Polymer Journal, Vol. 27, No. 4, pp 347—360 (1995)

Conformations of Synthetic Model Peptides for *Plasmodium falciparum* Circumsporozoite Protein in Me_2SO by ^1H NMR and Distance Geometry Calculations††

Kimiko UMEMOTO,† Jun KIKUCHI,* Mitsuaki NARITA,* Kenichi FUJITA,** and Tetsuo ASAKURA*

Department of Chemistry, International Christian University, Mitaka, Tokyo 181, Japan
** Department of Biotechnology, Faculty of Technology, University of Agriculture and Technology, Koganei, Tokyo 184, Japan*
*** ML Project, JEOL Co., Ltd., Akishima, Tokyo 197, Japan*

(Received September 6, 1994)

ABSTRACT: A series of terminally blocked peptides consisting of a tetrapeptide repeat, Boc-Asn-Ala-(Asn-Pro-Asn-Ala)$_{n=0,1,2,5,17}$-Asn-Pro-OBzl, has been synthesized and one and two dimensional ^1H NMR studies in dimethyl sulfoxide (Me_2SO)-d_6 as well as distance geometry calculations have been carried out. The repeating tetrapeptide units constitute the central area of the circumsporozoite coat protein of human malaria parasite *Plasmodium falciparum*. The two dimensional nuclear Overhauser effect (NOE) data observed in Me_2SO and the information from the temperature dependence of the amide proton chemical shifts were used as constraints in distance geometry calculations. The result suggests that each -(Asn-Pro-Asn-Ala)- tetrapeptide forms a structural unit, a considerable population of which exists as unique turnlike structures.

KEY WORDS Peptide Conformation / Circumsporozoite Coat Protein / Malaria Parasite / ^1H NMR / Distance Geometry / β-Turn / (Asn-Pro-Asn-Ala) Repeat /

Human malaria parasite *Plasmodium falciparum* produces, during its sporozoite stage, the circumsporozoite (CS) coat protein. The immunodominant component of CS is a tandem repeat consisting of 37 units of tetrapeptide, Asn-Pro-Asn-Ala (NPNA).[1] The secondary structure of such tetrapeptide repeat has attracted much attention from the immunogenical point of view, since synthetic peptides containing this repeat have been prepared and shown to be effective in generating antibodies.[2]

In addition to the immunogenical importance, the above specific sequence is also of stereochemical interest. According to the amino acid preference analysis of turns in proteins, Asn and Pro are among the residues that are most frequently found in tight turns.[3] Similar conclusions were arrived by statistical studies, where Asn and Pro were preferentially found at the $i+1$ and $i+2$ positions, respectively, of the four residues that constitute a β-turn.[4] Structural studies on cyclic model peptides have also shown that Pro-Asn combination promotes the formation of β-turns especially in non-aqueous solvents.[5]

Several studies have been conducted on the solution conformation of the model peptides

† To whom correspondence should be addressed.

†† *Abbreviations*: NMR, nuclear magnetic resonance spectroscopy; Boc, *tert*-butoxyloxycarbonyl; OBzl, benzyl ester; COSY, two-dimensional correlated spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; ROESY, rotating-frame NOESY spectroscopy; CD, circular dichroism.

that contain NPNA tetrapeptide repeats in aqueous media. The results, however, are somewhat inconclusive: Esposito *et al.*[6] found no stable secondary structure in $(NANP)_2$-NA and $(NANP)_6$ in aqueous solutions. In another NMR study of peptides with the similar sequence, Dyson *et al.*[7] found the presence of a substantial population of folded conformers together with the extended form in water. On the other hand, CD measurements reported by Fasman *et al.*[8] on aqueous $(NANP)_8$ indicated a series of β-turns (or β-sheet).

It has been shown that a peptide with this specific sequence acquires the preferred secondary structure when it is incorporated in a protein.[9] That is, the CD spectra of the NANP tetrapeptide repeat have indicated a significant conformational change induced when the repeat was linked to a large protein. At the same time, the peptide immunogenicity was greatly enhanced, suggesting that the immunogenical activity is closely related to the specific conformation of the repeat. Structural dependence of this sequence on the solvent polarity has also been suggested by energy minimization calculations.[10] It is possible, therefore, that the molecule has a considerable flexibility as a free state in aqueous solution, while a stable secondary structure is induced in more hydrophobic environment such as at the receptor site.

To provide a further insight into the structural preference of this immunogenically important peptide, we report here conformational studies conducted in dimethylsulfoxide (Me_2SO) on the following peptides by 2D ^1H NMR and distance geometry calculations:

Boc-N^1-A^2-N^3-P^4-OBzl **1**

Boc-N^1-A^2-N^3-P^4-N^5-A^6-N^7-P^8-OBzl **2**

Boc-N^1-A^2-N^3-P^4-N^5-A^6-N^7-P^8-N^9-A^{10}-N^{11}-P^{12}-OBzl **3**

Boc-N^1-A^2-$(N^{(3)}$-$P^{(4)}$-$N^{(5)}$-$A^{(6)})_{n=5}$-N^7-P^8-OBzl **4**

Boc-N^1-A^2-$(N^{(3)}$-$P^{(4)}$-$N^{(5)}$-$A^{(6)})_{n=17}$-N^7-P^8-OBzl **5**

Boc-N^1-P^2-N^3-A^4-OBzl **6***

(Residue numbers in brackets denote homologous residues in the repeat. *Some of the ^1H NMR studies on **6** have already been described in ref 11.) Me_2SO-d_6 was selected as the solvent for the present ^1H NMR study in order to achieve more hydrophobic conditions. It has been suggested that Me_2SO, being composed of both lipophilic and polar portions, can mimic the receptor environments better than aqueous solutions.[12,13] The terminally blocked peptides were used also to help stabilizing the secondary structures in Me_2SO.

In recent years, many attempts have been reported to determine the solution conformations of short linear peptides, using interproton distance restraints derived from NOESY spectra.[14,15] One of the major conclusions from these studies suggests that a more realistic description of the conformations of such linear peptides is an equilibrium among multiple structures, likely including an extended form. In some cases, however, characteristic NOE patterns could be used to predict preferential conformations to be present in solution. In the present study, one of the most probable structures of a dodecapeptide, $(NANP)_3$, has been derived from distance geometry calculations using NOESY and hydrogen bond possibilities as the constraining information.

EXPERIMENTAL

Materials

Details of the synthesis and chemical purity of Boc-NA$(NPNA)_n$NP-OBzl ($n=0, 1, 2, 5, 17$; **1**—**5**) and Boc-NPNA-OBzl (**6**) have been described elsewhere.[11,16] Homogeneity of all the peptides was confirmed by amino acid and elemental analyses, as well as by high-performance liquid chromatography. ^1H NMR samples were prepared by dissolving the

peptides in Me_2SO-d_6 at an approximate concentration of 30 mM. The absence of aggregation was confirmed by comparing the spectra with those taken at 3 mM.

