Folding – Unfolding Equilibrium of a Methylidene-Substituted β -Peptide

by Nathan Schmid, Bojan Zagrovic, and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, ETH, Swiss Federal Institute of Technology, CH-8093 Zürich (phone: +41446325501; fax: +4144632 1039; e-mail: wfvgn@igc.phys.chem.ethz.ch)

Dedicated to Professor Dieter Seebach on the occasion of his 70th birthday

 β -Peptides possess the ability to fold into secondary structure elements, and this property, together with resistance to biodegradation, makes these compounds interesting for pharmaceutical applications. Recently, a novel class of β -peptides containing methylidene moieties was described. The GROMOS 53A6 force field was used to simulate the folding equilibrium of a β^3 -hexapeptide with methylidene (CH₂=) groups at all six CA-atoms. Due to the rotational barriers induced by these methylidene groups, the helical secondary-structure elements, normally found in β^3 -peptides, are disfavored in this molecule. Simulations, started from fully extended and 3₁₄-helical conformations, showed that the molecule adopts a complete 2₈-helix for *ca.* 5% of the time and partial 2₈-helical conformations for *ca.* 20% of the time. Yet, as suggested by experiments, the folding equilibrium is dominated by unfolded conformations.

Introduction. – Foldamers are a class of non-biological molecules which have the ability to adopt secondary and/or tertiary structures defined by non-covalent interactions [1-3]. For β -peptides, which belong to this class of molecules, it was shown that even short oligopeptides (less than six residues) fold into different secondary-structure elements such as the pleated sheet [4] or the 3_{14} -helix [5]. In particular, the 3_{14} -helix seems to be one of the most dominant secondary structure elements in β -peptides. Due to the additional C-atom in the backbone of β -peptides compared to α -peptides, these compounds are remarkably resistant to proteolytic degradation [6]. Further, some of them have the ability to translocate across the cell membrane [7]. These properties make β -peptides promising candidates for pharmaceutical applications such as to function as antibacterial or antimicrobial agents [8], or as inhibitors of fat and cholesterol absorption [9].

More recently, an interesting type of β -peptides, namely those built from (*S*)-2methylidene-3-aminoalkanoic acids, have been synthesized [10]. In these molecules, due to the presence of the methylidene group at the C_A-atom (*Fig. 1*), the rotational barriers of the adjacent backbone torsional angles are different compared to nonsubstituted β^3 -peptides. It was expected from previous structural work that Boc-[β^3 -HVal(=CH₂)- β^3 -HAla(=CH₂)- β^3 -HLeu(=CH₂)]₂-OMe (**1**) folds into a β_{14} -helix, as non-substituted β -peptides would do. However, the preparation of crystals failed, and no significant NOE signals indicative of a stable secondary structure were observed. Consequently, it was concluded that this molecule is conformationally rather unstructured and does not fold into a β_{14} -helix [10].

The GROMOS force field was used previously to successfully reproduce the folding equilibria of different β -peptides by molecular-dynamics (MD) simulation [11–14].

^{© 2007} Verlag Helvetica Chimica Acta AG, Zürich



Fig. 1. The topology of **1**. The residues carry a methylidene group at C_A (C_A , C_B , C_G , C_D , C_E , and C_M correspond to the C-atoms $C(\alpha)$, $C(\beta)$, $C(\gamma)$, $C(\delta)$, $C(\varepsilon)$, and methylidene C-atoms, respectively). The carboxy end is protected by a Me group, the amino end by a (*tert*-butoxy)carbonyl (Boc) group.

The observation that **1** does not form any detectable secondary structure encouraged us to carry out a MD study of the molecule. First, the rotational profiles of the C_M-C_A-C-O dihedral angle of an unsubstituted and a methylidene-substituted β amino acid peptide were explored by using two simple models. The findings already predicted that a 3_{14} -helical secondary structure of **1** is rather unlikely. Second, two simulations with different starting structures, extended and 3_{14} -helical, were carried out to study the folding equilibrium. It was found that the molecule does fold into a helical geometry, but that this helix is of the 2_8 type, rather than 3_{14} . However, this helix is not very stable, the molecule being conformationally rather unstructured most (70%) of the time.

