Positive and negative transcriptional regulatory elements in the early H4 histone gene of the sea urchin, *Strongylocentrotus purpuratus*

Lin Tung, Insong J.Lee+, Howard L.Rice and Eric S.Weinberg*
Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA

Received September 4, 1990; Revised and Accepted November 16, 1990

ABSTRACT

The early H4 (EH4) histone gene of the sea urchin, *Strongylocentrotus purpuratus*, is shown to contain at least five positive-responding sequence elements and one negative-responding site which control the level of *in vitro* transcription in an embryonic nuclear extract. The positive acting elements are: 1) the UHF-1 region, located between −133 and −102 (the site of a strong footprint, due at least in part to the binding of an 85 kD protein factor termed UHF-1); 2) the H4 specific element (H4SE), situated between −62 and −39; 3) a sequence corresponding to a TATA box between −33 and −26 (TAACAATA); 4) the transcriptional initiation site; and 5) an internal sequence element found between +19 and +50. Deletion of, or base changes in, the UHF-1, H4SE, initiation, or internal sequence sites resulted in significant decreases in transcription. Base subtitutions in the TATA-like sequence had much less effect, resulting in no more than a 2-fold decrease in transcription, and there was no evidence that alternative initiation sites are utilized in the mutant templates. The negative element (termed the UHF-3 site) is contained within a footprinted region between nucleotides −75 and −56. Base subtitutions in this area result in templates that were transcribed at a level 1.2−2.0-fold higher than the wild-type gene. Transcription levels of double UHF-1/H4SE and UHF-1/INR mutants were those expected from additive effects of the individual mutations and there was no suggestion of synergism.

INTRODUCTION

Two types of H4 histone genes are transcribed during development of the sea urchin embryo: the early (EH4) and late (LH4) genes. The EH4 genes are part of the early gene set, consisting of tandem arrays of several hundred copies of a 6−7 kb unit, which contains genes for all five histone proteins (reviewed in refs. 1,2). The LH4 genes are part of a late histone gene complex of 5−12 genes for each nucleosomal histone (3−6) and at least two H1 genes (7−9), arranged in small irregular clusters or found singly. The early genes are first transcribed in the oocyte to yield a pool of stored maternal histone mRNAs, and then become active again in the 16-cell embryo. The level of early histone mRNA increases 10-fold from the 16-cell stage to the early blastula stage and then decreases rapidly (10−12). The early genes are evidently not reactivated until oogenesis in the adult. Late gene mRNAs are present in low levels in the egg and, depending on the gene, increase during the blastula and gastrula stages (4,8,9,13−18). The basis of the differential utilization of the gene sets appears to be transcriptional, as indicated by nuclear run-on experiments (16,19, Yager and Weinberg, unpublished experiments).

Sequence homologies often provide candidates for functional transcriptional elements. When sequences of the 5′ flanking regions of EH4 and LH4 genes were compared (5,20,21), the region found to have the highest degree of conservation was located just upstream from the TATA box. This sequence contains a number of motifs previously identified as typical of vertebrate and echinoderm H4 genes (1,22−31). This region, which we will refer to as the H4 Specific Element (H4SE), consists of two domains. The proximal domain of echinoderm, mammalian, and avian H4SEs, located 5−15 bases upstream from the TATA element, contains the sequence GTCCG. Sea urchin EH4 and LH4 genes share an even more extensive consensus in this region: GTCCGC$^{A}/_{T}$. (Curiously, sea urchin H1 histone genes also have a sequence identical to the H4SE proximal domain at the same position relative to the TATA box [32].) The distal domain of the H4SE, consisting of 12−20 nucleotides, is quite different for vertebrate and echinoderm H4 genes, although this region often contains the $^{G}/_{A}$TCA sequence motif first noted by Clerc et al. (23). The H4SE appears to have an important role in transcription since expression of a *Xenopus* H4 gene, after injection into a *Xenopus* oocyte (23), and expression of a human histone gene, after transfection into mouse C127 cells (33), was completely dependent on the presence of a region containing the element. Information on the proteins which interact with the H4SE comes exclusively from work on two human H4 genes. The full extent of the H4SEs of these genes is protected from

* To whom correspondence should be addressed

+Present address: Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA

DNase I digestion in *in vitro* footprinting experiments and specific proteins (H4TF-2 and HiNF-D) have been found to interact with the region (34−37).

Sequences upstream of the H4SE have also been shown to be important for H4 gene transcription. Wells (30) noted that a number of vertebrate H4 genes have purine-rich stretches on the sense strand, 50−100 nucleotides upstream of their H4SEs. A G-rich region located between −79 and −105 of the human Hu4A gene forms a strong footprint with HeLa cell extracts and is required for maximal transcriptional activity in a nuclear extract made from S phase cells (34,35,38). Proteins of 110 and 105 kD, termed H4TF-1, which bind to the site, have been purified from HeLa nuclear extract (34). A purine-rich region situated 75 to 110 nucleotides upstream of the cap site of another human H4 gene, FO108, was also found to be required for maximal *in vivo* (33) and *in vitro* (35) expression, and two factors, HiNF-A and HiNF-C (Sp1-like), which bind to this region, have been identified (36,37). Somewhat analogously, we have found that a G-rich region (termed the UHF-1 site), located upstream of the H4SE of *S. purpuratus* EH4 and LH4 genes, is required for maximal transcription of these genes in a sea urchin embryonic nuclear extract (Lee et al., manuscript submitted). These regions (−133 to −103 of the EH4 gene and −94 to −66 of the LH4 gene) produce sharp DNase I footprints and contain a consensus sequence, $AGG^{G}/_{C}GGC^{G}/_{T}CACTC$. Deletion of the site in either gene results in a 2−5-fold decrease in template activity in the extract. Embryonic expression of the EH4 gene lacking a UHF-1 site, after injection into eggs, is decreased to a similar extent. We have identified an 85 kD protein which binds to the site in both genes. The organization of regulatory elements in sea urchin and human H4 genes, therefore, appears to be quite similar.

We wished to determine whether there are additional sequence elements which have a role in transcription of the sea urchin EH4 gene. Particularly, we sought to determine if there were negatively responding sequences which might be involved in the shut-down of transcription. We have used an embryonic nuclear extract, capable of proper transcriptional initiation (39,40), to test a series of EH4 gene mutants. We used extracts from 18 hour embryos for this work since cleaner nuclear preparations with a higher yield per embryo could be obtained at this stage than at earlier stages, and the EH4 gene was transcribed with high efficiency (40). In addition to reporting results of experiments in which the H4SE and a negatively responding element (termed the UHF-3 site) were mutated, we also show that whereas changing the site corresponding to the TATA element has only a minor effect on the level of EH4 transcription, mutations in the initiation site and elimination of a site within the region coding for the untranslated leader cause dramatic decreases in transcription *in vitro*.

MATERIALS AND METHODS

In vitro transcription in an embryonic nuclear extract

Nuclei and nuclear extracts were prepared from *S. purpuratus* 18 hour hatching blastula embryos by procedures of Morris and Marzluff (41) and Morris et al. (39), as modified by Tung et al. (40). Transcription in the nuclear extracts followed the protocol of Tung et al. (40). Template DNAs were all derivatives of plasmid pEH4-EH2B(BK), which carries oppositely oriented EH4 and early H2B (EH2B) genes (see Figure 1A). Plasmids cut with *Nco*I were transcribed to produce run-off products of 322 and 252 nucleotides, corresponding to correctly initiated EH4 and EH2B RNAs. (The position of the 5′ end of the EH4 mRNA had previously been determined by S1 nuclease protection [Kaumeyer and Weinberg, unpublished results] and that of the EH2B mRNA by primer extension [42]). Quantitation of transcription was performed by scanning X-ray films of the gels with an LKB UltraScan laser enhanced densitometer and normalizing the amount of H4 transcript produced in each reaction by the level of EH2B transcript. Films with exposures in the linear response range were used for scanning.

