DMBT1 as an archetypal link between infection, inflammation, and cancer

J. Mollenhauer1,2, C. End1, M. Renner1, S. Lyer1,2, A. Poustka1

1Division of Molecular Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, Heidelberg, Germany.
2Department of Molecular Oncology, Institute of Medical Biology, University of Southern Denmark, Odense-C, Denmark.

DMBT1 COMO NEXO DE UNIÓN ARQUETÍPICO ENTRE INFECCIÓN, INFLAMACIÓN Y CÁNCER

Recibido: 7 Noviembre 2007
Aceptado: 21 Noviembre 2001

Epidemiological and molecular studies have pointed to links between infection, inflammation and cancer, which appear to converge at the molecular level in mechanisms associated with innate immunity. Here, the present knowledge about the secreted scavenger receptor cysteine-rich (SRCR) protein Deleted in Malignant Brain Tumors 1 (DMBT1), also known as glycoprotein-340 or salivary agglutinin, is summarized. DMBT1 is differentially expressed in various cancer types with most of these displaying a downregulation. As a lumenally secreted protein, it exerts functions in innate pathogen defense and the regulation of inflammation. By contrast, it may trigger epithelial and stem cell differentiation as an extracellular matrix protein. Its broad responsiveness to pathophysiological stimuli points to a general role in cell and tissue protection, which possibly is best circumscribed by linking pathogen defense and regulation of the inflammatory response to regenerative processes. Compelling similarities to the functions of SRCR proteins in primitive metazoa such as sponges and sea urchins exist, which support that its various functions may rely on an ancient and simple principle, i.e. the differential mediation of adhesion and anti-adhesion. Similar to NF-κB signaling pathways, which are also indirectly regulated by DMBT1, the present state of the art indicates that DMBT1 not only could exert disease-preventing, but probably also disease-promoting functions. Taken together, DMBT1 may represent a paradigm for an archetypal link between infection, inflammation, and cancer. Understanding its complex mode of action promises novel insights into the origin and the molecular basis of major human diseases.

KEY WORDS: Scavenger receptor/ DMBT1/ Inflammation/ Infection/ Cancer.
INTRODUCTION

Relationships between infection, inflammation and cancer are assumed for more than one century and have been supported by various paradigms since that time\(^1\). The consequence of Helicobacter pylori infection, acute and chronic gastritis and gastric cancer is one of the best understood links between infection, inflammation and cancer. Chronic viral hepatitis and papillomavirus infections resulting in liver and cervical cancer, respectively, are further examples that have intensely been investigated with regard to such relationships. Inflammatory bowel disease (IBD) is a chronic inflammatory disorder associated with an increased risk for colon cancer, and it is assumed that exogenous pathogenic bacteria and/or the endogenous bacterial microflora contribute to the pathogenesis of IBD.

For several reasons, innate defense mechanisms have come into focus as a link between infection, inflammation and cancer at the molecular level\(^2\). Innate immunity comprises a broad spectrum of non-adaptive (i.e. innate) mechanisms, classically considered to play a role in the first stages of defense after challenge by bacterial and viral pathogens. However, it has become evident that innate defense mechanisms also participate in the regulation of the inflammatory- and the adaptive-immune response. Dysfunction of innate immunity has been linked to chronic inflammatory disease and cancer at the genetic and the functional level. For example, mutations in the intracellular pattern recognition receptor Nucleotide-binding oligomerization domain 2 (NOD2), which is thought to sense invading bacteria, are associated with an increased risk for Crohn’s disease, one of the major subtypes of IBD\(^3-5\). Dysfunction of NOD2 has been shown to alter Nuclear factor kappa-B (NF-κB) signaling, which is a central pathway in infection, inflammation and cancer, as well as to result in decreased secretion of defensins by specialized epithelial cells in the intestine\(^6-8\).

Defensins and mucins represent epithelial-cell derived secreted factors playing a critical role as the first frontline of antimicrobial defense through direct bactericidal or trapping activities. Beyond that, however, it has become evident that some of these proteins also function in morphogenesis, the regulation of proliferation and migration of cells and/or as tumor suppressors\(^9\). Defensins are antimicrobial peptides that are produced by various cell types in the body, including epithelial cells. They play a crucial role in host defense by directly targeting and killing bacteria, viruses, and fungi. Although defensins are known to exert different functions and the purposes this processing might serve, their primary role is to provide a critical defense mechanism against infections. However, some defensins also have non-antimicrobial functions, such as signaling and tissue remodeling. Mucins, on the other hand, are glycoproteins that are secreted by epithelial cells lining the body's protective barriers and are involved in a variety of functions, including lubrication, protection, and regulation of cell signaling. Mucins can act as barriers to infection, especially against viruses, and they can also act as signaling molecules to control cell proliferation and migration. Both defensins and mucins are produced in the body as precursors that are then processed to their mature forms. The processing of defensins and mucins is an intricate and complex process that involves multiple enzymes and pathways. The mature forms of these proteins are then released into the extracellular space, where they can exert their functions. The processing of defensins and mucins is crucial for their proper function, and defects in this process can lead to disease. The impact of defensins and mucins on the development of cancer and other diseases is an area of active research, and a better understanding of their processing and function could lead to the development of new therapies for these diseases.

Here, the present knowledge about the secreted scavenger receptor cysteine-rich (SRCR) protein Deleted in Malignant Brain Tumors 1 (DMBT1) is summarized, which may represent an archetypal link between infection, inflammation and cancer and may provide important clues about how innate immunity relates to regenerative processes.

DMBT1 GENOMIC AND PROTEIN STRUCTURE

Historically, the DMBT1 gene, which locating at human chromosome 10q26.13, was recovered from a differential screen for genomic alterations in cancer by representational difference analysis, one of the earliest available genomics-based methods\(^10\). The gene spans a genomic region of about 80 kb (Fig. 1) and consists of 55 exons\(^1\). The largest transcript identified so far (DMBT1/8kb.2) consists of 7656 nucleotides (nt) containing exons 1-16 and 18-54 and giving rise to a protein with 2413 amino acids and a calculated molecular weight of 265 kDa\(^18-20\). At the amino terminus, a signal peptide is located followed by a motif with unknown function and without homologies to other proteins that is coded by 5 short repeated exons. Thirteen SRCR domains followed by short serine-threonine-proline-rich motifs that were referred to as SRCR interspersed domains (SIDs) build up a major part of the protein. The fourth SRCR domain, which is directly followed by the fifth SRCR domain without an intervening SID, represents the only exception from this uniform repetitive structure. Further, the protein contains two CUB (C1r/C1s-Uegf-Bmp1) domains flanking a fourteenth SRCR domain that has only limited homology to the other SRCR domains. At the carboxy-terminal end a zona pellucida (ZP) domain is located.

