Crystal Structure of the Cysteine-rich Domain of Scavenger Receptor MARCO Reveals the Presence of a Basic and an Acidic Cluster That Both Contribute to Ligand Recognition

Juha R. M. Ojala, Timo Pikkarainen, Ari Tuuttila, Tatyana Sandalova, and Karl Tryggvason

Division of Matrix Biology and Division of Molecular Structural Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden

MARC0 is a trimeric class A scavenger receptor of macrophages and dendritic cells that recognizes polyanionic particles and pathogens. The distal, scavenger receptor cysteine-rich (SRCR) domain of the extracellular part of this receptor has been implicated in ligand binding. To provide a structural basis for understanding the ligand-binding mechanisms of MARCO, we have determined the crystal structure of the mouse MARCO SRCR domain. The recombinant SRCR domain purified as monomeric and dimeric forms, and their structures were determined at 1.78 and 1.77 Å resolution, respectively. The monomer has a compact globular fold with a twisted five-stranded antiparallel β-sheet and a long loop covering a single α-helix, whereas the dimer is formed via β-strand swapping of two monomers, thus containing a large eight-stranded β-sheet. Calculation of the surface electrostatic potential revealed that the β-sheet region with several arginines forms a basic cluster. Unexpectedly, an acidic cluster was found in the long loop region. In the monomer, the acidic cluster is involved in metal ion binding. Studies with cells expressing various SRCR domain mutants showed that all of the arginines of the basic cluster are involved in ligand binding, suggesting a cooperative binding mechanism. Ligand binding is also dependent on the acidic cluster and Ca2+ ions whose depletion appears to affect ligand binding at least by modulating the electrostatic potential or relative domain orientation. We propose that the SRCR domain dimerization can contribute to the recognition of large ligands by providing a means for the MARCO receptor oligomerization.

MARCO belongs to a diverse group of scavenger receptors (SRs)3 expressed by macrophages, dendritic cells, and certain endothelial cells (1). These germ line-encoded receptors, also known as pattern recognition receptors due to their ability to recognize conserved pathogen-associated molecular patterns, are considered as an important part of innate immunity, the evolutionarily conserved, first line host defense mechanism. In addition to pathogen-associated molecular patterns, a long list of SR ligands, often polyanionic in nature, includes pollution particles, polyanucleotides, bacterial lipopolysaccharides, modified host molecules such as oxidized low density lipoprotein (LDL), and unmodified endogenous proteins (1, 2). The SRs are classified into several subgroups, of which class A SRs have primarily been associated with innate immunity. This class consists of five members: SR-A (SR-AI, -II, and -III/SCARA1) (3, 4), MARCO (macrophage receptor with collagenous domain)/SCARA2 (5), CSR1 (cellular stress response 1) and II/SCARA4 (7, 8), and Tsr (testis-expressed scavenger receptor)/SCARA5 (class A scavenger receptor 5) (9, 10). All of these are trimeric type II membrane proteins with a similar predicted tertiary structure consisting of a short intracellular domain, a transmembrane domain, and a large extracellular domain with an α-helical coiled-coil domain, a triple-helical collagenous domain, and a C-terminal cysteine-rich domain. MARCO differs from the others in that it has a long triple-helical collagenous domain but no α-helical coiled-coil domain. In addition, the C-terminal cysteine-rich domain can be either a C-type lectin-like domain, as it is in SRCL, or a scavenger receptor cysteine-rich (SRCR) domain, present in SR-AI, MARCO, and SCARA5. The splice variants, SR-AII/III, SRCL II, and CSR, are missing the cysteine-rich domain (3–10).

The SRCR domain is an ancient, highly conserved fold consisting of ~110 residues. It has been divided into two subclasses (A and B), depending on the number and position of cysteine residues (6–8), as well as based on the exon-intron organization (3, 11, 12). However, the structure of human hepsin reveals a third, atypical cysteine residue pattern for the SRCR domain.

* This study was supported in part by research grants from the Swedish Research Council (to K. T. and T. P.) and the Sigrid Juselius Foundation (to A. T.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2OY3 and 2OYA) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S8.

1 Present address: Institute of Biotechnology, Electron Microscopy Unit, University of Helsinki, P.O. Box 56, FIN-00014 Helsinki, Finland.

2 To whom correspondence should be addressed: Division of Matrix Biology, Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet, Scheele vág 2 B1, S-17177 Stockholm, Sweden. Tel.: 46-8-524-87720; Fax: 46-8-316165; E-mail: karl.tryggvason@ki.se.

3 The abbreviations used are: SR, scavenger receptor; SRCL, scavenger receptor cysteine-rich domain; SRCL, scavenger receptor with C-type lectin; LDL, low density lipoprotein; AcLDL, acetylated low density lipoprotein; DSP, diethylpolysorbate; CHO, Chinese hamster ovary; poly I, polyinosinic acid; PBS, phosphate-buffered saline; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PEG, polyethylene glycol; MME, monomethyl ether; CMV, cytomegalovirus; r.m.s., root mean square.
(class C) (13). The domain is found either as a single copy or as tandem repeats in several soluble or membrane-bound proteins often associated with the innate immune system. Strikingly, the genome of the purple sea urchin, Strongylocentrotus purpuratus, contains 1,095 tentative SRCR domain copies distributed to 218 genes. In comparison, only 81 copies in 16 genes are found in humans (14). The extensiveness of the SRCR gene family in the sea urchin genome is intriguing and indicates a major need for SRCR domains in the host defense of lower organisms, which solely rely on innate immunity in the absence of adaptive immunity.

