Linker Mutagenesis of the *Caulobacter crescentus* S-Layer Protein: toward a Definition of an N-Terminal Anchoring Region and a C-Terminal Secretion Signal and the Potential for Heterologous Protein Secretion

WADE H. BINGLE, JOHN F. NOMEILLINI, AND JOHN SMIT*

Department of Microbiology and Immunology, The University of British Columbia,
Vancouver, British Columbia, Canada

Received 30 July 1996/Accepted 10 November 1996

Linker insertion mutagenesis was used to modify the paracrystalline surface layer (S-layer) protein (RsaA) of the gram-negative bacterium *Caulobacter crescentus*. Eleven unique *Bam*HI linker insertions in the cloned *rsaA* gene were identified; at the protein level, these linker insertions introduced 4 to 6 amino acids at positions ranging from the extreme N terminus to the extreme C terminus of the 1,026-amino-acid RsaA protein. All linker-peptide insertions in the RsaA N terminus caused the secreted protein to be shed into the growth medium, suggesting that the RsaA N terminus is involved in cell surface anchoring. One linker-peptide insertion in the RsaA C terminus (amino acid 784) had no effect on S-layer biogenesis, while another (amino acid 907) disrupted secretion of the protein, suggesting that RsaA possesses a secretion signal lying C terminal to amino acid 784, near or including amino acid 907. Unlike extreme N- or C-terminal linker-peptide insertions, those more centrally located in the RsaA primary sequence had no apparent effect on S-layer biogenesis. By using a newly introduced linker-encoded restriction site, a 3' fragment of the *rsaA* gene encoding the last 242 C-terminal amino acids of the S-layer protein was expressed in *C. crescentus* from heterologous *Escherichia coli lacZ* transcription and translation initiation information. This C-terminal portion of RsaA was secreted into the growth medium, confirming the presence of a C-terminal secretion signal. The use of the RsaA C terminus for the secretion of heterologous proteins in *C. crescentus* was explored by fusing 109 amino acids of an envelope glycoprotein from infectious hematopoietic necrosis virus, a pathogen of salmonid fish, to the last 242 amino acids of the RsaA C terminus. The resulting hybrid protein was successfully secreted into the growth medium and accounted for 10% of total protein in a stationary-phase culture. Based on these results and features of the RsaA primary sequence, we propose that the *C. crescentus* S-layer protein is secreted by a type I secretion system, relying on a stable C-terminal secretion signal in a manner analogous to *E. coli* α -hemolysin, the first example of an S-layer protein secreted by such a pathway.

The outer membrane of the dimorphic gram-negative bacterium *Caulobacter crescentus* is covered by a protein surface layer (S-layer) composed of a single 1,026-amino-acid protein termed RsaA (13). The S-layer is paracrystalline in nature, exhibiting an array of ring-like subunits (each composed of six copies of RsaA) arranged on a lattice with p6 symmetry and interlinked at the threefold rotational axis (57); proper crystallization of the S-layer is dependent on Ca^{2+} ions (62, 63). The association between RsaA and the outer membrane is not completely understood, but the protein appears to be anchored to the outer membrane via noncovalent interactions with a specific smooth lipopolysaccharide (LPS) molecule (62). So far, the only known function of the *C. crescentus* S-layer is to protect cells against predation by a *Bdellovibrio*-like organism (38).

Although RsaA must traverse two membranes to arrive at its final location, the protein is not subjected to any posttranslational processing beyond the removal of N-formyl methionine, yielding the mature protein of 1,025 amino acids (31). In previous studies, we used deletion analysis to identify secretion-related regions of the protein (9, 10). These studies showed

that the N terminus is dispensable as far as secretion is concerned. Indeed, portions of the N terminus within the first 775 N-terminal amino acids could be removed from RsaA without preventing secretion of the remainder of the protein. In contrast to N-terminally deleted RsaA, C-terminally truncated versions of the protein were not secreted. Further, these truncated species were apparently degraded after synthesis because intracellular accumulation of secretion-incompetent molecules was not observed. Significantly, although deletion of the RsaA N terminus did not disrupt secretion, it did disrupt anchoring of the protein to the cell surface and its crystallization into a regular array.

A potential problem with the interpretation of deletion data is the possibility that such major modifications cause large-scale perturbations to protein structure or render the protein more susceptible to proteolytic degradation rather than simply removing regions responsible for S-layer secretion, anchoring, and crystallization. In this study, we have used linker insertion mutagenesis of the *C. crescentus* S-layer protein gene to more precisely identify functional regions of the protein. In contrast to deletion analysis, linker mutagenesis yields two- to four-amino-acid-insertion mutations and has the potential to produce only local disruptions of protein structure without serious polar effects. Using this method, we have more directly implicated the N terminus of the protein as important for surface anchoring and localized the secretion signal to the extreme C

* Corresponding author. Mailing address: Department of Microbiology and Immunology, The University of British Columbia, #300, 6174 University Blvd., Vancouver, B.C., Canada V6T 1Z3. Phone: (604) 822-4417. Fax: (604) 822-6041. E-mail: jsmit@unixg.ubc.ca.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Reference or source
Bacterial strains		
<i>C. crescentus</i>		
JS4000	Spontaneous RsaA [−] mutant of strain CB2 (ATCC 15252); formerly CB2A	56
JS3001	Shedder mutant of JS3000 (formerly CB2NY66R); lacks a smooth LPS species in the outer membrane, resulting in the shedding of RsaA into the growth medium; formerly CB2NY66Rmgl	62
<i>E. coli</i>		
DH5α	<i>recA endA</i>	Life Technologies 19
RB404	<i>dam</i>	
Plasmids		
pKT215 and pKT210	RSF1010-derived cloning vectors; Cm ^r Sm ^r	3
pWB9	pKT215-derived vector incorporating the <i>rsaA</i> promoter; Cm ^r Sm ^r	12
pBR322	ColE1 cloning vector; Tc ^r Ap ^r	51
pUC8 and pUC9	pBR322-derived cloning/ <i>lacZ</i> -based expression vectors; Ap ^r	61
pTZ18U	A phagemid version of pUC18; Ap ^r	Amersham Life Science
pUC13: <i>rsaA</i>	Source of <i>rsaA</i> gene; <i>Hind</i> III/ <i>Bam</i> HI A19 fragment of <i>C. crescentus</i> chromosomal DNA (4.3 kb) cloned into pUC13; Ap ^r	31
pUC1021K	pUC9 vector carrying a <i>Bam</i> HI linker tagged with a Km ^r gene from Tn903; Ap ^r Km ^r	11
pUC4 KSAC	pUC4 carrying a Km ^r gene from Tn903 flanked by a symmetrical polylinker; Ap ^r Km ^r	4
pKT215: <i>rsaA</i> (<i>Mlu</i> I93 <i>Bam</i> HI)	<i>rsaA</i> carrying a <i>Bam</i> HI linker insertion at an <i>Mlu</i> I site corresponding to amino acid 93 of the primary <i>rsaA</i> translation product; Cm ^r Sm ^r . In reference 9, the location of the linker insertion was designated with respect to the mature protein. To maintain a consistent naming system with respect to Gilchrist et al. (31), the location of the linker insertion is given here with respect to the primary translation product.	9
pCHR64	Source of a <i>Bgl</i> II fragment carrying the promoterless Cm ^r gene of Tn9.	53
pG8	Source of IHNVG gene	65

terminus. Further, we show that the C-terminal secretion signal of RsaA has the potential to facilitate the secretion of large heterologous proteins from *C. crescentus*.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Strains of *C. crescentus* and *Escherichia coli* (Table 1) were grown in liquid peptone-yeast extract (PYE) medium and modified Luria-Bertani medium, respectively, as previously described (9); agar was added to 1.2% for solid media.

C. crescentus and *E. coli* were prepared for electroporation as described by Gilchrist and Smit (32). All plasmids introduced into *C. crescentus* were constructed from the cloning vector pKT215 or pKT210 (Table 1) and retained both the streptomycin and chloramphenicol markers of these plasmids. In all electroporation experiments, plasmid-containing *C. crescentus* cells were selected initially on solid medium containing streptomycin, and a colony was then transferred to liquid medium containing chloramphenicol for subsequent work (12). Antibiotics were included at the following concentrations: ampicillin, 50 μ g/ml; kanamycin, 50 μ g/ml; and chloramphenicol, 2 μ g/ml for *C. crescentus* and 20 μ g/ml for *E. coli*.

