

# Targeting the DNA-binding activity of the human ERG transcription factor using new heterocyclic dithiophene diamidines

Raja Nhili<sup>1</sup>, Paul Peixoto<sup>1</sup>, Sabine Depauw<sup>1</sup>, Sébastien Flajollet<sup>2</sup>, Xavier Dezitter<sup>1</sup>, Manoj M. Munde<sup>3</sup>, Mohamed A. Ismail<sup>3</sup>, Arvind Kumar<sup>3</sup>, Abdelbasset A. Farahat<sup>3</sup>, Chad E. Stephens<sup>3</sup>, Martine Duterque-Coquillaud<sup>2</sup>, W. David Wilson<sup>3</sup>, David W. Boykin<sup>3</sup> and Marie-Hélène David-Cordonnier<sup>1,\*</sup>

<sup>1</sup>INSERM UMR837-JPARC, Team 4, Molecular and Cellular Targeting for Cancer Treatment, University of Lille North of France, IMPRT-IFR114, Institut pour la Recherche sur le Cancer de Lille, Place de Verdun, Lille F-59045, France, <sup>2</sup>CNRS UMR 8161, Institut de Biologie de Lille, University of Lille North of France, Institut Pasteur de Lille IFR 142, Lille F-59021, France and <sup>3</sup>Department of Chemistry, Georgia State University, Atlanta, GA 30303-3083, USA

Received July 15, 2012; Revised September 24, 2012; Accepted September 25, 2012

## ABSTRACT

Direct modulation of gene expression by targeting oncogenic transcription factors is a new area of research for cancer treatment. ERG, an ETS-family transcription factor, is commonly over-expressed or translocated in leukaemia and prostate carcinoma. In this work, we selected the di-(thiophene-phenyl-amidine) compound DB1255 as an ERG/DNA binding inhibitor using a screening test of synthetic inhibitors of the ERG/DNA interaction followed by electrophoretic mobility shift assays (EMSA) validation. Spectrometry, footprint and biosensor-surface plasmon resonance analyses of the DB1255/DNA interaction evidenced sequence selectivity and groove binding as dimer. Additional EMSA evidenced the precise DNA-binding sequence required for optimal DB1255/DNA binding and thus for an efficient ERG/DNA complex inhibition. We further highlighted the structure activity relationships from comparison with derivatives. *In cellulo* luciferase assay confirmed this modulation both with the constructed optimal sequences and the Osteopontin promoter known to be regulated by

ERG and which ERG-binding site was protected from DNaseI digestion on binding of DB1255. These data showed for the first time the ERG/DNA complex modulation, both *in vitro* and in cells, by a heterocyclic diamidine that specifically targets a portion of the ERG DNA recognition site.

## INTRODUCTION

Development of therapeutic strategies for inhibiting transcription is of major interest for modulating gene expression associated with various diseases. Transcription factors are key regulators of gene expression, and their deregulation, direct or indirect, is often associated with oncogenesis, cancer development, invasiveness and metastasis. However, in spite of their important cancer generation/progression roles, transcription factors have not been extensively evaluated as targets for cancer treatment strategies (1,2). As transcription factors are considered as “undruggable” targets because of difficulty to directly modulate protein/DNA binding, most drug development strategies act at the protein–protein interaction or protein degradation levels. An example is the treatment of acute promyelocytic leukaemia expressing the fusion protein ProMyelocytic Leukemia–Retinoic Acid Receptor alpha

\*To whom correspondence should be addressed. Tel: +33 3 20 16 92 23; Fax: +33 3 20 16 92 29; Email: marie-helene.david@inserm.fr  
Present addresses:

Paul Peixoto, Metastasis Research Laboratory, GIGA Cancer, LIEGE SART TILMAN, 4000, Belgium.

Mohamed A. Ismail, Department of Chemistry, College of Science, King Faisal University, Hofuf 31982 Saudi Arabia.

Abdelbasset A. Farahat, Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University, Mansoura 35516, Egypt.

Chad E. Stephens, Department of Chemistry and Physics, Augusta State University, Augusta, GA 30904, USA.

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

using retinoid acid derivatives that target the DNA binding activity of the RAR moiety (3). Alternatively, other approaches were recently developed to target protein–protein interactions using structurally specific competitive drugs such as nutlin-3 that binds to MDM2 and avoids p53 degradation resulting from p53/MDM2 complex formation in numerous cancers (4). Another approach was developed to target transcription factor activities using compounds that block protein–DNA interactions such as S3I-201 inhibiting Stat3/DNA binding (5), the isoquinolone alkaloid compound berberine interfering with TATA binding protein (6) or synthetic polyamides, specifically designed for transcription factor/DNA modulation through their sequence-selective binding to the minor groove of the DNA helix (7). Such targeted transcription factor/DNA complexes include NF- $\kappa$ B, ETV1 and ETS-1, leading to a decrease in the expression of controlled genes (8–10). Non-specific DNA targeting is a major limitation to the development of transcription factor modulators as illustrated by echinomycin that targets both HIF-1 and Myc/Max transcription factors binding DNA (11). To bypass this drawback, identifying new DNA-binding compounds and evaluating them for DNA-binding selectivity using molecular studies are essential to obtain more effective DNA sequence-specific compounds.

With this aim, we focussed on the synthesis and DNA-binding activities of heterocyclic diamidines for directly targeting the DNA minor groove in a sequence-selective manner. Previous work highlighted the ability of the phenyl-furan-benzimidazole diamidine DB293 to inhibit Pit-1 and Brn-3 transcription factor/DNA complex (12). Because the used TransSignal protein/DNA array also evidenced a much smaller effect on transcription factor interactions to the ETS-binding site (EBS) (12), we then focussed on the modulation of transcription factors that interact with EBS. The minimal EBS core is the consensus 5'-GGA(A/T)-3' known to be recognized by the ETS family of transcription factors through their highly conserved winged helix-turn-helix DNA-binding domain (ETS-domain) (13,14). The ETS family is divided in 12 subgroups based on structural homologies, among which ERG (ETS-related gene) is of particular interest for its oncogenic function. ERG, together with FLI1 and FEV, belongs to the ERG subgroup (15) on four recently defined subclasses based on their preferred ETS DNA-binding sequences (16). The ETS proteins have regulatory functions in embryonic development and physiological processes including proliferation, apoptosis, vasculogenesis, differentiation and haematopoiesis (17). However, aberrant expression could be associated with cancer diseases. In the case of ERG, fusion of the androgen-regulated gene TMPRSS2 to ERG sequences induces an over-expression of ERG associated with ~50% of prostate cancers with poor prognosis in >90% of TMPRSS2-ERG-positive prostate cancers (18,19). Other ETS fusion proteins (TMPRSS2-ETV1, TMPRSS2-ETV4, TMPRSS2-FLI1) are also detected in 5–10% of prostate cancers (18,20). Moreover, over-expression of ERG is observed in acute megakaryoblastic, myeloblastic and lymphoblastic leukaemia, associated

with poor prognosis and frequent relapses (21–23). Fusion proteins (FUS/TLS-ERG and ELF4-ERG) resulting from translocations have also been associated with those leukaemia, resulting in aberrant expression of ERG transcription factor (24,25). Furthermore, EWS-FLI1 and EWS-ERG fusion proteins are commonly observed in Ewing sarcoma (26,27). Despite their frequent implication in cancer disease, those ETS transcription factors are poorly studied in terms of inhibition and are currently not used in targeted therapy. A few studies investigated the inhibition of EWS-FLI1 in Ewing sarcoma by using, for instance, the mimetic peptide ESAP1 that directly binds EWS-FLI1 and alters its oncogenic function, or YK-4-279 that blocks EWS-FLI1 interaction with the RNA helicase A, leading to the decrease in EWS-FLI1 transcriptional activity. Moreover, YK-4-279 also targets ERG and ETV1 transcription activities reducing motility and invasion of cancer cells without targeting ETS/DNA binding (28–30). Alternatively, the DNA minor groove alkylating agent trabectedin (ET743) inhibits EWS-FLI1/DNA binding and modulates a known EWS-FLI1 downstream target in Ewing sarcoma cell lines (31). Using a high-throughput screening strategy, the same group evidenced similar *in vitro* and *in vivo* results with mithramycin (32). Alternatively, another targeting approach was based on the epigenetic modulation of ERG-positive tumors by histone deacetylase inhibitors (33,34).

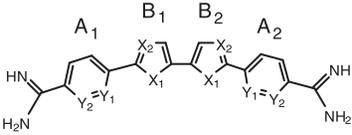
Our approach was to investigate ERG/DNA binding inhibition using heterocyclic diamidine compounds that target the minor groove of DNA. After selecting DB1255 as the most active compound, its DNA binding selectivity and that of ERG were studied to highlight their respective optimal binding site. The efficiency of DB1255 in ERG/DNA binding inhibition was identified at the molecular and cellular levels. The results of this research provide new directions in anticancer drug design.

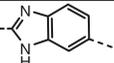
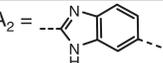
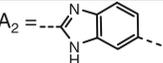
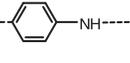
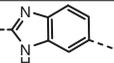
## MATERIAL AND METHODS

### Chemicals and plasmids

All DB compounds (Table 1) were prepared using methodologies previously reported (35,36) and prepared as 5 or 10 mM solutions in DMSO. CT-DNA was purchased from Sigma Aldrich (France). Oligonucleotides EBS-WT (5'-GATCTCGAGCAGGAAGTTCGA), EBS-Em (5'-GATCTCGAGCACCAAGTTCGA), EBS-Dm (5'-GATCTCGAGCAGGAAGTGGGA), EBS-m (5'-GATCTCGAGCAAGAAGTTCGA) and all EBS-mutated oligonucleotides were purchased from Eurogentec (Belgium). The ERG-sensitive reporter plasmid Pye-LUC, containing four copies of the artificial Pye-enhancer-element (Pye), has been described in Carrere *et al.* (37). pSG5, pSG5-ERG, pSG5-FLI1, Pye-Luc, murine Osteopontin (OPN) promoter reporter vectors OPN<sub>-136/+77</sub>-Luc and OPN<sub>-136/+77</sub>mutant1-Luc were obtained as previously described (38,39). pSG5-ETS-1 and -ETS-2 were generous gifts from Dr Marc Aumercier (40), whereas pRSV-ERM, PEA-3 and ETV1 were kind gifts from Dr Jean-Luc Baert (41). pGL3-EBS-WT-1R, pGL3-EBS-WT-6R,

**Table 1.** Structure of the chemical compounds. The general scheme of the compounds is presented on the top of the table. The external A1 and A2 and the central B1 and B2 cycles are as described in the following table