Despite their large number in the proteome, only a few SRCR domains have a known function. The lymphocyte cell surface receptor CD6 interacts with activated leucocyte cell adhesion molecule (ALCAM/CD166) via the SRCR3 domain (15); in gp-340/DMBT1 (deleted in malignant tumors 1), the peptide QGRVEVLYRGSWGTVC, present in eight of its 14 SRCR domains, binds and agglutinates Streptococcus mutans and various other bacterial strains (16, 17), and the first of the three basic cluster composed of several arginines and an acidic one (1.77 Å resolution, respectively. Calculation of the electrostatic potential shows that the domain contains two clusters, a basic cluster composed of several arginines and an acidic one with a bound metal ion. Evidence is provided indicating that both clusters are involved in ligand binding.

**EXPERIMENTAL PROCEDURES**

**Production and Purification of the MARCO SRCR Domain**—A previously generated human embryonic kidney epithelial 293/EBNA cell line (Invitrogen) stably expressing a soluble recombinant MARCO SRCR domain starting from Glu321 was used for protein production (20). The recombinant protein contains an N-terminal extension Ala-Pro-Leu-Ala (APLA), left from the BM-40 protein, the signal sequence of which was used to facilitate secretion of the protein. For protein collection, the cells were grown in serum-free Dulbecco’s modified Eagle’s medium/F-12 (without phenol red) for a maximum of 12 days or until the cells started to detach.

The cleared, protease inhibitor-supplemented (1 mM phenylmethylsulfonyl fluoride, 5 mM EDTA) conditioned medium was first dialyzed at +4 °C against 20 mM sodium acetate buffer, pH 4.7. After three buffer changes, the medium was cleared by centrifugation (10,000 × g, 20 min) and applied at +4 °C to a Fractogel EMD SO₃⁻ packed cation exchange column (Merck). Bound proteins were eluted with a linear NaCl gradient (0–800 mM), and fractions containing the recombinant protein were dialyzed against 20 mM Tris-HCl, pH 9.0, 5 mM NaCl, before injection to a Mono Q anion exchange column (GE Healthcare). Elution with a linear NaCl gradient (5–500 mM) yielded two main peaks, both containing the MARCO SRCR domain as assessed by SDS-PAGE and Western blotting. N-terminal sequencing with Edman degradation also confirmed the expected APLAQ sequence. Next, the two peaks were separately pooled, concentrated by ultrafiltration, and run on a calibrated Superdex 75 size exclusion chromatography column (GE Healthcare) with 20 mM Tris-HCl, pH 7.5, 150 mM NaCl (0.5 ml/min). The protein peaks were analyzed for homogeneity by mass spectrometry on a QTOF 1 API instrument (Waters Corp., Milford, MA) equipped with the standard Z-spray source. Prior to the electrospray, the samples were desalted with ZipTip U-C18 (Millipore) and resolved in 60% acetonitrile, 1% acetic acid.

**Crystal Structure of the MARCO SRCR Domain**

Chemical Cross-linking—The two MARCO SRCR domain pools from the size exclusion chromatography were chemically cross-linked with dithiobis(succinimidyl propionate) (DSP; Pierce). In this experiment, 125 μg/ml protein solutions in PBS were prepared and mixed with equal volumes of the DSP solution (1 mM DSP in PBS containing 10% Me₃SO). After incubation for 1 h at room temperature, the reactions were stopped by adding Tris-HCl, pH 7.5, to a final concentration of 50 mM, and the samples were analyzed by SDS-PAGE and silver staining under nonreducing or reducing conditions. DSP contains a cleavable disulfide bond; therefore, the cross-links can be reversed by reduction.

Crystallization and Data Collection—For crystallization trials by the hanging drop vapor diffusion method at +4 °C, the two forms of the MARCO SRCR domain were concentrated in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl. The monomeric form (14 mg/ml) was crystallized in 12% PEG 8000, 0.1 M BisTris, pH 6.0, 0.2 M CH₃COONH₄ by macroseeding from 20% PEG 8000, 0.1 M cacodylate, pH 6.5, 0.2 M magnesium acetate, whereas the dimeric form (7 mg/ml) crystallized in 7% PEG 4000, 0.1 M BisTris buffer, pH 6.0, 0.1 M Li₂SO₄. One single crystal of the monomeric SRCR domain was cryoprotected with 40% PEG 550 monomethyl ether (MME) and frozen in a nitrogen stream generated by a cryosystem (Xstream; Molecular Structure Corp., The Woodlands, TX). CuKα x-rays (1.5418 Å) from a Rigaku rotating anode generator and a MAR image plate system were used to collect a data set of 159 images with a 1.5° oscillation angle. The data were indexed, integrated, and scaled using the HKL software package (23). In case of the dimeric MARCO SRCR domain, the x-ray diffraction data were collected from a single crystal at beamline I711 in Lund (MaxLab) using a MARCCD detector. As for the monomeric SRCR domain crystals, a short soak in 40% PEG 550 MME was used for cryoprotection. The diffraction data of 321 images, collected with 1° of oscillation per image, were integrated with MOSFLM (24) and scaled with programs of the CCP4 suite (25).

Structure Determination and Refinement—Both structures were determined by molecular replacement. For the monomeric form of the MARCO SRCR domain, the first structure...
solution was obtained with AMORE (26) using the structure of the M2BP (Mac-2-binding protein) SRCR domain (Protein Data Bank code 1BY2) as a search model. For the dimeric form of the MARCO SRCR domain, the structure of the monomeric MARCO SRCR domain was used as a search model, and MOLREP found the first structure solution in two steps (29). The models were built iteratively with O and COOT (30, 31) and refined with REFMAC5 (32). At the final stage of refinement, solvent molecules were added to well defined peaks (over 3.0 σ) with ARP/warp (33) and checked by COOT (31). The final models were checked for errors by calculating simulated annealing omit maps in CNS (34) and by evaluation of the geometry with PROCHECK and COOT (31, 35).

SRCR Domain Modeling—The SRCR domain structures of mouse SR-A, human MARCO, and the mouse/human MARCO chimeras (M496–518, M449–473, and M453–518) were built iteratively in COOT by mutating the mouse MARCO and human M2BP SRCR domain structures.

