Structural Conservation of Neurotropism-associated VspA within the Variable *Borrelia* Vsp-OspC Lipoprotein Family*

Wolfram R. Zückert‡§, Tatiana A. Kerentseva‡, Catherine L. Lawson‡, and Alan G. Barbour‡§

Received for publication, September 14, 2000, and in revised form, October 2, 2000
Published, JBC Papers in Press, October 3, 2000, DOI 10.1074/jbc.M008449200

Vsp surface lipoproteins are serotype-defining antigens of relapsing fever spirochetes that undergo multiphasic antigenic variation to avoid the immune response. One of these proteins, VspA of *Borrelia turicatae*, is also associated with neurotropism in infected mice. Vsp proteins are highly polymorphic in sequence, which may relate to their specific antibody reactivities and host cell interactions. To determine whether sequence variations affect protein structure, we compared *B. turicatae* VspA with three related proteins: VspB of *B. turicatae*, Vsp26 of the relapsing fever agent *Borrelia hermsii*, and OspC of the Lyme disease spirochete *Borrelia burgdorferi*. Recombinant non-lipidated proteins were purified by affinity or ion exchange chromatography. Circular dichroism spectra revealed similar, highly α-helical secondary structures for all four proteins. *In vitro* assays demonstrated protease-resistant, thermostable Vsp cores starting at a conserved serine at position 34 (Ser34). All proteins aggregate as dimers in solution. *In situ* trypsin treatment and surface protein cross-linking showed that the native lipoproteins also form protease-resistant dimers. These findings indicate that Vsp proteins have a common compact fold and that their established functions are based on localized polymorphisms. Two forms of VspA crystals suitable for structure determination by x-ray diffraction methods have been obtained.

Relapsing fever and Lyme disease are bacterial infectious diseases caused by spirochetes of the genus *Borrelia*. In contrast to other spirochetes that affect humans, such as the syphilis agent *Treponema pallidum*, their transmission to vertebrate hosts depends on arthropod vectors (1). Throughout their vector-host life cycle, *Borreli* cells display abundant lipoproteins on their surfaces. The genomic sequence of the Lyme disease spirochete *Borrelia burgdorferi* revealed over 130 putative lipoprotein genes on its chromosome and plasmids (2, 3). With the exception of 3 matrix-binding proteins (4, 5), their specific functions have not been determined.

Included in this class of lipoproteins with unknown function is the tick-associated Outer Surface Protein A (OspA) (6), which is in current use as a Lyme disease vaccine for humans (7). OspA is anchored to the outer membrane via an N-terminal triacyl-modified cysteine (8, 9) and is intrinsically resistant to proteases such as trypsin, despite its surface accessibility and high lysine content (10–12). The OspA crystal structure revealed 21 antiparallel β-strands and one C-terminal α-helix, which fold into globular N- and C-terminal domains connected by an unusual, free-standing nonglobular β-sheet (13–15).

Less is known about the structure of the major outer membrane lipoproteins of relapsing fever *Borreliae*, which are immunodominant and determine serotype (16). Collectively called variable membrane proteins, they have been divided into two families: the variable large proteins (Vlps) of 36–40 kDa and the variable small proteins (Vsp)s of 20–23 kDa (17). Vlp and Vsp proteins appear unique to the genus *Borrelia*. They have been described in the relapsing fever species *Borrelia turicatae* (17–20), *Borrelia hermsii* (21), *Borrelia recurrentis* (22), and *Borrelia crocidurae* (23). *B. burgdorferi* OspC, a major outer surface protein in early Lyme disease (24), is phylogenetically related to the Vsp (17), which has led to the term Vsp-OspC family.

While their signal peptides necessary for proper translocation and processing are conserved, the mature Vlp and Vsp-OspC lipoproteins are highly polymorphic. OspC amino acid sequences of different Lyme disease *Borreliae* can vary as much as 25% (25, 26). There is evidence that this variation is maintained, even within a local population, by frequency-dependent balancing selection (27). Vlp and Vsp proteins diverge even more, with 40 to 80% amino acid identities among them (17, 28). In contrast to the single plasmid-encoded ospC of *B. burgdorferi* (29), several archival copies of *B. hermsii* vsp and vlp genes are maintained on linear plasmids and sequentially expressed from a promoter site after gene conversions or DNA rearrangements (21, 30, 31). The resulting multiphasic antigenic variation of Vsp and Vlps allows the spirochete to repeatedly evade the host’s immune response, which leads to recurrent spirochemia and the characteristic febrile episodes (16). This strategy is analogous to that for variant surface glycoproteins (VSG) of African trypanosomes, the PiEMP1 proteins of the malarial parasite *Plasmodium falciparum*, and the PilE/PilS and Opa surface proteins of *Neisseria* species (32, 33).

