Twin hydroxymethyluracil-A base pair steps define the binding site for the DNA-bending protein TF1

Anne Grove  
*University of California, San Diego*

Marxa L. Figueiredo  
*University of California, San Diego*

Aldo Galeone  
*Università degli Studi di Napoli Federico II*

Luciano Mayol  
*Università degli Studi di Napoli Federico II*

E. Peter Geiduschek  
*University of California, San Diego*

Follow this and additional works at: [https://repository.lsu.edu/biosci_pubs](https://repository.lsu.edu/biosci_pubs)

**Recommended Citation**

Grove, A., Figueiredo, M., Galeone, A., Mayol, L., & Geiduschek, E. (1997). Twin hydroxymethyluracil-A base pair steps define the binding site for the DNA-bending protein TF1. *Journal of Biological Chemistry, 272*(20), 13084-13087. [https://doi.org/10.1074/jbc.272.20.13084](https://doi.org/10.1074/jbc.272.20.13084)

This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Scholarly Repository. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Scholarly Repository. For more information, please contact [ir@lsu.edu](mailto:ir@lsu.edu).
Twin Hydroxymethyluracil-A Base Pair Steps Define the Binding Site for the DNA-bending Protein TF1*

(Received for publication, February 12, 1997)

Anne Grove‡, Marxa L. Figueiredo‡§, Aldo Galeone*, Luciano Mayol§, and E. Peter Geiduschek‡

From the Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093-0634 and Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli Federico II, Via Domenico Montesano, 49 80131 Napoli, Italy

The DNA-bending protein TF1 is the Bacillus subtilis bacteriophage SP01-encoded homolog of the bacterial HU proteins and the Escherichia coli integration host factor. We recently proposed that TF1, which binds with high affinity ($K_d$ was $\sim 3 \text{nM}$) to preferred sites within the hydroxymethyluracil (hmU)-containing phage genome, identifies its binding sites based on sequence-dependent DNA flexibility. Here, we show that two hmU-A base pair steps coinciding with two previously proposed sites of DNA distortion are critical for complex formation. The affinity of TF1 is reduced 10-fold when both of these hmU-A base pair steps are replaced with A-hmU, G-C, or C-G steps; only modest changes in affinity result when substitutions are made at other base pairs of the TF1 binding site. Replacement of all hmU residues with thymine decreases the affinity of TF1 greatly; remarkably, the high affinity is restored when the two hmU-A base pair steps corresponding to previously suggested sites of distortion are reintroduced into otherwise T-containing DNA. T-DNA constructs with 3-base bulges spaced apart by 9 base pairs of duplex also generate nM affinity of TF1. We suggest that twin hmU-A base pair steps located at the proposed sites of distortion are key to target site selection by TF1 and that recognition is based largely, if not entirely, on sequence-dependent DNA flexibility.

The genome of the Bacillus subtilis bacteriophage SP01 contains 5-hydroxymethyluracil (hmU) in place of thymine, and discrimination between T- and hmU-containing DNA is important during phage multiplication (1, 2). For one phase-encoded protein, the DNA-binding and -bending protein TF1, this discrimination is essentially absolute: the high affinity of TF1 ($K_d$ order of magnitude nM) for preferred sites within the hmU-containing phage genome is greatly reduced for the corresponding T-containing DNA (3, 4). We recently showed that the affinity of TF1 for hmU-DNA is matched in T-containing DNA constructs with bulge constructs were performed at 4 °C to prevent dissociation of duplexes. After electrophoresis, gels were dried and protein-DNA complexes were quantified using phosphorimaging. DNA migration between bands corresponding to free and complexed DNA was considered as free DNA (5) in calculating equilibrium dissociation constants ($K_d$) (5). Values of $K_d$ are reported as the average of at least three experiments ± S.E. (for closely related DNA constructs yielding suboptimal complex formation, triplicate determinations were deemed unnecessary).

RESULTS

The sequence of a 37-mer DNA construct corresponding to a preferred binding site within the phage SP01 genome (4) is shown in Fig. 1A. Because of the ability of TF1 to form nested complexes on longer DNA fragments, the length of the DNA constructs was chosen to accommodate only one TF1 dimer (5). The affinity of TF1 for the 37-mer hmU-containing duplex ($K_d$ was $\sim 3 \text{nM}$; Fig. 1A) was reduced 10-fold upon replacement of the two T-hmU-A base pair steps corresponding to the potential sites of distortion with A-hmU, G-C, or C-G steps (Figs. 1 and 2, panels B). Altering the sequence at the center of the binding site (Figs. 1 and 2, panels C) or on the flanks, distal to the proposed sites of bending (Figs. 1 and 2, panels D), did not significantly affect complex formation. The observed sequence specificity at proposed sites of distortion is consistent with a direct participation of hmU-A base pair steps in complex formation and DNA bending.

