International Journal of Biological Macromolecules

Determination of the capsular polysaccharide structure of the Klebsiella pneumoniae ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

--Manuscript Draft--

Manuscript Number:	IJBIOMAC_2020_583R1
Article Type:	Research Paper
Section/Category:	Carbohydrates, Natural Polyacids and Lignins
Keywords:	Klebsiella pneumoniae KPB-1; capsular polysaccharide structure; glycosyltransferases
Corresponding Author:	Paola Cescutti Universita degli Studi di Trieste Trieste, Italy
First Author:	Barbara Bellich
Order of Authors:	Barbara Bellich
	Cristina Lagatolla
	Roberto Rizzo
	Marco Maria D'Andrea
	Gian Maria Rossolini
	Paola Cescutti
Abstract:	Klebsiella pneumonia e strain KPB-1 was isolated in early 2011 from the pleural fluid of an inpatient admitted at an Italian hospital. It was characterized to produce the KPC-3 carbapenemase and to belong to sequence type 512, a derivative of sequence type 258 clade II characterized by the cps-2 gene cluster. The K-antigen of K. pneumoniae KPB-1 was purified and its structure determined by using GLC-MS of appropriate carbohydrate derivatives and 1D and 2D NMR spectroscopy of the native polysaccharide. All the collected data demonstrated the following repeating unit for the K. pneumonia e KPB-1 capsular polysaccharide: [3)-[α-L-Rha p -(1->4]-α-D-Gal p A-(1->2)-α-L-Rha p -(1->2)-α-L-Rha p -(1->2)-α-L-Rha p -(1->3)-β-D-Gal p -(1->]n The reactions catalyzed by each glycosyltransferase in the cps-2 gene cluster were assigned on the basis of structural homology with other Klebsiella K antigens.
Suggested Reviewers:	Rafal Mostowy Mostowy Rafal.mostowy@uj.edu.pl expert in microbial genomics
	Andrzej Gamian gamian@immuno.iitd.pan.wroc.pl Expert in polysaccharides structure determination
	Yuriy Knirel knirel@ioc.ac.ru Expert in the determination of polysaccharides structure
	Joanna Kubler-Kielb kielbj@mail.nih.gov Expert in the determination of polysaccharides primary structure
	Kathryn E. Holt Kathryn.Holt@monash.edu computational biologist specialised in infectious disease genomics
Opposed Reviewers:	
Response to Reviewers:	Dear Editor,

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

I read the responses of the reviewers and prepared our reply that you can find in the uploaded files.
In the manuscript, the section “authors contributions” was changed according to the instructions found on the website EVISE.
I hope that now the manuscript entitled: “Determination of the capsular polysaccharide structure of the Klebsiella pneumoniae ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions” is ready for publication in Int. J. Biol. Macromol.
Thanking you in advance for your attention, I send my best wishes
Sincerely,

Paola Cescutti

RESPONSE TO REVIEWERS
-Reviewer 1
The manuscript reports the results of the structural analysis of a novel capsular polysaccharide from pathogen Klebsiella pneumoniae and the assignment of glycosyltransferases functions inferred by comparison with similar genes producing structurally related polysaccharides from the other strains of the same organism. The work is well performed and described. I could not find any errors.
-Reply to Reviewer 1
We thank the reviewer. We did not change the manuscript since no criticism was raised.

-Reviewer 2
They have characterized a protein called Capsular polysaccharide structure from Klebsiella pneumoniae ST512 using GLC-MS and NMR. The same group has already investigated the CPS gene cluster of Klebsiella pneumoniae strain KPB-1. They also predicted the glycosyltransferases functions to a specific glycosidic linkage formation. They also compared and showed the structure of repeating units of CPS of KPB-1 with already reported strains. They have characterized the identified proteins in almost all required ways for its conformation. The graphs plotted with observed peaks are acceptable and easily understandable. The way in which they have presented the figures easy to follow the write up.
Though there is nothing novel in their findings, the manuscript adds one more structural information’s on Capsular Polysaccharide. Similar structural conformations have been already reported using different strains in this field. Hence, this can add a value to this particular field.
-Reply to Reviewer 2
Reviewer 2 gave a positive feedback overall, concluding his revision by saying that the manuscript can add value to the bacterial polysaccharide field. We did not complete understood some of his/her comments. For example, we have characterised a polysaccharide, and not a protein, but probably this is a typo. Moreover, it is the first time that we report the structure of the CPS produced by K. pneumoniae KPB-1, while he/she wrote that we have already investigated it. But maybe this is also a typo, because we have characterised the capsular polysaccharide of the strain KK207-2. Anyway, the latter strain has a completely different capsular polysaccharide than that of KKBP-1.
We do not agree with Reviewer 2 when he/she says that “…there is nothing novel in their findings”. In our paper we described for the first time the structure of the capsular polysaccharide produced by a K. pneumoniae clinical strain belonging to ST512; we also assigned the functions of the glycosyltransferases of the CPS gene cluster for the first time. Both these information are not present in the literature.
Since also Reviewer 2 did not ask for specific changes of the manuscript, we did not change it.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Dear Editor,

I read the responses of the reviewers and prepared our reply that you can find in the uploaded files.

In the manuscript, the section “authors contributions” was changed according to the instructions found on the website EVISE.

I hope that now the manuscript entitled: “Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions” is ready for publication in Int. J. Biol. Macromol.

Thanking you in advance for your attention, I send my best wishes

Sincerely,

Paola Cescutti

RESPONSE TO REVIEWERS

-Reviewer 1
The manuscript reports the results of the structural analysis of a novel capsular polysaccharide from pathogen Klebsiella pneumoniae and the assignment of glycosyltransferases functions inferred by comparison with similar genes producing structurally related polysaccharides from the other strains of the same organism. The work is well performed and described. I could not find any errors.

-Reply to Reviewer 1
We thank the reviewer. We did not change the manuscript since no criticism was raised.

-Reviewer 2
They have characterized a protein called Capsular polysaccharide structure from *Klebsiella pneumoniae* ST512 using GLC-MS and NMR. The same group has already investigated the CPS gene cluster of *Klebsiella pneumoniae* strain KPB-1. They also predicted the glycosyltransferases functions to a specific glycosidic linkage formation. They also compared and showed the structure of repeating units of CPS of KPB-1 with already reported strains. They have characterized the identified proteins in almost all required ways for its conformation. The graphs plotted with observed peaks are acceptable and easily understandable. The way in which they have presented the figures easy to follow the write up.

Though there is nothing novel in their findings, the manuscript adds one more structural information's on Capsular Polysaccharide. Similar structural conformations have been already reported using different strains in this field. Hence, this can add a value to this particular field.

-Reply to Reviewer 2
Reviewer 2 gave a positive feedback overall, concluding his revision by saying that the manuscript can add value to the bacterial polysaccharide field. We did not complete understood some of his/her comments. For example, we have characterised a polysaccharide, and not a protein, but probably this is a typo. Moreover, it is the first time that we report the structure of the CPS produced by K. pneumoniae KPB-1, while he/she wrote that we have already investigated it. But maybe this is also a typo, because we have characterised the capsular polysaccharide of the strain KK207-2. Anyway, the latter strain has a completely different capsular polysaccharide than that of KKBP-1.

We do not agree with Reviewer 2 when he/she says that "…there is nothing novel in their findings". In our paper we described for the first time the structure of the capsular polysaccharide produced by a K. pneumoniae clinical strain belonging to ST512; we also assigned the functions of the glycosyltransferases of the CPS gene cluster for the first time. Both these information are not present in the literature.
Since also Reviewer 2 did not ask for specific changes of the manuscript, we did not change it.

HIGHLIGHTS

- The CPS produced by a *K. pneumoniae* ST512 strain was characterised for the first time
- It consists of a branched hexasaccharide repeating unit rich in rhamnose
- Glycosyltransferases functions in the CPS gene cluster were identified

Abstract

Klebsiella pneumoniae strain KPB-1 was isolated in early 2011 from the pleural fluid of an inpatient admitted at an Italian hospital. It was characterized to produce the KPC-3 carbapenemase and to belong to sequence type 512, a derivative of sequence type 258 clade II characterized by the *cps-2* gene cluster. The K-antigen of *K. pneumoniae* KPB-1 was purified and its structure determined by using GLC-MS of appropriate carbohydrate derivatives and 1D and 2D NMR spectroscopy of the native polysaccharide. All the collected data demonstrated the following repeating unit for the *K. pneumoniae* KPB-1 capsular polysaccharide:

[3)-α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→]$_n$
4
↑
1
α-L-Rha*p*

The reactions catalyzed by each glycosyltransferase in the *cps-2* gene cluster were assigned on the basis of structural homology with other *Klebsiella* K antigens.

Manuscript File Click here to view linked References

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich[a], Cristina Lagatolla[a], Roberto Rizzo[a], Marco Maria D'Andrea[b,c] Gian Maria Rossolini[d,e] and Paola Cescutti[a*]

[a] *Department of Life Sciences, University of Trieste, 34127 Trieste, Italy*

[b] *Department of Medical Biotechnologies, University of Siena, Siena, Italy*

[c] *Department of Biology, University of Rome "Tor Vergata", Rome, Italy*

[d] *Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy*

[e] *Microbiology and Virology Unit, Florence Careggi University Hospital, Florence, Italy*

* Corresponding author. Department of Life Sciences, Bldg C11, University of Trieste, Via Licio Giorgieri 1, I-34127 Trieste, Italy. Tel.: +39 040 5588755.

