molecular oral microbiology

molecular oral microbiology

Conservation and revised annotation of the *Treponema denticola prcB-prcA-prtP* locus encoding the dentilisin (CTLP) protease complex

M.P. Goetting-Minesky, V. Godovikova, J.J. Li, S. Seshadrinathan, J.C. Timm, S.S. Kamodia and J.C. Fenno

Department of Biologic and Materials Sciences, School of Dentistry, University of Michigan, Ann Arbor, MI, USA

Correspondence: J. Christopher Fenno, Department of Biologic and Materials Sciences, School of Dentistry, University of Michigan, Ann Arbor, MI 48109-1078, USA Tel.: +1 734 763 3331; fax: +1 734 647 2110; E-mail: fenno@umich.edu

Keywords: gene annotation; periodontal disease; protease; Treponema denticola
Accepted 1 November 2012
DOI: 10.1111/omi.12013

SUMMARY

Interstrain differences in antigenic surface proteins may reflect immunological pressure or differences in receptor specificity of the antigen. *Treponema denticola* exhibits considerable interstrain variability in its major surface protein (Msp), but no studies have addressed this issue in dentilisin (CTLP), a surface protease complex that has a significant role in *T. denticola*–host interactions in periodontal disease. Furthermore, the genome annotation of the *prcB-prcA-prtP* operon encoding dentilisin contains apparent errors and lacks a deduced PrtP amino acid sequence. To address these issues we analysed the protease operon from diverse *T. denticola* strains, as well as clones of the ATCC 35405 Type strain from which the genome sequence and original GenBank *prtP* sequence were derived. 6xHis-tagging of the PrtP C-terminus in ATCC 35405 demonstrated absence of the 'authentic frameshift' in PrtP reported in the genome databases. We propose that *T. denticola* genome annotations be updated to reflect this new information. PrcB and the PrtP N-terminal region that includes the catalytic domain were highly conserved in common laboratory strains and clinical isolates of *T. denticola*. Dentilisin proteolytic activity varied considerably between strains. Antibodies against PrcB, PrcA and PrtP from the type strain recognized these proteins in most *T. denticola* strains. PrtP varied up to 20% over the C-terminal 270 residues between strains. The PrtP C-terminal eight-residues (DWFYVEYP) was present in all strains, with two strains containing an additional Y-residue preceding the stop codon. Such conserved PrtP domains may be required for interactions with PrcA and PrcB, or for substrate interactions.

INTRODUCTION

The *Treponema denticola* outer membrane lipoprotein–protease complex (dentilisin) is comprised of the polypeptide products of the monocistronic *prcB-prcA-prtP* operon (Ishihara *et al.*, 1996; Bian *et al.*, 2005; Godovikova *et al.*, 2010), which is both unique to and conserved in several oral *Treponema* species (Correia *et al.*, 2003). Dentilisin contributes to periodontal disease by degrading components of serum and extracellular matrix (Mäkinen *et al.*, 1995; Uitto *et al.*, 1995; McDowell *et al.*, 2009), by disrupting intercellular junctions (Ellen *et al.*, 2000; Chi *et al.*, 2003) and by dysregulation of tissue homeostasis control (Miao *et al.*, 2011). Interstrain differences in antigenic surface proteins may reflect responses to immunological pressure or differences in host cell or tissue receptor specificity of the antigen. While *T. denticola* exhibits

considerable interstrain variability in the Msp major surface protein (Fenno *et al.*, 1997), no studies have addressed interstrain differences in the *T. denticola* dentilisin complex (CTLP), which is antigenically prominent (Capone *et al.*, 2005) and plays a role in interactions with host tissue (reviewed in Fenno 2012). We aimed to characterize sequence variability in the locus encoding the protease complex and relate this to the levels of protease activity in diverse strains.

One of the significant challenges in the era of microbial genomics is confirmation of annotated genome information by demonstration of function and activity of products of annotated genes. Automated genome analysis and annotation provides unprecedented amounts of information at the genomic level but prediction of gene function, especially in the case of 'hypothetical proteins', remains problematic (Sivashankari & Shanmughavel, 2006). Experimental confirmation of predicted gene structure and function is necessarily even more of an issue because of the rapidly expanding amount of unconfirmed genome annotation in the databases. Genomes are annotated based on accepted best practices and the best information available at the time. It is the responsibility of the research community to identify annotation errors to prevent their propagation as future annotations are constructed.

The operon encoding the dentilisin complex (CTLP) is encoded by TDE0760–TDE0762 in the annotated *T. denticola* genome (Seshadri *et al.*, 2004). TDE0760, which we recently demonstrated to encode the acylated, 22-kDa outer membrane protein PrcB (Godovikova *et al.*, 2010), is annotated as a 17.6-kDa conserved hypothetical non-acylated periplasmic protein with limited homology to a very small Pfam group (PF01833) characterized as containing a domain with an immunoglobulin-like fold found in some cell surface receptor and intracellular transcription factors. TDE0761, which we have shown to encode the acylated outer membrane protein PrcA (Lee *et al.*, 2002; Godovikova *et al.*, 2011), is annotated differently by the Center for Microbial Resources at the J. Craig Venter Institute (CMR; cmr.jcvi.org) and the Oral Pathogen Sequence Database at Los Alamos National Laboratory (Oralgen; oralgen.lanl.gov). PrcA is described as a potential member of Pfam PF04773 (which includes the FecR sensor protein involved in dicitrate transport across the inner membrane; Oralgen annotation) and Pfam PF00041 (due to presence of a potential fibronectin Type III domain; CMR annotation). TDE0762 is annotated in both databases as a serine protease (PrtP, dentilisin) containing an 'authentic frameshift'. Although the graphic display of TDE0762 on the CMR site identifies a Type II signal peptide, no predicted PrtP amino acid sequence is shown and PrtP can be retrieved neither from protein databases by BLAST search with the PrtP amino acid sequence (Altschul *et al.*, 1990) nor by searching the *T. denticola* genome using an algorithm designed specifically to identify lipoproteins in spirochete genomes (Setubal *et al.*, 2006). We are not the first to note that significant annotation errors plague the genome databases (Brenner, 1999; Perrodou *et al.*, 2006). We believe it is particularly appropriate to address the issue of PrtP annotation because the dentilisin protease complex is a significant virulence determinant of *T. denticola* pathogenesis in periodontal disease.

