# Cloning, Expression, and Primary Structure of a Chlamydia trachomatis Binding Protein

RAVI KAUL,<sup>1</sup> KENNETH L. ROY,<sup>2</sup> AND WANDA M. WENMAN<sup>1,2\*</sup>

Departments of Pediatrics<sup>1</sup> and Microbiology,<sup>2</sup> University of Alberta, Edmonton, Alberta T6G 2R7, Canada

Received 2 June 1987/Accepted 24 August 1987

The gene encoding an 18,000-dalton eucaryotic cell-binding protein of *Chlamydia trachomatis* serovar  $L_2$  was cloned into *Escherichia coli*, and the nucleotide sequence of a 1,658-base-pair *Pst*I restriction endonuclease fragment encoding this protein was determined. The recombinant chlamydial gene consists of a 486-base-pair open reading frame encoding a polypeptide of molecular weight 18,314. The resultant polypeptide, comprising 162 amino acids, possesses a highly charged carboxy-terminal end. The expression of this recombinant protein is under the control of a vector promoter. The recombinant 18,000-dalton protein possessed the same eucaryotic cell-binding characteristics as did the native chlamydial 18,000-dalton protein when electrophoresed and transferred to nitrocellulose. Polyclonal antibodies to the recombinant protein exhibited neutralizing activity.

Chlamydiae are obligate intracellular parasites of eucaryotic cells which have been linked to an expanding spectrum of human disease and constitute an important public health problem (6, 19). The infectious form of the organism, elementary bodies (EB), attaches to host cells and, once ingested, initiates the subsequent developmental cycle consisting of intracellular, metabolically active reticulate bodies (25).

Recently, two proteins of molecular ratios  $(M_rs)$  18,000 and 31,000, associated with the extracellular EB have been found to bind to eucaryotic cells (5, 28). These Chlamydia trachomatis proteins are of particular interest since the attachment of the pathogen to its host cell represents a critical point of interaction and is necessary for successful invasion. Antibodies raised against these chlamydial cellbinding proteins inhibit association of EB to host cells and possess neutralizing activity as well (26, 28). These data suggest a protective role for such antibodies. The eucaryotic cell-binding proteins are present only on EB and not on the intracellular, noninfectious reticulate bodies. Moreover, the major 18,000-dalton binding protein is present on all C. trachomatis strains which have been examined to date (5, 26) and is a surface-exposed component. The biological importance of the 18,000-dalton protein, coupled with our interest in examining potential C. trachomatis vaccine candidates, prompted us to use a genetic approach to prepare quantities of this protein for subsequent characterization. We describe here the molecular cloning and expression in Escherichia coli of an 18,000-dalton C. trachomatis binding protein and report its complete nucleotide sequence. We named this protein chlanectin because of its cell adhesion properties.

### **MATERIALS AND METHODS**

**Organisms.** C. trachomatis serovar  $L_2$  ( $L_2$ /434/Bu) was used for preparation of chromosomal DNA. Chlamydiae were grown in HeLa 229 cells for 48 h and purified as described previously (9). E. coli JM83 and JM105 were used as recipients for DNA transformation or transfections (24), and E. coli P678-54 (1) was used as material for minicell

5152

preparations. The plasmids used were pUC8, pUC118, and pUC119, which were provided by J. Vieira. *E. coli* was grown routinely in  $2 \times$  YT medium (1.6% tryptone, 1% yeast extract, 0.5% sodium chloride). JM83 cells harboring pUC8 or transformants were grown in 100 µg of ampicillin per ml.

**DNA enzymes.** The restriction enzymes used, T4 DNA ligase, and the large fragment of *E. coli* DNA polymerase I (Klenow fragment) were purchased from New England BioLabs, Inc., or Boehringer Mannheim Biochemicals. These enzymes were used in conditions specified by the manufacturers.

