

# SCIENTIFIC REPORTS



OPEN

## Structural basis for fragmenting the exopolysaccharide of *Acinetobacter baumannii* by bacteriophage $\Phi$ AB6 tailspike protein

Received: 26 September 2016

Accepted: 12 January 2017

Published: 17 February 2017

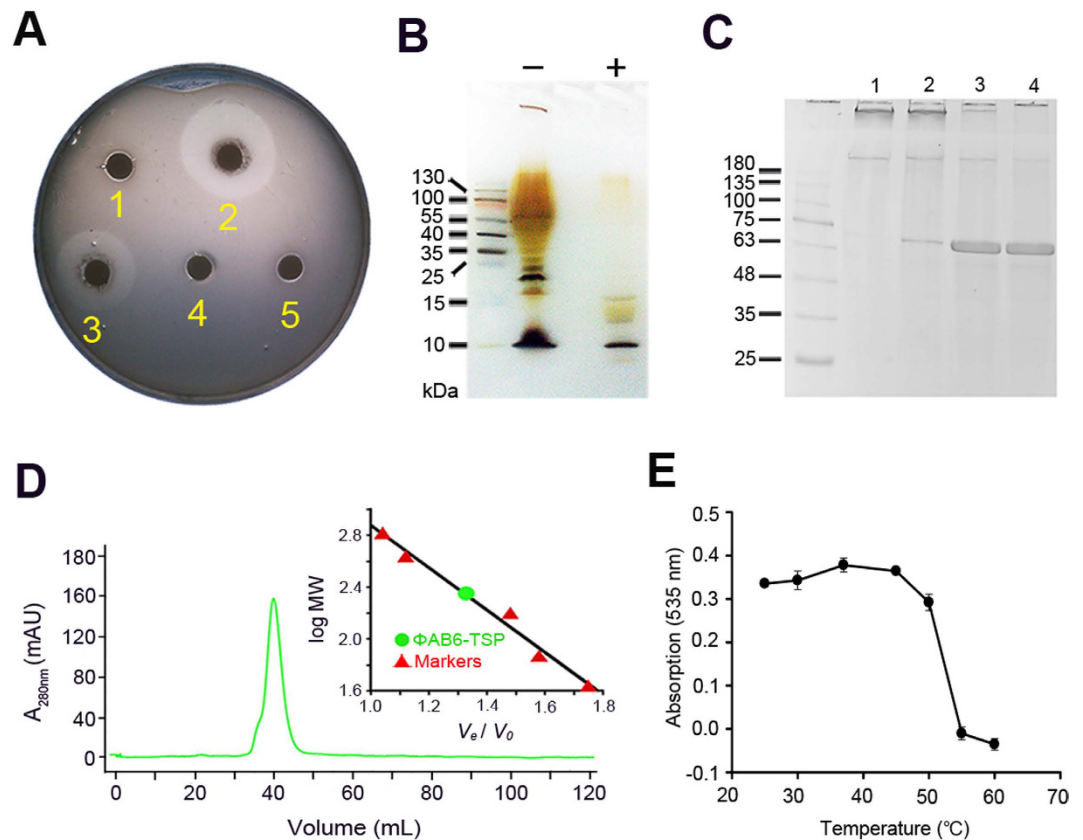
I-Ming Lee<sup>1,2</sup>, I-Fan Tu<sup>2</sup>, Feng-Ling Yang<sup>2</sup>, Tzu-Ping Ko<sup>2,3</sup>, Jiahn-Haur Liao<sup>2</sup>, Nien-Tsung Lin<sup>4</sup>, Chung-Yi Wu<sup>5</sup>, Chien-Tai Ren<sup>5</sup>, Andrew H.-J. Wang<sup>2,3</sup>, Ching-Ming Chang<sup>2</sup>, Kai-Fa Huang<sup>2,3</sup> & Shih-Hsiung Wu<sup>1,2,6</sup>

With an increase in antibiotic-resistant strains, the nosocomial pathogen *Acinetobacter baumannii* has become a serious threat to global health. Glycoconjugate vaccines containing fragments of bacterial exopolysaccharide (EPS) are an emerging therapeutic to combat bacterial infection. Herein, we characterize the bacteriophage  $\Phi$ AB6 tailspike protein (TSP), which specifically hydrolyzed the EPS of *A. baumannii* strain 54149 (*Ab-54149*). *Ab-54149* EPS exhibited the same chemical structure as two antibiotic-resistant *A. baumannii* strains. The  $\Phi$ AB6 TSP-digested products comprised oligosaccharides of two repeat units, typically with stoichiometric pseudaminic acid (Pse). The 1.48–1.89-Å resolution crystal structures of an N-terminally-truncated  $\Phi$ AB6 TSP and its complexes with the semi-hydrolyzed products revealed a trimeric  $\beta$ -helix architecture that bears intersubunit carbohydrate-binding grooves, with some features unusual to the TSP family. The structures suggest that Pse in the substrate is an important recognition site for  $\Phi$ AB6 TSP. A region in the carbohydrate-binding groove is identified as the determinant of product specificity. The structures also elucidated a retaining mechanism, for which the catalytic residues were verified by site-directed mutagenesis. Our findings provide a structural basis for engineering the enzyme to produce desired oligosaccharides, which is useful for the development of glycoconjugate vaccines against *A. baumannii* infection.

The gram-negative pathogen *Acinetobacter baumannii* (*Ab*) causes nosocomial urinary tract and bloodstream infections, ventilator-associated pneumonia, and meningitis<sup>1</sup>. In the past decade, estimated mortality of hospital outbreaks and community-acquired *Ab* infections was 26% to 68%<sup>2</sup>. While carbapenem and colistin are used in first-line treatments<sup>3</sup>, antibiotic-resistant strains are emerging<sup>2</sup>, likely due to this pathogen acquiring antibiotic-resistant genes from other *Acinetobacter* strains or from bacteria colonizing the same environment through horizontal gene transfer<sup>4</sup>. This represents a rigorous challenge to the treatment of *Ab* infections in the future, so there is a pressing need for the development of new therapies.

In addition to antibiotic treatments, another choice for combating bacterial infections is vaccination, particularly utilizing the bacterial exopolysaccharide (EPS) to generate glycovaccines<sup>5,6</sup>. Glycoconjugate vaccines (GCV), which are generated by conjugating the bacterial EPS to a carrier protein, can elicit a strong and long-lasting immune response compared to vaccines using whole polysaccharide<sup>7,8</sup>. This is likely due to GCVs preferentially stimulating the adaptive immune responses<sup>9–11</sup>. However, the heterogeneity of bacterial EPS extracts hampered the development of GCVs. The heterogeneous mixture of polymerization forms of EPS usually resulted in multi-point attachment of protein on the polysaccharide when conjugating to carrier proteins, leading to a considerable challenge to the reproducibility of glycoconjugates. It has been well known that many bacteriophages are capable of depolymerizing the EPS of their host bacteria through the tailspike proteins (TSPs)<sup>12</sup>. The digested products

<sup>1</sup>Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan. <sup>2</sup>Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan. <sup>3</sup>Core Facilities for Protein Structural Analysis (CFPSA), Academia Sinica, Taipei 115, Taiwan. <sup>4</sup>Master program in Microbiology and Immunology, Tzu Chi University, Hualien 970, Taiwan. <sup>5</sup>Genomics Research Center, Academia Sinica, Taipei 115, Taiwan. <sup>6</sup>Department of Chemistry, National Taiwan University, Taipei 106, Taiwan. Correspondence and requests for materials should be addressed to K.F.H. (email: huangkf@gate.sinica.edu.tw) or S.H.W. (email: shwu@gate.sinica.edu.tw)



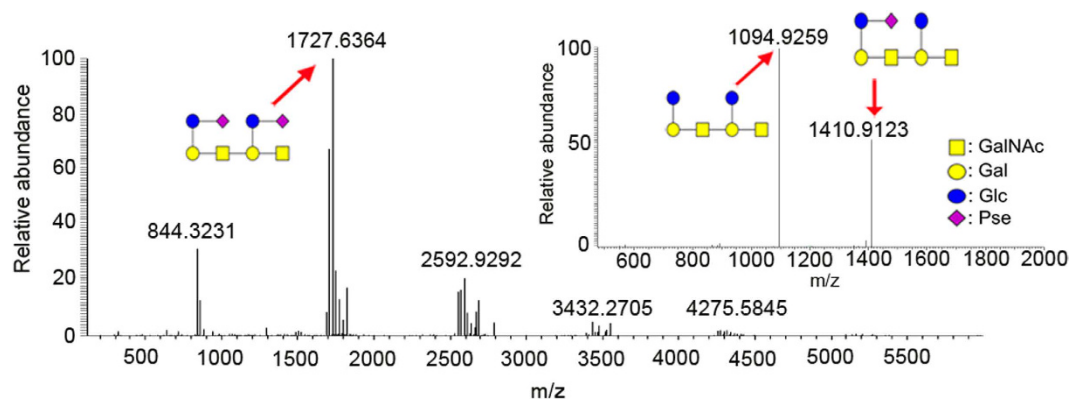
**Figure 1. Characterization of the recombinant ΦAB6 TSP.** (A) Evaluation of enzyme activity of wild-type and mutant ΦAB6 TSPs using top agar assay. The top agar was pre-inoculated with *Ab-54149* and then spotted with (1) buffer (25 mM Tris-HCl, 0.1 M NaCl, pH 7.5), (2) wild-type ΦAB6 TSP, (3) ΦAB6 TSPΔN, (4) mutant E425Q, and (5) mutant E447Q. (B) Analysis of the extract of *Ab-54149* exopolysaccharide by silver-stained SDS-PAGE. About 150 μg exopolysaccharide extract with (+) or without (–) the treatment of ΦAB6 TSP was loaded into the gel. The first lane was loaded with molecular weight markers. (C) Ability of resistance to the SDS denaturation. The enzyme was pre-heated at 80 °C for 0 (lane 1), 1 (lane 2), 5 (lane 3), and 10 (lane 4) min in the presence of 2% SDS. (D) Oligomerization status in solution. ΦAB6 TSP was eluted as a single peak from an analytical gel-filtration column. Based on the calibration curve of log MW versus  $V_e/V_0$  ( $V_e$ , elution volume;  $V_0$ , void volume) generated with five protein markers (the inset), the elution volume of ΦAB6 TSP corresponds to a molecular weight of 217.9 kDa, very close to the theoretic value (228.5 kDa) of the trimer calculated from the amino acid sequence. (E) Evaluation of thermal stability. The enzyme was heated at varied temperatures for 5 min and then the enzyme activity was evaluated immediately as described in the Methods. The experiment was performed in duplicates.

comprise up to two oligosaccharides with some repeat units in general<sup>13</sup>. In this way, TSPs can be employed to make EPS derivatives with preferred size and homogeneity (Supplementary Fig. 1).

