INFECTION AND IMMUNITY, Dec. 1998, p. 5620–5629 Vol. 66, No. 12
0019-9567/98/$04.00+0

Large-Scale Identification of Virulence Genes from *Streptococcus pneumoniae*

ALESSANDRA POLISSI,[1] ANDREA PONTIGGIA,[1] GEORG FEGER,[2] MARIO ALTIERI,[1] HARALD MOTTL,[1] LIVIA FERRARI,[1] AND DANIEL SIMON[1]*

Department of Microbiology, Medicine Research Centre, Glaxo Wellcome S.p.A., 37100 Verona, Italy,[1] *and Geneva Biomedical Research Institute, Glaxo Wellcome, Geneva, Switzerland*[2]

Received 6 May 1998/Returned for modification 22 June 1998/Accepted 18 August 1998

Streptococcus pneumoniae* is the major cause of bacterial pneumonia, and it is also responsible for otitis media and meningitis in children. Apart from the capsule, the virulence factors of this pathogen are not completely understood. Recent technical advances in the field of bacterial pathogenesis (in vivo expression technology and signature-tagged mutagenesis [STM]) have allowed a large-scale identification of virulence genes. We have adapted to *S. pneumoniae* the STM technique, originally used for the discovery of *Salmonella* genes involved in pathogenicity. A library of pneumococcal chromosomal fragments (400 to 600 bp) was constructed in a suicide plasmid vector carrying unique DNA sequence tags and a chloramphenicol resistance marker. The recent clinical isolate G54 was transformed with this library. Chloramphenicol-resistant mutants were obtained by homologous recombination, resulting in genes inactivated by insertion of the suicide vector carrying a unique tag. In a mouse pneumonia model, 1.250 candidate clones were screened; 200 of these were not recovered from the lungs were therefore considered virulence-attenuated mutants. The regions flanking the chloramphenicol gene of the attenuated mutants were amplified by inverse PCR and sequenced. The sequence analysis showed that the 200 mutants had insertions in 126 different genes that could be grouped in six classes: (i) known pneumococcal virulence genes; (ii) genes involved in metabolic pathways; (iii) genes encoding proteases; (iv) genes coding for ATP binding cassette transporters; (v) genes encoding proteins involved in DNA recombination/repair; and (vi) DNA sequences that showed similarity to hypothetical genes with unknown function. To evaluate the virulence attenuation for each mutant, all 126 clones were individually analyzed in a mouse septicemia model. Not all mutants selected in the pneumonia model were confirmed in septicemia, thus indicating the existence of virulence factors specific for pneumonia.

Streptococcus pneumoniae is the major cause of community-acquired bacterial pneumonia, and it is also responsible for otitis media and meningitis (2). Capsular polysaccharides were the first virulence factors to be identified. The capsule is thought to protect the bacteria from the host immune system by preventing phagocytosis (17). Purified capsular extracts do not have an inflammatory or toxic effect (31, 32). Among the proteins considered to be virulence factors (17, 45) are pneumolysin (3, 7, 12), autolysin (4, 12, 56), hyaluronidase (5), pneumococcal surface protein A (PspA) (8), PsaA (6), neuraminidase (10), immunoglobulin A1 (IgA1) protease (46, 59), and pyruvate oxidase (55), although for some of them a role in virulence has not been demonstrated.

Recent advances in the field of bacterial pathogenesis have allowed a large-scale identification of new virulence genes in different bacterial species. The methods developed are based on the concept that specific gene products are required for each stage of an infection process and that their expression is often regulated by the different environmental conditions encountered in the host. Mahan *et al.* (38) developed a system called IVET (in vivo expression technology) aimed at identifying bacterial genes that were preferentially expressed in the host during infection and were poorly transcribed under laboratory conditions. IVET was originally developed for use with *Salmonella typhimurium* (38) and then applied to *Vibrio cholerae* (11) and *Pseudomonas aeruginosa* (58). Hensel *et al.* (28) expanded the concept of tagging originally developed by Walsh and Cepko (57) to monitor the fate of clonally related neocortical cells during brain development, developing a strategy to identify virulence genes by negative selection. This system, called STM (signature-tagged mutagenesis), exploits a pool of transposons in which each transposon is tagged with a unique DNA sequence so that the resulting insertion mutants are marked with a different DNA sequence. This permits the identification of bacteria recovered from hosts infected with a mixed population of mutants.

Tagged insertion mutants are combined into pools, which are used to infect the animals. At a defined time point, bacteria are recovered from the animals. Tag sequences are amplified from each pool by using a radioactive label before and after the infection. These two labeled tag probes are hybridized to filters containing spotted genomic DNA from all mutants of the corresponding pool. Mutants whose tags are positive for hybridization with the probe from the original pool and negative with the one from the recovered bacteria are considered to be virulence attenuated. This system was originally used to identify genes involved in virulence in *S. typhimurium* (28) and recently applied to *Staphylococcus aureus* and *V. cholerae* (15, 42).

S. pneumoniae has been studied for many years, yet its virulence mechanisms are not fully understood (17). Therefore, we have modified the original STM methodology to discover novel virulence factors in *S. pneumoniae*. Initially, we attempted to use for this purpose the encapsulated type 3 strain GP119, obtained by transformation from the nonencapsulated avirulent laboratory strain Rx1. Surprisingly, when GP119 was tested for virulence in a mouse septicemia model, it was found to be not virulent (unpublished observations); this result indi-

* Corresponding author. Mailing address: Glaxo Wellcome S.p.A., Microbiology Department, via A. Fleming 4, 37135 Verona, Italy. Phone: (39) 45 9218554. Fax: (39) 45 9218196. E-mail: ds20886@glaxowellcome.co.uk.

cates that in factors in addition to the capsule are required for virulence. We then selected the encapsulated type 19F strain G54, a recent clinical isolate. This strain was chosen because of its high transformation rate (47) induced by using the 17-residue competence-stimulating peptide (26) and its virulence in a mouse pneumonia model, as assessed in our laboratory. The following two major changes were introduced in the original STM protocol (28).

(i) A library of tagged mutants was constructed by insertion-duplication mutagenesis using short random genomic DNA fragments inserted in a suicide plasmid vector bearing the molecular tag. The original transposon mutagenesis technique (28) was not applied to *S. pneumoniae*, since we did not expect transposons to insert randomly as shown for *Streptococcus mutans* (24) and *Lactococcus lactis* (49).

(ii) While in the original method (28) the filters corresponding to each pool had spots of genomic DNA from each mutant obtained by transferring bacterial colonies to the filters (colony hybridization), we used filters containing amplified tags from each mutant. This modification was necessary since in our hands colony hybridization generated a high background giving rise to false positives.

Using the modified STM technique, we identified 126 putative virulence genes from *S. pneumoniae*. Some of them corresponded to previously described pneumococcal virulence factors, while others showed homology to virulence genes found in other bacteria. In addition, we have identified genes not previously known to be involved in virulence.

MATERIALS AND METHODS

Media. *Escherichia coli* strains were grown in Luria broth (LB; Difco). *S. pneumoniae* strains were grown in Todd-Hewitt broth (Difco), and tryptic soy broth (TSB; GIBCO) or agar plates. When needed, these media were supplemented with ampicillin (100 μg/ml), chloramphenicol (8 μg/ml), or 5% defibrinated sheep blood (Unipath; Oxoid).

Animals. Male BALB/c mice weighing 16 to 18 g (Charles River, Como, Italy) were used.

Competence factor 1. The 17-amino-acid peptide used to stimulate competence was synthesized by Neosystem Laboratoire (Strasbourg, France) according to the published sequence (26).

Construction of the tagged library. A pool of single-stranded 89-bp DNA molecules containing a central stretch of 40 random bp flanked by two constant sequences was generated by oligonucleotide synthesis. The constant regions contained *Hin*dIII and *Kpn*I sites (Fig. 1). The sequence of the random 89-mer oligonucleotides used was 5′-CTAGGTACCTACAACCTCAAGCTTN$_{40}$AAG CTTGGTTAGAATGGGTACCATG-3′. The constant regions allowed the sequence tags to be amplified by PCR using the upper primer (5′-TACCTACAA CCTCAAGCT-3′) and the lower primer (5′-TACCCATTCTAACCAAGC-3′). The double-stranded library of tags was generated by extending the 18-mer primer P1 (5′-CATGGTACCCATTCTAAC-3′) annealed to the random 89-mer template. Ten micrograms of P1 primer was mixed with 3 μg of 89-mer random template in 14 μl of annealing buffer (T7 sequencing kit; Pharmacia) and annealed as indicated by the manufacturer. T7 DNA polymerase (8 U) and 3 μl of 5 mM deoxynucleoside triphosphates (dNTPs) were added and incubated for 30 min at 37°C. The final population of double-stranded DNA molecules was digested with *Kpn*I. The corresponding DNA band was recovered from a 7% polyacrylamide gel, subcloned into the *Kpn*I site of pR326 vector (3 μg) (16) (Fig. 1), and used to transform *E. coli* DH5α by electroporation with a Bio-Rad apparatus according to the manufacturer's instructions. Ten electrotransformation preparations, each containing 100 μl of cells and approximately 300 ng of vector DNA, were pooled, and a small aliquot was plated to assess transformation efficiency. The remainder was added to 500 ml of LB with ampicillin and grown for 8 h with shaking. Tagged DNA plasmids (tagged library) were then extracted and stored at −20°C.

