High-Affinity Binding of the Staphylococcal HarA Protein to Haptoglobin and Hemoglobin Involves a Domain with an Antiparallel Eight-Stranded β-Barrel Fold

Agnieszka Dryla,1‡ Bernd Hoffmann,2‡‡ Dieter Gelbmann,1 Carmen Giefing,1 Markus Hanner,1 Andreas Meinke,1 Annaliesa S. Anderson,4 Walter Koppensteiner,3 Robert Konrat,2 Alexander von Gabain,1 and Eszter Nagy1*

Intercell AG, Vienna Biocenter 6, A-1030 Vienna, Austria;1 Department of Biomolecular Structural Chemistry, University of Vienna, Vienna Biocenter 5, A-1030, Vienna, Austria;2 Biovertis AG, Vienna Biocenter 6, A-1030 Vienna, Austria;3 and Merck and Co. Incorporated, 440 Summertown Pike, WP16 100, Westpoint, Pennsylvania 194864

Received 29 August 2006/Accepted 3 October 2006

Iron scavenging from the host is essential for the growth of pathogenic bacteria. In this study, we further characterized two staphylococcal cell wall proteins previously shown to bind hemoproteins. HarA and IsdB harbor homologous ligand binding domains, the so-called NEAT domain (for “near transporter”) present in several surface proteins of gram-positive pathogens. Surface plasmon resonance measurements using glutathione S-transferase (GST)-tagged HarAD1, one of the ligand binding domains of HarA, and GST-tagged full-length IsdB proteins confirmed high-affinity binding to hemoglobin and haptoglobin-hemoglobin complexes with equilibrium dissociation constants (K_D) of 5 to 50 nM. Haptoglobin binding could be detected only with HarA and was in the low micromolar range. In order to determine the fold of this evolutionarily conserved ligand binding domain, the untagged HarAD1 protein was subjected to nuclear magnetic resonance spectroscopy, which revealed an eight-stranded, purely antiparallel β-barrel with the strand order (β1-β2-β3-β6-β5-β4-β7-β8), forming two Greek key motifs. Based on structural-homology searches, the topology of the HarAD1 domain resembles that of the immunoglobulin (Ig) fold family, whose members are involved in protein-protein interactions, but with distinct structural features. Therefore, we consider that the HarAD1/NEAT domain fold is a novel variant of the Ig fold that has not yet been observed in other proteins.

Staphylococcus aureus is a highly successful human pathogen implicated in a wide spectrum of diseases. The major medical concern is the potential of S. aureus, with multidrug-resistant strains, to cause severe infections in hospitals. Up to 65% of nosocomial staphylococcal infections are already resistant to methicillin (methicillin-resistant S. aureus [MRSA]), and clinical isolates that cannot be eliminated with available antibiotics (vancomycin-resistant S. aureus) are being reported worldwide (22). The epidemiology of MRSA is now changing; infections are no longer confined to the hospital setting but also appear in healthy community-dwelling individuals, most notably in children without established risk factors for the acquisition of MRSA (23, 27). As a consequence, there is a need to identify and characterize suitable targets for new treatment modalities.

Intervening in bacterial iron acquisition systems is a potential strategy to fight pathogens, since iron is the most important nutrient that limits growth inside the human host. Due to this dependence on iron availability, pathogens have to utilize high-affinity iron acquisition systems to gain access to this essential nutrient. Since free iron is highly toxic for living organisms, the majority of iron is compartmentalized intracellularly (by hemo-globin [Hb], myoglobin, ferritin, and iron-containing enzymes) or sequestered by plasma proteins (haptoglobin [Hp]-hemoglobin complexes, transferrin, and lactoferrin). Pathogens have evolved diverse mechanisms to gain access to the different iron sources that are available to them in their host organisms (2, 5, 6).

Iron acquisition by pathogens involves the recognition and capture of bacterial siderophores or iron-containing host proteins by surface receptors and transport of Fe or heme by membrane proteins into the bacterial cells (reviewed in reference 46). S. aureus uses both mechanisms and possesses redundant iron uptake systems. Analysis of the complete genome of sequenced S. aureus strains (e.g., COL, Mu50, and 8425-5) suggests that several putative membrane transport systems possessing homology to ATP-binding cassette (ABC)-type iron transporters exist (12, 29, 31, 42–44). In vivo data generated with gene deletion mutant strains suggest that staphylococci rely more on the abundant heme-containing host proteins as iron sources during the critical time of establishing disease, while siderophores are involved in iron acquisition once the bacteria have occupied niches in the host that are devoid of heme proteins (11, 43).

Several staphylococcal surface proteins have been reported to recognize iron-containing host proteins. First, the transferrin binding capacity of S. aureus was shown, and two laboratories associated this activity with two different proteins, GAPDH (glyceraldehyde-3-phosphate dehydrogenase) and SibA/IsdB (30, 45). More recently, IsdB has been suggested to have broad binding specificity, showing in vitro ligand binding activity for extracellular matrix proteins, such as fibrinogen and...
fibronectin, but also for fetuin and holo tranferrin, and even for hemoglobin, by surface plasmon resonance (SPR) analysis (9). It is somewhat controversial that in independent studies, transferrin binding by living S. aureus cells expressing IsaA could not be detected (34) and IsaA did not recognize hemoglobin as a ligand in filter binding assays (29). Our laboratory identified a novel cell wall-anchored protein, HarA, that binds haptoglobin, hemoglobin, and haptoglobin-hemoglobin complexes (15). The orthologous protein, IsdB, was also shown to interact with hemoglobin in a filter binding assay (29). IsaA, HarA (also called IsdH and SasI), and IsdB (also named FrpB and SirH) are iron-Fur-regulated surface proteins with an LPXTG cell wall anchor motif (15, 29, 32, 39, 45). LPXTG motif-containing proteins are covalently attached to the cell wall by the activities of sortases and are implicated in wide range of host-pathogen interactions (28, 41).

