

# Folding and Stability of the Three-Stranded $\beta$ -Sheet Peptide Betanova: Insights From Molecular Dynamics Simulations

Giorgio Colombo,<sup>1,2</sup> Danilo Roccatano,<sup>1,3</sup> and Alan E. Mark<sup>1\*</sup>

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

<sup>2</sup>Istituto di Biocatalisi e Riconoscimento Molecolare, CNR via Mario Bianco, Milano, Italy

<sup>3</sup>Department of Chemistry, University of Rome "La Sapienza" Rome, Italy

**ABSTRACT** The dynamics of the three-stranded  $\beta$ -sheet peptide Betanova has been studied at four different temperatures (280, 300, 350, and 450 K) by molecular dynamics simulation techniques, in explicit water. Two 20-ns simulations at 280 K indicate that the peptide remains very flexible under "folding" conditions sampling a range of conformations that together satisfy the nuclear magnetic resonance (NMR)-derived experimental constraints. Two simulations at 300 K (above the experimental folding temperature) of 20 ns each show partial formation of "native"-like structure, which also satisfies most of the NOE constraints at 280 K. At higher temperature, the presence of compact states, in which a series of hydrophobic contacts remain present, are observed. This is consistent with experimental observations regarding the role of hydrophobic contacts in determining the peptide's stability and in initiating the formation of turns and loops. A set of different structures is shown to satisfy NMR-derived distance restraints and a possible mechanism for the folding of the peptide into the NMR-determined structure is proposed. *Proteins* 2002; 46:380–392. © 2002 Wiley-Liss, Inc.

**Key words:** protein folding; molecular dynamics;  $\beta$ -sheet; peptide conformation in water; Betanova

## INTRODUCTION

To carry out their functions within living cells, proteins must fold to a well-defined three-dimensional (3-D) structure. Understanding the mechanism of peptide or protein folding is thus important from an experimental as well as a theoretical perspective. Much effort has been devoted to understanding how proteins attain their native structure, and a wide variety of experimental techniques have been applied to this problem. These range from spectroscopic techniques, such as circular dichroism (CD) and nuclear magnetic resonance (NMR), to experiments in which laser temperature jumps are used to synchronize folding.<sup>1–3</sup> Despite this sophistication, there is still no experimental technique that can provide insight at an atomic level into the folding process itself. To understand folding at an atomic level, we have little choice but to turn to theoretical approaches.

Theoretical methods to predict the folded structure of a peptide or a protein fall into three main classes: (a) statistical approaches, (b) conformational search techniques, and (c) molecular dynamics (MD) simulations. Statistical approaches relate an amino acid sequence to known 3-D structures and are widely used to predict the conformation of peptides in solution.<sup>4,5</sup> They provide, however, no information regarding the mechanism of folding nor any insight into possible alternative conformations. Conformational search methods can, in contrast, provide information on the relative population of different conformations in solution. A range of conformations is generated, and an energy function is used to discriminate between them.<sup>6,7</sup> However, the size of the conformational space potentially accessible to even a small peptide precludes any systematic search at atomic detail. Moreover, these methods cannot be used to characterize the dynamics of the folding process.

To characterize both the range of alternative states that can be sampled by a peptide under specific conditions and the dynamics of the process of (un)folding, we must turn to equilibrium MD simulations. MD simulations address the folding problem by only sampling "relevant" (low-energy) regions of conformational space. In cases that are characterized by a limited number of low-energy conformations in rapid equilibrium (on the MD timescale), MD simulations have been demonstrated to provide highly detailed information on the nature of alternative states in solution, and the dynamics of spontaneous folding and unfolding.<sup>8–14</sup> In particular, Daura et al.<sup>15–17</sup> have been able to demonstrate the reversible folding of a helix forming  $\beta$ -peptide in methanol from an arbitrary starting structure under a variety of conditions within 50 ns. These studies clearly showed how the study of small peptides could help better understand the nature of the unfolded state and the mechanism of protein folding. A variety of other computational studies on helix-forming peptides have also ap-

---

Grant sponsor: EC TMR Network Project "Structure and Dynamics of Intermediate States in Protein Folding"; Grant number: ERBFMRXCT960013.

\*Correspondence to: Alan E. Mark, Groningen Biomolecular Science's and Biotechnology, Department of Biophysical Chemistry, University of Groningen, Nijenborah 4, 9747 AG Groningen, The Netherlands. E-mail: mark@chem.rug.nl

Received 21 December 2000; Accepted 4 June 2001

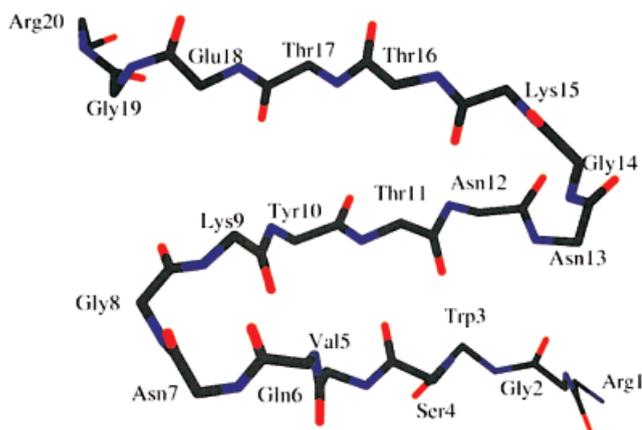


Fig. 1. Scheme depicting the backbone structure of the NMR model and residue types in Betanova. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

peared, and factors affecting helix formation and stability, at least in general terms, are well characterized.<sup>12,13,18,19</sup>

In contrast,  $\beta$ -sheets are complex structures. The connections between the  $\beta$ -strands can be well separated along the sequence, and the interactions that stabilize  $\beta$ -sheet formation are less well understood.<sup>20</sup> Although,  $\beta$ -hairpins, along with  $\beta$ -turns, have been proposed to act as initiation sites in early protein folding events,<sup>21,22</sup> the folding properties of  $\beta$ -hairpins have proved difficult to study experimentally primarily because of their tendency to aggregate. Nevertheless, in recent years,  $\beta$ -sheet formation has been the subject of intense research because of its importance in amyloid fibril formation, a factor in a wide variety of pathological disorders.<sup>23,24</sup> Peptides obtained from proteins<sup>25,26</sup> or from de novo design<sup>27–29</sup> which fold stably in solution have been isolated and characterized by using NMR, fluorescence, and other spectroscopic techniques. These studies indicate that  $\beta$ -hairpins have folding times in the range of several microseconds, thus much longer than the respective helix folding times. Computational studies of  $\beta$ -turn forming sequences,<sup>30–32</sup> as well as  $\beta$ -hairpin forming sequences,<sup>33–40</sup> have addressed different aspects of  $\beta$ -hairpin stability, such as the importance of the turn sequence, hydrogen bonding patterns and hydrophobic interactions.

To garner further insight into the atomic details of the folding mechanism of  $\beta$ -sheets, we have studied the folding and stability of the 20-residue designed peptide Betanova<sup>41</sup> by using MD simulation techniques. Betanova, whose sequence is RGWSVQNGKYTNNGKTTEGR was designed to form a monomeric, three-stranded, antiparallel  $\beta$ -sheet [Figs. 1 and 2(a)]. NMR spectroscopic data are consistent with Betanova containing two turns (residues Asn7-Gly8, and Asn13-Gly14) and three  $\beta$ -strands. Spectroscopic and physicochemical characterization has shown that the  $\beta$ -sheet conformation is stabilized by specific tertiary interactions and that the molecule shows evidence of a cooperative two-state folding-unfolding transition, which is a distinctive feature of natural proteins. Thermodynamically, the stability of Betanova is around 2.5–3.0

kJ/mol at 278 K. Betanova thus constitutes a useful model to help understand the mechanism of  $\beta$ -sheet folding and stability. Computationally, Bursulaya and Brooks<sup>36</sup> simulated Betanova for 2 ns in explicit water by using the CHARMM force field. The peptide remained close to the starting structure, which indicated some degree of stability. Using unfolding conditions and umbrella sampling techniques, these workers generated a free-energy map for folding and unfolding and highlighted the importance of intramolecular contacts in defining the folded state of the peptide.

Van der Vaart and coworkers<sup>42</sup> have investigated the role of many body effects in the stabilization of Betanova and concluded that, although many body effects are in principle important, the folding of Betanova could be accurately described by using effectively two body potentials. Very recently (after the completion of this manuscript) Ferrara and Caffisch<sup>37</sup> published extensive simulations on a closely related peptide by using a simplified solvent model and short interatomic cutoffs. To their credit, these workers were able to demonstrate folding from random starting conformations. At the same time, the work highlighted the problems associated with the use of implicit solvent models. Although the peptide folded, significant violations of experimentally derived nuclear Overhauser effect (NOE) constraints were still found despite very long simulations (200 ns) and  $\langle r^{-6} \rangle^{-1/6}$  averaging of the interproton distances  $r$ . Second, the simulations severely overestimated the stability of the peptide. The peptide rapidly folded at 360 K despite experimentally the peptide was at most 50% folded at 280 K.<sup>29</sup> Nevertheless, the authors claimed general agreement with the studies of Bursulaya and Brooks and used the simulations to draw general conclusions for the folding of Betanova-like molecules.

