Published online May 12, 2005

Nucleic Acids Research, 2005, Vol. 33, No. 8 ***e79***
doi:10.1093/nar/gni079

Proof of concept for microarray-based detection of DNA-binding oncogenes in cell extracts

Tanja Egener, Emmanuelle Roulet, Marc Zehnder, Philipp Bucher[1] and Nicolas Mermod*

Institute of Biotechnology, University of Lausanne, 1015 Lausanne, Switzerland and [1]Swiss Institute of Bioinformatics and Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland

Received February 4, 2005; Revised April 12, 2005; Accepted April 27, 2005

ABSTRACT

The function of DNA-binding proteins is controlled not just by their abundance, but mainly at the level of their activity in terms of their interactions with DNA and protein targets. Moreover, the affinity of such transcription factors to their target sequences is often controlled by co-factors and/or modifications that are not easily assessed from biological samples. Here, we describe a scalable method for monitoring protein–DNA interactions on a microarray surface. This approach was designed to determine the DNA-binding activity of proteins in crude cell extracts, complementing conventional expression profiling arrays. Enzymatic labeling of DNA enables direct normalization of the protein binding to the microarray, allowing the estimation of relative binding affinities. Using DNA sequences covering a range of affinities, we show that the new microarray-based method yields binding strength estimates similar to low-throughput gel mobility-shift assays. The microarray is also of high sensitivity, as it allows the detection of a rare DNA-binding protein from breast cancer cells, the human tumor suppressor AP-2. This approach thus mediates precise and robust assessment of the activity of DNA-binding proteins and takes present DNA-binding assays to a high throughput level.

INTRODUCTION

Transcription factors control most cellular processes by their coordinated expression and binding to specific regulatory DNA sequences. They therefore possess a great potential for mis-regulation leading to developmental aberrations or tumors. A large class of cancer-linked genes can be assigned to the DNA-binding proteins. Protein-binding microarrays (PBMs) have recently been successfully used to analyze the binding specificity (1) or to identify the binding sites genome-wide (2) of purified transcription factors. In this study, we developed a highly sensitive and robust functional assay for DNA-binding proteins in crude cellular extracts, using double-stranded DNA microarrays. This approach has very broad application, and here we have focused on the assessment of the binding of tumor-suppressors and oncogenes from cell or tissue extracts. In particular, we demonstrate the feasibility of this approach using the human transcription factor and tumor suppressor AP-2.

AP-2 is of special interest because of its versatile and essential functions in cell growth, cell differentiation and programmed cell death (3–7). AP-2 family members have been implicated in the progression, vascularization, metastasis and/or recurrence of tumors, and they constitute good prognostic factors for melanoma, breast and prostate cancers (8–11). Recent studies have revealed that the effect of AP-2 on cancer prognosis is not simply a question of expression level, but rather is defined by the regulation of the AP-2 protein itself, for instance by its interaction with other proteins (12) or its subcellular localization (9). Alternatively, it is the ratio of AP-2 relative to other competing transcription factors that is the relevant prognostic indicator, as it controls the expression of downstream cancer-linked genes (10,11). These post-translational effects are not readily detectable using conventional large-scale approaches such as expression profiling or proteomics, and the association of such proteins to the relevant target sequences must be probed directly from cell or tissue extracts. Here we propose a novel application for the high-throughput analysis of protein–DNA-binding activity with PBMs, allowing the rapid, sensitive and scalable detection of relative binding activities and specificities directly from cellular extracts.

*To whom correspondence should be addressed at Laboratory of Molecular Biotechnology, Station 6, FSB-ISP-ISIC, EPFL, 1015 Lausanne, Switzerland. Tel: +41 21 693 61 51; Fax: +41 21 693 76 10; Email: Nicolas.Mermod@unil.ch

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

MATERIALS AND METHODS

Generation of the AP-2 profile

The AP-2 profile constructed for this work (Figure 1A) is based on the SELEX dataset published by Mohibullah and co-workers (13) comprising 70 *in vitro* selected protein-binding sites. To estimate the weights of the profile we applied a computational protocol developed for the modeling of the CTF/NF-1 transcription factor (14), using programs from the software packages SAM version 1.3.3 (R. Hughey and A. Krogh, http://www.cse.ucsc.edu/research/compbio/sam.html) and pftools version 2.2 (P. Bucher, ftp://ftp.isrec.isb-sib.ch/pub/sib-isrec/pftools/). As part of this procedure, the sequences were realigned using an expectation-maximization algorithm (15). This analysis supported only the common spacing classes N3 and N4 (N1 and N2 in our model) but not the spacing classes N5 and N6 considered by the authors of the data and the half-site binding mode. The program pfsearch from the pftools package was used to compute relative binding scores for potential binding site sequences with the aid of the profile shown in Figure 1A.

Construction of DNA fragments

Double-stranded 25 bp DNA molecules containing AP-2 bindings sites were generated by hybridization of two complementary oligonucleotides (Supplementary Table 1) and cloned into pBAS. This vector corresponds to pBluescriptSKII+ (Amersham Bioscience), whose multiple cloning site (MCS) was modified by the deletion of SmaI and ApaI sites (Supplementary Table 1). Hybridized oligonucleotides were inserted into a HincII, EcoRV or NotI site and correct cloning and orientation were confirmed by sequence analysis.

Double-stranded DNA fragments for microarray spotting were generated by PCR amplification using standard protocols with increased concentrations (1 µM) of primers. Amplification with primers 120L (5′-GCTCCACCGCGGTGCG-3′) and 120R (5′-CTATAGGGCGAATTGGGTAC-3′) yielded fragments of 120 bp length (Figure 1B). Primer 120L carried a 5′-amino-modification for covalent linkage to the array surface. PCR products were purified by isopropanol precipitation in the presence of 0.6 M NH_4-acetate. The dried pellet was re-suspended in H_2O, and the DNA concentration was determined using Sybr Green I (Molecular Probes).

