Acta Crystallographica Section B
**Structural
Science**
ISSN 0108-7681
Editor: **Carolyn P. Brock**

Transferability and reproducibility in electron-density studies – bond-topological and atomic properties of tripeptides of the type L-alanyl-*X*-L-alanine

Simon Grabowsky, Roman Kalinowski, Manuela Weber, Diana Förster, Carsten Paulmann and Peter Luger

Acta Cryst. (2009). B**65**, 488–501

Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Transferability and reproducibility in electron-density studies – bond-topological and atomic properties of tripeptides of the type L-alanyl-*X*-L-alanine

Simon Grabowsky,[a] Roman Kalinowski,[a] Manuela Weber,[a] Diana Förster,[a] Carsten Paulmann[b] and Peter Luger[a]*

[a]Freie Universität Berlin, Institut für Chemie und Biochemie/Kristallographie, Fabeckstrasse 36a, 14195 Berlin, Germany, and [b]Universität Hamburg, Mineralogisch–Petrographisches Institut, c/o HASYLAB/DESY, Notkestrasse 85, 22607 Hamburg, Germany

Correspondence e-mail:
luger@chemie.fu-berlin.de

Received 19 January 2009
Accepted 5 May 2009

In the last decade three different data bank approaches have been developed that are intended to make electron-density examinations of large biologically important molecules possible. They rely on Bader's concept of transferability of submolecular fragments with retention of their electronic properties. Therefore, elaborate studies on the quantification of transferability in experiment and theory are still very important. Tripeptides of the type L-alanyl-*X*-L-alanine (*X* being any of the 20 naturally encoded amino acids) serve as a model case between amino acids and proteins. The two experimental electron-density determinations (L-alanyl-L-histidinyl-L-alanine and L-alanyl-L-phenylalanyl-L-alanine, highly resolved synchrotron X-ray diffraction data sets) performed in this study and theoretical calculations on all 20 different L-alanyl-*X*-L-alanine molecules contribute to a better estimation of transferability in the peptide case. As a measure of reproducibility and transferability, standard deviations from averaging over bond-topological and atomic properties of atoms or bonds that are considered equal in their chemical environments were calculated. This way, transferability and reproducibility indices were introduced. It can be shown that experimental transferability indices generally slightly exceed experimental reproducibility indices and that these larger deviations can be attributed to chemical effects such as changes in the geometry (bond lengths and angles), the polarization pattern and the neighboring sphere due to crystal packing. These effects can partly be separated from each other and quantified with the help of gas-phase calculations at optimized and experimental geometries. Thus, the degree of transferability can be quantified in very narrow limits taking into account experimental errors and chemical effects.

1. Introduction

About 20 years ago, Bader introduced the topological analysis of electron density $\varrho(\mathbf{r})$ (QTAIM = quantum theory of atoms in molecules) as a powerful tool for extracting quantitative chemical information from theoretically or experimentally derived $\varrho(\mathbf{r})$ (Bader, 1990). The application of QTAIM has since developed into a widely accepted standard method (Coppens, 2005*a*,*b*). The additivity of submolecular fragments to build up the whole system and the transferability of these fragments from one system to another with the fragment in the same chemical environment are the main propositions of the QTAIM. Bader stated that '. . . the use of the zero-flux surface for the definition of an atom or functional grouping of atoms maximizes the extent of the transferability of its properties between systems, a characteristic essential to the role of the atomic concept in chemistry' and 'It is the purpose of the present discussion to explore the limiting case where a group

Table 1
Multiple electron-density measurements of the same compound in the literature; reproducibility indices from experiment.

Compound name	No. of data sets	$\overline{\sigma}_{\rm rep,exp}(\varrho_{\rm bcp})$ in e Å^{-3}	$\overline{\sigma}_{\rm rep,exp}(\nabla^2\varrho_{\rm bcp})$ in e Å^{-5}	$\overline{\sigma}_{\rm rep,exp}(Q_{001})$ in e	$\overline{\sigma}_{\rm rep,exp}(V_{001})$ in Å^3	Ref.
cyclo-(D,L-Pro)$_2$-(L-Ala)$_4$	2	0.06	1.9	–	–	(*a*)
Strychnine	4	0.05	1.8	–	–	(*b*)
L-Ala-Gly-L-Ala	3	0.07	3.4	0.10	0.4	(*c*)
Acceptor-substituted aziridine	2	0.06	4.2	0.13	0.4	(*d*)
Gly-L-Thr	2	0.07	3.9	–	–	(*e*), (*f*)
L-Asp	2	0.12	4.7	–	–	(*g*), (*h*)
Mean values: $\overline{\sigma}^{o}_{\rm rep,exp}$		0.07	3.3	0.12	0.4	

References: (*a*) Dittrich *et al.* (2002); (*b*) Messerschmidt *et al.* (2005); (*c*) Förster *et al.* (2006); (*d*) Grabowsky *et al.* (2008); (*e*) Dittrich *et al.* (2000); (*f*) Benabicha *et al.* (2000); (*g*) Arnold *et al.* (2000); (*h*) Flaig *et al.* (1999*a*,*b*).

is transferable with little or no change in its properties and, correspondingly, its perturbation of the remainder of the system is minimized' (Bader, 1990).

This discussion is still ongoing and very important with regard to the data bank applications that have been developed in the last decade in the groups of Lecomte (Pichon-Pesme *et al.*, 1995; Jelsch *et al.*, 1998), Coppens (University at Buffalo Pseudoatom Data Bank = UBDB; Dominiak *et al.*, 2007; Volkov *et al.*, 2007), and our group (Invariom data bank; Dittrich *et al.*, 2004*a*,*b*, 2006). These data banks that are intended to make electron-density examination on proteins or other biologically important macromolecules possible by multipole parameter transfer (Volkov *et al.*, 2006; Hansen & Coppens, 1978) rely on the transferability concept. Therefore, extensive studies on the quantification of transferability of bond-topological and atomic properties from experimental and theoretical data are still indispensable. The multipoles of model compounds that can be transferred to the macromolecule of interest are derived from experimental multipole modeling (Lecomte's approach), from theoretical calculations at experimental geometries within the UBDB, or from theoretical calculations at fully optimized geometries within the Invariom approach. In the Invariom data bank, only the nearest or next-nearest neighbor influence is assumed when choosing suitable model compounds (Dittrich *et al.*, 2006; Hübschle *et al.*, 2007). The nearest-neighbor approximation (NNA) is used in most cases; the next-nearest-neighbor approximation (NNNA) is only used for delocalized systems such as the peptide bond, heavier atoms like S and P, and H atoms. For this study on transferability, we use the same approximations as in the Invariom data bank to judge whether the properties of two atoms or bonds have to be averaged or treated independently. The situation is simplified because all atoms in the peptide bond that are unequal to other atoms in the molecule in the NNNA are already unequal in the NNA. Also, $O_{\rm pep}$ and $O_{\rm carbox}$ are unequal in the NNA because they carry a different charge. Therefore, we always refer to the NNA in the following.

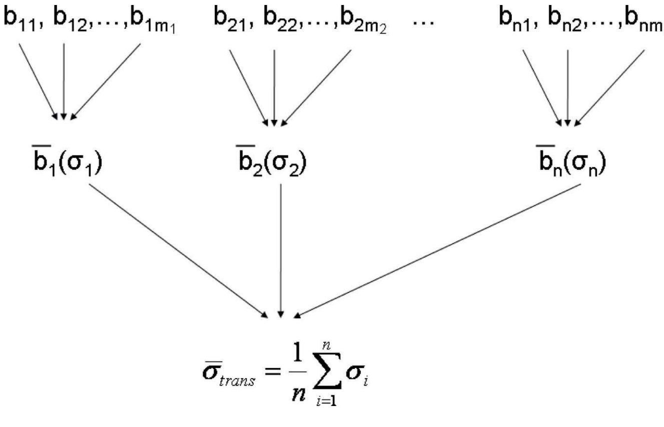


Figure 1
Averaging procedure to gain the transferability index $\overline{\sigma}_{\rm trans}$.

For the sake of clarity concerning the various statistical results of this study, we want to make a clear distinction between reproducibility and transferability. In the following, reproducibility describes the agreement of a property of the same atom or bond of the same compound in the same modification from different calculations or measurements. This way, reproducibility serves as a measure of experimental and/or methodological errors (McNaught & Wilkinson, 1997). Transferability describes the agreement of a property of atoms or bonds in the same chemical environment within one molecule and/or within different compounds. As a measure of agreement, the standard deviation σ $\{\sigma = [(1/n-1)\sum_{i=1}^{n}(x_i-\overline{x})^2]^{1/2}\}$ is used (McNaught & Wilkinson, 1997).

In the electron-density field, it is common practice to consider, amongst others, the electron density $\varrho_{\rm bcp}$ and the Laplacian $\nabla^2\varrho_{\rm bcp}$ at a bond-critical point (b.c.p.) as bond-topological properties. As atomic properties, atomic volumes V_{001} [$\varrho(\mathbf{r}) > 0.001$ a.u. at each point $\mathbf{r}$ in an atomic basin] and atomic charges Q_{001} are usually considered (Bader, 1990). In transferability studies, the above-mentioned properties are averaged for atoms or bonds that are considered equal in terms of the NNA, providing a mean value and a standard deviation (Kalinowski *et al.*, 2007). The applied procedure is illustrated by the example scheme in Fig. 1. Assume that transferability is to be examined for a class of compounds with n different types of bonds $b_1 \ldots b_n$. For each bond type b_i, m_i contributors (bonds that are considered equal in the NNA) to a property of the bond exist: $b_{i1}, b_{i2} \ldots b_{im_i}$. Averaging over all m_i values of the property of the bond type b_i gives mean values and corresponding standard deviations $\overline{b_i}(\sigma_i)$. To obtain a transferability index $\overline{\sigma}_{\rm trans}$ for all n bonds, we average over all σ_i of the class of compounds: $\overline{\sigma}_{\rm trans} = (1/n)\sum_{i=1}^{n}\sigma_i$. For a bond, $\overline{\sigma}_{\rm trans}$ is calculated for the properties $\varrho_{\rm bcp}$ and $\nabla^2\varrho_{\rm bcp}$; for an atom, $\overline{\sigma}_{\rm trans}$ is calculated for the properties V_{001} and Q_{001}. Accordingly, reproducibility indices $\overline{\sigma}_{\rm rep}$ can be obtained. Here, the contributors are identical

Table 2
Crystallographic and refinement data.