Here, we present the results of a study of the stability and folding of Betanova using MD simulations in explicit water under periodic boundary conditions at several temperatures (280, 300, 350, and 450 K). Two 20-ns simulations were performed at 280 K (40 ns total), and two 20 ns simulations at 300 K (40 ns total). The results of these simulations are compared with experimentally derived NMR structural data. At both 350 and 450 K, a single 20-ns trajectory was generated and analyzed. The simulations (120 ns total) show extensive sampling of the conformational space around the folded state and allow us to address issues related to the stability and folding of Betanova. Based on the analysis of the trajectories, the respective roles of single amino acids, loops, hydrogen bonding, and side-chain interactions in determining the folding-unfolding mechanism of Betanova are discussed. Moreover, a detailed analysis of the ability of the simulations to reproduce NMR-derived distance restraints is presented.

## MATERIALS AND METHODS

### Molecular Dynamics Simulations and Analysis

The starting structure for the simulations (Fig. 1) was the averaged NMR structure of Kortemme et al.<sup>41</sup> The

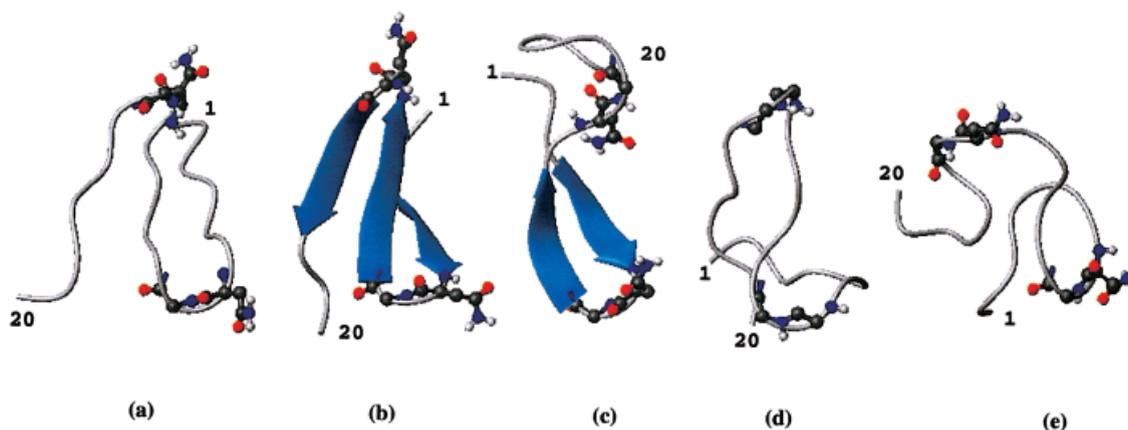


Fig. 2. Tertiary structure of the backbone of Betanova as obtained from: (a) the average between 20 NMR structures, (b) the most populated cluster at 280 K, (c) the most populated cluster for 300A, (d) the most populated cluster for 300B, (e) the most populated cluster for 350 K. The pictures were obtained by using MOLMOL.<sup>59</sup> The secondary structure definitions were based on Kabsch-Sander DSSP definition.<sup>43</sup> Residues involved in turn formation Asn7, Gly8, Asn13, and Gly14 are represented as ball and stick. The numbers indicate the initial and terminal residue. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

peptide was protonated to give a zwitterionic form (with N-terminal  $\text{NH}_3^+$  and C-terminal  $\text{COO}^-$  groups) in line with the experimental conditions (pH 5.0) at which the peptide was studied.<sup>41</sup> The peptide was solvated with water in a periodic truncated octahedron large enough to contain the peptide and 0.8 nm of solvent on all sides. All solvent molecules within 0.15 nm of any peptide atom were removed. The total charge on the peptide was +3. No counterions were added because water is a high dielectric and the inclusion of no counterions was considered a better approximation to the low-salt experimental conditions. The resulting system was composed of 220 peptide atoms and 2,940 water molecules. 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, simulations at 280, 300, 350, and 450 K were performed. In all simulations, the temperature was maintained close to the intended values by weak coupling to an external temperature bath<sup>51</sup> with a coupling constant of 0.1 ps. The peptide and the rest of the system were coupled separately to the temperature bath. The GROMOS96 force field<sup>52,53</sup> was used. The simple point charge (SPC)<sup>54</sup> water model was used. The LINCS algorithm<sup>55</sup> was used to constrain all bond lengths. For the water molecules, the SETTLE algorithm<sup>56</sup> was used. A dielectric permittivity ( $\epsilon = 1$ ) and a time step of 2 fs were used. A twin range cutoff was used for the calculation of the nonbonded interactions. The short-range cutoff radius was set to 0.8 nm, and the long-range cutoff radius was set to 1.4 nm for both coulombic and Lennard-Jones interactions. The cutoff values are the same as those used for the GROMOS96 force field parameterization.<sup>52</sup> Interactions within the short-range cutoff were updated at every time step, whereas interactions within the long-range cutoff were updated every five time steps together with the pairlist. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial

temperature. The density of the system was adjusted by performing the first equilibration runs under NPT condition by weak coupling to a bath of constant pressure ( $P_0 = 1$  bar, coupling time  $\tau_P = 0.5$  ps).<sup>51</sup> All the simulations, starting from the average NMR 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 using NVT conditions, after equilibration, were 20 ns long. Cluster analysis was performed by using the Jarvis-Patrick<sup>45</sup> method: a structure is added to a cluster when this structure and a structure in the cluster have each other as neighbors and they have at least  $P$  neighbors in common. The neighbors of a structure are the  $M$  closest structures or all the structures within a cutoff. In our case  $P$  is 3,  $M$  is 9, and the cutoff value is 0.1 nm. All the MD runs and the analysis of the trajectories were performed by using the GROMACS software package.<sup>57</sup>

The graphical representations of the peptide were realized with the program MOLSCRIPT<sup>58</sup> and MOLMOL.<sup>59</sup>

## RESULTS

### Betanova at 280 and 300 K

Figure 3 shows the secondary structure content as a function of time for the two simulations of the peptide at 280 K [Fig. 3(a) and (b)], labeled 280A and 280B, respectively. Figure 4 shows the same quantity for the two simulations of the peptide at 300 K [Fig. 4(a) and (b)], labeled 300A and 300B. Secondary structure assignments were based on the DSSP algorithm.<sup>43</sup> The secondary structure plot for the first of the 280 K trajectories shows that the protein maintains its starting triple  $\beta$ -sheet conformation for about 1.8 ns. Note that in the DSSP notation, the difference between secondary structure elements is based only on the backbone geometry and the presence of hydrogen bonds. The loss of  $\beta$ -sheet defining

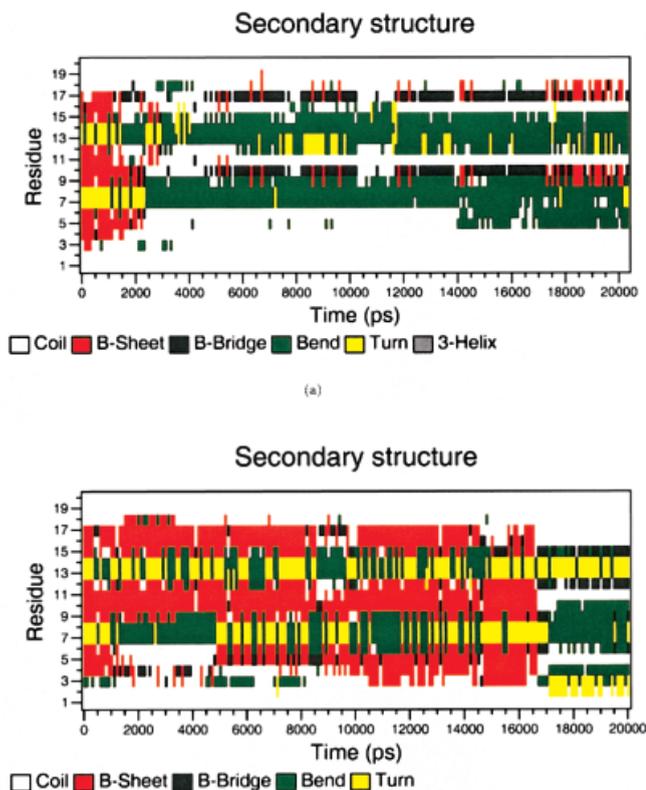


Fig. 3. Secondary structure plots for the two simulations at 280 K: (a) 280A, (b) 280B.

contacts between strands 2 and 3 is accompanied by the loss of the typical loop conformation and loop hydrogen bonds involving residues Asn13 and Gly14. The same is observed in the case of the contacts between strands 1 and 2. In both cases, a change in the conformation of the loop residues accompanies the loss of secondary structure. For brief periods, the loop geometry of residues Asn7 and Gly8 is recovered, but this is not accompanied by the formation of secondary structure in the strands. Figure 5(a) shows the root-mean-square positional deviation (RMSD) relative to the average NMR structure as a function of time for 280A. The RMSD calculated for the backbone atoms of the whole molecule (average value 0.20 nm) primarily reflects the behavior of the backbone atom RMSD calculated for only residues 9–18 (strands 2 and 3) (average value 0.15 nm). In contrast, the RMSD values for residues 3–12 (strands 1 and 2) remain close to their average value (0.14 nm) for the whole simulation.