Gel shift assay

Probes for the gel shift assays were radioactively labeled using T4 polynucleotide kinase (10 U) and 5 µl γATP for 15 min at room temperature (RT) for 200 ng double-stranded DNA, i.e. oligonucleotide or PCR product. Probes were separated from unincorporated ATP by size exclusion using a Sephadex g50 matrix, and stored in aliquots at −20°C until use. Unless stated otherwise, the assay mix per lane contained 2 ng radiolabeled

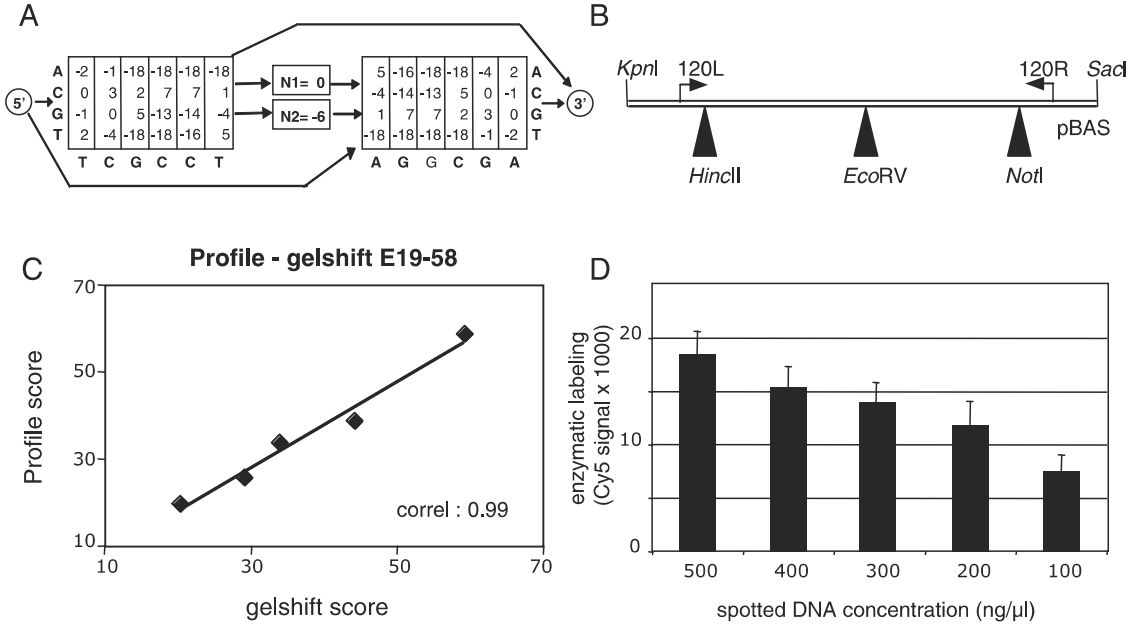

Figure 1. Bioinformatics prediction profile and microarray standardization tools for AP-2 binding studies. (**A**) Architecture of the AP-2 profile. The model was constructed from *in vitro* DNA binding sequences (13), as described in Materials and methods. It is composed of two palindromic half-sites, each containing 6 nt positions, with allowed spacing between the half-sites of 1 or 2 bp. The position-specific weights associated with the four possible nucleotides and with the two possible spacings are indicated within boxes. The resulting consensus sequence is given below. Arrows indicate alternative paths corresponding to different binding modes. The score of a potential binding site is computed as the sum of the corresponding weights. (**B**) Construction of microarray DNA probes. Double-stranded oligonucleotides coding for potential AP-2 binding sequences of various predicted affinities, with profile score of 19, 25, 33, 38 or 58, were cloned into the constant backbone of pBAS using restriction sites HincII, EcoRV or NotI (black triangles). PCR amplification with primers 120L and 120R (arrows) produced 120 bp products that were used to spot microarrays. (**C**) Correlation of computer-predicted binding scores with experimental binding values. Affinity scores computed from the AP-2 profile shown in part (A) were plotted against experimental gel shift binding values expressed as gel shift scores (see Materials and methods). The correlation coefficient of the profile score and gel shift binding values is 0.99. (**D**) Enzymatic quantification of microarray-spotted DNA. DNA solutions of indicated concentrations were spotted onto a microarray surface, immobilized DNA probes were subjected to enzymatic Cy5 labeling and fluorescence was scanned. Background fluorescence was subtracted from all values, which represent the average fluorescence and SD from four experiments.

probe (O58), 20 ng unlabeled probe (O58), 2 μg poly(dI–dC), 0.5× TEN and 175 ng GST-AP-2 in 1× AP-2 binding buffer. The mix was incubated at RT for 1 h to reach equilibrium of the binding reaction before separation on a 4% polyacrylamide gel. Autoradiograms were analyzed and quantified using a Storm 240 (Amersham Bioscience) and ImageQuant software version 1.2 (Molecular Dynamics). The probe occupancy was subsequently determined as the ratio of shifted band signal over total signal (Supplementary Figure 1). Absolute affinity measurements were performed using competitive gel shift experiments and calculated from Scatchard plots as described previously (14,16). Relative binding affinities were determined in competition assays, where the apparent affinity was estimated by calculating the competitor-to-probe molar ratio resulting in 50% probe occupancy (IC_{50} ratio). Relative binding affinities were converted to gel shift binding scores and normalized to that obtained for probe O58 using the principles described by Roulet *et al*. 2002 (16), according to the following calculation: gel shift score = $58 - \log_{10}(IC_{50}\ \text{ratio}) \times 50$.

Protein expression and purification

Human AP-2α full-length cDNA was cloned into GST-fusion vector pGEX3X (Amersham Bioscience), resulting in plasmid pGEX-AP-2, and overexpressed as a GST-fusion protein in *Escherichia coli* W3110. Protein expression was induced by the addition of IPTG during the late exponential growth phase and cells were harvested 2 h later. Cells were disrupted by sonication and GST-AP-2 enriched from soluble protein fraction by binding to a Glutathione Sepharose 4B column (Amersham Bioscience). After washing the pure fusion protein was eluted in 100 mM Glutathione, yielding typically 350 ng/μl GST-AP-2 according to the Bradford assay (BioRad Inc.), and aliquots were stored at −80°C.