Several of these features are reflected at the level of the genomic organization of the DMBT1 gene (Fig. 1). One exon coding for an SRCR domain and two small exons coding for one SID comprise a repeating unit of about 3-4 kb in length including the intronic sequences. These repeating units share an extraordinarily high degree of sequence homology of up to more than 99% including the intronic sequences\(^18\). Of note, there is only one exon coding for the amino-terminal half of a SID between the exons for SRCR4 and SRCR5, while the exon coding for the carboxy-terminal part of the SID is missing. It is conceivable that this interferes with splicing so that also this remaining exon is skipped and SRCR4 is directly followed by SRCR5 within the protein.

Exon 55 was identified based on sequence homologies to the corresponding mouse and rat genomic and cDNA sequences. It shares high similarity with the exons in the rodent orthologs of DMBT1, which code for a transmembrane domain\(^18,21,22\). In human transcripts, this exon has not been found yet, so that it could represent a relic from evolution that is not actively used anymore. Thus, only secreted protein variants of DMBT1 are known in humans. It was proposed for both Crp-ductin (mouse Dmbt1; also known as gp300, muclin, apactin, and vomeroglandin) and ebnerin (rat dmbt1) that the transmembrane variants are proteolytically processed so that probably the part comprising at least the SRCR and CUB domains are eventually released to the extracellular space\(^22-25\). As to whether these variants exert different functions and the purposes this processing might
serve for remains to be determined, because also secretory variants are expressed, in which the transmembrane domain coding exon has not yet been identified in human transcripts. The genomic DMBT1 locus contains two repetitive regions with high homologies in both exon and intron sequences. Four short repeating units (blue arrows) are present in the 5′-region, while 12 larger repeating units (red arrows) of 3-4 kb in length are present in the SRCR domain and SID coding region. The domain organization of DMBT1 is shown below the genomic structure. The prototype represents the protein conceptually translated from exons 1-54. DMBT1/8kb.2 and DMBT1/6kb.1 represent the domain organizations of the proteins coded by the largest and the smallest transcripts identified so far. Pink triangle: signal peptide; blue box: polypeptide sequence with unknown function lacking homology to other proteins; red circles: SRCR domains; orange circles: SIDs, threonine- (TTT) and serine-threonine-proline-rich (STP) domains with some similarities to SIDs; purple boxes: CUB domains; green circle: ZP domain. The box at the bottom depicts publicly available SNP-genotyping data (http://www.ncbi.nlm.nih.gov/SNP/GeneGI.cgi?geneID=1755). The position of four SNPs in relation to the genomic sequence is indicated by broken lines. There are two haplotype blocks (red triangles) present at the DMBT1 locus, which are separated by the repetitive region that harbors the SRCR domain and SID coding exons. This means that 5′-flanking and 3′-flanking SNPs are neither linked to each other nor to genetic alterations present in the repetitive region. Thus, a genetic association of an altered copy number of the SRCR and SID exons cannot be detected by analyzing flanking SNPs or microsatellites.

The smallest human DMBT1 variant known to date (DMBT1/6kb.1) is a transcript comprising 5802 nt that codes for a protein of 1785 amino acids with an estimated molecular weight of 196 kDa. This variant lacks five of the thirteen highly homologous SRCR domains and SIDs, including SRCR4 and SRCR5, which introduce a break into the uniform SRCR-SID repetitions (Fig. 1). Various further transcripts were identified by Northern blot analysis, cDNA cloning and mapping by restriction enzymes. At present there is only evidence for differences in the number of SRCR and SID coding exons27-29. Recently a shorter SRCR gene with high homology to DMBT1 was discovered at human chromosome 7q11.23. This gene, designated as S4D-SRCRB, codes for a protein with four SRCR domains and three SIDs organized in a similar manner as in DMBT128. S4D-SRCRB shows considerable sequence homologies to DMBT1, which requires consideration, when determining expression patterns for these two genes.
GENOMIC REARRANGEMENTS VERSUS DELETION POLYMORPHISMS VERSUS ALTERNATIVE SPlicing

Based on the original observation that there exist internal deletions within DMBT1 in brain cancer, the gene was designated as Deleted in Malignant Brain Tumors 1 (17). While a precise delineation of the genomic configuration within DMBT1 remains a challenging task due to the highly repetitive structure of the gene, present evidence suggests that genomic rearrangements may comprise pre-existing deletion polymorphisms uncovered by a loss of the normal allele in cancer cells (29-32). There are indications for some tumors displaying potential de novo rearrangements also including duplications of SRCR and SID exons (30,33). Presumably, however, there exists a great variety of different DMBT1 alleles with copy number variations of these exons in the population (30). Based on these findings it is unclear to which extent – if at all – alternative splicing contributes to the diversity of DMBT1 transcript variants identified so far. Alternatively, these could represent transcripts originating from individuals with different DMBT1 genotypes.

In conclusion, the molecular basis of its variability remains uncertain, but it is probably based on genetic polymorphism, at least in part. As discussed below, there is initial evidence that a reduced number of SRCR domains and SIDs could be important in terms of susceptibility to certain diseases.

