

# Protein Science

## A molecular dynamics study of the 41-56 beta-hairpin from B1 domain of protein G

D Roccatano, A Amadei, A Di Nola and HJ Berendsen

*Protein Sci.* 1999 8: 2130-2143

Access the most recent version at doi:[10.1110/ps.8.10.2130](https://doi.org/10.1110/ps.8.10.2130)

---

### References

Article cited in:

<http://www.proteinscience.org/cgi/content/abstract/8/10/2130#otherarticles>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

### Notes

---

To subscribe to *Protein Science* go to:  
<http://www.proteinscience.org/subscriptions/>

---

# A molecular dynamics study of the 41–56 $\beta$ -hairpin from B1 domain of protein G

DANILO ROCCATANO,<sup>1</sup> ANDREA AMADEI,<sup>2</sup> ALFREDO DI NOLA,<sup>2</sup>  
AND HERMAN J.C. BERENDSEN<sup>1</sup>

<sup>1</sup>Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

<sup>2</sup>Dipartimento di Chimica, Università di Roma (La Sapienza), P. le A. Moro 5, 00185 Rome, Italy

(RECEIVED February 19, 1999; ACCEPTED July 5, 1999)

## Abstract

The structural and dynamical behavior of the 41–56  $\beta$ -hairpin from the protein G B1 domain (GB1) has been studied at different temperatures using molecular dynamics (MD) simulations in an aqueous environment. The purpose of these simulations is to establish the stability of this hairpin in view of its possible role as a nucleation site for protein folding. The conformation of the peptide in the crystallographic structure of the protein GB1 (native conformation) was lost in all simulations. The new equilibrium conformations are stable for several nanoseconds at 300 K ( $>10$  ns), 350 K ( $>6.5$  ns), and even at 450 K (up to 2.5 ns). The new structures have very similar hairpin-like conformations with properties in agreement with available experimental nuclear Overhauser effect (NOE) data. The stability of the structure in the hydrophobic core region during the simulations is consistent with the experimental data and provides further evidence for the role played by hydrophobic interactions in hairpin structures. Essential dynamics analysis shows that the dynamics of the peptide at different temperatures spans basically the same essential subspace. The main equilibrium motions in this subspace involve large fluctuations of the residues in the turn and ends regions. Of the six interchain hydrogen bonds, the inner four remain stable during the simulations. The space spanned by the first two eigenvectors, as sampled at 450 K, includes almost all of the 47 different hairpin structures found in the database. Finally, analysis of the hydration of the 300 K average conformations shows that the hydration sites observed in the native conformation are still well hydrated in the equilibrium MD ensemble.

**Keywords:**  $\beta$ -hairpin; essential dynamics; molecular dynamics; peptide conformation in solution; peptide folding; protein G B1

Insight into the molecular mechanisms governing the stability and dynamics of secondary structure elements constitutes a very important step toward understanding the protein folding process. In fact, secondary structural elements are in general present in the unfolded state of proteins and their early appearance plays an important role in the folding process, providing nucleation sites for the unfolded part of the protein (Dyson & Wright, 1993; Dobson et al., 1998). For this reason many experimental and theoretical studies have been devoted to isolated protein fragments that form secondary structure elements in solution with the aim to characterize their stability and dynamics (Dyson & Wright, 1993; Blanco et al., 1998).

The  $\beta$ -sheet (in particular the antiparallel ones) is the most frequently occurring secondary structure component of proteins (Kabsch & Sander, 1983), and therefore the study of a context free

$\beta$ -hairpin peptide model could be very important for understanding the nature of  $\beta$ -sheet stabilizing interactions.  $\beta$ -hairpins are the simplest representation of an antiparallel  $\beta$ -sheet but it has not been possible to study these structures because of the difficulty to find stable peptides in solution. Only recently stable  $\beta$ -hairpin monomeric peptides obtained from proteins (Blanco et al., 1994; Blanco & Serrano, 1995; Searle et al., 1995) or from de novo design (Ramirez-Alvarado et al., 1996; Sieber & Moe, 1996; de Alba et al., 1997) have been discovered and characterized using NMR and other spectroscopic techniques (Muñoz et al., 1997). These studies revealed that such  $\beta$ -hairpins have a lifetime of several microseconds and therefore have a folding/unfolding rate at room temperature that is 30 times smaller (Muñoz et al., 1997) than that of a polyalanine  $\alpha$ -helix (Williams et al., 1996; Thompson et al., 1997). It has been shown that  $\beta$ -hairpin structures can be stable in solution and hence could act as nucleation sites for the folding of the protein (Blanco et al., 1994; Blanco & Serrano, 1995).

Theoretical studies on the dynamics of  $\beta$ -hairpin peptides in solution are limited to few examples. Different authors have

Reprint requests to: Herman J.C. Berendsen, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), Department of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands; e-mail: berendsen@chem.rug.nl.

studied  $\beta$ -hairpins folding/unfolding dynamics using simplified models with stochastic dynamics (Yapa et al., 1992; Klimov & Thirumalai, 1997). Molecular dynamics (MD) simulations with implicit solvation model have been recently used to study a synthetic  $\beta$ -hairpin forming peptide (Schaefer et al., 1998). Pugliese et al. (1995) studied with full-atoms MD simulations with explicit solvent the unfolding of a  $\beta$ -hairpin from barnase. They found that the peptide lost its native conformation at 450 K, adopting a collapsed conformation stabilized by hydrophobic interactions that required more aggressive denaturing conditions (600 K) to reach a complete unfolding. Their simulations were limited to short time scales (300 ps). Prévost and Ortman (1997) performed a refolding MD simulation of the same peptide using a simulated annealing procedure. They found that hydrophobic and H-bond interactions act as stabilizing factors in the refolding process. A long time scale MD simulation (10 ns) was performed by Constantine et al. (1996) using a different type of  $\beta$ -hairpin. These authors showed that the peptide lost its  $\beta$ -hairpin structure after about 2 ns of simulation at low temperature, where it was supposed to be stable in a  $\beta$ -hairpin conformation. However, the structural behavior in the first 2 ns was consistent with experimental data. In the paper they discuss the results of their long simulations to point out the effect of the force field or the simulation protocol on the stability of the peptide. This problem is still an open question that is matter of controversial discussions (Braxenthaler et al., 1997; Finkelstein, 1997; Daura et al., 1998). We generally agree with the fact that it is very important to have an extensive sampling of the system to obtain more reliable theoretical results that can be compared with the experimental observations.

In this paper, we want to extend the theoretical knowledge of the stability and dynamics of the  $\beta$ -hairpin structures by studying the isolated 41–56  $\beta$ -hairpin ( $\beta$ -hairpin II) from the protein GB1 using MD simulations at different temperatures. This peptide was recently studied using different experimental techniques (Blanco et al., 1994, 1996; Muñoz et al., 1997). From these studies the stability and the folding/unfolding rate have been assessed. Furthermore, Muñoz et al. (1998) have formulated a simple statistical mechanical model providing an explanation of the peptide stability and a possible mechanism of the folding/unfolding process. According to this model  $\beta$ -hairpins, in contrast to  $\alpha$ -helices where stability is the result of local residue interactions (mainly H-bonds) (Thompson et al., 1997), are stabilized by the combination of entropic effects, main-chain hydrogen bonds, and hydrophobic interactions between the two strands (Muñoz et al., 1997, 1998). The importance of hydrophobic interactions for the stability of  $\beta$ -hairpins in solution has been shown by other authors using different types of  $\beta$ -hairpins (Searle et al., 1995; Blanco et al., 1998; Maynard et al., 1998).

The general outline is as follows: in Materials and methods, the procedures used to perform the simulations and to analyze the trajectories are described. In Results and discussion, simulations at 280, 300, 350, and 450 K are described. In this section we show the results of structural analysis on the equilibrium peptide conformations in the different simulations. Furthermore, “essential dynamics” analysis of the trajectories is reported. Using the essential dynamics method, we also performed a comparison of  $\beta$ -hairpin fragments extracted from the Protein Data Base (PDB) (Bernstein et al., 1977) with structural fluctuations of the  $\beta$ -hairpin II obtained from the different simulations. The aim of this analysis was to compare the distribution of the crystal structures in the sub-

spaces of configurations spanned by the peptide during the simulations. The distribution of water molecules around the peptide at 300 K is also reported and discussed. Finally, in Conclusions, a summarizing discussion of the work is reported.

## Results and discussion

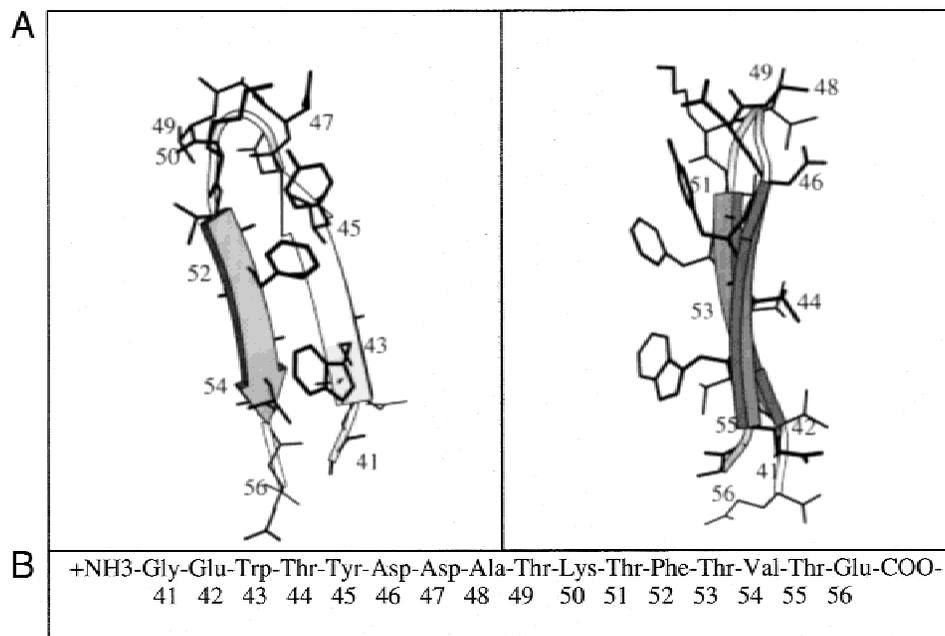
### Native conformation

In Figure 1 the crystal structure of the  $\beta$ -hairpin II is shown. Following Kabsch–Sander’s DSSP definition (Kabsch & Sander, 1983) for secondary structures, the residues 42–46 and 51–55 are in a strand conformation. Using the classification proposed by Sibanda and Thornton (1991), the  $\beta$ -hairpin is of type 4:4 because it has four residues in the turn region, and the distal strand residues (Asp46, Thr51) have two H-bonds. The four residues Trp43, Tyr45, Phe52, Val54 form a hydrophobic cluster that in protein GB1 contributes to stabilization of the protein structure by interactions with other hydrophobic protein residues. According to experimental results (Blanco & Serrano, 1995; Blanco et al., 1996; Muñoz et al., 1997), this hydrophobic cluster remains intact also in the isolated peptide and plays an important role in the stability of the peptide in solution. The residues in the  $\beta$ -strand region have three pairs of H-bonds, linking residues 42, 44, and 46 with residues 55, 53, and 51. In the turn region (residues 47–50), residues Ala48, Thr49, and Lys50 form main-chain H-bonds with Asp46. The carboxylate of Asp46 and the NH of Ala48 stabilize the turn region forming another H-bond. The Tyr45 hydroxyl group is H-bonded to the side chain of Asp47. Finally there are H-bonds between the hydroxyl groups of Thr44–Thr51 and Thr49–Thr53. In the turn region, Lys50 is in the left-handed helical conformation with  $\phi$  and  $\psi$  angles of about 60 and 40°. Furthermore, the side-chain charged groups of Asp47 and Lys50 are involved in a salt bridge interaction.

