Mapping of DNA-binding proteins along the yeast genome by UV-induced DNA-protein crosslinking

D.A. Papatsenko*, I.V. Priporova, S.V. Belikov, V.L. Karpov

W.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 117984 Moscow, Russia

Received 27 October 1995; revised version received 10 January 1996

Abstract UV-induced crosslinking of DNA-binding proteins to DNA in intact nuclei of *Saccharomyces cerevisiae* and subsequent 'protein image' hybridization were applied to map non-histone proteins along single-copy genes of yeast. We detected two polypeptides that most probably correspond to core subunits of yeast RNA-polymerase II in the coding region of transketolase gene (TKL2). Several non-histone proteins were also detected which bind to the upstream region of TKL2 gene, and to the intergenic spacer between calmodulin (CMD1) and β-mannosyl transferase (ALG1) genes.

Key words: UV-light; DNA-protein crosslinking; *Saccharomyces cerevisiae*; DNA-binding protein

1. Introduction

Successful development of the Human Genome Project and other genome sequencing programs has already generated huge amount of new sequences. They include specific regions that are associated in nuclei with regulatory proteins. It is essential to have a suitable method for mapping and characterization of DNA-binding proteins in vivo. However, the number of currently available methods are rather limited [1].

In our laboratory we have developed the 'protein image' hybridization method [2] which utilizes the ability of UV-light to generate (induce) DNA–protein crosslinks in a whole cell or in isolated nuclei [3–7]. It enables one to determine the molecular mass (Mm) of proteins associated with specific DNA sequences thus being a good alternative/complement to the gel-retardation assay. Our method could be useful for checking the difference in the pattern of proteins bound to the sequences of interest and/or bound to the same sequence with different activation status.

Here we describe mapping of DNA-binding proteins in the coding and promoter regions of the TKL2 gene [8] and inside the intergenic spacer between CMD1 and ALG1 genes [9–11].

2. Experimental

Yeast diploid cells (strain MR-2) were grown in a rich medium to the optical density of 2–3 A_{600}/ml. Nuclei were isolated as described earlier [12]. To induce DNA–protein crosslinks, purified nuclei were irradiated by conventional UV light (254 nm, irradiation dose about 0.5 J/cm^2). 'Protein imaging' was done as described by Papatsenko et al. [13] (see also [14]). DNA fragments were electrotransferred onto HYBOND N+ membrane (Amersham, GB) according to the manufacturer's recommendations and hybridized thereon with high-labeled single-stranded probes synthesized using bacteriophage M13 vectors [15].

'Protein' two-dimensional gel electrophoresis was done as described earlier [14].

*Corresponding author.

Abbreviations: UV (light), ultra-violet (light); TKL2, transketolase gene; CMD1, calmodulin gene; ALG1, β-mannosyl transferase gene; Mm, molecular mass; PAAG, polyacrylamide gel; 2D, two-dimensional.

3. Results

The general concept of the 'protein image' hybridization assay was published earlier [16]. Briefly: Yeast nuclei were UV-irradiated, lysed, and loaded onto CsCl density gradient to purify DNA–protein complexes from the most part of non-crosslinked proteins and other contaminations. DNA-containing fractions were collected after centrifugation and DNA was shared by sonication. Most of the uncrosslinked DNA was removed by the phenol-chloroform extractions; purified DNA–protein complexes were loaded onto PAAG, and two-dimensional (2D) gel electrophoresis was performed. In this system DNA–protein complexes were separated in the 1st dimension. The proteins were then digested directly in the gel by pronase, the resulting DNA was separated in the 2nd dimension and electrotransferred onto a nylon membrane. Finally, the filter was successively hybridized with the probes of interest.

In order to keep DNA and proteins intact, we had to use rather low doses of irradiation (0.5 J/cm^2). As a result, the yield of crosslinked proteins was often very low, especially in the case of a single copy gene. To overcome this problem, we used M13-derived ^{32}P-labeled probes substituted by three labeled nucleotides. Several types of nylon membranes were tested to find an optimal one in terms of the highest signal/background ratio and a possibility of numerous rehybridizations. HYBOND N+ (Amersham) was found to be the best.

Autoradiograms of the filter after 'protein image' hybridizations with the coding and the promoter regions of the TKL2 gene, and intergenic spacer between CMD1 and ALG1 genes are shown in Fig. 1. Hybridization with the coding region of TKL2 gene revealed three diagonals corresponding to the crosslinked proteins of Mms 17, 45, and 100 kDa, respectively (Fig. 1A). A different set of proteins was detected in the promoter region of the gene. The diagonal corresponding to the protein of 17 kDa was also present but a diffuse pattern appears instead of proteins 45 and 100 kDa (Fig. 1B). The pattern corresponds to several proteins with Mms of 100–140 kDa. Hybridization with the intergenic spacer between CMD1 and ALG1 genes allows us to detect the diagonal corresponding to 17 kDa protein, proteins of 100–120 kDa, and low intensity one, corresponding to a protein with Mm of 150 kDa (Fig. 1C).