Experimental relapsing fever in mice has demonstrated a role for Vsp proteins in differential tissue localization as well as for avoidance of the immune response. In a clonal population of *B. turicatae*, expression of VspB was associated with high densities of spirochetes in the blood, while expression of VspA led to early invasion and persistent infection of the central nervous system (17–19, 34, 35). VspA may help to evacuate the bacterium to an immunoprivileged niche, while VspB may facilitate efficient transmission to the next feeding tick. Tissue culture assays yielded clues about pathogenesis mechanisms. *B. turi-
cateae expressing VspA penetrated human umbilical vein epithelial cell monolayers more readily than those expressing VspB (18). On the other hand, VspB increased binding of the spirochete to mammalian endothelial and glial cells (36), predominantly by the direct interaction of VspB with host cell surface glycosaminoglycans (37).

Structural data for Vsp-OspC proteins are limited. NMR studies of B. burgdorferi strain B31 OspC have suggested that the core of OspC consists of four α-helices, while the N- and C-terminal sequences are highly flexible (38). OspC dimers were observed both with recombinant protein in solution (38) as well as with cell surface-exposed native lipoprotein (11). However, the effect of sequence variation on Vsp-OspC protein structure, and thus the structural basis of their established functions, remained unknown. To address this, we determined structural features of VspA and three other members of the Vsp-OspC protein family by biochemical and biophysical methods. A comparison indicates that Vsp-OspC proteins share a highly α-helical, compact fold, which includes a dimerization domain and confers protease resistance to a large central portion of the Vsp proteins. For high resolution x-ray structure determination, we obtained two forms of VspA crystals.

**EXPERIMENTAL PROCEDURES**

**Expression and Purification of Recombinant Histidine-tagged Fusion Proteins**—Non-lipidated, soluble fusion proteins with N-terminal histidine tags (hisVspA, hisVspB, hisVsp26, and hisOspC) were obtained as described for hisVspA (17). Briefly, gene segments lacking the N-terminal signal sequence and lipoprotein signal were amplified by polymerase chain reaction using Taq thermostable DNA polymerase (Roche Molecular Biochemicals). Forward and reverse primers for hisVspA, hisVsp26, and hisOspC are listed in Table I. The polymerase chain reaction products were ligated into pET15b (Novagen) and cloned in Escherichia coli BL21(DE3) (Novagen). Recombinant proteins were expressed after induction with isopropyl-β-D-thiogalactopyranoside and subsequently purified by nickel affinity chromatography. Recombinant HisOspC proteins were cleaved after induction using formaldehyde as described for B. burgdorferi OspC (11).

**Expression and Purification of N-terminal Truncated Core Proteins**—To obtain core proteins S34VspA and S34VspB, vspA and vspB sequences starting with the Ser24 codon and ending with the translational stop codon were amplified from plasmid clones (see above and Refs. 17 and 19) by polymerase chain reaction using Pfu polymerase (Roche Molecular Biochemicals). The forward primer for both sequences was S34vpsA-fwd. Reverse primers were S34vpsA-rev and S34vpsB-rev, respectively (see Table I). The polymerase chain reaction products were digested with NdeI and BamHI and ligated into pET29b (Novagen). Recombinant plasmids pET29::S34vpsA and pET29b::S34vpsB were obtained after transformation of E. coli InvA2F (Invitrogen); E. coli BL21(DE3) (Novagen) transformants containing the plasmids were grown overnight from single colonies at 30 °C in LB containing 100 μg/ml of ampicillin and 30 μg/ml of kanamycin. Recombinant purified proteins with N-terminal His tags (hisVspA, hisVspB, hisVsp26, and hisOspC) were treated separately with four proteases. The two serine proteases, trypsin and plasmin (Roche Molecular Biochemicals), were used at 10 to 400 μg/ml or 1 mg/ml, respectively, in 100 mM Tris-HCl, pH 8.5, while lysosomal cysteine protease cathepsin B (Calbiochem) and aspartyl protease cathepsin D (Calbiochem) were used at concentrations of 1.5 units/ml in 250 mM Na citrate, 1 mM EDTA, 2 mM dithiothreitol, pH 5.5. Reactions were incubated at 37 °C for 1 h and then stopped with Roche Complete protease inhibitor mixture supplemented with 1 μM peptatin.

For in situ protease treatment, B. tauricatae Oz1 ectotype A cells were grown in BSK II medium and harvested as described (40). Intact Borrelia cells were treated in situ with trypsin (Roche Molecular Biochemicals) as described (41). Cross-linking of surface-exposed proteins was performed using formaldehyde as described for B. burgdorferi OspC (11).

