Cell Research(1995),5,75–91

FT I, anovel positive myeloid-lineage-specific transcription regulatory element within the mouse myeloperoxidase gene enhancer, En 1[1]

ZHU JINGDE
CRC Department of Medical Oncology, CRC Beatson Laboratories, Glasgow University, Garscube Estate, Switchback Road, Glasgow G61 1BD, United Kingdom. Tel: 44141942 9361; Fax: 44141 330 4127. e-mail: gpma66@udcf.gla. ac. uk

ABSTRACT

FT I (AAAAGGGGAAGCAGAG), a poly purine element within the myeloid-lineage specific enhancer (En 1) of the mouse myeloperoxidase gene [1, 2] has been further characterised. 1, FT I functions as a myeloid-lineage specific transcription regulatory element; 2, WEHI 3BD+ cells have higher binding activity to FT I and express the proteins which could form the unique DNA-protein complex(es) of FT I;. 3, The essential sequence for the specific DNA-protein interactions of FT I is AAAAGGGGAAGC; 4, South-western analysis in conjunction with the competition assay of the proteins binding to FT I, has revealed a 28 kd protein in WEHI 3BD+ cells that displays the properties of the putative transcription factor which acts through FT I. These new findings have demonstrated both the functional myeloid-lineage specificity and the novelty of FT I.

Key words: *Mouse myeloperoxidase gene, poly purine element, transcription activation, South-western analysis.*

INTRODUCTION

Myeloid cell differentiation, in which multipotential progenitor cells are converted

1. This work is dedicated to the 80th anniversary of my supervisor Professor Zhen YAO.

into one of the six mature differentiated cells, i.e., erythrocytes, platelets (megakaryocytes), macrophages, neutrophils, eosinophils and basophils, involves temporal regulation of expression of a number of lineage- and differentiation stage-specific genes. Understanding the developmental specification of lineage as well as maturation stage associated patterns of gene expression in myeloid cell differentiation requires an insight into the control of individual lineage-specific 'marker' genes for each lineage. Within the myeloid lineages, extensive studies have illustrated a complex interplay of regulatory cis-elements ahd nuclear binding proteins, in particular for the β-globin enhancer 5' locus control region in the erythrocytic lineage[3]. On the other hand, there is a marked deficit of information, in this respect, on the granulocytic lineage, a major pathway of non-erythroid myeloid cell development.

Myloperoxidase (MPO) gene expression is regulated in both cell-lineage and differentiation-stage specific fashions. The MPO protein and bactericidal activities reach maximum at the promyelocytic stage of granulocytic differentiation; as these cells mature into granulocytes, levels of MPO mRNA rapidly decline, probably as a result of both a cessation of transcription and a decrease in stability of the MPO mRNA[4]. By a transient transfection assay, we have identified a myeloid-lineage-specific enhancer at approximately 3.1 kb upstream of the mouse MPO gene[2]. It acts and resides at a DNase I hypersensitive chromatin region in WEHI 3BD+ cells, a myelomonocytic cell line[4, 5] in which the mouse MPO gene is highly expressed, but not in two non-MPO expressing lymphocytic cell lines {pre-B lymphocytic cell line, 18.8[6] and pre-T lymphocytic cell line, EL-4[7]}. *In vitro* DNase I footprinting of the enhancer revealed three previously unreported protected sequences (FT I, II and III), which were associated with proteins enriched in WEHI 3BD+ cells; but not in both lymphocytic cell lines[l]. Functional analyses by means of deletion- and site-specific mutations followed by a transient-transfection assay have demonstrated the synergistic nature of these three modular cis-elements, FT I, II and III. Consistent with loss of function, the corresponding wild type *in vitro* DNase I footprints were also affected by site-specific mutations[l].

In this report, several important aspects of FT I, a poly purine motif of En 1 have been addressed. Firstly, does FT I have the same positive myeloid-lineage-specific transcription regulatory activity as En 1? Secondly, which nucleotides are essential for the specific DNA-protein interactions of FT I? And lastly, what are the molecular weight and binding characteristics of the proteins enriched in WEHI 3BD+ cells which bind to FT I?

MATERIALS AND METHODS

Plasmid constructs, cell cultures, and the transient transfection assay

Equal moles of the two complementary oligonucleotides of FT I (sense: 5'-gggAAAAGGGGAAGCAGAGCT-3'; and antisense: 5'-cccagCTCTGCTTCCCCTTTT-3') were 5'end phosphorylated by T4-DNA polynucleotide kinase, heat-denatured and annealed. After ligation at room temperature overnight, the polymers bigger than 4 mer were purified by spin column chromatography (using

Clontech, Sp-400 column). After filling their 3 recessive ends with Klenow fragment and dNTP, FT I polymers were cloned at Sma I site of Pt-109 vector {the firefly luciferase gene is controlled by Herpes thymidine kinase promoter spanning from -109 to +52, [8]}. The desired clones were identified by sequencing.

WEHI 3BD+, a myelomonocytic leukaemic cell line[5, 9], 18.8[6] (a pre-B cell line derived from BALB /c bone marrow infected with Abelson leukaemia virus, and EL-4 {a C57BL murine T lymphoma (ATCC TIB39)} cell line[7] were cultured as previously described[2]. Both transient iransfection by electroporation and luciferase assay were carried out as in the previously published procedures[2]

Band-shift assays for the sequence-specific protein-DNA interactions

The whole cell extracts were prepared f rom the late-log phase of cell cultures of WEHI 3BD+, 18.8 and EL-4 cell lines respectively, as previously described[1]. The protein extract was incubated on ice for 20 min with 10 fmole of ^{32} p-end-labelled double stranded FT I oligonucleotides, in a 10 μl solution containing 5 μg poly dI:dC (Pharmacia), 10 % polyvinyl alcohol (Sigma, p-8136), 12.5 mM Herpes, pH 7.9, 6.25 m*M* $MgC1_2$, 50 mM KC1 and 10 % glycerol. The DNA-protein complexes were separated by electrophoresis on a 1 % agarose gel with 0.5 X TBE (1 × TBE= 89 m *M* Tris, 89 m *M* boric acid, 8 m*M* EDTA, pH 8.0) at room temperature (voltage setting: 6 volts per cm). In the competition assay, 10-50 fold molar excess of the individual double stranded oligonucleotides were included in the binding reaction. After electrophoresis, the gels were dried onto a piece of DEAE paper (Whatman DE 81) before autoradiography was taken.

For the assay in polyacrylamide gel, the binding conditions were same as that for the agarose gel system, except that the amount of poly dI:dC was 1 μg.

