A carboxy-terminal processing protease gene is located immediately upstream of the invasion-associated locus from *Bartonella bacilliformis*

Samuel J. Mitchell and Michael F. Minnick

INTRODUCTION

*Bartonella bacilliformis* is a member of a group of unusual intracellular pathogens of growing medical importance. The bacteria within this group cause a severe haemolytic anaemia (*B. bacilliformis*; Reynafarje & Ramos, 1961), cat-scratch disease and bacillary angiomatosis (*B. henselae*; Regnery et al., 1992; Regnery & Tappero, 1995), and trench fever (*B. quintana*; Vinson & Fuller, 1961). Within the genus *Bartonella*, the ability to invade erythrocytes appears to be limited to *B. bacilliformis* (Reynafarje & Ramos, 1961) and *B. henselae* (Kordick & Breitschwerdt, 1995). We have previously identified a two-gene locus from *B. bacilliformis* (*ialAB*) which confers the ability to invade human erythrocytes upon non-invasive *Escherichia coli* (Mitchell & Minnick, 1995). Because so little is known about the molecular biology of this activity, we undertook the present study in an attempt to identify other genes which might interact with the *ialAB* locus. In addition, the possibility that *ialAB* might lie in the midst of a ‘pathogenicity island’ [Pais; recently reviewed by Falkow (1996)] led us to analyse chromosomal DNA flanking the *ialAB* genes. We have identified a homologue of an interesting class of carboxy-terminal processing proteases (*Ctpases*) immediately upstream of the *ialAB* locus.

As a whole, carboxy-terminal proteases are not well understood. To date, the best-characterized *CtpA* is that from the cyanobacterium *Synechocystis* sp. 6803 ([SYCCTPA; Shestakov et al., 1994]). This 34 kDa enzyme is responsible for cleavage of 16 C-terminal...
residues from its target protein, the D1 precursor polypeptide of photosystem II (PSII) (Nixon et al., 1992). The D1 precursor is translated as a polypeptide which binds to the PSII complex and is subsequently processed to its mature form by Ctpase activity (reviewed by Barber & Andersson, 1992). The complex is turned over very rapidly because of light-induced damage to D1 (Barber & Andersson, 1992), thus implying a vital role for CtpA in maintenance of PSII activity in response to environmental stress.

A protein with a conserved core sequence similar to that of the C-terminal proteases is the Tsp (tail-specific protease) of *E. coli* (Silber et al., 1992). Tsp was originally identified as a 76 kDa periplasmic protein involved in processing the penicillin-binding protein 3 (PBP 3) of *E. coli* (Hara et al., 1991). In the absence of Tsp, *E. coli* was extremely sensitive to heat and osmotic stress. More recently, Tsp has been shown to recognize and degrade several proteins with nonpolar C-termini, with strongest preference for alanine at the three C-terminal residues (Keiler et al., 1996). The Tsp of *E. coli* also functions as a universal degradative enzyme for aberrant proteins tagged by a C-terminal hydrophobic sequence (Keiler et al., 1996). Thus, Tsp appears to degrade aberrant proteins arising as a result of environmental stress or anomalous processing. A tsp homologue may also play a role in virulence and has been implicated in the survival of *Salmonella typhimurium* within macrophages (Bäumler et al., 1994).

In this study, we describe a * Bartonella* homologue of the carboxy-terminal processing proteases. Although the *ctpA* gene does not appear to directly affect the erythrocyte-invasion phenotype encoded by *ialAB*, the function of C-terminal proteases in response to stress in other organisms implies a significant role for the encoded protein. We also describe several intriguing features of the encoded protein: namely, a putative autocatalytic modification of its carboxy terminus, and a possible alternative translation start site which may give rise to two forms of the protein, perhaps targeted to different cellular locations.

**METHODS**

**Bacterial strains, plasmids and growth conditions.** Bacterial strains and plasmids are listed in Table 1. *B. bacilliformis* strains KC583 and KC584 (American Type Culture Collection) were grown on heart infusion agar (Difco) supplemented with 5% (w/v) defibrinated sheep erythrocytes and 5% (v/v) filter-sterilized sheep serum (Quad-Five) (HIBB plates) in a water-saturated atmosphere at 30 °C. *B. quintana*, *B. elizabethae* and *B. vinsonii* were obtained from the American Type Culture Collection in lyophilized form. *B. henselae* was obtained from the Centers for Disease Control & Prevention. *Br. abortus* was obtained as a frozen whole-cell suspension from Burroughs Wellcome.

*E. coli* strains were grown in Luria–Bertani (LB) medium at 37 °C with constant shaking. When required, ampicillin was added to the growth medium to a final concentration of 100 μg ml⁻¹. Haemolysin was assayed on HIBB plates supplemented with ampicillin as necessary. Collagenase activity was assayed in 5 ml nutrient gelatin tubes (3 g Bacto beef extract (Difco) 1⁻¹, 5 g peptone 1⁻¹, 120 g gelatin 1⁻¹, pH 6.8) with growth at 37 °C and gelling observed at 20 °C.

