Systematic Cloning of *Treponema pallidum* Open Reading Frames for Protein Expression and Antigen Discovery

Matthew McKevitt,[1] Krupa Patel,[1] David Smajs,[4] Michael Marsh,[2] Melanie McLoughlin,[5] Steven J. Norris,[4,5] George M. Weinstock,[3] and Timothy Palzkill[1,6]

[1]*Department of Molecular Virology and Microbiology,* [2]*Structural and Computational Biology and Molecular Biophysics Program, and* [3]*Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA;* [4]*Department of Microbiology and Molecular Genetics and* [5]*Department of Pathology and Laboratory Medicine, University of Texas—Houston Medical School, Houston, Texas 77030, USA*

A topoisomerase-based method was used to clone PCR products encoding 991 of the 1041 open reading frames identified in the genome sequence of the bacterium that causes syphilis, *Treponema pallidum* subsp. *pallidum*. Cloning the open reading frames into the univector plasmid system permitted the rapid conversion of the original clone set to other functional vectors containing a variety of promoters or tag sequences. A computational prediction of signal sequences identified 248 *T. pallidum* proteins that are potentially secreted from the cell. These clones were systematically converted into vectors designed to express the encoded proteins as glutathione-S-transferase fusion proteins. To test the potential of the clone set for novel antigen discovery, 85 of these fusion proteins were expressed from *Escherichia coli*, partially purified, and tested for antigenicity by using sera from rabbits infected with *T. pallidum*. Twelve of the 85 proteins bound significant levels of antibody. Of these 12 proteins, seven had previously been identified as *T. pallidum* antigens, and the remaining five represent novel antigens. These results demonstrate the potential of the *T. pallidum* clone set for antigen discovery and, more generally, for advancing the biology of this enigmatic spirochete.

[Supplemental material is available online at www.genome.org.]

The genome sequences of many infectious microorganisms have now been determined. The availability of genome sequence information will enable new approaches to be developed to determine the function of gene products and their possible role in pathogenesis. A proteomics approach is being developed to identify proteins important for the *Treponema pallidum* host-pathogen interaction. As a first step, an efficient high-throughput strategy has been used to clone 96% of the predicted *T. pallidum* open reading frames (ORFs) into a recombination-based vector system.

T. pallidum is the causative agent of syphilis. Syphilis is a multistage infection characterized by localized, disseminated, and chronic manifestations interspersed between periods of latent infection. The molecular mechanisms of *T. pallidum* pathogenesis are poorly understood (Norris et al. 2001). The organism is an obligate human pathogen that has not been cultured continuously in vitro (Norris et al. 2001), precluding the use of many experimental approaches, including direct genetic analysis. Thus, new methods are needed to address questions about the biology and pathogenesis of this organism.

The complete genome sequence of *T. pallidum* was determined in 1998 (Fraser et al. 1998). It consists of a circular chromosome of ~1100 kb and is therefore one of the smallest known prokaryotic genomes (Fraser et al. 1998). There are a total of 1041 predicted ORFs that encompass 93% of the total genomic DNA (Fraser et al. 1998). Biological roles were predicted for 55% of the ORFs based on matches to ORFs of known function from other organisms; 17% of the ORFs correspond to hypothetical proteins from other species, and 28% of the ORFs are novel genes (Fraser et al. 1998).

A number of factors make the *T. pallidum* genome an excellent model system for the development of functional genomic techniques. These factors include the small size of the genome, the correspondingly small number of ORFs, and the intractability of this organism to standard genetic approaches. Important questions regarding *T. pallidum* biology that a functional genomics approach can address include the identification of antigenic proteins that may aid in diagnostics and vaccine development, and identification of proteins important for attachment and invasion of human tissues.

RESULTS AND DISCUSSION

Construction of the *T. pallidum* Univector Clone Set

The purpose of this study was to develop a complete set of *T. pallidum* genes cloned into a variety of plasmids useful for protein expression and purification in *Escherichia coli*, as well as other functional vectors useful for methods such as phage

[6]Corresponding author.
E-MAIL timothyp@bcm.tmc.edu; FAX (713) 798-7375.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.288103. Article published online before print in June 2003.

display and two-hybrid system protein interaction studies. This goal was accomplished by using a recombination-based approach developed by Liu et al. (1998) to facilitate the cloning of a PCR product into a plasmid, followed by the rapid conversion of the plasmid to a number of different expression systems without the necessity of cloning the PCR product multiple independent times. The method, termed the univector plasmid fusion system (UPS), involves the insertion of a gene into the univector plasmid (in this case, pUNI-D) and subsequent fusion of the gene to a variety of promoters and tag sequences. The system is based upon plasmid fusion using the Cre-*loxP* site–specific recombination system of bacteriophage P1 (Fig. 1; Sternberg et al. 1981). The pUNI plasmid is used for the initial cloning of PCR products, whereas the pHOST plasmid contains the appropriate promoter or tag sequences for creating fusion proteins. The recombinant protein expression construct is made by fusion of the pUNI and pHOST plasmids via Cre-*loxP* site–specific recombination (Fig. 1; Liu et al. 1998).

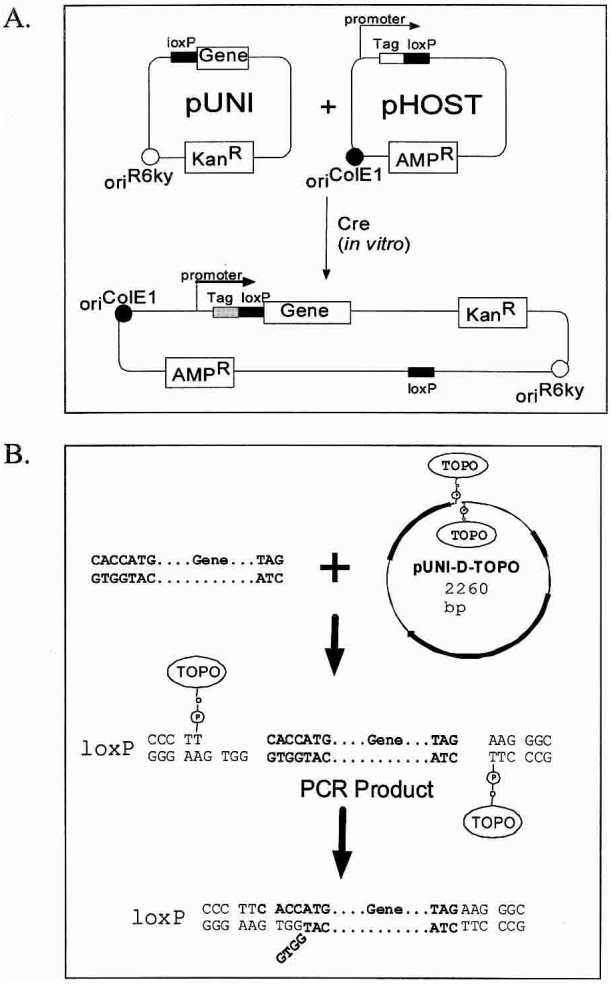

Figure 1 (*A*) Univector plasmid fusion system. Cre-*loxP*–mediated site-specific recombination fuses the pUNI and pHOST plasmids at the *loxP* site. As a result, the gene of interest is placed under the control of the pHOST promoter and fused to any tag sequences present in the pHOST plasmid. Figure adapted from Liu et al. (1998). (*B*) Schematic illustration of directional topoisomerase cloning into the pUNI-D vector. The sequence 5′-CACC is required at the 5′ end of the PCR product for directional topoisomerase-mediated cloning. The oval represents vaccinia virus topoisomerase I. In the example shown, the 5′-CACC sequence is appended immediately 5′ of the ATG start codon of the gene to be inserted.

