

# Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*

Robert J. Belland<sup>\*††</sup>, Guangming Zhong<sup>§</sup>, Deborah D. Crane<sup>¶</sup>, Daniel Hogan<sup>\*</sup>, Daniel Sturdevant<sup>\*</sup>, Jyotika Sharma<sup>§</sup>, Wandy L. Beatty<sup>||</sup>, and Harlan D. Caldwell<sup>\*¶</sup>

Laboratories of <sup>\*</sup>Human Bacterial Pathogenesis and <sup>¶</sup>Intracellular Parasites, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Disease, National Institutes of Health, Hamilton, MT 59840; <sup>§</sup>Department of Microbiology, University of Texas Health Science Center, San Antonio, TX 78229; and <sup>||</sup>Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110

Edited by Stanley Falkow, Stanford University, Stanford, CA, and approved April 29, 2003 (received for review February 26, 2003)

*Chlamydia trachomatis* is one of the most common bacterial pathogens and is the etiological agent of debilitating sexually transmitted and ocular diseases in humans. The organism is an obligate intracellular prokaryote characterized by a highly specialized biphasic developmental cycle. We have performed genomic transcriptional analysis of the chlamydial developmental cycle. This approach has led to the identification of a small subset of genes that control the primary (immediate-early genes) and secondary (late genes) differentiation stages of the cycle. Immediate-early gene products initiate bacterial metabolism and potentially modify the bacterial phagosome to escape fusion with lysosomes. One immediate early gene (CT147) is a homolog of the human early endosomal antigen-1 that is localized to the chlamydial phagosome; suggesting a functional role for CT147 in establishing the parasitophorous vacuole in a nonfusogenic pathway. Late gene products terminate bacterial cell division and constitute structural components and remodeling activities involved in the formation of the highly disulfide cross-linked outer-membrane complex that functions in attachment and invasion of new host cells. Many of the genes expressed during the immediate-early and late differentiation stages are *Chlamydia*-specific and have evolutionary origins in eukaryotic lineages.

The *Chlamydia trachomatis* bacterium is an obligate intracellular pathogen of humans that primarily infects columnar epithelial cells of the ocular and genital mucosae. Chlamydial infections of the eye and genital tract have a significant impact on human health worldwide, causing trachoma, the leading cause of preventable blindness, and sexually transmitted diseases (STD) that include pelvic inflammatory disease and tubal factor infertility (1, 2). Chlamydial STDs are also risk factors in cervical squamous cell carcinoma (3) and HIV infection (4, 5).

*C. trachomatis* has a small genome of  $\approx 1$  Mb encoding 893 chromosomal and 8 plasmid ORFs that share significant homology in both gene structure and order among strains that infect human and animal hosts (6, 7). Two distinguishing characteristics of this pathogen are its developmental cycle and predilection for causing persistent infections (8). The developmental cycle consists of infectious and noninfectious stages that exhibit unique morphological, biochemical, and biological properties. The infectious elementary body (EB) is a metabolically inactive particle with a rigid, disulfide cross-linked outer membrane (OM) (9–12) that enables the EB to attach to and enter host cells (13–15). After host cell entry, the EB is localized to a phagosome, and the primary differentiation process is initiated. This developmental process involves the commencement of bacterial metabolism and the conversion of the EB to the intracellular replicating form of the organism, termed the reticulate body (RB).

At the very early stage of infection (1–3 h) the parasite exerts profound effects on the host. Through an unknown mechanism, dependent on both bacterial transcription and translation (16), chlamydiae modify the properties of the phagosome and prevent its entry into the lysosomal pathway (17–19). Many obligate and facultative intracellular pathogens use this approach to avoid intracellular killing, by using different means to interfere with

cellular trafficking (20, 21). The early chlamydial inclusion is located in close apposition to, but does not fuse with, endocytic vesicles (18, 22). In a secondary process requiring bacterial transcription and translation, the modified phagosome is transported to the peri-Golgi region, where it interacts with a subset of exocytic vesicles, suggesting both passive and active mechanisms are necessary for avoidance of lysosomal fusion (23). This unique parasite strategy provides a continuously protected intracellular niche in which chlamydiae then replicate. The RB OM lacks the highly cross-linked OM structure, the chromosome is relaxed and transcriptionally active, and the bacterium multiplies by binary fission for a period of 24–36 h. After multiple rounds of replication the RB undergoes a secondary differentiation process back to an infectious EB. Expression of several late genes *hctA*, *hctB*, and *ompB*, *ompC*; that encode the histone-like proteins (24, 25), and the 60 and 12 kDa OM cysteine rich proteins (26), respectively, is associated with differentiation from RB to EB. At this late stage in development (40–60 h) the host cell lyses and releases mature EBs that then reinfect neighboring host cells. In this study we have used microarrays to analyze the temporal expression of chlamydial genes throughout the developmental cycle. These studies have provided powerful new findings about chlamydial gene products that control the differentiation stages of the developmental cycle and determine the nature of the host–pathogen interaction.

## Materials and Methods

**Chlamydia and Cell Culture.** *C. trachomatis* serovar D (strain UW-3/Cx) was grown in HeLa 229 cells cultivated at 37°C with 5% CO<sub>2</sub> in high glucose-containing DMEM (Cellgro, Mediatech, Washington, DC) supplemented with 10% heat-inactivated FBS. EBs were purified on density gradients of RenoCal-76 (Bracco Diagnostics) as described (9).

**RNA Purification, Labeling, and Hybridization.** Total RNA was purified from infected HeLa 229 cultures ( $8 \times 10^7$  cells cultivated as monolayers in 150-cm<sup>2</sup> tissue culture flasks) at various times postinfection (PI). Monolayers were pretreated with DEAE-Dextran (30  $\mu$ g/ml) and infected at a multiplicity of infection (MOI) of 1 or 100. Cultures inoculated at an MOI of 100 were used to prepare RNA for early time points (1 and 3 h PI), whereas cultures infected at a MOI of 1 were used for the 8-, 16-, 24-, and 40-h time points. PI timing began immediately after the addition of infectious EBs to the monolayers. At the designated times PI, the culture media was discarded, the cells were lysed in 10 ml of TRIzol (Invitrogen), and total RNA was isolated. Polyadenylated host mRNA was removed by using oligo(dT) columns (Oligotex, Qiagen, Valencia, CA). Bacterial mRNA was repurified (QIAquick

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

Abbreviations: STD, sexually transmitted disease; EB, elementary body; OM, outer membrane; RB, reticulate body; PI, postinfection; MOI, multiplicity of infection.

