.*[89] for multipole densities, with $G(BCP) = 12\,(2)\,10^3$ $\exp[-2.73\,(9)\,d(\mathrm{H}\cdots\mathrm{O})]$ and $V(BCP) = -50\,(1.1)\,10^3$ $\exp[-3.6\,d(\mathrm{H}\cdots\mathrm{O})]$.

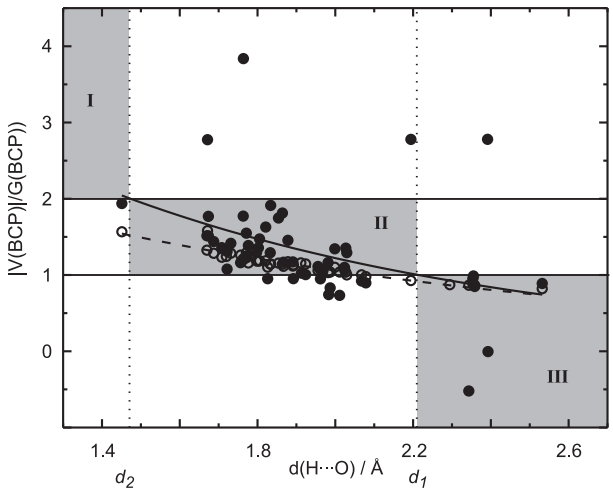

Figure 5.9: Ratio of potential and kinetic energy densities $|V(BCP)|/G(BCP)$ at BCPs of hydrogen bonds of MEM (filled circles) and PRIOR (open circles) densities in dependence on the distance $d(\mathrm{H}\cdots\mathrm{O})$. The solid line represents the fit to the MEM values with $|V(BCP)|/G(BCP) = 7.95\,(5.07)\,\exp[-0.94\,(34)\,d(\mathrm{H}\cdots\mathrm{O})]$. The dashed line represents the fit to the PRIOR values with $|V(BCP)|/G(BCP) = 4.17(30)\,\exp[-0.69\,(4)\,d(\mathrm{H}\cdots\mathrm{O})]$. $d_1 = 2.21$ Å and $d_2 = 1.47$ Å.

BCPs, which corresponds to closed-shell interactions. Values $|V(BCP)|/G(BCP) > 2$ indicate an accumulation of electrons at the BCP, which corresponds to a shared-shell interaction, *i.e.* a covalent bond. Values of $|V(BCP)|/G(BCP)$ between one and two describe bonds with partial covalent and partial ionic character (Espinosa *et al.*[90] and references therein).

In agreement with previous studies on multipole densities by Espinosa *et al.*,[89] we find for hydrogen bonds that both $|V(BCP)|$ and $G(BCP)$ increase on decreasing distance $d(\mathrm{H}\cdots\mathrm{O})$ (Fig. 5.8). However, the relation between $V(BCP)$ and $G(BCP)$ is not linear, such that $|V(BCP)|/G(BCP)$ increases with decreasing distance $d(\mathrm{H}\cdots\mathrm{O})$ (Fig. 5.9). These relations can again be described by exponential functions. From the average exponential dependence of $|V(BCP)|/G(BCP)$ of MEM densities on $d(\mathrm{H}\cdots\mathrm{O})$, two distances can be derived that describe the crossover between covalent, mixed-character and closed-shell types of hydrogen bonds. The distance $d_1 = 2.21$ Å is the distance at which $|V(BCP)|/G(BCP) = 1$ and $d_2 = 1.47$ Å is the distance at which $|V(BCP)|/G(BCP) = 2$. It is thus found that

long hydrogen bonds ($d(\mathrm{H}\cdots\mathrm{O}) > d_1$) are dominated by electrostatic interactions, while short hydrogen bonds ($d(\mathrm{H}\cdots\mathrm{O}) < d_2$) are covalent bonds. Most hydrogen bonds studied in the present work are of intermediate character ($d_2 < d(\mathrm{H}\cdots\mathrm{O}) < d_1$; see Fig. 5.9) and thus at least partly covalent.

The distances d_1 and d_2 coincide with the classification by Jeffrey,[78] who considers hydrogen bonds with $d(\mathrm{H}\cdots\mathrm{O}) > 2.2$ Å to be weak and hydrogen bonds with $d(\mathrm{H}\cdots\mathrm{O}) < 1.5$ Å to be strong. We have thus found that strong hydrogen bonds are covalent bonds, while weak hydrogen bonds possess mainly electrostatic character. Most hydrogen bonds are of intermediate strength and will have mixed covalent–electrostatic character.

A few outliers can be observed in Fig. 5.9, for which $|V(BCP)|/G(BCP) \gg 2$ or $|V(BCP)|/G(BCP) < 0$. These points are precisely those hydrogen bonds for which a negative Laplacian $\nabla^2\rho(BCP)$ has been obtained (Fig. 5.6). Because we interpret a negative Laplacian as unphysical for hydrogen bonds, these values of $|V(BCP)|/G(BCP)$ are most probable caused by inaccuracies of the MEM or the data. Deviations from a smooth dependence on the distance might also arise from the approximate character of Eqs. 5.6–5.8,[85] and from the fact that these relations have not been derived for dynamic densities.

5.3.4 Topological and energetic properties of covalent bonds

Densities at the BCPs of covalent bonds are found to depend exponentially on the bond length d(X–Y). The bonds C–O, C–N and C–C require a different function than the covalent bonds C–H and N–H and hydrogen bonds, while all bonds involving hydrogen atoms are described by a single curve [Fig. 5.10(a)]. A more close analysis of the values $\rho^{prior}(BCP)$ shows that slightly different exponential curves apply to values obtained from different bond types C–O, C–N and C–C, in agreement with the behavior of $\rho^{MP}(BCP)$ obtained from multipole densities by Dominiak *et al.*.[103] A similar analysis cannot be made for $\rho^{MEM}(BCP)$, because of the relatively few data points for each bond type and the larger scatter about the average exponential relation.

The exponential dependencies of $\rho^{MEM}(BCP)$ and $\rho^{prior}(BCP)$ on $d(\mathrm{H}\cdots\mathrm{O})$ as determined from hydrogen bonds only [$1.4 < d(\mathrm{H}\cdots\mathrm{O}) < 2.6$ Å; Fig. 5.6(a) and the dotted lines in Fig. 5.10(a)] extrapolate well towards values of $\rho^{MEM}(BCP)$ and $\rho^{prior}(BCP)$, respectively, for covalent bonds C–H and N–H [$1.0 < d(\mathrm{H}\cdots\mathrm{O}) <$

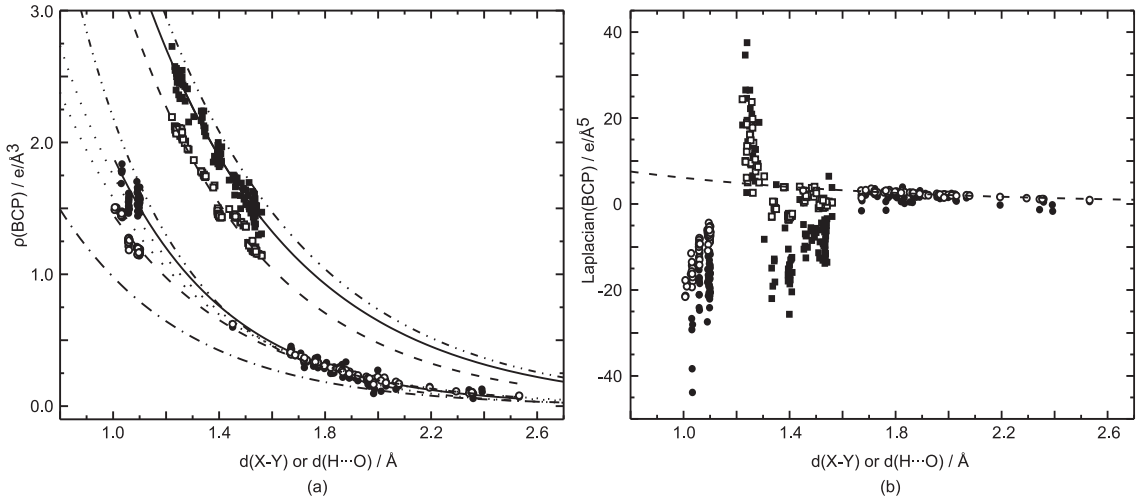

Figure 5.10: Topological properties at BCPs from MEM (filled symbols) and PRIOR (open symbols) densities in dependence on the bond length d(X–Y) or distance d(H$\cdots$O). Covalent bonds C–O, C–N, C–C (squares) are distinguished from covalent C–H and N–H and hydrogen bonds (circles). (a) $\rho(BCP)$, and (b) $\nabla^2\rho(BCP)$. Solid lines represent fits to MEM values with $\rho(BCP) = 23.21\,(95)\,\exp[-1.79\,(3)d(\mathrm{H}\cdots\mathrm{O})]$ for C–O, C–N and C–C bonds and $\rho(BCP) = 18.45\,(99)\,\exp[-2.28\,(5)d(\mathrm{H}\cdots\mathrm{O})]$ for C–H, N–H and hydrogen bonds. Dashed lines correspond to fits to PRIOR values with $\rho(BCP) = 23.47\,(98)\,\exp[-1.95\,(3)d(\mathrm{H}\cdots\mathrm{O})]$ for C–O, C–N and C–C bonds and $\rho(BCP) = 10.71\,(35)\,\exp[-2.00\,(3)d(\mathrm{H}\cdots\mathrm{O})]$ for C–H, N–H and hydrogen bonds. Dotted lines represent the fits to values for hydrogen bonds from Fig. 5.6(a). The dash-dotted line is the function determined by Espinosa *et al.*[88] from multipole values for hydrogen bonds [Fig. 5.6(a)]. Dash-dot-dotted lines are functions from Dominiak *et al.*[103] with $\rho^{MP}(BCP) = \exp[-1.74\,(4)(d(\mathrm{X–Y})-1.822(10))]$ for C–O, C–N and C–C bonds, and with $\rho^{MP}(BCP) = \exp[-2.61\,(5)(d(\mathrm{X–Y})-1.300(4))]$ for covalent C–H, N–H and O–H bonds and hydrogen bonds.

1.2 Å; Fig. 5.10(a)]. This remarkable feature indicates that a different trend exists in the values $\rho^{MEM}(BCP)$ and $\rho^{prior}(BCP)$ of hydrogen bonds, despite the fact that their magnitudes are within the same range and that the fitted curves hardly differ within the hydrogen-bond region [Fig. 5.6(a)]. Figure 5.10(a) shows that the similar magnitudes of $\rho^{MEM}(BCP)$ and $\rho^{prior}(BCP)$ of hydrogen bonds are a coincidental feature, because the different trends show that—for purposes of characterizing hydrogen bonds—the true density cannot be replaced by the procrystal density, as it has sometimes been suggested.[92,93]

Previously we have shown that for most bonds in trialanine and α-glycine $\rho^{MEM}(BCP) < \rho^{MP}(BCP)$, while the opposite is true for hydrogen bonds.[42,53] The extended analysis presented here shows that, on the average, this property is valid for all covalent bonds. The exponential dependence of $\rho^{MEM}(BCP)$ of H–X covalent and H$\cdots$O hydrogen bonds intersects the function determined for $\rho^{MP}(BCP)$ by Dominiak *et al.*[103] at a distance of 1.44 Å such that $\rho^{MEM}(BCP) < \rho^{MP}(BCP)$ when $\rho(BCP)$ is large and $\rho^{MEM}(BCP) > \rho^{MP}(BCP)$ when $\rho(BCP)$ is small. Figure 5.10(a) suggests that a similar property would be valid for the Van der Waals contacts, but an extrapolation towards large distances (d(X–Y) > 3 Å) does not seem permissable for the given accuracy and range of data points.

Eigenvalues of the Hessian matrix at the BCPs show a systematic variation with the bond length, which can be described by an exponential function within the limited range of distances d(H$\cdots$O) of hydrogen bonds (Figs. 5.7 and 5.11). While each type of bond seems to require its own curve, the limited number of data points does not allow these functions to be determined. Unlike $\rho(BCP)$, the values of distance dependencies of $\lambda_1(BCP)$, $\lambda_2(BCP)$ and $\lambda_3(BCP)$ of hydrogen bonds do not extrapolate well towards corresponding values for covalent N–H and C–H bonds. Instead, each type of covalent bond exhibits a large variation of values of the second derivatives, while being of almost a single length (Fig. 5.11). Not so dramatic, the distance dependencies exhibit similar features for the values of the second derivatives at the BCPs of other covalent bonds. This property is enhanced for the values of $\nabla^2\rho^{MEM}(BCP)$ and $\nabla^2\rho^{prior}(BCP)$ [Fig. 5.10(b)]. While a systematic dependence of $\nabla^2\rho(BCP)$ on the bond distance has been reported for values derived from multipole densities,[103] a close inspection of the published diagrams shows that they exhibit similar features as presently found for $\nabla^2\rho^{MEM}(BCP)$ and $\nabla^2\rho^{prior}(BCP)$, *i.e.* different values for bonds of nearly equal length.

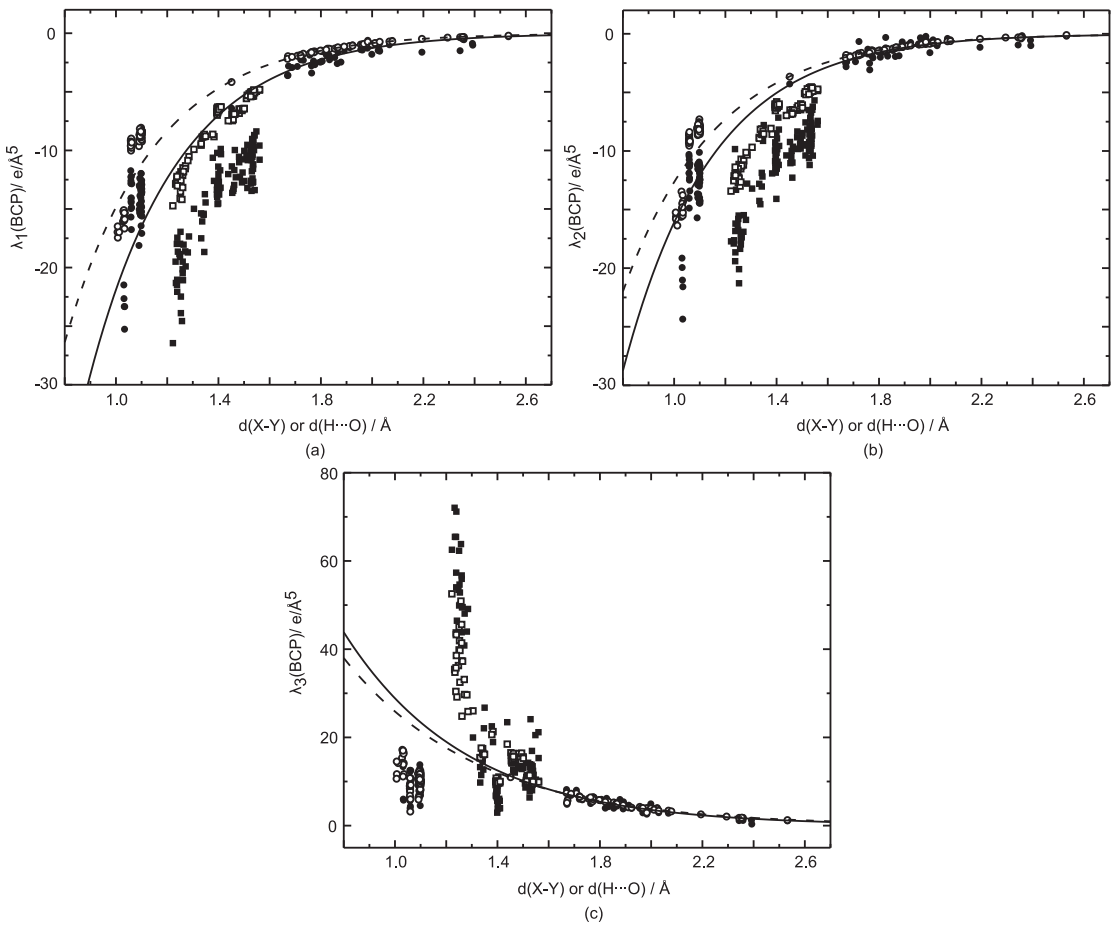

Figure 5.11: Curvatures at BCPs from MEM (filled symbols) and PRIOR (open symbols) densities in dependence on the bond length d(X–Y) or distance $d(\mathrm{H}\cdots\mathrm{O})$. Covalent bonds C–O, C–N, C–C (squares) are distinguished from covalent C–H and N–H and hydrogen bonds (circles). (a) $\lambda_1(BCP)$, (b) $\lambda_2(BCP)$, and (c) $\lambda_3(BCP)$. Lines represent fits to the values for hydrogen bonds from MEM (solid lines) and PRIOR (dashed lines) densities, as taken from Fig. 5.7.

The properties of the distance dependencies of $\rho(BCP)$ and $\nabla^2\rho(BCP)$ are transported towards the energy densities at the BCPs (Fig. 5.12). They are even more pronounced for the values of $|V(BCP)|/G(BCP)$. $|V^{MEM}(BCP)|/G^{MEM}(BCP)$ assumes a large range of values for covalent bonds C–C, C–N, C–H and N–H, while distances of these bonds cluster around a few values only (Fig. 5.13). This suggest a variation of the character of bonds of similar length, as it might be the result of variations of thermal motion and different environments of these bonds.

Polar C–O bonds appear as of mixed covalent/ionic character with $|V^{MEM}(BCP)|/G^{MEM}(BCP) < 2$, while bonds C–C, C–N, C–H and N–H appear as covalent with $|V^{MEM}(BCP)|/G^{MEM}(BCP) > 2$ (Fig. 5.13), in accordance with general chemical knowledge. For covalent bonds C–C and C–N $|V^{prior}(BCP)|/G^{prior}(BCP)$ is close to two and generally much smaller than $|V^{MEM}(BCP)|/G^{MEM}(BCP)$. This again shows that the true density should not be replaced by the procrystal density for the quantitative description of chemical bonding.

5.4 Conclusions

Accurate Charge Densities have been determined by the MEM from X-ray diffraction data on six different crystals of amino acids and tripeptides. Employing the previously proposed criterion of convergence for the iterations of the MEM,[42] the range of values of χ^2_{aim} has been found to encompass a factor of four. These values correlate with the GoF of the multipole refinements (Table 5.1), and they thus show the ability of the MEM to determine the correct scale of standard uncertainties of measured intensities (Section 5.2.1).

Electron densities $\rho^{MEM}(\mathbf{x})$ and $\rho^{MP}(\mathbf{x})$ exhibit similar features, with atomic maxima and BCPs at similar positions.[42,53] Distinctions are due to the differences between dynamic $[\rho^{MEM}(\mathbf{x})]$ and static $[\rho^{MP}(\mathbf{x})]$ densities as well as the peculiarities of each method. Electron densities in hydrogen bonds have been found to be better represented by the MEM than by multipole models (Section 5.3.1), as it is the result of the the inflexibility of the multipole model in describing densities at remote positions from atomic maxima.[3,4,5,6,7]

MEM electron densities at BCPs show an exponential dependence on the bond

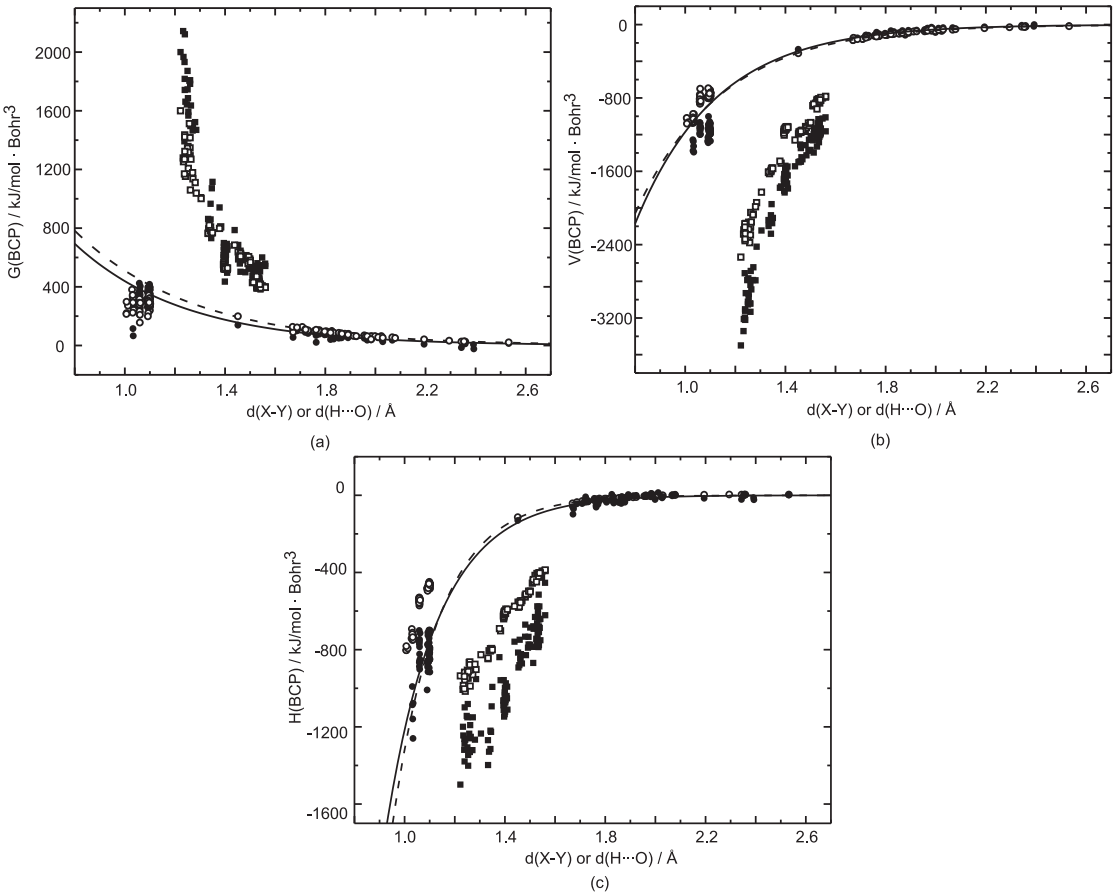

Figure 5.12: Energetic properties at BCPs from MEM (filled symbols) and PRIOR (open symbols) densities in dependence on the bond length d(X–Y) or distance d(H$\cdots$O). Covalent bonds C–O, C–N, C–C (squares) are distinguished from covalent C–H and N–H and hydrogen bonds (circles). (a) Kinetic energy density $G(BCP)$, (b) potential energy density $V(BCP)$, and (c) total energy density $H(BCP)$. Lines represent fits to the values for hydrogen bonds from MEM (solid lines) and PRIOR (dashed lines) densities, as taken from Fig. 5.8.

