ELSEVIER

FEBS Letters 344 (1994) 166–170

FEBS 13995

FEBS LETTERS

Nucleic acid-binding regions of the second-largest subunit of *Drosophila* RNA polymerase II identified by Southwestern blotting

Roland E. Kontermann**, Ekkehard K.F. Bautz*

Institute of Molecular Genetics, University of Heidelberg, Im Neuenheimer Feld 230, D-69120 Heidelberg, Germany

Received 30 March 1994

Abstract

Analysing overlapping bacterially expressed fragments of the second-largest subunit of *Drosophila melanogaster* RNA polymerase II in Southwestern DNA binding assays we have identified regions that have the potential to bind nucleic acids non-specifically. A region exhibiting strong DNA binding is located in the N-terminal part of the molecule (amino acids 357–504) and some weak DNA binding is observed for the C-terminal part (amino acids 860–1160). The non-specific DNA binding behavior of these regions is similar to that of the native enzyme. Most of the known mutations responsible for rifampicin resistance map to a region of the *Escherichia coli* β subunit corresponding to the N-terminal nucleic acid-binding region, indirectly supporting the notion that this region participates in interaction with the RNA transcript in ternary complexes.

Key words: Non-specific DNA binding; RNA polymerase; Fusion protein; Southwestern blotting; Transcription

1. Introduction

DNA-dependent RNA polymerases (EC 2.7.7.6.) have the potential to interact with DNA in a sequence-independent manner primarly based on electrostatic interactions. These non-specific binary complexes are not able to initiate RNA synthesis. Accurate and efficient initiation of transcription needs additional factors which position the RNA polymerase at initiation sites (for review see [1]). The electrostatic nature of non-specific DNA binding is characterized by a strong dependence on the salt concentration and a general affinity for nucleic acids and other polyanionic substances [2–5]. In contrast to non-transcribing binary complexes, in elongating ternary complexes the RNA polymerase interacts not only with the DNA template but also with the DNA–RNA hybrid and the nascent RNA. Binding of nucleic acids in these complexes is more resistant to high salt concentrations and negatively charged substances such as sarcosyl and heparin [6,7].

Eukaryotic nuclear RNA polymerases are composed of two non-identical large subunits with molecular weights between 120 and 240 kDa and several smaller subunits [8]. As shown by photochemical cross-linking experiments, Southwestern blotting techniques, and antibody studies, both large subunits are involved in the interaction with the DNA template and the nascent RNA chain [9–15]. Using Southwestern assays and nitrocellulose filter binding assays of bacterially expressed fusion proteins we have recently identified a region within the largest subunit of *Drosophila* RNA polymerase II with the potential to bind nucleic acids non-specifically with a similar binding behavior as the native enzyme [16]. Similarity of the region to part of the DNA-binding cleft of *E. coli* DNA polymerase I involved in binding of the newly synthesized DNA double strand, indirectly supports the idea that this region of the largest subunit participates in interaction with nucleic acids during transcription.

In the work reported here we have extended this study to the second-largest subunit. Analysing fusion proteins expressing overlapping parts of the second-largest subunit of RNA polymerase II of *Drosophila melanogaster* by Southwestern blotting we were able to identify at least one region within the second-largest subunit with the potential to bind DNA non-specifically.

2. Materials and methods

2.1. Enzymes and oligonucleotides

RNA polymerase II was purified in our laboratory from *Drosophila* embryos as described [17] and *E. coli* RNA polymerase was purchased from Boehringer-Mannheim (Germany). *Taq* DNA polymerase was purchased from Pharmacia (Freiburg, Germany). Oligonucleotide primers used for PCR were synthesized by Dr. R.W. Frank (ZMBH, Heidelberg).

2.2. Construction of β-galactosidase fusion proteins

Genomic DNA or cDNA fragments of the coding region of the gene encoding the second-largest subunit of *D. melanogaster* RNA polymerase II (DmRP140, 1,176 amino acids) ([17] and references therein)

*Corresponding author. Fax: (49) (6221) 562 678.