**Nucleic acid isolation and manipulations.** Chromosomal DNA was isolated from *B. bacilliformis* using a standard hexadecyl trimethyl ammonium bromide (CTAB) method (Ausbub et al., 1995). Individual DNA fragments were extracted from ethidium-bromide-stained agarose gels using GeneClean II (Bio101). Large-scale purifications of plasmid DNA were done with a Qiagen Midi-kit. Transformation of *E. coli* with plasmids was done by the methods of Chung et al. (1989). A bacteriophage λ vector, LambdaGem-11 (Promega), was used to construct a library of partial Sau3AI-digested chromosomal DNA fragments from *B. bacilliformis* KC583 by a standard protocol (Ausbub et al., 1995). The library was subsequently probed with the BamHI insert of pAL1 by standard hybridization protocol (Ausbub et al., 1995) to identify plaques containing the *ialAB* locus plus contiguous DNA. A LambdaGem-11 recombinant identified by this method was subsequently subcloned into the vector pUC19 and named pSAC2 (Fig. 1). After identification of the ORF encoding the putative *ctpA* gene (Fig. 2), subclones pCTPA and pCTPAA were constructed by digestion with restriction endonucleases at the appropriate sites (Fig. 1) and ligation into pUC19 by standard methods. Subclone pCTPAA was constructed by linker mutagenesis. Following endonuclease excision of an 18 bp EcoRV fragment from pCTPA, a 10 bp XhoI linker was inserted into the site by blunt-end ligation (Promega; XhoI linker: 5'-CCCTCGAGGG-3'). The resulting truncated ORF, *ctpA’s*, contains a mutated sequence beginning at base 520 of the sequence shown in Fig. 2. A subclone containing the two ORFs downstream of the *ialB* gene was constructed and will be described elsewhere (S. J. Mitchell & M. F. Minnick, unpublished).

**Nucleotide sequencing and computer analysis.** The SacI insert of pSAC2 was sequenced on both strands by the methods of Sanger et al. (1977) using an automated DNA sequencer (Applied Biosystems). In addition to M13 universal forward and reverse primers for pUC19, primers were made on an automated DNA synthesizer (Applied Biosystems). Computer analysis was done using pcgene 6.8 (Intelligenetics) for ORF identification and analysis; BLAST (Altschul et al., 1990) for multiple sequence alignments (Hofmann & Baron, 1996) for formatting of aligned sequences; BOXSHADE 3.2 (Hofmann & Baron, 1996) for multiple sequence alignments; and NUCLEOS.2D (Thompson et al. 1994) for multiple sequence alignments, and BOXSHADE 3.2 (Hofmann & Baron, 1996) for formatting of aligned sequences.

**Nucleic acid hybridization analysis.** Total DNA from *E. coli* HB101, and *B. bacilliformis* strains KC583 and KC584, was isolated and digested to completion with BamHI. The plasmid pCTPAA was digested with SacI and XhoI to release the entire *ctpA*-containing insert. The DNA samples were then separated on an ethidium-bromide-stained 1% (w/v) agarose gel, blotted to nitrocellulose (0.45 μm pore size, Schleicher & Schuell) by the method of Southern (1975) and baked for 1 h at 80 °C. A 1223 bp BamHI-EcoRI fragment of pCTPAA was purified from an agarose gel and labelled by random-primer extension using [α- 32P]dCTP and the Klenow fragment of *E. coli* polynucleotide I (Feinberg & Vogelstein, 1984). The blot was probed overnight at 50 °C with the 32P-labelled pCTPAA fragment and washed at high stringency (approximately 10% DNA mismatch) as previously described (Minnick et al., 1990). The blot was subsequently exposed for 6 h to X-ray film (X-Omat XR-5, Eastman Kodak) to visualize hybridized DNA fragments.
Removal Kit (Qiagen) and denatured at 100 °C for 5–10 min prior to addition to fresh hybridization solution. The nylon membrane was pre-hybridized at room temperature for 2 h in 0.25 M Na₂HPO₄, 7% (w/v) SDS, and 3 × Denhardt’s solution. Hybridization was carried out overnight at 45 °C with fresh hybridization solution plus the appropriate ³²P-labelled probe. The membrane was then washed twice for 15 min in 50 ml 1× SDS plus sufficient Na₂HPO₄ to give 40%, 30%, 20%, or 10% mismatch according to the DNA:DNA formula of Meinkoth & Wahl (1984) (40% mismatch was obtained with 0.115 M Na₂HPO₄, 45 °C; 30% mismatch with 0.115 M Na₂HPO₄, 55 °C; 20% mismatch with 0.029 M Na₂HPO₄, 55 °C; 10% mismatch with 0.029 M Na₂HPO₄, 65 °C). The membrane was kept damp at all times by wrapping in plastic-wrap, and was exposed to X-Omat XAR-5 film overnight at −70 °C with two intensifying screens. After each 10% mismatch experiment, the probe was stripped from the membrane by washing twice in 500 ml 0.005 M Na₂HPO₄ and 0.5% SDS at 95 °C for 20 min. Removal of the probe was confirmed by a 24 h exposure to X-ray film with two intensifying screens at −70 °C prior to applying the next probe.

To determine the distribution of the ctpA, ialA and ialB genes within Bartonella species, and to investigate their prevalence in other bacteria, a multiple-species hybridization experiment was carried out. Total chromosomal DNA was isolated as described above from five Bartonella species (B. bacilliformis KC583 and KC584, B. henselae Houston-1, B. quintana Fuller and B. vinsonii Baker), and E. coli (ATCC/Myers et al. 1979) as a negative control. The DNA samples were subjected to electrophoresis on an ethidium-bromide-stained agarose (1%, w/v) gel and vacuum-blotted to a positively charged nylon membrane (ZetaProbe GT; Bio-Rad) under alkaline conditions according to the manufacturer’s instructions. Transfer efficiency was verified by post-staining of the agarose gel with ethidium bromide. Probes were constructed as described above, using a 1223 bp BamHI–EcoRI fragment of ctpA, a 774 bp BamHI–SmaI fragment of ialB, and a 492 bp Hpal–ScaI fragment of ialA. Following a 2 h random-primer extension, probes were purified from unincorporated [α-³²P]dCTP using a QIAquick Nucleotide Removal Kit (Qiagen) and denatured at 100 °C for 5–10 min prior to addition to fresh hybridization solution. The nylon membrane was pre-hybridized at room temperature for 2 h in 0.25 M Na₂HPO₄, 7% (w/v) SDS, and 3 × Denhardt’s solution. Hybridization was carried out overnight at 45 °C with fresh hybridization solution plus the appropriate ³²P-labelled probe. The membrane was then washed twice for 15 min in 50 ml 1× SDS plus sufficient Na₂HPO₄ to give 40%, 30%, 20%, or 10% mismatch according to the DNA:DNA formula of Meinkoth & Wahl (1984) (40% mismatch was obtained with 0.115 M Na₂HPO₄, 45 °C; 30% mismatch with 0.115 M Na₂HPO₄, 55 °C; 20% mismatch with 0.029 M Na₂HPO₄, 55 °C; 10% mismatch with 0.029 M Na₂HPO₄, 65 °C). The membrane was kept damp at all times by wrapping in plastic-wrap, and was exposed to X-Omat XAR-5 film overnight at −70 °C with two intensifying screens. After each 10% mismatch experiment, the probe was stripped from the membrane by washing twice in 500 ml 0.005 M Na₂HPO₄ and 0.5% SDS at 95 °C for 20 min. Removal of the probe was confirmed by a 24 h exposure to X-ray film with two intensifying screens at −70 °C prior to applying the next probe.