Herein we provide experimental data demonstrating the identity and amino acid sequence of PrtP, including showing the absence of the putative 'authentic frameshift' that has resulted in exclusion of this significant microbial virulence determinant from genome-based databases. We then summarize our experimental results showing function and behavior of PrcB, PrcA and PrtP in contrast to the limited and incorrect information available in genomic databases. Furthermore, we characterize conservation, variability and expression of the *prcB-prcA-prtP* locus in *T. denticola*, demonstrating that this locus unique to a particular group of mammalian host-associated spirochetes encodes a highly conserved protease activity.

METHODS

Bacterial strains and growth conditions

Treponema denticola strains (Table 1), were grown in New Oral Spirochete (NOS) broth medium or NOS semi-solid medium under anaerobic conditions as previously described (Haapasalo *et al.*, 1991; Chan *et al.*, 1997), with erythromycin (40 μg ml^{-1}) added as appropriate. Cultures were examined by dark-field microscopy for purity and typical strain morphology.

Escherichia coli JM109 (Yanisch-Perron *et al.*, 1985) and *E. coli* Rosetta(DE3)/pLysS (Novagen Inc., Madison, WI) were used as hosts for cloning and

Table 1 *Treponema denticola* strains and plasmids used in this study

Strain	Features	Source (references)
35405	Type strain	ATCC (Chan *et al.*, 1993)
K1	*prtP* mutant	H. Kuramitsu (Ishihara *et al.*, 1998)
35404		ATCC (Chan *et al.*, 1993)
33520		ATCC (Jacob *et al.*, 1979)
33521		ATCC (Jacob *et al.*, 1979)
ASLM	Clinical isolate	W. Loesche (Ohta *et al.*, 1986)
SP82	Clinical isolate	W. Loesche (Salvador *et al.*, 1987)
OTK	Clinical isolate	R. Johnson (Fenno *et al.*, 1997)
CF417	*prcB*-6xHis	(Godovikova *et al.*, 2010)
CF646	*prtP*-6xHis	This study

ATCC, American Type Culture Collection, Rockville, MD.

Table 2 Oligonucleotide primers used in this study

Primer	Sequence[1]	Notes[2]
CX816	CGGGACTTCCATGGC AACCCCATTTG	mid-*prtP* w/*Nco*I [F]
CX822	GCTCGAGCGGATATT CTACATAGAACCAATC	*prtP* 3′ end w/*Xho*I [R]
CX834	GATAACTGCAGATCA AATGAAGGCTCTGC	mid-*prtP* w/*Pst*I P1 [F]
CX847	cttcttaattacTCAGTGGT GGTGGTGGT	*prtP6xHis-ermB* P2 [R]
CX848	CACCACTGAgtaattaag aaggagtgattac	*prtP6xHis-erm*B [F]
CX849	ATATTAACGcctaataatt tatctacattcc	*ermB*-TERM*prtP* [R]
CX850	gataaattattaggCGTTAA TATGGGTAATTAGG	*ermB*-TERM*prtP* [F]
CX819	CTTCAATGGGAAGAA GGAAG	past *prtP*–>*tuf* [R]
CX859	GATCAAATGAAGGCT CTGC	*mid*-prtP [F]
CX636	GCGCCGCCCTAATTAC	past *prtP* [R]
CX637	GTTTGGTACGGTCGAAAC	*prtP* 1188 [F]

[1]Engineered restriction enzyme sites underlined; Upper case: *prtP* homology; Lower case: *ermB* homology.
[2]Target; orientation with respect to *prtP*: forward [F]; reverse [R].

expression of recombinant proteins, respectively. The *E. coli* was grown on LB agar or broth medium with ampicillin (50 μg ml^{-1}), kanamycin (30 μg ml^{-1}) and chloramphenicol (34 μg ml^{-1}) as appropriate. Plasmid vector pSTBlue-1 (Novagen) was used for direct cloning of polymerase chain reaction (PCR) products, and 6xHis-tagged constructs were made in pET28b (Novagen Inc.).

Construction of plasmids for expression and mutagenesis studies

DNA encoding the C-terminal region of *prtP* was amplified from *T. denticola* genomic DNA using primers CX616 and CX822 (Table 2), and the resulting PCR product carrying 5′ *Nco*I and 3′ *Xho*I engineered restriction sites was cloned in pET28b (Novagen) such that in the resulting plasmid (pCF617), a partial *prtP* open reading frame including a C-terminal 6xHistidine tag (6xHis) was expressed from the vector-encoded T7 promoter. To construct a DNA molecule capable of transferring this tagged *prtP* to *T. denticola* we employed a variation on overlap extension (OE) PCR methodology described by Shevchuk *et al.* (2004). An OE-PCR product was generated by combining three PCR products: A: *prtP6xHis*, B: *ermB*, encompassing its Shine–Dalgarno sequence and coding region and C: downstream of *prtP* through the 5′ end of TDE0765. Primers used to generate these fragments (listed in Table 2) contain engineered overlapping 10–12 base pairs complementary to the adjacent PCR product. In the first step, a 100-μl PCR containing templates A, B and C in a molar ratio of 5 : 1 : 5 was carried out for 10 cycles in the absence of oligonucleotide primers. One microliter of this product was used as template for a 35-cycle PCR using primers CX859 and CX819 complementary to the 5′ end of fragment A and the 3′ end of fragment C, respectively. The resulting PCR product was purified and cloned in pSTBlue-1 (Novagen) yielding pCF640, which carries *ermB* inserted between the 3′ end of *prtP*-6xHis tag and DNA downstream of *prtP* including TDE0763, TDE0764 and the 5′ end of TDE0765.