**Protein Gel Electrophoresis and Immunoblot Analysis**—Proteins were separated on 12% polyacrylamide-SDS gels and visualized by Coomassie Blue staining. For immunoblots, proteins were electrophoretically transferred to nitrocellulose membranes (Immobilon-NC, Millipore) and blocked with 0.02% NaN3, pH 10.0 (AIEX-B), for S34VspB. Cleared lysates were obtained by ultrasonication, centrifugation at 15,000 × g for 10 min and passing through a 0.45-μm pore size filter (Nalgene). Cleared lysates were blocked for 1 h with 10% nonfat dry milk in 0.02% NaN3, pH 10.0 (AIEX-B) or - B column buffers containing 0 to 200 mM NaCl. S34VspA and S34VspB eluted at salt concentrations of 80 and 130 mM, respectively. Peak fractions containing the recombinant proteins were pooled, concentrated using Centricon Plus-20 centrifugal concentrators (Amicon), and dialyzed overnight against deionized H2O at 4 °C. Protein aliquots were flash-frozen on dry ice and stored at −80 °C.

**Circular Dichroism Spectroscopy and Secondary Structure Prediction**—Circular dichroism (CD) spectra were obtained using a Jasco 410 spectropolarimeter. Samples were diluted in 50 mM NaCl, pH 7.5 (TBS), and either air-dried or processed directly. 5% Dry milk in TBS with 0.05% Tween 20 was used for membrane blocking and subsequent incubations for 1 h each. TBS with 0.05% Tween 20 alone was used for washing. Rabbit antiserum was used at 1:500 dilutions, and mouse monoclonal antibody hybridoma supernatants at a 1:10 dilution. Alkaline phosphatase-conjugated Protein A/G (ImmunoPure Protein A/G CIP conjugate, Pierce) and a 1:5000 dilution was used as the second ligand, and a stabilized alkaline phosphatase substrate solution (1-Step NBT/BCIP, Pierce) was used for colorimetric detection.

**N-terminal Protein Sequence Analysis and Mass Spectrometry**—N-terminal protein sequences of recombinant proteins were determined on a Model 477A protein sequenator (Applied Biosystems) from protein samples previously purified by gel electrophoresis and recovery to a polyvinylidene difluoride membrane (43). Molecular masses of purified recombinant proteins were determined by matrix-assisted laser desorption ionization-time of flight on a Voyager biospectrometry workstation (Perseptive Biosystems) using sinapinic acid as matrix.

**TABLE I**

| Name | Nucleotide sequence (5’ to 3’)* |
|------|--------------------------------|
| hisVspA-fwd | AGTGCATTTATTGACTTATAATTATAATACATGCT6CTAATAA | |
| hisVspB-rev | CATATAATGACTCATTATTTATTTATCATGGAACAGTAG | |
| hisVsp26-rev | TATGGTGTTTCATATGGCCAATAATGGAGGGCC | |
| hisOspC-fwd | GGCATCTTATTGTTTTGAGAGGTTT | |
| hisOspC-rev | ATAGTGCGTTATTAATGACTTTATTTTTACATATGTCTAATAATAATGGAGGGCC | |
| S34vpsAB-fwd | CTCAGGATATTCATTAAAGGTITTTTGCACTCTGCG | |
| S34vpsA-rev | GCGGATCTTATTGTAATGGAACAGTAG | |
| S34vpsB-rev | GCGGATCTTATTGTAATGGAACAGTAG | |

* The underlined sequences were added to introduce restriction sites and are noncomplimentary to the Borrelia sequences.
ary structure were obtained using software available on the protein
sequence analysis server at the Biomolecular Engineering Research
Center of Boston University (45).

Gel Filtration Liquid Chromatography—For size determination, pro-
teins were separated on a HiPrep 16/60 Sephacryl S-200 high-resolution
column (Amersham Pharmacia Biotech). The column was first equilibrated with column running buffer (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 0.2% NaN3) and then calibrated separately with blue dextran 2000 (for void volume determination), albumin (67 kDa), chymotryp-
sinogen A (25 kDa), ovalbumin (43 kDa), and ribonuclease A (13.7 kDa) (Amersham Pharmacia Biotech). Approximately 10 mg of concentrated recombinant proteins previously purified by anion exchange chroma-
tography were loaded in 1 ml of column running buffer. For each protein, the partition constant K was calculated from the peak elution volume using the following formula: $K = \frac{V_e}{(V_t - V_e)}$, where $V_e$ = elution volume, $V_t$ = column void volume, and $V_c$ = total bed volume.

Crystallography of S34VspA and preliminary digestion analysis—Recombinant purified S34VspA was dialyzed against deionized H2O, then diluted to 8 mg/ml. Crystals suitable for structure determination by x-ray diffraction methods were obtained using standard vapor dif-
fusion in the hanging droplet method. These crystals were produced in conditions numbers 6 and 22 during initial screening with Hampton Research sparse matrix kit I. These conditions share precipitant and buffer conditions of 30% PEG 4000 and 0.1 M Tris-HCl, pH 8.5, but differ in salt additive of 0.2 M MgCl2 or 0.2 M NaCl, respectively. The precipitant, salt, and pH were varied systematically about these two conditions to optimize crystal size and quality (see "Results"). S34VspA crystals were characterized using a conventional rotating anode x-ray
source and native diffraction data were collected at beamline X12C of the Brookhaven National Laboratory National Synchrotron Light Source on a CCD detector. Diffraction images were processed with DENZO/SCALEPACK (46).

