Differential activity of *Rickettsia rickettsii* *ompA* and *ompB* promoter regions in a heterologous reporter gene system

Paul F. Policastro and Ted Hackstadt

Author for correspondence: Paul F. Policastro. Tel: +1 406 363 9343. Fax: +1 406 363 9204. e-mail: pfp@rml.niaid.nih.gov

Laboratory of Intracellular Parasites, Rocky Mountain Laboratories, NIAID, NIH, Hamilton, Montana 59840, USA

The outer membrane of the Gram-negative obligate intracellular parasite *Rickettsia rickettsii* contains two large surface protein antigens with approximate molecular masses of 200 and 135 kDa termed rOmpA and rOmpB, respectively. rOmpB is the most abundant protein in the outer membrane, while rOmpA is a relatively minor constituent. Densitometry of intrinsically radiolabelled protein profiles from *R. rickettsii*-infected Vero cells indicated a molar ratio of approximately 1:9 between rOmpA and rOmpB. The putative promoter–5' untranslated regions (5' UTR) from their recently characterized genes (*rompA* and *rompB*) were placed in the promoter assay vector pKK232-8 to test whether these elements conserve aspects of differential expression in a heterologous host-reporter system. Primer extension analysis of RNA from *Escherichia coli* clones containing the constructs indicated that *E. coli* RNA polymerase faithfully utilizes *rompA* and *rompB* transcription start sites identified previously in *R. rickettsii*. The *rompB* insert directs 28-fold higher levels of chloramphenicol acetyl transferase activity than the *rompA* insert.

Keywords: *Rickettsia rickettsii*, α -proteobacteria, outer membrane proteins, promoter, transcription

INTRODUCTION

Two high molecular mass surface-exposed protein antigens of *Rickettsia rickettsii* show structural and immunogenic features that warrant further investigation. These proteins have been termed rickettsial outer membrane proteins (rOmp) A and B, respectively (McDonald *et al.*, 1988; Anacker *et al.*, 1984), and the genes encoding them, proposed as *ompA* and *ompB* (Gilmore *et al.*, 1991; Hackstadt *et al.*, 1992), are referred to here as *rompA* and *rompB*. rOmpB is the predominant protein in *R. rickettsii* and may play a structural role as an S-layer protein (Palmer *et al.*, 1974; Dasch, 1981). The *rompB* gene sequence predicts a 168 kDa precursor (Gilmore *et al.*, 1991), while the *rompA* gene sequence predicts a 217 kDa protein, with a unique pattern of 13 tandemly repeated units of 72–75 amino acids each in the N-terminal end of the protein, and greater than 75% identity between repeat units (Anderson *et al.*, 1990). rOmpA is a minor constituent of the rickettsial cell surface, yet recombinant

rOmpA antigen serves as an effective immunogen in mice and guinea pigs against challenge with virulent *R. rickettsii* (McDonald *et al.*, 1987). Though rOmpB does not possess any sequence similarity to the amino acid repeats found in the rOmpA protein, rOmpA and rOmpB do show homology within their respective C-terminal 320 amino acids (> 36% identity), a region that may include a membrane anchor (Hackstadt *et al.*, 1992) in the last 10 amino acids of each protein. It appears that in the case of rOmpB, a precursor form of the protein is cleaved to yield a 135 kDa N-terminal polypeptide and a 32 kDa C-terminal polypeptide which remain non-covalently associated during radioimmunoprecipitation (Gilmore *et al.*, 1991; Hackstadt *et al.*, 1992).

In this study we investigate possible genetic mechanisms for the large difference in the relative amounts of the rOmpA and rOmpB proteins. This difference could potentially involve promoter recognition by RNA polymerase, half-life of mRNA, ribosome binding and prevention of ribonuclease inactivation, and efficiency of protein processing and maturation under specific growth conditions. This report focuses on the role played by transcription and translation. The *rompB* gene of *R.*

The GenBank accession numbers for the sequence data reported in this paper are M84167 (*rompB*) and M84168 (*rompA*).

rickettsii possesses a single transcription start site and a relatively long 5' untranslated region (UTR) of 130 nucleotides (nt), as determined by primer extension mapping on transcripts from rickettsia-infected Vero cells (Gilmore *et al.*, 1991). The *rompA* gene also contains a single transcription start site, while its 5' UTR is 31 nt in length (Policastro *et al.*, 1990). To examine these components of gene structure separately from other aspects of protein synthesis, we used *Escherichia coli* as a host for constructs containing the putative promoters, transcription start sites and 5' UTRs of the *rompA* and *rompB* genes inserted in a reporter gene construct. The chloramphenicol acetyl transferase (*cat*) reporter gene function is used because of the apparent toxicity of *romp* gene products in *E. coli*. *E. coli* was chosen as a surrogate system for analysis of rickettsial promoter activity because transformation and clonal isolation of genetically recombined rickettsiae is not feasible at present. Results using this approach indicate these segments of the *rompA* and *rompB* genes do function as promoter elements in *E. coli* and the relative transcriptional activity of these promoters is similar to levels for the intact genes in *R. rickettsii*.

METHODS

Bacterial strains and growth conditions. *E. coli* strain DH5 α (Gibco BRL) was used as host for plasmid preparation, for expression of reporter gene function, and as a source of high molecular mass genomic DNA. Yeast tryptone (YT) was used as plating medium and Luria-Bertani (LB) as liquid medium (Maniatis *et al.*, 1982). The R strain of *R. rickettsii* was grown in Vero cell monolayer culture, isolated and stored frozen in aliquots as previously described (Weiss *et al.*, 1975).

