Comparative Sequence Analysis of a Genus-Common Rickettsial Antigen Gene

BURT E. ANDERSON* AND THEODORE TZIANABOS

Viral and Rickettsial Zoonoses Branch, Division of Viral and Rickettsial Diseases, Center for Infectious Diseases, Centers for Disease Control, Atlanta, Georgia 30333

Received 23 March 1989/Accepted 10 June 1989

The genes encoding the 17-kilodalton genus-common antigen have been cloned and sequenced from Rickettsia conorii, Rickettsia prowazekii, and Rickettsia typhi. Compared with the R. rickettsii sequence, this sequence had a high degree of homology within the coding and control regions (R. conorii, 99.8%; R. prowazekii, 88.1%; R. typhi, 88.7%). The 5' flanking regions, including the promoter and the transcription initiation sites, were extremely well conserved for all four species, suggesting that control and expression of this locus are important to the survival of the rickettsiae.

The 17,000-molecular-weight-antigen (17K antigen) gene of Rickettsia rickettsii is one of the most thoroughly characterized loci from any rickettsial genome (2, 3; B. E. Anderson, Ph.D. dissertation, Georgia State University, Atlanta, 1988). Nucleotide sequences for transcription initiation and putative translation initiation, signal peptide, and lipid modification sites have been identified (2). To determine the degree of genetic conservation among these regions, as well as other areas within the coding region for the mature protein, we cloned and sequenced this genus-common antigen from three additional pathogenic species of rickettsiae: Rickettsia conorii, Rickettsia prowazekii, and Rickettsia typhi. In addition, nucleotide sequences were examined for regions of variability that may serve as group- or species-specific priming sites for polymerase chain reaction amplification of rickettsial DNA from clinical samples.

To determine the level of conservation of the 17K antigen throughout the genus Rickettsia, immunoblot analysis was performed. R. rickettsii Shelia Smith, R. conorii Moroccan, R. prowazekii Breinl, R. typhi Wilmington, Rickettsia bellii RML 360-C, Rickettsia akari Hartford, and Rickettsia canadensis McKiel were grown in Vero cells as previously described (3) and purified by Renografin density gradient centrifugation (14). Purified rickettsiae were solubilized at 100°C and subjected to electrophoresis by the method of Laemmli (8). Resolved proteins in the gel were electroblotted to nitrocellulose filters by the method of Towbin et al. (11). The resulting filters were reacted with rabbit antiserum raised to the synthetic peptide NH₄-Asp-Asn-Gly-Asn-Tyr-Gly-Tyr-Val-Thr-Pro-Asn-Lys-Thr-Tyr-Arg-COOH, which is derived from residues 106 to 120 of the deduced amino acid sequence of the 17K antigen of R. rickettsii (3; Anderson, Ph.D. dissertation). When the filters were reacted with a horseradish peroxidase-conjugated anti-rabbit serum, all species of rickettsiae tested exhibited an antigenic protein with an approximate molecular weight of 17,000 (Fig. 1). This antigenic protein has not been observed in members of other genera, including Bacillus, Proteus, Neisseria, Escherichia, and Chlamydia (data not shown). Thus, the 17K antigen described for R. rickettsii appears to be a genus-common antigen found in members of both the spotted fever and the typhus group rickettsiae.

DNA from R. rickettsii, R. conorii, R. typhi, and R. prowazekii was isolated by lysis of purified rickettsiae with proteinase K and Sarkosyl, followed by phenol-chloroform extractions and subsequent ethanol precipitation. A number of primers homologous to different regions of the coding and flanking sequences of the 17K-antigen gene from R. rickettsii were chosen for polymerase chain reaction amplification of 1 ng of genomic rickettsial DNA. Twenty-five cycles of amplification were performed with a thermal cycler and a Gene-Amp kit according to the directions of the manufacturer (Perkin-Elmer/Cetus, Norwalk, Conn.). Various combinations of primers from each of the four rickettsial species until the entire structural genes and flanking regions of each were obtained. The amplified fragments were cloned into plasmid vector pUC19 and transformed into Escherichia coli DH5α by the method of Hanahan (5), and the DNA sequence was determined for both strands by the Sanger chain termination method (10). The nucleotide sequences from overlapping fragments were connected to provide the sequence for the full-length coding region and promoter (Fig. 2). To ensure

FIG. 1. Immunoblot of solubilized rickettsial proteins reacted with rabbit antiserum specific for the 17K genus-common antigen. Antiserum was raised to the synthetic peptide corresponding to residues 106 to 120 of the R. rickettsii-deduced amino acid sequence. Lane A, R. akari; lane B, R. conorii; lane C, R. bellii; lane D, R. rickettsii; lane E, R. canadensis; lane F, R. typhi; lane G, R. prowazekii. Electrophoresis was performed on a 12.5% acrylamide gel by the method of Laemmli (8). Molecular mass standards (in kilodaltons) are shown at left.

* Corresponding author.
The obtained sequences were responsible for expressing the 17K genus-common antigens and that no polymerase incorporation errors were present, the entire sequence (encoding the full-length gene) from each of the four rickettsial species was reamplified, cloned, and sequenced. The nucleotide sequences from a second independent determination, for each of the four species of rickettsiae, agreed perfectly with those shown in Fig. 2. The cloned genes from each of the four species of rickettsiae directed the synthesis of the 17K genus-common antigen in E. coli DH5α as determined by immunoblotting techniques with the antipeptide serum already described (Fig. 3).

