The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease

Holger Brüggemann*, Sebastian Bäumer*, Wolfgang Florian Fricke*, Arnim Wiezer*, Heiko Liesegang*, Iwona Decker*, Christina Herzberg†, Rosa Martínez-Arias*, Rainer Merkl*, Anke Henne*, and Gerhard Gottschalk*†‡

*Göttingen Genomics Laboratory and †Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August-University, Grisebachstrasse 8, D-37077 Göttingen, Germany

Edited by Dieter Söll, Yale University, New Haven, CT, and approved December 11, 2002 (received for review September 27, 2002)

Tetanus disease is one of the most dramatic and globally prevalent diseases of humans and vertebrate animals, and has been reported for over 24 centuries. The manifestation of the disease, spastic paralysis, is caused by the second most poisonous substance known, the tetanus toxin, with a human lethal dose of ≈1 ng/kg. Fortunately, this disease is successfully controlled through immunization with tetanus toxoid; nevertheless, according to the World Health Organization, an estimated 400,000 cases still occur each year, mainly of neonatal tetanus. The causative agent of tetanus disease is *Clostridium tetani*, an anaerobic spore-forming bacterium, whose natural habitat is soil, dust, and intestinal tracts of various animals. Here we report the complete genome sequence of toxigenic *C. tetani* E88, a variant of strain Massachusetts. The genome consists of a 2,799,250-bp chromosome encoding 2,372 ORFs. The tetanus toxin and a collagenase are encoded on a 74,082-bp plasmid, containing 61 ORFs. Additional virulence-related factors could be identified, such as an array of surface-layer and adhesion proteins (35 ORFs), some of them unique to *C. tetani*. Comparative genomics with the genomes of *Clostridium perfringens*, the causative agent of gas gangrene, and *Clostridium acetobutylicum*, a nonpathogenic solvent producer, revealed a remarkable capacity of *C. tetani*: The organism can rely on an extensive sodium ion bioenergetics. Additional candidate genes involved in the establishment and maintenance of a pathogenic lifestyle of *C. tetani* are presented.

Neurotoxigenic clostridia have attracted considerable attention during the past decades. The mode of action of the strongest toxins known to mankind, the tetanus toxin produced by *Clostridium tetani* and the homologous botulinum toxins A–G produced by various *Clostridium botulinum* strains, is now well understood (1–4). The tetanus toxin (TeTx) blocks the release of neurotransmitters from presynaptic membranes of inhibitory neurons of the spinal cord and the brainstem of mammals; it catalyzes the proteolytic cleavage of the synaptic vesicle protein synaptobrevin (5). This leads to continuous muscle contractions, which are primarily observed in jaw and neck muscles (lockjaw). Since the introduction of a potent vaccine during World War II, cases of tetanus disease have been only sporadic in industrial countries. However, the disease, and in particular neonatal tetanus (NT), is still an important cause of death due to insufficient immunization (4, 6). NT is the second leading cause of death from vaccine-preventable diseases among children worldwide. It is considered endemic to 90 developing countries, with an estimated 248,000 deaths occurring in 1997 (World Health Organization; www.who.int/vaccines-diseases/diseases/NeonatalTetanus.shtml).

So far, genetic information on *C. tetani* is mainly restricted to the nucleotide sequences of the tetanus toxin TeTx and of its direct transcriptional activator TetR, both of which are encoded on a plasmid, designated pCL1 in the strain Massachusetts (7–10). The identification and analysis of all genes in the genome of *C. tetani* will contribute to our understanding of the lifestyle switch from a harmless soil bacterium to a potentially devastating neurosystem-damaging organism after entering and infecting a mammalian host. The presented information may also help to unravel the regulatory network governing tetanus toxin formation, which is still poorly understood (11).

Methods

Sequencing Strategy. Genomic DNA of *C. tetani* E88, a nonsporulating variant of strain Massachusetts used in vaccine production, was extracted and sheared randomly. Several shotgun libraries were constructed by using size fractions ranging from 1 to 3 kb. A cosmid library was constructed from *Tas*I partially digested genomic DNA cloned in the cosmid vector SuperCos1 (Stratagene). Insert ends of the recombinant plasmids and cosmids were sequenced by using dye-terminator chemistry with MegaBACE 1000 and ABI Prism 377 DNA automated sequencers (Amersham Biosciences and Applied Biosystems). Sequences were processed with PHRED and assembled into contigs by using the PHRAP assembling tool (12). Sequence editing was done by using GAP4, which is part of the STADEN package software (13). A coverage of 9.2-fold was obtained after the assembly of 44,500 sequences. To solve problems with misassembled regions caused by repetitive sequences and to close remaining sequence gaps, PCR-based techniques and primer walking on recombinant plasmids and cosmids were applied.

ORF Prediction and Annotation. Initial ORF prediction was accomplished by using the GLIMMER program (www.tigr.org/software/glimmer/). The result was verified and improved manually by using criteria such as the presence of a ribosome-binding site and codon usage analysis. Annotation was done first automatically by the ERGO annotation tool (Integrated Genomics, www.integratedgenomics.com), which was verified and refined by searching derived protein sequences against the public GenBank/European Molecular Biology Laboratory databases by using the BLAST program (14).