Intrinsic radiolabelling of *R. rickettsii* and quantitative densitometry. Monolayers of Vero cells were infected with the R strain of *R. rickettsii* and maintained at 34 °C for 5 d. The medium was replaced with RPMI-1640 (Roswell Park Memorial Institute medium 1640) containing 0.1 $\times$ the normal amount of amino acids, 1% (w/v) foetal bovine serum (FBS) and emetine (2.5 μ g ml⁻¹), and incubated at 34 °C for 4 h to inhibit host protein synthesis. This medium was replaced with fresh RPMI medium with amino acids, 1% FBS, emetine, plus 1 μ Ci (37 kBq) ¹⁴C-labelled amino acid mixture ml⁻¹ (New England Nuclear). After 48 h continued incubation at 34 °C, the infected cells were scraped into the medium and pelleted by low speed centrifugation (1000 r.p.m. for 5 min). All subsequent steps utilized chilled buffers and precautions were taken to keep the materials cold. The supernatant was saved and the cells resuspended in K36 buffer (Weiss, 1965) for disruption by sonication at 40 W for 40 s. The disrupted cells were again centrifuged for 5 min at 1000 r.p.m. and the supernatants pooled. Rickettsiae in the low speed supernatants were pelleted at 12000 r.p.m. for 15 min. The pellet was resuspended in K36, layered over a 30% (w/v) Renografin (Squibb) pad, and centrifuged at 18000 r.p.m. for 30 min in a Beckman SW 27 rotor. The pellet was washed once with K36 buffer, resuspended in BHI, and samples frozen at -70 °C. Purified organisms were thawed, resuspended after centrifugation in SDS sample buffer and either maintained at 20 °C or heated for 5 min at 100 °C before loading on polyacrylamide gels (Laemmli, 1970). Gels

were treated for fluorography, dried as recommended by the manufacturer (Entensify, New England Nuclear) and exposed at -70 °C to Kodak X-omat AR film. Quantitative densitometry was performed on a LKB Ultrosan XL laser densitometer using preflashed (Laskey & Mills, 1975) film.

Cloning of *rompA* promoter-5' UTR region. All DNA restriction and modification enzymes were purchased from New England Biolabs and used according to the manufacturer's recommendations. A HindIII digest of 10 μ g *R. rickettsii* high molecular mass DNA (Silhavy *et al.*, 1984) was separated by agarose gel electrophoresis, 1.2-1.0 kb fragments were isolated on powdered glass (GeneClean, Bio101), ligated to HindIII-linearized, alkaline phosphatase-treated pUC19, and transformed into competent (Maniatis *et al.*, 1982) *E. coli*. This library was screened (Maniatis *et al.*, 1982) with a previously described minus-strand oligonucleotide probe, 155.5, to nt 34-54 of the putative *rompA* ORF (Policastro *et al.*, 1990), labelled by T4 polynucleotide kinase and [α -³²P]ATP [New England Nuclear; 3000 Ci mmol⁻¹, (111 TBq mmol⁻¹)]. A single positive clone, pH5.1, was confirmed by restriction enzyme digest (HindIII, XbaI and PstI) and Southern blot hybridization to contain the 5' HindIII fragment of the *rompA* gene (Anderson *et al.*, 1990; Policastro *et al.*, 1990).

Amplification of *omp* gene 5' flanking UTR DNAs and insertion in *cat* reporter gene vector. The pH5.1 clone containing the 5' end of the *rompA* gene, a previously described λ -derived phage clone containing the 5' end of the *rompB* gene (Gilmore *et al.*, 1991), and total genomic *E. coli* DNA, prepared as described (Silhavy *et al.*, 1984), were subjected to *Taq* DNA polymerase amplification of specific segments extending from approximately 120 nt upstream of respective transcription start sites to apparent ribosome binding sites in the 5' UTR of the *R. rickettsii rompA*, *rompB* and *E. coli ompA* genes, respectively. The *ompA* gene encodes an abundant outer membrane protein, contains one promoter, and has a relatively long 5' UTR of 134 nt (Cole *et al.*, 1982). Oligonucleotides directed to these sequences (see Fig. 2a) were synthesized on a Milligen/Biosearch Cyclone Plus DNA synthesizer (Millipore). Upstream, plus strand oligonucleotide primers contained *Bam*HI linker sequences and downstream, minus strand primers contained *Sal*I linker sequences to permit directional ligation. In amplifying the *rompB* gene segment, a *Bam*HI site present in the native sequence was used in the upstream oligonucleotide primer and in the cloning step. PCR products were inserted upstream of the RBS and initiation codon of the *cat* reporter gene in plasmid pKK232-8 (Brosius, 1984) (Pharmacia). Cloned or total DNA was used as template for amplification (1 or 5 ng), with 20 pmol of each primer and 50 μ M dNTPs in a final volume of 200 μ l with 5 U *Taq* DNA polymerase in buffer recommended by the manufacturer (USB), for 30 cycles of 30 s at 94 °C, 30 s at 50 °C, and 1.5 min at 72 °C after an initial denaturation at 94 °C for 5 min in a Perkin Elmer-Cetus Thermal Cycler. Products were extracted with phenol/chloroform, precipitated with 2.5 M ammonium acetate and isopropyl alcohol to recover high molecular mass products, washed with 70% ethanol and digested with *Bam*HI and *Sal*I prior to ligation (King & Blakesley, 1986) to *Bam*HI-*Sal*I-digested pKK232-8. *E. coli* strain DH5 α was transformed with ligations and plated on YT containing 250 μ g carbenicillin ml⁻¹ and 30 μ g chloramphenicol ml⁻¹. Plasmid DNA extracted from clones by miniprep (Maniatis *et al.*, 1982) was analysed by *Hinf*I digestion and agarose gel electrophoresis for the presence of inserts of the expected size. Dideoxynucleotide sequence analysis (Sanger *et al.*, 1977; Sanger & Coulson, 1978) of the plasmid inserts (Sequenase; USB), using a minus strand oligonucleotide (232.2: TCTTTACGATGCCATTGGGA) directed against nt 40-59

of the *cat* gene ORF, was used to confirm correct orientation of inserts and fidelity to published sequences (Gilmore *et al.*, 1991; Policastro *et al.*, 1990; Anderson *et al.*, 1990; Beck & Bremer, 1980; Movva *et al.*, 1990) and to 5' flanking sequences of the *rompA* and *rompB* clones.