To evaluate the contribution of hmU residues to the DNA binding affinity of TF1, DNA constructs were designed in which central hmU-containing blocks of 13 or 11 bp, encompassing the proposed sites of DNA distortion (Fig. 3, constructs Th2 and Th3), or of 9 or 7 bp, excluding hmU-A steps that are candidate
Twin hmU-A Base Pair Steps Define TF1 Binding Site

FIG. 1. Sequences of 37-mer oligonucleotides. A, the sequence of the top strand corresponding to a preferred binding site within the SPO1 genome is shown, with hmU-A base pair steps that correspond to proposed sites of distortion in boldface. B, a 10-fold reduction in affinity is associated with replacement of target hmU-A steps, C and D, sequences of oligonucleotides with substitutions between (C) or distal to (D) hmU-A base pair steps (substitutions are underlined). Substitutions on the flanks target additional hmU-A base pair steps; additional substitutions were required to prevent introduction of new hmU-A steps. Top strands shown were annealed to complementary hmU-containing bottom strands to generate perfect duplexes. Dissociation constants, \( K_d \) nM, are shown at the right, wild-type.

| A | HA(wt) | 5’- CCAAGG(G/T)CAAC(C/T)GAGAAGGAACGMC-3’ | 3.1±0.6 |
| B | AH | 5’- CCAAGG(G/T)CAAC(C/T)GAGAAGGAACGMC-3’ | 31 |
| C | GC-rich | 5’- CCAAGG(G/T)CAAC(C/T)GAGAAGGAACGMC-3’ | 3.7±0.3 |
| D | wt-ctr | 5’- CCAAGG(G/T)CAAC(C/T)GAGAAGGAACGMC-3’ | 3.7±0.3 |


FIG. 2. Electrophoretic analysis of hmU-DNA titrated with TF1. The affinity of TF1 for the 37-mer duplex corresponding to a preferred binding site (panel A) is reduced upon replacement of hmU-A base pair steps with A-hmU (panel B). Altering the central base pairs (GC-rich; panel C) or the sequence distal to proposed sites of DNA distortion (wild-type (wt) + GC; panel D) has little effect on complex formation. Protein concentrations indicated below panel D are identical for all panels.

\[ K_d (\text{nM}) \]

| T/T | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 3.1±0.6 |
| h/h | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 26 |
| Th1 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 6.1±1.4 |
| Th2 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 4.5±1.1 |
| Th3 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 15 |
| Th4 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 31 |
| Th5 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 6.4±1.3 |
| Th6 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 3.6±0.5 |
| Th7 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 29 |
| Th8 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 40 |
| Th9 | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 3.8±0.9 |
| Th/T | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 31 |
| Th/Th | 5’- CCTAGGCTACACTCTTGGGAGAATTCGCTC-3’ | 27 |

FIG. 3. DNA duplexes with T or hmU content. Blocks of hmU content are underlined. DNA constructs are named at the left, and dissociation constants, \( K_d \) nM, are shown at the right. The symbol \( \gg \) denotes complex dissociation during electrophoresis that is too rapid to allow a reliable determination.

The high affinity of TF1 for hmU-DNA (\( K_d \approx 3 \text{nM} \)) was reproduced in T-DNA that contains two appropriately positioned flexible sites generated by sets of tandem mismatches (Ref. 5 and Fig. 5A). Improved complex formation relative to perfect duplex T-DNA would also be expected for T-DNA constructs in which the energetic cost of bending is lessened through the incorporation of bulge-loops, i.e. with nucleotides that are formally unopposed on the complementary strand (Fig. 5). Constructs with sets of 1- or 2-nt bulges combined with one mismatched base pair positioned 9 bp apart indeed generated high affinity for TF1 (Fig. 5B); note that sets of 2-nt loops (one mismatched base pair in the suboptimal binding compared with sets of 4-nt loops (two consecutive mismatches)). The high affinity of TF1 was fully restored for DNA constructs with sets of 2- or 3-nt bulges with suboptimal binding for DNA with a pair of 1-nt mismatches (Ref. 5 and Fig. 5C). Replacing the T-containing top strand of bulge constructs with the corresponding hmU-containing strand did not affect complex formation (data not shown). Bulge-loops are considered to generate directed kinks in the helix axis (8, 9). Consistent with the presence of such kinks, we observed a ~5 nM). Remarkably, TF1 was found to have essentially undiminished affinity for DNA in which the two hmU-A base pair steps are completely surrounded by T-containing DNA (\( K_d = 4 \text{nM} \); Fig. 3, construct Th7/6, and Fig. 4, A and B). We note also that the affinity of TF1 for hmU-DNA and for the Th7/Th6 construct was the same at room temperature and at 4 °C (Figs. 3 and 5 and data not shown). The T–10-fold reduction in affinity was observed if only one strand contained the requisite hmU-A steps (Fig. 3, constructs Th7/Th6 and T/Th7, and Fig. 4C). Our interpretation of these observations is that hmU content affords an increased flexibility that is recognized by TF1 and that hmU-A base pair steps symmetrically disposed about the center of the binding site are targets for specific distortion in the protein-DNA complex.
reduction in electrophoretic mobility of bulged DNA constructs, the magnitude of which depended on the size of the bulge (data not shown). The mobility of protein-DNA complexes was similar irrespective of the nature of the DNA construct.