	Code	Modifications <sup>a</sup>
Monofurans	DB75	no B <sub>2</sub> ; X <sub>1</sub> = O; Y <sub>1</sub> = Y <sub>2</sub> = X <sub>2</sub> = CH
	DB293	no B <sub>2</sub> ; X <sub>1</sub> = O; A <sub>2</sub> =  ; Y <sub>1</sub> = Y <sub>2</sub> = X <sub>2</sub> = CH
Difurans	DB914	X <sub>1</sub> = O; Y <sub>1</sub> = Y <sub>2</sub> = X <sub>2</sub> = CH
	DB1246	X <sub>1</sub> = O; Y <sub>1</sub> = N; Y <sub>2</sub> = X <sub>2</sub> = CH
	DB1579	X <sub>1</sub> = O; Y <sub>2</sub> = N; Y <sub>1</sub> = X <sub>2</sub> = CH
	DB1256	X <sub>1</sub> = O; X <sub>2</sub> = Y <sub>1</sub> = Y <sub>2</sub> = CH; A <sub>2</sub> = 
	DB1975	X <sub>1</sub> = O; X <sub>2</sub> = CH; A <sub>1</sub> = A <sub>2</sub> = 
	DB1324	X <sub>1</sub> = O; X <sub>2</sub> = CH; A <sub>1</sub> = A <sub>2</sub> = 
	DB832	X <sub>1</sub> = O; X <sub>2</sub> = Y <sub>1</sub> = Y <sub>2</sub> = CH; no A <sub>2</sub>
	DB934	X <sub>1</sub> = O; Y <sub>1</sub> = N; X <sub>2</sub> = Y <sub>2</sub> = CH; no A <sub>2</sub>
	DB1999	X <sub>1</sub> = O; no A <sub>2</sub> ; B <sub>1</sub> =  ; X <sub>2</sub> = Y <sub>1</sub> = Y <sub>2</sub> = CH
Dithiophenes	DB1255	X <sub>1</sub> = S; Y <sub>1</sub> = Y <sub>2</sub> = X <sub>2</sub> = CH
	DB1247	X <sub>1</sub> = S; Y <sub>1</sub> = N; Y <sub>2</sub> = X <sub>2</sub> = CH
	DB1357	X <sub>1</sub> = S; Y <sub>1</sub> = C-CH <sub>3</sub> ; Y <sub>2</sub> = X <sub>2</sub> = CH
	DB1578	X <sub>1</sub> = S; Y <sub>2</sub> = C-F; Y <sub>1</sub> = X <sub>2</sub> = CH
	DB1514	X <sub>1</sub> = S; X <sub>2</sub> = C-CH <sub>3</sub> ; Y <sub>1</sub> = Y <sub>2</sub> = CH
	DB1998	X <sub>1</sub> = S; X <sub>2</sub> = N; Y <sub>1</sub> = Y <sub>2</sub> = CH
	DB1974	X <sub>1</sub> = S; X <sub>2</sub> = CH; A <sub>1</sub> = A <sub>2</sub> = 
	DB1450	X <sub>1</sub> = S; X <sub>2</sub> = Y <sub>1</sub> = Y <sub>2</sub> = CH; no A <sub>2</sub>
Diselenophenes	DB1282	X <sub>1</sub> = Se; Y <sub>1</sub> = Y <sub>2</sub> = X <sub>2</sub> = CH
	DB1273	X <sub>1</sub> = Se; Y <sub>1</sub> = N; Y <sub>2</sub> = X <sub>2</sub> = CH

<sup>a</sup>All compounds are hydrochloride salts except DB1256, DB832, and DB934, which are acetates.

pGL3-EBS-Em-6R and pGL3-EBS-Dm-6R reporter plasmids were obtained on ligation between *Nhe*I and *Xho*I restriction sites of an oligonucleotide with single EBS-WT or 6-fold tandem repeat of EBS-WT, EBS-Em or EBS-Dm elements, respectively, positioned 5' upstream to the minimal SV40 promoter.

#### Electrophoretic mobility shift assay

All double-strand oligonucleotides were 5'-radiolabelled as described (12). The binding experiments were

performed as previously detailed (12), with some modifications. Briefly, 5 µg of HT29 nuclear extracts (Nuclear extraction kit, Panomics, Fremont, CA, USA) or 2 µl of ERG or other ETS proteins expressing reticulocyte lysates (TNT<sup>®</sup>-coupled reticulocyte lysate system, Promega, Madison, WI, USA) were incubated for 20 min in binding buffer at 4°C with the radiolabelled DNAs and DB compounds at specified concentrations or 50-fold of the same EBS oligonucleotide (specific) or EBS-m (non-specific) as unlabelled competitors. Free DNA was separated from protein-DNA complexes on a 6% non-denaturing polyacrylamide gel under electrophoresis for 3 h 30 min at 300 V in 0.5× TBE buffer. Gel revelations were performed using a Molecular Dynamics STORM 860 and analysed with the ImageQuant-3.3 software.

#### DNase I footprint assay

pGL3-EBS-WT-1R, pGL3-EBS-WT-6R, pGL3-EBS-Em-6R and pGL3-EBS-Dm-6R plasmids were digested by *Cla*I and *Bgl*II restriction enzymes and radiolabelled with [ $\alpha$ -<sup>32</sup>P]dGTP at 3'-termini as reported (42). The OPN promoter DNA sequence was obtained by *Hind*III and *Mlu*I restriction enzymes digestion of the OPN<sub>-136/+77</sub>-Luc vector. The generated DNA fragment was 3'-end radiolabelled using [ $\alpha$ -<sup>32</sup>P]dCTP as specified previously. DNA sequences were deduced from guanine positions cleaved on DMS and piperidine treatment (G-track).

#### ELISA-derived protein/DNA binding inhibition assay

Streptavidin-coated 96-well plates (ThermoFischer) were blocked with 200 µl of blocking solution Tris Buffer Saline (TBS: 10 mM Tris-HCl, pH 8, 150 mM NaCl, 0.5% Tween and 5% BSA) for 1 h at room temperature and 5.7 nM of biotinylated EBS-containing oligonucleotide was immobilized per well for 1 h at room temperature. After two washes for 5 min with 200 µl of washing solution (TBS-Tween 0.5%) at room temperature, ERG proteins expressed from reticulocyte lysate system (Promega, Madison, WI, USA) (2 µl of lysate previously diluted to one-fifth in sterile water) were incubated for 30 min at 4°C with/without 5 µM of the indicated DB compounds in binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, 75 mM NaCl, 6% glycerol, 10 µg BSA). After three washes, the ERG/DNA complex was incubated with ERG antibody (sc-354, Santa Cruz Biotechnology) in TBS-Tween 0.5%–0.5% BSA solution 1 h at room temperature. After three washes, ERG/DNA complexes were detected by the incubation of the rabbit horseradish peroxidase-conjugated secondary antibody (Santa Cruz Biotechnology) in TBS-Tween 0.5%–BSA 0.5% solution for 1 h at room temperature. After two washes of 5 min with 200 µl of washing solution and one with TBS alone, 100 µl of TMB (Promega, Madison, WI, USA) was added and incubated for 15 min at room temperature. The reaction was stopped with 100 µl of 0.5 M H<sub>2</sub>SO<sub>4</sub>, and the absorbance was measured at 450 nm on the microplate spectrometer (Molecular Devices Versamax). All collected data was analysed using SoftMax Pro software.

### Spectrometric analysis of DNA binding

DNA melting temperature, UV/Visible spectroscopy and circular dichroism experiments were performed as previously reported (43), with the following modifications. For melting temperature studies ( $\Delta T_m$ ), the indicated ratios are expressed as drug/base pair. For UV spectroscopy, DB1255 (20  $\mu$ M) was incubated with increasing concentrations of CT-DNA (0.1–200  $\mu$ M) or EBS-WT oligonucleotide (0.1–50  $\mu$ M). For circular dichroism spectrometry, 200  $\mu$ M (bp) of CT-DNA or EBS-WT oligonucleotide was incubated with increasing concentrations of DB1255 from 1 to 50  $\mu$ M and 0.01 to 10  $\mu$ M, respectively. All three experiments were performed in BPE buffer (6 mM  $\text{Na}_2\text{HPO}_4$ , 2 mM  $\text{Na}_2\text{H}_2\text{PO}_4$ , 1 mM EDTA, pH 7).

### Surface plasmon resonance analysis of the DNA binding

Surface plasmon resonance (SPR) measurements were performed with a four-channel Biacore T200 optical biosensor system (Biacore, GE Healthcare, Inc.). 5'-biotin labelled hairpin oligonucleotides EBS-WT 5'-G CGTCGAACTTCCTGCTtctcAGCAGGAAGTTCGAC GC-3' (the EBS-minimal sequence is underlined, loop in lower case and stem in caps letters) or EBS-Dm 5'-GCGT CCCACTTCCTGCTtctcAGCAGGAAGTGGGACGC -3' (mutated bases are in bold and italic letters) were immobilized onto streptavidin-coated sensor chips (Biacore SA) as previously described (44,45), with the difference that the compounds were injected in a solution containing 10% DMSO to avoid drug precipitation in the flow cells at the highest drug concentrations.

### Cell culture

Human cervical HeLa, colon HT29 and prostatic VCaP carcinoma cells were grown at 37°C under 5%  $\text{CO}_2$  in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% foetal bovine serum, penicillin (1000 U/ml) and streptomycin (10  $\mu$ g/ml).

### Transient transfections

Using HT29 cell line,  $15 \times 10^4$  cells/well were cultured overnight in 1 ml of supplemented medium in 12-well plate. Transfection was performed by adding a mixture of 3  $\mu$ l of lipofectamine 2000 (Invitrogen), 1.5  $\mu$ g of reporter plasmid (pGL3-EBS-WT-6R, pGL3-EBS-Em-6R, pGL3-EBS-Dm-6R or pGL3), 0.5  $\mu$ g of pmaxGFP as a normalising plasmid qsp 100  $\mu$ l of OptiMEM medium (Gibco, Invitrogen) previously incubated for 20 min at room temperature. After 6 h, the medium was changed by complete medium supplemented, or not, with DB1255 at the indicated concentrations, and cells were cultured for 24 h at 37°C. The transfected cells were collected, rinsed with PBS before lysis for 20 min at 4°C by lysis passive buffer (Promega, Madison, Wisconsin, USA). After a 15-min centrifugation at 13 000g, supernatants were collected. Luciferase activity was quantified using Luciferase assay system (Promega, Madison, Wisconsin, USA) according to manual recommendations, whereas the fluorescence of GFP protein was detected using a 96-well fluorimeter (Mithras, Berthold) at  $\lambda_{\text{excitation}} = 485$  nm and

$\lambda_{\text{emission}} = 535$  nm. For exogenous ERG expression in HeLa cells,  $2 \times 10^5$  cells/well were plated in 12-well plates 5 h before transfection using polyethylenimine (Eurogentec) according to the manufacturer's instructions. DNA mixture included 500 ng of the indicated firefly luciferase reporter gene with or without 200 ng of the expression vectors pSG5-ERG and 5 ng of a control plasmid tk-luciferase (tk-*Renilla*, Promega). After 6 h, cells were challenged with 2.5  $\mu$ M of DB1255 for 20 h. Cells were harvested and luciferase activities were detected using the Dual-Glo Luciferase Assay System (Promega). Luciferase activity was measured on a luminometer (Berthold Biolumat centro LB960). Firefly luciferase values were normalized to those of the control renilla luciferase.

