Volume 10 Number 15 1982 Nucleic Acids Research

Nucleotide sequence of the gene encoding adenovirus type 2 DNA binding protein

W.Kruijer, F.M.A.Van Schaik and J.S.Sussenbach*

Laboratory of Physiological Chemistry, State University of Utrecht, Utrecht, The Netherlands

Received 8 June 1982; Revised and Accepted 8 July 1982

ABSTRACT

We have determined the nucleotide sequence of the gene encoding adenovirus type 2 (Ad2) DNA binding protein (DBP). From the nucleotide sequence the complete amino acid sequence of Ad2 DBP has been deduced. A comparison of the amino acid sequences of Ad2 and Ad5 DBP, both 529 residues long, reveals that the C-terminal 354 residues of both sequences are identical. Within the N-terminal 175 amino acid residues Ad2 and Ad5 show nine differences.

The site of mutation in Ad2 ND1ts23, a mutant with a temperature-sensitive DNA replication, was mapped at the nucleotide level. A single nucleotide alteration in the DBP gene, resulting in a leucine → phenylalanine substitution at position 282 in the amino acid sequence is responsible for the temperature-sensitive character of this mutant.

Previously, we localized the mutation of another DBP mutant with a temperature-sensitive DNA replication (H5ts125) at position 413 in the amino acid sequence of the DBP molecule (Nucleic Acids Res. 9 (1981) 4439-4457). These mapping data are discussed in relation to the structure and function of the DBP molecule.

INTRODUCTION

The genomes of the homologous adenovirus types 2 and 5 consist of linear double-stranded DNA molecules with a MW of 23 x 10^6 daltons (1,2). The adenoviral DNA binding protein, which is encoded by early region E2A, located between co-ordinates 61.6 and 66.5 on the viral genome (Fig. 2) is one of the best characterized viral products (3-8). Analysis of Ad2 and Ad5 mutants with an altered DBP gene indicates that DBP is a multifunctional protein (8-10). Investigation of the temperature-sensitive mutant H5ts125 has revealed that this mutant does not replicate viral DNA (9), is unable to regulate transcription of early region E4 (11) and accumulates early viral mRNA (12,13) at the non-permissive temperature.

Analysis of another group of mutants (the host-range mutant H5hr404 and the related mutants H2hr400-hr403), which in contrast to wild-type virus grow efficiently in monkey cells, indicates that DBP is also involved in RNA proces-

0305-1048/82/1015-4493$2.00/0

sing (10,14). Recently, we established the nucleotide sequence of the Ad5 DBP gene and deduced from the nucleotide sequence that the DBP molecule (apparent MW 72,000 daltons) is 529 amino acid residues long (8). Ad2 and Ad5 DBPs can be cleaved by chymotrypsin into two fragments with a MW of 45 kD and 25 kD, respectively (15, 16).

The C-terminal 45 kD fragment, about 360 amino acids long, binds to single-stranded DNA (15, 16), is able to complement a defective *in vitro* DNA replication system of H5ts125 (17) and carries the H5ts125 mutation (8). The N-terminal 25 kD fragment, about 170 residues long, does not bind to single-stranded DNA, is extensively phosphorylated and contains the H5hr404 mutation (8,15,16). The structural and functional differences between these fragments strongly suggest that DBP is composed of two domains.

In this report we extend our studies on the structure and function of DBP by presenting the nucleotide sequence of the Ad2 DBP gene. We have also localized the site of the mutation in the DBP gene of another DNA synthesis-negative mutant Ad2 ND1ts23.

MATERIAL AND METHODS

Cells and viruses

Cell growth, propagation of Ad2 ND1ts123 in HeLa cells and purification of viral DNA were performed as described before (18). Ad2 ND1ts23 was kindly provided by J.F. Williams.

Plasmids and cleavage of DNA

The plasmid containing the Ad2 *Hin*dIII-A fragment inserted in the *Hin*dIII site of pBR322 (19) was kindly provided by M. Perricaudet. Plasmids, grown in HB101 were isolated according to Ish-Horowicz and Burke (20). For large-scale isolations, plasmid DNA was further purified by CsCl density gradient centrifugation for 20 h in a Beckman V50Ti rotor at 45,000 rpm. The Ad2 *Hin*dIII-A containing plasmid was double-digested with *Bgl*II and *Sac*I and the resulting *Bgl*II-fragment (coordinates 60.2 and 63.6) and *Bgl*II-*Sac*I-fragment (coordinates 63.6 and 69.1) were isolated (21). Subsequently the *Bgl*II-fragment was cleaved with *Rsa*I and *Hin*fI, respectively, and the *Bgl*II-*Sac*I-fragment was cleaved with *Dde*I, *Bst*NI and *Xho*I, respectively.

