Coiled-Coil Structure of Group A Streptococcal M Proteins. Different Temperature Stability of Class A and C Proteins by Hydrophobic−Nonhydrophobic Amino Acid Substitutions at Heptad Positions a and d[†]

Tommy Cedervall, Maria U. Johansson,[§] and Bo Åkerström*,[‡]

Department of Cell and Molecular Biology and Department of Physical Chemistry 2, Lund University, Lund, Sweden

Received December 4, 1996; Revised Manuscript Received February 6, 1997[⊗]

ABSTRACT: M proteins and M-like proteins, expressed on the surface of group A streptococci and binding to human plasma proteins, can be divided into two classes, A and C, depending on the structure of the central repeated regions. The class C proteins have been shown to be dimers with a coiled-coil structure. In this work, we have compared the structure and binding of a class A protein, Mrp4, and a class C protein, Arp4, expressed by the same bacterial strain. Circular dichroism spectra, gel filtration, and binding assays showed that both proteins had a coiled-coil dimer configuration and a high-affinity binding at 20 °C. However, striking differences were seen at 37 °C. The class A protein, Mrp4, was still a coiled-coil dimer with high affinity binding activity, whereas the class C protein, Arp4, had lost both the coiled-coil structure and binding activity. Raising the temperature even higher, Mrp4 retained the coiled-coil structure up to 70−90 °C. Furthermore, a recombinant protein, Mrp(C), in which the A-repeats of Mrp4 were replaced by the C-repeats of Arp4, lost its coiled-coil structure and fibrinogen-binding around 40−45 °C. These results suggest a high thermal stability of class A proteins and a low stability of class C proteins and that the structural basis for this can be found partly in the A- and C-repeats. Analysis of the amino acid sequences of the A- and C-repeats, revealed a large difference, 87% and 45%, respectively, in the content of hydrophobic amino acid residues in the positions regarded as important for the formation of the coiled-coil structure. In particular, several alanine residues in the A-repeats were replaced by serine residues in the C-repeats. Our results suggest that important structural and functional changes within the M protein family have evolved by specific hydrophobic−nonhydrophobic amino acid replacements.

Group A streptococci are important human pathogens causing widespread diseases such as pharyngitis, acute rheumatic fever, and glomerulonephritis [for a review see Bisno (1991)]. A well-known feature of the group A streptococcus is the binding of plasma proteins. Fibrinogen, IgG, IgA, C4b-binding protein, albumin, factor H, and plasminogen are examples of proteins bound by the bacteria (Kantor, 1965; Horstmann et al., 1988; Lindahl & Åkerström, 1989; Åkesson et al., 1990; Berge & Sjöbring, 1993; Frick et al., 1994; Thern et al., 1995). The plasma proteins are bound via cell surface proteins belonging to the M protein family, called M proteins and M-like proteins.

M proteins and M-like proteins are similar molecules with a homologous primary structure. N-terminally, they have a variable region which show little sequence identity among the M protein family. In the center of the molecule is a region with conserved repeated sequences. There are two different types of repeats, A- and C-repeats, which split the M protein family into class A and C proteins (O'Toole et al., 1992). C-terminally of the repeated region is a cell wall- and membrane-spanning region, homologous among all M proteins and M-like proteins. In contrast to class C proteins, all of the studied class A proteins have highly conserved amino acid sequences (Boyle et al., 1994), forming a more homologous group than the class C proteins.

M proteins and M-like proteins have typical amino acid repeats of seven residues, similar to the fibrillar mammalian proteins tropomyosin, myosin, and laminin (Manjula & Fischetti, 1980; Fischetti, 1989; Nilson et al., 1995). Physicochemical studies, such as ultracentrifugation and electron microscopy, have shown that class C proteins are fibrillar dimeric molecules (Phillips et al., 1981; Lindahl & Åkerström, 1989). Moreover, the secondary structure of several class C molecules was approximately 75% α-helix, as determined by circular dichroism (CD) analysis (Phillips et al., 1981; Nilson et al., 1995). This has led to the proposition that class C proteins of the M protein family have a coiled-coil configuration, similar to tropomyosin and myosin (Phillips et al., 1981; Nilson et al., 1995). In a coiled-coil molecule, two α-helical monomers slowly twist around each other, forming a homodimer. Hydrophobic side chains occupy positions 1 and 4 (called a and d) of the seven-residue repeats, called heptads. A heptad represents two turns of the α-helix of the monomers, and the hydrophobic groups

[†] This work was supported by grants from the Swedish Medical Research Council (Project No. 7144), King Gustav V's 80-Year Foundation, the Medical Faculty at Lund University, The Swedish Society for Medical Research, The Royal Physiographic Society in Lund, and the Foundations of Crafoord, Greta and Johan Kock, and Alfred Österlund.

* Corresponding author: Department of Cell and Molecular Biology, Section for Molecular Signalling, Box 94, S-221 00 Lund. Sweden Tel: +46 46 222 8578. FAX: +46 46 222 4022. E-mail: Bo.Akerstrom@medkem.lu.se.

[‡] Department of Cell and Molecular Biology.

[§] Department of Physical Chemistry.

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1997.

[1] Abbreviations: BSA, bovine serum albumin; CD, circular dichroism; Ig, immunoglobulin; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

S0006-2960(96)02971-6 CCC: $14.00

in position a and d will be located on the same side of the helix, forming a basis for the attraction between the two strands of the coiled-coil dimer. Furthermore, if amino acid side chains of opposite charge are found in heptad positions 5 and 7 (called e and g), they will contribute to the attraction between the two helices.

Class C proteins undergo a dramatic conformational change with increasing temperatures. Gel filtration, ultracentrifugation and circular dichroism studies showed that the proteins existed as coiled-coil dimers at 10 and 20 °C, but unfolded to random coil monomers at 37 °C, with a transition temperature around 30−35 °C (Åkerström et al., 1992; Cedervall et al., 1995; Nilson et al., 1995). In these reports the coiled-coil dimer conformation of the bacterial proteins was shown to be a prerequisite for the plasma protein binding, as the class C proteins have a high affinity for the plasma proteins at 10 and 20 °C but a very weak binding at 37 °C (Åkerström et al., 1992; Cedervall et al., 1995). The corresponding temperature dependencies of class A protein structures and binding properties are not known.

