HAL open science

Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response.

H Wakao, F. Gouilleux, B Groner

To cite this version:

H Wakao, F. Gouilleux, B Groner. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response.. EMBO Journal, 1994, 13 (9), pp.2182-2891. 10.1002/j.1460-2075.1994.tb06495.x . hal-02427591

HAL Id: hal-02427591

https://univ-tours.hal.science/hal-02427591

Submitted on 30 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The EMBO Journal vol.13 no.9 pp.2182–2191, 1994

Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response

Hiroshi Wakao, Fabrice Gouilleux and Bernd Groner[1]

Friedrich Miescher Institute, Schwarzwaldallee 215, PO Box 2543, CH-4002 Basel, Switzerland and [1]Institute for Experimental Cancer Research, Tumor Biology Center, Breisacher Strasse 117, D-79106 Freiburg, Germany
[1]Corresponding author

Communicated by B.Groner

Milk protein gene expression in mammary epithelial cells is regulated by the action of the lactogenic hormones insulin, glucocorticoids and prolactin. The mammary gland factor, MGF, has been shown to be a central mediator in the lactogenic hormone response. The DNA binding activity of MGF is hormonally regulated and essential for β-casein promoter activity. We have used Red A Sepharose- and sequence-specific DNA affinity chromatography to purify MGF from mammary gland tissue of lactating sheep. Proteins of 84 and 92 kDa were obtained, proteolytically digested and the resulting peptides separated by reverse phase high pressure liquid chromatography. The 84 and 92 kDa proteins yielded very similar peptide patterns. The amino acid sequence of two peptides was determined. The sequence information was used to derive oligonucleotide probes. A cDNA library from the mRNA of mammary gland tissue of lactating sheep was screened and a molecular clone encoding MGF was isolated. MGF consists of 734 amino acids and has sequence homology with the 113 (Stat113) and 91 kDa (Stat91) components of ISGF3, transcription factors which are signal transducers of IFN-α/β and IFN-γ. Two species of MGF mRNA of 6.5 and 4.5 kb were detected in mammary gland tissue of lactating sheep. Lower mRNA expression was found in ovary, thymus, spleen, kidney, lung, muscle and the adrenal gland. MGF cDNA was incorporated into a eukaryotic expression vector and cotransfected with a vector encoding the long form of the prolactin receptor into COS cells. A strong MGF-specific bandshift was obtained with nuclear extracts of COS cells induced with prolactin. Treatment of activated MGF with a tyrosine-specific protein phosphatase resulted in the loss of DNA binding activity. Prolactin-dependent transactivation of a β-casein promoter-luciferase reporter gene construct was observed in transfected cells.

Key words: cytokine receptors/signal transduction/transactivation/tyrosine phosphorylation

Introduction

Prolactin exerts pleiotropic biological effects in vertebrates and is involved in the regulation of reproduction, cell growth, lactation, steroidogenesis, maternal behaviour, osmoregulation, metabolism and the immune response (Russell, 1989; Kelly *et al.*, 1991). Prolactin is best known for its role in the development of the mammary gland and its control of lactation. It promotes milk protein gene expression by regulating the transcription and the stability of mRNA (Vonderhaar and Ziska, 1989). The role of prolactin in the immune system has also been studied extensively (Gagnerault *et al.*, 1993; Hooghe *et al.*, 1993; Murphy *et al.*, 1993).

Prolactin is a peptide hormone produced in the anterior pituitary and a member of the growth hormone/prolactin/placental lactogen gene family. It exerts its function via the prolactin receptor (PRLR), a member of the cytokine/hematopoietic growth factor receptor family. Growth hormone (GH) (Leung *et al.*, 1987), erythropoietin (EPO) (D'Andrea *et al.*, 1991), IL-2 (Hatakeyama *et al.*, 1989), IL-3 (Itoh *et al.*, 1990), IL-4 (Mosley *et al.*, 1989), IL-5 (Takaki *et al.*, 1990), IL-6 (Yamasaki *et al.*, 1988), IL-7 (Goodwin *et al.*, 1990), granulocyte-macrophage colony-stimulating factor (GM-CSF) (Gearing *et al.*, 1989), granulocyte colony-stimulating factor (G-CSF) (Fukunaga *et al.*, 1990), leukemia-inhibiting factor (LIF) (Gearing *et al.*, 1991) and ciliary neurotrophic factor (CNTF) (Davis *et al.*, 1991) act via receptors belonging to this family. The extracellular domains of these receptors share structural similarities (Cosman *et al.*, 1990) and contain the protein motif W-S-X-W-S. They also contain four paired cysteine residues possibly important for ligand binding (Bazan, 1990; Kelly *et al.*, 1991; Miyazaki *et al.*, 1991). Among the ligands, only growth hormone and prolactin are structurally related.

Prolactin triggers cellular responses by binding to a cell surface receptor. PRLR has been cloned from different species (Boutin *et al.*, 1988, 1989; Davis and Linzer, 1989; Edery *et al.*, 1989; Shirota *et al.*, 1990; Zhang *et al.*, 1990). There are isoforms of PRLR, a short form with 57 amino acids of intracellular domain, and a longer form with 358 amino acids. Both forms have identical extracellular and transmembrane domains and can be expressed in the same tissue (Shirota *et al.*, 1990). The comparison of the biological function as signal transducers has shown that the isoforms are distinguishable (Lesueur *et al.*, 1991). A third, intermediate receptor form was found in Nb2, pre-T rat lymphoma cells. Nb2 cells require prolactin for growth. This PRLR is lacking 198 amino acids in its cytoplasmic domain when compared with the long form (Ali *et al.*, 1991).

In contrast to PDGF and EGF receptors, which contain intrinsic tyrosine kinase activities (Schlessinger and Ullrich, 1993), members of the cytokine/hematopoietic growth factor receptor family do not have such an enzymatic activity. Nevertheless, protein tyrosine phosphorylation has been implicated in the signalling cascades of these receptors. The tyrosine-specific protein kinase JAK2 is associated with the EPO receptor and the GH receptor and is also activated by

© Oxford University Press

prolactin (Argetsinger *et al.*, 1993; Witthuhn *et al.*, 1993; Rui *et al.*, 1994).

The mechanism by which gene transcription is regulated by prolactin has been studied most extensively in mammary epithelial cells. A lactogenic hormone response element within the proximal 335 bp of the rat β-casein gene promoter has been found (Doppler *et al.*, 1989). Prolactin and glucocorticoids synergistically induce β-casein gene transcription. The binding site of a transcription factor (mammary gland factor, MGF) was identified in the β-casein gene promoter (Schmitt-Ney *et al.*, 1991). This sequence, 5′-ACTTCTTGGAATT-3′, is highly conserved in the promoter region of the casein genes (Yoshimura and Oka, 1989; Wakao *et al.*, 1992). MGF binding is indispensable for β-casein gene promoter activity. MGF is a transducer of lactogenic hormone and environmental signals. Its activity is highly regulated during gestation, lactation and post-lactation. Little DNA binding activity is detected at early stages of pregnancy. Its activity increases towards the end of pregnancy. At the end of gestation, high levels of MGF activity are observed and maintained during the lactation period. At post-lactation, very little activity is detected. During lactation the level of MGF activity is controlled in part through signals induced by the suckling of the pups (Schmitt-Ney *et al.*, 1992), most likely by the level of circulating prolactin.

In this paper, we describe the purification of MGF from mammary gland tissue of lactating sheep and the cloning of its cDNA. The protein sequence of MGF exhibits similarity to ISGF3 proteins, mediators of the interferon response (Fu *et al.*, 1992; Schindler *et al.*, 1992b). These proteins have been termed 'Stat' for signal transducer and activator of transcription (Shuai *et al.*, 1993a). After activation through ligand interaction with their respective receptors, the Stats translocate into the nucleus where they induce gene transcription by binding to DNA response elements. Another member of this family is possibly the acute phase response factor (APRF) activated by IL-6 (Wegenka *et al.*, 1993; Standke *et al.*, 1994). Activation of these signal transducers involves tyrosine phosphorylation (Shuai *et al.*, 1993b; Wegenka *et al.*, 1993). As in the case of ISGF3 proteins and APRF, phosphorylation on tyrosine is necessary for MGF DNA binding activity. Transfected MGF is activated by prolactin, and induces the transcription of a reporter gene containing an MGF binding site in its promoter. These data indicate that the sequence to which MGF binds is a prolactin response element (PLRE) and MGF is a PLRE binding protein.

Results

Purification of MGF from mammary gland tissue of lactating sheep

We have described previously the purification of MGF from rat mammary gland (Wakao *et al.*, 1992). Because the amount of protein obtained was too small to allow amino acid sequence determination, we have prepared nuclear extracts from the mammary tissue of lactating sheep as a source for the purification of MGF and modified the purification procedure. MGF was purified by Red A- and sequence-specific DNA affinity column chromatography. The purification was monitored by the bandshift activity of MGF.

Protein (212 mg) was obtained from 6000 g of sheep mammary gland after extraction of nuclei with a buffer containing 400 mM NaCl (buffer D). The detergent NP-40 was added to a final concentration of 0.1% (v/v) before dialysis and nearly 100% of the MGF activity was recovered. Very low MGF activity was recovered in the absence of NP-40. The specific MGF activity of the nuclear extract after dialysis was determined in a bandshift assay and found to be 1037 U/mg of protein. One unit of MGF activity was defined as the amount which causes the retardation of 1 fmol of labelled DNA probe.