E-mail address: pcescutti@units.it

Abstract

Klebsiella pneumoniae strain KPB-1 was isolated in early 2011 from the pleural fluid of an inpatient admitted at an Italian hospital. It was characterized to produce the KPC-3 carbapenemase and to belong to sequence type 512, a derivative of sequence type 258 clade II characterized by the *cps-2* gene cluster. The K-antigen of *K. pneumoniae* KPB-1 was purified and its structure determined by using GLC-MS of appropriate carbohydrate derivatives and 1D and 2D NMR spectroscopy of the native polysaccharide. All the collected data demonstrated the following repeating unit for the *K. pneumoniae* KPB-1 capsular polysaccharide:

[3)-α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→]$_n$
4
↑
1
α-L-Rha*p*

The reactions catalyzed by each glycosyltransferase in the *cps-2* gene cluster were assigned on the basis of structural homology with other *Klebsiella* K antigens.

Keywords: *Klebsiella pneumoniae* KPB-1; capsular polysaccharide structure; NMR; glycosyltransferases.

1. Introduction

Klebsiella pneumoniae are Gram-negative encapsulated rods of the *Enterobacteriaceae* family and are mainly responsible for urinary-tract infections, meningitis and pneumonia in hospitalised patients, infants and third-world populations [1]. The global emergence and spreading of carbapenem-resistant *K. pneumoniae* (CR-Kp) has recently become a serious health threat, since these antibiotics are among the last-resort drugs for treatment of infections caused by multi-resistant strains of this species. The major carbapenem resistance mechanism among CR-Kp consists in the production of different types of carbapenemases (e. g. those of the KPC-, VIM-, NDM- and OXA-48-lineages), with KPC producing strains representing one of the most challenging CR-Kp, given their extended antibiotic resistance phenotypes, the ability to rapidly disseminate in hospital settings and the association with significant morbidity and mortality rates. Indeed, in recent years KPC-producing *K. pneumoniae* (KPC-Kp) experienced a global dissemination largely sustained by a restricted number of clones, and are now endemic in several countries [2]. Different clones of KPC-Kp have been described, and members of clonal group (CG) 258 are probably the most successful and widespread, being reported in several areas of North and South America, Israel, Europe and Asia [3-7]. This observation has been also confirmed by a recent large-scale European survey which showed that about 19% of *K. pneumoniae* strains obtained from 455 hospitals across Europe and neighbouring countries were KPC-Kp of CG258 [8]. CG258 includes strains of the sequence type (ST) 258 and some related variants [2] and represent the most diffused KPC-Kp in several countries [8] including Italy [9].

K. pneumoniae has been classified in more than 80 serotypes based on the serological reaction with the capsular polysaccharide (CPS, or K-antigen), underlying the CPS structural diversity which characterises this species. Being exposed on the cell surface, different CPSs confer diverse antigenic properties within the same species, with relevant consequences to bacterial virulence. Apart from serotyping, more modern methods for classifying the different capsules are based on genotyping of single genetic targets such as the conserved *wzi* [10] and *wzc* [11] genes, or on the analysis of the full-length *cps* locus [12]. An extended analysis of the *cps* clusters sequences from 79 different K types of *K. pneumoniae* [13] identified more than 1,500 different genes that were grouped into 361 homology groups. Despite the high number of different CPSs biosynthesised by *K. pneumoniae* strains, KPC-Kp isolates show limited CPS structural variation, because of their clonal origin. Genotyping of KPC-Kp isolates belonging to ST258 [14, 15] identified the presence of at least two lineages, indicated as clade I and clade II, which differ from each other primarily in the *cps* gene clusters, named *cps-1* and *cps-2*

(corresponding to *cps*$_{207\text{-}2}$ and *cps*$_{BO\text{-}4}$ in ref 14, respectively), and are novel with respect to those described by Pan et al [13]. The *cps-2* gene cluster is identical or very similar to those present in most of the CG258 *K. pneumoniae* strains from different countries, including strains of ST512, and also very similar to that described in a ST442 KPC-Kp strain (Kp13) that caused an outbreak in Brazil [16]. Both the K-antigen end-product of *cps-1* gene cluster together with the assignment of glycosyltransferases functions [17], and the CPS encoded by the *cps-2* gene cluster from two representative strains of ST258 have been previously described [18].

The present work reports the primary structure of the CPS produced by *K. pneumoniae* KPB-1 [19], a ST512 representative, determined by using GLC-MS of appropriate carbohydrate derivatives, and NMR spectroscopy of the native CPS. Moreover, the CPS of KPB-1 was compared with those produced by two other Italian strains of CG258. Taking advantage of the sequenced *cps*$_{KPB\text{-}1}$ gene cluster [19] and of the structural knowledge gained in the present investigation, each glycosyltransferase (GT) in the CPS gene cluster was assigned to the corresponding catalysed reaction.

2. Material and methods

2.1. Bacterial strain, biofilm production and exopolysaccharide purification.

The strain *Klebsiella pneumoniae* KPB-1 was isolated from pleural fluid of an inpatient at the St. Orsola-Malpighi University Hospital of Bologna, Italy and was shown to produce KPC-3 beta-lactamase and to belong to ST512 [19]. Bacterial cells were grown on 10 Worfel-Ferguson agar plates for 4 days at 30°C. The bacterial lawn was collected with 0.9 % NaCl (about 6 mL per dish), gently stirred at 10 °C for 2 h, centrifuged at 22,400 x *g* at 6 °C for 30 min to separate the cells from the supernatant which was precipitated with 4 vol of cold ethanol. The precipitated material, named **CPS KPB-1-P1**, was recovered by centrifugation at 1,900 x *g* at 4 °C for 30 min, dissolved in 60 mL of water, dialyzed first against 0.1 M NaCl and then water, taken to pH = 7.3, filtered (Millipore membranes 0.45 μm) and lyophilized. The bacterial cells recovered in the pellet of the first centrifugation were stirred in 30 mL of 0.9 % NaCl at 10 °C for 2 h, centrifuged at 22,400 x *g* at 4 °C for 30 min and the supernatant was treated as the one reported above and the resulting polysaccharide was named **CPS KPB-1-P2**.

2.2. General procedure

Analytical GLC was performed on a Perkin–Elmer Autosystem XL gas chromatograph equipped with a flame ionisation detector and using He as carrier gas. An SP2330 capillary column (Supelco, 30 m) was used to separate alditol acetates (temperature program: 200–245 °C at 4 °C/min). A HP-1 capillary column (Agilent Technologies, 30 m) was used to separate partially methylated alditol acetates (temperature program: 150–245 °C at 2 °C/min), trimethylsilylated methyl glycosides (temperature program: 150–280°C at 3 °C/min) and trimethylsilylated (+)-2-butyl glycosides, for the determination of the absolute configuration of the sugar residues [20] (temperature program: 135–240 °C at 1 °C/min). GLC–MS analyses were carried out on an Agilent Technologies 7890A gas chromatograph coupled to an Agilent Technologies 5975C VL MSD. Native and permethylated polysaccharides were hydrolysed with 2 M trifluoroacetic acid (TFA) for 1 h at 125 °C. Alditol acetates were prepared as already described [21]. Trimethylsilyl methyl glycosides were obtained by derivatization with the reagent Sylon™ HTP (Sigma) after methanolysis of the polysaccharide with 2 M HCl in methanol at 85 °C for 16 h [22]. Permethylation of the polysaccharide was achieved following the protocol by Harris [23], after conversion of the carboxylate groups to carboxyl functions by use of IR H^+ 120 cation exchange resin. Integration values of the areas of the partially methylated alditol acetates were corrected by the effective carbon response factors [24].

2.3. NMR experiments

The capsular polysaccharide produced by *K. pneumoniae* KPB-1 strain was dissolved in water (1 g/L) and sonicated using a Branson sonifier equipped with a microtip at 2.8 Å, in order to decrease its molecular mass. The sample was cooled in an ice bath and sonicated using 5 bursts of 1 min each, separated by 1 min intervals. Afterwards the polysaccharide was exchanged two times with 99.9% D_2O by lyophilisation and subsequently dissolved in 0.6 mL of 99.96% D_2O. Spectra were recorded on a 500 MHz VARIAN spectrometer operating at 50 °C. 2D experiments were performed using standard VARIAN pulse sequences and pulsed field gradients for coherence selection when appropriate. HSQC spectra were recorded using 140 Hz (for directly attached ^1H–^{13}C correlations). HMBC experiments were recorded using a coupling constant of 8 Hz (for long-range ^1H–^{13}C correlations) and relaxation time 1.2 s. TOCSY spectra were acquired using 100 ms spin-lock time and 1.0 s relaxation time. NOESY experiments were recorded with 200 ms mixing time and 1.0 s relaxation time. Chemical

shifts are expressed in ppm using acetone as internal reference (2.225 ppm for ^{1}H and 31.07 ppm for ^{13}C). NMR spectra were processed using MestreNova software.

2.4. Analysis of sequence data

Comparison of the nucleotide and protein sequences was performed using the web interface of BLAST [25], available at the NCBI website. An E value $<1^{-50}$ was chosen as cut-off to consider two proteins as belonging to the same Homology Group (HG) [13]. Multiple protein sequences were aligned with Multalin [26]. Structural comparison of KPB-1 and K34 *cps* gene clusters was performed with EasyFig [27].

