Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage

Rubens López *, Ernesto García

Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu, 9, 28040 Madrid, Spain

Received 11 March 2004; received in revised form 19 May 2004; accepted 25 May 2004

First published online 7 June 2004

Abstract

Streptococcus pneumoniae has re-emerged as a major cause of morbidity and mortality throughout the world and its continuous increase in antimicrobial resistance is rapidly becoming a leading cause of concern for public health. This review is focussed on the analysis of recent insights on the study of capsular polysaccharide biosynthesis, and cell wall (murein) hydrolases, two fundamental pneumococcal virulence factors. Besides, we have also re-evaluated the molecular biology of the pneumococcal phage, their possible role in pathogenicity and in the shaping of natural populations of S. pneumoniae. Precise knowledge of the topics reviewed here should facilitate the rationale to move towards the design of alternative ways to combat pneumococcal disease.

Keywords: Streptococcus pneumoniae; Capsular polysaccharide; Cell wall hydrolases; Phage; Virulence factors

Contents

1. Introduction ......................................................... 554
2. Capsular polysaccharide ................................................. 555
   2.1. Genetic bases of CP biosynthesis in pneumococcus .............. 555
   2.2. Mechanisms of regulation and transport of the CP ............. 556
   2.3. Genetic variability in pneumococcal capsular genes .......... 557
3. Cell wall hydrolases .................................................... 558
   3.1. General characteristics ........................................ 558
   3.2. CWHs of the pneumococcus .................................. 558
   3.3. Domain organization of the pneumococcal CWHs: choline-binding proteins .... 560
   3.4. The crystal structure of the choline-binding domain (C-LytA) ................ 561
   3.5. Subcellular localization and biological role of CWHs ........... 562
   3.6. CWHs and genetic transformation ............................ 564
   3.7. CWHs and pathogenesis ..................................... 565
4. Bacteriophage of pneumococcus ........................................... 566
   4.1. General features of phage Dp-1, Cp-1, MM1, and EJ-1 .......... 566

*Corresponding author. Tel.: +34-91-8373112; fax: +34-91-5360432.
E-mail address: ruben@cib.csic.es (R. López).

Abbreviations: aa, amino acid(s); C-, carboxy; ChBD, choline-binding domain; ChBP, choline-binding protein; ChBR, choline-binding repeat; C-LytA, the C-terminal moiety of LytA; CSMtase, cytosine 5-methyltrasferase; CP, capsular polysaccharide; CSP, competence-stimulating peptide; CWH, cell wall hydrolase; IS, insertion sequence; N-, amino-; nt, nucleotide(s); SP, signal peptide; TA, teichoic acid(s); ORF, open reading frame; UDP-Glc, uridine diphosphoglucose.
1. Introduction

Infectious diseases are currently the third cause of death in the United States and the leading cause worldwide [1]. *Streptococcus pneumoniae* (the pneumococcus) is a major Gram-positive human pathogen with a historical background. At the beginning of the past century, ‘pneumococcal pneumonia was the leading cause of death’ and, as mentioned by McCarty [2], research directed against this specific medical problem resulted in a fundamental discovery for Molecular Biology. Early observations by application of specific genetic information were first linked [3,4]. Currently, the pneumococcus is the main cause of pneumonia, meningitis, and bloodstream infections in the elderly, the young, and the immunocompromised, and responsible for middle ear infections in children. In the United States more than 500,000 cases of pneumococcal pneumonia occur each year, causing almost 40,000 deaths in adults and children [5]. It is at present the principal cause of bacterial meningitis since the introduction of *Haemophilus influenzae* type Ib conjugate vaccine (6000 cases per year with a 30% mortality rate in the United States). Surveillance of invasive pneumococcal disease in the United States has revealed that the total disease incidence amounted up to 72–103 per 10^5 inhabitants/year, and has been reported to be the cause of disease in 10–20 per 10^5 inhabitants/year for four European countries (United Kingdom, Spain, Denmark, and Finland) [6]. However, the United States surveillance was not limited to hospitalized patients, as is the case of Europe, and therefore, the European data are certainly an underestimation of the incidence.

*S. pneumoniae* is a human commensal that localizes at the upper respiratory tract of individuals. The large scale prevalence of drug-resistant pneumococci and the limited therapeutic options to successfully treat these resistant organisms underline the need for alternative solutions to limit the perturbation of the normal microbiota with unpredictable long-term consequences. The global importance of *S. pneumoniae* as the cause of illness, sequelae, and death, as well as the emergence of drug resistance that is making these infections more difficult to treat, justify the interest of many groups of scientists to solve many pending questions concerning pneumococcal infections, like colonization and spread, control or eradication of the carrier state, genetic versatility of the genome (including mutation rate and its influence on increasing virulence), new diagnostic tests and improvement of the available vaccines as an alternative to multiple drug resistance, DNA transfer in nature, control of capsule production, etc. There is a limited efficacy of the available vaccines and to combat pneumococcal infections through generalized use of antibiotics looks unrealistic to the long-range because of the genetic plasticity of the pneumococcus that results in either capsular type shifting or in the rapid appearance and spreading of antibiotic-resistant isolates and antibiotic resistance ‘determinants’ [7].

This review focuses in the recent efforts to achieve a better knowledge on two factors that play fundamental roles in virulence as a way to provide a discussion on novel approaches to deal with the clinical problem that *S. pneumoniae* currently represents, that is to say: (1) the molecular organization and regulation of the capsular genes coding for the major virulence trait of *S. pneumoniae*; (2) the murein hydrolases, a family of choline-binding proteins (ChBPs), i.e. proteins that share a common choline-binding domain (ChBD), offer a noticeable example of protein versatility, and are potential new therapeutic targets. We also deal with the isolation and characterization of pneumococcal phage to study, on one hand, their putative impact in virulence, taking into account the abundant presence of temperate phage among clinical isolates (up to 70%), and, on the other hand, the use of phage products, as a complementary (or alternative) way to the use of antibiotics to fight pneumococcal infections.
2. Capsular polysaccharide

The pneumococcal capsules are polysaccharides excreted outside the cell and are composed, most times, of repeating units of simple sugars and remain attached, possibly in a covalent form, to the outer surface of the bacterium. Capsules are usually associated with increased virulence as they may function as adhesins, recognition molecules and/or by favouring the camouflage of the parasite against the host immune response. It should be emphasized that early studies on the capsular polysaccharides (CPs) of \textit{S. pneumoniae} allowed the identification of polysaccharides as immunogens [8]. Pneumococcal CPs appear to be dispensable for cell viability since unencapsulated mutants are not impaired in growth, at least, under laboratory conditions. However, whereas the pneumococcal capsule is absolutely required for virulence [9], in other streptococci the presence of a capsule clearly increases the pathogenicity of a particular isolate. Pneumococcal CPs display a complex composition, are highly variable in chemical structure, and frequently, also contain some nonsaccharidic components [10] explaining the large antigenic variability characteristic of these polymers.

2.1. Genetic bases of CP biosynthesis in pneumococcus

Early observations concluded that the genes coding for the biosynthesis of the pneumococcal capsule were clustered together [11], and recent achievements have confirmed this proposal. To date, the complete nucleotide (nt) sequence of 16 different capsular gene clusters (designated as \textit{cap} or \textit{cps}) have been published. Moreover, efforts to sequence the 90 pneumococcal serotypes already known [12] are currently being done at the Sanger Institute (\texttt{http://www.sanger.ac.uk/Projects/S_pneumoniae/CPs}). From the existing data there are three different organizational models of the capsular gene cluster in \textit{S. pneumoniae} (for a review, see [13]):

1) The most common capsular gene cluster organization corresponds to that of types 1, 2, 4, 6B, 8, 9V, 14, 18C, 19F, 19A, 19B, 19C, 23F, and 33F. The \textit{caplcs} gene cluster is located between \textit{dexB} and \textit{alilA}, two genes that do not participate in capsular biosynthesis (Fig. 1). In all these cases, a functional promoter is located immediately upstream of the gene cluster [14,15], and the first four open reading frames (ORFs) of the \textit{caplcs} operon are well conserved among types although only the first ORF is virtually identical in all of the cases analyzed so far. In spite of the sequence conservation of the two first ORFs among serotypes, these genes show enough polymorphism to allow serotyping of \textit{S. pneumoniae} isolates by PCR-based methods [16–18]. Except in the cases of types 3 and 37 (see below), CpsB, CpsC, and CpsD are essential for encapsulation. CpsB is a phosphotyrosine-protein phosphatase, CpsC is required for CpsD tyrosine phosphorylation, and CpsD is an autophosphorylating protein-tyrosine kinase [19]. The genes located downstream of the fourth gene of the cluster are considered serotype-specific although some of them may be shared among various types. Sequence conservation among serotype-specific genes is particularly evident in the case of group 19 pneumococci [20] though the fifth gene of the \textit{caplcs} locus, encoding a glucosyltransferase catalyzing the transfer of glucose-1-phosphate to a lipid carrier [21], is conserved at least in types 2, 4, 14, 19F/A/B/C, 23F, and 33F. Noticeably, the genes \textit{cap1K}, \textit{cps2K}, and \textit{cap3A}, encoding otherwise biochemically identical enzymes, namely a uridine diphosphoglucose (UDP-Glc) dehydrogenase, were clearly divergent in types 1, 2, and 3, respectively, although \textit{cap8L} that codes for the type 8 UDP-Glc dehydrogenase was 92% identical to \textit{cps2K} [22,23].

2) Type 3 is exceptional in that the four initial ORFs of the capsular operon are not involved in CP biosynthesis and are not expressed [24] (Fig. 1). A functional promoter is located immediately upstream of the first gene of the operon (\textit{cap3A}). In accordance with the simple chemical structure of the type 3 repeating unit (cellobiuronic acid units connected in a \(\beta(1 \rightarrow 3)\) linkage), only three complete genes were found in the capsular operon (Fig. 1). Moreover, the third gene (\textit{cap3C}; also referred to as \textit{cps3U}), is not required for CP biosynthesis since the biochemical function of its product (a UDP-Glc pyrophosphorylase) is compensated by that of the \textit{galU} gene located far apart in the \textit{S. pneumoniae} chromosome [25]. The GalU enzyme has been shown to be essential for CP synthesis. This enzyme is required for the interconversion of UDP-Glc and UDP-galactose by way of the Leloir pathway [26]. Prokaryotic UDP-Glc pyrophosphorylases are well conserved and although UDP-Glc pyrophosphorylases are also present in eukaryotes, these enzymes are completely unrelated to their prokaryotic counterparts [25, and references therein]. This finding suggests the possibility that putative inhibitors of the bacterial enzymes would not be harmful for the host. A similar requirement to that of GalU for CP biosynthesis has been recently reported for the Pgm phosphoglucomutase [27,28].

