Structure of the DNA binding wing of the gene-V encoded single-stranded DNA binding protein of the filamentous bacteriophage M13

J.P.M. van Duynhoven[+], P.J.M. Folkers[+], A.P.M. Stassen[*], B.J.M. Harmsen[+], R.N.H. Konings[+*] and C.W. Hilbers[+]

[+] *Laboratory of Biophysical Chemistry* and [*] *Laboratory of Molecular Biology, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands*

Received 10 November 1989

The structure in solution of a β-loop in mutant Y41H of the single-stranded DNA binding protein encoded by gene-V of the filamentous phage M13 has been elucidated using 2-dimensional ^{1}H-nuclear magnetic resonance techniques. Furthermore, these studies enabled us to demonstrate that an identical structural element is present in wild-type gene-V-protein and that this element intimately is involved in the binding of gene-V-protein to single-stranded DNA. It is shown that the structure of the DNA binding wing deviates from that proposed for the same amino acid sequence on the basis of X-ray diffraction data. The structure is, however, identical to that of the DNA binding wing present in the single-stranded DNA binding protein encoded by the genome of the evolutionary distantly related filamentous phage IKe. The latter observations support our current view that in the binding of these proteins to single-stranded DNA a common structural motif is involved.

DNA binding protein; Bacteriophage M13; Gene-V-protein; Phage IKe

1. INTRODUCTION

In the life-cycle of the filamentous phages, such as Ff (M13,f1, fd), IKe or Pf3, a single-stranded DNA binding protein encoded by gene-V (GVP) plays an indispensable role in the DNA replication process (for a review, see [1]. Recently, by using 2D-NMR techniques, the structure of part of the DNA binding domain of IKe GVP has been elucidated [2]. This structure consists of a β-loop and comprises residues 16 through 29 of the 88 amino acid residues long polypeptide. It is noted in passing that in solution these proteins occur as dimers (20 kDa) and their structure determination by NMR forms a considerable challenge. Alignment of the amino acid sequences of the ssDNA binding proteins encoded by IKe, M13 and Pf3 has demonstrated that the GVP of the latter two contains a stretch of amino acid residues that is highly homologous to that of the afore-mentioned β-loop sequence in IKe GVP [3] (fig.1). This homology becomes even more apparent when these amino acid sequences are also folded into a β-loop and the residues which are situated at the DNA binding face are compared.

Comparison of the hypothetical β-loop structure in M13 GVP with the structure postulated for the same region on the basis of X-ray crystallographic studies [4] revealed that they differ in two respects. First, the proposed β-loop structure is more regular than the β-loop in the crystal structure, and second, the residues involved in the formation of the loops are shifted by 4 amino acids with respect to one another.

In this communication it will be demonstrated that the β-loop structure, as proposed on the basis of the observed similarities with IKe GVP, is indeed present in both mutant (Y41H) and wild-type M13 GVP. This observation highly supports the thesis that in this class of DNA binding proteins a common structural motif is present for recognition of and binding to ssDNA.

2. MATERIALS AND METHODS

Wild-type and mutant (Y41H) GVP was prepared and isolated as described previously [5]. The gel filtration step was, however, replaced by chromatography on MonoS cation-exchange FPLC-column (Pharmacia). After dialysis against 1 mM cacodylate buffer (pH 6.9) containing 50 mM NaCl, the purified GVP was lyophilised and stored at -20°C. The final concentration of mutant GVP in the NMR-samples was about 1–4 mM, with no buffer components and only a few millimolar of NaCl. The pH was adjusted to pH 5.2 with diluted DCl.