**DNA isolation and manipulations.** Chromosomal DNA was isolated from purified EB as described previously (7, 27). Plasmid vector DNA was isolated by the procedure of Birnboim and Doly (2) and purified by cesium chloride gradient centrifugation in the presence of ethidium bromide (17). Chromosomal DNA was partially digested with restriction endonuclease *PstI* and sized on 0.8% agarose gels (11). DNA fragments (4 to 6 kilobases [kb]) were pooled and ligated to *PstI*-digested pUC8. Competent JM83 cells were prepared by calcium chloride and transformed with ligated DNA. White ampicillin-resistant recombinant colonies were identified on isopropyl- $\beta$ -D-thiogalactopyranoside and X-Gal plates.

**Expression of the recombinant 18,000-dalton protein.** Positive recombinants were grown in YT broth. The cells were harvested, suspended in electrophoresis buffer (0.0625 M Tris, 2% sodium dodecyl sulfate [SDS], 10% glycerol, 0.001% bromophenol blue, 5%  $\beta$ -mercaptoethanol), and boiled for 3 min. Proteins were separated by 12.5% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (10) and either stained with Coomassie blue or transferred to nitrocellulose (23). Blots were then reacted with <sup>125</sup>I-labeled HeLa membranes as described previously (28) or hyperimmune rabbit antiserum to recombinant protein, followed by probing with <sup>125</sup>I-labeled protein A. HeLa cell membranes were isolated by a protocol followed earlier (28) and labeled with <sup>125</sup>I by the lactoperoxidase method (14).

Analysis of recombinant proteins. E. coli minicells harboring pCT161/18 or pUC8 were grown to stationary phase and purified by sucrose density centrifugation. The cells were labeled with  $[^{35}S]$ methionine (4), and then labeled polypep-

<sup>\*</sup> Corresponding author.



FIG. 1. Construction of chlanectin expression plasmids. A linearized physical map of inserted sequences and their properties is shown. The expression of the recombinant polypeptides, as monitored by their binding to <sup>125</sup>I-labeled HeLa cell membranes or by immunoblot assay using hyperimmune rabbit antisera to cloned chlanectin, is indicated on the right. The cross-hatched bar indicates the protein-coding region of the chlanectin gene on the cloned DNA insert, and the arrow shows the direction of transcription from the vector promoter. Only the relevant restriction sites are shown. Abbreviations: P, *PstI*; H, *HaeIII*; X, *XbaI*.

tides were analyzed by SDS-PAGE, followed by fluorography with En<sup>3</sup>Hance (New England Nuclear Corp.) to increase the autoradiographic signal.

In vitro transcription translation studies were performed with 1 µg of DNA and 5 µCi of [<sup>35</sup>S]methionine with *E. coli* S-30 extract (Boehringer Mannheim). Labeled polypeptides were resolved by SDS-PAGE followed by fluorography and examined by autoradiography. In some experiments, the in vitro-synthesized gene products were reacted to <sup>125</sup>I-labeled HeLa cells. In these experiments, [<sup>35</sup>S]methionine was deleted from the reaction and replaced by unlabeled methionine. The products were resolved and transferred to nitrocellulose after electrophoretic separation and then reacted to <sup>125</sup>I-labeled HeLa cells as previously described (28).

Localization of recombinant protein. The inner and outer membrane fractions of *E. coli* JM83 harboring pCT161/18 were separated by the procedure of Yamato et al. (29). The French-pressed pellet (78,000 × g) was dissolved in 3 mM EDTA and layered on 44% (wt/vol) sucrose-3 mM EDTA (pH 7.2) and then centrifuged at 78,000 × g overnight. The upper layer, consisting of inner membrane fraction, was diluted threefold in 3 mM EDTA (pH 7.2), while the outer membrane pellet fraction was suspended and layered on 52%-56% (wt/vol) sucrose-EDTA gradients. After centrifugation at 78,000 × g for 12 h, the band on the 56% sucrose layer was collected and diluted threefold in 3 mM EDTA. Both inner and outer membrane fractions were pelleted and repurified through sucrose twice by centrifugation at 176,000 × g for 1 h.