	AHA	AFA
Crystal data		
Chemical formula	$C_{12}H_{19}N_5O_4{\cdot}C_3H_8O{\cdot}H_2O$	$C_{15}H_{21}N_3O_4{\cdot}C_3H_7NO$
M_r	375.43	380.44
Crystal system, space group	Monoclinic, $P2_1$	Monoclinic, $P2_1$
Temperature (K)	100	100
a, b, c (Å)	8.741 (2), 9.420 (2), 11.989 (2)	9.598 (2), 8.939 (1), 12.170 (1)
β (°)	95.49 (3)	108.75 (1)
V (Å^3)	982.7 (3)	988.7 (2)
Z	2	2
Radiation type, λ (Å)	Synchrotron, 0.5600 (2)	Synchrotron, 0.5600 (2)
μ (mm^{-1})	0.061	0.058
Crystal form, size	Block, 0.40 × 0.35 × 0.25	Block, 0.45 × 0.25 × 0.15
Data collection		
Diffractometer	κ-axis MarCCD 165 detector	κ-axis MarCCD 165 detector
Data collection method	φ scans	φ scans
Absorption correction	None	None
No. of measured, independent and observed reflections	152 583, 13 235, 11 918	180 324, 13 499, 12 060
Criterion for observed reflections	$F^2 > 2\sigma(F^2)$	$F^2 > 2\sigma(F^2)$
$R_{\rm int}$	0.027	0.059
$\theta_{\max}$ (°)	40.6	40.6
Multipole refinement		
Refinement on	F^2	F^2
$R(F)[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.020, 0.035, 1.96	0.022, 0.033, 1.95
No. of reflections	11 918	12 060
No. of parameters	783	733
H-atom treatment	See §2.1	See §2.1
$(\Delta/\sigma)_{\max}$ (e Å^{-3})	< 0.0001	< 0.0001
$\Delta\rho_{\max}$, $\Delta\rho_{\min}$ (e Å^{-3})	0.14, −0.22	0.33, −0.34

Computer programs used: *XDS* (Kabsch, 1993), *SHELXS*97 (Sheldrick, 2008), *XD*2006 (Volkov *et al.*, 2006).

bonds or atoms from different measurements or calculations.

In experimental studies, transferability and reproducibility cannot be separated from each other since experimental errors can never be avoided. If transferability is therefore examined with respect to reproducibility, one has a basis for chemical interpretations of differences and a justification for the application of multipole transfer using data bank approaches. Experimental studies on reproducibility are rare. Table 1 summarizes reproducibility indices derived from multiple experimental electron-density studies. If we average again over each column in Table 1 we obtain overall experimental $\overline{\sigma}^{o}_{\rm rep,exp}$ values which should serve in the following as a benchmark for the mean experimental error of each property (see last line in Table 1).

Reproducibility in terms of *ab initio* calculations depends on the variation of method and basis set. Flaig *et al.* (2002) carried out an extensive study on the reproducibility of bond-topological properties in serine using different methods (Hartree–Fock, B3LYP/ B3PW91 of density functional theory and MP2/ MP3/ MP4 of Møller–Plesset perturbation theory) and different basis sets (Pople basis sets with different numbers of diffuse and polarization functions). They found $\overline{\sigma}_{\rm rep,theo}(\varrho_{\rm bcp})$ = 0.04 e Å^{-3} and $\overline{\sigma}_{\rm rep,theo}(\nabla^2\varrho_{\rm bcp})$ = 3.1 e Å^{-5}. This study shall only hold as one example for numerous studies in the literature describing method and basis-set dependence of topological parameters. However, if one studies transferability between different systems using the same method and basis set, there should be no methodological error to be considered and any differences should be chemically interpretable. Therefore, we decided to use only one method and basis set [B3LYP/6–311++G(2d,2p)] to concentrate on the effects of transferability.

Mebs *et al.* (2006) summarized the experimental transferability results of the topological analysis of $\varrho(\mathbf{r})$ for the 16 amino acids for which multipole modelings on high-resolution X-ray diffraction data sets were carried out. They found $\overline{\sigma}_{\rm trans,exp}(\varrho_{\rm bcp})$ = 0.09 e Å^{-3} and $\overline{\sigma}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp})$ = 3.5 e Å^{-5}. As a first extensive work on transferability from theoretical calculations, Matta and Bader carried out a pilot study for all 20 genetically encoded amino acids (Matta & Bader, 2000, 2002, 2003). The transferability indices they found for the mean values of the density $\varrho(\mathbf{r})$ and the Laplacian $\nabla^2\varrho(\mathbf{r})$ at the bond-critical point of all common bonds are $\overline{\sigma}_{\rm trans,theo}(\varrho_{\rm bcp})$ = 0.02 e Å^{-3} and $\overline{\sigma}_{\rm trans,theo}(\nabla^2\varrho_{\rm bcp})$ = 0.5 e Å^{-5}. For the atomic properties of all common atoms, they found $\overline{\sigma}_{\rm trans,theo}(Q_{001})$ = 0.02 e and $\overline{\sigma}_{\rm trans,theo}$ (V_{001}) = 1.6 Å^3. Dittrich *et al.* (2003) performed experimental studies on dipeptides resulting in $\overline{\sigma}_{\rm trans,exp}(\varrho_{\rm bcp})$ = 0.10 e Å^{-3}, $\overline{\sigma}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp})$ = 3.0 e Å^{-5}, $\overline{\sigma}_{\rm trans,exp}(Q_{001})$ = 0.06 e and $\overline{\sigma}_{\rm trans,exp}$ (V_{001}) = 0.4 Å^3. No systematic theoretical study on dipeptides is available. However, studies on the transferability of multipole populations for some tripeptides exist (Pichon-Pesme *et al.*, 1995; Pichon-Pesme & Lecomte, 1998; Koritsanszky *et al.*, 2002) and one study (Pichon-Pesme *et al.*, 2000) comparing the experimental bond-topological properties of the tripeptide tyrosyl-glycyl-glycine with other small peptides [$\overline{\sigma}_{\rm trans,exp}(\varrho_{\rm bcp})$ = 0.09 e Å^{-3} and $\overline{\sigma}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp})$ = 2.6 e Å^{-5}].

In recent years, we have entered into a systematic study on the class of tripeptides L-alanyl-*X*-L-alanine (A*X*A). Measurements of six different A*X*A tripeptides with *X* being L-alanyl (Rödel *et al.*, 2006), L-tyrosyl (Cheçińska *et al.*, 2006), glycyl (Förster *et al.*, 2006), L-prolyl (Kalinowski *et al.*, 2007), L-histidinyl (this study) and L-phenylalanyl (this study), have been performed. The measurements were carried out using different synchrotron beamlines at HASYLAB/ DESY in Hamburg and at the Swiss Light Source (PSI, Villigen) as well as using the in-house diffractometer (Mo *Kα* radiation, 2 kW sealed tube) at temperatures ranging from 100 to 9 K. Different integration and data reduction software routines had to be employed, but the multipole modeling strategy of the main chain was identical. In addition to the comparison of bond-topological and atomic properties between the tripeptides of the type A*X*A, the transferability of bond-topological and atomic properties of the middle amino acids histidine and phenylalanine between different molecule types is examined. Dipeptides and corresponding free amino acids were chosen as systems for comparison because effects upon peptide bond formation can be studied. The hydrogen-bonding networks and the electrostatic potentials were compared for the two tripeptides of this study [L-alanyl-L-histidinyl-L-alanine (AHA) and L-alanyl-L-phenylalanyl-L-alanine (AFA)], in addition to bond-topological and atomic properties. Moreover, theoretical calculations on the currently available experimental geometries were performed as well as theoretical calculations on optimized geometries of all 20 different A*X*A molecules with *X* being any of the genetically encoded amino acids.

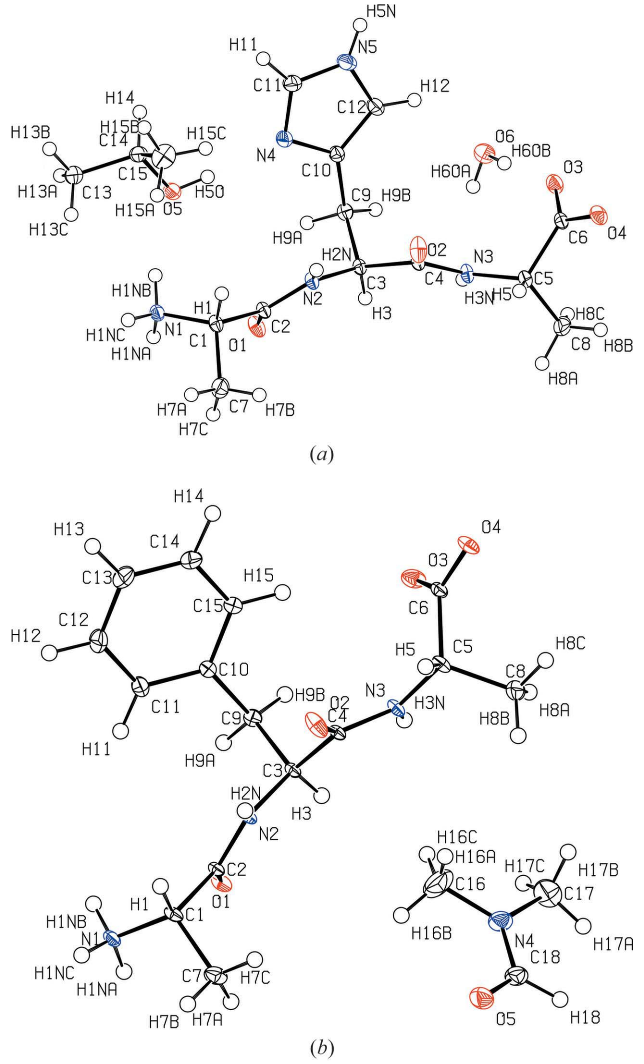


Figure 2
ORTEP representation (Burnett & Johnson, 1996) of the 100 K structures of (*a*) AHA and (*b*) AFA at 50% probability; the atom-numbering scheme is also shown.

2. Experimental

2.1. Data collection and density modeling

Data sets of AHA and AFA were measured at beamline F1 of storage ring DORIS III at HASYLAB/DESY in Hamburg. F1 is equipped with a Kappa-axis diffractometer and a marCCD 165 area detector. A wavelength of 0.5600 (2) Å was adjusted and the temperature was maintained at 100 K using an open-flow nitrogen cooling device. More than 150 000 reflections each could be measured to a resolution of 1.16 Å^{-1}. For more details on the measurements and the crystallographic data see Table 2.

Integration, scaling and merging of data were carried out using the program *XDS* (Kabsch, 1988*a*,*b*, 1993). The structures were solved with the program *SHELXS* (Sheldrick, 2008). For AHA, the asymmetric unit consists of one molecule of the tripeptide, one molecule of 2-propanol and one molecule of water (see Fig. 2); for AFA, the asymmetric unit is composed of one molecule of the tripeptide and one molecule of *N*,*N*-dimethyl-formamide (see Fig. 2). Conventional spherical refinement was carried out using *SHELXL* (Sheldrick, 2008) to establish the starting positional and displacement parameters (anisotropic for non-H atoms, isotropic for H atoms) for the aspherical refinement steps. Anisotropic displacement parameters for H atoms can be obtained by a rigid-body approximation using the SHADE approach (Madsen, 2006). We did not make use of this advantage because we wanted to enable a direct comparison to the earlier published data sets by maintaining the same refinement strategy.