DNA templates

The EH4 and EH2B genes used in these studies are derived from the *S. purpuratus* early repeat unit, originally cloned as plasmid pCO2 (43). The plasmids pEH4 and pEH2B (40) were subclones of pCO2, inserted in the vector pUC118 (44). The plasmid pEH4-EH2B(BK) containing oppositely oriented EH4 and EH2B genes (see Figure 1A), was derived from pEH4 and pEH2B (Lee

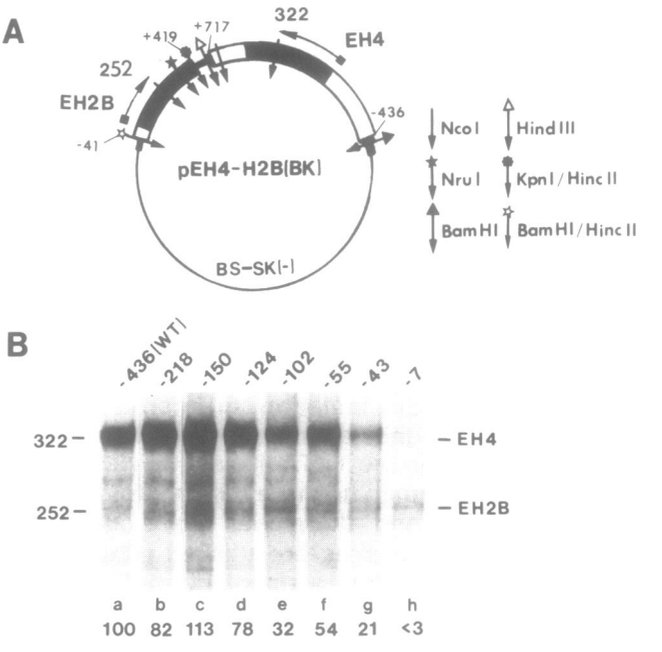

Figure 1. Effect of upstream deletions on transcription of the EH4 gene. (A) Map of pEH4-EH2B(BK), a construct containing an EH4 gene and an oppositely oriented EH2B gene. The thin line represents the pBluescript vector and a slightly thicker line indicates the Bluescript polylinker. An even thicker line represents the residual pUC118 polylinker at the ends of the EH4 segment. Filled-in boxed areas are the transcribed regions of the EH4 and EH2B genes and open boxed areas are regions 5′ and 3′ of these genes. The number of nucleotides upstream and downstream of the two transcriptional initiation sites and the restriction enzyme sites used for insertion of the EH4 and EH2B gene fragments are indicated. Cutting the plasmid with *Nco*I results in a template which supports the synthesis of a 322 base EH4 RNA and a 252 base EH2B RNA, as diagrammed. The position of the *Nru*I site used to prepare template for the experiment of Figure 4B is also shown. The *Bam*HI/*Hinc*II and *Kpn*I/*Hinc*II joints, formed by the blunt ended insertion of the EH2B fragment, no longer contain recognition sites for these enzymes. (B) *In vitro* transcription products produced with plasmid templates containing decreasing amounts of 5′ flanking sequence. *Bal*31 was used to generate a series of 5′ deletions of the EH4 promoter region. The values above each lane refer to the number of bases upstream of the EH4 transcriptional initiation site still present in the construct. The value below each lane is the percentage of wild-type transcription obtained with the EH4 template of that particular reaction. The amount of EH4 and EH2B transcripts of each lane were determined densitometrically and the EH4 RNA level was normalized by dividing by the level of EH2B RNA. The −436 construct is considered to be the wild-type (WT) gene in this and all other Figures.

et al., submitted for publication). pEH2B was digested with *Kpn*I, treated with S1 nuclease, then digested with *Bam*H1, and filled in with Klenow enzyme. The resulting 460 bp fragment, extending from −41 to +419 relative to the EH2B cap site, was blunt end ligated into the *Hinc*II site of Bluescript (SK−) to yield the plasmid pEH2B(BK). A 1.1 kb EH4 gene fragment, consisting of the EH4 transcribed gene plus 436 bp of sequence upstream from the mRNA cap and 281 bp downstream from the 3′ mRNA terminus, was obtained from pEH4 by digestion with *Bam*HI and *Hind*III, enzymes whose cutting sites flank the insert in the polylinker of pUC118. The fragment was inserted into the *Bam*HI and *Hind*III polylinker sites of pEH2B(BK), as diagrammed in Figure 1A, to yield the plasmid pEH4-EH2B(BK). Some of the mutant EH4 templates (including a *Bal*31 5′ deletion series [40] and a number of internal deletions, see below) originally had been created in pEH4. *Bam*HI-*Hind*III fragments of these mutant constructs were also recloned in pEH2B(BK) to give a series of pEH4-EH2B(BK) plasmids carrying altered EH4 genes.

Construction of the EH4 deletions lacking bases −135 through −103 (Δ−135/−103) and −135 through −44 (Δ−135/−44) utilized *Bal*31-produced derivatives of pEH4 which had 5′ flanking sequence removed up to bases −102 or −43. An *Xba*I-*Hind*III fragment of each of these plasmids (containing the EH4 gene and a small portion of the pUC118 polylinker on each side) was ligated to an *Nhe*I-*Hind*III fragment of pEH4 (which lacked the EH4 sequence except for the region upstream of −135), filled in with Klenow fragment, and then blunt end ligated to circularize the plasmid. (These constructs are termed deletions but actually carry 4 bp (AGTC) from the polylinker in place of 33 or 93 bp of EH4 upstream sequence.) *Bam*HI-*Hind*III fragments of these two plasmids were recloned in pEH2B(BK).

Additional mutant pEH4-EH2B(BK) plasmids were created by oligonucleotide priming of pEH4-EH2B(BK) single stranded DNA produced from *dut*⁻ *ung*⁻ *E. coli* strains (45) or by priming of single stranded DNA in the presence of dCTP-α-S and processing the heteroduplex by the method of Eckstein and colleagues (as described in the Amersham kit protocol). Double mutants were prepared, in which single stranded pEH4(Δ−135/−102)-EH2B(BK) DNA was primed with oligonucleotides containing the base substitutions.

We also prepared three constructs in which 3′ portions of the EH4 gene were removed. Sequences within the EH4 transcribed gene, in the plasmid pEH4-EH2B(BK), at nucleotides +20 to +25, +51 to +56, and +103 to +108 were individually converted to *Hind*III sites by oligonucleotide-directed mutagenesis using the method of Kunkel et al. (45). The stretch of sequence between the newly created *Hind*III site and the *Hind*III site in the polylinker, 3′ of the EH4 gene, was eliminated by digestion of pEH4-EH2B(BK) with *Hind*III and circularization of the larger *Hind*III fragment by ligation. The resulting constructs consisted of only the first 19, 50, or 102 nucleotides of transcribed region of the EH4 gene, oppositely oriented from the standard EH2B gene (see Figure 4A).

RESULTS

Multiple transcriptional elements in the EH4 promoter

An intitial *in vitro* transcriptional analysis of *Bal*31-produced upstream deletion mutant templates indicated that there were a number of distinct positively responding elements in the promoter of the *S. purpuratus* EH4 gene (40). One of these elements is coincident with a region between −133 and −103 which forms a strong footprint with embryonic nuclear extracts and which binds an 85 kD factor which we have termed UHF-1 (Lee et al., submitted for publication). Using the nuclear extract, we now have transcribed a more extensive series of *Bal*31 5′ nested deletions of the EH4 gene promoter region. Plasmids containing an EH2B gene as a reference standard, situated in opposite orientation from an EH4 gene (as diagrammed in Figure 1A), were used as templates in this experiment (Figure 1B). Removal of sequences between −436 and −124 resulted in no reproducible loss of template activity in the assay. Deletion to −102 (lane e), on the other hand, caused a three-fold loss of activity (see Table 1 for additional data), consistent with the loss of the UHF-1 site. A further deletion, resulting in a template with just 55 nucleotides of upstream sequence, led to a moderate increase in template activity (lane f). This observation was reproducible, and tests of additional mutant templates (see below) indicate that a negatively responding element is present in the promoter between nucleotides −102 and −56. When an additional 12 nucleotides are removed (including almost all of the H4SE), the template is approximately five-fold less active than the −436 construct (lane g). This result confirms our previous findings (40), and additional data on the −43 template are presented below (see Figure 5 and Table 1). A template from which all but 7 nucleotides upstream from the transcriptional start site have been

Table 1.

