Vol. 179, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 30, 1991 Pages 512-519

The presence of sequence-specific protein binding sites correlate with replication activity and matrix binding in a 1.7 Kb-long DNA fragment of the chicken α-globin gene domain

C. V. de Moura Gallo§, Y.S.Vassetzky§*, F. Recillas Targa§, G.P.Georgiev*, K. Scherrer§+ and S. V. Razin§*

§ Institut JACQUES MONOD, Université PARIS 7, 2 Place Jussieu 75251 PARIS Cedex 05, FRANCE

*Institute of Biology of the Gene, Academy of Sciences of the USSR, 34/5 Vavilov St, 117984 MOSCOW, USSR

Received July 22, 1991

Several recognition sites for novel sequence-specific DNA-binding proteins were found at the 5'-side of the chicken α-globin gene domain in a 1.7 Kbp DNA fragment. This fragment includes the replication origin, a non tissue-specific transcriptional enhancer, a DNAse I hypersensitive site and a permanent site of DNA attachment to the nuclear matrix. Most of the identified protein binding sites differ from previously known consensus sequences. Two sites coincide with MARs located at the 5'-end of the 1.7 kbp fragment. The proteins interacting with these two recognition sites were observed only in proliferating cells and were virtually absent in the extracts obtained from the non-replicating differentiated form of the same cells. One of them seems to belong to the GATA protein family, but its presence in nuclear extracts correlates with cell proliferation rather than expression of the domain.

The avian α-globin gene domain is one of the best known genomic domains in higher eukaryotes (Figure 1A, line a) . It is framed at a distance of 20-25 Kbp by a characteristic pattern of AT-rich and repetitive DNA segments (1,2) that include the start and the termination sites of a transcript running through the entire domain (3,4). The positions of genes and their individual transcripts have been mapped by several laboratories (3,5,6), and the existence of a transcription enhancer placed at the 3'-terminus discovered in the duck (7) was confirmed in the chicken (8). Previously, a 1.7 Kbp-long DNA fragment located about 4 Kbp upstream to the chicken α^{π}-globin gene was found to contain several functionally significant elements: a permanent site of DNA

+To whom correspondence should be addressed.

0006-291X/91 $1.50

interaction with the nuclear matrix (9-11), a replication origin (12,13), a constitutive site of DNase I hypersensitivity (14) and a transcriptional enhancer-like element that is not tissue-specific (15) (Figure 1A, line b). This suggested that specific DNA-protein interactions might be involved in some or all of the functions of the above mentioned elements. The whole area has been recently sequenced, and computer analysis demonstrated the presence of binding sites for some known transcription factors (16,17). However, only direct experiments could indicate whether sequence-specific interaction between proteins and DNA existed in this region. The present paper presents the results of mapping the DNA-protein interactions *in vitro* in the above 1.7 kb DNA fragment.

METHODS

Cells. The HD3 cells (18) were grown as described. The differentiation induction was performed by maintaining cells at 42°C four days, in presence of 20 μM 1-(5-Isoquinolinesulfonyl)-2-methylpiperazine. Normal differentiated erythroblasts were isolated from the blood of highly anaemic birds (19).
Preparation of protein extracts was carried out essentially as described (20). The concentration of proteins in the extracts was determined according to the procedure of Bradford and, if necessary, the solution was diluted with the dialysis buffer up to a final protein concentration of 5 mg/ml. The extracts were stored in aliquots at - 70°C.
DNA retardation experiments were carried out as described (21). In analytical experiments ~ 5 ng of end-labelled DNA, 5-10 μg of poly dI/dC (as a non-specific competitor), 10 μl of 2 x retardation buffer (40 mM Hepes (pH 7.6), 8% Ficoll, 10 mM $MgCl_2$, 80 mM NaCl, 0.2 mM EDTA) and the appropriate amount of bidistilled water were mixed in a final volume of 20 μl. Then 1 μl of extract (5 μg of protein) was added and the mixture was incubated for 45 min. at 0°C. The material was then loaded on a 6% acrylamide gel prepared in 0.2 x TBE buffer. The samples were run at 4°C.
Methylation interference and DNase I footprinting were carried out as described in (20).
Isolation of recombinant plasmids DNA, digestion of DNA with restriction enzymes and end-labelling reactions were carried out as described (22).

