Solution Structure of the DNA-Binding Domain of TraM[†,‡]

Thomas Stockner,[#] Carmela Plugariu,[#] Günther Koraimann,[§] Gregor Högenauer,[§] Wolfgang Bermel,[◇] Stefan Prytulla,[⊥] and Heinz Sterk*,[#]

Institute of Chemistry, Karl Franzens University Graz, Heinrichstrasse 28, A-8010 Graz, Austria, Institute of Molecular Biology, Biochemistry and Microbiology, Karl Franzens University Graz, Universitätsplatz 2, A-8010 Graz, Austria, Bruker Analytik GMBH, Silberstreifen, Rheinstetten/Karlsruhe, Germany, and Institute of Pharmazeutical Chemistry, Karl Franzens University Graz, Schubertstrasse 6, A-8010 Graz, Austria

Received August 28, 2000; Revised Manuscript Received January 17, 2001

ABSTRACT: The solution structure of the DNA-binding domain of the TraM protein, an essential component of the DNA transfer machinery of the conjugative resistance plasmid R1, is presented. The structure has been determined using homonuclear 2-dimensional NMR spectroscopy as well as ^{15}N labeled heteronuclear 2- and 3-dimensional NMR spectroscopy. It turns out that the solution structure of the DNA binding domain of the TraM protein is globular and dominantly helical. The very first amino acids of the N-terminus are unstructured.

TraM protein is an essential part of the DNA transfer machinery of the conjugative resistance plasmid R1 (IncFII). On the basis of mutational analyses, it was shown that the essential transfer protein TraM has at least two functions. First, a functional TraM protein was found to be required for normal levels of transfer gene expression (*1*). Second, experimental evidence was obtained that TraM stimulates efficient site-specific single-stranded DNA cleavage at the *oriT*, in vivo (*2*). Furthermore, a specific interaction of the cytoplasmatic TraM protein with the membrane protein TraD was demonstrated (*3*), suggesting that the TraM protein creates a physical link between the relaxosomal nucleoprotein complex and the membrane-bound DNA transfer apparatus.

The immediate translation product of the *traM* gene of plasmid R1 comprises 127 amino acids (Swissprot entry number P07294). Methionine, the first amino acid, is then cleaved off, and the mature TraM protein consists of 126 amino acids (residues 2−127).

Experimental evidence suggests that both the sequence specific DNA-binding and very likely also the autoregulatory capacity of the TraM protein is conferred by the N-terminal 22 amino acids of the protein (*1*, *2*, *4*).

Recently, it was demonstrated that TraM forms tetramers in solution and binds as a tetramer to its binding sites in the plasmid R1 *oriT* DNA region (*5*). Because of tetramerization, the TraM wild-type protein is unsuitable for structural determination using NMR methods.

Therefore, an N-terminal TraM variant that represents the DNA-binding domain was chosen for structural analysis, as it does not form tetramers (*5*). This C-terminal deletion mutant designated TraMM26 comprises amino acids 2−56 of TraM; although it does not produce detectable footprints (unpublished results) nor clear bandshifts (*3*), this protein retains the 55 amino acids of TraM enabling specific DNA recognition and binding (*1*, *4*). It lacks, however, the domain that is required for tetramer formation. The evidence for a biological function of TraMM26 in conjugation comes from studies on the conjugative plasmid R1-16 in which the *traM* wt gene was replaced by the *traMM26* allele. In this case, the transfer frequencies were approximately 2 orders of magnitude above the transfer frequencies when compared with the *traM* null allele (*1*). The functional relevance of TraMM26 is also underscored by the results of experiments in which the DNA binding activity of various TraM variants was measured indirectly by using a reporter gene construct containing the TraM binding sites and measuring β-galactosidase expression. Whereas with TraMM26 there was a significant detectable activity of TraM, the activity of TraM was completely lost when only three N-terminal amino acids in the wild type protein were removed (*1*).

As a result of the NMR analysis, we provide the first NMR structural information on the DNA-binding domain of a plasmid encoded relaxosomal protein.

MATERIALS AND METHODS

Protein Expression and Cell Growth. Expression and purification of the *Escherichia coli* plasmid pExtraMM26, a derivative of pExtraM with the M26 mutation in the traM gene (*1*), was performed as described by Verdino et al. (*5*). For the expression of TraMM26 in minimal medium, another TraMM26 expression construct was created by cloning the *Xba*I/*Hin*dIII segment from pExtraMM26 into the pET28a (+) vector (Novagen). This construct was transformed into *Escherichia coli* BL21 (DE3) and the cells were grown in minimal medium (*6*) containing ^{15}N ammonium chloride as sole nitrogen source, under kanamycin selection, at 37 °C, to an OD_{600} of 0.60. Protein expression was induced by

[†] This research was supported by the "Fonds zur Förderung der wissenschaftlichen Forschung" project numbers 11048, 11844, and 12881 as well as the Österreichische Nationalbank Nr. 7062.

[‡] PDB 10 1DP3, RCSB 10 RCSB010259.

[#] Institute of Chemistry, Karl Franzens University Graz.

[§] Institute of Molecular Biology, Biochemistry and Microbiology, Karl Franzens University Graz.

[◇] Bruker Analytik GMBH.

[⊥] Institute of Pharmazeutical Chemistry, Karl Franzens University Graz.

10.1021/bi002031c CCC: $20.00
Published on Web 02/23/2001

adding IPTG[1] to a concentration of 1 mM, and growth was continued for additional 3 h (final $OD_{600} = 1.10$). The cells were then kept for 10 min on ice, harvested, and stored at −20 °C.