3. Results and discussion

3.1. Purification and composition analysis of ***CPS KPB-1***

The strain *Klebsiella pneumoniae* KPB-1 was grown on carbohydrate rich Worfel-Ferguson agar medium for 4 days at 30°C. The bacterial lawn was collected with 0.9 % NaCl and subsequent ethanol precipitation of the supernatant gave 105 mg of **CPS KPB-1-P1**, while further dissolution of the remaining bacterial cells with 0.9 % NaCl followed by purification resulted in 19 mg of **CPS KPB-1-P2**. UV spectroscopy of a solution of both polysaccharides showed the absence of proteins and nucleic acids, and ^{1}H NMR spectroscopy revealed that the two samples were identical; thus, the polysaccharide was named **CPS KPB-1**.

Composition analysis as alditol acetate derivatives [21] revealed Rha and Gal in the molar ratios 2.5 : 1.0, while methanolysis followed by derivatization to TMS methyl glycosides [22] showed Rha : Gal: GalA in the molar ratios 4.0 : 1.4 : 1.0. The absolute configuration [20] was established to be D for Gal and GalA, and L for Rha. The position of the glycosidic linkages was determined by GLC and GLC-MS analysis on a HP-1 capillary column after derivatization of the sugar components to partially methylated alditol acetates: three peaks were detected and assigned to t-Rha*p*, 2-linked Rha*p*, and 3-linked Gal*p* in the molar ratios 0.6 : 2.4 : 1.0; since no branched hexose was found it was hypothesised that GalA is the site of branching.

3.2. NMR spectroscopy of ***CPS KPB-1***

The ^{1}H NMR spectrum (Figure 1) of a solution of **CPS KPB-1** showed five signals in the anomeric region at 5.40, 5.18, 5.14, 5.10, and 4.65 ppm; integration of the peak areas gave 1.0, 1.1, 1.2, 2.0 and 3.0, respectively. The resonance at 4.65 ppm is the sum of three overlapping peaks, of which only one belonged to an anomeric proton, as shown by the HSQC experiment (Figure 2). Therefore, the signals were named from **A** to **F** in order of decreasing chemical shift, with **D** and **E** overlapping at 5.10 ppm (**E** is visible as a shoulder, and from the COSY plot it was set at 5.09 ppm). Two resonances were detected in the high field region of the ^{1}H spectrum at 1.29 and 1.23 ppm, and their peak areas had integration values of 10.10 and 3.50, with respect to the anomeric signal at 5.40 ppm. They were attributed to the H-6 of the Rha residues and indicated the presence of three Rha residues with a similar electronic environment, and one with a slightly different surrounding. The anomeric configuration of the H-1 resonances were determined by measuring their $J_{C1,H1}$, after recording a coupled HSQC experiment. The values found (Table 1) showed that the residues **A** – **E** have the α anomeric configuration, while residue **F** has the β one [28]. Proton connectivities from H-1 to H-2 of spin systems **A** – **D** and from H-6 to H-5 of the deoxy sugars were determined from the COSY spectrum (Figure S1 in the Supplementary file). The TOCSY plot (Figure S2 in the Supplementary file) identified the remaining protons, starting both from the H-2's and from the methyl H-6's. The chemical shifts thus obtained assigned the spin systems **A** – **D** to Rha residues. Regarding residues **E** and **F**, proton connectivities from H-1 to H-4 were attributed in the COSY spectrum, while the TOCSY led to the determination of the remaining protons for these two spin systems. A HSQC experiment (Figure 2) gave most of the corresponding ^{13}C chemical shifts, except C-3 of **A**, and C-5 of **A**, **B**, **C** and **D** because of proton signals overlap in the 3.8 - 3.7 ppm region of the spectrum; they were obtained from a HMBC experiment (Figure 3). Full assignment of the ^{1}H and ^{13}C chemical shifts for each spin system is reported in Table 1. By comparison of the experimental chemical shifts for each spin system with literature values [29], the type of monosaccharide together with the position of glycosidic linkages were determined. Residue **A** was assigned to t-Rha, while residues **B, C** and **D** to 2-linked Rha. The spin systems **E** and **F** were attributed to 3,4-linked GalA and 3-linked Gal, respectively. The HMBC plot (Figure 3, Table 2) gave the carboxyl carbon chemical shift (data not shown) at 175.5 ppm, intra-residue as well as inter-residue resonances which determined the following sequences: H-1 α-Rha (**A**) to C-4 of α-GalA (**E**), H-1 α-Rha (**B**) to C-3 of β-Gal (**F**), and H-1 α-Rha (**C**) to C-2 of α-Rha (**B**). Since H-1 of α-Rha (**D**) and α-GalA (**E**) overlap, the crosspeaks to C-2 of α-Rha (**C**) and C-2 of α-Rha

(**D**) could not be unambiguously assigned to α-Rha (**D**) and α-GalA (**E**). Moreover, H-1 of α-Gal (**F**) showed a crosspeak at 80.4 ppm (data not shown) which can be both an intra-residue connectivity (C-3 of F) as well as an inter-residue one (C-3 of **E**). The NOESY plot (Figure 4, Table 2) showed intra-residues as well as the following inter-residues connectivities which unambiguously established the sequence of the monosaccharides in the repeating unit: H-1 α-Rha (**A**) to H-4 of α-GalA (**E**), H-1 α-Rha (**B**) to H-3 of β-Gal (**F**), H-1 α-Rha (**C**) to H-2 of α-Rha (**B**), H-1 α-Rha (**D**) to H-2 of α-Rha (**C**), H-1 α-GalA (**E**) to H-2 of α-Rha (**D**), and H-1 α-Gal (**F**) to H-3 of α-GalA (**E**).

In conclusion, all the experimental data collected demonstrated that the CPS produced by *Klebsiella pneumoniae* clinical strain KPB-1, belonging to the sequence type ST512, has a repeating unit composed of a pentasaccharide backbone and a rhamnosyl side chain, with the following structure:

```
       E                    D                  C                  B                 F
[3)-α-D-GalpA-(1→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→2)-α-L-Rhap-(1→3)-β-D-Galp-(1→]n
       4
       ↑
       1
  α-L-Rhap
       A
```

CPS KPB-1 is the first polysaccharide produced by a ST512 strain to be characterised; its structure is identical to that one of two representatives of an outbreak clone of ST258 KPC-Kp from USA [18] and also to the capsular polysaccharides of the epidemiologically unrelated CR-Kp of ST258 clade II strains KKBO-4 and KMn7 isolated from inpatients in Italian hospitals. Their CPSs were also purified and investigated as described for **CPS KPB-1** and as part of the same project. Within the same genus, the structure of **CPS KPB-1** is similar to *K. pneumoniae* K34 [30] and, outside the *Klebsiella* genus, to the O antigen produced by *Xanthomonas campestris* strain NCPPB 45 with which it differs only for a 3-Rha in place of a 2-Rha [31].

3.3. Prediction of the reactions catalyzed by the $cps_{BO\text{-}4}$ gene cluster glycosyltransferases

The *cps* gene cluster in *Klebsiella* spp. consists of some highly conserved genes together with a less conserved region [13]. Among the highly conserved there are six genes located at the 5' end, mainly involved in the export and polymerization of the repeating unit (*galF, cpsACP, wzi, wza, wzb, wzc)*, and two genes located downstream and encoding for glucose-6-phosphate dehydrogenase and UDP-glucose

dehydrogenase (*gnd* and *ugd*, respectively). In the cps_{KPB-1} cluster the less conserved region (Figure 5) contains: i) genes involved in the synthesis of activated α-L-Rha*p* (*rmlB-A-D-C*) and α-D-Gal*p*A (*uge-*1), located upstream and downstream of *ugd*, respectively; ii) a *wzx* gene coding for a flippase, located downstream *uge*-1 and a *wzy* gene, coding for the polymerase, located at the 3' end of the cluster; iii) five genes coding for glycosyltranferases (GT), four of which (*wbaP*, *orf8-10* which in the present study have been named *wcaA, wcuX* and *wclL*, respectively) are positioned in the central region between *wzc* and *gnd*, and one (*orf19*, now named *wcuY*) located between *wzx* and *wzy* [14]. *wbaP* is known to encode the initial GT which transfers the Gal-1-P from UDP-Gal to the lipid carrier undecaprenyl phosphate (Und-P), forming Gal-PP-Und [32].

As described above, the RU of the **CPS KPB-1** is very similar to that of K34 [30] (Figure 6), with which it shares three glycosidic bonds. However, comparison of the GTs coded in the two *cps* clusters retrieved homology only for two of them (see below) and it was not helpful for the prediction of the reactions catalysed. Therefore, the assignment of each GT catalysed reaction was carried out by comparison with other *Klebsiella* K antigens, taking advantage of the extensive *cps* gene cluster data published by Pan et al. [13], as hereafter described.