RESULTS

Vsp-OspC Protein Purification—While pathogenesis-associated
functions of Borrelia Vsp proteins, e.g. in antigenic vari-
ation and tissue tropism, have been established, the structural
basis for these functions remain unknown. One question is how
the observed primary protein sequence polymorphisms affect
secondary, tertiary, and quaternary protein structure. Begin-
ing to address this by biochemical and biophysical methods, we first expressed VspA of B. turicatae in E. coli. We chose this protein to focus on because of its association with nervous
system invasion. Two related proteins, VspB of the same strain
of B. turicatae and Vsp26 of B. hermsii, as well as the more
distantly related OspC of B. burgdorferi, were also expressed.

Their conserved signal peptides and overall amino acid identities
of 40 to 70% are representative of the entire Vsp-OspC protein
family (Fig. 1). The proteins were purified in an N-terminal
histidine-tagged form by standard nickel affinity chromatography. In these constructs (hisVspA, hisVspB, his-
Vsp26, and hisOspC), the Vsp-OspC sequences start with resid-
e Asn(20) (Fig. 1), i.e. they lack the signal peptide and acyla-
tion site at Cys(19). The recombinant proteins therefore were
soluble, but otherwise mimic the full-length processed form.

Protease Resistance of Vsp Core Proteins—Previous studies
showed that B. burgdorferi OspA and OspC are highly resis-
tant to trypsin in situ, i.e. when expressed in a native, lipidad
form on the surface of the bacteria (10, 11) and that protease
resistance is an intrinsic property of recombinant OspA (12). To
determine whether VspA and the related proteins are also
protease resistant, we incubated recombinant hisVspA, his-
VspB, hisVsp26, and hisOspC with trypsin at concentrations ranging from 12.5 µg to 400 µg/ml for 1 h at 37 °C. Recombin-
ant, non-lipidated OspA and casein (Sigma) were used as
controls (Fig. 2).

As expected, casein was digested to completion, while recom-
binant OspA was unaffected by trypsin. In contrast to the study
of in situ OspC (11), several independently purified recombi-
nant hisOspC preparations did not exhibit comparable trypsin
resistance in vitro. This difference in results suggests that
OspC is either stabilized by the lipid-mediated attachment to
the cell surface, or that the recombinant hisOspC obtained
from our construct is destabilized by the additional, non-native
sequence at its N terminus.

Although the processed VspA, VspB, and Vsp26 proteins
have 27 to 30 predicted trypsin sites distributed over the whole
length of their protein sequences, the three proteins were
merely truncated 5 to 8 kDa by trypsin digestion (Fig. 2).
Similar results were obtained with plasmin, which has trypsin-
like activity. Yet, to reach an equivalent cleavage efficiency,
plasmin concentrations had to be increased to 1 mg/ml, which
corresponds to about 10 times the physiological concentration
(47). This indicated that most of the sites were not accessible to
the protease, thus yielding a protease-resistant core. No fur-
ther truncations were observed with trypsin concentrations of
up to 10 mg/ml, and VspA maintained its protease-resistant core
even after being boiled for 10 min (not shown). Protein
sequencing of the trypsin-treated VspA, VspB, and Vsp26 pro-
tein cores revealed the N-terminal amino acids in all three

![Figure 1: Amino acid sequence alignment of VspA, VspB, Vsp26, and OspC. The protein sequence alignment was performed using CLUSTAL X.](image-url)
cases to be SDGTV. This sequence can be found starting at a conserved serine at position 34 (Ser34) (Fig. 1).

To purify the protease-resistant cores and for further structural analysis of VspA (see below), we constructed N-terminal truncated versions of VspA and VspB starting at Ser34 (S34VspA and S34VspB). A methionine codon was added to initiate translation. S34VspA and S34VspB proteins were purified to >95% purity from crude E. coli lysates in a single-step anion-exchange chromatography step. Trypsin truncated S34VspA and S34VspB by approximately 1 kDa, revealing a susceptible site at the C terminus as well (Fig. 2). The location of this site, i.e. the C terminus of the core protein, was derived from matrix-assisted laser desorption ionization-time-of-flight mass spectrometry data. The mass of undigested S34VspA was determined as a control, yielding a major peak of 18,507 Da. It corresponds to a protein starting with Ser34 and ending with Asn214 with a predicted size of 18,475 Da. As confirmed by N-terminal sequencing, the f-Met was removed, probably by the E. coli methionyl aminopeptidase (48). For trypsin-treated S34VspA, the mass of the major peak was 17,739 Da, which corresponds to a predicted 17,700 Da protein starting with Ser34 and ending with Lys236. The mass difference of 768 Da would also match the removal of 8 C-terminal amino acids, a predicted 792 Da.

