ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

PURIFICATION OF SPECIFIC DNA BINDING PROTEINS FROM HUMAN CELLS

Thesis Submitted for the Degree of
Magister Philosophiae

Candidate: Lidija Marušić

Supervisors:
Prof. Arturo Falaschi
Dr. Eva Csordas Toth

Academic Year 1989/90

TRIESTE

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

PURIFICATION OF SPECIFIC DNA BINDING PROTEINS FROM HUMAN CELLS

Thesis Submitted for the Degree of
Magister Philosophiae

Candidate:

Lidija Marušić

Supervisors:
Prof. Arturo Falaschi
Dr. Eva Csordas Toth

Academic Year 1989/90

ACKNOWLEDGMENTS

This work has been done at the International Centre for Genetic Engineering and Biotechnology. I would like to thank Prof. Arturo Falaschi who gave me the opportunity to explore this interesting field. This work could not have been possible without the unselfish help and advices of dr. Eva Csordas Toth in performing the experiments as well as in writing the thesis. I am also very grateful to dr. Georgine Faulkner, Alexander Ochem and Catherine Ludlow for useful suggestions.

CONTENTS

1. INTRODUCTION

1.1 Initiation of DNA replication in eukaryotic chromosomes

1.2 Origins: still elusive in mammalian cells

1.3 Transcription and replication

1.4 Sequence-specific DNA-binding proteins: active in transcription and DNA replication

1.5 The aim of the work

2. MATERIALS AND METHODS

3. RESULTS

3.1 Chromatographic steps

3.2 Characterization of purified fractions

4. DISCUSSION

4.1 More than one factor binds to the pB48 sequence homologous to the MLTF binding site

4.2 Is there a new family of regulatory proteins for MLTF binding site ?

5. BIBLIOGRAPHY

iate DNA synthesis at about 10^5 sites). One round of lication must be completed before starting another and iation must be regulated in a way that all regions of the omosome are replicated only once (3). This is very important ause growing cells can only maintain their genetic integrity if A replication is controlled precisely (4).

In eukaryotic cells, the chromosomes are replicated in a crete phase of the cell cycle, the S-phase. The eukaryotic omosome initiates replication at multiple sites along its length . Different sections of a chromosome replicate simultaneously, ming active sites that contain clusters of about 50 replication ks. The rate of elongation varies a little in different cell types (600 bases per minute) and the frequency of initiation of plication is the primary determinant of the length of S phase. idence for this came from studies in which different cell types a given organism were compared (5)

During cleavage in Drosophila embryos all the DNA is plicated in an interphase of 3-4 minutes. In somatic cells with S phase of 10 hours, replicating regions are infrequent; this can seen by autoradiography of whole chromosomes labelled with dioactive DNA precursors during S-phase or by fibre toradiography of DNA isolated cells pulse labelled with [^{3}H] ymidine (6,7). The embryo initiates replication at closer tervals along the chromosome, speeding up replication. Such a exible program of replication allows the differences in the length the S phase in different cells (4).

here are still questions to be answered:

ow can one be sure that each unit of DNA synthesis (replicon) itiates once in every cell cycle?

how is it possible to be sure that the replication process is complete?
are there specific DNA sequences that are used to initiate replication, in an analogous manner to prokaryotes where there is a single origin of replication in the chromosome?

1.2 ORIGINS: STILL ELUSIVE IN MAMMALIAN CELLS

The existence of origins in eukaryotic chromosomes has not been formally demonstrated. Numerous experiments have been performed often giving contrasting results. The evidence that initiation need not be sequence specific comes from experiments in which circular DNA molecules were injected into Xenopus laevis eggs. No specific DNA sequences were required for regulated replication (8). The same results were obtained in the experiments with extracts of such eggs (9).

Different results were provided by replication timing studies of the mouse immunoglobulin heavy chain constant region gene cluster and the dihydrofolate reductase (DHFR) gene region in Chinese hamster ovary (CHO) cells (10,11). The earliest replicating portion of an amplified DHFR domain in a mammalian cell line suggest the existence of two origins about 22 kb apart, both downstream of the DHFR gene (11). A new method using exponentially growing CHO cells containing the non-amplified, single copy DHFR gene locus could identify a chromosomal origin of DNA replication (12). The results showed that bidirectional DNA replication begins within an initiation zone about 2.5 kb long that is centered approximately 17 kb downstream of the DHFR gene.

The DNA sequence that contains the 2.5 kb initiation zone of the earliest replicating region has been examined for structural properties that may be related to origin function (13).

In the yeast Saccharomyces cerevisiae it has been found that certain sequence elements derived from chromosomal DNA allow plasmids containing these elements to transform recipient yeast cells at high frequency (14). The short sequence elements were called autonomously replicating sequences or ARS elements (14). Although there is not yet any direct cytological evidence that ARSs coincide with origins in chromosomes, this possibility has been widely assumed. There are approximately 400 such sequences per haploid genome and they replicate only during S phase and only once per cell cycle. Moreover replication in vitro initiates at ARSs (15,16).

After discovery of ARS elements in yeast numerous attempts have been made to detect similar elements in higher eukaryotic cells. In one attempt, a number of selectable plasmids was constructed. Human DNA sequence bearing a gene C418 for antibiotic resistance were inserted into these plasmids (17). No evidence of autonomous replication of recombinant plasmids could be detected. Similar results were obtained by many other investigators (18,19).

In a recent study, a new method has been employed for mapping origins (20). This method utilizes antibodies to bromodeoxiuridin (Brd U) to isolate nascent Brd U-labeled DNA strands. Nascent strand segments were amplified by polymerase chain reaction to examine replication initiated in vivo near the c-myc in human cells. The identified chromosomal DNA replication initiation zone coincides approximately with the regions of

autonomously replicating sequence (ARS) activity reported previously for plasmids transfected into HeLa or HL 60 cells, but the final conclusion cannot be drawn yet (21,22).

Different approaches have been tried to isolate mammalian replication origins by using physical means. In one experiment African green monkey kidney cells were arrested in the G1/S border using the DNA replication inhibitor aphidicolin (23). Nascent DNA from the onset of S-phase was isolated and cloned. The cloned fragments were examined for their time of replication by hybridization to cellular DNA fractions synthesized at various intervals of S phase. Four out of five studied sequences hybridized with early replicating fractions. Cloned fragments should be proximal to the origins and perhaps contain replication origins at or close to their centers. It was found that six "ors" sequences out of nine contain a 21 bp concensus sequence. There is an abundance of AT rich regions and a high frequency of potential "stem-loop" structures. "Stem-loop" structures show similarity between mammalian replication origins and prokaryotic and papaovavirus origins (23).

Using the synchronization of human promyelocytic HL 60 cells, sequences that replicated immediately after the onset of S-phase were isolated (24). Cells were synchronized by two blocks with aphidicolin. About 200 sequences synthesized immediately after the entry into the S-phase were cloned. The newly synthesized DNA showed a significant increase in snap-back DNA and no enrichment in repeated sequences. The two largest clones (pB48 - 1560 bp and pLC46 - 716 bp) were chosen for further analysis. They were completely sequenced and studied in detail. The longest early replicating DNA fragment, pB48, contains a number

of characteristic features showing its possible regulatory role in the cell (Figure 1). These are: i) an element of the human Alu family, present in the last 189 nucleotides

ii) a 600 bp long CpG- rich region with the properties of an HTF island (high frequency of Hpa II sites and CpG dinucleotide),

iii) three possible thermodynamically stable stem-loop structures with stability greater than -27Kcal/mol,

iv) three sequences homologous (70-100%) to the central (GC-rich) palindrome in the origin of human Papova virus JVC, two regions homologous (83-100%) to the binding site III of the human Papova virus BK T antigen,

v) some transcription signals in the CpG-rich region: three Sp 1 binding sites located within a short range, in the stem-and-loop area; a 9 bp sequence 70% homologous to the SV 40 and human Ig (κ chain) enhancer (24,25).

CAT assay experiments were performed to show a possible presence of an active promotor in pB48 (25). The pB48 CpG-rich area was inserted into the pCAT vector. Transcription occurs, but only in one orientation, towards the HTF island. In an *in vivo* experiment, using Nothern type hybridization, two bands were observed at 1250 and 1550 nucleotides when the mRNAs of actively growing HeLa cells were probed with pB48.

In the search for possible binding sites for specific nuclear factors by a band shift assay, the central 521 bp Ava I- Ava I fragment was used. After incubation with HeLa and HL 60 nuclear extract the appearance of retarded bands could be observed. Maximum shift was in the presence of 5mM Mg^{++} (24).