Type 3 pneumococci also differ from pneumococci of other serotypes in that only one enzyme, the polysaccharide synthase Cap3B, is needed for CP biosynthesis provided that the host strain was able to synthesize UDP-glucuronic acid (UDP-Glc is a universal precursor for all the pneumococcal capsules) [29]. In addition, the type 3 CP is polymerized through a processive mechanism by the addition of glucose and glucuronic acid to the nonreducing termini of pre-existing polysaccharide primers [30]. Moreover, it has been found that termination and release of the type 3 polysaccharide chain occurs in \textit{S. pneumoniae} via a specific enzymatic
mechanism, when the concentration of either of the sugar nt precursors (UDP-Glc and UDP-glucuronic acid) drops below a critical concentration [31]. Besides, experiments carried out using membranes of Escherichia coli containing the pneumococcal synthase strongly suggested that this enzyme initiates polysaccharide synthesis on a lipid primer [32].

(3) The most peculiar case among the pneumococcal capsular genes is provided by type 37 isolates. It has been found that these strains are genetically binary strains, this is, they contain a cap37 locus virtually identical to that of type 33F strains [33,34] but several mutations inactivate some of the ORFs and, consequently, this locus is actually silent. It was also demonstrated that type 37 capsulation is due to the presence of a single copy of a gene (tts) located far apart from the cap cluster (Fig. 1). The Tts synthase contains several motifs known to be characteristic of cellulose synthases and other glucosyltransferases. Cloning and expression of the tts gene in heterologous hosts, namely Streptococcus oralis, Streptococcus gordonii, and Bacillus subtilis demonstrated that tts is the only gene required for CP biosynthesis. Further works also showed that type 37 pneumococci isolated 60 years apart and in two different continents are very closely related and constitute a clonal complex [35]. In addition, the genetic relatedness between type 37 and type 33F pneumococci has been recently confirmed by sequence determination of the polymorphic locus galU [36]. Type 37 pneumococci offer a peculiar example of a genetically binary strain of potential clinical relevance as an alternative to the well-studied phenomenon of capsular shifting [37] to introduce variability in the CP biosynthetic pathways of S. pneumoniae [13].

2.2. Mechanisms of regulation and transport of the CP

Production of most CPs is achieved through the formation of a lipid-linked repeat unit that is synthesized on the intracellular face of the membrane, exported to the surface, and polymerized. Deletion of either cps2A, 2B, 2C, or 2D genes in the serotype 2 gene cluster, does not affect the transfer of the capsular polysaccharide to the cell wall [38]. Interestingly, the
correlation between tyrosine phosphorylation of CpsD and CP production is a matter of current debate. It was first reported that CpsD acts as a negative regulator of capsule biosynthesis [39]. On the other hand, Weiser et al. [40] observed a positive correlation between capsule production and CpsD phosphorylation in clinical pneumococcal isolates. Using the *S. pneumoniae* type 19F capsule expressed in the laboratory strain Rx1, Morona et al. [41] recently reported that pneumococcal mutants containing zero, one, or two tyrosine residues present in the carboxy-(C-) terminal (YGX)₄ repeat domain of CpsD exhibited a mucoid appearance although it was not due to overproduction of CPs, as previously proposed but, possibly, to physical and rheological differences between wild-type CP and mucoid CP. In addition, mutants that produced C-terminally truncated CpsD exhibited less CP production than the parent strain. On the other hand, recent analyses using two different strains of serotype 2 have concluded that there is a positive correlation between phosphorylation of CpsD and capsular production in *S. pneumoniae* [38]. Based on the observation that a different response in CP production was observed between D39 strain, a clinical virulent type 2 strain, and Rx1, a laboratory passaged D39 derivative containing multiple uncharacterized mutations, these authors concluded that a factor outside the capsular locus, missing in the Rx1 background, also controls CP synthesis. Whether the factor missing in Rx1 interacts with the phosphotyrosine system is currently unknown.

The pneumococcus undergoes spontaneous phase variations between an opaque and a transparent colony form, the former being more virulent in a murine sepsis model. The opaque phenotype appears to be associated with increased production of CP [42]. Unfortunately, the genetic mechanisms that contribute to each phenotype have not been described. Upregulation of pneumococcal CP biosynthesis in the human setting in response to environmental conditions has been noted [43]. It was observed that the expression of cps2A increased fourfold following intraperitoneal injection in a murine sepsis model. Interestingly, a pneumococcal gene (*regM*) putatively encoding a catabolite repressor protein has been implicated in the transcriptional control of the *cap/cps* locus [44]. It was found that *regM* mutants, which showed a pleiotropic phenotype, were attenuated in animal models as they were rapidly cleared from the blood of infected mice possibly as a consequence of producing reduced amounts of CP.

Capsule phase variation in several pneumococcal serotypes, under conditions that mimic nasopharyngeal carriage, has been detected using Sorbarod biofilms. Molecular analyses of a subpopulation of acapsular colonies (about 30%) generated in the biofilm revealed random tandem duplications of different nt size within *cap3A*, *cap8E*, or *tts* genes through an still unknown mechanism [45,46]. A significant proportion of these acapsular colonies (about 20% in the case of *cap3A* mutants) reverted to being capsular on subculture. These revertants possessed CP wild-type genes in the pneumococcal serotypes 3, 8, and 37 that have been independently analysed so far, indicating the precise excision of the duplication. Additionally, the frequency of phase reversion exhibited a linear relationship between (log) frequency of reversion and (log) length of duplication. The pneumococci studied in Sorbarod biofilms embrace serotypes with either simple (one or two genes) or complex (12 genes) capsule-coding loci. The observed duplication within *cap8E* suggested that capsule variation could also be controlled by tandem sequence duplication in *capE* homologues in other pneumococcal serotypes that construct their capsule through polymerization of lipid-linked intermediates. In sharp contrast with the quantitative change in encapsulation associated with the opacity phase types (see above), the *cap3A* switch associated with capsule phase variants is apparently all or nothing, with acapsular isolates possessing disrupted *cap3A* genes and being totally unresponsive to the Quellung reaction with type 3 antiserum [45].

2.3. Genetic variability in pneumococcal capsule genes

As already mentioned, the genes encoding the enzymes involved in CP biosynthesis in most serotypes discussed so far are clustered together in the genome of *S. pneumoniae* exhibiting a cassette-like organization flanked by the genes *dexB* and *aliA*. Generally speaking, this is a stable structure of the strain. Early studies showed that inter-type transformation (transformation of pneumococci to the heterologous type) implies the substitution of the CP of the recipient by that of the donor type and thus large pieces of DNA should be involved [47]. Furthermore, different recombinational replacements within the capsular biosynthesis operon have also been documented in the recent literature. This type of recombination takes advantage of genes within the capsular gene cluster exhibiting high similarity between serotypes [37,48].

Although the frequency of inter-type transformation in vivo does not appear to be very high [49] it is currently recognized that serotype switching does take place in vivo [50,51], as illustrated by the isolation of eight type 19F variants that were otherwise genetically identical to the major Spanish multiresistant 23F clone (Spain²³F⁻¹). It was demonstrated that recombination at the *cap* locus had taken place on at least four occasions [37]. In vitro construction of four isogenic capsular variants of the virulent strain TIGR4, whose genome has been completely sequenced [52], has rendered variants suitable for studies on the role of capsular
polysaccharide in immunity, colonization, and pathogenesis [53].

An alternative mechanism of serotype switching is illustrated by genetic analysis of serotypes 15B and 15C capsular clusters. The principal difference between these CPs is the presence of an O-acetyl group attached to the pentasaccharide repeating unit in the case of serotype 15B. In this serotype there is an intact ORF (cps15bM) that appears to code for an acetyltransferase [54]. This gene contains a short TA tandem repeat consisting of eight TA units whereas in serotype 15C this gene has nine TA units that result in a frameshift, and a truncated product. Moreover, the analysis of clinical isolates confirmed that the structural difference between these serotypes is based on variations in the short tandem TA repeat of cps15bM and that the transition between serotypes is due to slipped-strand mispairing with deletion or insertion of TA units [54].

Transposition-like events may also contribute to capsular diversity in S. pneumoniae. All the capsular gene clusters of S. pneumoniae are flanked by insertion sequence (IS) elements. This is particularly evident in the region downstream of dexB, in which copies of IS1167 or IS1202 are frequently found either in active or inactive form [55]. A copy of the IS1167 element is also located downstream of the capl operon followed by four ORFs putatively involved in the synthesis of dTDP-rhamnose (Fig. 1). Since the capsule of serotype 1 does not contain rhamnose it has been suggested that this serotype has arisen via a transposition event that took place in an ancestor strain that had rhamnose in its CP [15]. Interestingly, we have characterized a new insertion sequence (IS) element that is present in the genome of most type 1 pneumococci, and sometimes inactivates CP biosynthesis (ISestingly, we have characterized a new insertion sequence ancestor strain that had rhamnose in its CP [15]). Inter-

arisen via a transposition event that took place in an putatively involved in the synthesis of dTDP-rhamnose syntheses of the cell wall, the separation of the daughter cells at the end of the cell division, cell motility, etc. It has also been suggested that the CWHs are responsible for the irreversible effects caused by β-lactam antibiotics (bacteriolytic and bactericidal effects) that turned out to be the most relevant from the clinical point of view [57]. These enzymes show both substrate and bond specificities. The former characteristic is related to their interaction with the insoluble substrate whereas the latter determines the site of action. The bond specificity allows the classification of CWHs as: (a) glycosidases (including β-N-acetylmuramidases (lysozymes) and β-N-

acetylglucosaminidases, which hydrolyze the β-1-4 glycosidic bond between N-acetyluramuramic acid and N-acetylgalactosamine or vice versa, respectively), and transglycosylases, which cleave the same bond that lysozymes but that do not transfer the glycosyl moiety onto water but onto the C-6 hydroxyl group of muramic acid; (b) amidases, that cleave the amide bond between the lactyl group of N-acetyluramuramic acid and the α-amino group of L-alanine, the first amino acid (aa) of the stem peptide of the cell wall; and, (c) endopeptidases that cut within the peptide moiety of the peptidoglycan (Fig. 2). Collectively, these enzymes are also named peptidoglycan hydrolases. Besides, other proteins such as the pneumococcal phosphocholine esterase that releases phosphocholine residues from the cell wall teichoic acid (TA) (see below) is also a CWH.