Insertion of *T. pallidum* ORFs in a univector was accomplished by using a cloning method based on the mechanism of action of the vaccinia virus DNA topoisomerase I (Shuman 1994). In the reaction, the topoisomerase forms a covalent complex with the vector DNA, and PCR products that contain a 5′ hydroxyl tail complementary to the sequence on the covalent adduct can be inserted to create a recombinant molecule (Shuman 1994). The cloning method has recently been improved to allow directional cloning of PCR products (Invitrogen; Fig. 1). The only requirement for PCR primer design for this system is to include the sequence 5′-CACC in the primer sequence at the 5′ end of the PCR product. The 5′-CACC sequence is complementary to a sequence on the 5′ side of the plasmid-topoisomerase I adduct and, as such, controls the orientation of insertion of the PCR product (Fig. 1B).

The annotated *T. pallidum* genome sequence deposited in GenBank includes the sequence position of the start and stop codons of each ORF (Fraser et al. 1998). Frameshifts were identified in 10 ORFs, reducing the GenBank deposit from 1041 to 1031 ORFs. To amplify each of the 1031 ORFs, a computer program was written that used the GenBank annotation data to design primers complementary to the initial and final 25 bp of each ORF. The initial 240 nucleotides of each ORF were also translated and submitted to a signal sequence prediction program (SignalP http://www.cbs.dtu.dk/services/SignalP/). Conserved signal peptidase I or II signal sequences were predicted for 248 ORFs; however, only 183 predictions appeared reasonable based on length of the encoded gene and/or signal sequence. Primers were therefore designed to amplify only the portion of the ORF encoding the mature portion of these 183 proteins. The complete set of 2062 primers necessary to amplify all of the predicted *T. pallidum* genes were designed by using the program, synthesized, and used for PCR amplification. The template DNA for the PCR reactions consisted of *T. pallidum* genomic DNA and a set of large insert bacterial artificial chromosome (BAC) clones encompassing the entire *T. pallidum* genome (Smajs et al. 2002).

In the initial PCR reactions, 959 of the 1031 ORFs were successfully amplified by using the primers designed with 25 bp of complementary nucleotides. Inspection of primer pairs for the ORFs that failed to amplify indicated that below average T_m values for the primers might have been responsible for the lack of amplification of 61 of these ORFs. The region of complementarity in these primers was extended to 30 bp, and subsequent amplification yielded PCR product for 38 of the 61 genes. In addition to extending the length of primers, increasing the annealing/extension parameter of the PCR cycling program aided in acquiring PCR products for an additional 26 ORFs. To date, 1023 of the 1031 predicted genes have been successfully amplified by PCR.

Of the 1023 PCR products, 991 have been inserted into the pUNI-D vector by using the topoisomerase cloning method. Briefly, PCR products were inserted into pUNI-D (Fig. 1), transformed into *E. coli*, and selected on Luria Bertani (LB) agar plates containing kanamycin. Transformants were screened by using a PCR technique that identifies ORFs of the correct size cloned into pUNI-D in the correct orientation.

The *T. pallidum* ORFs that have not been cloned successfully are listed in a table available at www.genome.org as

Table 1. Mutation Sequence Data

	TP ORF	Number of base calls from 5′ end of gene	Result	Number of base calls from 3′ end of gene	Result
	11	651		696	error type 2[a]
	13[b]	592	error type 3[c]		
	74	664	error type 3[c]	325	error type 2[a]
	117	874		678	error type 3[d]
	174	718	error type 5[e]		
	212	230		315	error type 1[a]
	247[b]	970			error type 2[a]
	252	647	error type 6[e]	663	
	260[b]	1237			error type 2[a]
	309[b]	793			error type 2[a]
	324	769		383	error type 7[e]
	350[b]	1291	error type 3[c]		
	389	294		299	error type 1[a]
	415[b]	535	error type 6[e]		
	459	710	error type 4[e]		
	471	181		660	error type 3[c]
	536[b]	330	error type 8[e]		
	602	480	error type 4[e]		
	692	250		306	error type 1[a]
	708[b]	424	error type 3[c]		
	711	505	error type 3[c]		
	747	590	error type 3[a]		
	782[b]	385	error type 3[c]		
	792	270		400	error type 1[a]
	943	280	error type 4[e]	300	
	1031	736		657	error type 2[a]
Table count	26	15,406	13	5682	14
Total count	108	57,756	13	14,431	14
Total number of nucleotides sequenced					72,187
Total number of nucleotides cloned					1,044,638
Total percentage sequenced					6.9%

Error type 1 indicates last nucleotide deleted; error type 2, multiple nucleotides deleted; error type 3, nucleotide substitution; error type 4, single nucleotide deletion inside primer region; error type 5, single nucleotide deletion outside primer region; error type 6, multiple nucleotide deletions inside primer region; error type 7, nucleotide insertion outside primer region; and error type 8, nucleotide insertion inside primer region.
[a]Results in silent mutation.
[b]Entire gene sequenced.
[c]Results in mis-sense mutation.
[d]Results in non-sense mutation.
[e]Results in frameshift mutation.

Supplemental Materials. The average size of the ORFs that were not cloned is 1695 bp, whereas the average size ORF for the entire genome is 1013 bp. Therefore, a partial explanation for the failure to clone these ORFs may be related to the size of the PCR product. However, this is not a complete explanation in that several of the uncloned ORFs are average-sized or smaller. Another possibility is that these *T. pallidum* gene products are toxic when expressed in *E. coli*. Although the pUNI vector is designed to minimize transcription of the cloned insert in *E. coli*, it is possible that low levels of transcription of some genes occur, and these could have toxic effects (Liu et al. 1998).

Characterization of the *T. pallidum* Univector Clone Set

Because each of the ORFs was amplified by PCR prior to insertion into the univector, it is possible that errors exist in some of the cloned ORFs. To determine the frequency of errors among the cloned ORFs, DNA sequence information was obtained from 106 clones (data available in Supplemental Materials). Sequence data was acquired by sequencing into each gene from both the 5′ and 3′ ends with primers located ~100 nucleotides away from the cloned gene. A total of 72,187 nucleotides were sequenced, representing 6.9% of all nucleotides cloned. From this analysis, 27 mutations were identified, which corresponds to one error per 2673 bp sequenced (Table 1).

Of the 27 mutations identified by sequencing, nine are predicted to result in frameshift or nonsense mutations and, therefore, to encode nonfunctional protein products. Five of these nine mutations were single or multiple nucleotide deletions that occurred inside the region complementary to the 5′ PCR primers used to amplify ORFs TP0252, TP0415, TP459, TP0602, and TP0943. This resulted in a frameshift mutation in each of these ORFs. In addition, a frameshift mutation occurred in TP0536 due to a single nucleotide insertion within the region complementary to the 5′ primer used for amplification of this ORF. Two additional frameshift mutations occurred due to single base pair deletions within the coding sequences of TP0174 and TP0324. Finally, a single nucleotide substitution resulted in the generation of a STOP codon within the coding sequence of TP0117. The high frequency of mutations detected inside PCR primer regions (six of nine) indicates that quality control of PCR primer production is an important step that may improve the systematic cloning of any genome.