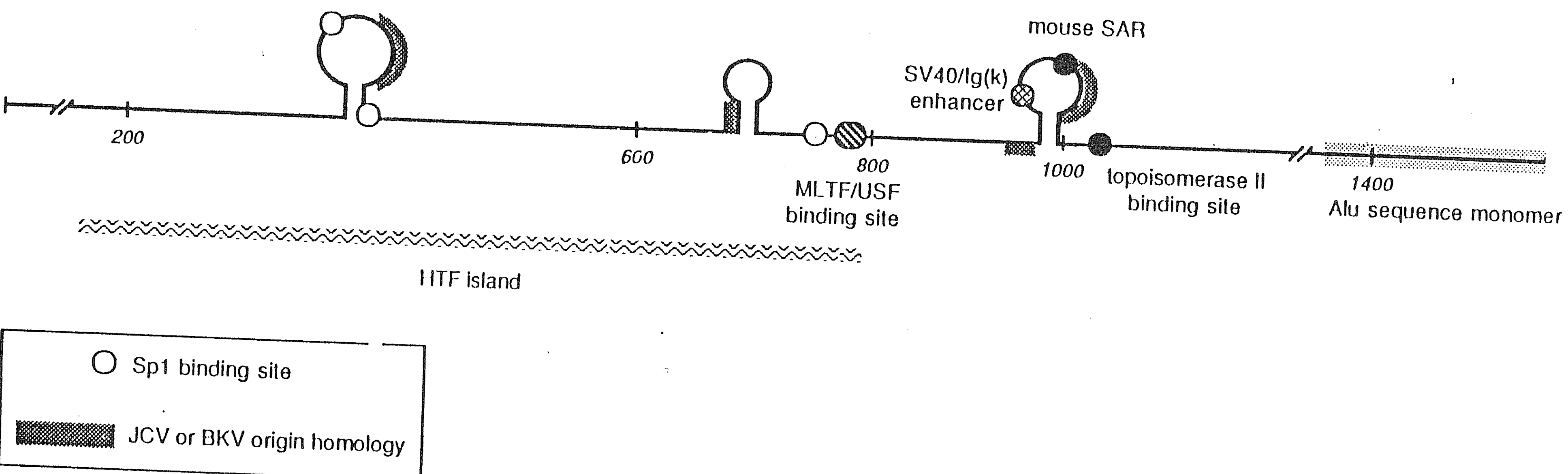


Figure 1. Schematic representation of pB48.

In order to analyze the binding site, the 195 bp fragment was submitted to DNase I footprinting analysis. A protected region of 17 nucleotides was observed on the L strand between nucleotides 801 and 785. The protected sequence 5'TTCGTCACGTGATGCGA3' is palindromic in the 12 central nucleotides except for one mismatch (24). It is nearly identical to the upstream element of the major late promoter (MLP) of adenovirus 2. Similar sequences can also be found in the long terminal repeat (LTR) of HIV-1 virus, in yeast centromeres, in the TRP1ARS1 element of yeast, in one of the Alu repeats, upstream of two homeobox genes of Xenopus and mouse and in the regulatory region of the chorion genes of several insects. The results of competition experiments showed that major late promoter of adenovirus and pB48 compete for the same factor. Lower but still significant competition was found with the sequence present in the LTR of HIV-1 (25).

Initiation of DNA replication uses the replication fork mechanism, as recognized in bacterial, plasmid, bacteriophage and animal virus genomes. The same mechanism has been assumed for the eukaryotic DNA replication, although only two of its characteristics have ever been demonstrated: The presence of Okazaki fragments and the presence of replication bubbles (26,27). In the attempt to sumarize all the data about mammalian replication origins, a model of replication initiation in higher eukaryotic cells has been proposed (28). According to the model a replication origin contains an origin of bidirectional replication with the sequence recognized by initiation proteins and a large initiation zone (10-30 kb) that is required for the initial

unwinding and strand synthesis events (29,30). But the complex mechanism of higher eukaryotic replication will be clarify only after the identification of all replication intermediates and proteins that interact with the origin.

1.3 TRANSCRIPTION AND REPLICATION

According to the data available there is a relationship between replication and gene expression. In one study a temporal order of replication of tissue-specific genes in nine types of differentiated cells has been examined (31). The results obtained showed that actively transcribed genes usually replicate during the first quarter of the S phase. Early replicating chromosomal domains consist of active tissue specific genes and flanking non-transcribed sequences. The domain, when inactive may replicate later in the S-phase of a different cell type, perhaps because of a different three-dimensional chromosomal configuration. Differentiated cell may use sequence specific origins, a specific, different subset of many potential origins existing in the mammalian genome. This might be important for the control of the timing and a site of replication. Unlike tissue-specific genes, house-keeping genes can replicate during any interval of the S-phase (31).

There is still little information about cellular origins of replication and the role of cis-acting transcriptional elements that might regulate cellular DNA replication. In eukaryotic viruses the situation is more clear (32). In the case of SV 40, polyoma virus (PyV) and adenovirus (Ad) there is a core component where replication begins and an auxiliary component with promoter or enhancer elements that may be involved in transcription as well

as replication. Enhancer and promoter elements function in either orientation, but they must be close to the AT-rich end (32). Ori cores from these viruses contain all cis-acting information necessary to initiate replication. But, as in the case of Ad 2 where the cellular protein called nuclear factor I (NF I) facilitates up to 10 fold the interaction of the initiation complex with ori core, proteins recognizing transcriptional elements may stabilize binding of initiation proteins to ori-core by promoting localized strand separation or modifying the chromatin structure of ori-core making it more accessible to the initiation protein complex (32). Bovine papilloma virus (BPV) and Epstein-Barr virus (EBV) enhancer and ori-core are separated by 300-1000 bp and function independently of their relative orientation and distance. In mitochondrial DNA, it is a transcription from an upstream promoter that provides RNA primers for initiation of DNA synthesis (32). There may be a role for cis-acting transcriptional elements in regulation of eukaryotic DNA replication. In yeast, repression of the yeast silent mating type loci requires cis-acting sequences located over 1 kb from regulated promoters. One of the silencers also behaves like an origin of DNA replication having ARS activity. It might be that DNA replication initiated at the silencer is also important for transcriptional repression (33).

The two longest sequences, pB48 and pLC46, isolated as putative replication origins from human cells contain numerous transcription signals (24,25). Future experiments may show if there is a correlation between transcription and activation of replication.

1.4 SEQUENCE-SPECIFIC DNA-BINDING PROTEINS: ACTIVE IN TRANSCRIPTION AND DNA REPLICATION

At the 5'-nontranscribed region of genes there are short, conserved sequences which are involved in regulating transcription of various genes. These so called recognition elements or cis-regulatory regions of genes interact with sequence-specific DNA binding proteins. Recently, many of them have been identified and they may be divided into two groups: ubiquitous factors and tissue-specific factors (34). Some of them can participate in both transcription and replication events (35,36). A great number of short DNA sequences that bind to nuclear proteins have been discovered using the footprinting technique (37). A DNA-protein complex is treated with DNase I to determine a zone protected by the protein against the enzyme action.

Different genes contain sets of regulatory sequences and the order in which they are located, as well as the distance between them, seem to be important (34). The stability of a DNA-protein complex depends on how protein and DNA contact surfaces fit together. It seems that the stability of binding may be enhanced by cooperative interactions between several proteins and DNA. Eukaryotic DNA binding proteins often bind to asymmetrical sequences, while prokaryotic regulatory DNA-binding proteins usually recognize palindromic sequences (38). One of the characteristics of eukaryotic DNA-binding proteins is that they recognize similar or identical DNA sequences (34). They are situated in the nucleus, but in a very low concentration and it is difficult to isolate them directly in quantities sufficient to

determine their complete amino acid sequences. The other approach involves first isolating the gene for a particular factor.

Analysis of the protein sequences of these factors showed some striking features. OTF-2 is a human transcription factor which participates in initiation of transcription of genes for immunoglobulin k and H chains (39). The cDNA sequence of OTF-2 contains a highly conserved structure called the homeobox, which is present in many genes that regulate embryonic development in Drosophila (40,41). The homeobox encodes a helix-turn-helix motif, the first and the best characterized structural motif for a DNA-binding domain. On the C-terminal end of OTF-2 there is a leucine zipper-like structure, probably responsible for the interaction with other factors and/or RNA polymesrase II.

Another DNA-binding motif is called the zinc-finger where cysteine and histidine residues interact with zinc atoms forming a tetrahedral coordination site for a single zinc ion, while the aminoacids between these coordination sites project out as fingers (42). Each finger motif binds to half a turn of DNA and the number of zinc fingers in different proteins ranges from two to ten or more (43).

There is a large number of nuclear factors that are Zn^{++} dependent like TFIIIA, a protein required for transcription of the 5SRNA genes by RNA polymerase III, Sp 1, an ubiquitous transcription factor, receptors for steroid hormones etc (43).Some transcription factors do not bind directly to DNA. S300-II is a protein which was isolated with the COUP, the factor that initiates transcription of the ovalbumin gene. Although S 300-II does not bind directly to DNA. it does stabilize the COUP-DNA complex (44).

Many of regulatory proteins contain at least two domains, one interacting with a certain DNA sequence while the other interacts with other proteins. Often, several protein factors are needed to regulate transcription of a gene, some of them interacting with the promoter and enhancer region of the gene and the others with other proteins of the transcriptional complex.

There are also factors like OTF-1 and NF 1 that may be involved in transcription as well as in DNA replication, depending on the location of the sequence they recognize and on their interaction with other factors (34). (Table1)

In recent years, much of the work has focused on specific DNA binding proteins that control transcription in eukaryotic cells but they may play a pivotal role in initiation of DNA replication as well.