***Present address:* MRC Centre for Protein Engineering, Hills Road, Cambridge, CB2 2QH, UK.

SSDI 0014-5793(94)00380-E

were cloned into pUR expression vectors [18]. Smaller fragments of the DmRP140 sequence (fusion proteins 140-11 to 140-14 covering amino acids 357–504, and fusion proteins 140-15 and 140-16 covering amino acids 912–993) were synthesized by PCR amplification under conditions as described [16] using the following primers: sense primer, 140-1 (GCGCGGATCCTTTCTGCGAGACCAAGAAGG; corresponding to amino acids 357–363), 140-2 (GCGCGGATCCCTTCCTCTTCCGTGGTCTGT; amino acids 403–409), 140-5 (GCGCGGATCCTACGCTCAACAGCGAGGGCT; amino acids 912–918) and 140-7 (GCGCG-GATCCAAGGTATTCCTTGGCCACTG; amino acids 446–452); antisense primer, 140-3 (GCGCAAGCTTAATACCTTAAACCATCAGTA; amino acids 441–447), 140-4 (GCGCAAGCTTAATTGTGAACGTGACGAGGT; amino acids 498–504), 140-6 (GCGCAAGCTTATGT-GAAGGCCATGTCCTCC; amino acids 953–959) and 140-8 (GCGCAAGCTTATCCCAG-TTTGCCCTGCAGA; amino acids 987–993).

2.3. *Purification of fusion proteins and Southwestern DNA binding assays*

Fusion proteins were expressed in RR1DM15 cells in the presence of 1 mM IPTG and were purified from inclusion bodies as described [16]. Proteins were solubilized in 20 mM Tris-HCl, pH 8.3, 8 M urea, 10 mM DTT. Southwestern assays were performed as described [16].

3. Results

3.1. *Localization of nucleic acid-binding regions*

In Southwestern DNA binding assays the second-largest subunit exhibits about 40–50% of the binding capacity of the largest subunit (Fig. 1). DNA binding is strongly reduced at higher salt concentrations and can be competed with by unlabeled DNA or RNA as well as by heparin [16]. Analysis of *E. coli* RNA polymerase under identical conditions shows, in contrast, only non-specific DNA binding of the β' subunit while the β subunit as well as the α and σ subunit do not bind DNA (Fig. 1). In order to identify DNA-binding regions of the second-largest subunit of RNA polymerase II, overlapping β-galactosidase fusion proteins were produced covering amino acids 18–1159 (constructs 140-1 to 140-10) and were analysed in Southwestern DNA binding assays. Two constructs bind DNA rather strongly (140-3 and

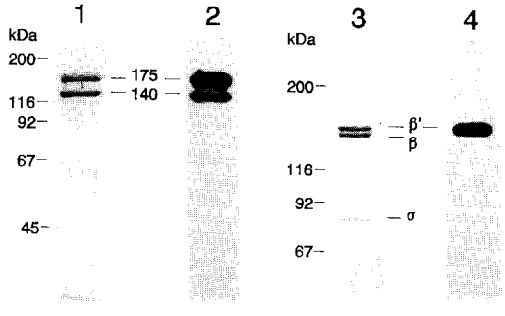

Fig. 1. Purified RNA polymerase II of *Drosophila* (2 μg) and *E. coli* RNA polymerase (2 μg) were separated on a 7% polyacrylamide gel by SDS-PAGE and either stained with Coomasie blue (lanes 1 and 3) or analysed for non-specific binding of labeled plasmid DNA in Southwestern assays (lanes 2 and 4). The positions of the largest subunit lacking the CTD (175 kDa) and the second-largest subunit (140 kDa) of RNA polymerase II and of the β', β, and σ subunits of *E. coli* RNA polymerase, respectively, are indicated.

