The Sequence of the *Mycoplasma arthritidis* Superantigen, MAM: Identification of Functional Domains And Comparison with Microbial Superantigens and Plant Lectin Mitogens

By Barry C. Cole,* Kevin L. Knudtson,* Arnold Oliphant,†
Allen D. Sawitzke,* Ann Pole,* Muniraj Manohar,* L. Scott Benson,*
Elsayed Ahmed,* and Curtis L. Atkin§

From the *University of Utah School of Medicine, Department of Internal Medicine, Division of Rheumatology; †Department of Medical Informatics; and §Department of Biochemistry, Salt Lake City, Utah, 84132

Summary

*Mycoplasma arthritidis*, an agent of chronic proliferative arthritis of rodents, secretes a potent soluble superantigen, MAM, that is active for both murine and human T and B lymphocytes. We now report the complete nucleotide and amino acid sequence of MAM and show it to be distinct from other proteins and not closely related phylogenetically to other superantigens. Two functional domains on MAM are identified based on the ability of peptides encompassing these regions to inhibit lymphocyte proliferation by the intact MAM molecule. One of these domains shares short sequences or epitopes with other microbial superantigens. The second domain contains the consensus legume lectin motif-β, which is important for T cell activation by concanavalin (Con) A. MAM and Con A peptides containing this motif are functionally cross reactive, suggesting a novel secondary pathway for T cell activation by MAM.

Superantigens are potent immunomodulatory molecules that are produced by a number of pathogenic bacteria (1, 2), endogenous murine retroviruses (3), and may be associated with HIV-1 (4, 5). Superantigens have been hypothesized to play a role in human diseases including Kawasaki’s disease, toxic shock syndrome, rheumatoid arthritis, and diabetes (6–8). The *Mycoplasma arthritidis* mitogen (MAM) (9) is a particularly interesting superantigen as it is produced by an organism that causes spontaneous chronic arthritis in rodents (10). Furthermore, evidence suggests that MAM triggers autoimmune arthritis in collagen-injected mice by activating the specific T cells that drive the inflammatory response (11).

MAM is a typical superantigen in that it is presented to murine and human T cells by direct binding to MHC molecules present on accessory cell surfaces and is recognized by specific \( V_\beta \) chain segments of the TCR without MHC restriction (9). Although MAM is much less potent in activating human T cells than are the *Staphylococcus aureus*-derived superantigens, unlike the latter it promotes a strong polyclonal B cell activation in human PBL. This property suggests that MAM may be an ideal model for the study of the role of superantigens in the pathogenesis of human autoimmune disease characterized by hypergammaglobulinenia (12). We show here that MAM is a unique protein, phylogenetically distinct from other superantigens, but which contains cross-reactive motifs found in other T cell mitogens.

Materials and Methods

Cloning and Sequencing of MAM. Degenerate oligonucleotide primers were designed corresponding to the NH2- and COOH-terminal ends of the previously sequenced peptide (13), based upon the codon usage of *Mycoplasma capricolum* (14). These primers were used to amplify, by PCR, chromosomal DNA from *M. arthritidis* strain PG6. The PCR product was labeled with \(^{32}\)P and used to screen an *M. arthritidis* PG6 genomic library constructed in EMBL3 using previously described methods (15). Three clones were detected, and restriction enzyme analysis of these clones showed that *mam* was contained within a 1.4-kbp XbaI fragment that was subcloned into the plasmid pTZ18R (Pharmacia Biotech Inc., Piscataway, NJ) and sequenced by use of the Sanger dideoxy chain termination method (15).

Analysis of Physical Properties of MAM. The algorithms of Garnier et al. (16) and Chou and Fasman (17) were used to predict the secondary structure for MAM. Hopp and Wood’s (18) method predicted hydrophilicity values. Secondary structure and hydrophilicity analyses were performed using the DNASIS v. 2.0 (Hitachi Software, San Bruno, CA) sequence analysis software.