Dialysed nuclear extracts were loaded onto a column of Red A resin (Amicon) and eluted with a linear gradient of KSCN in buffer D′ (same as buffer D except KSCN was replaced by NaCl). KSCN was used to elute the column because poor recovery resulted after NaCl elution. Elution with 3 M NaCl yielded <10% of the applied activity (data not shown). MGF activity was eluted between 100 and 180 mM of KSCN (Figure 1) and a 13-fold purification with a yield of 76% was obtained.

Sequence-specific DNA affinity chromatography was used for further purification. The promoter sequence of ovine β-casein is not yet known. Therefore we compared two MGF binding sequences, one from the rat β-casein gene (5′-GGA-CTTCTTGGAATTAAGGGA-3′) and one from the bovine β-casein gene (5′-AGATTTCTAGGAATTCAATCC-3′) (Gorodetsky *et al.*, 1988) for their ability to bind sheep MGF. The factor bound more efficiently to the bovine sequence (data not shown). The bovine β-casein MGF binding sequence was used in the bandshift and the affinity purification experiments. MGF activity was eluted from the sequence-specific DNA affinity column with buffer E containing 600 mM NaCl (Figure 2A) and a >14 464-fold purification resulted with a yield of 68% (Table I). The amount of purified protein was too small to be determined accurately. Two bands of 84 and 92 kDa were revealed when the eluates of 600 mM NaCl from the DNA affinity column were analysed on SDS–PAGE and after silver staining (Figure 2B). Both proteins possessed similar DNA binding affinity. No further purification was achieved by selection

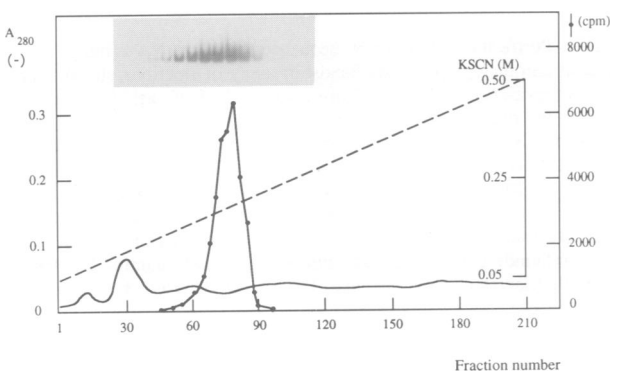

Fig. 1. Purification of MGF from mammary gland tissue of lactating sheep. Proteins of nuclear extracts were fractionated by Matrex Gel Red A column chromatography. Unbound proteins were washed away with NaCl and KSCN as described in Material and methods. After development of the column with a linear gradient of KSCN (broken line), MGF activity was assayed by bandshift analysis, shown in the insert. The radioactivity in the DNA–MGF complexes was determined (○). 1 ml of the concentrated fraction was mixed with poly[d(I,C)] and poly[d(A,T)] and applied to the DNA affinity column.

on a DNA affinity column containing a point mutated MGF binding sequence followed by a column containing the wild type sequence (data not shown).

Isolation of a cDNA clone encoding MGF

The purified proteins of 84 and 92 kDa were separated by SDS gel electrophoresis and transferred onto a nitrocellulose membrane. The proteins were visualized by staining and individually excised from the membrane. The proteins were digested *in situ* with the protease Lyn C. The resulting peptides were separated by reverse phase high pressure liquid chromatography. The peptide profiles of the two proteins obtained by reverse phase chromatography were almost identical (data not shown). Peaks were pooled and subjected to amino acid sequence analysis. The amino acid residues (0.6–2.2 pmol) were recovered after the sequencing reactions. Amino acid sequence was obtained for two peptides (underlined in Figure 3). Based on the bias of human codon usage, a 42 bp oligonucleotide guessmer (see Materials and methods) representing one of the peptides was synthesized and used as a probe to screen a cDNA library. The library was made from mRNA derived from mammary gland tissue of lactating sheep. Twenty positive clones were obtained (frequency of 0.004%).

Restriction enzyme analysis and partial DNA sequencing revealed that all clones represent the same mRNA. The sequence of clone 10 was determined. This clone contains a cDNA insert of 2816 bp and a coding region of 2202 nt. The cDNA and the deduced amino acid sequence of clone 10 is shown in Figure 3. An open reading frame begins at position 244 and ends at position 2415. The open reading frame comprises 734 amino acids. The predicted molecular weight corresponds well with the observed molecular weight of the purified proteins. The two peptide sequences identified by protein microsequencing are present in this open reading frame (Figure 3, underlined). No poly(A) sequence was observed within the 3′ non-translated region, suggesting the occurrence of a natural *Not*I site.

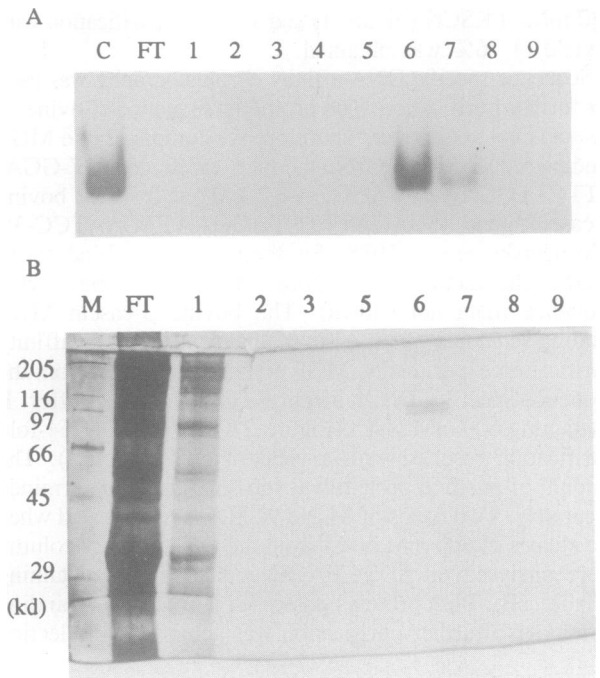

Fig. 2. Purification of MGF by sequence-specific DNA affinity column chromatography. (**A**) Bandshift assay of fractions eluted from the sequence-specific DNA affinity column. 1 μl of each fraction was used in the binding reaction. Lane C, pool of positive fractions from the Matrex Gel Red A column; lane FT, flowthrough after the third loading. Lanes 1–4, wash with 180 mM NaCl; lanes 5–9, eluate with 600 mM NaCl. (**B**) SDS–PAGE of fractions from the sequence-specific DNA affinity column. 100 μl of the wash and eluate fractions were precipitated with TCA (17%, v/v) and analysed on a 7.5% polyacrylamide gel. Proteins are stained with silver staining kit (BRL). M, molecular weight marker lane; lanes 1–9 correspond to the lanes shown in (A).

MGF has sequence homologies with the interferon-stimulated gene factors Stat91 and Stat113

Comparison of the MGF protein sequence with proteins in the EMBL database showed a distinct homology to the human ISGF3-a proteins (Fu *et al.*, 1992; Schindler *et al.*, 1992b). The overall identity between MGF and ISGF3 113 kDa and MGF and ISGF3 91 kDa is 28.5 and 29.2% respectively. Like the ISGF3 proteins, MGF contains a src homology 2 (SH2)-like domain and a src homology 3 (SH3)-like domain (Stahl *et al.*, 1988; Koch *et al.*, 1989). The SH2 and SH3 domains of MGF and those of ISGF3-a proteins are compared in Figure 4. Although MGF and the ISGF3-a proteins share homologies throughout, they also exhibit distinct differences. The leucine zipper motif found in the ISGF3-a proteins is not present in the MGF sequence. MGF contains a putative tyrosine phosphorylation site at position 694 which is in the same relative position as the Stat91 and Stat113 (VDGYVKP, Shuai *et al.*, 1993b). This tyrosine could be phosphorylated by a protein tyrosine kinase (PTK) and regulate the MGF DNA binding activity (see below). A highly conserved motif, GTFLLRFSXS, is found in the potential SH2 domain of MGF.