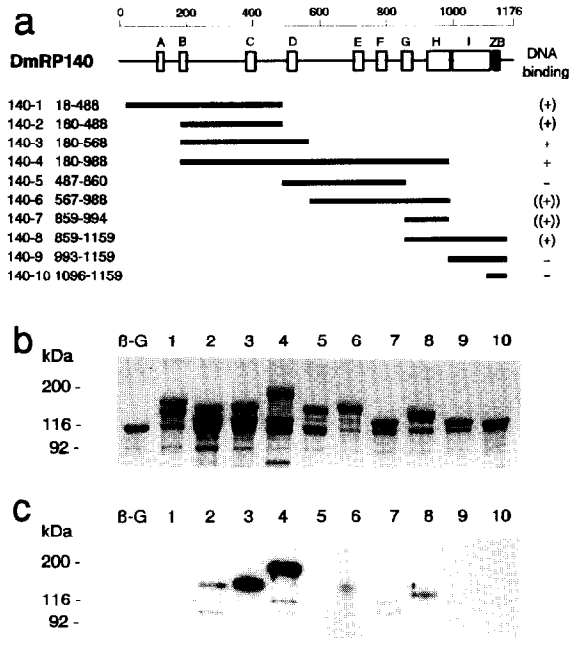

Fig. 2. (a) Location of fusion proteins of the second-largest subunit of RNA polymerase II (140-1 to 140-100. The positions of regions A–I, which are highly conserved within the second-largest subunits of RNA polymerases of eukaryotes and bacteria [8], as well as the zinc-binding motif at the C-terminus (ZB) found in the eukaryotic second-largest subunits are indicated. Fusion proteins as well as β-galactosidase without a fusion part (β-G) were separated by SDS-PAGE and were either stained with Coomassie blue (b) or analysed by Southwestern blotting for non-specific binding of labeled plasmid DNA (c). Lane numbers correspond to the assignment of the fusion proteins.

140-4), and relatively weak binding was observed for constructs 140-1, -2, -6, -7, and -8 (Fig. 2). β-Galactosidase expressed and purified under the same conditions was not able to bind DNA and no binding was observed with [^{32}P]dCTP. No differences in the binding pattern were observed using heat-denatured DNA (data not shown). The binding assays indicate that a region exhibiting significant DNA binding is located near the N-terminal portion of the subunit (region 1) and that a region with weak DNA binding capacity resides close to the C-terminus (region 2).

In order to further narrow down the positions of these DNA-binding regions, smaller β-galactosidase fusion proteins of the second-largest subunit were produced by cloning of PCR amplified fragments (Fig. 3). A fusion protein expressing amino acids 357–504 (140-11) exhibits about the same DNA-binding capacity as construct 140-3 (amino acids 180–568) (Fig. 4). Deletions of the C-terminal part of construct 140-11 (140-12: amino acids 357–447; 140-13: amino acids 403–447) diminish DNA binding, but DNA binding is observed for construct 140-14 expressing amino acids 403–504. However, densitometric analyses indicate that fusion protein 140-14 exhibits only about 50% of the DNA-binding capacity of constructs 140-11 and 140-3. Thus, the entire region

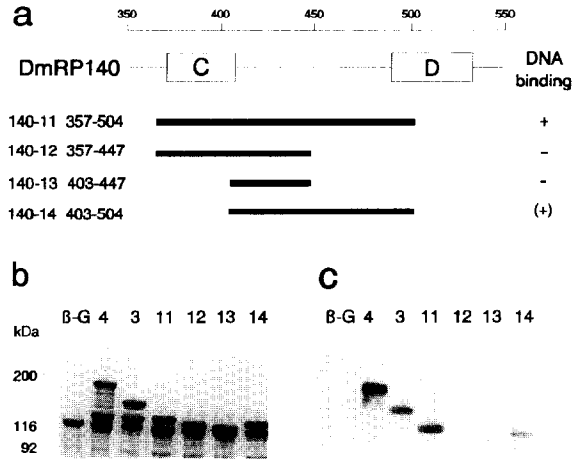

Fig. 3. (a) Location of fusion proteins 140-11 to 140-14 of the second-largest subunit of RNA polymerase II. The positions of the two conserved regions C and D are shown. (b) Coomassie blue stained fusion proteins separated on a 5% polyacrylamide gel by SDS-PAGE. Lane numbers correspond the assignment of the fusion proteins. For positions of fusion proteins 140-3 and 140-4 see figure 2. (c) Southwestern blot of the fusion proteins under standard conditions.

expressed by fusion protein 140-11 (amino acids 357–504) is necessary for efficient DNA binding.

Almost no DNA binding was observed for smaller fusion proteins of the C-terminal part indicating that the whole C-terminal region (represented by construct 140-8: amino acids 859–1159) is necessary for weak DNA binding (data not shown).