Tagged genomic library construction. *S. pneumoniae* chromosomal DNA was isolated from G54 (47) as reported by Dillard and Yother (19).

Chromosomal DNA (40 μg) was sheared with a French cell press (SLM Aminco) as indicated by the manufacturer. The DNA ends were repaired by using 25 U of Klenow polymerase, 15 U of T4 DNA polymerase, and 50 μM dNTPs. The blunt-ended DNA was then size fractionated by centrifugation in a 10 to 40% sucrose gradient (26,000 rpm for 18 h in an SW28 rotor (Beckman) (2a). The size of each fragment from the fractions was determined by agarose gel electrophoresis. Fractions from 400 to 600 bp were pooled, dialyzed, and ethanol precipitated. The fractions were resuspended in 200 μl of 1× ligase buffer and ligated for 16 h at 4°C with 20 μg of *Bam*HI phosphorylated linkers (New England Biolabs) and 1,000 U of T4 DNA ligase. This mixture was then digested with *Bam*HI and ligated to the *Bam*HI-digested tagged plasmid library (Fig. 1).

The ligation mixture was used directly to transform *S. pneumoniae* G54 as reported by Pozzi et al. (47). The bacteria were plated on tryptic soy agar plates containing 8 μg of chloramphenicol per ml and incubated for 16 h at 37°C and 5% CO_2. The multilayer plating technique described by Pozzi *et al.* (47) was used to ensure that each chloramphenicol-resistant (Cm^r) colony had arisen from one independent recombinational event.

The colonies were picked, inoculated in TSB containing 8 μg of chloramphenicol per ml, grown to an OD_{600} (optical density at 600 nm) of 0.6, and frozen individually with glycerol (10%, final concentration). Preparation of the pools was performed as follows: 50 mutants were individually inoculated in 3 ml of TSB containing 8 μg of chloramphenicol per ml, and the cells were grown up to OD_{600} of 0.6; 300 μl of each culture was pooled (*in vitro* pool) and frozen with 10% (final concentration) glycerol.

Three hundred microliters of the frozen *in vitro* pool was inoculated into 3 ml of TSB containing 8 μg of chloramphenicol per ml, and the cells were grown up to OD_{600} of 0.9, corresponding to a concentration of 10^9 cells/ml. An aliquot was removed for DNA extraction, while 50 μl of the remaining culture was used to inoculate the mice as described below.

Filter preparation. The individual DNA tag from each mutant was amplified by PCR as follows. The mutants were grown in 3 ml of TSB containing 8 μg of chloramphenicol per ml to an OD_{600} of 0.6; 1 ml of this culture was centrifuged, the bacterial pellet was resuspended in 100 μl of H_2O and then heated at 96°C for 5 min; 4 μl of lysate was used as the template. The PCR mixture (50 μl) contained the primers (1 μM each), dATP, dTTP, dGTP, and dCTP (0.2 mM each), 4 μl of DNA solution, *Taq* DNA polymerase (1 U), and 5 μl of *Taq* polymerase 10× buffer (Perkin-Elmer Cetus). Twenty cycles of denaturation (60 s at 94°C), annealing (60 s at 50°C), and polymerization (10 s at 72°C) were performed on a Perkin-Elmer Cetus thermocycler. The amplification products were checked on a gel, and the same amount of DNA tags was spotted onto duplicate filters.

Pneumonia and septicemia animal models. The animal model developed by Canvin *et al.* (12) was adapted for this work. The experiments were carried out in accordance with prevailing legislation (Decreto legislativo italiano no. 116), company policy on the care of use of animals, and related codes of practice.

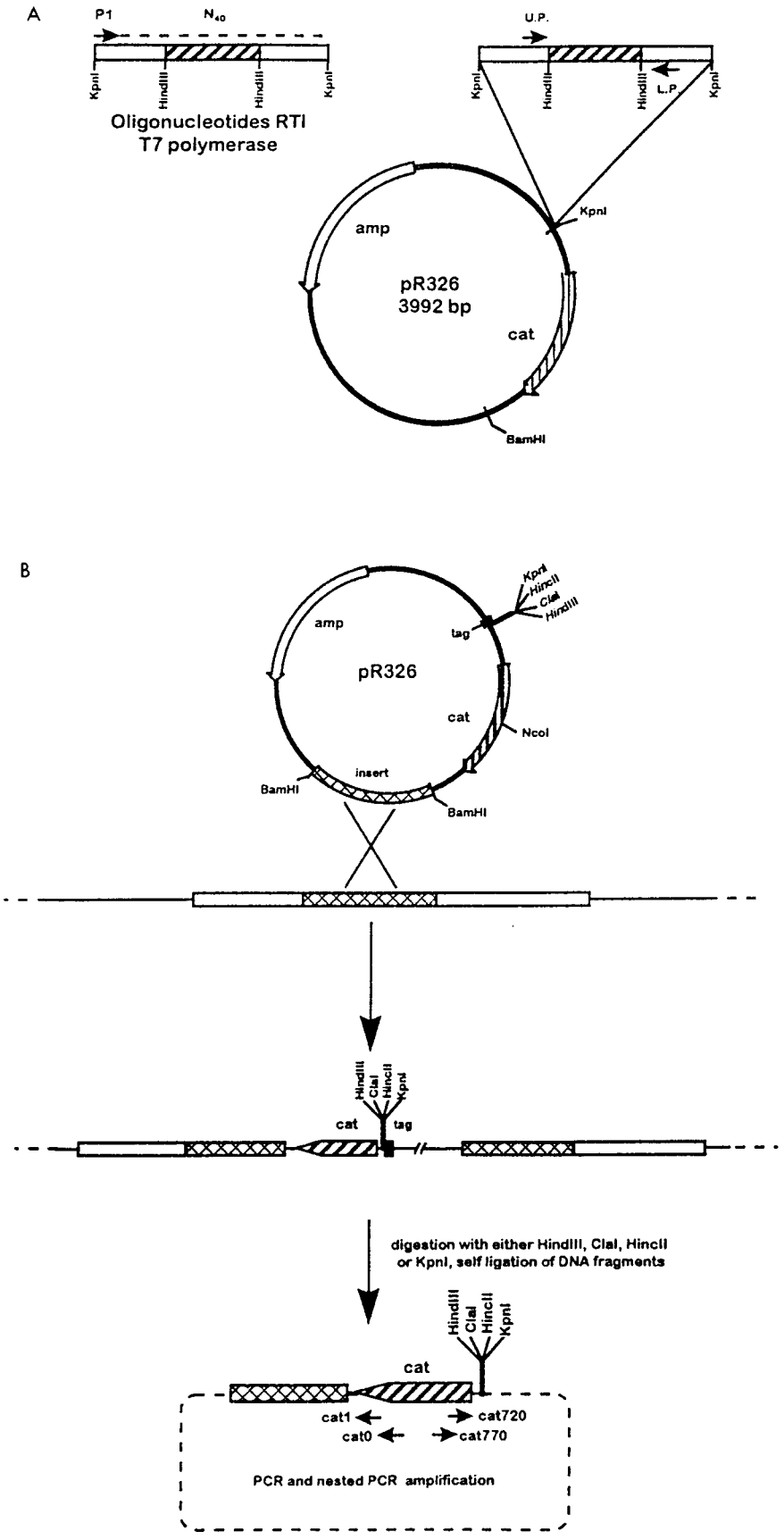
Enflurane (Abbott)-anesthetized mice were intranasally inoculated with 50 μl (~10^8 CFU) of a Todd-Hewitt broth suspension of a pool of mutants *S. pneumoniae*. Groups of six BALB/c mice were randomly assigned to each treatment. After 32 h of infection, the mice were sacrificed by cervical dislocation and their lungs were aseptically removed. The lungs from a single animal were washed free of blood and homogenized in sterile Potter homogenizers in 10 ml of sterile phosphate-buffered saline. The number of CFU in the supernatant (*in vivo* pool) was determined by serial dilution and plating in triplicate on 5% blood agar plates supplemented with chloramphenicol (8 μg/ml).