See commentary on page 8040.

<sup>\*</sup>R.J.B. and H.D.C. contributed equally to this work.

<sup>††</sup>To whom correspondence should be addressed. E-mail: rbelland@niaid.nih.gov.

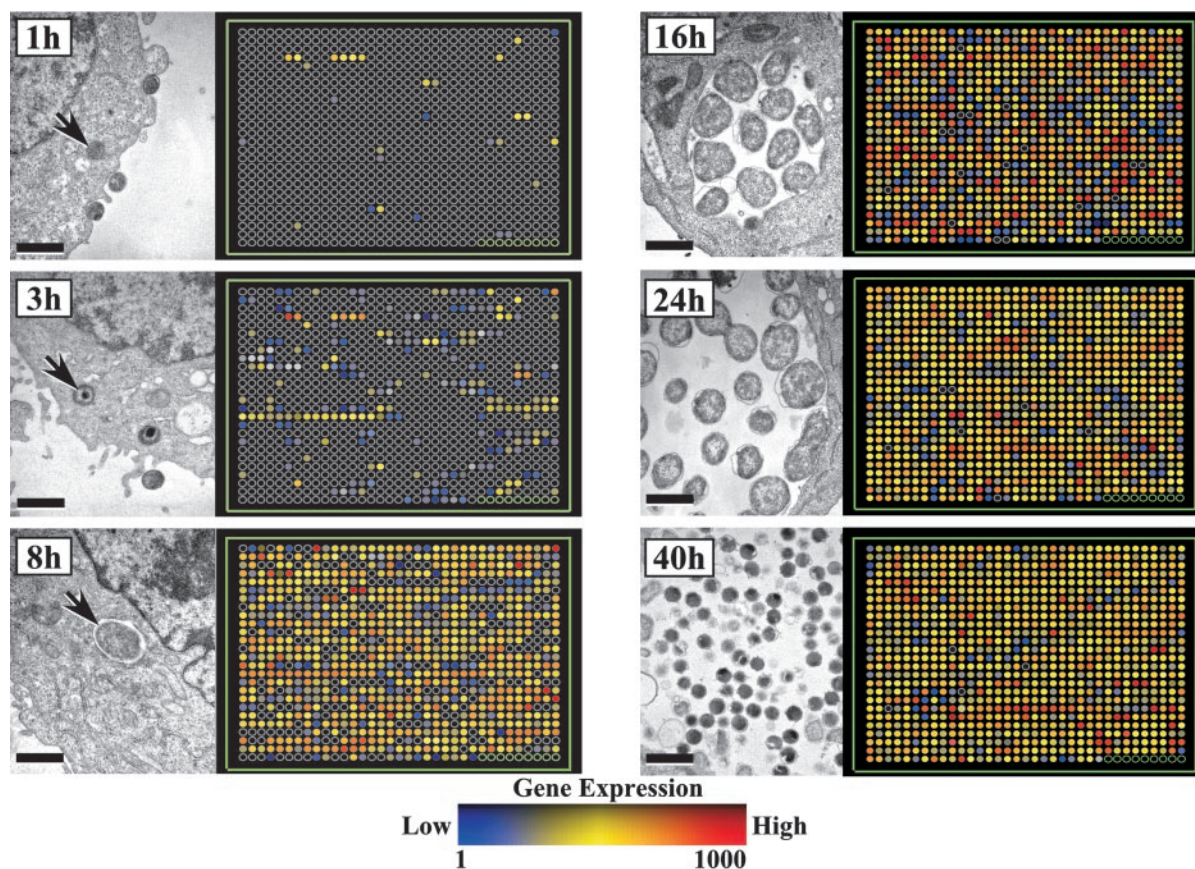
columns, Qiagen) and specifically primed for cDNA synthesis (SuperScript Choice System, Invitrogen) by using a complete set of complementary 3' oligonucleotides (the same 3' oligonucleotides used in the amplification of ORFs for the microarray). This step was essential for generation of bacterial probes because direct labeling of total RNA gave weak and inconsistent results. After cDNA synthesis, residual RNA was removed from the samples by alkaline hydrolysis and the remaining bacterial cDNA was repurified (QIAquick columns) and labeled with either Cy3 or Cy5 (FluoroLink, Amersham Pharmacia) by a random priming procedure (Invitrogen).

Labeled probe samples were cohybridized on microarray slides with fragmented, labeled genomic DNA from *C. trachomatis* serovar D, as described by Talaat *et al.* (27) and Xu *et al.* (28). For a given experiment, samples from each time point of the cycle were cohybridized with an aliquot of a single, labeled genomic DNA sample (gDNA), which provided an internal control for each microarray spot. The use of an internal control ("RNA versus gDNA") was necessary as the number of genes expressed at different time points varied substantially and no single gene was constitutively expressed. This method of normalization was particularly important in comparing early time points to mid and late-cycle time points. Fluorescence intensity data were determined by using a ScanArray 5000 scanner (Packard) and values at each time point were then normalized to the internal control (cDNA versus gDNA), allowing for independent measurements of the gene

expression levels throughout the cycle and eliminating the need for pairwise comparisons (as discussed in ref. 27 and 29). Analysis of microarray data and presentation of the data in an "array layout" format were done by using the GENESPRING software package from Silicon Genetics (Redwood City, CA). The developmental cycle experiment was done three times, and the hybridizations were done in duplicate. Thus, 18 measurements were done for each gene, as each gene was spotted on the microarray in triplicate.

