ELSEVIER

Chemical Physics 263 (2001) 203–219

Chemical Physics

www.elsevier.nl/locate/chemphys

Energetics and structure of glycine and alanine based model peptides: Approximate SCC-DFTB, AM1 and PM3 methods in comparison with DFT, HF and MP2 calculations

Marcus Elstner [a,b], Karl J. Jalkanen [b], Michaela Knapp-Mohammady [b], Thomas Frauenheim [a], Sándor Suhai [b,*]

[a] *Theoretische Physik, Universität Paderborn, D-33098 Paderborn, Germany*
[b] *Molekulare Biophysik, Deutsches Krebsforschungszentrum, D-69120 Heidelberg, Germany*

Received 18 July 2000; in final form 2 November 2000

Abstract

We calculate relative energies and geometries of important secondary structural elements for small glycine and alanine based polypeptides containing up to eight residues. We compare the performance of the approximate methods AM1, PM3 and self-consistent charge, density-functional tight-binding (SCC-DFTB) to density-functional theory (DFT), Hartree–Fock (HF) and MP2. The SCC-DFTB is able to reproduce structures and relative energies of various peptide models reliably compared to DFT results. The AM1 and PM3 methods show deficiencies in describing important secondary structure elements like extended, helical or turn structures. The discrepancies between different ab initio (HF, MP2) and DFT (B3LYP) methods for medium sized basis sets (6-31G*) also show the need for higher level calculations, since systematic errors found for small molecules may add up when investigating longer polypeptides.

1. Introduction

Semi-empirical methods like the AM1 [1] and PM3 [2] models are widely used for electronic structure calculations, since they are about two to three orders of magnitude faster than ab initio methods. The implementation of these theoretical models in linear scaling algorithms allows large-scale electronic structure calculations for biomolecular systems up to several thousand atoms [3–6].

Recently we developed an approximate quantum mechanical method for organic molecules [7–10]. This method is comparable in computational speed with the AM1 and PM3 methods and is derived from density-functional theory (DFT) by an expansion of the DFT total energy up to second order in the charge density fluctuations around a reference density. The subsequent approximations lead to a generalized eigenvalue problem which has to be solved self-consistently for atomic charges. The method can be seen as an extension of so-called tight-binding (TB) methods to charge self-consistency. All parameters of this model are calculated from DFT and the method is therefore called a self-consistent

* Corresponding author. Fax: +62-21-42-23-33.
E-mail address: s.suhai@dkfz.de (S. Suhai).

0301-0104/01/$ - see front matter
PII: S0301-0104(00)00375-X

charge, density-functional tight-binding (SCC-DFTB) method.

In this paper we want to compare the ability of the approximate methods SCC-DFTB, AM1 and PM3 to reproduce the energetics and structure of small model peptides in order to prove their reliability for the description of larger polypeptides and proteins.

In recent years the secondary structure of proteins has been intensely studied by simulating model peptides with different theoretical methods. Larger polypeptides have been examined mostly with empirical force fields, while ab initio and DFT studies have been performed for smaller model compounds, like the *N*-acetyl-L-alanine-*N'*-methylamide molecule, where the potential energy surface (PES) in vacuo has been studied with Hartree–Fock (HF), Møller–Plesset (MP) perturbation theory and density functional theory [11–15]. These data give a reliable basis for benchmarking other more approximate methods. However, the *N*-acetyl-L-alanine-*N'*-methylamide (Ac-Ala-NHMe) molecule does not show important secondary structure motifs, like α_R helical and β-turn conformers which appear to be stable only for larger polypeptides [18]. It has been shown that solvation plays a crucial role for the stability of the different conformations. By applying a quantum chemical reaction field model at the HF/6-31G* level of theory [14], the α_R conformation is significantly stabilized with respect to the C_7^{eq} conformer. Including water molecules on a quantum theory level explicitly [16], the α_R conformer is found to be a stable conformation supporting free energy calculations with empirical force fields, which also stabilize this conformer in solution [17]. In a previous publication, we focussed on the stabilization of common secondary structures like α and 3_{10} helices and extended conformations with respect to the peptide size [18] using the DFT-B3LYP/6-31G*, SCC-DFTB and AM1 methods.

Here, we consider various small peptide models based on glycine and alanine residues and systematically analyze the ability of approximate quantum methods to reproduce energetic and structural properties of secondary structure elements like the extended, turn and helical conformations for small glycine and alanine based peptide models and the *N*-acetyl-(L-alanine)$_n$-*N'*-methylamide (Ac-Ala$_n$-NHMe) molecule with $n = 1, 2, 3, 5$ and 8.

In Section 2, we give a short description of the applied computational methods. In the following sections we discuss results for blocked polypeptide chains of different length and at different levels of theory. We first consider small glycine and alanine diamide and triamide structures, which have been examined with DFT and on the HF and post-HF levels of theory. Typical secondary structural elements, like helical, extended and turn conformations are investigated for polyalanine molecules with 3, 5 and 8 residues and the stabilization of helical conformations with respect to extended ones is discussed.

2. Methods

DFT-B3LYP/6-31G*, AM1 and PM3 calculations were carried out by using the GAUSSIAN 98 package [19]. By applying the PM3 method we included the additional force field (as provided by the keyword PM3MM in the GAUSSIAN package) which corrects for the PM3 deficiencies in the description of the peptide linkage. AM1 and PM3 have been parametrized to reproduce heats of formation of small molecules.

In this work, we compare energy differences between different conformers of peptide models for different theoretical methods. We compare total energy differences at the ab initio, DFT and SCC-DFTB levels of theory to the differences in the heats of formations evaluated with the AM1 and PM3 models. Energy corrections at the ab initio level of theory with respect to zero point vibrations (ZPE) have been calculated e.g. by Möhle et al. [20]. The difference in the ZP energies for extended and cyclic conformers for a triamide are in the range of 1 kcal/mol. This is a small value, but different conformers often differ in energy by this order of magnitude. Therefore, benchmarking the AM1 and PM3 methods with respect to total energies of ab initio methods might not lead to a fair evaluation of those approximate methods. However, the main failure of the AM1 and PM3 methods occurs not due to deviations in relative

energies of this magnitude, but due to an insufficient description of structures of some important secondary elements, as will be shown below.

Since the SCC-DFTB approach has been developed recently, we want to describe this method briefly. A detailed discussion of the model has already been given elsewhere [7,8].

The SCC-DFTB method is derived from DFT by a second order expansion of the DFT total energy functional with respect to the charge density fluctuations $\Delta\rho$ around a given reference density ρ_0 [7,8]. The second order terms in the density fluctuations are approximated by a simple distribution of atom-centered point charges $\Delta q_\alpha = q_\alpha - q_\alpha^0$, estimated by a Mulliken charge analysis. The approximate DFT energy functional becomes:

$$E_{\text{tot}} = \sum_{i}^{\text{occ}} \sum_{\mu\nu} c_\mu^i c_\nu^i H_{\mu\nu}[\rho_0] + \frac{1}{2} \sum_{\alpha\beta} \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} + E_{\text{rep}}[\rho_0]. \quad (1)$$

The Hamilton matrix elements $H_{\mu\nu}[\rho_0]$ are calculated within DFT-GGA in a two-center approximation using a minimal basis of atomic-like wavefunctions ϕ_μ. The second term on the right hand side represents the long-range Coulomb interactions between point charges at different sites and includes the self-interaction contributions of the single atoms. $E_{\text{rep}}[\rho_0]$ is approximated as a sum of two-body interactions, $E_{\text{rep}} = \sum_{\alpha\neq\beta} U(R_{\alpha\beta})$ which are determined by comparing bond-stretching energies calculated from the SCC-DFTB method with those from DFT calculations.

The results for reaction energies, geometries and frequencies for small organic molecules have been presented elsewhere [7,8]. The mean average deviations from experimental values are comparable to DFT calculations. The method has also been benchmarked for biologically relevant molecules, H-bonded complexes, small peptides and DNA H-bonded and stacking interactions [9,10,18]. Vibrational frequencies for small model peptides have been compared with results of B3LYP/6-31G* and MP2/6-31G* calculations and the vibrational absorption and vibrational circular dichroism spectra have been evaluated within an SCC-DFTB/DFT hybrid scheme, leading to a very good agreement with the results from the ab initio methods [22]. The benchmarks performed so far have been quite satisfactory, showing that the SCC-DFTB method is able to give a reliable description of several biological model molecules.