Septicemia experiments were performed as follows. Each clone was grown in TSB to an OD_{600} of 0.2. The bacterial culture was serially diluted and each dilution was plated for counting the viable cells. Dilutions of 10^5, 10^4, and 10^3 CFU/ml were mixed with an equal volume of 8% (wt/vol) hog mucin, and 500 μl of each dilution was then injected intraperitoneally into five BALB/c mice. The infected mice were monitored over 4 days; each day the surviving mice were counted, and the results were compared to those for the control strains G54 and pR212 (avirulent noncapsulated G54; received from G. Pozzi).

Probe preparation and hybridization. Five-microgram aliquots of chromosomal DNA extracted from *S. pneumoniae in vitro* and *in vivo* pools (19) were used as PCR templates to generate labeled probes. Genomic DNA was amplified by PCR as described above. The amplified DNA was separated on a 3% low-melting-point agarose gel, and DNA bands of approximately 80 bp were recovered. Then 5 μl of melted agarose was used as a template for a second PCR to generate the labeled probes. The PCR mixture (30 μl) contained primers (1 μM each), dATP, dTTP, and dGTP (0.2 mM each), 0.05 mM dCTP, 5 μl of [^{32}P]dCTP (5,000 mCi/mmol), *Taq* DNA polymerase (1 U), and 3 μl of *Taq* polymerase 10× buffer (Perkin-Elmer Cetus). Twenty cycles of denaturation (30 s at 92°C), annealing (60 s at 50°C), and polymerization (10 s at 72°C) were performed. The amplified DNA was digested with *Hin*dIII; the 40-bp fragment was recovered from a 6% low-melting-point agarose gel and added to 5 ml of hybridization buffer (5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 2% blocking reagent [Boehringer], 0.02% sodium dodecyl sulfate). The filters were hybridized for 14 h at 50°C in a Hybaid hybridization oven range, washed twice in washing buffer (5× SSC, 0.02 sodium dodecyl sulfate) for 30 min at 42, 50, and 55°C, dried, and autoradiographed with Kodak XAR-5 film at −80°C for 24 to 72 h.

Southern blot analysis. Southern blot experiments and ^{32}P DNA probe labeling were performed according to standard methods (54).

Localization of the integration site in the virulence-attenuated mutants. Inverse PCR was used to amplify the genomic DNA segment flanking the chloramphenicol resistance (*cat*) gene in the *S. pneumoniae* mutants. Two micrograms of genomic DNA was digested separately with each of the enzymes *Kpn*I, *Hin*cII, *Cla*I, and *Hin*dIII for 8 h at 37°C. Digested genomic DNA was diluted 100-fold and self-ligated. One microliter of ligation mixture was used as a template for PCR using primers cat0 (5′-ACAGCTTCCAGATCCATATCC-3′) and cat770


A
P1
N40
KpnI
HindIII
HindIII
KpnI
Oligonucleotides RTI
T7 polymerase
U.P.
KpnI
HindIII
HindIII
L.P.
KpnI
amp
pR326
3992 bp
cat
KpnI
BamHI
B
amp
tag
KpnI
HincII
ClaI
HindIII
pR326
cat
NcoI
insert
BamHI
BamHI
HindIII
ClaI
HincII
KpnI
cat
tag
digestion with either HindIII, ClaI, HincII
or KpnI, self ligation of DNA fragments
HindIII
ClaI
HincII
KpnI
cat
cat1
cat720
cat0
cat770
PCR and nested PCR amplification

(5′-TATCCCACTTTATCCAATTTT-3′) (Fig. 1). The PCR mixture (50 μl) contained primers (1 μM each), dATP, dTTP, dGTP, and dCTP (0.2 mM each), *Taq* DNA polymerase (1 U), and 5 μl of *Taq* polymerase 10× buffer (Perkin-Elmer Cetus). Twenty cycles of denaturation (30 s at 92°C), annealing (60 s at 50°C), and polymerization (1.5 min at 72°C) were performed. One microliter of the PCR reaction was used as the template in a nested reaction performed under the same conditions as described above, using primers cat1 (5′-TGACGTTGA GCCTCGGAACCCATCG-3′) and cat720 (5′-GTTGAACCATTATATCACA TT-3′) (Fig. 1). The average size of the amplified DNA product was 500 bp. The PCR products were sequenced with an automated sequencer (ABI Prism 377) as instructed by the manufacturer. Predicted amino acid sequences from all six reading frames of the DNA sequences obtained above were subjected to similarity search of the SWISS-PROT TREMBL database, using the BLAST programs (1).

RESULTS

Construction of a tagged *S. pneumoniae* chromosomal DNA library. The pR326 vector (16) is unable to replicate in *S. pneumoniae* and carries two markers: one for ampicillin resistance and the other for chloramphenicol resistance (*cat*). The latter is under the control of the pneumococcal promoter *amiA* and therefore can be expressed in *S. pneumoniae*. A mixture of double-stranded DNA tags was ligated into plasmid pR326. The ligation mixture was used to transform *E. coli* DH5α to obtain a pool of pR326 plasmid DNA tags (Fig. 1A). The central variable sequence allows the potential generation of 10^{17} different tags. Thirty ampicillin-resistant colonies were picked at random, and the variable 40-bp region of the tag was sequenced by using both upper and lower primers. None of the 30 plasmids tested contained identical 40-bp sequences (data not shown). The tagged genomic library was constructed by cloning fragments of *S. pneumoniae* chromosomal DNA from the G54 strain in the pool of tagged plasmids (Fig. 1). Therefore each fragment was associated with a unique DNA sequence tag. The G54 strain was transformed with this tagged genomic library. Since pR326 cannot replicate in *S. pneumoniae*, Cm^r colonies could be generated only by integration of the donor vector at the region of the chromosome which was homologous to the pneumococcal DNA fragment inserted in the vector (Fig. 1B). Using this procedure, we selected 1,250 Cm^r mutants.

The DNA tag from each mutant was amplified by PCR; equal amounts of all DNA tags were spotted on duplicate filters. To verify that the Cm^r clones resulted from random homologous insertion of a single tagged plasmid, Southern blot analysis was performed. The chromosomal DNA was extracted from four clones and then subjected to a single digestion with *Eco*RV (which is absent from pR326) and to a double digestion with *Eco*RV and *Nco*I (which cuts once in the *cat* gene [Fig. 1]). The *cat* gene from pR326 was used as a probe. The results showed that each mutant had arisen from a single integration of the plasmid into a random site of the bacterial genome (Fig. 2).

Screening for virulence genes. The mutant library was analyzed for reduced virulence in a mouse pneumonia model system (12). The 1,250 mutants were grouped in 25 pools of 50 clones; for each pool, duplicate filters containing spots of the individually amplified tags were prepared. The pooled mutants were grown in liquid medium, an aliquot was then removed, and genomic DNA was extracted; this represented the inoculated *in vitro* pool (Fig. 3). The bacteria were inoculated in mice intranasally, and six animals were used for each pool. After 32 h, the mice were killed and the lungs were recovered; then the bacteria were titrated and cultured in liquid medium. The bacteria recovered (approximately 10^5 from each mouse) were pooled, and the DNA was extracted; this represented the recovered *in vivo* pool (Fig. 3). The tags from the *in vitro* and the *in vivo* pools were amplified by PCR using ^{32}P label in the reaction. The invariant arms of 20 bp were released by digestion with *Hin*dIII (Fig. 1), and the radiolabeled 40-bp central sequences were used as a probe against the duplicate filters. Figure 4 shows as an example two filters hybridized with the *in vivo* and *in vitro* tags; tags that hybridize to the probe from the inoculated in vitro pool but not to the probe from the recovered in vivo pool represent mutants with attenuated virulence. From the screen of our 1,250 *S. pneumoniae* mutants, we identified 200 clones whose tag did not hybridize to the probe obtained from the recovered pools, showing reduced virulence *in vivo*.

Sequence analysis of the virulence-attenuated clones. To identify the genes responsible for virulence attenuation, we first attempted to clone the DNA regions flanking the *cat* insertion point in *E. coli* either by recovering plasmid pR326 or by cloning the resistance marker and its flanking region into a low-copy-number vector. In many cases, the plasmids were not stable and were severely rearranged. Instability of plasmids containing streptococcal DNA in *E. coli* is well documented (14, 18, 40). To circumvent this problem, the regions flanking the *cat* insertion points were identified by inverse nested PCR and subsequently sequenced as described in Materials and Methods. Sequence analysis of the 200 clones revealed that they originated from insertions in 126 different genes. Some of the mutants resulted from different insertions within the same gene, whereas in other cases the mutants had identical insertion points within the same gene.