RESULTS AND DISCUSSION

The region was subdivided into a number of short DNA fragments, and their capacity to interact with sequence specific DNA-binding proteins, extracted from the nuclei of different chicken cells was studied by electrophoretic mobility shift experiments (21) and then by methylation interference (23) and DNAse I footprinting techniques (20). First gel retardation experiments demonstrated that all the subfragments of the 1.7 Kb a-globin DNA fragment interacted with some proteins from the nuclear extracts of chicken and duck erythroblasts (data not shown). Subsequent footprinting allowed to identify ten novel recognition sites for sequence-specific DNA-binding proteins that interact with this fragment . They are summarized in Figures 1 and 2. Three of

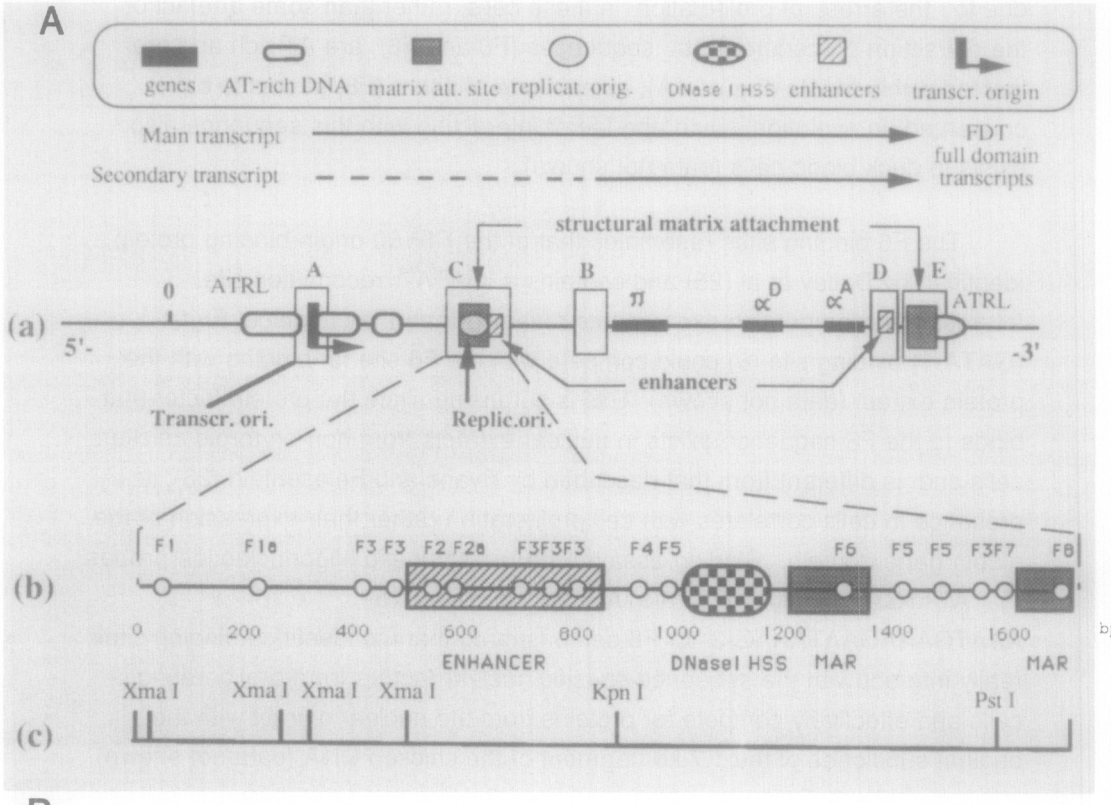
them, F2, F2a, and F3 are located in the area of the enhancer-like element identified by us earlier (15). One of the elements, F6, contains an element resembling the GATA-1 consensus binding site (Eryf 1, 8). Two of the elements, F6 and F8 are located within matrix attachment regions (MARs) recently identified in this area (24).

Since all the described protein factors interacted with the DNA fragment containing the origin of DNA replication and a structural site of DNA attachment to the nuclear matrix (9,10,11), it would be interesting to identify the factors that may be involved in cell proliferation, i.e. DNA replication and synthesis of structural proteins. To approach the problem we have compared the presence of all the above identified factors in the nuclear extracts from proliferating and non-proliferating chicken erythroid cells using the HD3 chicken erythroblast cell line transformed by a temperature sensitive avian erythroblastosis virus (tsAEV). These cells transcribe the domain of α-globin genes (4, Recillas Targa et al, unpublished observations) though do not produce hemoglobin at permissive temperature (36°); at 42°C they stop to proliferate, differentiate into essentially mature erythrocytes and start to express the globin genes by mRNA processing and translation (18).