Recombinant DNA methods, nucleotide sequencing, and PCR. Plasmids used as raw materials for the constructions described in this study are listed in Table 1. Crude miniprep plasmid DNA extracted from *E. coli* was used for electroporation (see above) or as substrates for DNA-modifying enzymes (9). Unless otherwise indicated, DNA-modifying enzymes were purchased from Life Technologies (Burlington, Ontario, Canada) or New England Biolabs (Mississauga, Ontario, Canada) and were used according to the manufacturer's instructions or other recommendations (51). Agarose gel electrophoresis was performed by standard methods (51); in preparative applications, DNA fragments were purified by using Gene Clean (Bio 101, Mississauga, Ontario, Canada) or QIAEX II (Qiagen Inc., Chatsworth, Calif.). Dideoxynucleotide sequencing (52) was accomplished with a Sequenase kit (Amersham Life Science, Mississauga, Ontario, Canada). Oligonucleotides were purchased from New England Biolabs or were synthesized in-house as described previously (31); oligonucleotides synthesized in-house were purified by butanol extraction (54). DNA colony hybridizations were done by using Whatman 541 filter paper as a support in place of nitrocellulose (45); radiolabeled probes were prepared by nick translation or fill-in of 5'-single-stranded extensions of annealed oligonucleotides with the Klenow fragment of *E. coli* DNA polymerase I (PolIK). PCR amplification was conducted by

using *Taq*I DNA polymerase, as instructed by the manufacturer (Life Technologies) and a Techne PH3 Thermal Cycler (Mandel Scientific Co., Guelph, Ontario, Canada).

SDS-PAGE and Western analysis. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and Western immunoblot analysis were conducted as previously described (63). All samples containing RsaA were solubilized at 37°C, rather than boiling, in SDS-PAGE solubilization buffer prior to electrophoresis (56). After transfer of proteins to nitrocellulose, the blots were probed with an anti-RsaA polyclonal antibody (63) or a monoclonal antibody (MAb 136J) against the 109-amino-acid segment of the envelope glycoprotein of infectious hematopoietic necrosis virus (IHNV) (65). Antibody binding was visualized by using goat anti-rabbit or mouse serum coupled to horseradish peroxidase and color-forming reagents (56).

Linker mutagenesis the *C. crescentus* S-layer protein gene. A version of the *rsaA* gene (*rsaA*ΔP) lacking its native promoter and carried on an 3.7-kb *Eco*RI/*Sst*I fragment in plasmid pTZ18UB was used for linker mutagenesis experiments (Fig. 1). To create target sites for linker insertions, pTZ18UB:*rsaA*ΔP and those propagated in *E. coli* RB404 was partially digested with *Taq*I (Fig. 2), and those molecules potentially corresponding to full-length linearized plasmid were recovered from agarose gels and ligated to a *Bam*HI linker (*Bam*HI-1021K) previously tagged with the 1.2-kb Km^r gene of Tn903 (11). The tagged *Bam*HI linker was excised from its carrier plasmid with *Acc*I to provide it with 5'-CG extensions compatible with those left by digestion of *rsaA*ΔP with *Taq*I. *E. coli* DH5 α transformant colonies exhibiting kanamycin and ampicillin resistance were pooled, and a library of plasmid DNA containing *Bam*HI linker insertions was prepared from this colony material. To separate linker insertions in the *rsaA*ΔP gene from those in the plasmid vector, the library of 4.9-kb *rsaA*ΔP *Eco*RI/*Sst*I fragments carrying tagged linker insertions was purified from agarose gels and ligated to fresh pTZ18UB (the Km^r gene of Tn903 contains no *Eco*RI or *Sst*I sites). Finally, the Km^r gene was removed by *Bam*HI digestion, and the plasmid was circularized by using T4 DNA ligase; neither pTZ18UB, *rsaA*ΔP, nor the Km^r gene of Tn903 contains *Bam*HI sites.

When RsaA was synthesized by *E. coli* DH5 α from the *lacZ* α promoter, colonies acquired a distinctive morphology, surface texture, and density; on a microscopic level, the cells became highly filamentous (sometimes 200 to 300 μ m in length) and RsaA accumulated as inclusion bodies. These features were used as an initial screen to ensure that the linker mutagenesis protocol had not resulted in disruption of the *rsaA* reading frame downstream of the linker insertion site through the introduction of small deletions that changed the reading frame. Plasmid DNA was prepared from 58 individual cultures of filamentous

FIG. 1. Construction of the *rsaADP* gene and the expression vector pWB9KSAC. (A) The *rsaA* gene was separated from its promoter and transcription start site by initially manipulating a 5' portion of the gene carried in plasmid pSa5 rather than the entire gene. Plasmid pSa5 was created by inserting the 1.6-kb *HindIII*/*AvaI* fragment of *rsaA* into the same sites of pBR322. pSa5 was linearized with *HindIII* and then partially digested with *Bss*HII. Those fragments resulting from cleavage at the central *Bss*HII site were purified, their cohesive termini were polished with PolIK, and the resulting blunt ends were ligated, forming pSa5ΔP. The *Clal* site of the pBR322 portion of pSaΔP (propagated in *E. coli* DH5α [Dam⁺]) was destroyed by digestion with *Clal*, polishing with PolIK, and religation, forming pSa5ΔP2. (The *Clal* site of *rsaA* is methylated in *dam*⁺ strains and is therefore protected from cleavage by *Clal*.) The *rsaADP* gene was constructed via a tripartite ligation in which the promoterless 5' *rsaA* *EcoRI*/*Clal* fragment from pSa5ΔP2 was mixed with a 3' *Clal*/*SstI* *rsaA* fragment from pUC13:*rsaA* and the large fragment of *EcoRI*/*SstI*-digested pTZ18UB. Finally, the remaining *Bam*HI site was destroyed by digestion with *Bam*HI, polishing with PolIK, and religation. The result of these manipulations was to produce a promoterless version of *rsaA* carried on an *EcoRI*/*SstI* fragment in pTZ18UB which contained no *Bam*HI sites. The *Bam*HI site of pTZ18UB was destroyed (and converted to a *Clal* site) by *Bam*HI digestion, polishing with PolIK, and religation, yielding pTZ18UB. (B) Vector pWB9KSAC is a pKT215-based expression vector carrying *rsaA* transcription initiation information on a *HindIII*/*EcoRI* fragment. pWB9KSAC was constructed from pWB9 by replacing its multiple cloning site with a portion of the symmetrical polylinker from pUC4KSAC. This was done by inserting the KSAC Km^r gene cassette into the *EcoRI* site of pWB9 and collapsing the gene cartridge by using *Bam*HI. This removed the *SstI* site of the pWB9 multiple cloning site and facilitated insertion of the *rsaADP* gene downstream of the *rsaA* promoter and transcription start site as an *EcoRI*/*SstI* fragment. Abbreviations: Bm, *Bam*HI; A, *AccI*; C, *Clal*; E, *EcoRI*; Hd, *HindIII*; M, *MspI*; P, *PstI*; Sm, *SmaI*; Ss, *SstI*; SD, Shine-Dalgarno; Pm, promoter; ts, transcription start; MCS, multiple cloning site.

cells exhibiting inclusion bodies, and the approximate locations of the *Bam*HI linker insertions in the *rsaA* gene were determined by restriction mapping. Of the 45 insertions subsequently found to map in the coding region of the gene, 35 either were associated with deletions or appeared to be represented more than once in the collection of linker insertion mutants. The positions of the remaining 10 linker insertions in the coding region were determined by nucleotide sequenc-

ing by using a previous collection of primers (31). Finally, mutant versions of the *rsaAΔP* gene carrying in-frame *Bam*HI linker insertions were transferred to expression vector pWB9KSAC as *Eco*RI/*Sst*I fragments and electroporated into *C. crescentus*.

Analysis of cell-associated and cell-free S-layer protein. Cell-associated S-layer protein was extracted from the cell surface by mild acid treatment (63).