Comparative Genomics. For comparative analysis, each ORF of *C. tetani* was searched against all ORFs of *Clostridium perfringens* and *Clostridium acetobutylicum* by using the BLAST program (14). Significant homology was defined as >30% amino acid identity over >60% of both the *C. tetani* protein sequence and the homologous sequence of *C. perfringens* or *C. acetobutylicum*, respectively.

Results and Discussion

General Features of the *C. tetani* Genome. The chromosome of *C. tetani* contains 2,799,250 bp with a G+C content of 28.6%. The 74,082-bp plasmid pE88 exhibits a remarkably low G+C content of 24.5%. The replication origin of the chromosome was

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: SLP, surface-layer protein.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database [accession nos. AE015927 (chromosome) and AF528097 (pE88)].

‡To whom correspondence should be addressed. E-mail: ggottsc@gwdg.de.

www.pnas.org/cgi/doi/10.1073/pnas.0335853100

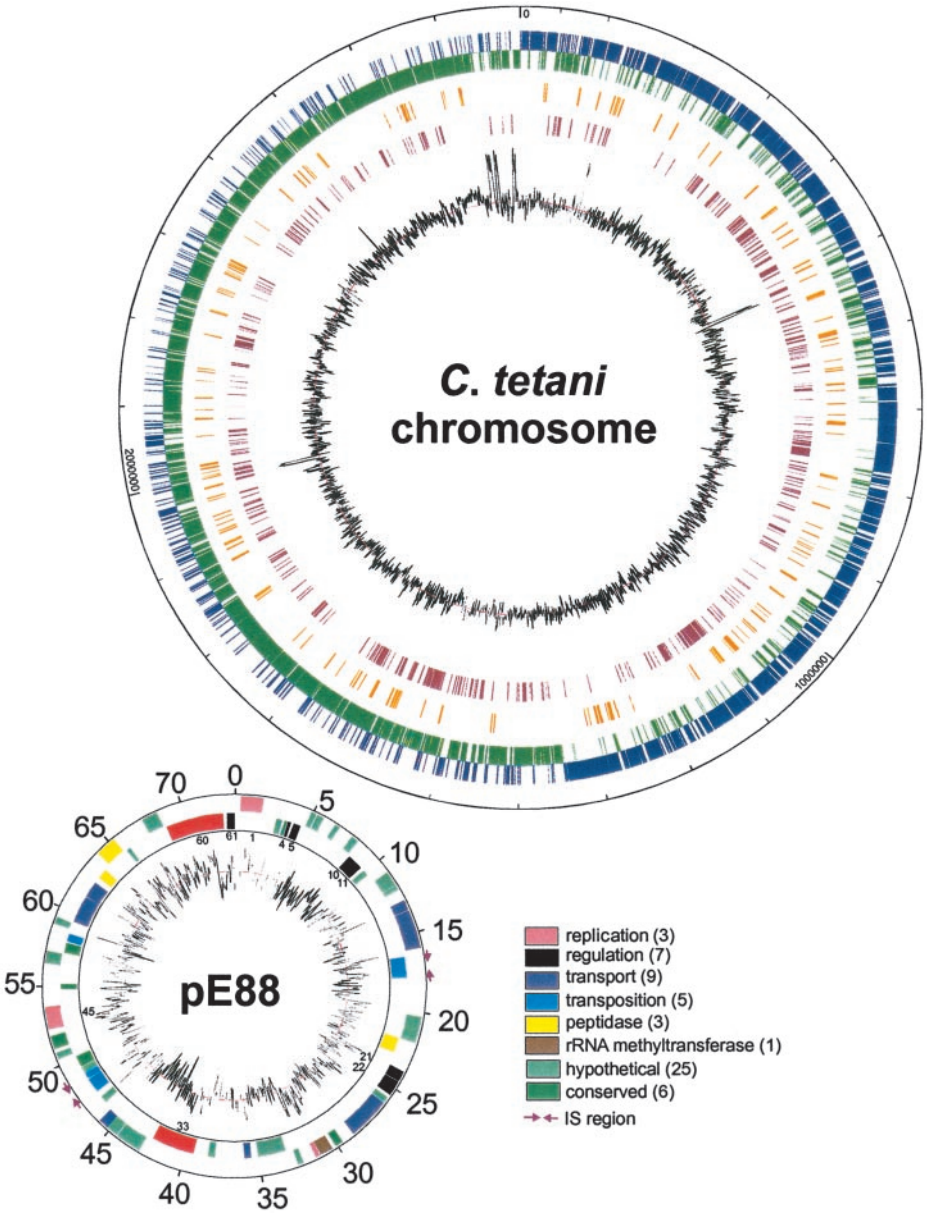

Fig. 1. Circular maps of the chromosome and the plasmid pE88 of *C. tetani*. The coding sequence of the chromosome is shown in blue or green, depending on strand orientation. ORFs of *C. tetani* that have homologous proteins in *C. perfringens* (15) but not in *C. acetobutylicum* (16) are shown in orange. ORFs of *C. tetani*, which are not present in the two other mentioned clostridial genomes, are shown in pink. The plasmid map shows ORFs color-coded according to their assigned functions. The gene of the tetanus toxin, *tetX* (CTP60), and the gene encoding a collagenase, *colT* (CTP33), are highlighted in red. Numbers at the inner ring refer to genes mentioned in the text. Most inner rings of both maps show the G+C content variation (higher values outward).

