

Technical University of Denmark



A simple and realistic model system for studying hydrogen bonds in beta-sheets

Rossmeisl, Jan; Hinnemann, Berit; Jacobsen, Karsten Wedel; Nørskov, Jens Kehlet; Olsen, Ole H.; Pedersen, Jan T.

Published in:
Journal of Chemical Physics

Link to article, DOI:
[10.1063/1.1570395](https://doi.org/10.1063/1.1570395)

Publication date:
2003

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Rossmeisl, J., Hinnemann, B., Jacobsen, K. W., Nørskov, J. K., Olsen, O. H., & Pedersen, J. T. (2003). A simple and realistic model system for studying hydrogen bonds in beta-sheets. *Journal of Chemical Physics*, 118(21), 9783-9794. DOI: 10.1063/1.1570395

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

A simple and realistic model system for studying hydrogen bonds in β -sheets

Jan Rossmesl, Berit Hinnemann, Karsten W. Jacobsen, and Jens K. Nørskov^{a)}
Center for Atomic-scale Materials Physics (CAMP), Department of Physics, Building 307,
Technical University of Denmark, DK-2800 Lyngby, Denmark

Ole H. Olsen
Novo Nordisk Park, Novo Nordisk A/S, DK-2760 Måløv, Denmark

Jan T. Pedersen
H. Lundbeck A/S, Ottiliavej 9, DK-2500 Valby, Denmark

(Received 5 December 2002; accepted 6 March 2003)

We investigate the interaction between peptide chains at the level of state-of-the-art *ab initio* density functional theory. We propose an interacting periodic polypeptide model for studying the interactions in β -sheets and apply this to glycine and alanine peptide chains in both parallel and antiparallel structures. The calculated structures of alanine are compared to x-ray structures of β -sheets and the model is found to reproduce the geometry of the hydrogen bonds very well both concerning parallel and antiparallel β -sheets. We investigate the structures of both the N-H \cdots O=C and the C $_{\alpha}$ -H \cdots O=C hydrogen bonds. The former is thoroughly investigated, whereas the structure of the latter still is the subject of much discussion. We show that the hydrogen bonds between peptide chains are considerably weaker than what is found in studies of smaller models, e.g., the N-methylacetamide molecule. By molecular mechanics calculations we study the effect of twisting, which is not included in our model. We estimate its contribution to the interaction energy to be small. © 2003 American Institute of Physics. [DOI: 10.1063/1.1570395]

I. INTRODUCTION

The β -sheet is the second most common secondary structure element in proteins. Even though its structure was first proposed by Pauling and Corey in 1953,¹ there are still open questions about the nature of the interactions in β -sheets. Hydrogen bonds between amides of the backbones are known to play a key role in the stabilization of β -sheets in proteins.² Native β -sheets have been studied statistically using data from x-ray crystallography and evidence has been found for two types of hydrogen bond, the strong N-H \cdots O=C type and the weaker C $_{\alpha}$ -H \cdots O=C bonds.³⁻⁶

Previous theoretical investigations of β -sheets fall in two categories: Quantum mechanical studies have concentrated on small molecules whose structure was assumed to resemble the β -sheet structure to a certain degree.⁷⁻¹⁴ The other strategy has been to analyze larger and realistic β -sheet structures in proteins with force field methods.^{7,15} Thus there is either a payoff in method accuracy or in number of atoms in the model.

In the present paper we suggest a different approach by introducing a simple and realistic model: the interacting periodic polypeptide model (IPPM) for studying the interactions in β -sheets. The periodicity makes it possible to study extended systems by quantum mechanical methods. We investigate this model using *ab initio* density functional theory (DFT). In order to validate the model we present calculations

for alanine peptide chains and show that experimentally observed hydrogen bond geometries of β -sheets are very well reproduced.

Having verified the validity of IPPM, we concentrate on the two different hydrogen bonds present in β -sheets, namely the strong N-H \cdots O=C type and the weaker C-H \cdots O=C type. We investigate the nature of these bonds for peptide chains consisting of glycine (Gly) and alanine (Ala), in both parallel and antiparallel structures.

Furthermore, we present DFT calculations on the N-methylacetamide (NMA) molecule, a model system which previously has been used to describe the hydrogen bonds in β -sheets. We show that there are significant differences between results obtained with the small molecule and with IPPM. This comparison is important, as some parameters in force fields like, e.g., Amber94 are based on quantum mechanical calculations on small molecules.¹⁶ We calculate the structural and energetic properties of our model with the Amber94 force field and show that it does not account for key structural and energetic features. Furthermore, we use Amber94 to estimate the energy gained by twisting, a feature of β -sheets that cannot be described using IPPM. We do this by simulating extended and terminated chains.

II. METHODS

A. Molecular system

The IPPM setup for a β -sheet structure is depicted in Fig. 1 for the case of Gly. Figure 1 and all color figures have been prepared with MOLSCRIPT¹⁷ and with RASTER3D2.6.¹⁸

^{a)}Author to whom correspondence should be addressed. Electronic mail: nørskov@fysik.dtu.dk

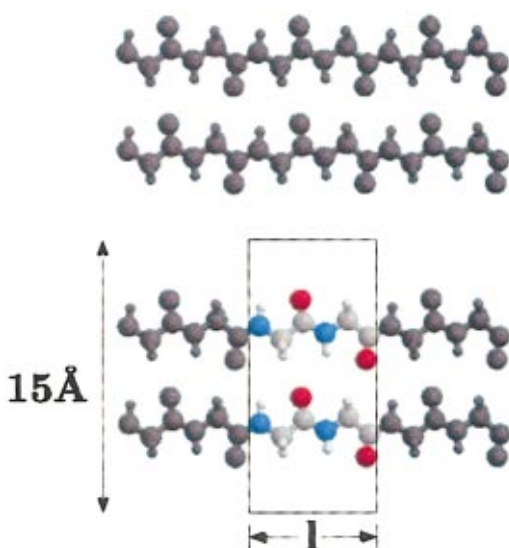


FIG. 1. (Color) The interacting periodic polypeptide model (IPP) for Gly. The colored part is one supercell. The supercell length is denoted by l and corresponds to the length of two residues.

The CPK color code (O: red, N: blue, C: gray, H: white) is used for all color figures. We consider periodically repeated peptide chains. The supercell consists of two peptide chains each with two residues; there is no terminating group as the one end of the supercell binds to the other. In this way the simulations are performed on infinitely long periodic interacting peptide chains. The x direction is chosen along the chains and the length of the supercell in the x direction is varied in order to determine the supercell length with the minimum energy. The supercell length along the x direction will be denoted by l in the following and is also defined in Fig. 1. The supercell length l is often referred to as the rise per two residues.² In the y and z directions, the supercell lengths are chosen sufficiently large to ensure that there are no interactions from one pair of chain to the next, i.e., 15 Å. The structure in IPPM resembles the β -pleated sheet structure originally proposed by Pauling and Corey.^{1,2} The model is a real β -sheet consistent of only a few different amino-acids, while this is not what is found in nature, the model gives a unique way of identifying the interactions associated with different amino acids. In this paper we only consider one species at a time, i.e., Gly or Ala. Another difference to naturally occurring β -sheets is that no twisting can be included as the period of a twist is much longer than the two residue unit used in this paper. In a later section, we investigate the energy gain by twisting with molecular mechanics and show that this is unimportant for the interactions. A periodic model of a polypeptide chain has previously been used by Improta *et al.* for the description of single peptide chains and helices.^{19,20} They have not used the model for the description of interacting peptide chains.