As shown in Table 1, the sequences obtained from the 126 mutants can be grouped in five classes. Only key examples of each class are discussed below.

The first class comprises insertions in previously identified *S. pneumoniae* virulence factors: hyaluronidase (SPN-673), neuraminidase (SPN-605), autolysin (SPN-1264), and IgA1 protease (SPN-1810). One more insertion (SPN-227) was found in a gene product that belongs to the family of choline binding proteins (CBPs) since it possesses at the C-terminus the repeated signature choline binding domain (52) (Table 1).

The second class includes insertions in genes involved in metabolic pathways: purine biosynthesis (SPN-1645, SPN-786, SPN1404, and SPN-2078), phosphate assimilation (SPN-1589), synthesis of glutamine (SPN-902), and anaerobic metabolism (SPN-181).

The third class includes an insertion in a gene that show similarity to ATP-dependent proteases. SPN-1055 maps in an

FIG. 1. (A) Construction of a pool of pR326 plasmid DNA tags. A pool of double-stranded DNA tags was generated by extending with T7 DNA polymerase the 18-mer primer P1 annealed to random 89-mer templates (see Materials and Methods). Each tag is a variable sequence of 40 bp (N_{40}), where N is A, C, G, or T, flanked by two constant arms of 20 bp which contain restriction sites for *Hin*dIII and *Kpn*I. These arms allow PCR amplification with the upper primer (U.P.) and lower primer (L.P.) of each tag. Tags were digested with *Kpn*I and ligated in the vector pR326, generating a pool of tagged plasmids. (B) Construction of a library of *S. pneumoniae*-tagged mutants by homologous recombination. *S. pneumoniae* chromosomal DNA was mechanically sheared; fragments ranging from 500 to 700 bp were size selected and ligated with the pool of pR326 plasmids DNA tags. The ligation mix was directly transformed in *S. pneumoniae* G54; homologous recombination events lead to insertion-duplication mutagenesis, with the plasmid carrying the *cat* gene and the tag inserted into the inactivated gene in the chromosome. Total genomic DNA was extracted from mutant clones, digested with a restriction endonuclease, and circularized by ligation. The genomic flanking region outside the *cat* gene was amplified with the inner set of primers cat0 and cat770; 4 ng of the product of the first PCR was reamplified with the nested pair of primers cat1 and cat720. The resulting single-PCR product was then sequenced.

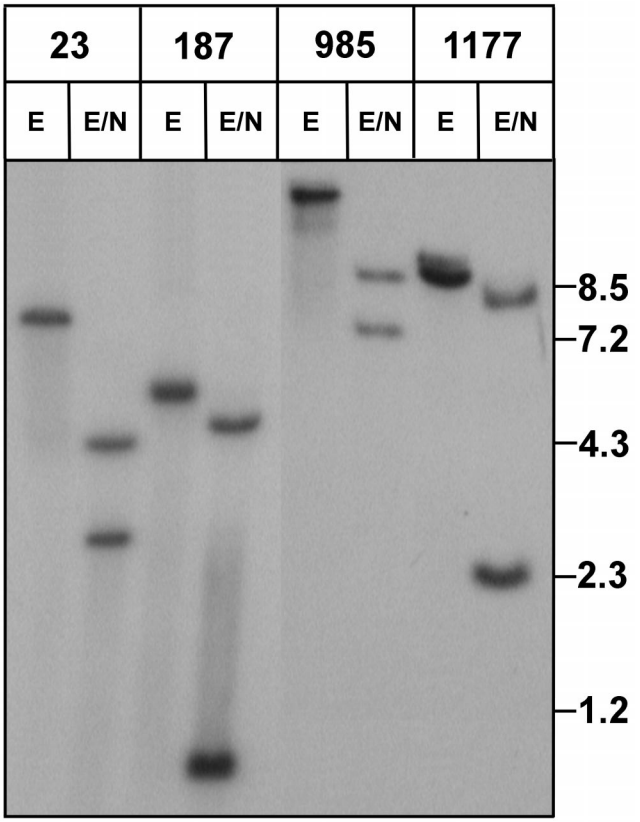


FIG. 2. Southern blot analysis of mutants. Genomic DNAs obtained from clones 23, 187, 985, and 1177 were digested with *Eco*RV (E), or with *Eco*RV and *Nco*I (E/N). The *Eco*RV site is absent in pR326; the *Nco*I site is unique to the *cat* gene. Five micrograms of digested DNA was then applied on 0.8% agarose gel, electrophoresed, blotted on nylon filter, and subjected to Southern hybridization analysis using the *cat* gene as a probe. G54 genomic DNA did not hybridize to the probe (data not shown). The positions of size markers are shown in kilobases on the right.

S. pneumoniae gene whose product shows similarity to ClpC, an ATP-dependent protease whose activity is essential for survival of *Listeria monocytogenes* in macrophages (51).

The fourth class includes insertions found in genes coding for polypeptides that show similarity to the ATP binding cassette (ABC) family of transporters, and all possess the conserved A and B Walker motifs (29). Two insertions were mapped in genes similar to *potA* (SPN-2041) and *potD* (SPN-924) belonging to the ABC transport system for spermidine and putrescine. Glutamine appears to be a key metabolite for virulence since two mutants were found to have insertions in genes homologous to *glnH* and *glnQ* (SPN-1364 and SPN-1452), two components of the well-characterized high-affinity glutamine transport system. Mutant SPN-1802 showed striking similarity to the MsmK of *S. mutans*, the ATP binding component of an ABC transporter involved in uptake of multiple sugars (53). SPN-948 showed similarity to the MesD protein (22), a member of the HlyB subfamily of ABC transporters involved in the transport of bacteriocins and toxins (20, 29).

The fifth class contains mutants showing insertions in genes involved in mutation avoidance: *recJ* and *mutX* (SPN-1411 and SPN-1634, respectively). Mutations which impair DNA recombination/repair have already been identified as essential for survival in macrophages in *S. typhimurium* (9, 27) and to be responsible for virulence attenuation in *S. aureus* (42).

Finally, 63 mutants whose insertions mapped within hypothetical genes with unknown function were identified. Some of these genes are described in Table 2.

Analysis of septicemia of the virulence-attenuated clones. To evaluate rapidly the attenuated virulence of the individual mutants in a different *in vivo* model, all 126 clones were individually analyzed in a mouse septicemia model. The mutants fell into three classes, as shown in Table 1. Those mutants showing the same virulence as G54 (intraperitoneal 50% lethal dose of 10^2) were classified as high, since all of the 15 treated mice died within 2 days. Mutants showing the same virulence attenuation as the nonencapsulated PR212 strain (intraperitoneal 50% lethal dose of $>10^5$) were classified as low, since all of the 15 treated mice survived the 4-day duration of the experiment. Mutants falling in between the two classes were defined as medium. Thirty-six mutants were shown to belong to the low-virulence class, 52 were classified as having medium virulence, and 38 were as virulent as the parent strain (G54) and thus classified as having high virulence. The analyzed mutants had different in *vitro* growth rates; however, we did not find a correlation between the loss of virulence and slow-growing bacteria since several mutants that grew very slowly *in vitro* showed the same virulence as G54 wild-type strain. We conclude that *in vivo* and *in vitro* growth rates do not correlate.

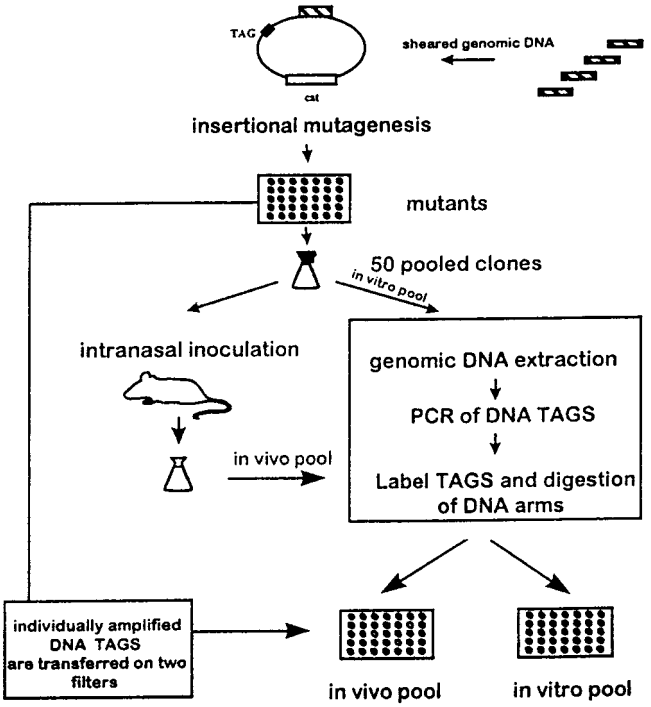


FIG. 3. Design of virulence gene screening. *S. pneumoniae* chromosomal DNA was fragmented, size selected, and ligated with the pool of pR326 plasmid DNA tags. Homologous recombination events lead to insertion-duplication mutagenesis. DNA tags were individually amplified from all mutants and applied to duplicate filters in groups of 50; 50 mutants representing the tags on each filter were pooled and grown in a flask (20 ml of broth). Fifteen milliliters was removed for DNA extraction (*in vitro* pool). Fifty microliters was used to inoculate mice intranasally; after 32 h, the mice were sacrificed and their lungs were removed and homogenized. The recovered bacteria (*in vivo* pool) were plated by serial dilutions. The bacteria recovered (approximately 50,000 CFU from each mouse) were pooled and grown in a flask, and then the genomic DNA was extracted. The tags of the *in vitro* and *in vivo* pools were amplified and labeled with [^{32}P]dCTP. Variable regions were recovered by *Hin*dIII digestion and used to probe twin filters; DNA tags that hybridize to the family of probes from the *in vitro* pool but not from the *in vivo* pool correspond to mutants negatively selected *in vivo*.