3. Diamides: Conformational energies and structures

3.1. Glycine dipeptide analogues and alanine dipeptide analogues

The simplest models for polypeptides are the glycine and alanine diamides (also called glycine dipeptide analogues (GDA) and alanine dipeptide analogues (ADA) [23], see also Fig. 1) which have been theoretically studied using HF, DFT and MP2 methods, respectively.

In this section we compare conformational energies, dipole moments and geometries (focusing on the backbone dihedral angles Φ and Ψ) of GDA and ADA evaluated with the SCC-DFTB, PM3 and AM1 methods in comparison with the higher level calculations at the DFT-B3LYP/6-31G* level of theory and with HF/6-31G*, MP2/6-31G* and DFT-BP (triple-ζ + polarization on heavy atoms, double-ζ + polarization on hydrogen) from Ref. [23].

For the GDA molecule, large deviations in relative energies among the ab initio methods occur, with HF being in qualitative disagreement with the other methods (Table 1). The values obtained at the SCC-DFTB and AM1 level of theory are satisfactorily compared to the deviations among the ab initio methods, while in the PM3 method both conformers relax into a single one.

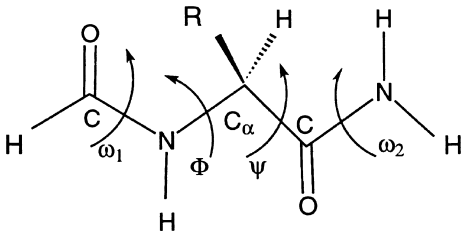

Fig. 1. The simplest models for polypeptides are the GDA (R = H) and ADA (R = CH_3).

Table 1
Conformational energies (kcal/mol) of GDA for the theoretical methods, as described in the text

Conformer	DFT-BP	HF	MP2	B3LYP	SCC-DFTB	AM1	PM3
C_7^{eq}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C_5^{ext}	0.45	−0.62	1.22	0.51	0.17	1.76	0.0

The DFT-BP, HF and MP2 results are from Ref. [23].

Table 2
Dihedral angles (degrees) of GDA for the theoretical methods, as described in the text

Conformer	Dihedral	DFT-BP	HF	MP2	B3LYP	SCC-DFTB	AM1	PM3
C_7^{eq}	Φ	−78.6	−85.2	−82.7	−80.2	−80.1	−80.3	−93.2
	Ψ	61.8	67.4	74.0	60.3	72.3	60.3	147.7
C_5^{ext}	Φ	−177.8	180.0	−178.2	180.0	−155.4	−114.4	−93.2
	Ψ	177.1	180.0	179.7	180.0	178.9	154.4	147.7

The DFT-BP, HF and MP2 results are taken from Ref. [23].

At the ab initio level, the dihedral angles Φ and Ψ deviate only within few degrees with the applied method (Table 2). However, for the C_5^{ext} conformer, MP2 and DFT-BP find a slight pyramidization at the *N* atoms, while HF and B3LYP give a planar conformation. This pyramidization is overestimated within the SCC-DFTB model, resulting in a deviation of the Φ angle of over 20° from the MP2 values. The AM1 and PM3 methods show the same tendency, where the deviations from planarity are significantly larger. In the PM3 method, both conformers relax into a single one, which in Φ–Ψ space is located between the C_7^{eq} and C_5^{ext} conformers, as can be seen from Table 2. A similar conformation is predicted by AM1 for the C_5^{ext} structure. In this conformation, the internal H-bond is broken, the O–H bond length has the value of 2.53 and 2.75 Å at the AM1 and PM3 level, respectively, while B3LYP predicts this H-bond to be 2.17 Å.

In Table 3 we compare the dipole moments for B3LYP, SCC-DFTB, AM1 and PM3 methods. In the SCC-DFTB method, the dipole moments are calculated from atomic charges, evaluated with the Mulliken charge approximation. Despite this approximation, the dipole moments compare quite well with those from B3LYP. The dipole moment for the C_7^{eq} conformer is well reproduced by the AM1 method. The large deviation for the C_5^{ext} conformer might be related to the distorted structure found for this conformation. The same is true for the dipole moment evaluated with PM3.

Table 3
Dipole moments (D) of GDA for the various conformers

Conformer	B3LYP	SCC-DFTB	AM1	PM3
C_7^{eq}	2.7	2.3	3.0	1.9
C_5^{ext}	3.9	3.4	1.6	1.9

For the ADA molecule, six stable conformers have been discussed [23]. The β_2 conformer is not a stable conformation in the SCC-DFTB, AM1 and PM3 models, it relaxes into a C_7^{eq} conformation. For the SCC-DFTB model the forces at the values of the dihedral angles in the β_2 region are small, as will be discussed below for the Ac-Ala-NHMe molecule.

All ab initio methods find the same energetic ordering of the conformers, although deviations of 1 kcal/mol in the relative energies appear (Table 4). The SCC-DFTB also reproduces the energetic ordering, but underestimates the energy differences consistently compared to the ab initio values. In the AM1 and PM3 models, the α_L conformation relaxes into a C_7^{ax} conformer. Further, the α_P relaxes into the C_7^{eq} conformer in the AM1 model. In the PM3 model, the C_7^{eq} and the C_5^{ext} relax into a single conformer, similar as in the case of GDA. Due to this, the relative energies at the AM1 and PM3 level are not satisfactory.

The dihedral angles for the five conformers at the various levels of theory are shown in Table 5.

Table 4
Conformational energies (kcal/mol) of ADA molecule for the theoretical methods, as described in the text

Conformer	DFT-BP	HF	MP2	B3LYP	SCC-DFTB	AM1	PM3
C_7^{eq}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C_5^{ext}	1.32	0.13	1.18	1.33	0.52	1.49	0.0
C_7^{ax}	2.07	2.52	2.17	2.36	0.94	0.54	3.02
α_L	3.15	4.57	4.25	5.46	3.19	0.54	3.02
α_P	6.53	5.71	5.14	6.44	4.30	0.0	2.86

The DFT-BP, HF and MP2 results are taken from Ref. [23].

Table 5
Dihedral angles (degrees) of ADA for the theoretical methods, as described in the text

Conformer	Dihedral	DFT-BP	HF	MP2	B3LYP	SCC-DFTB	AM1	PM3
C_7^{eq}	Φ	−80.0	−85.8	−82.9	−80.1	82.5	−82.2	−90.7
	Ψ	65.2	78.1	77.5	73.6	68.2	65.7	140.4
C_5^{ext}	Φ	−154.4	−155.6	−159.3	−159.9	−155.9	−113.0	−90.7
	Ψ	164.9	160.2	166.8	167.8	176.2	145.2	140.3
C_7^{ax}	Φ	69.3	75.1	73.9	72.1	73.3	75.1	67.9
	Ψ	−57.9	−54.1	−65.6	−58.5	−69.0	−61.8	−86.8
α_L	Φ	89.8	69.5	62.3	66.0	59.3	75.1	67.9
	Ψ	−0.9	24.9	37.3	29.4	23.0	−61.8	−78.7
α_P	Φ	−164.2	−165.6	−168.0	−171.3	−174.0	−82.2	−128.0
	Ψ	−41.0	−40.7	−36.2	−36.0	−44.6	65.7	−59.0

DFT-BP, HF and MP2 results from Ref. [23].

At the ab initio level, these values show variations of about 10° with the exception of the α_L conformer.

The SCC-DFTB dihedral angles compare well with the corresponding values at the higher theoretical levels. The largest error occurs for the α_L conformer, a structure which is not stabilized by an internal hydrogen bond. However, the energy hypersurfaces in the Ψ–Φ space are very shallow and stabilization effects could occur for larger polypeptides due to the internal hydrogen bonds. Compared to the deviations among the ab initio methods themselves (between 10° and 30°), the overall performance of the SCC-DFTB seems to be reliable.

