INFECTION AND IMMUNITY, July 1990, p. 2076–2084
0019-9567/90/072076-09$02.00/0

Vol. 58, No. 7

The 56-Kilodalton Major Protein Antigen of *Rickettsia tsutsugamushi*: Molecular Cloning and Sequence Analysis of the *sta*56 Gene and Precise Identification of a Strain-Specific Epitope

CHARLES K. STOVER,[1]† DAVID P. MARANA,[1]‡ J. MARK CARTER,[2] BRUCE A. ROE,[3] E. MARDIS,[3] AND EDWIN V. OAKS[1]§*

Department of Rickettsial Diseases[1] and Department of Immunology,[2] Walter Reed Army Institute of Research, Washington, D.C. 20307, and Department of Chemistry, University of Oklahoma, Norman, Oklahoma 73019[3]

Received 29 January 1990/Accepted 5 April 1990

Lasting immunity against *Rickettsia tsutsugamushi*, the causative agent of scrub typhus fever, has been demonstrated to be strain specific. Two protein antigens of 110 and 56 kilodaltons (kDa) have been shown to exhibit strain-specific epitopes. The 56-kDa scrub typhus antigen (Sta56) is an abundant outer membrane protein of *R. tsutsugamushi* and is an antigen often recognized by humans infected with this obligate intracellular bacterium. In this study the complete gene encoding Sta56 (strain Karp) was cloned into pBR322 on a 2.3-kilobase genomic *Hin*dIII DNA fragment and the complete 56-kDa polypeptide was expressed in *Escherichia coli*. DNA sequence analysis of the 2.3-kilobase *Hin*dIII fragment revealed an open reading frame large enough to encode a 56-kDa polypeptide. A putative signal sequence was identified at the deduced amino terminus of the Sta56 polypeptide, and pulse-chase analysis of maxicells labeled with [^{35}S]methionine demonstrated that a higher-molecular-weight precursor matures into the 56-kDa polypeptide. Epitope scanning analysis with synthetic peptides derived from the deduced amino acid sequence identified an octapeptide (located from amino acid residues 117 to 124) that was reactive with a Karp strain-specific monoclonal antibody (K13F88A). Other epitopes recognized by different monoclonal antibodies, including another Karp strain-specific monoclone (K1E106), were localized to different regions of the protein based on their reactivities with λgt11 recombinants expressing various portions of the *sta*56 gene.

Rickettsia tsutsugamushi, the causative agent of scrub typhus fever, is a slow-growing, obligate intracellular bacterium. The scrub typhus rickettsiae, which are transmitted to humans by infected chiggers, replicate and spread within the host. Symptoms of scrub typhus fever include fever, rash, and often an eschar at the site of the chigger bite. The immune response generated by infection with *R. tsutsugamushi* consists of both humoral and cellular components. Even with antibiotic (doxycycline or chloramphenicol) treatment, the immune response remains essential for clearance of the rickettsiae because these antibiotic treatments are not rickettsiacidal (33, 34). Although vaccines for human use are not presently available, previous studies in a mouse model have shown protection by whole irradiated rickettsiae (6). Because large-scale growth and purification of the scrub typhus rickettsiae are prohibitively expensive, future human vaccines will most likely be derived from recombinant organisms expressing *R. tsutsugamushi* protein antigens. To this end, eight different antigen genes (scrub typhus antigen genes *sta*150, -110, -72, -58, -56, -49, -47, and -20) have been cloned and expressed in *Escherichia coli* (22, 23, 37).

Serological surveys have indicated that three prototype strains (Karp, Kato, and Gilliam) are the most prominent and represent the majority of *R. tsutsugamushi* isolates. These strains can be distinguished by various degrees of virulence in animal models and unique antigenic characteristics. More important, previous studies in humans have shown that the protective immune response is long lasting (1 to 3 years) against homologous strains and short lived (3 months) against heterologous strains (25, 34). Although the protective antigens of *R. tsutsugamushi* have not been identified, efforts to characterize the strain-specific epitopes have identified two proteins (Sta110 and Sta56) that have strain-specific antibody reactivity and strain-specific electrophoretic mobilities on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (11, 22, 23, 38). Both of these proteins also contain additional epitopes that are cross-reactive with similar protein homologs in other *R. tsutsugamushi* strains, corroborating DNA hybridization data that identified homologous gene sequences in the Kato and Gilliam strains probed with Karp strain *sta*56 and *sta*110 gene probes (22).

The *R. tsutsugamushi* Karp Sta56 protein (also called the 60-kilodalton (kDa) antigen by Hanson [11] and the 56-kDa protein by Tamura et al. [38]) is a major antigen exposed on the surface of scrub typhus rickettsiae (11, 38), and it is often recognized by the humoral immune system of infected animals and humans (E. V. Oaks, unpublished data). It is present in high concentrations in the rickettsial cell and has solubility properties similar to those of major outer membrane proteins (10, 24). The immunogenicity of this protein, the presence of strain-specific epitopes, and its surface exposure suggest that it may be an important factor in the generation of a protective immune response.

In this study we report the molecular cloning, nucleotide sequencing, and expression of the complete *sta*56 gene in *E. coli*. A potentially important Karp strain-specific determinant was located within an 8-amino-acid span on the Sta56

* Corresponding author.
† Present address: Molecular Vaccines Incorporated, Gaithersburg, MD 20878.
‡ Present address: Division of Rickettsial Diseases, Naval Medical Research Institute, Bethesda, MD 20814.
§ Present address: Department of Enteric Infections, Walter Reed Army Institute of Research, Washington, DC 20307.

polypeptide. The data and recombinants generated in this study should facilitate the molecular cloning and expression of genes encoding the potentially protective Sta56 homologs from other strains of *R. tsutsugamushi* and will provide a source of Sta56 protein for future vaccine trials.

MATERIALS AND METHODS

Bacterial strains and vectors. *E. coli* BNN97, Y1089, and Y1090 were used as host strains for λgt11 (41). *E. coli* HB101 was used as the host for all plasmids (17). M13 recombinants were grown and maintained in *E. coli* JM101 (20). The *E. coli* vectors used in this study were λgt11, pBR322, M13mp18, and M13mp19.

Three plaque-purified reference strains (Karp, Kato, and Gilliam) of *R. tsutsugamushi* were used throughout this study. The passage histories of these strains are as follows: Karp strain passaged in embryonated chicken eggs 52 times, in L cells 3 times, and in eggs 6 times (E52-L3-E6); Kato strain (E162-L3-E8); and Gilliam strain (E164-L3-E8). Rickettsiae were grown in mouse fibroblasts (L-929 cells) incubated at 34°C in a humidified atmosphere of 5% CO_2 and 95% air. The rickettsiae were harvested 5 to 7 days postinfection and subsequently purified on Renografin (E. R. Squibb & Sons, New Brunswick, N.J.) density gradients as previously described (23).