Of the 18 remaining mutations, six were single nucleotide substitutions that resulted in amino acid substitutions within ORFs TP0013, TP0074, TP0350, TP0471, TP0711, and TP0782. An additional single nucleotide substitution was silent, whereas another occurred in the CACC sequence that is used for cloning but is not within the coding sequence of an ORF. The remaining 10 mutations were deletions at the 3′ end of cloned genes. Three single nucleotide deletions were identified in clones TP0212, TP0692, and TP0792 that resulted in silent mutations because the next nucleotide maintains the stop codon. The seven other deletions involved multiple (two to six) nucleotides; however, all of these deletions impact only the last amino acid and/or the stop codon in the translated protein of ORFs TP0011, TP0074, TP0247, TP0260, TP0309, TP0389, and TP1031.

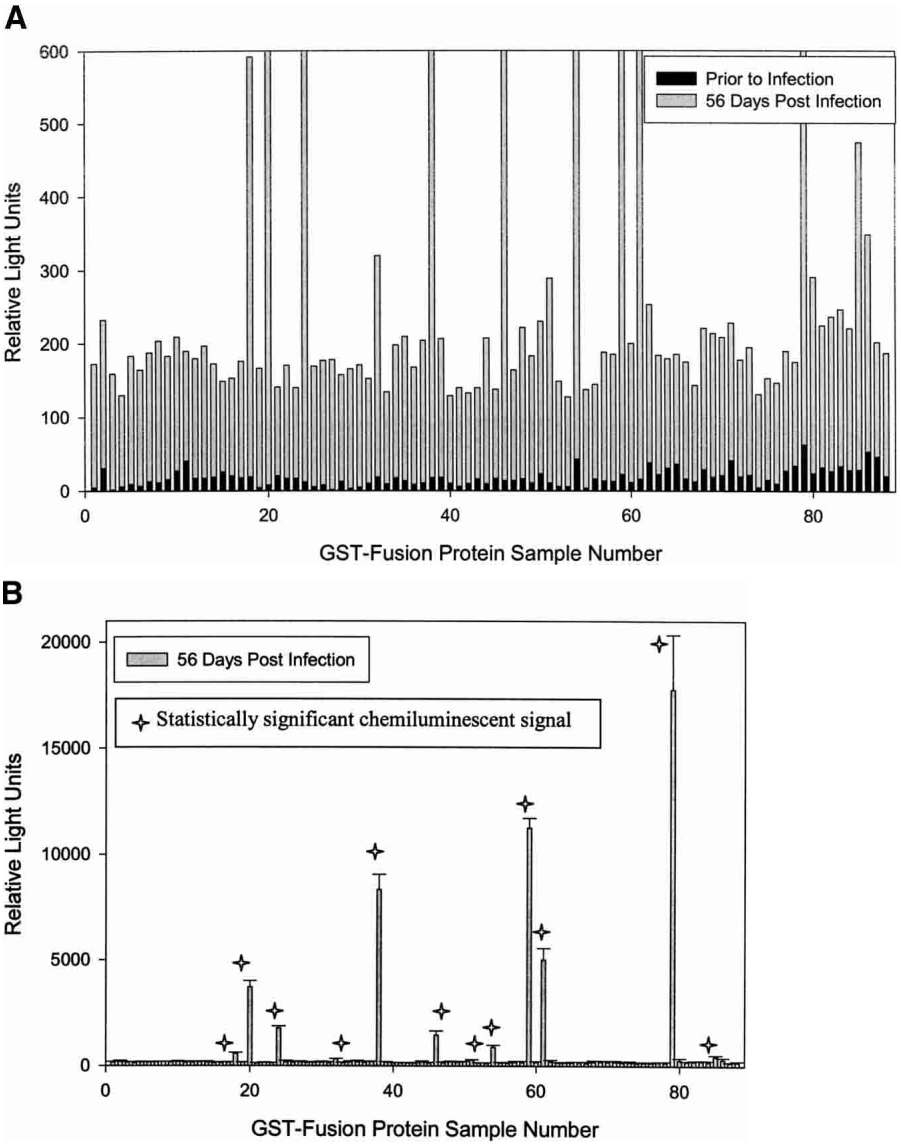

Figure 2 (*A*) Rapid identification of antigenic GST-fusion clones by a chemiluminescent immunoassay technique. Individual GST-fusion *Treponema pallidum* proteins from crude *Escherichia coli* extracts were selectively immobilized on glutathione-coated 96-well plates and were exposed to rabbit serum from rabbits either before or 2 months after intratesticular injection with virulent *T. pallidum*. Detection of binding interactions between rabbit IgG antibody and immobilized *T. pallidum* proteins was carried out with donkey, anti-rabbit IgG antibody conjugated to horseradish peroxidase. A plate reader monitored chemiluminescence, and the mean relative signal is presented for 85 *T. pallidum* fusion proteins, samples 3–87. Sample numbers 1 and 2 were prepared from *E. coli* that does not express a fusion protein. Sample 88 is the *Borrelia burgdorferi* fibronectin binding protein (BBK32) fused to GST expressed from pMM110-pUNI-D. (*B*) The mean relative signal and SD data for rabbit serum collected 2 months after intratesticular injection of *T. pallidum* is presented. A star has been placed above each of the 12 samples with a statistically significant signal.

Conversion of *T. pallidum* Clone Set to glutathione-S-transferase–Fusion Protein Expression Vector

As described above, the univector plasmids can be fused with pHOST plasmids to create functional vectors. In vitro Cre-*loxP* recombination was used to generate a set of protein expression vectors by fusing the *T. pallidum* pUNI-D clone set with a pHOST vector designed for high-level protein expression. The set of *T. pallidum* pUNI-D clones predicted to contain a signal sequence were fused in individual reactions with a pHOST vector, pMM110, containing the bacterial P_{tac} promoter and a glutathione-S-transferase (GST) tag sequence (Fig. 1). Properly fused pMM110-pUNI-D clones were verified by PCR using primers specific for the fusion constructs. Proteins that are fused to GST can be affinity-purified from crude *E. coli* lysates by using glutathione-coated ELISA plates and, subsequently, can be used to capture antigen-specific antibodies (SLehr et Lal. 2001).

Currently, there is no vaccine available for the prevention of syphilis (Weinstock et al. 1998). Although it is known that patients treated during the early stages of syphilis can be reinfected, other studies, such as the demonstration of complete protection against *T. pallidum* infection in rabbits, indicate a vaccine is feasible (Miller 1973). The likelihood of developing a practical immunization regime is also supported by reports of partial protection against intradermal challenge as a result of immunizations with protein 4D (Borenstein et al. 1988), endoflagella (Champion et al. 1990), TmpB (Wicher et al. 1991), TprK (Centurion-Lara et al. 1999), and Tpn92 (Cameron et al. 2000). Therefore, it may be possible to identify a combination of recombinant *T. pallidum* antigens that generate complete protection.

Over the past 25 years, identification and antigenic characterization of *T. pallidum* polypeptides has dominated *T. pallidum* research (Strugnell et al. 1990; Norris 1993). The use of sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and two-dimensional gel electrophoresis coupled with immunoblot technology advanced the discovery of *T. pallidum* antigenic proteins. In addition, improvements in outer membrane protein purification techniques (Blanco et al. 1994, 1999; Radolf et al. 1995) have led to the identification of potential outer membrane proteins (Blanco et al. 1997; Shevchenko et al. 1997) that may aid in vaccine development; however, there is considerable uncertainty regarding the purity of outer membrane preparations (Radolf 1995; Shevchenko et al. 1997). The use of recombinant DNA technology has played a role in advancing the antigenic analysis of *T. pallidum* (Cameron et al. 2000), but systematic cloning and expression studies have not been previously reported.