B. Density functional theory calculations

The density functional calculations^{21–23} presented here have been performed with the plane-wave pseudopotential

code DACAPO,²⁴ which uses a plane-wave expansion of the Kohn–Sham wave functions, and the generalized gradient approximation (PW91)²⁵ for the exchange–correlation term. The performance of the PW91 functional concerning nonbonded interactions between small molecules has been investigated previously,²⁶ and it was found that the interaction energies were more accurate than the ones obtained by other functionals. PW91 reproduced experimental binding energies for weakly bound systems with errors about 10%. A similar observation was made by²⁷ concerning the structure and energies of ice. All energies are calculated self-consistently. The main advantage of a plane-wave basis set is that convergence only depends on a single parameter, the cutoff energy, and can therefore be ensured by choosing a sufficiently large energy cutoff. In the present work we include plane waves with kinetic energies up to 25 Ry, and we have tested that the results are converged with respect to this basis set. Because of the plane-wave basis set, we use periodic supercells, and hence the model is computationally very efficient. Ultrasoft pseudopotentials have been used to describe scattering due to core electrons.²⁸ The Brillouin zone is sampled by four k points in the direction of the backbone chain, i.e., along the x axis. Fermi population of the Kohn–Sham orbitals with an electronic temperature of $k_B T = 0.01$ eV and Pulay mixing of the electron density have been used.²⁹ In order to find equilibrium structures, the atoms are relaxed according to their Hellman–Feynman forces until the sum of the absolute forces on all ions is less than 0.05 eV/Å.

C. Molecular mechanics calculations

All molecular mechanics calculations are carried out with the Molecular Modeling Toolkit (MMTK) using the Amber94 force field.^{16,30} Two models are used for the peptide chains: the first model is exactly the same as for the DFT calculations, which is depicted in Fig. 1. The second model does not use a periodic supercell, but includes terminated peptide chains in the vacuum. This model will be discussed later. All atoms are relaxed according to their forces, until the total absolute force is less than 0.05 eV/Å. For the periodic model, the minimum image convention has been used for determining the Lennard-Jones energy contribution. This introduces a small error for the energies, as the cutoff length, beyond which the Lennard-Jones energy contributions can be neglected, and has been tested to be about 10 Å, which is significantly larger than the smallest dimension of the unit cell. We have estimated this energy error to be smaller than 0.03 eV. The estimate was made by taking a particular configuration and calculating its energy in periodic systems with increasing supercell sizes, until the energy was converged. We emphasize, however, that for the calculation of the twisting energies, where we need high accuracy, the above-mentioned error has been corrected. This could be done by repeating the chain configuration and terminating it. By calculating the energy in a nonperiodic supercell for varying chain lengths, the correct energy per residue has been found. We note that this problem of course does not apply to the terminated chains in the extended model.

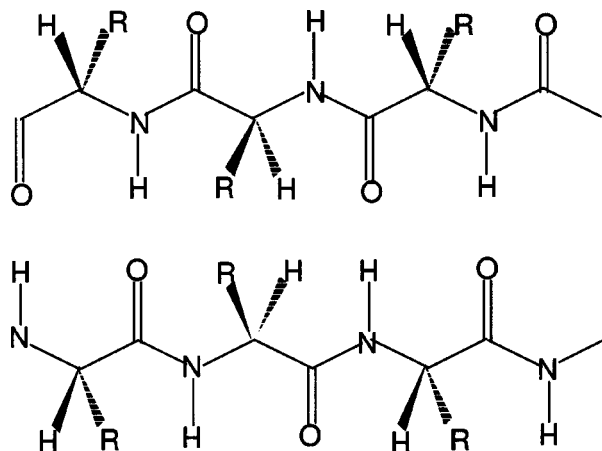


FIG. 2. Schematic drawing of the studied *antiparallel* configuration. There are two kinds of hydrogen bonds, the $C=O \cdots H-N$ and the $C=O \cdots H-C_\alpha$ hydrogen bonds. R represents the sidechain for the respective amino acid.

III. RESULTS AND DISCUSSION

A. DFT results for alanine chains

Due to its periodic geometry, IPPM calculations are very efficient. In this section we show that our model represents the local structure of a β -sheet, i.e., that experimental results for the local geometry of a β -sheet are reproduced. The available experimental data deduced from β -sheet structures are geometrical bond parameters from x-ray structures averaged over structures from different proteins. For a first study one has to choose some representative amino acids, as a calculation for all amino acids and averaging according to the probability of appearance is not possible in a first approach.

We choose to study the amino acid alanine (Ala), because its sidechain is small and without any special chemical features, e.g., charged or polar groups, and without special structural features like proline or cysteine. In addition we study the amino acid glycine (Gly). Gly is even simpler than Ala, but it has an unusually high conformational freedom

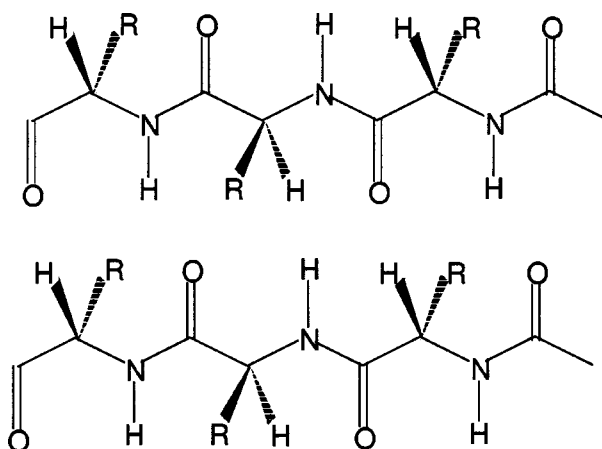


FIG. 3. Schematic drawing of the studied *parallel* configuration. There are two kinds of hydrogen bonds, the $C=O \cdots H-N$ and the $C=O \cdots H-C_\alpha$ hydrogen bonds. R represents the sidechain for the respective amino acid.

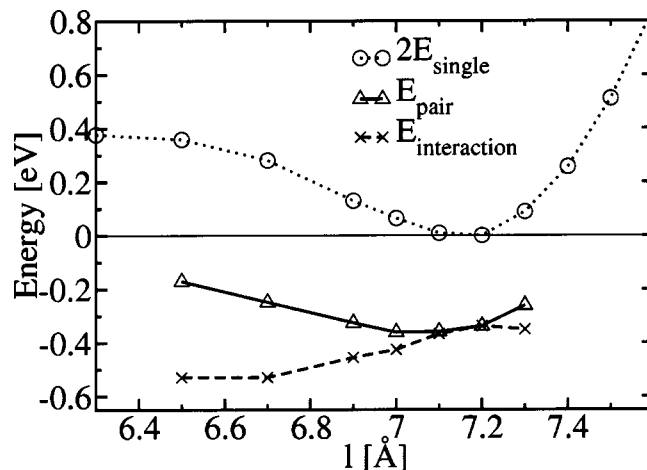


FIG. 4. The energies E_{single} , E_{pair} , and $E_{\text{interaction}}$ as a function of l for the *antiparallel* Ala-Ala chains.

compared to all other amino acids. We consider both the parallel and the antiparallel configurations, which are shown schematically in Figs. 2 and 3.

It should be noted that Ala is known to be a strong former of α helices, as well as a weak former of β -sheets.³¹ In IPPM the formation of an α helix is not possible, as the supercell consists of only two residues, however due to its simplicity Ala is an instructive model system for hydrogen bond formation in β -sheets.

The length of the two residue period l (see Fig. 1) is constant during a simulation and the energy dependence with respect to this degree of freedom for this reason has to be treated separately. Important insight into elastic and binding properties of the chains can be gained in this way. We investigate the energy dependence on the supercell length for a single chain ($E_{\text{single}}(l)$) and for the pair of interacting chains ($E_{\text{pair}}(l)$). The binding energy ΔE is defined as

$$\Delta E = E_{\text{pair}}(l_{\text{pair}}^{\text{min}}) - 2E_{\text{single}}(l_{\text{single}}^{\text{min}}). \quad (1)$$

In the energy diagrams Figs. 4 and 5 we have chosen the energy zero to be $2E_{\text{single}}(l_{\text{single}}^{\text{min}}) = 0$ so that ΔE is equal to $E_{\text{pair}}(l_{\text{pair}}^{\text{min}})$. As $l_{\text{single}}^{\text{min}}$ and $l_{\text{pair}}^{\text{min}}$ are not necessarily the same we separate the energy $E_{\text{pair}}(l)$ into two contributions:

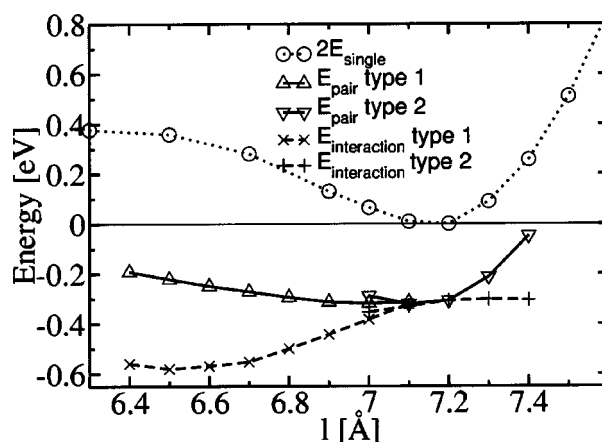


FIG. 5. The energies E_{single} , E_{pair} , and $E_{\text{interaction}}$ as a function of l for the *parallel* Ala-Ala chains. E_{pair} and $E_{\text{interaction}}$ are shown for both the type 1 and the type 2 structures.

$$E_{\text{pair}}(l) = 2E_{\text{single}}(l) + E_{\text{interaction}}(l), \quad (2)$$

where $E_{\text{single}}(l)$ is the energy of the single chain of length l . Since we have chosen $2E_{\text{single}}(l_{\text{single}}^{\text{min}}) = 0$, $E_{\text{interaction}}(l) \equiv E_{\text{pair}}(l) - 2E_{\text{single}}(l)$ is the interaction energy at a given l . A more negative $E_{\text{interaction}}(l)$ means a stronger bond.