The first GT was identified by comparing **CPS KPB-1** with the K types K52 [33], K79 [34] and K81 [35], which share only the glycosidic bond Rha(α1-3)Galβ with **CPS KPB-1** (Figure 6). Alignment of the protein sequences of the GTs encoded by cps_{KPB-1} gene cluster with those expressed by the three K types showed good identity values (65%) between the *orf8* product and the rhamnosyl transferase WcaA. Notably, although the *wcaA* gene has been detected in many different *cps* clusters [13], the glycosidic bond Rha(α1-3)Galβ is present only in about half of the corresponding CPSs, while the remaining structures contain the disaccharide Rha(α1-3)Glcβ, suggesting the existence of different variants of this enzyme that, even if belonging to the same homology group (HG), underwent substantial modifications in their acceptor binding domain, corresponding to the C-terminal region [36]. This hypothesis is supported by comparison of the phylogenetic distance of the C-terminal region of WcaA in *K. pneumoniae* KPB-1 with that of homologous enzymes coded by other K types: the analysis detected a higher similarity among enzymes catalysing exactly the same reaction, with respect to variants catalysing the addition of Rha to a Glc acceptor, as indicated by a higher phylogenetic distance (Figure S3). A successive comparison of **CPS KPB-1** with K48 structure (Figure 6) [37] led to the identification of the GT coded by *orf10*. In fact, these CPSs share only one more glycosydic bond, and thus one GT, besides the already identified WcaA. More precisely, they have in common the GalA(α1-2)Rhaα structure and the GT coded by *orf10* showed homology with WclL coded in the K48

gene cluster (44% identities).

The identification of the GT coded by *orf9* turned out to be more complex. The glycosidic bond Rha(α1-2)Rhaα was detected in many different *Klebsiella* K types which, however, did not share any GT or expressed GTs with E values too high to be assigned to the same HG. The attention was focussed on the K types K40 [38] and K70 [39] which share only the Rha(α1-2)Rhaα glycosidic bond with **CPS KPB-1** (Figure 6). Comparison of their respective GTs retrieved low similarity values between the enzyme coded by *orf9* and both WcmY of K70 (31% identities) and WcpS of K40 (26% identities), suggesting that the rhamnosyl transferases catalysing the addition of an α-L-Rha*p* residue to C-2 of an α-L-Rha*p* acceptor might be very different from each other, showing at most some common motifs. Indeed, multiple alignments of their sequences retrieved some conserved residues and motifs (Figure S4) that were found also in two other rhamnosyl transferases catalysing the same reaction in the biosynthesis of the O-antigen of *E. coli* O13 and O35 [40]. Therefore, these data strongly support the hypothesis that the formation of the two Rha(α1-2)Rhaα glycosidic bonds in the RU of **CPS KPB-1** are catalysed by the GT coded by *orf9* which does not belong to any of the HGs described till now [13] and for which we propose the name WcuX.

By exclusion, it can be inferred that the GT coded by *orf19* catalyses the addition of the α-L-Rha*p* to C-4 of the α-D-Gal*p*A residue, forming the side chain of the RU. Searching for this protein in public databases showed that it is widely distributed among KPC-producing *Klebsiellae*, but since no homologous protein was detected among those of the *cps* clusters described by Pan et al. [13], it was assigned to a new HG and named WcuY. All the data collected, summarized in Table 3, led to the biochemical pathway for the synthesis of the **CPS KPB-1** repeating unit which is reported in Figure 7.

Finally, taking account the gained information, comparison between the biosynthetic pathway of **CPS KPB-1** and K34 RUs was reconsidered. The major difference between their CPSs, consisting in the first sugar residue linked to the lipid carrier, a Gal*p* in KPB-1 and a Glc*p* in K34, is clearly due to the different initial GT expressed by the two strains: WbaP in KPB-1 and WcaJ in K34 (Figure 5). The two strains express homologous WcaA enzymes that catalyse the addition of Rha*p* to C-3 of the first residue: comparison of their C terminal domains suggested that they are variants of the same enzyme which can recognize either Gal (KPB-1) or Glc (K34) as acceptor. The third reaction, consisting in the formation of the Rha(α1-2)Rhaα glycosidic bond, is identical in both strains but is catalysed by different enzymes. In fact, the WcuX GT of KPB-1 displays a low similarity (28% identities) with the unnamed protein coded by the *orf16* of the cps_{K34} cluster while both of them share some conserved residues with other rhamnosyl transferases reported in the literature to catalyse the same reaction.

However, the similarity between these enzymes is too low to include them into the same HG. The synthesis of the backbone continues with the addition of another Rha*p* residue, but the carbon atom involved is different in the two strains. In KPB-1 the glycosidic bond formed is identical to the previous one, so the reaction is clearly catalysed by the same GT, WcuX; in K34 the Rha*p* is linked to C-3 of Rha*p*, so the responsible GT cannot be surely predicted. It can be hypothesized that the addition of Rha*p* to C-3 and C-2 of another Rha*p* is catalysed by the same GT (the product of *orf16*), as an identical bifunctional activity has already been described for the rhamnosyl trasferase RfbF coded in the O-antigen gene cluster of two *Escherichia coli* strains [40, 41]. Another possibility is the existence of a different GT, coded by a gene located downstream and yet unidentified, as the sequence of the cps_{K34} cluster is truncated (as described below). The synthesis of the backbone ends with the addition of a GalA*p* residue, catalysed by the WclL GT that is expressed by both strains. Finally, as proposed above, the addition of Rha*p* to C-4 of GalA*p* is catalysed by the new GT named WcuY and coded by *orf19* which is located at the 3' end of the $cps_{KBP\text{-}1}$ cluster. Interestingly, although the K34 RU shows an identical side chain, no protein homologous to WcuY was detected among those coded in the cps_{K34} cluster. A careful analysis of the sequence of its *cps* cluster (Genbank Accession Number: AB924572) revealed a *ugd* truncated gene at the 3' end, suggesting that other, still unidentified, genes involved in the synthesis of K34 RU might be coded downstream. In fact, genes coding for a galacturonate epimerase and for a flippase located downstream *ugd* in the $cps_{KPB\text{-}1}$ cluster (*uge-1* and *wzx*, respectively), are lacking in the cps_{K34} cluster, even if these enzymes are essential for the synthesis of the CPS K34. In conclusion, although KPB-1 and K34 show a very similar CPS structure, at the genomic level their *cps* clusters look quite different in the less conserved region, consisting of genes with low similarity level (i.e. the operon for the synthesis of the Rha*p* residue) even when coding for common functions. One might speculate that they have a common ancestor which underwent important rearrangements after phage infections, since phage genes are present in both clusters, but in order to confirm this hypothesis more information about the 3' end sequence of the cps_{K34} cluster should be acquired.

Author Contributions:

Barbara Bellich: Investigation, Validation, Writing - Reviewing & Editing. **Cristina Lagatolla:** Conceptualization, Validation, Formal analysis, Visualization, Writing - Original Draft. **Roberto Rizzo:** Investigation, Validation, Writing - Reviewing & Editing. **Marco Maria D'Andrea:** Resources, Validation, Visualization, Writing - Original Draft. **Gian Maria Rossolini:** Resources,

Writing - Reviewing & Editing. **Paola Cescutti:** Conceptualization, Validation, Visualization, Supervision, Funding acquisition, Writing - Original Draft.

Competing interests

Authors have no competing interests to declare.

Acknowledgments

The authors thank the University of Trieste (FRA 2015) for financial support.

References

[1] M.K. Paczosa, J. Mecsas, *Klebsiella pneumoniae*: Going on the Offense with a Strong Defense, Microbiol. Mol. Biol. Rev. 80(3) (2016) 629–661. https://doi.org/10.1128/MMBR.00078-15.

[2] C-R. Lee, J.H. Lee, K.S. Park, Y.B. Kim, B.C. Jeong, S.H. Lee, Global dissemination of carbapenemase-producing *Klebsiella pneumoniae*: epidemiology, genetic context, treatment options, and detection methods, Front. Microbiol. 7 (2016) 895. https://doi.org/10.3389/fmicb.2016.00895.

[3] P. Nordmann, G. Cuzon, T. Naas, The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria, Lancet Infect. Dis. 9(4) (2009) 228-236. https://doi.org/10.1016/S1473-3099(09)70054-4.

[4] L.S. Munoz-Price, L. Poirel, R.A. Bonomo, M.J. Schwaber, G.L. Daikos, M. Cormican, G. Cornaglia, J. Garau, M. Gniadkowski, M.K. Hayden, K. Kumarasamy, D.M. Livermore, J.J. Maya, P. Nordmann, J.B. Patel, D.L. Paterson, J. Pitout, M.V. Villegas, H. Wang, N. Woodford, J.P. Quinn, Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases, Lancet Infect. Dis. 13(9) (2013) 785–796. https://doi.org/10.1016/S1473-3099(13)70190-7

[5] J.D. Pitout, P. Nordmann, L. Poirel, Carbapenemase-Producing *Klebsiella pneumoniae*, a Key Pathogen Set for Global Nosocomial Dominance, Antimicrob. Agents Chemother. 59(10) (2015) 5873-84. https://doi.org/10.1128/AAC.01019-15. Epub 2015 Jul 1

[6] B. Kitchel, J.K. Rasheed, J.B. Patel, A. Srinivasan, S. Navon-Venezia, Y. Carmeli, A. Brolund, C.G. Giske, Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* isolates in the United States: clonal expansion of multilocus sequence type 258, Antimicrob. Agents Chemother. 53(8) (2009) 3365-70. https://doi.org/doi: 10.1128/AAC.00126-09.

[7] A. Adler, O. Hussein, D. Ben-David, S. Masarwa, S. Navon-Venezia, M.J. Schwaber, Y. Carmeli, Post-Acute-Care Hospital Carbapenem-Resistant Enterobacteriaceae Working Group, Persistence of *Klebsiella pneumoniae* ST258 as the predominant clone of carbapenemase-producing Enterobacteriaceae in post-acute-care hospitals in Israel, 2008-13, J. Antimicrob. Chemother. 70(1) (2015) 89-92. https://doi.org/10.1093/jac/dku333

[8] S. David, S. Reuter, S. R. Harris, C. Glasner, T. Feltwell, S. Argimon, K. Abudahab, R. Goater, T. Giani, G. Errico, M. Aspbury, S. Sjunnebo, the EuSCAPE Working Group, the ESGEM Study Group, E. J. Feil, G. M. Rossolini, D. M. Aanensen, H. Grundmann, Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread, Nat. Microbiol. 4, (2019) 1919–1929. https://doi.org/10.1038/s41564-019-0492-8.

