Nucleic Acids Research, Vol. 20, No. 16 4311–4318

Multiple octamer binding sites in the promoter region of the bovine αs2-casein gene

Martien A.M.Groenen, Rosilde J.M.Dijkhof, Jan J.van der Poel, Ria van Diggelen and Esther Verstege
Department of Animal Breeding, Agricultural University Wageningen, PO Box 338, 6700 AH Wageningen, The Netherlands

Received April 15, 1992; Revised and Accepted July 20, 1992

ABSTRACT

Using a set of overlapping oligonucleotides from the promoter region of the bovine αs2-casein gene we have identified two nuclear factors which probably are involved in expression of this gene and the related calcium sensitive αs1- and β-casein genes. One of these factors which was present in extracts of all tissues that have been tested including Hela cells turned out to be the octamer binding protein OCT-1. Oct-1 binds with different affinity to 4 sites at positions centred around −480, −260, −210 and −50. The strongest of these 4 binding sites, the one around position −50, is highly conserved in all calcium sensitive caseins of mouse, rat, rabbit and cattle. The other nuclear factor (MGF, mammary gland factor) which is specifically expressed in the mammary gland, binds to a site around position −90. This binding site is also highly conserved in all calcium sensitive caseins of mouse, rat, rabbit and cattle.

INTRODUCTION

In eukaryotes initiation of transcription by RNA polymerase II is dependent on the interaction of multiple transcription factors with their binding sites within the promoter and enhancer (1,2). These factors are believed to increase the formation of the preinitiation complex (3, 4) by interactions with general transcription factors such as TFIID (5, 6) and TFIIB (7) and probably with the RNA polymerase itself (8). Often both tissue specific and ubiquitous factors act in a consorted fashion to activate a certain gene. It is believed that the presence of activators or the absence of repressors/ silencing factors is a prerequisite for the expression of developmentally regulated genes.

One group of genes that is highly stage- and tissue specific is that of the milk protein genes. In cattle this group is comprised of the related calcium sensitive αs1-, αs2- and β-casein genes, the kappa-casein gene and the two genes encoding the whey proteins α-lactalbumin and β-lactoglobulin (for a review see reference 9). Expression of these genes is preceded by differentiation of the epithelial cells in the developing mammary gland and is strongly influenced by interactions of the milk protein expressing epithelial cells with the extracelluar matrix and the presence of several hormones such as prolactin, insulin and glucocorticoids (10, 11). Of these genes the four caseins are especially interesting since they are clustered within 200 Kbp on chromosome 6 (12, 13) and might be concurrently regulated by a locus control region. Furthermore greater than 80% of the milk protein is casein and since the caseins are the raw material for cheese production, these genes are of considerable economical importance.

Little is known about the transcription factors that regulate the expression of these genes. Of the caseins, the best studied example is the rat β-casein gene. In transgenic animals, a transgene containing nucleotides −524 to +490 from the rat β-casein gene, was correctly tissue-specific and developmentally regulated (14). In addition regions that probably contain elements involved in positive and negative regulation and hormone induction have been identified in rat and mouse β-casein genes using the mammary epithelial cell line HC11 (15, 16) or primary mammary epithelial cells (17). Recently it has been shown (16) that a mammary gland specific transcription factor (MGF) binds to the conserved sequences at −90 in the promoter region of the rat β-casein gene. A different putative mammary gland specific transcription factor (MPBF) has been found to bind to the promoter of the ovine β-lactoglobulin gene (18). Furthermore, in the same study, several binding sites for NF-1 were also identified within the promoter region of the ovine β-lactoglobulin gene.

We started to analyze the regulation of transcription of the casein genes, and have examined the binding of nuclear factors to the promoter region of the bovine αs2-casein gene, using electrophoretic mobility shift assays.

MATERIALS AND METHODS

Nuclear extracts and electrophoretic mobility shift assay

Nuclear extracts from mammary gland and liver were prepared essentially as described by Dignam (19) with minor modifications according to Hennighausen and Lubon (20). The buffers used in preparing the nuclei contained 0.5 mM phenylmethylsulfonyl fluoride and 2 μg/ml each of antipain, leupeptin and pepstatin A (sigma). Nuclear extract from HeLa cells was purchased from BRL/Life technologies inc. Tissue of mammary gland and liver was obtained from a local slaughter house. Immediately after the animals were killed, the tissues were put on ice, divided into several 25 g portions, and quickly frozen on liquid nitrogen.

For the binding studies, 4−6 μg of nuclear extract was combined with 1−3 μg poly(dI-dC) in 17 μl of a buffer containing 20 mM HEPES (pH 7.5), 75 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, 0.1% Triton X-100, 1 mM DTT and incubated at room temperature for 15 minutes. In competition experiments the competitor was added together with the poly (dI-dC). After adding 0.5 ng end labelled probe and incubating for 25 minutes at room temperature, the samples were separated on a 5% polyacrylamide gel using 0.5 times TBE running buffer, which was constantly recirculated. Supershift experiments with antibodies specific for the DNA binding domain of Oct-1 were done on ice. The Oct-1 antiserum was a gift from Peter Verrijzer.

Protein concentration of extracts were determined using the BIO-RAD protein assay. Quantification of the percentage residual binding to the probe in competition experiments was done by densitometric scanning using the Cybertech CS-1 image documentation system and image processing software.

Oligonucleotides used in the binding assay

Oligonucleotides were synthesized on a Pharmacia Gene Assembler. Oligonucleotides containing the binding sites for NF1, AP1, AP2, AP3 and SP1 were purchased from Stratagene and those for CREB and OCT-1 from Promega. Oligonucleotides used as probes were endlabled by phosphorylation of their 5′ ends with γ^{32}P-ATP and kinase. The oligonucleotides of the 5′ flanking region of the bovine αs2-casein gene, which were used in this study, are indicated in figure 1. The sequences of the upper strand of the other oligonucleotides used are shown below.