MGF expression in sheep organs

Northern blot analysis of mRNA from different tissues was performed. The highest expression of MGF mRNA was detected in sheep mammary gland. Two transcripts of 6.5 and 4.5 kb were found (Figure 5, lanes 10 and 11). The smaller form of 4.5 kb was predominant. MGF transcripts

Table I. Summary of the MGF purification from mammary gland of lactating sheep

Fraction	Protein (mg)	Activity (U)	Specific activity (U/mg)	Yield (%)	Purification (fold)
Mammary gland	6000.0 (g)	–	–	–	–
Nuclear extract	212.0	220 000	1037	100	1
Red A Sepharose	12.6	170 000	13 492	76	13
Specific DNA affinity	<0.01	150 000	>15 000 000	68	>14 464

Total yield and purification factors have been calculated based on the content of MGF in the nuclear extract.

```
   1  GAATTCCGTTGCTGTCGCGTGGGATGCCATCGACCTGGACAATCCCCAGGAGCCCCGGCCGGAGCGAGCGCCGGCCCCGT   80
  81  CGCCGACCGCCCGGCACGACCAGCCCTCCCTCCTGAACACCCCAGCCACGTGCCAGCGGTTGCCGAGCCGTCAGGGGAAA  160
 161  CCTTCGGTGACCCTAGCTTTGGGCTGGAATATCCCCGCGTTCCCAGAAACGGAGACTCGGAGTTCGAAGCAGGGTGAACG  240
 241  GACATGGCGGGCTGGATCCAGGCCCAGCAGCTGCAGGGAGATGCCCTGCGCCAGATGCAGGTGCTATACGGGCAGCACTT  320
   1     M  A  G  W  I  Q  A  Q  Q  L  Q  G  D  A  L  R  Q  M  Q  V  L  Y  G  Q  H  F   26

 321  CCCCATCGAGGTCCGGCATTACTTGGCGCAGTGGATTGAGAGCCAGCCGTGGGATGCCATCGACCTGGACAATCCCCAGG  400
  27   P  I  E  V  R  H  Y  L  A  Q  W  I  E  S  Q  P  W  D  A  I  D  L  D  N  P  Q  D   53

 401  ACCGGGCCCAGGTCACCCAGCTCCTGGAGGGCCTGGTGCAGGAGTTGCAAAAGAAGGCAGAGCACCAAGTCGGGGAAGAC  480
  54   R  A  Q  V  T  Q  L  L  E  G  L  V  Q  E  L  Q  K  K  A  E  H  Q  V  G  E  D   79

 481  GGGTTCCTACTGAAGATCAAGCTGGGGCACTACGTCCACGTCAGCTCCAGAACACGTACGACCGCTGCCCCATGGAGCTG  560
  80  G  F  L  L  K  I  K  L  G  H  Y  V  H  V  S  S  R  T  R  T  T  A  A  P  W  S  W  106

 561  GTTGCGCTGCATTCGCCACATTCTGTACAATGAACAGAGGCTGGTCCGAGAAGCCACCAATGGTAATTCTTCTGCTGGGA  640
 107   L  R  C  I  R  H  I  L  Y  N  E  Q  R  L  V  R  E  A  T  N  G  N  S  S  A  G  I  133

 641  TCCTGGTTGATGCCATGTCCCAGAAACACCTTCAGATCAACCAGACATTTGAGGAGCTGCGACTGGTCACACAAGACACA  720
 134   L  V  D  A  M  S  Q  K  H  L  Q  I  N  Q  T  F  E  E  L  R  L  V  T  Q  D  T   159

 721  GAGAATGAGCTGAAGAAGCTGCAGCAGACTCAAGAGTATTTCATCATCCAGTATCAGGAGAGCCTGAGGATCCAGGCTCA  800
 160  E  N  E  L  K  K  L  Q  Q  T  Q  E  Y  F  I  I  Q  Y  Q  E  S  L  R  I  Q  A  Q   186

 801  GTTTGCCCAGCTGGCCCAGCTGAACCCCCAGGAGCGACTGAGCCGGGAGACGGCCCTCCAGCAGAAGCAGGTGTCCCTGG  880
 187   F  A  Q  L  A  Q  L  N  P  Q  E  R  L  S  R  E  T  A  L  Q  Q  K  Q  V  S  L  E  213

 881  AGGCCTGGCTGCAGCGCGAGGCCCAGACGCTGCAGCAGTACCGCGTGGAGCTGGCTGAGAAGCACCAGAAGACCCTGCAG  960
 214   A  W  L  Q  R  E  A  Q  T  L  Q  Q  Y  R  V  E  L  A  E  K  H  Q  K  T  L  Q   239

 961  CTGCTGCGGAAGCAGCAGACCATCATCCTGGATGACGAGCTGATCCAGTGGAAGCGGCGGCACGACTGGCGGGGAATGGA 1040
 240  L  L  R  K  Q  Q  T  I  I  L  D  D  E  L  I  Q  W  K  R  R  H  D  W  R  G  M  E   266

1041  GGCCCCCCCGAGGAGCCTGGATGTGCTACAGTCCTGGTGTGAGAAGTTGGCGGAGATCATCTGGCAGAACCGGCAGCAGA 1120
 267   A  P  P  R  S  L  D  V  L  Q  S  W  C  E  K  L  A  E  I  I  W  Q  N  R  Q  Q  I  293

1121  TCCGCAGAGCCGAGCACCTCTGCCAGCAGCTGCCCATCCCCGGCCCCGTGGAGGAGATGCTGGCTGAGGTCAACGCCACC 1200
 294   R  R  A  E  H  L  C  Q  Q  L  P  I  P  G  P  V  E  E  M  L  A  E  V  N  A  T   319

1201  ATCACGGACATCATCTCAGCCCTGGTGACCAGCACATTCATCATCGAGAAGCAGCCCCCTCAGGTCCTGAAGACCCAGAC 1280
 320  I  T  D  I  I  S  A  L  V  T  S  T  F  I  I  E  K  Q  P  P  Q  V  L  K  T  Q  T   346

1281  CAAGTTCGCGGCCACCGTGCGCCTGCTGGTGGGCGGGAAGCTGAACGTGCACATGAACCCGCCCCAGGTGAAGGCCACCA 1360
 347   K  F  A  A  T  V  R  L  L  V  G  G  K  L  N  V  H  M  N  P  P  Q  V  K  A  T  I  373

1361  TCATCAGCGAGCAGCAGGCCAAGTCCCTGCTCAAGAACGAGAACACCCGCAATGAGTGCAGCGGGGAGATCCTGAACAAC 1440
 374   I  S  E  Q  Q  A  K  S  L  L  K  N  E  N  T  R  N  E  C  S  G  E  I  L  N  N   399

1441  TGCTGTGTGATGGAGTACCACCAGCGCACAGGCACCCTCAGCGCCCACTTCAGGAACATGTCCCTCAAGAGGATCAAGCG 1520
 400  C  C  V  M  E  Y  H  Q  R  T  G  T  L  S  A  H  F  R  N  M  S  L  K  R  I  K  R   426

1521  AGCTGACCGGCGAGGCGCAGAGTCTGTGACAGAAGAGAAGTTCACGGTCCTGTTTGAGTCTCAATTCAGTGTTGGCAGCA 1600
 427   A  D  R  R  G  A  E  S  V  T  E  E  K  F  T  V  L  F  E  S  Q  F  S  V  G  S  N  453

1601  ATGAGCTTGTGTTCCAGGTGAAGACCCTGTCCCTTCCCGTGGTTGTCATCGTTCACGGCAGCCAGGACCACAATGCTACC 1680
 454   E  L  V  F  Q  V  K  T  L  S  L  P  V  V  V  I  V  H  G  S  Q  D  H  N  A  T   479

1681  GCCACTGTGCTGTGGGACAATGCCTTTGCTGAGCCGGGCAGGGTGCCGTTTGCGGTGCCTGACAAAGTCCTGTGGCCGCA 1760
 480  A  T  V  L  W  D  N  A  F  A  E  P  G  R  V  P  F  A  V  P  D  K  V  L  W  P  Q   506

1761  GCTGTGCGAGGCGCTCAACATGAAATTCAAGGCCGAGGTGCAGAGCAACCGGGGCCTGACCAAGGAGAACCTGTTGTTTC 1840
 507   L  C  E  A  L  N  M  K  F  K  A  E  V  Q  S  N  R  G  L  T  K  E  N  L  L  F  L  533

1841  TGGCGCAGAAGCTGTTCAACAACAGCAGCAGCCACCTCGAGGACTACAACGGCATGTCTGTGTCCTGGTCCCAGTTCAAC 1920
 534   A  Q  K  L  F  N  N  S  S  S  H  L  E  D  Y  N  G  M  S  V  S  W  S  Q  F  N   559

1921  CGGGAGAACTTGCCCGGCTGGAACTACACCTTCTGGCAGTGGTTTGACGGGGTCATGGAGGTGCTGAAGAAACATCACAA 2000
 560  R  E  N  L  P  G  W  N  Y  T  F  W  Q  W  F  D  G  V  M  E  V  L  K  K  H  H  K   586

2001  GCCCCACTGGAATGACGGGGCCATCCTAGGTTTTGTGAACAAGCAACAGGCCCATGACCTGCTCATCAACAAGCCCGACG 2080
 587   P  H  W  N  D  G  A  I  L  G  F  V  N  K  Q  Q  A  H  D  L  L  I  N  K  P  D  G  613

2081  GTACCTTCTTGTTGCGCTTTAGCGACTCAGAAATTGGGGGCATCACCATTGCCTGGAAGTTTGACTCTCCTGACCGTAAC 2160
 614   T  F  L  L  R  F  S  D  S  E  I  G  G  I  T  I  A  W  K  F  D  S  P  D  R  N   639

2161  CTGTGGAATCTGAAGCCATTCACCACACGGGAGGGCTCCATCCGATCCCTGGCCGACCGGTTGGGGGACCTGAACTATCT 2240
 640  L  W  N  L  K  P  F  T  T  R  E  G  S  I  R  S  L  A  D  R  L  G  D  L  N  Y  L   666

2241  CATCTACGTGTTTCCCGACCGGCCCAAGGATGAGGTCTTCTCCAAGTACTACACTCCTGTGCTTGCCAAAGCAGTGGACG 2320
 667   I  Y  V  F  P  D  R  P  K  D  E  V  F  S  K  Y  Y  T  P  V  L  A  K  A  V  D  G  693

2321  GATACGTGAAGCCGCAGATCAAACAAGTGGTCCCTGAGTTTGTGAGCGCCTCTGCAGACTCTGCTGGAAGCCGCCACCTA 2400
 694   Y  V  K  P  Q  I  K  Q  V  V  P  E  F  V  S  A  S  A  D  S  A  G  S  R  H  L   719

2401  CATGGACCAGGCTCCCTCCCCAGCCGTGTGCCCCCAGCCTCACTATAACATGTACCCACAGAACCCTGACCCGGTGCTCG 2480
 720  H  G  P  G  S  L  P  S  R  V  P  P  A  S  L  *                                   734

2481  ACCAGGATGGAGAATTTGACCTGGACGAGACCATGGATGTGGCCCGGCACGTGGAAGAACTCCTCCGCCGCCCCAATGGA 2560
2561  CAGTCTGGACCCCTCTCTCCCCCGCCCGCTGGTCTCTTTACTCCTGCCAGAGGCTCGCTCTCCTGAATGTTTGTTCGAAC 2640
2641  ACTGCACTCCTCTGTGGGAACAATCCCCAGTGTGCAGGGTCCTATTCATTGTGATTTTGTGTTTGTATCTCTGTGCATAC 2720
2721  TGATGCCTTTGCAGGCAGCCCACGTACACGTGTAGACGTGCACGCGCATTTGTGTACGAGGTGTGCCCACTTCGCCTCCG 2800
2801  CAGTCCTAGGTGTGCG  2816
```