To identify novel antigens, binding interactions between recombinant *T. pallidum* proteins and IgG antibodies present in rabbits infected with *T. pallidum* were examined. Expression of 85 GST-fusion proteins (a sample set from the proteins with predicted signal sequences) was carried out in *E. coli* cells grown in a 96-well format. The GST-fusion proteins present in *E. coli* cell lysates were incubated in ELISA wells coated with immobilized glutathione (Pierce), and washes were performed to reduce nonspecific binding interactions between *E. coli* proteins and the ELISA wells. The immobilized GST-fusion proteins were then exposed to serum collected from three rabbits both before inoculation (normal rabbit serum [NRS]) and 2 months after intratesticular injection of *T. pallidum* (infected rabbit serum [IRS]). Detection of interactions between rabbit IgG antibody and recombinant fusion proteins was accomplished by chemiluminescence with donkey anti-rabbit IgG antibody conjugated to horseradish peroxidase (HRP).

The IgG binding data obtained by ELISA (Fig. 2A) indicates that rabbit serum collected prior to *T. pallidum* infection contains little or no antibody that specifically binds to *T. pallidum* proteins. Further, the data indicate that significant binding interactions occur between IgG antibodies present in rabbit serum collected 2 months after *T. pallidum* injection and several immobilized *T. pallidum* proteins (Fig. 2B). There is a striking difference between the binding interactions detected between the two serum types. For example, the mean level of binding of antibody infected rabbits to proteins in *E. coli* control lysates that were grown without *T. pallidum* clones (sample numbers 1 and 2) was 10 times greater than that obtained using sera from uninfected rabbits (Fig. 2A). The difference between the two samples may be a result of nonspecific binding of *E. coli* proteins to ELISA wells that in turn capture *E. coli*–specific IgG antibodies. Alternatively, it is possible that intratesticular injections of virulent *T. palldium* act as an adjuvant and generate an immune response against *E. coli* proteins that increases IgG antibody concentrations. Finally, an increase in overall IgG antibody concentration may occur as a result of *T. pallidum* infection, which results in increased nonspecific binding of rabbit IgG antibodies to ELISA wells. In an attempt to eliminate IgG antibodies that recognize *E. coli* proteins and thereby decrease the background binding levels in the assay, the rabbit serum samples were preincubated with soluble *E. coli* proteins (Methods). This subtraction technique, along with the fusion protein purification step, allowed the capture and detection of *T. pallidum*–specific IgG antibodies from serum collected from rabbits 2 months after infection.

A total of 12 GST-fusion protein samples with the highest mean chemiluminescence proteins were identified as statistically relevant *T. pallidum* antigenic proteins based on a *t* test for independent samples with *P* scores of $P < 0.001$ (Fig. 2B; Table 2). Seven of the 12 antigens have been previously identified by various methods. These proteins include TP0100 (thioredoxin-like protein; Radolf et al. 1995), TP0163 (TroA/Tromp1; Hardham et al. 1997), TP0257 (glycerophosphodiester phosphodiesterase [GlpQ]; Radolf et al. 1995), TP0298 (exported protein TpN38; Stamm et al. 1996), TP0319 (membrane protein [tmpC]; Blanco et al. 1988), TP0326 (outer membrane protein Tp92; Cameron et al. 2000), and TP0435 (17-kD lipoprotein; Radolf et al. 1988). Five of these proteins—including TP0163, TP0298, TP0319, TP0326, and TP0435—have been identified previously as surface-exposed proteins, although these classifications may not be accurate (Weinstock et al. 1998). Proteins TP0257 and TP0100 were initially identified in purified outer membrane proteins fractions, but there is also evidence that these protein are localized in the periplasmic space (Shevchenko et al. 1999). Partial protection with recombinant proteins TP0326 and TP0257 has been demonstrated (Cameron et al. 1998, 2000), whereas TP0435 is a known major antigen (Lukehart et al. 1986) but does not provide protection against infection (Centurion-Lara et al. 1997).

In contrast to the above clones, the GST-fusion of lipoprotein 15 kD (Tpp15), a previously identified major treponemal antigen, was not bound significantly by IgG antibody in the IRS (Radolf et al. 1988). Therefore, the GST-Tpp15 represents a false-negative result in this experiment. A possible ex-

Table 2. Identity and Statistical Analysis of the 15 *Treponema pallidum* Genes Yielding the Highest Signal in a Chemiluminescent Immunoassay

Gene	Name	Sample number	Predicted protein size (kD)	Mean[a]	St. dev	*P* value
TP0100	thioredoxin-like protein	20	—[c]	3709.2	294	<0.001
TP0163	ABC transporter (TroA, Tromp1)	32	—[c]	300.7	36	<0.001
TP0257	glycerophosphoryldiester phosphodiesterase (GlpQ)	46	—[c]	1447.7	94	<0.001
TP0298	exported protein TpN38	51	—[c]	277.7	16.5	<0.001
TP0319	membrane lipoprotein (TmpC)	59	—[c]	11259.1	452	<0.001
TP0326	outer membrane protein	61	—[c]	5017.7	550	<0.001
TP0435	lipoprotein, 17 kD (Tpp17)	79	—[c]	17765.6	2583	<0.001
TP0088	conserved hypothetical protein	18	23	572.4	49	<0.001
TP0133	hypothetical protein	24	43	1769.3	91	<0.001
TP0183	hypothetical protein	38	31	8338.6	705	<0.001
TP0307	conserved hypothetical protein	54	37	880.9	105	<0.001
TP0327[b]	hypothetical protein	62	20	214.1	58	0.198
TP0456[b]	hypothetical protein	80	47	265.7	102	0.023
TP0469	conserved hypothetical protein	85	54	444.4	69	<0.001
TP0470[b]	conserved hypothetical protein	86	43	294.7	101	0.0034

[a]Mean value units correspond to relative light units detected.
[b]Gene identified as statistically insignificant.
[c]Predicted protein size was not included for previously characterized genes.

