Crystal Structure of the N-terminal Domain of the Group B Streptococcus Alpha C Protein

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 280, No. 18, Issue of May 6, pp. 18245–18252, 2005
Printed in U.S.A.

Crystal Structure of the N-terminal Domain of the Group B *Streptococcus* Alpha C Protein*

Received for publication, November 2, 2004, and in revised form, February 25, 2005
Published, JBC Papers in Press, March 6, 2005, DOI 10.1074/jbc.M412391200

Thierry C. Aupérin‡§¶, Gilles R. Bolduc§‖, Miriam J. Baron‖, Annie Heroux, David J. Filman‡, Lawrence C. Madoff‖, and James M. Hogle‡ ‡‡**

*From the ‡Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, the ‖Brigham and Women's Hospital, Channing Laboratory, Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, and the **Biology Department, 463, Brookhaven National Laboratory, Upton, New York 11973-5000*

Group B *Streptococcus* (GBS) is the leading cause of bacterial pneumonia, sepsis, and meningitis among neonates and an important cause of morbidity among pregnant women and immunocompromised adults. Invasive diseases due to GBS are attributed to the ability of the pathogen to translocate across human epithelial surfaces. The alpha C protein (ACP) has been identified as an invasin that plays a role in internalization and translocation of GBS across epithelial cells. The soluble N-terminal domain of ACP (NtACP) blocks the internalization of GBS. We determined the 1.86-Å resolution crystal structure of NtACP comprising residues Ser^{52} through Leu^{225} of the full-length ACP. NtACP has two domains, an N-terminal β-sandwich and a C-terminal three-helix bundle. Structural and topological alignments reveal that the β-sandwich shares structural elements with the type III fibronectin fold (FnIII), but includes structural elaborations that make it unique. We have identified a potential integrin-binding motif consisting of Lys-Thr-Asp^{146}, Arg^{110}, and Asp^{118}. A similar arrangement of charged residues has been described in other invasins. ACP shows a heparin binding activity that requires NtACP. We propose a possible heparin-binding site, including one surface of the three-helix bundle, and nearby portions of the sandwich and repeat domains. We have validated this prediction using assays of the heparin binding and cell-adhesion properties of engineered fragments of ACP. This is the first crystal structure of a member of the highly conserved Gram-positive surface alpha-like protein family, and it will enable the internalization mechanism of GBS to be dissected at the atomic level.

Group B *Streptococcus* (GBS)[1] (*Streptococcus agalactiae*) remains the leading cause of invasive bacterial diseases in neonates, despite its decline in prevalence during the last decade because of intrapartum chemoprophylatic therapy (1, 2). It is also an important cause of morbidity in pregnant women and non-pregnant adults with underlying medical conditions (3–5). GBS colonizes the human gastrointestinal and genitourinary tracts and may cause chorioamnionitis and urinary tract infection in pregnant women and a range of invasive infections in elderly and immunocompromised adults (1, 6–8). During labor and delivery, GBS may be transmitted to neonates, causing pneumonia, sepsis, or meningitis. Four to six percent of all neonatal GBS infections result in death (6, 9). *In vitro*, GBS adheres to (10, 11), internalizes within (12–14), and translocates across (15) intact human epithelial and endothelial cells. Little is known about the bacterial components that allow this pathogen to adhere to and penetrate cellular membranes. Previous studies suggest that surface proteins are significantly involved in the process (11, 16). The surface-expressed GBS alpha C protein (ACP) has been shown to act as an invasin (15). ACP is the prototype of a family of surface-expressed proteins containing long tandem repeats (alpha-like proteins (Alp)). Members of this family of proteins have a high degree of sequence homology with each other and are thought to share similar function. Indeed, the Alp expressed in group A *Streptococcus*, R28, has been associated with cell adhesion (17) and that of *Enterococcus faecalis*, Esp, has been associated with virulence (18). Deletion of the ACP gene *bca* attenuates the virulence of GBS 5–7-fold and the internalization of GBS into cervical epithelial cells by 80% (19). Furthermore, the amount of internalized GBS within human cervical epithelial cells *in vitro* is inhibited by 75% in the presence of the N-terminal domain of ACP (NtACP) (15), suggesting that ACP is a major determinant of virulence. Recently, ACP has been reported to bind glycosaminoglycans (GAG) (20). ACP consists of an N-terminal domain (174 amino acids), a variable number of tandem repeats of 82 amino acids each, and a 45-amino acid C-terminal domain containing a LPXTG peptidoglycan-binding motif. NtACP mediates GBS internalization within human epithelial cells (15). NtACP in association with a repeat domain is necessary to bind GAG (20). Moreover, both the alpha C protein and the isolated N-terminal domain are immunogenic and elicit antibodies that protect against GBS infections in experimental animals (21, 22).

* This research was supported by Public Health Services Grant AI38424 (to L. C. M.) and the William Randolph Hearst Fund (to G. R. B.). 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 1YWM) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Both authors contributed equally to this work.

¶ Current address: Dept. of Biological Sciences, Columbia University, 703 Fairchild Center, M.C. 2452, 1212 Amsterdam Ave., New York, NY 10027.

‡‡ To whom correspondence should be addressed: Harvard Medical School, Dept. of Biological Chemistry and Molecular Pharmacology, 240 Longwood Ave., Boston, MA 02115. Tel.: 617-432-3918; Fax: 617-432-4360; E-mail: jhogle@hms.harvard.edu.

[1] The abbreviations used are: GBS, group B *Streptococcus*; ACP, alpha C protein; NtACP, N-terminal domain of ACP; Alp, alpha-like protein; GAG, glycosaminoglycan; SeMet, selenomethionine(yl); 9RR, 9-repeat region; BR, binding region; FnIII, type III fibronectin; PBS, phosphate-buffered saline.

To provide further clues concerning the role of NtACP in adherence, entry, virulence, and immunity, we have determined its three-dimensional structure at 1.86 Å resolution by x-ray crystallography. The mostly negative surface of NtACP includes a regular linear arrangement of positive charges that strongly suggested the location of the heparin-binding site and provided preliminary indications regarding their detailed intermolecular interaction. To confirm these structurally based predictions and define the limits of the binding site more precisely, we have conducted *in vitro* and *in vivo* binding assays using engineered fragments of ACP. This is the first crystal structure of a member of the highly conserved Gram-positive surface alpha-like protein family, and it will enable the internalization mechanism of GBS to be further dissected at the atomic level.

MATERIALS AND METHODS

All amino acid residues in this paper and in the coordinates file are numbered according to the alpha C protein numbering (23). Thus, the first methionine of the recombinant proteins is numbered Met^{39}, as the first residue of the alpha C protein[2] N-terminal domain in the GenBank™ data base is Ser^{52}.

Cloning and Site-directed Mutagenesis of NtACP—Cloning of the DNA encoding for the N-terminal domain of alpha C protein (pDEK14) has been described previously (22). Two leucine residues were substituted with methionine residues (L88M and L94M) using the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. Forward primer 5′-GATTTATATGATGTAAAAATGGGTAAAATAGATCCAATGCAATTAATTGTTTTAG-3′, reverse primer 5′-CTAAAACAATTAATTGCATTGGATCTATTTTACCCATTTTTACATCATATAAATC-3′, and pDEK14 template were used. The resulting construct, pGB1, was cloned into *Escherichia coli* BL21(DE3) for expression of the mutant form of NtACP.

