A* protein of bacteriophage φX174 carries an oligonucleotide which it can transfer to the 3′-OH of a DNA chain

A.D.M. van Mansfeld, H.A.A.M. van Teeffelen, J. Zandberg, P.D. Baas, H.S. Jansz, G.H. Veeneman+ and J.H. van Boom+

Institute of Molecular Biology and Laboratory for Physiological Chemistry, State University of Utrecht, Padualaan 8, 3584 CH Utrecht and +*Department of Organic Chemistry, State University of Leiden, 2300 RA Leiden, The Netherlands*

Received 28 September 1982

The bacteriophage φX174 gene A encodes two proteins: gene A protein and A* protein. Purified A* protein acts as a single-stranded, DNA-specific endonuclease which remains covalently attached to the 5′-end of the cleavage site. Incubation of A* protein with the synthetic heptamer CAACTTG or with oligonucleotides which yield this heptamer after cleavage with the A* protein yields oligonucleotides with the sequences CAACTTGAG, CAACTTGAGG and CAACTTGAGGA. This indicates that A* protein carries an oligonucleotide with the sequence −AG, −AGG or −AGGA. The oligonucleotide can be transferred to the 3′-end of the heptamer CAACTTG. This suggests that A* protein reacts with a specific DNA sequence in the infected cell.

Bacteriophage φX174 *A* protein* *DNA–protein complex* *Synthetic oligonucleotide* *DNA-ligating activity*

1. INTRODUCTION

The DNA replication of bacteriophage φX174 depends on the replication machinery of the host and only one viral gene, gene A. Gene A encodes two proteins, gene A protein and A* protein [1]. Gene A protein is a multifunctional protein in the DNA replication of φX [2]. Among others it triggers the rolling circle replication by cleavage of the viral strand of the supercoiled double-stranded replicative form DNA (RFI) at a specific site, the origin [3]. At the end of one replication round gene A protein cleaves and ligates the viral DNA strand such that a circular unit length molecule is produced.

The role of the second protein specified by gene A, the A* protein, is so far unknown. It is synthesized from an internal translational start within gene A, in the same reading frame as the gene A protein. It thus lacks the N-terminal part (~⅓rd of the polypeptide chain) of the gene A protein. However, experiments with purified A* protein have shown that it has retained some of the enzymatic activities of the gene A protein (see [4,5]). A* protein cleaves, like gene A protein, single-stranded φX viral DNA at the origin, but also at other sites [6]. A* protein and gene A protein remain covalently attached to the 5′-end of the cleavage site. A* protein can ligate DNA [7,8].

The synthetic hexadecadeoxyribonucleotide CAACTTGATATTAATA, which corresponds to nucleotides 4299–4314 in the φX DNA sequence [9], is cleaved next to the G-residue by gene A protein. This site corresponds to the cleavage site of gene A protein in RF DNA [3]. A* protein cleaves the hexadecamer at the same site but additional products are observed [10]. Here, we report that the additional products are the result of a transfer of oligonucleotides with the sequence AG, AGG or AGGA, respectively, to the heptamer CAACTTG. The oligonucleotides occur covalently linked to the A* protein in the A* protein preparation. They are presumably remnants of a DNA sequence which is cleaved intracellularly by the A* protein.

Published by Elsevier Biomedical Press
00145793/82/0000—0000/$2.75

2. MATERIALS AND METHODS

2.1. *Enzymes*

ϕX A* protein was purified from ϕX174 am3-infected *Escherichia coli* C cells as in [7]. Snake venom phosphodiesterase (SVD) and bacterial alkaline phosphatase (BAPF) were from Worthington (Freehold NJ), proteinase K from Merck (Darmstadt), T4-induced polynucleotide kinase from Boehringer (Mannheim), and terminal transferase from Enzo Biochem. Inc. (New York NY) ^{32}P-Labelled A* protein was purified following the same procedure as for A* protein from a 0.5 l culture of *E. coli* C in dephosphorylated 3XD medium [11] that had been infected with ϕX am3 in the presence of 10 mCi [^{32}P]orthophosphate (Amersham).