Results and Discussion. – In spite of intensive efforts, the secondary structure of **1** could not be determined by NMR spectroscopy or X-ray crystallography [10]. This has been considered as an indication that the molecule does not fold into any well-defined secondary structure. The search for molecules containing α -methylidene β -amino acids in the Cambridge crystallographic database revealed that the s-*trans*-conformation of the C_M-C_A-C-O torsional angle is preferred over s-*cis* (s-*cis/s*-*trans ca.* 1:4). This conformation combined with a planar amide group is not compatible with a β_{14} -helical or β -peptidic pleated-sheet structure [10], suggesting an explanation for the apparent lack of stable secondary structure in the peptide.

First, to investigate the features of the force field used, two simple models consisting of only four heavy (non-H) atoms were explored (*Fig.* 2). The first model was built of a Me group, two C-atoms and a carbonyl O-atom, mimicking one of the backbone elements of a β -peptide. The parameters for the *Van der Waals* interactions of the Me group and the carbonyl O-atom, and, for the bonds, the angles, and the dihedral-angle potential energy terms were taken from the 53A6 GROMOS force field [15]. This GROMOS parameter set has been tested in a number of different simulations [3][15–17]. The Me group was then rotated around the axis defined by the two backbone C-atoms (C_A and C), and the energy was calculated in steps of 0.01°, according to the following equation [18][19].

$$V_{\text{tot}}(\mathbf{r}_{C_{\text{M}}}, \mathbf{r}_{C_{\text{A}}}, \mathbf{r}_{\text{C}}, \mathbf{r}_{\text{O}}) = V_{LJ}(\mathbf{r}_{C_{\text{M}}}, \mathbf{r}_{C_{\text{A}}}, \mathbf{r}_{\text{C}}, \mathbf{r}_{\text{O}}) + V_{\text{dih}}(\mathbf{r}_{C_{\text{M}}}, \mathbf{r}_{C_{\text{A}}}, \mathbf{r}_{\text{C}}, \mathbf{r}_{\text{O}})$$
(1)



Fig. 2. Torsional energy profiles for rotation around the $C_A - C$ bond for a) methyl and b) methylidene β peptide. Torsional-angle potential energy: normal line, Lennard–Jones potential energy: dotted line, potential energy sum: bold line. In the methylidene case, the torsional-angle barrier separating the planar conformations is high due to resonance effects originating in the two double bonds.

1968

The bonds and bond angles were kept fixed. It is seen in *Fig. 2* that the main contributor to the overall potential energy is the *Lennard–Jones (Van der Waals)* term, which is seriously disfavoring the fully eclipsed conformation. The latter conformation is destabilized by more than 20 kJ mol⁻¹ compared to the s-*trans*-conformation. The second model is obtained by replacing the Me group by a methylidene group. In this case, the torsional dihedral-angle term is disfavoring the $\pm 90^{\circ}$ conformations due to resonance of the two double bonds. The *Lennard–Jones* energy raises the s-*cis*-minimum (torsional angle 0°) by 12 kJ mol⁻¹ compared to the s-*trans*-conformation.

