A XerD Recombinase with Unusual Active Site Motifs in *Streptococcus pneumoniae*

Peter Reichmann and Regine Hakenbeck*

Department of Microbiology, University of Kaiserslautern, Paul-Ehrlich Straße, D-67663 Kaiserslautern, Germany

Abstract

XerD belongs to the site specific recombinases of the integrase family of proteins that catalyze recombination events via a phosphotyrosine intermediate. Sequence alignments and crystal structure resolution of *E. coli* XerD and related enzymes demonstrated the importance of four conserved amino acids R-H-R-H that are spaced along the C-terminal domain in addition to a conserved K and the active site Y, all of which have been implicated in catalysis. The deduced amino acid sequence of the putative *S. pneumoniae* XerD contained three unique replacements at the conserved positions resulting in L-Q-R-L; moreover, the active site Y was the penultimate amino acid residue, and the extreme C-terminal region suggested to be involved in interaction of *E. coli* XerD with XerC was lacking. Severe growth defects in a loss-of-function xerD mutant are consistent with an important in vivo function of the *S. pneumoniae* XerD protein. Highly related xerD genes with similar unusual amino acid replacements were found in *S. mitis*, *S. mutans* and *S. pyogenes* but not in other Gram-positive bacteria, although the genetic environment was very similar in many species. There are at least another four genes in the *S. pneumoniae* KNR_7/87 genome encoding Xer related peptides, one of which was identified as the xerC homologue. The xerD and xerC genes were present in a sample of 20 *S. pneumoniae* strains whereas the other xer genes appear to be absent in some of the strains and are more closely related to integrases of phage and transposon origin.

Introduction

Chromosome segregation at cell division is an essential feature of all bacteria. In *E. coli*, this process is associated with the dif locus in the replication terminus region which is the substrate for two site specific recombinases XerC and XerD, converting dimeric chromosomes into monomers (Sherratt et al., 1995). These proteins are members of the integrase family of recombinases that become covalently linked to the DNA during catalysis through a tyrosine residue (Nunes-Düby et al., 1998; Recchia and Sherratt 1999). Their functions include recombination events of phage and plasmid DNA with the chromosome, maintenance of monomeric plasmid DNA, and regulatory DNA rearrangements.

Despite overall sequence similarity, the formation of the recombination synapse is believed to proceed essentially in a similar manner involving two pairs of recombinase molecules each acting on one of the two DNA strands to be recombined. Distinct regions that appear conserved in the integrase family have been implicated in DNA cleavage, DNA rejoining, and protein-protein interaction to ensure formation of the synaptic complex (for review, see (Craig 1988)). The crystal structures of five site specific recombinases, the enzymatic active C-terminal domains of λ Int and the *Haemophilus influenzae* phage HP1, the *E. coli* recombinase XerD, the Cre recombinase of bacteriophage P1 and the Flp recombinase from yeast complexed with DNA have been determined at 1.9 Å, 2.7 Å, 2.2 Å, 2.4 Å and 2.65 Å respectively (Guo et al., 1997; Subramanya et al., 1997; Hickman et al., 1997; Kwon et al., 1997; Chen et al., 2000). This information together with sequence alignment of 105 members of the Int family of recombinases allowed the definition of conserved amino acids and areas of similarity that relate to the activity of these enzymes (see (Nunes-Düby et al., 1998), and references within). These include, in addition to the active site tyrosine which is located in the C-terminal domain and a lysine in a β-hairpin turn, four conserved amino acids referred to as the “R-H-R-H sandwich” (Nunes-Düby et al., 1998). In the Cre protein, where one of the histidines is replaced by W315, these residues have been shown to be part of the positively charged catalytic pocket and donate hydrogen bonds to oxygen atoms of the scissile phosphate (Guo et al., 1997). Moreover, the site interacting with the major groove of the DNA is conserved in the alignment. The extreme C-terminal end with the active site tyrosine is also involved in protein-protein interaction in the synaptic complex. It shows major structural differences in the four known protein structures, possibly indicating distinct modes of activation of the tyrosine (Subramanya et al., 1997).

