

**SUPPLEMENTAL DATA TO:**

**CRYSTAL STRUCTURE OF THE MINOR PILIN FCTB REVEALS DETERMINANTS OF  
GROUP A STREPTOCOCCAL PILUS ANCHORING\***

**Linke, Christian<sup>1,2</sup>, Paul G. Young<sup>1,2</sup>, Hae Joo Kang<sup>1,2</sup>, Richard D. Bunker<sup>1,2</sup>, Martin J.  
Middleditch<sup>1,2</sup>, Tom T. Caradoc-Davies<sup>4</sup>, Thomas Proft<sup>2,3</sup> and Edward N. Baker<sup>1,2</sup>**

<sup>1</sup>School of Biological Sciences, <sup>2</sup>Maurice Wilkins Centre for Molecular Biodiscovery and <sup>3</sup>School of Medical Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand; <sup>4</sup>Australian Synchrotron, Clayton, Victoria 3168, Australia

Running title: Crystal structure of the minor pilin FctB

Address correspondence to: Edward N. Baker, School of Biological Sciences, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand. Phone +64-9-373-7599. Fax +64-9-373-7414, E-mail: [ted.baker@auckland.ac.nz](mailto:ted.baker@auckland.ac.nz)

**TABLE S1**

**Isopeptide bond between Spy0130 and Spy0128.** Daughter ions of the fragmentation of the parent ion LGETEK<sub>146</sub>SELIFK and DFEVPT<sub>311</sub> ( $m/z$  694.68<sup>3+</sup>) which are covalently linked through Lys-146 and Thr-311.

Observed (m/z)	Charge	Calculated (m/z) <sup>a</sup>	$\Delta_{\text{obs-calc}}$	Proposed structure	Ion type
147.11	+1	147.11	0	K	y <sub>1</sub>
171.12	+1	171.11	+0.01	LG	b <sub>2</sub>
235.10	+1	235.10	0	DF (-CO)	a <sub>2</sub> *
263.10	+1	263.10	0	DF	b <sub>2</sub> *
294.18	+1	294.18	0	FK	y <sub>2</sub>
300.16	+1	300.16	0	LGE	b <sub>3</sub>
392.13	+1	392.15	-0.02	DFE	b <sub>3</sub> *
407.25	+1	407.27	-0.02	IFK	y <sub>3</sub>
491.20	+1	491.21	-0.01	DFEV	b <sub>4</sub> *
520.35	+1	520.35	0	LIFK	y <sub>4</sub>
530.23	+1	530.25	-0.02	LGETE	b <sub>5</sub>
588.24	+1	588.27	-0.03	DFEVP	b <sub>5</sub> *
694.68	+3	694.70	-0.02	LGETEKSELIFK and DFEVPT (-H <sub>2</sub> O) <sup>b,c</sup>	parent
697.36	+2	697.38	-0.02	LGETEKSELIFK	parent-b <sub>6</sub> *
736.38	+1	736.42	-0.04	SELIFK	y <sub>6</sub>
747.87	+2	747.91	-0.04	LGETEKSELIFK and T (-H <sub>2</sub> O) <sup>b</sup>	parent-b <sub>5</sub> *
787.40	+2	787.44	-0.04	LGETEKSELIFK and PT (-H <sub>2</sub> O <sup>b</sup> , -H <sub>2</sub> O)	parent-b <sub>4</sub> * - H <sub>2</sub> O
796.39	+2	796.44	-0.05	LGETEKSELIFK and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-b <sub>4</sub> *
845.95	+2	845.98	-0.03	LGETEKSELIFK and VPT (-H <sub>2</sub> O) <sup>b</sup>	parent-b <sub>3</sub> *
856.43	+1	856.45	-0.02	LGETEK and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-y <sub>6</sub> -b <sub>4</sub> *
901.47	+2	901.50	-0.03	LGETEKSELIFK and EVPT (-H <sub>2</sub> O <sup>b</sup> , -H <sub>2</sub> O)	parent-b <sub>2</sub> * - H <sub>2</sub> O
910.47	+2	910.50	-0.03	LGETEKSELIFK and EVPT (-H <sub>2</sub> O) <sup>b</sup>	parent-b <sub>2</sub> *
1072.46	+1	1072.52	-0.06	LGETEKSE and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-y <sub>4</sub> -b <sub>4</sub> *
1185.55	+1	1185.61	-0.06	LGETEKSEL and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-y <sub>3</sub> -b <sub>4</sub> *
1298.62	+1	1298.69	-0.07	LGETEKSELI and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-y <sub>2</sub> -b <sub>4</sub> *
1445.64	+1	1445.76	-0.12	LGETEKSELIF and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-y <sub>1</sub> -b <sub>4</sub> *
1591.74	+1	1591.86	-0.12	LGETEKSELIFK and PT (-H <sub>2</sub> O) <sup>b</sup>	parent-b <sub>4</sub> *

<sup>a</sup>Monoisotopic mass, calculated using the web server

<http://db.systemsbiochemistry.net:8080/proteomicsToolkit/FragIonServlet.html>

<sup>b</sup>Loss of a water molecule to the covalent intersubunit linkage.