Chloramphenicol acetyl transferase (CAT) and β -lactamase (β LA) assays. Overnight cultures of transformants containing vector control or recombinant forms of pKK232-8 grown at 37 °C in LB plus 250 μ g carbenicillin ml⁻¹ (vector control) or LB plus 250 μ g carbenicillin ml⁻¹ and 30 μ g chloramphenicol ml⁻¹ (recombinants) were diluted 1:50 in LB, grown at 30 °C to an OD₆₅₀ of 0.4, and assayed for CAT and β LA activities essentially as described by Tomizawa (1985), except the chloramphenicol concentration in the CAT assay was 0.5 mM rather than 0.1 mM. [1-¹⁴C]Acetyl-coenzyme A [Amersham; 5.8 mCi mmol⁻¹ (214.6 MBq mmol⁻¹)] was used as the radiolabelled acetyl donor in the CAT assay. Nitrocefin used in the β LA assay was purchased from Becton Dickinson. Units of enzyme activity are expressed as nmol acetylated chloramphenicol formed min⁻¹ (ml cells)⁻¹ at an OD₆₅₀ of 1.0 (CAT assay) and nmol nitrocefin hydrolysed min⁻¹ (ml cells)⁻¹ at an OD₆₅₀ of 1.0 (β LA assay).

RNA extraction and analysis. Samples (8.0 ml) of cells harvested at the same time as the CAT assay samples (above) were chilled on ice, centrifuged and RNA was extracted and treated with DNaseI (Boehringer Mannheim; RNase-free) as described by Emory & Belasco (1990). Electrophoresis in formaldehyde-agarose gels (Fourney *et al.*, 1988), transfer of RNA to nitrocellulose (Thomas, 1980) and hybridization to [α -³²P]dATP (DuPont/NEN)-labelled, randomly primed (Boehringer Mannheim) restriction fragments containing *bla* and *cat* gene sequences were performed as described previously (Policastro *et al.*, 1989). The *bla* gene probe was a 1.0 kb *Hinf*I-*Eco*RI fragment from pBR322 and the *cat* gene probe was a 0.78 kb *Hind*III fragment from pCM7 (Pharmacia), each isolated on NA45 ion exchange membrane (Schleicher and Schuell) after agarose gel electrophoresis. X-omat AR film (Kodak) was exposed to hybridized filters; for quantitative densitometry, preflashed film was exposed with an intensifying screen, and autoradiograms were analysed as described above. Primer extension analysis on total RNA was formed as described (Policastro *et al.*, 1990), except that 2 μ g RNA were used and the total reaction volume was reduced to 5.6 μ l. The 232.2 minus strand oligonucleotide to the *cat* gene described above was used as primer in all reactions.

RESULTS

Molar ratio of rOmpA and rOmpB

R. rickettsii purified from Vero cell monolayers after intrinsic labelling by ¹⁴C-amino acids gave the SDS-PAGE profile shown in Fig. 1. Several protein mobilities are altered upon heating of rickettsial particles solubilized in SDS (Anacker *et al.*, 1984): rOmpA migrates near rOmpB in extracts prepared at 20 °C, but migrates at 200 kDa upon heating to 100 °C. Quantitative densitometry (see Methods) on the profile after heating to 100 °C indicates that the rOmpA:rOmpB peak area ratio is 1:6.3. When run in a 7.5% separating gel, the heated samples gave a ratio of 1:5.8, or a mean of 1:6 in the two determinations. To adjust this 1:6 ratio for the estimated mass of each mature protein, we used a value of 135 kDa from relative mobility on SDS-PAGE for rOmpB, and 200 kDa for rOmpA. Using the mature rOmpA:rOmpB

molecular mass ratio of 1:4.8 to adjust the densitometric ratio, we derive a molar ratio of 1:8.8 between rOmpA and rOmpB.

Cloning putative promoter-5' UTR sequences of *omp* genes

To test whether *rompA* and *rompB* sequences upstream of their transcription start sites are recognized in a heterologous host-reporter gene context, we amplified and then inserted elements of these rickettsial *omp* genes upstream of a *cat* gene for expression studies in *E. coli*. The length

Fig. 1. Autoradiogram of an SDS-polyacrylamide gel containing *R. rickettsii* R intrinsically labelled with a ¹⁴C-amino acid mix. Lanes: left, cells solubilized at 20 °C; right, cells solubilized at 100 °C. Positions of rOmpA and rOmpB proteins and molecular masses of prestained protein standards are shown on the right.