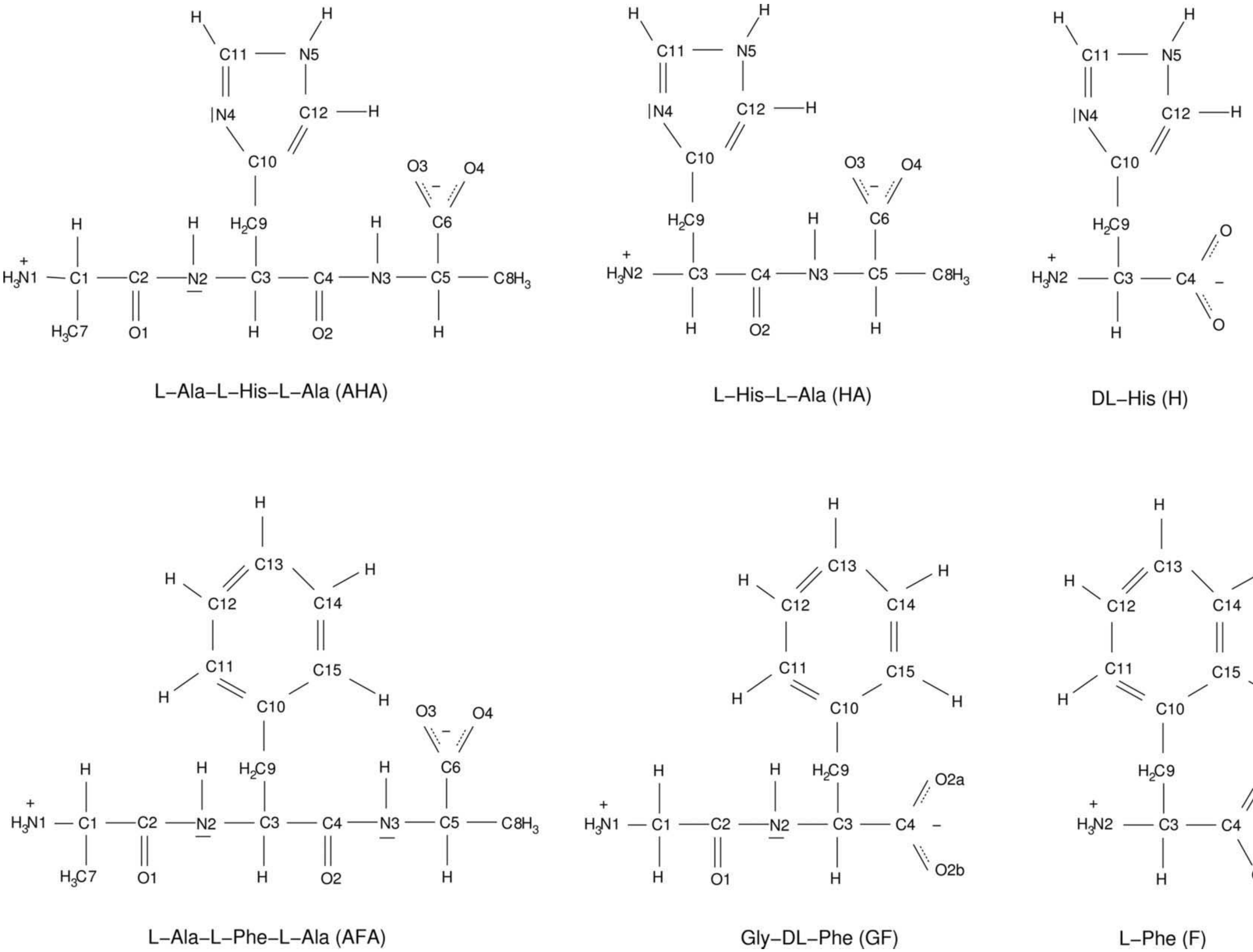
For aspherical refinement the Hansen–Coppens multipole formalism (Hansen & Coppens, 1978) as implemented in the program *XD2006* (Volkov *et al.*, 2006) was used. The usage of chemical site symmetries and constraints in the main chain is identical for all tripeptides under discussion. The density model can be summarized as follows: local threefold (3) symmetry was imposed on the ammonium N atom N1 and on the methyl C atoms C7 and C8 which were additionally chemically constrained to each other; local mirror (*m*) symmetry was imposed on the atoms N2, N3, C2 and C4 in the peptide bond region; all other main-chain atoms (C1, O1, C3, O2, C5, C6, O3, O4) were refined without local site symmetry. For the rest of the non-H atoms in AHA (C9, C10, C11, C12, N4 and N5), *m* symmetry was applied. For AFA, *m* symmetry was imposed on C9, but two mirror planes and one twofold axis (*mm*2) were imposed on the C atoms of the phenyl group. All non-H atoms were treated up to the hexadecapole level of expansion, while monopoles and bond-directed dipoles were introduced for all H atoms. The expansion–contraction parameter κ was refined independently for all non-constrained atoms, but for H atoms optimized values $\kappa = 1.13$ and $\kappa' = 1.29$ were used (Volkov *et al.*, 2001). Charge transfer was allowed between tripeptide and solvent molecules but an electro-

Table 3
Bond-topological properties of AHA compared with HA (Dittrich, 2002) and H (Coppens *et al.*, 1999) derived from experiment.

Labels according to Fig. 3.

	AHA		HA		H	
Bond labels	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})
N1−C1	1.56 (2)	−9.4 (1)	–	–	–	–
C1−C7	1.63 (2)	−9.8 (1)	–	–	–	–
C1−C2	1.65 (2)	−10.2 (1)	–	–	–	–
C2−O1	2.84 (4)	−36.6 (2)	–	–	–	–
C2−N2	2.43 (3)	−28.2 (2)	–	–	–	–
N2−C3	1.67 (2)	−7.8 (1)	1.81 (5)	−14.4 (2)	1.61 (1)	−7.3 (1)
C3−C9	1.52 (2)	−8.4 (1)	1.80 (4)	−12.2 (1)	1.56 (1)	−9.2 (1)
C9−C10	1.67 (2)	−8.9 (1)	1.81 (4)	−13.5 (1)	1.75 (2)	−11.3 (1)
C10−N4	2.10 (3)	−19.3 (1)	2.03 (6)	−12.6 (2)	2.16 (3)	−14.5 (1)
N4−C11	2.42 (3)	−25.4 (2)	2.56 (7)	−24.6 (3)	2.50 (3)	−22.7 (1)
C11−N5	2.14 (3)	−21.2 (2)	2.29 (6)	−16.1 (3)	2.27 (3)	−21.4 (1)
N5−C12	2.14 (3)	−17.9 (2)	2.26 (6)	−16.1 (2)	2.12 (3)	−16.9 (1)
C12−C10	2.13 (2)	−18.9 (1)	2.42 (5)	−23.7 (2)	2.23 (3)	−18.8 (1)
C3−C4	1.70 (2)	−11.5 (7)	1.72 (4)	−12.7 (1)	1.73 (1)	−13.2 (1)
C4−O2	2.66 (4)	−27.0 (3)	2.76 (6)	−24.7 (4)	–	–
C4−N3	2.30 (3)	−29.2 (2)	2.24 (6)	−21.9 (2)	–	–
N3−C5	1.72 (3)	−14.0 (1)	1.85 (5)	−13.2 (2)	–	–
C5−C8	1.55 (2)	−9.6 (1)	1.75 (5)	−15.4 (2)	–	–
C5−C6	1.62 (2)	−11.1 (7)	1.88 (4)	−15.3 (1)	–	–
C6−O3	2.58 (4)	−31.7 (2)	2.73 (6)	−29.2 (3)	–	–
C6−O4	2.62 (4)	−34.0 (3)	2.90 (6)	−36.8 (3)	–	–

neutrality constraint was applied for the asymmetric unit. All bonds involving H atoms were fixed to mean neutron diffraction values (Allen *et al.*, 1992). The figures of merit for data reduction and refinement are collected in Table 2. To analyze the electron-density distributions obtained, the program *XDPROP* of the *XD2006* program package (Volkov *et al.*, 2006) was used. Atomic integration led to maximum Lagrangian values of 0.008 for AHA and 0.009 for AFA. Electroneutrality of the asymmetric unit could be reproduced with an error of 0.07 e for both AHA and AFA. The atomic volumes add up to 99.1% of the crystallographic volume of the asymmetric unit for both AHA and AFA. Further discussions concentrate on the tripeptides;



L−Ala−L−His−L−Ala (AHA)

L−His−L−Ala (HA)

DL−His (H)

L−Ala−L−Phe−L−Ala (AFA)

Gly−DL−Phe (GF)

L−Phe (F)

Figure 3
Labeling schemes of AHA, HA, H (first row) and AFA, GF, F (second row).

the results for the co-crystallized solvent molecules, which behave as expected, are not considered further.

Table 4
Bond-topological properties of AFA compared with GF (Dittrich, 2002; Dittrich *et al.*, 2003) and F (Mebs *et al.*, 2006; first/second line refer to zwitterionic/cationic form in the asymmetric unit) derived from experiment.

Labels according to Fig. 3.

	AFA		GF		F	
Bond labels	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})
N1—C1	1.80 (2)	−11.1 (8)	1.82 (3)	−15.4 (1)	–	–
C1—C7	1.65 (2)	−10.8 (1)	–	–	–	–
C1—C2	1.74 (2)	−13.3 (1)	1.80 (2)	−14.4 (1)	–	–
C2—O1	2.80 (3)	−29.1 (2)	2.94 (3)	−37.5 (1)	–	–
C2—N2	2.45 (2)	−23.4 (1)	2.28 (3)	−23.2 (1)	–	–
N2—C3	1.80 (2)	−12.3 (1)	1.73 (2)	−11.9 (1)	1.83 (7) 1.72 (7)	−12.5 (3) −6.1 (2)
C3—C9	1.58 (2)	−9.5 (1)	1.62 (2)	−7.0 (1)	1.70 (6) 1.74 (5)	−10.4 (2) −10.0 (1)
C9—C10	1.73 (2)	−11.9 (1)	1.69 (2)	−11.7 (1)	1.84 (4) 1.84 (4)	−12.0 (1) −12.2 (1)
C10—C11	2.10 (1)	−16.3 (1)	2.18 (2)	−20.1 (1)	2.24 (4) 2.24 (4)	−18.2 (1) −18.2 (1)
C11—C12	2.09 (1)	−15.9 (1)	2.20 (1)	−20.2 (1)	2.21 (1) 2.21 (1)	−16.9 (1) −16.9 (1)
C12—C13	2.10 (1)	−16.4 (1)	2.21 (1)	−20.6 (1)	2.20 (1) 2.20 (1)	−16.7 (1) −16.7 (1)
C13—C14	2.10 (2)	−16.4 (1)	2.19 (1)	−20.0 (1)	2.20 (1) 2.19 (1)	−16.7 (1) −16.5 (1)
C14—C15	2.09 (1)	−16.0 (1)	2.20 (1)	−20.4 (1)	2.19 (2) 2.19 (2)	−16.6 (1) −16.6 (1)
C15—C10	2.10 (2)	−16.3 (1)	2.15 (2)	−18.3 (1)	2.17(4) 2.17 (4)	−16.9 (1) −17.0 (1)
C3—C4	1.73 (2)	−11.5 (1)	1.68 (2)	−12.7 (1)	1.67 (6) 1.62 (6)	−14.3 (2) −6.7 (2)
C4—O2	2.75 (3)	−24.4 (2)	–	–	–	–
C4—N3	2.44 (3)	−22.8 (1)	–	–	–	–
N3—C5	1.85 (2)	−13.0 (1)	–	–	–	–
C5—C8	1.61 (2)	−8.4 (1)	–	–	–	–
C5—C6	1.69 (2)	−10.3 (1)	–	–	–	–
C6—O3	2.61 (4)	−37.2 (3)	–	–	–	–
C6—O4	2.66 (3)	−31.0 (2)	–	–	–	–

2.2. Theoretical calculations

Tripeptides of the type A*X*A (zwitterionic form) were computed, with *X* being any of the 20 naturally encoded amino acids. Starting geometries were chosen according to ideal β-sheet structure. Subsequent geometry optimizations using the AMBER force field (Cornell *et al.*, 1995) were performed with the program *GAUSSIAN*98 (Frisch *et al.*, 1998). At the AMBER optimized geometries, determinations of the wavefunctions were carried out at the B3LYP/6–311++G(2d,2p) level of theory. Additionally, for the eight different molecules for which experimental coordinates after multipole modeling were available, determinations of the wavefunctions were carried out at the same level of theory at their experimental geometries. Topological analysis and integration of atomic properties were performed using the programs *MORPHY*98 (Popelier, 1996*a*,*b*) and *AIM*2000 (Biegler-König *et al.*, 2001).