---

**Table 1. Bacterial strains and plasmids**

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Description</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>F− 80dlacZΔM15 recA1 endA1 gyrA96 thi-1 hsdR17 (rK m') supE44 relA1 deoR Δ(lacZYA−argF)U169</td>
<td>BRL</td>
</tr>
<tr>
<td>HB101</td>
<td>F− thi-1 hsdS20 (rK m') supE44 recA13 ara-14 leuB6 proA2 lacY1 rpsL20 (Str') xyl-5 mil-1</td>
<td>Promega/Boyer &amp; Roulland-Dussoix (1969)</td>
</tr>
<tr>
<td>Y1090</td>
<td>Δ(lacU169) proA Δ(lon) araD139 supF [tpC22::Tn10(Tet')] hsdR (rK m') (pMC9)</td>
<td>Promega/Young &amp; Davis (1983)</td>
</tr>
<tr>
<td>Y1090-C</td>
<td>Same as Y1090 except cured of pMC9</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Bartonella</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. bacilliformis KC583</td>
<td>Neotype strain</td>
<td>ATCC/Brenner et al. (1991)</td>
</tr>
<tr>
<td>B. bacilliformis KC584</td>
<td>Peruvian isolate, 1963</td>
<td>ATCC/Brenner et al. (1991)</td>
</tr>
<tr>
<td>B. elizabethae</td>
<td>Type strain</td>
<td>ATCC/Daly et al. (1993)</td>
</tr>
<tr>
<td>B. henselae Houston-1</td>
<td>Type strain</td>
<td>CDC/Regnery et al. (1992)</td>
</tr>
<tr>
<td>B. quintana Fuller</td>
<td>Type strain</td>
<td>ATCC/Myers et al. (1979)</td>
</tr>
<tr>
<td><strong>Brucella abortus</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strain 19</td>
<td>Live attenuated vaccine strain</td>
<td>Burroughs Wellcome</td>
</tr>
<tr>
<td><strong>Proteus vulgaris</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LB-53</td>
<td>Laboratory strain</td>
<td>Int. Inst. Biochem. &amp; Biomed. Tech., Chicago, IL, USA</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUC19</td>
<td>M13mp19 derivative, Amp′</td>
<td>Yannisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>pAL1</td>
<td>pUC19 containing Bb KC583 ialA and ialB on a 1.5 kb BamHI fragment</td>
<td>Mitchell &amp; Minnick (1995)</td>
</tr>
<tr>
<td>pSAC2</td>
<td>pUC19 containing Bb KC583 ctpA, ialA and ialB on a 4.5 kb SacI fragment</td>
<td>This study</td>
</tr>
<tr>
<td>pCTPA</td>
<td>pUC19 containing Bb KC583 ctpA on a 1.7 kb SacI–(Hpal/Smal fusion) fragment</td>
<td>This study</td>
</tr>
<tr>
<td>pCTPAA</td>
<td>Same as pCTPA except 10 bp Xhol-linker inserted into EcoRV site in ctpA to create ctpA′</td>
<td>This study</td>
</tr>
</tbody>
</table>