Allelic replacement mutagenesis

Defined isogenic mutants were constructed as described previously (Li *et al.*, 1996; Fenno *et al.*, 1998b), by electroporation of *T. denticola* with linear DNA fragments consisting of the selectable *erm* cassette cloned between DNA fragments flanking the target gene. Plasmid pCF640 was digested with *Eco*RI before electroporation of *T. denticola* to separate the vector and insert fragments. Mutants were selected for resistance to erythromycin in NOS/GN agar (Chan *et al.*, 1997). Mutations were verified by PCR analysis and by DNA sequencing of the target region in genomic DNA of the mutants.

Preparation of *Treponema denticola* extracts

Treponema denticola cultures were harvested by centrifugation at 10,000 ***g*** (10 min, 4°C), washed once in phosphate-buffered saline (PBS) and suspended in PBS at an optical density of 0.2 at 600 nm. Whole cell lysates were prepared by sonication before suspension in sample buffer. For some experiments, Triton X-114 extraction and phase partitioning of outer membrane proteins were performed as described for *T. pallidum* (Cunningham *et al.*, 1988) with slight modifications (Fenno *et al.*, 1998a; Miao *et al.*, 2011). The final detergent phase extract was then precipitated in acetone and resuspended in electrophoresis sample buffer.

DNA sequence analysis

Templates for DNA sequencing, including plasmid DNA and PCR products generated from *T. denticola* genomic DNA, were sequenced at the University of Michigan DNA Sequencing Core Facility. Sequences of both DNA strands of regions of interest were obtained and analysed using Lasergene software (DNAStar Inc., Madison, WI). DNA sequences of the *prcB* open reading frame in *T. denticola* strains ATCC 33520 and OTK were previously assigned GenBank accession numbers FJ555200 and FJ555201, respectively (Godovikova *et al.*, 2010). DNA sequences encoding *prcB-prcA-prtP* of strains 33520, 33521, 35404, ASLM, SP82 and OTK have been assigned GenBank accession numbers JX984656–JX98466, respectively. The *prcB-prcA-prtP* locus in 35405 was determined to match that in the published genome sequence (GenBank AE017226.1 and data not shown). Annotation of the open reading frames was performed using the FGENESB algorithm trained to the *T. denticola* genome sequence (Softberry Inc., Mt Kisco, NY). A predicted σ-70 class promoter upstream of *prcB* was identified using BPROM software (Softberry Inc.). Identification of the potential signal peptidase cleavage sites was done using PSORT (Nakai & Kanehisa, 1991; Emanuelsson *et al.*, 2007), LipoP (Juncker *et al.*, 2003) and SpLip (Setubal *et al.*, 2006).

Protein electrophoresis and immunoblotting

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and Western immunoblotting were carried out as described previously (Fenno *et al.*, 1996). The *T. denticola* cultures were harvested in the presence of 2 mM phenylmethylsulfonyl fluoride (PMSF) by centrifugation at 10,000 ***g*** (10 min, 4°C), washed once in PBS and suspended in PBS at an optical density of 0.2 at 600 nm, then subjected to Triton X-114 extraction (Cunningham *et al.*, 1988) without partitioning to aqueous and detergent phases. Equal volumes of samples were prepared in standard SDS–PAGE sample buffer containing β-mercaptoethanol and 2 mM PMSF. Before electrophoresis, samples were either heated at 100°C for 5 min or held on ice. Proteins blotted to nitrocellulose membranes were detected with rabbit polyclonal antibodies raised against recombinant *T. denticola* proteins (Fenno *et al.*, 1996; Bian *et al.*, 2005; Godovikova *et al.*, 2010, 2011), followed by horseradish peroxidase (HRP) -conjugated goat anti-rabbit IgG (Thermo Scientific, Rockford, IL). 6xHis-tagged proteins were detected using HisProbe HRP reagent (Thermo Scientific). Protein bands of interest were visualized using SuperSignal West Pico chemiluminescent substrate (Thermo Scientific).

Protease activity assay

PrtP-dependent hydrolysis of the chromogenic substrate succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanine-*p*-nitroanilide (SAAPFNA) in 5-day *T. denticola* cultures was assayed by change in absorbance at 405 nm, as described previously (Uitto *et al.*, 1988; Bian *et al.*, 2005). SAAPFNA activity is expressed as a percentage of the SAAPFNA activity of strain 35405 defined as 100%. Error bars represent the standard error over four independent experiments including triplicate samples.

RESULTS AND DISCUSSION

Treponema denticola TDE0762 (*prtP*) does not contain an 'authentic frameshift'

The DNA sequence of *T. denticola prtP* was first reported in 1996 and assigned GenBank accession D83264 (Ishihara *et al.*, 1996). The *T. denticola* ATCC 35405 genome was published in 2004 (Seshadri *et al.*, 2004). In addition to the complete genome sequence (GenBank accession NC_002967), two

major websites maintain the annotated genome sequence: the Center for Microbial Resources at the J. Craig Venter Institute (cmr.jcvi.org) and the Oral Pathogen Sequence Database at Los Alamos National Laboratory (Oralgen; oralgen.lanl.gov). Whereas the DNA sequence of the genomic region including TDE0762 (*prtP*) has remained unchanged in the databases, the reported TDE0762 DNA sequence at these sites has varied over the last 6 years, and TDE0762 (*prtP)* is annotated as a 'pseudogene' (GenBank) or as containing an 'authentic frameshift' (CMR, Oralgen). In standard genome annotation methodology, open reading frames are detected using GLIMMER (Salzberg *et al.*, 1998). Detection of a mismatch in any particular open reading frame between the genomic sequence and a previously deposited GenBank sequence results in annotation of the gene as containing an 'authentic frameshift' (Nelson *et al.*, 2003). One consequence of this convention is that the amino acid sequence encoded by such a gene is not reported in the genome annotation and is generally absent from protein databases.

