Analysis by atomic force microscopy of Med8 binding to *cis*-acting regulatory elements of the *SUC2* and *HXK2* genes of *Saccharomyces cerevisiae*

F. Moreno-Herrero[a,b], P. Herrero[a], J. Colchero[b], A.M. Baró[b], F. Moreno[a,*]

[a] *Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Biotecnología de Asturias, Universidad de Oviedo, 33006 Oviedo, Spain*
[b] *Laboratorio de Nuevas Microscopías, Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

Received 3 September 1999

Abstract Med8 protein is a regulator that specifically binds to upstream activating sequences (UASs) of *SUC2* promoter, to downstream repressing sequences (DRSs) of the *HXK2* gene and to the carboxy-terminal domain of the RNA polymerase II. Atomic force microscopy has allowed for direct visualization of Med8 interactions with a 305 bp fragment of *SUC2* promoter and with a 676 bp fragment of *HXK2* gene, containing respectively the UASs and DRSs regulatory regions. This approach has provided complementary information about the position and the structure of the DNA-protein complexes. Med8 binding to DNA results in total covering of one of the two existing 7 bp motives (consensus, (A/C)(A/G)GAAAT) in the studied DNA fragments. No preference for binding either of the two UASs of *SUC2* promoter as well as for the two DRSs of *HXK2* gene has been found. We also discuss whether this protein works as dimer or as a monomer.

Key words: Atomic force microscopy; *SUC2*; *HXK2*; Med8; Transcriptional regulation; *Saccharomyces cerevisiae*

1. Introduction

The RNA polymerase II transcription initiation apparatus consists on a multisubunit complex, called the RNA polymerase holoenzyme. In this complex are included the RNA polymerase II, the general transcription factors [1], the products of five *SRB* genes, identified as suppressors of carboxy-terminal domain (CTD) truncation mutants [2], and six members of a protein family, termed Med (Med2–5, Med7 and Med8), required for transcriptional activation [3], which function through the CTD of RNA polymerase II [4]. One role of the CTD is to serve as an interaction site for multiple mediator proteins. The CTD heptapeptide repeat sequence (YSPTSPS) is rich in hydroxyl groups [5,6], providing multiple sites for these interactions. Once anchored to RNA polymerase II via the CTD, components of the holoenzyme may interact with other factors necessary for regulated transcription initiation, such as activator proteins. The CTD structure is highly conserved from yeast to man, the heptad sequence is repeated 52 times in mammals, 44 times in *Drosophila melanogaster* and 26 times in *Saccharomyces cerevisiae* [6]. The CTD of RNA polymerase II is essential for cell growth because truncation of half of the repeats in yeast causes cold sensitivity and inability to express several genes [7].

Recently, in a search to identify new factors required for expression of the *SUC2* gene in *S. cerevisiae*, we have partially purified the Med8 protein and demonstrated that it specifically binds to both the downstream repressing sequences (DRSs) of the *HXK2* gene and to the upstream activating sequences (UASs) of the *SUC2* gene [8]. Because Med8 has been described as a mediator protein interacting with the CTD of the RNA polymerase II, its role could be to act as coupling factor by linking activating and repressing transcription complexes to the RNA polymerase II holoenzyme transcriptional machinery. This suggests that Med8 may be important to couple the glucose repression pathway of the *SUC2* gene to the *HXK2* gene expression. Using a heterologous glutathione S-transferase (GST)-Med8 fusion protein, we have demonstrated, by band shift assays, that Med8 specifically binds to a synthetic, 34 bp, double-stranded oligonucleotide containing a 7 bp motif with the consensus sequence (A/C)(A/G)GAAAT [8], located both in the UAS region of the *SUC2* promoter [9,10] and in the DRS region of the *HXK2* gene [11,12]. However, this methodology was not able to demonstrate if the protein-DNA interaction was established through the 7 bp motif of the double-stranded oligonucleotide used as probe neither if the protein binds to one or two motives simultaneously per DNA molecule.