Based on the structure of NtACP, the DNA encoding for D2-R was produced using PCR forward primer 5′-CACCTTGAGGGATAAGATT-3′ and reverse primer 5′-TCATAGTTTATTTCCTTTACC-3′ to amplify a 504-base-pair portion of ACP including approximately one-third of the NtACP (*i.e.* Leu^{164}–Leu^{225}, most of domain 2), an adjacent repeat, and 13 amino acid residues of the C-terminal region from plasmid pT7LM39 (24). The forward primer includes a CACC sequence, and the reverse primer includes a TCA stop codon added as directed for use with the pET200 Directional TOPO Expression Kit (Invitrogen). The PCR product was cloned into the pET200/D-TOPO vector using the manufacturer's instructions, and the resulting construct was transformed into *E. coli* BL21(DE3).

Sample Preparation—The native and mutant recombinant NtACP were expressed in *E. coli* BL21(DE3) (Novagen) with a His_6 tag fused at their C-terminal end and purified as previously described (15). Selenomethionyl (SeMet) proteins were expressed overnight in *E. coli* BL21(DE3) at 25 °C following a protocol previously described for the use of non-auxotrophic cells (25). SeMet mutant proteins were purified in the presence of 0.05% β-mercaptoethanol. Complete substitution of methionine residues and loss of the N-terminal methionine in native and SeMet-labeled proteins were confirmed by mass spectrometry (matrix-assisted laser desorption ionization time-of-flight). Final preparations were dialyzed into 20 mM HEPES, pH 7.2, 10 mM dithiothreitol.

E. coli BLR (Novagen) was used to express full-length 9-repeat ACP, as well as the 9-repeat region (9RR, without N- or C-terminal regions). Full-length 9-repeat ACP, 9RR, and D2-R were expressed and purified as described previously (15, 24).

Crystallization—Crystals were grown by hanging drop vapor diffusion method at 21 °C using a 2-μl drop. Native and SeMet mutant proteins (18 mg/ml) were mixed separately with an equal volume of the well solution containing 100 mM sodium acetate, pH 4.6–5.1, and 10–15% (w/v) polyethylene glycol 4000 (well solution for SeMet proteins included 10 mM dithiothreitol). The resulting elongated crystals (typical dimensions 0.6 × 0.3 × 0.3 mm for the native and 2.0–3.0 × 0.2 × 0.2 mm for the SeMet derivative) grow in space group $P6_122$ with unit cell dimensions of $a = b = 56.5$ and $c = 271.6$ Å for the native and $a = b = 56.7$ and $c = 272.2$ Å for the SeMet derivative (see Table I). These crystals had 1 NtACP molecule/asymmetric unit (V_m = 2.88 Å^3/Da) and a solvent content of 57%. In advance of cryocooling the crystals, the native crystals were quickly soaked in 100 mM sodium acetate, 10% (w/v) polyethylene glycol 4000, and 25% glycerol, whereas the SeMet crystals were cryoprotected by bringing the mother liquor to a final concentration of 20% glycerol in small steps (3% increments every 30 min), and were frozen in liquid nitrogen.

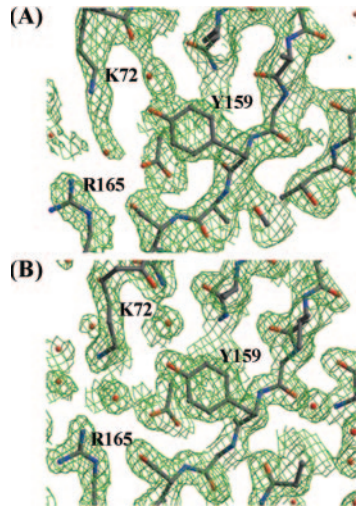

FIG. 1. **Experimental and final refined electron density maps.** The end of the β-sheet I is shown here. Tyr^{159} is buried between the two domains, with Lys^{72} and Arg^{165} closing the pocket, enclosing two water molecules. *A*, the refined SeMet model is shown in *stick form* on the initial experimental solvent flattened map at 2.60 Å resolution, contoured at 1 σ. *B*, the refined native structure is shown in *stick form* on the $2F_o - F_c$ electron density map at 1.86 Å resolution, contoured at 1 σ. Figures were visualized with XtalView (28) and drawn with RASTER3D (66).

Data Collection—Data for the native crystals were collected using a 1.01-Å radiation on a nine-element CCD detector at beamline 19ID at the Advanced Photon Source (APS), Argonne National Laboratory. A multiwavelength anomalous dispersion data set of the SeMet mutant crystal was collected at beamline X26C at the National Synchrotron Light Source, Brookhaven National Laboratory, at three different x-ray energies around the K absorption edge of selenium and recorded on an ADSC Quantum 4 CCD detector (Table I).

Structure Determination and Refinement—The present structure determination depended solely on the anomalous scattering from two selenium atoms (for 180 residues). The crystal structure of SeMet NtACP was determined by the single anomalous dispersion method, using its peak wavelength data set (Table I). The data were integrated and reduced using DENZO and SCALEPACK (26). Atomic positions for two selenium atoms ($SeMet^{88}$ and $SeMet^{94}$) were located by SOLVE (27), and the phases were calculated. A solvent-flattened map (RESOLVE) at 2.60 Å resolution was of high quality (see Fig. 1*A*). Clearly defined solvent boundaries were present, and most of the side chain densities were visible. Model building was performed using XtalView (28), and the atomic model was refined using REFMAC5 (29), with experimental phases included as restraints and with intermittent model rebuilding. The resulting model for residues 60–231 with most side chains present gave an R factor and R_{free} of 26.9 and 30.9%, respectively, *versus* the SeMet data (Table I). A rigid body fit of the SeMet model to the native data were performed by molecular replacement using EPMR (30) at 4 Å resolution. Further refinement of the atomic model and phase extension to 1.86 Å resolution were done with REFMAC5 and additional model rebuilding was performed with XtalView. Solvent molecules (water, dithiothreitol, and glycerol) were built based on expected hydrogen bond geometry and electron density. The final R factor and R_{free} of the NtACP (Fig. 1*B*) were 19.3 and 23.3%, respectively, at 1.86 Å nominal resolution (Table I). Analysis using PROCHECK (31) shows 97.1% of non-proline and non-glycine residues in the most favored region and 2.85% in the additionally allowed region of the Ramachandran plot. However, one residue (Asn^{76}) could not be built into the electron density envelope in a single low energy conformation, suggesting that multiple conformations of the main chain may be present.

Dot Blot Assay—Protein binding to heparin was studied with a modified dot blot technique (20) modeled after an immunoblot assay. Proteins (1 μg of each) were applied to a nitrocellulose membrane

[2] The nucleotide sequence for the alpha C protein gene has been deposited in the GenBank™ data base under accession number M97256 (23). The amino acid sequence of this protein can be accessed through NCBI Protein Database under NCBI accession number AAA26848 (23).