### Structural analysis

In Figure 2, the root-mean-square deviation (RMSD) of the peptide backbone atoms, with respect to the minimized crystallographic structure, are reported for the 280, 300, 350, and 450 K simulations. At 280 K a progressive increase of the RMSD in the first 1.5 ns, followed by a stabilization to  $0.25 \pm 0.03$  nm, can be observed. The 300 K simulation shows a rapid increase of the RMSD followed by large fluctuations up to 2 ns followed by a stabilization to  $0.30 \pm 0.02$  nm. In the case of the 350 K simulation, a value of  $0.28 \pm 0.04$  nm occurs after about 800 ps, although, in the last part of the trajectory, there are large fluctuations with a temporary recovery of the native structure. Finally, at 450 K there are large fluctuations ( $\pm 0.07$  nm) around an average value of 0.26 nm in the first 2.5 ns of the trajectory, followed by an abrupt increase up to 0.6 ns due to the start of the unfolding. In this paper we do not analyze the unfolding mechanism of the peptide and in the following analysis, if not explicitly reported, we use the first 2.5 ns of the 450 K trajectory.

The similarity between structures obtained in the different simulations was assessed by comparing (using the RMSD of backbone atoms) conformations from the 280, 350, and 450 K simulations with those from the 300 K simulation. The results of these comparisons, in the form of two-dimensional RMSD maps, are reported in Figure 3. The RMSD values of the 280 K simulation decrease with time showing a relaxation toward the 300 K equilibrium structures. In the case of the 350 K simulation, there is a



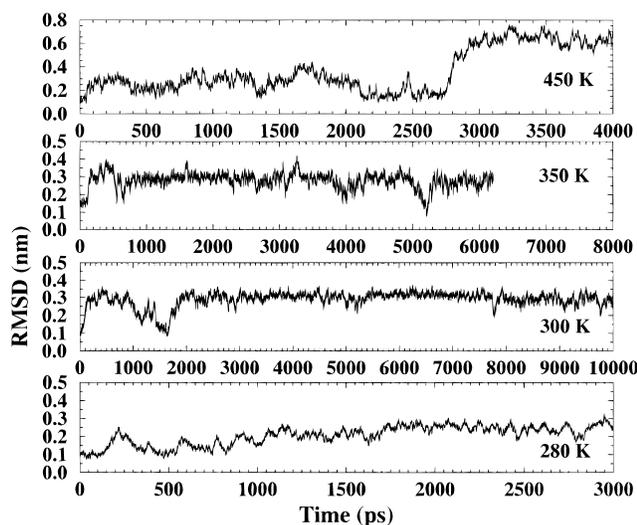
**Fig. 1.** **A:** Two views of the  $\beta$ -hairpin II in the crystallographic conformation. **B:** Amino acid sequence of the peptide. The pictures were created using the program MOLSCRIPT (Kraulis, 1991).

more uniform distribution of low RMSD values; the black bands correspond to configurations in which there is a temporary recovery of the native conformation as also shown in Figure 2. At 450 K the large fluctuations observed in Figure 2 (horizontal) are present as black bands of large RMSD values. From these comparisons the tendency of the peptide to adopt the same backbone conformation at the three different temperatures is evident. The average peptide radius of gyration in all simulations is 0.74 nm. This value, compared with the radius of gyration of the native structure (0.83 nm), shows an evident increase of the compactness of the peptide struc-

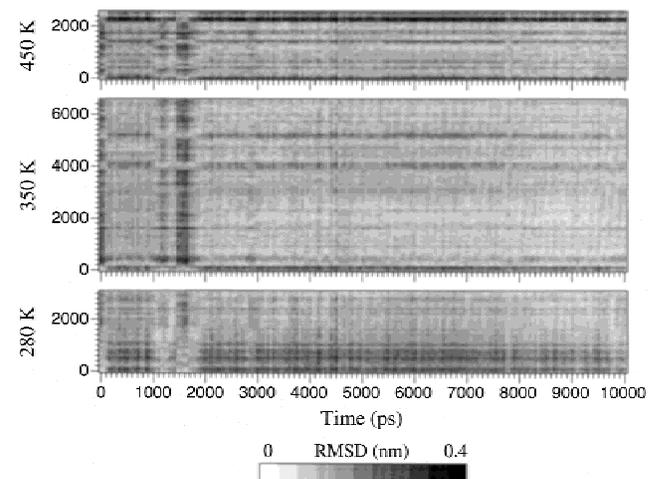
ture. In Figure 4, some examples of structures adopted by the peptide during the simulations are shown.

The 280 K simulation will be excluded from now on because, as shown by the previous RMSD plots and maps and as it will be shown by the essential dynamics (ED) analysis results in the next section, the dynamical behavior and the structural equilibrium configurations slowly converge to those observed at higher temperatures.

In Figure 5 the secondary structure, evaluated using the Kabsch-Sander algorithm (Kabsch & Sander, 1983) is reported for the three simulations. In all cases there is a loss of native secondary



**Fig. 2.** Backbone RMSD from the  $\beta$ -hairpin II crystal structure as function of the time for the 280, 300, 350, and 450 K simulations.



**Fig. 3.** Two-dimensional RMSD maps. Sampled structures along the 300 K simulation (horizontal scale) are compared with those along the 280, 350, and 450 K simulations (vertical scale) and their backbone RMSDs are represented using a gray scale color map.

**Table 1.** Amino acid sequence of 47 different 4:4  $\beta$ -hairpin fragments obtained from the not redundant PDB database of the WHATIF program<sup>a</sup>

| PDB code | Sequence number <sup>b</sup> | Amino acid composition            |
|----------|------------------------------|-----------------------------------|
| 1bpi     | 19–34                        | I R Y F Y N A K A G L C Q T F V   |
| 2bbk     | 52–67                        | V A S C Y N P T D G Q S Y L I A   |
| 2ebn     | 43–58                        | A N I N Y D A A N D K V F V S N   |
| 1clc     | 59–74                        | A T S M F D N D T K E T V Y I A   |
| 1aoc     | 148–163                      | R L V T Y N L E K D G F L C E S   |
| 1rge     | 68–83                        | R R I I T G E A T Q E D Y Y T G   |
| 1arb     | 104–119                      | T V K A T Y A T S D F T L L E L   |
| 1igd     | 46–61                        | G V W T Y D D A T K T F T V T E   |
| 7rsa     | 106–121                      | I I V A C E G N P Y V P V H F D   |
| 2phy     | 107–122                      | V H M K K A L S G D S Y W V F V   |
| 2mcm     | 71–86                        | S F Q A V V G A D G T P W G T V   |
| 1htr     | 15–30                        | F G E I S I G T P P Q N F L V L   |
| 1kap     | 430–445                      | A I L S Y D A A S K A G S L A I   |
| 1frd     | 4–19                         | Q V R L I N K K Q D I D T T I E   |
| 3cla     | 86–101                       | Q F T V F H Q E T E T F S A L S   |
| 2ltu     | 162–177                      | V V I A F N A A T N V L T V S L   |
| 1gof     | 298–313                      | N G E V Y S P S S K T W T S L P   |
| 1onc     | 65–80                        | S D C N V T S R P C K Y K L K K   |
| 2bbk     | 130–145                      | A V G V V D L E G K A F K R M L   |
| 2bbk     | 314–329                      | L L Y A L S T G D K T L Y I H D   |
| 1bur     | 100–115                      | R I I G F D N K R Q V L Q C I S F |
| 3tgl     | 60–75                        | A M V A R G D S E K T I Y I V F   |
| 1hbq     | 75–90                        | G T F T D T E D P A K F K M K Y   |
| 1ayl     | 189–204                      | N F V A F N L T E R M Q L I G G   |
| 1mml     | 146–161                      | A F E W R D P E M G I S G Q L T   |
| 1ytb     | 41–56                        | A V I M R I R E P K T T A L I F   |
| 1rsy     | 17–32                        | Y S L D Y D F Q N N Q L L V G I   |
| 2prd     | 28–43                        | N K Y E Y D P D L G A I K L D R   |
| 1pnk     | 125–140                      | N I L Q T D Q T T Q T A Y A K S   |
| 1oac     | 15–30                        | A D V Q W D D Y A Q L F T L I K   |
| 1oac     | 159–174                      | I E A V V D L Q N N K L L S W Q   |
| 1oac     | 260–275                      | L V A V V D L E Q K K I V K I E   |
| 1mpp     | 17–32                        | A I P V S I G T P G Q D F Y L L   |
| 1obp     | 40–55                        | R E L V F D D E K G T V D F Y F   |
| 1lts     | 84–99                        | K L C V W N N K T P N S I A A I   |
| 21lb     | 68–83                        | Y L S C V L K D D K P T L Q L E   |
| 1led     | 173–188                      | A H I T Y D A R S K I L T V L L   |
| 1mjc     | 7–22                         | I V K W F N A D K G F G F I T P   |
| 1frp     | 167–182                      | N C F M L D P A I G E F I L V D   |
| 1bbp     | 113–128                      | Y Y C K Y D E D K K G H Q D F V   |
| 1got     | 164–179                      | T C A L W D I E T G Q Q T T T F   |
| 1got     | 248–263                      | T C R L F D L R A D Q E L M T Y   |
| 1got     | 273–288                      | T S V S F S K S G R L L L A G Y   |
| 1csn     | 21–36                        | I F E G T N L L N N Q Q V A I K   |

<sup>a</sup>The first five fragments have an average backbone RMSD with respect to the equilibrated 300, 350, and 450 K configurations less than 0.15 nm.