Protein-binding microarrays

DNA concentration was adjusted to 250 μg/ml, unless otherwise noted, and mixed 1:1 with 2× spotting buffer (6× SSC; 3 M betaine). Each DNA was spotted in quadruplicate on different positions of the array. Spotting was performed using an ESI arrayer (BioRad) at 50–55% humidity and 23°C onto Nexterion aldehyde slides (Schott). Slides were blocked by washing twice for 2 min in TEN [40 mM Tris–HCl (pH 7.5), 1 mM EDTA, 150 mM NaCl], 10 min in aldehyde blocking solution, twice for 2 min in TEN, 1 h in 5% powdered skim milk in 1× phosphate-buffered saline (PBS) with 0.05% Tween-20 and twice for 2 min in TEN. Use of a high salt buffer such as TEN is required to retain the double-stranded structure of the nucleic acids on the slide surface.

The DNA on the array was fluorescently labeled using terminal transferase (TdT, NEB). Slides were pre-incubated in adenosine mono-phosphate (AMP, 100 ng/ml) in PBS for 1 h. Subsequently enzymatic treatment was carried out according to manufacturer's instructions in a total volume of 20 μl using 4 U TdT and 5 μM Cyanine5-ddATP (Perkin Elmer) under addition of 100 ng/μl AMP in the reaction mix. Slides were incubated at 37°C for 1.5 h, washed twice in TEN and spun dry.

The AP-2 binding assay was carried out essentially as described for use with Zn finger proteins (16), apart from the protein-binding buffer. Unless otherwise noted, 35 ng of purified GST-AP-2 was diluted in AP-2 binding buffer [2.5 mM Tris–HCl (pH 8.0), 15 mM KCl, 1 mM $MgCl_2$, 0.025 mM EDTA, 2.5% glycerol, 0.05% NP40, 50 mg/ml BSA] complemented with 2% powdered skim milk and 2 μg poly(dI–dC). After washing and antibody binding [first antibody: AP-2α (C-18): sc-184 from Santa Cruz Biotechnologies; second antibody: Cy3-labeled Alexa Fluor 546 goat anti-rabbit IgG (H + L) from Molecular Probes], the slides were spun dry and scanned using an Agilent G2566AA scanner (Supplementary Figure 2). Images were analyzed using ScanAlyze (www.rana.lbl.gov). After background signal subtraction, normalized binding activity was determined for each spot and defined as the Cy3 fluorescence signal (pixel) of the antibody for protein detection divided by the Cy5 fluorescence signal (pixel) of the enzymatic DNA labeling multiplied by 100 to correct for detection sensitivity differences between the two dyes. The ratio of microarray-bound DNA to protein on these arrays can be approximately calculated as follows: 35 ng GST-AP-2 binds to 150–200 spots of 1 nl (spot volume) with 125 pg/nl DNA on the general assumption that 20% of the DNA is covalently bound to the slide. This results in 8 ng GST-AP-2 binding 1 ng of DNA.

Cell culture and nuclear protein extractions

Hepatocellular carcinoma line HepG2 (ATCC-No. HB-8065) was cultured according to the supplier's instructions in Dulbecco's MEM with Glutamax-I (Gibco) containing 10% fetal bovine serum. MCF-7 breast carcinoma cells were cultured in RPMI 1640 medium with Glutamax-I (Gibco) with 10% fetal bovine serum. Transient transfections were carried out as co-transfections of the pEGFPN1 green fluorescent protein expression vector (BD Bioscience) and pCMX-AP-2 (17) using 3 μg of each plasmid for the transfection of 400 000 cells. Cells were tested for green fluorescent protein expression at day 2, reaching a transfection efficiency of usually 20–30%. Nuclear protein extractions were carried out according to standard techniques, yielding extracts of ∼1 mg total protein per milliliter as determined by a Bradford assay (BioRad).

RESULTS

Design of an AP-2 binding profile and DNA-binding sequences

Although a number of AP-2 binding sites have been identified in an *in vitro* binding site selection assay (13), an appropriate computer tool allowing the prediction of AP-2 binding sites and affinities is not yet available from these data. We therefore reanalyzed the published binding site data and constructed a quantitative binding site model based on a generalized profile (18). This type of descriptor and its construction method have been successfully applied to the quantitative prediction of the affinities of a transcription factor for DNA sequences of variable length, notably CTF/NF-1 (14) and members of the STAT protein family (19). The resulting binding site profile is shown in Figure 1A. AP-2α homodimer is known to interact with target sites of different length (13). It typically binds to a palindromic DNA sequence motif consisting of two reverse-complementary half-sites, separated by a short spacer of 1 or 2 bp (parameters N1 and N2 in the profile). Sequences

consisting of only a single half-site motif also appear to be weak binding targets according to our analysis. Each of the above-mentioned binding modes corresponds to a different profile path. The profile allows the computation of a relative binding score (supposed to be proportional to log K_d values) for a potential target site by adding up the weights corresponding to the bases at each position and the weight corresponding to the spacer class. For instance, the AP-2α profile shown in Figure 1A assigns a score of 58 to the consensus binding sequence with highest affinity, TCGCCTXAGGCGA, and a score of 29 to the optimal half-site, TCGCCT or AGGCGA. The medium-affinity binding sites referred to in this paper fall in the range of scores between 33 and 38, and low-affinity sites score from 19 to 25.

In order to assess the specific DNA-binding affinity of AP-2 experimentally, and to distinguish it from low-affinity background binding to any unspecific sequence, we introduced known binding sites into a common and defined DNA sequence background. We used a derivative of the pBluescript SKII+ plasmid, pBAS, for this purpose, first removing potential AP-2 binding sites that might interfere with our assay from the polylinker. Such sites were predicted by both the MatInspector (www.gene-regulation.com) and present (ISREC, www.isrec.isb-sib.ch) computer tools. Five 25 bp oligonucleotides containing binding sites of different profile scores (58, 38, 33, 25 and 19) were cloned into pBAS at each of the three restriction sites indicated in Figure 1B. Two PCR primers termed 120L and 120R were designed to amplify the 120 bp portion of the plasmid encompassing the binding sites, with the 120L primer incorporating an amino (N) link for immobilization onto the microarray surface. Binding sites introduced at the HincII restriction site are thus positioned close to the 120 bp fragment extremity used to attach the double-stranded DNA to the microarray surface; EcoRV marks the middle of the fragment, and NotI lies at the far end of the PCR product generated with the 120L/R primers (Figure 1B). All possible 120mer PCR-amplified probes were first tested for their affinity to AP-2 in gel electrophoresis mobility-shift assays prior to spotting on microarrays (Supplementary Figure 1), using purified AP-2 expressed as a fusion to the GST protein in *E.coli*. Binding affinities were found to range between 3 nM and 8 mM on the oligonucleotide probes, as estimated from Scatchard plot analysis. Interaction with the 120mer probes containing the binding sites in the central location correlated well with the binding scores determined from the profile (Figure 1C). Similar findings were obtained using probes with the binding sequences inserted closer to the probe extremities (data not shown), indicating that the position and DNA sequence context of the binding sites do not affect affinity in gel shift assays.