TABLE I
Data collection and refinement statistics

	Native	Mutant SeMet
Data collection		
Space group	$P6_122$	$P6_122$
Unit cell parameters		
a, *b*, *c*, Å	56.46, 56.46, 271.62	56.97, 56.97, 272.23
α, β, γ, °	90, 90, 120	90, 90, 120
Molecules/asymmetric unit	1	1
Wavelength, Å	1.0100	0.9763
Resolution range, Å	40–1.86 (1.93–1.86)	30–2.6 (2.69–2.60)
Completeness, %	95.1 (71.7)[a]	97.4 (99.2)
Observations total/unique	267,765/19,465 (1,462)	110,146/8.581 (828)
Redundancy	14.0	12.8
R_{sym}, %[b]	5.3 (33.9)	6.8 (13.4)
$\langle I\rangle/\langle\sigma(I)\rangle$	27.5 (4.7)	30.7 (24.7)
Heavy atoms sites (Se)	0	2
Refinement		
Ordered amino acid residues	180 (Ser^{57}–His^{236})	172 (Pro^{60}–Leu^{231})
No. of non-hydrogen atoms		
of the protein	1,410	1,300
of the solvent	528	51
Resolution range, Å	39.5–1.86 (1.96–1.86)	12–2.6 (2.73–2.60)
R_{cryst}, %[c]	19.3 (21.6)[a]	26.9 (32.7)
R_{free}, %[d]	23.3 (27.5)	30.8 (39.5)
Root mean square deviation bond length, Å	0.015	0.010
Root mean square deviation bond angle, °	1.406	1.552

[a] Highest resolution shell values are given in parentheses.
[b] $R_{sym} = 100*(\Sigma|I - \langle I\rangle|/\Sigma I)$, where I is the observed intensity of an individual reflection, and $\langle I\rangle$ is the mean intensity of that reflection.
[c] $R_{cryst} = 100*(\Sigma|F_o - F_c|/\Sigma|F_o|)$, where F_o and F_c are observed and calculated structure factors, respectively.
[d] R_{free}, R_{cryst} calculated for a subset of the reflections (10.1% for the native, 9.8% for the SeMet), which were omitted during the refinement and used to monitor its convergence.

followed by blocking for 1 h with 5% skim milk. The membrane was then incubated with heparin-albumin-biotin (Sigma) at 0.05 mg/ml for 1 h, followed by alkaline phosphatase-conjugated avidin (Pierce) for 1 h prior to washing and developing.

Flow Cytometry—Fluorescent labeling of proteins was performed using the AlexaFluor 488 Protein Labeling Kit (A-10235; Molecular Probes), to conjugate AlexaFluor 488 dye to protein samples, according to manufacturer's instructions as described previously (15).

ME180 human cervical epithelial cells (ATCC) were grown to monolayer confluence in 6-well plates with 2 ml of RPMI 1640 (Invitrogen), including 10% fetal calf serum (Invitrogen) and 1% penicillin/streptomycin (Invitrogen). The day prior to the assay, the medium was replaced with 1 ml of fresh medium, and the cells were incubated overnight at 37 °C with 5% CO_2. The next day, AlexaFluor-labeled protein was added to the wells to a final concentration of 0.1 μM. The plates were incubated at 37 °C with 5% CO_2 for 1.5 h. The medium was removed from the wells, and the monolayers were washed three times with 1 ml of PBS to remove unbound proteins. 350 μl of trypsin-EDTA (0.25% trypsin, 1 mM EDTA-4Na, Invitrogen) was added to the wells, and plates were incubated for 10 min at 37 °C. Cells were detached by repeated pipetting and harvested by centrifugation at 650 rpm (50 × *g*) for 8 min. Cells were washed with 1 ml of PBS and resuspended in 0.1 ml of 2% paraformaldehyde in PBS before incubation at 4 °C overnight. The samples were washed with 1 ml of PBS to remove the fixative, resuspended in 0.4 ml of PBS, filtered through a cell-strainer cap (Falcon), and analyzed by flow cytometry on the MoFLo (Cytomation) machine. The cell population of interest was identified by using the AlexaFluor-labeled bovine serum albumin sample to define nonspecific staining and/or autofluorescence levels. Positive cells were defined to have a fluorescent signal greater than that of 98.5% of the bovine serum albumin-treated control cell population.

RESULTS

Overall Structure—The crystal structure of the NtACP, residues 52–225, has been determined to 1.86 Å resolution by single anomalous dispersion (Table I and Fig. 2). The final model has acceptable geometry, with an *R* factor and R_{free} of 19.3 and 23.3%, respectively (Table I). The molecule is elongated with overall dimensions of 82 × 34 × 27 Å. It is composed of two major domains, the membrane-distal domain 1 (D1) is comprised of Ser^{57}–Asp^{160}, and domain 2 (D2) includes Ser^{161}–Leu^{225}. Domain 1 contains eight β-strands arranged into three β-sheets: a large flat sheet containing strands F-E-G-I and two

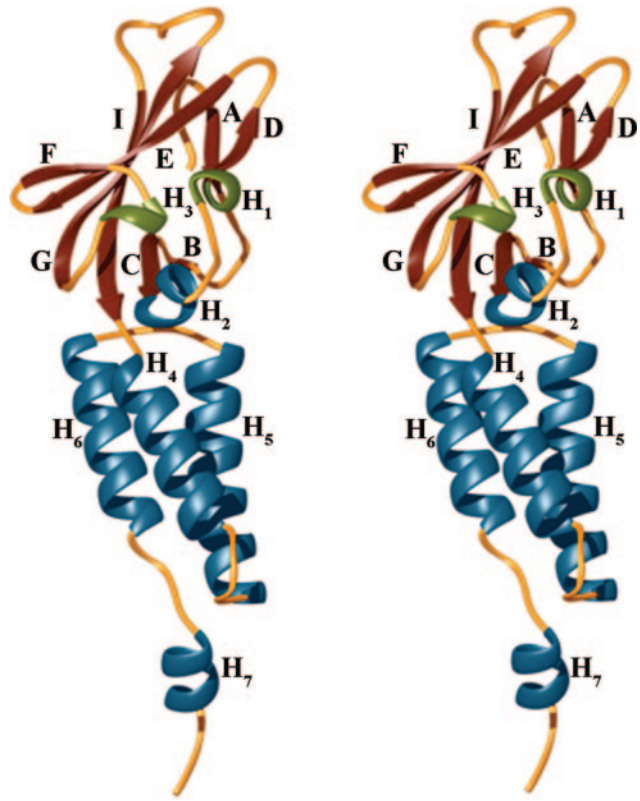

FIG. 2. **Stereo ribbon representation of the structure of the N-terminal domain of *S. agalactiae* Alpha C protein.** β-sheets are shown in *burgundy*, α-helices in *blue*, and 3_{10}-helices in *green*. Figures were drawn with RIBBONS (67) and rendered with POV-Ray.

smaller sheets, A-D and B-C (Fig. 2). These sheets enclose a hydrophobic core that is highly conserved among the Alp family members (Fig. 3). The F-E-G-I and the A-D sheets are arranged to form a typical β-sandwich, but the basic pattern is elaborated in a distinctive way by the addition of an α-helix (H_1, 6