To address these questions, we have used the recently highly improved technique called atomic force microscopy (AFM) [13]. AFM (also termed scanning force microscopy) appears as a powerful technique for imaging DNA and other biological assemblies (for reviews, see [14-16]). AFM is now a powerful tool that can reveal the assembly of complexes between proteins and DNA [17,18] and to monitor the activity of enzymes [19,20]. AFM can work under several conditions: in vacuum, in liquids [21,22] as well as in ambient air conditions at room temperature. Mica has been extensively used, as a substrate, for AFM imaging [23,24]. It is cleaved before sample deposition and the cleaved sheet exposes a very flat layer of basal oxygens, negatively charged because of a structural imbalance of charge. DNA binding to mica has been shown to be enhanced in the presence of millimolar concentrations of Mg^{2+} [25,26]. Details for reproducible imaging of individual DNA molecules have also been worked out independently in several laboratories and AFM images of DNA can be obtained quickly and routinely [27,28]. In this study, we use the atomic force microscope to examine, by direct visualization, Med8 protein-DNA complexes and the location of protein interactions with two DNA fragments from the *SUC2* promoter and from the *HXK2* gene both containing

*Corresponding author. Fax: (34) (98) 510 3157.
E-mail: fms@sauron.quimica.uniovi.es

PII: S0014-5793(99)01289-2

two repeats of the putative binding motif. We have also analyzed the binding stoichiometry of Med8 protein to the DNA fragments.

2. Materials and methods

2.1. Strains

S. cerevisiae strain AMW-13C$^+$ (*MATa trp1(fs) ura3(fs) leu2-3,112 his3-11,15 can1*) [29] was used throughout this work. Bacterial transformation, large scale propagation of plasmid DNA and fusion protein expression were performed in *Escherichia coli* BL21(DE3)-pLysS (Promega).

2.2. Heterologous expression of the S. cerevisiae MED8 gene in E. coli and purification of the fusion protein GST-Med8

A 675 bp DNA fragment, of chromosome II of *S. cerevisiae* containing the complete coding region of the *MED8* gene, was amplified from AMW-13C$^+$ genomic DNA by the polymerase chain reaction (PCR), using as primers OL1 = 5′-AAGGATCCATGTCACAATC-TACTGC-3′ and OL2 = 5′-GCGAATTCAATTACTAGATGATGT-TGA-3′. The PCR product was cleaved with *Eco*RI and *Bam*HI (underlined) and inserted into an *Eco*RI/*Bam*HI previously cleaved pGEX-2T plasmid. The resulting plasmid (pGEX-MED8) was used to transform *E. coli*.

E. coli cells containing the expression plasmid pGEX-MED8 were used to produce the *S. cerevisiae* protein Med8 as a fusion protein with the *Schistosoma japonicum* GST according the procedure previously described [8].

The fusion protein GST-Med8 was purified as in [8].

2.3. Preparation of free DNAs

Two different DNA fragments have been used to study specific DNA-protein interactions: (i) a 305 bp DNA fragment of the *SUC2* promoter containing two 7 bp motives with the consensus sequence (A/C)(A/G)GAAAT, located at −423 and −443 bp relative to the translational start codon, which are apparently required for gene derepression [9,10], and (ii) a 676 bp DNA fragment of the *HXK2* gene containing two 7 bp motives (consensus sequence (A/C)(A/G)GAAAT) identical to the ones contained in *SUC2* promoter. These two 7 bp motives are located at +144 bp relative to the translational start codon, in the DRS1 (from +139 to +167 bp) and at +240 bp relative to the translational start codon, in the DRS2 (between +230 and +249 bp). These elements are required for gene repression when cells are growing in media with ethanol as carbon source [11,12].

The 305 bp DNA fragment of the *SUC2* promoter was amplified by PCR, using as primer oligonucleotides OL3 = 5′(−487)-TTAT-TACTCTGAACAGGAA-3′ and OL4 = 5′(−182)-AAGTCGTCAA-ATCTTTCTT-3′.

The 676 bp DNA fragment of the *HXK2* gene was amplified by an identical procedure using PCR and oligonucleotides OL5 = 5′(+39)-GGTTCCATGGCCGATGTG-3′ and OL6 = 5′(+715)-GCACCATT-GACACCAGTACC-3′ as primers. In both cases, genomic DNA from the wild-type strain AMW-13C$^+$ has served as the *SUC2* promoter and *HXK2* gene-containing template. The amplified products were purified using the high pure PCR product purification kit (Boehringer Mannheim) following the indications of the manufacturer.