The resulting fragments were 5'-endlabeled according to Maxam and Gilbert (22) and digested with suitable restriction endonucleases to separate the two labeled ends of each fragment and to generate a set of overlapping subfragments. These subfragments were subjected to DNA sequence analysis.

DNA sequence determination

DNA sequence analysis was performed according to Maxam and Gilbert (22), modified slightly in order to determine sequences for fragments longer than 250 nucleotides (8).

Sequence ladders of corresponding restriction fragments of Ad2 and Ad2 ND1ts23 DNA were run side by side on the same sequence gel.

RESULTS AND DISCUSSION

Nucleotide sequence and organization of the Ad2 DBP gene

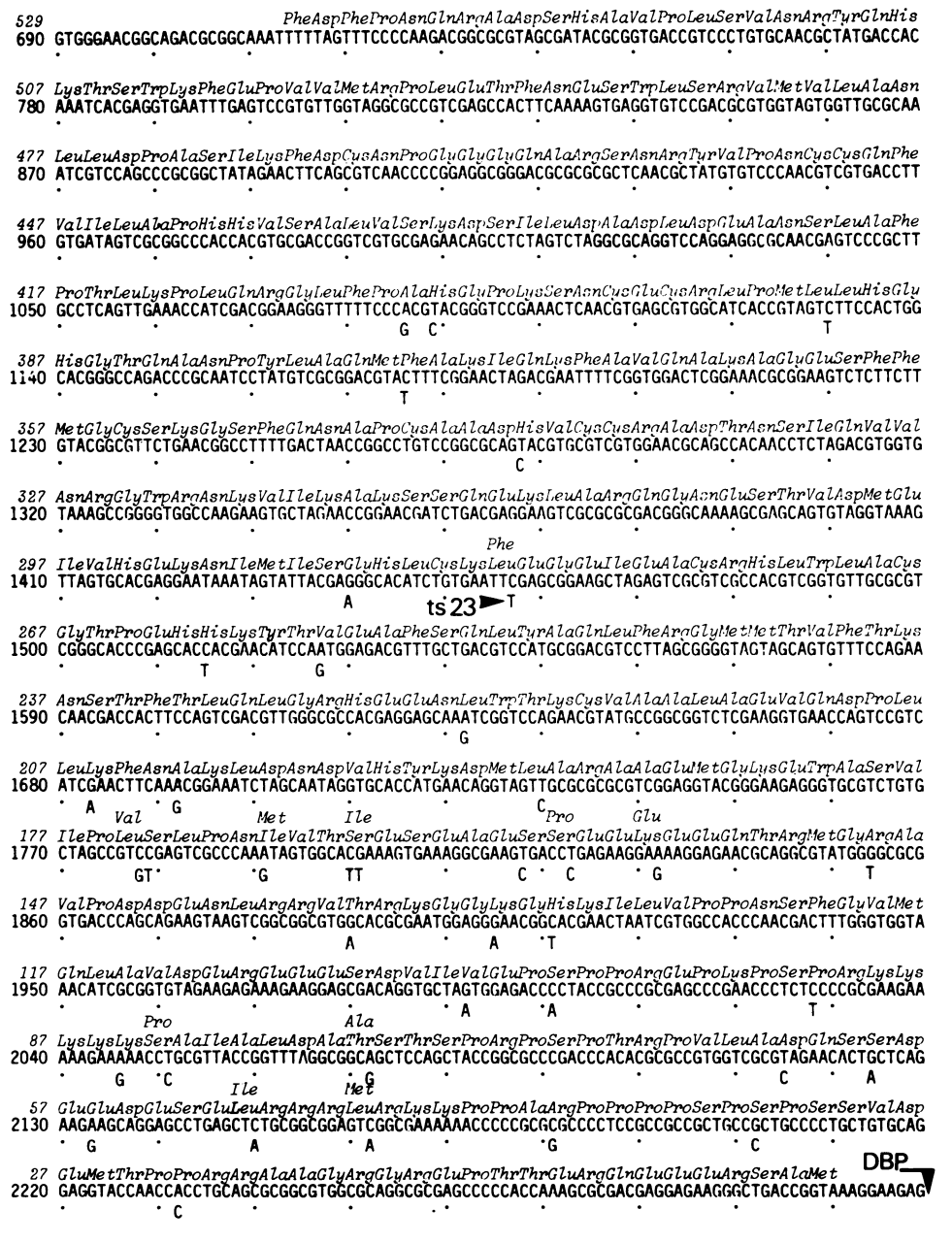
Figure 1 shows the established sequence of the Ad2 DBP gene and adjacent regions. The numbering of the residues is according to the previously established Ad5 sequence of this region (8).