Recombinant DNA procedures. *R. tsutsugamushi* genomic DNA was isolated from purified rickettsiae as previously described (23). All restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase, and bacterial alkaline phosphatase were obtained from New England BioLabs (Beverly, Mass.) and employed as per the instructions of the manufacturer. λgt11 libraries were constructed by using a modified method employing conversion adaptor oligonucleotides as described previously (36, 40). Antigen-positive recombinant phage were identified by screening with hyperimmune rabbit sera prepared as described by Oaks et al. (23).

Recombinant λgt11 phage DNA was purified as described by Silhavy et al. (32). Plasmid DNA was prepared by a modification (17) of the procedure of Birnboim and Doly (3). Plasmid DNA was transformed into *E. coli* HB101 competent cells (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) as specified by the manufacturer.

All nonpreparative agarose (FMC Bioproducts, Rockland, Maine) gel electrophoresis was performed in Tris-acetate buffer, and the DNA was visualized with ethidium bromide (Sigma Chemical Co., St. Louis, Mo.) (17). Cloned inserts were separated from recombinant λgt11 phage DNA by *Eco*RI digestion and agarose gel electrophoresis on 5% low-melting-point agarose (International Biotechnologies, Inc., New Haven, Conn.). The restriction fragment band was excised, melted at 60°C in 50 mM Tris (pH 8.0)–100 mM NaCl buffer, and purified with Elu-Tips (Schleicher & Schuell Co., Keene, N.H.) according to the specifications of the manufacturer.

Southern blot or colony hybridization was performed as described by Maniatis et al. under stringent hybridization (35°C in 50% formamide, 6× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 2× Denhardt solution) and washing (62°C, 100 mM NaCl, 0.1% SDS) conditions (17). DNA probes were α-^{32}P labeled with [α-^{32}P]dCTP (Dupont, NEN Research Products, Boston, Mass.) by nick translation with a nick-translation kit (Bethesda Research Laboratories). Unincorporated nucleotides were removed from the nick translation reaction by size fractionation on a 5-ml column of Sephadex G-50 medium (Pharmacia Fine Chemicals, Piscataway, N.J.) in 50 mM Tris–5 mM EDTA (pH 7.8). Autoradiography was accomplished with Kodak Blue Brand film and X-ray film cassettes with intensifying screens at −70°C.

A single-stranded template was constructed for DNA sequencing by cloning the complete 2.3-kilobase (kb) *Hin*dIII restriction fragment insert of pRTS56H2.3 and overlapping smaller internal subfragments in phage M13mp18 and M13mp19. Single-stranded M13 phage DNA was prepared as described previously (2). DNA sequencing was accomplished by a modification of the dideoxynucleotide chain-termination method of Johnston-Dow et al. (13, 28, 29). DNA sequencing of the M13mp18 and M13mp19 derivatives was initiated by priming with the M13 universal primer (Bethesda Research Laboratories). Both strands of the 2.3-kb restriction fragment were progressively sequenced to completion by priming with additional synthetic oligonucleotides based on previous 3′ experimental data. Synthetic oligonucleotides were prepared on a DNA synthesizer (Applied Biosystems, Foster City, Calif.).

Analysis of rickettsial antigens. Pooled rabbit anti-*R. tsutsugamushi* serum, exhaustively absorbed (23) with Y1089 and λgt11, was used for screening recombinant phage and plasmids and also for Western immunoblot experiments as previously described (21, 23). In addition, antibodies that were affinity purified against λgt11 recombinants expressing Sta56 epitopes were also used to evaluate plasmid recombinants for expression of the Sta56 protein either in Western blot procedures or colony blot assays with unlysed bacteria lifted onto nitrocellulose. Staphylococcal protein A (Pharmacia), radioiodinated by the chloramine T method (12) or conjugated with alkaline phosphatase (Organo Teknika Corp., West Chester, Pa.), was used to detect the antibody bound to antigens in the Western blot assay. Alkaline phosphatase-conjugated probes were developed with fast red TR salt and naphthol AS-MX phosphate (Sigma) (31). Monoclonal antibodies (MAbs) K13F88A, K1E106, Kt2A34, and KCH47 have been previously characterized (5, 11). Convalescent human sera (kindly provided by Gregory Dasch) from patients diagnosed with scrub typhus were diluted 1:200 in casein filler and reacted with lysates of Karp and Kato strains of *R. tsutsugamushi* in Western blot assays.

Protein electrophoresis and Western blotting. SDS-PAGE and Western blotting of rickettsial polypeptides from *R. tsutsugamushi* or recombinant proteins expressed in *E. coli* were performed as previously described (21). Most protein electrophoresis gels consisted of 13% acrylamide except those used for maxicell analysis (see below).

Identification of plasmid-encoded proteins. Plasmid-encoded proteins were identified in *E. coli* HB101 by a modified maxicell procedure (27, 35). Modifications included the use of HB101 instead of CSR603, the doubling of media volume with M9 plus 1% Casamino Acids (Difco Laboratories, Detroit, Mich.) immediately after UV irradiation, and the addition of freshly prepared cycloserine (Sigma) to 200 μg/ml 2 h after irradiation and 2 h before harvest and labeling of the irradiated CSR603 cells with [^{35}S]methionine (Dupont, NEN). Maxicells were labeled after incubating for 1 h in methionine assay medium (Difco) to deplete methionine. The labeled maxicell preparations were analyzed by discontinuous SDS-PAGE, using 17% polyacrylamide (Bio-Rad Laboratories, Richmond, Calif.) gels. Autoradiography was performed on gels after electroelution onto nitrocellulose with Kodak Blue Brand X-ray film. After autoradiography the nitrocellulose sheets were used in the Western blot procedures with an enzymatic probe.

TABLE 1. *R. tsutsugamushi* polypeptide antigens recognized by human convalescent sera[a]

Human serum	Recognition of polypeptide antigen[b]				
	47	56	58	72	110
104	+	+	−	+	−
105	−	+	+	−	−
106	−	+	+	−	+
110	+	+	+	−	−
111	+	+	−	−	−
112	+	+	+	−	+[c]
119	+	+	+	+	+

[a] All the human sera were obtained approximately 2 weeks after the onset of scrub typhus fever. They were reacted with the Karp and Kato strains of *R. tsutsugamushi* by Western blot analysis.

[b] +, Reactivity by Western blot analysis; −, negative reaction against both the Karp and Kato strains. Numbers in subheadings indicate molecular masses of proteins in kilodaltons.

[c] Serum 112 reacted only with the 110-kDa protein in the Kato strain of *R. tsutsugamushi*.

