EXPRESSION ANALYSIS AND ANTIBODY NEUTRALIZATION OF P44 MAJOR SURFACE PROTEINS OF ANAPLASMA PHAGOCYTOPHILUM DURING MAMMALIAN INFECTION

Dissertation

Presented in Partial Fulfillment of the Requirements for
The Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By
Xueqi Wang, M. S.

* * * * *

The Ohio State University
2006

Dissertation Committee:
Professor Yasuko Rikihisa
Professor James W. DeWille
Professor Young C. Lin
Associate Professor Roger W. Stich

Approved by

Adviser
Graduate Program in Veterinary Biosciences
**ABSTRACT**

*Anaplasma phagocytophilum* is an obligatory intracellular bacterium that causes human granulocytic anaplasmosis. The immunodominant polymorphic 44-kDa major surface proteins of *A. phagocytophilum* are encoded by more than 80 *p44* paralogs. Although diverse P44 species are shown to be expressed, what drives the P44 antigenic variation in mammals is unknown. Objectives of this study were to 1) characterize the temporal and dynamic expression of *p44* genes during the infection of horses, immunocompetent and immunocompromised mice, and in cell culture, 2) examine expression of genes involved in homologous recombination and polycistronic expression of the *p44*-expression locus, and 3) elucidate P44 antibody neutralization mechanisms in vitro. First, our results showed rapid switch-off of the initial dominant transcript *p44-18* occurred in the blood of horses during the logarithmic increase of bacteria growth. Each of the subsequently dominant *p44* transcript species was phylogenetically dissimilar from *p44-18*. When *A. phagocytophilum* was preincubated with infected horse plasma, the dominance of the *p44-18* transcript was rapidly suppressed in vitro and most of the newly emerged *p44* transcript species were previously undetected in this horse. Second, we demonstrated two monoclonal antibodies recognize bacterial surface-exposed epitopes of naturally folded P44 proteins and mapped these epitopes to specific peptide sequences. These results indicated that antibodies directed to certain epitopes of P44 proteins have a critical role in inhibiting *A. phagocytophilum* infection. Third, immunocompetent mice
cleared *A. phagocytophilum* infection by 3 weeks post inoculation (p.i.), whereas *A. phagocytophilum* persisted at least two months in C3H/HeJ (TLR4 deficient) and SCID mice. The development of antibodies against P44 N-terminus region was delayed in C3H/HeJ mice. Of seven genes involved in homologous recombination, only *recA* mRNA level in *A. phagocytophilum* was significantly greater in mice than in cell culture. Fourth, an electrophoretic migration shift assay showed a specific shift of the promoter region of *tr1* in the *p44*-expression locus upon incubation with *A. phagocytophilum* lysate. These studies are expected to facilitate the understanding of the mechanisms of P44 antigenic variation: interplay between bacterial recombination and transcriptional regulation and immunoclearance, and a new approach in designing vaccine candidate antigens for the control of HGA.
ACKNOWLEDGMENTS

I wish to thank my adviser, Yasuko Rikihisa, for intellectual support, encouragement, and enthusiasm which made this thesis possible, and for her patience in correcting both my scientific and stylistic errors.

I wish to thank my committee members, Dr. James W. DeWille, Dr. Young C. Lin and Dr. Roger W. Stich, for stimulating discussion of various aspects of this thesis.

I wish to thank Dr. Takane Kikuchi for the peptide synthesis and assistance with the peptide-pin ELISA analysis. I also wish to thank those who help me to solve various experimental problems, especially Drs. Yumi Kumagai and Takane Kikuchi.

This dissertation research was supported by grants from the National Institute of Health without which the dissertation would not have been possible.
VITA

January 6, 1972 ...............................................Born–Shanghai, China

1995 B. S. ......................................................Fudan University, China

1998 M. S.......................................................Fudan University, China

2001- Present ................................................the Ohio State University

PUBLICATIONS

Research Publications:


FIELDS OF STUDY

Major Field: Veterinary Biosciences
TABLE OF CONTENTS

Abstract ............................................................................................................................. ii

Dedication ........................................................................................................................ iv

Acknowledgments .......................................................................................................... v

Vita ................................................................................................................................... vi

List of Tables ................................................................................................................... x

List of Figures ................................................................................................................ xi

Chapters:

1. Introduction ........................................................................................................... 1

2. Rapid Sequential Changeover of Expressed p44 Genes during the Acute Phase of Anaplasma phagocytophilum Infection in Horses ...................... 18
   2.1 Introduction ....................................................................................... 18
   2.2 Materials and methods ...................................................................... 20
   2.3 Results ............................................................................................... 25
   2.4 Tables and Figures ............................................................................ 32
   2.5 Discussion ......................................................................................... 42
   2.6 Summary ........................................................................................... 45

3. Two Monoclonal Antibodies with Defined Epitopes of P44 Major Surface Proteins Neutralize Anaplasma phagocytophilum by Distinct Mechanisms ...... 46
   3.1 Introduction ....................................................................................... 46
   3.2 Materials and methods ...................................................................... 49
   3.3 Results ............................................................................................... 56
   3.4 Tables and Figures ............................................................................ 61
   3.5 Discussion ......................................................................................... 75
   3.6 Summary ........................................................................................... 79
4. Delayed Development of Antibody to P44N and Upregulation of RecA in Diverse P44 Expression and Delayed Clearance of Anaplasma phagocytophilum in C3H/HeJ Mice .................................................................80

4.1 Introduction ........................................................................................................ 80
4.2 Materials and methods ..................................................................................... 83
4.3 Results .................................................................................................................. 85
4.4 Tables and Figures ............................................................................................. 88
4.5 Discussion ............................................................................................................ 98
4.6 Summary .............................................................................................................. 102

5. Identification of a DNA-Binding Protein for Upstream Region of P44 Operon Encoding Major Outer Membrane Proteins of Anaplasma phagocytophilum .........................................................................................103

5.1 Introduction ........................................................................................................ 103
5.2 Materials and methods ..................................................................................... 105
5.3 Results .................................................................................................................. 109
5.4 Tables and Figures ............................................................................................. 112
5.5 Discussion ............................................................................................................ 118
5.6 Summary .............................................................................................................. 121

Bibliography ............................................................................................................. 122
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Sequences of oligonucleotides used in RT-PCR, PCR, colony hybridization, and cloning</td>
<td>32</td>
</tr>
<tr>
<td>2.6 Characteristics of ( p44 ) cDNA clonotype distribution after incubation of ( A. ) phagocytophilum with horse EQ005 day 31 and day 0 plasma</td>
<td>41</td>
</tr>
<tr>
<td>4.1 Sequences of oligonucleotides used in PCR and RT-PCR</td>
<td>88</td>
</tr>
<tr>
<td>4.4 ( p44 ) transcript species identified in C3H/HeN, C3H/HeJ mouse spleen and cell culture samples on day 12 p.i.</td>
<td>93</td>
</tr>
<tr>
<td>5.1 Sequences of oligonucleotides used in PCR amplification of EMSA probes</td>
<td>112</td>
</tr>
<tr>
<td>5.2 Homologous ( tr1 ) genes found in sequenced genomes of the family ( Anaplasmataceae )</td>
<td>113</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2</td>
<td><em>A. phagocytophilum</em> levels in the blood of two horses during the course of infection, as determined by <em>p44</em> C-PCR</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Relative proportion of each <em>p44</em> transcript in the blood of horses EQ005 and EQ006</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Development of anti-hvP44-18 antibody in the plasma of <em>A. phagocytophilum</em>-infected horses</td>
<td>37</td>
</tr>
<tr>
<td>2.5</td>
<td>Infected horse plasma reduced dominance of the <em>p44-18</em> transcript in cell culture</td>
<td>39</td>
</tr>
<tr>
<td>3.1</td>
<td>MAb 5C11 recognizes the P44 N-terminal conserved region</td>
<td>61</td>
</tr>
<tr>
<td>3.2</td>
<td>MAb 5C11 and MAb 3E65 label the surface of <em>A. phagocytophilum</em></td>
<td>63</td>
</tr>
<tr>
<td>3.3</td>
<td>MAb 5C11 and MAb 3E65 inhibit infection of <em>A. phagocytophilum</em> in HL-60 cells by inhibiting and intracellular development, respectively</td>
<td>65</td>
</tr>
<tr>
<td>3.4</td>
<td>Two views of the representative P44 protein three-dimensional structure predicted by the Robetta program and the MAb 5C11 epitope</td>
<td>69</td>
</tr>
<tr>
<td>3.5</td>
<td>Peptide mapping of MAb 5C11 and MAb 3E65 epitopes</td>
<td>71</td>
</tr>
<tr>
<td>3.6</td>
<td>ELISA analysis for the linear B-cell epitopes within P44N-N and P44-18hvC-C of <em>A. phagocytophilum</em>-infected horses and mice</td>
<td>73</td>
</tr>
<tr>
<td>4.2</td>
<td>Levels of <em>A. phagocytophilum</em> organisms in the spleen of C3H/HeN and C3H/HeJ; ICR and ICR/SCID mice as determined by <em>p44</em> cPCR</td>
<td>89</td>
</tr>
<tr>
<td>4.3</td>
<td>Development of anti-P44N and P44-18hv antibodies in the plasma of <em>A. phagocytophilum</em>-infected mice</td>
<td>91</td>
</tr>
</tbody>
</table>
4.5 Percentage of \textit{p44-18} transcript population in cell culture, C3H/HeJ, and C3H/HeN mouse spleen tissues at day 12 p.i. ................................................................. 94

4.6 \textit{recA} expression levels in cell culture, C3H/SCID and C3H/HeN mouse spleen samples on day 12 p.i.............................................................................. 96

5.3 Schematic representation of primer locations within the intergenic region of \textit{ndk} and \textit{tr1} genes at the \textit{p44} expression locus of \textit{A. phagocytophilum} and EMSA analysis of proteins bound upstream of \textit{tr1} upstream region with different DNA probes ................................................................. 114

5.4 Proteins bound to the R3-1 probe ........................................................................... 116
CHAPTER 1

INTRODUCTION

Taxonomy and Genome sequence


So far, the genome sequences of eight *Anaplasmataceae* members are completed. They are three representative zoonotic anaplasmal pathogens, *Anaplasma phagocytophilum* (CP000235), *Ehrlichia chaffeensis* (CP000236) and *Neorickettsia sennetsu* (CP000237); one insect parasite, *Wolbachia pipiensis* wMel (AE017196); the filarial nematode endosymbiont, *Wolbachia* sp. wBmsp; two bovine pathogens, *A. marginale* (CP000030) and *E. ruminantium* (CR767821); and one dog pathogen, *E. canis* (CP000107) (Brayton *et al.*, 2005; Collins *et al.*, 2005; Foster *et al.*, 2005; Wu *et al.*, 2004). Members of the order Rickettsiales have relatively small genomes (0.8-1.5 Mb) that have arisen through reductive evolution as they developed dependence on the host cell for essential functions. Based on the genome sequence data for *W. pipiensis* wMel,
*Ehrlichia* spp., and *Anaplasma* spp., which are most closely related, all have numerous repeats in their genomes. The repetitive nature of the *Ehrlichia* and *Anaplasma* genomes is exemplified by the expansion of outer membrane proteins of the OMP-1/P44/Msp2 family and other functionally important genes including those involved in type IV secretion and vitamin cofactor biosynthesis (Dunning Hotopp *et al.*, 2006).

Human Granulocytic Anaplasmosis

Human granulocytic anaplasmosis (HGA, formerly human granulocytic ehrlichiosis) is an emerging tick-borne zoonosis that has been reported in the United States and Europe (van Dobbenburgh et al., 1999; Petrovec et al., 1997; Bakken et al., 1994). HGA is characterized by chills, headache, myalgia, and hematological abnormalities, including leukopenia and thrombocytopenia, and increased serum transaminase activities suggesting mild to moderate liver injury. In addition, the gastrointestinal and respiratory systems are frequently affected, although skin rash is rare (Dumler, 2005). It frequently requires prolonged hospitalization, and when the treatment is delayed due to misdiagnosis or in elderly or immunocompromized patients, HGA can be fatal (Hardalo et al., 1995). However, most patients respond within 24 to 48 hours to treatment with oxytetracycline (Dumler, 2005). HGA is increasingly recognized as an important and frequent cause of fever after tick bite in the Upper Midwest, New England, parts of the mid-Atlantic states, northern California, and many parts of Europe (Dumler et al., 2005). These areas correspond to the distribution of Ixodes persulcatus complex ticks which includes I. scapularis (black-legged or deer tick) in the eastern United States (Madigan et al., 1995; Pancholi et al., 1995) and I. pacificus (western black-legged tick) in the western United States (Richter et al., 1996), I. ricinus in Europe (Ogden et al., 1998) and I. persulcatus in parts of Asia (Cao et al., 2003; Kim et al., 2003). From 1986 through 1997, about 450 cases of HGA have been confirmed in the United States: Connecticut (15.90 cases per million), Wisconsin (8.79 cases per million), Minnesota (3.90 cases per million), and New York (2.68 cases per million) (McQuiston et al., 1999). From 2001 through 2002, the average reported annual incidence for HGA was 1.4 cases
per million population (Demma et al., 2005). Since 1986, at least 2,189 cases of HGA have been reported to the CDC (Dumler, 2005). Small mammals, in particular white-footed mice (Peromyscus leucopus) are considered to be the major reservoir of A. phagocytophilum (Walls et al., 1997; Telford et al., 1996). Human coinfection with A. phagocytophilum and Borrelia burgdorferi occurs because both organisms are indigenous to the same areas and use the same tick vector (Nadelman et al., 1997; Aguero-Rosenfeld et al., 1996; Telford et al., 1996; Pancholi et al., 1995; Bakken et al., 1994).

The diagnosis of Anaplasmosis is affected by specific laboratory diagnostic tests, such as the direct microscopic examination of peripheral blood smears, detection of specific antibodies by IFA or enzyme-linked immunosorbent assay, culture isolation of the A. phagocytophilum by inoculating the affected animal buffy coat fraction of peripheral blood into cell culture, and polymerase chain reaction (PCR) amplification of specific 16S rRNA or p44 multigene family (Wang et al., 2004; Tajima et al., 2000; Zhi et al., 1997; Edelman and Dumler, 1996; Rikihisa, 1991).

**Anaplasma phagocytophilum**

*A. phagocytophilum*, etiologic agent of human granulocytic or granulocytotropic anaplasmosis (HGA), are gram-negative, coccoid to ellipsoidal, often pleomorphic, intracytoplasmic bacteria that infect cells of mammalian bone marrow derivation, predominantly cells in the myeloid lineage (Dumler et al., 2001). They replicate in membrane-bound vacuoles (parasitophorous vacuoles) in the cytoplasm of host cells (Rikihisa, 2003). In mammalian cells, the size of micro-colonies (morulae) are usually between 1.5-2.5 μm in diameter, may be as large as 6 μm (Popov et al., 1998; Rikihisa et al., 1991).
al., 1997). The bacterium has a relatively small genome size (1.47 Mb) including 1,369 ORFs (Dunning Hotopp et al., 2006). *A. phagocytophilum* is maintained through an enzootic cycle between wild animals and Ixodid ticks and infects humans through the bite of infected ticks. *A. phagocytophilum* requires horizontal and bidirectional transmission altering between mammals, or birds, and ticks (Goethert and Telford, 2003; Alekseev et al., 2001; Telford et al., 1996), because transovarial transmission of these bacteria is ineffective (Hodzic et al., 1998a; Macleod, 1936). Wild rodents, such as white-footed mice (*Peromyscus leucopus*) and dusky-footed wood rats (*Neotoma fuscipes*), are believed to be the major reservoirs of *A. phagocytophilum* in the eastern and western United States (Rikihisa, 2003).

**Pathogenesis of *A. phagocytophilum***

*A. phagocytophilum* is a very unusual bacterium. It is an obligatory intracellular bacterium, its major tropism is neutrophils which represent 50 to 60% of the total circulating leukocytes and constitute the mammalian “first line of defense”. In order to infect and replicate in host cells, *A. phagocytophilum* created their own intracellular compartment for replication. Furthermore, this bacterium acquired some unique features to avoid the host immunity and also to orchestrate a series of events that culminate in suppression of host innate immune response (Rikihisa, 2006).

**Internalization of host cells**

*A. phagocytophilum* lack pili or capsules (Rikihisa, 1991), these organisms are believed to bind the host cells via its outer membrane proteins (Rikihisa, 2003). Our
previous study showed the entry of *Ehrlichia, Neorickettsia* spp. (closely related to *Anaplasma*) and *Anaplasma* spp. into host phagocytes is independent of microfilaments, but dependent on transglutaminase activity, which is required for receptor-mediated endocytosis (Yoshiie *et al.*, 2000; Barnewall *et al.*, 1999; Rikihisa *et al.*, 1994; Messick and Rikihisa, 1993). It was reported that the binding of *A. phagocytophilum* to HL-60 cells was dependent on the expression of both P-selectin glycoprotein ligand-1 (PSGL-1) and an α1-3-fucosyltransferase (Herron *et al.*, 2000), whereas the binding of *A. phagocytophilum* to murine neutrophils requires expression of α1-3-fucosyltransferase, but not PSGL-1 (Carlyon *et al.*, 2003). We also demonstrated that the cholesterol-rich lipid rafts or caveolae and glycosylphosphatidylinositol (GPI)-anchored proteins were involved in the entry and intracellular infection of *E. chaffeensis* and *A. phagocytophilum* (Lin and Rikihisa, 2003b).

**Creation of a replicative compartment that does not fuse with lysosomes**

Our previous study showed that *E. chaffeensis* and *A. phagocytophilum* occupy unique cytoplasmic compartments which are distinct from endosomes or phagolysosomes (Barnewall *et al.*, 1999; Mott *et al.*, 1999). *E. chaffeensis* selectively accumulates transferrin receptors (TfRs) on its inclusion membrane (Barnewall *et al.*, 1999). This modification mimics *E. chaffeensis* inclusion transferrin endosomes which are separated from the lysosome fusion pathway. Another strategy to avoid lysosomal fusion is for the parasite to isolate itself from host endocytic and exocytic vesicular traffic. *E. chaffeensis* and *A. phagocytophilum* inclusions lack lysosome markers LAMP-1 and LAMP-2. Moreover, *A. phagocytophilum* inclusions do not colocalize with early endosomal antigen
1 (EEA1) and rab5, which are found in the early endosomes. *A. phagocytophilum* inclusions are deficient of vacuole-type H⁺-ATPase, which is responsible for phagosomal acidification (Mott *et al.*, 1999). Therefore, the inclusion of intracellular bacteria is sequestered from the host cell endocytic and exocytic traffic.

**Downregulation of reactive oxygen intermediate generation**

Neutrophils are the most powerful superoxide anion (O₂⁻) generators. In order to survive, *A. phagocytophilum* acquired the remarkable ability to prevent O₂⁻ generation by neutrophils. *A. phagocytophilum* decreased levels of p22phox in human peripheral blood neutrophils and HL-60 cells. The absence of colocalization of NADPH oxidase components with the inclusion further protects *A. phagocytophilum* from oxidative damage (Mott *et al.*, 2002; Mott and Rikihisa, 2000). Our previous study also showed *A. phagocytophilum* and *E. chaffeensis* contained sodB, which encodes superoxide dismutase. This superoxide dismutase may play an important role for the oxidative stress response, since sodB is cotranscribed with the type IV secretion system (Ohashi *et al.*, 2002). The genomic sequencing study showed that *E. chaffeensis*, *N. sennetsu* and *A. phagocytophilum* each contain the 12 transmembrane segments and six conserved histidine residues consistent with members of the heme-copper oxidase family. These orthologs may function as nitric oxide reductases. These three spp. also encode proteins with functional motifs similar to flavohemoglobins that may function as alkylhydroperoxide and nitric oxide reductases (Dunning Hotopp *et al.*, 2006).
Unusual cell wall

To avoid activating host cell innate immunity, *A. phagocytophilum* acquired unique membrane features. They lost all genes required for the biosynthesis of lipopolysaccharide (LPS), and most genes required for the biosynthesis of peptidoglycan. Instead, *A. phagocytophilum* incorporate host cell cholesterol into their membrane to compensate for the reduction in membrane stability by the loss of LPS and peptidoglycan (Lin and Rikihisa, 2003a). Therefore, the loss of LPS and peptidoglycan avoid recognition by host cells and subsequent activation of signal transduction pathway gene expression that triggers innate immune effects.

Inhibition of host cell apoptosis

*A. phagocytophilum* induces a significant delay in morphological apoptosis and the cytoplasmic appearance of histone-associated DNA fragments in granulocytes (Yoshiie *et al.*, 2000). Our study showed that *A. phagocytophilum* not only inhibits human neutrophil apoptosis via transcriptional upregulation of *bfl-1* and inhibition of mitochondria-mediated activation of caspase 3 (Ge *et al.*, 2005), but also inhibits fas-mediated neutrophil apoptosis. Another study also showed that *A. phagocytophilum* blocks CD95 (APO-1/Fas)-mediated apoptosis (Borjesson *et al.*, 2005).

Downregulation of leukocyte activation and differentiation signals

*A. phagocytophilum* inhibits or delays the crucial signaling in host phagocyte activation and differentiation pathways. *A. phagocytophilum* induced IL-1β, TNF-α, and IL-6 expression in monocytes in vitro as early as 2 h p.i. (Kim and Rikihisa, 2002).
However, in the hosts of *A. phagocytophilum*, neutrophils, delayed expression of these cytokines was detected (Borjesson *et al.*, 2005; Kim and Rikihisa, 2002). Moreover, p38 MAPK was not activated in neutrophils incubated with *A. phagocytophilum* (Kim and Rikihisa, 2002). INF-γ signaling was defective in infected HL-60 cells, because the binding of phosphorylated Stat 1 to the promoter of interferon regulatory factor-1 was impaired (Thomas *et al.*, 2005).