The group A streptococcal strain AP4 expresses two M-like proteins, protein Mrp4 and protein Arp4. Protein Mrp4 is a class A protein binding to fibrinogen and IgG, and protein Arp4 is a class C protein binding to IgA. The dimerization and IgA-binding of protein Arp4 is temperature dependent (Åkerström et al., 1992). The IgA-binding by whole bacteria is also temperature dependent. Surprisingly, the fibrinogen-binding by whole bacteria is equally strong at 20 and 37 °C (Cedervall et al., 1995), suggesting that the fibrinogen-binding protein, Mrp4, is stable at high temperatures.

In this work, we have compared the structure and binding properties of protein Mrp4 and protein Arp4 at various temperatures. In contrast to protein Arp4 and other class C proteins, the coiled-coil structure of protein Mrp4 as well as its fibrinogen-binding properties were stable at high temperatures. This suggest that the class A and class C proteins in general have different thermal properties. To study the role of the A- and C-repeats, we constructed a recombinant protein Mrp in which the A-repeats were replaced by C-repeats. This hybrid molecule, protein Mrp(C), lost its coiled-coil structure at low temperatures, suggesting that the A- and C-repeats are important for the thermal stability of the M proteins and M-like proteins.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. Plasmid pARP401, containing the *arp4* gene, has been described previously (Frithz et al., 1989). Plasmid pMRP41, containing the *mrp4* gene, is derived from one of the normal size clones described earlier (Lindahl, 1989; Stenberg et al., 1992; O'Toole et al., 1992). These plasmids were gifts from Dr. Gunnar Lindahl (Department of Medical Microbiology, Lund University, Sweden). Plasmid pHD389 (Dalböge et al., 1989), used for cloning and expression, was a kind gift from Dr. Lars Björck (Department of Molecular and Cell Biology, Lund University, Sweden).

Oligonucleotides and Proteins. Oligonucleotides were from the Biomolecular Resource Facility, Lund University. The DNA sequence for primer A was AAA CAA CAA GAT GCC TCT AAG TTA GAA GCA GAA AAA; primer B, TTC TTT TTC TTG GTT TTC TGC TAA GTC TGC; for primer C, CAG TCT AGA TTA TTA GTT GGT TGT TTC TTC GCC TGT TGA; and for primer D, CAG CAG GCG CCG GAG AGT CGT CGT TAT. Restriction enzymes *Nar*I and *Xba*I, and Taq polymerase were purchased from Promega, U.S.A., and T4 ligase was from Gibco BRL, U.S.A. Human fibrinogen, human IgG, BSA (bovine serum albumin), and ovalbumin were from Sigma Chemical Co., U.S.A. Protein Arp4 was purified from *Escherichia coli* cultures transformed with pARP401 as described by Stenberg et al. (1994) and kindly provided by Dr. Gunnar Lindahl.

Construction of Mrp4- and Mrp(C)-Containing Plasmids. The DNA fragment coding for the hybrid protein Mrp(C) was made using a three-step PCR method. In step 1 polymerase chain reaction (PCR) was performed with primer A and primer B and the plasmid pARP401 as a template. The PCR product was called fragment A−B and corresponded to the C-repeats of protein Arp4, flanked by the sequences immediately upstream and downstream of the A-repeats in the *mrp4* gene. The A−B fragment was then used as a primer in step 2 together with primer C, which is complementary to a 15-nucleotide sequence just upstream of the membrane-spanning region of protein Mrp4, and plasmid pMRP41 as a template. The resulting PCR fragment, called A−C, corresponded to fragment A−B plus the *mrp4* sequence down-stream of the A-repeats. In step 3, the primers were the A−C fragment and primer D and the template pMRP41. Primer D coded for the first 15 nucleotides on the *mrp4* gene. The resulting PCR fragment was called *mrp(C)*. The *mrp4* gene was amplified in one step using primer D and primer C and pMRP41 as a template. Cleaving sites for *Nar*I and *Xba*I were introduced to the *mrp(C)* and *mrp4* genes by including them in primer D and primer C. After digestion with *Nar*I and *Xba*I, the *mrp(C)* and *mrp4* sequences were ligated into pHD389. The recombinant plasmids were introduced by electroporation into electrocompetent *E. coli*. The plasmids, purified from transformed *E. coli*, were sequenced using fluorescence-based dideoxy termination cycle reactions with Taq DyeDeoxy TM Terminator Cycle Sequencing Kit, Applied Biosystems Inc., U.S.A., and analyzed on an automatic sequencing apparatus, model 373A DNA Sequencing System, Applied Biosystems Inc., U.S.A. Ninety percent of the *mrp4* insert was sequenced and corresponded to the *mrp4* gene previously published (O'Toole et al., 1992). The *mrp(C)* insert was partly sequenced and found to be as expected, except for a truncation in the 3′-end, not involving the C-repeats.

Purification of the Gene Products. E. coli with recombinant *mrp4* or *mrp(C)* plasmids were grown in LB broth with ampicillin, 100 μg/mL, at 30 °C to $OD_{600} \approx 0.6$. The temperature was increased to 41 °C, which induces the protein production, and incubated for an additional 4 h. After the bacteria were harvested, the pellet was resuspended in 0.5 M sucrose, 1 mM EDTA, and 0.1 M Tris-HCl, pH 8.0, and the bacteria were lysed with lysozyme (80 mg/L) for 10 min and centrifuged. The supernatants were subjected to fibrinogen-Sepharose affinity chromatography, and the bound proteins were eluted with 0.1 M glycine-HCl, pH 2.5. The eluted proteins were dialyzed against 0.05 M Tris-HCl, pH 8.0, and 0.01 M NaCl and applied to a DEAE-Sephadex A50 column (Pharmacia Fine Biochemicals, Sweden) equilibrated with 0.05 M Tris-HCl and 0.05 M NaCl, pH 8.0, and eluted with a continuous gradient of 0.01−0.2 M NaCl in 0.05 M Tris-HCl, pH 8.0. For protein Mrp4, the ion-exchange

chromatography was followed by IgG-Sepharose affinity chromatography. Fibrinogen and IgG were coupled to CNBr-activated Sepharose (Pharmacia Fine Chemicals, Sweden), according to the manufacturer's recommendations. The purified proteins were dialyzed against PBS (10 mM phosphate buffer, 0.12 M NaCl, 3 mM KCl, pH 7.4) prior to subsequent experiments.