1, FT I: 5′ gggAAAAGGGGAAGCAGAGct

2, FT I-1a 5′ gggTTAAGGGGAAGCAGAGct

3, △1 5′ gggAAGGGGGGAAGCAGAGct

4, FT I-2 5′ gggAAAACCCCAAGCAGAGct

5, FT I-2a 5′ gggAAAACCGGAAGCAGAGct

6, FT I-2c 5′ gggAAAAGGCGAAGCAGAGct

7, FT I-2d 5′ gggAAAAGGGCAAGCAGAGct

8, FT I-3 5′ gggAAAAGGGGTTOGAGAGct

9, FT I-3a 5′ gggAAAAGGGGTFGCAGAGct

10, FT I-3b 5′ gggAAAAGGGGAACGAGAGct

11, FT I-3c 5′ gggAAAAGGGGTAGCAGAGct

12, FT I-3d 5′ gggAAAAGGGGAJGCAGAGct

13, △Ia4 5′ gggAAAAGGGGAATCJGAGct

14, FT I-4 5′ gggAAAAGGGGAAGCTCTCct

15, △3c6 5′TCGACGTCCTAGTCGTCTGCGAAGCTACTCGTTGACCCC;

16, FT II, 5′ CCCTAAATAGAGTTGAGAAAGAGCAATC1TFTGCAACCAGT

17 Sp 1, 5′ CTAGAGCTAAGGCCCCGCCCCCACCAAGGT [10]

18,CTCF (2x) [10]

AGCT(TGAGCCCCCTCGGCCGCCCCCTCGCGGCGCGCCCTCCCCGC)ACTA

G(TGAGCCCCTCGGCCGCCCCCTCGCGGCGCGCCCTCCCCGC)T;

19, FT I (2x):

(GGGAAAAGGGGAAGCAGAGCT)(GGGAAAAGGGGAAGCAGAGCT)GGG;

N. B., 1, The sense strand sequences of the double stranded oligonucleotides are shown. In the case of FT .I, the CAPITAL letters refer to the protected sequence, revealed by the *in vitro* DNase I footprinting assay with the WEHI 3BD+ extract [l].

Myeloid-lineage—specific transcription regulatory element

2, The underlined sequences are the altered nucleotides in the mutants.
3, After re-annealing, a filling reaction with Klenow fragment and dNTP was carried out to eliminate the 5' protruding ends.

South—western analysis of the proteins binding to FT I

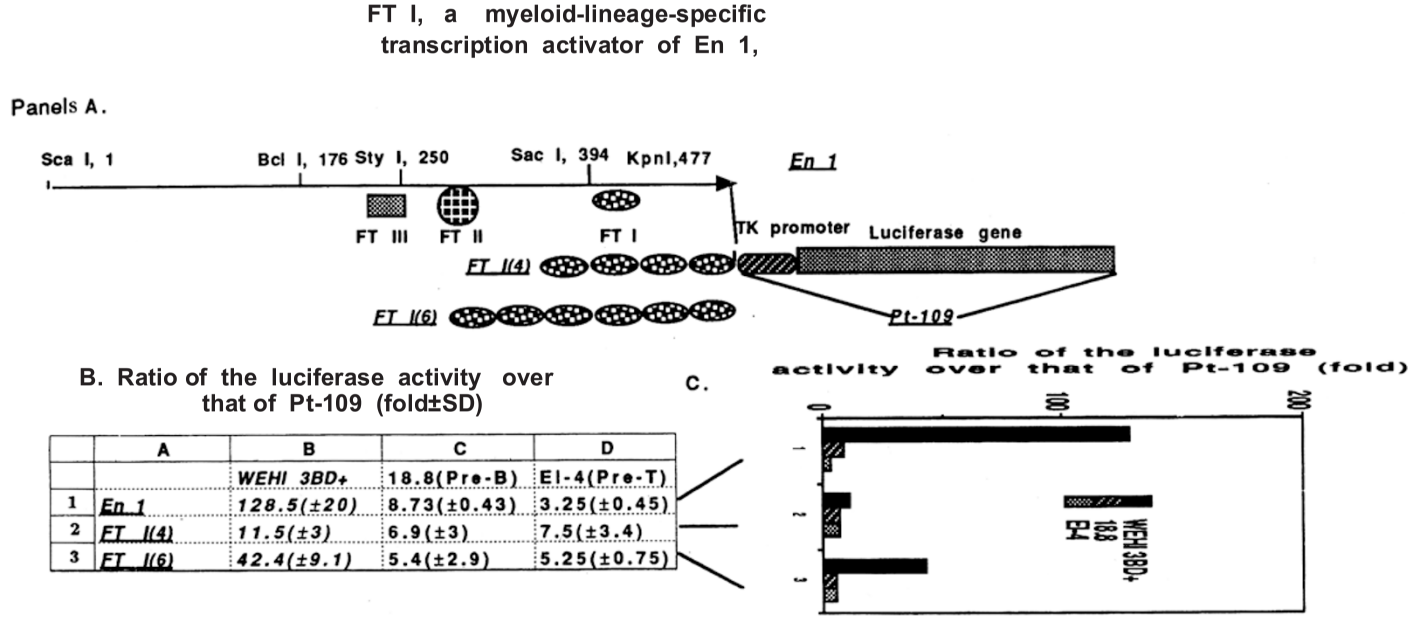
Protein extracts were denatured at 37℃ for 10 min in 1×sample buffer (5% SDS / 5% glycerol / 37. 5 mM Tris. HCl, pH 7: 0 / 5%2-mercaptoethanol), fractionated on a 10% SDS—polyacrylamide gel by electrophoresis[11] and electroblotted onto nitrocellulose membrane using the LKB Noval blotter under conditions recommended by the manufacturer. The molecular weight references were provided by the rainbow markers (Pharmacia). The following procedures were carried out in a cold room with gentle shaking. The membranes, on which the proteins were immobilised were incubated with prebinding buffer (20 mM Herpes, pH 7.9, 0.2%Triton 50 mM KCl, and 5%fat-free milk powder[Marvel], 10 μM ZnS_{04}, 1 mM DTT and 10% glycerol) for 20 min, followed by incubating with the DNA-binding buffer(20 mM Herpes, pH 7.9, 0.2%Trition, 50 mM KCl, 10 μM ZnS_{04}, 1 mM DTT+1 μg / ml denatured salmon sperm DNA+10 μg / ml poly dI: dC and 10%glycerol)for another 20 min, before the addition of the ^{32}P-labelled double stranded oligonucleotide probes (01 pmole). In the competition aSSays, 50—100 fold molar excess of double stranded oligonucleotides was included in the binding reaction. After 20 min incubation. the unbound probes were washed away with DNA binding bufief minus nucleic acid competitors for 10 min. The washing was repeated twice. The pattern of the DNA-binding proteins to the probes were visualised by either autoradiography or phosphor-imaging(Molecular Dynamics).