<sup>b</sup>According to WHATIF numbering of the sequence.

structure. The two residues at the end of the two strands rapidly lose their native conformations. The same tendency was also noticed for the two strand residues Asp46 and Thr51. In the loop region, three of the four loop residues preserve their native turn conformations while the fourth residue (Lys50) tends to assume a coil configuration. Note that in the DSSP notation the difference between turn and bend conformation is the presence or absence of a backbone–backbone H-bond. This conversion is evident in the

loop where the flexibility of the region makes this change easily possible. Finally, residues 43–45 and 52–54 maintain their strand native conformations at all temperatures.

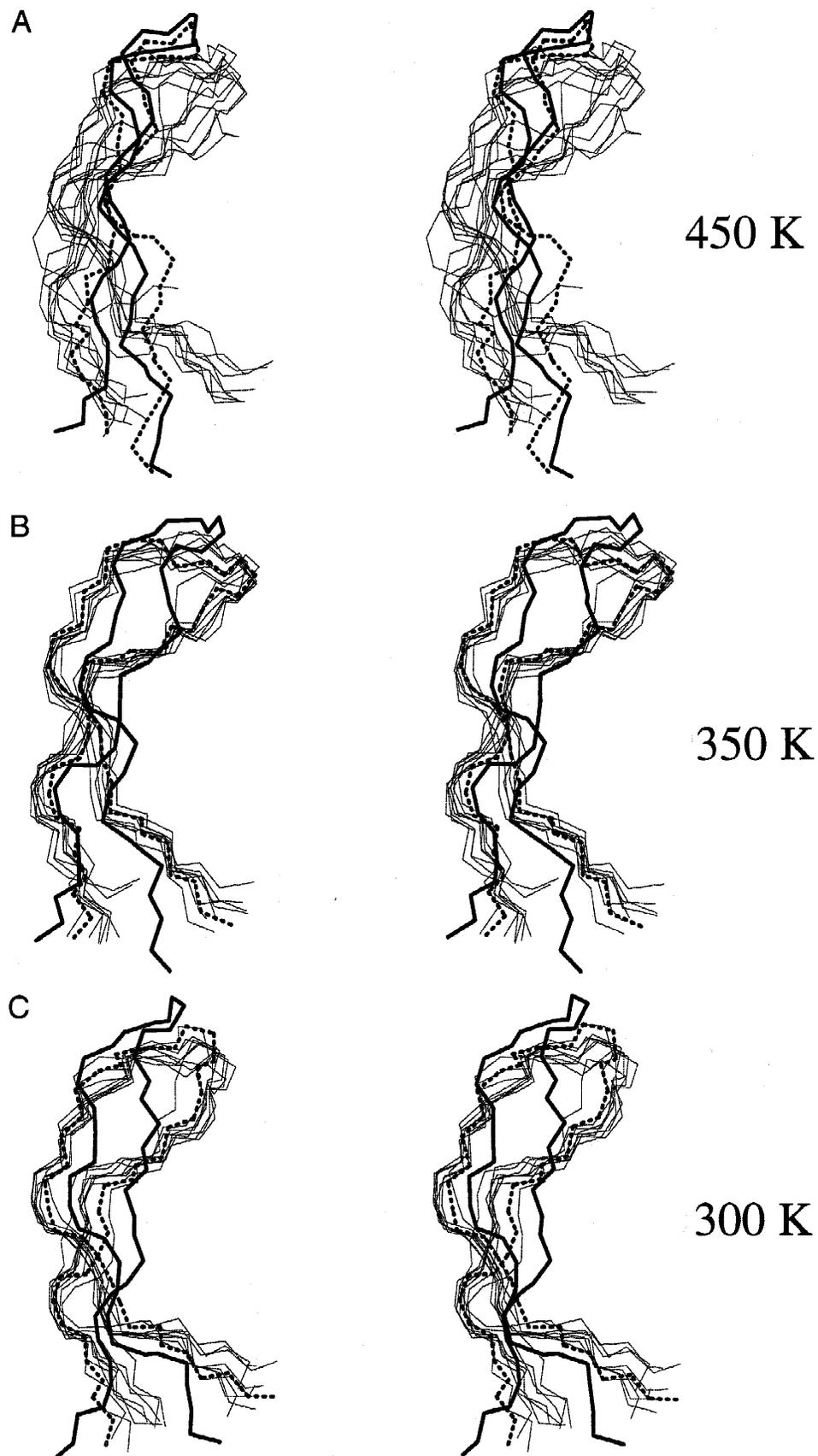
The Ramachandran plots of  $\phi$  and  $\psi$  adopted by different conformations during the last 8 ns of 300 K, 5 ns of the 350 K, and 2.5 ns of the 450 K simulations are reported in Figure 6. The four crystallographic loop residues and the ten strand ones are indicated with different symbols. The loop residues of the 300 K structures adopt conformations close to the native ones. At 350 K the general behavior resembles that observed at 300 K with the exception of the Lys50 residue that shows a high propensity for the  $\delta$  and  $\beta$  regions of the plot. The  $\phi$  and  $\psi$  angles of the strands show at 300 K a spread of the points toward the upper part of the  $\beta$ -sheet region (above the line  $\phi = \psi$  defining an ideal  $\beta$ -strand), due to the right-hand twisting deformations of the strands (Chothia, 1973). The same localization was obtained at 350 K although for the larger fluctuations the points are more uniformly distributed.

The  $\beta$ -strand structures are very flexible and this flexibility allows them to adopt different geometries (Chothia, 1973; Salemme, 1983). A systematic study of the  $\beta$ -sheet structures in the protein was made by various authors (Chothia, 1983; Salemme, 1983; Maccallum et al., 1995) and in these studies it was shown that almost all strands in the  $\beta$ -sheets are right-handed twisted. This twisting was explained as the results of steric interactions of the  $\beta$ -carbon atoms (Chou et al., 1983a, 1983b) and, more recently, of the stabilization due to electrostatic interactions between main-chain partially charged atoms (Maccallum et al., 1995). The extent of twisting and coiling in strands can be estimated from the position of the  $\phi$  and  $\psi$  angle in the Ramachandran plot (Chothia, 1983). According to Chothia (1983), the twisting of  $\beta$ -strands can be estimated using the  $\langle\phi + \psi\rangle$  values (where, in our case,  $\langle\dots\rangle$  indicates the average over five residues) and the coiling using the torsion angle  $\alpha_i$  defined by the approximate formula (Levitt, 1976):

$$\alpha_i = 180 + \psi_i + \phi_{i+1} + 20(\sin \phi_i + \sin \psi_{i+1}), \quad (1)$$

where the index  $i$  indicates the  $i^{\text{th}}$   $\alpha$ -carbon atom. According to Chothia (1983), a twisted and coiled strand is characterized by  $\alpha_i \neq \alpha_{i+1}$ . Furthermore, if one of the values is large and positive and the other is smaller and negative the coil of the chain will be right-handed.

The average values of  $\langle\phi + \psi\rangle$  calculated from 300, 350, and 450 K conformations, compared with the native one, and with some of the  $\beta$ -hairpin fragments listed in Table 1 (see Materials and methods) are reported in Table 2. The values were calculated using residue 42–46 and 51–55 for the two strands, respectively. It has to be noted that, according to the DSSP definitions, in the present case the strand conformation is limited to the six residues used to calculate the reported values; however, the  $\phi$  and  $\psi$  angles of the residues that are not in a strand configuration remain mainly in the  $\beta$ -region of the Ramachandran plot during the simulations. From Table 2 it is evident that the average equilibrium conformation of the peptide shows an increase of the right-handed twisting with respect to the crystal structure. The two strands are also twisted differently. In all the average structures, obtained from the simulations, the first strand is more twisted than the second one. At 450 K the strands are less twisted because of the larger structural fluctuations that allow the sampling of conformations more close to the native one. In the same table, the  $\langle\phi + \psi\rangle$  values for five different crystallographic  $\beta$ -hairpin fragments (see Materials and



**Fig. 4.** Stereo views of the superimposition of 10 backbone conformations sampled from the 300, 350, and 450 K simulations. In bold solid lines the initial conformations and in bold dotted lines the final conformations are shown.

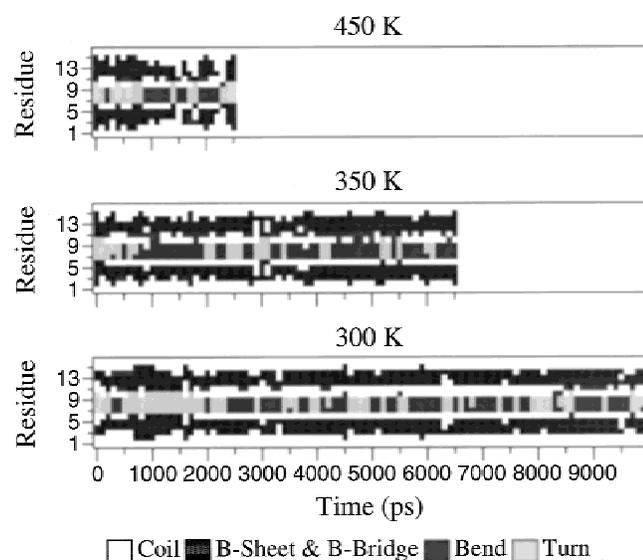


Fig. 5. Secondary structure during the 300, 350, and 450 K simulations.

methods) having large twisting and coiling deformations are reported. The values obtained from the two strands are in the range of those observed for other crystallographic  $\beta$ -hairpin fragments. The MD configurations have an average backbone RMSD from these structures lower than (less than 0.164 nm) that of the native  $\beta$ -hairpin II structure (up to 0.212 nm). In Figure 7 the backbone superimposition of the selected fragments with those of the  $\beta$ -hairpin II is shown.

In Table 3, the average values of  $\alpha_i$  obtained from conformations of the different simulations, compared with the native one, are reported. The alternation of positive and negative values gives a clear indication of the presence of coiling in the structures (as expected for strongly twisted antiparallel  $\beta$ -strands (Chothia, 1983)). In the same table, the  $\alpha_i$  values for the selected  $\beta$ -hairpin fragments are reported. The values  $\alpha_1$  and  $\alpha_2$  of the fragments are, for both strands, closer to those of the MD configurations than to those

of the native  $\beta$ -hairpin II conformation. For the  $\alpha_2$  value, the opposite trend is observed. The reported data clearly show that the  $\beta$ -hairpin structure in solution is quite different from the crystal structure. However, the average MD structure is close to the crystal conformations of other  $\beta$ -hairpins reported in the literature (Table 2).