DMBT1 AND CANCER

In tumors, few point mutations were discovered within DMBT1, of which none had an unambiguous inactivating character (30-38). Thus, DMBT1 does not share the feature of biallelic inactivation by mutation found in other classical tumor suppressors. Initial studies of DMBT1 transcript and protein levels suggested a downregulation in brain, gastric, colon, esophageal, skin, oral, lung, breast and liver cancer (17,20,27,29,33-35,38-45), while pancreatic, salivary gland and prostate tumors suppressors. Initial studies of DMBT1 transcript and protein levels suggested a downregulation in brain, gastric, colon, esophageal, skin, oral, lung, breast and liver cancer (17,20,27,29,33-35,38-45), while pancreatic, salivary gland and prostate tumors also displayed an upregulation of DMBT1 (46-49). It turned out that in some of these tissues, low or no DMBT1 expression is detectable under normal physiological conditions, i.e. in the absence of pathophysiological alterations. By contrast, inflamed tissues or tumor-flanking “normal” tissues display a strong upregulation (20,27,29,40,47). This has led to the proposal that DMBT1 is induced at early stages of tumorigenesis as part of a protective response. Tumor-flanking cells and tumor cells may show increased DMBT1 levels compared to tissues without disease, because the pathophysiological condition remained unresolved. Loss of DMBT1 expression may in turn favor tumor growth or progression (50). There are initial indications supporting this dynamic model. For example, a strong upregulation of Dmbt1 is observed in the rodent mammary gland epithelium shortly after exposure to the breast cancer-inducing agent 7,12-dimethylbenz(a)anthracene (DMBA), while a downregulation takes place in the resulting mammary adenocarcinoma (35,36). An induction of rat dmbt1 takes place during alachlor-induced olfactory mucosal tumorigenesis (52).

In the 2-aminoacetylfluorene (2-AAF) and partial hepatectomy model of rat liver damage and regeneration, dmbt1 is upregulated at early stages by oval cells, which represent resident stem cells. Increased dmbt1 levels are maintained during the regenerative processes (53). Human liver tumors display elevated DMBT1 levels compared to the normal liver, but DMBT1 levels are significantly reduced during liver cancer progression (40). Kang and Reid demonstrated that DMBT1 expression is negatively regulated in gastric cancer cells through the activity of the oncogene ERK1/2, which is frequently activated in various cancer types (44). By contrast, mouse Dmbt1 was found to be strongly upregulated in prostate cancer of mice transgenic for the Neu (ERBB2) oncogene, and immunohistochemical studies also pointed to DMBT1 upregulation in human prostate cancer (49).

Taken together, these data point to a more complex role of DMBT1 in cancer, which possibly splits up into two distinct functions: a participation in general protective responses as reflected by its induction during tumorigenesis, and a possible further function in more directly cancer-related processes such as prevention of progression.

DMBT1 IN EPITHELIAL AND STEM CELL DIFFERENTIATION

Early observations demonstrated that kidney epithelial cells reverse their polarity by translocation of transmembrane ion transporters, such as KAE1, from the apical to the basal membrane, when exposed to an acidic environment (54). An extracellular matrix (ECM) protein designated as hensin (Japanese for “change in body”) was purified and shown to induce this reversal of cell polarity (55,56). Cloning of the hensin gene and homology searches identified hensin as the rabbit ortholog of DMBT1 (29,57). In vitro, polymerization of rabbit DMBT1 in the ECM triggers reversal of cell polarity and terminal epithelial differentiation processes in kidney epithelial cells, including the development of microvilli (58,59). Polymerization requires interaction with galectin-3 and probably the activity of one or more cyclophilins (60,61). It was further demonstrated that ECM-localized rabbit DMBT1 decelerates the proliferation of mouse embryonic stem cells and initiates their differentiation into monolayered epithelia (62). A role in stem cell-related processes is further supported by expression of DMBT1 in the zones of the intestinal crypts thought to harbor the stem and transit-amplifying cells as well as by its upregulation in
liver stem cells during regeneration\(^{(21,29,53,63)}\). Based on these data, counteracting tumor growth by induction of differentiation depending on localization in the ECM offers as conceivable mechanism, by which DMBT1 could act tumorsuppressive. A switch of the secretion mode from luminal to the ECM was frequently observed in tumor cells\(^{(27,33)}\), which could represent a consequence of loss of cell polarity and the formation of a more solid tumor mass that lacks the luminal context. It is imaginable that this translocation to the ECM puts selection pressure on the tumor cells resulting in DMBT1 inactivation by loss of expression.

**RELATIONS TO DEVELOPMENTAL PROCESSES AND TUMOR SUPPRESSION SUGGESTED BY MOUSE MODELS**

In accordance with a critical role in differentiation and development, Dmbt1 knockout mice, in which the first exon was substituted by a LacZ reporter gene (designated as *hensin*\(^{-/-}\) mice) displayed early embryonic lethality due to severe developmental defects\(^{(62)}\). Embryonic lethality is a phenotype frequently observed for mice with knockouts of tumorsuppressors\(^{(64)}\). On the other hand, however, other data were not in favor of such straightforward function. *Dmbt1*\(^{-/-}\) mice generated by a similar targeting strategy (i.e. targeted deletion of parts of the promoter and the first exon within the same genetic background) did not display embryonic lethality nor could an increased rate of spontaneous tumorigenesis be observed\(^{(60)}\).

Genetic screens in *Tp53*\(^{+/–}\) mice identified Dmbt1 as a strong candidate for a modifier of Tp53-deficiency induced breast carcinogenesis. Consecutive studies revealed substantially lower, i.e. 4-fold decreased *Dmbt1* expression levels in the mammary gland epithelium of breast cancer susceptible versus breast cancer resistant *Tp53*\(^{+/–}\) mice. Likewise, women with breast cancer displayed significantly reduced DMBT1 levels in the normal mammary gland epithelium compared to women without breast cancer\(^{(66)}\).

While the molecular basis of the different phenotypes observed in Dmbt1 knockout mice remains to be determined, a more sophisticated role in tumor suppression probably has to be considered, which may depend on the inactivation of other tumorsuppressors such as *Tp53* and/or on the presence of certain tumor-initiating stimuli.

**DMBT1 IN INNATE IMMUNITY**

Simultaneous to the cloning of DMBT1, Holmskov and co-workers purified glycoprotein-340 (gp-340), a surfactant protein-D binding protein representing a putative receptor for SP-D opsonized pathogens, which turned out to be identical to DMBT1\(^{(19,67)}\). Soon afterwards, mass spectrometric analysis of salivary agglutinin (SAG) demonstrated that this protein represents the salivary variant of gp-340 and DMBT1\(^{(68,69)}\). Thus, all three proteins are coded by the same gene and therefore can be considered as identical. Variations in the cDNA sequences concerned numbers of SRCR and SID exons resulting from alternative splicing or from genetic polymorphisms. To avoid confusions, it will here be referred to the respiratory variant as DMBT1\(^{G340}\) and to the salivary variant as DMBT1\(^{SAG}\), when a distinction is appropriate, and the term DMBT1 will be used when referring to no specific variant.

