

# Imaging and Mapping Protein-Binding Sites on DNA Regulatory Regions with Atomic Force Microscopy

Fernando Moreno-Herrero,\* Pilar Herrero,† Jaime Colchero,\*  
Arturo M. Baró,\* and Fernando Moreno†

†*Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Biotecnología de Asturias, Universidad de Oviedo, 33006 Oviedo, Spain; and \*Laboratorio de Nuevas Microscopías, Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

Received November 27, 2000

**Regulation of gene expression is fundamental in biological systems. A systematic search for protein binding sites in gene promoters has been done in recent years. Biochemical techniques are easy and reliable when analysing protein interactions with short pieces of DNA, but are difficult and tedious when long pieces of DNA have to be analysed. Here we propose AFM as a reliable and easy technique for identifying protein interaction sites in long DNA molecules like gene promoters. We support this idea using a well-known model: the interaction of the Pho4 protein with the *PHO5* gene promoter. We have also applied the technique to demonstrate that Mig1 protein binds to two motifs in the promoter of *HXX2* gene. Our results allow us to define Mig1p as a new factor probably contributing to the carbon source-dependent transcription regulation of *HXX2* gene.** © 2001 Academic Press

**Key Words:** Atomic force microscopy; *Saccharomyces cerevisiae*; *PHO5*; *HXX2*; Mig1p; transcriptional regulation.

Since the invention of the Atomic Force Microscopy (AFM) (1) (also called Scanning Force Microscopy) in 1986, the technique has experienced an enormous growth and its use is, nowadays, routinely in many different fields such as physics, chemistry, biology, biochemistry, and molecular biology. AFM allows to image the surface of biological macromolecules (2) and their structures in native conditions, in three dimensions, without staining or shadowing, in air or under liquid conditions (see for reviews: 3–5). AFM has shown spectacular results not only imaging but also as a nanolithographic tool (6), for measuring antigen-antibody binding and unbinding forces (7, 8) and folding and unfolding protein forces (9). The technique has enough resolution to resolve DNA molecules in air and under physiological conditions. It has been also proved that small proteins are also visible

with AFM, and that interactions between DNA and proteins can be imaged (10–14). On the other hand, biochemical techniques can identify protein–DNA-specific interactions easily when the DNA molecules involved are of only a few dozens of base pairs, but the mapping of big DNA pieces of hundreds of base pairs is a tedious and difficult work. In this paper we propose the use of the AFM as a complementary tool to the classical biochemical techniques for identifying DNA-protein interaction sites when large fragments of DNA (hundreds of bp) have to be analysed. In this field the AFM can compete with techniques like DNaseI footprint, methylation interference or uracil interference assays in terms of simplicity, velocity, cost, and precision.

To illustrate this idea we have used two transcription factors involved in the phosphate metabolism and glucose repression signalling of the yeast *Saccharomyces cerevisiae*. We have first mapped, as a control system, the interaction of the Pho4p activator with the regulatory elements of the *PHO5* promoter. This first experiment demonstrates that AFM can identify restriction sites in long pieces of DNA, and have allowed us to define the standard conditions of binding and the precision of the technique. Then we have demonstrated that the repressor Mig1p interacts with two Mig1p-motifs located in the *HXX2* promoter. We have, previously, shown that Mig1p binds to the consensus Mig1p-motifs by mobility shift assays. The results obtained are biologically meaningful.

Transcriptional activators and repressors stimulate and repress, respectively, gene expression by binding enhancer or repressor elements of gene promoters and contacting, either directly or indirectly, components of the RNA polymerase II transcriptional machinery (15–18).

The transcription of *PHO5*, which encodes a secreted acid phosphatase, is tightly repressed when *S. cerevisiae* is grown in phosphate-rich medium. On the other hand, when yeasts are starved of phosphate, *PHO5* induction its more than 100-fold increased (19). Tran-

scriptional induction of *PHO5* requires the transcription factor Pho4p. This protein binds to the CACGTK sequences of the *PHO5* promoter and its activity is negatively regulated by phosphorylation (20).

Glucose regulates carbon utilisation in *S. cerevisiae* mainly by inhibiting the transcription of numerous genes. The enzymes affected include those involved in gluconeogenesis, in the Krebs cycle, in respiration, and in the early steps of galactose and other sugars utilisation (21, 22). This mechanism, called glucose repression, ensure the use of glucose in preference to other carbon sources. Although the mechanism of glucose repression is not completely understood, several genes involved have been defined. One of these genes, *MIG1*, has found to be very important in the mechanism of glucose repression (23). *MIG1* encodes a C<sub>2</sub>H<sub>2</sub> zinc finger protein, with structural homology to mammalian Sp1, Egr (early growth response) and Wilm's tumour proteins. Mig1p can bind to several glucose-regulated promoters and, with Tup1p and Cyc8p (Ssn6p) represses transcription during glucose growth (24). Mig1p binds to the motif WWWWN(G/C)(C/T)GGG that resembles to the GC-box that recognize the mammalian proteins Sp1, Egr and Wilm's, but with the difference that for the binding of Mig1p an AT-rich region, which is located in 5' of the GC-box, is also important (25). Studies on glucose repression of the *SUC2* and *GAL* genes identified Mig1p as a downstream factor in the Snf1 protein kinase signal-transduction pathway. Mig1 was found to repress those genes (26). Genome-wide monitoring of promoter region of genes and ORFs in yeast shows that almost 340 gene promoters contain the consensus sequence for Mig1p binding. However, the functionality of the protein in the control of expression of these genes is not still clear. In the last years, some progress has been achieved in this direction and Mig1p functional binding sites have been identified in the promoter region of *MAL63*, *MAL61*, *MAL62*, *FBP1*, and *HAP4* genes (27–29). However, much of the work remains and we think that the use of AFM to determine the functionality of Mig1 in those genes that contain the consensus binding sequence can be of invaluable help.