NMR Spectroscopy

Fourier transform ^1H NMR spectra were obtained on a Varian XL300, and a Varian VXR400 spectrometers. Temperature dependence of the NH chemical shifts were measured at 5 K intervals between 298—328 K on a Varian XL300 equipped with temperature controllers (± 1 K). All other spectra were run at room temperature (298 K). Chemical shifts were measured from Me_4Si as internal reference. The coupling constants were determined from 2D *J*-resolved spectra measured at 300 MHz using 4 K × 128 data points with a delay of 2 s, and processed as 4 K × 256. COSY and NOESY spectra were recorded at 300 MHz using 256 × 1024 data points for a spectral width of 2770 Hz and processed as 2 K × 2 K. The mixing times for NOESY spectra were 200 ms and 500 ms with a delay time 2 s. ROESY spectra were obtained at 400 MHz using a spectral width of 3550 Hz and mixing times of 60 ms for **3** and 50 and 100 ms for **6**. Typically 64 or 96 scans were collected with a relaxation delay of 3 s. NOESY and ROESY spectra were obtained in the phase-sensitive mode.

Structure Calculations

The possible structures of Boc-NA(NPNA)$_2$NP-OBzl were calculated using DADAS90 (MolScop, JEOL)[17] program developed by JEOL Co. Ltd., based on DISMAN[18] on a MIPS-RS3230 work station. The distance constraints were mainly derived from the relative peak area in the slices of NOESY spectra recorded at 200 ms and 500 ms mixing times on **3**. The NOE intensities were converted into upper limits of inter proton distances 2.5—5.0 Å at intervals of 0.5 Å, based on the NOE measured for the fixed distances between the protons of proline as the reference. Pseudo atom corrections[19] were applied to methylene and methyl groups, with an additional correction of 0.5 Å for methyl groups. All the lower distance bounds were set to van der Waals radii, using ECEPP standard sizes.[20] Hydrogen bond constraints were set to 1.5—2.3 Å for H–O, and 2.7—3.3 Å for O–N.[21] Total of 96 distance restraints were used for the calculation. Pyrrolidine ring of proline was fixed to the "down" configuration.[20] Those structures with violations in distance of >0.3 Å, or in hydrogen bond length (H–O) of >2.6 Å were eliminated.

The chemical shifts of the amide protons of Ala and Asn following Pro showed very low temperature dependence ($-d\delta/dT$, Table I, *vide infra*), indicating possible hydrogen bonding interactions. Only six types of hydrogen bond pairings are possible for these amide protons. Each of these possibilities was individually examined by the DADAS calculation by randomly generating 100 structures for each type of pairing, using the hydrogen bond constraints in addition to the NOE distance constraints. Ten each of the resulting structures with the lowest target function values were checked to see if the original hydrogen bond pairs are preserved. Only two hydrogen bond types remained. In the next step, 300 structures were randomly generated and then calculated using each of these two types of the hydrogen bond constraints and the same NOE distance constraints. Out of the 300 structures, 24 structures each with the lowest target function values were selected.

RESULTS AND DISCUSSIONS

Chemical Shifts and Coupling Constants

Figure 1 demonstrates 1D proton NMR spectra of the amide and α proton regions of series of model peptides, **1**—**5**. Individual assignments of the proton spin systems were obtained following the sequential connectivities between successive residues, based on the combined use of COSY (data not shown) and

Table I. Proton chemical shifts,[a] their temperature coefficients,[b] and $^3J_{NH\alpha}$ coupling constants[c] for the series of peptides, **1**—**4**, in Me_2SO-d_6 at 298 K

		δ_{NH} ($d\delta_{NH}/dT$)	δ_α	δ_β	$\delta_{\gamma,\delta}$, δ_{NH_2} ($d\delta_{NH_2}/dT$)	$^3J_{NH\alpha}$
1	Asn[1]	6.93 (−5.2)	4.22	2.35/2.33	7.3/6.87 (−4.56/−4.83)	8.3
	Ala[2]	7.83 (−3.77)	4.24	1.16		7.3
	Asn[3]	8.20 (−5.29)	4.82	2.46/2.28	7.3/6.87 (−4.29/−4.83)	7.6
	Pro[4]	—	4.34	2.16/1.82	1.90, 3.63	—
2	Asn[1]	6.93 (−5.5)	4.20	2.43/2.34	7.29/6.84 (−3.75/−5.21)	—
	Ala[2]	7.83 (−4.46)	4.25	1.16		7.4
	Asn[3]	8.28 (−5.33)	4.75	2.66/2.43	7.66/7.12 (−4.31/−5.06)	7.8
	Pro[4]	—	4.27	2.10/1.84	1.85, 3.68	—
	Asn[5]	7.85 (−2.79)	4.42	2.54/2.43	7.20/6.89 (−3.00/−5.63)	7.9
	Ala[6]	7.43 (−1.80)	4.17	1.21		7.7
	Asn[7]	7.86 (−5.68)	4.80	2.70/2.30	7.26/6.84 (−5.78/−5.21)	7.7
	Pro[8]	—	4.37	2.17/1.80	1.9, 3.50	—
3	Asn[1]	6.94 (−5.53)	4.20	2.46/2.38	7.29/6.87 (−4.68/—)	—
	Ala[2]	7.86 (−2.3)	4.23	1.16		7.8
	Asn[3]	8.27 (−5.31)	4.77	2.72/2.46	7.71/7.16 (−4.62/−5.06)	7.7
	Pro[4]	—	4.28	2.12/1.89	1.85, 3.71	—
	Asn[5]	7.86 (−2.27)	4.43	2.56/2.45	7.25/6.94 (−3.60/−5.53)	7.6
	Ala[6]	7.44 (−1.38)	4.18	1.21		8.0
	Asn[7]	7.94 (−4.89)	4.74	2.72/2.46	7.71/7.16 (−4.62/−5.06)	7.8
	Pro[8]	—	4.28	2.12/1.89	1.85, 3.71	—
	Asn[9]	7.86 (−2.27)	4.43	2.56/2.45	7.25/6.94 (−3.60/−5.53)	7.6
	Ala[10]	7.44 (−1.38)	4.18	1.21		8.0
	Asn[11]	7.85 (−4.68)	4.81	2.52/2.34	7.7/6.87 (−5.33/−4.68)	7.7
	Pro[12]	—	4.34	2.22/1.95	1.9, 3.62	—
4[d]	Asn[1]	6.9 —	4.2	2.5/2.4	7.27/6.85 (−3.5/−4.7)	
	Ala[2]	7.9 (−4.9)	4.2	1.2		
	Asn[(3)]	8.3[e] (−5.0)	4.8	2.7/2.4	7.7/7.2 (−4.6/−5.0)	
		7.9				
	Pro[(4)]	—	4.3	2.1/1.9	1.8, 3.7	
	Asn[(5)]	7.9 (−2.0)	4.4	2.5/2.4	7.2/6.9 (−3.3/−5.3)	
	Ala[(6)]	7.4 (−1.2)	4.2	1.2	(−3.3/−5.3)	
	Asn[7]	7.9 (−4.3)	4.8	2.5/2.3	7.7/6.9	
	Pro[8]	—	4.3	2.2/1.9	1.9, 3.6	

[a] δ in ppm from TMS used as internal reference. [b] The δ_{NH} temperature coefficients in ppb/K in parentheses. [c] $^3J_{NH\alpha}$ in Hz. [d] In this peptide, (Asn[(3)]-Pro[(4)]-Asn[(5)]-Ala[(6)]) are repeated five times; the parameters are the average, and the peaks were too broad to measure $^3J_{NH\alpha}$. [e] Amide shift of Asn[3]. All the other backbone amide protons of Asn[(3)] have the shift at 7.9 ppm.