On-chip estimation of immobilized DNA amounts

To compare the specific binding affinity of AP-2 to various DNA sequences, the amount of DNA effectively linked to the array and available to proteins must be determined. Quantitative DNA staining using Sybr Green I, as in previous studies (1,2), remains impractical, as dyes that intercalate between DNA bases often interfere with protein binding. Thus, we designed a direct enzymatic labeling of the DNA immobilized on the array surface. The terminal transferase enzyme was used to add one fluorescently labeled Cyanine5-ddATP to the 3′ end of the DNA strands. The resulting fluorescent signal of the DNA spots is 200-fold higher than the low-level background fluorescence of the array surface. Linearity of the enzymatic DNA labeling was assessed on microarrays where all the fragments were spotted at concentrations ranging from 100 to 500 ng/μl, indicating precise and reproducible determination of the relative amounts of the immobilized DNA (Figure 1D). Although the relationship between the concentration of the spotted DNA and the fluorescent labeling of the spots on the microarray is fairly linear, fluorescence is not fully proportional to the amount of spotted DNA. This may indicate that only a fraction of the spotted DNA is available for protein interaction once bound to the microarray surface or it might reflect saturation effects.

Sequence-specific binding of AP-2 to microarray DNA probes

In order to assess whether AP-2 binds to double-stranded DNA immobilized on a microarray following the affinities determined for free binding sites, we assayed binding to the five sequences with known affinities (profile score 58–19), with binding sequences inserted at the three above-mentioned positions into the constant backbone of pBAS (Figure 1B). The normalized relative binding activity of the purified AP-2 revealed highly specific interaction, as empty pBAS sequences used as a negative control showed low (background) signal, and binding reflected the affinities of AP-2 to the target sequences as determined by gel shift assays (Figure 2A). Incubation of the protein with a free specific binding site as a competitor reduced signal-to-background values, as did use of non-specific control antibodies that do not recognize bound AP-2 for fluorescent quantification, further illustrating sequence-specific interaction of the protein with the immobilized probes (data not shown).

Estimating protein binding ‘on a chip’ is what characterizes surface-bound binding assays. We considered the possibility that, in this type of assay, having one of the binding partners immobilized on a small surface might skew apparent binding affinities because of a high local concentration. Therefore we tested several possibly influential parameters to probe further the general applicability of this technology as well as to assess its ability to provide both qualitative and quantitative data.

To determine whether steric hindrance might restrict accessibility to the surface-bound DNA, the binding sites were tested at the different positions described before, H58–H19 being closest to the microarray-linked residue, E58–E19 being in central position and N58–N19 being in end-standing position (Figure 1B). Terminal positions revealed higher binding signals than the same sites in a central position (E58–E19). Lower binding to the central position of the DNA immobilized on the microarray may indicate that DNA linking to the surface occurred not solely by the terminal N-link, but also by amine residues within the DNA bases, as shown previously for DNA immobilization on aldehyde surfaces (20). Overall, binding to PCR-amplified probes was similar to that observed with the shorter oligonucleotides (O58–O19), except for probes which do not display significant binding affinity, where binding was somewhat higher with the longer probes (Figure 2A). Amplification of DNA probes of various length indicated that