The established Ad2 sequence between residues 690 and 991 is identical to the Ad2 sequence of this region as determined previously by Akusjärvi *et al.* (23). The positions of termination codons in the three reading frames of both complementary strands is indicated schematically in Fig. 2. Since r-strand transcripts are homologous to the l-strand of Ad2 DNA, the position of termination codons in r-strand transcripts are indicated in the l-strand sequence. Likewise, the termination codons of l-strand transcripts are indicated in the r-strand sequence.

Long open reading frames are present in the r-strand between ATG^{2300} and TAA^{713} and in the l-strand between ATG^{1747} and the right-hand end of the sequence shown in Fig. 1.

A comparison of the Ad2 and Ad5 sequences between residues 690 and 2346 shows a high degree of homology between both sequences and a limited number of base pair changes (Fig. 1). The region between 2233 and 2346 at the 5'-end of the DBP gene is identical in both serotypes. Since residue 2309 has been identified as the acceptor site at the 5'-end of Ad5 DBP encoding mRNAs (8), it is very likely that the main bodies of Ad2 and Ad5 DBP encoding mRNAs have identical 5'-ends. Also the 3'-ends of Ad2 and Ad5 DBP mRNA are the same, since the nucleotide sequences between residues 690 and 1085 are identical. The high degree of homology between the Ad2 and Ad5 sequences suggests very strongly that the open reading frame between ATG^{2300} and TAA^{713} in the Ad2 sequences encodes the Ad2 DBP.

Besides the long open reading frame encoding DBP, short open reading frames are found in frame 3 of the r-strand from residues 2326 to 1796 and from 1155 to 838.


ts 23
DBP

In the l-strand, both in Ad2 and Ad5 a long open reading frame is present downstream from ATG^{1747}, preceding the presumed initiation codon ATG^{2329} for the 100 kD protein. At present, however, there are no indications that these open reading frames encode proteins.

Amino acid sequence of Ad2 DBP

Figure 1 shows the amino acid sequence of Ad2 DBP as it can be deduced from the nucleotide sequence of the Ad2 DBP gene. The Ad2 and Ad5 DBP amino acid sequences differ at only nine positions (Fig. 1). These amino acid differences are clustered in the N-terminal 175 residues.

As outlined in the Introduction, Ad2 and Ad5 DBP can be cleaved by chymotrypsin into two structurally and functionally different domains of 45 kD and 25 kD, respectively. The exact cleavage site of chymotrypsin that generates these fragments is not known exactly, but there are indications that this site is located around position 170 (Asselbergs, Van der Vliet, personal communications). Therefore, the differences between Ad2 and Ad5 are probably all located in the N-terminal 25 kD fragment. Most amino acid differences between Ad2 and Ad5 are found between residue 157-175. The fact that these differences do not seem to influence any vital function of this multifunctional protein, suggests that the differences are located in a region of the DBP molecule which serves as a linker between the functional regions of the N- and C-terminal domains. However, further experiments are required to confirm this suggestion.

Analysis of the mutant Ad2 ND1ts23

The mutant Ad2 ND1ts23 was obtained after mutagenesis of the non-defective virus Ad2ND1 (24). Ad2 ND1ts23 has been mapped genetically in the same complementation group as H5ts125. The DBPs of both mutants are thermolabile and are sensitive to proteolytic degradation at the non-permissive temperature. Both mutants are defective in viral DNA replication at 39.5°C (25). The kinetics of shut off of viral DNA replication in mutant-infected cells after shift up to the non-permissive temperature are very similar for both mutants.

Figure 1. Nucleotide sequence of Ad2 DBP gene and adjacent regions. Only the sequence of one strand, the r-strand, is indicated. The expected position of the exon-intron junction at the 5'-end of the main body of DBP mRNA is indicated by a vertical arrow. The nucleotide sequence differences present in Ad5 DNA are indicated underneath the Ad2 sequence. When base pair differences result in amino acid differences, the corresponding amino acids present in Ad5 are indicated above the Ad2 amino acid sequence. A horizontal arrow marks the position of the mutation present in Ad2 ND1ts23 along with the amino acid substitution due to this lesion. The positions 690 to 991 in this sequence correspond to the positions 863 to 1164 in the nucleotide sequence determined by Akusjarvi et al. (23).

