Journal of Mammary Gland Biology and Neoplasia, Vol. 8, No. 2, April 2003 (© 2003)

The Nuclear Factor I (NFI) Gene Family in Mammary Gland Development and Function

Janice Murtagh,[1] Finian Martin,[1,3] and Richard M. Gronostajski[2,3]

Mammary gland development and function require the coordinated spatial and temporal expression of a large fraction of the mammalian genome. A number of site-specific transcription factors are essential to achieve the appropriate growth, branching, expansion, and involution of the mammary gland throughout early postnatal development and the lactation cycle. One family of transcription factors proposed to play a major role in the mammary gland is encoded by the Nuclear Factor I (NFI) genes. The NFI gene family is found only in multicellular animals, with single genes being present in flies and worms and four genes in vertebrates. While the NFI family expanded and diversified prior to the evolution of the mammary gland, it is clear that several mammary-gland specific genes are regulated by NFI proteins. Here we address the structure and evolution of the NFI gene family and examine the role of the NFI transcription factors in the expression of mammary-gland specific proteins, including whey acidic protein and carboxyl ester lipase. We discuss current data showing that unique NFI proteins are expressed during lactation and involution and suggest that the NFI gene family likely has multiple important functions throughout mammary gland development.

KEY WORDS: NFI; transcription; mammospheres; mammary gland; development; lactation.

INTRODUCTION

The mammary gland is one of the few mammalian organs to develop predominantly postnatally. The stages of its postnatal development, mammary tree expansion during puberty, alveolar epithelial cell proliferation during pregnancy, lactation-induced differentiation, and the apoptosis and remodeling associated with involution, provide a unique opportunity to study stage- and tissue-specific gene expression. Both ubiquitous and cell-type specific promoter DNA-protein interactions, those of NFI among them, would seem to be required for the coordination of tissue and stage-specific gene expression that underpins this development (reviewed in (1) and (2)).

The NFI family of transcription/replication genes encode one of the first site-specific transcription factors cloned (3). NFI was thought initially to be a single protein and was purified from human HeLa cell nuclear extracts as a factor that was essential for the in vitro replication of adenovirus (Ad) DNA (4). Subsequent studies revealed that NFI bound specifically to sites within the origin of replication of Ad DNA and that these sites were essential for DNA replication (5–7). NFI thus appears to function in DNA

[1] Conway Institute of Biomolecular and Biomedical Research and Department of Pharmacology, University College Dublin, Belfield, Dublin 4, Ireland.

[2] Department of Biochemistry, State University of New York at Buffalo, 140 Farber Hall, 3435 Main Street, Buffalo, New York, 14214.

[3] To whom correspondence should be addressed; e-mail: rgron@buffalo.edu; e-mail: finian.martin@ucd.ie.

Abbreviations used: NFI, Nuclear Factor I; CTF, CAAT-box transcription factor; DBD, DNA-binding domain; Ad, Adenovirus; KO, knock out; β4-GT, β1,4-galactosyltransferase; CEL, carboxyl ester lipase; MMTV, mouse mammary tumour virus; ECM, extracellular matrix; WAP, whey acidic protein; BLG, β-lactoglobulin; STAT, signal transducer and activator of transcription; GR, glucocorticoid receptor; EMSA, electrophoretic mobility shift assay; LTR, long terminal repeat; MAF, mammary cell-activated factor; ConA, concanavalin A; WGA, wheat germ agglutinin.

1083-3021/03/0400-0241/0 © 2003 Plenum Publishing Corporation

replication by interacting with viral replication proteins and recruiting them to the replication origin (8–10). Further studies identified a large number of NFI binding sites (~60,000) distributed throughout the human genome (11,12). This early work suggested that NFI might function in host DNA replication, but to date there is no evidence that NFI plays such a role in mammalian cells. Instead, current evidence suggests that adenoviruses have usurped a potent and widely expressed cellular transcription factor family for their replication (2). However, given the complexity of mammalian replication origins and the links between DNA transcription and replication (13), only time will tell whether NFI functions in cellular DNA synthesis. Here we focus on the structure and evolution of the NFI genes and the known roles of NFI proteins in the regulation of transcription, with specific emphasis on their proposed functions in mammary gland development and physiology.

EARLY IDENTIFICATION OF NFI TRANSCRIPTION FACTORS

Coincident with the identification of NFI as a site-specific DNA-binding protein essential for adenovirus DNA replication, the TGGCA binding protein was identified as a protein that bound specifically to multiple sites in the chicken lysozyme gene (14). Comparing the sequences of the binding sites of the chicken TGGCA protein and human NFI suggested that these proteins were related, and detailed biochemical analysis showed that both proteins bind the same DNA sequences and stimulate Ad DNA replication (15). These data, together with the identification of NFI binding sites in the promoter regions of multiple genes (16), were the first indications that NFI proteins play a major role in transcriptional regulation. Since the consensus sequence for NFI binding (TTGGCN$_5$ GCCAA) has some similarity to canonical CAAT box sequences, NFI was also identified as a CAAT box transcription factor (CTF) (17). However, NFI proteins bind to only a small subset of CAAT sequences, with other proteins binding the majority of CAAT sites (18,19). Since transcripts from the NFI-C gene were first cloned as CTF cDNAs, the NFI family is sometimes called the NFI/CTF family of transcription factors and NFI-C proteins are sometimes referred to as CTFs (3). The subsequent cloning of NFI cDNAs from a variety of organisms has shown that what was first thought to be a single protein is a multigene family of transcription factors.

EVOLUTION OF THE NFI MULTIGENE FAMILY

Initial studies identified four genes encoding NFI transcription factors in the chicken: NFI-A, NFI-B, NFI-C, and NFI-X (20,21) (Table I). These 4 genes had clear homologs in other vertebrates, including hamster (22), mouse (23,24), rat (25), and human (26,27). Sequencing of the *C. elegans* and *Drosophila* genomes revealed single NFI genes in these organisms (2), and no NFI homologs have been identified in prokaryotic, fungal or plant genomes. Thus the NFI transcription factor family appears to have arisen as a single gene during the evolution of multicellular animals and to have been amplified to a four gene family in vertebrates. The recent sequencing of the genomes of the mosquito *Anopheles* (28) and the urochordate *Ciona* (29) have verified these initial findings and have begun to shed light on the evolution of NFI gene structure (Fig. 1). All mammals whose genomes have been sequenced have four NFI genes that are highly homologous to the chicken NFI-A, -B, -C, and -X genes. These are *NFIA*, *NFIB*, *NFIC*, and *NFIX* in human (26) and *Nfia*, *Nfib*, *Nfic*, and *Nfix* in mouse (24) and rat. The four NFI genes are distributed throughout the genomes, with *NFIC* and *NFIX* being linked in human, and their homologs not linked in mouse and rat, while *Nfia* and *Nfib* are linked in mouse and rat, and their homologs not linked in humans (Table I). The mammalian and chicken genes all contain at least 11 exons (2), but since novel NFI transcripts continue to be obtained from these species it is possible that additional exons exist. Together with the alternative splicing of 3′ exons seen in all four vertebrate NFI genes (31), both *Nfia* and *Nfib* possess conserved 1st exons that are expressed specifically in brain (Table I) (2,32,33).

One unusual feature of the mammalian and chicken NFI genes is that the DNA-binding domains (DBDs) for the NFI proteins are encoded predominantly within a single large exon for each gene, and the 5′ splice acceptor and 3′ splice donor sites for these exons are at identical positions for all four genes in each species (34). This structural homology shows the relatedness of the four vertebrate NFI genes, and the size and distribution of introns within the DBD-encoding exons of NFI genes of other species yields insights into the evolution of the gene family (Fig. 1). For example, in the cephalochordate *Amphioxus* there appears to be a single NFI gene, as in worms and insects, but with the DBD encoded predominantly by a single large exon, as in chickens and mammals

Table I. Structure of NFI Genes in Various Organisms

Organism	No. NFIs	Gene names	Chr. loc.	No. exons[a]	No. introns in DBD[b]	Brain-specific isoform[c]	Mamm.-specific isoform
Human (*H. sapiens*)	4	*NFIA*	1p31.2-3	>11	0	P	?
		NFIB	9p24.1	>11	0	P	?
		NFIC	19p13.3	>11	0	?	?
		NFIX	19p13.3	>11	0	?	?
Mouse (*M. musculus*)	4	*Nfia*	4:45.8cM	>11	0	Y	N
		Nfib	4:38.6cM	>11	0	Y	N
		Nfic	10:43.0cM	>11	0	?	Y[d]
		Nfix	8:38.6cM	>11	0	?	N
Rat (*R. norvegicus*)	4	*Nfia*	5	>11	0	Y	?
		Nfib	5	>11	0	P	?
		Nfic	7	>11	0	?	?
		Nfix	19	>11	0	?	?
Chicken (*G. gallus*)	4	*NFI-A*	?	>11	0	P	NA
		NFI-B	?	>11	0	P	NA
		NFI-C	?	>11	0	?	NA
		NFI-X	?	>11	0	?	NA
Amphioxus (cephalochordate)	1[e]	NFI	?	?	0[f]	?	NA
Ciona intestinalis (urochordate)	1	NFI[g]	Scaf. 131[h]	>2[i]	1[i]	?	NA
Anopheles gambiae (insect)	1	NFI[g]	X	>3[j]	2[j]	?	NA
D. melanogaster (insect)	1	NFI	4	>6[k]	1[k]	?	NA
C. elegans (nematode)	1	*nfi-1*	II	>15[k]	4[k]	?	NA