Table 5
Atomic properties of AHA compared with HA (Dittrich, 2002) derived from experiment (V_{001} and Q_{001} refer to values cut at ϱ = 0.001 a.u.).

Labels according to Fig. 3.

	AHA		HA	
Atom labels	V_{001} (Å)	Q_{001} (e)	V_{001} (Å)	Q_{001} (e)
N1	14.3	−1.14	–	–
C1	7.3	0.14	–	–
C7	10.5	−0.05	–	–
C2	6.2	1.00	–	–
O1	15.9	−0.85	–	–
N2	13.4	−1.13	10.5	−0.67
C3	6.8	0.10	6.3	0.46
C9	7.9	0.10	7.6	0.28
C10	9.2	0.35	9.1	0.22
N4	14.3	−0.85	13.3	−0.62
C11	9.9	0.70	8.1	0.85
N5	14.5	−0.86	13.9	−0.89
C12	12.4	0.21	9.9	0.31
C4	6.3	1.33	6.4	1.15
O2	17.4	−0.90	16.9	−0.87
N3	12.9	−0.92	11.5	−0.47
C5	6.6	0.50	7.1	0.14
C8	10.9	−0.06	9.3	0.49
C6	5.8	1.47	6.3	1.20
O3	15.7	−0.86	15.3	−0.82
O4	15.3	−0.80	14.6	−0.65

3. Results and discussion

3.1. Experimental results for AHA and AFA, comparison with histidine (H) and phenylalanine (F) in dipeptides and as free amino acids

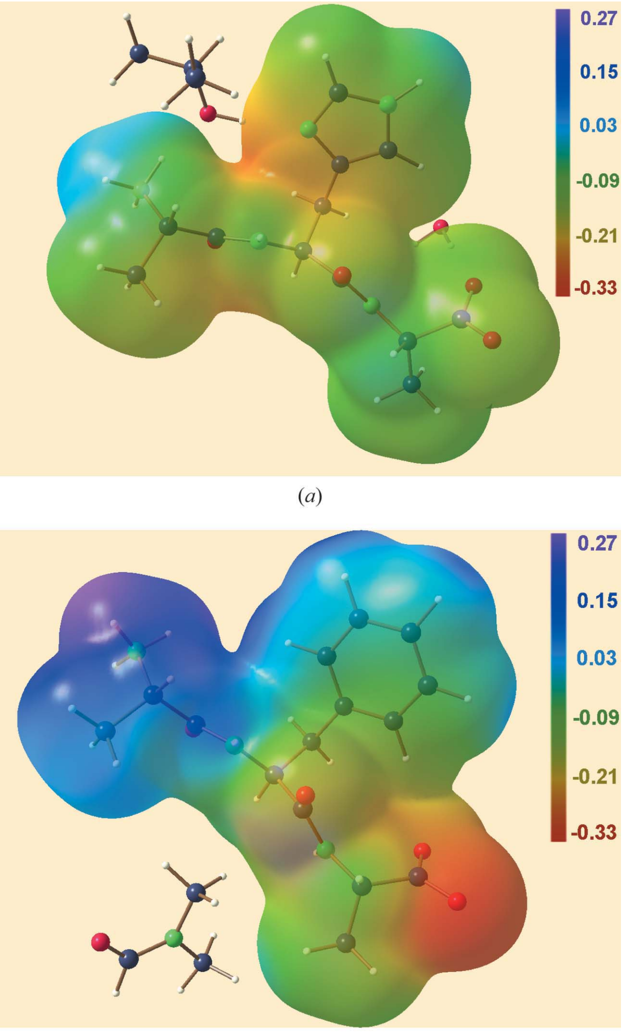
The bond-topological and atomic properties derived for the two tripeptides of this study (AHA and AFA) are listed in Tables 3, 4, 5 and 6. The results are compared with the corresponding properties of the histidinyl (H) fragment and the phenylalanyl (F) fragment in dipeptides and as free amino acids. AHA is compared with L-histidinyl-L-alanine (HA) (Dittrich, 2002) and DL-histidine (H) (Coppens *et al.*, 1999). Unfortunately, there is no experimental study on AH and, furthermore, for H atomic properties have not been published. AFA is compared with glycyl-DL-phenylalanine (GF) (Dittrich, 2002; Dittrich *et al.*, 2003) and L-phenylalanine (F) (Mebs *et al.*, 2006). There is no study on AF or FA, so GF was chosen because the difference is only one methyl group. For F, there is a zwitterionic and a cationic form in the asymmetric unit so that comparison with both forms is possible. The labeling schemes of all molecules that are compared with each other are given in Fig. 3.

Table 6
Atomic properties of AFA compared with GF (Dittrich, 2002; Dittrich *et al.*, 2003) and F (Mebs *et al.*, 2006; first/second line refer to zwitterionic/cationic form in the asymmetric unit) derived from experiment (V_{001} and Q_{001} refer to values cut at ϱ = 0.001 a.u.).

Labels according to Fig. 3.

	AFA		GF		F	
Atom labels	V_{001} (Å^3)	Q_{001} (e)	V_{001} (Å^3)	Q_{001} (e)	V_{001} (Å^3)	Q_{001} (e)
N1	15.0	−1.33	12.2	−0.94	–	–
C1	6.6	0.29	8.0	0.30	–	–
C7	9.9	0.16	–	–	–	–
C2	5.9	1.09	7.0	1.00	–	–
O1	15.3	−1.07	16.5	−1.01	–	–
N2	13.2	−1.10	12.8	−1.12	16.6	−1.59
					15.3	−1.56
C3	6.6	0.22	6.8	0.27	6.5	0.27
					6.6	0.01
C9	7.8	0.15	7.9	0.24	8.5	−0.17
					8.3	−0.17
C10	9.5	0.08	10.0	−0.18	9.6	−0.05
					9.5	−0.05
C11	10.9	0.08	11.8	0.01	10.9	0.00
					10.7	0.00
C12	11.1	0.08	11.6	0.02	10.6	0.01
					11.1	0.01
C13	11.1	0.07	11.7	0.03	11.1	0.01
					10.4	0.03
C14	11.4	0.07	12.0	0.02	10.4	0.01
					11.1	0.00
C15	11.0	0.08	11.3	0.03	11.1	−0.01
					10.7	−0.01
C4	5.8	1.09	5.8	1.32	5.4	1.20
					5.7	1.25
O2/O2*a*	16.4	−1.02	16.4	−1.02	17.0	−1.15
	–	–	–	–	16.6	−1.15
O2*b*	–	–	15.2	−0.95	14.3	−0.99
	–	–	–	–	16.7	−1.31
N3	12.3	−1.18	–	–	–	–
C5	6.9	0.24	–	–	–	–
C8	9.6	0.18	–	–	–	–
C6	6.1	1.29	–	–	–	–
O3	17.8	−1.07	–	–	–	–
O4	15.3	−1.02	–	–	–	–

For the following comparisons, the nearest-neighbor approximation (NNA) will be assumed (see §1). Two bonds are considered equal in the NNA if each atom of the first bond is equal in the NNA to its corresponding atom of the second bond. Therefore, not only the same bond in the three different molecule types (tripeptides, dipeptides, amino acids) is considered equal but there are also equal bonds within the same molecule, namely C1—C2 = C3—C4, C2—O1 = C4—O2, C2—N2 = C4—N3, N2—C3 = N3—C5, C6—O3 = C6—O4 in the main chain of the tripeptides and C10—C11 = C10—C15, C11—C12 = C12—C13 = C13—C14 = C14—C15 in the phenyl group of the F fragment. We note that the corresponding atoms were not subject to constraints in the multipole model. Thus, there are, for example, three values to be averaged for the bond type C_α—C_{pep}, namely C1—C2(tripeptide), C3—C4(tripeptide), C3—C4(dipeptide). All standard deviations from the averaging procedures are averaged again to obtain the transferability indices $\overline{\sigma}_{trans,exp}$ corresponding to the procedure outlined in Fig. 1. If this is done in the row AHA, HA, H (Table 3) it results in $\overline{\sigma}_{trans,exp}(\varrho_{bcp})$ = 0.10 e Å^{-3} and $\overline{\sigma}_{trans,exp}(\nabla^2\varrho_{bcp})$ = 2.8 e Å^{-5}. The corresponding quantities in the row AFA, GF, F(zwitterionic), F(cationic) (Table 4) are $\overline{\sigma}_{trans,exp}(\varrho_{bcp})$ = 0.05 e Å^{-3} and $\overline{\sigma}_{trans,exp}(\nabla^2\varrho_{bcp})$ = 2.2 e Å^{-5}. To judge whether these indices express successful transferability or not they have to be compared with the overall reproducibility indices $\overline{\sigma}^o_{rep,exp}$ [$\overline{\sigma}^o_{rep,exp}(\varrho_{bcp})$ = 0.07 e Å^{-3}, $\overline{\sigma}^o_{rep,exp}(\nabla^2\varrho_{bcp})$ = 3.3 e Å^{-5}; see last line in Table 1]. Thus, transferability is given in the F row but it is slightly reduced for the H row because the transferability index of the density at the b.c.p. (0.10 e Å^{-3}) slightly exceeds the corresponding experimental overall reproducibility index (0.07 e Å^{-3}). One possible explanation for this small difference between the F and the H row could be that the refinement strategy of the H molecule is different to that of HA and AHA as only octupole expansion was used as well as other κ' values for the non-H atoms, whereas the refinement strategies in the F row are identical. Two other possible explanations are:



(*a*)

(*b*)

Figure 4
MolIso representation (Hübschle & Luger, 2006) of the experimental electrostatic potential in e Å^{-1} mapped on an isosurface of the electron density at ϱ = 0.0067 e Å^{-3} (= 0.001 a.u.) of (*a*) AHA and (*b*) AFA.

Table 7
Hydrogen-bonding pattern of AHA and AFA (for labeling see Fig. 2).