3.2. Characterization of the DNA-binding behavior

DNA binding of fusion proteins 140-3 (region 1) and 140-8 (region 2) in Southwestern assays depends strongly on the concentration of mono- and divalent cations. Increasing the NaCl concentration from 50 mM to 100 mM reduces the DNA-binding capacity to about 5–10%, and DNA binding is totally diminished at 150 mM. A reduction of DNA binding is also seen by increasing the $MgCl_2$ concentration. The non-specificity and the general affinity of the fusion proteins for nucleic acids and polyanionic substances was shown by competition assays. Blotted fusion protein 140-3 was preincubated with unlabeled calf thymus DNA, calf liver RNA or with heparin at varying concentrations. 50% inhibition of DNA binding by construct 140-3 was observed at about 5 μg/ml DNA and about 15 μg/ml RNA, while 50% inhibition by heparin is already seen at 0.05 μg/ml. Similar results were obtained for construct 140-8. Competition experiments with excess amounts of ATP or ApU showed no reduction of non-specific DNA binding of any of the constructs. In summary, both regions behave like non-specific nucleic acid-binding regions.

3.3. Structural conservation of the nucleic acid-binding regions

An alignment of amino acids 446–504 of region 1 (which are necessary but not sufficient for DNA binding) with the corresponding region of the second-largest subunits of RNA polymerases I, II, and III of *Drosophila* and yeast, as well as the RNA polymerases of various archaebacteria, *E. coli* and tobacco chloroplasts, is shown in Fig 4. Between 25–45% of the amino acid residues are conserved within this part of the subunits. As shown in Fig. 3, region 1 includes region C and part of region D, which are highly conserved within the second-largest subunits of RNA polymerases [8]. Sequence analysis showed that region 1 is rich in basic amino acid residues (15–19%) resulting in an isoelectric point of about 11, whereas region 2 has about 14% basic amino acid residues and an isoelectric point of 8.7. The region of the *E. coli* β subunit (amino acids 503–552) corre-

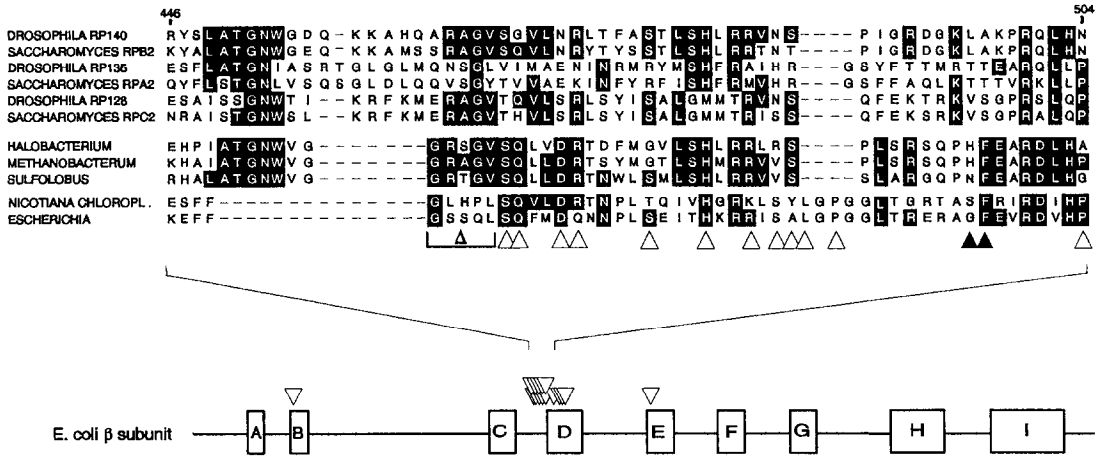

Fig. 4. Alignment of amino acids 446–504 of region 1 of *Drosophila* with the corresponding regions of the second-largest subunits of other eukaryotic and bacterial RNA polymerases. Amino acid residues with > 45% identity are marked as black squares. The positions of mutations responsible for rifampicin resistance (△) or streptolydigin resistance (▲) in the *E. coli* β subunit are shown. The location of the conserved regions A–I in the β subunit are indicated as open boxex. Sequences are from the SWISSPROT database.

sponding to the C-terminal part of region 1 contains most of the known mutations (amino acid substitutions, deletions) responsible for rifampicin resistance [19,20] as well as streptolydigin resistance [21] (Fig. 4).