2.4. Protein-DNA binding reaction and AFM sample preparation

The typical binding reactions contained 10 mM HEPES (pH 7.5), 1 mM dithiothreitol, 80 mM NaCl, 0.5 ng of DNA fragment and 10 ng of purified GST-Med8 fusion protein in a volume of 25 μl. After 30 min of incubation at room temperature, 3 μl of the binding reaction mixture was mixed with 1 μl of 110 mM $MgCl_2$. The 4 μl was deposited on a freshly cleaved mica and allowed to bind for 1 min. Preparations were then rinsed with 5–10 ml MilliQ water and the excess liquid wicked away at the mica edge with a tissue. The mica was then blown completely dry with a stream of nitrogen gas and stored into a vacuum desiccator until it could be imaged by AFM [24,26].

The samples of free DNA were prepared in the same way, by making special attention on the concentration to obtain a convenient molecule density (around 10 molecules per μm^2), except that the DNA was resuspended in 20 mM Tris-HCl, pH 8.0, containing 10 mM $MgCl_2$.

2.5. AFM imaging

The AFM images were obtained with a commercial microscope (Nanotec Electrónica S.L., Spain) operating in a non-contact tapping mode. Images were collected in ambient air at room temperature at about 30% humidity. We used Olimpus type cantilevers with a nominal force constant of 1 N/m, resonance frequency of 75 kHz and a tip curvature radius of less than 20 nm. Images were recorded at typical scan frequencies of 2–3 Hz and the set point was set as high as possible to keep tip-sample interaction the softest.

Images were processed by subtracting a general plane to remove the background slope and with a two dimensional Fourier transform to eliminate the low frequency noise.

3. Results and discussion

3.1. AFM imaging of linearized DNA; length measurements

The deposition protocol described in Section 2 leads to substrate-sample interactions strong enough to prevent the tip from moving and sweeping the molecules around, but at the same time not to strong to cause deformation. However, the shape and dimensions of the tip strongly affect the resolution in AFM imaging. The effect of the finite size of the tip is to overestimate the width of the DNA molecule. We measure a mean DNA diameter of 10 nm. Taking the known 2 nm diameter for B-form DNA [28], we found a tip radius of 12 nm [30]. However, the measured height of the DNA molecules is slightly less than 1 nm [22]. So, we must consider a tip radius of about 20 nm according to the specifications of the manufacturer. To explain the smaller height found for DNA, many reasons have been put forward [16,31]. A water layer, present in all surfaces in ambient air, the existence of un-dissolved salts as well as the dehydration of DNA molecules are the most common explanations for this result. A priori AFM length measurements of long 'one dimensional' molecules should be obtained within a determination which depends on the tip radius as well as on the method used for obtaining the length of the DNA molecules from the images.

To address these questions, we have studied eight different types of DNA with well-known lengths, in terms of the number of bp. The DNA fragments were obtained using restriction endonucleases or PCR and agarose gel electrophoresis (Fig. 1A, a–h). We have worked a wide size range going from 271 to 5188 bp. Exhaustive studies of samples were done. An average length, measured with the WSxM software that controls the microscope, and a S.D. for each type of molecule was obtained (Table 1). We have calculated the linear regression to obtain a typical length per pair of bases (Fig. 1B) and we have deduced a value for the axial rise of 0.34 ± 0.01 nm/pb, which is consistent with the data previously reported [25]. Moreover, we find that the linear regression passes through the origin with a precision of 8 bp, which corresponds to about 2.5 nm. These results confirm the high sensitivity and linearity of the method developed and utilized in this work for determining DNA lengths from AFM images and that is consistent with small as well as with long DNA strands.