**Quantitative RT-PCR.** Primer/probe sets were designed for selected genes by using the PrimerExpress software (Applied Biosystems). Standard curves were performed for each gene by using purified chromosomal template DNA at concentrations ranging from 10 to 0.001 ng/ml (data not shown). Assays were performed (OneStep RT-PCR System, Applied Biosystems) by using RNA samples used from infections at a MOI of 1 (before cDNA generation) by using bacterial 16S rRNA as controls (7700 Sequence Detector, ABI Prism). Relative copy numbers were determined for each sample by converting the mean critical threshold values to DNA concentrations by using standard curves. DNA concentrations were converted to copy numbers by using the calculated molecular mass of the bacterial genome (6). Copy number values were normalized to the transcript level per bacterial cell by using the 16S rRNA per bacterial cell ratio.

**Electron Microscopy.** *Chlamydia*-infected cells were fixed in 2% paraformaldehyde/2.5% glutaraldehyde (Polysciences, War-

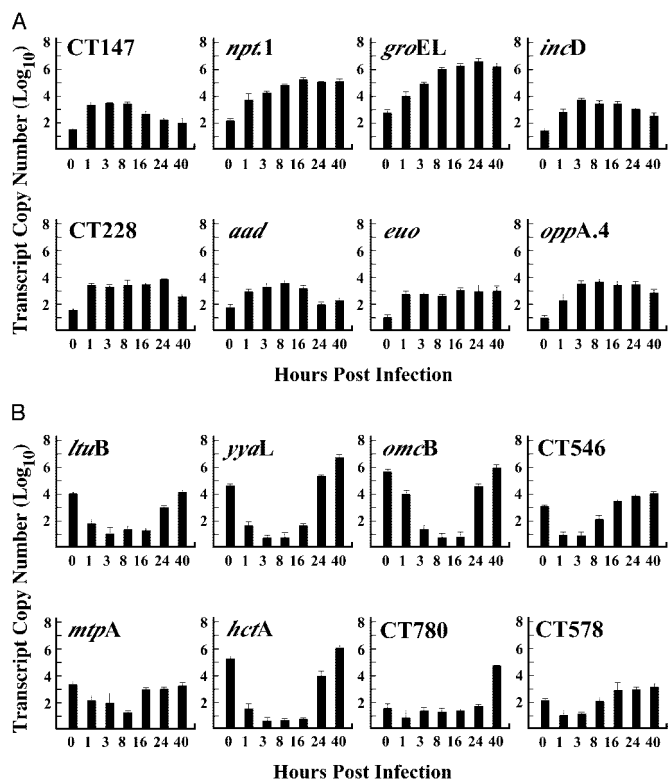


**Fig. 1.** Transcriptional profiling of the developmental cycle of *C. trachomatis* serovar D in HeLa 229 cells with accompanying transmission electron microscopy of bacterial inclusions at 1, 3, 8, 16, 24, and 40 h PI. Results are shown in an array layout format that is a linear representation of the bacterial genome (i.e., row 1, column 1 is ORF CT001; row 1, column 2 is CT002, etc. . .). The data shown are the normalized expression values for three independent experiments done in duplicate. The 1- and 3-h analyses were from infections using a MOI of 100, whereas the 8-, 16-, 24-, and 40-h analyses were from infections using a MOI of 1. Black circles represent genes that are considered transcriptionally inactive at the time point shown. For example, the 29 immediate-early ORFs are shown in the 1-h PI analysis, and their expression level, CT numbers, and chromosomal location are indicated by their position in the array layout. The complete data set is listed in Table 2, which is published as supporting information on the PNAS web site, [www.pnas.org](http://www.pnas.org).

**Table 1. Expression of immediate early and late genes by *C. trachomatis***

Immediate early genes				Late genes			
CT no.	Gene	Function	qRt	CT no.	Gene	Function	qRt
CT035	<i>bp1</i>	Biotin protein-ligase	Yes	CT046	<i>hctB</i>	Histone-like protein 2	ND
CT065	<i>npt1</i>	ADP/ATP translocase	Yes	CT080	<i>ltuB</i>	Late transcription unit protein B	Yes
CT110	<i>groEL</i>	Chaperonin (60 kDa)	Yes	CT214		Conserved hypothetical	ND
CT111	<i>groES</i>	Chaperonin (10 kDa)	ND	CT356	<i>yyaL</i>	Multidomain protein	Yes
CT115	<i>incD</i>	Inclusion membrane protein	Yes	CT443	<i>omcB</i>	Cysteine-rich OMP (60 kDa)	Yes
CT116	<i>incE</i>	Inclusion membrane protein	ND	CT444	<i>omcA</i>	Cysteine-rich OMP (9 kDa)	ND
CT117	<i>incF</i>	Inclusion membrane protein	ND	CT546		Predicted OMP	Yes
CT118	<i>incG</i>	Inclusion membrane protein	ND	CT565		Conserved hypothetical	ND
CT133	<i>rrm</i>	Predicted methylase (rRNA)	Yes	CT576	<i>lcrH.1</i>	Predicted Type III chaperone	Yes
CT147		Predicted EEA-1-like protein	Yes	CT578		Conserved hypothetical	Yes
CT228		Inclusion membrane protein	Yes	CT579		Conserved hypothetical	ND
CT229		Inclusion membrane protein	ND	CT659		Predicted RNA binding protein	ND
CT288		Conserved hypothetical	ND	CT660	<i>gyrA.2</i>	DNA gyrase subunit A protein 2	Yes
CT365		Conserved hypothetical	ND	CT694		Conserved hypothetical	ND
CT375	<i>aad</i>	D-amino acid dehydrogenase	Yes	CT712		Conserved hypothetical	ND
CT376	<i>mdhC</i>	Malate dehydrogenase	ND	CT743	<i>hctA</i>	Histone-like protein 1	Yes
CT446	<i>euo</i>	DNA binding protein	Yes	CT780		Thioredoxin disulfide isomerase	Yes
CT473	<i>yidD</i>	Predicted $\alpha$ -hemolysin	ND	CT783		Thioredoxin disulfide isomerase	ND
CT474		Conserved hypothetical	ND	CT798	<i>glgA</i>	Glycogen synthase	Yes
CT480	<i>oppA.4</i>	Oligopeptide permease	Yes	CT814		Conserved hypothetical	ND
CT495	<i>npt2</i>	Nucleoside phosphate transporter	Yes	CT814.1		Conserved hypothetical	ND
CT529		Conserved hypothetical	ND	CT847		Conserved hypothetical	Yes
CT648		Conserved hypothetical	ND	CT867	<i>mtpA</i>	Membrane thiol protease	Yes
CT734		Conserved hypothetical	ND	CT868	<i>mtpB</i>	Membrane thiol protease	ND
CT735	<i>dagA</i>	D-alanine/glycine permease	ND	CT871	<i>pmpG</i>	Polymorphic membrane protein G	ND
CT774	<i>cysQ</i>	Biphosphate phosphatase	Yes	CT872	<i>pmpH</i>	Polymorphic membrane protein H	ND
CT795		Conserved hypothetical	ND				
CT850		Conserved hypothetical	ND				
CT851	<i>map</i>	Methionine aminopeptidase	ND				

qRT, quantitative RT-PCR; ND, not determined.