Protein Purification. All the procedures described below were performed at 4 °C. Typically, protein extracts were prepared from cell pellets of 500-mL cultures using a french pressure cell for disrupting the cells. The soluble material containing from 30 to 50 mg of total protein was then used for the subsequent chromatography steps. Protein separations were performed with a FPLC system from Pharmacia Biotech, Sweden. The cytoplasm extract was loaded onto a MonoS HR10/10 cation-exchange column, preequilibrated with 50 mM Tris-phosphate buffer, pH 6.4, at a flow rate of 3.9 mL/min. The protein was eluted using a 200-mL NaCl linear gradient from 0 to 250 mM. The fractions containing the protein were eluted between 70 and 110 mM NaCl, and they were pooled and dialyzed overnight against 50 mM NaH_2PO_4-phosphate buffer, pH 3.8. The dialyzed material was concentrated by ultrafiltration (Amicon). Approximately 3 mg of protein was loaded onto a Superose 12 XK 16/70 gel-filtration column and eluted with 50 mM NaH_2PO_4 buffer, pH 3.8, at a flow rate of 0.8 mL/min. After this step, the protein was of high purity as was judged by the appearance of a single band corresponding to the expected molecular weight of TraMM26 on a SDS−polyacrylamide gel (data not shown).

Protein solutions from several runs were pooled and concentrated to 4 mg/mL using a YM3 membrane in an Amicon-stirred cell. Because of the low solubility within a range of the pH from 4.5 to 10.5, attempts to concentrate the protein at pH 6.4 led to the loss of the product.

HPLC and MALDI Analyses. The authenticity and purity of the protein was checked by HPLC and MALDI-TOF. The analyses were performed by Pi-Chem Graz, Austria. Protein samples were loaded onto a Lichrosphere 100 RP-18 (125 × 4 mm, 5 μm) or Vydac RP-18 (150 × 4 mm) column and were eluted in reverse phase with water (0.l% TFA)/acetonitrile (0.1% TFA) using a gradient of acetonitrile from 10 to 90% in 30 min, flow rate 1 mL/min. The protein was eluted as a broad peak on Lichrosphere and as a sharp peak on the Vydac column. The eluted protein was separated in two fractions and analyzed by MALDI-TOF, using insulin as an internal standard (5.734 kDa). The mass spectroscopy indicated a mixture of two proteins with molecular weight of 6.392 and 6.262 kDa, respectively. The difference in mass corresponds to a methionine (*7*, *8*). However, within these two components a slight mass heterogeneity can be detected, as deduced from the splitting of the mass peaks (data not shown). In our opinion, these differences stem from the alkaline counterions that are attached to Asp and Glu and give rise to [AA-Na_x-H^+, AA-H^+].

Size-Exclusion Chromatography. Size-exclusion chromatography under physiological pH conditions was performed on a Superose 12 HR 10/30 column (flow rate of 0.4 mL/min) for TraM wild type and the TraMM26 form. TraMM26 was eluted also with 50 mM NaH_2PO_4-phosphate buffer, pH 4.0. The calibration curve of the column was obtained using bovine serum albumin (Pharmacia) 66.2 kDa, β-lactoglobulin (ICN) 14.4 kDa, cytochrome C (Sigma) 12.4 kDa, aprotinin (Boehringer) 6.5 kDa, and vitamin B12 (Merck) 1.35 kDa as standards.

The TraMM26 appeared as a single peak in the size-exclusion chromatography corresponding to a dimer with a molecular mass of 12.6 kDa. At pH 4.0, TraMM26 eluted as a protein with approximately 6.2 kDa molecular mass, corresponding to the monomeric form.

Circular Dichroism. Circular dichroism spectra were recorded on a Jasco J-715 spectrometer using a 1.5 mg/mL protein solution in 50 mM NaH_2PO_4-phosphate buffer, pH 4.0 and 6.4. The spectra were the average of three scans run at 50 nm/ min with a 1 s response time. The thermal stability of the protein at pH 4.0 was evaluated by monitoring the ellipticity at 222 nm in the interval from 25 to 70 °C, in 5 °C steps. These spectra were taken with one scan, 50 nm/min speed and 64 ms response time.

The protein was found to be extremely stable up to a temperature of 55 °C, when major changes in the ellipticity value at 222 nm occurred (data not shown). These changes are due to the loss of the helical structure and the start of defolding.

NMR Spectroscopy. The protein samples for NMR spectroscopy consisted of 800 μL 4 mg/mL unlabeled or 1.3 mg/mL ^{15}N-labeled protein solution in 50 mM NaH_2PO_4-phosphate buffer, pH 4.0, containing 5% D_2O (Aldrich) and 0.02% sodium azide. All NMR measurements were performed at 300 K. ^1H chemical shifts were referenced to the H_2O resonance at 4.78 ppm.

The spectra were recorded using a Bruker Avance 800 and 600 MHz spectrometer as well as a Varian 600 MHz Inova Unity spectrometer. TOCSY and NOESY experiments (*9*, *10*) were acquired using spectral widths of 8000 or 10 000 Hz in each dimension. For the ^{15}N dimension, a 1785-Hz spectral width was used, and the carrier was positioned at 117.112 ppm. Homonuclear shift correlations resulted from TOCSY experiments with 30, 40, 60, and 70 ms mixing time and MLEV17 spin-lock (*11*) and NOESY experiments with 150 and 300 ms mixing time.

The ^{15}N resonances were assigned using 2D ^{15}N-HSQC spectra (*12*, *13*), 3D ^{15}N-TOCSY-HSQC spectra (*14*), and 3D ^{15}N-NOESY-HSQC spectra (*15*, *16*). In homonuclear experiments, the quadrature detection was obtained by using States TPPI in the case of heteronuclear experiments Echo-AntiEcho-TPPI (*17*) was used. The spin-lock time in the 3D ^{15}N-TOCSY-HSQC was set to 30 ms, and the mixing time for the 3D ^{15}N-NOESY-HSQC was 150 ms. To get the values of the scalar coupling constants, an HNHA experiment (*18*) was performed. In all experiments, water suppression was performed using the WATERGATE technique (*19*).