In Table 4 the percentage of H-bonds is reported. The native main-chain H-bonds shows high occurrence values with the exception of those in the loop region and at the ends of the strands. Note that, according to the Sibanda and Thornton (1991) classification of  $\beta$ -hairpins, the absence of the Thr51 NH CO Asp46 H-bond changes the peptide to a 4:6  $\beta$ -hairpin type. It is possible to observe a relatively low conservation of the native side-chain H-bonds. Different side-chain H-bond interactions, not present in the native structure, were formed during the simulations. In particular those between residues Thr49 and Asp46 show high occurrence values even at high temperatures.

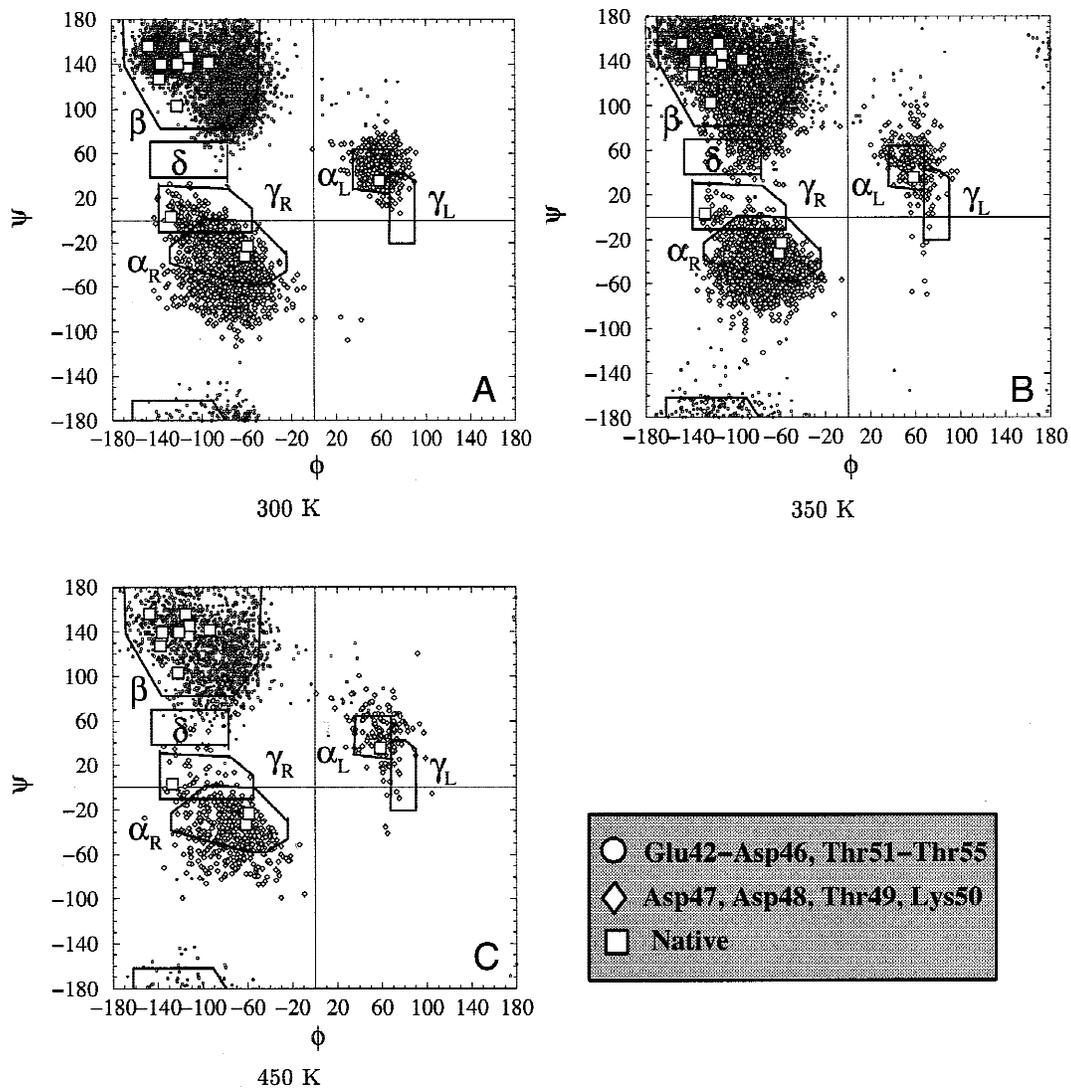
Percentages of occurrence of salt bridges (using a 0.5 nm cut off) are the opposite, charged side-chain groups of Lys50 and Asp47 are 70, 36, 39, and 54 at 280, 300, 350, and 450 K, respectively. The charge groups of the terminal Gly41 and Glu56 residues also form salt bridges during the simulation with percentages of occurrence of 62, 38, 44, and 66 at the same temperatures.

#### Essential dynamics analysis

The analysis of the trajectory using the mean-squared inner product (MSIP) index (see Materials and methods) was performed as follows. The 280 K simulation was divided into three parts of 1 ns each, and an ED analysis was performed for each of them. The eigenvectors obtained from this analysis were compared with those obtained from the last two 4 ns sections of the 300 K simulation. In Table 5 the results of these comparisons are reported. The highest MSIP value is obtained from the two 4 ns sections of the 300 K simulation. This indicates that a rather good convergence in the essential subspace definition can be reached within 4 ns. The MSIP values of the different 280 K trajectory parts with the 300 K sections increase from the first to the last. This means that the peptide is slowly changing toward a conformation with motions very close to those of the equilibrated 300 K simulation. It has to be noticed that in general the obtained MSIP values are quite high, and hence

**Table 2.**  $\langle \phi + \psi \rangle$ , obtained from residues 42–46 (strand I) and 51–55 (strand II), and backbone RMSD values (calculated on residues 42–55) of the MD configurations at the different temperatures (average values), the five  $\beta$ -hairpin fragment having backbone RMSD values, with respect to the three sets of MD configurations, less than 0.15 nm and the native  $\beta$ -hairpin II conformation (1pgb)

| PDB code | Strand I (deg) | Strand II (deg) | zRMSD (300 K) (nm) | RMSD (350 K) (nm) | RMSD (450 K) (nm) | RMSD (cry) (nm) |
|----------|----------------|-----------------|--------------------|-------------------|-------------------|-----------------|
| 2bbk     | 19             | 32              | 0.164              | 0.136             | 0.123             | 0.164           |
| 1bpi     | 31             | 36              | 0.160              | 0.130             | 0.139             | 0.212           |
| 1aoc     | 31             | 30              | 0.159              | 0.130             | 0.133             | 0.190           |
| 2ebn     | 20             | 25              | 0.155              | 0.129             | 0.121             | 0.184           |
| 1clc     | 27             | 26              | 0.147              | 0.114             | 0.104             | 0.174           |
| 1pgb     | 14             | 14              | 0.290              | 0.265             | 0.230             | 0.00            |
| 300 K    | 43             | 36              | 0.000              | 0.080             | 0.110             | 0.290           |
| 350 K    | 40             | 31              | 0.080              | 0.000             | 0.092             | 0.265           |
| 450 K    | 37             | 23              | 0.110              | 0.092             | 0.000             | 0.230           |

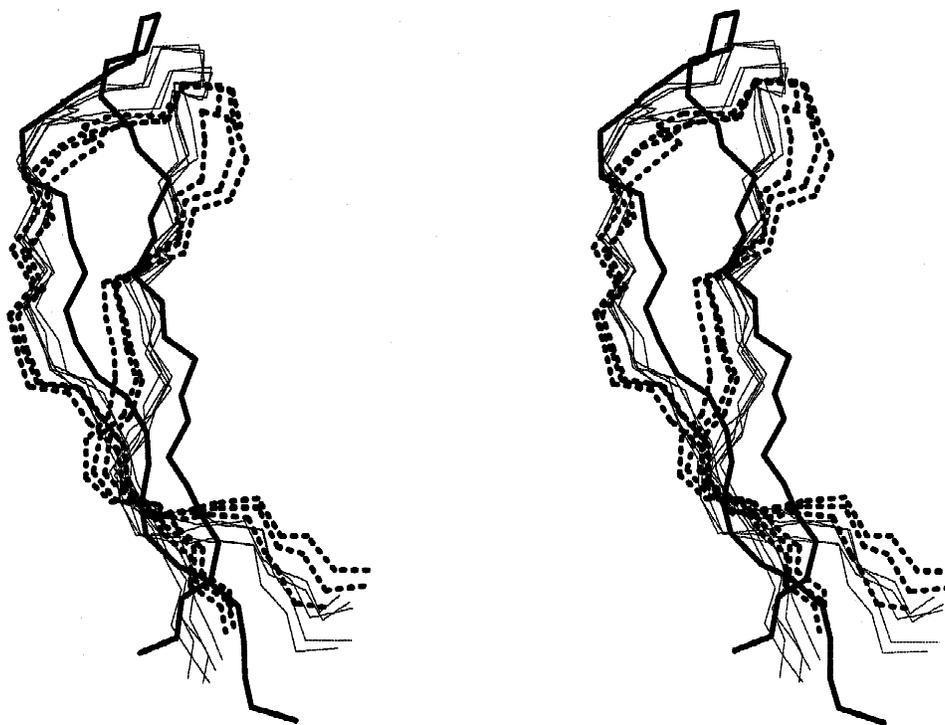


**Fig. 6.** Ramachandran plots of the backbone  $\phi, \psi$  angles of sampled conformations from (A) the last 8 ns of the 300 K, (B) the last 5 ns of the 350 K, and (C) from the 450 K simulations. The regions in the plots are approximately defined according to Efimov (see Sibanda et al., 1989).

that also in the case of the lower overlap (0.64), where the first nanosecond of the 280 K simulation is compared with the two 4 ns sections of the 300 K simulation, there is still much similarity in the structural fluctuations. For the 350 K and 450 K simulations, the comparisons with the two sections of the 300 K simulation show values larger than 0.7, indicating a very high overlap of the essential subspaces, and the lower overlap of the 450 K trajectory could be due to a slight change of the dynamical behavior at very high temperature (large fluctuations). Interestingly, the comparisons of the first part of the 280 K simulation with the 350 and 450 K simulations give higher MSIP values than those obtained from the comparison with the 300 K simulation. The reason for this behavior is probably the large fluctuations observed in the 350 K and especially in the 450 K dynamics. These fluctuations allow the peptide to sample more configurations close to the crystal structure, which are mostly present in the first part of the 280 K simulation.