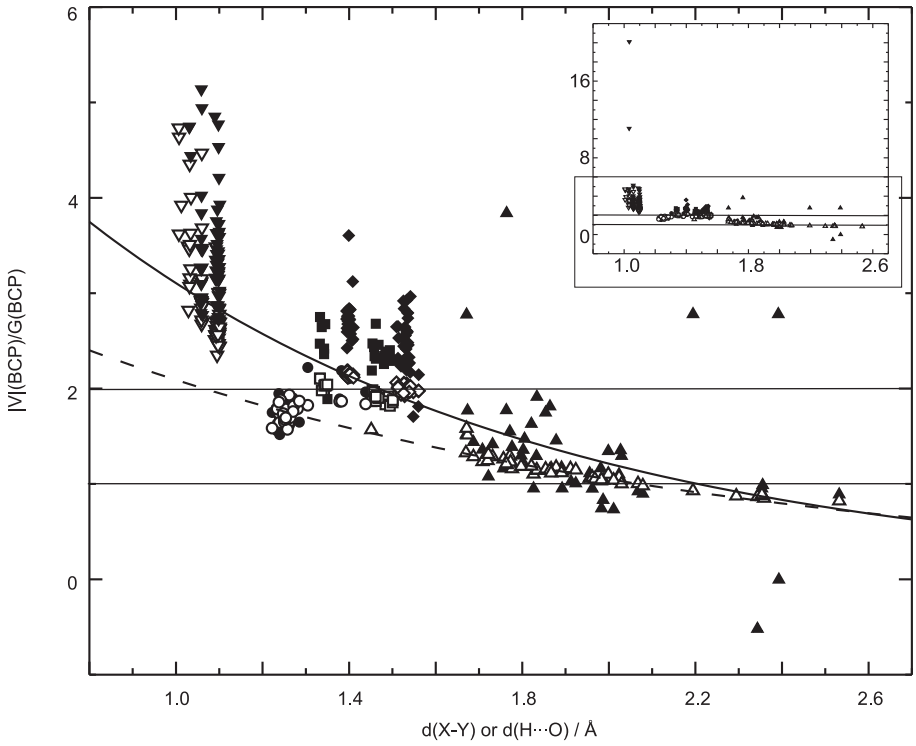

Figure 5.13: The ratio $|V(BCP)|/G(BCP)$ from MEM (filled symbols) and PRIOR (open symbols) densities in dependence on the bond length d(X–Y) or distance $d(\mathrm{H}\cdots\mathrm{O})$. Indicated are values for C–O (circles), C–N (squares), C–C (diamonds), C–H (down triangles) and N–H (down triangles) covalent bonds, and for hydrogen bonds (up triangles). The solid line (MEM values) and the dashed line (PRIOR values) represent fits to values for hydrogen bonds from Fig. 5.9. The inset shows all values of $|V(BCP)|/G(BCP)$, including two very large ratios for N–H bonds.

length with individual functions for covalent bonds and hydrogen bonds. These functions differ from the functions that have been determined for corresponding values derived from multipole densities [Figs. 5.6(a) and 5.10(a)].[88,103] In general, $\rho^{MEM}(BCP) <$ and $\rho^{MP}(BCP)$ for covalent bonds, while the opposite is true for hydrogen bonds (Section 5.3.4).

Values of $\nabla^2\rho^{MEM}(BCP)$ exhibit much more scatter than corresponding values from multipole densities. Nevertheless, it has proved possible to establish systematic dependencies of energetic properties at BCPs on the bond length (Fig. 5.12). In particular the ratio between potential and kinetic energy density shows two kinds of behavior.

For hydrogen bonds, $|V(BCP)|/G(BCP)$ derived from MEM densities allows a classification of hydrogen bonds according to their distance $d(\mathrm{H}\cdots\mathrm{O})$ (Section 5.3.3). Short hydrogen bonds ($d(\mathrm{H}\cdots\mathrm{O}) < d_1 = 1.47$ Å) are covalent bonds, hydrogen bonds of intermediate length ($d_1 < d(\mathrm{H}\cdots\mathrm{O}) < d_2 = 2.21$ Å) possess mixed covalent-ionic character, while long hydrogen bonds ($d(\mathrm{H}\cdots\mathrm{O}) > d_2$) are mainly stabilized by closed-shell type of interactions. This classification coincides with the usual classification of strong ($d(\mathrm{H}\cdots\mathrm{O}) < 1.5$ Å), intermediate ($1.5 < d(\mathrm{H}\cdots\mathrm{O}) < 2.2$ Å), and weak ($d(\mathrm{H}\cdots\mathrm{O}) > 2.2$ Å) hydrogen bonds.[78]

For covalent bonds, the ratio $|V^{MEM}(BCP)|/G^{MEM}(BCP)$ assumes values within a large range for each type of bond with a narrow range of bond lengths (Fig. 5.13). This feature indicates that the character of covalent bonds of a single type (*e.g.* C(sp^3)–C(sp^3) bonds) varies despite almost equal bond lengths. A classification of bonds according to their length can therefore capture at most part of the chemistry.

The procrystal PRIOR is only rarely considered in Accurate Charge Density studies.[93] It has been shown here that topological properties at BCPs exhibit similar dependencies on bond lengths when derived from MEM and procrystal densities, while the latter values show much less scatter. In particular, most of the density in the BCP is already described by the procrystal density [Fig. 5.10(a)], which illustrates the difficulties of multipole refinements and MEM calculations in establishing an Accurate Charge Density beyond the procrystal model. Differences between MEM and procrystal densities are more pronounced in the derivatives $\nabla^2\rho(BCP)$ and energy densities, such that only Accurate Charge Densities may lead to an correct interpretation of the character of bonds.

Different trends could be identified in the distance dependencies of $\rho^{MEM}(BCP)$ and $\rho^{prior}(BCP)$ of hydrogen bonds, despite almost equal values of $\rho^{MEM}(BCP)$ and $\rho^{prior}(BCP)$ for these bonds [Figs. 5.6(a) and 5.10(a)]. This remarkable feature stresses the differences between MEM and procrystal densities, and it shows once more that—for purposes of characterizing chemical bonding—the true density cannot be replaced by the procrystal density, as it has sometimes been suggested.[92,93]

Acknowledgements

We are greatly indebted to R. Destro and P. Luger and coworkers for providing the diffraction data. R. Destro and L. Lo Presti are thanked for providing a yet unpublished plot of a difference Fourier map of L-alanine [Fig. 5.2(c)]. Financial support was obtained from the German Science Foundation (DFG) within the framework of the priority program SPP1178.

Table 5.1: Crystallographic data of α-glycine (Gly),[74] L-alanine (Ala),[79] L-phenylalanine formic acid complex (Phe),[72] trialanine (Ala-Ala-Ala),[14] L-alanyl-L-tyrosyl-L-alanine with water (Ala-Tyr-Ala$_{(H_2O)}$),[83] and L-alanyl-L-tyrosyl-L-alanine with ethanol (Ala-Tyr-Ala$_{(EtOH)}$),[83] together with summaries of the ISAM refinements and MEM calculations (present work). Reflections with $I/\sigma(I) > 3$ are classified as observed, with the exception of the criterion $F_{obs}/\sigma(F_{obs}) > 4$ for trialanine.[14]

Compound	Gly	Ala	Phe	Ala-Ala-Ala	Ala-Tyr-Ala$_{(H_2O)}$	Ala-Tyr-Ala$_{(EtOH)}$
Chemical formula	$C_2O_2NH_5$	$C_3O_2NH_7$	$C_9H_{11}NO_2$ $\cdot$ $C_9H_{12}NO_2^+ \cdot CHO_2^-$	$C_9H_{17}N_3O_4 \cdot H_2O$	$C_{15}H_{21}N_3O_5 \cdot 2.634\ H_2O$	$C_{15}H_{21}N_3O_5 \cdot C_2H_5OH$
Space group	$P2_1/n$	$P2_12_12_1$	$P2_1$	C2	$P2_1$	$P2_1$
Z	4	4	2	8	2	2
a (Å)	5.0866	5.9279	11.4585	18.4408	8.121	8.845
b (Å)	11.7731	12.2597	5.5941	5.2153	9.299	9.057
c (Å)	5.4595	5.7939	14.2147	24.8543	12.532	12.364
β (deg)	111.99	90.00	99.46	98.76	91.21	94.56
V (Å^3)	303.16	421.1	898.8	2362.4	946.2	987.3
F(000)	160	192	400	1031.8	397	396
Temperature (K)	23	23	25	20	9	20
Wavelength λ (Å)	0.71073	0.71073	0.71073	0.71073	0.50000	0.71073
$[\sin(\theta)/\lambda]_{\max}$ (Å^{-1})	1.15	1.08	1.18	1.15	1.24	1.11
Unique reflections (obs/all)	3483/3822	2328/2535	8971/10981	12928/14895	12875/14111	10901/11703
Multipole refinement[74, 79, 72, 83, 14]						
R_F(obs/all)	-/1.29	-/2.03	2.72/3.50	1.83/2.47	2.93/3.51	2.23/2.64
R_{wF}(obs/all)	-/-	-/1.59	3.07/-	1.53/-	2.08/-	1.77/-
GoF	1.04	1.17	1.06	0.67	2.06	1.63
ISAM refinement						
R_F(obs/all)	2.33/2.60	2.85/3.16	4.04/4.80	-	3.99/4.55	3.60/4.00
R_{wF}(obs/all)	5.25/5.35	3.73/3.77	4.73/4.87	-	4.78/4.94	4.19/4.25
GoF(obs/all)	2.02/1.96	1.81/1.75	1.43/1.33	-	1.67/1.64	2.27/2.22
MEM calculation[53, 42]						
Number of pixels	$64 \times 144 \times 72$	$72 \times 162 \times 72$	$144 \times 72 \times 192$	$216 \times 64 \times 324$	$96 \times 108 \times 162$	$108 \times 108 \times 162$
χ^2_{aim}	0.3131	0.7600	0.8300	0.4250	1.2750	1.3081
R_F/R_{wF}	1.04/1.53	1.99/1.90	3.55/3.43	2.63/1.84	3.30/3.42	2.48/2.22

Chapter 6

Summary

The multipole method provides a description of the crystal structure that goes beyond the simple independent spherical atom model (ISAM). The effects on the density due to chemical bonding can be recovered with the aspherical electron density provided by the multipole refinement against X-ray diffraction data. However, due the generally large number of highly-correlated parameters, *ad hoc* constraints on these parameters may lead to an incomplete description of the electron density. This problem is aggravated for increasing size of the unit cell. The Maximum Entropy Method (MEM) reconstructs electron densities model-independently and does not suffer from correlated parameters. The present thesis shows that the MEM is well suited for Accurate Charge Density studies.

Accurate Charge Densities of several amino acids and tripeptides have been reconstructed by the MEM to study chemical bonds. The study of trialanine is presented in Chapter 3 and the study of α-glycine is reported in Chapter 4. An extensive study of electron densities of six amino acids and tripeptides including 52 hydrogen bonds is reported in Chapter 5. In Chapter 2 the procedure for the reconstruction of the MEM densities is described. The MEM has been employed to select the most probable electron density, which simultaneously maximizes the informational entropy, fulfills the normalization constraint and fits the diffraction data. To determine the most probable electron density, an iterative search is necessary because an analytical solution does not exist. The computer program BayMEM which incorporates the Cambridge algorithm has been employed for the reconstruction of the densities optimized by the MEM.

The quality of the MEM densities is enhanced by several extensions to the MEM such as the use of a non-uniform prior density, the employment of the prior-derived F-constraint, static weighting, and the choice of the optimal value of χ^2_{aim} as stopping criterion for the MEM iterations. These extensions are explained in Chapter 2. The procedure of the determination of the optimal χ^2_{aim} has been developed within the present thesis (Chapters 3 and 4). It is demonstrated that via inspection of dynamic deformation maps and difference Fourier maps, the point of convergence of the MEM iterations is determined and thus the optimal χ^2_{aim} is chosen. The results of the studied compounds indicate that the values of χ^2_{aim} depend on the scale of the standard uncertainties of the intensities, allowing an estimation of their accuracy.

Analysis of the reconstructed MEM densities has been performed according to Bader's Atoms in Molecules (AIM) theory to derive information about chemical bonds, in particular about covalent bonds and hydrogen bonds. Local maxima of the densities, their associated atomic basins and charges, bond critical points (BCPs) and their densities and second derivatives, *i.e.* the eigenvalues and the Laplacians, have been determined according to the AIM theory.

For all studied compounds it is shown that, providing the employment of the extensions to the MEM, the densities obtained from the MEM exhibit similar properties as those obtained from the multipole method. However, it is demonstrated that the fine features of hydrogen bonds are described more convincingly by the MEM than by the multipole method.

It is shown that the MEM densities of α-glycine and trialanine indicate the same chemical features, while differences between densities at BCPs of similar bonds of the two compounds are caused by the different chemical environment of the bonds compared. By comparison of densities at BCPs from the MEM with the corresponding values from multipole refinement and from quantum chemical calculations it is shown that each method produces similar densities at BCPs. For all studied compounds it is demonstrated that for MEM densities, densities at BCPs of hydrogen bonds possess a larger magnitude than corresponding values from densities by the multipole method, while the opposite is true for covalent bonds. The values of the Laplacians at BCPs, especially of C–O bonds, show larger discrepancies between values from the MEM and from multipole refinement. These differences are caused by thermal motion which is present in dynamic MEM densities, but absent in static densities produced by the multipole method.

The results of the extensive study of electron densities of six amino acids and tripeptides show that densities and energetic properties at BCPs of covalent bonds and hydrogen bonds depend exponentially on their bond lengths. The functions of the dependencies of the densities on the bond lengths differ from the corresponding functions fitted to values from the multipole method. It is demonstrated that the ratio of the potential and kinetic energy densities at BCPs of hydrogen bonds reveals the possibility to classify them according to the distance between hydrogen atom and acceptor atom. Short hydrogen bonds possess covalent character, hydrogen bonds with intermediate distance have a mixed covalent-ionic character and long hydrogen bonds are ionic. This classification coincides with the usual classification of strong, intermediate and weak hydrogen bonds as proposed in the literature. The studied hydrogen bonds are classified as possessing mainly a mixed covalent-ionic character. However, for covalent bonds a classification according to the bond lengths does not suffice to characterize them.

The results of the six amino acids and tripeptides indicate that the prior density contributes a large part to the densities at BCPs. However, for the Laplacians and the energy densities at BCPs the differences between MEM and prior densities are larger than for the densities at BCPs. Densities at BCPs of hydrogen bonds from MEM densities show a different trend in their dependence on the bond distance than corresponding trends from prior densities. Thus, it is demonstrated that the analysis of the true density instead of the prior or procrystal density is recommended in order to extract information about chemical bonding.

It is concluded from the results of the Accurate Charge Density studies reported in the present thesis, that the MEM allows a good characterization of chemical bonds and describes the electron density of hydrogen bonds more realistic than the multipole model. The future application of the MEM is to reconstruct and study charge densities of proteins, where its model-independency provides an advantage.

Chapter 7

Zusammenfassung

Die Multipolmethode liefert eine Beschreibung der Kristallstruktur, die über das einfache spherische Atommodell (ISAM) hinausgeht. Die Effekte auf die Dichte aufgrund chemischer Bindungen können mit der aspherischen Dichte aus der Multipolverfeinerung gegen Röntgenbeugungsdaten erhalten werden. Aufgrund der normalerweise großen Zahl hoch korrelierter Parameter, müssen *ad hoc* Zwangsbedingungen zwischen diesen Parameter eingeführt werden, was zu einer unvollständigen Beschreibung der Elektronendichte führen kann. Dieses Problem wird für größere Elementarzellen verstärkt. Die Maximum-Entropie-Methode (MEM) rekonstruiert Elektronendichten modellunabhängig und leidet nicht unter korrelierten Parametern. Die vorliegende Dissertation zeigt, daß die MEM gut für Akkurate Ladungsdichtestudien geeignet ist.

Akkurate Ladungsdichten mehrerer Aminosäuren und Tripeptide wurden mit der MEM rekonstruiert um chemische Bindungen zu untersuchen. Die Untersuchung von Trialanin wird in Kapitel 3 beschrieben und die Untersuchung von α-Glycin in Kapitel 4. Über eine ausführliche Untersuchung der Elektronendichten von sechs Aminosäuren und Tripeptiden einschließlich 52 Wasserstoffbrückenbindungen wird in Kapitel 5 berichtet. In Kapitel 2 wird der Ablauf bei der Rekonstruktion der MEM-Dichten beschrieben. Die MEM wurde verwendet, um die wahrscheinlichste Elektronendichte zu ermitteln, welche simultan die Informationsentropie maximiert, die Normalisierungsbeschränkung erfüllt und an die Beugungsdaten anpaßt. Für die Bestimmung der wahrscheinlichsten Elektronendichte ist eine iterative Suche nötig, da eine analytische Lösung nicht existiert. Das Computerprogramm BayMEM,

inklusive des Cambridge-Algorithmus, wurde für die Rekonstruktion der optimierten MEM-Dichten verwendet.

Die Qualität der MEM-Dichten wird durch mehrere Erweiterungen der MEM verbessert. Diese Erweiterungen umfassen die Verwendung eines ungleichförmigen Priors (engl. non-unform prior), die Methode der prior-abgeleiteten F-Zwangsbedingung (engl. prior-derived F-constraint), statischer Gewichtung (engl. static weighting) sowie die Wahl des optimalen Stopkriteriums χ^2_{aim} für das Ende der MEM-Iterationen. Diese Erweiterungen werden in Kapitel 2 beschrieben. Der Ablauf für die Bestimmung des optimalen χ^2_{aim} wurde innerhalb dieser Dissertation entwickelt (Kapitel 3 und 4). Es wird demonstriert, daß durch Inaugenscheinnahme der dynamischen Deformationskarten sowie der Differenz-Fourier-Karten der Punkt der Konvergenz der MEM-Iterationen und somit das optimale χ^2_{aim} bestimmt wird. Die Ergebnisse der untersuchten Verbindungen zeigen, daß die Werte von χ^2_{aim} von der Größenordnung der Standardabweichungen der Intensitäten abhängen, wodurch eine Abschätzung ihrer Genauigkeit erlaubt wird.

Die Analyse der rekonstruierten MEM-Dichten wurde entsprechend Baders Theorie über Atome in Molekülen (AIM) durchgeführt, um Informationen über chemische Bindungen, insbesondere über kovalente und Wasserstoffbrückenbindungen, zu erhalten. Lokale Maxima der Dichten, deren assoziierte Atomgebiete (engl. atomic basins) und -ladungen, bindungskritische Punkte (BCPs) und deren Dichten und zweite Ableitungen (Eigenwerte und Laplace-Operator) wurden entsprechend der AIM-Theorie bestimmt.

Für alle untersuchten Verbindungen wird gezeigt, daß, unter Verwendung der Erweiterungen der MEM, die Dichten, die durch die MEM erhalten werden, ähnliche Eigenschaften aufweisen wie die Dichten aus der Multipolmethode. Jedoch wird demonstriert, daß die detailierten Eigenschaften von Wasserstoffbrückenbindungen überzeugender durch die MEM beschrieben werden als durch die Multipolmethode.

Es wird gezeigt, daß die MEM-Dichten von α-Glycin und Trialanin die gleichen chemischen Eigenschaften aufweisen, während Abweichungen der Dichten in BCPs gleicher Bindungstypen beider Verbindungen verursacht werden durch die unterschiedliche chemische Umgebung beider, miteinander verglichenen Bindungen. Durch Vergleich der Dichten in BCPs der MEM mit den entsprechenden Werten der Multipolmethode und von quantenchemischen Berechnungen wird gezeigt, daß jede Methode ähnliche Dichten in BCPs liefert. Für alle untersuchten Verbindungen wird

demonstriert, daß für die MEM-Dichten die Dichten in BCPs von Wasserstoffbrückenbindungen größer sind, als die entsprechenden Werte der Dichten aus der Multipolmethode. Für kovalente Bindungen wird das Gegenteil festgestellt. Die Werte des Laplace-Operators in BCPs, besonders von C–O-Bindungen, zeigen größere Diskrepanzen für die Werte der MEM und der Multipolverfeinerung. Diese Unterschiede werden durch thermische Bewegung verursacht, die in dynamischen MEM-Dichten enthalten ist, jedoch in statischen Dichten, die durch die Multipolmethode erzeugt werden, abwesend ist.

Die Ergebnisse der umfangreichen Untersuchung der Elektronendichten von sechs Aminosäuren und Tripeptiden zeigt, daß Dichten und energetische Eigenschaften in BCPs kovalenter und Wasserstoffbrückenbindungen exponentiell von ihren Bindungslängen abhängen. Die Funktionen der Abhängigkeiten der Dichten von den Bindungslängen unterscheiden sich von den entsprechenden Funktionen, die an die Werte aus der Multipolmethode angepaßt wurden. Es wird demonstriert, daß das Verhältnis der potentiellen und kinetischen Energiedichten an BCPs von Wasserstoffbrückenbindungen die Möglichkeit eröffnet, diese entsprechend der Abstände zwischen Wasserstoffatom und Akzeptoratom zu charakterisieren. Kurze Wasserstoffbrückenbindungen besitzen kovalenten Charakter, Wasserstoffbrückenbindungen mittlerer Länge haben einen gemischt kovalent-ionischen Charakter und lange Wasserstoffbrückenbindungen sind ionisch. Diese Charakterisierung stimmt mit der üblichen Einteilung nach starken, intermediären und schwachen Wasserstoffbrückenbindungen überein, die in der Literatur veröffentlicht wurde. Die untersuchten Wasserstoffbrückenbindungen tragen hauptsächlich einen gemischt kovalent-ionischen Charakter. Jedoch reicht eine Einteilung kovalenter Bindungen entsprechend der Bindungslängen nicht aus, um diese zu charakterisieren.

Die Ergebnisse der sechs Aminosäuren und Tripeptide zeigen, daß die Prior-Dichte einen großen Teil zu den Dichten in BCPs beiträgt. Jedoch sind die Unterschiede zwischen MEM- und Prior-Dichten für die Werte des Laplace-Operators und Energiedichten in den BCPs größer als für die Dichten in BCPs. Dichten in BCPs von Wasserstoffbrückenbindungen der MEM-Dichten zeigen einen anderen Verlauf der Abhängigkeit vom Abstand zwischen Wasserstoffatom und Akzeptoratom als die der Prior-Dichten. Somit wird gezeigt, daß die Untersuchung der wahren Dichte anstelle eines Priors oder Prokristalls für die Gewinnung von Informationen über chemische Bindungen zu empfehlen ist.

Aus den Ergebnissen der Akkuraten Ladungsdichte-Untersuchungen, über die in der vorliegenden Dissertation berichtet wurde, läßt sich schließen, daß die MEM eine gute Charakterisierung chemischer Bindungen erlaubt und die Elektronendichte von Wasserstoffbrückenbindungen realistischer beschreibt als die Multipolmethode. Zukünftige Anwendungen der MEM umfassen die Rekonstruktion und Untersuchung von Dichten von Proteinen, wobei die Modellunabhängigkeit der MEM einen Vorteil bietet.

Appendices

Appendix A

Crystallographic data of amino acids and peptides

Single-crystal x-ray diffraction data of L-alanine (Ala),[79] an L-phenylalanine formic acid complex (Phe),[72] L-alanyl-L-tyrosyl-L-alanine (Ala-Tyr-Ala) with water as solvent[83] and Ala-Tyr-Ala with ethanol as solvent[83] were kindly provided by Destro and Luger and co-workers, who have already reported multipole refinements against these data (Tables A.1,A.2).[79,72,83] Crystallographic data of the studied compounds and a summary of the ISAM refinements and the MEM calculations are given in Tables A.1,A.2.

Table A.1: Crystallographic data of α-glycine (Gly),[74] L-alanine (Ala),[79] phenylalanine formic acid complex (Phe),[72] and summary of the ISAM refinement (present work). Reflections with $I/\sigma(I) > 3$ are classified as observed.