To assess the resistance to other physiologically relevant proteases, we treated the proteins with two lysosomal proteases. Cathepsin B is a broad specificity cysteine endopeptidase with preference for R-R bonds, but also has peptidyl dipeptidase activity, liberating C-terminal dipeptides. Cathepsin D is an aspartic endopeptidase with pepsin A-like activity, cleaving F-V, Q-A, A-L, L-Y, Y-L, G-F, F-F, and F-Y bonds (49). Incubation conditions were derived from the ones used to study antigen presentation of ovalbumin (50). Casein was digested by both proteases, while OspA appeared unaffected. Neither VspA, VspB, Vsp26, nor OspC was cleaved by cathepsin D, although they contain 1 to 4 predicted endopeptidase sites (Fig. 1).

Cathepsin B treatment resulted in several distinct protein bands for all Vsp proteins. Interestingly, the smallest bands for VspA and VspB, presumably the proteolytic end products, were similar if not identical in size to the ones obtained with trypsin. We therefore performed mass spectrometry on cathepsin B-digested S34VspA. The mass of the final proteolytic product, 17,733 Da, was almost identical to trypsin-digested VspA. This suggests that cathepsin B removes 8 amino acids from the C terminus of S34VspA, which would be consistent with its peptidyl dipeptidase activity. Together, these data indicate that Vsp proteins, although differing in primary amino acid sequence, have a conserved core that is resistant to proteases with different specificities.

Vsp-OspC Protein Dimerization In Solution—Huang et al. (38) have recently shown that a truncated recombinant form of B. burgdorferi OspC forms a 37-kDa dimer in solution. To see if this is a common structural feature of the Vsp-OspC protein family, we determined the size of the recombinant proteins by gel filtration. Purified recombinant proteins were loaded on a high-resolution Sephacryl column, which had been calibrated with protein size standards. The apparent molecular masses were calculated from the partition coefficients to be as follows: hisVspA, 43.5 kDa; hisVspB, 42.5 kDa; hisVsp26, 49.2 kDa; hisOspC, 49.2 kDa; S34VspA, 38.5 kDa; S34VspB, 40.5 kDa (see Fig. 3). The predicted monomer sizes for the recombinant proteins range from 18.6 kDa (S34VspA) to 23 kDa (hisVsp26). This shows that VspA, like the other relapsing fever Vsp proteins and the related OspC, forms dimers in solution.

VspA Dimerization and Protease Resistance In Situ—In a next step, we investigated whether the dimerization and protease resistance of Vsp proteins could also be observed in situ, i.e. on the surface of intact Borrelia cells. Dimerization has recently been shown for in situ B. burgdorferi OspC (11). We applied the same approach to VspA expressed by B. turicatae Oz1 serotype A. Cross-linking of closely associated surface proteins by formaldehyde revealed that VspA dimers form also in situ (Fig. 4A).

We next assessed whether the native VspA protein dimers also exhibited the protease-resistant properties observed with the recombinant protein. Accessibility of the C-terminal trypsin sites would lead to a truncated, yet cell-associated protein. If the N-terminal sites in both dimer subunits were accessible as well, the protein parts containing the two lipid anchors would be cleaved and the cores would be released into the supernatant. We therefore incubated intact B. turicatae cells with trypsin and then analyzed the pelletted cell-associated proteins and solubilized proteins in the supernatant. VspA was detected using an anti-VspA polyclonal antibody (Fig. 4B). In untreated B. turicatae cell pellets, VspA was found in its full-length lipidated form running at about 22 kDa. No protein was detectable in the supernatant. In the cell-associated fraction of trypsin-treated B. turicatae cells however, two additional forms of VspA approximately 20 and 18 kDa in size were detected, with the 20-kDa band being the most prominent. In the supernatant, only the 18-kDa band was detected. The 20-kDa band likely represents the C-terminal truncated lipidated protein, while the 18-kDa band corresponds in size to the VspA core. This indicates that both the N- and C-terminal domains were accessible to proteases such as trypsin. It also suggests that the VspA core can still be anchored to the surface via association with a lipidated dimer subunit, and that release into the surrounding environment occurs only after cleavage of both membrane anchor-containing peptides.

Conserved Vsp-OspC Secondary Structure—We used CD spectroscopy as a first approach to compare the secondary structures of the recombinant Vsp and OspC (Fig. 5A). The obtained mean residue ellipticity values were plotted against the wavelength. Peaks at 192 nm and two troughs at 208 and 222 nm are indicative of predominantly α-helical secondary structures. The CD spectra of VspA, VspB, Vsp26, and OspC overlap extensively, thus the high helix content appears conserved among the Vsp-OspC protein family.
protein standard values.