Labels	H···*A* (Å)	*D*···*A* (Å)	*D*—H···*A* (°)	$\varrho_{\rm bcp}$ (e Å^{-3})	$\nabla^2\varrho_{\rm bcp}$ (e Å^{-5})	$V_{\rm tot}$ (Å^3)	$Q_{\rm tot}$ (e)	Symmetry codes
AHA								
N1—H1N*A*···O2	1.95	2.82	165.4	0.20 (2)	3.2 (1)	3.1	0.46	$2 - x, -\frac{1}{2} + y, 1 - z$
N1—H1N*B*···O5	1.94	2.79	158.7	0.24 (2)	3.3 (1)	3.2	0.42	$1 - x, -\frac{1}{2} + y, 1 - z$
N1—H1N*C*···O3	1.88	2.76	169.6	0.26 (2)	3.5 (1)	2.8	0.49	$x, y, 1 + z$
N2—H2N···O1	2.08	2.99	158.8	0.12 (1)	2.1 (1)	3.4	0.43	$2 - x, \frac{1}{2} + y, 1 - z$
N3—H3N···O4	2.05	2.89	165.4	0.16 (1)	2.7 (1)	3.2	0.41	$2 - x, -\frac{1}{2} + y, -z$
N5—H5N···O6	1.92	2.81	166.9	0.26 (2)	3.4 (1)	4.0	0.34	x, y, z
O6—H6O*A*···O4	1.98	2.79	165.2	0.18 (2)	3.1 (1)	2.3	0.57	$-1 + x, y, z$
O6—H6O*B*···O3	1.87	2.76	172.3	0.22 (2)	3.4 (1)	2.7	0.50	$1 - x, \frac{1}{2} + y, -z$
O5—H5O···N4	1.93	2.79	175.9	0.24 (2)	3.1 (1)	2.6	0.53	$1 - x, \frac{1}{2} + y, 1 - z$
C1—H1···O1	2.43	3.26	137.5	0.07 (1)	1.0 (1)	6.2	0.03	$2 - x, \frac{1}{2} + y, 1 - z$
C9—H9*B*···O4	2.46	3.29	141.9	0.07 (1)	0.9 (1)	6.8	−0.17	$2 - x, -\frac{1}{2} + y, -z$
C11—H11···O5	2.59	3.55	157.6	0.05 (1)	0.7 (1)	8.5	−0.03	x, y, z
AFA								
N1—H1N*A*···O2	1.85	2.74	176.0	0.25 (1)	4.1 (1)	2.6	0.52	$1 - x, -\frac{1}{2} + y, -z$
N1—H1N*B*···O5	1.91	2.80	174.1	0.24 (2)	3.6 (2)	2.2	0.54	$1 + x, y, z$
N1—H1N*C*···O4	1.93	2.78	159.5	0.23 (2)	3.2 (1)	2.5	0.54	$x, y, -1 + z$
N2—H2N···O1	2.25	2.98	152.9	0.10 (1)	1.8 (1)	3.0	0.48	$1 - x, \frac{1}{2} + y, -z$
N3—H3N···O4	2.06	2.89	168.0	0.14 (1)	2.8 (1)	3.0	0.47	$1 - x, -\frac{1}{2} + y, 1 - z$
C1—H1···O1	2.34	3.03	129.7	0.10 (1)	1.4 (1)	7.4	0.10	$1 - x, \frac{1}{2} + y, -z$
C12—H12···O5	2.54	3.51	163.3	0.04 (1)	0.5 (1)	9.7	0.01	$1 - x, \frac{1}{2} + y, -z$
C13—H13···O3	2.54	3.38	149.4	0.05 (1)	0.8 (1)	7.6	0.00	$2 - x, \frac{1}{2} + y, 1 - z$
C15—H15···O3	2.60	3.61	169.2	0.03 (1)	0.6 (1)	7.4	0.00	x, y, z
C18—H18···O1	2.43	3.16	129.0	0.07 (1)	1.0 (1)	7.3	−0.01	$-1 + x, y, z$

Table 8
Mean standard deviations of bond-topological and atomic properties in the nearest-neighbor approximation from different studies on transferability from experiment, units: $[\varrho_{\rm bcp}]$ = 1 e Å^{-3}, $[\nabla^2\varrho_{\rm bcp}]$ = 1 e Å^{-5}, $[V_{001}]$ = 1 Å^3 and $[Q_{001}]$ = 1 e.

Study	$\overline{\sigma}_{\rm trans,exp}(\varrho_{\rm bcp})$	$\overline{\sigma}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp})$	$\overline{\sigma}_{\rm trans,exp}(Q_{001})$	$\overline{\sigma}_{\rm trans,exp}(V_{001})$	Ref.
Amino acids	0.09	3.5	–	–	(*a*)
Dipeptides	0.10	3.0	0.06	0.4	(*b*)
Small peptides	0.09	2.6	–	–	(*c*)
A and H fragment in AHA, HA, H	0.10	2.8	0.17	0.8	(*d*)
A and F fragment in AFA, GF, F	0.05	2.2	0.09	0.7	(*d*)
Eight A*X*A molecules	0.08	2.7	0.11	0.7	(*d*)
Mean values: $\overline{\sigma}^{o}_{\rm trans,exp}$	0.09 (2)	2.8 (4)	0.11 (4)	0.7 (2)	(*e*)

References: (*a*) Mebs *et al.* (2006); (*b*) Dittrich (2002), Dittrich *et al.* (2003); (*c*) Pichon-Pesme *et al.* (2000); (*d*) this work; (*e*) all studies.

(i) the H data sets summarized in Table 3 are of lower quality than the F data sets in Table 4 and

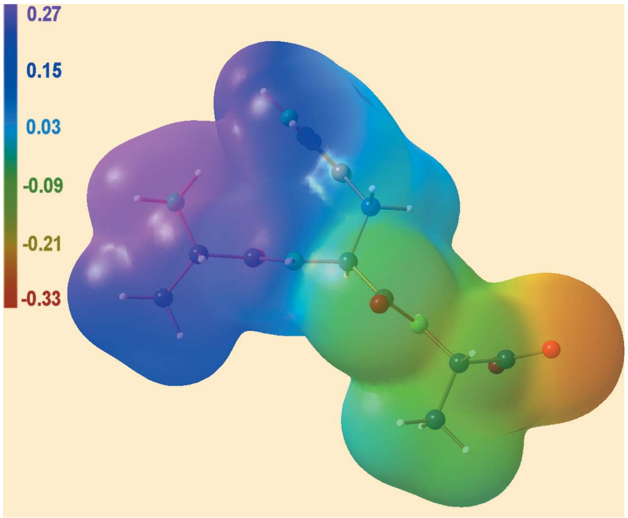
(ii) the existence of many more donor and acceptor groups for hydrogen bonding in the H fragment than in the F fragment (see Table 7) which is not accounted for in the NNA leads to greater differences for the H row in different crystalline environments.

The effect of hydrogen bonds on the bond-topological and atomic properties of the acceptor and donor groups in the main chain of the tripeptides will be discussed later.

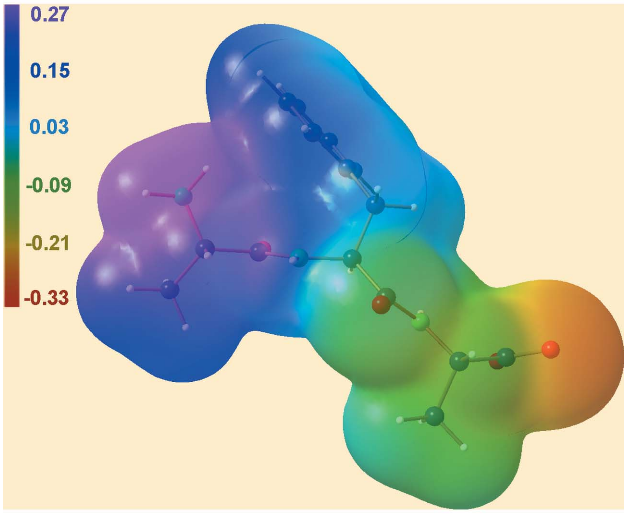
The transferability indices by means of bond parameters found here compare favorably with those found in the literature which were mentioned earlier (see Table 8). Amazingly, the transferability indices are very close together regardless of which systems in the field of peptides are scrutinized or if different systems from amino acids to tripeptides are compared as was the case in this study. We note that the transferability indices seem to be slightly reduced compared with the overall experimental reproducibility indices concerning the density at the b.c.p. but not concerning the Laplacian at the b.c.p.

For the atomic properties, the NNA is also assumed and as for the bonds, not only the same atom in the three different molecule types (tripeptides, dipeptides, amino acids) is considered equal but there are also equal atoms within the same molecule (C1 = C3 = C5, C2 = C4, O1 = O2, N2 = N3, O3 = O4 in the main chain of the tripeptides, N3 = N5 in HA, N2 = N3 = N5 in AHA and C11 = C12 = C13 = C14 = C15 in the F fragment). An averaging procedure corresponding to that used for the bonds was applied (see Fig. 1). The resulting transferability indices for AHA and HA (Table 5) are $\overline{\sigma}_{\rm trans,exp}$ (Q_{001}) = 0.17 e and $\overline{\sigma}_{\rm trans,exp}$ (V_{001}) = 0.8 Å^3. For the corresponding atomic properties in the row AFA, GF, F(zwitterionic) and F(cationic) (Table 6), $\overline{\sigma}_{\rm trans,exp}(Q_{001})$ = 0.09 e and $\overline{\sigma}_{\rm trans,exp}(V_{001})$ = 0.7 Å^3 were obtained. These values again have to be compared with the experimental overall reproducibility indices [$\overline{\sigma}^{o}_{\rm rep,exp}(Q_{001})$ = 0.12 e and $\overline{\sigma}^{o}_{\rm rep,exp}(V_{001})$ = 0.4 Å^3; last line in Table 1]. As for the bond-topological properties, the transferability indices found for the H and the F row are very close to the overall reproducibility indices. However, the degree of transferability is again slightly reduced. This may point to chemical differences of the H fragment and the F fragment in different molecules (amino acids without peptide bonds, dipeptides with one peptide bond, tripeptides with two peptide bonds) or differences of the crystal environment (hydrogen bonds, overall polarization, geometrical distortion, see below) that can be detected in the atomic parameters.

There might be different reasons for the fact that transferability with respect to reproducibility is somewhat reduced in most studies for most bond-topological and atomic parameters. The NNA might not always be sufficient, even for atoms that are not in delocalized systems. Next-nearest neighbors should be accounted for if the chemical environment of the nearest-neighbor changes dramatically, for example in the case of the formation of N_{pep} from N_{amm}. Moreover, the intermolecular interaction pattern and crystal environment might have a non-negligible influence on the bond and atomic properties due, for example, to changes in the bond lengths, angles and conformation (Koritsanszky *et al.*, 2002) so that new nearest neighbors and overall polarization effects would have to be accounted for. Fig. 4 shows the experimental electrostatic potentials of AHA and AFA on a molecular surface of ϱ = 0.001 a.u giving an impression of the molecules' overall polarization patterns on this surface. The scales are adjusted to each other. The two molecules have a very similar molecular shape, but the polarization is quite different. For AHA, the polarization is moderate: the ammonium group is most positively polarized, the most negatively polarized region is around the N4 atom of the histidinyl residue where the hydrogen bond with the 2-propanol molecule as the donor is present. AFA is much more strongly polarized: the ammonium group is positively polarized to a much higher degree than in AHA and the most negatively polarized region is the negatively charged carboxylate group. Strong differences in the polarization of the main chain as found here might be a reason for the reduction of the degree of transferability of atoms between the two different tripeptides that are equal in the NNA. These differences are not found for the theoretical electrostatic potentials which were obtained from AMBER optimized geometries in the gas phase, in which there are no intermolecular interactions. As Fig. 5 shows, the distribution of the electrostatic potentials on the molecular surfaces is nearly identical for AHA and AFA. Here the ammonium group is most strongly positive, the carboxylate group is most strongly negative and there is a gradient between these two termini. Hence, the comparison between Figs. 4 and 5 makes the strong effect of intermolecular interactions visible and the importance of the crystallographic experiment especially for biological compounds where the understanding of intermolecular interactions is of key importance.



(*a*)



(*b*)

Figure 5
MolIso representation (Hübschle & Luger, 2006) of the theoretical electrostatic potential (isolated molecule) in e Å^{-1} derived from optimized geometries mapped on an isosurface of the electron density at ϱ = 0.0067 e Å^{-3} (= 0.001 a.u.) of (*a*) AHA and (*b*) AFA.