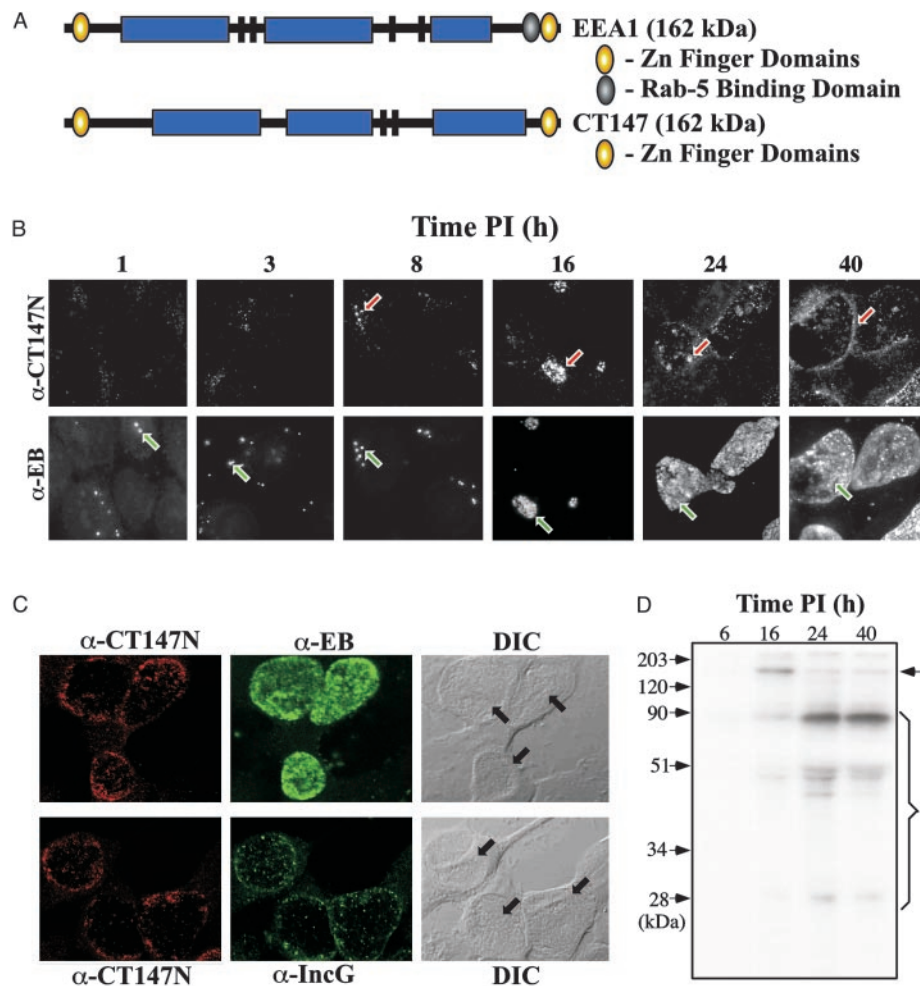


**Fig. 2.** Expression levels of immediate-early and late genes detected by quantitative RT-PCR. Quantitative RT-PCR (qRT) assays were performed by using TaqMan primer/probe sets specific for each gene by using RNA samples purified from infections done at a MOI of 1. The efficiency of amplification was comparable for all primer/probe sets as estimated by standard curves. Transcript levels

increased for each immediate-early gene (8/8) by 1 h PI, demonstrating that new transcription had occurred. In contrast, all late genes analyzed showed decreased transcript levels at 1 h PI, indicating that EBs contained significant levels of mRNA but that these levels decreased during the early part of the developmental cycle. Late gene expression for all eight genes tested was up-regulated between 16 and 40 h PI. RNA samples from purified EBs (0 time) were used to control for “carry over” transcript levels.

**Fluorescence Microscopy.** Immunofluorescent staining and microscopy were carried out as described (30). HeLa cells infected with *C. trachomatis* were processed for antibody staining at various intervals after infection. For temporal expression studies, the processed cell samples were stained with a primary rabbit  $\alpha$ -EB and a primary mouse  $\alpha$ -CT147 antibody, and visualized with secondary goat  $\alpha$ -rabbit IgG conjugated with Cy2 (green) or goat  $\alpha$ -mouse IgG conjugated with Cy3 (red). Nuclei were visualized by staining with Hoechst dye 33258 (blue). Images were acquired by using an AX70 fluorescence microscope (Olympus, Melville, NY) equipped with SIMPLE PCI (Compix, Tualatin, OR) imaging software. For colocalization studies, the  $\alpha$ -CT147 mouse antibody was paired with either the  $\alpha$ -EB or an  $\alpha$ -IncG antibody. After the corresponding secondary antibody staining (green for rabbit IgG and red for mouse IgG), images were acquired by using an Olympus confocal microscope.

increased for each immediate-early gene (8/8) by 1 h PI, demonstrating that new transcription had occurred. In contrast, all late genes analyzed showed decreased transcript levels at 1 h PI, indicating that EBs contained significant levels of mRNA but that these levels decreased during the early part of the developmental cycle. Late gene expression for all eight genes tested was up-regulated between 16 and 40 h PI. RNA samples from purified EBs (0 time) were used to control for “carry over” transcript levels.