The periplasmic proteins were isolated by the cold-shock procedure of Dougan et al. (3), while the cytoplasmic proteins were prepared from spheroplasts obtained as described above. The spheroplasts were dissolved in cold water and French pressed three times at 12,000 lb/in<sup>2</sup>. Unlysed cells were removed by centrifugation at 8,000  $\times$  g for 15 min, while the supernatant was centrifuged for 1 h at 100,000  $\times$  g to remove any crude membrane fraction.

Neutralization assay. Hyperimmune rabbit antiserum to recombinant 18,000-dalton protein purified from SDS-PAGE samples was raised as described previously (7, 28). These antisera were then used in neutralization assays as described earlier (26, 28). In each assay, normal rabbit serum was used as a control.

Sequence determination. The nucleotide sequence of both strands of the 1.7-kb PstI fragment of plasmid pCT161/18

was determined by the dideoxy nucleotide chain termination method (18) with M13 mp18 and M13 mp19 recombinant clones (12). The DNA fragments used for sequencing were generated by restriction endonuclease digestion, followed by electrophoresis on 5% polyacrylamide gels. Subclones were subsequently generated, on the basis of sequencing information, either through single or double digests with restriction endonucleases. The 17-base-pair (bp) oligonucleotide primer was purchased from the Regional DNA Synthesis Facility, University of Calgary. Amersham Corp. provided [ $\alpha$ -<sup>32</sup>P]dATP (3,000 Ci/mmol).

## RESULTS

Identification of recombinant chlanectin. A recombinant DNA library comprising 4- to 6-kb fragments of C. trachomatis serovar L<sub>2</sub> DNA was constructed by using the expression vector pUC8. From 300 recombinants identified on the basis of ampicillin resistance and insertional inactivation of the lacZ gene, one clone, designated pCT161, encodes a polypeptide of  $M_r$  18,000, identical in size to a chlamydial eucaryotic cell-binding protein. The chimeric pCT161 contains approximately 4.7 kb of C. trachomatis DNA. Complete digestion of pCT161 with the restriction endonuclease PstI generated three fragments, one corresponding to vector DNA and the other two representing the insert. The latter fragments were 3 and 1.7 kb. Subcloning experiments revealed that the sequence encoding the 18,000-dalton protein is confined to the small 1.7-kb PstI fragment (Fig. 1). This subclone, designated pCT161/18, was used in further experiments.

Expression and biological features of the recombinant 18,000-dalton protein. Evidence that the recombinant 18,000-dalton polypeptide is the *C. trachomatis* binding protein was derived from its binding to  $^{125}$ I-labeled HeLa cell membranes (Fig. 2, lane 5). Labeled HeLa cells also reacted with native



FIG. 2. Identification of the chlanectin gene product synthesized by pCT161/18 both in vivo and in vitro. (A) Coomassie blue-stained SDS-PAGE analysis of the polypeptides synthesized by the vector pUC8 and the constructed plasmid pCT161/18. Lanes 1 to 4 represent standard molecular weight markers, pCT161/18 in *E. coli* K-12 strain JM83, pUC8 in JM83, and serovar L<sub>2</sub> EB, respectively. (B) Parallel blot of the gel shown in panel A reacted to <sup>125</sup>I-labeled HeLa cell membranes; lanes 5 to 7 are identical to lanes 2 to 4. (C and D) Gel electrophoresis pattern of <sup>35</sup>S-labeled gene products synthesized in minicells (C) and the in vitro transcription-translation coupled system (D). Lanes 8 and 10 represent pCT161/18, and lanes 9 and 11 represent pUC8-directed polypeptides. (E) Parallel blot of an unlabeled gel similar to the gel shown in panel D reacted to <sup>125</sup>I-labeled HeLa cell membranes (lanes 12 and 13). 18 K, 18,000dalton protein.