ROESY (and/or NOESY) spectra (see below). The assignments of the amide protons of the Asn side-chain (NH_2) were mainly obtained based on ROESY spectra, since NOESY spectra did not produce good cross-peaks from some of the protons of the terminal residues, or from the Asn NH_2 protons. Otherwise, reasonably good NOESY spectra were observed except for tetrapeptides, **1** and **6**. The assignments for peptides **1**—**4** are summarized in Table I, which also contains the NH–α coupling constants.

Since both terminals are protected in the present peptides, the proton chemical shifts of Asn[1] and of the C-terminal Pro are affected by the protecting groups, and so will be excluded

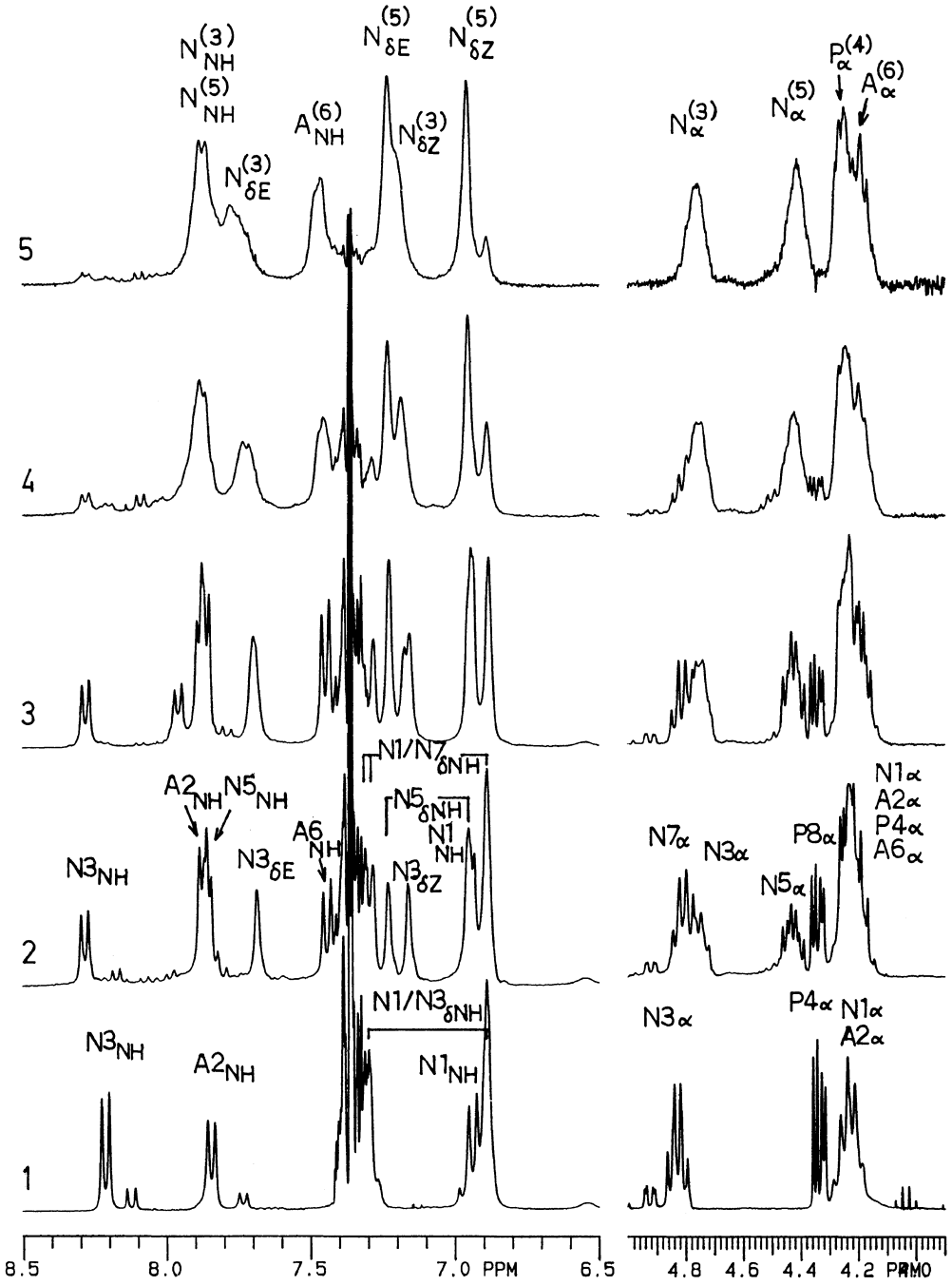

Figure 1. One-dimensional ^{1}H NMR spectra of the model peptides observed in Me_2SO-d_6 at 298 K. **1**, Boc-NANP-OBzl; **2**, Boc-NANPNANP-OBzl; **3**, Boc-NANPNANPNANP-OBzl; **4**, Boc-NA(NPNA)$_5$NP-OBzl; **5**, Boc-NA(NPNA)$_{17}$NP-OBzl. Note that the gross spectral pattern stays the same by extending the chain length by the NPNA unit.

from the following discussion.

Careful examination of the backbone amide and α-proton chemical shifts of the series of model peptides, **1**—**5** (Figure 1), shows that by extending the peptide chain by the tetrapeptide unit, NANP, Pro$^{(4)}$, Asn$^{(5)}$, and Ala$^{(6)}$ display constant shifts that are similar to the corresponding values of Pro^4, Asn^5, and Ala^6 of **2**, respectively. The amide proton shift of Asn$^{(3)}$ is unique; its first member of the series, Asn^3, shows its amide proton unusually shifted to downfield at 8.2—8.3 ppm, the same shift as Asn^3 amide proton of **1**, while the shifts of all the other protons of Asn^3 are different from those of **1** and have the same values for peptides **2**—**5**. Note that the amide proton shift of Ala$^{(6)}$ is markedly shifted upfield at 7.4 ppm compared with the shift of 7.7—7.9 ppm for Ala^2, or the random-coil chemical shift in Me_2SO-d_6 of Ala, 8.04 ppm.[22] Another case of a large shift is noted for one of the side-chain amide protons; dNH_E of the Asn$^{(3)}$ is found at 7.7 ppm, while all the other Asn side-chain amide protons are at 6.9—7.3 ppm. These unusual shifts are quite persistent throughout the present series of the peptides, **2**—**5**.