**Interaction between organisms and host environment**

*A. phagocytophilum* is cycled between mammalian hosts and tick vectors. In order to adapt to these different environments, this organism uses two component systems and a type IV secretion system to sense and respond to environmental changes. The type IV secretion system is one of the few sets of genes syntenic between all of the Rickettsiales sequenced, indicating that tight coordination of expression of these genes is critical (Dunning Hotopp *et al.*, 2006). Transcription of the three *virB* genes in two *Anaplasma* and *Ehrlichia* spp. is regulated by factors that influence the *sodB* gene expression. This unique regulation of gene expression for the type IV secretion system may be associated with intracellular survival and replication of *Anaplasma* and *Ehrlichia* spp. in granulocytes or monocytes (Ohashi *et al.*, 2002). Our previous study also indicated the developmental regulation of expression of components of the type IV secretion system during *A. phagocytophilum* intracellular life cycle (Niu *et al.*, 2006).

The genomes of *E. chaffeensis* and *A. phagocytophilum* were each predicted to encode three pairs of two component systems that were conserved in some members of the order Rickettsiales. A number of *E. chaffeensis* genes, including the six two
component system genes, were down-regulated with closantel treatment. These results suggest that these two component systems play an essential role in infection and survival of *E. chaffeensis* and *A. phagocytophilum* in human leukocytes (Cheng et al., 2006).

**Animal models and pathogenicity**

The horse is not only the natural host of *A. phagocytophilum*, but it serves as a useful animal model for human granulocytic anaplasmosis (Wang et al., 2004; Kim et al., 2002; Zhi et al., 2002b). During the horse infection, IL-1β, TNF-α and IL-8 mRNA expression was upregulated, suggesting generation of these three cytokine during *A. phagocytophilum* infection has a primary role in *A. phagocytophilum* pathogenesis and immunomodulation (Kim et al., 2002). In the horse model, the bacterial load showed a monotopic, but not oscillating, pattern of rickettsemia: an initial logarithmic increase followed by a precipitous decline during 1 month post i.v. inoculation or post tick placement (Wang et al., 2004; Kim et al., 2002; Zhi et al., 2002b).

Laboratory mice can be infected with *A. phagocytophilum* and serve as a model for studying of granulocytic anaplasmosis (Borjesson and Barthold, 2002). INF-γ has a protective role in mice infected with *Ehrlichia* spp. (Bitsaktsis et al., 2004; Ismail et al., 2004). IFN-γ-deficient mice had a markedly elevated *A. phagocytophilum* burden (Akkoyunlu and Fikrig, 2000). One study showed that infection of *A. phagocytophilum* induces neutrophil secretion of interleukin-8 or murine homologs and perpetuates infection by recruiting susceptible neutrophils (Akkoyunlu et al., 2001). Therefore, blockade of CXCR2 inhibits the *A. phagocytophilum* infection (Akkoyunlu et al., 2001). The infection of *A. phagocytophilum* MRK strain in TLR2-/-, TLR4-/-, MyD88-/-, TNFα-
/-, inducible nitric oxide synthase, or phagocyte NADPH oxidase (gp91phox/-) knockout mice does not affect bacterial burden (von Loewenich et al., 2004).

**Major outer membrane protein P44**

**Gene structure**

The *A. phagocytophilum* genome has three *omp-1*, one *msp2*, two *msp2* homolog, one *msp4* and 113 *p44* loci belonging to the OMP-1/MSP2/P44 superfamily. The largest expansion of this family is that of *p44* genes in *A. phagocytophilum*. *p44*s, the members of *p44* multigene family, consist of a central hypervariable region of approximately 280 bp and conserved flanking sequences longer than 50 bp (Dunning Hotopp et al., 2006). The *p44*s were annotated as full-length, silent/reserve, truncated, and fragments. There are 22 full-length *p44*s identified that have ORFs longer than 1.0 kb with conserved start and stop codons. Sixty-four shorter *p44*s, which lack a translational start codon and likely serve as reserve/silent *p44*s, were identified by locating highly conserved 5′ and 3′ flanking sequences and signature sequences within the hypervariable region. The full-length and silent/reserve *p44* genes are preferentially located near the replication origin and symmetrically located around the *p44* expression locus. Localization near the origin, where multiple replication forks coexist, may facilitate recombination between the expression locus and the reserve/silent *p44* genes. In addition to the full-length and silent/reserve *p44* genes, 21 5′ and 3′ fragments and six truncations of *p44* genes larger than 60 nucleotides have been identified in the genome. Truncations include portions of a hypervariable region; fragments did not include a hypervariable region (Dunning Hotopp et al., 2006).
Diverse p44 paralogs (p44–1 to p44–65) are expressed in mammals and ticks and confer antigenic environmental adaptation, especially during tick transmission (Felek et al., 2004; Wang et al., 2004; Lin et al., 2002; Zhi et al., 2002b). Another 23 p44 genes (p44-66 to p44–88) have not yet been experimentally identified as being expressed. Total P44s are upregulated in A. phagocytophilum at 37°C than at 24°C (Zhi et al., 2002b) and in mice compared to tick salivary glands (Zhang et al., 2006). The specific p44-expression locus was identified in A. phagocytophilum (Barbet et al., 2003; Lin et al., 2003), which consists of four tandem genes, tr1, omp-1X, omp-1N (p44ESup1) and p44E. The polycistronic transcriptional start site was detected 20 bp upstream of tr1 (Lin et al., 2003). tr1 has been predicted to possess a typical −35/−10 σ70 promoter (Lin et al., 2003). Barbet et al. showed that a -40 bp promoter region of the tr1 is capable of driving the expression of green fluorescent protein (gfp) reporter gene in a surrogate E. coli background (Barbet et al., 2005). Additional transcriptional start sites are present upstream of omp-1N, omp-1X, and p44E (Zhang et al., 2006; Barbet et al., 2005; Barbet et al., 2003; Lin et al., 2003).

P44 recombination

Previous studies showed a unique p44 expression locus (polymorphic p44 expression locus) was discovered downstream of three tandem genes (tr1, omp-1X, and omp-1N), from which diverse p44 genes have been proposed to be transcribed as a result of gene conversion (Barbet et al., 2003; Lin et al., 2003). Gene conversion at a single
$p44$-expression locus leads to P44 antigenic variation. Homologs of genes for the RecA-dependent RecF pathway, but not RecBCD or RecE pathways of recombination were detected in the *A. phagocytophilum* genome (Lin *et al*., 2006; Lin *et al*., 2003). The recombination intermediate structure between a donor $p44$ and the $p44$-expression locus of *A. phagocytophilum* was detected in HL-60 cell culture by Southern blot analysis followed by sequencing the band and in blood samples from infected SCID mice by PCR followed by sequencing (Lin *et al*., 2006). Using a double-origin plasmid carrying the $p44$-expression locus and a donor $p44$ locus, the recombination intermediate was recovered in an *E. coli* strain with active RecF recombination pathway, but not in strains with deficient RecF pathway (Lin *et al*., 2006). $p44$ recombination at the $p44$-expression was demonstrated in SCID mice and in cell culture, indicating that an acquired immune response is not essential for $p44$ recombination (Lin and Rikihisa, 2005).

**P44 neutralization**

The $p44$ multigene family encodes various surface exposed immunodominant 44 kDa major outer membrane protein P44s (Zhi *et al*., 1999). P44s are major antigens recognized by patients' sera (Zhi *et al*., 1999; Zhi *et al*., 1998). HGA patients, unless immunocompromised, generally develop antibodies to P44s; thus, P44s are considered useful antigens for serological diagnosis of HGA (Ijdo *et al*., 2002; Lin *et al*., 2002; Tajima *et al*., 2000; Zhi *et al*., 1998). Horses and mice experimentally infected with *A. phagocytophilum* also develop antibodies to P44s (Wang *et al*., 2004; Kim *et al*., 2002; Kim and Rikihisa, 1998). In experimentally infected horses sequential changeovers of expressed $p44$ transcripts occur in parallel with the development of anti-P44 antibodies.
Passive immunization of monoclonal antibodies against P44 inhibited the \textit{A. phagocytophilum} infection in mice (Kim and Rikihisa, 1998). Infected animals develop antibodies directed against the N-terminal conserved region as well as against the hypervariable region (Zhi \textit{et al.}, 2002b; Kim and Rikihisa, 1998). Two anti-Msp2 (P44) monoclonal antibodies (MAbs) and a recombinant Msp2 weakly block \textit{A. phagocytophilum} binding and infection of HL-60 cells (Park \textit{et al.}, 2003). Our MAb 3E65, obtained through screening by immunofluorescence followed by Western blot analysis, recognizes a linear epitope (Kim and Rikihisa, 1998). MAb 5C11 reacts with a linear epitope within the recombinant N terminal half of P44 N terminal conserved (Kim and Rikihisa, 1998; Zhi \textit{et al.}, 1998). Passive immunization with MAbs 5C11 and 3E65 partially protects naïve mice from infection with \textit{A. phagocytophilum} HZ (Kim and Rikihisa, 1998), indicating that P44 proteins contain at least two in vivo neutralizable B-cell epitopes.

**Role of P44 in pathogenesis**

\textit{A. phagocytophilum} is a bacterium that lacks pili, a capsule (Rikihisa, 1991), lipopolysaccharides, and peptidoglycans (Lin and Rikihisa, 2003a), suggesting that the outer membrane proteins play an important role in its interaction with host granulocytes. P44s were shown to be present in the isolated outer membrane fraction of \textit{A. phagocytophilum} by Western blot analysis and on the surface of \textit{A. phagocytophilum} within the inclusion by immunogold labeling of postembedded electron microscopy specimens (Kim and Rikihisa, 1998). Despite its small genome size (1.47 Mb) due to the ongoing reductive genome evolution among members of the family \textit{Anaplasmataceae}, the
*A. phagocytophilum* genome contains approximately 90 *p44* paralogues, suggesting that this large expansion of *p44* paralogues has given *A. phagocytophilum* a survival advantage, perhaps by allowing it to escape host immunoclearance. P44 proteins consist of a single central hypervariable region of approximately 94 amino acid residues, an N-terminal conserved region of approximately 186 amino acids, and a C-terminal conserved region of approximately 146 amino acids; the N- and C-terminal regions flank the central hypervariable region containing a signature of four conserved amino acid regions (C, C, WP, A) (Lin *et al.*, 2002; Zhi *et al.*, 1999). P44s undergo antigenic variation during infection in human granulocytic anaplasmosis patients (Barbet *et al.*, 2003). Our previous study also showed that P44s play a role in inducing the proinflammatory cytokines (Kim *et al.*, 2002; Kim and Rikihisa, 2000).

**Objectives of this study**

1. **Rapid Sequential Changeover of Expressed *p44* Genes during the Acute Phase of *Anaplasma phagocytophilum* Infection in Horses**

1) To investigate in the horse model of infection with *A. phagocytophilum* strain HZ, within-host dynamics of *p44* expression and the P44-18 variant-specific humoral immune response.

2) To determine the suppression of dominance of immuno-cross-reactive *p44* variants in cell culture by incubating host cell-free *A. phagocytophilum* with horse plasma followed by infection of HL-60 cells.
2. Two Monoclonal Antibodies with Defined Epitopes of P44 Major Surface Proteins Neutralize \textit{Anaplasma phagocytophilum} by Distinct Mechanisms

1) To define the two neutralization sites on P44 molecules by epitope peptide mapping and used the MAbs to delineate their bacterial surface exposure and inhibitory mechanisms of infection of host cells.

2) To perform ELISA analysis for the linear B- cell epitope with N terminal half of P44 N terminal conserved region (P44N-N) and P44-18 hypervariable region between two absolutely conserved amino acids Cysteine and Cysteine (P44-18hvC-C) of \textit{A. phagocytophilum}-infected horses and mice.

3. Delayed Development of Antibody to P44N and Upregulation of \textit{RecA} in Diverse \textit{P44} Expression and Delayed Clearance of \textit{Anaplasma phagocytophilum} in C3H/HeJ Mice

1) To study the persistence of bacteria in C3H/HeJ mice using \textit{A. phagocytophilum} HZ strain isolated from human patient.

2) To analyze the temporal development of antibodies against conserved P44N and P44hv C-C regions and \textit{p44} transcript populations during infection of mice.

3) To investigate how \textit{p44} recombination is involved in changes in \textit{p44} transcript population and \textit{A. phagocytophilum} persistence during mammalian infection.

4. Identification of a DNA-Binding Protein for Upstream Region of \textit{P44} Operon Encoding Major Outer Membrane Proteins of \textit{Anaplasma phagocytophilum}

1) To examine DNA binding proteins that may regulate polycistronic \textit{p44E} transcription by electrophoretic migration shift assay (EMSA).

2) To purify the DNA bound protein purification by streptavidin-magnetic beads.
Our results showed that the rapid and synchronized switch of expression is an intrinsic property of $p44$s, and the coordination necessary for this change might be provided by the host. The present results support the idea that recombination and large expansion of $p44$ paralogs allows $A. phagocytophilum$ to escape neutralizing antibodies. Future identification of the $A. phagocytophilum$ protein that binds to $cis$ element potentially involved in $tr1$ regulation would help in analyzing transcriptional regulation.

Thereby, the new knowledge generated in my Ph.D. study is expected to facilitate the understanding of the mechanisms of P44 antigenic variation: dynamic interplay between bacterial recombination and transcriptional regulation and immunoclearance. The results are expected to facilitate a new approach in designing vaccine candidates against HGA.
CHAPTER 2

RAPID SEQUENTIAL CHANGEOVER OF EXPRESSED P44 GENES
DURING THE ACUTE PHASE OF ANAPLASMA PHAGOCYTOPHILUM
INFECTION IN HORSES

2.1 Introduction

*Anaplasma phagocytophilum* is a tick-borne obligatory intracellular pathogen that causes persistent infections in various mammals (Dumler *et al.*, 2001). In humans, *A. phagocytophilum* can cause an acute systemic and potentially fatal disease, human granulocytic anaplasmosis (Bakken *et al.*, 2002; Dumler *et al.*, 2001). Antigenic variation is a survival strategy of the pathogen to enhance phenotypic variation within its hosts, prolonging duration of infection and the potential for transmission. Despite several mathematical models for the dynamics of antigenic variation (Gatton *et al.*, 2003; Antia *et al.*, 1996), the recent report of within-host dynamics of *var* gene expression by *Plasmodium falciparum* during the acute phase of human infection (Peters *et al.*, 2002) underscores the importance and paucity of experimental data. In the bovine intraerythrocytic agent, *A. marginale*, researchers demonstrated antigenic variation of MSP2 major surface proteins between peak rickettsemia that occurred at 6- to 8-week intervals (French *et al.*, 1999; French *et al.*, 1998). However, little is known about the
dynamics of antigenic variation within each rickettsemic cycle. The *A. phagocytophilum* genome encodes a large number of immunodominant major surface protein P44s (homologs of Msp2). Expression of diverse *p44* in human patients, horses, mice, and ticks has been documented (Felek *et al.*, 2004; Ijdo *et al.*, 2002; Lin *et al.*, 2002; Zhi *et al.*, 2002b). A unique *p44* expression locus (polymorphic *p44* expression locus) was discovered downstream of three tandem genes (*tr1*, *omp-1X*, and *omp-1N*), from which diverse *p44* genes have been proposed to be transcribed as a result of gene conversion (Barbet *et al.*, 2003; Lin *et al.*, 2002). Although *p44* have been hypothesized to go through antigenic variation like *msp2* of *A. marginale*, within-host dynamics of *p44* gene expression has not been demonstrated due in part to difficulty in monitoring switching behavior in a single infected human or in a laboratory mouse model of infection.

The horse is not only the natural host of *A. phagocytophilum*, but it serves as a useful animal model for human granulocytic anaplasmosis (Kim *et al.*, 2002; Zhi *et al.*, 2002b). The HZ strain, a human isolate of *A. phagocytophilum*, was well characterized in previous infection studies in horses, mice, ticks, and in cell culture (Lin *et al.*, 2003; Zhi *et al.*, 2002a; Zhi *et al.*, 2002b; Zhi *et al.*, 1999). This strain predominantly expresses *p44-18* in cell culture at 37°C and in the early stages of infection in the blood of horses and mice (Zhi *et al.*, 2002b). Therefore, in the present study, in the horse model of infection with *A. phagocytophilum* strain HZ, within-host dynamics of *p44* expression and the P44-18 variant-specific humoral immune response were investigated. Furthermore, since *A. phagocytophilum* is readily cultivable using the human promyelocytic leukemia HL-60 cell line, in contrast to *A. marginale* (French *et al.*, 1999; French *et al.*, 1998), it may provide the opportunity for in vitro investigation of antigenic
variation. Therefore, we also determined the suppression of dominance of immuno-cross-reactive p44 variants in cell culture by incubating host cell-free *A. phagocytophilum* with horse plasma followed by infection of HL-60 cells. Expressed p44 compositions were determined through a newly developed variant-specific probe hybridization method. Our results suggest that the rapid and synchronized switch of expression is an intrinsic property of p44s, and the coordination necessary for this change might be provided by the host.

2.2 Materials and methods

*A. phagocytophilum and horse infection.* The *A. phagocytophilum* HZ strain was cultured in HL-60 cells as previously described (Rikihisa *et al.*, 1997). Specimens from three infected horses were analyzed. Horse EQ005, a 5-year-old male, was infected by attaching 89 laboratory-reared *Ixodes scapularis* adult ticks infected with *A. phagocytophilum* as nymphs. A 12-year-old male horse (EQ006) was inoculated intravenously (iv.) with 10⁷ HL-60 cells infected with *A. phagocytophilum*. Horse EQ002 was infected by tick attachment in the previous study (Zhi *et al.*, 2002b). All horses were confirmed to be seronegative for *A. phagocytophilum* by the indirect fluorescent antibody test (Rikihisa *et al.*, 1997) prior to tick placement or iv. inoculation. On various days post-tick placement (p.t.) or post-iv. inoculation (p.i.), blood samples (50 to 500 ml) were collected from the jugular vein into EDTA tubes or acid citrate dextrose anticoagulant bottles, and plasma and peripheral blood leukocytes (PBLs) were prepared as previously described (Kim *et al.*, 2002).

**C-PCR.** Total DNA was extracted from the PBLs with a QIAamp blood kit
To determine levels of *A. phagocytophilum* organisms in the horse blood, a sensitive *p44* competitive PCR (C-PCR) assay was developed to amplify *p44* paralogs using primer set 1 (Table 2.1) (Zhi *et al.*, 1999). A 463-bp competitor for *p44* homologs was prepared by ligating the two PCR products with primer sets 2 and 3 (Table 2.1), using *A. phagocytophilum* HZ chromosomal DNA as the template after digestion of *Bam*HI sites included in primers pCompI and pCompII (Table 2.1). Primer set 4 (Table 2.1) was used to amplify the *A. phagocytophilum* 16S rRNA gene, resulting in a 361-bp product, whereas a competitor for the 16S rRNA gene (H. Niu and Y. Rikihiisa, unpublished data) with the same primer pair yielded a 323-bp product. Densitometric analysis of PCR products was performed as previously described (Kim and Rikihiisa, 2000). To normalize the amount of PBL DNA across samples, PCR amplification for the horse β-actin gene was performed with primer set 5 in Table 2.1 over a linear range, as described previously (Kim and Rikihiisa, 2000). The number of PBL cells was estimated based on the predetermined ratio of cell number to the density of the β-actin gene PCR product.

**Sequence analysis of *p44* cDNA and the polymorphic *p44* expression locus.**

The RNasea kit (QIAGEN) was used to extract total RNA from PBLs of each horse and from *A. phagocytophilum*-infected HL-60 cells. The RNA was treated with DNase I (Invitrogen, Carlsbad, Calif.), and cDNA was synthesized using SuperScript II reverse transcriptase (Invitrogen) with random hexamer (Invitrogen) as a primer, as described previously (Lin *et al.*, 2002). The RNA sample without reverse transcriptase treatment served as a negative control. PCR was performed with primer set 1 (Table 2.1) to amplify cDNA encoding *p44* paralogs, and the resulting cDNA was cloned into the pCRII vector.
by using a TA cloning kit (Invitrogen) for sequencing the insert.

To determine the sequence of the polymorphic *p44* expression locus in the DNA specimen, PCR with primer set 6 (Table 2.1) derived from the 5′-upstream region sequence of the polymorphic *p44* expression locus and the 3′-end conserved region primers of *p44* paralogs (Lin et al., 2003; Zhi et al., 1999) was performed, yielding a ~1,230-bp product. Amplified PCR products were cloned. Each insert was sequenced by the dideoxy termination method using the ABI Prism BigDye terminator v3.0 cycle sequencing reaction kit and an ABI 310 or 3730 sequencer. The deduced amino acid sequences were aligned using the CLUSTALV method in the MegAlign program (DNAStar, Madison, Wis.). Phylogenetic analysis was performed with the PHYLIP software package (version 3.6).

**Cloning and expression of rhvP44-18.** Primer set 9 (Table 2.1) was designed to clone the DNA fragment encoding the unique P44-18 hypervariable (hv) region. The resulting PCR products were digested with *Eco*RI and *Hind*III and ligated into the *Eco*RI- and *Hind*III-digested pET33b(+) vector (Novagen Inc., Madison, Wis.), and the plasmid was amplified in *Escherichia coli* Novablue cells (Novagen). *E. coli* BL21 (DE3) cells (Novagen) were transformed with the recombinant plasmid and induced to express the recombinant P44-18 hypervariable region (rhvP44-18) with isopropyl-thio-β-D-galactoside. The recombinant protein was affinity purified with a His-Bind Quick column (Novagen).