Compound	Gly	Ala	Phe
Chemical formula	$C_2O_2NH_5$	$C_3O_2NH_7$	$C_{18}H_{23}N_2O_4^+ \cdot CHO_2^-$
Space group	$P2_1/n$	$P2_12_12_1$	$P2_1$
Z	4	4	2
a (Å)	5.0866	5.9279	11.4585
b (Å)	11.7731	12.2597	5.5941
c (Å)	5.4595	5.7939	14.2147
β (deg)	111.99	90.00	99.46
V (Å^3)	303.16	421.1	898.8
F(000)	160	192	400
Temperature (K)	23	23	25
Wavelength λ (Å)	0.71073	0.71073	0.71073
$[\sin(\theta)/\lambda]_{\max}$ (Å^{-1})	1.15	1.08	1.18
Crystal size (mm)	0.26	0.18	$0.3 \times 0.4 \times 0.55$
Redundancy	-	-	4.9
R_{int}	-	-	4.02
Completeness	-	-	86.2
Integration software	-	-	-
Refinement software	VALRAY	VALRAY	SHELX, XD
Unique reflections (obs/all)	3483/3822	2328/2535	8971/10981
Multipole refinement[74, 79, 72]			
Number of F_{obs}/number of parameters	-	13.5	14.9
R_F(obs/all)	-/1.29	-/2.03	2.72/3.50
R_{wF}(obs/all)	-/-	-/1.59	3.07/-
GoF	1.04	1.17	1.06
$\Delta\rho$ (electrons/Å^3)	-	-	-0.241/0.293
ISAM refinement			
Number of F_{obs}/number of parameters	73.5	46.0	44.6
R_F(obs/all)	2.33/2.60	2.85/3.16	4.04/4.80
R_{wF}(obs/all)	5.25/5.35	3.73/3.77	4.73/4.87
GoF(obs/all)	2.02/1.96	1.81/1.75	1.43/1.33
$\Delta\rho$ (electrons/Å^3)	-0.33/0.39	-0.29/0.35	-0.33/0.51
MEM calculation[53]			
Number of pixels	$64 \times 144 \times 72$	$72 \times 162 \times 72$	$144 \times 72 \times 192$
Pixel size (Å^3)	0.080 ×0.082 ×0.076	0.082 ×0.076 ×0.080	0.080 ×0.078 ×0.074
χ^2_{aim}	0.3131	0.7600	0.8300
R_F/R_{wF}	1.04/1.53	1.99/1.90	3.55/3.43

Table A.2: Crystallographic data of trialanine (Ala-Ala-Ala),[14] alanyl-tyrosyl-alanine in water (Ala-Tyr-Ala$_{(H_2O)}$)[83] and ethanol (Ala-Tyr-Ala$_{(EtOH)}$)[83] and summary of the ISAM refinement (present work). Reflections with $I/\sigma(I) > 3$ are classified as observed, exception: $F_o/\sigma(F_o) > 4$ for trialanine.[14]

Compound	Ala-Ala-Ala	(Ala-Tyr-Ala$_{(H_2O)}$)	(Ala-Tyr-Ala$_{(EtOH)}$)
Chemical formula	$C_9H_{17}N_3O_4 \cdot H_2O$	$C_{15}H_{21}N_3O_5$ $\cdot$ 2.634 H_2O	$C_{15}H_{21}N_3O_5$ $\cdot$ C_2H_5OH
Space group	C2	$P2_1$	$P2_1$
Z	8	2	2
a (Å)	18.4408	8.121	8.845
b (Å)	5.2153	9.299	9.057
c (Å)	24.8543	12.532	12.364
β (deg)	98.76	91.21	94.56
V (Å^3)	2362.4	946.2	987.3
F(000)	1031.8	397	396
Temperature (K)	20	9	20
Wavelength λ (Å)	0.71073	0.50000	0.71073
$[\sin(\theta)/\lambda]_{\max}$ (Å^{-1})	1.15	1.24	1.11
Crystal size (mm)	$0.3 \times 0.4 \times 0.5$	$0.54 \times 0.25 \times 0.13$	$0.4 \times 0.25 \times 0.2$
Redundancy	5.7	14.4	9.6
R_{int}	2.95	5.35	4.15
Completeness	93.3	91.2	99.4
Integration software	-	XDS	SAINT, SADABS
Refinement software	SHELX, XD	SHELX, XD	SHELX, XD
Unique reflections (obs/all)	12928/14895	12875/14111	10901/11703
Multipole refinement[83,14]			
Number of F_{obs}/number of parameters	13.1	18.9	15.1
R_F(obs/all)	1.83/2.47	2.93/3.51	2.23/2.64
R_{wF}(obs/all)	1.53/-	2.08/-	1.77/-
GoF	0.67	2.06	1.63
$\Delta\rho$ (electrons/Å^3)	-	-0.360/0.329	-0.235/0.273
ISAM refinement			
Number of F_{obs}/number of parameters	-	55.1	48.7
R_F(obs/all)	-	3.99/4.55	3.60/4.00
R_{wF}(obs/all)	-	4.78/4.94	4.19/4.25
GoF(obs/all)	-	1.67/1.64	2.27/2.22
$\Delta\rho$ (electrons/Å^3)	-	-0.30/0.56	-0.26/0.55
MEM calculation[42]			
Number of pixels	$216 \times 64 \times 324$	$96 \times 108 \times 162$	$108 \times 108 \times 162$
Pixel size (Å^3)	0.085 $\times$0.081 $\times$0.077	0.085 $\times$0.086 $\times$0.077	0.082 $\times$0.084 $\times$0.076
χ^2_{aim}	0.4250	1.2750	1.3081
R_F/R_{wF}	2.63/1.84	3.30/3.42	2.48/2.22

Appendix B

Topological properties of α-glycine

The procedure to obtain accurate charge density of α-glycine and its analysis are described in detail in Chapter 4.[53] Crystallographic data and a summary of the ISAM refinement and the MEM calculation of α-glycine are given in Table A.1. Table B.1 displays the geometrical, topological and energetic properties of hydrogen bonds of α-glycine, which are discussed in detail in Chapter 5. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance d(H$\cdots$O).

Table B.1: Geometrical, topological and energetic properties of hydrogen bonds of α-glycine (Figure 4.1). d(H$\cdots$O) and d(D$\cdots$O) in Å, $\rho(BCP)$ in electrons/Å^3, eigenvalues and $\nabla^2\rho(BCP)$ in electrons/Å^5, G, V and H in kJ/mol·Bohr^3.

	$\rho(BCP)$	λ_1	λ_2	λ_3	$\nabla^2\rho(BCP)$	d(H$\cdots$O)	d(D$\cdots$O)	G	V	$\lvert V\rvert/G$	H
O(1)$^6\cdots$H(1)^4–N^4	0.303	-1.87	-1.32	6.17	2.98	1.756	2.764	96.75	-112.37	1.16	-15.62
O(2)$^1\cdots$H(2)^3–N^3	0.249	-1.93	-0.31	6.18	3.94	1.826	2.833	102.36	-97.36	0.95	5.00
O(2)$^1\cdots$H(3)^2–N^2	0.201	-1.39	-1.02	3.30	0.90	2.027	3.006	37.84	-51.18	1.35	-13.34
O(1)$^3\cdots$H(3)^2–N^2	0.108	-1.02	-1.01	0.36	-1.67	2.393	2.937	-22.75	-0.10	0.00	-22.85
O(1)$^6\cdots$H(4)^1–C(2)^1	0.129	-0.87	-0.55	1.23	-0.19	2.392	3.304	6.78	-18.84	2.78	-12.06
O(2)$^5\cdots$H(4)^1–C(2)^1	0.124	-1.49	-0.95	1.14	-1.31	2.343	3.218	-14.12	-7.35	-0.52	-21.47

Appendix C

Accurate charge density of L-alanine by the Maximum Entropy Method (MEM)

C.1 MEM calculations

Single-crystal X-ray diffraction data of L-alanine (Ala)[79] (Figure C.1) were kindly provided by Destro and co-workers, who have already reported multipole refinement against these data (Table A.1).[79] We have used these data to perform refinements of the independent spherical atom model (ISAM) with the computer program JANA2000.[60]

The coordinates and atomic displacement parameters (ADP) of Ala obtained by the multipole refinement were used as starting model for the ISAM refinement. Hydrogen atoms were fixed at distances known from neutron diffraction[76] (X_3–C–H: 1.099 Å, C–C–H_3: 1.059 Å, C_2–C–H_2: 1.092 Å, -NH_3^+: 1.033 Å, X_2–N–H: 1.009 Å). The riding model with $U_{iso}(H) = 1.2U_{eq}(N,C)$ was employed to calculate the ADP of hydrogen atoms. Coordinates of hydrogen atoms were obtained by attachment to their neighbor atoms with tetrahedral angle restraints according to chemical meaning.

According to a procedure by Bagautdinov,[75] the ISAM refinement was employed to obtain phased and scaled structure factors which are corrected for anomalous scattering and used for the MEM calculations. The coordinates and ADP obtained by the ISAM refinement were used to compute the procrystal electron density (prior density $\rho^{prior}(\mathbf{x})$) with the computer program PRIOR.[105] This prior density $\rho^{prior}(\mathbf{x})$ is calculated on a grid over the unit cell and used as reference density in the MEM calculations. The grid for the prior and MEM density has to be chosen in that way that the pixel size does not exceed

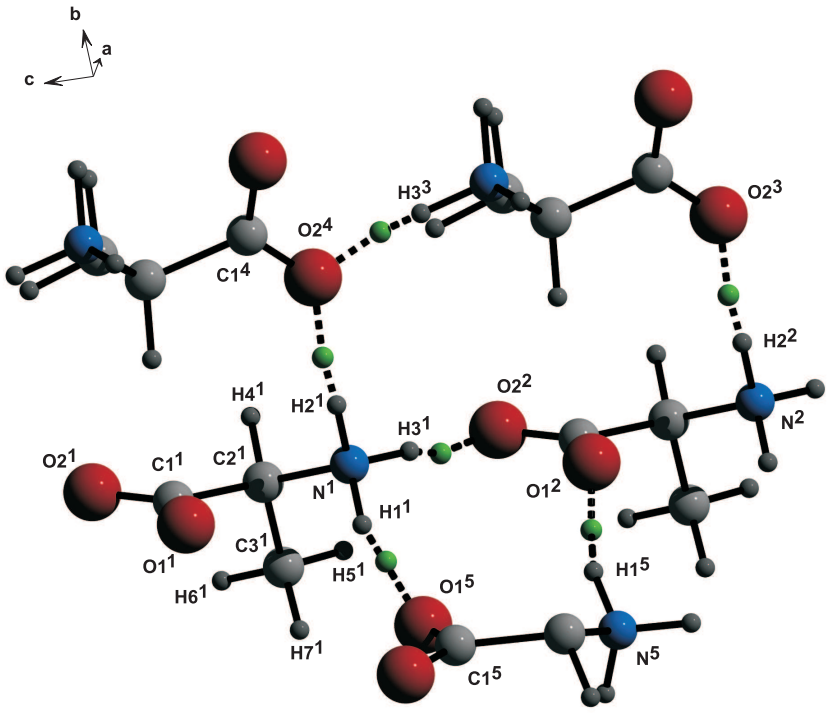

Figure C.1: Perspective view of L-alanine obtained with atomic coordinates from the ISAM. Hydrogen bonds are indicated by dotted lines, green dots indicate BCPs of hydrogen bonds. Symmetry operators are denoted by superscripts: $1:(x, y, z)$; $2:(x, y, -1+z)$; $3:(0.5+x, 0.5-y, -z)$; $4:(0.5+x; 0.5-y, 1-z)$; $5:(1.5-x; -y, -0.5+z)$.

0.1×0.1×0.1 Å^3 (Table A.1).

For the iterative procedure of the historical MEM the summation of Eq. 4.2 has been extended towards all reflections up to $\sin(\theta)/\lambda = 2.5$ Å^{-1} according to the method of prior-derived F-constraint (PDC).[51] Static weights of H4 have been chosen according to de Vries *et al.*[52] employing Eq. 4.3 (Chapter 4). MEM calculations have been performed with the computer program BayMEM[45] employing the Cambridge maximum entropy algorithm.[46,106]

For Ala $\chi^2_{aim} = 0.76$ has been determined by the classical MEM calculation[53] and used as χ^2_{aim} for the historical MEM calculation. The resulting difference Fourier map shows to be featureless, while the difference map $\rho^{MEM} - \rho^{prior}$ has smooth contour lines indicating that no noise of the data has been fitted (Figure C.2). According to the inspection of the maps $\chi^2_{aim} = 0.76$ has been found applicable for Ala. The determination of χ^2_{aim} for Ala followed the procedure applied to α-glycine[53] (Chapter 4).

C.2 Analysis of the MEM density

Atomic positions obtained from ρ^{MEM} are given in Table C.1. Hydrogen atom H(2) does not exhibit an atomic maximum. Its atomic volume and charge is incorporated into the basin of the neighboring nitrogen atom (Table C.2). Volumes of the atomic basins add up to 421.1 Å^3 which is in perfect agreement with volume of the unit cell (Table A.1). The number of electrons of the unit cell of 192 indicates that all electrons are incorporated into atomic basins. The net charge Q of the unit cell is determined zero fulfilling the required electrical neutrality.

The density at BCPs of covalent bonds from the MEM density is in general smaller than the values from multipole refinement (Table C.3), which has also been observed for trialanine[42] (Chapter 3) and α-glycine[53] (Chapter 4). The values of the Laplacian of C–C and C–N bonds from the MEM have the same magnitude as from multipole refinement, whereas the values of C–O bonds differ significantly (Table C.3) as in case of trialanine[42] (Chapter 3) and α-glycine[53] (Chapter 4). The Laplacians of C–H and N–H bonds from the MEM have a larger magnitude than from multipole refinement. The geometrical, topological and energetic properties for hydrogen bonds are given in Table C.4 and are discussed in detail in Chapter 5. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance d(H$\cdots$O).

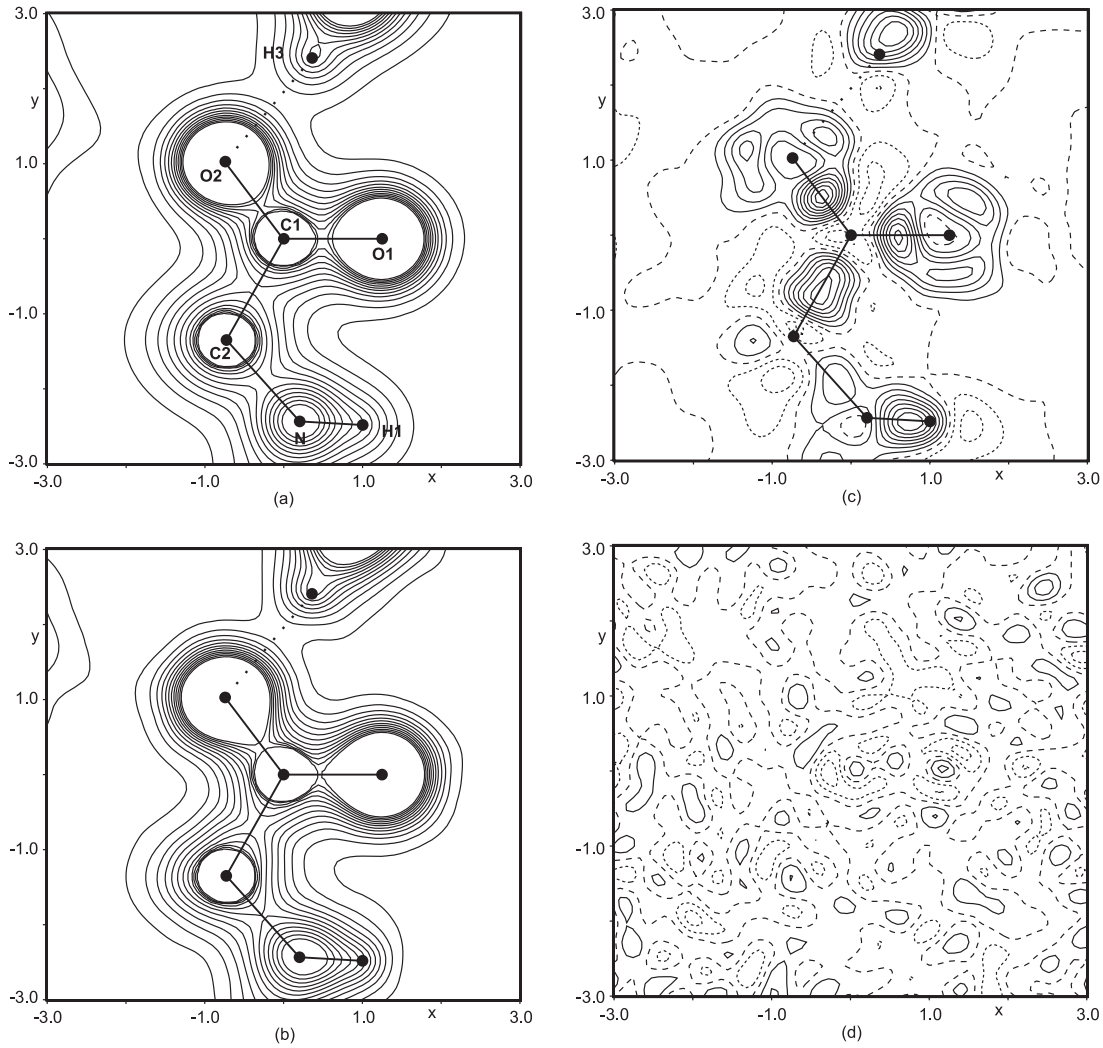

Figure C.2: Sections of 6×6 Å^2 through the C1–O1–O2 plane of L-alanine. (a) ρ^{prior}. (b) ρ^{MEM}. (c) Difference map with $\Delta\rho$(min/max) = -0.14/0.44 electrons/Å^3. (d) Difference Fourier map with $\Delta\rho$(min/max) = -0.14/0.13 electrons/Å^3. (a) and (b): Contour of equal density are given from 0.2 to 2.5 electrons/Å^3 in steps of 0.2 electrons/Å^3. (c) and (d): Contour lines are at intervals of 0.05 electrons/Å^3. Solid lines are contours of positive value, dotted lines are negative contours, and dashed lines represent the contour of zero value.

Table C.1: Atomic positions derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms.

Atom	x	y	z
C(1)	0.945657	0.859016	0.099304
C(2)	0.533506	0.660989	0.145180
C(3)	0.739831	0.590723	0.196519
N	0.647472	0.137340	0.182459
O(1)	0.773361	0.916237	0.124519
O(2)	0.941407	0.315748	0.238046
H(1)	0.692835	0.064971	0.194546
H(3)	0.588951	0.150138	0.029854
H(4)	0.578373	0.744577	0.161457
H(5)	0.201235	0.108027	0.145340
H(6)	0.861761	0.605640	0.077335
H(7)	0.697833	0.512264	0.191743

Table C.2: Volumes of the atomic basins and the number of electrons in the atomic basins and the net atomic charge (Q) in the asymmetric unit for $\rho^{MEM}(\mathbf{x})$. The basin of nitrogen encompasses the hydrogen atom H(2) because the latter atom does not constitute a local maximum in the density.

Atom	Volume	Electrons	Q
C(1)	4.96	4.54	+1.46
C(2)	9.58	5.94	+0.06
C(3)	12.69	6.57	−0.57
NH_1	21.97	9.12	−1.12
O(1)	15.54	8.90	−0.90
O(2)	13.96	8.98	−0.98
H(1)	2.21	0.37	+0.63
H(3)	1.81	0.46	+0.54
H(4)	4.71	0.70	+0.30
H(5)	6.83	0.86	+0.14
H(6)	6.04	0.87	+0.13
H(7)	4.97	0.69	+0.31
Total	105.27	48.00	0.00

Table C.3: Electron densities and Laplacians at the BCPs of covalent bonds of L-alanine. Values are given for $\rho(BCP)$ (electrons/Å^3; first line) and $\nabla^2\rho(BCP)$ (electrons/Å^5; second line) derived from $\rho^{prior}(\mathbf{x})$, $\rho^{MEM}(\mathbf{x})$ and the static electron density of the multipole model.[74]

Bond	Prior	MEM	Multipole
C1–O1	2.08	2.36	3.02
	12.39	18.36	-39.0
C1–O2	2.01	2.31	2.86
	7.58	12.69	-29.5
C1–C2	1.17	1.50	1.76
	0.38	-9.09	-10.9
C2–C3	1.20	1.51	1.67
	0.18	-13.08	-10.1
C2–N	1.38	1.57	1.70
	2.65	-7.72	-11.0
C2–H4	1.19	1.56	1.66
	-5.29	-14.24	-11.0
C3–H5	1.27	1.60	1.66
	-8.09	-18.48	-11.0
C3–H6	1.27	1.52	1.66
	-7.93	-13.32	-11.0
C3–H7	1.28	1.62	1.66
	-8.39	-21.09	-11.0
N–H1	1.47	-	1.80
	-14.28	-	-13.2
N–H2	1.47	1.84	1.80
	-14.94	-43.83	-13.2
N–H3	1.46	1.78	1.80
	-13.51	-29.22	-13.2

Table C.4: Geometrical, topological and energetic properties of hydrogen bonds of L-alanine (Figure C.1). d(H$\cdots$O) and d(D$\cdots$O) in Å, $\rho(BCP)$ in electrons/Å^3, eigenvalues and $\nabla^2\rho(BCP)$ in electrons/Å^5, G, V and H in kJ/mol·Bohr^3.

	$\rho(BCP)$	λ_1	λ_2	λ_3	$\nabla^2\rho(BCP)$	d(H$\cdots$O)	d(D$\cdots$O)	G	V	$\lvert V\rvert/G$	H
N–H3$\cdots$O2	0.326	-3.40	-3.08	5.01	-1.46	1.764	2.785	21.65	-83.10	3.84	-61.45
N–H2$\cdots$O2	0.298	-2.27	-1.91	4.31	0.14	1.834	2.815	44.09	-84.32	1.91	-40.24
N–H1$\cdots$O1	0.251	-1.91	-1.46	4.91	1.54	1.833	2.831	59.08	-76.34	1.29	-17.26

Appendix D

Accurate charge density of the L-phenylalanine formic acid complex by the Maximum Entropy Method (MEM)

D.1 MEM calculations

Single-crystal X-ray diffraction data of an L-phenylalanine formic acid complex (Phe)[72] (Figures D.1,D.2) were kindly provided by Luger and co-workers, who have already reported multipole refinements against these data (Table A.1).[72] We have used these data to perform refinements of the independent spherical atom model (ISAM) with the computer program JANA2000.[60]

For the ISAM refinement of Phe the coordinates and ADP obtained by the multipole refinement were used starting model. Hydrogen atoms were fixed at distances known from neutron diffraction[76,72] (X_3–C–H: 1.099 Å, C_2–C–H_2: 1.092 Å, -NH_3^+: 1.033 Å, O(2b)–H(2b): 1.05 Å). The riding model with U_{iso}(H)= 1.2U_{eq}(N,C) and U_{iso}(H)= 1.5U_{eq}(O) was employed to calculate the ADP of hydrogen atoms. The coordinates of H(2b) of the hydroxyl group were refined freely being subject to the distances known from neutron diffraction, since reasonable geometrical restraint was not available. Coordinates of all other hydrogen atoms were obtained by attachment to their neighbor atoms with tetrahedral or trigonal angle restraints according to chemical meaning.

According to a procedure by Bagautdinov,[75] the ISAM refinement was employed to

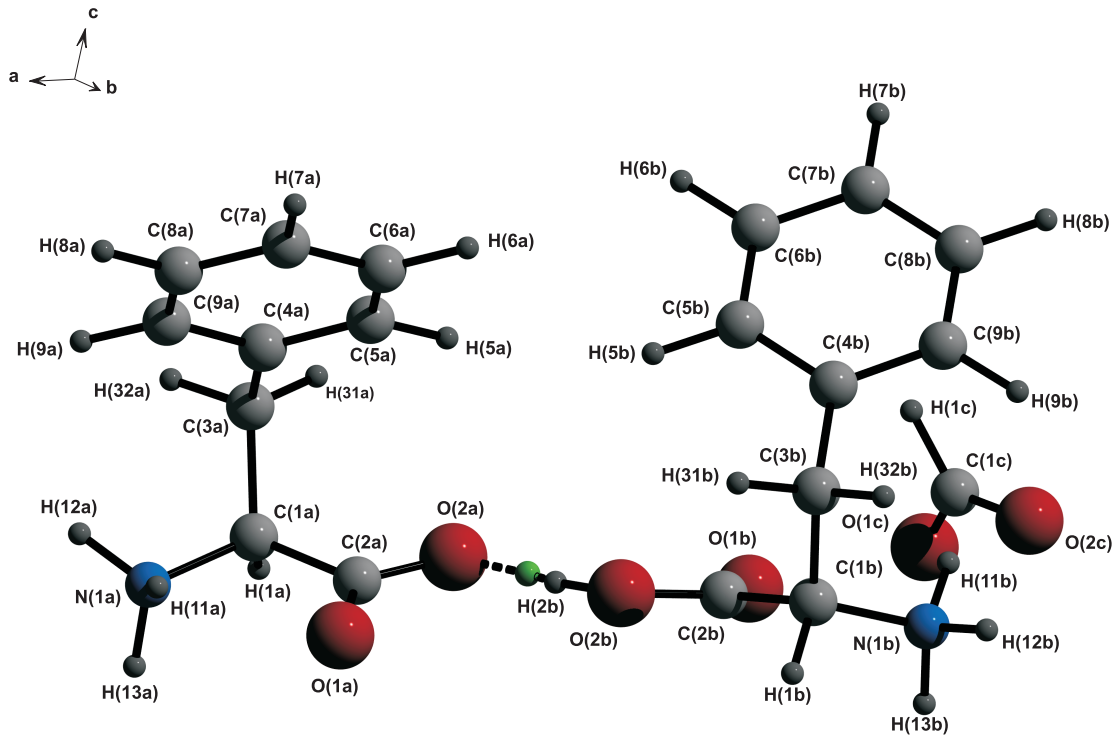

Figure D.1: Perspective view of the L-phenylalanine formic acid complex obtained with atomic coordinates from the ISAM, displaying the atoms of the asymmetric unit. The strongest hydrogen bond determined for this structure is indicated by dotted lines; a green dot indicates the BCP of that hydrogen bond.