These observations suggest that the internal tetrapeptide NPNA units exist in a repeating structure, while the N-terminal Asn^1-Ala^2- and the C-terminal -Asn^7-Pro^8 residues are apparently in a different state. These terminal residues show their proton shifts very similar to those of **1** and are likely in random coil. According to the same argument, the α-proton of Asn^3 is a part of the internal repeating structure, but not its amide proton; also, the amide proton of Asn^7 is included, but not its α-proton.

Temperature Dependence of Chemical Shifts

Table I also contains the temperature coefficients, $-d\delta/dT$, for the amide protons observed in Me_2SO-d_6 in the temperature range of 298—328 K. The temperature coefficients of the backbone amide protons are plotted in Figure 2. Note the particularly small coefficients $-d\delta/dT$ for both Asn and Ala residues following Pro. A small value for this parameter is usually taken to indicate that the NH proton is shielded from the solvent by, for example, being involved in intramolecular hydrogen-bonding. Also note the periodic pattern of the variation which is quite similar for peptides **2**, **3**, and **4**; the same pattern is repeated every tetrapeptide unit, NPNA, suggesting the repeating nature of the structure. The coefficients also indicate that Asn^1 and Ala^2 are not included in the same repeats.

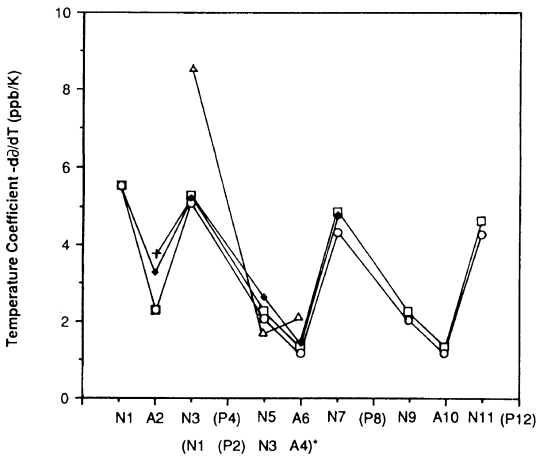

Figure 2. Temperature coefficients of peptide proton chemical shifts, $-d\delta_{NH}/dT$, in ppb/K (*i.e.*, 10^{-3} ppm K^{-1}) are plotted against residues, for **1**, Boc-NANP-Obzl; **2**, Boc-NANPNANP-Obzl; **3**, Boc-NANPNANPNANP-OBzl; **4**, Boc-NA(NPNA)$_5$NP-OBzl; and **6**, Boc-NPNA-OBzl (taken from ref 12). For **4**, only the average coefficients were plotted for N7 to A22 as N7—A10. For **6**, the plot has been shifted, so that the residues match with those of the other peptides. Note that the coefficients for the Asn and Ala that follow Pro are consistently small, suggesting possible participation of these peptide protons in hydrogen bonding. Even for the small tetrapeptide, **6**, the variation is similar. The periodic nature of the conformation is evident with 'NPNA' tetrapeptide as the repeating unit.
+, NANP (**1**); ◆, NANPNANP (**2**); □, NANPNANP-NANP (**3**); ○, (NANP)$_6$ (**4**); △, NPNA (**6**)*.

It is surprising that small tetrapeptide, **6**, also exhibits low temperature coefficients for the amide protons of Asn^3 and Ala^4, comparable to the corresponding protons in the longer chains. This strongly suggests that, even with

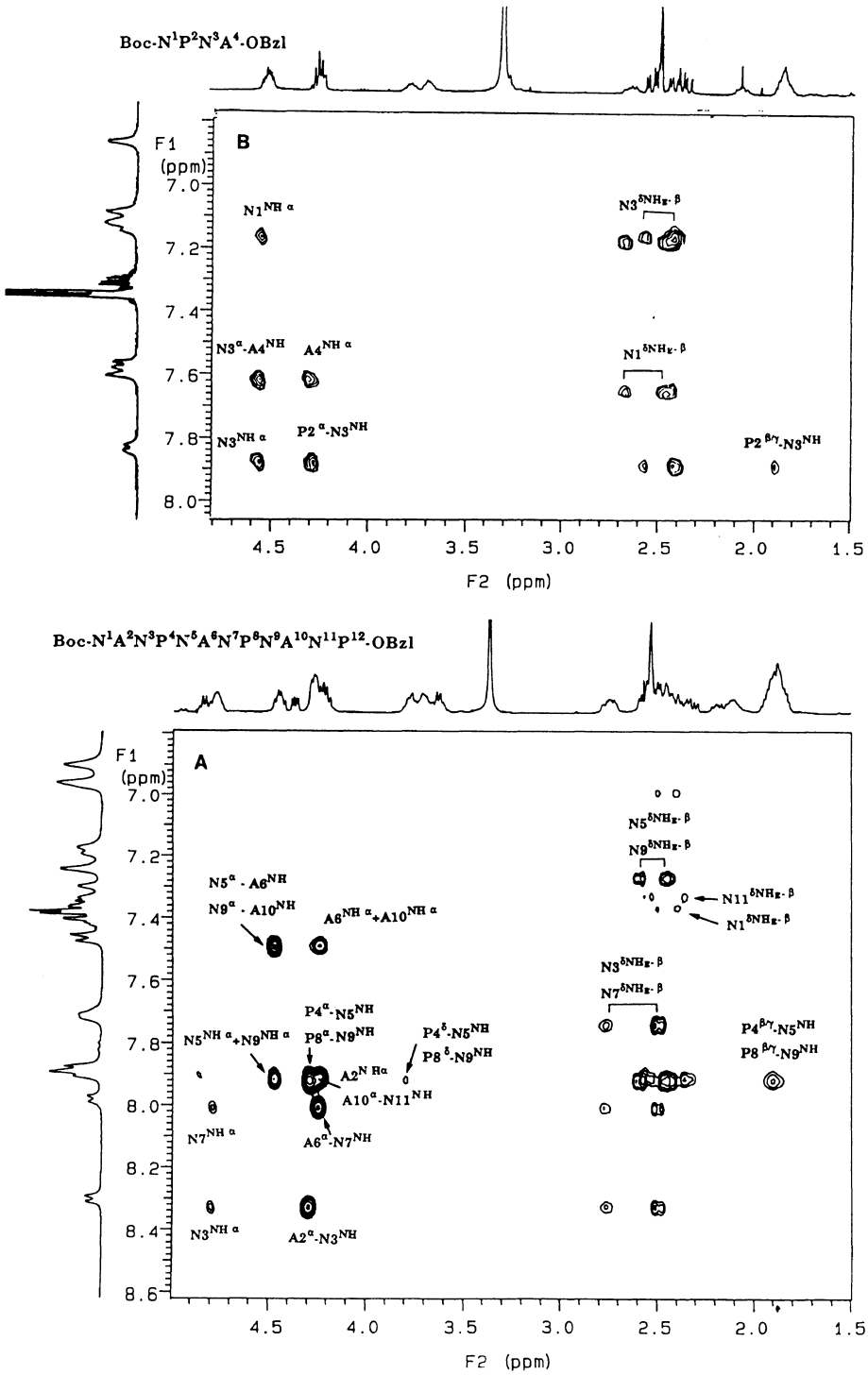

Figure 3. ROESY cross sections showing the connectivities between the peptide protons and the α—δ protons observed for (A) Boc-NANPNANPNANP-OBzl, **3**, and for (B) Boc-NPNA-OBzl, **6**. The mixing times are 60 ms for **3** and 100 ms for **6**, both measured at 298 K in Me_2SO-d_6.

