Available online at www.sciencedirect.com
SCIENCE DIRECT®
Genomics 84 (2004) 10–22
GENOMICS
www.elsevier.com/locate/ygeno

Probing the S100 protein family through genomic and functional analysis ☆

Timothy Ravasi,[a,b,*] Kenneth Hsu,[c] Jesse Goyette,[c] Kate Schroder,[a] Zheng Yang,[c] Farid Rahimi,[c] Les P. Miranda,[b,1] Paul F. Alewood,[b] David A. Hume,[a,b] and Carolyn Geczy[c]

[a] *SRC for Functional and Applied Genomics, CRC for Chronic Inflammatory Diseases, University of Queensland, Brisbane 4072, QLD, Australia*
[b] *Institute for Molecular Bioscience, University of Queensland, Brisbane 4072, QLD, Australia*
[c] *Cytokine Research Unit, School of Medical Sciences, University of New South Wales, Sydney 2052, NSW, Australia*

Received 25 November 2003; accepted 2 February 2004

Available online 10 May 2004

Abstract

The EF-hand superfamily of calcium binding proteins includes the S100, calcium binding protein, and troponin subfamilies. This study represents a genome, structure, and expression analysis of the S100 protein family, in mouse, human, and rat. We confirm the high level of conservation between mammalian sequences but show that four members, including S100A12, are present only in the human genome. We describe three new members of the S100 family in the three species and their locations within the S100 genomic clusters and propose a revised nomenclature and phylogenetic relationship between members of the EF-hand superfamily. Two of the three new genes were induced in bone-marrow-derived macrophages activated with bacterial lipopolysaccharide, suggesting a role in inflammation. Normal human and murine tissue distribution profiles indicate that some members of the family are expressed in a specific manner, whereas others are more ubiquitous. Structure–function analysis of the chemotactic properties of murine S100A8 and human S100A12, particularly within the active hinge domain, suggests that the human protein is the functional homolog of the murine protein. Strong similarities between the promoter regions of human S100A12 and murine S100A8 support this possibility. This study provides insights into the possible processes of evolution of the EF-hand protein superfamily. Evolution of the S100 proteins appears to have occurred in a modular fashion, also seen in other protein families such as the C2H2-type zinc-finger family.

The S100 proteins belong to the superfamily of EF-hand calcium binding proteins that contribute to key cellular pathways involving calcium as a second messenger. Some S100 proteins have strong associations with human disease [4,23,38]. S100 proteins are encoded by a multigene family, and 19 members have been reported in the human genome [3,26]. The major S100 gene cluster is located on chromosome 1q21.3 [3–5] in human, 3f2 in the mouse [15,37], and 2q34 in the rat (Ensembl 13.2.1). The S100B locus, which is linked to the neuronal pathology associated with Alzheimer disease and Down syndrome, is located outside these clusters [5,10] on human chromosome 21q22.3 and mouse chromosome 10b5.3. S100B has long been a marker of malignant melanoma and is implicated in regulation of the cell cycle and blockade of the tumor suppressor activity of p53.

Most S100 proteins have a mass of ~9–14 kDa with a conserved C-terminal EF hand with high affinity for calcium and an N-terminal noncanonical EF hand, containing two additional amino acids, which, in some S100s, binds Ca^{2+} with low affinity. S100 dimerization is important for Ca^{2+}-dependent conformational changes that may facilitate recognition of target proteins [reviewed in 4] within cells, and formation of homo- or heterodimers may provide a second tier of functional regulation. The diverse and numerous functions attributed to S100 proteins are likely to depend on their particular binding partner, type of posttranslational modification (such as phosphorylation), and the metal ions bound. The differing affinities of the EF hands for Ca^{2+} may mediate Ca^{2+}-dependent responses in the intracellular and extracellular compartments. Intracellular functions include regulation of Ca^{2+} homeostasis, cytoskeletal dynamics, and regulation of key

☆ Supplementary data for this article may be found on ScienceDirect.

* Corresponding author. Fax: +61-7-3365-4388.
E-mail address: t.ravasi@imb.uq.edu.au (T. Ravasi).
[1] Current address: Gryphon Therapeutics, Inc., 600 Gateway Boulevard, South San Francisco, CA 94080, USA.

0888-7543/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2004.02.002

enzymes, including phosphorylation/dephosphorylation of target proteins by various kinases and phosphatases, energy metabolism, cell differentiation, regulation of the cell cycle, and metastatic transformation [22]. Despite the lack of signal sequences for secretion, numerous S100s occur extracellularly and functions include trophic and toxic effects on the nervous system [4], cell migration (chemotaxis), and proinflammatory functions affecting leukocyte adhesion, transmigration, and activation [30], some of which are mediated via signaling following binding to the receptor for advanced glycosylation end products (RAGE) [4,12]. S100A8 has antioxidant activity [33]. Some S100s bind zinc, and copper may bind at the same site. These cations may regulate important intracellular homeostatic functions and extracellular functions such as the antiparasitic and antifungal activities of S100A12 and S100A9 [28]. Although the S100 proteins are structurally related, differences in expression profiles and subtle differences in activity may underlie nonredundant functions. Gene deletion of S100A8 in the mouse results in rapid and synchronous embryo resorption by day 9.5 of development, indicating a necessary function in trophoblast invasion and/or maternal recognition [32].

The "myeloid-associated" S100s, S100A8, S100A9, and S100A12 ("calgranulins"), are constitutively expressed at high levels in neutrophils and are associated with various inflammatory conditions, including cystic fibrosis, rheumatoid arthritis, psoriasis, and infection [21,36]. Human S100A8 and S100A9 may represent epidermal differentiation genes, and their expression is associated with hyperproliferative responses [7] and wound healing [43]. The human genes are overexpressed in psoriatic lesions of 1q-linked families and may be associated with the disease-susceptibility locus, PSORS4 [40].

The murine (m) S100A8 gene is regulated by pro- and anti-inflammatory mediators in macrophages, microvascular endothelial cells, epithelial cells, and fibroblasts. Our earlier studies demonstrated a potent chemotactic activity of mS100A8 for monocytes and neutrophils in vitro and in vivo. This activity was not found with human S100A8, which has only 57% amino acid identity, posing the question of whether the two proteins were orthologs [20] and, by extension, whether there existed an equivalent chemotactic peptide in humans. We subsequently isolated a functional S100A8 homolog from human neutrophils, S100A12, that is a potent monocyte chemoattractant [25,48]. The functional and sequence divergence suggests complex evolution of the S100 family in mammals.

Earlier studies indicated that S100 proteins are phylogenetically recently evolved and linkage relationships between the S100A8–S100A9 and the S100A3–S100A4–S100A5–S100A6 genes were identified [37]. YAC cloning indicated that most genes encoding members of the S100 family are structured as a conserved chromosal cluster [37,39]. The availability of the mouse, human, and rat genome sequences allows a comprehensive comparison of S100 cluster evolution in mammals. We now confirm that four S100 genes present in the human genome are absent from the mouse and rat genomes and have identified three new genes in the S100 family within the same chromosomal cluster in the three species. We have carried out gene expression, structure, and functional analysis to establish functional onthology, as well as phylogenetic analysis of the genes encoding EF-hand proteins. Based on this information we present a revised nomenclature proposal.

Results and discussion

The S100 proteins

The EF-hand superfamily includes proteins such calmodulin, parvalbumin, troponins, CaBPs,[1] and S100s, all of which share a structurally related modular Ca^{2+} binding motif [3,4,10]. Most S100 proteins also have an S100/ICaBP-type Ca^{2+} binding motif (IPR001751) not present in other EF-hand-containing proteins. This represents the low-affinity site, whereas the canonical EF hand is a high-affinity Ca^{2+} binding site shared by most members of the superfamily; two exons (2 and 3) encode each of the two calcium binding motifs. Because these structurally different proteins lie within the same chromosomal cluster they are likely to be evolutionarily related and are designated as members of the S100 family.

Based on previous reports [4,37], and on sequences deposited in NCBI, we collected sequences corresponding to 21 human, 19 mouse, and 19 rat S100 proteins. Of the human genes, 17 were located within the 1q21.3 cluster and 14 each within the 3f2 or 2q34 cluster of the mouse and rat chromosomes, respectively (Table 1); exceptions include S100zeta, S100P, S100B, and calbindin D-9k [37] (Table 1, Figs. 3A and 3B). Profillagrin, trichohyalin [45] and repetin [18] are larger proteins containing an S100/ICaBP-type motif and represent "fused genes" of S100 protein followed at the 5′ end by tandem peptide repeats typical of multifunctional epidermal matrix proteins [45]. Profillagrin and trichohyalin are located on the same chromosome, but outside the core S100 cluster, and possibly represent a separate branch of the EF-hand superfamily, whereas repetin and calbindin lie within the S100 subfamily (Fig. 1).

No sequences related to S100A2, S100A7, or S100A12 were found within the S100 cluster or anywhere else within the genome sequences of rat and mouse. There was also no apparent homolog of S100P. We performed phylogenetic analysis of the EF-hand superfamily. The full-length protein sequences of its members were clustered

[1] *Abbreviations used*: CaBP, calcium binding protein subfamily; SSAHA, sequence search and alignment by hashing algorithm; HPRT, hypoxanthine phosphoribosyltransferase; SD, standard deviation; READ, Riken Expression Array Database; N-J, neighbor-joining.