While none of GenBank, CMR and Oralgen currently post the deduced amino acid sequence (CDS) encoded by TDE0762, several have been posted at various times over the past few years. Except for the full-length 766-residue PrtP briefly posted on the TIGR (now CMR) site in 2005, the others appear to have been truncated to approximate the length of *prtP* found in the first submitted *prtP* sequence (GenBank D83264), which reports *prtP* as encoding a 722-residue protein. Our DNA sequencing results in ATCC 35405 confirmed the reported genomic DNA sequence. Both our sequence and the genome databases show three differences compared with GenBank D83264: two to three-base substitutions (1414–1416:TAT vs. ATA; 1494–1497: GAA vs. CGA) and a single additional 'G' in the D83264 sequence (position 2109). At the protein level, this results in I^{472} (D83264) vs. V^{472}, E^{499} (D83264) vs. R^{499} and a frameshift in D83264 resulting in mismatches beyond residue 703 of the protein sequences deduced from the databases (Fig. 1). It must be noted that the *T. denticola* genome sequence does not contain an in-frame stop codon at the point identified as the end of the coding sequence by both CMR and Oralgen, but that both databases arbitrarily truncate the *prtP* coding region after codon 721 (Oralgen) or 722 (CMR). However, both the genome sequence and our sequencing results suggest that, rather than the 722-residue PrtP reported in D83264 and implied in the genome databases, PrtP is a 766-residue protein whose sequence beyond residue 703 differs from that reported in D83264.

To ensure that the mismatch between the original GenBank submission and the genomic sequence was not the result of a mutation acquired during subculture of ATCC 35405 in separate laboratories, we next determined the DNA sequence of the 3′ region of *prtP* in *T. denticola* K1, an isogenic mutant of *T. denticola* ATCC 35405 that carries an antibiotic resistance marker inserted in the 5′ region of *prtP* (Ishihara *et al.*, 1998). The K1 strain is derived from the ATCC 35405 clone that was the source of the D83264 *prtP* sequence. The DNA sequences of *prtP* from base 1290 through the end of the predicted *prtP* open reading frame shown in our 35405 clone and in the genomic sequence were identical in *T. denticola* K1 (data not shown). This strongly suggests that the *prtP* sequence deposited as GenBank D83264 contains sequencing errors, resulting in prediction of premature C-terminal truncation of the PrtP open reading frame.

Finally, to provide experimental evidence of the lack of an 'authentic frameshift' in *prtP*, we constructed isogenic *T. denticola* mutant strain CF646 carrying a C-terminal 6xHis tag immediately before the *prtP* stop codon at base 2199 (following deduced amino acid codon 766 in the TDE0762 open reading frame). If native PrtP is truncated at residue 722, as shown in the original GenBank record and as suggested by current genome databases, then PrtP in the CF646 mutant would not include the C-terminal 6xHis tag. As shown in Fig. 2 (left panel) the presence of 6xHis tagged full-length PrtP in CF646 clearly demonstrates that TDE0762 encodes a 766-residue PrtP protein and that the reported 'authentic frameshift' in the genome databases is most likely the result of a sequencing error in the original GenBank entry. We believe that these data support revised annotation of the *T. denticola* dentilisin locus in curated genome databases. We have submitted requests to Oragen, CMR and the Human Oral Microbiome Database (homd.org) for updates reflecting this new information.

```
                 .                   .
681 KGGVFQGGNYALKIDRTPLNSNGCKYRRYGKKSQQCFNERQP 722
    |||||||||||||||||||||||       +
681 KGGVFQGGNYALKIDRTPLNSNGVNIDDTARSPNNASTNDSH 722
    |||||||||||||||||||||||||||||||||||||||||||
681 KGGVFQGGNYALKIDRTPLNSNGVNIDDTARSPNNASTNDSHEDDDTPAFAKVKGNAWGQTKGCNLVAHPIGSNNVDLDWFYVEYP 766
                 .                   .                   .                   .
```

Figure 1 Comparison of the *Treponema denticola* ATCC 35405 deduced PrtP amino acid sequences from residue 681. The top line is PrtP as reported in GenBank D83264. The middle line is the amino acid sequence deduced from TDE0762 as reported in the CMR genome database. Deduced translation of the Oralgen TDE0762 results in an amino acid sequence identical to that of CMR, except that the C-terminal histidine residue is absent. The bottom line is from the *T. denticola* genome sequence including DNA 3′ to identified TDE0762 until an in-frame stop codon. This sequence is confirmed in the present study. The difference between the sequences shown results from inclusion of an additional guanosine at nucleic acid residue 2109 in D83264, as discussed in the text.

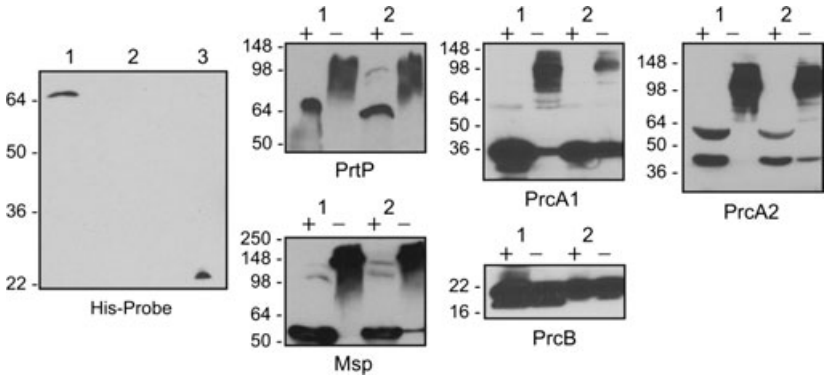

Figure 2 Outer membrane protein expression in *Treponema denticola* 35405 and CF646. The panels consist of blots probed with HisProbe reagent (left panel) or with antibodies to the *T. denticola* proteins or polypeptides indicated. Molecular standards in kDa are shown to the left of each blot. Lane 1, *T. denticola* CF646; lane 2, *T. denticola* 35405; lane 3, *T. denticola* CF417. In blots probed with antibodies, samples were either heated (+) or not heated (–) before electrophoresis.