Epitope scanning of Sta56 polypeptide. Overlapping octamer peptides were prepared essentially as originally described by Geysen et al. (8). Briefly, prederivatized polyethylene pins (Cambridge Research Biochemicals, Valley Stream, N.Y.) were deprotected, washed, neutralized, washed, and amino acylated (with blocked 9-fluorenomethoxycarbonyl amino acids [Milligen; Millipore Corp., Bedford, Mass.]) repeatedly until the octamers were completed. The peptides were then acetylated, deblocked, and washed thoroughly once more. Finally, the peptide pins were subjected to three rounds of ultrasonic disruption before the first enzyme-linked immunosorbent assay was performed. The pins were screened with MAbs K13F88A (previously described to recognize a strain-specific epitope on the Sta56 protein), KCH47, and Kt2A34. The antibodies (final dilution, 1:200) were incubated with the immobilized peptides overnight. Next the pins were washed with 10 mM Tris buffer (pH 7.4) containing 0.1% Tween 20, incubated with alkaline phosphatase-conjugated goat anti-mouse immunoglobulin (Kirkegaard & Perry Laboratories, Inc., Gaithersburg, Md.) for 4 h, washed again, and then incubated with phosphatase substrate (*p*-nitrophenylphosphate [1 mg/ml] in diethanolamine buffer [pH 9.8]) for 1 h. The conjugate was also incubated with the pins to determine the degree of nonspecific binding by the conjugate. The sequence selected for epitope mapping (amino acid residues 72 through 134) was derived from data indicating that the strain-specific epitope was encoded by a 0.37-kb *Sau*3A fragment within the *sta*56 gene.

RESULTS

Recognition of the Sta56 protein by human convalescent sera. Sera from humans previously infected with *R. tsutsugamushi* contained antibodies that were reactive with the Sta56 protein antigen (Table 1). In addition, Sta56 was recognized more frequently (in all seven sera tested) than other *R. tsutsugamushi* antigens, including the Sta58, Sta47, and Sta110 proteins.

Molecular cloning of the *sta*56 major antigen gene. Recombinant λgt11 clones expressing incomplete segments of the *R. tsutsugamushi* 56-kDa antigen gene were identified and isolated from genomic λgt11 libraries as described previously by Stover et al. (37). Purified cloned inserts from the λgt11-*sta*56 recombinants were used as *sta*56 gene probes in Southern blot analysis of *R. tsutsugamushi* genomic DNA to identify larger restriction fragments potentially containing a complete *sta*56 gene (Fig. 1). Restriction fragments homologous to the λgt11-*sta*56 gene probes were observed for each of three *R. tsutsugamushi* prototype strains. This analysis and previous work with restriction enzymes recognizing 4-base-pair palindromes revealed a moderate degree of restriction fragment length polymorphism (22). Genomic libraries were constructed in plasmid vector pBR322 with *Eco*RI-, *Hin*dIII-, *Cla*I-, and *Bgl*II-digested *R. tsutsugamushi* chromosomal DNA (Karp strain). *E. coli* HB101 recombinants were selected by colony hybridization with the *sta*56 gene probes and assayed for reactivity with anti-*R. tsutsugamushi* (Karp strain) hyperimmune rabbit serum and anti-56-kDa protein polyvalent antisera that had been affinity purified with λgt11-*sta*56 recombinant phage lysates (23). Recombinant colonies selected by colony hybridization with the λgt11-*sta*56 probe also reacted with the affinity-purified α-Sta56 sera without the lysis of colonies by chloroform or SDS to expose cytoplasmic proteins. One recombinant plasmid, designated pRTS56H2.3, was selected for further analysis because it contained the smallest DNA insert expressing determinants of the Sta56 antigen. Plasmid pRTS56H2.3 was shown by Southern blot hybridization (data not shown) to carry the same 2.3-kb *Hin*dIII fragment previously identified by genomic Southern blot analysis. Western blot analysis of the HB101(pRTSH2.3) transformant showed the expression of an antigenic protein comigrating with the 56-kDa antigen of *R. tsutsugamushi* (Fig. 2, lanes A and B). This 56-kDa recombinant antigen was easily distinguished from the native *R. tsutsugamushi* 58-kDa antigen, designated Sta58 (Fig. 2, lane A), and a recombinant Sta58 encoded by plasmid pRTS58H2.9 (36) (Fig. 1, lane C). Similar Western blots with MAbs K13F88A and K1E106 (data not shown) to the Sta56 antigen confirmed that the 56-kDa protein encoded on plasmid pRTS56H2.3 is the immunodominant Sta56 antigen of *R. tsutsugamushi*.

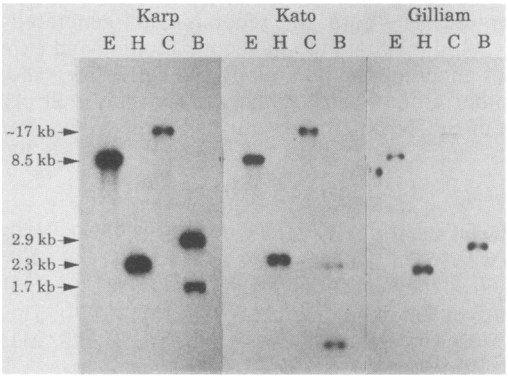

FIG. 1. Southern blot analysis of *R. tsutsugamushi* genomic DNA with *sta*56 gene probes. *R. tsutsugamushi* genomic DNAs of the three prototype strains (Karp, Kato, and Gilliam) were digested with *Eco*RI (E), *Hin*dIII (H), *Cla*I (C), and *Bgl*II (B), electrophoresed on a 0.7% agarose gel, and subjected to Southern blot hybridization with *sta*56 gene probes derived from cloned inserts of a λgt11 recombinant (λgt11 Rts12) expressing determinants of the Sta56 antigen (23). Restriction fragment patterns obtained with the λgt11-*sta*56 gene probes for the *Eco*RI, *Hin*dIII, and *Cla*I digests of the Karp, Kato, and Gilliam DNAs are identical, whereas restriction fragment length polymorphisms are shown in the *Bgl*II digestion.

Nucleotide sequence of pRTS56H2.3. Nucleotide sequencing of the pRTS56H2.3 2.3-kb *Hin*dIII fragment disclosed an open reading frame large enough to encode a 57-kDa protein

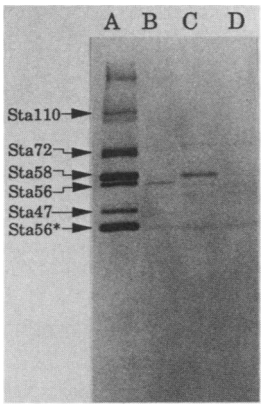

FIG. 2. Expression of Sta56 antigen by an *E. coli* HB101 (pRTS56H2.3) recombinant. Lysates of *E. coli* HB101(pBR322) and recombinant derivatives were subjected to SDS-PAGE and analyzed by Western blotting with rabbit anti-*R. tsutsugamushi* Karp for comparison with the antigen profile of purified *R. tsutsugamushi*. Lanes: A, *R. tsutsugamushi* Karp; B, *E. coli* HB101(pRTS56H2.3); C, *E. coli* HB101(pRTS58H2.9); D, *E. coli* HB101(pBR322). Five immunodominant scrub typhus antigens (Sta) are indicated. Sta56* indicates the 43-kDa heat-modifiable form of the Sta56 antigen (10).