The spectra were analyzed using ANSIG, version 3.3 (*20*, *21*). Besides the system-specific software, all spectra were processed using nmrPipe (*22*). The structure calculation was done using X-PLOR 3.851 (*23*) using standard structure

[1] Abbreviations: Standard abbreviations are used for the common amino acids. MALDI-TOF: matrix-assisted laser desorption time-of-flight mass spectrometer, FPLC: fast protein liquid chromatography, HPLC: high-pressure liquid chromatography, HSQC: hetero single quantum coherence, HNHA: *J* correlation between the NH and H_α proton, TOCSY: total correlation spectroscopy, NOE: nuclear Overhauser effect, NOESY: Nuclear Overhauser spectroscopy, TPPI: time proportional phase incrementation, MLEV17: Malcom Levitt pulse cycle, DQF-COSY: double quantum-filtered correlation spectroscopy, IPTG: isopropyl -β-D-thio-galactopyranoside, SDS: sodium dodecyl sulfate, TFA: trifluoroacetic acid.

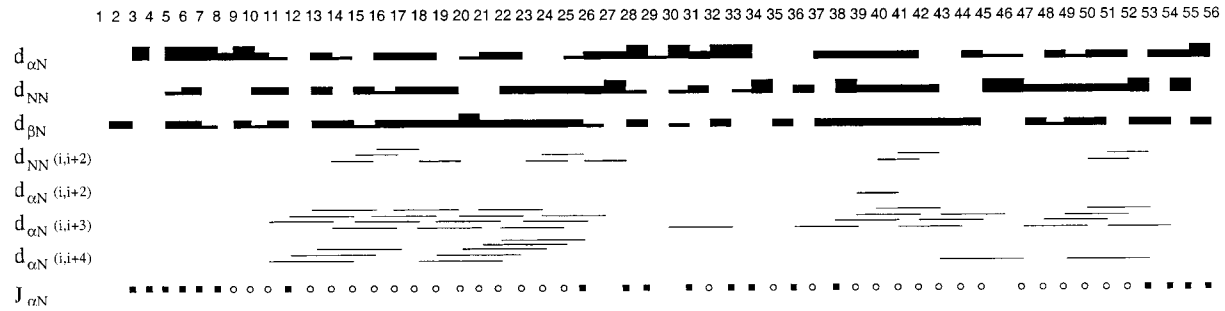

FIGURE 1: Summary of all NOEs (sequential and secondary structure) involving HN, $C_\alpha H$ and $C_\beta H$ protons, and $^3J_{HNHA}$ coupling constants. The relative strengths of the NOEs are classified into strong, medium, and weak, as indicated by the thickness of the lines. The strength of the $^3J_{HNHA}$ coupling constants are indicated by filled squares ($8 > {}^3J_{HNHA} > 5$) and open circles ($^3J_{HNHA} < 5$).

calculation protocols (*24*−*26*). The analyses of the structures were done using MOLMOL 2.6 (*27*) and Procheck 3.5 (*28*, *29*).

RESULTS

Assignment of the NMR Spectra. (a) NOE-Derived Distance Restraints, $^3J_{HN-H\alpha}$ Scalar Couplings, Relaxation Behavior, and CD Spectra. A first approach to the sequence-specific assignment of TraMM26 was the achievement of the 2D-TOCSY and NOESY spectra. The availability of the completely ^{15}N labeled protein allowing the recording of 3D ^{15}N-TOCSY-HSQC and ^{15}N-NOESY-HSQC spectra was of crucial importance for the assignment because the proton chemical shift dispersion of TraMM26 is very poor and severe overlap prohibited unambiguous assignment of the homonuclear spectra.

The ^{15}N dimension facilitated the assignment and the correct identification of the spin systems by the combined use of 2D-TOCSY, 2D-^{15}N-HSQC, and 3D-^{15}N-TOCSY-HSQC spectra. The spin system and sequential assignment was performed according to the procedure described by Wüthrich (*30*). Sequential assignment was done using 2D-NOESY and 3D-^{15}N-NOESY-HSQC spectra. The $d_{\alpha N}(i,i+1)$ connectivities were used as starting points. These connectivities, together with a large number of sequential d_{NN} and $d_{\beta N}$ NOEs cross-peaks, served to reveal the whole sequence. NOE assignment was done in the 3D-^{15}N-NOEDY-HSQC and in addition as many NOEs as possible were extracted from the homonuclear 2D-NOESY spectrum. The NOE intensities were calibrated by referring to known distances in α-helices of the spectrum (Figure 1). The intensities were converted into the distance restraints classes: short (1.8−2.8 Å), medium (1.8−4.0 Å), and long (1.8−6.0 Å).

The dihedral $^3J_{HN-H\alpha}$ coupling constants have been extracted from the HNHA spectrum in the usual way by comparing diagonal and cross-peak intensities (Figure 1). Afterward, the coupling constants were converted into dihedral angle restraints, where a deviation of ± 20 degrees within the α-helices and a deviation of ± 40 degrees in the loop part of the structure was allowed. At the N-terminal end (first nine residues) and at the C-terminal end (last three residues), no dihedral restraints were used for the structure calculation as all results (NOEs and coupling constants) indicated clearly a random coil structure. For a comparison of the $^3J_{HN-H\alpha}$ coupling constants, see Figure 1. In principle, the coupling constants within the helical regions are between 3 and 5 c/s, whereas for the N-terminal regime coupling constants between 5 and 8 c/s are found (typical random coil structures values). The coupling constants of the amino acids in the region between 28 and 33 differ from amino acid to amino acid and therefore cannot be attributed to a regular secondary structure like a helix or sheet, but it could indicate a loop-like structure. Within the second helix, the $^3J_{HN-H\alpha}$ coupling constant of Thr^{40} is extremely low. In our opinion, there are two reasons for this finding. First, the intensity of the cross-peak is very low, which hampers the accuracy of the integral determination; second, a deviation from the ideal helical fold can be expected, since the sequence Ser^{39}−Thr^{40} shows strong steric effects due to the β branching of Thr and a possible interaction of the Ser^{39}−Thr^{40} side chains with the carbonyl oxygen atoms of the preceding helical turn (*31*). However, it must be stated that the scalar coupling constants corroborate all secondary structure elements indicated by the short and medium range NOEs.

