Clostriolysin S, a Post-translationally Modified Biotoxin from Clostridium botulinum

Received for publication, February 26, 2010, and in revised form, June 5, 2010. Published, JBC Papers in Press, June 25, 2010. DOI 10.1074/jbc.M110.118554

David J. Gonzalez, a Shaun W. Lee, b Mary E. Hensler, c Andrew L. Markley, c Samira Dahesh, c Douglas A. Mitchell, d Nuno Bandeira, e Victor Nizet, c,f Jack E. Dixon, a,g,h,i and Pieter C. Dorrestein a,g,i

From the a Department of Chemistry and Biochemistry, Skaggs School of Pharmacy and Pharmaceutical Sciences, 1Center for Marine Biotechnology and Biomedicine, bDepartment of Pharmacology, cDepartment of Cellular and Molecular Medicine, dDivision of Pharmacology and Drug Discovery, Department of Pediatrics, and eDepartment of Computer Science and Engineering, University of California at San Diego, La Jolla, California 92093, the f Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, the g Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556, and the h Departments of Chemistry and Microbiology and the Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Through elaboration of its botulimum toxins, Clostridium botulinum produces clinical syndromes of infant botulism, wound botulism, and other invasive infections. Using comparative genomic analysis, an orphan nine-gene cluster was identified in C. botulinum and the related foodborne pathogen Clostridium sporogenes that resembled the biosynthetic machinery for streptolysin S, a key virulence factor from group A Streptococcus responsible for its hallmark β-hemolytic phenotype. Genetic complementation, in vitro reconstitution, mass spectral analysis, and plasmid intergrational mutagenesis demonstrate that the streptolysin S-like gene cluster from Clostridium sp. is responsible for the biogenesis of a novel post-translationally modified hemolytic toxin, clostriaolysin S.

Microbial virulence and survival are often defined by metabolic output. Among the many molecular species used to give a competitive advantage are hydrogen peroxide, genetically encoded small molecules, siderophores, proteins, and bacteriocins (1, 2). Streptolysin S (SLS)2 is a well known hemolytic/cytolytic, ribosomally encoded bacteriocin and virulence factor group A Streptococcus (GAS) (3–5). To this day, the characteristic β-hemolytic phenotype observed on blood agar plates is used as a clinical diagnostic tool for GAS identification. GAS is best known as the agent of acute pharyngitis (strep throat) but also may cause invasive infections, including necrotizing fasciitis and toxic shock syndrome. In the 100-year history of our knowledge of streptococcal bacteria, the precise chemical structure of this toxin has remained elusive, although the discovery of its biosynthetic gene cluster more than a decade ago (4) has guided investigations into its post-translational modification and likely heterocyclic nature (6, 7).

Through recent comparative genomic analysis, an SLS-type gene cluster was identified in clostridia species including Clostridium botulinum and Clostridium sporogenes, two disease-causing bacteria known to endanger food supplies (8–10). Similar gene clusters were found in Staphylococcus aureus RF122, Bacillus thuringiensis, Streptococcus iniae, and Listeria monocytogenes 4b. The L. monocytogenes 4b strain is the primary serotype and causative agent for outbreaks of listeriosis (11). The S. aureus RF122 strain is responsible for bovine mastitis. S. iniae is a cytotoxic fish pathogen. This suggests that a shared metabolic output of these pathogens including SLS-like toxins may contribute to their pathogenic potential (11).

Each of these SLS family gene clusters contains a similar set of genes. For instance, in C. botulinum and C. sporogenes, the cloSA–I genes are related by sequence to sagA–I from GAS (see Fig. 1A). Of these genes, CloS is a protein of unknown function. CloG, -H, and -I are ABC transporters and therefore are likely to be responsible for exporting the mature hemolytic product. CloSA is the prepropeptide (structural peptide) that is post-translationally modified to form the propeptide. CloS may be annotated as an immunity protein but is similar to the CaaX protease superfamily in sequence. Thus, CloS may be the enzyme responsible for proteolytic processing of the post-translationally modified propeptide to the mature toxin. CloSBCD resembles the McbBCD-modifying proteins found in the Escherichia coli microcin B17 system and the Prochloron didemni patDG genes of the patellamide biosynthetic gene cluster. This family of natural product toxins was recently shown by several groups to include the thioproteptides, which are nearly ubiquitous in soil-dwelling Bacillus and Streptomyces bacteria (12–16). In the microcin B17 system, McbB, -C, and -D were shown to be necessary and sufficient for the installation of heterocyclic moieties on C-terminal serine and cysteine residues of the McBA substrate. Chemical transformations installed on the McBA substrate were assayed by mass spectrometry and were shown to result in the loss of multiples of 20 Da (see Fig. 3A). Based on the comparative genomics with the above-mentioned clusters, CloSBCD also are likely involved in
the post-translational conversion of serine, threonine, and cysteine residues of ClosA to their corresponding thiazole and methylxazole heterocycles to generate a modified peptide. BLAST analysis indicates that ClosB is similar to a FMN-dependent oxidase, ClosC is similar to the ATP-dependent enzymes E1, ThiF, and MoeB, which contain a structural zinc that is not involved in catalysis (17, 18). The E1/ThiF/MoeB protein family utilizes ATP to activate the C-terminal end of ubiquitin or an ubiquitin-like protein. ClosD belongs to a biochemically uncharacterized protein family YcaO (DUF181).

In this paper, we characterize the SLS-type toxin from C. botulinum and C. sporogenes. We demonstrate that the gene clusters in clostridia are functionally equivalent to the SLS biosynthetic pathway by complementing targeted GAS SLS operon knock-outs with the corresponding Clostridium genes. Once the genome became available (10), plasmid integrational mutants of the closA and closC gene products also were generated in C. sporogenes to verify that the clostridiolysin S gene cluster was indeed responsible for a hemolytic phenotype. Hemolytic activity of this gene cluster was reconstituted in vitro, allowing the direct mass spectral confirmation that the clostridiolysin toxin is post-translationally modified to contain heterocyclic moieties and that this process required structural zinc bound to the putative cyclodehydratase, ClosC. Finally, we demonstrate that the synthetase enzymes install heterocycles on ClosA using nanocapillary LC-MS/MS in conjunction with the mass spectral data analysis programs InSpecT (19) and Spectral Networks (20), to establish that the nontoxic precursor (ClosA) is converted into the final post-translationally modified toxin, clostridiolysin S.