Bacteriophage TSP is found at the tail of the viral particle<sup>12</sup>. They comprise elongated monomers with β-helical topology, which associate into trimers. These proteins bind to the EPS of specific bacteria and digest the EPS to facilitate the attachment of virions to the host cells. Previous studies of TSP structures, such as TSPs from bacteriophages P22, Sf6 and HK620, including oligosaccharide complexes have shed light on the substrate binding modes and catalytic mechanisms<sup>14–16</sup>. Here we present the crystal structures of the TSP from bacteriophage ΦAB6 and its complex with cleaved EPS fragments of *Ab-54149*, a clinical *A. baumannii* strain in Taiwan. ΦAB6 was reported to belong to the *Autographivirinae* subfamily and to specifically recognize *Ab-54149*<sup>17</sup>. In the present study, the substrate binding mode and catalytic mechanism of ΦAB6 TSP were subsequently elucidated by structural analysis and site-specific mutagenesis. These results can serve as a basis for producing better EPS fragments as the starting materials of GCV.

## Results

**Characterization of ΦAB6 TSP.** The open reading frame encoding the tailspike protein of ΦAB6 (ΦAB6 TSP) has recently been identified<sup>17</sup>. Here, a recombinant ΦAB6 TSP was overexpressed in *E. coli* cells and purified to near homogeneity. The purified ΦAB6 TSP was shown to generate a translucent halo when spotted on a top agar inoculated with *A. baumannii* strain 54149 (*Ab-54149*). The enzyme also degraded the extracts of *Ab-54149* exopolysaccharide (EPS) in a time-dependent manner (Fig. 1A and Supplementary Fig. 2A), manifesting its catalytic activity. As analyzed by silver-stained SDS-PAGE, ΦAB6 TSP appeared to specifically degrade the



**Figure 2.** ESI-MS analysis of the  $\Phi$ AB6 TSP-digested products. The major peak in the ESI-MS spectrum, as indicated by a red arrow, which corresponds to an  $m/z$  value of 1727.6, was subjected to ESI-MS-MS analysis (the inset). The results indicated that the main digestion product comprises oligosaccharide of two repeat units with two pseudaminic acids, as described in the text.

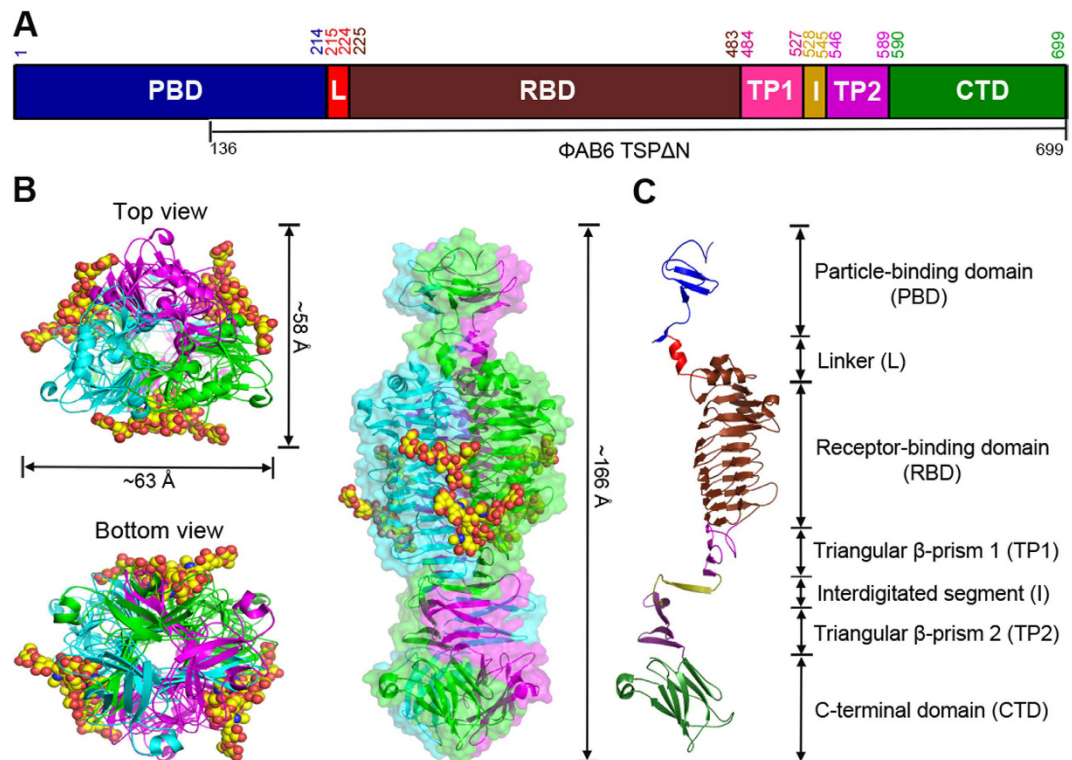
lipopolysaccharide of *Ab*-54149 (Fig. 1B), presumably to the O-antigen<sup>18</sup>. Similar to the reported bacteriophage TSPs<sup>13,14-16</sup>,  $\Phi$ AB6 TSP was shown to be homotrimeric in solution and resistant to SDS denaturation (Fig. 1C and D and Supplementary Fig. 2B). By contrast, its thermal stability ( $T_m \sim 55^\circ\text{C}$ ) was relatively low compared to other TSPs (Fig. 1E and Supplementary Fig. 2C)<sup>14</sup>.

**Characterization of hydrolyzed products.** We analyzed the whole extract and  $\Phi$ AB6 TSP-digested products of *Ab*-54149 EPS by 1D ( $^1\text{H}$  and  $^{13}\text{C}$ ) and 2D NMR (COSY, HSQC and HMBC). The results indicated the EPS to be made up of repeat units of  $\rightarrow 3$ - $\beta$ -GalNAc $p$ -(1  $\rightarrow$  3)- $[\beta$ -Glc $p$ -(1  $\rightarrow$  6)]- $\beta$ -Gal $p$ -(1  $\rightarrow$  ), and a majority of the repeat units contained a pseudaminic acid (Pse) connected to Glc $p$  via a  $\beta$ -Pse-(2  $\rightarrow$  6)- $\beta$ -Glc $p$  linkage (Supplementary Figs 3–8, Supplementary Table 1). The hydrolyzed products were further analyzed by LC-ESI-MS, which revealed a major product ( $m/z = 1727.6324$ ) of two repeat units with two Pse residues (Fig. 2). Moreover, the MS spectra also showed two minor peaks ( $m/z = 1411.50$  and  $1095.51$ ) corresponding to the fragments of two repeat units, one with a Pse and other without Pse, (Supplementary Fig. 9), in agreement with the findings from our NMR assignment.

**Overall structure of  $\Phi$ AB6 TSP $\Delta$ N.** Despite testing numerous crystallization conditions, we did not obtain suitable crystals using the full-length  $\Phi$ AB6 TSP. This is most likely due to the flexible linker between its N-terminal particle-binding domain and the central receptor-binding domain, as described for other TSPs<sup>14-16</sup>. We therefore produced an N-terminally-truncated form (residues 136–699, named  $\Phi$ AB6 TSP $\Delta$ N) based on sequence alignment with other TSPs (Fig. 3A).  $\Phi$ AB6 TSP $\Delta$ N retained nearly all of the enzymatic activity, homotrimeric assembly, and thermal stability of  $\Phi$ AB6 TSP (Fig. 1A and Supplementary Fig. 2B and C).  $\Phi$ AB6 TSP $\Delta$ N was then crystallized successfully and the crystal structure was solved by the SAD-phasing method via Se-labeled  $\Phi$ AB6 TSP $\Delta$ N and refined to 1.48-Å resolution, with one  $\Phi$ AB6 TSP $\Delta$ N trimer in the asymmetric unit (Table 1). The trimeric  $\Phi$ AB6 TSP $\Delta$ N is a compact, elongated molecule with an overall length of  $\sim 166$  Å and the diameters between 26 and 63 Å (Fig. 3B). The three subunits are tightly packed together to form a parallel, left-handed, superhelical twist. Each subunit exhibits a parallel  $\beta$ -helix in the main body and can be divided into the N-terminal particle-binding domain (PDB), the linker region (L), the receptor-binding domain (RBD), two triangular  $\beta$ -prisms (TP1 and TP2) connected by a highly interdigitated segment (I), and the C-terminal domain (CTD) (Fig. 3C).

**Structure of the  $\Phi$ AB6 TSP $\Delta$ N monomer.** The N-terminal PBD region (Ser136-Ile214) of each subunit comprises a three-stranded, antiparallel  $\beta$ -sheet flanked by twisted loops and a  $\beta$ -turn. It is followed by a  $\beta$ -hairpin of  $\sim 16$  Å in length. The loop connecting the second and third strands extrudes prominently from the  $\beta$ -sheet and contacts the N-terminal loops via two hydrogen bonds (H-bonds) (Fig. 3C). The N-terminal region appears to be the most flexible part as judged by the high average B factor ( $19.7 \text{ \AA}^2$ ) compared to that of the whole subunit ( $14.6 \text{ \AA}^2$ ). This likely led to the broken electron densities around the N-terminal 18 residues in all subunits. The  $\beta$ -hairpin is succeeded by a short  $\alpha$ -helix (Thr215-Ile224) linking PBD and RBD. This linker is a common motif in the TSP structures reported thus far<sup>14-16</sup>.