The TF1 double mutant protein TF1 (E15G/T32I) binds DNA with ~40-fold higher affinity than does wild-type TF1 (10) due to interactions outside the central 25-bp region of the binding site (5). The presence of bulge-loops in T-DNA affected the affinity of TF1 (E15G/T32I) and wild-type TF1 in similar ways, consistent with unaltered interactions for the mutant protein with a central DNA segment that includes proposed sites of bending (data not shown).

The structure of TF1 reveals the presence of flexible β-ribbon arms that are directly involved in DNA binding (11), as does the structure of the closely related Hu protein (12, 13) and the recently reported structure of integration host factor (IHF) in a complex with DNA (14). At the tip of the DNA-binding arms, TF1 contains a Phe residue at position 61 in place of the Arg found in all other members of this protein family. TF1(F61R) has reduced preference for hmU-DNA over T-DNA (15) but was seen to retain high affinity for hmU-DNA (Kd = 9.6 ± 0.7 nM). Substituting Phe-61 with Tyr had little effect on affinity (Kd = 6.8 ± 0.9 nM). Replacing Phe-61 with a polar, uncharged residue (Ser or Gln) significantly reduces both DNA binding affinity and site-selectivity (15). Notably, the affinity for hmU-DNA of both TF1(F61S) (Kd = 236 ± 19 nM) and TF1(F61Q) (Kd = 685 ± 48 nM) increased more than 5-fold for T-DNA with 2 3-base bulges (Kd = 36 ± 2 and 90 ± 7 nM, respectively; Fig. 5). This is in contrast to wild-type TF1 that bound hmU-DNA and bulge-DNA constructs (1+2, 1+3, 0+2, and 0+3) with approximately the same affinity (Fig. 5). T-DNA with sets of 4-nt loops, which also reproduces the affinity of wild-type TF1 for hmU-DNA (Fig. 5A and Ref. 5), did not significantly enhance complex formation by TF1(F61S) (Kd = ~200 nM) or TF1(F61Q) (Kd = ~570 nM). Similarly, we observed with hmU-DNA. Apparently hmU content and flexible loops in T-DNA contribute similarly to complex formation for these TF1 mutant proteins, which are deficient in binding and bending of hmU-DNA (15, 16), whereas DNA with a static kink generates an increase in affinity.

2 M. Silva, L. Pasternack, and D. Kearns, submitted for publication.

DISCUSSION

Our previous conjecture that TF1 principally identifies its preferred binding sites through recognition of sequence-dependent DNA flexure was based on an analysis of affinities for DNA constructs with sets of 4-nt loops (5). For this mode of target site selection, the prediction would be for the sequence at the sites of distortion to significantly affect complex formation. This prediction appears to be borne out both in terms of preferred base pair steps in the context of hmU-DNA (Figs. 1 and 2) but particularly by the significant effect of hmU-A base pair steps in a T-containing surrounding sequence (Figs. 3 and 4). We propose that hmU content confers a site-specific flexibility on the DNA duplex that is recognized by TF1. hmU-A base pair steps that are situated at presumptive sites of DNA bending make the principal contributions to complex formation with TF1. hmU residues at other positions that generate measurable, yet suboptimal affinities, presumably do so either due to longer range effects on DNA flexure or because of alternative placement of TF1. An effect of hmU residues on the deformability of the double helix is corroborated by the increased affinity observed for other members of the family of type II DNA-binding proteins, Hu and IHF, and for the unrelated eukaryotic protein, HMGI (17). Changes in complex formation by both major and minor groove-binding proteins and antibiotics induced by introduction of other modified bases have also been interpreted in terms of effects on DNA deformability by exocyclic substituents (18–20).