---

To determine the distribution of the ctpA, ialA and ialB genes within Bartonella species, and to investigate their prevalence in other bacteria, a multiple-species hybridization experiment was carried out. Total chromosomal DNA was isolated as described above from five Bartonella species (B. bacilliformis KC583 and KC584, B. henselae, B. quintana, B. elizabethae and B. vinsonii). Br. abortus and E. coli were used as controls. The DNA samples were subjected to electrophoresis on an ethidium-bromide-stained agarose (1%, w/v) gel and vacuum-blotted to a positively charged nylon membrane (ZetaProbe GT; Bio-Rad) under alkaline conditions according to the manufacturer’s instructions. Transfer efficiency was verified by post-staining of the agarose gel with ethidium bromide. Probes were constructed as described above, using a 1223 bp BamHI–EcoRI fragment of ctpA, a 774 bp BamHI–SmaI fragment of ialB, and a 492 bp Hpal–ScaI fragment of ialA. Following a 2 h random-primer extension, probes were purified from unincorporated [α-³²P]dCTP using a QIAquick Nucleotide Removal Kit (Qiagen) and denatured at 100 °C for 5–10 min prior to addition to fresh hybridization solution. The nylon membrane was pre-hybridized at room temperature for 2 h in 0.25 M Na₂HPO₄, 7% (w/v) SDS, and 3 × Denhardt’s solution. Hybridization was carried out overnight at 45 °C with fresh hybridization solution plus the appropriate ³²P-labelled probe. The membrane was then washed twice for 15 min in 50 ml 1× SDS plus sufficient Na₂HPO₄ to give 40%, 30%, 20%, or 10% mismatch according to the DNA:DNA formula of Meinkoth & Wahl (1984) (40% mismatch was obtained with 0.115 M Na₂HPO₄, 45 °C; 30% mismatch with 0.115 M Na₂HPO₄, 55 °C; 20% mismatch with 0.029 M Na₂HPO₄, 55 °C; 10% mismatch with 0.029 M Na₂HPO₄, 65 °C). The membrane was kept damp at all times by wrapping in plastic-wrap, and was exposed to X-Omat XAR-5 film overnight at −70 °C with two intensifying screens. After each 10% mismatch experiment, the probe was stripped from the membrane by washing twice in 500 ml 0.005 M Na₂HPO₄ and 0.5% SDS at 95 °C for 20 min. Removal of the probe was confirmed by a 24 h exposure to X-ray film with two intensifying screens at −70 °C prior to applying the next probe.
Analysis of mRNA by Northern blotting was done by two standard methods (Ausubel et al., 1995): formaldehyde (67% w/v)/MOPS gel electrophoresis, or glyoxal/DMSO gel electrophoresis, followed by vacuum-blotting to nylon membranes (ZetaProbe GT; Bio-Rad), and probing of the membranes with an isotopically labelled DNA probe. Briefly, an RNaseasy kit (Qiagen) was used to extract total RNA from B. bacilliformis strains KC583 and KC584; and from E. coli DH5α containing either no plasmid, or pUC19, pSAC2, pCTPA or pCTPAA. Ten micrograms of RNA from each sample was denatured and run on 12% agarose gels under appropriate conditions for formaldehyde/MOPS or glyoxal/DMSO gels (Ausubel et al., 1995). RNA was transferred by vacuum-blotting to positively charged nylon membranes (ZetaProbe GT) and UV-cross-linked with a total treatment of 150 mJ. A 32P-labelled probe was constructed by random-primer extension on a DNA template constructed from a 1223 bp BamHI–EcoRI fragment of pCTPA (Ausubel et al., 1995). The membranes were probed using the manufacturer’s recommendations (Bio-Rad) for pre-hybridization, overnight hybridization (45 °C, 50% formamide), and high-stringency washing (final washes of 2 x 20 min at 65 °C in 50 ml 0.1 x SSC, 0.5% SDS). Transfer of RNA to the nylon membranes was verified after probing the membrane by the methylene-blue staining method of Herrin & Schmidt (1988).

**Invasion assays.** These were done as previously reported (Mitchell & Minnick, 1995). Briefly, overnight cultures of E. coli DH5α containing the appropriate recombinant plasmid (pUC19, pSAC2, pCTPA or pCTPAA) were standardized to a concentration of 2.5 x 10⁶ cells ml⁻¹. Recently outgrown human erythrocytes (American Red Cross) were washed in PBS and added to a final concentration of 5 x 10⁷ cells to yield a final multiplicity of infection of 50 bacteria:1 erythrocyte. Samples (0.5 ml total volume) were incubated for 1 h at 37 °C, centrifuged at 1320 g for 20 s to associate bacteria and erythrocytes, and incubated for a further 2 h at 37 °C. Samples were then washed, treated with 200 µg gentamicin ml⁻¹ for 2 h at 37 °C, and washed again to remove gentamicin. Intra-cellular bacteria were then released by lysing the erythrocytes with 100 µl deionized water. Samples were then plated onto LB/ampicillin plates for overnight growth. At the end of this period, colonies were counted the following morning. All experiments were done in triplicate.

**Protein expression in vitro and in vivo.** Expression of protein products in vitro was done using two commercially available E. coli-derived in vitro transcription/translation (IVTT) kits (Amersham and Promega). Briefly, 2 µg purified plasmid DNA was incubated with 10-µCi (37 MBq) [35S]Met/Cys (EX-PRESS, NEN Dupont) and the IVTT kit for 2 h at 37 °C, followed by immediate addition of 80 µl Laemmli sample buffer (LSB) (Laemmli, 1970), heating to 100 °C for 10 min, and storage at −20 °C. Aliquots of 15 µl were heated to 100 °C for 10 min, quickly centrifuged to collect the supernatant, and subjected to 0.1% SDS-PAGE (12.5%, w/v, acrylamide) analysis. The gel was vacuum-dried and subjected to autoradiography overnight (X-Omat XAR-5).

Because of the autolytic processing observed in vitro (see Fig. 5), two separate protease-inhibitor treatments were applied during subsequent IVTT incubations. In the first experiment, a mixture of inhibitors was made with a final concentration of 6 µg aprotinin ml⁻¹ (Sigma), 0.5 µg leupeptin ml⁻¹ (Sigma), 0.2 µg pepstatin ml⁻¹ (Sigma). An IVTT reaction was conducted as described above and divided into two equal aliquots. The protease-inhibitor mix was applied to one aliquot at 0 min and to the other aliquot at 10 min, and the reaction allowed to proceed for a total time of 2 h. The samples were then treated as described above and analysed by SDS-PAGE and autoradiography. The second protease inhibitor mixture was prepared using a Complete Protease Inhibitor Cocktail Tablet (Boehringer Mannheim, no. 1697498). One tablet was dissolved in 2 ml deionized water and 1 µl of this solution was used per inhibition reaction, to give a final concentration of 10 µg aprotinin ml⁻¹, 0.5 µg leupeptin ml⁻¹, 2 µmol Pefablock ml⁻¹ (a water-soluble PMSF analogue), and 1 µmol EDTA ml⁻¹. The experiment was done in an identical fashion to that described above, with one half of the IVTT sample treated at 0 min and the second at 10 min, followed by a 2 h incubation at 37 °C. The samples were then treated as described above and analysed by SDS-PAGE and autoradiography.

Expression of the ctpA-encoded protein in vivo was accomplished by the Maxicell procedure of Sancar et al. (1979). E. coli HB101 containing recombinant plasmids were irradiated with 195 µW cm⁻² UV light for 50 s at a 15 cm exposure distance. Maxicells containing the 35S-labelled proteins were then pelleted, resuspended in 75 µl LSB, and heated to 100 °C for 10 min. Samples were centrifuged at 16000 g for 1 min, and 20 µl of each supernatant was then subjected to 0.1% SDS-PAGE (12.5%, w/v, acrylamide) and visualized by overnight autoradiography of the dried gel.