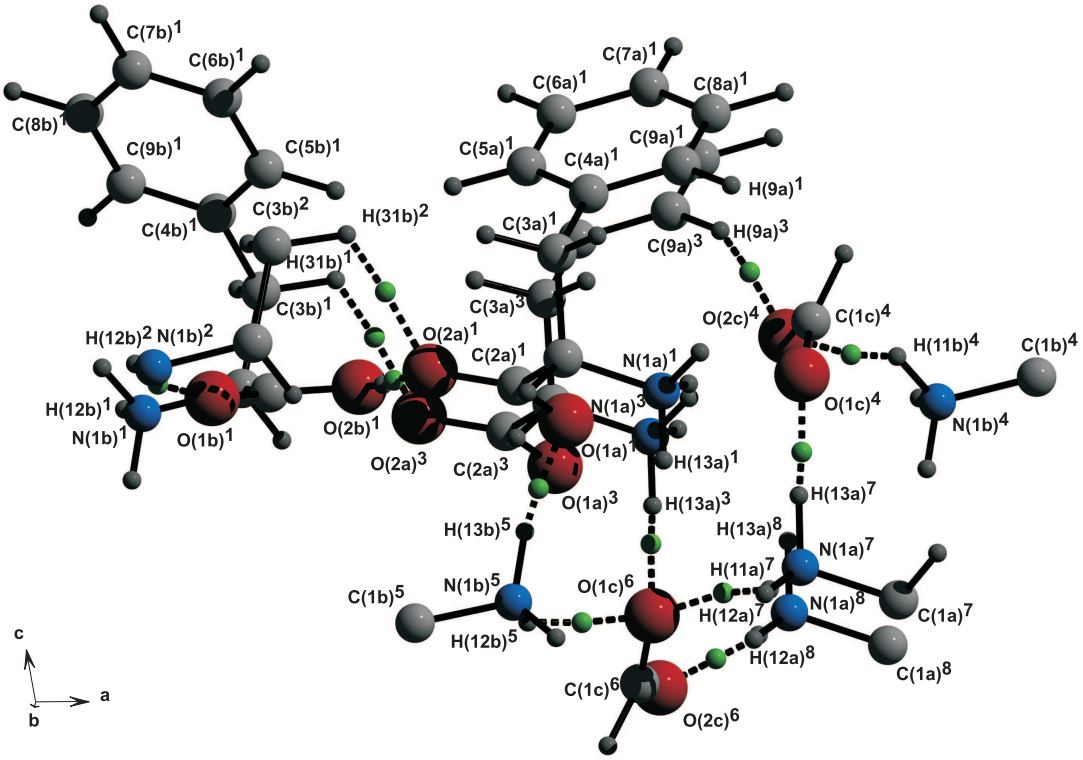

Figure D.2: Perspective view of the L-phenylalanine formic acid complex obtained with atomic coordinates from the ISAM, displaying hydrogen bonds. Hydrogen bonds are indicated by dotted lines, green dots indicate BCPs of hydrogen bonds. Symmetry operators are denoted by superscripts: $1:(x,y,z)$; $2:(x,-1+y,z)$; $3:(x,1+y,z)$; $4:(1+x;y,z)$; $5:(1-x;-0.5+y,-z)$; $6:(1-x;0.5+y,-z)$; $7:(2-x;0.5+y,-z)$; $5:(2-x;1.5+y,-z)$.

obtain phased and scaled structure factors which are corrected for anomalous scattering and used for the MEM calculations. The coordinates and ADP obtained by the ISAM refinement were used to compute the procrystal electron density (prior density $\rho^{prior}(\mathbf{x})$) with the computer program PRIOR.[105] This prior density $\rho^{prior}(\mathbf{x})$ is calculated on a grid over the unit cell and used as reference density in the MEM calculations. The grid for the prior and MEM density has to be chosen in that way that the pixel size does not exceed $0.1\times0.1\times0.1$ Å^3 (Table A.1).

For the iterative procedure of the historical MEM the summation of Eq. 4.2 has been extended towards all reflections up to $\sin(\theta)/\lambda = 2.5$ Å^{-1} according to the method of prior-derived F-constraint (PDC).[51] Static weights of H4 have been chosen according to de Vries *et al.*[52] employing Eq. 4.3 (Chapter 4). MEM calculations have been performed with the computer program BayMEM[45] employing the Cambridge maximum entropy algorithm.[46,106]

By the classical MEM[53] for Phe, $\chi^2_{aim} = 0.83$ has been determined and used for the historical MEM[53] calculation. After inspection of the difference Fourier map and the difference map $\rho^{MEM} - \rho^{prior}$ (Figure D.3), $\chi^2_{aim} = 0.83$ has been found to be applicable for Phe. χ^2_{aim} for Phe could be determined the same way as for Ala (Appendix C). That procedure followed the one applied to α-glycine[53] (Chapter 4).

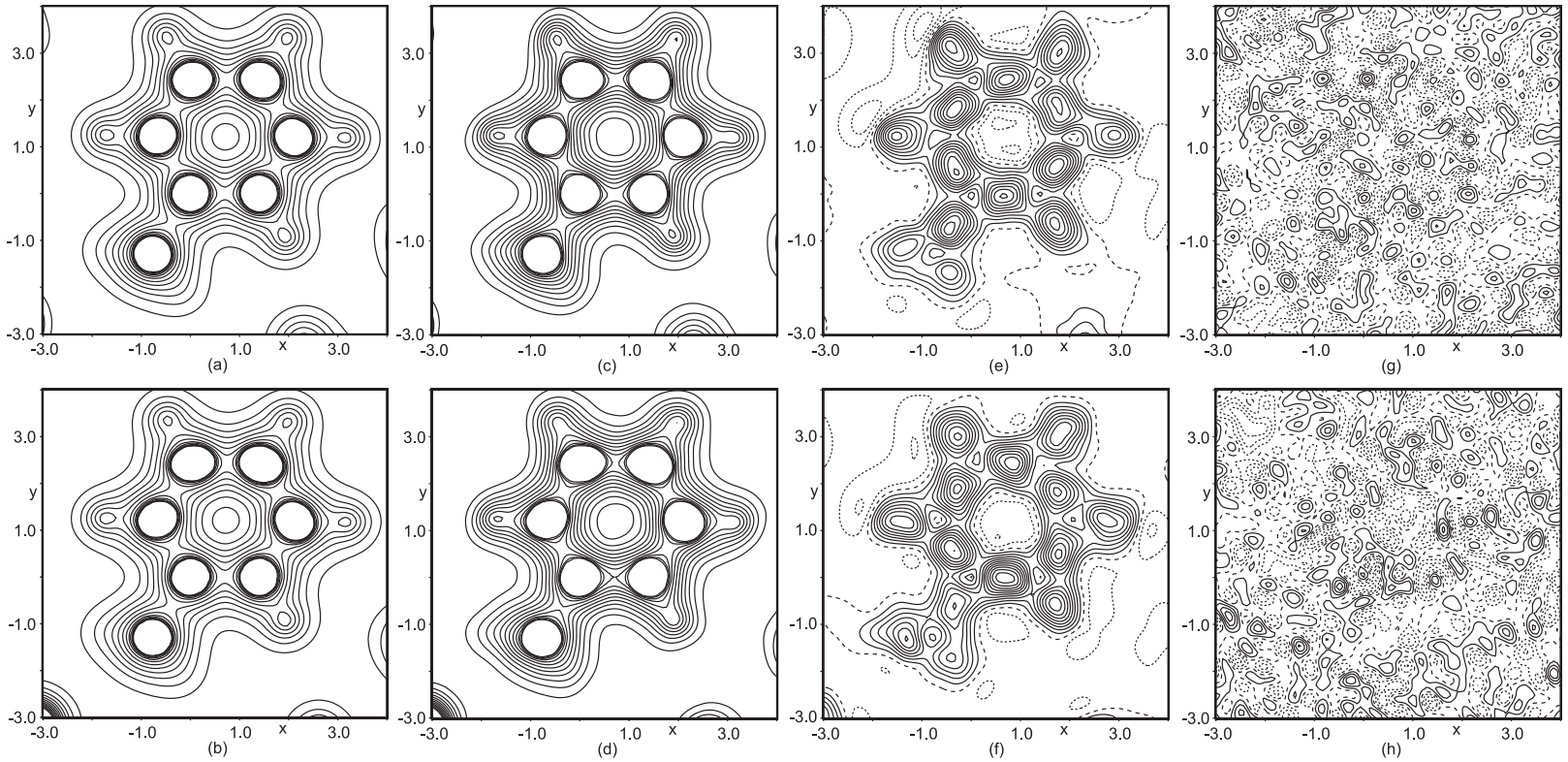

Figure D.3: Sections of 7×7 Å^2 through the phenyl ring plane of the L-phenylalanine formic acid complex. (a) ρ^{prior} of molecule A. (b) ρ^{prior} of molecule B. (c) ρ^{MEM} of molecule A. (d) ρ^{MEM} of molecule B. (e) Difference map of molecule A with $\Delta\rho$(min/max) = -0.18/0.48 electrons/Å^3. (f) Difference map of molecule B with $\Delta\rho$(min/max) = -0.11/0.54 electrons/Å^3. (g) Difference Fourier map of molecule A with $\Delta\rho$(min/max) = -0.22/0.19 electrons/Å^3. (h) Difference Fourier map of molecule B with $\Delta\rho$(min/max) = -0.21/0.25 electrons/Å^3. (a) - (d): Contour of equal density are given from 0.2 to 2.5 electrons/Å^3 in steps of 0.2 electrons/Å^3. (e) - (h): Contour lines are at intervals of 0.05 electrons/Å^3. Solid lines are contours of positive value, dotted lines are negative contours, and dashed lines represent the contour of zero value.

D.2 Analysis of the MEM density

Atomic positions obtained from ρ^{MEM} are given in Tables D.1,D.2,D.3. Hydrogen atoms of ammonium groups, atoms H(2b) and H(6b) do not exhibit atomic maxima. Their atomic volumes and charges are incorporated into the basins of their neighboring atoms (Table D.4). Volumes of the atomic basins add up to 898.8 Å^3 which is in perfect agreement with volume of the unit cell (Table A.1). The number of electrons of the unit cell of 400 indicates that all electrons are incorporated into atomic basins. The net charge Q of the unit cell is determined -0.02, which is almost zero, fulfilling the required electrical neutrality.

The density at BCPs of covalent bonds from the MEM density is in general smaller than the values from multipole refinement (Table D.5) as also reported for trialanine[42] (Chapter 3) and α-glycine[53] (Chapter 4). The calculated values for the density at BCPs are almost always smaller than from multipole refinement, indicating that multipole refinement might overestimate $\rho(BCP)$. This trend can also be observed in case of heterogenous bonds of trialanine (Chapter 3). In general, the values of the Laplacian of C–C and C–N bonds from the MEM have the same magnitude as from multipole refinement, whereas the values of C–O bonds differ significantly (Table D.5). In case of C–O bonds, the magnitude of the Laplacian from multipole refinement is larger than the calculated values. The geometrical, topological and energetic properties for hydrogen bonds are given in Table D.6 and are discussed in detail in Chapter 5. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance d(H$\cdots$O).

Table D.1: Atomic positions for molecule A derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms.

Atom	x	y	z
C(1a)	0.811676	0.539455	0.115762
C(2a)	0.752999	0.785325	0.099240
C(3a)	0.826381	0.461259	0.220893
C(4a)	0.884671	0.646428	0.290818
C(5a)	0.817244	0.820540	0.328027
C(6a)	0.871215	0.996409	0.390043
C(7a)	0.993909	0.999820	0.416267
C(8a)	0.061542	0.825437	0.380349
C(9a)	0.007479	0.649695	0.318224
N(1a)	0.927234	0.543349	0.080376
O(1a)	0.810684	0.951868	0.070292
O(2a)	0.647648	0.797785	0.116050
H(1a)	0.758481	0.414829	0.076287
H(5a)	0.729434	0.819637	0.309468
H(6a)	0.821000	0.120599	0.414991
H(7a)	0.033029	0.125404	0.461655
H(8a)	0.149983	0.828163	0.400180
H(9a)	0.056887	0.522979	0.292797
H(31a)	0.745691	0.420838	0.237185
H(32a)	0.874301	0.305478	0.229716

Table D.2: Atomic positions for molecule B derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms.

Atom	x	y	z
C(1b)	0.424593	0.503760	0.134453
C(2b)	0.451917	0.240916	0.117130
C(3b)	0.473026	0.577283	0.238003
C(4b)	0.431683	0.419972	0.312277
C(5b)	0.497110	0.218614	0.347771
C(6b)	0.457697	0.069808	0.414973
C(7b)	0.352201	0.121786	0.447655
C(8b)	0.286392	0.321954	0.412950
C(9b)	0.326492	0.469955	0.346010
N(1b)	0.295489	0.546385	0.112114
O(1b)	0.375057	0.087334	0.104934
O(2b)	0.565216	0.208811	0.120099
H(1b)	0.463423	0.608852	0.088950
H(5b)	0.574049	0.181278	0.324403
H(7b)	0.323766	0.015884	0.496392
H(8b)	0.211737	0.360207	0.436074
H(9b)	0.278593	0.610931	0.320758
H(31b)	0.562402	0.577616	0.247216
H(32b)	0.450148	0.749671	0.248170

Table D.3: Atomic positions for molecule C derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms.

Atom	x	y	z
C(1c)	0.143737	0.053748	0.174525
O(1c)	0.116318	0.895604	0.111903
O(2c)	0.113910	0.269224	0.169078
H(1c)	0.197327	0.999703	0.236999

Table D.4: Volumes of the atomic basins and the number of electrons in the atomic basins and the net atomic charge (Q) in the asymmetric unit for $\rho^{MEM}(\mathbf{x})$. The basins of C(6b), N(1a,b) and O(2b) encompass their bonded hydrogen atoms because the latter atoms do not constitute local maxima in the density.

	Molecule A			Molecule B		
Atom	Volume	Electrons	Q	Volume	Electrons	Q
C(1)	6.06	5.74	+0.26	6.63	5.95	+0.05
C(2)	5.58	4.74	+1.26	5.36	4.74	+1.25
C(3)	12.54	6.62	−0.62	14.18	6.85	−0.85
C(4)	11.61	6.26	−0.26	10.68	6.25	−0.25
C(5)	15.48	6.33	−0.33	12.64	6.06	−0.06
C(6)	10.63	6.08	−0.08	-	-	-
$C(6)H_1$	-	-	-	20.52	6.94	+0.06
C(7)	11.45	6.17	−0.17	11.34	6.27	−0.27
C(8)	13.49	6.33	−0.33	12.95	6.23	−0.23
C(9)	10.94	6.04	−0.04	11.25	6.15	−0.15
$N(1)H_3$	23.65	9.71	+0.29	21.79	9.93	+0.37
O(1)	16.05	8.88	−0.88	16.85	8.85	−0.85
O(2)	13.27	8.95	−0.95	-	-	-
$O(2)H_1$	-	-	-	17.02	9.37	−0.37
H(1)	7.51	0.92	+0.08	7.23	0.85	+0.15
H(5)	4.55	0.61	+0.39	6.26	0.66	+0.34
H(6)	7.67	0.70	+0.30	-	-	-
H(7)	9.01	0.97	+0.03	10.21	0.85	+0.15
H(8)	4.17	0.58	+0.42	4.59	0.65	+0.35
H(9)	7.17	0.92	+0.08	5.95	0.72	+0.28
H(31)	3.88	0.70	+0.30	3.36	0.60	+0.30
H(32)	6.58	0.78	+0.22	4.65	0.70	+0.30
Total A,B	201.28	88.02	−0.02	202.76	88.34	+0.65
Molecule C						
C(1c)	7.65	4.88	+1.12			
O(1c)	14.97	8.90	−0.90			
O(2c)	14.61	8.93	−0.93			
H(1c)	8.11	0.92	+0.08			
Total C	45.34	23.63	−0.63			
Total	**449.39**	**200**	**-0.01**			

Table D.5: Electron densities and Laplacians at the BCPs of covalent bonds and RCPs (ring critical points) of the L-phenylalanine formic acid complex. Values are given for $\rho(BCP)$ (electrons/Å^3; first line) and $\nabla^2\rho(BCP)$ (electrons/Å^5; second line) derived from $\rho^{prior}(\mathbf{x})$, $\rho^{MEM}(\mathbf{x})$, the static electron density of the multipole model[72] and DFT calculations (B3LYP/6-311++G**).[72]

Bond	Prior	MEM	Multipole	Calc.
C(2a)–O(1a)	2.11	2.51	2.95	2.59
	18.24	16.13	-31.1	-11.2
C(2a)–O(2a)	2.00	2.42	3.02	2.48
	10.39	5.33	-31.7	-12.5
C(2b)–O(1b)	2.19	2.73	3.33	2.75
	24.38	18.38	-53.5	-9.6
C(2b)–O(2b)	1.87	2.20	2.47	2.25
	6.43	-8.21	-20.2	-10.6
C(1c)–O(1c)	2.11	2.48	3.00	2.55
	23.73	20.90	-47.0	-11.1
C(1c)–O(2c)	2.08	2.33	2.71	2.51
	16.18	22.19	-22.3	-8.4
C(1a)–N(1a)	1.37	1.65	1.83	1.54
	1.69	-7.50	-12.5	-10.7
C(1b)–N(1b)	1.39	1.60	1.72	1.65
	2.87	-5.52	-6.1	-14.2
C(1a)–C(2a)	1.18	1.45	1.67	1.69
	0.26	-3.68	-14.3	-14.2
C(1b)–C(2b)	1.19	1.40	1.62	1.71
	-0.08	-6.51	-6.7	-14.6
C(1a)–C(3a)	1.18	1.47	1.70	1.60
	-0.03	-4.98	-10.4	-12.5
C(1b)–C(3b)	1.18	1.49	1.74	1.60
	0.45	-3.42	-10.0	-12.5
C(3a)–C(4a)	1.23	1.70	1.84	1.69
	-0.03	-12.59	-12.0	-14.2
C(3b)–C(4b)	1.24	1.60	1.84	1.70
	-0.99	-11.34	-12.2	-14.3
C(4a)–C(5a)	1.45	1.89	2.17	2.05

Continued on Next Page...

Table D.5 – Continued

Bond	Prior	MEM	Multipole	Calc.
	-3.17	-14.10	-16.9	-20.0
C(4b)–C(5b)	1.45	1.90	2.17	2.05
	-2.40	-16.46	-17.0	-20.0
C(5a)–C(6a)	1.46	1.92	2.19	2.08
	-2.71	-17.03	-16.6	-20.6
C(5b)–C(6b)	1.47	1.93	2.19	2.07
	-3.64	-17.47	-16.6	-20.4
C(6a)–C(7a)	1.47	1.92	2.20	2.09
	-2.65	-17.92	-16.7	-20.9
C(6b)–C(7b)	1.50	1.98	2.19	2.08
	-2.78	-15.22	-16.5	-20.7
C(7a)–C(8a)	1.48	1.84	2.20	2.09
	-3.59	-10.49	-16.7	-20.8
C(7b)–C(8b)	1.48	1.90	2.20	2.08
	-2.73	-16.49	-16.7	-20.7
C(8a)–C(9a)	1.45	1.90	2.21	2.07
	-2.43	-17.80	-16.9	-20.5
C(8b)–C(9b)	1.46	1.92	2.21	2.08
	-3.54	-15.11	-16.9	-20.7
C(9a)–C(4a)	1.46	1.92	2.24	2.07
	-2.29	-16.32	-18.2	-20.2
C(9b)–C(4b)	1.46	2.00	2.24	2.07
	-3.50	-16.99	-18.2	-20.4
C(1a)–H(1a)	1.18	1.66	-	-
	-5.16	-20.65	-	-
C(1b)–H(1b)	1.19	1.57	-	-
	-7.28	-20.54	-	-
C(1c)–H(1c)	1.15	1.51	-	-
	-5.70	-11.09	-	-
C(3a)–H(31a)	1.20	1.66	-	-
	-7.52	-18.09	-	-
C(3b)–H(31b)	1.19	1.65	-	-
	-6.22	-21.39	-	-
C(3a)–H(32a)	1.20	1.62	-	-

Continued on Next Page...

Table D.5 – Continued

Bond	Prior	MEM	Multipole	Calc.
	-6.36	-18.23	-	-
C(3b)–H(32b)	1.20	1.53	-	-
	-6.80	-15.19	-	-
C(5a)–H(5a)	1.16	1.58	-	-
	-6.25	-16.03	-	-
C(5b)–H(5b)	1.18	1.56	-	-
	-8.06	-16.11	-	-
C(6a)–H(6a)	1.16	1.58	-	-
	-7.97	-15.86	-	-
C(6b)–H(6b)	1.14	1.53	-	-
	-6.84	-22.09	-	-
C(7a)–H(7a)	1.15	1.56	-	-
	-6.28	-18.81	-	-
C(7b)–H(7b)	1.15	1.55	-	-
	-8.71	-16.07	-	-
C(8a)–H(8a)	1.15	1.48	-	-
	-6.49	-15.48	-	-
C(8b)–H(8b)	1.16	1.53	-	-
	-8.92	-17.13	-	-
C(9a)–H(9a)	1.17	1.59	-	-
	-7.67	-19.47	-	-
C(9b)–H(9b)	1.17	1.59	-	-
	-6.37	-24.13	-	-
Phenyl ring A	0.32	0.23	-	-
	2.08	1.98	-	-
Phenyl ring B	0.32	0.24	-	-
	2.06	2.39	-	-

Table D.6: Geometrical, topological and energetic properties of hydrogen bonds of L-phenylalanine formic acid complex (Figure D.2). d(H$\cdots$O) and d(D$\cdots$O) in Å, $\rho(BCP)$ in electrons/Å^3, eigenvalues and $\nabla^2\rho(BCP)$ in electrons/Å^5, G, V and H in kJ/mol·Bohr^3.

	$\rho(BCP)$	λ_1	λ_2	λ_3	$\nabla^2\rho(BCP)$	d(H$\cdots$O)	d(D$\cdots$O)	G	V	$\lvert V\rvert/G$	H
O(2b)–H(2b)$\cdots$O(2a)	0.599	-6.40	-4.28	11.00	0.32	1.451	2.490	138.93	-269.25	1.94	130.31
N(1b)–H(13b)$\cdots$O(1a)	0.402	-2.84	-2.38	8.03	2.81	1.708	2.729	119.61	-162.55	1.36	-42.94
N(1a)–H(13a)$\cdots$O(1c)	0.331	-2.38	-2.00	5.85	1.48	1.805	2.820	76.25	-112.30	1.47	-36.04
N(1a)–H(12a)$\cdots$O(2c)	0.316	-2.68	-2.05	6.48	1.75	1.776	2.762	77.72	-107.79	1.39	-30.08
N(1b)–H(11b)$\cdots$O(2c)	0.216	-1.26	-1.03	5.39	3.09	1.892	2.818	80.42	-76.56	0.95	3.87
N(1a)–H(11a)$\cdots$O(1c)	0.129	-0.93	-0.60	3.53	2.00	1.987	2.908	46.56	-38.69	0.83	7.86
N(1b)–H(12b)$\cdots$O(1c)	0.128	-0.98	-0.40	2.82	1.43	2.067	2.834	36.19	-33.35	0.92	2.84
C(9a)–H(9a)$\cdots$O(2c)	0.111	-0.51	-0.43	1.86	0.91	2.355	3.371	24.57	-24.26	0.99	0.31
C(3b)–H(31b)$\cdots$O(2a)	0.069	-0.30	-0.21	1.08	0.58	2.532	3.110	14.10	-12.54	0.89	1.56
N(1b)–H(12b)$\cdots$O(1b)	0.058	-0.55	-0.16	1.18	0.47	2.359	3.167	11.26	-9.73	0.86	1.53

Appendix E

Topological properties of trialanine

The procedure to obtain accurate charge density of trialanine and its analysis are described in detail in Chapter 3.[42] Crystallographic data and a summary of the MEM calculation of trialanine are given in Table A.2. Table E.1 displays the geometrical, topological and energetic properties of hydrogen bonds of trialanine, which are discussed in detail in Chapter 5. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance d(H$\cdots$O).