Table 1
Human, mouse, and rat EF-hand superfamily

Proposed gene symbol	Mouse nucleotide accessions No.	Mouse Chr/band	Rat nucleotide accessions No.	Rat Chr/band	Human nucleotide accessions No.	Human Chr/band
S100A1	NM_011309	3f2	S68809	2q34	NM_006271	1q21.3
S100A2	No match[a]	No match	No match	No match	NM_005978	1q21.3
S100A3	NM_011310	3f2	NM_053681	2q34	NM_002960	1q21.3
S100A4	XM_283861	3f2	NM_012618	2q34	NM_019554	1q21.3
S100A5	NM_011312	3f2	AF087469	2q34	NM_002962	1q21.3
S100A6	NM_011313	3f2	NM_053485	2q34	NM_014624	1q21.3
S100A7a	Kulski J.K. et al. 2003	3f2	Kulski J.K. et al. 20003	2q34	XM_048124	1q21.3
S100A7b	Kulski J.K. et al. 2004	3f2	Kulski J.K. et al. 20004	2q34	XM_060509	1q21.3
S100A7c	Kulski J.K. et al. 2005	3f2	Kulski J.K. et al. 20005	2q34	NM_002963	1q21.3
S100A7d	Kulski J.K. et al. 2006	3f2	Kulski J.K. et al. 20006	2q34	XM_060508	1q21.3
S100A7e	Kulski J.K. et al. 2007	3f2	Kulski J.K. et al. 20007	2q34	Kulski J.K. et al. 20003	1q21.3
S100A8	NM_013650	3f2	NM_053822	2q34	NM_002964	1q21.3
S100A9	NM_009114	3f2	NM_053587	2q34	NM_002965	1q21.3
S100A10	NM_009112	3f2	NM_031114	2q34	NM_002966	1q21.3
S100A11	NM_005620	3f2	XM_215598	3q34	MM41341	1q21.3
S100A12	No match	No match	No match	No match	NM_005621	1q21.3
S100A13	NM_009113	3f2	ENSRNOT00000016519	2q34	X99920	1q21.3
S100A14	NM_025393	3f2	RNOR01075109	2q34	NM_020672	1q21.1
S100A15 (2300002L21Rik)[b]	NM_026416	3f2	ENSRNOT00000016499	2q34	NM_080388	1q21.3
S100A17 (5430400H23Rik)[b]	NM_027762	3f2	XM_227375	2q34	XM_060104	1q21.3
S100A16 (A530063N20Rik)[b]	AK041026	3f2	ENSRNOT00000012771	2q34	AL356504	1q21.3
Profilaggrin	AF510860	3f2	AY102923	2q34	M60500	1q21.3
Repetin	NM_009100	3f2	XM_227371	2q34	AL589986	1q21.3
Calneuron1	NM_021371	5f	ENSRNOT00000001184	12q16	NM_031468	7q11.22
Calcium binding protein-1 (CaBP1)	NM_013879	5f	NM_133529	12q16	NM_004276	12q24.31
Troponin T3	NM_011620	7f5	M15202	1q37	NM_006757	11p15.5
Troponin I	NM_009406	7a1	NM_017144	1q12	NM_000363	19q13.42
Troponin T1	NM_011618	7a1	NM_134388	No match	NM_003283	19q13.42
Calcium binding protein-5 (CaBP5)	NM_013877	7a2	ENSRNOT00000020090	1q12	NM_019855	19q13.33
Calcium binding protein-2 (CaBP2)	NM_013878	19a	ENSRNOT00000024770	1q41	NM_031204	11q13.3
Calcium binding protein-4 (CaBP4)	NM_144532	19a	Not available	No match	AC005849	11q13.3
Troponin T2	NM_011619	1f	NM_012676	13q13	NM_000364	1q32.1
Troponin C	NM_009394	2h3	ENSRNOT00000020348	3q42	NM_003279	20q13.12
Hippocalcin	NM_010471	4d2.3	NM_017122	6q14	NM_002143	1p35.1
S100 beta	NM_009115	10b5.3	NM_013191	20p12	NM_006272	21q22.3
Calcium binding protein-7 (CaBP7)	NM_138948	11a1	ENSRNOT00000010439	14q21	ENST00000216144	22q12.2
Calmodulin	NM_007589	12f1	NM_012518	6q31	NM_001743	14q32.11
S100zeta	XM_176783	13d1	ENSRNOT00000024202	2q12	NM_130772	5q13.3
1700028N11Rik[b]	NM_029341	15a2	ENSRNOT00000023378	2q16	BC017596	5p13.2
Parvalbumin-alpha	NM_013645	15e2	NM_002854	22q12.3	NM_022499	7q34
Calbindin D-9k	NM_009789	Xf3	NM_012521	Xq32	NM_004057	Xp22.13
Trichohyalin	XM_195464	3f2	XM_227373	2q34	L09190	1q21
S100P	No match	No match	No match	No match	NM_005980	4p16.1

The genomic location was determined by blasting the corresponding transcripts against the assembly of mouse whole-genome sequence data as described in Materials and methods. Newly described S100 proteins are presented in boldface.

[a] "No match" indicates those genes that we failed to detect in the respective genome.

[b] Newly described members of the EF-hand superfamily.

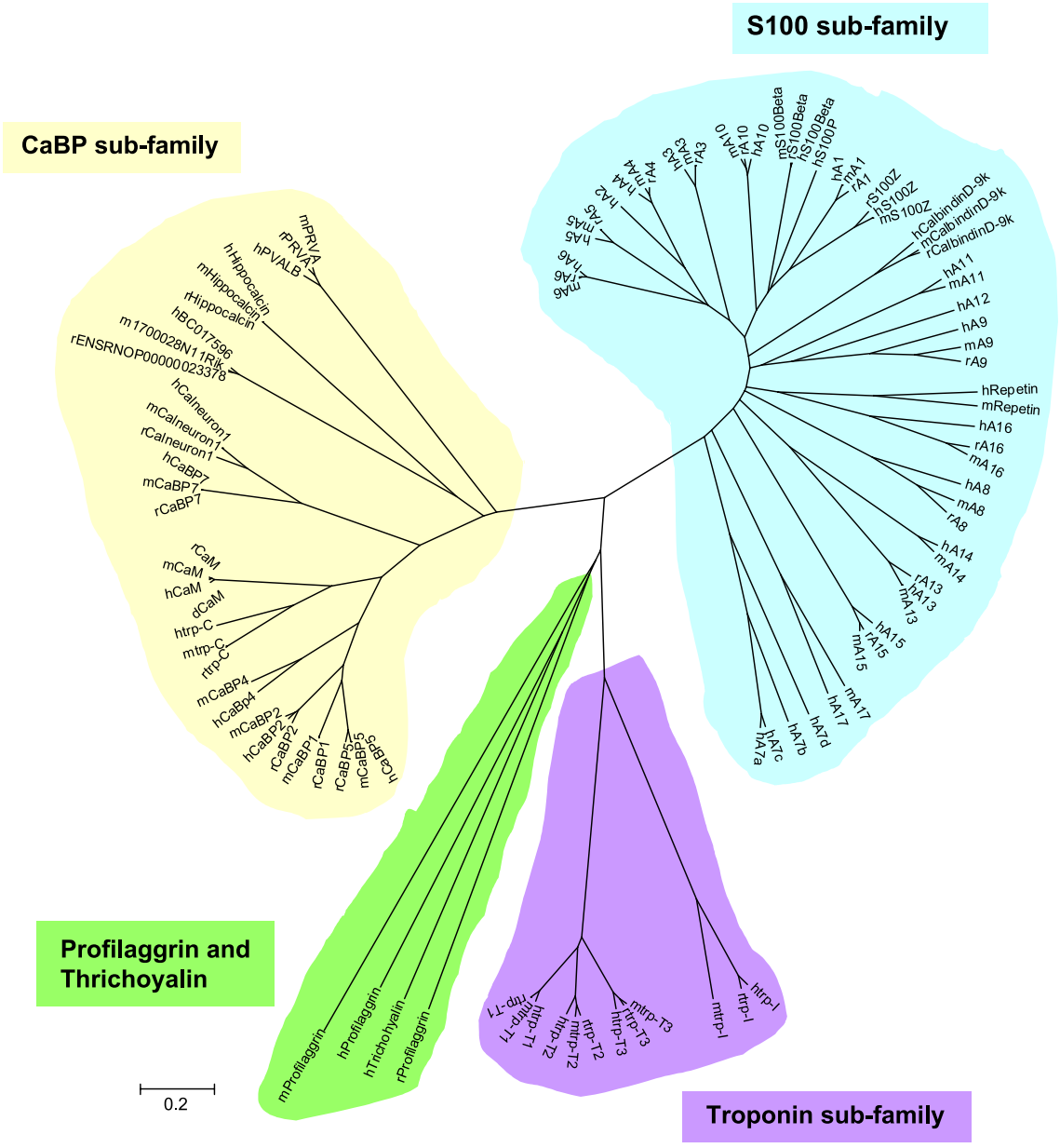
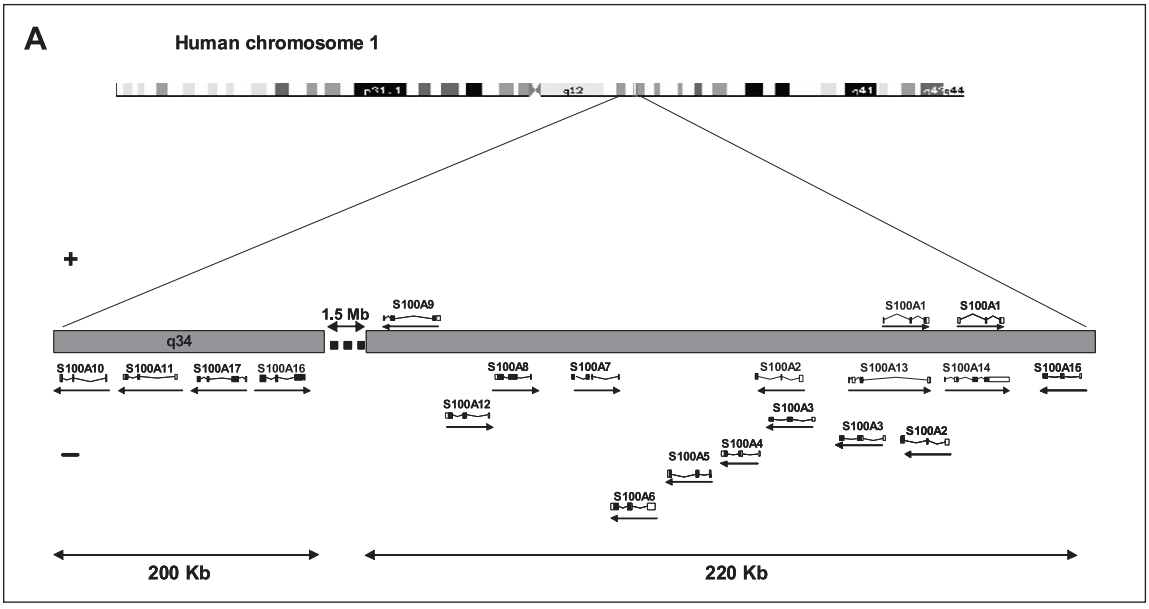