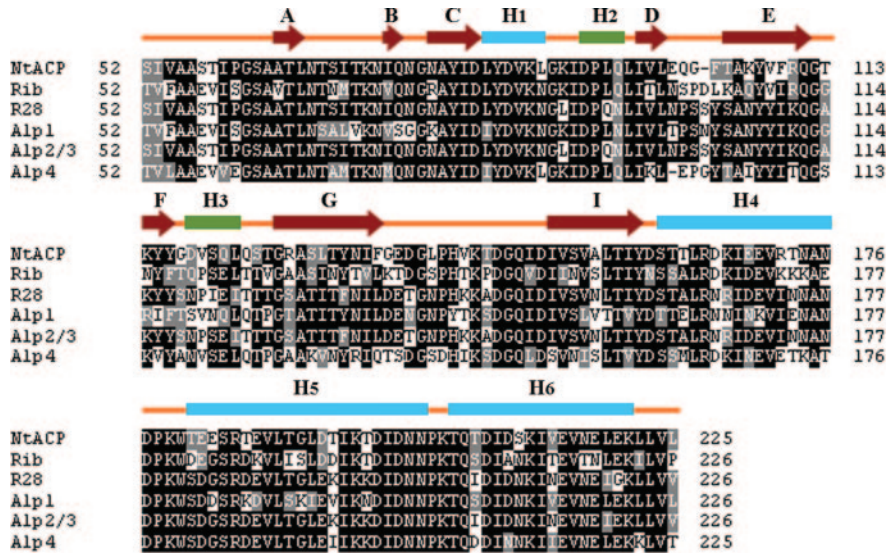

FIG. 3. **Sequence alignment of the N-terminal domain of the Alp family members.** The N-terminal domain sequences of ACP (NCBI accession number AAA26848) (23), Rib (NCBI accession number AAC44468) (39), R28 (NCBI accession number AAD39085) (17), Alp1 (NCBI accession number AAR08144, L. C. Madoff, unpublished data), Alp2 and Alp3 (NCBI accession numbers AAG01392 and AAG02097, respectively) (37), and Alp4 (NCBI accession number CAD32934) (68) are aligned. Highly conserved residues are shaded in *black* and *gray*. The secondary structure elements of NtACP are shown as *arrows* (β-strands), *bars* (*blue* and *green*, α-helices and 3_{10}-helices, respectively), and *lines* (connecting loops). The sequences were aligned with ClustalW (69).

residues) and two 3_{10}-helices (H_2, 4 residues, and H_3, 5 residues). Domain 2 is composed of three antiparallel α-helices (H_4, H_5, and H_6) arranged in a ~35-Å long left-handed three-helix bundle, a very common structural motif that delimits a highly conserved hydrophobic core (Figs. 2 and 3). The three α-helices are composed of 16 amino acids (Ser^{161}–Asn^{176}), 22 amino acids (Thr^{181}–Asn^{202}), and 17 amino acids (Thr^{205}–Lys^{221}), respectively. An additional, though biologically irrelevant, α-helix (H_7, 9 residues), corresponding to the linker from the pET24a vector connecting the protein to the His_6 tag, immediately follows domain 2.

Structural Alignment—Data base searches were carried out to establish how domain 1 of NtACP might be related to other proteins of known sequence and structure. Except for the Alps, homology searches at the nucleotide and amino acid levels did not identify any other proteins with meaningful similarities to NtACP. The structure comparison program DALI (32) found no protein that aligned well with domain 1 in its entirety. However, several protein structures align with a root mean square deviation of <3.0 Å with a subset of the C_α comprising 70% of the NtACP domain 1, which includes the F, E, G, and I strands of the large β-sheet and the A and D strands of the smaller sheet of the β-sandwich. The most significant structural homologues include several proteins known to be involved in cell adhesion, including MADCAM-1 (PDB code 1bqs (33)), type III fibronectin domain of tenascin (PDB code 1qr4 (34)), and module 10 of type III human fibronectin (FnIII10; PDB code 1fnf (35)). The closest structural similarity is seen when the large β-sheet F-E-G-I of NtACP is aligned with the corresponding β-sheet C′-C-F-G of FnIII10 (root mean square deviation of 1.1 Å on C_αs (Fig. 4*A*)).

The relationship between the two proteins is made clearer by protein folding topology diagrams of FnIII10 and NtACP domain 1 (Fig. 4*B*), which were calculated using the TOPS algorithm (36). The diagrams confirm the topological and structural correspondence between the large β-sheet of NtACP and FnIII10 and the correspondence of the A and D strands of NtACP with the A and B strands of the smaller three-stranded sheet of FnIII10. Interestingly, in NtACP a short helix (H_3) occupies the same position in the structure and folding pattern as the third strand of the small sheet of FnIII10 (strand E) (Fig. 4). NtACP domain 1 is significantly different from other proteins of known structure because of the presence of a single large and elaborate insertion into the common fold. In the three-dimensional structure of the protein, the inserted elements, β-strands B and C and helices H_1 and H_2, are clustered

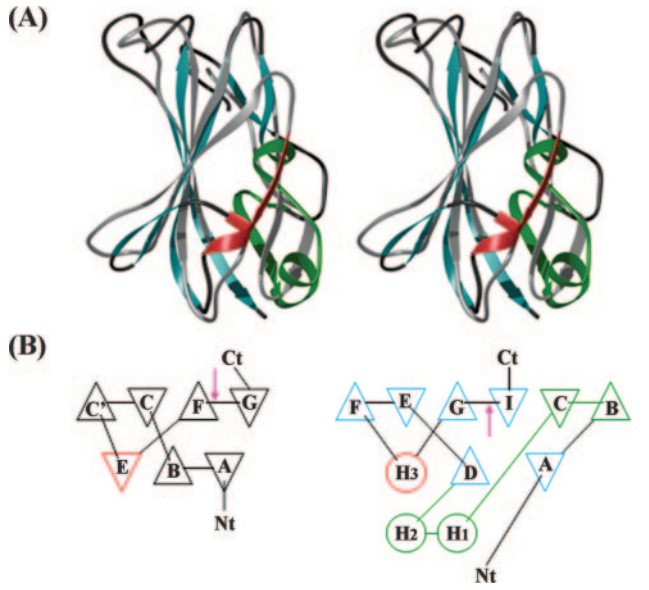

FIG. 4. **Structural alignment of NtACP domain 1 *versus* human type III fibronectin module 10.** *A*, the NtACP and FnIII10 backbones are superimposed in stereo. *B*, TOPS (36) protein folding topology cartoons of FnIII10 (*left*) and NtACP domain 1 (*right*). β-Strands are represented as *triangles* and α-helices as *circles*. The NtACP fold can be understood as a modification and elaboration of the FnIII pattern. The topologically similar secondary structural elements (shown in *gray* and *blue* for FnIII and NtACP, respectively) are structurally similar as well. Elements in *red* occupy analogous positions. Structural elements that are unique to NtACP (*green*) are found in a single large insertion. *Pink arrows* indicate the position of the known integrin-binding motif RGD and the putative integrin-binding motif KTD of FnIII10 and NtACP, respectively. Stereo figures were drawn with RIBBONS (67) and rendered with POV-Ray.

together on one face of the molecule (Fig. 4*A*). The sequence alignment of the N-terminal domains of Alp family members shows that this large insertion is present and highly conserved in the family members (Fig. 3). The conservation and the pronounced clustering of this large insertion suggest that it may play an essential functional role.