identified on the basis of GC-skew analysis, the distinctive inflection point in the coding strand and the vicinal presence of characteristic replication proteins such as DnaA (Fig. 1; also published as Figs. 5 and 6 in supporting information on the PNAS web site, www.pnas.org). Eighty-two percent of the 2,368 predicted ORFs on the chromosome are transcribed in the same direction as DNA replication. A similar figure (83%) was calculated from the sequence data of *C. perfringens* (15). This distinctive codirectionality of replication and transcription is not seen in the genomes of pathogens such as *Vibrio cholerae* or *Yersinia pestis* or in archaeal genomes such as *Methanosarcina mazei* (17–19). Only a few mobile elements could be identified in the genome of *C. tetani*: 16 transposase genes are present, which can be classified in four different insertion sequence families. Most of these genes seem to be nonfunctional because of insertions, deletions, and point mutations resulting in frame-shifted or degenerated ORFs. G+C variation within the genome is very low; the only regions with a significantly higher G+C content (≈50%) are those harboring the six rRNA gene clusters and genes encoding ribosomal proteins. This lack of G+C fluctuation indicates that recent events of gene acquisition by lateral gene transfer have not taken place, and that the *C. tetani* genome is more stable than the genomes of enteropathogens.

The Tetanus Toxin-Encoding Plasmid pE88. It was known that the 74-kb *C. tetani* plasmid pE88 harbors the genes for the tetanus toxin (*tet*X) and its direct transcriptional regulator TetR (7–10). We have now identified 61 ORFs, 28 of which show similarities to known proteins with assigned functions (Fig. 1). Interestingly,

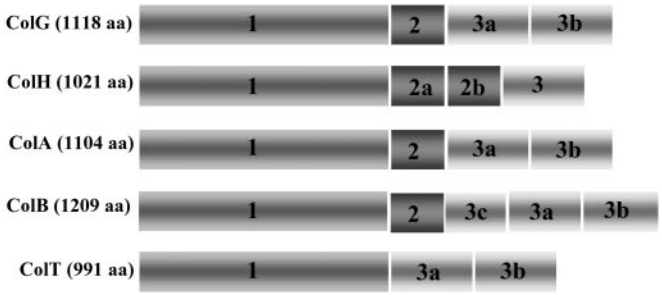

Fig. 2. Comparison of clostridial collagenases. Segment comparison of ColG and ColH of *C. histolyticum* (21), ColA of *C. perfringens* (22), ColB of *C. botulinum* (Sanger Centre, *C. botulinum* sequencing project; online access: ftp://ftp.sanger.ac.uk/pub/pathogens/cb/), and ColT of *C. tetani*. Bacterial collagenases consist of three different segments (21). Segment 1 represents the catalytic domain containing the consensus motif for zinc proteases, HEXXH. Segment 2 contains the so-called PKD domain (PF00801) of unknown function. The name refers to the polycystic kidney disease protein PKD1, in which the domain was detected in multiple copies. Segment 3 is thought to be the collagen-binding domain.

pE88 encodes an additional virulence factor, a 114-kDa collagenase, designated ColT (CTP33), an enzyme that may play an important role in *C. tetani* pathogenesis because of its activity to destroy tissue integrity in the infected host (20). A multiple sequence alignment with other known clostridial collagenases of *Clostridium histolyticum* (21), *C. perfringens* (22) and the one encoded on the genome of *C. botulinum* (Sanger Centre, Cambridge, UK, *C. botulinum* sequencing project), designated ColB, revealed a different domain structure: ColT lacks segment 2, the so-called PKD domain of unknown function (Fig. 2). Among clostridial collagenases, ColT shows the strongest similarity to ColB of *C. botulinum* with 79% amino acid identity over segment 1.

Furthermore, nine transport-related genes similar to ATP-binding cassette (ABC) transporters or associated permeases are encoded on pE88 as well as eight additional genes coding for proteins with multiple transmembrane helices. Together, these proteins are apparently part of five different multisubunit ABC transporters; most of them show strongest similarity to protein/peptide-transporting systems.

Of special interest for an understanding of the regulatory network governing the tetanus toxin formation are seven regulatory proteins on pE88. Among them is a two-component system (CTP21/CTP22), which shows high similarities to systems of other Gram-positive bacteria such as *Bacillus cereus*, *Bacillus halodurans*, and *C. botulinum*. Further regulatory proteins, including TetR, are summarized in Table 1. They show some similarity to each other, as well as to TxeR from *Clostridium difficile*, the alternative σ-factor regulating the synthesis of the toxins ToxA and ToxB (23), and to UviA of *C. perfringens*, a putative positive regulator of the UV-inducible bacteriocin gene (24). CTP05 seems to be a pseudogene disrupted by an authentic frameshift.

The origin of pE88 remains unclear. Over 50% of all ORFs on pE88 are unique to *C. tetani*. A mobile element, present in two copies, shows strongest similarity to a transposase of *B. cereus* (60% identity over 355 amino acids). The putative replication protein of pE88 is duplicated (CTP1 and CTP45). It shows homology only to a protein (40% identity over 438 amino acids) encoded on the plasmid of the alkaliphilic *Bacillus* strain KSM-KP43 and to the replication protein (39% identity over 418 amino acids) on pIP404, the bacteriocin- and UviA-encoding 10-kb plasmid of *C. perfringens* CPN50 (25).