**Western blot analysis.** *A. phagocytophilum* organisms purified as previously described (Zhi et al., 1998) and rhvP44-18 were subjected to sodium dodecyl sulfate-15% polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a nitrocellulose
membrane. The membrane was incubated with either 5C11 (against pan-P44) or 3E65 (against hvP44-18) monoclonal antibodies (MAbs) (Kim and Rikihisa, 1998), or with plasma samples from each horse that were preabsorbed with *E. coli* BL21(DE3) lysates and diluted to 1:50. The peroxidase-conjugated affinity-purified anti-mouse or anti-horse immunoglobulin G secondary antibodies (Kirkegaard & Perry Laboratories Inc., Gaithersburg, Md.) were used at a dilution of 1:1,000. Finally, colorimetric reactions were developed with a 4-chloro-naphthol substrate solution (Sambrook and Russell, 2001).

**Analysis of p44 species expressed in cell culture after incubation with horse plasma.** All horse plasma and sera were filtered through 0.22-µm mixed cellulose ester membrane filters (Fisher Scientific, Pittsburgh, Pa.) and inactivated by incubation at 56°C for 30 min. Host cell-free *A. phagocytophilum* was freshly prepared from 2 x 10⁶ *A. phagocytophilum*-infected HL-60 cells (>90% cells infected) per treatment, as previously described (Yoshiie *et al.*, 2000). The organisms were suspended in 1 ml of culture medium and incubated at 4°C for 30 min with 0.5 ml of plasma collected from horse EQ005 either on day 0 (before tick placement) or day 22 or 31 p.t. Control samples were incubated with 0.5 ml of heat-inactivated fetal bovine serum (FBS) or RPMI 1640 medium. Each of the mixtures was added to 10⁶ uninfected HL-60 cells in 0.5 ml of RPMI 1640 medium supplemented with 5% FBS and 2 mM L-glutamine (final concentration, 5 x 10⁵ cells/ml) and incubated at 37°C in 95% air-5% CO₂. After 12 h, the medium was replaced with fresh 5% FBS-RPMI medium, and the medium was replaced every 2 days thereafter. About 10⁶ infected cells were harvested from each sample 2.5, 6, or 12 days postculture (p.c.) and immediately preserved in RNAlater (QIAGEN). p44
transcripts were amplified by reverse transcription-PCR (RT-PCR), and the polymorphic
*p44* expression locus was amplified by DNA PCR as described above and cloned.

**Colony hybridization to detect p44-18 and total p44 cDNA and DNA clones.**

*E. coli* colonies harboring pCRII vector with *p44* homolog inserts were randomly picked
and plated onto two nitrocellulose membranes placed on Luria-Bertani plates. After
incubation at 37°C overnight, the membranes were peeled off and treated with denaturing
solution and UV cross-linked at 1,200 µJ for 50 s with a Stratalinker UV cross-linker
(Stratagene Cloning Systems, La Jolla, Calif.) (Sambrook and Russell, 2001). The *p44-18*
hv region probe (*p44-18*hv; 123 bp) was prepared by PCR using primer set 7 (Table 2.1)
and digoxigenin labeled with a DIG DNA labeling and detection kit (Roche Molecular
Biochemicals, Mannheim, Germany). The pan-*p44*-specific probe was prepared by
amplification of a C-terminal fragment from the conserved *p44* region (102 bp) with
primer set 8 (Table 2.1) and digoxigenin labeled. The pan-*p44* and *p44-18*hv probes were
hybridized to each of the duplicate membranes, and hybridization was detected according
to the manufacturer's instructions.

**Complexity analysis.** To quantify complexity, the Shannon entropy (\(H\))
calculation (Stewart *et al.*, 1997; Shannon, 1948) was applied. Shannon entropy
incorporates both the number of *p44* species and the number of cloned cDNAs in each
*p44* species. It is defined as

\[
H = - \sum_{i=1}^{N} P(i) \ln[P(i)],
\]

where \(N\) is the total number of *p44* species and \(P(i)\) is the number of clones represented in each *p44* species. A
normalized value of \(H\), denoted \(H'\), was defined as \(H/\ln(N)\). This resulted in a range of
possible values for \(H'\) from 0 to 1, representing 1 to \(N\) distinct *p44* species, respectively.

\(H'\) values of specimens incubated in immune horse plasma and preimmune horse plasma
were compared with a paired Student's t test. A P value of <0.05 was considered significant.

**Nucleotide sequence accession numbers.** GenBank accession numbers of sequences newly identified in the present study are as follows: p44-43, AY147269; p44-46, AY147263; p44-47, AY147264; p44-48, AY147267; p44-49, AY147268; p44-58, AY279320; p44-59, AY279321; p44-60, AY279319. The remaining p44 sequences and GenBank numbers are described elsewhere (Felek et al., 2004; Lin et al., 2002; Zhi et al., 2002b).

### 2.3 Results

**Logarithmic rise and decline in rickettsemia in horses.** In order to determine the pattern of initial rickettsemia in naive hosts, we developed a p44 C-PCR assay to detect low levels of infection. We chose C-PCR over other PCR methods because a competitor coexists with the target DNA in the same reaction tube and serves as an internal control for variation of each PCR. The p44 gene is composed of a central hv region of approximately 280 bp flanked by conserved sequences ranging from 100 to 500 bp in length (Lin et al., 2003; Lin et al., 2002; Zhi et al., 1999). Primer set 1 (Table 2.1), which hybridizes to the conserved 5' and 3' regions, was used to amplify a group of p44 paralogs in the genome of *A. phagocytophilum*. In order to estimate the number of organisms detected by p44 C-PCR, we also developed C-PCRs specific for the 16S rRNA gene, which is a single copy within the *A. phagocytophilum* genome (www.tigr.org). Both p44 paralog and 16S rRNA gene C-PCRs were performed with the same amount of *A. phagocytophilum* chromosomal DNA to determine the ratio of p44 paralogs to the 16S rRNA gene. The finding from C-PCR indicated that this ratio was 48 under the assay
conditions used. Thus, the number of \textit{p44} paralogs detected by \textit{p44} C-PCR divided by 48 represents the approximate \textit{A. phagocytophilum} genome equivalent. The amount of horse PBLs in each specimen was normalized based on the horse \textit{ß}-actin gene, as described previously (Kim and Rikihisa, 2000). Figure 2.2 shows similar levels and a similar monotopic, but not oscillating, pattern of rickettseemia: an initial logarithmic increase followed by a precipitous decline during 1 month p.t. in EQ005 and p.i. in EQ006.

\textbf{Sequential appearance of dominant \textit{p44} transcript species.} To determine the within-host dynamics of \textit{p44} transcript species, RNA specimens from the PBLs from three experimentally infected horses were subjected to \textit{p44}-specific RT-PCR followed by sequencing of the \~{}550-bp amplicon as previously described (Zhi \textit{et al.}, 2002b). Two horses, EQ005 and EQ002, were infected by tick transmission of the HZ strain. One horse, EQ006, was infected by i.v. inoculation of the cultured HZ strain. No amplicon was detected in any of the samples that lacked reverse transcriptase, indicating that the fragments were not due to amplification of contaminating DNA (data not shown). A total of 249 \textit{p44} cDNA clones (20 to 35 clones for each horse at each time point) were randomly selected for sequencing after cloning of RT-PCR products. \textit{p44} species can be identified based on the sequences of the central hv regions (Felek \textit{et al.}, 2004; Lin \textit{et al.}, 2002; Zhi \textit{et al.}, 2002b). A previous study showed that the HZ strain predominantly expresses \textit{p44-18} in cell culture and day 8 postinfection in the blood of both tick- and i.v.-infected horses (Zhi \textit{et al.}, 2002b). \textit{p44} transcripts could not be detected in the blood of tick-infected EQ005 on day 8 p.t., perhaps due to the very low level of \textit{A. phagocytophilum} infection (1.6 rickettsiae/10\textsuperscript{4} PBLs, based on \textit{p44} C-PCR) (Fig. 2.2). On day 12 p.t., all \textit{p44} transcripts detected were \textit{p44-18}. On day 17 p.t., \textit{p44-30}, \textit{p44-2}, and
other new transcript species emerged. On day 22 p.t., the dominance of the \textit{p44-18} transcript was replaced with new dominant \textit{p44} transcript species (\textit{p44-30} and \textit{p44-2}). On day 31 p.t., \textit{p44-30} and \textit{p44-2} were replaced with the new dominant \textit{p44} transcript species (\textit{p44-55} and \textit{p44-16}) (Fig. 2.3A). In horse EQ002, which was also infected by tick transmission, \textit{p44-18} was the major transcript species detected on day 8 p.t. (Zhi et al., 2002b). When we analyzed RNA specimens from this horse that were preserved in the –80°C freezer, we found that on day 16 p.t. the dominance of the \textit{p44-18} transcript was replaced with a new transcript species, \textit{p44-1} (25 of 25 cDNA clones). In i.v.-inoculated horse EQ006, results for days 6, 14, 18, and 22 p.i. are shown (Fig. 2.3B), since \textit{p44} transcript species changeover was faster in this horse than in EQ005 and on days 30 and 37 p.i. \textit{p44} transcripts could not be detected in the blood, perhaps due to the very low level of \textit{A. phagocytophilum} infection (below 3.5 rickettsiae/10^5 PBLs, based on \textit{p44} C-PCR) (Fig. 2.2). Nonetheless, in EQ006 a similar changeover pattern of the initially dominant \textit{p44-18} transcript species to that of tick-transmitted EQ005 was observed, i.e., traveling waves of sequential population changeovers of the \textit{p44} transcript species within a single peak of rickettsemia.

To determine the similarity among these sequentially dominant \textit{p44} transcript species, phylogenetic analysis was performed by defining the dominant \textit{p44} transcript species as those that represented greater than 20% of the total number of cDNA clones at each time point for each horse sample. Investigators from our laboratory previously showed that hv region sequences of \textit{p44} paralogs are clustered into three major groups (\textit{α}, \textit{β}, and \textit{γ}) (Lin et al., 2002). \textit{p44-18}, which is located in the \textit{α} cluster, was the dominant transcript initially detected in all six horses infected with \textit{A. phagocytophilum} HZ,
including EQ001 to EQ004 in the previous study (Zhi et al., 2002b), regardless of whether inoculation was i.v. or through ticks. All dominant \( p44 \) transcript species in horses EQ005 and EQ006 that subsequently appeared within the initial rickettsemia period belonged to the \( \beta \) or \( \gamma \) clusters. The \( p44-1 \) transcript detected in horse EQ002 on day 16 p.t. also belonged to the \( \beta \) cluster. These results suggest suppression of dominance of the \( \alpha \) cluster \( p44 \) transcript species during subsequent \( p44 \) transcript species changeovers by immunologic cross-reactivity.

To investigate whether the sequential appearance of dominant \( p44 \) transcript species is due to sequential change of the bacterial genetic population having corresponding \( p44 \) species in the polymorphic \( p44 \) expression locus, \( p44 \) genes in the locus were cloned by DNA PCR. DNA fragments were amplified with primer set 6 (Table 2.1) from EQ005 PBL specimens obtained on days 17 and 22 p.t. PBL specimens obtained on days 12 and 31 p.t. did not yield significant products due to the insufficient levels of \( A. \) phagocytophilum DNA present in these specimens. The resulting PCR products were cloned, and 13 and 12 clones from days 17 and 22 p.t., respectively, were randomly selected for DNA sequencing. The DNA clones in the PBL specimen from day 17 p.t. were \( p44-18 \) (11 clones), \( p44-30 \) (1 clone), and \( p44-2 \) (1 clone). The major DNA clones in the PBL specimen from day 22 p.t. were \( p44-2 \) (6 clones), \( p44-30 \) (5 clones), and \( p44-18 \) (1 clone). The major \( p44 \) cDNA species roughly correlated with the \( p44 \) species present within the \( p44 \) polymorphic expression locus, suggesting that sequential appearance of dominant \( p44 \) transcript species is due to sequential recombination at the \( p44 \) polymorphic expression locus and growth of the newly recombined population.

**Development of P44 variant-specific antibodies during the course of \( A. \)**
**Phagocytophilum infection.** To explore the possibility that the suppression of dominance of the p44 transcript species is associated with immunologic cross-reactivity, we determined the kinetics of P44 variant-specific antibody. Since P44-18 is an initially dominant P44 variant, the DNA fragment encoding the P44-18 unique hv region (83 amino acids), based on the alignment of all p44 paralogs, was cloned into the pET33b (+) vector. *E. coli* transformed with the plasmid expressed a 124-amino-acid (13,096-Da) fusion protein, including His6 residues at its N terminus. The expressed rhvP44-18 was purified through Ni-affinity chromatography, and a single band of approximately 13 kDa was detected with SDS-PAGE (Fig. 2.4A). Western blot analysis revealed that rhvP44-18 specifically reacted with MAb 3E65, which binds only to the 44-kDa protein expressed by strain HZ in HL-60 cell culture and not to 44-kDa proteins expressed by the other five different *A. phagocytophilum* strains in HL-60 cells (Kim and Rikihisa, 2000). rhvP44-18 did not react with MAb 5C11, which is specific to the invariable N-terminal regions of P44s (Zhi et al., 2002b; Kim and Rikihisa, 1998) (Fig. 2.4B). Immunoreactivity to native P44s was first detectable in the plasma at day 8 p.t. in EQ005 and day 10 p.i. in EQ006. Anti-rhvP44-18 immunoglobulin G became detectable in the plasma of EQ005 and EQ006 starting on day 17 p.t. and day 10 p.i., respectively (Fig. 2.4C). This was roughly the time point when the p44-18 transcript population declined in each horse. This result also confirmed the expression of the P44-18 protein by *A. phagocytophilum* in horses and recognition of the hv region by the horse immune system.

**Infected horse plasma reduced dominance of the p44-18 variant in cell culture.** To test whether the infected horse plasma was involved in the suppression of p44-18 transcript dominance, host cell-free *A. phagocytophilum* was preincubated with
the plasma from infected horses and then the mixture was coincubated with HL-60 cells. Since there is no efficient and reliable method to determine a specific \( p44 \) transcript population in a large number of specimens, we developed the colony hybridization method. In this method, we verified 100 \( p44 \) cDNA clones using the pan-\( p44 \) probe and determined the percentage of cDNA clones which had the \( p44-18 \) insert among \( p44 \) cDNA clones by using the \( p44-18 \) species-specific probe (Fig. 2.5A). Two cDNA clones with \( p44-18 \) and \( p44-40 \) inserts (confirmed by sequencing) were used as positive and negative controls, respectively, in each hybridization reaction. PCRII vector plasmid with a non-\( p44 \) insert was also used as a negative control. An advantage of this method over real-time PCR, for example, is that it is not influenced by primer sensitivity, plasmid stability, or variation of PCR. The variation in this assay in duplicate specimens was less than 10%. Using this method it was found that when incubated with infected horse plasma (EQ005 day 22 or 31 p.t.) the dominance of the \( p44-18 \) transcript rapidly declined in vitro (Fig. 2.5). The inhibitory effect was specific to infected horse plasma, since incubation with EQ005 day 0 (pre-tick attachment) plasma, FBS, or RPMI medium did not reduce the dominance of the \( p44-18 \) transcript (Fig. 2.5).

The colony hybridization method was confirmed by sequencing a total of 226 \( p44 \) cDNA clones (21 to 29 clones each per culture at each time point for two independent experiments) at 2.5, 6, and 12 days p.c. (Table 2.6). At 2.5 days p.c. with day 31 p.t. horse plasma samples, \( p44-18 \) represented 58% of the total clones and \( p44-13 \) (3 clones), \( p44-47 \) (2 clones), \( p44-29 \) (2 clones), \( p44-15 \) (2 clones), \( p44-32 \) (1 clone), and \( p44-49 \) (1 clone) were also detected in 24 cDNA clones. Whereas, when incubated with day zero plasma \( p44-18 \) represented 86% of all clones and only \( p44-2 \) (1 clone), \( p44-16 \) (1 clones),
and $p44-32$ (1 clone) were also detected in 21 cDNA clones. These four P44 variants were detected in horse blood by day 31 p.t. However, none of the cDNA clones were detected after incubation with day 31 plasma, except $p44-13$, which was a minor transcript at one time point (1 of 24 cDNA clones on day 17 p.t.) in the horse blood. We applied Shannon entropy analysis (Stewart et al., 1997; Shannon, 1948) to measure $p44$ species population complexity. With immune plasma the entropy value was significantly greater than that with preimmune plasma (Table 2.6).

Day 31 p.t. plasma from horse EQ005 also reduced the percentage of the *A. phagocytophilum* genetic population that contained $p44-18$ at the polymorphic $p44$ expression locus: relative to the culture incubated with day zero plasma from horse EQ005, the average percentage of the population with $p44-18$ at the polymorphic $p44$ expression locus was 20% (the average of 10 and 30%, obtained in two independent experiments). These results suggest that reduction of $p44-18$ transcript frequencies in vitro is also accompanied by genetic population change at this locus.
**Table 2.1 Sequences of oligonucleotides used in RT-PCR, PCR, colony hybridization, and cloning**

<table>
<thead>
<tr>
<th>Primer set no.</th>
<th>Primer name</th>
<th>Target</th>
<th>Purpose</th>
<th>Size (bp)</th>
<th>5’ primer</th>
<th>3’ primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>p3709-p4257</td>
<td>p44 paralogs</td>
<td>C-PCR, RT-PCR</td>
<td>~549</td>
<td>GTAAAGGGATGTTAGCTATGA</td>
<td>AGAAGATCAATAACAAGCATTTG</td>
</tr>
<tr>
<td>2</td>
<td>p3709- pCompI</td>
<td></td>
<td></td>
<td>252</td>
<td>GTAAAGGGATGTTAGCTATGA</td>
<td>CTGGGAATCCACTAGCTGCG</td>
</tr>
<tr>
<td>3</td>
<td>pCompII-p4257</td>
<td></td>
<td></td>
<td>221</td>
<td>CTGGGAATCCGTAGCTGAG</td>
<td>AGAAGATCAATAACAAGCATTTG</td>
</tr>
<tr>
<td>4</td>
<td>pComp16SI-pComp16SI</td>
<td>A. phagocytophilum 16S rRNA</td>
<td>C-PCR</td>
<td>323</td>
<td>CGGGGGAAGATTATACGTATTA</td>
<td>CGCTTGCCCCCTCGGTATTAG</td>
</tr>
<tr>
<td>5</td>
<td>HB-ACTIN5-HB-ACTIN3</td>
<td>horse β-actin</td>
<td>DNA PCR</td>
<td>628</td>
<td>CGTGCAACACCTCTACAACCGAG</td>
<td>GTACAAAGGCTCTACGGATGT</td>
</tr>
<tr>
<td>6</td>
<td>p5u37-p4257</td>
<td>p44 upstream sequence (p44 expression locus)</td>
<td>First step DNA PCR</td>
<td>~1232</td>
<td>CGTTATTGGTCTCAGAGAAAAG</td>
<td>AGAAGATCTAATAACAAGCATTTG</td>
</tr>
<tr>
<td>7</td>
<td>p18-5’-p18-3’</td>
<td>p44 paralog C-terminus</td>
<td>Colony hybridization</td>
<td>123</td>
<td>GTGATAGTAGCAGATGCTG</td>
<td>CATCACTAACGGTGTTAGAA</td>
</tr>
<tr>
<td>8</td>
<td>p44con- p5247</td>
<td>p44 paralog C-terminus conserved region</td>
<td></td>
<td>102</td>
<td>AGTAGCAGGGTTACTAGC</td>
<td>AGAAGATCAATAACAAGCATTTG</td>
</tr>
<tr>
<td>9</td>
<td>18RECS-18RECA</td>
<td>p44-18 hypervariable region</td>
<td>cloning</td>
<td>271</td>
<td>TCCGAATCCTGGAGAAGAAGC</td>
<td>CGCAAGCTCTACAGCTT</td>
</tr>
</tbody>
</table>

*a BamH I site is underlined.*
Fig. 2.2 Levels of *A. phagocytophilum* organisms in the blood of two horses during the course of infection, as determined by *p44* C-PCR. EQ005 was infected by tick attachment, and EQ006 was inoculated i.v.
Fig. 2.2

A. phagocytophilum chromosome equivalent/10^5 PBLs

- EQ005 infected by tick transmission
- EQ006 infected by IV inoculation

days PT and PI

0 5 10 15 20 25 30 35

10^6

10^5

10^4

10^3

10^2

10^1

10^0

A. phagocytophilum chromosome equivalent/10^5 PBLs
Fig. 2.3 Relative proportion of each of the *p44* transcript species in the blood of horses EQ005 (A) and EQ006 (B). The vertical axes show the percentage of cDNA clones of each *p44* species at each time point. The horizontal axes show the various *p44* cDNA clone species detected.
Fig. 2.3

A. EQ005

B. EQ006

dl day 12 PT

dl day 17 PT

% cDNA clone number

dl day 22 PT

% cDNA clone number

dl day 31 PT

% cDNA clone number

% cDNA clone number

% cDNA clone number

% cDNA clone number
Fig. 2.4 Development of anti-hvP44-18 antibody in the plasma of *A. phagocytophilum*-infected horses. (A) Three micrograms of affinity-purified hvP44-18 (AP-hvP44-18) was subjected to SDS-PAGE followed by Coomassie blue staining. M, molecular size marker. The recombinant hvP44-18 is indicated by the arrow. (B) Purified hvP44-18 (2.5 µg/lane) was separated by SDS-PAGE and blotted to the membrane. The membrane was incubated with the hvP44-18-specific MAb (3E65) or pan-P44-specific MAb (5C11). (C) Purified hvP44-18 (2.5 µg/lane) and *A. phagocytophilum* (3 µg/lane) were separated by SDS-PAGE and blotted to the membrane. The membrane was incubated with horse sera collected at the indicated days p.t. or p.i. Numbers on the right indicate molecular masses (in kilodaltons) based on the broad-range prestained standards (Bio-Rad).
Fig. 2.4