[a] Number of exons estimated from cDNA and genomic sequences for vertebrates.
[b] Number of introns within highly-conserved DNA-binding and dimerization domain (DBD). See Fig. 1 for locations.
[c] Does gene contain conserved 1st exon expressed specifically in brain (Y = yes, P = probably based on sequence conservation, ? = unknown).
[d] Mammary-specific glycosylation of NFI-C protein seen during involution (see below) (Y = yes, N = no, NA = not applicable, ? = unknown).
[e] Based on PCR analysis of genomic DNA, multiple species obtained but may be polymorphisms.
[f] Based on PCR analysis of genomic DNA.
[g] Provision gene name.
[h] Designation of *Ciona intestinalis* genome browser (http://genome.jgi-psf.org).
[i] Based on TBlastN analysis with *C. elegans* gene and ESTs.
[j] Based on TBlastN analysis with *C. elegans* gene. No known ESTs or transcripts.
[k] Based on PCR, ESTs and cDNAs.

(Fig. 1). However, in the urochordate *Ciona*, the single NFI gene contains a DBD-encoding region with the same 5′ and 3′ splice acceptor and donor sites seen in the mammalian and chicken genes, but with an additional intron interrupting this exon that is absent in *Amphioxus* and mammals but present in *Anopheles* (Fig. 1, gray triangle, intron 4). In *Anopheles*, although NFI cDNAs have not been isolated, the genomic sequence indicates that the region encoding the DBD shares the conserved 5′ splice acceptor and 3′ splice donor sites with the mammalian genes but contains two additional introns: one that is present in both *Drosophila* and *C. elegans* and one that present in the *Ciona* gene (Fig. 1, introns 2 and 4). In *Drosophila*, the DBD-encoding exons are interrupted by a single intron (Fig. 1, intron 2) that is also present in *C. elegans*, but here the conserved 3′ splice acceptor site is lost while the conserved 5′ splice donor site is retained (2). Finally, in *C. elegans* the DBD-encoding region is interrupted by four introns, one that is present in both *Drosophila* and *Anopheles* and one that is present only in *Anopheles* (2,34) (Fig. 1, introns 1, 2, 3, and 5). This *C. elegans* DBD-encoding region contains the initial 5′ splice acceptor site seen in all other species and a 3′ splice donor site shared with *Drosophila*. All five introns in the nonmammalian species are in the same reading frame (phase 0), suggesting a possible relationship.

These shared features strongly suggest that there was a single primordial NFI gene, whose DBD was

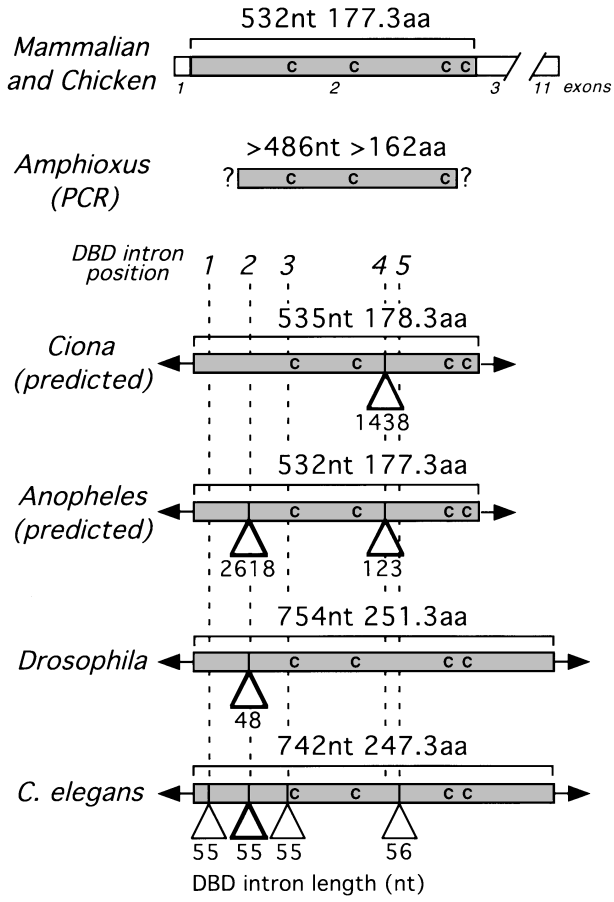

Fig. 1. Introns within the DBDs of nonmammalian NFI genes. The DBD-encoding exons of Mammals and Chickens, *Amphioxus*, *Ciona*, *Anopheles*, *Drosophila*, and *C. elegans* are shown as gray boxes with introns indicated by triangles below. Intron sizes are given below the triangles. The Cs within the exon boxes indicate conserved cysteine residues in the NFI-DBD (30). The thin black triangles indicate exons present only in *C. elegans*, the thick black triangles denote introns present in *C. elegans*, *Drosophila*, and *Anopheles* while the gray triangles show introns shared by *Anopheles* and *Ciona*. The positions of the introns are shown by dashed lines labeled 1–5. Numbers above the exon boxes show the length of the DBD-encoding region in nucleotides and amino acids and brackets indicate that complete exon sequences are known or predicted. The question marks and lack of a bracket on *Amphioxus* reflect that PCR was used to show the absence of introns and that the full DBD exon size is unknown. Arrows at the end of the nonmammalian exons indicate additional exons are known or predicted.

encoded by at least six exons and arose in the ancestors of nematodes, insects, and mammals. This primordial gene became four NFI genes at some time between the divergence of cephalochordates, and vertebrates. It seems less likely that multiple NFI genes were lost from insects, worms, urochordates, and cephalochordates, leading to the single NFI gene in these species. However, the recently sequenced pufferfish genome (*Fugu rubripes*) (35) contains what appear to be five NFI genes (R.M.G., unpublished), one containing some of the introns seen in the DBDs of the worm, insect, and urochordate genes, and four that appear to contain no introns within the DBD region, as seen in other vertebrates. Blast analysis shows that the four NFI genes lacking introns in their DBDs are most homologous to the four mammalian NFIs, while the 5th gene is slightly more distant. Since few *Fugu* NFI transcripts have been isolated, it will be important to determine whether any of these five apparent *Fugu* NFI genes represent pseudogenes or whether *Fugu* is an anomaly among vertebrates and contains five functional NFI genes. Of particular relevance to this review, since nonmammals (chickens and pufferfish) contain multiple NFI genes, is the fact that the NFI family diversified prior to the evolution of the mammary gland. The structural similarities of the NFI genes in mammals, insects and worms should make insects and worms useful model systems in which to study NFI function, possibly even as it relates to mammary gland-specific functions of NFI proteins.

NFI PROTEINS AS MODULAR TRANSCRIPTION FACTORS

Binding sites for NFI proteins have been found in many mammalian genes, including those expressed specifically in brain (36–38), liver (39–41), lung (42), mammary gland (43), and other tissues. Mutation of NFI binding sites in genes has led to either repression (44,45) or activation of gene expression (46–48), depending on the specific gene studied. Thus NFI proteins appear to both activate and repress gene expression (see (2) for a review). Transient transfection studies have shown that NFI proteins appear modular, with N-terminal domains that mediate DNA-binding and dimerization, and C-terminal domains that mediate transactivation and repression (30,49–50). Since products of the four NFI genes are differentially spliced (31) and can form both homo- and heterodimers (51,52), there may be many unique dimeric NFI proteins present within a single tissue or cell. In addition, since the same NFI proteins can either activate or repress expression from the same promoter (depending on the cell-type used) (52–54), the forms of NFI present in a given tissue or cell-type must be determined if the role of NFI in that tissue is investigated. Since many studies proposing a role for NFI

binding sites in a given gene have failed to identify which NFI proteins function at those sites, it has been difficult to determine the precise roles of individual NFI proteins in the expression of these genes.

NFI GENES IN MOUSE DEVELOPMENT

The four NFI genes are expressed in a complex and overlapping pattern during mouse development and in adult tissues (24). However, targeted deletion of single murine NFI genes has yielded specific developmental phenotypes, suggesting that each NFI gene has unique and essential roles in development. Mutation of three of the four murine *Nfi* genes has been reported (32,55,56). Loss of *Nfia* function causes severe neuroanatomical defects, including agenesis of the corpus callosum and hippocampal commissure, two major axonal tracts connecting the two hemispheres of the brain (32,57). Most *Nfia*−/− animals die within a day after birth, but rare surviving *Nfia*−/− mice develop hydrocephalus, which may be related to the neuroanatomical defects seen in these animals. Recently it has been shown that three midline glial cell structures appear to be aberrantly formed in *Nfia*−/− animals, suggesting that they may play a role in commissural axon guidance or survival (57). The molecular pathways connecting the loss of *Nfia* to these neuroanatomical defects remain to be determined. Also, the perinatal lethality of *Nfia*−/− mice has prevented the analysis of any role for *Nfia* in mammary gland function, but this problem could be overcome by use of mammary gland transplants or mammary-directed gene deletions.