[9] V. Conte, M. Monaco, T. Giani, F. D'Ancona, M. L. Moro, F. Arena, M. M. D'Andrea, G. M. Rossolini, A. Pantosti, on behalf of the AR-ISS Study Group on Carbapenemase-Producing *K. pneumoniae*, Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* from invasive infections in Italy: increasing diversity with predominance of the ST512 clade II sublineage, J. Antimicrob. Chemother, 71(12) (2016) 3386–3391. https://doi.org/10.1093/jac/dkw337.

[10] S. Brisse, V. Passet, A.B. Haugaard, A. Babosan, N. Kassis-Chikhani, C. Struve, D. Decré, *wzi* gene sequencing, a rapid method for determination of capsular type for *Klebsiella* strains, J. Clin. Microbiol. 51 (2013) 4073–4078. https://doi.org/10.1128/JCM.01924-13.

[11] Y.J. Pan, T.L. Lin, Y.H. Chen, C.R. Hsu, P.F. Hsieh, M.C. Wu, J.T. Wang, 2013. Capsular types of *Klebsiella pneumoniae* revisited by *wzc* sequencing. PLoS ONE 8(12), e80670. https://doi.org/10.1371/journal.pone.0080670.

[12] K.L. Wyres, R.R. Wick, C. Gorrie, A. Jenney, R. Follador, N.R. Thomson, K.E. Holt, 2016. Identification of *Klebsiella* capsule synthesis loci from whole genome data. Microbial genomics 2(12), e000102. https://doi.org/10.1099/mgen.0.000102.

[13] Y.J. Pan, T.L. Lin, C.T. Chen, Y.Y. Chen, P.F. Hsieh, C.R. Hsu, M.C. Wu, J.T. Wang, 2015. Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of *Klebsiella* spp., Scientific Reports 5, 15573. https://doi.org/10.1038/srep15573.

[14] M.M. D'Andrea, F. Amisano, T. Giani, V. Conte, N. Ciacci, S. Ambretti, L. Santoriello, G.M. Rossolini, 2014. Diversity of capsular polysaccharide gene clusters in Kpc-producing *Klebsiella pneumoniae* clinical isolates of sequence type 258 involved in the italian epidemic, PLoS ONE 9(5), e96827. https://doi.org/10.1371/journal.pone.0096827.

[15] F.R. Deleo, L. Chen, S.F. Porcella, C.A. Martens, S.D. Kobayashi, A.R. Porter, K.D. Chavda, M.R. Jacobs, B. Mathema, R.J. Olsen, R.A. Bonomo, J.M. Musser, B.N. Kreiswirth, Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*, Proc. Natl. Acad. Sci. 111(13) (2014) 4988–4993. https://doi.org/10.1073/pnas.1321364111.

[16] P.I.P. Ramos, R.C. Picão, L.G.P. de Almeida, N.C.B Lima, R. Girardello, A.C.P. Vivan, D.E. Xavier, F.G Barcellos, M. Pelisson, E.C. Vespero, C. Médigue, A.T.R. de Vasconcelos, A.C. Gales, M.F. Nicolás (2014). Comparative analysis of the complete genome of KPC-2-producing *Klebsiella pneumoniae* Kp13 reveals remarkable genome plasticity and a wide repertoire of virulence and resistance mechanisms, BMC Genomics 15: 54. https://doi.org/10.1186/1471-2164-15-54.

[17] B. Bellich, N. Ravenscroft, R. Rizzo, C. Lagatolla, M.M. D'Andrea, G.M. Rossolini, P. Cescutti, Structure of the capsular polysaccharide of the KPC-2-producing *Klebsiella pneumoniae* strain KK207-2 and assignment of the glycosyltransferases functions, Int. J. Biol. Macromol. 130 (2019) 536–544. https://doi.org/10.1016/j.ijbiomac.2019.02.128.

[18] J. Kubler-Kielb, E. Vinogradov, W.I. Ng, B. Maczynska, A. Junka, M. Bartoszewicz, A. Zelazny, J. Bennett, R. Schneerson, The capsular polysaccharide and lipopolysaccharide structures of two carbapenem resistant *Klebsiella pneumoniae* outbreak isolates, Carbohydr. Res. 369 (2013) 6–9. https://doi.org/10.1016/j.carres.2012.12.018.

[19] A. Cannatelli, V. Di Pilato, T. Giani, F. Arena, S. Ambretti, P. Gaibani, M.M. D'Andrea, G.. M. Rossolini, In vivo evolution to colistin resistance by PmrB sensor kinase mutation in KPC-producing *Klebsiella pneumoniae* is associated with low-dosage colistin treatment, Antimicrob. Agents Chemother. 58(8) (2014) 4399-4403. https://doi.org/10.1128/AAC.02555-14.

[20] G.J. Gerwig, J.P. Kamerling, J.F.G. Vliegenthart, Determination of the d and l configuration of neutral monosaccharides by high-resolution capillary g.l.c., Carbohydr. Res. 62 (1978) 349–357. https://doi.org/10.1016/S0008-6215(00)80881-2.

[21] P. Albersheim, D.J. Nevins, P.D. English, A. Karr, A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography, Carbohydr. Res. 5 (1967) 340–345. https://doi.org/10.1016/S0008-6215(00)80510-8.

[22] K. Kakehi, S. Honda, Silyl ethers of carbohydrates, in: C.J. Biermann, G.D. McGinnis (Eds.), Analysis of carbohydrates by GLC and MS, CRC Press, Boca Raton (FL), 1989, pp. 43–85.

[23] P.J. Harris, R.J. Henry, A.B. Blakeney, B.A. Stone, An improved procedure for the methylation analysis of oligosaccharides and polysaccharides, Carbohydr. Res. 127 (1984) 59–73. https://doi.org/10.1016/0008-6215(84)85106-X.

[24] D.P. Sweet, R.H. Shapiro, P. Albersheim, Quantitative analysis by various g.l.c. response-factor theories for partially methylated and partially ethylated alditol acetates, Carbohydr. Res. 40 (1975) 217–225. https://doi.org/10.1016/S0008-6215(00)82604-X.

[25] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T. L. Madden, NCBI BLAST: a better web interface, Nucleic Acids Res. 36 (2008) W5–W9. https://doi.org/10.1093/nar/gkn201.

[26] F. Corpet, Multiple sequence alignment with hierarchical clustering, Nucl. Acids Res. 16 (1988) 10881–10890. https://doi.org/10.1093/nar/16.22.10881.

[27] M.J. Sullivan, N.K. Petty, S.A. Beatson, Easyfig: a genome comparison visualizer, Bioinformatics 27(7) (2011) 1009-10. https://doi.org/10.1093/bioinformatics/btr039.

[28] K. Bock, I. Lundt, C. Pedersen, Assignment of anomeric structure to carbohydrates through germinal ^{13}C-H coupling constants, Tetrahedron Lett. 14 (1973) 1037–1040. https://doi.org/10.1016/S0040-4039(01)95898-8.

[29] P.E. Jansson, L. Kenne, G. Widmalm, Computer-assisted structural analysis of polysaccharides with an extended version of CASPER using ^{1}H-and ^{13}C-NMR data, Carbohydr. Res. 188 (1989) 169–191. https://doi.org/10.1016/0008-6215(89)84069-8.

[30] J.P. Joseleau, F. Michon, M. Vignon, Structural investigation of the capsular polysaccharide from *Klebsiella* serotype K-34 and its characterization by N.M.R. spectroscopy, Carbohydr. Res. 101 (1982) 175–185. https://doi.org/10.1016/S0008-6215(00)80998-2.

[31] A.V. Bukharov, I.M. Skvortsov, V.V. Ignatov, A.S. Shashkov, Y.A. Knirel, J. Dabrowski, Structure of the O-specific polysaccharide of *Xanthomonas campestris* NCPPB 45 lipopolysaccharide, Carbohydr. Res. 241 (1993) 309–316. https://doi.org/10.1016/0008-6215(93)80121-T.

[32] H.Y. Shu, C.P. Fung, Y.M. Liu, K.M. Wu, Y.T. Chen, L.H. Li, T.T. Liu, R. Kirby, S.F. Tsai, Genetic diversity of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* clinical isolates, Microbiology 155 (2009) 4170–4183. https://doi.org/10.1099/mic.0.029017-0.

[33] H. Bjorndal, B. Lindberg, J. Lonngren, M. Meszaros, J.L. Thompson, Structural studies of the capsular polysaccharide of *Klebsiella* type 52, Carbohydr. Res. 31 (1973) 93-100. https://doi.org/10.1016/S0008-6215(00)82321-6.

[34] G.G. Dutton, A.V. Lim, Structure of the capsular polysaccharide of *Klebsiella* serotype K79, Carbohydr. Res. 144 (1985) 263-276. https://doi.org/10.1016/S0008-6215(00)90674-8.

[35] M. Curvall, B. Lindberg, J. Lonngren, Structural studies of the capsular polysaccharide of *Klebsiella* type 81, Carbohydr. Res. 42 (1975) 73-82. https://doi.org/10.1016/S0008-6215(00)84101-4.