Figure 3(b) shows the secondary structure content as a function of time for the second 20-ns simulation at 280 K (280B). This simulation was started from the structure obtained after 1.5 ns of simulation 280A by using a different set of velocities obtained from a Maxwellian velocity distribution for a temperature of 280 K. The three-stranded  $\beta$ -sheet geometry is maintained during the first 1 ns of this simulation. The  $\beta$ -strand geometry of the first strand is then lost for approximately 3 ns, in association with the loss of the ideal turn geometry of residues Asn7 and Gly8. The original  $\beta$ -strand geometry for strand

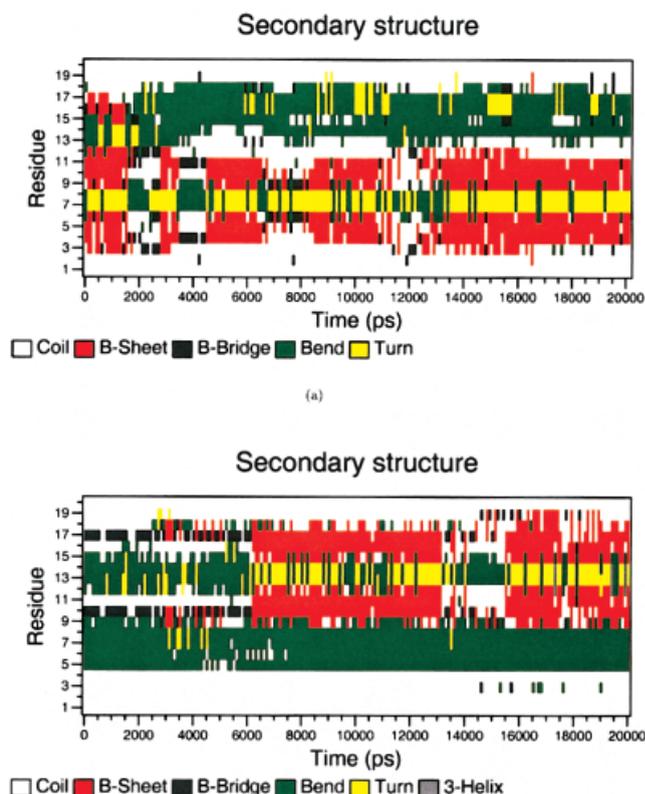
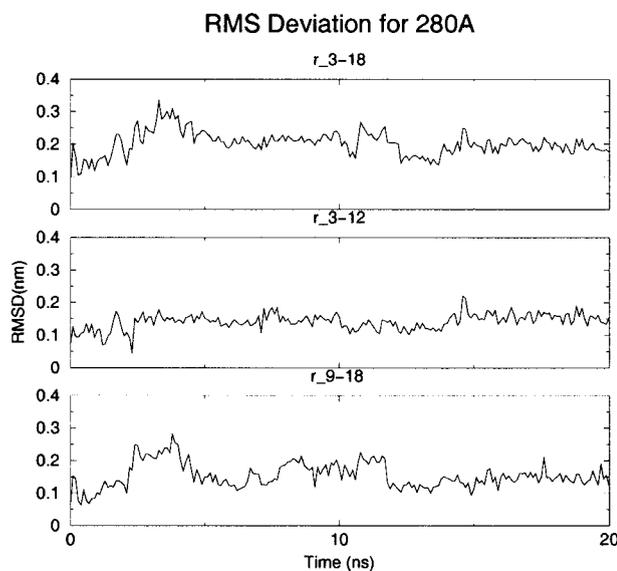


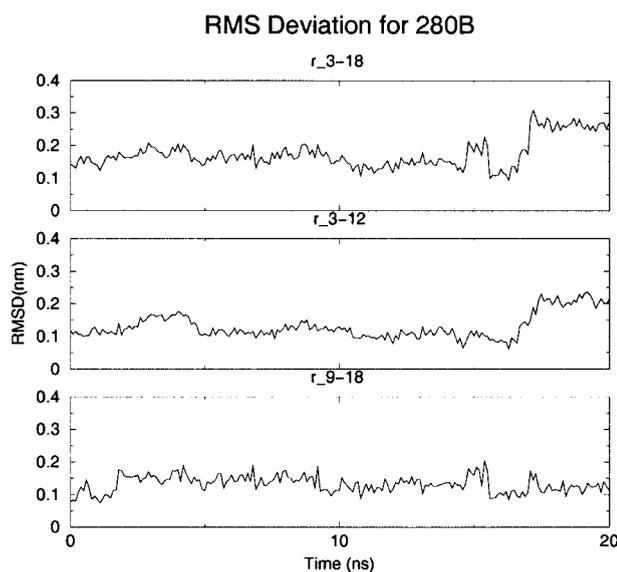
Fig. 4. Secondary structure plots for the two simulations at 300 K: (a) 300A and (b) 300B.

1 recovers at around 5 ns and is maintained until 17 ns. Strands 2 and 3 essentially maintain their secondary structure during the first 17 ns, although the geometry of the turn region fluctuates. At 17 ns, a transition leading to the loss of structure in all strands is observed. The  $\beta$ -turn geometry of residues Asn13 and Gly14 is present until the end of the simulation. However, strands 2 and 3 do not regain their original conformation. Figure 5(b) shows the RMSD relative to the average NMR structure as a function of time for 280B. The RMSD plots of the whole peptide (average value 0.18 nm) and of residues 3–12 (average value 0.14 nm) and residues 9–18 (average value 0.13 nm) separately show that the increase in the RMSD values at 17 ns parallels the increase in the RMSD value of residues 3–12 (strands 1 and 2), whereas the RMSD for residues 9–18 (strands 2 and 3) remains largely constant during this simulation.

Figure 4(a) shows the secondary structure content for the first simulation at 300 K. In this case, the structure is mostly folded for about 2 ns, before conformational changes involving the loop residues Asn7 and Gly8 and Asn13 and Gly14 are observed and the secondary structure is lost. The  $\beta$ -sheet conformation for strands 1 and 2 is recovered repetitively during the simulation, along with the conformation and hydrogen-bonding pattern for the loop residues Asn7 and Gly8.  $\beta$ -strands 1 and 2 remain centered around the Asn7 and Gly8 pair. The loss of secondary structure involving strand 3 is associated with a conformational change in residues Asn13



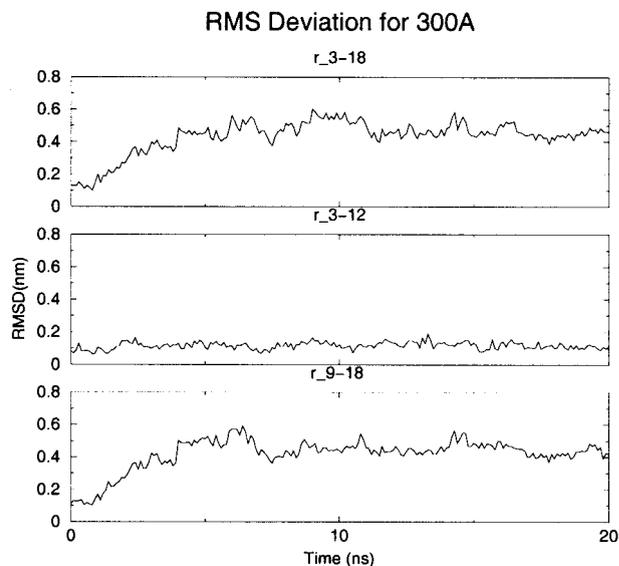
(a)



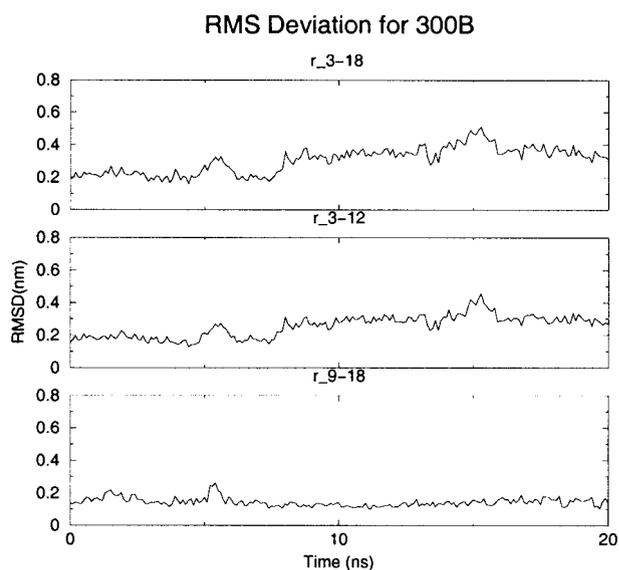
(b)

Fig. 5. RMSD for the two simulations at 280 K of the conformations found in the simulations to the NMR model: (a) 280A and (b) 280B. From top to bottom: RMSD for backbone of residues between 3 and 18, RMSD for backbone of residues between 3 and 12 (hairpin 1), RMSD for backbone of residues between 9 and 18 (hairpin 2).

and Gly14 resulting in the structure of the loop to first be identified as a bend and then as a coil at around 5 ns. Strand 3 moves out of its original plane, losing its contacts with the rest of the molecule, and fluctuates between a series of different conformations without regaining the original hydrogen bonds and conformation. The RMSD calculations on separate parts of Betanova reflect this behavior: the average RMSD values for the whole peptide and for residues 9–18 are 0.43 nm



(a)



(b)

Fig. 6. RMSD for the two simulations at 300 K, of the conformations found in the simulations to the NMR model: (a) 300A and (b) 300B. From top to bottom: RMSD for backbone of residues between 3 and 18, RMSD for backbone of residues between 3 and 12 (hairpin 1), RMSD for backbone of residues between 9 and 18 (hairpin 2).

and 0.41 nm, respectively, whereas the value for residues 3–12 is just 0.11 nm [Fig. 6(a)].