3.2. CWHs of the pneumococcus

The activity of some pneumococcal CWHs appears to be constrained, at the post-translational level, by the membrane lipoteichoic acid [58]. Pneumococcal TA and lipoteichoic acids contain choline in their structure, an aminoalcohol that plays a fundamental biological role in the physiology of pneumococcus converting choline in the surface signature of S. pneumoniae [59]. Choline residues of the cell wall act as binding ligands of ChBPs. Up to 15 ChBPs have been found in the genome of the TIGR4 strain recently sequenced [52]. The four CWHs identified so far in S. pneumoniae are ChBPs [52] (see below) and include: two glycosidases: LytC, a β-N-

acetyluramidase (lysozyme) [60] and LytB, a β-N-

acetylglucosaminidase [61], an amidase, LytA, the major autolysin of this bacterium [62], and the Pce phospho-
ylcholine esterase (also designated as CbpE) acting at the TA side chains of the cell wall [63,64] (Fig. 2). Transglycosylases and endopeptidases have not been identified yet in S. pneumoniae.

3.2.1. N-acetyluramoyl-L-alanine amidase: LytA

The lytA gene encodes the major S. pneumoniae autolysin (amidase) and represents the first example of a bacterial autolytic gene that was cloned and expressed [65]. The lytA gene codes for a 318-aa protein with a predicted M, of 36,532 (Fig. 3). The primary translation
product of *lytA* is the low active form (E-form) of the amidase [65] that is converted, in vitro and in vivo, to the fully active C-form in a process named ‘conversion’ that requires the presence of choline [66]. A *lytA* promoter and a typical rho-independent terminator has been also characterized [67]. Moreover, in noncompetent cells the *lytA* gene is also expressed from the upstream located *recA* promoter [68].

3.2.2. β-N-acetylglucosaminidase: *LytB*

The *lytB* gene encodes a 76.4 kDa putative glucosaminidase (658 aa residues). LytB contains a 23-aa-long,
The first direct experimental evidence showing that a pneumococcal CWH required the presence of choline for activity came from the early work carried out with the LytA amidase at Tomasz’s laboratory [72]. At present, we know that all the CWHs identified so far are choline-dependent for activity in the pneumococcal system. However, it was not until 1988 when this requirement could be explained on structural bases once the gene encoding a choline-dependent lysozyme (Cpl-1) from the pneumococcal bacteriophage Cp-1 was characterized (Fig. 3) [73]. Sequence comparisons revealed that the C-terminal parts of Cpl-1 and LytA shared extensive sequence similarities whereas the respective N-terminal moieties were completely different. As both enzymes only shared the choline requirement for activity it was assumed that the C-terminal domain of both proteins was involved in choline binding whereas the active center should reside at the N-terminal domain. Further experimental support to this proposal was achieved with additional research on several new phage lytic enzymes, and construction of functional chimeric proteins (for a review, see [74]) as well as by the independent functional expression of the two domains [75,76]. The C-terminal domain of LytA and Cpl-1 is constituted by a tandem of about 20-aa repeats, known as ChBRs (Fig. 3). Similar repeats (or motifs) have been now found in a large family of cell wall-binding proteins ( Pfam ID code PF01473; CW_binding_1 repeat) that currently has more than 1400 members although only few of them appear to be related with choline binding. Also, it recently became clear that all bacterial CWHs appear to have a two-domain architecture [77].

Before the simultaneous characterization of LytB and LytC, all previously described ChBPs (including those from S. pneumoniae phage, as discussed below) possessed the C-terminal part of Cpl-1 and LytA shared extensive sequence similarities whereas the respective N-terminal moieties were completely different. As both enzymes only shared the choline requirement for activity it was assumed that the C-terminal domain of both proteins was involved in choline binding whereas the active center should reside at the N-terminal domain. Further experimental support to this proposal was achieved with additional research on several new phage lytic enzymes, and construction of functional chimeric proteins (for a review, see [74]) as well as by the independent functional expression of the two domains [75,76]. The C-terminal domain of LytA and Cpl-1 is constituted by a tandem of about 20-aa repeats, known as ChBRs (Fig. 3). Similar repeats (or motifs) have been now found in a large family of cell wall-binding proteins ( Pfam ID code PF01473; CW_binding_1 repeat) that currently has more than 1400 members although only few of them appear to be related with choline binding. Also, it recently became clear that all bacterial CWHs appear to have a two-domain architecture [77].

3.3. Domain organization of the pneumococcal CWHs: choline-binding proteins

The first direct experimental evidence showing that a pneumococcal CWH required the presence of choline for activity came from the early work carried out with the LytA amidase at Tomasz’s laboratory [72]. At present, we know that all the CWHs identified so far are choline-dependent for activity in the pneumococcal system. However, it was not until 1988 when this requirement could be explained on structural bases once the gene encoding a choline-dependent lysozyme (Cpl-1) from the pneumococcal bacteriophage Cp-1 was characterized (Fig. 3) [73]. Sequence comparisons revealed that the C-terminal parts of Cpl-1 and LytA shared extensive sequence similarities whereas the respective N-terminal moieties were completely different. As both enzymes only shared the choline requirement for activity it was assumed that the C-terminal domain of both proteins was involved in choline binding whereas the active center should reside at the N-terminal domain. Further experimental support to this proposal was achieved with additional research on several new phage lytic enzymes, and construction of functional chimeric proteins (for a review, see [74]) as well as by the independent functional expression of the two domains [75,76]. The C-terminal domain of LytA and Cpl-1 is constituted by a tandem of about 20-aa repeats, known as ChBRs (Fig. 3). Similar repeats (or motifs) have been now found in a large family of cell wall-binding proteins ( Pfam ID code PF01473; CW_binding_1 repeat) that currently has more than 1400 members although only few of them appear to be related with choline binding. Also, it recently became clear that all bacterial CWHs appear to have a two-domain architecture [77].

Before the simultaneous characterization of LytB and LytC, all previously described ChBPs (including those from S. pneumoniae phage, as discussed below) possessed the C-terminal part of Cpl-1 and LytA shared extensive sequence similarities whereas the respective N-terminal moieties were completely different. As both enzymes only shared the choline requirement for activity it was assumed that the C-terminal domain of both proteins was involved in choline binding whereas the active center should reside at the N-terminal domain. Further experimental support to this proposal was achieved with additional research on several new phage lytic enzymes, and construction of functional chimeric proteins (for a review, see [74]) as well as by the independent functional expression of the two domains [75,76]. The C-terminal domain of LytA and Cpl-1 is constituted by a tandem of about 20-aa repeats, known as ChBRs (Fig. 3). Similar repeats (or motifs) have been now found in a large family of cell wall-binding proteins ( Pfam ID code PF01473; CW_binding_1 repeat) that currently has more than 1400 members although only few of them appear to be related with choline binding. Also, it recently became clear that all bacterial CWHs appear to have a two-domain architecture [77].

Before the simultaneous characterization of LytB and LytC, all previously described ChBPs (including those from S. pneumoniae phage, as discussed below) possessed the C-terminal part of Cpl-1 and LytA shared extensive sequence similarities whereas the respective N-terminal moieties were completely different. As both enzymes only shared the choline requirement for activity it was assumed that the C-terminal domain of both proteins was involved in choline binding whereas the active center should reside at the N-terminal domain. Further experimental support to this proposal was achieved with additional research on several new phage lytic enzymes, and construction of functional chimeric proteins (for a review, see [74]) as well as by the independent functional expression of the two domains [75,76]. The C-terminal domain of LytA and Cpl-1 is constituted by a tandem of about 20-aa repeats, known as ChBRs (Fig. 3). Similar repeats (or motifs) have been now found in a large family of cell wall-binding proteins ( Pfam ID code PF01473; CW_binding_1 repeat) that currently has more than 1400 members although only few of them appear to be related with choline binding. Also, it recently became clear that all bacterial CWHs appear to have a two-domain architecture [77].

Before the simultaneous characterization of LytB and LytC, all previously described ChBPs (including those from S. pneumoniae phage, as discussed below) possessed the C-terminal part of Cpl-1 and LytA shared extensive sequence similarities whereas the respective N-terminal moieties were completely different. As both enzymes only shared the choline requirement for activity it was assumed that the C-terminal domain of both proteins was involved in choline binding whereas the active center should reside at the N-terminal domain. Further experimental support to this proposal was achieved with additional research on several new phage lytic enzymes, and construction of functional chimeric proteins (for a review, see [74]) as well as by the independent functional expression of the two domains [75,76]. The C-terminal domain of LytA and Cpl-1 is constituted by a tandem of about 20-aa repeats, known as ChBRs (Fig. 3). Similar repeats (or motifs) have been now found in a large family of cell wall-binding proteins ( Pfam ID code PF01473; CW_binding_1 repeat) that currently has more than 1400 members although only few of them appear to be related with choline binding. Also, it recently became clear that all bacterial CWHs appear to have a two-domain architecture [77].
sharp contrast with all other pneumococcal CWHs [79] (see below).

Choline, as a structural component of the pneumococcal TA, is essential for the physical attachment of the CWHs to the cell wall. Binding to the cell wall is essential for subsequent hydrolysis as long as the substrate is insoluble, necessitating an attachment step prior to positioning correctly for the catalysis. As a matter of fact, soluble choline- or ethanolamine-containing muropeptides were indeed hydrolyzed by LytA at similar rates [80]. On the other hand, cell walls isolated from choline-independent pneumococcal mutants grown in the absence of any aminoalcohol, and lacking either choline or ethanolamine in the TA, were also resistant to degradation by LytA [81,82]. However, a complete and rapid hydrolysis was observed when pneumococcal TA-free peptidoglycan was used as substrate [82]. This observation is consistent with the finding that murein from Gram-negative bacteria (i.e., *E. coli* and *Pseudomonas putida*), which do contain neither choline nor TA in their cell walls, were also degraded by LytA [70]. Apparently, free peptidoglycans are good substrates for the pneumococcal autolysin and TA chains may actually block access of the enzyme molecules to the peptidoglycan substrate unless this inhibitory effect is somehow neutralized through an interaction of TA-linked choline residues with choline-binding sites at the ChBD of the enzyme [82].

