Journal of Cellular Biochemistry 80:606–616 (2001)

# Molecular Analysis of the Mouse S100A9 Gene and Evidence that the Myeloid Specific Transcription Factor C/EBPepsilon is not Required for the Regulation of the S100A9/A8 Gene Expression in Neutrophils

**Wolfgang Nacken,[1]* Julie A. Lekstrom-Himes,[2] Clemens Sorg,[2] and Marie Pierre Manitz[2]**

[1]Institute of Experimental Dermatology, University of Münster, D-48149 Münster, Germany
[2]Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892

**Abstract** The genomic locus of the mouse S100A9 (MRP14) gene, a myeloid expressed gene belonging to the S100 family, is split in three exons and two introns. Insertions of B1 like and LINE elements as well as several sequence repeat structures are scattered over the gene suggesting that this region of the S100 gene cluster has been the subject of a high mutational activity in mouse evolution. The insertions may represent molecular footprints of a recently postulated inversion event, which resulted in a rearrangement of the S100 gene cluster in mouse compared to man. Deletion analysis of the promoter reveals, that a 1200 bp fragment is able to direct a cell type-specific expression of a reporter gene in granulocytic 32D cells. Unexpectedly, the myeloid-specific transcription factor C/EBPepsilon is not needed for the transcriptional upregulation of the S100A9 and S100A8 genes in neutrophils. The data described here provide further insights into the evolution of the S100 gene cluster and into the myeloid-specific regulation of the murine S100A9 gene expression. J. Cell. Biochem. 80:606–616, 2001. 

**Key words:** S100A9; MRP14; calgranulin; S100 proteins; C/EBPepsilon; myeloid cells

The calcium-binding S100 protein family comprises more than 13 members, which share structural characteristics, but which display a surprisingly diverse expression pattern [Ridinger et al., 1998; Schafer et al., 1996]. S100A9, a member of this gene family (synonyma: CF-Ag, L1-Ag, calgranulin B, MRP14) is massively expressed in developing as well as mature cells of the myeloid lineage. S100A9 and S100A8 are particularly upregulated in neutrophils making up to 40% of the cytosolic proteins [Nacken et al., 2000]. Upon inflammation their expression is also induced in endothelial cells [Kerkhoff et al., 1998; Passey et al., 1999]. S100A9 is usually co-expressed with S100A8 (MRP8) and it is known, that both proteins form heterodimeric complexes [Kerkhoff et al., 1998; Nacken et al., 2000]. In the course of an inflammatory process S100A8/A9-positive cells appear very early at the site of inflammation [Kerkhoff et al., 1998]. Although there are a number of hypotheses, the exact function of S100A9 is not yet known. Recent data indicate that in human granulocytes S100A9 may modulate the affinity of the integrins to fibronectins and may thus be involved in adhesion of granulocytes to inflammatory stimulated endothelial cells [Newton and Hogg, 1998]. It could also be shown that the S100A8/S100A9 complex specifically binds unsaturated fatty acids especially

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number AJ250496 (mS100A9).

Grant sponsor: Deutsche Forschungsgemeinschaft; Grant number: SFB293.

*Correspondence to: Dr. Wolfgang Nacken, Institute of Experimental Dermatology, University of Münster, Von-Esmarch-Str. 56, 48149 Münster, Germany.
E-mail: nacken@uni-muenster.de

Received 10 July 2000; Accepted 22 August 2000


This article published online in Wiley InterScience, December XX, 2000.

arachidonic acid [Klempt et al., 1997; Siegenthaler et al., 1997]. These and other data indicate that S100A8 and S100A9 are involved in the modulation of inflammatory processes.

Most of the S100 genes are organized in a cluster in mouse and man. However, recent data indicate that both clusters are only partially collinear and that gross rearrangements have occurred during mouse evolution [Ridinger et al., 1998]. The molecular analysis of S100 genes from different species may provide new insights into the evolution of this gene family and into the transcriptional regulation of the highly specific expression of these genes. Here, the genomic locus of the murine S100A9 gene has been molecularly analyzed. We performed a promoter deletion analysis using murine 32D cells as a model system to study the upstream regulatory region of the S100A9 gene in granulocytic cells and we finally present evidence that C/EBPepsilon, a candidate regulatory factor for myeloid-expressed genes, is not involved in the regulation of the S100A9/A8 gene expression.

## MATERIAL AND METHODS

### Molecular Cloning and Sequencing

A lambda library made from genomic DNA of SVJ129 mice was screened with the mouse S100A9 (MRP14) cDNA as probe [Lagasse and Weissman, 1992]. Several cross-hybridizing clones were isolated and purified following standard procedures. *XbaI* restriction of the DNA of phage 14.7 resulted in a 6- and 2.4-kb long DNA fragment hybridizing with the S100A9 cDNA. These fragments were subcloned (p14.7XG, p14.7XK) and further analyzed (Fig. 1). Sequence analysis was performed according to standard methods. Additionally, a mouse P1 phage library (from the RZPD, Resource Centre of the Deutsches Genomforschungs Projekt, Max-Planck-Institute Berlin, Germany) was screened with a murine S100A8 and S100A9 probe. Several phages were isolated and purified, which hybridized with both probes. One of them was analyzed in detail by Southern blot analysis to prove that the P1 phage insert contained both mS100A8 and S100A9 genes (data not shown).

### Southern Blotting

Genomic DNA isolated from different mouse strains was enzymatically restricted as indicated. Subsequently, the fragments were separated by 0.8% agarose gel electrophoresis and blotted onto nylon membranes (Hybond, Pharmacia Biotech, Freiburg, Germany) according to standard procedures. A 4.5-kb long BgII DNA fragment from clone p14XG was radioactively labeled and used as probe to hybridize the DNA-filter.