In *S. pneumoniae*, we identified a putative xerD gene encoding a protein with unusual amino acid alterations at the conserved sites R-H-R-H. The two histidines were replaced by leucine and glutamine, and one arginine was converted into a leucine; in...
Further Reading

Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at caister.com

• MALDI-TOF Mass Spectrometry in Microbiology
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof

• Aspergillus and Penicillium in the Post-genomic Era
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2

• The Bacteriocins: Current Knowledge and Future Prospects
  Edited by: RL Dott, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins

• Omics in Plant Disease Resistance
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr

• Acidophiles: Life in Extremely Acidic Environments
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles

• Climate Change and Microbial Ecology: Current Research and Future Trends
  Edited by: J Marxsen (2016)
  www.caister.com/climate

• Biofilms in Bioremediation: Current Research and Emerging Technologies
  Edited by: G Lear (2016)
  www.caister.com/biorem

• Microalgae: Current Research and Applications
  Edited by: MN Tsiaouli (2016)
  www.caister.com/microalgae

• Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma

• Virus Evolution: Current Research and Future Directions
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusvol

• Arboviruses: Molecular Biology, Evolution and Control
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo

• Shigella: Molecular and Cellular Biology
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella

• Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment
  Edited by: AM Romani, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms

• Alphaviruses: Current Biology
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha

• Thermophilic Microorganisms
  Edited by: F Li (2015)
  www.caister.com/thermophile

• Flow Cytometry in Microbiology: Technology and Applications
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow

• Probiotics and Prebiotics: Current Research and Future Trends
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics

• Epigenetics: Current Research and Emerging Trends
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015

• Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications
  Edited by: A Burkovski (2015)
  www.caister.com/cory

• Advanced Vaccine Research Methods for the Decade of Vaccines
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines

• Antifungals: From Genomics to Resistance and the Development of Novel Agents
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals

• Bacteria-Plant Interactions: Advanced Research and Future Trends
  www.caister.com/bacteria-plant

• Aeromonas
  Edited by: J Graf (2015)
  www.caister.com/aeromonas

• Antibiotics: Current Innovations and Future Trends
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics

• Leishmania: Current Biology and Control
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish

• Acanthamoeba: Biology and Pathogenesis (2nd edition)
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba

• Microarrays: Current Technology, Innovations and Applications
  Edited by: Z He (2014)
  www.caister.com/microarrays

• Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications
  Edited by: D Marco (2014)
  www.caister.com/n2

Order from caister.com/order
addition, the active site tyrosine preceded a C-terminal arginine. The gene was located in a gene cluster recognizable in a large variety of Gram positive bacteria, but similar changes of the deduced XerD protein were found only in related *Streptococcus* spp.. The genome of the type 4 *S. pneumoniae* strain KNR_7/87 contains another four xer related genes, three of which were reported previously (Chalker *et al*., 2000). One of them could now be assigned as the bacterial xerC gene since it was conserved in a genomic comparison involving a collection of 20 *S. pneumoniae* strains similar to xerD (Hakenbeck *et al*., 2001). In contrast, the other xer genes appeared to be absent and part of a larger mobile gene cluster in some of the strains.

**Results**

**The xerD locus in Streptococcus pneumoniae**

A Blast search of the DNA sequence of a 2.8 kb DNA fragment obtained from a Sall/Kpnl digest of chromosomal DNA of *S. pneumoniae* R6, initially isolated in an attempt to clone antibiotic resistance determinants, revealed homology to xerD homologues. Since only xerC-related genes have been described in *S. pneumoniae* so far (Chalker *et al*., 2000), we investigated the fragment in detail. There were four complete ORFs flanked by one incomplete ORF, all with the same orientation (Figure 1). orfB encoded a putative 246 amino acid protein with homology to the β-integrase family of recombinases, especially to the XerD proteins. orfA shows some similarities to inosine monophosphate dehydrogenases, orfE is related to rluB genes encoding pseudouridine synthases that modify rRNA, and conserved homologues were ubiquitously found in all prokaryotes and in yeast as well. The function of orfC and orfD remains unknown.

The 738 bp xerD started at A544TG sequence and terminated at T128AA of the 2808 bp DNA fragment. There is another possible start codon at A550TG, but the distance to the Shine-Dalgarno sequence makes the second possible translation start less likely. No putative promoter or terminator sequences were found, and investigation of the *S. pneumoniae* KNR_7/87 genome database (The Institute for Genomic Research TIGR; http://www.tigr.org) revealed no promoter consensus or other obvious regulatory sequences, indicating the genes described here are part of an operon structure. The xerD genes of both strains KNR_7/87 and R6 were identical.