C T G C A G C A A A A A T T C C T T A T G T A T T T G G A A A T A T G G G T A T T G C C G T T C I T T C G A C T C C T C
10 20 30 40 50 60 A A G G G G T T T T A G A A G G C T C T G T A G C A A G G G C C T A A G A A T G T T G C G C G A A T G C T T G T T
70 80 90 100 110 120 MET SER ARG LYS ALA ARG ASP PRO
T G G T T T G G T A G C A A A T T A A A A G A T T <u>I A G G A</u> ]C G G T A A C G A A T G T C T C G T A A A G C T C G A G A C C 130 140 150 160 170 180
ILE VAL LEU PRO CLN CLV VAL GLU VAL SER ILE OLN A5N A5P GLU ILE SER VAL LYS GLY CTATTGTGCTTCCTCAAGGCGTAGAGGTCTCTAATTGAAAGGAAATCTCAGTAAAG 1800 200 200 220
PROLYS GLY SER LEU THE GLN VAL LEU ALA LYS GLU VAL GLU ILE ALA VAL LYS GLY ASN GTCCTAAAGGGTCTTTGACGCAGGTATTGGCTAAAGAAGTGAGATTGCCGTTAAAGGTA 250 280 290 300
GLU VAL PHE VAL ALA PRO ALA ALA HIS VAL VAL ASP ARG PRO GLY ARG MET GLN GLY LEU A TGA GGTGTTTGTTGCTCCTGCGGCTCA CGTTGTA GA CA CGA CCTGGTCGTA TGCA A GG GC 310 320 340
TYR TRP ALA LEU ILE ALA ASH MET VAL LYS GLY VAL HIS THR GLY PHF GLU LYS ARG LEU TTTATTGGGCCTTAATAGCAAATATGGTCAAAGGTGTCCATACTGGATTTGAGAAGCGT 370 - 410 - 420
CALU MET LLE MET VAL ALA SER PRO 1HBP CALY LYS ALA SER ALA HIS LEU HIS SER VAL LEU TAGAAATGATCATGGTAGCITCCCAACGGGGAAAGCTTCCGCTCATCTGCATAGCGTAA 430 450 460 460 470 480
THRE SER CUN CLY ILE VAL CLY ASP SER VAL CLU VAL VAL THRE ILE HIS LYS PHE LEU LYS TAACGTCTCAGGGAATCGTTGGAGATCGTAGAAGTCGTTACGATTCATAAATTTTTAA 490 500 510 520 520 540
ASP MET ARG ARG CLY ARG SER PRO VAL ASP HIS PRO ILE ARG ALA SER TYR ARG LYS ARG A A G A T A T G C G T A G A G G G C G C T C C C C I G T A G A T C C C C A A T A C G A G C A T C C T A T C G T A G A SEO SEO SEO SEO SEO SEO SEO SEO SEO SEO
TRP MET TVR SER CVS CVS SER ARG ARG ARG ARG GLU ASN LEU **** GGTGGATGGATCGTGCTGTAGTAGTAGTAGAAGAGAAAAACCTTTGAGCGCAGACA 610 620 630 640 650 660
TRP   MET   TVR   SER   CVS   SER   ARG   ARG   ARG   GLU   ASN   LEU   ****   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G   G
TAP   NET   TYR   SER   CYS   SER   AG   AG   ARG   ARG   ARG   ARG   CU   ASN   LEU   ***     G G T G G A T G T A T T C G T G C T G T A G T A G T A G A A G A G A A G A G
TAP   NET   TYR   SER   CYS   SER   AG   AG   ARG   ARG   ARG   CLU   ASN   LEU   ****     G G G G A T G T A T T C G T G C T G T A G T A G T A G A A G A G A A G A A G A C T T T G A G C T A G A G A C A G   640   650   660   660   660   660   660   660   660   660   660   660   660   660   660   670   660   670   660   670   660   670   660   670   670   720   720   720   720   720   720   720   720   720   720   720   720   720   720   720   720   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800   800
TBP NET TYR SER CYS SER A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A A A C C T T G A G C T T G G G G A G A G A A A C C T T G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G
TBP NET TYR SER CYS SER A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A A A C C T T T A G A G A G A A A C C T T T G A G A C A A A A C C T T T G A G A C A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
TBP NET TYR SER CYS SER A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A G A A A C C T T T G G G A G A G A G A A A A A C C T T G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G <t< td=""></t<>
TBP NET TYR SER CYS SER A G A G A G A G A G A G A G A G A G A G A G A G A G A G A A A C C T T G A G C T C T T G A G A G A G A A A A A C C T T T G G C A A C A G A A A A A C C T T T G G G G T T A G A A A C C T T T G G G G T T T T T T G G G G T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T T
TAP NOT TAP A G A A G A A G A A G A A G A A G A A G A A G A A G A A G A A G A A C C T T T G G G G A T T T T G G G G G A T G T A G A G
TBP NET TYR SEC TSC CSC SEC AGG ARG ARG ARG ARG ARG ARG AC CT TSC SEC
TBP NET TW SEC TSC SEC SE
TBP NO <t< td=""></t<>
TBP NGC 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 1
TBP NGC 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 1
TBP NO 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 10
TBP NGC TGC T
TBP NGC 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 1000 100 100