Template	Mean[a]	S.E.[b]	N[c]
Wild-type (−436)	100		
INR-1	6.5	3.7	6
TA-1	77	23	4
TA-2	88	32	4
TA-3	74	19	5
H4SE-1	80	26	5
H4SE-2	16	5	9
H4SE-3	75	20	4
H4SE-4	24	6	6
H4SE-5	27	18	2
H4SE-6	54	14	2
UHF-3-1	137	18	5
UHF-3-2	138	22	5
Δ−135/−103[d]	31	9	8
−102[e]	31	3	3
Δ−135/−44[d]	18	3	4
−43[e]	17	3	5
Δ−135/−103[d] + INR-1	1.2	0.8	2
Δ−135/−103[d] + H4SE-2	3.4	0.9	6
Δ−135/−103[d] + UHF-3−1	49	11	3
Δ−135/−103[d] + UHF-3−1	42	16	3

[a] The amount of EH4 transcript produced using a particular mutant template was normalized by the amount of EH2B transcript produced in the same transcription reaction, and then compared to the amount of EH4 transcript made using the wild-type template, also normalized for EH2B transcription. The wild-type EH4 template was assigned a value of 100 and the amount of transcript produced by each mutant template was expressed as a percentage relative to the wild-type transcript, as in Figures 1−3 and 5. The mean of the relative amount of transcription for a particular mutant was calculated using the percentage values determined in independent experiments.

[b] Standard Error

[c] N = the number of independent determinations. The experiments of Figures 1−3 and 5 are included, as are other experiments that tested various subsets of mutant templates.

[d]Numbers refer to the position of bases deleted from the template.

[e]Numbers refer to the position of the 5′-most nucleotide still present in the template.

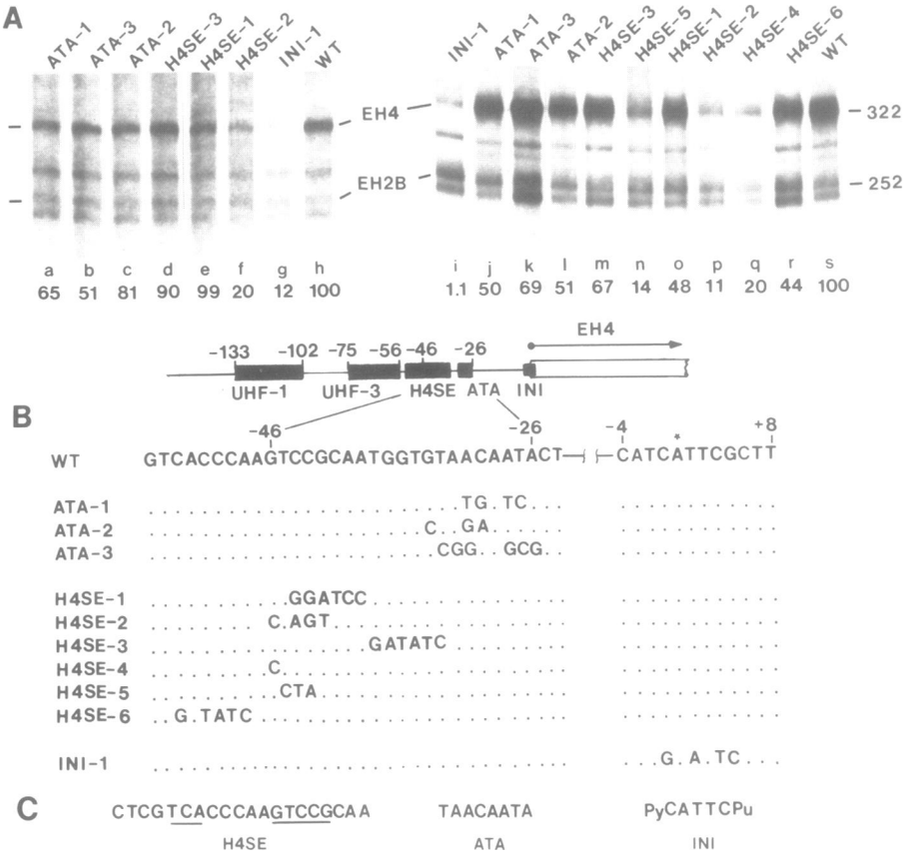

Figure 2. Effect of base substitutions in the USE1, ATA, and initiation site regions on transcription of the EH4 gene. (A) Individual base substitution mutants were transcribed in a nuclear extract in two separate experiments, presented in the two panels. All constructs carried a reference EH2B gene as diagrammed in Figure 1A. The percentage of wild-type (−436) transcription obtained with the mutant template tested in that particular reaction is noted below each lane. All values are normalized for EH2B transcription. (B) The wild-type sequence for the region tested and the actual mutational changes are indicated. Dots indicate unchanged sequence. (C) Consensus sequences for the sea urchin EH4 H4SE, EH4 ATA box, and histone gene initiation site. Underlined in the H4SE sequence are the GTCCG motif, conserved in all sea urchin and most vertebrate H4 genes, and the ATA motif found in the more variable portion of the H4SE of almost all H4 genes. The conserved sea urchin EH4 H4SE actually extends an additional 4 bases further upstream to position −62. The H4SE distal domain shared in most sea urchin early and late H4 genes and in the USE0 sequence of early and late H1 genes extends from position −46 to −39 (GTCCGCAA). The H4SE proximal domain present in sea urchin EH4 genes extends from −62 to −47.

removed appears to be very weakly transcribed with less than 3% of the transcriptional activity of the −436 construct (lane h). However, in many cases we could not detect any specifc transcript with the −7 template.

The −436 construct is considered to be the 'wild-type' template in these experiments since it is transcribed as well as templates with considerably less upstream sequence. Furthermore, we have shown in DNA injection experiments that an EH4 gene containing 436 bp of upstream sequence is expressed *in vivo* as efficiently as an EH4 gene in a complete early repeat unit containing all five early histone genes. Removal of all sequence upstream of −436 and downstream of +720, relative to the EH4 transcriptional start site, had no effect on the level or timing of EH4 gene expression (Lee et al., submitted for publication).

Transcription of base subtitution mutants in the ATA, H4SE, and initiator regions

The results of transcription of the 5′ deletion constucts suggested that in addition to the UHF-1 site, there are at least two additional positively responding *cis*-acting elements in the EH4 promoter. One of these elements is located between nucleotides −55 and −43, the other between −42 and −7. We therefore prepared base substitution mutants (shown in the lower part of Figure 2) in the H4SE and the ATA regions, the two sequences which we thought likely to have a role in transcription. The early H4 genes of several sea urchin species contain a highly conserved but non-standard sequence at the exact position expected of a TATA box. This sequence, TAACAATA, referred to here as an 'ATA' motif, is not found in the promoters of other early or late sea urchin histone genes. We prepared three mutants which had base substitutions in this motif (ATA-1, 2, and 3). In the experiments illustrated in Figure 2A, these base changes had only a small effect on the level of EH4 transcription. The decrease observed in these experiments was no more than two-fold, and in experiments using other extracts (summarized in Table 1), the effect of these mutations varied from a two-fold decrease to no reduction at all. The elimination of the ATA site did not appear to have an effect on the position of the RNA initiation site since the run-off EH4 product was exactly the same size as the wild-type control and no additional bands were seen. The effect of the ATA site mutations was considerably less pronounced than had been expected from the difference between transcriptional efficiency of the −43 and −7 templates. Either an additional positive-responding element must reside in this region, or the lack of upstream sequence in the deletion constructs results in a dependence on the ATA sequence. Our reasons for favoring the latter alternative will be noted in the Discussion section.

Figure 2A also illustrates the results of transcription of H4SE

site mutations. The sequence between −62 and −39 of the *S. purpuratus* EH4 gene (46,47) is almost identical to the corresponding sequences from other sea urchin species (22,23,48). The major part of the sea urchin early H4SE is shown in Figure 2C. The sequence between −46 and −42, GTCCGC, is conserved in early and late sea urchin H4 genes, in all reported mammalian and chicken H4 genes, and in at least one *Xenopus* H4 gene. Four mutants (H4SE-1, -2, -4, and -5), containing base subtitutions in this most conserved part of the H4SE, were tested in the *in vitro* transcription system. We also tested one mutant (H4SE-6) with changes in the distal domain of the H4SE, specific for sea urchin EH4 genes, and one mutant (H4SE-3) containing base subtitutions just downstream of the conserved sequence. Although some effect was seen with all the mutant templates we used, only mutants H4SE-2, -4, and -5 showed large and reproducible decreases in transcriptional efficiency. In the two experiments shown in Figure 2A, these mutants were transcribed at 11−20% of the wild-type level. Although the most meaningful way to look at the data is to note the relative transcriptional activity within a single experiment, averages of multiple experiments using different extracts were also significant (Table 1). The effects of H4SE-1, -3, and -6 were not as dramatic. H4SE-1 and -3 templates were transcribed at wild-type levels in the experiment shown in the left panel of Figure 2A and at somewhat reduced levels in the experiment in the right panel. In the averages presented in Table 1, the decreases are not significant. The H4SE-6 template was transcribed at 44% and 64% of wild-type levels in two separate experiments, a moderate decrease.