Similar differences in the experimental electrostatic potentials between tripeptide, dipeptide and free amino acid of the H or the F row as found above for two different tripeptides might also explain the reduction of transferability in these rows. The hydrogen-bonding patterns between tripeptide, dipeptide and free amino acid are different too, mainly because there are other solvent molecules in the crystal packing. AHA crystallizes with 2-propanol and water, HA crystallizes as dihydrate and H crystallizes without solvent molecules. AFA crystallizes with *N*,*N*-dimethyl-formamide, GF without solvent molecules and F with formic acid.

The hydrogen-bonding patterns of AHA and AFA are summarized in Table 7. AHA exhibits nine classical donor—H···acceptor bonds and three C—H···acceptor bonds, whereas AFA exhibits only five classical donor—H···acceptor bonds but also five C—H···acceptor bonds. The AFA molecule can be divided into one half with donor functions and one half with acceptor functions which explains the strong polarization. This is not the case for AHA.

3.2. Comparison of experimental results for A*X*A tripeptides

Eight different tripeptide molecules of the type A*X*A from seven different measurements are compared here: two independent molecules in the asymmetric unit of L-alanyl-L-alanyl-L-alanine (AAA; Rödel *et al.*, 2006), two different modifications of L-alanine-L-tyrosyl-L-alanine (AYA; Cheçińska *et al.*, 2006), L-alanyl-glycyl-L-alanine (AGA; Förster *et al.*, 2006), L-alanyl-L-prolyl-L-alanine (APA; Kalinowski *et al.*, 2007) and finally AHA and AFA of this work. The transferability indices for bond-topological and atomic properties of these eight different molecules are $\overline{\sigma}_{\rm trans,exp}(\varrho_{\rm bcp}) = 0.08$ e Å^{-3}, $\overline{\sigma}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp}) = 2.7$ e Å^{-5}, $\overline{\sigma}_{\rm trans,exp}(Q) = 0.11$ e and $\overline{\sigma}_{\rm trans,exp}(V_{001}) = 0.7$ Å^3. Again, these quantities should be compared with the benchmarking overall reproducibility indices (Table 1): $\overline{\sigma}^{o}_{\rm rep,exp}(\varrho_{\rm bcp}) = 0.07$ e Å^{-3}, $\overline{\sigma}^{o}_{\rm rep,exp}(\nabla^2\varrho_{\rm bcp}) = 3.3$ e Å^{-5}, $\overline{\sigma}^{o}_{\rm rep,exp}(Q) = 0.12$ e and $\overline{\sigma}^{o}_{\rm rep,exp}(V_{001}) = 0.4$ Å^3. As found for the H and the F row and also several references (*cf.* Table 8), the degree of transferability is slightly reduced because the density at the b.c.p. and the volume exceed the margins given by reproducibility. However, the deviations are small (0.01 e Å^{-3} for $\varrho_{\rm bcp}$ and 0.3 Å^3 for V_{001}) and exactly in the range always found in transferability studies. Table 8 summarizes all the mentioned experimental transferability studies from the literature and the three different experimental transferability studies from this work to obtain a reliable estimate for the general degree of transferability in electron-density studies in the peptide field as overall transferability indices: $\overline{\sigma}^{o}_{\rm trans,exp}(\varrho_{\rm bcp}) = 0.09\,(2)$ e Å^{-3}, $\overline{\sigma}^{o}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp}) = 2.8\,(4)$ e Å^{-5}, $\overline{\sigma}^{o}_{\rm trans,exp}(Q_{001}) = 0.11\,(4)$ e and $\overline{\sigma}^{o}_{\rm trans,exp}(V_{001}) = 0.7\,(2)$ Å^3. The results confirm that transferability is generally given in a range that is comparable to reproducibility.

To find out why the degree of transferability slightly exceeds reproducibility in some cases, the independent bonds and atoms of the main chain of the tripeptides are examined in detail. Fig. 6 shows the mean values and the standard deviations of all bond-topological parameters for the independent main-chain bonds; Fig. 7 shows the same for the atomic parameters. Bonds shown in the same color are equal in the nearest-neighbor approximation (NNA). It is striking that transferability is given for all bonds ($\rm C_{pep}{-}N_{pep}$, $\rm C_{\alpha}{-}C_{pep}$, $\rm N_{pep}{-}C_{\alpha}$, $\rm C_{\alpha}{-}C_{carb}$) and atoms ($\rm C_{pep}$, $\rm C_{\alpha}$, $\rm N_{pep}$, $\rm C_{carb}$) in the rigid inner main chain because the standard deviations of all parameters are smaller or only negligibly larger than the margins from reproducibility. On the other hand, transferability is not given for all bonds ($\rm N_{amm}{-}C_{\alpha}$, $\rm C_{pep}{-}O_{pep}$, $\rm C_{carb}{-}O_{carb}$) and atoms ($\rm N_{amm}$, $\rm O_{pep}$, $\rm O_{carb}$) in the termini containing heteroatoms and the peptide O atoms, because the densities at the b.c.p. and the volumes again exceed the margins given by reproducibility. The reasons are obvious. Firstly, these end groups are less rigid and their parameters might therefore be less accurately determined. Secondly, the charged termini are polarized strongly (compare Fig. 4) and are subjected to the strongest electrostatic interactions in the molecule that change the bond-topological and atomic parameters. Thirdly, these groups are donors and acceptors of classical hydrogen bonds that are not accounted for at averaging in the NNA. These results show that the concept of transferability is valid for peptide chains and deviations can be explained by chemical effects. In proteins, for which the data bank applications have mainly been developed, the percentage of charged termini in contrast to rigid main-chain atoms or bonds is reduced drastically compared with tripeptides. Therefore, the systematic error introduced by multipole

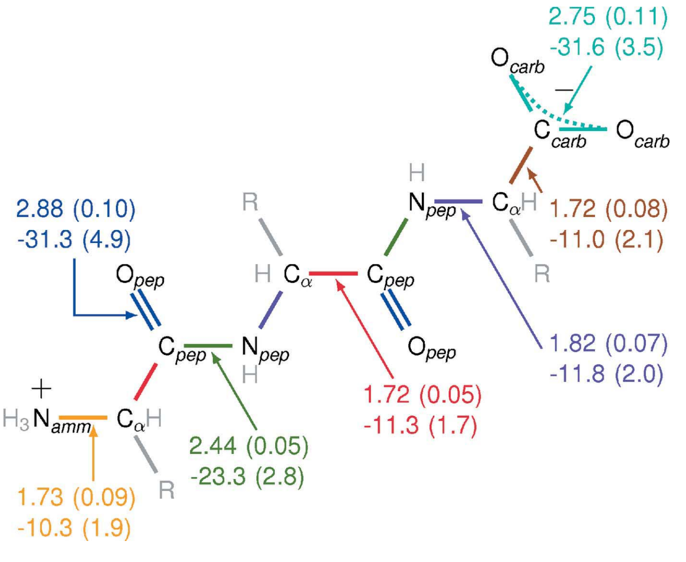


Figure 6
Mean values and standard deviations of bond-topological properties from experiment. First line: electron density $\varrho_{\rm bcp}$ in e Å^{-3}; second line: Laplacian $\nabla^2\varrho_{\rm bcp}$ in e Å^{-5}. The same color indicates equality in the nearest-neighbor approximation. Values from eight tripeptides AHA (this study), AFA (this study), two independent molecules in the asymmetric unit of AAA (Rödel *et al.*, 2006), two modifications of AYA (Cheçińska *et al.*, 2006), APA (Kalinowski *et al.*, 2007) and AGA (Förster *et al.*, 2006).

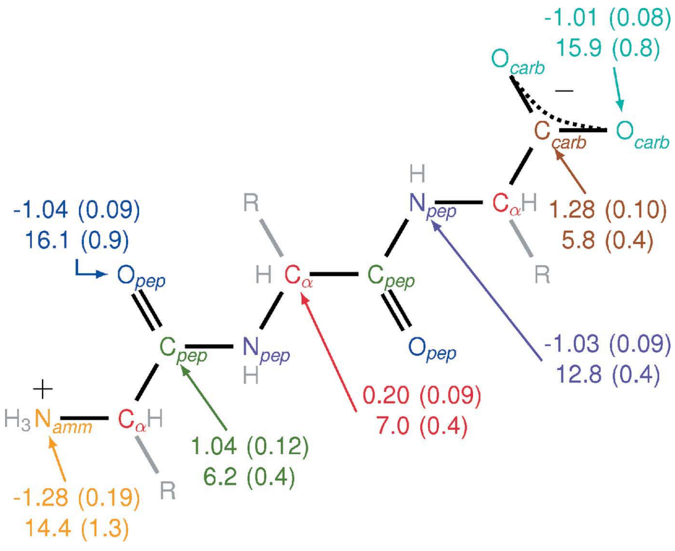


Figure 7
Mean values and standard deviations of atomic properties from experiment. First line: atomic charge Q_{001} in e, second line: atomic volume V_{001} in Å^3. The same color indicates equality in the nearest-neighbor approximation. Values from the eight tripeptides as listed in the legend of Fig. 6.

population transfer is minimized. For peptide O atoms, for which the percentage in contrast to rigid main-chain atoms or bonds is not reduced, the NNA should be extended towards neighbor approaches accounting for H atoms of hydrogen bonds as new nearest neighbors.

3.3. Comparison of theoretical results from optimized geometries

Bond-topological and atomic properties derived from theoretical calculations with the same method and basis set should give the opportunity to concentrate on transferability because we assume that no methodological error can occur this way. This is why we have examined bond-topological and atomic parameters for all 20 different tripeptides of the type A*X*A. Gas-phase optimizations with the AMBER force field starting from ideal β-sheet geometry were chosen because full *ab initio* geometry optimizations of zwitterions are challenging and can only be obtained satisfactorily by joint quantum mechanical/molecular mechanics calculations, *e.g.* with the ONIOM approach (Vreven *et al.*, 2003). The differences in the optimized geometries of the main chains between the different tripeptide molecules after AMBER optimization are negligibly small so that the influence of different bond lengths and angles can be discarded.

Theoretical transferability indices for bond-topological and atomic properties of these 20 different molecules are $\overline{\sigma}_{\rm trans,theo}(\varrho_{\rm bcp})$ = 0.005 e Å^{-3}, $\overline{\sigma}_{\rm trans,theo}(\nabla^2\varrho_{\rm bcp})$ = 0.14 e Å^{-5}, $\overline{\sigma}_{\rm trans,theo}(Q_{001})$ = 0.019 e and $\overline{\sigma}_{\rm trans,theo}(V_{001})$ = 0.29 Å^3. As expected, the theoretical transferability indices are much smaller than the experimental ones, but they are not zero. Thus, transferability is also not perfect in the gas phase without any intermolecular interactions and with identical geometries in the main chain. The differences must be attributed to chemical differences in the variable amino acid *X* and to chemical differences of the bonds or atoms in the same molecule that are considered equal but are next to a positive or a negative charge, respectively. These two effects influence the bond-topological and atomic parameters of the main-chain bonds or atoms. Therefore, the standard deviations shown above are the error one introduces by neglecting atoms further away from the atom or bond of interest in the nearest or next-nearest-neighbor approximations of the Invariom data bank approach.

Figs. 8 and 9 show mean values and standard deviations of all bond-topological and atomic parameters for the independent main-chain bonds or atoms to evaluate whether a systematic effect is again striking: bond-topological and atomic properties show no features, the deviations are uniformly distributed. Therefore, no special chemical effects of termini or rigid inner bonds or atoms can be found. The variation of the middle amino acid leads to small changes of the bond-topological and atomic properties in each bond or atom of the main chain to the same extent so that perfect transferability is not given. However, it can also be clearly seen that the general concept of transferability in the NNA is properly fulfilled because the deviations are very small.