FIG. 2. Linker insertion mutagenesis of *rsaA*ΔP, locations of linker-peptide insertions in RsaA, and their effects on S-layer biogenesis. A *Bam*HI linker was tagged with a *Km^r* gene, and *TaqI* sites of known location were used as target sites. The tagged *Bam*HI linker was excised from its carrier plasmid with *AccI* to provide it with 5'-CG extensions compatible with those created by digestion of *rsaA*ΔP with *TaqI*. After insertion of the linker, the *Km^r* gene was removed by using *Bam*HI. The effect of each linker peptide insertion on S-layer biogenesis, as well as the deduced amino acid sequence of each linker-peptide insertion, is indicated in each case. The positions of the linker-peptide insertions in the primary translation product are specified by the number of the amino acid codon corresponding to the first base of the target restriction site. The superscript (a, b, or c) associated with each insertion point indicates which of the three linker peptides was inserted into RsaA. The shaded region represents the aspartate-containing glycine-rich repeat portion of RsaA. ND, not determined.

Briefly, samples of cells (1.5 ml) obtained from early-stationary-phase cultures (optical density at 600 nm of 1.5) were collected by centrifugation (12,000 × g, 5 min), resuspended in 100 μl of 100 mM HEPES (pH 2) by vortexing, and held at room temperature. After 5 min, the cells were pelleted by centrifugation and the supernatant was analyzed for the presence of RsaA. When low-pH extraction was used, 2 μl of 1 N NaOH was added per 20 μl of extracted material prior to analysis by SDS-PAGE.

When RsaA fails to attach to the cell surface, the protein forms an insoluble macroscopic particulate material in the culture fluids (9, 62). The formation of this particulate material is promoted by a large culture volume relative to flask volume (e.g., 30%), slow shaking speeds (e.g., 100 rpm), and growth to stationary phase. Particulate RsaA present in such cultures was recovered by gravity filtration through a disk of nylon mesh (2-cm diameter) fitted into a microfiltration apparatus; the mesh contained square pores (diagonal dimension, 350 μm). The particulate material collecting on the mesh was washed with distilled water in the filtration apparatus, and then the nylon mesh disc was transferred to a 1.5-ml microcentrifuge tube, covered with 1 ml of distilled water, and vortexed briefly to dislodge the protein from the disc. The tube was subjected to a brief (10 s; 12,000 × g) centrifugation to pellet the particulate RsaA protein dislodged from the filter. After removal of the filter and supernatant from the tube, the pelleted material was solubilized in 4 volumes of 8 M urea-100 mM Tris-HCl (pH 8.5) and used directly for SDS-PAGE analysis and Western immunoblotting.

To simplify analysis of cell-associated and cell-free RsaA, cells recovered from solid medium were used in some experiments. Initially, cells from PYE-chloramphenicol liquid medium were plated on PYE-streptomycin solid medium to produce between 300 and 500 colonies per plate. After 3 days of growth, a thick cell paste was formed on the surface of the agar by mixing the colonial material with 200 μl of 100 mM Tris-HCl-10 mM Na₂EDTA-500 μg of lysozyme per ml (pH 8), using a bent glass rod plate spreader. After cell lysis, 2 ml of 8 M urea-100 mM Tris-HCl (pH 8.5) was added to the lysate (still on the surface of the solid medium) to solubilize the cell debris (including any particulate RsaA

present). Before analysis by SDS-PAGE and Western immunoblotting, the viscosity of lysate was reduced by passing it through an 18-gauge needle.

Protein was estimated by the method of Markwell et al. (42), using bovine serum albumin as a standard. Whole cells and proteins were solubilized in 1% SDS at 65°C prior to assay. If samples contained urea, it was removed by overnight dialysis (Spectra/Por dialysis tubing; 3,500-molecular-weight cutoff; VWR Scientific, Toronto, Ontario, Canada) against distilled water (4°C).

N-terminal amino acid microsequencing. N-terminal amino acid sequences were determined by Edman degradation, using an Applied Biosystems 47XA automatic sequencer. Proteins for amino acid sequencing were first purified by SDS-PAGE, electrophoretically transferred to a polyvinylidene difluoride membrane, and stained with Coomassie blue R-250 as described by the manufacturer.

In vitro crystallization assays and electron microscopy. The ability of specific modified RsaA constructs to crystallize into a hexagonal array was assessed by an in vitro assay (43). In short, the assay combined vesicles, composed of phospholipids and the species of *C. crescentus* LPS involved with S-layer anchoring (62), RsaA protein prepared as described above, and calcium ions. Vesicles exhibiting the crystallized S-layer were viewed by negative-stain (2% ammonium molybdate) electron microscopy as previously described (57, 63).

RESULTS

Locations of linker-peptide insertions in RsaA. A version of the *rsaA* gene lacking its native promoter was used for linker mutagenesis (Fig. 1). In past studies from this laboratory, plasmid-based expression of *rsaA* has always involved the native gene possessing an unmodified transcription-translation initiation region. For the present work, we thought it prudent to use a promoterless version of *rsaA* in the event that insertions

of linker-encoded amino acids at certain positions in the protein were lethal. A promoterless version of the gene would permit further study of such mutants by placing *rsaA*ΔP downstream of a regulated promoter.

To permit the direct selection and rapid localization of linker insertion events in the *rsaA*ΔP gene, an oligonucleotide linker (*Bam*HI) was tagged with an antibiotic resistance gene, and *Taq*I sites of known location in *rsaA*ΔP were used as target sites (Fig. 2). Of a possible 28 linker insertions (there are 28 *Taq*I sites in *rsaA*), 10 different *Bam*HI linker insertions in the *rsaA* coding region, corresponding to amino acid positions 29, 154, 169, 188, 229, 250, 316, 510, 784, and 907, were recovered (Fig. 2). Previously, a *Bam*HI linker was introduced into the native *rsaA* gene (i.e., the gene carrying its native promoter) at an *Mlu*I restriction site corresponding to amino acid 93 in order to create an N-terminal deletion of RsaA (9). This construction was studied along with the insertions made at *Taq*I sites.

Wild-type and mutant versions of the *rsaA*ΔP gene were expressed in *C. crescentus* by using pWB9KSAC. Although only 11 unique linker insertions were evaluated, five different effects on S-layer biogenesis which exhibited clustering in the RsaA primary structure were identified (Fig. 2). Because *Taq*I did not cleave between the first and second nucleotides of any sequence specifying an RsaA amino acid codon, *Bam*HI linker insertion mutagenesis yielded only two of the possible three linker-encoded tetrapeptides: N-Asp-Gly-Ser-Val-C or N-Thr-Asp-Pro-Ser-C (the hexapeptide sequence N-Asp-Gly-Ser-Ala-Ser-Ala-C was inserted at amino acid 93). There was, however, no correlation between the peptide eventually inserted into RsaA and the effects on S-layer biogenesis described below.

Effects of linker-peptide insertions in the RsaA N terminus on S-layer biogenesis. Of the 11 linker-peptide insertions, 9 were located in the N-terminal half of the protein (Fig. 2). Cultures of *C. crescentus* synthesizing RsaA proteins carrying linker-peptide insertions at amino acid positions 29, 93, and 154 exhibited macroscopic particulate material in the culture fluids. As expected from previous deletion studies of RsaA (9), urea solubilization and analysis by SDS-PAGE indicated that this precipitate was substantially pure RsaA protein (Fig. 3A). Thus, these linker-peptide insertions prevented the cell surface anchoring of the protein.

When *rsaA* *C. crescentus* JS4000 cells expressing plasmid-borne copies of *rsaA*(*Taq*I29*Bam*HI), *rsaA*(*Taq*I93*Bam*HI), or *rsaA*(*Taq*I154*Bam*HI) were subjected to low-pH extraction to release S-layer protein anchored to the cell surface and the extracts were analyzed by SDS-PAGE and Coomassie blue staining, the protein profiles were identical to the profile resulting from low-pH extraction of strain JS4000 lacking a plasmid-borne copy of the *rsaA* gene (data not shown). Thus, if any S-layer protein remained anchored to the cell surface, it constituted a very minor fraction of the total. The anchoring-defective S-layer proteins were unable to crystallize onto the surface of whole cells in vivo probably because these species had precipitated following release into the culture fluids.