Based on fluorescence-activated cell sorter analysis of a gene knockout S. aureus strain, HarA was proven to be the sole haptoglobin-binding surface component of S. aureus (15). Importantly, it was also demonstrated that ligand binding was mediated by two homologous domains, each containing 145 amino acid residues (15). The HarA ligand binding domains largely overlap with the 125-amino-acid-long NEAT (for “near transporter”) domain that is present in different copy numbers in proteins of several gram-positive bacteria, such as Bacillus halodurans, Clostridium perfringens, Streptococcus pyogenes, and several Listeria species. Notably, these proteins are encoded by genes that are in the vicinity of putative Fe³⁺ transporters, and for that reason, the domain was called NEAT (1). harA seems to be an exception, since the gene does not cluster with iron ABC transporter genes on the chromosome, unlike isdB and isdA, which are neighbors of the isdDEF genes (29). IsdB shares 68% identity with HarA in half of its sequence, containing two NEAT domains. However, the single NEAT domain in IsdA has lower homologies to the NEAT domains in HarA and IsdB (20 to 24% identity and 38 to 42% similarity). Members of the Isd family of proteins have been shown to be important for iron metabolism in S. aureus; IsaA, IsdA, -B, -C, -D, -E, and -G bind heme; IsdB and IsdA also bind hemoproteins; IsaB and -G are suggested to remove the heme molecule; and IsdDEF transport heme iron through the cell wall and plasma membrane (29, 44).

Since the abilities of pathogens, especially that of S. aureus, to develop resistance to drug treatment are enormous, a promising alternative approach to the management of staphylocccal diseases is to increase host defenses. Induction of specific antibodies that not only enhance the uptake of S. aureus by professional phagocytic cells, which are often reduced in number and function in hospitalized patients, but also interfere with bacterial growth and potential to cause disease is a viable approach. The best targets for such antibodies are bacterial surface molecules with functions important for the survival of the pathogens in their hosts, such as iron uptake receptors. It has been demonstrated very recently that immunization with IsdB provides protection in animal models mimicking human S. aureus-induced infections (24). IsdB and the closely related HarA, as well as IsaA, antigens were identified in our laboratory using a comprehensive genome-based approach combined with human serology. They all induce high levels of antibodies during human infections, suggesting a high level of in vivo expression (16).

The present study was designed to further characterize the ligand binding specificity and affinity of HarA by surface plasmon resonance measurements. The isolated single N-terminal domain, HarAD1, that was able to bind hemoproteins in the low nanomolar range was subjected to nuclear magnetic resonance (NMR) spectroscopy. The NMR fold determination confirmed the in silico-predicted β-strands and identified a novel β-barrel fold that resembled the previously identified immunoglobulin fold. The functional overlap among proteins containing the NEAT domain was demonstrated by the similar binding affinities of HarA and IsdB for hemoglobin and haptoglobin-hemoglobin proteins, although IsdB lacked direct haptoglobin-binding capacity.

**MATERIALS AND METHODS**

**Cloning and expression of recombinant HarA and IsdB proteins.** Full-length HarA, HarAD1, and HarAD2 domains were generated as described previously (15). cDNA encoding IsdB was amplified from S. aureus strain COL genomic DNA with the gene-specific oligonucleotides AATGGATCCGCCAGTTGAAG AACAGGTTGATACAAA and TATGTGCACTAAGTTGTAATAGA TTTTGCTTTATTTTCTTG, with incorporated BamHI and SalI sites (underlined), respectively. The restriction enzyme-digested PCR product was cloned into the BamHI/SalI-cleaved pGEX4T-3 vector downstream of sequences coding for the glutathione S-transferase (GST) tag (Amersham Biosciences). The resulting gene lacked sequences corresponding to the signal peptide and the C-terminal part downstream from the source cleavage site (LPKT+G). The recombinant protein was purified from bacterial extracts of IPTG (isopropyl-β-D-thiogalactopyranoside)-induced Escherichia coli BL21 cells by sonication in buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA) and purified on a gluthatione-Sepharose 4B affinity column (Amersham Biosciences) from soluble bacterial fractions. Recombinant proteins were eluted either by thrombin digestion (50 U/mL at 3 h room temperature) or with 10 mM glutathione.