(a)

Rr *rompA* (pRRA)

BamHI linker CCGGATCC
 -118
 CGAGCCATGCAACAAGGCCAACCCGCTACGGAATGACATATAATCTATACATAAAAAATA
 -35 -10 +1 RBS
 TCATCTATGTTATTAATTATAATAGACATATTAATAAAGTTGCGTATTAACACTTTT**AGTGA**TTTAATTTT**GATTAA**GGTTTACATATG
 GTCGACGG **SaII linker**

Rr *rompB* (pRRB)

BamHI site CC
 -142
 TTTTTTTCGATCCCGGTCAGGCCACGGGATGACACTGATAAGCATAGCGCCAATGTGGTAAAAATATAACATTTATTAATAAAGTTTATAA
 -35 -10 +1
 AAAAACTATAGCGATGCTTGTAAATTATATACATATATAGTATGCTCAGCCGTAGTTTAGAACTATTGAAACAAAATATTAGGTTATTTCTT
 TATCAAATGTGGGATATCTTGACTCACGTTTGATAAGCTTGTTTTAATACTAGGTACTAAATTTTAACTTAAATATAGGACCAAAAATTATG
 RBS
 GTCGACGG **SaII linker**

Ec *ompA* (pECA)

BamHI linker CCGGAT
 -127
 CCGCGATTCCCTCTTCTGTAAATTGTCGCTGACAAAAAAGATTAAACATACCTTATACAAGACTTTTTTTCATATGCCTG
 -35 -10 +1
 ACGGAGTTCACACTTGTAAAGTTTCAACTACGTTGTAGACTTTACATCGCCAGGGGTGCTCGGCATAAGCCGAAGATATCGGTAGAGTTAATAT
 TGAGCAGATCCCCGGTGAAGGATTTAACCGTGTATCTCGTTGSAGATATTCATGGCGTATTTTGGATGATAACGAGGCGCAAAAAATG
 RBS
 GTCGACGG **SaII linker**

(b)

Fig. 2. (a) Sequences of *omp* promoter-5' UTR segments amplified for insertion in the pKK232-8 promoter detection vector. Gene designations (Rr, *rompA*, *rompB*; Ec, *ompA*) and resultant pKK232-8 constructs (pRRA, pRRB, pECA) are indicated for each sequence. Putative -35 and -10 promoter sequences, transcription start sites (+1) and ribosome binding sites (RBS) are underlined; sequences used in oligonucleotides for *Taq* polymerase amplification of the inclusive sequences from template DNA are shown in bold type (complementary strand in the case of the 3' sequence), including *Bam*HI and *Sa*II linker sequences. The 3' flanking gene sequences not amplified for pKK232-8 inserts, including RBS and ATG initiation codons of the ORFs, are shown above *Sa*II linkers. (b) Schematic of the *omp* promoter-5' UTR insertions in pKK232-8. PCR amplification products (PCR fragment) were purified from amplification reactions, cleaved with *Bam*HI and *Sa*II and ligated to appropriately cleaved vector in the multiple cloning site, MCS, as described in Methods. The pKK232-8 vector contains the ribosome binding site (RBS) and initiator methionine codon (ATG) for the *cat* gene downstream from the MCS. Other components of the vector have been described (Brosius, 1984).

of the segment upstream of the start codon in bacterial mRNAs (5' UTR) has been suggested as an important factor in determining mRNA stability (Emory & Belasco, 1990). Therefore the cloned sequences included approximately 120 nt 5' of the reported transcription start sites and 5' UTRs extending to the nucleotide 5' of the ribosome binding site (RBS) of the rickettsial *rompA* and *rompB* genes (Fig. 2a). To compare the relative strength of these foreign *omp* gene elements to a native *E. coli* gene, an equivalent segment of the *ompA* gene of *E. coli* was also

amplified. Target sequences were cloned upstream of the reporter gene in pKK232-8 to yield pRRA, pRRB and pECA, a series of clones containing the *R. rickettsii rompA*, *rompB*, and *E. coli ompA* sequences, respectively (Fig. 2b). This vector contains several transcription terminators flanking the *cat* gene, and the ColE1 origin of replication and β -lactamase (*bla*) gene of pBR322. Clones were analysed by restriction endonuclease mapping and DNA sequencing to confirm accurate copies of the target sequences (data not shown).

Table 1. CAT and β LA activities and mRNA levels of reporter plasmid constructs

Plasmid construct*	Enzyme activity [nmol min ⁻¹ (ml cells) ⁻¹] [†]			mRNA level [‡]		
	CAT	β LA	CAT: β LA ratio (mean $\pm$ SD)	<i>cat</i>	<i>bla</i>	<i>cat:bla</i> ratio (mean $\pm$ SD)
pKK232-8	0.035	200.5	ND	0.3	7.05	ND
pECA (<i>rompA</i>)	25.2	315.0	0.08 $\pm$ 0.002	2.25	11.55	0.195 $\pm$ 0.007
pRRA (<i>rompA</i>)	5.5	281	0.02 $\pm$ 0.001	2.0	11.33	0.175 $\pm$ 0.007
pRRB (<i>rompB</i>)	194.7	349.5	0.559 $\pm$ 0.063	8.7	16.5	0.53 $\pm$ 0.085

ND, Not determined.

* DH5 α containing plasmid constructs. Measurements made on two independent isolates for each plasmid.

[†] CAT values are the mean of two separate determinations and β LA values are the mean of three determinations for each isolate. SD for CAT values were < 2% and for β LA, < 10%.

[‡] Peak areas for full-length mRNA signals. Values for *bla* include both mRNAs seen in Fig. 4(b).