B. turicatae tact the log of the proteins’ predicted molecular weights (\( \text{MW} \)) were calculated based on the peak elution volumes and plotted against the log of the proteins’ predicted molecular weights (\( \text{MW} \), in kDa). The diagonal line represents a logarithmic regression derived from the protein standard values.

**Fig. 3.** Dimerization of Vsp-OspC proteins in solution. Recombinant proteins previously purified by anion exchange chromatography were separated by gel filtration liquid chromatography. Albumin, chymotrypsigen A, ovalbumin, and ribonuclease A (Amersham Pharmacia Biotech) served as size standards. \( K_a \), partition coefficient values were calculated based on the peak elution volumes and plotted against the log of the proteins’ predicted molecular weights (\( \text{MW} \), in kDa). The diagonal line represents a logarithmic regression derived from the protein standard values.

**Fig. 4.** VspA dimerization and protease resistance in situ. Intact B. turicatae serotype A (B.T. A) cells expressing VspA were incubated with either (A) 1% formaldehyde (form.) or (B) 200 \( \mu \text{g/mL} \) trypsin (tryp.). Total cellular proteins were separated on a 12% SDS-polyacrylamide electrophoresis gel, transferred to nitrocellulose membranes, and probed with anti-VspA polyclonal rabbit antisera. For the trypsin protease assay, both the pelleted, cell-associated proteins (p) and the reaction supernatant (s) were loaded in equivalent volumes. S34VspA protein was also included in the reactions as a size control. Sizes of protein markers in kDa are indicated at the left of each panel. In panel A, labeled arrows to the right indicate the sizes of VspA monomers and dimers derived from the size standards. In panel B, L, indicates lipated, NL, non-lipidated forms of VspA.

Spectral data lead to estimates of secondary structure ratios. For histidine-tagged proteins, the mean values for \( \alpha \)-helical, \( \beta \)-sheet, and coiled secondary structure were 71, 2, and 27%, respectively. For the N-terminal truncated proteins, the coiled structure content decreased to 16%, with \( \alpha \)-helix and \( \beta \)-sheet ratios of 80 and 4%.

As we currently lack further secondary or tertiary structure information on the VspAs, we resorted to computer algorithms for further comparisons. While a comparison of predicted secondary structure cores confirmed the experimental data of highly \( \alpha \)-helical proteins, they also revealed further conservatism. All studied proteins are calculated to fold into 4 helices separated by turns (Fig. 5B). Furthermore, the N and C termini likely form flexible, unordered tails of approximately 20 and 10 amino acids, respectively. Both these structural features are in accordance with the NMR-derived data for OspC (38).

**Crystallization of S34VspA**—To date, B. burgdorferi OspA is the only Borrelia lipoprotein whose crystal structure has been determined (13–15). Initial attempts to crystallize the histidine-tagged forms of VspA and VspB were unsuccessful. We surmised that the predicted flexible, unordered N- and C-terminal sequences could inhibit crystallization. In our next set of experiments, we therefore used S34VspA, which lacks most of the flexible N-terminal tail.

Large single crystals of S34VspA were obtained at either room temperature or 4 °C using 28% (w/v) polyethylene glycol 4000, 80 mM Tris/HCl, pH 8.5, 15% (v/v) glycerol, and either 100–200 mM nickel chloride (Fig. 6) or magnesium formate (not shown). Nucleation was controlled by reducing the initial concentration of S34VspA to 6–8 mg/mL. S34VspA/nickel chloride crystals belong to the monoclinic space group C2, and have unit cell dimensions \( a = 241.8 \, \text{Å}, \, b = 69.1 \, \text{Å}, \, c = 87.6 \, \text{Å}, \, \text{and} \, \beta = 104.9^\circ \). S34VspA/magnesium formate crystals belong to space group P2(1), and have unit cell dimensions \( a = 121.8 \, \text{Å}, \, b = 68.7 \, \text{Å}, \, c = 87.5 \, \text{Å}, \, \text{and} \, \beta = 103.4^\circ \). The volume of the asymmetric unit is quite similar in both crystal forms. We anticipate that four VspA dimers are accommodated with approximately 50% solvent content.

We have collected two complete diffraction data sets at 100 K. Dataset I (S34VspA/nickel chloride) has 38,032 reflections merged from 130,123 individual measurements and is 98.9% complete to a resolution of 2.7 Å, with a linear R-merge of 4.2% (14.7% for the highest resolution shell, 2.8–2.7 Å). Dataset II (S34VspA/magnesium formate) has 27,615 reflections merged from 84,782 individual measurements and is 96.2% complete to a resolution of 3.0 Å, with a linear R-merge of 6.7% (20.9% for the highest resolution shell, 3.11 - 3.0 Å). In both cases the mosaic spread of the crystal was determined to be 0.95°. We have recently succeeded in obtaining selenomethionine-substituted crystals of S34VspA for structure determination using the multiple anomalous dispersion method (51).