The central RBD region (Thr225-Ser483) folds into a right-handed, parallel  $\beta$ -helix composed of 8 complete rungs, with an extended  $\alpha$ , $\beta$ -mixed turn on top, where a 10-residue  $\alpha$ -helix caps the hydrophobic core of  $\beta$ -helix domain. This “cap” is preceded by a  $3_{10}$ -helix and an  $\alpha$ -turn (Fig. 3C). The main body of the  $\beta$ -helix is collapsed to an L-shaped (rungs 1–4) or a kidney-shaped (rungs 5–8) cross section with the first complete rung starting at Val270. Rungs 1–4 are organized into three strands, B1, B2, and B3, separated by turns T1, T2, and T3, as seen in other TSPs, while rungs 5–8 contain an extra strand (B1a) within the T1 region (Fig. 4A). The extra strands were also identified in the  $\beta$ -helix structures of several galacturonases<sup>19</sup>, but not in other TSPs (Supplementary Fig. 10). The three sets of parallel B1, B2, and B3 strands merge into the b1, b2 and b3  $\beta$ -sheets, respectively, and form the three faces of the  $\beta$ -helix domain. In the 8 complete rungs, turns T1 and T3 vary in length ranging between 2 and

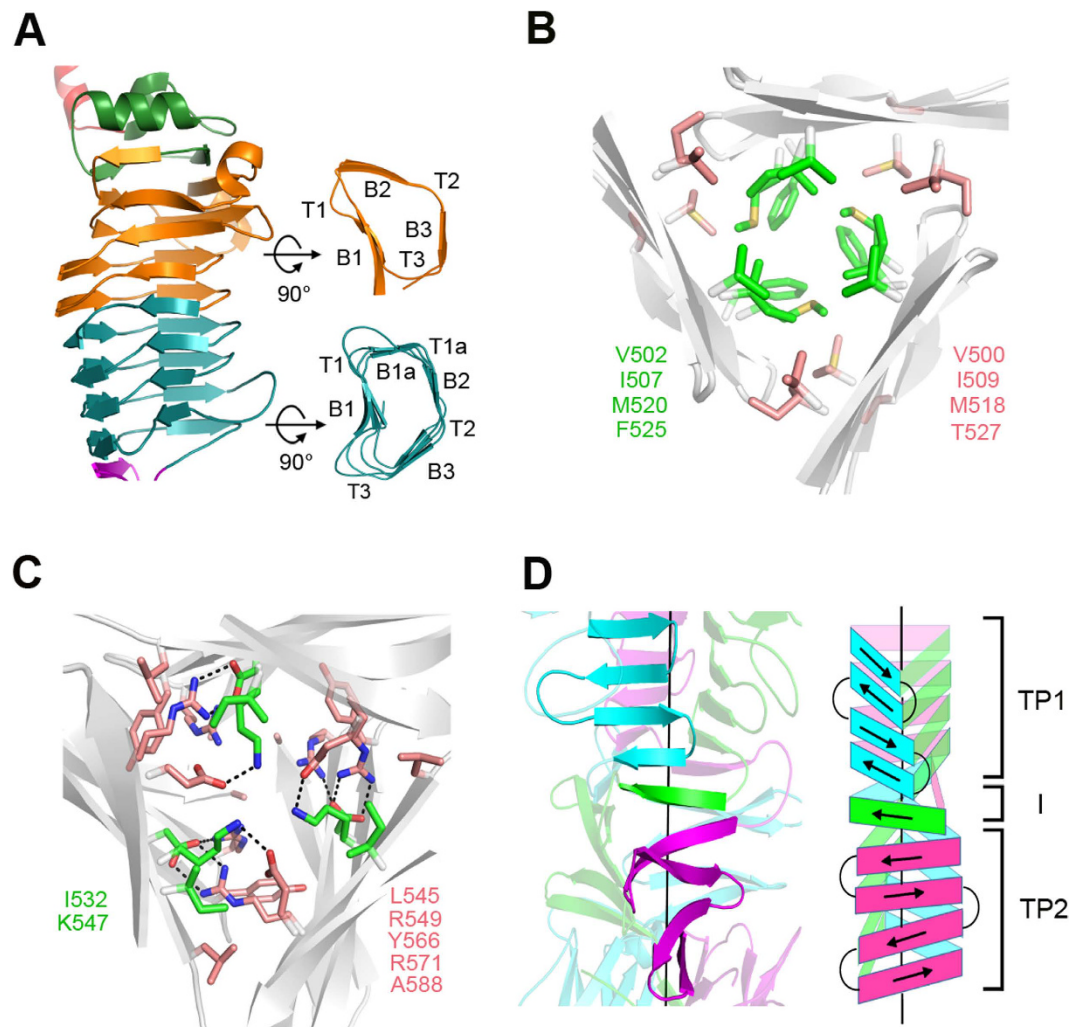


**Figure 3. Overall structure.** (A) Domains in full-length  $\Phi$ AB6 TSP are shown in a schematic representation. The numberings for amino acid residues at the start and end of each domain are further indicated. TSP $\Delta$ N includes residues 136–699. (B) The TSP $\Delta$ N trimer is shown as three-colored ribbons in three different views, with bound oligosaccharides shown as space-filled models. The view on the right is perpendicular to the trimer axis onto the intersubunit groove with the N-terminal end on top, and a transparent surface of the structure is shown as well. (C) The seven domains in a TSP $\Delta$ N monomer are colored as in (A).

10 residues and, together with strands B1, form a concave solvent-exposed groove (Fig. 4A). T3 of rung 1 contains a 5-residue  $3_{10}$ -helix. By contrast, the length of T2 is restricted to 2 residues, except for the second T2, which contains 14 residues, protrudes from the border of the  $\beta$ -helix, and forms a 7-residue  $\alpha$ -helix near the B3. The hydrophobic interior of the  $\beta$ -helix contains only two water molecules bound to the side chains of Gln355 and Ser375 and the main chains of Leu372, Ser375, Gly405, and His406. Moreover, there are a number of inward-orientated side chain stacks of cysteine (Cys417/Cys439), isoleucine (e.g. Ile335/Ile356/Ile380), valine (Val414/Val436), leucine (e.g. Leu349/Leu372), phenylalanine (e.g. Phe395/Phe419/Phe441/Phe471), and mixed-type amino acids (e.g. Phe243/Val270/Phe310/Ile335/Ile356/Ile380/Met408/Ile431/Leu458). These aliphatic and aromatic stacks, together with 116 H-bonds, presumably stabilize the  $\beta$ -helical conformation.

Lys484-Pro589 is a  $\beta$ -sheet region that forms the triangular  $\beta$ -prisms in the trimer. The polypeptide chain makes a nearly U-shaped turn towards the C-terminus and then folds into a four-stranded antiparallel  $\beta$ -sheet. This  $\beta$ -sheet extends  $\beta$ -sheet b of the  $\beta$ -helix domain towards the C-terminus. After this  $\beta$ -sheet, the polypeptide chain loops out with respect to the subunit longitudinal axis, makes a  $\sim 90^\circ$  bend and a sharp turn, and then forms the second four-stranded antiparallel  $\beta$ -sheet  $\sim 22$  Å away from the first four-stranded  $\beta$ -sheet, with a nearly parallel orientation. The segment (Gly528-Leu545) connecting the two four-stranded  $\beta$ -sheets tightly interdigitates with the neighboring subunits once the monomers assemble into a trimer as described below. In the CTD region (Met590-Ser699), the polypeptide chain is rotated by  $\sim 60^\circ$  around the longitudinal axis with respect to the second four-stranded  $\beta$ -sheet, tilted toward the axis by  $\sim 30^\circ$ , and organized as an independent  $\beta$ -sandwich formed by three-stranded and four-stranded antiparallel  $\beta$ -sheets. A short  $\alpha$ -helix is included in the connecting loop between the fifth and sixth strands.

**Structure of the  $\Phi$ AB6 TSP $\Delta$ N trimer.** In a trimer, three closely-packed subunits are related by a non-crystallographic triad, and  $\sim 39.8\%$  (11453 out of 28710 Å<sup>2</sup>) of the solvent-accessible surface area in each subunit is buried. In the N-terminal PBD region, three  $\beta$ -sheets, each from a different subunit, are laterally associated and stabilized by side chain H-bonds. The N-terminal twisted loops project into the other subunits and further strengthen the subunit association. Moreover, the  $\beta$ -hairpin of each subunit inserts into a cleft of its neighboring subunit, formed by the “linker” and “cap”  $\alpha$ -helices and a  $3_{10}$ -helix on top of  $\beta$ -helix domain, and makes extensive contact between the subunits. In the linker region, three  $\alpha$ -helices from three different subunits form a helix bundle stabilized by 6 sidechain H-bonds between the three pairs of Gln216 and Asn220, rather than hydrophobic contacts observed in other TSP structures<sup>14–16</sup>. In RBD, the three  $\beta$ -helices are packed laterally through the b2 and b3  $\beta$ -sheets to form a parallel, left-handed superhelix. At the interface of two adjacent  $\beta$ -helices, 35 residues



**Figure 4. Structures of the  $\beta$ -helix domain and the triangular  $\beta$ -prisms.** (A) The central, right-handed, parallel  $\beta$ -helix in  $\Phi$ AB6 TSP $\Delta$ N monomer. The first four rungs, the last four rungs, and the extended  $\alpha$ , $\beta$ -mixed turn on top are colored orange, deep teal, and green, respectively. (B) The interior of triangular  $\beta$ -prism 1 (TP1). The aliphatic side chain stacks at the center and the corners are colored green and salmon, respectively. The residues participating in the stacks are indicated on both sides. (C) The interior of the interdigitated segment and TP2. The black dotted lines depict the side chain hydrogen bonds between these residues. (D) TP1 and TP2 are connected by a highly interdigitated segment (I). The interdigitated segment merges into the  $\beta$ -sheets of TP1 and TP2 from different neighboring subunits, resulting in the formation of a nine-stranded, highly twisted, mixed  $\beta$ -sheet, as illustrated by cartoon on the right. The three subunits in  $\Phi$ AB6 TSP $\Delta$ N are colored cyan, green, and magenta.

are involved in the intersubunit interactions, of which 58% are H-bonds, 11% salt bridges, and 21% hydrophobic interactions. The exposed surface along the interface forms an elongated solvent-accessible groove of  $\sim 40$  Å in length and  $\sim 10$  Å in width, where the hydrolyzed products of *Ab*-54149 EPS were bound as described below. The central triangular channel created by the three  $\beta$ -helices is hydrophilic and accommodates many solvent molecules. Each wall of the channel contains two asparagine ladders, i.e. Asn343/Asn366 and Asn369/Asn397/Asn421.