-50 CONS	5′-AAAATTAGCATGCCATTAA-3′
-50 MUT	5′-AAAATTAGAATGCCATTAA-3′
MGF-MUT	5′-CAGAGTTGTTACAATTCAAA-3′
O6	5′-CAGACTTGCATATGGCTA-3′
O5	5′-TTAAATAAGCATCTAAAATT-3′
AP1	5′-CTAGTGATGAGTCAGCCGGATC-3′
AP2	5′-GATCGAACTGACCGCCCGCGGCCCGT-3′
AP3	5′-CTAGTGGGACTTTCCACAGATC-3′
NF1	5′-ATTTTGGCTTGAAGCCAATATG-3′
SP1	5′-GATCGATCGGGGCGGGGCGATC-3′
SRE	5′-TTGGATGTCCATATTAGGACATCT-3′
GRE	5′-TTCGGTACAATGTGTTCTAGC-3′
CREB	5′-AGAGATTGCCTGACGTCAGAGAGCTAG-3′
OCT-1	5′-TGTCGAATGCAAATCACTAGAA-3′

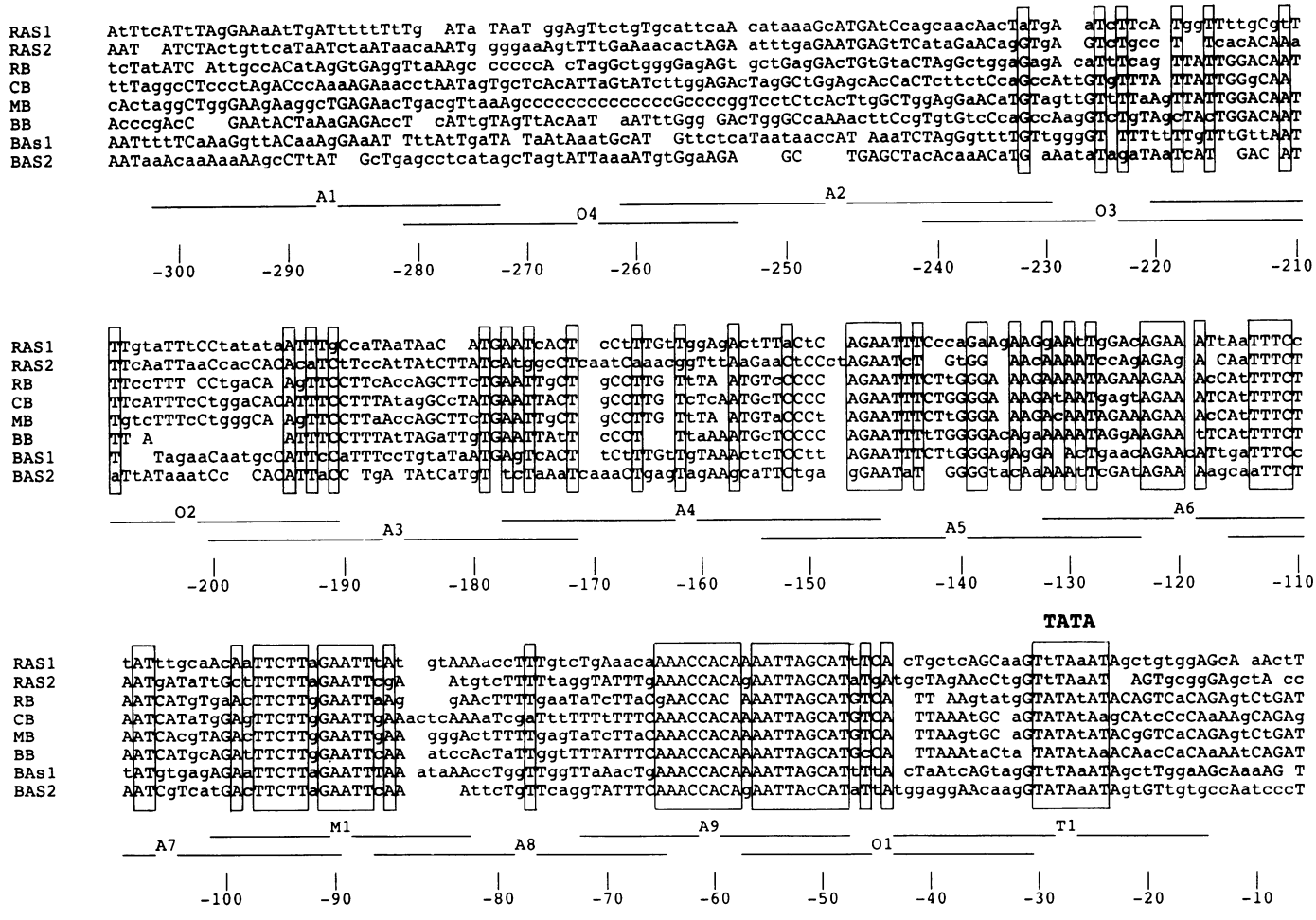

Figure 1. Alignment of the promoter region of 8 calcium sensitive caseins from 4 different organisms. The sequence of the bovine αs2-casein gene is shown at the bottom. The nucleotides of the promoter region of the bovine αs2-casein gene are numbered. The 15 overlapping oligonucleotides spanning the region from −305 to −15 that were used in this study are indicated by horizontal bars. BAS1 = Bovine αS1-casein (21, 22), BAS2 = Bovine αS2-casein (23), BB = Bovine β-casein (24), CB = Rabbit β-casein (25), MB = Mouse β-casein (26), RAS1 = Rat αS1-casein (21),RAS2 = Rat αS2-casein (27), RB = Rat β-casein (28). Capitals are used for nucleotides that are present at that position in at least 5 of the 8 sequences. Nucleotides that are conserved in at least 7 of the 8 sequences are boxed. The TATA box is also boxed.

RESULTS

To examine the interaction of transcription factors with the promoter region of the bovine αs2-casein gene, we performed gel mobility shift assays with a set of overlapping oligonucleotides spanning the region from −305 to −15. This region contains several stretches of nucleotides which are highly conserved in all calcium sensitive caseins sequenced thus far (9, 21) especially in the region up to −170 (see figure 1). The 15 oligonucleotides spanning this region of the bovine αs2-casein gene which were used in this study are also shown in figure 1.