Virulence Factors on the Chromosome. In addition to pE88, the *C. tetani* chromosome was searched for virulence genes (Table 2). ORFs for tetanolysin O, hemolysin, and fibronectin-binding proteins were detected exhibiting homology to ORFs in other clostridial species including *C. perfringens* (15, 26). A difference from *C. perfringens* is the presence of an array of surface-layer proteins (SLPs) in *C. tetani*, all of which are absent in *C. perfringens*: 19 homologues of Cwp66, an adhesin of *C. difficile*, were found, all of which possess multiple copies of the putative cell wall-binding domain PF04122 identified in characterized SLPs (27, 28). Two proteins with multiple leucine-rich repeat domains similar to Internalin A could be detected, a protein mediating binding of *Listeria monocytogenes* to host epithelial cells (29). In addition, 11 SLP-related ORFs were identified, which have no homologues in other sequenced clostridial genomes but contain characterized domains of surface proteins such as bacterial Ig-like domains (PF02368, PF02369). These proteins, as well as the Cwp66 homologues, possess a characteristic leader peptide, suggesting that these proteins are exported. The genes of the SLPs were found to be clustered; otherwise, pathogenicity islands or mobile elements in flanking regions of virulence genes could not be detected.

Comparative Genomics. Which other gene products could contribute to the pathogenic lifestyle of *C. tetani*? To approach this question, the genome of *C. tetani* was compared with the genomes of both *C. perfringens* (15), and the nonpathogenic *C. acetobutylicum* (16) and homologous ORFs were identified. Focusing on *C. tetani*, four groups of ORFs could be distinguished (Fig. 3). Two of them are of special interest: the 199 ORFs that are common to both *C. tetani* and *C. perfringens*, and the 516 ORFs of *C. tetani*, which have no homologues in *C. perfringens* and *C. acetobutylicum* (see Tables 3 and 4, which are in supporting information on the PNAS web site, for the complete ORF lists of these groups). Remarkably, ORFs encoding 27 peptidases were found within these two groups, many of them being predicted zinc-metalloproteases. Seven ORFs for lipid degradation and 15 ORFs for ethanolamine utilization similar to the *eut* genes of the genus *Salmonella* (30) were detected as well as 21 ORFs for amino acid decomposition, including glutamate mutase, tryptophanase, histidine ammonia lyase, and lysine-2,3-aminomutase. The apparent capability of degrading many amino acids is in contrast to the lack of genes for sugar utilization, which are abundant in *C. acetobutylicum* and *C. perfringens*. Striking is the

Table 1. Sequence similarities of regulatory elements encoded on pE88 to the characterized regulatory proteins UviA of *C. perfringens*, TxeR of *C. difficile*, and BotR/A of *C. botulinum*

ORF name and size*	Annotation	UviA (185 aa)	TxeR (184 aa)	BotR/A (178 aa)
CTP5 (172 aa)	Frameshift, UviA homolog	32% (48%) >185 aa	31% (49%) >170 aa	26% (43%) >178 aa
CTP10 (179 aa)	Putative σ-factor/DNA-binding protein	24% (43%) >184 aa	28% (47%) >82 aa	n.s.
CTP11 (184 aa)	Putative σ-factor/DNA-binding protein	25% (44%) >180 aa	24% (44%) >179 aa	n.s.
TetR (177 aa)	Positive regulator of *tet*X	27% (49%) >183 aa	26% (47%) >184 aa	67% (85%) >177 aa

Percentages indicate identical (in parentheses: similar) amino acids. n.s., not significant.

*Not listed here is the two-component system CTP21/CTP22 and CTP4 (71 aa), which shows homology only to UviB (64 aa) of *C. perfringens* (40% identity over 67 aa).

Table 2. Virulence-related genes in *C. tetani*

ORF	Product	Similar sequences	Comment
tetX	Tetanus toxin	Botulinum toxin (*Cb*)	Unique in *C. tetani* and *Cb*
colT	Collagenase	*Cp, Cb, Ba, C. histolyticum*	Different domain structure of ColT (Fig. 2)
CTC164, CTC1606	Fibronectin-binding protein	*Cp, Cb, Cd, Lm, Li, Ba*	Enables bacteria to colonize wound tissues and blood clots
CTC594	Putative PAF acetylhydrolase	*S. pneumoniae, S. mutans, S. pyogenes*	Inactivates PAF through cleavage of an acetyl group
CTC586	Hemolysin III	*Cp, Cb, Ca, Ba, B. cereus*	Predicted membrane protein with hemolytic activity
CTC1814	Periplasmic immunogenic protein	*Ba, Brucella melitensis*	Outer membrane protein
CTC1855	Myosin-crossreactive antigen	*Cp, Lm, Li, Cb, S. pyogenes*	Similar with class II major histocompatibility antigens
CTC1888	Tetanolysin O	Perfringolysin O (*Cp*), *S. pyogenes, Lm*	Catalyzes the lysis of cholesterol-containing membranes
CTC2255, CTC2261	Virulence factor MviN	*Cb, Ca, Cd, Neisseria meningitidis*	Membrane protein
CTC462, CTC491, CTC504, CTC515*	SlpA/Cwp66 homologue	*Cd* (29 homologues), *Cb* (8 homologues)	SLP with adhesive properties, mediates adherence to host tissue
CTC494, CTC495	Internalin A homologue	*Lm, Li, Cb, Ca*	Internalin A binds to E-cadherin of host epithelial cells
CTC465, CTC747, CTC749, CTC750†	SLP and/or adhesion protein	—	Contain domains found in surface proteins such as intimin