RsaA proteins carrying N-terminal linker-peptide insertions were evaluated for the ability to attach to the surface of phospholipid vesicles containing *C. crescentus* LPS by using an in vitro S-layer crystallization assay (43). To do this, proteins were first solubilized in 8 M urea–100 mM Tris-HCl (pH 8.5), diluted fivefold, and dialyzed against distilled water (4°C). As expected, RsaA(*Taq*I29*Bam*HI) did not crystallize onto the surface of the vesicles, and no hexameric units (Fig. 4B) were visible in the RsaA-vesicle preparations (data not shown). Instead, the majority of the protein formed amorphous aggre-

FIG. 3. SDS-PAGE and Western analysis of RsaA proteins with linker-peptide insertions. (A) SDS-PAGE–Coomassie blue analysis of particulate material recovered from cultures of *C. crescentus* JS4000 cells synthesizing RsaA defective in cell surface anchoring. Lanes 1, 2, 3, 5, and 6 represent RsaA with a linker-peptide insertions at amino acids 29, 93, 154, 316, and 510, respectively. (B) Western analysis of *E. coli* DH5α cell lysates containing RsaA with linker-peptide insertions. The linker-mutated *rsaA*ΔP genes were expressed in *E. coli* DH5α from the *lacZ*α promoter and the *rsaA* ribosome binding site. When RsaA is synthesized in *E. coli* DH5α, cells become filamentous and the protein forms inclusion bodies. These cells were lysed with a Tris-EDTA-lysozyme mixture (8) and subjected to Western analysis. Lanes 5 to 7 represent RsaA with linker-peptide insertions at amino acids 316, 510, and 907, respectively. (C) SDS-PAGE–Coomassie blue analysis of cell surface-localized RsaA with linker-peptide insertions. RsaA was extracted from the surface of whole *C. crescentus* JS4000 cells by mild acid treatment. Lanes 1, 2, 3, 5, and 6 represent RsaA with a linker peptide insertions at amino acids 169, 188, 229, 250, and 784, respectively. Lane 4 in each panel contains native RsaA either released from mutant shedder strain JS3001 (A) or derived from expression of the *rsaA*ΔP gene (B and C). Only the region of the gel containing RsaA is shown.

gates free of the vesicle surfaces. The lack of crystallinity in the protein unattached to the vesicles was consistent with earlier attempts to crystallize the S-layer in vitro. In the absence of a template, only occasional and poor-quality patches of crystallized S-layer are seen (63). In contrast to RsaA (*Taq*I29*Bam*HI), RsaA proteins with linker-peptide insertions at amino acids 93 and 154 were capable of forming an S-layer on the vesicle surfaces even though they were shed from the cell surface in vivo (Fig. 4A).

The specific sensitivity of the RsaA N terminus to linker-peptide insertions was highlighted by a cluster of four linker-peptide insertions immediately C terminal to the insertions yielding shed RsaA. These linker-peptide insertions produced no apparent disruption of S-layer biogenesis (Fig. 2). Linker-peptide insertions at amino acids 169, 188, 229, and 250 yielded protein that remained anchored to the cell surface. That is, no particulate S-layer protein was observed in the culture fluids, the mutant S-layer proteins could be extracted from the cell surface of cells in amounts comparable to wild-type amounts (Fig. 3C), and examination of whole cells from colonies by negative-stain electron microscopy revealed cell wall fragments and ghosts carrying crystallized S-layer in all cases (data not shown).

Expression of two *rsaA*ΔP genes with linker insertions corresponding to amino acid positions 316 and 510 also yielded macroscopic particulate material in the culture fluids. SDS-PAGE and Western analysis of this material yielded a protein immunoreactive with anti-RsaA antibody (data not shown) but of lower molecular mass (70 kDa) than native RsaA (Fig. 3A). Because we sequenced only those regions of mutated *rsaA*ΔP genes where linkers had been inserted, it was possible that the mutagenesis protocol had produced undetected alterations in the gene, leading to the production of modified proteins. To explore this possibility, the expression of the mutated *rsaA*ΔP genes was evaluated in *E. coli* DH5α by expressing *rsaA*ΔP from the *lacZ*α promoter of pTZ18UB. *E. coli* DH5α cells expressing linker-mutated *rsaA*ΔP genes were lysed in a Tris-HCl-EDTA-lysozyme solution, and the cell lysates were sub-

FIG. 4. (A) Electron microscopy of in vitro-crystallized RsaA carrying a linker-peptide insertion at amino acid 154. RsaA(*Taq*I154*Bam*HI) exhibits a defect in cell-surface anchoring and crystallization in vivo. Bar, 0.2 μ m. (B) Diagrammatic representation of the *C. crescentus* S-layer showing the array of ring-like hexameric subunits, each composed of six copies of RsaA, arranged on a lattice with p6 symmetry and interlinked at the threefold rotational axis. $\blacktriangle$, threefold axis; $\blacklozenge$, sixfold axis.

jected to Western analysis. In both cases, a protein with a molecular mass characteristic of native RsaA which was immunoreactive with anti-RsaA antibody was synthesized (Fig. 3B). We conclude that linker peptide insertions at amino acids 316 and 510 alter RsaA folding, exposing a site recognized by a cellular protease of *C. crescentus*. Because of this degradation, it is not possible to conclude that these linker-peptide insertions disrupt cell surface anchoring of RsaA. All known deletions of RsaA, regardless of their location, fail to anchor to the cell surface (9).

Effects of linker-peptide insertions in the RsaA C terminus on S-layer biogenesis. Two linker-peptide insertions were recovered in the C-terminal half of RsaA, one at amino acid 784 and another at amino acid 907. The insertion of a linker peptide at amino acid 784 had no effect on cell surface anchoring and crystallization of the S-layer protein; the protein was found crystallized into an S-layer on the cell surface where it could be extracted by mild acid treatment of whole cells (Fig. 3C).

In contrast, the insertion of the same linker peptide at amino acid 907 produced a severe disruption of S-layer biogenesis. RsaA(*Taq*I907*Bam*HI) could not be detected in mild acid extracts of whole cells, and no particulate material characteristic of shed RsaA protein was observed in culture fluids. Again, this result was not due to unforeseen alterations to *rsaA* Δ P during the linker mutagenesis protocol because RsaA(*Taq*I907*Bam*HI) could be detected in cell lysates of *E. coli* DH5 α expressing *rsaA*(*Taq*I907*Bam*HI) (Fig. 3B). We conclude that the linker-peptide insertion at amino acid 907 disrupts RsaA secretion. RsaA(*Taq*I907*Bam*HI) did not accumulate intracellularly and presumably was degraded after synthesis; no evidence of this protein could be found associated with cells recovered from solid medium when colonial material was solubilized in urea and analyzed by Western immunoblotting (data not shown). Similarly treated control samples prepared from *C. crescentus* synthesizing surface-anchored and shed RsaA exhibited an immunoreactive band of the expected molecular mass.

Taken together, the results of the linker-peptide insertions

in the C terminus and our previous studies (9, 10) suggested the presence of a stable (uncleaved) secretion signal residing C terminal to amino acid 784, near amino acid 907. To confirm that all of the information necessary to interact with the RsaA secretion apparatus was located in the RsaA C terminus, a deletion of RsaA retaining only the last 242 C-terminal amino acids [RsaA(Δ 1-784)] was constructed. A 1.3-kb *Bam*HI/*Hind*III fragment of *rsaA* Δ P(*Taq*I784*Bam*HI) encoding the last 242 amino acids was inserted into the same sites of pUC8, forming pUC8:*rsaA*242C (Fig. 5). This resulted in an in-frame translational fusion between the first 9 amino acids of LacZ α and the last 242 C-terminal amino acids of RsaA. In such constructions, synthesis of the fusion proteins was directed by *lacZ* α transcription and translation signals. Plasmid maintenance and replication functions were provided by fusing pUC8:*rsaA*242C to broad-host-range vector pKT210 at their common *Hind*III sites, forming pKT210:*rsaA*242C.

As with full-length RsaA protein, cultures of *C. crescentus* synthesizing only the last 242 C-terminal amino acids of RsaA exhibited macroscopic particulate material in the culture fluids. Analysis by SDS-PAGE showed a protein of 31 kDa that was immunoreactive with an anti-RsaA antibody (Fig. 6), confirming that all of the information necessary to interact with the RsaA secretion apparatus was located within the last 242 amino acids of the RsaA C terminus.