## MATERIALS AND METHODS

**Strains.** Bacterial transformation, large scale propagation of plasmid DNA and fusion protein expression were performed in *Escherichia coli* BL21(DE3)pLysS (Promega).

**Preparation of free DNAs.** Two different DNA fragments have been used to study specific DNA-protein interactions: (i) Oligonucleotides 5'-CAAGAGACTCCGTCCTT-3' (OL1) and 5'-TCATTTTCGAC-AATTCAAAGATG-3' (OL2) were used to generate by PCR a 775 bp DNA fragment of the *PHO5* promoter containing two 6 bp Pho4 protein motifs (CACGTK), located at 132 and 155 bp from one of the ends of the DNA fragment; (ii) oligonucleotides 5'-ACTGAACGCCATAG-AAGAGC-3' (OL3) and 5'-GATAAGACAGTGGCGAAGGT-3' (OL4) were used to generate by PCR a 570 bp DNA fragment of the *HXK2* promoter containing two putatives Mig1p binding motives (WWWWN-

SYGGGG), located at 128 and 163 bp from one of the ends of the DNA fragment.

The PCR was performed by using 0.1 µg of the primer pairs 1 + 2 or 3 + 4, 1 µg of genomic DNA as template, 2.5 U of *Taq* Polymerase (Promega), 0.2 mM dNTPs (Pharmacia) in a total reaction volume of 25 µl in reaction buffer provided by the manufacturer for 30 cycles at 94°C for 30 s, 55°C for 90 s, and 72°C for 60 s. In both cases, genomic DNA from the wild-type strain AMW-13C<sup>+</sup>, has served as the *PHO5* and *HXK2* promoters-containing template. The amplified products were purified using the high pure PCR product purification kit (Boehringer Mannheim) following the indications of the manufacturer.

**Subcloning *PHO4* gene into an *E. coli* expression vector.** A 939-bp DNA fragment of chromosome VI of *S. cerevisiae* containing the complete coding region of the *PHO4* gene was amplified from AMW-13C<sup>+</sup> genomic DNA by PCR, using as primers 5'-AAGGA-TCCATGGGCCGTACAACCTCTGA-3' (OL5) and 5'-GCAAGC-TTTCACGTGCTCACGTTCTGCT-3' (OL6). The PCR product was cleaved with *Bam*HI and *Hind*III, and inserted into an *Bam*HI/*Hind*III previously cleaved pQE-30 plasmid. The resulting plasmid (pQE30-PHO4), was used to transform *E. coli*. The clone used was verified by sequence analysis of fusion points. The pQE-30 plasmid is a His<sub>6</sub>-tag gene fusion vector (Quiagen) and was used to produce a His<sub>6</sub>-tagged Pho4 protein.

**Subcloning *MIG1* gene into an *E. coli* expression vector.** A 1515-bp DNA fragment of chromosome VII of *S. cerevisiae* containing the complete coding region of the *MIG1* gene was amplified from AMW-13C<sup>+</sup> genomic DNA by PCR, using as primers 5'-AAGGATCCATGCAAAGCCCATATCCAATG-3' (OL7) and 5'-GC-GGATCCTCAGTCCATGTGTGGGAAGG-3' (OL8). The PCR product was cleaved with *Bam*HI and inserted into an *Bam*HI previously cleaved pGEX-4T plasmid. The resulting plasmid (pGEX-MIG1), was used to transform *E. coli*. The clone used was verified by restriction and sequence analysis of fusion points. The pGEX-4T plasmid is a GST (glutathione S-transferase) gene fusion vector (Pharmacia) and was used to produce a GST-Mig1 fusion protein. As control, the original vector (pGEX-4T) containing the GST gene was also used to produce the glutathione S-transferase protein.

**Media, growth conditions, and heterologous expression of *S. cerevisiae* *PHO4* and *MIG1* genes in *E. coli*.** *E. coli* cells containing the expression plasmids pQE30-PHO4 or pGEX-MIG1 were grown at 37°C in 100 ml TB-phosphate medium containing ampicillin (100 µg/ml) to an optical density of 0.6 at 600 nm. At this point, isopropyl-1-thio-D-galactopyranoside (IPTG) was added to the medium to a final concentration of 0.2 mM, and the culture temperature was lowered to 30°C. After 3 h, the cells were harvested and washed once by centrifugation with a sodium phosphate solution (20 mM, pH 7.5) that contained NaCl (150 mM), the cell pellet was resuspended in 10 ml of lysis buffer (150 mM NaCl, 20 mM KCl; 101 mM Na<sub>2</sub>HPO<sub>4</sub>, 18 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.3) and kept on ice until the evaluation of fusion protein production had been performed. With this aim, 10 µl aliquots of the cells suspension were boiled in 30 µl of SDS-PAGE loading buffer for 3 min and production of the fusion protein was determined by SDS-PAGE analysis. Analysis of the proteins by Western blotting also confirmed that the 34-kDa protein band detected in Coomassie stained gels reacts with antibodies against Pho4p and that the 81-kDa protein band detected in Coomassie stained gels reacts with antibodies against GST (data not shown). Once it was demonstrated that the pQE30-PHO4 and pGEX-MIG1 transformed *E. coli* cells synthesised, respectively, the His<sub>6</sub>-Pho4 and GST-Mig1 fusion proteins in response to the presence of IPTG in the culture medium, the cells were lysed by sonication in 10 ml of lysis buffer per g of wet cell pellet in the presence of protease inhibitor (2 mM phenylmethylsulphonyl fluoride). Cell debris were removed by centrifugation at 17,000g for 30 min at 4°C, the supernatant was used as crude protein extract.