For averaging, the atom N2 and bonds incorporating N2 in the molecule APA are left out because N2 is in a different chemical environment and cannot be considered equal to N2 of other tripeptides in the NNA. This chemical effect can clearly be seen in the bond-topological and atomic properties: $\varrho({\rm bcp}_{\rm C2-N2})$ = 2.238 e Å^{-3}, $\nabla^2\varrho({\rm bcp}_{\rm C2-N2})$ = −25.41 e Å^{-5}, $Q_{001}({\rm N2})$ = −0.985 e and $V_{001}({\rm N2})$ = 10.96 Å^3. These values differ from the mean values of $\rm C_{pep}{-}N_{pep}$ and $\rm N_{pep}$ most

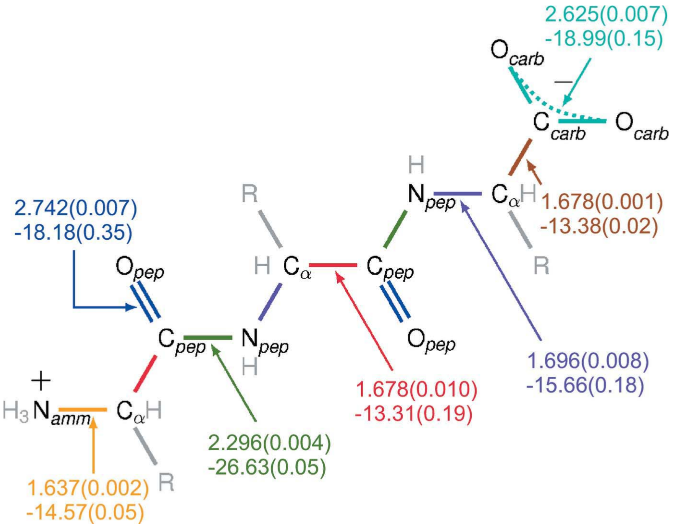


Figure 8

Mean values and standard deviations of bond-topological properties from theory at optimized geometries. First line: electron density $\varrho_{\rm bcp}$ in e Å^{-3}; second line: Laplacian $\nabla^2\varrho_{\rm bcp}$ in e Å^{-5}. The same color indicates equality in the nearest-neighbor approximation. Values from 20 tripeptides of the type A*X*A derived from AMBER (Cornell *et al.*, 1995) optimization and wavefunction calculation at the B3LYP/6–311++G(2d,2p) level of theory.

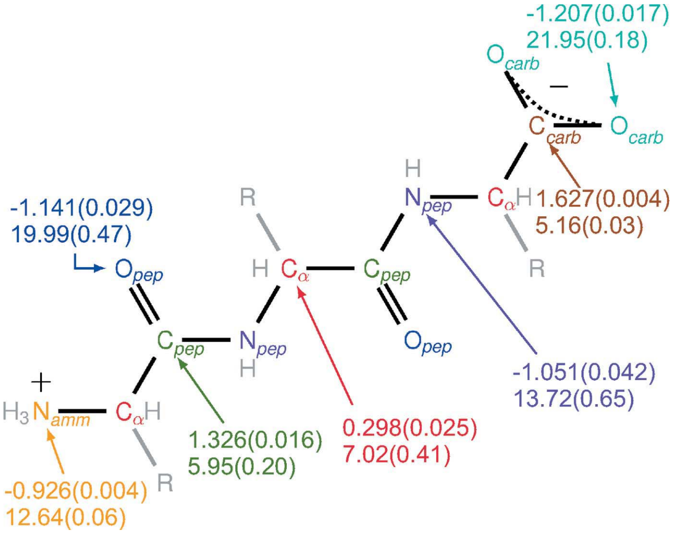


Figure 9

Mean values and standard deviations of atomic properties from theory at optimized geometries. First line: atomic charge Q_{001} in e; second line: atomic volume V_{001} in Å^3. The same color indicates equality in the nearest-neighbor approximation. Values from 20 tripeptides of the type A*X*A derived from AMBER (Cornell *et al.*, 1995) optimization and wavefunction calculation at the B3LYP/6–311++G(2d,2p) level of theory.

strongly for the A*X*A row. Therefore, it can be concluded that upon ring formation in proline, N2 becomes less negatively charged and its volume decreases. Furthermore, the C2—N2 bond loses electron density and the Laplacian becomes less negative.

3.4. Comparison between experimental and theoretical results

In the preceding chapters, theoretical and experimental transferability indices have been discussed. The values are much smaller for theory and give hints of chemical effects. Now, the differences of the mean values found for the independent bonds and atoms of the main chain between theory at optimized geometries and experiment are to be discussed. For the most polar bonds, *i.e.* the C—O bonds, the deviations between experiment and theory are most pronounced. The density and the negative Laplacian values are significantly larger in experiment. This effect is known and is due to the insufficient flexibility of the radial functions used in the multipolar fitting procedure (Volkov *et al.*, 2000; Henn *et al.*, 2004). Without C—O bonds, the mean differences are $\Delta(\overline{\varrho_{\rm bcp}}) = 0.09$ e Å^{-3}, $\Delta(\overline{\nabla^2\varrho_{\rm bcp}}) = 3.2$ e Å^{-5}, $\Delta(\overline{Q_{001}}) = 0.20$ e and $\Delta(\overline{V_{001}}) = 1.9$ Å^3. These deviations are in the range commonly found in comparisons between experimental topological properties and those derived from gas-phase calculations (Grabowsky *et al.*, 2007).

The transferability indices from theoretical calculations at experimental geometries are $\overline{\sigma}_{\rm trans,theoexp}(\varrho_{\rm bcp}) = 0.03$ e Å^{-3}, $\overline{\sigma}_{\rm trans,theoexp}(\nabla^2\varrho_{\rm bcp}) = 0.7$ e Å^{-5}, $\overline{\sigma}_{\rm trans,theoexp}(Q_{001}) = 0.05$ e and $\overline{\sigma}_{\rm trans,theoexp}(V_{001}) = 0.6$ Å^3. As expected, these deviations are in between those of the experimental case where the deviations are larger and the theoretical case with optimized geometries where the deviations are smaller. This can be seen in detail in Figs. 10 and 11.

The differences of the transferability indices between theory at optimized geometries and theory at experimental geometries are $\Delta\overline{\sigma}_{\rm trans}(\varrho_{\rm bcp}) = 0.025$ e Å^{-3}, $\Delta\overline{\sigma}_{\rm trans}(\nabla^2\varrho_{\rm bcp}) = 0.56$ e Å^{-5}, $\Delta\overline{\sigma}_{\rm trans}(Q_{001}) = 0.031$ e and $\Delta\overline{\sigma}_{\rm trans}(V_{001}) = 0.32$ Å^3. These differences are a measure of the impact of geometry changes on the bond-topological and atomic properties. As these values are quite large, it can be shown that geometry changes are a very important effect to change the bond-topological and atomic parameters.

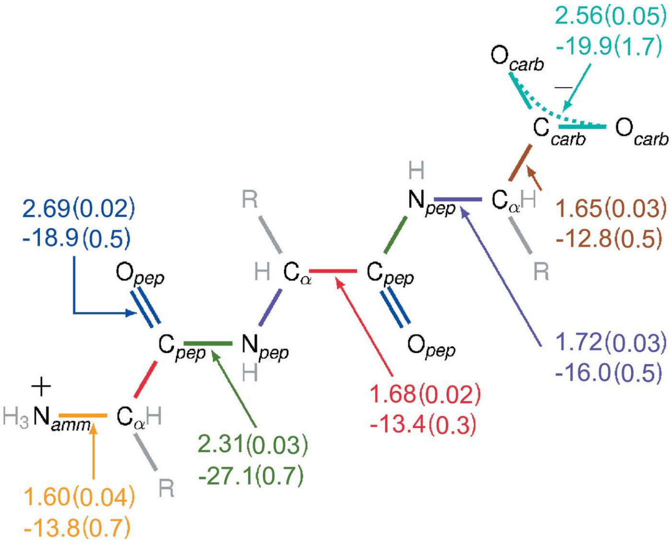


Figure 10
Mean values and standard deviations of bond-topological properties from theory at experimental geometry. First line: electron density $\varrho_{\rm bcp}$ in e Å^{-3}; second line: Laplacian $\nabla^2\varrho_{\rm bcp}$ in e Å^{-5}. The same color indicates equality in the nearest-neighbor approximation. Values from eight tripeptides AHA (this study), AFA (this study), two independent molecules in the asymmetric unit of AAA (Rödel *et al.*, 2006), two modifications of AYA (Cheçińska *et al.*, 2006), APA (Kalinowski *et al.*, 2007) and AGA (Förster *et al.*, 2006) derived from wavefunction calculations at the B3LYP/6–311++G(2d,2p) level of theory at experimental geometries.

4. Conclusion

High-resolution X-ray diffraction data sets of the tripeptides AHA and AFA at 100 K were measured at the synchrotron beamline F1 of HASYLAB/DESY. Electron-density modeling was carried out in the same way as was done for the other five existing data sets in the row A*X*A. Theoretical calculations on the experimental geometries of the tripeptides of these seven data sets and on AMBER optimized geometries of all 20 different tripeptides A*X*A were performed at the B3LYP/6–311++G(2d,2p) level of theory. For experiment and theory, bond-topological and atomic properties were derived to study the transferability of atomic fragments in the nearest-neighbor approximation (NNA). To have a basis for transferability considerations, studies on reproducibility of small molecules that were available in the literature were analyzed to give an

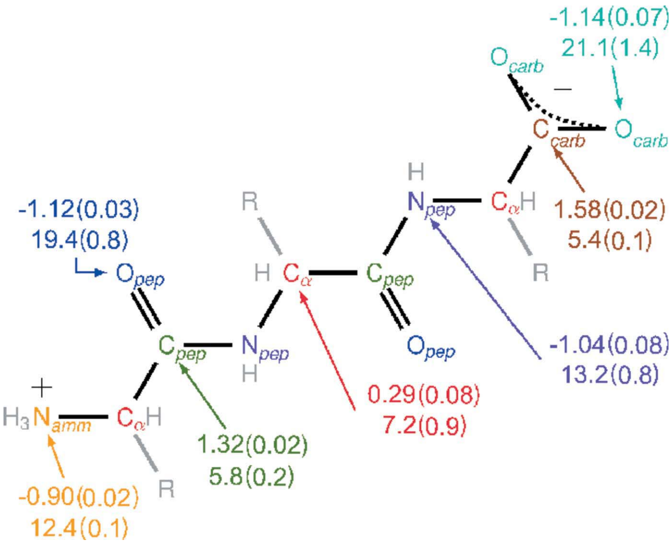


Figure 11
Mean values and standard deviations of atomic properties from theory at experimental geometry. First line: atomic charge Q_{001} in e; second line: atomic volume V_{001} in Å^3. The same color indicates equality in the nearest-neighbor approximation. Values obtained as detailed in the legend of Fig. 10.

estimate of the methodological or experimental error of X-ray diffraction experiment and multipole modeling. For this purpose, overall reproducibility indices were introduced which were found to be $\overline{\sigma}^{o}_{\rm rep,exp}(\varrho_{\rm bcp})$ = 0.07 e Å^{-3}, $\overline{\sigma}^{o}_{\rm rep,exp}(\nabla^2\varrho_{\rm bcp})$ = 3.3 e Å^{-5}, $\overline{\sigma}^{o}_{\rm rep,exp}(Q_{001})$ = 0.12 e and $\overline{\sigma}^{o}_{\rm rep,exp}(V_{001})$ = 0.4 Å^3 for bond-topological and atomic properties (last line in Table 1) and which were used as benchmarks.