**Haemolysis, collagenase and caseinase assays.** Haemolysis assays were done by streaking E. coli HB101 containing either no plasmid, or pUC19, pSAC2, pCTPA or pCTPAA, onto HIBB solid medium and growing overnight at 37 °C. Single plates were scored visually for haemolytic clearing.

Collagenase assays were done by stabbing the above E. coli HB101 strains into 5 ml nutrient gelatin tubes in duplicate. A positive control was done with Proteus vulgaris. Samples were grown at 37 °C for 7 d and assayed daily by chilling the tubes to 20 °C for 30 min to test for liquefaction or gelling.

Caseinase assays were done using the Lon⁻ E. coli strain Y1090 (Promega) after curing the strain of pMC9 (Amp⁺, Promega). Curing was done by growing the strain at 37 °C in LB for 30 h with periodic replacement of LB to maintain exponential growth. At the end of this period, colonies were identified which had become Amp⁻ and curing was confirmed by loss of the plasmid. The strain Y1090-C was then transformed with pUC19, pSAC2, pCTPA or pCTPAA as described above, and transformants confirmed by plasmid isolation. E. coli Y1090-C transformants containing the above plasmids were then plated in duplicate onto skim-milk plates and grown at 37 °C overnight to assay for casein proteolysis.

**RESULTS**

A 4.3 kb Sad fragment of the B. bacilliformis chromosome contains ORFs flanking ialAB

In order to determine if the invasion-associated locus (ialAB) of B. bacilliformis is flanked by other virulence factors in a 'pathogenicity island', we isolated regions upstream and downstream of the original 1469 bp BamHI fragment containing ialAB. By probing a bacteriophage λ library of Sau3AI-derived B. bacilliformis KC583 fragments with a 32P-labelled 1469 bp BamHI fragment conducting the ialAB locus, approximately 10 kb of the Bartonella chromosome was isolated. A 4.3 kb SacI fragment of the clone was found to contain
The 5' 1.8 kb Sad-HpaI fragment of pSAC2 contains a putative C-terminal protease by computer database searches and subsequently named ialA. Two minor ORFs of unknown significance were found between the ialA and tsp genes (Fig. 1). We have continued to sequence downstream of the 3' SacI site of pSAC2 and have recently identified an inorganic pyrophosphatase gene which we will describe elsewhere.

The region upstream of the putative start codon of ctpA (nucleotide 1, Fig. 2) contains hexamers with similarity to the E. coli consensus promoters (McLure, 1983). The putative −35 (TTTACA) and −10 (TGAATT) hexamers are followed by a putative ribosome-binding site (AGGA) with perfect identity to the E. coli consensus sequence (Gold et al., 1981) located 6 nt upstream of the start codon (Stormo et al., 1982). A possible rho-independent transcription terminator (Rosenberg & Court, 1979) is found 21 nt downstream of the stop codon and is composed of a 14-base inverted repeat separated by 3 nt. The hairpin predicted to form by this repeat has a thermostable free energy of −16.5 kcal mol⁻¹ (−69.0 kJ mol⁻¹) (PCGENE) and is immediately followed by four thymidines, further supporting the notion that this is the transcription-termination site.

Computer analysis of the ORF indicated that the predicted protein has a length of 434 amino acid residues, with a molecular mass of approximately 47 kDa. A probable secretory signal cleavage site (Von Heijne, 1986) was found between residues 23 and 24, which would result in a secreted protein (411 residues) (PCGENE) and is immediately followed by four thymidines, further supporting the notion that this is the transcription-termination site.

Further analysis of the predicted full-length CtpA protein (434 residues) was done by searching various databases with BLAST (Altschul et al., 1990). Several high-probability alignments were found to a group of carboxy-terminal processing proteases, including the 46 kDa CtpA protease of Synechocystis 6803 (SYCCTPA; Shestakov et al., 1994) and the 73 kDa Tsp of Bartonella bacilliformis C-terminal protease.
of *E. coli* (Silber et al., 1992). Alignments of all three predicted protein sequences are shown in Fig. 3, and indicate multiple areas of strong homology. The individual alignment of the *B. bacilliformis* CtpA with SYCCTPA had an identity of 30% (FASTA) and overall similarity of 69% (PCGENE). Alignment of the *B. bacilliformis* CtpA with the *E. coli* Tsp revealed 19% identity (FASTA) and 46% similarity (PCGENE). When the N- and C-terminal overhanging portions of Tsp were excluded from the analysis, the alignment of the *B. bacilliformis* CtpA with Tsp increased to 22% identity (FASTA) and 62% overall similarity (PCGENE). The BLAST analysis also revealed moderate homology of the *B. bacilliformis* CtpA to several interphotoreceptor retinol-binding proteins (IRBPs).

DNA hybridization analyses indicate that the ctpA, *ialA* and *ialB* genes are highly conserved within *Bartonella* species, but not in a related intracellular pathogen or in *E. coli*

In order to verify that the ctpA-containing DNA fragment was of *Bartonella* origin, DNA hybridization analysis was done using BamHI-digested DNA from *B.
**Fig. 3.** Multiple alignment of the *B. bacilliformis* CtpA (Bb_CtpA) with the Synechocystis sp. 6803 CtpA (Sy_CtpA) and the *E. coli* Tsp (Ec_Tsp) proteins (all are shown in their immature form). Identical amino acid residues are noted in black, conserved residues in grey, and introduced gaps by hyphens. The three catalytic residues and four structurally essential residues of Tsp (Keiler & Sauer, 1995) are indicated above the sequence by black circles and squares, respectively. The GenBank accession numbers for Bb_CtpA, Sy_CtpA, and Ec_Tsp are L37094, A53964, and A41798, respectively.
Fig. 4. Detection of the ctpA gene in the B. bacilliformis chromosome by DNA hybridization. (a) Ethidium-bromide-stained agarose gel (1%, w/v) containing: 1, λ/HindIII DNA size standards; 2–4, BamHI-digested chromosomal DNA of B. bacilliformis KC583 (2), B. bacilliformis KC584 (3) and E. coli HB101 (4); and 5, SacI-XbaI digested pCTPA. (b) The corresponding autoradiograph following DNA hybridization with a 32P-labelled BamHI-EcoRI subfragment of the pCTPA DNA insert. Lanes are the same as in (a). The double arrow points to the pCTPA fragment producing the hybridization signal. ctpA-containing chromosomal fragments of the two B. bacilliformis strains are bracketed.