(Fig. 3). Multiple methionine codons were identified within the first 35 bases of this open reading frame. Only the first (most 5′) of the three potential start codons was preceded by a consensus sequence resembling a ribosomal binding site. The potential ribosomal binding site (AGGA) was located within 7 bases of the first potential start codon. Three reasonable matches to the consensus sequence for procaryotic promoters were found within 500 bases upstream from the putative start codon of the *sta56* open reading frame. Two of the candidate promoters exhibited spacing between the −10 and −35 regions that was somewhat less than or greater than the 15 to 21 bases thought to separate most −10 and −35 regions. The most distant candidate promoter exhibited reasonable spacing (16 bases) between the −10 and −35 regions. A potential hairpinlike loop with an inverted repeat forming a stem with an energy of −21 kcal, and characteristic of a rho-independent transcription terminator, was located 3′ to the *sta56* open reading frame (Fig. 3). In addition to these potential control sequences, a 19-base direct repeat of unknown significance was found within the *sta56* open reading frame between bases 322 and 364.

The 40% G+C content of the sequenced 2.3-kb *Hin*dIII restriction fragment is comparable with the *R. tsutsugamushi* experimentally determined genomic G+C content of 35% (unpublished data). Because the Sta56 protein is abundant in *R. tsutsugamushi* and because the DNA sequence of the *R. tsutsugamushi sta56* gene is among the first determined for this A+T-rich rickettsia, a table (Table 2) summarizing codon usage was constructed. The codon frequencies for the *sta56* gene were compared with those of highly expressed genes in *E. coli* (14). The codon usage in the 60% A+T-rich *sta56* gene was substantially different from that of the highly expressed *E. coli* genes, with a general bias for adenine or thymine at the third base of each codon. For example, the CTG (Leu) and ATC (Ile) codons, which are among the most common in *E. coli*, were rarely used in the sequenced scrub typhus *sta56* gene and were replaced by Leu codons TTA, CTA, CTT, and TTG and Ile codons ATA and ATT, which are rarely employed in *E. coli*.

Amino acid sequence and putative signal peptide of the Sta56 protein. The deduced amino acid sequence of the *sta56* gene product indicated a capacity for coding a 56,749-Da protein. The Sta56 protein was extremely rich in glutamine (7.9%). In particular, one region encoded between bases 1010 and 1060 had 11 of 15 amino acids as glutamine. The 5.57 (without the signal peptide) or 5.87 (with the signal peptide) pI value calculated for the deduced Sta56 protein was also in agreement with the native Sta56 pI value of 5.5 to 5.8 determined by preparative isoelectric focusing of detergent-solubilized proteins of *R. tsutsugamushi* (E. V. Oaks and C. K. Stover, unpublished data). Hydrophobicity profiles calculated for the deduced amino acid sequence of the Sta56 protein indicated a polypeptide backbone consisting of alternating hydrophobic and hydrophilic domains (Fig. 4). One notable hydrophobic domain was found at the N terminus of the deduced amino acid sequence. The deduced amino acid sequence of the N-terminal 22 amino acids strongly resembles that of a signal peptide for membrane insertion (38). The signal peptide-like nature of the N terminus was indicated by the two positively charged lysine residues at the extreme N terminus (residues 2 and 3) followed by a stretch of hydrophobic or uncharged amino acids. The putative signal peptidase cleavage site is between Ala-22 and Ile-23, based on the amino acid sequence determination of purified 56-kDa antigen (24). This cleavage site is in agreement with other signal peptide cleavage sites (39). Interestingly, the Sta56 signal peptide has a high degree of identity (47%) with the listeriolysin O signal peptide of *Listeria monocytogenes* and to a lesser extent with the first 165 amino acids of listeriolysin O (Fig. 5) (19). The calculated molecular weight of the deduced *sta56* gene product without the putative signal sequence was in very close agreement with the molecular weight extrapolated by SDS-PAGE migration. Of the three potential start codons mentioned above, only initiation from the first (most 5′) methionine codon was consistent with the integrity of the proposed signal peptide.

Previous studies have provided evidence suggesting that the Sta56 protein is located in the outer membrane and exposed on the surface of *R. tsutsugamushi* (10, 38). Because the presence of an N-terminal signal peptide was indicated by sequence data, pulse-chase experiments were performed to detect evidence of Sta56 posttranslational modification for the removal of the putative signal peptide. Pulse-chase experiments in maxicells labeled with [^{35}S]methionine did show a rapid shift from a slightly larger 57-kDa band to the characteristic 56-kDa form of the Sta56 protein (Fig. 6).

Strain-specific epitope mapping of the Sta56 antigen. Previous studies of the 56-kDa polypeptide indicated small apparent molecular weight differences and strain-specific reactivity with monoclonal antibodies among the Karp, Kato, and Gilliam strains of *R. tsutsugamushi* (5, 11, 22, 23, 38). To locate the strain-specific determinant on this polypeptide, we constructed λgt11 subclones of the *sta56* gene from the pRTS56H2.3 plasmid. Resulting incomplete *sta56* gene products were examined for reactivity with a set of MAbs to the Sta56 antigen. This set of MAbs included strain-specific K13F88A and K1E106 MAbs described by Eisemann et al. (5). The Western blot data are summarized in Table 3. The termini of the recombinant λgt11 cloned insert DNAs were nucleotide sequenced to disclose the boundaries of the gene segments encoding epitopes. This analysis identified and located three (K13F88A, Kt2A34, and K1E106) unique determinants on the Sta56 polypeptide, including the strain-specific epitopes recognized by MAbs K13F88A and