For the antiparallel configuration these energies are shown in Fig. 4. It is seen that for a single chain the supercell length with the minimum energy $l_{\text{single}}^{\text{min}}$ is 7.2 Å and that the supercell length with the minimum energy for two interacting chains $l_{\text{pair}}^{\text{min}}$ is 7.0 Å. Generally, the structure of the interacting chains can be interpreted as a compromise between optimizing the hydrogen bonds and minimizing the elastic energy $E_{\text{single}}(l)$. The interaction energy becomes more negative when l decreases, as the hydrogen bond geometry can be optimized, but on the other hand the compression of the chains introduces a penalty in elastic energy E_{single} . Hence although the interaction energy is most favorable for supercell lengths below 6.8 Å, the penalty in elastic energy is too high for those structures to be favorable.

For the parallel structure the energy curves are shown in Fig. 5. Again the interactions are most favorable for supercell lengths below 6.8 Å, whereas the E_{single} is at a minimum at 7.2 Å. There is, however, an important difference to the antiparallel chains: the geometry of the structures at the two ends of the broad minimum is very different. As the chains are compressed, they undergo a “transition” from one structure to another. At the “transition” supercell length at 7.1 Å there are thus two different structures with the same energy. The rightmost curve in Fig. 5 marks the structures which have a low elastic energy, but a poor interaction (in the following referred to as “type 2”). The leftmost curve in Fig. 5 corresponds to structures for which the interactions are stronger but the elastic energy is higher (in the following referred to as “type 1”). The type 2 structures are stable for long supercells, whereas the type 1 structure is stable for shorter cells, as the long parallel chains are sterically unable to form optimal hydrogen bonds.

Let us now take a closer look at the structural differences between the type 1 and type 2 structures. In Fig. 6 the type 2 and type 1 structures are shown for a supercell length of $l = 7.1$ Å, where both structures are equally stable and with the same $E_{\text{interaction}}$ and E_{single} . The type 1 structure is represented by the atoms colored according to the CPK color scheme, whereas the type 2 structure is represented by the green atoms. The type 2 structure resembles very much the structure of a single Ala chain and only forms weak and relatively long hydrogen bonds. In order to improve the hydrogen bond, i.e., to go to the type 1 structure, the chains move the sidechains such that the backbones can come closer together. The C=O and the N-H group move slightly to form a tighter hydrogen bond, see Fig. 6. The difference between the two structures can also be seen in the dihedral angles in Table I, as the angles change significantly from the type 1 to the type 2 structure. The dihedral angles of the type 1 structure are also closer to experimentally observed values than for the type 2 structure.

The energy penalty related to the structural conversion from type 1 to type 2 can be illustrated by considering the

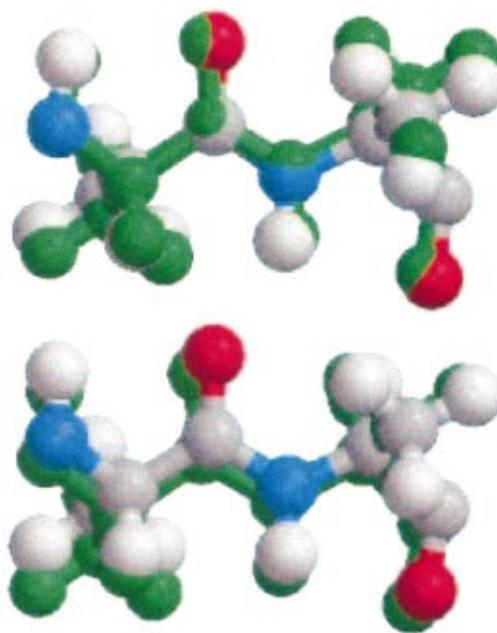


FIG. 6. (Color) The structures type 1 and type 2 for the *parallel* Ala–Ala chains at a supercell length $l = 7.1$ Å. The atoms colored according to the CPK color scheme belong to the type 1 structure, where the hydrogen bonds are close to the experimental geometry. The green atoms belong to the type 2 structure which forms only weak interactions.

energy penalty of the structural deformations. This is the energy difference between a single chain fixed in the structure of one of the interacting chains and a relaxed single chain. This calculation has been performed for both chains and the sum of the energies correspond to the reorganization energy. At the supercell length 7.1 Å, where the two structures are equally stable, the reorganization energy is 0.16 eV for type 1 and 0.03 eV for type 2, which shows that the type 1 structure undergoes reorganization but gains this energy by forming stronger hydrogen bonds, whereas type 2 neither deforms nor interacts strongly. It is interesting to notice that at $l = 6.5$ Å the reorganization energy for the type 1 structure is reduced to 0.03 eV, which means that at this length the chains do not need to modify the structure much in order to form the hydrogen bonds.

TABLE I. The dihedral angles in degrees for the parallel and antiparallel Ala–Ala chains. The numbering is such that A and B denote the two chains and that the residues which belong to angles with the same subscript face each other between the two chains.

	ϕ_1	ψ_1	ϕ_2	ψ_2
Antiparallel ($l = 7.0$ Å)				
Chain A	−147	150	−145	151
Chain B	−148	150	−145	152
Parallel type 1 ($l = 6.9$ Å)				
Chain A	−140	139	−137	142
Chain B	−136	141	−140	137
Parallel type 2 ($l = 7.1$ Å)				
Chain A	−155	159	−154	160
Chain B	−154	159	−155	160

TABLE II. Comparison between IPPM for Ala–Ala and experimental results from x-ray protein structures. The geometry for both types of hydrogen bonds for parallel and antiparallel structures. Each value occurs twice within one supercell. The computed values differ less than $<0.02 \text{ \AA}$ and $<2^\circ$ and thus the average is listed here.

	Antiparallel (expt.)	Antiparallel	Parallel (Expt.)	Parallel type 1	Parallel type 2
Length l	...	7.0 \AA	...	6.9 \AA	7.1 \AA
$\Delta E/2$...	-0.36 eV	...	-0.31 eV	-0.32 eV
NH \cdots OC bond					
NH \cdots OC	1.96 \pm 0.16 \AA^a	2.10 \AA	1.97 \pm 0.15 \AA^a	2.07 \AA	2.62 \AA
N \cdots O	2.91 \pm 0.14 \AA^a	3.07 \AA	2.92 \pm 0.14 \AA^a	3.03 \AA	3.49 \AA
\angle N–H \cdots O	160 \pm 10 $^\circ$ a	157 $^\circ$	161 \pm 9 $^\circ$ a	155 $^\circ$	143 $^\circ$
\angle C–O \cdots H	150 \pm 12 $^\circ$ a	161 $^\circ$	155 \pm 11 $^\circ$ a	150 $^\circ$	141 $^\circ$
CH \cdots OC bond					
CH \cdots O	2.44 \pm 0.14 \AA^b	2.35 \AA	2.41 \pm 0.14 \AA^b	2.20 \AA	2.25 \AA
	2.37 \pm 0.13 \AA^c		2.36 \pm 0.13 \AA^c		
\angle C–H \cdots O	141 \pm 6 $^\circ$ c	138 $^\circ$	146 \pm 6 $^\circ$ c	152 $^\circ$	141 $^\circ$
\angle C–O \cdots H	139.7 \pm 8.8 $^\circ$ b	136 $^\circ$	134.9 \pm 8.1 $^\circ$ b	143 $^\circ$	160 $^\circ$

^aFrom Ref. 3.