Cell Assays—The ligand-binding activity of CHO cells transfected with expression plasmids encoding various MARCO mutants was compared with that of cells transfected with a negative control plasmid (included vectors for an immunoglobulin superfamily member nephrin and a type II membrane protein asialoglycoprotein receptor H1 subunit) and wild-type MARCO expressed either from a standard CMV promoter/ enhancer-containing plasmid (high MARCO, or H-MARCO) or from a plasmid lacking the CMV enhancer (low MARCO, or L-MARCO). This latter plasmid gives lower expression levels than the standard plasmid and was used because some combinations of the Arg-to-Ala substitutions affected the cell surface expression levels of MARCO (some mutants were expressed at lower levels than wild-type MARCO, and they were excluded from the analysis).

For the cell adhesion assay, cells transfected on 100-mm dishes with the calcium phosphate method were prepared as follows. At 24 h post-transfection, cells were trypsinized and seeded on 60-mm dishes in complete growth medium containing ascorbic acid (100 μg/ml). The cells were cultured overnight and then detached by a brief trypsin/EDTA-treatment at 40–42 h post-transfection. After inactivation of trypsin with complete growth medium, the cells were centrifuged and washed once with Dulbecco’s modified Eagle’s medium, 20 mM HEPES, pH 7.5, 0.2% bovine serum albumin for 10 min, followed by a 45-min incubation with anti-MARCO antibodies diluted in the same solution. After several washes with PBS, the cells were fixed and incubated with Alexa Fluor 488-conjugated secondary antibodies. Coverslips were mounted using DAKO fluorescent mounting medium.

Similar adhesion assays were also performed on polyinosinic acid (poly I)-coated coverslips. For the coating, a 200 μg/ml solution (Sigma P-4154; prepared in diethyl pyrocarbonate-treated H2O) was incubated on coverslips overnight at room temperature. Before plating the cells, the coverslips were treated as above.

When testing the dependence of MARCO-mediated cell adhesion to gelatin or poly I on divalent cations, cells were rapidly washed after trypsin inactivation with 2 mM EDTA in Hanks’ balanced salt solution (without Ca2+ and Mg2+), containing 20 mM HEPES, 2 mg/ml glucose, and 0.2% bovine serum albumin (divalent cation-free plating solution). Thereafter, the cells were washed three times with the same solution lacking EDTA, resuspended into the same solution, and allowed to recover before plating on coverslips. Immediately prior to plating, cell suspension aliquots were supplemented with 2 mM CaCl2 (final concentration), 1 mM MgCl2, and both CaCl2 and MgCl2, or were left without any extra cations. Cells were plated for 1 h and then processed as described above.

The binding of AcLDL was tested as described previously (19). For these assays, transfectants were seeded on coverslips 24 h post-transfection. The binding assays were performed in duplicates 45–48 h post-transfection. A third set of cell platings was used for monitoring the cell surface levels of the MARCO proteins by incubating with the anti-MARCO antibodies as described above. In most of the antibody incubation assays, we used polyclonal anti-MARCO antibodies raised against the extracellular part of mouse MARCO (the end of the collage- nous domain and the SRCR domain) (5). Alternatively, one of the following monoclonal antibodies was used: ED31 (37) and IBL-12 (38), rat monoclonal antibodies recognizing the SRCR domain of mouse MARCO, or PLK-1 (39), a mouse monoclonal antibody recognizing the human MARCO SRCR domain.

RESULTS AND DISCUSSION

Purification and Crystal Structure of Monomeric and Dimeric Forms of the Recombinant MARCO SRCR Domain—Purification of the recombinant mouse MARCO SRCR domain by conventional ion exchange and size exclusion chromatography yielded two forms with the same N-terminal sequence and mass value (11,320 Da). Their elution positions from the calibrated size exclusion column suggested the presence of both monomers and dimers. This was verified by chemical cross-linking with DSP (data not shown). The purified dimers were very stable in physiological buffers but disintegrated into monomers during the desalting step prior to mass analysis and when incubated in hydrophobic solutions (not shown).

The x-ray diffraction data for the monomeric form of the SRCR domain were collected with a laboratory source to a resolution of 1.78 Å, whereas the data for the dimeric form were
collected at the synchrotron to a resolution of 1.77 Å. Both crystal structures were solved by molecular replacement. After the final refinement, the \( R_{\text{work}} \) and \( R_{\text{free}} \) factors for the monomer are 16 and 21%, respectively, whereas the values for the dimer are 15 and 19%. Both models have good stereochemistry. The monomer contains 67 water molecules and two ions modeled as Mg\(^{2+}\), since this ion was present in the macroseeding solution. The dimer has three bound sulfate ions that originate from the crystallization media and 192 water molecules (Table 1 and Fig. S1). The electron density for the APLA sequence and the side chain of Arg422 could not be seen in the monomeric structure. Analysis of the contacts at the PISA interface (1403 Å\(^2\)) revealed that the monomers do not form parallel fold (30 Å in diameter) with N-terminal twisted antiparallel \( \beta \)-strands A–C, followed by a single \( \alpha \)-helix (I) and a \( \beta \)-strand D (Fig. 1, A and B). The polypeptide chain continues with a short \( \beta \)-strand E and a long loop containing two \( 3_{10} \)-helices (II and III) considered as loops. Finally the polypeptide chain ends with a C-terminal \( \beta \)-strand F, which lies antiparallel between \( \beta \)-strands D and B. Altogether, the \( \beta \)-strands A–D and F form a twisted five-stranded antiparallel \( \beta \)-sheet wrapped around one side of the \( \alpha \)-helix, whereas the long loop region is arching over to the other side. This organization is held together by extensive hydrophobic contacts and stabilized by three disulfide bonds: Cys\(^{446}\)–Cys\(^{507}\) connects the long loop region to the end of the \( \beta \)-strand C, a 10-residue long turn in the loop region is stabilized by Cys\(^{487}\)–Cys\(^{497}\), and Cys\(^{459}\)–Cys\(^{517}\) connects the C-terminal \( \beta \)-strand F to the \( \alpha \)-helix.