Relative CAT activity of *omp-cat* promoter-reporter gene constructs

Table 1 shows CAT and β LA activities in *E. coli* transformants harbouring the pKK232-8 reporter gene vector with and without *omp* promoter inserts. Independent isolates of each of the *omp* constructs conferred resistance to 30 μ g chloramphenicol ml⁻¹ in broth or solid media, while the parental pKK232-8 plasmid did not. The *rompB* constructs clearly direct greater CAT activity than either the *E. coli ompA*, or the *rompA* constructs. The data in Table 1 include β LA activities for the same extracts, allowing normalization of CAT activities to an independently expressed, plasmid-encoded function (CAT: β LA mean ratio). On this basis the pRRB constructs give 28-fold higher CAT activity than the pRRA constructs. The pRRB constructs were also sevenfold higher in CAT activity than the pECA constructs containing the promoter-5' UTR of the endogenous *ompA* gene of *E. coli*. These CAT activity ratios suggest that the cloned rickettsial sequences contain promoters recognized by the *E. coli* host. These segments of the *rompA* and *rompB* genes, expressed in a novel context, show significant differences in their capacity to effect protein synthesis.

Primer extension analysis on *omp-cat* mRNAs

The CAT activities seen in the above transformants are dependent on inserts placed upstream of the *cat* gene coding sequences, but do not indicate whether specific transcription start sites previously identified for the intact *romp* genes are being utilized in their recombinant host. We therefore isolated RNA from *E. coli* clones to confirm that *cat* mRNA was originating from the same transcription start sites utilized in the native genes. RNA subjected to primer extension analysis was compared to sequence ladders generated with the same *cat* gene oligonucleotide annealed to corresponding plasmid DNA

constructs. Fig. 3 shows that for each *omp* construct, a single major primer extension product characterizes the mRNA population encoding the CAT enzyme. No product is observed in the lanes containing RNA from the parental pKK232-8 plasmid. Comparison of these products to the sequence ladders shows that previously identified transcription start sites, as illustrated in Fig. 2, are in fact the major transcription start sites of the recombinant sequences. Specifically, the *ompA* start site utilized in pECA.1 maps to a G (3' TGTA $\underline{G}$ CGGT 5') in the antisense strand of the sequencing ladder, corresponding to a C (5' ACAT $\underline{C}$ GCCA 3') in the sense strand as reported previously; no evidence was found of transcription from an upstream *ompA* start site seen *in vitro* (Cole *et al.*, 1982).

The *rompB* start site utilized in pRRB.4 maps to a C (5' G $\underline{C}$ CCGTAGT 3') in the sense strand as reported previously (Gilmore *et al.*, 1991). Several additional, minor transcription start sites appear to be utilized in this system that were not observed in the analysis of RNA from *R. rickettsii*. Alternatively, these may be premature reverse transcriptase products in regions of high thymidine content (Shelness & Williams, 1985). In the case of the *rompA* construct pRRA.1, the start site maps to an A (5' TTTTA $\underline{A}$ AGTGA 3') in the sense strand as identified previously (Policastro *et al.*, 1990). Thus for both *romp* gene segments, the predominant transcript species initiates at the same site as initially observed in intact rickettsia.

Quantitative analysis of *cat* transcripts

The CAT activity and primer extension results do not differentiate between promoter strength, message stability and translational efficiency as possible sources of the differences in CAT activity in cell extracts. However, relative intensity of the primer extension products in Fig. 3 suggest that the pRRB and pECA constructs generate

Fig. 3. Primer extension analysis on RNA from *E. coli* transformants bearing pKK232-8 or *omp* promoter-5' UTR derivatives. Of each reaction, 1.3% was loaded on a 6% sequencing gel adjacent to dideoxynucleotide sequence reactions of the respective plasmid constructs pKK232-8.1, pECA.3, pRRB.4 and pRRA.1 as indicated above, primed with the 232.2 oligonucleotide. Lanes: A, G, C, T refer to the dideoxynucleotide used in the respective sequence lanes (the sequence ladders are complementary to the promoter insert sequences shown in Fig. 2); PE, primer extension products. RNA for primer extension was isolated from cells grown in the presence (+) or absence (-) of 250 μ g carbenicillin ml^{-1} (cb) or 30 μ g chloramphenicol ml^{-1} (cm).

higher levels of transcript than the pRRA constructs. To quantify these apparent differences in steady-state mRNA population, a Northern blot of RNA isolated at the same time as clones were extracted for enzyme assays was hybridized to a full-length *cat* gene cassette probe to

Fig. 4. Autoradiogram of RNA from *E. coli* bearing pKK232-8 or *omp* promoter-5' UTR derivatives on Northern blot hybridized to (a) a probe for *cat* gene mRNA or (b) a probe for *bla* gene mRNA. Only that segment of the autoradiogram showing full-length transcripts is shown. The clonal source of the RNAs is shown above each lane. Migration of a 1.35 kb RNA marker is shown on the left.

compare levels of *cat* transcripts (Fig. 4a). Prior to this hybridization, the blot was hybridized to a full-length *bla* gene probe (Fig. 4b) to provide a standard for comparison of *cat* transcript signals. Some residual *bla* probe is apparent at a different mobility than the *cat* transcripts after stripping and re-hybridization (Fig. 4a). The signals indicate that *cat* mRNA is more abundant in cells containing the pRRB constructs than those with the pECA and pRRA constructs. The slower mobility of the pRRB and pECA *cat* transcripts relative to the pRRA transcripts results from the longer 5' UTRs in the former mRNAs. Fig. 4(b) shows that *bla* transcripts are present at approximately equal amounts in all extracts. Two different *bla* mRNAs are to be expected from this pBR322-derived gene, as previously reported (von Gabain *et al.*, 1983). A densitometric comparison of these steady-state transcript levels is shown in Table 1. The normalized values for *cat* transcript levels are represented by the ratio of *omp-cat* to *bla* mRNA (*cat:bla*) and indicate that pRRB clones contain approximately twice the level of *cat* transcript found with pRRA clones or pECA clones. These data, when compared to normalized CAT activities of the cell extracts, show that CAT enzyme levels in these clones are not directly proportional to *cat* transcript levels. Utilization of mRNA to produce protein may be influenced by the different 5' UTRs of these transcripts, such that the pRRB mRNA generates more CAT enzyme than either of the