Next to the RBD there are two triangular  $\beta$ -prisms, TP1 and TP2, connected by a highly interdigitated segment. In TP1, three four-stranded  $\beta$ -sheets from the three subunits assemble into the first  $\beta$ -prism with an exclusively hydrophobic interior, reminiscent of the  $\beta$ -prism in P22 TSP structure<sup>20</sup>. The triangular interior of TP1 is filled with regular stacks of aliphatic side chains: V502/I507/M520/F525 at the center and V500/I509/M518/T527 at each corner (Fig. 4B). Presumably these stacks stabilize the structure of TP1. Adjacent to TP1, the three polypeptide chains swap twice with the neighboring counterparts around the threefold axis and then assemble into the second twisted  $\beta$ -prism TP2. The interdigitated region and TP2 are stabilized by van der Waals association of side chains in the interior: I532/K547 at the center and L545/R549/Y566/R571/A588 at each corner (Fig. 4C), plus several H-bonds between Glu534, Lys547, Arg549, Asp564 and Arg571. In fact, the interdigitated segment is a 6-residue  $\beta$ -strand, and in a twisted manner this strand bridges the C- and N-terminal  $\beta$ -strands of TP1 and TP2, each from a different subunit in the trimer (Fig. 4D). As a consequence, the  $\beta$ -strands on each side of the  $\beta$ -prisms are indeed merged into a mixed, nine-stranded, strongly twisted  $\beta$ -sheet. Notably,  $\sim 44\%$  of intersubunit

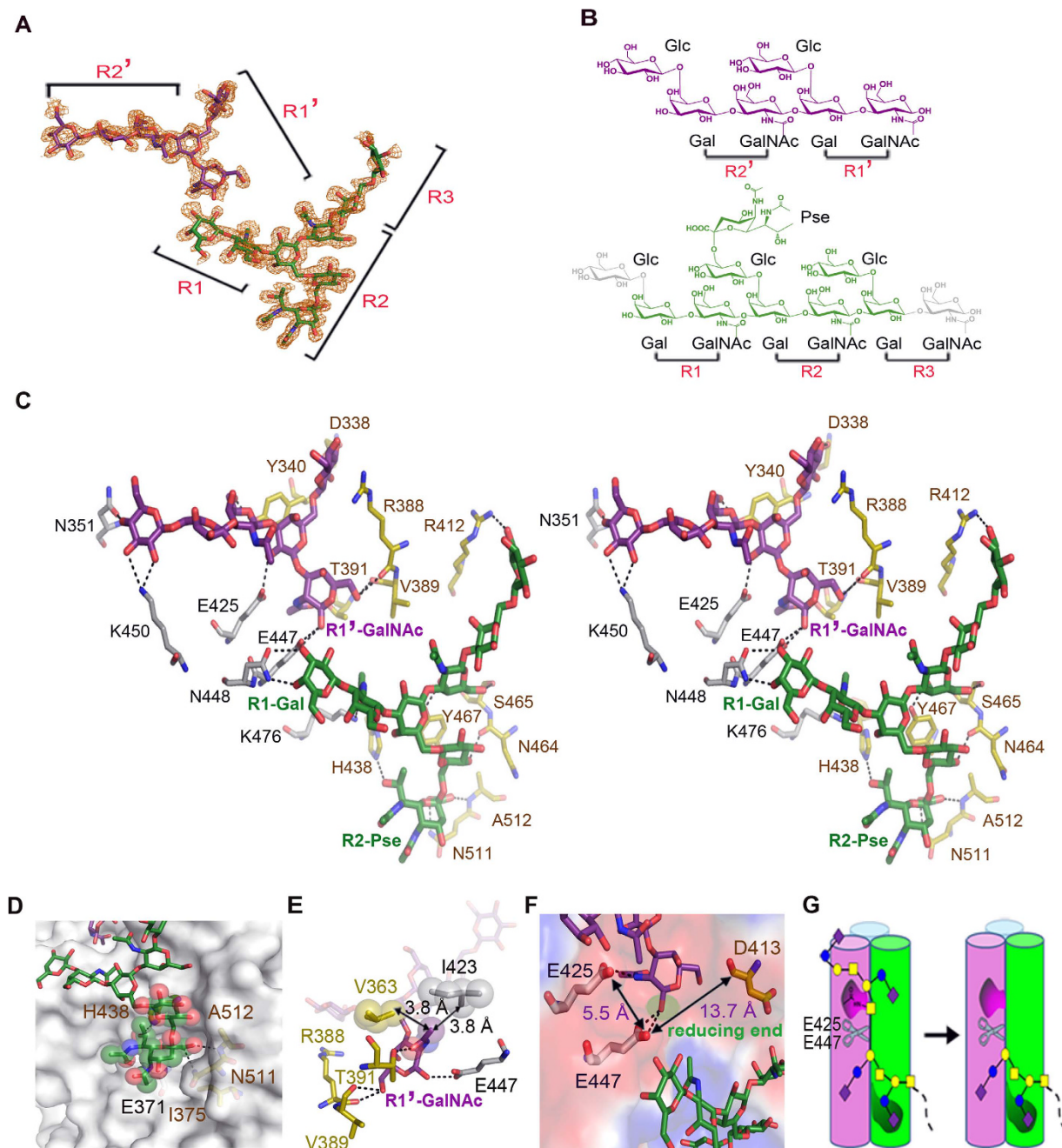
	Se- $\Phi$ AB6 TSP	$\Phi$ AB6 TSP	$\Phi$ AB6 TSP -5 R.U.	$\Phi$ AB6 TSP -3 R.U.
<b>Data collection</b>				
Wavelength (Å)	0.979299 Å	1.0 Å	1.0 Å	1.0 Å
Resolution (Å)	30–2.69 (2.79–2.69)	20–1.48 (1.53–1.48)	20–1.48 (1.53–1.48)	30–1.89 (1.96–1.89)
Unit cell dimensions				
<i>a</i> (Å)	135.45	135.02	135.81	134.93
<i>b</i> (Å)	78.21	77.98	78.39	77.95
<i>c</i> (Å)	248.18	248.08	247.74	247.74
$\alpha, \beta, \gamma$ (°)	90.0, 100.5, 90.0	90.0, 100.5, 90.0	90.0, 100.5, 90.0	90.0, 100.5, 90.0
Total observations	356,683	2,847,907	1,121,692	1,031,426
Unique reflections	69,784 (6,408)	406,268 (40,050)	399,721 (41,740)	203,248 (20,345)
Multiplicity	5.1 (5.0)	7.0 (7.6)	2.8 (2.9)	5.1 (4.0)
Completeness (%)	98.5 (91.3)	96.6 (95.3)	94.1 (99.0)	99.8 (99.9)
<i>I</i> / $\sigma$ ( <i>I</i> )	43.7 (21.0)	34.3 (3.3)	20.8 (2.0)	18.4 (2.7)
<i>R</i> <sub>merge</sub> (%)	8.0 (11.6)	5.4 (60.6)	5.4 (50.1)	13.1 (79)
<b>Refinement</b>				
Resolution (Å)		19.98–1.48	19.96–1.48	24.89–1.89
Reflections [ $>0\sigma(F)$ ], working/test		355,526/19,849	335,547/18,826	180,320/9,852
<i>R</i> factor/ <i>R</i> <sub>free</sub>		0.136/0.156	0.136/0.164	0.122/0.162
R.m.s.d., bond lengths (Å)/ angles (°)		0.004/1.062	0.005/1.175	0.005/1.127
Average <i>B</i> factor (Å <sup>2</sup> )/No. of atoms				
Protein		14.61/12537	14.66/12579	20.88/12564
Sugar			28.08/537	48.11/318
Malonate		26.71/63	28.99/21	43.59/42
Water		32.63/2786	32.88/2883	38.85/2138
Ramachandran plot, residues in (%)				
Most favored regions		88.0	88.2	88.4
Additionally allowed regions		11.0	10.8	10.6
Generously allowed regions		0.6	0.6	0.6
Disallowed regions		0.4	0.4	0.4
PDB code		5JS4	5JSD	5JSE

**Table 1. Data-collection and refinement statistics.** These crystals belong to the space group C2 and one asymmetric unit comprises one  $\Phi$ AB6 TSP trimer. Values in parentheses correspond to the highest resolution shell. R.U. is repeat units of *A. baumannii* exopolysaccharide.

interactions are concentrated in this region, despite the involvement of only ~19% residues, suggesting a critical role in maintaining the trimer structure. In the CTD, three  $\beta$ -sandwiches pack together to form a dome-like structure via the loop between the fourth and fifth  $\beta$ -strands in each  $\beta$ -sandwich (Fig. 3B,C). This loop slightly projects into the neighboring subunit and makes tight contacts with the fifth  $\beta$ -strand and the  $\alpha$ -helix. The interface between the  $\beta$ -sandwiches is mainly stabilized by side chain H-bonds.