The AT content of xerD with 57% resembles that of structural genes of *S. pneumoniae* (Baltz *et al*.,

---

**Figure 1. Genetic organization of the xerD locus in different bacterial species.** The xerD genes are indicated by black arrows. Homologues of *S. pneumoniae* orfCDE are marked gray. The *B. subtilis* xerD homologue is named ripX (Sciochetti *et al*., 1999).
The xerD Locus in Streptococcus mitis
Using oligonucleotides primers specific for xerD, PCR analysis revealed the presence of the xerD gene in two isolates of the closely related species S. mitis. Since gene transfer events between S. mitis and S. pneumoniae are common, we investigated the relatedness of xerD between the two species. In S. mitis 476, the xerD gene was located within the same operon organization as in S. pneumoniae (not shown), but the DNA sequence differed by 9% and the deduced peptide sequence by 7%, and the adjacent genes were also different (10–5% on the DNA level, and 11 to 2% on the amino acid level). The region most likely reflects the difference between the two genomes.

The xerD Locus in Other Bacteria
XerD homologues are present in the majority of the bacterial species and form a distinct subfamily of site specific tyrosine recombinases (Recchia and Sherratt 1999; Esposito and Scocca 1997). The organization of the putative S. pneumoniae xerD locus with the three downstream orfCDE is conserved in S. mitis, S. pyogenes, S. mutans and Enterococcus faecalis (Figure 1). Other Gram positive cocci of low GC content investigated here all contained orfCDE in close vicinity of a xerD homologue. In Bacillus subtilis, orfCDE homologues were present but not linked to the xerD (ripX) gene (Figure 1). Interestingly, Mycobacterium leprae contained the parA gene important in partitioning of circular plasmids as part of the orfCDE operon, while Streptomyces coelicolor, also a high GC Gram positive bacterium which has maintained the parA-orfCDE organization (Figure 1). Homologues of orfC and orfD were also detected in the genomes of Synechocystis spp. (str1577), the spirochaete Treponema pallidum (AE001243, AE001223), and of the three archaea Archaeoglobus fulgidus (AF1580, AF1559); Pyrococcus horikoshii (PHCY016, PHBB010); Methanococcus jannaschii (MJ1134), but not in Gram negative bacteria.

The Putative Streptococcal xerD Genes Encode Unusual XerD
The deduced XerD proteins of S. pneumoniae, S. mitis, S. mutans and S. pyogenes all contained the S(T)RQ motif at the N-terminus of the DNA recognition helix which is characteristic for the XerD subfamily of site-specific recombinases. However, alignment of the C-terminal domain of XerD proteins revealed striking differences in the streptococcal proteins at positions that have been described as extremely conserved among all members of the entire Int family of recombinases, located in the conserved regions named box I and box II (see (Nunes-Düby et al., 1998)) (Figure 2).

In box I (L119 – D142 in the putative S. pneumoniae XerD), an invariable R is replaced by L129, (or T respectively S in S. pyogenes and S. mutans). In box II (S208 – R246), R213 and the catalytic Y245 are conserved, but both histidine residues common to most recombinases are replaced by Q210 and L236 in case of the pneumococcal protein. Furthermore, the strictly conserved K172 which is flanked by G or M in the XerD subgroup (see (Nunes-Düby et al., 1998)) is part of a SKA triad in S. pneumoniae (QKA in S. mitis, and TKK in S. pyogenes). Moreover, the last 14 to 19 amino acids that are part of a distinct three dimensional arrangement in the four recombinases of known structure are missing in the streptococcal proteins, and the catalytic Y is the penultimate or the third last amino acid (Figure 2).

The putative streptococcal XerD proteins differed from the E. coli XerD by 77 to 75% on the amino acid level, whereas XerD of other Gram positive bacteria including E. faecalis and S. aureus differed only by 56 to 59% from the E. coli protein (Table 2). The phylogenetic tree shown in Figure 3 confirms the existence of two distinct groups within the XerD family, one of which is represented by the streptococcal proteins.