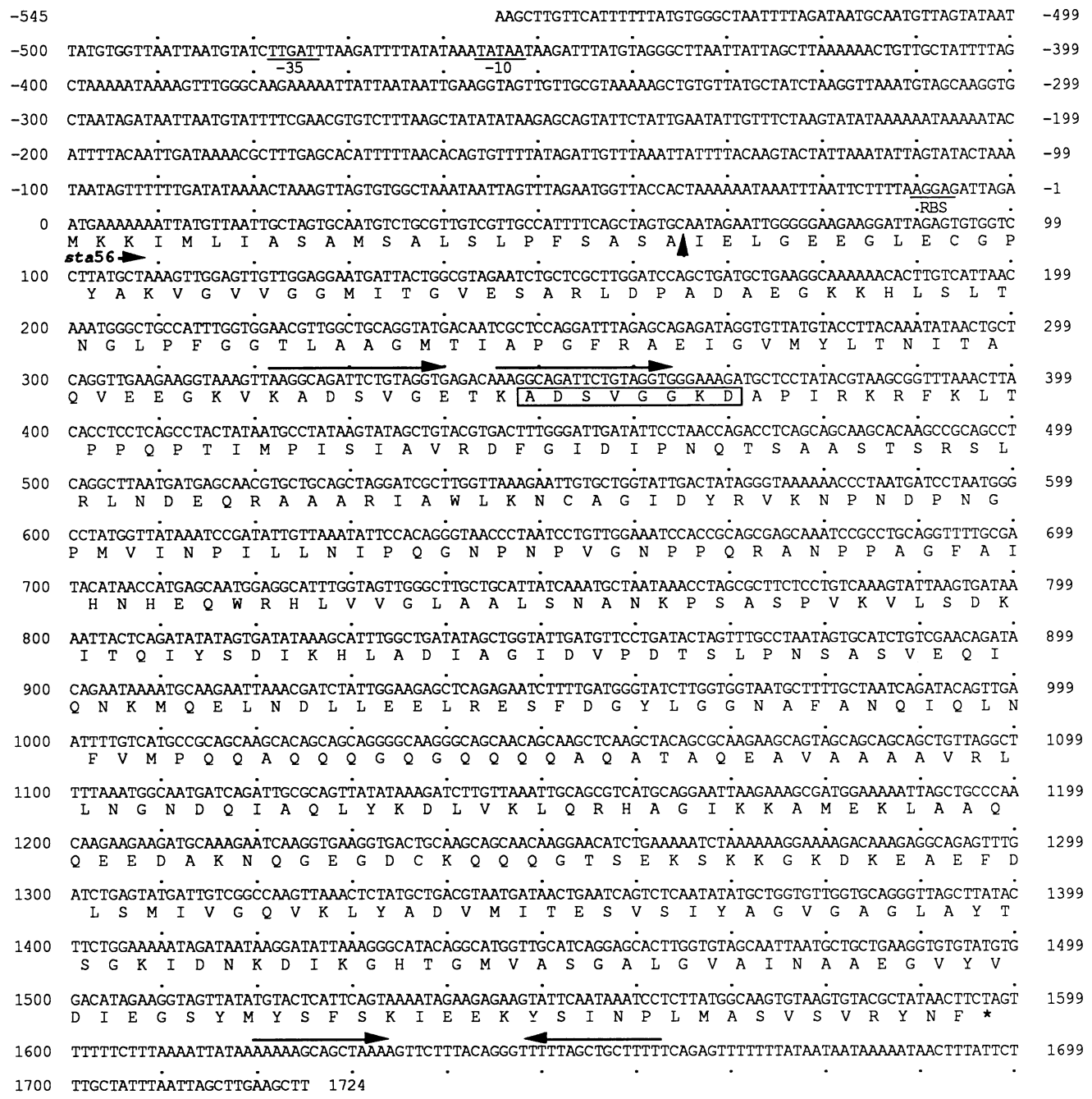

FIG. 3. Nucleotide sequence and deduced amino acid sequence of the *R. tsutsugamushi* Karp *sta*56 gene and flanking 5′ and 3′ regions. The first base of the presumed *sta*56 initiation codon is numbered 0. The presumed initiation codon for the Sta56 protein is denoted by a rightward arrow (5′ to 3′). Sequences resembling the consensus sequences for ribosomal binding sites (RBS) and promoter −10 and −35 regions are underlined. A 19-base-pair direct repeat is indicated by two long rightward arrows. Indirect repeats presumed to be rho-independent transcription terminators are also indicated by one set of opposing arrows. An 8-amino-acid residue region identified in this study as a Karp strain-specific determinant (see Fig. 7) is boxed. The GenBank accession number for the *sta*56 sequence is M33004.

K1E106. These data indicate that at least two distinct strain-specific epitopes are present on the Sta56 polypeptide. The reactivity of MAb Kt2A34, which was produced against the Kato strain of *R. tsutsugamushi*, also localizes a scrub typhus reactive epitope between bases 593 and 1138 in the *sta*56 gene.

The region encoding the MAb K13F88A epitope was narrowed down to a 377-base *Sau*3A restriction fragment (bases 160 through 537) by the analysis described above. To more precisely define the MAb K13F88A epitope encoded within the 377-base *Sau*3A fragment, overlapping octameric peptides were synthesized from amino acid 72 to amino acid 127 on polyethylene pins. This region was selected because it was within the *Sau*3A fragment and because it had a region of high hydrophilicity (Fig. 4) that overlapped a region of high antigenic index scores (4). Combined, these parameters often are predictive of antibody-reactive epitopes. The peptide series was reacted with MAb K13F88A in an enzyme-

TABLE 2. Comparison of codon usage in the *sta*56 gene and *E. coli*

Amino acid	Codon	Codon fraction		Amino acid	Codon	Codon fraction	
		*sta*56	*E. coli*			*sta*56	*E. coli*
Ala	GCG	0.08	0.31	Asn	AAT	0.78	0.24
	GCA	0.41	0.23		AAC	0.22	0.76
	GCT	0.49	0.28				
	GCC	0.02	0.19	Pro	CCG	0.14	0.65
					CCA	0.21	0.20
Cys	TGT	0.67	0.42		CCT	0.64	0.09
	TGC	0.33	0.58		CCC	0.00	0.06
Asp	GAT	0.78	0.51	Gln	CAG	0.57	0.73
	GAC	0.22	0.49		CAA	0.43	0.27
Glu	GAG	0.30	0.27	Arg	AGG	0.31	0.00
	GAA	0.70	0.73		AGA	0.13	0.01
					CGG	0.06	0.03
Phe	TTT	0.83	0.43		CGA	0.06	0.02
	TTC	0.17	0.57		CGT	0.25	0.58
					CGC	0.19	0.35
Gly	GGG	0.23	0.07				
	GGA	0.25	0.04	Ser	AGT	0.33	0.06
	GGT	0.42	0.48		AGC	0.11	0.22
	GGC	0.10	0.41		TCG	0.03	0.06
					TCA	0.25	0.08
His	CAT	0.86	0.39		TCT	0.28	0.26
	CAC	0.14	0.61		TCC	0.00	0.26
Ile	ATA	0.57	0.01	Thr	ACG	0.06	0.20
	ATT	0.38	0.37		ACA	0.50	0.06
	ATC	0.05	0.62		ACT	0.39	0.24
					ACC	0.06	0.50
Lys	AAG	0.28	0.23				
	AAA	0.72	0.77	Val	GTG	0.06	0.27
					GTA	0.35	0.23
Leu	TTG	0.32	0.08		GTT	0.44	0.37
	TTA	0.30	0.06		GTC	0.15	0.13
	CTG	0.05	0.69				
	CTA	0.03	0.02	Tyr	TAT	0.86	0.41
	CTT	0.22	0.09		TAC	0.14	0.59
	CTC	0.08	0.07				

linked immunosorbent assay (Fig. 7). By this method the K13F88A epitope was precisely located to amino acids 117 through 124 (ADSVGGKD). This region partially overlaps one of the two hexameric amino acid repeats encoded by the 19-base direct repeat between amino acids 322 and 364. MAbs KCH47 and Kt2A34, which react with the Sta47 and Sta56 polypeptides, respectively (but further downstream [Table 3]), did not react with the synthetic peptides used in these experiments.