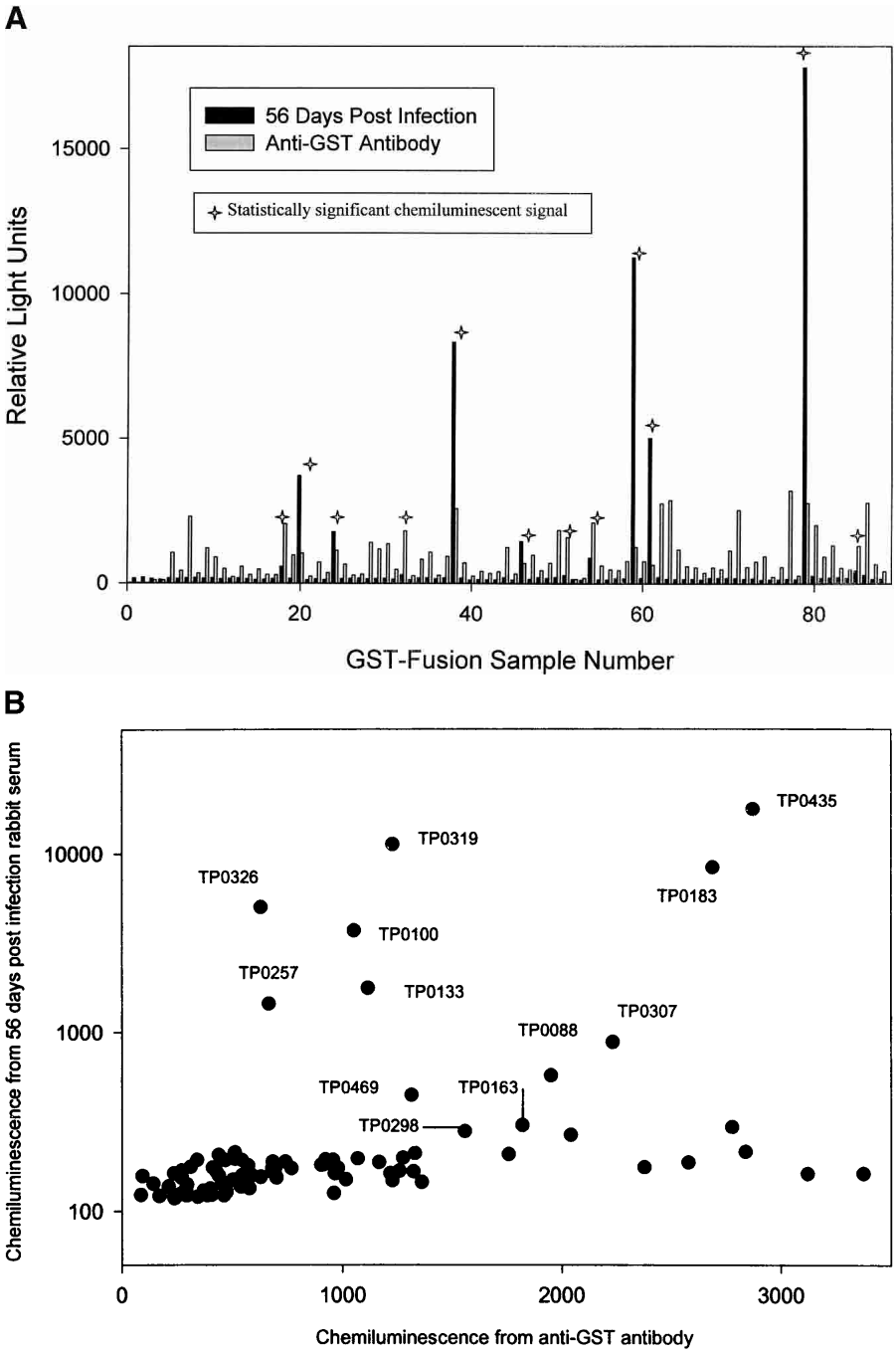

Figure 3 (*A*) Comparison of the chemiluminescent signal generated from exposing immobilized GST-fusion proteins to either anti-GST-HRP or serum from rabbits 56 d after *Treponema pallidum* infection. As described in Fig. 2A, GST-fusion proteins from crude *Escherichia coli* extracts were selectively immobilized on glutathione-coated 96-well plates. To compare relative amounts of immobilized GST-fusion protein, anti-GST-HRP antibody was incubated in the wells instead of rabbit serum. A plate reader monitored chemiluminescence and the mean relative signal is presented next to the chemiluminescent signal measured in the immunoassay technique (Fig. 2B). A star has been placed above each of the 12 samples with a statistically significant signal (Table 2). (*B*) A scatter plot presents the absence of a correlation between GST-fusion protein amount and antigenic properties. The mean relative chemiluminescent signal detected from captured anti-GST-HRP antibody is plotted along the *X*-axis. The mean relative chemiluminescent signal detected from captured IgG antibodies present in rabbit serum 56 d after *T. pallidum* infection is plotted along the *Y*-axis in log scale. Samples with a statistically significant signal have been labeled (Table 2).

planation for this result is that the fusion protein is poorly expressed. However, DNA sequence data on the Tpp15 (ORF TP0171) univector construct indicated that no mutations are present in the cloned gene. In addition, immunoblot experiments using an anti-GST antibody indicated the full-length GST-Tpp15 fusion protein was present in *E. coli* BL21*DE3 cell lysates separated by SDS-PAGE (data not shown). Finally, the immobilized GST-fusion protein was detected by anti-GST antibody in the ELISA wells of glutathione-coated plates (Fig. 3A). Thus, it is unclear why this known antigen did not capture more IgG antibody.

Five previously unreported antigens (TP088, TP0133, TP0183, TP0307, and TP0469) were found to bind significant amounts of IgG in the sera from infected rabbits. Interestingly, none of these proteins have been assigned a biological function. Protein TP0088, a conserved hypothetical protein, shares identity to only one known protein, the *Borrelia burgdorferi* conserved hypothetical protein, BB0662. Residues 55–170 are 24% identical to BB0662, but neither TP0088 nor BB0662 has a predicted biological function.

Protein TP0133, a hypothetical lipoprotein, shares significant identity to four other *T. pallidum* hypothetical lipoproteins, including TP0134, TP0136, and TP0462. Proteins TP0133 and TP0134 appear to be closely related because residues 8–140 of protein TP0133 are 49% identical to residues 6–376 of protein TP0134. Protein TP0133 also appears related to proteins TP0462, sharing 38% identity between residues 12–410 of protein TP133 and residues 1–392 of protein TP0462. Protein TP0462 was also among the 85 proteins assayed in this experiment, and despite the sequence homology with TP0133, protein TP0462 did not bind significant amounts of antibody from the sera from infected rabbits. However, the sequence of clone TP0462 has not been analyzed; therefore, it is possible that a mutation in this construct may reduce protein expression. Lastly, residues 6–151 of protein TP0133 are 51% identical to residues 3–146 of protein TP0136.

The binding interactions de-

tected between immobilized protein TP0183 and rabbit IgG generated the third highest chemiluminescent signal after that of proteins TP0453 and TP0319 (Fig. 2B). It is difficult to make quantitative conclusions about binding interactions because this screen does not normalize for the quantity of immobilized fusion protein for each gene examined, but the observed robustness of the chemiluminescent signal generated by proteins TP0183 and TP0435, the latter of which is known to elicit a strong immune response in rabbits (Lukehart et al. 1986), indicates that TP0183 may also be a major antigenic protein. It is possible that protein TP0183 is expressed at low levels in *T. pallidum*, thereby escaping prior detection on *T. pallidum* immunoblots. The TP0183 protein, however, does not exhibit significant sequence homology with other proteins, and so, no putative biological function can be assigned.

Although the most significant sequence homology of TP0307 is with a *B. burgdorferi* hypothetical protein of unknown function, other sequence alignment data for protein TP0307 indicates a putative function as a kinase, as well as a possible subcellular location on the outer membrane. Residues 33–191 of protein TP0307 are 30% identical to residues 441–593 of the *Streptomyces toyocaensis* eukaryotic-type serine/threonine kinase. In addition, residues 5–201 of protein TP0307 are 22% identical to residues 12–214 of a *Leptospira interrogans* putative outer membrane protein. The fact that TP0307 reacts with the sera and has homology with an outer membrane protein makes this an interesting candidate for immune protection studies.

Of all the previously unreported antigens identified here, protein TP0469 appears to be the most conserved. Of the 989 sequence alignments identified in database searches, 186 possess a low enough *E* value to qualify as significant. Multiple alignments to proteins such as *Methanothermobacter thermautotrophicus* O-linked *N*-acetylglucosamine transferase indicate this as a putative function of protein TP0469. In addition, protein TP0469 has sequence identity to several proteins containing tetratricopeptide repeats. Lastly, residues 233–472 of protein TP0469 are 21% identical to residues 719–949 of *B. burgdorferi* surface-located membrane protein 1 (lmp1).