A. AP-rhp44-18

B. M (kDa)

C. EC005

rhp44-18

A. phago.

EQ006

rhp44-18

A. phago.

Days PT

Days PI
Fig. 2.5 Infected horse plasma reduced dominance of the p44-18 transcript in cell culture. (A) Colony hybridization of p44 cDNA clones from A. phagocytophilum HZ incubated with horse (EQ005) day zero plasma (pre-tick attachment) and day 22 plasma at 6 days p.c. hvp44-18 and pan-p44 probes each were hybridized to one of the duplicated membranes each with 100 p44 cDNA clones. Two cDNA clones with p44-18 and p44-40 inserts were used as a positive and a negative control, respectively. A pCRII plasmid with a non-p44 insert was also used as a negative control. Approximately 14% of p44 cDNA clones had the p44-18 insert with A. phagocytophilum incubated with day 22 p.t. plasma, and approximately 80% of p44 cDNA clones had the p44-18 insert with those incubated with day zero plasma. (B) Temporal changes in the p44-18 transcript population in A. phagocytophilum incubated with plasma obtained on day zero (preinfection) and day 31 p.t. from EQ005, FBS, or RPMI 1640 medium. Colony hybridization analysis was performed on 100 p44 cDNA clones in each specimen at each day p.c. to determine the percentage of cDNA clones with p44-18 inserts.
Fig. 2.5

A

day 0 horse plasma

PCRII vector with non-p44 insert

PCRII vector with non-p44 insert

PCRII vector with non-p44 insert

PCRII vector with non-p44 insert

hvp44-18 probe

pan p44 probe

day 22 horse plasma

B

Percentage of p44-18 cDNA clone/100 clones

Days PC

- 0 day EQ005 plasma
- 31 day EQ005 plasma
- Fetal bovine serum
- RPMI 1640
Table 2.6 Characteristics of \textit{p44} cDNA clonetype distribution after incubation of \textit{A. phagocytophilum} with horse EQ005 day 31 and day 0 plasma

<table>
<thead>
<tr>
<th>Days PC</th>
<th>horse plasma</th>
<th>\textit{p44} clone type distribution$^b$</th>
<th>No. of clonotypes$^c$</th>
<th>Entropy$^{d,e}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>day 31</td>
<td>14, 3, 2, 2, 1, 1, 1, 1</td>
<td>7</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>day 0</td>
<td>18, 1, 1, 1</td>
<td>4</td>
<td>0.41</td>
</tr>
<tr>
<td>6</td>
<td>day 31</td>
<td>10, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1</td>
<td>13</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>day 0</td>
<td>19, 2, 1, 1, 1, 1, 1, 1</td>
<td>9</td>
<td>0.58</td>
</tr>
<tr>
<td>12</td>
<td>day 31</td>
<td>10, 4, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1</td>
<td>10</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>day 0</td>
<td>20, 2, 1, 1, 1</td>
<td>5</td>
<td>0.48</td>
</tr>
<tr>
<td>12$^a$</td>
<td>day 31</td>
<td>6, 4, 3, 3, 3, 3, 3, 2, 2, 1, 1, 1</td>
<td>11</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>day 0</td>
<td>18, 1, 1, 1, 1, 1</td>
<td>7</td>
<td>0.52</td>
</tr>
</tbody>
</table>

$^a$ Independent experiment  
$^b$ Each series shows the number of clones comprising each clonetype for the given sample  
$^c$ Number of sequence indistinguishable cloned cDNAs  
$^d$ Normalized Shannon entropy (H'), a calculated value that incorporated the number and distribution of clone types  
$^e$ A paired $t$-test was used to compare the difference of entropy between samples of day 31 and day 0 horse plasma, $P<0.003$
2.5 Discussion

This study is the first, to our knowledge, to report the within-host dynamics of transcribed \( p44 \) genes during the acute phase of \( A. phagocytophilum \) infections. During the logarithmic increase of rickettsiae in the blood, the rapid switch-off of the initial dominant transcript \( p44-18 \) occurred regardless of whether the bacterium was transmitted by ticks or by i.v. inoculation. \( p44-18 \) was not expressed by \( A. phagocytophilum \) strain HZ in the salivary gland of ticks that transmitted \( A. phagocytophilum \) to naïve horses (Zhi et al., 2002b). Thus, this result suggests timed or programmed \( p44-18 \) expression after transmission to mammalian hosts and reinitiation of rapid switch-off of \( p44-18 \) expression. This is similar to the \( var \) gene expression by \( P. falciparum \) reported during the acute phase in two human volunteer infections (Peters et al., 2002). Although in the \( P. falciparum \) study subsequent changes in the same individual were not investigated, in our study two subsequent population changeovers in the \( p44 \) transcripts were observed in each horse: the second and third synchronized waves of dominant \( p44 \) transcripts appeared during peak rickettsemia and the declining phase of rickettsemia, respectively, suggesting that the switch is not dependent on rickettsial density. It is unclear how the entire population of \( A. phagocytophilum \) within the host can synchronize \( p44 \) expression. These changeovers do not resemble the simple clearance of the previous \( p44 \) variant population by the immune system and the rapid growth of new \( p44 \) escape variants, since the characteristic oscillating pattern of rickettsemia was not evident during these changeovers.

Each of the subsequently dominant \( p44 \) transcripts was phylogenetically dissimilar from \( p44-18 \). Concomitant development of an immune response to the hv
region of P44-18 suggests that antibodies against the \textit{p44} hv regions in the previously expressed \textit{p44}s prevent dominance of similar \textit{p44} transcripts, and thus shape the next \textit{p44} transcript population in infected horses. Our work suggests that immune responses to multiple \textit{p44} genes with diverse antigenic properties also shape the \textit{A. phagocytophilum} genetic population, since the changeover of \textit{p44} transcript species was accompanied by a corresponding changeover of \textit{p44} species at the polymorphic \textit{p44} expression locus.

The present study is the first to demonstrate that immune plasma contributes to changes in the phenotype and the genotype of the \textit{A. phagocytophilum} population in cell culture. This supports the hypothesis that antibodies specific to the P44 hv region are responsible for preventing the reemergence of \textit{A. phagocytophilum} cells that have a similar \textit{p44} hv region at the polymorphic expression locus. Furthermore, the emergence of more diverse \textit{p44} transcript species in the presence of immune plasma suggests that the immune plasma not only changes transcription levels but also promotes \textit{p44} recombination. The data also showed that a 12-h treatment with heat-inactivated immune horse plasma is sufficient to cause a continuous decline of the \textit{p44}-18 transcript population for up to 12 days (corresponding to approximately 40 bacterial fission events, at an estimated doubling time of 7 h). Thus, this \textit{p44} transcript changeover appears to be an intrinsic property of \textit{p44}s which is rapidly reinitiated and sustained when introduced into mammals or exposed to immune plasma, but not by some chance event. The in vitro system developed in this study would provide a useful and simple model for studying mechanisms of \textit{p44} antigenic variation, as an alternative to the use of experimental animals.
The present and previous studies showed that low levels of bacteria persisted in the blood of horses that were infected with *A. phagocytophilum* HZ by either tick attachment or by the i.v. route for more than 1 month. *A. phagocytophilum* can cause persistent infections in ruminants and horses (Rikihisa, 1991). This bacterium was detected by PCR in the serum of an untreated human 30 days after the onset of illness (Lepidi *et al.*, 2000). The persistence of *A. phagocytophilum* in a reservoir rodent host would be an important adaptation which allows greater access for uninfected tick populations to an infectious blood meal. The present study suggests that dynamic interactions with the host innate and adaptive immune systems orchestrate a prolonged *A. phagocytophilum* infection by spreading the expansion of different *p44* antigenic variants over time to avoid rapidly exhausting the *p44* repertoire.
2.6 Summary

*A. phagocytophilum* immunodominant polymorphic major surface protein P44s have been hypothesized to go through antigenic variation, but the within-host dynamics of *p44* expression has not been demonstrated. In the present study we investigated the composition and changes of *p44* transcripts in the blood during the acute phase of well-defined laboratory *A. phagocytophilum* infections in naïve equine hosts. Three traveling waves of sequential population changeovers of the *p44* transcript species were observed within a single peak of rickettsemia of less than 1 month. During the logarithmic increase, the rapid switch-off of the initial dominant transcript *p44-18* occurred regardless of whether the bacterium was transmitted by ticks or by intravenous inoculation. Each of the subsequently dominant *p44* transcript species was phylogenetically dissimilar from *p44-18*. Development of antibody to the hypervariable region of P44-18 during the rickettsemia suggests the suppression of dominance of immuno-cross-reactive *p44* populations. When *A. phagocytophilum* was preincubated with plasma from the infected horse and then coincubated with HL-60 cells, the dominance of the *p44-18* transcript was rapidly suppressed in vitro and most of the newly emerged *p44* transcript species were previously undetected in this horse. This work provides experimental evidence of within-host *p44* antigenic variation. Results suggest that the rapid and synchronized switch of expression is an intrinsic property of *p44s* reinitiated after transmission to naïve mammalian hosts and shaped upon exposure to immune plasma.
CHAPTER 3

TWO MONOCLONAL ANTIBODIES WITH DEFINED EPITOPEs OF P44 MAJOR SURFACE PROTEINS NEUTRALIZE ANAPLASMA PHAGOCYTOPHILUM BY DISTINCT MECHANISMS

3.1 Introduction

Human granulocytic anaplasmosis (formerly human granulocytic ehrlichiosis) is an emerging tick-borne zoonosis that has been reported in the United States and Europe (van Dobbenburgh et al., 1999; Petrovec et al., 1997; Bakken et al., 1994). Human granulocytic anaplasmosis is caused by infection with an obligatory intracellular bacterium, *Anaplasma phagocytophilum*, in the family Anaplasmataceae. This bacterium lacks pili, a capsule (Rikihisa, 1991), lipopolysaccharides, and peptidoglycans (Lin and Rikihisa, 2003a), suggesting that the outer membrane proteins play an important role in its interaction with host granulocytes. A family of proteins with molecular sizes of the 44-kDa range (P44s, also called Msp2s) was shown to be present in the isolated outer membrane fraction of *A. phagocytophilum* by Western blot analysis and on the surface of *A. phagocytophilum* within the inclusion by immunogold labeling in the postembedded electron microscopy specimens (Kim and Rikihisa, 1998).

P44 proteins are encoded by the polymorphic *p44 (msp2)* multigene family.
Despite its small genome size (1.47 Mb) due to the ongoing reductive genome evolution among members of the family Anaplasmataceae, the A. phagocytophilum genome contains approximately 90 p44 paralogues, suggesting that this large expansion of p44 paralogues has given A. phagocytophilum a survival advantage, perhaps by allowing it to escape host immunoclearance. P44 proteins consist of a single central hypervariable region of approximately 94 amino acid residues, an N-terminal conserved region of approximately 186 amino acids, and a C-terminal conserved region of approximately 146 amino acids; the N- and C-terminal regions flank the central hypervariable region (Lin et al., 2002; Zhi et al., 1999). There are three short conserved segments including two absolutely conserved cysteines within the hypervariable region of all predicted P44 proteins (Lin et al., 2002).

Infected animals develop antibodies directed against the N-terminal conserved region as well as against the hypervariable region (Wang et al., 2004; Zhi et al., 2002b; Kim and Rikihisa, 1998). P44s undergo antigenic variation during infection in human granulocytic anaplasmosis patients and in experimentally infected horses (Wang et al., 2004; Barbet et al., 2003). The hypervariable region of P44 molecules has been assumed to be exposed on the bacterial surface and involved in antigenic variation and immune evasion (Wang et al., 2004; Barbet et al., 2003; Lin et al., 2002; Zhi et al., 1999; Kim and Rikihisa, 1998). However, since epitopes of anti-P44 antibodies have never been defined, whether or which part of the hypervariable region or any other regions of naturally folded P44 molecules is exposed to the surface of the intact bacterium has been unknown.

Human granulocytic anaplasmosis patients, unless immunocompromised,
generally develop antibodies to P44s; thus, P44s are considered useful antigens for serological diagnosis of human granulocytic anaplasmosis (Ijdo et al., 2002; Lin et al., 2002; Tajima et al., 2000; Zhi et al., 1998). Horses and mice experimentally infected with *A. phagocytophilum* also develop an antibody to P44s (Wang et al., 2004; Kim et al., 2002; Kim and Rikihisa, 1998). It is less clear whether antibodies to P44s are protective from infection. Ijdo et al. (Ijdo et al., 2002) reported lack of protection on day 15 postchallenge in mice immunized with a recombinant P44 protein. Two anti-Msp2 (P44) monoclonal antibodies (MAbs) and a recombinant Msp2 only weakly block *A. phagocytophilum* binding and infection of HL-60 cells (Park et al., 2003). The passive immunization of naïve mice with MAbs directed against P44s partially protects mice from infection (Kim and Rikihisa, 1998). The results of these studies have given an overall impression that antibodies directed to P44 (Msp2) do not have a significant role in immunoprotection.

However, the previous studies did not define epitopes bound by the MAbs or the epitopes bound by antibodies developed by immunization with the recombinant P44 protein. Furthermore, the *p44* species predominantly expressed by the *A. phagocytophilum* population used to infect the mice or HL-60 cells were not defined in previous work. Thus, it is unclear whether this poor protection in mice or HL-60 cells is simply due to (i) poor neutralization ability of particular anti-P44 antibodies involved, (ii) lack of surface exposure of the target epitope on the intact bacteria, or (iii) epitope mismatch between anti-P44 antibodies and P44 proteins expressed by the organisms used for infection.

Our MAb 3E65 was obtained through screening by immunofluorescence followed
by Western blot analysis (Kim and Rikihisa, 1998) and recognizes a linear epitope within the recombinant hypervariable region of P44-18 protein (Wang et al., 2004). MAb 5C11 reacts with a linear epitope within the recombinant partial P44-1 protein, which consists of most of the conserved N-terminal region and a part of the hypervariable region of P44-1 (Kim and Rikihisa, 1998; Zhi et al., 1998), with the *A. phagocytophilum* HZ strain cultured in HL-60 cells at 37°C, which expresses various *p44* s, but not *p44-1* (Zhi et al., 1999), and with diverse P44s derived from several other strains of *A. phagocytophilum* so far examined (Kim and Rikihisa, 1998). Thus, the MAb 5C11 epitope has been considered to be within the conserved P44 N terminus, but not within the hypervariable region of P44-1.

Passive immunization with MAbs 5C11 and 3E65 partially protects naïve mice from infection with *A. phagocytophilum* HZ (Kim and Rikihisa, 1998), indicating that P44 proteins contain at least two in vivo neutralizable B-cell epitopes. In the present study, we defined the two neutralization sites on P44 molecules by epitope peptide mapping and used the MAbs to delineate their bacterial surface exposure and inhibitory mechanisms of infection of host cells. The present results support the idea that large expansion of *p44* paralogues allows *A. phagocytophilum* to escape neutralizing antibodies.

### 3.2 Materials and methods

**Bacteria.** *A. phagocytophilum* HZ strain isolated from a human granulocytic anaplasmosis patient in 1995 (Rikihisa et al., 1997) and *Ehrlichia chaffeensis* Arkansas were cultured in HL-60 cells (a human promyelocytic leukemia cell line) (American Type
Culture Collection, Manassas, VA) in RPMI 1640 medium (Gibco-BRL, Grand Island, NY) supplemented with 5% fetal bovine serum (FBS; Bio-technologies Inc., Parker Ford, PA), and 2 mM L-glutamine (Gibco-BRL). Cell cultures were incubated at 37°C in a humidified 5% CO₂-95% air atmosphere. The percentage of infected cells was determined by Diff-Quik staining (Baxter Scientific Products, Obetz, OH).

**Cloning and expression of rP44N and rP44N-N.** Primer pairs TCCGAATTCCGACTGGTGTTGCAGGGAT and CGCAAGCTTCTAATCAGTCTGCCCC, TCCGAATTCCGACTGGTGTTGCAGGGAT and CGCAAGCTTCTACTTAAACC-CAATCCGA were designed to clone the DNA fragments encoding recombinant P44N (rP44N; P44N terminal from position 35 to 162, 128 amino acids) and recombinant P44N-N (rP44N-N; N-terminal portion of rP44N, from positions 35 to 98, 64 amino acids) based on the full-length P44-18 sequence at the \text{p44} expression locus (Zhi \textit{et al.}, 1999). The \text{Eco}RI and \text{Hind}III sites are underlined. PCR products of the expected sizes were obtained and digested with \text{Eco}RI and \text{Hind}III. These fragments were then ligated into the \text{Eco}RI and \text{Hind}III sites of the pET33b (+) vector (Novagen Inc., Madison, WI). The resulting plasmid was amplified in \textit{Escherichia coli} Novablue cells (Novagen). \textit{E. coli} BL21 (DE3) cells (Novagen) were transformed with the recombinant plasmid and induced to express rP44N and rP44N-N with isopropyl-thio-β-D-galactoside. The recombinant proteins include 41 amino acids at the N terminus derived from the pET33b (+) vector. The rP44N and rP44N-N inclusion bodies were purified using the B-PER II bacterial protein extraction reagent (Pierce, Rockford, IL). The recombinant proteins were then Ni-affinity purified with a His-Select cartridge (Sigma, St. Louis, MO).

**Surface labeling of host cell-free \textit{A. phagocytophilum} with MAb 5C11 and**
Host cell-free organisms were isolated as previously described (Yoshiie et al., 2000) and cytocentrifuged on glass slides. Organisms were fixed with 2% paraformaldehyde in phosphate-buffered saline (PBS; 137 mM NaCl, 10 mM Na$_2$HPO$_4$, 2.7 mM KCl, 1.8 mM KH$_2$PO$_4$; pH 7.2) for 1 h at room temperature and then incubated with 5C11, 3E65, or normal mouse immunoglobulin G (IgG) only or with both 5C11 and a rabbit antibody directed against the recombinant *A. phagocytophilum* NtrX transcription factor (T.-H. Lai, Y. Kumagai, and Y. Rikihisa, unpublished data) for 1 h at 37°C. After being washed twice with PBS, cells were incubated with Alexa Fluor 555-conjugated (red fluorescence) goat anti-mouse IgG antibody (Molecular Probes, Eugene, OR) or with both Alexa Fluor 555-conjugated goat anti-rabbit IgG and Alexa Fluor 488-conjugated (green fluorescence) goat anti-mouse IgG antibody (Molecular Probes) for 1 h at 37°C. As the control, host cell-free organisms were fixed with methanol at –20°C for 5 min to permeabilize the organism membrane and then double labeled with both MAb 5C11 and anti-NtrX, with both MAb 3E65 and anti-NtrX, or with both MAb 5C11 and rabbit preimmune serum for 1 h at 37°C. After being washed twice with PBS, cells were incubated with both Alexa Fluor 555-conjugated goat anti-rabbit IgG and Alexa Fluor 488-conjugated goat anti-mouse IgG antibody for 1 h at 37°C. Bacteria were then washed and observed under an Eclipse E400 epifluorescence microscope with a xenon–mercury light source (Nikon, Melville, NY).

**Effects of P44 MAbs on binding, internalization, and infection in vitro.** For the binding analysis, the host cell-free *A. phagocytophilum* was freshly prepared from 2.5 x 10$^4$ *A. phagocytophilum*-infected HL-60 cells (>90% cells infected) as previously described (Yoshiie et al., 2000). The preparation was then incubated with MAbs 5C11
(IgG2b), 3E65 (IgG1) (Kim and Rikihisa, 1998), or normal mouse IgG (Sigma)-nonrelated mouse MAb with the isotype of IgG2b (hybridoma culture supernatant) (final concentration, 1 mg/ml), or with RPMI 1640 medium at room temperature for 30 min with gentle shaking. Each mixture was then added to 2 x 10^5 uninfected HL-60 cells (final concentration, 1.3 x 10^6 cells/ml) in RPMI 1640 medium supplemented with 5% FBS and 2 mM L-glutamine. After incubation at room temperature for 15 min with shaking, each mixture was incubated at 37°C in 95% air-5% CO₂ for 45 min. Cells were harvested by centrifugation at 750 x g for 5 min and washed with PBS to remove the unbound A. phagocytophilum; this washing procedure was then repeated twice. The mixture was cytocentrifuged and fixed with cold methanol for 5 min and sequentially incubated with horse anti-A. phagocytophilum plasma and Cy3-conjugated anti-horse IgG (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA). The number of bound organisms was scored in 100 HL-60 cells in triplicate samples.

For analysis of infection, the host cell-free A. phagocytophilum freshly prepared from each sample of 4 x 10^5 A. phagocytophilum-infected HL-60 cells (>90% cells infected) as previously described (Yoshiie et al., 2000), was incubated with various concentration of either MAb 5C11 or 3E65 (Kim and Rikihisa, 1998), nonrelated mouse MAb with the isotype of IgG2b, nonrelated mouse MAb with the isotype of IgG1, normal mouse IgG, or RPMI 1640 medium at room temperature for 30 min with gentle shaking. As a negative control, the same number of host cell-free Ehrlichia chaffeensis cells was incubated with MAb 5C11, MAb 3E65, normal mouse IgG or RPMI 1640 medium under the same conditions. Each of the mixtures was added to 2 x 10^5 uninfected HL-60 cells (final concentration, 10^6 cells/ml) in RPMI 1640 medium supplemented with 5% FBS and
2 mM L-glutamine and incubated at 37°C in 95% air-5% CO₂. After 12 h, the medium was replaced with fresh RPMI 1640 medium supplemented with 5% FBS and 2 mM L-glutamine. The host cell-free organisms were then coincubated with HL-60 cells at 37°C for several days to allow the growth of internalized *A. phagocytophilum* (Lin and Rikihisa, 2004). The infection of *A. phagocytophilum* and *Ehrlichia chaffeensis* was evaluated by Diff-Quik staining. The infectivity was scored in 100 HL-60 cells in triplicate samples.