TABLE 1. Characterization of *S. pneumoniae* genes identified by STM

Class	Strain	Gene[a]	Similar to[b]:	% Amino acid identity (homology)	Function	Role in pathogenesis	Septicemia[c]
Known virulence factors	SPN-673	*yhа*			Hydrolysis of hyaluronic acid	Host colonization?	High
	SPN-605	*nanA*			Hydrolysis of *N*-acetylneuraminic acid	Host colonization?	High
	SPN-1264	*lytA*			Autolysis	Release of pneumolysin and cell wall products	High
	SPN-227	CBP-like			Unknown	Unknown	Medium
	SPN-1810	*iga*			Cleavage of human IgA1	Counteracts mucosal defense mechanisms	Low
Metabolic functions	SPN-2078		*purK* (Bsu)	48 (65)	Purine biosynthesis	In vivo survival in *Salmonella* and *Pseudomonas*	Low
	SPN-786		*purL* (Bsu)	24 (40)	Purine biosynthesis	In vivo survival in *Salmonella* and *Pseudomonas*	Medium
	SPN-1404		*purE* (Bsu)	64 (78)	Purine biosynthesis	In vivo survival in *Salmonella* and *Pseudomonas*	Medium
	SPN-1645		*purC* (Bsu)	54 (72)	Purine biosynthesis	In vivo survival in *Salmonella* and *Pseudomonas*	Medium
	SPN-181		*adh* (Bsu)	34 (53)	Anaerobic metabolism	Unknown	Low
	SPN-902		*ghuA* (Bsu)	70 (82)	Glutamine biosynthesis	In vivo survival in *Salmonella*	Medium
	SPN-1589		*phoH* (Bsu)	62 (78)	Phosphate assimilation	Unknown	High
Proteases	SPN-1055		*clpC* (Lmo)	43 (65)	ATP-dependent protease	Required for survival of *Listeria* in macrophages	Medium
ABC transporters	SPN-1364		*glnH* (Bsu)	33 (53)	Glutamine transport	In vivo survival of *Salmonella*	Medium
	SPN-1452		*glnQ* (Bsu)	56 (76)	Glutamine transport	In vivo survival of *Salmonella*	Medium
	SPN-924		*potD* (Eco)	41 (58)	Spermidine/putrescine transport	Unknown	Medium
	SPN-2041		*potA* (Eco)	64 (75)	Spermidine/putrescine transport	Unknown	Low
	SPN-1802		*msmK* (Smsu)	75 (84)	Sugar transport?	Unknown	High
	SPN-948		*msiD* (Lme)	23 (44)	Unknown	Unknown	High
DNA recombination/repair	SPN-1411		*recJ* (Hin)	34 (56)	Methylation-directed mismatch repair	Survival to oxidative burst of macrophages?	Low
	SPN-1634	*mutX*			Degradation of 8-oxo-dGTP	Survival to oxidative burst of macrophages?	Low

[a] Gene already identified in *S. pneumoniae*.

[b] Determined by comparison of sequences and coding regions with the EMBL GenBank DNA, and SWISS-PROT databases. BLAST, BLASTX, and FASTA network services were used. Eco, Bsu, Hin, Lmo, Lme, Smu denote *Escherichia coli*, *Lactococcus lactis*, *Bacillus subtilis*, *Haemophilus influenzae*, *Listeria monocytogenes*, *Leuconostoc mesenteroides*, and *Streptococcus mutans*, respectively.

[c] For definitions of high, medium, and low, see Materials and Methods.

FIG. 4. Identification of avirulent mutants. Two identical filters containing the same amount of individually amplified DNA tags (from A1 to L5) were hybridized to labeled DNA tags from the *in vitro* pool and the *in vivo* pool. The filters shown exemplify the many analyzed. Tags that weakly hybridized to the probes from both the *in vitro* pool and the *in vivo* pool (D1, E1, F1, G1, D2, C5, and E3) probably corresponded to slow-growing mutants and were not considered. Tags that hybridized to the probe from the *in vitro* pool but not to the probe from the *in vivo* pool (I5, L5, F5, L1, A2, and C4) corresponded to mutants with attenuated virulence and therefore were unable to survive in the host.

Interestingly, the mutants having insertions in the previously identified virulence factors (with the exception of SPN-227 and SPN-1810) did not show reduction in virulence in this animal model. In contrast, most of the mutants in genes involved in metabolic functions, transport, and DNA modification showed a significant reduction in virulence.

DISCUSSION

Characterization of microbial genes that are specifically required for the establishment and maintenance of an infection is an important aspect of understanding the mechanism by which microbial pathogens cause diseases. Large-scale isolation of virulence genes has been performed by either the IVET (38, 39) or STM (15, 28, 42) method.

The approach that we used to identify new virulence genes in *S. pneumoniae* was validated by the fact that among the mutants unable to survive in mice, insertions were found in previously identified or predicted pneumococcal virulence

TABLE 2. Hypothetical genes with unknown function identified by STM

Strain	Gene[a]	Similar to[b] (source and/or product)	% Amino acid identity (homology)	Notes, Prosite motif(s)	Septicemia[c]
SPN-1031		*yejD* (*Escherichia coli*)	40 (63)	Unknown function Motif: PS01149, hypothetical YciL/YejD/YjbC family signature	High
SPN-1101		*ykrA* (*Bacillus subtilis*)	40 (57)	Unknown function	Medium
SPN-1119		*PHBQ002* (*Pyrococcus horikoshii*)	30 (48)	Unknown function	High
SPN-113		*TPTC* (*Streptococcus crista*, ABC protein)	34 (59)	Hypothetical ABC transporter Motifs: PS00017, ATP/GTP binding site motif A (P-loop); PS00211, ABC transporter family signature	High
SPN-1145		*MTCY21C12.02* (*Mycobacterium tuberculosis*, putative cation-transporting ATPase)	34 (54)	Hypothetical cation-transporting ATPase Motif: PS00154, E1-E2 ATPase phosphorylation site	High
SPN-1200		*Dictyostelium discoideum* α-L-fucosidase precursor	26 (49)	Unknown function	High
SPN-1338		No homology		Unknown function Motif: PS00343, gram-positive coccus surface protein-anchoring hexapeptide	Medium
SPN-1471		*iga* (*Streptococcus sanguis*)	40 (60)	Hypothetical IgA1 protease (different from the IgA1 protease identified in *Streptococcus pneumoniae*) Motif: PS00142 zinc protease	High
SPN-1583		*yqeT* (*B. subtilis*, probable methyltransferase)	42 (59)	Unknown function	Medium
SPN-162		No homology		Unknown function	Low
SPN-1631		*ydeG* (*Schizosaccharomyces pombe*)		Unknown function	High
SPN-1808		*BB0831* (*Borrelia burgdorferi*, xylose operon regulatory protein)	30 (49)	Unknown function	High
SPN-1818		*ydhP* (*B. subtilis*)	58 (75)	Hypothetical β-glucosidase Motifs: PS00548, ribosomal protein S3 signature 1; PS00449, ATPase A signature	ND
SPN-224		*yidA* (*E. coli*)	48 (69)	Unknown function	High
SPN-627		*ydiL* (*B. subtilis*)	31 (64)	Unknown function	High
SPN-631		*ypsA* (*B. subtilis*)	39 (56)	Unknown function	High
SPN-633	*ypoA*			Unknown function	High
SPN-636		*radA* (*B. subtilis*, possible DNA repair protein)	65 (82)	Hypothetical DNA repair protein (by homology) Motif: PS00017, ATP/GTP binding site motif A (P-loop)	ND
SPN-655		*ORF8* (*Enterococcus faecalis*)	45 (63)	Unknown function	Medium
SPN-233		*ybhL* (*E. coli*, hypothetical membrane protein)	41 (60)	Unknown function	High
SPN-634		No homology		The insertion is located in an intergenic region close to regulatory sequences BoxA and BoxB	High
SPN-641		*yitS* (*B. subtilis*)	27 (48)	Unknown function	Medium
SPN-962		*N*5,*N*10-Methylenetetrahydromethanopterin reductase gene (*Staphylococcus carnosus*)	54 (75)	Hypothetical *N*5,*N*10-methyltetrahydromethanoterin reductase (by homology)	High
SPN-1297		*pstB* (*Methanococcus jannaschii*, hypothetical ABC transporter)	58 (80)	Hypothetical ABC transporter (by homology) Motifs: PS00017, ATP/GTP binding site motif A (P-loop); PS00211 ABC transporter family signature	High