Further investigations into these putative antigenic proteins are necessary to determine their relative antigenicity and potential protective activity. However, the fact that several well-characterized *T. pallidum* antigens were identified by the same method lends support to the findings and indicates that these antigens should be treated as potential vaccine candidates.

Antigenic characterization experiments were accomplished by exposing *T. pallidum* proteins fused to GST to serum from rabbits 56 d after *T. pallidum* infection. Current protein quantitation techniques are not sensitive enough to accurately quantify the amount of fusion protein immobilized to the glutathione-coated plate. However, it is important to determine whether there is a relationship between the amounts of protein immobilized and the amounts of rabbit IgG antibody captured by the GST-fusion protein. For example, if the proteins that resulted in the strongest binding signal with antibody were also present at much higher levels than other proteins tested, it would indicate that the signal being detected is simply an increase in nonspecific binding due to increased amount of protein. Similarly, if a protein is not expressed, it cannot bind antibody and will yield a false-negative result in the ELISA experiment.

The relative amount of *T. pallidum* fusion proteins immobilized in the ELISA wells was estimated by using an anti-GST antibody to detect the GST tag that is common to all of the fusions. The methods used to immobilize *T. pallidum*–GST fusion proteins on glutathione-coated wells for quantitation were identical to those used in the antigenic characterization experiments. However, instead of exposing immobilized protein to rabbit serum, an anti-GST antibody was used. Nonspecific antibody binding interactions were washed away before the HRP substrate was added to each well in order to detect chemiluminescence generated from captured HRP. The chemiluminescent signal was compared with the signal generated from captured IgG antibody present in serum from rabbits 56 d after *T. pallidum* infection (Fig. 3A). A scatter plot (Fig. 3B) demonstrates that there is no correlation between the amount of GST-fusion protein immobilized and the amount of IgG antibody captured by the immobilized GST-fusion proteins. Many of the *E. coli* extracts tested contained relatively low quantities of the *T. pallidum*-GST fusion proteins, based on low anti-GST reactivities (Fig. 3). This low reactivity could be due to poor transcription and or translation of the fusion product, or to proteolytic degradation of the product. Those fusion proteins that are present at low levels could potentially yield false-negative results. Thus one of the goals in future studies will be to achieve more uniform levels of the fusion products. However, it is notable that two of the gene products, TP0326 and TP0257, were highly reactive with sera from *T. pallidum*–infected rabbits and demonstrated relatively low reactivity with anti-GST (Fig. 3B). Furthermore, several proteins are expressed at high levels but do not exhibit marked reactivity with anti–*T. pallidum* antibodies (Fig. 3B).

Taken together, these results indicate that high reactivity with anti-*T. pallidum* antibodies is a reliable indicator of high antigenicity in this system, whereas low reactivity could be due to either low antigenicity or low fusion protein levels. This is a common limitation in most recombinant antigen screening techniques (e.g., colony immunoblot assays), and this report represents the first attempt to approach this problem quantitatively.

The above experiments indicate that the *T. pallidum* clone set will be a powerful reagent for functional genomics studies of *T. pallidum* biology and pathogenesis. The availability of the clone set permits the systematic evaluation of antigenicity of *T. pallidum* proteins. In addition, the ELISA experiments described above can be used to detect other types of protein–protein or protein ligand interactions. The clone set can also be used to examine binding interactions in vivo by using two-hybrid approaches to generate large-scale protein–protein interaction maps. Such tools are important for an organism that cannot be cultured in vitro and is therefore recalcitrant to most molecular genetic approaches for the study of gene function and pathogenesis.

METHODS

Bacterial Strains and Plasmids

Chemically competent *E. coli* PIR1 (Liu et al. 1998) were provided by Invitrogen, Inc., and were used to propagate the univector plasmid DNA containing the cloned *T. pallidum* ORFs. *E. coli* BL21* DE3 (Invitrogen, Inc.) was used to propagate univector-host fusion plasmid DNA after Cre-*lox*P recombinase reactions and to express all of the GST-fusion proteins. The vector pUNI-D/V5-His-TOPO (referred to as pUNI-D in Results and Discussion) was provided by Invitrogen, Inc., and was used for the insertion of all *T. pallidum* PCR products. The

vector pMM110 was created by inserting a *lox*P site into the *Xho*I-*Not*I site of pGex-4t-2 (Amersham Biosciences) and was used as the host vector for expression of all GST-tagged *T. pallidum* proteins.

Media

Immediately after heat-shock transformation, *E. coli* PIR1 was grown in media containing 5 g Bacto-Yeast, 20 g Bacto-Tryptone, 5M NaCl (2 mL), 0.19 g KCl, 10 mL, 1 M $MgCl_2$, 10 mL 1 M $MgSO_4$, and 3.6 g glucose per liter (SOC). *E. coli* PIR1 cells containing pUNI-D with or without *T. pallidum* DNA inserts were grown overnight on LB agar plates containing 25 µg/mL kanamycin; isolated colonies were expanded by incubation in 3 mL of the corresponding liquid medium. Immediately after heat-shock transformation, *E. coli* BL21*DE3 cells were grown in media containing 16 g Bacto-Tryptone, 10 g Bacto-Yeast, and 5 g NaCl per liter (2YT). *E. coli* BL21*DE3 cells containing the pMM110-pUNI-D plasmid fusion construct were selected on LB agar plates containing 25 µg/mL kanamycin and 100 µg/mL ampicillin. Clones were selected and cultured in the liquid media form of the same medium prior to storage.

PCR Reactions

Pfu DNA polymerase (Stratagene) was used to amplify *T. pallidum* ORFs for cloning into pUNI-D. AdvanTaq DNA polymerase (Clontech) was used to PCR amplify inserts for screening. Research Genetics provided all of the primers used to amplify the ORFs. The template DNA for the PCR reactions consisted of a set of large insert BAC clones encompassing the entire *T. pallidum* genome (Smajs et al. 2002). For the majority of cloning and screening PCR reactions, the thermocycle program was 1 min at 96°C and 4 min at 64°C for 30 cycles with a final 10-min 64°C step.

Topoisomerase Cloning

Topoisomerase cloning reactions were carried out with 4 µL of PCR product, mixed with 1 µL of salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) and 1 µL of topoisomerase-adapted pUNI-D for 5 min at room temperature. This entire reaction was added to chemical-competent *E. coli* PIR1 cells prepared and supplied by Invitrogen, Inc. Cells and DNA were incubated for 30 min on ice, heat-shocked for 30 s at 42°C, and returned for 5 min on ice. A total of 250 µL of SOC media was added to the cells, followed by a 1-h recovery period at 37°C. A total of 200 µL of transformed cells was spread onto LB-kanamycin agar plates, which were then incubated overnight at 37°C.

PCR Screen for Recombinant Clones

Kanamycin-resistant *E. coli* PIR 1 transformants were screened for the presence of recombinant *T. pallidum* genes in pUNI-D by using colony PCR. Autoclaved pipette tips were used to pick three to 12 colonies for each *T. pallidum* ORF to be tested. The tips were used to inoculate individual colonies into PCR tubes containing a PCR mixture that included a pUni-forward (Invitrogen) primer within the vector and a 3′ (bottom strand) primer specific to the *T. pallidum* ORF being screened. Visualization of a PCR product corresponding to the size of the ORF plus 120 bp indicated the insertion of the ORF in the correct orientation.