Nuclear factor 1 (NF-1) is a cellular protein that enhances the initiation of adenovirus DNA replication (45). NF-1 binds a specific nucleotide sequence within the viral origin of replication. The consensus recognition sequence for NF-1 is TGG(A/C)N5GCCAA. It has been identified by analysis of adenovirus origin mutants and by comparison of NF-1 binding site formed in several viral and cellular DNAs (46,47). NF-1 has been purified using the DNA recognition site affinity chromatography (48). The affinity matrix was made by using a plasmid that contains 88 copies of the NF-1 binding site from the adenovirus origin of replication. SDS polyacrylamide gel electrophoresis of the most purified fraction, followed by silver stainig showed several major polypeptides with nuclear weights between 52 and 66 KDa (48).

MAMMALIAN PROTEIN FACTORS AFFECTING REGULATION OF TRANSCRIPTION AND OF DNA REPLICATION

FACTOR	*SEQUENCE MOTIF BOUND*	*TRANSCRIPTION*		*REPLICATION INITIATION*	
		GENE	FUNCTION	GENOME	FUNCTION
Sp1	CCGCCC	late genes of papova many housekeeping genes	upstream element binding	Papovavirus	stimulation of origin activation
T Antigen	GAGGC	early genes of papova	essential for initiation of transcription	Papovavirus	essential for origin activation
CTF/NF-I	GCCAAT	globin genes, genes stimulated by TGF-beta and many RNA pol II - transcribed genes	upstream element binding	Adenovirus 2 and 4	essential for initiation
NF-III	ATGCAAAT	H2B histone genes Immunoglobulin genes U1 and U2 snRNA genes SV40	upstream element binding upstream element and enhancer binding enhancer binding	Adenovirus 2 and 4	essential for initiation
MLTF/USF	PuPyCACGTGPuPy	Adenovirus late genes mouse metallothionein gene rat gamma-fibrinogen gene HIV-I LTR	upstream element binding negative regulation (?)	Human (HL60)	(?)

Table1.

Transcription factor CTF is responsible for selective recognition of eukaryotic promoters that contain sequence CCAAT (36). This sequence matches a portion of the concensus recognition sequence for the NF-1. CTF has been purified by sequence specific DNA affinity chromatography with CAAT-box elements from the human a-globin and Ha-ras genes (36). A potential NF-1 recognition element was present in the Ha-ras binding site. This similarity suggested that CTF and NF-1 might be structuraly and functionally related. Purified CTF contains a heterologous population of polypeptides of molecular weight similar to that of NF-1. The possible explanation for this is that the different polypeptides are generated by post-translational modification or are encoded by different but related genes, or they might be generated by proteolytic degradation of a larger CTF/NF-1 protein (36). A 160 KDa polypeptide has been identified in highly purified preparations of NF-1 (49). A 33-nucleotide synthetic oligomer corresponding to the expected coding sequence of CTF peptide 1 was selected as a specific DNA probe for screening a HeLa cDNA library (50). Structural analysis of ten independent cDNA clones suggested that a single gene in human cells gives rise to at least three distinct mRNA species which are most likely generated by alternative RNA splicing.

The regions of the CTF/NF-1 protein which are responsible for its DNA binding, dimerization, transcriptional and replication activities have been identified (51). CTF-1 does not contain any of the previously characterized motifs of various DNA binding proteins nor does it possess a leucine zipper dimerization motif. CTF/NF-1 proteins probably recognize their binding site by using

a different type of a DNA binding structure. A C- terminal 100 aminoacids portion of a protein is needed for transcriptional activation. It contains approximately 25% proline residues. It is possible that the CTF proline domain interacts directly with some component of the general transcription apparatus (e.g. TFIIA,-B,-D,-E,-F, subunits of RNA polymarase). An N-terminal portion of 185 aminoacids is required for DNA recognition, protein dimerization and adenovirus replication. These results suggest that interactions and mechanisms governing transcriptional activation by CTF are distinct from those mediating DNA replication, although CTF/NF-1 function both as a transcription selectivity factor and as an initiation factor for DNA replication (51).

Major late promoter (MLP) is a viral promoter and is required late in adenovirus infection (52). The promoter contains a TATA box at position -31 to -25, and a unique upstream sequence at positions -63 to -52 (53). Point mutation in the -57 to -50 region significally reduce transcriptional efficiency (54). A transcription factor from HeLa cells binds specifically to the region in the Ad MLP which is essential for efficient transcription of the MLP (54). Using a gel electrophoresis DNA binding assay, in which extracts from unefected HeLa cells were incubated with an end labeled DNA fragment containing MLP, factors that specifically bind a MLP upstream sequence have been identified. DNase I footprinting analysis revealed the proctected 17 bp sequence (5'GTAGGCCACGGACCGG3') located from -50 to -66 bp upstream of the transcriptonal initiation site (54). *In vivo* transcription from the Ad 2 MLP depended on this sequence and required the binding a MLP transcription factor (55).

The factor called MLTF has been purified from HeLa cells using DNA affinity chromatography (56). Proteins in the molecular weight range of 44 and 48 kDa, which were eluted and renatured from an SDS-polyacrylamide gel, exibited MLTF binding and transcrption-stimulatory activities. The mechanism by which MLTF activates transcription from the MLP is unclear. It is able to bind with a high affinity to its recognizing sequence, even in the absence of the TATA box binding factor (54). Since it possesses transcription and a DNA binding activity, it must have a DNA binding domain and a domain to contact another transcription factor (57).

The protein called upstream stimulatory transcription factor (USF) has been purified from HeLa cells (57). It is shown to be required, both *in vivo* and *in vitro* , for maximal expression of the MLP of adenovirus. It has been purified using a combination of classical purification techniques and fast-flow protein liquid chromatogrphy. Purified USF was found to be stable at the temperatures as high as 100 C. There are two forms of the USF protein, each having a slightly different mobility in SDS gel electrophoresis. Their molecular weights are respectively 44 and 43 kDa. These two forms of USF displayed identical affinities for the MLP upstream sequence (58).

Recently a novel protein has been isolated by screening a bacteriophage l expression library with a probe containing the binding site for MLTF (59). The cDNA-encoded portion is approximately 60 KDa. There is a presence of a newly established DNA binding and a protein dimerization motif called helix-loop-helix and there is a proline rich region at the carboxil terminus. Several other proteins belong to the helix-loop-helix family,

namely immunoglobulin k light chain gene enhancer binding proteins E 47 and E 12, proteins specific for muscle cell differentiation MyoD and myogenin, a protein that recognizes sequences in the heavy-chain gene enhancer TFE 3 (3) and protooncogenes c-, N- and L-myc (60,61). The E-box sequence in the heavy-chain immunoglobin enhancer is very similar to the binding site of MLTF (59). Also, proteins that bind to these sequences are closely related.

1.5 THE AIM OF THE WORK

As I mentioned before, human DNA sequences synthesized at the onset of S phase were isolated (24). The longest of these fragments (pB48) was found to exhibit a lot of interesting characteristics, such as possible regulatory signals, and it is believed to play a role in the regulation of transcription and initiation of DNA replication.

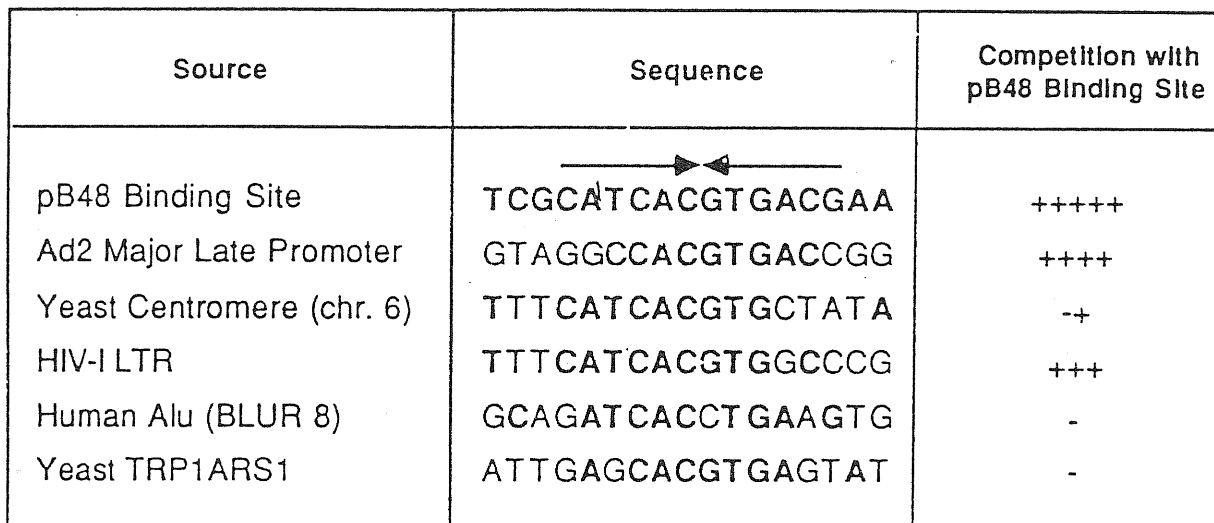
Using band shift and DNase footprinting assays, a specific 17 bp recognition sequence (5'TTCGTCACGTGATGCGA3') for one (or more) nuclear factors was identified (24). The binding sequence is very similar to an upstream element in the major late promoter of Ad 2. The factor MLTF/USF that binds to this sequence has been purified by different investigators (56,58). It is a 46 kDa protein and is able to activate transcription both *in vivo* and *in vitro*. (56,62) Since it is present in uninfected HeLa cell extracts it is probably involved in the transcription of cellular genes. It was found that it selectively stimulates transcription from the rat γ-fibrinogen promoter (63). A binding site for this factor also exists in the promoter of the mouse matallothionein I gene. Competition experiments showed that the MLP of Ad 2 competes with the pB48 binding site for the same factor. Lower, but still significant competition was found for the LTR of HIV-1 (25).