Expression of *Treponema denticola* outer membrane complex proteins in CF646

Molecular 'tags' can interfere with expression and function of some proteins. To determine whether inclusion of the 6xHis tag at the C-terminus of PrtP affected expression of the outer membrane complexes in *T. denticola*, we assayed the expression of individual protease complex components and the major surface protein Msp in parent and mutant strains. Samples were prepared for electrophoresis both with and without heating to detect both denatured proteins and protein complexes. As shown in Fig. 2B, all proteins tested are expressed similarly in both 35405 and CF646, and there are no apparent differences in migration of outer membrane protein complexes between parent and mutant strains. Taken together, this suggests that the PrtP C-terminal 6xHis tag does not interfere with expression and assembly of the protease complex.

Conservation of the *prcB-prcA-prtP* locus in *Treponema denticola*

To characterize interstrain conservation of the protease locus, we sequenced the *prcB-prcA-prtP* region encoding the three lipoproteins comprising the dentilisin complex in well-established laboratory strains and in more recent isolates. As shown in Fig. 3, the deduced amino acid sequences of all three proteins show high levels of conservation, particularly in PrcB (89% identical at the amino acid level among the seven strains). PrcA is 77% identical among the seven strains, and much of the variability is in three regions (151–180, 328–362 and 512–545). In PrtP, the predicted catalytic triad (Asp^{203}, His^{258}, Ser^{447}) is conserved in all strains. Consistent with conservation of proteolytic function, most PrtP interstrain variations are in the C-terminal 250 residues. Several strains showed insertions or deletions of between one and eight residues compared with the Type strain ATCC

35405. Apparent 'hot spots' for this type of variation are in the regions of residues 235, 295, 562, 692, 724 and 754. Strains ASLM and SP82 also contained a C-terminal Y in addition to the conserved –DWFYVEYP C-terminus present in all strains. PrtP in 35405 is comprised of 766 amino acid residues; the PrtP amino acid content ranged from 759 residues in 33521 to 776 residues in ASLM.

PrcB, PrcA and PrtP in all strains contain Type II signal peptidase cleavage sites. Our previous studies showed that each protein segregates to the detergent phase of Triton X-114 extracts (Fenno *et al.*, 1998a; Lee *et al.*, 2002; Godovikova *et al.*, 2010, 2011). In the current genome annotation, only PrcA is annotated as a lipoprotein. PrcB is reported as a cytoplasmic protein whose N-terminus is at residue 38 of the PrcB protein as reported here and previously (Godovikova *et al.*, 2010). As noted earlier, the absence of annotation of PrtP protein is most likely the result of sequencing errors in previous studies (Ishihara *et al.*, 1996). From a functional standpoint the variability in the PrtP C-terminal region is suggestive of at least two possibilities. Variability in regions of a protein not essential for enzymatic function may result from selective immunological pressure, resulting in antigenic variation between strains. Alternatively, sequence variations may reflect functional differences in binding to host substrates or other protein interactions. The dentilisin complex (PrcA, PrcB and PrtP) is part of a very high molecular weight outer membrane complex that also includes the oligomeric Msp protein (Godovikova *et al.*, 2011). The gene encoding the major surface protein (Msp) in each strain examined here falls within one of the three previously identified *msp* types (Fenno *et al.*, 1997) as do those of more than 30 clinical isolates (Fenno *et al.*, 2001 and data not shown). Although the data are somewhat limited, phylogenetic trees for each protein do not reflect any consistency in strain-dependent relationships between PrtP, PrcA and PrcB sequences (data not shown), nor is there any discernable relationship between the patterns of interstrain PrtP homology (Fig. 3) and interstrain Msp homology (Fenno *et al.*, 1997). Hence, while it is clear that Msp and dentilisin components interact at the molecular level (Godovikova *et al.*, 2011), these interactions are likely between highly conserved domains of the relevant proteins. The protease complex is reported to specifically bind and degrade fibrinogen (Bamford *et al.*, 2007), and its ability to degrade a range of other proteins and bioactive peptides (Uitto *et al.*, 1988; Mäkinen *et al.*, 1995) is suggestive of specific binding of these or other substrates. We speculate that the C-terminal region of PrtP is involved in substrate binding and that interstrain differences in binding or proteolytic activity of the dentilisin complex is due to strain-dependent sequence differences in this region. To date, no studies have systematically examined the function of the PrtP C-terminus.

As part of the Human Oral Microbiome Project (Dewhirst *et al.*, 2010; Human Microbiome Project Consortium, 2012), the genomes of most of the strains examined here are being sequenced by the Broad Institute of Harvard and MIT (http://www.broadinstitute.org/). We have compared our DNA sequences with those of the provisional genome contigs and find them generally in close agreement. GenBank genome accession numbers of *T. denticola* strains in this study to date are as follows: NC_002967 (ATCC 35405), AGDU01000000 (ATCC 35404), AGDS01000000 (ATCC 33520), AGDT01000000 (ATCC 33521) and AGDR01000000 (ASLM), AGDY01000000 (OTK).

Protease complex protein expression

In addition to signal peptide cleavage and acylation, two of the three proteins of the protease complex undergo further post-translational processing. In *T. denticola* 35405, PrtP is cleaved at residue 159 (Ishihara *et al.*, 1996) to yield an acylated 16-kDa N-terminal polypeptide (PrtP-N) and 65-kDa mature PrtP (Godovikova *et al.*, 2011). PrcA undergoes PrtP-dependent cleavage (Lee *et al.*, 2002) to acylated PrcA1 (approximately 30 kDa) and non-acylated PrcA2 (approximately 40 kDa). To determine the polypeptide profile of the dentilisin complex in diverse *T. denticola* strains, we probed immunoblots with antibodies specific for individual components of the *T. denticola* 35405 dentilisin complex.