As already discussed, the MALDI-TOF mass spectrometry analysis evidences the presence of two forms of the TraMM26 protein in the sample (see Materials and Methods). This finding is in agreement with the NMR studies. It can be unambiguously observed that in one peptide chain, the starting amino acid Met^1 is still present, whereas in the other case Met^1 is cleaved and the sequence starts with Ala^2 (*7*, *8*). For the structural assignment, this leads to the observation that the first three amino acids Ala^2, Lys^3, and Val^4 show a reasonable chemical shift difference. Accordingly, the two amino acid chains also differ in the HN behavior. The protein which includes the leading Met^1 shows the Ala^2 HN couplings and NOE cross-peaks. However, as it turned out that the N-terminal end of the protein is unstructured and no long-range NOEs can be observed, no huge effect of this difference can be expected. To have one notation for both chains, we designate the first amino acid as Met^1 and the second as Ala^2. The sequence-specific assignment of TraMM26 is reported in the Supporting Information. The structure calculations will only be done on the shorter sequence beginning with Ala^2.

To get some information concerning the relaxation, ^{15}N-HSQC experiments with a variable delay to evolve spin lattice relaxation [compare Ziegler and Sterk (*32*)] were measured (data shown in the Supporting Information). The fairly long ^{15}N T_1 relaxation time of ∼ 0.6−0.7 [s] for most of the backbone is, at least in the case of a globular protein, best explained by assuming a molecular mass between 6 and 8 kDa, which corresponds with the TraMM26 monomer.

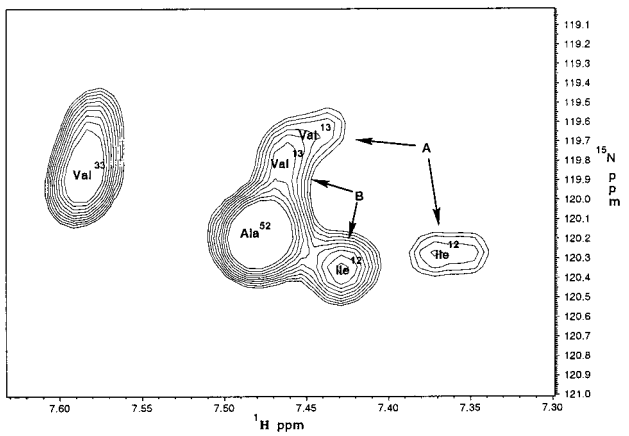

FIGURE 2: An expansion of the ^{15}N-HSQC spectrum is depicted, showing the cross-peaks of Ile[12] and Val[13]. A and B belong to the conformers A and B, respectively.

However, TraMM26 is too small to lead to pronounced effects arising from a different flexibility along the amino acid chain. Solely, the ^{15}N relaxation time of Phe[35] and Tyr[7] are remarkably short. We think that the existence of a conformer (will be discussed later in greater detail), where the aromatic side chains of Phe[35] and Tyr[7] are in close contact and thereby reduce the segmental mobility, is responsible for this observation. Figure 1 shows the short- and medium-range NOEs that were found for TraMM26 protein. The combination of characteristic medium-range d_{NN} (i,j) and $d_{N\alpha}$-(i,j) NOEs and the coupling constants indicate that an α-helix is starting at Asp[10] and extends up to Glu[26]. The C-cap of this helix (Gly[27]) is formed by a Gly Schellman motif (*31*, *33*−*35*). The N-cap amino acid of the second α-helix is Ser[34], and the motif is of box Ia type (*34*, *36*). Unfortunately, the serine side chains and the first medium range $d_{\alpha N}$ connections (the H_α signals of Ser[34] and Phe[35] residues) of this helix are in close proximity to the water resonance, which prevents them from being used as restraints. Nevertheless, the starting point of helix 2 can be defined unambiguously as the 34 HN−37 HN, 34 HN−38 HG2#, and the 35 HB−38 HG2# cross-peaks are clearly visible. According to the typical i → i + 3 NOE cross-peaks, the second α-helix extends up to Leu[45]. The C-cap of the second α-helix forms also a Schellman motif. Helix 3 starts almost immediately after helix 2 with Leu[47]. A series of $d_{\alpha N}$ cross-peaks as well as the values of the $^3J_{HN-H\alpha}$ coupling constants indicate that the helix extends up to Gln[53]. The loop connecting helix 1 and helix 2 lacks medium range connectivities characteristic for helices or *β*-sheet structures. No long-range NOEs can be found. Only a reduced number of side chain contacts is present, the coupling constants are of variable size, and the conformation of the calculated structures are poorly defined. The last three amino acids as well as the very first eight amino acids are also random coils. Their i → i + 1 NOE cross-peaks are strong, the scalar coupling constants are within the typical random coil range, and no evidence of a secondary structure can be found.