### Cell Culture

32D cells were cultured in RPMI 1640 medium containing 100 pg/ml IL-3 (ICN Biochemicals, Eschwege, Germany) to sustain the promyeloic status. Granulocytic differentiation was induced by incubating the cells in 60 pg/ml G-CSF (ICN Biochemicals) as described [Nacken et al., 2000; Valtieri et al., 1987]. After 5 days of G-CSF-induced differentiation 32D cells were electroporated (see below).

### Northern Blotting

Total RNA was isolated as described earlier using the TRIZOL reagent (Gibco/BRL via Life Technologies, Karlsruhe, Germany). Twenty micrograms of RNA was loaded onto a formaldehyde containing 1.2% agarose gel. After electrophoresis the gels were blotted onto a nylon membrane (Hybond, Amersham-Pharmacia Biotech) according to standard procedures and hybridized with a radioactively labeled DNA probe.

### Polymerase Chain Reaction and RFLP Analysis

Genomic DNA was isolated from mouse tails and E14 ES-cells, respectively. Tail biopsies were digested with Proteinase K at 56°C and subsequently extracted with phenol/chloroform. Polymerase chain reaction (PCR) was performed using Ex TAQ$^{TM}$-Polymerase (Takara, BioWhittaker, Belgium) and the appropriate primers to amplify a 4.5 kb fragment of the S100A9 5′ upstream region. The amplified DNA was digested with various restriction enzymes and separated by 1.2% agarose gel electrophoresis to detect possible restriction fragment length polymorphism.

### Promoter Analysis

We used the green fluorescent protein (GFP) as a reporter to study the S100A9 promoter. The suitability to use the green fluorescent protein as reporter gene in quantitative pro-

moter studies has been shown earlier [Lissemore et al., 2000; Meng et al., 1997; Scholz et al., 2000]. The CMV promoter/enhancer was removed from pEGFP-C1 (Clontech, Dreieich, Germany), which encodes a modified form of the GFP. Promoter fragments were generated via PCR using the appropriate oligonucleotides as primers. The nucleotide sequences of the PCR products were verified by standard sequencing methods. The construct containing the 1200-bp upstream regulatory region (URR) including the short untranslated exon 1 was defined as "full length" promoter construct 14P. The deletion constructs "delta 2," "delta 3" and "delta 4" contain exon 1 and 600 bp, exon 1 and 360 bp, and exon 1 and 170 bp of the upstream regulatory region, respectively (Fig. 6). The fragments were cloned into the *Nhe*I site of pEGFP-C1 (Clontech, Dreieich, Germany) just 5′ of the transcriptional start site of the EGFP gene. The CMV (cytomegalovirus) enhancer was amplified by PCR and cloned 5′ of the "full length" promoter (CMV14P) and 5′ of the "delta 4" promoter fragment (CMVdelta 4). The constructs "delta 4intron" and "14Pintron" additionally contain the intron1 (Fig. 6). Typically 20 μg of plasmid DNA was added to $10^6$ cells and electroporated (960 μF, 250 V) as described earlier [Scholz et al., 2000]. Cells were harvested after 48 h, a cytosolic extract was prepared and fluorescence was measured using a fluoromax spectrophotometer (Spex, Model Fluoromax II spectrophotometer, Instruments SA, Munich, Germany). As an internal transfection control a CMV-β-galactosidase construct was co-transfected. Each construct was at least eight times electoporated. Additionally, three different plasmid preparations per construct were used, to avoid variations caused by potential impurities or differences in the degree of supercoiling. All values are shown in relation to the CMV-driven pEGFP-C1 ( = 100%) construct.

## RESULTS

### Molecular Structure of the Genomic Murine S100A9 Locus

To investigate the genomic structure of the murine S100A9 gene 7 kb were sequenced. The comparison with the cDNA revealed [Lagasse and Weissman, 1992], that the murine S100A9 gene consists of three exons (Fig. 1). The first exon is at least 23 bp long and represents most of the 5′-untranslated leader sequence according to the homology to the cDNA [Lagasse and Weissman, 1992]. Intron 1 spans over 373 bp. The translational start codon is located 10 bp downstream of the beginning of exon 2, which encodes the N-terminal half of the S100A9 protein and is followed by a 2-kb long intron (intron 2). The third exon encodes the C-terminal half of the S100A9 protein and includes the 3′-untranslated trailer region and a consensus polyadenylation site. All intron splice sites agree well with the typical GT-AG consensus splice site (Fig. 2).

The expression of the S100A9 gene is regulated by a classical TATA and CCAAT box containing promoter. Twenty-three bp downstream of the TATA box the putative transcriptional start site was identified by homology to the cDNA sequence [Lagasse and Weissman, 1992]. A stretch of 500 bp upstream of the TATA box is highly conserved between murine and human S100A9 [Lagasse and Clerc, 1988] sharing 60% nucleotide sequence identity (Fig. 3). Between positions 500 and 1200 the

