Refined Structure, DNA Binding Studies, and Dynamics of the Bacteriophage Pf3 Encoded Single-Stranded DNA Binding Protein[†,‡]

Rutger H. A. Folmer,[§] Michael Nilges,[||] Christina H. M. Papavoine,[§] Ben J. M. Harmsen,[§] Ruud N. H. Konings,[§] and Cornelis W. Hilbers*[,§]

Nijmegen SON Research Center, Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld 6525 ED Nijmegen, The Netherlands, and European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, Germany

Received February 3, 1997; Revised Manuscript Received May 15, 1997[⊗]

ABSTRACT: The solution structure of the 18-kDa single-stranded DNA binding protein encoded by the filamentous *Pseudomonas* bacteriophage Pf3 has been refined using 40 ms ^{15}N- and ^{13}C-edited NOESY spectra and many homo- and heteronuclear *J*-couplings. The structures are highly precise, but some variation was found in the orientation of the β-hairpin denoted the DNA binding wing with respect to the core of the protein. Backbone dynamics of the protein was investigated in the presence and absence of DNA by measuring the R_1 and R_2 relaxation rates of the ^{15}N nuclei and the ^{15}N−^{1}H NOE. It was found that the DNA binding wing is much more flexible than the rest of the protein, but its mobility is largely arrested upon binding of the protein to d(A)$_6$. This confirms earlier hypotheses on the role of this hairpin in the function of the protein, as will be discussed. Furthermore, the complete DNA binding domain of the protein has been mapped by recording two-dimensional TOCSY spectra of the protein in the presence and absence of a small amount of spin-labeled oligonucleotide. The roles of specific residues in DNA binding were assessed by stoichiometric titration of d(A)$_6$, which indicated for instance that Phe43 forms base stacking interactions with the single-stranded DNA. Finally, all results were combined to form a set of experimental restraints, which were subsequently used in restrained molecular dynamics calculations aimed at building a model for the Pf3 nucleoprotein complex. Implying in addition some similarities to the well-studied M13 complex, a plausible model could be constructed that is in accordance with the experimental data.

The filamentous *Pseudomonas* phage Pf3 codes for a 78-residue protein that belongs to the classical family of single-stranded DNA (ssDNA)[1] binding protein (ssDBPs). Proteins in this class bind nonspecifically and stoichiometrically to ssDNA (and ssRNA) and are capble of destabilizing dsDNA. They are presumed to function by binding to transient ssDNA formed during replication, recombination, and repair. Many ssDBPs bind with positive cooperativity (Alberts & Frey, 1970), which allows them to saturate a stretch of single-stranded nucleic acid at relatively low protein concentration. This ability is thought to be necessary to protect the DNA from the action of nucleases and to hold the DNA in a conformation which facilitates the function of other replication, recombination, or repair enzymes (Alberts & Sternglanz, 1977).

Recently, three-dimensional structures of several ssDBPs have been reported: the Ff (=M13, fd, f1) gene V protein (Folkers et al., 1994; Skinner et al., 1994; Prompers et al., 1995), the T4 gene 32 protein (Shamoo et al., 1995), the Pf3 ssDBP (Folmer et al., 1995), and the nucleic acid binding domain of adenovirus (Ad) ssDBP (Tucker et al., 1994). M13 GVP and Pf3 ssDBP both form a five-stranded antiparallel β-sheet from which protrude two β-hairpins and a large loop, and both proteins occur in solution as homodimers. Notwithstanding the similar overall folds and the positions of the loops being strikingly similar, there is very limited sequence homology between these two proteins. Only when their secondary structures were elucidated was a reliable sequence alignment possible. This revealed that they have 16 identical residues, 7 of which are situated in the β-hairpin denoted the DNA binding loop (Folmer et al., 1994a). The M13 GVP is one of the best characterized ssDBPs and is considered a paradigm for the study of nonspecific protein−ssDNA interactions. DNA binding studies using fluorescence (Alma et al., 1983; Bulsink et al., 1985), CD (Kansy et al., 1986), neutron scattering (Gray et al., 1982), EM (Gray, 1989; Olah et al., 1995), NMR (Folkers et al., 1993), and mass spectrometry (Cheng et al., 1996) have provided much insight into the binding characteristics and the macroscopical structure of the nucleoprotein complex formed with phage DNA. So far, however, a high-resolution structure of the M13 protein−ssDNA complex could not be determined. Instead, reliable models were constructed that

[†] This work was supported by Stichting Scheikundig Onderzoek Nederland (SON) under the auspices of the Netherlands Organization of Scientific Research (NWO) and in part by a European Union Access to Large Scales Facilities grant (ERBCHGECT940062) to EMBL.

[‡] The coordinates of the energy-minimized average structure of the (uncomplexed) protein have been deposited with the Brookhaven Protein Data Bank, along with the related NMR constraints. Entry codes are 1PFS and R1PFSMR, respectively.

* Corresponding author: telephone, xx31-243652678; FAX, xx31-243652112; email, ceesh@sci.kun.nl.

[§] University of Nijmegen.

[||] EMBL Heidelberg.

[⊗] Abstract published in *Advance ACS Abstracts,* July 1, 1997.

[1] Abbreviations: CPMG, Carr−Purcell−Meiboom−Gill; EM, electron microscopy; DBW, DNA binding wing; dsDNA, double-stranded DNA; GVP, gene V protein; MD, molecular dynamics; NOE, nuclear Overhauser effect; nt, nucleotide; R_1, longitudinal relaxation rate; R_2, transversal relaxation rate; RMD, restrained molecular dynamics; rms, root mean square; SA, simulated annealing; ssDNA, single-stranded DNA; ssDBP, ssDNA-binding protein; TEMPO, 4-hydroxy-2,2,6,6-tetramethylpiperidine-1-oxyl; XNOE, ^{15}N−^{1}H heteronuclear NOE.

S0006-2960(97)00251-1 CCC: $14.00

are in accordance with all studies mentioned above (Folmer et al., 1994b; Guan et al., 1994, 1995).

Similarly extensive studies have been performed on the Ad ssDBP (Tucker et al., 1994, and references therein) and the T4 gene 32 protein [for a review, see Chase and Williams (1986)]. In contrast, much less data are available on the DNA binding properties of Pf3 ssDBP. Some parameters of the *in vivo* precursor complex were assessed by EM (Casadevall & Day, 1985), and CD studies have been reported on complexes formed with various polynucleotides (Powell & Gray, 1993, 1995). The latter indicate that Pf3 ssDBP can use different binding modes in binding to polynucleotides, as binding ratios of $n = 2$, $n = 3$, and $n = 4$ nucleotides per monomer were observed.

Here, we report the refined 3D solution structure of a Pf3 ssDBP mutant (Phe36 → His) as determined from short-mixing-time NOESYs and several homo- and heteronuclear *J*-couplings. This so-called "solubility mutant" was designed to reduce the tendency to aggregate observed for the wild-type protein (Folmer et al., 1994a). It was shown earlier that the wild-type and mutant proteins have identical tertiary structures (Folmer et al., 1994a). In addition, we describe the DNA binding domain of the protein, which was mapped using spin-labeled oligonucleotides. These results will be related to the dynamics data of the protein, which were obtained from backbone ^{15}N relaxation and heteronuclear NOE measurements, performed in the absence and presence of the oligonucleotide $d(A)_6$. Finally, from a combination of all data presented, and imposing some similarities to the M13 complex, we were able to build a plausible model for the Pf3 nucleoprotein complex.

MATERIALS AND METHODS

NMR Measurements. The ^{15}N-labeled and ^{13}C/^{15}N doubly labeled Pf3 ssDBP F36H were isolated as described previously (Folmer et al., 1994a, 1995). The concentration of the NMR samples was about 1.8 mM for the ^{15}N-labeled and 1.4 mM for the doubly labeled protein. Samples were unbuffered and contained only a few milimolar salt.

NMR experiments were performed at 500 and 600 MHz on a Varian Unity+ and Bruker AMX spectrometer, respectively, and were carried out at 300 K. Homonuclear NMR experiments and ^{15}N, ^{13}C, and ^{1}H resonance assignments have been described earlier (Folmer et al., 1994a, 1995). Backbone carbonyl resonances were assigned from 3D COCAH and CT-HNCO experiments (Dijkstra et al., 1994; Grzesiek & Bax, 1992), while shifts of the carbonyl carbons in the side chains of Asx and Glx were found in a 2D LRCC spectrum (Bax et al., 1992). All three experiments were performed at 500 MHz. Quantitative distance constraints were obtained from a 3D gradient-enhanced ^{15}N NOESY-HSQC spectrum, recorded at 600 MHz with a mixing time of 40 ms. In addition, a 40 ms 3D ^{13}C NOESY-HMQC (Ikura et al., 1990) was acquired at 600 MHz. Details on these spectra were reported earlier (Folmer et al., 1997).

The following homo- and heteronuclear *J*-couplings were measured at 500 MHz: $^3J_{\mathrm{HNH}\alpha}$ in both an HMQC-J (Kay & Bax, 1990) and 3D HNHA (Vuister & Bax, 1993), [see Folmer et al. (1995)], $^3J_{\mathrm{H}\alpha\mathrm{H}\beta}$ in a 3D HACAHB-COSY (Grzesiek et al., 1995), $^3J_{\mathrm{NH}\beta}$ in a 3D CT-HNHB (Archer et al., 1991), $^3J_{\mathrm{C}\gamma\mathrm{N}}$ in a 2D {^{15}N} spin−echo difference HSQC (Vuister et al., 1993), and $^3J_{\mathrm{C}\gamma\mathrm{C}'}$ in both a 2D {^{13}CO} spin−echo difference HSQC (Grzesiek et al., 1993) and a 2D LRCC (Bax et al., 1992). The latter experiment also allowed measurement of the $^3J_{\mathrm{C}\alpha\mathrm{C}\delta}$ couplings in (iso)leucines. Peak volumes in the two subspectra of the spin−echo experiments were integrated and divided to obtain values for the couplings, as described by Vuister et al. (1993). To compensate for systematic errors reducing the apparent coupling as outlined by these authors, calculated values were multiplied with 1.05 and 1.075, for $^3J_{\mathrm{CC}}$ and $^3J_{\mathrm{CN}}$, respectively, to obtain the final *J*-couplings. The LRCC experiment was recorded with 64 transients and evolution times of 22.0 (t_1, ^{13}C) × 77.6 (t_2) ms. Details on the other *J*-coupling experiments are reported elsewhere (Folmer et al., 1997).

Backbone ^{15}N relaxation experiments were carried out at 500 MHz and 25 °C, to measure the R_1 and R_2 rates and the heteronuclear NOE (XNOE). The pulse sequences are based on those described by Davie and Wagner (1994), but we implemented the following improvements: instead of dephasing the solvent resonance, it is returned to the +*z*-axis prior to acquisition (Stonehouse et al., 1994). During the relaxation delay of the R_2 experiment a 3.7 kHz CPMG-type sequence was used (Kay et al., 1992; Palmer et al., 1992). In particular, we used $\{[\Delta-180°_{\phi}(^{15}\mathrm{N})-\Delta]_3-180°_{x}(^{1}\mathrm{H})-[\Delta-180°_{\phi}(^{15}\mathrm{N})-\Delta]_6-180°_{-x}(^{1}\mathrm{H})-[\Delta-180°_{\phi}(^{15}\mathrm{N})-\Delta]_3\}_N$, where $\Delta = 500\ \mu s$ and $\phi = x, -x$. In our experience, it was necessary to apply two proton pulses with opposite phase in each CPMG cycle to prevent saturation of the H_2O resonance, which otherwise is quite efficient at longer relaxation delays (>200 ms). Similarly, we used $\{\delta-180°_{x}(^{1}\mathrm{H})-2\delta-180°_{-x}(^{1}\mathrm{H})-\delta\}_N$ in the R_1 experiments ($\delta =$ 2 ms). Here, the phase of the 90° pulse prior to the relaxation delay is alternated between *y* and −*y* to ensure that the N_z magnetization properly relaxes as $\exp(-R_1T)$ (Kay et al., 1989). Furthermore, a constant duty cycle (i.e., sample temperature) was enforced in the series of R_2 experiments by applying an appropriate number of ^{15}N pulses during the recycle delay.

Recycle delays in the R_1, R_2, and XNOE experiments were 1.7, 1.8, and 6.0 s, respectively, and they were all acquired with evolution times of 177 (t_1,^{15}N) × 117 (t_2) ms. To sample R_1, ten relaxation delays were recorded, $T = 16$, 48, 97, 202, 403, 702, 1000, 1403, 2000, and 2597 ms. An equal number of R_2 delays was sampled: 14, 27, 41, 68, 95, 123, 163, 231, 313, and 408 ms. All experiments were performed twice, under identical conditions: once for the Pf3 protein free in solution and once for a stoichiometer mixture of protein and the olignonucleotide $d(A)_6$.

Spectra were processed using the MNMR program (PRONTO Software Development and Distribution, Copenhagen, Denmark) and analyzed with the XEASY pacakge (Bartels et al., 1995), both running on a Silicon Graphics Indigo workstation.

Structure Generation and Restraints. Distance constraints were obtained from the 3D ^{15}N- and ^{13}C-edited NOESY spectra, recorded with 40 ms mixing times. Peak volumes were corrected for the HMQC building block in the 3D pulse sequences as described earlier (Folmer et al., 1995). Integrals were converted into distances by a plain R^{-6} relation, using the prerefined average structure for calibration. Upper distance restraint limits were set by lengthening all determined distances *r* by $r \times \max[0.15, (0.15 + (r - 2.6)0.08)]$, and lower limits were put at 0.85*r*.