**Purification of His<sub>6</sub>-tagged Pho4 protein.** The His<sub>6</sub>-tagged Pho4 protein was purified by incubating 1 ml of the bacterial crude lysate containing the fusion protein with 0.5 ml of Ni-NTA agarose (Quia-

gen), in a slurry 1:1 in binding buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , pH 8.0, 300 mM NaCl, 20 mM imidazole), for 60 min at 4°C with gentle rocking and then poured into a 1 ml column. The column was washed five times with 1 ml of binding buffer at 4°C. The specifically bound proteins were eluted with 0.5 ml of elution buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , pH 8.0, 300 mM NaCl, 200 mM imidazole) and 100  $\mu\text{l}$  fractions were collected. Eluted proteins in the different fractions were analysed by Coomassie blue staining after sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

**Purification of GST-Mig1 fusion protein.** The fusion protein GST-Mig1 was purified by incubating 1 ml of the bacterial crude lysate containing the fusion protein with 0.5 ml of glutathione-Sepharose beads, in a slurry 1:1 in binding buffer (150 mM NaCl, 20 mM KCl; 100 mM  $\text{Na}_2\text{HPO}_4$ , 18 mM  $\text{NaH}_2\text{PO}_4$ , pH 7.3), for 60 min at 4°C with gentle rocking and then poured into a 1 ml column. The column was washed five times with 1 ml of binding buffer at 4°C. The specifically bound proteins were eluted with 0.5 ml of elution buffer (50 mM Tris-HCl, pH 8.5 containing 10 mM glutathione) and 100  $\mu\text{l}$  fractions were collected. Eluted proteins in the different fractions were analysed by Coomassie blue staining after SDS-PAGE.

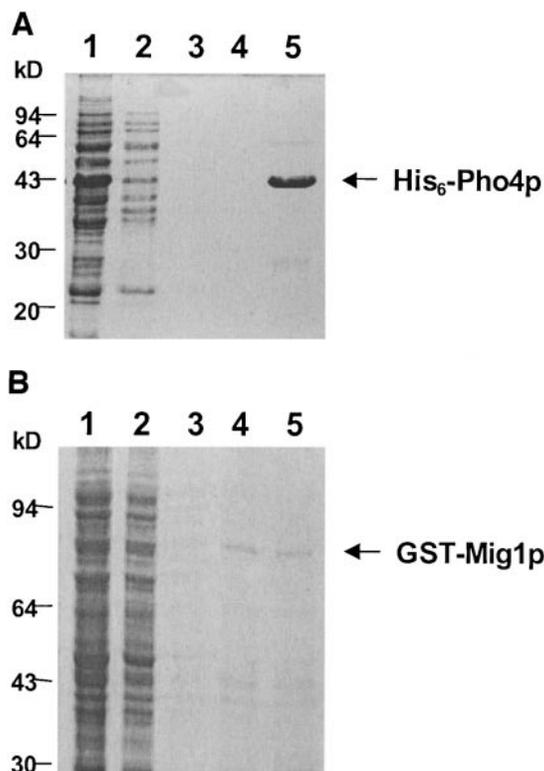
**DNA probes.** Oligonucleotides, corresponding to both strands of the Mig1p binding motives located in the *HXK2* gene promoter, were synthesized with an added TCGA nucleotide overhang at the 5'-terminal end: *HXK2*<sub>MIG1</sub> sense, 5'-tcgaAAAAAAGTCGGGG-3' and *HXK2*<sub>MIG1</sub> antisense, 5'-tcgaCCCCGCACTTTTTT-3'.

The complementary strands were annealed and either end-labelled with [ $\alpha$ - $^{32}\text{P}$ ]dCTP using the Klenow fragment of DNA polymerase I or used as unlabelled competitors in protein binding experiments.

**Gel retardation assay.** The typical binding reactions contained 10 mM Hepes (pH 7.5), 1 mM dithiothreitol, 1 to 5 mg of poly(dI-dC) and 0.5 ng of end-labelled probe DNA in a volume of 25  $\mu\text{l}$ . Six microliters (10  $\mu\text{g}$ ) of purified GST-Mig1 fusion protein was added. The amount of unlabeled competitor DNA added is indicated in the figure legends. After 30 min of incubation at room temperature the binding reaction mixtures were loaded onto a 4% nondenaturing polyacrylamide gel. Electrophoresis was allowed to proceed at 10 V/cm of gel for 45 min to 1 h in 0.5 $\times$  TBE buffer. Gels were dried and autoradiographed at -70°C with an intensifying screen.