Table E.1: Geometrical, topological and energetic properties of hydrogen bonds of trialanine (Figure 3.1). d(H$\cdots$O) and d(D$\cdots$O) in Å, $\rho(BCP)$ in electrons/Å^3, eigenvalues and $\nabla^2\rho(BCP)$ in electrons/Å^5, G, V and H in kJ/mol·Bohr^3.

	$\rho(BCP)$	λ_1	λ_2	λ_3	$\nabla^2\rho(BCP)$	d(H$\cdots$O)	d(D$\cdots$O)	G	V	$\lvert V\rvert/G$	H
N11–H11a$\cdots$O9b	0.454	-3.60	-2.81	4.83	-1.57	1.671	2.685	55.24	-153.27	2.77	-98.03
N1–H1a$\cdots$O19b	0.432	-2.80	-2.41	5.98	0.77	1.674	2.699	91.20	-161.35	1.77	-70.15
N11–H11c$\cdots$O31	0.370	-2.38	-1.87	5.66	1.41	1.771	2.784	85.25	-132.06	1.55	-46.81
O31–H31a$\cdots$O9a	0.335	-2.46	-1.87	5.92	1.59	1.878	2.637	79.36	-115.47	1.45	-36.10
N1–H1c$\cdots$O19a	0.334	-2.30	-1.78	5.86	1.77	1.731	2.764	82.54	-116.83	1.42	-34.29
O21–H21a$\cdots$O19a	0.273	-1.79	-1.61	4.94	1.54	1.999	2.697	63.76	-85.69	1.34	-21.93
N4–H4a$\cdots$O16a	0.214	-1.23	-0.94	4.62	2.45	1.913	2.885	68.50	-70.18	1.02	-1.67
N1–H1b$\cdots$O21	0.207	-1.47	-0.74	4.09	1.88	2.024	2.844	56.78	-62.43	1.10	-5.66
N7–H7a$\cdots$O13a	0.196	-1.28	-0.91	3.62	1.43	1.981	2.947	46.66	-54.29	1.16	-7.62
N11–H11b$\cdots$O9a	0.193	-1.40	-0.96	3.96	1.60	1.956	2.876	49.21	-54.72	1.11	-5.51
N14–H14a$\cdots$O6a	0.193	-0.99	-0.74	3.92	2.19	1.924	2.878	59.91	-60.22	1.01	-0.32
N17–H17a$\cdots$O3a	0.154	-0.62	-0.52	3.24	2.11	2.079	3.027	51.99	-46.64	0.90	5.35

Appendix F

Accurate charge density of Ala-Tyr-Ala with water by the Maximum Entropy Method (MEM)

F.1 MEM calculations

Single-crystal X-ray diffraction data of Ala-Tyr-Ala with water[83] (Figures F.1,F.2) were kindly provided by Luger and co-workers, who have already reported multipole refinements against these data (Table A.2).[83] We have used these data to perform refinements of the independent spherical atom model (ISAM) with the computer program JANA2000.[60]

For the ISAM refinement of Ala-Tyr-Ala with water as solvent the coordinates and ADP obtained by the multipole refinement were used starting model. Hydrogen atoms were fixed at distances known from neutron diffraction[76,94,95] (X_3–C–H: 1.099 Å, C_2–C–H_3: 1.059 Å, -NH_3^+: 1.033 Å, X_2–N–H: 1.009 Å, C–O–H: 0.967 Å, O–H (water): 0.96 Å). For all three water molecules of the asymmetric unit the angle H–O–H was restrained to 104.5°. The riding model with U_{iso}(H)= 1.2U_{eq}(N,C) and U_{iso}(H)= 1.5U_{eq}(O) was employed to calculate the ADP of hydrogen atoms. The coordinates of H(15) of the hydroxyl group and hydrogen atoms of all water molecules (H(61), H(62), H(71), H(72), H(81) and H(82)) were refined freely being subject to the distances known from neutron diffraction, since a reasonable geometrical restraint was not available. Coordinates of all other hydrogen atoms were obtained by attachment to their neighbor atoms with tetrahedral or trigonal angle restraints according to chemical meaning. The occupancy a_i of one water molecule (O(8), H(81) and H(82)) has been determined as different from one. The value of a_i=0.6344 from multipole refinement was used as starting value for the ISAM refinement, which yielded

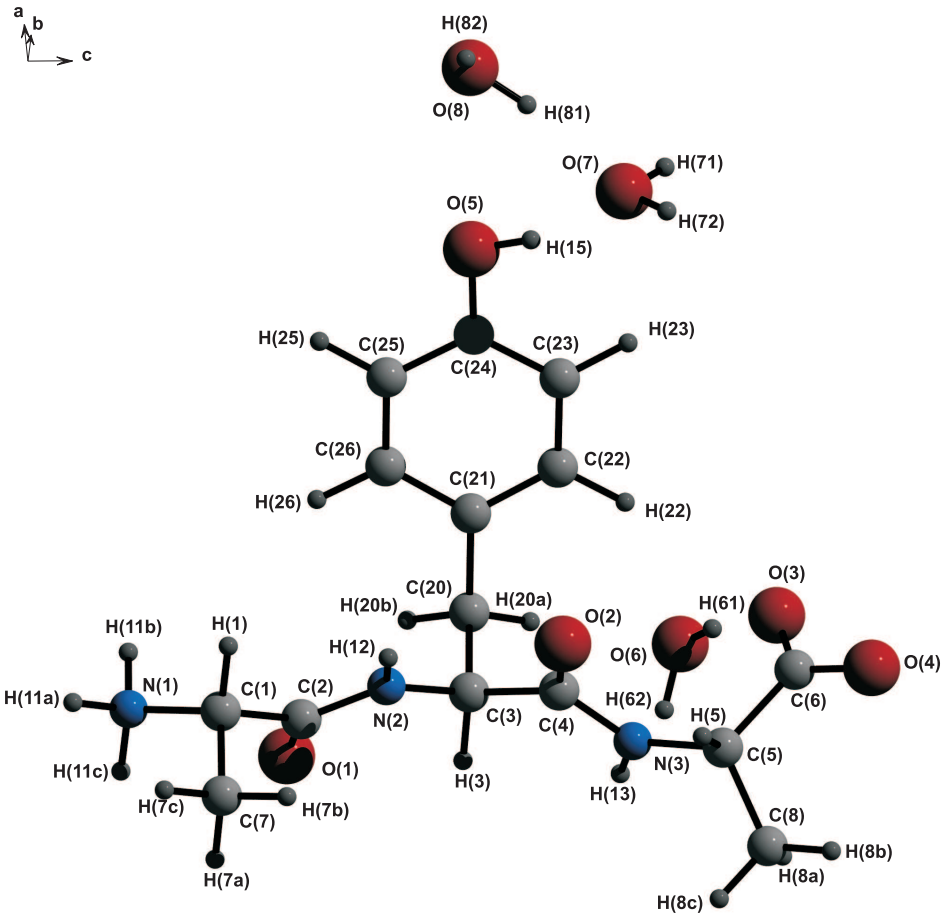

Figure F.1: Perspective view of Ala-Tyr-Ala with water obtained with atomic coordinates from the ISAM, displaying the atoms of the asymmetric unit.

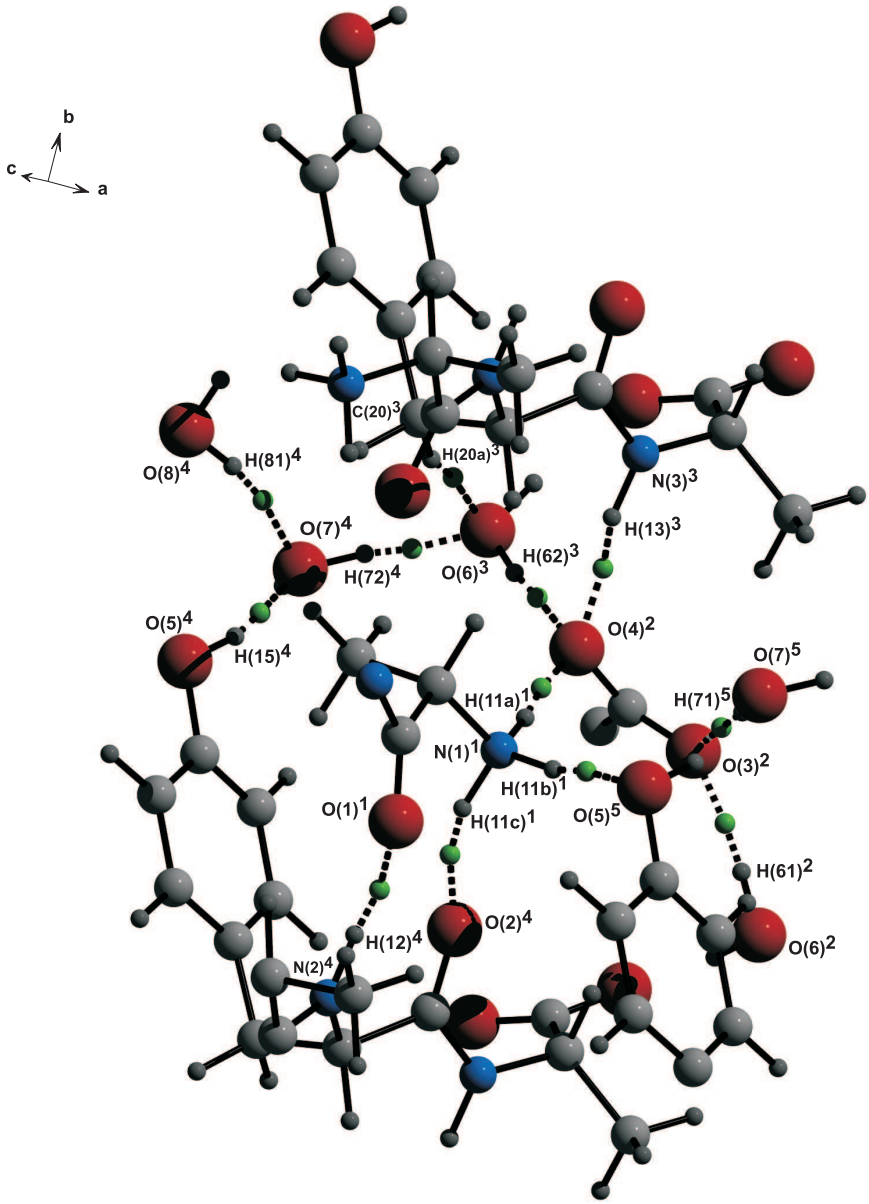

Figure F.2: Perspective view of Ala-Tyr-Ala with water obtained with atomic coordinates from the ISAM, displaying hydrogen bonds. Hydrogen bonds are indicated by dotted lines, green dots indicate BCPs of hydrogen bonds. Symmetry operators are denoted by superscripts: $1:(x,y,z)$; $2:(x,y,-1+z)$; $3:(-x,0.5+y,1-z)$; $4:(-x;-0.5+y,1-z)$; $5:(1-x;-0.5+y,1-z)$.

an occupancy of that molecule of a_i=0.6225. A weighting factor

$$w_{F_{obs}} = \frac{1}{[\sigma^2_{F_{obs}}|F_{obs}| + (kF_{obs})^2]} \tag{F.1}$$

with k=0.02 was employed for weighting of the structure factors in the ISAM refinement.

According to a procedure by Bagautdinov,[75] the ISAM refinement was employed to obtain phased and scaled structure factors which are corrected for anomalous scattering and used for the MEM calculations. The coordinates and ADP obtained by the ISAM refinement were used to compute the procrystal electron density (prior density $\rho^{prior}(\mathbf{x})$) with the computer program PRIOR.[105] This prior density $\rho^{prior}(\mathbf{x})$ is calculated on a grid over the unit cell and used as reference density in the MEM calculations. The grid for the prior and MEM density has to be chosen in that way that the pixel size does not exceed 0.1×0.1×0.1 Å^3 (Table A.2).

For the iterative procedure of the historical MEM the summation of Eq. 4.2 has been extended towards all reflections up to $\sin(\theta)/\lambda$ = 2.5 Å^{-1} according to the method of prior-derived F-constraint (PDC).[51] Static weights of H4 have been chosen according to de Vries *et al.*[52] employing Eq. 4.3 (Chapter 4). MEM calculations have been performed with the computer program BayMEM[45] employing the Cambridge maximum entropy algorithm.[46,106]

For Ala-Tyr-Ala with water as solvent $\chi^2_{aim} = 1.25$ has been determined by the classical MEM and used for the historical MEM.[53] The difference map $\rho^{MEM} - \rho^{prior}$ shows noisy contour lines indicating that noise has already been fitted (Figure F.3(e)). To determine the optimal χ^2_{aim}, calculations of the historical MEM with $\chi^2_{aim} = 1.00$ and $\chi^2_{aim} = 1.50$ have been performed. By inspection of the resulting difference maps, $\rho^{MEM} - \rho^{prior}$ of $\chi^2_{aim} = 1.00$ shows slightly more noisy contour lines (Figure F.3(b)), whereas $\rho^{MEM} - \rho^{prior}$ of $\chi^2_{aim} = 1.50$ exhibits smoother contour lines (Figure F.3(h)). Thus, it has been concluded that the optimal χ^2_{aim} for Ala-Tyr-Ala with water as solvent lies between 1.25 and 1.50. Therefore, the $\chi^2_{aim} = 1.25$ has successively been increased to values of 1.2750, 1.2875 and to 1.3000 to pinpoint the optimal χ^2_{aim}. With increasing χ^2_{aim}, $\rho^{MEM} - \rho^{prior}$ maps become smoother and the difference Fourier maps contain more residual density which has not been fitted (Figure F.4). The compromise of a smooth $\rho^{MEM} - \rho^{prior}$ and a most featureless difference Fourier map has shown to be satisfied by $\chi^2_{aim} = 1.2750$. A further increase of its value would results only in a larger residual density in the difference Fourier map but no further smoothing of $\rho^{MEM} - \rho^{prior}$ (Figure F.4). Thus, $\chi^2_{aim} = 1.2750$ has been chosen for the final historical MEM calculation. The determination of χ^2_{aim} for this compound was performed as in the case of trialanine[42] (Chapter 3).

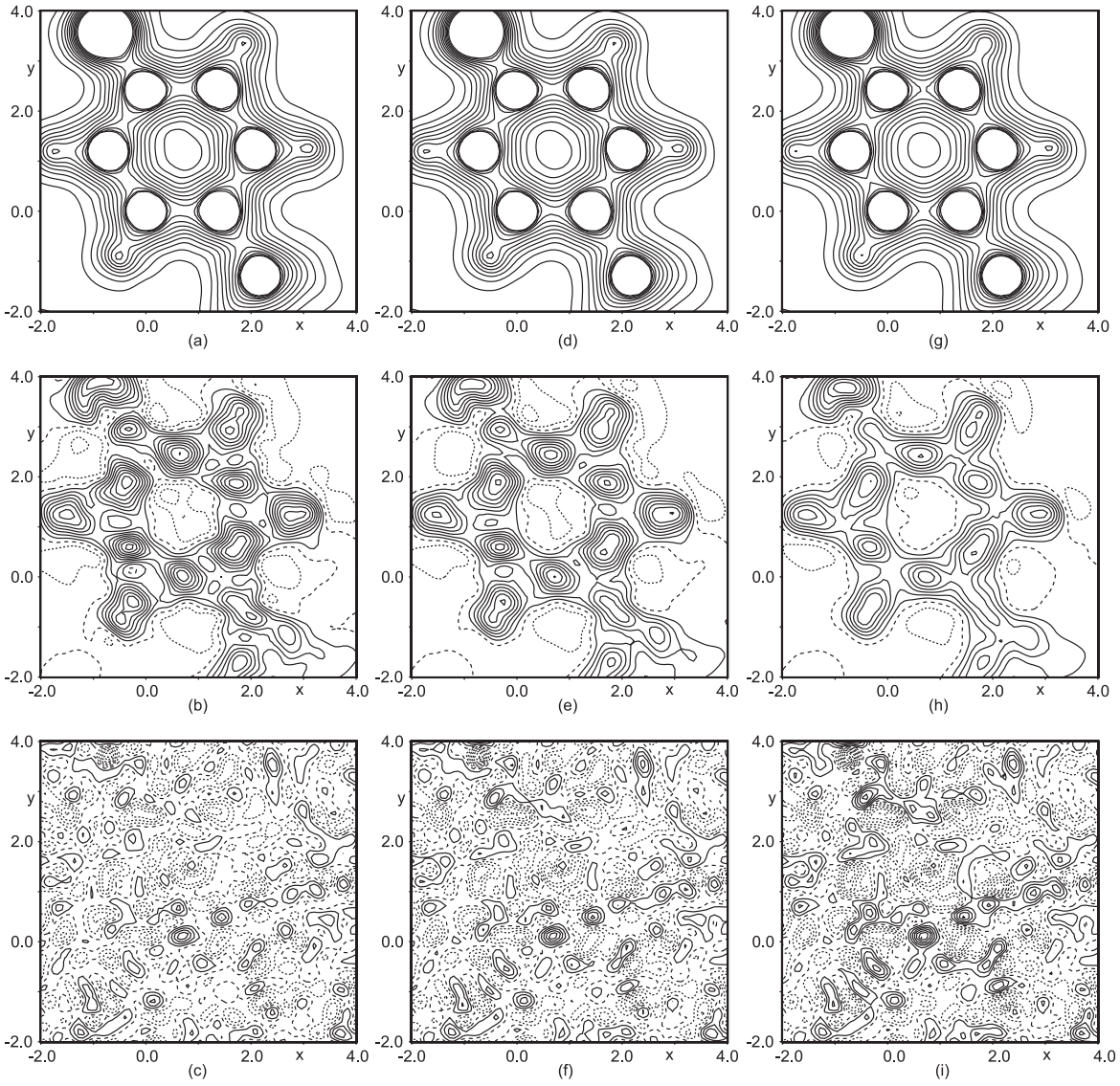

Figure F.3: Sections of 6×6 Å^2 through the phenyl ring plane of Ala-Tyr-Ala with water. (a) ρ^{MEM}, $\chi^2_{aim} = 1.00$. (b) Difference map, $\chi^2_{aim} = 1.00$, with $\Delta\rho$(min/max) = -0.15/0.47 electrons/Å^3. (c) Difference Fourier map, $\chi^2_{aim} = 1.00$, with $\Delta\rho$(min/max) = -0.31/0.24 electrons/Å^3. (d) ρ^{MEM}, $\chi^2_{aim} = 1.25$. (e) Difference map, $\chi^2_{aim} = 1.25$, with $\Delta\rho$(min/max) = -0.13/0.44 electrons/Å^3. (f) Difference Fourier map, $\chi^2_{aim} = 1.25$, with $\Delta\rho$(min/max) = -0.36/0.28 electrons/Å^3. (g) ρ^{MEM} with $\chi^2_{aim} = 1.50$. (h) Difference map, $\chi^2_{aim} = 1.50$, with $\Delta\rho$(min/max) = -0.12/0.44 electrons/Å^3. (i) Difference Fourier map, $\chi^2_{aim} = 1.50$, with $\Delta\rho$(min/max) = -0.42/0.35 electrons/Å^3. (a), (d) and (g): Contours of equal density are given from 0.2 to 2.5 electrons/Å^3 in steps of 0.2 electrons/Å^3. (b), (c), (e), (f), (h) and (i): Contour lines are at intervals of 0.05 electrons/Å^3. Solid lines are contours of positive value, dotted lines are negative contours, and dashed lines represent the contour of zero value.

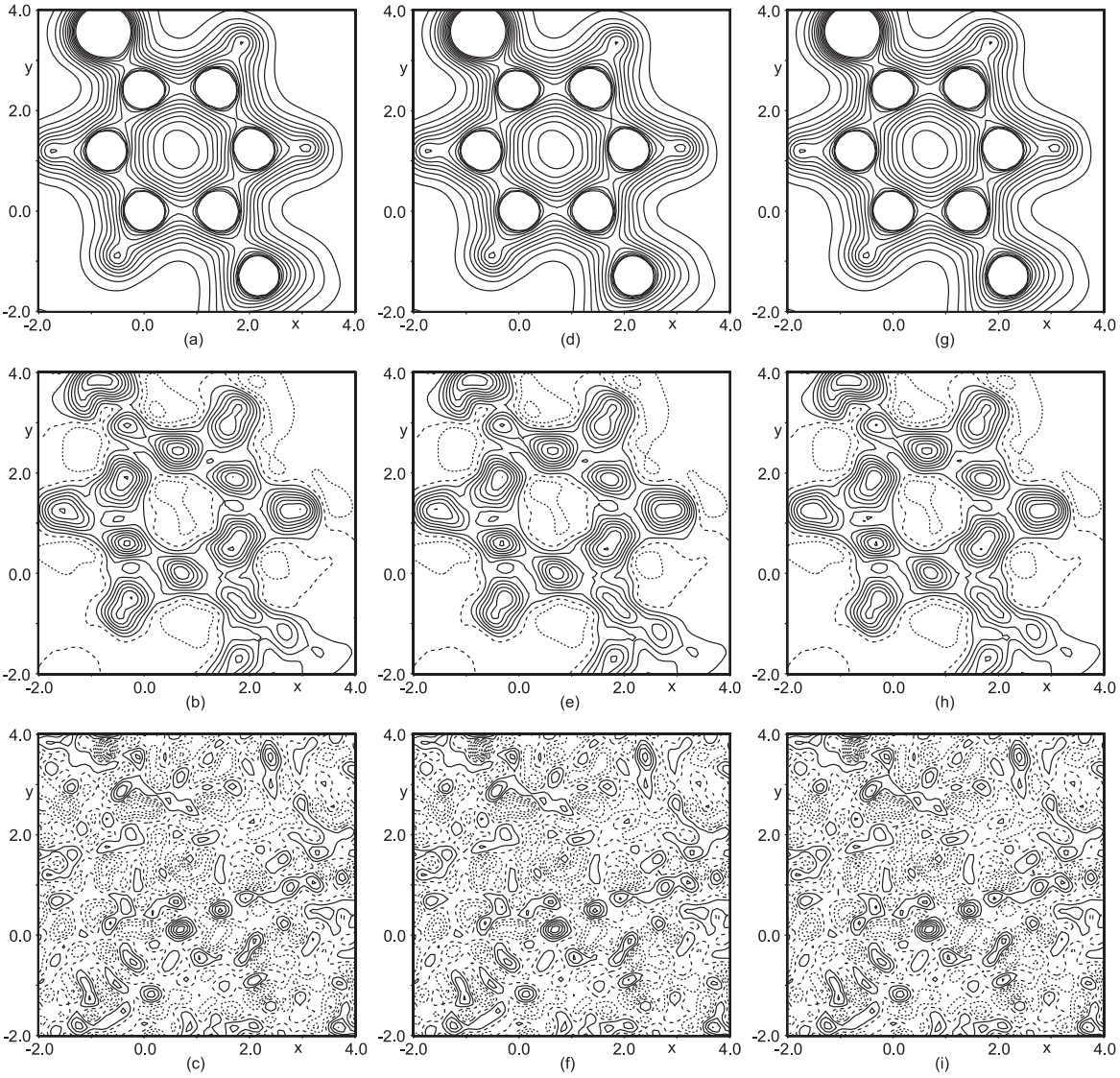

Figure F.4: Sections of 6×6 Å^2 through the phenyl ring plane of Ala-Tyr-Ala with water. (a) ρ^{MEM}, $\chi^2_{aim} = 1.2750$. (b) Difference map, $\chi^2_{aim} = 1.2750$, with $\Delta\rho$(min/max) = -0.12/0.44 electrons/Å^3. (c) Difference Fourier map, $\chi^2_{aim} = 1.2750$, with $\Delta\rho$(min/max) = -0.37/0.29 electrons/Å^3. (d) ρ^{MEM}, $\chi^2_{aim} = 1.2875$. (e) Difference map, $\chi^2_{aim} = 1.2875$, with $\Delta\rho$(min/max) = -0.12/0.44 electrons/Å^3. (f) Difference Fourier map, $\chi^2_{aim} = 1.2875$, with $\Delta\rho$(min/max) = -0.37/0.29 electrons/Å^3. (g) ρ^{MEM} with $\chi^2_{aim} = 1.3000$. (h) Difference map, $\chi^2_{aim} = 1.3000$, with $\Delta\rho$(min/max) = -0.12/0.44 electrons/Å^3. (i) Difference Fourier map, $\chi^2_{aim} = 1.3000$, with $\Delta\rho$(min/max) = -0.38/0.29 electrons/Å^3. Contour lines as defined in Figure F.3.