Stereospecific assignments and angle restraints were obtained from a qualitative analysis of the various *J*-coupling spectra. The χ_1 angles of most residues were accessible

through two or more experiments so that the extraction of angle restraints was rather straightforward. The χ_2 angles of isoleucines were measured in the LRCC experiment (Bax et al., 1992). If the *J*-couplings were diagnostic of near purely staggered rotamer conformations, the angles were restrained with 30° upper and lower limits. Otherwise, 40° or 50° limits were used, depending on the apparent degree of rotamer averaging.

Structures were calculated using simulated annealing starting from conformations with random backbone torsion angles. Calculations were performed with the program X-PLOR, version 3.1 (Brünger, 1992), with an extension for floating chirality and with a modified version of the "parallhdg" geometric force field which now more closely matches the geometric parameters reported by Engh and Huber (1991). The protocol used contains several modifications with respect to the basic ideas published previously (Nilges et al., 1988, 1991). In particular, we used a reduced atom representation for nonbonded interactions during the high-temperature phase to improve efficiency. Furthermore, all prochiral centers for which no stereospecific assignments were available were treated with a floating chirality procedure (Weber et al., 1988; Holak et al., 1989). The entire protocol has been amply described elsewhere (Folmer et al., 1997).

The ensemble of structures calculated with this protocol will be noted ⟨SA⟩, and they were used for most of the analyses mentioned in the present paper. However, the force field used does not include electrostatic interactions, which are needed, e.g., for a proper evaluation of hydrogen bonds. Therefore, the SA structures were subjected to an additional RMD calculation: they were immersed in a 7 Å shell of TIP3P water molecules (Jorgensen et al., 1983), the electrostatic interactions were switched on, and the purely repulsive nonbonded interaction was replaced by the OPLS Lennard-Jones potential (Jorgensen & Tirado-Rives, 1988; LeMaster et al., 1988). This refinement step consisted of 14 ps of dynamics at 500 K, using 5 fs time steps and uniform masses of 100 amu, followed by a cooling phase of 8 ps during which time the temperature was brought to 100 K. The resulting "water-refined" structures are denoted ⟨WR⟩.

Analysis of Relaxation Data. The R_1, R_2, and XNOE data sets were processed using Lorentzian-to-Gaussian apodization in the acquisition dimension and a squared cosine in the indirect dimension. After zero filling, digital resolutions of 1.46 and 0.68 Hz/point were obtained in F_1 and F_2, respectively. We used peak volumes rather than intensities to determine the relaxation rates, which was possible as nearly all resonances are well separated in the [^{15}N,^{1}H] correlation spectrum. For partly overlapping peaks, the integration region was set to comprise only the nonoverlapping part. This does not influence the measured rates as the line shapes of the individual resonances are independent of the relaxation delay *T*. Uncertainties in the peak volumes due to (thermal) noise were assessed by integrating noise areas similar in size to the cross peaks. The R_1 and R_2 rates were obtained by nonlinear least squares fitting the experimental data to a two-parameter single-exponential function using the MATLAB package (MathWorks, Inc.) (Levenburg−Marquardt algorithm). Uncertainties in the rate constants were determined via Monte Carlo simulations using Modelfree 3.1 (Palmer et al., 1991). The XNOE was determined by dividing the peak volumes of the spectra recorded with and without saturation of the ^{1}H reservoir.

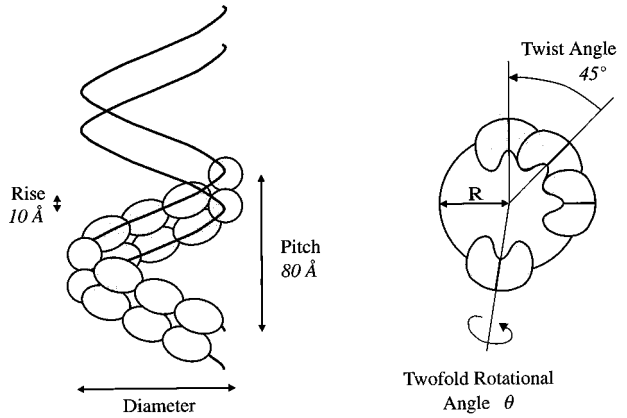

FIGURE 1: Schematic drawing of the helical and orientational parameters determining the superhelical nucleoprotein complex. The left-hand picture is a side view of the helix, built of two strands of ssDNA and eight protein dimer molecules. The right-hand picture provides a view along the helix axis. *R* is the distance from the helix axis to the center of mass of the protein molecule, and the angle θ measures the rotation around the protein's symmetry axis. From symmetry considerations, we implied this 2-fold axis to be perpendicular to and to intersect the helix axis (Folmer et al., 1994b). The values of the diameter, the twist angle, and helix pitch (printed in italics) have been proposed on the basis of the apparent similarity with the M13 nucleoprotein complex (see text).

The experimental relaxation parameters were interpreted in terms of the reduced spectral density mapping method first described by Farrow et al. (1995) and Ishima and Nagayama (1995a,b).

DNA Binding Studies. The spin-labeled trinucleotide, three adenyl residues to which a 4-hydroxy-2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) group is attached at both the 5′ and 3′ ends, was a gift from Prof. G. I. Tesser. We will refer to this oligonucleotides as *d(A)$_3$*, where the asterisks denote the TEMPO groups containing the unpaired electron. A binding experiment was performed by recording 30 ms homonuclear TOCSY spectra (Bax & Davis, 1985; Griesinger et al., 1988) before and after addition of a small amount of *d(A)$_3$* to a 1.6 mM solution of Pf3 F36H in D_2O. Spectra were acquired at 500 MHz with evolution times of 43 × 171 ms.

Furthermore, a stoichiometric titration series was carried out, adding d(A)$_6$ to a 1.2 mM protein solution in D_2O. Resonance shifts were followed by measuring homonuclear 2D NOESY (120 ms mixing time) and TOCSY (29 ms) spectra at each titration point. Experiments were performed at 600 MHz, with evolution times of 35 × 148 ms. At the end of the series, when the DNA/protein ratio was 3.6 mononucleotides/monomer, a 500 ms 2D NOESY was recorded.

Model Building of the Nucleoprotein Complex. The procedure followed to build a model for the Pf3 nucleoprotein complex was essentially the same as was described earlier for the complex of the M13 gene V protein and ssDNA (Folmer et al., 1994b). Here, the three-stage procedure will be outlined only briefly.

Preliminary results from electron microscopy indicate that the Pf3 nucleoprotein complex forms a superhelical structure similar to that of the M13 complex (K. R. Leonard, private communication). As shown in Figure 1, four parameters are necessary to describe this helix: a twist angle (related to the number of protein molecules per helix turn), a helical rise distance (related to the helix pitch), the distance *R* (related to the diameter), and the rotational angle θ.

Unfortunately, the EM pictures so far obtained do not allow reliable estimates of these parameters. Therefore, at this stage we simply assume a helix pitch of 80 Å and a twist angle of 45° (close to the M13 values) merely to investigate whether it is possible at all to construct a plausible model for the Pf3 complex. The sign of the twist angle determines the handedness of the helix which has not been established experimentally for the Pf3 complex. In principle, both possibilities could be investigated, but the current study is restricted to the left-handed conformation, which is the one exclusively observed in the M13 complex (Gray, 1989).

The first stage consisted of a 2D grid search varying R and θ to uncover the combination of these parameters that allows the energetically most favorable packing of the (rigid) molecules in the protein superhelix [see Folmer et al. (1994b)]. The grid search calculations were done using the energy-minimized average Pf3 structure, from which all atoms having a root mean square (rms) fluctuation in the ensemble of structures larger than 3 Å had been removed. In this way, the influence of the mobile regions of the molecule on the protein packing is reduced. R was varied from 20 to 50 Å, corresponding to an outer diameter of the complex of about 75−105 Å, with 2 Å intervals. The angle θ was scanned from 0 to 180° in steps of 4°.

In the second stage, two 14 nucleotides of arbitrary sequence were added to a protein superhelix built of three dimer molecules using the parameters found in the grid search. At the same time, the histidine residue at position 36 was replaced by a phenylalanine to restore the wild-type protein sequence. The resulting nucleoprotein complex was relaxed by restrained MD in a geometric force field, allowing the protein structure to adapt to the presence of the neighboring protein molecules and the ssDNA. The complete set of experimental restraints defining the protein structure was used, following removal of the NOEs connecting to residues which are believed to play an important role in DNA binding *(vide infra)*. In addition, the DNA was restrained to follow approximately the binding path outlined by the spin-label experiment described above. This was achieved by including distance restraints from the broadened residues to the DNA phosphorus atoms (Fomer et al., 1994b). We used the following empirical criteria to classify the spin-label restraints: if a TOCSY cross peak was broadened by 80−100%, the effective distance between the protons responsible for the cross peak and the nearest phosphorus atoms was assumed to be smaller than 6 Å. Likewise, 0−8.5 and 6.5−11 Å distances were implied for resonances broadened by 40−80% and 20−40%, respectively. If no broadening was observed, the distance was set to be larger than 10 Å. It is noted that these restraints are ambiguous in two ways: the individual contributions of the protons constituting the broadened cross peak are unknown, and it is unclear which is the nearest phosphorus atom. Therefore, the effective distance can only be calculated properly with the method known as "sum averaging" in X-PLOR ($R_{ij} = \{\sum_{i,j} r_{ij}^{-6}\}^{-1/6}$) (Nilges, 1993).

Because we used a force field in which the van der Waals interactions are represented by a repulsive potential, and in which no electrostatic interactions are incorporated, there are no attractive forces between the protein molecules in the superhelix. To prevent the molecules from being pushed away from each other as a result of possible steric hindrance, a parabolic restraint was added, keeping the well-determined atoms (rmsd <0.4 Å) near their original positions (Bruccoleri & Karplus, 1986).

The third stage consisted of a 30 ps restrained MD run at 500 K with explicit inclusion of solvent. The electrostatics were incorporated, and the OPLS force field was used for the nonbonded interactions (Jorgensen & Tirado-Rives, 1988; LeMaster et al., 1988). The parameters for the DNA were always from the AMBER force field (Weiner et al., 1984). Now, the force field does contain attractive components, so that the harmonic restraints could be released. Thus, the molecules are allowed to move or rotate in order to minimize the intermolecular nonbonded energy, hence optimizing the protein−protein interface.

RESULTS

Restraints for Structure Calculation. The prerefined structures reported earlier (Folmer et al., 1995) were calculated from 75 ms ^{13}C- and ^{15}N-edited 3D NOESY spectra, which were completely assigned. Assignment of the 40 ms equivalents used for the current refinement was therefore trivial. In total, 1145 independent NOE restraints were obtained, 543 of which are intraresidue, 210 are sequential, 74 are medium range (two to five residues apart in the sequence), and 318 are long range (more than five residues apart). Figure 2a shows how the restraints are distributed over the 78 residues of the protein. Out of these 1145 NOEs, 1018 were assigned to be intramonomer, while 22 intermonomer NOEs were counted. The remaining 105 distance restraints connect to residues close to the symmetry axis of the molecule. These NOEs will have both intra- and intermonomer contributions and hence were treated with the ambiguous NOE approach of Nilges (1993).

Aside from the 33 ϕ angles restrained earlier, several J-couplings could be measured by assessing the χ_1 and χ_2 dihedral angles. This analysis yielded stereospecific assignment and approximate χ_1 angles for all valines but Val11, whose J-couplings suggest the presence of rotamer averaging. In addition, χ_1 angles could be measured for all four isoleucines in the protein and for Thr8 and Thr23, while rotamer averaging was observed for Thr6 and Thr15. Stereospecific assignment of the methylene groups was rather straightforward using the HNHB and HACAHB-COSY spectra, which allowed stereospecific assignment and measurement of χ_1 angles for 30 of the 43 β-methylene groups that give rise to resolved ^{1}H NMR signals. The remaining 13 prochiral β-pairs could not be unambiguously assigned due to spectral overlap (five residues) or because the J-couplings suggest some level of rotamer averaging (eight residues).

χ_2 restraints could be measured for all four isoleucines in the LRCC spectrum, which was possible because all $^3J_{C\alpha C\delta}$ couplings corresponded to *trans* conformations (apparent coupling ∼3.5 Hz). This experiment furthermore showed that rotamer averaging is present in χ_2 of Leu29 and that the low-field shifted Cδ of Leu68 is *trans* to the Cα. Resonance overlap prohibited analysis of Leu31 and Leu38 in this spectrum.