**EXPERIMENTAL PROCEDURES**

**Plasmid Integrational Mutagenesis of closA and closC**—PCR was used to amplify the central regions of C. sporogenes closA and closC genes. PCR products were recovered by T-A cloning in the vector pCR2.1-TOPO (Invitrogen) and then subcloned into the temperature-sensitive erythromycin (ErmR) plasmid, pHY304. The knock-out plasmids were introduced by electroporation into C. sporogenes, and transformants were recovered on THA-Erm 500 µg/ml. Single recombination events were identified by shifting to the nonpermissive temperature (37 °C) while maintaining Erm selection and were later confirmed by PCR. Intragenic plasmid integrational mutants were streaked onto blood agar plates (5% sheep blood, Hardy Diagnostics) to assess hemolytic activity.

**Cloning and Protein Purification of closA, closB, closC, closD, and sagB**

Cloning of closA, closC, and closD—Single synthetic gene copies of closA–D were subcloned into pGS-21a KpnI/HindIII restriction enzyme sites were purchased from GenScript Corp. The pGS-21a vectors harboring the closA–D genes contain a GST tag and His6 tag at the N terminus of each clos gene for affinity purification purposes, followed by an enterokinase cleavage site. Using the closA–D/pGS-21a vectors as templates, closA–D were PCR-amplified by BamHI/NotI flanking sequences. The closA–D PCR amplicons were verified by agarose (1%) gel electrophoresis. Positively identified closA–D amplicons were then excised and purified using the Qiagen QIAquick PCR purification kit. The purified amplicons and the pET28 vector containing an N-terminal maltose-binding protein (MBP) tag were then sequentially digested by allowing the NotI digest to proceed overnight followed by a 2-h BamHI digestion at 37 °C (New England Biolabs). Following the BamHI/NotI digest of the pET28-MBP-modified vector, a calf intestine phosphatase reaction was performed (New England Biolabs). The digested amplicons and vector were then separated by agarose gel electrophoresis, excised, and purified with a Qiaquick gel extraction kit. Purified closA–D amplicons were then ligated into the similarly digested vector using a rapid ligation kit (Roche Applied Science). Ligation products of MBP-closA–D were then transformed into one-shot Top10 DH5α chemically competent cells (Invitrogen) and plated on Kan60 Luria Broth (LB) plates. Single colonies of each transformant were picked and grown overnight in preparation for plasmid recovery. The recovered closA–D plasmids were then verified by PCR amplification and sequenced by Eton Biosciences (La Jolla, CA). The validated plasmids were then transformed into Invitrogen E. coli BL21(DE3) cells by heat shock and recovered in super-optimal broth with catabolic repression (SOC) medium for 1 h (Invitrogen). Cells were harvested by centrifugation at 13,000 rpm for 1 min and then plated on Kan60 LB agar plates at 37 °C overnight. Single colonies were picked and verified for the presence of the closA–D genes by colony PCR. Colonies that contained the closA–D genes were grown and then stored at −80 °C until used in the hemolytic/cytolytic assays. Preparation of the MBP-SagB inducible construct was described previously (6).

**Protein Purification of MBP-ClosA–D and MBP-SagA–D**—All MBP-tagged proteins used in the in vitro hemolytic assays were purified as described previously (6). Upon purification, all proteins were immediately stored at −80 °C until further use.

**Cytolytic Activity Assay**

**In Vitro Synthetase Reactions**—Protein preparation and synthetase reactions employing MBP-tagged substrate ClosA and the BCD synthetases were performed as described previously for the in vitro reconstitution of SLS activity (6). In every case, omission of the substrate or the synthetase resulted in no detectable hemolytic activity. All assays were internally normalized and baseline-adjusted using two positive controls (Tripton X-100 and wild-type SagA treated with SagBCD) and two negative controls (substrate and SagBCD alone).

**Clos Single Gene Complementation into Group A Streptococcus ΔsagA–D Strains**

Preparation of GAS ΔsagA–D Strains Harvesting the closA–D Genes—The construction of GAS ΔsagA, ΔsagB, ΔsagC, and ΔsagD allelic exchange mutants in GAS strain NZ131 (M49 serotype) was described previously (3). To introduce the individual closA–D genes in trans to the corresponding sagA–D allelic exchange mutants, closA–D were individually PCR-amplified with XbaI/BamHI flanking sequences from pGS-21a closA–D gene containing vectors (GenScript). Each PCR amplion was verified on a 1.0% gel, and the positively identified closA–D amplicons were then excised and purified using a Pro-
Clostridiolysin S, a Post-translationally Modified Biotoxin

mega PCR purification wizard kit. Sequential restriction enzyme digests, BglII/XbaI for the pDCerm plasmid (3) and BamHI/XbaI for the closA–D amplicons followed (New England Biolabs). Restriction enzyme products were separated on a 1.0% agarose gel, excised, and purified by the use of a Promega gel extraction wizard kit. Ligation was performed by the use of a 3:1 insert to vector ratio with a ligation kit (Roche Applied Science). The resulting ligation product was then transformed into E. coli MC1061 competent cells via electroporation and then plated onto LAermR plates. Single colonies of each transformant were then picked and grown overnight in preparation for plasmid recovery. The recovered pDCerm-closA–D plasmid was then verified by PCR amplification and sequenced by Eton Biosciences. The validated pDCerm-closA–D plasmids were then transformed into the corresponding GAS ΔsagA–D strains via electroporation and plated on THBermR plates (Todd Hewitt Broth). Single colonies were picked and verified for the presence of the closA–D genes by colony PCR. Complementation of the constructed GAS ΔsagA–D harvesting pDCerm-closA–D was assessed by quantification of SLS-like hemolytic activity.