Mice containing an insertion into the *Nfib* locus die at birth, apparently due to defects in lung development (55). The lungs of *Nfib*−/− animals appear to be arrested at an immature stage in development, leading to asphyxiation at birth. No other anatomical defects were noted in the *Nfib* KO animals. It will be of interest to determine whether this apparent arrest in lung development is absolute, or whether lung maturation can be induced in these animals, as is done in premature human infants. Again, the perinatal lethality of this mutation has hampered the analysis of *Nfib* function in mammary gland development.

Disruption of *Nfic* results in severe defects in postnatal tooth development (56). *Nfic*-deficient mice have weakened, dysmorphic incisors, and the roots of the molars fail to develop. However, the molar crowns appear normal in *Nfic*−/− animals; thus, this is the first mutation generated that affects primarily postnatal tooth root formation in the mouse. These tooth defects cause wasting and premature death if the animals are reared on standard laboratory chow, but *Nfic*−/− animals can survive and breed if maintained on a soft dough-like diet postweaning. *Nfic*−/− animals are currently being examined for possible defects in mammary gland development.

Given that disruption of three of the NFI genes has yielded specific developmental phenotypes, it is clear that NFI genes play essential roles in mammalian development. Studies on double and triple NFI-null mutants and disruption of the *Nfix* locus are essential for determining whether NFI genes play roles in developmental systems other than brain, lung and tooth. In particular, for analysis of NFI function in mammary glands it will be essential to generate mammary-specific knockouts and/or dominant negative forms of NFI or to perform mammary gland transplants from *Nfi*-null mice to wild-type mice to overcome the perinatal lethality seen in some *Nfi* knockout mice.

NFI IN MAMMARY GLAND DEVELOPMENT AND FUNCTION

NFI proteins have been suggested to contribute to different stages of mammary gland development. NFI has been implicated in regulating the transcription of milk protein genes such as whey acidic protein (58), α-lactalbumin (179) and β-lactoglobulin (60). Lactation-associated transcription of β1,4-galactosyltransferase (β4-GT) and carboxyl ester lipase (CEL) genes have also been shown to be regulated by NFI (61,62). Furthermore, a mammary cell-specific enhancer in the mouse mammary tumor virus (MMTV) promoter includes an NFI binding site that is required for function (63). In addition, involution-associated transcription of the TRPM-2/clusterin gene has been proposed to be controlled by a unique *N*-glycosylated 74 kDa NFI-C (64,65). Finally, it has been demonstrated that NFI activity is induced, as judged by increased NFI–DNA interactions, when primary mouse mammary epithelial cells are maintained on a laminin-rich ECM, a prerequisite for cell survival and differentiation (66). Thus, a clear potential exists for multiple contributions by different NFI isoforms to gene expression in the different phases of the cycle of postnatal mammary gland development, in particular in the terminally differentiated alveolar epithelial cells in late pregnancy and lactation, but also in involution. A summary of mammary-associated expressed genes that are known

Table II. Mammary-Specific Genes Regulated by NFI

Gene	NFI isoform implicated	Reference
α-Lactalbumin	—	59
β-Lactoglobulin	—	60
MMTV	NFI-B>NFI-X	63
WAP	NFI-B>NFI-X	43,79,82
β1,4-Galactosyltransferase	—	61
TRPM2	NFI-C	64,65,84
CEL	NFI-C	62,83

to be regulated by NFI transcription factor activity is given in Table II.

The study of milk protein gene expression has provided an attractive system for examining hormonal and developmental regulation of gene expression. Furthermore, it has offered an insight into mammary gland-specific control of transcription. α-Lactalbumin is one of the most abundant nutritional proteins and is essential for lactose synthesis. α-Lactalbumin is a component of the enzyme complex, lactose synthase, which catalyzes the synthesis of lactose. Expression of the rat α-lactalbumin gene becomes induced during pregnancy and lactation (67). The lactogenic hormones prolactin (68), insulin (69), and glucocorticoids (70–73) have been shown to play a role in this regulation. These observations stimulated the comparison of the 5′ flanking regions of α-lactalbumin and six other milk protein genes, revealing a conserved sequence that might play a role in their activation (74). To determine whether this conserved sequence in the α-lactalbumin promoter was a target for transcription factor binding during lactation, Lubon and Hennighausen performed DNase I footprinting analyses using nuclear extracts from virgin, pregnant and lactating rats (59). Their results showed sequence-specific interactions by mammary gland nuclear proteins with the −125 to −85 region of the rat α-lactalbumin promoter. Nuclear extracts from all three stages of mammary gland development provided a similar protection pattern. This region contains at least two sites with a similar sequence motif, TTGGCAG, which contains part of the consensus recognition site for NFI protein binding (12,17,75). Purified NFI from HeLa cells protected this sequence in a similar fashion to proteins in mammary gland extracts. These results suggested that the binding activities of nuclear proteins from mammary glands share characteristics with NFI or are indeed due to NFI–DNA interactions. It was hypothesized, therefore, that the proteins binding to the α-lactalbumin promoter are members of the NFI transcription factor family and that this binding contributes to driving the cell-specific α-lactalbumin gene expression seen in differentiated mammary epithelial cells.

β-lactoglobulin (BLG) is the major whey protein found in ruminant milk, but is absent from mouse milk. At least 5 NFI binding elements were identified in the BLG promoter, four of which are only half sites and would presumably bind an NFI monomer rather than an NFI dimer. When transgenic mice were generated that expressed a genomic BLG construct carrying 4.3 kb of 5′ flanking sequence immediately upstream of the transcription start-site, the transgene was efficiently and exclusively expressed in the pregnant/lactating mammary gland, reflecting the tissue-specific control of the activation of this promoter (58,76). The region between bases −149 and −406 of this proximal promoter was essential for efficient, tissue-specific expression, as deletion of this region resulted in a dramatically reduced level of transgene expression (77). In an attempt to identify DNA-protein interactions in this promoter region that are involved in transcriptional regulation of the BLG gene, Watson *et al.* performed in vitro binding studies (60). Using electromobility shift analysis, lactating mammary gland extracts exhibited at least two types of NFI–DNA interactions with the BLG promoter, each with distinctly different mobility patterns and recognition site affinities. Using competitor titration analysis, one of the NFI complexes was shown to have highest affinity for a consensus palindromic NFI binding element, while the other was shown to bind preferentially to a half-site element. These data suggest that a variant NFI is present in the lactating mammary gland that binds to a half site better than to the palindromic site. The importance of this observation is further emphasized by the multiplicity of NFI half-sites present in various milk protein gene promoters (63). Therefore, it is possible that a mammary gland-specific form of the transcription factor NFI is involved in the regulation of BLG transcription, and that it binds to the NFI half-site in preference to the classical palindromic NFI site.

Whey acidic protein (WAP), the major whey protein in rodents, is exclusively expressed at high levels in the mammary gland during late pregnancy and lactation (78). WAP gene expression is regulated by a number of complex interactions between transcription factors (42,79), as well as by cell-extracellular matrix (ECM) interactions (80). Transgenic mice were generated that contain a 3.0 kb rat WAP transgene consisting of 949 bp of immediate

5′-promoter sequence and 70 bp of 3′-flanking DNA sequence of the structural gene. This construct was expressed specifically in the mammary gland in a copy number-dependent, position-independent manner (81). Though the transgene and the endogenous WAP gene are expressed at maximum levels during lactation in the mammary gland, they are not expressed in most other tissues, e.g. the liver. In an attempt to identify critical regulatory promoter elements that facilitate tissue-specific, high level expression of WAP during lactation, Li and Rosen performed DNase I footprinting analyses on the more distal of two mammary-specific DNase I hypersensitivity sites on this promoter segment using extracts from lactating mammary glands and livers of these transgenic animals (82). Two mammary-associated protected regions were observed in this segment of the WAP promoter. Comparison of the sequences of the protected regions with the known consensus binding sites for a number of transcription factors revealed that the first footprint contained a half-palindromic binding element for NFI. The other protected region also showed extensive homology to an NFI binding element. To confirm that the transcription factors binding to the protected regions were indeed NFI family members, the protected sequence from the transgene was used as a radiolabelled oligonucleotide probe in EMSA. A consensus NFI binding element proved to be an efficient inhibitor of the multiple complexes observed in EMSA performed with lactating mammary gland nuclear extracts. Deletion of this distal mammary-specific hypersensitivity site region resulted in loss of rat WAP gene expression in transgenic mice. Since NFI was the principal factor found binding to this region in vitro, it suggested that NFIs were critical for WAP gene expression during lactation. Subsequent studies by Li and Rosen reported that point mutations introduced into the WAP promoter NFI binding sites dramatically decreased rat WAP gene expression in the transgenic mice (79). In addition, a binding element for STAT5, a mediator of prolactin-induced milk protein gene expression, was identified immediately proximal to the NFI binding sites. Point mutations introduced into this region reduced WAP transgene expression by 90% per gene copy, but did not alter tissue specificity of expression, suggesting that WAP gene expression is regulated by a cooperative interaction between STAT5 and NFI proteins.

