



# Crystal structure of the choline-binding protein CbpJ from *Streptococcus pneumoniae*

Qian Xu, Jun-Wei Zhang, Yuxing Chen, Qiong Li<sup>\*\*</sup>, Yong-Liang Jiang<sup>\*</sup>

Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, 230026, China

## ARTICLE INFO

### Article history:

Received 29 April 2019

Accepted 6 May 2019

Available online 16 May 2019

### Keywords:

Crystal structure

*Streptococcus pneumoniae*

Choline-binding proteins

Adhesion

## ABSTRACT

The choline-binding proteins play essential roles in pneumococcal colonization and virulence. It has been suggested that the choline-binding protein J (termed CbpJ; encoded by the gene *sp\_0378*) from *Streptococcus pneumoniae* TIGR4 involves in the colonization in host and contributes to evasion of neutrophil killing. Here we report the 2.0 Å crystal structure of CbpJ in complex with choline. CbpJ consists of an N-terminal putative functional domain (N-domain) followed by a C-terminal choline-binding domain (CBD). The N-domain harbors four degenerated choline-binding repeats (CBRs) that lose the capacity of binding to choline, whereas the CBD is composed of seven typical CBRs. Further functional assays showed that the CBD contributes to the pneumococcal adhesion to human lung epithelial cell A549. These findings provide insights into the pneumococcal pathogenesis and broaden our understanding on the functions of choline-binding proteins.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

The *Streptococcus pneumoniae* is a severe human pathogen that causes serious life-threatening infections, such as sepsis, meningitis and pneumonia [1]. The morbidity and mortality of *S. pneumoniae* infections remain high worldwide. The interaction between *S. pneumoniae* and the host mucosal epithelial cells is a prerequisite for pneumococcal disease development. This process is usually mediated by *S. pneumoniae* external components, including the capsular polysaccharide, the cell wall and the proteins anchored on the cell wall [2]. The adaption of *S. pneumoniae* into the nasopharynx and blood requires the capsule biosynthesis [3], which reduces entrapment in mucus thereby allowing the pneumococcus to access the epithelial surface [4]. Once binding to the epithelial surface, *S. pneumoniae* changes from opaque to transparent variants thereby exposing their surface proteins on the cell surface that mediate host-pathogen interactions [5].

Surface proteins play a central role owing to the direct interaction with host that has the potential as vaccine antigens [6]. *S. pneumoniae* encodes various surface proteins, which can be

classified into three major groups distinguished by their mode of attachment to the host cell: the proteins harboring the LPxTG motif, lipoproteins and choline-binding proteins (CBPs) [7,8]. Eighteen proteins with LPxTG motif have been found in *S. pneumoniae* TIGR4. Structural and functional studies demonstrated that the LPxTG-anchored proteins mostly function as enzymes and adhesins [9]. Totally 42–47 of the lipoproteins were found in *S. pneumoniae* [10,11], which perform diverse functions, such as the substrate-binding proteins of ABC transporters, the adhesion proteins and chaperons. Notably, the CBPs, only found in *Streptococci*, such as *S. pneumoniae*, *Streptococcus mitis* and *Streptococcus oralis*, and their bacteriophages, play roles in cell wall remodeling and host recognition [12].

The CBPs are usually characterized by two modular structures: a highly conserved choline-binding domain (CBD) and a functional domain. The CBD generally consists of two to 12 tandem repeats rich in aromatic residues, termed the choline-binding repeats (CBRs), which allow the binding to phosphocholine molecules via non-covalent interactions. The functional domains vary a lot in 3-D structure and physiological function. The number of CBPs varies from 13 to 16 depending on different pneumococcal strains. These CBPs play important roles in cell wall physiology (the amidase LytA [13], the peptidoglycan hydrolase LytB [14]), the colonization process (the adhesin PspC [15]) and contributing to virulence by fratricide (the murein hydrolase CbpD and LytC [16–18]). To date, only

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [liqiong@ustc.edu.cn](mailto:liqiong@ustc.edu.cn) (Q. Li), [jyl@ustc.edu.cn](mailto:jyl@ustc.edu.cn) (Y.-L. Jiang).

4 full-length structures of *S. pneumoniae* CBPs, namely LytA, Pce, CbpF and LytC, are available [13,17,19,20]. LytA, Pce and LytC have a globular N-terminal catalytic domain followed by an elongated CBD. In contrast, CbpF is an autolysis regulator that shows a new modular structure assembled entirely by CBRs through its N and C-terminal domains [20].