Cre-*loxP* Recombinase Reactions

Cre recombinase (New England Biolabs) was used to fuse the pUNI-D clones containing *T. pallidum* ORFs with the pMM110 plasmid. Host vector DNA was purified by using a cesium chloride gradient (Sambrook et al. 1989) and diluted to 50 ng/µL in water. Individual pUNI-D clones containing *T. pallidum* ORFs were used in Cre reactions for the construction of the GST fusion clone set; 150 ng of pMM110 and 3 µL of univector DNA were mixed with 1 µL of Cre and 1.5 µL of 10× Cre Buffer (New England Biolabs, Inc.). Cre reactions were incubated for 1 h at 37°C, heat-killed for 10 min at 70°C, and chilled to 4°C. Chemically competent *E. coli* BL21*DE3 cells were added to the Cre reaction for transformation and selection. Cells and DNA were incubated for 30 min on ice, heat-shocked for 90 sec at 42°C, and returned for 5 min at ice; 200 µL of 2YT media was added to the cells, followed by a 1-h recovery period at 37°C; and then 200 µL was then spread on LB-kanamycin-ampicillin plates that were incubated overnight at 37°C. Colony PCR was used to identify *E. coli* BL21*DE3 cells containing Cre-*lox*P recombined pMM110-pUNI-D fusion plasmids. Autoclaved pipette tips were used to pick one to 12 colonies for each cloned *T. pallidum* ORF. The tips were used to inoculate PCR tubes containing a PCR mixture that included a top-strand primer specific to the terminus of GST in pMM110 pGextop (5′-GGGCTGGCAAGCCAC GTTTGGTG-3′) and pUni-reverse (Invitrogen). Visualization of a PCR product corresponding to the size of the ORF plus 250 bp indicated the correct fusion of a *T. pallidum* ORF-pUNI-D plasmid with pMM110.

DNA Sequence Data Collection

Sequence data was acquired from PCR products generated from PCR colony screening of the pMM110-pUNI-D constructs. The PCR primers pGextop and pUni-reverse (Invitrogen) were used in sequencing reactions in order to sequence the 5′ and 3′ end of each ORF. An Applied Biosystems 3100 Genetic Analyzer was used to analyze sequencing reactions. BLAST searches and some manual inspections were used to analyze sequence data. Twenty-three of the clones in which a mutation was found were resequenced in order to confirm the presence of the mutation.

Rabbit Serum Preparation

Rabbit serum was collected from adult male New Zealand White rabbits obtained from Myrtle's Rabbitry (Thompson Station, TN). Rabbits were housed individually at 18°C to 20°C and given antibiotic-free food and water. Rabbits were tested for evidence of *Treponema paraluiscuniculi* infection by using a Macro-Vue RPR card test (Becton-Dickinson). Only seronegative animals were included in this study. Three rabbits were each infected with 4×10^8 *T. pallidum* Nichols strain organisms by intratesticular injection. Sera was collected 48 h before infection and 56 d after infection and were stored at −80°C until use.

Absorption of Anti–*E. coli* Protein Antibodies

E. coli BL21*DE3 was grown overnight at 37°C in 150 mL of LB liquid media. Cells were pelleted and freeze-thawed two times at −80°C. The cell pellet was resuspended in 5 mL TE (pH 8.0) and subjected to cell lysis with a French press. The lysate was centrifuged for 10 min at 13K and 4°C. A mixture of 45 µL rabbit serum, 405 µL PBS (pH 7.4), and 250 µL BL21*DE3 cell lysate was rocked for 5 h at room temperature. Absorbed serum was stored at −20°C until use.

ELISA Protocol

E. coli BL21*DE3 hosting the pMM110-pUNI-D constructs containing *T. pallidum* ORFs were inoculated into 1 mL LB media containing 25 µg/mL kanamycin, 100 µg/mL ampicillin, and 2% glucose in a 96-well format. Cultures were grown overnight with shaking at 37°C. The next day, a 100 µL aliquot of the overnight culture was added to 1.5 mL 2YT containing 25 µg/mL kanamycin and 100 µg/mL ampicillin. Cultures were grown for 4 h at 37°C on a shaker before IPTG was added to a final concentration of 0.5 mM. A 5-h induction period continued on a shaker at 37°C. Cells were pelleted and stored at −80°C. Each cell pellet was freeze-thawed three

times at −80°C before 0.3 mL B-PER (Pierce bacterial protein extraction reagent) containing 0.375 mg/mL lyzozyme and 420 ng/mL DNase I was used to lyse the pellet. The resuspended pellets were shaken vigorously for 10 min at room temperature while Reacti-Bind glutathione-coated white 96-well pPlates (Pierce) were washed two times with 0.21 mL each of PBS (pH 7.4) and 0.05% Tween 20 (wash buffer); 0.19 mL of cell lysates were then added to the glutathione-coated plates and incubated at room temperature for 2 h. Plates were washed three times with 0.21 mL of wash buffer and then blocked in 0.21 mL PBS (pH 7.4) and 5% dry milk for 1 h at room temperature. Prepared absorbed serum (0.75 mL) was diluted into 20 mL PBS (pH 7.4; final serum dilution is 1:444), and 0.19 mL of serum solution was added to each well for 3 h at room temperature. Plate wells were washed three times with 0.21 mL of wash buffer. A 1:2500 dilution of anti-rabbit Ig, HRP-linked F(ab′)2 fragment from donkey (Amersham Pharmacia) in wash buffer was added to each plate well for 1 h at room temperature. Each plate well was washed five times before 0.19 mL of SuperSignal ELISA Pico Chemiluminescent Substrate (PIERCE) was added. Light emission from each plate well was monitored at a 5- and 15-min interval with a Genios Plate Reader for 200 ms. Each experiment was repeated three times. The resulting six measurements were used to generate a mean and SD for each sample. The same experimental procedures were used to measure the amount of immobilized GST-fusion proteins, except a 1:50,000 dilution of anti-GST-HRP (Pierce) was diluted in PBS (pH 7.4) and incubated in the ELISA wells for 1 h at room temperature before washing.

ACKNOWLEDGMENTS

We thank Invitrogen, Inc., for providing the pUNI-D vector adapted with topoisomerase and Research Genetics for providing the PCR primer set. In addition, we thank Steve Elledge for helpful advice. M.M. is supported in part through the Keck Center for Computational and Structural Biology, NLM Training Grant T15 LM07093. This work was supported by NIH grants AI45842 to T.P. and AI49552 to S.N.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Blanco, D.R., Champion, C.I., Miller, J.N., and Lovett, M.A. 1988. Antigenic and structural characterization of *Treponema pallidum* (Nichols strain) endoflagella. *Infect. Immun.* **56:** 168–175.

Blanco, D.R., Reimann, K., Skare, J., Champion, C.I., Foley, D., Exner, M.M., Hancock, R.E., Miller, J.N., and Lovett, M.A. 1994. Isolation of the outer membranes from *Treponema pallidum* and *Treponema vincentii*. *J. Bacteriol*. **176:** 6088–6099.

Blanco, D.R., Miller, J.N., and Lovett, M.A. 1997. Surface antigens of the syphilis spirochete and their potential as virulence determinants. *Emerg. Infect. Dis*. **3:** 11–20.