Fig. 3. Nucleotide sequence and deduced amino acid sequence of MGF. The amino acid sequences derived from Lys C-generated peptides and obtained by microsequencing are underlined. The asterisk indicates the stop codon. The amino acid sequence starts at nucleotide position 244 and is indicated below the nucleic acid sequence. EMBL Database accession number X78428.

were also detected in ovary, thymus, spleen, kidney, lung, muscle, the adrenal gland (Figure 5, lanes 1–9) and cytotoxic T cells (not shown). MGF expression is not restricted to the mammary gland and the factor may play a wider role than anticipated originally.

In vitro transcription and translation of MGF cDNA

We have performed *in vitro* transcription and translation of the MGF cDNA using the T7 RNA polymerase and the rabbit reticulocyte lysate. The size of the major *in vitro* translated protein is ~90 kDa (Figure 6, lanes 4 and 5). The *in vitro* translation efficiency depends absolutely on the 5′ non-translated region. Very little *in vitro* translated protein was observed when the 5′ non-translated region comprised 243 nt (Figure 6, lane 3). Deletion up to position −83 from the AUG initiation codon increased the translational efficiency (Figure 6, lane 4). Deletion to position −23 resulted in even more efficient *in vitro* translation (Figure 6,

```
A)
                 **** * ***αA2 *   **      ******** *   ****   *    * *
ISGF3-α/113kd WNDGRIMGFVSRSQERRLLKKTMSGTFLLRFSESS.EGGITCSWVEH.QDDDK      622
ISGF3-α/91kd  WNDGCIMGFISKERERALLKDQQPGTFLLRFSESSREGAITFTWVERSQNGGE     614
MGF/92kd      WNDGAILGFVNKQQAHDLLINKPDGTFLLRFSDSEI.GGITIAWKFDSPDRN.      639

                       * *  * *   * *       *            *   *  ****
                 VLIYSVQPYTKEVLQSLPLTEIIRHYQLLTEENIPENPLRFLYPRIPRDE      672
                 PDFHAVEPYTKKELSAVTFPDIIRNYKVMAAENIPENPLKYLYPNIDKDH      664
                 ..LWNLKPFTTREGSIRSLADRLGDLNYLI..........YVFPDRPKDE      677
                                                  region B
B)
                 *   *    ***    *   *  ***        *      **  *  * * *  *  **  * *
ISGF3-α/113kd FSSYV.. GRGLNSDQLS MLRNKLFGQN CRTE.DPL.. .LSWADFTKR ESPPGKLPFW TWLD 558
ISGF3-α/91kd  FSS.VT. KRGLNVDQLN MLGEKLLGPN A.SP.DGL.. .IPWTRFCKE NINDKNFPFW LWIE 548
MGF/92kd      FKAEVQS NRGLTKENLL FLAQKLFNNS SSHLEDYNGM SVSWSQFNRE NLPGWNYTFW QWFD 575
                 region C
```

Fig. 4. Sequence alignment between ISGF3-a proteins and MGF. (**A**) Sequence alignment of the SH2 domains of ISGF3-a proteins and MGF. Identical amino acids are indicated by an asterisk. A putative tyrosine phosphorylation site is underlined. (**B**) Sequence alignment of SH3 domains of ISGF3-a proteins and MGF. The symbols are the same as in (A). The regions B and C (Fu, 1992; Fu *et al.*, 1992; Schindler *et al.*, 1992b) are indicated under the sequence. The αA2 position is shown on the sequence (Waksman *et al.*, 1993).

lane 5). The *in vitro* translated MGF was tested in a bandshift assay and shown to have no specific DNA binding activity (data not shown). This is most probably due to the lack of protein phosphorylation (see below).

Activation of recombinant MGF by the long form of the prolactin receptor

We have transfected an MGF expression vector into COS cells to confirm that the cloned cDNA encodes a sequence-specific DNA binding protein. Forty-eight hours after transfection, cells were harvested, nuclear extracts were prepared and bandshift assays were performed (Figure 7). Endogenous MGF activity was not detected in COS cells (lanes 5 and 6). Low MGF activity was observed when COS cells were transfected with MGF cDNA (lane 1). The treatment of the cells with prolactin did not significantly enhance MGF binding (lane 2).

Specific DNA binding of endogenous MGF in cultured mammary epithelial cells is directly regulated by prolactin (Standke *et al.*, 1994). We investigated the prolactin responsiveness of the recombinant MGF expressed in COS cells. COS cells do not express prolactin receptor endogenously. For this reason we transfected MGF cDNA and two different forms of the prolactin receptor, the short form and the long form. Twenty-four hours after transfection the cells were induced with prolactin. Nuclear extracts were prepared 24 h later and analysed in bandshift assays. COS cells cotransfected with MGF cDNA and the long form of the prolactin receptor show a strong MGF DNA binding activity upon prolactin induction (Figure 7, lane 4). Low MGF activity was found in the absence of prolactin (lane 3). The cotransfection of the short form of the prolactin receptor failed to enhance MGF binding, independent of prolactin stimulation (lanes 7 and 8). The transfection of long and short form prolactin receptor alone resulted in no MGF binding (Figure 7, lane 5, 6, 9 and 10). These observations are consistent with transfection experiments, which have shown that only the long form of the prolactin receptor confers milk protein induction and induces β-lactoglobulin gene transcription (Lesueur *et al.*, 1991).

The specificity of the DNA binding was confirmed by oligonucleotide competition experiments. In the presence of a 100-fold molar excess of the wild type oligonucleotide, representing the MGF binding sequence, the MGF–DNA complex was efficiently competed. In the presence of a mutated oligonucleotide, not capable of MGF binding, no competition was observed (data not shown). These data further confirm that the cloned cDNA encodes MGF and that the DNA binding of the recombinant protein can be activated directly by the long form of the prolactin receptor.

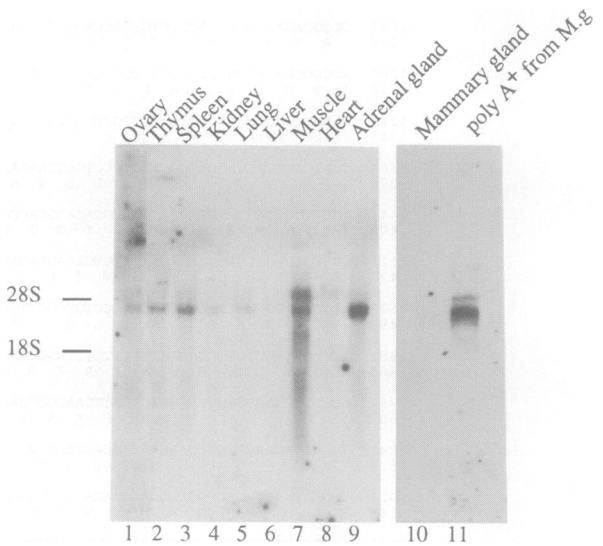

Fig. 5. Northern blot analysis of MGF mRNA. Total RNA from different tissues of sheep was separated by gel electrophoresis and transferred onto a nylon membrane. 15 mg of total RNA and 2 μg of poly(A)$^+$ RNA from lactating mammary gland were used. The RNA blot was probed with a *Bst*BI and *Bgl*I (223–804, corresponding to the N-terminal coding region) fragment from clone 10. After washing, RNA blots from mammary gland were exposed for 16 h; RNA blots from the other organs were exposed for 14 days at −70°C. The position of 28S and 18S rRNA is indicated.