F.2 Analysis of the MEM density

Atomic positions obtained from ρ^{MEM} are given in Table F.1. Hydrogen atoms H(11c), H(12), H(13), H(15) and of all water molecules do not exhibit atomic maxima. Their atomic volumes and charges are incorporated into the basins of their neighboring atoms (Table F.2). Volumes of the atomic basins add up to 946.2 $Å^3$ which is in perfect agreement with volume of the unit cell (Table A.2). The number of electrons of the unit cell of 396.7 indicates that all electrons are incorporated into atomic basins. The net charge Q of the unit cell is determined -0.18, which is close to zero, fulfilling the required electrical neutrality.

The density at BCPs of covalent bonds from the MEM density is always smaller than the values from multipole refinement (Table F.3) as also reported for trialanine[42] (Chapter 3) and α-glycine[53] (Chapter 4). The theoretical values for the density at BCPs of heterogenous covalent bonds are always smaller than from multipole refinement, indicating that multipole refinement might overestimate $\rho(BCP)$. This trend can also be observed in case of trialanine[42] (Chapter 3). The theoretical values for C–C bonds have the same magnitude as from multipole refinement. The density at BCPs from ρ^{MEM} are always smaller than the theoretical values. The values of the Laplacian of C–O bonds from the MEM deviate highly from the values of multipole refinement (Table F.3), which has also been found with trialanine[42] (Chapter 3) and α-glycine[53] (Chapter 4). The geometrical, topological and energetic properties for hydrogen bonds are given in Table F.4 and are discussed in detail in Chapter 5. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance d(H$\cdots$O).

Table F.1: Atomic positions for Ala-Tyr-Ala with water derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms. †Occupancy $a_i = 0.6225$.

Atom	x	y	z
C(1)	0.979599	0.597638	0.387691
C(2)	0.031913	0.510367	0.486654
C(3)	0.915177	0.011928	0.321463
C(4)	0.984033	0.100364	0.229021
C(5)	0.100938	0.089903	0.050476
C(6)	0.030474	0.631093	0.035447
C(7)	0.791870	0.606218	0.377400
C(8)	0.769601	0.490240	-0.001788
C(20)	0.725372	0.989577	0.308980
C(21)	0.626244	0.127517	0.309566
C(22)	0.591444	0.202289	0.214459
C(23)	0.508574	0.334290	0.215173
C(24)	0.461672	0.393682	0.312836
C(25)	0.490110	0.318116	0.408037
C(26)	0.572373	0.185964	0.406447
N(1)	0.046517	0.523069	0.291649
N(2)	0.959306	0.084449	0.421251
N(3)	0.030278	0.021875	0.144149
O(1)	0.060601	0.379612	0.479397
O(2)	0.993761	0.233060	0.233592
O(3)	0.177580	0.593387	0.025654
O(4)	0.975162	0.702652	0.114750
O(5)	0.388533	0.527585	0.319249
O(6)	0.676513	0.841564	0.059766
O(7)	0.415688	0.708807	0.159467
O(8)†	0.285172	0.876195	0.335165
H(1)	0.027689	0.702728	0.391735
H(3)	0.972057	0.910899	0.324134
H(5)	0.154916	0.184886	0.080298
H(7a)	0.745692	0.507639	0.375405
H(7b)	0.749407	0.656616	0.440284
H(7c)	0.758322	0.661079	0.310683

Continued on Next Page...

Table F.1 – Continued

Atom	x	y	z
H(8a)	0.821362	0.399690	0.023058
H(8b)	0.718180	0.537754	0.060735
H(8c)	0.317937	0.968019	0.056268
H(11a)	0.017386	0.576931	0.228749
H(11b)	0.166306	0.517584	0.298369
H(20a)	0.699945	0.933301	0.238782
H(20b)	0.687291	0.922000	0.369961
H(22)	0.626045	0.159945	0.143665
H(23)	0.481291	0.388558	0.144937
H(25)	0.452251	0.361761	0.478886
H(26)	0.594886	0.131346	0.477163

Table F.2: Volumes of the atomic basins, the number of electrons in the atomic basins and the net atomic charge (Q) in the asymmetric unit for $\rho^{MEM}(\mathbf{x})$. The basins of N(1,2,3) and O(5,6,7,8) encompass their bonded hydrogen atoms because the latter atoms do not constitute local maxima in the density. †Occupancy $a_i = 0.6225$.

Atom	Volume	Electrons	Q
C(1)	8.28	5.76	+0.24
C(2)	6.85	4.84	+1.16
C(3)	6.58	5.70	+0.30
C(4)	7.21	4.84	+1.16
C(5)	8.78	5.86	+0.14
C(6)	5.77	4.59	+1.41
C(7)	11.09	6.31	-0.31
C(8)	12.18	6.29	-0.29
C(20)	9.01	6.24	-0.24
C(21)	10.83	6.15	-0.15
C(22)	12.99	6.08	-0.08
C(23)	13.59	6.21	-0.21
C(24)	11.10	5.70	+0.30
C(25)	11.98	6.02	-0.02
C(26)	13.86	6.27	-0.27
$N(1)H_1$	19.82	8.84	-0.84
$N(2)H_1$	15.37	8.31	-0.31
$N(3)H_1$	17.31	8.61	-0.61
O(1)	16.28	9.01	-1.01
O(2)	17.17	8.92	-0.92
O(3)	15.43	8.83	-0.83
O(4)	14.05	8.85	-0.85
$O(5)H_1$	19.37	9.63	-0.63
$O(6)H_2$	23.67	10.12	-0.12
$O(7)H_2$	24.28	9.80	+0.20
$O(8)H_2$†	27.72	6.24	-0.01
H(1)	5.13	0.82	+0.18
H(3)	5.42	0.89	+0.11
H(5)	7.26	0.86	+0.14
H(7a)	5.59	0.80	+0.20

Continued on Next Page...

Table F.2 – Continued

Atom	Volume	Electrons	Q
H(7b)	8.97	0.98	+0.02
H(7c)	8.20	0.90	+0.10
H(8a)	6.83	0.89	+0.11
H(8b)	9.12	0.98	+0.02
H(8c)	8.44	0.87	+0.13
H(11a)	3.11	0.53	+0.47
H(11b)	1.04	0.37	+0.63
H(20a)	6.97	0.99	+0.01
H(20b)	7.83	1.04	-0.04
H(22)	7.85	0.90	+0.10
H(23)	6.72	0.78	+0.22
H(25)	6.74	0.83	+0.17
H(26)	7.30	0.87	+0.13
Total	473.08	198.34	−0.09

Table F.3: Electron densities and Laplacians at the BCPs of covalent bonds and RCPs (ring critical points) of Ala-Tyr-Ala with water. Values are given for $\rho(BCP)$ (electrons/Å^3; first line) and $\nabla^2\rho(BCP)$ (electrons/Å^5; second line) derived from $\rho^{prior}(\mathbf{x})$, $\rho^{MEM}(\mathbf{x})$, the static electron density of the multipole model[83] and DFT calculations (B3LYP/6-311++G(3df,3pd)).[83]

Bond	Prior	MEM	Multipole	Calc.
C(2)–O(1)	2.08	2.46	2.95	2.72
	5.05	14.50	-25.6	-17.3
C(4)–O(2)	2.08	2.40	2.99	2.71
	6.12	2.67	-28.8	-16.8
C(6)–O(3)	2.07	2.52	2.97	2.61
	14.81	6.10	-34.5	-17.1
C(6)–O(4)	1.95	2.16	2.64	2.51
	5.00	19.00	-24.1	-21.8
C(24)–O(5)	1.67	1.90	-	-
	3.93	-5.58	-	-
C(1)–N(1)	1.36	1.53	1.70	1.59
	2.51	-5.67	-8.9	-13.2
C(2)–N(2)	1.74	2.08	2.44	2.32
	-1.14	-18.14	-20.6	-26.3
C(3)–N(2)	1.44	1.60	1.75	1.72
	2.11	-6.96	-9.4	-15.8
C(4)–N(3)	1.74	2.05	2.39	2.31
	-1.17	4.51	-19.2	-26.1
C(5)–N(3)	1.44	1.66	1.81	1.71
	1.91	-12.54	-9.6	-15.3
C(1)–C(2)	1.17	1.36	1.67	1.72
	0.46	0.70	-8.9	-15.2
C(1)–C(7)	1.19	1.43	1.71	1.64
	0.20	-0.53	-9.9	-13.6
C(3)–C(4)	1.17	1.48	1.64	1.71
	0.25	-11.51	-7.8	-15.0
C(3)–C(20)	1.15	1.31	1.61	1.55
	0.37	3.84	-7.3	-11.9
C(5)–C(6)	1.15	1.30	1.59	1.67

Continued on Next Page...

Table F.3 – Continued

Bond	Prior	MEM	Multipole	Calc.
	0.54	6.46	-6.7	-14.2
C(5)–C(8)	1.17	1.38	1.60	1.62
	0.32	-4.36	-8.3	-13.3
C(20)–C(21)	1.23	1.50	-	-
	-0.47	-8.81	-	-
C(21)–C(22)	1.43	1.83	-	-
	-2.43	-13.03	-	-
C(22)–C(23)	1.46	1.82	-	-
	-2.97	-16.72	-	-
C(23)–C(24)	1.45	1.89	-	-
	-3.04	-17.65	-	-
C(24)–C(25)	1.44	1.85	-	-
	-2.61	-16.18	-	-
C(25)–C(26)	1.46	1.89	-	-
	-2.95	-25.66	-	-
C(21)–C(26)	1.43	1.84	-	-
	-2.58	-20.43	-	-
C(1)–H(1)	1.19	1.55	-	-
	-5.50	-16.36	-	-
C(3)–H(3)	1.19	1.59	-	-
	-5.64	-20.68	-	-
C(5)–H(5)	1.19	1.54	-	-
	-5.36	-13.89	-	-
C(7)–H(7a)	1.26	1.50	-	-
	-9.62	-17.26	-	-
C(7)–H(7b)	1.25	1.56	-	-
	-9.41	-17.95	-	-
C(7)–H(7c)	1.25	1.46	-	-
	-9.04	-12.17	-	-
C(8)–H(8a)	1.27	1.55	-	-
	-9.96	-12.34	-	-
C(8)–H(8b)	1.24	1.52	-	-
	-9.41	-14.56	-	-
C(8)–H(8c)	1.26	1.49	-	-

Continued on Next Page...

Table F.3 – Continued

Bond	Prior	MEM	Multipole	Calc.
	-9.28	-11.42	-	-
C(20)–H(20a)	1.20	1.48	-	-
	-6.26	-12.18	-	-
C(20)–H(20b)	1.20	1.52	-	-
	-6.52	-17.00	-	-
C(22)–H(22)	1.16	1.57	-	-
	-6.55	-14.85	-	-
C(23)–H(23)	1.16	1.47	-	-
	-6.51	-11.71	-	-
C(25)–H(25)	1.17	1.54	-	-
	-6.70	-16.28	-	-
C(26)–H(26)	1.16	1.48	-	-
	-6.07	-10.98	-	-
N(1)–H(11a)	1.44	1.69	-	-
	-15.69	-26.66	-	-
N(1)–H(11b)	1.47	1.79	-	-
	-17.00	-28.05	-	-
N(1)–H(11c)	1.46	1.77	-	-
	-16.25	-38.33	-	-
N(2)–H(12)	1.51	-	-	-
	-19.16	-	-	-
N(3)–H(13)	1.51	-	-	-
	-17.74	-	-	-
Phenyl ring	0.31	0.25	-	-
	1.92	4.09	-	-

Table F.4: Geometrical, topological and energetic properties of hydrogen bonds of Ala-Tyr-Ala with water (Figure F.2). d(H$\cdots$O) and d(D$\cdots$O) in Å, $\rho(BCP)$ in electrons/Å^3, eigenvalues and $\nabla^2\rho(BCP)$ in electrons/Å^5, G, V and H in kJ/mol·Bohr^3.

	$\rho(BCP)$	λ_1	λ_2	λ_3	$\nabla^2\rho(BCP)$	d(H$\cdots$O)	d(D$\cdots$O)	G	V	$\|V\|/G$	H
O(5)–H(15)$\cdots$O(7)	0.382	-2.93	-1.87	6.86	2.07	1.687	2.630	100.31	-144.32	1.44	-44.01
N(1)–H(11b)$\cdots$O(5)	0.364	-2.78	-2.51	5.87	0.57	1.763	2.792	68.53	-121.50	1.77	-52.96
O(6)–H(62)$\cdots$O(4)	0.315	-2.59	-1.94	4.88	0.36	1.864	2.820	52.23	-94.65	1.81	-42.42
N(1)–H(11a)$\cdots$O(4)	0.306	-2.42	-1.53	5.75	1.80	1.802	2.826	76.04	-103.15	1.36	-27.11
O(7)–H(71)$\cdots$O(3)	0.286	-2.47	-1.54	5.89	1.88	1.797	2.751	72.92	-94.54	1.30	-21.63
O(7)–H(72)$\cdots$O(6)	0.248	-1.84	-1.57	3.97	0.55	1.821	2.772	40.72	-66.32	1.63	-25.61
N(1)–H(11c)$\cdots$O(2)	0.219	-1.62	-1.12	4.45	1.71	1.867	2.824	55.91	-65.32	1.17	-9.41
O(6)–H(61)$\cdots$O(3)	0.213	-1.49	-0.69	3.79	1.61	1.891	2.844	53.04	-62.16	1.17	-9.13
N(3)–H(13)$\cdots$O(4)	0.153	-1.53	-0.74	2.94	0.67	2.030	3.023	25.80	-33.35	1.29	-7.55
N(2)–H(12)$\cdots$O(1)	0.151	-1.04	-0.60	2.91	1.26	1.968	2.952	36.35	-38.31	1.05	-1.96
C(20)–H(20a)$\cdots$O(6)	0.100	-0.48	-0.30	1.64	0.86	2.354	3.428	22.34	-21.26	0.95	1.08
O(8)–H(81)$\cdots$O(7)	0.095	-0.61	-0.22	2.61	1.78	1.983	2.914	38.46	-28.58	0.74	9.89

Appendix G

Accurate charge density of Ala-Tyr-Ala with ethanol by the Maximum Entropy Method (MEM)

G.1 MEM calculations

Single-crystal X-ray diffraction data of Ala-Tyr-Ala with ethanol[83] (Figures G.1,G.2) were kindly provided by Luger and co-workers, who have already reported multipole refinements against these data (Table A.2).[83] We have used these data to perform refinements of the independent spherical atom model (ISAM) with the computer program JANA2000.[60]

For the ISAM refinement of Ala-Tyr-Ala with ethanol as solvent the coordinates and ADP obtained by the multipole refinement were used as starting model. Hydrogen atoms were fixed at distances known from neutron diffraction[76] (X_3–C–H: 1.099 Å, X_2–C–H_2: 1.092 Å, C_2–C–H_3: 1.059 Å, -NH_3^+: 1.033 Å, X_2–N–H: 1.009 Å, C–O–H: 0.967 Å). The riding model with U_{iso}(H)= 1.2U_{eq}(N,C) and U_{iso}(H)= 1.5U_{eq}(O) was employed to calculate the ADP of hydrogen atoms. The coordinates of H(15) and H(16) of the hydroxyl groups were refined freely being subject to the distances known from neutron diffraction, since a reasonable geometrical restraint was not available. Coordinates of all other hydrogen atoms were obtained by attachment to their neighbor atoms with tetrahedral or trigonal angle restraints according to chemical meaning.

According to a procedure by Bagautdinov,[75] the ISAM refinement was employed to obtain phased and scaled structure factors which are corrected for anomalous scattering and used for the MEM calculations. The coordinates and ADP obtained by the ISAM refinement were used to compute the procrystal electron density (prior density $\rho^{prior}(\mathbf{x})$)

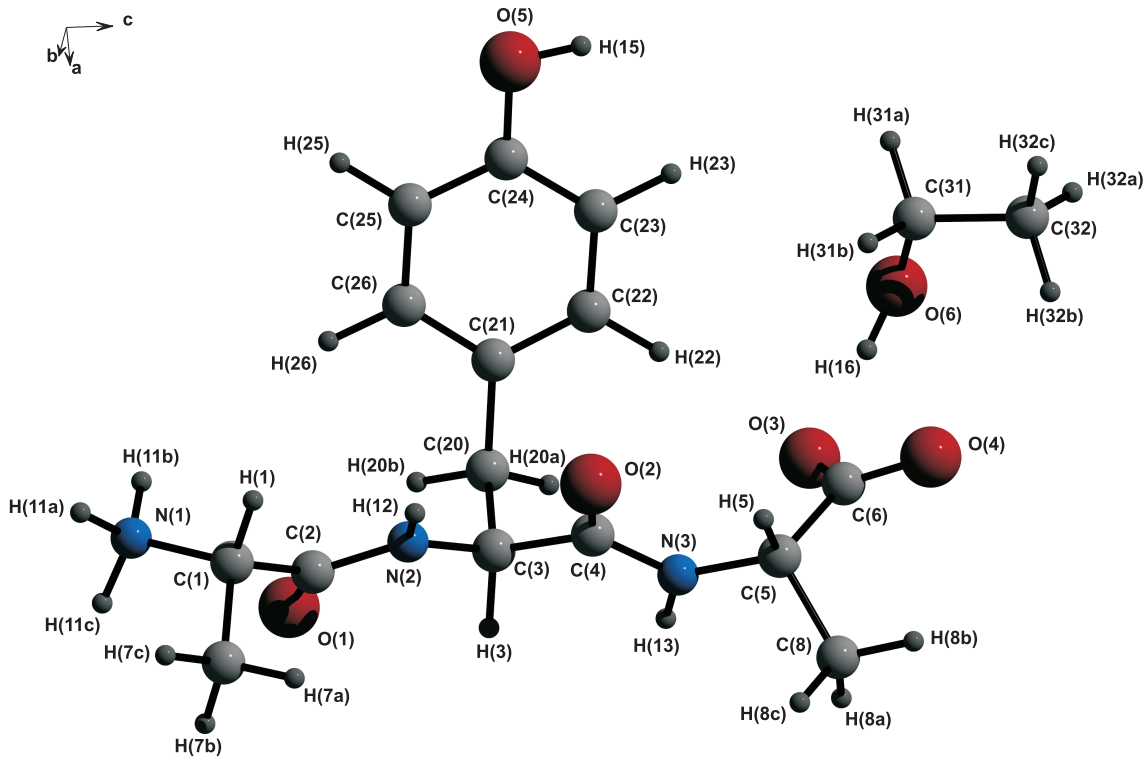

Figure G.1: Perspective view of the Ala-Tyr-Ala with ethanol obtained with atomic coordinates from the ISAM, displaying the atoms of the asymmetric unit.

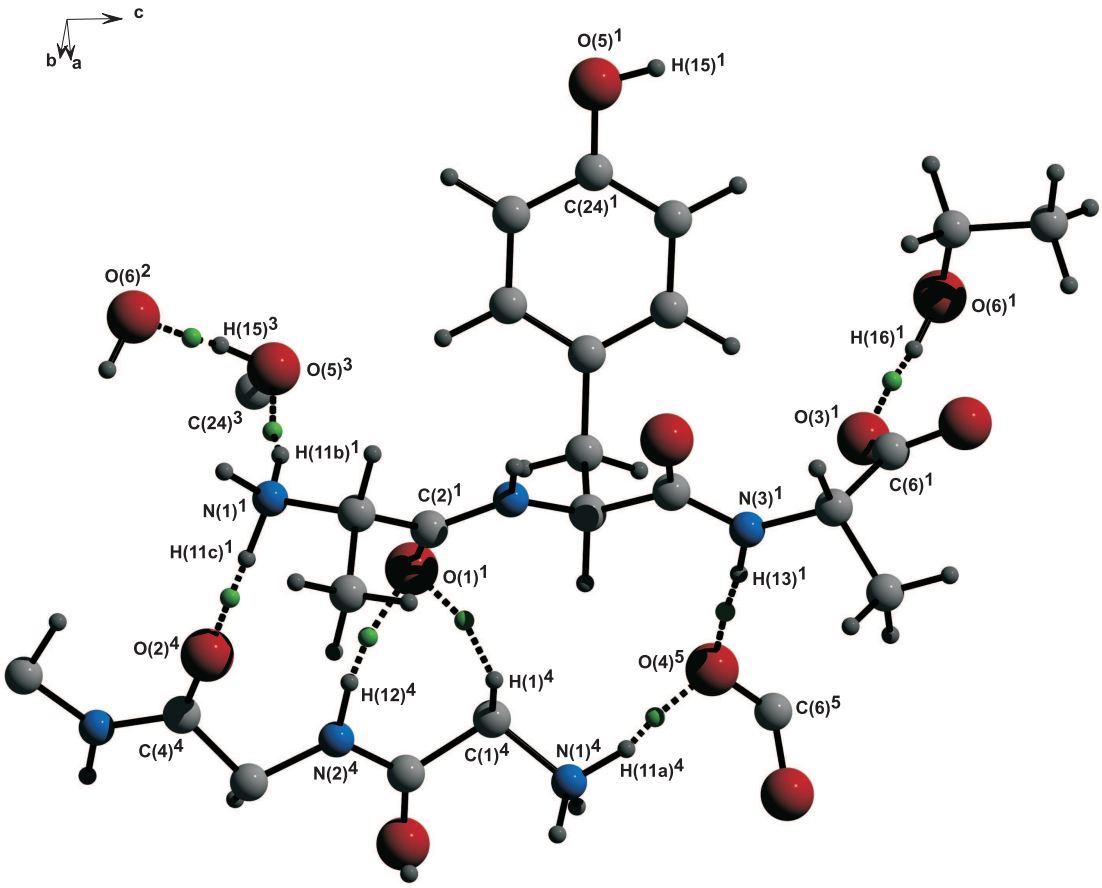

Figure G.2: Perspective view of Ala-Tyr-Ala with ethanol obtained with atomic coordinates from the ISAM, displaying hydrogen bonds. Hydrogen bonds are indicated by dotted lines, green dots indicate BCPs of hydrogen bonds. Symmetry operators are denoted by superscripts: $1:(x,y,z)$; $2:(x,y,-1+z)$; $3:(-x,0.5+y,1-z)$; $4:(1-x;0.5+y,1-z)$; $5:(1-x;0.5+y,2-z)$.

with the computer program PRIOR.[105] This prior density $\rho^{prior}(\mathbf{x})$ is calculated on a grid over the unit cell and used as reference density in the MEM calculations. The grid for the prior and MEM density has to be chosen in that way that the pixel size does not exceed 0.1×0.1×0.1 Å^3 (Table A.2).