Salivary agglutinin was discovered in the early 80s as major non-immunoglobulin bacteria binding component in saliva\(^{(70)}\). Based on its ability to aggregate cariogenic bacteria, such as *Streptococcus mutans*, DMBT1\(^{SAG}\) was considered to function in prevention of caries through hindrance of bacterial adhesion. However, based on its ability to also mediate bacterial adhesion to surfaces\(^{(71-73)}\), DMBT1\(^{SAG}\) was suspected to play a role in caries promotion. While this discussion is still prevalent, these data suggested that DMBT1 can also directly interact with bacteria without the necessity of opsonization by other proteins, such as SP-D. Indeed, up to now DMBT1 was found to interact with a great diversity of Gram-positive and Gram-negative bacteria (Table I), including, for example, *Escherichia coli*, *Lactobacillus casei*, *Haemophilus influenzae*, *Klebsiella oxytoca*, *Helicobacter pylori*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Bacteroides fragilis*, *Salmonella* and many more\(^{(68,69,74-79)}\). Basically, this function is shared by human DMBT1 and its mouse ortholog\(^{(26)}\). DMBT1 binding to *Salmonella enterica* serovar Typhimurium and subsequent DMBT1-mediated bacterial aggregation substantially reduces bacterial invasion into intestinal epithelial cells *in vitro*\(^{(75)}\). Beyond that, DMBT1 interacts with at least two different virus types, namely HIV and influenza A viruses, which results in inhibition of viral infection *in vitro*\(^{(67,76)}\). Similar to bacteria, this includes aggregation of influenza A virus particles\(^{(77)}\), while inhibition of HIV infection probably relies on a distinct mechanism as will be discussed below.

Further indications for an important function in innate defense result from two major lines of evidence. Firstly, DMBT1 interacts also with a considerable range of other molecules playing a role in innate or adaptive defense mechanisms (Table I), which include SP-D, SP-A, MUC5B, C1q, lactoferrin, galectin-3, and IgA/sIgA\(^{(19,26,60,67-82)}\). Of note, porcine DMBT1 was also shown to interact with one of the trefoil factors, namely TFF2, which play a critical role in the regulation of wound healing\(^{(83)}\). Secondly, DMBT1 was found to be upregulated in response to bacterial and viral exposure in different organs and in different organisms, including chronic viral hepatitis in humans,
H. pylori infection in mice, as well as exposure of germ-free mice and zebrafish to the normal gut microflora

DMBT1 in inflammation and inflammatory disease

Early observations pointed to an upregulation of DMBT1 in expression in the human lung mainly caused by a gradual increase of DMBT1-GP340-positive alveolar type II cells, which correlated with the severity of inflammation (27). DMBT1 is further upregulated in the epithelial airway (non-small cell lung cancer) cell line A549 upon exposure to pro-inflammatory phorbol myristate acetate and in the rat lung upon exposure to cigarette smoke (89). It was demonstrated that DMBT1-GP340 reduces SP-D caused promotion of neutrophil oxidant response in vitro, and inhibited SP-D mediated influenza A virus uptake by neutrophils (80). Furthermore, a strong upregulation of DMBT1 expression takes place in human nasal polyposis, a chronic inflammatory disease of the sinuses (90).

Recent studies supported a role for DMBT1 in the regulation of inflammation and further pointed to an involvement in the pathogenesis of chronic inflammatory disease, i.e. Crohn’s disease. DMBT1 is specifically upregulated in the intestinal surface epithelial and Paneth cells of patients with Crohn’s disease and ulcerative colitis, and its levels correlate with the disease activity (80). Stimulation of wild type NOD2 with muramyl dipeptide results in DMBT1 upregulation and secretion via NF-κB activation, which involves NF-κB-responsive elements in the DMBT1 promoter (79). In vitro, this effect is abolished by mutations of NOD2, which frequently occur in Crohn’s disease patients. Crohn’s disease patients with risk-promoting NOD2 mutations show significantly decreased DMBT1 levels in the inflamed mucosa compared to patients with wild type NOD2. In addition, stimulation of Toll-like receptor 4 (TLR4) with its ligand lipopolysaccharide (LPS) results in DMBT1 upregulation and secretion via NF-κB activation. In turn, DMBT1 inhibits bacterial invasion into epithelial cells as well as LPS induced TLR4-mediated NF-κB activation (79). Thus, DMBT1 is able to regulate its own expression as an extracellular homeostatic element, which represents an elegant autoregulatory loop for maintaining a balanced response (Fig. 2). Hindrance of infection through pathogen aggregation and anti-inflammatory effects are a hallmark of sIgA, which is referred to as anti-inflammatory immune exclusion. Accordingly, in this particular sense DMBT1 functions very similar to mucosal antibodies by conferring anti-inflammatory immune exclusion. As discussed below, a function as a kind of “primitive antibody” would be in accordance with the evolutionary roots of DMBT1.