Table 6 shows the dipole moments evaluated with the B3LYP, SCC-DFTB, AM1 and PM3 methods. The SCC-DFTB dipole moments are slightly but consistently smaller than those from B3LYP and, therefore, can reproduce the trends for the various conformers as indicated by the B3LYP values very well. The AM1 and PM3 dipole moments suffer from the same deficiency as discussed for the GDA molecule. The geometries do not compare well with those determined at the ab initio level of theory, therefore, the dipole moments deviate significantly.

Table 6
Dipole moments (D) of ADA for the various conformers

Conformer	B3LYP	SCC-DFTB	AM1	PM3
C_7^{eq}	2.5	2.2	2.8	1.9
C_5^{ext}	3.7	3.3	1.5	1.9
C_7^{ax}	3.0	2.6	2.9	2.3
α_L	5.2	4.9	2.8	2.3
α_P	4.5	4.3	2.8	6.0

3.2. Ac-Gly-NHMe and Ac-Ala-NHMe

Next, we discuss the *N*-acetyl-glycine-*N*′-methylamide (Ac-Gly-NHMe) and Ac-Ala-NHMe molecules which deviate from the GDA and ADA molecules only through the capping groups.

In Ac-Gly-NHMe, the C_5^{ext} conformer is 0.65 and 0.85 kcal/mol higher in energy than C_7^{eq} at the SCC-DFTB and B3LYP/6-31G* levels of theory, respectively. The SCC-DFTB model predicts, similar to the case of ADA, a nonplanar structure. However, the potential energy surface is very flat in this region. The planar conformation is only 0.1 kcal/mol higher in energy than the nonplanar conformation. At the B3LYP level of theory we find the planar configuration favored as for ADA.

Applying the AM1 and PM3 models we find the same problems with these two conformers as discussed for GDA and ADA. Therefore, we will not discuss this molecule in more detail.

Ac-Ala-NHMe has six stable conformers on the B3LYP/6-31G* and MP2/6-31G* PES [12], which have been taken as starting structures for further geometry optimizations with the SCC-DFTB, AM1 and PM3 methods. The three lowest energy conformers have internal H-bonds, whereas the higher energy conformers do not.

Based on the discussion of Saebo et al. [24] of the Local MP2 (LMP2) method, Beachy et al. [15] estimated the internal basis set superposition error (BSSE) in the correlation energy by taking the difference of the LMP2 and MP2 energies (with cc-pVTZ(-f) basis set at the MP2/6-31G* geometry). This BSSE estimates were subtracted from MP4/cc-pVTZ(-f)//MP2/6-31G* energies and Beachy et al. argue this 'MP4-BSSE' model to be their best estimation of the relative energies of the Ac-Ala-NHMe molecule.

The relative energies are given in Table 7 for different levels of ab initio theories (HF, MP2 and B3LYP with 6-31G* basis set from Ref. [12] and 'MP4-BSSE' from Ref. [15]), and for the approximate methods SCC-DFTB, AM1 and PM3. Internal BSSE favors cyclic conformers over more extended ones [15]. Therefore, a large effect of BSSE can be seen in the relative energies of the β and C_5 conformers which are more extended compared to the more compact C_7 conformation. The C_7 conformation is stabilized relative to the C_5 conformer due to BSSE which is reflected in the difference of the MP2 and 'MP4-BSSE' results. In this case, the internal BSSE is 0.5 kcal/mol in favor of the C_7 structure (difference of MP2 and LMP2). The effect of correlation going from the MP2 to the MP4 level is to decrease the C_7–C_5 energy difference by another 0.2 kcal/mol [15].

The performance of the approximate SCC-DFTB, AM1 and PM3 methods for this molecule is similar to the examples discussed above. However, it can be seen that the performance of the SCC-DFTB model appears to be much better when compared to 'MP4-BSSE'. Therefore, similar effects might also be expected for the smaller peptide models as discussed above.

Despite this good performance of the SCC-DFTB model, two shortcomings are obvious: in the case of the C_7^{ax} conformer, the steric interaction of this methyl group with the hydrogen bonded ring seems to be underestimated in the SCC-DFTB (also for AM1 and PM3) method, which results in an underestimation of its relative energy. The β_2 conformer is found to be unstable in the SCC-DFTB model (as for the ADA molecule), but the maximum force at the dihedral angles given for the

Table 7
Relative energies (kcal/mol) of the different conformers of Ac-Ala-NHMe for different methods as described in the text

Conformer	MP4-BSSE	B3LYP	MP2	HF	SCC-DFTB	PM3	AM1
C_7^{eq}	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C_5^{ext}	0.89	1.43	1.76	0.41	0.99	−1.55	1.72
C_7^{ax}	2.55	2.58	2.61	2.82	1.03	0.87	0.72
β_2	2.56	3.18	3.37	2.58	2.20[a]	–	–
α_L	4.21	5.82	4.60	4.72	3.70	3.53	–
α_P	5.17	6.85	6.34	5.74	4.78	1.10	3.29

The MP4-BSSE results are taken from Ref. [13] while the B3LYP, MP2 and HF results are from Ref. [10].
[a] The β_2 conformer is not stable within the SCC-DFTB, but the maximum force at the B3LYP geometry is very small. The energy is given for the geometry, when the maximum force is smaller than 0.00065 a.u.

Table 8
Dihedral angles (degrees) of the Ac-Ala-NHMe conformers for different methods, as described in the text

	C_7^{eq}		C_7^{ax}		C_5^{ext}		β_2		α_L		α_P	
	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ
B3LYP	−81.9	72.3	73.8	−60.0	−157.3	165.3	−135.9	23.4	68.5	24.5	−169.4	−37.8
SCC	−81.3	72.0	74.6	−66.1	−153.2	176.6	−136.7	24.9[a]	65.6	13.0	−172.5	−51.1
AM1	−84.4	68.5	76.6	−64.0	−117.7	141.5	–	–	–	–	−115.5	−55.2
PM3	−71.4	77.7	68.8	−67.9	−93.9	147.9	–	–	62.3	39.6	−137.6	−60.5

[a] The β_2 conformer is not stable within the SCC-DFTB method. The Φ, Ψ values refer to a conformation where the maximum force is lower than 0.00065 a.u.

SCC-DFTB in Table 8 is very small. This conformer may be stabilized due to internal H-bonds in larger polypeptides.

As mentioned above, AM1 and PM3 give a very distorted C_5^{ext} conformation, where the internal H-bond is broken, as can be seen from the dihedral angles in Table 8. The β_2 conformer is also not stable in both methods and the α_L is unstable in the AM1 model.

Dipole moments show similar trends as reported above for the other peptide models, and are therefore not presented here. The SCC-DFTB consistently underestimates them compared the B3LYP/6-31G* results while PM3 and AM1 predict dipole moments quite well, as long as the geometry of the conformer is described well.

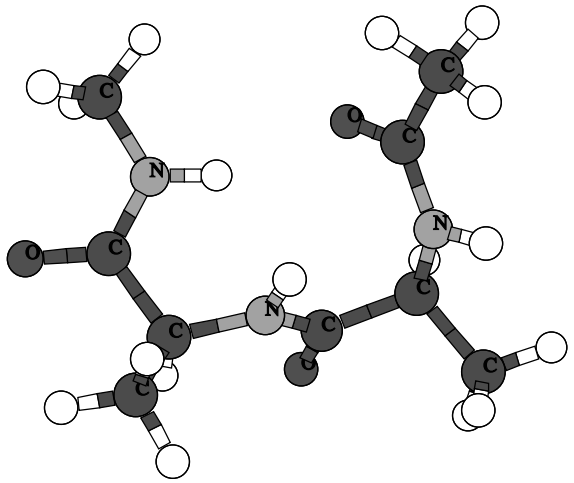

Fig. 2. In the dipeptide shown the oxygen atom of the acetyl group can form a H-bond with the nitrogen atom of the methylamide group, a so-called $i \rightarrow i+3$ H-bond, since this H-bond connects the ith residue (here modeled by the acetyl group) along a polypeptide chain with the residue $i+3$ (modeled by the NHMe group). Protein α-helices form an $i \rightarrow i+4$ H-bond pattern, therefore, they cannot appear in this dipeptide.