RESULTS

FT I , in the tandem repeat form, functions as positive myeloid lineage-specific cis-element for transcription

En 1, like many previously identified and characterised eukaryotic enhancers[12], is a complex enhancer; in this case, composed of three modular cis—elements which are all required for the potent myeloid-lineage-specific transcription activation function[1]. It has been shown for the majority of eukaryotic enhancers, one single modular cis-element can mimic the function of the natural enhancer when it is in a polymer form, suggesting that both relative position and the orientation of the modular components in the natural enhancer are not critical[12]. It is,therefore, important to determine whether En 1 belongs to this category.

Four and six tandem repeats of FT I were cloned upstream of the Herpes viral thymidine kinase promoter (in the same orientation) in Pt-109 and transfected by electroporation into WEHI 3BD+, 18.8 and EL-4 cells, respectively (Fig 1). The ratio of the luciferase activity between the testing constructs and Pt-109 was used to represent the potency of FT I and En 1 transcription activation function. It has been found that: four and six tandem repeats of FT I were able to enhance the luciferase activity more dramatically(11.5 and 42.9 fold)in myeloid WEHI 3BD+cells than in lymphoid 18.8 (6.9 and 5.4 fold)or EL-4 cells (7.5 and 5.2 fold, respectively). Also, a six mer of FT I was mors potent than a four met in WEHI 3BD+cells, while the reverse was true for lymphoid 18.8 and EL-4 cells. Although the mecha—nism underlying this effect is unclear. the result clearly shows that FT I can function as a positive myeloid-lineage-specific transcription regulatory cis-element on its own.



	A	B	C	D
		WEHI 3BD+	18.8(Pre-B)	EI-4(Pre-T)
1	En 1	128.5(±20)	8.73(±0.43)	3.25(±0.45)
2	FT I(4)	11.5(±3)	6.9(±3)	7.5(±3.4)
3	FT I(6)	42.4(±9.1)	5.4(±2.9)	5.25(±0.75)

Myeloid-lineage-specific transcription regulatory element

Fig 1. Myeloid-lineage-specific transcription activation function of FT I, in the tandem repeat form.
Panel A, Schematic illustration of the constructs used in the study, Pt-109, En 1, FT I[4], and FT I[6]. The three WEHI 3BD+ cell type-specific in vitro DNase I footprints (FT I, II, and III) are shown.
Panel B, The ratios of the luciferase activities of each constructs over that of Pt-109, in transfected cells, representing as fold+ SD;
Panel C, Graphic illustration of the data presented in panel B.

Quantitative and qualitative differences in the specific DNA-protein interactions of FT I among WEHI 3BD+, 18.8 and EL-4 cells

The previous work showed that three WEHI 3BD+ cell-specific *in vitro* DNase I footprints (FT I, II, and III) are likely to represent the essential protein-DNA interactions for the enhancer function of En 1 *in vivo*[1]. Since the *in vitro* DNase I footprinting assay is relatively insensitive, I used the band-shift assay to reassess the previous observations, by which the sequence-specific DNA-protein interactions at a much lower level can be detected.

Bearing in mind that the differential transcription of a given gene can be dictated by quantitative variations of specific DNA-protein interactions with cell types[13], I have controlled the protein input of the whole cell extracts from different cell lines, by taking the binding activity of a common transcription factor, Sp-1, as the refer-

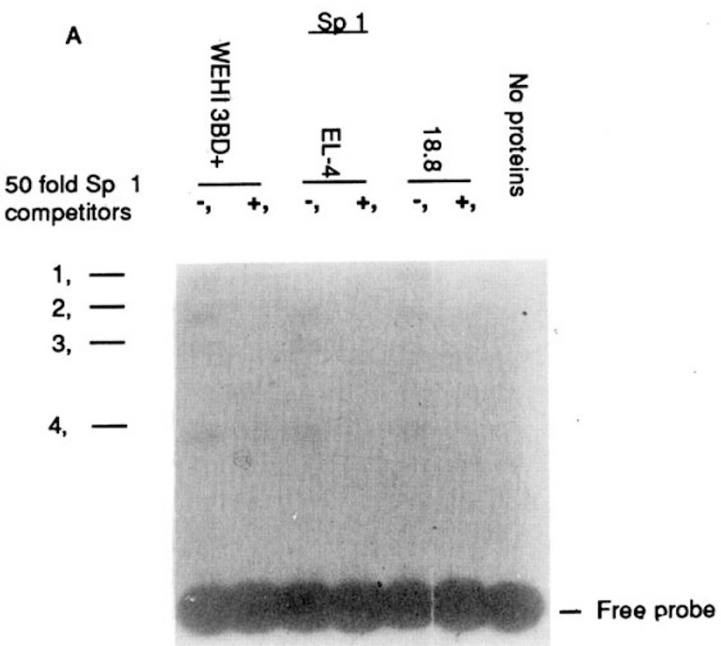


Fig 2A. The relative binding activity of Sp 1 transcription factor
The Sp 1 binding activity. Approximately 0.5 μg protein extracts from each cell samples indicated at the top of the autorad were incubated with 10 fmole of ^{32}p-labelled double stranded oligonucleotides of Sp 1 binding sequence[10] without and with 50 fold molar excess of the cold same oligonucleotides and analysed by band-shift assay in 1 % agarose gel. Both the shifted bands (1-4) and the sense stranded sequence of Sp 1 are indicated.