We wanted to purify a nuclear protein that specifically recognizes the pB48 binding site. We used HeLa cells to prepare a nuclear extract. Purification was followed by band shift gel electrophoresis and the active fractions were tested for the presence of the pB48BS protein using south-western transfer.

A southwestern experiment using crude nuclear extract revealed the presence of more than one protein that specifically binds our binding site, a 17 bp sequence homologous to the MLTF binding site. There is a band in the ~44 kDa position that might correspond to the already purified MLTF/USF protein. A second specifically recognized protein was about 70 kDa and third one was ~90 kDa. These factors are probably involved in the regulation of transcription, but a possibility that they also mediate DNA replication cannot be excluded. An example is CTF/NF-1, a factor that activates transcription from several eukaryotic promoters and is required for the initiation of adenovirus DNA replication (36).

Table2.

COMPETITION FOR A NUCLEAR FACTOR
OF SEQUENCES ANALOGOUS TO pB48 BINDING SITE



Source	Sequence	Competition with pB48 Binding Site
pB48 Binding Site	TCGCATCACGTGACGAA	+++++
Ad2 Major Late Promoter	GTAGGCCACGTGACCGG	++++
Yeast Centromere (chr. 6)	TTTCATCACGTGCTATA	-+
HIV-I LTR	TTTCATCACGTGGCCCG	+++
Human Alu (BLUR 8)	GCAGATCACCTGAAGTG	-
Yeast TRP1ARS1	ATTGAGCACGTGAGTAT	-

2. MATERIALS AND METHODS

Buffers: All the buffers and solutions were made using sterile water and were sterilized by filtration through 0.2 µm cellulose acetate filters (Nalgene).

Buffers for nuclear extract preparation: Buffer A contained 10 mM Hepes pH 7.9 at 4 ^{o}C, 1.5 mM $MgCl_2$, 10 mM KCl; buffer C contained 20 mM Hepes pH 7.9, 20% glycerol, 0.42 M NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA and 10% sucrose; Buffer D contained 20 Mm Hepes, 20% glycerol and 0.1 M NaCl. Protease inhibitors: 1 mM DTT, 1 mM PMSF and 1 mM sodium metabisulphite were added to all the buffers immediately before the use.

Buffers for BIO-REX 70: the buffer for column equilibration and washing contained 20 mM Hepes pH 7.9, 20% glycerol and 0.1 M NaCl. The composition of the elution buffers was the same except that they contained either 0.6 or 1.0 M NaCl. Protease inhibitors: 1 mM DTT, 0.5 mM PMSF, 1 mM sodium metabisulfite.

Buffers for HiLoad S Sepharose: Buffer A contained 20 mM Hepes pH 7.9, 20% glycerol, 12.5 mM $MgCl_2$, 0.2 mM EDTA, 0.1 M NaCl. Buffer B was the same except that it contained 1.0 M NaCl. Protease inhibitors: 1 mM DDT, 0.5 mM PMSF, 1 mM sodium metabisulfite.

The buffers were degassed before use.

Buffers for FPLC Mono Q: Buffer A contained 20 mM Tris pH 8, 20% glycerol, 12.5 mM $MgCl_2$, 0.2 mM EDTA, 0.1 M NaCl. Buffer B was identical to buffer A except that it contained 1.0 M NaCl.

Protease inhibitors: 1 mM DDT, 0.5 mM PMSF, 1 mM sodium metabisulfite. The buffers were degassed before use.

Buffers for specific DNA-affinity chromatography: for column equilibration and washing: 20 mM Hepes pH 7.9, 20% glycerol, 12.5 mM $MgCl_2$, 0.1% NP-40, 0.1 M NaCl; the composition of the elution buffers was the same except they contained different molarities of NaCl: 0.2, 0.4, 0.6, 1.0 and 2.0 M NaCl. Protease inhibitors used: 1 mM DTT, 0.1 mM PMSF, 1 mM sodium metabisulfite.

Nuclear extract preparation: HeLa cells were grown in spinner flasks at 37^oC in Dulbecco's Minimal Essential Medium with addition of gentamycin (100 mg/ml) and 10% fetal calf serum + glutamine. Nuclear extracts were prepared as described by Dignam et al. HeLA cells were harvested from cell culture media by centrifugation at room temperature for 30 min at 2000 rpm in a Sorvall HG4L rotor. Pelleted cells were then suspended in five volumes of 4^oC phosphate buffered saline and collected by centrifugation as above. Subsequent steps were performed at 4^oC. The cells were suspended in five packed cell pellet volumes of buffer A and allowed to stand for 10 min and then centrifuged as above. The pellet was suspended in two packed cell pellet volumes (prior to the initial wash with buffer A) of buffer A and lysed by 40 strokes of a Knotes all glass Dounce homogenizer (B type pestle). The homogenate was checked microscopically for cell lysis and centrifuged for 10 min at 2000 rpm in a Sorvall HG4L rotor to pellet nuclei. The pellet obtained from the low speed centrifugation of the homogenate was subjected to a second

centrifugation for 20 min at 15000 rpm in Sorvall SS34 rotor to remove residual cytoplasmic material and this pellet was designated as crude nuclei. These crude nuclei were resuspended in buffer C (3ml per 2 g of cells) with a Knotes all glass Dounce homogenizer (20 strokes, B type pestle) The resulting suspension was stirred at 4°C with a magnetic stirring bar for 30 min and then centrifuged for 30 min at 15000 rpm (Sorvall SS 34 rotor).

Ammonium sulfate precipitation: The resulting supernatant (after the final centrifugation) was transfered into a small bottle. 0-35% precipitation was carried out at 4°C by gradually adding small quantities of ammonium sulfate within 30 min. The solution was stirred gently. After the addition of ammonium sulfate the suspension was left in cold for another 30 min. The precipitated proteins were spun at 15000 rpm for 30 min in a Sorvall SS34 rotor. The pellet was dissolved in buffer D and dialysed.

Chromatography

BIO-REX 70 chromatography: BIO-REX 70 resin was equilibrated to 0.1 M NaCl and packed into the column. Nuclear extract (215 ml) was loaded on the column at a flow rate 13 ml/hr. The column was washed with buffer containing 0.1 M NaCl and the DNA binding activity was eluted with a linear gradient from 0.1-0.6 M NaCl, followed by a 1.0 M NaCl step elution. The fractions that contained activity were pooled and dialysed.

HiLoad S Sepharose chromatography: The diluted pool of activity (62 ml) was loaded onto a HiLoad S Sepharose column (Pharmacia) at a flow rate 130 ml /hr. The proteins were eluted

with a linear gradient from 0.1-1.0 M NaCl. The specific DNA binding activity was eluted between 0.25 and 0.37 M NaCl.

Chromatography on Mono Q: The dialysed pooled fractions were filtered through 0.45 mm syringe filter (Millipore) and loaded at 1 ml/min onto a 1ml Mono Q column (Pharmacia HR5/5) preequilibrated in buffer A containing o.1 M NaCl. The column was then washed with 2 ml of the same buffer and eluted at 0.5 ml/min with a 0.1-1.0 M NaCl gradient (40 ml).

Specific DNA-affinity chromatography: Coupling of the DNA to Sepharose: Resins were prepared according to Kadonaga and Tjian (65). CNBr-Sepharose (3 g) was suspended in ~50 ml of 1 mM HCl and left for a few minutes to swell and then transfered to a sintered-glass funnel and washed with 600 ml 1 mM HCl under slight vacuum. While in the funnel, it was washed with 300 ml of ice-cold water and 100 ml of ice-cold 10 mM potassium phosphate. The resin was transfered into a 15 ml srew-capped plastic tube and 4ml of 10 mM K-phosphate was added together with the ligated oligonucleotides. It was incubated overnight (16 hrs) at room temperature. The slurry was transfered to a sintered-glass funnel and washed with 200 ml of H_2O and 100 ml of 1 M Ethanolamine-HCl (pH 8). The first few drops were collected and checked on bound DNA, because the DNA is labelled with ^{32}P and the efficiency of DNA attachment to the Sepharose can be estimated by comparing the amount of radioactivity that is retained on the resin with the amount of radioactivity that remains in the solution after the coupling reaction. The resin was transfered to a tube and 4 ml of 1 M Ethanolamine-HCl was added. It was incubated on a rotating wheel at room temperature for 4 hrs. After that, the resin was washed in a sintered-glass funnel

successively with: 100 ml 10 mM K-phosphate (pH 8.0), 100 ml 1M K-phosphate (pH 8.0), 100 ml 1M KCl, 100 ml H_2O, 100 ml column storage buffer (0.3 M NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and 0.02% NaN_3). The resin was store at 4^oC in column storage buffer.