DISCUSSION

The problems associated with growing large quantities of the obligate intracellular bacterium *R. tsutsugamushi* have impeded the identification of protective antigens and the development of a vaccine. This hindrance is compounded by the fact that there are numerous strains of *R. tsutsugamushi* and that lasting immunity to scrub typhus infection is strain specific (25). In this paper we report the molecular cloning and DNA sequence analysis of a gene encoding the 56-kDa outer membrane major protein antigen of *R. tsutsugamushi*. This protein appears to be an immunodominant antigen in humans infected with scrub typhus rickettsiae. All human immune sera tested to date, from individuals naturally infected with scrub typhus rickettsiae, have shown reactivity with the 56-kDa antigen. The Sta56 antigen is also one of two *R. tsutsugamushi* protein antigens (the other being the Sta110 protein) identified that exhibit strain-specific characteristics (11, 22, 24, 38).

```
1'  MKKIMLVFITLILVSLPIAQQTEAKDASAFNKENSISSMAPPASPPASPKTPIEKKHADE
    :::::::. .. .::::.    . .: ....:. .. . .. .. : .:. . :
1"  MKKIMLIASAMSALSLPF-----SASAIELGEEGLECGPYAKVGVVGGMITGVESARLDP

61'  IDKYIQGLDYNKNNVLVYHGDAVTNVPPRKGYKDGNEYIVVEKKKKSINQNNADIQVVNA
     :     . . .: : . :. .....  :.... . . ... . ..... . :..
56"  ADA-EGKKHLSLTNGLPFGGTLAAGMTIAPGFRAEIGVMYLTNITAQVEEGKVKADSVGE

121'  ISSLTYPG--ALVKANSELVENQPDVLPVKRDSLTLSIDLPGMTNQDKNIVVKNATKSNV
      ... . : : .. . .:. ::...::.. . ...:.:. :.
115"  TKADSVGGKDAPIRKRFKLTPPQPTIMPISIAVRDFGIDIPNQTSAASTSRSLRLNDEQR
```

FIG. 5. Sequence homology between amino-terminal ends of *L. monocytogenes* listeriolysin O and the Sta56 polypeptide. The alignment was performed with the Fast P program (16). The top sequence is listeriolysin O, and the bottom sequence is Sta56. A colon represents amino acid identity, and a period represents a conserved or similar amino acid. The putative signal peptides are the first 25 and the first 22 amino acids of the listeriolysin and Sta56 proteins, respectively.

Previous studies have provided evidence that the Sta56 protein antigen has solubility properties similar to those of major outer membrane proteins of gram-negative bacteria (10, 24). Our analysis of the recombinant Sta56 protein in *E. coli* supports these previous findings in that the protein was accessible to antibody without lytic treatment of the host *E. coli*. The identification of a putative signal peptide and evidence of posttranslational processing provide further support for membrane insertion of the Sta56 protein. Comparison of the deduced amino acid sequence of the recombinant protein with recent N-terminal sequence data for purified Sta56 protein antigen definitively identifies the site of signal peptidase cleavage in *R. tsutsugamushi* between amino acid residues 22 and 23 of the deduced Sta56 amino acid sequence, since residues 23 through 56 of the recombinant Sta56 protein are identical to the 34 N-terminal residues determined by amino acid sequencing (24). The posttranslational processing of the recombinant Sta56 protein observed in *E. coli* pulse-chase experiments probably removes the N-terminal 22-residue signal peptide identified in Fig. 5. The function of the abundant Sta56 major outer membrane protein in *R. tsutsugamushi* is not known. The alternating hydrophobic-hydrophilic motif suggests that the polypeptide backbone includes a number of transmembrane domains,

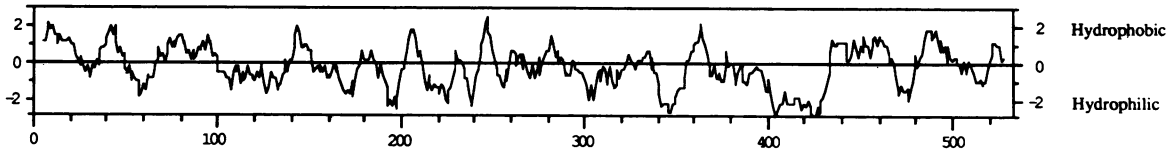

FIG. 4. Hydrophobicity profile of the Sta56 outer membrane protein antigen. Hydrophobicity was plotted by the method of Kyte and Doolittle (15).

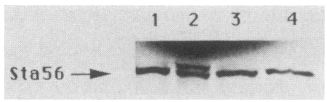

FIG. 6. Pulse-chase analysis of recombinant Sta56 by using maxicells. Maxicell cultures of recombinant Sta56 were harvested 5, 15, and 60 min after the maxicells were pulsed with [^{35}S]methionine. Casamino Acids were added to the culture at 5 min postpulse but before collection of the 5-min sample. Lanes 2, 3, and 4 are the samples taken at 5, 15, and 60 min, respectively. Lane 1 is a maxicell preparation of recombinant Sta56 that was harvested at 60 min without a pulse-chase. In lane 2 there is a doublet that gradually matured into only one lower-molecular-weight band by 15 min (lane 3) and was still there after 60 min (lane 4). The representative segment shown is from a 17% acrylamide gel to which 25,000 acid-precipitable counts of each sample were added per lane.

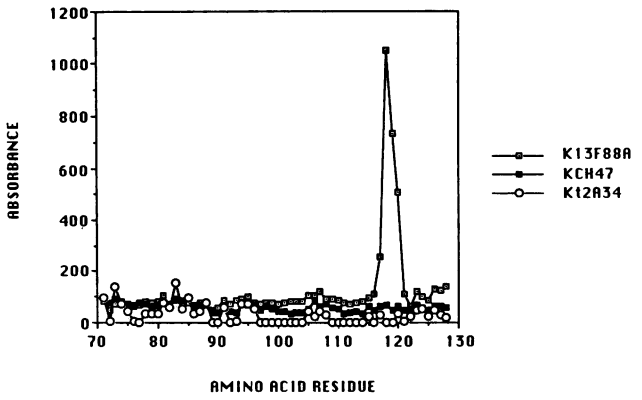

FIG. 7. Identification of Karp strain-specific epitope recognized by MAb K13F88A with synthetic peptides. Synthetic octamer peptides of the Sta56 sequence from amino acid residues 71 through 128 were tested for reactivity with MAbs K13F88A, KCH47, and Kt2A34 as described in Materials and Methods. The absorbance value represents the actual absorbance value multiplied by 1,000. Amino acid residue refers to the first amino acid of the synthetic octapeptide.