^bFrom Ref. 5.

^cFrom Ref. 6.

From Fig. 5 it is impossible to predict which of the structures occur under the conditions in real proteins, due to the small difference in energy.

In Table II we compare the DFT results for the Ala–Ala chains with results from statistical analysis of x-ray data from protein structures. For the N–H \cdots O=C bond we list the results from Ref. 3, which agree within the uncertainties with the results in Ref. 6. For the C–H \cdots O=C bond we list the results from Refs. 6 and 5. The error bars of all experimental results are large, as averages of the different amino acids are measured.

For the antiparallel chains, one can see that all geometrical features of the hydrogen bonds are very well described by the IPPM model, all bond lengths and angle agree with experiments within the experimental uncertainties. This suggests that IPPM mimics the hydrogen bonds in antiparallel chains very well. For the parallel chains, the type 1 structure compares equally well. The type 2 structure has much longer bond lengths and different angles compared to the experimental structure.

The conclusion of this section is that the hydrogen bond between both parallel and antiparallel chains is very well described in IPPM and that results obtained using this model would apply to natural β -sheets.

Comparing the binding energies ΔE for the parallel and antiparallel structures one notices that the binding is slightly stronger for the antiparallel structure. This is not because the interaction is generally weaker for the parallel chains. Actually the interaction can be stronger in short parallel chains. The minimum for $E_{\text{interaction}}$ is -0.58 eV for the parallel and -0.53 eV for the antiparallel. This is probably due to larger repulsion between the sidechains in antiparallel chains. The explanation for the stronger binding is rather that the antiparallel chains can have low elastic energy without a high penalty in interaction energy. This can also be seen in the reorganization energy. At $l=6.5 \text{ \AA}$ the penalty is the same 0.03 eV for both parallel and antiparallel structures. The antiparallel structure is fairly optimal also for longer chains (at $l=7.1 \text{ \AA}$ the penalty is only 0.07 eV compared to 0.16 eV for the parallel chains).

B. The nature of the two hydrogen bonds

The nature of the N–H \cdots O=C hydrogen bond has been intensively studied, and it is well established that it contains both an electrostatic and a quantum mechanical contribution.^{32,33} Furthermore, it has become accepted that the C=O \cdots H–C $_{\alpha}$ hydrogen bond exists and that it is an important force in secondary structure formation (Ref. 34, and references therein). This hydrogen bond has been studied theoretically^{11–14,35} and experimental evidence from x-ray structures exists.⁵ In this section, the C=O \cdots H–N and the C=O \cdots H–C $_{\alpha}$ hydrogen bonds are investigated through a study of the interaction-induced change in the electron density. Figures 7 and 8 depict the change of the electron density induced by the interaction between the chains, for the parallel (type 1) and the antiparallel Ala–Ala chains, respectively. A polarization of both kinds of hydrogen bonds is observed. One notices that the electron density distribution around the H $_{\alpha}$ atom is changed, and that the C $_{\alpha}$ –H $_{\alpha}$ bond becomes polarized along the direction defined by O and H $_{\alpha}$. The same is seen for C=O \cdots H–N where the polarization is along the N–H direction. The induced charge on, e.g., the H $_{\alpha}$, can be estimated to be of the order -0.05 to -0.10 of an electronic charge, which is a positive charge. It should be noted that this quantity cannot be uniquely determined. The supercell can be divided into four equally sized rectangular boxes each containing one peptide bond and the C $_{\alpha}$ atom on the carboxylic end of the peptide bond. The total induced charged on such a region is in the parallel case ± 0.020 electrons and in the antiparallel case ± 0.025 electrons. This could be interpreted as a transfer of charge between the interacting chains from the N–H and C–H groups to the C=O group in the hydrogen bond. The results are not very dependent on the exact position of the box. It is clear that this polarization of the chains must be important for the interaction. Apart from the induced static polarization shown here, there would also be a dynamic contribution giving rise to the van der Waals interactions, but this effect is not described explicitly in the present DFT, GGA calculations.

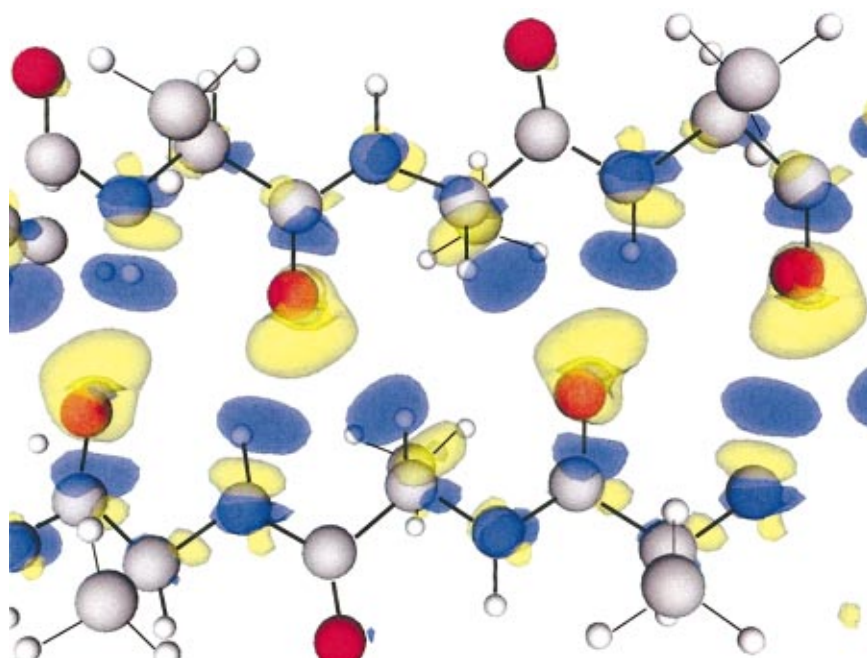


FIG. 7. (Color) Change in electron density induced by the interaction of two *antiparallel* Ala chains. The CPK color code is used for the atoms. Yellow regions refer to a concentration and blue regions to a depletion of electron density, relative to the isolated chain. The isosurfaces at are ± 0.007 eV/Å³.

C. Glycine–glycine chains

Here we present results for the interaction of antiparallel and parallel Gly chains. The results were obtained in the same way as the Ala results presented earlier. The Gly chain has additional structural freedom compared to other amino acids as the sidechains consist of only a hydrogen atom. Therefore it cannot be expected that its hydrogen bond geometry is similar to the experimental values, which are averages over all amino acids. However, modeling the Gly chain gives us a first estimate of the influence of the size of the sidechain. Another reason to analyze Gly chains is that Gly peptides have been used to model interactions in β -sheets and thus we can compare the results of IPPM to results in Ref. 36. The energies $E_{\text{total}}(l)$, $E_{\text{interaction}}(l)$, and

$E_{\text{single}}(l)$ are depicted in Fig. 9 for the antiparallel chains and in Fig. 10 for the parallel chains. See also Table III.

First we focus on the elastic energy, which is twice the energy of a single chain as before. Comparing E_{single} for Gly and Ala one notices that for small supercell lengths the curvature of E_{single} is smaller for Ala than for Gly. This means that it costs more energy to compress a Gly chain than an Ala chain, whereas the energy required to stretch a chain is approximately equal for both amino acids. Another difference is that the minimum of E_{single} for the Gly chain occurs at 7.25 Å whereas for the Ala chain it is located at approximately 7.15 Å.