**Circular dichroism measurements** were performed on a Jasco J-815 spectropolarimeter. The temperature in the sample container was maintained using a Jasco PFD temperature stabilizer. Freshly purified MBP-ClosC and MBP-ClosC CxxC mutants were buffer-exchanged into 10 mM phosphate buffer using equilibrated PD-10 columns (GE Healthcare). Thereafter, MBP-ClosC and all MBP-CxxC mutants were prepared to a final protein concentration of 1.2 mg/ml. Before running the samples on a CD spectrometer, the machine was flushed with nitrogen gas for 1 h. Once the CD spectrometer was equilibrated, each sample was scanned 10× within a wavelength range of 185–300 nm. Data were processed by applying baseline subtraction, deconvolution, and noise reduction using Spectra Manager® (version 2). A xenon lamp was used as the light source. Sensitivity of the machine was set at 100 millidegrees. The scanning method was set to step fashion with a response time of 2 s. Bandwidth was set at 1, and two open channels, CD and dynode voltage, were operated. CD data were verified by repeating each experiment with different protein preparations, which resulted in consistent data. The CD results were expressed as a molar ellipticity ([θ]) in degrees cm² dmol⁻¹, as defined: [θ] = θ_{obs}/10nC_{p}, where θ_{obs} is the CD in millidegrees, n is the number of amino acids, l is the path length of the cell (cm), and C_{p} is the mole fraction. The [θ] value obtained at θ_{222} was used to calculate the helical content using the following equation: % helix = ([θ])_{222} − 4000)/33,000 × 100. The following secondary structure prediction programs were used to obtain a theoretical value of helical content for comparative purposes: PSIPRED, SOPMA, and nnPredict.

**Detection of Heterocyclized ClosA and Saga Peptides via Bottom-up Mass Spectrometry Analysis**

Sample Preparation—*In vitro* ClosA and synthetase samples, with the ability to lyse red blood cells, first were chemically derivatized by the addition of 50 mM 2-bromoethylamine (BrEA) in 200 mM Tris buffer (pH 8.8). The BrEA-lytic sample reaction was allowed to proceed for 16 h, at which time the samples were desalted by passing 70 μl of reaction mixture through a Bio-Rad Spin6 column. 50 μl of the flow-through was then trypsin-digested for 10 min, 20 min, 1 h, 4 h, and overnight by the use of a trypsin singles proteomic grade kit (Sigma). HPLC—Tryptic fragments were separated off-line by the use of an Agilent 1200 HPLC system equipped with an analytical C4
Clostridiolysin S, a Post-translationally Modified Bacteriocin

Preparation of ClosA Ala Mutant at Identified Heterocycle—A QuikChange kit was used to create the ClosA mutants using pET28b-MBP closA plasmid as a template for PCR. Further processing of the construct follows the standard preparation described for CxxC mutants.

Protein Purification of MBP-ClosA Ala Mutants—ClosA mutant was MBP-tagged and purified as described previously (6). Upon purification, the protein was immediately stored at −80 °C until needed for assays.

RESULTS

Comparative Bioinformatics of Clos and Sag Locus—Previously, we and others (6, 10–12) have identified a large, conserved family of bacteriocin gene clusters present across six phyla of microorganisms. These were identified by similarity to the microcin B17 (21–23) biosynthetic gene cluster and include the important human pathogens GAS, L. monocytogenes, S. aureus, and C. botulinum. Comparison of the gene clusters between the SLS biosynthetic operon in GAS and its ortholog in C. botulinum, displayed high levels of similarity both in gene content and organization (Fig. 1A). The SLS gene cluster contains the gene encoding the SLS precursor peptide, designated SagA, as well as genes encoding the SLS thiazole/oxazole synthetase (SagB–D), a putative immunity protein (SagE), a protein of unknown function (SagF), and three proteins containing the putative secretion system (SagG–I). All of these genes are present in the gene cluster of C. botulinum (ATCC 3502) in identical gene order and orientation to the SLS gene cluster, which we hereby refer to as the clostridiolysin S gene cluster.

ClosA–D Genetic Complementation Studies in Group A Streptococcus—To determine whether the clostridiolysin S genes are biochemically equivalent to the sag gene cluster in GAS, genetic complementation studies were performed using the sagA–D deletion strains of GAS NZ131 (M49 serotype) (ΔsagA, ΔsagB, ΔsagC, and ΔsagD) (3). We transformed these strains with a plasmid containing the corresponding wild-type copies of the clos genes (closA, closB, closC, and closD), to determine the extent of genetic tolerance between GAS and C. botulinum. The ability of the complemented strains to lyse erythrocytes was quantified by measuring the amount of hemoglobin released into the supernatant by measuring the absorbance at 450 nm (Fig. 1B). These were compared with the hemolytic activity of wild-type GAS, ΔsagA complemented with pDCerm-saga, and ΔsagA–D complemented with the pDCerm plasmid alone. The closA, closC, and closD complemented strains were able to lyse erythrocytes, whereas complementation of the GAS ΔsagB mutant with a plasmid containing the closB gene was not sufficient to produce a hemolytic phenotype, even though closB is 60% similar to SagB. Of the genes that were able to reconstitute the hemolytic phenotype, closC partially complemented, whereas the closA and closD complemented strains restored, near wild-type levels of hemolysis. This experiment

Protein Purification of MBP-ClosA Ala Mutants—ClosA mutant was MBP-tagged and purified as described previously (6). Upon purification, the protein was immediately stored at −80 °C until needed for assays.