Tyrosine phosphorylation is required for the DNA binding activity of MGF

We have investigated the effect of phosphorylation on MGF DNA binding activity using purified MGF from sheep mammary gland. MGF purified by DNA affinity chromatography was treated with the serine/threonine-specific phosphatases PP1C and PP2A. No effect on the DNA binding activity of MGF was observed (data not

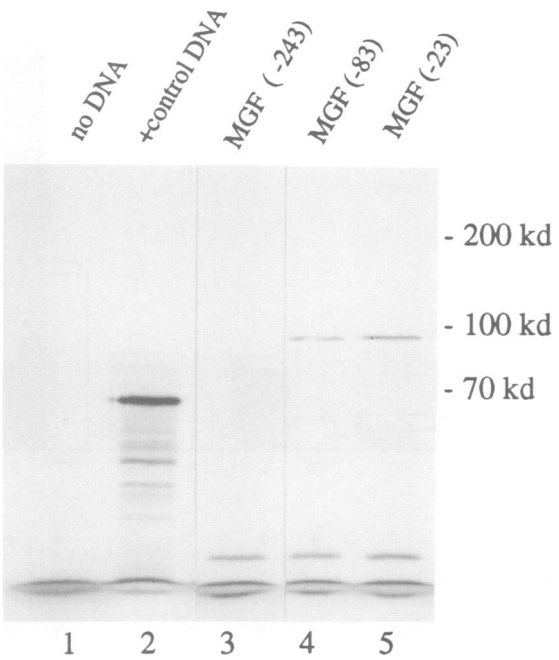

Fig. 6. *In vitro* transcription and translation of MGF cDNA. Three constructs with different 5′ non-translated regions were tested. The numbering refers to the ATG initiator codon (lanes 3–5). DNA encoding the luciferase gene was used as a control. This DNA yields a 67 kDa product (lane 2). 2 μl from a 25 μl reaction were used for SDS–PAGE. The gel was dried and exposed for 16 h at room temperature. Molecular weight markers are indicated.

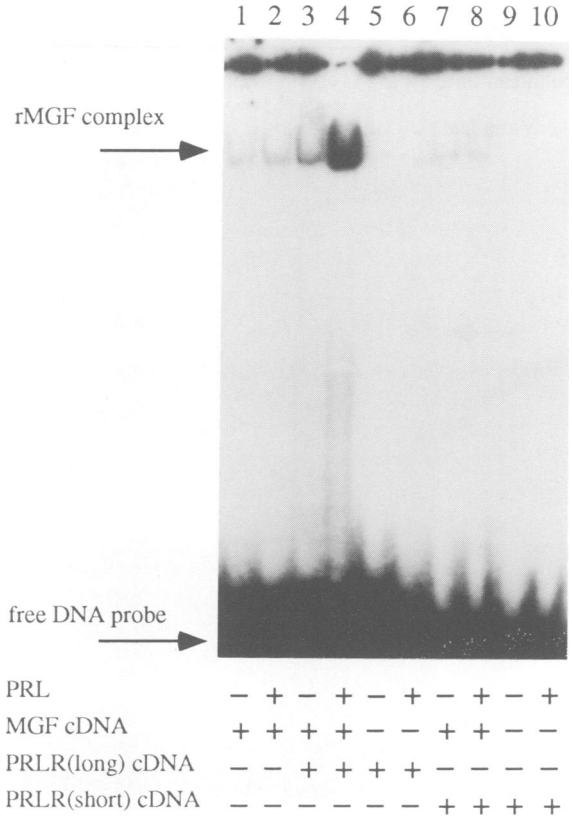

Fig. 7. Prolactin induces DNA binding of MGF via the long form of the prolactin receptor. COS 7 cells were transfected with MGF cDNA (lanes 1 and 2), MGF cDNA and prolactin receptor cDNA (long form) (lanes 3 and 4), MGF cDNA and prolactin receptor (short form) (lanes 7 and 8) and prolactin receptor cDNA alone long form (lanes 5 and 6) and short form (lanes 9 and 10) and induced with prolactin (lanes 2, 4, 6, 8 and 10) or not (lanes 1, 3, 5, 7 and 9).

shown). Treatment of MGF with the tyrosine-specific phosphatase PTP1B, however, abolished DNA binding (Figure 8, lanes 3, 5, 7 and 9). Inclusion of vanadate, a potent protein tyrosine phosphatase inhibitor, in the dephosphorylation reaction prevented the loss of DNA binding activity (Figure 8, lanes 4, 6, 8 and 10). Recombinant MGF activated by prolactin from transfected COS cells was also tested. Tyrosine-specific phosphatase treatment of nuclear extracts prevented the formation of the MGF–DNA complex (data not shown). These data show that MGF phosphorylation on tyrosine is essential for its DNA binding activity.

Prolactin-dependent transactivation of a reporter gene construct in CHO cells transfected with MGF cDNA

The activation of the prolactin receptor induces the DNA binding activity of recombinant MGF in transfected cells (Figure 7). We have tested the ability of recombinant MGF to transactivate a reporter gene construct. For this purpose MGF cDNA was cotransfected with a β-casein promoter-luciferase gene construct into CHO cells (Figure 9). A very slight increase in luciferase activity was detected when the β-casein promoter construct was cotransfected with the MGF cDNA (lanes 5 and 6) or the long form of the prolactin receptor (lanes 3 and 4). Cotransfection of the promoter construct, the prolactin receptor and the MGF cDNA resulted in a substantial increase in luciferase activity which could be enhanced further by prolactin treatment of the transfected cells (lanes 7 and 8). These data demonstrate that recombinant MGF has the potential to activate transcription in transfected cells.

Discussion

Purification of MGF from mammary tissue of lactating sheep

MGF from rat mammary glands has been purified previously to homogeneity by anion exchange chromatography on Bio-Rex 70, non-specific DNA and sequence-specific DNA affinity chromatography (Wakao *et al.*, 1992). MGF from sheep mammary gland does not bind to the Bio-Rex 70 resin under the same conditions used for the purification of rat MGF. This indicates different charge properties of MGF from rat and sheep. NaCl was not sufficient to elute sheep MGF from the Red A resin — KSCN was used instead. KSCN shows a much stronger negative elution coefficient (Robinson *et al.*, 1981), suggesting that sheep MGF was bound tightly to the Red A resin through ionic and hydrophobic interactions.

After the sequence-specific DNA affinity column, two bands of 92 and 84 kDa were detected by SDS gel electrophoresis. Only one band of 89 kDa was found after purification from rat tissue. This band was sufficient to form a specific protein–DNA complex in a bandshift experiment (Wakao *et al.*, 1992). The 84 kDa protein is closely related to the 92 kDa protein and might originate from alternative splicing or proteolytic clipping. Both proteins have very similar DNA binding specificities and bind to the wild type MGF binding sequence and not to a mutated version. The

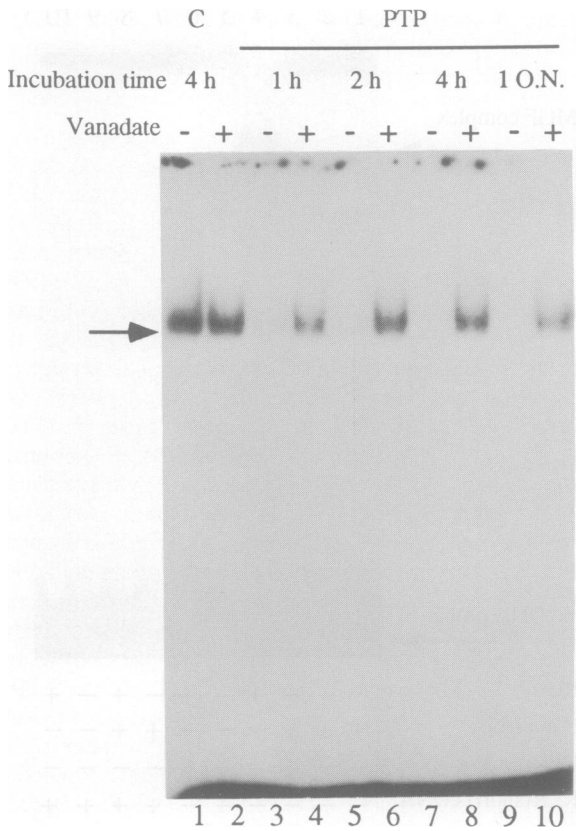

Fig. 8. Treatment of purified MGF with tyrosine-specific protein phosphatase causes loss of DNA binding. 1 μl of purified MGF were treated with 0.1 U of protein tyrosine phosphatase (PTP) for the times indicated. The MGF–DNA complex is indicated by an arrow. No PTP was added in lanes 1 and 2. Incubations were carried out with (+) or without (−) vanadate (lane 3–10). The concentration of vanadate was 100 μM.