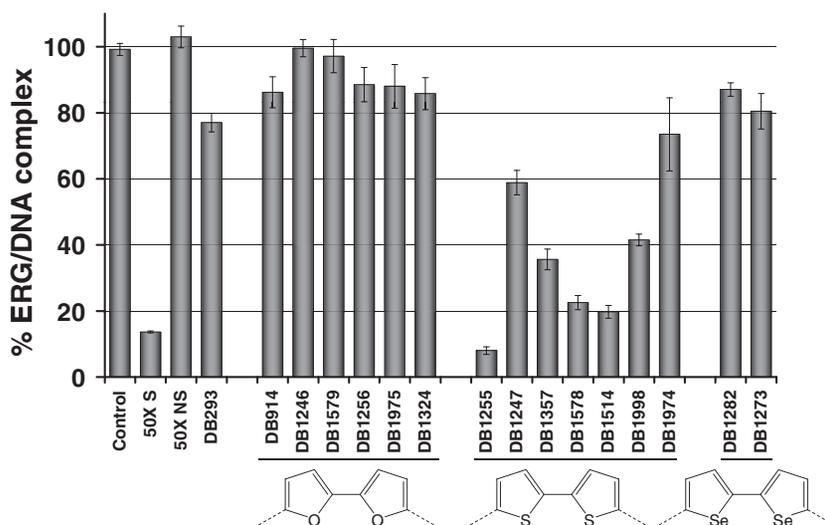
### Western blotting

Transfected cells were washed once with PBS and incubated in ice-cold lysis buffer (50 mM Tris-HCl, 10 mM EDTA, 1% SDS, pH 8) for 2 h at 4°C. After a 20-min centrifugation at 4°C and 13 000g, the protein extracts were collected, quantified (Bradford) and equal amounts of proteins was denatured in loading buffer (100 mM Tris-HCl, pH 6.8, 200 mM DTT, 4% SDS, 0.2% bromophenol blue and 20% glycerol) 3 min at 90°C and loaded on SDS-PAGE. After gel electrophoresis, the proteins were transferred onto a nitrocellulose membrane (GE Healthcare, France), which was blocked for 1 h at room temperature with TBS blocking solution (Tris-HCl 10 mM pH 8, NaCl 150 mM, Tween 0.5%) containing 5% of BSA before immunodetection using polyclonal anti-ERG antibody (sc-353, Santa Cruz, Biotechnology) and revelation by ECL detection system (GE Healthcare) using an HRP-conjugated secondary antibody (Santa Cruz, Biotechnology). Anti- $\beta$ -actin monoclonal antibody (Sigma Aldrich) was used as an internal control.

## RESULTS

### Screening heterocyclic diamidines for inhibition of ERG/DNA binding

Because previous researches highlighted some inhibitory effects of DB293 on the ability of proteins to complex with EBS sequence present on TransSignal protein/DNA array, we first investigated a series of transcription factors belonging to the ETS family (ETS-1, ETS-2, ERG, FLI1, ERM, PEA-3 and ETV1) for binding to this EBS-containing sequence. Only ERG, and to a much lesser extent FLI1, shows impressive band shift (data not shown) and was therefore selected for inhibition with the lead compound, DB293. The band shift was inhibited in a similar dose-response manner using both ERG-expressing reticulocyte lysate and HT29 nuclear extracts with full inhibition using 25  $\mu$ M of DB293 (Supplementary Figure S1A). In a DNaseI footprinting assay (Supplementary Figure S1B), binding of DB293 at a site (grey box) located on the 3'-side of the minimal EBS (underlined in black) was observed, as exemplified by a strong negative differential cleavage from 1.25  $\mu$ M. The minimal EBS site was not sufficient for ERG/DNA binding, which is known to require specific 5' and 3' adjacent bases and the



**Figure 1.** EPDBi for screening the ERG/EBS-WT binding inhibitors. EPDBi was performed on streptavidin-coated plates in which biotinylated EBS-WT oligonucleotides were fixed. ERG proteins were added together with the indicated DB compounds at 5  $\mu$ M. Control corresponds to 100% of ERG/DNA complex. EBS-WT DNA-binding specificity was validated by addition of 50-fold excess of EBS-WT (50 $\times$  S) or non-specific oligonucleotide (50 $\times$  NS) in the binding buffer. Results are means  $\pm$  s.e.m. from two experiments both performed in triplicate.

3' binding site of DB293 is in agreement with interference of ERG/DNA binding.

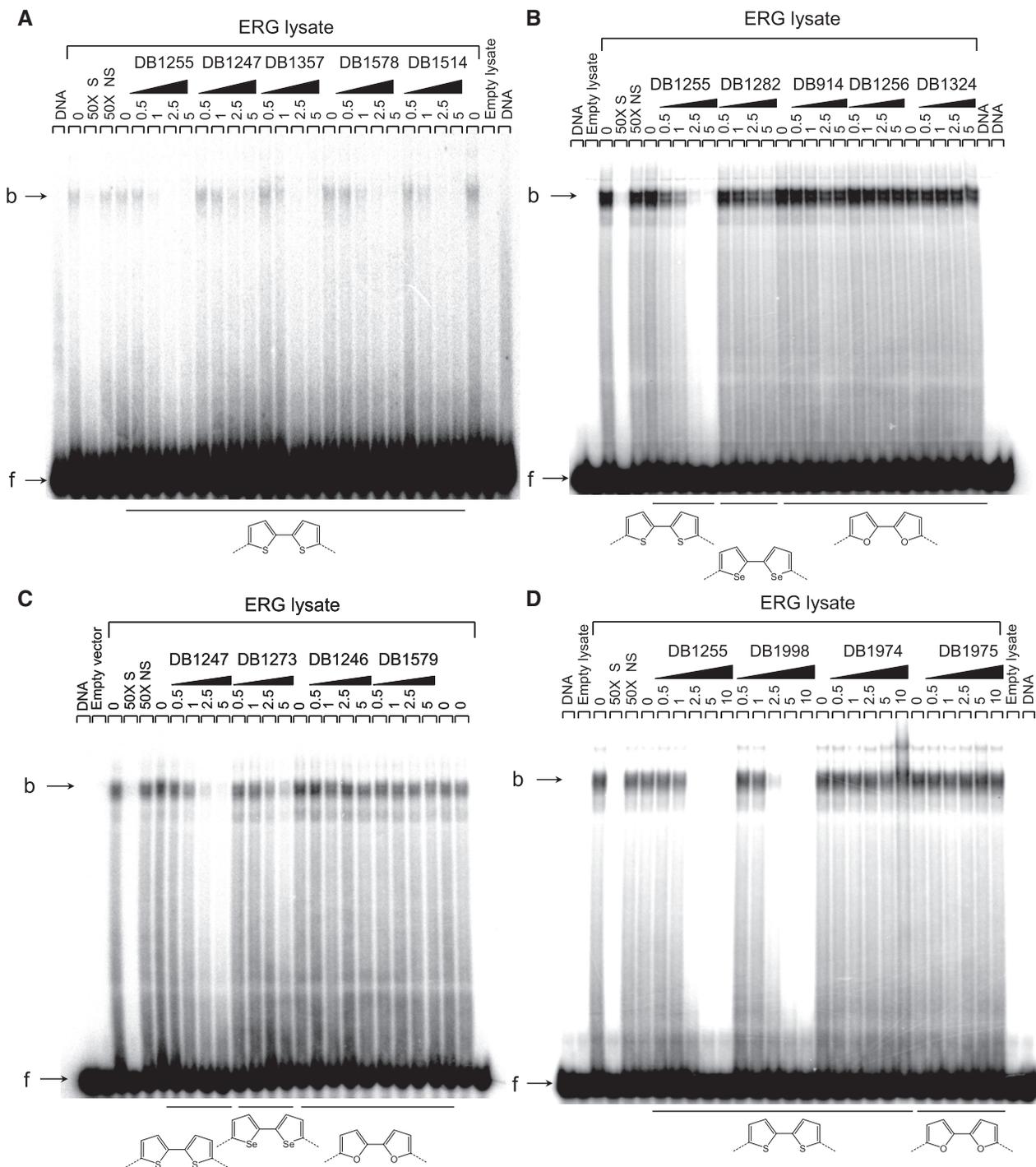
We then challenged a new series of heterocyclic diamidines of varied structure for ERG/DNA binding inhibition using an ELISA-derived Protein/DNA binding inhibition assay (which we call here EPDBi). This method can screen and quantify in one step using 96-well plates the inhibitory effect of a large series of compounds for protein/DNA complex formation. A new structural sub-type of diamidine compounds (Table 1) gave promising inhibitory results (Figure 1). Structure/activity analyses showed that di-thiophene derivatives were more active than their di-furan or di-selenophene counterparts. Among di-thiophenes, the di-thiophene-phenyl-amidine DB1255 was a particularly efficient inhibitor with more than 90% of complex inhibited using 5  $\mu$ M of DB1255. At the same concentration, DB293 only induced a 20% inhibition. Such relative inhibition efficiency was confirmed using EMSA (Figure 2). Results for ERG/DNA binding inhibition using EMSA were consistent with those resulting from EPDBi assay where the more efficient competitors present two thiophenes as central rings rather than furans or selenophenes. In the di-thiophene series, the substitution of the di-phenyl by di-pyridine (DB1247) or di-benzimidazole (DB1974) heterocycles decreases the inhibition efficiency by only 40–60%. In conclusion, the DB1255 is selected as the best inhibitor of the ERG/DNA binding. The protein/DNA complex obtained with HT-29 nuclear extracts was also inhibited on addition of increasing concentrations of DB1255 (Figure S2).

#### DNA-binding properties of DB1255 and derivatives

DB1255 was selected for DNA binding studies using UV-visible spectrometry, DNA melting temperature studies and circular dichroism analysis of the mode of binding. First, interaction of DB1255 with DNA was analysed

using spectroscopic analyses in the presence of EBS oligonucleotide (EBS-WT) or CT-DNA (Figure 3 and Supplementary Figure S3). The variation of the DNA melting temperature indicates strong binding of DB1255 to the EBS-containing oligonucleotide, even at low drug/DNA ratio (Figure 3A). UV/visible absorbance spectroscopy evidenced hypochromic effects at low concentration (0–4  $\mu$ M) of EBS-WT oligonucleotide (high drug/DNA ratios) that is probably due to stacking of the compound outside of the double helix. The spectra then evidenced bathochromic and hyperchromic effects from 4 to 50  $\mu$ M with an isosbestic point at 460 nm that argues for a stronger, single binding mode at the lower drug/DNA ratios (Figure 3B). Circular dichroism spectrometry was then performed to get an insight in the orientation of the compound relatively to the DNA helix (groove binding or intercalation) on addition of increasing concentrations of DB1255. In the presence of DB1255 (Figure 3C), the biphasic intrinsic CD signal is modified with the appearance of a strong bisignate CD around 420 nm that argues for a groove binding interaction of DB1255 and is consistent with dimer binding to the EBS-containing oligonucleotide (46). Similar conclusions could be obtained on binding to CT-DNA (Figure S3).