Inferences about increased flexure of DNA containing modified bases stress the need for analysis of the associated structural and conformational energetics. Structural nuclear magnetic resonance analysis of an A-hmU base pair flanked by G-C nearest neighbors does not significantly reduce DNA binding affinity, and site-selectivity (15). Notably, the affinity for hmU-DNA of both TF1(F61S) (Kd = 236 ± 19 nM) and TF1(F61Q) (Kd = 685 ± 48 nM) increased more than 5-fold for T-DNA with 2 3-base bulges (Kd = 36 ± 2 and 90 ± 7 nM, respectively; Fig. 5). This is in contrast to wild-type TF1 that bound hmU-DNA and bulge-DNA constructs (1+2, 1+3, 0+2, and 0+3) with approximately the same affinity (Fig. 5). T-DNA with sets of 4-nt loops, which also reproduces the affinity of wild-type TF1 for hmU-DNA (Fig. 5A and Ref. 5), did not significantly enhance complex formation by TF1(F61S) (Kd = ~200 nM) or TF1(F61Q) (Kd = ~570 nM). Similarly, we observed with hmU-DNA. Apparently hmU content and flexible loops in T-DNA contribute similarly to complex formation for these TF1 mutant proteins, which are deficient in binding and bending of hmU-DNA (15, 16), whereas DNA with a static kink generates an increase in affinity.

Inferences about increased flexure of DNA containing modified bases stress the need for analysis of the associated structural and conformational energetics. Structural nuclear magnetic resonance analysis of an A-hmU base pair flanked by G-C nearest neighbors does not significantly reduce DNA binding affinity, and site-selectivity (15). Notably, the affinity for hmU-DNA of both TF1(F61S) (Kd = 236 ± 19 nM) and TF1(F61Q) (Kd = 685 ± 48 nM) increased more than 5-fold for T-DNA with 2 3-base bulges (Kd = 36 ± 2 and 90 ± 7 nM, respectively; Fig. 5). This is in contrast to wild-type TF1 that bound hmU-DNA and bulge-DNA constructs (1+2, 1+3, 0+2, and 0+3) with approximately the same affinity (Fig. 5). T-DNA with sets of 4-nt loops, which also reproduces the affinity of wild-type TF1 for hmU-DNA (Fig. 5A and Ref. 5), did not significantly enhance complex formation by TF1(F61S) (Kd = ~200 nM) or TF1(F61Q) (Kd = ~570 nM) beyond that observed with hmU-DNA. Apparently hmU content and flexible loops in T-DNA contribute similarly to complex formation for these TF1 mutant proteins, which are deficient in binding and bending of hmU-DNA (15, 16), whereas DNA with a static kink generates an increase in affinity.

For a protein that bends DNA through a large angle (16), it is not surprising to find a greater affinity for DNA whose axis is kinked by unpaired bases (Fig. 5). Indeed, similar results have been previously reported for Hu (27). We note that T-DNA constructs with pairs of suitably placed 4-nt loops or 3-nt bulge-loops generate the same affinity for wild-type TF1. However, different structures and energetics of deformation are indicated by the significantly higher affinity for bulge constructs of the TF1 mutant proteins TF1(F61S) and TF1(F61Q), which are impaired in DNA binding and bending. These observations are consistent with a static kink introduced by bulge-loops compared with a looser more flexible structure imposed by symmetric loops (8, 9, 24, 28); we suggest that pre-bent DNA bulge-loops also permit complex formation with certain TF1 mutant proteins, whereas binding and bending of DNA with sets of 4-nt loops (or hmU content) require a protein-mediated introduction and stabilization of DNA kinks.

Replacement of Phe-61 at the tip of DNA-binding arms with a polar, uncharged residue may render TF1 mutant proteins either incapable of producing specific DNA kinks, a process which would be aided locally by the presence of DNA bulge-loops, or deficient in securing the resulting DNA conformation,
demands on which would be lessened in DNA with a predisposed orientation. If TF1, like IHF, employs a totally conserved Pro residue to intercalate between specific base pairs of the binding site (14), then the deficiency of TF1(F61S) and TF1(F61Q) in DNA binding and bending may reside in their inability to stabilize a severely bent DNA conformation. Phe-61 contributes to the striking preference of TF1 for hmU-DNA over T-DNA (15). TF1(F61R) also has high affinity for hmU-DNA but a reduced preference for hmU-DNA over T-DNA (15). In addition, IHF (which has higher affinity for a specific binding site in hmU-DNA compared with T-DNA (17)) relies heavily on indirect readout to recognize its binding sites (14). If TF1 and IHF engage their DNA targets in an isomorphous fashion, the preference for hmU-DNA characteristic of TF1 may not be due to specific interactions between Phe-61 and hmU residues of the TF1 binding site. Rather, the inability of TF1 to stabilize a T-DNA duplex that is severely kinked through intercalation of Pro residues may be alleviated by the more pliable hmU-DNA. On the other hand, the possibility that TF1 does use Phe-61 directly for inducing DNA kinks at two hmU-A base pair steps remains to be established or disproved by direct structure determination.