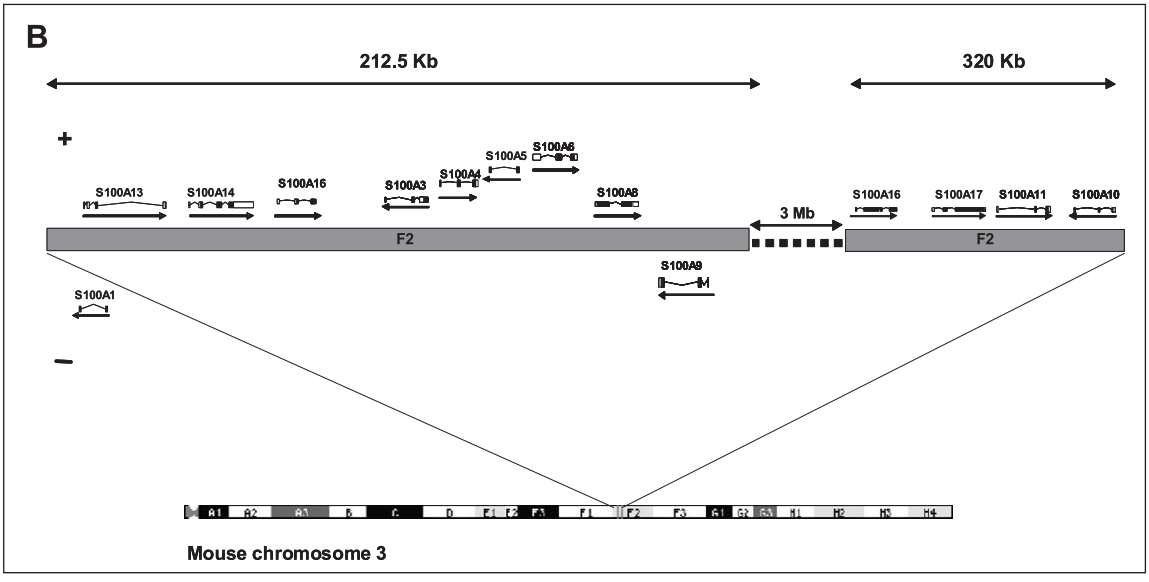
Fig. 1. Phylogeny of the EF-hand superfamily. Unrooted phylogeny among the EF-hand superfamily. The entire human, mouse, and rat protein sequences of the EF-hand superfamily (Table 1) were aligned and subjected to neighbor-joining with 1000-bootstrap analysis. The subfamilies are highlighted in different colors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

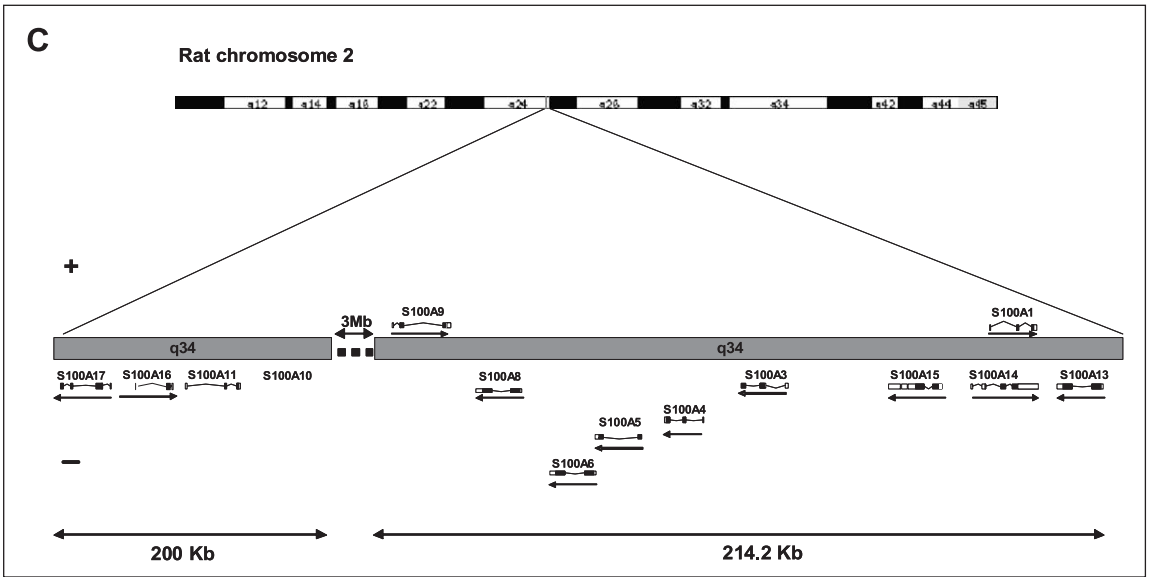
using ClustalX and then bootstrapped (1000 replicates). An N-J tree was calculated (Fig. 1). This analysis indicates that human S100A2, S100A7, S100A12, and S100P also lack obvious mouse or rat ortholog at the protein sequence level. Our analysis confirmed that no S100 proteins occur in the genomes of *Arabidopsis thaliana, Drosophila melanogaster, Caenorhabditis elegans, or Saccharomyces cerevisiae*. In evolutionary terms, the lowest organism in which an S100 protein was found is the Chondricthyes (AF421551) (C. Wang and G.V. Callard, unpublished results) *Squalus acanthias*. In this case the most closely related protein is S100A1, suggesting it to be the ancestral member of this family.

The S100 gene cluster

The 1q21.3 human S100 core cluster is composed of S100A1, S100A2, S100A3, S100A4, S100A5, S100A6, S100A9, S100A13, and S100A14 and is separated from a second subcluster by 4.2 kb containing S100A7 and isoforms S100A8, S100A9, and S100A12 in close proximity to the repetin locus. S100A10 and S100A11 lie outside these, positioned 1.5 Mb away from the core S100 cluster (Fig. 2A). Clustering is extremely well conserved in mouse, in which the core cluster containing S100A1, S100A3, S100A4, S100A5, S100A6, S100A13, and S100A14 is adjacent to a secondary cluster containing S100A8, S100A9, and


A
Human chromosome 1
q12
q41
+
1.5 Mb
S100A9
S100A1
S100A1
q34
S100A10
S100A11
S100A17
S100A16
S100A8
S100A7
S100A2
S100A13
S100A14
S100A16
S100A12
S100A3
S100A3
S100A2
S100A4
S100A5
S100A6
−
200 Kb
220 Kb


B
212.5 Kb
320 Kb
+
S100A6
S100A5
S100A4
S100A3
S100A13
S100A14
S100A16
S100A8
S100A16
S100A17
S100A11
S100A10
3 Mb
F2
F2
S100A9
S100A1
−
Mouse chromosome 3


C
Rat chromosome 2
+
3Mb
S100A9
S100A1
q34
q34
S100A17
S100A16
S100A11
S100A10
S100A8
S100A3
S100A15
S100A14
S100A13
S100A4
S100A5
S100A6
−
200 Kb
214.2 Kb

repetin. Again S100A10 and S100A11 are distant, in this case ~3 Mb from the core cluster (Fig. 2B). To determine if the chromosomal organization was conserved in the rat, syntenic with region 3f2 of the mouse (human 1q21.3), we screened the recently published ensemble of the *Rattus norvegicus* genome (Ensembl v11.2.1). The structures of individual S100 loci are conserved across species [29,49] and conservation was confirmed between mouse and rat (Figs. 2B and 2C).