<sup>c</sup>Unfragmented parent ion.

**TABLE S2**

**C-terminal sequences preceding the sortase motif of pilin proteins.** Sequences are shown for the pilin proteins from *Streptococcus pyogenes*, *S. agalactiae* and *Corynebacterium diphtheriae*. The sortase motifs are highlighted in bold face and proline residues in red.

<i>Streptococcus pyogenes</i> M1 strain SF370		
Spy_0125 <sup>1</sup>	NATVSKTGITSDETLAFENNKEP <b>VVPTG</b>	Minor pilin, tip
Spy_0128 <sup>2</sup>	STEQETSTDKDMTITFTNKKDFE <b>VPTG</b>	Major pilin
Spy0130 <sup>3</sup>	<b>PEPHQP</b> DTTEKEK <b>PQ</b> KKRNGIL <b>PSTG</b>	Minor pilin, cell wall anchor
<i>Streptococcus pyogenes</i> serotype M3 MGAS315		
Cpa <sup>4</sup>	KATKASVKEDETVAFENRKDL <b>VPTG</b>	Minor pilin, tip
FctA <sup>5</sup>	KTDESADEIVVTNKRDT <b>QVPTG</b>	Major pilin
FctB <sup>6</sup>	<b>VKPIPPRQP</b> NIP <b>KTPLPLAG</b>	Minor pilin, putative cell wall anchor
<i>Streptococcus pyogenes</i> serotype M5 strain Manfredo		
Cpa <sup>7</sup>	SASENVTADKEVTFENRKDL <b>VPTG</b>	Minor pilin, tip
FctA <sup>8</sup>	KTDESADEIVVTNKRDT <b>QVPTG</b>	Major pilin
FctB <sup>9</sup>	<b>VKPIPPRQP</b> NIP <b>KTPLPLAG</b>	Minor pilin, putative cell wall anchor
<i>Streptococcus agalactiae</i> strain NEM316		
PilA <sup>10</sup>	EEGDKHLITNTHI <b>PPKGIIP</b> MTG	Minor pilin, tip
PilB <sup>11</sup>	IAYDKGSVKKDAQVQNKKV <b>TIPQ</b> TG	Major pilin
PilC <sup>12</sup>	ET <b>PPPTNPKPSQ</b> PL <b>FPQSFLPK</b> TG	Minor pilin, cell wall anchor
<i>Streptococcus agalactiae</i> strain 2603V/R		
GBS52 <sup>13</sup>	<b>VTPK</b> V <b>PSR</b> GGL <b>IPK</b> TG	Minor pilin (Seq. ident. 29 % to PilC)
GBS80 <sup>14</sup>	DITVDSADAT <b>PDTIK</b> NNKR <b>PSIP</b> NTG	Major pilin (Seq. ident. 22 % to PilB)
GBS104 <sup>15</sup>	IGYLEGNGKHLITNT <b>PKRPP</b> GV <b>FPK</b> TG	Minor pilin (Seq. ident. 43% to PilA)

GBS59 <sup>16</sup>	DYVANSNQKDATRVENKKVT <b>IPQ</b> TG	Major pilin
GBS67 <sup>17</sup>	YHEEGDKHLITNTH <b>IPPKGIIP</b> MTG	Minor pilin
GBS150 <sup>18</sup>	ET <b>PPPTNPKPSQPLFPQSFLPK</b> TG	Minor pilin, cell wall anchor

---

***Corynebacterium diphtheriae* strain NCTC 13129**

---

SpaA <sup>19</sup>	NESTNVLVEQKVKIDNKKKNAGFEL <b>PL</b> TG	Major pilin
SpaB <sup>20</sup>	<b>PGAPNVPSVPSPPSVTSPAPKKTPPRLA</b> FTG	Minor pilin, cell wall anchor
SpaC <sup>21</sup>	SGLITVEHP <b>QGKPWL</b> IKVANVSASTL <b>PL</b> TG	Minor pilin, tip
SpaD <sup>22</sup>	KDKFEGDDEVTLVSEIKNIKQGT <b>PKLP</b> MTG	Pilin, major pilin
SpaE <sup>23</sup>	<b>PSTPPP</b> GHT <b>PPLRETP</b> GSGDEKEREQGD <b>LAL</b> TG	Pilin, putative cell wall anchor
SpaF <sup>24</sup>	SGQVQAKGEGDKMILTVADTTAGEL <b>PK</b> TG	Pilin
SpaG <sup>25</sup>	FTISKNGEEIVAGAFKNELGKGV <b>KLPL</b> TG	Pilin
SpaH <sup>26</sup>	AVKIGQTATTTYDAKVENVKRDT <b>PDLPL</b> TG	Pilin, major pilin
SpaI <sup>27</sup>	<b>VP</b> GT <b>PKTPGK</b> PDL <b>PEKFR</b> KEVTDRLG <b>NT</b> G	Pilin, putative cell wall anchor