**Structure of the carbohydrate-binding groove.** To obtain the structures of  $\Phi$ AB6 TSP bound to different products,  $\Phi$ AB6 TSP $\Delta$ N crystals were soaked in reservoir solution containing the whole extract of *Ab*-54149 EPS for 2, 5, 8, 12, 15, 18, and 24 hours. Finally, structures obtained at the soaking times of 12 and 15 hours, but not others, showed 14 and 8 sugar residues corresponding to 5 and 3 repeat units, respectively (Figs 3B and 5A,B and Supplementary Fig. 11). In both structures, the bound oligosaccharides lie in an elongated groove near the interface of two adjacent  $\beta$ -helices, reminiscent of the intersubunit carbohydrate-binding site in Sf6 TSP<sup>15</sup>. No significant conformational change was found in comparison with the free-form structure (r.m.s. deviation = 0.068 Å and 0.055 Å, respectively, for all C $\alpha$  atoms), indicating a rigid architecture of the enzyme. In the structure with 5-repeats, the oligosaccharide spans the groove from rung 4 of the  $\beta$ -helix to the TP1 of a neighboring subunit. The density for one Pse was observed (Figs 3B and 5A). Unexpectedly, the GalNAcp-(1  $\rightarrow$  3)-Galp glycosidic bond between the second and third repeat of the oligosaccharides indeed broke, resulting in two separate fragments in the binding groove, denoted R1-R2-R3 and R1'-R2' (R: repeat unit) (Fig. 5A,B). The retention of stereochemistry at the reducing end (R1'-GalNAcp) indicated a retaining hydrolysis mechanism of  $\Phi$ AB6 TSP<sup>21</sup>.

The  $\Phi$ AB6 TSP-carbohydrate interaction is primarily mediated by H-bonds. Twelve direct H-bonds from the R1'-R2' fragment to the residues Asn351, Glu425, Glu447 and Lys450 in one subunit and Asp338, Tyr340, Arg388, Val389, and Thr391 in a neighboring subunit stabilize R1'-R2' in the binding groove (Fig. 5C). In addition, there are 8 water-mediated H-bonds between R1'-R2' and the enzyme (Supplementary Fig. 12A). Towards the N-terminus, the groove is restricted by the T2  $\alpha$ -helix of rung 2, resulting in a ~70° bend between R1' and R2' (Figs 3B and 5A). As a consequence, R1' is parallel to the longitudinal axis while R2' is nearly perpendicular. The non-reducing end of R1'-R2' is completely outward oriented (Fig. 5A), suggesting that fragments



**Figure 5. Substrate binding.** (A) The  $1\sigma$   $2F_o - F_c$  electron density map around the bound oligosaccharides is overlaid with the refined model, colored green for the R1-R2-R3 fragment and purple for R1'-R2'. (B) Chemical structure of the bound oligosaccharides. In this diagram, the unobserved R1-Glc and R3-GalNAc residues are colored gray. R2-Glc is linked to a Pse. (C) Stereo view of the interactions between  $\Phi$ AB6 TSP and the bound oligosaccharides. The fragments R1-R2-R3 and R1'-R2' are colored as in (B). Direct hydrogen bonds contributed by two different subunits (colored yellow and gray) are drawn with dotted lines. (D) The interaction of R2-Pse (spheres) and the enzyme (surface) is viewed close-up. (E) R1'-GalNAc shows both hydrogen bonds and nonpolar interactions. (F) An acidic patch is formed by candidate residues for hydrolyzing the glycosidic bond of R1'-R1. (G) Schematic diagram for polysaccharide binding to the intersubunit groove of  $\Phi$ AB6 TSP. In this diagram, E447 and E425 form the catalytic dyad. R2-Pse serves as the recognition site. The N-acetyl group of R1'-GalNAc assists in positioning the anomeric carbon towards the nucleophilic residue Glu447. The representation of sugars is the same as in Fig. 2.

beyond R2' would not contact the enzyme. Likewise, the backbone of R1-R2-R3 fragment adopts an L-shaped conformation stabilized by 11 direct H-bonds to the residues Glu447, Asn448 and Lys476 in one subunit and Arg412, His438, Asn464, Ser465, Tyr467, Asn511 and Ala512 in a neighboring subunit (Fig. 5C). Besides, there

		$k_{cat}$ (min <sup>-1</sup> )	$K_m$ (mg • ml <sup>-1</sup> )	$k_{cat}/K_m$ (ml • mg <sup>-1</sup> • min <sup>-1</sup> )	Relative activity (%)
<b>ΦAB6 TSP</b>	WT	244 ± 14	3.79 ± 0.77	64.38 ± 18.1	100
	E425Q	11.8 ± 1.85	10.9 ± 4.7	1.08 ± 0.39	1.6
	E447Q	9.76 ± 1.43	7.89 ± 3.56	1.23 ± 0.4	1.9
	D413N	220 ± 12.5	3.9 ± 0.9	56.4 ± 13.8	87.6

**Table 2. Kinetic data of wild-type and mutant ΦAB6 TSPs.**

are 12 water-mediated H-bonds between R1-R2-R3 and enzyme (Supplementary Fig. 12A). Notably, three CH/π interactions were observed<sup>22</sup>, two from R1'-R2' to the residues Tyr340 and Tyr374 (Supplementary Fig. 12B,C) and one from R1-R2-R3 to Tyr467 (Supplementary Fig. 12D). Presumably the H-bond network and the CH/π interactions confer the EPS-binding specificity of ΦAB6 TSP.

The densities for R1-Glcp and R3-GalNAcp were not visible, most likely due to their outward locations. The other Glcp residues are also outwardly oriented, except for R2-Glcp. The Pse connects to R2-Glcp and fits well in a wide pocket made by Glu371, Ile375, His438, Asn511 and Ala512 from two different subunits (Fig. 5D). Three H-bonds to His438, Asn511, and Ala512 also stabilize the Pse in the pocket (Fig. 5C), implying that this Pse serves as a recognition site for ΦAB6 TSP. In addition, the N-acetyl group of R1'-GalNAcp is inserted into a small cavity stabilized by two H-bonds to Thr391 and Glu425 from different subunits (Fig. 5E). The methyl moiety of this N-acetyl group is in further contact with Val363 and Ile423 (Fig. 5E). Consequently, this orientation allows the anomeric carbon of R1'-GalNAcp to be in the proximity of Glu447 (Fig. 5F), a possible nucleophilic residue as described below.

The R1'-R2' fragment was released in the structure obtained at the 15-hour soaking time, leaving the R1-R2-R3 fragment in the binding groove (Supplementary Fig. 11), suggesting that this structure represents the post cleavage stage. Interestingly, the release of the R1'-R2' fragment is consistent with the observation of the digested products as revealed by MS spectra (Supplementary Fig. 9), implying that R1'-R2' represents a terminal fragment of the polysaccharides (Fig. 5G).

**Catalytic center.** It has been known that retaining glycoside hydrolases operate through a two-step mechanism by utilizing two carboxylate residues, one acting as a nucleophile and the other as an acid/base<sup>21</sup>. In the complex crystal structure, the distance between the anomeric carbon of R1'-GalNAcp and the side chain of Glu447 is only 3.3 Å, implying that Glu447 is the nucleophilic residue during catalysis. In fact, R1'-GalNAcp lies on a prominent acidic surface patch made up of the residues Glu425 and Glu447 in one subunit and Asp413 in a neighboring subunit (Fig. 5F). The distance from Glu447 to Glu425 is 5.5 Å; to Asp413, 13.7 Å (Fig. 5F). Considering the typical distance from the catalytic nucleophile to the general acid/base in a retaining glycoside hydrolase is between 4.5 and 5.5 Å<sup>23</sup>, Glu425 is a prime candidate to be the acid/base. To verify their catalytic roles, Asp413, Glu425 and Glu447 were individually mutated to their respective amide residues. Mutations at Glu425 and Glu447 drastically reduced the turnover rate of the enzyme with little effect on the binding affinity (Table 2 and Supplementary Fig. 13), whereas mutation at Asp413 only slightly decreased the enzyme activity. Consistent results were also obtained when spotting these mutant enzymes on a top agar inoculated with *Ab*-54149 (Fig. 1A). These results confirm that Glu425 and Glu447 make substantial contributions to catalysis. Consequently, the catalytic carboxylates of ΦAB6 TSP reside in the same subunit, unlike those of Sf6 TSP found in two different subunits.

## Discussion

*Acinetobacter baumannii* has captured significant attention recently in clinical and epidemic researches owing to a significant increase of its antibiotic-resistant strains. Using bacterial exopolysaccharide (EPS) to generate glycovaccines is one of the new treatments to combat bacterial infections, which have been applied to reduce the spreads of *Haemophilus influenzae type b*, *Streptococcus pneumoniae*, and *Neisseria meningitidis*<sup>24</sup>. By conjugating polysaccharide to carrier protein, glycoconjugate vaccine (GCV) can elicit a strong and long-lasting immune response compared to whole EPS<sup>11</sup>, making GCV useful in combating antibiotic-resistant bacteria<sup>25,26</sup>. A critical problem to GCV development is heterogeneity. It may be overcome by obtaining proper carbohydrate repeat units, which can induce stronger immune responses than whole bacterial EPS<sup>27</sup>. Because chemical synthesis of polysaccharide can be time-consuming with low yields<sup>28</sup> and chemical cleavage for bacterial EPS usually turns out a mixture of different sizes<sup>29</sup>, enzymatic digestion may provide an alternative. In this regard, bacteriophage TSPs can be employed to produce more homogenous fragments of bacterial EPS. The homogeneity and preferred size of oligosaccharides may be achieved through protein engineering.

In the present study, the TSP from the bacteriophage ΦAB6 (ΦAB6 TSP), which specifically hydrolyzed the EPS of *A. baumannii* strain 54149 (*Ab*-54149), was characterized. The whole extract and the ΦAB6 TSP-digested products of *Ab*-54149 EPS were also analyzed by NMR and LC-ESI-MS. Recently, the chemical structures of EPS of two antibiotic-resistant *A. baumannii* strains have been reported; one is aminoglycoside-resistant and another carbapenem-resistant<sup>30,31</sup>. Both polysaccharides indeed have the same structure. Interestingly, the structure of *Ab*-54149 EPS characterized here is the same as the two antibiotic-resistant *A. baumannii* strains, implying that substrates produced using the TSP may be useful for a GCV.