Structures and Statistics. A total of 1145 NOE restraints and 79 dihedral restraints were used per monomer to calculate 80 (dimer) structures. The 30 structures with the lowest total energies were selected, and this ensemble is denoted ⟨SA⟩. None of these structures had NOE violations larger than 0.3 Å or dihedral violations larger than 2°. The structures were

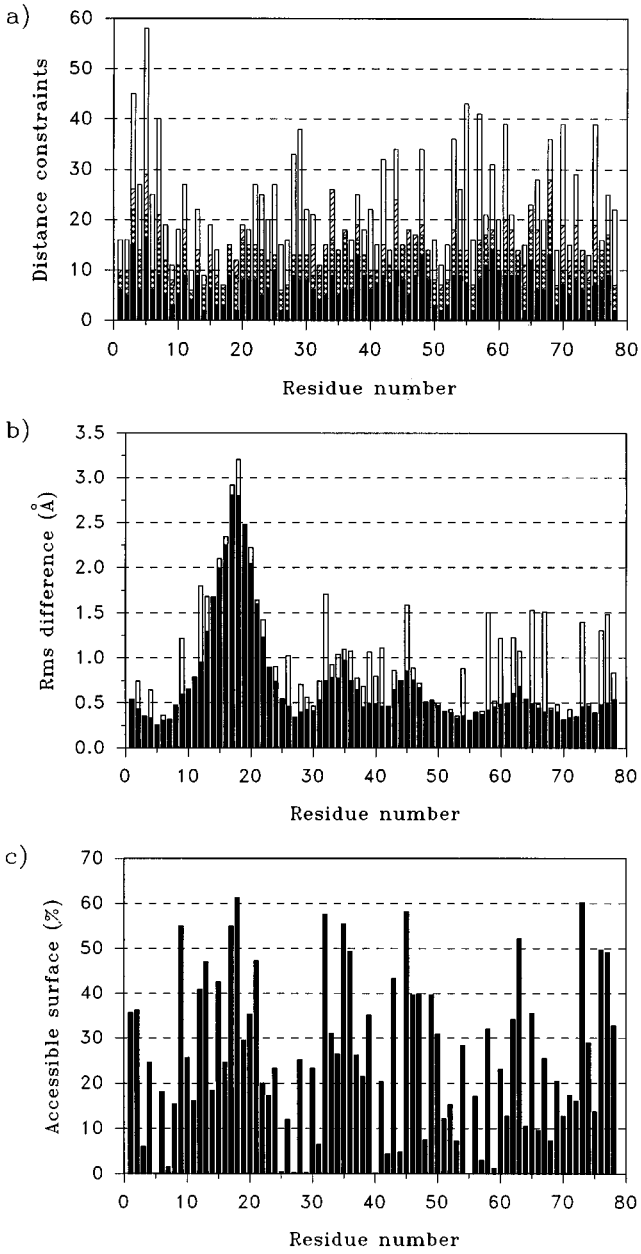

FIGURE 2: (a) Histogram showing the number of NOE-derived distance constraints per residue. The bars represent intraresidue (filled), sequential (cross-hatched), medium-range (hatched), and long-range (open) NOEs. The number of intraresidue restraints reflects independent restraints, whereas the sequential, medium-range ($2 \leq |i - j| \leq 4$), and long-range ($|i - j| \geq 5$) restraints were countered for both residues involved. (b) Plot of the residue-based root mean square deviation of $\langle SA \rangle$ from $\langle \overline{SA} \rangle$ for the backbone (N, C, Cα; solid bars) and all non-hydrogen atoms (ppen bars). (c) Histogram showing the relative solvent accessibility per residue of the dimeric P3 F36H energy-minimized average structure. Numbers were calculated with the program MOLMOL (Koradi et al., 1996).

superimposed for the backbone atoms (N, Cα, C) of residues 1−11 and 25−78, and the coordinates were averaged to create the average structure, $\langle \overline{SA} \rangle$. This was energy minimized through 1500 steps of conjugate gradient minimization to give the energy-minimized average structure, $\langle \overline{SA} \rangle_r$, which is schematically shown in Figure 3. Each monomer consists of seven β-strands (β1−β7) of which β4, β3, β1, β5, and β7 form a five-stranded antiparallel sheet across the midsection of the molecule. From this core protrude two prominent β-hairpins and a large loop. The hairpins are composed of residues 11−25 and 57−70 and are designated

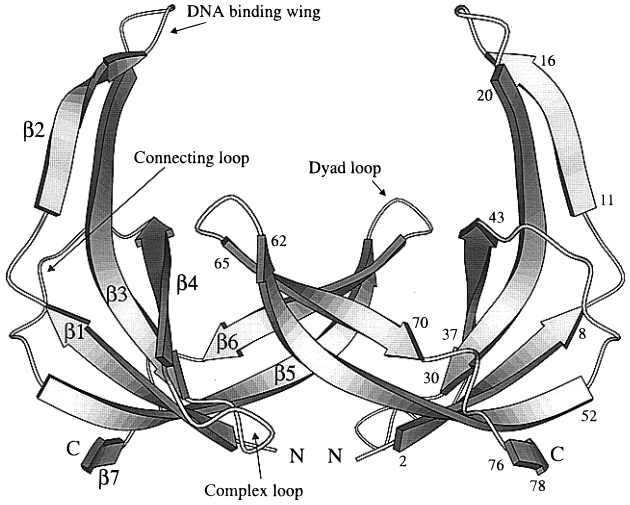

FIGURE 3: Schematic drawing of $\langle \overline{SA} \rangle_r$. The four major loops are denoted on the left-handed monomer. The seven β-strands are indicated with their strand numbers and first and last residue numbers. The molecule was drawn with MOLSCRIPT (Kraulis, 1991).

the "DNA binding loop" and "dyad loop", respectively (Folkers et al., 1991b; Folmer et al., 1994a). The large loop consists of residues 31−36 and is referred to as the "complex loop". In addition, there is a loop connecting strands β4 and β5, which is rather tightly bound to the core of the protein and which will be denoted the "connecting loop". A detailed description of the structure was given earlier (Folmer et al., 1995).

Figure 2b affords rms deviations per residue with respect to the mean structure. Note that the high values around residues 12−24 (the DNA binding wing) are partly due to the fact that these residues were not included in overlaying the ensemble. The best-fit superposition of the family of 30 structures is shown in Figure 4. A solvent-accessible surface of 25% (see Figure 2c) was used as a threshold to classify solvent-exposed and buried residues, which are separately shown in red and blue, respectively. The rms deviation about the mean coordinate positions excluding residues 12−24 is 0.50 ± 0.17 Å for the backbone atoms and 0.86 ± 0.12 Å for all heavy atoms. If one furthermore excludes the side chains of the solvent-exposed residues, the latter number drops to 0.57 ± 0.14 Å, which indicates that the interior of the protein is well-defined. The Lennard-Jones energy, calculated with the CHARMm force field, is large and negative, indicating that the van der Waals packing is favorable. A summary of the statistics of the ensemble of structures and their average is given in Table 1. Figure 5 illustrates the excellent stereochemical quality of the ensemble of structures, as basically 100% of the backbone dihedral angles fall within the allowed regions of the Ramachandran map. The ϕ angle of one residue (Phe70) in one structure (number 16 if ordered according to total energy) apparently got trapped in a positive conformation (+59°), while this angle averages to −58.3 ± 4.4° in the other 29 molecules. According to the PROCHECK program (Laskowski et al., 1993), version 3.4.4, 82.8% of the non-glycine backbone dihedral angles fall within the most favored regions of the map. Angular order parameters (S) (Hyberts et al., 1992) are lower than 0.95 only for five ϕ and three ψ angles (lowest values: $S_\phi = 0.86$ for Leu38 and $S_\psi = 0.88$ for Pro37), indicating that the structures are highly converged.

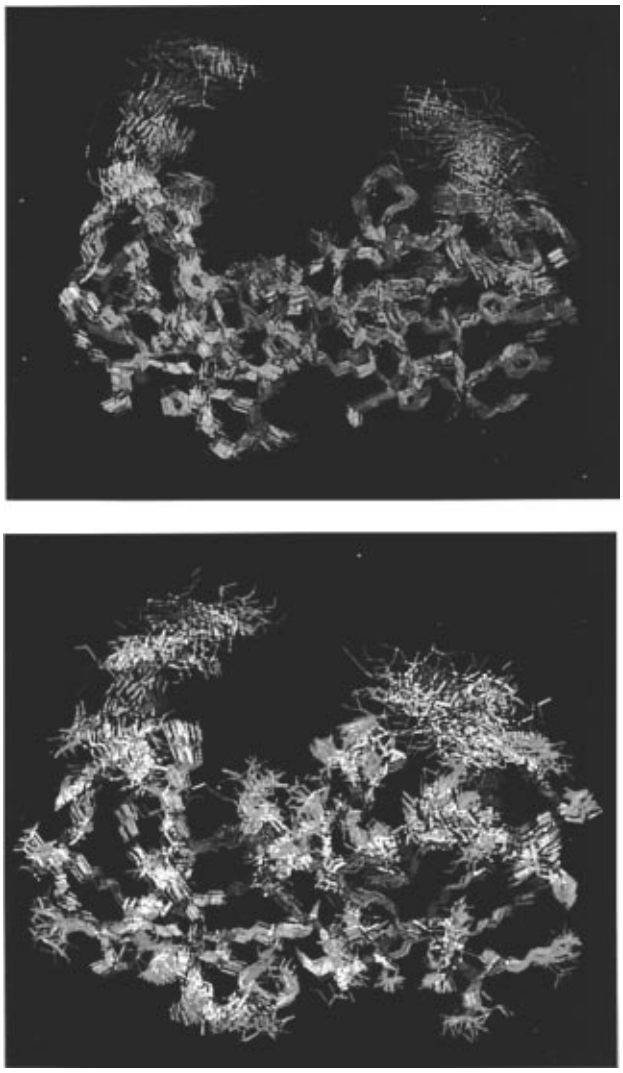

FIGURE 4: Best-fit superposition of the ensemble of 30 calculated structures. Coordinates were superimposed for the backbone atoms, excluding those of the DNA binding loop (residues 12−24). The orientation of the molecules differs from that of Figure 3 by a rotation of 30° around the horizontal axis. (a, top panel) Side chains are shown in blue of residues whose surfaces are less than 25% solvent accessible (see Figure 2c). (b, bottom panel) Solvent-exposed residues are shown in red (>25% accessible surface; see also Table 1). The picture was prepared with MOLMOL (Koradi et al., 1996).

This high precision and presumed concomitant accuracy should allow a reliable classification of the various β-turns in the protein. A classical analysis was performed as first described by Venkatachalam (1968), the results of which are compiled in Table 2. The amino acids connecting strands β1 and β2 constitute a β-bulge rather than a turn. The carbonyl oxygen of Glu26 is hydrogen bonded to the amide proton of Thr8 (average energy -1.9 ± 0.2 kcal/mol), while the amide proton of Glu26 is very stable hydrogen bonded to the carbonyl oxygen of Ser10 (-3.0 ± 0.2 kcal/mol). According to the classifications of Chan et al. (1993), the three residues opposite to Glu26 form an S3 bulge.

Backbone Dynamics. The backbone nitrogen R_1, R_2, and XNOE parameters were measured for the Pf3 protein in the absence and presence of a stoichiometric amount of the oligonucleotide $d(A)_6$. This hexanucleotide was added to a ratio of one molecule per protein dimer, i.e., three nucleotides per monomer. Below, the subscripts "free" and "complex" will be used to denote parameters of the free protein and the protein in complex with $d(A)_6$, respectively. Quantitative data could be obtained for all non-proline residues (except Met1) in the free protein, but resonance overlap in the ^{1}H−

Table 1: Structural Statistics

	$\langle SA \rangle^a$	$\langle \overline{SA} \rangle_r$
X-PLOR energies (kcal mol^{-1})		
E_{tot}	193 ± 14	159
E_{bond}	9.2 ± 1.0	7.5
E_{angle}	75.1 ± 4.8	67.1
E_{impr}	15.2 ± 0.9	12.2
E_{vdW}	14.7 ± 2.9	10.7
E_{NOE}	79.2 ± 9.2	61.2
E_{dihe}	0.1 ± 0.1	0.0
E_{NCS}[b]	(5.4 ± 3.9) × 10^{-3}	1.5 × 10^{-6}
$E_{L\text{-}J}$[c]	−459 ± 49	−163
rms deviations from ideal geometry		
bonds (Å)	0.0020 ± 0.0001	0.0017
angles (deg)	0.33 ± 0.01	0.31
impropers (deg)	0.36 ± 0.007	0.35
rms deviations from exptl restraints		
NOEs (Å)	0.027 ± 0.002	0.023
dihedrals (deg)	0.067 ± 0.037	0.031

atomic rms differences	backbone	non-hydrogen
$\langle SA \rangle$ vs $\langle \overline{SA} \rangle$	0.81 ± 0.36	1.07 ± 0.28
$\langle SA \rangle$ vs $\langle \overline{SA} \rangle$ not residues 12−24	0.50 ± 0.17	0.86 ± 0.12
$\langle SA \rangle$ vs $\langle \overline{SA} \rangle$ excluding solvent-exposed residues[d]	0.49 ± 0.17	0.57 ± 0.14
$\langle \overline{SA} \rangle$ vs $\langle \overline{SA} \rangle_r$	0.12	0.40

[a] $\langle SA \rangle$ is the ensemble of the 30 final structures. $\langle \overline{SA} \rangle$ is the Cartesian coordinates obtained by averaging $\langle SA \rangle$ following a least squares superposition of the backbone atoms for residues 1−11 and 25−78. $\langle \overline{SA} \rangle_r$ is the energy-minimized averaged structure. [b] Typically, the rms difference between the two monomer (all residues, all atoms) is smaller than 0.002 Å in the final structures. [c] Lennard-Jones energies were calcualted with QUANTA, release 4.0, using the CHARMm22 force field. [d] Sice chains (all but Cβ) of residues whose solvent-accessible surface was larger than 25% were excluded (see Figure 2c). These were M1, N2, Q4, D9, S10, R12, Q13, T15, S16, K18, G19, N20, P21, F28, E32, D33, K34, P35, H36, P37, N39, F43, E45, S46, V47, P49, Q54, R58, N62, N63, R65, E67, K73, K76, and R77.

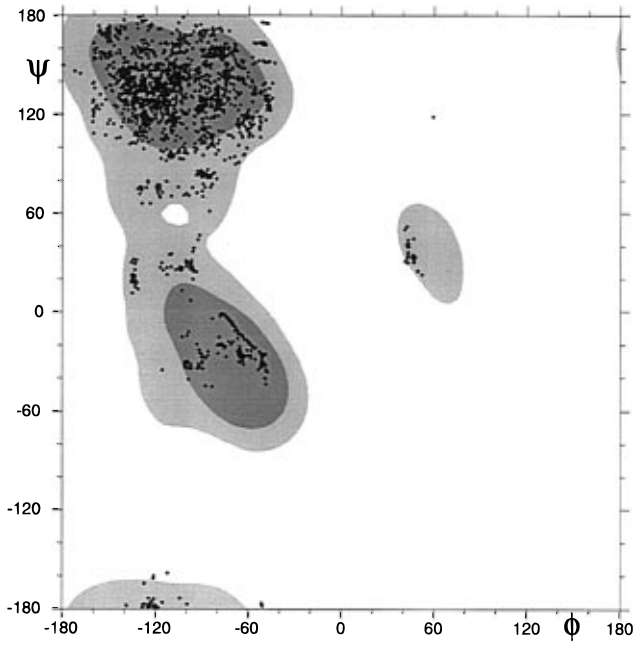

FIGURE 5: Ramachandran plot of $\langle SA \rangle$ showing all non-glycine residues. There is basically only one ϕ/ψ pair out of 2160 in a disallowed region, which is that of Phe70 in one of the 30 structures. The plot was generated using MOLMOL (Koradi et al., 1996).