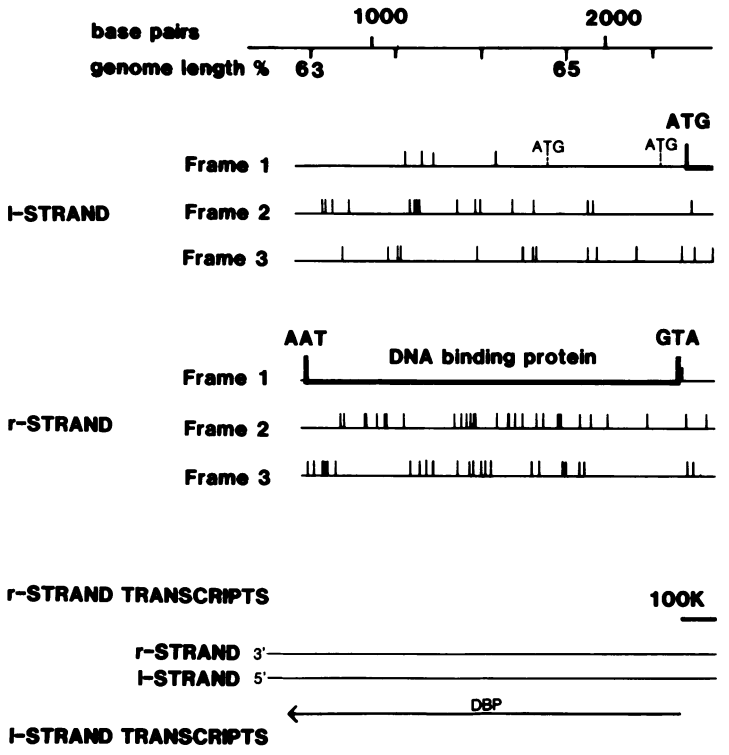


Figure 2. Distribution of initiation and termination codons between co-ordinates 61.9 and 66.6 of Ad2 DNA. Termination codons are designated by tick marks for each of three possible reading frames in r- and l-strand sequences. Open reading frames encoding the DBP and the 100 kD protein are indicated by heavy lines.

These results indicate that H5ts125 and Ad2 ND1ts23 are both DBP mutants, which display a very similar phenotype. This prompted us to determine the site of mutation in Ad2 ND1ts23.

By comparing the complete sequences of the genes encoding Ad2 and Ad2 ND1ts23 DBPs, we found a single-nucleotide alteration in the gene encoding DBP. At position 1457 a C-residue in Ad2 has been changed into a T-residue in the gene of Ad2 ND1ts23. This results in the substitution of a leucine residue at position 282 in Ad2 DBP by phenylalanine in DBP of the mutant Ad2 ND1ts23 (Fig. 3). In H5ts125 the mutation is located at position 413 in the amino acid sequence leading to a proline → serine substitution. This clearly established that H5ts125 and Ad2 ND1ts23 are both mutants with lesions in the gene encoding DBP, but that these lesions are located at different positions in the C-termi-

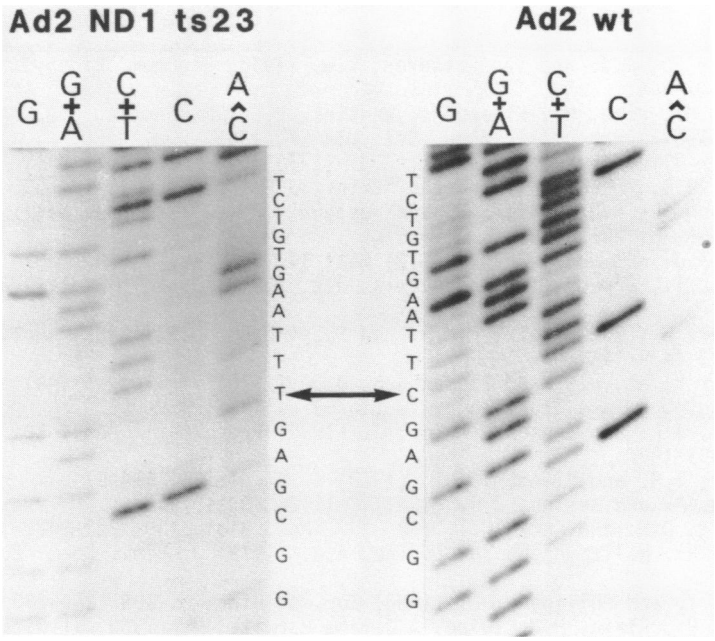


Figure 3.Sequencing gels showing the site of mutation in Ad2 ND1ts23. At position 1457 (Fig. 1) C in Ad2 is altered into T in Ad2 ND1ts23 as indicated by arrows.

nal 45 kD domain.