3.4. The crystal structure of the choline-binding domain (C-LytA)

Physicochemical techniques such as analytical ultracentrifugation and differential scanning calorimetry have been used to study the structure of LytA. These experimental approaches revealed that LytA forms dimers in solution through the C-terminal part of the molecule, and that the shape of the LytA dimer appears to be a prolate ellipsoid [83,84]. The recently reported crystal structure (solved in the presence of saturating amounts of choline) of a recombinant, shortened form of C-LytA (C-LytA*) revealed important traits on the protein domain that recognizes choline in the cell envelope. C-LytA* exhibits a protein fold, the ββ3-solenoid spiral staircase, which consists exclusively of 6 β-hairpins that stack to form a left-handed superhelix maintained by choline molecules at hydrophobic cavities on the protein surface [85]. Each hairpin consists of two antiparallel β-strands connected by a short internal loop region (Fig. 4(a)). All the β-strands have the same length and character (five residues and predominantly hydrophobic). Consecutive hairpins are connected by loops of 8–10 residues that contain a type I+G1 β-bulge turn, plus 4–6 residues mostly in an extended conformation. The hairpins extend perpendicularly from the axis towards the surface of the cylinder. The C-terminal hairpin (ChBR6) in the C-LytA* solenoid is responsible for the formation of the homodimer, which has a boomerang-like shape with arm lengths of 50 Å and an angle between the arms of ~85° (Fig. 4(a)). Inspection of the buried surface area in the C-LytA* crystal showed that aromatic and hydrophobic residues of ChBR5 and 6 make up the most extensive interactions in both monomers. Preliminary results on the three-dimensional structure of the complete ChBD of LytA (C-LytA), strongly suggest the existence of an additional ChBR
located immediately N-terminal to the ChBR1 of C-LytA* [86]. Dimerization has been related to regulation of the catalytic activity of LytA because, when lacking the C-terminal hairpin, it becomes monomeric and undergoes a significant decrease (90%) in catalytic activity [87] but apparently not in choline-binding efficiency [78]. ChBPs having ChBDs with more than 7 ChBRs seem to lack the last hairpin (previously designated as ‘tail’) probably because their affinity for the cell wall might be high enough, as they may contain a higher number of choline-binding sites. In a similar way, the ChBDs of LytB, LytC, and Pce, whose function must also be tightly regulated, contain different aa changes that may be likely related to the regulation of their biological activities (see above).

More recently, two new crystal forms of C-LytA* were obtained also in the presence of choline, and their structures solved and refined to 2.4 and 2.8 Å resolution [88]. The structures derived from both crystal forms are similar to one another and also to the one previously reported [85]. Nevertheless, significant variations arise from a different orientation of the most C-terminal hairpin with respect to the rest of the structure that produces a boomerang structure having an internal angle of ~110° in the case of tetragonal crystals. This finding indicates that LytA dimers posses a certain degree of flexibility that provides internal motion features to the protein that might be involved in determining the distance between the N-terminal domains. Depending on the integrity of the cell wall and/or the membrane, TA embracing the ChBD of LytA autolysin may move and determine the relative distance of the catalytic domains, regulating in this way the catalytic function of the enzyme. However, to get a more complete understanding of the mechanism(s) of choline binding, dimerization, and regulation of the pneumococcal LytA further studies must be undertaken. Our current efforts are aimed to determine the three-dimensional structure of LytA crystallized either in the presence or in the absence of choline.

3.5. Subcellular localization and biological role of CWHs

The large number of CWHs found in many bacterial systems (up to 18 in E. coli) [89] and the functional redundancy of several of these enzymes, most probably to compensate the loss of these fundamental proteins, illuminate on the biological importance of CWHs but cast difficulties to precisely determine their biological functions. However, we have succeeded to settle on the role of the pneumococcal CWHs:

1. It has been suggested that LytA associates to the membrane of choline-grown pneumococci [90] but it is mostly cytoplasmic when ethanolamine was used instead of choline in the growth medium [91], which suggests an interaction between the enzyme and the choline residues of lipoteichoic acid. However, the possibility that the cytoplasmic autolysin could bind to the choline-containing lipoteichoic acid in the membrane upon cell fractionation or lysis of the protoplasts cannot be completely ruled out. The finding that release of PspA (also a ChBP), but not LytA, from the cell surface could be effected by 2% choline [92] or upon incubation of a choline-independent pneumococcal mutant in the absence of this aminoalcohol [81] suggested that LytA and PspA differ in their cellular localization. It should be noted, however, that LytA has been located in an external position at the equatorial growing zone of intact pneumococci, or bound to the outer face of the cytoplasmic membrane when synthesized in E. coli [93].

In the holomogous system, a key point of the regulation of LytA activity will be, most likely, linked to the elucidation of the timing of the transport, as well as the triggering event of the general disorganization of the LytA-lipoteichoic acid complex leading to the uncontrolled action of LytA and subsequent lysis. This mechanism appears to control the activity of the enzyme even in two extreme situations such as the hyperproduction of LytA (from inside) [94] or by the addition of the enzyme to the culture medium (from outside) [95]. In both cases, culture lysis takes place shortly after cells enter the stationary phase in contrast with the normal lysis-prone behavior observed at the late stationary phase. Opposite to the other CWHs (LytB, LytC, and Pce) identified in pneumococcus, LytA has neither a cleavable SP [93] (Fig. 3) nor a cell wall anchoring motif LPXTG characteristic of many surface proteins of Gram-positive bacteria [96]. Consequently, the mechanism(s) of transport of LytA through the membrane to reach the peptidoglycan substrate remains an open question.

Many lytic enzymes from phage infecting Gram-negative bacteria also lack SPs and gain access to the peptidoglycan with the co-operation of small proteins (holins) that produce membrane openings enabling the passage of large molecules [97]. An analogous mechanism has been recently reported in Staphylococcus aureus where a 14.7 kDa protein (CidA) acts as a holin by increasing both extracellular murein hydrolase activity and sensitivity to penicillin-induced killing [98]. Although cidA homologs appear to exist in many Gram-positive and Gram-negative bacteria, a pneumococcal protein with a significant sequence similarity to CidA has not been found. On the other hand, two autolytic enzymes (p60 and NamA) of Listeria monocytogenes are secreted by using an auxiliary protein secretion system (SecA2) [99]. Although both p60 and NamA contain SPs, other SecA2-dependent proteins without SPs were also found. Interestingly, SecA2-like systems have been identified in other Gram-positive bacteria [100,101] including the S. pneumoniae strain TIGR4 [32]. However, the fact that the 37-kb region encompassing the entire
secA2 locus is missing in the genome of the pneumococcal R6 strain [102] casts some doubts on the possible contribution of SecA2 in the secretion of the LytA amidase. Recently, a SecA-dependent mechanism for the export and control of the endolysin from the E. coli phage P1 and other phage infecting Gram-negative bacteria has been reported [103]. It has been shown that the N-terminal domain of several phage lytic enzymes is both necessary and sufficient not only for export of the endolysin to the membrane but also for its release into the periplasmic space. Nevertheless, since no amino acid similarity exists between the N-terminal domain of those phage and that of LytA, the question on whether a similar mechanism might be functioning in vivo for the secretion of the LytA autolysin is currently a matter of investigation.

It has been proposed that a zinc metalloprotease, ZmpB, was a key factor to control translocation to the cell surface of several ChBPs including LytA [104]. The authors claimed that, in a zmpB background, LytA formed an SDS-resistant 80 kDa complex with CinA, a protein induced when cells become competent for genetic transformation. Nevertheless, this assumption revealed inconsistent when the purported zmpB mutant from Novak and co-workers was demonstrated to exhibit characteristic traits of the viridans group streptococci, including lack of capsule and resistance to optochin [105].

(2) The contribution of CWHs to cell separation has been a subject of continuous interest. Whereas wild-type (lytA⁺) pneumococcal strains exhibit a typical 'diplo' morphology, lytA mutants (either containing point mutations or having a complete deletion of the gene) grow forming small chains of 8–10 cells [94,106]. However, the construction of lytB mutants by insertion-duplication mutagenesis resulted in the formation of long chains (more than 100 cells) directly demonstrating the fundamental role that this enzyme plays in cell separation (Fig. 5(a), (c), and (d)) [61]. These long chains could be dispersed, in a dose-dependent way, by the addition of purified LytB (Fig. 5(b)) [79]. The preparation of a translational fusion between the gfp gene (coding for the green fluorescent protein) and lytB allowed the purification of a fusion protein GFP-LytB. This protein was used to demonstrate the specific localization of LytB at the poles of the cell (Fig. 5(f)). It has been proposed that cell poles provide a mechanism by which proteins, once inserted, are maintained at this location [107]. In this case GFP-LytB only would bind to the specific target (the old wall zones or poles) to which it remains bound for a long time. The presence of choline is essential for the successful cell septation by LytB, but initial binding to ethanolamine-containing cell walls could also be achieved [79] indicating that LytB has a wider range of substrate recognition to bind the pneumococcal cell walls, in contrast to the choline requirement of LytA.

Fig. 5. Microscopic observations of the dispersion of R6B (a lytB mutant) chains by purified LytB glucosaminidase, and localization of LytB in the surface of S. pneumoniae. (a) R6B cells, (b) R6B (2.5 × 10⁸ cells ml⁻¹) treated with 2.7 μg ml⁻¹ (30.4 units mg⁻¹) at 37 °C during 30 min. Cells were observed using phase-contrast microscopy. (c, d) Electron micrographs of ultrathin sections of R6B cells. (e, f) A culture of R6B cells received a pure preparation of a fusion protein between the green fluorescent protein and LytB (10 μg/ml), and was incubated at 37 °C during 20 min. Pictures were taken with a Nikon Eclipse inverted microscope in phase contrast (e) and using fluorescence (f). Bar in (c) and (d) represents 1 and 0.5 μm, respectively.
(3) In sharp contrast with what occurs at 37 °C, all the lytA mutants lysed when incubated overnight at 30 °C [108]. The insertional inactivation of lytC in the ΔlytA strain M31 (M31C) makes the cells resistant to autolysis upon incubation at 30 °C. Addition of purified LytC lysozyme to this deficient strain showed that this enzyme was kept under regulatory control of the ‘cured’ cells until the culture reached the stationary phase of growth and then the culture started to lyse at 30 °C. Besides, the lytC mutants exhibited a normal growth rate and an average chain length similar to that of the parental strain. Combined cell fractionation and Western blot analysis showed that the unprocessed LytC protein is located in the cytoplasm whereas the processed, active form of LytC is tightly bound to the cell envelope [60]. This binding strength is probably related to the number of repeating motifs and could be a mechanism for the regulation and control of the potentially suicidal activity of this enzyme.

(4) In vitro experiments carried out with purified Pce confirmed that this enzyme is a TA phosphorylcholine esterase that only remove a maximum of 20% phosphorylcholine residues from the cell wall TA, in agreement with early results [109]. As it is also bound to the envelope, Pce should play its role only after its secretion through the membrane, although it is currently difficult to assign a defined function. It has been suggested that the about 20% of residues removable by the enzyme might exist either in an anatomically unique position in the cell wall or might represent terminal residues in the TA chains [63,64]. This esterase activity might regulate the availability of choline residues required for activity (attachment) of its own and/or of other ChBPs.