However, it is of interest that within the area of the unstructured N-terminus two different slowly (lifetime ≫ 10 ms) converting conformers (conformer A and conformer B) can be found. Both conformers can be clearly identified and show distinct NOE patterns (Figure 2). But while no NOE contacts can be found for conformer A (roughly 30−40%)

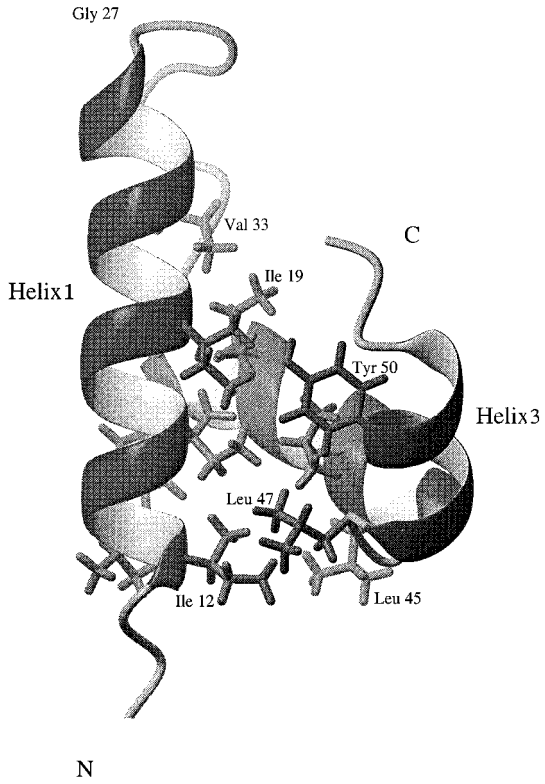

FIGURE 3: Ribbon diagram of a final structure of TraMM26. The secondary structure elements are defined as follows: coil: residues 1−10; helix 1: residues 10−26; loop: residues 26−34; helix 2: residues 34−45; turn: residues 46−47; helix 3: residues 47−53; coil: residues 53−56. N: N-terminus, C: C-terminus, all side chains of the hydrophobic core are shown. The core holds the three helices together. This figure was created with MOLMOL (*27*).

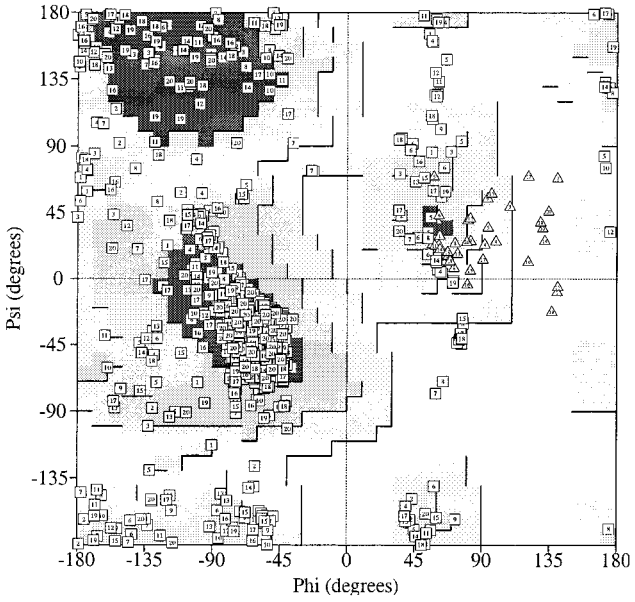

FIGURE 4: Distribution of the backbone dihedral angles ϕ and ψ of conformer B of the final 20 TraMM26 structures calculated by X-PLOR is shown. The Ramachandran plot of conformer A is very similar. The Gly residues are represented by triangles. The figure was created with PROCHECK (*28*, *29*).

between the unstructured region and the globular protein domain, in conformer B (roughly 60−70%) three additional NOE contacts form Val[6] and Tyr[7] to Ser[39] and Thr[40] which hold the N-terminal tail down to the helix 2 can be identified. This structural arrangement brings the aromatic moieties of