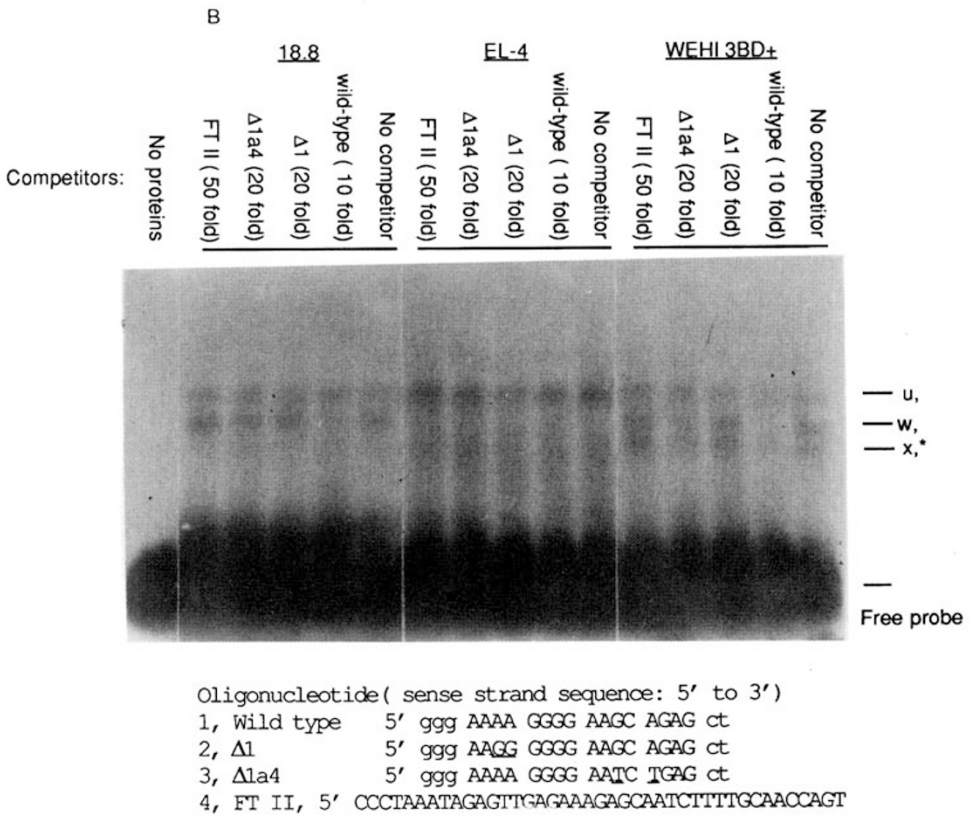
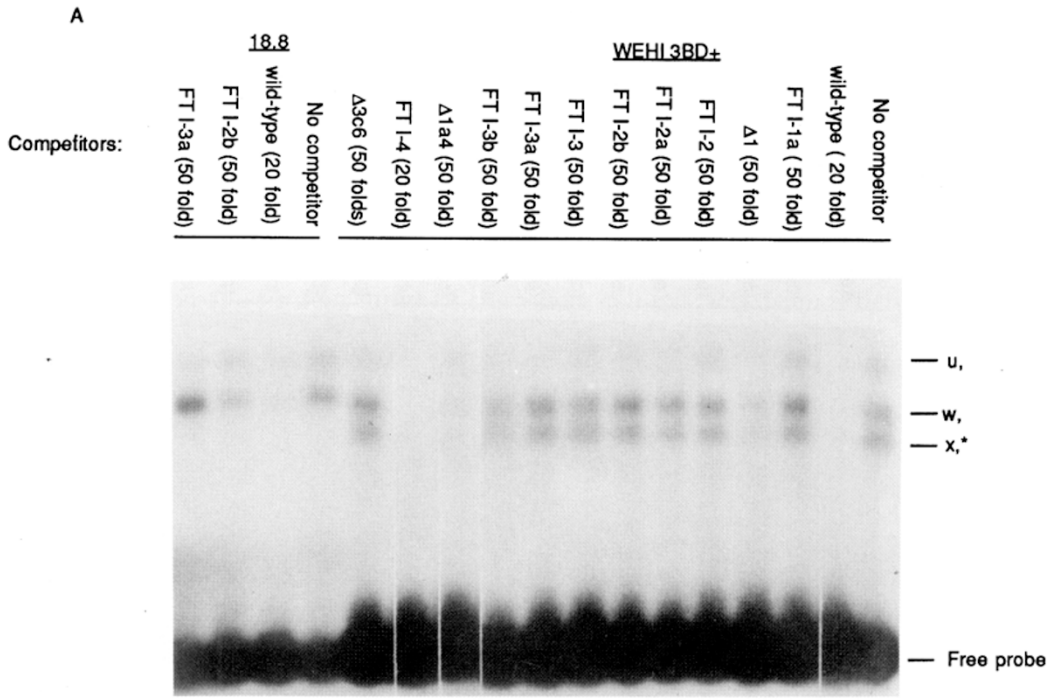


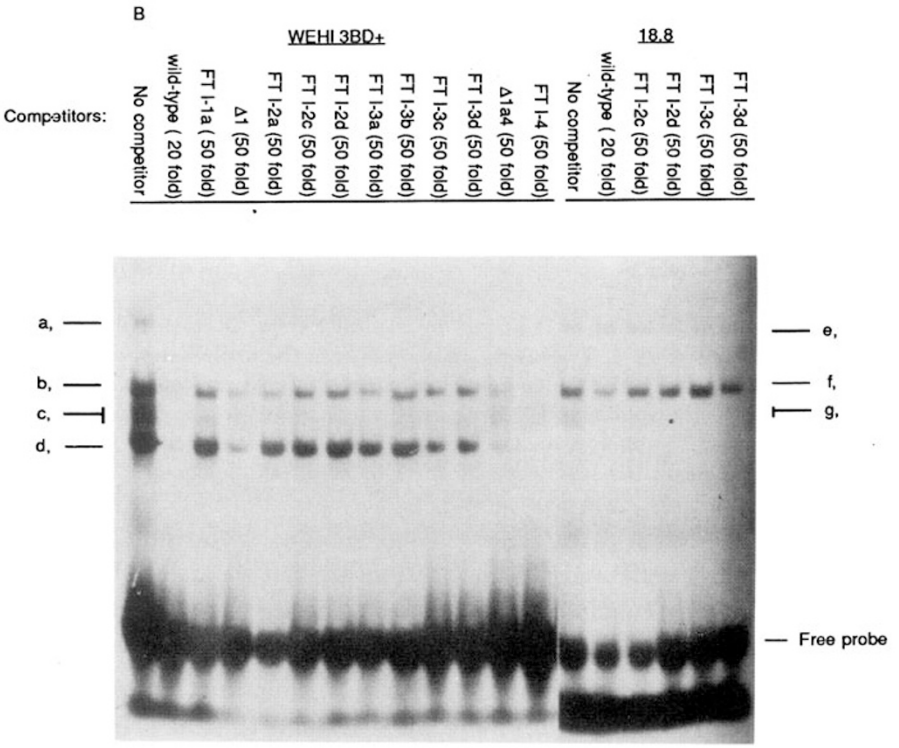
Fig 2B. FT I binding proteins among WEHI 3BD+, 18.8 and EL-4 cells
Cell-type specificity of the sequence-specific DNA protein interactions of FT I
0.5 μg WEHI 3BD+, 7.5 μg EL-4 and 3 μg 18.8 whole cell extracts as indicated at the top of the autorad were respectively, incubated with ^{32}p-labelled double stranded oligonucleotides of FT I sequence, without and with 10 fold of FT I, 20 fold of $\triangle$1, 20 fold of $\triangle$la4, 50 fold of FT II oligonucleotides (which sense strand sequences are indicated at the bottom of the figure) and analysed by band shift assay in 1 % agarose gel. The shifted band (u, w, x*) and free probe band are indicated.