The procedures used for the construction and expression of a library of M13 GVP mutants will be described elsewhere [6]. The GVP mutant, Y41H, whose solution structure has been investigated in the 2D-NMR studies described, contained a histidine instead of a

Correspondence address: C.W. Hilbers, Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

Abbreviations: 2D-NMR, two-dimensional nuclear magnetic resonance spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; TPPI, time proportional phase incrementation; GVP, gene-V-protein; ssDNA, single-stranded DNA

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

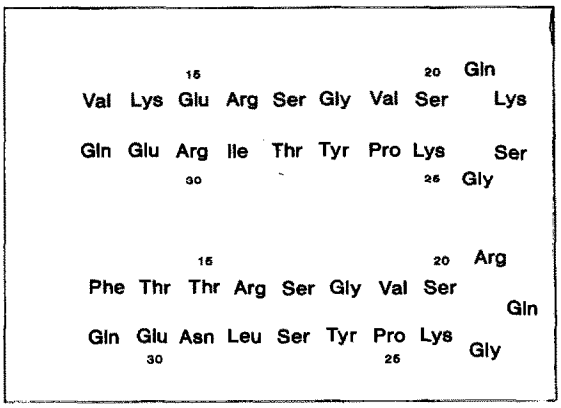

Fig.1. Comparison of the β-loop structures of (top) IKe GVP (residues 13–32) and (bottom) M13 GVP (residues 13–31).

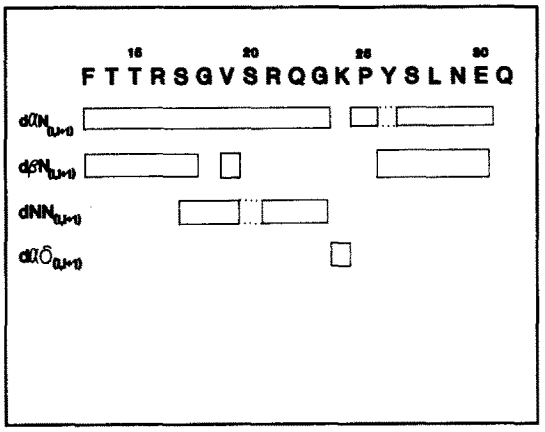

Fig.2. The sequentially assigned residues of Y41H GVP witha summary of dαN, dβN, dNN, dα-(δ,δ') contacts.

tyrosine residue at position 41 of wild-type M13 GVP.

The NMR experiments were performed at 600 MHz on a Bruker AM 600 spectrometer interfaced to an ASPECT 3000 computer. Clean TOCSY (MLEV17) [7,8], DQF-COSY [9] and NOESY [10] experiments were conducted at 298 K. The TOCSY spectra were recorded with mixing times ranging from 24 to 40 ms and the NOESY spectra were recorded with mixing times of 100 and 175 ms. In NOESY(H_2O) experiments the solvent resonance was suppressed by applying a semi-selective jump-return [11] pulse sequence. In the TOCSY(H_2O) measurements both continuous irradiation and a semi-selective jump-return pulse sequence [12] were used for this purpose. For experiments in D_2O, only continuous irradiation was applied to suppress the water resonance. In all experiments the carrier was placed at the position of the water signal and TPPI [13] was used for signal accumulation in the tl-dimension. After multiplication with shifted square sine bell or Gaussian functions, the data were Fourier transformed in the phase-sensitive mode. Zero-filling was used for improving the digital resolution (8.0 Hz/pt). Additional baseline corrections were performed using polynomial functions of standard Bruker software installed on a Bruker X32 computer.