The structure of the N-terminally-truncated ΦAB6 TSP determined here revealed a trimeric β-helix architecture. The structure exhibits an organization similar to the structures of other *Podoviridae* TSPs, despite the lack of homology in amino acid sequence (Supplementary Fig. 14). A close comparison of these TSP structures revealed ΦAB6 TSP to be more elongated in both the monomer and trimer, likely due to the extensive



structure of triangular  $\beta$ -prisms as well as the fewer coils and smaller loop insertions in the  $\beta$ -helix domain (Supplementary Fig. 10). In contrast to the  $\beta$ -helix of other TSPs, which typically contains 13 rungs, the  $\beta$ -helix in  $\Phi$ AB6 TSP has only 8 rungs, and 4 of them contains the fourth  $\beta$ -strand never before seen in other TSPs. TP1 of  $\Phi$ AB6 TSP shares the structural feature of the C-terminal intertwined region of P22 TSP, i.e. with regular aliphatic side chain stacks in the center and corners of the triangular inner space. TP2 is unique in the TSP family, as it is more twisted and stabilized by mixed aliphatic/polar side chain stacks and side chain H-bonds. The N-terminal particle-binding domain of  $\Phi$ AB6 TSP is dominated by a  $\beta$ -stranded scaffold, similar to the N-terminal domain of P22 TSP, but the  $\beta$ -sheets in these two structures are nearly perpendicular to each other. By comparison, the N-terminal domain of  $\Phi$ AB6 TSP is significantly different from those of Sf6 TSP and HK620 TSP, which are organized into  $\alpha$ -helical bundles (Supplementary Fig. 10). In addition, the C-terminal  $\beta$ -sandwich domain of  $\Phi$ AB6 TSP resembles the C-terminal domain of Sf6 TSP and HK620 TSP, but the orientations of the  $\beta$ -sandwiches are quite different in these structures. As a consequence, the buried surfaces between the  $\beta$ -sandwiches are also different.

The structures of  $\Phi$ AB6 TSP in complexes with the products revealed the intersubunit carbohydrate-binding grooves. To date, although more than 30 right-handed  $\beta$ -helix structures have been reported<sup>15</sup>, the intersubunit carbohydrate-binding site can only be identified in the structures of SF6 TSP and an inulin fructotransferase from *Bacillus sp. snu-7*<sup>32</sup>. Notably, the inulin fructotransferase depolymerizes inulin by successively removing the terminal difructosaccharide units<sup>32</sup>. In the complex structure of  $\Phi$ AB6 TSP, the N-terminal side of the carbohydrate-binding groove is closed by the T2  $\alpha$ -helix of rung 2, and the distance between the reducing end of R1'-R2' fragment and the  $\alpha$ -helix is  $\sim 17$  Å. Because the backbone of a 3-repeat unit fragment is more than 22 Å in length, the binding groove in this N-terminal area could only accommodate 2 repeat units. This finding is also in agreement with the sizes of digested products as mentioned above. In this regard, this area of carbohydrate-binding groove represents a target for rational design and engineering of the enzyme to produce desired products. Indeed, in P22 TSP and HK620 TSP, the size of digested product also reflects the dimension of this N-terminal area of carbohydrate-binding groove<sup>16,33</sup>. Conceivably, these TSPs might first bind to the terminal fragments of bacterial surface polysaccharides and catalyze their successive removal to gain access to the host cell.

The branched Glcp residues in R1'-R2' and R1-R2-R3 are outwardly oriented and flexible in the complex structures, with the exception of R2-Glcp. As a consequence, the densities for the Pse residues were not visible, except for the one connected to R2-Glcp, indicating that most Pse residues are not involved in the binding to  $\Phi$ AB6 TSP. By contrast, the Pse at R2-Glcp fits well to a pocket in the binding groove, suggesting a recognition site in the substrate for  $\Phi$ AB6 TSP (Fig. 5G). This is reminiscent of the recognition site in *Salmonella* O-antigens for P22 TSP, which relies on a branched sugar residue as well<sup>34,35</sup>. It is notable that we did not observe the degradation products when the extract of *Ab*-SK44 EPS, which contains  $\beta$ -GalNAcp-(1  $\rightarrow$  3)- $\beta$ -Galp linkages but no Pse, was treated with  $\Phi$ AB6 TSP (data not shown)<sup>36</sup>. This also supports the recognition role of the Pse residue. On the other hand, the stoichiometric Pse in the main digestion product of *Ab*-54149 EPS correlates very well with the high Pse content of the polysaccharide, as estimated by using 1D NMR spectra (Supplementary Fig. 15).

Regarding the catalysis mechanism, mutations at the catalytic carboxylate residues of P22 TSP and Sf6 TSP to their respective amide reduced the enzyme activity to less than 0.1%<sup>15,34</sup>. By contrast,  $\sim 1.6\%$  and  $\sim 1.9\%$  activity were still detectable for the mutants E425Q and E447Q of  $\Phi$ AB6 TSP, respectively (Table 2), implying that an alternative but minor hydrolysis mechanism might coexist. By comparing the chemical structure of substrate of  $\Phi$ AB6 TSP with those of P22 TSP and Sf6 TSP, it is reasonable to assume that the N-acetyl group next to the anomeric carbon can act as an intramolecular nucleophile to proceed with the reaction. Indeed, this is common for certain N-acetylhexosaminidases<sup>37</sup>.

In conclusion, a new bacteriophage tailspike protein that specifically hydrolyzed the EPS of *Ab*-54149 was characterized. *Ab*-54149 EPS exhibited the same chemical structure as two other antibiotic-resistant *A. baumannii* strains. The structures of  $\Phi$ AB6 TSP in complexes with the semi-hydrolyzed products provide deep insights into the substrate recognition and product specificity of the enzyme and also elucidate a retaining hydrolysis mechanism, for which the catalytic residues have been verified by site-directed mutagenesis. These results constitute a structural basis for engineering the enzyme to produce desired oligosaccharides, which can be useful for the development of GCVs against *A. baumannii* infections.

## Methods

**Protein expression and purification.** The DNA fragments encoding the amino acid sequences 1–699 ( $\Phi$ AB6 TSP) and 136–699 ( $\Phi$ AB6 TSP $\Delta$ N) of the  $\Phi$ AB6 tailspike protein, respectively, were amplified from the phage genomic DNA and inserted into the vector pET28a (Novagen) via NdeI and XhoI cloning sites. After the sequences were confirmed, the vectors were transformed into the *E. coli* BL21(DE3) (Novagen). The cells were grown in Luria-Bertani (LB) medium supplemented with 50  $\mu$ g/mL kanamycin at 37 °C until the cell density reached OD<sub>600</sub> of 0.4–0.6. The cultured cells were induced with 0.1 mM IPTG at 20 °C overnight, and the cells were harvested by centrifugation (6,000 rpm) at 4 °C for 30 min and resuspended in buffer A (25 mM Tris-HCl and 100 mM NaCl, pH 7.5). The cells were lysed by passing through a French Press (Constant System Ltd, Constant System TS 2.2kw) three times and the lysate was clarified by centrifugation (20,000 rpm) at 4 °C for 60 min. The supernatant was loaded onto an open column filled with nickel-charged chelating resin (Qiagen) and pre-equilibrated with buffer A. The recombinant protein was eluted with 100–300 mM imidazole and the eluted fractions were pooled and then dialyzed against buffer A at 4 °C overnight. The recombinant proteins were further purified by a Superdex-200 gel-filtration column (GE-Healthcare), leading to near homogeneity. The protein was concentrated to  $\sim 10$  mg/mL by using a 30 K cut-off centrifuge filter (Millipore).

The expression vectors for the mutants D413N, E425Q, and E447Q were constructed using the QuikChange II site-directed mutagenesis Kit (Stratagene) following the manufacturer's instructions. The protein expression

and purification procedures were the same as described above. The Se-labeled  $\Phi$ AB6 TSP $\Delta$ N was prepared on the basis of a nonauxotrophic protocol using the commercially available Se-Met medium (Molecular Dimensions)<sup>38</sup>.

**Extraction of bacterial surface polysaccharides.** The crude extracts of bacterial surface polysaccharides were obtained on the basis of protocol reported by Zamze *et al* with several modifications<sup>39</sup>. Briefly, the bacterial cells were cultured with LB medium (for *A. baumannii*) or grown on LB plate (for *K. pneumoniae*) at 37 °C for 15 h and the cultured cells were collected. The cells were suspended in d.d. water and heated to 100 °C for 20 min to lyse the cells. The cell lysate was clarified by centrifugation at 10,000 rpm for 20 min, and the supernatant containing the bacterial surface polysaccharides was incubated with 80% acetone overnight to precipitate the polysaccharides. The precipitate dissolved in 10 mM Tris-HCl and 1 mM CaCl<sub>2</sub>, pH 7.5, was treated with ribonuclease (Sigma) and deoxyribonuclease I (Roche) at 37 °C for 6 h and then treated with proteinase K (Bioshop) for 12 hr. Subsequently, the sample was dialyzed against d.d. water by using a 1 kDa-cutoff membrane and then lyophilized. Finally, the crude polysaccharide extracts were further purified by a HW-65F gel-permeation column (TSK-GEL) to remove the contamination of bacterial organisms. The presence and concentration of the extracted polysaccharides were determined by the phenol-sulfuric acid method<sup>40</sup>.

**Digestion of *A. baumannii* surface polysaccharide by  $\Phi$ AB6 TSP.** Twenty mg crude extract of *Ab-54149* surface polysaccharide dissolved in 25 mM Tris-HCl and 100 mM NaCl, pH 7.5 was incubated with 500  $\mu$ g of purified  $\Phi$ AB6 TSP at 37 °C for 6 h, and then the digestion reaction was terminated by heating to 100 °C for 15 min. The denatured proteins were removed by centrifugation. Subsequently, the digested products were loaded onto a P-6 column (Bio-Rad) and the oligosaccharides were eluted with d.d. water. The eluted fractions were pooled and lyophilized for subsequent enzyme activity and chemical structure analyses.