Despite the significant structural differences between NtACP and FnIII10, NtACP contains a large highly exposed loop between the G and I strands of the large sheet that is topologically and structurally related to the integrin-binding loop of FnIII10. In FnIII10, this loop contains a canonical Arg-Gly-Asp (RGD) integrin-binding motif, in which the Arg

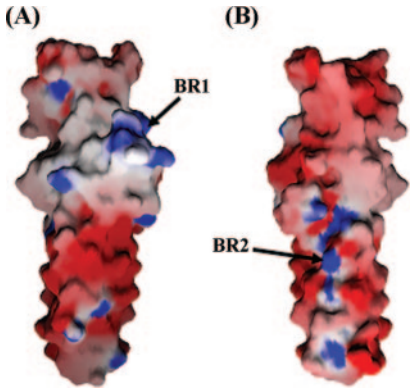

FIG. 5. **Molecular surface representation of the N-terminal domain of alpha C protein.** The views are related by a rotation of 180° about the vertical axis. The structure exhibits a highly acidic surface (*red*). However, we have identified two positively charged clusters, *BR1* and *BR2* (*blue*), as potential glycosaminoglycan-binding sites. Figures were drawn with GRASP (70).

and Asp side chains are highly exposed. In NtACP, the structurally analogous residues are Lys^{144}-Thr^{145}-Asp^{146} (KTD). The Lys^{144} and Asp^{146} side chains are exposed and are highly conserved in the Alp family. Although this sequence has not been previously described as an integrin-binding sequence, variations on the RGD motifs have been described for other integrin-binding proteins, raising the possibility that integrins may serve as ligands for NtACP-mediated cell adhesion.

Distribution of Surface Charges—The NtACP surface is dominated by acidic residues, except for two small positively charged clusters (Fig. 5). The distribution of positive charges may be important for understanding the structural basis of the ability of ACP to bind heparin (20). One basic cluster (BR1) is located in domain 1. It includes three basic residues from the vicinity of the E-F loop, Arg^{110}, Lys^{114}, and Lys^{106}, and two amide side chains, Gln^{121} and Asn^{133}. The positively charged atoms typically lie ~8–9 Å apart and only two of them, Arg^{110} and Lys^{114}, show charge conservation among Alp family members (Figs. 3 and 5*A*). The second, much larger, cluster (BR2) includes basic residues Lys^{72} and Lys^{90} from domain 1, and Arg^{165}, Lys^{196}, Arg^{172}, Arg^{185}, and the amide side chain of Asn^{176} in domain 2. These residues are aligned down the cleft between α-helices H4 and H5 and exposed to the solvent, forming a well defined positively charged band of ~35 Å in length (Fig. 5*B*). Residues included in BR2 are on average 6 Å apart, and most of them are conserved (Lys^{72}, Arg^{165}, Lys^{196}, and Arg^{185}) or conserved by charge (Lys^{90}) within the Alp family (Fig. 3).

Heparin Binding Activity of ACP—Prior work has shown binding of full-length 1-repeat and 9-repeat ACP to ME180 cells in culture and to heparin in a dot blot assay (20). It was reported that the NtACP and the 9RR, taken separately, bind only minimally in these assays and that both the NtACP and the repeat-region domains are required for the GAG binding activity of ACP (20). The NtACP crystal structure shows a cluster of positively charged residues (BR2) in the three-helix bundle that is predicted to lie near the junction of NtACP with the first of the ACP repeats.

To test whether the BR2 region contributes to the GAG binding activity of full-length ACP, we expressed and purified D2-R, comprised of the three-helix bundle of the NtACP (starting at Leu^{164}), one 82-amino acid repeat, and 13 amino acids in the C-terminal region. It thus includes most of the charged residues in BR2, excluding only Lys^{72} and Lys^{90}, due to their location distant from D2 in the linear sequence of NtACP. D2-R was tested for heparin binding activity by dot blot analysis as described in Ref. 20. It showed a level of activity similar to that

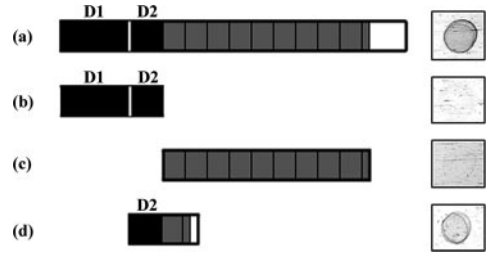

FIG. 6. **Localization of the heparin binding activity of ACP by dot blot assay.** *a–d*, constructs corresponding to various portions of ACP. *a*, full-length ACP includes NtACP (*black*), a series of tandem repeats (*gray*), and a C-terminal domain (*white*). *b*, NtACP alone. *c*, 9 repeats (9RR). *d*, D2-R is responsible for most of the heparin binding activity of full-length ACP.

of full-length 9-repeat ACP (Fig. 6). In contrast, neither NtACP nor the 9RR region alone bound to heparin appreciably.

D2-R was also evaluated for its interaction with ME180 human cervical epithelial cells in flow cytometry assays performed as described previously (20). We found that D2-R associates with 91.6% of these cells, which is comparable with the binding activity of full-length ACP. Furthermore, this binding was inhibited in a concentration-dependent way in the presence of soluble heparin, as was previously seen for ACP (20). In marked contrast, fewer than 5% of the cells bound either the NtACP or 9RR construct alone (20).

DISCUSSION

Adhesion to human epithelial and endothelial cells is a critical step for GBS colonization and invasion. ACP binds to human epithelial cells and plays a role in the internalization and translocation of GBS across epithelial cells (15). Similarly, group A *Streptococcus* Alp protein R28, identical to GBS Alp3, binds human epithelial cells (17, 37). ACP expressed by GBS strain A909 contains a 56-amino-acid residue signal sequence at the N terminus, highly conserved among Alps. It is followed by a 174-residue N-terminal domain (Ser^{52}–Leu^{225}), nine 82-residue tandem repeats, and a 45-residue C-terminal domain containing a LPXTG peptidoglycan-binding motif. Immunogenic and protective epitopes have been mapped to both the NtACP and repeat regions (22). The number of tandem repeats in ACP expressed by other GBS strains varies from 1 to 16, affecting the antigenicity and presumably the protein structure (38). NtACP competitively inhibits binding of ACP to and internalization of GBS within the human cervical epithelial cell line ME180 (15), suggesting that the N-terminal region of ACP is involved in binding to one or more receptors on the surface of these cells and that binding facilitates the internalization of GBS. The N-terminal regions of Alps share 60–100% similarity at the amino acid level (Fig. 3), suggesting that they may function similarly (17, 23, 37, 39).