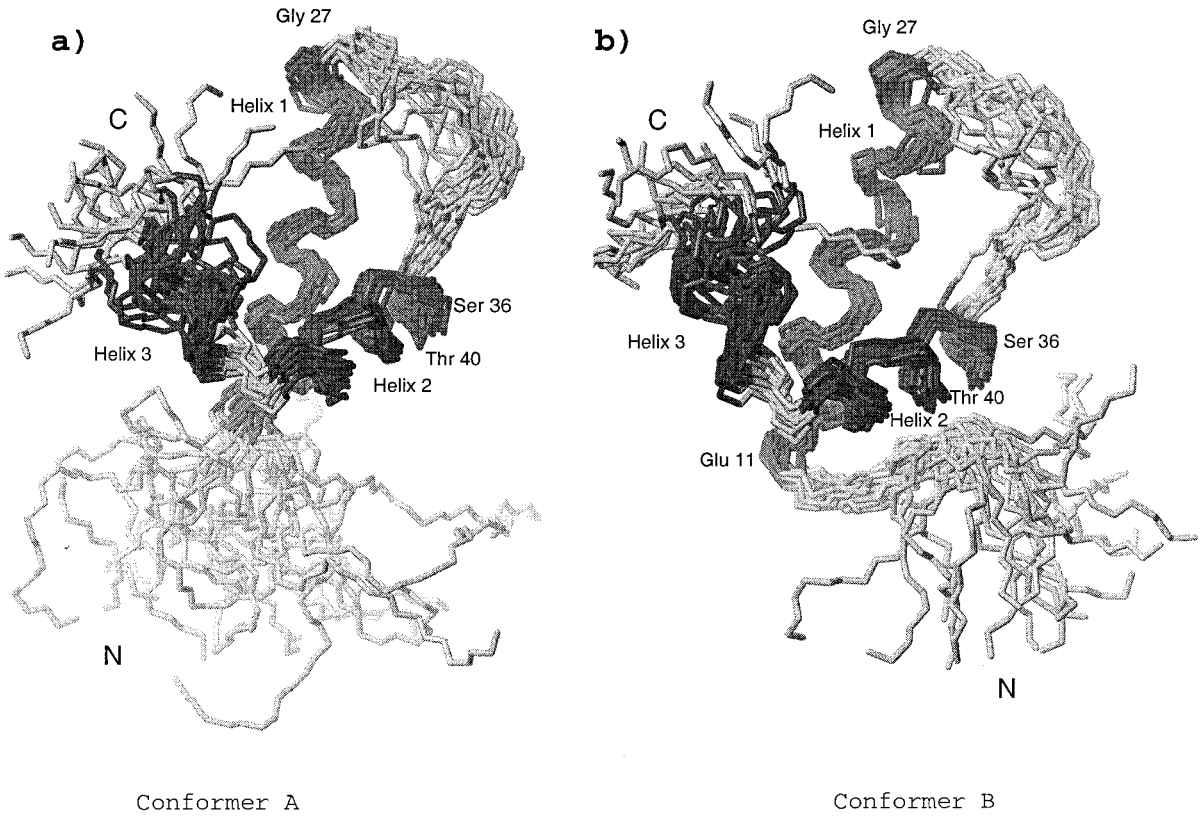

FIGURE 5: Overlay of 20 calculated structures. The three helices i.e., residues 10−26, 34−45, and 47−53 are shaded. The structures were calculated by X-PLOR. (a) Overlay of 20 structures of conformer A without contacts between the N-terminus and the second helix. (b) Overlay of 20 structures of conformer B with NOE contacts between the N-terminus and the second helix. The conformational freedom of the N-terminal part (N) in conformer B is reduced by the interaction with the second helix.

Tyr[7] and Phe[35] in close contact, providing the basis for the chemical shift differences of the aromatic protons Tyr[7] and Phe[35] and also accounts for the large shift difference observed for Val[8], and some minor shifts of the neighbors, as well as the relaxation anomaly mentioned above. The amino acids and their chemical shift values that belong to the conformer A and to the conformer B are listed in the Supporting Information.

CD spectra, "chemical shift indexing" (*37*) (data shown in the Supporting Information), and "secondary structure prediction methods" (*38*) corroborate this result in a perfect manner. So, for example, the shapes of the CD spectra as well as the high value of the mean residue ellipticity are consistent, using the convex constraint analysis from Percel et al. (*39*), with large (60−70%) α-helical contents. Thereby, the monomeric TraMM26 at pH 4.0 [the molar ellipticity reads at 192 nm = 65 000, at 209 nm = −28 500, and at 222 nm = −27 500] was found to have higher α-helical content as compared to the, at least, partial dimer at pH 6.4. The "chemical shift indexing", which has been performed using the H_α protons, revealed also three helices from Ser[9] to Glu[26], from Ser[36] to Glu[44], and from Leu[47] to Glu[55]. The results of "secondary structure prediction methods" are also in line with our NMR results showing the helical parts and random coils at the proper position within the TraMM26.

Structure Calculations. The assigned NOE cross-peaks were converted into distance restraints and the HNHA intensities into dihedral restraints. For the calculation of the final structure, an extensive calculation was performed calculating 100 structures using the standard X-PLOR protocols of Distance Geometry (*24*, *25*) and Simulated Annealing (*26*). The CHARMM all hydrogen force field was used (*40*). The best 20 structures were selected for further statistical analysis. As the two conformers show clear and separated cross-peak sets and since no exchange peaks could be found, distance restraints sets were extracted separately for both conformers A and B, and structure calculation was done for each conformer.

The calculated structures reflect perfectly the indications extracted from the different NMR and CD experiments. They show the three α-helices from Asp[10] to Glu[26], from Ser[34] to Leu[45], and from Leu[47] to Gln[53]: the Schellmann motifs at the end of helix 1 and 2 with their characteristic positive ϕ angle, the loop connecting helix 1 and helix 2, and the globular domain formed by the three helices (Figure 3). The ϕ and ψ angles of both conformers of the 20 selected structures fit quite well into the allowed regions of the Ramachandran plot. As an example, the Ramachandran plot of conformer B is shown in Figure 4. The globular domain of both conformers appears to be very similar. Comparison of Figure 5, panels a and b, indicates that the differences are at the conjunction of the globular domain and the unstructured N-terminus. Conformer A (Figure 5, panel a), which has no NOE contacts between the random coil region and the second helix, shows a larger unstructured part before the beginning of the first helix. In conformer B (Figure 5, panel b), the additional three NOE contacts to the second helix limit the conformational freedom before the first helix (from Val6 to N-cap of the first helix). The very first amino acids (Ala2−Ala5) remain unstructured. The statistical data of the DNA binding domain of TraMM26 are summarized in Table 1.

Table 1: Structural Statistics for the TraMM26 (Residues 2−56)[a]

	conformer A	conformer B
Distance Restraints		
intraresidue	233	233
medium range	240	232
long range	81	84
φ angle restraints	42	42
Mean Values of Restraints Violations		
NOE violations > 0.3 (Å)	0	0
NOE violations > 0.1 (Å)	3.30	3.85
NOE distance violations (Å)	0.015	0.16
dihedral angle violations > 3 (Å)	0	0
dihedral angle violations > 1 (Å)	0.8	0.9
torsion angle violations (Å)	0.25	0.28
Ramachandran plot distribution		
residues in most favored region (%)	70.1	70.2
residues in allowed region (%)	22.9	22.7
residues in generously allowed region (%)	5.3	5.3
residues in disallowed region (%)	1.7	1.8
RMS Deviation from Idealized Geometry		
bonds (Å)	0.0015 ± 0.0002	0.0016 ± 0.0003
angles (Å)	0.467 ± 0.014	0.469 ± 0.017
improper (Å)	0.362 ± 0.009	0.364 ± 0.012
X-PLOR energy		
total (kcal/mol)	74.2 ± 7.1	76.7 ± 12.6
bond (kcal/mol)	2.1 ± 0.5	2.2 ± 0.7
angle (kcal/mol)	54.1 ± 3.3	54.5 ± 4.0
improper (kcal/mol)	8.8 ± 0.4	9.0 ± 0.6
vdw (kcal/mol)	2.6 ± 1.5	3.5 ± 2.9
noe (kcal/mol)	6.3 ± 2.3	7.3 ± 4.8
dihedral (kcal/mol)	0.2 ± 0.1	0.3 ± 0.3

RMS Deviation between the Two Conformers	residues 2−56	residues 10−53
backbone (Å)	6.302	0.722
heavy atoms (Å)	6.380	0.974

RMS Deviation from mean Structure	conformer A		conformer B	
	residue 2−56	residue 10−53	residue 2−56	residue 10−53
backbone (Å)	2.693	1.256	2.315	1.269
heavy atoms (Å)	3.303	1.800	2.876	1.858

[a] 20 structures/conformer.