For the iterative procedure of the historical MEM the summation of Eq. 4.2 has been extended towards all reflections up to $\sin(\theta)/\lambda = 2.5$ Å^{-1} according to the method of prior-derived F-constraint (PDC).[51] Static weights of H4 have been chosen according to de Vries *et al.*[52] employing Eq. 4.3 (Chapter 4). MEM calculations have been performed with the computer program BayMEM[45] employing the Cambridge maximum entropy algorithm.[46,106]

The classical MEM[53] for Ala-Tyr-Ala with ethanol as solvent yielded $\chi^2_{aim} = 1.27$ which has been used for the historical MEM.[53] The noisy contour lines in the difference map (Figure G.3(b)) and fifteen spurious maxima which cannot be identified with atoms, indicate that noise has been incorporated into the refinement. Calculations of the historical MEM with $\chi^2_{aim} = 1.00$ and $\chi^2_{aim} = 1.50$ have been performed to search for the optimal χ^2_{aim}. As in the case of Ala-Tyr-Ala with water as solvent, the difference map shows smoother contour lines with $\chi^2_{aim} = 1.50$ and only one spurious maxima was still present, but the difference Fourier map contains a much larger amount of residual density not fitted (Figure G.3). Whereas, in $\rho^{MEM} - \rho^{prior}$ with $\chi^2_{aim} = 1.00$ the distortion of the contour lines has dramatically increased (Figure G.3). Again, it could be concluded that the optimal χ^2_{aim} lies between 1.27 and 1.50. Thus, a successive increase of χ^2_{aim} to 1.3012, 1.3081 and to 1.3208 has been done to perform the historical MEM. For $\chi^2_{aim} = 1.3012$ ten spurious maxima were found present, whereas for $\chi^2_{aim} = 1.3081$ and $\chi^2_{aim} = 1.3208$ only six were present. The larger χ^2_{aim}, the more residual density was present in the difference Fourier map (Figure G.4). The compromise for a most featureless difference Fourier map and a smooth $\rho^{MEM} - \rho^{prior}$ was found satisfied by $\chi^2_{aim} = 1.3081$ (Figure G.3), which has therefore been chosen for the final historical MEM. The procedure to determine an optimal χ^2_{aim} for Ala-Tyr-Ala with ethanol as solvent followed the one applied on trialanine[42] (Chapter 3).

G.2 Analysis of the MEM density

Atomic positions obtained from ρ^{MEM} are given in Table G.1. Hydrogen atoms of the ammonium group and of the methyl group of ethanol, H(12), H(13), H(15) and H(16) do not exhibit atomic maxima. Their atomic volumes and charges are incorporated into the basins of their neighboring atoms (Table G.2). Volumes of the atomic basins add up to

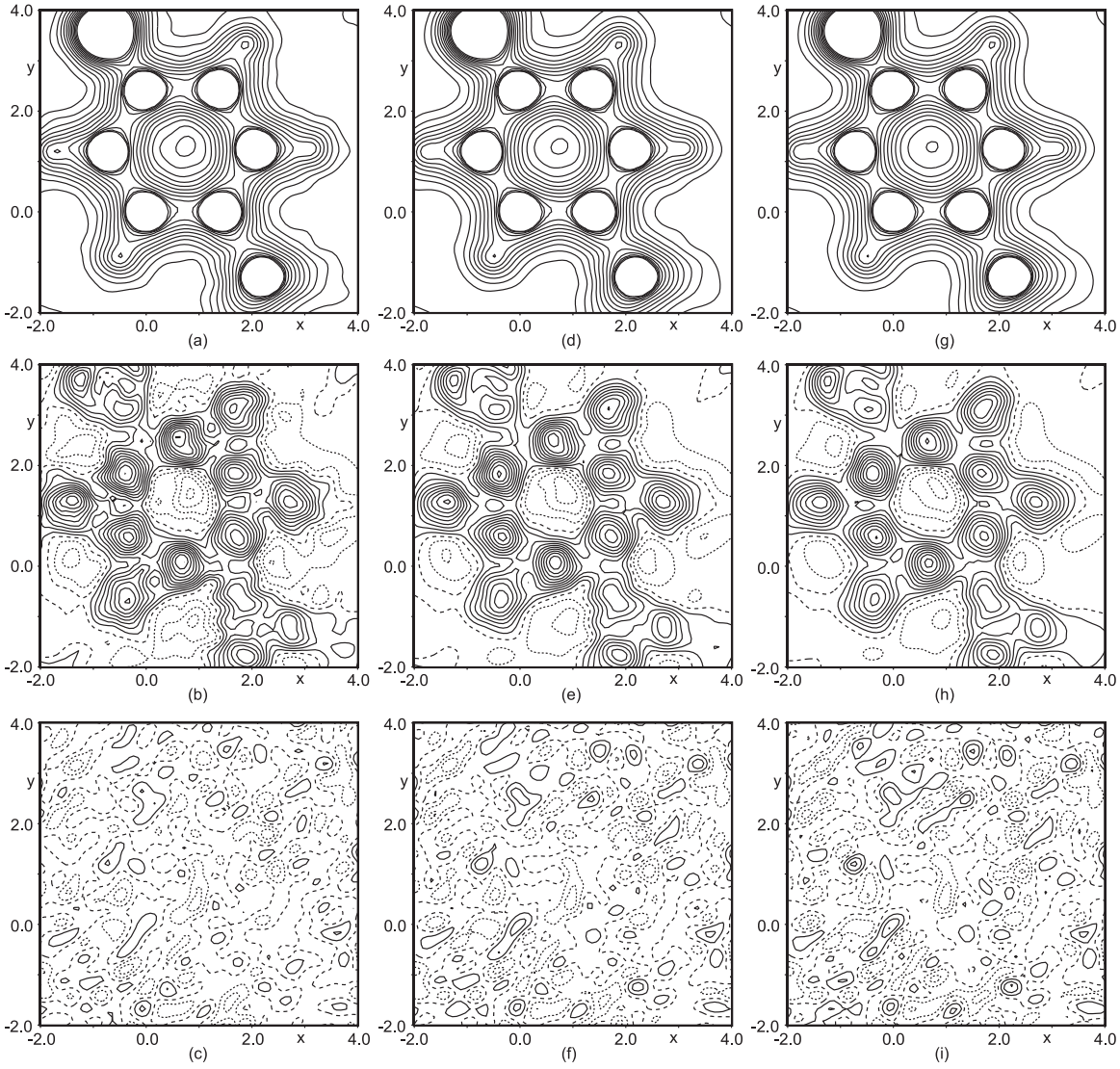

Figure G.3: Sections of 6×6 Å^2 through the phenyl ring plane of Ala-Tyr-Ala with ethanol. (a) ρ^{MEM}, $\chi^2_{aim} = 1.00$. (b) Difference map, $\chi^2_{aim} = 1.00$, with $\Delta\rho$(min/max) = -0.21/0.55 electrons/Å^3. (c) Difference Fourier map, $\chi^2_{aim} = 1.00$, with $\Delta\rho$(min/max) = -0.14/0.12 electrons/Å^3. (d) ρ^{MEM}, $\chi^2_{aim} = 1.27$. (e) Difference map, $\chi^2_{aim} = 1.27$, with $\Delta\rho$(min/max) = -0.18/0.54 electrons/Å^3. (f) Difference Fourier map, $\chi^2_{aim} = 1.27$, with $\Delta\rho$(min/max) = -0.14/0.14 electrons/Å^3. (g) ρ^{MEM} with $\chi^2_{aim} = 1.50$. (h) Difference map, $\chi^2_{aim} = 1.50$, with $\Delta\rho$(min/max) = -0.16/0.52 electrons/Å^3. (i) Difference Fourier map, $\chi^2_{aim} = 1.50$, with $\Delta\rho$(min/max) = -0.16/0.16 electrons/Å^3. Contour lines as defined in Figure F.3.

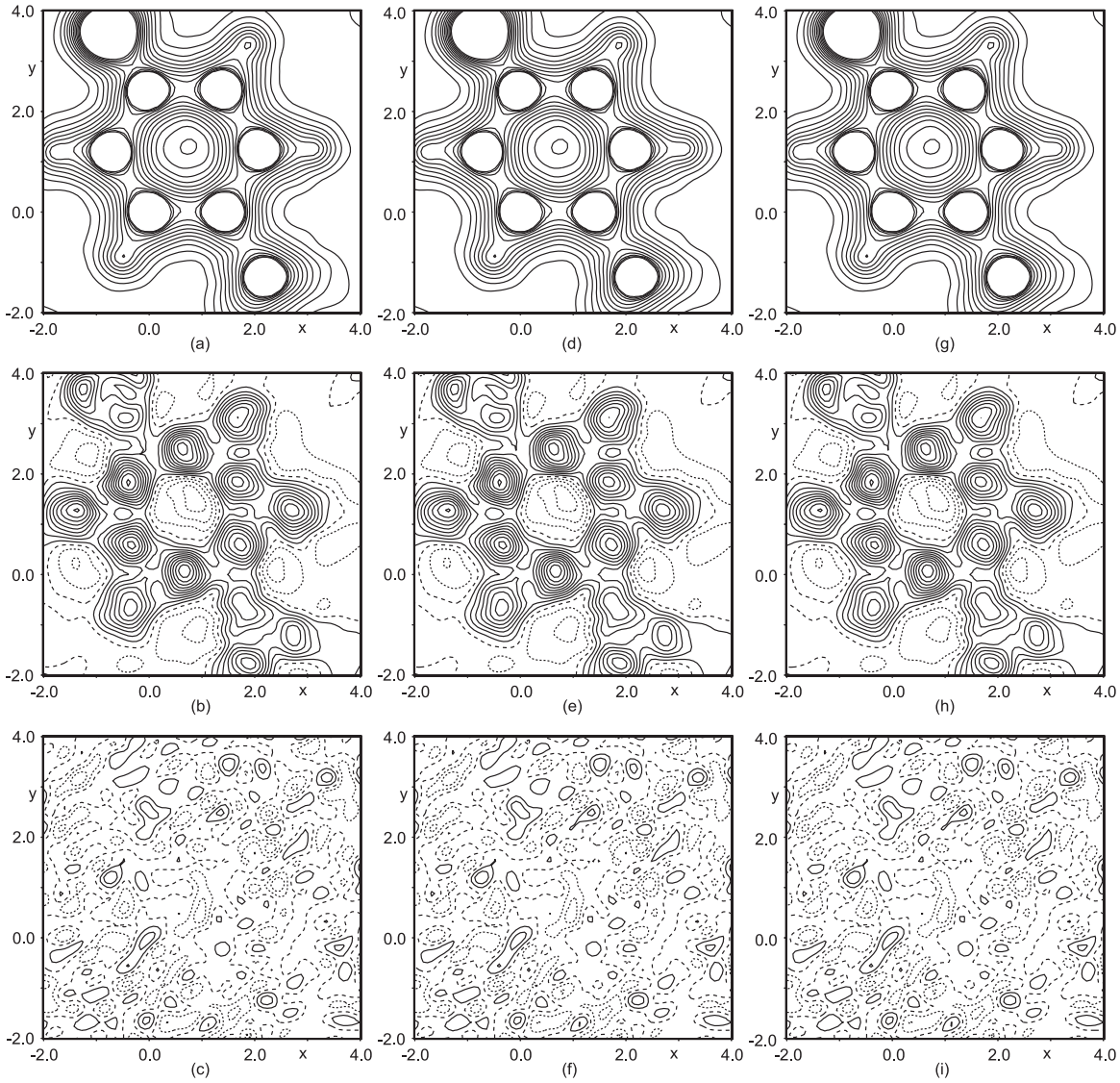

Figure G.4: Sections of 6×6 Å^2 through the phenyl ring plane of Ala-Tyr-Ala with ethanol. (a) ρ^{MEM}, χ^2_{aim} = 1.3012. (b) Difference map, χ^2_{aim} = 1.3012, with $\Delta\rho$(min/max) = -0.18/0.54 electrons/Å^3. (c) Difference Fourier map, χ^2_{aim} = 1.3012, with $\Delta\rho$(min/max) = -0.15/0.15 electrons/Å^3. (d) ρ^{MEM}, χ^2_{aim} = 1.3081. (e) Difference map, χ^2_{aim} = 1.3081, with $\Delta\rho$(min/max) = -0.18/0.54 electrons/Å^3. (f) Difference Fourier map, χ^2_{aim} = 1.3081, with $\Delta\rho$(min/max) = -0.15/0.15 electrons/Å^3. (g) ρ^{MEM} with χ^2_{aim} = 1.3208. (h) Difference map, χ^2_{aim} = 1.3208, with $\Delta\rho$(min/max) = -0.18/0.54 electrons/Å^3. (i) Difference Fourier map, χ^2_{aim} = 1.3208, with $\Delta\rho$(min/max) = -0.15/0.15 electrons/Å^3. Contour lines as defined in Figure F.3.

978.0 Å^3. Six spurious maxima which cannot be identified as atomic maxima are present. The volumes of these spurious maxima add up to 9.3 Å^3, which is 0.94 % of the volume of the unit cell (Table A.2). The number of electrons distributed in the atomic basins is 395.8. The number of electrons of the spurious maxima is 0.20, constituting 0.03 % of the total number of electrons (Table A.2). The net charge Q of the unit cell is determined +0.18, which is close to zero, fulfilling the required electrical neutrality.

The density at BCPs of covalent bonds from the MEM density is always smaller than the values from multipole refinement (Table G.3) as also reported for trialanine[42] (Chapter 3) and α-glycine[53] (Chapter 4). The theoretical values for the density at BCPs of heterogenous covalent bonds are always smaller than from multipole refinement, indicating that multipole refinement might overestimate $\rho(BCP)$. This trend can also be observed in case of trialanine[42] (Chapter 3). The theoretical values for C–C bonds have the same magnitude as from multipole refinement. The density at BCPs from ρ^{MEM} are in general smaller than the theoretical values. The values of the Laplacian of C–O bonds from the MEM deviate most prominently from the values of multipole refinement (Table G.3). It can be observed that the magnitudes of the Laplacian from ρ^{MEM} of Ala-Tyr-Ala with ethanol are larger than from Ala-Tyr-Ala with water (Appendix F), indicating that ρ^{MEM} from the latter is more smooth. The geometrical, topological and energetic properties for hydrogen bonds are given in Table G.4 and are discussed in detail in Chapter 5. Since some hydrogen atoms do not constitute atomic maxima, the corresponding coordinates of hydrogen atoms from the ISAM were employed to calculate the distance d(H$\cdots$O).

Table G.1: Atomic positions for Ala-Tyr-Ala with ethanol derived from the ρ^{MEM}. x, y, z denote the coordinates of the atoms.

Atom	x	y	z
C(1)	0.497356	0.199119	0.392160
C(2)	0.456416	0.284391	0.492888
C(3)	0.563306	0.775220	0.310427
C(4)	0.504677	0.668043	0.220400
C(5)	0.431327	0.639845	0.025627
C(6)	0.440670	0.140017	0.052023
C(7)	0.665856	0.217957	0.374218
C(8)	0.716538	0.195400	0.034895
C(20)	0.735173	0.808256	0.302105
C(21)	0.833923	0.671670	0.306347
C(22)	0.860137	0.595718	0.210154
C(23)	0.944686	0.464980	0.213029
C(24)	0.003961	0.407877	0.313398
C(25)	0.983535	0.484490	0.409572
C(26)	0.898642	0.615321	0.405784
C(31)	0.080796	0.980251	0.132946
C(32)	0.133317	0.917044	0.244022
N(1)	0.401359	0.257728	0.296152
N(2)	0.534002	0.706649	0.413934
N(3)	0.471094	0.729493	0.122650
O(1)	0.420363	0.416504	0.485795
O(2)	0.496407	0.534193	0.239520
O(3)	0.328261	0.224188	0.032818
O(4)	0.461563	0.055823	0.135077
O(5)	0.081813	0.275992	0.320114
O(6)	0.083883	0.139010	0.132901
H(1)	0.473632	0.087322	0.400274
H(3)	0.503489	0.874742	0.305899
H(5)	0.417565	0.533035	0.055042
H(7a)	0.726180	0.176229	0.435885
H(7b)	0.687368	0.320041	0.368032
H(7c)	0.691426	0.166151	0.308763

Continued on Next Page...

Table G.1 – Continued

Atom	x	y	z
H(8a)	0.707632	0.294972	0.053544
H(8b)	0.738810	0.140072	0.099922
H(8c)	0.203990	0.685662	0.009615
H(20a)	0.746765	0.863635	0.231334
H(20b)	0.771375	0.878743	0.363221
H(22)	0.817273	0.635790	0.138751
H(23)	0.964116	0.412471	0.142368
H(25)	0.029345	0.444232	0.481242
H(26)	0.883356	0.670183	0.475612
H(31a)	0.977197	0.947015	0.110690
H(31b)	0.142416	0.944256	0.080241

Table G.2: Volumes of the atomic basins, the number of electrons in the atomic basins and the net atomic charge (Q) in the asymmetric unit for $\rho^{MEM}(\mathbf{x})$. The basins of N(1,2,3), C(32) and O(5,6) encompass their bonded hydrogen atoms because the latter atoms do not constitute local maxima in the density.

Atom	Volume	Electrons	Q
C(1)	8.06	5.89	+0.11
C(2)	5.61	4.74	+1.26
C(3)	6.70	5.74	+0.26
C(4)	5.98	4.78	+1.22
C(5)	12.46	6.07	-0.07
C(6)	5.85	4.68	+1.32
C(7)	16.39	6.79	-0.79
C(8)	16.40	6.94	-0.94
C(20)	9.36	6.47	-0.47
C(21)	9.91	5.99	+0.01
C(22)	10.94	6.08	-0.08
C(23)	14.72	6.20	-0.20
C(24)	10.45	5.76	+0.24
C(25)	14.52	6.15	-0.15
C(26)	15.10	6.41	-0.41
C(31)	13.10	6.12	-0.12
$C(32)H_3$	37.49	8.81	+0.19
$N(1)H_3$	22.07	9.71	+0.29
$N(2)H_1$	17.00	8.47	-0.47
$N(3)H_1$	17.27	8.49	-0.49
O(1)	18.05	9.14	-1.14
O(2)	17.44	9.23	-1.23
O(3)	16.52	8.75	-0.75
O(4)	13.50	9.03	-1.03
$O(5)H_1$	19.43	9.62	-0.62
$O(6)H_1$	16.41	9.40	-0.40
H(1)	5.26	0.66	+0.34
H(3)	6.42	0.88	+0.12
H(5)	6.76	0.73	+0.27
H(7a)	9.28	0.76	+0.24

Continued on Next Page...

Table G.2 – Continued

Atom	Volume	Electrons	Q
H(7b)	3.33	0.48	+0.52
H(7c)	6.79	0.76	+0.24
H(8a)	4.35	0.53	+0.47
H(8b)	8.50	0.83	+0.17
H(8c)	7.54	0.63	+0.37
H(20a)	7.66	0.80	+0.20
H(20b)	7.87	0.94	+0.06
H(22)	7.80	0.76	+0.24
H(23)	6.52	0.75	+0.25
H(25)	7.73	0.75	+0.25
H(26)	7.67	0.77	+0.23
H(31a)	6.24	0.60	+0.40
H(31b)	8.53	0.82	+0.18
Total	488.98	197.90	+0.09

Table G.3: Electron densities and Laplacians at the BCPs of covalent bonds and RCPs (ring critical points) of Ala-Tyr-Ala with ethanol. Values are given for $\rho(BCP)$ (electrons/Å^3; first line) and $\nabla^2\rho(BCP)$ (electrons/Å^5; second line) derived from $\rho^{prior}(\mathbf{x})$, $\rho^{MEM}(\mathbf{x})$, the static electron density of the multipole model[83] and DFT calculations (B3LYP/6-311++G(3df,3pd)).[83]

Bond	Prior	MEM	Multipole	Calc.
C(2)–O(1)	2.10	2.50	2.83	2.72
	13.47	37.54	-35.7	-17.3
C(4)–O(2)	2.07	2.55	2.98	2.71
	12.17	24.58	-42.8	-16.8
C(6)–O(3)	2.08	2.55	2.81	2.61
	19.79	17.78	-30.1	-17.1
C(6)–O(4)	1.95	2.41	2.75	2.51
	8.65	9.41	-34.7	-21.8
C(24)–O(5)	1.66	1.85	-	-
	3.94	3.75	-	-
C(31)–O(6)	1.49	1.71	-	-
	4.04	1.04	-	-
C(1)–N(1)	1.36	1.64	1.71	1.59
	3.09	-6.64	-9.3	-13.2
C(2)–N(2)	1.74	2.11	2.43	2.32
	-0.96	-0.80	-24.4	-26.3
C(3)–N(2)	1.43	1.70	1.84	1.72
	1.60	-7.34	-12.5	-15.8
C(4)–N(3)	1.75	2.24	2.53	2.31
	-0.87	-12.80	-27.0	-26.1
C(5)–N(3)	1.43	1.73	1.88	1.71
	2.22	-10.06	-14.4	-15.3
C(1)–C(2)	1.18	1.56	1.75	1.72
	-0.07	-13.80	-13.6	-15.2
C(1)–C(7)	1.19	1.61	1.79	1.64
	-0.21	-10.68	-13.7	-13.6
C(3)–C(4)	1.17	1.57	1.76	1.71
	0.57	-12.09	-12.9	-15.0
C(3)–C(20)	1.14	1.47	1.63	1.55

Continued on Next Page...

Table G.3 – Continued

Bond	Prior	MEM	Multipole	Calc.
	0.35	-2.90	-10.9	-11.9
C(5)–C(6)	1.16	1.52	1.77	1.67
	-0.36	-13.55	-14.2	-14.2
C(5)–C(8)	1.17	1.52	1.68	1.62
	0.53	-6.24	-10.9	-13.3
C(20)–C(21)	1.22	1.51	-	-
	-0.21	-6.02	-	-
C(21)–C(22)	1.43	1.86	-	-
	-2.79	-12.42	-	-
C(22)–C(23)	1.45	1.88	-	-
	-2.89	-14.28	-	-
C(23)–C(24)	1.44	1.92	-	-
	-2.46	-12.77	-	-
C(24)–C(25)	1.44	1.95	-	-
	-2.98	-18.60	-	-
C(25)–C(26)	1.44	1.91	-	-
	-2.89	-16.52	-	-
C(21)–C(26)	1.43	1.97	-	-
	-2.28	-15.92	-	-
C(31)–C(32)	1.25	1.63	-	-
	0.85	-8.88	-	-
C(1)–H(1)	1.18	1.58	-	-
	-6.25	-19.82	-	-
C(3)–H(3)	1.18	1.54	-	-
	-6.09	-13.06	-	-
C(5)–H(5)	1.18	1.56	-	-
	-6.43	-17.56	-	-
C(7)–H(7a)	1.24	-	-	-
	-9.37	-	-	-
C(7)–H(7b)	1.25	1.55	-	-
	-10.74	-24.12	-	-
C(7)–H(7c)	1.24	1.52	-	-
	-10.72	-17.13	-	-
C(8)–H(8a)	1.26	1.58	-	-

Continued on Next Page...

Table G.3 – Continued

Bond	Prior	MEM	Multipole	Calc.
	-12.48	-21.13	-	-
C(8)–H(8b)	1.21	1.49	-	-
	-10.85	-15.33	-	-
C(8)–H(8c)	1.24	1.59	-	-
	-12.96	-24.70	-	-
C(20)–H(20a)	1.20	1.56	-	-
	-7.10	-17.62	-	-
C(20)–H(20b)	1.19	1.65	-	-
	-6.17	-18.33	-	-
C(22)–H(22)	1.16	1.55	-	-
	-6.07	-21.39	-	-
C(23)–H(23)	1.16	1.55	-	-
	-7.04	-15.83	-	-
C(25)–H(25)	1.15	1.57	-	-
	-6.41	-17.35	-	-
C(26)–H(26)	1.16	1.54	-	-
	-7.17	-15.86	-	-
C(31)–H(31a)	1.15	1.44	-	-
	-9.43	-16.52	-	-
C(31)–H(31b)	1.15	-	-	-
	-10.86	-	-	-
C(32)–H(32b)	1.18	-	-	-
	-14.18	-	-	-
N(1)–H(11a)	1.43	-	-	-
	-17.95	-	-	-
N(1)–H(11c)	1.44	-	-	-
	-19.33	-	-	-
N(2)–H(12)	1.49	-	-	-
	-21.37	-	-	-
N(3)–H(13)	1.49	-	-	-
	-21.58	-	-	-
Phenyl ring	0.30	0.16	-	-
	2.98	5.93	-	-
H(11c)-N(1)-C(1)-C(7)-H(7b)	-	0.15	-	-

Continued on Next Page...