ence. As shown in Fig 2A, Sp 1 binding activity did not vary significantly with cell lines, where 0.5 μg protein extracts were used. The only observed difference is that among the four shifted bands (designated 1, 2, 3 and 4, in ascending order of the distance shifted) for WEHI 3 BD+ and 18.8 cells, only three corresponding bands



1, Wild type	5′ ggg AAAA GGGG AAGC AGAG ct
2, Δ1	5′ ggg AAGG GGGG AAGC AGAG ct
3, FT 1-1a	5′ ggg TTAA GGGG AAGC AGAG ct
4, FT 1-2	5′ ggg AAAA CCCC AAGC AGAG ct
5, FT 1-2a	5′ ggg AAAA CCGG AAGC AGAG ct
6, FT 1-2b	5′ ggg AAAA GGCC AAGC AGAG ct
7, FT I-3	5′ ggg AAAA GGGG TTCG AGAG ct
8, FT 1-3a	5′ ggg AAAA GGGG TTGC AGAG ct
9,FT 1-3b	5′ ggg AAAA GGGG AACG AGAG ct
10,Δ1a4	5′ ggg AAAA GGGG AATC TGAG ct
11,FT 1-4	5′ ggg AAAA GGGG AAGC TCTC ct
12,Δ3c6	5′TCGACGTCCTAGTCGTCTGCGAAGCTACTCGTTGACCCC

Fig 3A. The effect of the mutations on the wild-type FT I-DNA-protein interactions
A, 0.5 μg WEHI 3BD+ and 3 μg 18.8 whole cell extracts as indicated at the top of the autorad were respectively, incubated with ^{32} p-labelled double stranded oligonucleotides of FT I sequence, without and with 20 fold of FT I, as well as 50 fold molar excess of FT I mutants and △-3c6 double stranded oligonucleotides (as indicated at the top of the autorad) and analysed by band shift assay in 1% agarose gel. The shifted bands (u, w, x*) and free probe band are indicated. The sense strand sequence of the oligonucleotides are indicated, and the underlined nucleotides are the altered nucleotides in FT I mutants.



1, Wild type	5′ ggg AAAA GGGG AAGC AGAG ct
2, FT 1-1a	5′ ggg TTAA GGGG AAGC AGAG ct
3, Δ1	5′ ggg AAGG GGGG AAGC AGAG ct
4, FT 1-2a	5′ ggg AAAA CCGG AAGC AGAG ct
5, FT 1-2c	5′ ggg AAAA GGCG AAGC AGAG ct
6, FT 1-2d	5′ ggg AAAA GGGC AAGC AGAG ct
7, FT 1-3a	5′ ggg AAAA GGGG TTGC AGAG ct
8, FT I-3b	5′ ggg AAAA GGGG AACG AGAG ct
9, FT 1-3c	5′ ggg AAAA GGGG TAGC AGAG ct
10,FT 1-3d	5′ ggg AAAA GGGG ATGC AGAG ct
11,Δ1a4	5′ ggg AAAA GGGG AATC TGAG ct
12,FT 1-4	5′ ggg AAAA GGGG AAGC TCTC ct

Fig 3B. The effect of the mutations on the wild-type FT I-DNA-protein interactions 0.5 μg WEHI 3BD+ and 3 μg 18.8 whole cell extracts as indiacted at the top of the autorad were respectively, incubated with ^{32}p-labelled double stranded oligonucleotides of FT I sequence, with or without 20 fold of FT I, as well as 50 fold molar excess of various FT I mutants in the form of double stranded oligonucleotides (as indicated at the top of the autorad) and analysed by band shift assay in the 4 % polyacrylamide gel. The shifted bands (a to g) and free probe band are indicated.

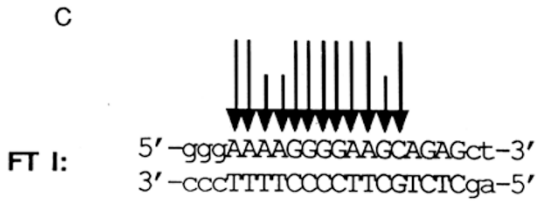


Fig 3C. The effect of the mutations on the wild-type FT I-DNA-protein interactions
Summary of the essential nucleotide sequence (sense) for the wild-type FT I DNA-protein interaction.
The 12 nucleotide sequence indicated by capital letters, is essential to the specific protein binding of FT I. The longer arrows indicate the nucleotides, upon mutation, led to a dramatic loss of the competition capacity of the corresponding mutants in comparison to the wild-type FT I sequence with respect to the DNA-protein interaction of FT I. The shorter arrows indicate the nucleotides whereby upon mutation elicited a significant but less dramatic effect on competition capacity.

(2, 3 and 4) were present with EL-4 cells (Fig 2A). Although the cause of this missing band in EL-4 cells and its implication are not clear, the results (Fig 2A) confirmed the previous conclusion drawn from the *in vitro* DNase I footprinting assay[1], namely, Sp 1 binding activity is invariant with these three cell lines.

On the contrary, for a comparable level of FT I binding, approximately 6 fold (3 μg) more 18.8 and 15 fold (7.5 μg) more EL-4 extracts than WEHI 3 BD+ extracts (0.5 μg) were required (Fig 2B). Furthermore, the pattern of protein binding to FT I for these three cell line lines was different. There were three discrete shifted bands (designated as u, w, and x*) associated with WEHI 3BD+ cells, but only band u was asociated with EL-4 cell, while both bands u and w were present in the 18.8 extract. Therefore, band x* represented a WEHI 3 BD+ cell-specific DNA-protein interaction of FT I.

In competition studies, both bands w and x* were abolished by 10 fold molar excess of FT I oligonucleotides, while the full elimination of band u needed more FT I competitor (Fig 3). It has been previously shown that $\triangle$1 and $\triangle$1a4 mutations had deleterious effects on En 1 function and abolished the occurrence of the wild-type FT I in an *in vitro* Dnase I footprinting assay[2]. In agreement with this finding, 20 fold molar excess of both mutants were not sufficient to abolish bands u, w, and x* in band-shift assay. Furthermore, the formation of bands u, w, and x* were not affected by 50 fold excess of FT II which was used as an unrelated sequence control.