**AFM sample preparation.** AFM images of bare DNA were performed first of all in order to check the purity and concentration of the sample. A final concentration of 5–15 adsorbed DNA molecules per  $\mu\text{m}^2$  was used for all the experiments. Typical binding reactions contained 50 mM Hepes (pH 7.5), 1 mM dithiothreitol, 80 mM NaCl, the DNA fragment and the protein under study in a volume of 25  $\mu\text{l}$ . DNA molecules-protein ratios were adjusted to 5/1 keeping the concentration of adsorbed DNA molecules as mentioned before. After 30 min of room temperature incubation, 3  $\mu\text{l}$  of the binding reaction was mixed with 1  $\mu\text{l}$  of 110 mM  $\text{MgCl}_2$  to assist binding to mica through  $\text{Mg}^{2+}$  ions. The final 4  $\mu\text{l}$  were deposited on a freshly cleaved mica, allowed to bind for a minute, rinsed with  $\text{d}_2\text{H}_2\text{O}$ , and dried with a gentle stream of nitrogen gas.

**AFM imaging.** AFM images were obtained with a commercial microscope (Nanotec Electronica S.L., Spain) operating in noncontact tapping mode following the conditions described in detail elsewhere (30). To image the surface in noncontact AFM, the cantilever is oscillated at its free resonance frequency (typically 80 kHz) and the normal force signal is processed to measure the amplitude and the relative phase of the oscillation. Our digital control electronics then establishes feedback at a certain oscillation amplitude which is slightly lower than the free resonance amplitude. As discussed in detail elsewhere (31, 32), we found that for appropriate working conditions, which are essentially using small oscillation amplitude (5–10 nm), very small reduction factors for the set point ( $r = 0.8$ – $0.99$  where  $r = 1$  corresponds to the free amplitude), and soft cantilevers (<1 N/m), it is possible to work without mechanical contact between tip and sample. We note that with these experimen-



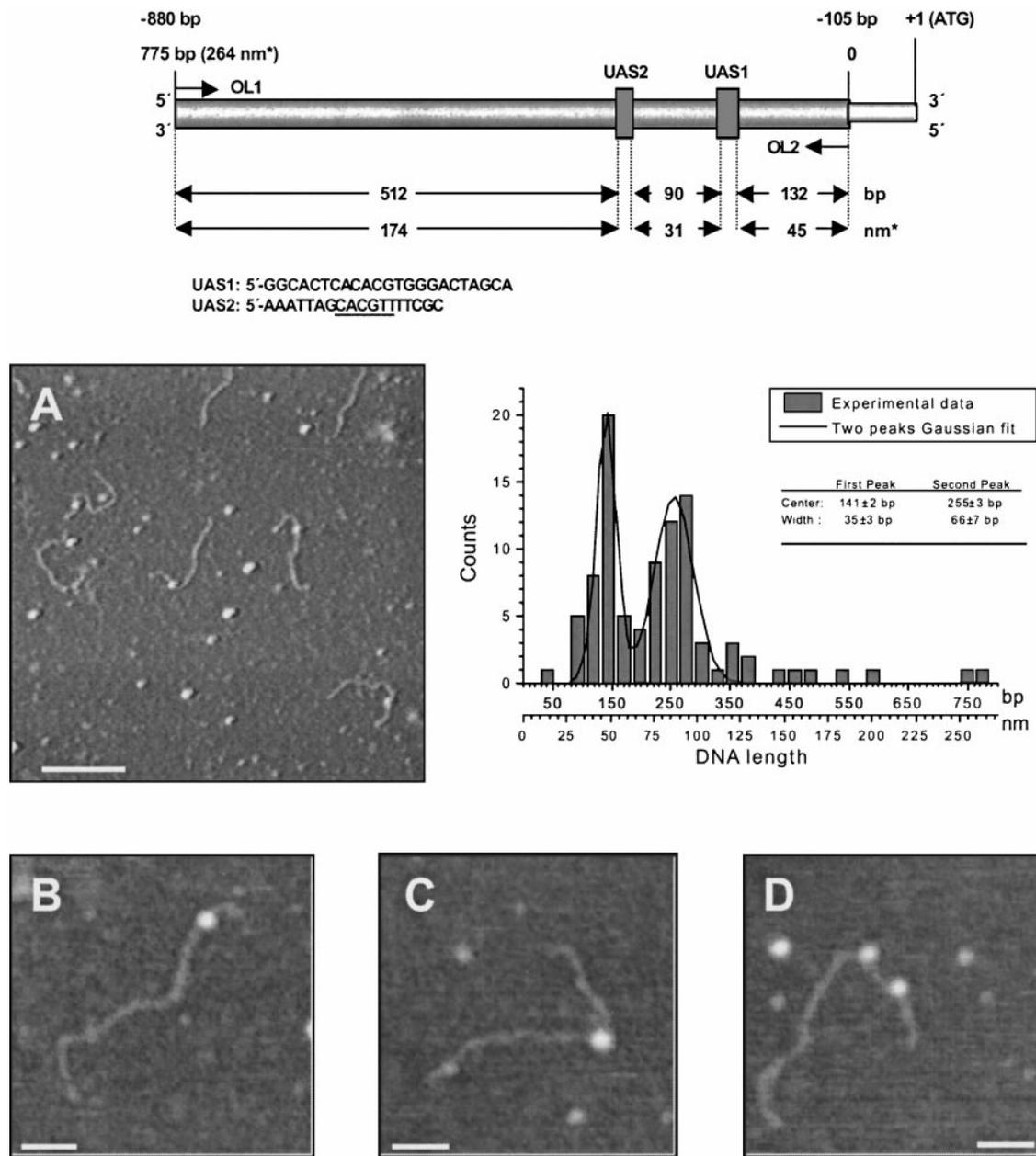
**FIG. 1.** Purification of His<sub>6</sub>-Pho4p (A) and GST-Mig1 fusion protein by affinity chromatography (B). (A) Proteins were eluted from Ni-NTA agarose with binding buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , pH 8.0, 300 mM NaCl, 20 mM imidazole) until no proteins were detected (lane 4), then the specifically bound proteins were eluted with 0.5 ml of elution buffer (50 mM  $\text{NaH}_2\text{PO}_4$ , pH 8.0, 300 mM NaCl, 200 mM imidazole) and 100- $\mu\text{l}$  fractions were collected (lane 5). The arrow points to the 43-kDa His<sub>6</sub>-Pho4 protein. (B) Proteins were eluted from glutathione-Sepharose beads with binding buffer (150 mM NaCl, 20 mM KCl; 100 mM  $\text{Na}_2\text{HPO}_4$ , 18 mM  $\text{NaH}_2\text{PO}_4$ , pH 7.3) until no proteins were detected (lane 3), then 500  $\mu\text{l}$  of elution buffer (50 mM Tris-HCl, pH 8.5 containing 10 mM glutathione) were added and 100  $\mu\text{l}$  fractions were collected (lane 4 and 5). The arrow points to the 81-kDa GST-Mig1 fusion protein.