The increased stiffness of the Gly chains seems at first glimpse counterintuitive as Gly is known to be more flexible

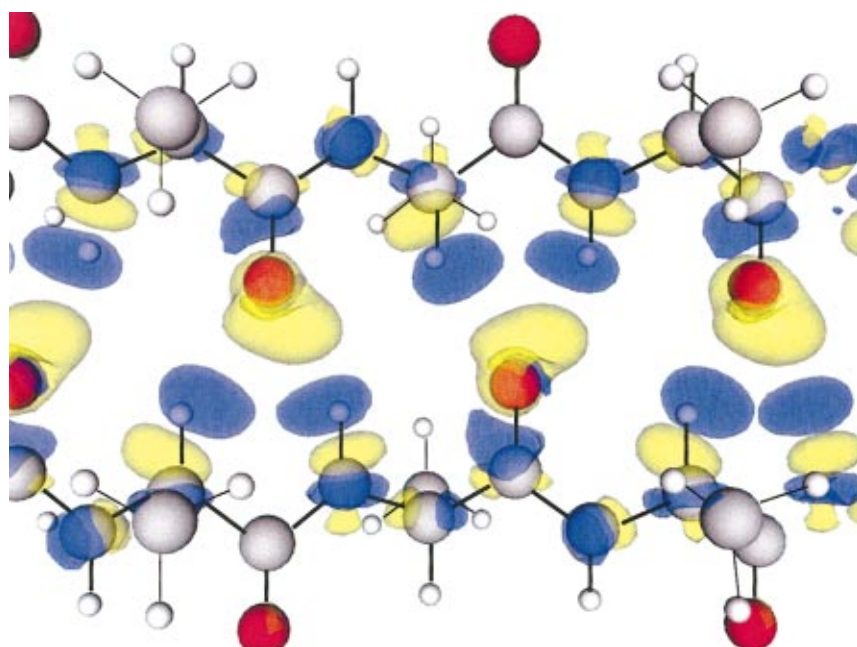


FIG. 8. (Color) Change in electron density induced by the interaction of two *parallel* Ala chains. The CPK color code is used for the atoms. Yellow regions refer to a concentration and blue regions to a depletion of electron density, relative to the isolated chain. The isosurfaces are at ± 0.007 eV/Å³.

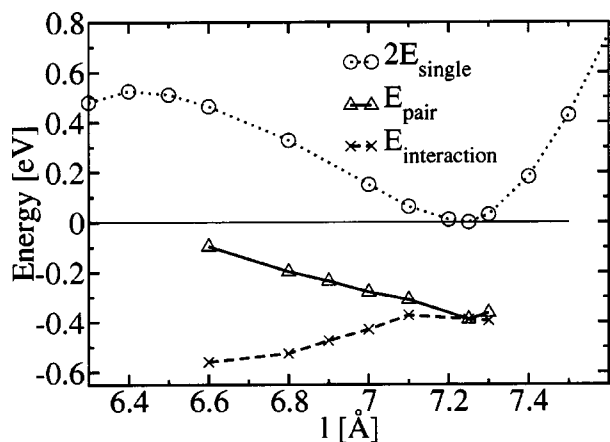


FIG. 9. The energies E_{single} , E_{pair} , and $E_{\text{interaction}}$ as a function of l for the antiparallel Gly-Gly chains.

in the sense that it has more sterically allowed regions than other amino acids. This flexibility allows the Gly chain to occupy a structure not allowed for the Ala backbone, and the result is that the chain is harder to compress. Another effect is that the Ala sidechain is less perturbed by the backbone for shorter chains than for the longer ones, making the Ala chain softer and shorter than the Gly chain. This is illustrated by calculations performed on the single Gly chain frozen in the backbone structure of Ala. In Fig. 11 E_{single} for the three cases is depicted: Gly chain relaxed, Gly chain frozen in the structure of Ala backbone, and Ala chain relaxed. The difference between the first two curves illustrates the effect of the flexible Gly backbone and the difference between the last two curves is related to the structural perturbation of the Ala sidechain. The zero energy point of the Ala chain is somewhat arbitrary, in this case it has been chosen such that the energy of the short Ala chain coincides with the energy for Gly frozen in the Ala backbone structure at the shortest l where the perturbation of the Ala sidechain is smallest. Or in other words: the energy of smallest perturbation of the Ala sidechain is chosen to be zero.

For the antiparallel Gly chains the binding energy ΔE is in the same range as for Ala, but $l_{\text{pair}}^{\text{min}}$ is different for the two

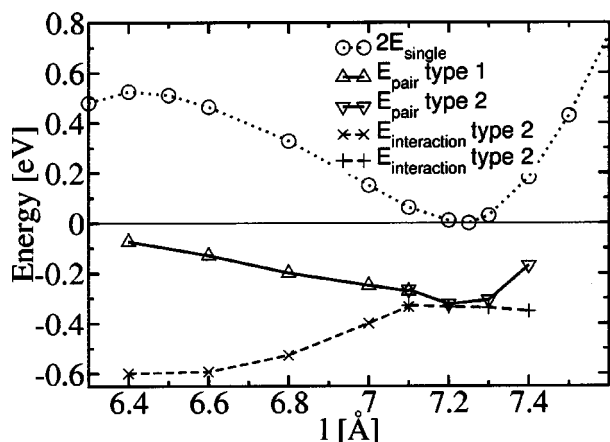


FIG. 10. The energies E_{single} , E_{pair} , and $E_{\text{interaction}}$ as a function of l for the parallel Gly-Gly chains. E_{pair} and $E_{\text{interaction}}$ are shown for both the type 1 and the type 2 structures.

TABLE III. The geometry for both types of hydrogen bonds and for parallel and antiparallel structures of the Gly chains. For the antiparallel structure, the geometry of the $\text{CH}\cdots\text{OC}$ bond alternates.

	Antiparallel	Parallel type 1	Parallel type 2
Length l	7.25 Å	7.0 Å	7.2 Å
ΔE	-0.39 eV	-0.25 eV	-0.32 eV
$\text{NH}\cdots\text{OC}$ bond			
$\text{NH}\cdots\text{OC}$	2.13 Å	2.18 Å	2.73 Å
$\text{N}\cdots\text{O}$	3.08 Å	3.12 Å	3.58 Å
$\angle\text{N}-\text{H}\cdots\text{O}$	154°	152°	140°
$\angle\text{C}-\text{O}\cdots\text{H}$	164°	145°	139°
$\text{CH}\cdots\text{OC}$ bond			
$\text{CH}\cdots\text{O}$	2.58/2.74 Å	2.17 Å	2.36 Å
$\angle\text{C}-\text{H}\cdots\text{O}$	114/102°	151°	126°
$\angle\text{C}-\text{O}\cdots\text{H}$	129/134°	144°	162°

systems, resulting in different structural features for the stable structure, e.g., the $\text{C}=\text{O}\cdots\text{H}-\text{C}_\alpha$ bond distance, where the Gly chain has a longer distance than Ala, corresponding to no significant $\text{C}=\text{O}\cdots\text{H}-\text{C}_\alpha$ bond for the stable Gly chains.

For the parallel Gly chains this increase in E_{single} renders the type 2 structure more stable than the type 1 structure, whereas for Ala the two structures were equally stable. This change occurs, because $E_{\text{interaction}}$ does not change notably from Ala to Gly. The binding energy ΔE for type 2 in Gly is the same as for Ala indicating that the small sidechains of Gly and Ala allow the chains to stay in a structure close in energy to the relaxed single chains. The structural changes between the type 2 and the type 1 structures are analogous to the parallel Ala-Ala chain: the sidechains move to the sides so that the backbones can come closer and form shorter hydrogen bonds. This can be clearly seen in the downshift of the dihedral angles in Table IV.

The conclusion is that the Gly chain lacks some of the features of Ala chains and β -sheets due to the missing sidechains. In both the parallel and antiparallel cases most of the difference between Gly and Ala has to do with the difference in E_{single} more than in $E_{\text{interaction}}$, i.e., the elastic properties of the isolated single chain strongly determines

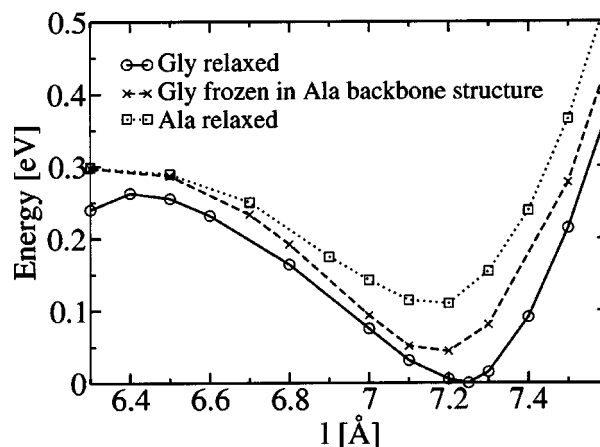


FIG. 11. The energy as a function of l for the single Gly chain, for the Gly chain frozen in the structure of Ala backbone, and for the single Ala chain.