The  $\beta$ -hairpin fragments extracted from the PDB (see Material and methods) were compared with structural fluctuations obtained

from the 300, 350, and 450 K simulations using the essential dynamics method. Essential dynamics has already been used to extract information from a collection of crystallographic structures (van Aalten et al., 1997). We use the same methodology to compare the MD essential subspaces with that obtained from the analysis of the ensemble of the crystallographic fragments. Also in this case we limited the analysis to 14 of the 16 residues, eliminating the terminal ones. The first three eigenvectors obtained from the ED analysis of the fragments contain 67% of the overall structural fluctuations. We compared these eigenvectors with the first ten eigenvectors obtained from the 300, 350, and 450 K simulations. In Figure 8 the values of the inner products of the three crystallographic eigenvectors with the first ten of the three simulations are reported. From Figure 8 it is clear that the first two crystallographic eigenvectors are mostly within the essential subspace obtained from the MD trajectories. Interestingly, as the temperature increases, the similarity of the first two crystallographic eigenvectors with the first two eigenvectors obtained from dynamics be-



**Fig. 7.** Stereo view of the superimposition of the average 300, 350 and 450 K conformations (dotted lines), native crystal structure (bold solid line), and the five  $\beta$ -hairpin protein fragments (thin solid lines) having low RMSD (see Table 2).

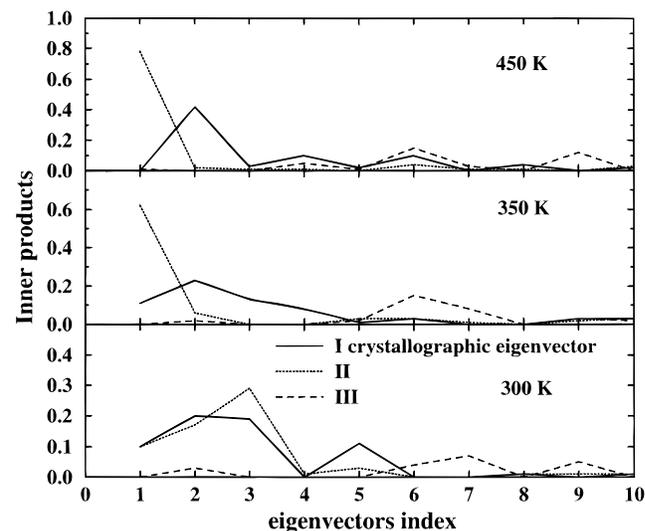
comes remarkable high. In Figure 9 the crystal structure with low RMSD from the average MD conformation have been projected onto the planes defined by the first two eigenvectors from the MD trajectories at the three temperatures. In the same planes the cor-

responding trajectories and the native structure are also shown. From these last figures, it is clear that the ensemble of structures obtained from the dynamics at 300 and 350 K are close to the five twisted  $\beta$ -hairpins conformations reported in Table 2.

At the highest temperature (450 K), the ensemble of structures obtained from MD is enlarged and encompasses most of the crystallographic configurations, including even the native  $\beta$ -hairpin II.

**Table 3.** Average  $\alpha_i$  values, obtained from the  $\phi$  and  $\psi$  values of residues 42–46 (strand I) and 51–55 (strand II), of MD configurations at the different temperatures, the five  $\beta$ -hairpin fragment having backbone RMSD values, with respect the three set of MD configurations, less than 0.15 nm and the native  $\beta$ -hairpin II conformation (Ipgb)

| PDB code | Strand I              |                     |                     | Strand II           |                     |                     |
|----------|-----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|          | $\alpha_1^a$<br>(deg) | $\alpha_2$<br>(deg) | $\alpha_3$<br>(deg) | $\alpha_1$<br>(deg) | $\alpha_2$<br>(deg) | $\alpha_3$<br>(deg) |
| 2bbk     | -117                  | 178                 | -125                | -131                | 183                 | -114                |
| 1bpi     | -116                  | 185                 | -106                | -112                | 161                 | -101                |
| 1aoc     | -99                   | 176                 | -134                | -120                | 166                 | -114                |
| 2ebn     | -122                  | 156                 | -131                | -137                | 159                 | -97                 |
| 1clc     | -129                  | 168                 | -115                | -113                | 167                 | -132                |
| Ipgb     | -165                  | 178                 | -157                | -162                | 184                 | -162                |
| 300 K    | -122                  | 130                 | -90                 | -103                | 141                 | -108                |
| 350 K    | -111                  | 150                 | -107                | -97                 | 157                 | -116                |
| 450 K    | -111                  | 147                 | -117                | -108                | 157                 | -130                |



**Fig. 8.** Inner product of the first three crystallographic eigenvectors with the first ten of the 300, 350, and 450 K simulations.

**Table 4.** H-bonds occurrences<sup>a</sup>

| Donor                     | Acceptor                    | %<br>(300 K) | %<br>(350 K) | %<br>(450 K) |
|---------------------------|-----------------------------|--------------|--------------|--------------|
| <b>Glu42 HN</b>           | <b>Thr55 O</b>              | <b>5</b>     | <b>9</b>     | <b>11</b>    |
| <b>Thr44 HN</b>           | <b>Thr53 O</b>              | <b>81</b>    | <b>92</b>    | <b>65</b>    |
| <b>Asp46 HN</b>           | <b>Thr49 O</b>              | <b>&lt;5</b> | <b>&lt;5</b> | <b>&lt;5</b> |
| <b>Asp46 HN</b>           | <b>Thr51 O</b>              | <b>86</b>    | <b>89</b>    | <b>69</b>    |
| <b>Ala48 HN</b>           | <b>Asp46 O<sup>δ1</sup></b> | <b>&lt;5</b> | <b>38</b>    | <b>15</b>    |
| <b>Lys50 HN</b>           | <b>Asp47 O</b>              | <b>6</b>     | <b>5</b>     | <b>11</b>    |
| <b>Thr51 HN</b>           | <b>Asp46 O</b>              | <b>8</b>     | <b>6</b>     | <b>18</b>    |
| <b>Thr53 HN</b>           | <b>Thr44 O</b>              | <b>86</b>    | <b>99</b>    | <b>64</b>    |
| <b>Thr55 HN</b>           | <b>Glu42 O</b>              | <b>71</b>    | <b>83</b>    | <b>51</b>    |
| Glu42 HN                  | Thr55 O <sup>γ1</sup>       | 55           | 42           | 16           |
| Trp43 HN                  | Glu42 O <sup>ε1</sup>       | 7            | 10           | 7            |
| Trp43 HN                  | Glu42 O <sup>ε2</sup>       | 12           | 11           | 7            |
| Thr44 HN                  | Thr51 O                     | <5           | <5           | <5           |
| Thr44 H <sup>γ1</sup>     | Glu42 O                     | 14           | 12           | 7            |
| Ala48 HN                  | Asp46 O <sup>δ2</sup>       | 15           | 15           | 18           |
| Ala48 HN                  | Asp47 O <sup>δ1</sup>       | 11           | <5           | <5           |
| Thr49 HN                  | Asp46 O <sup>δ1</sup>       | 10           | 65           | 34           |
| Thr49 HN                  | Asp46 O <sup>δ2</sup>       | 53           | 25           | 40           |
| Thr49 H <sup>γ1</sup>     | Asp46 O <sup>δ1</sup>       | 52           | 43           | 40           |
| Thr49 H <sup>γ1</sup>     | Asp46 O <sup>δ2</sup>       | 58           | 19           | 43           |
| Lys50 HN                  | Asp46 O                     | 17           | 20           | 24           |
| Lys50 HN                  | Asp46 O <sup>δ2</sup>       | 32           | <5           | 7            |
| Lys50 H <sup>ε1,2,3</sup> | Asp47 O                     | 26           | 17           | 20           |
| Thr51 HN                  | Thr44 O                     | <5           | <5           | <5           |
| Thr51 HN                  | Asp46 O <sup>δ1</sup>       | 56           | <5           | 11           |
| Thr51 HN                  | Asp46 O <sup>δ2</sup>       | 16           | <5           | 8            |
| Thr51 H <sup>γ1</sup>     | Asp46 O <sup>δ1</sup>       | 74           | <5           | 12           |
| Thr51 HN                  | Thr49 O                     | <5           | 36           | <5           |
| Thr51 H <sup>γ1</sup>     | Thr49 O <sup>γ1</sup>       | 13           | 20           | 29           |
| Thr53 HN                  | Glu42 O                     | <5           | <5           | <5           |
| Thr53 H <sup>γ1</sup>     | Asp46 O <sup>δ1</sup>       | 28           | 6            | 5            |
| Thr53 H <sup>γ1</sup>     | Asp46 O <sup>δ2</sup>       | <5           | 20           | <5           |
| Thr53 H <sup>γ1</sup>     | Thr51 O                     | 31           | 9            | <5           |
| Thr55 H <sup>γ1</sup>     | Thr44 O <sup>γ1</sup>       | <5           | 20           | <5           |

<sup>a</sup>The residues in bold characters indicate native H-bond interactions.

**Table 5.** MSIP values for different parts of MD trajectories at different temperatures

| Simulation             | 300 K | 280 K | 280 K | 280 K | 350 K | 450 K <sup>g</sup> |
|------------------------|-------|-------|-------|-------|-------|--------------------|
|                        | II    | I     | II    | III   |       |                    |
| 300 K I <sup>a</sup>   | 0.89  | 0.64  | 0.68  | 0.72  | 0.79  | 0.70               |
| 300 K II <sup>b</sup>  |       | 0.65  | 0.71  | 0.74  | 0.76  | 0.71               |
| 280 K I <sup>c</sup>   |       |       | 0.87  | 0.76  | 0.74  | 0.78               |
| 280 K II <sup>d</sup>  |       |       |       | 0.82  | 0.76  | 0.79               |
| 280 K III <sup>e</sup> |       |       |       |       | 0.78  | 0.75               |
| 350 K <sup>f</sup>     |       |       |       |       |       | 0.83               |

<sup>a</sup>2–6 ns

<sup>b</sup>6–10 ns.