tal parameters, the mode usually termed "tapping mode" should rather be named "noncontact dynamic AFM," and that tip-sample contact, and thus damage to the sample, can be excluded. For our experiments the microscope was operated in an airtight chamber in order to modify and control the RH by introducing dry or humid nitrogen. Typically, the microscope was operated in ambient conditions at rather low relative humidity (30%). Some images were processed by subtracting a general plane to remove the background slope and filtered to eliminate the low frequency noise.

## RESULTS AND DISCUSSION

### Purification of His<sub>6</sub>-Tagged Pho4 and GST-Mig1 Fusion Protein

The expression plasmids PQE30-PHO4 and pGEX-MIG1, which contained respectively the full-length *PHO4* and *MIG1* genes, were constructed as indicated under Materials and Methods. Sequence analysis confirmed that *PHO4* and *MIG1* genes were fused, respectively, in frame with a His<sub>6</sub>-tag and the glutathione



**FIG. 2.** Schematic picture of the *PHO5* promoter showing the two UASs regions and histogram of the position of Pho4p in the Pho4p-*PHO5* complexes analysed ( $n = 100$ ). (A) AFM image of complexes of the 775-bp DNA fragments of *PHO5* promoter with the (His)<sub>6</sub>-tagged Pho4 protein (bar size is 200 nm). (B, C, and D) Examples of the three types of DNA-protein complexes observed (bar size is 50 nm). For details see Materials and Methods.

*S*-transferase. SDS-PAGE and Western blot analysis confirmed that the resultant full-length *PHO4* and *MIG1* expression plasmids, respectively, expressed the recombinant proteins His<sub>6</sub>-Pho4 (~43 kDa) and GST-Mig1 (~81 kDa).

Once it was demonstrated that the PQE30-*PHO4* and pGEX-*MIG1* transformed *E. coli* cells synthesise the His<sub>6</sub>-Pho4 and GST-Mig1 proteins in response to the presence of IPTG in the culture medium, our goal was to purify the recombinant proteins. The nickel and glutathione binding activity of the His<sub>6</sub>-Pho4 and GST-

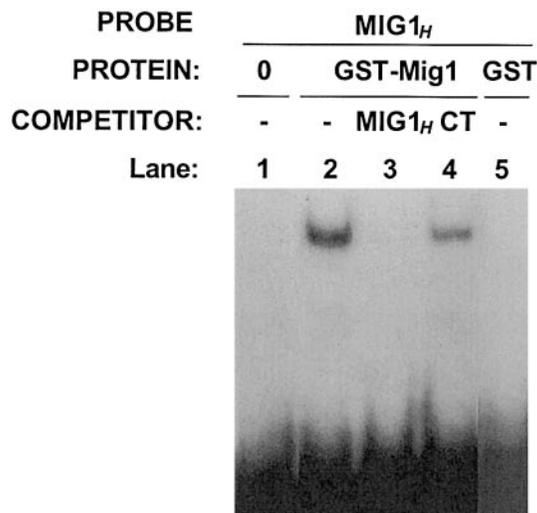
Mig1 proteins was examined by affinity chromatography (Fig. 1). The His<sub>6</sub>-Pho4 and GST-Mig1 proteins were bound, respectively, to Ni-NTA agarose and glutathione-Sepharose beads using a binding buffer (see Material and Methods). A high number of protein bands were observed by Coomassie staining in the two first fractions eluted with binding buffer (Figs. 1A and 1B, lanes 1 and 2), but in the last wash fraction no more protein bands were detected (Fig. 1A, lanes 3 and 4; Fig. 1B, lane 3). The bound protein was eluted with an elution buffer containing 200 mM imidazole and a

43-kDa protein was detected (Fig. 1A, lane 5) or 10 mM glutathione and a 81-kDa protein was detected (Fig. 1B, lanes 4 and 5). The identity of the proteins was confirmed by Western blotting, with antibodies against Pho4p and GST (data not shown).

#### *Images of Pho4p-DNA Complexes and Mapping of the Protein along the DNA*

A 775-bp DNA fragment of the *PHO5* promoter and an His<sub>6</sub>-tagged Pho4 protein were used to study their capacity to form DNA-protein complexes by AFM (Fig. 2). A typical image of Pho4p, which was purified by affinity chromatography on Ni-NTA-agarose (Fig. 1A), interacting with *PHO5* promoter DNA is shown in Fig. 2A. Under the binding conditions described in the previous section, most of the DNA molecules appear with no protein attached. Those with a protein attached, were analysed in the same manner of previous works (13). Three types of DNA-protein complexes have been found as can be seen in Figs. 2B, 2C, and 2D. Complexes of type B and C are very common and were detected with equal probability. Complex of type D is not very often due to the low protein concentration employed. Experiments with higher concentrations of protein were not possible because of the obvious surface character of the SFM: an excess of protein will completely cover the surface so that imaging of the DNA adsorbed on the mica is extremely difficult. The experimental data are consistent with the idea that two protein complexes appear only when saturation of one site is achieved.