*S. pneumoniae* TIGR4 CbpJ has been described as a putative adhesin [10]. Microarray analysis revealed that the expression of *pneumococci* gene *cbpJ* is up-regulated upon adhering to the epithelial cells, indicating that CbpJ might assist pneumococcal adhesion at the host-pathogen interaction stage [21]. *In vitro* and *in vivo* assays indicated that CbpJ functions as a virulence factor by contributing to evasion of neutrophil killing [22]. Notably, mice intranasally infected with *cbpJ* knockout strain showed an improved survival rate compared to those infected with *S. pneumoniae* TIGR4. Sequence analysis revealed that CbpJ has a domain organization similar to CbpF from *S. pneumoniae* R6. However, the structure and function of CbpJ remain unknown. Here we solved the full-length structure of CbpJ and found its CBD is required for the adhesion to the host cells. The findings provide new insights into pneumococcal virulence and pathogenesis mediated by CBPs.

## 2. Materials and methods

### 2.1. Protein preparation

The coding region of *cbpJ* without the signal peptide was cloned into the expression vector pET28a and overexpressed in *Escherichia coli* strain BL21 (DE3). The transformed cells were grown at 37 °C to an  $A_{600nm}$  of 0.8, followed by induction with 0.2 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside at 16 °C for another 22 h before harvesting. The cells were collected by centrifugation at 8000  $\times$ g for 10 min and resuspended in 30 mL lysis buffer (20 mM Tris-HCl, pH 8.0, 50 mM NaCl). The cells are disrupted by 20 min of sonication and 30 min of centrifugation at 16,000  $\times$ g. Then the clarified cell supernatant was loaded onto a column packed with DEAE sefinose equilibrated with the binding buffer (20 mM Tris-HCl, pH 8.0, 50 mM NaCl). The target protein was eluted with binding buffer containing 250 mM choline chloride and further loaded onto a Superdex 75 column pre-equilibrated with 20 mM Tris-HCl, pH 8.0, 100 mM NaCl and 10 mM choline chloride. Fractions containing the target protein were collected and concentrated to 15 mg/mL for crystallization.

### 2.2. Crystallization, data collection and processing

Crystals of CbpJ were grown at 289 K using the hanging-drop vapor-diffusion method by mixing 1  $\mu$ L protein solution with an equal volume of reservoir solution (10% PEG 20K and 0.1 M MES, pH 6.5). Crystals were transferred to a cryoprotectant solution (reservoir solution supplemented with 20% glycerol) and flash-cooled with liquid nitrogen. The diffraction data of a single crystal were collected at 100 K in a liquid-nitrogen gas stream on beamline 17U at the Shanghai Synchrotron Radiation Facility using an ADSC

Quantum 315r CCD (MAR Research). All diffraction data were integrated and scaled with HKL2000 [23].

### 2.3. Structure determination and refinement

The crystal structure of CbpJ was determined by molecular replacement with Molrep [24]. The homologous search model was the structure of CbpF from *S. pneumoniae* R6 (PDB code 2v04), which shares a 56% sequence identity over 338 residues with CbpJ. The multiple-sequence alignment between CbpJ and CbpF was performed using ClustalW [25]. The initial model was further refined by the maximum likelihood method implemented in Refmac5 [26] as part of the CCP4i program suite [27], and rebuilt interactively using the program Coot [28]. The final model was evaluated with the web service Molprobtity [29]. The crystallographic parameters are presented in Table 2. All figures showing the structure were prepared with PyMOL (<http://www.pymol.org>).

### 2.4. Construction of *S. pneumoniae* knockout, mutants and complementation strains

All *S. pneumoniae* strains used in this study are listed in Table 1. The strains were cultured at 37 °C with 5% CO<sub>2</sub> in Todd-Hewitt broth containing 0.5% yeast extract or on tryptic soy agar plates supplied with 5% sheep blood.

The in-frame *cbpJ* knockout strain was generated from *S. pneumoniae* TIGR4 by allelic replacement as described previously [30]. In brief, the upstream and downstream flanking regions of *cbpJ* were separately amplified from the genomic DNA of strain TIGR4, respectively. The Janus cassette was amplified from the genomic DNA of *S. pneumoniae* strain ST588 [31]. The three PCR products were purified from agarose gels using the DNA gel purification kit (Qiagen), and ligated by overlap PCR. The ligation fragments were then transformed into *S. pneumoniae* TIGR4 to select kanamycin-resistant colonies on blood agar plates as described previously [32]. The replacement of the whole *cbpJ* coding region with the Janus cassette was detected by PCR and DNA sequencing. The strains with in-frame deletion of the coding regions of *cbpJ* N-domain and CBD were generated and detected in the same manner as *cbpJ* knockout strain.

A Zn<sup>2+</sup> inducible pJWV25 plasmid [33] was used for complementation. The coding regions of full-length *cbpJ* was cloned into pJWV25. Then the plasmid was transformed into the *cbpJ* knockout strain similar to the PCR products. The strain was then grown on blood agar plates supplemented with 1  $\mu$ g/mL tetracycline. Colonies harboring the target plasmids were further confirmed by PCR and DNA sequencing.