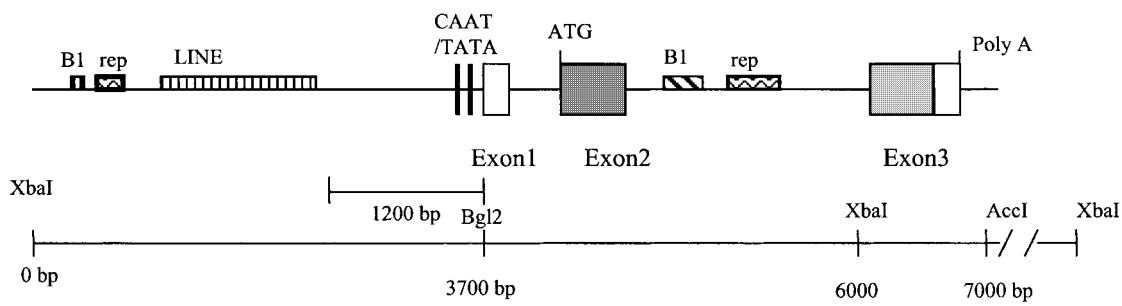
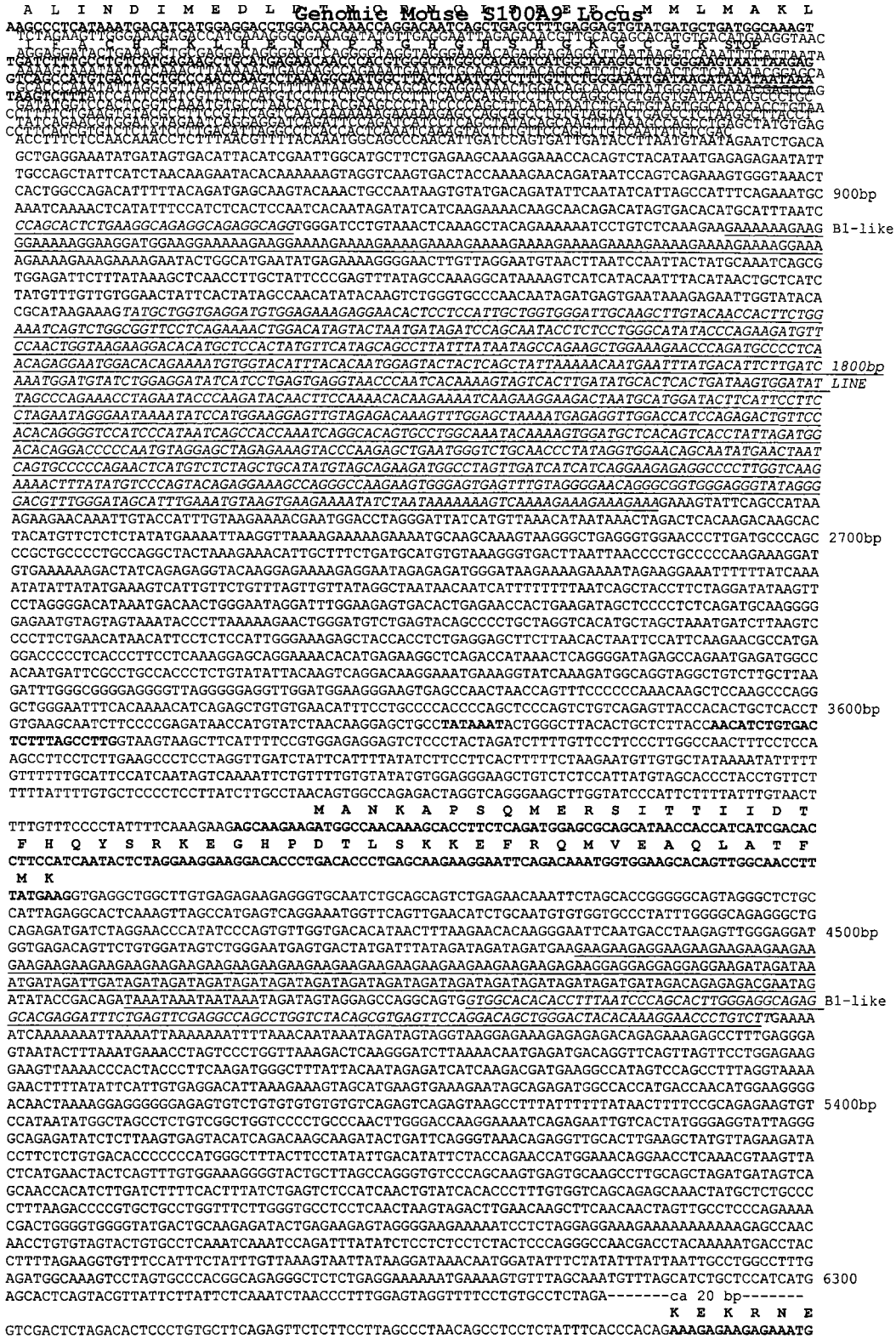



**Fig. 1.** Schematic representation of the genomic locus of the mouse S100A9 gene. The bars represent the three exons (unfilled is the untranslated, shaded the coding sequence), the position of the CCAAT box, TATA box, Start codon and the polyadenylation site are indicated. The LINE insertion, the B1-like sequences as well as the repeat structures are shown. The position of the 1200 bp region used in the promoter deletion analysis is indicated.


Genomic Mouse S100A9 Locus
900bp
B1-like
1800bp
LINE
2700bp
3600bp
4500bp
B1-like
5400bp
6300
ca 20 bp