3.6. CWHs and genetic transformation

Genetic transformation is an important mechanism for gene exchange between streptococci in nature that implies the ability to take up free DNA from the medium and the incorporation of this DNA into the recipient genome by recombination. This ability of cells to take up DNA, called competence, is a transitory property in S. pneumoniae which develops suddenly during the exponential phase of growth. A role for CWHs in genetic transformation has been repeatedly proposed. Seto and Tomasz [110] first reported an increase in autolytic activity during competence development. Although our preliminary analyses had suggested that the lytA gene was constitutively expressed along the exponential phase of growth [67], more recent data have revealed, however, that the 5.7-kb competence-specific, cinA recA transcript also includes the lytA gene, and that transcription of lytA is strongly induced during competence from a promoter located upstream of cinA [68]. More recently, the kinetics of global changes in transcription patterns during competence development in S. pneumoniae were analyzed by using macro and microarrays [111–114], and differential fluorescence induction (DFI) [115]. Competence was induced by the addition of the competence-stimulating peptide (CSP) to S. pneumoniae cultures grown to the early exponential phase. These analyses fully confirmed that lytA is overexpressed during competence and provided additional evidence that lytA is part of a competence-induced operon. The reason(s) for this finding is still unclear since it is well known that the LytA amidase is not required for transformation as every lytA mutant tested is fully transformable [106,116,117]. In addition, CWHs do not appear to be required to achieve the competent state for transformation since S. pneumoniae mutants lacking all the four CWHs are still transformable at, or near, normal levels [69].

Charpentier and co-workers recently claimed that a mutant deficient in the ClpC ATPase, a subfamily of HSP100/Clp molecular chaperones-regulators of proteolysis, formed long chains, and failed to be transformed and to undergo lysis after treatment with penicillin or vancomycin [118]. Moreover, the clpC mutant apparently failed to express LytA and other ChBPs suggesting that the heat shock protein ClpC might play an essential complex pleiotropic role in pneumococcal physiology. Nevertheless, more recent approaches using a ΔclpC mutant constructed by gene deletion/Replacement have established that ClpC does not play a major role in autolysis [119]. Besides, it has also been reported that transcription of the lytA gene is induced in a S. pneumoniae clpP mutant (ClpP, ATP-dependent protease) [120], although no significant variations in autolysis and/or LytA levels were found in this and other clp mutants affected in the heat-shock response [119–121]. Most likely, the induction of lytA in the clpP mutant results from competence induction [119].

During a search for β-lactam resistant determinants in S. pneumoniae the two-component system (TCS) CiaRH was identified [122]. An increased rate of autolysis in the stationary phase was observed in a ciaR mutant, and attributed to a constitutive regulation of competence operon (comCDE) [123,124] and to a competence-independent expression of lytA [125]. The existence of a connection between the CiaRH two-component system (TCS) and autolysis was suggested since the R6ciaHΔ306 mutant was resistant to sodium deoxycholate-induced autolysis and to treatments that induce protoplast formation [126]. However, CSP was demonstrated to trigger growth arrest and stationary phase autolysis in ciaR cells in a more recent report establishing a direct connection between competence and autolysis [113].

Under natural conditions pneumococcus might participate in gene exchange when located in biofilms in places such as the nasopharynx, the common carrier state of this microorganism [46,106]. In fact, there are
data supporting the idea that DNA-mediated spontaneous transformations may occur in nature, even interspecific transformation between pneumococci and streptococci, and that they may play a role in bacterial infections [127,128]. Interestingly, type 9 pneumococci were transformed by DNA released from living unencapsulated pneumococci sprayed onto the pharynx in a healthy human carrier [128]. As early as in 1960, release of transforming DNA to the medium was reported to take place in pneumococci relatively early in the exponential phase of growth [129,130] but the mechanism(s) of release remained unknown for many years. Recently, it was documented the spontaneous release of DNA either from competence-induced pneumococcal cells or when grown in competence-permissive C medium [113], or after induction of the competent state [129–131], thus establishing a causal relationship between competence and DNA release. However, we do not know yet whether the source of donor DNA is the autolysed dead cell or implies the existence of an active process. Since the pores in the Gram-positive cell wall are too small to allow DNA to pass through it has been assumed that the liberation of these molecules from the cytoplasm to the medium involves degradation of the peptidoglycan. The hypothesis focusing on autolysis as a possible explanation to the mechanism of DNA release was favoured since lytA mutants exhibited a two- to four-fold decrease in the amount of DNA available in the medium [132]. Also, a role for several pneumococcal CWHs, mainly the LytA amidase, has been proposed for DNA release from a subfraction (5–20%) of the actively growing bacterial population, probably by cell lysis. The other fraction will act as recipient of the DNA from the medium [131]. More recent results have suggested that competence-induced pneumococci lyed competence-deficient cells of the same strain during co-cultivation [133]. Apparently, LytA and, probably, additional CWHs are involved in this process where the lysins made by competent cells are not released into the growth medium. More likely, the lysins are anchored to the surface of competent cells by their ChBDs and cause lysis of noncompetent pneumococci through cell-to-cell contact (heterolysis) rather than by autolysis. However, it cannot be ruled out that both mechanisms operate simultaneously. Although Steinmoen et al. [133] proposed that the autolytic lysozyme LytC is not involved in the competence-activated lysis mechanism, other authors suggest the combined action of the autolytic amidase and lysozyme is required to promote DNA release in competent cultures of S. pneumoniae [M. Moscoso and J.P. Claverys 2004, in preparation].

3.7. CWHs and pathogenesis

LytA is the best characterized CWH, and has been implicated in the pathogenicity of pneumococci. The release of the components of the cell wall after bacterial cell death has been shown to be highly inflammatory in some animal models [134]. In pneumococcus, the indirect implication of LytA in pathogenesis also involves the release of cytoplasmic proteins like pneumolysin and other intracellular toxins [135]. Furthermore, early studies carried out using lytA mutants revealed a tolerant response of these strains against β-lactam antibiotics [136], this is, the classical bacteriolytic response of the wild-type strain against penicillin changed to a bacteriostatic one. Phenotypic antibiotic tolerance in all bacterial species also arises as growth rate decreases [137]. Tolerance appears to be of clinical significance since the inability to eradicate tolerant bacteria leads to failure of antibiotic therapy [138,139] and, moreover, tolerance is thought to promote the development of antibiotic resistance because it creates survivors of antibiotic therapy, this is, many penicillin-resistant pneumococci are also tolerant [140]. Besides, it has been reported that CP negatively influences the lytic process and contributes to antibiotic tolerance in clinical S. pneumoniae isolates [141]. In this study it was reported that, in agreement with what was found in different clinical isolates, mutant strains not expressing CP were more lytic after β-lactam and/or vancomycin treatment that their corresponding encapsulated progenitors of serotypes 2, 4, and 9V. Besides, spontaneous autolysis in stationary phase also was negatively affected by capsule expression in a LytA-dependent manner. This behavior was not apparently due to a defect (or alteration) in the transcription and/or expression of the lytA gene. Whether the CP interferes in some way with the in situ binding or activity of the LytA amidase is currently unknown.

Insertion-duplication mutagenesis was employed to construct lytB, lytC, or pce (cbpE) mutants of the type 4 pneumococcal strain TIGR4. All of them showed a significant reduction in the colonization of the nasopharynx by S. pneumoniae using an infant rat model but no changes in virulence were observed when tested in a sepsis model [142]. It has been proposed that bacterial chain formation limits the dissemination of the bacteria during infection [143]. Since a single murein hydrolase (LytB) appears to be fundamental for cell separation in S. pneumoniae, this enzyme might be an interesting target for the design of a future conjugate vaccine. Besides, loss of function of LytC or Pce reduced adherence to Detroit cells at 30 °C to about 70% [142]. On the other hand, Vollmer and Tomasz [64] have reported that inactivation of Pce in pneumococci producing type 3 capsule caused a significant increase in virulence when tested in an intraperitoneal mouse model. It may be that the increase in the number of choline residues (due to the inactivation of the pce gene) may facilitate interactions with the platelet-activating factor receptor during infection.
4. Bacteriophage of pneumococcus

Bacteriophage play central roles both in the shaping of natural populations of bacteria and in the development of genetics. Historically, phage were profoundly important for seminal studies illuminating the nature of the gene and in early studies of gene regulation. Bacteriophage have also been essential tools for therapeutic purposes [144,145]. Recently, phage have again come to prominence due to the discovery of the remarkable dynamics underlying their evolution, as well as due to their role in virulence, and use as therapeutic agents.

Pneumococcal phage (pneumophage) were first isolated in 1975 from throat swabs of healthy children, by two independent groups [146,147]. A drawback to the study of these entities has been the lack of simple and reproducible techniques that facilitate the visualization of phage plaques. Genetic analyses have been also handicapped by difficulties to distinguish phenotypic plaque variations at a glance. However, the molecular characterization of several pneumophage, from different origins, facilitated in recent years the study of the peculiarities and diversity of these phage [148]. Pneumophage have also represented a useful way for expanding the information on the mechanisms of genetic interchange in pneumococcus as well as to gain information on the relationship between bacterium and virus. A paradigmatic example on how to get a better knowledge on the molecular biology of pneumococcus using this experimental setting has been achieved through the study of the lytic enzymes [74]. Most interestingly, phage lysins are currently being studied as a mean for bacterial control using animal models which require a double lysins are currently being studied as a mean for bacterial control using animal models which require a double stranded DNA phage specific for the pathogen of interest to serve as a lysin source [149]. Bacteriophage are the most abundant entities in the biosphere (about 10^{31}) and, consequently, a stringent selection of a variety of phage appropriate to combat a given pathogen is a manageable aim.

The abundance of temperate phage in clinical isolates of pneumococcus was suggested some years ago [150] and, recently, it has been proposed that they account for as much as 75% of the samples analyzed [151]. The complete nt sequence of four pneumococcal phage have been published [152–155]. The comparative genomic analyses have revealed interesting clues like the presence of highly similar prophage in several clinical isolates of pneumococcus of different capsular types, which indicates the widespread presence of closely related pneumococcal phage in virulent strains [156]. Detailed studies carried out in different bacterial species have shown that phage are important vehicles for the transmission of virulence genes within bacterial populations [157]. In some pathogens, like *Streptococcus pyogenes*, more than 90% of the isolates are lysogenic and many of them are polylysogenic [158]. There are many examples of toxin genes present in phage or prophage genome [159]. Most interesting, recent observations based on comparative studies of the genomes of several mycobacteriophage have suggested a much broader contribution of phage in bacterial virulence and the host response to bacterial infections than that previously thought [160]. Genomic analyses also appear to be a fundamental step in furthering our understanding on the contribution of phage genes to the virulence of a clinically outstanding microorganism like *S. pneumoniae* and to appreciate the role that phage play in bacterial diversity.