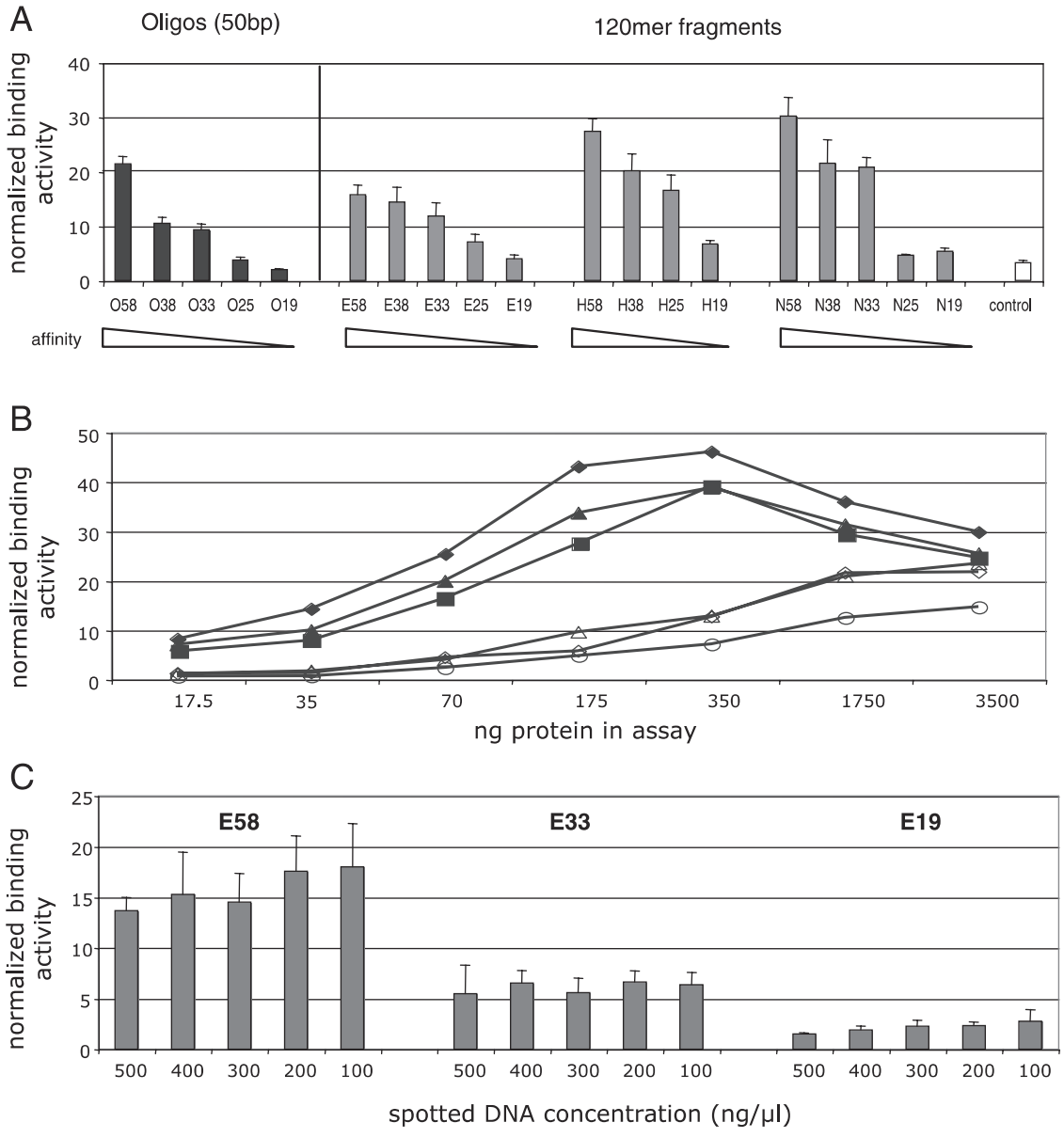

Figure 2. AP-2 binding to double-stranded DNA microarray probes. (**A**) AP-2 DNA-binding specificity. Microarrays were spotted with probes in the form of 50 bp oligonucleotides (O58–O19, black columns) or of 120 bp PCR products (grey columns) generated as described in the legend to Figure 1. Binding sites are located either in central (EcoRV site, E58–E19) or in terminal (HincII, H58–H19 and NotI, N58–N19) position on the 120 bp probes. Probe numbers indicate the profile score of the binding sites computed as in Figure 1. Microarrays were incubated with purified GST-AP-2 fusion and bound proteins were detected with Cy3-labeled polyclonal antisera as described in Materials and methods. Binding activities determined from the Cy3 fluorescence were normalized to the Cy5 signal corresponding to the spotted DNA. The white column indicates background binding to a control PCR fragment into which no AP-2 site has been inserted. Values depict the average and SD of three data points. (**B**) Protein dose–response curves. Normalized binding activities were determined as above using the indicated amounts in nanograms of GST-AP-2 protein. High-affinity sequences (matrix values >33) are given as closed symbols (diamonds N58, triangles N38, squares N33), and low-affinity sequences (matrix values <25) are shown as open symbols (triangles N25, diamonds N19, circles control). (**C**) Binding activity as a function of DNA-spotting concentration. DNA probes containing high- (E58), intermediate- (E33) or low-affinity (E19) sites were spotted in various concentrations as in Figure 1D and incubated with 70 ng of purified GST-AP-2. The average and SD of three data points of normalized binding activities are displayed as for (A).

this residual binding increases with probe length, as expected from non-specific interactions (data not shown).

Normalized binding to the high- and medium-affinity probes increased with protein concentration and reached saturation with ~350 ng of purified protein, illustrating the specific interaction of AP-2 with its binding sites (Figure 2B). Fragments with little or no affinity (profile scores of 25 and lower) also exhibited a gradual increase of signal with rising protein concentrations but did not reach saturation, as expected from non-specific interactions. Using amounts of AP-2 corresponding to 50% occupancy of the probe, we estimate that 8 ng of AP-2 binds 1 ng DNA with the microrarray, and 16 ng is required with the gel shift assay. Thus, the two types of assays are of similar sensitivity.

Next we assessed whether AP-2 would access and bind equally well to various amounts of immobilized nucleic acids

in the standard range of DNA concentrations used for spotting (100–500 ng/µl), or whether 'crowding effects' would be observed. Binding normalized to the amount of DNA effectively immobilized on the microarray shows little bias over the concentration range (Figure 2C and Supplementary Figure 2). This indicates that a similar fraction of the DNA is accessible over the range of DNA concentrations, regardless of the actual affinity of the site, and that robust measurements can be achieved irrespective of the spotting conditions.

AP-2 binding to microarrays correlates with profile prediction and gel shift results

To further evaluate the quality of the microarray-assessed AP-2 binding to DNA, binding values were compared with those predicted *in silico* from the profile score or those determined by the gel shift assay. For the 120 bp fragments with target sequences in central position (E58–E19), high correlation coefficients were obtained when comparing the microarray data with the profile score (0.96, Figure 3A) or with binding values from gel shift experiments (0.97, Figure 3B). Microarray data also compared well with *in silico* determined profile scores when the binding sites were in end-standing positions (H58–H19 and N58–N19), with correlation coefficients of 0.98–90.99 (Figures 3C and D). The higher correlation coefficients obtained from probes displaying binding sites at their extremity are consistent with the higher binding affinities observed in microarray measurements.

Sequence-specific DNA binding was further confirmed by conducting microarray competition assays in analogy to the gel shift competition assays described above. Low amounts (5 ng) of the specific competitor DNA O58 abolished detectable binding (Supplementary Figure 3A), while 1 µg of non-specific competitor DNA (O19) was unable to abolish binding (Supplementary Figure 3B). These competition assays indicate that binding to immobilized DNA on a microarray reflects specific protein–DNA interaction by AP-2 as for the gel shift assay, while providing a much higher data yield as multiple DNA sequences are probed simultaneously.