RESULTS

Comparative Bioinformatics of Clos and Sag Locus—Previously, we and others (6, 10–12) have identified a large, conserved family of bacteriocin gene clusters present across six phyla of microorganisms. These were identified by similarity to the microcin B17 (21–23) biosynthetic gene cluster and include the important human pathogens GAS, L. monocytogenes, S. aureus, and C. botulinum. Comparison of the gene clusters between the SLS biosynthetic operon in GAS and its ortholog in C. botulinum, displayed high levels of similarity both in gene content and organization (Fig. 1A). The SLS gene cluster contains the gene encoding the SLS precursor peptide, designated SagA, as well as genes encoding the SLS thiazole/oxazole synthetase (SagB–D), a putative immunity protein (SagE), a protein of unknown function (SagF), and three proteins containing the putative secretion system (SagG–I). All of these genes are present in the gene cluster of C. botulinum (ATCC 3502) in identical gene order and orientation to the SLS gene cluster, which we hereby refer to as the clostridiolysin S gene cluster.

ClosA–D Genetic Complementation Studies in Group A Streptococcus—To determine whether the clostridiolysin S genes are biochemically equivalent to the sag gene cluster in GAS, genetic complementation studies were performed using the sagA–D deletion strains of GAS NZ131 (M49 serotype) (ΔsagA, ΔsagB, ΔsagC, and ΔsagD) (3). We transformed these strains with a plasmid containing the corresponding wild-type copies of the clos genes (closA, closB, closC, and closD), to determine the extent of genetic tolerance between GAS and C. botulinum. The ability of the complemented strains to lyse erythrocytes was quantified by measuring the amount of hemoglobin released into the supernatant by measuring the absorbance at 450 nm (Fig. 1B). These were compared with the hemolytic activity of wild-type GAS, ΔsagA complemented with pDCerm-saga, and ΔsagA–D complemented with the pDCerm plasmid alone. The closA, closC, and closD complemented strains were able to lyse erythrocytes, whereas complementation of the GAS ΔsagB mutant with a plasmid containing the closB gene was not sufficient to produce a hemolytic phenotype, even though closB is 60% similar to SagB. Of the genes that were able to reconstitute the hemolytic phenotype, closC partially complemented, whereas the closA and closD complemented strains restored, near wild-type levels of hemolysis. This experiment

Protein Purification of MBP-ClosA Ala Mutants—ClosA mutant was MBP-tagged and purified as described previously (6). Upon purification, the protein was immediately stored at −80 °C until needed for assays.

RESULTS

Comparative Bioinformatics of Clos and Sag Locus—Previously, we and others (6, 10–12) have identified a large, conserved family of bacteriocin gene clusters present across six phyla of microorganisms. These were identified by similarity to the microcin B17 (21–23) biosynthetic gene cluster and include the important human pathogens GAS, L. monocytogenes, S. aureus, and C. botulinum. Comparison of the gene clusters between the SLS biosynthetic operon in GAS and its ortholog in C. botulinum, displayed high levels of similarity both in gene content and organization (Fig. 1A). The SLS gene cluster contains the gene encoding the SLS precursor peptide, designated SagA, as well as genes encoding the SLS thiazole/oxazole synthetase (SagB–D), a putative immunity protein (SagE), a protein of unknown function (SagF), and three proteins containing the putative secretion system (SagG–I). All of these genes are present in the gene cluster of C. botulinum (ATCC 3502) in identical gene order and orientation to the SLS gene cluster, which we hereby refer to as the clostridiolysin S gene cluster.

ClosA–D Genetic Complementation Studies in Group A Streptococcus—To determine whether the clostridiolysin S genes are biochemically equivalent to the sag gene cluster in GAS, genetic complementation studies were performed using the sagA–D deletion strains of GAS NZ131 (M49 serotype) (ΔsagA, ΔsagB, ΔsagC, and ΔsagD) (3). We transformed these strains with a plasmid containing the corresponding wild-type copies of the clos genes (closA, closB, closC, and closD), to determine the extent of genetic tolerance between GAS and C. botulinum. The ability of the complemented strains to lyse erythrocytes was quantified by measuring the amount of hemoglobin released into the supernatant by measuring the absorbance at 450 nm (Fig. 1B). These were compared with the hemolytic activity of wild-type GAS, ΔsagA complemented with pDCerm-saga, and ΔsagA–D complemented with the pDCerm plasmid alone. The closA, closC, and closD complemented strains were able to lyse erythrocytes, whereas complementation of the GAS ΔsagB mutant with a plasmid containing the closB gene was not sufficient to produce a hemolytic phenotype, even though closB is 60% similar to SagB. Of the genes that were able to reconstitute the hemolytic phenotype, closC partially complemented, whereas the closA and closD complemented strains restored, near wild-type levels of hemolysis. This experiment
produced the initial indication that the orphan clostridium gene cluster indeed is functionally equivalent to the SLS cluster. During our investigations into the function of the SLS-like gene cluster from *C. botulinum*, the genome sequence of *S. pyogenes* became available (10). *S. pyogenes*, an opportunistic foodborne disease-causing bacteria, also contains the SLS-like gene cluster. In *S. pyogenes*, this gene cluster is 85–98% identical to the gene cluster found in *C. botulinum*. Because *S. pyogenes* is a BSL1 level organism, it was possible to generate integrated plasmid mutants of this gene cluster within our laboratories, enabling the direct assessment of the involvement of this gene cluster in the generation of an SLS-type toxin. Indeed, when closA and closC were inactivated, *S. pyogenes* no longer retained its hemolytic activity (Fig. 1C). Results of both the complementation of GAS with the *C. botulinum* genes and the targeted mutagenesis in *S. pyogenes* confirm that the SLS-like gene clusters from clostridia are involved in the bio-genesis of a hemolysin.