From experimental studies in the peptide field, it was found that the degree of transferability can be very reliably calculated to give overall transferability indices: $\overline{\sigma}^{o}_{\rm trans,exp}(\varrho_{\rm bcp})$ = 0.09 (2) e Å^{-3}, $\overline{\sigma}^{o}_{\rm trans,exp}(\nabla^2\varrho_{\rm bcp})$ = 2.8 (4) e Å^{-5}, $\overline{\sigma}^{o}_{\rm trans,exp}(Q_{001})$ = 0.11 (4) e and $\overline{\sigma}^{o}_{\rm trans,exp}(V_{001})$ = 0.7 (2) Å^3 (last line in Table 8). These values can therefore generally be used as a good estimate for the degree of transferability in experimental electron-density studies. They are very close to the reproducibility margins and show that transferability is given in very narrow limits. However, as the transferability indices of the densities at the bond-critical points and the atomic volumes exceed the margin of reproducibility, chemical effects were looked for that lower the degree of transferability with respect to reproducibility. Crystal packing effects are responsible for these differences mainly in terms of geometrical deformations, intermolecular interactions like hydrogen bonds and different polarization patterns. An increase in the degree of transferability might therefore be reached by introducing next-nearest neighbors to account for polarization effects and further nearest neighbors to account for hydrogen bonds.

To separate different effects that influence the degree of transferability, theoretical gas-phase calculations were performed. At optimized geometries in the gas phase, no crystal packing effects occur and also geometrical distortions by packing forces are not present. Even so, transferability is not perfect as the standard deviations are not negligible but much smaller than in the experimental case. As long as it is assumed that no methodological errors are introduced when the method and basis set are not varied, these differences can therefore only be attributed to the chemical differences in the vicinity of the positively or negatively charged termini and in the middle amino acids *X*. Thus, the theoretical transferability indices [$\overline{\sigma}_{\rm trans,theo}(\varrho_{\rm bcp})$ = 0.005 e Å^{-3}, $\overline{\sigma}_{\rm trans,theo}(\nabla^2\varrho_{\rm bcp})$ = 0.14 e Å^{-5}, $\overline{\sigma}_{\rm trans,theo}(Q_{001})$ = 0.019 e and $\overline{\sigma}_{\rm trans,theo}(V_{001})$ = 0.29 Å^3] show how large the errors are that are introduced when influences of atoms further apart are neglected in the NNA. Theoretical calculations at experimental geometries gave the possibility to quantify the impact of geometrical differences on bond-topological and atomic properties.

It was shown in this study that transferability in the peptide case is valid in very narrow limits and is a reliable basis for data bank applications. However, it could also be shown that chemical effects that lower the degree of transferability are present in the crystal but can be accounted for in the individual case. This study was based on bond-topological and atomic properties as these values are the most commonly used ones to describe molecules in electron-density studies. However, there are other approaches that describe atomic transferability in a less practical but more general way by integrating isosurfaces of molecular properties (Popelier *et al.*, 2004). It is therefore planned in the future to repeat this study on tripeptides with other measures of transferability and other ways of determining the electron density (Jayatilaka & Dittrich, 2008) to obtain a better overview on the nature of transferability.

The authors thank the Deutsche Forschungsgemeinschaft (DFG) for financial support within the special priority program SPP1178, grant Lu222/29–2. Furthermore, Dr Birger Dittrich is acknowledged for his help with the dipeptide data sets.

References

Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1992). *International Tables for X-ray Crystallography*, Vol. C, ch. 9.5, pp. 685–706. Amsterdam: Kluwer Academic Publishers.

Arnold, W. D., Sanders, L. K., McMahon, M. T., Volkov, A. V., Wu, G., Coppens, P., Wilson, S. R., Godbout, N. & Oldfield, E. (2000). *J. Am. Chem. Soc.* **122**, 4708–4717.

Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory*,1st ed. No. 22 in the International Series of Monographs on Chemistry. Oxford: Clarendon Press.

Benabicha, F., Pichon-Pesme, V., Jelsch, C., Lecomte, C. & Khmou, A. (2000). *Acta Cryst.* B**56**, 155–165.

Biegler-König, F., Schönbohm, J. & Bayles, D. (2001). *J. Comput. Chem.* **22**, 545–559.

Burnett, M. N. & Johnson, C. K. (1996). *ORTEP*III. Report ORNL-6895. Oak Ridge National Laboratory, Tennessee, USA.

Chęcińska, L., Mebs, S., Hübschle, C. B., Förster, D., Morgenroth, W. & Luger, P. (2006). *Org. Biomol. Chem.* **4**, 3242–3251.

Coppens, P. (2005*a*). *Angew. Chem. Int. Ed.* **44**, 6810–6811.

Coppens, P. (2005*b*). *Angew. Chem.* **117**, 6970–6972.

Coppens, P., Abramov, Y., Carducci, M., Korjov, B., Novozhilova, I., Alhambra, C. & Pressprich, M. R. (1999). *J. Am. Chem. Soc.* **121**, 2585–2593.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R. Jr, Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). *J. Am. Chem. Soc.* **117**, 5179–5197.

Dittrich, B. (2002). PhD thesis. Freie Universität Berlin, Berlin.

Dittrich, B., Flaig, R., Koritsánszky, T., Krane, H.-G., Morgenroth, W. & Luger, P. (2000). *Chem. Eur. J.* **6**, 2582–2589.

Dittrich, B., Hübschle, C. B., Luger, P. & Spackman, M. A. (2006). *Acta Cryst.* D**62**, 1325–1335.

Dittrich, B., Koritsánszky, T., Grosche, M., Scherer, W., Flaig, R., Wagner, A., Krane, H. G., Kessler, H., Riemer, C., Schreurs, A. M. M. & Luger, P. (2002). *Acta Cryst.* B**58**, 721–727.

Dittrich, B., Koritsánszky, T. & Luger, P. (2004*a*). *Angew. Chem.* **43**, 2773–2776.

Dittrich, B., Koritsánszky, T. & Luger, P. (2004*b*). *Angew. Chem. Int. Ed.* **43**, 2718–2721.

Dittrich, B., Scheins, S., Paulmann, C. & Luger, P. (2003). *J. Phys. Chem. A*, **107**, 7471–7474.

Dominiak, P. M., Volkov, A., Li, X., Messerschmidt, M. & Coppens, P. (2007). *J. Chem. Theory Comput.* **3**, 232–247.

Flaig, R., Koritsánszky, T., Dittrich, B., Wagner, A. & Luger, P. (2002). *J. Am. Chem. Soc.* **124**, 3407–3417.

Flaig, R., Koritsánszky, T., Janczak, J., Krane, H.-G., Morgenroth, W. & Luger, P. (1999*a*). *Angew. Chem. Int. Ed.* **38**, 1397–1400.

Flaig, R., Koritsánszky, T., Janczak, J., Krane, H.-G., Morgenroth, W. & Luger, P. (1999*b*). *Angew. Chem.* **111**, 1494–1497.

Förster, D., Wagner, A., Hübschle, C. B., Paulmann, C. & Luger, P. (2006). *Z. Naturforsch. B*, **62**, 696–704.

Frisch, M. J. *et al.* (1998). *Gaussian*98. Revision A.11.3. Gaussian, Inc., Pittsburgh PA.

Grabowsky, S., Pfeuffer, T., Checińska, L., Weber, M., Morgenroth, W., Luger, P. & Schirmeister, T. (2007). *Eur. J. Org. Chem.* **17**, 2759–2768.

Grabowsky, S., Pfeuffer, T., Morgenroth, W., Paulmann, C., Schirmeister, T. & Luger, P. (2008). *Org. Biomol. Chem.* **6**, 2295–2307.

Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.

Henn, J., Ilge, D., Leusser, D., Stalke, D. & Engels, B. (2004). *J. Phys. Chem. A*, **108**, 9442–9452.

Hübschle, C. B. & Luger, P. (2006). *J. Appl. Cryst.* **39**, 901–904.

Hübschle, C. B., Luger, P. & Dittrich, B. (2007). *J. Appl. Cryst.* **40**, 623–627.

Jayatilaka, D. & Dittrich, B. (2008). *Acta Cryst.* A**64**, 383–393.

Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.

Kabsch, W. (1988*a*). *J. Appl. Cryst.* **21**, 67–72.

Kabsch, W. (1988*b*). *J. Appl. Cryst.* **21**, 916–924.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Kalinowski, R., Dittrich, B., Hübschle, C. B., Paulmann, C. & Luger, P. (2007). *Acta Cryst.* B**63**, 753–767.

Koritsanszky, T., Volkov, A. & Coppens, P. (2002). *Acta Cryst.* A**58**, 464–472.

Madsen, A. Ø. (2006). *J. Appl. Cryst.* **39**, 757–758.

Matta, C. F. & Bader, R. F. W. (2000). *Proteins*, **40**, 310–329.

Matta, C. F. & Bader, R. F. W. (2002). *Proteins*, **48**, 519–538.

Matta, C. F. & Bader, R. F. W. (2003). *Proteins*, **52**, 360–399.

McNaught, A. D. & Wilkinson, A. (1997). IUPAC Compendium of Chemical Terminology, 2nd ed. Oxford: Blackwell Science.

Mebs, S., Messerschmidt, M. & Luger, P. (2006). *Z. Kristallogr.* **221**, 656–664.

Messerschmidt, M., Scheins, S. & Luger, P. (2005). *Acta Cryst.* B**61**, 115–121.

Pichon-Pesme, V., Lachekar, H., Souhassou, M. & Lecomte, C. (2000). *Acta Cryst.* B**56**, 728–737.

Pichon-Pesme, V. & Lecomte, C. (1998). *Acta Cryst.* B**54**, 485–493.

Pichon-Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.* **99**, 6242–6250.

Popelier, P. L. A. (1996*a*). *Mol. Phys.* **87**, 1169–1187.

Popelier, P. L. A. (1996*b*). *Comput. Phys. Commun.* **93**, 212–240.

Popelier, P. L. A., Devereux, M. & Rafat, M. (2004). *Acta Cryst.* A**60**, 427–433.

Rödel, E., Messerschmidt, M., Dittrich, B. & Luger, P. (2006). *Org. Biomol. Chem.* **4**, 475–481.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Volkov, A., Abramov, Y. A. & Coppens, P. (2001). *Acta Cryst.* A**57**, 272–282.

Volkov, A., Abramov, Y., Coppens, P. & Gatti, C. (2000). *Acta Cryst.* A**56**, 332–339.

Volkov, A., Macchi, P., Farrugia, L. J., Gatti, C., Mallinson, P. R., Richter, T. & Koritsánszky, T. S. (2006). *XD2006*, Rev. 5.34. University of New York at Buffalo, NY, USA.

Volkov, A., Messerschmidt, M. & Coppens, P. (2007). *Acta Cryst.* D**63**, 160–170.

Vreven, T., Morokuma, K., Farkas, Ö., Schlegel, H. B. & Frisch, M. J. (2003). *J. Comput. Chem.* **24**, 760–769.