1 Abbreviations used in this paper: MAM, *Mycoplasma arthritidis* mitogen; SEA, SEB, SEC, staphylococcal enterotoxin A, B, and C, respectively.
Peptide Blocking of Lymphocyte Proliferation. Proliferation of BALB/c spleen cells was measured in a microtiter plate assay at 72 h after a 24 h pulse with [3H]TdR as detailed previously (19). For blocking, MAM peptides at 0.25–1 mg/ml were incubated with lymphocytes for 1 h at 37°C before addition of a 1:20,000 dilution (approximately ≤50 pg/ml) of pure MAM in complete RPMI 1640 medium. Due to insolubility, peptide 10 was not tested. To determine whether the inhibitory peptides were toxic, they were incubated at the above concentrations with a B cell lymphoma (2PK-3), two T cell hybridomas (one MAM reactive, Vγ8,2, and one MAM nonreactive, Vγ3) and I.29 fibroblasts in a 28-h assay using an 8-h pulse with [3H]TdR. Peptides and nucleotides were synthesized by the University of Utah DNA/Amino Acid Core Facility. Peptides based on the MAM sequence were as follows: peptide 1, amino acid 1–17; 2A, 15–31; 2B, 11–38; 3, 33–57; 4, 51–74; 5, 71–95; 6, 89–118; 8, 136–164; 9, 160–180; 10, 173–195; and 11, 189–213.

Vγ8 Usage by Native and Recombinant MAM. Vγ8 usage by MAM and other control superantigens was measured by IL-2 production by specific Vγ8 TCR-bearing T hybridomas in the presence of γ-irradiated 2PK-3 AC (both at 10⁶ cells/ml), Vγ1 (KSEA), Vγ3 (5Kc), and Vγ11 (1BvB11-17.7) were kindly provided by Dr. P. Marrack (National Jewish Center, Denver, CO), Vγ8 from Dr. B Huber (Tufts University, Boston, MA), and Vγ8.2 (2Hd-11.2) by Dr. B. Araneo (University of Utah, Salt Lake City, UT). Methods have been described (20). Staphylococcus enterotoxins A and B (SEA, SEB) were obtained from Toxin Technology, Inc. (Sarasota, FL) and Con A from Sigma Chemical Co. (St. Louis, MO).

Sequence Analysis. Superantigen amino acid sequences were obtained from the current PIR and Swiss-protein databases, from Ito et al. (21), and our sequence data for MAM. All were aligned using the ClustalV algorithm (D. Higgins, European Molecular Biology Laboratory) using default settings. The DDMATRIX application of the IntelliGenetics (IntelliGenetics, Inc., Mountain View, CA) package was used to search for regions of homology within the GCG sequence analysis software package. The best-fit unrooted dendogram was saved.

Results and Discussion

Previously, MAM was purified to homogeneity, and the sequence of the first 53 amino acids of the mature protein was obtained by Edman degradation (13). This sequence was used to design primers to generate a PCR product of the PIP. and Swiss-protein databases, from Ito et al. (21), and our sequence data for MAM. All were aligned using the ClustalV algorithm (D. Higgins, European Molecular Biology Laboratory) using default settings. The DDMATRIX application of the IntelliGenetics (IntelliGenetics, Inc., Mountain View, CA) package was used to search for regions of homology between MAM and other superantigens/mitogens using the Jimenez-Montano similarity alphabet for amino acid homology. A search for motifs in MAM was conducted using the Motifs program in the GCG sequence analysis software package (Madison, WI). Aligned sequences were examined using an heuristic search frame giving the amino acid sequence previously determined by Edman degradation. MAM begins with the methionine codon at nucleotide 127 and ends at nucleotide 766. A ρ-independent transcriptional termination region was predicted using the Termination program contained within the GCG sequence analysis software package. The three UGA universal stop codons at amino acid positions 132, 177, and 178 were read as tryptophans based on the codon usage for the Mycoplasmatales (14). After converting these stop codons to UGG tryptophan codons, the full-length, mature MAM protein encoded by nucleotides 127–766 was expressed in Escherichia coli and purified by methods to be described (Knudston, K.L., M. Manohar, and B.C. Cole, manuscript in preparation). This recombinant protein (rMAM) was biologically similar to native MAM and showed an identical Vγ8 usage that was distinct from that of other superantigens (Table 1).