Putative XerC Proteins in Streptococcus spp.
The function of the E. coli XerD protein depends on the presence of at least one other protein, the related XerC recombinase (Blakely et al., 1993). A database search in the pneumococcal genome sequence revealed the presence of another four members of site-specific recombinases of the integrase family. Recently, we have carried out a genomic comparison between 20 different S. pneumoniae strains representing 14 distinct clonal groups, using an oligonucleotide microarray representing the genome of a type 4 isolate (Hakenbeck et al., 2001). Only one of the xerC related genes was conserved in all 20 S. pneumoniae strains and was therefore considered to represent the xerC homologue, whereas the other three appeared absent in at least one of the strains considering the very low hybridization signals obtained (not shown). The putative XerC corresponds to Xer1 of three Xer proteins described recently by Chalker et al. (Chalker et al., 2000). Also, the relatedness between the S. pneumoniae XerC and those of closely related Streptococcus spp. was high (> 78% identical amino acids) (Figure 4), forming a closely related cluster compared to XerC proteins of other bacteria in contrast to the previously described XerC2 and XerC3 peptides (Figure 5).

Unlike the situation concerning the putative xerD gene, the chromosomal location of xerC was not conserved when compared to the closely related S. mutans, S. equi and S. pyogenes (Figure 6).

Mutational Analysis of the S. pneumoniae xerD Region
The xerD gene was inactivated by insertion duplication mutagenesis. An internal gene fragment was cloned into the vector pJDC9 that is not able to replicate in S. pneumoniae but contains a selectable
Figure 2. Comparison of streptococcal Xer D with those of other bacterial species and the Cre recombinase of Phage P1. The C-terminal domain of XerD (XD) was aligned using the Clustal X 1.8. Dashed lines correspond to the structurally conserved integrase domains box I and box II and the XerD DNA recognition α-helix (αRH). Patch I and patch II refer to conserved hydrophobic regions among integrase family members as described (Nunes-Duby et al., 1998). The invariant amino acids of the integrase family and the XerD specific motif (SRQ/TRQ) are shaded black, Lys and Arg residues in box II of the streptococcal proteins are shaded gray. The secondary structure assignment of the helices (B, C, D, E, F, G, H) of the E. Coli XerD, and those of the P1 Cre protein (H, I, F, J, K, L, M, N) are placed according to the published structures (Boo et al., 1997; Subramanya et al., 1997). Ec, Escherichia coli; Ml, Mycobacterium leprae; Sm, Streptococcus mitis; Sp, Streptococcus pneumoniae; Smu, Streptococcus mutans; Spy, Streptococcus pyogenes; P1Cre, P1 phage Cre protein.
erythromycin resistance gene, resulting in plasmid pJPR1 (Chen and Morrison 1988). Transformation of S. pneumoniae R6 with pJPR1 and selection with erythromycin resulted in the expression of a putative truncated XerD-peptide of 210 residues that lacks the four amino acids in box II implicated in catalysis (H-R-H-Y in the integrase family, but Q-R-L-Y in the putative S. pneumoniae protein, see Figure 2). The xerD mutants grew with a considerably longer generation time compared to the parental R6 strain of 48 min versus 30 min (Figure 7A). They had a tendency to grow in longer chains of misshaped cells especially at the end of the exponential growth phase (Figure 7B). In contrast, insertional inactivation of the orfC-gene resulted in a strain that grew with almost the same generation time (33 min) as the R6 strain, demonstrating that the phenotype of the xerD mutant is not due to a polar effect on downstream genes (Figure 7A).

Discussion

We present here members of the XerD subfamily of site specific recombinases with unusual features when compared to the over 100 known related proteins of the λ Int family of tyrosine recombinases (Nunes-Düby et al., 1998). The extreme C-terminal amino acids are lacking, and they have been shown to be important for inter-subunit interactions in E. coli XerD using mutants with various C-terminal deletions (Spiers and Sherratt 1997); this has also been confirmed for the Cre structure (Guo et al., 1997). Furthermore, the four positively charged residues RRRH considered to be strictly conserved in the integrase family, appear to be crucial for DNA interaction and are represented in the putative S. pneumoniae XerD by the residues LQRL. Both Arg residues are invariably conserved in the 105 members of the integrase family investigated in detail, and in very few cases one of the His residues is represented by a functionally equivalent amino acid, Asn, Arg, Lys, Tyr or Trp (Nunes-Düby et al., 1998). In Cre where these four residues are represented by R-H-R-W they are coordinating non-bridging oxygen atoms by five hydrogen bonds in the Cre-loxA complex (Guo et al., 1997). This proton-cradle clearly cannot be maintained in case of the streptococcal proteins with two leucine replacements. The particular amino acid makeup of the putative streptococcal XerD indicates either distinct kinetic properties of the overall recombination reaction in the synapse, or other amino acids as part of the catalytic center. It is noticeable that in all cases, the catalytic Tyr residue is preceded by a Lys, and that box II is unusually rich in positively charged residues (Figure 2), potential important candidates for the interaction with DNA.