The essential sequence for the specific DNA-protein interaction of FT I (AAAA-GGGGAAGCAGAG) is AAAAGGGGAAGC

In order to determine the nucleotide sequences essential for the specific DNA-protein interactions of FT I, a set of FT I mutants were made in which one to four nucleotides were individually mutated and assayed for their competition potency on the wild-type DNA-protein interactions of FT I. By using 20 to 50 fold molar excess of FT I mutants, the relative importance of the individual nucleotides to the squence-specific binding have been approximately estimated by band-shift assays (in agkrose gel, Fig 3A, or in a 4 % polyacrylamide gel, Fig 3B).

Although DNA-protein complexes bigger than the pore size of 4 % polyacrylamide gel fail to enter the gel and can not be analysed, both quantitative and qualitative aspects of the myeloid-lineage-specificity of the protein binding to FT I shown in 1% agarose gel (Fig 2B, and Fig 3A.) stillheld (Fig 3B, the EL-4 data are not shown). Four shifted bands (a to d) associated with WEHI 3 BD+ cells can be effectively competed away by 20 fold of wild type FT I oligonucleotide and 50 fold molar excesss of mutant FT I-4. Among the three shifted bands (e, f and g) with 18.8, bands f and g are likely to be similar if not identical to bands b and c of WEHI 3 BD+ cells, according to their mobility and response to the specific competitors. Band c in WEHI 3BD+ cells was also present at a lower level in 18.8 cells (as band g), suggesting that this DNA-protein complex may not have an important role in the positive myeloid-lineage-specific-transcription regulatory function of FT I. Among the two bands (a and d) representing the DNA-protein interactions of FT I specific to WEHI 3BD+ cells, band a might not be functionally important, sa it can be abolished by all the FT I mutants tested. On the contrary, band d was resistant to various extent to the excess FT I mutants (except for FT I-4), including mutants △1 and △la4 which had deleterious effects on En 1 function[1]. This indicates that band d is likely to represent the WEHI 3 BD+ cell-specific-DNA-protein interaction of FT I with a direct functional relevance. At the present time, there is very little known on how the protein-DNA complexes behave in the agarose band-shift system, therefore, the relationship between band x* in agarose gel (Fig 3A) and band d in polyacrylamide gel (Fig 3B) is not clear.

As shown in Fig 3, the sequence, AAAAGGGGAAGC is essential for the sequence-specific DNA-protein interactions of FT I. However, the relative effect of the mutations varies with the nucleotides being affected on the wild type DNA-protein interactions of FT I (for a summary, see Fig 3C).

The putative transcription factor for FT I: a 28 kd protein in WEHI 3 BD+ cells ?

Having characterised the sequence of FT I essential for the specific DNA-protein interaction, I used South-western analysis in conjunction with competition assay to address the following three important aspects: the cell type-specificity, the binding characteristics and the molecular weight of the WEHI 3BD+ specific FT I binding proteins.

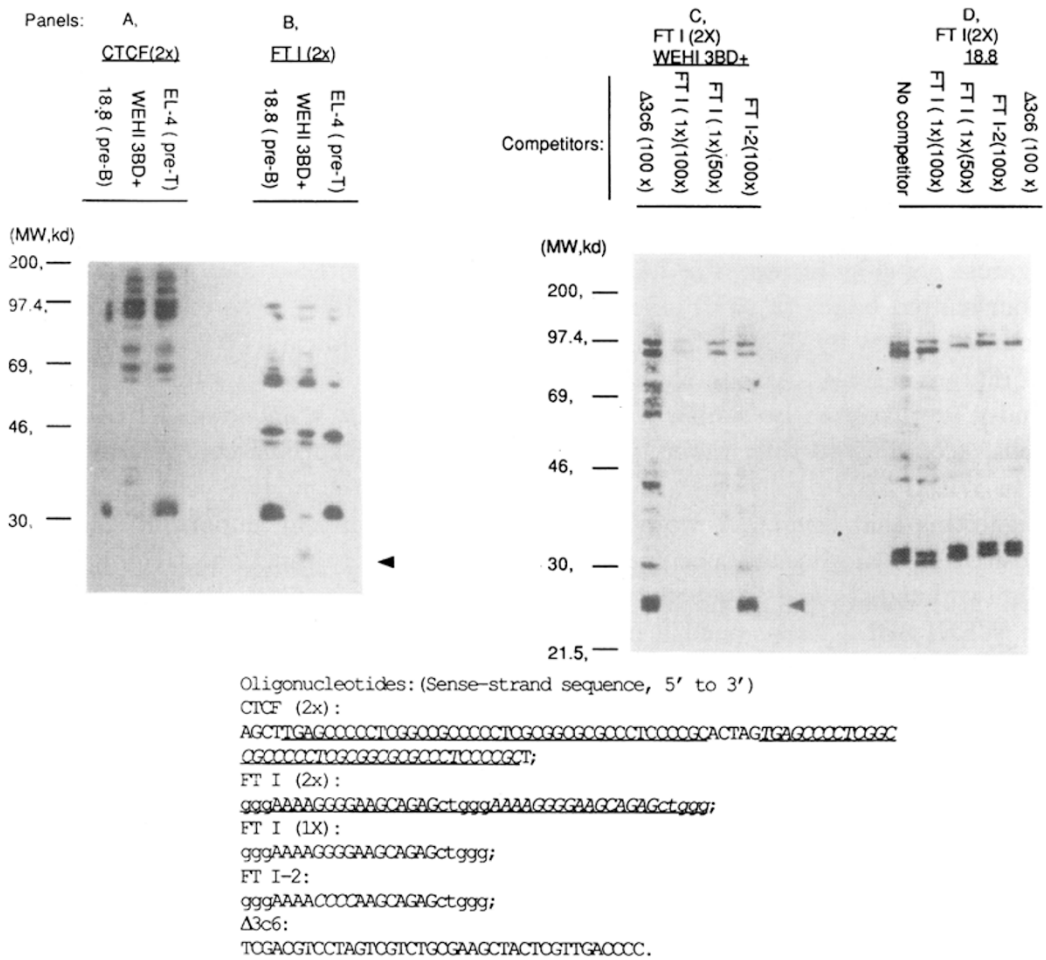


Fig 4. South-western analysis of the DNA-binding proteins specific to FT I sequence
An approximately 7.5 μg proteins from each cell lines (as indicated at top of the Figure) were separated by electrophoresis on a 10 % SDS-polyacrylamide gel and electroblotted onto nitrocellulose membrane. The immobilised proteins were bound to ^{32}p-labelled CWCF(2x)(panel A) or FT I(2x) (panels B-D) double-stranded oligonucleotides. For the competition assays, 50-100 fold molar excess of specific competitors (as indicated at the bottom of the Figure) were included in the binding reaction as indicated at the top of the Figure. Size markers were provided by Rainbow markers (Pharmacia). The WEHI 3BD+ cell type-specific 28 kd protein bound to FT I probe is indicated with an arrow. The sense-strand sequences of the double stranded oligonucleotides are shown at the bottom of the gel.