4.1. General features of phage Dp-1, Cp-1, MM1, and EJ-1

Pneumophage have been isolated from a variety of sources and geographical locations [148]. Four of these phage have been thoroughly characterized, including the determination of the complete genome sequences [161]. Dp-1 and Cp-1 are lytic phage whereas MM-1 and EJ-1 are temperate ones. The host range of Cp-1 phage is not restricted to *S. pneumoniae*, since it can infect and replicate in *S. oralis*, which shares a common habitat with pneumococcus in humans [162]. Several relevant physicochemical characteristics of the DNAs from pneumococcal phage and some of the derived proteins have been recently summarized [163]. Electron micrographs revealed the presence of three different viral morphologies in the isolates studies so far: Siphoviridae, Podoviridae, and Myoviridae (see insets in Fig. 3). Interestingly, choline is required not only for the activity of lytic enzymes of the pneumophage but also for Dp-1 binding to the cell wall receptors [164], suggesting the presence in the phage tail of a protein that also binds to the choline-TA. Analysis of the Dp-1 genome has revealed the presence of an *orf* coding for such a protein. The putative product of orf55 (the anti-receptor) codes for a 1230 aa protein whose 424 amino- (N-) terminal aa do not share significant similarity with any known protein. From aa position 425 to the end of the protein, gp55 share high similarity (43% identity, 60% similarity) with the anti-receptor protein (gp38) of the *Streptococcus thermophilus* phage 7201 and its relatives. As many other phage proteins, gp55 contains 6 groups of collagen-like motifs (Gly-X-Y) (37 repeats) that have been proposed to mediate binding of bacteriophage proteins to other macromolecules [165]. The collagen-like motifs are interspersed among additional repeated motifs. One of them (motif 3) is particularly rich in aromatic residues and similar in sequence to that characteristic of ChBDs of pneumococcal proteins [85]. Gp55 contains 4 copies of this motif, which appears to be the minimum required for efficient binding to the Ch residues of the *S. pneumoniae* cell wall [78].
A remarkable characteristic is the presence in Cp-1, and related phage, of a protein covalently linked to the 3'-ends of their DNAs. This protein is fundamental for phage DNA replication by a protein-priming mechanism that has been well documented for phage φ29 infecting Bacillus subtilis [166,167]. In this mechanism for initiation of DNA replication the primer, instead of being the 3'-OH group of a nt, is the –OH group of a serine, threonine, or tyrosine residue of the terminal protein. In Cp-1 the aa covalently linked to the DNA was identified as threonine [168]. Whereas all the phage DNAs of the φ29 family have a short inverted terminal repeat of 6–8 nt, the Cp family exhibits very long, inverted terminal repeats of about 250–350 nt, depending on the phage [169,170]. The possible role of these repeats appears to be the formation of a panhandle structure of the parental DNA strand that could be displaced in the process of replication [167]. On the other hand, protein-linked DNAs in temperate phage MM1 and HB-3 have also been reported but the detailed biological role of this protein as well as the physicochemical peculiarities of the link between protein and DNA remain to be established. This property is unique since the protein, detected in the DNA isolated from purified virions, must be lost prior to phage DNA integration into the host chromosome [154,171].

4.2. Sequence analysis of the genome of lytic pneumophage: functional hints

The complete genome sequences for the four pneumococcal phage described above have been determined, some genes functionally characterized, and many others putatively identified by a variety of bioinformatic analyses. The length of the genomes of the four pneumophage vary from the 19,343 bp of Cp-1 [152] to the 56,506 bp of Dp-1 [153] whereas the unit genome sizes for MM1 and EJ-1 were 40,248 bp [154] and 42,935 bp [155], respectively, in the prophage state. The average G+C content ranges from 38.4% to 40.3%, which is similar to the 39.7% reported for the S. pneumoniae genome [52].

Genomic maps of the four pneumophage are shown in Fig. 6. The genome of Dp-1 appears to be organized into four major clusters (replication, packaging, structural, and lytic) in agreement with that found for other lytic siphoviruses infecting lactic acid bacteria [172]. However, Dp-1 DNA shows two remarkable features that appear to be unique. First, it contains 11 ORFs that are apparently transcribed leftwards (orf27 to orf37). With the exception of the predicted product of orf34 (gp34), similar to other hypothetical ORFs, none of them showed a significant similarity with sequences

![Fig. 6. Schematic representation of the genomes of the pneumococcal phage Cp-1, Dp-1, MM1, and EJ-1. Genes are represented as arrows indicating the direction of transcription. The putative products of selected ORFs are indicated below each scheme. Yellow, dark blue, green, light blue, and red arrows correspond to genes involved in lysogeny regulation, DNA replication, packaging and head-assembly, tail morphogenesis, and lysis of the host cell, respectively.](image-url)
included in the data banks. Second, a set of at least 5 genes (orf1 to orf5), potentially encoding proteins involved in the synthesis of 6-pyruvoyltetrahydropterin, is located at the most left part of the genome (Fig. 6). These include a GTP cyclohydrolase I (gp5) and a synthase (gp3). In Dp-1, these genes appear to be organized as a single transcription unit. Dp-1 DNA is not cleaved by many restriction endonucleases although the cleavage sites are present. Physicochemical studies suggested that Dp-4, a member of the Dp-1 family, possesses a highly modified DNA [173]. As the cluster of genes embracing from orf1 to orf5 are located inside the replication cluster of Dp-1 it is tempting to speculate whether they (and possibly other genes) might be involved in DNA modification.

As documented above, Cp-1 has a linear double-stranded DNA genome with a terminal protein (gp4) covalently linked to its 5’ ends, and a 236-bp-long perfect inverted repeat. On the basis of its temporal expression [174], transcription of the Cp-1 genome takes place in two stages: genes involved in DNA replication are expressed early after infection, whereas those coding for structural components of the phage particle, morphogenesis, and host lysis are expressed subsequently. One promoter located upstream of orf1 and two more that mapped between orf2 and orf3 direct the transcription of early genes (orf1 to orf7). A 1.3-kb early right-to-left transcript was detected by Northern blotting that is responsible for the translation of orfa to orfc [152,174]. Transcription of late genes is directed by 7 additional promoters. In summary, Cp-1 DNA transcribes almost exclusively one strand, from left to right, and it is unlikely that the shift from early to late transcription was controlled in a unique region of DNA. One of the peculiarities of φ29 and related phage comes from the unique role carried out by a short RNA (pRNA; 174 nt in φ29) in the packaging process [152,167,174]. Primer extension analysis of the Cp-1 genome also suggested the transcription of a small RNA (169 nt) showing a secondary structure similar to that of φ29 [152]. The pRNA from Cp-1, however, did not show any sequence similarity with that from φ29, in a similar way to that reported for the Bacillus phage GA-1 and B103. In addition, the promoter for pRNA is located at the far-left end of the φ29, B103, and GA-1 genomes whereas, in Cp-1 DNA, it is located at the rightmost part, partially overlapping orfb (Fig. 6).

Following a detailed set of experiments, the initiation and the first elongation steps of Cp-1 DNA replication, have been studied [175]. It could be concluded that: (a) formation of terminal protein-dAMP is template-instructed; (b) single-stranded DNA molecules can support protein-primed replication; (c) a specific DNA sequence at the origin is required for efficient template recognition; (d) terminal protein-dAMP formation is directed by the third nt from the 3’-end of the template molecule, although the two terminal nt are recovered during the first steps of elongation; (e) a stepwise sliding-back mechanism accounts for the maintenance of Cp-1 DNA ends. These results reinforce the hypothesis that sliding-back must be a common feature in all genomes that use protein-priming to initiate replication [167].

4.3. Sequence analysis of the genome of temperate pneumophage

The DNA isolated from mature phage particles of the temperate phage MM1, like that of HB-3 [171] and their relatives [176], appears to possess a covalently bound protein, as reported for Cp-1. It has been shown that MM1 DNA is circularly permuted, terminally redundant, and packaged via a headful mechanism. In the prophage state, the MM1 genome is organized into five major clusters and starts with the gene encoding the integrase [154]. The five leftmost genes comprise the lysogeny cluster. This region contains a characteristic gene (orf4) that codes for a 120-aa protein (CI), which exhibits high similarity to several phage-related transcriptional repressors, and two divergent promoters (Pr and P1). We have recently characterized for the first time the lysogenic control region where CI prevents elongation of the transcripts controlled by Pr and P1 [177]. Analyses comparing the MM1 genome with those of other bacteriophage revealed similarities, mainly with genomes of phage infecting Gram-positive bacteria, which suggests recent exchange of genes between species colonizing the same habitat. Remarkably, the modular organization of the MM1 genome is similar to those of other temperate streptococcal phage, where genes belonging to the lysogeny cluster were the only genes carried in the opposite DNA strand of the phage genome. The noticeable similarity between genes of phage infecting S. pyogenes and S. pneumoniae suggests a frequent genetic interchange between either species or a recent divergence from a common ancestor phage [154].

Bacteriophage EJ-1 was purified from mitomycin C-induced cultures of the atypical S. pneumoniae strain 101 [178]. Recent efforts to establish the correct phylogenetic position of this isolate have revealed the difficulties to ascertain whether strain 101 should be considered as S. pneumoniae or S. mitis species [155]. The report of the complete genomic sequence of EJ-1 represents the first description of the genome of a myovirus infecting a low G+C content Gram-positive bacterium [155]. The predicted proteins encoded by the EJ-1 genome revealed similarities in the lysogeny, DNA replication, regulation, packaging, and head morphogenesis protein clusters with those from several siphoviruses infecting lactic acid bacteria. However, the proteins encoded by genes orf53 to orf64, corresponding to putative tail proteins of the virion, were very similar to those of the defective B. subtilis myovirus PBSX with the notable exception of
the gene product of *orf56* (the tape measure tail protein) that was similar to proteins from phage infecting Gram-negative bacteria [155]. Several of the gene products corresponding to the lysogeny, replication, and regulation clusters appear to be mostly related to the *S. pyogenes* SF370.3 defective prophage [179,180] although sequence comparisons suggested the existence of frequent recombination events as various gene products appear to be less related to SF370.3 than to other phage. Besides, the putative excisionase gene (*orf16*) is located in the replication cluster far apart from the lysogeny module. This appears to be unique since the distance between the genes coding for the integrase and the excisionase usually ranges from zero to 6 ORFs in other temperate bacteriophage from both Gram-positive and Gram-negative bacteria, or even the excisionase gene is missing as in phage Sfi21 [181]. This finding suggests the insertion in the EJ-1 genome of intervening genes from a foreign source and/or genome rearrangements. The global image that arises from the EJ-1 genomic organization is that this phage is a mosaic one sharing sequence similarities with various different phage genomes although, quite unexpectedly, it only showed homology with a few ORFs from the pneumococcal siphovirus MM1.