Second, two simulations of 1, one starting from the fully extended and one from the \mathcal{J}_{14} -helical conformation, were carried out. The backbone atom-positional root-meansquare-difference (RMSD) with respect to an ideal 3_{14} -helix was very high for both simulations, 0.66 ± 0.02 nm starting from the extended conformation and $0.64 \pm$ 0.02 nm starting from the β_{14} -helix. In the simulation started from the extended conformation, the 3_{14} -helix was never populated. In the other simulation, the 3_{14} -helix starting conformation unfolded in the equilibration period, and the molecule diverged from the \mathcal{J}_{14} -helix. Upon visual analysis of the trajectories, we have noticed that the molecules tend to adopt a different kind of helix, namely the 2_8 -helix characterized by NH(i) to CO(i-2) H-bonds. The backbone atom-positional RMSD with respect to an ideal 2_8 -helix is shown in Fig. 3, a. In the simulation started from the 3_{14} -helix, 2_8 -helical structures are populated more often than in the other simulation. Since the rotational barrier of the $C_M - C_A - C - O$ torsional angle is high, it is expected that the helical structures are less populated in the simulation started from the extended structure: in this case, all the backbone torsional angles were set to 180° , so first an isomerization would have to occur for the molecule to fold into a helix, be it 3_{14} or 2_8 . Altogether, it is seen that the conformation of the molecule is heavily fluctuating which results in a wide range of RMSD values and large occasional RMSD excursions. The 28-helix was populated in both simulations, but only in the simulation starting from the \mathcal{J}_{14} -helix, the 2_8 -helix was stable for longer periods of time (between 30 and 40 ns, and between 145 and 150 ns). From the time series of the distance between the C_A -atom of the (tertbutoxy)carbonyl (Boc) protecting group on the N-terminus and the carboxy C-atom (C) at the C-terminus (Fig. 3,b), it is clearly seen that the molecule is fluctuating in both simulations: the hexamer is bending, compacting and stretching back to more extended conformations. The head-to-tail distance fluctuates between 0.5 and 2.5 nm, which is consistent with the 2_8 -helical geometry (head to tail distance of 1.85 nm).

The six C_M-C_A-C-O dihedral angles were monitored, and their distributions are shown in *Fig. 4*. In the simulation started from the extended conformation, where initially all the dihedral angles were set to 180°, almost no s-*cis*-conformation is observed, except for the fifth residue. This is because of the large energy barrier for s*trans-cis* rotation. In the simulation started from the 3_{14} -helix, the situation is different: the s-*cis*-conformation is populated more often than in the extended case. The finding that both simulations produce different distributions in the C_M-C_A-C-O dihedral angle is an indication that the folding equilibrium is still not fully sampled. In the simulation started from the extended conformation, the C_M-C_A-C-O dihedral angle of residue 3, which is in the middle of the molecule, is mostly in the s-*trans*conformation. This fact is an explanation for the low head-to-tail distance observed in this simulation: a large fraction of trajectory structures show a V-shaped geometry. The



Fig. 3. a) *Time series and distributions of the atom-positional backbone RMSDs with respect to an ideal* 2₈-*helix. Black:* starting from an extended conformation. *Red:* starting from a 3₁₄-helical conformation.
b) *Time series of the head-to-tail distance.* The head to tail distance of an ideal 2₈-helix is 1.85 nm.

distribution of φ and ψ angles gives further insight into the folding equilibrium (*Fig. 5*). The dense areas of the contour plot histogram show the angles which are adopted most frequently in the simulation. In the simulation started from the extended conformation, the dihedral angle values characteristic of the 2₈-helical conformation (indicated by the arrow) are only significantly populated by residues 2, 3, and 4. Residues 5 and 6 also adopt these values, but only to a very low extent. Residue 1, finally, does not populate this area at all. In the simulation started from the 3₁₄-helix, the situation is quite similar. Again, the φ and ψ values characteristic for the 2₈-helical conformation are never populated by residue 1. From these observations, one may conclude that helical conformations are populated by the inner residues in both simulations. The residues at the N- and C-termini do not exhibit helical conformations. However, one should keep in mind that in β -peptides a third backbone degree of freedom is present. So, the a



Fig. 4. Populations of the six $C_M - C_A - C - O$ backbone torsional angles of the β -hexapeptide





population of particular φ and ψ angles is necessary but not sufficient evidence for particular helical structures.