Simulation 300B [Fig. 4(b)] was started from the last conformation obtained from 280A, which does not show defined secondary and tertiary structure elements. During the first 6.5 ns, there is no formation of well-defined secondary structure. Nevertheless, residues Asn7 and Gly8 as well as Asn13 and Gly14 do, for short periods of time, adopt a  $\beta$ -turn conformation. After 6.5 ns, there is a

**TABLE I. NOE Connectivities of Betanova From NMR Experiments in Water Compared with Calculations Based on Simulations 280A and 280B<sup>†</sup>**

No.	NOE <i>i</i> - <i>j</i> atom	Calc. Dist. 280A		Calc. Dist. 280B		
		Exp NOE	r3	r6	r3	r6
1	CeH Y10-NH Q6	w	0.3	0.3	0.6	0.6
2	NH Q6-NH K9	m	0.6*	0.5*	0.4	0.4
3	NH N7-NH K9	w-m	0.5	0.5	0.5	0.5
4	CeH Y10-CαH N7	w-m	0.4	0.4	0.9*	0.8*
5	CeH Y10-CαH S4	w	0.6	0.5	0.5	0.5
6	CeH Y10-CαH G8	w	0.5	0.4	0.7*	0.6
7	CeH Y10-CβH W3	w-m	0.5	0.4	0.4	0.4
8	CeH Y10-Cββ'H N12	w	0.8*	0.7*	0.7*	0.5
9	CeH Y10-CγH V5	m	0.2	0.2	0.2	0.2
10	CδH Y10-CαH V5	w-m	0.5	0.5	0.4	0.4
11	CδH Y10-CγH V5	s	0.3	0.2	0.2	0.2
12	Cη2H W3-CαH V5	m	0.4	0.4	0.9*	0.9*
13	Cη2H W3-CγH V5	Hairpin 1 w	0.2	0.2	0.6	0.5
14	Cξ2H W3-CαH V5	Res 3–12 w	0.5	0.5	0.8*	0.8*
15	Cξ2H W3-CγH V5	w	0.3	0.3	0.2	0.2
16	NδH N7-CγH V5	w	0.5	0.2	0.6	0.6
17	NH T11-CαH V5	m	0.6*	0.5*	0.4	0.4
18	NH K9-CαH N7	m	0.4	0.2	0.4	0.4
19	NH G8-CγH V5	s	0.2	0.2	0.2	0.2
20	NH Q6-CαH Y10	w	0.5	0.4	0.3	0.3
21	CαH Y10-CγH V5	w	0.5	0.3	0.2	0.3
22	CαH Y10-CαH V5	w-m	0.6*	0.6*	0.4	0.4
23	NH K9-CγH V5	m	0.4	0.4	0.2	0.2
24	NH T17-Cββ'H K15	w-m	0.4	0.4	0.3	0.3
25	NH K15-CαH N13	Hairpin 2 w-m	0.5	0.5	0.3	0.4
26	NH T17-CγH K15	Res 9–18 w	0.5	0.5	0.4	0.4
27	NH T16-CβH T11	w	0.5	0.4	0.5	0.5
28	NH R20-CαH N7	m	0.4	0.4	0.4	0.3
29	NH N12-CαH T16	Interstrand w	0.5	0.5	0.5	0.4
30	NH N13-CγH T11	w	0.3	0.3	0.2	0.2
31	NH T16-CβH T11	w	0.5	0.5	0.5	0.5

<sup>†</sup>Experimentally, NOEs are classified according to the following upper distance restraints: strong (s) = 0.25 nm, medium (m) = 0.35 nm, weak-medium (w-m) = 0.5 nm, weak (w) = 0.55 nm. Column r3 refers to  $\langle r^{-3} \rangle^{-1/3}$  averaging, and r6 refers to  $\langle r^{-6} \rangle^{-1/6}$  averaging.

\*Asterisk: violations >0.05 nm.

formation of secondary structure involving strands 2 and 3, which assume a well-defined  $\beta$ -sheet conformation centered around residues Asn13 and Gly14. This can also be seen in the low value of the RMSD for residues 9–18 (average value 0.14 nm), with respect to the NMR model of Betanova. The RMSD value for the whole peptide (average value 0.30 nm) reflects the behavior of residues 3–12 (average value 0.25 nm) [Fig. 6(b)]. As in the previous 300 K simulation, the system tends to populate an ensemble of conformational states in which a part of the conformation deduced from NMR studies is attained, whereas the rest of the molecule (strands 1 or 3) shows a higher flexibility and mobility. The average radius of gyration ( $R_{gyr}$ ) is 0.76 nm for 280A, 0.74 nm for 280B, 0.77 nm for 300A and 0.76 nm for 300B, indicating the sampling of primarily compact conformations.

### NMR-Derived Structural Properties

Tables I and II show the 31 experimental NOE connectivities defining the  $\beta$ -sheet conformation of Betanova, together with the distance values calculated from the

simulations at 280 and 300 K, respectively. The interproton distances were calculated from the simulations as  $\langle r^{-3} \rangle^{-1/3}$  and  $\langle r^{-6} \rangle^{-1/6}$  averages.<sup>44</sup> The averages were calculated for each trajectory separately. From the results in Table I for simulation 280A, it can be seen that most of the NOE constraints are satisfied when averaged over the trajectory. The experimental data are satisfied even if the DSSP analysis indicates that the peptide does not assume a defined three-stranded  $\beta$ -sheet conformation. The molecule in this simulation does, nevertheless, sample a series of conformations basically characterized by the presence of three-strands and two-loop (turn) regions corresponding to residues Asn7 and Gly8, and Asn13 and Gly14, respectively. Violations using  $\langle r^{-3} \rangle^{-1/3}$  distance weighting (calculated values >0.05 nm higher than the experimental bound) are observed for entries 2, 8, 17, and 22. These NOEs correspond to interactions within hairpin 1 between strands 1 and 2. There are no violations in hairpin 2. The NMR constraints are also largely satisfied in simulation 280B. Nevertheless, in this case, despite the long persistence of a well-defined three-stranded  $\beta$ -sheet structure,

**TABLE II. NOE Connectivities of Betanova From NMR Experiments in Water Compared with Calculations Based on Simulations 300A and 300B<sup>†</sup>**

No.	NOE <i>i</i> - <i>j</i> atom	Calc. Dist. 280A		Calc. Dist. 280B		
		Exp NOE	r3	r6	r3	r6
1	CεH Y10-NH Q6	w	0.6	0.4	0.6	0.5
2	NH Q6-NH K9	m	0.4	0.4	0.7*	0.7*
3	NH N7-NH K9	w-m	0.5	0.5	0.6*	0.5
4	CεH Y10-CαH N7	w-m	0.7*	0.6*	0.6*	0.4
5	CεH Y10-CαH S4	w	0.6	0.5	0.6	0.5
6	CεH Y10-CαH G8	w	0.6	0.5	0.6	0.5
7	CεH Y10-CβH W3	w-m	0.4	0.3	0.5	0.4
8	CεH Y10-Cββ'H N12	w	1.1*	0.9*	0.6	0.4
9	CεH Y10-CγH V5	m	0.2	0.2	0.4	0.2
10	CδH Y10-CαH V5	w-m	0.4	0.4	0.7*	0.6*
11	CδH Y10-CγH V5	s	0.4*	0.2	0.7*	0.7*
12	Cη2H W3-CαH V5	m	0.8*	0.7*	0.4	0.3
13	Cη2H W3-CγH V5	w	0.4	0.2	0.2	0.2
14	Cξ2H W3-CαH V5	w	0.9*	0.7*	0.5	0.5
15	Cξ2H W3-CγH V5	w	0.3	0.2	0.3	0.2
16	NδH N7-CγH V5	w	0.6	0.6	0.7*	0.7*
17	NH T11-CαH V5	m	0.4	0.4	0.8*	0.7*
18	NH K9-CαH N7	m	0.4	0.3	0.5*	0.5*
19	NH G8-CγH V5	s	0.2	0.2	0.7*	0.8*
20	NH Q6-CαH Y10	w	0.4	0.4	0.5	0.5
21	CαH Y10-CγH V5	w	0.3	0.2	0.7*	0.7*
22	CαH Y10-CαH V5	w-m	0.4	0.4	0.7*	0.7*
23	NH K9-CγH V5	m	0.2	0.2	0.9*	0.8*
24	NH T17-Cββ'H K15	w-m	0.3	0.3	0.4	0.4
25	NH K15-CαH N13	w-m	0.5	0.4	0.4	0.4
26	NH T17-CγH K15	w	0.4	0.3	0.5	0.4
27	NH T16-CβH T11	w	0.8*	0.6	0.4	0.4
28	NH R20-CαH N7	m	0.4	0.4	0.5*	0.4
29	NH N12-CαH T16	w	0.6	0.5	0.4	0.4
30	NH N13-CγH T11	w	0.2	0.2	0.3	0.2
31	NH T16-CβH T11	w	0.8*	0.3	0.4	0.3

<sup>†</sup>Experimentally, NOEs are classified according to the following upper distance restraints: strong (s) = 0.25 nm, medium (m) = 0.35 nm, weak-medium (w-m) = 0.5 nm, weak (w) = 0.55 nm. Column r3 refers to  $\langle r^{-3} \rangle^{-1/3}$  averaging, and r6 refers to  $\langle r^{-6} \rangle^{-1/6}$  averaging.