2.2. *Protein electrophoresis*

The final step in the purification of ^{32}P-labelled A* protein was a gradient elution from a heparin–Sepharose column [7]. Fourty fractions of 550 μl were collected. The protein of 200 μl of one fraction was precipitated with trichloroacetic acid, dissolved and subjected to electrophoresis as in [12] on 12.5% polyacrylamide gel. After electrophoresis the gel was stained with silver as in [13]. Then the gel was dried and radioactive bands were visualized by autoradiography. Treatment with alkali was performed as follows. Protein was precipitated with trichloroacetic acid, the precipitate was washed twice with 400 μl ethanol/ether (1:1) and once with 100 μl ethanol/ether (1:3). The precipitate was dried, dissolved in 50 μl 0.1 M KOH and incubated for 3 h at 37°C. Then the protein was analyzed on a gel as above.

2.3. *The synthetic hexadecadeoxyribonucleotide*

The characterization of the synthetic hexadecadeoxyribonucleotide (hexadecamer) CAACTTGATATTAATA was described in [10]. It was labelled at the 5'-end using [γ-^{32}P]ATP (2000–3000 Ci/mmol; Amersham) and T4 polynucleotide kinase following the protocol in [14]. The hexadecamer and ATP were separated by chromatography over a Sephadex G-50 column.

2.4 *Incubation with A* protein*

In a standard assay 1.5 ng A* protein is incubated with 0.015 pmol labelled hexadecamer in 10 mM Tris–HCl (pH 7.6), 1 mM EDTA, 5 mM $MgCl_2$, 5 mM DTT, 150 mM NaCl, 2% glycerol in 16 μl for 30 min at 37°C. The reaction is stopped by the addition of 1.5 μl 100 mM EDTA (pH 8.0) and incubation for 30 min at 37°C with 1 μl of a solution of 5 mg proteinase K/ml 1 mM Tris (pH 7.5), 0.1 mM EDTA that had been preincubated for 30 min at 37°C. The samples were analyzed on a 25% polyacrylamide gel as in [10].

2.5. *Isolation of the products after incubation of hexadecamer with A* protein*

5'-Labelled hexadecamer (1.5 pmol) was incubated with 300 ng A* protein in the above buffer, in 400 μl for 30 min at 37°C. Then 40 μl 100 mM EDTA (pH 8.0) and 25 μl of 5 mg proteinase K/ml, 1 mM Tris–HCl (pH 7.6) and 0.1 mM EDTA that had been pre-incubated for 30 min at 37°C were added. After 30 min at 37°C this mixture was extracted with an equal volume of phenol. After separation the phenol phase was extracted with 50 μl H_2O and the combined water phases were extracted with ether. The sample was diluted with H_2O to 1.5 ml and applied to a small DEAE-cellulose (Whatman DE52) column that had been packed in a blue Eppendorf tip and had been washed twice with 1 ml 2 M NH_4HCO_3 and 3 times with 1 ml 0.1 M NH_4HCO_3. The radioactive oligonucleotides bound to the DEAE-cellulose. The column was washed 3 times with 0.5 ml 0.1 M NH_4HCO_3 and the radioactive material was eluted with 4 times 0.1 ml 2 M NH_4HCO_3. The samples were dried at 37°C, the residue was dissolved twice in 50 μl H_2O and dried again. Then the samples were dissolved in 20 μl H_2O, mixed with 20 μl 98% formamide, 10 mM NaOH, 0.05% bromophenol blue, 0.05% xylenecyanol F, heated and applied to a 25% polyacrylamide gel as above.

After electrophoresis the radioactive bands were excised and the oligonucleotides were eluted as in [14]. The eluates (1 ml) were diluted to 5 ml with H_2O applied to a small DEAE column, washed and eluted. Water and the volatile salt were removed as described above. Finally, the samples were dissolved in 50 μl H_2O.