The most dominant stable H-bonds in both simulations are from amide H-atoms H-(i) to carbonyl O-atoms O-(i-2) (*Table*). These H-bonds are dominant in both simulations, yet they break and reform during the simulation. In the simulation started from the extended conformation, there is a period where only a few of these H-bonds are present (80–100 ns). During this period, the molecule is rather unstructured. The H-bonds which define the 3_{14} -helix are not seen.

| Donor | Acceptor | Occurrence [%] |
|---------------------------|----------|----------------|
| Initially extended | | |
| 2-Ala-H | 1-Boc-O | 91 |
| 3-Leu-H | 1-Boc-O | 34 |
| 3-Leu-H | 1-Val-O | 42 |
| 4-Val-H | 2-Ala-O | 85 |
| 5-Ala-H | 3-Leu-O | 47 |
| 6-Leu-H | 4-Val-O | 10 |
| Initially 3_{14} -helix | | |
| 2-Ala-H | 1-Boc-O | 55 |
| 3-Leu-H | 1-Val-O | 50 |
| 4-Val-H | 2-Ala-O | 70 |
| 5-Ala-H | 3-Leu-O | 24 |
| 6-Leu-H | 4-Val-O | 7 |

Table. The Most Dominant H-Bonds

The structures from our trajectories were conformationally clustered to find the most dominant conformations in the ensemble (*Fig. 6*). The clusters obtained in the simulation which was started from the extended configuration are very diverse, and are populated to a rather low degree: the five biggest clusters contain together only 42% of structures. Among the central structures of the clusters, there is a significant helical structure (2_8 -helix, cluster 3). In the simulation started from the 3_{14} -helix, the situation is different. Here, more helical clusters (1 and 3) are present, but the populations are even smaller: the five biggest clusters represent only 28% of all the structures.

From the cluster analysis, it is again seen that the molecule does fold into a 2_8 -helical geometry but that this helix is not very stable.

Finally, the experimental finding that no significant NOEs could be assigned between non-neighboring H-atoms [10] seems at odds with our simulations. The NOEs calculated from the trajectories are shown in *Fig.* 7. The total number of predicted NOE signals is the same for both simulations, but the signal intensities differ slightly. The predicted NOE signals do not bridge the atoms in a way indicative of a helical structure: many different structures may fit these calculated NOE signals. Furthermore, it is worth mentioning that the experimental observation of amide H-atom chemical shifts may be difficult [10]. Without the amide H-atoms, the number of predicted NOE signals would decrease significantly, and the resulting NOE signals would all be rather weak (data not shown). It should also be kept in mind that other reasons exist for the

non-observation of NOEs predicted by the simulation, including, for example, spin diffusion [20].

Conclusions. – Although methylidene containing β^3 -peptides may seem interesting with an eye to pharmaceutical application, their usefulness seems questionable: within the limits of sampling employed, the methylidene β -peptide studied here does not seem to form any permanent secondary structure except for a complete 2₈-helix for 5% of the time and partial 2₈-helical conformations for *ca.* 20% of the time. This makes the design of molecules based on such β^3 -amino acids, which would adopt particular secondary and tertiary structure, complicated. It is questionable that such a flexible and unstructured molecule can carry out a particular function like inhibition of a cellular process or translocation over a membrane. This is most probably the reason why **1** shows neither antibacterial activity nor hemolytic activity against erythrocytes [21].

Experimental Part

Simulation Setup. Both simulations were performed using the GROMOS96 biomolecular simulation package [18][19] and the 53A6 GROMOS force field [15]. For the non-standard atoms, the following parameters were used: $m(C_M) = 14.027$ u; $q(C_M) = 0$ e; Lennard–Jones, $C_6^{1/2}(C_M) = 0.08642$ (kJ·mol⁻¹· nm⁶)^{1/2}, $C_{12}^{1/2}(C_M) = 5.828 \times 10^{-3}$ (kJ·mol⁻¹· nm⁶)^{1/2}; bonds, $b_0(C_A, C_M) = 0.133$ nm; bond angles, $K_{\theta}(C_B, C_A, C_M) = K_{\theta}(C_M, C_A, C) = 685$ kJ·mol⁻¹, $\theta_0(C_B, C_A, C_M) = \theta_0(C_M, C_A, C) = 121^{\circ}$, $K_{\theta}(C_B, C_A, C) = 500$ kJ·mol⁻¹, $\theta_0(C_B, C_A, C) = 120^{\circ}$; torsional angles, $K_{\varphi_n}(N, C_B, C_A, C) = 5.92$ kJ·mol⁻¹, $\cos(\delta_n)(N, C_B, C_A, C) = +1.0$, $m_n(N, C_B, C_A, C) = 3$, $K_{\varphi_n}(C_B, C_A, C_N) = 16.7$ kJ·mol⁻¹, $\cos(\delta_n)(C_B, C_A, C, N) = -1.0$, $m_n(C_B, C_A, C, N) = 2$; improper dihedral angles, $K_{\xi}(C_A, C_B, C, C_M) = 0.051$ kJ·mol⁻¹·degree⁻², $\xi_0(C_A, C_B, C, C_M) = 0^{\circ}$. Bond lengths were constrained using the SHAKE algorithm [22] and a relative tolerance of 10^{-4} .