[a] Gene already identified in *S. pneumoniae*.
[b] Determined obtained by comparison of sequences and coding regions with the EMBL, GenBank DNA and SWISS-PROT databases. BLAST, BLASTX, and FASTA network services were used.
[c] For definitions of high, medium, and low, see Materials and Methods. ND, not determined.

genes (Table 1). There are contradictory reports about the role of autolysin in virulence; some authors reported that autolysin-negative mutants were less virulent than the parent strain (4), while others claimed that autolysin played no role (56). Our first screen demonstrates the importance of autolysin for establishing pneumonia, while intraperitoneal inoculation of the same mutant shows that it is not necessary for septicemia. The discrepancies between the published data and ours could be due to differences in the genetic methodologies. The role of neuraminidase in streptococcal pathogenesis has been inferred from immunization studies showing that the purified protein partially protected mice against pneumococcal infection (36). A hypothetical role of hyaluronidase in virulence has been suggested since this enzyme hydrolyzes hyaluronic acid, an important component of connective tissue (5). Our results are the first genetic evidence that these genes play an important role during pneumonia infection. No significant reduction in virulence was observed when mutants SPN-605, SPN-673, and SPN-1264 were analyzed in a mouse septicemia model. In this model, the adhesion and colonization steps are not required since bacteria are introduced directly in the bloodstream. It is thus likely that these genes are involved in steps specific for pneumonia infection. We did not find in our screen genes involved in capsule biosynthesis. The mutagenesis that we performed was not saturating since we screened only 1,250 insertional mutants. However, we were surprised not to find at least one mutant mapping in the 20-kb-long capsule cassette.

Neither the pneumolysin gene nor a pneumolysin homologue which can be identified by sequence homology in the genome (this gene might code for a second hemolysin recently identified in *S. pneumoniae* [13]) was found in our screen. This could be due to the fact that mutants deficient in virulence proteins acting extracellularly such as the hemolysins might not be detected by this methodology since they will be complemented by the other strains in the pool.

We found a CBP whose inactivation (SPN-227) resulted in

virulence attenuation in both pneumonia and septicemia models. LytA and PspA are the best-characterized CBPs: LytA is the major cell wall amidase and is required for cell autolysis (50); PspA has been characterized as a protective antigen, although its function is unknown (60, 61). Additional CBPs have been identified recently in *S. pneumoniae* (25, 52): SpsA described by Hammerschmidt *et al.* (25), CbpA described by Rosenow *et al.* (52), and PcpA (GenBank accession no. Z82001) identified by A. Sanchez-Beato, and J. Garcia. From genomic analysis, we know that our CBP is different from the above-mentioned genes that map at other loci on the G54 chromosome.

From our *in vivo* screen, we identified a mutant (SPN-1810) with an insertion in the IgA1 protease; this enzyme belongs to the class of extracellular endopeptidases which specifically cleave human (52) but not mouse (48) IgA1. It has recently been found that in *Neisseria gonorrhoeae*, IgA1 protease is important for survival of this pathogen within epithelial cells due to its ability to cleave LAMP1, a major integral membrane protein of lysosomes (35). Since inactivation of IgA1 leads to reduced virulence in a mouse model, we suggest that IgA1 protease in *S. pneumoniae* may play a role in pathogenesis by cleaving some substrate(s) other than immunoglobulins. Further support for this hypothesis comes from the fact that SPN-1810 has attenuated virulence in both pneumonia and septicemia models; thus, the IgA1 protease could play roles in addition to counteracting the mucosal defenses (a specific step in pneumonia infection). Alternatively, the phenotype observed could be ascribed to a polar effect on the expression of downstream genes. From our genome analysis, it was possible to identify only one downstream open reading frame (ORF) since our contig ends with this ORF. Surprisingly, this ORF codes for a gene showing homology to *S. pneumoniae iga* (29% identity and 49% similarity). This *iga*-like ORF belongs to the same operon as *iga*, and we cannot tell if there are termination signals downstream. This gene, the one identified by STM (mutant SPN-1471 [Table 2], and *iga* (59) constitute a family of *iga*-like genes stressing the importance of its function for *S. pneumoniae*.

The ability to adapt to the host environment is a key component of pathogenicity. Many genes whose products are involved in metabolic pathways and nutrient uptake have indeed been identified following our negative *in vivo* selection. *De novo* synthesis of purines has been shown to be a strict requirement for bacterial growth and persistence under the nucleotide-limiting conditions in the host (38); also purine auxotrophy attenuates many pathogenic bacteria (58). Accordingly, insertions in four genes belonging to the purine biosynthetic pathway (*purE*, *purK*, *purC*, and *purL*) were found. Three insertions in genes involved in glutamine metabolism, *glnA* (glutamine synthase), *glnQ*, and *glnH* (high-affinity glutamine transport system), were identified, suggesting that reduced availability of glutamine synthesis or loss of high-affinity glutamine transport leads to virulence attenuation in *S. pneumoniae*. Klose and Mekalanos (34) reported that only the simultaneous inactivation of glutamine synthesis (*glnA*) and high-affinity transport (*glnH*) reduced virulence in *S. typhimurium*. Since *salmonellae* also have a low-affinity glutamine transport system, a mutant defective in both *glnA* and *glnH* can survive in vitro when supplemented with a high level of glutamine but not *in vivo* due to the low level of this metabolite in animal tissues. Glutamine is a key metabolite since it is the major donor of amino groups in the amino acid biosynthetic pathway and it is required in the first two steps of purine biosynthesis. The virulence defects of *S. pneumoniae* support the hypothesis that the concentration of glutamine in the host is a critical parameter for *in vivo* survival and show that import and biosynthesis are both required *in vivo*.

Nutrient and amino acid uptake seems to be a critical factor for *in vivo* survival of *S. pneumoniae*. Two mutants unable to grow in mice had insertions in *potA* and *potD*, which encode the ATPase component and substrate binding lipoprotein, respectively, of the spermidine/putrescine ABC transport system. Spermidine and putrescine are polyamines whose synthesis uses arginine as a precursor. A mutant in arginine transport in *L. monocytogenes* (33) and more recently a mutant in arginine biosynthesis in *V. cholerae* (11) were both shown to be attenuated in virulence. One mutant (SPN-1802) had an insertion in a gene with strong homology to *msmK*, a component of a multiple-sugar uptake system of *S. mutans* (53). *S. pneumoniae* is limited in sugar fermentation ability, being able to use only glucose, raffinose, lactose, threalose, and maltose as sources of energy (30, 44). The MsmK homologue in *S. pneumoniae* is likely to play a role in the transport of those sugars.

One gene involved in anaerobic metabolism, *adh* (encoding alcohol dehydrogenase) was identified in this study. *S. pneumoniae* is a facultative anaerobe, and its energy-yielding metabolism is fermentative. Alcohol dehydrogenase catalyzes the reduction of acetyl coenzyme A first to acetaldehyde and then to ethanol. These reactions are needed to regenerate reducing ability (NADH) from glycolysis under anaerobic conditions. Our data suggest that as infection progresses, streptococci must adapt to low-oxygen tension environments, and a fully active anaerobic metabolism is required for survival in the host.