The dimeric structure of the MARCO SRCR domain is in principle a \( \beta \)-strand-swapped form of two monomers. Strand swapping is a form of domain swapping, a mechanism first described at the atomic level for diphtheria toxin; the topology of the domains remains the same, but some parts are swapped (41). Investigations on diphtheria toxin initially led to the proposal that dimerization by domain swapping requires a decrease in pH that converts the monomeric toxin to an open form (41). Later, however, a domain-swapped dimer of the protein was also found to be formed by ligand binding at a neutral pH (50). In the case of the MARCO SRCR domain, it is true that the protein is exposed to an acidic pH during the purification procedure for crystallization, but we have also been able to detect dimers when the protein is purified under neutral conditions (not shown). In the dimeric structure of MARCO, the A \( \beta \)-strands from two monomers have changed places with each other. The hinge movement of the four-residue A–B loop (Gly\(^{428}\)–Arg\(^{431}\)) has resulted in the formation of a long AB \( \beta \)-strand, which interacts with the corresponding segment of another monomer, thereby restoring the topology and the A–B \( \beta \)-strand interactions seen in the monomeric SRCR fold. Notably, the AB-D and F \( \beta \)-strands of the monomers are brought together, resulting in the formation of a large eight-stranded antiparallel \( \beta \)-sheet with a surface area of ~26 × 15 Å (Fig. 1C).

<table>
<thead>
<tr>
<th>Data collection and refinement statistics</th>
<th>Data set</th>
<th>Monomer</th>
<th>Dimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength</td>
<td>CuK(_x), 1.5418 Å</td>
<td>1.09700 Å</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P12,1</td>
<td>P1</td>
<td></td>
</tr>
<tr>
<td>Unit cell (Å)</td>
<td>( a = 32.3, b = 28.1, c = 42.7 )</td>
<td>( a = 31.1, b = 36.6, c = 44.6 )</td>
<td></td>
</tr>
<tr>
<td>Unit cell (degrees)</td>
<td>( \alpha = \gamma = 90, \beta = 99.2 )</td>
<td>( \alpha = 116.3, \beta = 90.4, \gamma = 96.4 )</td>
<td></td>
</tr>
<tr>
<td>Data range (Å)</td>
<td>23.7–1.78</td>
<td>14.91–1.77</td>
<td></td>
</tr>
<tr>
<td>Measured reflections</td>
<td>56,258</td>
<td>104,490</td>
<td></td>
</tr>
<tr>
<td>Unique reflections</td>
<td>7155</td>
<td>16,364</td>
<td></td>
</tr>
<tr>
<td>( R_{\text{merge}} ) (p.e.)</td>
<td>4.8 (37.5) (1.81–1.78 Å)(^a)</td>
<td>3.7 (7.4) (1.86–1.77 Å)(^a)</td>
<td></td>
</tr>
<tr>
<td>Data completeness (%)</td>
<td>96.2 (78.8)</td>
<td>95.4 (89.4)</td>
<td></td>
</tr>
<tr>
<td>( \langle l \rangle / \langle \sigma l \rangle )</td>
<td>32.42 (3.17)</td>
<td>13.8 (9.1)</td>
<td></td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>23.7–1.78</td>
<td>39.8–1.77</td>
<td></td>
</tr>
<tr>
<td>No. of reflection working set</td>
<td>6252</td>
<td>14,720</td>
<td></td>
</tr>
<tr>
<td>No. of reflection test set</td>
<td>903</td>
<td>1644</td>
<td></td>
</tr>
<tr>
<td>( R_{\text{work}} ) (%)</td>
<td>16.1</td>
<td>15.0</td>
<td></td>
</tr>
<tr>
<td>( R_{\text{free}} ) (%)</td>
<td>21.0</td>
<td>19.1</td>
<td></td>
</tr>
<tr>
<td>r.m.s. deviation in bond lengths (Å)</td>
<td>0.013</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>r.m.s. deviation in bond angles (degrees)</td>
<td>1.35</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>Mean B value (Å(^2))</td>
<td>19.7</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>67</td>
<td>192</td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of residues in the most favored/additional allowed regions (%)</td>
<td>88/12</td>
<td>89/11</td>
<td></td>
</tr>
</tbody>
</table>

\( R_{\text{merge}} \) = 100% \times \Sigma \Delta F_{h0} = \langle (I_{ih})^2 \rangle \Sigma F_{h0}^2, \) where \( I_{ih} \) is the intensity of the \( i \)-th observation of reflection \( h, \) and \( \langle I_{ih} \rangle \) is the average intensity of redundant measurements of \( h \) reflections.

\( R_{\text{work}} \) = 100% \times \Sigma |F_o| – |F_c| / \Sigma |F_o|, where \( F_o \) and \( F_c \) are the observed and calculated structure factor amplitudes.

\( ^a \) Values in parentheses are for the highest resolution shell.
residues and the opposite side with a cluster of negatively charged residues (Fig. 1, D and E). Second, dimerization brings the positively charged clusters together, whereas the negative clusters are pointing away from each other, resulting in a tripo- lar charge distribution (Fig. 1F). The presence of the negative cluster, which is located at the loop region of the molecule, was unexpected, because most of the ligands thus far shown to bind to SRs, such as modified LDL, poly I, dextran, and lipopolysaccharide, are negatively charged molecules. Nevertheless, the presence of the positive cluster may well explain the binding of the negatively charged ligands. This possibility is supported by our earlier finding that, in C-terminally truncated versions of MARCO containing only short segments of the SRCR domain, an RXR motif can provide avid bacterial binding (21).