More recent analyses of the distal region of the WAP gene promoter identified several glucocorticoid receptor (GR) binding elements (GRE) flanking these NFI binding sites (43). To investigate the synergy between GR, STAT5 and NFI activity on binding to this mammary gland specific and hormonally-regulated composite response element (CoRE) in the transcriptional regulation of WAP gene expression, transient cotransfection studies were preformed in JEG-3 choriocarcinoma cells because this cell line has low endogenous NFI and GR content. NFI-A4, NFI-B2, and NFI-X1 were the predominant splice-variant transcripts shown to be present during lactation, although studies from other laboratories also suggest the significant presence of NFI-C isoforms ((65), (83), and (84) (and see below)). These NFI isoforms were expressed from the mammalian expression vector pCH, and their activity on a transfected WAPtk-Luc promoter-reporter construct was determined. Prolactin receptor (PrlR), GR, and STAT5 were coexpressed with the different NFI isoforms. Synergy between STAT5 and GR was reflected by a 47-fold upregulation in WAP promoter-mediated reporter gene expression when both factors were expressed in the JEG-3 cells on stimulation with prolactin and hydrocortisone. When GR and STAT5 were coexpressed with NFI-A4, WAP promoter activity was significantly repressed. However, when GR and STAT5 were coexpressed with NFI-B2, WAP promoter activity was increased 7.3-fold, while with NFI-X1, a 5.3-fold increase in activity was observed. This work further supports the hypothesis that WAP gene expression in the lactating mammary gland is controlled by these specific transcription factor interactions on the WAP promoter, with prolactin-induced activated STAT5, hydrocortisone-activated GR and NFI-B2 and/or NFI-X1 all working cooperatively to regulate WAP gene transcription. NFI-B2 exhibited the highest level of cooperativity with STAT5 at the WAP distal composite response element and also displayed the highest cooperativity with GR in regulating both this WAP distal element and MMTV LTR-driven reporter activity in these transient transfection assays. These studies suggest that the NFI-B2 isoform might interact more efficiently with GR and STAT5 than do other lactation-expressed NFI isoforms. These studies have, for the first time, assigned differing activities to NFI isoforms in milk protein promoter activation, and interestingly, also showed that the DNA-binding specificities of the three NFI isoforms studied (A4, B1, and X1) were not the same on the WAP CoRE palindromic NFI site, in contrast to their interaction with the NFI consensus binding site from the adenovirus promoter.

β1,4-Galactosyltransferase (β4-GT) is a trans-Golgi resident, type II membrane glycoprotein that is widely distributed in the vertebrate kingdom (61). It catalyzes the transfer of galactose to *N*-acetylglucosamine residues, forming the β4-*N*-acetyllactosamine (Galβ4-GlcNAc) or poly-*N*-acetyllactosamine structures found in the *N*- and *O*-inked side chains of glycoproteins and proteoglycans (85). This enzyme is considered to have a housekeeping function, since glycoprotein biosynthesis occurs in all tissues. β4-GT also has a tissue specific function in the lactating mammary gland, where it is necessary for lactose production. The synthesis of lactose is catalyzed by the heterodimer, lactose synthetase, which is assembled from β4-GT and α-lactalbumin (85). Enzymatic levels of β4-GT increase in the mammary gland from mid-pregnancy through lactation, in order to facilitate lactose production (86).

The 5′ end of the β4-GT gene contains three transcription start sites. The most distal start site is male germ cell-specific (87). The middle start site is the predominant site used in a wide range of somatic cells (88). The proximal start site, located 200 bp 3′ of the middle start site, is preferentially used in the mid- to late pregnant and lactating mammary gland (88). Using DNase I footprinting analysis and electromobility shift assays, Rajput *et al.* showed that the region immediately upstream of the middle start site is occupied mainly by the ubiquitous transcription factor Sp1 (61). In contrast, the region adjacent to the proximal start site is bound by multiple transcription factors, including AP2, Sp1, a putative mammary gland specific form of NFI and a candidate negative regulatory complex that represses the transcription from this particular start site. A protected region seen adjacent to the proximal start site was shown to be due to a putative mammary gland-specific form of NFI binding to an NFI half site. Rajput *et al.* concluded that the proximal start site in the β4-GT gene promoter had evolved in order to facilitate mammary gland-specific production of β4-GT for lactose biosynthesis and that NFI, along with other transcription factors, regulates its transcription. However, the proposed mammary-specific NFI isoform involved has to date not been further characterized. There is the possibility that this is the same NFI isoform that had been partially characterized by Watson *et al.* (60) as binding to NFI half sites in the BLG promoter.

The mammary gland and pancreas are the main tissues that synthesize and secrete a bile salt-stimulated lipase, carboxyl ester lipase (CEL). This gene is constitutively expressed in the pancreas but its expression is developmentally regulated in the mammary gland where it becomes upregulated during pregnancy in parallel with, but to a much lower degree, than does the WAP gene. A fragment of the mouse CEL promoter, some 2.2 kb of DNA upstream of the translational initiation site, was isolated from the genomic mouse CEL1 clone (89). Transfection studies suggested that the region between bases −1696 and −1831 was important for high expression of the CEL gene (62). The DNA sequence was determined and revealed the presence of a number of transcription factor binding sites, including those for NFI, Oct-1, C/EBP-β, and the glucocorticoid receptor. To investigate whether any of these factors were interacting with this region, Kannius-Janson *et al.* performed DNase I footprinting on nuclear extracts from the mouse mammary-derived cell line, HC11 (62). One distinct footprint was detected extending from bases −1792 to −1764, a region containing a palindromic NFI binding site and a GR half site. When a mutation was introduced into the NFI binding site, the protein–DNA interaction was prevented, as reflected by a loss of protection. This result showed that an NFI family member or members is responsible for the DNA–protein interaction. Footprints with nuclear extracts from a pancreatic cell line gave a different binding pattern, suggesting that formation of the NFI–DNA complex was a mammary-specific event. It was concluded therefore that NFI–DNA interactions are important for the mammary-specific activation of CEL gene expression. A subsequent report from the same group suggested that the NFI isoform responsible for triggering mammary-specific CEL gene expression was a 50 kDa NFI-C, characterized by western blot analysis using an anti-NFI-C specific antibody (83). CEL promoter-luciferase reporter gene assays performed in HC11 cells suggested that the presence and binding of this 50 kDa NFI-C protein to the NFI-binding site in the CEL gene promoter activates CEL gene expression. Mutation of the NFI-binding site precluded binding of NFI-C, and, accordingly, reporter gene expression was reduced to 15% (83). Subsequent RT-PCR analysis on RNA extracted from HC11 cells suggested that the NFI-C transcript responsible for the expression of the 50 kDa NFI-C was NFI-C2, an isoform from which exon 9 is spliced. The binding of the NFI-C2 to the NFI-binding site in the CEL gene promoter increased and decreased in parallel with the expression levels of the NFI-C2 protein, and NFI-C2 interacted with the NFI binding element in the CEL promoter with a higher apparent affinity than did other NFI

isoforms. Kane *et al.* had previously reported that the NFI-C2 and NFI-C5 transcripts were the predominant NFI-C transcripts detectable in differentiated mammary epithelial cells (84). This characterization of the NFI/CEL promoter interactions further emphasizes a role for tissue-specific, isoform-specific actions of NFI in mammary epithelial cells.

Mouse Mammary Tumor Virus (MMTV) is a latently transforming retrovirus that is responsible for the neoplastic transformation of mammary epithelial cells. It is transmitted either exogenously as a retroviral particle in milk during nourishment of newborn mice or via the germ-line as an integrated provirus (90). MMTV gene expression is highest during lactation and is controlled by sequences in the 1.3 kb long terminal repeat (LTR) region in the 5′ half of the proviral DNA (90). In addition to WAP and β-casein gene expression, the expression of MMTV has also been considered a marker for mouse mammary epithelial cell differentiation (91).