TABLE IV. The dihedral angles in degrees for the parallel and antiparallel Gly–Gly chains. The numbering is such that A and B denote the two chains and that the residues which belong to angles with the same subscript face each other between the two chains.

	ϕ_1	ψ_1	ϕ_2	ψ_2
	Antiparallel ($l=7.25$ Å)			
Chain A	-177	179	-178	177
Chain B	-175	176	-175	176
	Parallel type 1 ($l=7.0$ Å)			
Chain A	145	-147	142	-149
Chain B	144	-149	144	-148
	Parallel type 2 ($l=7.2$ Å)			
Chain A	173	-174	173	-175
Chain B	171	-175	173	-173

the structure of the interacting pair of chains. One could speculate that E_{single} also determines features like β -sheet propensities of the amino acids.

The comparison to results obtained in previous quantum mechanical studies is somewhat complicated by the fact that the IPPM geometry and method used differs from the previous ones^{7,36} in several ways. We emphasize that IPPM reproduces experimental β -sheet structures without any constraints other than the periodicity. Furthermore the method is *ab initio* concerning both functional and basis set.

D. Small model molecules

Here we present calculations of the interaction between two N-methylacetamide (NMA) molecules both with density functional theory and with the molecular mechanics force field Amber94. Due to its resemblance with an amino acid residue, the interactions between two N-methylacetamide (NMA) molecules have been assumed to be very similar to the interactions between two peptide chains, and partly due to this the NMA dimer has been studied extensively both with force fields and quantum mechanical methods.^{13,37,38} Our focus is not to find the globally most stable structure of the *trans*-NMA dimer, but to estimate the strength of the hydrogen bond that resembles the N–H···O=C bond in β -sheets. We start with two structures, a parallel and an antiparallel, and let them relax to their local energy minima. The structure of the parallel *trans*-NMA dimer is depicted in Fig. 12. There is, however, an important difference between an NMA dimer and a peptide chain, namely that the NMA molecules can reorient freely in order to create optimal N–H···O=C hydrogen bonds, whereas a peptide chain within a β -sheet is subject to constraints. Thus one would expect that the binding energy between two *trans* N-methylacetamide molecules is larger than between two peptide chains, as energy can be gained by reorientation.

The DFT binding energies for the *trans*-NMA dimer shown in Table V are indeed much higher than the binding energies found for IPPM. Notice that it is the binding energy per residue, i.e., $\Delta E/2$ which is always less than 0.2 eV, that should be compared to the results in Table V. The NMA binding energies could be seen as the upper limit for the energy

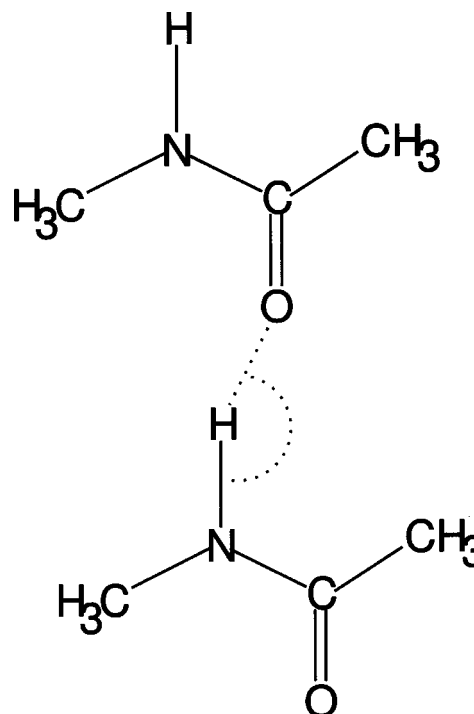


FIG. 12. Hydrogen bond between two molecules of *trans*-N-methylacetamide (NMA). The dashed line indicates the hydrogen bond and the dashed angle denotes the N–H–O angle.

gained by forming a N–H···O=C bond. But used as a model of the interactions in proteins the NMA dimer is clearly too small and too flexible. The DFT results in Table V give similar energies compared to other quantum mechanical methods^{13,38} even though the equilibrium structures are somewhat different.

The Amber94 binding energies and geometry for the *trans*-NMA dimer agree reasonably with the DFT results, and with those of a recent paper by Buck and Karplus.³⁷ They also find a minimum binding distance of 1.9 Å for *trans*-NMA.

An interesting fact is that the *trans*-NMA dimer has been used to calculate some of the partial charges in force fields.^{16,39} Complementary to fitting force field parameters to systems containing hydrogen bonds, as done for Amber94, there have been attempts to account for hydrogen bonds with a separate energy term.^{40,41} As we in the following present Amber94 results for our β -sheet model, the results of this section will help one to understand the differences between the density functional and the molecular mechanics results.

TABLE V. The interaction energy and the N–H···O=C distance for the *trans*-N-methylacetamide dimer calculated with different approaches.

	Parallel		Antiparallel	
	DFT	Amber94	DFT	Amber94
ΔE	-0.39 eV	-0.38 eV	-0.36 eV	-0.36 eV
N–H···O=C	1.94 Å	1.90 Å	2.05 Å	1.88 Å
\angle N–H–O	178°	162°	174°	173°

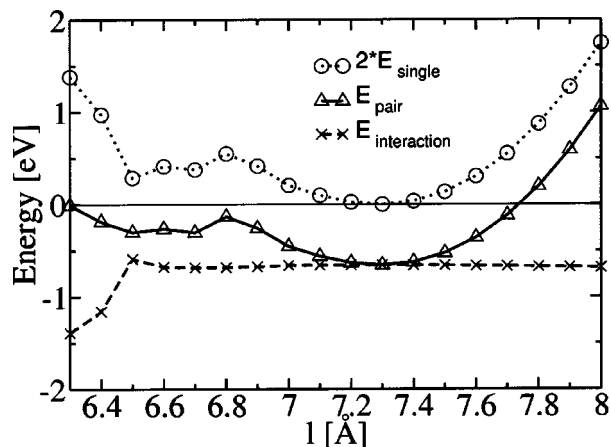


FIG. 13. The energies E_{single} , E_{pair} , and $E_{\text{interaction}}$ as a function of l for the antiparallel Ala-Ala chains from the molecular mechanics calculations.

E. Force field calculations on interacting periodic polypeptide model

One could pose the question, how many of the DFT results using the IPPM could have been obtained using a molecular mechanics force field, e.g., Amber94, which would be much faster. Force fields are extensively used for describing polypeptides or proteins, e.g., for protein dynamics and folding problems. Therefore it is interesting to know, how many features of the DFT characterization of our model can be obtained using a force field. Thus, in this section we present calculations of the interaction between two Ala-Ala chains obtained using Amber94.

The total binding energy between the two peptide chains is separated in the same way as for the DFT calculations, i.e., according to Eqs. (1) and (2). For the case of the Ala-Ala chains, these energies are plotted in Fig. 13 for the antiparallel configuration and in Fig. 14 for the parallel configuration. The first thing to notice is that the binding energies are significantly larger than the DFT results obtained for the same systems. The binding energy is found to be almost as large as the energies for the NMA dimer.

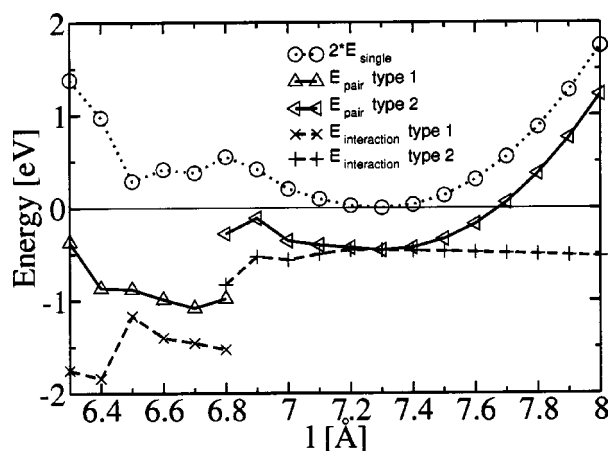


FIG. 14. The energies E_{single} , E_{pair} , and $E_{\text{interaction}}$ as a function of l , for the parallel Ala-Ala chains from the molecular mechanics calculations. E_{pair} and $E_{\text{interaction}}$ are shown for both the type 1 and the type 2 structures.