4.4. Methylases

It has been well established that methylases (methyltransferases) provide functions that are important to the bacterial cells since they play essential roles in gene regulation, DNA mismatch repair, and restriction-modification [182–184]. It has been found that *L. monocytogenes* strains derived from separate outbreaks were invariably resistant to digestion by *Sau3AI* and other restriction endonucleases sensitive to cytosine methylation at 5′ GATC 3′ sites [185]. A DNA cytosine 5-methyltrasferase (C5Mtase) has been recently identified in pneumococcus [182]. Remarkably, genes putatively coding for C5Mtases have been documented in several pneumococcal phage [154,186]. The MM1 genome encodes two gene products, *gp14* and *gp15*, that are very similar respectively to the α and β polypeptides of the C5Mtase of the *S. pneumoniae* transposon Tn5252 [182]. The structure of the genes coding for these proteins is very peculiar and noticeably similar in both methylases since *orf15*, coding for 141 aa (β polypeptide) is included within *orf14* encoding a 385-aa protein (α subunit) using alternative reading frames. C5Mtase genes encoding two overlapping subunits have also been reported in *E. coli* [187]. Noteworthy, the G+C contents of *orf14* (41.4%) and *orf15* (44.1%) were clearly different from the average of the MM1 genome (38.4%) suggesting a recent horizontal transfer of the methylase-coding genes. To further support this hypothesis the analysis of other MM1-related phage was accomplished. As pointed out above, the pneumococcal phage HB-3 shares many peculiarities with MM1. Moreover, preliminary sequence data indicated a global nt sequence identity of more than 90% between these phage genomes (unpublished data). PCR analysis of HB-3 DNA using specific primers surrounding the region encoding the C5Mtases of MM1 revealed the absence of a C5Mtase gene in HB-3 DNA although a remnant corresponding to the region coding for the four N-terminal aa of *gp14* from MM1 DNA was found [186]. This observation indicated that the HB-3 genome lacks the genes coding for the C5Mtase found in MM1. Likewise, the temperate bacteriophage VO1, which has been isolated from the *S. pneumoniae* type 19A multiresistant epidemic 8249 strain (South African strain), provided some interesting clues about the origin of these C5Mtases [186]. The structural analyses of the specific integration site, protein composition, restriction patterns, and molecular dissection of the lytic system of this phage also revealed a high sequence similarity with MM1. Taking advantage of this similarity, when the VO1 region that might contain the C5Mtase genes was analyzed it was observed that, at the same genome position, there was a single gene (*orfmet*) coding for a putative C5Mtase smaller than and different from that encoded by MM1. This observation strongly suggested that the corresponding genes had been incorporated from another source(s) rather than being the result of divergent evolution. Orfmet is a protein exhibiting a 73% identity with a truncated *M. NgoMorf2P* C5Mtase encoded by *Neisseria gonorrhoeae* MS11 [188], and 34% with the putative C5Mtase coded by phage MAV1 of *Mycoplasma arthritidis* [189]. It is noteworthy that the Tn5252 C5Mtase was also detected in the genome of strain 949, the host of MM1 phage, demonstrating that this strain contains two similar C5Mtases, one encoded by the phage and the other by the transposon [186]. Whether both C5Mtases are equally active and functional has to be determined but it was observed by restriction analyses that the phage and host genomes are methylated. The co-existence of both C5Mtases put forward the possibility of a co-evolution by horizontal transfer. Interestingly, Tn5252 is present neither in the genomes of strains carrying phage VO1 nor HB-3.

4.5. Lysis cassette: evolutionary insights

A dominant feature of the phage-lysis mechanisms is to be a system based on a CWB and a second protein, a holin, required for the access of the hydrolase to its murein substrate. The combined activities of these proteins allow the liberation of the phage progeny to the medium. Holins cause non-specific lethal lesions (holes) of the host cytoplasmic membrane, most likely, through
the oligomerization of several molecules when the levels of such proteins reach a critical mass [190]. In the pneumococcal system there is evidence that this holin-lysin strategy is present in every phage analysed so far [186].

The genes coding for lytic enzymes have been cloned by taking advantage of their similarities with the host LytA amidase. Two different classes of lytic enzymes have been found in pneumococcal phage: all the reported temperate phage (HB-3, MM1, EJ-1, and VO1) as well as the lytic phage Dp-1 and ω-1 possess amidases whereas phage belonging to the Cp family synthesize lysozymes (Fig. 3). As for the pneumococcal CWHSs, all the phage lytic enzymes display a two-domain structure, where the N-terminal half is responsible of the catalytic activity and the C-terminal part functions as the ChBD. Most remarkably, Cp-7, a phage highly similar to Cp-1, codes for a lysozyme (Cpl-7) able to degrade not only choline- but also ethanolamine-containing pneumococcal cell walls [191]. Cp-7 contains a completely different C-terminal domain, this is, this lysozyme has a 2.8 identical 48-aa repeats instead of the characteristic 20-aa long ChBRs (Fig. 3). This peculiar finding provided early support to our proposal about the fundamental role of the C-terminal domains of LytA and Cpl-1 in choline recognition. On the other hand, the N-terminal domains of the Cpl-1 and Cpl-7 lysozymes are almost identical (Fig. 3) [191].

It is noteworthy that amidases coded by temperate phage phage share a high degree of similarity among them and also compared to the host LytA in their catalytic domain whereas the Pal amidase, coded by phage Dp-1, is a natural chimeric enzyme of intergeneric origin, since its N-terminal domain is similar to that of the murein hydrolase coded by the Streptococcus mitis phage SM1 [192] and the Lactococcus lactis phage BK5-T [179,193]. In the last decade, we have rehashed, by genetic engineering, the domains of different host and phage endolysins allowing the creation of chimeric cell wall degrading enzymes with new properties [74]. Thus, it is conceivable that the laboratory model of domain swapping analyzed in the pneumococcal system might have its counterparts in nature (see below).

The molecular characterization of the lysis cassette of EJ-1, Cp-1, and MM1 [148] has been carried out by cloning and expression of the two genes involved in the release of phage progeny to the medium. EJ-1 (Fig. 6) codes for a holin (Ejh) of 85 aa residues with two strongly hydrophobic regions predicted to form amphipathic α-helices. This phage also codes for an amidase (Ejl) with 85% identity to the pneumococcal LytA amidase. Expression of the ejh gene results in inhibition of growth of a variety of Gram-negative bacteria belonging to the main subgroups of Proteobacteria. Growth arrest was caused by a cell-killing activity and correlated with the formation of lesions in the cytoplasmic membrane detected by electron microscopy of ultrathin sections. Simultaneous expression of ejh and ejl genes resulted in cell lysis of cultures of E. coli and Pseudomonas putida, as well as by a dramatic change in cell morphology, characterized by empty cells and cellular debris [70]. Recently, a deeper insight into the structure–function relationship of Ejh has been achieved by using a synthetic fragment of the protein [194]. It was shown that the predicted N-terminal transmembrane helix is the potential active region of the holin molecule. This region, together with the charged N-terminus (1–32 sequence region), folds into a transmembrane helix displaying self-assembly properties that permitted membrane permeation to dextrans of various sizes. Also the study of peptide-induced bilayer lesions by atomic force microscopy allowed the visualization of polydisperse defects as holes.

To better understand the mechanism by which murein hydrolases bind and degrade the peptidoglycan, a coordinated effort between our group and other laboratories have recently revealed for the first time the three-dimensional structure of an entire pneumophage ChBP, the Cpl-1 lysozyme [195]. The polypeptide chain consists of the catalytic and the ChBDs joined by an acidic linker comprising residues 189–199 (Fig. 4(b)). The catalytic domain (residues 1–188) resembled a flattened ellipsoid that folds into an irregular (β/α)5 β5 barrel. As in regular TIM barrels (the eight-stranded α/β barrels), the first five β strands and α-helices alternate, but the α5 helix is followed by strands β6 to β8, which are connected by loops lacking any helices (Fig. 4(b)). All β strands are arranged in a parallel fashion, except β8, which runs antiparallel to the other strands. A small hairpin and a short helix (αA) cap the C-terminal side of the β sheet. The well-defined structure of the linker unambiguously identifies the beginning of the ChBD. Each repeat (p1–p6) comprises a symmetrical β hairpin followed by a loop and a coiled region and are arranged into two well-defined structural regions referred to as CI and CII (Fig. 4(b)). In addition, after the 6th repeat, there is also a short 16-aminoacid short tail. The CI structural domain (residues 200–281) folds following a left-handed superhelical arrangement of the initial four repeats with the hairpins extending perpendicularly to the axis of the superhelix, each CI repeat being located at a 120° counterclockwise rotation as in the ChBD of C-LytA* (see above). The CII structural domain (residues 282–339) folds completely differently as an almost antiparallel six-stranded β sheet formed by the last two repeats (p5–p6) and the C-terminal tail. This region is responsible for the interaction between Cpl-1 domains. Overall, the ChBD of Cpl-1 constitutes a novel and distinct arrangement of six conserved supersecondary repeats.
5. Bacteriophage therapy

The emergency of pathogenic bacteria resistant to most available antimicrobial agents has become a fundamental problem in modern medicine and some authors alerted on the possibility that we are re-entering the preantibiotic era [196]. Examples of antibiotic resistance spreading are known in virtually all human pathogens and vancomycin-resistant enterococci provide an illustrative model of an endemic disease that predispaces the individuals to bacteremia and/or endocarditis in many hospitals. The ability of virulent phage to rapidly kill and lyse infected bacteria and the skill of phage to increase in number during the infection process have made phage excellent potential agents for fighting bacterial diseases. However, temperate phage are of limited use in this therapy. Phage therapy is actually an old idea and was widely employed to treat various bacterial diseases in people and animals at the beginning of the last century [144,145]. However, it has been only recently when more rigorous assays have been performed, mainly in animal models [197,198]. Phage have been successfully used as an alternative therapy to rescue mice bacteremic from a vancomycin-resistant Enterococcus faecium clinical isolate [199]. An increased prevalence of S. pneumoniae strains resistant to important anti-pneumococcal drugs, including β-lactams and macrolides, has been observed in both developing and developed countries [200]. Moreover, it has been anticipated that pneumococcal strains resistant to both penicillin and erythromycin are increasing faster than those strains singly resistant to either [7]. Unfortunately, the lytic pneumophage currently available (Dp-1 and Cp-1 families) are not suitable for clinical use since they do not infect capsulated strains of pneumococcus either in vitro or in vivo (unpublished observations).

By means of an elegant variation of the classical phage therapy approach, Fischetti and co-workers have exploited the use of specific phage-encoded murein hydrolases to kill the species in which they were produced, defining a new class of antibiotics totally distinct from any other previously described. Such lytic enzymes have been named ‘enzbyiotics’ and successfully used to prevent and eliminate the upper respiratory colonization of mice by group A streptococci [201], to kill pneumococci from nasopharyngeal carriage of mice [202], and to rapidly detect and kill both vegetative cells and germinating spores of Bacillus anthracis [149]. In the S. pneumoniae report, those authors dealt with Pal, the amidase coded by Dp-1, which was very effective to eliminate pneumococci of several serotypes demonstrating that the capsule did not interfere with the enzyme’s access to the cell wall. It is important to emphasize that, in this case, the rise of resistant strains appears a rare event since Pal is strictly dependent on the presence of choline-containing TA of S. pneumoniae and this aminoalcohol (or an analogue) is necessary for pneumococcal viability in vivo. Hence, it is likely that, over evolution, the ChBD of these lytic enzymes evolved to target a unique and essential molecule in the wall.