Sequences between bases −1094 and −739 in the 5′ MMTV LTR have been reported to mediate a mammary epithelial cell-specific response (92). The region was therefore further investigated for DNA–protein interactions using DNase I footprinting analysis. Protected elements in mammary epithelial cells contained a half-palindromic binding site for NFI and a recognition site for the putative transcription factor, mammary cell-activated factor (MAF; now STAT5). These two transcription factors acted cooperatively when their binding sites were close to each other or when the binding sites were multimerized. As single entities, they were transcriptionally inactive. Bioinformatic analyses performed on the promoter regions of a number of milk protein genes showed similar clusters of STAT5 and NFI binding sites in the whey milk protein genes α-lactalbumin (bovine, human and rat), β-lactoglobulin (ovine), and WAP (mouse and rat). The NFI and STAT5 binding element sequences are concentrated in the first 200 bases upstream from the transcription start sites in each of these whey milk proteins. In some of the whey milk protein genes, such as WAP and α-lactalbumin (human and rat), an ACAAAG sequence is found in close proximity to the NFI and STAT5 recognition sites. The ACAAAG sequence was a determinant of mammary cell expression specificity for the MMTV–LTR. Therefore, Mink *et al.* concluded that differential mammary cell specific MMTV gene expression is achieved by a concerted action of multiple bound transcription factors, as has been determined for the milk protein genes (63). In agreement with other studies (60,88), these authors emphasized the possible presence in differentiated mammary epithelial cells of a mammary gland-specific NFI that binds to these half-palindromic sites. NFI participation in response to glucocorticoid signaling has been demonstrated on the MMTV–LTR in a range of studies.

In addition to being vital for the transcription of milk proteins and other differentiation-associated genes during pregnancy and lactation, NFI has been implicated in regulating the transcription of an involution-associated gene, TRPM2/clusterin/SGP2. Increased expression of TRPM2/clusterin has been shown to be associated with mammary gland involution and apoptosis of a number of different cell types (93–95). TRPM2/clusterin mRNA levels increase rapidly and substantially once involution is initiated in both the mouse mammary gland (96) and in the androgen-deprived rat ventral prostate (97). Furlong *et al*. hypothesized that the rapid stimulation of transcription of this gene during involution might permit the detection of an involution-enhanced association of transcription factors with its promoter (64). Comparative DNase I footprinting analyses performed with lactating and 2 day involuting nuclear mouse mammary gland extracts revealed two involution-associated, protected regions on this promoter. The first footprint, at bases −309 to −335, lies over a classical NFI binding element $TGGN_7CCA$ (12,98). The second protected region, which was principally detected when the assay was performed with involuting prostate extracts, covered a potential NFI binding site $TGGN_9CCA$. Subsequent EMSA analysis using a 46-mer double-stranded oligonucleotide based on the sequence of the twin NFI binding element, revealed changes in complex formation as the mammary gland proceeded from lactation to involution. The principal change seen with the involuting extracts was the appearance of a complex of intermediate size not seen in lactation. Western analysis of the proteins in the involution-associated complex revealed a 74 kDa NFI. Northern analysis demonstrated an NFI transcript of 1.6 kb that hybridized strongly and exclusively to RNA from involuting glands. Such a small transcript could not encode a 74 kDa protein, suggesting to Furlong *et al.* that the involution-associated 74 kDa NFI was significantly posttranslationally modified.

This proposal was further investigated by Kane *et al.* (65). Western analysis on 2-day involuting nuclear extracts showed that the 74 kDa NFI was an NFI-C isoform, characterized using an anti-NFI-C specific antibody α-8199 (65). A previous report by

Kane *et al.* indicated that increased expression of an NFI-C transcript accompanied involution (84). It was hypothesized that the involution-specific 74 kDa NFI-C was posttranslationally modified, as it had previously been reported that NFI-C can be modified by a unique type of *O*-glycosylation (a single *N*-acetylglucosamine residue bound by *O*-linkage to serine or threonine (99)). Bioinformatic analysis performed on the amino acid sequence of the NFI-C protein revealed a number of potential *N*- and *O*-glycosylation sites. The lectin concanavalin A (ConA) recognizes α-linked mannose, which occurs in the core structure of *N*-glycans, whereas wheat germ agglutinin (WGA) recognizes *O*-GlcNAc, found in *O*-linked modifications. By western analysis with the NFI-C-specific antibody α-8199 as a probe, it was shown that under nondenaturing conditions, the 74 kDa NFI-C protein in 2-day involuting nuclear extracts binds to ConA, but not to WGA. The 74 kDa NFI-C did not bind to ConA in the presence of an excess of the specific competing sugars, suggesting that it is *N*-glycosylated. In vivo studies in mice and in vitro studies using a cell culture model system were subsequently performed using tunicamycin, a potent inhibitor of *N*-glycosylation. Tunicamycin depots (Elvax) were placed in the right-hand fourth inguinal mammary gland immediately after initiating involution, and the glands were removed 24 h later. Systemic, as well as local effects were seen, reflected by a significant increase in fluid intake over the 24-h period. Western analysis of nuclear extracts from the right and left mammary glands of mice carrying the Elvax pellets showed the loss of the 74 kDa NFI-C. From these results it was concluded that involution-associated expression of the 74 kDa NFI could be inhibited in vivo by preventing *N*-glycosylation using tunicamycin. Importantly, the effect of tunicamycin was specific to the 74 kDa NFI-C, as no changes were reported with other factors, such as STAT3 and AP1. Primary mouse mammary epithelial cells grown as 3D spherical structures in a laminin-rich extracellular matrix (Matrigel) expressed the 74 kDa NFI-C for up to 7 days in culture. Addition of tunicamycin to these cultures again led to the inhibition of the expression of the 74 kDa NFI-C, further suggesting that this protein is posttranslationally modified by *N*-glycosylation (65). To our knowledge, this is the first *N*-glycosylated transcription factor that has been detected. It is clear that this modification does not affect the ability of the 74 kDa NFI-C to translocate to the nucleus and bind to NFI–DNA binding elements. It is likely that this *N*-linked modification induces a functional change at a recognition level and/or at an activity level. The *N*-glycosylation target site(s) is in the C-terminal region of NFI-C that houses the transcriptional activating/repressing domains. It is therefore likely that the function of this region will be most affected by the modification.

Streuli *et al.* reported the induction of expression of an NFI, reflected by the appearance of a complex in an EMSA analysis, by culturing primary mammary epithelial cells on a laminin-rich ECM substrate (66). These conditions of culture should support cell survival and terminal differentiation. Our recent studies now link expression of an NFI isoform to organization of mammary epithelial cells as 3D-spheres that simulate epithelial alveoli that, on terminal differentiation, are the source of milk components in lactation. Primary mouse mammary epithelial cells cultured on a concentrated laminin-rich EHS-derived ECM will organize to form 3D cellular aggregates (mammospheres (100)). These alveolus-like structures generate a central lumen, become polarized and, in the presence of lactogenic hormones, secrete milk proteins vectorially into the central luminal space. EMSA analysis demonstrated the formation of a smaller, faster-migrating NFI–DNA complex with mammosphere nuclear extracts, but not with extracts from cells cultured as EHS-associated monolayers (Fig. 2(A)). That this complex was NFI DNA-binding site-specific was shown by competing away its formation with a modest excess of unlabeled site-specific oligonucleotide while a nonspecific oligonucleotide failed to compete (Fig. 2(B)). Using an anti-NFI-C specific antibody, the mammosphere-associated complex was super-shifted, suggesting that it is at least partially comprised of an NFI-C isoform (Fig. 2(C)). In mammospheres, in addition to the expression of the previously described lactation-specific ~50 kDa NFI-C (83,84), we also observed the expression of a smaller, approximately 35 kDa NFI-C by western blot analysis using an anti-NFI-C-specific antibody α-8199 (Fig. 2(D)). This 35 kDa NFI-C is similarly found in extracts from lactating mammary gland (data not shown). We hypothesize that either one or both of these NFI-C isoforms is responsible for mammosphere-associated NFI–DNA complex formation. Formation of this complex is not prolactin-dependent (Fig. 2(A)). Thus, culturing mammary epithelial cells in an ECM that induces mammosphere formation triggers the production of an NFI-C isoform(s). Further studies on NFI-C in this model system should shed light on the role of NFI in terminally differentiated mammary epithelial cells.