Secretion of the IHNV envelope protein from *C. crescentus* by using the RsaA C terminus. To evaluate the use of RsaA C-terminal secretion information for the secretion of heterologous proteins, an internal hydrophilic segment of the IHNV envelope glycoprotein (IHNVG; 508 amino acids) was used. PCR was used to amplify DNA encoding a 109-amino-acid segment by using plasmid pG8 carrying the entire IHNVG gene as a template (65). The PCR primers incorporated recognition sequences for the restriction enzymes *Xho*I and *Stu*I so that the PCR product could be passed through plasmid pUC9CXS to provide it with *Bam*HI termini for fusion to the DNA encoding RsaA secretion information contained in

FIG. 5. General strategy for fusing the C-terminal secretion signal of RsaA to a heterologous protein. Plasmid pUC8:*rsaA*242C (bottom) carries a 3' fragment of *rsaA* encoding a secretion signal that resides in the last 242 amino acids of RsaA. Synthesis of the RsaA C-terminal fragment is directed by pUC-based *lacZα* transcription-translation initiation information. The plasmid contains *EcoRI* and *BamHI* sites between *lacZα* and *rsaA* DNA for producing in-frame fusions with heterologous DNA. To provide DNA encoding a heterologous protein with *BamHI* termini for fusion to RsaA secretion information in pUC8:*rsaA*242C, the DNA of interest is first passed through pUC9CXS (top), a pUC9 derivative carrying a modified multiple cloning site. pUC9CXS was constructed by replacing the *PstI/BamHI* portion of the pUC9 multiple cloning site with another synthetic polylinker formed by annealing two complementary oligonucleotides. This polylinker encodes five restriction sites: *BamHI*, *XhoI*, *SalI*, *StuI*, and *BglII*. To complete the plasmid, a promoterless chloramphenicol resistance (*Cm^RΔP*) gene derived from Tn9 was inserted into the new *BglII* site so that its expression is directed by the *lacZα* promoter (Pm). The DNA of interest provided is with *XhoI/StuI* termini (middle) and inserted into the *XhoI/StuI* sites of pUC9CXS in the same orientation as the *Cm^RΔP* gene. It is then excised as a *BamHI* fragment and fused to DNA encoding RsaA secretion information via the *BamHI* site of pUC8:*rsaA*242C. Proper orientation of the heterologous DNA fragment in pUC8:*rsaA*242C is ensured by using the *Cm^RΔP* gene, which will be expressed only if the DNA of interest is inserted into the plasmid in the correct orientation. Finally, the *Cm^RΔP* gene is removed using *BglII*, and the plasmid is sealed. For introduction into *C. crescentus*, pUC8:*rsaA*242C carrying the DNA of interest is fused to broad-host-range vector pKT210 at their common *HindIII* sites.

pUC8:*rsaA*242C (Fig. 5). PCR primers and pUC9CXS were designed to maintain the proper translational frame between *lacZα* and IHNVG DNA and between IHNVG DNA and *rsaA*.

For introduction into *C. crescentus*, pUC8 carrying the *lacZα*:IHNVG:*rsaA* hybrid gene was fused to pKT210. As with cultures of *C. crescentus* synthesizing the last 242 C-terminal amino acids of RsaA alone, cultures synthesizing the LacZα:IHNVG:RsaA fusion protein also exhibited macroscopic particulate material in the culture fluids. Analysis by SDS-PAGE showed a protein of the expected molecular mass immunoreactive with both an anti-RsaA polyclonal antibody and a MAb against the 109-amino-acid IHNVG segment (Fig. 6). Thus, it was clear that the 242-amino-acid C-terminal fragment could promote secretion of heterologous proteins from *C. crescentus*.

In addition to the RsaA:IHNVG hybrid protein of the expected molecular mass and immunoreactivity, a protein with a molecular mass slightly larger than that of the 242 C-terminal portion of RsaA was detected in the particulate material recovered from the culture fluids. This 34-kDa protein was not immunoreactive with an anti-IHNVG MAb but was reactive with an anti-RsaA antibody, indicating that the majority of the

protein was derived from the RsaA C terminus. N-terminal sequencing of this protein showed that its N terminus was derived from IHNVG and that the 34-kDa protein was composed of 21 IHNVG amino acids linked to the 242-amino-acid RsaA C-terminal fragment. Further, the N-terminal amino acid lay immediately after a methionyl residue, raising the possibility that the 34-kDa protein was due to internal translational initiation within the IHNVG segment (there are three in-frame IHNVG AUG codons predicted to lie between 21 and 27 codons upstream of the IHNVG:*rsaA* fusion joint in the hybrid mRNA). Although internal translation initiation is an attractive explanation for the origin of the 34-kDa protein, we cannot rule out proteolytic cleavage within the IHNVG portion of the hybrid protein and secretion of the RsaA C-terminal fragment; little is known about the proteases of *C. crescentus* or their cleavage specificities.

In addition to using RsaA(Δ1-784), the IHNVG segment could also be secreted from *C. crescentus* as part of two RsaA internal deletions, RsaA(Δ93-784) and RsaA(Δ188-784). Each of the internal deletions was constructed from two mutant *rsaA*ΔP genes by fusing the 5' portion of one mutant gene to the 3' portion of another gene by using their common *BamHI*

FIG. 6. Secretion of RsaA deletions and RsaA:IHNVG hybrid proteins from *C. crescentus* JS4000 by using the RsaA C-terminal secretion signal and secretion machinery, monitored by SDS-PAGE analysis of particulate material recovered from the culture fluids of *C. crescentus*. (A) Proteins visualized by staining with Coomassie blue R-250; (B) Western immunoblot probed with an anti-RsaA antibody; (C) Western immunoblot probed with an anti-IHNVG MAb. Lanes: 1, native RsaA from shedding strain JS3001; 2, RsaA(Δ 1-784); 3, RsaA(Δ 1-784):IHNVG; 4, RsaA(Δ 93-784); 5, RsaA(Δ 93-784):IHNVG; 6, RsaA(Δ 188-784); 7, RsaA(Δ 188-784):IHNVG. Native RsaA (A, lane 1) was not included in the samples analyzed by Western immunoblotting. The values at the bottom represent the percentage of total culture protein recovered as particulate protein.

linker sites. Following construction of the internal deletions, the *Bam*HI fragment carrying the IHNVG DNA was inserted into the *Bam*HI site of the internal deletion by using the approach described above (Fig. 5). Hybrid proteins composed of internally deleted RsaA carrying the 109-amino-acid segment of IHNVG also formed macroscopic precipitates in the culture fluids. Following recovery by filtration through nylon mesh, urea solubilization, and Western analysis, hybrid proteins of the predicted molecular mass and immunoreactive with both the anti-IHNVG MAb and anti-RsaA antibody were detected (Fig. 6). The presence of the 34-kDa protein was also noted; this was not unexpected considering the common features of all of the *rsaA*:IHNVG gene fusions.

When RsaA was synthesized from a chromosomal copy of the *rsaA* gene in RsaA shedder strain JS3001, it accounted for approximately 8% of total culture protein of stationary-phase cultures, as judged by the Lowry assay. Deletions and IHNVG hybrid proteins incorporating the last 242 C-terminal amino acids of RsaA resulted in 3 to 10% of total culture protein as precipitated protein (Fig. 6). Although these data could be expressed on a molar basis for comparative purposes, the effects of different RsaA constructions on the amount of protein recovered cannot be easily ascertained because each RsaA deletion or hybrid protein was synthesized by using different transcription translation initiation signals.

DISCUSSION

Although S-layers were discovered over 40 years ago, the current state of knowledge concerning their structure is somewhat primitive (6, 55). So far, the highest spatial resolution models (ca. 2 nm) of S-layers have been derived from negatively stained preparations, electron microscopic tilt series, and

computer-based three-dimensional image reconstruction techniques (5, 37); no S-layer protein has yet been crystallized in a manner suitable for X-ray crystallography. Further, not one tentative model of S-layer protein folding based on a primary structure deduced from the nucleotide sequence of a cloned gene has been proposed. This is in sharp contrast to the situation with outer membrane proteins, where the positions and conformations of surface-exposed and transmembrane regions can be predicted with some confidence from such data. Biochemical studies of soluble monomeric S-layer proteins (14, 23, 46, 58) and computer-based structure prediction programs (31) have not been very informative either. The concern has always been that the structural information derived from these studies may not describe the properties of crystallized S-layer proteins anchored to cell surface.

Recently, deletion mutagenesis (27, 58), proteolytic modification of S-layer proteins (21), epitope accessibility determination and mapping (22, 44), and primary sequence comparisons (41) have been applied to the study of S-layer structure. In this study, we have investigated the use of linker mutagenesis in an effort to introduce localized disruptions of S-layer proteins. This technique seems to be a useful one warranting more extensive use. If the data derived from *in vitro* crystallization assay are considered along with the *in vivo* data, 11 linker-peptide insertions yielded five different effects on S-layer biogenesis.