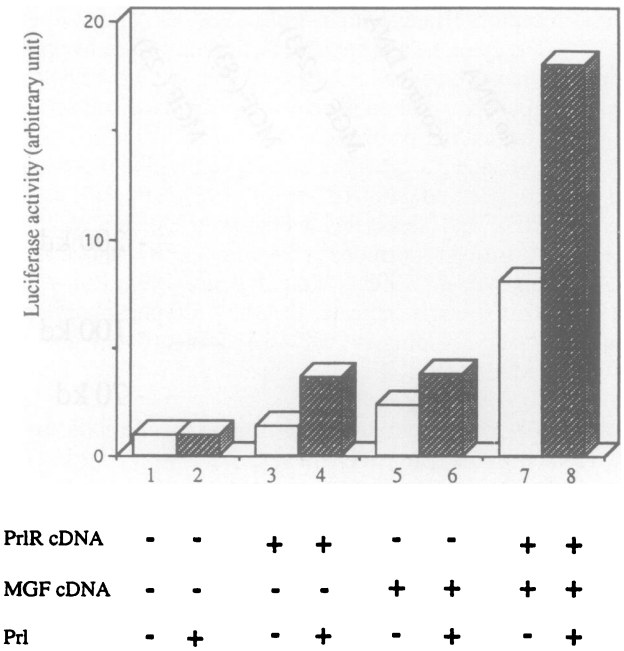

Fig. 9. MGF transactivates the β-casein gene promoter in transfected CHO cells. CHO cells were transfected with the prolactin receptor (long form) cDNA, MGF cDNA and MGF cDNA and the prolactin receptor (long form). Cells were treated with prolactin for 24 h before harvesting (columns 2, 4, 6 and 8). All transfections include a β-casein promoter–luciferase construct. Luciferase and β-galactosidase activities were measured. The luciferase activity divided by the β-galactosidase activity is shown. No cDNA was transfected (columns 1 and 2), prolactin receptor cDNA was transfected (columns 3 and 4), MGF cDNA was transfected (columns 5 and 6) and prolactin receptor and MGF cDNA was transfected (columns 7 and 8).

relative abundance of the two proteins varied in different preparations. The 92 kDa protein was usually more abundant than the 84 kDa protein. The peptide patterns after *in situ* lysine C digestion were similar. Both proteins were observed after *in vitro* transcription and translation of MGF cDNA.

The primary structure of MGF

The cDNA sequence indicates that MGF belongs to a new family of transcription factors involved in the cytokine responses, the Stats (signal transducers and activators of transcription). The presence of SH2 and SH3 domains and the homologies to the Stat91 and the Stat113 of ISGF3 are noteworthy. SH2 domains play a crucial role in protein–protein interactions and mediate the binding to phosphorylated tyrosine residues (Anderson *et al.*, 1990; Hirai and Varmus, 1990; Mayer and Hanafusa, 1990; Matsuda *et al.*, 1991; Mayer *et al.*, 1992). The presence of an SH2 domain within MGF suggests that MGF could interact directly with PTKs. Within the SH2 domain of MGF there is a basic amino acid, lysine, at position αA2, as in other src family proteins (Figure 4A). A basic amino acid at this position is highly conserved amongst different types of SH2, and is directly involved in the interaction with phosphotyrosine (Waksman *et al.*, 1992, 1993). MGF and ISGF3 proteins do not contain a region A. The region B present in ISGF3 proteins is absent in the SH2 domain of MGF (Figure 4A). Both regions are distinctive sequences for SH2-containing proteins (Fu *et al.*, 1992; Schindler *et al.*, 1992b). It has been suggested that these sequences might be targets for PTKs.

The SH3 domain of MGF also shares extensive homologies with the SH2 domains of the ISGF3 (Figure 4B). MGF does not contain the YDYEE sequence which is conserved in many other SH3-containing proteins. Although the role of the SH3 domain is not unequivocally established, several functions have been proposed such as cytoskeleton assembly and membrane rearrangement (Drubin *et al.*, 1990; Rodaway *et al.*, 1990). The DNA binding domain of MGF has yet to be identified. The DNA sequence to which MGF binds contains an imperfect palindrome, 5′-TTCNNN-GAA-3′, which might indicate that MGF binds to DNA as a homodimer. The DNA motif 5′-TTC/ACNNNAA-3′ is also present in the GAS element, the sequence to which Stat91 binds (Pearse *et al.*, 1993), and in the APRF response element (Wegenka *et al.*, 1993).

The activation of MGF by prolactin

Only the long form of prolactin receptor activates the 92 kDa MGF. The short form of the prolactin receptor is unable to do so (Figure 7). The intermediate form of the prolactin receptor, missing 198 amino acids at its C-terminus, does not affect its signal transducing capacity (Russell, 1989; Ali *et al.*, 1991). One hundred and sixty amino acid residues of cytoplasmic domain are sufficient to transduce a signal. Similar deletion mutants have been reported for the IL-2 and

EPO receptors (Hatakeyama *et al.*, 1989; D'Andrea *et al.*, 1991). The cytokine and hematopoietic cell surface receptors do not have enzymatic kinase domains. Nevertheless, tyrosine phosphorylation is involved in cytokine cell surface receptor signalling pathways. Prolactin receptor activation induces tyrosine phosphorylation of several cellular proteins, e.g. p120, p97 and p40 (Rui *et al.*, 1992). The PTK JAK2 is 130 kDa and associates with erythropoietin receptor (EPOR), growth hormone receptor (GHR) and PRLR (Argetsinger *et al.*, 1993; Witthuhn *et al.*, 1993; Rui *et al.*, 1994). This strongly indicates that the p120 phosphorylated upon prolactin binding is JAK2. JAK2 is also activated by IL-3, GM-CSF and IFN-γ (Silvennoinnen *et al.*, 1993a; Witthuhn *et al.*, 1993). The similarity in molecular weight of the 97 kDa protein and MGF might also be indicative.

In IFN-γ signalling, JAK1 and JAK2 are responsible for the tyrosine phosphorylation of Stat91. A tyrosine residue at position 701 has been shown to be important for the activation of Stat91 protein and for the translocation of this factor into the nucleus (Shuai *et al.*, 1993a,b; Silvennoinnen *et al.*, 1993a,b). MGF has a putative tyrosine phosphorylation site at position 694. The sequence around position 694 (VDGYVKP) is similar to that of Stat91 around position 701 (GTGYIKT). Mutation of this tyrosine will reveal whether this phosphorylation affects MGF DNA binding activity or translocation into the nucleus, similar to that reported for ISGF3 proteins upon IFN receptor activation (Shuai *et al.*, 1993b).

It might become interesting to analyse the phosphotyrosine residues in JAK2 and their interaction with SH2-containing molecules. It is possible that MGF interacts directly with JAK2 via its SH2 domain. An anti-phosphotyrosine mAb detected three proteins, including p97, upon prolactin receptor activation. When antibodies against the prolactin receptor were used, p97 was not detected (Rui *et al.*, 1992). This might support our hypothesis that MGF interacts with JAK2.

The electrophoretic migration of the DNA–protein complex formed by recombinant MGF from transfected COS cells is slightly different from that of the purified protein (data not shown). This difference may be due to post-translational modifications of MGF. MGF contains several putative PKC and CKII sites in addition to the potential phosphotyrosine site. The suppression of MGF activation and β-casein promoter induction by a selective inhibitor of PKC indicates additional regulatory phosphorylation events (Marte *et al.*, 1994). Other post-translational modifications, glycosylation and myristylation, and their effects on MGF function, will also have to be considered.

MGF activates a reporter gene in a prolactin-dependent manner. The promoter of the reporter construct is a truncated rat β-casein gene comprising positions -334 to -1 (Schmitt-Ney *et al.*, 1991). The cotransfection of MGF cDNA and the prolactin receptor into CHO cells and prolactin induction resulted in an 18-fold increase in the transcription of the reporter construct. MGF cDNA transfection or prolactin receptor transfection by themselves slightly increased the reporter gene promoter activity. Our bandshift analysis showed that slightly increased levels of MGF–DNA complex could be observed when MGF cDNA was transfected without the prolactin receptor construct (Figure 7, lanes 1 and 2). This might be attributed to the presence of small amounts of prolactin in the serum (Lesueur *et al.*, 1991). Our results demonstrate that MGF can be activated directly by prolactin and can transactivate a target gene. The MGF binding sequence is a prolactin response element.

The sequence to which MGF binds is highly conserved in milk protein gene promoters (Yoshimura and Oka, 1989). Similar sequences are also found in IL-4, IL-6, IFN-α and IFN-γ response elements (Hocke *et al.*, 1992; Kotanides and Reich, 1993; Wegenka *et al.*, 1993; Pine *et al.*, 1994; Standke *et al.*, 1994). IFN-γ activation factor (GAF) and acute-phase response factor (APRF) bind to GAS and IL-6RE. Both factors are recognised by an antibody against the Stat91 component of ISGF3 (Kanno *et al.*, 1993). The GAS consensus sites to which Stat91 binds is TTC/ACNNNAA (Pearse *et al.*, 1993). The MGF consensus binding sequence is TTCNNNGAA. The G in this sequence element is essential for MGF binding (Schmitt-Ney *et al.*, 1992), but not for Stat91. It is possible that MGF recognizes a subset of GAS core sites. MGF also recognizes the APRF binding sequence and shares DNA binding and transactivation potential with this IL-6-induced factor (Standke *et al.*, 1994).The molecular weights of APRF and MGF are similar (90–92 kDa). We suggest that Stat91, Stat113, MGF and possibly APRF are products from related, but distinct, genes and that MGF is a novel member of this Stat gene family. Additional receptors might be coupled to members of this family. An 84 kDa protein was detected in immunoprecipitates of activated GHR (Argetsinger *et al.*, 1993). EPO and IL-4 have been shown to induce phosphorylation of 91 and 92 kDa proteins, respectively (Izuhara and Harada, 1993; Witthuhn *et al.*, 1993). These observations suggest that MGF and/or related molecules transduce the effects of various cytokines such as IL-6, EPO, IFN-γ and prolactin. MGF may play a wider physiological role than anticipated originally and may be renamed as a member of the Stat family.