Ba, *B. anthracis*; *Ca*, *C. acetobutylicum*; *Cb*, *C. botulinum*; *Cd*, *C. difficile*; *Cp*, *C. perfringens*; *Li*, *L. innocua*; *Lm*, *L. monocytogenes*; *S*, *Streptococcus*; PAF, *platelet activating factor*.
*Additional ORFs: CTC518, CTC520, CTC521, CTC523, CTC691, CTC696, CTC748, CTC774, CTC775, CTC1202, CTC1364, CTC1872, CTC1980, CTC2092, and CTC2093.
†Additional ORFs: CTC767, CTC769, CTC770, CTC771, CTC772, CTC776, and CTC777.

presence of 35 genes for sodium ion-dependent systems (Fig. 4). The primary source for the ion-motive force at the cytoplasmic membrane of *C. tetani* is a vacuolar(V)-type ATPase similar to the ATPase of *Enterococcus hirae*, which is a Na^+ pump (35). *C. tetani* does not harbor the genes of the F_0F_1-type ATPase; these are present in *C. acetobutylicum* and *C. perfringens*. The lack of the latter genes in *C. tetani* is highly exceptional. Therefore, part of the ATP synthesized in the course of the fermentative metabolism of *C. tetani* is hydrolyzed by the V-type ATPase to generate a Na^+-motive force at the membrane. This force is used to drive various transport processes: several sodium-dependent substrate cotransporters are present, and at least six of the multidrug-resistance exporters seem to be driven by Na^+ intrusion, whereas there are apparently no proton-dependent systems present. As a common coupling device between proton and sodium ion gradients, several H^+/Na^+ antiporters were detected, challenging the search for a primary proton pump: a proton-translocating pyrophosphatase showing similarity to systems in plant vacuoles is present in *C. tetani* but absent in other clostridia. Furthermore, the genes for a membrane-bound electron transport chain could be identified in the genomes of *C. tetani* and *C. perfringens*. This system shows strong similarity to the *Rhodobacter*-specific nitrogen fixation system of *Rhodobacter capsulatus* as well as to the system NADH:ubiquinone oxidoreductase found in many aerobic pathogens such as *V. cholerae*, *Salmonella typhimurium*, *Y. pestis*, and *Haemophilus influenzae*, in which it apparently functions as a Na^+-translocating NADH dehydrogenase (32, 36). Because there are no reports for quinones in clostridia, such a NADH dehydrogenase could be involved in the fermentation of amino acids by *C. tetani*, as indicated in Fig. 4. A highly active NADH dehydrogenase was detected in washed membranes of the closely related *Clostridium tetanomorphum* (W. Buckel, Philipps-University, Marburg, Germany, personal communication). This may be the system that is missing for an understanding of electron flow from reduced ferredoxin, generated in the pyruvate-ferredoxin oxidoreductase reaction, via NADH to the NADH-consuming dehydrogenases of the butyrate pathway.

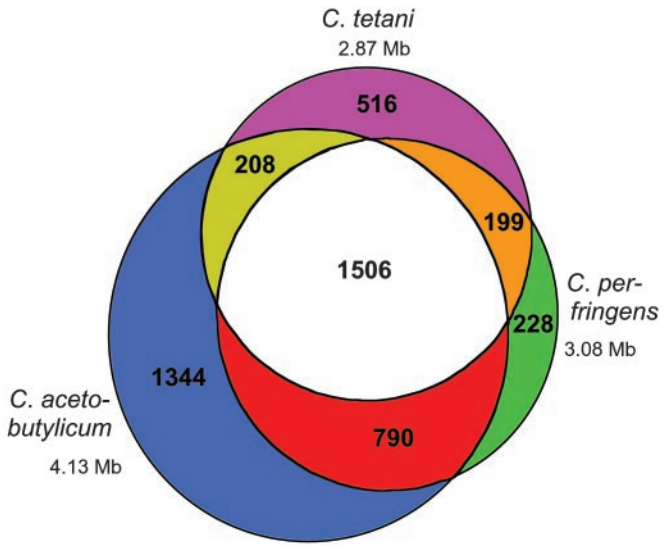

Fig. 3. Occurrence of homologous ORFs in *C. tetani*, *C. perfringens*, and *C. acetobutylicum*. Clostridial ''backbone'' ORFs, white; ORFs in *C. tetani* with homologues in *C. perfringens* but not in *C. acetobutylicum*, orange; *C. tetani* ORFs present in *C. acetobutylicum* but absent in *C. perfringens*, red; *C. tetani* ORFs not present in *C. perfringens* and *C. acetobutylicum*, pink. The size of the circles is proportional to the number of ORFs in each organism.