These results support the notion that in particular the C-terminal domain is involved in DNA replication, as was also suggested by the complementation of a defective *in vitro* DNA replication system of H5ts125 by the C-terminal 45 kD fragment (17).

*To whom correspondence should be addressed.

ACKNOWLEDGEMENTS

The stimulating discussions with Drs. H.S. Jansz and P.C. Van der Vliet as well as the expert technical assistance of R. Schiphof and M.G. Kuijk are gratefully acknowledged.

This work was supported in part by the Netherlands Foundations for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

REFERENCES

1. Van der Eb, A.J. and Van Kesteren, L.W. (1966) Biochem. Biophys. Acta 129, 441-444.
2. Green, M., Pina, M., Kimes, R., Wensink, P.C., Machattie, L.A. and Thomas, C.A. (1967) Proc. Natl. Acad. Sci. USA 57, 1302.
3. Van der Vliet, P.C. and Levine, A.J. (1973) Nat. New Biol. 246, 170.
4. Lewis, J.B., Anderson, C.W. and Atkins, J.F. (1977) Cell 12, 37.
5. Kitchingman, G.R., Lai, S.P. and Westphal, H. (1977) Proc. Natl. Acad. Sci. USA 74, 4392-4395.
6. Berk, A.J. and Sharp, P.A. (1978) Cell 14, 695-711.
7. Chow, L.T., Broker, T.R. and Lewis, J.B. (1979) J. Mol. Biol. 134, 265-303.
8. Kruijer, W., Van Schaik, F.M.A. and Sussenbach, J.S. (1981) Nucl. Acids Res. 9, 4439-4457.
9. Van der Vliet, P.C. and Sussenbach, J.S. (1975) Virology 67, 415-426.
10. Klessig, D.F. and Grodzicher, T. (1979) Cell 17, 957-966.
11. Nevins, J.R. and Jensen-Winkler, J. (1980) Proc. Natl. Acad. Sci. USA 77, 1893-1897.
12. Carter, T.H. and Blanton, R.A. (1978) J. Virol. 25, 664-674.
13. Babich, A. and Nevins, J.R. (1981) Cell 26, 371-379.
14. Klessig, D.F. and Chow, L.T. (1980) J. Mol. Biol. 139, 221-242.
15. Klein, H., Maltzman, W. and Levine, A.J. (1979) J. Biol. Chem. 254, 11051-11060.
16. Linné, T. and Philipson, L. (1980) Eur. J. Biochem. 103, 259-270.
17. Ariga, H., Klein, H., Levine, A.J. and Horwitz, M. (1980) Virology 101, 307-310.
18. Kruijer, W., Van Schaik, F.M.A. and Sussenbach, J.S. (1980) Nucl. Acids Res. 8, 6033-6042.
19. Stenlund, A., Perricaudet, M., Tiollais, P. and Pettersson, U. (1980) Gene 10, 47-52.
20. Ish-Horowicz, P. and Burk, J.F. (1981) Nucl. Acids Res. 9, 2989-2998.
21. Wieslander, L. (1979) Anal. Biochem. 98, 305-309.
22. Maxam, A.M. and Gilbert, W. (1980) in Methods of Enzymology, Eds. L. Grossman and K. Moldave, Academic Press, New York, 65, 499-560.
23. Akusjärvi, G., Zabielski, J., Perricaudet, M. and Pettersson, U. (1981) Nucl. Acids Res. 9, 1-7.
24. Lewis, A.M., Levin, M.J., Wiese, W.H., Crumpacker, C.S. and Henny, P.H. (1969) Proc. Natl. Acad. Sci. USA 63, 1128.
25. Van der Vliet, P.C., Blanken, W.M., Zandberg, J. and Jansz, H.S. (1977) in DNA Synthesis, Present and Future, Eds. I. Molineux and M. Kohiyama, Plenum Press, New York, pp. 869-885.