The position of the Pho4p along the DNA fragment was determined by measuring the contour length of the DNA molecule from the centre of the Pho4p to its closest DNA end. In this way a background error due to non-specific binding is always introduced because it is not possible to identify which of the two strand ends is analysed, but right complexes are always correctly placed because the protein motifs are asymmetrically located in the DNA molecules. Statistical analysis of the complexes position show two gaussian peaks distribution. The two peaks are centered at  $48 \pm 1$  nm and  $87 \pm 1$  nm from one end of the DNA molecule in complete agreement with the position of the two UAS elements of the *PHO5* promoter (33). The widths at half height of the peaks are 12 nm (~35 bp) and 22 nm (~65 bp) respectively. This value gives us the resolution in positioning the protein in the DNA molecule. Our results show that His<sub>6</sub>-Pho4p is able to form correct DNA-protein complexes in agreement with the position of the two regulatory elements of the *PHO5* promoter. This result confirms AFM data published previously where left and right-handed configurations of His<sub>6</sub>-RNAP-DNA complexes were equally populated (34). This experiment also demonstrates that the AFM technique is a very adequate technique for imaging and mapping protein-DNA interaction using large DNA frag-



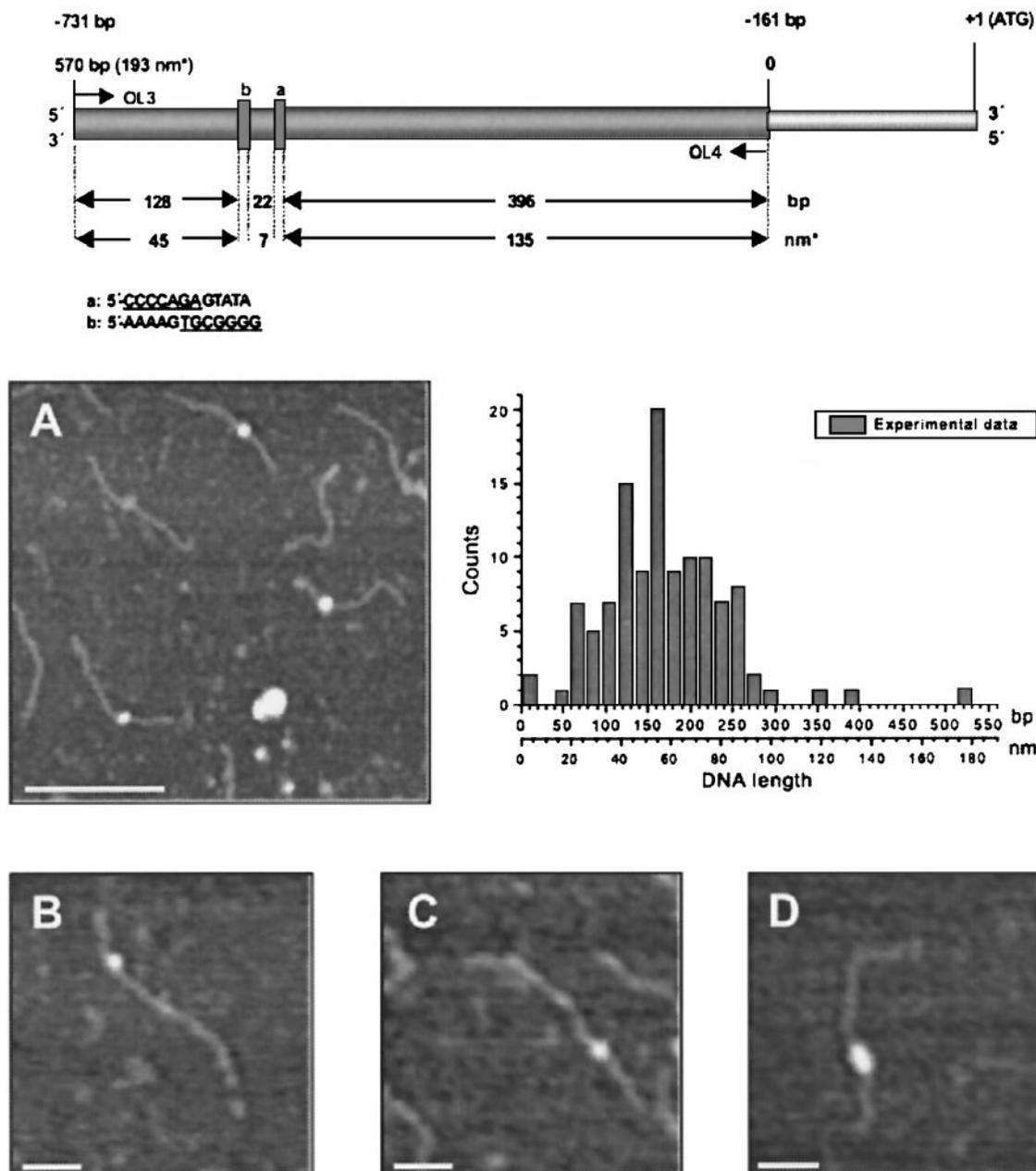
**FIG. 3.** Gel retardation analysis of complexes formed with the MIG1<sub>HXX2</sub> sequence. Each reaction included 0.5 ng of <sup>32</sup>P MIG1<sub>HXX2</sub> (lanes 1 to 5) DNA probe and, except for lane 1 and 5, 6 μl (10 μg) of GST-Mig1 purified fusion protein. For lanes 2 to 4 the protein GST-Mig1 was obtained by elution with 10 mM glutathione buffer from a GSH-affinity chromatography column while for lane 5 the protein eluted with the glutathione based buffer was GST. The competitor for binding was 50 ng of unlabeled MIG1<sub>HXX2</sub> (lane 3) or calf thymus DNA (lane 4). For the control lane 5, 10 μg of GST purified protein were used.