<table>
<thead>
<tr>
<th>Pathogen</th>
<th>Associated diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streptococcus mutans</td>
<td>Caries, endocarditis</td>
</tr>
<tr>
<td>Streptococcus gordonii</td>
<td>Caries, endocarditis</td>
</tr>
<tr>
<td>Streptococcus sobrinus</td>
<td>Caries, endocarditis</td>
</tr>
<tr>
<td>Streptococcus mitis</td>
<td>Caries, endocarditis</td>
</tr>
<tr>
<td>Streptococcus oralis</td>
<td>Endocarditis</td>
</tr>
<tr>
<td>Streptococcus intermedius</td>
<td>Endocarditis</td>
</tr>
<tr>
<td>Streptococcus anginosus</td>
<td>Endocarditis</td>
</tr>
<tr>
<td>Actinobac. actinomyc.</td>
<td>Periodontitis</td>
</tr>
<tr>
<td>Prevotella intermedia</td>
<td>Periodontitis</td>
</tr>
<tr>
<td>Peptostreptococcus micros</td>
<td>Periodontitis</td>
</tr>
<tr>
<td>Moraxella catarrhalis</td>
<td>Respiratory tract infections</td>
</tr>
<tr>
<td>Streptococcus pyogenes</td>
<td>Tonsilitis, pharyngitis, scarlet and rheumatoid fever, cellulitis, etc.</td>
</tr>
<tr>
<td>Streptococcus agalactiae</td>
<td>Meningitis, pneumoniae, wound infections, etc.</td>
</tr>
<tr>
<td>Streptococcus sanguis</td>
<td>Endocarditis</td>
</tr>
<tr>
<td>Streptococcus pneumoniae</td>
<td>Pneumonia</td>
</tr>
<tr>
<td>Klebsiella oxytoca</td>
<td>Biliary tract infections</td>
</tr>
<tr>
<td>Escherichia coli (F7)</td>
<td>Entero-/uropathogenic</td>
</tr>
<tr>
<td>Haemophilus influenza</td>
<td>Meningitis, pneumonia etc.</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>Endocarditis, wound infections, pneumonia, etc.</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>Gastritis, gastric cancer</td>
</tr>
<tr>
<td>Neisseria meningitidis</td>
<td>Septicemia, meningitis</td>
</tr>
<tr>
<td>Bacteroides fragilis</td>
<td>Diarrhoea</td>
</tr>
<tr>
<td>Lactobacillus casei</td>
<td>Normal microflora</td>
</tr>
<tr>
<td>Salmonella enterica</td>
<td>Gastroenteritis / diarrhoea</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Viruses</th>
<th>Associated diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV</td>
<td>AIDS</td>
</tr>
<tr>
<td>Influenza A viruses</td>
<td>Influenza</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Host ligand</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgA/sIgA</td>
<td>Adaptive/mucosal immunity</td>
</tr>
<tr>
<td>SP-D</td>
<td>Innate immunity</td>
</tr>
<tr>
<td>SP-A</td>
<td>Innate immunity</td>
</tr>
<tr>
<td>Lactoferrin</td>
<td>Innate immunity</td>
</tr>
<tr>
<td>MUC5B</td>
<td>Innate immunity</td>
</tr>
<tr>
<td>C1q</td>
<td>Complement cascade</td>
</tr>
<tr>
<td>TFF2</td>
<td>Wound healing</td>
</tr>
<tr>
<td>Galectin-3</td>
<td>Innate immunity, cancer, differentiation</td>
</tr>
</tbody>
</table>

Initial genetic analyses revealed a significant association of a DMBT1 deletion polymorphism with increased risk for Crohn’s disease, while no significant association could be found for ulcerative colitis (65). This deletion variant lacks 5 of the 13 amino-terminal SRCR domains and the intervening SIDs but otherwise is intact. It corresponds to the shortest transcript variant identified so far (Fig. 1), supporting the view that genetic polymorphism is one of the reasons for the apparent
INMUNOLOGÍA J. MOLLENHAUER ET AL.

Moreover, Dmbt1–/– mice displayed a significantly enhanced susceptibility to dextran sulfate sodium-triggered microflora-mediated colitis compared to Dmbt1+/+ mice (65). Taken together, these data indicated a role for DMBT1 in the pathogenesis of Crohn’s disease.

DMBT1 AND OTHER DISEASES

Dmbt1 was found to be upregulated in the pancreas and intestine of a cystic fibrosis (Cftr–/–) mouse model, in which the mice develop pancreatic and intestinal mucus plugs. Biochemical analyses identified Dmbt1 (gp300, mucin) as the major sulfated glycoprotein in these plugs (92,93). An SRCR protein designated as bovine gallbladder mucin (BGM) was shown to promote the cholesterol crystal formation in vitro (94,95). By using a recombinantly expressed part of the BGM gene, which comprised part of the SRCR coding exons, it was demonstrated that these domains are able to interact with cholesterol and other hydrophobic agents in vitro (96). Homology comparisons point to BGM as the cattle homolog of DMBT1 (20). In humans, DMBT1 was found to be upregulated in tissues with lithogenesis (40,97). These data point to a potential role in further diseases such as cystic fibrosis and gallstone formation.

Simultaneously, they indicate that in certain scenarios DMBT1 could also exert potential disease-promoting effects (Fig. 3A).

DMBT1 REPRESENTS A STEALTH GENE

Intriguingly, previous genome-wide scans utilizing microsatellites or single nucleotide polymorphisms (SNPs) had not pointed to the DMBT1 locus at 10q26.13 as a candidate susceptibility locus for Crohn’s disease (65). The relative risk conferred by the short DMBT1 allele is 1.75 (65) and thus ranges between the values arbitrarily defining high penetrance genes (relative risk >2.0) and low penetrance genes (relative risk >1.0 to 1.5). Accordingly, low penetrance is unlikely to be the reason for the DMBT1 locus slipping through the meshes of these screening methods. Genetic association studies demonstrated that while the deletion polymorphism is significantly associated with Crohn’s disease, none of the 5’- or 3’-flanking SNPs within DMBT1 displayed a significant association with this chronic inflammatory disorder. Also, flanking SNPs were neither linked to the deletion polymorphism, nor were 5’-flanking SNPs linked to 3’-flanking SNPs. Publicly available data confirm this configuration: 5’-flanking SNPs and 3’-flanking SNPs build up distinct haplotype blocks not linked to SNPs within the repetitive region containing the SRCR domain and SID coding exons (Fig. 1). The repetitive region therefore separates these two haplotype blocks, presumably, because it displays an increased recombination rate (65). As a consequence, flanking microsatellites and SNPs are not able to reveal a genetic association of copy number changes within the SRCR coding exons of DMBT1 with a disease. This represents a parallel to the HBD-2 locus at chromosome 8p23.1, which codes for human beta defensin 2 and likewise escaped genome-wide association studies with microsatellites and SNPs. This locus contains multiple repeats of the HBD-2 gene and decreased copy numbers are associated with an approximately 3-fold increased risk for Crohn’s disease (99).

In conclusion, DMBT1 can be considered as a paradigm for a stealth gene, whose association with disease can probably not be uncovered by genome-wide association studies.