These special XerD proteins are found exclusively in Streptococcus spp. in a highly conserved gene cluster containing at least one gene orfA and three downstream genes orfCDE, some of which may be involved in nucleotide modification. The chromosomal organization clearly identifies the gene as xerD homologue, in addition to the SRQ triad of the deduced protein which is typical for XerD proteins and occurs at the N-terminus of the DNA recognition helix. It is curious that even in E. faecalis where the chromosomal context of xerD is identical to that of the streptococcal genes, the deduced XerD protein clearly belongs to a different XerD subfamily containing all the recombinase-specific conserved amino acid residues.

The orfCDE homologues do not necessarily form an operon in other species but they are present in most Gram positive bacteria (Figure 1). In the high GC Gram positive Streptomyces coelicolor and Mycobacterium leprae, orfC is located downstream of parA which is required for partitioning of circular plasmids. The S. coelicolor DNA sequence of the unfinished genome does not contain a xerD homologue similar to Borrelia burgdorferi, both organisms with a linear genome where chromosome dimer resolution is not required (Recchia and Sherratt 1999).

The streptococci all possess XerC homologues, the proposed partner in Xer site-specific recombination events. In the deduced XerC proteins all conserved amino acid residues of the λ integrase family are present. The xerC gene was detected in a genomic comparison using microarray technology involving a sample of 20 strains representing 14 distinct clonal groups, in contrast to another three xerC related genes (Hakenbeck et al., 2001). This gene is identical to the one described as xer1 recently (Chalker et al., 2000). In agreement with an important role for the organism as expected for XerC, loss of function mutants in xer1 showed significant increase in chain length and a severe
Figure 4. Comparison of \textit{S. pneumoniae} XerC homologues with those of other bacterial species. The proteins were aligned using Clustal X 1.8. The invariant amino acids of the integrase family and the N-terminal triad of the DNA recognition helix are shaded black. Se, \textit{S. equi}; Spy, \textit{S. pyogenes}; Smu, \textit{S. mutans}; Sp, \textit{S. pneumoniae}; Li, \textit{Lactobacillus leichmanii}; E1, \textit{E. faecalis}; Bsub, \textit{B. subtilis}; Sa, \textit{S. aureus}; Mi, \textit{M. leprae}; Ec, \textit{E. coli}. 