From the overlap of substitutions in mutants H4SE-2, -4, and -5, and the lack of a strong effect of H4SE-1, it appears that the most crucial nucleotides for transcription are the G and T at positions −46 and −45. Surprisingly, transcription of the H4SE-1 template, in which the next six positions (CCGCAA) have been changed, is not reproducibly decreased. Since the sequence GTCCGCAA is highly conserved in the H4SEs of sea urchin early and late H4 genes, it might have been expected that changes anywhere in this motif would have resulted in a strong impairment of transcription. The decrease in transcription observed with mutant H4SE-6, with base changes in the region of the H4SE specific for sea urchin early genes, is not as profound as was observed with H4SE-2, -4, or -5. Finally, the transcription of mutant H4SE-3, which contains changes in the region between the H4SE and the ATA region, was not reproducibly lowered.

We also tested a base substitution mutant in which the sequence at the site of transcriptional initiation was altered. Sea urchin histone genes have a consensus sequence at the cap site, PyCATTCPu, where A is the first nucleotide in the transcript (1,42,49). The *S. purpuratus* EH4 initiation region conforms exactly to the consensus sequence. Transcription of this mutant (INR-1) was very poor, reproducibly yielding only 1−8% of wild-type template activity (an average of 6.5% in 6 experiments, Table 1). The initiator region, therefore, is extremely important for EH4 transcription.

UHF-3, a negative-responding element between −62 and −73

The results of transcription of the *Bal*31 deletion series (Figure 1) suggested that there might be a sequence element between −55 and −102 which responded negatively to a factor in the extract. To further explore this possibility, we performed gel mobility-shift and footprinting assays with fractions of the extract obtained from chromatography on DEAE cellulose (DE52) (data not shown). The 0.1 M KCl flow-through fraction from the column

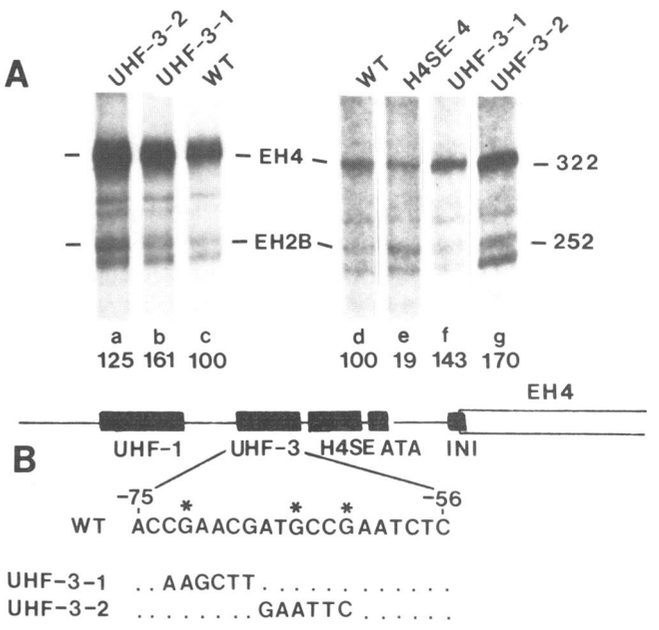

Figure 3. Effect of base subtitutions in the UHF-3 region on transcription of the EH4 gene. (A) Experiments using two different extracts are presented in the two panels. In each case, templates containing base subtitutions in the UHF-3 region were compared with the wild-type (−436) EH4 template. In the panel on the right, a H4SE-4 mutation was also tested. Percentages of wild-type EH4 transcription are shown below each lane. (B) The sequence protected by protein in a DNase I footprint experiment is shown below a diagram of the EH4 promoter region. Asterisks indicate Gs which interfere with binding after methylation by dimethylsulfate. Below the sequence are indicated the base substitutions carried by the two mutants in the UHF-3 region.

contained a protein which specifically bound to EH4 promoter fragments which included the region between −151 and +27. Competition experiments with various EH4 fragments showed that the factor was bound between −102 and −43. DNase I footprinting with labeled sense strand resulted in a strongly protected region between −75 and −56. Methylation of Gs at positions −72, −65, and −62, was found to strongly interfere with binding of the factor. A direct repeat of the sequence CCG-AA, separated by 5 bp, was located in the footprinted region. The protein which interacts with this region has been termed UHF-3 and the footprinted region is denoted the UHF-3 site. These findings are summarized in Figure 3.

To determine if this region of the promoter was involved in negative regulation, we prepared two mutant templates (UHF-3−1 and −2) containing base substitutions in the footprinted region. The base changes, noted in Figure 3, include the G positions which showed methylation interference. The results of two separate transcription experiments are illustrated in Figure 3. In both cases, the mutant templates were more effectively transcribed than the wild-type template (although the UHF-3−1 mutant is a better template in one experiment and the UHF-3−2 template is transcribed more efficiently in the other). An increased transcriptional level of these mutant templates is also seen in the experiment of Figure 5. The average increase for both mutants in five experiments is about 40% over wild-type template activity (Table 1). These results indicate that the sequence between −73 and −62 is the site of a negatively-responding transcriptional element, a finding consistent with the transcription of 5′ deletion templates. Although the effect of either

of the two mutations is not that large, the results are reproducible. Furthermore, additional evidence using deletions, to be presented below (Figure 5), is consistent with our interpretation of UHF-3 as a negatively-responding element.

An internal sequence element with a positive transcriptional effect

To examine the possibility that sequence elements downstream from the transcriptional initiation site have a role in the control of EH4 gene expression, we prepared three templates which lacked sequences downstream of nucleotides +19, +50, and +103 (see Materials and Methods and Figure 4A for details of their construction and structure). These plasmids contain one *Nru*I site and one *Nco*I site, located in the oppositely oriented EH2B gene (the *Nco*I site of the EH4 gene is no longer present in the 3′ deletion constructs). Templates for *in vitro* transcription were prepared by cutting each plasmid with either *Nru*I or *Nco*I (as diagrammed in Figure 4A). The EH2B run-off transcript is either 350 or 252 nucleotides in legnth, depending on the enzyme used to cut the plasmid. The EH4 transcripts include various amounts of EH4 sequence (19, 50, or 103 nucleotides in the A, B, and C constructs, respectively), the adjacent newly created *Hind*III site, a portion of the polylinker, and the antisense sequence of the 3′ end of the EH2B gene, up to the *Nru*I or *Nco*I site. The sizes of the run-off RNAs which would be expected from transcription of the three templates are shown in Figure 4A.