### 2.5. Cell adhesion assays

Pneumococcal adhesion assays with epithelial cell A549 were performed as described previously [34]. Human lung epithelial cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 2 mM L-Glutamine, 10% v/v fetal bovine serum (FBS), 5  $\mu$ g/mL penicillin and 100  $\mu$ g/mL

**Table 1**  
Bacterial strains used in this study.

Strains	Description
TIGR4	<i>S. pneumoniae</i> clinical isolate, serotype 4, encapsulated; <i>cbpJ</i> <sup>+</sup> , Strep <sup>s</sup> , Kan <sup>s</sup>
KO	<i>S. pneumoniae</i> TIGR4 derivative; $\Delta$ <i>cbpJ</i> ::Janus cassette; Kan <sup>r</sup>
$\Delta$ NT	<i>S. pneumoniae</i> TIGR4 derivative; $\Delta$ <i>cbpJ</i> -N domain::Janus cassette; Kan <sup>r</sup>
$\Delta$ CT	<i>S. pneumoniae</i> TIGR4 derivative; $\Delta$ <i>cbpJ</i> -CBD::Janus cassette; Kan <sup>r</sup>
Complementation	KO strain complemented with CbpJ-pJWV25; Kan <sup>r</sup> , Tet <sup>r</sup>

**Table 2**  
Crystal parameters, data collection, and structure refinement.

	Cbpj
<b>Data collection</b>	
Wavelength (Å)	0.97,923
Space group	$P2_12_12_1$
Unit cell parameters	
$a, b, c$ (Å)	54.52, 77.38, 187.85
$\alpha, \beta, \gamma$ (°)	90.00, 90.00, 90.00
Resolution range (Å) <sup>a</sup>	50.0–2.0 (2.1–2.0)
Unique reflections	51,107 (3946)
Completeness (%)	93.4 (73.5)
$\langle I/\sigma(I) \rangle$	10.3 (3.4)
$R_{\text{merge}}^b$ (%)	7.2 (19.9)
<b>Structure refinement</b>	
Resolution range (Å)	48.7–2.0
$R_{\text{factor}}^c/R_{\text{free}}^d$ (%)	21.4/25.6
Number of protein atoms	4909
Number of water atoms	319
RMSD <sup>e</sup> bond lengths (Å)	0.013
RMSD bond angles (°)	1.289
Mean B factors (Å <sup>2</sup> )	42.37
Ramachandran plot (residues, %) <sup>f</sup>	
Most favored (%)	96.88
Additional allowed (%)	3.12
Protein Data Bank entry	6JYX

<sup>a</sup> The values in parentheses refer to statistics in the highest bin.

<sup>b</sup>  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the intensity of an observation, and  $\langle I(hkl) \rangle$  is the mean value for its unique reflection. Summations are over all reflections.

<sup>c</sup>  $R_{\text{factor}} = \sum_h |F_o(h) - F_c(h)| / \sum_h F_o(h)$ , where  $F_o$  and  $F_c$  are the observed and calculated structure-factor amplitudes, respectively.

<sup>d</sup>  $R_{\text{free}}$  was calculated with 5% of the data excluded from the refinement.

<sup>e</sup> RMSD from ideal values.

<sup>f</sup> Categories were defined by MolProbity.

streptomycin. After growing in 5% CO<sub>2</sub> at 37 °C for 24 h, the cell monolayers were washed five times with sterile PBS and then incubated with  $\sim 10^7$  CFU/mL *pneumococci*, which were grown in THY medium to A<sub>620 nm</sub> of 0.3–0.4 and resuspended in the DMEM medium without serum. After 3 h of incubation at 37 °C, the infected monolayers were washed five times with sterile PBS in order to remove unbound bacteria. Then, the cells were treated with 200  $\mu$ L 0.25% trypsin and 0.02% EDTA, and lysed by adding 0.025% Triton X-100. The number of adhered bacteria were counted after plating serial dilutions of the bacterial suspensions on agar plates containing 5% sheep blood.

### 3. Results

#### 3.1. Overall structure of Cbpj-choline complex

The 2.0 Å resolution structure of Cbpj-choline complex was determined by molecular replacement using CbpF as a search model [20]. Each molecule contains residues Gly32–Arg332, except for the N-terminal signal peptide (Met1–Glu31) and a segment covering residues Glu139–Glu146, which are not modeled due to the poor electron density. Cbpj folds into two well-defined modules: a shell-shaped N-domain (Trp33–Thr163), and a canonical C-terminal CBD (Gly164–Arg332) previously described in other CBPs (Fig. 1A). The overall structure shows an elongated shape and spans 110 Å along its sequence. The cylindrical CBD consists of seven CBRs termed P1 to P7 (Gly164–Val315), followed by a C-terminal tail of 17 amino acids (Asn316–Arg332). Each choline molecule binds to the cavity between two neighboring CBRs. As a result, totally seven choline molecules were found in the structure. The N-domain of Cbpj is assembled entirely by four modified CBRs, termed dP1–dP4, each of which folds into a twisted  $\beta$ -hairpin flanked by flexible loops. The dP1 and dP2 consist of 20 and 21 residues, respectively,

similar to the canonical CBRs in length, whereas the dP3 and dP4 are much longer in length that consist of 44 and 38 residues, respectively (Fig. 1B).