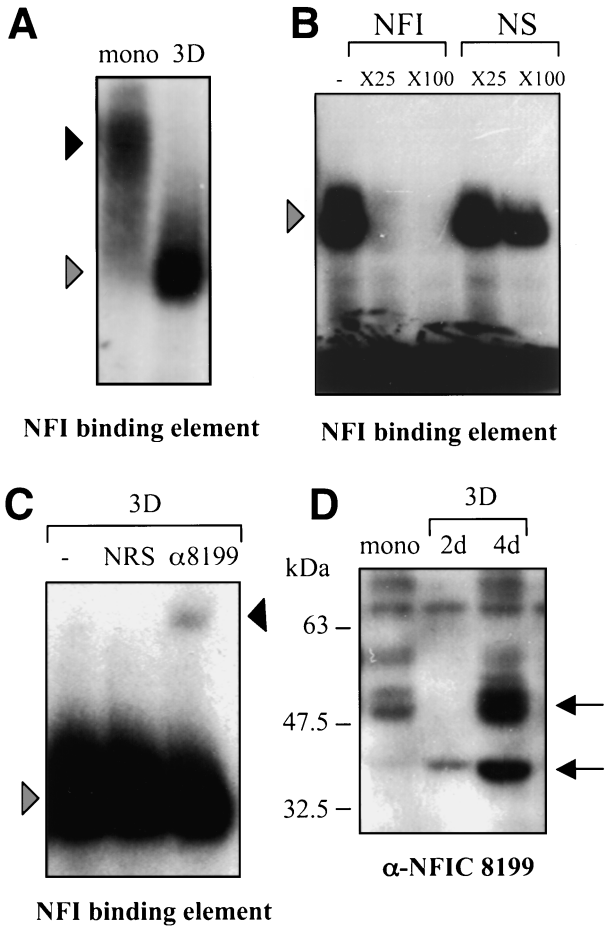

Fig. 2. (A) EMSA analysis of a radiolabelled NFI-binding element (twin NFI-binding element from proximal clusterin promoter (64)) with whole cell extracts of mouse mammary epithelial cells grown as a monolayer on ECM [Matrigel] diluted 1:1.4 with Ham's F12 (mono) or grown as mammospheres on concentrated ECM (3D). In this and subsequent panels the mammosphere-associated small NFI complex is indicated with a lightly-shaded arrowhead. (B) EMSA of extracts from mammospheres (conditions as in (A)) carried out in the absence or presence of a 25- or 100-fold excess of the unlabeled NFI DNA-binding-site-specific oligonucleotide or a 25- or 100-fold excess of a nonspecific binding element (NS, an AP-1 binding element). (C) Supershift analysis on mammosphere whole cell extracts. Prior to EMSA, mammosphere extracts were incubated with normal rabbit serum (NRS) or an anti-C-terminal NF1-C specific antibody (α-8199). The supershifted complex NFI is indicated with a black arrowhead. (D) Western analysis using the anti-C-terminal NFI-C specific antibody (α-8199) of whole-cell extracts of cells grown as a monolayer on plastic (mono), mammospheres grown on ECM with EGF and FCS for 2 days (2d) and mammospheres grown on ECM with prolactin for a further 2 days (4d).

CONCLUSIONS

The NFI transcription factors have been shown to be important for the regulated expression of several mammary-gland specific genes. NFI gene knockout studies emphasize the critical importance of the four individual NFI gene products in the development of various tissue structures. Complementing these are the interesting isoform-specific activities that have been reported as associated with mammary-specific activated promoters and with the different stages of postnatal mammary gland development. These data strongly suggest that the NFI transcription factor family plays multiple roles in mammary gland development and function. Future studies should focus on determining the roles of individual NFI isoforms in mammary gland function using mammary-specific NFI knockouts, mammary-inducible dominant-negative NFI transgenes, mammosphere cultures, and/or transplantation studies from NFI knockout mice into normal mice. Such analyses should clarify the specific roles of each NFI protein in mammary gland function.

ACKNOWLEDGMENTS

This work was supported in part by The Health Research Board, Ireland, and Enterprise Ireland (F.M.) and National Institutes of Health grants HD34908, DK58401, and DK48796 (R.M.G.). F.M. thanks Dr N. Tanese, NYU, for gifts of antibodies.

REFERENCES

1. P. J. Mitchell and R. Tjian (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245:**371–378.
2. R. M. Gronostajski (2000). Role of NFI/CTF gene family in transcription and development. *Gene* **249:**31–45.
3. C. Santoro, N. Mermod, P. Andrews, and R. Tjian (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: Cloning and expression of multiple cDNAs. *Nature* **334:**218–224.
4. K. Nagata, R. Guggenheimer, T. Enomoto, J. Lichy, and J. Hurwitz (1982). Adenovirus DNA replication *in vitro* Identification of a host factor that stimulates synthesis of the preterminal protein-dCMP complex. *PNAS* **79:**6438–6442.
5. K. Nagata, R. A. Guggenheimer, and J. Hurwitz (1983). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc. Natl. Acad. Sci. U.S.A.* **80:**6177–6181.
6. K. Nagata, R. A. Guggenheimer, and J. Hurwitz (1983). Adenovirus DNA replication in vitro: Synthesis of full-length DNA with purified proteins. *Proc. Natl. Acad. Sci. U.S.A.* **80:**4266–4270.
7. R. T. Hay (1985). Origin of adenovirus DNA replication: Role of the nuclear factor I binding site *in vivo*. *J. Mol. Biol.* **186:**129–136.

8. J. Bosher, I. R. Leith, S. M. Temperley, M. Wells, R. T. Hay (1991). The DNA-binding domain of nuclear factor I is sufficient to cooperate with the adenovirus type 2 DNA-binding protein in viral DNA replication. *J. Gen. Virol.* **72:**2975–2980.
9. R. N. de Jong and P. C. van der Vliet (1999). Mechanism of DNA replication in eukaryotic cells: Cellular host factors stimulating adenovirus DNA replication. *Gene* **236:**1–12.
10. M. T. Armentero, M. Horwitz, and N Mermod (1994). Targeting of DNA polymerase to the adenovirus origin of DNA replication by interaction with nuclear factor I. *Proc. Natl. Acad. Sci. U.S.A.* **91:**11537–11541.
11. R. M. Gronostajski, K. Nagata, and J. Hurwitz (1984). Isolation of human DNA sequences that bind to nuclear factor I, a host protein involved in adenovirus DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* **81:**4013–4017.
12. R. M. Gronostajski, S. Adhya, K. Nagata, R. A. Guggenheimer, and J. Hurwitz (1985). Site-specific DNA binding of nuclear factor I: Analyses of cellular binding sites. *Mol. Cell. Biol.* **5:**964–971.
13. J. A. Bogan, D. A. Natale, and M. L. Depamphilis (2000). Initiation of eukaryotic DNA replication: Conservative or liberal? *J. Cell Physiol.* **184:**139–150.
14. U. Borgmeyer, J. Nowock, and A. E. Sippel (1984). The TGGCA-binding protein: A eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded linear DNA. *Nucleic. Acids Res.* **12:**4295–4311.
15. P. A. J. Leegwater, P. C. van der Vliet, R. A. W. Rupp, J. Nowock, and A. E. Sippel (1986). Functional homology between the sequence-specific DNA-binding proteins nuclear factor I from HeLa cells and the TGGCA protein from chicken liver. *EMBO J.* **5:**381–386.
16. L. Hennighausen, U. Siebenlist, D. Danner, P. Leder, D. Rawlins, P. Rosenfeld, *et al.* (1985). High-affinity binding site for a specific nuclear protein in the human IgM gene. *Nature* **314:**289–292.
17. K. Jones, J. Kadonaga, P. Rosenfeld, T. Kelly, and R. Tjian (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48:**79–89.
18. A. Dorn, J. Bollekens, A. Staub, C. Benoist, and D. Mathis (1987). A multiplicity of CCAAT box-binding proteins. *Cell* **50:**863–872.
19. M. Raymondjean, S. Cereghini, and M. Yaniv (1988). Several distinct "CCAAT" box binding proteins coexist in eukaryotic cells. *Proc. Natl. Acad. Sci. U.S.A.* **85:**757–761.
20. R. Rupp, U. Kruse, G. Multhaup, U. Gobel, K. Beyreuther, and A. Sippel (1990). Chicken NFI/TGGCA proteins are encoded by at least three independent genes: NFI-A, NFI-B and NFI-C with homologues in mammalian genomes. *Nucleic. Acids Res.* **18:**2607–2616.
21. U. Kruse, F. Qian, and AE. Sippel (1991). Identification of a fourth Nuclear Factor I gene in chicken by cDNA cloning: NFIX. *Nucleic. Acids Res.* **19:**6641.
22. G. Gil, J. R. Smith, J. L. Goldstein, C. A. Slaughter, K. Orth, M. S. Brown, *et al.* (1988). Multiple genes encode nuclear factor 1-like proteins that bind to the promoter for 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc. Natl. Acad. Sci. U.S.A.* **85:**8963–8967.
23. T. Inoue, T. Tamura, T. Furuichi, and K. Mikoshiba (1990). Isolation of complementary DNAs encoding a cerebellum-enriched nuclear factor I family that activates transcription from the mouse myelin basic protein promoter. *J. Biol. Chem.* **265:**19065–19070.
24. A. Z. Chaudhry, C. E. Lyons, and R. M. Gronostajski (1997). Expression patterns of the four Nuclear Factor I genes during mouse embryogenesis indicate a potential role in development. *Dev. Dyn.* **208:**313–325.
25. M. Xu, S. Osada, M. Imagawa, and T. Nishihara (1997). Genomic organization of the rat nuclear factor I-A gene. *J. Biochem.* **122:**795–801.
26. F. Qian, U. Kruse, P. Lichter, and A. E. Sippel (1995). Chromosomal localization of the four genes NFIA, B, C, and X for the human transcription factor Nuclear Factor I by FISH. *Genomics* **28:**66–73.
27. D. Apt, Y. Liu, and H.-U. Bernard (1994). Cloning and functional analysis of spliced isoforms of human nuclear factor I–X: Interference with transcriptional activation by NFI/CTF in a cell-type specific manner. *NAR* **22:**3825–3833.
28. R. A. Holt, G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab, D. R. Nusskern, *et al.* (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298:**129–149.
29. P. Dehal, Y. Satou, R. K. Campbell, J. Chapman, B. Degnan, A. De Tomaso, *et al.* (2002). The draft genome of Ciona intestinalis: Insights into chordate and vertebrate origins. *Science* **298:**2157–2167.
30. S. Bandyopadhyay and R. M. Gronostajski (1994). Identification of a conserved oxidation-sensitive cysteine residue in the NFI family of DNA-binding proteins. *JBC* **269:**29949–29955.
31. U. Kruse and A. Sippel (1994). The genes for transcription factor nuclear factor I give rise to corresponding splice variants between vertebrate species. *J. Mol. Biol.* **238:**860–865.
32. L. das Neves, C. Duchala, F. Godinho, M. Haxhiu, C. Colmenares, W. Macklin, *et al.* (1999). Disruption of the murine Nuclear Factor I-A gene (*Nfia*) results in perinatal lethality, hydrocephalus and agenesis of the corpus callosum. *PNAS* **96:**11946–11951.
33. H. Baumeister, R. M. Gronostajski, G. E. Lyons, and F. L. Margolis (1999). Identification of NFI-binding sites and cloning of NFI-cDNAs suggest a regulatory role for NFI transcription factors in olfactory neuron gene expression. *Mol. Brain Res.* **72:**65–79.
34. C. F. Fletcher, N. A. Jenkins, N. G. Copeland, A. Z. Chaudhry, and R. M. Gronostajski (1999). Exon structure of the nuclear factor I DNA-binding domain from *C. elegans* to mammals. *Mamm. Genome* **10:**390–396.
35. S. Aparicio, J. Chapman, E. Stupka, N. Putnam, J. M. Chia, P. Dehal, *et al.* (2002). Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297:**1301–1310.
36. G. A. Elder, Z. Z. Liang, S. E. Snyder, and R. A. Lazzarini (1992). Multiple nuclear factors interact with the promoter of the human Neurofilament-M gene. *Mol. Brain Res.* **15:**99–107.
37. F. K. Bedford, D. Julius, and H. A. Ingraham (1998). Neuronal expression of the 5HT3 serotonin receptor gene requires nuclear factor 1 complexes. *J. Neurosci.* **18:**6186–6194.
38. T. Tamura, M. Masayuki, K. Ikenaka, and K. Mikoshiba (1988). Analysis of transcription control elements of the mouse myelin basic protein gene in HeLa cell extracts: Demonstration of a strong NFI-binding motif in the upstream region. *Nucleic Acids Res.* **16:**11441–11459.
39. P. Leahy, D. R. Crawford, G. Grossman, R. M. Gronostajski, and R. W. Hanson (1999). CREB binding protein coordinates the function of multiple transcription factors including