The protein extracts were prepared in the same manner for both differentiated and non-differentiated HD3 cells. Then the subfragments of the studied area including each of the identified recognition sites were probed with both extracts. After induction of differentiation, the concentration of two of the factors drastically decrease. The retarded bands caused by factors F6 and F8 could be observed only when the extracts from proliferating cells were used as a source of DNA-binding proteins, while all the other factors were represented in approximately the same amounts in both extracts (Figure 3). The presence of most factors in both extracts serves as an internal control of the quality of the extract preparation in both types of cells. This suggests that

---

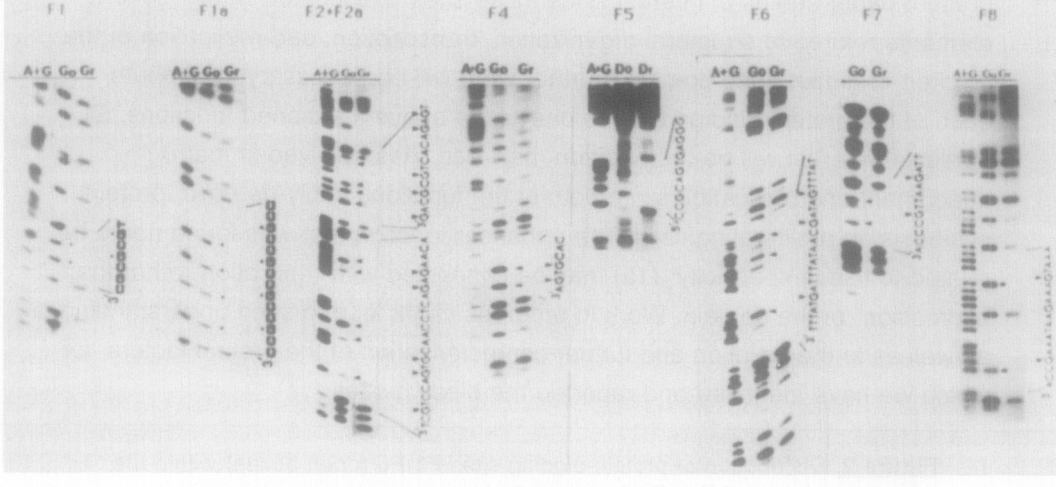
Figure 1. **(A)** General scheme of the chicken α-globin gene domain showing the position of the studied DNA fragment, the pattern of subclones constructed in the present investigation and the position of recognition sites F1-F8 within the fragment studied. In the upper part of the figure (line a), the known features of the chicken α-globin gene domain are represented. Open boxes show: (A, B) origins of full domain transcripts (3,4); (C) origin of replication (12,13), matrix attachment site (9,11) and 5' side enhancer-like element (15); (D) 3'-side enhancer (7,8); (E) termination of transcription (4) and 3'- side matrix attachment site (9,11,17). (Line b) the open circles on the map indicate positions of the protein-binding sites identified in the present investigation. Large boxes show positions of identified functional and structural elements within the 1.7 Kbp fragment. (Line c), restriction map of the fragment. **(B)** the identified recognition sequences F1-F8. **(C)** the results of methylation interference footprinting experiments. (A+G) the purine pattern; (G_O) control G pattern; (G_r) G pattern of protein-associated DNA fragments.



B

F1	GCGCGC	F4	CCTCACGTGCCC
F1a	GCGGCGGC	F5	GAGGG
F2	TCACGTTGGAGTCTCTTGT	F6	TATTTGATAGCAATATTTAGTATT
F2a	CTGCGCCAGTGTCTCA	F7	TTAGAATTG
F3	TGCPyGGG	F8	GAAAGAATAAAG