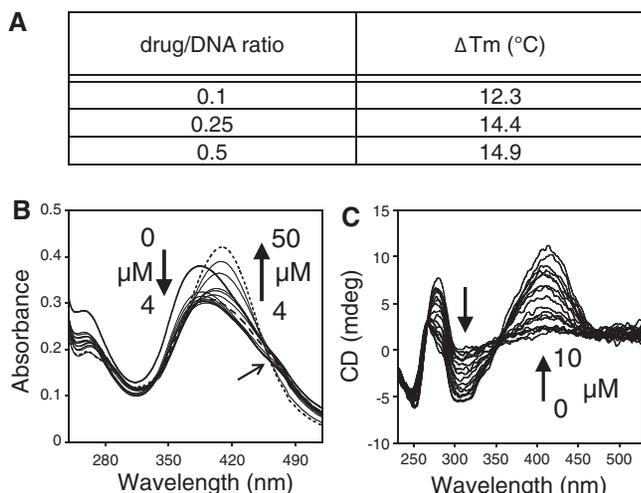
The precise binding localization of DB1255 or derivatives relative to the EBS site was determined using DNase I footprinting assays as described previously for the DB293 (Figure 4). For this purpose, we used the *Clal*-*Bgl*II DNA fragment containing the 6-fold repeated EBS site from pGL3-EBS-WT-6R reporter vector. The radiolabelled DNA was incubated with increasing concentrations of the various selected compounds before DNase I mild digestion. The densitometric analyses were performed for each gel. A representative gel is presented in Figure 4 using the active compounds DB1255 and DB1998 for comparison with the inactive DB1974 or DB1975. The portion of the gel that is covered by DB1255 and DB1998



**Figure 2.** Validation of the ERG/EBS complex inhibition by the evaluated diamidine compounds. EMSA was performed with the EBS-WT radiolabelled oligonucleotide incubated with ERG protein expressed in reticulocyte lysate systems or empty lysate made with the equivalent empty vector, in the presence of thiophene (A), selenophene or furan (B–D) derivatives. Bound ERG/EBS-WT complexes (*b*, indicated by a black arrow) were competed with increasing concentration of DB compounds ( $\mu\text{M}$ ) and separated to free EBS-WT oligonucleotide (*f*, indicated by black arrow) by electrophoresis migration. DNA binding specificity was determined by addition of 50-fold excess of the specific unlabelled EBS-WT ( $50\times$  S) or a non-specific unlabelled oligonucleotide EBSm ( $50\times$  NS).

encompasses the 5'-gAAGTTC site. The other compounds that appeared to be less active from EPDBi (Figure 1) were less selective or bound with lower affinity to the same portion of the DNA (Supplementary Figure S4). Such binding only partially covers (underlined) the EBS

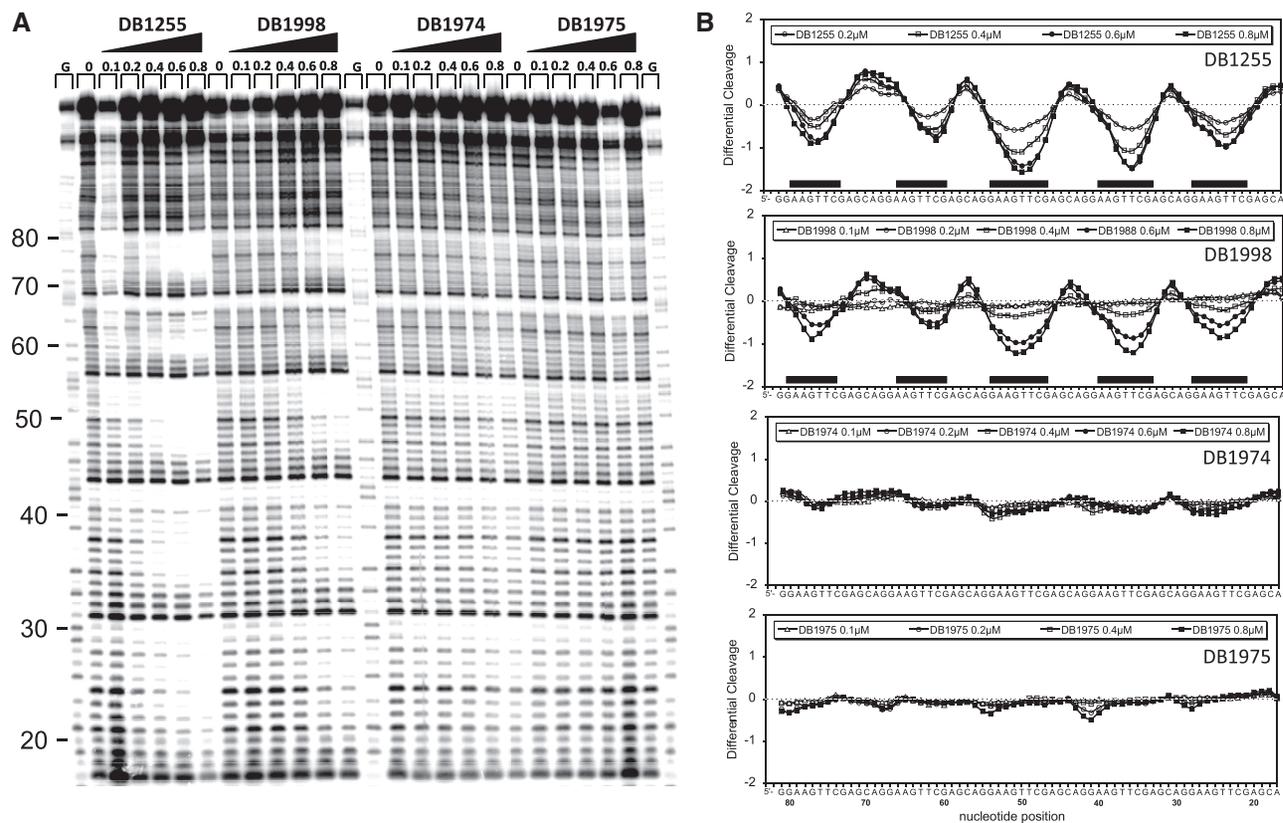
minimal binding site 5'-GGAA, but also spans the important EBS binding sequence, 5'-GTT, located 3' after it. In conclusion, active compounds determined by EPDBi showed an excellent DNA-binding recognition and affinity for the EBS binding site.



**Figure 3.** DNA-binding characteristics of the selected compound DB1255. (A) Temperature melting assay of the EBS-WT oligonucleotide in presence of the DB1255.  $\Delta T_m$  values were obtained from the equation:  $\Delta T_m = (T_m(\text{drug/DNA}) - T_m(\text{DNA}))$ . (B) UV-spectroscopy properties of the complex DB1255/DNA was observed using a fixed concentration of DB1255 and increasing concentrations of EBS-WT oligonucleotide. (C) Circular dichroism properties of EBS-WT oligonucleotide in presence of increasing concentrations of DB1255. Major changes in spectra (B and C) were indicated by black arrows.

### Looking for DB1255 and ERG optimal binding sites

To get an insight in the precise bases that are crucial for the DNA interaction, DNA binding capacity of the selected compound DB1255 was evaluated in terms of sequence selectivity using a DNase I footprint assay on DNA fragments containing the wild-type EBS sequence or EBS mutated at one position per experiment to get information with all four nucleotides options at each position. All nucleotides from positions  $-1$  to  $+7$  relative to the minimal EBS site  $\text{G}^{-1}\text{G}^0\text{A}^{+1}\text{A}^{+2}$  were mutated and evaluated for DB1255 binding selectivity (Figure 5A and B). Radiolabelled DNA contained the three mutated EBS surrounded with two wild-type EBS. At the external positions  $-1$ ,  $0$  and  $+7$ , no major changes in the differential cleavages from densitometric analyses were observed whatever the nature of the nucleotide at those positions suggesting that DB1255 has no selective binding to those bases (Figure S5). By contrast, mutations in central positions from  $+1$  to  $+5$  affected (decreased or abolished) DB1255 binding. Globally, DNase I footprinting data demonstrate the importance of the  $\text{A}^{+1}$ ,  $\text{A}^{+2}$ ,  $\text{T}^{+4}$  and  $\text{T}^{+5}$  bases for DB1255 selective binding to DNA. Mutations at position  $+3$  and  $+6$ , however, affected the compound binding to a lesser extent (Figure S5). Values from the densitometric analyses were used to deduce the



**Figure 4.** Sequence selectivity of diphenyl di-thiophene diamidine compounds. (A) DNase I footprint assays showed the interaction sequences of the DB compounds. The radiolabelled DNA fragment containing the six repeats of the EBS-WT sequence was incubated with the indicated concentrations of compounds ( $\mu\text{M}$ ) and submitting to electrophoresis migration. The track labelled “G” is as in Supplementary Figure S1. (B) Densitometric analysis derived from the gel and protected sites observed are indicated by black boxes.

predicted logo of the frequencies of DNA recognition with 0.8  $\mu$ M of DB1255 using each single base pair mutation (Figure 5C).