In monomers and dimers, the positively charged cluster resides mainly within the five-stranded (A–D and F) and the eight-stranded (AB–D and F) β-sheet, respectively. It contains arginines 422, 424, 431, 433, 466, and 468 (Fig. 1), and even arginines 460 and 488 are located in close proximity. The cluster contains three linear RXR motifs and two RER motifs, Arg$_{424}$-Glu$_{435}$-Arg$_{433}$ and Arg$_{431}$-Glu$_{516}$-Arg$_{468}$, which can be regarded as structural RXR motifs. The fact that most of these motifs lie within the large β-sheet structure (especially in the case of the dimer) supports the idea of this being the ligand-binding region, since β-sheet structures are often found as ligand-docking sites. Interestingly, in the dimeric structure, the Arg$_{431}$-Glu$_{516}$-Arg$_{468}$ triplet binds, together with Arg$_{460}$ from a symmetry-related molecule, a sulfate ion, whereas Arg$_{488}$ par-
In the B-factors for the ions, they were modeled as Mg\(^{2+}\) in the structure, and because there was no significant difference in the crystallization solutions. Based on the bond lengths seen between the monomeric SRCR domain, being roughly as large as the lipid bilayer, ready to bind various ligands.

In the structure of the monomeric MARCO SRCR domain, there are two peaks in the electron density, which represent either Mg\(^{2+}\) or Na\(^{+}\), the only cations present. This is due to crystal packing and contacts within the area; Arg\(^{222}\) from the symmetry molecule forms hydrogen bonds with Asp\(^{447}\) and Glu\(^{511}\), and the carbonyl oxygen of Ala\(^{417}\) from the APLA sequence is hydrogen-bonded to the side chain oxygen of Asp\(^{488}\).

Comparison of the MARCO SRCR Domain Structure with Other Structures—When comparing the structure of the MARCO SRCR domain with the other published SRCR domain structures, those of human M2BP (Protein Data Bank code 1BY2) and human hepsin (Protein Data Bank code 1P57) (secondary structure matching, superposition of C\(^{\alpha}\) traces in COOT, r.m.s. deviation of 1.07 and 2.07 Å, respectively), it is evident that the main differences are in the lengths of loops AB and DE as well as in the long loop region (Figs. 1, A and B, and 2A) (13, 27, 31, 42, 46, 47). A 10-residue turn in the long loop region stabilized by the disulfide bond in MARCO and M2BP does not need any stabilization in hepsin, mainly due to the lack of a destabilizing Glu residue (Glu\(^{291}\) in MARCO). Instead, hepsin has a stabilizing disulfide bond at the base of the entire long loop region. It is also interesting to note that in hepsin and M2BP, the region corresponding to the negative cluster and ion-binding site of MARCO is rather positively charged (Fig. 3, A and B). The Glu\(^{511}\) participating in hydrogen bonding of the bound ion in MARCO is replaced by an Arg residue, whose positively charged side chain occupies the ion binding site. Some additional Arg residues in the region make the whole area positively charged. On the other hand, comparison of the MARCO SRCR domain sequence with that of SR-A, the protein with the most similar ligand-binding pattern, suggests that the ion-binding site is conserved. However, the predicted surface electrostatic potential of the SR-A SRCR domain shows that otherwise the region appears to be positively charged. Furthermore, the region corresponding to the positive cluster in MARCO appears to be negatively charged in SR-A (Fig. 3, A and B). These differences may explain the fact that, for SR-A, the ligand-binding site has

![Crystal Structure of the MARCO SRCR Domain](https://example.com/crystal_structure.png)
been mapped to the collagenous sequence preceding the SRCR domain (43–45).

A search for similar structures in the Protein Data Bank indicates that the SRCR fold resembles folds in some other immunologically interesting proteins, such as toxin II from the scorpion Androctonus australis (Protein Data Bank code 1PTX) (46) and the antibacterial protein sapecin from Sarcophaga perigrina (Protein Data Bank code 1L4V) (47). Both of these are small, biologically active proteins whose only secondary structure elements are a two-stranded β-sheet and a single α-helix. Indeed, folds of these proteins align well with the β-strands D and F and the α-helix I (Fig. 2B), a positively charged region in MARCO indicated in ligand binding (Fig. 1, A and B). The r.m.s. deviation calculated by secondary structure matching in COOT is 2.45 Å for both structures (Protein Data Bank codes 1PTX and 1L4V), when Cα traces were superimposed with those of the MARCO SRCR domain (31, 42). Related to these findings, it is worth noting that the interaction between the basic residues of sapecin and the phosphate groups of the acidic phospholipids was found to be of crucial importance for the antibacterial function of the peptide (49).

![Surface electrostatic potentials and sequence alignment of the SRCR domains of MARCO, M2BP, hepsin, and SR-A.](image)