With oligonucleotides O1 and O4 specific complexes were seen with all extracts tested, whereas with oligonucleotide M1 specific complexes could only be clearly detected with extracts made from the mammary gland (figure 2). Although extracts made from the mammary gland of a lactating cow contained higher concentrations of the protein which binds to M1, it was also detected in extracts made from the mammary gland of a non-lactating 1½ year old animal. In addition to the prominent complex detected with oligonucleotide O1 (marked C in figure 2) several other minor bands (marked X in figure 2) were also seen with all extracts. The identity of these bands is not known. With some of the other oligonucleotides (e.g. O2, O3 and A7) very faint bands were seen but it was not clear whether these represented true complexes (not shown).

Binding of the mammary gland specific nuclear factor MGF

Oligonucleotide M1 contains the sequence ANTTCTTRGAATTNA (where N is any nucleotide and R is a purine), which is extremely conserved in all calcium sensitive caseins at position −90. Binding of a mammary gland specific nuclear factor to this sequence has recently been shown for the rat β-casein gene (16). If this sequence is mutated at two positions to the sequence ANTTGTTRCAATTNA, it no longer competes for binding with M1 (figure 3). Similarly a complex no longer is formed if this oligonucleotide is used as a probe (figure 3). Although this factor was only present at a high concentration in nuclear extracts from the mammary gland of a lactating animal, it was also present in nuclear extracts made from the mammary gland of a 1½ year old non-lactating animal (figure 2). In this perspective it is interesting that small amounts of the mRNAs for caseins can be

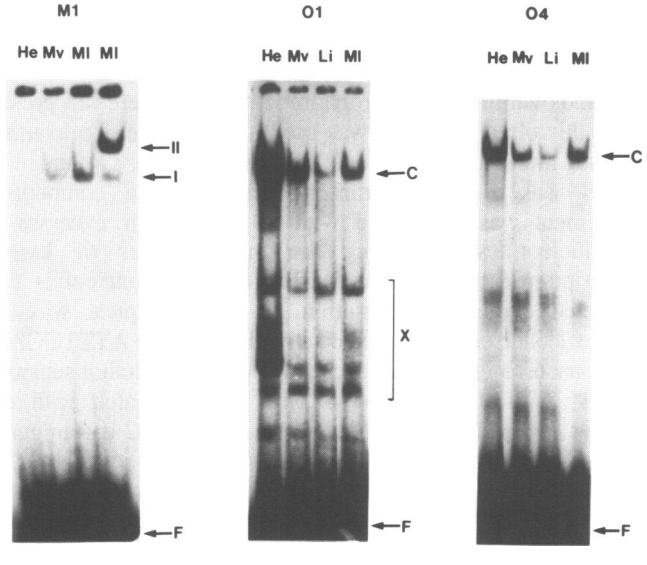

Figure 2. Gel mobility shift assay with different nuclear extracts and oligonucleotides M1, O1 and O4 as the probe. The two different complexes formed with the nuclear extract from the mammary gland on probe M1 are indicated (I and II). F=free probe, c=specific complex formed with all extracts on probe O1 and O4, x=other unidentified bands found with probe O1. Extracts used are: (He) Hela, (Mv) mammary gland of a 1½ year old non lactating animal, (Ml) mammary gland of a lactating animal, (Li) liver.

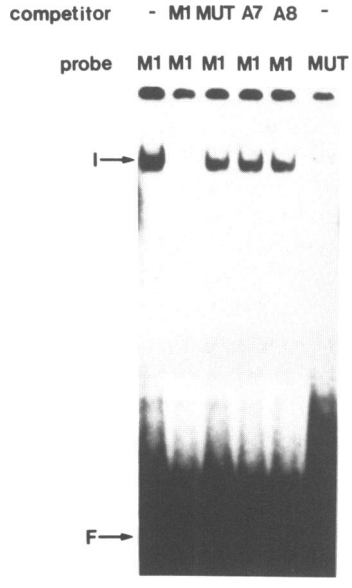

Figure 3. Gel mobility shift assay competition experiment. The probe used was M1 and the nuclear extract used was from the mammary gland of a lactating cow. The competitors used are indicated above each lane, and were added at a 50 fold molar excess before adding the probe.

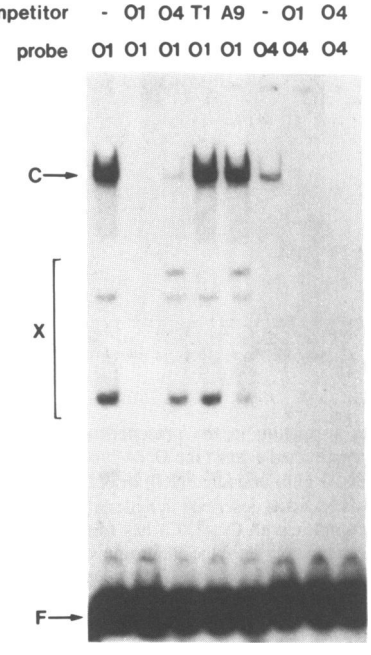

Figure 4. Gel mobility shift assay competition experiment. The probes and competitors used are indicated above each lane. The competitors were used at a 50 fold molar excess and the nuclear extract used was made from Hela cells. c=specific complex, F=free probe, x=other unidentified bands.

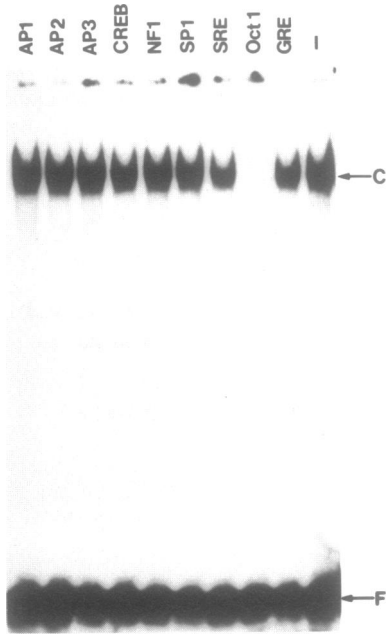

Figure 5. Gel mobility shift assay competition experiment to identify the binding site on oligonucleotides O1 and O4. Oligonucleotide −50 CONS which binds the same protein as O1 and O4 was used as the probe. The competitors used are indicated above each lane and were used at a 50 fold molar excess. The nuclear extract used was made from Hela cells. F=free probe, c=specific complex.