#### 3.2. The CBD of Cbpj

The CBD consists of seven CBRs, with the consensus GW-X<sub>6</sub>-WYY- $\Phi$ -X<sub>3–4</sub>-GXXMX<sub>2</sub>, where X represents any amino acid and  $\Phi$  represents a hydrophobic amino acid (Fig. 1B), in agreement with previous reports [19]. Each repeat comprises a symmetrical  $\beta$ -hairpin, followed by a loop region, in which one choline molecule is recognized at the cleft between two consecutive CBRs. Generally, three conserved aromatic residues, including two tryptophan residues from one CBR and one tyrosine from the next CBR stabilize the choline molecule by cation- $\pi$  interactions (Fig. 2A). In addition, a methionine is usually located at the bottom of the pocket. Notably, the C-terminal tail also adopts a  $\beta$ -hairpin structure that contributes to the binding of a choline molecule via Tyr322.

The CBD of Cbpj is structurally similar to CbpF from *S. pneumoniae* R6 [20], with a root mean square deviation (r.m.s.d.) of 1.6 Å over 169 C $\alpha$  atoms (Fig. 2B). However, the CBD of Cbpj contains six canonical CBRs in addition to one non-canonical CBR, whereas CbpF contains five canonical and two non-canonical CBRs. Apart from the seven typical CBRs, there is an insertion between P1 and P2 of Cbpj (Fig. 2B). This insertion of 7 residues (Glu179–Leu186) is rather hydrophilic and has a  $\beta$ -hairpin structure that points outwards from the cylinder, leaving no interaction with the choline molecule (Fig. 1A). Notably, CbpF also has an insertion at a similar position, but forms a  $\beta$ -turn structure (Fig. 2B).

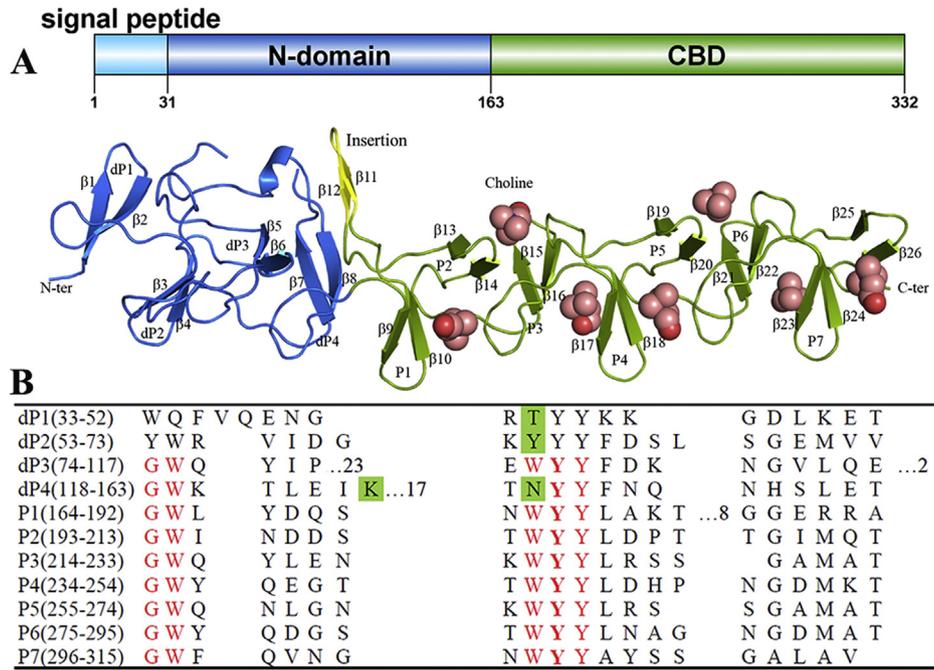
#### 3.3. The N-domain of Cbpj

A DALI [35] search revealed that the N-domain of Cbpj resembles the CBDs of other CBPs. The top hits include the N-domain of *S. pneumoniae* R6 CbpF [20] (PDB code 2v05, Z-score 18.1, r.m.s.d. of 1.5 Å over 114 C $\alpha$  atoms), the CBD of *S. pneumoniae* R6 CbpE [19] (PDB code 2bib, Z-score 9.3, r.m.s.d. of 2.3 Å over 80 C $\alpha$  atoms) and the CBD of *S. pneumoniae* TIGR4 LytA [13] (PDB code 4x36, Z-score 8.7, r.m.s.d. of 2.3 Å over 81 C $\alpha$  atoms). The N-domain of Cbpj is superimposable to these CBRs except that it does not fulfill the choline-binding requirements. At least one of the crucial tryptophan residues is replaced by residues Thr42, Tyr61 and Asn152 in dP1, dP2 and dP4, respectively (Fig. 1B). Despite that the cavity between dP3 and dP4 fulfills all of the choline-binding requirements, it lost the choline-binding capability due to the blockage of the choline-binding site (CBS) by Lys125 (Fig. 2C).