**FIGURE 3.** Surface electrostatic potentials and sequence alignment of the SRCR domains of MARCO, M2BP, hepsin, and SR-A. A, surface electrostatic potentials of mouse MARCO, human M2BP (Protein Data Bank code 1BY2) (27), human hepsin (Protein Data Bank code 1P57) (13), and mouse SR-A (modeled based on the structures of MARCO and M2BP) SRCR domains. The structures are in the orientation in which the basic (blue) and acidic (red) cluster of MARCO can be seen. B, the sequence of the mouse MARCO SRCR domain (numbered above the sequence) is aligned with that of its human homolog, mouse and human SR-A, human M2BP (Protein Data Bank code 1BY2), and human CD6 (the third SRCR domain; residues 265–361) as well as with human hepsin (Protein Data Bank code 1P57) (12, 13, 22, 27, 55). The SRCR domains of MARCO, SR-A, and M2BP belong to group A, and they share the same disulfide bond pattern (marked with green numbers under the alignment). The SRCR domain of CD6 belongs to group B and contains one additional disulfide bond (the cysteines marked with magenta). On the other hand, in hepsin (group C SRCR domain), the disulfide bond 3 has changed its place (here, too, the cysteines marked with magenta) (3, 11–13). Residues similar in the group A SRCR domains and in all groups are shown, respectively, in red letters and as yellow boxes. Red boxes denote exact matches. The secondary structure alignment for the SRCR domain structures of MARCO, M2BP, and hepsin is shown above the sequences. The β-strands are shown as arrows (βA–βF), and the α-helix/310-helices are shown as spirals (α1 and α1–α2). The turns are marked T. The solvent accessibility (acc) of the residues in the mouse MARCO SRCR domain is indicated with blue color in the bar below the alignment. The alignment was made with ClustalW and further processed with ESPript (56, 57).
Charged clusters. First, we examined whether the basic cluster contributes to the ability of MARCO to support CHO cell adhesion to degraded, denatured type I collagen (gelatin). Gelatin, with a net negative charge, is found at the site of inflammation and tissue damage and needs to be removed (e.g. by macrophages). Preliminary experiments indicated that the SRCR domain is of major importance for the MARCO-mediated cell adhesion to this substrate. A study with a series of MARCO mutants harboring single or multiple Arg-to-Ala substitutions within the cluster indicates that all of the Arg residues of the basic cluster appear to be involved in the interaction of MARCO with immobilized gelatin. The results are summarized in Table 2. None of the single substitutions tested (424, 433, or 460) affected the adhesive property of MARCO, whereas a variable degree of loss of function was seen when more than one Arg residue was mutated. In particular, cells expressing the quadruple mutant 431/433/466/468 exhibited as low binding as those transfected with an irrelevant control plasmid. The triple mutant 433/466/468 and even the double mutant 431/433 also exhibited very low levels of adhesion (Figs. 4 and S2). Importantly, all of the mutants were expressed at equal or higher levels than their wild-type MARCO control (expressed from the L-vector) (Fig. 5). Moreover, unlike the other MARCO transfectants, cells expressing this mutant retained the rounded morphology of wild-type CHO cells.

In another set of experiments, the same transfectants were assayed for ApoA-I binding. Regarding this assay, we found in previous studies avid binding to cells expressing wild-type mouse MARCO, whereas binding was undetectable to cells expressing a MARCO form lacking the SRCR domain (19). The double mutant 431/433, which exhibited severe defects in adhesion to gelatin, was found to have a slightly lower ability than wild-type MARCO to bind ApoA-I (Fig. S4). The only mutant with a dramatically impaired binding ability was the quadruple mutant 431/433/466/468. In fact, this mutant did not exhibit any ApoA-I binding at all, although it was expressed at higher levels on the cell surface than its wild-type MARCO control (expressed from the L-vector) (Fig. 5). Moreover, unlike the other MARCO transfectants, cells expressing this mutant retained the rounded morphology of wild-type CHO cells.

**Table 2**

<table>
<thead>
<tr>
<th>Expression plasmid</th>
<th>Adhesion on gelatin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrelevant control</td>
<td>0</td>
</tr>
<tr>
<td>wt MARCO</td>
<td>++++</td>
</tr>
<tr>
<td>424, 431</td>
<td>+ + + +</td>
</tr>
<tr>
<td>424, 468</td>
<td>+ + +</td>
</tr>
<tr>
<td>460, 468</td>
<td>+ / + +</td>
</tr>
<tr>
<td>424, 466, 468</td>
<td>+ / + / +</td>
</tr>
<tr>
<td>431, 433</td>
<td>+</td>
</tr>
<tr>
<td>433, 466, 468</td>
<td>+ / +</td>
</tr>
<tr>
<td>431, 433, 466, 468</td>
<td>0 / +</td>
</tr>
<tr>
<td>433</td>
<td>+ + + + +</td>
</tr>
<tr>
<td>460</td>
<td>+ + + + +</td>
</tr>
<tr>
<td>424</td>
<td>+ + + + +</td>
</tr>
<tr>
<td>D447, D448</td>
<td>+</td>
</tr>
<tr>
<td>E511</td>
<td>+ / + +</td>
</tr>
<tr>
<td>wt MARCO no ion</td>
<td>0</td>
</tr>
<tr>
<td>wt MARCO plus Mg</td>
<td>+</td>
</tr>
<tr>
<td>wt MARCO plus Ca</td>
<td>+ + + + +</td>
</tr>
<tr>
<td>D447, D448, E511 no ion</td>
<td>+ / + +</td>
</tr>
</tbody>
</table>

*Adhesive activity of the mutant was compared with that of wild-type MARCO expressed from a vector lacking the CMV enhancer.*

*Adhesive activity of the mutant was compared with that of wild-type MARCO expressed from a standard CMV promoter/enhancer-containing vector.*

**FIGURE 4.** The basic cluster within the SRCR domain has a major role in the ability of MARCO to mediate cell adhesion to gelatin. CHO cells transfected with the indicated plasmids (ctrl, an irrelevant control plasmid; wt, wild-type MARCO expressed from the L-vector; for the MARCO mutants, the numbers indicate the Arg-to-Ala substitutions) were trypsinized, allowed to recover, and plated on gelatin-coated coverslips for 1 h. Thereafter, the cells were washed, fixed, and stained for F-actin. The results demonstrate the major importance of the Arg residues of the basic cluster for this adhesive function of MARCO.
We also tested whether mutations in the acidic cluster involved in metal ion binding affect the above studied functional activities of MARCO. In one of the mutants, Asp residues 447 and 448 were replaced by Ala, whereas in the other one, Ala replaced Glu511. Both mutants were well expressed on the cell surface (Figs. 6 and S5). Strikingly, both of them supported cell adhesion to gelatin at much lower levels than wild-type MARCO (Fig. S5 and Table 2). They both also exhibited very low binding of AcLDL (Fig. 6). Furthermore, the mutants failed to induce any significant changes in cell morphology (Fig. 6). We also generated a triple mutant where Ala replaced all these three residues of the acidic cluster. Surprisingly, this mutant behaved very differently from its parental mutants. Thus, it supported cell adhesion to gelatin at much lower levels than wild-type MARCO (Fig. 7 and Table 2). The transfectants also exhibited very low binding of AcLDL (Fig. 6). Moreover, it is worth noting that these transfectants have a rounded cell morphology.