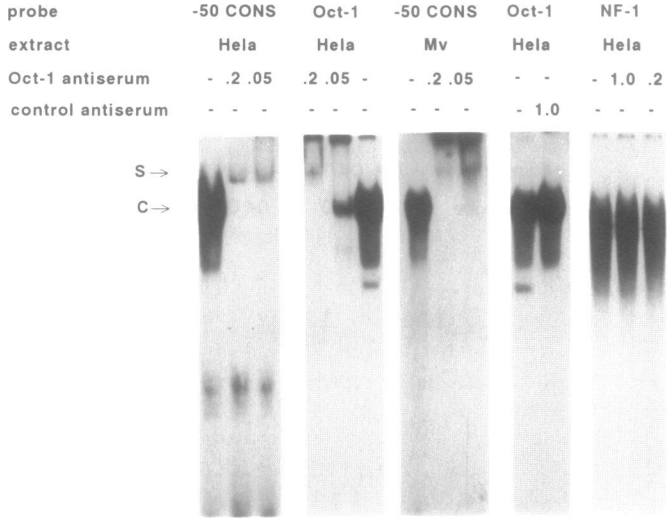

Figure 6. Complex formation in the presence of Oct-1 specific antiserum. Polyclonal antibodies directed against the DNA-binding domain of human Oct-1 were raised in rabbits. 1 μl of undiluted (1.0) or of 5-fold (0.2) or 20-fold (0.05) diluted antiserum was added as described in material and methods. The antiserum is added after the complex with Oct-1 (C) has formed, resulting in a complex of lower mobility (S). However since the antiserum is specific for the DNA binding domain of Oct-1 a large proportion of the Oct-1 will no longer be able to bind to the DNA, resulting in much lower levels of complex than without antiserum. Mv=mammary gland of a 1½ year old non lactating animal.

detected in this tissue. Another striking observation was that sometimes two different complexes (figure 2A and 3, complexes I and II) were seen. Different extracts made from the same tissue of the same animal either gave both complexes or only complex I. Furthermore if complex II was present, complex I was present at a very low concentration. Incubation of the extracts in which both of the complexes were detected for variable times at 37 or 45°C did not result in the disappearance of complex II. Also, repeated freezing and thawing had no effect on the relative intensity of the two complexes. Finally, the addition of several different nuclear extracts (Hela, liver) to a mammary gland nuclear extract that only gave complex I, did not result in the formation of complex II (data not shown).

Oligonucleotides O1 and O4 bind the octamer binding protein Oct-1

With oligonucleotides O1 and O4 a complex with similar electrophoretic mobility was detected with all the extracts used (figure 2). In addition several other faster migrating bands were seen with oligonucleotide O1 (marked X in figure 2). Competition experiments (figure 4) showed that both oligonucleotides were binding the same protein and that O1 contained a stronger binding site than O4. The complex formed with the weaker binding site on O4 was efficiently competed away with a 50 fold molar excess of either O1 or O4 (figure 4 lanes g and h). In contrast the complex formed with the stronger binding site on O1 was competed away much more efficiently by a given molar excess of O1 than O4 (figure 4). Oligonucleotides A9 and T1, which partially overlap with O1, did not compete. The minor bands (X, figure 4)) were only competed away by O1 and not by the other competitors. Using O4 and A9 as competitor even resulted in the appearance of another minor band.

Oligonucleotide O1 contains part of a highly conserved sequence found between nucleotides −65 and −40 of the calcium sensitive casein genes. To determine whether this sequence was the protein recognition site, an oligonucleotide homologous to nucleotides −30 to −48 of the bovine β-casein gene (−50 CONS) was synthesized. It was found that the complex formed efficiently with this oligonucleotide (figure 5), indicating that this indeed is the sequence recognized by the protein. Furthermore, the bands with faster electrophoretic mobility were not present when this oligonucleotide was used as the probe. To examine whether the ubiquitous protein that binds to O1, O4 and −50 CONS was a known transcription factor, a competition experiment was performed with 50 fold molar excesses of oligonucleotides containing the binding sites of known transcription factors (figure 5). The only oligonucleotide that competed with the formation of the complex was an oligonucleotide containing the octamer binding site ATTTGCAT. The consensus for the sequence at −50 of the calcium sensitive caseins, ATTAGCAT, shows only one mismatch with the octamer sequence whereas that of the bovine αS2-casein gene, ATTACCAT, shows two mismatches with this sequence.

As an extra control to determine whether the protein that binds to −50 CONS is indeed Oct-1, supershift experiments were performed with antiserum highly specific for the DNA binding domain of Oct-1 (Figure 6). Addition of as little as 0.05 μl of the antiserum to the reaction of −50 CONS with Hela nuclear extract, resulted in the complete disappearance of the complex (C in figure 6) and the appearance of more slowly migrating band (S in figure 6). The same result was obtained when either Oct-1 was used as the probe or when a nuclear extract from the mammary gland was used. High amounts of the Oct-1 antiserum