FIG. 3. Complete nucleotide sequence of the 1,658-bp chlamydial DNA encoding chlanectin (pCT161/18). The complete nucleotide and predicted amino acid sequence of chlanectin DNA. The proposed translation site is the amino acid Met at nucleotides 159 to 161. The boxed sequence is the postulated Shine-Dalgarno region. The numbers below each line refer to nucleotide positions.

EB binding proteins (lane 7) but not with any polypeptides from the host E. coli strain harboring pUC8 (lane 6). Thus, pCT161/18 directs the production of a protein in E. coli which possesses the same eucaryotic cell-binding determinants as the native chlamydial binding protein. The presence of this overproduced protein on gels could be visualized with Coomassie blue stain (lane 2) but was absent with pUC8derived products (lane 3). The chimeric plasmid pCT161/18 was transformed into the E. coli minicell strain P678-54. The recombinant plasmid encodes a protein of  $M_r$  approximately 18,000, which was absent from minicells containing pUC8 (Fig. 2C). The recombinant gene product was also identified by using an E. coli-derived coupled transcription-translation system. Figure 2D shows the synthesis of an  $18,000-M_r$ polypeptide using pCT161/18, compared with pUC8, with which no similar polypeptide was visualized. Expression of chlanectin relative to that of  $\beta$ -lactamase was greater in an in vitro system (Fig. 2D) than in in vivo systems (Fig. 2A and C). Unlabeled products synthesized by the in vitro transcrip-tion-translation system bound to <sup>125</sup>I-labeled HeLa cells after resolution on SDS-PAGE and subsequent ligand blotting (Fig. 2E).

Neutralization of chlamydial infectivity by antisera. Synthesis of chlanectin by cloned and subcloned DNA fragments was examined by immunoblot assay using hyperimmune rabbit antiserum to the cloned 18,000-dalton protein (Fig. 1). This antiserum was also used in neutralization assays. An

inoculum of  $3.5 \times 10^4$  inclusion-forming units produced approximately 50% infection of HeLa cell monolayers. Preincubation of EB with antisera before inoculation was approximately 90% inhibitory, compared with 80% with antibodies to the native 18,000-dalton protein of EB. Normal rabbit serum also exerted a slight inhibitory effect (approximately 10%). Antiserum raised to whole EB react only weakly to the recombinant 18,000-dalton protein in immunoblots, compared with a strong reaction with native EB proteins as reported earlier (28).

Cellular localization of chlanectin. The cellular localization of chlanectin was examined in *E. coli* harboring pCT161/18 by the immunoblot technique using polyclonal antibodies to recombinant chlanectin, as well as by its binding characteristics to  $^{125}$ I-labeled HeLa membranes on nitrocellulose blots. *E. coli* cells were fractionated as described above and separated on 12.5% polyacrylamide gels. This analysis revealed that most of the chlanectin was in the outer membrane fraction, as is the case with the native *C. trachomatis* binding protein (28).