We also tested whether mutations in the acidic cluster involved in metal ion binding affect the above studied functional activities of MARCO. In one of the mutants, Asp residues 447 and 448 were replaced by Ala, whereas in the other one, Ala replaced Glu511. Both mutants were well expressed on the cell surface (Figs. 6 and S5). Strikingly, both of them supported cell adhesion to gelatin at much lower levels than wild-type MARCO (Fig. S5 and Table 2). They both also exhibited very low binding of AcLDL (Fig. 6). Furthermore, the mutants failed to induce any significant changes in cell morphology (Fig. 6). We also generated a triple mutant where Ala replaced all these three residues of the acidic cluster. Surprisingly, this mutant behaved very differently from its parental mutants. Thus, it supported cell adhesion to gelatin at much lower levels than wild-type MARCO (Figs. 7 and S5 and Table 2). The transfectants also exhibited efficient spreading, which was very evident already 2–3 h after plating on gelatin- or fetal calf serum-coated coverslips. Moreover, the cells exhibited avid AcLDL binding (Fig. 6). These findings indicate that the acidic cluster plays a role in the binding functions of MARCO, possibly through both a direct and indirect mechanism.

These observations led us to examine whether chelation of divalent metal ions from wild-type MARCO affects its interaction with gelatin. The cell surface of wild-type MARCO transfectants was depleted of divalent cations by a brief EDTA treatment, and the cell adhesion assay was then carried out in the absence of exogenous cations or in the presence of Ca\(^{2+}\) or Mg\(^{2+}\), both of the cations. Remarkably, cell adhesion was completely abrogated in the absence of cations. The addition of Ca\(^{2+}\) or both Ca\(^{2+}\) and Mg\(^{2+}\) restored the adhesion to the level seen in the assay performed in Dulbecco’s modified Eagle’s medium (Fig. 7A, a and b, and Table 2). A similar stimulatory effect was seen by Mn\(^{2+}\), a nonphysiologic divalent cation (data not shown). The addition of Mg\(^{2+}\) alone also supported cell adhesion but at clearly lower levels. In this case, cells were not

![Crystal Structure of the MARCO SRCR Domain](http://www.jbc.org/)

**FIGURE 5.** The quadruple Arg-to-Ala mutant 431/433/466/468 does not bind AcLDL, a prototypic scavenger receptor ligand. Here, cells were grown on coverslips for 1 day and then assayed for AcLDL binding (right panels). Another set of samples was used for monitoring the cell surface levels of MARCO (left panels). For this, the cells were incubated on ice with the polyclonal anti-MARCO antibodies and then fixed and incubated with fluorescently labeled secondary antibodies. A and D, cells transfected with an irrelevant control plasmid; B and E, the quadruple mutant; C and F, wild-type MARCO expressed from the L-vector. The results demonstrate that the quadruple mutant does not bind AcLDL, although it is well expressed on the cell surface. It is also noteworthy that these transfectants have a rounded cell morphology.

**FIGURE 6.** Effects of mutations in the acidic cluster on the capability of MARCO for AcLDL binding. In this assay, we tested whether mutation of the residues in the acidic cluster within the SRCR domain affects the ability of MARCO for AcLDL binding (right panels). To examine the cell surface levels of the MARCO proteins, the transfectants were incubated on ice with the monoclonal antibody ED31 recognizing the SRCR domain of mouse MARCO and then fixed and incubated with fluorescently labeled anti-rat antibodies (left panels). A and F, cells transfected with an irrelevant control plasmid; B and G, the double mutant 447/448; C and H, the single mutant 511; D and I, the triple mutant 447/448/511; E and J, wild-type MARCO expressed from the L-vector. Mutation of the Asp residues 447 and 448 or Glu511 has a dramatic effect on the AcLDL binding ability of MARCO. However, simultaneous mutation of all three of these residues surprisingly largely restores the AcLDL binding function of MARCO. Moreover, it is worth noting that cells expressing the triple mutant also spread well, whereas those expressing the double and single mutant have a more rounded cell morphology.
In light of these findings, it is of interest to note two recent studies. In one report, bacterial binding to SpA, a macrophage-expressed soluble protein, was found to be markedly higher in the presence of Ca$^{2+}$ than in the presence of EDTA (18). Further, a study with thioglycollate-elicited macrophages from wild-type and SR-A knock-out mice indicated that SR-A mediates cell adhesion to human smooth muscle cell-deposited extracellular matrix and that this function is partially dependent on divalent cations (51). Curiously, sequence alignment indicates that the residues involved in ion binding in MARCO (Asp$^{447}$/Asp$^{448}$, Glu$^{511}$) are present in the SRCR domain of SR-A (Fig. 3B) as well as in the first and third SRCR domain of SpA (data not shown).

In accordance with the finding that soluble MARCO binds the polyanionic macromolecule poly I with a high affinity in the BIACore system (19), expression of the full-length protein conferred CHO cells also with the ability to adhere to this polyanion immobilized on glass coverslips. Cell adhesion was in fact somewhat more effective (20–25% more cells adhered) to poly I than to gelatin (Fig. 7, A (b) and B (b)). Of the Arg mutants, we tested only those that exhibited severe defects in adhesion to gelatin. These mutants were also found to have defects in adhesion to poly I. In particular, cells expressing the quadruple mutant 431/433/466/468 exhibited marginal binding (data not shown). The acidic cluster single (Glu$^{511}$) and double (Asp$^{447}$/448) mutants also exhibited significant defects in their adhesion, whereas the triple mutant mediated strong cell adhesion (Fig. 7B, d; data not shown). Further, binding of cells expressing wild-type MARCO was markedly affected by the absence of Ca$^{2+}$, Mg$^{2+}$ (Fig. 7A, d and B), whereas the triple mutant mediated strong cell adhesion (Fig. 7B, d; data not shown).