This method of measuring DNA length can also be used to identify the nucleotides involved in protein binding. About the precision in positioning the protein, we want to stress the following aspect: the binding interaction involves a few amino acids and bases. With the resolution of our AFM, we cannot resolve this small area, but we discern a blob of 13–14 nm in diameter as a result of the tip convolution. From the diameter imaged, we calculate a typical protein diameter of 3 nm. Us-

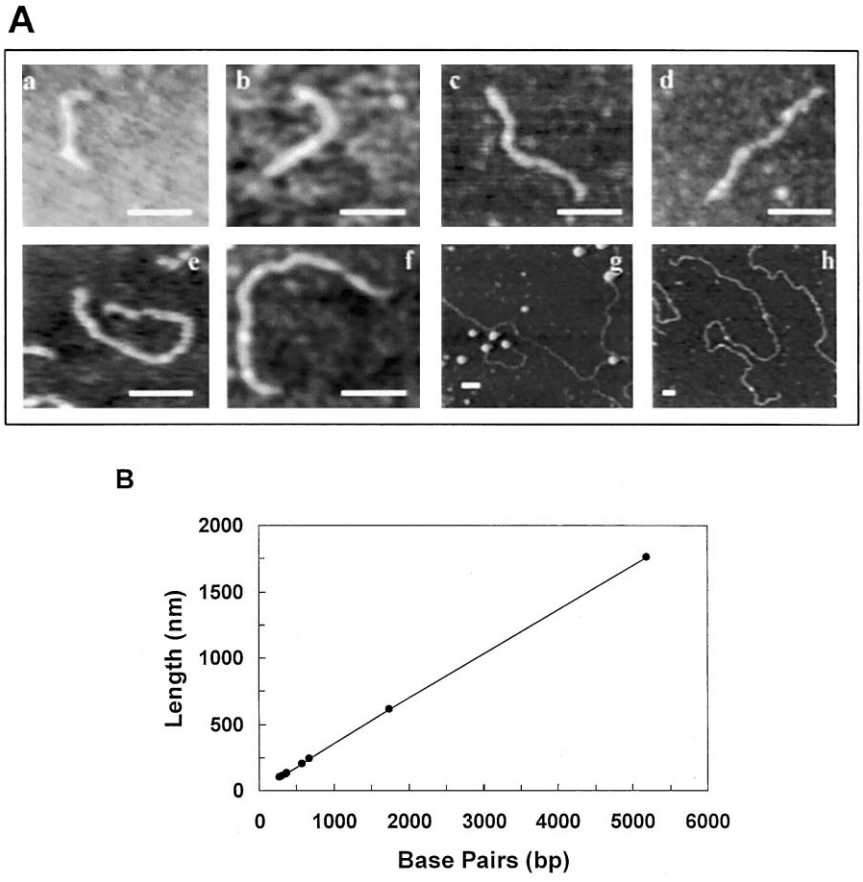

Fig. 1. (A) AFM images of DNA fragments of different lengths showing the structure of DNA bound to mica. In all images, the bar size is 50 nm. The (a) and (g) DNA fragments were obtained by *Rsa*I digestion of pUC18 and were isolated and purified by 0.8% agarose gel electrophoresis. The (b) DNA fragment is a PCR product amplified from the *SUC2* promoter. The (c) DNA fragment was obtained from *PHO5* promoter by digestion with *Cla*I/*Sal*I. The (d) DNA fragment was obtained from *HXK2* gene by digestion with *Eco*RI/*Hin*dIII. The (e) and (f) DNA fragments are PCR products amplified from the *HXK2* gene. The (h) DNA fragment is the episomal plasmid YEp352 digested with the restriction enzyme *Eco*RI until completion as judged by 0.8% agarose electrophoresis observation. (B) Correlation of the measured lengths with the number of bp of the DNA fragments. The linearity of the measurements is demonstrated even for long strands. We obtain a value for the axial rise of 0.34 ± 0.01 nm/pb.

ing the constant previously mentioned, the area of the image covered by the protein is about 40 bp. So, we find it reasonable to assume this value as the experimental error.

3.2. AFM visualization of GST-Med8/DNA complexes

The fusion protein GST-Med8 was produced by heterologous expression of the *S. cerevisiae MED8* gene in *E. coli* and purified by affinity chromatography as was previously described [8]. The GST-Med8/DNA complexes were prepared using the binding reaction and sample preparation described in Section 2.