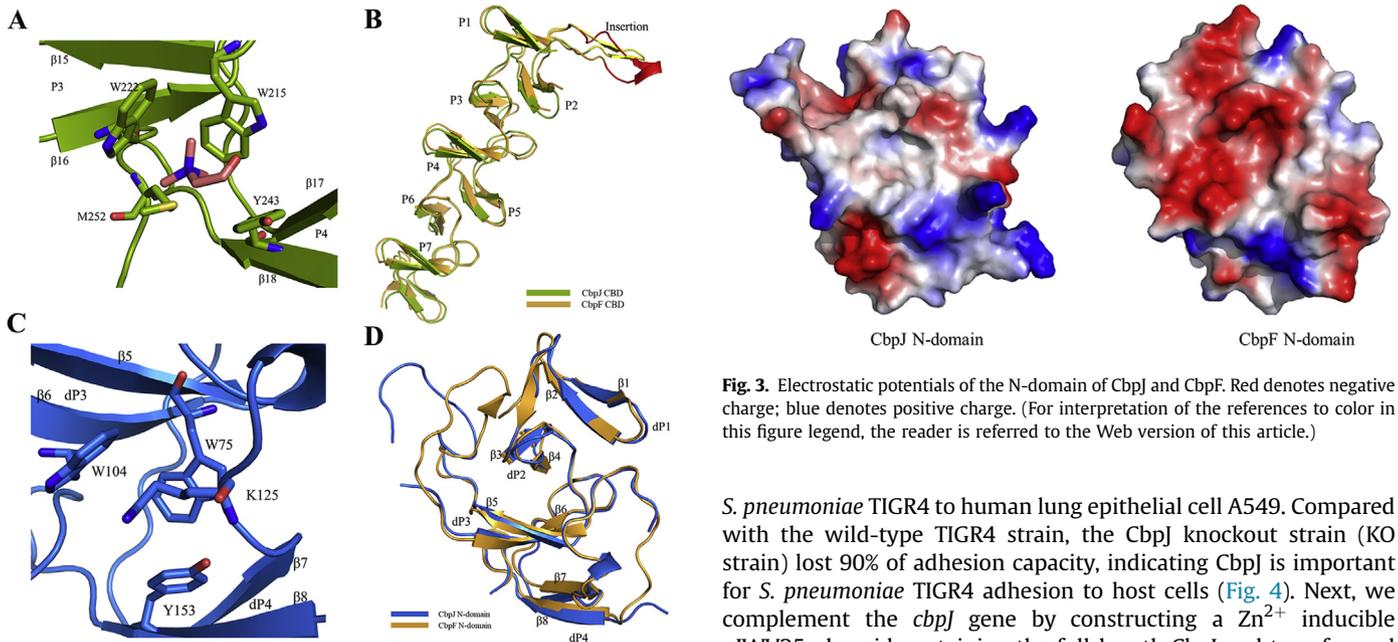
Superposition of the N-domains of Cbpj and CbpF reveals that they share a highly similar overall structure (Fig. 2D). Both of the N-domains are formed by a series of non-consensus CBRs modified by insertions (Fig. 2C). However, they differ dramatically by the various conformations of the connecting loops. Moreover, the cavity of N-domain of Cbpj is mainly hydrophobic (Fig. 3), whereas the cavity of N-domain of CbpF is mostly negatively charged to favor its binding to the peptidoglycan [20].

#### 3.4. The CBD is crucial for *S. pneumoniae* adhesion to human lung epithelial cell A549

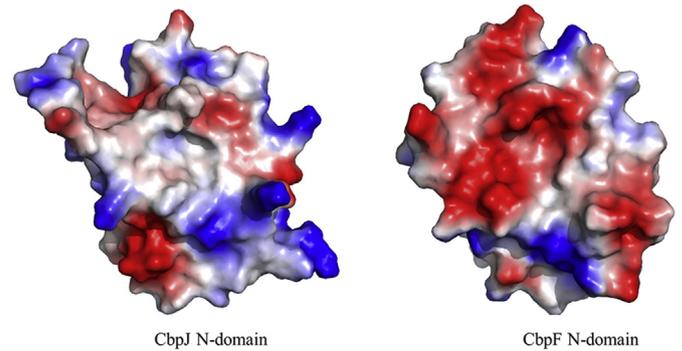
The structure of Cbpj reveals distinct features among CBPs of known structure. *In vitro* and *in vivo* assays indicated that Cbpj functions as an adhesin and a virulence factor by contributing to evasion of neutrophil killing. Despite the important structural and functional insights provided by these studies, the function of *S. pneumoniae* TIGR4 Cbpj remains elusive. To elucidate the molecular function of Cbpj, we detected the adhesion ability of