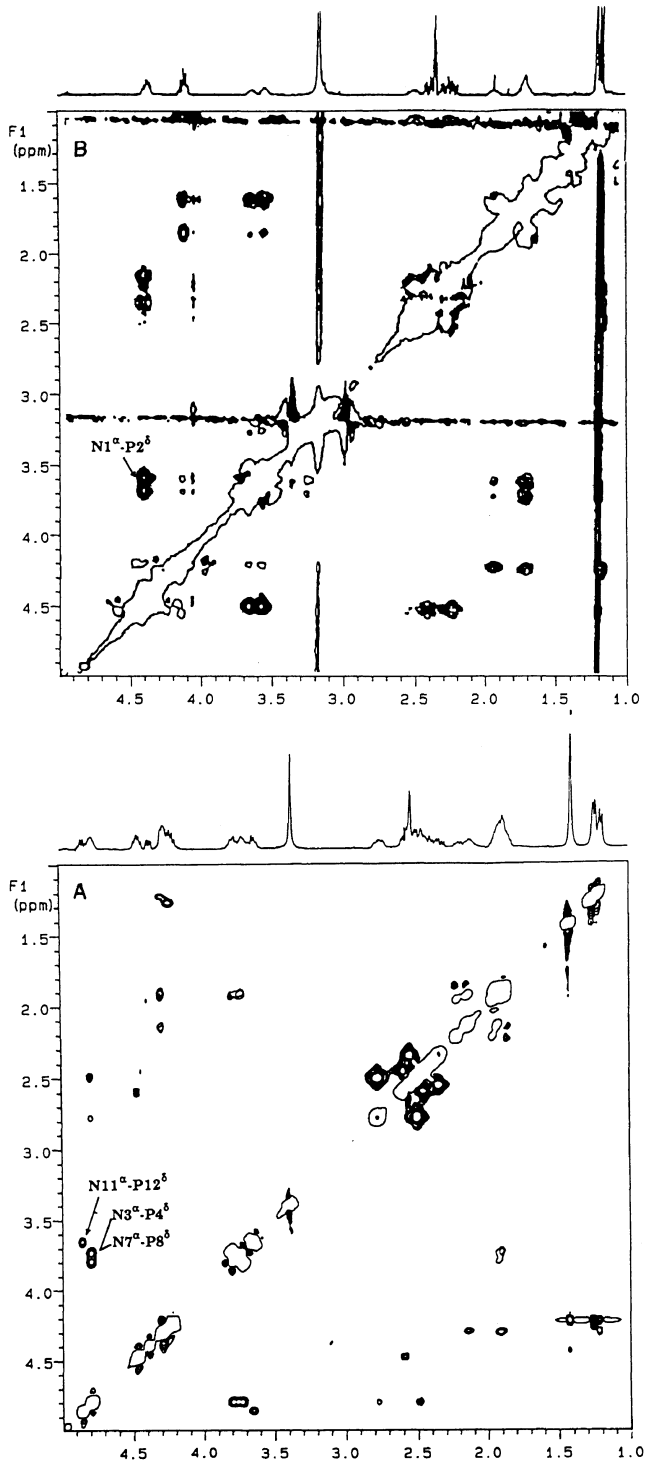

Figure 4. ROESY cross sections showing the connectivities among α—δ protons observed for (A) Boc-NANPNANPNANP-OBzl, **3**, and for (B) Boc-NPNA-OBzl, **6**. The mixing times are 60 ms for **3** and 100 ms for **6**, both measured at 298 K in Me_2SO-d_6. Note the strong cross-peak between the α proton of Asn and the δ of Pro that follows it.

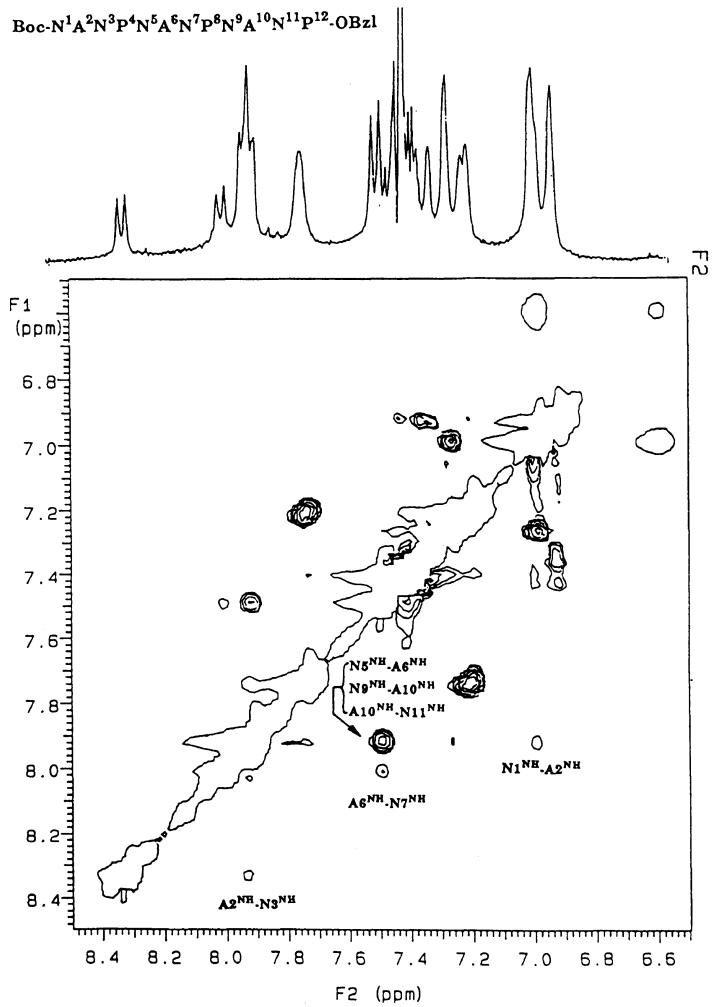

Figure 5. ROESY cross section of the peptide proton region observed for Boc-NANPNANPNANP-OBzl, **3** in Me_2SO-d_6 at a mixing time of 60 ms.

its small size, a significant proportion of its conformers exists in a folded state similar to the central part of the longer homologs. In contrast, the amide protons of another tetrapeptide, **1**, do not show such small temperature coefficients.