nuclear factor I to regulate phosphoenolpyruvate carboxykinase (GTP) gene transcription. *J. Biol. Chem.* **274:**8813–8822.
40. D. M. Faust, A. M. Catherin, S. Barbaux, L. Belkadi, T. Imaizumi-Scherrer, and M. C. Weiss (1996). The activity of the highly inducible mouse phenylalanine hydroxylase gene promoter is dependent upon a tissue-specific, hormone-inducible enhancer. *Mol. Cell. Biol.* **16:**3125–3137.
41. B. Bois-Joyeux and J. L. Danan (1994). Members of the CAAT/enhancer-binding protein, hepatocyte nuclear factor-1 and nuclear factor-1 families can differentially modulate the activities of the rat alpha-fetoprotein promoter and enhancer. *Biochem. J.* **301:**49–55.
42. C. J. Bachurski, S. E. Kelly, S. W. Glasser, and T. A. Currier (1997). Nuclear factor I family members regulate the transcription of surfactant protein-C. *J. Biol. Chem.* **272:**32759–32766.
43. S. S. Mukhopadhyay, S. L. Wyszomierski, R. M. Gronostajski, and J. M. Rosen (2001). Differential interactions of specific nuclear factor I isoforms with the glucocorticoid receptor and STAT5 in the cooperative regulation of WAP gene transcription. *Mol. Cell. Biol.* **21:**6859–6869.
44. T. Grewal, M. Theisen, U. Borgmeyer, T. Grussenmeyer, R. A. Rupp, A. Stief, *et al.* (1992). The -61-kilobase chicken lysozyme enhancer is a multifactorial complex containing several cell-type-specific elements. *Mol. Cell. Biol.* **12:**2339–2350.
45. Y. Xie, V. Madelian, J. Zhang, G. Ling, and X. Ding (2001). Activation of the NPTA element of the CYP2A3 gene by NFI-A2, a nasal mucosa-selective nuclear factor 1 isoform. *Biochem. Biophys. Res. Commun.* **289:**1225–1228.
46. K. Kido, H. Bannert, R. M. Gronostajski, and R. M. Flugel (2002). Bel1-mediated trans-activation of the spumaretroviral internal promoter is repressed by nuclear factor I. *J. Biol. Chem.*
47. P. Jethanandani and E. Goldberg (2001). ldhc expression in non-germ cell nuclei is repressed by NF-I binding. *J. Biol. Chem.* **276:**35414–35421.
48. C. S. Song, M. H. Jung, P. C. Supakar, B. Chatterjee, and A. K. Roy (1999). Negative regulation of the androgen receptor gene promoter by NFI and an adjacently located multiprotein-binding site. *Mol. Endocrinol.* **13:**1487–1496.
49. N. Mermod, E. O'Neill, T. Kelly, and R. Tjian (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* **58:**741–753.
50. G. Nebl and A. Cato (1995). NFI/X proteins: A class of NFI family of transcription factors with positive and negative regulatory domains. *Cell. Mol. Biol. Res.* **41:**85–95.
51. U. Kruse and A. E. Sippel (1994). Transcription factor nuclear factor I proteins form stable homo- and heterodimers. *FEBS Lett.* **348:**46–50.
52. A. Z. Chaudhry, A. D. Vitullo, and R. M. Gronostajski (1998). Nuclear factor I (NFI) isoforms differentially activate simple versus complex NFI-responsive promoters. *J. Biol. Chem.* **273:**18538–18546.
53. A. Z. Chaudhry, A. D. Vitullo, and R. M. Gronostajski (1999). Nuclear factor I-mediated repression of the mouse mammary tumor virus promoter is abrogated by the coactivators p300/CBP and SRC-1. *J. Biol. Chem.* **274:**7072–7081.
54. B. Gao and G. Kunos (1998). Cell type-specific transcriptional activation and suppression of the alpha1B adrenergic receptor gene middle promoter by nuclear factor 1. *J. Biol. Chem.* **273:**31784–31787.
55. A. Grunder, T. T. Ebel, M. Mallo, G. Schwarzkopf, T. Shimizu, A. E. Sippel, *et al.* (2002). Nuclear factor I-B (*Nfib*) deficient mice have severe lung hypoplasia. *Mech. Dev.* **112:**69–77.
56. G. Steele-Perkins, K. G. Butz, G. E. Lyons, M. Zeichner-David, H.-J. Kim, M. I. Cho, *et al.* (2003). Essential role for NFI-C/CTF transcription-replication factor in tooth root development. *MCB* **23:**1075–1084.
57. T. Shu, K. G. Butz, C. Plachez, R. M. Gronostajski, and L. J. Richards (2003). Abnormal development of forebrain midline glia and commissural projections in *Nfia* knock-out mice. *J. Neurosci.* **23:**203–212.
58. J. P. Simons, M. McClenaghan, and A. J. Clark (1987). Alteration of the quality of milk by expression of sheep beta-lactoglobulin in transgenic mice. *Nature* **328:**530–532.
59. H. Lubon and L. Hennighausen (1988). Conserved region of the rat alpha-lactalbumin promoter is a target site for protein binding in vitro. *Biochem. J.* **256:**391–396.
60. C. J. Watson, K. E. Gordon, M. Robertson, and A. J. Clark (1991). Interaction of DNA-binding proteins with a milk protein gene promoter in vitro—Identification of a mammary gland-specific factor. *Nucleic Acids Res.* **19:**6603–6610.
61. B. Rajput, N. L. Shaper, and J. H. Shaper (1996). Transcriptional regulation of murine beta1,4-galactosyltransferase in somatic cells. Analysis of a gene that serves both a housekeeping and a mammary gland-specific function. *J. Biol. Chem.* **271:**5131–5142.
62. M. Kannius-Janson, U. Lidberg, K. Hulten, A. Gritli-Linde, G. Bjursell, and J. Nilsson (1998). Studies of the regulation of the mouse carboxyl ester lipase gene in mammary gland. *Biochem. J.* **336**(Pt 3):577–585.
63. S. Mink, E. Hartig, P. Jennewein, W. Doppler, and A. C. B. Cato (1992). A mammary cell-specific enhancer in MMTV DNA is composed of multiple regulatory elements including binding sites For CTF/NFI and a novel transcription factor, Mammary Cell-Activating Factor. *Mol. Cell. Biol.* **12:**4906–4918.
64. E. Furlong, N. K. Keon, F. D. Thornton, T. Rein, and F. Martin (1996). Expression of a 74-kDa nuclear factor 1 (NF1) protein is induced in mouse mammary gland involution. Involution-enhanced occupation of a twin NF1 binding element in the testosterone-repressed prostate message-2/clusterin promoter. *J. Biol. Chem.* **271:**29688–29697.
65. R. Kane, J. Murtagh, D. Finlay, A. Marti, R. Jaggi, D. Blatchford, *et al.* (2002). Transcription factor NFIC undergoes N-glycosylation during early mammary gland involution. *J. Biol. Chem.* **277:**25893–25903.
66. C. H. Streuli, G. M. Edwards, M. Delcommenne, C. B. Whitelaw, T. G. Burdon, C. Schindler, *et al.* (1995). Stat5 as a target for regulation by extracellular matrix. *J. Biol. Chem.* **270:**21639-21644.
67. P. K. Qasba, A. M. Dandekar, T. M. Horn, I. Losonczy, M. Siegel, K. A. Sobiech, *et al.* (1982). Milk protein gene expression in the rat mammary gland. *Crit. Rev. Food Sci. Nutr.* **16:**165–186.
68. J. W. Perry and T. Oka (1984). The study of differentiative potential of the lactating mouse mammary gland in organ culture. *In Vitro* **20:**59–65.
69. Y. J. Topper, K. R. Nicholas, L. Sankaran, and J. Kulski (1984). Insulin as a developmental hormone. *Prog. Clin. Biol. Res.* **142:**63–77.
70. M. Ono and T. Oka (1980). The differential actions of cortisol on the accumulation of alpha-lactalbumin and casein in mid-pregnant mouse mammary gland in culture. *Cell* **19:**473–480.