**In Vitro Reconstitution of Clostridiolysin S Activity**—Because the above genetic complementation experiments indicated that the *C. botulinum* gene cluster is functionally equivalent to the SLS gene cluster, we attempted to reconstitute hemolytic activity *in vitro*. To determine whether the precursor protein from *C. botulinum* ClosA could be transformed into an active cytolysin in a synthetase-dependent manner, ClosA was produced as an MBP-fusion protein for solubility and then subjected to post-translational modification by the addition of ClosBCD. Incubating ClosA with purified ClosBCD did not produce a lytic product (data not shown). The individual enzymes were analyzed and ClosB, the flavin mononucleotide (FMN)-dependent oxidoreductase, did not exhibit a yellow color upon purification. Several attempts to purify an active ClosB (yellow) failed under different affinity purification systems. The anticipated yellow color is indicative of FMN-binding proteins and adding FMN to the assay also failed to reconstitute the hemolytic activity. We surmised that the ClosB, as purified, was in an inactive apoprotein and hypothesized that this enzyme may undergo an additional processing event that could not be reconstituted in an *E. coli* system. Based on genetic complementation studies and insertion mutants, ClosB is active *in vivo*. We hypothesize that *in vitro* reconstitution of a hemolytic SLS-like toxin from *C. botulinum* to gauge the relative *in vitro* hemolytic activity of the SLS system, reconstitution of SLS activity was assayed in parallel. Based on *in vitro* hemolysis, SLS and SLS have comparable bioactivities (Fig. 2A). This observation is supported by the genetic complementation studies as the closD and closA genes produced a natural product with potency at near wild-type levels. To provide additional confirmation of the *in vitro*-reconstituted reaction, we set out to characterize ClosC, which is modified by the ClosBCD synthetase complex. ClosE and -F are potentially involved in immunity.

**FIGURE 1. Comparison of the streptolysin S-associated genetic cluster (Sag) with the SLS-like clostridiolysin S genetic cluster in *Clostridium botulinum* A. The Sag cluster in GAS contains a SLS precursor sequence (SagA), which is modified by the SagBCD synthetase complex. SagE and -F are potentially involved in immunity. SagA–I have homology to ABC transporters. The clostridiolysin S gene cluster contains genes similar to the Sag locus, in identical order, and is present in *C. botulinum* ATCC 3502, and *S. pyogenes* ATCC 15579. B, shown is the genetic complementation of Clos genes in GAS M1 Sag mutants. BSA-stabilized extracts were assayed for lytic activity. Bar 1, M1 wtGAS; bar 2, M1 GAS ΔsagA and wtsagA; bar 3, M1 GAS ΔsagA and wtclosA; bar 4, M1 GAS ΔsagB and wtclosB; bar 5, M1 GAS ΔsagC and wtclosC; bar 6, M1 GAS ΔsagD and wtclosD; bar 7, M1 GAS ΔsagA and plasmid alone; bar 8, M1 GAS ΔsagB and plasmid alone; bar 9, M1 GAS ΔsagC and plasmid alone; and bar 10, M1 GAS ΔsagD and plasmid alone. Hemolytic activity was normalized against a Triton X-100 positive control. C, *S. pyogenes* exhibits a strong β-hemolytic phenotype when grown on blood agar plates. ΔclosA and ΔclosC S. pyogenes allelic exchange mutants lose the ability to lyse erythrocytes. Bacteria were grown for 36 h at 37 °C under anaerobic conditions.
Clostridiolysin S, a Post-translationally Modified Biotoxin

Figure 2. Hemolytic activity of clostridiolysin is detectable in vitro. A, in vitro reconstitution of ClosA hemolytic activity. Synthetase reactions using MBP-ClosA and SagB/Clos/ClosD produce a hemolytic toxin. Hemolytic activity was normalized against a Triton X-100 positive control. Bar 1, wt SagA and SagBCD; bar 2, wt ClosA and SagB/Clos/ClosD; bar 3, wt ClosA only; bar 4, SagB/Clos/ClosD only; bar 5, wt ClosA and SagB/Clos/ClosD; bar 6, wt ClosA and ClosC/ClosD; bar 7, wt ClosA and SagB/ClosD; bar 8, SagA alone. B, in vitro reconstitution of ClosA hemolytic activity using CxxC mutants as wild-type ClosC surrogates. Bar 1, wt ClosA and SagB/Clos/ClosD; bar 2, wt ClosA and SagB/ALEC/ClosD wt; bar 3, ClosA and SagB/CLEA/ClosD; bar 4, wt ClosA and SagB/APAC/ClosD wt; bar 5, ClosA and SagB/APAC/ClosD; bar 6, wt ClosA and SagB/CPAA/ClosD. Hemolytic activity was normalized against a Triton X-100 positive control. C, CD spectra of MBP-ClosC and MBP-tagged CxxC mutants. CD spectra of MBP-ClosC (a negative minimum at 215 and 225 nm confirms α-helical integrity); □, CD for CPAA mutant; ◇, CD for CLEA mutant; ×, CD for APAC mutant; +, CD for APAA mutant; ○, CD for ALEC mutant. CD was repeated three times with different protein purifications and resulted in consistent results.

An interesting observation in the microcin system is that Zn2+ bound Zn2+ tightly to ClosC (Fig. 3B). However, sequence alignments of ClosC with the E1-type family of proteins, did not catalyze the hemolytic activity of ClosC (supplemental Fig. 1). However, sequence alignments with E1-type family members, did not predict the biochemical reaction consistent results.