As shown in Fig. 4A, there is considerable variation both in the level of detection obtained and in the relative molecular weights of some of the dentilisin complex polypeptides. PrcB, which is highly conserved in all strains, was detected equally in all strains. This strongly suggests that the dentilisin complex proteins are expressed at similar levels in all strains and is consistent with the presence of an identical sigma-70

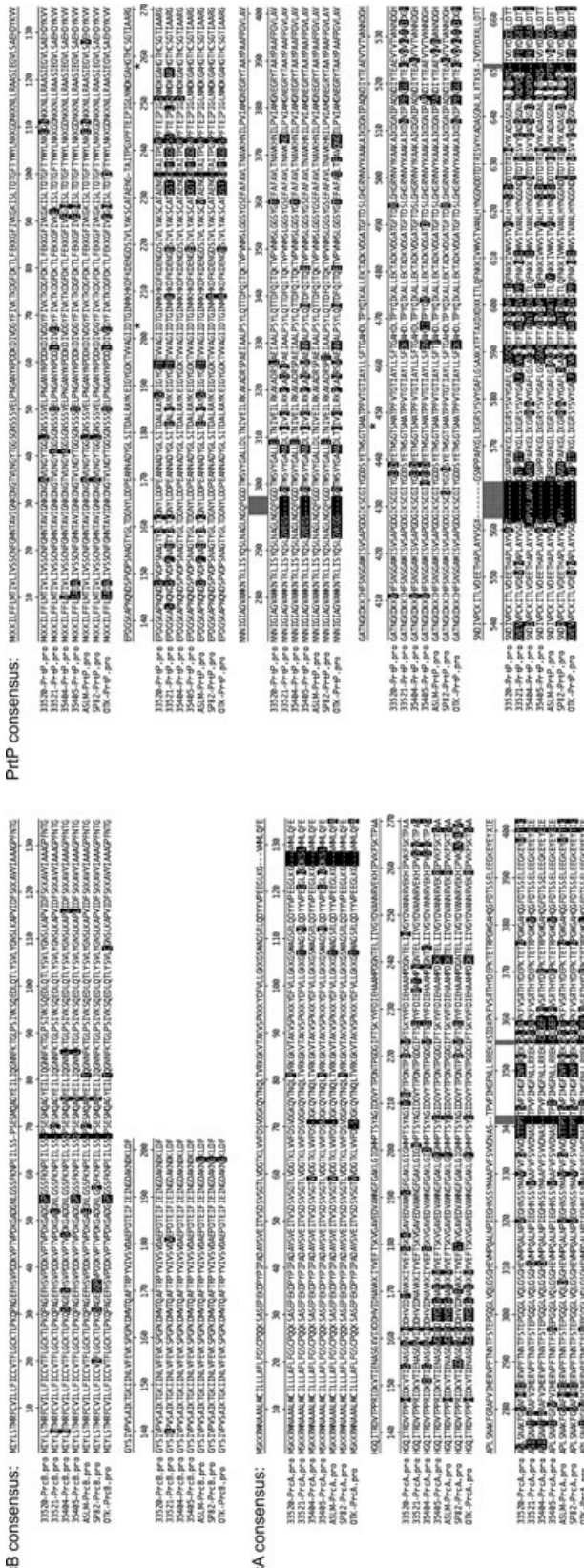

Figure 3 Conservation of PrcB, PrcA and PrtP in *Treponema denticola* strains. Deduced amino acid sequences were aligned using CLUSTAL-W (http://www.clustal.org). The consensus sequence of each protein, numbered as for 35405, is shown above the aligned sequences. Mismatches from the consensus are shaded black, and gaps in the consensus are shaded grey. Predicted PrtP active site residues (Asp^{203}, His^{258} and Ser^{447}) are indicated by asterisks below the consensus sequence.

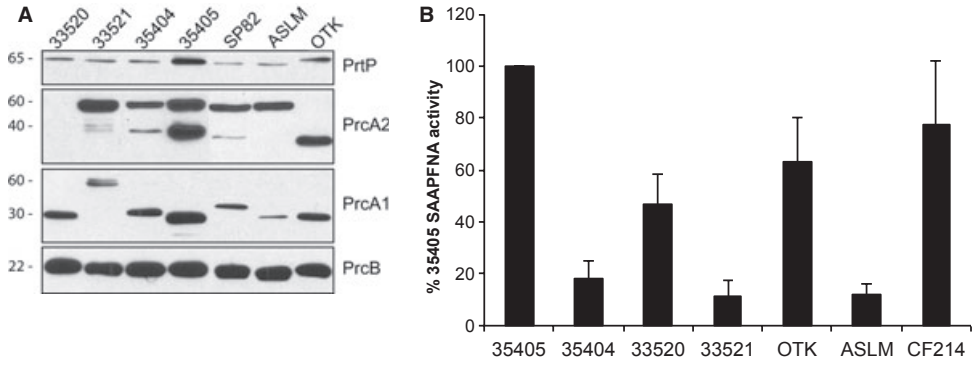

Figure 4 Protease expression and activity in *Treponema denticola* strains. (A) Immunoblots of *T. denticola* strains probed with antibodies specific for components of the dentilisin complex. Blots were probed with polyclonal antibodies specific for the proteins indicated to the right of each blot. All antibodies were raised against *T. denticola* 35405 proteins expressed in *Escherichia coli*. Strain names are listed above each lane. Approximate sizes in kDa of the reactive bands are indicated to the left of each blot. (B) PrtP protease activity in 5-day *T. denticola* cultures. Protease activity of individual strains, assayed by change in absorbance at 405 nm, is expressed as a percentage relative to the SAAPFNA activity of strain 35405. Error bars represent the standard error of four independent experiments including triplicate samples.

type predicted promoter sequence upstream of *prcB* in all seven strains (data not shown). PrtP was detected in all strains, although reactivity was relatively low in all but strain 35405, which was the source of the immunizing antigen. This is consistent with the interstrain variation evident in the PrtP C-terminal 250 residues. The N-terminal (PrcA1) and C-terminal (PrcA2) polypeptides resulting from PrtP-dependent processing of PrcA (Lee *et al.*, 2002) were detected by at least one of the two specific antibodies tested in all strains. Interestingly, while PrcA2 was not recognized in strain 33520, anti-PrcA1 recognized the expected 30-kDa band. In all strains except 33520, anti-PrcA2 recognized either or both PrcA2 (approximately 40-kDa) and the full-length PrcA molecule. Similarly, anti-PrcA1 recognized a PrcA1 of approximately 30-kDa in all strains except 33521, in which full-length PrcA was detected. Anti-PrtP-N, raised against the acylated 16-kDa N-terminal polypeptide released upon PrtP activation (Ishihara *et al.*, 1996; Godovikova *et al.*, 2011), reacted with all strains (data not shown).