Materials and methods

Purification of MGF

Mammary gland tissue was obtained from sheep lactating for 4 days. The tissue was frozen in liquid nitrogen and stored at $-70°C$. All purification steps were performed at 4°C. Nuclei were prepared as described previously (Wakao *et al.*, 1992) and extracted with buffer D containing 400 mM NaCl. This fraction contained most of the MGF activity and was used for the purification. NP-40 was added to a final concentration of 0.1% (v/v) and the fraction was centrifuged at 18 000 r.p.m. in an SW21 rotor for 16 h. After centrifugation, the fraction was dialysed against buffer D [20 mM HEPES–NaOH pH 7.5, 50 mM NaCl, 2 mM EDTA, 1 mM DTT, 10% (v/v) glycerol, 0.1% (v/v) NP-40 and 0.2 mM PMSF]. After dialysis the proteins were applied to a Matrex gel Red A (Amicon) column (radius 15 × 200 mm), pre-equilibrated with buffer D. After loading, the column was washed with three bed volumes of buffer D containing 300 mM NaCl and subsequently with the same volume of buffer D containing 50 mM KSCN. The column was developed with 800 ml of a linear gradient of 50–500 mM KSCN. 4 ml fractions were collected and 2 μl from each fraction were assayed for MGF DNA binding activity. Positive fractions (180 ml) were collected and dialysed against 2 l buffer D. Dialysed fractions were concentrated with Red A column (1/5 volume of the first column). MGF was eluted with 500 mM KSCN and dialysed against buffer E [10 mM Tris–HCl pH 7.5, 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 5% (v/v) glycerol, 0.1% NP-40 and 0.2 mM PMSF]. The eluate was concentrated to 20 ml after dialysis and aliquots of 1 ml were stored at $-70°C$. A sequence-specific DNA affinity column was prepared as described (Wakao *et al.*, 1992), except that the bovine β-casein MGF binding sequence

was used. For the sequence-specific DNA affinity column, 1 ml of material concentrated on Red A was mixed with the poly[d(I,C)] and poly[d(A,T)] (50 μg of each per mg protein), incubated on ice for 40 min and loaded onto 1 ml of the affinity column pre-equilibrated with buffer E. The sample was mixed with the DNA affinity resin and incubated for 45 min. The flowthrough was reloaded three times and the column washed with 20 vols buffer E containing 180 mM NaCl. MGF was eluted with 5 ml buffer E containing 600 mM NaCl. 1 ml fractions were collected and 1 μl of each fraction was assayed for MGF activity.

Internal amino acid sequencing

The eluate of the DNA affinity column was precipitated with TCA (trichloroacetic acid) at a final concentration of 17% (v/v) on ice for 2 h. The precipitate was collected by centrifugation and washed twice with ethanol/diethylether (1/1). The pellet was dried and the proteins were dissolved and separated by SDS gel electrophoresis. Then they were transferred onto a nitrocellulose membrane. The filter wash with NaOH was omitted. The filter was stained with Ponceau S and the two bands of 92 and 84 kDa were cut out. Protein sequencing was carried out at the Harvard Microchemistry facility. The proteins were digested *in situ* with lysine C and the resulting peptides were separated by reverse phase chromatography. The amino acid sequence of individual peptides was determined with an ABI 477 A protein sequencer connected to a 120 A on-line PTH amino acid analyser.

Construction of a cDNA library from the mRNA of lactating mammary tissue

A cDNA library was constructed from the RNA of mammary gland of sheep lactating for 4 days. RNA was prepared and poly(A)$^+$ RNA was selected. cDNA was synthesized with the Super Script lambda system cDNA synthesis kit (BRL), except *Eco*RI adapters were used instead of *Sal*I adapters and λgt 11 phage instead of λgt 22A phage for the construction of the phage library.

Isolation of cDNA clones

6×10^5 phage clones were screened with a guessmer derived from the 14 amino acid sequence of a peptide. The sequence of this oligonucleotide was 5′-TTCCACTGGATIAGCTCITCITCIAGGATGATGGTCTGCTGT-3′ (I indicates inosine). The synthetic oligonucleotide was end-labelled with [γ-^{32}P]ATP and T4 polynucleotide kinase and used to screen clones on Hybond N$^+$ (Amersham). Prehybridization was performed in 25 ml of solution containing 5× SSPE, 6% (w/v) polyethyleneglycol (mol. wt 6000), 5× Denhard's solution, 0.5% SDS (w/v) and 500 μg salmon sperm DNA at 37°C under gentle agitation. Hybridization was carried out for 16 h at 45°C in the same solution containing the radioactive probe. Filters were washed twice at room temperature in 2× SSPE and 0.1% (w/v) SDS for 10 min, then washed at 37°C in 1× SSPE and 0.1% (w/v) SDS for 15 min. The washing temperature was increased to 48°C in 1 × SSPE and 0.1% (w/v) SDS. After the wash at 48°C for 15 min, filters were autoradiographed and kept at −70°C for 16 h. After a second screening using the same oligonucleotide, 20 positive clones were isolated. These clones were hybridized with another oligonucleotide (5′-ATCTGIGGYTTCACI-CCICCGTCCACIGCYTT-3′; Y indicates pyrimidines and I indicates inosine) derived from the second peptide sequence obtained by micro-sequencing. All 20 clones hybridized with both oligonucleotides and were isolated as single plaques. These clones were subjected to restriction enzyme analysis and the DNA sequences of their 5′ and 3′ ends were determined. The longest cDNA (clone 10) was inserted in Blue Script KS$^+$ using *Eco*RI and *Not*I sites.

DNA sequencing

The cDNA clones were sequenced using Sequenase™ kits from US Biochemicals. Both strands of the clone 10 cDNA were sequenced. Synthetic oligonucleotides were used as primers. 5′ and 3′ deletion mutants were sequenced with T7 and M13 universal primers. These deletion mutants were generated by *Exo*III limited digestion using Erase Base kits from Promega. 5′ deletion mutants were used for the *in vitro* transcription and translation experiments and for the construction of the eukaryotic expression vector.

Northern blot analysis

Total RNA or poly(A)$^+$ RNA was prepared from different organs of lactating sheep and separated on 1% agarose gels. RNA was transferred on the nylon filter and hybridized with the radiolabelled clone 10 fragment. Washing was performed with 0.1× SSC and 0.1% SDS at 60°C.

In vitro transcription and translation

In vitro transcription was performed with T_NT T7 polymerase (Promega) and translation with a rabbit reticulocyte lysate according to the manufacturer's protocol. 1 μg cDNA in Blue Script KS$^+$ was used. The *in vitro* translated proteins were separated by SDS gel electrophoresis and autoradiographed.

Transient expression of cDNA clones in COS and CHO cells

The clone 10 cDNA (*Bst*BI cleaved, comprising nucleotides 222−2634) was inserted into the pXM eukaryotic expression vector. COS cells were grown in RPMI medium with 10% FCS. COS cells were transfected with 5 μg pXM-MGF, 0.5 μg of the SV-40-β-galactosidase expression vector and 5 μg of the pcDNAI vector containing the long or short form of the murine prolactin receptor (R.Ball, Basel, Switzerland, unpublished data). The total amount of DNA was adjusted with salmon sperm DNA to 20 μg. Transfection was performed by the calcium phosphate precipitation technique. 24 h after transfection, COS cells were induced with 5 μM ovine prolactin. The cells were harvested 48 h after transfection and nuclear extracts were prepared. Gel retardation assays were performed (Wakao *et al.*, 1992). CHO cells were grown in Optimem 1 medium (BRL) in the presence of 2% FCS. Transfection into CHO cells was performed in the same way as COS cells. 1.5 μg of the reporter gene (β-casein promoter-luciferase) was included.

Protein tyrosine phosphatase (PTP) treatment of purified MGF

DNA affinity-purified MGF was incubated with beads coupled to GST-fused PTP 1B (UBI) in buffer E for the indicated times and was used for bandshift analysis.

Acknowledgements

We thank Mr Willem and Dr Strittmatter (Ciba-Geigy at St Aubin) for providing mammary gland tissue of lactating sheep; Drs B.Lane and W.Lane (Harvard University Microchemistry facility) for the peptide sequencing; Dr K.Takabayashi (Gencia Ltd, San Diego, CA) for his interest and encouragement; Dr B.Happ, Mr P.Hofer and Mrs F.Maurer for their help with the experiments; Drs P.Matthias, N.E.Hynes, Y.Nagamine and K.Ballmer for helpful discussions; Dr R.K.Ball for providing mouse prolactin receptors; Dr B.Favre for providing phosphatases; Dr H.Rui for communicating results; and Mr G.Achenbach and Mr P.Müller for the synthesis of oligonucleotides.

References

Ali,S., Pellegrini,I. and Kelly,P.A. (1991) *J. Biol. Chem.*, **266**, 20110−20117.

Anderson,D., Koch,C.A., Grey,L., Ellis,C., Moran,M. and Pawson,T. (1990) *Science*, **250**, 979−982.

Argetsinger,L.S., Campell,G.S., Yang,X., Witthuhn,B.A., Silvennoinen,O., Ihle,J.N. and Carter-Su,C. (1993) *Cell*, **74**, 237−244.

Bazan,J.F. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 6934−6938.