(Fig. 2)

```
mS100A9 TGGGAAAGAGCTACC-ACCTCTGAGGAGCTTCTTAACACTAATTC-CATTCAAGAACGCC
hS100A9 TGCAAGAGGGTTGCCCACCTCTGAGAACCCTTCTAACCCCAAATCTCACCCTATGAATCT
        **  * ** * * ** ********* * * *  **** * ** ** **  * *  *  *
                                                          c-Myc
mS100A9 ATGAGGAC--CCCCTCACCCTTCC-----TCAAAGGAGCAGGAAAACACATGAGAAGGCT
hS100A9 AAGAACACAACCCCTCGCCATCCTAAGTATCACAGAGCCAGGCAAGCATGGGTGAGAGCT
        * **  **  ****** ** * *      *** **   **** ** **   * **  ***
                                GATA-2/1/3       Oct-1
mS100A9 CAGACCATAA-------ACTCAGGGGATAGAGCCAGAATGAGATGG-------CCACAAT
hS100A9 CAGACCATCCTTGTTGGACTAAAAGGA-AGGGGCAGACTGCCATGGGGGGCAGCCGAGAG
        ********         *** *  *** ** * **** **  ****       **   *

mS100A9 GATTCGCCTGCTAC----CCTCTGTATATTACAAG-TCAGGACAAGGAAATGAAAGGTAT
hS100A9 GGTCAGGCCCCCATAGGTCCTCAGCCTGCTTCAACCTCA--A—AGGGGATGGG-GGGCT
        * *  * *  * *     **** *  *  * ***  ***  *  ***  ***   **  *
                                                  SP1   MZF-1
mS100A9 CAAAGATGGCAG-GTAG--GCTGTCTTGCTTAAGATTTGGGCGGGGAGGGGTTAGGGGGA
hS100A9 GAGTGGTGCCAGAGGAGCAGCAGGCTCGCT----------CGGGGAGAG--TAGGGCCT
         *  * ** *** * **  ** * ** ***           ******* *  *****

mS100A9 GGTTGGATGGAAGGGAAGTGAGCCAACTAACCAGTTTCCCCCCAAACAAGCTCCAAGCCC
hS100A9 --TAGGATAGAAGGGAAATGAACTAAACAACCGCTT--CCTGCAAACCAGTTTCA-GGCC
        * **** ******** *** * **  ****  **  **  ***** ** * ** * ****
          NF-kappa C/EBPb
mS100A9 AGGGCTGGGAATTTCACAAAACATCAGA----GCTGTGTGAACATTTCCTGCCCCACCCC
hS100A9 AGGGCTGGGAATTTCACAAAAAAGCAGAAGGCGCTCTGTGAACATTTCCTGCCCCGCCCC
        ********************* * ****    *** ******************* ****
                          C/EBPb              CAAT-Box C/EBPb
mS100A9 AGCTCCCAGTCTGTCAGAGTTACCACACTGCTCACCTGTGAAGCAATCTTCC-----CCG
hS100A9 AGCCCCCTTCCTGGCAGCATTAGCACACTGCTCACCTGTGAAGCAATCTTCCGGAGACAG
        *** ***   *** ***  *** ******************************    * *
         GATA2                              TATA-Box
mS100A9 AGATAAC---CATGTATCTAACA-AGGAGCTGCCTATAAATACTGGGCTTACACTGCTCT
hS100A9 GGCCAAAGGGCAAGTGCCCCAGTCAGGAGCTGCCTATAAATGCCGAGCCTGCACAGCTCT
         *  **    ** **  *  *   ***************** * * ** * *** *****
```

**Fig. 3.** Comparison of the nucleotide sequence located 500 bp upstream, the TATA box of both the human [Lagasse and Clerc, 1988], and mouse S100A9 promoters. Putative transcription factor binding sites and homologous regions are indicated. mS100A9, murine S100A9 promoter region; hS100A9, human S100A9 promoter region.

degree of homology decreases significantly the more upstream the nucleotide sequence is located. About 1200 bp upstream of the TATA box we detected a sequence 1050 bp in length sharing 90% homology to the murine repetitive LINE retrotransposon (Fig. 2). Computational analysis implicates that this sequence represents a pseudogene originating from a mRNA encoding ORF1 of the LINE element. Additionally, we discovered about 2.7 kb upstream of the TATA box a 31-bp long stretch of DNA with homology to the B1 repetitive transposable element followed by a $(GAAA)_n$ repeat structure. Intron 2 also contains a 124 bp long piece of DNA, which shares 90% sequence identity with B1 elements. Additionally, a $(GAG)_n$ and $(TAG)_n$ repeat sequence is located in this intron (Figs. 1 and 2).

A restriction length polymorphism was recently discovered in the murine S100A8 (MRP8) locus [Lagasse and Clerc, 1988]. Therefore, southern blots were performed with genomic

**Fig. 2.** Nucleotide sequence of the mouse S100A9 genomic locus. Exons are written in fat and underlined letters; polyadenylation site, TATA and CCAAT box in italic letters and underlined. The repeat structures are underlined only, insertions originating from B1 and LINE elements are shown underlined and in italic letters.

DNA from BALB/c, C57/Bl6 and SJL mice. DNA was digested with different restriction enzymes to detect potential restriction length polymorphisms in the S100A9 locus. The restriction pattern of different mouse lines seemed to be identical (data not shown). However, high background signals due to the presence of repetitive DNA in the probe significantly affect the clarity of the signals. To circumvent this problem genomic DNA from ES-cells and FVB mice were used as templates to amplify a 5.5 kb genomic fragment spanning the murine S100A9 locus by PCR. Restriction analysis of the fragments—even with *Dde*I, a frequently cutting enzyme—revealed an identical restriction pattern (Fig. 4). Hence, no polymorphism was observed between these mouse lines. Additionally, screening a P1-phage based mouse genomic library, several clones were isolated, whose inserts hybridized with both S100A8 and S100A9 cDNAs indicating that both genes are located within a stretch of less than 70 kb of genomic DNA (data not shown). This finding has recently been confirmed by Ridinger et al. [1998].

### 32D Cells Express S100A9 Upon G-CSF-Induced Myeloid Differentiation

To study the strong expression of S100A9 in granulocytes, the 32D cell line was established as a model system. Cultivation of 32D cells with IL-3 maintains the undifferentiated promyelocytic status of these cells [Meng et al., 1997]. As shown by a Northern blot S100A9 expression is absent or very low in promyeloic cells, but induced by G-CSF-mediated granulocytic differentiation (Fig. 5). Recent experiments confirmed that S100A9 was immunodetectable in only 5–10% of the IL-3 propagated cells [Nacken et al., 2000]. Upon induction with G-CSF 32D cells differentiate to a granulocytic phenotype [Valtieri et al., 1987] and after 8 days almost 100% of the cells expressed S100A9 [Nacken et al., 2000]. Thus, murine 32D cells appeared to be a suitable cell line to study the expression of S100A9 during granulocytic differentiation.