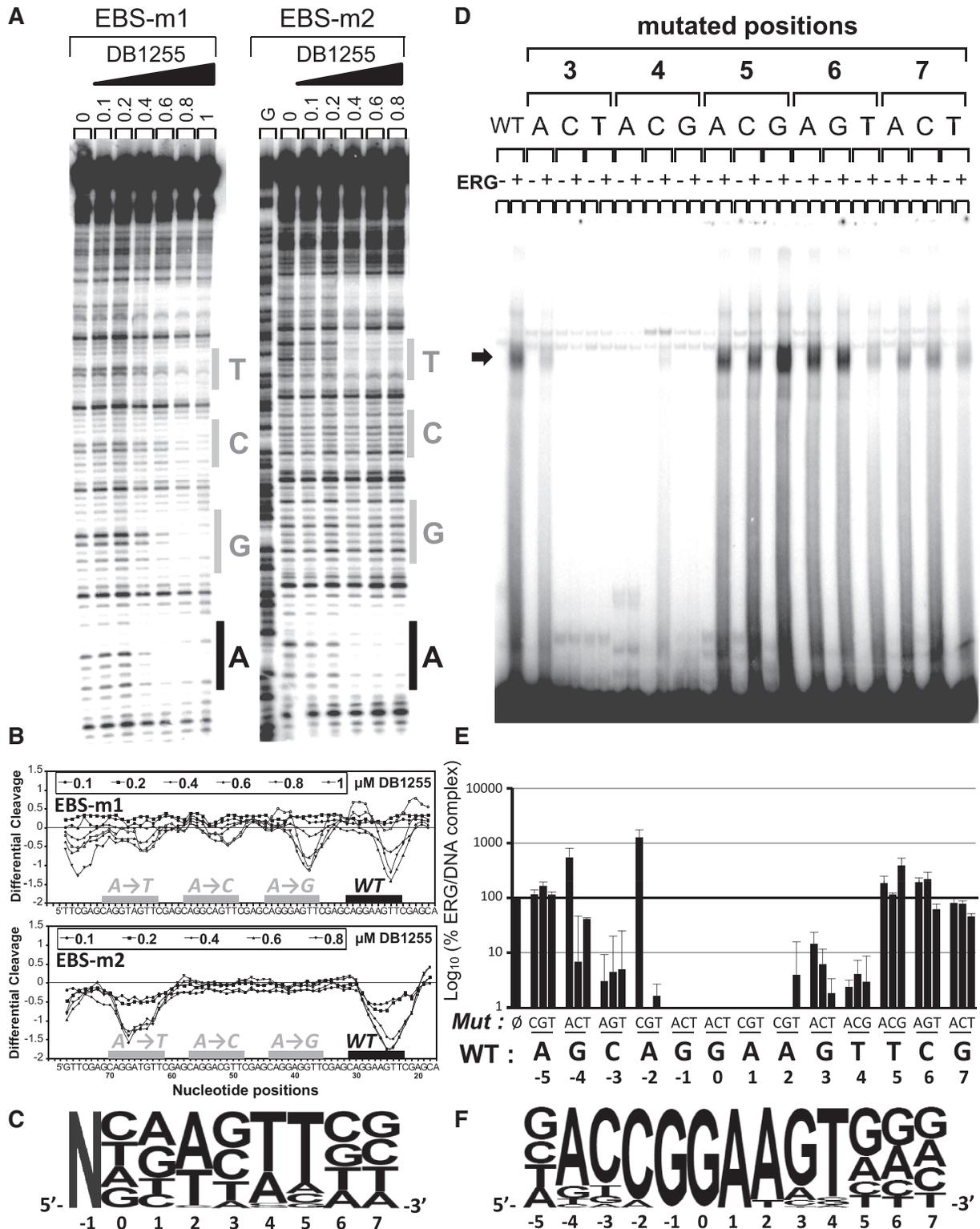
In parallel, the DNA binding selectivity of the ERG transcription factor was evaluated in EMSA using 5'-end labelled oligonucleotides containing the EBS-WT sequence or mutants at each single position from -5 to +7 as indicated in Figure 5D and Supplementary Figure S6. Globally, the mutations affecting the EBS minimal site abolish the ERG binding as expected. In addition, the same mutations at positions -2 (A→C) or -4 (G→A) and to a lesser extent +5 (T→G) increases ERG/DNA binding intensity. For the surrounding positions -3, +3 and +4 the relative percentage of ERG/DNA binding decreased relatively to EBS-WT suggesting their relevance for the ERG/DNA binding. By contrast, changing bases at the most external positions -5, +6 and +7 did not significantly modify ERG/DNA binding propensity. From those data, the predicted logo for ERG/DNA binding selectivity is presented in Figure 5F. For each mutated sequence that could form ERG/DNA complexes (mutations 3A, all mutated points at positions +5 to +7), we evaluated the propensity of DB1255 at a fixed concentration (1  $\mu$ M) to inhibit ERG/DNA binding. A strong decrease of the inhibition potency is observed as expected for all three mutations (A, C and G) at +5 position (Figure 6A). This result is in agreement with the consensus DNA binding site for DB1255 showing that the presence of a T<sup>+5</sup> is essential for the binding (Figure 5) as well as for the ERG/DNA complex inhibition (Figure 6A). In conclusion, comparison of DB1255 and ERG consensus binding sites exemplifies the necessity of the presence of a common 5'-AA(G/N)T-3' site that encompasses the minimal EBS for optimal DB1255 interaction and for best ERG/DNA complex inhibition.

To evaluate the inhibition of ERG/DNA binding in cells using luciferase assays, we selected from previous analyses the ERG/DNA binding mutated sequences (EBS-Em) and the DB1255/DNA binding mutated sites (EBS-Dm) as being 5'-GGAA → 5'CCAA and 5'-GGAA GTTCGA → 5'-GGAAGTGGA, respectively (mutated bases are underlined). The efficiency for binding of DB1255 (Figure 6B) and other derivatives (data not shown) to EBS-Em but not EBS-Dm sites was confirmed from comparison with the EBS-WT sites using DNaseI footprinting assays on 6-fold repeated sequences generated from the pGL3-EBS-6R, pGL3-EBS-Em-6R and pGL3-EBS-Dm-6R reporter vectors as described in the Materials and Methods section. As expected, inactive compounds such as DB1974 and DB1975 did not present any footprint on the three DNA fragments (data not shown), whereas the active compounds DB1255 and DB1998, and to a lesser extent DB1357 and DB1578 (Figure 6B and data not shown), displayed huge footprints characterized by negative differential cleavages for EBS-WT and EBS-Em but not for EBS-Dm sequences. To get more quantitative analysis of the sequence binding and selectivity of DB1255 on EBS-WT, SPR experiments were conducted with DNA hairpin duplexes containing EBS-WT or EBS-Dm sequences (Figure 6C). Because of the poor solubility of DB1255, the experiments

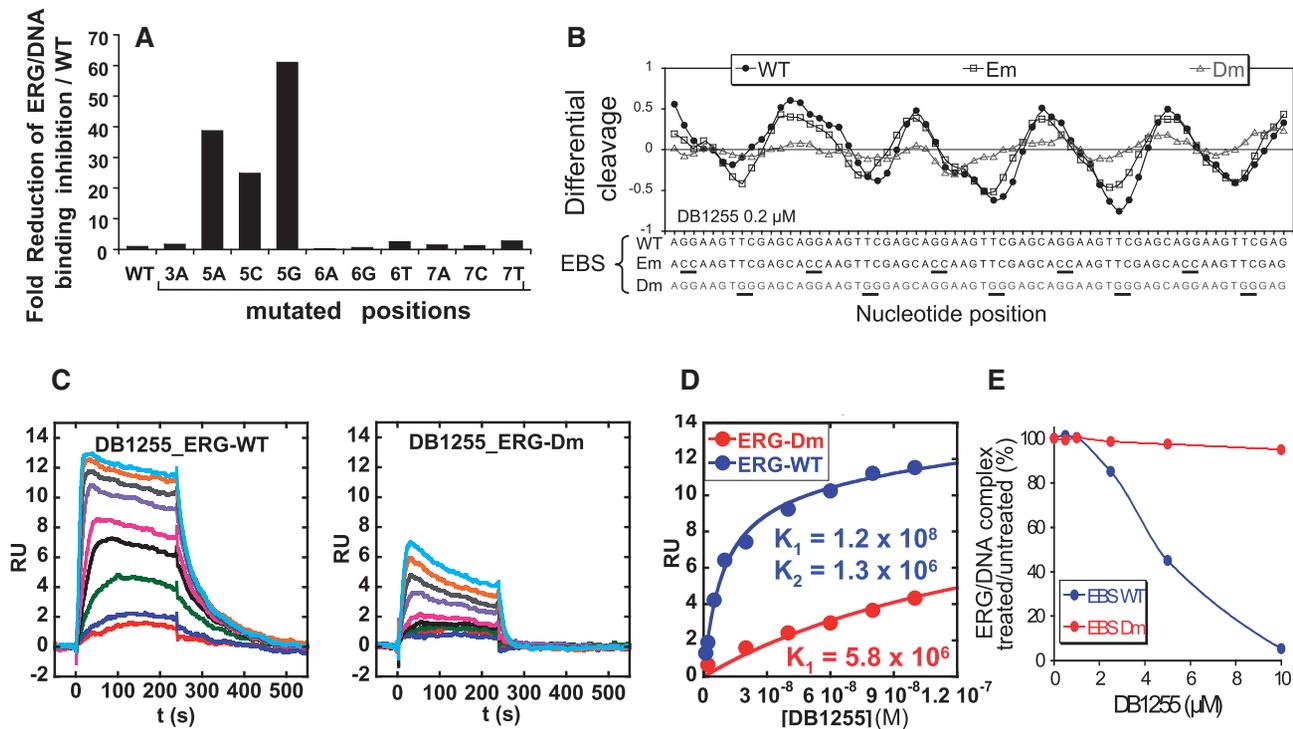
were performed in a buffer containing 10% DMSO during the whole experimental process. The impact of DMSO in this context was evaluated using the DB1998 derivative, which is more water soluble. The equilibrium and kinetic constants obtained in the presence of 10% DMSO (Figure S7) and in low concentrations of DMSO (data not shown) were roughly the same and any correction is less than experimental error. The sensorgrams for DB1998 (Figure S7A) showed fast kinetics of association and dissociation, whereas that for DB1255 presented slower association/dissociation rates, suggesting better binding kinetics for DB1255 over DB1998. The steady state RU values were plotted against *C<sub>f</sub>* (free compound concentration, in M) and fitted to a single or two sites model (Figures 6D and S7B) and showed that both DB1255 and DB1998 present similar equilibrium constants for EBS-WT but different ones for EBS-Dm sequence. Both compounds have a single strong binding site for the WT sequence (*K*<sub>1</sub>), with a second site that is 100–200 times weaker (*K*<sub>2</sub>). The binding of DB1255 to the EBS-Dm sequence is much weaker than that to EBS-WT, suggesting that such mutations strongly affect the DNA binding propensity of DB1255 as also observed earlier using DNase I footprinting assays (Figure 6B). By contrast, DB1998 binds strongly to EBS-Dm sequence, and *K*<sub>1</sub> is only reduced by a factor of 5 compared with EBS-WT, suggesting lower binding selectivity (Figure S7B). Therefore DB1255 appears to be the most interesting compounds in this series and was therefore retained for further cellular experiments. As part of our sequence validation, we finally showed that the ERG protein binds to EBS-WT and EBS-Dm, but not EBS-Em oligonucleotide (data not shown), which confirms the above data showing that the two guanines at positions -1 and 0 are essential for the ERG binding but not the bases at positions +5 and +6. In competition studies, ERG binding to DNA is abolished by increasing concentrations of DB1255 for the EBS-WT sequence but not for the EBS-Dm sequences (Figure 6E). Thus, those three sequences could be used for luciferase assays.