Table G.3 – Continued

Bond	Prior	MEM	Multipole	Calc.
	-	0.34	-	-
H(20a)-C(20)-C(21)-C(22)-H(22)	-	0.11	-	-
	-	1.77	-	-

Table G.4: Geometrical, topological and energetic properties of hydrogen bonds of Ala-Tyr-Ala with ethanol (Figure G.2). d(H$\cdots$O) and d(D$\cdots$O) in Å, $\rho(BCP)$ in electrons/Å^3, eigenvalues and $\nabla^2\rho(BCP)$ in electrons/Å^5, G, V and H in kJ/mol·Bohr^3.

	$\rho(BCP)$	λ_1	λ_2	λ_3	$\nabla^2\rho(BCP)$	d(H$\cdots$O)	d(D$\cdots$O)	G	V	$\lvert V\rvert/G$	H
O(5)–H(15)$\cdots$O(6)	0.435	-3.62	-2.51	8.18	2.05	1.670	2.628	115.37	-174.82	1.52	-59.45
N(1)–H(11c)$\cdots$O(2)	0.330	-2.12	-1.64	5.99	2.23	1.720	2.750	89.70	-118.72	1.32	-29.02
N(1)–H(11b)$\cdots$O(5)	0.298	-2.07	-1.42	3.95	0.46	1.854	2.870	49.96	-87.30	1.75	-37.33
O(6)–H(16)$\cdots$O(3)	0.293	-2.27	-0.66	6.48	3.55	1.722	2.688	104.87	-112.95	1.08	-8.07
N(1)–H(11a)$\cdots$O(4)	0.291	-2.56	-1.81	6.56	2.20	1.773	2.786	79.83	-99.88	1.25	-20.05
N(3)–H(13)$\cdots$O(4)	0.165	-1.43	-0.34	3.78	2.00	1.962	2.961	51.87	-49.27	0.95	2.60
C(1)–H(1)$\cdots$O(1)	0.150	-1.62	-1.15	2.52	-0.25	2.194	3.030	8.72	-24.23	2.78	-15.51
N(2)–H(12)$\cdots$O(1)	0.113	-0.90	-0.52	3.90	2.47	2.011	2.974	53.12	-38.93	0.73	14.18

Appendix H

Histograms of $\Delta F(\mathbf{H}_i)/\sigma_i$

The distribution of residuals $\Delta F(\mathbf{H}_i)/\sigma_i$ according to Eq. 2.11 for α-glycine, L-alanine, the L-phenylalanine formic acid complex, and Ala-Tyr-Ala with water and Ala-Tyr-Ala with ethanol are displayed in Figure H.1. All residuals show a Gaussian distribution. The histogram of trialanine is given in Chapter 3, Fig. 3.2.

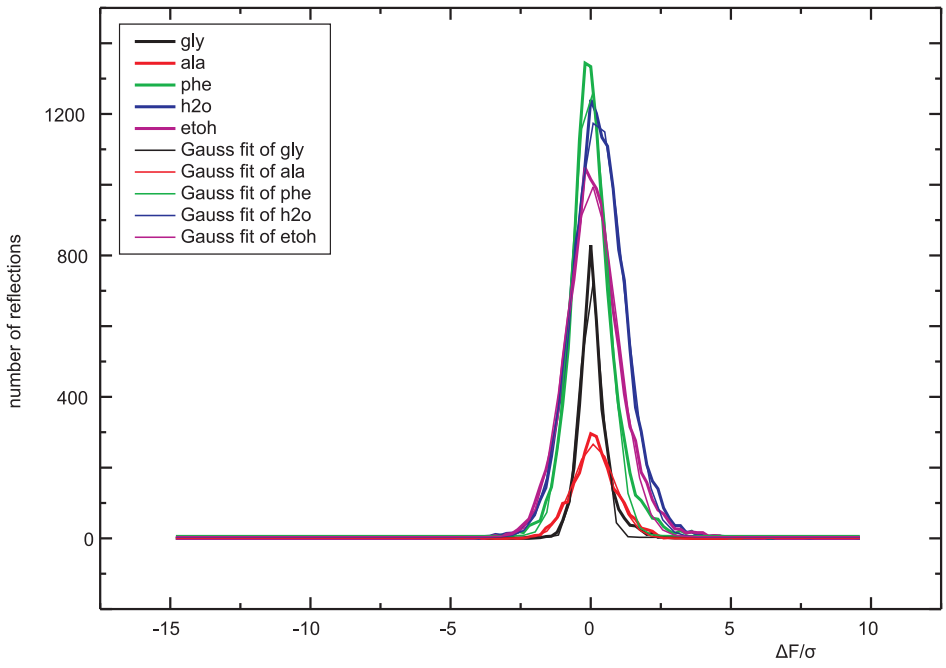

Figure H.1: Distribution of residuals $\Delta F(\mathbf{H}_i)/\sigma_i$ for α-glycine (gly), L-alanine (ala), L-phenylalanine formic acid complex (phe), and Ala-Tyr-Ala with water (h2o) and ethanol (etoh).

Bibliography

[1] N. K. Hansen and P. Coppens. *Acta Crystallogr., Sect. A*, 1978, **34**, 909–921.

[2] P. Coppens. *X-ray Charge Densities and Chemical Bonding*. Oxford University Press, 1997.

[3] A. Volkov, Y. Abramov, P. Coppens and C. Gatti. *Acta Crystallogr., Sect. A*, 2000, **56**, 332–339.

[4] A. Volkov, Y. Abramov and P. Coppens. *Acta Crystallogr., Sect. A*, 2001, **57**, 272–282.

[5] A. Volkov and P. Coppens. *Acta Crystallogr., Sect. A*, 2001, **57**, 395–405.

[6] A. O. Madsen, H. O. Sørensen, C. Flensburg, R. F. Stewart and S. Larsen. *Acta Crystallogr., Sect. A*, 2004, **60**, 550–561.

[7] T. Koritsanszky. In *Hydrogen bonding—new insights* (edited by S. J. Grabowski), Springer, Berlin, 2006, (pp. 441–470).

[8] W. Jauch and A. Palmer. *Acta Crystallogr., Sect. A*, 1993, **49**, 590–591.

[9] W. Jauch. *Acta Crystallogr., Sect. A*, 1994, **50**, 650–652.

[10] R. Y. de Vries, W. J. Briels and D. Feil. *Phys. Rev. Lett.*, 1996, **77**, 1719–1722.

[11] M. Takata and M. Sakata. *Acta Crystallogr., Sect. A*, 1996, **52**, 287–290.

[12] P. Roversi, J. J. Irwin and G. Bricogne. *Acta Crystallogr., Sect. A*, 1998, **54**, 971–996.

[13] R. F. W. Bader. *Atoms in Molecules - A Quantum Theory*. Oxford University Press, Oxford, 1994.

[14] E. Rödel, M. Messerschmidt, B. Dittrich and P. Luger. *Org. Biomol. Chem.*, 2006, **4**, 475–481.

[15] E. T. Jaynes. *Phys. Rev.*, 1957, **106**, 620–630.

[16] E. T. Jaynes. *The Maximum Entropy Formalism.* 15–118 pp., Cambridge: MIT Press, 1979.

[17] C. J. Gilmore. *Acta Crystallogr., Sect. A*, 1996, **52**, 561–589.

[18] C. Bricogne. *Acta Crystallogr., Sect. A*, 1988, **44**, 517–545.

[19] C. Bricogne. *Acta Crystallogr., Sect. D*, 1993, **49**, 37–60.

[20] D. M. Collins. *Nature*, 1982, **298**, 49–51.

[21] M. Sakata and M. Sato. *Acta Crystallogr., Sect. A*, 1990, **46**, 263–270.

[22] M. Takata, B. Umeda, E. Nishibori, M. Sakata, Y. Saito, M. Ohno and H. Shinohara. *Nature*, 1995, **377**, 46–49.

[23] P. Luger, M. Messerschmidt, S. Scheins and A. Wagner. *Acta Crystallogr., Sect. A*, 2004, **60**, 390–396.

[24] Y. Kuroiwa, N. Sato, S. Negishi, H. Negishi and S. Aoyagi. *J. Phys. Soc. Jpn.*, 2003, **72**(11), 2811–2815.

[25] Y. Kuroiwa, S. Aoyagi and A. Sawada. *Phys. Rev. Lett.*, 2001, **87**(21), 2176011–2176014.

[26] E. Nishibori, T. Nakamura, M. Arimoto, S. Aoyagi, H. Ago, M. Miyano, T. Ebisuzakic and M. Sakata. *Acta Crystallogr., Sect. D*, 2008, **64**, 237–247.

[27] M. Takata. *Acta Crystallogr., Sect. A*, 2008, **64**, 232–245.

[28] S. Israel, R. Saravanan, N. Srinivasan and R. K. Rajaram. *J. Phys. Chem. Solids*, 2003, **64**, 43–49.

[29] K. S. S. Ali, R. Saravanan, S. Israel and R. K. Rajaram. *Bull. Mater. Sci.*, 2006, **29**(2), 107–114.

[30] N. Ishimatsu, Y. Ohishi, M. Takata, E. Nishibori, M. Sakata, J. Hayashi, I. Shirotani and O. Shimomura. *J. Phys.: Condens. Matter*, 2002, **14**, 11143–11146.

[31] F. S. Sarrat, R. B. Guimaraes, M. A. Continentino and J. C. Fernandes. *Phys. Rev. B*, 2005, **71**, 2244131–2244136.

[32] F. Cargnoni, E. Nishibori, P. Rabiller, L. Bertini, G. J. Snyder, M. Christensen, C. Gatti and B. B. Iversen. *Chem. Eur. J.*, 2004, **10**, 3861–3870.

[33] H. Yamada, W. S. Shi and C. N. Xu. *J. Appl. Cryst.*, 2004, **37**, 698–702.

[34] R. E. Dinnebier, M. Schneider, S. van Smaalen, F. Olbrich and U. Behrens. *Acta Crystallogr., Sect. B*, 1999, **55**, 35–44.

[35] S. W. Kim, S. Matsuishi, T. Nomura, Y. Kubota, M. Takata, K. Hayashi, T. Kamiya, M. Hirano and H. Hosono. *Nano Letters*, 2007, **7**, 1138–1143.

[36] L. Palatinus and S. van Smaalen. *Z. Kristallogr.*, 2004, **219**, 719–729.

[37] L. Palatinus, M. Amami and S. van Smaalen. *Acta Crystallogr., Sect. B*, 2004, **60**, 127–137.

[38] A. Hishikawa, N. Igawa, H. Yamauchi and Y. Ishii. *J. Phys. Chem. Solids*, 2005, **66**, 1810–1814.

[39] A. Hishikawa, N. Igawa, H. Yamauchi and Y. Ishii. *J. Chem. Phys.*, 2006, **125**, 034505.

[40] A. Bentien, B. B. Iversen, J. D. Bryan, G. D. Stucky, A. E. C. Palmqvist, A. J. Schultz and R. W. Henning. *J. Appl. Phys.*, 2002, **91**(9), 5694–5699.

[41] C. E. Shannon. *Bell Sys. Tech.*, 1948, **27**, 379–423.

[42] A. Hofmann, J. Netzel and S. van Smaalen. *Acta Crystallogr., Sect. B*, 2007, **63**, 285–295.

[43] J. Skilling and R. K. Bryan. *Mon. Not. R. Astr. Soc.*, 1984, **211**, 111–124.

[44] H. Tanaka, M. Takata, E. Nishibori, K. Kato, T. Iishi and M. Sakata. *J. Appl. Crystallogr.*, 2002, **35**, 282–286.

[45] S. van Smaalen, L. Palatinus and M. Schneider. *Acta Crystallogr., Sect. A*, 2003, **59**, 459–469.

[46] S. F. Gull. In *Maximum Entropy and Bayesian Methods* (edited by J. Skilling), Kluwer, Dordrecht, 1989, (pp. 53–71).

[47] J. Skilling. In *Maximum Entropy and Bayesian Methods* (edited by J. Skilling), Kluwer, Dordrecht, 1989, (pp. 45–52).

[48] S. Kumazawa, Y. Kubota, M. Takata and M. S. Y. Ishibashi. *J. Appl. Crystallogr.*, 1993, **26**, 453–457.

[49] S. F. Gull and J. Skilling. MEMSYS5 *v1.2 program package.* Maximum Entropy Data Consultants Ltd., Suffolk, U.K., 1999.

[50] L. Palatinus and S. van Smaalen. *Acta Crystallogr., Sect. A*, 2002, **58**, 559–567.

[51] L. Palatinus and S. van Smaalen. *Acta Crystallogr., Sect. A*, 2005, **61**, 363–372.

[52] R. Y. de Vries, W. J. Briels and D. Feil. *Acta Crystallogr., Sect. A*, 1994, **50**, 383–391.

[53] J. Netzel, A. Hofmann and S. van Smaalen. *CrystEngComm*, 2008, **10**, 335–343.

[54] B. Buck and V. A. Macaulay (Eds.). *Maximum Entropy in Action.* Clarendon Press, Oxford, 1991.

[55] E. T. Jaynes. *Maximum Entropy and Bayesian Methods in Applied Statistics.* 26–58 pp., Cambridge University Press, 1986.

[56] P. Roversi. *Electron, Spin and Momentum Densities and Chemical Reactivity.* Dordrecht: Kluwer Academic Publishers, 2002.

[57] B. B. Iversen, J. L. Jensen and J. Danielsen. *Acta Crystallogr., Sect. A*, 1997, **53**, 376–387.

[58] S. F. Gull and J. Skilling. *Quantified Maximum Entropy,* MEMSYS5 *Users' Manual.* Maximum Entropy Data Consultants Ltd., Suffolk, U.K., 1999.

[59] E. Rödel. Diplomarbeit, Freie Universität Berlin, 2003.

[60] V. Petříček, M. Dušek and L. Palatinus. *JANA*2000*, The Crystallographic Computing System.* Institute of Physics, Academy of Sciences of the Czech Republic, Praha, 2000.

[61] T. Steiner and W. Saenger. *Acta Crystallogr., Sect. A*, 1993, **49**, 379–384.

[62] F. Benabicha, V. Pichon-Pesme, C. Jelsch, C. Lecomte and A. Khmou. *Acta Crystallogr., Sect. B*, 2000, **56**, 155–165.

[63] V. Pichon-Pesme, H. Lachekar, M. Souhassou and C. Lecomte. *Acta Crystallogr., Sect. B*, 2000, **56**, 728–737.

[64] B. Dittrich, C. B. Hübschle, M. Messerschmidt, R. Kalinowski, D. Girnt and P. Luger. *Acta Crystallogr., Sect. A*, 2005, **61**, 314–320.

[65] A. Bach, D. Lentz and P. Luger. *J. Phys. Chem. A*, 2001, **105**, 7405–7412.

[66] C. Gatti, R. Bianchi, R. Destro and F. Merati. *J. Mol. Struct. (Theochem)*, 1992, **255**, 409–433.

[67] V. Pichon-Pesme, C. Jelsch, B. Giullot and C. Lecomte. *Acta Crystallogr., Sect. A*, 2004, **60**, 204–208.

[68] T. S. Koritsanszky and P. Coppens. *Chem. Rev.*, 2001, **101**, 1583–1627.

[69] B. Zarychta, V. Pichon-Pesme, B. Guillot, C. Lecomte and C. Jelsch. *Acta Crystallogr., Sect. A*, 2007, **63**, 108–125.

[70] P. M. Dominiak, A. Volkov, X. Li, M. Messerschmidt and P. Coppens. *J. Chem. Theory Comput.*, 2007, **3**, 232–247.

[71] B. Dittrich, C. B. Hübschle, P. Luger and M. A. Spackman. *Acta Crystallogr., Sect. D*, 2006, **62**, 1325–1335.

[72] S. Mebs, M. Messerschmidt and P. Luger. *Z. Kristallogr.*, 2006, **221**, 656–664.

[73] A. Hofmann, R. Kalinowski, P. Luger and S. van Smaalen. *Acta Crystallogr., Sect. B*, 2007, **63**, 633–634.

[74] R. Destro, P. Roversi, M. Barzaghi and R. E. Marsh. *J. Phys. Chem. A*, 2000, **104**, 1047–1054.

[75] B. Bagautdinov, J. Luedecke, M. Schneider and S. van Smaalen. *Acta Crystallogr., Sect. B*, 1998, **54**, 626–634.

[76] A. J. C. Wilson (Ed.). *International Tables for Crystallography*, vol. C. Kluwer Academic Publishers, 1995.

[77] U. Koch and P. L. A. Popelier. *J. Phys. Chem.*, 1995, **99**, 9747–9754.

[78] G. A. Jeffrey. *An Introduction to Hydrogen Bonding.* Oxford University Press, Oxford, 1997.

[79] R. Destro, R. E. Marsh and R. Bianchi. *J. Phys. Chem.*, 1988, **92**, 966–973.

[80] A. Wagner and P. Luger. *J. Mol. Struct.*, 2001, **595**, 39–46.

[81] R. Flaig, T. Koritsanszki, B. Dittrich, A. Wagner and P. Luger. *J. Am. Chem. Soc.*, 2002, **124**, 3407–3417.

[82] S. Scheins, B. Dittrich, M. Messerschmidt, C. Paulmann and P. Luger. *Acta Crystallogr., Sect. B*, 2004, **60**, 184–190.

[83] L. Checinska, S. Mebs, C. B. Hübschle, D. Förster, W. Morgenroth and P. Luger. *Org. Biomol. Chem.*, 2006, **4**, 3242–3251.

[84] R. Kalinowski, B. Dittrich, C. B. Hübschle, C. Paulmann and P. Luger. *Acta Crystallogr., Sect. B*, 2007, **63**, 753–767.

[85] Y. A. Abramov. *Acta Crystallogr., Sect. A*, 1997, **53**, 264–272.

[86] A. Cremer and E. Kraka. *Croat. Chem. Acta*, 1984, **57**(6), 1259–1281.

[87] A. Cremer and E. Kraka. *Angew. Chem.*, 1984, **96**(8), 612–614.

[88] E. Espinosa, M. Souhassou, H. Lachekar and C. Lecomte. *Acta Crystallogr., Sect. B*, 1999, **55**, 563–572.

[89] E. Espinosa, E. Molins and C. Lecomte. *Chem. Phys. Lett.*, 1998, **285**, 170–173.

[90] E. Espinosa, I. Alkorta, J. Elguero and E. Molins. *J. Chem. Phys.*, 2002, **117**(12).

[91] E. Espinosa, C. Lecomte and E. Molins. *Chem. Phys. Lett.*, 1999, **300**, 745–748.

[92] M. A. Spackman. *Chem. Phys. Lett.*, 1999, **301**, 425–429.

[93] R. T. Downs, G. V. Gibbs, M. B. Boisen Jr. and K. M. Rosso. *Phys. Chem. Minerals*, 2002, **29**, 369–385.

[94] W. H. Baur. *Acta Crystallogr., Sect. B*, 1972, **28**, 1456–1465.

[95] N. Ohtomo and K. Arakawa. *Bull. Chem. Soc. Jpn.*, 1995, **51**(6), 1649–1654.

[96] R. Destro, R. Soave and M. Barzaghi. *J. Phys. Chem. B*, 2008, **112**, 5163–5174.

[97] R. Destro and L. Lo Presti. Personal communication; the deformation map is based on the final model of L-alanine described in *J. Phys. Chem B*, 2008, **112**, 5163–5174. (For the previous model of L-alanine see the publication of Destro *et al.*, *J. Phys. Chem.*, 1988, **92**, 966–973), 2008.

[98] P. Roversi and R. Destro. *Chem. Phys. Lett.*, 2004, **386**, 472–478.

[99] A. E. Whitten, P. Turner, W. T. Klooster, R. O. Piltz and M. A. Spackman. *J. Phys. Chem. A*, 2006, **110**, 8763–8776.

[100] K. A. Lyssenko, D. V. Lyubetsky, A. B. Sheremetev and M. Y. Antipin. *Russ. Chem. Bull., Int. Ed.*, 2005, **54**(4), 924–932.

[101] S. Grabowsky, T. Pfeuffer, L. Checinska, M. Weber, W. Morgenroth, P. Luger and T. Schirmeister. *Eur. J. Org. Chem.*, 2007, .

[102] A. Wagner, R. Flaig, B. Dittrich, H. Schmidt, T. Koritsanszky and P. Luger. *Chem. Eur. J.*, 2004, **10**, 2977–2982.

[103] P. M. Dominiak, A. Makal, P. R. Mallinson, K. Trzcinska, J. Eilmes, E. Grech, M. Chryszcz, W. Minor and K. Wozniak. *Chem. Eur. J.*, 2006, **12**, 1941–1949.

[104] Y. Marechal. *The Hydrogen Bond and the Water Molecule*. Elsevier, Radarweg 29, PO Box 211, 1000 Amsterdam; The Boulevard, Langford Lane, Kidlington, Oxford Ox5, UK., 1st edn., 2007.

[105] L. Palatinus and S. van Smaalen. *PRIOR - A computer program for calculation electron densities from analytical atomic form factors*. Laboratory of Crystallography, University of Bayreuth, D-95440 Bayreuth, Germany, 2005.

[106] S. F. Gull and J. Skilling. MEMSYS5 *v1.2 program package*. Maximum Entropy Data Consultants Ltd., United Kingdom, 1999.

Acknowledgements

Danksagung

Ich möchte mich ganz besonders bedanken bei Professor Sander van Smaalen, der als geduldiger Doktorvater und Mentor meine Dissertation betreut hat. In meiner Zeit am Lehrstuhl habe ich dank seiner Hilfe Wissen über Kristallographie, wissenschaftliches Arbeiten, und das Schreiben wissenschaftlicher Texte angereichert. Ich habe mich außerdem mit der Mathematik anfreunden können. Besonders nach Auslaufen meiner Doktorandenstelle hat er mir mit Rat bezüglich persönlicher Dinge zur Seite gestanden.

Ein sehr großer und lieber Dank gilt meinen Eltern Brigitte und Hans-Joachim, meiner Schwester Madeleine und meinem Schatz Christoph für ihre besondere emotionale und finanzielle Unterstützung, die ich gerade in schwierigen Zeiten gebraucht und sehr gerne in Anspruch genommen habe. Ohne ihren unschätzbaren Beistand hätte ich die Doktorandenzeit nicht so gut gemeistert und wäre heute nicht an dem Punkt angekommen, an dem ich stolz auf mich bin. Glücklich und frohen Mutes ist der, der geliebte Menschen an seiner Seite hat. Ihnen widme ich meine Doktorarbeit.

Bei meinen Kollegen am Lehrstuhl möchte ich mich sehr herzlich bedanken für prompte Hilfe und Ratschläge bei allen möglichen Problemen, die ich ihnen angetragen habe, sowie auch für die Gespräche zur geistigen Entspannung zwischendurch. Ich habe mich als Mitarbeiterin am Lehrstuhl für Kristallographie sehr wohl gefühlt.

Professor Luger danke ich für die Begutachtung meiner Dissertation sowie die Überlassung von Beugungsdaten.
Professor Destro möchte ich danken für die privaten wissenschaftlichen Gespräche bezüglich Glycin und die Überlassung seiner Beugungsdaten.
Für die schnelle und äußerst hilfreiche Unterstützung in Fragen bezüglich JANA2000 oder bei Computergeschichten möchte ich mich bei L. Palatinus, M. Dusek und V. Petricek bedanken.

Meinen Freunden gilt ebenfalls ein großes Dankeschön für ihre Unterstützung während meiner Zeit in Bayreuth, in der sie für mich da waren und sind.

Erklärung

Dieses Dokument wurde vom Autor selbständig und nur mit den angegebenen Quellen und Hilfsmitteln erstellt. Der Autor hat bisher keine Promotionsversuche unternommen.

Bayreuth, 2008

Jeanette Netzel