The ctpA gene does not affect the erythrocyte-invasion phenotype conferred by ialAB

To determine if the ctpA gene had a direct effect upon erythrocyte invasion, invasion assays were done using the modified gentamicin-protection method we described previously (Mitchell & Minnick, 1995). Experiments with pCTPA (containing only ctpA) indicated that this gene does not confer an invasive phenotype upon E. coli (data not shown). An additional experiment with pSAC2 (containing ctpA, ialA, ialB, ORF1 and ORF2) revealed that there was no statistically

<table>
<thead>
<tr>
<th>Target DNA</th>
<th>ctpA</th>
<th>ialA</th>
<th>ialB</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. bacilliformis KC583</td>
<td>++ ++ ++ ++ ++</td>
<td>++ ++ ++ ++</td>
<td>++ ++ ++ ++</td>
</tr>
<tr>
<td>B. bacilliformis KC584</td>
<td>++ ++ ++ ++ ++</td>
<td>++ ++ ++ ++</td>
<td>++ ++ ++ ++</td>
</tr>
<tr>
<td>B. quintana</td>
<td>++ ++ ++ ++ ++</td>
<td>++ ++ ++ ++</td>
<td>++ ++ ++ ++</td>
</tr>
<tr>
<td>B. henselae</td>
<td>+ + - - - - - -</td>
<td>+ + - - - - - -</td>
<td>+ + - - - - - -</td>
</tr>
<tr>
<td>B. vinsonii</td>
<td>+ + - - - - - -</td>
<td>+ + - - - - - -</td>
<td>+ + - - - - - -</td>
</tr>
<tr>
<td>B. elizabethae</td>
<td>- - - - - - - -</td>
<td>- - - - - - - -</td>
<td>- - - - - - - -</td>
</tr>
<tr>
<td>Br. abortus</td>
<td>- - - - - - - -</td>
<td>- - - - - - - -</td>
<td>- - - - - - - -</td>
</tr>
<tr>
<td>E. coli</td>
<td>- - - - - - - -</td>
<td>- - - - - - - -</td>
<td>- - - - - - - -</td>
</tr>
<tr>
<td>pSAC2</td>
<td>+ + + + + + + +</td>
<td>+ + + + + + + +</td>
<td>+ + + + + + + +</td>
</tr>
</tbody>
</table>

ND, Not done; ND*, not done due to membrane degradation.
† Done in a previous study (Mitchell & Minnick, 1995).
significant increase in the level of erythrocyte invasion conferred by ctpA, as compared to pIAL1 (containing only the ialA and ialB genes) (data not shown).

The protein product of ctpA does not affect ialA or ialB, but may autolytically process itself in vitro where it cannot be easily inhibited

Because ctpA is so closely linked to ialA in the genome, we conducted several experiments to determine if ctpA might affect ialA and/or ialB gene products in a more subtle fashion than was detectable by the invasion assay. As it is possible that the ctpA gene product might function incorrectly in the E. coli environment, we analysed both the transcription and translational products. Two separate analyses of RNA transcription products from B. bacilliformis KC583 and KC584, as well as from E. coli DH5α containing pSAC2, pCTPA, pIAL1 or pUC19, were done. In both cases, probing with a single-stranded 32P-labelled DNA probe constructed from an internal ctpA template revealed that the transcript exists in both B. bacilliformis strains and in E. coli containing the ctpA gene cloned within pSAC2 or pCTPA (data not shown). No signal was seen on lanes containing control RNA from E. coli DH5α, E. coli DH5α(pUC19) or E. coli DH5α(pIAL1) in either experiment. However, because of anomalous migration of the transcripts, presumably caused by association with the 23S and 16S rRNA, we were unable to size the transcripts accurately. Transfer of total RNA in all lanes was verified by post-staining the membrane with methylene blue after probe hybridization and washing.

In order to analyse the expected protein products from ctpA, a prokaryote-directed IVTT kit was used to express the ctpA and/or ialA and ialB genes. As shown in Fig. 5, several protein products were produced by the ctpA gene (lanes 3 and 4) that were not found in either a control (pUC19, lane 1), or clones containing other Bartonella DNA (pIAL1, lane 2). One group of bands in the 43–46 kDa range appears to correspond to the full-length 1302 bp ORF, whereas a less intense group of bands at approximately 36–37 kDa may represent the protein encoded by the 993 bp ORF beginning with the alternative start site (Fig. 2). Because SYCCTPA has been reported to possibly undergo post-translational processing in addition to removal of its leader peptide (Anbudurai et al., 1994), we were interested to see if the Bartonella CtpA might be modifying itself or other CtpA molecules. Therefore, we mutated the ctpA gene in pCTPA by enzymically removing an 18 bp EcoRV fragment and inserting a 10 bp XhoI linker fragment in its place (Fig. 1). This insertion resulted in a nonsense mutation in the ctpA gene, thus creating ctpA′. The truncated protein product encoded by the 549 bp ctpA′ contains only 173 N-terminal amino acid residues, followed by a 10-residue missense sequence (NH2-PSRYHKCKCC-COOH) and an early stop codon, and would have a predicted molecular mass of 20 kDa. As shown in Fig. 5, lane 5, it is apparent that the multiple protein products seen in both the 36–37 kDa and 43–46 kDa ranges all disappear when only ctpA′ is expressed. Subsequent analysis of the ctpA′ product on a longer SDS-PAGE system revealed that it was a single polypeptide migrating at approximately 16–18 kDa (data not shown). The analysis was repeated with a second prokaryote-directed IVTT kit from a different manufacturer with identical results (data not shown). Neither ORF1 nor ORF2 appeared to produce visible protein products (predicted molecular masses of 11 and 12 kDa, respectively) with either IVTT system (Fig. 5, lane 3).