Protein Secretion. Only little is known about protein secretory systems in *Clostridia*. Until now, it has not been understood how the tetanus toxin, which lacks a typical N-terminal signal peptide, is exported. The export of toxA and toxB of *C. difficile* remains obscure as well, thus the presence of an alternative, so far unknown mechanism for toxin export was postulated (37). Because protein secretion is an important part in establishing a pathogenic phenotype, it is worth having a closer look at secretory systems and secreted proteins of *C. tetani*.

MICROBIOLOGY

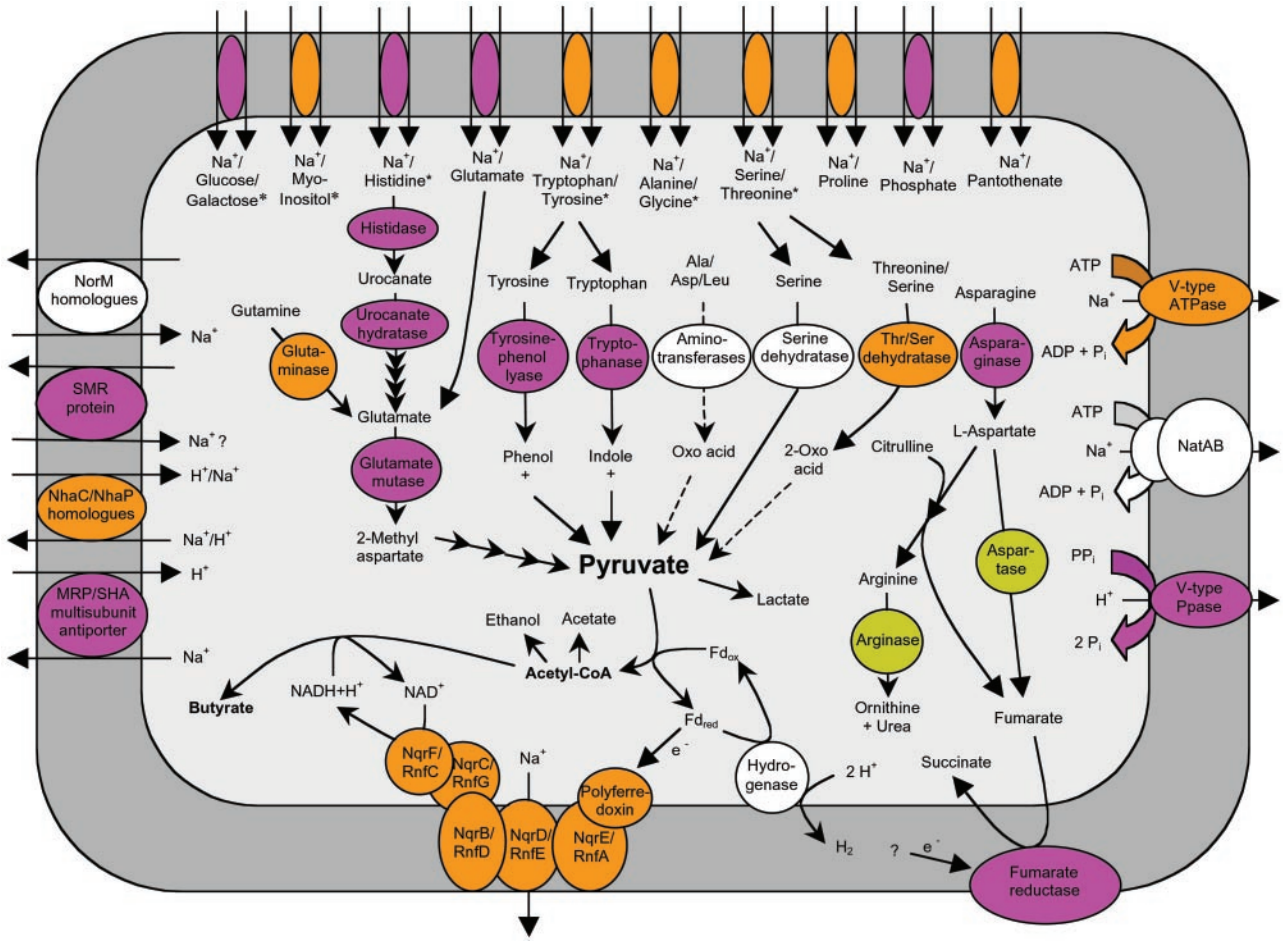

Fig. 4. Role of Na^+ in amino acid fermentations by *C. tetani*. The color code is as in Fig. 3. Information on the NADH:ubiquinone oxidoreductase/*Rhodobacter*-specific nitrogen fixation electron transport system is published as Fig. 7 in supporting information on the PNAS web site. The substrate specificity of many Na^+-dependent symporters (tagged with *) remains obscure (31). V-type ATPase, vacuolar-type ATPase; V-type Ppase, vacuolar-type pyrophosphatase; NatAB, an ATP-binding cassette transporter (32); Fd, ferredoxin; NorM, a Na^+-translocating multidrug transporter characterized in *Vibrio parahaemolyticus* (33); SMR, small multidrug resistance family; NhaC and NhaP, one-subunit secondary H^+/Na^+ antiporters; MRP/SHA, this seven-subunit H^+/Na^+ antiporter with similarities to hydrophobic components of the respiratory Complex I has been characterized in *Bacillus subtilis* and *Staphylococcus aureus* (34).