Table 2: β-Turns in Pf3 ssDBP F36H

residues	hydrogen bond[a]	energy[b]		ϕ	ψ	class[c]
16−19	Gly19 HN−Ser16 O	−2.21 ± 0.32	Ala17	−52.7 ± 5.3	−35.9 ± 7.9	type I
			Lys18	−80.1 ± 5.1	−0.5 ± 10.5	
31−34	Lys34 NH−Leu31 O	−1.32 ± 0.46	Glu32	−61.7 ± 1.0	−30.5 ± 10.9	type I
			Asp33	−99.2 ± 9.2	9.3 ± 9.3	
62−65	Asn62 NH−Arg65 O	−2.72 ± 0.39	Asn63	51.8 ± 5.8	34.4 ± 8.4	type I′
			Gly64	84.4 ± 6.5	5.7 ± 11.5	
71−74	Ala74 HN−Asp71 O$^{\delta}$	−3.13 ± 0.98	Phe72	−78.3 ± 4.7	−1.0 ± 9.6	type I (?)
			Lys73	−90.7 ± 10.0	−40.5 ± 6.1	
71−74[d]	Ala74 HN−Asp71 O	−1.80 ± 0.48	Phe72	−64.1 ± 3.3	−25.5 ± 2.6	type I
			Lys73	−73.6 ± 3.8	−6.2 ± 8.1	

[a] Hydrogen bonds were evaluated in the structures refined in water; ⟨WR⟩. No distance restraints were used to enforce the formation of hydrogen bonds during any of the structure calculations. [b] In kcal·mol^{-1}, calculated with the program DSSP (Kabsch & Sander, 1983). [c] Turns were classified as described by Richardson (1981). [d] These numbers refer to the structures refined in the geometric force field; ⟨SA⟩. They are included for this turn because in ⟨WR⟩ the backbone angles of residues 72 and 73 deviate rather significantly from the predicted values of the common turns (Venkatachalam, 1968). This is likely to be related to the amide proton of Ala74 uncommonly being hydrogen bonded to a side-chain oxygen of Asp71. In contrast, the more classical hydrogen bond to the backbone carbonyl of Asp71 is formed in ⟨SA⟩, and here the ϕ and ψ angles are actually quite close to the expected values for a type I turn.

^{15}N correlation spectra of the complex prohibited analysis of Asp9, Thr15, Leu38, and Asp71. The signal to noise ratio in the R_1 and R_2 correlation spectra was excellent; peak volumes in the spectra recorded with the shortest T delays were 50−100 times larger than the standard deviation calculated from an randomly selected set of noise−area integrals. These standard deviations were determined for both experiments and both samples (with and without DNA). The R_1 and R_2 rate constants were obtained from a nonlinear least squares fitting procedure. Here, the noise−area standard deviations were used as uncertainties in the peak volumes. $R_{1,\mathrm{free}}$ values range from 1.33 (Ser46) to 1.69 (Phe7) s^{-1}, with typical numbers around 1.6 s^{-1} for core residues. Very similar rate constants were measured for the protin−d(A)$_6$ complex. $R_{2,\mathrm{free}}$ constants vary between 6.6 (Asn20) and 13.2 (Glu26) s^{-1}, with values around 11−12 s^{-1} for most core residues. R_2 values measured for the protein complexed to the oligonucleotide are about 2 s^{-1} larger for residues in the interior of the protein and up to 3 s^{-1} larger in the DNA binding loop. Subsequent Monte Carlo simulations yielded the following uncertainties in these relaxation rates, averaged over all residues analyzed: 3.0 ± 0.8% for $R_{1,\mathrm{free}}$, 1.9 ± 0.5% for $R_{1,\mathrm{complex}}$, 2.0 ± 0.6% for $R_{2,\mathrm{free}}$, and 3.5 ± 0.9% for $R_{2,\mathrm{complex}}$. The values of the heteronuclear NOE in the free protein were about 0.7−0.8 for core residues while values below 0.5 were found for residues 14−22 (the DNA binding loop; 0.07 for Asn20) and for His36 (0.43) and Ala78 (0.45). Numbers for core residues increased only slightly upon DNA binding, but those of residues in the DNA binding loop increased by up to 0.35. Notably residues 15−20, whose XNOEs were 0.07−0.30 in the free protein, all have values above 0.4 when the protein is bound to d(A)$_6$. Uncertainties in the heteronuclear NOE were 5.7 ± 2.6% for XNOE$_{\mathrm{free}}$ and 5.5 ± 0.8% for XNOE$_{\mathrm{complex}}$. Values of R_1, R_2, and XNOE were subsequently used to calculate spectral densities, employing the reduced spectral density mapping method (Farrow et al., 1995; Ishima & Nagayama, 1995a). Figure 6 shows the spectral densities, which were calculated at frequencies zero, ω_{N} and ω_{H} (50.7 and 500 MHz, respectively).

The mean value of R_2/R_1 was used to assess the overall rotational correlation time (τ_{m}) (Kay et al., 1989). In a first step, the residues with XNOE <0.6 were excluded from this procedure, as their R_1 and R_2 relaxation rates are likely to contain significant contributions from fast internal motions. Then, an average R_2/R_1 was calculated, and the resulting standard deviation (σ) was used to further exclude residues whose R_2/R_1 deviated more than 2σ from the mean (only Glu26 and Phe43), as these are probably influenced by exchange processes. The average ratios calculated from the remaining residues were $(R_2/R_1)_{\mathrm{free}} = 7.37 \pm 0.59$ and $(R_2/R_1)_{\mathrm{complex}} = 7.94 \pm 0.76$, from which correlation times were calculated of 9.7 and 10.2 ns, respectively. These numbers are somewhat lower than what was expected on the basis of the molecular weights, which are 17.8 kDa for the free protein and ∼20 kDa for the complex (assuming an average binding of three nucleotides per monomer). Previously reported values for systems of similar size are, e.g., 13 ns at 27 °C for 19-kDa human granulocyte colony-stimulating factor (Zink et al., 1994) and 12.7 ns at 30 °C for the 18-kDa complex of *Escherichia coli* topoisomerase I (Yu et al., 1996). The low number of the uncomplexed Pf3 protein might be explained by a high relative mobility of the two monomers in the dimer molecule. Also, it emphasizes the success of the substitution of Phe36 for His and accounts for the high quality of the NMR spectra that could be obtained. The fact that the correlation time increased by only 0.5 ns upon DNA binding indicates that the protein−d(A)$_6$ complex is a relatively mobile system, which in turn is in accordance with the fast exchange observed in binding experiments *(vide infra)*.

DNA Binding Domain. Previously, we have demonstrated that the use of spin-labeled oligonucleotides is a powerful method to map DNA binding domains of proteins that bind to DNA in fast exchange (Van Duynhoven et al., 1993; Folkers et al., 1993). The first step in this study was to determine the proper amount of *d(A)$_3$* that has to be added to the Pf3 protein so that an optimal dynamic range is obtained with respect to the degree of broadening of the respective protein resonances. This is dependent on the rate of exchange, the specificity of binding, and the efficiency of dipolar relaxation of the paramagnetic moiety. We established that one *d(A)$_3$* molecule per 34 protein monomers was sufficient to completely broaden the most sensitive TOCSY cross peak, which is the Hα−Hβ of Ala69. The addition of more spin-label led to the complete broadening of increasingly more residues which prohibited differentiation between the most sensitive residues.

A regular homonuclear TOCSY spectrum was recorded on the free protein, and a second spectrum was recorded after addition of $^1/_{34}$ molar equiv of *d(A)$_3$*. To quantify the broadening effects, relative numbers were calculated for each

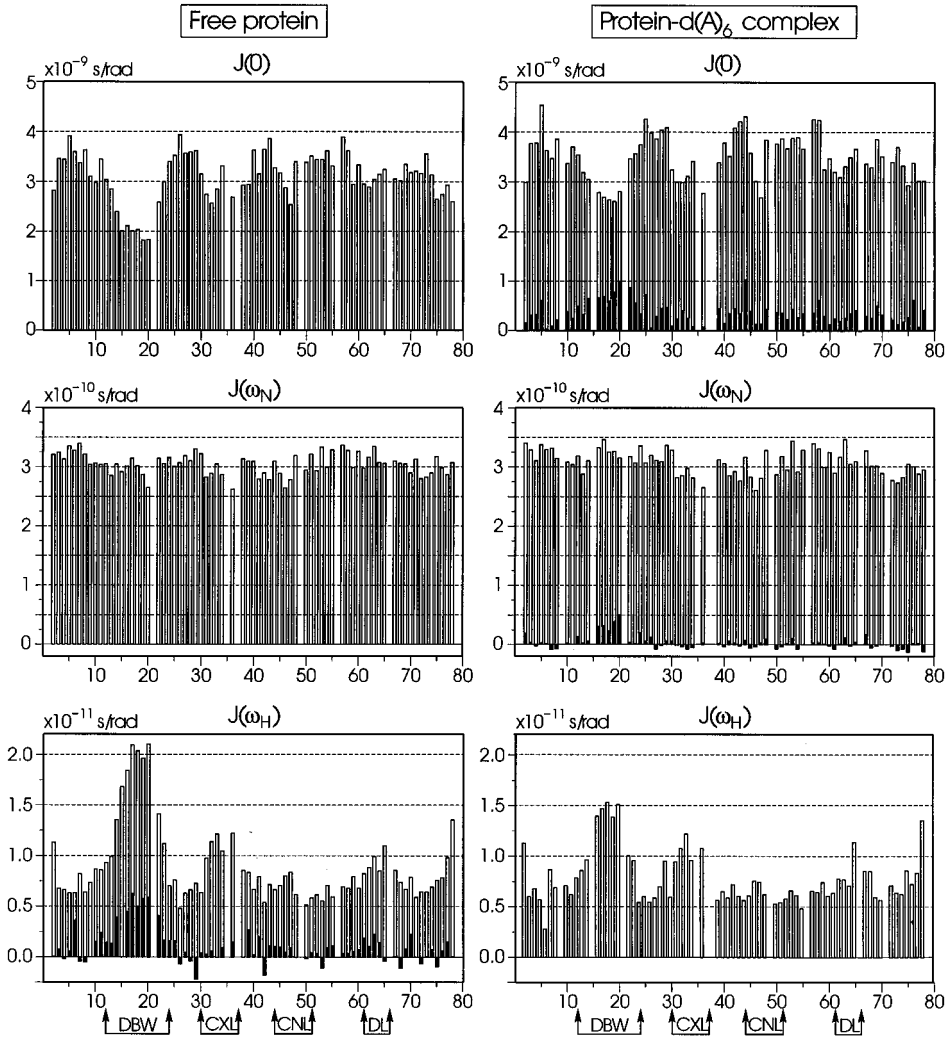

FIGURE 6: Histograms showing the spectral densities (open bars) at three frequencies for the free protein (left) and the protein−$d(A)_6$ complex (right), plotted against the residue number. Solid bars show the difference between free and bound protein. For $J(0)$ and $J(\omega_N)$ these are plotted in the right-hand graph; the spectral densities at these frequencies generally increased upon DNA binding, and the bars show $J_{complex} - J_{free}$. Spectral densities at frequency ω_H mostly decreased, and the solid bars show $J_{free} - J_{complex}$. The four major loops, designated in Figure 3, are indicated below the graphs: the DNA binding wing (DBW), the complex loop (CXL), the connecting loop (CNL), and the dyad loop (DL).

Table 3: Relative Dipolar Broadening Effects and Classification of Protein−ssDNA Distance Constraints

broadening (%)	most affected cross peak (per residue)[a]	distance restraint (Å)
80−100	R12$\beta\delta$, S16$\beta\beta$, Y22$\alpha\beta$, Y22$\delta\epsilon$, F43$\alpha\beta$, F43$\delta\zeta$, A69$\alpha\beta$[b]	0−6
40−80	G14$\alpha\alpha$, A17$\alpha\beta$, N20$\alpha\beta$, P21$\alpha\beta$, F24$\alpha\beta$, F42$\alpha\beta$, E45$\alpha\gamma$, N63$\alpha\beta$, G64$\alpha\alpha$, R65$\alpha\delta$, P66$\gamma\delta$	0−8.5
20−40	T15$\alpha\gamma$, K18$\alpha\epsilon$, T23$\alpha\gamma$, E26$\alpha\gamma$, F42$\delta\epsilon$, V44$\alpha\beta$, N60$\alpha\beta$, N62$\alpha\beta$, L68$\beta\delta$	6.5−11

[a] The aromatic residues Tyr22, Phe42, and Phe43 are listed for both the aliphatic and aromatic resonances. Spectra overlap prohibited accurate measurements of the aromatic cross peaks in Phe24, but it is clear these are affected as well. [b] The Ala69$\alpha\beta$ cross peak was broadened by 95%, but initial calculations on the protein−DNA complex showed that this residue cannot be as close to the DNA as this number suggests. This particular distance restraint was therefore released (see text).

cross peak by dividing their integrals in the two spectra:

$$\text{relative broadening} = [1 - f(V_{presence}/V_{absence})] \times 100\%$$

where $V_{presence}$ and $V_{absence}$ are the peak volumes in the presence and absence, respectively, of the spin-labeled oligonucleotide and f is a correction factor. This factor is used to correct for systematic deviations in cross peak intensities caused by (1) sample dilution upon oligonucleotide addition, (2) nonspecific resonance broadening induced by unbound spin-labeled DNA, and (3) increased transversal relaxation due to the larger rotation correlation time of the protein−oligonucleotide complex. We used $f = 1.06$, a number easily calibrated by "nilling" the effect on residues that are not perturbed by the spin-label (i.e., the majority of cross peaks). Table 3 lists these relative broadening effects, and Figure 7 shows where the affecte residues are located in the three-dimensional structure of the protein. Basically these residues are situated on the DNA binding wing (residues 10, 12−24, 26), strand β4 (residues 40, 42−45), and the dyad loop (residues 60, 62−69).