**Fig. 1.** Overall structure of CbpJ. (A) Domain organization and cartoon representation of the full-length CbpJ in complex with choline. The N-domain and CBD of CbpJ were colored in blue and green, respectively. Individual choline molecules bound to the CBSs are shown as space-filled balls (salmon). Canonical CBRs are labelled as P1–P7 and modified CBRs are dP1–dP4. The seven repeats (P1–P7) are folded in a super helical arrangement, allowing binding to choline, whereas P1 is modified by an insertion (yellow). (B) Multiple-sequence alignment of CBRs in CbpJ. Green frames depict residues Thr42, Tyr61, Lys125 and Asn152 in dP1, dP2, dP3 and dP4, respectively. Five conserved residues of consensus CBRs are shown in red. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



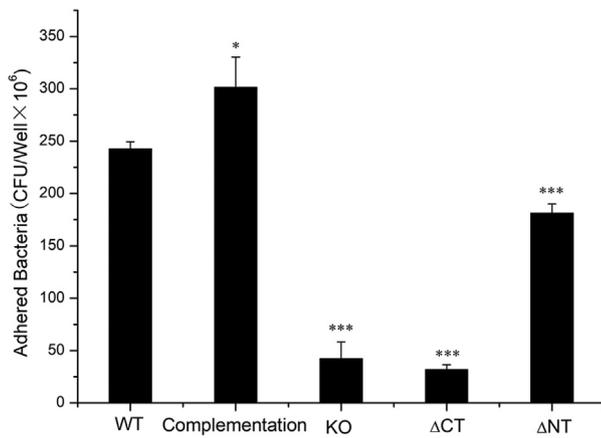
**Fig. 2.** The N-domain and CBD of CbpJ. (A) Three-dimensional structure of a canonical CBS between P2 and P3 of CbpJ. The choline molecule and the interacting residues are displayed as sticks. (B) Superposition of the CBD of CbpJ (green) against that of CbpF (orange, PDB code 2vyu). The insertion of 7 residues (Glu179–Leu186) between P1 and P2 of CbpJ is colored in yellow, whereas the corresponding insertion of CbpF is colored in red. (C) Three-dimensional structure of a degenerated CBS between dP3 and dP4. Three residues fulfilling the choline-binding requirements (Trp 75, Trp 104 and Try 153) and Lys125 blocking the CBS are displayed as sticks, respectively. (D) Superposition of the N-domain of CbpJ (blue) against that of CbpF (yellow, PDB code 2v05). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 3.** Electrostatic potentials of the N-domain of CbpJ and CbpF. Red denotes negative charge; blue denotes positive charge. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

*S. pneumoniae* TIGR4 to human lung epithelial cell A549. Compared with the wild-type TIGR4 strain, the CbpJ knockout strain (KO strain) lost 90% of adhesion capacity, indicating CbpJ is important for *S. pneumoniae* TIGR4 adhesion to host cells (Fig. 4). Next, we complement the *cbpJ* gene by constructing a Zn<sup>2+</sup> inducible pJWV25 plasmid containing the full-length CbpJ and transformed into the KO strain. The results showed that the complementary strain recovered its adhesion ability (Fig. 4).

To decipher which domain of CbpJ contributes to the adhesion, we deleted the regions coding for the N- and C-terminal domains of CbpJ, respectively, for comparing the adhesion ability. The strain with an in-frame deletion of N-domain coding region ( $\Delta$ N) still retains approximately 70% adhesion activity compared to the wild-type, whereas the strain with an in-frame deletion of CBD coding region ( $\Delta$ C) lost 90% adhesion activity compared to the wild-type. All these results suggested that both the two domains of CbpJ are



**Fig. 4.** Quantification of *S. pneumoniae* TIGR4 and its isogenic *cbpJ* mutants adhesion to human lung epithelial A549 cells. The strains used are listed in Table 1. WT, wild-type *S. pneumoniae* TIGR4, used as the positive control; KO, the *cbpJ* knockout strain; ΔCT: the strain with in-frame deletion of the coding region of *cbpJ* CBD; ΔNT: the strain with in-frame deletion of the coding region of *cbpJ* N-domain. The data are presented as means ± standard deviation from the results of three independent assays. Two-tailed student's *t*-test was used to determine significance of the difference. The *p* values of <0.05 and <0.05 are indicated with \* and \*\*\*, respectively.

required for the adhesion to the host cells, with the CBD of a more significant role. Altogether, these data demonstrated for the first time that, in addition to favor the attachment of CBPs to its cell wall, the CBD can assist pneumococcal adhesion to respiratory epithelial cells.