Detection of AP-2 from extracts of breast cancer cells

Downregulation of AP-2 activity and/or its re-localization from the cell nucleus to the cytoplasm correlate with increased tumor malignancy and unfavorable prognosis for a wide variety of cancer types, including breast tumors (8–11). We therefore assessed whether the microarray technique might be sensitive enough to detect low levels of AP-2 binding activity from tumor cell nuclear extracts. In a first approach, we transiently co-transfected HepG2 cells with CMV-promoter driven AP-2 expression vector, as these cells express little or no endogenous AP-2 (21). Nuclear proteins were extracted and used for binding studies on microarrays as described above. Whereas non-transfected cells showed little signal above background as expected, transfected HepG2 cells showed a binding pattern over the various DNA probes that was similar to that observed using the purified AP-2 protein (Figure 4A). Mixing of the nuclear extract of non-transfected cells with the purified protein also yielded binding patterns and levels similar to those obtained with the purified protein alone, indicating that

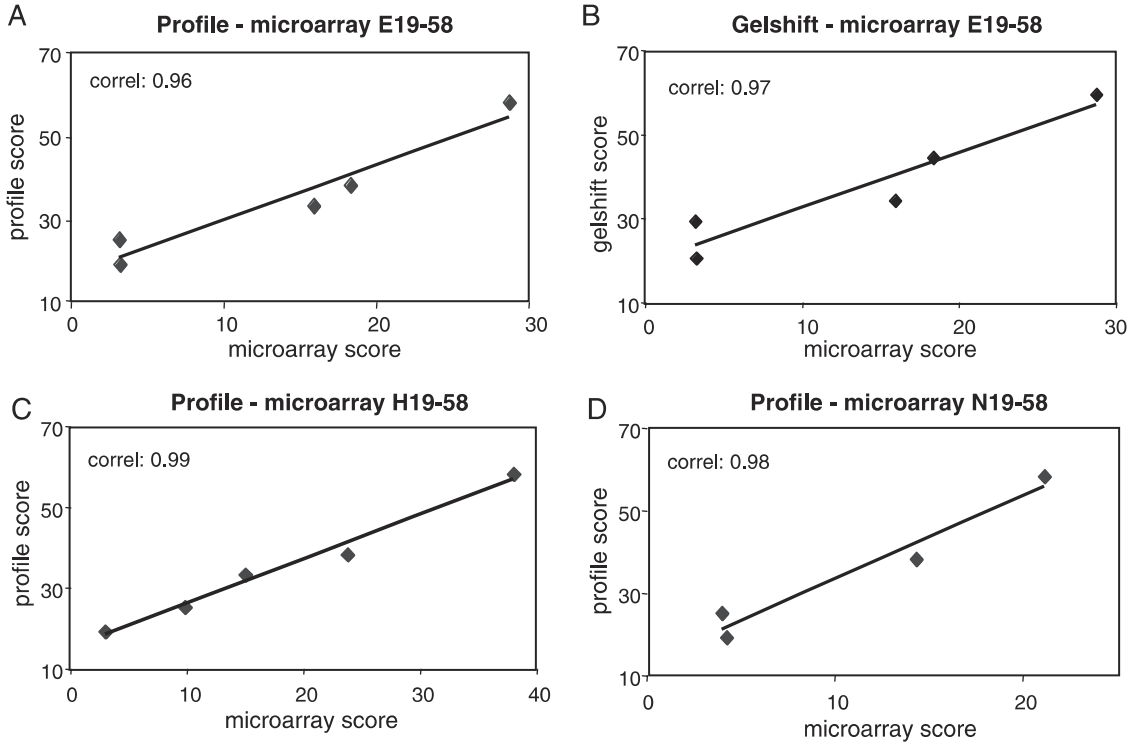

Figure 3. Comparisons of relative AP-2 binding activities using *in silico*, *in vitro* and on-chip assays. Comparison between profile prediction scores and microarray binding values using E19–E58 (**A**), H19–H58 (**C**) and N19–N58 (**D**) probes, determined as detailed in the legends to Figures 1 and 2. (**B**) displays a comparison of gel shift and microarray binding values for the E19–E58 probes. Correlation coefficients are given in the top left corner of (A–D).

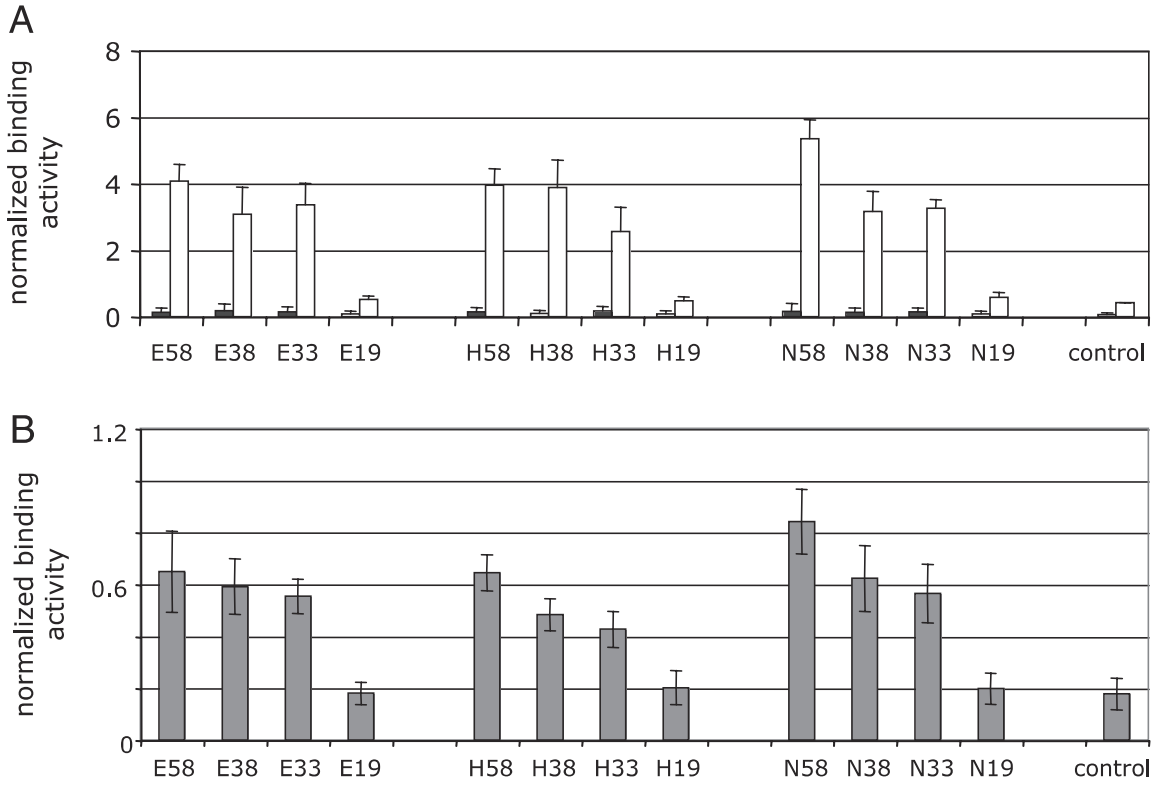

Figure 4. Microarray-based detection of sequence-specific AP-2 DNA-binding activity from human cell nuclear extracts. Normalized binding activities were determined for the indicated microarray probes incubated with AP-2-containing nuclear extracts. Nuclear extracts consisted of 2 µg total proteins of HepG2 hepatocellular carcinoma cells, either non-transfected (closed bars) or transiently transfected with an AP-2 expression vector (open bars) [four data points, (**A**)], or of 10 µg total proteins from untransfected MCF-7 mammary gland carcinoma cells [$n = 20$, (**B**)].