The addition of EDTA did not affect the structure of ClosC. The structure of ClosC mutants, indicating the striking loss of α-helical content. Based on the molar ellipticity at 222 nm, the helical content of MBP-ClosC was determined to be 49%, in close agreement with the predicted content of 38–42%. The ClosC CxxC mutants, lacking the ability to bind Zn2+, showed the loss of the negative minima at 215 and 225 nm (Fig. 2C). The resulting CD spectra obtained by the CxxC mutants, indicating the striking loss of α-helical integrity, were unexpected results because the dramatic loss of α-helical integrity was caused by a single amino acid point mutation on the ClosC portion of the MBP-ClosC construct. The amylase resin purification of the MBP-ClosC CxxC mutant proteins was verified by SDS-PAGE (supplemental Fig. 4). Experiments were performed to rescue hemolytic activity by zinc titration of the CxxC proteins. Addition of zinc did not restore a hemolytic phenotype. Zinc titration experiments of the CxxC proteins were also assayed by CD. The CD showed a minor restoration of α-helical content, but the extent was not sufficient for hemolytic reconstitution (supplemental Fig. 5). These findings indicate that the tetracysteine coordination of Zn2+ is required for a properly folded ClosC, which can function in vitro and further solidify the in vitro reconstitution of a hemolytic toxin.

Detection of Heterocyclized ClosA Peptides via Off-line Separation and LC-MS/MS Methods—From the first reports of SLS activity in the early 1900s, structural characterization attempts have failed, and all attempts to characterize other hemolytic biotoxins from Listeria, Streptococci, or Clostridia have met the same fate (6, 7, 11, 28, 29). Therefore, any direct structural insight into a SLS-type toxin would be a major advancement. For the past three years, we also were unsuccessful when attempting to analyze this toxin by traditional approaches. Using intact protein mass spectrometry, observation of the ClosA protein by Fourier transform-ion cyclotron mass spectrometry was successful. However, upon addition of the synthetase enzymes (SagB, ClosC, and ClosD), the ClosA protein only and Cys residues, we cannot predict the biochemical reaction catalyzed by ClosC (supplemental Fig. 1). However, sequence alignments of ClosC with the E1-type family of proteins, displays that ClosC contains two conserved CxxC motifs that typically bind Zn2+ (Fig. 3B). In the microcin system, Zn2+ was originally proposed to serve as a Lewis acid, activating the amide carbonyl for cyclodehydration (22). Although that mechanism is certainly plausible, the crystal structures of ThiF, MoeB, and MccB have become available and show that the tetracysteine coordinated Zn2+ is 20 Å away from the active site pocket where ATP binds (25, 26). In addition to the CxxC motifs, the crystal structures of ThiF and MoeB resolved the amino acid residues critical for ATP binding. Of these residues, 43% are conserved in ClosC, indicating that it requires ATP for catalysis, which is in agreement with the requirement of ATP in the in vitro reconstitution assay (6). Based on the structural arguments, we predict that the tetracysteine coordinated Zn2+ in ClosC plays a structural role rather than a catalytic role. We therefore anticipated that the addition of EDTA could disrupt the structure of ClosC. The addition of EDTA did not affect the structure of ClosC as judged by CD nor did it kill the hemolytic activity (supplemental Fig. 2), similar to the observations by the homolog McbB in the microcin B17 pathway (22). Therefore, we determined the effect of the cysteine mutations on the Zn2+ content, in vitro reconstitution, and structural integrity of ClosC.

Point mutations in the CxxC motifs resulted in the abolishment of hemolytic activity and minimal levels of Zn2+ binding compared with wild-type ClosC (Fig. 2B and supplemental Fig. 3). CD measurements indicated that wild-type MBP-ClosC had two distinct negative minima at 215 and 225 nm, indicative of α-helical content. Based on the molar ellipticity at 222 nm, the helical content of MBP-ClosC was determined to be 49%, in close agreement with the predicted content of 38–42%. The ClosC CxxC mutants, lacking the ability to bind Zn2+, showed the loss of the negative minima at 215 and 225 nm (Fig. 2C). The resulting CD spectra obtained by the CxxC mutants, indicating the striking loss of α-helical integrity, were unexpected results because the dramatic loss of α-helical integrity was caused by a single amino acid point mutation on the ClosC portion of the MBP-ClosC construct. The amylase resin purification of the MBP-ClosC CxxC mutant proteins was verified by SDS-PAGE (supplemental Fig. 4). Experiments were performed to rescue hemolytic activity by zinc titration of the CxxC proteins. Addition of zinc did not restore a hemolytic phenotype. Zinc titration experiments of the CxxC proteins were also assayed by CD. The CD showed a minor restoration of α-helical content, but the extent was not sufficient for hemolytic reconstitution (supplemental Fig. 5). These findings indicate that the tetracysteine coordination of Zn2+ is required for a properly folded ClosC, which can function in vitro and further solidify the in vitro reconstitution of a hemolytic toxin.
could be observed as a heterogeneous species of overlapping ions in the Fourier transform-ion cyclotron mass spectrum, which provided little insight into the modified structure (supplemental Fig. 6). One of the limitations of intact protein mass spectrometry, especially when there are many forms of one protein, is that the signal is dispersed over a large number of different and overlapping masses, which are difficult, if not impossible, to characterize. Tandem mass spectrometry also consistently failed to identify the toxin portion of this protein. For other proteins of this size or larger such as BSA, GFP, a freestanding AT domain on the Group B Streptococcus hemolytic pathway, it was possible to observe the intact proteins with relative ease, indicating that our problem most likely was due to the troublesome biophysical properties of SLS-like biotoxins. In the case of clostridiolysin S, the overlap of heterologous isotopic distributions hindered the assignment of an accurate mass, rendering the top-down data uninterpretable. Therefore, we set out to capture the modifications on ClosA by a bottom-up mass spectrometry strategy.