	-50	-35	-10	+1	+20
Rr <i>rompA</i>	TGTTATTAATTATAATAGACATATTA	AAAAAGTTGCGTTATAAC	ACTTTTAAAGTGATTAAATTTTGATTA		
		5/6	17nt	5/6	7nt
Rr <i>rompB</i>	TAAAAATAGCGATGCTTGTAAATTATA	TACAATATATAGTATGCT	CAGCCCGTAGTTTAGAACTATTGAA		
		4/6	17nt	4/6	6nt
Rr 17 kDa	GAATCTAGTTTTAGCTTTACAAAATT	CTAAAAACCATATACTTATTAATTATATATTAATTTAGAGAGAA			
		5/6	15nt	5/6	8nt
Ec <i>ompA</i>	TGACGGAGTTCACACTTGTAAAGTTTT	CAACTACGTTGTAGACT	TTACATCGCCAGGGGTGCTCGGCATA		
		4/6	16nt	4/6	6nt
Consensus	-----a-----tcTTGACat--t-	-----t-tg-TATAAAT	-----cat-----		
		17nt	7nt		

Fig. 5. Promoter comparisons to *E. coli* consensus sequences. Comparison of the promoter sequences for the *omp* genes in this study and for the *R. rickettsii* 17 kDa lipoprotein surface antigen (Anderson *et al.*, 1988), whose rickettsial transcription start site is also recognized in *E. coli* transformants. Sequences are aligned with respect to conserved -35 and -10 regions, which are shown in bold type, and to +1 transcription start sites, which are underlined. Fractions indicate the number of conserved nucleotides relative to the consensus sequences for -35 and -10 regions. Spacing between -35 and -10 regions and between the -10 region and transcription start site are indicated. The consensus sequence (Hawley & McClure, 1983) shows only those nucleotides which are conserved: dashes represent non-conserved nucleotides. Lower case nucleotides are less conserved positions.

other transcripts, and the pECA mRNA generates more than the pRRA mRNA.

Rickettsial gene comparison to consensus *E. coli* promoter sequences

The *rompA* and *rompB* gene promoters cannot be identified by classical methods of point mutation and homologous RNA polymerase recognition because transformation of rickettsiae is not feasible at present. However, sequences upstream of their transcription start sites, and comparable sequence from the *R. rickettsii* 17 kDa genus-specific lipoprotein gene (Anderson *et al.*, 1988), are compared in Fig. 5 to the *E. coli ompA* promoter and a consensus *E. coli* RNA polymerase promoter (Hawley & McClure, 1983). The sequence from *rompA* has an imperfect match to the most conserved (Hawley & McClure, 1983) TTGAC sequence at -35 and TA-A-T sequence at -10. The *rompB* sequence shows less similarity to the *E. coli* consensus in both of these regions. The rickettsial 17 kDa lipoprotein gene sequence shows as strong a match to the *E. coli* consensus -35 and -10 regions as does the *rompA* gene, but a suboptimal 15 nt gap between -35 and -10 regions. These comparisons may explain *E. coli* RNA polymerase recognition of rickettsial transcription start sites, but the higher CAT activities of the *rompB* and the *ompA* constructs relative to the *rompA* construct appear to be the result of factors in addition to conservation of -35 and -10 sequences. Definitive promoter mapping awaits introduction of altered rickettsial gene sequences into the cognate species.

DISCUSSION

This study began with our interest in understanding the basis for the relative levels of *R. rickettsii* rOmpA and

rOmpB synthesis. The ratio of approximately one rOmpA molecule per nine rOmpB molecules measured late in the rickettsia-host cell interaction (Fig. 1) may not be conserved during all phases of the rickettsial life cycle, and we are presently investigating the dynamics of their relative expression under different culture conditions. Neither of these proteins has a defined role in the colonization by *R. rickettsii* of tick or vertebrate hosts, but their localization to the rickettsial outer membrane implicates them in the first line of interaction with eukaryotic cell components.

We have demonstrated that the sequences upstream of *rompA* and *rompB* serve as accurately recognized targets for expression in *E. coli*. The primer extension results in Fig. 3 demonstrate that *E. coli* RNA polymerase is utilizing native *rompA*, *rompB* and *ompA* transcription start sites in the context of the pKK232 multicopy plasmid, and thus represent accurately transcribed mRNAs containing the rickettsial and *E. coli* 5' UTRs. Conservation of transcription start site usage has been reported previously for the 17 kDa surface protein of *R. rickettsii*, expressed in intact rickettsia in Vero cells and in *E. coli* transformed with the cloned gene (Anderson *et al.*, 1988). It is therefore not surprising that homology exists between *E. coli* σ^{70} promoter regions and sequences upstream of *R. rickettsii* transcripts (Fig. 5), which may indicate some conservation of RNA polymerase recognition sites between these two species. RNA polymerase holoenzymes extracted from *R. prowazekii* and from *E. coli* appear to initiate transcription at identical start sites in two rickettsial genes *in vitro* (Ding & Winkler, 1993). However, promoter sequence preferences may differ between the genus *Rickettsia*, part of the α -2 subgroup of purple bacteria, and *E. coli*, a member of the γ -3 subgroup of the purple bacteria (Weisburg *et al.*, 1985, 1989). Further characterization of rickettsial RNA poly-

merase sequence preferences will be necessary to determine differences in promoter sequence recognition.