The human, mouse, and rat S100A1, S100A3, S100A4, S1006, S100A8, S100A9, and S100A10 genes all contain three exons, the first of which is short and contains the 5′ untranslated sequences. The second exon contains the ATG start codon and encodes the N-terminal EF hand; the third exon encodes the carboxy-terminal EF hand [10,29]. The remaining genes contain four exons; the additional exon is located in the 5′ UTR and, like exon 1, is noncoding. In agreement with Zimmer et al. [49], we found that intron/exon lengths were also conserved between species and within the family. Interestingly, our structural predictions (using InterPro and SMART databases) of human and mouse S100A10 indicate a transmembrane segment distinct from the hydrophobic domains in helix 4 thought to be involved in target recognition and dimeric packing of the other S100 members (1 and 4). Human S100A10 is associated with the plasma membrane in basal and spinous layers of the normal epidermis and as a cytoplasmic protein [2], supporting the possibility that it may be processed at the cell surface and that there may be two functional isoforms.

Five gene copies for human S100A7 occur (S100A7a–S100A7d) on the same locus [19]. S100A7a, S100A7b, and S100A7c have clear open reading frames that reach 90% homology, whereas those of S100A7d and S100A7e are fragmented and therefore may be noncoding. The 5′ UTR and proximal promoters of the three duplicated genes S100A7a–c are divergent, suggesting that they may be differentially expressed [19]. Our analysis also indicates possible duplications of other members of the human S100 cluster (S100A1, S100A2, and S100A3), but not the corresponding mouse or rat genes (not shown). Human S100A2 is most closely related to S100A4 in all species, as indicated by phylogenetic analysis (Fig. 1). In the human S100 chromosome cluster, the S100A2 locus lies adjacent to the S100A4 locus and it is possible that S100A2 arose through gene duplication.

The S100 gene family is found only in vertebrates, and not in those below the Chondrichthynes, and the sequence divergence such as occurs in the S100A8 and S100A9 mouse and human genes underscores the rapid evolution of these genes, despite their presumed recent origin.

Proteins evolve by modular evolution, whereby conserved domains are swapped by homologous recombination of exons, and by gene duplication. Using this mechanism a protein's, family can evolve rapidly and at the molecular level it is possible to gain a protein with a new function by combination of already well-"tested" effector domains [27,34]. The structure of the classical S100 proteins is reminiscent of that of calmodulin, which contains four canonical EF hands and is the most ancient protein of the EF-hand superfamily [15,17]. S100 proteins may have evolved from a calmodulin-type ancestor by domain swapping [8,31] and subsequent loss of two of four EF hands. Although one of the S100 EF-hand motifs has low affinity for Ca^{2+}, reflecting less sequence similarity with the classical calmodulin EF hands, the evolutionary divergence of this EF hand may be relevant to the regulation of Ca^{2+} binding properties of the S100 subfamily in intracellular and extracellular environments. The duplication and divergence seen in the human genome support the proposal of rapid evolution and expansion of this family and could account for the absence of S100A2, S100A7, S100A12, and S100P in mouse and rat. A duplicated gene, such as occurs with human S100A7, may diverge in the future to create new S100 family members. Similar but distinct mechanisms of evolution can extend to the other EF-hand subfamilies. For example, the CaBP subfamily has a modular structure more similar to the family ancestor, calmodulin, although some members have different numbers of EF hands although these retain strong sequence similarities to the calmodulin EF-hand motif (Fig. 1. and Supplementary Fig. 1).

Identification of novel members of the S100 subfamily

In a search for additional S100 proteins in mouse, we screened the mouse RIKEN FANTOM2 transcriptome database (http://www.genome.gsc.riken.go.jp) for full-length proteins containing the S100/ICaBP conserved motif. To verify that these sequences were truly full length we mapped the respective cDNA sequences against the Ensembl mouse genome browser using Hashing SSAHA. We identified four candidates; BLAST analysis did not identify homologous proteins with known functional annotations. Three of the proteins were found within the 3f2 or 2q34 S100 cluster of the mouse and rat chromosome, respectively, and the 1q21.3 cluster of the human chromosome (Table 1).

Our analysis revealed that the human RefSeq NM_176823 (S100A15) sequence appears to be the same as one of the S100A7 duplications (S100A7c) described

Fig. 2. S100 chromosomal clusters. (A) Human chromosome 1q21 S100 clusters. (B) Mouse chromosome 3f3 S100 clusters. (C) Rat chromosome 2q34 S100 clusters. The figures are graphical representations of the S100 cluster organization; the loci are represented by white box, untranslated UTR, and back box, coding exon, and the lines represent introns. The arrows indicate the direction of the loci from the 5′ UTR to the 3′ UTR. The cluster structures were determined using human, mouse, and rat Ensembl genome browsers as described under Materials and methods.

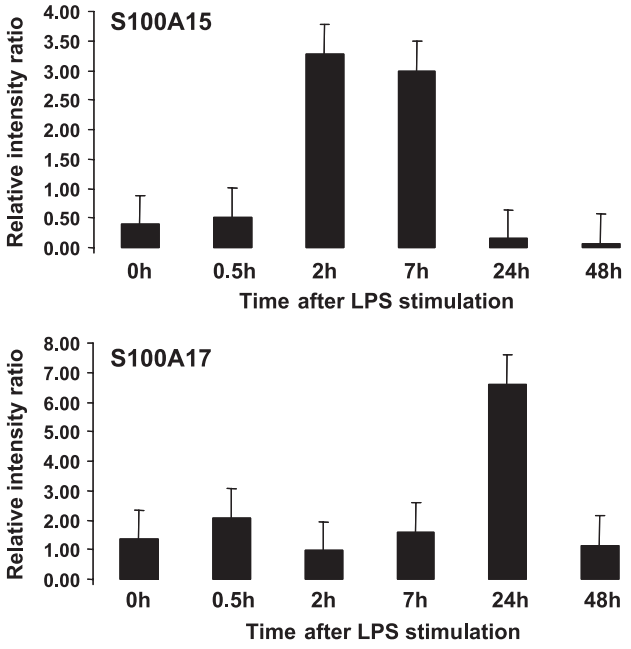


Fig. 3. Expression profiles of the newly described S100A15 and S100A17 in LPS-stimulated BMM. Microarray expression profiles of S100A15 and S100A17 shown inducibility of these two transcripts in BMM after LPS stimulation over a time series. Data points represent the means of duplicates ± SD normalized intensity values from two independent biological replicas.

previously. For this reason we have sequentially annotated the three newly described S100 genes as S100A15, S100A16, and S100A17.

Murine S100A15 (2300002L21Rik) encodes a 124-amino-acid protein of predicted mass 14.33 kDa. Expression profile analysis showed induction of S100A15 mRNA in murine bone marrow-derived macrophages (BMM) 2 h post-stimulation with lipopolysaccharide (LPS; *S. minnesota,* 10 ng/ml). Expression was maintained over 7 h but decreased to baseline thereafter (Fig. 3). S100A16 (A530063N20Rik) is a 147-amino-acid protein of predicted mass 16.85 kDa, and S100A17 (5430400H23Rik) is a 638-amino-acid protein with a predicted molecular mass of 70.76 kDa. S100A17 mRNA was induced by LPS (Fig. 3) in BMM, 24 h after stimulation.

These newly identified S100s all have N-terminal S100/ICaBP and EF-hand motifs (Supplementary Fig. 1). The murine S100A15, S100A16, and S10017 loci are close to the S100A11 locus, which is 3 Mb from the S100 core cluster. The locus structures are also conserved, with three exons, the first encoding 5′ untranslated mRNA, followed by two coding exons and two introns. The structure and physical locations of the three ortholog were maintained in the human and rat genomes (Figs. 2A, 2B, and 2C; Table 1).

Many of the early S100 members were named according to their functional associations or tissue of origin, resulting in considerable confusion in the field. This is particularly relevant to S100A4, S100A8, S100A9, and S100A12. On the basis of their chromosomal locations (Table 1) and phylogenetic (Fig. 1) and structural (Supplementary Fig. 1) analyses, a revised nomenclature for the S100 family is proposed (Table 2).