Linker-peptide insertions in the RsaA N terminus (amino acid positions 29, 93, and 154) disrupt *in vivo* anchoring of the protein to the cell surface. This finding supports the results of our previous study (9) showing a similar effect resulting from the genetic deletion of RsaA N-terminal amino acids 2 to 29. In the present study, we had hoped to use *in vitro* crystallization to confirm and supplement *in vivo* results; however, we encountered problems in a straightforward interpretation of the data. For example, the simplest explanation for the effect of linker-peptide insertions in the RsaA N terminus is that these modifications disrupt a site in the protein required for cell surface anchoring. We expected, therefore, that if these proteins failed to become anchored to the surface of whole cells *in vivo* for this reason, they would also fail to attach to the surface of vesicles designed to simulate the surface of cells. This was not the case for two of the three RsaA proteins carrying N-terminal linker-peptide insertions. Both RsaA(*Taq*I93*Bam*HI) and RsaA(*Taq*I154*Bam*HI) attached to phospholipid vesicles containing *C. crescentus* LPS and crystallized into S-layers.

There are at least two possible explanations for these results. The first possibility relates to the different properties of *in vitro*- and *in vivo*-crystallized S-layers. Early experiments with several S-layers showed that isolated or *in vitro*-crystallized S-layers could be more easily disrupted than native S-layers, suggesting some difference in their structures (see reference 7 for a discussion). More recently, it has been shown that S-layer proteins form an association with the underlying cell envelope layers *in vivo* that can be different from that formed when S-layers are assembled *in vitro* (24, 30). Presumably *in vitro*-crystallized S-layers may sometimes provide misleading data concerning the *in vivo* situation. Not even the gross appearance of a crystallized S-layer in negatively stained preparations can be taken as evidence that an *in vitro*-crystallized S-layer possesses the same structure as an S-layer assembled *in vivo*. Alterations to structure can go undetected unless electron micrographs are subjected to computer image processing (7). If, as seems possible, the data from the *in vitro* crystallization assay cannot be applied to *in vivo* cell surface anchoring, we can still conclude that linker-peptide insertions at amino acids 93 and 154 do disrupt a region of the protein required for cell

surface anchoring in vivo. The finding that the N terminus of RsaA is important for cell surface anchoring would be consistent with the results of other studies with both the gram-negative bacterium *Campylobacter fetus* (27) and several gram-positive organisms (41, 44), suggesting a common role for the N termini of S-layer proteins.

On the other hand, if the in vitro-crystallized S-layers are anchored to vesicles and crystallized in the same manner as native S-layers, this would mean that linker-peptide insertions at amino acids 93 and 154 disrupt the process of in vivo anchoring and crystallization rather than a contact point between RsaA and the outer membrane. This could be due to disruption of aspects of the in vivo folding pathway or simply the rate of folding of an S-layer protein domain important in cell surface anchoring. The same linker insertion would not necessarily be expected to disrupt in vitro attachment where the protein is refolded after urea solubilization and presented to the external surface of vesicles that are not simultaneously enlarging.

Thus, in several ways in vitro crystallization involves S-layer formation in a manner substantially different from that naturally employed by the cell. At the moment, we cannot resolve these competing possibilities; more work needs to be done with regard to comparing properties of in vivo-assembled S-layers to those assembled in vitro. Parenthetically, the data from these linker insertions highlight the general difficulty that we currently have in distinguishing positions within RsaA responsible for surface anchoring versus key contact positions used for crystallization. This is due the strong template dependence of S-layer crystallization; that is, attachment to a suitable surface is a prerequisite for crystallization.

A useful by-product of the linker mutagenesis approach is that genes carrying linker insertions can be easily manipulated to generate targeted deletions. This property was used to isolate a 3' fragment of the *rsaA* gene encoding the 242 C-terminal amino acids of RsaA. Because this portion of RsaA can be secreted from *C. crescentus* by using heterologous *E. coli lacZ* transcription-translation initiation information, it is now unambiguous that all of the information necessary to interact with the RsaA secretion apparatus is located within the C terminus of the protein.

Bacterial proteins carrying stable C-terminal secretion signals are secreted via type I dedicated secretion machinery rather than the general type II secretion pathway of the cell (47, 48). Type I secretion involves release of the protein directly into the medium without a periplasmic intermediate. Secretion is mediated by a complex of three membrane proteins, one of which belongs to the ATP-binding cassette (ABC) superfamily of transport proteins (36). These ABC transporters show considerable similarity at the primary and secondary structural levels to cognate proteins in the superfamily, whereas, oddly, the C termini of the substrate proteins do not. The genes encoding type I secretion machinery are usually clustered near the gene for the protein that is the substrate of the secretion proteins. Recently, we have begun nucleotide sequencing of DNA adjacent to the *rsaA* gene and have identified an open reading frame which shows a high degree of similarity to the ABC transporter for the alkaline protease of *Pseudomonas aeruginosa* (2, 25). The location of this gene relative to *rsaA* is consistent with the hypothesis that RsaA is secreted by a type I system.

It has been clear since the completion of *rsaA* sequencing (31) that the *C. crescentus* S-layer protein is probably secreted via a system related to, but perhaps not totally analogous to, the type I variety. Inspection of the predicted amino acid sequence of RsaA indicated that (i) neither N- nor C-terminal processing occurred during its transport to the cell surface; (ii)

the protein contained several C-terminal aspartate-containing glycine-rich repeat sequences; and (iii) RsaA lacked multiple cysteine residues, eliminating the potential for disulfide bond formation. All of these features are generally diagnostic for proteins secreted by a type I system. However, at the same time RsaA sequence analysis was proceeding, gene fusion/hybrid protein studies indicated that the N terminus of RsaA could promote secretion of passenger proteins to the periplasm, suggesting that stable secretion information was located in the N terminus (9, 10). These results were clearly at odds with the picture emerging for type I secretion, as exemplified by *E. coli* α -hemolysin, although the presence of N-terminal secretion information and a periplasmic intermediate were not without precedence in type I systems. Indeed, colicin V appeared to be secreted via the periplasm, relying on N-terminal secretion information, yet somehow used the type I secretion machinery (33). Further, the α -hemolysin type I secretion machinery could be used to secrete colicin V from *E. coli* (29), indicating that type I secretion signals did not necessarily have to lie at the C terminus to be functional. These findings opened up the possibility that some type I secretion systems (including that for RsaA secretion) are more complex than the α -hemolysin variety, perhaps involving both N- and C-terminal secretion signals (see reference 48 for a discussion). Interestingly, it was recently demonstrated that the N terminus of colicin V is cleaved during secretion (35), and although an ABC transporter is involved in its secretion, it is a member of different family of transporters (34) than those used to secrete α -hemolysin and similar proteins. As the confusion surrounding secretion of colicin V was being resolved, we eliminated a role for the RsaA N terminus in secretion of the protein by showing that N-terminally deleted RsaA proteins could still be secreted by *C. crescentus* (9). We concluded that the use of hybrid proteins (see above) had provided misleading information concerning RsaA secretion and hypothesized that once separated from the context of the native protein, the RsaA N terminus functioned as a cryptic export signal mediating hybrid protein export from the cytoplasm via the general secretory pathway of the cell.

In a recent review, Blight and Holland (17) reported that 13 proteins were known to be secreted by type I systems. Since that review, at least four additional proteins (1, 26, 28, 39) have been confirmed or are suspected to be secreted via type I systems. The *C. crescentus* S-layer protein is the only S-layer protein and the first structural protein from a gram-negative bacterium known to be secreted by a type I system. The only other gram-negative S-layer protein secretion systems investigated to date are those of *Aeromonas salmonicida*, *Aeromonas hydrophila*, and *Campylobacter fetus* (18). Like the *C. crescentus* S-layer protein, the *C. fetus* S-layer protein (15) is probably secreted by a type I secretion system (18). On the other hand, the S-layer proteins from the *Aeromonas* species are synthesized with classical N-terminal signal leader peptides (21, 59, 60) and are secreted via the periplasm by the general secretory pathway of the cell (47).