Labeling of Proteins. Human IgG, fibrinogen, and bacterial proteins were radiolabeled with ^{125}I (Nordion Inc., Canada) to a specific activity of approximately 0.6 MBq/μg using the chloramin T-oxidation method (Greenwood et al., 1963).

SDS−PAGE, Western Blotting, and Binding Experiments. Protein Mrp4 and Mrp(C) were separated by 12% SDS−PAGE according to Laemmli (1970). The proteins were stained with Kenacid Blue or transferred to Immobilon membranes (Millipore, U.S.A.) and blotted with [^{125}I]-fibrinogen. The bound radioactivity was analyzed in a Fujix BAS 2000 Bio-imaging analyzer (Fuji films Co., Japan). Human fibrinogen or IgA was coupled to Immuno-Beads (Bio-Rad Laboratories, U.S.A.) according to the manufacturer's recommendations. 100 μL of radiolabeled protein Mrp4 or Arp4 (approximately 850 Bq) in PBS + 0.02% BSA and 0.05% Tween 20 were mixed with fibrinogen- or IgA-beads in the same buffer. After incubation at 4, 20, or 37 °C for 3 h, 1 mL of PBS, having the same temperature, was added and the mixture centrifuged. Nonspecific binding was measured by mixing radiolabeled protein Mrp4 or Arp4 with empty Immuno-Beads. The amount of bound radioactivity on the beads was analyzed in a γ-counter.

Gel Filtration. Purified protein Mrp4 or Mrp(C) (approximately 100 μg) or radiolabeled protein Arp4 (approximately 12 kBq) was subjected to gel chromatography on a column (1 × 40 cm) packed with Sephacryl S-300 (Pharmacia Fine Biochemicals, Sweden) and eluted with 20 mM Tris-HCl, pH 8.0, + 0.15 M NaCl + 0.02% NaN_3 at a flow rate of 3 mL/h. The gel chromatography experiments were performed at 10, 37, and 50 °C. The void and total volumes were estimated as the elution volumes of Blue Dextran and DNP-alanine, respectively. Fractions of 1 mL were collected and analyzed for protein content by absorbance at 230 nm, and fractions containing proteins were also analyzed by SDS−PAGE. Fractions from gel filtrations of protein Arp4 were analyzed in a γ-counter.

The molecular size of denatured protein Mrp4 was determined by gel chromatography on a column (2.1 × 1000 cm) packed with Sepharose CL-6B (Pharmacia Fine Chemicals, Sweden), equilibrated and eluted with 50 mM sodium acetate (pH 4.8) + 6 M guanidine hydrochloride. Radiolabeled protein Mrp4, BSA, and lysozyme were reduced and alkylated in the elution buffer as described (Ekström & Berggård, 1977) before application to the column. BSA and lysozyme were included as markers with known molecular weights. The void and total volumes were determined as above.

Circular Dichroism Spectroscopy. Protein Mrp4, protein Mrp(C), and protein Arp4 were subjected to CD spectroscopy. The proteins were dialyzed against PBS. Experiments with protein Mrp4 were performed at concentrations of 4.2, 2.0, and 0.4 μM and with proteins Arp4 and Mrp(C) at concentrations of 4.1 and 0.4 μM, respectively. Concentrations were estimated from the absorbance at 280 nm. The CD spectra were recorded in the far UV region (190−260 nm) on a Jasco J-720 spectropolarimeter at a scan speed of 10 nm/min and with a 4 s response time. CD spectra for each protein sample were obtained at 20, 37, 20, 50, and 20 °C (in this order). The ratio of ellipticity was calculated from the value at the minimum at 222 nm divided by the value at the minimum at 210 nm (protein Arp4 and protein Mrp4) or at 209 nm [protein Mrp(C)]. A temperature scan at 222 nm was also performed from 5 to 95 °C with a scan rate of 50 °C/h.

Computational Coiled-Coil Prediction and Analysis of the Content of Hydrophobic Amino Acids. The amino acid sequences of protein Mrp4, Arp4, and Mrp(C) were analyzed for the probability to form coiled-coils, using the program PAIRCOIL (version 1.0) (Berger et al., 1995). A cutoff of 50% was used. The program also predicted the probable heptad repeat structures, in which the seven individual amino acid positions were denoted as a−g. These heptads were used to analyze the identity and number of hydrophobic amino acids in positions a and d and charged amino acids in positions e and g.

RESULTS

Purification of Protein Mrp4 and Protein Mrp(C). The structures and binding properties of protein Mrp4, Arp4, and Mrp(C) were compared at various temperatures. Protein Arp4 has previously been examined by binding assays, electrophoresis, and gel chromatography at 10, 20, and 37 °C (Åkerström et al., 1992). In this study, protein Mrp4 and protein Mrp(C) were initially isolated. The recombinant *mrp4* and *mrp(C)* genes were expressed in *E. coli.* The protein products were found in the supernatant after lysozyme digestion of the bacterial cell wall and release of the periplasmic proteins. The production of protein Mrp4 and Mrp(C) was approximately 25 mg/L of bacterial culture. The proteins were isolated by fibrinogen-affinity chromatography and ion-exchange chromatography; protein Mrp4 required an additional IgG-affinity chromatography step. The purified protein Mrp4 migrated on SDS−PAGE as a 39 kDa molecule (Figure 1) which is close to the molecular weight (37.987) calculated from the gene sequence of *mrp4* (O'Toole et al., 1992). Similar to other M proteins and M-like proteins, minor bands of protein Mrp4 were seen around 36 kDa, most likely representing cleavage products at the C-terminus of the major protein product. The molecular weight of protein Mrp4 was also determined by gel chromatography in the presence of 6 M guanidine-HCl and found to be 38 kDa. Protein Mrp(C) migrated as a major 34 kDa band and two additional bands with the M_r values 29 and 26 kDa. The three bands were seen in repeated expression experiments. The N-terminal sequence of protein Mrp4 and the 34 and 29 kDa bands of protein Mrp(C) was RYQAPP, which corresponds to amino acid residues 4−8 of the previously reported sequence of protein Mrp4 (O'Toole et al., 1992). This suggests a nonidentical signal peptide processing in *E. coli* and the streptococcal strain AP4, the natural expression system of protein Mrp4. Some of the biochemical and physicochemical properties of protein Mrp4, Mrp(C), and Arp4 are summarized in Table 1.