DNA affinity chromatography: after the dialysis the sample (active pool from FPLC Mono Q) was diluted with a buffer without salt and loaded on the column. The column was washed with a 2.0 M NaCl buffer and then equilibrated with a buffer containing 0.1 M NaCl. After loading, the column was washed with a buffer+ 0.1 M NaCl. Bound proteins were eluted with 0.2, 0.4, 0.6, 1.0 and 2.0 M NaCl step elutions All operations were performed at 4^oC.

Protein analysis: Protein concentrations were determined according to the method of Bradford using a Bio-Rad protein assay kit (66).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresisSDS-PAGE) was carried out essentially as described by Laemmli (67). 3 x sample buffer contained 189 µl Tris pH 6.8, 300 µl glycerol, 300 µl 20% SDS, 150 µl 2-mercaptoethanol and 50 µl bromophenolblue/X-cianol in1 ml of H_2O. The gel standards were myosin (200,000), β galactosidase (116,250), phosphorylase b (97,400), BSA (66,200), ovalbumin (42,699), carbonic anhydrase (31,000) and lysosime (14,400). Silver staining was performed using a kit from Bio-Rad.

Gel retardation assay-Band Shift: Proteins that bind DNA fragments retard the migration of these fragments in an electrophoretic gel. They form a complex that migrates slowly. DNA fragments are incubated with a mixture of proteins. If there

are protein(s) that bind(s) that fragment, it is possible to detect it by loading the reaction mixture on an acrylamide gel and observing the presence of a retarded band (band shift), with respect to the band of the same DNA fragment not treated with the protein mixture. A labeled fragment or probe is a target molecule to which protein(s) bind. Poly[d(I-C)] is a polynucleotide and therefore it is "DNA" in all effects. It is supposed to act as an unspecific competitor against the DNA fragments.and is usually added in excess with respect to the DNA fragments. If any protein binds to the probe DNA in the presence of this excess competitor, then the reaction between DNA and protein is specific.

Band Shift experiments were performed as described by Tribioli et al (24). 1 ng (10^4 cpm) of end-labelled probe (pUF4 [BglII-HindIII]) containing the binding site for the human nuclear factor was incubated with 2-4 µl of protein fractions and 1-3 µg of poly[d(I-C)] in 60 mM Hepes pH 7.9, 150 mM NaCl, 15 mM $MgCl_2$, 6 mM DTT, 0.6 mM EDTA and 15% glycerol.(final volume of the reaction mixture was 20 µl) After 30 min of incubation at room temperature the samples were loaded directly onto a 5% polyacrylamide gel in 0.5xTBE (25 mM Trizma base, 25 mM boric acid, 1 mM EDTA). After electrophoresis the gel was dried and analysed by autoradiography.

South-Western experiments were carried out essentially as described by Silva et al.,.with some modifications (68). The proteins were separated by SDS gel electophoresis. The gel was incubated for 1 hour in renaturation buffer (10 mM Tris pH 7.2, 50 mM NaCl, 20 mM EDTA, 4 M urea) to remove the SDS. The proteins were transfered to a nitrocellulose (0.2 mm thick) filter

by electroblotting in Tris-glycine buffer (25 mM Tris, 190 mM glycine) without methanol. The nitrocellulose was then incubated for 1 hour in binding buffer (10 mM Tris pH 7.2, 50 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT) containing 5% powdered lean milk. The filter was treated for 1 hour in binding buffer with the reaction mixture: ^{32}P end labelled oligonucleotide,~10^6 cpm/ml and 50 µg/ml of single-stranded and 50 µg/ml of double-stranded sonicated salmon sperm DNA for unspecific binding. The filter was washed 3 times within 30 min with binding buffer, dried briefly, covered with a thin foil and exposed with an intensifying screen for autoradiography.

DNA fragments and oligonucleotides:

pUF4 is a plasmid that contais a tandem repeat of 4 binding site sequences of pB48 cloned in the polylinker region of the vector pUC18. pUF4 was linearized with BglII, ^{32}P labeled with a Klenow fragment of E.coli polymerase I and cut with HindIII. The labeled fragment was purified on a 5% plyacrylamide gel and used as a probe in band-shift assays.

Oligonucleotides were synthesyzed using an Applied Biosystems 380B DNA synthesizer.

"Super" probe for south-western experiments: an oligonucleotide containing four copies of pB48 binding site.(64 bases).was synthesized together with its complementary strand. Complementary oligostrands were annealed, ligated and nick translated.

Oligonucleotides for coupling to CNBr-activated Sepharose:
Synthesized complementary oligostrands:

5'GATCTCGCATCACGTGACGAA3'
3'AGCGTAGTGCACTGCTTCTAG5'

were annealed, 5'-phosphorylated and ligated as described by Kadonaga and Tjian (65). The resulting DNA was analysed by agarose gel gelelectrophoresis.

3. RESULTS

The purification of the protein(s) has been done following the scheme in Fig.3. The activity of our protein was monitored by a band shift assay. As a specific DNA probe we used a fragment containing 4 binding sites of pB48. A representative assay of the active fractions after each chromatographic step is shown in Fig.4. Details of the entire purification are given in Table 3.

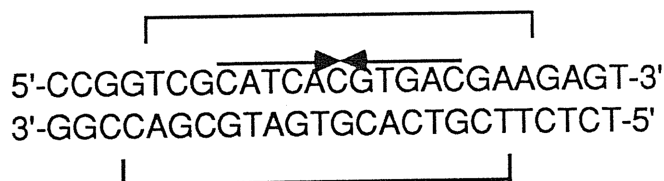


Figure2. Sequence of the pB48 binding site.

3.1 CHROMATOGRAPHIC STEPS

Nuclear extract was prepared according to Dignam(64). We started from 340g of HeLa cells which corresponded to 550 ml of nuclear pellet extract. This relatively large volume of starting material was necessary because of the low abundance of the desired protein. The second step of purification, ammonium sulfate precipitation was designed to decrease the total protein and the volume to be handled at the first chromatographic step.

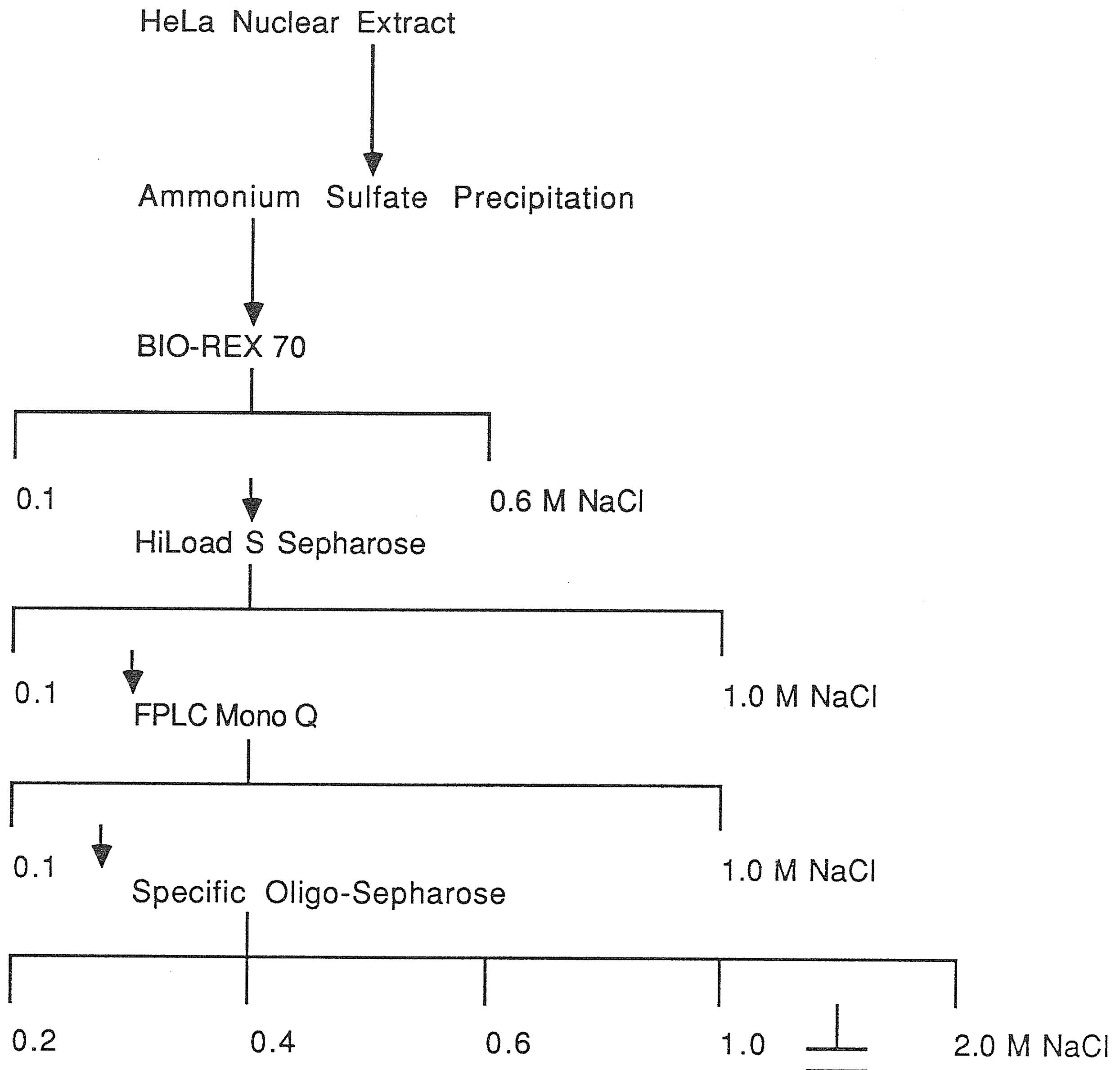


Figure 3. Purification scheme. Extract preparation was performed essentially as described by Dignam et al. Each column was loaded in 0.1 M NaCl. The binding activity from the specific DNA column was eluted with step elutions.