**Fig. 3.** Confocal fluorescence microscopy and radioimmuno precipitation of *C. trachomatis* CT147. (A) Schematic comparison of human EEA1 and *C. trachomatis* CT147. The proteins have nearly identical predicted molecular masses ( $\approx 162$  kDa) and share N- and C-terminal Zn finger domains. The EEA1 C-terminal domain is of the C<sub>2</sub>H<sub>2</sub> type and is termed a FYVE domain (39). The central regions of the proteins have extensive coiled-coil domains (shown as blue rectangles) and two (CT147) or more leucine zipper motifs (shown as black bars). CT147 lacks homology with the region of EEA1 that has been shown to bind Rab5 (shown as the blue filled oval). (B) CT147 expression can be seen early in the developmental cycle (i.e., 8 h PI as indicated by red arrows) and continues throughout the cycle. Localization of RBs within the inclusion in infected cells is shown for comparison (i.e., at all times in the cycle, as indicated by green arrow). (C) CT147 is localized to the inclusion membrane as shown by immunofluorescent staining in comparison to RBs. (Upper) Costaining results using  $\alpha$ -CT147 and  $\alpha$ -EB antisera are shown.  $\alpha$ -CT147 staining is predominantly at the periphery of the inclusion, whereas the  $\alpha$ -EB staining is found throughout the inclusion. The inclusions are shown in the differential interference contrast (DIC) microscopic images and are indicated with arrows. (Lower) The results of costaining with  $\alpha$ -CT147 and  $\alpha$ -IncG antisera are shown. Both antisera react with antigens found predominantly at the periphery of the inclusion, indicating that CT147 is localized to the inclusion membrane. (D) CT147 is first detected by RIP at 16 h PI. CT147 is indicated by the arrow (predicted Mr = 162 kDa). At subsequent time points the quantity of the full-length product decreased and the quantity and number of breakdown products increased (indicated by the bracket), indicating increased levels of posttranslational modification.

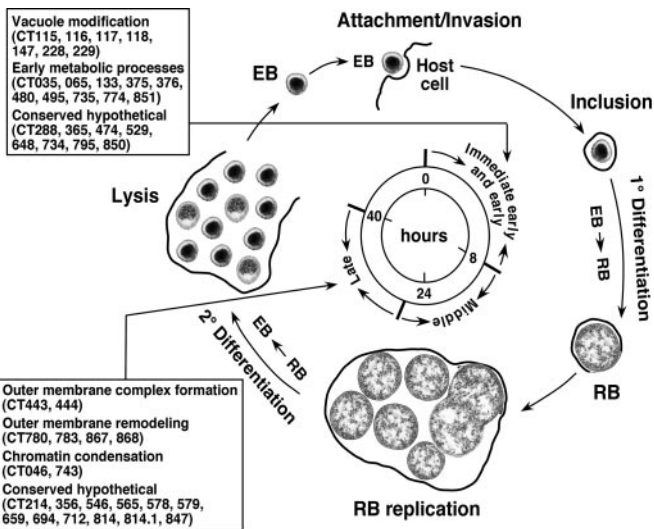
**Immune Precipitation.** The radioimmunoprecipitation assay was carried out as described (31). Briefly, HeLa cells infected with *C. trachomatis* were pulse labeled with [<sup>35</sup>S]methionine/cysteine for 1 h at different time points after infection. The cells were harvested by scraping and lysed by using a Mg<sup>2+</sup>-containing lysis buffer containing 1% Igepal in 25 mM Hepes (pH 7.4). Cell lysates were incubated with protein G-agarose precoated with mouse  $\alpha$ -CT147 antiserum. After washing to remove non-specifically bound lysate components, bound proteins were eluted from the pellets and analyzed by SDS/PAGE followed by autoradiography.

## Results and Discussion

**Gene Expression Profile of the Chlamydial Developmental Cycle.** The genomic transcriptional profiles of *C. trachomatis* serovar D (893 chromosomal and 8 plasmid ORFs) obtained from RNAs isolated at 1, 3, 8, 16, 24, and 40 h PI and accompanying transmission electron micrographs of the chlamydial developmental cycle taken at identical time points are shown in Fig. 1. The arrays show the expression levels of each gene analyzed throughout the growth cycle starting from CT001 (*Top Left*) and then proceeding in a collinear manner to CT893, followed by the 8 plasmid ORFs, and 7 negative controls (*Bottom Right*). Chlamydiae were transcriptionally active as early as 1 h PI, with a total of 29 genes (3.2% of the genome) exhibiting moderate to strong fluorescent signals by microarray analyses. By 3 h PI, a time corresponding to microscopic observation of EB conversion to RBs by chromatin decondensation (arrows at 1, 3, and 8 h in Fig. 1), an additional 200 genes were transcriptionally active. We have termed this transcriptome subset as early

genes (1–3 h PI) and the 1 h subset as immediate-early genes. A remarkable increase in transcriptional activity was evident by 8 h PI when the transition of EB to RB is complete. This intense level of transcriptional activity was maintained throughout the mid-portion of the developmental cycle (16–24 h), a time period when RBs were actively replicating by binary fission. At this mid-point in the cycle, gene transcription involves virtually every chromosomal and plasmid ORF, indicating that the small chlamydial genome has very little facultative capacity. A subset of 28 genes (3.1% of the genome) were specifically expressed only during the late stages of the growth cycle (40 h) when the majority of RBs had differentiated to infectious EBs. A complete list of genes transcribed over the entire chlamydial developmental cycle is given in Table 2, which is published as supporting information on the PNAS web site.