2.6. *Sequence analysis of the oligonucleotides*

The oligonucleotides obtained by the above procedure were digested partially by incubation at 37°C with 200 pg SVD in 10 mM Tris–HCl

(pH 7.6), 1 mM EDTA, 5 mM $MgCl_2$ in 30 μl. At 1, 10 and 30 min samples were taken and added to phenol to stop the reaction. Suitable samples were mixed and subjected to two-dimensional separation; i.e., high-voltage electrophoresis at pH 3.5 on a cellulose acetate strip and after transfer of the oligonucleotides to a DEAE thin-layer, chromatography with a 3% 'homomix' [15].

The radioactive products were detected by autoradiography. One of the oligonucleotides was also subjected to the chemical method for sequencing DNA [14], which was carried out exactly as in [10].

3. RESULTS

Fig. 1 shows the reaction products obtained after incubation of the 5'-labelled hexadecamer CAACTTGATATTAATA with A* protein. Besides the heptamer CAACTTG some other products are present. Based on its mobility in the urea-containing polyacrylamide gel the predominant extra product has an apparent length of 10 nucleotides. It was therefore called the 10* product. Two other products which are present in far lower amounts migrated as a nonamer and an undecamer and were called 9* and 11*, respectively. The 9*, 10* and 11* products could be dephosphorylated by alkaline phosphatase and elongated by terminal transferase with an efficiency comparable to the 5'-labelled hexadecamer (not shown). This and the fact that they can be digested with SVD (see below) indicates that the 9*, 10* and 11* products have free 5'-terminal phosphate and free 3'-hydroxyl groups.

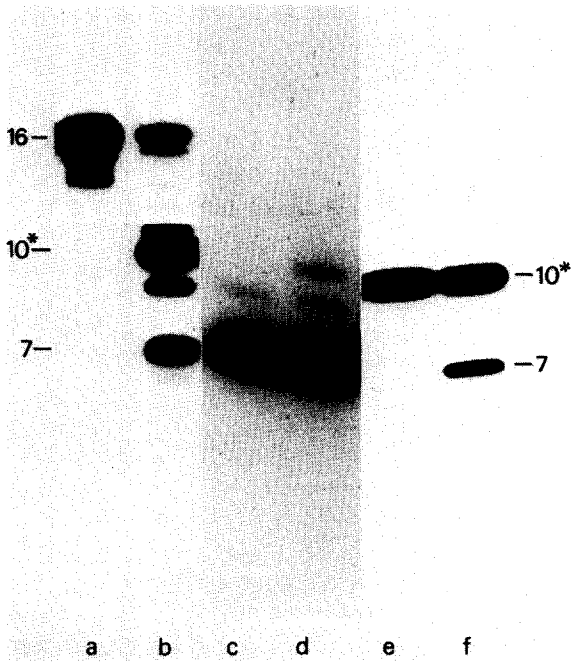

Fig. 1. Incubations of A* protein with the hexadecamer, the heptamer and 10* product, autoradiograph after gel electrophoresis: hexadecamer without (a) and with A* protein (b); heptamer without (c) and with A* protein (d); the 10* product without (e) and with A* protein (f).

The 9*, 10* and 11* products were isolated and nucleotide sequences were determined. This was done by partial digestion from the 3'-end with SVD and separation of the digestion products in the two-dimensional system following the 'mobility-shift' method (fig. 2). The sequence of the 10* product was confirmed by the procedure in [14].

The sequences are compared with the sequence of the starting hexadecamer in table 1.

Table 1

The sequences of the oligonucleotides

Hexadecamer	CAACTTGATATTAATA
Heptamer	CAACTTG
10*, major extra product	CAACTTGAGG
9*, minor extra product	CAACTTGAG
11*, minor extra product	CAACTTGAGGA

Surprising are the sequences –AG, –AGG and –AGGA at the 3'-ends of the 9*, 10* and 11* products, respectively. The starting hexadecamer does not contain the sequence –AG, –AGG or –AGGA, so the oligonucleotide material with these sequences must have been added with another component of the incubation mixture; i.e., the A* protein preparation. The final steps in the purification procedure of A* protein are a DNA–cellulose column, followed by a dialysis and an heparin–Sepharose column [7]. If free oligonucleotide material was present after the DNA-cellulose column this would probably have been lost by the dialysis and subsequent column chromatography. As a consequence –AG, –AGG