The result of transcribing the *Nru*I-digested templates is illustrated in Figure 4B. Constructs C and B, containing 102 and 50 bases of EH4 sequence respectively, were transcribed efficiently to yield run-off RNAs of approximately the correct size. No decrease in transcriptional efficiency resulted from removal of nucleotides between +102 and +50. In contrast, in this experiment, no run-off product could be detected when construct A, containing only 19 bases of EH4 sequence, was used as a template. The specific run-off RNAs were all sensitive to α-amanitin. The degree of decrease in transcriptional efficiency

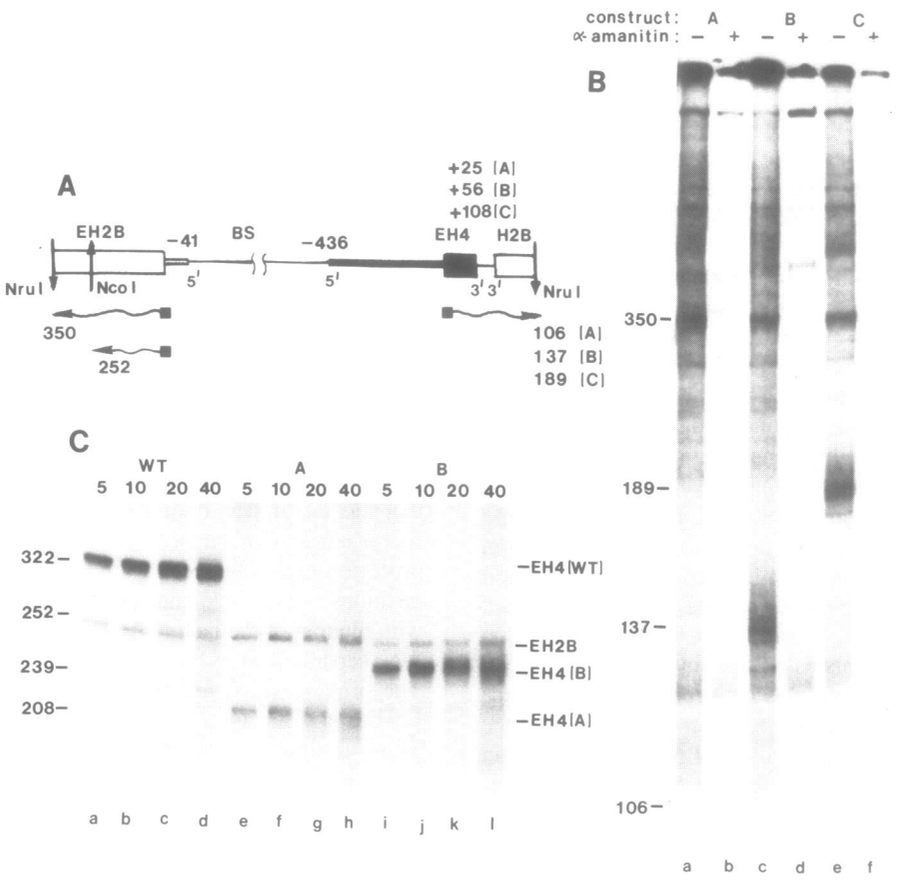

Figure 4. The effect of 3′ deletions on EH4 gene transcription. (A) A diagram of constructs A, B, and C, lacking EH4 sequence downstream of positions +19, +50, and +102, respectively. The plasmids contain one *Nru*I site and one *Nco*I site, both within the EH2B gene. As diagrammed, the *Nru*I-cut plasmid serves as a template for a 350 base EH2B RNA and for 106, 137, or 189 base RNAs initiated from the EH4 promoter. Transcription of the *Nco*I-cut template results in formation of a 252 base EH2B RNA and of 208, 239, or 291 base RNAs initiated from the EH4 promoter. In the diagram, the DNA derived from the EH4 gene is shown as a dark box or dark thick line, that derived from the EH2B gene as an open box or open thick line, and vector DNA as a thin line. The 5′ ends of the EH2B and EH4 genes are noted as −41 and −436, respectively. The position of the 3′-most nucleotide in each of the deleted EH4 genes is noted directly above the EH4 gene in the diagram. (B) Transcription of the three deletion constructs cut with *Nru*I. As was the case in all of the previously presented transcription experiments, the reactions were carried out for 30 minutes. The whole gel is shown to indicate that the expected bands are specifically produced. Lanes b, d, and f are products of reactions that were carried out in the presence of 2.5 μg/ml α-amanitin. (C) Time course of transcription with the wild-type (−436) template and with constructs A and B, all cut with *Nco*I. Reactions were allowed to procede for 5, 10, 20, or 40 minutes, as noted above each lane. The wild-type template was the same as used in the experiments of Figures 1−3 and 5. Constructs A and B were the same as used in panel B, except that they were cut with *Nco*I and therefore yielded transcripts of different sizes, as described above. Diffferent extracts were used for the experiments of panels B and C. (D) Time course of transcription exactly as in panel (C), except that at 20 minutes after the start of the transcription, the concentration of UTP in the reaction mixes was increased from 30 μM to 600 μM by addition of unlabeled nucleotide (note arrows). The reactions were allowed to continue for an additional 20 minutes.

of construct A compared to constructs B and C was variable from extract to extract. In three extracts, no transcription could be detected using construct A. In two extracts, some transcription of construct A could be measured, but the deletion of sequence between +19 and +50 still caused a profound decrease in template activity (as in the experiments of Figure 4C and D).

In contrast to the experiments using templates with alterations in the 5′ flanking region of the EH4 gene, the use of the 3′ deletions resulted in the production of different RNA species. We considered the possibility that the removal of different parts of the EH4 transcribed sequence could result in the formation of RNAs with different stabilities in the nuclear extract. We therefore performed two time course experiments with constructs A and B, as well as the wild-type pEH4-EH2B(BK) template: continuous labeling over a 40 minute period, and a pulse-chase experiment (plasmids used in for these reactions were cut with *Nco*I, resulting in run-off RNAs with different sizes from those of Figure 4B). In the first experiment, the reaction was carried out, for 5, 10, 20, and 40 minutes (Figure 4C) and the intact EH4 gene was transcribed efficiently over the whole period (lanes a−d). Construct B is transcribed as efficiently as the wild-type EH4 gene during the first 10 minutes of the reaction (lanes j and i). However, the amount of transcript does not increase from 10 to 20 minutes (lane k), and decreases at 40 minutes with the transcript size becoming somewhat heterogeneous (lane l). The most likely explanation is that the construct B transcript is not as stable in the extract as the wild-type EH4 transcript. The earlier time points, however, clearly show that the level of transcriptional initiation is not decreased by removing the sequences downstream of +50. Construct A is transcribed to a much lower level than construct B over the whole time course (lanes e−h). The accumulation curves of the two reactions, however, are parallel, indicating that the stabilities of the transcripts are comparable. In the pulse-chase experiment (Figure 4D), an excess of unlabeled UTP was added 20 minutes after the start of the reaction. As can be seen, at 10 and 20 minutes after the addition of unlabeled nucleotide, there is a decrease in reaction product in all cases. As expected from the continuous labeling experiment, the products of constructs A and B are less stable than those of the wild-type gene. The key point, however, is that they are equally unstable. The decrease in the level of transcription due to removal of the sequence between +50 and +19 does not, therefore, appear to be due to a difference in RNA stability. Rather, the presence of this sequence is necessary to achieve efficient transcription at the EH4 promoter.

Effects of double mutants and deletions

We were interested in whether UHF-1 and the presumed factors which had effects via the H4SE and initiator sites acted additively or interdependently to stimulate transcription. By analogy to the human H4 gene, we would expect that a factor will be found that binds to the sea urchin H4SE (the name UHF-2 is reserved for this protein, if identified). There is no direct evidence for such a factor in the sea urchin at present, and there is also no evidence for a protein which binds at the initiator region. Nevertheless, to see if there was interaction between these sites, double mutants (ΔUHF-1/H4SE-2 or ΔUHF-1/INR-1) were constructed and their transcription compared with the individual mutants. Figure 5 shows the results of such an experiment. The values directly below each gel are the relative transcription efficiencies for that particular experiment. As can be seen, the single ΔUHF-1 (Δ−135/−102) and H4SE-2 mutants are transcribed at 28% and 11% of the wild-type template level, respectively (lanes b and d). The construct carrying the ΔUHF-1/H4SE-2 double mutation is transcribed at a much lower level, 2.7% (lane e). As noted in Table 2 (Experiment I), if the effect of the two mutations were completely independent, introduction of the H4SE-2 mutation on the ΔUHF-1 template should decrease the transcriptional level from 28% to 3.1% of that of wild-type (0.28×0.11 = 0.031). The actual value, 2.7%, is very close to the value expected for independent effects. Similar findings were made in four additional experiments (Experiments II−V of Table 2). These results indicate that UHF-1 does not

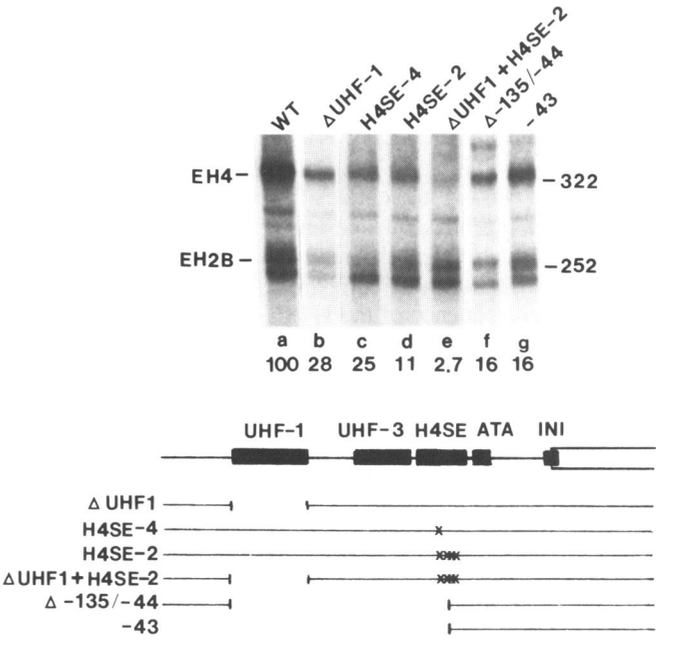

Figure 5. The effect of double mutants and deletions on EH4 transcription. Transcription of a series of mutant templates was compared with the wild-type (−436) EH4 template as in Figures 1−3. The percentages of wild-type transcription obtained with each mutant EH4 gene is shown below each lane. At the bottom of the Figure is a schematic presentation of the various mutants used in the experiment.