Table 2
S100 nomenclature and synonyms

Protein	Synonyms
S100A1	S100α, S100
S100A2	S100L, CaN19
S100A3	S100E
S100A4	CAPL, p9Ka, 42A, pEL98, mts1, metastasin, calvasculin, 18A2, murine placental calcium protein, fibroblast-specific protein (fsp)
S100A5	S100D
S100A6	CACY, PRA, 2A9, 5B10, calcyclin
S100A7 (isoforms a–c)	PSOR1, psoriasin, BDA111, CAAF2, S100A15 (RefSeq NM_176823: identity with S100A7c)
S100A8	MRP-8, p8, B8Ag, CP-10 (murine), calgranulin A (CAGA), cystic fibrosis antigen (CFAg), 60B8Ag, L1 Ag L1 light chain (L_{1L}), p7 (bovine A8), S100A8/S100A9 complex, calprotectin
S100A9	MRP-14, p14, B8Ag, L1Ag, calgranulin B (CAGB), 60B8Ag, L1 heavy chain (L_{1L}), BEE22 and p23 (bovine), S100A8/S100A9 complex calprotectin
S100A10	Calpactin light chain, p10, p11, 42C, CLP11, CAL12
S100A11	S100C, calgizzarin
S100A12	MRP-6, calgranulin C (CAGC), calgranulin-related protein (CGRP), CaBP in amniotic fluid-1 (CAAF1), p6, cornea-associated antigen (CoAg), extracellular newly identified RAGE-binding protein (EN-RAGE)
S100A13	
S100A14	
S100A15	(2300002L21Rik)
S100A16	(A530063N20Rik)
S100A17	(5430400H23Rik)
S100B	S100β, neural extension factor (NEF), S100
S100P	
S100Z	
Repetin	
Calbindin-D9K	CaBP9K, calbindin 3 (CALB3), ICaBP, cholecalcin

Finally, the RIKEN full-length cDNA 1700028N11Rik encodes a 158-amino-acid protein of predicted mass 18.63 kDa that contains three calmodulin-type EF hands. (Supplementary Fig. 1). Analysis indicates that this protein is more structurally homologous to hippocalcin and calmodulin (Table 1, Fig. 1) and should be included in the CaBP subfamily. 1700028N11Rik maps to mouse chromosome 15a2 and is composed of four exons and three introns.

Expression profiles of S100s

S100A8, S100A9, and S100A12 are increasingly associated with acute and chronic inflammation [44,48] and S100A8, S100A9, and S100A4 have suggested roles in myeloid cell differentiation [21,41]. They are expressed in neutrophils and monocytes. To gain additional insight into anthology relationships we performed quantitative expression analysis. Fig. 4A compares the constitutive mRNA levels in human neutrophils and peripheral blood mononuclear cells. S100A9 mRNA levels in neutrophils were some fourfold greater than those of S100A8 and both were significantly more abundant than S100A4 and S100A12 mRNAs. The broader distribution of these genes in mouse tissue was analyzed using the READ database and compared to expression of S100A1 and S100B. Fig. 5 shows similar expression profiles for S100A8 and S100A9 in all tissues. The human tissue expression profiles from the Laboratory of System Biology and Medicine Database were similar (Supplementary Fig. 2) and in both cases S100B and S100A1 exhibited more restricted expression. The extremely high levels of expression of the S100A8 and S100A9 genes in neutrophils may contribute in some cases to the expression found in those tissues containing blood cells. In comparison, S100A12 constitutive gene expression in human tissue is restricted to spleen, neutrophils, and monocytes and overall levels are significantly lower in these cells than S100A8 and S100A9 levels (Fig. 4). S100 proteins were originally thought to be expressed in a tissue-specific manner [16,20], such as that seen for S100B, S100A1, and S100A12, but other S100 genes apparently do not have the same tissue-restricted expression (Supplementary Fig. 2).

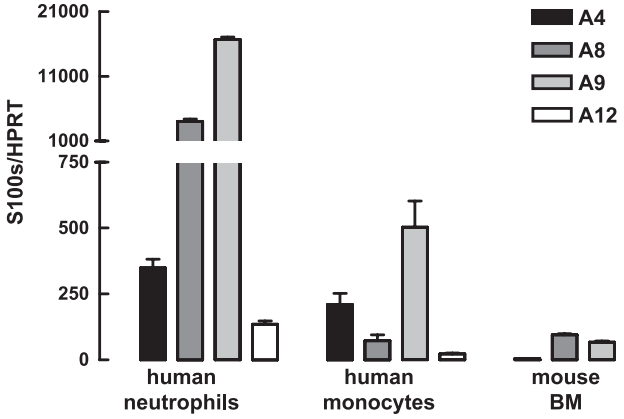


Fig. 4. S100A4, S100A8, S100A9, and S100A12 expression profiles in human monocytes and neutrophils and in murine bone marrow cells. Comparative expression levels of human S100A4, S100A8, S100A9, and S100A12 mRNA and mouse S100A4, S100A8, and S100A9 mRNA relative to HPRT mRNA in neutrophils and monocytes (human) and bone marrow (BM; mouse). Assays for the human or murine genes were performed concurrently. Data points represent the means of triplicates ± SD, and similar results were obtained in two independent experiments.

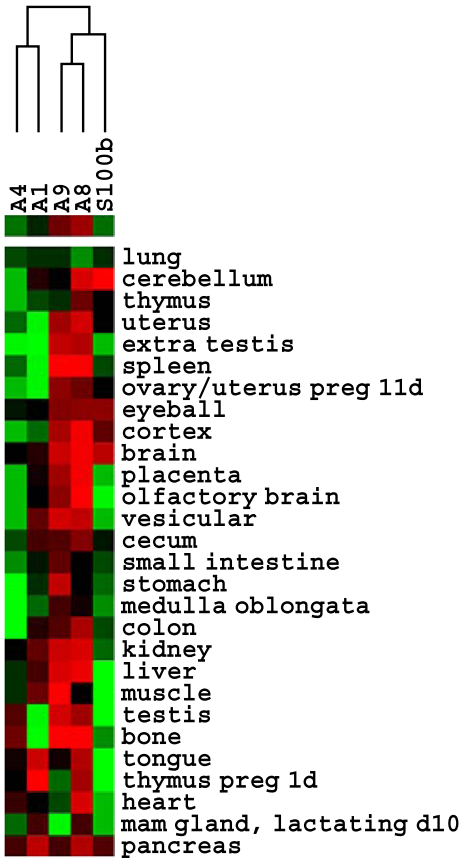


Fig. 5. Hierarchical cluster of the murine S100A1, S100A4, S100A8, S100A9, and S100 expression profiles in mouse tissues. The values correspond to the $\log_2$ signal/control ratio of the respective S100 full-length cDNAs from the READ database [24] and were used to generate hierarchical clusters. The color code is green, low expression compared to a common reference (17.5 day whole embryo RNA); black, same expression compared to the reference; red, high expression compared to the reference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Interestingly, two of the newly described S100s, S100A15 and S100A17, were up-regulated during macrophage activation by LPS in a time-dependent manner (Fig. 3), suggesting a role for these proteins in inflammation/infection, although more experiments are required to characterize these proteins functionally in this context.

Do hS100A12 and mS100A8 share common functions?

To investigate possible similarities with respect to regulation of the "inflammation-associated" S100 genes, potential regulatory elements present in human and mouse

S100A4, S100A8, and S100A9 and human S100A12 were compared. Putative proximal promoter regions were obtained and transcription factor binding site predictions and promoter alignments performed. The entire transcription factor consensus binding sites predicted for human and mouse S100A4, S100A8, and S100A9 and human S100A12 promoters are provided as a cluster analysis (Supplementary Fig. 3). All promoters contain consensus binding sites for transcription factors normally required for gene regulation in macrophages and include Sp1, Pu.1, MZF-1, AP1, and HMG-I/Y [6,11,13,35,47]. Based on this prediction, and consistent with expression profiles, the human S100A12 promoter is most similar to the murine S100A8 and S100A4 promoters, supporting the concept that mS100A8 and S100A12 may be functional chemoattractant equivalents and that S100A12 may have arisen by duplication of either human S100A8 or human S100A9. The S100A12 locus lies between S100A8 and S100A9 (Fig. 2) and the intron/exon structure more closely resembles that of S100A8 even though the amino acid sequence shares greater identity with S100A9. Amino acid sequences have diverged rapidly between species in the S100A8 and S100A9 genes (19), making it difficult to draw any conclusion as to which is the ancestral gene giving rise to S100A12.