To study the *SUC2* DNA-protein complexes, we used a 305 bp DNA linearized fragment containing the UAS1 and UAS2 regulatory elements of the *SUC2* promoter (Fig. 2). The 305 bp DNA fragments complexed with GST-Med8 protein were prepared as described above and imaged by AFM (Fig. 2A). Bright globular blobs represent the fusion protein GST-Med8. Two populations of complexed blobs, coating two different regions of the DNA molecules, were detected. The 150×150 nm images show the two populations of complexed blobs imaged (Fig. 2B,C). In Fig. 2B, we can see the typical image of the DNA-protein complexes with the protein coating the end of the DNA molecules. Because the diameter of the fusion protein given by the AFM is about 14 nm, the coated region of DNA is the expected if the protein binds to DNA through the UAS1 sequence of the *SUC2* promoter. In the same way, Fig. 2C represents a typical image of the second type population. In this case, we find the position of the maximum height of the protein at a mean value of 17 nm from the edge of the DNA molecule, which corresponds to about 50 bp, in good agreement with the fusion protein binding to DNA through the UAS2 sequence of the *SUC2* promoter.

A complementary study was performed using a 676 bp DNA linearized fragment containing the DRS1 and DRS2

Table 1
Measured length of DNA samples

DNA sample	Size (bp)	Length (nm)
a	271	98 ± 6
b	305	109 ± 5
c	355	122 ± 5
d	366	131 ± 7
e	574	203 ± 6
f	676	237 ± 6
g	1739	616 ± 9
h	5188	1760 ± 6