[36] B. Henrissat, G. Sulzenbacher, Y. Bourne, Glycosyltransferases, glycoside hydrolases: surprise, surprise!, Curr. Opin. Struct. Biol. 18 (2008) 527–533. https://doi.org/10.1016/j.sbi.2008.09.003.

[37] J.P. Joseleau, M.-F. Marais, The structural repeating-unit of the capsular polysaccharide from *Klebsiella* serotype K48, Carbohydr. Res. 179 (1988) 321–326. https://doi.org/10.1016/0008-6215(88)84128-4.

[38] P. Cescutti, R. Toffanin, B.J. Kvam, S. Paoletti, G.G.S. Dutton, Structural determination of the capsular polysaccharide produced by *Klebsiella pneumoniae* serotype K40. NMR studies of the oligosaccharide obtained upon depolymerisation of the polysaccharide with a bacteriophage-associated endoglycanase, Eur. J. Biochem. 213(1) (1993) 445–453. https://doi.org/10.1111/j.1432-1033.1993.tb17780.x.

[39] G. G.S. Dutton, K. L. Mackie, Structural investigation of *Klebsiella* serotype K70 polysaccharide. Carbohydr. Res. 62 (1978) 321–335. https://doi.org/10.1016/S0008-6215(00)80879-4.

[40] A.S. Shashkov, Y. Zhang, Q. Sun, X. Guo, S.N. Senchenkova, A.V. Perepelov, Y.A. Knirel, Structure and gene cluster of the O-antigen of *Escherichia coli* O133, Carbohydr. Res. 430 (2016) 82-84. https://doi.org/10.1016/j.carres.2016.04.028.

[41] B. Liu, F. Wu, D. Li, L. Beutin, M. Chen, B. Cao, L. Wang, Development of a serogroup-specific DNA microarray for identification of *Escherichia coli* strains associated with bovine septicemia and diarrhea, Vet. Microbiol. 142 (2010) 373-378. https://doi.org/10.1016/j.vetmic.2009.10.019.

[42] T. Rundlöf, A. Weintraub, G. Widmalm, Structural determination of the O-antigenic polysaccharide from *Escherichia coli* O35 and cross-reactivity to *Salmonella arizonae* O62, Eur. J. Biochem. 258 (1998) 139-143. https://doi.org/10.1046/j.1432-1327.1998.2580139.x.

[43] A.V. Perepelov, S.D. Shevelev, B. Liu, S.N. Senchenkova, A.S. Shashkov, L. Feng, Y.A. Knirel, L. Wang, Structures of the O-antigens of *Escherichia coli* O13, O129, and O135 related to the O-antigens of *Shigella flexneri*, Carbohydr. Res. 345 (2010) 1594-1599. https://doi.org/10.1016/j.carres.2010.04.023.

[44] B. Liu, Y.A. Knirel, L. Feng, A.V. Perepelov, S.N. Senchenkova, Q. Wang, P.R. Reeves, L. Wang, Structure and genetics of *Shigella* O antigens, FEMS Microbiol Rev. 32 (2008) 627-653. https://doi.org/10.1111/j.1574-6976.2008.00114.x.

Table 1

^{1}H and ^{13}C chemical shift assignments of the **CPS KPB-1**

Residues[a]	Nucleus	Chemical shifts (ppm)[b]					
		1	**2**	**3**	**4**	**5**	**6**
A [176.45]	^{1}**H**	5.40	4.06	3.81	3.37	3.79	1.23
α-L-Rha*p*-(1→	^{13}**C**	100.7	71.3	70.8	73.2	69.9	17.5
B [174.93]	^{1}**H**	5.18	4.08	3.95	3.50	3.81	1.30
→2)-α-L-Rha*p*-(1→	^{13}**C**	101.6	78.6	70.9	73.0	70.7	17.5
C [174.42]	^{1}**H**	5.14	4.11	3.88	3.48	3.72	1.28
→2)-α-L-Rha*p*-(1→	^{13}**C**	101.6	79.2	70.6	73.0	70.1	17.5
D [174.93]	^{1}**H**	5.10	4.14	3.88	3.51	3.75	1.29
→2)-α-L-Rha*p*-(1→	^{13}**C**	100.5	76.5	70.6	73.0	70.1	17.5
E [174.93]	^{1}**H**	5.09	4.08	4.21	4.67	4.65	
→3,4)-α-D-Gal*p*A-(1→	^{13}**C**	98.2	68.2	80.4	77.4	72.6	175.5
F [162.76]	^{1}**H**	4.65	3.72	3.72	4.00	3.71	3.75-3.87
→3)-β-D-Gal*p*-(1→	^{13}**C**	105.2	72.0	80.4	69.5	75.9	61.9

[a] $^{1}J_{C1,H1}$ in square brackets.

[b] Chemical shifts are given relative to internal acetone (2.225 ppm for ^{1}H and 31.07 ppm for ^{13}C).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 2: Inter- and intra-residues NOE contacts of the anomeric protons of **CPS KPB-1**

Residue	NOE contact (ppm)	assignment
A1	4.66	**E4**
	4.06	**A2**
B1	4.08	**B2**
	3.72	**F3**
C1	4.11	**C2**
	4.08	**B2**
D1	4.14	**D2**
	4.11	**C2**
E1	4.14	**D2**
	4.08	**E2**
F1	5.40	**A1**
	4.21	**E3**
	3.72	**F2**

Table 3.

Assignments of glycosyltransferases catalysed reactions on the basis of CPS structural homology and protein sequence homology.

GT	Glycosidic bond	Accession no.	K type and O-antigen	References	
				Structure	*cps* cluster
WcaA	Rha(α1-3)Galβ	CCI88020.1	KPB-1	This work	[14]
		BAT23633.1	K34	[30]	[13]
		BAT23965.1	K52	[33]	[13]
		BAT24385.1	K79	[34]	[13]
		BAT24429.1	K81	[35]	[13]
WcuX	Rha(α1-2)Rhaα	CCI88021.1	KPB-1	This work	[14]
WcpS		BAT23740.1	K40	[38]	[13]
WcmY		BAT24307.1	K70	[39]	[13]
Orf 16		BAT23642.1	K34	[30]	[13]
WcnY		ACV67288.1	*E. coli* O35	[42]	[41]
RfbF		ACD37163.1	*E. coli* O13	[43]	[44]
WclL	GalA(α1-2)Rhaα	CCI88022.1	KPB-1	This work	[14]
		BAT23641.1	K34	[30]	[13]
		BAT23896.1	K48	[37]	[13]
WcuY	Rha(α1-4)GalAα	CCI88031.1	KPB-1	This work	[14]

1
2
3
4 **FIGURE LEGENDS**
5
6
7 Figure 1: Expansion of the ^{1}H-NMR spectra recorded in D_2O at 50°C of **CPS KPB-1**. Anomeric
8
9 signals are named **A** to **F**, as reported in Table 1.
10
11 Figure 2: Expansion of the HSQC plot recorded in D_2O at 50 °C of **CPS KPB-1**. C-H cross peaks
12
13 assignments are shown (residues nomenclature as reported in Table 1).
14
15
16 Figure 3: Expansion of the HMBC plot recorded in D_2O at 50 °C of **CPS KPB-1**. The corresponding
17
18 part of the ^{1}H C NMR spectrum is shown along the horizontal axis. Proton/carbon crosspeaks have
19
20 been labelled according to the corresponding residue (**A** to **F**).
21
22 Figure 4: Expansion of the NOESY plot recorded in D_2O at 50 °C of **CPS KPB-1** showing intra- and
23
24 inter-residue connectivities. Cross-peaks are labeled as reported in Tables 1 and 2.
25
26
27 Figure 5. Comparison of the *cps* gene clusters of *K. pneumoniae* KPB-1 (cps_{KPB-1}) and the K34
28
29 serotype reference strain (Accession number AB924572). Genes are connected by areas of different
30
31 colors, reflecting the degree of protein identity. Open reading frames encoding glycosyltransferases are
32
33 in yellow while those encoding phage genes are in red. The sequence of cps_{KPB-1} differs to that of
34
35 cps_{KKBO-4} (Accession number HE866751.1) by two non-synonymous single-nucleotide polymorphisms
36 occurring in the genes *wzy* (T794C; L265S) and Mu-like (T1337G; G446V). Sequence of the *ugd* gene
37
38 in cps_{K34} appears to be truncated and therefore is indicated as *Δugd*.
39
40
41 Figure 6: Structure of the repeating units of different CPS: **KPB-1** (this article), **K52** [28], **K79** [29],
42
43 **K81** [30], **K48** [31], **K40** [32], **K70** [33] and **K34** [26] *Klebsiella pneumoniae* capsular
44
45 polysaccharides. Circles of different colors highlight common glycosidic linkages. For **K48** the red
46
47 circle has a dashed line to highlight that the GT WcaA catalyses the addition of a Rha residue to a Glc
48
49 instead of a Gal. These structural similarities were used to assign the corresponding GTs genes.
50
51 Figure 7: Proposed glycosyltransferase and polymerase activity of the *Klebsiella pneumoniae* KPB-1
52
53 *cps* gene cluster. Glycosyltransferases responsible for each elongation step are listed above the
54
55 respective glycosidic linkage. The polymerization site is marked by an arrow.
56
57
58
59
60
61
62