One insertion (SPN-1055) maps in a gene whose product shows similarity to ClpC, one of the regulatory subunits of Clp proteases. Clp proteases are a family of highly conserved proteins involved in ATP-dependent proteolysis (23). The Clp protease in *E. coli* is a two-component system where the larger regulatory subunit (ClpA, ClpB, or ClpX) has ATPase activity and the smaller subunit (ClpP) has proteolytic activity. The Clp protease is involved not only in degradation of toxic denatured proteins but also in controlling the half-life of unstable proteins to modulate gene expression. In fact, Clp proteases have been shown to be important in modulation of stress response, stationary phase, and competence in *B. subtilis* (41). Recent evidence suggests that Clp proteases also may be involved in modulating virulence. A ClpP homologue has been shown to be essential for survival of *S. typhimurium* in the spleen (28), the regulatory subunit ClpC is required for survival of *L. monocytogenes* in macrophages (51), and more recently a ClpX homologue has been found to be important for virulence in *S. aureus* (42). It is not clear what role is played by these Clp proteins in the virulence process or what pathways are affected, but it is tempting to suggest a common mechanism of action. Identification of the gene products that are regulated by the *S. pneumoniae* ClpC homologue will shed light on the role of this regulatory protein in virulence.

The ability to keep spontaneous mutation levels low seems to be crucial for in vivo survival. A mutant (SPN-1634) showing insertions in the *S. pneumoniae mutX* gene (the homologue of *E. coli mutT*) (43) and a mutant (SPN-1411) whose insertion maps in a gene similar to *recJ* were identified. MutX is an NTPase whose role is to hydrolyze 8-oxo-dGTP to 8-oxo-dGMP, thus preventing its incorporation into DNA (21, 43). RecJ is a single-stranded DNA binding protein involved in DNA recombination/repair (37). It is known that during infection, pathogenic bacteria may encounter a variety of oxygen radicals produced by macrophages and neutrophils. These oxygen species are likely to induce mutations in the DNA or to oxidize the NTP pool. The ability to repair DNA damage has indeed been shown to be essential for full *S. typhimurium* (27)

and *S. aureus* (42) virulence. The recombination-deficient *recA* and *recBC* mutants of *S. typhimurium* have been shown to be sensitive to the oxidizing compounds generated by the oxidative burst of macrophages or phagocytes (9, 27).

The results presented in this paper clearly show that *S. pneumoniae* virulence is a multifactorial process and that, in addition to factors involved in specific steps in pneumonia infection, many other gene products are required to confer to the bacterial cells the ability to grow and survive in the host. Most of the mutants isolated that bore genes involved in metabolic functions, transport, and DNA modification showed reduction in virulence in both the pneumonia and septicemia models. This finding is consistent with the hypothesis that these functions are needed for bacterial adaptation to the different host environments and not for specific steps during infection. The adaptation of IVET and STM technologies to several distinct bacterial pathogens is revealing that different pathogens causing different diseases may have virulence determinants in common. Indeed, many of the genes identified in this work as important for *S. pneumoniae* survival in the host were also identified as virulence determinants in other pathogens (*S. typhimurium*, *S. aureus*, *V. cholerae*, and *P. aeruginosa*).

Insertions in several genes that show no homology to known genes in databases or that have similarity to hypothetical proteins could potentially encode novel factors that may play important roles in the bacterial infection process. The identification of the entire ORF and search for specific motifs or domains may give clues about the putative function of these proteins.

Despite the possibility that some of our mutants show polar effects, our data provide strong evidence that the ability to sense and respond to complex and overlapping signals encountered in the host is a key point in pathogenicity. This ability results in the elaboration of regulatory, metabolic, and virulence functions that contribute to growth and persistence in the host. We are currently constructing nonpolar mutations to exclude the involvement of neighboring genes in virulence. We believe that the approach described here can serve as a basis for a more comprehensive understanding of *S. pneumoniae* pathogenicity.

ACKNOWLEDGMENTS

We thank F. Faggioni for excellent technical assistance. We are in debt to G. Pozzi for providing strain PR212 and for helpful discussions. We thank J. P. Claverys for providing pR326 and for critical reading of the manuscript.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Austrian, R.** 1984. Pneumococcal infections, p. 257–288. *In* R. Germanier (ed.), Bacterial vaccines. London, Academic Press.

2a. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1996. Current protocols in molecular biology. Wiley Interscience, New York, N.Y. (CD-Rom version.)

3. **Benton, K. A., M. P. Everson, and D. E. Briles.** 1995. A pneumolysin-negative mutant of *Streptococcus pneumoniae* causes chronic bacteremia rather than acute sepsis in mice. Infect. Immun. **63:**448–455.
4. **Berry, A. M., R. A. Lock, D. Hansman, and J. C. Paton.** 1989. Contribution of autolysin to virulence of *Streptococcus pneumoniae*. Infect. Immun. **57:** 2324–2330.
5. **Berry, A. M., R. A. Lock, S. M. Thomas, D. Prasana Rajan, D. Hansman, and J. C. Paton.** 1994. Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*. Infect. Immun. **62:**1101–1108.
6. **Berry, A. M., and J. C. Paton.** 1996. Sequence heterogeity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*. Infect. Immun. **64:**5255–5262.
7. **Berry, A. M., J. Yother, D. E. Briles, D. Hansman, and J. C. Paton.** 1989. Reduced virulence of a defined pneumolysin-negative mutant of *Streptococcus pneumoniae*. J. Gen. Microbiol. **138:**249–259.
8. **Briles, D. E., J. Yother, and L. McDaniel.** 1988. Role of pneumococcal surface protein A in the virulence of *Streptococcus pneumoniae*. Rev. Infect. Dis. **10:**S372–S374.
9. **Buchmeier, N. A., C. J. Lipps, M. Y. H. So, and F. Heffron.** 1993. Recombination-deficient mutants of *Salmonella typhimurium* are avirulent and sensitive to the oxidative stress of macrophages. Mol. Microbiol. **7:**933–936.
10. **Camara, M., G. J. Boulnois, P. W. Andrew, and T. J. Mitchell.** 1994. A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. Infect. Immun. **62:**3688–3695.
11. **Camilli, A., and J. J. Mekalanos.** 1995. Use of recombinase gene fusion to identify *Vibrio cholerae* genes induced during infection. Mol. Microbiol. **18:** 671–683.
12. **Canvin, J. R., A. P. Marvin, M. Sivakumaran, J. C. Paton, G. J. Boulnois, P. W. Andrew, and T. J. Mitchell.** 1995. The role of pneumolysin and autolysin in the pathogenicity of pneumonia and septicemia in mice infected with type 2 pneumococcus. J. Infect. Dis. **172:**119–123.
13. **Canvin, J. R., J. C. Paton, G. J. Boulnois, P. W. Andrew, and T. J. Mitchell.** 1997. *Streptococcus pneumoniae* produces a second haemolysin that is distinct from pneumolysin. Microb. Pathog. **22:**129–132.
14. **Chen, J. D., and D. A. Morrison.** 1987. Cloning of *Streptococcus pneumoniae* DNA fragments in *E. coli* requires vectors protected by strong transcriptional terminators. Gene **55:**179–187.
15. **Chiang, S. L., and J. J. Mekalanos.** 1998. Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. Mol. Microbiol. **27:**797–805.
16. **Claverys, J. P., A. Dintilhac, E. V. Pestova, B. Martin, and D. A. Morrison.** 1995. Construction and evaluation of a new drug-resistance cassette for gene disruption mutagenesis in *Streptococcus pneumoniae*, using an *ami* test platform. Gene **164:**123–128.
17. **DeVelasco, E. A., A. F. M. Verheul, J. Verhoef, and H. Snippe.** 1995. *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines. Microbiol. Rev. **59:**591–603.
18. **Dillard, J. P., and J. Yother.** 1991. Analysis of *Streptococcus pneumoniae* sequences cloned into *Escherichia coli*: effect of promoter strength and transcription terminators. J. Bacteriol. **173:**5105–5109.
19. **Dillard, J. P., and J. Yother.** 1994. Genetic and molecular characterization of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 3. Mol. Microbiol. **12:**959–972.
20. **Fath, M. J., and R. Kolter.** 1993. ABC transporters: bacterial exporters. Microbiol. Rev. **57:**995–1017.
21. **Fowler, R. G., and R. M. Schaaper.** 1997. The role of *mutT* gene of *Escherichia coli* in maintaining replication fidelity. FEMS Microbiol. Rev. **21:**43–54.
22. **Fremaux, C., Y. Hechard, and Y. Cenatiempo.** 1995. Mesentericin Y105 gene clusters in *Leuconstoc mesenteroides* Y105. Microbiology **141:**1637–1645.
23. **Gottesman, S., C. Squires, E. Pichersky, M. Carrington, M. Hobbs, J. S. Mattick, B. Dalrymple, H. Kuramitsu, T. Shiroza, T. Foster, W. P. Clark, B. Ross, C. L. Squires, and M. R. Maurizi.** 1990. Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. Proc. Natl. Acad. Sci. USA **87:**3513–3517.
24. **Hallet, B., R. Rezsohazy, J. Mahillon, and J. Delcour.** 1994. IS231A insertion specificity: consensus sequence and DNA bending at the target site. Mol. Microbiol. **14:**131–139.
25. **Hammerschmidt, S., S. R. Talay, P. Brantzaeg, and G. S. Chhatwal.** 1997. SpsA, a novel pneumococcal surface protein with specific binding to secretory Immunoglobulin A and secretory component. Mol. Microbiol. **25:**1113–1124.
26. **Håvarstein, L. S., G. Coomarswamy, and D. A. Morrison.** 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **92:**11140–11144.
27. **Heitoff, D. M., C. P. Conner, P. C. Hanna, S. M. Julio, U. Hentschel, and M. J. Mahan.** 1997. Bacterial infection as assesses by *in vivo* gene expression. Proc. Natl. Acad. Sci. USA **94:**934–939.
28. **Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269:**400–403.
29. **Higgins, C. F.** 1992. ABC transporters: from microorganism to man. Annu. Rev. Cell Biol. **8:**67–113.
30. **Holt, J. G., N. R. Krieg, P. H. A. Sneath, J. T. Staley, and S. T. Williams.** 1994. Gram-positive cocci, p. 555. *In* W. R. Hensyl (ed.), Bergey's manual of determinative bacteriology. Williams & Wilkins, Baltimore, Md.
31. **Jann, K., and G. Westphal.** 1975. The antigens. Academic Press, London.
32. **Johnston, R. B., Jr.** 1991. Pathogenesis of pneumococcal pneumonia. Rev. Infect. Dis. **13:**S509–S517.
33. **Klarsfeld, A. D., P. L. Goossens, and P. Cossart.** 1994. Five *Listeria monocytogenes* genes preferentially expressed in infected mammalian cells: *plc*A, *pur*H, *pur*D, *pyr*E and an arginine ABC transporter gene, *arp*J. Mol. Microbiol. **13:**585–597.
34. **Klose, K. E., and J. J. Mekalanos.** 1997. Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates *Salmonella typhi-*