In order to eliminate tne background binding by the abundant non-specific DNA binding protein in cells, a common problem encountered in South-western analysis, the protein binding pattern to an unrelated sequence, i.e. the chicken CTCF element
{upstream of the chicken c-myc gene[10]}, was established (panel A, Fig 4). By identifying the common bands, the bands representing the specific proteins bound to either CTCF or FT I sequences can be detected. In addition, both 18.8 and EL-4 cell extracts were analysed in parallel, WEHI 3BD+ cell-type-specificity of binding could also be addressed. Although there were some protein bands showing a certain degree of binding specificity to the FT I probe (panels A and B, Fig 4), only one protein band had a sufficient WEHI 3BD+ cell-type-specificity. This protein was approximately 28 kd in molecular weight (panel B, Fig 4), and did not bind to CTCF sequence (panel A, Fig 4). Its specificity to FT I sequence has been further demonstrated by the competition assay with FT I (wild-type) FT I-2 (a FT I mutant which failed to compete with wild-type FT I for the formation of FT I-protein complexes in the band shift analyses) and △3c 6, an unrelated sequence (panel C, Fig 4). The FT I binding by this 28 kd protein was abolished by 100 fold molar excess, and affected significantly by 50 fold molar excess of FT I oligonucieotide, but not affected at all by 100 fold mutant FT I-2 as well as 100 fold molar excess of △3c6 sequences (panel C).

DISCUSSION

FT I, a novel positive myeloid-lineage-specific transcription regulatory element

The modular nature of the enhancer was initially observed in the enhancer of the SV-40 early gene, in which duplication of one wild type module can fully compensate for the loss of other modules[14]. Since then, this phenomenon has been reported in many eukaryotic enhancers, where the function of the natural enhancer can be mimicked by one of such elements when multimerised and placed upstream of a test promoter (for a review see[15]). Both relative positions and orientation of the modules in the natural enhancer are not essential for the function of these so called non-stereo-specific enhancers[8]. However, there are several recent exceptions, including the mouse T cell receptor a gene enhancer[16] and the virus-inducible enhancer of the human interferon β gene[17, 18], in which the function of the natural enhancer can not be reproduced by the tandem repeats of its individual modular components. In both cases, one minor groove binding transcription factor is involved. This type of protein specifically binds to its cognate recognition sequence and induces a sharp bend in DNA to the extent that other transcription factors bound to the flanking cis-elements can interact with each other to form the functional, stereo-specific enhancer complexes[12]. It naturally follows that once a complex enhancer (in this case, En 1) consisting of several integral modular elements is identified, the first priority is to determine whether or not it is stereo-specific.

As shown in Fig 1, FT I, like the natural enhancer, En 1, from which it is derived,

was able to activate the transcription of the luciferase gene guided by the Herpes thymidine kinase promoter in a myeloid-lineage specific manner, suggesting that En 1 is a nonstereo-specific enhancer. However, the tandem repeat of FT I is much less potent. Four and six mers of FT I were able to enhance the luciferase expression by approximate 8 % and 30 % of that of En 1, respectively. In light of the fact that FT I is the weakest of three WEHI 3 BD+ cell-specific in vitro DNase I footprints within En 1[1], and both FT II and FT III polymers are much more potent at this aspect (unpublished observation), there is an interesting quantitative correlation between the DNA-protein interactions and the function of FT I.

Previous database searching[l] have shown of no significant sequence homology between the sequences of FT I, FT II, as well as FT III and the binding sites recognised by the known transcription factors [16-18], an indication of the novelty of the three cis elements, FT I, FT II, and FT III. In the case of FT I, the 16 nucleotide DNase I protected sequence: AAAAGGGGAAGCAGAG was used for this analysis. In this report, by assaying the mutation effects on the wild-type DNA-protein interactions of FT I, it has been shown that the critical nucleotide sequence of FT I is a 12 nucleotide sequence: *AAAAGGGGAAGC* (the underlined nucleotides are the relatively more important sequences, Fig 3C). By searching the same database[19-21] with this 12 nucleotide sequence, I am still not able to identify any previously reported cis-elements which has more than 50% homology with FT I.

Observations from band shift assays as well as South-western analyses are also supportive of the novelty of FT 1. A WEHI 3 BD+ cell specific shifted band of FT I were both found in both agarose gel and 4 % polyacrylamide gel system (Fig 3), representing a WEHI 3 BD+ cell-specific protein binding to the essential sequence of FT I. And the size of this protein is 28 kd (panels C and D, Fig 4). Again, to the best of my knowledge, there is no previous report of any transcription factor with these characteristics, including size, expression pattern and binding specificity.

Binding characteristics of the FT I binding proteins enriched in WEHI 3BD+ cells

Three methods are available for assessing the molecular weight of the protein(s) which bind specifically to a given DNA sequence. These are, 1, UV-cross-linking of ^{32}P-labelled double stranded target sequences to the proteins prior to SDS-polyacrylamide gel electrophoresis[22]; 2, 'South-western blotting' in which gel fractionated proteins are transferred from the gel to a membrane and, after renaturation *in situ*, are probed with a ^{32}P-labelled double stranded target sequences[23]; and 3, elution of size separated proteins from polyacrylamide slices[11] followed by an analysis of each fraction for the binding activity to the target sequence by band-shift assays[24]. However, not all methods can provide the information on a very important question, whether the target protein binds to the DNA sequence in a heterodimer,

or homodimer and monomer forms. In addition, the measurement from UV-cross linking may suffer from the fact that the binding kinetics and specificty of the target protein to DNA can be altered by substitution of dT by BrdU in the target DNA sequences[25]. On the other hand, recovering the functional proteins from the gel slices is labour-intensive and error-prone.

Establishing the protein-binding-pattern to an unrelated cis-element, CTCF as a control, has made it possible to eliminate the background binding to FT I by the abundant non-specific DNA binding proteins in cell. Moreover, setting both 18.8 and EL-4 cell-extracts as controls in this assay, has led to an unambiguous identification of the WEHI 3 BD+ specific FT I binding proteins. The binding specificity to FT I of the 28 kd protein in WEHI 3 BD+ cells has been finally confirmed by the competition assay with the specific oligonucleotides, including the wild-type and mutant FT I (FT I-2) (panels B and C, Fig 4). By this assay, it has also been shown that WEHI 3 BD+ cell-specific 28 kd protein binds to FT I in either the monomer or homodimer form and can survive through the denaturation and renaturation of South-western analysis. These findings will be invaluable for the cloning of the gene encoding this 28 kd protein in the near future.