64
65

Figure 1 Click here to access/download;Figure;Fig 1 1HNMR.tif

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

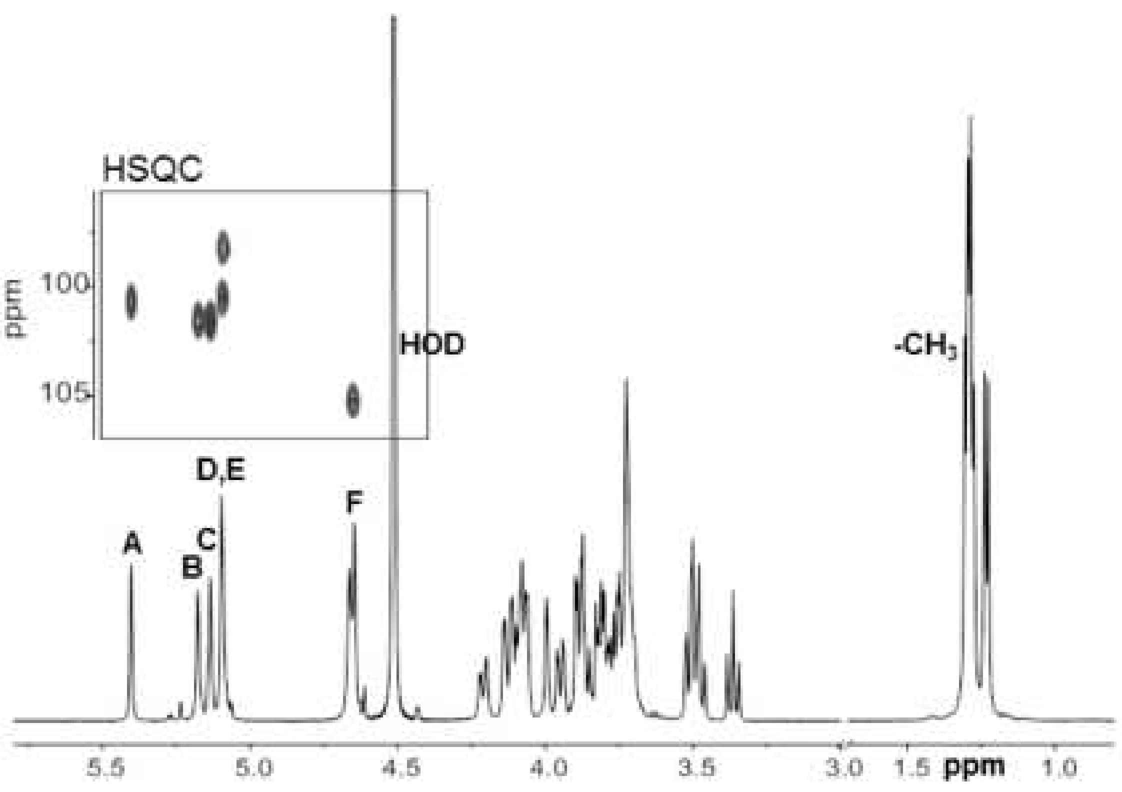

Figure 1

Click here to access/download;Figure;Fig 2 HSQC.tif

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

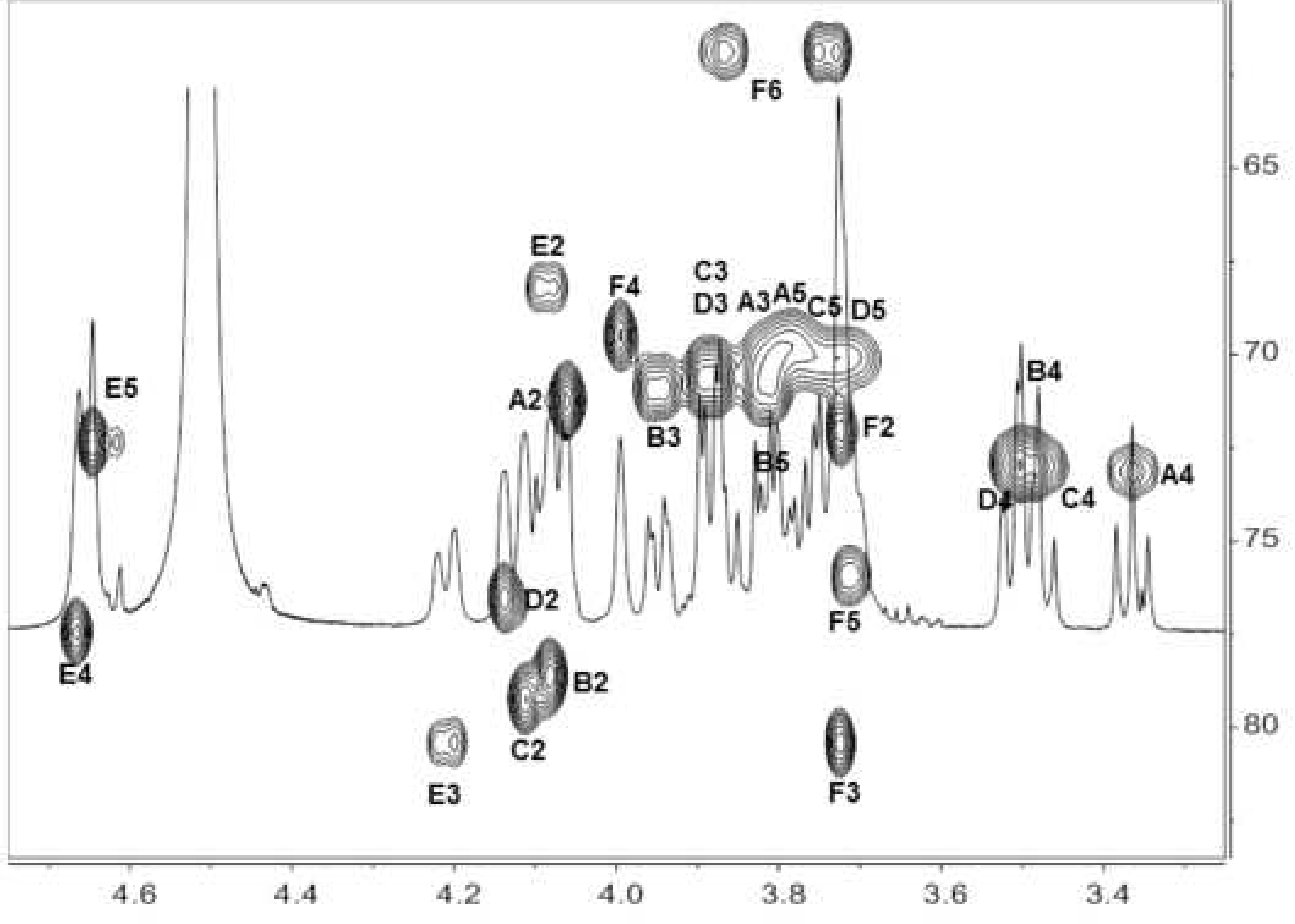

Figure 2

 Click here to access/download;Figure;Fig 3 HMBC.tif

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

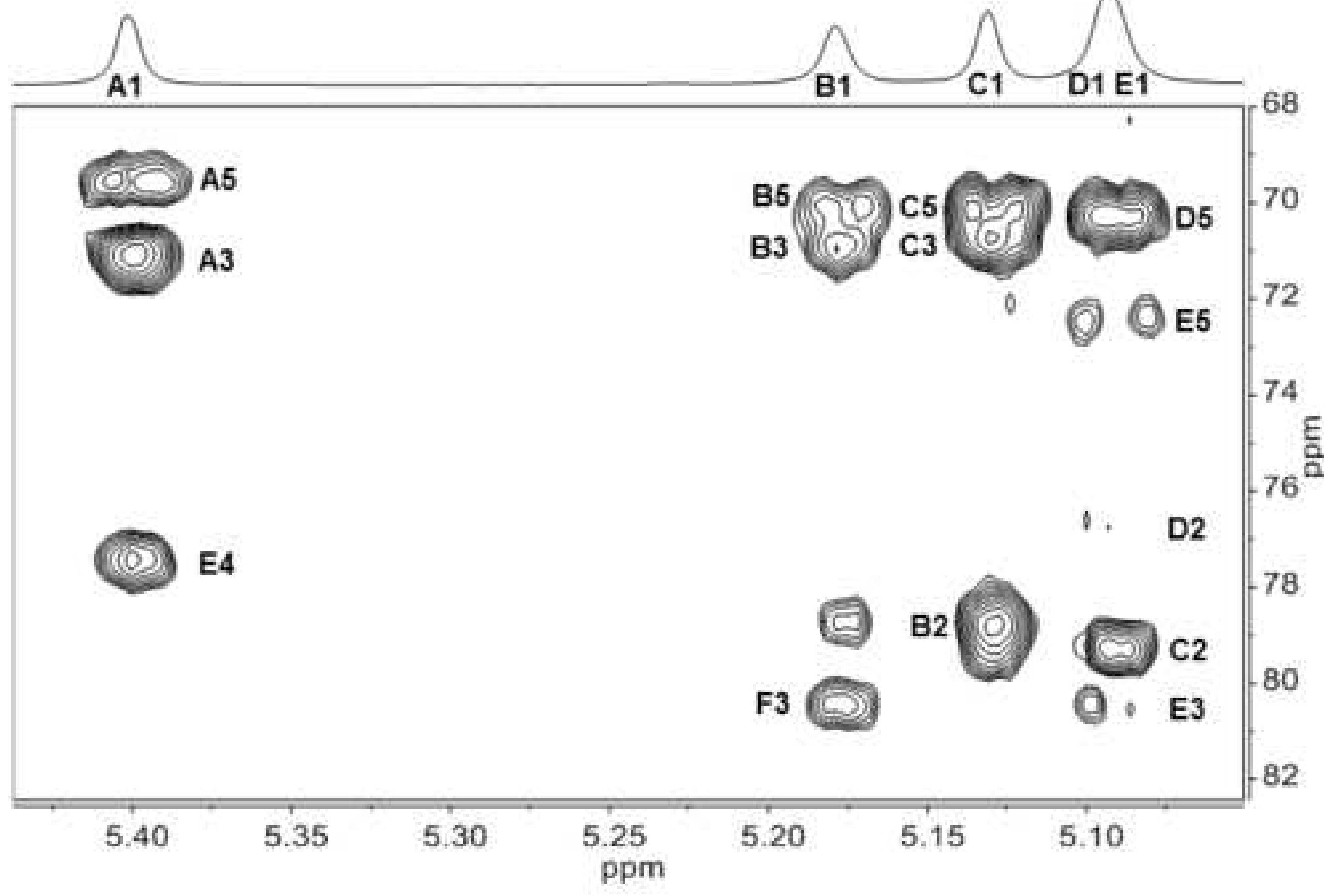

Figure 3

Click here to access/download;Figure;Fig 4 NOESY.tif

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

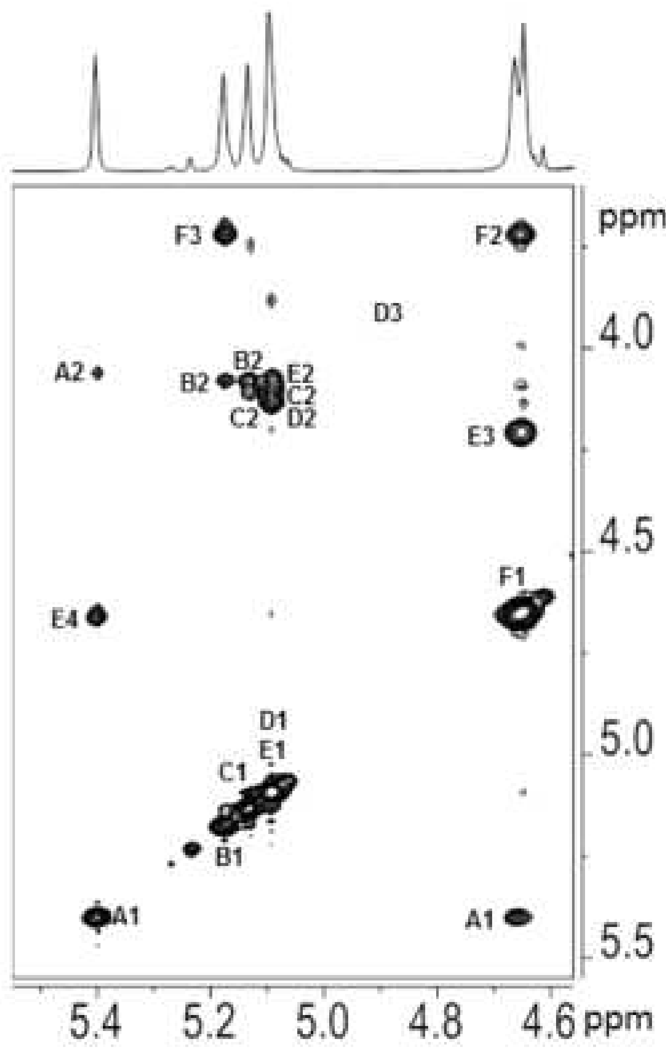

Figure 4

Click here to access/download;Figure;Fig 5 CLUSTERS.tif

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

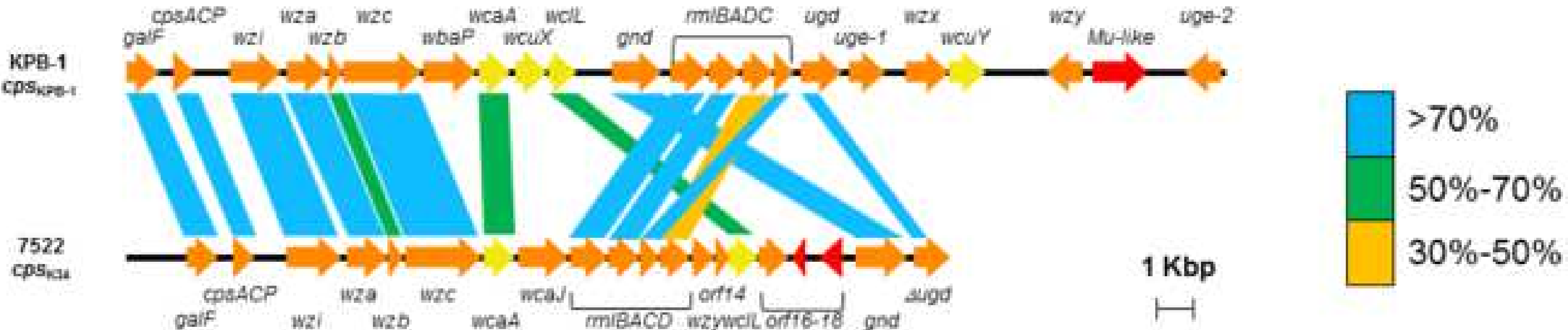

Figure 5

Figure 6

KKBO-4

[3)-α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→]$_n$
4
↑
1
α-L-Rha*p*

K52

[2)-α-L-Rha*p*-(1→4)-β-D-Glc*p*A-(1→3)-α-D-Gal*p*-(1→4)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→]$_n$
2
↑
1
α-D-Gal*p*

K79

[3)-β-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→]$_n$
4
↑
1
α-D-Glc*p*
6
↑
1
α-D-Glc*p*

K48

[3)-α-L-Rha*p*-(1→4)-α-D-Glc*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Glc*p*-(1→]$_n$
2
↑
1
α-D-Gal*p*A

K81

[2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→4)-β-D-Glc*p*A-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-β-D-Gal*p*-(1→]$_n$

K40