\begin{verbatim}
-----boxI---------- patchII patchIII

EfXC 128: EQFLDLRNLQALLELYTGIGYVCSCANLTVADDVFAQ---------- SVILVHSGGKDYVFPGFSQDALKDYLENRAGLMTKYYKHHYVFVGHSG
BsubXC 129: QSPGMRQDLALLELYTGIGYVCSCSITINDLFLM---------- DTIVYHGFQKRQFYIFGSYREALKVMNSGRQCLL.MKAEHPHLLFLVNRGQ
SaXC 42: DTSDKLNRDLALLELYTGIGYVCNAVNLQKQIDFYA---------- NGVTVLQGKEKERVFPGFGAQCRIELEHPFQOCS---HDFLIVNMRK
Ml XC 157: GDLALRDRLIVEMLITGIGYVCSELGLIDIDDDVTRH---------- RLRVRLGQKNQRTAFPFGVASADLRLGWDGPRALVMTB---SGFALLLGRG
EoXC 129: -DPLAVRDRAMEVYGAQGLSELVGLLDKHLLLES---------- -GWIWMMGKEKRLFKPRIMAIWAEWHLDELRDLSGD---DALFLSKLG
SeXC 190: FRKKERDLIAILLASGIDLSEAVNLDKVDNLNM---------- MLVEYRTQGKMDSVWNAFAPFKHEAYLSVWYKDYQAEK---QDUAPFLIA1YK
SpyXC 190: FRKKERDLIAILLASGIDLSEAVNLDKVDNLNM---------- MIIEVIRQGKDGNVAGPAGKFQELYSLAVQRYKAEK---QDAPFLLTEYR
SmuXC 190: FRKKERDLIAILLASGIDLSEAVNLDKVDNLNM---------- MIIEVIRQGKDGNVAGPAGKFQELYSLAVQRYKAEK---QDAPFLLTEYR
SpXC1 190: FNNKRDLIAILLASGIDLSEAVNLDKVDNLNM---------- MVIDTRGCRKDSVNAFAFPKQELNYLTAIRNQYNTKEK---TDAPFLLSEYR
SpXC2 164: ----LRLDLAIIALLASGIDLSEAVNLDKVDNLNM---------- RECVVPEQIJKERPVYFDARTKIHRLRNLYDRLD---S-----IFAPFVILYVG
SpXC3 71: ----DGILKWFVWFLGATQVSEILKLVHEVIEG---------- --YDFIYQQKIRRLYIFKILANISCLESLWEBNR---R-----SYGLFLMKFN
SpXC4 202: SNEYNLPDVVLKTTYLATSCEPSEALALEWSDIDLESGVI3NKMLNYQBEINSFSDKSSHDYDIDPKDATLLLKQYKQRQIQSWKLG---RSTVVFSPFT

-----\textbf{QRH}----- boxII

EfXC: -EQUFTTIYEILYVNLQIKK-SLSAENITHFANTHLLLNGAMRMTVQELSGILANLTTQ-IQAHTKXRSKQNYRTFHPA* : 299
BsubXC: -GPIPLGACHIRLLSGLYQAKLASLHTFHLLNGADRLGQELLSQMLNLESTQ-IRTHVSKBLRTNTMSHPAFFKKN* : 304
SaXC: -EAILTKLVRNLQIKK-SLSAENITHFANTHLLLNGAMRMTVQELSGILANLTTQ-IQAHTKXRSKQNYRTFHPA* : 215
Ml XC: -RRRVRKQARTVQVHTVAV-VAQADMPGHLAHSATHLLEGGDLRVQELLSSLATQ-IQHTANVRVLRVSHQDHAPRK* : 326
EoXC: -KTRSCARVQKREASWQST-KQILNLSVHFIQKLSFPHAMLSSQGIDLCVQVQELLSQMLNLESTQ-IQAHTKXRSKQNYRTFHPA* : 298
SeXC: -GLPNRASDEIKMKYQK*-EDFKIRTVETGKLTHLRYDTASGQVLQHSLQGASSTQT*-LTHVNDXKNALDNL* : 355
SpyXC: -GVPVRKASDEIKMKYQK*-EDFKIRTVETGKLTHLRYDTASGQVLQHSLQGASSTQT*-LTHVNDXKNALDNL* : 356
SmuXC: -GVPVRKASDEIKMKYQK*-EDFKIRTVETGKLTHLRYDTASGQVLQHSLQGASSTQT*-LTHVNDXKNALDNL* : 356
SpXC1: -GVPVRKASDEIKMKYQK*-EDFKIRTVETGKLTHLRYDTASGQVLQHSLQGASSTQT*-LTHVNDXKNALDNL* : 356
SpXC2: -QECRKTWIATIKEMKVQYK**-AAFFHETLPKHTLRITVESLAYQKDVQMVADQQLLGTKTSATD-ETHYDQKQRAALNSAKQPKSSDD* : 308
SpXC3: -EKTCRKTWIATIKEMKVQYK**-AAFFHETLPKHTLRITVESLAYQKDVQMVADQQLLGTKTSATD-ETHYDQKQRAALNSAKQPKSSDD* : 321
SpXC4: -EKKKCRKTWIATIKEMKVQYK**-AAFFHETLPKHTLRITVESLAYQKDVQMVADQQLLGTKTSATD-ETHYDQKQRAALNSAKQPKSSDD* : 229

\end{verbatim}
reduction in growth rate, a phenotype that resembles that seen in the xerD mutants described in the present publication. Moreover, xer1 mutants were highly attenuated in a murine respiratory tract infection model (Chalker et al., 2000).