C



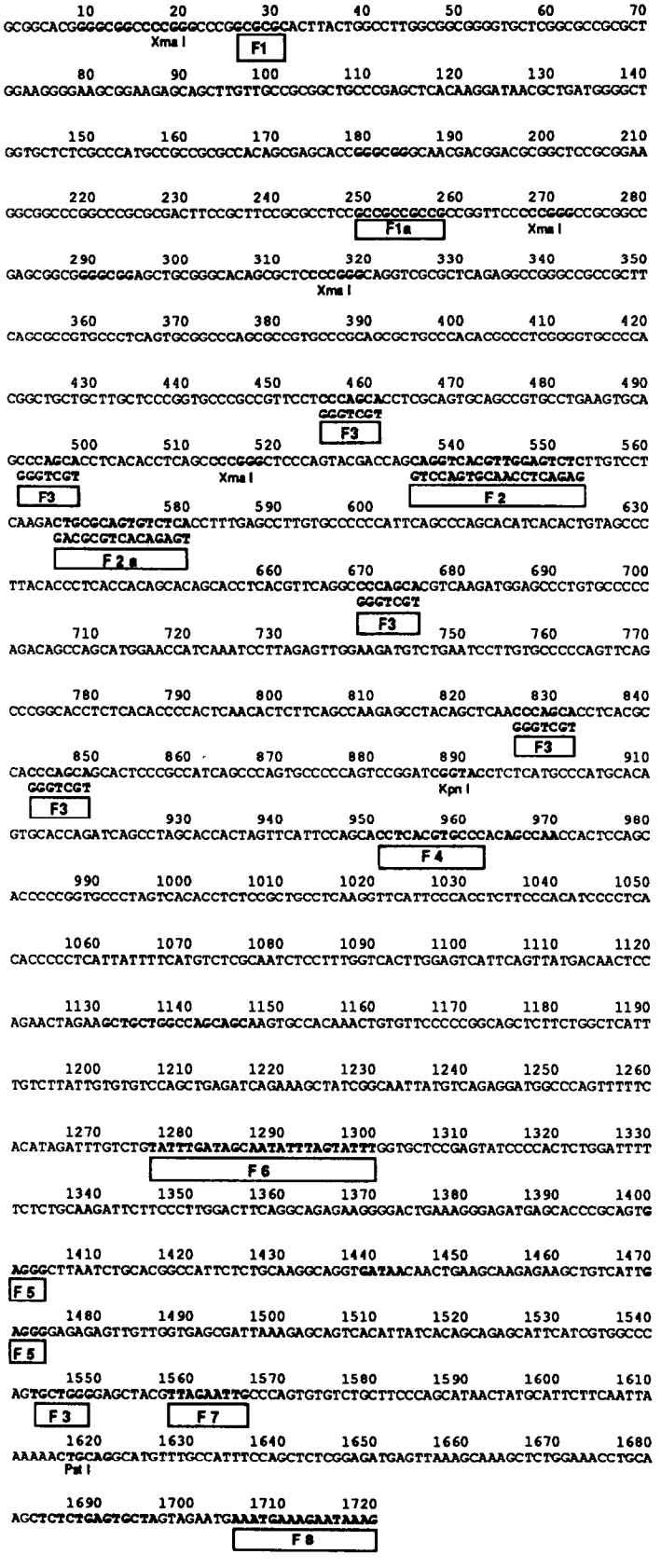
the absence of factors F6 and F8 in the extract from non-proliferating cells is due to the arrest of proliferation in these cells, rather than some artefact of the extraction procedure. Both sequences (F6 and F8) are AT-rich and are located within MARs (Figure 1A); at least one of them (F8) seems to be conserved in evolution, since the factor interacting with this sequence also exists in duck blood cells (data not shown).

The F6 binding sites resembles that of the RIP 60 origin-binding protein identified by Dailey et al (25) and contains a GATA-1 recognition site. Interestingly, competition experiments revealed that the canonical Eryf-1 (GATA-1) binding site (8) could compete with the F6 site for binding with the protein extract (data not shown). This is surprising since the protein factor that binds to the F6 sequence exists in nuclear extracts from non-erythroid chicken cells and is different from that described by Evans and Felesenfeld (26); its presence in cells correlates with cell proliferation rather than expression of the globin gene domain. Binding assays with the synthetic oligonucleotide probes (5'-TATTTGATAGCAATATTTAGTATTT-3' for F6 and 5'-AAATGAAAGAATAAAG-3' for F8 demonstrated that the identified binding sites really interact with the sequence-specific nuclear factors from proligerating cells and effectively compete for proteins from the nuclear extract with the original subclones of the 1.7 kb fragment of the chicken DNA (data not shown).

We should also point out that the DNA fragment containing the replication origin and matrix attachment site studied in the present paper, is transcribed in normal and transformed erythroblasts within the full-domain transcript (3,4) but could be also transcribed locally.