The C-terminal secretion signals characteristic of type I secretion are located beyond aspartate-containing glycine-rich repeat regions usually within last 40 to 60 residues of the substrate protein. Sequence analysis is not a productive approach for the identification of these signals because they possess no clear consensus primary sequence (48) or identifiable secondary structural elements. The linker-peptide insertion at amino acid 907 which disrupts RsaA secretion lies at the C-terminal end of the aspartate-containing glycine-rich repeat region (amino acids 860 to 905) of RsaA. If this linker-peptide insertion disrupts the secretion of RsaA by directly interrupt-

ing the secretion signal, the RsaA secretion signal may be slightly more N terminal than other similar signals. However, it is also possible that the linker-peptide insertion at amino acid 907 disrupts secretion by a more indirect effect. There is some indication that the aspartate-containing glycine-rich regions of other type I proteins act as spacer regions between the secretion signal and the rest of the protein and may be important for presentation of the secretion signal to the secretion apparatus (40). Thus, the linker-peptide insertion at amino acid 907 may disrupt secretion of RsaA by restricting access of the secretion signal to the secretion machinery.

We have not been able to detect RsaA(*Taq*I907*Bam*HI) associated with cells recovered from solid medium, indicating that this secretion-incompetent version of RsaA does not accumulate intracellularly but is degraded after synthesis. A similar result was found with a C-terminal deletion of RsaA lacking the last 21 C-terminal amino acids of the protein (10). Our inability to detect intracellular accumulation of RsaA carrying extreme C-terminal modifications is not without precedence. Degradation of proteins secreted by type I systems as a result of defects or absence in the secretion signal or secretion apparatus seems to be a common feature of such systems (20).

One practical consequence of this work has been the demonstration that the C-terminal secretion signal of RsaA can be used to secrete heterologous proteins. This finding indicates that the RsaA secretion system is tolerant of different substrates so long as the C-terminal secretion signal is present. Further, the tendency of shed RsaA to precipitate makes recovery and purification of hybrid proteins a facile task. The quantity of precipitated protein so far recovered from cultures has amounted to between 3 and 10% of total culture protein, comparable or exceeding that reported for the *E. coli* α -hemolysin secretion system (17). We are currently investigating the versatility of the *C. crescentus* S-layer protein secretion system for secretion of a number of heterologous proteins. If the *E. coli* α -hemolysin experience is an indicator, it should be possible to link a wide variety of proteins to the RsaA secretion signal and achieve high-level secretion of hybrid proteins into the growth medium. So far, only the S-layer protein of *Bacillus brevis* has been exploited for heterologous protein secretion; the secretion of this protein relies on a classical N-terminal signal leader peptide (64).

Finally, we note that many of the linker-peptide insertions, especially in the central region of the protein, produce no apparent effect on S-layer biogenesis. These sites may be capable of accepting additional mass, such as small epitopes through the introduction of heterologous DNA into the linker-encoded restriction sites of the *rsaA* gene. Epitopes so introduced would be presented in a high-density repeating pattern on the cell surface (current estimates suggest that there are 60,000 RsaA monomers per cell [57]). These epitopes could be used for a variety of purposes, including whole-cell vaccine development or mapping of surface-exposed and buried regions of the S-layer.

ACKNOWLEDGMENTS

We thank Khai Le for performing the PCR cloning of the IHNVG segment and Faten Borchani for preparing samples for amino acid microsequencing. MAb 136J was kindly provided by Jo-Ann C. Leong of Oregon State University, Corvallis.

This work was supported by grants from the British Columbia Health Care Research Foundation, the National Sciences and Engineering Research Council of Canada, and the Canadian Department of Fisheries and Oceans.

REFERENCES

1. Akatsuka, H., E. Kawai, K. Omori, and T. Shibatani. 1995. The three genes *lipB*, *lipC*, and *lipD* involved in the extracellular secretion of the *Serratia marcescens* lipase which lacks an N-terminal signal peptide. *J. Bacteriol.* **177**:6381–6389.
2. Awram, P., and J. Smit. Unpublished data.
3. Bagdasarian, M., R. Lurz, B. Ruckert, F. C. H. Franklin, M. M. Bagdasarian, and K. N. Timmis. 1981. Specific-purpose cloning vectors. II. Broad-host-range, high copy number RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**:237–247.
4. Barany, F. 1985. Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. *Gene* **37**:111–123.
5. Baumeister, W., and H. Engelhardt. 1987. Three-dimensional structure of bacterial surface layers, p. 109–154. In J. R. Harris, and W. Horne (ed.), *Electron microscopy of proteins*, vol. 6. Membrane structure. Academic Press, London, England.
6. Beveridge, T. J. 1994. Bacterial S-layers. *Curr. Opin. Struct. Biol.* **4**:204–212.
7. Bingle, W. H., H. Engelhardt, W. J. Page, and W. Baumeister. 1987. Three-dimensional structure of the regular tetragonal surface layer of *Azotobacter vinelandii*. *J. Bacteriol.* **169**:5008–5015.
8. Bingle, W. H., H. D. Kurtz, Jr., and J. Smit. 1993. An all-purpose cellulase reporter for gene fusion studies and application to the paracrystalline surface(S)-layer protein of *Caulobacter crescentus*. *Can. J. Microbiol.* **39**:70–80.
9. Bingle, W. H., K. D. Le, and J. Smit. 1996. The extreme N-terminus of the *Caulobacter crescentus* S-layer protein directs export of passenger proteins from the cytoplasm but is not required for the secretion of the native protein. *Can. J. Microbiol.* **42**:672–684.
10. Bingle, W. H., and J. Smit. 1994. Alkaline phosphatase and a cellulase reporter are not exported from the cytoplasm when fused to large N-terminal portions of the *Caulobacter crescentus* S-layer protein. *Can. J. Microbiol.* **40**:777–782.
11. Bingle, W. H., and J. Smit. 1991. A method of tagging specific-purpose linkers with an antibiotic resistance gene for linker mutagenesis using a selectable marker. *BioTechniques* **10**:150–153.
12. Bingle, W. H., and J. Smit. 1990. High-level expression vectors for *Caulobacter crescentus* incorporating the transcription/translation initiation regions of the paracrystalline surface-layer protein gene. *Plasmid* **24**:143–149.
13. Bingle, W. H., S. G. Walker, and J. Smit. 1993. Definition of form and function for the S-layer of *Caulobacter crescentus*, p. 181–194. In T. J. Beveridge and S. F. Koval (ed.), *Advances in paracrystalline bacterial surface layers*. Plenum Press, New York, N.Y.
14. Bingle, W. H., P. W. Whippley, J. L. Doran, R. G. E. Murray, and W. J. Page. 1987. Structure of the *Azotobacter vinelandii* surface layer. *J. Bacteriol.* **169**:802–810.
15. Blaser, M. J., and E. C. Gotschlich. 1990. Surface array protein of *Campylobacter fetus*: cloning and gene structure. *J. Biol. Chem.* **265**:14529–14535.
16. Blight, M. A., C. Chervaux, and I. B. Holland. 1994. Protein secretion pathways in *Escherichia coli*. *Curr. Opin. Biotechnol.* **5**:468–474.
17. Blight, M. A., and I. B. Holland. 1994. Heterologous protein secretion and the versatile *Escherichia coli* haemolysin translocator. *Trends Biotechnol.* **12**:451–455.
18. Boot, H., and P. H. Pouwels. 1996. Expression, secretion and antigenic variation of bacterial S-layer proteins. *Mol. Microbiol.* **21**:1117–1123.
19. Brent, R., and M. Ptashne. 1980. The *lexA* gene product represses its own promoter. *Proc. Natl. Acad. Sci. USA* **77**:1932–1936.
20. Chervaux, C., and I. B. Holland. 1996. Random and directed mutagenesis to elucidate the functional importance of helix II and F-989 in the C-terminal secretion signal of *Escherichia coli* hemolysin. *J. Bacteriol.* **178**:1232–1236.
21. Chu, S., S. Cavaignac, J. Feutrier, B. M. Phipps, M. Kostrzynska, W. W. Kay, and T. J. Trust. 1991. Structure of the tetragonal surface virulence array protein and gene of *Aeromonas salmonicida*. *J. Biol. Chem.* **266**:15258–15265.
22. Doig, P., W. D. McCubbin, C. M. Kay, and T. J. Trust. 1993. Distribution of surface-exposed and non-accessible amino acid sequences among the two major structural domains of the S-layer protein of *Aeromonas salmonicida*. *J. Mol. Biol.* **233**:753–765.
23. Dooley, J. S. G., W. D. M. McCubbin, C. M. Kay, and T. J. Trust. 1988. Isolation and biochemical characterization of the S-layer protein from a pathogenic strain of *Aeromonas hydrophila*. *J. Bacteriol.* **170**:2631–2638.
24. Doran, J. L., W. H. Bingle, and W. J. Page. 1987. Role of calcium in the assembly of the *Azotobacter vinelandii* surface array. *J. Gen. Microbiol.* **133**:399–413.
25. Duong, F., A. Lazdunski, B. Cami, and M. Murgier. 1992. Sequence of a cluster of genes controlling synthesis and secretion of alkaline protease in *Pseudomonas aeruginosa*: relationships to other secretory pathways. *Gene* **121**:47–54.
26. Duong, F., C. Sosia, A. Lazdunski, and M. Murgier. 1994. The *Pseudomonas fluorescens* lipase has a C-terminal secretion signal and is secreted by a three-component bacterial ABC-exporter system. *Mol. Microbiol.* **11**:1117–1126.
27. Dworkin, J., M. K. R. Tummuru, and M. J. Blaser. 1995. A lipopolysaccharide-binding domain of the *Campylobacter fetus* S-layer protein resides within