**DISCUSSION**

This study has shown that the neurotropism-associated VspA of B. turicatae shares secondary, tertiary, and quaternary structural characteristics with other Borrelia Vsp-OspC family proteins, despite as little as 40% identity in primary sequence. CD spectroscopy demonstrated conserved, highly \( \alpha \)-helical secondary structures that are predicted to fold into a four-helix bundle. Supporting this model were the findings that postswitch mutations in expressed vsp alleles are mainly conservative (30, 52), and that non-conservative amino acid changes, insertions, and deletions cluster in regions outside the predicted helices (Fig. 1). We conclude that localized polymorphisms, e.g. variable loops, rather than differences in overall domain structures determine specific Vsp-OspC functions such as antigenic variation and niche selection.

Further validating our model of structural conservation were experiments showing that the proteins had same size protease-resistant cores and dimerized in a physiologic solution as well as on the bacterial surface. These shared characteristics are likely a consequence of a common protein fold and conserved residues. Supporting this, the Vsp cores coincide with the predicted flexible unordered N- and C-terminal tails appear susceptible to proteolysis. In the case of trypsin, the enzyme active site must obtain access to the target peptide bond (K-X) to accomplish cleavage. This can only be achieved in regions where the main chain atoms are not involved in formation of stable secondary structural elements (53). In OspA, the lysines are in antiparallel \( \beta \)-sheet hydrogen bonding that make these trypsin sites inaccessible. On the other hand,
lysines in the loops may be shielded from proteolytic attack by dimer subunit interactions. A comparable mechanism has been observed with two *B. burgdorferi* surface proteins: OspA closely interacts with an integral membrane protein, P66, thereby protecting a surface-exposed loop of P66 from proteolysis (11). The common, intrinsic property of Vsp-OspC proteins to dimerize appears mediated by sequences present in the protein cores. This conservation despite significant divergence between the Vsp and OspC proteins suggests that some of the invariant core residues participate in this process. In the absence of further structural information, a stretch of constant amino acids found at the C-terminal end of the predicted first a-helix (Fig. 1) is currently our strongest candidate for a dimerization domain.

The effect of Vsp-OspC core features on the proteins’ biological functions could be manifold. Various tick organs have proteases (54, 55), thus protease resistance would be of benefit during the migration of the spirochetes from the midgut through the hemolymph to the salivary glands (1). The resistance to trypsin and plasmin might also serve the spirochetes in their vertebrate hosts. After host plasminogen acquisition and activation, surface bound plasmin facilitates the dissemination of bacteria in the tick and fosters heart and brain invasion in infected mice (56, 57).

While the lipid moiety of *Borrelia* lipoproteins such as *B. burgdorferi* OspA is a determinant of mitogenicity (58), the common structural features of Vsp-OspC proteins might modulate host immune responses as well. An efficient peptide antigen presentation to T cells via the major histocompatibility complex requires the limited proteolytic processing of exogenous proteins by endosomal and lysosomal proteases (59). Experiments with specific protease inhibitors and protease gene knockout mice or the elimination of protease sensitive sites have shown this (60). *Borrelia* spirochetes may have evolved the latter strategy, rendering their major antigens resistant to proteolytic attack and thus leading to the persistence of Vsp-OspC core proteins in infected mammals. Furthermore, the multimerization of the highly abundant Vsp-OspC proteins might produce an array that is analogous to viral capsid particles (61). Both antigen persistence and repetitive structure have been associated with a T-independent response (62, 63). Intriguingly, this type of immune response has been observed in the clearance of *B. turicatae* infection (64) and was important for protective immunity and resolution of Lyme disease in an animal model (65). The current study has focused on the activity of a limited selection of proteolytic enzymes in vitro, and further immunological studies will be needed to comprehensively address the possible effects of Vsp protease resistance on antigen presentation.

The mechanisms of antigenic variation in African trypanosomes and relapsing fever *Borrelia* are strikingly similar (66). A comparison of the known structural features of these proteins reveals further convergences. VSGs also share common, highly a-helical folds and form stable dimer units in solution (67). Both VSGs and Vsp-Vlps are anchored to the cell’s surface via post-translational modifications, the VSGs via a C-terminal glycosylphosphatidylinositol anchor (68), and the Vsp-Vlps, like other *Borrelia* lipoproteins (8, 9), via an N-terminal triacyl-modified cysteine. Interestingly, VSG glycosylphosphatidylinositol anchors are hydrolyzed by an endogenous trypanosomal phospholipase under stress conditions, which leads to VSG...
sheding (69). This may expedite the exchange of coat proteins during an antigenic switch, and at the same time divert the host immune system from live trypanosomes expressing the same VSG. A similar mechanism has been recently observed with the oral pathogen Streptococcus mutans, which expresses an enzyme that releases the adhesin P1 from its cell wall (70). This activity might allow the bacterium to shed P1-antibody complexes or to recolonize after detaching from an adherent surface.