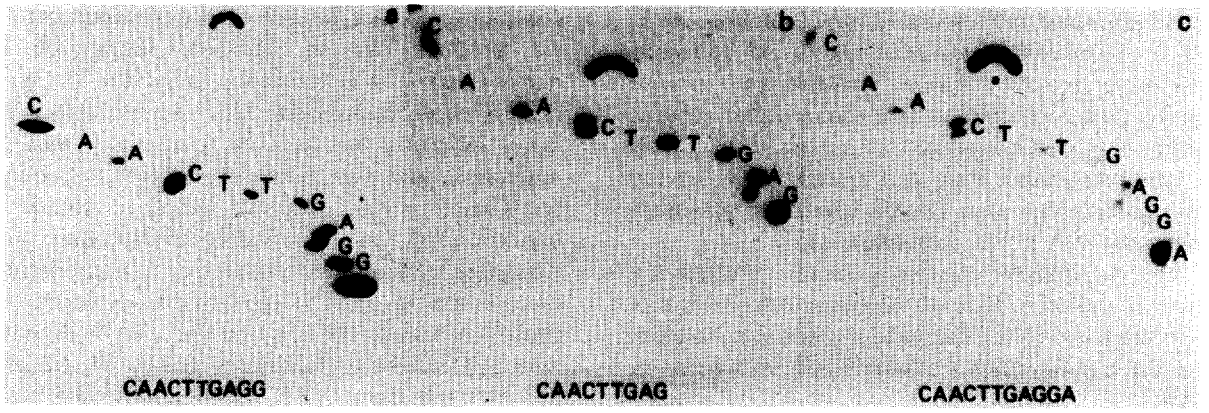

Fig. 2. Mobility shift analysis of 10*, 9* and 11* products (a, b and c, respectively) Electrophoresis was from the left to the right and chromatography was ascending. Starting from the first nucleotide, the nucleotide that causes the shift to the next longer product has been indicated.

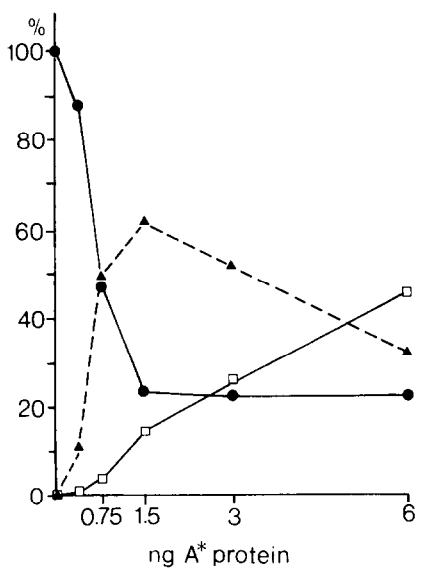

Fig. 3. The amounts of hexadecamer, heptamer and 10* product after incubation of the hexadecamer with varying amounts of A* protein; the hexadecamer was incubated with different amounts of A* protein for 60 min. The products were analyzed by gel electrophoresis. After autoradiography the relative amounts of the products were calculated by excising the bands corresponding to the hexadecamer, the 10* product (together with the 9* and 11* product) and the heptamer, and determining the radioactivity; 100% is the sum of the radioactivity of the 3 bands; (●—●) relative amount of hexadecamer; (▲—▲) relative amount of heptamer; (□—□) relative amount of 10* product.