The stoichiometric titration series with the oligonucleotide $d(A)_6$ proved useful in two respects. First, it allowed an estimate of the binding ratio, and second, it provided some insight into the role of specific amino acids in nucleic acid binding. Figure 8 displays the shifts in resonance positions of some characteristic protons. In particular, the ring protons

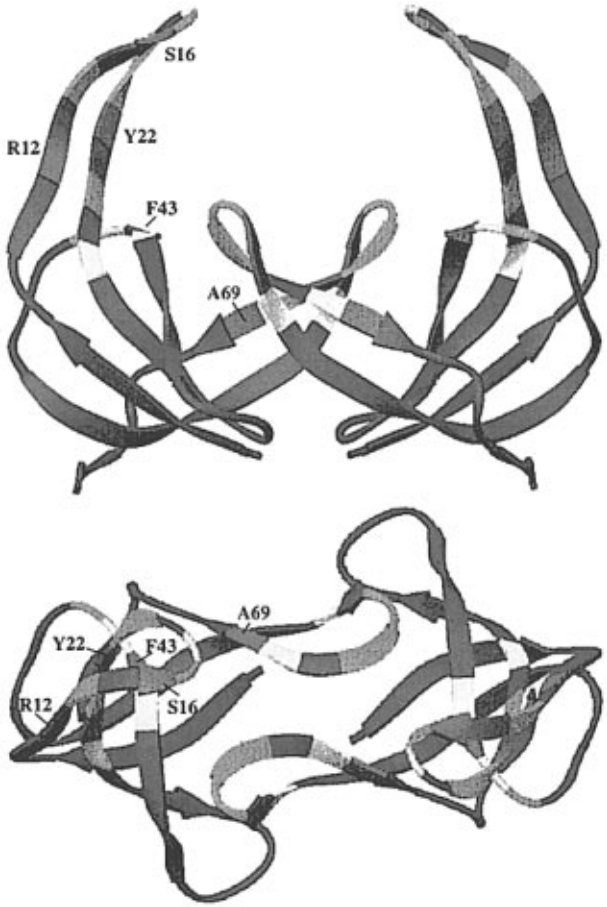

FIGURE 7: Ribbon plot of the Pf3 protein in two orientations, color-coded with respect to the degree of broadening by the spin label: blue, no broadening (<20%); yellow, 20−40%; orange, 40−80%; red, 80−100%. The plot was made with RIBBONS (Carson, 1991).

of Phe43 shift significantly upon $d(A)_6$ addition, which suggests that this aromatic ring stacks or (partly) intercalates with the bases of the DNA. Chemical shift differences on Tyr22 and Phe24, the other aromatic residues implicated in binding, are much less pronounced. The 500 ms 2D NOESY recorded at the end of the titration series shows further evidence for the proposed interactions of Phe43 as its ring protons make clear NOE contacts to many ring protons of the oligonucleotide (data not shown). In addition, NOE effects occurred between all ring protons of Phe43 on the one hand and many of the H1′ sugar protons on the other. The aromatic protons of Tyr22 and Phe24 show contacts very similar to those of Phe43, suggesting that these three residues are about equally involved in binding. So far, the only DNA−protein NOE contacts assigned in this NOESY spectrum are between H1′, H2, and H8 and the ring protons of Tyr and Phe. There is no indication of aromatic residues other than Tyr22, Phe24, and Phe43 involved in DNA binding.

Modeling of the Pf3 ssDBP Nucleoprotein Complex. The result of the two-dimensional grid search assessing the distance R and 2-fold rotational angle θ in the Pf3 complex (see Figure 1) is shown in Figure 9 and compared to that of its M13 analogue. Electron microscopy and neutron scattering experiments revealed that R must be close to 30 Å in the M13 complex (Gray et al., 1982; Gray, 1989), which at the time allowed a straightforward estimate of the angle θ in the grid search. We chose $\theta = 64°$, and subsequent MD calculations showed that within a few degrees this is indeed the optimal orientation to build a compact, energetically stable protein superhelix (Folmer et al., 1994b). On the basis of a qualitative comparison of the overall shapes of the Pf3 and M13 energy profiles, we established that the center of mass of the Pf3 protein may be 1−2 Å closer to the helix axis and that the 2-fold rotational angle is some 10° smaller (see Figure 9). Hence, (R, θ) = (29 Å, 55°) was used to build the protein superhelix that served as the starting structure for the molecular dynamics calculations.

Prior to the MD calculations it had to be decided which binding stoichiometry was to be used. Fluorescence experiments suggest $n = 3$ (Folmer, 1997), which was imposed in the calculations by applying distance restraints (3.5 Å) between the ring protons of Phe43 and bases of the DNA, with which Phe43 is proposed to have stacking interactions. Thus, bases 1, 4, 7, etc. stack with the Phe43s of successive protein molecules.

Ten complexes were calculated in the geometric force field, out of which the five lowest energy models were subsequently refined in explicit solvent. In all five resulting complexes the protein molecules had rotated counterclockwise by 5−9°, suggesting that the initial value for θ (55°) was ∼7° too small. We therefore repeated the calculations using $\theta = 62°$, and the resulting five protein structures now showed orientations scattered around this value. Figure 10 presents two representations of the lowest energy conformation: Figure 10a,b shows a protein superhelix consisting of 12 dimer molecules, built by repeated symmetry operations on the one protein molecule explicitly simulated in the MD simulation. Figure 10c,d displays ribbon diagrams of a complex of two strands of ssDNA bound to a minihelix of three protein dimers.

Aside from protein reorientation there is also the possibility of movement, typically in the direction perpendicular to the helix axis. This was quantified by calculating radii of gyration. These averaged to 33.1 ± 1.2 Å for the protein in the five models. This number was 31.9 Å in the superhelix built prior to the MD simulations; hence the protein has moved slightly outward. The backbone rms deviation between the protein structures in the five calculated complexes is 1.36 Å for the core residues (i.e., excluding the flexible DNA binding loop and the complex loop) and 2.10 Å for all residues. These numbers were calculated following superpositioning of the structures, so they reflect the protein structures themselves rather than the protein's location in the superhelix. When the molecules are not overlaid, the rms deviations are 2.1 and 2.7 Å, respectively, indicating that the five protein helices are reasonably similar. There is more variation in the structures of bound DNA strands in these complexes. This is due to the fact that the protein−ssDNA interaction is only described by the rather loose restraints obtained from the spin-label study. What is important, however, is that all protein−DNA restraints could be fulfilled in these models. This indicates that the DNA binding domain mapped by the spin-label experiment is completely exposed and available to the ssDNA *also* when the Pf3 protein is in a superhelix. The spin-label studies were done using *$d(A)_3$* to which only one protein molecule can bind at the time. The complete protein surface is thus available to this oligonucleotide, which is not the case when the protein is in a superhelix. Ala69, however, forms the only and remarkable exception. Its Hα−Hβ cross peak is the most affected signal in the TOCSY spectrum, indicating that the *$d(A)_3$* molecule must have been truly close to either an Hα and Hβ proton. The RMD calculations showed,

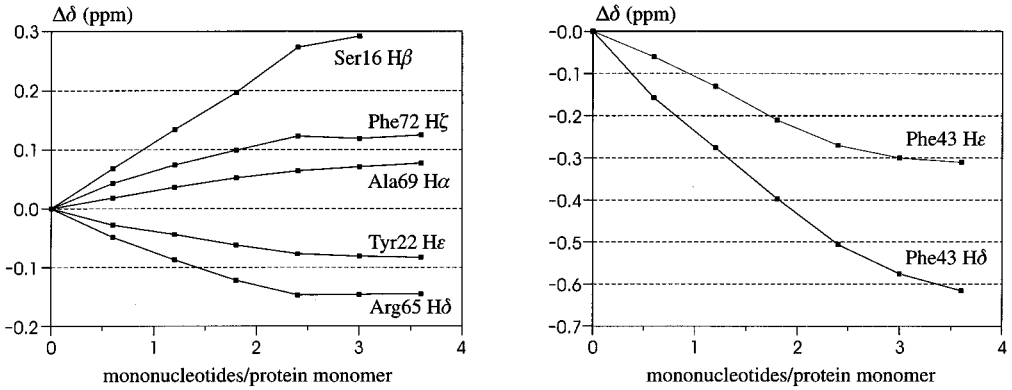

FIGURE 8: Plots of the chemical shift differences ($\Delta\delta$) of several protons in Pf3 ssDBP F36H, as a function of added oligonucleotide $d(A)_6$.

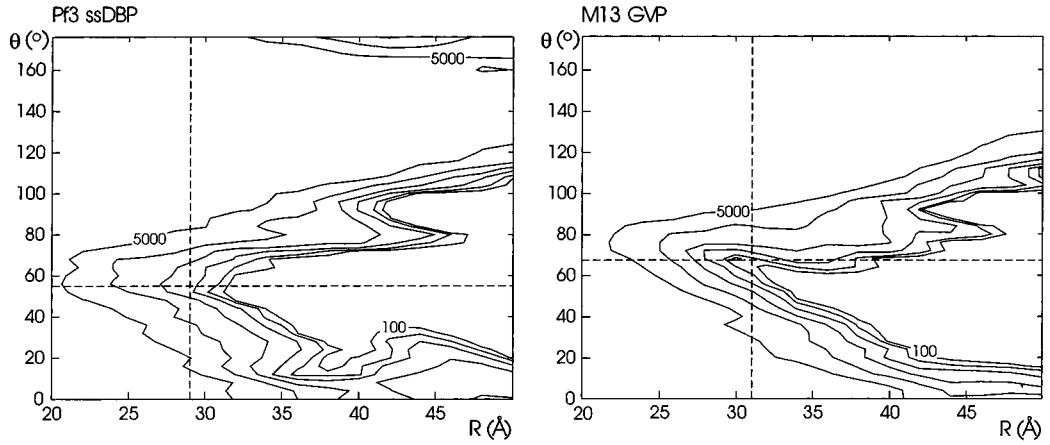

FIGURE 9: Contour plots of the energy matrices calculated in the two-dimensional grid searches for Pf3 ssDBP and M13 GVP. Energies were computed using a simple repulsive nonbonded energy term, in a two-step procedure as described earlier (Folmer et al., 1994b). The normal van der Waals type potential would be much too sensitive for such a global evaluation of the interaction energy. Contour lines are drawn from 100 to 5000 kcal mol^{-1} with a logarithmic distribution. Our experience with the M13 complex is that both higher energies (as R decreases) and lower energies (as R increases and the molecules no longer interact) most likely correspond to unlikely conformations. The dashed lines in the diagram of M13 GVP show the conformation that was measured in the final model, while those in the Pf3 plot show the initial guesses of R and θ.

nonetheless, that in the superhelices built Ala69 cannot be closer than approximately 10−12 Å to the nearest phosphorus atom. To avoid excessive geometric stress in the DNA and in the protein near Ala69, the corresponding spin-label-deduced distance restraint was not incorporated in the final calculations.

DISCUSSION

Structure and Dynamics. In our opinion, the current structure refinement approaches the limit of accuracy that can be obtained in NMR structure determination of proteins of this size (rotational correlation time around 10 ns). This is mainly suggested by the way the distance restraints were collected and used in the RMD calculations. First, all observable protons and the ^{15}N and ^{13}C nuclei to which they are attached had been assigned, and the favorable dispersion of the $^{15}N-^{1}H$ and $^{13}C-^{1}H$ correlation spectra allowed complete and unambiguous assignment of the 3D ^{15}N- and ^{13}C-edited NOESY spectra. Second, the NOESYs were recorded with the exceptionally short mixing time of 40 ms, and NOE buildup curves indicate that this is too short for spin diffusion to be really effective. Consequently, assuming a plain R^{-6} distance dependence on the NOE becomes reasonable, particularly when the molecule is known to be well-structured and rigid. To make optimally use of these accurate spectra, we manually integrated them, ensuring that overlapping peaks are treated properly and only reliable integrals are obtained. No corrections whatsoever were applied (e.g., for pseudoatoms), which was possible as "sum averaging" was used in X-PLOR to calculate effective distances (Nilges, 1993; Folmer et al., 1997). Note that the use of both upper and lower limits constrains the distances to a relatively small region (basically $0.85r$ to $1.15r$), in contrast with most commonly used procedures that effectively do not use lower limits. Finally, the accuracy of the NOESY spectra was further exploited by the use of a floating chirality procedure to treat prochiral centers for which no stereospecific assignments were available.

The many stereospecific assignments and J-couplings furthermore improved the quality of the structures. In this respect, we should mention that out of the several different experiments we performed to assess the ϕ and χ_1 angles, only the ones based on evaluation of cross peak amplitudes were of sufficient signal to noise ratio to permit accurate measurement of the couplings. In particular, the spin−echo difference HSQCs (Grzesiek et al., 1993; Vuister et al., 1993) were very sensitive and allowed determination of χ_1 in Val, Thr, and Ile and permitted stereospecific assignments of the methyl groups in Val without having to resort to fractional labeling (Neri et al., 1989). The structures are in perfect agreement with the NMR data. Moreover, they exhibit excellent covalent bond geometry as indicated by the various rms deviations presented in Table 1 and the Ramachandran plot shown in Figure 5.