\*Asterisk: violations above 0.05 nm.

violations of the NMR restraints are observed for entries 4, 6, 8, 12, and 14 ( $\langle r^{-3} \rangle^{-1/3}$ ) averaging. All violations again occur within hairpin 1. The NOEs within hairpin 2 are satisfied although we note that the calculated distances for hairpin 2 are significantly shorter in 280B compared to 280A. Hairpin 1 shows a well-defined  $\beta$ -sheet conformation of the backbone throughout most of the trajectory, and the largest violations are observed for NOEs involving the side chains of Val5 and Trp3. Averaging as  $\langle r^{-6} \rangle^{-1/6}$ , only NOEs 4, 12, and 14 would be violated by more than 0.05 nm. Taking simulations 280A and B together, only one weak NOE constraint, from the side chain of Tyr 10 to the side chain of Asn 12 (entry 8) is still violated and only when using  $\langle r^{-3} \rangle^{-1/3}$  as opposed to  $\langle r^{-6} \rangle^{-1/6}$  averaging.

Table 2 lists the average distances and violations of the experimentally measured NOEs at 280K for the simulations 300A and 300B. Clearly, there are more violations in the simulations at 300 K than at 280 K. Nevertheless, most of the NOE-derived distance constraints are satisfied in both simulations. In the simulation 300A, violations higher than 0.05 nm assuming  $\langle r^{-3} \rangle^{-1/3}$  averaging occur for entries 4, 8, 11, 12, 14, 27, and 31. Entries 4, 8, 11, 12,

and 14 occur in hairpin 1. In Figure 4(a) it is evident that throughout simulation 300A residues 2–12 primarily adopt a well-defined  $\beta$ -hairpin. The observed violations involve mostly NOEs in the side-chain atoms of residues Trp3 and Tyr10. Using  $\langle r^{-6} \rangle^{-1/6}$  averaging only entries 8, 12, and 14 show violations. The simulation 300B was started from a conformation that lacked well-defined secondary structure. After 6.5 ns, there was spontaneous formation of a  $\beta$ -hairpin between strands 2 and 3 but not between strands 1 and 2. Violations of the NOE-derived distances are observed for entries 2–4, 10, 11, 16–19, 21–23, and 28. Applying  $\langle r^{-6} \rangle^{-1/6}$  averaging entries 2, 11, 16–19, and 21–23 would still show violations. The violations are concentrated in hairpin 1 primarily involving backbone as opposite to side-chain atoms. There are no violations in hairpin 2. Importantly, only entries 4 and 11 show significant violations in both simulations and only when using  $\langle r^{-3} \rangle^{-1/3}$  as opposed to  $\langle r^{-6} \rangle^{-1/6}$ . Taken together, the comparison between the experimental NOEs and the distance averaged from the simulations indicate that overall the simulations satisfy the experimental constraints. It is important to note that different structures

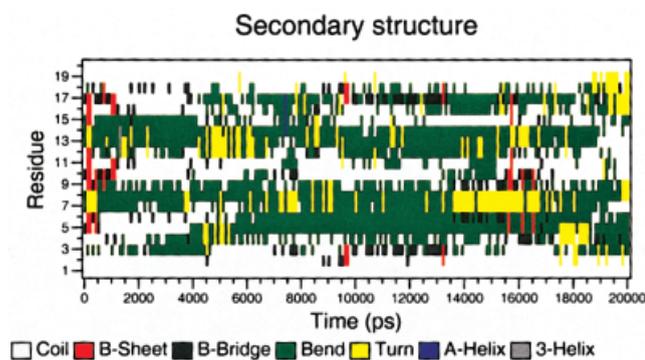


Fig. 7. Secondary structure plot for the simulation at 350 K.

from the simulation satisfy different sets of NOEs. In these simulations, the system can be considered to sample an ensemble of compact states characterized by the presence of fully or partially formed secondary structure elements that together satisfy the experimental data.

### Betanova at 350 and 450 K

The unfolding of Betanova at 350 K is illustrated in Figure 7. At 350 K, the initial triple stranded  $\beta$ -sheet structure is maintained for only a few hundred picoseconds and not recovered for any extended period during the remainder of the simulation. It is of interest that the loop conformation is recovered several times during the first 10 ns at 350 K and is present continuously between 13 and 18 ns. The second  $\beta$ -turn is observed for shorter periods. The RMSD values for the turn residues are low during the simulation: the average backbone RMSD value for residues 6–9 (the first  $\beta$ -turn) is 0.08 nm, whereas for residues 12–15 (the second  $\beta$ -turn) this value is 0.1 nm. This calculation shows that these residues tend to maintain or to recover their conformation giving rise to the turn formation. The average backbone RMSD value for the whole peptide is 0.40 nm, for strands 1 and 2 is 0.28 nm, and for strands 2 and 3 is 0.25 nm. Even at this high temperature, some secondary structure can, in fact, be observed for short periods, indicating that the conformational space sampled under these conditions includes the conformational space accessible to the molecule at lower temperatures. In the simulation at 450 K, the peptide rapidly loses its initial structure: the strands become highly flexible and the well-defined turn regions are lost. There is no evidence for the formation of a triple stranded  $\beta$ -sheet. The presence of stable  $\beta$ -turn sequences in the simulation at 350 K is reflected in the value and in the variation of the  $R_{gyr}$  with respect to time. The average value is 0.76 nm, very similar to that calculated for the simulations at 280 and 300 K. Even though the oscillations at 350 K are more pronounced and values of 1.0 nm for the  $R_{gyr}$  are obtained, the molecule primarily maintains a compact structure.

### Hydrogen Bonding

Experimentally, the hydrogen bonds between Ser4 and Thr11, and Tyr10 and Thr17 have been considered impor-

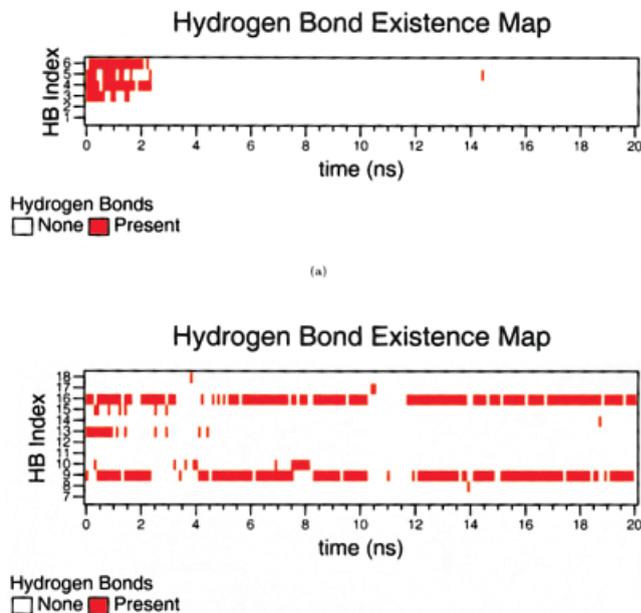


Fig. 8. Interstrand hydrogen bond persistence maps for simulation 280A. (a) Strands 1 and 2. Entry 1: N—H of Trp3 to C=O of Thr11; Entry 2: N—H of Ser4 to C=O of Tyr10; Entry 3: N—H of Ser4 to C=O of Thr11; Entry 4: N—H of Gln6 to C=O of Lys9; Entry 5: N—H of Lys9 to C=O of Gln6; Entry 6: N—H of Thr11 to C=O of Ser4. (b) Strands 2 and 3. Entry 7: N—H of Lys9 to C=O of Thr16; Entry 8: N—H of Tyr10 to C=O of Thr16; Entry 9: N—H of Tyr10 to C=O of Thr17; Entry 10: N—H of Tyr10 to C=O of Glu18; Entry 11: N—H of Thr11 to C=O of Thr16; Entry 12: N—H of Thr11 to C=O of Thr17; Entry 13: N—H of Asn12 to C=O of Lys15; Entry 14: N—H of Lys15 to C=O of Thr11; Entry 15: N—H of Lys15 to C=O of Asn12; Entry 16: N—H of Thr17 to C=O of Tyr10; Entry 17: N—H of Thr17 to C=O of Thr11; Entry 18: N—H of Thr17 to C=O of Asn12. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

tant in defining a three-stranded structure for the Betanova peptide.<sup>41</sup> Figure 8(a) shows the hydrogen bonds between strands 1 and 2 in simulation 280A. A hydrogen bond was considered to exist if the distance between the hydrogen atom and the acceptor is  $<0.25$  nm and if the angle donor-hydrogen-acceptor is  $<60^\circ$ . Hydrogen bond 3 (N—H of Ser4 to C=O of Thr11), 4 (N—H of Gln6 to C=O of Lys9), 5 (N—H of Lys9 to C=O of Gln6), and 6 (N—H of Thr11 to C=O of Ser4) are present only in the first 2 ns of the simulation. Figure 8(b) shows the backbone-backbone hydrogen bonds between strands 2 and 3. In this case, two hydrogen bonds are present for most of the simulation, namely, hydrogen bond 9 (N—H of Tyr10 to C=O of Thr17) and 16 (N—H of Thr17 to C=O of Tyr10). Hydrogen bond 13 (N—H of Asn12 to C=O of Lys15) is present for about 2 ns. The presence of a stable hydrogen bond between Tyr10 and Thr17 is indicative of a structure consistent with the NOE experimental data. This would not be inferred by examination of the DSSP graph in Figure 3(a) alone.