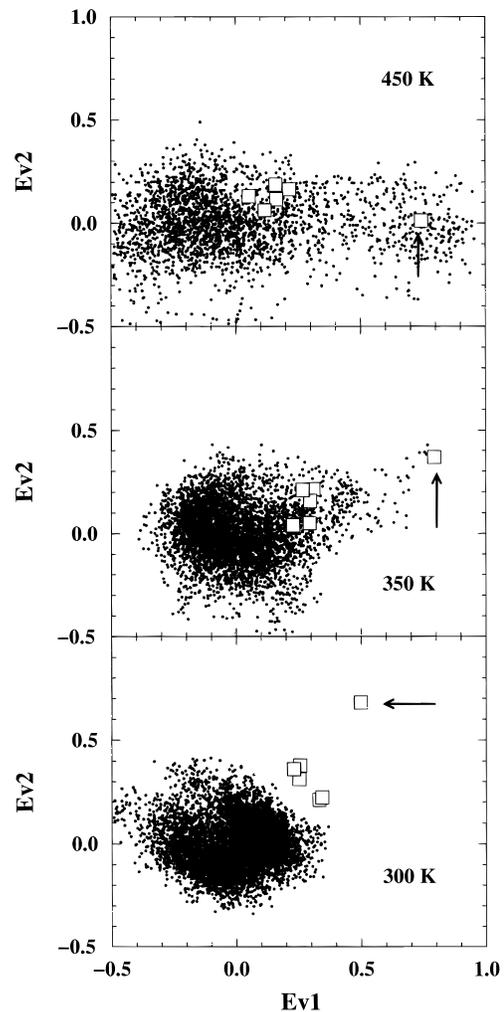
<sup>c</sup>0–1 ns.

<sup>d</sup>1–2 ns.

<sup>e</sup>2–3 ns.

<sup>f</sup>Last 5 ns.

<sup>g</sup>2.5 ns.

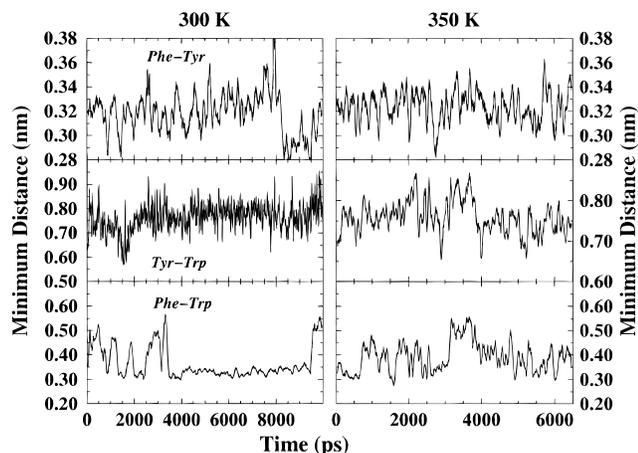


**Fig. 9.** Projections of samples of the MD trajectories at 300, 350 and 450 K onto the planes defined by the first two eigenvectors of each trajectory (dotted points). Squared points: the projections of five selected crystallographic  $\beta$ -hairpin fragments (see Table 2). The arrows indicate the position of the native  $\beta$ -hairpin II structure.

This last result indicates that although the essential subspace does not change significantly with temperature, the structural fluctuations in it can be extremely different at different temperatures. Such large structural fluctuations at higher temperature may be associated with an initial unfolding process that slightly changes the essential coordinates definition.

#### Comparison with experimental data

Experimental data on the stability of the  $\beta$ -hairpin II in solution were obtained by different authors (Blanco et al., 1994; Blanco & Serrano, 1995; Muñoz et al., 1997). First, Blanco et al. (1994) and Blanco and Serrano (1995) have assessed, with NMR techniques, the tendency of the peptide to maintain the  $\beta$ -hairpin conformation in solution. Later, Muñoz et al. (1997) have performed kinetic studies of the folding/unfolding rate of the same peptide confirming the NMR data and providing the value for the folding/unfolding rate constant.



**Fig. 10.** Minimal distances between aromatic side-chain H-atoms along the 300 and 350 K simulations. The curves are running averages for clarity.

The nuclear Overhauser effect spectroscopy (NOESY) experiments, in fact, revealed the presence of long-range nuclear Overhauser effects (NOEs) among the side chains of those residues (Blanco et al., 1994; Blanco & Serrano, 1995). Furthermore, the increase of the fluorescence of the Trp43 residue in the Muñoz et al. (1997) experiments is due to the stabilization of this residue in a hydrophobic cluster formed by three other residues (Tyr45, Phe52, Val54).

In our simulations we have investigated the behavior of the hydrophobic cluster to find a possible correspondence with the experimental results. In Figure 10 the minimal distances between the hydrogens of the aromatic side chains during the 300 and 350 K simulations are reported. The minimal distance between the aromatic H atoms in the Trp43 and Tyr45 residues is in agreement with the absence of NOE contacts observed experimentally (Blanco & Serrano, 1995; Blanco et al., 1996). Furthermore, the small distances between the Trp43 and Phe52 and Tyr45 agree well with the presence of long-range NOEs between these couples of residues. The short distances ( $<4 \text{ \AA}$ ) between Trp43–Val54 and Tyr45 and Phe52  $C_{\alpha}$  atoms occurring during all the simulations are also in agreement with the observed  $d_{\alpha\alpha}(i,j)$  NOEs.

In the central region of the peptide  $d_{NN}(i, i + 1)$ , NOEs of medium intensity were found. We verified the distances of these contacts: up to 350 K these distances were  $<4 \text{ \AA}$  for more than 90% of the simulation time (the only exception at 350 K was the Thr49NH–Lys50NH distance being  $<4 \text{ \AA}$  for 60% of the time).

#### Comparison with other MD simulations of $\beta$ -hairpin peptides

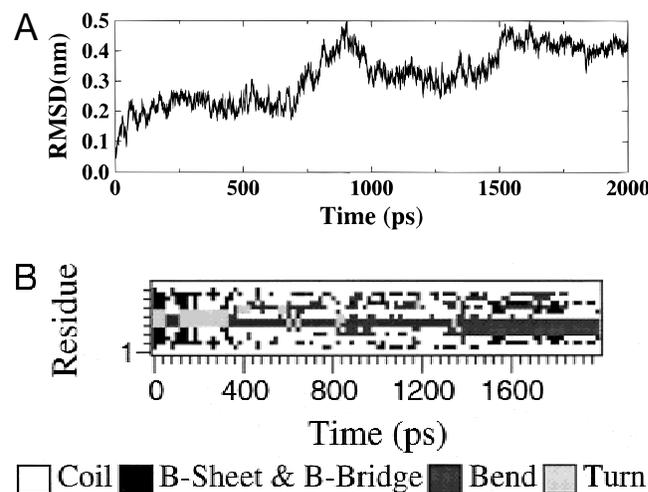
Our results can be qualitatively compared with those from simulations with explicit solvent of different  $\beta$ -hairpin peptides. In the case of the  $\beta$ -hairpin comprising residues 85–102 of barnase (Pugliese et al., 1995), the short time scale simulations (300 ps) cannot be used to draw any conclusions about the long-term stability of the peptide. However, it is interesting to note that at 450 K the peptide has a backbone RMSD comparable to that observed in our simulations (Pugliese et al., 1995). Furthermore, the refolding simulations of the same  $\beta$ -hairpin (Prévost & Ortmans, 1997) provides  $\beta$ -hairpin like conformations and the main factors in-

involved in the stabilization of the peptide were identified as the main-chain H-bonds and the hydrophobic interactions among side chains (Prévost & Ortmans, 1997). In our simulations, the partial conservation of the main-chain H-bonds and the agreement of the distances of H-contacts of hydrophobic residue side chain with NOE measurements indicate that these factors play an important role also in the present case. Preliminary results on the simulation at 350 K of one mutant having all the aromatic residues substituted with alanines have shown an evident destabilization of the structure and the starting of unfolding after only 800 ps. In Figure 11 the backbone RMSD with respect to the native conformation and the secondary structure during the simulation are reported.

The nine residue  $\beta$ -hairpin studied by Constantine et al. (1996), using long time scale MD simulations (10 ns) at low temperature (278 K) is less stable than our peptide, as it unfolds after about 2 ns. The authors explain this lack of stability as a possible consequence of a not well-optimized force field. We are investigating the effect of the force field and MD protocols performing long time scale simulations of both the peptides. In the case of  $\beta$ -hairpin from barnase preliminary results are in good agreement with simulations reported in the literature.

#### Density distribution of the water molecules around the peptide at the 300 K simulation

The average number of water molecules inside a 0.35 nm thick shell (first hydration shell) was  $89 \pm 5$ , and they form  $75 \pm 3$  H-bonds on average with the peptide. The average positions of the water molecules involved in these H-bonds were obtained by calculating the local density of water molecules around the peptide. This approach was recently used by other authors to study the hydration of an helical peptides (Hummer et al., 1996; Garcia et al., 1997) and a  $\beta$ -hairpin from bovine pancreatic trypsin inhibitor in MD simulations (Hummer et al., 1996). The average number of bulk water molecules at 300 K in a  $1 \text{ \AA}$  grid is  $\rho_0 = 0.033$  molecules. In the high density regions, the water density was up to



**Fig. 11.** **A:** Backbone RMSD deviation from  $\beta$ -hairpin II crystal structure during the 350 K simulation of a mutant with all the aromatic residues changed into alanines. **B:** Secondary structure during the simulation.

four times larger than  $\rho_0$ . In Figure 12 the stereo plot of the isosurfaces of the regions having densities  $\sim 2.5$  times larger than the bulk, together with the average peptide conformation, is shown. It is possible to identify four different large blobs. The first and third are located in the part of the molecule where the aromatic side chains are exposed; the water molecules interact with the NH groups of the Phe52 and Val54 residues. The second is localized close to the loop region and involves interactions with the side chains of residues Thr49, Thr51, and Asp46. Finally, the fourth blob is localized in the groove formed by the coiling of the structure between the Thr44 and Thr53. The other smaller blobs are close to polar side-chain residues. In the crystal structure, the  $\beta$ -hairpin presents two high resolution water molecules localized in the threonine-rich outer surface of the  $\beta$ -sheet (Gallagher et al., 1994). The presence of water molecules with long residence times around this region was shown by NMR studies (Clare & Gronenborn, 1992). The stability of these water molecules is a consequence of the favorable interactions with the main-chain H-bonds and the O $\gamma$  atoms of the three threonines. The blob structures in Figure 12 indicate that these hydration sites are present in the simulated peptides.