As discussed in the introduction, we were interested in identifying the functional ortholog of mS100A8, which is a potent chemoattractant. Because of the structural divergence of S100 proteins within the hinge and C-terminal motifs, these regions are postulated to exhibit functional specificity [16,20]. The hinge region of mS100A8 is implicated in its chemotactic activity, whereas an equivalent peptide segment corresponding to the hinge sequence of hS100A8 is inactive. First, we compared structural models of the three proteins. The hS100A12 and hS100A8 templates produced successful models for mS100A8 although the total amino acid identity between hS100A8 and mS100A8 (57%) is higher than between S100A12 and mS100A8 (32%). Because of higher overall identity (45), and the large amount of surface hydrophobicity evident in the S100A12 model, the hS100A8 template was considered more accurate. However, the hinge regions of human and murine S100A8 have low homology, particularly in the C-terminal half (Fig. 6). In the hS100A8 hinge, Gly_{49} allows a level of flexibility not afforded by Asp_{49} in the mouse protein and consequently the murine hinge is unlikely to form the same kinked conformation described for hS100A8 [14]. Importantly, electrostatic potential maps indicate that the hinge regions of mS100A8 and S100A12 are structurally more similar. In contrast to

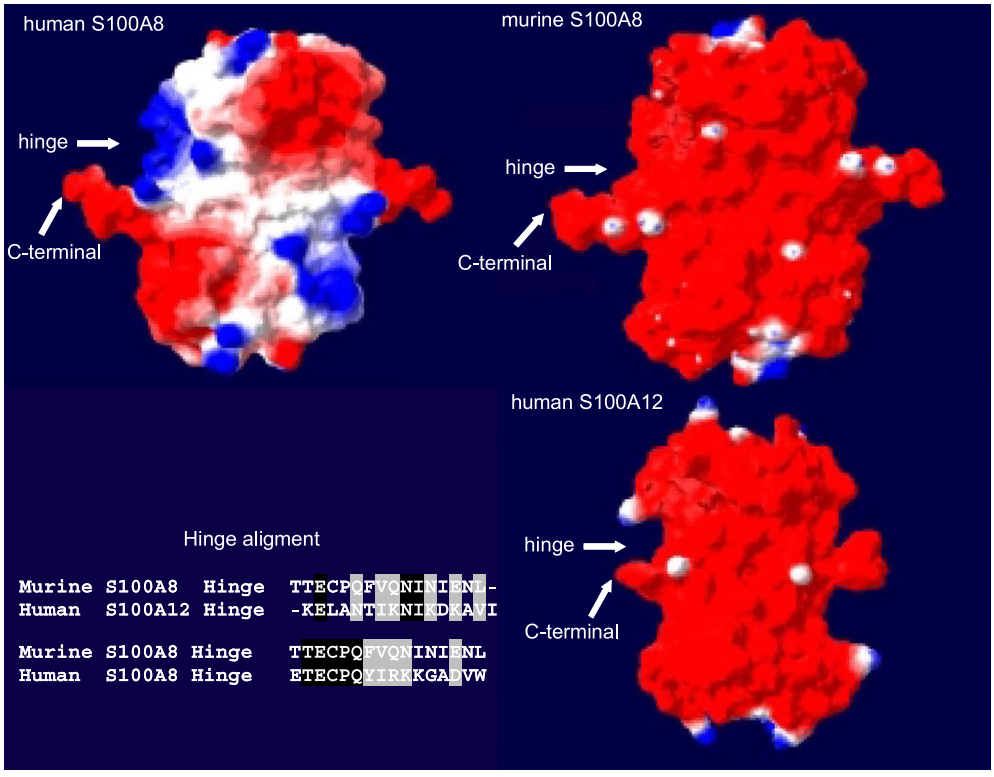


Fig. 6. Electrostatic surface potentials of human S100A8 and S100A12 and murine S100A8. Electrostatic surface potentials of hS100A8 and S100A12 and those derived from the 3D-JIGSAW mS100A8 model: blue, positive regions; white, neutral; red, negative. The hinge binding region and C-terminal are labeled. Alignments of the hinge regions (amino acids 38–53) are given. Identities are indicated as black shaded residues and similarities as gray. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the hinge regions of S100A12 and mS100A8, which are electronegative, this region in hS100A8 is positively charged (Fig. 6).

We prepared a synthetic peptide based on the S100A12 hinge sequence, $S100A12_{38-53}$ (Fig. 6), which has chemotactic activity for human THP-1 monocytoid cells (Table 3) and for human monocytes (not shown). The response provoked by 10^{-10} M was reduced by 85% when equal concentrations of the peptide were incubated on each side of the membrane and by 20% with 5×10^{11} M, indicating chemotactic rather than chemokinetic activity. The activity was similar to the effective levels of full-length S100A12 [25,48] and of $mS100A8_{42-55}$ tested using murine monocytoid cells [20]. The potency of the S100A12 hinge peptide was similar to that of the classical chemoattractant C5a (Table 3).

In addition to its chemotactic activity, mS100A8 may have a protective function in inflammatory lesions. By virtue of a highly reactive Cys and of Met residues, mS100A8 is a potent scavenger of reactive oxygen intermediates (ROI), particularly hypochlorite and peroxide [9]. This function is preserved in human S100A8. In contrast to murine and human S100A8, S100A12 has no Cys or Met residues, making it less susceptible to oxidation. In this respect, the chemotactic activity of human S100A12, in contrast to mS100A8 (in which disulfide bond formation causes loss of activity), would be resistant to covalent modification by oxidants. Interestingly, a number of classical chemoattractants are susceptible to oxidative inactivation [9], suggesting that S100A12 may be important in human diseases by propagating monocyte migration into chronic inflammatory lesions, such as in rheumatoid arthritis [48], where ROI are continuously generated. This may not be the case in lower species, however; rabbit S100A12 has a single Cys residue at position 30, although porcine and bovine S100A12 have no Cys residues. Interestingly, it is proposed that one mechanism whereby S100A12 may kill parasites is via its strong affinity for copper ions and generation of ROI and hydroxyl radicals [28]. In overview, we propose that a duplication and divergence of S100A8 in humans, compared to rodents, have permitted the separation of two functions. In mice, S100A8 is abundant and acts an antioxidant, whereas it is chemotactic within the picomolar range, but its chemotactic activity can be compromised by sensitivity to ROI [9]. In human, and in addition to the numerous functions ascribed to the S100A8/A9 complex, S100A8 may function as an antioxidant and Ca^{2+} binding protein, whereas S100A12 may chiefly act as a proinflammatory mediator by binding and activating cells expressing RAGE [12]. The absence of the S100A12 gene in laboratory rodents makes studies using these animals as model systems difficult to interpret [12,32] and makes it difficult to relate particular functions exhibited by the human proteins to the calgranulins in these species. The evolution of the S100 proteins appears to have occurred in a modular fashion, also seen in other protein families such as the C2H2-type zinc-finger and the bHLH families [27,34]. New proteins evolve not only by point mutation, but also by adding and swapping domains to already structured proteins. The S100 protein family is a good example of amplification occuring by gene duplication followed by divergence from the original duplicated gene.

Table 3
Chemotaxis of THP-1 monocytoid cells to $A12_{38-53}$[a]

Lower chamber contents	Upper chamber $A12_{38-53}$ concentration			
	0	10^{-9} M	10^{-10} M	5×10^{-11} M
	(Cell number $\times 10^2$)			
$A12_{38-53}$ 10^{-10} M	198 ± 12*	151 ± 11**	159 ± 20**	189 ± 17*
C5a, 10^{-9} M	212 ± 12*			
Medium control	152 ± 4			

THP-1 monocytoid cells were used to test the chemotactic activity of S100 $A12_{38-53}$. The peptide concentration in the lower chamber was 10^{-10} M, which was the predetermined optimal dose (activity range 10^{-9} – 10^{-11} M). To reduce the chemotactic gradient, peptide was added to the upper chambers at the concentrations given. The response was compared with that of the classical chemoattractant C5a. Results are means of six samples ± SD.
[a] Results of one experiment, representative of three experiments, are given.
* $p < 0.01$, relative to unstimulated control.
** $p < 0.01$, compared with maximal chemotaxis with 10^{-10} M $A12_{38-53}$ in lower chamber.

Materials and methods

Alignments and phylogenetic construction

Protein GenBank accession numbers used for alignments and phylogenetic trees for the S100 subfamily are listed in Table 1. ClustalX version 1.6.6 [42] was used to generate alignments and bootstrap (1000 replicates) N-J trees. TreeView software (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html) was used to obtain the N-J trees presented in Fig. 1.

Mapping of the new S100 protein genes

The genomic mapping of the human, mouse, and rat S100 genes performed in this study was carried out using SSAHA (http://www.sanger.ac.uk/Software/analysis/SSAHA/), against Ensembl of the human (version 11.31.1), mouse (version 11.3.1), and rat (version 11.2.1) genome browsers (http://www.ensembl.org/). DNA sequences used for this mapping are listed in Table 1. BLAST search was performed using the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/).

The names of the newly described S100 proteins (S100A15, S100A16, and S100A17) were proposed based on homologies of already characterized members of this family. The full-length cDNAs of the newly described

S100s are stored in the RIKEN Mouse cDNA Encyclopedia, Genome Science Laboratory, Yokohama, Japan (http://genome.gsc.riken.go.jp), under the following accession numbers: S100A15 (2300002L21Rik), S100A16 (A530063N20Rik), S100A17 (5430400H23Rik), and 1700028N11Rik.

RNA extraction and quantitative PCR

Bone marrow from femurs of adult C57BL6/J mice were collected by flushing femurs with RPMI 1640 medium (Life Technologies, Paisley, UK) containing 10% fetal calf serum, penicillin/streptomycin, and glutamine. Human peripheral blood neutrophils and mononuclear cells were purified according to established methods [48]. Total cell RNA was extracted and purified using the Qiagen RNAeasy mini-kit (Quiagen, Hilden, Germany) according to the manufacturer's instructions.