**ELISA.** Hyperimmune rabbit sera against HarAD1 and HarAD2 were generated as described previously (15). The enzyme-linked immunosorbent assay (ELISA) was performed according to standard protocols. Briefly, 96-well plates (Maxisorb; NUNC, Denmark) were coated with recombinant proteins (50 μL) diluted in 0.1 M NaHCO₃ buffer (pH 9.3) to a concentration of 5 μg/mL. The plates were incubated overnight at 4°C. After blocking of nonspecific sites with 2% bovine serum albumin in phosphate-buffered saline (PBS), rabbit sera were added at 5,000-, 25,000-, and 125,000-fold dilutions. The plates were incubated for 1.5 h at 37°C. After the plates were washed with PBS containing 0.1% Tween 20, horseradish peroxidase-conjugated donkey anti-rabbit antibodies (Amersham Biosciences, England) were added to each well in 1,000-fold dilution and the plates were incubated for 1 h at 37°C. After the last washing step with PBS containing 0.1% Tween 20, a signal was developed with 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (Sigma) for 30 min at 37°C, and optical densities were measured at 405 nm on an ELISA reader (Tecan Sunrise, Tecan, Switzerland).

**Haptoglobin-binding assays.** The solid-phase binding assay was performed by using the recombinant staphylococcal proteins HarA, HarAD1, HarAD2, and IsdB (without GST tags) for coating ELISA plates at 5 μg/mL concentration in coating buffer (0.1 M sodium carbonate, pH 9.3). After the plates were washed and blocked with 2% bovine serum albumin in PBS, haptoglobin purified from pooled human plasma (Sigma and Fluka) was added in different concentrations (0.1 to 10 μg/well). Binding was detected with polyclonal anti-haptoglobin antibody (Sigma) at 1:1,000 dilution and horseradish peroxidase-labeled anti-rabbit immunoglobulin G (IgG) and quantified by measuring the conversion of the substrate [2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid; Sigma] to the colored product based on optical density readings at 405 nm in an automated ELISA reader (Tecan Sunrise).

**SPR.** The kinetics of interaction between binding partners were analyzed using biomolecular interaction analysis with a biosensor based on SPR at Bioline, Kassel, Germany. Anti-GST antibodies (Biacore AB, Uppsala, Sweden) were coupled covalently to the sensor chip surface using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride/N-hydroxy succinimide) chemistry. The resulting affinity surfaces were used to capture GST-tagged HarAD1 and IsdB proteins. An “empty” flow cell was used as a reference for detection of nonsp-
cific binding and bulk effects. Dilution series of the ligands Hp, Hb, and Hp-Hb complexes were prepared, and binding was measured by injecting between 8 μM and 1 nM ligands in the running buffer (PBs, pH 7.4, 0.005% Tween 20) with injection volumes of 60 to 150 μl at 10 to 20 μl/min. Binding curves were acquired for every concentration by first capturing the GST-tagged protein on the surface and then incubating it with ligand solution. Dissociation was initiated by switching to the running buffer without the ligand. The signals were measured with a Biacore 3000 instrument (Biacore AB, Uppsala, Sweden).

NMR measurements and data processing. All NMR experiments were performed on 500-MHz and 800-MHz Varian UNITY INOVA spectrometers equipped with pulsed-field gradient unit triple-resonance probes with actively shielded z gradients. The sample was kept in a 250-μl Shigemi tube and contained approximately 0.8 mM uniformly 13C/15N-labeled HarAD1 in 20 mM sodium phosphate buffer, pH 6.5, containing 0.5 mM EDTA, 0.5 mM dithiothreitol in H2O/D2O (9:1). All spectra were recorded at 25°C. The following experiments were used in the present study for spin system identification and sequential assignment: two-dimensional transverse-relaxation-optimized (36)15N HSQC, three-dimensional (3D) HNCA, 3D HN(CO)CA, 3D HNCACB, and 3D CBCA(CO)NH, as well as 2H- and 13C-edited 3D nuclear Overhauser effect spectroscopy (NOESY)-HSQC (reference 8 and references therein). Experiments involving amide proton detection included pulsed-field gradients for coherence transfer pathway selection and making use of an enhanced-sensitivity approach (reference 8 and references therein). Spectra were processed using NMRPipe software (13) and visualized and assigned using NMRView (20) software. The number of data points was doubled by linear prediction in the indirect dimensions prior to zero filling and Fourier transformation.

Input restraints. Interproton distance restraints were derived from 1H- and 15N-edited NOESY-HSQC spectra, in which each assigned NOE translates into one distance restraint employed in subsequent structure calculations. Since a highly resolved NMR solution structure was beyond the scope of this study, NOE peak volumes were not integrated; however, for the sake of simplicity, upper and lower bounds of NOE-derived distance restraints were uniformly set to 1.8 and 5.0 Å, respectively. Redundant, as well as meaningless, intraresidue restraints were omitted from the input list. Backbone dihedral-angle restraints (Φ and Ψ) were obtained by employing TALOS software (10), that is, a database system for secondary structure prediction: two-dimensional transverse-relaxation-optimized 15N HSQC, three-dimensional (3D) HNCA, 3D HN(CO)CA, 3D HNCACB, and 3D CBCA(CO)NH, as well as 2H- and 13C-edited 3D nuclear Overhauser effect spectroscopy (NOESY)-HSQC (reference 8 and references therein). Experiments involving amide proton detection included pulsed-field gradients for coherence transfer pathway selection and making use of an enhanced-sensitivity approach (reference 8 and references therein). Spectra were processed using NMRPipe software (13) and visualized and assigned using NMRView (20) software. The number of data points was doubled by linear prediction in the indirect dimensions prior to zero filling and Fourier transformation.