The identification of ten novel protein binding motifs communicated in the present paper seems of interest since these sites are located within several elements related to structural organization, transcription, and replication of the chicken α-globin gene domain. Further work will be necessary to attribute each of the protein-binding sites to one of the above mentioned functions. Of particular interest will be the partition between sites involved in matrix attachment and replication since both of the functions imply *de novo* protein synthesis in proliferating cells. The enhancer function that was found not to be related to tissue-specificity (15) may be connected with replication rather than expression of the domain. Work in progress might lead to such understanding, as well as to the isolation and further characterization of the protein factors for which we have identified and reported the binding sites.

Figure 2. Distribution of protein-binding sites F1-F8 (open boxes) within the sequence of the 1.7 Kbp DNA fragment of the α-globin gene domain containing the replication origin.


Xma I
F1
F1a
Xma I
Xma I
GGGTCGT
F3
GGGTCGT
Xma I
GTCCAGTGCAACCTCAGAG
F3
F2
GACGCGTCACAGAGT
F2a
GGGTCGT
F3
GGGTCGT
F3
GGGTCGT
Kpn I
F3
F4
F6
F5
F5
F3
F7
Pst I
F8

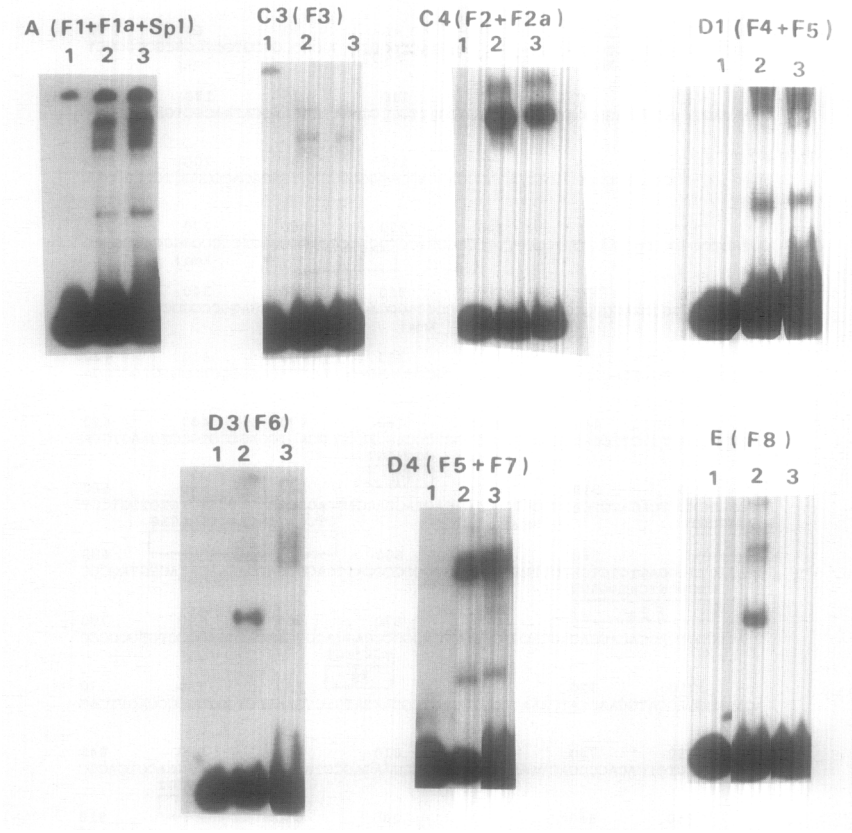


Figure 3. Electrophoretic mobility shift experiments with the nuclear protein extracts from proliferating and non-proliferating HD3 cells. Five nanograms of the probes (different ^{32}P-end-labeled subfragments of the 1.7 Kbp DNA fragment containing the binding sites for each of the identified factors) were incubated in the presence of 5µg poly(dI/dC) with 1 3-6 µg of nuclear protein extract from non-differentiated and differentiated HD3 cells (differentiation was induced by incubation in the DMEM medium containing 10% fetal calf serum, 4% chicken serum and 20 mM 1-(5-Isoquinolinesulfonyl)-2-methylpiperazine dihydrochloride at 42°C) . Each panel contains (1) control (incubation without extract); (2) incubation with an extract from proliferating HD 3 cells; (3) incubation with an extract from non-proliferating differentiated HD 3 cells.