The expression cloning using the lambda gt11 vector based on the specific DNA-protein interaction[26] has the advantages of simplicity, speed and of a function-related screening procedure, but is not applicable for cloning the genes encoding proteins which either can not survive the denaturing and renaturing steps of the screening procedure or bind to DNA in heterodimer form. But, it is obvious that this approach should be fully applicable to the cloning of the gene encoding the WEHI 3BD+ cell-type-specific 28 kd protein.

Since South-western analysis has its own limitations as previously discussed, the possibility can not be excluded at the moment that there axe other transcription factors binding to FT I with functional importance in WEHI 3 BD+ cells. However, the following findings do argue strongly otherwise. Band shift analyses (Fig 2 and 3), showed a single WEHI 3BD+ cell specific shifted band, i.e., band d in polyacrylamide gel and band x* in agarose gel (Fig 3). The possibility that both shifted bands may be associated with a single WEHI 3BD+ cell type specific protein is very appealing. There is only one protein (28 kd in molecular weight) exhibiting the all essential characteristics of the key transcription factor vital to the myeloid-lineage-specificity of the FT I function (Fig 4). Nevertheless, the final answer of this question as well as the proof of the novelty of FT I can only be provided by the detailed analyses of the cloned 28 kd protein at aspects of both in vivo function and *in vitro* binding characteristics to FT I.

ACKNOWLEDGEMENTS

This work is supported by Cancer Research Campaign, U. K. Thanks are due to E. Klenova for providing Sp 1 and CTCF (2x) oligonucleotides, and Ashworth A,

Myeloid-lineage-specific transcription regulatory element

Davis F, Liou A, Reid A, and Zelent A. for critical reading the manuscript.

REFERENCES

[1] Zhu JD, The myeloid-lineage-specific enhancer of the mouse myeloperoxidase gene consists of three cis-elements defined as *in vitro* DNase I footprints. FEBS Letters 1994; **339**:243-8.

[2] Zhu JD, Bennett CA, MacGregor AD, Greaves MF, Goodwin GH, and Ford AM. A myeloid-lineage-specific enhancer upstream of the mouse myeloperoxidase gene. Leukaemia 1994; **8(5)**:717-23.

[3] Orkin SH. Globin gene regulation and switching: circa 1990. Cell 1990; **63**:665-72.

[4] Tobler A, Miller CW, Johnson KR, Selsted ME, Rovera G, Koeffler HP. Regulation of gene expression of myeloperoxidase during myeloid differentiation. J Cell Physiol, 1988; **136(2)**:215-25.

[5] Warner NL, Moore MAS, Metcalf D. A transplantable myelomonocytic leukaemia in BALB /c mice: cytology, karyotype, and muramidase content. J Natl Cancer Inst. 1969; **43**:963-82.

[6] Alt F, Rosenberg N, Lewis S, Thomas E, Baltimore D. Organisation of immunoglobulin genes in A-MuLV-transformed cells: rearrangement of heavy but not light chain gene. Cell 1981; **27**:381-90.

[7] Herberman RB. Serological analysis of cell surface antigens of tumours induced by murine leukaemia virus. J Natl Cancer Inst 1972; **48**:265-.

[8] McKnight SL, Kingsbury RC, Spence A, Smith M. The distal transcription signals of the herpes virus tk gene share a common hexanucleotide control sequence. Cell 1984; **37**:253-62.

[9] Metcalf D, Nicola NA. Autoinduction of differentiation in wehi-3b leukaemia cells. Int J Cancer 1982; **30**:773-80.

[10] Lobanenkov VV, Nicolas RH, Adler VV et al. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-Flankin sequence of the chicken c-myc gene. Oncogene 1990; **5**:1743-53.

[11] Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature 1970; **227**:680-5.

[12] Tjian R, Maniatis T. Transcriptional activation: a complex puzzle with few easy pieces. Cell 1994; **77**:5-8.

[13] McKnight S, Yamamoto Y. ed. Transcriptional regulation. Cold Spring Harbour Laboratory Press: New York. 1992.

[14] Herr W, Clarke J. The SV40 enhancer is composed of multiple functional elements that can compensate for one another. Cell 1986; **45**:461-70.

[15] Dynan WS. Modularity in promoters and enhancers. Cell 1989; **58**:1-4.

[16] Giese K, Cox J, Grosschedl R. The HMG domain of lyphoid enhancer factor 1 bends DNA and facilitatesassemnly of functional nucleoprotein structure. Cell 1992; **69**:185-96.

[17] Du W, Thanos D, Maniatis T. Mechanisms of transcriptional synergism between distinct virus-inducible enhancer elements. Cell 1993; **74**:887-98.

[18] Thanos D, Maniatis T. The high mobility group protein HMG I (Y) is required NF-kb-dependent virus induction of the human IFN-beta gene. Cell 1992; **71**:777-89.

[19] Faisst S, Meyer S. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res 1992; **20 (1)**:3-26.

[20] Locker J. Transcriptional controls: cis-elements and trans-factors, in Gene transcription, a practical approach. BD Hames, SJ Higgins, Eds, IRL Press: Oxford **1993**: pp. 321-45.

[21] Lopex M, Oettgen P, Akbarali Y, Dendorfer U, Libermann TA. ERP, a new member of the ets transcription factor/oncoprotein family: cloning, characterisation, and differential expression during B-lymphocyte development. Mol Cell Biol 1994; **14 (5)**:3292-309.

[22] Chodosh LA, Carthew RW, Sharp PA. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. Mol Cell Biol 1986; **6**:4723-33.

[23] Bowen B, Steinber J, Laemmli UK, and Weintraub H. The detection of DNA-binding proteins by protein-blotting. Nucleic Acids Res 1980; **8**:1-2.
[24] Briggs MR, Kadonaga JT, Bell SP, Tjian R. Purification and biochemical characterisation of the promoter-specific transcription factor, Sp 1. Science 1986; **234**:47-52.
[25] Smith KC, O'Leary ME. The pitfalls of measuring DNA synthesis kinetics as exemplified in ultraviolet radiation studies. Biochem Biophys Res Commun 1969; **34**:354-7.
[26] Singh H. Specific recognition site probes for isolating genes encoding DNA-binding proteins. In Methods in Enzymology. R Wu Ed. Academic Press: New York **1993**: pp. 551-67.

Received 7-4-1995 *Accepted 23-5-1995*