Input restraints. Interproton distance restraints were derived from 1H- and 15N-edited NOESY-HSQC spectra, in which each assigned NOE translates into one distance restraint employed in subsequent structure calculations. Since a highly resolved NMR solution structure was beyond the scope of this study, NOE peak volumes were not integrated; however, for the sake of simplicity, upper and lower bounds of NOE-derived distance restraints were uniformly set to 1.8 and 5.0 Å, respectively. Redundant, as well as meaningless, intraresidue restraints were omitted from the input list. Backbone dihedral-angle restraints (Φ and Ψ) were obtained by employing TALOS software (10), that is, a database system for empirical prediction of Φ and Ψ angles using a combination of five kinds (1Hα, 1Hβ, 13Cα, 13Cβ, and 15N) of chemical shift assignments for a given protein sequence. According to the empirical “rules of thumb” implemented in the TALOS software, for 100 HarAD1 residues, the Φ and Ψ predictions were classified to be unambiguous and were consequently used as dihedral-angle restraints in subsequent structure calculations. Careful analysis of NOE patterns in the 1H- and 15N-edited NOESY-HSQC spectrum and of 1Hα, 1Hβ, 13Cα, and 13Cβ secondary chemical shifts allowed the definition of 63 hydrogen bonds in elements of regular secondary structure. For each hydrogen bond, two distance restraints were defined: the first was between the backbone amide hydrogen and the carbonyl oxygen, and the second was between the amide nitrogen and the carbonyl oxygen, corresponding to the limits of 1.8 to 2.5 Å and 2.8 to 3.5 Å, respectively.

Determination of ligand binding affinities of the HarAD1 domain by surface plasmon resonance. Previously, we used an ELISA-based binding assay to demonstrate haptoglobin, hemoglobin, and haptoglobin-hemoglobin binding by HarA and its subdomains, designated HarAD1 and HarAD2 (15). The two domains, each consisting of 145 amino acid residues, share 53% identity and 72% similarity in their primary sequences. These domains also share biological functions, and both bind Hp and Hb-Hb complexes, as was demonstrated with recombinant HarAD1 and HarAD2 proteins (15). We found that antibodies generated with recombinant HarAD1 recognized HarAD2 very efficiently, and conversely, HarAD2-specific antibodies showed...
high reactivity with HarAD1 protein in ELISA (Fig. 1A) and Western blot (not shown) analyses. Importantly, the antibodies generated with one domain inhibited the function of the other, as well as that of the full-length protein, suggesting the involvement of conserved amino acid residues in ligand binding (Fig. 1B).

In order to characterize the binding of HarA to its ligands more accurately, we employed the SPR technology, which allows the sensitive detection of molecular interactions in real time and the determination of the association and dissociation rate constants. For the SPR analysis, the 16-kDa ligand binding domain was used, since the 100-kDa full-length recombinant HarA could not be generated without smaller products. GST-tagged HarAD1 and IsdB proteins were immobilized on an anti-GST antibody-coated surface, and the three ligands, Hp, Hb, and Hp-Hb complexes, were added in increasing concentrations. Haptoglobin binding to HarAD1 showed a complex nature with pronounced involvement of more than one component in the binding affinity. Apart from the fast increase in binding within the first seconds (the first binding event), a slower association was observed within the next 900 seconds, suggesting secondary binding events (Fig. 2A). This additional component could be due to a second (or even more) binding site, oligomerization of the ligand, or a conformational change within one of the binding partners upon interaction. Given that the natures of the two binding events were not clear, only the stable slow dissociation was considered for the calculation of the equilibrium dissociation constant ($K_D$), and it was deter-

FIG. 2. SPR measurements with HarAD1 and haptoglobin, hemoglobin, and Hp-Hb complexes. Data were generated by Biacore measurements using GST-HarAD1-coated surfaces and (A) haptoglobin, (B) hemoglobin, and (C) haptoglobin-hemoglobin complexes in the indicated concentration ranges. Binding curves were acquired by first capturing the GST-HarAD1 protein on the surface and then incubating it with ligand solution at the indicated ligand concentrations. The data are presented as real-time graphs of response units (RU) against time (left) and as SPR signals against ligand concentration in the late dissociation phase (right).
mined to be in the low μM range. Using the lowest molecular masses of the haptoglobin molecule consisting of one beta (27 kDa) and one alpha (9 and 16 kDa for α1 and α2; both phenotypes were present in the preparation) subunit, the $K_D$ was determined to be ~5 μM, while calculations with the high-molecular-mass-oligomer suggested a value for $K_D$ of 0.5 to 1 μM.

We also determined the affinity for Hb by Biacore analysis and measured a stronger interaction of HarAD1 with hemoglobin ($K_D = 30$ nM) relative to Hp (Fig. 2B) that was not anticipated from previous ELISA-based assays (15). Using Hb as ligand, the binding curves suggested that only one binding event occurred, since after fast association, the binding reached equilibrium. The SPR results generated with Hp-Hb complexes confirmed our previous observation that HarA bound this ligand with the highest affinity, with a $K_D$ in the low nanomolar range (Fig. 2C) ($K_D = 5$ nM). These binding curves, like the ones detected for HarAD1-haptoglobin interaction, also showed strong biphasic kinetics and suggested a complex interaction with the ligand, most likely due to oligomerization of Hb.