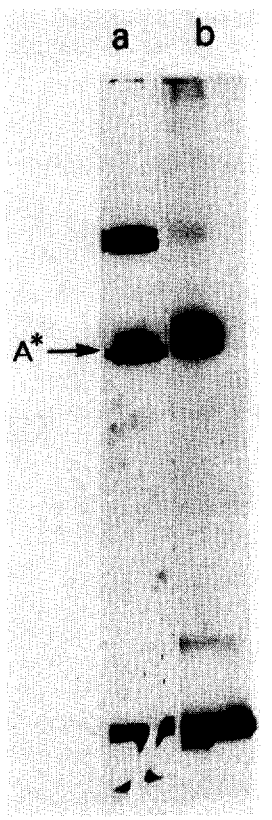

Fig. 4. SDS–polyacrylamide gel analysis of A* protein purified from cells infected with phage in the presence of ^{32}P: A* protein stained; (b) autoradiograph of the same lane.

and –AGGA must be present bound to the protein. This suggests that A* protein can be ^{32}P-labelled via the oligonucleotide. Therefore, we isolated A* protein from cells that had been infected with phage in the presence of [^{32}P]phosphate. Purified A* protein was analyzed on a SDS–polyacrylamide gel, the gel was stained to localize the A* band. An autoradiograph of the dried gels shows a radioactive band at the position of A* protein (fig. 4). A sample of the A* preparation was treated with alkali as in section 2 and analyzed on a gel. After the incubation with alkali radioactive material still comigrates with the A* protein. This indicates that the ^{32}P was not present as phosphate esterified to serine residues since this band is alkali labile.

So phosphate-containing material, possibly oligonucleotides, occurs firmly bound to a protein that migrates at the position of A* protein.

The sequences of the 9*, 10* and 11* products suggest that –AG, –AGG and –AGGA are transferred to the heptamer CAACTTG, which is generated by cleavage of the hexadecamer by A* protein. This transfer indeed takes place, for in an incubation of purified heptamer with A* protein 10* was produced (fig. 1). Formation of 10* product in an incubation of A* protein with the heptamer CAACTTG also shows that the transfer is not necessarily coupled to the cleavage reaction. Cleavage and transfer can proceed in two separate steps. Fig. 3 shows how much 10* is formed after 60 min incubation of the hexadecamer with five different amounts of A* protein (a constant portion of 20% of the hexadecamer seems not to be nicked). The amount of 10* product increases if more A* protein is added. Concomitantly, the amount of heptamer decreases. Some of the 10* product can, on its turn, be cleaved by A* protein and yields the heptamer. This is shown by incubation of purified 10* product with A* protein (fig. 1).

4. DISCUSSION

In incubations of purified A* protein with the synthetic hexadecamer CAACTTGATATTAATA, a heptamer with the sequence CAACTTG, a decamer CAACTTGAGG (10* product) and, to a lower extent, a nonamer CAACTTGAG (9* product) and an undecamer CAACTTGAGGA (11* product) are generated. The heptamer is formed by cleavage next to the G-residue. The presence of the 10* product (and 9* and 11*) can be explained by assuming that the A* protein preparation contains oligonucleotide material with the sequences (predominantly) –AGG and also –AG and –AGGA, which are transferred to the heptamer, CAACTTG. In analyses of incubations of A* protein with 5′-labelled single-stranded DNA fragments, the bands representing the cleavage products are often accompanied by bands that migrate slightly slower. The positions of these bands correspond to an increase in length of about three nucleotides (unpublished).

The 10* product arises mainly if a large amount of A* protein is added. Together with the observation that 10* can also be generated by incubation of the purified heptamer with A* protein, this indicates that 10* is formed in two steps:

(1) Cleavage of the hexadecamer which yield the heptamer;
(2) Transfer of the oligonucleotide to the heptamer.

The role of the A* protein during the ϕX infection cycle is unknown. Possibly it is concerned with the shut off of the host DNA synthesis [17–19]. The presence of oligonucleotides in the A* protein preparation is a new observation that may be the trace to the function of A* protein. Therefore it is important to know how the oligonucleotide(s) that can be transferred is (are) present and whether they are derived from host DNA or from viral DNA. Free oligonucleotides would have been lost during the purification procedure so they probably occur bound to the protein. Possibly the oligonucleotides occur covalently bound to the A* protein. This is supported by the observation of a radioactive band at the position of A* protein in an experiment in which A* protein was isolated from cells that had been infected with ϕXX am3 phage in the presence of ^{32}P. *In vitro* A* protein binds covalently to the 5′-end after cleavage of the DNA [6,16]. So the oligonucleotides which are carried by A* protein may indicate an earlier cleavage by A* protein. Since divalent cations, which are essential for the cleavage activity, are trapped by adding EDTA, A* protein cannot nick during the isolation procedure. So the A* molecules that carry an oligonucelotide have nicked DNA in a previous stage; i.e., in the infected cell.