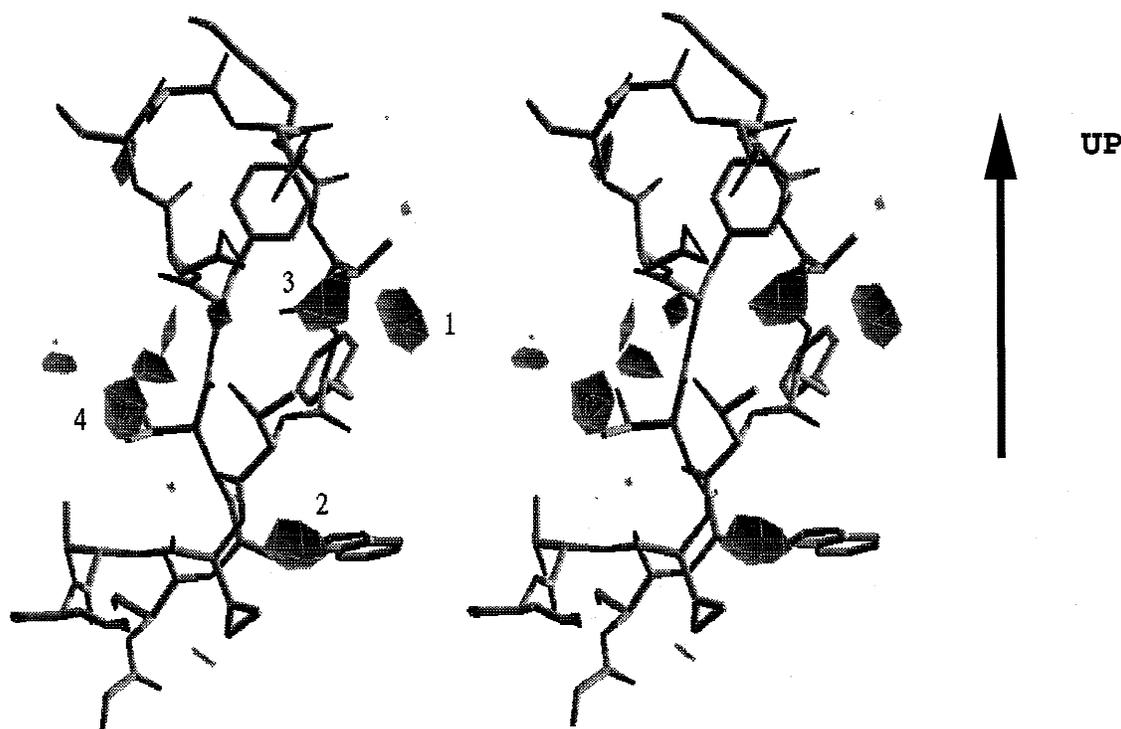
### Conclusion

Molecular dynamics simulations of the 41–56  $\beta$ -hairpin from protein GB1 at different temperatures showed that the peptide in solution loses its crystal conformation as a consequence of the excision of the peptide from the protein. The new equilibrium conformations observed at the different temperatures are very similar. These structures are more compact than the native one and

conserve part of the strand conformation in the aromatic residues region. Furthermore, the topological properties of the strands (twisting and coiling) are similar to those of highly twisted  $\beta$ -hairpin fragments of the same type present in other proteins. The aromatic side chains of the residues in this region show H-contacts in agreement with NOE measurements.

The hydration analysis of the 300 K equilibrium conformation shows that the threonine rich region maintains its solvation state as in the crystal structure. Finally, the essential dynamics analysis showed that the dynamical behavior of the peptide at the different temperatures, once equilibrated, is very similar and characterized by large motions of the turn and end residues. Remarkably, up to 450 K the essential subspace is basically unaltered although with larger eigenvalues. The comparison with the set of  $\beta$ -hairpin crystal fragments showed that these are mainly distributed in such essential subspace, but in a region which is reached only by the high temperature (450 K) dynamics of the  $\beta$ -hairpin studied. This observation can be considered a sort of further indication of the  $\beta$ -hairpin character of the peptide in solution.

The results of these simulations provide further evidence of the importance of H-bonds and hydrophobic interactions on the stability of  $\beta$ -hairpin forming peptides. These results are in agreement with the experimental data (Blanco et al., 1994, 1996; Muñoz et al., 1997) and theoretical model on the same peptide (Muñoz et al., 1997, 1998). However, many open questions have to be answered. In particular, the effect of mutations on the stability of these structures could be studied to assess the role of the residue type on the stability in analogy to the MD study of the whole protein GB1 (Ceruso et al., 1999). Furthermore, the characterization of the unfolding pathways also can provide further insights to



**Fig. 12.** Stereo view representation of the water density isosurfaces of the regions having an average density 2.5 times the bulk one together with the average peptide structure from the last 2 ns of the 300 K simulation.

the mechanism governing the peptide stability. In this sense, these long time scale simulations can be used as a reference for future investigations of these questions.

## Materials and methods

### Molecular dynamics simulations

The starting structure for our simulations (Fig. 1) was taken from the 0.18 nm resolution refined crystal structure of the protein GB1 (PDB entry 1pgb) (Gallagher et al., 1994). The C-terminal part of the protein GB1, starting from the residues Gly41 was cut out. The fragment was protonated to give a zwitterionic form (with N-terminal  $\text{NH}_3^+$  and C-terminal  $\text{COO}^-$  groups) to reproduce the experimental conditions at which the peptide was studied (Blanco & Serrano, 1995; Blanco et al., 1996). The peptide was solvated in a rectangular box with water, by stacking equilibrated boxes of solvent molecules to form a box of  $3.1 \times 3.6 \times 3.9$  nm, large enough to contain the peptide and 0.8 nm of solvent on all sides. All solvent molecules with any atom within 0.15 nm of the peptide were removed. Since the resulting protonation state (at pH 7) of the peptide gives a total charge of  $-3$ , sodium counterions were added by replacing water molecules at the most negative electrical potential to provide a neutral simulation cell. The resulting system was composed of the peptide, 1,414 water molecules, and three sodium counterions subjected to periodic boundary conditions. The system was subsequently energy minimized with a steepest descent method for 100 steps.

To compare the dynamical behavior of the peptide at different temperatures, four simulations at 450, 350, 300, and 280 K were carried out. All the MD simulations were performed using an isothermal-isochoric simulation algorithm (Berendsen et al., 1984). The temperature was kept constant to the reference values by weak coupling to an external temperature bath with a coupling constant of 0.1 ps. The peptide and the rest of the system were coupled separately to the temperature bath. The Gromos87 force field (van Gunsteren & Berendsen, 1987) was used with modification as suggested in van Buuren et al. (1993) and explicit hydrogen atoms in aromatic rings (van der Spoel et al., 1996). For the solvent the SPC/E (Berendsen et al., 1987), water model was used. The SHAKE algorithm (Ryckaert et al., 1977) was used to constrain all bond lengths. For the water molecules, we used the SETTLE algorithm to constrain the bond lengths as well as the bond angle (Miyamoto & Kollman, 1992). A dielectric permittivity,  $\epsilon = 1$ , and a time step of 2 fs were used. The cutoff radius for the nonbonded interactions was set to 1.2 nm. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial temperature. The density of the system was adjusted performing the first equilibration runs at NPT condition by weak coupling to a bath of constant pressure ( $P_0 = 1$  bar, coupling time  $\tau_P = 0.5$  ps) (Berendsen et al., 1984). In the case of the 450 K simulations, the density was fixed to a 10% lower value to reduce the increase of pressure that could affect the results of the simulation (Daggett & Levitt, 1993; Finkelstein, 1997).

All the simulations, starting from the crystallographic structure, were equilibrated by 50 ps of MD runs with position restraints on the peptide to allow relaxation of the solvent molecules. These first equilibration runs were followed by other 50 ps runs without position restraints on the peptide. The production runs, after equilibration, were 4, 6.5, 10, and 3 ns long for the 450, 350, 300, and 280 K simulations, respectively.

All the MD runs and the analysis of the trajectories were performed using the GROMACS software package (van der Spoel et al., 1994) on a Silicon Graphics Power Challenge.

The local density of the solvent molecules around the peptide was calculated in a cubic grid with 0.1 nm grid spacing. The analysis was performed using the last 2 ns of the 300 K simulation. The procedure consisted of a translational-rotational fit of the  $\alpha$ -carbon atoms of each peptide configuration with a reference configuration (we used the first configuration of the trajectory used for the analysis). Then the obtained rotation matrices were applied to all oxygen atoms of the solvent molecules surrounding the peptide conformations inside a 0.5 nm thick shell. Finally the oxygen positions of the selected water molecules were registered with respect to a cubic grid centered on the geometric center of the reference peptide. The density matrix obtained was analyzed using the program SciAn (Pepke & Lyons, 1993a, 1993b).

### Essential dynamics

The conformational space spanned by a macromolecule during its equilibrium dynamics can be evaluated performing a principal component analysis on the atomic coordinates, ED analysis (Amadei et al., 1993). This method allows the characterization of a configurational subspace, the "essential subspace," in which the principal protein motions occur. The essential subspace obtained from this analysis is defined by an orthonormal set of vectors, which are the eigenvectors with the largest eigenvalues of the atomic positional fluctuations covariance matrix. In proteins and peptides, the first 10 eigenvectors give a good representation of the essential subspace (Amadei et al., 1993; de Groot et al., 1996). The comparison of essential subspaces of different simulated systems can provide a valid method to assess their dynamical similarity. A good estimate of the overlap of two essential subspaces can be obtained by using the MSIP (de Groot et al., 1996) defined by

$$\text{MSIP} = \frac{1}{N} \sum_i^N \sum_j^N (\xi'_j \cdot \xi_i)^2, \quad (2)$$

where  $\xi_i$  and  $\xi'_j$  are the  $i^{\text{th}}$  and  $j^{\text{th}}$  eigenvectors of the two sets and  $N$  is the essential subspace dimension. The ED analyses were performed on the Cartesian coordinates of the  $\text{C}_\alpha$  atoms as described elsewhere (Amadei et al., 1993) using the WHATIF program (Vriend, 1990).

### Selection of $\beta$ -hairpin peptide fragments from PDB

The equilibrium average peptide backbone conformations obtained from the simulations were compared with those of analogous peptide fragments present in the crystal structure of different proteins. The searching procedure of those peptides was performed using the WHATIF program. Using the SCAN3D option of the program, all possible fragments having the same length (16 residues), strand conformations at residues 3–5 and 13–15, turn or bend conformations at residues 8–10, and arbitrary conformation at the remaining residues, were selected. This first selection resulted in 142 structures. Then we refined the search results, eliminating all fragments that did not have a  $\beta$ -hairpin structure. After this selection, 47 hairpins were obtained. Table 1 reports a summary of these fragments.