Initial coordinates of the 3_{14} -helical conformation were generated from an ideal 3_{14} -helix by the addition of the methylidene group and the appropriate side chain to each amino acid. The initial coordinates of the extended conformation were generated by setting all backbone dihedral angles to 180° .

The peptide was solvated in a cubic box containing 961 (3_{14}) and 964 (extended) MeOH molecules (initial box edge length of 4 nm). Periodic boundary conditions were applied. Initially, the energy was minimized, and both simulations were initiated using the following equilibration scheme. First, the initial velocities were randomly generated from a Maxwell-Boltzmann distribution at 50 K. All solute atom positions were restrained to their positions in the initial structures through a harmonic potential energy term with a force constant of $2.5 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$. The system was simulated with these settings for 20 ps. Second, the temp. was raised in 50 K steps during five additional 20-ps equilibration steps, with the positional restraints being reduced by $5 \times 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ at each step. Next, a production simulation was performed. The temp. of 300 K and atmospheric pressure were kept constant using a weak-coupling approach [23] with relaxation times $\tau_T = 0.1$ and $\tau_p = 0.5$ ps, and an isothermal compressibility of $2.0 \times$ 10^{-3} (kJ·mol⁻¹·nm⁻³)⁻¹. Non-bonded interactions were calculated using a triple-range cutoff scheme. The interactions within a cutoff distance of 0.8 nm were calculated at every step from a pair list which was updated every fifth time step. At this point, interactions between atoms (of charge groups) within 1.4 nm were also calculated and were kept constant between updates. To account for the influence of the homogeneous medium outside the cutoff sphere with a radius of 1.4 nm, a reaction field contribution with a relative dielectric permittivity ε of 18.6 was added.

Analysis. Atom-positional root-mean-square-differences (RMSD) between the backbone atoms of the indicated (*e.g.*, 2_8 -helical) conformation, and the simulated structures were calculated after superposition of the backbone atoms. Conformational clustering analysis was performed as described in [24] on a set of 1000 structures taken at 0.15-ns intervals from the simulation. The atom-positional backbone RMSD was used as similarity criterion. A similarity cut-off (maximum cluster radius) of







0.1 nm was chosen [24]. The presence of a H-bond was determined by geometrical criteria. If the H-acceptor distance was less than 0.25 nm and the donor -H – acceptor angle was at least 135°, the H-bond was considered to be present. Distances between two H-carrying united atoms or explicit H-atoms were averaged using $r = (\langle r^{-6} \rangle)^{-1/6}$ averaging. A NOE signal was considered to be present if the average distance *r* was smaller than 0.5 nm. NOEs between covalently bound neighboring atoms (first and second neighbors) were neglected. The intensity *I* of a NOE signal was calculated from the distance *r* by $I = r^{-6}$. The relative intensities were calculated as I/I_{max} , where I_{max} is the maximum intensity observed.

Software and Hardware. All simulation and energy-minimization computations were carried out using GROMOS96 [18][19]. For analysis, either GROMOS ++ 0.2.4 [25] or esra¹) were used. Additional analysis, conversion, and batch programs were written either in Perl or Java. The Java esra analysis programs for the prediction of NOE signals²) was published on the esra CVS server.