Fibrinogen Binding by Protein Mrp4 and Mrp(C). The binding of fibrinogen to protein Mrp4 and protein Mrp(C) was first tested by Western blotting (not shown). Both the 39 and the 36 kDa bands of protein Mrp4 bound to [^{125}I]-

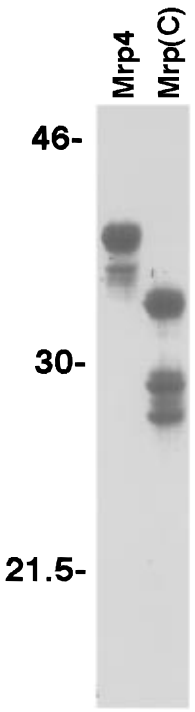

FIGURE 1: SDS−PAGE of purified protein Mrp4 and protein Mrp-(C). The *mrp4* and *mrp(C)* constructs were expressed in *E. coli* and purified as described in Materials and Methods.

Table 1: Some Properties of Protein Mrp4, Protein Mrp(C), and Protein Arp4

parameter	Mrp4	Mrp(C)	Arp4
molecular mass deduced from amino acid sequence (Da)	37 987[a]		39 544[a]
molecular mass by SDS−PAGE (kDa)	39	34[b]	41
molecular mass by gel chromatography in 6 M guanidine-HCl (kDa)	38		41[c]
N-terminal sequence	RYQAPP	RYQAPP[d]	
S−S bonds	none	none	none
ratio of ellipticity at 20 °C[e]	1.08	1.08	1.08

[a] Processed protein Mrp4 (O'Toole et al., 1992) and protein Arp4 (Frithz et al., 1989). [b] Molecular mass of the major band seen in the SDS−PAGE. [c] From Åkerström et al. (1991). [d] Both the 34 and 29 kDa bands. [e] Ratio of ellipticity was calculated as described in Materials and Methods.

fibrinogen, and the relative binding intensity was approximately the same as for the stained bands. All three bands of protein Mrp(C) bound fibrinogen. The binding of radiolabeled protein Mrp4 to fibrinogen was investigated at 20 and 37 °C. The fibrinogen binding by ^{125}I-labeled protein Mrp4 was found to be almost equally strong at 37 and 20 °C (Figure 2). In contrast, ^{125}I-labeled protein Arp4 bound strongly to IgA at 20 but weakly at 37 °C. These results show that the ligand binding by protein Mrp4 is temperature independent, whereas the ligand-binding by protein Arp4 is temperature dependent.

Gel Filtration of Protein Mrp4, Mrp(C), and Arp4 at Different Temperatures. Protein Mrp4 and Mrp(C) were subjected to gel filtration at 20, 37, and 50 °C (Figure 3). Protein Mrp4 was eluted as a large molecule at all three temperatures. The elution volumes for the protein was not changed with increasing temperature, suggesting that the molecule is stable even at 50 °C. Protein Mrp(C) was also eluted as a large molecule at 20 and 37 °C but as a smaller

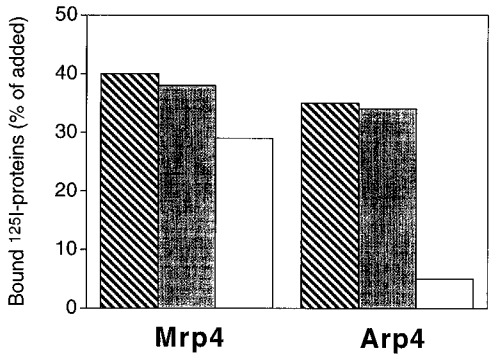

FIGURE 2: Binding of protein Mrp4 and protein Arp4 to fibrinogen and IgA, respectively, at 4 °C (striped), 20 °C (filled), and 37 °C (empty). Radiolabeled protein Mrp4 or Arp4 was incubated at different temperatures with fibrinogen or IgA coupled to Immuno-Beads. Nonspecific binding was measured to noncoupled Immuno-Beads and subtracted from the values.

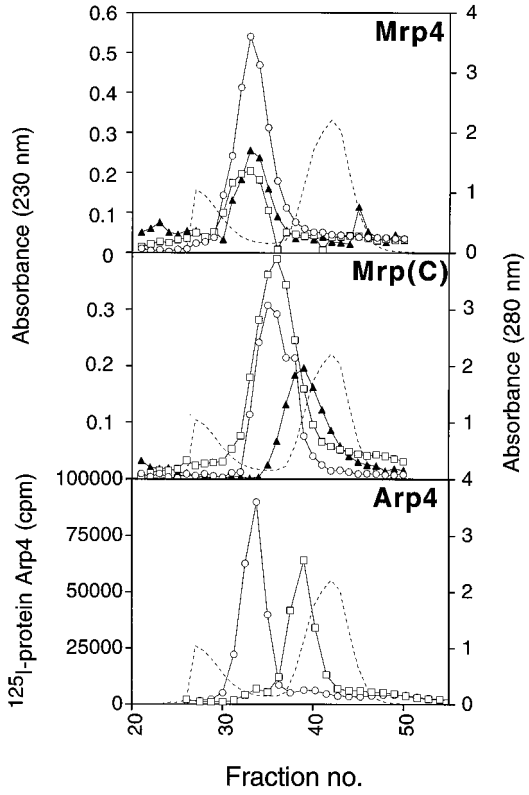

FIGURE 3: Gel filtration of protein Mrp4, protein Mrp(C) and radiolabeled protein Arp4 at 20 °C (○), 37 °C (□), and 50 °C (▲). The eluted protein Mrp4 and Mrp(C) were detected by measuring absorbance at 230 nm. Blue dextran and ovalbumin (dashed line) are shown as markers of elution volumes, and were detected by measuring absorbance at 280 nm.

molecule at 50 °C. Protein Arp4 was subjected to gel filtration at 20 and 37 °C under the same conditions, and was eluted as a large molecule at 20 °C and as a small molecule at 37 °C. Protein Arp4 was previously shown to be a dimer at 20 °C and a monomer at 37 °C (Åkerström et al., 1992), suggesting that protein Mrp(C) is a dimer at 20 and 37 °C and a monomer at 50 °C, while protein Mrp4 appears to be a dimer even up to 50 °C. Finally, in a gel filtration at 20 °C after pre-incubation at 50 °C for 3 h, protein Mrp(C) was eluted at the same position as when loaded directly to the column at 20 °C, demonstrating that the conformational change between monomer and dimer was reversible.