NMR fold determination of the ligand binding domain of HarA. The recombinant HarAD1 protein was labeled uniformly by incorporating $^{13}$C and $^{15}$N amino acids and analyzed by NMR spectroscopy. The NMR spectra, in particular, the good dispersion of $^{15}$N HSQC-detected cross peaks, indicated that HarAD1 adopted a well-defined structure in aqueous solution. For the 137 nonproline residues of HarAD1, signals for 133 backbone amide functions were observed in the two-dimensional $^{15}$N HSQC spectrum. For four backbone groups, a signal was either not seen, due to fast exchange with bulk water and/or rapid conformational change, or could not be unambiguously identified because of signal overlap, i.e., shift degeneracy of backbone amide moieties in both the $^1$H and $^{15}$N dimensions.

Sequential backbone signal assignment was primarily accomplished by combined analyses of a series of 3D triple-resonance spectra [HNCA, HN(CO)CA, HN(CACB), and CBCA(CO)NH] (reference 8 and references therein) and was further supported by observation of sequential nuclear Overhauser effects (NOEs) in a $^{15}$N-edited NOESY-HSQC spectrum. This strategy allowed the complete HN, N, Cα, and Cβ resonance assignment of all $^{15}$N-HSQC-detected residues. Nearly complete (89% of the HarAD1 sequence) Hα assignment was subsequently achieved based on $^{13}$C- and $^{15}$N-edited NOESY data, as well as knowledge of sequential Cα chemical shifts. Six asparagines and two glutamine side chain amides were detected and unambiguously assigned in HNCA, HN(CO)CA, and $^{15}$N-edited NOESY spectra. The indolic NH group of the only HarAD1 Trp residue could be clearly discerned by its presence in the $^{15}$N-HSQC and $^{15}$N-NOESY-HSQC spectra and its absence in triple-resonance spectra. Since we were interested in a fold determination of HarAD1 rather than a refined NMR solution structure, experiments devised for complete side chain assignment, such as HCCH-TOCSY, were not performed.

The number and type of secondary-structure elements and their succession along the primary sequence were determined on the basis of specific cross-peak patterns observed in the 3D $^{15}$N-NOESY-HSQC. The secondary structure was also reflected by secondary ΔHα/ΔCα/ΔCβ shifts (47) (i.e., the difference in chemical shift between experimental Hα/Cα/Cβ and random coil values). There was a good agreement between secondary-shift values and the assignment of secondary-structure elements from NOE patterns along the sequence. This analysis revealed that HarAD1 contained two α-helices and eight β-strands. Our experimentally determined secondary structure corresponds well with the secondary-structure prediction previously reported by Andrade et al. (1).

Strand pairing could be unambiguously identified by assigning numerous interresidue NOEs arising from backbone HN and Hα protons. In total, 354 nonredundant interresidue HN/HN and HN/Hα, as well as Hα/Hα, NOEs were assigned for all experimentally observed HarAD1 residues. When displayed in a correlation plot, these NOEs clustered with specific patterns that were indicative of antiparallel strand pairing (Fig. 3). Each HarAD1 β-strand was paired on both sides with another strand in an antiparallel manner.

The NMR fold of HarAD1 was determined in an iterative procedure by using restrained molecular dynamics and simulated annealing as described in Materials and Methods. In the last structure calculation run, a family of 100 HarAD1 models was calculated by employing a final data input set with a total of 693 experimental NMR restraints (giving an average of 5 restraints per residue), comprising 367 nonredundant, interresidue NOE-derived distance restraints; 200 backbone dihedral-angle restraints; and 126 hydrogen bond restraints. Of these 100 models, the 25 best with respect to restraint violation and total energy were used for structure analysis.

HarAD1 topology. A schematic representation of the HarAD1 topology, as well as a ribbon drawing of one representative NMR-derived HarAD1 model, is shown in Fig. 4A and B. The HarAD1 protein featured an eight-stranded, purely antiparallel β-barrel with the strand order (-β1, -β2, -β3, -β6, -β5, -β4, -β7, -β8 -) (β1, residues Q111 to D118; β2, S130 to T140; β3, K143 to I150; β4, K153 to E163; β5, V170 to
The structural model of HarAD1 was compared to all protein structures deposited in the Protein Data Bank as described in Materials and Methods. Its topology resembled that of the Ig fold family, whose members share a common structural motif of a seven- or eight-stranded β-sandwich (18). The members of this fold are highly diverse in both sequence and structure, and most of the proteins are involved in protein-protein interactions. According to the SCOP classification, the Ig fold superfamilies with the highest similarity to the HarAD1 protein were the Cu,Zn superoxide dismutases and the clathrin adaptor appendage domains. Figure 5 shows the structural superimposition of HarAD1 with its closest structural neighbor from the PDB: human Cu,Zn dismutase (PDB entry code, 1fm) (17). The latter, although assigned to the immunoglobulin fold family in the SCOP data bank, has structural features of both a β-sandwich and a β-barrel. While a β-barrel forms a closed structure in which the first strand of a single β-sheet comes around to hydrogen bond to the last strand, a stereotypical β-sandwich consists of two pleated β-sheets packed against each other.