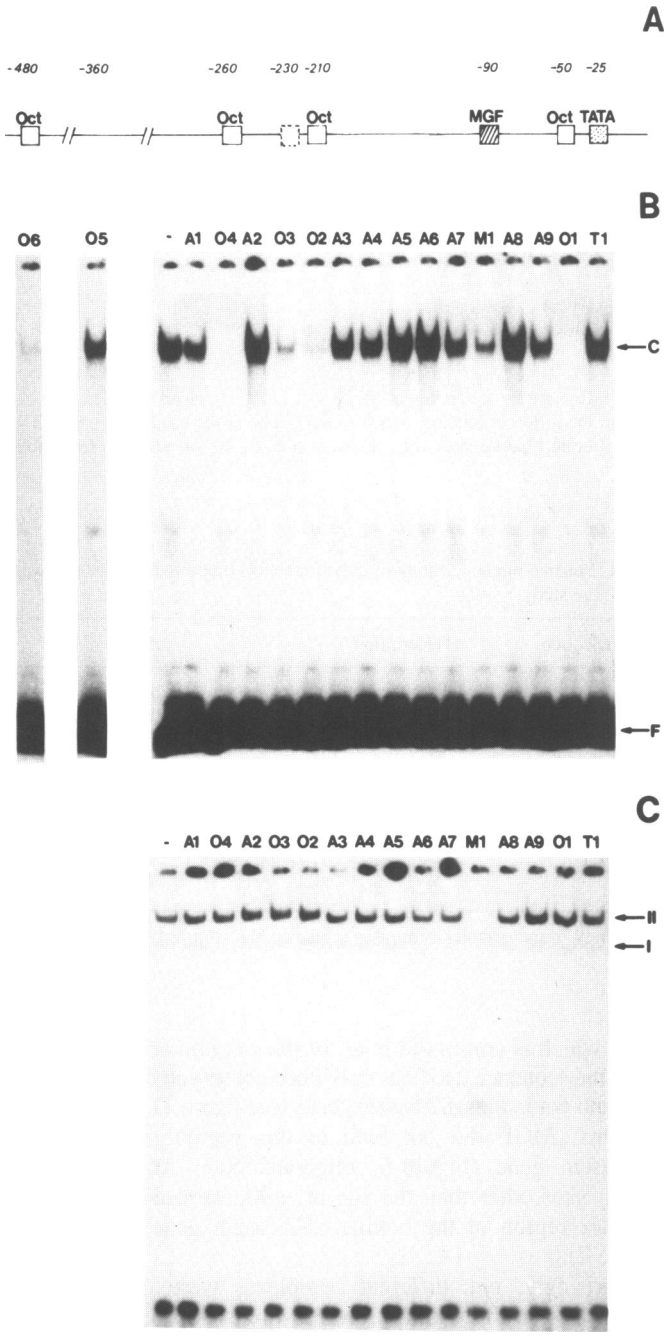

Figure 7. Gel retardation experiment to determine other potential MGF and octamer binding sites in the promoter region of the bovine αS2-casein gene. (A) Schematic representation of the position of the TATA box, the MGF binding site and the octamer binding sites. The position of the very weak octamer binding site O4 is also shown. The approximate position of the binding sites is shown at the top of the figure. (B) Gel mobility shift assay competition experiment to determine potential octamer binding sites. The probe used was −50 CONS and the nuclear extract was derived from Hela cells. The competitor oligonucleotides used are indicated above each lane, and these were added at a 100 fold molar excess before adding the probe. The complex between −50 CONS and Oct-1 (c) and the free probe (F) are indicated. (C) Gel mobility shift assay competition experiment to determine other potential MGF binding sites. The probe used was M1 and the nuclear extract was derived from the mammary gland of a lactating cow. The competitor oligonucleotides used are indicated above each lane, and these were added at a 100 fold molar excess before adding the probe. In the extract used in this experiment both complexes (I and II) are formed. The free probe has just run of the gel. Lanes indicated with −; no competitor has been added.

(1.0 μl) had no effect on the binding of NF-1 to its binding site. Similarly, 1.0 μl of a control antiserum had no effect on binding of Oct-1 to its binding site. These results show that the protein that binds to −50 CONS is indeed Oct-1, in both the Hela extract and the extract from the mammary gland.

Multiple octamer binding sites in the promoter of the bovine αS2-casein gene

To examine the possible occurrence of other potential binding sites for oct-1 within the promoter region of the bovine αs2-casein gene, all oligonucleotides spanning the region from −305 up to −15 were used in a competition experiment with −50 CONS as the probe (figure 7B). In addition two other oligonucleotides (O6 and O5), homologous to sequences −484 to −467 and −371 to −352 of the bovine αS2-casein gene respectively, were also used. These sequences show considerable homology with the octamer binding site namely ACTTGCAT and ATAAGCAT, respectively. As can be seen in figure 7B, in addition to oligonucleotides O1 and O4, three other oligonucleotides (O6, O3 and O2) competed, although less efficiently, with the binding of Oct-1 to −50 CONS. Of these three, competition by O3 was very weak and it is not clear if this is a real octamer binding site. However, regarding the significance for expression of the αs2-casein gene, no conclusions can be drawn from the results presented in this study concerning the binding of Oct-1 to any of the weak Oct-1 binding sites (see discussion). Furthermore, oligonucleotide M1 also slightly competes with binding of Oct-1 to −50 CONS. Given the strong affinity of this site for MGF and the very low affinity for Oct-1 we do not think that this is of any significance.

A similar competition experiment was done with oligonucleotide M1 as the probe and the nuclear extract isolated from the mammary gland of a lactating animal (figure 7C). The only oligonucleotide that competed with the binding of MGF to M1 was M1 itself, indicating that no other binding sites for MGF occur within the promoter region of the bovine αS2-casein gene up to nucleotide −305. Furthermore, it can be seen that complex I and II behave the same in competition experiments.

Binding affinities and nucleotide sequences of the different octamer binding sites in the bovine αs2-casein promoter

The binding affinities of the different binding sites were determined in competition experiments with oligonucleotide −50 CONS as the probe and a 3, 10, 50 and 200 fold molar excess of the binding site to be tested as competitor (figure 8). The amount of complex still present in the different reactions was determined by scanning the autoradiogram, and the relative binding affinity was calculated relative to that of the consensus Oct-1 binding site. The relative binding affinities for the oligonucleotides used were: Oct-1 = O1 = −50 CONS > O4 > O2 = O6 > −50 MUT = O3 > O5. The quantitative analysis of the relative binding affinities is shown in table 1, the affinity of Oct-1 has arbitrarily been set to 100.