We have solved the three-dimensional structure of NtACP of GBS strain A909 by x-ray crystallography at high resolution to further understand the role(s) that the N-terminal region of Alps may play during streptococcal infection. NtACP is composed of a membrane-distal β-sandwich at its amino end, and a three-helix bundle at its carboxyl end. In the complete ACP, the carboxyl domain links to the first of the series of repeat regions. Highly conserved residues on the NtACP surface were identified by combining NtACP structural information with a structure-based sequence alignment of the N-terminal regions of the Alp family members. A particularly high degree of sequence conservation was seen along one side of a cleft on top of the molecule, in the G-I loop. The conservation of these residues does not appear to be essential for maintaining the integrity of the three-dimensional fold, which suggests that they may play a functional role instead.

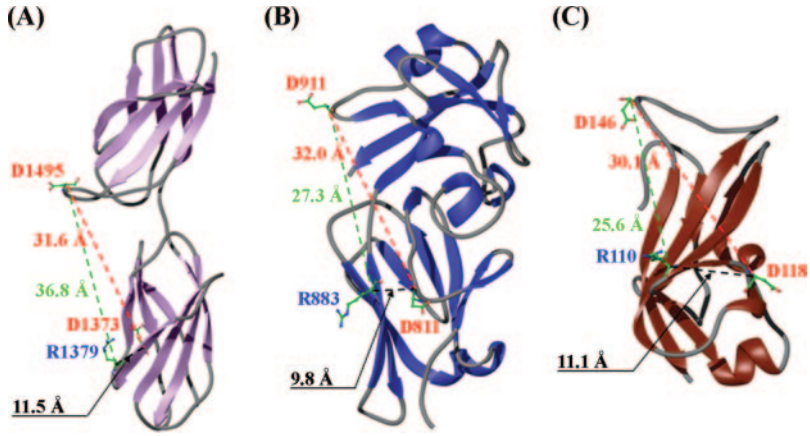

FIG. 7. **A structurally conserved triangle of exposed charges may be implicated in integrin binding.** Ribbon representations of FnIII9–10 (*A*) (35), Inv497 D4-D5 (*B*) (51), and domain 1 of NtACP (*C*) (this paper). Distances reported are measured between C_α. Residues (RG)Asp^{1495} (*D1495*), Asp^{1373} (*D1373*), and Arg^{1379} (*D1379*) of FnIII are known to be required for integrin binding (35, 40, 50). Residues Asp^{911} (*D911*) and Asp^{811} (*D811*) of *Yersinia* Inv497 are known to be involved in integrin binding (54–56,71) and residue Arg^{883} (*R883*) has been proposed as a structural analogue of Arg^{1379} (*R1379*) of FnIII based on interresidue distances (51).

Potential Integrin-binding Site—A number of proteins involved in cell adhesion or which serve as cell surface receptors share a common folding pattern. A close structural homology between NtACP β-sandwich (domain 1) and FnIII10 (35) was identified (Fig. 4), wherein each secondary structural element identified as analogous in a topological alignment was also structurally similar. The integrin binding activity of FnIII10 requires the presence of an Arg^{1493}-Gly^{1494}-Asp^{1495} (RGD) sequence motif (40, 41) on an exposed F-G loop of the structure (35). Although NtACP lacks such an RGD sequence, there is an exposed Lys^{144}-Thr^{145}-Asp^{146} (KTD) motif at the top end of domain 1. The KTD motif of NtACP is part of the G-I loop that structural and topological alignment procedures have identified as analogous to the loop containing the RGD motif of FnIII10 (35, 40). Although integrins generally bind RGD sequences, several proteins have been shown to bind a number of related motifs, such as KTS (42, 43), MLD (44), or MGD(W) (45) of disintegrins obtustatin, EC3, and EMF-10, respectively. Similarly, COL15, the largest collagenous domain of type XVII collagen, binds to the $\alpha_5\beta_1$ and $\alpha_V\beta_1$ integrins through a KGD motif (46). The sequence alignment of the N-terminal domains of the members of the Alp family shows that Lys^{144} and Asp^{146} are conserved in all members, and although Thr^{145} is not conserved, other members display uncharged residues with small side chains (Fig. 3). The conserved residues surrounding the RGD loop of FnIII10 are acidic, as are those surrounding the KTD loop of NtACP.

In addition to the RGD motif of FnIII10, FnIII also requires module 9 for optimal integrin-binding affinity (47). Two charged residues in module 9 (FnIII9) have been shown to be critical for binding to $\alpha_5\beta_1$ integrin. These are Arg^{1379}, which belongs to a synergy domain with basic residues (35, 48, 49), and Asp^{1373} (35, 50). Both Arg^{1379} and Asp^{1373} are located ~30 Å away from the RGD motif of FnII10 (Fig. 7*A*). Invasin from *Yersinia pseudotuberculosis* (Inv497) is a non-RGD protein that competes with FnIII by binding to $\alpha_5\beta_1$ integrin at the same site as fibronectin (51–53) with ~100 times greater affinity. In reporting the invasin structure, Hamburger *et al.* (51) proposed that Inv497 (PDB code 1cwv), although structurally dissimilar from FnIII, shares with FnIII a structurally conserved triangular arrangement of charged residues (Fig. 7, *A* and *B*). The proposed integrin-binding triangle in Inv497 includes Asp^{911} and Asp^{811}, which were identified by mutational analysis (54–56), and Arg^{883}, which was suggested by structural considerations (51).

We have found that the structure of NtACP has a very similar arrangement of exposed aspartate and arginine residues (Fig. 7*C*), wherein the proposed KTD sequence of NtACP and the known RGD sequence of FnIII10 occupy analogous positions, and NtACP residues Arg^{110} and Asp^{118} are each positioned within 30 Å of the KTD motif and are structurally analogous to Arg^{1379}, and Asp^{1373} of FnIII9–10. We hypothesize that NtACP and *Y. pseudotuberculosis* invasin may mimic the integrin binding activity of FnIII in similar ways. Further analysis will be necessary to determine which of the integrins or other molecules of the extracellular matrix are bound by NtACP.

Potential Heparin/Heparan Sulfate-binding Site—Several bacterial pathogens express adhesins that bind heparin (57–59). Heparin, an anticoagulant GAG, is composed of repeating disaccharide units of D-glucosamine and uronic acid linked by α1→4 interglycosidic bonds. Several hydroxyl groups of these monosaccharide residues become sulfated giving rise to a highly negatively charged polymer. These negatively charged sulfates in GAG molecules bind to positively charged groups on proteins (60–62). It has previously been shown that ACP binds to host cell GAG, *in vitro*, in a dose-dependent manner, suggesting the presence of a specific binding site on the surface of the protein (20). High affinity binding requires the presence both of NtACP and at least one of the ACP repeat domains, suggesting that the complete binding site spans both domains. Knowledge of the detailed molecular structure of NtACP now provides the opportunity to identify which surface residues participate in forming the heparin-binding site(s), as a way to begin to understand the molecular basis of infection.

Pure preparations of homogeneously sulfated heparin have well determined structures forming helical ribbons with an ~5-Å regular spacing of the sulfates along the edge of the ribbon and ~11 Å across the ribbon (63). The binding of heparan sulfate by proteins of known structure can involve a distribution of positively charged side chains on the protein surface that is complementary to the arrangement of the negative sulfate groups along one or both edges of the polymer (64, 65). To see whether such a regular arrangement of positive charges would be recognizable in NtACP, the distribution of charges on the molecular surface was calculated (Fig. 5). The surface of NtACP is dominated by acidic residues, except for two small positively charged clusters, BR1 and BR2. Residues of BR1 belong exclusively to domain 1 of NtACP, and exhibit an 8–9-Å spacing between residues. BR2 spans both domain 1 and domain 2 of NtACP, and the average spacing between residues is ~6 Å.