**TABLE 1. Structural statistics**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental restraints</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>693</td>
</tr>
<tr>
<td>Interresidue, sequential ((i - j = 1))</td>
<td>207</td>
</tr>
<tr>
<td>Interresidue, medium range ((2 \leq i - j \leq 4))</td>
<td>42</td>
</tr>
<tr>
<td>Interresidue, long range ((i - j &gt; 4))</td>
<td>118</td>
</tr>
<tr>
<td>Dihedral-angle restraints ((F,Y))</td>
<td>200</td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>126</td>
</tr>
<tr>
<td>Average RMSD from experimental restraints</td>
<td></td>
</tr>
<tr>
<td>Distance restraints ((Å))</td>
<td>0.112 ± 0.017</td>
</tr>
<tr>
<td>Dihedral-angle restraints ((degree))</td>
<td>0.264 ± 0.028</td>
</tr>
<tr>
<td>Average RMSD from idealized covalent geometry</td>
<td></td>
</tr>
<tr>
<td>Bond ((Å))</td>
<td>0.0037 ± 0.0004</td>
</tr>
<tr>
<td>Angles ((degree))</td>
<td>0.58 ± 0.06</td>
</tr>
<tr>
<td>Improper ((degree))</td>
<td>0.59 ± 0.07</td>
</tr>
<tr>
<td>CNX energies</td>
<td></td>
</tr>
<tr>
<td>Overall ((kcal mol^{-1}))</td>
<td>790 ± 84</td>
</tr>
<tr>
<td>Backbone atomic RMSD ((Å)^a)</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>2.65 ± 0.74</td>
</tr>
<tr>
<td>Barrel, including loops</td>
<td>1.71 ± 0.24</td>
</tr>
<tr>
<td>Barrel, strands only</td>
<td>0.93 ± 0.21</td>
</tr>
</tbody>
</table>

**FIG. 4.** HarAD1 adopts a β-barrel fold. (A) Topology diagram of HarAD1. The eight-stranded antiparallel β-barrel harbors two four-stranded Greek key motifs that mutually overlap with three strands. (B) Ribbon plot of an averaged HarAD1 model calculated from the 25 lowest-energy models. The plot was generated by fitting heavy atoms of residues located in strands \(β1\) to \(β8\). Helix \(α1\) adopts random orientations relative to the backbone atoms of residues located in strands \(β1\) to \(β8\). Helix \(α1\) was preceded by a helix, \(α1\), comprising residues L90 to D96. Another short helix, \(α2\), consisting of a single turn (residues Y125 to F128), was inserted between strands \(β1\) and \(β2\). The precise orientation of helices \(α1\) and \(α2\) relative to the barrel is not yet determined.

**Structural-similarity searches.** The structural model of HarAD1 was compared to all protein structures deposited in the Protein Data Bank as described in Materials and Methods. Its topology resembled that of the Ig fold family, whose members share a common structural motif of a seven- or eight-stranded β-sandwich (18). The members of this fold are highly diverse in both sequence and structure, and most of the proteins are involved in protein-protein interactions. According to the SCOP classification, the Ig fold superfamilies with the highest similarity to the HarAD1 protein were the Cu,Zn superoxide dismutases and the clathrin adaptor appendage domains. Figure 5 shows the structural superimposition of HarAD1 with its closest structural neighbor from the PDB: human Cu,Zn dismutase (PDB entry code, 1fm) (17). The latter, although assigned to the immunoglobulin fold family in the SCOP data bank, has structural features of both a β-sandwich and a β-barrel. While a β-barrel forms a closed structure in which the first strand of a single β-sheet comes around to hydrogen bond to the last strand, a stereotypical β-sandwich consists of two pleated β-sheets packed against each other.
The X-ray structure of human superoxide dismutase (at 1.02-Å resolution) reveals that the enzyme consists of a single eight-stranded antiparallel β-sheet adopting the shape of a partly closed barrel. Its peripheral strands (β5 and β6) are close enough for interstrand hydrogen bond formation at one end but diverge from each other toward the opposite end so that hydrogen bond formation is impeded. According to the present study, HarAD1 features a closed β-barrel; therefore, we consider HarAD1 to be a new variant of the Ig fold that has not yet been observed in other proteins. The RMSD values of β-strand heavy backbone atoms of the human superoxide dismutase and that of a representative NMR-derived model of HarA were determined to be 3.1 Å.

The orthologous staphylococcal protein IsdB does not bind haptoglobin but recognizes hemoglobin and Hp-Hb complexes as its ligands. The orthologous cell wall protein of *S. aureus*, IsdB, shares substantial homology with HarA, especially in a 340-amino-acid-long region that harbors the second ligand binding domain of HarA (Fig. 6). In silico analysis of the ligand binding domain sequences in HarA and IsdB suggested similar secondary structures with predicted β-strands in good alignment (1, 15). Importantly, IsdB has already been reported by

---

**FIG. 5.** Structural similarity of HarAD1 to known structures. Shown is a superimposition of the HarAD1 protein with the most similar structure found in the latest version of the Protein Data Bank: human Cu,Zn superoxide dismutase (PDB code, 1mfm). The HarAD1 protein is displayed in yellow, with superoxide dismutase in magenta. The eight strands of both proteins are in structurally equivalent positions and have the same sequential order. In HarAD1, all eight strands build a closed barrel, whereas the dismutase structure forms a barrel-like β-sandwich with hydrogen bonds missing between its peripheral strands. The figure was prepared using the PyMol program.