Nucleotide sequence. The entire 1,658-bp PstI fragment of plasmid pCT161/18 was sequenced in both directions by using the dideoxy method. The DNA sequence of the chlanectin gene and the deduced primary structure of the protein encoded by this gene are shown in Fig. 3. The complete *C. trachomatis* 18,000-dalton cell-binding protein was located on this 1,658-bp fragment. The proposed trans-

lation initiation site is the amino acid methionine at nucleotides 159 to 161. The untranslated sequence 10 positions upstream of the ATG start codon containing a sequence AGGA represents a putative ribosome-binding site (20). Stop codons were found in all three reading frames preceding the start codon. No other significant open reading frames were found on this fragment. Codon usage of the chlanectin gene is strongly reflective of the low G+C content of C. trachomatis DNA (44%), with almost 70% of codons favoring U or A in position 3.

Amino acid sequence. The open reading frame of the recombinant gene was translated and, beginning at the initial methionine, comprises 162 amino acids, specifying a polypeptide of  $M_r$  18,314. The derived chlanectin sequence contains highly charged residues toward the carboxy-terminal end. The presence of two short clusters of positively charged arginines—one between amino acids 146 and 148 and other between amino acids 156 and 159—is striking. The sequence contains two cysteine residues that are adjacent to each other. There is a short hydrophobic region toward the amino-terminal end which is preceded by two positively charged amino acids. The protein has a homogeneous distribution of charged, polar, and hydrophobic residues without clustering of hydrophobic groups, although chlanectin exhibits 43% hydrophobicity overall.

**Promoter studies.** Sequences similar to the *E. coli* consensus promoter were not found upstream of the Shine-Dalgarno region. Sequence data demonstrated that the chlanectin gene in pCT161/18 is located in the same orientation as the *lacZ* promoter of pUC8. To test the dependence of the chlamydial gene on the *lacZ* promoter, the 650-bp *PstI-XbaI* (nucleotides 1 to 650) fragment was subcloned into the *PstI-XbaI* sites of pUC118 and pUC119. Expression of a biologically active gene product based on its binding characteristics to <sup>125</sup>I-labeled HeLa membranes was obtained only with pUC119, where the open reading frame is in the same orientation as the *lac* promoter (Fig. 1).

### DISCUSSION

The data presented here indicate that we isolated an *E.* coli recombinant that expresses an 18,000-dalton *C. trachomatis* binding protein identical in size and activity to a native *C. trachomatis* polypeptide. The deduced molecular size of the recombinant chlanectin (18,314) agrees closely with the published value of 18,000 (5, 28) for the native protein. <sup>125</sup>I-labeled HeLa cells react with this recombinant polypeptide as well as they react with native EB binding proteins, reported earlier (5, 26, 28).

The ability of the cloned pCT161/18 plasmid to direct the synthesis of chlamydial protein was examined in *E. coli* minicells, as well as in an S-30 in vitro transcription-translation system. Polypeptides with similar molecular sizes were expressed in both systems, although the ratio of 18,000-dalton protein synthesis relative to that of  $\beta$ -lacta-mase was significantly higher in the in vitro system. Low levels of expression by minicells may be due to regulatory functions which are inoperable in an in vitro system (8). Further, the ability of in vitro products to bind to <sup>125</sup>I-labeled HeLa cells may suggest either that the protein is fully active with normal structure and function, as has been observed with cloacin DF13 (8), or, if final conformation is not achieved, that the overall conformation of the protein is not crucial for eliciting its cell-binding properties.

The cellular localization of chlanectin on the outer membrane of *E. coli* and the location of the native binding protein on the outer membrane of C. trachomatis suggest the presence of a leader peptide (21). Most bacterial proteins inserted into the outer membrane have an amino-terminal extension with one or more basic residues, including one Arg or Lys (13). The presence of a hydrophobic region consisting of five amino acids preceded by a charged amino acid may represent a signal sequence although the characteristic hydrophobic core of 12 to 20 amino acids representing a typical signal sequence is lacking.