To study the effects of MAbs on the internalization and development of *A. phagocytophilum* in HL-60 cells, the cell-free organisms were prepared as described above and then incubated with MAb 3E65 (final concentration, 1 mg/ml), or with RPMI 1640 medium at room temperature for 30 min with gentle shaking. Each mixture was then added to uninfected HL-60 cells as described above. After incubation at room temperature for 15 min with shaking, each mixture was incubated at 37°C for 1, 4, 16, and 32 h. At 12 h postinoculation, the medium was replaced with fresh RPMI 1640 medium supplemented with 5% FBS and 2 mM l-glutamine. Cells were harvested and washed as described above, followed by fixation with 2% paraformaldehyde at room temperature for 1 h. The extracellular bacteria were first stained with horse anti-*A. phagocytophilum* plasma and Cy3-conjugated goat anti-horse IgG in the absence of saponin, whereas the total extracellular and intracellular bacteria were stained with the same antibodies in the presence of 0.3% saponin.

In order to estimate lysosomal fusion, double immunofluorescence labeling was performed using phycoerythrin-MAb against human CD63 (IgG1; BioLegend, San Diego, CA) and horse anti-*A. phagocytophilum* plasma (detected with fluorescein
isothiocyanate-conjugated goat anti-horse IgG [Jackson ImmunoResearch]) in the presence of 0.3% saponin. All cells were washed three times with PBS to remove unbound antibodies before observation with a Nikon Eclipse E400 fluorescence microscope. Bacteria or inclusions per 100 HL-60 cells were scored in three independent experiments. Statistical analyses were performed by using analysis of variance and the Tukey honestly significant differences test or by Student's \( t \) test, and a \( P \) of <0.05 was considered significant.

**Peptide synthesis and peptide-pin ELISA analysis of MAbs 5C11 and 3E65 and infected horse, mouse, and human plasma.** Octamer peptide libraries were synthesized for the P44N-N region (64 amino acids) and the domain bordered by two cysteines in the P44-18 hypervariable (P44-18hv C-C) region (30 amino acids) with three and four overlapping amino acids, respectively, using noncleavable multipin synthesis technology and fluorenylmethoxycarbonyl (Fmoc) chemistry (Mimotopes Pty Ltd., Victoria, Australia) (Geysen, 1990). After disruption of the peptide-pins with 0.1 M sodium phosphate buffer containing 1% sodium dodecyl sulfate and 0.1% 2-mercaptoethanol (pH 7.2) and washing with hot water, nonspecific binding sites of pins were blocked with 200 µl of blocking solution containing 2% (wt/vol) bovine serum albumin in PBS-Tween 20 in 96-well plates for 1 h at room temperature.

Sets of peptide-bound pins were washed once with PBS containing 0.1% Tween 20 for 10 min and then incubated with either MAb 5C11 (1:600 dilution), MAb 3E65 (1:100 dilution), or one of the following in blocking solution at 4°C overnight: a 1:50 dilution of horse preimmune plasma from experimentally infected horses, EQ001, EQ005, and EQ006; plasma from seven horses from a region where human granulocytic
anaplasmosis is nonendemic (Columbus, Ohio); plasma from horse EQ001 (immunofluorescence titer, 1:320) on day 16 postinoculation of *A. phagocytophilum* (Kim *et al.*, 2002); plasma from horse EQ005 (immunofluorescence assay titer, 1:5,000) on day 31 after infected-tick placement; plasma from horse EQ006 (immunofluorescence titer, 1:640) on day 22 postinoculation (Wang *et al.*, 2004); one of three pooled mouse preimmune plasma samples (for each pool, there were >5 mice); and one of three plasma samples pooled from three infected ICR, C3H/HeN, and C3H/HeJ strain mice (immunofluorescence titer, 1:1,000) on day 31 postinoculation. Immunofluorescence titers of these plasma against *A. phagocytophilum* were determined as previously described (Kim and Rikihisa, 1998).

After being washed four times as described above, horseradish peroxidase-labeled goat anti-mouse or horse IgG (Kirkegaard & Perry Laboratories, Gaithersburg, MD) secondary antibodies were used. All secondary antibodies were diluted 1:500 or 1:200 in 1% (vol/vol) sheep serum (Sigma). The peptide pins were placed in the wells filled with the corresponding secondary antibodies and incubated for 1 h at room temperature. Samples were washed four times, and then the horseradish peroxidase substrate azino-di-3-ethyl-benzthiazolin-sulfonate (Sigma) in 70 mM citrate buffer (pH 4.2) was applied to a new plate and incubated with peptide pins for 10 min at room temperature. The optical densities at 415 nm (OD$_{415}$) and 492 nm were measured in an enzyme-linked immunosorbent assay (ELISA) plate reader (Molecular Device, Sunnyvale, CA). Each assay was repeated more than twice. The cutoff OD$_{415-492}$ value for positive reaction was the mean OD$_{415-492}$+3 standard deviations of negative control plasma.

### 3.3 Results
Expression of rP44N and rP44N-N, and Western blot analysis. To map the MAb 5C11 epitope, rP44N and rP44N-N were cloned and expressed in *E. coli*. This expression yielded a 169-amino-acid rP44N (18,550 Da) and a 105-amino-acid rP44N-N (11,635 Da), each of which included 41 amino acids derived from the pET33b (+) vector in the N terminus. The rP44N and rP44N-N were detected as single bands of approximately 19 and 12 kDa, respectively, on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Western blot analysis (Fig. 3.1) showed that both recombinant proteins were recognized by MAb 5C11, indicating that the 5C11 epitope was located within the 64 amino acids encompassing the P44N-N conserved region. MAb 5C11 did not react with *E. coli* that had been transformed with vector alone.

**MAb 5C11 and 3E65 epitopes were exposed to the surface of host cell-free bacteria.** In order to determine whether the MAb 5C11 and 3E65 epitopes are exposed to the bacterial surface, host cell-free-organisms were fixed with paraformaldehyde to prevent the penetration of the antibody into the organisms. Both MAbs 5C11 and 3E65 preferentially stained the surface of individual bacteria, resulting in a ring-like labeling pattern (Fig. 3.2 A and B). Normal mouse IgG did not label *A. phagocytophilum* (Fig. 3.2D). A rabbit antibody directed against the recombinant *A. phagocytophilum* NtrX transcription factor (a cytoplasmic protein) did not label prefixed bacteria (Fig. 3.2C).

When fixed cells were permeabilized by methanol fixation, bacteria were positively labeled with both MAb 5C11 and the anti-NtrX antibody or MAb 3E65 and anti-NtrX (Fig. 3.2E and F). NtrX (red fluorescence) was in the cytosol and both MAb epitopes (green fluorescence) were on the membrane (Fig. 3.2E and F). Rabbit preimmune serum did not label organisms (Fig. 3.2G). These results indicate that both the
MAb 5C11 and the MAb 3E65 epitopes are exposed to the bacterial surface. The majority of the *A. phagocytophilum* population expressed the P44-18 protein. This result was consistent with predominant expression of the *p44-18* transcript by the *A. phagocytophilum* HZ strain in cell culture (Wang et al., 2004; Zhi et al., 2002b).

**Inhibition of binding, development, and infection of *A. phagocytophilum* with MAbs in vitro.** To analyze whether MAb 5C11 and MAb 3E65 block infection of *A. phagocytophilum*, host cell-free organisms were preincubated with MAbs and then coincubated with HL-60 cells at 37°C for several days to allow the growth of internalized *A. phagocytophilum* (Lin and Rikihisa, 2004). As shown in Fig. 3.3A, both MAb 5C11 and 3E65 blocked infection, whereas IgG from normal mice, or unrelated mouse MAb isotype IgG2b and IgG1 had no effect on *A. phagocytophilum* infection. As a negative control, host cell-free *Ehrlichia chaffeensis*, which is the human monocytic ehrlichiosis agent and belongs to the family *Anaplasmataceae*, was not labeled by either MAb (data not shown), and the infection of HL-60 cells with *Ehrlichia chaffeensis* was not inhibited by either of the two MAbs (Fig. 3.3B).

To analyze whether MAb 5C11 and MAb 3E65 block binding of *A. phagocytophilum* to host cells, host cell-free organisms were preincubated with MAbs and then coincubated with HL-60 cells at room temperature for 15 min with gentle shaking and at 37°C for 45 min to allow binding and internalization of *A. phagocytophilum* into HL-60 cells. MAb 5C11 almost completely inhibited the binding and internalization of *A. phagocytophilum* to the host cells compared to normal mouse IgG or a nonrelated mouse MAb (isotype of IgG2b), but MAb 3E65 did not inhibit the *A. phagocytophilum* binding (Fig. 3.3C and D). The numbers of extracellular bound bacteria
and total host cell-associated bacteria at 1 and 4 h postinoculation were similar with or without MAb 3E65 (Fig. 3.3E). This result suggested that MAb 3E65 did not inhibit internalization. However, after internalization, transformation from individual bacteria to microcolonies (morulae) was inhibited in HL-60 cells at 16 and 32 h postinoculation (Fig. 3.3F). The lysosomal fusion was undetectable in any of these bacteria at these time points by double immunofluorescence labeling with anti-CD63 at this time postinoculation (not shown).

**Epitope mapping of MAbs 5C11 and 3E65.** The epitopes of MAbs 5C11 and 3E65 were determined using peptide-pin ELISA (Conlan et al., 1988). To access the P44 protein hypervariable region that contains the MAb 3E65 epitope, we analyzed the 3D structure of several P44 proteins (P44-18, P44-28, and P44-30) predicted by the Robetta full-length protein prediction program (Kortemme and Baker, 2004; Kuhlman and Baker, 2004). Robetta is a full-chain protein structure prediction server. It parses protein chains into putative domains with the Ginzu protocol, and models those domains either by homology modeling or by *ab initio* modeling.

A representative predicted structure of the P44 proteins is shown in Fig. 3.4. Three major domains were predicted in the P44 protein full-length query by the Robetta program. (i) Approximately amino acid positions 237 to 437, including the C-terminal half of the central hypervariable region and the C-terminal conserved region, were highly homologous to a β-barrel structure of *Neisseria meningitidis* outer membrane protein NspA (for P44-30 and P44-28, which have PDB BLAST confidence levels of 3.30 and 4.10, respectively, a confidence level of ≥3.0 predicts almost certainly the right protein folding) or *E. coli* outer membrane protein A (OmpA) (for P44-18, which has a PDB-
BLAST confidence level of 2.27, the confidence level between 2.0 and 3.0 usually predicts the right protein folding). (ii) From the N terminus, an approximately 115-amino-acid sequence, which includes the P44N-N region (approximately from amino acid positions 35 to 98) was characterized by a four-stranded $\beta$-sheet based on multiple sequence alignment (Sadreyev and Grishin, 2004) (for P44-18 the confidence level was 3.06). (iii) The region approximately from amino acid positions 105 to 240, which includes the region between two absolutely conserved cysteines (C-C region, approximately from amino acid positions 195 to 229) was highly hydrophilic and represents the most variable (nonidentical) domain in the P44s (for P44-18, P44-28, and P44-30, PDB-BLAST confidence levels were 5.14, 16.27, and 16.30, respectively) (Lin et al., 2002). This C-C region was primarily comprised of $\beta$ turns.

Based on this computer-predicted model of P44 protein structure, overlapping peptides corresponding to the P44-18hv C-C region (30 amino acids) were synthesized for MAb 3E65 epitope analysis. For MAb 5C11 epitope analysis, overlapping peptides corresponding to the P44N-N region (64 amino acids) were synthesized. The MAb 5C11 epitopes were determined to be KFDWTNPDB by peptide-pin ELISA analysis (Fig. 3.5). However, MAb 3E65 bound FAKYIVGA and NRFAKY in the peptide-pin ELISA analysis (Fig. 3.5). This suggests that the best fit epitope may lie within the core FAKY, with affinity perhaps defined by the flanking regions. We further performed a detailed analysis on NRFAKYIVGA. Comparative analysis using FAKY, FAKYIV, and NRFAKY showed FAKYIV was the strongest epitope peptide recognized. By BLAST search, FAKY was detected only in P44-18 but not in any of 86 other P44 proteins.
predicted to be encoded by *A. phagocytophilum* HZ (www.tigr.org).

**P44N-N and the P44-18hv C-C region epitope mapping of infected horse and mouse plasma.** To investigate whether plasma from infected immunocompetent mammals react to these P44 peptides, a peptide-pin ELISA analysis was performed using overlapping peptides derived from the P44N-N region and P44-18hv C-C region. Compared with horse preimmune or normal plasma, all plasma from pooled infected mice, or some plasma from infected horses recognized the MAb 5C11 epitope, and the MAb 3E65 epitope was recognized by all infected mouse and horse plasma (Fig. 3.6). This suggests that P44-18 was expressed by *A. phagocytophilum* infecting both host species and that the neutralizing epitope can be recognized in vivo by at least some infected hosts.
Fig. 3.1 MAb 5C11 recognizes the P44 N-terminal conserved region. (A) Affinity-purified rP44N (3 µg) and rP44N-N (3 µg) were subjected to SDS-PAGE followed by Coomassie blue staining. M, molecular size markers. rP44N is indicated by the arrow. (B) A duplicate gel was subjected to Western blot analysis with MAb 5C11.
Fig. 3.1
Fig. 3.2 MAb 5C11 and MAb 3E65 label the surface of *A. phagocytophilum*. Organisms were prefixed in paraformaldehyde (A to D) or in methanol (E to G) and subjected to immunofluorescence labeling with (A) MAb 5C11, (B) MAb 3E65, (C) MAb 5C11 and NtrX, (D) normal mouse IgG, (E) MAb 5C11 and NtrX, (F) MAb 3E65 and NtrX, or (G) MAb 5C11 and rabbit preimmune serum. Note ring-like surface labeling of individual organisms with two MAbs more clearly seen in the small Figures to the right of panels (A and B, red; C and E to G, green). Note red (anti-NtrX) labeling of the cytoplasm of methanol-permeabilized bacteria (E and F). Bar 5 µm.
Fig. 3.2
Fig. 3.3 MAb 5C11 and MAb 3E65 inhibit infection of *A. phagocytophilum* in HL-60 cells by inhibiting binding and intracellular development, respectively. (A and B) Inhibition of infection of *A. phagocytophilum* or *Ehrlichia chaffeensis* with MAb 5C11, MAb 3E65, normal mouse IgG, or isotype-matched mouse IgG control. The percent inhibition of infection is expressed as described below. (A) Inhibition of *A. phagocytophilum* infection. (B) Inhibition of *Ehrlichia chaffeensis* infection. (C and D) Inhibition of binding of the host cell-free *A. phagocytophilum* with MAb 5C11, MAb 3E65, normal mouse IgG, isotype-matched mouse IgG control, or RPMI 1640 medium. (C) Immunofluorescence micrographs of *A. phagocytophilum* bound to HL-60 cells. Bar: 5 µm. (D) Percent inhibition of binding to HL-60 cells. Numbers of bound or internalized *A. phagocytophilum* were scored in 100 HL-60 cells in triplicate samples. The percent inhibition of binding or internalization is expressed as the number of bacteria per HL-60 cell incubated with RPMI medium minus the number of bacteria per HL-60 cell under the indicated conditions divided by the number of bacteria per HL-60 cell incubated with RPMI medium multiplied by 100. (A and D) Significantly different from the remaining groups by the Tukey honestly significant differences test (*, *P* < 0.01; **, *P* < 0.05). (E and F) Inhibition of transformation from individual bacterium to a microcolony (morula) by MAb 3E65 in HL-60 cells. (E) Number of total and extracellular bacteria in HL-60 cells at 4 h postinoculation; (F) size of bacteria or inclusion at 16 h postinoculation. (E and F) Significantly different between the two groups by Student's *t* test (*, *P* < 0.002).
Fig. 3.3

A

% inhibition of the number of A. phagocytophilum

-80
-60
-40
-20
0
20
40
60
80
100

5C11 288 µg/ml
5C11 58 µg/ml
5C11 12 µg/ml
3E65 576 µg/ml
3E65 115 µg/ml
3E65 23 µg/ml
Normal mouse IgG 600 µg/ml
Normal mouse IgG2b 300 µg/ml
Normal mouse IgG1 600 µg/ml

B

% inhibition of the number of E. chaffeensis

0
20
40
60
80
100

5C11 288 µg/ml
3E65 576 µg/ml
Normal mouse IgG 600 µg/ml
% inhibition of the number of *A. phagocytophilum* binding to the host cells

- RPM medium
- 3E65
- 5C11

67
\textbf{E} 

4h p.i.

\begin{figure}[h]
\centering
\begin{tikzpicture}
\begin{axis}[
    title={Total Extracellular},
    ylabel={A. phagocytophilum or inclusion/100 HL-60 cells},
    xtick={0,1,2,3,4},
    xticklabels={Total, Extracellular},
    ybar, ymajorgrids=true,
    enlarge x limits=0.5,
]
\addplot+[fill=gray!20] coordinates {
(0,120)
(1,100)
(2,80)
(3,60)
};
\addplot+[fill=gray!40] coordinates {
(0,140)
(1,120)
(2,100)
(3,80)
};
\end{axis}
\end{tikzpicture}
\caption{Graph showing the total extracellular counts of A. phagocytophilum or inclusions per 100 HL-60 cells.}
\end{figure}

\textbf{F} 

16h p.i.

\begin{figure}[h]
\centering
\begin{tikzpicture}
\begin{axis}[
    title={Size of bacteria or inclusions (µm)},
    xlabel={Size of bacteria or inclusions (µm)},
    ylabel={A. phagocytophilum or inclusion/100 HL-60 cells},
    xtick={0,1,2,3,4},
    xticklabels={0.2-2,2-4,4-10,>=10},
    ybar, ymajorgrids=true,
    enlarge x limits=0.5,
]
\addplot+[fill=gray!20] coordinates {
(0,25)
(1,20)
(2,15)
(3,10)
};
\addplot+[fill=gray!40] coordinates {
(0,30)
(1,25)
(2,20)
(3,15)
};
\addplot+[fill=gray!60] coordinates {
(0,35)
(1,30)
(2,25)
(3,20)
};
\addplot+[fill=gray!80] coordinates {
(0,5)
(1,5)
(2,5)
(3,5)
};
\end{axis}
\end{tikzpicture}
\caption{Graph showing the size distribution of bacteria or inclusions in A. phagocytophilum or inclusions per 100 HL-60 cells.}
\end{figure}

\begin{itemize}
\item [\textbullet] RPMI medium
\item [\textbullet] 3E65
\end{itemize}
Fig. 3.4 Two views of the representative P44 protein three-dimensional structure predicted by the Robetta program and the MAb 5C11 epitope. A. Open hinge structure; B. closed hinge structure. The P44N-N and P44hv C-C region are shown in blue (the MAb 5C11 epitope is green) and red, respectively. From the N terminus, approximately 242 amino acids which include the whole N-terminal conserved region and approximately half of the central hypervariable region form an $\alpha$-helix- and $\beta$-turn-rich domain which is expected to be surface exposed on the outer membrane of *A. phagocytophilum*. Another half of the central hypervariable region and most of the C-terminal conserved region (amino acid positions 243 to 440) are predicted to make a $\beta$-barrel structure embedded in the outer membrane of *A. phagocytophilum*. The P44N-N region (amino acid positions 35 to 98) is characterized by a four-stranded $\beta$ sheet. The central hypervariable region (amino acid positions 195 to 229) between two absolutely conserved cysteines mainly contains $\beta$ turns. The outer surface protruded domain and the membrane-embedded $\beta$-barrel are connected by the flexible hinge (amino acid positions 236 to 242, yellow).
Fig. 3.4
Fig. 3.5 Peptide mapping of MAb 5C11 and MAb 3E65 epitopes. Peptide-bound pins were incubated with either MAb 5C11 (upper panel) or 3E65 (lower panel). The amino acid sequences of overlapping synthetic peptides within P44N-N (A) and P44-18hvC-C (B) are arranged from the N to the C terminus and indicated on the x axis. Relative absorbance as determined by ELISA analysis is shown on the y axis. A representative result of triplicate assays is shown.
Fig. 3.5

Relative absorbance

A

5C11

B

3E65

TGGAGYFY
YFYYGLDY
LDYSPAFS
AFSKIRDFRDFSIRES
RESNGETK
ETKAYYPY
YPYLKDGK
DGKSYKLE
KLESHKFD
KFDW
NTPD
TPDPRIGF
IGFK
ETKRKDGD
KDGDTTNR
TTNRFAKY
FAKYIVG
IVG
AGDSS
GDSSNAGT
NAGTSL
Fig. 3.6 ELISA analysis immune plasma or sera for linear B-cell epitopes within P44N-N and P44-18hvC-C of *A. phagocytophilum*. Preimmune and immune plasma from horse EQ001, EQ005, and EQ006; plasma from seven control horses; pools of plasma from three infected ICR, three infected C3H/HeN, and three infected C3H/HeJ strain mice; and three pooled uninfected plasmas from >5 mice each were allowed to react with the various synthesized peptides. The amino acid sequences of overlapping synthetic peptides are arranged from the N to the C terminus and indicated on the x axis. The y axis shows the difference in the absorbance (OD$_{415-492}$) of immune plasma (individual reactions are shown) and that of control plasma (means and standard deviations are shown). The 5C11 and 3E65 epitopes are indicated in square boxes. A reaction was considered to be positive when the immune plasma yielded an OD$_{415-492}$ value larger than the mean OD$_{415-492}$ + 3 standard deviations of negative control plasma or sera (shown as dash with solid triangle). A representative result of three to five assays is shown.
3.5 Discussion

The results from this study provide a significant advancement in our understanding of the structure of the P44 protein and neutralizing B-cell epitopes. Park et al. (Park et al., 2003) reported dimerization and oligomerization of P44/MSP2 mediated by a disulfide bond. Robetta, using amino acid sequence alone, as with other structure prediction methods, cannot predict influences of other interacting proteins. However, at least three different P44 protein amino acid sequences in our analysis provided the consistent P44 structural model, and the present data were in agreement with the Robetta model indicating that two neutralizing epitopes present within the P44N-N and P44hv C-C regions of naturally folded P44 molecules embedded in the bacterial membrane (thus perhaps in an oligomerized condition) are exposed on the intact bacterial surface and accessible by the MAbs.