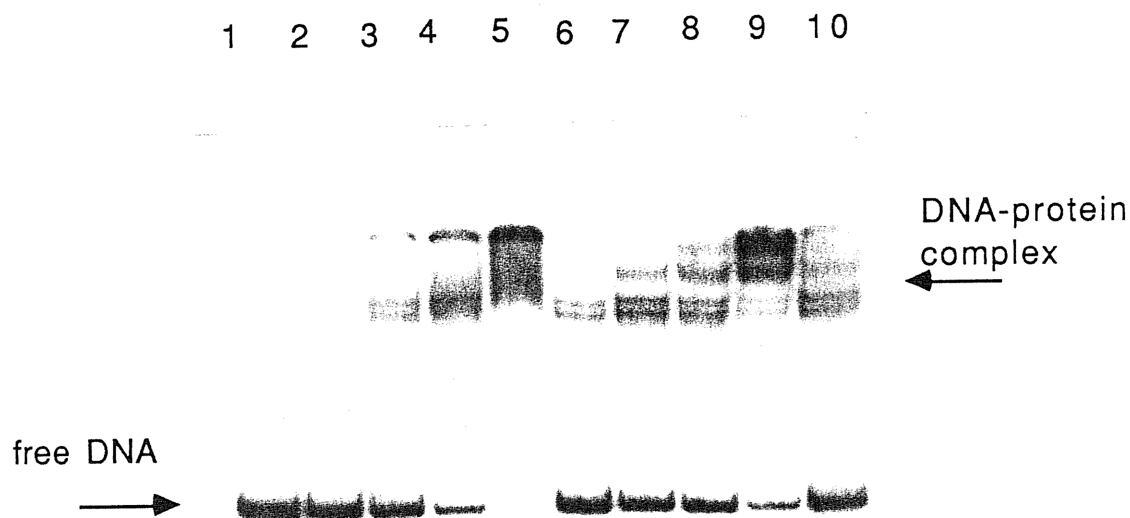


Figure 4. Band shift assay. 1 ng of end-labeled probe (pUF[BglII-HindIII]) and 3 μg of competitor poly[d(I-C)] was mixed with aliquots of protein fractions in a 20 ml volume. After 30 min of incubation at room temperature the mixture was electrophoresed on 5% polyacrylamide gel in a low-ionic strength buffer in the cold. Lane 1. fragment alone. Lanes 2. & 3. nuclear extract 1 & 2 μl. Lanes 4. & 5. BIO-REX active pool 1 & 2 μ l. Lanes 6. & 7. HiLoad S Sepharose active pool 1 & 2 μl. Lanes 8. & 9. FPLC Mono Q active pool I 1 & 2 μl. Lane 10. FPLC Mono Q active pool III 2 μl.

Table 3

Fraction	Volume ml	Total protein mg	Purification fold
I. HeLa Nuclear Extract	550	1850	
II. Ammonium Sulfate Ppt	215	1440.5	1.3
III. BIO-REX 70	570	208	8.9
Amm Sulf Ppt	31	176.64	10.47
IV. HiLoad S Sepharose	143	37.13	49.8
V. FPLC Mono Q	12.7	5.48	337.5
VI. Specific oligo-Sepharose	2.2	<0.02	not determined

We performed 0-35% ammonium sulfate pecipitation. The extract after the ammonium sulfate cut (215 ml) was loaded onto the high capacity, weak cation exchange resin BIO-REX 70. The specific binding activity was eluted by a continuous salt gradient (0.1-0.6 M NaCl), followed by 1 M NaCl step elution. This step removed 90% of the loaded proteins and significantly reduced the matrix required for subsequent purification steps. The specific DNA-binding activity eluted between 0.25 and 0.35 M NaCl. No activity was detected in the loading and washing fractions. The BIO-REX 70 active gradient pools (BRAP) were precipitated with ammonium sulfate (40%), dissolved in buffer without salt, dialysed and loaded onto the HiLoad S Sepharose High Performance column. HiLoad S Sepharose is strong cation exchanger based on a rigid highly cross-linked beaded agarose. The functional group on the matrix is the sulfonate group. We used this column as a preparative purification step since it has a high loading capacity. The proteins were eluted from the column with a gradient from 0.1 to 1.0 M NaCl. Active fractions were eluted in two regions, at approximately 0.19-0.28 M NaCl (1st gradient pool) and 0.3-0.38 M NaCl (2nd gradient pool). Active fractions of the 2nd gradient pool were dialysed and applied to a strong anion exchanger fast flow protein liquid chromatography (FPLC) column. The functional group on the matrix for Mono Q is the quaternary amino group. This column provides rapid purification and high resolution. The proteins were eluted with a gradient from 0.1 to 1.0 M NaCl. The active fractions in which salt concentration varied approximately from 0.16-0.26 M NaCl were divided into three pools in order to determine where the different

specific proteins were eluted. The first pool containing 0.16- 0.2 M NaCl fractions was dialysed and applied to a sequence specific DNA-affinity column. The DNA affinity chromatography exploits the specific chromatographic behaviour of the binding factor on a DNA resin containing the specific binding site. The matrix for specific column was prepared by coupling a synthetic oligonucleotide to CNBr-activated Sepharose. The binding site had been previously determined by DNase I footprinting analysis (24) The column was step eluted with 0.2, 0.6, 1.0 and 2.0 M NaCl. The specific proteins were eluted at a high salt concentration..

3.2 CHARACTERIZATION OF PURIFIED FRACTIONS

Band shift assays: All the fractions during the purification were checked for the specific DNA-binding activity using band shift assay. Poly [d(I-C)] (1-3 μl) was used in the reaction mixture as a nonspecific inhibitor. We observed different types of band retardation. One DNA-protein complex, designated as a "normal" band shift, probably corresponded to the 40 kDa proteins. "Abnormal" retarded bands are possibly due to the other proteins.

South-Western experiments: After each purification step the active pools were tested for specific DNA-binding activity by south-western hybridization. The radioactive probe contained an oligonucleotide with four binding sites, annealed to its complementary strand and ligated. Proteins were transferred to the nitrocellulose filter and treated with the specific probe. Salmon sperm DNA was used as a competitor. The south-western hybridization using crude nuclear extract revealed the presence of

more proteins that specifically bind to the probe.The band observed around 70 kDa was of particular interrest for us. It probably corresponds to the novel protein that specifically binds to the sequence within pB48. (Figure 5.)

Following the purification steps, we observed enrichment of the proteins at around 40 kDa and 70 kDa. We tried to purify the 70 kDa protein and after FPLC Mono Q we succeeded in separating it from the 40 kDa proteins. (Figure 6.)

Further purification of the active fraction containing mainly the 40 kDa proteins revealed the presence of two bands possibly corresponding to already purified MLTF/USF. (Figure 7.), but unfortunately, it was not possible to detect it on the SDS gel. This means that the final protein amount was extremely low, but still active, because it bound specifically to the probe in the south western experiment.

SDS-PAGE: Fractions at various steps of purification were analyzed for their polypeptide content by electrophoresis on SDS-gels. After silver staining we found a large number of proteins in the first fractions. In the final preparation, after the specific chromatography, it was still possible to see a few protein bands. This means that the preparation was not homogenuous. Further purification steps are necessary to obtain the pure fraction of the protein.