4. Discussion

We have identified two regions within the second-largest subunit of *Drosophila* RNA polymerase II with the potential to bind DNA non-specifically. The similar binding behavior of these regions and that of the separated subunit in Southwestern assays and of the native enzyme in nitrocellulose filter binding assays [16] indicates that non-specific DNA binding is based on the same mechanisms. A high content of basic amino acid residues, especially of region 1 exhibiting strong DNA binding, is in agreement with an electrostatic nature of non-specific DNA binding [5].

The major non-specific DNA binding site of the second-largest subunit is located in the N-terminal part and covers conserved region C and part of region D. Region C alone is not able to bind DNA, but the requirement of region C for binding is indicated by a reduced DNA binding of fragment 140-14 lacking this region. A 10 amino acid deletion of the β subunit of *E. coli* RNA polymerase (amino acids 435–444), which resides in the conserved region C, results in an interrupted transcription cycle and a lethal phenotype [22]. The mutation decreases the affinity of the enzyme to the promoter and the stability of the open complex, but does not affect activity in the 'moving' mode of transcription analysed in a slippage assay with single-stranded template. The authors suggested that region C is involved in DNA binding in the 'static' mode of transcription initiation. This finding agrees with our observation that region C of the *Drosophila* equivalent of the β subunit contributes to nucleic acid binding.

Photochemical cross-linking experiments with *E. coli* RNA polymerase showed that the β', β, and σ subunit become cross-linked to DNA in non-specific binary complexes, while in specific promoter complexes only β and σ were labeled [9]. However, no DNA binding can be observed for the β subunit alone, consistent with our results from Southwestern assays, but sub-assembled $\alpha_2\beta$ complexes bind to limited sites on DNA [23]. Thus, formation of a DNA-binding site of the β subunit needs assembly into $\alpha_2\beta$ complexes or the core enzyme, respectively, while the second-largest subunit of *Drosophila* RNA polymerase II by itself has DNA-binding potential.

Besides a possible participation of the identified regions in non-specific DNA template binding in binary complexes, these regions might be involved in interaction with nucleic acids in the transcribing complex. Binding of the *E. coli* β subunit or of the second-largest subunit of various eukaryotic RNA polymerases (including *Drosophila* RNA polymerase II) to DNA and RNA in transcribing complexes has been previously demonstrated by crosslinking experiments [10,13,14,24]. The observation that most of the known mutations responsible for rifampicin resistance and streptolydigin resistance resides in a region of the *E. coli* β subunit corresponding to nucleic acid-binding region 1 points to a functional importance of this region. Rifampicin interferes with binding of RNA chains longer than 2–3 nucleotides to the enzyme [25], and interaction with RNA and rifampicin was previously demonstrated for the sub-assembled *E. coli* $\alpha_2\beta$ complex [26]. These findings indirectly support the proposal that region 1 participates in the interaction with the RNA transcript.

Recently, similarity between the second-largest subunits of various eukaryotic and bacterial RNA polymerases and bacterial RNases has been reported and the authors suggested an RNase activity of this region and that this part of the molecule plays a role in RNA binding [27]. The region of similarity maps to amino acid residues 287–399 of the second-largest subunit of *Drosophila* RNA polymerase II. Thus, this sequence overlaps with the nucleic acid-binding region 1 (amino acids 357–504). A 3′-5′ exonuclease activity has been recently found to be an intrinsic property of human RNA polymerase II [28].

A second nucleic acid-binding region is located in the highly conserved C-terminal part. This region of the molecule forms part of the active centre and is responsible for binding of the initiating dinucleotide [29]. The physiological relevance of this region for DNA binding, however, is unclear since only weak binding is observed in Southwestern assays. Further analyses will be needed to show to what extent the two identified nucleic acid-binding regions of the second-largest subunit participate in non-specific binding of the DNA template in binary complexes and interaction with DNA and RNA in transcribing complexes.