The 5′ flanking control regions as well as the coding regions for the 17K genus-common antigen genes from each of the various species of rickettsiae were well conserved. The observed homology to the R. rickettsii 17K-antigen gene was 99.8% for R. conorii (Fig. 2, lane 2), 88.7% for R. typhi (Fig. 2, line 3), and 88.1% for R. prowazekii (Fig. 2, line 4). The −35 (nucleotides 1 to 6) and −10 (nucleotides 22 to 27) promoter regions previously identified for R. rickettsii were entirely conserved among all four species of rickettsiae (Fig. 2). The adenine residue (indicated as +1 nucleotide in Fig. 2), where transcription initiates in R. rickettsii, was also present in all the rickettsiae sequenced. Likewise, the presumed ribosome-binding site was well conserved, with only slight variations. The actual coding regions of the genes from R. rickettsii (Fig. 2, nucleotides 60 to 536), R. conorii (Fig. 2, nucleotides 60 to 536), R. typhi (Fig. 2, nucleotides 56 to 532), and R. prowazekii (Fig. 2, nucleotides 57 to 530) were also well conserved. At the amino acid level, the sequence of the 17K genus-common antigen of R. conorii was 100% homologous to the R. rickettsii sequence; R. typhi and R. prowazekii were 91.8% homologous (Fig. 4). The signal peptide sequences identified from the R. rickettsii-deduced amino acid sequence (residues 1 to 20) contain three amino acid substitutions for R. typhi (Fig. 4, line 3) and one for R. prowazekii (Fig. 4, line 4). Each of these changes results in substitution with similar hydrophobic amino acids. The tetrapeptide sequence Lys-Gln-Ala-Cys that targets the recombinant expressed antigen for lipid modification was found in the deduced amino acid sequence of all four rickettsiae (Fig. 4, residues 17 to 20).

The high degree of homology found in the coding regions for this genus-common antigen indicates a strong selective pressure for continued expression of this gene among a diverse group of rickettsiae. Furthermore, near-perfect conservation of regions controlling gene expression suggests that a specific level of expression from this genus-common

![FIG. 3. Immunoblot showing cloned rickettsial 17K genus-common antigens expressed in E. coli. Samples were electrophoresed and reacted with antiseraum as described in the legend to Fig. 1. The rickettsial sources of the gene expressed in E. coli are as follows: lane A, R. rickettsii; lane B, R. conorii; lane C, R. typhi; lane D, R. prowazekii; lane E. coli DH5α (pUC19). Molecular mass standards (in kilodaltons) are shown at left.](http://jb.asm.org/)

![FIG. 4. Deduced amino acid sequences for the rickettsial 17K genus-common antigens. Sequences were aligned for maximal homology, with deletions indicated (−). The homology to the R. rickettsii amino acid sequence is indicated at the end of each sequence. Rickettsial sequences are those of R. rickettsii (Rr), R. conorii (Rc), R. typhi (Rt), and R. prowazekii (Rp).](http://jb.asm.org/)
antigen locus is beneficial for the rickettsiae. However, the role of this gene product in the life cycle of the rickettsiae has not been established.

The nucleotide sequence for the *R. typhi* 17K genus-common antigen gene contains 50 separate deletions or base changes compared with the *R. rickettsii* sequence. However, only 13 amino acid residues are affected, indicating that most of the mutations are silent. Similarly, only 13 amino acids from the *R. prowazekii*-deduced amino acid sequence differ from those of the 17K antigen of *R. rickettsii* despite 55 nucleotide changes or deletions. The vast majority of amino acid substitutions between the spotted fever and the typhus group 17K genus-common antigens occur in either neutral or hydrophilic regions. These may represent determinants found on the exterior portion of the antigen and may be surface-exposed epitopes that have undergone antigenic variation. The deduced amino acid sequence for *R. prowazekii* contains a deletion at residue 140; the glutamine residue found in *R. rickettsii, R. conorii, and R. typhi* is absent at this position in the sequence of *R. prowazekii* (Fig. 4).

Comparing nucleotide sequences for a common locus among a number of species of bacteria has assisted attempts at understanding the phylogeny of these organisms (4, 13). In the case of the 17K genus-common antigen, the gene is well conserved among members of both the spotted fever and the typhus group rickettsiae. For this reason, an ancestral form of the gene should have been the source from which this genus-common antigen evolved. The sequence comparisons provide data consistent with the fact that *R. prowazekii* and *R. typhi* evolved from a common source and are much more closely related to each other than to *R. rickettsii or R. conorii*. When the *R. rickettsii* and *R. conorii* 17K genus-common antigen genes and flanking regions were compared, only one nucleotide differed, despite the fact that these two rickettsiae were isolated on different continents. These data, coupled with the fact that DNA-DNA hybridizations between the entire genomes of the two species show greater than 90% homology (12), suggest that a separate species designation for *R. conorii* may not have been warranted. The high degree of similarity between *R. rickettsii* and *R. conorii* is also evident in serological cross-reactivity (1, 9) and cross-reactive T-cell responses (6, 7). Regardless, from the data presented in this report concerning the 17K genus-common antigen gene, it is obvious that *R. rickettsii* and *R. conorii* are highly related. In contrast, areas of divergence within the nucleotide sequence between the members of the spotted fever group (*R. rickettsii* and *R. conorii*) and the typhus group (*R. typhi* and *R. prowazekii*) can be seen in the 17K genus-common antigen gene (Fig. 2). These divergent areas are currently being used as group-specific priming sites for polymerase chain reaction amplification of subpicogram quantities of rickettsial DNA (T. Tzianabos, B. E. Anderson, and J. E. McDade, submitted for publication). This technique should allow detection of rickettsial DNA in blood from infected individuals and provide a much-needed specific and sensitive test for early diagnosis of rickettsial diseases.

We are grateful to Brian Holloway for synthesizing the oligonucleotide primers and the synthetic peptide used to generate specific antiserum to the genus-common antigen. We also thank William Bellini for helpful advice and discussion and Russell Regnery for providing *R. conorii* genomic DNA.

**LITERATURE CITED**