In this heterologous system, the relative CAT enzyme activities and *cat* mRNA levels obtained with *rompA* and *rompB* gene sequences show the same trend as rOmpA and rOmpB levels observed in intact rickettsiae. Promoter activity, operationally defined, is at least one factor affecting differential CAT activity from the rickettsial DNA inserts, but the activity of each *omp-cat* construct in *E. coli* (Table 1) reflects both production and degradation of transcripts, and efficiency of translation initiation. It must be recognized that the *rompA:rompB* activity ratio obtained with this reporter gene system may be influenced by host factors specific to *E. coli*, and cannot be extrapolated to conditions in *R. rickettsii*. The half-lives of mRNAs for two *R. prowazekii* proteins in infected mouse L929 cells and in *E. coli* bearing clones of the genes (Cai & Winkler, 1993) indicate that the *E. coli* environment may not be comparable to the native rickettsial environment with regard to the relative stability of full-length rickettsial transcripts. The linkage of a reporter gene to the putative *rompA* and *rompB* gene promoter segments in the present study allowed us to focus on transcription events. Preliminary data suggest that the half-life of *rompB* gene transcripts in *R. rickettsii* is several-fold longer than *rompA* transcripts, which may be a factor in higher *in vivo* levels of rOmpB expression (P. Policastro, unpublished). Half-life differences may exist for *romp-cat* gene fusions in *E. coli*, since the *rompA* and *rompB* 5' UTRs are present in the transcripts. These segments of untranslated RNA are proposed as determinants of message half-life in several cases, including the *E. coli ompA* gene (Emory & Belasco, 1990; Movva *et al.*, 1990; Chen *et al.*, 1991).

Genetic manipulation of rickettsiae has been precluded by the difficulty in obtaining directionally mutated variants of organisms that maintain a strictly obligate intracellular life cycle. Analysis of rickettsial gene transcription is limited to cell-free systems or to testing putative regulatory regions of the structural genes in heterologous systems. Likewise, testing specific functional properties of rickettsial gene products requires the expression of these genes in heterologous organisms. This has been possible in *E. coli* transformants for some rickettsial gene products which complement mutations (Wood *et al.*, 1983) or provide novel catalytic properties (Krause *et al.*, 1985). However, functional expression of outer membrane proteins like the *rompA* and *rompB* gene products may require secretory pathway components that are only present in members of the genus *Rickettsia*. The activity of their putative promoter-5' UTR segments in *E. coli* provides an explanation for the instability of full-length *rompA* and *rompB* genes in this host. In the case of the *rompB* gene, a 120 kDa product can be expressed in *E. coli* from a 3' segment of the gene in pUC19 (Gilmore *et al.*, 1989), but the 5' portion of the gene has only been stably maintained as an insert in a lytic bacteriophage λ vector (Gilmore *et al.*, 1991). Transcription from the intact *rompB* gene in *E. coli* presumably results in toxic levels of rOmpB, as a result of *E. coli* recognition of the rickettsial promoter. This information regarding transcription ac-

tivity will help us to devise alternate strategies for the stable recombinant expression of these important rickettsial antigens.

ACKNOWLEDGEMENTS

We thank members of the Laboratory of Intracellular Parasites for helpful discussions and comments. We thank Stephen Lory for timely help with the β -lactamase assay, Janet Sager for sequence analyses on some of the constructs, and Esther Lewis for her secretarial assistance.

REFERENCES

- Anacker, R. L., Philip, R. N., Williams, J. C., List, R. H. & Mann, R. E. (1984). Biochemical and immunochemical analysis of *Rickettsia rickettsii* strains of various degrees of virulence. *Infect Immun* **44**, 559–564.
- Anderson, B. E., Baumstark, B. R. & Bellini, W. J. (1988). Expression of the gene encoding the 17-kilodalton antigen from *Rickettsia rickettsii*: transcription and posttranslational modification. *J Bacteriol* **170**, 4493–4500.
- Anderson, B. E., McDonald, G. A., Jones, D. C. & Regnery, R. L. (1990). A protective protein antigen of *Rickettsia rickettsii* has tandemly repeated, near-identical sequences. *Infect Immun* **58**, 2760–2769.
- Beck, E. & Bremer, E. (1980). Nucleotide sequence of the gene *ompA* coding the outer membrane protein II of *Escherichia coli* K-12. *Nucleic Acids Res* **8**, 3011–3024.
- Brosius, J. (1984). Plasmid vectors for the selection of promoters. *Gene* **27**, 151–160.
- Cai, J. & Winkler, H. H. (1993). Identification of *tlc* and *gltA* mRNAs and determination of *in situ* RNA half-life in *Rickettsia prowazekii*. *J Bacteriol* **175**, 5725–5727.
- Chen, L.-H., Emory, S. A., Bricker, A. L., Bouvet, P. & Belasco, J. G. (1991). Structure and function of a bacterial mRNA stabilizer: analysis of the 5' untranslated region of *ompA* mRNA. *J Bacteriol* **173**, 4578–4586.
- Cole, S. T., Bremer, E., Hindennach, I. & Henning, U. (1982). Characterisation of the promoters for the *ompA* gene which encodes a major outer membrane protein of *Escherichia coli*. *Mol & Gen Genet* **188**, 472–479.
- Dasch, G. A. (1981). Isolation of species-specific protein antigens of *Rickettsia typhi* and *Rickettsia prowazekii* for immunodiagnosis and immunoprophylaxis. *J Clin Microbiol* **14**, 333–341.
- Ding, H.-F. & Winkler, H. H. (1993). Characterization of the DNA-melting function of the *Rickettsia prowazekii* RNA polymerase. *J Biol Chem* **268**, 3897–3902.
- Emory, S. A. & Belasco, J. G. (1990). The *ompA* 5' untranslated RNA segment functions in *Escherichia coli* as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency. *J Bacteriol* **172**, 4472–4481.
- Fourney, R. M., Miyakoshi, J., Day, R. S., III & Paterson, M. C. (1988). Northern blotting: efficient RNA staining and transfer. *Focus* **10**, 5–7.
- von Gabain, A., Belasco, J. G., Schottel, J. L., Chang, A. C. Y. & Cohen, S. N. (1983). Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. *Proc Natl Acad Sci USA* **80**, 653–657.
- Gilmore, R. D., Jr, Joste, N. & McDonald, G. A. (1989). Cloning, expression and sequence analysis of the gene encoding the 120 kD surface-exposed protein of *Rickettsia rickettsii*. *Mol Microbiol* **3**, 1579–1586.