DISCUSSION

The structure of the TraM mutant TraMM26 consists of three helical regions, from Asp[10] to Glu[26], from Ser[34] to Leu[45], and from Leu[47] to Gln[53]. A loop connects helix 1 and 2. The N-terminal end of the protein is unstructured. Helices 2 and 3 are directly connected by Gly[46], which changes the propagation direction. The core of the globular part of the protein, where the three helices contact each other, consists entirely of hydrophobic amino acids (Figure 3). [For comparison see Dill (*41*); Sharp (*42*)]. This hydrophobic interaction gives stability to the tertiary structure and holds the helices together. The same forces stabilize also the major number of the helix caps. Only the loop and the N-terminal end are not involved in these hydrophobic core interactions. The surface of the protein consists predominately of hydrophilic amino acids; however, at the interface between helix 2 and helix 3 close to the N-cap of the first helix, hydrophobic amino acids (Ile[12], Val[13], Leu[43], and Leu[47]) are partially exposed to the surface (Figure 3). Whether this is relevant for dimerization is not clear, but one could speculate that these amino acids form the pH-dependent dimerization area. From our structural model, monomer−monomer interactions in this region seem to be possible since a steric hindrance antagonizing dimer formation is not apparent.

The C-caps of the first and second helix are formed by Gly−Schellman motifs. This particular C-capping motif is often found in protein structures. It shows the carbonyl groups of the last amino acids of a helix involved in hydrogen bonds with the amide proton of the amino acid directly following the C-cap in the sequence. The steric freedom of Gly is particularly favorable for this kind of folding. At the end of helix 1 this is evidenced by a series of $d_{\alpha N}$ NOEs to Gly[27] and Ala[28], which directly follow the C-cap. Moreover, Gly[27] has a positive ϕ angle as expected for a Gly−Schellman motif. For helix 2, the same situation is corroborated by the NOE cross-peaks between the last helix turn and the following amino acids, and again Gly[46] shows the positive ϕ angle.

Unfortunately, the N-cap of the first helix cannot be exactly characterized because of severe spectral overlap of the side chain NOEs. Glu[11] shows medium range cross-peaks and the small $^3J_{HN-H\alpha}$ coupling constant typical for a helix; however, there is no clear evidence for such a cross-peak coming from Asp[10]. We suggest therefore that Asp[10] is the N-cap of helix 1, which forms H-bonds using its own side chain carbonyl group to the free HNs from the first helical turn. The N-cap of helix 2 is Ser[34], it is of the box Ia motif, and it is stabilized by hydrophobic contacts between Ile[38]

and Val[33] as well as Val[20]. The N-cap of the third helix is in direct conjunction with the C-cap of helix 2. In fact, Leu[45] is the C-cap of helix 2, whereas Leu[47] is the N-cap of helix 3. Between is the Gly[46] residue with higher conformational freedom and no side chain that could cause steric hindrance. This Gly[46] allows a strong deviation from the chain elongation direction, stabilizes the C-cap of helix 2 and the N-cap of helix 3, and allows the formation of the H-bonds and hydrophobic interactions mentioned earlier. The hydrophobic interactions between the side chains of Leu[42], Leu[43], Leu[45] with Leu[47], Val[49], and Val[13], as well as Ile[16], are interactions characteristic for N-cap and C-cap motifs, and part of the hydrophobic core that stabilizes not only the N- and C-cap, but the entire structure.

The N-terminal part of the TraMM26 is unstructured as was already stated. Nevertheless, this part together with the first helix plays an important role in the DNA binding capacity of the TraMM26 as well as the TraM wild type. It is the very first 22 amino acids that give specificity to the DNA binding activity (*2*). Many cases have been described in the literature in which flexible elements can interact, change conformation, and be of crucial importance for the DNA binding (*43*−*49*). Our previous investigations (*50*, *51*) showed that this 22-amino acid N-terminal peptide of the TraM sequence, displays no secondary structure in water. In trifluoroethanol, this peptide was helical as expected, and in SDS an amphiphilic helix beginning with Ser[9] was formed. Whereas in the case of TraMM26, two different conformations (see Figures 5, panels a and b) were found for this protein segment. To get a better understanding of this finding, further studies, for example, DNA binding studies will be carried out. At this stage, no conclusions about a possible "structure−mechanism relation" of this flexible part of the protein can be drawn. However, the orientation of the amino acids within helix 1 suggests that this helix could be an interaction interface. The end of this helix contains the sequence VERRRAE with Val[20] as a part of the hydrophobic core and Glu[26] as the C-cap. Two of the arginines are in positions where they could interact with the DNA backbone. In addition, two more positively charged amino acids in favorable positions can be found, e.g., Lys[15] and Lys[18]. Also Tyr and Asn residues are often involved in the specific recognition of the DNA. Here Tyr[14] and Asn[17] are present in a position allowing specific contacts to the bases in the DNA-binding site. The other parts of TraMM26 do not show such characteristics.