Altogether, the results in the ligand-binding studies may indicate that there are differences in the mode of recognition of the various negatively charged ligands. This notion is supported by the finding that there was slightly more cell adhesion to poly I than to gelatin. Additionally, comparing the cell adhesion and AcLDL binding assays, differences in assay sensitivities may have an impact on the results. It is conceivable that during the washing step after the 1-h cell adhesion or AcLDL incubation with a cell monolayer, a large “ligand” (cell) is subjected to a higher shear stress than a smaller one (AcLDL).

Finally, it has been reported that human MARCO, whose SRCR domain exhibits 74% amino acid identity with that of mouse MARCO, does not bind AcLDL (48). The sequence identity includes the Arg residues except for Arg$^{466}$, which is replaced by Lys. Thus, it appears that the Arg cluster is necessary but not sufficient for AcLDL binding. In our assay, cells transfected with a construct encoding human MARCO were found to exhibit low but detectable binding (data not shown). Similar low binding was also exhibited by transfectants expressing mouse MARCO with the human MARCO SRCR domain (Fig. S6, human MARCO SRCR sequence after Val$^{423}$; according to the mouse MARCO numbering, residues 420–518 comprise the SRCR domain). We then replaced different segments of the human MARCO SRCR domain in this chimera with corresponding segments from mouse MARCO to examine which...
Asp⁴⁹⁶ (Ser, Thr, and Ser in humans) are located at the tip of the SRCR domain. Reveals only two regions where the amino acid changes affect AcLDL binding. The assembly could proceed to the formation of oligomers (the yellow molecule has swapped domains with three other molecules), thus resulting in the creation of a large surface capable of interacting with large ligands, such as modified LDL (250 Å in diameter) or bacteria (0.2–2 μm). Top, left, surface electrostatic potential of the oligomeric MARCO model viewed from above. The structure models were generated in CCP4mg (58, 59).

Concluding remarks—It is currently unclear why MARCO depends on metal ions for its ligand-binding function. Similarly, it is not easy to understand the reason for the dramatically different binding properties of the acidic cluster triple mutant and the corresponding single and double mutants. All of these mutants have probably lost the calcium ion coordinated by the triplet Asp⁴⁴⁷, Asp⁴⁴⁸, Glu⁵¹¹, but the proteins are still well expressed on the cell surface. Therefore, it appears that the ion is not needed for structural integrity of the SRCR domain fold. A more likely explanation is that the calcium ion is directly involved in ligand binding by interacting, for example, with the phosphate group of poly I or the acidic side chains of gelatin (i.e. the same structures in these macromolecular ligands that the basic cluster interacts with). Further, it is very possible that the binding of the calcium ion and the decrease in the negative electrostatic potential of the acidic cluster can, as such, have a positive effect on the interaction between MARCO and a polyanionic ligand. Another possibility is that the change in the electrostatic potential may also affect the relative orientation of the domains in the trimeric MARCO molecule, which, in turn, may affect the avidity of ligand binding. The loss of the calcium ion could have a similar effect if residues from the SRCR domains of different subunit chains participate in the ability to swap, it can lead to the formation of large oligomeric structures easily capable of effectively binding such large objects (Fig. 8). Supporting the possibility of oligomerization, in the rotary-shadowed images of the soluble version of MARCO, molecules were often found to be associated via their globular heads (i.e. via the C-terminal SRCR domain and/or the N-terminal spacer domain) (20).

Ligands—Since MARCO is a trimeric molecule, the finding that a large proportion of the soluble SRCR domains were dimers is intriguing. Due to physical restrictions, such SRCR domain dimerization via β-strand swapping very likely occurs between separate MARCO molecules, not within a single trimeric molecule. One can envision a physiological function for this phenomenon. Thus, in addition to being able to bind low molecular weight ligands, the MARCO receptor binds bacteria, commonly 0.2–2.0 μm in diameter, and large macromolecules, such as modified LDL, which has also a diameter several times that of a single SRCR domain (25 nm versus 3 nm). If each SRCR domain of a trimeric MARCO has
metal ion coordination. Therefore, it would be of significant interest to examine the structure of the trimeric soluble MARCO either by protein crystallography or by such a technique as electron tomography, with or without bound metal ions.

Acknowledgments—We are grateful to Ekaterina Margunova (Karolinska Institutet) for help and advice in the initial phase of this work. We greatly acknowledge Péter Balogh (University of Pécs) and Lester Kobzik (Harvard School of Public Health) for providing, respectively, the anti-mouse- and -human MARCO monoclonal antibodies IBL-12 and PLK-1. We thank Martin Spiess (Biocentrum Basel) for the expression plasmid for the human asialoglycoprotein receptor H1 subunit. We also acknowledge access to synchrotron radiation at beamline 1711, Maxlab (Lund, Sweden).

REFERENCES
9. Kobzik (Harvard School of Public Health) for providing, respectively, the expression plasmid for the human asialoglycoprotein receptor H1 subunit. We also acknowledge access to synchrotron radiation at beamline 1711, Maxlab (Lund, Sweden).

Crystal Structure of the MARCO SRCR Domain

Crystal Structure of the MARCO SRCR Domain

Crystal Structure of the Cysteine-rich Domain of Scavenger Receptor MARCO Reveals the Presence of a Basic and an Acidic Cluster That Both Contribute to Ligand Recognition
Juha R. M. Ojala, Timo Pikkarainen, Ari Tuuttila, Tatyana Sandalova and Karl Tryggvason

doi: 10.1074/jbc.M701750200 originally published online April 3, 2007

Access the most updated version of this article at doi: 10.1074/jbc.M701750200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2007/04/03/M701750200.DC1

This article cites 57 references, 20 of which can be accessed free at
http://www.jbc.org/content/282/22/16654.full.html#ref-list-1