Traditional proteomic approaches, such as iodoacetamide alkylation, trypsin digestions, and capillary LC-MS/MS analysis also failed to recover any of the peptides covering the C-terminal end of the protoxin. We assumed that the modified peptides (ClosA and SagA) were hydrophobic, and therefore, we applied approaches usually reserved for membrane proteins, such as dissolving in 75% formic or acetic acid. This too failed to recover any of these peptides. Because ClosA and SagA have an array of contiguous cysteines within the primary sequence (CCCCSCCCC), another approach was taken based on some classic enzymological studies carried out on acetoacetate decarboxylase (30). This particular enzyme decarboxylates acetoacetate via the formation of an imine with an active site lysine. One approach used to determine that the lysine was critical for the activity of acetoacetate decarboxylase was to replace it with a cysteine. This rendered the protein inactive. However, upon 2-bromoethylamine treatment, the cysteine is converted to S-aminoethyl cysteine, an unnatural side chain with steric and electronics complementary to lysine. To observe heterocycles on ClosA and SagA, alkylation of the cysteines with BrEA was performed. BrEA alkylation has three significant advantages: it 1) improves water solubility; 2) improves the ionization by introducing an amine, which is positively charged under our conditions; and 3) enables tryptic digestion after the modified cysteine, generating a ladder of peptides. Because the SLS proteins were available in large quantities, we first optimized this protocol using the in vitro reconstituted system of SagA and SagBCD. The in vitro-reconstituted sample was alkylated with 2-bromoethylamine, then trypsin-digested, and analyzed by nanocapillary LC-MS/MS. The resulting data were subjected to the tandem mass spectral processing algorithms, InSpecT, and Spectral Networks (19, 20). Examination of the reconstituted system, consisting of SagA and SagBCD, uncovered a peptide containing two oxazole moieties observed at positions Ser46 and Ser48 (see Fig. 5B, peptide 2), which were absent in the control samples that lacked SagBCD or a reaction mixture that excluded one of

![FIGURE 3. Alignment of ClosC in C. botulinum ATCC 3503 and MoeB/ThiF and proposed mechanisms.](image-url)
Indeed, as demonstrated in previous mutational analyses in the SLS system, these identified sites of heterocyclic residues contribute to the hemolytic properties of the toxin 

in vitro (7).

Similarly, the in vitro-reconstituted system containing ClosA, digested, and 2-bromoethylamine-alkylated, was subjected to nanocapillary LC-MS/MS analysis on an LTQ instrument. Spectral Networks processed 576,339 total spectra and identified 133,435 clusters. Of the identified clusters, 42 were observed in the protoxin-encoding, C-terminal half of ClosA (CCVSVS region; the N terminus is a leader peptide). Using this approach, Spectral Networks and InSpecT identified several reoccurring peptides as containing −20 Da losses that were absent in the control assays. The identification of −20 Da species in both the Sag and Clos in vitro-reconstituted samples represents the first direct information that SLS-type protoxin substrates accept the installation of heterocycles by the synthetase enzymes (Fig. 4). In the clostridiolysin S-reconstituted system, peptides with modifications at Thr\textsuperscript{11}, Thr\textsuperscript{12}, and Thr\textsuperscript{46} (Fig. 5B, peptide 1; supplemental Figs. 7 and 8) were identified as a heterogeneous mixture, in agreement with top-down mass spectrometry data. Although the data do not provide a complete structure, the post-translationally modified peptides confirm that the clostridiolysin S precursor undergoes extensive post-translational modification in vitro and contains a series of methyloxazole motifs. Because a heterocycle was identified at Thr\textsuperscript{46} (Fig. 5B, peptide 1) within ClosA, which follows the classical paradigm in the microcin B17 system of C-terminal amino acid residue heterocyclization, a mutant bearing an individual point mutation was generated to assess the importance of heterocyclic conversion at this residue on the overall cytolytic properties of clostridiolysin S.

In vitro synthetase reactions using the purified ClosA point mutant (T46A) and the synthetase enzymes were assessed for hemolytic activity. Substitution of an alanine at Thr\textsuperscript{46} had a dramatic effect on the in vitro hemolytic phenotype (Fig. 5A). In a similar fashion to in vitro reconstitution experiments in microcin B17 system, where an additional heterocycle was produced relative to the known structure (23), we believe heterocycles at positions Thr\textsuperscript{11} and Thr\textsuperscript{12} on the leader peptide are results of excessive in vitro processing because they are found on the proposed leader peptide of the protoxin (31).

**DISCUSSION**

Based on the striking genetic and functional similarities with the key GAS virulence factor SLS, we demonstrate that two
Clostridial species produce a related hemolytic factor. This article has provided evidence that pathogenic bacteria utilize synthetase enzymes to produce chemical changes in ribosomally encoded precursors to biosynthesize biologically active toxins, which gain activity from the installation of heterocyclic moieties. Given the functional similarities and equivalent potencies, it is likely that the metabolic output of clostridiolysin S by C. botulinum can play a role in host tissue injury and contribute to clostridial virulence. While recognizing its broader classification as a bacteriocin-like peptide, it also could be fruitful to explore whether clostridiolysin S may aid C. sporogenes, a known component of the human gut microbiome, to establish a competitive niche within the vast number of microbes comprising the gut microflora. The mass spectrometry approach outlined herein is the first step toward the characterization of this class of hemolytic/cytolytic toxins. Preliminary results on the structural characterization of the SLS biosynthetic pathway have further solidified the described mass spectral method to be significant and reproducible for this family of toxins. Our findings on the structure/function of the ClosC subunit are likely to extend to other cyclodehydratases from this family of bacteriocin synthetase enzyme complexes. Because targeting virulence factors is becoming an attractive anti-infective strategy, selective inhibition of this family of enzymes may well become an attractive therapeutic target or could be used as an approach to prevent future food spoilage.

Acknowledgments—We greatly appreciate Jane Yang for critically reading the manuscript, Ted Huston for performing ICP-MS analysis, Theresa Smith and Gary Xie for performing comparative sequence alignments of the clos cluster against their available clostridia genomes, and Richard Kleman and Mark Lortie (University of California, San Diego, Pathology Biomarker and Diagnostic Discovery Center) for use of the LTQ-MS in the validation experiments.