4. Ac-Gly$_2$-NHMe and Ac-Ala$_2$-NHMe

Next we discuss longer amides which allow the formation of other secondary structural motifs due to the possibility of additional internal H-bond formation, beginning with the triamides (dipeptides) Ac-Gly$_2$-NHMe and Ac-Ala$_2$-NHMe. In these models, containing two peptide groups blocked with the acetyl and methylamide groups, the formation of turn structures is possible. The β-turn structures reverse the direction of a polypeptide and are, therefore, a frequently occurring structural motif in proteins. In the dipeptide (see Fig. 2), the oxygen atom of the acetyl group can form an H-bond with the nitrogen atom of the methylamide group, a so-called $i \rightarrow i+3$ H-bond, since this H-bond connects the ith residue (here modeled by the acetyl group) along a polypeptide chain with the residue $i+3$ (modeled by the NHMe group). Protein α-helices form an $i \rightarrow i+4$ H-bond pattern, therefore, they cannot appear in this dipeptide. Longer polypeptides, starting from the Ac-Ala$_3$-NHMe model would, in principle, allow the formation of such H-bonds.

Turn structures have been studied at the HF/3-21G [25], HF, DFT-B3LYP and MP2 level of theory [20,26,27]. While Böhm [26] evaluated the energies at the MP2/DZP level for HF/DZP geometries, Möhle et al. [20,27] performed full geometry optimizations at the MP2/6-31G* and B3LYP/6-31G* levels of theory, respectively. Here

we focus on the C_7^{eq}, C_5^{ext} linear repeat conformers and on two β-turn structures, the type I turn, β_I, and type II turn, β_{II} as in the mentioned previous publications. β-turns are classified by idealized backbone dihedral angels [28]. A β_I type turn structure is characterized by the values $\Phi_1 = -60°$ and $\Psi_1 = -30°$, $\Phi_2 = -90°$ and $\Psi_2 = 0°$, respectively. A β_{II} type turn is classified by the ideal dihedral angles $\Phi_1 = -60°$, $\Psi_1 = 120°$, $\Phi_2 = 80°$ and $\Psi_2 = 0°$. A β_{III} type turn structure is characterized by the values of $\Phi_1 = -60°$, $\Psi_1 = -30°$, $\Phi_2 = -60°$ and $\Psi_2 = -30°$, but this turn is not a stable conformer in this dipeptide. In case of Ac-Ala_2-NHMe, we also consider the inversed $\beta_{I'}$ and $\beta_{II'}$ turns, classified by $\Phi_1 = 60°$ and $\Psi_1 = 30°$, $\Phi_2 = 90°$ and $\Psi_2 = 0°$ ($\beta_{I'}$), $\Phi_1 = 60°$, $\Psi_1 = -120°$, $\Phi_2 = -80°$ and $\Psi_2 = 0°$ ($\beta_{II'}$), respectively. All these turn structures form $i \to i+3$ H-bonds, but only the type β_{III} turns are usually referred to as 3_{10} helices.

Table 9 shows the relative energies of the four conformers of Ac-Gly_2-NHMe (MP2/6-31G* and HF/6-31G* values from Ref. [20,27]). For this molecule, MP2 and HF are in qualitative disagreement. This has been pointed out in detail by Möhle et al. [27]. Although the β-turns form only one internal H-bond, while the extended conformer forms two, the β-turns are favored energetically over the extend one at the MP2 level of theory. The β-turns are more compact conformations than the extended C_5^{ext} conformer. Möhle et al. [20] discuss several other examples from the literature, where the inclusion of correlation energy seems to favor more compact conformers over the extended ones. They argue that this effect might also occur in this example and might explain the reversed stabilities at the MP2 level. On the other hand, as discussed above for the Ac-Ala-NHMe molecule, the internal BSSE at the MP2 level might cause a large part of the stabilization of the more compact conformations [15].

Compared to the B3LYP value of 1.37 kcal/mol, the 3.42 kcal/mol relative energy of the C_5^{ext} conformer at the MP2 level is very large. However, a similarly large difference has already been found for the GDA model (Table 1). Judging from the results for Ac-Ala-NHMe, the MP2/6-31G* level seems to have the tendency to predict the C_5^{ext} conformations to be very high in energy, which might be even pronounced for glycine based peptides. Although HF, B3LYP and MP2 agree well on the relative stabilities of the two β-turns, no conclusion can be drawn here on their relative stability with respect to the extended conformer, since it is not clear to which extent correlation contributions are responsible for these effects and to which extent BSSE (not only in the correlation energy) is responsible. Furthermore, MP2 on the one hand, and HF and B3LYP on the other hand disagree in the structure of the extended conformer, which is slightly non planar at the MP2 level, and planar at the HF [20] and B3LYP (Table 10) levels of theory. This disagreement appeared already for the GDA molecule and might be another source for the disagreement in the relative stabilities. We did not calculate vibrational frequencies for this conformer as has been done for Ac-Gly-NHMe, but we performed a second geometry optimization by taking the backbone geometry of Ac-Ala_2-NHMe (which is nonplanar, see below) as a starting point. This was also relaxed into a planar conformation. For the other conformers B3LYP/6-31G* dihedral angles for Ac-Gly_2-NHMe deviate only slightly from those obtained at the MP2/6-31G* level.

In the SCC-DFTB model, the energy differences between the conformers are smaller than at the B3LYP level of theory, as has been found also for

Table 9
Relative energies (kcal/mol) for the Ac-Gly_2-NHMe conformers at the different levels of theory as discussed in the text

Ac-Gly_2-NHMe	B3LYP	SCC-DFTB	PM3	AM1	MP2	HF
C_7	0.0	0.0	0.0	0.0	0.0	0.0
C_5	1.37	0.80	1.55	3.89	3.42	−0.65
β_I	2.11	1.60	0.36	0.0	0.74	0.76
β_{II}	1.24	0.53	0.50	2.24	−0.07	−0.36

Table 10
Dihedral angles (in degrees) of Ac-Gly_2-NHMe at different levels of theory

Conformer	Method	Φ_1	Ψ_1	Φ_2	Ψ_2
C_7^{eq}	MP2	−83.2	66.4	−85.2	67.7
	B3LYP	−81.2	63.7	−82.7	64.7
	SCC-DFTB	−80.2	66.4	−81.9	65.0
	PM3	−70.5	67.5	−73.4	72.7
	AM1	−81.7	59.4	−81.3	59.6
C_5		−171.2	−176.9	−179.8	−179.8
		179.9	179.9	180.1	180.1
		151.5	176.4	159.2	178.0
		91.0	−164.1	177.3	−168.7
		111.0	−158.5	110.0	−157.8
β_I		−72.1	−21.2	−99.6	15.3
		−73.6	−13.7	−105.3	16.3
		−69.2	−9.0	−112.9	23.7
		−116.1	−86.8	−86.3	169.1
		−81.7	59.3	−81.3	59.6
β_{II}		−58.6	139.8	92.7	−14.0
		−62.3	130.3	102.5	−16.7
		−66.2	96.2	121.9	−15.0
		−60.0	118.4	105.0	−34.9
		−71.5	86.7	104.2	17.1

the other peptide models discussed so far. But, as in the case of the Ac-Ala-NHMe molecule, correction for internal BSSE could change the picture and make the comparison more favorable for the SCC-DFTB model. The values of the dihedral angles compare reasonably well with those of B3LYP, showing that the β type conformers are intrinsically stable also at the SCC-DFTB level of theory. Larger errors occur for the C_5^{ext} and β_{II} conformers. As in the case of GDA and Ac-Gly-NHMe, the SCC-DFTB seems to favor a larger pyramidalization of the nitrogen atom, leading to the deviation form planarity. However, the PES is extremely shallow at this point, the conformation given in Table 10 is only 0.1 kcal/mol more stable than the planar one.

The large deviations in relative energies, evaluated with AM1 and PM3, may again result from the very distorted structures predicted by these methods. The C_5^{ext} conformer shows the same deviations as described above for the smaller molecules. Furthermore, both β type structures are not intrinsically stable at the AM1 level, the β_I conformer relaxes into the C_7^{eq} conformer and the β_{II} relaxes into a C_7^{eq} conformation at the N-terminus, which has been reported before [26]. PM3 describes the β_{II} structure reasonably well, but the β_I conformer is very distorted. There, the internal H-bond is broken, leading to an H-bond length of 3.3 Å.