SSDI S0014-5793(96)00091-9

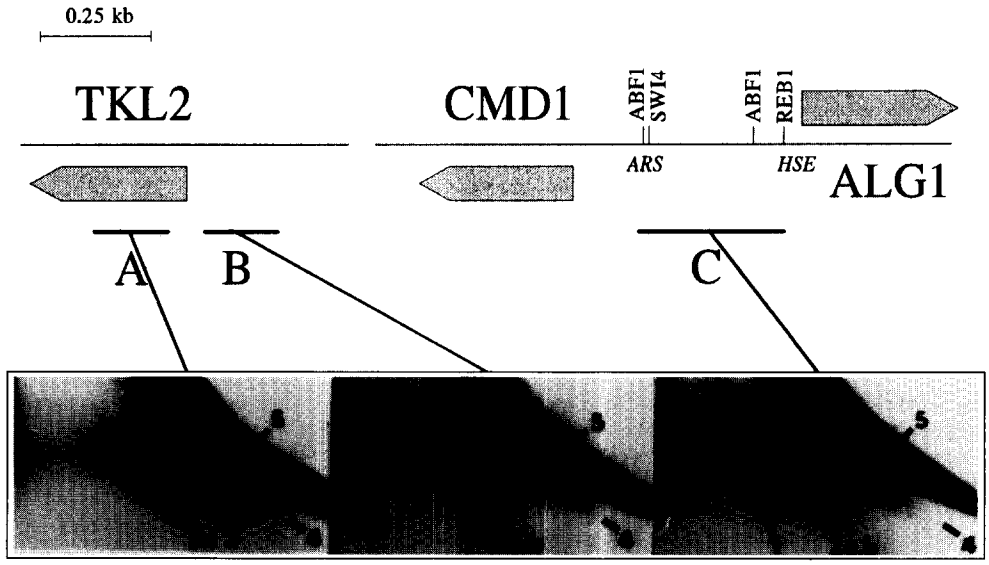

Fig. 1. 'Protein image hybridization' with different regions of TKL2, CMD1-ALG1 genes. A: coding region of TKL2 gene; B: promoter region of TKL2 gene; C, intergenic spacer between CND1 and ALG1 genes. 1, protein of Mm 150 kDa; 2a–b, proteins of Mm within the range 100–140 kDa; 3, protein of Mm 45 kDa; 4, protein of Mm 17 kDa; 5, uncrosslinked DNA. Genetic map of the genes and homology sites for DNA-binding proteins (see text) presented above.

4. Discussion

We have initiated DNA–protein interactions studies with the final aim to combine data at the nucleotide sequence level with biochemical data on the arrangement of DNA-binding proteins along DNA. Here the 'protein image' hybridization assay was used to probe DNA–protein interactions in the coding and upstream regions of the TKL2 gene, and in the intergenic spacer between CMD1 and ALG genes of yeast.

We believe the crosslinked proteins of 17 kDa Mm correspond to the core histones. The attribution of this diagonal to the core histones is based on calibrations of the 2D gels [17]. Also we observed proteins corresponding to the core histones after 2D 'protein' electrophoresis (Fig. 2).

It was logical to expect that the transcribed region of an active gene would contain RNA-polymerase molecules. Basing on the molecular masses of the proteins crosslinked in the coding region we assumed that they probably correspond to the RNA-polymerase II subunits (Fig. 1A). The core subunits of yeast RNA-polymerase II include polypeptides RPB1, RPB2, RPB3 with molecular masses of 220, 150, and 45 kDa, respectively [18]. Mm of the crosslinked 45 kDa protein coincides with that of the RPB3 subunit; 100 kDa polypeptide could be a product of specific degradation of either RPB1, or RPB2. However, the diagonals that we observed in the coding region may correspond to some other specific proteins.

We observed a rather diffuse pattern corresponding to non-histone proteins bounded to the promoter region of TKL2 gene (Fig. 1B). Separation of the crosslinked complexes in 2D system was inefficient, and the diffuse pattern did not permit us to distinguish between proteins of close molecular masses.

Genes CMD1 and ALG1 are transcribed in opposite directions and their promoters are separated by a 0.5 kb-long region. The intergenic spacer between these genes has significant homology to the ARS sequence and it contains several potential binding sites for non-histone proteins: REB1, ABF1, SWI4 [9–11]. It is not clear now whether or not binding of these proteins takes place in vivo [19]. However, the Mms of the proteins we observed in the intergenic spacer are close to those of ABF1 (100 kDa) and REB1 (120 kDa) (Fig. 1C). Precise identification of crosslinked proteins could be done using antibodies against proteins of interest.