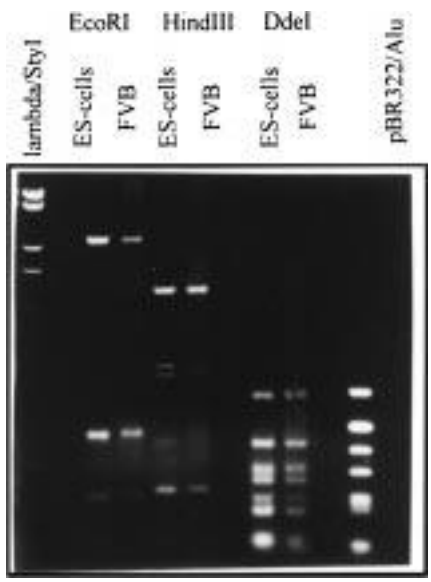



**Fig. 4.** RFLP analysis: Genomic DNA of FVB mice and ES-cells served as templates to generate a 4.5 kb long PCR-fragment, which was digested with the enzymes indicated. No polymorphism could be detected.

### Analysis of the Upstream Regulating Region of the Murine S100A9 Gene

To investigate whether the region downstream of the LINE insertion and upstream of the putative transcriptional start site contains all the necessary elements to drive the expression of a gene in granulocytes, a promoter analysis was performed. The expression of a CMV-promoter-driven construct (pEGFP-C1) was set as 100% in each cell line. The delta 4 construct (Fig. 6) containing 170 bp upstream of the putative transcriptional start site was sufficient to drive the reporter gene in 32D cells (Fig. 6) as well as in COS cells (data not shown) and was defined as the basic or minimal

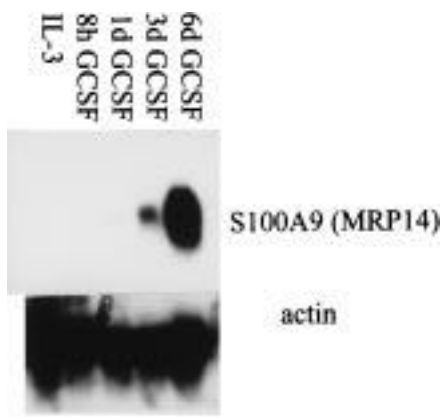



**Fig. 5.** Northern blot: 20 µg total RNA of IL3-32D cells (left) and after 4 h, 8 h, 1, 3, and 6 days of G-CSF-propagated 32D cells was loaded on a formaldehyde containing agarose gel, blotted and probed with the murine S100A9 cDNA and actin as indicated.

promoter. The insertion of the CMV enhancer fragment upstream of the "minimal promoter" delta 4 strongly enhanced the expression in both cell lines. Addition of the CMV enhancer to the "full length" 14P construct (CMV14P), however, did not significantly stimulate the expression in whichever cell type compared to the 14P promoter alone (Fig. 6). At least in S100A9 negative cells the upregulating effect of the CMV enhancer was obviously suppressed by upstream regulating sequences of the murine S100A9 gene.

An enhancing effect of the first intron of the human S100A9 gene has recently been reported [Valtieri et al., 1987]. Here, the addition of the first exon/intron did not stimulate expression either in IL-3 sustained cells or in G-CSF-induced 32D cells (Fig. 6).

Using the IL-3 sustained promyeloic 32D cells, which do not express S100A9, the expression level hardly exceeded 10% of the positive CMV driven control irrespective of the construct used. Only the CMVdelta 4 construct mediated a high expression level (Fig. 6). In contrast, in G-CSF-differentiated 32D cells the full 1200-bp long promoter fragment (14P) mediated a reporter gene activity, which is as strong as the CMVdelta 4 reaching 50% versus 60% of the pEGFP activity (Fig. 6) suggesting that this region mediates cell type-specific expression of the S100A9 gene.

### C/EBP Epsilon −/− Mice Express S100A8 and A9

A search for potential transcription factor binding sites in the murine S100A9 promoter revealed several putative C/EBP (CCAAT enhancer binding protein) binding sites (Fig. 1). It has been shown, that members of the C/EBP family are involved in the regulation of the human S100A9 gene expression [Klempt et al., 1998; Kuruto-Niwa et al., 1998]. A recently isolated member of this family, C/EBP epsilon is specifically upregulated in the myeloid cell lineage and has been shown to play a positive regulatory role for a number of genes, whose expression is induced during granulopoesis as those encoding secondary and tertiary granule proteins (Fig. 7) [Lekstrom-Himes et al., 1999; Yamanaka et al., 1997]. Since the upregulation of these genes correlates well with that of S100A9 and its dimerisation partner S100A8, we hypothesized, that C/EBP epsilon might be a candidate regulatory factor for the latter. A Northern analysis showed that C/EBPepsilon is present in both IL-3 as well as in G-CSF cells (Fig. 7). However, the C/EBPepsilon message seems not to be upregulated during differentiation in these cells in contrast to the murine S100A9, S100A8 and a secondary granule protein (Fig. 7). Subsequently, we analyzed the expression of S100A9 and S100A8 in the C/