Evidence that the other three xerC genes are of phage origin or related to transposable elements comes from the genomic comparison with other pneumococcal strains (Hakenbeck et al., 2001). The xerC2 and xerC3 genes showed extremely low hybridization signals in one respectively five out of 14 different clones, and the xerC4 gene identified in the present study gave a uniformly low signal in all other except the reference strain, strongly suggesting that they are not part of the pneumococcal chromosome in general. All of them were located within gene clusters that showed similarly reduced signals, indicating that they may be part of defective prophages or other mobile genetic elements that are frequently found in S. pneumoniae (Baltz et al., 1998). The relatedness of XerC2 and XerC3 to an enzyme of transposon origin has been noted (Chalker et al., 2000; Severina et al., 1999).

It is curious that all streptococcal species that contain the unusual XerD proteins are either naturally transformable such as S. mitis and S. mutans, and S. pyogenes contains at least the “late competence genes” important for recombination events (Hàvarstein 1998). We have observed a 10- to 50-fold reduction in transformation efficiency in the S. pneumoniae xerD mutants, and addition of the competence stimulating peptide CSP (Hàvarstein et al., 1995) did not restore this phenotype, indicating that recombination is affected rather than competence development per se (not shown). E. coli XerD mutants, lacking a functional active site, have been shown to mediate certain recombination reactions in partnership with XerC (Spiers and Sherratt 1997; Arciszewska et al., 2000), and it is possible that the Streptococcus spp. putative XerD proteins are also able to function in complex with XerC. In vitro studies using purified proteins are required to confirm the function of the streptococcal XerD enzymes.

**Experimental Procedures**

**Bacterial Strains, Plasmids and Culture Conditions**
The bacterial strains and plasmids used are listed in table 1. E. coli strains were grown in Luria Bertani broth at 37°C with aeration. Streptococci were grown in C medium (Lacks 1970) supplemented with 0.1% yeast extract at 37°C without aeration.

**Insertion Duplication Mutagenesis**
Internal gene fragments of xerD and orfC were amplified by PCR and cloned first into pCR™ (Invitrogen, Groningen, The Netherlands). The xerD fragment was amplified using the oligonucleotide primers 5'-CGGCCCTGTAACCAATTTC and 5'-AGACTTGAGCTGATAAGG covering codons 99 to 210, and for the orfC fragment, 5'-ACGATGTGCCCATTACGG and 5'-TCAATGGTCGTCTTGTC (codons 32 to 139).
were used. Inserts were isolated after EcoRI restriction of recombinant plasmids and cloned into pJDC9 to create pJPRDI and pJPRD7, respectively. The plasmid pJDC9 cannot replicate in S. pneumoniae, but contains as selectable erythromycin resistance marker (Chen and Morrison 1988). After transformation into S. pneumoniae R6, erythromycin resistant colonies were selected at a concentration of 1 μg/ml. Six to eight independent transformants were isolated and insertion of the plasmid into the target genes verified by PCR analysis.

**DNA Sequencing**

DNA was amplified by PCR with Goldstar Taq Polymerase (Eurogentec, Seraing, Belgium) using 30 cycles and the following conditions: 30 sec denaturation at 96°C; 30 sec annealing at 50–54°C depending on the oligonucleotides used, 60 sec extension at 72°C; followed by a 10 min delay period after the last cycle. DNA sequences of plasmids were determined with the T7™ sequencing kit from Pharmacia (Uppsala, Sweden); direct sequencing of PCR products was performed with the Sequenase™ PCR Sequencing Kit (USB, Cleveland Ohio, USA). Cycle Sequencing was performed with the ABI PRISM™ Dye Terminator Ready Reaction Kit (Perkin Elmer, Foster City, USA).

**Database Searching**

Sequence database searches were carried out using BLAST (Altschul et al., 1990) and FASTA (Pearson and Lipman 1998). Sequence alignments were constructed using the GCG 9 package or Clustal X 1.8 and Treeview 1.5.2. Sequences for microbial genomes were obtained from The Institute of Genomic Research website at http://www.tigr.org..