Compared to the B3LYP values, the SCC-DFTB again systematically underestimates the dipole moments, while the dipole moments evaluated with AM1 and PM3 are reasonable only for those conformers where the structure is well described, i.e. the C_7^{eq} and the β_{II} conformers (Table 11).

In contrast to the situation in Ac-Gly_2-NHMe, the β_{II} conformer is higher in energy than the β_I conformer in Ac-Ala_2-NHMe at the B3LYP/6-31G* level of theory, which is in agreement with HF/6-31G* and MP2/6-31G*//HF/6-31G* results from Möhle et al. [21], shown in Table 12 (the energy difference of C_7^{eq} and C_5^{ext} at the MP2/6-31G*//HF/6-31G* and HF/6-31G* levels, which is not included in [21] has been calculated in our laboratory). However, as in the case of Ac-Gly_2-NHMe, the C_5^{ext} conformer is higher in energy than the β_I and β_{II} turns at the MP2 level, while it is vice versa at the B3LYP and HF levels of theory. After

Table 11
Dipole moments (D) for the Ac-Gly_2-NHMe conformer at the B3LYP, SCC-DFTB, PM3 and AM1 levels of theory

Ac-Gly_2-NHMe	B3LYP	SCC-DFTB	PM3	AM1
C_7	6.2	5.5	6.6	5.6
C_5	6.6	5.6	3.3	3.3
β_I	8.3	7.6	4.8	5.6
β_{II}	6.6	6.2	6.4	6.4

Table 12
Relative energies (kcal/mol) for the Ac-Ala_2-NHMe conformers at the different levels of theory, as discussed in the text

Ac-Ala_2-NHMe	B3LYP	SCC-DFTB	PM3	AM1	MP2	HF
C_7	0.0	0.0	0.0	0.0	0.0	0.0
C_5	2.13	1.51	3.27	3.02	2.66	0.23
β_I	2.59	1.81	1.15	2.01	1.27	1.14
β_{II}	4.32	2.83	3.17	4.44	2.23	2.28
$\beta_{I'}$	6.25	3.82	7.13	5.95	3.73	4.89
$\beta_{II'}$	3.97	1.53	2.16	1.94	2.18	2.66

MP2/6-31G*//HF-6-31G*.

the discussion so far, relative energies have to be taken with some care. At the level of theory we use here, it is not clear to which extent they are contaminated with internal BSSE. Discrepancies in the structures of the extended conformers between MP2 and B3LYP seem to occur for the glycine based peptides, for alanine based models like the ADA and Ac-Ala-NHMe molecules, the structures are very similar on the two levels which is also the case for Ac-Ala_2-NHMe (MP2/6-31G* geometries are given in Ref. [21]).

Inspection of the relative energies of the Ac-Ala-NHMe molecule in comparison with the relative energies of longer peptides (although they will be discussed in more detail below), might give some further insight. In Table 13 we show the energy difference of the C_7^{eq} and C_5^{ext} conformers per residue (i.e. total energy difference divided by the number of residues) for various chain lengths. The energy difference per residue slightly decreases with increasing number of residues. This might have two sources. First, there are small cooperative effects, i.e. the energy gain for inserting a residue into a chain containing already n residues depends on n. These effects are large for helical structures, but small for C_7^{eq} and C_5^{ext} conformers with more than 3 residues. For example, inserting one (C_7^{eq} or C_5^{ext}) residue into a chain with $n = 3$ leads to a binding energy, which is 0.1 kcal/mol larger than inserting the residue to a chain with $n = 2$ (evaluated at the B3LYP level). Interestingly, the B3LYP values for the energy difference converge to the value of 1 kcal/mol, which is close the energy difference of the MP4-'BSSE' model in Ac-Ala-NHMe, and closer to the energy difference of 0.8 kcal/mol at the B3LYP/6-311+G** [27] level of theory. One possible interpretation is that in the longer chains the basis functions of neighboring residues lead to a more complete basis for the particular residues.

Table 13
Energy difference (kcal/mol) of the C_7^{eq} and C_5^{ext} conformers divided by the number of residues n $(= 1, 2, 3, 5)$ for Ac-$(Ala)_n$-NHMe

	1	2	3	5
B3LYP	1.43	1.07	1.00	0.98
SCC-DFTB	0.99	0.75	0.76	0.74
HF	0.41	0.12	0.12	–
MP2	1.55	1.33	1.35	–

B3LYP denotes B3LYP/6-31G*, HF denotes HF/6-31G* and MP2 denotes MP2/6-31G*//HF/6-31G*.

The table also shows that effects found for Ac-Ala-NHMe may persist as trends in the longer chains. At the HF and SCC-DFTB levels, the energy difference is underestimated for Ac-Ala-NHMe, and this trend remains valid for the longer chains. Therefore, we expect that certain deficiencies of particular methods (like the small energy

difference in HF) may add up when going to larger chains.

This finding may help in interpreting trends in relative energies for longer peptides by inspecting the relative energies of Ac-Ala-NHMe. Compared to the MP4 values, the MP2 strongly underestimates the stability of the C_5^{ext} conformer in Ac-Ala-NHMe while HF overestimates it. Due to this, we expect the Ac-Ala_2-NHMe C_5^{ext} conformer to be too stable at the HF level and too high in energy at the MP2 and B3LYP (6-31G*) levels of theory.

Large differences between B3LYP and MP2 appear for the β-type conformations, similar to the case of Ac-Gly_2-NHMe. Comparing to the results of Ac-Ala-NHMe, we find an agreement between B3LYP and MP2 for the relative energy of the β_2 conformation (but higher in both cases compared to MP4-BSSE), but a large discrepancy is found for the α-helical regions in Ac-Ala-NHMe, where B3LYP seems to lead to too high energies by about 1 kcal/mol. Since the values of the Φ–Ψ angles of the β_I conformer at the N-terminal are close to the α_R helical region and the Φ–Ψ angles of the β_{II} at the C-terminal are close to the α_L values in Ac-Ala-NHMe, this might lead to higher energies at the B3LYP than at the MP2 level.

The trends discussed for Ac-Ala_2-NHMe will also hold for longer peptides, presented below. Compared to the B3LYP results for relative energies (Table 12), geometries (Table 14) and dipole moments (Table 15), the same tendencies for the SCC-DFTB as described above are found. The energy differences and the dipole moments are slightly smaller than the B3LYP values but the dihedral angles are close to the B3LYP results. Deviations in the dihedral angles occur up to 20°, as for the β_{II} conformer, but similar deviations may be expected by comparing different ab initio methods, since in Φ–Ψ space the energy surface is

Table 14
Dihedral angles (in degrees) of Ac-Ala_2-NHMe at different levels of theory

Conformer	Method	Φ_1	Ψ_1	Φ_2	Ψ_2
C_7^{eq}	B3LYP	−82.7	69.5	−84.6	70.0
	SCC-DFTB	−81.3	68.8	−83.3	68.7
	PM3	−76.8	70.7	−76.9	76.1
	AM1	−84.4	66.9	−84.0	65.4
C_5		−158.7	165.5	−159.4	165.0
		−157.4	175.7	−160.5	178.5
		−91.0	151.3	−94.4	150.7
		−109.0	151.2	−107.7	148.5
β_I		−74.7	−12.3	−105.5	13.1
		−70.6	−5.4	−111.0	18.1
		−129.7	−66.3	−87.6	140.3
		−84.7	70.2	113.0	−49.2
β_{II}		−60.8	128.8	69.8	15.1
		−59.2	110.2	64.1	20.9
		−59.4	123.7	65.9	21.3
		−72.8	92.2	87.6	29.8
$\beta_{I'}$		62.1	31.0	65.7	18.6
		59.3	25.4	63.6	14.4
		53.2	51.3	65.2	11.8
		46.1	56.0	77.9	−56.4
$\beta_{II'}$		55.9	−128.1	−105.2	18.2
		62.1	−96.1	−115.7	12.9
		57.4	−103.1	−84.7	−22.5
		69.5	−78.2	−102.1	−32.8

Table 15
Dipole moments (D) for the Ac-Ala_2-NHMe conformer at the B3LYP, SCC-DFTB, PM3 and AM1 levels of theory

Ac-Ala_2-NHMe	B3LYP	SCC-DFTB	PM3	AM1
C_7	5.8	5.3	6.6	5.6
C_5	6.3	5.8	2.3	3.3
β_I	8.1	7.5	3.8	6.4
β_{II}	6.6	6.3	6.1	5.8
$\beta_{I'}$	9.0	8.3	8.3	6.9
$\beta_{II'}$	8.1	6.4	6.5	6.2

very shallow around the local minima. SCC-DFTB reproduces the B3LYP tendency to favor the linear structures over the cyclic ones, which should be expected, since the SCC-DFTB method is an approximation to DFT.