TABLE VI. The hydrogen bond geometry for the antiparallel and parallel Ala-Ala chains from molecular mechanics Amber94 results. Each interaction occurs two times in the supercell. For the parallel type 2 configurations those values differ significantly and are both indicated. For all other configurations these values differ very little (<0.02 Å and $<1^\circ$) and they are averaged. ΔE is the interaction energy per residue.

	Antiparallel	Parallel type 1	Parallel type 2
Length l	7.3 Å	6.7 Å	7.3 Å
ΔE	-0.66 eV	-1.08 eV	-0.46 eV
NH \cdots OC bond			
NH \cdots OC	1.95 Å	1.88 Å	2.44 Å/2.67 Å
N \cdots O	2.93 Å	2.87 Å	3.32 Å/3.56 Å
\angle N-H \cdots O	162°	164°	146°/146°
\angle C-O \cdots H	161°	142°	146°/138°
CH \cdots OC bond			
CH \cdots OC	2.52 Å	2.45 Å	2.36 Å/2.36 Å
\angle C-H \cdots O	126°	140°	132°/140°
\angle C-O \cdots H	126°	113°	154°/152°

Beachy *et al.* have performed quantum mechanical studies of di- and tetra-alanine dimers and shown that most force fields significantly overestimate the interaction energy between alanine dipeptides.⁷ According to this study, most force fields even estimate the interaction energy between alanine dipeptides to be stronger than the one in a *cis*-N-methylacetamide dimer. The authors argue that the reason for the poor performance of the force fields is that the atom-centered partial charge approach fails to describe the electrostatic interactions correctly. The reason could be that atom centered partial charges always prefer the more linear bond in β -sheets over the bond in a *cis*-N-methylacetamide dimer. But in the case of the *trans*-NMA dimer the hydrogen bond is very linear and this argument does not really explain the too strong β -sheet structures compared to the *trans*-NMA dimer. This example shows that there is a need for more realistic models of the β -sheet structure, as they cannot be modeled correctly by small molecules with the same type of interactions.

The case of the antiparallel Ala-Ala chain is straightforward to interpret; over a large range of l (~ 6.8 – 8 Å) the interaction energy is almost constant, and the total energy E_{pair} is only determined by E_{single} . Thus the minimum of E_{pair} at a supercell length of $l_{\text{min}} = 7.3$ Å is only caused by the

TABLE VII. The dihedral angles in degrees for the parallel and antiparallel Ala-Ala chains. The numbering is such that A and B denote the two chains and that the residues which belong to angles with the same subscript face each other between the two chains.

	ϕ_1	ψ_1	ϕ_2	ψ_2
	Antiparallel ($l=7.3$ Å)			
Chain A	-154	161	-153	162
Chain B	-154	161	-153	162
	Parallel type 1 ($l=6.7$ Å)			
Chain A	-82	146	-152	84
Chain B	-152	85	-82	146
	Parallel type 2 ($l=7.3$ Å)			
Chain A	-154	165	-152	164
Chain B	-153	162	-154	164

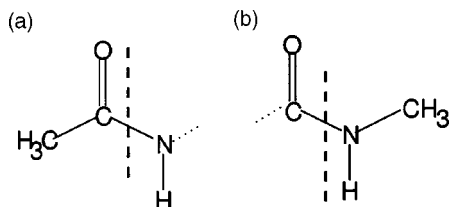


FIG. 15. Termination molecules for the extended chains. To the left (a) is the termination for the N-end and to the right (b) the termination for the C-end.

minimum in E_{single} . The hydrogen bond geometries and the binding energy for $l_{\text{min}}=7.3 \text{ \AA}$ are listed in Table VI. It should be noted that the decrease in interaction energy for l smaller than 6.5 \AA corresponds to structures which do not resemble a β -sheet any longer, as the chains start to curl up and interact internally. In this study we are not interested in this regime.

The case of the parallel Ala–Ala chains (depicted in Fig. 14) is more complicated, in analogy with the DFT results two structures with different hydrogen bond geometry emerge. The type 2 structure dominates for supercell lengths larger than 6.8 \AA and is characterized by a constant and comparatively small interaction energy. The local minimum in this region, which is located at $l_{\text{min}}=7.3 \text{ \AA}$, is exclusively caused by the minimum in E_{single} . The type 1 structure dominates for supercell lengths below 6.8 \AA with a minimum at 6.7 \AA . The minimum at 6.4 \AA is not considered here, as the structure does not resemble a β -sheet anymore. The supercell length 6.8 \AA is the only point where both structures are stable. At this point the starting configuration determines the relaxed structure. For any other supercell length, the result is independent of the initial configuration. The structural changes between the type 1 and the type 2 structures are similar to the ones for the DFT results shown in Fig. 6. The sidechains move in such a way that the chains can approach each other. Regarding the hydrogen bond geometry, we choose to analyze the structures for the supercell lengths 7.3 \AA (the type 2 structure) and 6.7 \AA (the type 1 structure)

which are both local minima of E_{pair} . The results are shown in Table VI.

Comparing the hydrogen bond geometry for the Ala–Ala chains in Table VI to the experimental values listed in Table II one notes that the values for the antiparallel Ala chains compare well. Also the dihedral angles, which are listed in Table VII, seem reasonable and are in the range of the typical dihedral angles in an antiparallel β -sheet. The energy plot of the antiparallel Ala chains Fig. 13 resembles very much the DFT plot and there are no discontinuities. For the parallel Ala chains, however, none of the two structures in Table VI compares well with the experimental structure. The type 2 structure has too weak interactions and corresponds structurally to two noninteracting single chains. The type 1 structure shows reasonable binding distances, but the angles do not match. Also the dihedral angles in Table VII are outside the range of typical dihedral angles for a β -sheet. This shows that the compressed chains do not model a β -sheet well.

This is an interesting conclusion, since the force fields are designed to give the experimentally observed structures. The fact that they do not do this in the present setup is a strong indication that the potentials are not robust enough to handle even the very mild constraints set up by the periodicity. Generally, force fields are optimized to ensure fast convergence to the equilibrium structure and are not optimized to describe nonequilibrium structures.

Finally, we observe in Fig. 14 that there are discontinuities, and the structures change dramatically, whereas this change is much smoother in DFT. There are even discontinuities in E_{single} , which are completely absent in DFT. In that regime, where the chains are compressed, the force field Amber94 is not suitable for describing the chains. One cannot help wondering if such unphysical behavior could be a problem when describing, e.g., the folding process, which could involve a local compression of the polypeptide chains.

We have performed an analogous investigation for antiparallel and parallel Gly chains. We observe the same kind of discontinuities in energies and structures as for the Ala chains, and therefore we do not reiterate this discussion. The

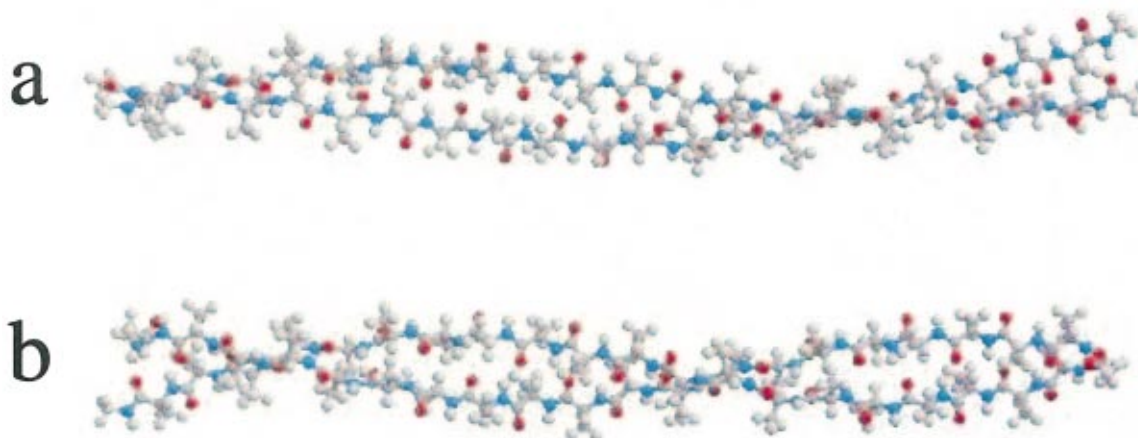


FIG. 16. (Color) Upper (a): Two extended antiparallel interacting Ala chains. Lower (b): Two extended parallel interacting Ala chains. Both chains consist of 20 residues and the termination at both ends.