Figures 9(a) and (b) shows the hydrogen bonds for strands 1 and 2 and 2 and 3, respectively, for simulation 280B. Figure 9(a) shows the presence of three long-lived hydrogen bonds between strands 1 and 2: hydrogen bonds 3, 4, and 6. Hydrogen bond 5 appears intermittently.

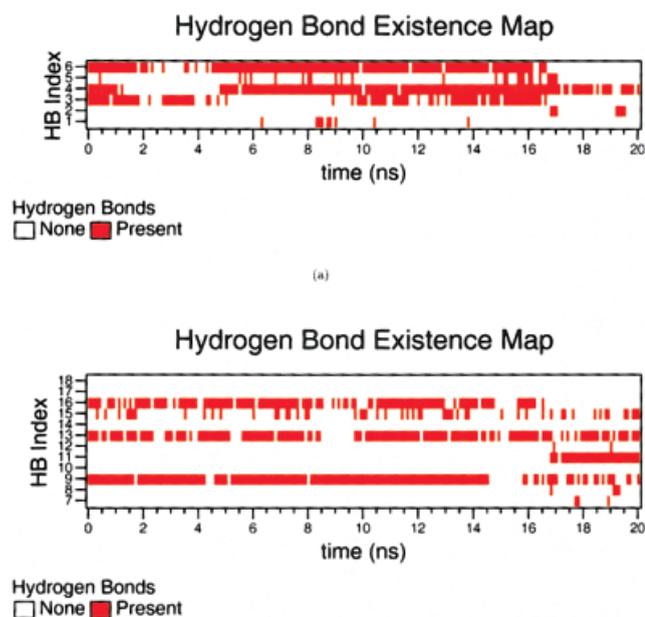


Fig. 9. Interstrand hydrogen bond persistence maps for simulation 280B. See Figure 8 for legend. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

Figure 9(b) shows the hydrogen bond interactions between residues Tyr10 and Thr17, namely, hydrogen bonds 9 and 16 are again observed. Hydrogen bonds 13 and 15 (N—H of Lys15 to C=O of Asn12) also persist for most of the simulation. The presence of a higher number of interstrand hydrogen bonds in simulation 280B is consistent with a more defined three-stranded  $\beta$ -sheet structure, as seen in the DSSP-based analysis [Fig. 3(b)].

### Side-Chain Contacts

Kortemme and coworkers<sup>41</sup> reported that the packing of aromatic side chains in  $\beta$ -sheets contributes significantly to their stability, as well as establishing important conformational constraints. The distances between the centers of mass of the side chains of hydrophobic-aromatic residue pairs, involving residues Trp3, Val5, and Tyr10 were thus monitored to determine the presence of hydrophobic interactions in Betanova. Two side chains were considered to be in contact if the distance between their respective centers of mass was  $<0.65$  nm.

At 280 K the average distance between the centers of mass of the side chains of Trp3 and Tyr10 is 0.59 nm for simulation 280A and 0.50 nm for simulation 280B. In 280B the distance stabilizes after 5 ns together with the formation of secondary structure involving strands 1 and 2. A large oscillation is observed at 17 ns, coincident with the loss of tertiary structure previously noted. In 300A, this distance oscillates during the whole simulation. On average, it is around 0.65 nm, indicating the presence of a hydrophobic contact. Larger distances corresponded to when strands 1 and 2 lose secondary structure. In 300B, the average distance between the side chains is 0.86 nm. Higher values were observed in the latter part of the simulation, in which strand 1 fluctuates, whereas strands

2 and 3 form a well-defined  $\beta$ -hairpin structure. At 350 K (average distance 1.03 nm) the Trp3-Tyr10 contact is present for only short periods that do not coincide either with the formation of secondary structure involving strands 1 and 2 or with the formation of a stable turn involving Asn7 and Gly8.

The contact between Val5 and Tyr10 was present in all the simulations at lower temperature. In particular, in 280B, in which the peptide is mostly folded, the value of this distance was around 0.4 nm. In 300A, low inter-residue distances characterize parts of simulation in which secondary structure in strands 1 and 2 is present. In 300B, there was no contact between Val5 and Tyr10 (average value 0.87 nm).

Other important interstrand packing interactions are between Asn12 and Thr17, and between Asn12 and Trp3. During simulation 280B, the average distance between the centers of mass of the side chains of Asn12 and Thr17 was 0.52 nm. Fluctuations occur at the beginning of the simulation (in which there is partial loss of secondary structure) and around 17 ns. Nonetheless, after this transition, the contact was reestablished and the peptide maintains a compact state. The two side chains of Asn12 and Trp3 remained in contact (distance around 0.55 nm) until 17 ns. In 280A, the contact between Asn12 and Thr17 was partially lost, whereas the contact between Asn12 and Trp3 was maintained for most of the simulation. The contact Asn12-Thr17 was intermittent in 300A, in which there was partial loss of structure for residues 9–18. The Asn12-Trp3 contact was stable. In 300B, in which the  $\beta$ -hairpin between strands 2 and 3 was present, the situation is reversed. In the 350 K simulation, the side chain-side chain contacts involving Asn12-Tyr10 and Asn12-Val5 appeared to be stable for long intervals, as were the hydrophobic contacts between Val5 and Tyr10. The structure of the hydrophobic core of the peptide in the 350 K simulation is illustrated in Figure 10.

## DISCUSSION

The NOE distance-bound violations, summarized in Tables I and II, show that the simulations sample regions of conformational space in which the peptide fulfills the available NMR-derived constraints. It is of interest that even when the DSSP algorithm is unable to identify a well-defined three-stranded  $\beta$ -sheet conformation (280A), the constraints are mostly satisfied. A combination of configurations from both trajectories 280A and B was required to satisfy experimental NOE data. The overall agreement is very good in comparison to other studies.<sup>37</sup> The peptide does not remain fixed in a triple-stranded  $\beta$ -sheet configuration. After approximately 2 ns at 280 K, the peptide begins to deviate from the idealized starting structure. In this regard, we note the danger of inferring the stability of the peptide from only a few nanosecond simulations. Bursulaya and Brooks,<sup>36</sup> for example, concluded that the peptide was stable within the CHARMM force field because no significant structural changes were observed within 2 ns. Experimentally, however, Betanova is estimated to be just 30%  $\beta$ -sheet on average<sup>41</sup> [even as low as 10% based on recent experimental

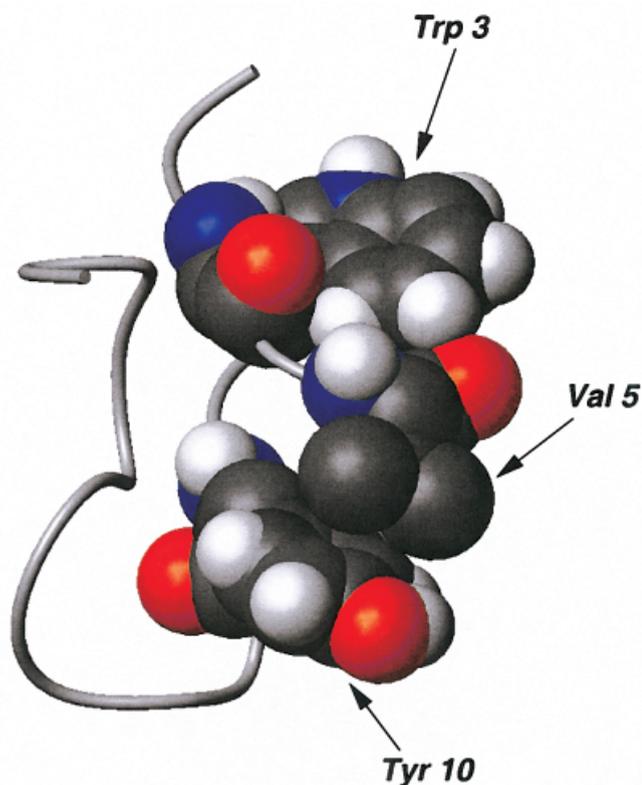


Fig. 10. Hydrophobic core formation in the 350 K simulation. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

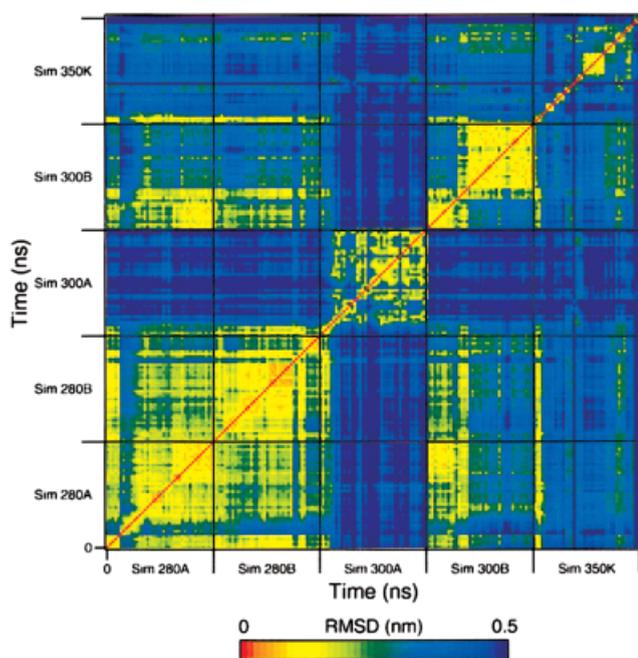


Fig. 11. RMSD matrix for each structure from the simulations at 280, 300, and 350 K.

measurements on Betanova and some of its mutants (Serano, personal communication)].