#### Modulation of the ERG/EBS complex inhibition *in cellulo*

This experiment was performed in HT29 cell line, which expresses the ERG transcription factor as visualized by a Western blot (Supplementary Figure S8A). The cell uptake and nuclear localization of DB1255 was evidenced using fluorescence microscopy based on DB1255 intrinsic fluorescence properties in comparison with the cytoplasmic localization of MitoFluor Red (a mitochondrial tracer) (Supplementary Figure S8B). Such nuclear localization is in agreement with genomic DNA as a target for DB1255. Therefore, the luciferase assays were performed in cells transfected with pGL3-EBS-WT-6R, pGL3-EBS-Em-6R and pGL3-EBS-Dm-6R for 6 h before the addition of DB1255 from 1 to 5  $\mu$ M during 24 h, conditions for which DB1255 is not cytotoxic (Figure S8C). After normalization relatively to GFP expression and empty vectors, we evidenced relative ERG-dependent transactivation using EBS-WT and EBS-Dm but not EBS-Em element, highlighting the



**Figure 5.** Sequence selectivity of ERG and DB1255 DNA binding. (A) Sequence selectivity of the DB1255/DNA interaction. DNase I footprint assays were performed with a DNA fragment containing three EBS sequences mutated at the indicated positions (positions -1 to +7, DNase I footprint gels and densitometric analyses for the position -1 and 0 and position +3 to +7 are presented in Figure S5A). For each mutation incubation with DB1255 at the indicated concentrations (μM) revealed protected sites of DB1255 interaction. The track labelled “G” is as in Figure S1. (B) Densitometric analyses derived from the gels and EBS-WT (WT) sequence or EBS-mutated sequence are indicated by black boxes or a grey boxes, respectively. (C) The logos were drawn using enoLOGOS (47), where the size of the letter is directly proportional to the binding affinity at 0.8 μM of the compound resulting for densitometric analysis (data shown in Figure S5B). (D) Sequence selectivity of the ERG binding to the EBS-WT sequence. ERG binding sequence selectivity are observed by EMSA using radiolabelled EBS oligonucleotide WT (EBS WT) or mutated (EBS mXN) at the indicated position (position X of base N) incubated with ERG protein expressed in reticulocyte lysate system or empty lysate. EBS sequences with mutated positions from -5 to +7 were evaluated but only the EMSA with mutated positions +3 to +7 are shown here (EMSA of mutated position -5 to +2 are shown in Figure S6A). (E) Representation of the percentage of ERG/DNA complex derived from the gel quantification as means ± s.e.m. relatively to EBS-WT. (F) The logos were drawn as above (full data presented in Figure S6B).



**Figure 6.** Specificity of ERG/DNA and DB/DNA binding on mutated on EBS sequences or on DB binding sequences. **(A)** ERG/DNA binding inhibition depending on mutated EBS sequences. The capacity of a fixed concentration of DB1255 (1 μM) to inhibit ERG/DNA binding was evaluated by EMSAs on mutated EBS sequences when a binding was possible. Binding complexes were quantified and expressed as following. ERG/DNA binding inhibition by DB1255 with mutated EBS were measured as a percentage of the inhibition obtained with EBS-WT. ERG/DNA binding inhibition by DB1255 obtained with EBS-WT sequence was scaled up to 1 and the percentage of inhibition with mutated EBS relative to EBS-WT were plotted as fold-reduction of 1255 inhibition capacity. **(B)** Densitometric analysis of DB/DNA interaction derived from a DNase I footprint (gels not shown). DNase I footprint assays were performed with DNA fragments containing six repeats of EBS-WT, EBS-Em (mutated on the EBS element) or EBS-Dm (mutated on the compound site interaction) and compounds at the indicated concentrations. **(C)** SPR Sensorgrams for DB1255 on the EBS-WT (left panel) or EBS-Dm (right panel) hairpin oligonucleotides. **(D)** Fitting analysis of SPR sensorgrams. RU values from the steady-state region of SPR sensorgrams are plotted against free DB1255 concentration (M). The corresponding  $K_1$  and  $K_2$  values are calculated from one or two site(s) fitting models for EBS-Dm (Red) and EBS-WT (Blue) sequences. **(E)** ERG/EBS-WT and ERG/EBS-Dm binding modulation by DB1255. Quantification of the percentage of ERG/DNA complex formation was performed from EMSA gels (not shown).

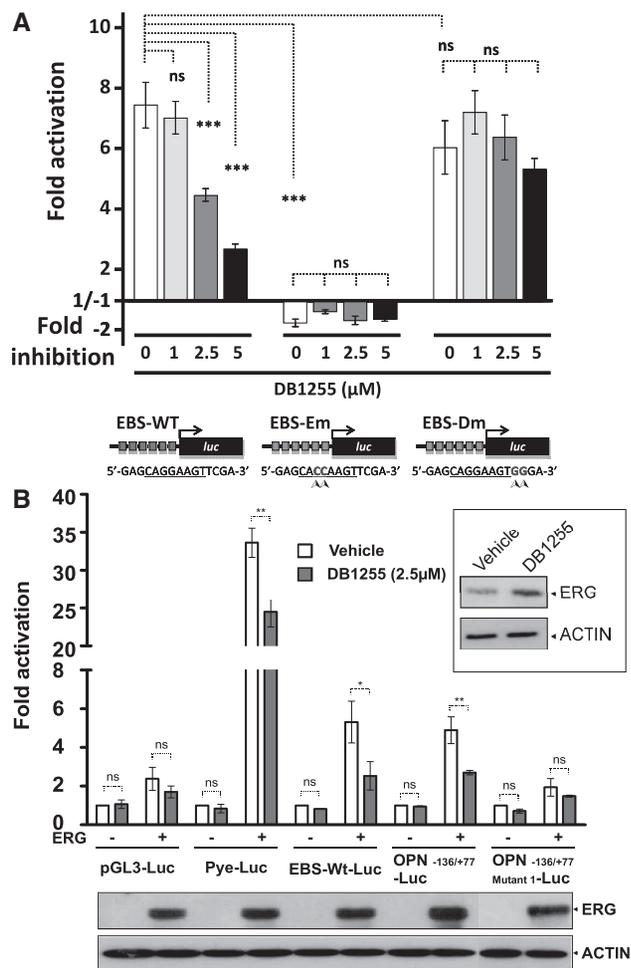
crucial role of the GG bases of the EBS site that are mutated in EBS-Em sequences (Figure 7A). In the presence of increasing concentrations of DB1255, the ERG-mediated transactivation at EBS-WT sites decreased in a dose-dependent manner, consistent with the inhibition of ERG/DNA binding observed *in vitro* in EMSA (Figure 6E). No inhibitory effect is obtained using EBS-Dm-containing reporter vector in agreement with previous EMSA analyses (Figure 6E).

To confirm this modulation we finally evaluated the competitive activity of DB1255 on promoters that have been described to be regulated by ERG protein such as the OPN promoter (39). The DNA recognition of DB1255 at the EBS site was addressed using DNase I footprinting assays (Figure S9). The luciferase assay was performed in HeLa cells as described previously (39). Because HeLa cells do not express ERG protein, those cells were co-transfected with the ERG-expressing vector or empty vector as control (Figure 7B, bottom panel). The pGL3-EBS-WT-6R (EBS-WT-Luc) previously evaluated in HT-29 cells and the reference polyomavirus enhancer (Pye) were first evaluated from comparison with the control to validate the model in HeLa cell line. An inhibition of the ERG-mediated transactivation by DB1255

was obtained for both EBS-WT and Pye models. Such inhibition could not be attributed to a decrease of the transcription of ERG-expressing vector itself as visualized in the embedded Western blot panel from comparison with actin expression. Moreover, DB1255 also inhibits ERG-mediated transactivation from the OPN promoter (OPN<sub>-136/+77</sub>-Luc) in a manner that is dependent on the presence of the ERG-binding site because no significant effect is obtained using the sequence mutated at the EBS at position -119 to -115 (5'-GGAGGAAG mutated to 5'-GGTAAAAG in OPN<sub>-136/+77</sub> Mutant1-Luc) (Figure 7B). Both luciferase assays bring out the inhibitory effect of DB1255 on ERG/DNA binding in cells.

## DISCUSSION

Transcriptional modulation of oncogenic transcription factors is a new and promising area to develop targeted therapies, particularly in cancer. Despite the difficulty to target this class of proteins, many groups have investigated their inhibition by using several approaches. One method under development consists of targeting the protein-DNA binding interaction using small DNA-binding compounds.



**Figure 7.** *In cellulo* inhibition of the ERG/DNA complex by DB1255. **(A)** Modulation of the ERG/DNA transcriptional activity on a 6-fold repeated EBS-containing reporter vector. HT29 cells transfected with luciferase vector with or without 6 repeat of the EBS-WT, EBS-Em or EBS-Dm elements (point mutations indicated with arrow heads) located upstream the cDNA of the luciferase (*luc*) gene, as indicated, were incubated with increasing concentrations of DB1255 (24h). The fold-activation values correspond to the normalized luciferase activity in comparison with cells transfected with empty vector. Results are means  $\pm$  s.e.m. from three independent experiments performed in triplicate. Statistical analyses were performed using analysis of variance test. **(B)** *In cellulo* inhibition of the transcriptional activity of ERG-target gene promoter by the DB1255 compound on the OPN promoter. Transfections were performed in HeLa cells using ERG-expressing vector or empty control vector. For transcriptional modulation, different reporter vectors were used: pGL3-Luc as empty reporter, Pye-Luc as the reference of polyomavirus enhancer, EBS-WT-Luc similar to pGL3-EBS-WT-6R, OPN<sub>-136/+77</sub>-Luc as the OPN promoter and OPN<sub>-136/+77</sub> Mutant1-Luc containing the OPN promoter mutated on EBS. After transfection, HeLa cells were treated or not with DB1255 (2.5  $\mu$ M, 24h). ERG and  $\beta$ -actin expressions were detected by Western blotting (bottom and embedded panels). Results are means  $\pm$  s.e.m. from at least two experiments. Statistical significance was determined using one-way analysis of variance, followed by Bonferroni's multiple comparison post test. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , ns, non-specific ( $P > 0.05$ ).

Such compounds specifically interact with the same DNA site and consequently decrease the transcription factor/DNA complexes leading to the diminution of gene transcription. By using this competition approach of

transcription factor inhibition, we first showed the possibility of this inhibition concept using the heterocyclic diamidine DB293, which targets Pit-1 and Brn-3 (12). In that project, we showed that if compound binding to a sequence (ATGA) common to that bound by a transcription factor was important, it does not mean that all transcription factors containing that sequence would be affected. Indeed, the IRF-1 transcription factor consensus interaction site also contains an ATGA sequence, but that complex was not affected by the presence of DB293. This result shows that the mode of binding of each transcription factor needs to be taken into account in inhibition design. Based on previous data using TranSignal Protein/DNA arrays (12), we focussed here on the ERG transcription factor and showed that the inhibition of ERG binding to the DNA is obtained through the recognition of a sequence that partially spans the minimal ERG-binding site. It is known that the binding of the ETS-domain to DNA requires direct recognition of conserved arginine residues with the core 5'-GGA(A/T)-3' in the major groove as well as contacts with the phosphate backbone within the minor groove at bases flanking the minimal 5'-GGA(A/T)-3' core (48). Thus, targeting one portion of the EBS consensus may be sufficient to impede ERG/DNA binding. This is what we observed here using DB1255 and related derivatives. In terms of structure/activity relationships, changing di-thiophene rings in DB1255 for di-furans or di-selenophenes abolishes sequence-selectivity and the ERG/DNA binding inhibitory effect.