Pf3 ssDBP is a well-structured and compact protein. In fact, no less than 54 of its 78 residues reside in β-strands

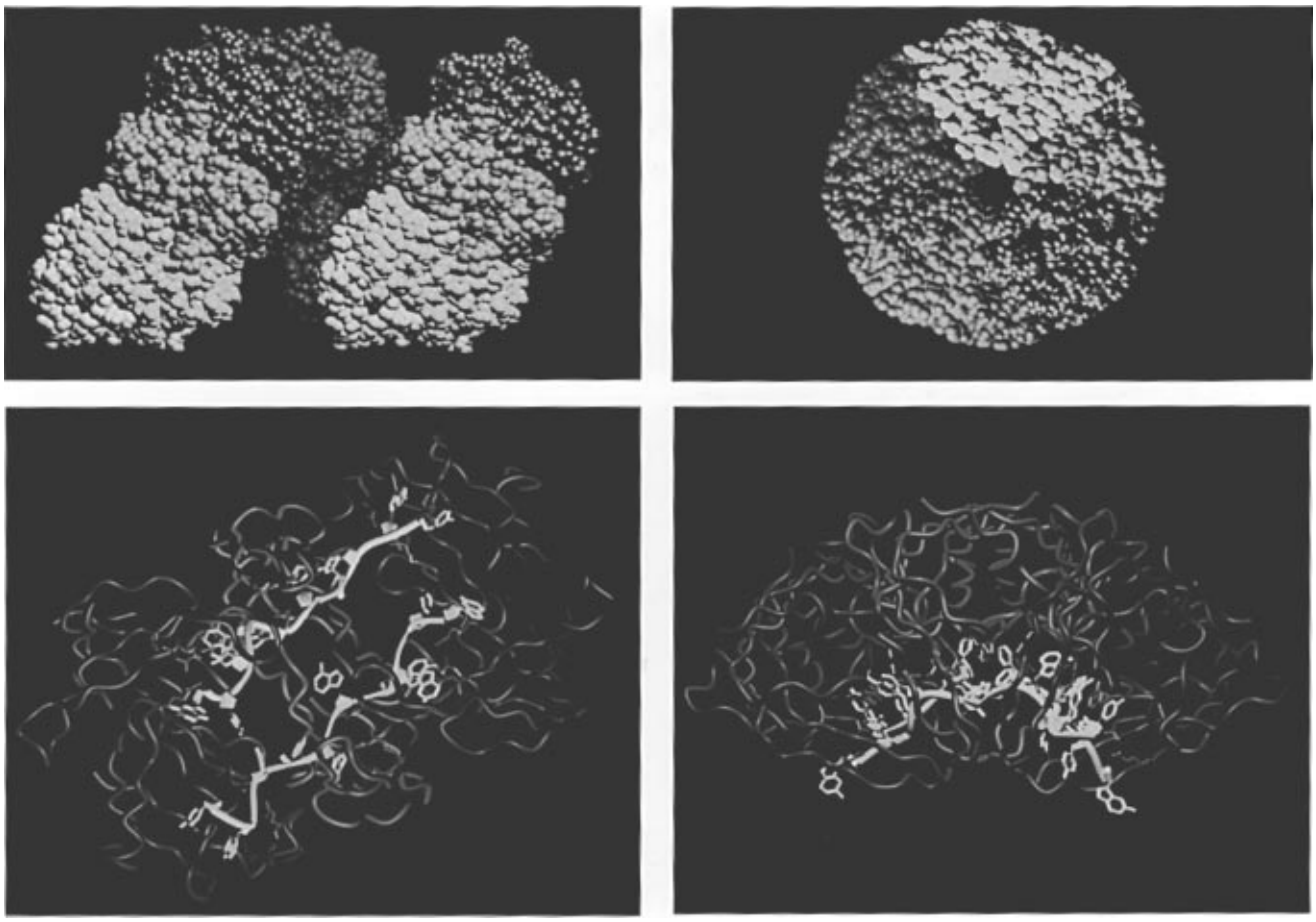

FIGURE 10: (a, top left panel) Space-filling representation of 1.5 turn of the Pf3 ssDBP superhelix. Different colors are used for the eight dimers constituting one helix turn. The DNA is not included in this picture. (b, top right panel) As in (a), now viewed along the helix axis. (c, bottom left panel) Ribbon diagram of three protein molecules (light and dark blue) in the helical conformation and two strands of ssDNA (white and yellow). The residues shown in red are Arg12, Arg65, and Phe43. Note how the latter stacks with the bases of the DNA. The protein−DNA complex is positioned such that the helix axis has a vertical orientation in the plane of drawing, while in (d, bottom right panel) the axis is perpendicular to the plane.

and another 13 are part of β-turns. The type I′ turn connecting β5 and β6 (see Figure 3) is particularly tight, considering that Asn62 and Arg65 still form two β-sheet hydrogen bonds with each other. This is nicely reflected by the dynamics data shown in Figure 6 and 11; Asn63 and Gly64 do not show increased values for $J(\omega_H)$. Rather, the dyad loop as a whole (residues 59−68) appears to be somewhat mobile. The two terminal residues of the protein are also involved in secondary structure interactions. Hydrogen bonds are formed between Met1 O and Tyr57 HN (-1.7 ± 0.7 kcal/mol) and between Ala78 HN and Ser52 O (-3.2 ± 0.2 kcal/mol). Values of $J(0)$ and $J(\omega_H)$ indeed show that Asn2 already belongs to the core of the molecule. The compact fold of the Pf3 protein is furthermore illustrated by the S3 β-bulge formed by residues 8−10. No increased mobility is measured for these amino acids, in accordance with their presumed rigidity. Although forming a tiny β-strand that is part of the five-stranded sheet, the three C-terminal residues are clearly more flexible than the core of the protein.

The most striking feature of Pf3 ssDBP is of course the β-hairpin formed by residues 12−24, the DNA binding wing (DBW). Figure 3 shows that it very much protrudes from the core of the molecule, and not surprisingly, its exact position could not be determined (see Figure 2b). Earlier, we have proposed that this is mainly due to the high flexibility of this loop (Folmer et al., 1995), a suggestion supported by the absence of slowly exchanging amide hydrogens in this region of the protein (Folmer et al., 1994a). The relaxation data presented here confirm this: dramatically different values of $J(0)$ and $J(\omega_H)$ were observed for residues 14−22 compared to the majority of protein residues (Figures 6 and 11). Still, many (homonuclear) NOEs could be observed in the DBW, and locally its structure is actually very well defined. The rms deviation of the backbone atoms of residues 12−24, when overlaying only this region, is as low as 0.26 Å. This suggests that the high mobility is due to hinge motion rather than motion within the hairpin itself. rms difference calculations in which a floating window of five amino acids is fitted to the average coordinates show that the local structure at residues 12−14 is significantly less well determined than at residues 15−20 (data not shown). This suggests that Arg12, Gln13, and Gly14 are the key residues responsible for this hinge motion. The mobility of the DBW has always been postulated to be important for the protein's function; one could envision the loops opening up, presenting the binding cleft to the DNA, and subsequently closing again, truly embracing the bound nucleic acid. Indeed, Figure 11 shows that the flexibility is much reduced upon complexation with $d(A)_6$, to which Pf3 ssDBP only binds in fast exchange and in a noncooperative manner. Even larger effects may be anticipated in the case of cooperative binding in slow exchange, characteristics valid for polynucleotide binding during phage replication. It is attractive to speculate on Arg12 being the key residue in this DNA embracing process. Located in the hinge region it could

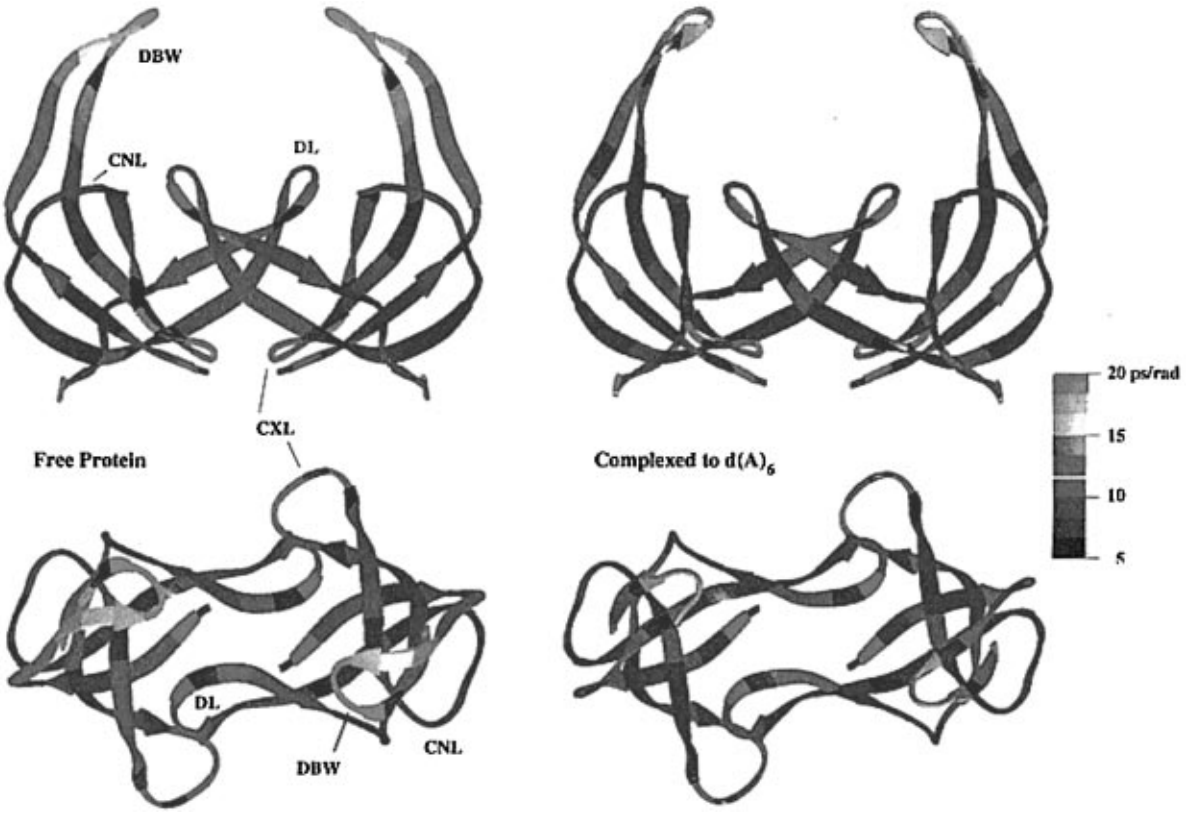

FIGURE 11: Ribbon diagrams of Pf3 ssDBP color-coded with respect to the spectral density $J(\omega_H)$ at 500 MHz (cf. Figure 6). The orientations of the top and bottom pictures differ by 90°. The coordinates of $\langle \overline{SA} \rangle_r$ were used to draw the ribbon of the uncomplexed protein, while the right-handed ribbon was generated from the coordinates of the protein in the model of the nucleoprotein complex (see Figure 10). Black was used to color residues for which no parameters could be obtained, either due to spectral overlap or because the residue is a proline.

somehow induce opening of the loops if triggered by electrostatic interactions with DNA that is nearby but not yet bound. Likewide, once the DNA has been "guided" into the binding cleft, interactions with Arg12 may again result in closure. Interestingly, there are arginines at exactly the same (relative) position also in the DBWs of the M13 and IKe gene V proteins (Folkers et al., 1991b; De Jong et al., 1989). Replacing this arginine in the M13 GYP with Cys, Asp, or His leads to nonfunctional or partially functional proteins, based both on an assay for inhibition of *E. coli* growth (Terwilliger et al., 1994) and on the ability of R16C and R16H mutants to inhibit packaging of phage ssDNA and to repress translation of gene II mRNA (Stassen et al., 1992). However, most *putative* DNA binding loops found in sequence alignment studies in a number of other ssDBPs (e.g., ϕ29 p5, If1 DBP, Pf1 and I2-2 GVP) lack a positively charged residues at this position (Stassen et al., 1995). Human vimentin, on the other hand, whose N-terminal half contains a ssDNA binding domain that is structurally and functionally similar to the M13 GVP (Traub et al., 1992), does contain the arginine; sequence alignment by Stassen (1994) suggests that vimentin features a DNA binding wing as well, in which Arg28 is at the corresponding position to Arg12 in Pf3 ssDBP.

The large difference in mobility measured for the DNA binding wing in the presence and absence of DNA is in sharp contrast with a recent dynamics study on the C-terminal ssDNA binding domain of *E. coli* topoisomerase I (Yu et al., 1996). There, the binding of 12-mer ssDNA to the free protein caused only a slight decrease in the mobility of a few specific residues, while in fact the mobility of the majority of residues slightly increased. Although *E. coli* topoisomerase I has some structural similarity to M13 GVP (and Pf3 ssDBP) (Yu et al., 1995), its DNA binding mechanism is probably quite different. This is further illustrated by the fact that its DNA binding loop (residues 41−46) is actually quite rigid (Yu et al., 1996). To our knowledge, no other studies have been reported on backbone dynamics of both the free and complexed form of DNA binding proteins.