The sequences of the octamer binding sites described in this study have been aligned in figure 9 together with other previously identified octamer binding sites, the octamer consensus sequence and the TAATGARAT consensus sequence. The TAATGARAT sequence is a cis-regulatory element common to herpesvirus IE gene enhancers, and is a target for trans-activation by the viral gene product VP16 (1). The action of VP16 to this element appears to be mediated by cellular proteins that bind to this element. Given the high degeneracy of some of the previously

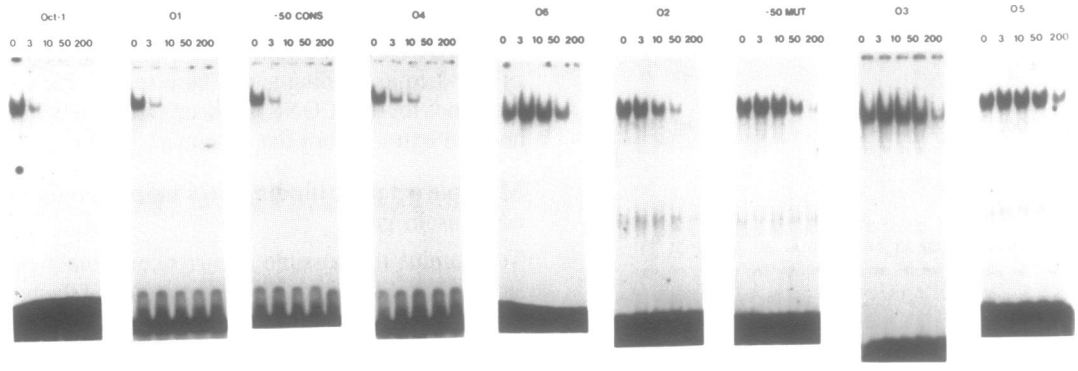

Figure 8. Gel mobility shift assay competition experiment to determine the relative affinity of the different binding sites for Oct-1. The probe used was −50 CONS and the nuclear extract was derived from Hela cells. The oligonucleotides containing the different binding sites were added at a 0, 3, 10, 50 and 200 fold molar excess before adding the probe.

identified octamer binding sites (29) it was not directly obvious how to align the binding sites present on O2 and O4. On the basis of a comparison with the other sequences shown in figure 9, the sequences shown for O2 and O4 probably are the sequences recognized by Oct-1.

DISCUSSION

Regulation of eukaryotic gene transcription involves the simultaneous interaction of several transcription factors with specific sites in the promoter and enhancer of a particular gene. To understand how a particular gene is regulated at the molecular level it is, therefore, essential to know what transcription factors interact with the promoter and enhancer of that gene. In this study we have used gel mobility shift assays to examine the interaction of transcription factors with the promoter region of the hormonal and developmentally regulated bovine αS2-casein gene. Two factors have been identified that bind to sequences that are conserved in the promoter region of all calcium sensitive caseins sequenced thus far: The mammary gland specific factor MGF, and the ubiquitous factor Oct-1.

The mammary gland specific factor MGF

This factor recognizes the sequence ANTTCTTRGAATTNA between nucleotides −85 to −99 in the promoter region of the bovine αS2-casein gene. This sequence has been conserved completely at that position in all calcium sensitive casein genes sequenced thus far. However this sequence is not found in the promoter of the genes encoding the other major milk protein genes; k-casein, β-lactoglobulin and α-lactalbumin. That this sequence is involved in the binding of MGF is further supported by the fact that simultaneous mutations of the G and C residues in this sequence completely abolishes binding of MGF. Binding of MGF to this sequence in the promoter of the rat β-casein gene has recently been described (16). In that study mutations of the G and C residues also inhibited binding of MGF. Furthermore they showed that these mutations affected transcription of the gene after transfection of the mutant gene to the mouse mammary gland epithelial cell line HC11. In the promoter of the rat β-casein gene a second, weaker, MGF binding site was found at position −130 to −150 (16). This region contains the sequence ANTTCTTGGNA which closely resembles the conserved sequence at position −90. Although this sequence occurs in a region which is conserved in all of the calcium sensitive casein genes, the sequence itself has only been conserved in the β-casein genes and not in the αS2-casein genes (see figure 1). In agreement with this, MGF did not bind to this region of the bovine αS2-casein gene (Figure 6, oligonucleotide A6). No MGF binding sites,other than the site at −90, were detected in the promoter region of the bovine αS2-casein gene up to −300 (Figure 7).

Surprisingly, two different complexes were detected with oligonucleotide M1. Both of these complexes probably involve the binding of MGF to the sequence ANTTCTTRGAATTNA for the following reasons; (i) competition experiments affect both complexes in the same way; (ii) mutations of the G and C residue in this sequence prevents the formation of both complex I and II; (iii) in different extracts made from the mammary gland of the same animal sometimes only complex I is seen, and if complex II is present, only small amounts of complex I are found. Although we can not rule out the possibility of proteolytic clipping in an early stage of preparation of the extract, we think this is not very likely given the number and amount of protease inhibitors included in the extraction buffers that were used. A possible explanation for the occurrence of an extra, more slowly migrating, complex might be that a second protein interacts specifically with MGF. If this should turn out to be the case, then this other protein probably is not one of the general transcription factors since adding different extracts to mammary gland extracts that only gave complex I, never resulted in the

Table 1. Relative binding affinity of the different octamer binding sites that were used in this study.

Oligonucleotide	Description	affinity
Oct-1	octamer consensus	100
O1	αS2-casein −35/−50	100
−50 CONS	casein −50 consensus	100
O4	αS2-casein −250/−280	25
O2	αS2-casein −191/−220	7
O6	αS2-casein −476/−484	7
O3	αS2-casein −210/−240	2
−50 MUT	−50 consensus mutant	2
O5	αS2-casein −352/−371	0.5

The affinity of the octamer consensus sequence has arbitrarily been set to 100.