AM1 and PM3 reproduce for this molecule the trends reported up to now. The β-turn structures are not stable or very distorted. AM1 has the tendency to relax these conformers into C_7 conformations, while PM3 finds very distorted structures for the β_I conformers, where the internal H-bond is broken, leading to a H-bond length of about 2.7–3.3 Å.

We now discuss larger peptide models and focus on alanine based polypeptides, with 3, 5 and 8 alanine residues, blocked with the acetyl group at the N-terminus and the methylamide group at the C-terminus.

5. Ac-Ala_n-NHMe, $n = 3, 5, 8$

For Ac-Ala_3-NHMe, we built up all permutations of the eight possible conformations for Ac-Ala-NHMe, i.e C_7^{eq}, C_5^{ext}, C_7^{ax}, β_2, α_R, α_L, α_D, α_P as they are determined by the dihedral angles given in Ref. [12], resulting in 512 structures for the Ac-Ala_3-NHMe molecule. These structures have been optimized using the Amber4.1 force field [30]. We chose all structures which were lower in energy than the repeated C_7^{ax} conformer for further investigations, 14 structures in total. We reoptimized the 14 structures with the B3LYP/6-31G*, SCC-DFTB, AM1 and PM3 methods. None of these structures contains a residue in the C_5^{ext} conformation. Further, the α_R and α_L conformations are also not present in this set. Therefore, we considered additionally the repeat-C_5, α_R and α_L structures.

The α_R and β_2 conformations are both not stable at the B3LYP level of theory, they relax into a 3_{10} type helix, while the α_L conformer relaxes into a left-handed 3_{10} type helix, 3_{10}^L.

From the grid search, we found a structure which is lower in energy than the repeat-C_7^{eq} structure and will be called global minimum (gm). This structure can be labelled as $C_7^{eq}C_7^{ax}C_7^{eq}$, the second residue is in a C_7^{ax} conformation, while the first and the third residues assume C_7^{eq} conformations. In the following, we compare B3LYP, SCC-DFTB, AM1 and PM3 relative energies (Table 16) and geometries (Table 17) for the gm, C_7^{eq}, C_5^{ext}, C_7^{ax}, 3_{10} and 3_{10}^L conformations. The 3_{10} conformers relax at the C-terminus into a β_I type conformation, as can be seen from the dihedral angles in Table 17.

Beachy et al. [15] studied ten conformers of this molecule, generated by a limited conformational search. These conformers were optimized at the

Table 16
Relative energies (kcal/mol) of Ac-Ala_3-NHMe conformers as described in the text

Conformer	B3LYP	SCC-DFTB	PM3	AM1	LMP2
gm: $C_7^{eq}C_7^{ax}C_7^{eq}$	0.0	0.0	0.0	0.0	0.0
$[C_7^{eq}]_3$ linear	0.1	1.2	1.3	0.3	–
$[C_5]_3$ linear	3.1	3.5	−3.2	5.5	2.7
3_{10}	2.6	2.6	4.2	5.4	–
$[C_7^{ax}]_3$ linear	6.3	3.7	3.3	3.3	6.9
3_{10}^L	7.1	4.8	11.8	11.0	–
$3_{10}^L C_7^{ax}$	8.5	5.5	9.7	7.6	7.0

The 6-31G* basis set was used at the B3LYP level, LMP2 refers to local MP2/cc-pVTZ(-f) calculations at HF/6-31G* geometries as described in the text.

Table 17
Dihedral angles for the Ac-Ala_3-NHMe molecule, as described in the text

Conformer	Method	Φ_1	Ψ_1	Φ_2	Ψ_2	Φ_3	Ψ_3
gm	B3LYP	−75.8	91.9	74.5	−56.7	−74.7	82.9
	SCC-SFTB	−75.2	83.1	74.6	−55.8	−71.9	82.6
	PM3	−77.1	106.3	62.2	−66.1	−69.7	77.9
	AM1	−78.7	81.6	75.6	−58.8	−74.9	84.8
C_7^{eq}		−82.4	70.1	−84.2	66.3	−84.8	70.3
		−81.1	70.0	−82.7	67.1	−83.4	66.3
		−77.2	71.3	−78.4	70.1	−77.5	75.8
		−84.5	67.1	−84.3	64.3	−84.4	65.7
C_5		−158.7	167.1	−158.3	166.2	−159.3	164.9
		−156.0	175.2	−162.8	175.3	−162.3	176.4
		−90.9	150.9	−91.9	152.7	−94.4	150.6
		−109.7	150.7	−109.9	151.9	−107.7	148.7
C_7^{ax}		73.6	−55.6	73.7	−53.3	73.1	−54.1
		75.0	−65.2	75.9	−63.7	75.7	−63.4
		70.0	−65.3	69.3	−65.0	68.3	−67.7
		77.0	−64.0	77.0	−62.1	76.8	−61.7
3_{10}		−68.7	−22.4	−69.4	−7.4	−104.0	10.9
		−63.1	−21.6	−68.3	−5.4	−107.5	16.0
		−82.0	141.1	−106.9	−53.2	−85.2	137.8
		−84.8	70.8	−105.0	−40.8	−82.9	67.9
α_L		60.6	32.1	57.2	25.0	64.2	22.6
		57.9	26.4	57.3	23.7	66.5	12.2
		52.7	46.2	59.1	18.4	59.6	31.7
		75.5	−25.0	92.9	6.9	77.6	−63.2
$3_{10}^L C_7^{ax}$		61.7	31.1	67.4	14.3	71.4	−54.1
		60.8	19.4	51.5	29.8	72.8	−59.9
		75.5	−25.0	93.0	6.9	77.6	−63.2
		59.6	40.7	49.7	43.3	59.5	76.1

The first row labels the conformer, the second row the theoretical method, always in the order B3LYP, SCC-DFTB, PM3 and AM1. The dihedral angles are labeled starting from the Φ-angle at the N-terminus.

HF/6-31G** level of theory and energies were calculated for those geometries with LMP2/cc-pVTZ(-f). Three of their structures are similar to ours, the dihedral angles deviating only by about 5° from our B3LYP values: Their structure 3 corresponds to our gm (and is also the lowest energy structure), their structure 1 corresponds to the repeat-C_5 conformer and structure 9 corresponds to the repeat-C_7^{ax} conformation. We also included their conformer 10 into our comparison, which is described by a left-handed 3_{10}^L conformation at the N-terminus and a C_7^{ax} conformation at the C-terminus. This conformer will we labeled in the following as $3_{10}^L C_7^{ax}$. The relative energies for these four conformers at the LMP2/cc-pVTZ(-f) level from Ref. [15] are also given in Table 17.

B3LYP and LMP2 agree in the energy difference between the gm and linear C_7^{ax} conformer very well. The energy difference between C_7^{eq} and C_7^{ax} for Ac-Ala-NHMe is well reproduced at the B3LYP level compared to the results of the 'MP4-BSSE' model. From the results of Ac-Ala-NHMe (Table 7) we expect the C_5^{ext} to be slightly too high in energy, which can be seen from Table 17. From the Ac-Ala-NHMe results we expect the turn and helical structures β, α_L and α_P to be slightly too high in energy at the B3LYP level of theory, e.g. the α_L is more than 1 kcal/mol higher than at the

MP2 and MP4-BSSE levels. The higher energy of $3_{10}^{L}C_7^{ax}$ at the B3LYP compared to the LMP2 level might be explained by a lower intrinsic stability of these helical conformations at the B3LYP level. Therefore, the stability of this conformer as well as the stability of the 3_{10} conformation might be underestimated at the B3LYP level of theory.