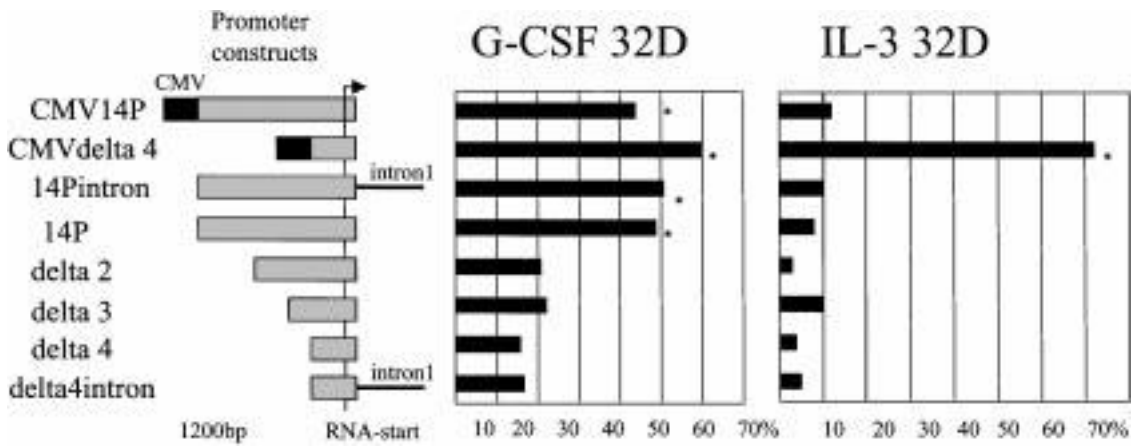



**Fig. 6.** Promoter analysis: Left panel shows the promoter mS100A9 constructs used for the transfection studies: shaded bars represent the mS100A9 sequence, black box represents the CMV enhancer. The arrow marks the putative transcriptional start site. Right panels: Promoter analysis of the mS100A9 upstream regulating region using 32D cells. IL-3: Interleukin 3 maintained 32D cells; G-CSF: granulocyte colony stimulating factor differentiated 32D cells. Stars indicate significant differences compared to the ''minimal promoter'' delta 4. Each experiment was repeated at least eight times. Hundred percent is the normalized activity for the pEGFP-C1 construct.

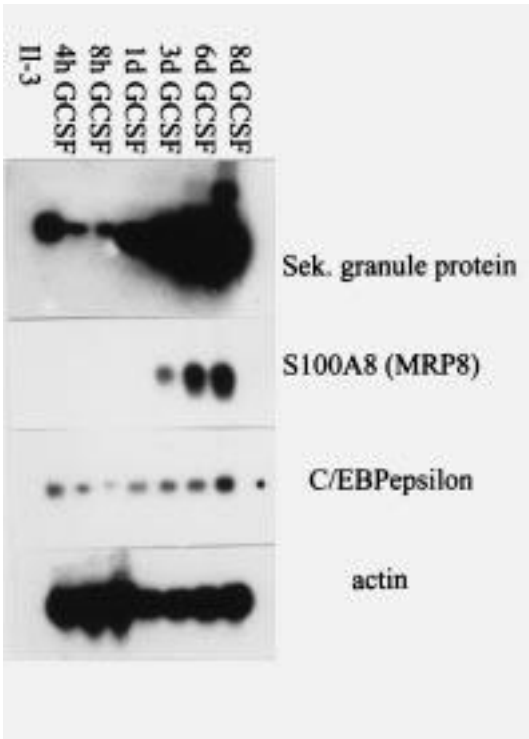



**Fig. 7.** Northern analysis: Total RNA from IL3 and G-CSF-treated 32D cells was loaded onto a denaturing agarose gel and probed with the cDNAs of the secondary granule protein bactenecin (EMBL Ac.No: MMU95002), S100A8, C/EBPepsilon and actin as indicated.


**Fig. 8.** Northern RNA from bone marrow from C/EBPepsilon −/− mice and +/+ mice was probed with the S100A9, S100A8 and an actin cDNA probe. The weak signal of the S100A8 probe is due to weaker radioactive labeling. Lower panel: Western blot of spleen extracts of +/+ and C/EBPepsilon −/− mice probed with a rabbit polyclonal anti-murine S100A9 antibody.

EBPepsilon −/− mice. Expression of both S100 genes in bone marrow is indistinguishable in mutant and wild-type mice (Fig. 8). In agreement with these results S100A9 was also detected in the spleen via Western blotting experiments in both the mutant and the wild-type mice (Fig. 8).

## DISCUSSION

The S100 proteins have several structural features in common with respect to gene and protein structure [Nacken et al., 2000; Schafer and Heizmann, 1996]. They consist of two calcium-binding EF-hand domains bridged by a hinge region. The genomic characterization of the murine S100A9 mirrors this bipartite structure in that each domain is encoded by a separate exon. The coding region spans over exon 2 and exon 3, each of which encodes one EF-hand domain [Lagasse and Clerc, 1988; Nacken et al., 1996; Zimmer et al., 1996]. The short exon 1 encodes most of the untranslated leader sequence. The exon–intron structure of the murine S100A9 gene is highly homologous to the human S100A9 gene (Fig. 1). Even the length of the introns is comparable. Thus, the murine S100A9 locus confirms earlier findings, that the genomic structure of S100 genes is quite similar at the molecular level [Zimmer et al., 1996].

However, the molecular characterization of the murine S100A9 gene revealed some peculiarities, namely several repeat structures and a number of bits and pieces originating from retrotransposable elements scattered over the whole gene (Fig. 2). The identification of such an accumulation of insertions and repeat

structures suggests that the murine S100A9 locus may represent a mutational hot spot. The generation of a gene family is usually explained by sequential duplication events of an ancestor gene and subsequent evolutionary divergence of the copies. This divergence is driven by molecular mutations within the individual gene copies. Interestingly, the nucleotide sequence of the S100A9 and the S100A8 cDNA of mouse and man are $< 80\%$ identical, lower than expected between "identical" genes of mouse and man [Lagasse and Clerc, 1988; Nacken et al., 1996]. Other S100 proteins share up to 95% homology among different vertebrate species as S100A1 and S100A10 [Zimmer et al., 1996]. The insertions and repeat structures present molecular evidence of high mutational activity specifically in this region of the mouse genome, which may have led to the limited relationship between the human and murine S100A9/A8 genes.