Blanco, D.R., Champion, C.I., Lewinski, M.A., Shang, E.S., Simkins, S.G., Miller, J.N., and Lovett, M.A. 1999. Immunization with *Treponema pallidum* outer membrane vesicles induces high-titer complement-dependent treponemicidal activity and aggregation of *T. pallidum* rare outer membrane proteins (TROMPs). *J. Immunol*. **163:** 2741–2746.

Borenstein, L.A., Radolf, J.D., Fehniger, T.E., Blanco, D.R., Miller, J.N., and Lovett, M.A. 1988. Immunization of rabbits with recombinant *Treponema pallidum* surface antigen 4D alters the course of experimental syphilis. *J. Immunol*. **140:** 2415–2421.

Cameron, C.E., Castro, C., Lukehart, S.A., and Van Voorhis, W.C. 1998. Function and protective capacity of *Treponema pallidum* subsp. *pallidum* glycerophosphodiester phosphodiesterase. *Infect. Immun*. **66:** 5763–5770.

Cameron, C.E., Lukehart, S.A., Castro, C., Molini, B., Godornes, C., and Van Voorhis, W.C. 2000. Opsonic potential, protective capacity, and sequence conservation of the *Treponema pallidum* subspecies *pallidum* Tp92. *J. Infect. Dis*. **181:** 1401–1413.

Centurion-Lara, A., Arroll, T., Castillo, R., Shaffer, J.M., Castro, C., Van Voorhis, W.C., and Lukehart, S.A. 1997. Conservation of the 15-kilodalton lipoprotein among *Treponema pallidum* subspecies and strains and other pathogenic treponemes: Genetic and antigenic analyses. *Infect. Immun.* **65:** 1440–1444.

Centurion-Lara, A., Castro, C., Barrett, L., Cameron, C., Mostowfi, M., Van Voorhis, W.C., and Lukehart, S.A. 1999. *Treponema pallidum* major sheath protein homologue Tpr K is a target of opsonic antibody and the protective immune response. *J. Exp. Med*. **189:** 647–656.

Champion, C.I., Miller, J.N., Borenstein, L.A., Lovett, M.A., and Blanco, D.R. 1990. Immunization with *Treponema pallidum* endoflagella alters the course of experimental rabbit syphilis. *Infect. Immun*. **58:** 3158–3161.

Fraser, C.M., Norris, S.J., Weinstock, G.M., White, O., Sutton, G.G., Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A., et al. 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281:** 375–388.

Hardham, J.M., Stamm, L.V., Porcella, S.F., Frye, J.G., Barnes, N.Y., Howell, J.K., Mueller, S.L., Radolf, J.D., Weinstock, G.M., and Norris, S.J. 1997. Identification and transcriptional analysis of a *Treponema pallidum* operon encoding a putative ABC transport system, an iron-activated repressor protein homolog, and a glycolytic pathway enzyme homolog. *Gene* **197:** 47–64.

Liu, Q., Li, M.Z., Leibham, D., Cortez, D., and Elledge, S.J. 1998. The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. *Curr. Biol*. **8:** 1300–1309.

Lukehart, S.A., Baker-Zander, S.A., and Sell, S. 1986. Characterization of the humoral immune response of the rabbit to antigens of *Treponema pallidum* after experimental infection and therapy. *Sex. Transm. Dis*. **13:** 9–15.

Miller, J.N. 1973. Immunity in experimental syphilis, VI: Successful vaccination of rabbits with *Treponema pallidum*, Nichols strain, attenuated by irradiation. *J. Immunol*. **110:** 1206–1215.

Norris, S.J. 1993. Polypeptides of *Treponema pallidum*: Progress toward understanding their structural, functional, and immunologic roles. Treponema Pallidum Polypeptide Research Group. *Microbiol. Rev*. **57:** 750–779.

Norris, S.J., Cox, D.L., and Weinstock, G.M. 2001. Biology of *Treponema pallidum*: Correlation of functional activities with genome sequence data. *J. Mol. Microbiol. Biotechnol*. **3:** 37–62.

Radolf, J.D. 1995. *Treponema pallidum* and the quest for outer membrane proteins. *Mol. Microbiol*. **16:** 1067–1073.

Radolf, J.D., Chamberlain, N.R., Clausell, A., and Norgard, M.V. 1988. Identification and localization of integral membrane proteins of virulent *Treponema pallidum* subsp. *pallidum* by phase partitioning with the nonionic detergent triton X-114. *Infect. Immun.* **56:** 490–498.

Radolf, J.D., Robinson, E.J., Bourell, K.W., Akins, D.R., Porcella, S.F., Weigel, L.M., Jones, J.D., and Norgard, M.V. 1995. Characterization of outer membranes isolated from *Treponema pallidum*, the syphilis spirochete. *Infect. Immun*. **63:** 4244–4252.

Sambrook, J., Maniatis, T., and Fritsch, E.F. 1989. *Molecular cloning : A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sehr, P., Zumbach, K., and Pawlita, M. 2001. A generic capture ELISA for recombinant proteins fused to glutathione S-transferase: Validation for HPV serology. *J. Immunol. Meth*. **253:** 153–162.

Shevchenko, D.V., Akins, D.R., Robinson, E.J., Li, M., Shevchenko, O.V., and Radolf, J.D. 1997. Identification of homologs for thioredoxin, peptidyl prolyl *cis-trans* isomerase, and glycerophosphodiester phosphodiesterase in outer membrane fractions from *Treponema pallidum*, the syphilis spirochete. *Infect. Immun.* **65:** 4179–4189.

Shevchenko, D.V., Sellati, T.J., Cox, D.L., Shevchenko, O.V., Robinson, E.J., and Radolf, J.D. 1999. Membrane topology and cellular location of the *Treponema pallidum* glycerophosphodiester phosphodiesterase (GlpQ) ortholog. *Infect. Immun.* **67:** 2266–2276.

Shuman, S. 1994. Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J. Biol. Chem*. **269:** 32678–32684.

Smajs, D., McKevitt, M., Wang, L., Howell, J.K., Norris, S.J., Palzkill, T., and Weinstock, G.M. 2002. BAC library of *T. pallidum* DNA in *E. coli*. *Genome Res*. **12:** 515–522.

Stamm, L.V., Hardham, J.M., and Frye, J.G. 1996. Expression and sequence analysis of a *Treponema pallidum* gene, tpn38(b), encoding an exported protein with homology to *T. pallidum* and *Borrelia burgdorferi* proteins. *FEMS Microbiol. Lett*. **135:** 57–63.

Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M., and Hoess, R. 1981. Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harb. Symp. Quant. Biol.* **45:** 297–309.

Strugnell, R., Cockayne, A., and Penn, C.W. 1990. Molecular and antigenic analysis of treponemes. *Crit. Rev. Microbiol.* **17:** 231–250.

Weinstock, G.M., Hardham, J.M., McLeod, M.P., Sodergren, E.J., and Norris, S.J. 1998. The genome of *Treponema pallidum*: New light on the agent of syphilis. *FEMS Microbiol. Rev.* **22:** 323–332.

Wicher, K., Schouls, L.M., Wicher, V., Van Embden, J.D., and Nakeeb, S.S. 1991. Immunization of guinea pigs with recombinant TmpB antigen induces protection against challenge infection with *Treponema pallidum* Nichols. *Infect. Immun.* **59:** 4343–4348.

WEB SITE REFERENCES

http://www.cbs.dtu.dk/services/SignalP/; SignalP program.

Received March 19, 2002; accepted in revised form April 22, 2003.