ACKNOWLEDGMENTS

The authors are grateful to François STRAUSS for valuable discussion on the technical approach and the results of this work. We would like to thank Geoff PARTINGTON from the laboratory of Dr GOODWIN for the kind gift of HD3 cells and Michèle HUESCA for cell culture. The help of Richard SCHWARTZMANN in the preparation of the manuscript is gratefully acknowledged. C.M.G. is a recipient of a fellowship from the CAPES-BRAZIL and from the UERJ-BRAZIL. Y.V . received the C.I.E.S. fellowship of the M.R.T. F.R.T. is a recipient of a fellowship from the CONACYT, México. S.R. had short term FEBS and EMBO fellowships. This work was supported by the French CNRS, the Fondation pour la Recherche Médicale, the INSERM, the Ministère de la Recherche et de l'Enseignement Supérieur, the Association pour la Recherche sur le Cancer (ARC) and the Ligue Nationale Française contre le Cancer.

REFERENCES

1.Moreau J., Marcaud L., Maschat F., Kejzlarowa-Lepesant J., Lepesant J.-A., and Scherrer K. (1982) Nature 295, 260-262

2.Kretsovali A., Marcaud L., Moreau J., and Scherrer K. (1986) Mol. Gen. Genet. 203, 193-201
3.Broders F., and Scherrer K. (1987) Mol. Gen. Genet. 209, 210-220
4.Broders F., Zahraoui A., and Scherrer K. (1990) Proc. Natl. Acad. Sci. USA 87, 503-507
5.Dodgson J.B., McCure K.C., Krust A., and Engel J.D. (1981) Proc. Natl. Acad. Sci. USA 78, 5998-6002
6.Engel J.D., and Dodgson J.B. (1980) Proc. Natl. Acad. Sci. USA 77, 2596-2600
7.Kretsovali A., Huesca M., and Marcaud L. (1988) C.R. Acad. Sci. Paris 307, 563-568
8.Knezetic J.A., and Felsenfeld G. (1989) Mol. Cell. Biol. 9, 893-901
9.Razin S.V., Rzeszowska-Wolny J., Moreau J., and Scherrer K. (1985) Molekularnaya Biologia USSR 19, 456-466
10.Razin S.V. (1987) BioEssays 6, 19-23
11.Farache G., Razin S.V., Rzeszowska-Wolny J., Moreau J., Recillas Targa F., and Scherrer K. (1990b) Mol. Cell. Biol. 10, 5349-5358
12.Razin S.V., Kekelidze M.G., Lukanidin E.M., Scherrer K., and Georgiev G.P. (1986) Nucl. Acids Res. 14, 8189-8207
13.Umek R.M., Linskens M.H., Kowalski D., and Huberman J.A. (1988) Biochim. Biophys. Acta 1007, 1-14
14.Weintraub H., Larsen A., and Groudine M. (1981) Cell 24, 333-344
15.Razin S.V., Vassetzky Y.S., Grinenko N.F., and Georgiev G.P. (1991) J. Mol. Biol. 217, 523-527
16.Kalandadze A.G., Bushara S.A., Vassetzky Y.S., and Razin S.V. (1990) Biochem. Biophys. Res. Comm. 168, 9-15
17.Farache G., Razin S.V., Recillas-Targa F., and Scherrer K. (1990a) Nucl. Acids Res. 18, 401-409
18.Beug H., Von Kirchbach A., Doderlin G., Conscience J.F., and Graf T. (1979) Cell 18, 375-390
19.Scherrer K., Marcaud L., Zajdela F., London I.M., and Gros F. (1966) Proc. Natl. Acad. Sci. USA 56, 1571-1578
20.Plumb M.A., Lobanenkov V.V., Nicolas R.H., Wright C.A., Zavou S., and Goodwin G. (1986) Nucl. Acids Res. 14, 7675-7693
21.Fried M., and Chrothers D.M. (1981) Nucl. Acids Res. 9, 6505-6525
22.Maniatis T., Fritsch E.F., and Sambrook J. (1982) Molecular cloning , a laboratory manual .Cold Spring Harbor Laboratory Press
23.Sienbelist V., and Gilbert W (1980) Proc. Natl. Acad. Sci. USA 77, 122-126
24. Razin S., Vassetzky Y., and Hancock R. (1991) Bichem. Biophys. Res. Comm. 177, 265-270
25.Dailey L., Caddle M.S., Heintz N., and Heintz N.H. (1990) Mol. Cell. Biol. 10, 6225-6235
26.Evans T., Reitman M., and Felesenfeld G. (1988) Proc. Natl. Acad. Sci. USA 85, 5796-5980.
27.DePamphilis M.L. (1988) Cell 52, 635-638