the affinity of AP-2 for the different probes is not affected by other cellular proteins. In further titrations, the detection limit for AP-2 in the presence of total nuclear protein was set to a 2-fold signal increase between high-affinity binding sites (E58) and the noise obtained with non-binding control probes (Supplementary Figures 2B and 4). The detection limit for AP-2 was found to lie around ~100–200 pg, corresponding to ~2 fmol of the purified protein in the assay. Therefore, meaningful and sensitive measurements can be performed from crude nuclear cell extracts.

Next, the naturally occurring endogenous levels and affinities of AP-2 were assessed in MCF-7 human breast cancer cells, representing a tumor type where the AP-2α family member used here is known to play a prominent role in carcinogenesis. Although transcription factors are of low abundance, the endogenous level was readily detectable using 10 µg of total nuclear extract in a volume of 20 µl (Supplementary Figure 2B). The relative binding affinities of non-purified MCF-7 AP-2 to the various specific and non-specific probes were analogous to those observed for the HepG2-expressed or purified proteins (Figure 4B).

Considering that ~20 pg of total protein is obtained from one cell nucleus, there are ~200 000 AP-2 molecules per cell in this rather aggressive tumor type, as estimated from the calibration curve shown in Supplementary Figure 4. Thus, such microarrays may allow sensitive estimations of the amounts of active DNA-binding proteins from minute samples.

DISCUSSION

Gene expression is only one of the relevant aspects of cell physiology, another being the misregulation of protein-mediated cellular signaling pathways. The localization and activity of low-abundance oncogenes such as DNA-binding regulatory factors remain difficult to assess with DNA-only approaches. Although proteomic analysis provides useful complementary data, detection is still limited to a fraction of the proteins present in crude cell extracts. Several recent studies have used surface plasmon resonance (SPR) or PBM technologies in conjunction with microarray-based binding assays to analyze and model the DNA-binding sequence specificity of purified transcription factors (1,2,22,23). Here, we show that PBMs can also be used to develop sensitive and reliable techniques for functional assays of DNA-binding proteins from unfractionated cellular extracts.

The specificity of the assay can be assessed by comparing the pattern of binding of the purified protein with those obtained from transfected cells or with endogenous AP-2 in tumor cells. A similar DNA-binding pattern was obtained from all types of samples, illustrating the high specificity of the microarray-based assay. Detection specificity relies on the monoclonal antibody used to detect the transcription factor, coupled to the analysis of a collection of binding sites of known affinities to that protein. Therefore, spurious cross-reactivity of the antibody with another DNA-binding protein can be ruled out by analyzing the pattern of signals over a collection of DNA sequences of known affinities, and the specificity of the

detection method does not rely merely on the interaction of the antibody with the protein as in immunological assays.

The sensitivity of the assay allows the detection of very low protein amounts, i.e. femtomoles of the AP-2 protein. Also, a large number of DNA sequences can be analyzed in parallel from culture or tissue samples. Thus, a few nanograms of purified protein or micrograms of total protein extract suffice for an assay of several hundred data points, while one sequence is analyzed using low-throughput measurements such as gel shift assays. In the dynamic range of protein concentrations, microarray-mediated binding measurements are commensurate with the affinities determined by conventional assays or computed *in silico*. Thus, double-stranded DNA microarrays can provide an estimate of the relative affinities of a purified transcription factor for a collection of DNA sequences.

The AP-2 transcription factor analyzed in this study has a protective role in several cancer types. For instance, it acts by competing with activating factors for overlapping binding sites in the promoters of oncogenes. In human melanoma, transcriptional activator Sp1 is predominantly bound to the PAR-1 promoter, a contributor to tumor invasion and metastasis, whereas AP-2 occupies this promoter in non-metastatic cells (10). Similarly, AP-2 acts as a tumor suppressor in prostate cancer, where it competes with the transcriptional activator Sp3 for binding to the VEGF promoter, thereby repressing its expression (11). Changes in the localization of AP-2 also modulate tumorigenicity. AP-2 proteins can be relocated from the nucleus to the cytoplasm, and low nuclear levels of AP-2 are associated with disease progression and increased metastatic capability in breast cancer patients (9,12). PBM testing of the relative activities of AP-2 and competing proteins could therefore be used to rapidly analyze prevalent cancer types.

Cancer type identification is currently based mainly on histology, although identical histological features can be observed from tumors that progresses rapidly in one patient and slowly in another (e.g. prostate or breast cancer, or some types of sarcoma). PBMs carry a potential predictive value for functional/behavioral diagnosis with direct clinical impact.

Coupled to the detection of other DNA-binding proteins, more general applications for the analysis of normal and diseased cells can be envisioned. For instance, the aberrant activity, location or expression of various transcription factors has been linked to a growing number of diseases (24). In normal cells, transcription factors usually interface signaling pathways to the regulation of gene expression. This technique, if coupled to the analysis of a collection of DNA-binding proteins, may therefore probe the interface of the proteomic and genomic worlds, opening a window on the cell's regulome and providing a snapshot of its regulatory status.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr R. Buettner for kindly providing the AP-2 expression vector, Dr P. Descombes for custom-made microarrays and helpful discussions and Dr T. Williams and Dr I. Stamenkovic for useful comments on the manuscript. Equipment and expertise for microarray analysis were kindly provided by the DNA array facility at Lausanne. This research was supported by a special grant from the program in genomics of the University of Lausanne, the FNRS and by the ISREC, Switzerland. Funding to pay the Open Access publication charges for this article was provided by the University of Lausanne.