FUNCTIONS OF DMBT1 DOMAINS

DMBT1 is secreted as large high molecular weight oligomers consisting of about 20 monomers, so that the total molecular weight may sum up to roughly 7000 kDa (67,100). ZP domains were demonstrated to function in protein oligomerization (101), so that it is likely that this process is mediated through the carboxy-terminal ZP domain of DMBT1. The amino-terminal part of DMBT1 comprising the polypeptide sequence up to

---

Figure 2. DMBT1 function in immune exclusion and as extracellular part of an anti-inflammatory homeostatic loop. Bacteria (blue circles) may elicit NF-κB-activation through NOD2, which senses muramylpeptide upon bacterial invasion, and through TLR4, which recognizes lipopolysaccharide of extracellular bacteria. NF-κB-activation in turn elevates DMBT1 expression and secretion. DMBT1 hinders bacterial invasion and simultaneously inhibits LPS-triggered TLR4-mediated NF-κB-activation, which would be predicted to also result in abolishing DMBT1 upregulation. Thus, DMBT1 probably acts as extracellular regulator of an autoregulatory homeostatic loop, which serves to provide a balanced response against pathogenic bacteria and/or the microflora at mucosal surfaces.
SRCR13 was demonstrated to be involved in bacterial interactions. Consecutive studies identified a minimal peptide motif of 11 amino acids (NH2-GRVEVLYRGSW-COOH), which is able to bind to the same broad spectrum of bacteria as the protein does. In addition, synthetic peptides containing this minimal bacterial binding motif (designated as DMBT1pathogen binding site 1; DMBT1pbs1) also mimicked the ability of the protein to aggregate Gram-positive and Gram-negative bacteria in vitro.102

Further evidence indicates that the SRCR domains of DMBT1 are also involved in interactions with various other ligands. Wu and co-workers demonstrated that a recombinantly expressed fragment comprising the first SRCR domain and part of the first SID interacts with HIV and is able to mimic DMBT1-mediated suppression of HIV infection in vitro.103 Ligtjenberg and colleagues and Oho and co-workers demonstrated that the bacterial recognition motif also interacts with IgA and with lactoferrin.104,105 As mentioned above, recombinantly expressed SRCR domains of the bovine homolog of DMBT1 were shown to interact with cholesterol in vitro.96

Thus, the general function of the SRCR domains of DMBT1 appears to be the mediation of multiple ligand interactions.
The functions of the two CUB domains and of the fourteenth SRCR domain, which only displays limited homology to the other SRCR domains of DMBT1, have not yet been elucidated.

**MODE OF PATHOGEN RECOGNITION**

Comprehensive efforts have been spent to unravel the mode of DMBT1-mediated pathogen recognition and interactions. Interaction with *H. pylori* was proposed to be based on bacterial binding to Lewis antigens attached to DMBT1[106], while interactions with *S. mutans* appear not to be affected by competition with Lewis groups[107]. On the other hand, both bacterial strains are bound and aggregated through the DMBT1pbs1 motif present in the SRCR domains[74,102].

To this end, *Streptococcus* antigen I/II family members such as Pas, SspB/SspA, and SpaP (also known as P1 or Pac), but also other bacterial surface proteins such as Hsa and pili of *Neisseria meningitidis* were proposed to interact with DMBT1SAG and/or DMBT1[108-118]. It was repeatedly shown that Hsa-deficient bacteria display a substantially decreased binding activity for immobilized DMBT1, while SspB/SspA-deficient bacteria displayed a moderately decreased binding activity for immobilized DMBT1. The interactions of these proteins with DMBT1 were sialidase sensitive, suggesting a role for sialic acid residues on DMBT1[77,123]. It further was observed that pathogen interaction with IAV was proposed to depend on sialic acid residues on DMBT1[77]. Moreover, it was repeatedly shown that interactions of these proteins with DMBT1 were sialidase sensitive, suggesting a role for sialic acid residues on DMBT1[77,123].

Intriguingly, also endogenous host ligands such as IgA and lactoferrin were demonstrated to interact with DMBT1[104,108]. While IgA did not compete for bacterial binding of this motif[108], lactoferrin exerted inhibitory effects on the bacterial interaction of DMBT1[108]. Synthetic peptides comprising amino acids 480-492 (NH3-SCAFDEFFSQCA-COOH) of lactoferrin were shown to compete for the interaction of DMBT1pbs1 with SpaP of *S. mutans*[106]. Moreover, SP-D was also demonstrated to compete for interactions of DMBT1 with pathogens, i. e. with influenza A viruses[77].

On HIV, a highly conserved peptide motif (NH3-CTRPNYNKKR-COOH) near the stem of the V3 loop of gp120 critical for chemokine receptor interactions on the host cells was identified as binding site for DMBT1[121,122]. By contrast, interaction with IAV was proposed to depend on sialic acid residues on DMBT1[77,123]. It further was observed that pathogen interactions might vary according to the source the protein was purified from. DMBT1SAG purified from saliva displayed stronger anti-viral activity against certain IAV types than DMBT1[27,74] purified from the lung mucus of the same donor[128]. Both anti-viral activities against IAV and bacterial binding properties were demonstrated to also vary in an interindividual manner[123,128]. Jonasson and colleagues recently demonstrated that small protein variants of DMBT1SAG in saliva correlate with caries susceptibility and mediate stronger adhesion of *S. mutans* to hydroxyapatite beads. Enhanced bacterial adhesion was shown to depend on interactions with the Streptococcus Agl/II polypeptides SpaP and Pac[128]. On the other hand, the size variants did not alter the aggregation of *S. mutans* or of Lactobacilli expressing Pac[124].

In summary, while substantial progress was achieved in understanding the pathogen interactions of DMBT1, there are still numerous enigmas. Variations of bacterial interactions and anti-viral activities between protein variants derived from lung or saliva and interindividual variability may depend on differential glycosylation, which, however, may alter interactions with some but not all pathogens. It has not yet been taken into consideration to a sufficient extent that genetic polymorphisms altering the number of pathogen-interacting SRCR domains may contribute to interindividual differences in size and pathogen interactions. Such polymorphisms would also alter the number of SIDs, which are potential targets of O-glycosylation, and therefore could have a quantitative or qualitative impact on glycosylation. To incorporate such effects and to ascertain an improved comparability of the data, it could be helpful to conduct a parallel genotyping of DMBT1 in the donors used for such experiments. Moreover, to study the influence contributed by genetic alterations and by the different pathogen interacting structures on DMBT1, recombinant protein variants expressed in standardized sources, i. e. mammalian cells, may be a valuable tool[126].