- the conserved N terminus of a family of silent and divergent homologs. *J. Bacteriol.* **177**:1734–1741.
28. Ertesvåg, H., B. Doseth, B. Larsen, G. Skjåak-Braek, and S. Valla. 1994. Cloning and expression of an *Azotobacter vinelandii* mannuronan C-5-epimerase gene. *J. Bacteriol.* **176**:2846–2853.
 29. Fath, M. J., R. C. Skvirsky, and R. Kolter. 1991. Functional complementation between bacterial MDR-like export systems: colicin V, alpha-hemolysin, and *Erwinia* protease. *J. Bacteriol.* **173**:7549–7556.
 30. Gardiño, R. A., B. M. Phipps, and W. W. Kay. 1995. Physical and functional S-layer reconstitution in *Aeromonas salmonicida*. *J. Bacteriol.* **177**:2684–2694.
 31. Gilchrist, A., J. Fisher, and J. Smit. 1992. Nucleotide sequence analysis of the gene encoding the *Caulobacter crescentus* paracrystalline surface layer protein. *Can. J. Microbiol.* **38**:193–202.
 32. Gilchrist, A., and J. Smit. 1991. Transformation of freshwater and marine caulobacters by electroporation. *J. Bacteriol.* **173**:921–925.
 33. Gilson, L., H. K. Mahanty, and R. Kolter. 1990. Genetic analysis of an MDR-like export system: the secretion of colicin V. *EMBO J.* **9**:3875–3884.
 34. Håvarstein, L. S., D. B. Diep, and I. F. Nes. 1995. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol. Microbiol.* **16**:229–240.
 35. Håvarstein, L. S., H. Holo, and I. F. Nes. 1994. The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by Gram-positive bacteria. *Microbiology* **140**:2383–2389.
 36. Higgins, C. F. 1992. ABC-transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* **8**:67–113.
 37. Hövmöller, S., S. Sjögren, and D. N. Wang. 1988. The structure of crystalline bacterial surface layers. *Prog. Biophys. Mol. Biol.* **51**:131–163.
 38. Koval, S. F., and S. H. Hynes. 1991. The effect of paracrystalline protein surface layers on predation by *Bdellovibrio bacteriovorus*. *J. Bacteriol.* **173**:2244–2249.
 39. Létoffé, S., J. M. Ghigo, and C. Wandersman. 1994. Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *J. Bacteriol.* **176**:5372–5377.
 40. Létoffé, S., and C. Wandersman. 1992. Secretion of CyaA-PrtB and HylA-PrtB fusion proteins in *Escherichia coli*: involvement of the glycine-rich repeat domain of *Erwinia chrysanthemi* protease B. *J. Bacteriol.* **174**:4920–4927.
 41. Lupas, A., H. Engelhardt, J. Peters, U. Santarius, S. Volker, and W. Baumeister. 1994. Domain structure of the *Acetogenium kivui* surface layer revealed by electron crystallography and sequence analysis. *J. Bacteriol.* **176**:1224–1233.
 42. Markwell, M. K., S. M. Hass, L. L. Bieber, and N. E. Tolbert. 1978. A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Anal. Biochem.* **87**:206–210.
 43. Nomellini, J. F., S. Küpcü, U. B. Sleytr, and J. Smit. Unpublished data.
 44. Olabarria, G., J. L. Carrascosa, M. A. De Pedro, and J. Berenguer. 1996. A conserved motif in S-layer proteins is involved in peptidoglycan binding in *Thermus thermophilus*. *J. Bacteriol.* **178**:4765–4772.
 45. Paddock, G. V. 1993. Rapid colony hybridization on Whatman 541 paper using oligonucleotide probes. *Methods Enzymol.* **217**:340–346.
 46. Phipps, B. M., T. J. Trust, E. E. Ishiguro, and W. W. Kay. 1983. Purification and characterization of the cell surface virulence A-protein from *Aeromonas salmonicida*. *Biochemistry* **22**:2934–2939.
 47. Pugsley, A. P. 1993. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **57**:50–108.
 48. Pugsley, A. P. 1992. Superfamilies of bacterial transport systems with nucleotide binding components. *Symp. Soc. Gen. Microbiol.* **47**:223–248.
 49. Quigley, N. B., and P. R. Reeves. 1987. Chloramphenicol resistance cloning vector based on pUC9. *Plasmid* **17**:54–57.
 50. Salmond, G. P. C., and P. J. Reeves. 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. *Trends Biochemical Sci.* **18**:8–12.
 51. Sambrook, J., E. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 52. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 53. Saskawa, C., and M. Yoshikawa. 1987. A series of Tn5 variants with various drug-resistance markers and suicide vector for transposon mutagenesis. *Gene* **56**:283–288.
 54. Sawadogo, M., and M. W. Van Dyke. 1991. A rapid method for the purification of deprotected oligodeoxynucleotides. *Nucleic Acids Res.* **19**:674.
 55. Sleytr, U. B., P. Messner, D. Pum, and M. Sára. 1993. Crystalline bacterial cell surface layers: general principles and application potential. *J. Appl. Bacteriol. Symp. Suppl.* **74**:21S–32S.
 56. Smit, J., and N. Agabian. 1984. Cloning of the major protein of the *Caulobacter crescentus* periodic surface layer: detection and characterization of the cloned peptide by protein expression assays. *J. Bacteriol.* **160**:1137–1145.
 57. Smit, J., H. Engelhardt, S. Volker, S. H. Smith, and W. Baumeister. 1992. The S-layer of *Caulobacter crescentus*: three-dimensional image reconstruction and structure analysis by electron microscopy. *J. Bacteriol.* **174**:6527–6538.
 58. Thomas, S., J. W. Austin, W. D. McCubbin, C. M. Kay, and T. J. Trust. 1992. Roles of structural domains in the morphology and surface anchoring of the tetragonal paracrystalline array of *Aeromonas hydrophila*: biochemical characterization of the major structural domain. *J. Mol. Biol.* **228**:652–661.
 59. Thomas, S. R., and T. J. Trust. 1995. A specific PulD homolog is required for the secretion of the paracrystalline surface array subunits in *Aeromonas hydrophila*. *J. Bacteriol.* **177**:3932–3939.
 60. Thomas, S. R., and T. J. Trust. 1995. Tyrosine phosphorylation of the tetragonal paracrystalline array of *Aeromonas hydrophila*: molecular cloning and high level expression of the S-layer protein gene. *J. Mol. Biol.* **245**:568–581.
 61. Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259–268.
 62. Walker, S. G., D. N. Karunaratne, N. R. Ravenscroft, and J. Smit. 1994. Characterization of mutants of *Caulobacter crescentus* defective in surface attachment of the paracrystalline surface layer. *J. Bacteriol.* **176**:6312–6323.
 63. Walker, S. G., S. H. Smith, and J. Smit. 1992. Isolation and comparison of the paracrystalline surface layer proteins of freshwater caulobacters. *J. Bacteriol.* **174**:1783–1792.
 64. Yamagata, H., K. Nakahama, Y. Suzuki, A. Kakinuma, N. Tsukagoshi, and S. Ueda. 1989. Use of *Bacillus brevis* for efficient synthesis and secretion of human epidermal growth factor. *Proc. Natl. Acad. Sci. USA* **86**:3589–3593.
 65. Xu, L., D. V. Mourich, H. M. Engleking, S. Ristow, J. Arnzen, and J. C. Leong. 1991. Epitope mapping and characterization of the infectious hematopoietic virus glycoprotein, using fusion proteins synthesized in *Escherichia coli*. *J. Virol.* **65**:1611–1615.