In conclusion, UV-mapping technology allows rapid preliminary screening of DNA-binding proteins along genomic DNA and it could be very helpful for projects dealing with mapping of DNA–protein contacts.

Acknowledgements: We are grateful to Prof. Andrei Mirzabekov for advice and stimulating discussions. We thank Dr. Olga Preobrazhenskaya and Mrs. Elena Novikova for useful comments on the manuscript. This work was supported by grants from the International Science Foundation (Grant # MGL 000), Volkswagen-Stiftung (Germany), by the Russian Foundation for Fudamental Research (Grant no. 897) and by the Russian State Program 'Frontiers in Genetics'.

References

[1] Wang, M.X. and Church, G.M. (1993) Nature 360, 606–610.

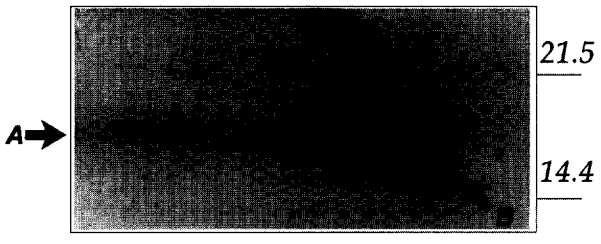

Fig. 2. Proteins crosslinked to yeast DNA. DNA–protein complexes were prepared as described above. After the 1st dimension DNA component of the complexes were digested accordingly to [14] and proteins were resolved in the 2nd dimension. A, core histones; B, diagonal of uncrosslinked proteins. Positions of marker proteins are given to the right.

[2] Karpov, V.L., Preobrazhenskaya, O.V. and Mirzabekov, A.D. (1984) Cell 36, 423–431.
[3] Belikov, S.V., Dzherbashyajan, A.R., Preobrazhenskaya, O.V., Karpov, V.L. and Mirzabekov, A.D. (1990) FEBS Lett. 273, 205–207.
[4] Belikov, S.V., Belgovsky, A.I., Preobrazhenskaya, O.V., Karpov, V.L. and Mirzabekov, A.D. (1993) Nucl. Acids Res. 21, 1031–1034.
[5] Belikov, S.V., Belgovsky, A.I., Partolina, M.P., Karpov, V.L. and Mirzabekov, A.D. (1993) Nucl. Acids Res. 21, 4796–4802.
[6] Hockensmith, J., Kubashek, W., Vorachek, W., Evertsz, E. and von Hippel, P. (1991) Methods Enzymol. 208, 211–235.
[7] Papatsenko, D.A., Priporova, I.V., Kukarin, A.V., Preobrazhenskaya, O.V., Karpov, V.L. and Mirzabekov, A.D. (1994) Molecular Biology (Russia) 26, 1376–1382.
[8] Fletcher, T.S., Kwee, I.L., Nakada, T., Largman, C. and Martin, B.M. (1992) Biochemistry 31, 1892–1896.
[9] Davis, T.N., Urdea, M.S., Masiarz, F.R. and Thornen, J. (1986) Cell 47, 423–431.
[10] Albright, C.F. and Robbins, P.W. (1990) J. Biol. Chem. 265, 7042–7049.
[11] Mannhaupt, G., Stucka, R., Ehnle, S., Vetter, A. and Feldmann, H. (1992) Yeast 8, 397–408.
[12] Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J. and Struhl, K. (1990) Current Protocols in Molecular Biology. New York: John Wiley and Sons.
[13] Papatsenko, D.A., Belikov, S.V., Preobrazhenskaya, O.V. and Karpov, V.L. (1995) Methods Mol. Cell. Biol. 5, 171–177.
[14] Mirzabekov, A.D., Bavykin, S.G., Belyavsky, A.V., Karpov, V.L., Shick, V.V., Preobrazhenskaya, O.V. and Ebralidse, K.K. (1989) Methods Enzymol. 170, 386–408.
[15] Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991–1995.
[16] Nacheva, G., Guschin D., Preobrazhenskaya, O., Karpov, V., Ebralidze K. and Mirzabekov A. (1989) Cell 58, 27–36.
[17] Belyavsky, A.V., Bavykin, S.G., Goguadze, E.G. and Mirzabekov, A.D. (1980) J. Mol. Biol. 139, 519–536.
[18] Young, R.A. (1991) Annu. Rev. Biochem. 60, 689–715.
[19] Mannhaupt, G., Stucka, R., Ehnle, S., Vetter, A. and Feldmann, H. (1994) Yeast 10, 1363–1381.