Protease complex activity

To determine relative levels of dentilisin activity in *T. denticola* strains, we assayed cleavage of a chromogenic substrate (SAAPFNA) by *T. denticola* cultures in late logarithmic growth phase. We previously reported that transcription of *prcA* and *prtP* and protease activity increased as the growth phase progressed, likely in response to depletion of available nutrients (Bian *et al.*, 2005). As shown in Fig. 4B, all strains exhibited chymotrypsin-like activity, though activity varied considerably between strains with the Type strain (35405) and strains OTK and SP82 exhibiting the highest activity. Interestingly the three other ATCC strains and strain ASLM showed markedly lower activity. Proteolytic activity does not appear to be related specifically to overall homology with 35405 PrtP (Fig. 3). It is not apparent whether protease activity has any relationship with passage number or total length of time in culture. Compared with the four ATCC strains, strains OTK, ASLM and SP82 are relatively low-passage isolates. Whereas the putative promoter sequence upstream of *prcB* is identical in all seven strains, it is not known what specific environmental sensing and signaling mechanisms influence expression of the protease operon. Studies are in progress to characterize transcription of the *prcB-prcA-prtP* operon in these strains under a range of environmental conditions.

ACKNOWLEDGEMENTS

This work was supported by Public Health Service grant DE018221 (National Institute of Dental and Craniofacial Research), and by the Office of the Vice President for Research (University of Michigan). The authors thank Brian T. Foley of Los Alamos National Laboratories for helpful discussions.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Bamford, C.V., Fenno, J.C., Jenkinson, H.F. and Dymock, D. (2007) The chymotrypsin-like protease (CTLP) complex of *Treponema denticola* ATCC 35405 mediates fibrinogen adherence and degradation. *Infect Immun* **75**: 4364–4372.

Bian, X.-L., Wang, H.-T., Ning, Y., Lee, S.Y. and Fenno, J.C. (2005) Mutagenesis of a novel gene in the *prcA-prtP* protease locus affects expression of *Treponema denticola* membrane complexes. *Infect Immun* **73**: 1252–1255.

Brenner, S.E. (1999) Errors in genome annotation. *Trends Genet* **15**: 132–133.

Capone, R., Wang, H.-T., Ning, Y., Sweier, D.G., Lopatin, D.E. and Fenno, J.C. (2005) Recognition of *Treponema denticola* outer membrane proteins by human antibodies. *J Dent Res* **84**(Spec Iss A): 2841.

Chan, E.C., Siboo, R., Keng, T. *et al.* (1993) *Treponema denticola* (ex Brumpt 1925) sp. nov., nom. rev., and identification of new spirochete isolates from periodontal pockets. *Int J Syst Bacteriol* **43**: 196–203.

Chan, E.C.S., DeCiccio, A., McLaughlin, R., Klitorinos, A. and Siboo, R. (1997) An inexpensive solid medium for obtaining colony-forming units of oral spirochetes. *Oral Microbiol Immunol* **12**: 372–376.

Chi, B., Qi, M. and Kuramitsu, H.K. (2003) Role of dentilisin in *Treponema denticola* epithelial cell layer penetration. *Res Microbiol* **154**: 637–643.

Correia, F.F., Plummer, A.R., Ellen, R.P. *et al.* (2003) Two paralogous families of a two-gene subtilisin operon are widely distributed in oral treponemes. *J Bacteriol* **185**: 6860–6869.

Cunningham, T.M., Walker, E.M., Miller, J.N. and Lovett, M.A. (1988) Selective release of the *Treponema pallidum* outer membrane and associated polypeptides with Triton X-114. *J Bacteriol* **170**: 5789–5796.

Dewhirst, F.E., Chen, T., Izard, J. *et al.* (2010) The human oral microbiome. *J Bacteriol* **192**: 5002–5017.

Ellen, R.P., Ko, K.S., Lo, C.M., Grove, D.A. and Ishihara, K. (2000) Insertional inactivation of the *prtP* gene of *Treponema denticola* confirms dentilisin's disruption of epithelial junctions. *J Mol Microbiol Biotechnol* **2**: 581–586.

Emanuelsson, O., Brunak, S., von Heijne, G. and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**: 953–971.

Fenno, J.C. (2012) *Treponema denticola* interactions with host proteins. *J Oral Microbiol* **4**: 9929.

Fenno, J.C., Müller, K.-H. and McBride, B.C. (1996) Sequence analysis, expression and binding activity of recombinant major outer sheath protein (Msp) of *Treponema denticola*. *J Bacteriol* **178**: 2489–2497.

Fenno, J.C., Wong, G.W.K., Hannam, P.M., Müller, K.-H., Leung, W.K. and McBride, B.C. (1997) Conservation of *msp*, the gene encoding the major outer membrane protein of oral *Treponema* spp. *J Bacteriol* **179**: 1082–1089.

Fenno, J.C., Hannam, P.M., Leung, W.K., Tamura, M., Uitto, V.-J. and McBride, B.C. (1998a) Cytopathic effects of the major surface protein (Msp) and the chymotrypsinlike protease (CTLP) of *Treponema denticola*. *Infect Immun* **66**: 1869–1877.