Among the di-thiophene series in which the di-phenyl rings of DB1255 were changed for di-benzimidazole or in which di-methyl-thiophene or di-thiazole substituted the di-thiophene central rings, DB1255 remains the most active. Such modifications may increase the steric hindrance relative to DB1255, thus leading to the loss of DNA recognition. Sulphur atoms are less electronegative than oxygen atoms and develop weaker hydrogen bonding but present better stacking properties than furan groups. However, the size of the sulphur atoms is larger than that of oxygen (but smaller than selenium atoms) and this can increase the C-S-C bond angle of the thiophene rings of DB1255 in comparison with the C-O-C angle in furan rings (DB914). The selenium atoms may induce a much larger angle that is no longer compatible with proper DNA binding (DB1282). The optimal recognition seems to require a balance between thiophene properties and the curvature of this type of molecule for matching the minor groove of the double helix of DNA. The different capacity of inhibition was confirmed by an EMSA assay (Figure 2) where the more active compounds detected in the EDPBi assay decrease the ERG/DNA complex in contrast with the inactive compounds. As observed for DB293, the inhibition potency of active compounds strictly correlates with their ability to partially cover the EBS minimal 5'-GGA(A/T) and to span the 5'-GTT sequence located 3' to the minimal EBS sequence (Figure 4). The active di-thiophene heterocyclic diamidine compounds inhibit ERG/DNA complex formation through DNA interaction on one portion of the EBS, leading to the destabilization of ERG/DNA binding. As DB1255 is clearly the best ERG/DNA complex inhibitor, we focused on that

compound for fine analysis of sequence and cellular validation. As usual for this class of compound, DB1255 is a (minor) groove DNA binder as evidenced using both non-specific CT-DNA and a sequence-specific EBS-containing oligonucleotide (Figures 3 and Supplementary Figure S3). The strong bisignate CD spectra observed for this compound suggest a groove interaction as a dimer similar to DB293 (46,49). The SPR results suggest that the difference in inhibition effects of DB1255 relative to other compounds such as DB1998 is due to several factors. DB1255 is more specific than DB1998 and should therefore have more DB1255 molecules focused at specific sites and not lost in non-specific binding. DB1255 also dissociates from the specific site more slowly than DB1998 and this may help in modulation of protein–DNA complexes and function.

About the sequence-selective DNA recognition by DB1255, we brought out the importance of the nucleotides 5'-(g)TT following the core 5'-GGAA that showed an AT-rich sequence interaction separated by a G base as observed for other compounds (49). In parallel, studies on ERG selective binding showed, as expected, the necessity of the core 5'-GGAA-3' for optimal ERG/DNA complex formation. These results also showed the importance of adjacent bases for binding, particularly the 5'-GT dinucleotide following the core 5'-GGAA (DNaseI footprinting and SPR, Figure 6).

Interestingly, such consensus binding site strictly correlates with that deduced earlier by Wei et al. (16) from ChIP-sequence analyses. This reinforces the therapeutic interest in our compound with the objective to target *in vivo* the strong and selective ERG/DNA binding sites as defined using ChIP-sequencing. At the cellular level, the transient transfection assays confirmed the gradual decrease by DB1255 of ERG/DNA binding in the cellular model using both a synthetic promoter model and the OPN gene promoter (Figure 7).

As a transcriptional factor, ERG interacts with the DNA to positively or negatively regulate target gene expression depending on the cell context. Thus, our selected compound may positively or negatively interfere with the transcription of ERG-controlled genes, but in a way that may induce a change in the phenotype of ERG-dependent cancers. However, DB1255 interacts with DNA on one portion of the defined ERG-binding site, suggesting that it may not modulate all ERG-regulated genes but a portion of those that include the proper DB1255 binding site as defined here (Figure 5C). Thus, further experiments will be required to characterize the implication of the ERG/DNA binding, modulated by the selected compound regarding changes in gene expression. Genome-wide analyses (ChIP-seq and transcriptome analysis) in physiological conditions, disease conditions and in presence of the DNA ligand would be of major interest and could be compared with cells invalidated for ERG expression using RNA interference. One part of the question is provided by recent data of the invalidation of ERG expression in the TMPRSS2-ERG positive VCaP prostate cancer cell line evidencing the upregulation of genes associated with differentiated luminal prostate

epithelial cells (50). In HUVEC endothelial cells, similar transcriptome profiling evidenced the cytosolic histone-deacetylase-6 gene (*HDAC6*) as a direct ERG transcriptional target implicated in cell migration (51). The vascular endothelial adhesion molecule *VE-cadherin* was also described as a target for ERG, associated with angiogenesis (52). Those genes and approaches would be of major interest to validate the effect of ERG/DNA binding inhibitory compounds such as DB1255.

In conclusion, we selected a di-(phenyl-thiophene-amidine) compound as a DNA-binding modulator that specifically targets the ERG/DNA complex. To our knowledge, this is the first *in cellulo* modulation of the ERG/DNA binding by a small molecule targeting the protein/DNA interaction. As a strong inhibitor, DB1255 thus offers a therapeutic window for targeting ERG oncogenicity both in leukaemia disease and in prostate cancer.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–9.

## ACKNOWLEDGEMENTS

We thank the CHRU de Lille and the Région Nord/Pas-de-Calais for a PhD fellowship (to R.N.); the Institut National du Cancer (INCa) for post-doctoral fellowships (to X.D. and S.F.), and the Institut pour la Recherche sur le Cancer de Lille (IRCL) for technical expertise (to S.D.). The IMPRT-IFR114 is acknowledged for giving access to the Storm 860 equipment.

## FUNDING

The Fonds Européen de Développement Régional (FEDER, European Community) together with the Région Nord/Pas-de-Calais (to M.-H.D.-C. and M.D.-C.); the Ligue Nationale Contre le Cancer (Comité du Pas-de-Calais, Septentrion), the Association Laurette Fugain and the Association pour la Recherche sur le Cancer (to M.-H.D.-C.); NIH NIAID [064200 to W.D.W. and D.W.B.]; the CHRU de Lille and the Région Nord/Pas-de-Calais for a PhD fellowship (to R.N.); the Institut National du Cancer (INCa) for post-doctoral fellowships (to X.D. and S.F.); the Institut pour la Recherche sur le Cancer de Lille (IRCL) for technical expertise (to S.D.). Funding for open access charge: IRCL.

*Conflict of interest statement.* None declared.

## REFERENCES

- Darnell, J.E. Jr (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 740–749.
- Koehler, A.N. (2010) A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol.*, **14**, 331–340.
- Huang, M.E., Ye, Y.C., Chen, S.R., Chai, J.R., Lu, J.X., Zhao, L., Gu, L.G. and Wang, Z.Y. (1988) Use of all-trans retinoic acid in the treatment of acute promyelocytic leukaemia. *Blood*, **72**, 567–572.