We focused on the immediate-early and late genes because they represent a subset of genes that are important in understanding key events in the subversion of host response and the differentiation processes that control the developmental cycle. Immediate-early and late genes are listed in Table 1 according to CT number, gene designation, putative function, and confirmation of gene expression by quantitative RT-PCR (qRT in Table 1). The same RNA samples used for microarray analysis were also used for Taqman quantitative RT-PCR assays, with the exception of the 1- and 3-h PI times (MOIs of 1 were used instead of 100). Taqman primer-probe sets specific for each target gene were used to establish standard curves to ascertain that the efficiency of amplification was comparable for all primer-probe sets (results not shown). The transcript copy numbers for eight immediate-early and eight late genes, normalized to 16s RNA and assayed throughout the chlamydial developmental



**Fig. 4.** Schematic representation of the *C. trachomatis* developmental cycle. The infectious EB, after attachment and internalization by the eukaryotic cell, undergoes a 1° differentiation to the metabolically active RB. This phase of the developmental cycle is characterized by expression of immediate-early and early genes and involves cellular processes associated with vacuole modification and the initiation of bacterial metabolism (as indicated in the boxed region). Bacterial transcription and translation follow decondensation of the chromosome, leading to bacterial replication by binary fission (between  $\approx 16$  and 40 h PI). The majority of genes are expressed during this highly active stage of the cycle. Late in the developmental cycle the organism undergoes a 2° differentiation from RB to EB. This phase of the cycle is associated with expression of late genes and their corresponding products that give rise to the highly disulfide cross-linked OM complex (remodeling disulfide isomerases and proteases) and chromosome condensation (the histone-like proteins) as indicated in the boxed region. The eukaryotic cell then lyses, releasing infectious EBs to infect neighboring cells and continue the infectious process.

cycle, are shown in Fig. 2. The copy number for each of the eight immediate-early genes (Fig. 2A) increased 10- to 100-fold by 1 h PI confirming that new transcription had indeed occurred. Importantly, these findings confirm the accuracy and reliability of the results obtained by genomic microarray analysis shown in Fig. 1. The copy number of all eight immediate-early genes increased during the early phase of the growth cycle and either increased slightly or remained constant throughout the remaining stages of the cycle. A similar Taqman analysis of late genes is shown in Fig. 2B. There was a significant level of transcript detected for all eight late genes at the zero time point; values derived from RNAs obtained from EBs before infection. However, in contrast to the profiles of early genes, late gene transcripts decreased to near base-line levels at 1, 3, and 8 h PI. A dramatic increase in late gene transcripts was seen at late time points PI (24 and 40 h). This increase ranged from a low of  $10^2$  copies for *mtpA* and CT578 to a high of  $10^6$  for *yyaL*, *omcB*, and *hctA*. We believe that the transcripts present at time 0 represent mRNA that is “carried over” to the infectious EB from RBs during the late stages of secondary differentiation. The mRNA in EBs does not appear to be translationally active. Hatch *et al.* (32) and Douglas and Hatch (33) have shown that *omcB*-specific mRNA, that is carried over in EBs, does not result in synthesis of the OmcB protein early in infection. Thus, this “carry over” of mRNA is perhaps not unexpected, and may in fact be characteristic of late genes; particularly those that are present at very high copy numbers late in the developmental cycle. For this reason we have omitted “carry over” transcripts from the immediate-early genes, to focus our analysis on genes that are newly transcribed after infection.

**Expression of Immediate-Early Genes.** Twenty-nine immediate-early genes were identified by microarray analysis and eight were con-

firmed by qRT-PCR analysis. Seven of 29 genes had been described as early genes in other reports. These are a family of inclusion membrane proteins genes (*incD*, E, F, and G; ref. 34), the early upstream ORF gene (*euo*, ref. 35), and the chaperonin genes *groEL* and *groES* (34). This finding supported the validity of the microarray approach in the identification of new immediate-early genes in that it confirmed previous reports using different experimental approaches. Newly identified immediate-early genes encode proteins that function in the translocation of metabolites into the bacterial cell (e.g., ADP/ATP translocase, nucleotide phosphate transporter, oligopeptide permease, and a D-alanine/glycine permease) and enzymes involved in metabolite interconversions (e.g., malate dehydrogenase, nucleoside phosphohydrolase, and a methionine aminopeptidase). Two additional predicted inclusion-membrane protein genes, CT228 and CT229 (34), were also transcribed at 1 h PI. The number of Inc-like proteins expressed argues strongly for an important functional role in modifying the inclusion membrane to support chlamydial intracellular survival and growth. Eight immediate-early genes encoded conserved (among chlamydial species) hypothetical proteins of unknown function.

A novel predicted protein described here is encoded by CT147. The protein has a significant level of homology to the 162-kDa human early endosomal antigen 1 (EEA1) that is involved in endosomal trafficking and fusion in mammalian cells (36–38). CT147 and EEA1 share structural properties characterized by a similar mass (162 kDa) with coiled-coil domains, N- and C-terminal Zn-finger motifs, and two or more leucine zipper domains (Fig. 3A). The EEA-1 C-terminal end contains a Zn-finger domain referred to as a FYVE domain (39) that interacts specifically with phosphatidylinositol 3-phosphate (PtdIns-3-P) and a Rab5 GTPase binding domain that is independent of the FYVE motif (40). EEA1 mediated endosome tethering depends on the binding of EEA1 to the PtdIns-3-P domain, whereas the regulation of endosomal fusion depends on its interaction with Rab5 (38, 40).

CT147 lacks the region of homology to EEA1 that binds to Rab5 but has maintained the C-terminal Zn-finger, suggesting a modified function for the protein that could theoretically tether endosomes but circumvent endosomal fusion, a novel strategy for chlamydial endosomes to avoid fusion with lysosomes. Such a function for CT147 is consistent with the properties of early chlamydial inclusions that associate, but do not fuse with endocytic vesicles (18). Clearly identifiable CT147 orthologs with *E* values better than  $10^{-50}$  have been found in each of the chlamydial species sequenced to date (6, 7), consistent with their shared ability to avoid lysosomal fusion.

We hypothesized that if CT147 interacted with host vesicular organelles and its function was important to the pathogens intracellular survival the protein should be (i) expressed early during infection and (ii) localized to the inclusion membrane or host cytosol. We tested this hypothesis by analyzing the temporal expression of CT147 by immunofluorescence and radioimmune precipitation of  $^{35}\text{S}$  amino acid intrinsically radiolabeled organisms. Immunofluorescence staining by using hyperimmune  $\alpha$ -CT147 antiserum showed specific staining of chlamydiae by 8 h PI (Fig. 3B, red arrows). CT147 staining intensified as infection progressed and was localized to the periphery of the inclusion. This staining pattern was most easily visualized at the 24- and 40-h PI time points. In contrast,  $\alpha$ -EB serum stained only EB and RB particles that were contained within the lumen of the inclusion. These results indicated that CT147 might be localized to the inclusion membrane. To test this possibility we double stained infected cells by using antiserum against CT147 and antiserum specific to the well characterized inclusion membrane protein IncG (41). Nearly identical staining reactions were observed with antibodies specific to CT147 and IncG (Fig. 3C) confirming that CT147 colocalizes to the inclusion membrane.