Boutin,J.M. *et al.* (1988) *Cell*, **53**, 69−77.

Boutin,J.M. *et al.* (1989) *Mol. Endocrinol.*, **31**, 1455−1461.

Cosman,D., Lyman,S.D., Idzerda,R.L., Beckman,M.P., Park,L.S., Goodwin,R.G. and March,C.J. (1990) *Trends Biochem. Sci.*, **15**, 265−270.

D'Andrea,A.D., Yoshimura,A., Youssoufian,H., Zon,L.I., Koo,J.-W. and Lodish,H.F. (1991) *Mol. Cell. Biol.*, **11**, 1980−1987.

Davis,J.A. and Linzer,D.I.H. (1989) *Mol. Endocrinol.*, **3**, 674−680.

Davis,S., Aldrich,T.H., Valenzuela,D.M., Wong,D.M., Furth,M.E., Squinto,S.P. and Yancopoulos,G.D. (1991) *Science*, **253**, 59−63.

Doppler,W., Groner,B. and Ball,R.K. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 104−108.

Drubin,D.G., Mulholland,J., Zhu,Z. and Botstein,D. (1990) *Nature*, **343**, 288−290.

Edery,M. *et al.* (1989) *Proc. Natl Acad. Sci. USA*, **86**, 2112−2116.

Fu,X.Y. (1992) *Cell*, **70**, 232−335.

Fu,X.Y., Schindler,C., Improta,T., Aebersold,R. and Darnell,J.E.,Jr (1992) *Proc. Natl Acad. Sci. USA*, **89**, 7840−7843.

Fukunaga,R., Ishizuka-Ikeda,E., Seto,Y. and Nagata,S. (1990) *Cell*, **61**, 341−350.

Gagnerault,M.-C., Touraine,P., Savino,W., Kelly,P.A. and Dardenne,M. (1993) *J. Immunol.*, **150**, 5673−5681.

Gearing,D.P., King,J.A., Gough,N.M. and Nocola,N.A. (1989) *EMBO J.*, **8**, 3667−3676.

Gearing,D.P. *et al.* (1991) *EMBO J.*, **10**, 2839–2848.
Goodwin,R.G. *et al.* (1990) *Cell*, **60**, 941–951.
Gorodetsky,S.I., Tkach,T.M. and Kapelinskaya,T.V. (1988) *Gene*, **66**, 87–96.
Hatakeyama,M., Tsudo,M., Minamoto,S., Kono,T., Doi,T., Miyata,T., Miyasaka,M. and Taniguchi,T. (1989) *Science*, **244**, 551–556.
Hirai,H. and Varmus,H.E. (1990) *Mol. Cell. Biol.*, **10**, 1317.
Hocke,G.M., Barry,D. and Fey,G.H. (1992) *Mol. Cell. Biol.*, **12**, 2282–2294.
Hooghe,R., Delhase,M., Vergani,P., Malur,A. and Hooghe-Peters,E.L. (1993) *Immunol. Today*, **14**, 212–214.
Itoh,N. *et al.* (1990) *Science*, **247**, 324–237.
Izuhara,K. and Harada,N. (1993) *J. Biol. Chem.*, **268**, 13097–13102.
Kanno,Y., Kozac,C.A., Schindler,C., Driggers,P.H., Ennist,D.L., Gleason,S.L., Darnell,J.E.,Jr and Ozato,K. (1993) *Mol. Cell. Biol.*, **13**, 3951–3963.
Kelly,P.A., Djiane,J., Postel-Vinay,M.C. and Edery,M. (1991) *Endocrine Rev.*, **12**, 235–251.
Koch,C.A., Moran,M., Sadowski,I. and Pawson,T. (1989) *Mol. Cell. Biol.*, **9**, 4131–4140.
Kotanides,H. and Reich,N.C. (1993) *Science*, **262**, 1265–1267.
Lesueur,L., Edery,M., Ali,S., Paly,J., Kelly,P.A. and Djiane,J. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 824–828.
Leung,D.W. *et al.* (1987) *Nature*, **330**, 537–543.
Marte,B.M., Meyer,T., Stabel,S., Standke,G.J.R., Jaken,S., Fabbro,D. and Hynes,N.E. (1994) *Cell Growth Diff.*, in press.
Matsuda,M., Mayer,B.J., Fukui,Y. and Hanafusa,H. (1991) *Science*, **248**, 1537–1539.
Mayer,B.J. and Hanafusa,H. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 2638–2642.
Mayer,B.J., Jackson,P.K., van Etten,R.A. and Baltimore,D. (1992) *Mol. Cell. Biol.*, **12**, 609–618.
Miyazaki,T., Maruyama,M., Yamada,G., Hatakeyama,M. and Taniguchi,T. (1991) *EMBO J.*, **10**, 3191–3197.
Mosley,B. *et al.* (1989) *Cell*, **59**, 335–348.
Murphy,W., Durum,S.K. and Longo,D.L. (1993) *J. Exp. Med.*, **178**, 231–236.
Pearse,R.N., Feiman,R., Shuai,K., Darnell,J.E.,Jr and Ravetch,J.V. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 4314–4318.
Pine,R., Canova,A. and Schindler,C. (1994) *EMBO J.*, **13**, 158–167.
Robinson,J.B.,Jr, Strottmann,J.M. and Stellwagen,E. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 2287–2291.
Rodaway,A.R.F., Teahan,C.G., Casimir,C.M., Segal,A.W. and Bentley,D.L. (1990) *Mol. Cell. Biol.*, **10**, 5388–5396.
Rui,H., Djeu,J.Y., Evans,G.A., Kelly,P.A. and Farrar,W.L. (1992) *J. Biol. Chem.*, **267**, 24076–24081.
Rui,H., Kirken,R.A. and Farrar,W.L. (1994) *J. Biol. Chem.*, **269**, 1–5.
Russell,D.H. (1989) *Trends Pharmacol. Sci.*, **10**, 40–44.
Schindler,C., Fu,X.Y., Improta,T., Aebersold,R. and Darnell,J.E.,Jr (1992) *Proc. Natl Acad. Sci. USA*, **89**, 7836–7839.
Schlessinger,J. and Ullrich,A. (1993) *Neuron*, **9**, 383–391.
Schmitt-Ney,M., Doppler,W., Ball,R.K. and Groner,B. (1991) *Mol. Cell. Biol.*, **11**, 3745–3755.
Schmitt-Ney,M., Happ,B., Ball,R.K. and Groner,B. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 3130–3134.
Shirota,M., Banville,D., Ali,S., Jolicoeur,C., Boutin,J.-M., Edery,M., Djiane,J. and Kelly,P.A. (1990) *Mol. Endocrinol.*, **4**, 1136–1143.
Shuai,K., Schindler,C., Prezioso,V.R. and Darnell,J.E. (1992) *Science*, **258**, 1808–1812.
Shuai,K., Ziemiecki,A., Wilks,A., Harpur,A.G., Sadowski,H.B., Gilman,M.Z. and Darnell,J.E. (1993a) *Nature*, **366**, 580–583.
Shuai,K., Stark,G.R., Kerr,I.M. and Darnell,J.E.,Jr (1993b) *Science*, **261**, 1744–1746.
Silvennoinnen,O., Witthuhn,B., Quelle,F.W., Cleveland,J.L., Yi,T. and Ihle,J.N. (1993a) *Proc. Natl Acad. Sci. USA*, **90**, 8429–8433.
Silvennoinnen,O., Ihle,J.N., Schlessinger,J. and Levy,D.E. (1993b) *Nature*, **366**, 583–585.
Stahl,M.L., Ferenz,C.R., Kelleher,K.L., Kritz,R.W. and Knopf,J.L. (1988) *Nature*, **332**, 269–272.
Standke,G., Meier,V. and Groner,B. (1994) *Mol. Endocrinol.*, **8**, in press.
Takaki,S., Tominaga,A., Hitoshi,Y., Mita,S., Sonoda,E., Yamaguchi,N. and Takatsu,K. (1990) *EMBO J.*, **9**, 4367–4374.
Vonderhaar,B. and Ziska,S.E. (1989) *Annu. Rev. Physiol.*, **51**, 641–652.
Yamasaki,K. *et al.* (1988) *Science*, **241**, 825–828.
Yoshimura,M. and Oka,T. (1989) *Gene*, **78**, 267–275.
Wakao,H., Schmitt-Ney,M. and Groner,B. (1992) *J. Biol. Chem.*, **267**, 16365–16370.
Waksman,G. *et al.* (1992) *Nature*, **358**, 646–653.
Waksman,G., Shoelson,S.E., Pant,N., Cowburn,D. and Kuriyan,J. (1993) *Cell*, **72**, 779–790.
Wegenka,U.M., Buschmann,J., Lutticken,C., Heinrich,P.C. and Horn,F. (1993) *Mol. Cell. Biol.*, **13**, 276–288.
Witthuhn,B.A., Quelle,F.W., Silvennoinen,O., Yi,T., Tang,B., Miura,O. and Ihle,J.N. (1993) *Cell*, **74**, 227–236.
Zhang,R., Buczko,E., Tsai-Moris,C.H., Hu,Z.Z. and Dufau,M.L. (1990) *Biochem. Biophys. Res. Commun.*, **168**, 415–422.

Received on January 24, 1994; revised on February 22, 1994