REFERENCES

1. Yang, Y. L., Xu, Y., Straight, P., and Dorrestein, P. C. (2009) Nat. Chem. Biol. 5, 885–887
2. Severinov, K., Semenova, E., Kazakov, A., Kazakov, T., and Gelfand, M. S. (2007) Mol. Microbiol. 65, 1380–1394
3. Datta, V., Myaskowski, S. M., Kwinn, L. A., Chiern, D. N., Varik, N., Kansal, R. G., Koth, M., and Nizet, V. (2005) Mol. Microbiol. 56, 681–695
4. Nizet, V., Beall, B., Bast, D. I., Datta, V., Kilburn, L., Low, D. E., and De Azavedo, J. C. (2000) Infect. Immun. 68, 4245–4254
5. Humar, D., Datta, V., Bast, D. J., Beal, B., De Azavedo, J. C., and Nizet, V. (2002) Lancet 359, 124–129
6. Lee, S. W., Mitchell, D. A., Markley, A. L., Hensler, M. E., Gonzalez, D., Wohlrab, A., Dorrestein, P. C., Nizet, V., and Dixon, J. E. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 5879–5884
7. Mitchell, D. A., Lee, S. W., Pence, M. A., Markley, A. L., Limm, J. D., Nizet, V., and Dixon, J. E. (2009) J. Biol. Chem. 284, 13004–13012
8. Hara-Kudo, Y., Ogura, A., Noguchi, Y., Terao, K., and Kumagai, S. (1997) Microb. Pathog. 22, 31–38
9. Savani, J., and Harris, N. D. (1978) J. Food Sci. 43, 222–224
10. Sebaia, M., Peck, M. W., Minton, N. P., Thomson, N. R., Holden, M. T., Mitchell, W. J., Carter, A. T., Bentley, S. D., Mason, D. R., Crossman, L., Paul, C. J., Ivens, A., Wells-Bennik, M. H., Davis, I. J., Cerdeño-Tárraga, A. M., Churcher, C., Quail, M. A., Chillingworth, T., Feltwell, T., Fraser, A., Goodhead, I., Hance, Z., Jagels, K., Larke, N., Maddison, M., Moule, S., Mungall, K., Norbertczak, H., Rabinowitsch, E., Sanders, M., Simmonds, M., White, B., Whitehead, S., and Parkhill, J. (2007) Genome Res. 17, 1082–1092
11. Cotter, P. D., Draper, L. A., Lawton, E. M., Daly, K. M., Groeger, D. S., Casey, P. G., Ross, R. P., and Hill, C. (2008) PLoS Pathog. 4, e1000144
12. Schmidt, E. W., Nelson, J. T., Rasko, D. A., Sudek, S., Eisen, J. A., Haygood, M. G., and Ravel, J. (2005) Proc. Natl. Acad. Sci. U.S.A. 102, 7315–7320
13. Liao, R., Duan, L., Lei, C., Pan, H., Ding, Y., Zhang, Q., Chen, D., Shen, B., Yu, Y., and Liu, W. (2009) Chem. Biol. 16, 141–147
14. Kelly, W. L., Pan, L., and Li, C. (2009) J. Am. Chem. Soc. 131, 4327–4334
15. Morris, R. P., Leed, J. A., Naegeli, H. U., Oberer, L., Memmert, K., Weber, E., LaMarche, M. J., Parker, C. N., Burrer, N., Esterow, S., Hein, A. E., Schmitt, E. K., and Krastel, P. (2009) J. Am. Chem. Soc. 131, 5946–5955
16. Wieland Brown, L. C., Acker, M. G., Claridy, J., Walsh, C. T., and Fischer, M. A. (2009) Proc. Natl. Acad. Sci. U.S.A. 106, 2549–2553
17. Duda, D. M., Walden, H., Sondouris, J., and Schullman, B. A. (2005) J. Biol. Biol. 349, 774–786
18. Jurgenson, C. T., Burns, K. E., Begley, T. P., and Elacki, S. E. (2008) Biochemistry 47, 10354–10364
19. Tanner, S., Shu, H., Frank, A., Wang, L. C., Zandi, E., Mummy, M., Pevzner, P. A., and Bafna, V. (2005) Anal. Chem. 77, 4626–4639
20. Bandeira, N., Tsur, D., Frank, A., and Pevzner, P. A. (2007) Proc. Natl. Acad. Sci. U.S.A. 104, 6140–6145
21. Li, Y. M., Milne, J. C., Madison, L. L., Kolter, R., and Walsh, C. T. (1996) Science 274, 1188–1193
22. Zamble, D. B., McClure, C. P., Penner-Hahn, J. E., and Walsh, C. T. (2000) Biochemistry 39, 16190–16199
23. Zamble, D. B., Miller, D. A., Heidelberg, J. G., Maxwell, A., Walsh, C. T., and Hoffelder, F. (2001) Proc. Natl. Acad. Sci. U.S.A. 98, 7712–7717
24. Roush, R. F., Nolan, E. M., Löh, F., and Walsh, C. T. (2008) J. Am. Chem. Soc. 130, 3603–3609
25. Regni, C. A., Roush, R. F., Miller, D. J., Nourse, A., Walsh, C. T., and Schuffman, B. A. (2009) EMBO J. 28, 1953–1964
26. Lehmann, C., Begley, T. P., and Elacki, S. E. (2006) Biochemistry 45, 11–19
27. Igarashi, Y., Kan, Y., Fujii, K., Fujita, T., Harada, K., Naoki, H., Tabata, H., Onaka, H., and Furumai, T. (2001) J. Antibiok. 54, 1045–1053
28. Cinader, B., and Pillaire, L. (1950) J. Exp. Med. 92, 219–237
29. Tsaihong, J. C., and Wennerstrom, D. E. (1983) Curr. Microbiol. 9, 333–338
30. Hopkins, C. E., O’Connor, P. B., Allen, K. N., Costello, C. E., and Tolan, D. R. (2002) Protein Sci. 11, 1591–1599
31. Oman, T. J., and van der Donk, W. A. (2010) Nat. Chem. Biol. 6, 9–18