ments. The resolution in positioning proteins in DNA templates is mainly limited by the tip radius, and by the nonspecific binding which always introduces a background noise. Typical tip radius are between 10 and 20 nm, which is consistent with the widths at half height of the gaussian peaks found. The non-specific binding is always present and depends on the binding conditions.

#### *Images of Mig1p-DNA Complexes and Mapping of the Protein along the DNA*

Once we have demonstrated that AFM can identify protein-DNA interactions in large DNA-fragments with enough resolution, we have studied whether the GST-Mig1 fusion protein can bind to the MIG1 motives present in the *HXX2* promoter. We expressed the *MIG1* yeast gene in *E. coli* as a fusion protein with GST, thus eliminating the possibility of contamination by other yeast proteins. We first analysed if the heterologous synthesised fusion protein binds specifically to a double stranded synthetic oligonucleotide containing the MIG1 motif. Then, we studied if GST-Mig1p also binds to the MIG1 motives present in the *HXX2* promoter.

<sup>32</sup>P-labelled MIG1<sub>HXX2</sub> (MIG1<sub>H</sub>) probe was incubated with the purified GST-Mig1 fusion protein and subjected to native polyacrylamide gel electrophoresis as described under Materials and Methods. As can be seen in Fig. 3, the presence of the GST-Mig1 fusion protein changed the mobility of the radiolabeled probe (Fig. 3; lane 2). Specific and non-specific binding of



**FIG. 4.** Schematic picture of the *HXK2* promoter showing the two putative Mig1p binding motives and histogram of the position of Mig1p in the Mig1p-*HXK2* complexes analysed ( $n = 117$ ). (A), AFM image of complexes of the 570 bp DNA fragments of *HXK2* promoter with the GST-Mig1 fusion protein (bar size is 200 nm). (B, C, and D) examples of the three types of DNA-protein complexes observed (bar size is 50 nm). For details see Materials and Methods.

GST-Mig1 fusion protein to the MIG1<sub>*HXK2*</sub> oligonucleotides were analysed by competition assays with the corresponding non-labelled oligonucleotides (Fig. 3; lane 3) or with unspecific calf thymus DNA (Fig. 3; lane 4). The presence of GST protein does not change the mobility of the radiolabeled probe (Fig. 3; lane 5).

Once demonstrated, by band shift experiments, that the GST-Mig1 fusion protein binds to the MIG1 motif present in the *HXK2* promoter. We were interested in the analysis of this protein-DNA interaction in the

whole *HXK2* promoter context. Thus we have used SFM to study the capacity of a 570 bp DNA fragment of the *HXK2* promoter and a GST-Mig1p to form DNA-protein complexes (Fig. 4). A typical image of Mig1p interacting with *HXK2* promoter DNA is shown in Fig. 4A. Similar experiments to those described above for His<sub>6</sub>-Pho4p and *PHO5* promoter, were performed with GST-Mig1p and the promoter of the *HXK2* gene. In this case, the two putative Mig1p binding motives are separated only by 22 bp. Despite this fact, AFM images

prove that Mig1p bind to both sites with equal probability (Figs. 4B and 4C). Two-proteins complexes have also been found but with much lower probability (Fig. 4D). This fact can be explained in the terms pointed out before. The statistical analysis of the protein-DNA complexes ( $n = 117$ ) shows that the higher probabilities correspond to the intervals centered in  $42 \pm 3$  nm and  $55 \pm 3$  nm in complete agreement with the position of the Mig1p binding motives (Fig. 4, scheme of the *HXK2* promoter). In this experiment we measure a high non-specific binding background, probably due to some protein contamination. Purification of GST-Mig1p leads to a lower yield than His<sub>6</sub>-Pho4p.

We have demonstrated that AFM is a suitable technique for identifying protein binding sites in long DNA molecules like gene promoters. We have tested the viability of the idea with a well-known system like Pho4p-*PHO5*. Pho4p-binding sites were mapped in the promoter sequence of *PHO5* gene at base pairs  $-245$  to  $-255$  (UASp2), and  $-354$  to  $-375$  (UASp1) in complete accordance with previous results (33, 35). AFM allows to locate restriction sites with a resolution of tens of base pairs, mainly limited by tip radius and non-specific binding. Once the system parameters were known AFM mapping was applied to the *HXK2* promoter looking for Mig1p-DNA interactions. Mig1p-binding sites were identified in the promoter sequence of *HXK2* gene at base pairs  $-563$  to  $-571$ , called site "a," and  $-601$  to  $-611$ , called site "b." In both cases the nucleotide number indicate the distance from the gene ORF. The interaction between Mig1p and their binding motifs was demonstrated with mobility shift assays. The results were confirmed with AFM analysis in the whole promoter and allow us to define Mig1p as a new factor probably contributing to the carbon source dependent transcription regulation of *HXK2* gene (36).