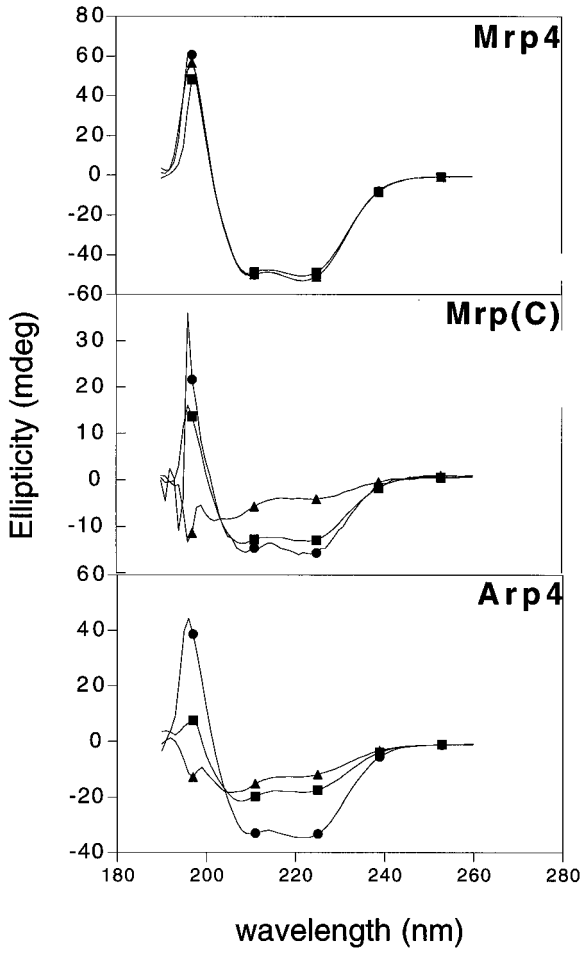

FIGURE 4: Circular dichroism spectra of protein Mrp4, protein Arp4, and protein Mrp(C) at 20 °C (●), 37 °C (■), and at 50 °C (▲). The experiments were also made with a 10-fold dilution of the proteins, with similar temperature dependencies (data not shown). After experiments at 37 and 50 °C, the temperature was decreased to 20 °C again, and a similar spectrum was obtained as before elevation of the temperature (data not shown).

Secondary Structure of the Bacterial Proteins at Different Temperatures. The secondary structure of protein Mrp4, Mrp(C), and Arp4 was qualitatively investigated by CD spectroscopy. At 20 °C, the CD spectra of all three proteins displayed two distinct minima at 207−210 and 222 nm (Figure 4, circles), typical of an α-helical secondary structure. The spectrum of protein Mrp4 was identical at 20, 37, and 50 °C. In contrast, the spectrum of protein Arp4 was dramatically changed at higher temperatures, suggesting unfolding of this protein at 37 (squares) and at 50 °C (triangles). Protein Mrp(C) was only slightly affected at 37 °C but had unfolded at 50 °C. The unfolding processes were reversible, since an identical spectrum was obtained at 20 °C, whether or not the proteins were first heated to 50 °C (data not shown). Furthermore, the ratio between the ellipticity at 207−210 and 222 nm was calculated from the values obtained at 20 °C (Table 1). All three proteins had similar ratios as previously reported for coiled-coil molecules (Lau et al., 1984; Hodges et al., 1988, 1990; Zhou et al., 1992a,b).

The CD signal at 222 nm was measured at different temperatures and plotted as a function of the temperature (Figure 5). Distinct transition regions between 30−40 and 40−50 °C were seen for proteins Arp4 and Mrp(C), respectively. The transition curve for protein Mrp4 was more complicated, with no region comparable to the sudden drop of α-helical content seen for protein Arp4 and Mrp(C). At 75 °C, 50% of the initial α-helical content remained and the unfolding of the protein was still not complete at 90 °C.

Predictions of the Secondary Structure of Protein Mrp4, Mrp(C), and Arp4. The amino acid sequences of protein Mrp4, Mrp(C), and Arp4 were analyzed for the probability of coiled-coil structure (Figure 6). High coiled-coil probability (70−100%) was found for the major parts of all three proteins, including the A- and C-repeats. Low coiled-coil probability (0−50%) was found for the N-terminal 27, 27, and 15 amino acids of protein Mrp4, Mrp(C), and Arp4, respectively. Low coiled-coil probability was also found in the C-terminal 67 and 58 amino acids of protein Mrp4 and protein Arp4, respectively. Medium coiled-coil probability (50−70%) was found in a region between the N-terminal low-probability and central high-probability regions, in all three proteins. In the central region of protein Mrp(C), the coiled-coil probability dropped slightly to 87% where the regions derived from protein Mrp4 and Arp4 were connected. Within the coiled-coil region, the heptad formation was broken at amino acid residue 54 in protein Mrp4, at residues