MAM was predicted to be a protein of 213 amino acids with a calculated molecular weight of 25,193. This is similar to the values reported for the staphylococcal and streptococcal superantigens (6). However, the calculated pI of 10.1 makes MAM the most basic of the known bacterial superantigens. The MAM sequence does not contain cysteine residues, and so the disulfide bond, present in most

M. arthritidis
Mycoptasma Superantigen MAM: Sequence and Functional Domains

Figure 1. Nucleotide and amino acid sequence of the MAM gene (mam). The nucleotide positions are numbered on the left and the amino acid positions on the right. The double underline denotes the probable ribosomal binding site (rbs). The single arrows indicate an inverted repeat representing the predicted transcriptional termination region. The single underline indicates the complementary bases within the inverted repeat. The asterisks represent the translational stop codon for MAM. This sequence will appear in the GenBank under the accession number U33151.
other superantigens, cannot be formed. Since MAM is a secreted product (9), we sought evidence of a signal peptide. Upstream of the mature protein start site there is encoded a 25–amino acid peptide that possesses most of the characteristics that define bacterial signal peptides (22) (Fig. 1). Moreover, just further upstream is a potential ribosomal binding site, the sequence of which complements the 3'-OH end of the 16S rRNA of M. arthritidis (23).

Previously, it was reported that culture supernatants from 32 strains of M. arthritidis all possessed mitogenic activity (13). The presence of the \textit{mare} gene in these strains was confirmed by the ability of primers to the 5' and 3' ends of the gene to generate the appropriate ~650-bp product by PCR (data not shown). The presence of \textit{mare} or \textit{mam}-like genes in other \textit{Mycoplasma} species was examined by DNA–DNA hybridization using \textit{mam} as a probe. Genomic DNA from human mycoplasmas including \textit{M. pneumoniae}, \textit{M. hominis}, \textit{M. penetrans}, \textit{M. fermentans}, \textit{M. salivarium}, \textit{M. orale}, and \textit{M. buccale} failed to hybridize with the \textit{mam} probe, even under conditions of low stringency (data not shown). Thus, \textit{mam} may occur only in \textit{M. arthritidis}.

Table 1. \textit{Similar V/3 TCR Usage by Native and rMAM}

<table>
<thead>
<tr>
<th>Inducer*</th>
<th>V/31</th>
<th>V/33</th>
<th>V/3.8.2</th>
<th>V/311</th>
<th>V/36</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MAM</td>
<td>0</td>
<td>0</td>
<td>232</td>
<td>0</td>
<td>640</td>
</tr>
<tr>
<td>rMAM</td>
<td>0</td>
<td>0</td>
<td>428</td>
<td>0</td>
<td>612</td>
</tr>
<tr>
<td>SEA</td>
<td>691</td>
<td>1,883</td>
<td>0</td>
<td>1,181</td>
<td>NT*</td>
</tr>
<tr>
<td>SEB</td>
<td>0</td>
<td>1,832</td>
<td>608</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Con A</td>
<td>310</td>
<td>2,240</td>
<td>315</td>
<td>2,560</td>
<td>539</td>
</tr>
</tbody>
</table>

*Native MAM and recombinant MAM (rMAM) were tested at \(\approx 50\) ng/ml, SEA and SEB at 5 \(\mu\)g/ml, and Con A at 15 \(\mu\)g/ml.

\(\dagger\) Not tested.

MAM and its gene were found to be unique because protein and gene database searches failed to identify similar sequences. The 27% G + C content of \textit{mam} is consistent with that of most mycoplasma genomes (14), but is also similar to that of the staphylococcal and streptococcal toxin genes. Considering a common origin for these genes, we conducted a phylogenetic analysis comparing the amino acid sequences of MAM and other superantigens. MAM was not closely related to other superantigens and alone occupied one of three main branches on an unrooted phylogram (not shown). The closest similarity was between MAM and streptococcal pep M5 (24) at 24% sequence similarity.