- Gilmore, R. D., Jr, Cieplak, W., Policastro, P. F. & Hackstadt, T. (1991). The 120 kilodalton outer membrane protein (rOmp B) of *Rickettsia rickettsii* is encoded by an unusually long open reading frame: evidence for protein processing from a large precursor. *Mol Microbiol* **5**, 2361–2370.
- Hackstadt, T., Messer, R., Cieplak, W. & Peacock, M. G. (1992). Proteolytic cleavage of the 120 kDa outer membrane protein of *Rickettsiae*: identification of an avirulent mutant deficient in processing. *Infect Immun* **60**, 159–165.
- Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res* **11**, 2237–2255.
- King, P. V. & Blakesley, R. W. (1986). Optimizing DNA ligations for transformation. *Focus* **8**, 1–3.
- Krause, D. C., Winkler, H. H. & Wood, D. O. (1985). Cloning and expression of the *Rickettsia prowazekii* ADP/ATP translocator in *Escherichia coli*. *Proc Natl Acad Sci USA* **82**, 3015–3019.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Laskey, R. A. & Mills, A. D. (1975). Quantitative film detection of ³H and ¹⁴C in polyacrylamide gels by fluorography. *Eur J Biochem* **56**, 335–341.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- McDonald, G. A., Anacker, R. L. & Garjian, K. (1987). Cloned gene of *Rickettsia rickettsii* surface antigen: candidate vaccine for Rocky Mountain spotted fever. *Science* **235**, 83–85.
- McDonald, G. A., Anacker, R. L. & Mann, R. E. (1988). Extraction of protective components of *Rickettsia rickettsii* with *n*-octyl- β -D-glucopyranoside. *Rev Infect Dis* **10**, S382–S385.
- Movva, N. R., Nakamura, K. & Inouye, M. (1990). Regulatory region of the gene for the OmpA protein, a major outer membrane protein of *Escherichia coli*. *Proc Natl Acad Sci USA* **77**, 3845–3849.
- Palmer, E. L., Martin, M. L. & Mallavia, L. (1974). Ultrastructure of the surface of *Rickettsia prowazekii* and *Rickettsia akari*. *Appl Microbiol* **28**, 713–716.
- Policastro, P. F., Fredholm, M. & Wilson, M. C. (1989). Truncated gag products encoded by *Gp-1*-responsive endogenous retrovirus loci. *J Virol* **63**, 4136–4147.
- Policastro, P. F., Anderson, B. E. & McDonald, G. A. (1990). Promoter structure and expression of the 155-kDa surface antigen gene of *Rickettsia rickettsii*. In *Rickettsiology: Current Issues and Perspectives*, pp. 468–477. Edited by K. E. Hechemy, D. Paretsky, D. H. Walker and L. P. Mallavia. New York: New York Academy of Sciences.
- Sanger, F. & Coulson, A. R. (1978). The use of thin acrylamide gels for DNA sequencing. *FEBS Lett* **87**, 107–110.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Shelness, G. S. & Williams, D. L. (1985). Secondary structure analysis of apolipoprotein II mRNA using enzymatic probes and reverse transcriptase. *J Biol Chem* **260**, 8637–8646.
- Silhavy, T. J., Berman, M. L. & Enquist, L. W. (1984). *Experiments with Gene Fusions*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Thomas, P. S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc Natl Acad Sci USA* **77**, 5201–5205.
- Tomizawa, J.-I. (1985). Control of ColE1 plasmid replication: initial interaction of RNA 1 and the primer transcript is reversible. *Cell* **40**, 527–535.
- Weisburg, W. G., Woese, C. R., Dobson, M. E. & Weiss, E. (1985). A common origin of rickettsiae and certain plant pathogens. *Science* **230**, 556–558.
- Weisburg, W. G., Dobson, M. E., Samuel, J. E., Dasch, G. A., Mallavia, L. P., Baca, O., Mandelco, L., Sechrest, J. E., Weiss, E. & Woese, C. R. (1989). Phylogenetic diversity of the rickettsiae. *J Bacteriol* **171**, 4202–4206.
- Weiss, E. (1965). Adenosine triphosphate and other requirements for the utilization of glucose by agents of the psittacosis-trachoma group. *J Bacteriol* **90**, 243–253.
- Weiss, E., Coolbaugh, J. C. & Williams, J. C. (1975). Separation of viable *Rickettsia typhi* from yolk sac and L cell host components by Renografin density gradient centrifugation. *Appl Microbiol* **30**, 456–463.
- Wood, D. O., Atkinson, W. H., Sikorski, R. S. & Winkler, H. H. (1983). Expression of the *Rickettsia prowazekii* citrate synthase gene in *Escherichia coli*. *J Bacteriol* **155**, 412–416.

Received 7 March 1994; revised 1 July 1994; accepted 7 July 1994.