A second region of structural variability in the ensemble of structures, though much less pronounced, is the complex loop (Figure 2b). Here, the relaxation data indicate as well that this is due to dynamic effects rather than an intrinsic lack of structural restraints. Spectra densities at frequencies zero and ω_N are clearly smaller than average, and $J(\omega_H)$ is higher (see also Figure 11). While motion of the DBW was significantly diminished upon oligonucleotide binding, this was not observed for the complex loop. This leads to two important conclusions: first, the complex loop is not involved in DNA binding. We knew already from the spin-label study that residues in the complex loop do not make direct contact with bound DNA, and now the relaxation data suggest, in addition, that allosteric effects are probably negligible as well; the dynamics of the DBW appears to be independent from that of the complex loop (at the time scales studied). Second, cooperativity indeed seems to be fully absent in binding to $d(A)_6$, as the concomitant protein−protein interactions would have affected the mobility of the complex loop. This absence of cooperativity is explained not only by the short length of the oligonucleotide but also by experiments being conducted using the Phe36 → His "solubility" mutant. On the basis of the homologies with M13 GVP, Phe36 is expected to be one of the crucial residues in the formation of protein multimers, not only in solution (causing aggregates) but also when bound to DNA (causing cooperat-

ivity). Thus, the use of the F36H mutant in combination with the short oligonucleotide enabled us to study dynamical effects solely to be attributed to the protein−DNA interaction. In a future study, the relaxation parameters of wild-type Pf3 protein bound to larger DNA fragments will be measured, which should provide information on the protein−protein interactions responsible for cooperativity.

The role of residue 36 in protein−protein interactions can be rationalized on the basis of the three-dimensional structure and the existing similarity with M13 GVP. Figure 2c shows that His36 is one of the most solvent exposed residues in the F36H protein, and the near-identity of the NMR spectra of wild-type and F36H protein (Folmer et al., 1994a) strongly suggests that Phe36 is similarly exposed in the wild-type protein. In M13 GVP, Tyr41 is at the position corresponding to Phe36, and X-ray analyses showed that Tyr41 intercalates in a cavity formed by Tyr34, Tyr41, and Pro42 on the adjacent molecule in the crystal (Guan et al., 1994). The Pf3 structure reveals that a similar cavity could be formed by Phe28, Phe36, and Pro37, three hydrophobic residues which are all relatively much exposed. We believe therefore that the hydrophobic interactions between Phe36 on one dimer and this cavity on the other are the main driving force for the observed aggregation in solution. Consequently, it is perfectly reasonable that substituting Phe36 with the much less hydrophobic histidine (which, in addition, is charged under our NMR conditions) suppresses aggregation. Phe36 will most likely interact in a different manner with adjacent protein molecules when bound to DNA *(vide infra)*, but similar arguments can be used to explain its importance there.

The structural variability in the DBW and the complex loop has been attributed to dynamic effects, and this appears to be true as well for the connecting loop and dyad loop (note the correlation between Figure 2a and the plot of $J(\omega_H)$ in Figure 6). The relaxation data thus explain the higher rms values obtained for some parts of the protein. Furthermore, in a qualitative manner, changes in dynamics upon oligonucleotide binding have been discussed. Here, we refrain from an analysis of the relaxation data in terms of the popular "model-free" approach for macromolecules with isotropic overall reorientation (Lipari & Szabo, 1982) because of the highly asymmetric form of the Pf3 ssDBP dimer. The atomic coordinates of the average structure reveal a relative ratio of 1:1.3:1.9 for the principal components of its inertia tensor. This indicates that the molecule is totally anisotropic; hence the shape of the molecule can be approximated neither by a sphere nor by an ellipsoid. We are currently in the process of fitting models based on anisotropic diffusion tensors (Woesner, 1962), for which we think Pf3 ssDBP forms a challenging model system (L. M. Horstink, R. H. A. Folmer, R. N. H. Konings, and C. W. Hilbers, in preparation).

Nucleic Acid Binding. Despite the biological importance of ssDBPs, structural studies on this type of protein have not advanced at the same pace as their double-stranded counterparts. From an NMR point of view, this is almost completely due to the strong tendency present in these proteins to form aggregates at concentrations needed for structural analysis. The ssDBP of phage *ϕ29*, for instance, forms a gel at concentrations ≥0.5 mM, while that of I2-2 precipitates already at concentrations above 0.1 mM (unpublished results). The two solution structures that have been reported for this class of proteins were in fact determined using solubility mutants (Folkers et al., 1991a; Folmer et al., 1994a). Athough X-ray analysis is in principle more tolerant to this phenomenon, no more than five structures could be solved: M13 GVP (Skinner et al., 1994), adenovirus ssDBP (Tucker et al., 1994), T4 gp32 (Shamoo et al., 1995), human replication protein A (Bochkarev et al., 1997), and human mitochondial ssDBP (Yang et al., 1997). So far, only one high-resolution structure of a protein−ssDNA complex has been reported, the 2.4 Å X-ray structure of replication protein A in complex with $d(C)_8$ (Bochkarev et al., 1997). This scarceness is largely due to problems of technical sort such as too high molecular weights (NMR) or the inability to grow suitable cocrystals (X-ray). However, a more intrinsic property may hinder structural studies as well, which is the nature of the protein−ssDNA interaction itself. The general idea is that ssDBPs bind to DNA via hydrophobic and electrostatic interactions, both of which are rather nonspecific. Not only will an ssDBP bind with similar preference to various sequences, it is even conceivable that binding to the same sequence may occur in a variety of (DNA) conformers that are relatively isoenergetic with respect to each other. This has been proposed to explain the apparent ssDNA disorder in the cocrystals of T4 gp32 (Shamoo et al., 1995). It also helps to explain the fast exchange in DNA binding of Pf3 ssDBP and the fact that the aromatic protons of Phe43 make NOE contacts to all bases of the $d(A)_6$. It should be noted that both in the T4 gp32 crystals and in the Pf3−$d(A)_6$ complex cooperativity is absent. A more rigid complex may be formed in case a long stretch of ssDNA is saturated with protein molecules. Then, the DNA will probably be prevented from constantly dissociating and associating, but sliding along the protein "lattice" is still very well feasible. This latter property is likely to be related to a number of functions generally ascribed to ssDBPs, such as DNA replication, repair, and recombination. In the case of Pf3 ssDBP, sliding mechanisms may be important in its ability to rearrange along the growing ssDNA chain during rolling circle replication or in the process of exchange with the coat proteins during phage assembly.

Functionally relevant or not, this suppleness forms a major obstacle in the study of protein−ssDNA interaction, both by X-ray and by NMR. In particular, NOE contacts are very difficult to interpret as they represent an average situation, both in time and in population, of the many possible conformations. Therefore, we had to use spin-labels to study the complexes. As shown in Figure 7, the protein clearly contains two distinct DNA binding domains, each formed by conjunction of the DNA binding wing and the dyad loop of different monomers. The tip of the DBW, however, appears to be quite distant from the bound DNA, as broadening effects were measured of only 40%, 20%, and 10% on Ala17, Lys18, and Gly19, respectively (see Table 3). The DNA binding domain of Pf3 ssDBP is very similar to that of M13 GVP, which was mapped using spin-labeled oligonucleotides as well (Folkers et al., 1993). Affected residues on M13 GVP are also situated on the DBW, the fourth β-strand, and the dyad loop, and approximately the same degrees of broadening were observed for equivalent residues. There are, however, two differences, the first of which is in the dyad loop. The secondary structure comparison already showed that this region is anyway quite different in the two proteins. M13 GVP contains a 3_{10} helix in the N-terminal half of this β-hairpin, while there is a continuous β-strand in Pf3 ssDBP, and there is no apparent

sequence homology. Asp79 and Arg80, two residues at the C-terminal half of the dyad hairpin of GVP, were significantly broadened in the spin-label experiment (Folkers et al., 1993), but there are no obvious counterparts for these residues in Pf3 ssDBP: not only are the sequences too divergent to be aligned for this region but a best-fit superposition of the three-dimensional structures simply shows that the C-terminal halves of the respective dyad hairpins are relatively far separated. In fact, Arg65 of Pf3 ssDBP may be considered the equivalent of Arg80, although they are several residues separated both in sequence and in structure. Both residues are prominently present in the DNA binding cleft, and there is no positively charged residue at the tip of the dyad hairpin in M13 GVP that would be the obvious counterpart of Arg65. The second difference involves Phe43 in the Pf3 protein, which is one of the most affected residues in the spin-label experiment. Secondary and tertiary structure comparisons show that Thr48 in M13 GVP is at the position corresponding to Phe43, but Thr48 was less broadened. Instead, Phe73 at the tip of the dyad hairpin in GVP was completely broadened, and this residue appears to be the true equivalent of Phe43: both residues are solvent-exposed and make stacking interactions with the bases of bound ssDNA (King & Coleman, 1988; Folkers et al., 1991a).

The DNA binding data presented in this work are in full agreement with the general idea on protein−DNA interaction for this class of proteins, i.e., hydrophobic interactions between specific amino acids and the DNA bases or electrostatic interactions between the phosphate groups and arginine and lysine side chains. The hydrophobic interactions include base stacking by exposed phenylalanine residues such as Phe73 in M13 GVP, Phe43 in Pf3 ssDBP, and presumably Phe43 as well in *E. coli* topoisomerase I (Yu et al., 1996). Figure 12a illustrates the importance of the electrostatic component in the Pf3−ssDNA interaction. Arg12, Lys18, and Arg65 cluster together to form a large positive potential that will attract the DNA backbone and direct it into the binding cleft. Lys18 is likely to be an essential residue in this respect, although it is not so close to the bound DNA, given the relatively mild broadening in the spin-label experiment.

Model of the Nucleoprotein Complex. Summarizing the results discussed so far, we may conclude that the Pf3 protein structure is highly refined, its DNA binding domain has been quite precisely mapped, the role of certain specific amino acids (notably Phe43) has been elucidated, and there is some idea as to how many nucleotides are bound per protein molecule. These data justified the attempt to build a model for the Pf3 protein−ssDNA complex. Although clear macroscopical data on this nucleoprotein complex have not been reported to date, we believe the similarities to the well-studied M13 complex should allow proper estimates for the parameters describing the Pf3 superhelical complex (Figure 1). The results from the grid search confirmed this suggestion: the contour plots in Figure 9 reveal a high level of similarity in the nonbonded energy profiles of the M13 and Pf3 proteins in their respective superhelices. Apparently, the proposed values for the helix pitch (80 Å) and twist angle (45°) of the Pf3 complex are quite reasonable. Furthermore, this similarity provides a firm enough basis to deduce proper values for distance R and angle θ, the final two parameters necessary to build the superhelix. Obviously, it is uncertain whether these numbers represent the deepest minimum in this four-dimensional energy surface, but the conformation they describe is at least energetically similar to that of the M13 complex, for which the model has been independently derived by two groups on the basis of a large variety of biophysical and biochemical data (Folmer et al., 1994b; Guan et al., 1994). Anyway, the precise values of R and θ are not too critical in this stage of the procedure, as the protein is allowed to move (changing R) or reorient (changing θ) while adjusting to the neighboring protein molecules during the MD simulation. Indeed, in the initial round of MD calculations it was found that the protein consistently rotated counterclockwise and θ was adjusted by 7°.

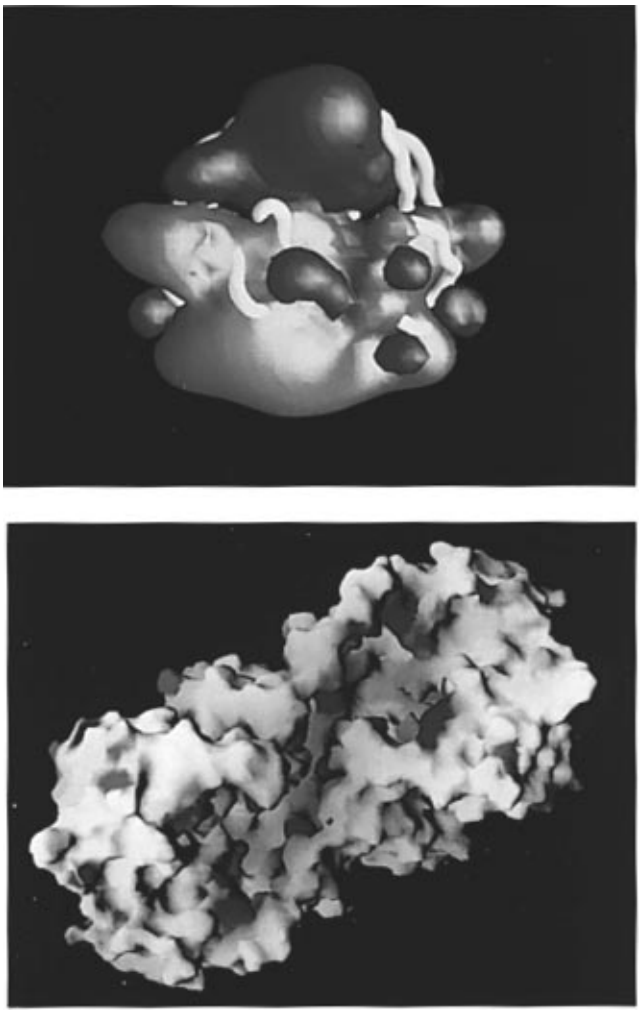

FIGURE 12: (a, top panel) Plot of the isopotential surface of Pf3 ssDBP at −1 kT (red) and +1 kT (blue). The protein backbone is represented by a white "worm". The orientation of the protein is identical to that in Figure 3. (b, bottom panel) View of the electrostatic surface (−12 to +12 kT) of four protein molecules in the superhelix conformation. The helix axis runs vertically in the plane. Both plots were made with GRASP (Nicholls et al., 1991).

Firm quantitative electron microscopy data for the Pf3 complex being absent, it would be premature to discuss in great detail the quaternary interactions of specific amino acids in the current model. On a subresidue level, however, it is apparent that the complex loops of adjacent molecules in the superhelix interact with one another. Similarly, the dyad loops are close to each other, as are the DNA binding loops. Also, the coil region connecting strands β6 and β7 appears to be involved in quaternary contacts. Detailed interactions are of course much dependent on the precise values of the four helix parameters, which will have to be unraveled by future studies. Still, we do not expect these parameters to

deviate drastically from those of the M13 complex, so that the model reported here is likely to be correct in its general features.