More recently, we have expanded such experimental approach using a murine sepsis model to study the ability of pneumococcal phage amidases and lysozymes to cure bacteremia produced by an antibiotic-multiresistant 6B strain, the most common pneumococcal serotype isolated from children with bacteremia. Mice challenged with 5 × 10^7 CFU of this strain alone died within 72 h whereas a single intraperitoneal injection of 100 μg (1100 units) of Cpl-1, or Pal administered 1 h after the bacterial challenge was sufficient to effectively protect the animals. Remarkably, a synergistic effect in vivo was also observed with the combined use of 2.5 μg each of Cpl-1 and Pal [203].

It has been recently emphasized that during the long history of using phage administered by different routes in several countries there have been virtually no reports of serious complications related to their use. As phage are common entities in the environment and regularly consumed in foods, in theory, the development of neutralizing antibodies should not be a significant obstacle during the initial treatment of acute infections because the kinetics of phage action or lytic enzymes [197]. These results together with the feasibility to construct chimeric enzymes in which bacterial specificity and activities can be combined by domain swapping (see above) constitute a further step in the fight against the pneumococcus.

6. Concluding remarks and outlook

The continuous worldwide increase in streptococcal antibiotic resistance, particularly in the case of the pneumococcus, has promoted an impressive effort to develop improved vaccines able to induce a protective immune response especially in children, the elderly, and the immunocompromised. The current 23-valent polysaccharide [204] as well as conjugate pneumococcal vaccines [205], although useful, still have several limitations. In the case of the pneumococcal conjugate vaccine, for example, the number of CPs that can be conjugated to the protein moiety is a limiting factor. In addition, the pneumococcal conjugate vaccine significantly reduces the carriage of pneumococcal serotypes included in the vaccine but nonvaccine serotypes were found more frequently in vaccinees [206]. In fact, the available 7-valent conjugate vaccine has been capable to decrease episodes of otitis media due to serotypes contained in the vaccine but the proportion of nonvaccine serotypes increased after vaccination [207,208]. Similar results have also been found in preliminary trials carried out with a nonavalent conjugate vaccine [209,210]. Besides, capsular shifting is a major concern since, if
confirmed in the future, this phenomenon may preclude (or limit) the usefulness of these type of vaccines. For these reasons, efforts to determine the role in pathogenicity of several pneumococcal proteins are currently being done [211–213]. Basically, proteins to be implemented in a vaccine should lack toxicity and be relatively nonpolymorphic to allow a wide coverage. Recently, it has been shown that combinations of three different S. pneumoniae virulence proteins generate an increased protective immune response when tested in a mouse model of infection [214]. The LytB glucosaminidase might be also an immune response when tested in a mouse model of infection shown that combinations of three different S. pneumoniae virulence proteins to control pneumococcal infections. LytA seems to play a basic biological function in pathogenesis and DNA-release. The regulation of the subtle mechanisms of selected lysis requires a finely tuned control. Furthermore, DNA exchanges have been postulated to happen in natural environments in complex multispecies biofilms and LytC might also contribute to the liberation of DNA in natural environment like the upper respiratory tract during autolysis at temperatures close to 30 °C [60]. The role of biofilms in the pathogenesis of some chronic human infections is currently widely accepted. Infectious kidney stones, bacterial endocarditis, and cystic fibrosis are the best illustrative examples of infections caused by bacterial biofilms [216]. In oral streptococci such as S. mitis and S. gordonii the close contact of the cells in the biofilm is mediated by surface-exposed proteins called co-aggregation adhesins [217]. It has been suggested that the contact in pneumococcal or mixed biofilms should rely on similar surface proteins and diffusion of the CSP is, presumably, more restricted and competence should develop among small clusters of cells expressing (and responding to) the same CSP. These cells could lyse neighbouring cells (heterolysis) by expressing cell-wall degrading enzymes [133]. These observations might have medical importance since pneumococci and/or a mix of oral streptococci forming a biofilm could be eventually a normal situation during the carrier state at the human upper respiratory tract. Interestingly, this process should be also facilitated by the strategy of pneumococcus of using specific antibacterial IgA antibodies to enhance colonization on respiratory cells [218]. After degradation by a bacterial protease, IgA-Fab (antigen-binding) fragments induce the selective unmasking of the bacterial ligand (phosphorylcholine) used in the interaction with the host cell via the platelet-activating factor receptor. In this sense, the Pce esterase may play a fundamental role in regulating the number of choline residues present in the TA. These findings might challenge conventional views on vaccine design for mucosal pathogens [218].

Bacteriophage, the most abundant entities in the planet, and their corresponding lytic enzymes, may represent an enormous untapped pool of agents with which to control human pathogens. In addition, it is also most challenging to consider the current efforts to elucidate the bacteriophage control of bacterial virulence. The scarlatinal and diphtheria toxins were early and historical examples illustrating that bacterial virulence is modified by bacteriophage. The relevant genes responsible for these diseases were located in the prophage. Currently, phage infection of bacteria is increasingly associated with additional effects on bacterial virulence. Toxins are the best recognized examples of phage-encoded virulence factors in several bacterial pathogens like those causing cholera, botulism, and diphtheria. Interestingly, phage are also involved in alteration of the susceptibility of bacteria to drugs as well as on the enhancement of the capacity of bacterial transmission among humans [157]. The putative participation of phage-coded genes in pneumococcal virulence remains elusive so far. However, the availability of the total or partial nt sequence of the genomes of three similar temperate pneumococcal phage (VO1, MM1, and HB-3) can illustrate on phage variability affecting the presence (or absence) of some supernumerary gene(s), like those putatively coding for a C5Mtase, in the same region of the genomes. It is interesting the observation that genes coding for C5Mtases have been reported in the genomes of the two pneumococcal virulent strains sequenced recently [52,219] but not in the avirulent laboratory strain R6 [102]. These findings open a new window to search on the possible contribution of the phage of S. pneumoniae in genetic variation in natural populations. Further information on phage genome versatility in streptococci has been provided by a temperate bacteriophage (φscl) infecting Streptococcus canis that was induced in fluoroquinolone-treated dogs [220]. The φscl phage contains a gene (scm) coding for a novel mitogen that might be responsible for the development of canine streptococcal toxic shock syndrome and necrotizing fasciitis. The scm gene maps inside a gene that seems to code for a minor tail protein and, on the basis of its different G+C content, might have been acquired from a nonstreptococcal source. Remarkably, the genes flanking scm are highly similar to ORFs from the pneumophage MM1.

Globally, preliminary analyses of pneumococcal phage genomes together with those carried out in close bacterial systems abound with previous suggestions arguing that bacteriophage can evolve by recombinational reassortment of genes and by the acquisition of novel genetic elements supporting the so-called ‘moron.
accretion hypothesis' for the origin of phage [183]. Furthermore, the existence of regions of high nt similarity between phage and bacterial DNAs promote site-specific recombination which should allow restructuring and evolution of the bacterial and phage genomes [221]. The data discussed here suggest that the phage and bacterial lytic genes of the pneumococcal system provide examples on how both short and large DNA regions, encompassing entire domains, can be shuffled, through recombinatory events, during the evolutionary process to generate new genes coding for active but impaired proteins whose catalytic efficiencies would be improved by selective pressures in the course of evolution. This, in turn, might be a variation of the so-called short-sighted evolution that should improve the mechanisms of virulence of pathogenic microorganisms [222].

It has been well documented that the repeated switch between the lytic/lysogenic state of temperate phage results, in some cases, in the incorporation of new ORFs. In the pneumococcal model, the genes coding for lytic enzymes illustrate on these prophage excisions. However, most of these genes yield interrupted ORFs designed as remnants [223] and it has been proposed that this biological scenario might favour gene rearrangement, through recombination, which can be selected during the course of evolution, under appropriate environmental conditions. The peculiarities of these genes provide interesting insights on mechanisms of horizontal transfer and lysogenic state co-evolution.

The abundance of temperate phage in freshly pneumococcal isolates from clinical samples has been repeatedly claimed. The expression of genes involved in pathogenesis (including adhesion, invasion, and evasion from immunity) are in some cases the result of induction of a prophage harboring these genes. Since the human body is full with prophage-inducing chemicals [157] the development of appropriate technical strategies might be illuminating to study the role of phage-coded genes in lysogenic strains of the pneumococcus. Recently, autolysis in Pseudomonas aeruginosa has been examined within the context of multicellular bacterial biofilm development. During growth of biofilm a bacteriophage capable of superinfecting and lysing P. aeruginosa was detected in the fluid effluent from a continuous culture. This cell death process turned out to be prophage-mediated by a subpopulation of viable cells and is considered as an important mechanism of differentiation inside microcolonies which facilitate dispersal of a subpopulation of surviving cells [224]. Multicellular interactions in pneumococcus, either in the carrier state like in the human throat or in other scenarios in the human body, could be a common situation for this microorganism. In this circumstance, prophage of a subpopulation of S. pneumoniae can also participate in an autolytic process similar to that proposed for P. aeruginosa, suggesting a tempting model for studying mechanisms of prophage-mediated control of pneumococcal virulence. Interchange of phage and bacterial genes coding for lytic enzymes of phage and host has been well documented in the pneumococcal system (see above) but autolysis has not yet been examined within the context of multicellular bacterial biofilm development. In summary, the outlook on S. pneumoniae and its phage discussed here has been focussed on a systematic evaluation of the topics that have been the subject of our work during many years as well as to provide insights on the many exciting matters that still remain to be investigated in the near future.

Acknowledgements

Our collaborators P. García, C. Ronda, and J.L. García, who have dedicated many years of efforts and commitment to work in pneumococcus and its phage, and A. Tomasz, from The Rockefeller University, mentor of one of us (RL), deserve a special and warm recognition of long-lasting gratitude. We sincerely thank to our former and current students who have contributed through 30 years to the development of the subjects reviewed in this work. This work was supported by grants from the Dirección General de Investigación Científica y Técnica (BCM2003-00074) and from Redes Temáticas de Investigación Cooperativa (G03/103 and C03/14) (Ministerio de Sanidad y Consumo).

References


choline-binding domain of pneumococcal murein hydrolases. Gene 89, 69–75.


temperate *Streptococcus thermophilus* bacteriophage Sf621. Virolology 233, 136–148.