## ACKNOWLEDGMENTS

We thank Susana Cuervo for generous help in performing some of the GST-Mig1p production experiments. F. Moreno-Herrero was supported by a fellowship from the Community of Madrid and J. Colchero 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

- Binnig, G., Quate, C. F., and Gerber, C. (1986) *Phys. Rev. Letters* **56**, 930-933.
- van Noort, S. J. T., van der Werf, K. O., Eker, A. P. M., Wyman, C., de Grooth, B. G., van Hulst, N. F., and Greve, J. (1998) *Biophys. J.* **76**, 2840-2849.
- Bustamante, C., Rivetti, C., and Keller, D. J. (1997) *Curr. Opin. Struct. Biol.* **7**, 709-716.
- Engel, A., Schoenenberger, C.-A., and Muller, D. J. (1997) *Curr. Opin. Struct. Biol.* **7**, 279-284.
- Shao, Z., Mou, J., Czajkowsky, D. M., Yang, J., and Yuan, Y. (1996) *Adv. Phys.* **45**, 1-86.
- Dai, H., Franklin, N., and Han, J. (1998) *Appl. Phys. Lett.* **73**, 1508.
- Ros, R., Schwesinger, F., Anselmetti, D., Kubon, M., Schäfer, R., Plückthun, A., and Tiefenauer, L. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 7402-7405.
- Schwesinger, F., Ros, R., Strunz, T., Anselmetti, D., Güntherodt, H.-J., Honegger, A., Jermutus, L., Tiefenauer, L., and Plückthun, A. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 9972-9977.
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M., and Gaub, H. E. (1997) *Science* **276**, 1109-1112.
- Bustamante, C., and Rivetti, C. (1996) *Annu. Rev. Biophys. Biomol. Struct.* **25**, 396-429.
- Muller-Reichert, T., and Gross, H. (1996) *Scanning Microsc. Suppl.* **10**, 111-120.
- Palecek, E., Vlk, D., Stankova, V., Brazda, V., Vojtesek, B., Hupp, T. R., Schaper, A., and Jovin, T. M. (1997) *Oncogene* **15**, 2201-2209.
- Moreno-Herrero, F., Herrero, P., Colchero, J., Baró, A. M., and Moreno, F. (1999) *FEBS Lett.* **459**, 427-432.
- Hansma, H. G., Pietrasanta, L. I., Auerbach, I. D., Golan, R., and Holden, P. A. (2000) *J. Biomater. Sci. Polym.* **11**, 675-683.
- Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., and Tjian, R. (1993) *Cell* **75**, 519-530.
- Joliet, V., Demma, M., and Prywes, R. (1995) *Nature* **373**, 632-635.
- Lin, Y., Nomura, T., Cheong, J., Dorjsuren, D., Iida, K., and Murakami, S. (1997) *J. Biol. Chem.* **272**, 7132-7139.
- Chaves, R. S., Herrero, P., and Moreno, F. (1999) *Biochem. Biophys. Res. Commun.* **254**, 345-350.
- Oshima, Y. (1982) in *The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression*, pp. 159-180 Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- O'Neill, E. M., Kaffman, A., Jolly, E. M., and O'Shea, E. K. (1996) *Science* **271**, 209-212.
- Gancedo, J. M. (1998) *Microbiol. Mol. Biol. Rev.* **62**, 334-361.
- Trumbly, R. J. (1992) *Mol. Microbiol.* **6**, 15-21.
- Ronne, H. (1995) *Trends Genet.* **11**, 12-17.
- Tzamaras, D., and Struhl, K. (1994) *Nature* **369**, 758-760.
- Lundin, M., Nehlin, J. O., and Ronne, H. (1994) *Mol. Cell. Biol.* **14**, 1979-1985.
- Smith, F. C., Davies, S. P., Wilson, W. A., Carling, D., and Hardie, D. G. (1999) *FEBS Lett.* **453**, 219-223.
- Wang, J., Sirenko, O., and Needleman, R. (1997) *J. Biol. Chem.* **272**, 4613-4622.
- Zaragoza, O., Rodriguez, C., and Gancedo, C. (2000) *J. Bacteriol.* **182**, 320-326.
- Bourgarel, D., Nguyen, C. C., and Bolotin-Fukuhara, M. (1999) *Microbiology* **31**, 1205-1215.
- Moreno-Herrero, F., de Pablo, P. J., Colchero, J., Gómez-Herrero, J., and Baró, A. M. (2000) *Surf. Sci.* **453**, 152-158.
- Luna, M., Colchero, J., and Baró, A. M. (1998) *Appl. Phys. Lett.* **72**, 3461.
- de Pablo, P. J., Colchero, J., Luna, M., Gómez-Herrero, J., and Baró, A. M. (2000) *Phys. Rev. B* **61**, 14179.
- Magbanua, J. P. V., Fujisawa, K., Ogawa, N., and Oshima, Y. (1997) *Yeast* **13**, 1299-1308.
- Rivetti, C., Guthold, M., and Bustamante, C. (1999) *EMBO J.* **18**, 4464-4475.
- Ogawa, N., and Oshima, Y. (1990) *Mol. Cell. Biol.* **10**, 2224-2236.
- Herrero, P., Galindez, J., Ruiz, N., Martínez-Campa, C., Moreno, F. (1995) *Yeast* **11**, 137-144.