To identify active domains on MAM, a series of overlapping MAM peptides were synthesized (Fig. 2) and tested for their ability to inhibit lymphocyte proliferation by native MAM (Fig. 3). Two candidate functional domains were identified. The first, represented in peptides 2A (MAM1_31) and 2B (MAM1_38) inhibited MAM-induced proliferation in a dose-dependent manner (Fig. 3), but had no effect on proliferation induced by Con A, indicating that MAM peptide 2 was not toxic (data not shown). The second, peptide 5 (MAM71_95) also blocked MAM activity (Fig. 3), but also inhibited proliferation induced by Con A by 50% (Fig. 4). The inhibition by MAM71_95 was not due to toxicity, as subsequent addition of Con A restored the proliferative responses to the values obtained with Con A alone (Fig. 4). In addition, neither MAM15_31 nor MAM71_95 suppressed the growth of B or T hybridomas or L929 cells (data not shown).

Although MAM was not closely related to other superantigens, searches for common sequences revealed that several superantigens shared potential epitopes with the functional MAM peptide 2 (Fig. 5A). Potential epitopes were defined to contain at least four adjacent identical or conserved amino acids followed by additional shared residues separated from the former by no more than three dissimilar residues. Two overlapping alignments of SEB with MAM1_38 are shown. The SEB region that interacts with MHC molecules comprises a hydrophobic area that includes important residues at SEB44, 45, 47 (25). This is followed by a hydrophilic or polar region containing SEB67 that is also
important for MHC binding (26) and which is situated within the shared MAM epitope. As is apparent from Fig. 2, the MAM peptide 2 has a similar profile, with a hydrophobic area followed by one of hydrophilicity, which lies in a predicted β sheet. The region of similarity between MAM and the conserved retroviral sequence of the murine mammary tumor virus 7 (MMTV-7s0) (27) is noteworthy since a similar retroviral peptide bound to MHC molecules and competitively blocked the binding of SEA to the same MHC molecules (28). Regions similar to the retroviral sequence were also present in the SEC superantigens (residues 127–139) (29) and in streptococcal pep M5 toxin (residues 101–112) (24), although there is no data to suggest that these sequences play a role in the activity of these latter superantigens. A sequence similarity to MAM was also present in the HIV-1 gp160 envelope glycoprotein, which has recently been shown to exhibit superantigen-like properties (5) and which selectively expands T cells bearing those Vβ chain segments that are used by MAM (5, 30). Regions of similarity between pep M5 and other bacterial superantigens have also been noted (24).

Whereas few sequence similarities were seen between MAM peptide 5 and other superantigens, a search for motifs revealed the striking result that MAM, but not other superantigens, contained the seven-residue lectin legume motif β consensus sequence present in all legume lectins including Con A. Shown in Fig. 5 B is the consensus legume lectin motif β, the Con A2-16 and MAM peptides, and the motif (shaded) that they contain. Peptide Con A2-16 containing this motif strongly inhibited Con A–induced lymphocyte proliferation (Fig. 4) as well as that mediated by
Published March 1, 1996

Published by Lippincott Williams & Wilkins

We thank Raymond A. Daynes for his helpful review of this manuscript, and Stacee Clayton for careful manuscript preparation.

The Utah DNA/Amino Acid Core Facility was supported in part by National Cancer Institute grant 5P30 CA42014. The work was supported by grants from the National Institutes of Health (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant AR 02255 and National Institute of Allergy and Infectious Disease grant AI 12103) and by a grant from the Nora Eccles Treadwell Foundation.

Address correspondence to Barry C. Cole, University of Utah School of Medicine, Department of Internal Medicine, Division of Rheumatology, 50 N. Medical Drive, Salt Lake City, UT 84132. A. Oliphant's present address is Department of Biology, University of Pennsylvania, Philadelphia, PA 19104-6018.

Received for publication 22 September 1995 and in revised form 3 November 1995.

References


In conclusion, MAM is not phylogenetically or structurally closely related to other superantigens on the basis of its linear sequence and a preliminary analysis of its predicted secondary structure. Thus, MAM is the first representative of a new class of superantigen. Despite these major differences, we have presented evidence that short regions of sequence homology may predict functional domains that are common to several superantigens. The domains so identified require confirmation of cross-reactive function. The presence of the legume lectin motif in MAM may indicate that this superantigen alone exhibits a unique interaction site that contributes to lymphocyte activation. Analysis of the x-ray crystallographic structure of MAM and of a library of deletion mutants are now required to more precisely define the molecular interactions that lead to T and B cell superantigenic activity.