#### 4. Discussion

In this study, we presented the crystal structure of CbpJ-choline complex from *S. pneumoniae* TIGR4. The N-domain and CBD are structurally similar to each other, despite that the N-domain lost the choline-binding capability. Both domains of CbpJ contribute to the adhesion of *S. pneumoniae* TIGR4 to the respiratory epithelial cells, which is the key step of pneumococcal colonization and pathogenesis. Moreover, structural characterization of some CBPs revealed that modular arrangement in CBDs are critical for their specific features, such as their choline-binding pattern, the catalytic properties and their localization on the cell surface [10,13,17].

Despite variations and insertions of the sequence, the N-domain of CbpJ possesses the modified CBRs without choline-binding capability. Here we found that the N-domain also contributes a part to the adhesion of *S. pneumoniae* TIGR4 to the host. These findings indicated that the N-domain and CBD might evolve from the same ancestor by duplication. During evolution, sequence variations resulted in a degenerated CBD at the N-terminus.

Structural analysis showed that CbpJ structurally resembles CbpF, but differ in the N-domains. The sequence similarity of CbpF and CbpJ indicated that a common ancestral *S. pneumoniae* acquired the genes by duplication [22]. Sequence analysis also showed CbpC in *S. pneumoniae* TIGR4 is likely to show an architecture similar to that of CbpJ or CbpF, which possesses a typical CBD and an N-domain of degenerated CBRs [20]. We propose that these CbpJ homologs might represent a new subfamily of CBPs. Altogether, these findings provide new insights into the pneumococcal pathogenesis and broaden our understanding of the functions of CBPs.

#### Funding

This work is supported by the National Natural Science

Foundation of China (Grants No. 31670732 and 31770787).

#### Acknowledgements

We thank the staff at the Shanghai Synchrotron Radiation Facility for technical assistance.