REFERENCES
1. Lee, G., Hannett, N. M., Korman, A., and Pero, J. (1980) J. Mol. Biol. 139, 497–422
2. Choy, H. A., Romeo, J. M., and Geiduschek, E. P. (1986) J. Mol. Biol. 191, 59–73
3. Johnson, G. G., and Geiduschek, E. P. (1977) Biochemistry 16, 1473–1485
4. Greene, J. R., and Geiduschek, E. P. (1985) EMBO J. 4, 1345–1349
5. Grove, A., Galeone, A., Mayol, L., and Geiduschek, E. P. (1996) J. Mol. Biol. 260, 196–206
6. Hard, T., Hsu, V., Geiduschek, E. P., Appelt, K., and Kearns, D. (1989) Biochemistry 28, 396–406
7. Conte, M. R., Galeone, A., Avizonis, D., Hsu, V. L., Mayol, L., and Kearns, D. R. (1992) Bioorg. Med. Chem. Lett. 2, 79–82
8. Galeke, C., Murchie, A. I. H., Lilley, D. M. J., and Clegg, R. M. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 11660–11664
9. Lilley, D. M. J. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7140–7142
10. Anders, L., Spangler, C. J., Galeone, A., Mayol, L., and Geiduschek, E. P. (1994) J. Mol. Biol. 236, 139–150
11. Jia, X., Grove, A., Ivancic, M., Hsu, V. L., Geiduschek, E. P., and Kearns, D. R. (1996) J. Mol. Biol. 263, 259–268
12. White, S. W., Appelt, K., Wilson, K. S., and Tanaka, I. (1989) Proteins Struct. Funct. Genet. 5, 281–288
13. Vis, H., Mariani, M., Vorgias, C. E., Wilson, K. S., Kaptein, R., and Boelens, R. (1995) J. Mol. Biol. 254, 692–703
14. Rice, P. A., Yang, S.-w., Mizuuchi, K., and Nash, H. (1996) Cell 84, 1295–1306
15. Sayre, M. H., and Geiduschek, E. P. (1990) J. Mol. Biol. 216, 819–833
16. Schneider, G. J., Sayre, M. H., and Geiduschek, E. P. (1991) J. Mol. Biol. 221, 777–784
17. Grove, A., Galeone, A., Mayol, L., and Geiduschek, E. P. (1996) J. Mol. Biol. 260, 120–125
18. Diekmann, S., von Kitzing, E., McLaughlin, L., Ott, J., and Eckstein, F. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 8257–8261
19. Bailly, C., Payet, D., Travers, A. A., and Waring, M. J. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 13623–13628
20. Mellac, S., Fazakerley, G. V., and Sowers, L. C. (1993) Biochemistry 32, 7779–7786
21. Pasternack, L. B., Bramham, J., Mayol, L., Galeone, A., Jia, X., and Kearns, D. R. (1996) Nucleic Acids Res. 24, 2740–2747
22. Millar, D. P., Robbins, R. J., and Zewail, A. H. (1982) J. Chem. Phys. 76, 2080–2094
23. Kahn, J. D., Yun, E., and Crothers, D. M. (1992) Nature 359, 163–166
24. Hunter, C. A. (1993) J. Mol. Biol. 239, 1025–1054
25. Wolfe, A. P., and Drew, H. R. (1997) Frontiers in Molecular Biology, IRL Press, Oxford, in press
26. Bhattacharyya, A., and Lilley, D. M. J. (1989) Nucleic Acids Res. 17, 6821–6840

Fig. 5. The nM affinity of TF1 for hmU-DNA is reproduced by T-DNA with base bulges separated by 9 bp of duplex. The affinity of TF1 for perfect duplex hmU-DNA is reproduced by T-DNA with a set of 4-nl loops (A). 1- or 2-base bulges (lowercase c, underlined) in combination with a mismatch reproduce the nM affinity for TF1 (B), as do 2- or 3-base bulges (C). >>, the same as in the legend to Fig. 3. All experiments were carried out at 4 °C.