The present data further indicate that DMBT1-mediated binding, i. e. adhesion, and aggregation seem to involve different mechanisms. These possibly overlap in some regards, because aggregation requires prior interactions by binding. On the other hand, certain binding sites within DMBT1 may be differentially accessible if the protein is phase-associated, so that its binding properties could differ (Fig. 4). It is also somewhat surprising that some endogenous innate defense factors as, for example lactoferrin, appear to utilize the same binding motif within the SRCR domains that is involved in bacterial interactions. This could mean that DMBT1 serves as a kind of scaffolding protein, which organizes a larger multi-component pathogen defense complex. It is conceivable that this complex is re-arranged upon pathogen contact in a manner, which allows each of the components to find its individual
Indeed, its multi-ligand interactions offer several theoretical options for differentiated pathogen responses, which may depend on the pathogen structure, on the differential availability of host ligands and on as to whether DMBT1 is present in fluid phase or phase-associated (Fig. 4).

Finally, the present state of the art suggests that different structures on DMBT1 interact with different structures on bacteria and viruses, including protein-protein interactions between the SRCR domains of DMBT1 and pathogen surface proteins as well as interactions of DMBT1-attached glycosyl structures with pathogen receptors. A comparison of the peptide sequences within lactoferrin, SpaP, and HIV-gp120 does not reveal major homologies at the level of the primary amino acid sequence. In addition, the mode of interaction seems to vary from pathogen to pathogen. While these sophisticated interactions obviously play a role, one may expect that a kind of superimposed default pathogen recognition mechanism is of more simple nature, which raises the possibility
that the pathogen-associated pattern recognized by DMBT1 remains to be identified. As discussed below, evolutionary aspects would support this in principle. The ancient origins of DMBT1 may favor simple and generalized mechanisms.

PATHOGEN ESCAPE AND ABUSE MECHANISMS

In terms of bacteria, two possible pathogen escape mechanisms were pinpointed up to date. Hardwidge and co-workers recently found that enteropathogenic E. coli (EPEC) with type III secretion systems, which inject bacterial factors into epithelial target cells, downregulate NOD2 by a factor of about 2 and DMBT1 by a factor of about 8\(^{[127]}\). Thus, DMBT1 is directly or indirectly downregulated through type III secretion system-delivered factors, which may aid these bacteria to escape DMBT1-mediated defense. Intriguingly, one has to consider that by aiming at inactivation of the innate defense functions of DMBT1, the bacteria may simultaneously target its functions in the ECM, which would be predicted to alter processes of epithelial differentiation. It is worthwhile to take into account that this could represent a mechanism, by which bacteria could contribute to cancer without the route via chronic inflammation, similar to the hit-and-run strategy of oncogenic viruses.

Enterohemorrhagic E. coli (EHEC) with type II secretion system secrete the zinc metalloprotease SteE. Grys and co-workers identified mucin 7 and DMBT1\(^{[134]}\) as additional targets for this protease\(^{[128]}\). Because SteE also reduced the viscosity of saliva, it was proposed that EHEC might digest a gap in the protective mucous layer, which consecutively allows their adherence to target cells\(^{[128]}\). It would be of interest to map the proteolytic cleavage site(s) within DMBT1 and to determine as to whether this eliminates or alters its anti-microbial effects. Based on the determined size differences, it is conceivable that proteolytic cleavage destroys the oligomeric structure of DMBT1, potentially by releasing the part containing the aminoterminal SRCR domains from the core region. As noted above, it was shown that the amino-terminal SRCR domain of DMBT1 is sufficient to suppress HIV-infection of epithelial and peripheral blood cells \textit{in vitro}\(^{[133]}\).

The mode of inhibition of HIV-infection conferred by DMBT1 and the N-terminal SRCR domain was proposed to be based on trapping the viral particles on the cell surface by interfering with chemokine co-receptor interactions\(^{[103,121]}\). Recent data, however, suggest that this DMBT1-mediated function might rather facilitate transmission of HIV from epithelial to peripheral blood target cells as determined by co-cultivation experiments\(^{[129]}\). This could point to a scenario, in which a virus abuses the DMBT1-mediated trapping mechanism for its purposes, i.e. for facilitating its transmission to target cells (Fig. 3B).

These latter findings may in fact also relate to the discussion of as to whether DMBT1\(^\text{SAG}\) may enhance caries susceptibility by increasing bacterial adhesion as tooth surface associated protein or may decrease caries susceptibility by aggregating the bacteria in fluid phase and thereby prevent adhesion to the tooth surface. Utilization of DMBT1 by pathogens for adhesion or transmission would point to the possibility that DMBT1 could also promote certain diseases (Fig. 3). However, it can be imagined that counteracting mechanisms exist, which may depend on the interaction with endogenous host ligands (Fig. 4), so that its functions could include more than mediation of aggregation and/ or adhesion. Configurations, in which a protein can exert disease-promoting or disease-preventing effects, were documented for other scenarios. Mutations predisposing for hemoglobinopathies simultaneously protect from malaria, while CFTR mutations resulting in cystic fibrosis protect from typhoid\(^{[130]}\). NF-κB may promote or prevent liver carcinogenesis depending on the environmental trigger and the downstream pathomechanisms\(^{[131,132]}\). Thus, it should be taken into consideration that DMBT1 may neither act as a solely good nor act as a solely bad molecule. As a consequence, it would be predicted that genetic polymorphisms reducing its ligand-interacting SRCR domains might increase the susceptibility for certain diseases but decrease the risk for other diseases (Fig. 3).

EVOLUTIONARY ASPECTS POINT TO REALIZATION OF ARCHETYPAL AND SIMPLE PRINCIPLES

While its precise mode of function seems to become more complex the more details are unraveled, one can conceptually simplify DMBT1 functions to the mediation of adhesive and anti-adhesive effects, which essentially may also apply to the putative functions of DMBT1 in the ECM and downstream signaling processes. In this regard, it might be instructive to consider the functions of related proteins in primitive organisms and aspects of metazoan evolution, which point to some parallels and suggest that the DMBT1 functions could be of archaic origin.