The structure-function analysis of the Vsp and Vlp proteins will likely yield important clues in our understanding of antigenic variation and tissue tropism. By integrating the structural information gained from the VspA crystals obtained here with data from in vitro immunological, biochemical, and cell biological studies, we will be able to define the domains interacting with antibodies and other host factors.

Acknowledgements—We are grateful to Naomi Sayano, Sharlene Lim, and Andrew Glenn for expert technical assistance. We also thank Bob Huestis (Griffith water, PA) for recombinant OspA protein, Saith Jayasinghe (Dept. of Physiology and Biophysics, University of California at Irvine) for instructions regarding CD spectroscopy, and Vance Cao (Biological Sciences Core Facility, University of California at Irvine) for performing the N-terminal sequencing and mass spectroscopy. C. L. Huebner and H. M. Berman for generously providing laboratory space and access to equipment at Rutgers University. The Brookhaven National Laboratory beamline X12C of the National Synchrotron Light Source is part of a macromolecular crystallography resource supported jointly by the Department of Energy, National Institutes of Health, and the National Science Foundation.

REFERENCES

1. Barbour, A. G., and Hayes, S. F. (1986) Microbiol. Rev. 50, 381–400
2. Genn, S. Palmer, N., van Vuig, R., H. N., L. B., D., Roux, P., A., M., L., B., R., T., T., R., D., H., D., Hickey, G., E., Gwinn, M., White, O., and Fraser, C. M. (1998) Methods 35, 155–171
3. Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R., Lathigra, R., Lathigra, R., S., L., D., D., H., D., Hickey, G., E., Gwinn, M., White, O., and Fraser, C. M. (1998) Microbiol. Rev. 62, 283–305
4. Hendrickson, W. A., and Ogata, C. M. (1997) in Methods in Enzymology (Carter, C. W., and Sweet, R. M., eds) Vol. 276, pp. 307–326, Academic Press, San Diego
5. Coleman, J. L., Roemer, E. J., and Benach, J. L. (1999) Infect. Immun. 67, 3929–3936
6. Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., and Blanquet, S. (1989) J. Bacteriol. 170, 2837–2847
7. Stultz, C. M., Nambudripad, R., Lathrop, R. H., and White, J. V. (1997) in Methods in Enzymology (Carter, C. W., and Sweet, R. M., eds) Vol. 276, pp. 307–326, Academic Press, San Diego
8. Freymann, D., Down, J., Carrington, M., Roditi, I., Turner, M., and Wiley, D. E. (1998) J. Mol. Biol. 281, 2874–2883
9. Soutschek, E., Will, G., and Wilke, B. (1993) Med. Microbiol. Immunol. 182, 57–50
10. Lin, L., Gibbs, C. P., Schuster, R., and Dorner, F. (1995) Mol. Microbiol. 18, 257–269
11. Wang, N. I., Dykhuijzen, D. E., Qin, W. G., Dunn, J. J., Bosler, E. M., and Luft, B. J. (1999) Genetics 151, 15–30
12. Dunn, J. J., Lade, V., and Barbour, A. G. (1990) Infect. Immun. 66, 1219–1225
13. Pennington, P. M., Cadavid, D., Bunikis, J., Norris, S. J., and Barbour, A. G. (1999) Mol. Microbiol. 34, 1120–1132
14. Barbour, A. G., and Restrepo, B. I. (2000) Emerg. Infect. Dis. 6, 449–457
15. Deitsch, K. W., Moxon, E. R., and Wellemes, T. E. (1997) Microbiol. Mol. Biol. Rev. 61, 281–293
16. Cadavid, D., Bunodic, V., and Barbour, A. G. (1993) J. Infect. Dis. 168, 445–448
17. Magoun, L., Zuckert, W. R., Robbins, D., Parveen, N., Alugapalli, K. R., Schwan, T. G., Barbour, A. G., and Leong, J. M. (2000) Mol. Microbiol. 36, 866–877
18. Huang, X., L., Link, K., Koea, D., Dunn, J. J., Lu, B. J., and Koide, S. (1999) J. Biomol. NMR 14, 283–284
19. Thomas, D. M., Cadavid, D., and Barbour, A. G. (1994) J. Infect. Dis. 169, 445–448
20. Ras, N. M., Postic, D., Ave, P., Huerre, M., and Baranton, G. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 155–171
21. Pennington, P. M., Allred, C. D., West, C. S., Alvarez, R., and Barbour, A. G. (1997) Infect. Immun. 65, 285–292
22. H. M., Berman for generously providing laboratory space and access to equipment at Rutgers University. The Brookhaven National Laboratory beamline X12C of the National Synchrotron Light Source is part of a macromolecular crystallography resource supported jointly by the Department of Energy, National Institutes of Health, and the National Science Foundation.