Fenno, J.C., Wong, G.W.K., Hannam, P.M. and McBride, B.C. (1998b) Mutagenesis of outer membrane virulence determinants of the oral spirochete *Treponema denticola*. *FEMS Microbiol Lett* **163**: 209–215.

Fenno, J.C., Lee, S.Y., Bayer, C.H. and Ning, Y. (2001) *Treponema* msp's from subgingival plaque and cultivated strains. *J Dent Res* **80**: S53.

Godovikova, V., Wang, H.T., Goetting-Minesky, M.P. *et al.* (2010) *Treponema denticola* PrcB is required for expression and activity of the PrcA-PrtP (dentilisin) complex. *J Bacteriol* **192**: 3337–3344.

Godovikova, V., Goetting-Minesky, M.P. and Fenno, J.C. (2011) Composition and localization of *Treponema denticola* outer membrane complexes. *Infect Immun* **79**: 4868–4875.

Haapasalo, M., Singh, U., McBride, B.C. and Uitto, V.-J. (1991) Sulfhydryl-dependent attachment of *Treponema denticola* to laminin and other proteins. *Infect Immun* **59**: 4230–4237.

Human Microbiome Project Consortium (2012) A framework for human microbiome research. *Nature* **486**: 215–221.

Ishihara, K., Miura, T., Kuramitsu, H.K. and Okuda, K. (1996) Characterization of the *Treponema denticola prtP* gene encoding a prolyl-phenylalanine-specific protease (dentilisin). *Infect Immun* **64**: 5178–5186.

Ishihara, K., Kuramitsu, H.K., Miura, T. and Okuda, K. (1998) Dentilisin activity affects the organization of the outer sheath of *Treponema denticola*. *J Bacteriol* **180**: 3837–3844.

Jacob, E., Allen, A.L. and Nauman, R.K. (1979) Detection of oral anaerobic spirochetes in dental plaque by the indirect fluorescent-antibody technique. *J Clin Microbiol* **10**: 934–936.

Juncker, A.S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. and Krogh, A. (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**: 1652–1662.

Lee, S.Y., Bian, X.L., Wong, G.W., Hannam, P.M., McBride, B.C. and Fenno, J.C. (2002) Cleavage of *Treponema denticola* PrcA polypeptide to yield protease complex-associated proteins PrcA1 and PrcA2 is dependent on PrtP. *J Bacteriol* **184**: 3864–3870.

Li, H., Ruby, J., Charon, N. and Kuramitsu, H. (1996) Gene inactivation in the oral spirochete *Treponema denticola*: construction of an *flgE* mutant. *J Bacteriol* **178**: 3664–3667.

Mäkinen, P.L., Mäkinen, K.K. and Syed, S.A. (1995) Role of the chymotrypsin-like membrane-associated proteinase from *Treponema denticola* ATCC 35405 in inactivation of bioactive peptides. *Infect Immun* **63**: 3567–3575.

McDowell, J.V., Huang, B., Fenno, J.C. and Marconi, R.T. (2009) Analysis of a unique interaction between the complement regulatory protein factor H and the periodontal pathogen *Treponema denticola*. *Infect Immun* **77**: 1417–1425.

Miao, D., Fenno, J.C., Timm, J.C., Joo, N.E. and Kapila, Y.L. (2011) *Treponema denticola* chymotrypsin-like protease (dentilisin) induces MMP-2-dependent fibronectin fragmentation in periodontal ligament cells. *Infect Immun* **79**: 806–811.

Nakai, K. and Kanehisa, M. (1991) Expert system for predicting protein localization sites in gram- negative bacteria. *Proteins* **11**: 95–110.

Nelson, K.E., Fleischmann, R.D., DeBoy, R.T. *et al.* (2003) Complete genome sequence of the oral pathogenic pacterium *Porphyromonas gingivalis* strain W83. *J Bacteriol* **185**: 5591–5601.

Ohta, K., Mäkinen, K.K. and Loesche, W.J. (1986) Purification and characterization of an enzyme produced by *Treponema denticola* capable of hydrolyzing synthetic trypsin substrates. *Infect Immun* **53**: 213–220.

Perrodou, E., Deshayes, C., Muller, J. *et al.* (2006) ICDS database: interrupted CoDing sequences in prokaryotic genomes. *Nucleic Acids Res* **34**: D338–D343.

Salvador, S.L., Syed, S.A. and Loesche, W.J. (1987) Comparison of three dispersion procedures for quantitative recovery of cultivable species of subgingival spirochetes. *J Clin Microbiol* **25**: 2230–2232.

Salzberg, S.L., Delcher, A.L., Kasif, S. and White, O. (1998) Microbial gene identification using interpolated Markov models. *Nucleic Acids Res* **26**: 544–548.

Seshadri, R., Myers, G.S., Tettelin, H. *et al.* (2004) Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc Natl Acad Sci USA* **101**: 5646–5651.

Setubal, J.C., Reis, M., Matsunaga, J. and Haake, D.A. (2006) Lipoprotein computational prediction in spirochaetal genomes. *Microbiology* **152**: 113–121.

Shevchuk, N.A., Bryksin, A.V., Nusinovich, Y.A., Cabello, F.C., Sutherland, M. and Ladisch, S. (2004) Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Res* **32**: e19.

Sivashankari, S. and Shanmughavel, P. (2006) Functional annotation of hypothetical proteins – a review. *Bioinformation* **1**: 335–338.

Uitto, V.-J., Grenier, D., Chan, E.C. and McBride, B.C. (1988) Isolation of a chymotrypsinlike enzyme from *Treponema denticola*. *Infect Immun* **56**: 2717–2722.

Uitto, V.-J., Pan, Y.M., Leung, W.K. *et al.* (1995) Cytopathic effects of *Treponema denticola* chymotrypsin-like proteinase on migrating and stratified epithelial cells. *Infect Immun* **63**: 3401–3410.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103–119.