Because SYCCTPA has been reported to possibly undergo removal of not only its N-terminal secretory leader peptide, but also additional residues during maturation (Anbudurai et al., 1994), the in vitro results in Fig. 5 (lanes 3 and 4) may indicate a similar C-terminal processing event for CtpA. It is known that SYCCTPA cleaves a known target (protein D1) between two alanine residues located nine residues from the C-terminus. In addition, the E. coli Tsp shows strong
preference for cleaving target proteins containing alanine near the C-terminus (Keiler & Sauer, 1996). As shown in Fig. 2, the B. bacilliformis CtpA contains several alanine pairs near its C-terminus, which could serve as targets for either autolytic processing or proteolysis of other CtpA molecules in vitro.

Because the ctpA gene is predicted to contain a secretory signal leader peptide (Fig. 2), we performed a maxicell analysis to better define the mature form of CtpA. As shown in Fig. 6, lane 3, the mature CtpA is approximately 41–42 kDa in mass, slightly less than the predicted molecular mass of the mature form of the secreted protein (45 kDa). The CtpA protein appears to be insert-specific, as shown by the lack of a corresponding band in E. coli HB101 alone or E. coli HB101 containing pUC19 (Fig. 6, lanes 1 and 2, respectively).

The CtpA protein does not appear to be an extracellular haemolysin, collagenase or caseinase

Computer analysis revealed strong homology between the Bartonella CtpA, the Synechocystis 6803 CtpA, and the E. coli Tsp. Therefore, several assays were done to determine if the Bartonella CtpA might have similar functions to these proteases, or unique features reflecting the haemotrophic lifestyle of B. bacilliformis. A haemolysin assay revealed that the products encoded by ctpA (and ialAB) did not possess haemolytic activity when expressed in E. coli. In addition, a collagenase assay indicated that the gene products did not act to break down gelatin in the growth medium. A proteolysis assay was done by introducing the recombinant plasmids into the Lon+ E. coli strain Y1090-C. Overnight growth at 37 °C on skim-milk plates did not reveal any clearing around colonies, indicating that CtpA is not secreted extracellularly.

**DISCUSSION**

We have identified a putative carboxy-terminal processing protease gene, ctpA, immediately upstream of a two-gene locus previously implicated in erythrocyte invasion by B. bacilliformis (Mitchell & Minnick, 1995). Our original goal was to explore the sequences flanking ialAB to determine if the locus was part of a larger pathogenicity island. By exploring the distribution of ctpA, ialA and ialB, we found that these genes are conserved throughout the genus Bartonella, but do not appear to be highly conserved across genera. However, computer analysis of the predicted CtpA protein product indicates that it belongs to a novel class of C-terminal proteases, enzymes which all appear to be involved in stress response. Therefore, despite the lack of a direct connection to the erythrocyte-invasion virulence mechanism, this gene may be an essential component for the bacteria to survive the stressful lifestyle imposed upon an intracellular pathogen.

We have determined that ctpA does not directly impact the ialAB locus. Invasion, haemolysin, collagenase and caseinase assays all demonstrated that CtpA is probably not active extracellularly. Despite the close physical proximity of the genes, CtpA does not appear to modify either IalA or IalB directly, as demonstrated by the constant molecular mass of these proteins in the presence or absence of CtpA (Figs 5 and 6). This is less surprising in light of the recent description of the Synechocystis sp. 6803 genome (Churin et al., 1995), in which mapping of ctpA revealed that it was over 300 kbp away from its nearest known target-encoding gene (psbA-2). However, one caveat is that ialA expression appears to be strikingly enhanced in the presence of ctpA, as shown in Fig. 5, lane 3 (ialA with ctpA upstream), versus Fig. 5, lane 2 (ialA without ctpA). Although there is a potential rho-independent terminator between the genes (Fig. 2), it is possible that this hairpin may function as a transcriptional pause site (Landick et al., 1996). Alternatively, there may simply be a basal level of transcription read-through in the IVTT system. If there is transcriptional pausing between the genes, it may imply the existence of an operon.

Based upon computer analysis of the predicted CtpA protein, it seems clear that this enzyme probably acts by binding substrates with a hydrophobic C-terminus and enzymically cleaving them. As seen in Fig. 3, the Bartonella CtpA contains several highly conserved motifs found in a group of enzymes which all bind.
hydrophobic ligands (IRBPs, Tsp, and the photosystem II Ctpases). Recent work by Keiler & Sauer (1995) identified the active-site residues of mature E. coli Tsp protease as Ser-430, Asp-441 and Lys-435. All three amino acids lie within a region of IRBP homology (Silber et al., 1991). As seen in Fig. 3, each active-site residue of Tsp has an identical counterpart in the Bartonella CtpA sequence. In addition, three of the four residues identified as playing a structural role essential for Tsp activity are identical in Bartonella CtpA, while the fourth residue is conservatively substituted. These data suggest that the Bartonella CtpA contains a similar active site, and that the enzyme may function by the serine-lysine dyad mechanism proposed for Tsp (Keiler & Sauer, 1995).