To further study the expression of CT147 we performed immunoprecipitation of intrinsically radiolabeled chlamydiae at different times PI. Anti-CT147 antibodies specifically precipitated a polypep-

tide of 160 kDa that was first detected at 16 h PI (Fig. 3D). These findings provide additional evidence for the expression of CT147 as well as confirming the correct predicted molecular mass of the protein. We did observe differences in the temporal kinetics of expression between immunofluorescence and immunoprecipitation. This likely reflects differences in the sensitivity of the two assays with immunofluorescence being more sensitive because of the inability to intrinsically radiolabel chlamydiae to a high degree of specific activity (typically only 1 dpm per 100 organisms are recovered). Interestingly, CT147 appears to be posttranslationally modified. This conclusion is supported by the disappearance of the predominant 162-kDa polypeptide at the 24- and 40-h time points with the concomitant appearance of several immunoreactive polypeptides of a lower mass, the most prominent being a 90-kDa polypeptide. It is not known whether this modification is the result of chlamydial or host specific proteolytic cleavage. The fact that the protein is posttranslationally modified however strongly implies a biological significance linked to its function.

**Expression of Late Genes.** A total of 26 late genes were identified (Table 1) by microarray. Among these were genes that have been previously characterized as being expressed late in the growth cycle such as *hctA* and *hctB* that encode for the chlamydial histone-like proteins and mediate chromosomal condensation in the differentiation of RBs to EBs (24, 25). An important aspect of the secondary differentiation process (RB to infectious EB) is the expression of genes that encode proteins that form the highly disulfide cross-linked bacterial OM complex (26). The late genes *omcA* and *omcB* encode two cysteine-rich OM proteins that interact with the major OM protein (OmpA) to form this complex. The formation of the OM complex involves extensive intra- and interprotein cross-linking through the formation of cysteine bonds. Two newly identified late genes encode thioredoxin disulfide isomerases (CT780 and CT783), which may function in the exchange of disulfide bonds among cysteine-rich proteins that form the OM complex. Two membrane thiol-proteases genes (*mtpA* and *mtpB*), are expressed late and encode proteins which may play a proteolytic role in the formation of the OM complex. These genes encode proteins with homology to adenoviral proteases (as described on the Los Alamos STD Sequence Database at [www.std-gen.lanl.gov](http://www.std-gen.lanl.gov)), which have proteolytic roles in maturation of virus particles (42). A key protein component of the OM complex, the OmcB protein, has been shown to undergo posttranslational proteolytic processing (43–45). Although it is unknown whether the

observed processing occurs through the membrane thiol proteases described here, we hypothesize that their coexpression, late in the developmental cycle, may be important in OM complex formation. Thus, several late cycle genes encode proteins with potential roles in the formation and maturation of the OM complex, a crucial step in the development of infectious EBs.

**Conclusion.** Genomic-level transcriptional analysis of the chlamydial developmental cycle has identified genes that potentially provide important functions in the interactions of chlamydiae with its host and that coordinate the highly specialized cellular differentiation process that is unique to chlamydial biology (Fig. 4). These genes encode products that are chlamydial-specific and in certain cases, have phylogenetic signatures that suggest eukaryotic origins (46). It is intriguing that an immediate early gene and two late genes have homology to the mammalian gene EEA-1 and the adenovirus C5 proteases, respectively. Moreover, these genes are highly conserved among chlamydiae, implicating a common biological function. It is interesting to speculate that chlamydiae, in the processes of evolving an obligatory lifestyle of intracellular parasitism, may have acquired these genes from the host and have used them functionally to define their highly unique biology.

Our ability to successfully analyze chlamydial temporal gene expression patterns by microarray was linked to both the unique biological characteristics of the pathogen and the experimental design of our work. For example, the identification of immediate early genes was facilitated by infection with high MOI (100) of transcriptionally silent EBs coupled with the rapid isolation of chlamydial RNA from infected host cells. This strategy allowed for synchronization of chlamydial transcription after initial interaction with host cells and the isolation of large quantities of chlamydial RNA without extensive manipulation; hence minimizing RNA degradation. The successful use of microarrays is a particularly valuable and powerful tool for the study of chlamydial pathogenesis because there are no genetic exchange tools for the pathogen. Therefore, many of the differentially expressed chlamydial genes identified in this work would likely remain undiscovered without the use of microarray analysis. The unique nature, intriguing putative function(s), and conservation among chlamydiae of many of the genes describe herein also makes them attractive new targets for therapeutic intervention.

We thank Drs. J. Musser, J. Portis, and D. Nelson for critical review of the manuscript, and Dr. S. Hua and G. Hettrick for help with data presentation.