A characteristic signal peptide, recognized by the Sec-dependent secretory mechanism, could be detected in 419 deduced proteins of *C. tetani*. Of these, 101 proteins are involved in transport, inclusively periplasmic substrate-binding proteins. About 160 are conserved proteins of unknown function, most of them strongly hydrophobic. Virulence and surface proteins, peptidases, sporulation-specific proteins, as well as sensory proteins like sensor histidine kinases and chemotaxis proteins have a predicted N-terminal signal peptide. Thus, the well-studied prokaryotic Sec-system (38) is apparently the major protein translocation apparatus in *C. tetani*. However, not all *sec* genes of *Escherichia coli* and related Gram-negative species are present in the genomes of *C. tetani* and other clostridia. SecA, SecD, SecE, SecF, and SecY could be identified; SecB, a translocation-specific chaperone and SecG, usually part of the SecYEG integral membrane complex seems to be missing in the clostridial group, indicating the existence of a modified Sec-system. Alternatively, other proteins involved in protein secretion may take over the SecBG functions. A homologue of YajC (CTG2208) is present, which can form an integral membrane heterotrimeric complex together with SecD and SecF (39). A protein in the immediate vicinity of the putative *ori* (CTG0096) exhibits some homology to YidC of *E. coli*, which has recently been shown to interact with the SecDFyajC complex (39). Furthermore, CTG166 and CTG458, present exclusively in clostridia, show homologies to the C-terminal domain of SecA.

In addition, *C. tetani* possesses genes of a signal-recognition particle (SRP)-like pathway, a protein translocation system similar to mammalian SRP. This system is thought to play an essential role in protein translocation and membrane protein insertion (40, 41). It consists of the SRP54 homologue Ffh (CTG1249) and a small cytoplasmic RNA (upstream of CTG0075), together forming the SRP, which binds to the signal sequence when it emerges from the ribosomes. FtsY (CTG1246) is part of the SRP receptor, which is thought to mediate the reception and insertion of a subset of proteins at the cytoplasmic membrane. Whether Ffh and FtsY are part of an independent protein secretion pathway or whether these proteins take over functions of the Sec-dependent pathway, as proposed previously, has to be elucidated (42, 43). In *C. tetani* as well as in all other sequenced clostridia, there are genes clustered with FtsY and Ffh, which are putatively involved in the SRP-dependent secretory mechanism: a SRP-associated protein (CTG1247), a RNA-binding protein (CTG1248), a 16S rRNA processing protein (CTG1250), a tRNA (Guanine-N1)-methyltransferase (CTG1251), two ribosomal proteins (L19P and S16P), a signal peptidase (CTG1253), a GTP-binding protein (CTG1254), and a ribonuclease (CTG1255).

The Sec-independent TAT (twin arginine translocation) system is absent from the *C. tetani* genome and all clostridial genomes known so far. Instead, a cluster of 10 genes (CTC1596-CTC1605) could be identified with similarities to subunits of type II secretion and type IV pilus biogenesis systems. Two ORFs (CTC1604, CTC1603) show strong similarities to the subunits EpsE and F of the type II secretion system in *V. cholerae*, which is responsible for the export of the cholera toxin (44). Three additional ORFs within the *C. tetani* cluster show some homology to EpsG, -H, and -I. The other proteins are hypothetical, occurring only in *C. tetani, C. acetobutylicum*, and *C. botulinum*.

Conclusion

It can be concluded that *C. tetani* is an invasive anaerobic pathogen possessing a wide arsenal of virulence factors; this arsenal is exacerbated by the ability to produce the tetanus toxin and possibly by the presence of various SLPs of yet unknown function. The performance of an extensive sodium ion bioenergetics might be an additional factor for the successful invasion of infected tissue. The genome sequence of *C. tetani* now provides the information on regulatory systems so that the signal transduction pathway leading to tetanus toxin expression can be studied.

Discussions with Wolfgang Buckel (Philipps-University, Marburg, Germany) on the biological function of the NADH:ubiquinone oxidoreductase/*Rhodobacter*-specific nitrogen fixation system are gratefully acknowledged. This work was supported by a grant of the Niedersächsisches Ministerium für Wissenschaft und Kultur to the Göttingen Genomics Laboratory and by funds from the Akademie der Wissenschaften zu Göttingen.