**Top agar assay.** The top agar assay was performed according to the protocol reported previously<sup>41</sup>. Briefly, the LB agar in a petri dish was overlaid with 10 mL top agar pre-inoculated with the fresh culture of *Ab-54149*. After the top agar was solidified, 3  $\mu$ L (~1  $\mu$ g/ $\mu$ L) of either the wide-type, truncated, or mutant  $\Phi$ AB6 TSP were spotted on the petri dish and incubated overnight at 37 °C. The enzyme activity was evaluated by measuring the presence and dimension of translucent halos on the surface of the agar.

**Structure determination of *Ab-54149* surface polysaccharide by NMR.** The digested products or whole polysaccharide of *Ab-54149* surface polysaccharide were dissolved in 99.95% D<sub>2</sub>O. NMR experiments were performed using Bruker Avance 500 MHz spectrometer equipped with a cryoprobe, and the spectra were acquired at 298 K. 1D and 2D spectra were obtained using standard Bruker software, and Bruker TopSpin 2.1 program was used to process the NMR data. All two-dimensional homo and heteronuclear experiments (correlation spectroscopy, COSY; Overhauser effect spectroscopy, NOESY; heteronuclear multiple quantum coherence, HSQC; heteronuclear multiple bond correlation, HMBC) were carried out with the standard pulse sequences provided by Bruker. The mixing time for One-dimensional TOCSY experiment is 120 ms. The NOESY spectra were recorded at mixing time at 60 ms in order to identify genuine NOE effects. For homonuclear experiment, 256 FIDS of 2048 complex data point were collected with 20 scans per FID. For HSQC and HMBC spectra, 256 FIDS of 2048 complex data point were collected with 32 scans per FID, respectively. The assignment of protons chemical shifts were achieved by COSY, NOESY and One-dimensional TOCSY. In addition, the assignment of carbon chemical shifts was performed by HSQC and HMBC.

**Mass spectrometry analysis of digested products of *Ab-54149* exopolysaccharide.** LC-ESI-MS and LC-ESI-MS-MS analyses were done on a LTQ Orbitrap XL ETD mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with standard ESI ion source. 5  $\mu$ L of sample was injected at a flow rate of 50  $\mu$ L/min in 80% ACN/H<sub>2</sub>O with 0.1% FA by Ultimate 3000 RSLC system from Dionex (Dionex Corporation, Sunnyvale, CA). The conditions for full-scan MS are as follows: mass range m/z 0–6000 and resolution 60,000 at m/z 400. The target ions were sequentially isolated for MS2 by LTQ. Electrospray voltage was maintained at 4 kV and capillary temperature was set at 275 °C.

**Enzyme kinetic assay.** The hydrolytic activity of  $\Phi$ AB6 TSP towards polysaccharides was evaluated by quantifying the production of reducing end with the reagent 3,5-dinitrosalicylic acid (DNS), as described previously<sup>42</sup>. For common activity assay, the extract of *Ab-54149* surface polysaccharide dissolved in 20 mM HEPES/MES/sodium acetate, pH 5.0, was treated with the purified  $\Phi$ AB6 TSP at 37 °C. The final concentration of the polysaccharides was 5 mg/mL. The reaction was quenched at different time intervals by heating to 100 °C for 15 min, and then the denatured enzyme was removed by centrifugation. Subsequently, an aliquot of the digested products was mixed with an equal volume of DNS reagent (20 mg/mL) in 0.7 M NaOH, and the mixture was heated to 100 °C for 5 min. The hydrolytic activity of the enzyme was evaluated by measuring the absorption at 535 nm using an UV spectrophotometry. For kinetic study of the enzyme, the concentrations of polysaccharides were in the range of 2–60 mg/mL. The kinetic parameters were obtained by fitting the data of initial velocities versus polysaccharide concentrations with the Michaelis–Menten equation.

**Crystallization and X-ray data collection.** We failed to obtain crystals using the full-length  $\Phi$ AB6 TSP or using  $\Phi$ AB6 TSP $\Delta$ N in Tris-HCl buffer, so we tried to grow the crystals of  $\Phi$ AB6 TSP $\Delta$ N in a different protein buffer. The purified  $\Phi$ AB6 TSP $\Delta$ N was dialyzed against 25 mM HEPES and 100 mM NaCl, pH 7.5 and then concentrated to ~10 mg/mL. The initial crystallization screening of ~1,000 conditions was accomplished in the Core Facilities for Protein Structural Analysis (CFPSA), Academia Sinica (Taipei, Taiwan). The resulting initial conditions were further refined manually. Finally, two crystallization conditions were selected, i.e., (i) 1.0 M sodium malonate, pH 7.0 and (ii) 0.9 M sodium citrate and 0.1 M sodium cacodylate, pH 6.5. The crystals were grown

at 20 °C by mixing the  $\Phi$ AB6 TSP $\Delta$ N solution with equal volume of crystallization buffers via the hanging-drop vapor-diffusion method. The rhombus-shaped crystals with dimensions reaching 0.15 × 0.15 × 0.2 mm appeared within two days. The crystals for product-bound forms of  $\Phi$ AB6 TSP $\Delta$ N were obtained by soaking the free-form crystals into reservoir solution containing 5 mM surface polysaccharide of Ab-54149 at 20 °C for various periods between 1 and 24 hours. The crystals for Se-labeled  $\Phi$ AB6 TSP $\Delta$ N could only be obtained by using the second crystallization buffer. The single-wavelength anomalous diffraction (SAD) data at 2.69-Å resolution for crystals of Se-labeled  $\Phi$ AB6 TSP $\Delta$ N was collected at the beamline 12B2 of SPring-8 (Hyogo, Japan). The high-resolution data and those for different product-bound forms were collected at the beamlines 15A1, 13B1, or 13C1 of National Synchrotron Radiation Research Center (Hsinchu, Taiwan). Before being mounted on the goniometer, the crystals were briefly immersed in reservoir solution containing 12% (v/v) glycerol as cryoprotectant. All diffraction data were processed and scaled with the *HKL-2000* package<sup>43</sup>. The data collection statistics are listed in Table 1. The space group of the crystals is  $C_2$  with the typical unit cell dimensions of  $a = 135.5 \text{ \AA}$ ,  $b = 78.0 \text{ \AA}$ ,  $c = 248.0 \text{ \AA}$ ,  $\alpha = 90.0^\circ$ ,  $\beta = 100.5^\circ$ , and  $\gamma = 90.0^\circ$ . The asymmetric unit comprises a  $\Phi$ AB6 TSP $\Delta$ N trimer with an estimated solvent content of 65.25%.

**Structure determination and refinement.** The crystal structure of  $\Phi$ AB6 TSP $\Delta$ N was solved by the SAD-phasing method with the program *AutoSol* within the *PHENIX* software suite<sup>44</sup> and using the 2.69-Å resolution data collected at the wavelength of absorption peak. The positions of the 30 Se sites, with occupancies between 0.63 and 1.0, for the  $\Phi$ AB6 TSP $\Delta$ N trimer in the asymmetry unit were determined. The Se sites were then refined and the initial phases were improved by density modification with *AutoSol*. Approximate 93% model was automatically traced into the Se-phased electron density map with the program *AutoBuild*, and the remainder was manually built with *Coot*<sup>45</sup>. The resulting model was subjected to computational refinement with the program *REFMAC5*<sup>46</sup>. Throughout refinement, a randomly selected 5% of the data was set aside as a free data set, and the model was refined against the remaining data with  $F > 0$  as a working data set. The parameters for ideal protein geometry of Engh and Huber were used during refinement<sup>47</sup>. Subsequently, iterative rounds of model adjustment with *Coot* and refinement with *REFMAC5* were performed using the 1.48-Å resolution data set to improve the quality and completeness of the structure. The well-ordered malonate and water molecules were located with *Coot*. Finally, the refinement converged at a final  $R$  factor and  $R_{\text{free}}$  of 0.136 and 0.156, respectively, using anisotropic temperature factors. The stereochemical quality of the refined structure was checked with the program *PROCHECK*<sup>48</sup>. The final refinement statistics are listed in Table 1. With respect to the structures of product-bound  $\Phi$ AB6 TSP $\Delta$ N, the initial difference Fourier maps were obtained by using the refined structure of free-form  $\Phi$ AB6 TSP $\Delta$ N, and the subsequent refinements were the same as described above. The molecular figures were generated with *PyMOL* (Schrödinger, New York, USA).

**Analytical ultracentrifugation analysis.** Ultracentrifugation sedimentation experiments were performed at 20,000 rpm in a Beckman XL-A analytical ultracentrifuge (Beckman Instruments, Fullerton, Calif, USA) equipped with the standard double-sector centerpieces and using an An-60 Ti rotor. The UV absorption at 280 nm was scanned every 4 min for 250 scans. The data were analyzed with the software SEDFIT. The protein sample was in 25 mM Tris-HCl and 300 mM NaCl, pH 8.0. The sample was visually checked for clarity after ultracentrifugation. No precipitation was observed.