The sequences of the oligonucleotides that are transferred, predominantly –AGG, but also –AG and –AGGA, demonstrate that previous cleavage and binding takes place with a high specificity. Possibly –AG, –AGG and –AGGA are the remnants of a same, longer DNA fragment that has been broken down by unspecific nucleolytic events. In vitro the sequence –CTTACTTGAGGA– (nucleotides 983–994 in the complete ϕX sequence [9]) in ϕX single-stranded viral DNA is efficiently cleaved next to the first G-residue by A* protein [16]. Cleavage in vivo at this site would yield A* protein bound to DNA with the sequence AGGA. There is no experimental evidence that this occurs. At the moment, we do not know whether the oligonucleotides are derived from host DNA or from viral DNA. Since the answer to this question will shed light on the function of A* protein, this is the subject of further investigation.

REFERENCES

[1] Linney, E. and Hayashi, M. (1973) Nature New Biol. 245, 6–8.
[2] Eisenberg, S., Griffith, J. and Kornberg, A. (1977) Proc. Natl. Acad. Sci. USA 74, 3198–3202.
[3] Langeveld, S.A., Van Mansfeld, A.D.M., Baas, P.D., Jansz, H.S., Van Arkel, G.A. and Weisbeek, P.J. (1978) Nature 271, 417–420.
[4] Van der Ende, A., Langeveld, S.A., Van Arkel, G.A. and Weisbeek, P.J. (1982) Eur. J. Biochem. 124, 245–252.
[5] Brown, D.R. and Hurwitz, J. (1982) in: Nucleases, Cold Spring Harbor Laboratory Press, NY, in press.
[6] Langeveld, S.A., Van Mansfeld, A.D.M., De Winter, J.M. and Weisbeek, P.J., (1979) Nucleic Acids Res. 7, 2177–2188.
[7] Langeveld, S.A., Van Arkel, G.A. and Weisbeek, P.J. (1980) FEBS Lett. 114, 269–272.
[8] Eisenberg, S. and Finer, M. (1980) Nucleic Acids Res. 8, 5305–5315.
[9] Sanger, F., Coulson, A.R., Friedmann, T., Air, G.M., Barrell, B.G., Brown, N.L., Fiddes, J.C., Hutchison, C.A. iii, Slocombe, P.M. and Smith, M. (1978) J. Mol. Biol. 125, 225–246.
[10] Van Mansfeld, A.D.M., Langeveld, S.A., Baas, P.D., Jansz, H.S., Van der Marel, G.A., Veeneman, G.H. and Van Boom, J.H. (1980) Nature 288, 561–566.
[11] Vereijken, J.M., Van Mansfeld, A.D.M. and Jansz, H.S. (1975) Virology 68, 221–233.
[12] Laemmli, U.K. (1970) Nature 227, 680–685.
[13] Boulikas, T. and Hancock, R. (1981) J. Biochem. Biophys. Methods 5, 219–228.
[14] Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[15] Brownlee, G.G. and Sanger, F. (1969) Eur. J. Biochem. 21, 395–399.
[16] Langeveld, S.A., Van Mansfeld, A.D.M., Van der Ende, A., Van de Pol, J.H., Van Arkel, G.A. and Weisbeek, P.J. (1981) Nucleic Acids Res. 9, 545–562.
[17] Martin, D.G. and Godson, G.N. (1975) Biochem. Biophys. Res. Commun. 65, 323–330.
[18] Funk, F.D. and Snover, D. (1976) J. Virol. 18, 141–150.
[19] Eisenberg, S. and Ascarelli, R. (1981) Nucleic Acids Res. 9, 1991–2002.