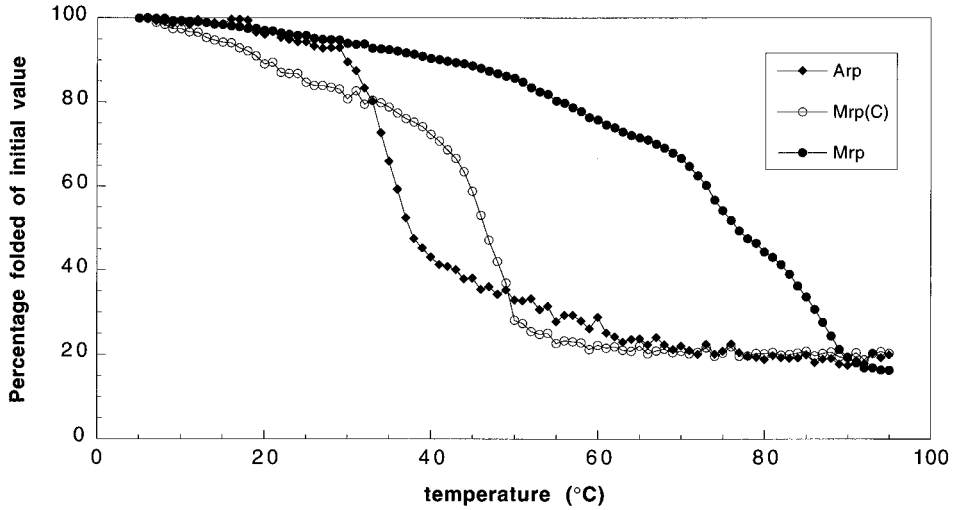

FIGURE 5: Circular dichroism temperature scan of protein Mrp4 (●), protein Arp4 (◆) and protein Mrp(C) (○) between 5 and 95 °C. The ellipticity was measured at 222 nm. The values are expressed as percent of the ellipticity obtained at 5 °C. Similar results were obtained in experiments with a 10-fold dilution of the proteins.

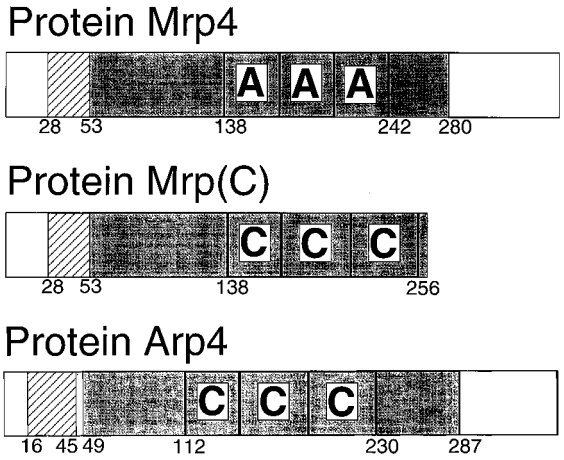

FIGURE 6: Schematic primary structure and coiled-coil prediction of protein Mrp4, Mrp(C), and Arp4. Protein Mrp4 has three A-repeats containing 35 amino acids, and protein Arp4 has three C-repeats containing 35, 42, and 42 amino acids. Empty fields represent regions with less than 50% coiled-coil probability, dashed fields represent 50−70%, and shaded fields represent more than 70% probability for coiled-coil structure.

54, 135, and 153 in protein Mrp(C), and at residues 96 and 110 in protein Arp4.

Melting Temperatures Are Related to the Structure of the Heptad Regions. The regions of protein Mrp4 and Arp4 with a coiled-coil probability above 50% were analyzed for the seven residue periodicity typical for coiled-coil sequences. Figure 7 shows the amino acid residues occupying the heptad positions a and d in the two proteins. In protein Mrp4, hydrophobic amino acids (alanine, leucine, isoleucine, and valine) were found in 76% of the a positions and 75% of the d positions. In the A-repeats, 87% and 87%, respectively, of both positions a and d carried hydrophobic amino acids. In protein Arp4, only 30% and 56% of the a and d positions, respectively, were hydrophobic, and in the C-repeats the values were 31% and 59%, respectively. Thus, protein Mrp4, a class A protein, had a much higher degree of hydrophobic amino acids in positions a and d than protein Arp4. The higher content of hydrophobic heptad positions a and d is thus one possible explanation for the higher temperature stability of protein Mrp4.

Table 2 shows a comparison of the hydrophobic amino acid content in the a and d positions of M proteins and M-like proteins for which the CD spectrum at different temperatures is known. This value was compared to the temperature at which 50% unfolding of the protein was seen. The table shows that there is a correlation between the number of hydrophobic amino acids in positions a and d and the unfolding temperature. Protein Mrp4 has the highest number of hydrophobic amino acids and the highest unfolding temperature. Protein Mrp(C) and M1 protein have intermediate amounts of hydrophobic amino acids and intermediate melting temperatures, whereas protein H and protein Arp4 have the lowest amounts of hydrophobic amino acids and the lowest melting temperatures.

A closer examination of the a and d positions within the A- and C-repeats revealed a similarity in the d position between the two classes of proteins (Table 3). The order of amino acids in the d positions in both types of repeats was L, L, X, L, X in both proteins, with an additional valine at the end of the two last C-repeats in protein Arp4. X was alanine in the two first A-repeats and threonine in the third,

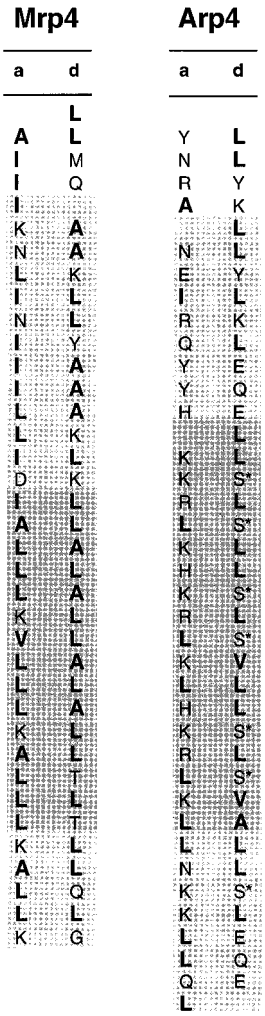

FIGURE 7: The heptad structures of protein Mrp4 and protein Arp4. The amino acid residues in position a and d in each heptad are listed. Amino acids A, L, I, and V are marked throughout the molecule with bold letters. Serines within the coiled-coil region of protein Arp4 are marked as S*. Regions with more than 70% probability of coiled-coil structure, and the A-repeats (amino acid residues 138−242) and the C-repeats (amino acid residues 112−230) are marked with light and dark shading, respectively. The last 67 amino acids (protein Mrp4) and 58 amino acids (protein Arp4) have low coiled-coil probability, and are not included in the figure.