- Schachter, J. (1978) *N. Engl. J. Med.* **298**, 428–434.
- Brunham, R. C., Binns, B., Gujion, F., Danforth, D., Koseim, M. L., Rand, F., McDowell, J., & Rayner, E. (1988) *J. Infect. Dis.* **158**, 510–517.
- Anttila, T., Saikku, P., Koskela, P., Bloigu, A., Dillner, J., Ikaheimo, I., Jellum, E., Lehtinen, M., Lerner, P., Hakulinen, T., et al. (2001) *J. Am. Med. Assoc.* **285**, 47–51.
- Chesson, H. W. & Pinkerton, S. D. (2000) *J. Acquired Immune Defic. Syndr.* **24**, 48–56.
- Mbizo, E. M., Msuya, S. E., Stray-Pedersen, B., Sundby, J., Chirenje, M. Z. & Hussain, A. (2001) *Int. J. STD AIDS* **12**, 524–531.
- Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R. L., Zhao, Q., Koonin, E. V. & Davis, R. W. (1998) *Science* **282**, 754–759.
- Read, T. D., Brunham, R. C., Shen, C., Gill, S. R., Heidelberg, J. F., White, O., Hickey, E. K., Peterson, J., Utterback, T., Berry, K., et al. (2000) *Nucleic Acids Res.* **28**, 1397–1406.
- Moulder, J. W. (1991) *Microbiol. Rev.* **55**, 143–190.
- Caldwell, H. D., Kromhout, J. & Schachter, J. (1981) *Infect. Immun.* **31**, 1161–1176.
- Newhall, W. J. & Jones, R. B. (1983) *J. Bacteriol.* **154**, 998–1001.
- Hatch, T. P., Allan, I. & Pearce, J. H. (1984) *J. Bacteriol.* **157**, 13–20.
- Bavoil, P., Ohlin, A. & Schachter, J. (1984) *Infect. Immun.* **44**, 479–485.
- Ting, L. M., Hsia, R. C., Haidaris, C. G. & Bavoil, P. M. (1995) *Infect. Immun.* **63**, 3600–3608.
- Su, H., Raymond, L., Rockey, D. D., Fischer, E., Hackstadt, T. & Caldwell, H. D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11143–11148.
- Stephens, R. S., Koshiyama, K., Lewis, E. & Kubo, A. (2001) *Mol. Microbiol.* **40**, 691–699.
- Scidmore, M. A., Rockey, D. D., Fischer, E. R., Heinzen, R. A. & Hackstadt, T. (1996) *Infect. Immun.* **64**, 5366–5372.
- Heinzen, R. A., Scidmore, M. A., Rockey, D. D. & Hackstadt, T. (1996) *Infect. Immun.* **64**, 796–809.
- Scidmore, M. A., Fischer, E. R. & Hackstadt, T. (1996) *J. Cell Biol.* **134**, 363–374.
- Taraska, T., Ward, D. M., Ajioka, R. S., Wyrick, P. B., Davis-Kaplan, S. R., Davis, C. H. & Kaplan, J. (1996) *Infect. Immun.* **64**, 3713–3727.
- Finlay, B. B. & Falkow, S. (1989) *Microbiol. Rev.* **53**, 210–230.
- Duclos, S. & Desjardins, M. (2000) *Cell Microbiol.* **2**, 365–377.
- Raulston, J. E. (1997) *Infect. Immun.* **65**, 4539–4547.
- Fields, K. A. & Hackstadt, T. (2002) *Annu. Rev. Cell Dev. Biol.* **14**, 14.
- Brickman, T. J., Barry, C. E., 3rd & Hackstadt, T. (1993) *J. Bacteriol.* **175**, 4274–4281.
- Barry, C. E., III, Brickman, T. J. & Hackstadt, T. (1993) *Mol. Microbiol.* **9**, 273–283.
- Hatch, T. P. (1999) in *Chlamydia: Intracellular Biology, Pathogenesis, and Immunity*, ed. Stephens, R. S. (Am. Soc. Microbiol., Washington, DC), pp. 29–67.
- Talaat, A. M., Howard, S. T., Hale, W. T., Lyons, R., Garner, H. & Johnston, S. A. (2002) *Nucleic Acids Res.* **30**, e104.
- Xu, Q., Dziejman, M. & Mekalanos, J. J. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 1286–1291.
- Dudley, A. M., Aach, J., Steffen, M. A. & Church, G. M. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 7554–7559.
- Zhong, G., Fan, P., Ji, H., Dong, F. & Huang, Y. (2001) *J. Exp. Med.* **193**, 935–942.
- Zhong, G., Castellino, F., Romagnoli, P. & Germain, R. N. (1996) *J. Exp. Med.* **184**, 2061–2066.
- Hatch, T. P., Miceli, M. & Sublett, J. E. (1986) *J. Bacteriol.* **165**, 379–385.
- Douglas, A. L. & Hatch, T. P. (2000) *Gene* **247**, 209–214.
- Shaw, E. I., Dooley, C. A., Fischer, E. R., Scidmore, M. A., Fields, K. A. & Hackstadt, T. (2000) *Mol. Microbiol.* **37**, 913–925.
- Wichlan, D. G. & Hatch, T. P. (1993) *J. Bacteriol.* **175**, 2936–2942.
- Christoforidis, S., McBride, H. M., Burgoyne, R. D. & Zerial, M. (1999) *Nature* **397**, 621–625.
- Mu, F., Callaghan, J. M., Steele-Mortimer, O., Stenmark, H., Parton, R. G., Campbell, P. L., McCluskey, J., Yeo, J., Tock, E. P. C. & Toh, B. (1995) *J. Biol. Chem.* **270**, 13503–13511.
- Simonsen, A., Lippe, R., Christoforidis, S., Gaullier, J. M., Brech, A., Callaghan, J., Toh, B. H., Murphy, C., Zerial, M. & Stenmark, H. (1998) *Nature* **394**, 494–498.
- Stenmark, H., Aasland, R., Toh, B. H. & D'Arrigo, A. (1996) *J. Biol. Chem.* **271**, 24048–24054.
- Lawe, D. C., Chawla, A., Merithew, E., Dumas, J., Carrington, W., Fogarty, K., Lifshitz, L., Tuft, R., Lambright, D. & Corvera, S. (2002) *J. Biol. Chem.* **277**, 8611–8617.
- Hackstadt, T., Scidmore-Carlson, M. A., Shaw, E. I. & Fischer, E. R. (1999) *Cell Microbiol.* **1**, 119–130.
- Greber, U. F. (1998) *Rev. Med. Virol.* **8**, 213–222.
- Allen, J. E. & Stephens, R. S. (1989) *J. Bacteriol.* **171**, 285–291.
- Everett, K. D. & Hatch, T. P. (1991) *J. Bacteriol.* **173**, 3821–3830.
- Batteiger, B. E., Newhall, W. J., 5th, & Jones, R. B. (1985) *Infect. Immun.* **50**, 488–494.
- Brinkman, F. S., Blanchard, J. L., Cherkasov, A., Av-Gay, Y., Brunham, R. C., Fernandez, R. C., Finlay, B. B., Otto, S. P., Ouellette, B. F., Keeling, P. J., et al. (2002) *Genome Res.* **12**, 1159–1167.