[3)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→4)-α-D-Glc*p*A-(1→2)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→3)-α-D-Gal*p*-(1→]

K70

pyr
|
[2)-α-L-Rha*p*-(1→4)-β-D-Glc*p*A-(1→4)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→2)-α-D-Glc*p*-(1→2)-β-D-Gal*p*-(1→]

K34

[3)-α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Glc*p*-(1→]$_n$
4
↑
1
α-L-Rha*p*

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

Figure 7
Click here to access/download;Figure;Fig 7 GT.tif

Determination of the capsular polysaccharide structure of the *Klebsiella pneumoniae* ST512 representative strain KPB-1 and assignments of the glycosyltransferases functions

Barbara Bellich, Cristina Lagatolla, Roberto Rizzo, Marco Maria D'Andrea, Gian Maria Rossolini and Paola Cescutti

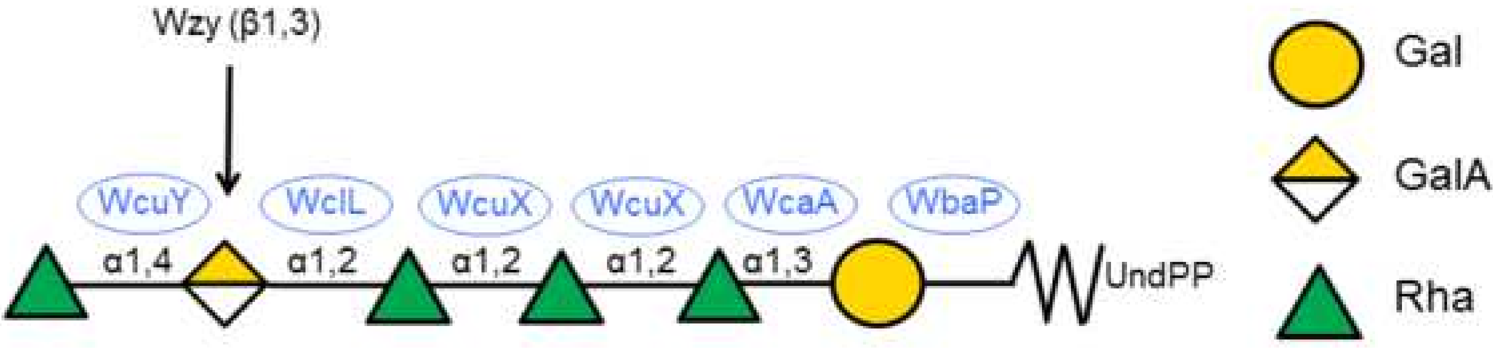

Figure 7

Author Contributions

Barbara Bellich: Investigation, Validation, Writing - Reviewing & Editing. **Cristina Lagatolla:** Conceptualization, Validation, Formal analysis, Visualization, Writing - Original Draft. **Roberto Rizzo:** Investigation, Validation, Writing - Reviewing & Editing. **Marco Maria D'Andrea:** Resources, Validation, Visualization, Writing - Original Draft. **Gian Maria Rossolini:** Resources, Writing - Reviewing & Editing. **Paola Cescutti:** Conceptualization, Validation, Visualization, Supervision, Funding acquisition, Writing - Original Draft.

Click here to access/download
Supplementary Material
200109_Supplementary material.pdf

graphical abstract

Click here to access/download
e-Component
Graph Abs.tif