Antibodies directed against the recombinant 18,000-dalton protein exert a neutralizing effect in HeLa cell infection assays, presumably by binding to chlanectin, which then becomes unavailable for subsequent attachment to host cells. These findings agree with earlier results with polyclonal antisera to native chlamydial 18,000- and 31,000-dalton polypeptides (28). The weak cross-reactivity of recombinant chlanectin with polyclonal antibodies to whole EB may represent posttranslational modification in chlamydiae or alternate mechanisms for assembly of protein which are lacking in *E. coli*.

The entire PstI fragment located on pCT161/18 was sequenced in both directions. The sequence revealed an open reading frame beginning at nucleotide 159 and extending up to nucleotide 644, comprising 162 amino acids. A sequence, AGGA, 10 positions upstream probably represents the ribosome-binding site (20). However, no consensus promoter sequences (-10 and -35) similar to E. coli promoters were detected. The possibilities that the promoter is present far upstream of the start codon or that chlamydial promoters are unlike those of E. coli cannot be addressed without additional molecular and biochemical characterization. Recently, major outer membrane protein genes sequenced from both servvars  $L_2$  and  $L_1$  also failed to reveal consensus promoter sequences comparable to E. coli promoters (15a. 22). However, although chlanectin uses the vector promoter for its expression, a fused gene product was not produced, since stop codons were found in all three reading frames preceding the start codon. The absence of a functional promoter on the 650-bp PstI-XbaI fragment was confirmed by the lack of expression of this insert in pUC118, where the lacZ promoter is oriented opposite to the inserted gene.

Hydrophilicity has been used extensively to predict the antigenicity of various proteins and peptides. On the basis of the hydrophilic-algorithm model of Parker et al. (15), it was found that amino acids 20 to 30 and 125 to 162 are mostly surface accessible and hence antigenic. Whether any cell attachment role is played by these residues remains undetermined. Moreover, the presence of a Gly-Asp-Ser-Val amino acid residue between amino acids 115 and 118 and Arg-Gly-Arg-Ser-Pro residues between 132 and 136, where four of five amino acids resemble the cell attachment domain of fibronectin (16), is notable. However, a computer-assisted comparison between chlanectin and E. coli revealed no significant homology. Interestingly, the percent hydrophobicity and homogeneous distribution of polar, hydrophobic, and charged residues on chlanectin bear some similarities to those of the serovar  $L_2$  major outer membrane protein as reported recently by Stephens et al. (22).

The cloning and expression of the gene for the functional part of the chlamydial binding protein chlanectin in E. coli will facilitate further structural analysis. We should have sufficient quantities of the 18,000-dalton protein to study its structure-function properties, and especially its eucaryotic cell-binding ability, in greater detail. The deduced amino acid sequence of chlanectin, derived from nucleotide sequencing, will assist in identification of the domain respon-

sible for cell binding. This domain may be a candidate for future chlamydial vaccine development.

#### ACKNOWLEDGMENTS

This work was supported by the Canadian Medical Research Council (grant MA7951) and the Natural Sciences and Engineering Research Council (grant A6484). R.K. is a fellow and W.M.W. is a scholar of the Alberta Heritage Foundation for Medical Research.

We thank Renate Meuser, Karen Chong, Matthew Duncan, and Jamie Doran for assistance.