with certain regions of the protein exposed on the surface of the organism. Surface exposure of the Sta56 protein has been previously demonstrated with surface-specific iodinations of intact *R. tsutsugamushi* (38) and by antibody binding to strain-specific epitopes of Sta56 on scrub typhus rickettsiae (11). It is not known whether the Sta56 protein is involved in the attachment and invasion of host cells, but it has been demonstrated that antibody-mediated blockage of invasion by *R. tsutsugamushi* is strain specific (9). The moderate degree of homology (47% identity in the signal peptide and 18% identity plus 44% similar or conserved amino acids in the first 165 amino acids of the two proteins) with *L. monocytogenes* listeriolysin O (which is thought to be involved in the lysis of phagocytic vesicles by listeriae) (7) also indicates that the Sta56 protein is a candidate for a rickettsial protein involved in the invasion of eucaryotic host cells. However, since it is still not possible to genetically manipulate viable rickettsiae, we have not been able to generate mutations in various antigen genes to determine their role in pathogenesis.

The Sta56 protein is among the most abundant membrane proteins in *R. tsutsugamushi*, but expression of the cloned *sta56* gene in *E. coli* was relatively poor. However, even with the marginal expression of the recombinant *sta56* gene, the Sta56 antigen was easily detected by Western blot and maxicell analyses. The low level of expression of an *R. tsutsugamushi* gene in *E. coli* is not a general phenomenon, since another scrub typhus antigen gene cloned in this laboratory (*sta47*) is expressed at substantially higher levels in *E. coli* than in *R. tsutsugamushi* and is under the control of a native rickettsial promoter (unpublished data). Sequences with reasonable homologies to the consensus *E. coli* promoter (−10 and −35 regions) and ribosomal binding sites were identified for *sta56*. *E. coli*-like promoters and ribosomal binding site consensus sequences have also been observed in other cloned rickettsial genes (40), and the use of an *E. coli*-like promoter was verified in the case of the 17-kDa antigen gene of *Rickettsia rickettsii* (1). Codon usage in the *sta56* gene resembled that determined for six other *R. tsutsugamushi* genes (36; unpublished data) in that adenines and thymines were more frequently found in the third base of each codon, as would be expected for a 60% A+T-rich gene. The amino acid composition determined for the deduced sequence agrees with the chemically determined amino acid composition (24), except that our sequence predicts only 4% glycine residues compared with 14% determined by amino acid analysis. Further study of the Sta56 antigen as a protective immunogen will require better expression of this gene. Because lethality due to elevated expression of recombinant outer membrane proteins is a common phenomenon, it is plausible that uncontrolled overproduction of the Sta56 outer membrane protein (at levels comparable to that seen in *R. tsutsugamushi*) would be lethal to *E. coli*. However, deleterious effects to the *E. coli* HB101 host were not observed at the level of expression observed with the pRTS56H2.3 plasmid.

The three epitope regions, including the strain-specific epitopes recognized by MAbs K13F88 and K1E106, defined on the Sta56 protein indicate that the strain specificity associated with this antigen is complex and not localized in one linear domain of the protein. The likelihood of other strain-specific epitopes is good, since it has been shown that substantial differences exist between the Sta56 homologs of the Karp, Kato, and Gilliam strains of *R. tsutsugamushi* within the first 35 N-terminal amino acid residues of the mature proteins (24). Whereas this study has focused on the mapping of B-cell epitopes, which may be involved in

TABLE 3. Reactivity of MAbs with *sta56* subclones

Clone[a]	DNA fragment location in *sta56*[b]	Clone product[c]	MAb reactivity[d]		
			K13F88A	Kt2A34	K1E106
λgt11RtsS3A2	160–537	FP	+	−	−
λgt11Rts50	0–890	35 kDa	+	−	−
λgt11RtsS3A15	924–1138	FP	−	+	−
λgt11Rts303	593–1313	FP	−	+	+
λgt11Rts12	593–1598	FP	−	+	+

[a] The clones listed were constructed as described by Stover et al. (37). Clone λgt11Rts12 was described previously (23).

[b] DNA fragment location was determined by DNA sequence analysis of the ends of each fragment. The base pair location refers to the *sta56* open reading frame.

[c] Clone product refers to the size of the protein antigen synthesized by a lysogen that was reactive with rabbit anti-Karp serum. The apparent size was determined by SDS-PAGE and Western blotting. FP indicates that the reactive lysogen product was a fusion protein in that it reacted with both rabbit anti-Karp and rabbit anti-β-galactosidase. The apparent molecular sizes of the fusion proteins were all greater than 116 kDa, the size of unfused β-galactosidase.

[d] MAbs K13F88A, Kt2A34, and K1E106 were previously characterized (5, 11). MAbs K13F88A and K1E106 are strain specific for the Karp strain of *R. tsutsugamushi* (5).

opsonization of extracellular rickettsiae and possibly in blocking rickettsiae from entering cells (9), previous studies on the host immune response to *R. tsutsugamushi* have indicated that clearing intracellular rickettsiae requires cellular immune responses (30). Therefore an effective scrub typhus vaccine will have to stimulate the appropriate cellular and humoral immune systems. Computer algorithms (18, 26) have been employed to identify a number of potential T-cell epitopes on the Sta56 protein antigen (unpublished data). Additional analysis of these regions is required to determine their actual reactivity with protective T cells.

If the Sta56 protein shows potential as a protective vaccine for scrub typhus, it would be necessary to clone and express the genes encoding the related antigens from other strains of *R. tsutsugamushi* to provide heterologous protection. Ohashi et al. have shown that the amino acid differences at the N termini of the Karp, Kato, and Gilliam Sta56 homologs can be explained by codon wobble (24). We have shown by Southern blot hybridization that strong DNA homology exists between the *sta56* genes of the Karp strain and the genes of the Kato and Gilliam strains, although restriction fragment length polymorphism existed for this gene (22). Therefore it should be possible to rapidly amplify, isolate, clone, and express the genes encoding the Sta56 protein homologs from other strains of *R. tsutsugamushi* by using probes and/or sequence data furnished by this study in conjunction with polymerase chain reaction technology.

ACKNOWLEDGMENTS

We thank Jerry Buysse and Carole Hickman for their careful review of the manuscript.