The SCC-DFTB method reproduces structures very well. However, the main deficiency is the overstabilization of the C_7^{ax} conformation. The gm and repeat-C_7^{eq} differ by only one C_7^{ax} conformation and are energetically nearly degenerate at the B3LYP level. SCC-DFTB predicts the C_7^{ax} in Ac-Ala-NHMe to be more that 1 kcal/mol too low, which explains the energy difference of gm and repeat-C_7^{eq} of 1.2 kcal/mol. If we assume that the energy of every C_7^{ax} conformation is underestimated by 1.2 kcal/mol and add this as a correction to the SCC-DFTB relative energies, the repeat-C_7^{ax} conformation would have an energy of 6.1 kcal/mol which is close to the B3LYP value of 6.3 kcal/mol relative energy. Both, gm and $3_{10}^{L}C_7^{ax}$, contain one C_7^{ax} conformation. The smaller energy difference of 5.5 kcal/mol at the SCC-DFTB level compared with the 7.0 kcal/mol at the LMP2 level might be explained by the lower intrinsic energy of the turn and helical structures at the SCC-DFTB level already found for Ac-Ala-NHMe.

Both, AM1 and PM3, find very distorted structures for the linear C_5^{ext} repeat conformation, which causes the deviations in the relative energies. Further, for the helical 3_{10} conformer PM3 predicts a distorted structure where the helix is partially unwound, leading to a breaking of the H-bonds. The same is true for the 3_{10}^{L}. However, the first O–H bond, from the acetyl oxygen at the N-terminus to the amide hydrogen at residue 2, assumes a bond length of 1.85 Å, i.e., this partial structure is predicted reasonably, but the second H-bond, from the oxygen of residue 1 to the methylamide hydrogen at the C-terminus, is very long, $r(\text{O–H}) = 2.7$ Å. A similar behavior is found for the gm structure. There, the internal H-bond at the N-terminus is also broken, i.e., the C_7^{eq} conformation is deformed similar to the case of GDA and ADA. As can be seen from the dihedral angles, the AM1 method relaxes the first and last residue of the 3_{10} helix into C_7^{eq} conformations, i.e., this helix is not stable. Further, the third residue of the 3_{10}^{L} is relaxed into a C_7^{ax} conformation.

The H-bond lengths vary with different conformations and will not be given in detail. At the B3LYP level of theory, these bond lengths are about 2 Å. In the C_7 conformers, they are slightly shorter, in C_5 and 3_{10} they are longer. The SCC-DFTB method gives H-bond lengths consistently shorter by 0.05–0.1 Å than B3LYP, the same tendency is found with the PM3 method for those structures, which are not distorted. AM1 predicts consistently longer H-bonds by 0.1–0.15 Å.

The other conformers mentioned above which were generated by the grid search using the Amber4.1 force field are a mixture of C_7, β, and 3_{10} secondary structure motifs (not shown). The dihedral angles for these conformers evaluated with the SCC-DFTB model, show similar deviations from those at the B3LYP/6-31G* level of theory, as they appear for the conformers given in Table 17. However, the B3LYP results deviate from MP2 and LMP2 results in the range of 2–3 kcal/mol, and from the discussion so far it seems that helical conformers are similarly overstabilized in SCC-DFTB as they are understabilized at the B3LYP level. The SCC-DFTB model, therefore, predicts structures accurately, the energy differences, as in all other cases discussed so far, are smaller than at the B3LYP level of theory. In AM1 and PM3, the β type secondary structure motifs are not stable, they relax into C_7 conformations. The 3_{10} helices at the PM3 level are unwound as described above.

For $n = 5$, 8, we used a genetic algorithm in combination with the Amber4.1 [30] force field to determine the gm of these peptides [29]. We optimized the gm, the C_7 and C_5 and the helices 3_{10} and α_R with B3LYP, SCC-DFTB, AM1 and PM3. Relative energies are shown in Table 18, dihedral angles will not be given, since they are similar to those in the smaller peptides discussed so far.

The gm structures for $n = 5$, 8 are compact structures. For $n = 8$, it is basically a mixture of the two C_7 type secondary structure elements where an extra H-bond is formed between the two chain ends, in the case of $n = 5$ different helical and turn conformations appear along the chain.

From the discussion of the energetics of Ac-Ala-NHMe above, we expect the energy differ-

Table 18
Relative energies (kcal/mol) with respect to the C_7^{eq} conformer for several conformations of Ac-Ala$_n$-NHMe, $n = 5$, 8 at the B3LYP/6-31G*, SCC-DFTB, PM3 and AM1 levels of theory

	B3LYP	SCC-DFTB	AM1	PM3
$n = 5$				
gm	−7.7	−7.9	2.0	0.6
C_7^{eq}	0.0	0.0	0.0	0.0
C_5^{ext}	4.8	3.7	8.0	7.1
C_7^{ax}	9.6	3.9	2.4	3.0
$3_{10}/\beta_{II}$	0.8	−1.1	5.5	1.5
$n = 8$				
gm	−7.8	−7.2	6.1	6.5
C_7^{eq}	0.0	0.0	0.0	0.0
$3_{10}/\beta_{II}$	−3.6	−6.4	5.8	4.5

ences between the C_7^{eq} and C_7^{ax} to be reproduced reliably at the B3LYP level of theory, at the SCC-DFTB level the C_7^{ax} is too low in energy by more than 1 kcal/mol per residue. The C_5 conformation is expected to be slightly too high at the B3LYP level, and slightly too low at the SCC-DFTB level of theory, while for the 3_{10} helix we expect the reverse to be true. The 3_{10} relaxes at the C-terminus into a β_{II} type conformation at the B3LYP and SCC-DFTB level of theory. The geometry at the C-terminus is similar to that in the case of $n = 3$. 3_{10} and α_R helices have been investigated in detail on the B3LYP/6-31G*, SCC-DFTB and AM1 level of theory [18]. The α_R was not found to be a stable structure at both the B3LYP and SCC-DFTB levels of theory. It relaxes into a 3_{10} conformation for $n = 5$ and into a mixture of α_R and 3_{10} for $n = 8$, where only one $i \to i + 4$ H-bond pattern is formed in the middle of the chain, whereas the ends assume 3_{10} type conformations.

AM1 and PM3 show similar deficiencies as mentioned for the smaller compounds. One exception is the 3_{10} conformation. AM1 reproduces the structure for $n = 5$, 8 very well compared to B3LYP (deviations of only several degrees), while it predicts C_7 type configurations for smaller peptides. The extended C_5 conformations, however, do not improve for longer chains at the AM1 and PM3 levels of theory. PM3 again predicts a totally distorted structure. For both helix types, 3_{10} and α_R, the helices are unwound leading to H-bonds lengths of about 2.7 Å. For $n = 5$ and 8 AM1 relaxes the α_R into a configuration, which is in between the 3_{10} and α_R conformers. This structure shows both $i \to i + 3$ and $i \to i + 4$ H-bonding patterns, i.e., bifurcated H-bonds, where both H-bond lengths are about 2.3 and 2.4 Å, respectively [18].

6. Discussion

We calculated relative energies, structures and dipole moments for glycine and alanine based model peptides with up to eight residues at various levels of theory. For small peptides with one and two residues, we compared the SCC-DFTB, AM1 and PM3 methods with DFT methods and MP2 calculations. For longer peptides, we focussed on a comparison with the DFT-B3LYP method.

Although B3LYP relative energies and structures compare well with those evaluated at the MP2 level for various dipeptides, we found qualitative discrepancies for specific cases, e.g. for GDA and Ac-Gly$_2$-NHMe. First, B3LYP as well as HF predict a planar conformation for the C_5^{ext} conformer, while at the MP2 level the conformation is nonplanar. B3LYP contains the exact exchange energy, mixed with the exchange functional of Becke and the correlation functional of Lee, Parr and Yang (LYP). Here, it seems that B3LYP follows the tendencies of HF. However, the PES in this region is very shallow and the energy difference between planar and nonplanar conformation is in the order of 0.1 kcal/mol.