**Nucleotide Accession Number**

The xerD sequences have been deposited in the EMBL Nucleotide Sequence Database under the following accession numbers:

- igA1, Y10286-1
- rnh, Z99112-76
- lplA, AP001509-121
- t1M, L18759-2
- t1S, NP_070538
- t1R, AE000505
- pepX, M35865-2
- glnA, AAC44800
- hss, AAA97245
- t1S*, AAG18734
- t1S**, S06097
- t1M*, Q47282
- t1R*, Q07736
- glrA, AAC17172
- tnr5, P27451
- fh, Q54431
- guaA, AA15805
- gntR, AA15462
- mucA, AA154059
- hk, CAB54567
- gntR*, AAK34062
- mucA*, AAC98440
- citC, P77390
- gid, P39815
- hslV, P39070
- gid**, BAB06185
- hslU, BAB42349
- rpsB, P81289
- tsf, P02997
- pyrB, X78999

**Figure 6. Chromosomal location and flanking regions of the xerC homologues in different Gram-positive bacteria.** Genes were designated according to the best matching database entry. Digits (1–7) mark ORFs with no significant homology to known proteins. "t" denotes type 1 restriction modification system; R, S, M correspond to the respective subunits. Accession numbers of the genes depicted in Fig. 5 are as follows: igA1, Y10286-1; rnh, Z99112-76; lplA, AP001509-121; t1M, L18759-2; t1S, NP_070538; t1R, AE000505; pepX, M35865-2; glnA, AAC44800; hss, AAA97245; t1S*, AAG18734; t1S**, S06097; t1M*, Q47282; t1R*, Q07736; glrA, AAC17172; tnr5, P27451; fh, Q54431; guaA, AA15805; gntR, AA15462; mucA, AA154059; hk, CAB54567; gntR*, AAK34062; mucA*, AAC98440; citC, P77390; gid, P39815; hslV, P39070; gid**, BAB06185; hslU, BAB42349; rpsB, P81289; tsf, P02997; pyrB, X78999.
Table 1. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. pneumoniae R6</td>
<td>Δ cap A-H</td>
<td>(Avery et al., 1944)</td>
</tr>
<tr>
<td>S. mitis 476</td>
<td>β-lactam resistant clinical isolate</td>
<td>(Reichmann et al., 1997)</td>
</tr>
<tr>
<td>E. coli Inv3F'</td>
<td>endA1, recA1, F, lacZ Δ M15</td>
<td>Invitrogen</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC18/19 cloning vector</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>(Yanish-Perron et al., 1985)</td>
</tr>
<tr>
<td>pJDC9 cloning vector</td>
<td>Ery&lt;sup&gt;R&lt;/sup&gt;</td>
<td>(Chen and Morrison 1988)</td>
</tr>
<tr>
<td>pCR&lt;sup&gt;TM&lt;/sup&gt; cloning vector for PCR products</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pUP138</td>
<td>pUC18 containing the Sal I, KpnI 2.8 kb xer D fragment</td>
<td>This work</td>
</tr>
<tr>
<td>pJPR1</td>
<td>pJDC9 containing a 440 bp internal fragment of the pneumococcal xerD</td>
<td>This work</td>
</tr>
<tr>
<td>pJPR7</td>
<td>pJDC9 containing an 317 bp internal fragment of the pneumococcal orf C</td>
<td>This work</td>
</tr>
</tbody>
</table>

Figure 7. Growth and morphology of xerD mutant. A. Mutant and wild type cells were grown in C-medium: xerD::pJPRDI; orfC::pJPRD7; R6 (wild type). Growth was monitored by nephelometry (N). B. Light micrographs of the xerD mutant strain (left) and wild type R6 (right).
Table 2. Amino acid identities (similarities) between XerD Proteins form different species (%).

<table>
<thead>
<tr>
<th>Species</th>
<th>S. pneumoniae</th>
<th>S. mitis</th>
<th>S. pyogenes</th>
<th>S. mutans</th>
<th>E. faecalis</th>
<th>S. aureus</th>
<th>B. subtilis</th>
<th>M. leprae</th>
</tr>
</thead>
<tbody>
<tr>
<td>accession numbers</td>
<td>110</td>
<td>56</td>
<td>43</td>
<td>37</td>
<td>30</td>
<td>22</td>
<td>28</td>
<td>24</td>
</tr>
</tbody>
</table>

Acknowledgements

This work was supported by the Stiftung Rheinland Pfalz für Innovation.

References