MAb 5C11 and 3E65 were found to act as infection neutralizing antibodies in vitro. The result is partially in agreement with a previous work that showed a weak neutralization with MAbs (Park et al., 2003). In this study, the neutralization mechanisms were found distinct between two MAbs: MAb 5C11 neutralized at the level of binding to HL-60 cells, whereas MAb 3E65 neutralized after internalization of *A. phagocytophilum* into HL-60 cells. Thus, the P44hv region may be associated with the critical signals for development of internalized bacteria into a replicating stage to form a morula. Since MAb 3E65 did not block binding, this region does not appear to be involved and may not even be sterically close to the ligand for binding and internalization of *A. phagocytophilum* into host cells. On the other hand, the epitope of MAb 5C11 appears to be either critical for binding or in close proximity to the ligand of *A. phagocytophilum*. 
Results from the Robetta full-length protein prediction program suggested that the surface-exposed P44 hypervariable C-C region and the P44N-N region are connected to the membrane-embedded β-barrel by a short chain of amino acids (positions 236 to 242) that may serve as flexible hinge (Fig. 3.4A and B). The predicted flexibility of these two exposed regions of the P44 molecule is expected to facilitate intermolecular interaction.

Antibodies directed against the hypervariable region of major surface antigens of bacteria are generally considered to have a strong neutralizing activity; therefore, such antibodies clear this specific antigenic population and allow bacteria with different hypervariable region sequences to become the next dominant population (Barbour, 1990). Our previous study showed that the initial P44-18 phenotype is cleared and that new P44 phenotypes sequentially emerged when horses are infected with *A. phagocytophilum* HZ (Wang *et al.*, 2004). The present study suggests that the C-C region of P44hv serves as a target for antigenic variation and clearance.

By alignment of multiple P44 proteins, the C-C region is only approximately one-third of the P44 hypervariable region (Lin *et al.*, 2002). According to the Robetta model, the remaining P44hv region is a part of the β-barrel structure and thus predicted not to be exposed to the bacterial surface (Fig. 3.4). The role of this potentially nonexposed hypervariable region in *A. phagocytophilum* infection remains to be studied. Our antigenic index analysis, using the Protean program in DNASstar, predicted that the non-surface-exposed region of the P44 hypervariable region is also highly antigenic (Lin *et al.*, 2002); thus, this region as well as a part of the P44 C-terminal region with predicted high antigenic index may serve as decoy antigens in *A. phagocytophilum* infection. Similarly, the immunodominant C-terminal domain and one invariable region in the
central variable domain (IR6) of *Borrelia burgdorferi* VlsE are not exposed at the surface of the intact spirochete and thus proposed to serve as decoy epitopes to divert immune responses away from the variable regions (Liang *et al*., 1999).

Our study showed that the MAb 5C11 and 3E65 epitopes can be recognized by infected outbred horses and three genetically different mouse strains. It remains to be determined whether these epitopes can be universally recognized by animal and human populations and/or serve as protective immunogens. In addition, several other peptides in the P44N-N region were recognized, and whether these peptides can serve as neutralizing epitopes remains to be analyzed. Furthermore, for more than 80 different P44 hypervariable C-C region amino acid sequences, neutralizing epitopes remain to be verified.

In a previous study, we showed that the *p44-18* sequence and genomic locus are conserved among 14 *A. phagocytophilum* strains from a horse, ticks, and human granulocytic anaplasmosis patients in northeastern states, Wisconsin, and California (Lin *et al*., 2004). *p44-18* is the major transcript species of *A. phagocytophilum* HZ that is initially detected in experimentally infected horses and mice, regardless of whether transmission occurs by syringe or by tick attachment (Wang *et al*., 2004; Zhi *et al*., 2002b). Current data further extended these observations, showing specifically recognized peptides within the P44-18hvC-C region by immune plasma.

The structure of the major surface protein 2 (MSP2) of the bovine erythrocytic parasite *A. marginale* is similar to P44, having highly conserved N- and C-terminal regions flanking a central hypervariable region (French *et al*., 1998). It was reported that B-cell epitopes are present predominantly in the central hypervariable region, and only a
few B-cell epitopes are present in the N-terminal conserved region of MSP2 of the *A. marginale* Florida strain (Abbott *et al.*, 2004). The native MSP2 protein isolated from *A. marginale* was reported to block hemagglutination of bovine erythrocytes with *A. marginale* in vitro (Magnarelli *et al.*, 2001). However, to our knowledge, there have been no reports on the neutralizing B-cell epitopes of MSP2.

In *A. marginale*, MSP1a and MSP1b are considered ligands for infection of ticks and bovine erythrocytes (Garcia-Garcia *et al.*, 2004; de la Fuente *et al.*, 2001; McGarey and Allred, 1994; McGarey *et al.*, 1994). Passive immunization of cattle with a monoclonal antibody directed against MSP1a neutralizes infection of bovine erythrocytes with *A. marginale* (Palmer *et al.*, 1986). The serum from rabbits immunized with recombinant MSP1a and MSP1b, either individually or in combination, reduces infection of tick cells with *A. marginale* (Blouin *et al.*, 2003). However, orthologues of *msp1a* or *msp1b* have not been detected in the *A. phagocytophilum* HZ genome (www.tigr.org). It is possible that in the absence of *msp1a* and *msp1b*, *p44 (msp2)* may have evolved to have a more significant role in the *A. phagocytophilum* infection of neutrophils.

In summary, our data imply that two separate neutralizing epitopes of P44 proteins are involved in the binding, internalization, and infection of *A. phagocytophilum* in HL-60 cells, and this may represent the basis for in vivo neutralization with P44N-N and P44hv C-C-specific antibodies. A better understanding of P44 neutralizing epitopes and domains may contribute to the development of new vaccines that can elicit a protective response against *A. phagocytophilum*. 
3.6 Summary

*Anaplasma phagocytophilum* is an obligatory intracellular bacterium that causes human granulocytic anaplasmosis. The polymorphic 44-kDa major outer membrane proteins of *Anaplasma phagocytophilum* are dominant antigens recognized by patients and infected animals. However, the ability of anti-P44 antibody to neutralize the infection has been unclear due to a mixture of P44 proteins with diverse hypervariable region amino acid sequences expressed by a given bacterial population and lack of epitope-defined antibodies. Monoclonal antibodies (MAbs) 5C11 and 3E65 are directed to different domains of P44 proteins, the N-terminal conserved region and P44-18 central hypervariable region, respectively. Passive immunization with either MAb 5C11 or 3E65 partially protects mice from infection with *Anaplasma phagocytophilum*. In the present study, we demonstrated that the two monoclonal antibodies recognize bacterial surface-exposed epitopes of naturally folded P44 proteins and mapped these epitopes to specific peptide sequences. The two MAbs almost completely blocked the infection of the *Anaplasma phagocytophilum* population that predominantly expressed P44-18 in HL-60 cells by distinct mechanisms: MAb 5C11 blocked the binding, but MAb 3E65 did not block binding or internalization. Instead, MAb 3E65 inhibited internalized *Anaplasma phagocytophilum* to develop into microcolonies called morulae. Some plasma from experimentally infected horses and mice reacted with these two epitopes. Taken together, these data indicate the presence of at least two distinct bacterial surface-exposed neutralization epitopes in P44 proteins. The results indicate that antibodies directed to certain epitopes of P44 proteins have a critical role in inhibiting *Anaplasma phagocytophilum* infection of host cells.
CHAPTER 4

DELAYED DEVELOPMENT OF ANTIBODY TO P44N AND UPREGULATION OF RECA IN DIVERSE P44 EXPRESSION AND DELAYED CLEARANCE OF ANAPLASMA PHAGOCYTOPHILUM IN C3H/HEJ MICE

Introduction

Anaplasma phagocytophilum is a tick-borne obligatory intracellular pathogen that causes an acute systemic and potentially fatal disease, human granulocytic anaplasmosis (HGA) (Dumler et al., 2001; Michalek et al., 1980). Although A. phagocytophilum infection has been detected in a variety of mammals, Peromyscus leucopus (the white-footed mouse) is considered a primary reservoir of A. phagocytophilum infection in the Northeastern United States (Telford et al., 1996). The A. phagocytophilum genome encodes more than 80 immunodominant major surface protein P44s (Dunning Hotopp et al., 2006; Lin and Rikihisa, 2005). P44 proteins consist of a single central hypervariable (hv) region of approximately 94 amino acid residues, an N-terminal conserved region of approximately 186 amino acids, and a C-terminal conserved region of approximately 146 amino acids; the N- and C-terminal regions flank the central hv region (Lin et al., 2002; Zhi et al., 1999). There are three short conserved segments including two absolutely conserved cysteines within the hv region of all predicted P44 proteins (Lin et al., 2002).
A part of P44 proteins, including conserved N-terminus regions and the hv region flanked by C-C, are exposed to the bacterial surface (Wang et al., 2006). Surface-exposed segments of P44 proteins render the bacteria susceptible to epitope-defined monoclonal antibody neutralization in vitro (Wang et al., 2006; Wang et al., 2004). In experimentally infected horses, sequential changeovers of expressed p44 transcripts occur in parallel with the development of anti-P44 antibodies (Wang et al., 2004). While P44s may stimulate a humoral immune response (Zhi et al., 2002b), little is known about how P44 N-terminus- and P44hv-specific antibodies develop and in turn influence A. phagocytophilum clearance and the P44E allele population during mammalian infection. Thus our primary objective of the present study was to analyze the temporal development of antibodies against conserved P44N and P44hv regions and p44 transcript populations during infection of different strains of mice.

Gene conversion at a single p44-expression locus leads to P44 antigenic variation. Homologs of genes for the RecA-dependent RecF pathway, but not RecBCD or RecE pathways of recombination were detected in the A. phagocytophilum genome (Lin and Rikihisa, 2005; Lin et al., 2003). The recombination intermediate structure between a donor p44 and the p44-expression locus of A. phagocytophilum was detected in HL-60 cell culture by Southern blot analysis followed by sequencing the band and in blood samples from infected SCID mice by PCR followed by sequencing (Lin et al., 2006). Using a double-origin plasmid carrying the p44-expression locus and a donor p44 locus, the recombination intermediate was recovered in an E. coli strain with active RecF recombination pathway, but not in strains with deficient RecF pathway (Lin et al., 2006). p44 recombination at the p44-expression was demonstrated in SCID mice and in cell
culture, indicating that an amnestic immune response is not essential for \textit{p44} recombination (Lin and Rikihisa, 2005). Therefore, our second objective of the present study was to investigate whether \textit{p44} recombination is related to changes in the \textit{p44} transcript population and \textit{A. phagocytophilum} persistence during mammalian infection.

TLR4 recognizes LPS which is an integral component of the outer membranes of Gram-negative bacteria and turns on the down stream MAPK kinase signal pathway (Akira \textit{et al.}, 2001). Although \textit{A. phagocytophilum} persists less than 3 weeks in various strains of immunocompetent laboratory mice (von Loewenich \textit{et al.}, 2004; Bunnell \textit{et al.}, 1999), it infects C3H/HeJ with \textit{Lps}^d allele and deficient in Toll-like receptor (TLR)4 response for a longer period with some strains (or perhaps passage history) of \textit{A. phagocytophilum} (Hodzic \textit{et al.}, 1998b; Telford and Dawson, 1996). \textit{A. phagocytophilum} persists in severe combined immunodeficiency (SCID) mice (Lin and Rikihisa, 2005; Hodzic \textit{et al.}, 1998b), indicating importance of a competent immune response in clearance of \textit{A. phagocytophilum}. An immunocompetent mouse model in which the bacterium can persist for a longer period of time would facilitate the investigation of involvement of \textit{p44} recombination and development of antibodies against conserved and hypervariable regions of P44 proteins during emergence of diverse P44 variants and clearance of \textit{A. phagocytophilum} infection. Therefore, in the present study we performed quantitative analysis of temporal \textit{A. phagocytophilum} HZ strain infection in C3H/HeJ mice and SCID mice relative to wild-type mouse strains.

Our results revealed delayed development of anti-P44N antibody relative to anti-P44hv and upregulation of \textit{A. phagocytophilum recA} in C3H/HeJ mice, suggesting the delayed development of antibodies to the P44 conserved region in the host and enhanced
bacterial homologous recombination may facilitate emergence of diverse P44 variants and delayed clearance of *A. phagocytophilum* in TLR4-deficient mice.

4.2 Material and method

*A. phagocytophilum and mouse infection.* The *A. phagocytophilum* HZ strain was cultured in HL-60 cells as previously described (Rikihisa *et al.*, 1997). Three-week-old C3H/SCID male mice (The Jackson Laboratory, Bar Harbor, ME), three-week-old C3H/HeN (Harlan Sprague Dawley Inc., Indianapolis, IN) and C3H/HeJ male mice (The Jackson Laboratory) and four-week-old ICR and ICR SCID male mice (Taconic Farm Inc. Germantown, NY) were each inoculated intraperitoneally each with $10^6$ HL-60 cells infected with *A. phagocytophilum* (50–80% infected cells). The plasma and spleen specimens were collected from each of three mice on designated days post inoculation (p.i.).

**Competitive (C)-PCR, RT-PCR, cDNA cloning and sequence analysis.** To determine amounts of *A. phagocytophilum* organisms in the mouse spleen tissue, DNA was extracted and *p44* cPCR were performed for mouse spleen specimens as previously described (Wang *et al.*, 2004). To normalize the amount of mouse spleen DNA input across the samples, mouse G3PDH gene cPCR was performed by primer pair set 1 (Table 4.1). The mouse G3PDH competitor was constructed as described previously for construction of the *p44* competitors (Wang *et al.*, 2004) using primer sets 2 and 3 (Table 4.1). RNA was extracted from each spleen specimen and *A. phagocytophilum*-infected HL-60 cells and cDNA was synthesized as described previously (Wang *et al.*, 2004). PCR was performed with primer pair p3709 and p4257 (Wang *et al.*, 2004) to amplify cDNA encoding *p44* paralogs. Sequencing of *p44* cDNA PCR products and subsequent
analyses were performed as previously described (Wang et al., 2004). Colony hybridization was performed to analyze the percentage of the p44-18 transcripts over the total p44 transcripts using respective DNA probes as previously described (Wang et al., 2004).

**Western blot analysis.** Recombinant proteins rP44N (Wang et al., 2006) and rP44-18hv (Wang et al., 2004) were subjected to sodium dodecyl sulfate-15% polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a nitrocellulose membrane. The membrane was incubated at a dilution of 1:50 with pooled plasma samples from each of three mice collected on designated days p.i. The peroxidase-conjugated affinity-purified anti-mouse secondary antibodies (Kirkegaard & Perry Laboratories Inc., Gaithersburg, MD) were used at a dilution of 1:500. Finally, colorimetric reactions were developed with a 4-chloro-naphthol substrate solution. Images were captured and densitometric analysis was performed using LAS3000 image documentation system and the densities of target bands were analyzed by the Multi Gauge software (FUJIFILM Medical Systems USA, Stamford, CT).

**Quantitative RT-PCR.** Real-time RT-PCR was used to quantify mRNA levels of recA, recF, recJ, recG, ruvA, ruvB and ruvC. A. phagocytophilum 16S rRNA level was used to normalize input A. phagocytophilum RNA across specimens. Thus, relative target gene units were expressed in dimensionless units, and calculated according to the following formula: relative target gene transcript amount = $2^{-(\text{Avg. target gene transcript Ct-Avg. 16S rRNA Ct})} \times 10^4$ (Livak and Schmittgen, 2001). Primers were designed to produce an amplicon of 100 to 150 bp (Table 4.1). Quantitative PCR was performed in a real-time instrument, the MX3000P (Stratagene, Austin, TX), using a Brilliant SYBR Green QPCR
Core Reagent Kit (Cedar Creek, TX). For each real-time RT-PCR, the dissociation curve was examined to confirm the absence of primer dimers. The RT-PCR was carried out in a 25-μl reaction tubes containing 150 nM of each primer, 1.5 mM of MgCl₂, 200 nM each of dNTP, 8% glycerol, 3% DMSO, 30 nM reference dye, 0.5X SYBR Green I dye, 2.5 U of SureStart Taq DNA polymerase (Stratagene), and 4 μl of cDNA template. The cycling condition was 94°C heating for 10 min, and then 70 cycles of denaturing at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 30 s. Dissociation curve detection began with a 30 s incubation at 55°C followed by 40 successive 30 s plateaus in which the temperature was increased by 1.0°C for each plateau. For each cDNA sample, an RT– sample was used to monitor DNA contamination. Each specimen was independently repeated in different batch or individual mouse specimens to assess biological variation. Statistical analyses were performed using the unpaired Student t-test. A P value of <0.05 was considered significant.

4.3 Results

Temporal levels of *A. phagocytophilum* infection in ICR, ICR/SCID, C3H/HeN, and C3H/HeJ mouse spleen tissues. Temporal levels of *A. phagocytophilum* genome equivalents in ICR, ICR/SCID, C3H/HeN, and C3H/HeJ mouse spleen tissues were determined by competitive PCR and expressed as means and standard deviations of three ICR, ICR/SCID, C3H/HeN or C3H/HeJ mice at each day p. i. Both ICR and C3H/HeN cleared infection by 3 weeks to a PCR-undetectable level (Fig. 4.2). Rickettsemia persisted in both ICR/SCID and C3H/HeJ mice for at least two months. The levels of infection in ICR/SCID remained constant, while infection levels progressively
declined in C3H/HeJ mice. This result suggests a requirement of an acquired immune response and TLR4 for *A. phagocytophilum* HZ clearance in mice.

**Development of P44N and P44-18hv-specific antibodies during the course of mouse infection.** To determine the kinetics of mouse antibody development against the P44N-terminus conserved and P44hv regions, Western blot analysis was performed using rP44N and rP44-18hv. The *A. phagocytophilum* HZ inoculum predominantly expressed P44-18 (Zhi et al., 2002b), thus anti-P44-18hv is used as the representative initial antibody against P44hv region. Western blot analysis revealed that the kinetics of development of antibody against rP44N peaked at day 12 p.i. in C3H/HeN and ICR mice (Fig. 4.3). However, the development of antibody against P44N was slow and did not peak until day 31 p.i. in C3H/HeJ mice. The kinetics of antibody development against the hypervariable region of P44-18 (P44-18hv) was similar among C3H/HeN, C3H/HeJ and ICR mice, albeit the ICR strain had stronger reactions than the C3H strain of mice. This result indicates that TLR4 influences the kinetics of P44 antibody development in mice.

**Characterization of p44 transcript species following time course of infection.** Delayed clearance and slower development of anti-P44N-terminus antibody relative to the development of P44hv region antibody, may facilitate appearance of more diverse P44 variants in C3H/HeJ mice. Twenty five to 33 *p44* cDNA clones from each of the three C3H/SCID, C3H/HeN and C3H/HeJ mouse spleens and infected HL-60 cell specimens at day 12 p.i. were randomly picked up for sequencing the inserts. *p44-18* was the major transcript population in infected HL-60 cells, and *p44-18* became the second major *p44* transcript species detected in the spleens of C3H/SCID mice at 12 day p.i. (Table 4.4). In the spleens of both C3H/HeN and C3H/HeJ mice, no *p44-18* transcript
was detected and \textit{p44-2} became the major transcript type. Although C3H/SCID mice are deficient in adaptive immunity, the \textit{p44-18} transcript, which was the major transcript in cell culture (Wang \textit{et al.}, 2004), was significantly decreased compared with that in cell culture (Fig. 4.5) in agreement with previous study using ICR/SCID (Lin and Rikihisa, 2005). These results indicate appearance of more divergent P44 variants in C3H/HeJ mice at an earlier stage of infection than in C3H/SCID or cell culture.