**FIG. 6.** Prediction of the ligand binding surface in HarA. (A) Structural and sequence homologies of HarA, IsdB, and IsdA NEAT domains. The schematic drawing shows the locations of the ligand binding domains (15) and predicted NEAT domains (1) by amino acid residue numbers, with the initiating methionine assigned to 1. SP, signal peptide, dotted bar; LPXTG cell-sorting signal, black bar; hydrophobic transmembrane domain, hatched horizontal line; region of high similarity (80%), hatched decreasing line. (B) Multiple-sequence alignment of the ligand binding domains of HarA, IsdB, and IsdA. The sequences were derived from the COL strain genomes SACOL1781, SACOL1138, and SACOL1140 (http://cmr.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi) and analyzed with ClustalW (http://www.ebi.ac.uk/clustalw/). If at least three amino acids are similar, they are colored according to their physicochemical properties as follows: ILVAM aliphatic/hydrophobic residues, yellow; FWY aromatic, light blue; KRH positively charged, blue; DE negatively charged, gray; STNQ hydrophilic residues, green; PG conformationally special, violet. In addition, conserved, solvent-exposed amino acids are highlighted in red.
indicated ligand concentrations. The data are presented as real-time protein on the surface and then incubating it with haptoglobin at the surfaces. Binding curves were acquired by first capturing the GST-IsdB resonance signals were detected on anti-GST antibody-coated chip coated with recombinant proteins, and binding to purified ligands was detected by monoclonal anti-haptoglobin antibodies. OD405, optical density at 405 nm. (B) Surface plasmon resonance signals were detected on anti-GST antibody-coated chip surfaces. Binding curves were acquired by first capturing the GST-IsdB protein on the surface and then incubating it with haptoglobin at the indicated ligand concentrations. The data are presented as real-time graphs of response units (RU) against time.

Mazmanian et al. to bind hemoglobin, based on filter binding assays (29).

To determine whether HarA and IsdB, with their orthologous domains, share evolutionarily conserved functions, we determined the haptoglobin binding of the IsdB protein compared to those of HarA, HarAD1, and HarAD2 in the previously developed solid-phase binding assay. ELISA plates were coated with recombinant proteins, and binding to purified ligands was detected by monoclonal anti-haptoglobin antibodies. Full-length HarA and the subdomains HarAD1 and HarAD2 captured haptoglobin; however, IsdB failed to bind even at high ligand concentrations (Fig. 7A). Lower-affinity binding of haptoglobin by the HarAD2 domain relative to that of HarAD1 was detected with high consistency. The lack of binding to haptoglobin observed for IsdB in the ELISA-based assay was confirmed by SPR measurements using GST-tagged full-length IsdB protein (Fig. 7B), since no signal was detected even with the highest concentration of haptoglobin (8 μM).

However, IsdB showed a high affinity for hemoglobin and haptoglobin-hemoglobin complexes similar to that observed with HarAD1 in SPR measurements, with calculated \( K_{D} \) of 55 nM and 6 nM, respectively (Fig. 8). Due to the complex nature of the binding with haptoglobin-hemoglobin complexes also observed with HarAD1, only the more stable, more slowly binding event was used for affinity calculations.

DISCUSSION

In this study, we have characterized two S. aureus surface proteins, HarA and IsdB, by measuring their ligand binding affinities with surface plasmon resonance and determining the fold of the homologous ligand binding domain of HarA by NMR.

To further define the biochemical mechanism involved in hemoglobin utilization by S. aureus, we have characterized the binding interactions of HarA and IsdB with three different ligands: Hb, Hb, and Hp-Hb complexes, by Biacore analysis. We found that HarAD1, encompassing the first predicted NEAT domain of HarA, as well as full-length IsdB, containing two NEAT domains, bound hemoglobin and Hp-Hb complexes with comparable high affinities, with \( K_{D} \) of \( \sim 10^{-8} \) M and \( \sim 10^{-9} \) M, respectively. However, haptoglobin-binding activity could be demonstrated only with HarA and its subdomains. This is quite interesting, since both HarAD1 and HarAD2 proteins, which contain the first and second NEAT domains of HarA, respectively, recognize haptoglobin as their ligand (reference 15 and this study), yet HarAD2 shows a higher level of homology to IsdB than to HarAD1 (6 versus 53% identity, respectively). Antibodies induced by HarAD1 efficiently recognized HarAD2 in ELISA, and vice versa, and cross-inhibited ligand binding, suggesting that the amino acid residues important for recognition of all three ligands are conserved between HarAD1 and HarAD2. Our data suggest that in spite of the higher homology between HarAD2 and IsdB, critical amino acids for haptoglobin binding are not conserved in IsdB. IsdA was also tested for haptoglobin binding and was found to be negative (9). These biochemical data strongly support our previous results with an isogenic single mutant harA strain that lost haptoglobin binding, suggesting that HarA is the only Hp receptor in S. aureus (15).