3. RESULTS AND DISCUSSION

2D-NMR studies of wild-type M13 GVP are rather difficult to perform because the protein is not soluble in millimolar concentrations. This is particularly harmful for spectra recorded of H_2O solutions because it impairs sufficient suppression of the H_2O resonance. Fortunately, mutation of the tyrosine residue at position 41 of wild-type M13 GVP (the protein itself is 87 amino acids long) into a histidyl residue (Y41H) caused a significant improvement in the solubility characteristics of this protein. As a result, measurements of good quality NOESY and TOCSY spectra of H_2O-samples became possible, thus allowing reliable sequential analysis of the protein spectrum. The interpretation of the 2D-NMR data was carried out using standard procedures [14] and, as a result, an unambiguous sequential assignment could be performed for residues 13 through 31. The sequential connectivities observed for this part of the molecule are summarized in fig.2, the resonance positions of the amino acid spin systems are listed in table 1. Frequently, a straightforward sequential analysis procedure is interrupted by a proline residue; this is also true in our situation, namely interruption by proline residue P25. Fortunately, the assignment could be continued by the observation of a sequential dαN connectivity between P25 and Y26 and a dα-(δ,δ') connectivity between K24 and P25. The latter observation is diagnostic for the presence of a normal *trans* conformation for P25. From the available NOESY data it can furthermore be concluded that the sequence encompassing residues 13 through 31 consists of a β-loop composed of an antiparellel β-ladder (residues 13–20 and 24–31) and a turn (residues 21–23). The β-ladder is defined by strong sequential dαN contacts and 3 long range d$\alpha\alpha$ contacts. Because of overlap of α-proton resonances a fourth d$\alpha\alpha$ contact between S17 and S27 could not be observed. Further evidence for the existence of the β-loop is provided by numerous long range non-sequential inter-residue dαN and dNN contacts and by long range NOEs between the aromatic protons fo F13 and Y26 and side-chain protons of the β-ladder. A summary of these long range NOEs is presented in fig.3. The presence of the turn is reflected in the sequential dNN walk from residue 20 to 24. A strong dNN contact between S20 and K24, together with a non-sequential dαN connectivity between N-S20 and α-P25, indicates the presence of a hydrogen bond between the amide of S20 and the carbonyl group of P25. This suggests that the β-ladder structure includes residues 20 and 24.

As pointed out in the beginning of this section, because of its lower solubility, we have not been able to make sequential assignments for wild-type M13 GVP. The spectra obtained for wild-type GVP solutions in D_2O are, however, of sufficient quality to permit a detailed comparison with those obtained for the Y41H mutant. From these comparisons it could unambiguously be concluded that the cross-peaks arising from the β-loop in the Y41H mutant are present in the spectra of the wild-type M13 GVP at almost identical positions. The shifts observed for different peak positions are in

Table 1

Resonance positions of protons in amino acid residues 13–31 of mutant (Y41H) M13 GVP at 298 K and pH 5.2

Residue	Chemical shifts (ppm)			
	$N^{\alpha}H$	$C^{\alpha}H$	$C^{\beta}H$	Others
F13	7.83	5.40	3.26, 3.68	$C^{\delta}H$ 6.93, $C^{\epsilon}H$ 7.27, C^{ζ} 7.27
T14	9.21	4.73	4.23	$C^{\gamma}H$ 1.28
T15	8.87	5.35	3.92	$C^{\gamma}H$ 1.23
R16	9.36	4.75	1.80, 1.88	$C^{\gamma}H$ 1.60
S17	8.56	5.01	3.81	
G18	7.66	3.54, 4.07		
V19	8.41	4.57	1.89	$C^{\gamma}H$ 0.88
S20	9.15	4.70	4.38, 4.08	
R21	9.13	4.26	1.93	$C^{\gamma}H$ 1.76, $C^{\delta}H$ 3.26
Q22	7.93	4.42	1.93, 2.34	$C^{\gamma}H$ 2.50, 2.39
G23	8.19	3.64, 4.20		
K24	7.67	4.92	1.93, 1.77	
P25		5.04	1.95, 2.34	$C^{\gamma}H$ 2.15, 2.04, $C^{\delta}H$ 4.09, 3.78
Y26	8.37	4.94	2.98	$C^{\delta}H$ 6.89, $C^{\epsilon}H$ 6.48
S27	8.39	4.97	3.69, 3.63	
L28	8.94	4.65	1.60	$C^{\gamma}H$ 1.60, $C^{\delta}H$ 0.90, 0.85
N29	9.06	5.47	2.48, 1.62	
E30	8.93	5.54	1.99[a]	
Q31	9.31	4.90	2.03[b]	

[a] A second β-resonance might be present but so far the COSY data do not allow discrimination of this resonance from a γ-resonance

[b] NMR data obtained so far could not provide unambiguous evidence for the non-degeneracy of this resonance

the order of the small shifts one expects from slight variations in sample conditions.