Additionally, the structure of the S100 cluster of mouse and man appears to be only partially colinear [Ridinger et al., 1998]. An inversion event which involves the S100A8/A9 genes has been postulated to explain the different organization of the S100 gene cluster in mouse and man [Ridinger et al., 1998]. The activity of transposons is frequently correlated with inversions and/or gross rearrangements. Here, we may have found the molecular footprints originating from such an event. We could not detect any restriction fragment length polymorphism among the different mouse lines in the S100A9 gene (Fig. 4) indicating that mutational changes occurred very early in mouse evolution.

Although a number of insertions and repeat structures were found in the upstream regulatory region of the murine S100A9 gene, the expression pattern of both the human and the murine S100A9 gene has been found to be identical. A 500-bp long region upstream of the putative transcriptional start site is highly homologous to the corresponding human promoter region (Fig. 3). Interestingly, this stretch of DNA is not affected by insertions or repeat structures. Deletion constructs were used to define the basic promoter and to confirm the importance of the highly conserved promoter region for the expression. A 170 bp fragment was able to drive the reporter gene at a low level in COS cells (data not shown) and in IL-3 as well as in G-CSF-sustained 32D cells (Fig. 6). The addition of a CMV enhancer (CMVdelta 4) dramatically increased this expression in the cell lines as expected. We conclude that this fragment contains all the elements of a non-cell type-specific basic promoter.

Intron 1 of the human S100A9 gene has been shown to stimulate gene expression manifold [Melkonyan et al., 1998]. The first intron of the murine gene seems to be a transcriptional enhancer less potent than the corresponding intron of the human S100A9. However, the biggest enhancing effect of the human intron I was observed in heterologous fibroblasts using the heterologous herpes simplex thymidine kinase (HSV tk) promoter. The effect was far lower (less than two fold) when the authors used the homologous promoter [Melkonyan et al., 1998]. Here, we could not find any enhancement of the murine intron 1 on reporter gene expression in 32D cells. By using the weak HSV-tk promoter, the effect of the intron in the homologous environment may have been overestimated in this earlier study.

We hypothesized, that the 1200 bp fragment between the LINE insertion and the putative transcriptional start site, is most likely responsible for the massive expression of the S100A9 gene in myeloid cells. The CMV-enhanced minimal promoter construct CMVdelta4 mediated the strongest expression of the reporter gene in all cell lines used. However, the expression of the "full length" 1200 bp promoter-driven construct (14P) in GCSF-propagated 32D cells is remarkably high and reaches the level of the CMVdelta 4 promoter (50% vs. 60% of the pEGFP-C1 level), whereas it dramatically drops in IL-3-sustained 32D cells (9% vs. 70%) (Fig. 6) suggesting that the sequences upstream of the basic promoter have a strong positive effect on S100A9 expression in granulocytes and are likely involved in the cell-type specific upregulation of the S100A9 gene transcription. However, the threefold upregulation of the 1200 bp 14P promoter compared to the 170 bp "minimal" promoter cannot solely explain the massive upregulation of this gene in granulocytes. Nevertheless, the comparison between the S100A9 negative IL-3 32D and the S100A9 positive GCSF-32D cells suggests that the 1200 bp region is mainly involved in the cell-type specificity of the S100A9 expression.

Interestingly, even the low level of the "full length" 14P expression observed in the S100A9

negative in IL-3 32D cells (Fig. 6) remained low upon CMV enhancer insertion, whereas it was strongly upregulated by the addition of the CMV enhancer to the basic promoter (CMVdelta 4). We conclude, that the region upstream of the minimal promoter may suppress the enhancer effect in S100A9 negative cells. Further experiments are needed to investigate whether a similar mechanism is involved in mediating the cell-type specificity of S100A9 expression.

Four members of the CCAAT/enhancer binding protein family (C/EBPs), C/EBPalpha, C/EBPbeta, C/EBPdelta, and C/EBPepsilon, play a decisive role in the myeloid differentiation and regulation of myeloid-specific genes [Yamanaka et al., 1998]. Recently, Kuruto-Niwa et al. [1998] demonstrated that C/EPBalpha and C/EBPbeta were able to bind to the C/EBP motif at position 81 in the human S100A9 promoter, which is also present at the homologous position in the murine promoter. C/EBPepsilon, however, is the only member which is specifically upregulated in the myeloid cell lineage [Yamanaka et al., 1997]. C/EBPepsilon has been shown to induce the expression of genes as those encoding secondary granule proteins (Fig. 7) [Lekstrom-Himes et al., 1999], whose upregulation correlates well with that of S100A9/A8. Therefore, we speculated that this factor may be directly involved in the regulation of the granulocytic expression of both genes. The analysis of the C/EBPepsilon −/− mice [Lekstrom-Himes et al., 1999; Verbeck et al., 1999], however, indicates that C/EBPepsilon is not required for the massive expression of both S100A9 and A8 genes during myeloid differentiation (Fig. 8). A pathway different from the secondary granule protein encoding genes has to be postulated for the granulocytic upregulation of both S100 genes.

In summary, we report the genomic structure of the murine S100A9 and its chromosomal linkage with S100A8. Some striking insertions originating from retrotransposable elements and some repeat structures as well as a restriction length polymorphism analysis suggest that this locus may have been subject of a high mutational activity early in mouse evolution. A promoter analysis and the position of a pseudogene insertion suggest that the granulocytic expression of S100A9 is mainly regulated by a 1200-bp long region upstream of the putative transcriptional start site box. Evidence is provided that the myeloid-specific transcription factor C/EBPepsilon is not involved in the upregulation of the S100A9 gene expression.