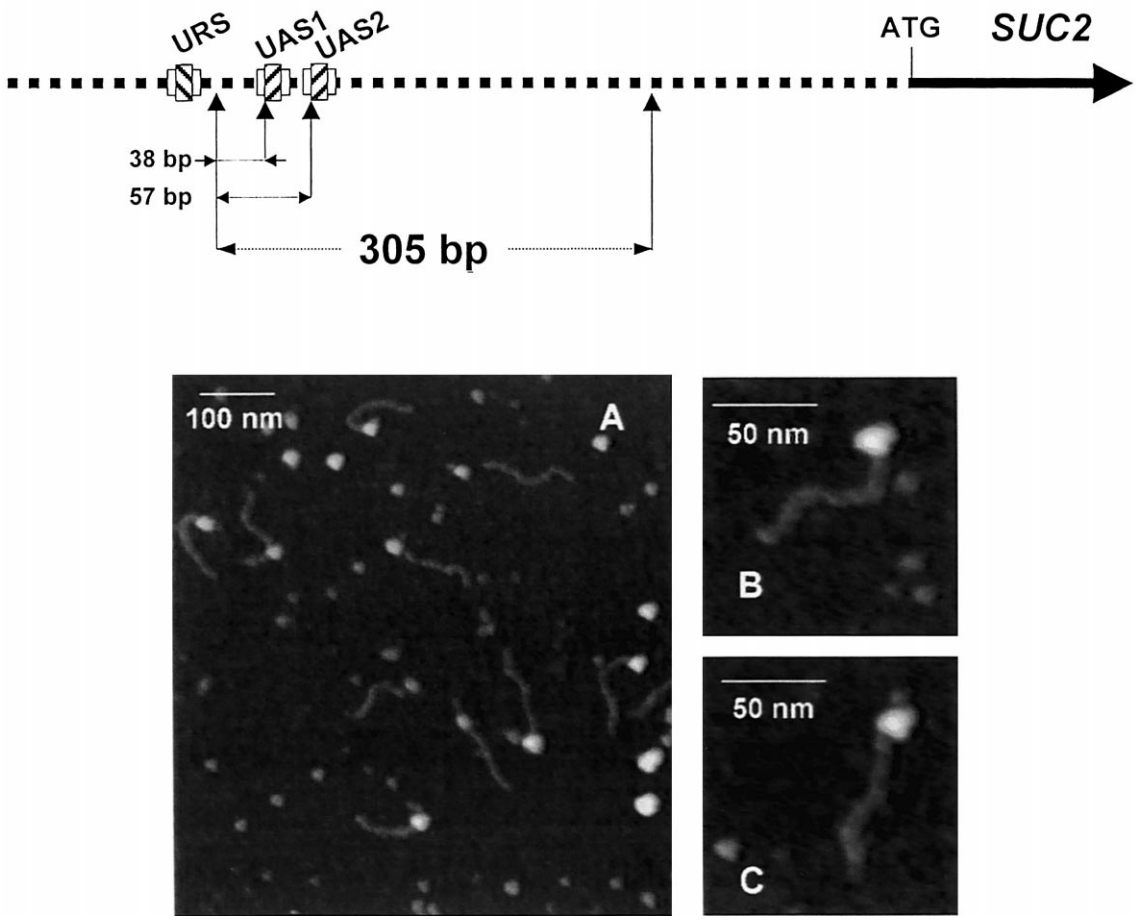

Fig. 2. Schematic picture of the *SUC2* gene showing the UASs regions, the 7 bp motif and its relative positions from the 5′ end of the 305 bp DNA fragment utilized. (A) AFM image of complexes of the 305 bp DNA fragments of *SUC2* promoter with the GST-Med8 protein. Two populations clearly differentiated can be seen. B and C: examples of the two types of DNA-protein complexes observed. For details, see Section 2.

regulatory elements of the *HXK2* gene at 100 and 191 bp from one of the ends of the fragment. Visualization of the 676 bp DNA fragments complexed with GST-Med8 protein by AFM has demonstrated that the fusion protein coated two different regions of the DNA molecules (Fig. 3A). The 300×150 nm images show the two populations of complexed blobs detected (Fig. 3B,C). In Fig. 3B, we can see the binding of the protein with the DRS1 sequence. From the statistical analysis of this type of images, we found a mean position of 113 bp. This result is consistent with the protein binding to the DNA through the DRS1 sequence of the *HXK2* gene. In the same way, we can calculate the position of the fusion protein for the second type of images (Fig. 3C). In this case, we found a mean position of the protein of 182 bp that means that, within the experimental errors, the coated region of the DNA is the expected if the protein binds to the DNA through the DRS2 sequence of the *HXK2* gene.

Our results strongly confirm that the Med8 protein binds to regulatory elements controlling the expression of *SUC2* and *HXK2* genes and that the protein binds through sequences containing a 7 bp motif with the consensus sequence (A/C)(A/G)GAAAT. Another significant observation reached by AFM is that in both cases, using the *SUC2* or the *HXK2* DNA fragments to form DNA-protein complexes, only one binding site per DNA molecule is occupied and that it does not matter what the length between the two binding sites is. Moreover, we found that the protein affinity is the same for the two binding sites of the DNA molecule.

In addition to the main results of the present work regarding binding sites, we have tried to use the data acquired to determine whether this protein appears as a monomer or as a dimer. The Med8 protein has a leucine zipper motif between amino acids 69 and 90 which can work as a protein interaction domain that allows for proteins to either hetero- or homodimerize [32], thus, the protein could bind to *cis*-regulatory elements of gene promoters either as a monomer or as a dimer. Attempting to resolve if this protein works as a monomer or as a dimer, the dimensions of the free and DNA-complexed proteins were analyzed. Proteins appear with similar dimensions in diameter as well as in height within the precision of the technique. From the molecular mass of the protein (55 kDa) and assuming a spherical shape, we can deduce a radius of about 5 nm. The mean height of the proteins analyzed was 3.8 ± 0.4 nm. Thus, by comparing both values, we suggest that these proteins appear as monomers in both cases, when complexed to DNA as well as when adsorbed on mica.

Direct visualization of these DNA-protein complexes by AFM constitutes a conceptually easy and reproducible method for confirming that the protein is a transcription factor