#### References

- [1] D. Bogaert, R. de Groot, P. Hermans, *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease, *Lancet Infect. Dis.* 4 (2004) 144–154.
- [2] I. Pérez-Dorado, S. Galan-Bartual, J. Hermoso, Pneumococcal surface proteins: when the whole is greater than the sum of its parts, *Mol. Oral Microbiol.* 27 (2012) 221–245.
- [3] J.N. Weiser, R. Austrian, P.K. Sreenivasan, H.R. Masure, Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization, *Infect. Immun.* 62 (1994) 2582–2589.
- [4] A.L. Nelson, A.M. Roche, J.M. Gould, K. Chim, A.J. Ratner, J.N. Weiser, Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance, *Infect. Immun.* 75 (2007) 83–90.
- [5] E.I. Tuomanen, R. Austrian, H.R. Masure, Pathogenesis of pneumococcal infection, *N. Engl. J. Med.* 332 (1995) 1280–1284.
- [6] A. Kadioglu, J.N. Weiser, J.C. Paton, P.W. Andrew, The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease, *Nat. Rev. Microbiol.* 6 (2008) 288.
- [7] S. Bergmann, S. Hammerschmidt, Versatility of pneumococcal surface proteins, *Microbiology* 152 (2006) 295–303.
- [8] A. Zomer, P.W. Hermans, H.J. Bootsma, Non-adhesive Surface Proteins of *Streptococcus pneumoniae*, *Streptococcus pneumoniae*, Elsevier, 2015, pp. 231–244.
- [9] J. Löfling, V. Vimberg, P. Battig, B. Henriques-Normark, Cellular interactions by LPxTG-anchored pneumococcal adhesins and their streptococcal homologues, *Cell Microbiol.* 13 (2011) 186–197.
- [10] H. Tettelin, K.E. Nelson, I.T. Paulsen, J.A. Eisen, T.D. Read, S. Peterson, J. Heidelberg, R.T. DeBoy, D.H. Haft, R.J. Dodson, Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*, *Science* 293 (2001) 498–506.
- [11] J. Hoskins, W.E. Alborn, J. Arnold, L.C. Blaszcak, S. Burgett, B.S. DeHoff, S.T. Estrem, L. Fritz, D.-J. Fu, W. Fuller, Genome of the bacterium *Streptococcus pneumoniae* strain R6, *J. Bacteriol.* 183 (2001) 5709–5717.
- [12] J.L. GARCÍA, A.R. SÁNCHEZ-BEATO, F.J. MEDRANO, R. LÓPEZ, Versatility of choline-binding domain, *Microb. Drug Res.* 4 (1998) 25–36.
- [13] Q. Li, W. Cheng, C. Morlot, X.-H. Bai, Y.-L. Jiang, W. Wang, D.I. Roper, T. Vernet, Y.-H. Dong, Y. Chen, Full-length structure of the major autolysin LytA, *Acta Crystallogr. D Biol. Crystallogr.* 71 (2015) 1373–1381.
- [14] X.-H. Bai, H.-J. Chen, Y.-L. Jiang, Z. Wen, Y. Huang, W. Cheng, Q. Li, L. Qi, J.-R. Zhang, Y. Chen, Structure of pneumococcal peptidoglycan hydrolase LytB reveals insights into the bacterial cell wall remodeling and pathogenesis, *J. Biol. Chem.* 289 (2014) 23403–23416.
- [15] S. Dave, A. Brooks-Walter, M.K. Pangburn, L.S. McDaniel, PspC, a pneumococcal surface protein, binds human factor H, *Infect. Immun.* 69 (2001) 3435–3437.
- [16] J.-P. Claverys, L.S. Håvarstein, Cannibalism and fratricide: mechanisms and reasons d'etre, *Nat. Rev. Microbiol.* 5 (2007) 219.
- [17] I. Pérez-Dorado, A. González, M. Morales, R. Sanles, W. Striker, W. Vollmer, S. Mobashery, J.L. García, M. Martínez-Ripoll, P. García, Insights into pneumococcal fratricide from the crystal structures of the modular killing factor LytC, *Nat. Struct. Mol. Biol.* 17 (2010) 576.
- [18] L. Kausmally, O. Johnsborg, M. Lund, E. Knutsen, L.S. Håvarstein, Choline-binding protein D (CbpD) in *Streptococcus pneumoniae* is essential for competence-induced cell lysis, *J. Bacteriol.* 187 (2005) 4338–4345.
- [19] J.A. Hermoso, L. Lagartera, A. González, M. Stelter, P. García, M. Martínez-Ripoll, J.L. García, M. Menéndez, Insights into pneumococcal pathogenesis from the crystal structure of the modular teichoic acid phosphorylcholine esterase Pce, *Nat. Struct. Mol. Biol.* 12 (2005) 533.
- [20] R. Molina, A. González, M. Stelter, I. Pérez-Dorado, R. Kahn, M. Morales, S. Campuzano, N.E. Campillo, S. Mobashery, J.L. García, Crystal structure of CbpF, a bifunctional choline-binding protein and autolysis regulator from *Streptococcus pneumoniae*, *EMBO Rep.* 10 (2009) 246–251.
- [21] C.J. Orihuela, J.N. Radin, J.E. Sublett, G. Gao, D. Kaushal, E.I. Tuomanen, Microarray analysis of pneumococcal gene expression during invasive disease, *Infect. Immun.* 72 (2004) 5582–5596.
- [22] M. Yamaguchi, K. Goto, Y. Hirose, Y. Yamaguchi, T. Sumitomo, M. Nakata, K. Nakano, S. Kawabata, Identification of evolutionarily conserved virulence factor by selective pressure analysis of *Streptococcus pneumoniae*, *Commun. Biol.* 2 (2019) 96.
- [23] Z. Otwinowski, W. Minor, [20] Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276 (1997) 307–326.
- [24] A. Vagin, A. Teplyakov, MOLREP: an automated program for molecular replacement, *J. Appl. Crystallogr.* 30 (1997) 1022–1025.
- [25] J.D. Thompson, D.G. Higgins, T.J. Gibson, W. CLUSTAL, Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic*

- Acids Res. 22 (1994) 4673–4680.
- [26] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D Biol Crystallogr.* 53 (1997) 240–255.
- [27] C.P. Collaborative, The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D Biol Crystallogr.* 50 (1994) 760.
- [28] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol Crystallogr.* 60 (2004) 2126–2132.
- [29] I.W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, L.W. Murray, W.B. Arendall 3rd, J. Snoeyink, J.S. Richardson, D.C. Richardson, MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Res.* 35 (2007) W375–W383.
- [30] C. Sung, H. Li, J. Claverys, D. Morrison, An *rpsL* cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae*, *Appl. Environ. Microbiol.* 67 (2001) 5190–5196.
- [31] L. Lu, Y. Ma, J.-R. Zhang, *Streptococcus pneumoniae* recruits complement factor H through the amino terminus of CbpA, *J. Biol. Chem.* 281 (2006) 15464–15474.
- [32] A.L. Bricker, A. Camilli, Transformation of a type 4 encapsulated strain of *Streptococcus pneumoniae*, *FEMS Microbiol. Immunol.* 172 (1999) 131–135.
- [33] A. Eberhardt, L.J. Wu, J. Errington, W. Vollmer, J.W. Veening, Cellular localization of choline-utilization proteins in *Streptococcus pneumoniae* using novel fluorescent reporter systems, *Mol. Microbiol.* 74 (2009) 395–408.
- [34] A. Ring, J.N. Weiser, E.I. Tuomanen, Pneumococcal trafficking across the blood-brain barrier. Molecular analysis of a novel bidirectional pathway, *J. Clin. Investig.* 102 (1998) 347–360.
- [35] L. Holm, L.M. Laakso, Dali server update, *Nucleic Acids Res.* 44 (2016) W351–W355.