One of these features is that the superhelix of Pf3 molecules forms two continuous, positively charged paths on its inner surface, ideal targets for two ssDNA strands (Figure 12b). Another aspect of the model is the binding stoichiometry. CD measurements on complexes formed by wild-type Pf3 ssDBP and various polynucleotides indicated a stoichiometry of $n = 2$ (Powell & Gray, 1993), but the authors also found evidence for $n = 4$ binding in gel electrophoresis experiments. A recent CD study on a mutant (Tyr22 → Phe) confirms the presence of this $n = 4$ mode (Powell & Gray, 1995). The model shown in Figure 10 is clearly incompatible with an $n = 2$ stoichiometry: the average distance between phosphorus atoms in the calculated complexes is 5.9 ± 0.9 Å. Here, three nucleotides are bound per protein monomer (see Results), which means that the binding path measures 3×5.9 or ~18 Å per protein unit. This distance is too large to be spanned by only two nucleotides. We feel, therefore, that it is unlikely that the Pf3 protein binds to polynucleotides covering only two nucleotides per monomer. Presumably, the $n = 2$ ratios found in the CD measurements are the result of unsaturated protein (or "oversaturated" nucleic acid) at the end points of the titrations, as indeed the authors propose themselves (Powell & Gray, 1993). Our calculations cannot distinguish between $n = 3$ or higher ratios, but $n = 3$ is clearly the minimum stoichiometry. Higher ratios simply infer that the DNA will be more compressed.

The main benefit of the presented model is that it demonstrates the consistency of the various results reported in this paper. Apparently, it is possible to construct a superhelix that features two continuous binding paths for ssDNA, by rotation and translation of the protein coordinates. What is more, the residues constituting these paths are those broadened by the spin-labeled oligonucleotide, whereas the residues that are distant from the ssDNA in the model were unperturbed (except Ala69) in the spin-label experiment. Furthermore, the model shows that the position of Phe43 indeed allows its side chain to stack with the bases of the DNA. Therefore, we believe that these complex calculations provide a suitable working model that will help to understand the structure−function relationship of the Pf3 ssDNA binding protein.

CONCLUSIONS

Previously, we have reported the three-dimensional structure, DNA binding domain, binding characteristics, and model of the nucleoprotein complex of M13 gene V protein. Here, we present the same set of results obtained for a different and less well-known member of the class of ssDNA binding proteins, the ssDBP encoded by bacteriophage Pf3. The M13 and Pf3 phages share only little sequence homology, and it is even possible that they have evolved from different ancestral viruses. Nevertheless, we have demonstrated that their ssDBPs share many homologies: the structures possess the same overall fold, their DNA binding domains are very similar, and comparable models can be built for the respective nucleoprotein complexes. Clearly, the two proteins have many basic principles in common, and we expect that by finding and understanding these essentials, also in other proteins, insight will eventually be gained into how this important class of proteins functions.

ACKNOWLEDGMENT

The NMR experiments were performed at the SON Hf-NMR facility (Nijmegen, The Netherlands). We thank Jos Joordens and Jan van Os for expert technical assistance. Dr. Kevin Leonard is acknowledged for performing the electron microscopy experiments.

SUPPORTING INFORMATION AVAILABLE

A list of ^{1}H, ^{15}N, and ^{13}C resonance assignments of Pf3 ssDBP F36H (2 pages). Ordering information is given on any current masthead page.

REFERENCES

Alberts, B., & Frey, L. (1970) *Nature 227*, 1313−1318.
Alberts, B., & Sternglanz, R. (1977) *Nature 269*, 655−661.
Alma, N. C. M., Harmsen, B. J. M., de Jong, E. A. M., Van de Ven, J., & Hilbers, C. W. (1983) *J. Mol. Biol. 163*, 47−62.
Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636−641.
Bartels, C., Xia, T., Billeter, M., Güntert, P., & Wüthrich, K. (1995) *J. Biomol. NMR 5*, 1−10.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−366.
Bax, A., Max, D., & Zax, D. (1992) *J. Am. Chem. Soc. 114*, 6923−6925.
Bochkarev, A., Pfuetzner, R. A., Edwards, A. M., & Frappier, L. (1997) *Nature 385*, 176−181.
Bruccoleri, R. E., & Karplus, M. (1986) *J. Comput. Chem. 7*, 165−175.
Brünger, A. T. (1992) *X-PLOR. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
Bulsink, H., Harmsen, B. J. M., & Hilbers, C. W. (1985) *J. Biomol. Struct. Dyn. 3*, 227−247.
Carson, M. (1991) *J. Appl. Crystallogr. 24*, 958−961.
Casadevall, A., & Day, L. A. (1985) *Virology 145*, 260−272.
Chan, A. W. E., Hutchinson, E. G., Harris, D., & Thornton, J. M. (1993) *Protein Sci. 2*, 1574−1590.
Chase, J. W., & Williams, K. R. (1986) *Annu. Rev. Biochem. 55*, 103−106.
Cheng, X., Harms, A. C., Goudreau, P. N., Terwilliger, T. C., & Smith, R. D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 7022−7027.
Dayie, K. T., & Wagner, G. (1994) *J. Magn. Reson., Ser. A 111*, 121−126.
De Jong, E. A. M., Van Duynhoven, J. P. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1989) *J. Mol. Biol. 206*, 119−132.
Dijkstra, K., Kroon, G. J. A., Van Nuland, N. A. J., & Scheek, R. M. (1994) *J. Magn. Reson., Ser. A 107*, 102−105.
Engh, R., & Huber, R. (1991) *Acta Crystallogr. A47*, 392−400.
Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A., & Kay, L. E. (1995) *J. Biomol. NMR 6*, 153−162.
Folkers, P. J. M., Stassen, A. P. M., Van Duynhoven, J. P. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1991a) *Eur. J. Biochem. 200*, 139−148.
Folkers, P. J. M., Van Duynhoven, J. P. M., Jonker, A. J., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1991b) *Eur. J. Biochem. 202*, 349−360.
Folkers, P. J. M., Van Duynhoven, J. P. M., Van Lieshout, H. T. M., Harmsen, B. J. M., Van Boon, J. H., Tesser, G. I., Konings, R. N. H., & Hilbers, C. W. (1993) *Biochemistry 32*, 9407−9416.
Folkers, P. J. M., Nilges, M., Folmer, R. H. A., Konings, R. N. H., & Hilbers, C. W. (1994) *J. Mol. Biol. 236*, 229−246.
Folmer, R. H. A. (1997) Understanding single-stranded DNA. Nuclear magnetic resonance and restrained molecular dynamcis studies on the M13 and Pf3 proteins and their nucleoprotein complexes, Ph.D. Thesis, Nijmegen, The Netherlands.
Folmer, R. H. A., Folkers, P. J. M., Kaan, A., Jonker, A. J., Aelen, J. M. A., Konings, R. N. H., & Hilbers, C. W. (1994a) *Eur. J. Biochem. 224*, 663−676
Folmer, R. H. A., Nilges, M., Folkers, P. J. M., Konings, R. N. H., & Hilbers, C. W. (1994b) *J. Mol. Biol. 240*, 341−357.
Folmer, R. H. A., Nilges, M., Konings, R. N. H., & Hilbers, C. W. (1995) *EMBO J. 14*, 4132−4142.

Folmer, R. H. A., Konings, R. N. H., Hilbers, C. W., & Nilges, M. (1997) *J. Biomol. NMR 9*, 245−258.
Gray, C. W. (1989) *J. Mol. Biol. 208*, 57−64.
Gray, D. M., Gray, C. W., & Carlson, R. D. (1982) *Biochemistry 21*, 2702−2713.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
Grzesiek, S., & Bax, A. (1992) *J. Magn. Reson. 96*, 432−440.
Grzesiek, S., Vuister, G. W., & Bax, A. (1993) *J. Biomol. NMR 3*, 487−493.
Grzesiek, S., Kuboniwa, H., Hinck, A. P., & Bax, A. (1995) *J. Am. Chem. Soc. 117*, 5312−5315.
Guan, Y., Zhang, H., Konings, R. N. H., Hilbers, C. W., Terwilliger, T. C., & Wang, A. H.-J. (1994) *Biochemistry 33*, 7768−7778.
Guan, Y., Zhang, H., & Wang, A. H.-J. (1995) *Protein Sci. 4*, 187−197.
Holak, T. A., Nilges, M., & Oschkinat, H. (1989) *FEBS Lett. 242*, 649−654.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.
Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204−209.
Ishima, R., & Nagayama, K. (1995a) *Biochemistry 34*, 3162−3171.
Ishima, R., & Nagayama, K. (1995b) *J. Magn. Reson., Ser. B 108*, 73−76.
Jorgensen, W. L., & Tirado-Rives, J. (1988) *J. Am. Chem. Soc. 110*, 1657−1666.
Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983) *J. Chem. Phys. 79*, 926−935.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
Kansy, J. W., Clack, B. A., & Gray, D. M. (1986) *J. Biomol. Struct. Dyn. 3*, 1079−1110.
Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110−126.
Kay, L. E., Torchia, D. A., & Bax, A. (1989) *Biochemistry 28*, 8972−8979.
Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., & Torchia, D. A. (1992) *J. Magn. Reson. 97*, 359−375.
King, G. C., & Coleman, J. E. (1988) *Biochemistry 27*, 6947−6953.
Koradi, R., Billeter, M., & Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
LeMaster, D. M., Kay, L. E., Brünger, A. T., & Prestegard, J. H. (1988) *FEBS Lett. 236*, 71−76.
Lipari, G., & Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
Neri, D., Szyperski, T., Otting, G., Senn, H., & Wüthrich, K. (1989) *Biochemistry 28*, 7510−7516.
Nicholls, A., Sharp, K. A., & Honig, B. (1991) *Proteins 11*, 281−296.
Nilges, M. (1993) *Proteins 17*, 297−309.
Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988) *Protein Eng. 2*, 27−38.
Nilges, M., Kuszewski, J., & Brünger, A. T. (1991) in *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy* (Hoch, J. C., Poulsen, F. M., & Redfield, C., Eds.) pp 451−455, Plenum Press, New York.
Olah, G. A., Gray, D. M., Gray, C. W., Kergil, D. L., Sosnick, T. R., Mark, B. L., Vaughan, M. R., & Trewhella, J. (1995) *J. Mol. Biol. 249*, 576−594.
Palmer, A. G., Rance, M., & Wright, P. E. (1991) *J. Am. Chem. Soc. 113*, 4371−4380.
Palmer, A. G., Skelton, N. J., Chazin, W. J., Wright, P. E., & Rance, M. (1992) *Mol. Phys. 75*, 699−711.
Powell, M. D., & Gray, D. M. (1993) *Biochemistry 32*, 12538−12547.
Powell, M. D., & Gray, D. M. (1995) *Biochemistry 34*, 5635−5643.
Prompers, J. J., Folmer, R. H. A., Nilges, M., Folkers, P. J. M., Konings, R. N. H., & Hilbers, C. W. (1995) *Eur. J. Biochem. 232*, 506−514.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
Shamoo, Y., Friedman, A. M., Parsons, M. R., Konigsberg, W. H., & Steitz, T. A. (1995) *Nature 376*, 362−366.
Skinner, M. M., Zhang, H., Leschnitzer, D. H., Guan, Y., BEellamy, H., Sweet, R. M., Gray, C. W., Konings, R. N. H., Wang, A. H.-J., & Terwilliger, T. C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2071−2075.
Stassen, A. P. M. (1994) Studies of the relationship between structure and function of the single-stranded DNA binding protein of the filamentous bacteriophage M13, Ph.D. Thesis, Nijmegen, The Netherlands.
Stassen, A. P. M., Zaman, G. J. R., Van Deursen, J. M. A., Schoenmakers, J. G. G., & Konings, R. N. H. (1992) *Eur. J. Biochem. 204*, 1003−1014.
Stassen, A. P. M., Folmer, R. H. A., Hilbers, C. W., & Konings, R. N. H. (1995) *Mol. Biol. Rep. 20*, 109−127.
Stonehouse, J., Shaw, G. L., Keeler, J., & Laue, E. D. (1994) *J. Magn. Reson., Ser. A 107*, 178−184.
Terwilliger, T. C., Zabin, H. B., Horvath, M. P., Sandberg, W. S., & Schlunk, P. M. (1994) *J. Mol. Biol. 236*, 556−571.
Traub, P., Mothes, E., Shoeman, R., Kühn, S., & Scherbarth, A. (1992) *J. Mol. Biol. 228*, 41−57.
Tucker, P. A., Tsernoglou, D., Tucker, A. D., Coenjaerts, F. E. J., Leenders, H., & Van der Vliet, P. C. (1994) *EMBO J. 13*, 2994−3002.
Van Duynhoven, J. P. M., Nooren, I. M. A., Swinkels, D. W., Folkers, P. J. M., Harmsen, B. J. M., Konings, R. N. H., Tesser, G. I., & Hilbers, C. W. (1993) *Eur. J. Biochem. 216*, 507−517.
Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425−1436.
Vuister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
Vuister, G. W., Wang, A. C., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 5334−5335.
Weber, P. L., Morrison, R., & Hare, D. (1988) *J. Mol. Biol. 204*, 483−487.
Weiner, S. J., Kollmann, P. A., Case, D. A., Singh, U., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765−784.
Woesner, D. E. (1962) *J. Chem. Phys. 36*, 1−4.
Yang, C., Curth, U., Urbanke, C., & Kang, C.-H. (1997) *Nat. Struct. Biol. 4*, 153−157.
Yu, L., Zhu, C.-X., Tse-Dinh, Y.-C., & Fesik, S. W. (1995) *Biochemistry 34*, 7622−7628.
Yu, L., Zhu, C.-X., Tse-Dinh, Y.-C., & Fesik, S. W. (1996) *Biochemistry 35*, 9661−9666.
Zink, T., Ross, A., Luers, K., Cieslar, C., Rudolph, R., & Holak, T. A. (1994) *Biochemistry 33*, 8453−8463.

BI970251T