Visualization was performed with the Visual Molecular Dynamics (VMD) [26] software.

Financial support from the National Center of Competence in Research (NCCR) Structural Biology of the Swiss National Science Foundation (SNSF) is gratefully acknowledged. We further thank Zrinka Gattin for proposing this project.

REFERENCES

- [1] S. H. Gellman, Acc. Chem. Res. 1998, 31, 173.
- [2] D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, Chem. Rev. 2001, 101, 3893.
- [3] 'Foldamers: Structure, properties and applications', Eds. S. Hecht, I. Huc, Wiley-VCH, Weinheim, 2007.
- [4] D. Seebach, S. Abele, K. Gademann, B. Jaun, Angew. Chem., Int. Ed. 1999, 38, 1595.
- [5] D. Seebach, P. E. Ciceri, M. Overhand, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta* 1996, 79, 2043.
- [6] D. F. Hook, F. Gessier, C. Noti, P. Kast, D. Seebach, ChemBioChem 2004, 5, 691.
- [7] M. Rueping, Y. Majahan, M. Sauer, D. Seebach, ChemBioChem 2002, 3, 257.
- [8] Y. Hamuro, J. P. Schneider, W. F. DeGrado, J. Am. Chem. Soc. 1999, 121, 12200.
- [9] M. Werder, H. Hauser, S. Abele, D. Seebach, Helv. Chim. Acta 1999, 82, 1774.
- [10] D. J. Bierbaum, D. Seebach, Aust. J. Chem. 2004, 57, 859.
- [11] R. Baron, D. Bakowies, W. F. van Gunsteren, X. Daura, Helv. Chim. Acta 2002, 85, 3872.
- [12] A. Glaettli, X. Daura, D. Seebach, W. F. van Gunsteren, J. Am. Chem. Soc. 2002, 124, 12972.
- [13] A. Glaettli, D. Seebach, W. F. van Gunsteren, Helv. Chim. Acta 2004, 87, 2487.
- [14] A. Glaettli, X. Daura, P. Bindschaedler, B. Jaun, Y. R. Mahajan, R. I. Mathad, M. Rueping, D. Seebach, W. F. van Gunsteren, *Chem.-Eur. J.* 2005, 11, 7276.
- [15] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, J. Comput. Chem. 2004, 25, 1656.
- [16] C. Oostenbrink, T. A. Soares, N. F. A. van der Vegt, W. F. van Gunsteren, Eur. Biophys. J. 2005, 34, 273.
- [17] D. Trzesniak, B. Jaun, R. I. Mathad, W. F. van Gunsteren, Biopolymers 2006, 83, 636.
- [18] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, 'Biomolecular simulation: The GROMOS96 manual and user guide', Hochschulverlag AG, ETH-Zürich, 1996.
- [19] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, W. F. van Gunsteren, J. Phys. Chem., A 1999, 103, 3596.
- [20] B. Zagrovic, W. F. van Gunsteren, Proteins 2006, 63, 210.
- [21] P. I. Arvidsson, N. S. Ryder, H. M. Weiss, D. F. Hook, J. Escalante, D. Seebach, *Chem. Biodiv.* 2005, 2, 401.
- [22] J. P. Ryckaert, G. Giccotti, H. J. C. Berendsen, J. Comput. Phys. 1977, 23, 327.

²) programs.predict_noe.

¹) Java analysis package written by *Mika A. Kastenholz* and *Vincent Kräutler*.

- [23] H. J. C. Berendsen, J. P. M. Postma, A. DiNola, W. F. van Gunsteren, J. R. Haak, J. Chem. Phys. 1984, 81, 3684.
- [24] X. Daura, W. F. van Gunsteren, A. E. Mark, Proteins: Struct., Funct., Genet. 1999, 34, 269.
- [25] M. Christen, P. H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Kräutler, C. Oostenbrink, C. Peter, D. Trzesniak, W. F. van Gunsteren, J. Comput. Chem. 2005, 26, 1719.
- [26] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graphics 1996, 14, 33.

Received July 19, 2007