Figure 11 shows the RMSD matrix calculated from the

simulations at 280, 300, and 350 K. The presence of light-colored off-diagonal regions indicates the sampling of similar regions of conformational space. The starting conformations and the initial velocities are different for the simulations at 280 K and at 300 K. Different initial conditions were used to test the influence of the starting conformation on the final results and to increase the amount of sampled conformational space. The slow folding times of  $\beta$ -sheet peptides means that some dependence of the simulation on the starting conditions is expected. However, despite using different velocities and geometries for all of the four simulations at low temperature (280 and 300 K), regions of strong overlap can be seen between simulations 280A, 280B, and 300B. There is also significant overlap with the simulation at 350 K. Simulation 300A shows the least overlap with the other simulations. It is clear, nevertheless, that in each simulation the peptide samples a range of conformations with RMSD values differing by  $>0.5$  nm and yet still returns close to the initial model. The rate of folding of  $\beta$ -sheet peptides, however, precludes the possibility to simulate the folding of native Betanova from, for example, an  $\alpha$ -helix or a completely extended structure on the timescale investigated, 20–40 ns.

At 280 K (folding conditions), the system can be considered to populate an ensemble of different states fluctuating around the experimentally determined structure. Based on the union of the two 280K trajectories, a cluster analysis using the Jarvis-Patrick algorithm was performed.<sup>45</sup> The structure representative of the most populated cluster, that is, the most probable structure in the simulations, shows a more regular three-stranded  $\beta$ -sheet conformation than the original NMR-based model [Fig. 2(b)]. The population of the most populated cluster, which comprise three-stranded  $\beta$ -sheet structures, is about 15% of the total population for the peptide. This is in rough accord with experimentally based estimations of the percentage of folded structures in water solution.<sup>41</sup> Experimentally, Kortemme and coworkers<sup>41</sup> determined that the stability of Betanova to be around  $-2.4$  kJ/mol at 278 K. Classifying all conformations with a backbone RMSD from the most probable structure lower than 0.15 nm to be folded, and all the conformations with an RMSD higher than 0.2 nm as unfolded, we obtained an estimate for the free-energy difference between the unfolded and the folded states of  $\Delta G(280 \text{ K}) = -0.8$  kJ/mol. Clearly, such free-energy estimates must be treated with caution because reversibility has not been demonstrated. Nevertheless, it underlines the fact that experimental studies indicate that at 280 K Betanova populates a range of conformations in solution and that the proportion of folded states observed in the simulation is in agreement with that inferred from the experimental data.

### Folding Mechanism

What can these simulations teach us about stability and folding mechanism of  $\beta$ -sheet peptides? They demonstrate that long MD simulations can reproduce much of the available experimental data. The fact that NMR restraints

are satisfied even in the absence of optimal hydrogen bond interactions underlines the probable importance of hydrophobic contacts in the folded peptide. The role of hydrophobic interactions in defining a compact state for Betanova is particularly evident in the 350 K simulation. Cluster analysis of the 350 K trajectory indicates that the predominate structure [Figs. 2(e) and Fig. 10] has a well-defined hydrophobic cluster comprising Trp3, Val5, and Tyr10. The turn conformation for residues Asn7 and Gly8, and Asn13 and Gly14 is obtained when these stabilizing hydrophobic interactions are present. Once this nucleus is formed, the loop residues are forced to adopt the turn conformation and the  $\beta$ -strands are able to form.<sup>46</sup>

At 300 K only two stable  $\beta$ -strands form. Betanova was designed on the basis of stable  $\beta$ -hairpin templates. These structural motifs are reproduced by the simulations. The presence of the different  $\beta$ -hairpin structures in the different simulations suggests that the actual formation of a complete triple-stranded  $\beta$ -sheet may not be required to satisfy the experimental NOE data [Figs. 2(c) and (d)].

It has been proposed that secondary structure formation involves formation of interstrand stabilizing interactions provided by the hydrophobic contacts, followed by formation of hydrogen bonds and development of stable  $\beta$ -turn conformations.<sup>47,48</sup> In each of the 300 K simulations, one strand of the peptide remains unstructured without assuming a definite  $\beta$ -strand conformation. Experimentally, the peptide was built on the basis of a stable hairpin involving strands 2 and 3. The formation of this hairpin was also observed in a previous simulation study of this system conducted on a shorter time range.<sup>36</sup> However, our simulations show that the formation of an intermediate hairpin structure can also involve strands 1 and 2. It is worth noting that residues Trp3, Val5, and Tyr10 are located on these two strands, and they are in contact for most of 300A simulations. This hydrophobic cluster may be an important factor in driving the formation and stabilization of a stable  $\beta$ -hairpin motive between strands 1 and 2. Moreover, the mutation of the residue in position 5 in both experiments (Serrano, personal communication) and simulations (Soto, manuscript in preparation) to a hydrophilic one has revealed a destabilizing effect on the peptide that can be traced back to the disruption of the hydrophobic core.

This observation, together with the analysis of the NMR violations, shows that both the intermediate structures at 300 K are plausible. Both fulfill most of the NOE-derived restraints. This implies that to fold, the peptide first forms a partially ordered structure (hairpin) and then the experimentally determined three-stranded antiparallel pattern. Bursulaya and Brooks,<sup>36</sup> using Betanova and Ferrara and Caffisch<sup>37</sup> (in studies of a closely related system), proposed that the formation of the contacts between strands 2 and 3 was the critical folding event.

Our simulations show that, in addition to this mechanism, a pathway involving the formation of a stable hairpin between strands 1 and 2 is not only possible but highly likely. In both cases, a free-energy barrier must still be overcome to form a triple-stranded  $\beta$ -sheet.<sup>49,50</sup> The

orientation of two peptide bonds must be fixed to form one interstrand hydrogen bond. This process is expected to be energetically unfavorable in the absence of stabilizing side-chain interactions. In the simulations, two  $\beta$ -strands are readily formed, but not the third. This is also in line with the data of deAlba et al.<sup>29</sup> for the related peptide simulated by Ferrara and Caffisch.<sup>37</sup> Chemical shift data indicate only very marginal cooperativity in the folding of the two hairpins to form the triple strand. The population of  $\beta$ -sheet in between strands 1 and 2 doubles with the addition of strand 3, whereas the population of  $\beta$ -sheet between strands 2 and 3 is unaffected by the presence or absence of strand 1. Both hairpins form in solution with almost identical propensities. The observation of the formation of the two different hairpins also implies that these two structures can coexist in equilibrium in solution, contributing to the determination of the experimentally observed spectrum. At lower temperature (280 K), conformations that are basically three stranded become favored, even though the observation of high conformational flexibility makes it more appropriate to consider an ensemble of NMR-satisfying structures than to define a single model to account for all experimental constraints.

## CONCLUSIONS

We have examined in atomic detail the folding and dynamics of the three-stranded  $\beta$ -sheet peptide Betanova in explicit water at several temperatures. Experimentally, Betanova is marginally stable at 280 K. The simulations suggest that at 280 K the peptide populates a series of different conformational states that overall satisfy the NOE constraints. Although a three-stranded  $\beta$ -sheet was the predominate structure in the simulations at this temperature, at 300 K the predominate conformations contained either one or the other of the two potential hairpin structure. The peptide was in an equilibrium between closely related states in which distinct secondary structure elements formed. At a higher temperature, the simulations suggest that side-chain hydrophobic contacts are important in determining a compact state and driving the peptide to its final structure.

MD simulations in explicit water remain computationally expensive. Given the slow rate of  $\beta$ -sheet folding (4–14  $\mu$ s), it is not yet possible to demonstrate reversible folding in this system under realistic conditions. Nevertheless, the simulations have given a detailed atomic picture of the conformational dynamics of Betanova and an improved understanding of the factors that play a role in folding stability. The combination of simulation techniques, together with experimentally derived structural and kinetic data on progressively more complex systems, has the potential to provide an ever-deeper understanding into the mechanism of peptide and protein folding.

## ACKNOWLEDGMENTS

The authors thank Dr. L. Serrano and Dr. M. Lopez de la Paz for providing the NMR-NOE data and for helpful discussion.