ROESY and NOESY Spectra

In Figures 3—5, ROESY cross sections are compared between **3** and **6**. Only the ROESY spectra are presented here for the purpose of comparison between the spectra of **3** and **6**, since only ROESY was obtained for **6**, and also because ROESY spectra were in general of better quality than the NOESY spectra. However, reasonably good NOESY spectra were obtained for peptides **2**—**5**, and were essentially the same as the corresponding ROESY spectra, except some of the cross-peaks from side chains were missing in the former.

For the amide proton-α proton connectivities in Figure 3, all the expected $d_{\alpha N}(i, i+1)$ and $d_{\alpha N}(i)$ cross-peaks are observed, except for the

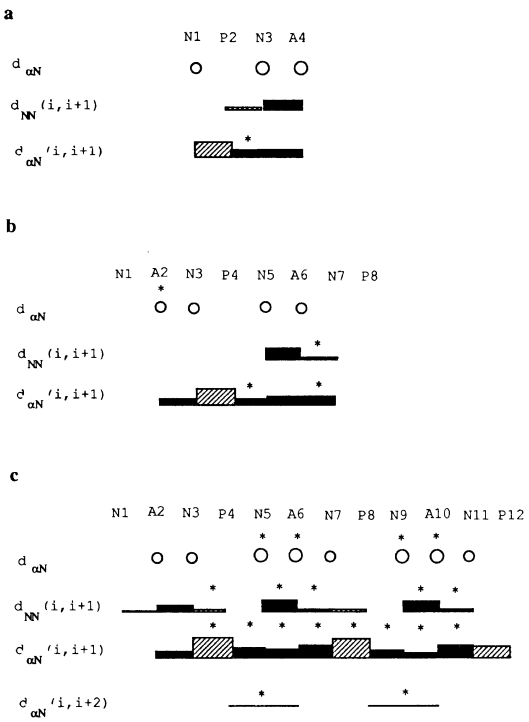

Figure 6. Summary of NOE connectivities observed for a. Boc-NPNA-OBzl, **6**, b. Boc-NANPNANP-Obzl, **2**, and Boc-NANPNANPNANP-OBzl, **3**. Data for **6** were obtained from ROESY spectra at mixing time of 100 ms and those for **2** and **3** were from NOESY spectra at mixing time of 500 ms. Asterisk indicates NOE intensities that are not quantitative due to overlapping of the NOE cross-peaks. Hatched boxes are the NOE connectivities due to the δ proton of Pro in place of the NH protons for the other residues.

N-terminal $d_{\alpha N}(1, 2)$ and $d_{\alpha N}(1)$ of **3**. The intraresidue $\delta NH_2{}^E$–β cross-peaks for all the Asn residues are also observed both for **3** and **6**. In NOESY spectra, these $\delta NH_2{}^E$–β cross-peaks were not observed.

In Figure 4, particularly strong NOEs are observed for connectivities $d_{\alpha\delta}(i, i+1)$ between the Asn residues preceding Pro and the Pro, both for **3** and **6**. All the $d_{NN}(i, i+1)$ are also observed in **3** (Figure 5) with those between the Asn-Ala connectivities particularly strong. For **6**, only one $d_{NN}(i, i+1)$ cross-peak is observed between Asn^3 and Ala^4.[11] The sequential NOE connectivities, based on ROESY spectra of **6** and NOESY spectra of **2** and **3**, are summarized in Figure 6.

Medium range cross-peaks, such as $d_{\alpha N}(i,i+2)$ or $d_{NN}(i, i+2)$, were difficult to distinguish because of overlap. This was especially so between the Asn residues corresponding to i and $i+2$. Observations of $d_{\alpha N}(i, i+2)$ between Pro and Ala were also hampered because α protons of Pro are mostly overlapped by those of Ala.

It is noted that, while all the expected $d_{\alpha N}(i)$ and $d_{\alpha N}(i, i+1)$ cross peaks are observable for **6**, those due to the terminal residues such as Asn^1, Asn^{11}, and Pro^{12} of **3** are either missing or very weak in ROESY spectra (Figure 3). This indicates that the structural unit is NPNA; these terminal residues of **3** do not constitute the repeating structure and accordingly are quite flexible, while all the four residues of **6** are involved in the structural unit. In Figure 4, particularly strong NOEs are observed for $d_{\alpha\delta}(i, i+1)$ connectivities both in **3** and **6** between the Asn residues preceding Pro and the Pro. The two δ protons of Pro show NOE of similar size with Asn-α; except, however, only one δ proton of Pro^{12} shows NOE in case of the terminal Asn^{11}-Pro^{12} sequence in **3**.

Cross-peaks due to chemical exchange were observed for some of the terminal residues: That is, Asn^1 NH proton at 6.9 ppm shows an exchange cross-peak with a small peak at 6.6 ppm due likely to the presence of the minor isomer with *cis*-Asn^1 (Figure 5). Also the two protons of δNH_2 of the C-terminal Asn residues, Asn^{11} of **3**, show an exchange cross-peak with each other, as well as those of both Asn^1 and Asn^3 in case of **6**, indicating the presence of amide bond rotation in the side groups of the terminal Asn residues. These are likely due to bond rotation and not from proton chemical exchange, since no inter-amide cross-peaks were found. These peaks have a positive phase in ROESY spectra. On the other hand, all the other cross-peaks have a negative phase indicating that they are due to NOE effects.

The repeating characteristics of the NOE

data, taken together with that of chemical shifts and their temperature dependence throughout the present series, suggest that the model peptides Boc-(NANP)$_n$-OBzl consist of a tandem repeat with tetrapeptide NPNA as the repeating unit, and that extending the chain length by the NPNA unit does not change the overall structure. Moreover, the NMR results presented above show evidences that each NPNA unit has a significant population of structured forms, in which the backbone amide protons of Asn$^{(5)}$ and Ala$^{(6)}$ either hydrogen-bonded or shielded from the solvent. Strong $d_{\alpha N}$ cross-peaks between Pro$^{(4)}$ and Asn$^{(5)}$ (Figure 3), and d_{NN} between Asn$^{(5)}$ and Ala$^{(6)}$ (Figure 5) indicate a high possibility of such structured forms to be turns of the β- or γ-type.

The present results are basically similar to those reported for the analogous peptides in aqueous solvents,[6,7] suggesting that they have contributions from similar conformers in both solvents. A difference may be pointed out, however, in the intensity of the $d_{\delta N}$ cross-peak between Pro$^{(4)}$ and Asn$^{(5)}$ (Figure 3); a much weaker cross-peak compared with the corresponding peaks in aqueous solutions shows a less possibility of a β-turn of type I in the present solvent system.