In addition to the sequence-predicted binding of hydrophobic target proteins, the Bartonella CtpA shares other unusual characteristics with Tsp and SYCCTPA. It has been shown that the common protease inhibitors PMSF, iodoacetamide, EDTA and pepstatin have little effect upon Tsp (Miller, 1996) or SYCCTPA (Bowyer et al., 1992), as we observed in attempting to inhibit the autolytic process of CtpA in vitro. In combination with the lack of homology between the C-terminal proteases and other known proteases, this insensitivity to inhibitors has led at least one group to postulate that these enzymes possess an unusual catalytic site (Taguchi et al., 1995). It is also worth noting that both SYCCTPA and Tsp are believed to be stress-response proteases. It is clear that the extremely rapid degradation of D1 by light energy must be balanced by an equally rapid proteolytic maturation of nascent D1 by SYCCTPA (Barber & Andersson, 1992). Similarly, E. coli mutants lacking a functional Tsp are sensitive to both heat shock and osmotic stress (Hara et al., 1991). A Tsp homologue from Salmonella typhimurium has been implicated in intracellular survival of the bacteria within the acidic environment of macrophages (Baumler et al., 1994).

Very recently, Tsp has also been shown to be responsible for degrading aberrant proteins synthesized from damaged mRNA in E. coli (Keiler et al., 1996). The proteins are modified by addition of a short hydrophobic C-terminal marker (AANDENYALAA) from the structural RNA molecule ssrA, after which they are specifically recognized by Tsp and degraded. In all of these cases, it is apparent that this class of C-terminal proteases acts to remove or replace 'stressed' protein(s). Although the target of the Bartonella CtpA has not yet been identified, the pathogen's passage from insect vector to human circulation and subsequently to host cells must subject it to a variety of stressors.

Because the target of the other C-terminal proteases was identified prior to the actual protease, generalized techniques for finding Ctpase targets have not been developed. A further complication is that no method of directly manipulating Bartonella genomes exists, so all work must necessarily be done in E. coli. Therefore, at this time we can only speculate upon possible targets. However, the autolytic in vitro behaviour bears further investigation. Anbudurai et al. (1994) proposed that SYCCTPA may be further modified after loss of the secretory-signal leader peptide, based upon the difference between the predicted molecular mass of the mature protein (43 kDa) and the observed mass (39 kDa). We obtained similar results with CtpA, wherein the predicted molecular mass of the mature protein (45 kDa) was considerably larger than that observed in vivo by SDS-PAGE (42 kDa; Fig. 6, lane 3). Although this may be due to aberrant migration in the SDS-PAGE system, the multiple in vitro protein products (Fig. 5, lanes 3 and 4) support the notion that the C-terminus may be modified. It is apparent that these products arise only from an ORF encoding the full C-terminal end, as an early truncation mutant, ctpA5', did not produce any of these proteins. In addition to the full-length, presumably periplasmically directed, CtpA the alternative start site may create a population of cytoplasmically bound proteases. In support of this hypothesis, Tsp activity has been noted both periplasmically and cytoplasmically, although there is some disagreement as to whether the cytoplasmic activity is artifactual (Miller, 1996). Although the products of both the full-length and alternative start site were clearly observed in vitro (Fig. 5, lanes 3 and 4), only a single product was found in vivo (Fig. 6, lane 3). It is possible that the alternative start site is insignificant within the cellular milieu. It is not clear whether CtpA is truly autocatalytic during folding and maturation, or whether it is the target of its fraternal neighbours in vitro, or both.

Sequence analysis also supports a possible autocatalytic activity of CtpA (see Fig. 2, putative C-terminal cleavage sites). The site of D1 cleavage by SYCCTPA has been the subject of intense investigation, beginning with identification of a scissile bond between two alanine residues 9 amino acids from the C-terminus of spinach (Takahashi et al., 1988). In the green alga Scenedesmus obliquus, the mature form of CtpA appears to cleave at a conserved motif consisting of a small nonpolar residue, a larger nonpolar residue, a negatively charged residue, a nonpolar residue, and a nearby proline, respectively (Bowyer et al., 1992). We note a slightly different, but conserved, domain on the carboxy side of both putative scissile bonds in Fig. 2: A-AFVPRD, and A-AFPDP. In both cases, cleavage may occur between the two alanine residues, and the C-terminal phenylalanine, proline and aspartic acid may be responsible for correct site geometry. The CtpA protein from Synechocystis sp. 6803 cleaves between an Ala-Ser bond 16 residues from the carboxy terminus (Nixon et al., 1992), and Ala-Ala mutants were cleaved equally well. Nixon et al. (1992) demonstrated that SYCCTPA can also cleave the D1 precursor protein from other organisms (including higher plants), indicating that there is probably a conserved structural motif in the target that is necessary for scission, rather than an exact primary sequence. The substrate specificity of Tsp appears to be based upon two critical factors: a free carboxyl group at the target protein's C-terminus, and one to three small hydrophobic residues (A > S > V preferred) at the C-terminus (Keiler et al., 1995).
S. but the bulky phenylalanine group at the CtpA terminus Tsp. CtpA and SYCCTPA may function in a similar, but not identical, fashion to Tsp, as the CtpA enzymes recognize residues immediately surrounding the scissile bond, whereas Tsp appears to recognize the C-terminal residues.

Due to its varied lifestyle, *B. bacilliformis* is subjected to many environmental stresses. The *ctpA* gene we have identified in this study is probably involved in maintenance of a highly stressed or rapidly degraded protein. Several residues.

many environmental stresses. The *ctpA* gene we have identified in this study is probably involved in maintenance of a highly stressed or rapidly degraded protein. Several residues.

host environment. Studies on the possible role of the protease in intracellular survival are presently under way.

**ACKNOWLEDGEMENTS**

We thank Joan Strange (The University of Montana Murdoch Molecular Biology Facility) for excellent technical assistance with nucleic acid sequencing, Kay Crull (American Red Cross) for providing us with outdated human erythrocytes, and Dr Russell Regnery (CDC-P) for the kind gift of *B. henselae*. This work was supported by Public Health Service grants AI4050 and RR10169 from the National Institutes of Health (NIAID).

**REFERENCES**


Received 22 July 1996; revised 2 November 1996; accepted 12 November 1996.