Bindingsite	Sequence	Reference	Binding
	1 2 3 4 5 6 7 8 9 10 11 12 13 14		
Octamer consensus	A T T G C A T		
Adenovirus ITR	C T C A T T A T C A T A T T	30	S
Histone H2B	C T C A T T T G C A T A A G	31	S
U2 snRNA	C G A A T T T G C A T G C C	32	S
IgH enhancer	G T A A T T T G C A T T T C	33	S
IgVH promoter	A G G A T T T G C A T A T T	34	S
SV40 site I	A T G C T T T G C A T A C T	29, 35	S
Bos T. αS2-cas -50	A G A A T T A C C A T A T T	This study	S
Consensus cas -50	A R A A T T A G C A T N Y C	This study	S
TAATGARAT consensus	T A A T G A R A T		
SV40 site II	C T A A T T G A G A T G C A	29, 35	I
Bos T. αS2-cas -260	C A C A T T T T A A T A C T	This study	I
HSV ICP4 promoter	G G T A A T G A G A T G C C	29	I
Bos T. αS2-cas -210	A A T C A T G A C A T T T A	This study	W
Mut cons cas -50	A A A A T T A G A A T G C C	This study	W
Bos T. αS2-cas -230	C T A T A T T T C A T G T T	This study	W
Synthetic 1	C T A A T T G A G A T G C A T	29	I
Bos T. αS2-cas -480	C A G A C T T G C A T A T G	This study	W
Synthetic 2	C T A A T T G A G A T G C T A		-
Bos T. αS2-cas -360	T A A A T A A G C A T C T A		-

Figure 9. Alignment of different octamer binding sites. S=strong binding site, I=intermediate strength binding site, W=weak binding site, −=no significant binding. The three nucleotides that are conserved in all the sites that bind Oct-1 are boxed. Nucleotides that match the octamer consensus sequence are underlined, and those that match with the TAATGARAT sequence are indicated with a dot.

appearance of complex II and a decrease in the amount of complex I. Furthermore, why complex II is only occasionally found in the extracts of the mammary gland, even if tissue is used from the same animal, is at present not clear.

Although not at as high a concentration as in nuclear extracts from a lactating cow, MGF activity was found in the nuclear extract made from the mammary gland of a 1½ year old non-lactating animal. Since this tissue contains more connective tissue and less epithelial cells, the concentration of MGF is probably the same in epithelial cells from a lactating and a non-lactating animal. Therefore, MGF probably is a mammary gland specific transcription factor but not a lactation specific transcription factor.

Octamer binding sites

In the promoter region of the bovine αS2-casein gene several octamer binding sites occur. These sites bind the ubiquitous transcription factor Oct-1 with different affinities. Two of these binding sites, at positions −50 and −260, exhibit strong and intermediate strength affinity to Oct-1 respectively. The other sites, at positions −480, −210 and −230, are very weak octamer binding sites. From this study it is not clear whether these weak binding sites are indeed functional *in vivo*, the functional significance of the extremely weak binding site at position −230 is particularly doubtful. However, since cooperative binding to adjacent binding sites might facilitate binding to the weaker sites, the occurrence of multiple weak binding sites could be important for activation of gene expression *in vivo*. Examination of the sequences of the promoter regions of the calcium sensitive genes (figure 1) shows that the weak octamer binding sites are not conserved, however, since many different sequences can act as weak octamer binding sites (29, this study), other sequences may have evolved as weak octamer binding sites in those genes (e.g the sequences CCATTGAATCC at position −450 in the 5′ upstream region of the rat αS2-casein gene).

Contrary to these weak octamer binding sites the strong octamer binding site at position −50 has been conserved in all calcium sensitive caseins sequenced thus far and, therefore, probably is essential for expression of these genes. Interestingly, the octamer sequence ATTAGCATNT is part of the larger sequence AAACCACARAATTAGCATNT which has been conserved in all of these genes around position −60/−50. A similar situation is seen in all the histone H2B promoters, where the octamer motif is embedded in an extended consensus sequence of 15 bp (31). It has been suggested (37) that the sequence context of the octamer motifs helps to specify either Oct-1 or Oct-2 binding, and that in order to be able to activate expression *in vivo*, Oct-1 has to be supplemented by another factor(s). These supplementary factors could either interact only with Oct-1 itself as has been shown for the viral trans-activator VP16 or in addition to interaction with Oct-1, they could also make specific contacts with certain nucleotides adjacent to the octamer motif. The high conservation of sequences adjacent to the octamer motif both in the histone H2B promoter and the promoters for the calcium sensitive caseins make this an attractive hypothesis, despite the fact that binding of factors to these sequences has so far not been detected.

Another similarity between the promoters of the calcium sensitive casein genes and the Histone H2B genes is that in both situations the octamer sequence is located 25 bp upstream of the TATA box.

The octamer motif is interesting not only because it is recognized by two very different proteins, the lymphocyte-specific transcription factor Oct-2 and the ubiquitous transcription factor Oct-1, but also because very divergent sequences can bind Oct-1 (29, this study). Two different consensus sequences bind Oct-1 : the octamer motif ATTTGCAT and the TAATGARAT sequence. Of these two sequences the octamer sequence has the higher affinity for Oct-1 (29). In figure 9 a number of different octamer binding sites have been aligned with the two consensus sequences. All the octamer binding sites in the promoter region of the bovine αS2-casein gene have greater homology with the octamer consensus except for the binding site at −210, which most closely resembles the TAATGARAT sequence. Three nucleotides have been conserved in all sequences that bind Oct-1; the T at position 6, the A at position 10 and the T at position 11. Furthermore, a pyrimidine is always found at position 13. Oct-1 probably interacts strongly with these nucleotides. It has been reported previously that nucleotides adjacent to the octamer sequence ATTTGCAT are important for the binding of Oct-1 (29). This is also supported by our finding that oligonucleotide A9, which contains the complete octamer sequence ATTACCAT at position −50 but which lacks sequences at the 3′ end did not bind Oct-1. This oligonucleotide did not compete with oligonucleotide −50 CONS for binding of Oct-1 in a competition experiment with concentrations of this competitor as high as a 100 fold molar excess.