Two factors argue that the basic residues of BR2, rather than BR1, are the most likely to be responsible for the NtACP contribution to heparin binding activity. First, heparin binding requires at least one of the ACP repeats to be present (20). The crystal structure shows that some of the charged residues in BR2 (but none of the residues in BR1) belong to domain 2 and are located quite close to the carboxyl end of NtACP, where the first of the repeats is attached in the full-length ACP. It is easy

to imagine a spatially contiguous binding site that includes residues from both domains of NtACP and residues from one or more of the repeat domains as well. A second important factor is that minor changes in the conformations of flexible Lys and Arg side chains could easily reduce the average 6-Å spacing between positive charges that was seen in the crystal structure to the 5-Å spacing required for optimal complementarity. Flexibility in basic side chains of the protein may well favor GAG binding by compensating for heterogeneity in the sequence, conformation, or a sulfation pattern of the ligand. The heparin binding activity of D2-R, which includes the majority of the residues of BR2, supports our hypothesis that the BR2 region of NtACP contributes to the heparin binding activity of the protein. Furthermore, fluorescently labeled D2-R binds to epithelial cells, and this binding is inhibited by soluble heparin. These data, in conjunction with the absence of heparin binding activity in the isolated NtACP or repeat-region constructs alone, suggest that a GAG binding domain involves a junctional epitope including residues in the BR2 domain as well as residues in the adjacent repeat region.

The relevance of heparin binding to the process of cell entry by GBS is strongly suggested by the correlation between results in the dot blot and cell binding assays. Specifically, both samples that showed similarly high levels of binding activity in the dot blot assay also showed high levels in binding epithelial cells. Conversely, both ACP fragments showing negligible levels of binding did so in both assays.

The binding of ACP to cells, and subsequent steps in cell entry, could involve one or more families of receptor molecules bound simultaneously, alternatively, or in sequence. Possible binding sites for integrin and GAG have been identified, and it is interesting to note that the proposed binding sites for the two ligands lie on opposite faces of the protein. Knowledge of the high-resolution structure of the distal end of ACP provides a framework for mutational studies to identify cellular receptors and their binding sites on the ACP surface. Structural determination of the repeat regions would clearly contribute to further characterize the heparin binding domain of ACP.

Acknowledgments—We thank the staff of APS beamline 19ID (S. Ginell) and the staff of NSLS beamline X26C (S. Myers) for their tremendous help during data collection. We acknowledge the staff of NSLS beamline X12C (R. Sweet and A. Saxena), and the staff of BioCARS (R. Henning) for assistance in developing the crystallization and freezing protocols and during attempts to find isomorphous derivatives. We thank the members of the Hogle laboratory (especially L. Guogas, B. Appleton, and V. Weiss) and members of the Ellenberger laboratory (especially E. Toth and B. Eichman) at the Harvard Medical School for their help and discussion during this work. We thank Meghan Gilmore for assistance in preparing D2-R and Hope Hamrick for assistance with flow cytometry assays.