4. Vassiley, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kamalott, U., Lukacs, C., Klein, C. *et al.* (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science*, **203**, 844–848.
5. Siddiquee, K., Zang, S., Guida, W.C., Blaskovich, M.A., Greedy, B., Lawrence, H.R., Yip, M.L., Jove, R., McLaughlin, M.M., Lawrence, N.J. *et al.* (2007) Selective chemical probe inhibitor of stat3, identified through structure-based virtual screening, induces antitumor activity. *Proc. Natl Acad. Sci. USA*, **104**, 7391–7396.
6. Wang, Y., Kheir, M.M., Chai, Y., Hu, J., Xing, D., Lei, F. and Du, L. (2011) Comprehensive study in the inhibitory effect of the berberine on gene transcription, including TATA box. *PLoS One*, **6**, e23495.
7. Doss, R.M., Marques, M.A., Foister, S., Chenoweth, D.M. and Dervan, P.B. (2006) Programmable oligomers for the minor groove DNA recognition. *J. Am. Chem. Soc.*, **128**, 9074–9079.
8. Raskatov, J.A., Meier, J.L., Puckett, J.W., Yang, F., Ramakrishnan, P. and Dervan, P.B. (2012) Modulation of the NF- $\kappa$ B-dependent gene transcription using programmable DNA minor groove binders. *Proc. Natl Acad. Sci. USA*, **109**, 1023–1028.
9. Zhang, Y., Sicot, G., Cui, X., Vogel, M., Wuertzer, C.A., Lezon-Geyda, K., Wheeler, J., Harki, D.A., Muzikar, K.A. *et al.* (2011) Targeting a DNA binding motif of the EVI1 protein by a pyrrole-imidazole polyamide. *Biochemistry*, **50**, 10431–10441.
10. Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M. and Dervan, P.B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc. Natl Acad. Sci. USA*, **95**, 12890–12895.
11. Vlamincik, B., Toffoli, S., Ghislain, B., Demazy, C., Raes, M. and Michiels, C. (2007) Dual effect of echinomycin on hypoxia-inducible factor-1 activity under normoxic and hypoxic conditions. *FEBS J.*, **274**, 5533–5542.
12. Peixoto, P., Liu, Y., Depauw, S., Hildebrand, M.P., Boykin, D.W., Bailly, C., Wilson, W.D. and David-Cordonnier, M.H. (2008) Direct inhibition of the DNA-binding activity of POU transcription factors Pit-1 and Brn-3 by a selective binding of a phenyl-furn-benzimidazole dication. *Nucleic Acids Res.*, **10**, 3341–3353.
13. Karim, F.D., Urness, L.D., Thummel, C.S., Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C., Maki, R.A., Gunther, C.V. and Nye, J.A. (1990) The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev.*, **4**, 1451–1453.
14. Nye, J.A., Petersen, J.M., Gunther, C.V., Jonsen, M.D. and Graves, B.J. (1992) Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.*, **6**, 975–990.
15. Hollenhorst, P.C., Shah, A.A., Hopkins, C. and Graves, B.J. (2007) Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev.*, **21**, 1882–1894.
16. Wei, G.H., Badis, G., Berger, M.F., Kivioja, T., Palin, K., Enge, M., Bonke, M., Jolma, A., Varjosalo, M., Gehrke, A.R. *et al.* (2010) Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.*, **29**, 2147–2160.
17. Oikawa, T. and Yamada, T. (2003) Molecular biology of the Ets family of transcription factors. *Gene*, **303**, 11–34.
18. Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R. *et al.* (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*, **310**, 644–648.
19. Demichelis, F., Fall, K., Perner, S., Andren, O., Schmidt, F., Setlur, S.R., Hoshida, Y., Mosquera, J.M., Pawitan, Y., Lee, C. *et al.* (2007) TMPRSS2:ERG gene fusion associated with lethal prostate cancer watchful waiting cohort. *Oncogene*, **26**, 4596–4599.
20. Rajput, A.B., Miller, M.A., De Luca, A., Boyd, N., Leung, S., Hurtado-Coll, A., Fazli, L., Jones, E.C., Palmer, J.B., Gleave, M.E. *et al.* (2007) Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. *J. Clin. Pathol.*, **60**, 1238–1243.
21. Salek-Ardakani, S., Smootha, G., de Boer, J., Sebire, N.J., Morrow, M., Rainis, L., Lee, S., Williams, O., Israeli, S. and Brady, H.J. (2009) ERG is a megakaryocytic oncogene. *Cancer Res.*, **69**, 4665–4673.
22. Baldus, C., Burmeister, T., Martus, P., Schwartz, S., Gökbüget, N., Bloomfield, C.D., Hoelzer, D., Thiel, E. and Hofmann, W.K. (2006) High expression of the ETS transcription factor ERG predicts adverse outcome in acute T-lymphoblastic leukemia in adults. *J. Clin. Oncol.*, **24**, 4714–4720.
23. Metzler, K.H., Dufour, A., Benthaus, T., Hummel, M., Sauerland, M.C., Heinecke, A., Berdel, W.E., Büchner, T., Wörmann, B., Mansmann, U. *et al.* (2009) ERG expression is an independent prognostic factor and allows refined risk stratification in cytogenetically normal acute myeloid leukemia: a comprehensive analysis of ERG, MN1, and BAALC transcript levels using oligonucleotide microarrays. *J. Clin. Oncol.*, **27**, 5031–5038.
24. Moore, S.D., Offor, O., Ferry, J.A., Amrein, P.C., Morton, C.C. and Dal Cin, P. (2006) ELF4 is fused to ERG in a case of acute myeloid leukemia with a t(X;21)(q25-26;q22). *Leuk. Res.*, **30**, 1037–1042.
25. Kim, J., Park, T.S., Song, J., Lee, K.A., Hong, D.J., Min, Y.H., Cheong, J.W. and Choi, J.R. (2009) Detection of FUS-ERG chimeric transcript in two cases of acute myeloid leukemia with t(16;21)(p11.2;q22) with unusual characteristics. *Cancer Genet. Cytogenet.*, **194**, 111–118.
26. Sorensen, P.H.B., Lessnick, S.L., Lopez-Terrada, D., Liu, X.F., Triche, T.J. and Denny, C.T. (1994) A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat. Genet.*, **6**, 146–151.
27. Delattre, O., Zucman, J., Melot, T., Garau, X.S., Zucker, J.M., Lenoir, G.M., Ambros, P.F., Sheer, D., Turc-Carel, C., Triche, T.J. *et al.* (1994) The Ewing family of tumors—a subgroup of small-round-cell tumors defined by specific chimeric transcripts. *N. Engl. J. Med.*, **331**, 294–299.
28. Erkizan, H.V., Scher, L.J., Gamble, E., Barber-Rotenberg, J.S., Sajwan, K.P., Üren, A. and Toretsky, J.A. (2011) Novel peptide binds EWS-FLI1 and reduces the oncogenic potential in Ewing tumors. *Cell Cycle*, **10**, 3397–3408.
29. Erkizan, H.V., Kong, Y., Merchant, M., Schlottman, S., Barber-Rotenberg, J.S., Yuan, L., Aboan, O.D., Chou, T.H., Dakshanamurthy, S., Brown, M.L. *et al.* (2009) A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. *Nat. Med.*, **15**, 750–756.
30. Rahim, S., Beauchamp, E.M., Kong, Y., Brown, M.L., Toretsky, J.A. and Üren, A. (2011) YK-4-279 inhibits ERG and ETV1 mediated prostate cancer cell invasion. *PLoS One*, **6**, e19343.
31. Grohar, P.J., Griffin, L.B., Yeung, C., Chen, Q.-R., Pommier, Y., Khanna, C., Khan, J. and Helman, L.J. (2011) Ecteinascidin 743 interferes with the activity of EWS-FLI1 in Ewing sarcoma cells. *Neoplasia*, **13**, 145–153.
32. Grohar, P.J., Woldemichael, G.M., Griffin, L.B., Mendoza, A., Chen, Q.R., Yeung, C., Currier, D.G., Davis, S., Khanna, C., Khan, J. *et al.* (2011) Identification of an inhibitor of the EWS-FLI1 oncogenic transcription factor by high-throughput screening. *J. Natl Cancer Inst.*, **103**, 962–978.
33. Björkman, M., Iljin, K., Halonen, P., Sara, H., Kaivanto, E., Nees, M. and Kallioniemi, O.P. (2008) Defining the molecular action of HDAC inhibitors and synergism with androgen deprivation in ERG-positive prostate cancer. *Int. J. Cancer*, **123**, 2774–2781.
34. Fortson, W.S., Kayarthodi, S., Fujimura, Y., Xu, H., Matthews, R., Grizzle, W.E., Rao, V.N., Bhat, G.K. and Reddy, E.S. (2011) Histone deacetylase inhibitors, valproic acid and trichostatin-A induce apoptosis and affect acetylation status of p53 in ERG-positive prostate cancers cells. *Int. J. Oncol.*, **39**, 111–119.
35. Ismail, M.A., Boykin, D.W. and Stephens, C.E. (2006) An efficient synthesis of 2,5'-diarylbiarylphosphines. *Tetrahedron Lett.*, **47**, 795–797.
36. Ismail, M.A., El Bialy, S.A., Brun, R., Wenzler, T., Nanjunda, R., Wilson, W.D. and Boykin, D.W. (2011) Dicationic phenyl-2, 2'-bichalcophenes and analogues as antiprotozoal agents. *Bioorg. Med. Chem.*, **19**, 978–984.
37. Carrere, S., Verger, A., Flourens, A., Stehelin, D. and Dutertre-Coquillaud, M. (1998) Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. *Oncogene*, **16**, 3261–3268.

38. Duterque-Coquillaud, M., Niel, C., Plaza, S. and Stehelin, D. (1993) New human erg isoforms generated by alternative splicing are transcriptional activators. *Oncogene*, **8**, 1865–1873.
39. Flajollet, S., Tian, T.V., Flourens, A., Tomavo, N., Villers, A., Bonnelye, E., Aubert, S., Leroy, X. and Duterque-Coquillaud, M. (2011) Abnormal expression of the ERG transcription factor in prostate cancer cells activates osteopontin. *Mol. Cancer Res.*, **9**, 914–924.
40. Baillat, D., Leprivier, G., Régnier, D., Vintonenko, N., Bègue, A., Stehelin, D. and Aumercier, M. (2006) Stromelysin-1 expression is activated in vivo by Ets-1 through palindromic head-to-head Ets binding sites present in the promoter. *Oncogene*, **25**, 5764–5776.
41. Baert, J.L., Monte, D., Verreman, K., Degerny, C., Coutte, L. and De Launoit, Y. (2010) The E3 ubiquitin ligase complex component COP1 regulates PEA3 group member stability and transcriptional activity. *Oncogene*, **29**, 1810–1820.
42. Genès, C., Lenglet, G., Depauw, S., Nhili, R., Prado, S., David-Cordonnier, M.H., Michel, S., Tillequin, F. and Porée, F.H. (2011) Synthesis and biological evaluation of N-substituted benzo[c]phenanthrolines and benzo[c]phenanthrolinones as antiproliferative agents. *Eur. J. Med. Chem.*, **46**, 2117–2131.
43. Racane, L., Tralic-Kulenivic, V., Kraljevic Pavelic, S., Ratkaj, I., Peixoto, P., Nhili, R., Depauw, S., Hildebrand, M.-P., David-Cordonnier, M.-H., Pavelic, K. *et al.* (2010) Novel diamidino-substituted derivatives of phenyl benzothiazolyl and dibenzothiazolyl furans and thiophenes: synthesis, antiproliferative and DNA binding properties. *J. Med. Chem.*, **53**, 2418–2432.
44. Liu, Y., Kumar, A., Depauw, S., Nhili, R., David-Cordonnier, M.H., Lee, M.P., Ismail, M.A., Farahat, A.A., Say, M., Chackal-Catoen, S. *et al.* (2011) Water-mediated binding of agents that target the DNA minor groove. *J. Am. Chem. Soc.*, **133**, 10171–10183.
45. Rahimian, M., Kumar, A., Say, M., Bakunoy, S.A., Boykin, D.W., Tidwell, R.R. and Wilson, W.D. (2009) Minor groove binding compounds that jump a GC base pair and bind to adjacent AT base pair sites. *Biochemistry*, **48**, 1573–1583.
46. Eriksson, M. and Norden, B. (2001) Linear and circular dichroism of drug-nucleic acids complexes. In: Chaires, J.B. and Waring, M.J. (eds), *Drug-nucleic Acids Interactions*, Vol. 340. 1 edn, pp. 68–98.
47. Workman, C.T., Yin, Y., Corcoran, D.L., Ideker, T., Stormo, G.D. and Benos, P.V. (2005) enoLOGOS: a versatile web tool for energy normalized sequence logos. *Nucleic Acids Res.*, **33**, W389–W392.
48. Szymczyna, B.R. and Arrowsmith, C.H. (2000) DNA binding specificity studies of four ETS proteins support an indirect read-out mechanism of protein-DNA recognition. *J. Biol. Chem.*, **275**, 28363–28370.
49. Bailly, C., Tardy, C., Wang, L., Armitage, B., Hopkins, K., Kumar, A., Schster, G.B., Boykin, D.W. and Wilson, W.D. (2001) Recognition of ATGA sequences by the unfused aromatic dication DB 293 forming stacked dimers in the DNA minor groove. *Biochemistry*, **40**, 9770–9779.
50. Tomlins, S.A., Laxman, B., Varambally, S., Cao, X., Yu, J., Helgeson, B.E., Cao, Q., Prensner, J.R., Rubin, M.A., Shah, R.B. *et al.* (2008) Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia*, **10**, 177–188.
51. Birdsey, G.M., Dryden, N.H., Shah, A.V., Hannah, R., Hall, M.D., Haskard, D.O., Parsons, M., Mason, J.C., Zvelebil, M., Gottgens, B. *et al.* (2012) The transcription factor Erg regulates expression of histone deacetylase 6 and multiple pathways involved in endothelial cell migration and angiogenesis. *Blood*, **119**, 894–903.
52. Birdsey, G.M., Dryden, N.H., Amsellem, V., Gebhardt, F., Sahnan, K., Haskard, D.O., Dejana, E., Mason, J.C. and Randi, A.M. (2011) Transcription factor Erg regulates angiogenesis and endothelial apoptosis through VE-cadherin. *Blood*, **111**, 3498–3506.