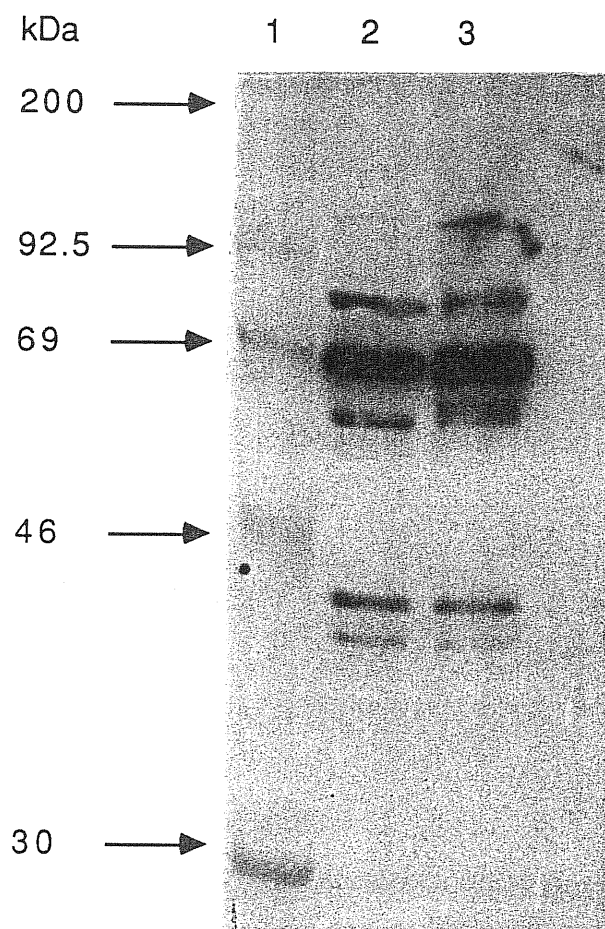


Figure 5. South-Western hybridization. The proteins were separated on a 12% gel, transferred to a nitrocellulose filter and incubated with a specific probe.(see Materials and Methods).
Lane 1, radioactive rainbow marker.
Lanes 2 and 3 nuclear extract.

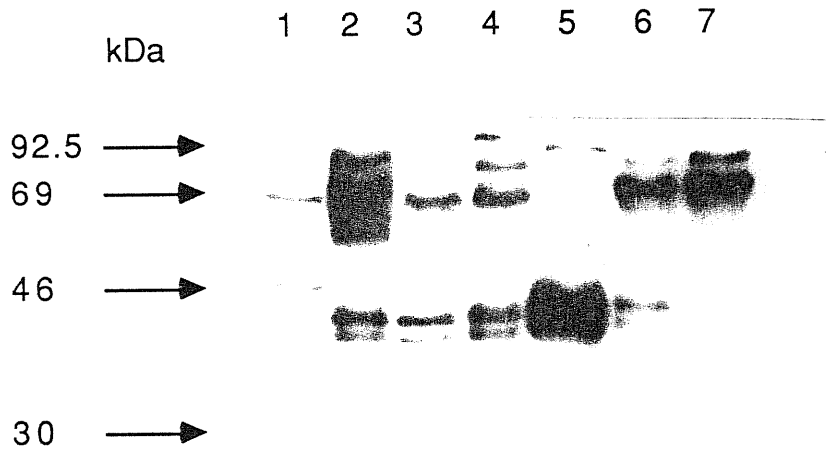


Figure 6. South-Western hybridization. Lane 1, radioactive rainbow marker. Lane 2, nuclear extract. Lanes 3&4, BIO-REX 70 active pools. Lane 5, Mono Q gradient pool 1. Lane 6, Mono Q gradient pool 2. Lane 7, Mono Q gradient pool 3.

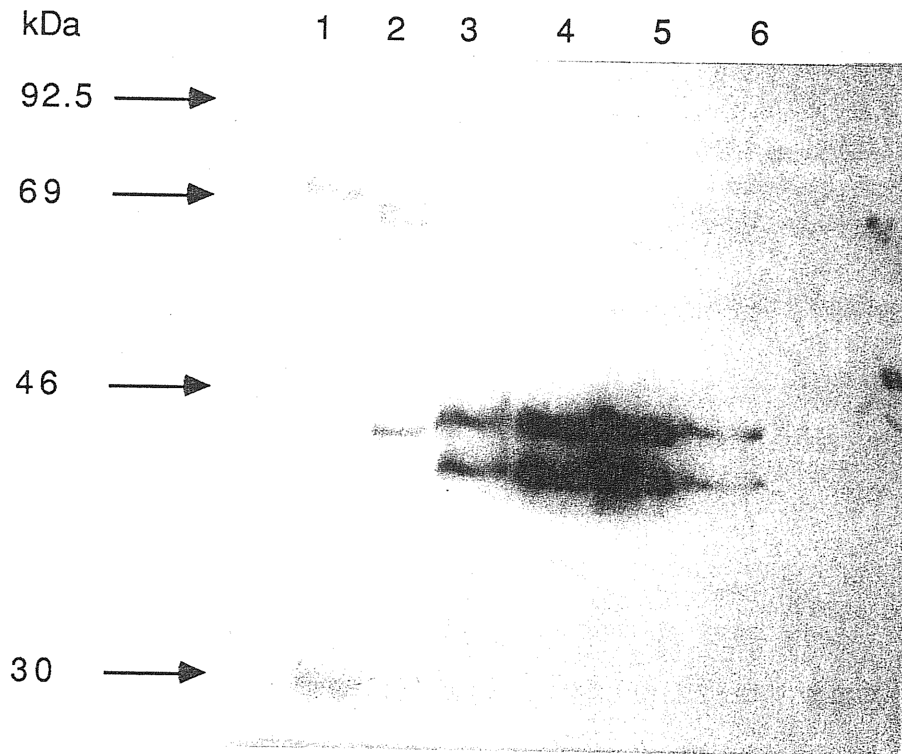


Figure 7. South-Western hybridization. Lane 1, radioactive rainbow marker. Lane 2, nuclear extract. Lane 3, Mono Q gradient pool 1. Lanes 4,5,6, specific oligo-Sepharose active pools.

4. DISCUSSION

4.1 MORE THEN ONE FACTOR BINDS TO THE pB48 SEQUENCE HOMOLOGOUS TO THE MLTF BINDING SITE

Sequence-specific DNA binding proteins are involved in regulation of fundamental cellular processes such as gene expression and DNA replication (38). In prokaryotes and viruses the identification and characterization of such regulatory proteins has been much easier then in eukaryotes because of the complexity of eukaryotic systems.

Recently, many nuclear proteins have been found to interact with short conserved sequences, but molecular mechanisms involved in the interaction of nuclear factors with DNA are not yet fully understood. Most of the specific DNA-binding proteins are present in very low amounts and therefore it is difficult to isolate them in sufficient quantities to determine their complete aminoacid sequence.

The identification and purification of the proteins that specifically recognize such sequences iis an important step towards understanding genetic regulatory mechanisms in eukaryotes.

We tried to purify proteins that specifically bind to a DNA fragment homologous to the upstream element of the major late promoter of adenovirus. This sequence has been identified by band retardation and footprinting experiments with a DNA fragment obtained from newly synthesized DNA in HL60 cells (24). Data obtained from south western experiments showed that there was more then one protein that bound specifically to the

probe. In the trial to purify these factors we used conventional chromatography, fast flow protein liquid chromatography (FPLC) and DNA recognition site affinity chromatography. Initial chromatography steps were needed to remove the majority of proteins and to concentrate the fractions containing our factor of interest. During the purification we succeeded in following both the 40 kDa and the 70 kDa proteins observed by south-western assay. Moreover, by using FPLC Mono Q column we obtained separation of the two factors since they eluted at different salt concentration. One of the problems we faced was the follow up of the 70 kDa protein, because it did not appear to give a retarded band in our experimental conditions. As a final step we used sequence-specific DNA-affinity chromatography. This method is useful for the isolation of nuclear regulatory proteins that are present in low amounts. An affinity matrix contained the DNA recognition site which is highly specific for our nuclear factor. We found that even if the matrix is maximized for specific DNA sequences, fractionation by DNA recognition site affinity chromatography may still result in only a partially purified preparation.

The other critical point was the starting material. It is necessary to use high quantities of cells (>500g), because the factors are not abundant. After the final purification step the quantity of the sequence specific protein obtained was not enough for microsequencing. Therefore in order to have enough protein for biochemical studies we should start with a higher amount of cells. We suppose that our factor has a role in transcription initiation of an early replicated chromosomal region, but a possible role in DNA replication cannot be excluded. Further

investigation will hopefully provide new data about the structural and biochemical properties of our factor and about its role in the mechanisms of transcription and replication initiation.

4.2 IS THERE A NEW FAMILY OF RELATED REGULATORY PROTEINS FOR MLTF BINDING SITE ?

There are a lot of regulatory proteins that bind to specific short DNA recognition sequences. Some of them recognize similar or identical DNA sequences. For instance, the factors NFI, CTF, CBF belong to a family of proteins which recognize the CCAAT core sequence (69) Several proteins bind to the octamer ATTTGCAT (70).

MLTF/USF is a transcription factor which binds to and stimulates transcription from the major late promoter (MLP) of adenovirus (54). MLTF/USF has been purified and identified as a 46 kDa polypeptide (56,58). The MLTF/USF binding site (5'GTAGGCCACGTGACCGGG3') is highly conserved among the adenoviruses. It is possible that MLTF/USF binding site represents a conserved sequence for a number of eukaryotic factors. A new factor TFEB (~60kDa) has been isolated by screening a bacteriophage library with a probe containing the binding site for MLTF/USF (59). The binding site for MLTF/USF is very similar to the E-box in the heavy chain immunoglobin enhancer. TFEB was found to be closely related to the TFE 3 protein which binds specifically to the E-box sequence. These data suggest that there is a family of factors with related binding specificities. The binding site for our protein is located in an early replicated sequence synthesized at the onset of the S phase in human cells and is

highly homologous to the MLTF/USF binding site. Thus, our factor might represent a new member of this family.