Scrutiny of the NOE data shows that the long range NOE connectivities are, with the same intensity ratios, also present in the NOESY spectrum of the wild-type M13 GVP (fig.4). These results thus demonstrate that in the wild-type GVP the same β-loop structure is present as in the Y41H mutant. This conclusion can even be extended to the overall conformations/structures of the two protein molecules. For instance, most connectivities between other α-proton resonances have almost identical positions in both the mutant and the wild-type GVP spectra or, in two cases, are shifted only to a minor extent (<0.10 ppm). The same goes for the

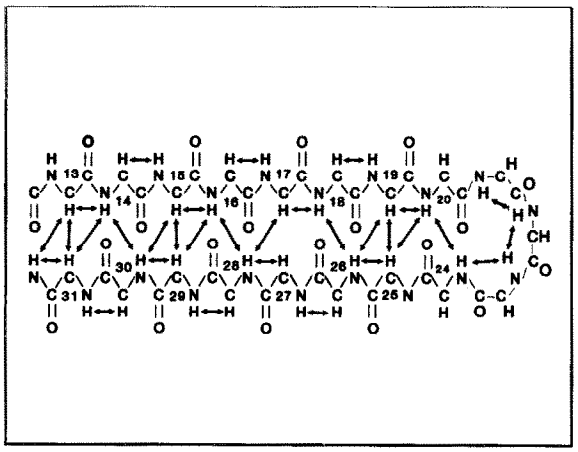

Fig.3. Schematic representation of the structure of the β-loop encompassing residues 13 through 31. Sequential and non-sequential $d\alpha N$ contacts and long-range dNN and $d\alpha\alpha$ contacts are indicated by arrows. They are characteristic of an antiparallel β-sheet. Within the turn only sequential connectivities were observed.

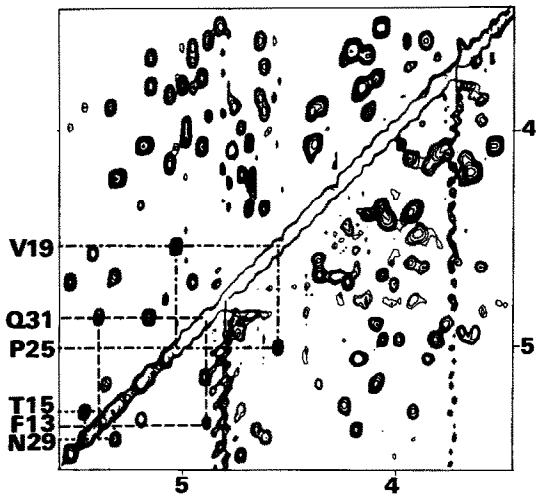

Fig.4. 600 MHz NOESY spectra of mutant (Y41H) GVP (upper diagonal) and wild-type GVP (lower diagonal). Mixing times were 100 ms. Samples contained 4 and 1 mM GVP, respectively. Long-range $d\alpha\alpha$-contacts which define the β-loop (see fig.3) have been indicated with dashed lines.

major part of the NOE connectivities of the aromatic ring proton resonances. Thus, the results suggest that substitution of a histidyl residue for a tyrosine residue at amino acid position 41 of wild-type M13 GVP has only a minor effect upon the overall protein structure. Indeed, in an earlier study the resonance of the 3,5-Tyr protons, which now can be assigned to Y41 because of its absence in the Y41H spectrum, showed a strong photo CIDNP effect [15]. This indicates that Y41 is situated at the outer surface of the protein. In line with this observation a pH-titration of the mutant GVP protein Y41H demonstrates that the pK_a value of the histifdyl residue is about 6.5. This value is characteristic of histidines exposed to the solvent. Thus, substitution of the tyrosine residue for the histidyl residue, which both are situated at the protein surface is expected to cause only minimal changes, if at all, in the overall protein structure.

Previous NMR studies have demonstrated that in the binding of M13 GVP to ssDNA both aromatic as well as charged residues are involved [16,17]. With the aid of NMR-studies on other GVP mutants (data not shown) and the above-mentioned sequential assignments some of these residues can now be identified as Y26 and R21, respectively. Other NMR studies, both with spin-labeled and non-labeled oligonucleotides, have furthermore indicated that also the amino acid residues, R16, G18, S20 [18], K24 [17] and L28 [19], form part of the DNA binding domain. Comparison of these data with the results of the NMR studies performed on IKe GVP [3] shows that the same (or similar) amino acids are also involved in the binding of the latter protein to ssDNA.