LITERATURE CITED

1. **Anderson, B. E., B. R. Baumstark, and W. J. Bellini.** 1988. Expression of the gene encoding the 17-kilodalton antigen from *Rickettsia rickettsii*: transcription and posttranslational modification. J. Bacteriol. **170**:4493–4500.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. E. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
3. **Birnboim, H. C., and J. Doly.** 1979. Rapid plasmid extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1525.
4. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. University of Wisconsin genetic computer group sequence analysis software package. Nucleic Acids Res. **12**:387–395.
5. **Eisemann, C. S., and J. V. Osterman.** 1985. Identification of strain-specific and group reactive antigenic determinants on the Karp, Gilliam, and Kato strains of *Rickettsia tsutsugamushi*. Am. J. Trop. Med. Hyg. **34**:1173–1178.
6. **Eisenberg, G. H. G., and J. V. Osterman.** 1978. Gamma-irradiated scrub typhus immunogens: development and duration of immunity. Infect. Immun. **22**:80–86.
7. **Gaillard, J.-L., P. Berche, J. Mounier, S. Richard, and P. Sansonetti.** 1987. In vitro model of penetration and intracellular growth of *Listeria monocytogenes* in the human erythrocyte-like cell line Caco-2. Infect. Immun. **55**:2822–2829.
8. **Geysen, H. M., R. H. Meloen, and S. J. Barteling.** 1984. Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. Proc. Natl. Acad. Sci. USA **81**:3998–4002.
9. **Hanson, B.** 1983. Effect of immune serum on infectivity of *Rickettsia tsutsugamushi*. Infect. Immun. **42**:341–349.
10. **Hanson, B.** 1984. Characterization of *Rickettsia tsutsugamushi* outer membrane proteins, p. 240–243. *In* L. Lieve and D. Schlessinger (ed.), Microbiology—1984. American Society for Microbiology, Washington, D.C.
11. **Hanson, B.** 1985. Identification and partial characterization of *Rickettsia tsutsugamushi* major protein immunogens. Infect. Immun. **50**:603–609.
12. **Hunter, W. M., and F. C. Greenwood.** 1962. Preparation of iodine-131 labelled human growth hormone of high specific activity. Nature (London) **194**:495–496.
13. **Johnston-Dow, L., E. Mardis, C. Heiner, and B. A. Roe.** 1987. Optimized methods for fluorescent and radiolabeled DNA sequencing. BioTechniques **5**:754–765.
14. **Konigsberg, W., and G. N. Godson.** 1983. Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:687–691.
15. **Kyte, J., and F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**: 105–132.
16. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
17. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Margalit, H., J. L. Spouge, J. L. Cornette, K. B. Cease, C. Delisi, and J. A. Berzofsky.** 1987. Prediction of immunodominant helper T cell antigenic sites from the primary sequence. J. Immunol. **138**:2213–2229.
19. **Mengaud, J., M. Vicente, J. Chenevert, J. M. Pereira, C. Geoffroy, B. Gicquel-Sanzey, F. Baquero, J. Perez-Diaz, and P. Cossart.** 1988. Expression in *Escherichia coli* and sequence analysis of the listeriolysin O determinant of *Listeria monocytogenes*. Infect. Immun. **56**:766–772.
20. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
21. **Oaks, E. V., T. L. Hale, and S. B. Formal.** 1986. Serum immune response to *Shigella* protein antigens in rhesus monkeys and humans infected with *Shigella* spp. Infect. Immun. **48**:124–129.
22. **Oaks, E. V., R. M. Rice, D. J. Kelly, and C. K. Stover.** 1989. Antigenic and genetic relatedness of eight *Rickettsia tsutsugamushi* antigens. Infect. Immun. **57**:3116–3122.
23. **Oaks, E. V., C. K. Stover, and R. M. Rice.** 1987. Molecular cloning and expression of *Rickettsia tsutsugamushi* genes for two major proteins antigens in *Escherichia coli*. Infect. Immun. **55**:1156–1157.
24. **Ohashi, N., A. Tamura, M. Ohta, and K. Hayashi.** 1989. Purification and partial characterization of a type-specific antigen of *Rickettsia tsutsugamushi*. Infect. Immun. **57**:1427–1431.
25. **Rights, F. L., J. E. Smadel, and E. B. Jackson.** 1948. Studies on scrub typhus (tsutsugamushi disease). III. Heterogeneity of strains of *Rickettsia tsutsugamushi* as demonstrated by cross vaccination studies. J. Exp. Med. **87**:339–351.
26. **Rothbard, J. B., and W. R. Taylor.** 1988. A sequence pattern common to T cell epitopes. EMBO J. **7**:93–100.
27. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**: 692.
28. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
30. **Shirai, A., P. J. Cantanzaro, S. M. Phillips, and J. V. Osterman.** 1976. Host defenses in experimental scrub typhus: role of cellular immunity in heterologous protection. Infect. Immun. **14**:39–46.
31. **Sidberry, H., B. Kaufman, D. C. Wright, and J. Sadoff.** 1985. Immunoenzymatic analysis by monoclonal antibodies of bacterial lipopolysaccharides after transfer to nitrocellulose. J. Immunol. Methods **76**:299–305.
32. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Smadel, J. E., C. A. Bailey, and F. H. Diercks.** 1950. Chloramphenicol (chloromycetin) in the chemoprophylaxix of scrub typhus (tsutsugamushi disease). IV. relapses of scrub typhus in treated volunteers and their prevention. Am. J. Hyg. **51**:229–241.
34. **Smadel, J. E., H. L. Ley, F. H. Diercks, and R. Traub.** 1950.

Immunity in scrub typhus: resistance to induced reinfection. Arch. Pathol. **50**:847–867.

35. **Stover, C. K., J. Kemper, and R. C. Marsh.** 1988. Molecular cloning and characterization of *supQ/newD*, a gene substitution system for the *leuD* gene of *Salmonella typhimurium*. J. Bacteriol. **170**:3115–3124.
36. **Stover, C. K., D. P. Marana, G. A. Dasch, and E. V. Oaks.** 1990. Molecular cloning and sequence analysis of the Sta58 major antigen of *Rickettsia tsutsugamushi*: sequence homology and antigenic comparison to the 60-kilodalton family of stress proteins. Infect. Immun. **58**:1360–1368.
37. **Stover, C. K., M. H. Vodkin, and E. V. Oaks.** 1987. Use of conversion adaptors to clone antigen genes in λgt11. Anal. Biochem. **163**:398–407.
38. **Tamura, A., K. Takahashi, T. Tsuruhara, H. Urakami, S. Miyamura, H. Sekikawa, M. Kenmotsu, M. Shibata, S. Abe, and H. Nezu.** 1985. Analysis of polypeptide composition and antigenic components of *Rickettsia tustsugamushi* by polyacrylamide gel electrophoresis and immunoblotting. Infect. Immun. **48**:671–675.
39. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
40. **Wood, D. E.** 1989. A preview of rickettsial gene structure and function, p. 93–103. *In* J. Moulder (ed.), Intracellular parasitism. CRC Press, Inc., Boca Raton, Fla.
41. **Young, R. A., and R. W. Davis.** 1983. Yeast RNA polymerase II genes: isolation with antibody probes. Science **222**:778–782.