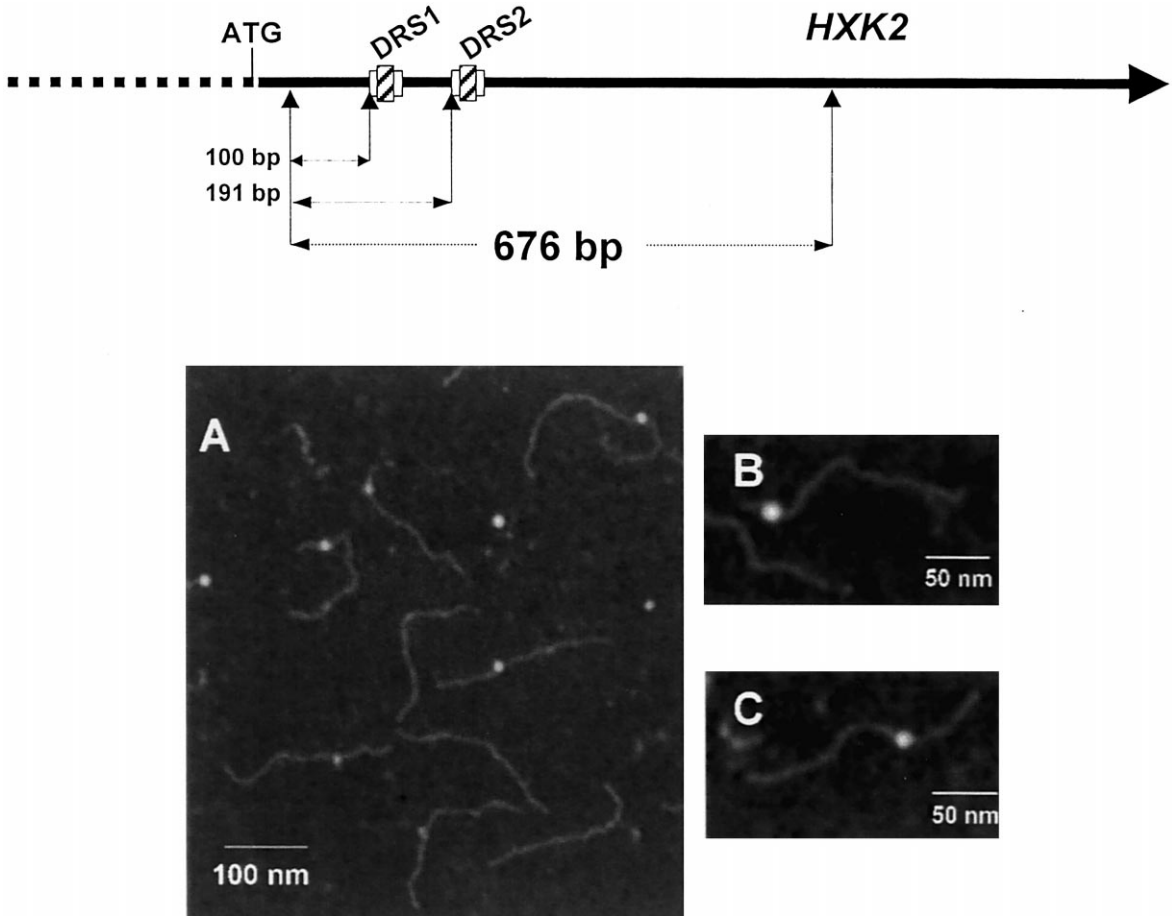

Fig. 3. Schematic picture of the *HXK2* gene showing the two DRSs regions, the 7 bp motif and its relative positions from the 5′ end of the 676 bp DNA fragment utilized. (A) AFM image of complexes of the 676 bp DNA fragments of the *HXK2* gene with the GST-Med8 protein. B and C: examples of the two types of DNA-protein complexes observed. For details, see Section 2.

that directly binds to DNA. In addition, AFM is a valuable tool for determining the position of the DNA region coated by the protein. In this work, we have showed that Med8 binds through sequences containing a 7 bp motif with the consensus sequence (A/C)(A/G)GAAAT. We have also found that only one binding site per DNA molecule is occupied and that the protein affinity is the same for the two binding sites. We estimate that the protein works as a monomer based on a very simple calculation. To make further progress in this field, a sharper tip would be needed.

Acknowledgements: Moreno-Herrero F. was supported by a fellowship from FICYT (Asturias, Spain) and Colchero J. by a contract of the Spanish Ministry of Education and Culture. This work was supported by Grants PB97-1213-C0202 and PB95-0169 from the DGESIC.