The following primers were used: human S100A4 (GB NM_019554) 5′ primer CAGCGCTTCTTCTTTCTTGG, 3′ primer CGAGTACTTGTGGAAGGTGGA; human S100A8 (GB NM_002964) 5′ primer GGGATGACCTGAAGAAATTGCTA, 3′ primer TGTTGATATCCAACTCTTTGAACCA; human S100A9 (GB NM_002965) 5′ primer GTGCGAAAAGATCTGCAAAATTT, 3′ primer GGTCCTCCATGATGTGTTCTATGA; human S100A12 (GB NM_005621) 5′ primer CTGCTTACAAAGGAGCTTGCAA, 3′ primer GGCCTTGGAATATTTCATCAATG; human HPRT (GB M26434) (used as internal standard for normalization) 5′ primer TCAGGCAGTATAATCCAAAGATGGT, 3′ primer AGTCTGGCTTATATCCAACACTTCG; mouse S100A4 (GB XM_283861) 5′ primer CAAGGAGCTACTGACCAGGG, 3′ primer GTCCCTGTTGCTGTCCAAGT; mouse S100A8 (GB NM_013650) 5′ primer CCGTCTTCAAGACATCGTTTGA, 3′ primer GTAGAGGGCATGGTGATTTCCT; mouse S100A9 (GB NM_009114) 5′ primer AGGAATTCAGACAAATGGTGGAA, 3′ primer TCCAGGTCCTCCATGATGTCA; mouse HPRT (GB J00423) (used as internal standard for normalization) 5′ primer GCAGTACAGCCCCAAAATGG, 3′ primer AACAAAGTCTGGCCTGTATCCAA.

Total RNA (3.5 μg) of each sample was treated with DNase I (Ambion, Austin, TX, USA) and reverse transcribed using a 17-mer oligo(dT) and the Superscript III RNase H$^-$ reverse transcriptase kit or Superscript III First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's instructions. Negative control samples (no first-strand synthesis) were prepared by performing reverse transcription reactions in the absence of reverse transcriptase. The reaction after reverse transcription (20 μl) was diluted to 100 μl by adding 80 μl of water. Five microliters of diluted cDNA was used for quantitative real-time PCR performed using the LightCycler–DNA Master SYBR Green I kit (Roche) following the manufacturer's instructions. The PCR was performed using an ABI Prism machine (Applied Biosystems) as follows: 1 min hot start at 94°C, followed by 45 cycles of 1 s at 94°C, 10 s at 60°C, and 15 s at 72°C. cDNA levels during the linear phase of amplification were normalized against HPRT controls. Assays were in triplicate and means ± SD determined.

Chemotactic analysis

A peptide of the hinge domain of S100A12 (sequence position 38–53) was synthesized on an Applied Biosystems Model 430A peptide synthesizer (Foster City, CA, USA), the crude peptide purified by preparative RP-HPLC, and the molecular mass confirmed by mass spectrometry. Freeze-dried samples were diluted in RPMI 1640 tissue culture medium for testing. Chemotaxis assays were as described in detail by Yang et al. [48].

Tissue array expression profiles

Murine S100 tissue microarray expression profile data were retrieved from the READ (http://www.read.gsc.riken.go.jp) as described previously [24]. The human S100 tissue microarray expression profile was retrieved from the System Biology and Medicine Database, Tokyo, Japan (http://www.lsbm.org). The hierarchical cluster presented in Fig. 4 was generated using the Cluster program from Stanford University (default set-up) and visualized using the TreeView program from Stanford University (http://genome-www5.stanford.edu/MicroArray/SMD/) from microarray data retrieved from the READ database.

S100A15 and S100A17 Expression profiles in mouse bone marrow-derived macrophage microarrays shown in Fig. 2 were obtained from experimental data generated as described in [46].

Structural comparisons of human and murine S100A8 and S100A12

For structural comparisons, the following sequences were obtained from NCBI for alignments and modeling: hS100A8 (Accession No. NP_002955), mS100A8 (NP_038678), and S100A12 (NP_005612). Homology modeling of the three-dimensional structure of mS100A8 l was generated using the 3D-JIGSAW Comparative Modeling Server (www.bmm.icnet.uk/~3djigsaw/) [1], which automatically selected human S100A8 as a template. Electrostatic potentials of the S100s were calculated with SWISS-Pdb Viewer (version 3.5, http://www.expasy.ch/swissmod/SWISS-MODEL.html) (26) using atomic partial charges and assuming vacuum.

T-COFFEE (http://www.ch.embnet.org/software/Tcoffee.html) was used for sequence alignments of hinge domains and corresponds to amino acids 38–53 of the proteins.

Promoter analysis

The putative proximal promoters of human and mouse S100A1, S100A4, S100A8, S100A9, and S100B and human S100A12 were predicted using ElDorado genome prediction software (Genomatix); a region of 600 bp (100 bp overlapping the 5′ UTR of the full-length cDNA) was retrieved. Transfac professional software (Biobase) was used to predict transcription factor binding sites within putative promoters. The core similarity was set at 0.75 and the matrix similarity at 0.85 for all our predictions.

For cluster analysis of transcription factor binding sites, matrix similarity >0.9 was chosen as an arbitrary cut-off as a measure of trust. Scores >0.9 were assigned a value of "1" if present or 0 if absent in the promoter sequences. Based on this binary code, hierarchical cluster analysis was performed using Cluster and visualized using TreeView software (Stanford University).

Acknowledgments

This work was supported by grants from the National Health and Medical Research Council of Australia (to C.L.G. and D.A.H.) and the Cooperative Research Center for Chronic Inflammatory Diseases. We thank Mr T.Tao for performing chemotaxis assays.

References

[1] P.A. Bates, M.J. Sternberg, Model building by comparison at CASP3: using expert knowledge and computer automation, Proteins Suppl. 3 (1999) 47–54.
[2] A.M. Broome, D. Ryan, R.L. Eckert, S100 protein subcellular localization during epidermal differentiation and psoriasis, J. Histochem. Cytochem. 51 (2003) 675–685.
[3] R. Donato, S-100 proteins, Cell Calcium 7 (1986) 123–145.
[4] R. Donato, S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles, Int. J. Biochem. Cell Biol. 33 (2001) 637–668.
[5] D. Engelkamp, B.W. Schafer, M.G. Mattei, P. Erne, C.W. Heizmann, Six S100 genes are clustered on human chromosome 1q21: identification of two genes coding for the two previously unreported calcium-binding proteins S100D and S100E, Proc. Natl. Acad Sci. USA 90 (1993) 6547–6551.
[6] L.F. Fowles, K.J. Stacey, D. Marks, J.A. Hamilton, D.A. Hume, Regulation of urokinase plasminogen activator gene transcription in the RAW264 murine macrophage cell line by macrophage colony-stimulating factor (CSF-1) is dependent upon the level of cell-surface receptor, Biochem. J. 347 (Pt 1) (2000) 313–320.
[7] T.O. Gabrielsen, I. Dale, P. Brandtzaeg, P.S. Hoel, M.K. Fagerhol, T.E. Larsen, P.O. Thune, Epidermal and dermal distribution of a myelomonocytic antigen (L1) shared by epithelial cells in various inflammatory skin diseases, J. Am. Acad. Dermatol. 15 (1986) 173–179.
[8] M. Hakansson, S. Linse, Protein reconstitution and reconstitution and 3D domain swapping, Curr. Protein Pept. Sci. 3 (2002) 629–642.
[9] C.A. Harrison, M.J. Raftery, J. Walsh, P. Alewood, S.E. Iismaa, S. Thliveris, C.L. Geczy, Oxidation regulates the inflammatory properties of the murine S100 protein S100A8, J. Biol. Chem. 274 (1999) 8561–8569.
[10] C.W. Heizmann, G. Fritz, B.W. Schafer, S100 proteins: structure, functions and pathology, Front (Biosci) 7 (2002) d1356–d1368.
[11] S.R. Himes, H. Tagoh, N. Goonetilleke, T. Sasmono, D. Oceandy, R. Clark, C. Bonifer, D.A. Hume, A highly conserved c-fms gene intronic element controls macrophage-specific and regulated expression, J. Leukoc. Biol. 70 (2001) 812–820.
[12] M.A. Hofmann, S. Drury, C. Fu, W. Qu, A. Taguchi, Y. Lu, C. Avila, N. Kambham, A. Bierhaus, P. Nawroth, M.F. Neurath, T. Slattery, D. Beach, J. McClary, M. Nagashima, J. Morser, D. Stern, A.M. Schmidt, RAGE mediates a novel proinflammatory axis: a central cell surface receptor for S100/calgranulin polypeptides, Cell 97 (1999) 889–901.
[13] D.A. Hume, X. Yue, I.L. Ross, P. Favot, A. Lichanska, M.C. Ostrowski, Regulation of CSF-1 receptor expression, Mol. Reprod. Dev. 46 (1997) 46–52 (discussion 52-53).
[14] K. Ishikawa, A. Nakagawa, I. Tanaka, M. Suzuki, J. Nishihira, The structure of human MRP8, a member of the S100 calcium-binding protein family, by MAD phasing at 1.9 Å resolution, Acta Crystallogr. D Biol. Crystallogr. 56 (Pt 5) (2000) 559–566.
[15] H. Kawasaki, S. Nakayama, R.H. Kretsinger, Classification and evolution of EF-hand proteins, Biometals 11 (1998) 277–295.
[16] D. Kligman, D.C. Hilt, The S100 protein family, Trends Biochem. Sci. 13 (1988) 437–443.
[17] R.H. Kretsinger, S.E. Rudnick, D.A. Sneden, V.B. Schatz, Calmodulin, S-100, and crayfish sarcoplasmic calcium-binding protein crystals suitable for X-ray diffraction studies, J. Biol. Chem. 255 (1980) 8154–8156.
[18] P. Krieg, M. Schuppler, R. Koesters, A. Mincheva, P. Lichter, F. Marks, Repetin (Rptn), a new member of the "fused gene" subgroup within the S100 gene family encoding a murine epidermal differentiation protein, Genomics 43 (1997) 339–348.
[19] J.K. Kulski, C.P. Lim, D.S. Dunn, M. Bellgard, Genomic and phylogenetic analysis of the S100A7 (Psoriasin) gene duplications within the region of the S100 gene cluster on human chromosome 1q21, J. Mol. Evol. 56 (2003) 397–406.
[20] M. Lackmann, P. Rajasekariah, S.E. Iismaa, G. Jones, C.J. Cornish, S. Hu, R.J. Simpson, R.L. Moritz, C.L. Geczy, Identification of a chemotactic domain of the pro-inflammatory S100 protein CP-10, J. Immunol. 150 (1993) 2981–2991.
[21] E. Lagasse, I.L. Weissman, Mouse MRP8 and MRP14, two intracellular calcium-binding proteins associated with the development of the myeloid lineage, Blood 79 (1992) 1907–1915.
[22] A. Mackay, C. Jones, T. Dexter, R.L. Silva, K. Bulmer, A. Jones, P. Simpson, R.A. Harris, P.S. Jat, A.M. Neville, L.F. Reis, S.R. Lakhani, M.J. O'Hare, cDNA microarray analysis of genes associated with ERBB2 (HER2/neu) overexpression in human mammary luminal epithelial cells, Oncogene 22 (2003) 2680–2688.
[23] N.S. McNutt, The S100 family of multipurpose calcium-binding proteins, J. Cutan. Pathol. 25 (1998) 521–529.
[24] R. Miki, K. Kadota, H. Bono, Y. Mizuno, Y. Tomaru, P. Carninci, M. Itoh, K. Shibata, J. Kawai, H. Konno, S. Watanabe, K. Sato, Y. Tokusumi, N. Kikuchi, Y. Ishii, Y. Hamaguchi, I. Nishizuka, H. Goto, H. Nitanda, S. Satomi, A. Yoshiki, M. Kusakabe, J.L. Eisen, M.B. Eisen, V.R. Iyer, P.O. Brown, M. Muramatsu, H. Shimada, Y. Okazaki, Y. Hayashizaki, Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays, Proc. Natl. Acad Sci. USA 98 (2001) 2199–2204.
[25] L.P. Miranda, T. Tao, A. Jones, I. Chernushevich, K.G. Standing, C.L. Geczy, P.F. Alewood, Total chemical synthesis and chemotactic activity of human S100A12 (EN-RAGE), FEBS Lett. 488 (2001) 85–90.
[26] N.D. Moncrief, R.H. Kretsinger, M. Goodman, Evolution of EF-hand