Table 2.

Experiment	ΔUHF-1[a]	H4SE-2[a] or INR-1	Double mutant[a]	Expected[b]
ΔUSE-1 + H4SE-2				
I	0.28	0.11	0.027	0.031
II	0.19	0.15	0.045	0.029
III	0.27	0.15	0.025	0.041
IV	0.27	0.14	0.039	0.038
V	0.38	0.16	0.044	0.061
ΔUSE-1 + INR-1				
II	0.19	0.039	0.006	0.0074
V	0.38	0.071	0.018	0.027

[a] Amount of transcript for single ΔUHF-1 (Δ-135/−103), H4SE-2, or INR-1 mutants or double mutants containing the ΔUHF-1 and one of the other two changes, expressed as a fraction of wild-type EH4 transcription. Each roman numeral represents an independent experiment done with a different extract. Experiment I is shown in Figure 5.

[b] The values shown are the products of the fractions listed in the single mutant columns. The values represent the amount of transcript relative to wild-type template which would be obtained if each of the two mutants tested had completely independent effects.

require a completely functional H4SE site for its positive effect on transcription and similarly, the H4SE site (and whatever factor may interact with this site) does not require the UHF-1 site for its positive effect. If the sequence elements required one another for a postive transcriptional effect, the double mutant would have been transcribed at the same level as the less active of the single mutants, in this case the H4SE mutant. The results also indicate that the UHF-1 and H4SE sites do not carry out a redundant function. If this were the case, the activity of the double mutant template would have been considerably lower than was observed.

An identical situation was found when transcription of the ΔUHF-1/INR-1 double mutant was compared with each of the two individual mutants. Table 2 presents two separate experiments in which the level of transcription of the double mutant is very close to that expected from independent effects of the individual mutations. In one experiment the double mutant is transcribed at 0.6% of the wild-type template and the expected value for independent effects was 0.74%. In the other experiment the observed value was 1.8% and the expected value, 2.7%.

Templates lacking nucleotides from −135 to −44 (Δ−135/−44) or lacking all sequence upstream of −43 ('−43') are transcribed much more efficiently than the ΔUHF-1/H4SE-2 double mutants (Figure 5, compare lanes e, f, and g). Since the most essential sequence of the H4SE is upstream of −43 (e.g, mutants H4SE-4 and −5 are almost as strong as H4SE-2), this result is at first glance unexpected. However, the UHF-3 site between −75 and −55 is also missing in these particular deletion templates and when this is considered, the increase in level of transcription can be explained by the removal of a negatively acting transcription element. Nevertheless, the increase relative to the UHF-1/H4SE double mutants is actually greater than expected if the negative effect of UHF-3 was independent of the positive effects of UHF-1 and the H4SE. The double mutants are transcribed at an average of 3.4% of wild-type level, whereas the −43 and Δ-135/−44 mutants are transcribed at an average of 17−18% of wild-type level (Table 1), a 5−6 fold difference. Since disruption of the UHF-3 site itself only results in an average increase of 40% over wild-type levels, one might have expected much lower transcription of the −43 and Δ-135/−44 mutants than was actually obtained.

DISCUSSION

The nuclear extract provides a reliable *in vitro* assay for transcriptional elements

We have previously shown that the nuclear extract is effective in the assay of the UHF-1 site in the EH4 and LH4 gene promoters (Lee et al., submitted for publication). Moreover, the average three-fold decrease in *in vitro* transcription observed with templates lacking the UHF-1 site is very close to the decrease in gene expression obtained when the same constructs are injected into eggs. Preliminary findings with *in vivo* expression of the *Bal*31 series of constructs used for the experiment of Figure 1, also confirm the close parallel between the effects of deletion on *in vitro* transcription and *in vivo* expression, in the region downstream of −135 (Bumcrot and Weinberg, unpublished results). However, additional *cis*-acting sequences located upstream of −135, mediating both positive and negative responses, and responsible for the temporal regulation of the gene, can only be assayed *in vivo* by expression of injected DNA (Bumcrot and Weinberg, unpublished results). The data presented above indicate that the sea urchin embryo nuclear extract is useful in assaying both positive and negative sequence elements involved in transcription, if they are reasonably close to the site of transcriptional initiation. The system has permitted us to identify important regions upstream and downstream of the start site as well as at the initiation site itself.

Individual experiments readily provide information on the relative strength of templates carrying different mutations in a particular sequence element. For example, the H4SE-2 mutant was always transcribed at a lower level than H4SE-4 and −5. The level of expression, however, does vary from extract to extract. This was observed most often with the ATA and H4SE-1 and −3 mutants, and the +19 3′ deletion. Nevertheless, the average values obtained in multiple experiments are revealing for most of the mutant templates, as shown in Table 1. The effects of some of the mutants are extremely reproducible from extract to extract, as can be deduced from the low standard errors. The internal EH2B standard is essential for the analysis; without the ability to normalize in each reaction, the results would be less clear. The extract has also been used to perform competition experiments in which *cis*-acting elements present on two different templates can be identified (40). The EH4 and LH4 genes compete for one another in the formation of a stable transcription complex, and the region of the EH4 promoter involved in binding a common factor was shown to be between −436 and −102. Subsequently, we identified the UHF-1 site as the common sequence element (Lee et al., manuscript submitted).

The H4 specific element (H4SE)

The results presented here on transcription of templates carrying mutations in different regions of the H4SE are the first report of the effects of specific sequence changes in the element on H4 gene expression. Previous studies with *Xenopus* (23) or human (33) H4 genes carrying complete or partial deletions of the H4SE, did indicate that the element had an important role in transcription. Although base substitution and insertion mutations in the corresponding region of a human H4 gene did not have an effect on transcription of the gene in a standard HeLa nuclear extract, addition to the extract of H4TF-2, a 65 kD protein which footprints the sequence, stimulated expression of the wild-type (but not mutant) gene templates (34). In the data presented above, we showed the most critical region of the EH4 H4SE for transcription was the GT sequence of the proximal domain, GTCCGCAA. Changing the first G (at −46) to a C (H4SE-4) or a change in the adjacent TCC to a CTA (H4SE-5) both resulted in an average four-fold decrease in template activity. The greatest decrease (6−7-fold) was obtained with a 4 base substitution at positions −46 and −44 to −42 (H4SE-2). In contrast, less than 50% reduction (and sometimes no decrease at all) was obtained with mutants containing 6 base subtitutions from −44 to −39 (H4SE-1) and from −38 to −33 (H4SE-3). An average two-fold decrease was observed with mutant H4SE-6 which had substitutions of five bases at −52 and −50 to −47. These results indicate that the most conserved sequence of the H4SE, GTCC, is essential for activity of the element, but that regions upstream and downstream of this sequence, corresponding to H4SE domains conserved in sea urchin EH4 genes, have a role in the function of the element.