Acknowledgements: We thank R. Baier (Heidelberg) for helping with scanning densitometry. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 229) and the EC (SCI/CT91/7202) to E.K.F.B.

References

[1] Conaway, J.W. and Conaway, R.C. (1991) J. Biol. Chem. 266, 17721–17724.
[2] Chamberlin, M.J. (1976) in: RNA polymerases (Losick, R. and Chamberlin, M.J. eds.) pp. 18–67, Cold Spring Harbor Laboratory Press, New York.
[3] deHaseth, P.L., Lohman, T.M., Burgess, R.R. and Record Jr., M.T. (1978) Biochemistry 17, 1612–1622.
[4] Shaner,S.L., Melancon, P., Lee, K.S., Burgess, R.R. and Record Jr., M.T. (1982) Cold Spring Harbor Sym. Quant. Biol. 47, 463–472.

[5] von Hippel, P.H., Bear, D.G., Morgan, W.D. and McSwiggen, J.A. (1984) Annu. Rev. Biochem. 54, 389–446.
[6] Arndt, K.M. and Chamberlin, M.J. (1990) J. Mol. Biol. 213, 79–108.
[7] Kadonaga, J.T. (1990) J. Biol. Chem. 260, 2624–2631.
[8] Young, R.A. (1991) Annu. Rev. Biochem. 60, 689–715.
[9] Hillel, Z. and Wu, C.-W. (1978) Biochemistry 17, 2954–2961.
[10] Gundelfinger, E.D. (1983) FEBS Lett. 157, 133–138.
[11] Carroll, S.B. and Stollar, B.D. (1983) J. Mol. Biol. 170, 777–790.
[12] Huet, J., Phalente, L., Buttin, G., Sentenac, A. and Fromageot, P.F. (1982) EMBO J. 1, 1193–1198.
[13] Hanna, M.M. and Meares, C.F. (1983) Proc. Natl. Acad. Sci. USA 80, 4238–4242.
[14] Bartholomew, B., Dahmus, M.E. and Meares, C.F. (1986) J. Biol. Chem. 261, 14226–14231.
[15] Bartholomew, B., Durkovich, D., Kassavetis, G.A. and Geiduschek, E.P. (1993) Mol. Cell. Biol. 13, 942–952.
[16] Kontermann, R.E., Kobor, M. and Bautz, E.K.F. (1993) Protein Sci. 2, 223–230.
[17] Seifarth, W., Petersen, G., Kontermann, R., Riva, M., Huet, J. and Bautz, E.K.F. (1991) Mol. Gen. Genet. 228, 424–432.
[18] Rüther, U. and Müller-Hill, B. (1983) EMBO J. 2, 1791–1794.
[19] Jin, D.J., Walter, W.A. and Gross, C.A. (1988) J. Mol. Biol. 202, 245–253.
[20] Landick, R., Stewart, J. and Lee, D.N. (1990) Genes Dev. 4, 1623–1636.
[21] Lisitsyn, N.A., Monastyrskaya, G.S. and Sverdlov, E.D. (1988) Eur. J. Biochem. 177, 363–369.
[22] Martin, E., Sagitov, V., Burova, E., Nikiforov, V. and Goldfarb, A. (1992) J. Biol. Chem. 267, 20175–20180.
[23] Fukuda, R. and Ishihama, A. (1974) J. Mol. Biol. 87, 523–540.
[24] Stackhouse, T.M. and Meares, C.F. (1988) Biochemistry 27, 3038–3045.
[25] Schulz, W. and Zillig, W. (1980) Nucleic Acids Res. 9, 6889–6907.
[26] Mi, H. and Hartmann, G.R. (1983) Eur. J. Biochem. 131, 113–118.
[27] Shirai, T. and Go, M. (1991)Proc. Natl. Acad. Sci. USA 88, 9056–9060.
[28] Wang, D. and Hawley, D.K. (1993) Proc. Natl. Acad. Sci. USA 90, 843–847.
[29] Grachev, M.A., Lukhtanov, E.A., Mustaev, A.A., Zaychikov, E.F., Abdukayumov, M.N., Rabinov, I.V., Richter, V.I., Skoblov, Y.S. and Chistyakov, P.G. (1989) Eur. J. Biochem. 180, 577–585.