Based on comparisons of the molecular repertoires, it was proposed that multicellular organisms of the animal kingdom (metazoa), plants and fungi emerged separately from unicellular eukaryotes during evolution. In accordance with this view, SRCR proteins are a unique hallmark of metazoa and were not found in plants or fungi\(^{[133,134]}\). Several SRCR proteins have been described in sponges, the most archaic metazoan organisms still alive. The so-called sponge aggregation receptor (AR) is a cell surface SRCR group A protein, while the multiadhesive protein of Geodia cydonium (MAP_GEOCY) represents a group B SRCR protein secreted to the ECM\(^{[134,137]}\). Group A and group B SRCR proteins differ in the number and spacing of highly
conserved cysteine residues within their SRCR domains and were reviewed elsewhere\(^\text{138}\). Three different transcripts were identified for the sponge AR, of which two code for transmembrane variants and the third one would be predicted to give rise to a secreted protein\(^\text{135,136}\). The AR recapitulates a number of properties of DMBT1. Through interaction with a sponge galectin and the so-called aggregation factor (AF), which represents a soluble multiprotein complex, the AR triggers the first step in regeneration of an intact sponge body after dissociation, i.e. re-aggregation of the dissociated sponge cells\(^\text{133,134}\). Thus, interaction with a galectin, the property to aggregate cells and a putative involvement in regenerative processes are shared by the AR and DMBT1, although they belong to different subgroups within the SRCR superfamily. The mRNA of MAP_GEOCY codes for a protein containing a fibronectin, a group B SRCR and a short consensus repeat (SCR) domain\(^\text{137}\). Based on the absence of transmembrane domain-coding stretches and on the domain composition, this protein is assumed to be secreted and to function as multidextrins in the ECM. Providing that this can be confirmed, MAP_GEOCY could represent a sponge SRCR protein that reflects certain functions of DMBT1 in the ECM.

Analyzing molecular mechanisms that are involved in the establishment and maintenance of sponge photosymbiosis, Steindler and co-workers identified two novel genes that were upregulated in the presence of endosymbiotic cyanobacteria. One of these genes, designated as PfSym2, codes for an SRCR protein\(^\text{139}\), which therefore resembles upregulation of DMBT1 in germ-free zebrafish and mice upon exposure to the normal gut microflora\(^\text{86,87}\).

Finally, comprehensive analyses in sea urchins pointed to an extremely broad repertoire of SRCR genes, which was estimated to comprise a set of about 150 genes. Among other functions, the SRCR proteins of sea urchins are thought to build up the primitive defense system\(^\text{140}\), which is reminiscent of the functions of DMBT1 in innate immunity.

Thus, it seems that many of the various functions of DMBT1 are reflected by functions of archaic SRCR proteins in primitive metazoa or vice versa. The processes, in which these ancient SRCR proteins participate, may even raise the question as to...
whether this group of molecules was actively involved in metazoan evolution. Cell-cell aggregation and cell-substratum adhesion were critical events in the formation of multicellular and sessile organisms, respectively.

In summary, there is probably much to learn about DMBT1, SRCR proteins, and general disease mechanisms from primitive metazoa. From the evolutionary perspective there is a considerable chance that DMBT1 represents an archetypal protein, which has maintained the generalism of its archaic precursors and resembles their involvement in the mediation of adhesive (and anti-adhesive) effects.

CONCLUDING REMARKS

DMBT1 is an archetypal SRCR protein involved in infection, inflammation, and cancer. As a lumenal protein, it may function in the hindrance of infection by bacterial and viral pathogens, homeostasis of the normal microflora and in the regulation of the inflammatory response. As extracellular matrix protein it may play a role in regulating aspects of epithelial and stem cell differentiation. At the present state of the art, there remain numerous pertinent issues to be solved, including questions such as to whether DMBT1 may prevent certain diseases while promoting other diseases. What is the pathogen-associated pattern recognized by DMBT1 and does this mode of pattern recognition relate to its other functions as well? How exactly is DMBT1 transducing signals into the cell when acting in the ECM, what pathways are involved and would this affect other processes except for differentiation, such as cell adhesion, migration, or proliferation of cancer cells? Does DMBT1 act as a scaffolding protein for a defense complex and may it exert similar functions in the ECM? How are its functions modulated by its various interaction partners and what is the net outcome in terms of health and disease?

While many questions could be added, one probably can conceptually simplify its mode of function to the mediation of adhesion and anti-adhesion, which is modified by the pool of available interaction partners in the respective microenvironment. As proposed earlier, it would be plausible to assume that DMBT1 links innate immune defense to the regulation of the inflammatory response and of regenerative processes, during which the compartment DMBT1 is secreted to and its available interaction partners determine the functional readout (Fig. 5). Combining innate defense with regeneration at the level of proteins with overlapping functions in both processes would represent an economical way to ascertain proper functioning of a tissue and this is reflected by the functions of archaic SRCR proteins.

Using DMBT1 as a paradigm for an archetypal link between infection, inflammation, and cancer may aid to understand fundamental mechanisms underlying the molecular basis of human diseases, some of which can possibly be reduced to problems resulting from altered adhesion and anti-adhesion. A detailed understanding of these complex adhesive and anti-adhesive processes promises to provide novel insights for the prevention and/or treatment of major diseases.

ACKNOWLEDGMENTS

This work was supported by the Health Sciences Future Award 2005 of the HGF, BMBF Functional Nutritional Research Program grant no. FKZ 0313845, and the Wilhelm Sander-Stiftung, grant no. 99.018.3.

DISCLOSURES

The authors declare no financial conflict of interest.

CORRESPONDENCE TO:

Dr. Jan Mollenhauer
Division of Molecular Genome Analysis
Deutsches Krebsforschungszentrum
Im Neuenheimer Feld 280
69120 Heidelberg, Germany
Phone: +49-6221-424746. Fax: +49-6221-423454
E-mail: j.mollenhauer@dkfz.de and jmollenhauer@health.sdu.dk

BIBLIOGRAFÍA


56. Takito J, Hikita C, Al-Awqati Q. Hensin, the polarity reversal protein, is encoded by DMBT1, a gene located in the extracellular matrix can induce apical endocytosis and reverse the polarity of intercalated cells. J Biol Chem 1999; 274: 1053-1061.