## References

- Antunes, L. C. S., Visca, P. & Towner, K. J. *Acinetobacter baumannii*: evolution of a global pathogen. *Pathog Dis* **71**, 292–301 (2014).
- Maragakis, L. L. & Perl, T. M. *Acinetobacter baumannii*: epidemiology, antimicrobial resistance, and treatment options. *Clin. Infect. Dis.* **46**, 1254–1263 (2008).
- Vila, J. & Pachon, J. Therapeutic options for *Acinetobacter baumannii* infections: an update. *Expert Opin. Pharmacother.* **13**, 2319–2336 (2012).
- Fournier, P. E., Vallenet, D., Barbe, V., Audic, S., Ogata, H., Poirel, L., Richet, H., Robert, C., Mangeot, S., Abergel, C., Nordmann, P., Weissenbach, J., Raoult, D. & Claverie J. M. Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. *PLoS genet.* **2**, 62–72 (2006).
- Yaqub, O., Castle-Clarke, S., Sevdalis, N. & Chataway, J. Attitudes to vaccination: a critical review. *Soc. Sci. Med.* **112**, 1–11 (2014).
- Garcia-Quintanilla, M., Pulido, M. R., Lopez-Rojas, R., Pachon, J. & McConnell, M. J. Emerging therapies for multidrug resistant *Acinetobacter baumannii*. *Trends Microbiol.* **21**, 157–163 (2013).
- Pace, D. Glycoconjugate vaccines. *Expert Opin. Biol. Ther.* **13**, 11–33 (2013).
- Vella, M. & Pace, D. Glycoconjugate vaccines: update. *Expert Opin. Biol. Ther.* **15**, 529–546 (2015).
- Avci, F. Y. & Kasper, D. L. How bacterial carbohydrates influence the adaptive immune system. *Annu. Rev. Immunol.* **28**, 107–130 (2010).
- Berti, F. & Adamo, R. Recent mechanistic insights on glycoconjugate vaccines and future perspectives. *ACS Chem. Biol.* **8**, 1653–1663 (2013).
- Avci, F. Y., Li, X., Tsuji, M. & Kasper, D. L. A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nat. Med.* **17**, 1602–1609 (2011).
- Salmond, G. P. C. & Fineran, P. C. A century of phage: past, present and future. *Nat. Rev. Microbiol.* **13**, 777–786 (2015).
- Freiberg, A. *et al.* The tailspike protein of *Shigella* phage Sf6. A structural homolog of *Salmonella* phage P22 tailspike protein without sequence similarity in the beta-helix domain. *J. Biol. Chem.* **278**, 1542–1548 (2003).
- Steinbacher, S. *et al.* Crystal structure of P22 tailspike protein: Interdigitated subunits in a thermostable trimer. *Science* **265**, 383–386 (1994).
- Muller, J. J. *et al.* An intersubunit active site between supercoiled parallel beta helices in the trimeric tailspike endorhamnosidase of *Shigella flexneri* Phage Sf6. *Structure* **16**, 766–775 (2008).
- Barbirz, S. *et al.* Crystal structure of *Escherichia coli* phage HK620 tailspike: podoviral tailspike endoglycosidase modules are evolutionarily related. *Mol. Microbiol.* **69**, 303–316 (2008).
- Lai, M. *et al.* The tail associated protein of *Acinetobacter baumannii* phage  $\Phi$ AB6 is the host specificity determinant possessing exopolysaccharide depolymerase activity. *PLoS ONE* **11**, e0153361 (2016).
- Fomsgaard, A., Freudenberg, M. A. & Galanos, C. Modification of the silver staining technique to detect lipopolysaccharide in polyacrylamide gel. *J. Clin. Microbiol.* **28**, 2627–2631 (1990).

19. Pickersgill, R., Smith, D., Worboys, K. & Jenkins, J. Crystal structure of polygalacturonase from *Erwinia carotovora* ssp. *carotovora*. *J. Biol. Chem.* **273**, 24660–24664 (1998).
20. Kreisberg, J. F., Betts, S. D. & King, J.  $\beta$ -Helix core packing within the triple-stranded oligomerization domain of the P22 tailspike. *Protein Sci.* **9**, 2338–2343 (2000).
21. Vuong, T. V. & Wilson, D. B. Glycoside hydrolases: Catalytic base/nucleophile diversity. *Biotechnol. Bioeng.* **107**, 195–205 (2010).
22. Tsuzuki S. CH/ $\pi$  interactions. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **108**, 69–95 (2012)
23. Zechel, D. L. & Withers, S. G. Glycosidase mechanism: anatomy of a finely tuned catalysis. *Acc. Chem. Res.* **33**, 11–18 (2000).
24. Astronomo, R. D. & Burton, D. R. Carbohydrate vaccines: developing sweet solutions to sticky situations? *Nat. Rev. Drug Dis.* **9**, 308–324 (2010).
25. Terra, V. S., Mills, D. C., Yates, L. E., Abouelhadid, S., Cuccui, J. & Wern, B. W. Recent developments in bacterial protein glycan coupling technology and glycoconjugate vaccine design. *J. Med. Microbiol.* **61**, 919–926 (2012).
26. Schumann, B., Pragani, R., Anish, C., Pereira, C. L. & Seeberger, P. H. Synthesis of conjugation-ready zwitterionic oligosaccharides by chemoselective thioglycoside activation *Chem. Sci.* **5**, 1992–2002 (2014).
27. Johnson, M. A. & Bundle, D. R. Designing a new antifungal glycoconjugate vaccine. *Chem. Soc. Rev.* **42**, 4327–4344 (2013).
28. Adamo, R., Nilo, A., Castagner, B., Boutureira, O., Berti, F. & Bernardes, G. J. Synthetically defined glycoprotein vaccines: current status and future directions. *Chem. Sci.* **4**, 2995–3008 (2013).
29. Perciani, C. T. *et al.* Conjugation of polysaccharide 6B from *Streptococcus pneumoniae* with pneumococcal surface protein A: PspA conformation and its effect on the immune response. *Clin. Vaccine Immunol.* **20**, 858–866 (2013).
30. Kenyon, J. J., Marzaioli, A. M., Hall, R. M. & De Castro C. Structure of the K2 capsule associated with the KL2 gene cluster of *Acinetobacter baumannii*. *Glycobiology* **24**, 554–563 (2014).
31. Senchenkova, S. N. *et al.* Structure of the capsular polysaccharide of *Acinetobacter baumannii* ACICU containing di-N-acetylpsudaminic acid. *Carbohydr. Res.* **391**, 89–92 (2014).
32. Jung, W. S. *et al.* Structural and functional insights into intramolecular fructosyl transfer by inulin fructotransferase. *J. Biol. Chem.* **282**, 8414–8423 (2007).
33. Andres, D., Baxa, U., Hanke, C., Seckler, R. & Barbirz, S. Carbohydrate binding of *Salmonella* phage P22 tailspike protein and its role during host cell infection. *Biochem. Soc. Trans.* **38**, 1386–1389 (2010).
34. Steinbacher, S. *et al.* Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proc. Natl. Acad. Sci.* **93**, 10584–10588 (1996).
35. Andres, D. *et al.* An essential serotype recognition pocket on phage P22 tailspike protein forces *Salmonella enterica* serovar Paratyphi A O-antigen fragments to bind as nonsolution conformers. *Glycobiology* **23**, 486–494 (2013).
36. Lees-Miller, R. G. *et al.* A common pathway for O-linked protein glycosylation and synthesis of capsule in *Acinetobacter baumannii*. *Mol. Microbiol.* **89**, 816–830.
37. Lemieux, M. J. *et al.* Crystallographic Structure of Human  $\beta$ -Hexosaminidase A: Interpretation of Tay-Sachs Mutations and Loss of GM2 Ganglioside Hydrolysis. *J. Mol. Biol.* **359**, 913–929 (2006)
38. Guerrero, S. A., Hecht, H. J., Hofmann, B., Biebl, H. & Singh, M. Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl. Microbiol. Biotechnol.* **56**, 718–723 (2001).
39. Zamze, S. *et al.* Recognition of bacterial capsular polysaccharides and lipopolysaccharides by the macrophage mannose receptor. *J. Biol. Chem.* **277**, 41613–41623 (2002).
40. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. Colorimetric method for determination of sugars and related substances. *Anal. Chem.* **28**, 350–356 (1956).
41. Verma, V., Harjai, K. & Chhibber, S. Characterization of a T7-like lytic bacteriophage of *Klebsiella pneumoniae* B5055: a potential therapeutic agent. *Curr. Microbiol.* **59**, 274–281 (2009).
42. Robert, S., Albert F., Jonathan, K. & Rainer J. Reconstitution of the thermostable trimeric phage P22 tailspike protein from denatured chains *in vitro*. *J. Biol. Chem.* **264**, 11750–11753 (1989).
43. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymol.* **276**, 307–326 (1997).
44. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D.* **66**, 213–221 (2010).
45. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst. D.* **60**, 2126–2132 (2004).
46. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures *Acta Cryst. D.* **67**, 355–367 (2011).
47. Engh, R. A. & Huber, R. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A.* **47**, 392–400 (1991).
48. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).

## Acknowledgements

This work was supported by grants from Ministry of Sciences and Technology (NSC101-2923-B-001-005-MY3, MOST-104-2811-M-001-065 and MOST104-2325-B-001-003 to SHW and NSC-101-2319-B-001-003 to AH-JW), Academia Sinica, and the Taiwan Protein Project (Grant No. MOST105-0210-01-12-01). We acknowledge the use of instruments in the Biophysics Core Facility, Scientific Instrument Center at Academia Sinica. LC-ESI-MS was performed in the Mass Core of Genomic Research Center, Academia Sinica. We thank the staffs of the beamlines 15A1, 13B1, and 13C1 in the National Synchrotron Radiation Research Center (Hsinchu, Taiwan) and the staffs of the beamlines 44XU and 12B2 in SPring-8 (Japan) for assistance in X-ray data collection and access to the synchrotron radiation centers.

## Author Contributions

Wu S.H. and Huang K.F. conceived and guided this project. Lin N.T. provided the expression vector for  $\Phi$  AB6 TSP. Lee I.M., Tu I.F. and Huang K.F. executed the crystallization, X-ray data collection, and structure determination. Lee I.M., Yang F.L., Liao J.H., Wu C.Y., Ren C.T. and Chang C.M. performed the characterization of  $\Phi$  AB6 TSP, structural analysis of Ab-54149 surface polysaccharide. Huang K.F., Lee I.M., Ko T.P., Wang A.H., and Wu S.H. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Lee, I.-M. *et al.* Structural basis for fragmenting the exopolysaccharide of *Acinetobacter baumannii* by bacteriophage  $\Phi$  AB6 tailspike protein. *Sci. Rep.* **7**, 42711; doi: 10.1038/srep42711 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017