Putting all the results together, we are forced to conclude that the observed structure of the β-loop in wild-type as well as mutant (Y41H) M13 GVP is analogous to the β-loop structure present in IKe GVP. The structure of this DNA binding wing deviates, however, significantly from the structure proposed for this stretch of amino acids on the basis of X-ray crystallographic studies of M13 GVP [4]. Comparison of these structures reveals that the β-loop structure found in solution is more regular than the one found in the crystal. Furthermore in the solution structure of the DNA binding wing the β-sheet is shifted 4 amino acid residues with respect to that of the crystal structure: specifically, in the crystal structure residue Y26 is positioned at the tip of a β-loop structure, while in the solution structure this residue forms part of the antiparallel β-sheet (fig.3). In a previous paper, anomalous observations towards crystallographic data have already been reported [20]. In this paper, however, the NMR results provide unambiguous evidence that the conformation of M13 GVP in solution differs significantly from that in the crystal.

Acknowledgements: The authors wish to thank Mrs C. Prinse and Mrs A. Jonker for their efforts in isolating and purifying wild type and mutant GVP.

REFERENCES

[1] Model, P. and Russel, M. (1988) in: The Bacteriophages (R. Calendar, ed.) pp. 2099–2107, Plenum, New York.
[2] de Jong, E.A.M., van Duynhoven, J.P.M., Harmsen, B.J.M., Konings, R.N.H. and Hilbers, C.W. (1989) J. Mol. Biol. 206, 119–132.
[3] de Jong, E.A.M., van Duynhoven, J.P.M., Harmsen, B.J.M., Konings, R.N.H. and Hilbers, C.W. (1989) J. Mol. Biol. 206, 133–152.
[4] Brayer, G.D. and McPherson, A. (1983) J. Mol. Biol. 169, 565–596.
[5] Garssen G.J., Hilbers, C.W., Schoenmakers, J.G.G. and van Boom, J.H. (1977) Eur. J. Biochem. 81, 453–463.
[6] Stassen, A.P.M., Deursen, J., Zaman, G. and Konings, R.N.H., in preparation.
[7] Bax, A. and Davis, D.G. (1985) J. Magn. Res. 65, 355–366.
[8] Griesinger, C., Otting, K., Wüthrich, K. and Ernst, R.R. (1988) J. Am. Chem. Soc. 110, 7870–7872.
[9] Rance, M., Sorensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) Biochem. Biophys. Res. Commun. 117, 479–485.
[10] Jeener, J., Meier, B.H., Bachman, P. and Ernst, R.R. (1979) J. Chem. Phys. 71, 4546–4553.
[11] Plateau, P. and Geuron, M. (1982) J. Am. Chem. Soc. 104, 7310–7311.
[12] Bax. A., Sklenar, V., Clore, G.M. and Gronenborn, A.M. (1987) J. Am. Chem. Soc. 109, 6511–6513.
[13] Marion, D. and Wütrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974.
[14] Wüthrich, K. (1986) in: NMR of Proteins and Nucleic Acids, Ch. 7–10, Wiley & Sons, New York.
[15] Garssen, G.J., Kaptein, R., Schoenmakers, J.G.G. and Hilbers C.W. (1978) Proc. Natl. Acad. Sci. USA 75, 5281–5285.
[16] Alma, N.C.M., Harmsen, B.J.M., Hilbers, C.W., van der Marel, G. and van Boom, J.H. (1981) FEBS Lett. 135, 15–20.
[17] Dick, L.R., Sherry, A.D., Newkirk, M.M. and Gray, D.M. (1988) J. Biol. Chem. 263, 18864–18872.
[18] Folkers, P.J.M., Duynhoven, J.P.M., Harmsen, B.J.M., Konings, R.N.H. and Hilbers, C.W., in preparation.
[19] King, G.C. and Coleman, J.E., (1987) Biochemistry 26, 2929–2937.
[20] Gray, D.M., Sherry, A.D., Teherani, J. and Kansky, J.W. (1984) J. Biomol. Struct. Dynam. 2, 77–91.