calcium-modulated proteins. I. Relationships based on amino acid sequences, J. Mol. Evol. 30 (1990) 522–562.
[27] B. Morgenstern, W.R. Atchley, Evolution of bHLH transcription factors: modular evolution by domain shuffling? Mol. Biol. Evol. 16 (1999) 1654–1663.
[28] O.V. Moroz, A.A. Antson, S.J. Grist, N.J. Maitland, G.G. Dodson, K.S. Wilson, E. Lukanidin, I.B. Bronstein, Structure of the human S100A12-copper complex: implications for host-parasite defence, Acta Crystallogr. D Biol. Crystallogr. 59 (2003) 859–867.
[29] T. Nakamura, M. Hayashi, A. Kato, T. Sawazaki, H. Yasue, T. Nakano, T. Tanaka, A unique exon-intron organization of a porcine S100C gene: close evolutionary relationship to calmodulin genes, Biochem. Biophys. Res. Commun. 243 (1998) 647–652.
[30] R.A. Newton, N. Hogg, The human S100 protein MRP-14 is a novel activator of the beta 2 integrin Mac-1 on neutrophils, J. Immunol. 160 (1998) 1427–1435.
[31] M. Ostermeier, S.J. Benkovic, Evolution of protein function by domain swapping, Adv Protein Chem. 55 (2000) 29–77.
[32] R.J. Passey, K. Xu, D.A. Hume, C.L. Geczy, S100A8: emerging functions and regulation, J. Leukocyte Biol. 66 (1999) 549–556.
[33] M.J. Raftery, Z. Yang, S.M. Valenzuela, C.L. Geczy, Novel intra- and inter-molecular sulfinamide bonds in S100A8 produced by hypochlorite oxidation, J. Biol. Chem. 276 (2001) 33393–33401.
[34] T. Ravasi, T. Huber, M. Zavolan, A. Forrest, T. Gaasterland, S. Grimmond, D.A. Hume, Systematic characterization of the zinc-finger-containing proteins in the mouse transcriptome, Genome Res. 13 (2003) 1430–1442.
[35] T. Ravasi, C. Wells, A. Forest, D.M. Underhill, B.J. Wainwright, A. Aderem, S. Grimmond, D.A. Hume, Generation of diversity in the innate immune system: macrophage heterogeneity arises from gene-autonomous transcriptional probability of individual inducible genes, J. Immunol. 168 (2002) 44–50.
[36] W. Renaud, M. Merten, C. Figarella, Increased coexpression of CFTR and S100 calcium binding proteins MRP8 and MRP14 mRNAs in cystic fibrosis human tracheal gland cells, Biochem. Biophys. Res. Commun. 201 (1994) 1518–1525.
[37] K. Ridinger, E.C. Ilg, F.K. Niggli, C.W. Heizmann, B.W. Schafer, Clustered organization of S100 genes in human and mouse, Biochim Biophys Acta 1448 (1998) 254–263.
[38] D. Rowden, G. Lovas, W. Shafer, K. Sheikh, Langerhans cells in verruciform xanthomas: an immunoperoxidase study of 10 oral cases, J. Oral Pathol. 15 (1986) 48–53.
[39] B.W. Schafer, R. Wicki, D. Engelkamp, M.G. Mattei, C.W. Heizmann, Isolation of a YAC clone covering a cluster of nine S100 genes on human chromosome 1q21: rationale for a new nomenclature of the S100 calcium-binding protein family, Genomics 25 (1995) 638–643.
[40] S. Semprini, F. Capon, A. Tacconelli, E. Giardina, A. Orecchia, R. Mingarelli, T. Gobello, G. Zambruno, A. Botta, G. Fabrizi, G. Novelli, Evidence for differential S100 gene over-expression in psoriatic patients from genetically heterogeneous pedigrees, Hum Genet 111 (2002) 310–313.
[41] K. Takenaga, Y. Nakamura, H. Endo, S. Sakiyama, Involvement of S100-related calcium-binding protein pEL98 (or mts1) in cell motility and tumor cell invasion, Jpn. J. Cancer Res. 85 (1994) 831–839.
[42] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.
[43] I.S. Thorey, J. Roth, J. Regenbogen, J.P. Halle, M. Bittner, T. Vogl, S. Kaesler, P. Bugnon, B. Reitmaier, S. Durka, A. Graf, M. Wockner, N. Rieger, A. Konstantinow, E. Wolf, A. Goppelt, S. Werner, The Ca^{2+}-binding proteins S100A8 and S100A9 are encoded by novel injury-regulated genes, J. Biol. Chem. 276 (2001) 35818–35825.
[44] P. Unger, K. Hoffman, D. Pertsemlidis, S. Thung, D. Wolfe, M. Kaneko, S100 protein-positive sustentacular cells in malignant and locally aggressive adrenal pheochromocytomas, Arch. Pathol. Lab. Med. 115 (1991) 484–487.
[45] A. Volz, B.P. Korge, J.G. Compton, A. Ziegler, P.M. Steinert, D. Mischke, Physical mapping of a functional cluster of epidermal differentiation genes on chromosome 1q21, Genomics 18 (1993) 92–99.
[46] C.A. Wells, T. Ravasi, R. Sultana, K. Yagi, P. Carninci, H. Bono, G. Faulkner, Y. Okazaki, J. Quackenbush, D.A. Hume, P.A. Lyons, Continued discovery of transcriptional units expressed in cells of the mouse mononuclear phagocyte lineage, Genome Res 13 (2003) 1360–1365.
[47] Y. Xie, C. Chen, D.A. Hume, Transcriptional regulation of c-fms gene expression, Cell. Biochem. Biophys. 34 (2001) 1–16.
[48] Z. Yang, T. Tao, M.J. Raftery, P. Youssef, N. Di Girolamo, C.L. Geczy, Proinflammatory properties of the human S100 protein S100A12, J. Leukoc. Biol. 69 (2001) 986–994.
[49] D.B. Zimmer, J. Chessher, W. Song, Nucleotide homologies in genes encoding members of the S100 protein family, Biochim Biophys Acta 1313 (1996) 229–238.