## ACKNOWLEDGMENTS

This work was supported by a grant of the Deutsche Forschungsgemeinschaft (SFB293) to W.N. and C.S. We thank Dieter Wiesmann for excellent technical help and Dr. Claus Kerkhoff for critical reading of the manuscript.

## REFERENCES

Kerkhoff C, Klempt M, Sorg C. 1998. Novel insights into structure and function of MRP8 (S100A8) and MRP14 (S100A9). Biochim Biophys Acta 1448:200–211.

Klempt M, Melkonyan H, Hofmann HA, Eue I, Sorg C. 1998. The transcription factors c-myb and C/EBP alpha regulate the monocytic/myeloic gene MRP14. Immunobiology 199:148–151.

Klempt M, Melkonyan H, Nacken W, Wiesmann D, Holtkemper U, Sorg C. 1997. The heterodimer of the $Ca^{2+}$-binding proteins MRP8 and MRP14 binds to arachidonic acid. FEBS Lett 408:81–84.

Kuruto-Niwa R, Nakamura M, Takeishi K, Nozawa R. 1998. Transcriptional regulation by C/EBP alpha and C/EBP beta in the expression of the gene for the MRP14 myeloid calcium binding protein. Cell Struct Funct 23: 109–118.

Lagasse E, Clerc RG. 1988. Cloning and expression of two human genes encoding calcium-binding proteins that are regulated during myeloid differentiation. Mol Cell Biol 8:2402–2410.

Lagasse E, Weissman IL. 1992. Mouse MRP8 and MRP14, two intracellular calcium-binding proteins associated with the development of the myeloid lineage. Blood 79:1907–1915.

Lekstrom-Himes JA, Dorman SE, Kopar P, Holland SM, Gallin JI. 1999. Neutrophil-specific granule deficiency results from a novel mutation with loss of function of the transcription factor CCAAT/enhancer binding protein epsilon. J Exp Med 189:1847–1852.

Lissemore JL, Jankowski JT, Thomas CB, Mascotti DP, deHaseth PL. 2000. Green fluorescent protein as a quantitative reporter of relative promoter activity in *E. coli*. Biotechniques 86:88–89.

Melkonyan H, Hofmann HA, Nacken W, Sorg C, Klempt M. 1998. The gene encoding the myeloid-related protein 14 (MRP14), a calcium-binding protein expressed in granulocytes and monocytes, contains a potent enhancer element in the first intron. J Biol Chem 273:27026–27032.

Meng A, Tang H, Ong BA, Farrell MJ, Lin S. 1997. Promoter analysis in living zebrafish embryos identifies a *cis*-acting motif required for neuronal expression of GATA-2. Proc Natl Acad Sci USA 94:6267–6272.

Nacken W, Manitz MP, Sorg C. 1996. Molecular characterisation of the genomic locus of the mouse MRP8 gene. Biochim Biophys Acta 1315:1–5.

Nacken W, Sopalla C, Pröpper C, Sorg C, Kerkhoff C. 2000. Biochemical characterization of the murine S100A9

(MRP14) protein suggests that it is functionally equivalent to its human counterpart despite its low degree of sequence homology. Eur J Biochem 2000 267:560–565.

Newton RA, Hogg N. 1998. The human S100 protein MRP-14 is a novel activator of the beta 2 integrin Mac-1 on neutrophils. J Immunol 160:1427–1435.

Passey RJ, Xu K, Hume DA, Geczy CL. 1999. S100A8: emerging functions and regulation. J Leukoc Biol 66:549–556.

Ridinger K, Ilg EC, Niggli FK, Heizmann CW, Schafer BW. 1998. Clustered organization of S100 genes in human and mouse. Biochim Biophys Acta 1448:254–263.

Schafer BW, Heizmann CW. 1996. The S100 family of EF-hand calcium-binding proteins: functions and pathology. Trends Biochem Sci 21:134–140.

Scholz O, Thiel A, Hillen W, Niederweis M. 2000. Quantitative analysis of gene expression with an improved green fluorescent protein. Eur J Biochem 267:1565–1570.

Siegenthaler G, Roulin K, Chatellard-Gruaz D, Hotz R, Saurat JH, Hellman U, Hagens G. 1997. A heterocomplex formed by the calcium-binding proteins MRP8 (S100A8) and MRP14 (S100A9) binds unsaturated fatty acids with high affinity. J Biol Chem 272:9371–9277.

Valtieri M, Tweardy DJ, Caracciolo D, Johnson K, Mavilio F, Altmann S, Santoli D, Rovera G. 1987. Cytokine-dependent granulocytic differentiation. Regulation of proliferative and differentiative responses in a murine progenitor cell line. J Immunol 138:3829–3835.

Verbeek W, Lekstrom-Himes J, Park DJ, Dang PM, Vuong PT, Kawano S, Babior BM, Xanthopoulos K, Koeffler HP. 1999. Myeloid transcription factor C/EBPepsilon is involved in the positive regulation of lactoferrin gene expression in neutrophils. Blood 94: 3141–3150.

Yamanaka R, Kim GD, Radomska HS, Lekstrom-Himes J, Smith LT, Antonson P, Tenen DG, Xanthopoulos KG. 1997. CCAAT/enhancer binding protein epsilon is preferentially up-regulated during granulocytic differentiation and its functional versatility is determined by alternative use of promoters and differential splicing. Proc Natl Acad Sci USA 94:6462–6467.

Yamanaka R, Lekstrom-Himes J, Barlow C, Wynshaw-Boris A, Xanthopoulos KG. 1998. CCAAT/enhancer binding proteins are critical components of the transcriptional regulation of hematopoiesis (Review). Int J Mol Med 1:213–221.

Zimmer DB, Chessher J, Song W. 1996. Nucleotide homologies in genes encoding members of the S100 protein family. Biochim Biophys Acta 1313:229–238.