1. Tizzoni, G. & Cattani, G. (1890) *Zentralbl. Bakt.* **8,** 69–73.
2. Montecucco, C. & Schiavo, G. (1993) *Trends Biochem. Sci.* **18,** 324–327.
3. Swaminathan, S. & Eswaramoorthy, S. (2000) *Nat. Struct. Biol.* **7,** 693–699.
4. Arnon, S. S. (1997) in *The Clostridia: Molecular Biology and Pathogenesis*, eds. Rood, J. I., McClane, B. A., Songer, J. G. & Titball, R. W. (Academic, San Diego), pp. 95–115.
5. Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R. & Montecucco, C. (1992) *Nature* **359,** 832–835.
6. Gasse, F. (1995) *Int. J. Gynaecol. Obstet.* **50,** S67–S72.
7. Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hudel, M., Habermann, E. & Niemann, H. (1986) *EMBO J.* **5,** 2495–2502.
8. Fairweather, N. F. & Lyness, V. A. (1986) *Nucleic Acids Res.* **14,** 7809–7812.
9. Marvaud, J.-C., Eisel, U., Binz, T., Niemann, H. & Popoff, M. R. (1998) *Infect. Immun.* **66,** 5698–5702.
10. Finn, C. W., Silver, R. P., Habig, W. H. & Hardegree, M. C. (1984) *Science* **224,** 881–884.
11. Marvaud, J.-C., Raffestin, S., Gibert, M. & Popoff, M. R. (2000) *Biol. Cell* **92,** 455–457.
12. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998) *Genome Res.* **8,** 175–185.
13. Staden, R., Beal, K. F. & Bonfield, J. K. (2000) *Methods Mol. Biol.* **132,** 115–130.
14. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
15. Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., Ogasawara, N., Hattori, M., Kuhara, S. & Hayashi, H. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 996–1001.
16. Nölling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., *et al.* (2001) *J. Bacteriol.* **183,** 4823–4838.
17. Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Umayam, L., *et al.* (2000) *Nature* **406,** 477–483.
18. Parkhill, J., Wren, B. W., Thomson, N. R., Titball, R. W., Holden, M. T., Prentice, M. B., Sebaihia, M., James, K. D., Churcher, C., Mungall, K. L., *et al.* (2001) *Nature* **413,** 523–527.
19. Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., Henne, A., Wiezer, A., Bäumer, S., Jacobi, C., *et al.* (2002) *J. Mol. Microbiol. Biotechnol.* **4,** 453–461.
20. Harrington, D. J. (1996) *Infect. Immun.* **64,** 1885–1891.
21. Matsushita, O., Jung, C. M., Katayama, S., Minami, J., Takahashi, Y. & Okabe, A. (1999) *J. Bacteriol.* **181,** 923–933.
22. Matsushita, O., Yoshihara, K., Katayama, S. I., Minami, J. & Okabe, A. (1994) *J. Bacteriol.* **176,** 149–156.
23. Mani, N. & Dupuy, B. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 5844–5849.
24. Garnier, T. & Cole, S. T. (1988) *Mol. Microbiol.* **2,** 607–614.
25. Garnier, T. & Cole, S. T. (1988) *Plasmid* **19,** 151–160.
26. Rood, J. I. (1998) *Annu. Rev. Microbiol.* **52,** 333–360.
27. Waligora, A. J., Hennequin, C., Mullany, P., Bourlioux, P., Collignon, A. & Karjalainen, T. (2001) *Infect. Immun.* **69,** 2144–2153.
28. Calabi, E. & Fairweather, N. F. (2002) *J. Bacteriol.* **184,** 3886–3897.
29. Mengaud, J., Ohayon, H., Gounon, P., Mege, R.-M. & Cossart, P. (1996) *Cell* **84,** 923–932.
30. Kofoid, E., Rappleye, C., Stojiljkovic, I. & Roth, J. (1999) *J. Bacteriol.* **181,** 5317–5329.
31. Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R. & Saier, M. H. (2000) *J. Mol. Biol.* **301,** 75–100.
32. Häse, C. C., Fedorova, N. D., Galperin, M. Y. & Dibrov, P. A. (2001) *Microbiol. Mol. Biol. Rev.* **65,** 353–370.
33. Morita, Y., Kataoka, A., Shiota, S., Mizushima, T. & Tsuchiya, T. (2000) *J. Bacteriol.* **182,** 6694–6697.
34. Ito, M., Guffanti, A. A. & Krulwich, T. A. (2001) *FEBS Lett.* **496,** 117–120.
35. Kakinuma, Y. & Igarashi, K. (1995) *FEBS Lett.* **359,** 255–258.
36. Schmehl, M., Jahn, A., Meyer zu Vilsendorf, A., Hennecke, S., Masepohl, B., Schuppler, M., Marxer, M., Oelze, J. & Klipp, W. (1993) *Mol. Gen. Genet.* **241,** 602–615.
37. Mukherjee, K., Karlsson, S., Burman, L. G. & Akerlund, T. (2002) *Microbiology* **148,** 2245–2253.
38. Driessen, A. J., Fekkes, P. & van der Wolk, J. P. (1998) *Curr. Opin. Microbiol.* **1,** 216–222.
39. Nouwen, N. & Driessen, A. J. (2002) *Mol. Microbiol.* **44,** 1397–1405.
40. Lutcke, H. (1995) *Eur. J. Biochem.* **228,** 531–550.
41. Zwieb, C. & Samuelsson, T. (2000) *Nucleic Acids Res.* **28,** 171–172.
42. van Wely, K. H., Swaving, J., Freudl, R. & Driessen, A. J. (2001) *FEMS Microbiol. Rev.* **25,** 437–454.
43. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., *et al.* (1997) *Nature* **390,** 249–256.
44. Sandkvist, M., Michel, L. O., Hough, L. P., Morales, V. M., Bagdasarian, M., Koomey, M., DiRita, V. J. & Bagdasarian, M. (1997) *J. Bacteriol.* **179,** 6994–7003.