A similar trend is found with respect to the energetic ordering of the cyclic and linear conformers for this molecule. The relative stabilities evaluated at the HF and B3LYP level of theory are in qualitative disagreement with MP2 results. A more detailed test of DFT functionals is desirable. Further, the effects of internal BSSE have not been addressed in this study. Beachy et al. [15] found that BSSE related only to the correlation part (i.e., comparing MP2 with localized MP2, LMP2) favors the more compact structures compared the extended ones by about 0.5 kcal/mol for the diamide. For the triamide, this effect might be even enhanced. In order to draw conclusions about the relative stability of cyclic versus extended

structures, further investigations are necessary. For all other conformers B3LYP predicts geometries and relative energies in very good agreement with MP2 results. The structures presented here may therefore be a helpful test set to benchmark approximate methods, although the energetic ordering, especially of extended versus turn and helical like conformers, should be taken with some care. The energetic ordering within the subset of cyclic and helical conformers is described reliably with B3LYP, we found a good agreement of all ab initio methods compared so far. Higher level calculations like the 'MP4-BSSE' results for small model compounds help to interpret energetic trends for longer polypeptides evaluated with HF, DFT or MP2 methods, since the latter methods may show systematic deviations for the relative energies of certain secondary structure elements.

The SCC-DFTB method underestimates consistently energy differences between different conformers compared to not BSSE corrected ab initio results. Concerning the relative ordering of the conformers the SCC-DFTB follows the prediction of the DFT-B3LYP method. Taking the correction for BSSE into account, the comparison might be more favorable for the SCC-DFTB model as the comparison with the available 'MP4-BSSE' data suggest. Dipole moments are slightly (and consistently by about 0.5 D) smaller than B3LYP/6-31G* values, which is a satisfactory result, since dipole moments are not subject to any parametrization in the SCC-DFTB scheme. Structures are reliably reproduced. In particular, all secondary structure elements which were shown to be intrinsically stable at ab initio levels of theory are also stable at the SCC-DFTB method (with the only exception of the β type conformer in Ac-Ala-NHMe). The most severe failures occur for the extended conformations of glycine based peptides, where the pyramidalization at the nitrogen atom is overestimated. But we find qualitative differences already at the ab initio levels of theory and the PES is very shallow that energy differences in the order of 0.1 kcal/mol may lead to deviations in the dihedral values of 10–20°. Further, the relative energy of the C_7^{ax} conformer is underestimated. This might be of minor importance, since this conformation is seldom occurring in proteins.

The AM1 and PM3 methods show severe deficiencies in describing extended, helical and turn conformations. For many important secondary structure elements, these structures are either predicted to be very distorted (compared to the ab initio results) or are predicted to be intrinsically unstable. Especially turn structures seem to cause problems and the extended structures turn out to be very distorted. Further, 3_{10} and α helices are not described well. The PM3 method shows the tendency to unwind these helices by breaking the internal H-bonds. AM1, which is known to favor bifurcated H-bonds in hydrogen bonded complexes, seems to keep this feature also in this case. It predicts helices which are in Φ–Ψ space between the 3_{10} and α helical conformation by forming bifurcated H-bonds, i.e., $i \rightarrow i+3$ and $i \rightarrow i+4$ type H-bonding pattern. Therefore AM1 and PM3 should be used with care in applications for polypeptides and proteins. Although we compared the heats of formation from AM1 and PM3 to potential energies at the ab initio level, we do not think that inclusion of thermodynamic corrections at the ab initio level would make the comparison more favorable for AM1 and PM3.

To a certain extent, the performance of approximate methods can be evaluated from smaller molecules like GDA, ADA and Ac-Ala-NHMe, concerning geometries or relative energies for conformers which are stable for this peptide size. Errors inherent in these methods, occur already for the small compounds and seem to persist for the larger polypeptides. Therefore, these molecules serve as a good starting point for benchmarking approximate methods. However, the performance for more complicated secondary structure elements like β-turns or helices can only be evaluated by explicitly considering such conformations in larger peptide models. Recent developments of ab initio methods and increasing computer capacities make a high level treatment of larger systems and, therefore, a better evaluation of approximate methods possible. The discrepancies between different ab initio (HF,MP2) and DFT (B3LYP) methods for medium sized basis sets (6-31G*) show the need for higher level calculations, since systematic errors found for small molecules may add up when investigating longer polypeptides.

References

[1] J.S. Dewar, E. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.
[2] J.J.P. Stewart, J. Comp. Chem. 10 (1989) 209, 221.
[3] W. Yang, Phys. Rev. Lett. 66 (1991) 1438.
[4] W. Yang, T.-S. Lee, J. Chem. Phys. 103 (1996) 667.
[5] D.M. York, T.-S. Lee, W. Yang, J. Am. Chem. Soc. 118 (1996) 10940.
[6] J.J. Stewart, Int. J. Quant. Chem. 58 (1996) 133.
[7] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, Phys. Rev. B 58 (1998) 7260.
[8] M. Elstner, D. Porezag, G. Jungnickel, T. Frauenheim, S. Suhai, G. Seifert, Tight-binding approach to computational materials science, in: P. Turchi, A. Gonis, L. Colombo (Eds.), MRS Symp. Proc. no. 491, Materials Research Society, Pittsburgh, 1998, p. 131.
[9] M. Elstner, D. Porezag, T. Frauenheim, S. Suhai, G. Seifert, Multiscale modelling of materials, in: T. Diaz de la Rubia, T. Kaxiras, V. Bulatov, N.M. Ghoniem, R. Phillips (Eds.), MRS Symp. Proc. no. 538, Materials Research Society, Warrendale, PA 15086, 1999, p. 243.
[10] M. Elstner, T. Frauenheim, E. Kaxiras, G. Seifert, S. Suhai, Phys. Stat. Sol. (b) 217 (2000) 357.
[11] T. Head-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III, J.A. Pople, J. Am. Chem. Soc. 113 (1991) 5989.
[12] K.J. Jalkanen, S. Suhai, Chem. Phys. 208 (1996) 81.
[13] Z. Deng, P.L. Polavarapu, S.J. Ford, L. Hecht, L.D. Barron, C.S. Ewing, K.J. Jalkanen, S. Suhai, J. Chem. Phys. 100 (1996) 2025.
[14] K. Rommel-Möhle, H.-J.K. Hofmann, J. Mol. Struct. (Theochem) 285 (1993) 211.
[15] M.D. Beachy, D. Chasman, R.B. Murphy, T.A. Halgren, R.A. Friesner, J. Am. Chem. Soc. 119 (1997) 5908.
[16] W. Han, K. Jalkanen, M. Elstner, S. Suhai, J. Phys. Chem. B 102 (1998) 2587.
[17] C.L. Brooks III, D.A. Case, Chem. Rev. 93 (1993) 2487.
[18] M. Elstner, K. Jalkanen, M. Knapp-Mohammady, T. Frauenheim, S. Suhai, Chem. Phys. 256 (2000) 15.
[19] GAUSSIAN 98, Revision A.5, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian, Inc., Pittsburgh, PA, 1998.
[20] K. Möhle, M. Gussmann, A. Rost, R. Cimiraglia, H.-J. Hofmann, J. Phys. Chem. A 101 (1997) 8571.
[21] K. Möhle, M. Gussmann, H.-J. Hofmann, J. Comp. Chem. 18 (1997) 1415.
[22] H.G. Bohr, K.J. Jalkanen, M. Elstner, K. Frimand, S. Suhai, Chem. Phys. 246 (1999) 13.
[23] A. St-Amant, Density functional methods in biomolecular modeling, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 7, New York, 1996, p. 217.
[24] S. Saebo, W. Tong, P. Pulay, J. Chem. Phys. 98 (1993) 2170.
[25] A. Perczel, M.A. McAllister, P. Csaszar, I.C. Csizmadia, J. Am. Chem. Soc. 115 (1993) 4849.
[26] H.-J. Böhm, J. Am. Chem. Soc. 115 (1993) 6152.
[27] K. Möhle, H.-J. Hofmann, J. Mol. Model. 4 (1998) 53.
[28] C.M. Venkatachalam, Biopolymers 6 (1968) 1425.
[29] F. Herrman, S. Suhai, J. Comp. Chem. 11 (1995) 1434.
[30] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Fergusen, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179.