Table 2: Hydrophobic Amino Acids in Heptad Positions a and d and Unfolding Temperature of M Proteins and M-Like Proteins

	% hydrophobic amino acids[a]			
		repeats[c]		
protein	coiled-coil[b]	a	d	temperature[d] (°C)
protein Mrp4	75	87	87	73
protein Arp4	43	31	59	36
protein Mrp(C)	56	31	59	44
M1 protein[e]	58	47	62	41
protein H[e]	44	22	67	29

[a] Percent of the a and d positions occupied by the amino acids A,V, L, or I. [b] Regions with more than 50% probability of coiled-coil structure. [c] A- and C-repeats. [d] Temperature at which the CD signal at 222 nm decreased to 50% of the value between 5 and 90 °C. [e] Heptad repeats and temperatures are taken from Nilson et al. (1995).

but in the C-repeats all X were serine. Replacement of alanine by serine in the d position has been shown to decrease the melting temperature of synthetic helical peptides by 11 °C, and the decrease by several substitutions was additive (Merutka & Stellwagen, 1990). This suggests a possible

Table 3: Amino Acid Residues in Heptad Positions d of the A- and C-Repeats in Protein Mrp4 and Protein Arp4

Mrp4[a] A-repeats			Arp4[b] C-repeats		
1st	2nd	3rd	1st	2nd	3rd
L	**L**	**L**	**L**	**L**	**L**
L	**L**	**L**	**L**	**L**	**L**
A	**A**	T	S	S	S
L	**L**	**L**	**L**	**L**	**L**
A	**A**	T	S	S	S
				V	**V**

[a] Protein Mrp4 has three A-repeats, each repeat consisting of 35 amino acids forming five heptads. [b] Protein Arp4 has three C-repeats, consisting of 35, 42, and 42 amino acids forming 5, 6, and 6 heptads, respectively.

explanation for the difference in the temperature stabilities of protein Mrp4 and protein Arp4.

Amino acid side chains with opposite charges in positions e and g can contribute to the attraction between the two helices in a coiled-coil protein. In protein Mrp4, M1 protein, protein Arp4, and protein H were found 10, 26, 18, and 22 such attractive interhelical ionic interactions, respectively. Repulsive interhelical ionic interactions, *i.e.*, amino acid side chains with the same charge in the e and g positions, were found in 6, 8, 6, and 12 interactions, respectively. Thus, no clear relationship was seen between the number of interhelical ionic interactions and the temperature stability of these proteins.

DISCUSSION

M-like proteins of the class C category have been shown to be α-helical dimeric coiled-coil molecules (Phillips et al., 1981; Nilson et al., 1995), but the stability of the coiled-coil structure is temperature sensitive. Thus, the dimers of protein Arp4, protein H, protein Sir22, and M1 protein dissociate into monomers when the temperature is raised to 37 °C (Åkerström et al., 1992; Cedervall et al., 1995). The monomerization of the class C proteins is accompanied by a decrease of the α-helical secondary structure (Nilson et al., 1995; this report) and the strength of the plasma protein binding (Åkerström et al., 1992; Cedervall et al., 1995). In this report we show that the structure and binding properties of a class A protein, protein Mrp4, is largely independent of the temperature. In contrast to the class C proteins, the α-helical conformation and dimerization of protein Mrp4 were stable at temperatures above 37 °C and the fibrinogen-binding was still strong at 37 °C. These results reveal a striking structural and functional difference between protein Mrp4 and protein Arp4, which are closely related proteins, located on the same bacterial surface.

Coiled-coil structures have been shown to be destabilized by the replacement of a hydrophobic amino acid with a less hydrophobic amino acid in the a and d positions of the heptad repeats (Hu et al., 1990; Zhou et al., 1992a,b). An examination of the amino acid sequences of protein Mrp4 and protein Arp4 showed a higher degree of hydrophobic amino acids in the a and d positions of protein Mrp4 than in protein Arp4. This is probably one of the reasons for the difference in temperature stability between protein Mrp4 and protein Arp4. The difference in hydrophobic residue content was even more pronounced in the centrally located repeats, which have been proposed as a framework for the formation of the coiled-coil structure (Nilson et al., 1995). In the d positions, the hydrophobic−nonhydrophobic replacements were accomplished by a specific substitution of alanines in four of the heptads of the A-repeats for serines in the four corresponding heptads of the C-repeats. The consistency of these amino acid substitutions suggests that the serine residues in the C-repeats are important for the temperature sensitivity of protein Arp4. This is supported by observations by others that a substitution of alanines by serines decreases the melting temperature in synthetic helical peptides (Merutka & Stellwagen, 1990). The A- and C-repeats are distinct regions within the M proteins, suggesting that they are used as genetic packages to facilitate the spreading of temperature stability or instability among M proteins and M-like proteins. Indeed, when C-repeats were introduced to protein Mrp, forming the recombinant protein Mrp(C), the thermal stability was dramatically reduced.

The variable N-terminal regions in protein Mrp4 and protein Arp4 also displayed a probability for coiled-coil structure. Thus, these regions are likely to contribute to the overall stability of the molecules. A comparison of the hydrophobic residues in this region between protein Mrp4 and class C proteins showed, as expected, that protein Mrp4 had more hydrophobic amino acids than protein Arp4 and protein H in the heptad a and d positions. Interestingly, M1 protein, a class C protein, had the same amounts of hydrophobic residues as protein Mrp4 in the N-terminal region. Consequently, the unfolding of M1 protein, although it was temperature-sensitive, required higher temperatures than protein Arp4 and protein H and it had a less distinct temperature transition curve than these proteins (Nilson et al., 1995). Furthermore, protein Mrp(C), containing the N-terminal part of protein Mrp4 had a higher melting temperature than protein Arp4 and protein H. Taken together, these results suggest that the melting temperature of class C proteins, to some degree, is correlated to the ratio of hydrophobic residues in the a and d positions both in the N-terminal and central regions of the proteins.