The negatively responding UHF-3 site

We have presented several lines of evidence indicating the existence of a negative transcription element between −62 and −73 of the *S. purpuratus* EH4 gene. First, a template lacking all but 55 bp upstream of the transcriptional initiation site is transcribed at a higher level than a template with 102 5′ flanking

nucleotides. Second, two different templates, each containing 6 bp changes in the footprinted region between −76 and −56, are transcribed more efficiently than the wild-type gene. Third, deletions which either remove all sequence upstream of −43 or which extend from −135 to −44 (thus removing the positively acting UHF-1 site and the most crucial nucleotides of the H4SE site) are transcribed at a level 5−6-fold higher than that obtained with a UHF-1/H4SE double mutant. A fourth piece of evidence is that a UHF-1/UHF-3 double mutant is more active than the UHF-1 mutant (see Table 1). Lastly, the −55 construct is expressed in embryos at a higher level than the −102 construct, after injection of the DNAs into eggs (Bumcrot and Weinberg, in preparation). With the exception of the data involving the Δ−135/−44 and −43 deletions, the increases observed are less than two-fold. Nevertheless, we have demonstrated that the effects of UHF-3 mutations are reproducible and completely consistent. Although the degree of increase varies from extract to extract (resulting in a fairly large standard error when results using different extracts are averaged, see Table 1), all templates containing deletions of, or lesions in, the UHF-3 site are transcribed to a higher level than the appropriate control templates. The Δ−135/−44 and −43 deletions, however, are transcribed at a higher level than would be expected from the independent effects of removal of the UHF-1, H4SE, and UHF-3 sites. Since we have shown that the UHF-1 and H4SE sites have independent effects, we are left with a number of possibilities to explain the high level of expression of the deletions: a) a deletion of the UHF-3 site might be much more effective in abolishing activity of the element than either of the base substitution mutants tested, b) there might be an additional negative site between −135 and −44, so that its removal along with the UHF-3 site is more profound than mutation of the UHF-3 site alone, c) there could be interference between factor binding at the UHF-3 and H4SE sites, or d) the deletions to −43 may not completely obliterate the function of the H4SE (although this is unlikely from the results of transcription of the H4SE base substitution mutants).

The ATA region

Sea urchin EH4 genes have an unusual sequence at the site usually occupied by a TATA box. The *S. purpuratus* EH4 gene (46,47) and the closely related EH4 gene of clone h19 from *Psammechinus miliaris* (22) have TAACAATAC at this position. The h22 clone of *P. miliaris* (48) shares all but the first T of this sequence, and the EH4 gene of *Paracentrotus lividus* (reported in ref. 23) contains the ACAATA core. We refer to this site as the ATA region, since the complete TATA motif does not appear in any of these genes. Since any part of the sequence noted could have the role of a TATA box, we created three mutations at different parts of the site. The effect of these mutations on EH4 gene transcription was shown here to be minimal. We observed no more than a 50% reduction in transcription in the mutants ATA-1, −2, or −3, and in a number of experiments, no decrease was detected. In none of the experiments with the mutant templates was there any indication of the utilization of alternative initiation sites. In itself, this result would not strongly support the existence of the equivalent of a TATA site in the EH4 gene. However, recent work has shown that *in vitro* transcription and *in vivo* expression of an EH4 gene lacking the UHF-1 site are completely dependent on the presence of the TAACAATA sequence (Tung, Bumcrot, and Weinberg, unpublished experiments). The site is therefore cryptic, and dramatic effects of its mutation are seen only when some other part of the transcription apparatus is made limiting.

The initiator and internal sequence element

We presented evidence that the sequence at the transcriptional start site of the EH4 gene was essential for proper expression of the EH4 gene. The average level of transcription of the INR-1 template was only 4% that of the wild-type template. The transcriptional initiation site of most histone genes is found within a highly conserved 5′-PyCATTCPu-3′ sequence, with the transcripts beginning at the A (1,42,49). There is surprisingly sparse evidence for the functional importance of the histone start consensus sequence. In the case of a sea urchin H2A gene tested for activity by injection into *Xenopus* oocytes, deletion of this region and a portion of the gene encoding most of the mRNA leader was found to result in a somewhat decreased level of expression (50). Deletion of the region containing the initiation site of the human pHu4A H4 gene resulted in a much stronger decrease in transcription, as assayed *in vitro* in a HeLa nuclear extract (38). In both cases, the transcripts derived from the deleted templates were initiated at new start sites, located at about the same distance from the TATA box as in the intact promoters. In the experiment performed here, we changed only 4 bp within the consensus sequence, but the 5′ A terminus was not altered. The sharp decrease in transcription cannot, therefore, be due to a sequence further downstream of the site of initiation (as might be the case in the deletion experiments with the sea urchin H2A and human H4 genes). We did not detect transcripts with alternative 5′ termini, but the resolution of the run-off products on the gel would not have been sufficient to distinguish transcripts differing in size by only several bases.

Base changes and deletions near the transcriptional start site of a number of genes resulted in a decrease in transcription or the use of new initiation sites (51−60) whereas no effect on promoter strength is found in other cases. In yeast genes transcribed by polymerase II, the sequences around the transcriptional start site are important and this region has been termed an 'initiator' (61). This term has been recently used to describe an 8 bp stretch of sequence overlapping the start site of the terminal deoxynucleotidyltransferase (TdT) gene which is sufficient to support accurate basal transcription of this gene (56,60). We note that a sequence CCTCATTC from −4 to +4 of the TdT gene differs at only one base from the sea urchin EH4 gene sequence at the same position. Specific base changes within this sequence of the TdT gene result in a decrease in transcription, and the initiator can be activated by either an upstream TATA box or by SV40 21 bp repeats (56,60). It is quite possible that the sea urchin EH4 gene utilizes the same initiator to support basal transcription. It is interesting that the −7 deletion mutant (Figure 1, lane h) appears to support a very low level of EH4 transcription, which would imply that the EH4 initiator region can suffice as a basal transcriptional element.

Maximal transcription of the EH4 gene also appears to require the presence of a sequence within the region coding for the 5′ untranscribed leader. The magnitude of the loss of expression due to elimination of the sequence between +19 and +50 varies somewhat from extract to extract (see Figure 4), but we have shown that the effect is not due to loss of stability of the transcript in the nuclear extract. There are two other reports of histone gene transcriptional elements located within the transcribed gene unit. Hurt et al. (62) found that a 110-nucleotide region located within the coding region of a murine H3 gene was required for high expression, and Mous et al. (63) showed that a sea urchin H2B

gene required a region in the untranslated leader and a sequence located near the 3′ end for maximal expression (the latter region may be downstream of the 3′ mRNA terminus). A number of other genes, including the Drosophila *Ultrabithorax* (64), *Antennapedia* (65), and *engrailed* (66) genes, the human glial acidic protein gene (57), and the LTRs of HIV-1 and HIV-2 (54) also have positive sequence elements in their untranslated leaders with some sequence homology to the sea urchin EH4 gene between +19 and +50. A number of mouse ribosomal protein genes (67), the Drosophila hsp22 gene (68), and the sea urchin Spec genes (69) also have transcriptional sequence elements in the untranslated leader region, but sequence homology with the EH4 gene is not evident. Our experiments do not address whether the sequence between the initiator and +19 is important for transcription. Further work is necessary to determine if the initiator and the required internal sequence are distinct transcriptional elements.

ACKNOWLEDGMENTS

This investigation was supported by National Institutes of Health grant RO1-GM27322 (to E. W.). The authors are grateful to Jau-Nian Chen and Paul Chamberlin for the preparation of a number of the mutant templates used in this study.