SUPPORTING INFORMATION AVAILABLE

The sequence specific assignment of TraMM26 is reported (Table 1). The chemical shifts of TraMM26 of conformer A and B are listed. Those amino acids that belong to conformer B are marked in bold. Figure 1a (conformer A) and Figure 1b (conformer B) show a sketch of the various ^{15}N-T_1 relaxation times. Figure 2 contains the assignment of the secondary structure based on the "chemical shift indexing". Negative bars symbolize α-helical parts, and positive bars symbolize *β*-sheets. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Pölzleitner, E., Zechner, E. L., Renner, W., Fratte, R., Jauk, B., Högenauer, G., and Koraimann, G. (1997) *Mol. Microbiol. 25*, 495−508.
2. Kupelwieser, G., Schwab, M., Högenauer, G., Koraimann, G., and Zechner, E. L. (1998) *J. Mol. Biol. 275*, 81−94.
3. Disque-Kochem, C., and Dreiseikelmann, B. (1997) *J Bacteriol. 179*, 6133−6137.
4. Schwab, M., Reisenzein, H., and Högenauer, G. (1993) *Mol. Microbiol. 7*, 795−803.
5. Verdino, P., Keller, W., Strohmaier, H., Bischof, K., Lindner, H., and Koraimann, G. (1999) *J. Biol. Chem. 274*, 37421−37428.
6. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
7. Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., and Blanquet, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8247−8251.
8. Lathrop, B. K., Burack, W. R., Biltonen, R. L., and Rule, G. S. (1992) *Protein Expr. Purif. 3*, 512−517.
9. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
10. Rance, M. (1987) *J. Magn. Reson. 74*, 557−564.
11. Levitt, M. H., Freeman, R., and Frenkiel, T. (1982) *J. Magn. Reson. 47*, 328−330.
12. Sklenar, V., and Bax, A. (1987) *J. Magn. Reson. 71*, 379−383.
13. Glushka, J., and Cowburn, D. (1987) *J. Am. Chem. Soc. 109*, 7879−7881.
14. Clore, G. M., and Gronenborn, A. M., (1991) *Prog. NMR Spectrosc. 23*, 32−92.
15. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.
16. Zuiderweg, E. R. P., and Fesik, S. W. (1989) *Biochemistry 28*, 2387−2391.
17. Marion, D., Ikura, M., Tschudin, R., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *J. Magn. Reson. 85*, 255−270.
18. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
19. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. A 102*, 241−245.
20. Kraulis, P. J. (1989) *J. Magn. Reson. 84*, 627−633.
21. Kraulis, P. J., Domaille, P. J., Campbell-Burk, S. L., Van Aken, T., and Laue, E. D. (1994) *Biochemistry 33*, 3515−3531.
22. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
23. Brünger,A. T. (1992) X-PLOR, Version 3.1, A System for X-ray Crystallography and NMR, Yale University, New Haven, CT.
24. Nilges, M., Kuszewski, J., and Brünger, A. T. (1991) In *Computational Aspects of the Study of Biological Macromolecules by NMR* (Hoch, J. C., Ed.) New York, Plenum Press.
25. Kuszewski, J., Nilges, M., and Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33−56.
26. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
27. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
28. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) *Proteins 12*, 345−364.
29. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
30. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp 130, J. Wiley, New York.
31. Richardson, J. S., and Richardson, D. C. (1988) *Science 240*, 1648−1652.
32. Ziegler, G., and Sterk, H. (1992) *Magn. Reson. Chem. 30*, 387−92.
33. Schellmann, C. (1980) In *Protein Folding: The Alpha -L Conformation at the Ends of Helices.* (Jaenicke, R., Ed.) pp 53−61, New York, Elsevier/North-Holland.
34. Aurora, R., and Rose, G. D. (1998) *Protein Sci. 7*, 21−38.
35. Aurora, R., Srinivasan, R., and Rose, G. D. (1994) *Science 264*, 1126−1130.

36. Harper, E. T., and Rose, G. D. (1993) *Biochemistry 32*, 7605−7609.
37. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
38. Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000) *Trends Biol. Sci. 25*, 147−150.
39. Percel, A., Hollosi, M., Tusnady, G., and Fasman, G. D. (1991) *Protein Eng. 4*, 669−679.
40. Brooks, R. E., Bruccoleri, B. D., Olafson, D. J., States, D. J., Swaminathan, S., and Karpuls, M. (1983) *J. Comput. Chem. 4*, 187−217.
41. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
42. Sharp, K. A. (1991) *Curr. Biol. 1*, 171−179.
43. Bonvin, A. M., Vis, H., Breg, J. N., Brugering, M. J., Boelens, R., and Kaptein, R. (1994) *J. Mol. Biol. 236*, 328−341.
44. Fraenkel, E., and Pabo, C. O. (1998) *Nat. Struct. Biol. 5*, 692−697.
45. Fraenkel, E., Rould, M. A., Chambers, K. A., and Pabo, C. O. (1998) *J. Mol. Biol. 284*, 351−361.
46. Jin. C., Marsden, I., Chen., X., and Liao, X. (1999) *J. Mol. Biol. 289*, 683−690.
47. Mo, Y., Vaessen, B., Johnston, K., and Marmorstein, R. (1998) *Mol. Cell. 2*, 201−212.
48. Raumann, B. E., Rould, M. A., Pabo, C. O., and Sauer, R. T. (1994) *Nature 367*, 754−757.
49. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell 81*, 705−714.
50. Plugariu, C., Stockner, T., Moskau, D., Koraimann, G., Högenauer, G., and Sterk, H. (1998) *J. Pep. Res. 51*, 244−250.
51. Stockner, T., Plugariu, C., and Sterk, H. (1998) *Magn. Reson. Chem. 36*, 161−168. ‡PDB ID1DP3, RCSB ID RCSB010259.

BI002031C