The structural and functional instability of the class C proteins makes them well adapted for a role as sensors of small changes in the near environment of the bacteria. The degree of folding of coiled-coil proteins in general is dependent on factors such as pH (Lau et al., 1984; Hodges et al., 1988, 1990; Zhou et al., 1994) and the ionic strength (Monera et al., 1994; Yu et al., 1996), and the degree of unfolding of the class C proteins is dependent on the temperature and the concentration of ligands (Nilson et., 1995; Cedervall et al., 1995; this work). Moreover, the expression of M proteins is regulated by environmental factors such as CO_2 pressure (Podbielski et al., 1992; Caparon et al., 1992; Okada et al., 1993), temperature, osmolarity, and iron concentration (McIver et al., 1995). Therefore it seems plausible that class C proteins have the role of a biosensor, detecting small changes in the environment by a structural shift of its coiled-coil structure, induced by these changes. Tentatively, the structural shift is then transmitted through the cell-wall to the interior of the bacterium. This hypothesis is presently investigated in our laboratory by analyzing the influence of different environmental factors at temperatures near 37 °C on the coiled-coil structure of class A and C proteins.

ACKNOWLEDGMENT

Maria Allhorn is acknowledged for excellent technical assistance as is Dr. Gunnar Lindahl for helpful discussions.

REFERENCES

Åkerström, B., Lindqvist, A., & Lindahl, G. (1991) *Mol. Immunol. 28*, 349−357.
Åkerström, B., Lindahl, G., Björck, L., & Lindqvist, A. (1992) *J. Immunol. 148*, 3238−3243.
Åkesson, P., Cooney, J., Kishimoto, F., & Björck, L. (1990) *Mol. Immunol. 27*, 523−531.
Berge, A., & Sjöbring, U. (1993) *J. Biol. Chem. 268*, 25417−25424.
Berger, B., Wilson, D. B. Wolf, E., Tonchev, M. M., & Kim, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8259−8263.
Bisno, A. L. (1991) *N. Engl. J. Med. 325*, 783−793.
Boyle, M. D. P., Hawlitzky, J., Raeder, R., & Podbielski, A. (1994) *Infect. Immun. 62*, 1336−1347.
Caparon, M. G., Geist, R. T., Perez-Casal, J., & Scott, J. R. (1992) *J. Bacteriol. 174*, 5693−5701.
Cedervall, T., Åkesson, P., Stenberg, L., Herrmann, A., & Åkerström, B. (1995) *Scand. J. Immunol. 42*, 433−441.
Dalböge, H., Bech Jensen, E., Tøttrup, H., Grubb, A. Abrahamson, M., Olafsson, I., & Carlsen, S. (1989) *Gene 79*, 325−332.
Ekström, B., & Berggård, I. (1977) *J. Biol. Chem. 252*, 8048−8057.
Fischetti, V. A. (1989) *Clin. Microbiol. Rev. 2*, 285−314.
Frick, I.-M., Åkesson, P., Cooney, J., Sjöbring, U., Schmidt, K.-H., Gomi, H., Hattori, S., Tagawa, C., Kishimoto, F., & Björck, L. (1994) *Mol. Microbiol. 12*, 143−151.
Frithz, E., Hedén, L.-O., & Lindahl, G. (1989) *Mol. Microbiol. 3*, 1111−1119.
Greenwood, F. C., Hunter, W. M., & Glover, J. S. (1963) *Biochem. J. 89*, 114−123.
Hodges, R. S., Semchuck, P. D., Taneja, A. K., Kay, C. M., Parker, J. M. R., & Mant, C. T. (1988) *Peptide Res. 1*, 19−30.
Hodges, R. S., Zhou, N. E., Kay, C. M., & Semchuck, P. D. (1990) *Peptide Res. 3*, 123−137.
Horstmann, R. D., Sievertsen, H. J., Knobloch, J., & Fischetti, V. A. (1988) *Proc. Natl. Acad. Sci. U. S. A 85*, 1657−1661.
Hu, J. C., O'Shea, E. K., Kim, P. S., & Sauer, R. T. (1990) *Science 250*, 1400−1403.
Kantor, F. S. (1965) *J. Exp. Med. 121*, 849−859.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Lau, S. Y. M., Taneja, A. K., & Hodges, R. S. (1984) *J. Biol. Chem. 259*, 13253−13261.
Lindahl, G. (1989) *Mol. Gen. Genet. 216*, 372−379.
Lindahl, G., & Åkerström, B. (1989) *Mol. Microbiol. 3*, 239−247.
Manjula, B. N., & Fischetti, V. A. (1980) *J. Exp. Med. 151*, 695−708.
McIver, K. S., Heath, A. S., & Scott, J. R. (1995) *Infect. Immun. 63*, 4540−4542.
Merutka, G., & Stellwagen, E. (1990) *Biochemistry 29*, 894−898.
Monera, O. D., Kay, C. M., & Hodges R. S. (1994) *Biochemistry. 33*, 3862−3871.
Nilson, B. H. K., Frick, I.-M., Åkesson, P., Forsén, S., Björck, L., Åkerström, B., & Wikström, M. (1995) *Biochemistry 34*, 13688−13698.
Okada, N., Geist, R. T., & Caparon, M. G. (1993) *Mol. Microbiol. 7*, 893−903.
O'Toole, P., Stenberg, L., Rissler, M., & Lindahl, G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8661−8665.
Phillips, G. N., Jr., Flicker, P. F., Cohen, C., Manjula, B. N., & Fischetti, V. A. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4689−4693.
Podbielski, A., Peterson, J. A., & Cleary, P. (1992) *Mol. Microbiol. 6*, 2253−2265.
Stenberg, L., O'Toole, P., & Lindahl, G. (1992) *Mol. Microbiol. 6*, 1185−1194.
Stenberg, L., O'Toole, P. W. Mestecky, J., & Lindahl, G. (1994) *J. Biol. Chem. 269*, 13458−13464.
Thern, A., Stenberg, L., Dahlbäck, B., & Lindahl, G. (1995) *J. Immunol. 154*, 375−386.
Yu, Y., Monera, O. D., Hodges, R. S., & Privalov, P. L. (1996) *J. Mol. Biol. 255*, 367−372.
Zhou, N. E., Kay, C. M., & Hodges, R. S. (1992a) *J. Biol. Chem. 267*, 2664−2670.
Zhou, N. E., Kay, C. M., & Hodges, R. S. (1992b) *Biochemistry 31*, 5739−5746.
Zhou, N. E., Kay, C. M., & Hodges, R. S. (1994) *J. Mol. Biol. 237*, 500−512.

BI962971Q