REFERENCES

1. Hentschel,C.C. and Birnstiel,M.L. (1981) Cell **25**, 301–313.
2. Maxson,R.E., Mohun,T.J., Cohn,R.J., and Kedes,L. (1983) Ann. Rev. Genet. **17**, 239–277.
3. Childs,G.C., Nocente-McGrath,C., Lieber,T., Holt,C., and Knowles,J. (1982) Cell **31**, 383–393.
4. Maxson,R.E., Mohun,T.J., Childs,G. and Kedes,L.H. (1983) Nature **301**, 120–125.
5. Kaumeyer,J.F. and Weinberg,E.S. (1986) Nucleic Acids Res. **14**, 4557–4576.
6. Kemler,I. and Busslinger,M. (1986) Mol. Cell. Biol. **6**, 3746–3754.
7. Knowles,J.A. and Childs,G. (1986) Nucleic Acids Res. **14**, 8121–8133.
8. Knowles,J.A., Lai,Z., and Childs,G. (1987) Mol.Cell. Biol. **7**, 478–485.
9. Lai,Z.-C. and Childs,G. (1988) Mol. Cell. Biol. **8**, 1842–1844.
10. Mauron,A., Kedes,L., Hough-Evans,B.R., and Davidson,E.H. (1982) Dev. Biol. **94**, 425–434.
11. Maxson,R.E., and Wilt,F.H. (1982) Dev. Biol. **94**, 435–440.
12. Weinberg,E.S., Hendricks,M.B., Hemminki,K., Kuwabara,P.E., and Farrelly,L.A. (1983) Dev. Biol. **98**, 117–129.
13. Grunstein, M. (1978) Proc. Natl. Acad. Sci. USA. **75**, 4135–4139.
14. Hieter, P.A., Hendricks,M.B., Hemminki,K., and Weinberg,E.S. (1979) Biochemistry **18**, 2707–2715.
15. Childs,G., Maxson,R. and Kedes,L.H. (1979) Dev. Biol. **73**, 153–173.
16. Knowles,J.A. and Childs,G. (1984) Proc. Natl. Acad. Sci. USA. **81**, 2411–2415.
17. Mohun,T.J., Maxson,R., Gormezano,G., Kedes,L. (1985) Dev. Biol. **108**, 491–502.
18. Busslinger,M. and Barberis,A. (1985) Proc. Natl. Acad. Sci. USA **82**, 5676–5680.
19. Uzman, J.A. and Wilt,F.H. (1984) Dev. Biol. **106**, 174–180.
20. Roberts,S.B., Weisser,K.E., and Childs,G. (1984) J. Mol. Biol. **174**, 647–662.
21. Maxson,R., Mohun,T., Gormezano,G., and Kedes,L. (1987) Nucleic Acids Res. **15**, 10569–10582.
22. Busslinger,M., Portmann,R., Irmiger,J.C., and Birnstiel,M.L. (1980) Nucleic Acids Res. **8**, 957–977.
23. Clerc, R.G., Bucher,P., Strub,K., and Birnstiel,M.L. (1983) Nucleic Acids Res. **11**, 8641–8657.
24. Sugarman,B.J., Dodgson,J.B., and Engel,J.D. (1983) J. Biol. Chem. **258**, 9005–9016.
25. Zhong,R., Roeder,R.G., and Heintz,N. (1983) Nucleic Acids Res. **11**, 7409–7425.
26. Sierra,F., Stein,G., and Stein,J. (1983) Nucleic Acids Res. **11**, 7069–7086.
27. Wang,S.W., Robbins,A.J., d'Andrea,R., and Wells,J.R.E. (1985) Nucleic Acids Res. **13**, 1369–1387.
28. Perry,M., Thomsen,G.H., and Roeder,R.G. (1985) J. Mol. Biol. **185**, 479–499.
29. Winkfein,R.J., Conner,W., Mezquita,J., and Dixon,G.D. (1985) J. Mol. Evol. **22**, 1–19.
30. Wells,D.E. (1986) Nucleic Acids Res. **14**, Supplement, r119–r149
31. Howell,A.M., Cool,D., Hewitt,J., Ydenberg,B., Smith,M.J., and Honda,B.M. (1987) J. Mol. Evol. **25**, 29–36.
32. Lai,Z.-C., Maxson,R., and Childs,G. (1988) Genes Dev. **2**, 173–183.
33. Kroeger,P., Stewart,C., Schaap,T., van Wijnen,A., Hirshman,J., Helms,S., Stein,G., and Stein,J. (1987) Proc. Natl. Acad. Sci. USA **84**, 3982–3986.
34. Dailey, L., Roberts,S.B., and Heintz,N. (1988) Genes Dev. **2**, 1700–1712.
35. Dailey, L., Hanly,S.M., Roeder,R.G., and Heintz,N. (1986) Proc. Natl. Acad. Sci. USA. **83**, 7241–7245.
36. van Wijnen,A.J., Stein,J.L., and Stein,G.S. (1987) Nucleic Acids Res. **15**, 1679–1698.
37. van Wijnen,A.J., Wright,K.L., Lian,J.B., Stein,J.L., and Stein,G.S. (1989) J. Biol. Chem. **264**, 15034–15042.
38. Hanly,S.M., Bleecker,G.C., and Heintz,N. (1985) Mol. Cell. Biol. **5**, 380–389.
39. Morris,G.F., Price,D.H., and Marzluff,W.F. (1986) Proc. Natl. Acad. Sci. USA **83**, 3674–3678.
40. Tung,L., Morris,G.F., Yager,L.N., and Weinberg,E.S. (1989) Mol. Cell. Biol. **9**, 1476–1487.
41. Morris,G.F. and Marzluff,W.F. (1983) Biochemistry **22**, 645–653.
42. Sures,I., Levy,S., and Kedes,L.H. (1980) Proc. Natl. Acad. Sci. USA **77**, 1265–1269.
43. Overton,G.C. and Weinberg,E.S. (1978) Cell **14**, 247–257.
44. Vieira,J. and Messing,J. (1987) Methods. Enzymol. **153**, 3–11.
45. Kunkel,T.A., Roberts,J.D., and R.A.Zakour. (1987) Methods Enzymol. **154**, 367–383.
46. Grunstein,M., Diamond,K.E., Knoppel,E., and Grunstein,J.E. (1981) Biochemistry **20**, 1216–1223.
47. Yager,L.N., Kaumeyer,J.F., and Weinberg,E.S. (1984) J. Mol. Evol. **20**, 215–226.
48. Schaffner,W., Kunz,G., Daetwyler,H., Telfore,J., Smith,H.O., and Birnstiel,M.L. (1978) Cell **14**, 655–617.
49. Hentschel,C., Irminger,J.-C., Bucher,P., and Birnstiel,M.L. (1980) Nature **285**, 147–151.
50. Grosschedl,R. and Birnstiel,M.L. (1980) Proc. Natl. Acad. Sci. USA **77**, 1432–1436.
51. Talkington,C.A. and Leder,P.A. (1982) Nature **298**, 192–195.
52. Concino,M.F., Lee,R.F., Merryweather,J.P., and Weinmann,R. (1984) Nucleic Acids Res. **12**, 7423–7433.
53. Tokunaga,K., Hirose,S., and Suzuki,Y. (1984) Nucleic Acids Res. **12**, 1543–1558.
54. Jones,K.A., Luciw,P.A., and Duchange,N. (1988) Genes Dev. **2**, 1101–1114.
55. Tebb,G. and Mattaj,I.W. (1988) EMBO J. **7**, 3785–3792.
56. Smale,S.T. and Baltimore,D. (1989) Cell **57**, 103–113.
57. Nakatani,Y., Brenner,M., and Freese,E. (1990) Proc. Natl. Acad. Sci. USA **87**, 4289–4293.
58. Hariharan,N. and Perry,R.P. (1990) Proc. Natl. Acad. Sci. USA **87**, 1526–1530.
59. Carcamo,J., Maldonado,E., Cortes,P., Ahn,M.-H., Ha,I., Kasai,Y., Flint,J., and Reinberg,D. (1990) Genes Dev. **4**, 1611–1622.
60. Smale,S.T., Schmidt,M.C., Berk,A.J., and Baltimore,D. (1990) Proc. Natl. Acad. Sci. USA **87**, 4509–4513.
61. Chen,W., and Struhl,K. (1985) EMBO J. **5**, 3273–3280.
62. Hurt,M.M., Pandey,N.B., and Marzluff,W.F. (1989) Proc. Natl. Acad. Sci. USA **86**, 4450–4454.
63. Mous,J., Stunnenberg,H., Georgiev,O., and Birnstiel,M.L. (1985) Mol. Cell. Biol. **5**, 2764–2769.
64. Biggin,M.D. and Tjian,R. (1988) Cell **53**, 699–711.
65. Perkins,K.K., Dailey,G.M., and Tjian,R. (1988) Genes Dev. **2**, 1615–1626.
66. Soeller,W.C., Poole,S.J., and Kornberg,T. (1988) Genes Dev. **2**, 68–81.
67. Hariharan,N., Kelley,D., and Perry,R.P. (1989) Genes Dev. **3**, 1789–1800.
68. Hultmark,D., Klemenz,R., and Gehring,W.J. (1986) Cell **44**, 429–438.
69. Tomlinson,C.R., Kozlowski,M.T., and Klein,W.H. (1990) Development (in press).