Structure Determinations

Linear short peptides in general are considered as conformationally flexible, especially in water. Even if they show characteristics of some secondary structure, they may exist as a mixture of conformational populations, including extended forms.[5]

In order to see if the observed NOE cross-peaks and the hydrogen-bond information alone can produce any reasonable structure, we applied the distance geometry calculations to the present system without using the energy term. This is equivalent to searching a structure which satisfies all the NMR data but may not necessarily be energetically favorable. If there exist a rapid exchange between several major conformers, we obtain either a virtual structure as the weighted average of all the conformers, or, more likely, no convergence at all. In more favorable and also possible situation, however, where only one major conformer is contributing to the majority of the middle-range NOE's, then we may obtain a reasonable structure with good convergence.

Table II. The structures obtained in the first step of DADAS90 calculations using NOE data and six different hydrogen bond constraints

H-Bond constraint[a]	No. of structures[b]	Possibility
$N_iC'O$–$N_{i+2}HN$	10	+
$N_iC'O$–$A_{i+3}HN$	7	+
$N_{iy}CO$–$N_{i+2}HN$	0	–
$N_{iy}CO$–$A_{i+3}HN$	0	–
$A_{i-1}C'O$–$N_{i+2}HN$	0	–
$A_{i-1}C'O$–$A_{i+3}HN$	0	–

[a] Hydrogen bond pairings of the tetrapeptide unit, NPNA, used as the constraints. [b] Number of acceptable structures generated from 100 starting structures.

Table II specifies the two hydrogen bond types remained out of possible six as the result of DADAS calculations. These two types will be denoted as the "γ-turn type", listed as $N_iC'O$–$N_{i+2}NH$, where the backbone carbonyl of Asn$^{(3)}$ is hydrogen-bonded to the amide proton of Asn$^{(5)}$, and the "β-turn type", listed as $N_iC'O$–$A_{i+3}NH$, when the same carbonyl is hydrogen-bonded to the amide proton of Ala$^{(6)}$. These two "acceptable structures" had no NOE violations over 0.3 Å or no van der Waals violations over 0.3 Å, and the hydrogen bond (H–O) distances were less than 2.6 Å. Using each of the two types of hydrogen bonding and the NOE information as the constraint, 300 structures were randomly generated, and 24 structures each with the lowest target function values were selected.

The superposed stereo view of the "γ-turn type" is shown in Figure 7(a) using VIEWER (MolSkop, JEOL). The convergence is poor, because the hydrogen-bond constraint affects only the very local part of the structure. This structure will not be considered further since

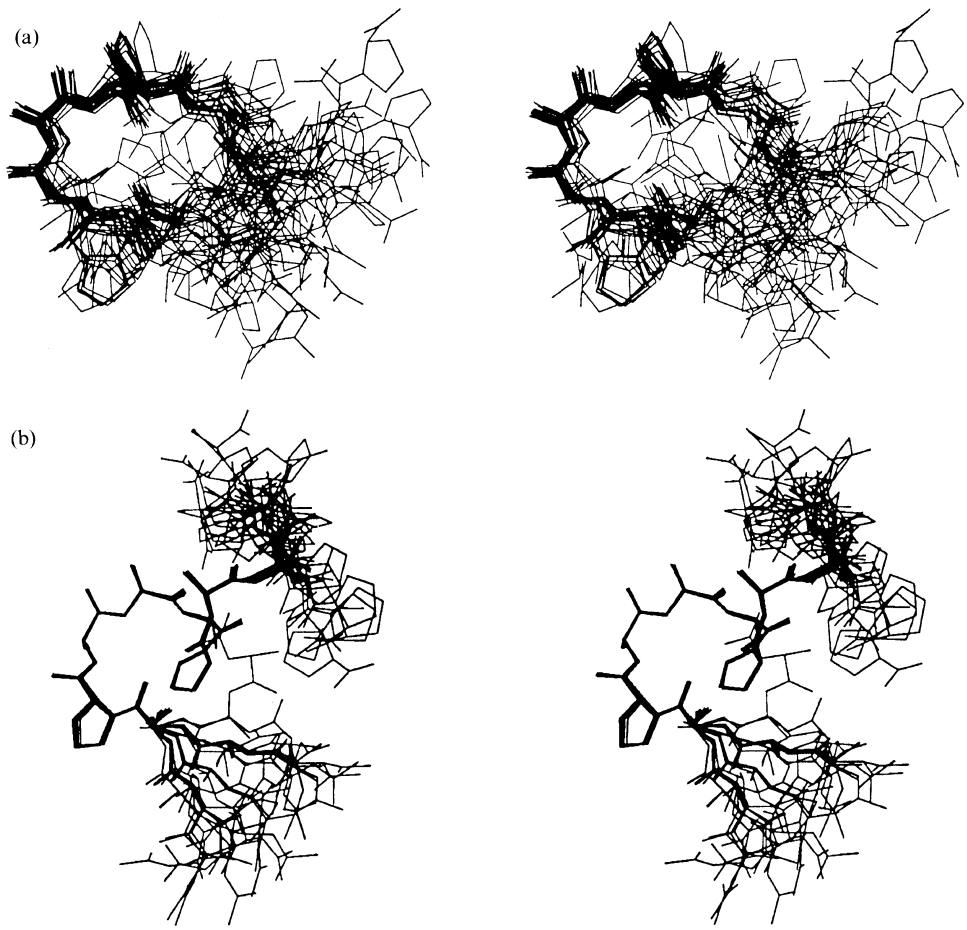

Figure 7. Stereo-views of 24 superposed structures each obtained from the DADAS90 calculations using, as the constraints, NMR data and hydrogen bond pairs; (a). $N_iC'O$–$N_{i+2}NH$ and (b). $N_iC'O$–$A_{i+3}NH$.

the structure is not consistent with the small temperature coefficient found in Ala[(6)] amide proton, and also, as discussed below, similar local structure is actually included in the following "β-turn type" structure.

For the "β-turn type", the 24 structures well converged and their stereo views are superposed in Figure 7(b). Although the three residues each at the N- and C-terminals are flexible, the central six residues exhibit a structure of good convergence, with the atomic root-mean-squared-distance difference from the averaged structure was 0.27 Å for the backbone atoms from residues 3 to 10. The main-chain dihedral angles of these central residues are as follows: Pro[4] ϕ $-75°$, ψ $60°$; Asn[5] ϕ $-120°$, ψ $-30°$; Ala[6] ϕ $165°$, ψ $75°$; Asn[7] ϕ $-175°$, ψ $-50°$; Pro[8] ϕ $-75°$, ψ $40°$; Asn[9] ϕ $-120°$, ψ $-30°$. The structure is turn-like, but it does not belong to any of the classic β-turns. The backbone carbonyl of Asn[(3)] is within the hydrogen-bonding distance to the amide proton of Ala[(6)], as in the case of any β-turn, but also makes a close contact with the amide proton of Asn[(5)]. In fact, the amide N–H bond of Asn[(5)] is extended toward the carbonyl of Asn[(3)], with the Pro[(4)] carbonyl pointing outward from the turn.