## REFERENCES

- Fersht AR. Characterizing transition states in protein folding: An essential step in the puzzle. *Curr Opin Struct Biol* 1995;5:79–84.
- Neira J. Towards the complete structural characterization of a protein folding pathway: The structures of the denatured, transition and native states for the association/folding of two complementary fragments of cleaved chymotrypsin inhibitor. *Folding Design* 1996;1:189–208.
- Jones C. Fast events in protein folding initiated by nanosecond laser photolysis. *Proc Natl Acad Sci USA* 1993;90:11860–11864.
- Thomas P, Dill K. Statistical potentials extracted from protein structures: How accurate are they? *J Mol Biol* 1996;257:457–469.
- Schulz G. A critical evaluation of methods for prediction of protein secondary structures. *Annu Rev Biophys Chem* 1988;17:1–21.
- Leach A. A survey of methods for searching the conformational space of small and medium sized molecules. *Rev Comput Chem* 1991;2:1–55.
- Pedersen J, Moult J. Protein folding simulations with genetic algorithms and a detailed molecular description. *J Mol Biol* 1997;269:240–259.
- Shakhnovich EI. Theoretical studies of protein-folding thermodynamics and kinetics. *Curr Opin Struct Biol* 1997;7:29–40.
- Lazaridis T, Karplus M. New view of protein folding reconciled with the old through multiple unfolding trajectories. *Science* 1997;278:1928–1931.
- Karplus M, Sali A. Theoretical studies of the protein folding and unfolding. *Curr Opin Struct Biol* 1995;5:58–73.
- Ladurner AG, Itzhaki LS, Daggett V, Fersht AR. Synergy between simulations and experiment in describing the energy landscape of protein folding. *Proc Natl Acad Sci USA* 1998;95:8473–8478.
- Brooks CL III, Boczek EM. First-principles calculation of the folding free energy of a three helix bundle protein. *Science* 1995;269:393–396.
- Zhou Y, Karplus M. Protein folding thermodynamics of a model three-helix-bundle protein. *Proc Natl Acad Sci USA* 1997;94:14429–14432.
- Duan Y, Kollman PA. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 1998;282:740–749.
- Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Reversible peptide folding in solution by molecular dynamics simulation. *J Mol Biol* 1998;280:925–932.
- Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide folding: When simulation meets experiment. *Angew Chemie Int Ed* 1999;38:236–240.
- Daura X, van Gunsteren WF, Mark AE. Folding-unfolding thermodynamics of a  $\beta$ -heptapeptide from equilibrium simulations. *Proteins* 1999;34:269–280.
- Lacroix E, Viguera A, Serrano L. Elucidating the folding problem of  $\alpha$ -helices: Local motifs, long-range electrostatics, ionic strength dependence and prediction of nmr parameters. *J Mol Biol* 1998;284:173–191.
- Munoz V, Cronet P, Lopez-Hernandez E, Serrano L. Analysis of the effect of local interactions in protein stability. *Folding Design* 1996;1:167–178.
- Smith C, Regan L. Guidelines for protein design: The energetics of  $\beta$ -sheet side-chain interaction. *Science* 1995;270:980–982.
- Dyson HJ, Wright PE. Peptide conformation and protein folding. *Curr Opin Struct Biol* 1993;3:60–65.
- Dobson CM, Sali A, Karplus M. Protein folding: A perspective from theory and experiment. *Angew Chem Int Ed* 1998;37:869–893.
- Kelly J. Amyloid fibril formation and protein miniassembly: A structural quest for insight into amyloid and prion diseases. *Structure* 1997;5:595.
- Booth D. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* 1997;385:787.
- Blanco FJ, Jiménez MA, Pineda A, Rico M, Santoro J, Nieto JL. NMR solution structure of the isolated *n*-terminal fragment of protein-G B<sub>1</sub> domain. Evidence of trifluoroethanol induced native-like  $\beta$ -hairpin formation. *Biochemistry* 1994;33:6004–6014.
- Searle MS, Williams DH, Packman LC. A short linear peptide derived from the *n*-terminal sequence of ubiquitin folds into a water-stable non native  $\beta$ -hairpin. *Nat Struct Biol* 1995;2:999–1006.
- Ramirez-Alvarado M, Blanco FJ, Serrano L. De novo design and structural analysis of a model  $\beta$ -hairpin peptide system. *Nat Struct Biol* 1996;3:604–612.
- de Alba E, Jimenez MA, Rico M. Turn residue sequence determines  $\beta$ -hairpin conformation in designed peptides. *J Am Chem Soc* 1997;119:175–183.
- deAlba E, Santoro J, Rico M, Jimenez AA. De novo design of a monomeric three-stranded antiparallel  $\beta$ -sheet. *Protein Sci* 1999;8:854–865.
- Tobias DJ, Mertz JE, Brooks CL III. Nanosecond time scale folding dynamics of a pentapeptide in water. *Biochemistry* 1991;30:6054–6058.
- Scully J, Hermans J. Backbone flexibility and stability of reverse turn conformation in a model system. *J Mol Biol* 1994;235:682–694.
- Demchuck E, Bashford D, Case D. Dynamics of a type VI turn in a linear peptide in aqueous solution. *Folding Design* 1997;2:35–46.
- Roccatano D, Amadei A, Di Nola A, Berendsen HJC. A molecular dynamics study of the 41–56  $\beta$ -hairpin from b1 domain of protein g. *Protein Sci* 1999;8:2130–2143.
- Bonvin AMJJ, van Gunsteren WF.  $\beta$ -hairpin stability and folding: Molecular dynamics simulations of the first  $\beta$ -hairpin of tendamistat. *J Mol Biol* 2000;296:255–268.
- Ma B, Nussinov R. Molecular dynamics simulations of a  $\beta$ -hairpin fragment of protein g: balance between side-chain and backbone forces. *J Mol Biol* 2000;296:1091–1104.
- Bursulaya B, Brooks CL III. Folding free energy surface of a three-stranded  $\beta$ -sheet protein. *J Am Chem Soc* 1999;121:9947–9951.
- Ferrara P, Caffisch A. Folding simulations of a three-stranded antiparallel beta-sheet peptide. *Proc Natl Acad Sci USA* 2000;97:10780–10785.
- Ibragimova GT, Wade XRC. Stability of the  $\beta$ -sheet of the ww domain: A molecular dynamics simulation study. *Biophys J* 1999;77:2191–2198.
- Wang H, Sung S. Molecular dynamics simulations of three-strand- $\beta$ -sheet folding. *J Am Chem Soc* 2000;122:1999–2009.
- Sung SS. Monte Carlo simulations of  $\beta$ -hairpin folding at constant temperature. *Biophys J* 1999;76:164–175.
- Kortemme T, Ramirez-Alvarado M, Serrano L. Design of a 20-amino acid three-stranded  $\beta$ -sheet protein. *Science* 1998;281:253–256.
- van der Vaart A, Bursulaya BD, Brooks CL III, Merz KM Jr. Are many body effects important in protein folding? *J Phys Chem B* 2000;104:9554–9563.
- Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983;22:2576–2637.
- Tropp J. Dipolar relaxation and nuclear overhauser effects in non-rigid molecules: The effect of fluctuating internuclear distances. *J Chem Phys* 1980;72:6035–6043.
- Jarvis RA, Patrick EA. Clustering using a similarity measure based on shared near neighbors. *IEEE Trans Comp* 1973;22:1025–1034.
- Martinez J, Pisabarro M, Serrano L. Obligatory steps in protein folding and the conformational diversity of the transition state. *Nat Struct Biol* 1998;5:721–729.
- Docholyan N, Buldyrev S, Stanley H, Shakhnovich E. Identifying the protein folding nucleus using molecular dynamics. *J Mol Biol* 2000;296:1183–1188.
- Ramirez-Alvarado M, Daragan VA, Serrano L, Mayo KH. Motional dynamics in a  $\beta$ -hairpin peptide measured by <sup>13</sup>C-NMR relaxation. *Prot Sci* 1998;7:720–729.
- Munoz V, Thompson PA, Hofrichter J, Eaton WA. Folding dynamics and mechanism of  $\beta$ -hairpin formation. *Nature* 1997;390:196–199.
- Eaton W, Munoz V, Thompson PA, Henry ER, Hofrichter J. Kinetics and dynamics of loops,  $\alpha$ -helices,  $\beta$ -hairpins and fast folding proteins. *Acc Chem Res* 1998;31:745–753.
- Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys* 1984;81:3684.

52. van Gunsteren WF, Daura X, Mark AE. GROMOS force field. *Encycl Comput Chem* 1998;2:1211–1216.
53. van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG. Biomolecular simulation: The GROMOS96 manual and user guide. vdf Hochschulverlag, ETH Zürich, Switzerland, 1996.
54. Berendsen HJC, Grigera JR, Straatsma TP. The missing term in effective pair potentials. *J Phys Chem* 1987;91:6269–6271.
55. Hess B, Bekker H, Fraaije J, Berendsen H. A linear constraint solver for molecular simulations. *J Comp Chem* 1997;18:1463–1472.
56. Miyamoto S, Kollman PA. Settle: An analytical version of the shake and rattle algorithms for rigid water models. *J Comp Chem* 1992;13:952–962.
57. van der Spoel D, van Drunen R, Berendsen HJC. Groningen machine for chemical simulations. Department of Biophysical Chemistry, BIOSON Research Institute Nijenborgh 4 NL-9717 AG Groningen 1994. e-mail to gromacs@chem.rug.nl.
58. Kraulis PJ. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 1991;24:946.
59. Koradi R, Billeter M, Wuthrich K. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graphics* 1996;14:51–55.