\textbf{RecA expression level of \textit{A. phagocytophilum} in C3H/SCID, C3H/HeN and \textit{A. phagocytophilum} infected HL-60 cells.} \textit{A. phagocytophilum} genome sequence data revealed that \textit{A. phagocytophilum} has \textit{recA}, \textit{recF}, \textit{recO}, \textit{recR}, \textit{recG}, \textit{ruvA}, \textit{ruvB ruvC} and \textit{recJ} homologue \textit{E. coli} genes that could be used for the RecA-dependent RecF homologous recombination pathway (Dunning Hotopp \textit{et al.}, 2006; Lin \textit{et al.}, 2003). Real time PCR was performed to analyze the expression levels of \textit{A. phagocytophilum recA} and six other genes involved in RecF homologous recombination in C3H/HeN and C3H/SCID mice, and \textit{A. phagocytophilum} in HL-60 cells. This result showed that \textit{A. phagocytophilum recA} expression level was significantly higher in the spleens of C3H/HeN and C3H/SCID mice than in HL-60 cells (FIG. 4.6). The significant difference of the expression level of other recombination genes, \textit{recF}, \textit{recJ}, \textit{recG}, \textit{ruvA}, \textit{ruvB} and \textit{ruvC}, could not be detected by this method due to their low expression level (relative transcript amount < 0.4).
<table>
<thead>
<tr>
<th>Primer set no.</th>
<th>Primer name</th>
<th>Target</th>
<th>Purpose</th>
<th>size (bp)</th>
<th>5' primer</th>
<th>3' primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MG3PDH5'-MG3PDH3'</td>
<td>mouse G3PDH</td>
<td>Mouse G3PDH gene QC-PCR competitor constructio</td>
<td>983</td>
<td>TGAAGGTCGGTGTGA</td>
<td>GACGGATTTG GC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CATGTAGGGCCAT</td>
<td>GAGGTCACCAC</td>
</tr>
<tr>
<td>2</td>
<td>MG3PDH5'-MG3PDHComp3'</td>
<td>mouse G3PDH</td>
<td>competitor constructio</td>
<td>387</td>
<td>TGAAGGTCGGTGTGA</td>
<td>GACGGATTTG GC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CATGAATTCTGCAAG</td>
<td>GCTTGGGGCAA b</td>
</tr>
<tr>
<td>3</td>
<td>MG3PDHComp5'-MG3PDH3'</td>
<td>mouse G3PDH</td>
<td>competitor constructio</td>
<td>364</td>
<td>TATGAATTCTGCAAG</td>
<td>GCTTGGGGCAA b</td>
</tr>
<tr>
<td>4</td>
<td>recA5'-recA3'</td>
<td>recA A</td>
<td>real time PCR</td>
<td>133</td>
<td>TGC CGAACACATGCTTTG A</td>
<td>GATACG</td>
</tr>
<tr>
<td>5</td>
<td>recF5'-recF3'</td>
<td>recF</td>
<td>real time PCR</td>
<td>141</td>
<td>GTGTACATATCGCCA AAATC A</td>
<td>CATGAATTCTGCAAG</td>
</tr>
<tr>
<td>6</td>
<td>recJ5'-recJ3'</td>
<td>recJ</td>
<td>real time PCR</td>
<td>116</td>
<td>TCCGTGATAGAAAGCC AGAT A</td>
<td>CACCGCACCCCGA</td>
</tr>
<tr>
<td>7</td>
<td>recG5'-recG3'</td>
<td>recG A</td>
<td>real time PCR</td>
<td>142</td>
<td>AAAAGTTGCGCTAAA GTCTATTCA</td>
<td>CTGTTGGTTATCTA TCAAG</td>
</tr>
<tr>
<td>8</td>
<td>ruvA5'-ruvA3'</td>
<td>ruvA A</td>
<td>real time PCR</td>
<td>119</td>
<td>GTCAGCGGAATAAA CTACAGAAC</td>
<td>GCTGTTGGTTATCTA TCAAG</td>
</tr>
<tr>
<td>9</td>
<td>ruvB5'-ruvB3'</td>
<td>ruvB A</td>
<td>real time PCR</td>
<td>141</td>
<td>TTGAATCCCGGCTTT TACATTTAG</td>
<td>CACCGCACCCCGA</td>
</tr>
<tr>
<td>10</td>
<td>ruvC5'-ruvC3'</td>
<td>ruvC A</td>
<td>real time PCR</td>
<td>120</td>
<td>TGGGCTGACTTA CGAAGAC GAGAAC</td>
<td>AAAGATCAA GCTGATACAC</td>
</tr>
<tr>
<td>11</td>
<td>16S F-16S R</td>
<td>16S rRNA</td>
<td>real time PCR</td>
<td>108</td>
<td>GGTGAGTAATGCATA GGGTAC</td>
<td>GCGATACAA GCTGATACAA A</td>
</tr>
</tbody>
</table>

a *Bam*HI site is underlined.
b *Eco*RI site is underlined.

Table 4.1. Sequences of oligonucleotides used in PCR and RT-PCR
Fig. 4.2 Levels of *A. phagocytophilum* organisms in spleens of C3H/HeN and C3H/HeJ mice (A); ICR and ICR/SCID (B) mice as determined by *p44* cPCR. The amount of each mouse spleen tissue was normalized based on mouse G3PDH DNA levels by G3PDH gene cPCR. N=3 mice per each group at each time p.i.
Fig. 4.2

A

\[ \text{Log} \left( \frac{\text{organism number}}{(\text{G3PDH DNA amount} \times 10^8)} \right) \]

\[ \text{Days post inoculation} \]

B

\[ \text{Log} \left( \frac{\text{organism number}}{(\text{G3PDH DNA amount} \times 10^8)} \right) \]

\[ \text{Days post inoculation} \]
Fig. 4.3 Development of anti-P44N and P44-18hv antibodies in the plasma of *A. phagocytophilum*-infected mice. Affinity-purified rP44N and rhvP44-18 (3µg/lane) were subjected to SDS-PAGE and blotted to the membrane. The membrane was incubated with pooled plasma of three mice at the indicated day p.i. Ratios of P44N to hvP44-18 band densities for each plasma sample are indicated below each panel. Numbers on the right of each panel indicate molecular masses (in kilodaltons) based on the broad-range prestained standards (Bio-Rad).
Fig. 4.3

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>5</th>
<th>12</th>
<th>17</th>
<th>22</th>
<th>31</th>
<th>Day p.i.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
<tr>
<td>19 kD</td>
<td>rP44N</td>
<td>0.14</td>
<td>0.91</td>
<td>0.86</td>
<td>0.68</td>
<td>0.19</td>
<td>rP44N/ rhwP44-18</td>
</tr>
<tr>
<td>13 kD</td>
<td>rhwP44-18</td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
</tr>
<tr>
<td>0.14</td>
<td>0.50</td>
<td>0.57</td>
<td>1.20</td>
<td>1.30</td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
<tr>
<td>C3H/HeN</td>
<td>rP44N/ rhwP44-18</td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
<td><img src="image21.png" alt="Image" /></td>
<td><img src="image22.png" alt="Image" /></td>
</tr>
<tr>
<td>0.32</td>
<td>1.47</td>
<td>1.01</td>
<td>1.05</td>
<td>1.06</td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
<td><img src="image25.png" alt="Image" /></td>
</tr>
<tr>
<td>ICR</td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
<td><img src="image28.png" alt="Image" /></td>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
<td><img src="image31.png" alt="Image" /></td>
<td><img src="image32.png" alt="Image" /></td>
</tr>
<tr>
<td>p44 transcript type</td>
<td>% of cDNA clones</td>
<td>p44 transcript type</td>
<td>% of cDNA clones</td>
<td>p44 transcript type</td>
<td>% of cDNA clones</td>
<td>p44 transcript type</td>
<td>% of cDNA clones</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>C3H/HeN</td>
<td>C3H/HeJ</td>
<td>C3H/SCID</td>
<td>cell culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p44-2</td>
<td>36.7</td>
<td>p44-2</td>
<td>48.5</td>
<td>p44-35</td>
<td>27.6</td>
<td>p44-18</td>
<td>76.0</td>
</tr>
<tr>
<td>p44-48</td>
<td>20.0</td>
<td>p44-7</td>
<td>12.1</td>
<td>p44-18</td>
<td>20.7</td>
<td>p44-16</td>
<td>4.0</td>
</tr>
<tr>
<td>p44-36</td>
<td>16.7</td>
<td>p44-16</td>
<td>9.1</td>
<td>p44-12</td>
<td>13.8</td>
<td>p44-40</td>
<td>4.0</td>
</tr>
<tr>
<td>p44-15</td>
<td>10.0</td>
<td>p44-15</td>
<td>9.1</td>
<td>p44-24</td>
<td>6.9</td>
<td>p44-15</td>
<td>4.0</td>
</tr>
<tr>
<td>p44-14</td>
<td>6.7</td>
<td>p44-30</td>
<td>6.1</td>
<td>p44-15</td>
<td>6.9</td>
<td>p44-13</td>
<td>4.0</td>
</tr>
<tr>
<td>p44-12</td>
<td>6.7</td>
<td>p44-26</td>
<td>3.0</td>
<td>p44-28</td>
<td>6.9</td>
<td>p44-48</td>
<td>4.0</td>
</tr>
<tr>
<td>p44-1</td>
<td>3.3</td>
<td>p44-32</td>
<td>3.0</td>
<td>p44-36</td>
<td>6.9</td>
<td>p44-47</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4 p44 transcript species identified in C3H/HeN, C3H/HeJ, C3H/SCID mouse spleen and cell culture samples on day 12 p.i.
Fig. 4.5 Percentage of $p44-18$ transcript population in the cell culture, C3H/SCID C3H/HeJ, and C3H/HeN mouse spleen tissues at day 12 p.i. The total RNA was extracted from spleen specimens and $A. \textit{phagocytophilum}$ infected HL-60 cells, and cDNA was synthesized. PCR was performed to amplify cDNA encoding $p44$ paralogs and colony hybridization was performed to analyze the percentage of $p44-18$ cDNA clone (N=3).
Fig. 4.5

![Graph showing the percentage of p44-18 transcript in total transcript across different cell cultures. The graph compares cell culture (C3H/S CID, C3H/HeN, C3H/HeJ).]
Fig. 4.6 recA expression levels in cell culture, C3H/SCID and C3H/HeN mouse spleen samples on day 12 p.i. RNA was extracted from spleen specimens and infected HL-60 cells and cDNA was synthesized. Real time PCR was performed to analyze recA expression levels. Real time PCR amplification for the A. phagocytophilum 16S rRNA was performed to normalize the inputs of RNA across samples. Relative recA transcript amount = $2^{-(\text{Avg. recA transcript Ct} - \text{Avg. 16S rRNA Ct})} \times 10^4$ Significant (*, $P < 0.05$; **, $P < 0.01$) different compared with cell culture by the unpaired Student $t$-test.
Fig. 4.6

![Graph showing relative recA transcript amount for C3H/HeN, C3H/SCID, and cell culture. The graph includes error bars and asterisks for significance.](image-url)
4.5 Discussion

This study showed an important role of host TLR4 for the defense against *A. phagocytophilum* HZ infection. Since *A. phagocytophilum* is LPS deficient (Lin and Rikihisa, 2003a), TLR4 may recognize other components of the bacterium. Heat shock protein could be one of the candidate proteins recognized by TLR4. *Chlamydia* heat shock protein 60 activates macrophages and endothelial cells through TLR4 (Bulut et al., 2002). It was reported that TLR4 and CD14 were constitutively expressed on monocytes at a high level, while neutrophils express these two molecules at low levels. In contrast, basophils express TLR4 but not CD14, while eosinophils express none of them (Sabroe et al., 2002). Previous work showed *A. phagocytophilum* induced the IL-1β, TNF-α, and IL-6 expression in human monocytes in vitro as early as 2 h p.i. (Kim and Rikihisa, 2002), whereas the expression of these cytokines was delayed in neutrophils (Borjesson et al., 2005; Kim and Rikihisa, 2002). This suggests that the deficiency of TLR4 could impair the activation of monocytes by *A. phagocytophilum*, thus likely its clearance.

In the present study, although antibody against P44N and hvP44-18 was developed in all ICR, C3H/HeN and C3H/HeJ mice, P44N antibody development was delayed in C3H/HeJ mice. In addition to control the innate immunity, another important function of TLRs is to provide a linkage to adaptive immune responses (Akira et al., 2001). Dendritic cells (DCs) are central to T lymphocyte activation and differentiation. DCs take up antigens and present the antigenic peptides on the relevant MHC molecules (Iwasaki and Medzhitov, 2004). Immature DCs express surface TLRs that bind to microbes or microbial products, which are then internalized and processed by the DCs. DCs are heterogeneous in terms of TLR expression. TLR4 is expressed on the monocyte-
derived human and mouse DCs (Iwasaki and Medzhitov, 2004; Akira et al., 2001). It was reported that the reduced number of CD4+ T cells responding to Salmonella were observed in TLR4-deficient mice compared with wild-type (Vazquez-Torres et al., 2004). The reduced number of CD4+ cell may be responsible for the delay the P44N antibody development. The P44 hypervariable region is immunodominant and speculated to divert immune responses away from conserved regions of P44 (Wang et al., 2006). The development of P44-18hv antibody seems earlier and stronger than that of P44N antibody in all infected mice tested and in infected horses (unpublished data). A significant reduction in the development of P44-18hv antibody in TLR4 deficient mice was not detected. The P44N and P44-18hv C-C epitope recognition pattern and IFA titers were not different between C3H/HeJ and C3H/HeN mice (Wang et al., 2006).

A competent immune response was expected to narrow down p44 transcript diversity as it cleared organisms expressing specific p44 species. The original inoculum phenotype p44-18 transcript disappeared from both C3H/HeJ and C3H/HeN spleen specimens by day 12 p.i. This may be due to the efficient development of P44-18hv antibody in both mouse strains. These data suggest that antibody to the P44 hypervariable region plays an important role for the P44 phenotype population change, whereas anti-P44N antibody is critical for bacterial clearance. In SCID mice, which are T- and B- cell deficient, the p44-18 transcript population was also decreased at day 12 p.i.

In our recent study (Lin et al., 2006), we detected the recombination intermediate structure of A. phagocytophilum p44s in blood samples from infected SCID mice and HL-60 cells by Southern blot and/or PCR analysis followed by sequencing. These sequences were half-crossover structures consisting of the donor p44 locus connected to a
3’ conserved region of the recipient *p44* in the *p44*-expression locus in direct orientation. *A. phagocytophilum* has homologs of genes for the RecF pathway, but not RecBCD or RecE pathways of recombination. To determine whether the *p44* recombination intermediate structure can be generated in a RecF-active *E. coli* strain, we constructed a double-origin plasmid carrying the *p44*-expression locus and a donor *p44* locus and introduced this into isogenic mutant *E. coli* strains. The recombination intermediate was recovered in an *E. coli* strain with an active RecF recombination pathway, but not in strains with a deficient RecF pathway (Lin et al., 2006). Our results suggested that the *p44* gene conversion occurs through two successive half-crossover events, initially at a 3’- conserved region and then at a 5’-conserved region of the molecule. *p44* sequence variations seen at the border within the conserved 5’ and 3’ regions are compatible with previous reports of homologous recombination (Kowalczykowski et al., 1994; Kobayashi, 1992). Of seven *A. phagocytophilum* genes homologous to genes involved in RecA-dependent RecF pathway of recombination in *E. coli* (Kowalczykowski et al., 1994), only *recA* expression was upregulated in the immune competent mice as well as in SCID mice compared with in cell culture, suggesting the recombination rate of *p44* homolog genes may increase in the immunocompetent mammalian infection. In *Neisseria gonorrhoeae* transformed with inducible RecA, there was Rec-A dependent recombination of *PilE/S*. Pilin antigenic variation occurs by RecA-dependent unidirectional transfer of DNA sequences from a silent pilin locus to the expressed pilin gene through high-frequency recombination events that occur at limited regions of homology (Criss et al., 2005; Koomey et al., 1987). It appears that SCID mouse *recA*
expression levels were slightly less than in immunocompetent mice. Further studies are necessary to compare RecA activities of *A. phagocytophilum* in immune and nonimmune environments.

These results suggest that TLR4 is involved in differential development of anti-P44N and anti-P44hv antibodies. Delayed development of P44N allows evasion of the humoral immune response and persistence of *A. phagocytophilum* in C3H/HeJ mice. *recA* upregulation facilitates more divergent P44 population expression, possibly allowing evasion of the humoral immune response in immunocompetent mice.
4.6 Summary

Expressed populations of polymorphic \( p44 \) paralogs encoding the 44-kDa immunodominant major outer membrane proteins of \( A. \) phagocytophilum drastically changes during mammalian infection, but not in HL-60 cell culture. In the present study we investigated whether transcription of genes involved in homologous recombination and development of antibodies against conserved P44 N-terminus region (P44N) and P44 central hypervariable (hv) region (P44hv) are related to the diverse P44 expression and clearance of \( A. \) phagocytophilum infection in mice. C3H/HeN and ICR mice cleared \( A. \) phagocytophilum infection by 3 weeks post inoculation (p.i.), whereas \( A. \) phagocytophilum persisted at least two month p.i. in ICR/SCID and TLR4-deficient C3H/HeJ mice. In cell culture, \( A. \) phagocytophilum \( p44-18 \) was dominant transcript population, but C3H/HeJ and C3H/HeN mice cleared \( p44-18 \) transcript population by day 12 p.i., and SCID mice had reduced \( p44-18 \) transcript population. Numbers of different \( p44 \) species expressed were greatest in C3H/HeJ > C3H/SCID > C3H/HeN. Development of antibodies against P44-18hv were similar between C3H/HeJ and C3H/HeN mice, but the development of antibodies against P44N was delayed in C3H/HeJ mice compared with that in C3H/HeN and ICR mice. Of seven genes involved in homologous recombination only \( recA \) mRNA levels in \( A. \) phagocytophilum were significantly greater in mice than in cell culture. These results suggest that slower development of anti-P44N relative to anti-P44hv, and up-regulation of \( recA \), may be involved in appearance of more diverse P44 variants and delayed clearance of \( A. \) phagocytophilum in TLR4-deficient mice.
CHAPTER 5

IDENTIFICATION OF A DNA-BINDING PROTEIN FOR UPSTREAM REGION OF P44 OPERON ENCODING MAJOR OUTER MEMBRANE PROTEINS OF ANAPLASMA PHAGOCYTOPHILUM

5.1 Introduction

*Anaplasma phagocytophilum*, the rickettsial agent of Human granulocytic anaplasmosis (HGA), is maintained in nature via a complex enzootic cycle involving *Ixodes scapularis* ticks and small rodents (Telford *et al.*, 1996). During transmission, the bacterium differentially expresses many of its constituent proteins for adaptation to its diverse host environments. Among those differentially regulated in this manner are major outer surface membrane protein P44s (Felek *et al.*, 2004; Wang *et al.*, 2004; Ijdo *et al.*, 2002; Lin *et al.*, 2002; Zhi *et al.*, 2002b). P44s are upregulated in *A. phagocytophilum* from at 24º C to 37ºC than at (Zhi *et al.*, 2002b) and in mice compared in tick salivary glands (Zhang *et al.*, 2006), and these proteins are essential for infection of mammalian host cells in vitro (Wang *et al.*, 2006).

Given the importance of P44s in the life cycle of *A. phagocytophilum*, and the pathogenesis of HGA (Wang *et al.*, 2006), the elucidation of the regulatory networks that govern the differential expression of P44s has become a central focus for understanding
the molecular mechanisms by which *A. phagocytophilum* adapts to its disparate host environments. However, the discernment of the molecular basis of gene regulation in *A. phagocytophilum* generally has been hampered by a lack of systems for genetically manipulating the obligatory intracellular bacteria. Nonetheless, recent advances in *A. phagocytophilum* genome sequence (Dunning Hotopp et al., 2006) provide us with tools to analyze potential *A. phagocytophilum* virulence factors (Cheng et al., 2006; Niu et al., 2006). The discovery of differential temperature-dependent and host–dependent P44 expression (Zhang et al., 2006; Zhi et al., 2002b) led to an important question concerning how P44 expression is regulated. The specific p44-expression locus was identified in *A. phagocytophilum* (Barbet et al., 2003; Lin et al., 2003), which consists of four tandem genes, *tr1*, *omp-1X*, *omp-1N* (*p44ESup1*) and *p44E*. The polycistronic transcriptional start site was detected 20 bp upstream of *tr1* (Lin et al., 2003). Additional transcriptional start sites are present upstream of *omp-1N*, *omp-1X*, and *p44E* (Zhang et al., 2006; Barbet et al., 2005; Barbet et al., 2003; Lin et al., 2003).

Based on determinations of transcriptional initiation by 5’RACE, *tr1* has been predicted to possess a typical −35/−10 σ70 promoter (Lin et al., 2003). One possibility is that σ70 controls *tr1*-p44 expression via an unidentified transactivator, which could bind to the regulatory region for the activation of *tr1*. Additional experiments are therefore warranted to define whether a potential DNA binding protein transactivates a σ70 promoter. Regarding initial efforts to investigate *tr1* promoter activity, Barbet *et al.* (Barbet et al., 2005) first showed that a -40 bp promoter region of the *tr1* is capable of driving the expression of a green fluorescent protein (*gfp*) reporter gene in a surrogate *E. coli* background. Despite these efforts, a direct examination of *A. phagocytophilum* DNA
binding protein has been lacking. To this end, we performed electrophoresis mobility shift assay (EMSA) with a DNA probe within the upstream regulatory region of tr1; this approach has allowed us to identify an A. phagocytophilum protein that binds to cis element potentially involved in tr1 regulation. The implications of our findings relative to the control of p44 expression are discussed.