71. M. Ono, J. W. Perry, and T. Oka (1981). Concentration-dependent differential effects of cortisol on synthesis of alpha-lactalbumin and of casein in cultured mouse mammary gland explants: Importance of prolactin concentration. *In Vitro* **17:**121–128.
72. S. J. Quirk, J. E. Gannell, M. J. Fullerton, and J. W. Funder (1986). Mechanisms of biphasic action of glucocorticoids on alpha-lactalbumin production by rat mammary gland explants. *J. Steroid Biochem.* **24:**413–416.
73. N. Terada, L. J. Leiderman, and T. Oka (1983). The interaction of cortisol and prostaglandins on the phenotypic expression of the alpha-lactalbumin gene in the mouse mammary gland in culture. *Biochem. Biophys. Res. Commun.* **111:**1059–1065.
74. L. Hall, D. C. Emery, M. S. Davies, D. Parker, and R. K. Craig (1987). Organization and sequence of the human alpha-lactalbumin gene. *Biochem. J.* **242:**735–742.
75. R. Gronostajski (1987). Site-specific DNA binding of nuclear factor I: Effect of the spacer region. *Nucleic. Acids Res.* **15:**5545–5559.
76. S. Harris, M. McClenaghan, J. P. Simons. S. Ali, and A. J. Clark (1991). Developmental regulation of the sheep beta-lactoglobulin gene in the mammary gland of transgenic mice. *Dev. Genet.* **12:**299–307.
77. C. B. Whitelaw, S. Harris, M. McClenaghan, J. P. Simons, and A. J. Clark (1992). Position-independent expression of the ovine beta-lactoglobulin gene in transgenic mice. *Biochem. J.* **286**(Pt 1):31–39.
78. R. M. McKenzie and B. L. Larson (1978). Isolation and partial characterization of rare casein proteins. *J. Dairy Sci.* **61:**885–889.
79. S. Li and J. M. Rosen (1995). Nuclear factor I and mammary gland factor (STAT5) play a critical role in regulating rat whey acidic protein gene expression in transgenic mice. *Mol. Cell. Biol.* **15:**2063–2070.
80. J. F. Levine and F. E. Stockdale (1985). Cell-cell interactions promote mammary epithelial cell differentiation. *J. Cell. Biol.* **100:**1415–1422.
81. T. C. Dale, M. J. Krnacik, C. Schmidhauser, C. L. Yang, M. J. Bissell, and J. M. Rosen (1992). High-level expression of the rat whey acidic protein gene is mediated by elements in the promoter and 3′ untranslated region. *Mol. Cell. Biol.* **12:**905–914.
82. S. Li and J. M. Rosen (1994). Distal regulatory elements required for rat whey acidic protein gene expression in transgenic mice. *J. Biol. Chem.* **269:**14235–14243.
83. M. Kannius-Janson, E. M. Johansson, G. Bjursell, and J. Nilsson (2002). Nuclear factor 1-C2 contributes to the tissue-specific activation of a milk protein gene in the differentiating mammary gland. *J. Biol. Chem.* **277:**17589–17596.
84. R. Kane, D. Finlay, T. Lamb, and F. Martin (2000). Transcription factor NF 1 expression in involuting mammary gland. *Adv. Exp. Med. Biol.* **480:**117–122.
85. R. L. Hill, K. Brew, T. C. Vanaman, I. P. Trayer, and P. Mattock (1968). The structure, function, and evolution of alpha-lactalbumin. *Brookhaven Symp. Biol.* **21:**139–154.
86. R. W. Turkington, K. Brew, T. C. Vanaman, and R. L. Hill (1968). The hormonal control of lactose synthetase in the developing mouse mammary gland. *J. Biol. Chem.* **243:**3382–3387.
87. A. Harduin-Lepers, N. L. Shaper, J. A. Mahoney, and J. H. Shaper (1992). Murine beta 1,4-galactosyltransferase: Round spermatid transcripts are characterized by an extended 5′-untranslated region. *Glycobiology* **2:**361–368.
88. A. Harduin-Lepers, J. H. Shaper, and N. L. Shaper (1993). Characterization of two cis-regulatory regions in the murine beta 1,4-galactosyltransferase gene. Evidence for a negative regulatory element that controls initiation at the proximal site. *J. Biol. Chem.* **268:**14348–14359.
89. A. S. Lidmer, M. Kannius, L. Lundberg, G. Bjursell, and J. Nilsson (1995). Molecular cloning and characterization of the mouse carboxyl ester lipase gene and evidence for expression in the lactating mammary gland. *Genomics* **29:**115–122.
90. N. E. Hynes, B. Groner, and R. Michalides (1984). Mouse mammary tumor virus: Transcriptional control and involvement in tumorigenesis. *Adv. Cancer Res.* **41:**155–184.
91. B. Munoz and F. F. Bolander Jr. (1989). Prolactin regulation of mouse mammary tumor virus (MMTV) expression in normal mouse mammary epithelium. *Mol. Cell Endocrinol.* **62:**23–29.
92. S. Mink, H. Ponta, and A. C. Cato (1990). The long terminal repeat region of the mouse mammary tumour virus contains multiple regulatory elements. *Nucleic Acids Res.* **18:**2017–2024.
93. J. A. Sensibar, M. D. Griswold, S. R. Sylvester, R. Buttyan, C. W. Bardin, C. Y. Cheng, *et al.* (1991). Prostatic ductal system in rats: Regional variation in localization of an androgen-repressed gene product, sulfated glycoprotein-2. *Endocrinology* **128:**2091–2102.
94. I. S. Sawczuk, G. Hoke, C. A. Olsson, J. Connor, and R. Buttyan (1989). Gene expression in response to acute unilateral ureteral obstruction. *Kidney Int.* **35:**1315–1319.
95. S. Bettuzzi, L. Troiano, P. Davalli, F. Tropea, M. C. Ingletti, E. Grassilli, *et al.* (1991). In vivo accumulation of sulfated glycoprotein 2 mRNA in rat thymocytes upon dexamethasone-induced cell death. *Biochem. Biophys. Res. Commun.* **175:**810–815.
96. R. Strange, F. Li, S. Saurer, A. Burkhardt, and R. R. Friis (1992). Apoptotic cell death and tissue remodelling during mouse mammary gland involution. *Development* **115:**49–58.
97. R. Buttyan, Z. Zakeri, R. Lockshin, and D. Wolgemuth (1988). Cascade induction of c-fos, c-myc, and heat shock 70K transcripts during regression of the rat ventral prostate gland. *Mol. Endocrinol.* **2:**650–657.
98. R. T. Hay (1985). The origin of adenovirus DNA replicaton: Minimal DNA sequence requirement *in vivo*. *EMBO J.* **4:**421–426.
99. S. P. Jackson and R. Tjian (1988). O-glycosylation of eukaryotic transcription factors: Implications for mechanisms of transcriptional regulation. *Cell* **55:**125–133.
100. D. R. Blatchford, L. H. Quarrie, E. Tonner, C. McCarthy, D. J. Flint, and C. J. Wilde (1999). Influence of microenvironment on mammary epithelial cell survival in primary culture. *J. Cell. Physiol.* **181:**304–311.