5. BIBLIOGRAPHY

1. Kornberg, A. (1980) DNA Replication, W.H. Freeman & Co., San Francisco

2. Huberman, J.A. (1987) Cell 48, 7-8

3. Harland, R. (1981) TIBS 6, 71-74

4. Jackson, D.A. (1990) BioEssays 12, 87-89

5. Blumenthal, A.B., Kriegstein, H.J. & Hogness, D.S. (1974) Cold Spring Harbour Symp. Quant. Biol. 38, 205-284

6. Edenberg, H.J. & Huberman, J.A. (1975) Annu. Rev. Genet. 9, 245-325

7. Hand, R. (1978) Cell 15, 317-325

8. Harland, R.M. & Laskey, R.A. (1980) Cell 2, 761-771

9. Blow, J.J. & Laskey, R.A. (1986) Cell 47, 577-587

10. Brown, E.H., Iqbal, M.A., Stuart, S., Hatton, K.S., Valinsky, j. & Schildkraut, C.T. (1987) Mol.Cell.Biol. 7, 450-457

11. Burhans, W.C., Selegue, J.E. & Heintz, N.H. (1986) Proc.Natl.Acad.Sci.USA 83, 7790-7794

12. Vassilev, L.T., Burhans, W.C.& DePamphilis, M.L. (1990) Mol.Cell.Biol. 10, 4685-4689

13. Caddle, M.S., Lussier, R.H. & Heintz, N.H. (1990) J.Mol.Biol. 211, 19-33

14. Stuhl, K., Stinchcomb, D.T., Scherer, S. & Davis, R.W. (1979) Proc.Natl.Acad.Sci.USA 76, 1035-1039

15. Hieter, P., Mann, C., Snyder, M. & Davis, R.W. (1985) Cell 40, 381-392

16. Koshland, D., Kent, J.C. & Hartwell, L.H. (1985) Cell 40, 393-403

17. Biamonti, g., Della Valle, g., Talarico, D., Cobianchi, F., Riva, S. & Falaschi, A. (1985) NucleicAcidsRes. 13, 5545-5561

18. Kaufmann,G., Zannis-Hadjopoulos, M. & Martin, J.G. (1985) Mol.Cell.Biol. 5, 721-727

19. Zannis-Hadjopoulos, M., Kaufmann, G., Wang, S., Lechner, R.L., Karawya, E., Hesse, J. & Martin, R.G. (1985) Mol.Cell.Biol. 5, 1621-1629

20. Vassilev, L.T. & Johnson, E.M. (1990) Mol.Cell.Biol. 10, 4899-4904

21. Iguchi-Ariga, S.M., Okazaki, M.T., Itani, T., Ogata, m., Sato, y. & Ariga, H. (1988) EMBO J. 7, 3135-3142

22. McWhinney, C. & Leffak, M. (1990) NucleicAcidsRes. 18, 1233-1242

23. Kaufmann, G., Zannis-Hadjopolous, M. & Martin, R.G. (1985) Mol.Cell.Biol. 5, 721-727

24. Tribioli,C., Biamonti, G., Giacca, M., Collona,M., Riva,S. & Falaschi,A. (1987) Nucleic Acids Res. 15, 10211-10232

25. Falaschi, A., Biamonti, G., Cobianchi, F., Csordas Toth, E., Faulkner, G., Giacca, M., Pedacchia, D., Perini, G., Riva, S. & Tribioli, C. (1988) Biochim.Biophys. Acta 951, 430-442

26. DePamphilis, M.L. & Wassarman, P.M. (1980) Annu.Rev.Biochem. 49, 627-666

27.Umek, R.M., Linskens, M.H.K., Kowalsky, D. & Huberman, J.A. (1989) Biochim.Biophys.Acta 1007, 1-14

28. Linskens, M.H.K. & Huberman, J.A. (1990) Cell 62, 845-847

29. Burhans, W.C., Vassilev, L.T., Caddle, M.S., Heintz, N.H. & DePamphilis, M.L. Cell (1990) 62, 955-965

30. Vaughn, J.P., Dijkwel, P.A. & Hamlin, J.L. (1990) Cell 61, 1075-1087

31. Hatton, K.S., Dhar, V., Brown, E.H., Iqbal, M.A., Stuart, S., Didamo, V.T. & Schildkraut, C.L. Mol.Cell.Biol. (1988) 8, 2149-2158

32. DePamphilis, M. Cell (1988) 52, 635-638

33. Brandt, A.H., Micklem, G. & Nasmyth, K. (1987) Cell 51, 709-719

34. Polyanovsky, O.L. & Stepchenko, A.G. (1990) BioEssays 12, 205-210

35. Prujin, G.J.M., Van Driel, W. & van der Vilet, P.C. Nature 322, 656-659

36. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. & Tjian, R. (1987) Cell 48, 79-89

37. Wingender, E. (1988) Nucleic Acids Res. 16, 1879-1902

38. Pabo, C.O. & Sauer, R.T. (1984) Annu.Rev.Biochem. 53, 293-321

39. Scheidereit, C., Heguy, A> & Roeder, R.G. (1987) Cell 51, 783-793

40. Ko, H.S., Fast, P., McBride, W. & Staudt, L. (1988) Cell 55, 135-144

41. Mueler, M.M., Ruppert, S., Schafner, W.& Matthias, P. (1988) 366, 544-551

42. Miller, J., McLachlan, A.D. 7 Klug, A. (1985) EMBO J. 4, 1609-1614

43.Rhodes, D. & Klug, A. (1986) Cell 46, 123-132

44. Tsai, S.Y., Sagami, I., Wang,H., Tsai, M.J. & O'Malley, B.W. (1987) Cell 50, 701-709

45. Nagata, K., Guggenheimer, R.A., Enomoto, T., Lichy, J.H. & Hurwitz, J. (1982) Proc.Natl.Acad.Sci.USA 79, 6438-6442

46. Siebenlist, V., Henninghausen, L., Battey, J. & Leder, P. (1984) Cell 44, 381-391

47. Gronostajski, R.M., Adhya, S., Nagata, K., Guggenheimer, R.A. & Hurwitz, J. (1985) Mol.Cell.Biol. 5, 964-971

48. Rosenfeld, P.J. & Kelly, T.J. (1986) J.Biol.Chem. 261, 1398-1408

49. Diffley, J.F. & Stillman, B. (1986) Mol.Cell.Biol. 6, 1363-1373

50. Santoro, C., Mermod, N., Andrews, P.C.& Tjian,R. (1988) Nature 334, 218-224

51. Mermod, N., O'Neill, E.A., Kelly, T.J. & Tjian, R. (1989) Cell 58, 741-753

52. Logan, J. & Schenk, T. (1986) Nucleic Acids Res. 14, 6329-6335

53. Hu, S.L. & Mannley, J.L. (1981) Proc.Natl.Acad.Sci.USA 78, 820-824

54. Carthew, R.W., Chodosh, L.A. & Sharp, P.A. (1985) Cell 43, 439-448

55. Hen, R., Sassone-Corsi,P., Corden, J., Gaub, M.P. & Chambon, P. (1982) Proc.Natl.Acad.Sci.USA 79, 7132-7136

56. Chodosh, L.A., Carthew, R.W. % Sharp,P.A. (1986) Mol.Cell.Biol. 6, 4723-4733

57. Bent, R. & Ptashne, M. (1985) Cell 43, 729-736

58. Sawadogo, M., Van Dyke, M.V., Gregor, P.D. & Roeder, R.G. (1988) J.Biol.Chem. 263, 11985-11993

59. Carr, C.S. & Sharp, P.A. (1990) Mol.Cell.Biol. 10, 4384-4388

60. Murre, C., McCaw, P.S. & Baltimore, D. (1989) Cell 56, 777-783

61. Beckmann, H., Su,L.K. & Kadesh, T. (1990) Genes.Dev. 4, 167-179

62. Sawadogo, M. (1988) J.Biol.Chem. 263, 11994-12001

63. Chodosh, L.A., Carthew, R.W., Morgan, J.G., Crabtree, G.R. & Sharp, P.A. (1987) Science 238, 684-688

64. Digman, J.D., Lebovitz, R.M. & Roeder, R.G. (1983) Nucleic Acids Res. 11, 1475-1491

65. Kadonaga, J.T. & Tjian, R. (1986) Proc.Natl.Acad.Sci.USA 83, 5889-5893

66. Bradford,M.M. (1976) Anal.Biochem. 72, 248-254

67. Laemmli, U.K. (1970) Nature 227, 680-685

68. Silva, C.M., Tully,D.B., Petch, L.A., Jewell, C.M. & Cidlowski, J.A. (1987) Proc.Natl.Acad.Sci.USA 84, 1744-1748

69. Dorn, A., Bollekens, J., Staub, A., Benoist, C. & Mathis, D. (1987) Cell 50, 863-872

70. Fletcher, C., Heintz, N. & Roeder, R.G. (1987) Cell 52, 773-781