5.2 Materials and methods

A. phagocytophilum and host cell-free bacteria preparation. The A. phagocytophilum HZ strain was cultured in HL-60 cells as previously described (Rikihisa et al., 1997). Host cell-free A. phagocytophilum was prepared by sonicating A. phagocytophilum-infected HL-60 cells (> 95% infected cells, 30% organisms start to release) for 8 s at an output setting of 2 with an ultrasonic processor W-380 (Heat Systems, Farmingdale, N.Y.). After low-speed centrifugation to remove nuclei and unbroken cells, in order to remove the host cell debris, the supernatant was sequentially passed through 5 µm (Whatman Inc. Clifton, NJ) and 0.8 µm (MILLIPORE Cork, Ireland) filters for EMSA analysis, or 5 µm (Whatman) filter only for isolation of the DNA binding protein. The filtrate was centrifuged at 10,000×g for 10 min, and the pellet enriched with host cell-free A. phagocytophilum was washed by 2×PBS (274 mM NaCl, 20 mM Na2HPO4, 5.4 mM KCl, and 3.6 mM KH2PO4 [pH 7.2]) and centrifuged at 10,000×g for 10 min. The pellet was stored at -80°C until use.

DNA probe synthesis. Total DNA was extracted from the A. phagocytophilum infected HL-60 cells with a QIAamp blood kit (QIAGEN, Valencia, Calif.). PCRs were performed in 100-µl reactions containing 50 ng of DNA, 20 pmol of each primer, 0.4
mM of each of the four deoxynucleoside triphosphates, 5 U of Taq DNA polymerase, and 1.5 mM MgCl₂, with 2 min of denaturation at 94°C, followed by 40 cycles of 45 sec of denaturation at 94°C, 1 min of annealing at 55°C, and 30 sec of extension at 72°C. The transcriptional start site of tr1 was previously determined at -20 of the translational start site (ATG) by 5’ RACE (Lin et al., 2003). To amplify the upstream DNA fragments of tr1, primer pairs shown in Table 1 were used. Primer set 1 amplified region 2 (R2) from -165 bp to -435 bp upstream of the tr1 transcriptional start site (271 bp, Fig. 5.3A). Primer set 2 amplified region 3 (R3) from 40 bp down stream to -215 bp upstream of the tr1 transcriptional start site (255 bp). Primer set 3 amplified the region 3-1 (R3-1), the 5’ segment of R3, from -65 to -215 bp upstream of tr1 transcriptional start site (151 bp). Primer set 4 amplified the 3’ segment of R3 (R3-2), from 40 bp downstream to -85 bp upstream of the tr1 transcriptional start site (125 bp).

To make the biotin-labeled probe for EMSA analysis, approximately 5 pmol of the heat denatured PCR product was biotinylated by the Biotin 3’ End DNA Labeling Kit (Pierce Biotechnology, Inc. Rockford, IL). The biotinylated probes were denatured at 94°C for 3 min, then slowly cooled down to the room temperature to allow proper annealing of probes. To make the biotin-labeled probe for the purification of DNA binding proteins, 5’ biotin-labeled of primer set 3 (Table 5.1) was used for PCR amplification of the R3-1 probe. PCR was performed under the same conditions described above. PCR product was precipitated by first adding 1/10 volume of 3M sodium acetate, pH 5.2 (final concentration of 0.3 M). After mixing, 3 volumes of cold 100% ethanol were added followed by gentle mixing. The mixture was placed at -20°C overnight. The precipitated DNA was centrifuged at 18,800×g for 30 min at 4°C. The
pellet was rinsed with 70% ethanol twice to remove residual salt. After air drying at room temperature for 30 min, the pellet was dissolved in 25 µl water.

**Electrophoretic mobility shift assay (EMSA).** Bacterial pellets were suspended in binding buffer containing 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM DTT, 0.4% NP-40 and 1% protease inhibitor cocktail (Calbiochem, San Diego, CA). Bacterial lysates were prepared by sonication four times using an ultrasonic processor at setting 2 for 5 s and centrifuged at 18,800×g for 30 min at 4°C. Bacterial lysates supernatants, bovine serum albumin or HL-60 cell lysates (5 µg protein) were incubated with biotinylated DNA probes (0.1 pmol) for 30 min at 4°C in 20 µl of binding buffer (10 mM Tris-HCl, pH 7.5; 50 mM KCl; 5 mM MgCl₂; 1 mM DTT; 2.5% glycerol; 0.1% NP-40; 50 ng/µl of poly[dI:dC] or salmon sperm DNA). As a control, the sample was incubated with an excess (15×) of unlabeled probe. Samples were loaded onto a 6% native polyacrylamide gel in 0.5× Tris–borate–EDTA (TBE) (0.44 M Tris base; 0.44 M boric acid and 0.01 M EDTA, pH 8.0) that was prerun for 1 h, and then electrophoresed at 100 V for 1.5 h at 4°C. The binding reactions were transferred to a nylon membrane at 380 mA for 45 min. Transferred DNA was UV cross-linked to the membrane at 1,200 µJ for 50 s with a Stratalinker UV cross-linker (Stratagene Cloning Systems, La Jolla, Calif.). The Biotin-labeled DNA was detected by the LightShift Chemiluminescent EMSA kit (Pierce).

**DNA binding protein purification.** About 1 mg of bacterial pellets isolated by filtration through a 5 µm filter was used for the preparation of bacterial lysates as described above. Lysate supernatants were pre-absorbed by incubation with 15 µg salmon sperm DNA at 4°C for 15 min, followed by incubation with 50 pmol biotinylated probe for 30 min at 4°C in 600 µl of the same binding buffer described above. The DNA
bound proteins were affinity-purified by the μMACS streptavidin kit (Miltenyi Biotec, Berglsch Gladbach, Germany). The column was set in the magnetic field, and, after equilibration of the column, about 600 µl of binding reaction mixture was applied on the column. The column was washed sequentially by 400 µl binding buffer (wash 1) and 300 µl binding buffer (wash 2) supplemented with 1 M NaCl. The DNA binding proteins were eluted by removing the column out of the magnetic field and applying 100 µl binding buffer on the top of column. The proteins in wash 2 and the final elution were precipitated by TCA and re-suspended in 10 µl binding buffer. Ten µl of each purification fraction was mixed with 2×SDS PAGE gel loading buffer, boiled for 5 min, and subjected to 15% SDS-PAGE gel electrophoresis followed by staining with GelCode Blue Stain Reagent (Pierce Biotechnology).

**Mass spectroscopy.** After electrophoresis, the gel was washed off from the glass plate and soaked in a solution contain 50% ethanol (Fisher Scientific, Pittsburg PA ) and 10% acetic acid (Fisher Scientific) for 1 h followed by soaking the gel in another solution containing 50% methanol (Fisher Scientific) and 10% acetic acid (Fisher Scientific) over night at room temperature with gentle shaking. After washing 15 min in water, the gel was stained with GelCode Blue Stain Reagent (Pierce Biotechnology) at room temperature for 1 h. The gel was then de-stained in water for 3–4 h. The bands were cut from gel with a clean scalpel. The cut bands were washed twice in 200 µl solution containing 50% methanol (Fisher Scientific) and 10% acetic acid (Fisher Scientific) at room temperature for 1 h. The gel was then soaked in 200 µl acetonitrile for 5 min at room temperature and then dried in a speed vacuum for 2 min. The dried gel pieces were then incubated in 75 µl of solution containing 100 mM ammonium bicarbonate and 32.5
mM DTT at room temperature for 30 min followed by incubated in another 75 μl of solution containing 100 mM ammonium bicarbonate and 80 mM iodoacetamide in the dark for 30 min at room temperature. After the gel was sequentially incubated twice in 200 μl of 100 mM ammonium bicarbonate for 5 min at room temperature and then in 200 μl of acetonitrile for 5 min at room temperature, gel slices were dried again in a speed vacuum for 2 min. Gel pieces were then digested in 50 μl of enzyme solution containing 20 μg/ml trypsin, 25 μg/ml chymotrypsin and 50 mM ammonium bicarbonate on ice for 40 min. The enzyme solution was removed and another 30 μl of cold 50 mM ammonium bicarbonate was added. After overnight digestion at room temperature, the digested peptides were collected by incubating each sample with 30 μl of extraction solution (50% acetonitrile, 10% formic acid) at room temperature for 15 min followed by collecting extract by vacuum. The collection step was repeated three times. All extracts were pooled and concentrated to final volume of 30 μl by vacuum. Following digestion, mass spectroscopy was performed on a Thermo Finnigan Linear Quadrupole Ion Trap MSn (Thermo, San Jose, CA). Liquid samples were introduced by Capillary-LC (LC Packings, Sunnyvale, CA).

5.3 Results

*Trl* is conserved in all sequenced members of the family *Anaplasmatacae*. 

TR1 (hypothetical transcriptional regulators, *un1*) has a helix turn helix DNA binding motif that is upstream of 22 tandem *omp-1* (*p28*) loci (Ohashi *et al.*, 2001), and these *omp-1*s were shown to be polycistrionically transcribed from the *tr1* promoter (Ohashi *et al.*, 2001). The *tr1* homolog was detected upstream of the *p44*-expression locus (Lin *et
and *A. marginale* major surface protein 2 (*msp2*) operon-associated proteins (*opags*) expression locus (Brayton *et al.*, 2005), and *E. ruminantium map-1* locus (van Dobbenburgh *et al.*, 1999). *omp-1, msp2, p44, wsp* themselves belong to the OMP-1/Msp2/P44 super family (Dunning Hotopp *et al.*, 2006). Searching the genome sequences revealed that *tr1s* were conserved among all sequenced members of the family *Anaplasmataceae*, including *Neorickettsia* and *Wolbachia* species, but genes homologous to *tr1* (E-value>10⁻¹⁵) were not detected in any other bacteria including the bacteria belong to the closest family, the *Rickettsiaceae* (Table 5.2).

**EMSA analysis to identify the tr1 promoter regions and DNA binding proteins.** The *tr1* transcription site starts at base -20 of the translational start site (ATG) (Lin *et al.*, 2003). The presence of putative transcription factors (proteins) that can bind to the *tr1* upstream DNA sequences was investigated by EMSA. To assess whether an additional 400 bp immediately upstream of the −35 consensus sequence plays a role in *tr1* regulation, the DNA fragments R3 (255 bp) and R2 (271 bp) were PCR amplified (Fig. 5.3A). By using the lysate of *A. phagocytophilum* cultured in HL-60 cells at 37 °C, the DNA fragments, R3 and R3-1, but not R2 or R3-2 were specifically shifted (Fig. 5.3B). This indicates the potential cis elements in the promoter region was between -65 to -215 bp upstream of the *tr1* transcriptional start site and -30 to -180 from the −35 consensus sequence. The unlabeled DNA probe competed for the binding of the labeled probe, demonstrating binding specificity, while BSA or HL-60 cell lysate did not bind this same DNA fragment, indicating the presence of specific binding proteins in the *A. phagocytophilum* lysate.

**Purification of proteins bound to R3-1 probe.** Proteins bound to R3-1 probe
were affinity-purified with the μMACS streptavidin kit. Each purification fraction was subjected to 15% SDS-PAGE gel electrophoresis. The SDS-PAGE electrophoresis revealed two protein bands (10–15kDa) enriched in the final elution (Fig. 5.4).
Bio-, the primer is 5’ biotinylated.

Table 5.1 Sequences of oligonucleotides used in PCR amplification of EMSA probes

<table>
<thead>
<tr>
<th>primer set no.</th>
<th>primer name</th>
<th>Target size (bp)</th>
<th>5’ primer</th>
<th>3’ primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R2</td>
<td>271</td>
<td>TACCTGCACTACGAC GGGTC TGCACTCTTTGTGCCC CCTTT (Bio-)TGCACTCTTTG-TGCCCCCTTT</td>
<td>CTGTTCGTGTGTTAACA TATT GCTTGAGAAGTCATAA ACAT AAACCGTGAATAGCAT GCTAG GCTTGAGAAGTCATAA ACAT</td>
</tr>
<tr>
<td>2</td>
<td>R3</td>
<td>255</td>
<td>TGCACTCTTTGTGCCC CCTTT GCTTGAGAAGTCATAA ACAT</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>R3-1</td>
<td>151</td>
<td>CTAGCATGCTATTCA CGGTTT</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>R3-2</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteria and GenBank No.</td>
<td>GenBank accession No.</td>
<td>Annotation</td>
<td>Identity</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------</td>
<td>------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Anaplasma phagocytophilum HZ CP000235</td>
<td>YP_505749.1</td>
<td>DNA-binding protein (Tr1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. phagocytophilum NY</td>
<td>AY_137510.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaplasma marginale St Maries CP000030</td>
<td>AA_N08711.1</td>
<td>hypothetical protein AM1138</td>
<td>82%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>YP_154239.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ehrlichia chaffeensis Arkansas CP000236</td>
<td>AA_C02808.2</td>
<td>Tr1 putative transcriptional regulator</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>YP_507902.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ehrlichia chaffeensis Sapulpa AAIF00000000</td>
<td>ZP_00545135.1</td>
<td>helix-turn-helix motif</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td>Ehrlichia ruminantium Welgevonden CR767821</td>
<td>AA_R10929.1</td>
<td>hypothetical transcriptional regulator</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YP_197785.1</td>
<td>putative transcriptional regulator</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>YP_180720.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ehrlichia ruminantium Gardel CR925677</td>
<td>CA_I28351.1</td>
<td>hypothetical transcriptional regulator</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>YP_196825.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ehrlichia canis Jake CP000107</td>
<td>YP_303523.1</td>
<td>transcriptional regulator</td>
<td>59%</td>
<td></td>
</tr>
<tr>
<td>Neorickettsia sennetsu Miyayama CP000237</td>
<td>YP_506124.1</td>
<td>DNA-binding protein</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>Wolbachia pipiens wMel wMel AE017196</td>
<td>ZP_00374080</td>
<td>transcriptional regulator</td>
<td>32%</td>
<td></td>
</tr>
</tbody>
</table>

*GenBank accession numbers are given. Percent amino acid sequence identities of the entire open reading frames are relative to the *Anaplasma phagocytophilum* HZ TR1. *E* value of $<10^{-15}$ based on amino acid sequences was used as cut-off.

Table 5.2 Homologous *tr1* genes found in sequenced genomes of the family *Anaplasmataceae*.
Fig. 5.3 EMSA analysis of proteins bound upstream of *trl* upstream region with different DNA probes. (A) Schematic representation of primer locations within the intergenic region of *ndk* and *trl* genes at the *p44* expression locus of *A. phagocytophilum*. (B) DNA binding protein bound at the potential cis elements in the promoter region of *trl*. Lanes 1, 3 and 6 show the free biotin-labeled DNA probes of R3, R3-1 and R3-2, respectively. When purified *A. phagocytophilum* lysate was added, positions of probes R3 (lane 2) and R3-1 (lane 4) were shifted, but R3-2 (lane 7) did not shift. For probe binding competition, a 15-fold excess amount of unlabelled R3-1 probe was used (lane 5). BSA (lane 8) or the uninfected HL-60 cell lysate (lane 9) did not cause the shift with probe R3-1. The Figure is representative of three independent experiments.
Fig. 5.3

A

ndk

R2
271 bp

654 bp

-35--10

R3
255 bp

R3-1
151 bp

R3-2
125 bp

B

bp

R3  R3+AP  R3-1  R3-1+AP  R3-1+AP+Comp.  R3-2  R3-2+AP  R3-1+BSA  R3-1+HL-60

255

151

125

1  2  3  4  5  6  7  8  9
Fig. 5.4 Proteins bound to the R3-1 probe. The bacterial lysate prepared from 1 mg of purified host cell-free *A. phagocytophilum*. The proteins were streptavidin affinity-purified. About 600 μl binding reaction mixture was applied on the column. The column was washed sequentially by binding buffer (wash 1) and binding buffer supplemented with 1 M NaCl (wash 2). The bound proteins were eluted by removing the column out of the magnetic field and applying binding buffer on the column. The proteins in the wash 2 and the final elution was precipitated by TCA and re-suspended in binding buffer. Each purified fraction was subjected to 15% SDS-PAGE gel electrophoresis. M, molecular size markers. Proteins enriched in elution are indicated by the arrow.
5.5 Discussion

The *A. phagocytophilum* genome encodes about 90 *p44* genes comprising about 7% of the total *A. phagocytophilum* genomic coding capacity (Dunning Hotopp et al., 2006). This remarkable feature distinguishes *A. phagocytophilum* from virtually all other pathogenic bacteria and has engendered the contention that the membrane *p44* likely play an important role in the adaptation of *A. phagocytophilum* to both its arthropod and mammalian hosts (Lin et al., 2003). In this regard, it is increasingly well documented that a number of these *p44*s indeed are differentially expressed during the transmission of *A. phagocytophilum* between ticks and mammals (Felek et al., 2004; Zhi et al., 2002b); recent demonstration of the essential roles for the P44s in *A. phagocytophilum* mammalian cell infection in vitro (Wang et al., 2006) further underscores the importance of *p44*s in the complex life cycle of *A. phagocytophilum*. Continuing efforts to elucidate the molecular mechanisms that modulate the expression of *A. phagocytophilum* P44s thus likely will hold the key for understanding how *A. phagocytophilum* survives in nature via its complex adaptive responses.

Previously, we and others identified a novel polycistronic expression locus of *p44E* that may govern the expression of all *p44*s through gene conversion (Barbet et al., 2003; Lin et al., 2003). In the present study, we exploited *tr1* upstream of the operon for understanding the mechanism(s) governing the expression of *p44* genes regulated by the σ70 type promoter based on sequence analysis and primer extension studies (Lin et al., 2003). Presence of a cis-acting element upstream of the *tr1* gene potentially constituted a binding site(s) for a putative transactivator of *tr1* regulation. As such, one attractive
hypothesis has been that \textit{p44}s actually are regulated indirectly via the induction of a requisite transactivator, followed by expression of \textit{trl} via a \(\sigma^{70}\)-like promoter.

By constructing different DNA probes, herein we now have provided several lines of evidence that the \textit{trl} gene likely is regulated directly by the binding of -30 to -180 from the $\sim$35 consensus sequence to its $\sigma^{70}$-dependent promoter. The minimal -30 to -180 from the $\sim$35 \textit{trl} promoter defined in our study is suspected to be responsive to key environmental stimuli typically associated with the regulation of \textit{p44} (e.g., temperature and host).

The housekeeping sigma factor, \(\sigma^{70}\) (RpoD), and the consensus promoter sequence for E\(\sigma^S\) is similar to that used by E\(\sigma^{70}\) (Loewen and Hengge-Aronis, 1994). In fact, a typical \(\sigma^S\)-dependent promoter binds to both E\(\sigma^S\) and E\(\sigma^{70}\) in vitro. However, the \textit{A. phagocytophilum} genome contains only \(\sigma^{70}\) and \(\sigma^{32}\) thereby precluding the regulation with E\(\sigma^S\). Barbet \textit{et al.} (Barbet \textit{et al.}, 2005) recently performed an analysis of promoter elements involved in the expression of \textit{trl} and other upstream genes of \textit{p44E} using a GFP reporter system in \textit{E. coli}. Although our results regarding \textit{trl} expression are largely in agreement with those of Barbet \textit{et al.} (Barbet \textit{et al.}, 2005), some differences are noteworthy. Barbet \textit{et al.} (Barbet \textit{et al.}, 2005) showed that -40 bp upstream of \textit{trl} promoter activity in \textit{E. coli}. In contrast, our study in this region by EMSA indicated that \textit{A. phagocytophilum} lysates did not bind to this region. The reason for this discrepancy is unclear but may be grounded in the fact that we assayed for native \textit{A. phagocytophilum} binding, whereas Barbet \textit{et al.}, (Barbet \textit{et al.}, 2005) used an \textit{E. coli} system that lacks \textit{A. phagocytophilum} proteins.
Therefore, this discrepancy may be due to differences in $\sigma^{70}$ function, plasmid topology, or other *trans*-acting factors that exist between the two species. Although surrogate systems can be valuable, studying *A. phagocytophilum* gene regulation in the relevant native background is preferable when feasible.

*Tr1* homologs were found in all sequenced members of the family *Anaplasmataceae*, and found polycistronically transcribed with downstream genes encoding outer membrane proteins in *Anaplasma* and *Ehrlichia* species, which appear to be regulated by similar environmental cues (Zhang *et al.*, 2006; Zhi *et al.*, 2002b; Unver *et al.*, 2001). Therefore, the fact that *tr1* of *A. phagocytophilum* is regulated by $\sigma^{70}$ and a transactivator protein may be applicable to the regulation of other outer membrane protein genes. On the other hand, such extrapolation to other outer membrane proteins will require further experimental corroboration, as presented herein for *A. phagocytophilum tr1*. This is particularly important given the fact that although *E. canis* *p30* paralogs are induced by elevated temperature, *P30-10* responsiveness to temperature differs from that of other *p30* paralogs (Unver *et al.*, 2001). Additional transcriptional start sites were detected upstream of *p44E* and *OMP-1X* (Zhang *et al.*, 2006; Barbet *et al.*, 2003). It is therefore conceivable that another layer of gene regulation, yet to be elucidated, remains for the regulation of *p44* genes.
5.6 Summary

*A. phagocytophilum* 44-kDa immunodominant major outer membrane proteins are encoded by *p44* paralogs, which can recombine with *p44E* at the *p44*-expression locus, the primary expression locus. To characterize transcriptional regulation of *p44* genes, the present study examined DNA binding proteins that may regulate polycistronic *p44E* transcription. Electrophoretic migration shift assay showed that specific shift of the promoter region of *tr1* which was previously shown to have a polycistronic transcriptional start site upon incubation with *A. phagocytophilum* protein lysate, using biotinylated DNA probes. Using streptavidin-magnetic beads, bound proteins were purified. *tr1* homologs were found in all sequenced members of the family *Anaplasmataceae* in genomes of *Ehrlichia* and *Anaplasma* species, *tr1* is found upstream of the expression loci of genes encoding major outer membrane proteins. Identification of transcriptional factors that regulate the expression of immunodominant major outer membrane proteins will facilitate our understanding the regulation of major outer membrane protein gene expression in members of *Anaplasma* and *Ehrlichia* species.