Our attempt to examine in vitro the involvement of HarA and IsdB in iron acquisition by using single- and double-knockout strains has so far been unsuccessful. A likely explanation is that hemoglobin as a sole iron source in iron-depleted growth medium supported bacterial growth very efficiently, independent of the presence of the harA and isdB genes, most likely due to the dominant activity of the siderophore-dependent iron acquisition system (reference 15 and unpublished data). The individual contributions of HarA and IsdB to iron acquisition by S. aureus through hemoprotein binding are likely to largely depend on their expression during disease and their surface densities, since they seem to have comparable high affinities for Hb and Hp-Hb. However, since we have determined the binding affinity of only one of the ligand binding domains of HarA by surface plasmon resonance measurements, it is likely that full-length HarA, containing three NEAT domains, has higher affinity than full-length IsdB, with two of them. ELISA-based binding assays consistently show a rank order of “full-length” HarA > HarAD1 > HarAD2 for the three ligands. IsdA seems to possess broader-specificity binding to host proteins with lower affinity for hemoglobin (\( K_{D} \) \( \sim 0.75 \) μM) than HarA and IsdB (9).

Hemoglobin, in its free form, is not directly available for invading staphylococci unless massive hemolysis is induced by cytotoxins or Hp-deficient patients are infected, since released Hb is immediately and almost irreversibly complexed by the
abundant plasma protein haptoglobin through an extremely strong interaction with $K_{D}$ in the picomolar range (19). Thus, Hp-Hb complexes are the most physiologically relevant iron source during invasive staphylococcal disease.

What could be the advantage for *S. aureus* in having HarA bind haptoglobin in addition to hemoglobin and haptoglobin-hemoglobin complexes? First, an obvious biochemical advantage could be higher-affinity binding. The binding curves obtained with haptoglobin and Hp-Hb complexes always showed a complex nature, suggesting more than one component in the binding event. It is well known that haptoglobin exists in oligomeric forms, and this feature is likely to contribute to the phenomenon. However, the slow dissociation of Hp-Hb binding was most obvious when we used HarAD1 as bait, suggesting that Hp binding contributes to longer-lasting interaction between the receptor and the Hp-Hb ligand. Secondly, one can envision that the most efficient way for *S. aureus* to acquire iron is the capture of hemoglobin immediately released from erythrocytes due to staphylococcal hemolysins produced by *S. aureus* cells coated with haptoglobin, which has an approximately $10^{6}$ times higher affinity for Hb than HarA or IsdB and forms an almost irreversible interaction. Further, being covered with haptoglobin, which is always highly abundant in plasma and tissue fluids, might provide *S. aureus* with a means to hide from the host immune system. It has also been suggested that haptoglobin has multiple physiological functions, such as protection of tissues against oxidative damage, especially that related to the presence of ferric iron, and the regulation of phagocytic activity (reviewed in references 14 and 40). Therefore, it could be beneficial for *S. aureus* to ensure high concentrations of Hp while inside phagocytic cells in order to survive by protection from oxidative damage.

We do not know at present which of these possibilities are operational and have the most important roles in staphylococcal pathogenesis. However, it is has very recently been shown, in an animal model that mimics human invasive staphylococcal infection, that IsdB protein is protective, highlighting its potential as a viable vaccine candidate antigen (24) and a possible target for therapeutic monoclonal antibody treatments.

Based on the data presented here and data from other laboratories (9, 29), it is obvious that the NEAT domain is highly relevant for the interaction of *S. aureus* with its host. Andrade et al. compared the sequences of predicted NEAT domains in available genomes of gram-positive bacteria and suggested that in spite of the high divergence in amino acid sequences, the predicted secondary structures are very similar, dominated by seven β-strands (1). In order to determine the fold of this evolutionarily conserved ligand binding domain, the untagged HarAD1 protein was subjected to NMR spectroscopy, which revealed an eight-stranded, purely antiparallel β-barrel structure with the strand order ($\beta_{1}$–$\beta_{2}$–$\beta_{3}$–$\beta_{6}$–$\beta_{5}$–$\beta_{4}$–$\beta_{7}$–$\beta_{8}$), forming two Greek key motifs with strands $\beta_{3}$ to $\beta_{6}$ and $\beta_{4}$ to $\beta_{7}$. The positions of the experimentally defined β-sheets correspond well with the predicted ones. Based on the results of structural-homology searches, the topology of the HarAD1 domain resembles that of the Ig fold family, whose
members are involved in protein-protein interactions. However, due to distinct structural features, we consider that the HarAD1/NEAT domain is a novel variant of the Ig fold that has not yet been observed in other proteins. From a functional point of view, it is difficult to make inferences from the observed similarities, since the Ig folds detected in proteins are highly diverse in amino acid sequence and function. Moreover, the similarity that results from the structure superimposition is marginal. Nevertheless, the fact that many proteins adopting an Ig fold are involved in protein-protein interaction is in agreement with the experimental observations for the NEAT domains presented in this paper and elsewhere. To our knowledge, we have determined for the first time a NMR-based structure model of a pathogen-derived protein domain that is capable of interacting with haptoglobin and hemoglobin.

Our data from this study, together with the immune protection shown with IsdB (1), suggest an attractive new strategy against the group of pathogens containing NEAT domain proteins in general and against S. aureus in particular.

REFERENCES


48. Reference deleted.