### Note added in proof

Two recent papers have appeared in the literature describing molecular dynamics and Monte Carlo simulations of the 41–56  $\beta$ -hairpin from protein G: Pande SV, Rokhsar DS. 1999. Molecular dynamics simulations of unfolding and refolding of a  $\beta$ -hairpin fragment of protein G. *Proc Natl Acad Sci USA* 96:9062–9067; and Dinner AR, Lazaridis T, Karplus M. 1999. Understanding  $\beta$ -hairpin formation. *Proc Natl Acad Sci USA* 96:9068–9073.

### Acknowledgments

This work was supported by the EC TMR Network Project No. ERBFMRXCT960013 and by Italian Ministero dell'Università e della Ricerca Scientifica e Tecnologica, MURST (National Project "Structural Biology"). A.A. thanks Istituto Pasteur Fondazione Cenci Bolognetti for financial support.

### References

- Amadei A, Linssen ABM, Berendsen HJC. 1993. Essential dynamics of proteins. *Proteins Struct Funct Genet* 17:412–425.
- Berendsen HJC, Grigera JR, Straatsma TP. 1987. The missing term in effective pair potentials. *J Phys Chem* 91:6269–6271.
- Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690.
- Bernstein FC, Koetzle TF, Williams GLB, Meyer JEF, Brice MD, Rodgers JR, Kennard O, Schimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol* 112:535–542.
- Blanco FJ, Jiménez MA, Pineda A, Rico M, Santoro J, Nieto JL. 1994. NMR solution structure of the isolated N-terminal fragment of protein G B1 domain. Evidence of trifluoroethanol induced native-like  $\beta$ -hairpin formation. *Biochemistry* 33:600–6014.
- Blanco F, Ramirez-Alvarado M, Serrano L. 1998. Formation and stability of  $\beta$ -hairpin structures in polypeptides. *Curr Opin Struct Biol* 8:107–111.
- Blanco F, Rivas G, Serrano L. 1996. A short linear peptide that folds into a native stable  $\beta$ -hairpin in aqueous solution. *Nat Struct Biol* 1:584–590.
- Blanco F, Serrano L. 1995. Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur J Biochem* 230:634–649.
- Braxenthaler M, Unger R, Auerbach D, Given JA, Moulton J. 1997. Chaos in protein dynamics. *Proteins Struct Funct Genet* 29:417–425.
- Ceruso M-A, Amadei A, Di Nola A. 1999. Mechanics and dynamics of B1 domain of protein G: Role of packing and surface hydrophobic residues. *Protein Sci* 8:147–160.
- Chothia C. 1973. Conformations of  $\beta$ -pleated sheets in proteins. *J Mol Biol* 75:295–302.
- Chothia C. 1983. Coiling of  $\beta$ -pleated sheets. *J Mol Biol* 163:107–117.
- Chou K-C, Nemethy G, Sheraga HA. 1983a. The effect of amino acid composition on the twist and relative stability of parallel and antiparallel  $\beta$ -sheets. *Biochemistry* 22:6213–6221.
- Chou K-C, Nemethy G, Sheraga HA. 1983b. The role of inter-chain interaction in the stabilization of the right-handed twist of  $\beta$ -sheet. *J Mol Biol* 168:389–407.
- Clore G, Gronenborn A. 1992. Localization of bound water in the solvation structure of immunoglobulin binding domain of streptococcal protein G. Evidence for solvent-induced helical distortion in solution. *J Mol Biol* 223:853–856.
- Constantine KL, Friedrichs MS, Stouch TR. 1996. Extensive molecular dynamics simulations of a  $\beta$ -hairpin forming peptide. *Biopolymers* 39:591–614.
- Daggett V, Levitt M. 1993. Protein unfolding pathways explored through molecular dynamics simulations. *J Mol Biol* 232:600–619.
- Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE. 1998. Reversible peptide folding in solution by molecular dynamics simulation. *J Mol Biol* 280:925–932.
- de Alba E, Jiménez MA, Rico M. 1997. Turn residue sequence determines  $\beta$ -hairpin conformation in designed peptides. *J Am Chem Soc* 119:175–183.
- de Groot BL, van Aalten DMF, Amadei A, Berendsen HJC. 1996. The consistency of large concerted motions in proteins in molecular dynamics simulations. *Biophys J* 71:1007–1713.
- Dobson CM, Sali A, Karplus M. 1998. Protein folding: A perspective from theory and experiment. *Angew Chem Int Ed* 37:869–893.
- Dyson HJ, Wright PE. 1993. Peptide conformation and protein folding. *Curr Opin Struct Biol* 3:60–65.
- Finkelstein AV. 1997. Can protein unfolding simulate protein folding? *Protein Eng* 10:843–845.
- Gallagher T, Alexander P, Bryan P, Gilliland G. 1994. Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. *Biochemistry* 33:4721–4729.
- Garcia AE, Hummer G, Soumpasis DM. 1997. Hydration of an  $\alpha$ -helical peptide: Comparison of theory and molecular dynamics simulation. *Proteins Struct Funct Genet* 27:471–480.
- Hummer G, Garcia AE, Soumpasis DM. 1996. A statistical mechanical description of biomolecular hydration. *Faraday Discuss* 103:175–189.
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2576–2637.
- Klimov DK, Thirumalai D. 1997. Viscosity dependence of the folding rates of proteins. *Phys Rev Lett* 79:317–320.
- Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 24:946–950.
- Levitt M. 1976. A simplified representation of protein conformations for rapid simulation of protein folding. *J Mol Biol* 104:59–107.
- Maccallum PH, Poet R, Milner-White EJ. 1995. Coulombic attractions between partially charged main-chain atoms stabilise the right-handed twist found in most  $\beta$ -strands. *J Mol Biol* 248:374–384.
- Maynard AJ, Sharman GJ, Searle MS. 1998. Origin of  $\beta$ -hairpin stability in solution: Structural and thermodynamic analysis of the folding of a model peptide supports hydrophobic stabilization in water. *J Am Chem Soc* 120:1996–2007.
- Miyamoto S, Kollman PA. 1992. SETTLE: An analytical version of the shake and rattle algorithms for rigid water models. *J Comp Chem* 13:952–962.
- Muñoz V, Henry ER, Hofrichter J, Eaton WA. 1998. A statistical mechanical model for  $\beta$ -hairpin kinetics. *Proc Natl Acad Sci USA* 95:5872–5879.
- Muñoz V, Thompson PA, Hofrichter J, Eaton WA. 1997. Folding dynamics and mechanism of  $\beta$ -hairpin formation. *Nature* 390:196–199.
- Pepke E, Lyons J. 1993a. *SciAn user's manual*. Tallahassee, Florida: Supercomputer Computations Research Institute, Florida State University.
- Pepke E, Lyons J. 1993b. *SciAn reference manual*. Tallahassee, Florida: Supercomputer Computations Research Institute, Florida State University.
- Prévost M, Ortmans I. 1997. Refolding simulations of an isolated fragment of barnase into a native-like  $\beta$ -hairpin: Evidence for compactness and hydrogen bonding as concurrent stabilizing factors. *Proteins Struct Funct Genet* 29:212–227.
- Pugliese L, Prévost M, Wodak SJ. 1995. Unfolding simulations of the 85–102  $\beta$ -hairpin of barnase. *J Mol Biol* 251:432–447.
- Ramirez-Alvarado M, Blanco FJ, Serrano L. 1996. De novo design and structural analysis of a model  $\beta$ -hairpin peptide system. *Nature Struct Biol* 3:604–612.
- Ryckaert JP, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J Comp Phys* 23:327–341.
- Salemme FR. 1983. Structural properties of protein  $\beta$ -sheets. *Prog Biophys Mol Biol* 42:95–133.
- Schaefer M, Bartels C, Karplus M. 1998. Solution conformations and thermodynamics of structured peptides: Molecular dynamics simulation with implicit solvation model. *J Mol Biol* 284:835–848.
- Searle MS, Williams DH, Packman LC. 1995. A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non native  $\beta$ -hairpin. *Nature Struct Biol* 2:999–1006.
- Sibanda BL, Blundell TL, Thornton JM. 1989. Conformation of  $\beta$ -hairpin in protein structures. A systematic classification with applications to modeling by homology, electron density fitting and protein engineering. *J Mol Biol* 206:759–777.
- Sibanda BL, Thornton JM. 1991. Conformation of  $\beta$  hairpins in protein structures: Classification and diversity in homologous structures. *Methods Enzymol* 202:59–82.
- Sieber V, Moe G. 1996. Interactions contributing to the formation of a  $\beta$ -hairpin-like structure in a small peptide. *Biochemistry* 35:181–188.
- Thompson PA, Eaton WA, Hofrichter J. 1997. Laser temperature jump study of the helix  $\rightleftharpoons$  coil kinetics of an alanine peptide interpreted with a "kinetic zipper" model. *Biochemistry* 36:9200–9210.
- van Aalten DMF, Conn DA, de Groot BL, Berendsen HJC, Findlay JBC, Amadei A. 1997. Protein dynamics derived from clusters of crystal structures. *Biophys J* 73:2891–2896.
- van Buuren AR, Marrink SJ, Berendsen HJC. 1993. A molecular dynamics study of the decane/water interface. *J Phys Chem* 97:9206–9212.
- van der Spoel D, van Buuren AR, Tieleman DP, Berendsen HJC. 1996. Molec-

- ular dynamics simulations of peptides from BPTI: A closer look at amide-aromatic interactions. *J Biomol NMR* 8:229–238.
- van der Spoel D, van Drunen R, Berendsen HJC. 1994. *GROningen MAchine for chemical simulations*. Department of Biophysical Chemistry, BIOSON Research Institute Nijenborgh 4 NL-9717 AG Groningen. e-mail: gromacs@chem.rug.nl.
- van Gunsteren WF, Berendsen HJC. 1987. *Gromos-87 manual*. Biomos BV Nijenborgh 4, 9747 AG Groningen, The Netherlands.
- Vriend G. 1990. WHAT IF: A molecular modeling and drug design program. *J Mol Graph* 8:52–56.
- Williams S, Causgrove TP, Gilmanshin R, Fang KS, Callender RH, Woodruff WH, Dyer RB. 1996. Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* 35:691–697.
- Yapa K, Weaver DL, Karplus M. 1992.  $\beta$ -sheet coil transition in a simple polypeptide model. *Proteins Struct Funct Genet* 12:237–265.