It is possible that the obtained structure is either a highly distorted β-turn or an averaged

structure resulted from β-turns of types I and II. If the structure is a distorted β-turn, both the amide protons of Asn[5] and Ala[6] are hydrogen-bonded and, therefore, the resulting structure is consistent with the observed small temperature coefficients of the amide proton shifts of these residues. The $^3J_{HN\alpha}$ value of Asn[5], 7.6—7.9 Hz is, however, slightly too small for ϕ (the literature value of $^3J_{HN\alpha}$ for $\phi = -120°$ is 9.7 Hz).[25]

Possible participation of Asn side chains to the formation of the secondary structure in the present system has been suggested.[6,7,11,26] If the distorted β-turn structure proposed here is existent in solution, it is likely to be stabilized greatly by the Asn side-chain interactions. Such side-chain hydrogen bond interaction was not included as an explicit constraint in the present study, since the number of NOE data involving the side-chain protons was insufficient to specify any particular hydrogen bond. As the result, the side chains did not converge and, therefore, were not shown in Figure 7.

In the other possibility, the derived structure shown in Figure 7(b) is an average resulted from a fast interconversion between type I and type II β-turn structures. Previous studies have suggested the presence of such equilibrium in cyclic peptide systems in solution, and observed the equilibrium to be shifted toward type II β-turn when Pro-Asn sequence is introduced in the $i+1$ and $i+2$ positions of the turn, respectively.[5] Such equilibrium alone does not explain the small temperature coefficient of Asn(5) amide proton shift, however. The $^3J_{HN\alpha}$ value, 7.6—7.9 Hz, of Asn[5] is in between the values of type I β-turn ($\approx$9 Hz) and type II β-turn ($\approx$5 Hz).[25]

In conclusion, a set of NOE data and information from the amide-shift temperature coefficients are used as constraints in distance geometry calculations. A good convergence is obtained and suggests the presence of either a unique β-turn like structure or an equilibrium between type I and type II β-turn conformations for the tetrapeptide unit in the NPNA repeat of the coat sporozoite protein of the malaria parasite.

REFERENCES

1. J. B. Dame, W. T. Hockmeyer, W.L. Maloy, J. D. Haynes, I. Schneider, D. Roberts, G. S. Sanders, E. P. Reddy, C. L. Diggs, and L. H. Miller, *Science*, **225**, 593 (1984); V. Enea, J. Ellis, F. Zavala, D. E. Arnot, A. Asavanich, A. Masuda, I. Wuaki, and R. S. Nussenzweig, *Science*, **225**, 628 (1984).
2. E. R. Clough, M. Jolivet, F. Audibert, J. W. Barnwell, D. H. Schlesinger, and L. Chedid, *Biochem. Biophys. Res. Commun.*, **131**, 70 (1985); E. R. Clough, F. M. Audibert, J. W. Barnwell, D. H. Schlesinger, and L. A. Chedid, *Infect. Immun.*, **48**, 839 (1985); L. D. Lise, D. Mazie, M. Jolivet, F. Audibert, L. Chedid, and D. H. Schlesinger, *Infect. Immun.*, **55**, 2658 (1987); E. A. Herrington, D. F. Clyde, G. Losonsky, M. Cortesia, J. R. Murphy, J. Davis, S. Baqar, A. M. Felix, E. P. Heimer, D. Gillessen, E. Nardin, R. S. Nussenzweig, V. Nussenzweig, M. R. Hollingdale, and M. M. Levine, *Nature*, **328**, 257 (1987).
3. P. Y. Chou and G. D. Fasmen, *Adv. Enzymol.*, **47**, 45 (1978).
4. C. M. Wilmot and J. M. Thornton, *J. Mol. Biol.*, **203**, 221 (1988).
5. S. J. Stadley, J. Rizo, M. D. Bruch, A. N. Stroup, and L. M. Gierasch, *Biopolymers*, **29**, 263 (1990), and references cited therein.
6. G. Esposito, A. Pess, and A. S. Verdini, *Biopolymers*, **28**, 225 (1989).
7. H. J. Dyson, A. C. Satterthwait, R. A. Lerner, and P. E. Wright, *Biochemistry*, **29**, 7828 (1990).
8. G. D. Fasman, K. Park, and D. H. Schlesinger, *Biopolymers*, **29**, 123 (1990).
9. G. E. Armah, S. Nishikawa, S. Miki, Y. Omata, T. Nakabayashi, and K. Tomita, *Mol. Biochem. Parasit.*, **38**, 135 (1990).
10. K. D. Gibson and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 5649 (1986).
11. S. Isokawa, K. Umemoto, and M. Narita, *Makromol. Chem.*, **194**, 3247 (1993).
12. G. A. Hashim, *Immunol. Rev.*, **39**, 60 (1978).
13. J. Saulitis, D. F. Mierke, B. Gerardo, C. Gilon, and H. J. Kessler, *J. Am. Chem. Soc.*, **114**, 4818 (1992).
14. H. J. Dyson and P. E. Wright, *Annu. Rev. Biophys. Biophys. Chem.*, **20**, 519 (1991), and references cited therein.
15. M. P. Williamson and J. P. Waltho, *Chem. Soc. Rev.*, 227 (1992).
16. M. Narita, J. Takegahara, S. Ono, and H. Sato, *Bull. Chem. Soc. Jpn.*, **63**, 484 (1990).
17. S. Endo, H. Wako, K. Nagayama, and M. Go, in "Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy", J. C.

Hoch, F. M. Poulsen, and C. Redfileld, Ed., Plenum Publishing Corp., New York, N.Y., 1991.
18. W. Brown and N. Go, *J. Mol. Biol.*, **186**, 611 (1985).
19. K. Wuthrich, M. Billeter, and W. Braun, *J. Mol. Biol.*, **169**, 949 (1983).
20. G. Nemethy, M. S. Pottle, and H. A. Scheraga, *J. Phys. Chem.*, **87**, 1883 (1983).
21. J. D. Forman-Kay, G. M. Clore, P. T. Wingfield, and A. M. Gronenborn, *Biochemistry*, **30**, 2685 (1991).
22. A. Bundi, C. Grathwohl, J. Hochmann, R. M. Keller, G. Wagner, and K. J. Wuthrich, *Mag. Reson.*, **18**, 191 (1975).
23. J. S. Richardson, *Adv. Protein Chem.*, **34**, 167 (1981).
24. G. Wagner, A. Pardi, and K. Wuthrich, *J. Am. Chem. Soc.*, **105**, 5948 (1983).
25. A. Pardi, M. Billeter, and K. Wuthrich, *J. Mol. Biol.*, **180**, 741 (1984).
26. A. C. Satterthwait, T. Arrhenius, R. A. Hagopian, F. Zavala, V. Nussenzweig, and R. A. Lerner, *Phil. Trans. R. Soc. Lond.*, **B323**, 565 (1989).