Conflict of interest statement. None declared.

REFERENCES

1. Linnell,J., Mott,R., Field,S., Kwiatkowski,D.P., Ragoussis,J. and Udalova,I.A. (2004) Quantitative high-throughput analysis of transcription factor binding specificities. *Nucleic Acids Res.*, **32**, e44.
2. Mukherjee,S., Berger,M.F., Jona,G., Wang,X.S., Muzzey,D., Snyder,M., Young,R.A. and Bulyk,M.L. (2004) Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nature Genet.*, **36**, 1331–1339.
3. Mitchell,P.J., Wang,C. and Tjian,R. (1987) Positive and negative regulation of transcription *in vitro*: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. *Cell*, **50**, 847–861.
4. Zeng,Y.X., Somasundaram,K. and el-Deiry,W.S. (1997) AP2 inhibits cancer cell growth and activates p21WAF1/CIP1 expression. *Nature Genet.*, **15**, 78–82.
5. Bosher,J.M., Totty,N.F., Hsuan,J.J., Williams,T. and Hurst,H.C. (1996) A family of AP-2 proteins regulates c-erbB-2 expression in mammary carcinoma. *Oncogene*, **13**, 1701–1707.
6. Jean,D., Gershenwald,J.E., Huang,S., Luca,M., Hudson,M.J., Tainsky,M.A. and Bar-Eli,M. (1998) Loss of AP-2 results in up-regulation of MCAM/MUC18 and an increase in tumor growth and metastasis of human melanoma cells. *J. Biol. Chem.*, **273**, 16501–16508.
7. Gaubatz,S., Imhof,A., Dosch,R., Werner,O., Mitchell,P., Buettner,R. and Eilers,M. (1995) Transcriptional activation by Myc is under negative control by the transcription factor AP-2. *EMBO J.*, **14**, 1508–1519.
8. Gee,J.M., Robertson,J.F., Ellis,I.O., Nicholson,R.I. and Hurst,H.C. (1999) Immunohistochemical analysis reveals a tumor suppressor-like role for the transcription factor AP-2 in invasive breast cancer. *J. Pathol.*, **189**, 514–520.
9. Pellikainen,J., Kataja,V., Ropponen,K., Kellokoski,J., Pietilainen,T., Bohm,J., Eskelinen,M. and Kosma,V.M. (2002) Reduced nuclear expression of transcription factor AP-2 associates with aggressive breast cancer. *Clin. Cancer Res.*, **8**, 3487–3495.
10. Tellez,C., McCarty,M., Ruiz,M. and Bar-Eli,M. (2003) Loss of activator protein-2alpha results in overexpression of protease-activated receptor-1 and correlates with the malignant phenotype of human melanoma. *J. Biol. Chem.*, **278**, 46632–46642.
11. Ruiz,M., Pettaway,C., Song,R., Stoeltzing,O., Ellis,L. and Bar-Eli,M. (2004) Activator protein 2alpha inhibits tumorigenicity and represses vascular endothelial growth factor transcription in prostate cancer cells. *Cancer Res.*, **64**, 631–638.
12. Aqeilan,R.I., Palamarchuk,A., Weigel,R.J., Herrero,J.J., Pekarsky,Y. and Croce,C.M. (2004) Physical and functional interactions between the Wwox tumor suppressor protein and the AP-2gamma transcription factor. *Cancer Res.*, **64**, 8256–8261.
13. Mohibullah,N., Donner,A., Ippolito,J.A. and Williams,T. (1999) SELEX and missing phosphate contact analyses reveal flexibility within the AP-2α protein: DNA binding complex. *Nucleic Acids Res.*, **27**, 2760–2769.
14. Roulet,E., Bucher,P., Schneider,R., Wingender,E., Dusserre,Y., Werner,T. and Mermod,N. (2000) Experimental analysis and computer prediction of CTF/NFI transcription factor DNA binding sites. *J. Mol. Biol.*, **297**, 833–848.
15. Durbin,R., Eddy,S., Krogh,A. and Mitchinson,G. (1988) Biological sequence analysis. In *Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, Cambridge, UK.
16. Roulet,E., Busso,S., Camargo,A.A., Simpson,A.J., Mermod,N. and Bucher,P. (2002) High-throughput SELEX SAGE method for quantitative modeling of transcription-factor binding sites. *Nat. Biotechnol.*, **20**, 831–835.

17. Hilger-Eversheim,K., Moser,M., Schorle,H. and Buettner,R. (2000) Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. *Gene*, **260**, 1–12.
18. Bucher,P., Karplus,K., Moeri,N. and Hofmann,K. (1996) A flexible motif search technique based on generalized profiles. *Comput. Chem.*, **20**, 3–23.
19. Ehret,G.B., Reichenbach,P., Schindler,U., Horvath,C.M., Fritz,S., Nabholz,M. and Bucher,P. (2001) DNA binding specificity of different STAT proteins. Comparison of *in vitro* specificity with natural target sites. *J. Biol. Chem.*, **276**, 6675–6688.
20. Diehl,F., Beckmann,B., Kellner,N., Hauser,N.C., Diehl,S. and Hoheisel,J.D. (2002) Manufacturing DNA microarrays from unpurified PCR products. *Nucleic Acids Res.*, **30**, e79.
21. Williams,T., Admon,A., Luscher,B. and Tjian,R. (1988) Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. *Genes Dev.*, **2**, 1557–1569.
22. Bulyk,M.L., Huang,X., Choo,Y. and Church,G.M. (2001) Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc. Natl Acad. Sci. USA*, **98**, 7158–7163.
23. Kyo,M., Yamamoto,T., Motohashi,H., Kamiya,T., Kuroita,T., Tanaka,T., Engel,J.D., Kawakami,B. and Yamamoto,M. (2004) Evaluation of MafG interaction with Maf recognition element arrays by surface plasmon resonance imaging technique. *Genes Cells*, **9**, 153–164.
24. Zaidi,S.K., Young,D.W., Choi,J.Y., Pratap,J., Javed,A., Montecino,M., Stein,J.L., van Wijnen,A.J., Lian,J.B. and Stein,G.S. (2005) The dynamic organization of gene-regulatory machinery in nuclear microenvironments. *EMBO Rep.*, **6**, 128–133.