References

[1] Conaway, R.C., Reines, D., Garrett, K.P., Powell, W. and Conaway, J.W. (1996) Methods Enzymol. 273, 194–207.
[2] Koleske, A.J. and Young, R.A. (1994) Nature 368, 466–469.
[3] Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M.H. and Kornberg, R.D. (1994) Cell 77, 599–608.
[4] Myers, L.C., Gustafsson, C.M., Bushnell, D.A., Lui, M., Erdjument-Bromage, H., Tempst, P. and Kornberg, R.D. (1998) Genes Dev. 12, 45–54.
[5] Young, R.A. (1991) Annu. Rev. Biochem. 60, 689–715.
[6] Corden, J.L. (1990) Trends Biochem. Sci. 15, 838–840.
[7] Barberis, A., Pearlberg, J., Simkovich, N., Farrel, S., Reinagel, P., Bamdad, C., Sigal, G. and Ptashne, M. (1995) Cell 81, 359–368.
[8] Chaves, R.S., Herrero, P. and Moreno, F. (1999) Biochem. Biophys. Res. Commun. 254, 345–350.
[9] Sarokin, L. and Carlson, M. (1986) Mol. Cel. Biol. 6, 2324–2333.
[10] Bu, Y. and Schmidt, M.C. (1998) Nucleic Acids Res. 26, 1002–1009.
[11] Martínez-Campa, C., Herrero, P., Ramírez, M. and Moreno, F. (1996) FEMS Lett. 137, 69–74.
[12] Herrero, P., Ramírez, M., Martínez-Campa, C. and Moreno, F. (1996) Nucleic Acids Res. 24, 1822–1828.
[13] Binning, G., Quate, C.F. and Gerber, C. (1986) Phys. Rev. Lett. 56, 930–933.
[14] Shao, Z., Mou, J., Czajkowsky, D.M., Yang, J. and Yuan, J.Y. (1996) Adv. Phys. 45, 1–86.
[15] Bustamante, C., Rivetti, C. and Keller, D.J. (1997) Curr. Opin. Struct. Biol. 7, 709–716.
[16] Engel, A., Schoenenberger, C.A. and Müller, D.J. (1997) Curr. Opin. Struct. Biol. 7, 279–284.
[17] Yokota, H., Nickerson, D.A., Trask, B.J., Engh, G., Hirst, M., Sadowski, I. and Aebersold, R. (1998) Anal. Biochem. 264, 158–164.
[18] Guthold, M. et al. (1994) Proc. Natl. Acad. Sci. USA 91, 12927–12931.
[19] Van Noor, S.J.T., van der Werf, K.O., Eker, A.P.M., Wyman, C., de rooth, B.G., van Hulst, N.F. and Greve, J. (1998) Biophys. J. 74, 2840–2849.

[20] Kasas, S., Thomson, N.H., Smith, B.L., Hansma, H.G., Zhu, X., Guthold, M., Bustamante, C., Kool, E.T., Kashlev, M. and Hansma, P.K. (1997) Biochemistry 36, 461–468.
[21] Lyubchenko, Y.L. and Shlyakhtenko, L.S. (1997) Proc. Natl. Acad. Sci. USA 94, 496–501.
[22] Hansma, H.G., Laney, D.E, Bezanilla, M., Sinsheimer, R.L. and Hansma, P.K. (1995) Biophys. J. 68, 1672–1677.
[23] Hansma, H.G., Revenko, I., Kim, K. and Laney, D.E. (1996) Nucleic Acids Res. 24, 713–720.
[24] Bustamante, C., Zuccheri, G., Leuba, S.H., Yang, G. and Samori, B. (1997) Methods Enzymol. 12, 73–83.
[25] Rippe, K., Mücke, N. and Langowski, J. (1998) Nucleic Acids Res. 25, 1736–1744.
[26] Allen, M.J., Bradbury, E.M. and Balhorn, R. (1997) Nucleic Acids Res. 25, 2221–2226.
[27] Delain, E., Fourcade, A., Poulin, J.C., Barbin, A., Coulaud, D., Le Cam, E. and Paris, E. (1992) Microsc. Microanal. Microstruct. 3, 457–470.
[28] Révet, B. and Fourcade, A. (1998) Nucleic Acids Res. 26, 2092–2097.
[29] Fernández, E., Moreno, F. and Rodicio, R. (1992) Eur. J. Biochem. 204, 983–990.
[30] Margeat, E., LeGrimellec, C. and Royer, C.A. (1998) Biophys. J. 75, 2712–2720.
[31] Vesenka, J., Guthold, M., Tang, C.L., Keller, D., Delaine, E. and Bustamante, C. (1992) Ultramicroscopy 42, 1243–1249.
[32] Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988) Science 240, 1759–1764.