REFERENCES

1. Baker, C. J., and Edwards, M. S. (1995) in *Infectious Diseases of the Fetus and Newborn Infant* (Remington, J., and Klein, J. O., eds) 4th Ed., pp. 980–1054, W.B. Saunders, Philadelphia
2. Schrag, S., Gorwitz, R., Fultz-Butts, K., and Schuchat, A. (2002) *MMWR Recomm. Rep.* **51,** 1–22
3. Jackson, L. A., Hilsdon, R., Farley, M. M., Harrison, L. H., Reingold, A. L., Plikaytis, B. D., Wenger, J. D., and Schuchat, A. (1995) *Ann. Intern. Med.* **123,** 415–420
4. Tyrrell, G. J., Senzilet, L. D., Spika, J. S., Kertesz, D. A., Alagaratnam, M., Lovgren, M., and Talbot, J. A. (2000) *J. Infect. Dis.* **182,** 168–173
5. Dahl, M. S., Tessin, I., and Trollfors, B. (2003) *Int. J. Infect. Dis.* **7,** 113–119
6. Zangwill, K. M., Schuchat, A., and Wenger, J. D. (1992) *MMWR CDC Surveill. Summ.* **41,** 25–32
7. Schrag, S. J., Zywicki, S., Farley, M. M., Reingold, A. L., Harrison, L. H., Lefkowitz, L. B., Hadler, J. L., Danila, R., Cieslak, P. R., and Schuchat, A. (2000) *N. Engl. J. Med.* **342,** 15–20
8. Bliss, S. J., Manning, S. D., Tallman, P., Baker, C. J., Pearlman, M. D., Marrs, C. F., and Foxman, B. (2002) *Clin. Infect. Dis.* **34,** 184–190
9. Centers for Disease Control (1997) *MMWR* (*Morb. Mortal. Wkly. Rep.*) **46,** 473–477
10. Zawaneh, S. M., Ayoub, E. M., Baer, H., Cruz, A. C., and Spellacy, W. N. (1979) *Infect. Immun.* **26,** 441–447
11. Tamura, G. S., Kuypers, J. M., Smith, S., Raff, H., and Rubens, C. E. (1994) *Infect. Immun.* **62,** 2450–2458
12. Rubens, C. E., Smith, S., Hulse, M., Chi, E. Y., and van Belle, G. (1992) *Infect. Immun.* **60,** 5157–5163
13. Hulse, M. L., Smith, S., Chi, E. Y., Pham, A., and Rubens, C. E. (1993) *Infect. Immun.* **61,** 4835–4841
14. Valentin-Weigand, P., Jungnitz, H., Zock, A., Rohde, M., and Chhatwal, G. S. (1997) *FEMS Microbiol. Lett.* **147,** 69–74
15. Bolduc, G. R., Baron, M. J., Gravekamp, C., Lachenauer, C. S., and Madoff, L. C. (2002) *Cell. Microbiol.* **4,** 751–758
16. Bulgakova, T. N., Grabovskaya, K. B., Ryc, M., and Jelinkova, J. (1986) *Folia Microbiol.* (*Praha*) **31,** 394–401
17. Stalhammar-Carlemalm, M., Areschoug, T., Larsson, C., and Lindahl, G. (1999) *Mol. Microbiol.* **33,** 208–219
18. Shankar, N., Lockatell, C. V., Baghdayan, A. S., Drachenberg, C., Gilmore, M. S., and Johnson, D. E. (2001) *Infect. Immun.* **69,** 4366–4372
19. Li, J., Kasper, D. L., Ausubel, F. M., Rosner, B., and Michel, J. L. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 13251–13256
20. Baron, M. J., Bolduc, G. R., Goldberg, M. B., Auperin, T. C., and Madoff, L. C. (2004) *J. Biol. Chem.* **279,** 24714–24723
21. Gravekamp, C., Kasper, D. L., Michel, J. L., Kling, D. E., Carey, V., and Madoff, L. C. (1997) *Infect. Immun.* **65,** 5216–5221
22. Kling, D. E., Gravekamp, C., Madoff, L. C., and Michel, J. L. (1997) *Infect. Immun.* **65,** 1462–1467
23. Michel, J. L., Madoff, L. C., Olson, K., Kling, D. E., Kasper, D. L., and Ausubel, F. M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 10060–10064
24. Gravekamp, C., Horensky, D. S., Michel, J. L., and Madoff, L. C. (1996) *Infect. Immun.* **64,** 3576–3583
25. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) *J. Mol. Biol.* **229,** 105–124
26. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
27. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 849–861
28. McRee, D. E. (1993) *Practical Protein Crystallography*, Academic Press, New York
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D* **53,** 240–255
30. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 484–491
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
32. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
33. Tan, K., Casasnovas, J. M., Liu, J. H., Briskin, M. J., Springer, T. A., and Wang, J. H. (1998) *Structure* (*Lond.*) **6,** 793–801
34. Bisig, D., Weber, P., Vaughan, L., Winterhalter, K. H., and Piontek, K. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 1069–1073
35. Leahy, D. J., Aukhil, I., and Erickson, H. P. (1996) *Cell* **84,** 155–164
36. Flores, T. P., Moss, D. S., and Thornton, J. M. (1994) *Protein Eng.* **7,** 31–37
37. Lachenauer, C. S., Creti, R., Michel, J. L., and Madoff, L. C. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9630–9635
38. Gravekamp, C., Rosner, B., and Madoff, L. C. (1998) *Infect. Immun.* **66,** 4347–4354
39. Wastfelt, M., Stalhammar-Carlemalm, M., Delisse, A. M., Cabezon, T., and Lindahl, G. (1996) *J. Biol. Chem.* **271,** 18892–18897
40. Main, A. L., Harvey, T. S., Baron, M., Boyd, J., and Campbell, I. D. (1992) *Cell* **71,** 671–678
41. Hynes, R. O. (1992) *Cell* **69,** 11–25
42. Monleon, D., Moreno-Murciano, M. P., Kovacs, H., Marcinkiewicz, C., Calvete, J. J., and Celda, B. (2003) *J. Biol. Chem.* **278,** 45570–45576
43. Moreno-Murciano, M. P., Monleon, D., Calvete, J. J., Celda, B., and Marcinkiewicz, C. (2003) *Protein Sci.* **12,** 366–371
44. Marcinkiewicz, C., Calvete, J. J., Marcinkiewicz, M. M., Raida, M., Vijay-Kumar, S., Huang, Z., Lobb, R. R., and Niewiarowski, S. (1999) *J. Biol. Chem.* **274,** 12468–12473
45. Marcinkiewicz, C., Calvete, J. J., Vijay-Kumar, S., Marcinkiewicz, M. M., Raida, M., Schick, P., Lobb, R. R., and Niewiarowski, S. (1999) *Biochemistry* **38,** 13302–13309
46. Nykvist, P., Tasanen, K., Viitasalo, T., Kapyla, J., Jokinen, J., Bruckner-Tuderman, L., and Heino, J. (2001) *J. Biol. Chem.* **276,** 38673–38679
47. Altroff, H., van der Walle, C. F., Asselin, J., Fairless, R., Campbell, I. D., and Mardon, H. J. (2001) *J. Biol. Chem.* **276,** 38885–38892
48. Redick, S. D., Settles, D. L., Briscoe, G., and Erickson, H. P. (2000) *J. Cell Biol.* **149,** 521–527
49. Takagi, J., Strokovich, K., Springer, T. A., and Walz, T. (2003) *EMBO J.* **22,** 4607–4615
50. Bowditch, R. D., Hariharan, M., Tominna, E. F., Smith, J. W., Yamada, K. M., Getzoff, E. D., and Ginsberg, M. H. (1994) *J. Biol. Chem.* **269,** 10856–10863
51. Hamburger, Z. A., Brown, M. S., Isberg, R. R., and Bjorkman, P. J. (1999) *Science* **286,** 291–295
52. Isberg, R. R., Hamburger, Z., and Dersch, P. (2000) *Microbes Infect.* **2,** 793–801
53. Stebbins, C. E., and Galan, J. E. (2001) *Nature* **412,** 701–705
54. Leong, J. M., Fournier, R. S., and Isberg, R. R. (1990) *EMBO J.* **9,** 1979–1989
55. Leong, J. M., Fournier, R. S., and Isberg, R. R. (1991) *Infect. Immun.* **59,** 3424–3433
56. Leong, J. M., Morrissey, P. E., Marra, A., and Isberg, R. R. (1995) *EMBO J.* **14,** 422–431
57. Menozzi, F. D., Mutombo, R., Renauld, G., Gantiez, C., Hannah, J. H., Leininger, E., Brennan, M. J., and Locht, C. (1994) *Infect. Immun.* **62,** 769–778
58. Menozzi, F. D., Rouse, J. H., Alavi, M., Laude-Sharp, M., Muller, J., Bischoff, R., Brennan, M. J., and Locht, C. (1996) *J. Exp. Med.* **184,** 993–1001
59. Freissler, E., Meyer auf der Heyde, A., David, G., Meyer, T. F., and Dehio, C. (2000) *Cell. Microbiol.* **2,** 69–82
60. Shimazaki, K., Tazume, T., Uji, K., Tanaka, M., Kumura, H., Mikawa, K., and Shimo-Oka, T. (1998) *J. Dairy Sci.* **81,** 2841–2849

61. Delacoux, F., Fichard, A., Geourjon, C., Garrone, R., and Ruggiero, F. (1998) *J. Biol. Chem.* **273,** 15069–15076
62. Smith, S. A., Mullin, N. P., Parkinson, J., Shchelkunov, S. N., Totmenin, A. V., Loparev, V. N., Srisatjaluk, R., Reynolds, D. N., Keeling, K. L., Justus, D. E., Barlow, P. N., and Kotwal, G. J. (2000) *J. Virol.* **74,** 5659–5666
63. Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J., and Rees, D. C. (1996) *Science* **271,** 1116–1120
64. DiGabriele, A. D., Lax, I., Chen, D. I., Svahn, C. M., Jaye, M., Schlessinger, J., and Hendrickson, W. A. (1998) *Nature* **393,** 812–817
65. Sharma, A., Askari, J. A., Humphries, M. J., Jones, E. Y., and Stuart, D. I. (1999) *EMBO J.* **18,** 1468–1479
66. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
67. Carson, M. (1997) *Methods Enzymol.* **277,** 493–505
68. Creti, R., Fabretti, F., Orefici, G., and von Hunolstein, C. (2004) *J. Clin. Microbiol.* **42,** 1326–1329
69. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
70. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
71. Leong, J. M., Morrissey, P. E., and Isberg, R. R. (1993) *J. Biol. Chem.* **268,** 20524–20532

Crystal Structure of the N-terminal Domain of the Group B *Streptococcus* Alpha C Protein

Thierry C. Aupérin, Gilles R. Bolduc, Miriam J. Baron, Annie Heroux, David J. Filman, Lawrence C. Madoff and James M. Hogle

J. Biol. Chem. 2005, 280:18245-18252.
doi: 10.1074/jbc.M412391200 originally published online March 6, 2005