Potential binding sites for other transcription factors

In addition to the octamer binding site at −50 and the MGF binding site at −90, other sequences in the promoter regions of the calcium sensitive caseins have also been conserved (figure 1). This suggests that other transcription factors interact with these sequences. One of these sequences, the one at position −60 adjacent to the octamer binding site at −50, has already been discussed above. Another example is the sequence around position −115 (oligonucleotide A6), which contains the palindromic sequence AGAAANNNTTTCT, suggesting that it might bind a dimeric protein. The concentration of these proteins either might have been to low to be detected or the proteins might have lost their activity during preparation of the extracts. Furthermore, in non-lactating animals the mammary gland contains only a small number of epithelial cells. Therefore it is possible that the concentration of certain transcription factors that have a negative effect on transcription of the calcium sensitive caseins is to low to be detected. The use of a bovine mammary gland epithelial cell line that can differentiate in culture and subsequently express these genes, such as the bovine MAC-T cell line (38), might provide a good model system to study these genes and the transcription factors involved into more detail.

ACKNOWLEDGMENTS

We like to thank Peter Verrijzer for kindly providing the Oct-1 antiserum and Chris Davies for critically reading the manuscript.

REFERENCES

1. Johnson, P.F. and McKnight, S.L. (1989) Ann. Rev. Biochem. 58, 799−839
2. Mitchell, J.P. and Tijan, R, (1989) Science 245, 371−378
3. Ptashne, M. (1988) Nature 335, 683−689
4. Roeder, R.G. (1991) TIBS 16, 402−408
5. Horikoshi, N., Maguire, K., Kralli, A., Maldonado, E., Reinberg, D., and Weinmann, R. (1991) Proc. Natl. Acad. Sci. USA 88, 5124−5128
6. Lee, W.S., Kao, C., Bryant, G.O., Liu, X. and Berk, A.J. (1991) Cell 67, 365−376
7. Lin, Y-S. and Green, M.R. (1991) Cell 64, 971−981
8. Scafe, C., Chao, D., Lopes, J., Hirsh, J.P., Henry, S. and Young, R.A. (1990) Nature 347, 491−494
9. Mercier, J.C., Vilotte J.L. and Provot, C. 1990. In 'Genome analysis in domestic animals'. (Geldermann, H. Ed.) VCH, Weinheim, Germany
10. Vonderhaar, B.K. and Ziska, S.E. (1989) Ann. Rev. Physiol. 51, 641−652
11. Aggeler, J., Park, C.S. and Bissell, M.J. (1988) J. Dairy Sci. 71, 2830−2842
12. Threadgill, D. W. and Womack, J. E. (1990) Nucl. Acid Res. 18, 6935−6942
13. Ferretti, C., Leone, P. and Sgaramella, V. (1990) Nucl. Acid Res. 18, 6829−6833
14. Lee, K-F., Atiee, S.H. and Rosen, J.M. (1989) Mol. Cell. Biol. 9, 560−565
15. Doopler, W., Groner, B. and Ball, R.K. (1989) Proc. Natl. Acad. Sci. USA 86, 104−108
16. Schmitt-Ney, M., Doppler, W., Ball, R.K. and Groner, B. (1991) Mol. Cell. Biol. 11, 3745−3755
17. Yoshimura, M. and Oka, T. (1990) Proc. Natl. Acad. Sci. USA 87, 3670−3674
18. Watson, C.J., Gordon, K.E., Robertson, M. and Clark, A.J. (1991) Nucl. Acids Res. 19, 6603−6610
19. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucl. Acids Res. 11, 1475−1489
20. Hennighausen, L and Lubon, H. (1987) In 'Guide to molecular cloning techniques' Eds. Berger, S.B. and Kimmel, A.R.; Academic press inc. pp 721−735
21. Yu-Lee, L.Y., Richter-Mann, L., Cough, C.H., Stewart, A.F., MacKinlay, A.G., Rosen, J.M. (1986) Nucl. Acids Res. 14, 1883−1902
22. Koczan, D. and Seyfert, H-M. (1991) Nucl. Acids Res. 19, 5591−5596
23. Groenen, M.A.M., Dijkhof, R.J.M. and van der Poel, J.J. (1990) Proc. 4th world congr. gen. appl. lives. prod. XIII, 79−82
24. Gorodetsky, S.L., Tkach, T.M. and Kapelinskaya, T.V. (1988) Gene 66, 87−96
25. Thépot, D., Devinoy, E., Fontaine, M.L. and Houdebine, L.M. (1991) Gene 97, 301−306
26. Yoshimura, M. and Oka, T. (1989) Gene 78, 267−275
27. Yu-Lee, L.Y. and Rosen, J.M. (1983) J. Biol. Chem. 258, 10794−10804
28. Jones, W.K., Yu-Lee, L.Y., Clift, S.M., Brown, T.L. and Rosen, J.M. (1985) J. Biol. Chem. 260, 7042−7050
29. Baumruker, T., Sturm R. and Herr, W. (1988) Genes Dev. 2, 1400−1413
30. Pruijn, G.J.M., van Driel, W. and van der Vliet P.C. (1986) Nature 322, 656−659
31. Harvey, R.P., Robins, A.J. and Wells, J.R.E. (1982) Nucl. Acids Res. 10, 7851−7863
32. Ares, M. Jr., Mangin, M. and Weiner, A.M. (1985) Moll. Cell. Biol. 5, 1560−1570
33. Ephrussi, A., Church, G.M., Tonegawa, S. and Gilbert, W. (1985) Science 227, 134−138
34. Mason, j.O., Williams, G.T. and Neuberger, M.S. (1985) Cell 41, 479−487
35. Sturm, R., Baumruker, T., Franza, B.R.Jr. and Herr, W. (1987) Genes Dev. 1, 1147−1160
36. Renkawitz (1990) TIG 6, 192−196
37. Schaffner, W. (1989) TIG 5, 37−38
38. Turner, J.D. and Hung, H.T. (1989) J. Dairy Sci. 72, Suppl. 1, 317