#### LITERATURE CITED

- 1. Adler, H. I., W. D. Fisher, A. Cohen, and A. A. Hardigree. 1967. Miniature *Escherichia coli* cells deficient in DNA. Proc. Natl. Acad. Sci. USA 57:321-326.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513–1523.
- 3. Dougan, G., G. Dowd, and M. Kehoe. 1983. Organization of K88ac-encoded polypeptides in the *Escherichia coli* cell envelope: use of minicells and outer membrane protein mutants for studying assembly by pili. J. Bacteriol. 153:364–370.
- Gill, R. E., F. Heffron, and S. Falkow. 1979. Identification of the protein encoded by the transposable element *Tn3* which is required for its transposition. Nature (London) 282:797-801.
- Hackstadt, T. 1986. Identification and properties of chlamydial polypeptides that bind eucaryotic cell surface components. J. Bacteriol. 165:13-20.
- 6. Holmes, K. K. 1981. Chlamydia epidemic. J. Am. Med. Assoc. 245:1718–1723.
- 7. Kaul, R., and W. M. Wenman. 1985. Cloning and expression in *Escherichia coli* of a species specific *Chlamydia trachomatis* outer membrane antigen. FEMS Microbiol. Lett. 27:7–12.
- Konings, R. N. H., P. M. Andreoli, E. Veltkamp, and H. J. J. Nijkamp. 1976. Clo DF13 plasmid deoxyribonucleic aciddirected in vitro synthesis of biologically active cloacin DF13 and Clo DF13 immunity protein. J. Bacteriol. 126:861–868.
- Kuo, C.-C., S.-P. Wang, and J. T. Grayston. 1977. Growth of trachoma organisms in HeLa 229 cell culture, p. 328–336. *In D.* Hobson and K. K. Holmes (ed.), Nongonococcal urethritis and related infections. American Society for Microbiology, Washington, D.C.
- 10. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.
- 11. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 12. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20-78.
- Michaelis, S., and J. Beckwith. 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. Annu. Rev. Microbiol. 36:435–465.
- 14. Morrison, M. 1980. Lactoperoxidase catalyzed iodination as a

tool for investigation of proteins. Methods Enzymol. 70: 214-220.

- 15. Parker, J. M. R., D. Guo, and R. S. Hodges. 1986. New hydrophilicity scale derived from high performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray derived accessible sites. Biochemistry 25:5425-5432.
- 15a. Pickett, M. A., M. E. Ward, and I. N. Clarke. 1987. Complete nucleotide sequence of the major outer membrane protein gene from *Chlamydia trachomatis* serovar L. FEMS Microbiol. Lett. 42:185–190.
- Pierschbacher, M. D., and E. Ruoslahti. 1984. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Nature (London) 309:30-33.
- Radloff, R., W. Bauer, and J. Vinograd. 1967. A dye buoyant method for the detection and isolation of closed circular duplex DNA in HeLa cells. Proc. Natl. Acad. Sci. USA 57:1514–1521.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 19. Schachter, J., and M. Grossman. 1981. Chlamydial infections. Annu. Rev. Med. 32:45-61.
- Shine, J., and L. Dalgarno. 1974. The 3' terminal sequence of Escherichia coli 16s ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA 71:1342-1346.
- Silhavy, T. J., S. A. Benson, and S. D. Emr. 1983. Mechanism of protein localization. Microbiol. Rev. 47:313-344.
- Stephens, R. S., G. Mullenbach, R. Sanchez-Pescador, and N. Agabian. 1986. Sequence analysis of the major outer membrane protein gene from *Chlamydia trachomatis* serovar L<sub>2</sub>. J. Bacteriol. 168:1277-1282.
- Towbin, H., T. Staehlin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350-4354.
- Vieira, J., and J. Messing. 1982. The pUC plasmids: an M13 mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259–268.
- Ward, M. E. 1983. Chlamydial classification, development and structure. Br. Med. Bull. 39:109–115.
- 26. Wenman, W. M., R. Kaul, and R. U. Meuser. 1986. Eukaryotic cell binding proteins of *Chlamydia trachomatis* and *Chlamydia psittaci*, p. 15–18. *In* D. Oriel, G. Ridgway, J. Schachter, D. Taylor Robinson, and M. Ward (ed.), Chlamydial infections. Cambridge University Press, Cambridge.
- Wenman, W. M., and M. A. Lovett. 1982. Expression in E. coli of Chlamydia trachomatis antigen recognized during human infection. Nature (London) 296:68-70.
- Wenman, W. M., and R. U. Meuser. 1986. Chlamydia trachomatis elementary bodies possess proteins which bind to eucaryotic cell membranes. J. Bacteriol. 165:602-607.
- Yamato, I., Y. Anraku, and K. Hirosawa. 1975. Cytoplasmic membrane vesicles of *Escherichia coli*. I. A simple method for preparing the cytoplasmic and outer membranes. J. Biochem. 77:705-718.