murium virulence. Infect. Immun. **65:**587–596.
35. **Lin, L., P. Ayala, J. Larson, M. Mulks, M. Fukuda, S. R. Carlsson, C. Enns, and M. So.** 1997. The Neisseria type 2 IgA1 protease cleaves LAMP1 and promotes survival of bacteria within epithelial cells. Mol. Microbiol. **24:** 1083–1094.
36. **Lock, R. A., J. C. Paton, and D. Hansman.** 1988. Comparative efficacy of pneumococcal neuraminidase and pneumolysin as immunogenes protective against Streptococcus pneumoniae. Microb. Pathog. **5:**461–467.
37. **Lovett, S. T., and R. D. Kolodner.** 1989. Identification and purification of a single stranded DNA specific exonuclease encoded by the *recJ* gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86:**2627–2631.
38. **Mahan, M. J., J. M. Slauch, and J. J. Mekalanos.** 1993. Selection of bacterial virulence genes that are specifically induced in host tissues. Science **259:**686–688.
39. **Mahan, M. J., J. W. Tobias, J. M. Slauch, P. C. Hanna, R. J. Collier, and J. J. Mekalanos.** 1995. Antibiotic based selection for bacterial genes that are specifically induced during infection of a host. Proc. Natl. Acad. Sci. USA **92:** 669–673.
40. **Martin, B., G. Alloing, C. Boucraut, and J. P. Claverys.** 1989. The difficulty of cloning *Streptococcus pneumoniae* mal and ami loci in *Escherichia coli*: toxicity of malX and amiA gene products. Gene **80:**227–238.
41. **Masadek, T., F. Kunst, and G. Rapoport.** 1994. MecB of Bacillus subtilis, a member of the ClpC ATPase family, is a pleiotropic regulator controlling competence gene expression and growth at high temperature. Proc. Natl. Acad. Sci. USA **91:**5788–5792.
42. **Mei, J. M., F. Nourbakhsh, C. W. Ford, and D. W. Holden.** 1997. Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. Mol. Microbiol. **26:**399–407.
43. **Mejean, V., C. Salles, L. C. Bullions, M. J. Bessman, and J. P. Claverys.** 1994. Characterization of the *mutX* gene of *Streptococcus pneumoniae* as a homologue of *Escherichia coli mutT* and tentative definition of a catalytic domain of the dGTP pyrophosphohydrolases. Mol. Microbiol. **11:**323–330.
44. **Nieto, C., M. Espinosa, and A. Puyet.** 1997. The maltose/maltodextrine regulon of *Streptococcus pneumoniae*. J. Biol. Chem. **232:**30860–30865.
45. **Paton, J. C., P. W. Andrew, G. J. Boulnois, and T. J. Mitchell.** 1993. Molecular analysis of the pathogenicity of *Streptococcus pneumoniae*: the role of streptococcal proteins. Annu. Rev. Microbiol. **47:**89–115.
46. **Poulsen, K., J. Reinholdt, and M. Kilian.** 1996. Characterization of the *Streptococcus pneumoniae* immunoglobulin A1 protease gene (*iga*) and its translation product. Infect. Immun. **64:**3957–3966.
47. **Pozzi, G., L. Masala, F. Iannelli, R. Manganelli, L. S. Havarstein, L. Piccoli, D. Simon, and D. A. Morrison.** 1996. Competence for genetic transformation in encapsulated strains of *Streptococcus pneumoniae*: two allelic variants of the peptide pheromone. J. Bacteriol. **178:**6087–6090.
48. **Proctor, M., and P. J. Manning.** 1990. Production of immunoglobulin A protease by *Streptococcus pneumoniae* from animals. Infect. Immun. **58:** 2733–2737.
49. **Renault, P., J. F. Nogrette, N. Galleron, J. J. Gordon, S. D. Erlich.** 1995. Specificity of insertion of *Tn*1545 transposon family in *Lactococcus lactis* subsp. lactis. Dev. Biol. Stand. **85:**535–541.
50. **Ronda, C., J. L. Garcia, E. Garcia, J. Sanchez-Puelles, and R. Lopez.** 1987. Biological role of the pneumococcal amidase. Cloning of the *lytA* gene in *Streptococcus pneumoniae*. Eur. J. Biochem. **164:**621–624.
51. **Roquette, C. M. T. Ripio, E. Pellegrini, J. M. Bolla, R. I. Tascon, J. A. Vasquez-Boland, and P. Berche.** 1996. Identification of a ClpC ATPase required for stress tolerance and in vivo survival of *Listeria monocytogenes*. Mol. Microbiol. **21:**977–987.
52. **Rosenow, C., P. Ryan, J. N. Weiser, S. Johnson, P. Fontan, A. Ortqvist, and R. Masure.** 1997. Contribution of novel cholin-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. Mol. Microbiol. **25:**819–829.
53. **Russel, R. R., J. Aduse-Opoku, L. Tao, and J. J. Ferretti.** 1992. A bonding protein-dependent transport system in *Streptococcus mutans* responsible for multiple sugar metabolism. J. Biol. Chem. **267:**4631–4637.
54. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laborabory Press, Cold Spring Harbor, N.Y.
55. **Spellerberg, B., D. R. Cundell, J. Sandros, B. J. Pearce, I. Idanpaan-Heikkila, C. Rosenow, and H. R. Masure.** 1996. Pyruvate oxidase as a determinant of virulence in *Streptococcus pneumoniae*. Mol. Microbiol. **19:**803–813.
56. **Tomasz, A., P. Moreillon, and G. Pozzi.** 1988. Insertional inactivation of the major autolysin gene of *Streptococcus pneumoniae*. J. Bacteriol. **170:**5931–5934.
57. **Walsh, C., and C. L. Cepko.** 1992. Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science **255:**434–440.
58. **Wang, J. A., Mushegian, S. Lory, and S. Jin.** 1996. Large-scale isolation of candidate virulence genes of *Pseudomonas aeruginosa* by *in vivo* selection. Proc. Natl. Acad. Sci. *USA* **93:**10434–10439.
59. **Wani, J. H., J. V. Gilbert, A. G. Plaut, and J. N. Weiser.** 1996. Identification, cloning and sequencing of the immunoglobulin A1 protease gene of *Streptococcus pneumoniae*. Infect. Immun. **64:**3967–3974.
60. **Yother, J., and D. E. Briles.** 1992. Structural properties and evolutionary relationships of PspA, a surface protein of *Streptococcus pneumoniae*, as revealed by sequences analysis. J. Bacteriol. **174:**601–609.
61. **Yother, J., G. L. Handsome, and D. E. Briles.** 1992. Truncated forms of PspA that are secreted from *Streptococcus pneumoniae* and their use in functional studies and cloning of the *pspA* gene. J. Bacteriol. **174:**610–618.

Editor: V. A. Fischetti