TABLE VIII. The energies gained by twisting for the Ala chains. All energies are per double residue. As the chains gain energy by twisting, all energies are negative.

Chain	Twist energy (eV)
Single Ala	-0.06
Antiparallel Ala-Ala	-0.08
Parallel Ala-Ala	-0.15

conclusion of this section is that Amber94 does not describe the periodic polypeptides well, with the possible exception of the antiparallel Ala-Ala chains. There are discontinuities in the energy curves which lead to the suspicion that Amber94 may not capture the chemistry of the peptide chain well enough.

F. Molecular mechanics calculations on twisting

Molecular mechanics offers the possibility to study extended force fields, as much larger systems can be handled than for *ab initio* methods. Here we study the influence of twisting, which is not included in the IPPM, by using a system of two extended and terminated peptide chains. The extended chains are constructed by repeating a unit containing two residues and by terminating them in a suitable way. The termination is depicted in Fig. 15. Simulations are performed for different chain lengths in order to determine the interaction energy per residue eliminating the effect of the terminations. The relaxed structures of the antiparallel and parallel Ala chains each consisting of 20 residues are shown in Fig. 16. One can see that the chains twist right-handedly, as has been known for a long time for β -sheets.^{10,42} A single Ala chain also twists slightly. The Gly chains do not twist—neither the parallel nor the antiparallel chains.

The extended chains can be used to estimate the energy contributions of twisting, which are not included in the IPPM. This is done by determining the energy difference between two interacting chains in the IPPM (which by definition are not twisted) and two interacting chains in the extended model. We emphasize that the energies obtained in the periodic model are corrected for the aforementioned small energy error. The energies gained by twisting are shown in Table VIII.

From those energies we can infer that the calculated interaction energies of the IPPM actually are quite accurate, as the energy deviations from the pair chains and the single chains cancel each other. For the antiparallel chains the energy deviation caused by twisting is -0.04 eV (energy for double chain minus twice the energy for single chain). For the parallel chains the energy deviation is 0.03 eV. Thus we make only very small errors in the energy by not including twisting. In Sec. III A we furthermore have seen by comparison to experiments that the geometry of the IPPM configurations also agrees well with the experimental findings.

IV. CONCLUSION

In conclusion we have proposed an interacting periodic polypeptide model (IPPM) for studying interactions in

β -sheets. We use the model to perform *ab initio* calculations with the PW91 exchange-correlation functional. We compared calculated structures of parallel and antiparallel alanine sheets to x-ray structures. We find that the geometry of the hydrogen bonds is well reproduced by our model. We emphasize that our model does not contain any constraints other than the periodicity. We have shown that smaller models are not suitable for investigating the interactions in β -sheets, as they predict considerably stronger interactions. By means of molecular mechanics calculations we estimate the effect of twisting and find it to be small.

- ¹L. C. Pauling and R. B. Corey, Proc. R. Soc. London, Ser. B **141**, 21 (1953).
- ²G. E. Schulz and R. H. Schirmer, *Principles of Protein Structure* (Springer, New York, 1979).
- ³E. N. Baker and R. E. Hubbard, Prog. Biophys. Mol. Biol. **44**, 97 (1984).
- ⁴R. Taylor and O. Kennard, J. Am. Chem. Soc. **104**, 3063 (1982).
- ⁵Z. S. Derewenda, L. Lee, and U. Derewenda, J. Mol. Biol. **252**, 248 (1995).
- ⁶G. F. Fabiola, S. Krishnaswamy, V. Nagarajan, and V. Pattabhi, Acta Crystallogr., Sect. D: Biol. Crystallogr. **53**, 316 (1997).
- ⁷M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, and R. A. Friesner, J. Am. Chem. Soc. **119**, 5908 (1997).
- ⁸N. Gresh, H. Guo, D. R. Salahub, B. P. Roques, and S. A. Kafafi, J. Am. Chem. Soc. **121**, 7885 (1999).
- ⁹J. K. Kang, J. Mol. Struct.: THEOCHEM **546**, 183 (2001).
- ¹⁰I. L. Shamovsky, G. M. Ross, and R. J. Riopelle, J. Phys. Chem. B **104**, 11296 (2000).
- ¹¹S. Scheiner, T. Karr, and Y. Gu, J. Biol. Chem. **276**, 9832 (2001).
- ¹²R. Vargas, J. Garza, D. A. Dixon, and B. P. Hay, J. Am. Chem. Soc. **122**, 4750 (2000).
- ¹³R. Vargas, J. Garza, R. A. Friesner, H. Stern, B. P. Hay, and D. A. Dixon, J. Phys. Chem. A **105**, 4963 (2001).
- ¹⁴R. Vargas, J. Garza, B. P. Hay, and D. A. Dixon, J. Phys. Chem. A **106**, 3213 (2002).
- ¹⁵A. Yang and B. Honig, J. Mol. Biol. **252**, 366 (1995).
- ¹⁶W. D. Cornell, P. Cieplak, C. I. Bayly *et al.*, J. Am. Chem. Soc. **117**, 5179 (1995).
- ¹⁷P. J. Kraulis, J. Appl. Crystallogr. **24**, 946 (1991).
- ¹⁸E. A. Merrit and D. J. Bacon, Methods Enzymol. **277**, 505 (1997).
- ¹⁹R. Improtta, W. Barone, K. N. Kudin, and G. E. Scuseria, J. Chem. Phys. **114**, 2541 (2001).
- ²⁰R. Improtta, W. Barone, K. N. Kudin, and G. E. Scuseria, J. Am. Chem. Soc. **123**, 3311 (2001).
- ²¹P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
- ²²W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- ²³R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- ²⁴The DFT calculations were performed with the DACAPO pseudopotential code, which is available at <http://www.fysik.dtu.dk/campus>
- ²⁵J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
- ²⁶S. Tsuzuki and H. P. Lü Thi, J. Chem. Phys. **114**, 3949 (2000).
- ²⁷D. R. Hamann, Phys. Rev. B **55**, 10157 (1997).
- ²⁸D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
- ²⁹G. Kresse and J. Furthmüller, J. Comput. Mater. Sci. **6**, 15 (1996).
- ³⁰K. Hinsen, J. Comput. Chem. **21**, 79 (2000).
- ³¹P. Y. Chou and G. D. Fasman, Biochemistry **13**, 222 (1974).
- ³²G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer, Berlin, 1991).
- ³³S. Scheiner, *Hydrogen Bonding: A Theoretical Perspective* (Oxford University Press, New York, 1997).
- ³⁴G. A. Jeffrey, J. Mol. Struct. **485**, 293 (1999).
- ³⁵Y. Gu, T. Karr, and S. Scheiner, J. Am. Chem. Soc. **121**, 9411 (1999).
- ³⁶Y.-L. Zhao and Y.-D. Wu, J. Am. Chem. Soc. **124**, 1570 (2002).
- ³⁷M. Buck and M. Karplus, J. Phys. Chem. B **105**, 11000 (2001).

- ³⁶D. A. Dixon, K. D. Dobbs, and J. J. Valentini, *J. Phys. Chem.* **98**, 13435 (1994).
- ³⁷C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, *J. Phys. Chem.* **97**, 10269 (1993).

- ³⁸P. K. C. Paul and C. Ramakrishnan, *J. Biomol. Struct. Dyn.* **2**, 879 (1985).
- ³⁹J. R. Damewood, Jr., R. A. Kumpf, W. C. F. Mühlbauer, R. J. Urban, and J. E. Eksterowicz, *J. Phys. Chem.* **94**, 6619 (1990).
- ⁴⁰C. Clothia, *J. Mol. Biol.* **75**, 295 (1973).