---

NCBI accession numbers: <sup>1</sup>BAC20340.1, <sup>2</sup>NP\_268517.1, <sup>3</sup>NP\_268519.1, <sup>4</sup>NP\_663902.1, <sup>5</sup>NP\_663904.1, <sup>6</sup>NP\_663906.1, <sup>7</sup>YP\_001127700.1, <sup>8</sup>YP\_001127702.1, <sup>9</sup>YP\_001127704.1, <sup>10</sup>NP\_735915.1, <sup>11</sup>NP\_735914.1, <sup>12</sup>NP\_735911.1, <sup>13</sup>NP\_687666.1, <sup>14</sup>NP\_687665.1, <sup>15</sup>NP\_687669.1, <sup>16</sup>NP\_688405.1, <sup>17</sup>NP\_688406.1, <sup>18</sup>NP\_688402.1, <sup>19</sup>NP\_940344.1, <sup>20</sup>NP\_940342.1, <sup>21</sup>NP\_940341.1, <sup>22</sup>NP\_938626.1, <sup>23</sup>NP\_938628.1, <sup>24</sup>NP\_938629.1, <sup>25</sup>NP\_940534.1, <sup>26</sup>NP\_940533.1, <sup>27</sup>NP\_940530.1

## FIGURE LEGENDS

FIGURE S1. **Superimposition of the hydrophobic core of FctB (turquoise) and the N-terminal domain of the major pilin Spy0128 (3B2M, brown) in a stereoview.** Sidechains of residues of the hydrophobic core are drawn in a stick diagram.

FIGURE S2. **Circular dichroism (CD) spectra of FctB and Spy0130.** The final spectra are the average of 5 measurements.

FIGURE S3. **Sequence alignment of FctB from *S. pyogenes* strain 90/306S (cloned in this work) with the mature form of Spy0130 from strain SF370 (UniProt identifier Q9A1S0).** N15 and Q69 participating in a hydrogen bridge in the hydrophobic core are highlighted by boxes, K110 by a grey box. Sequences were aligned using ClustalW.

## FIGURES

### FIGURE S1:

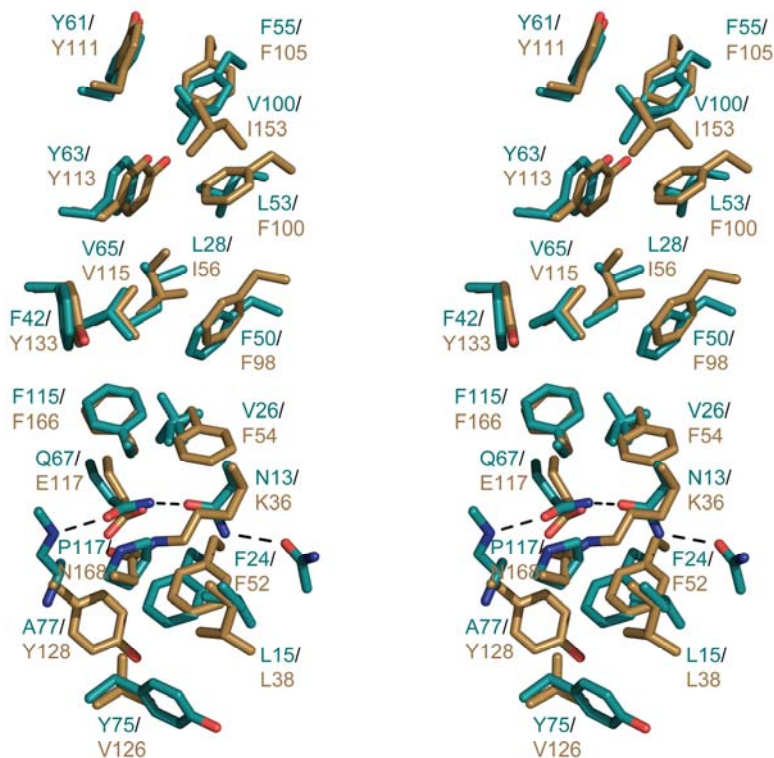


FIGURE S2:

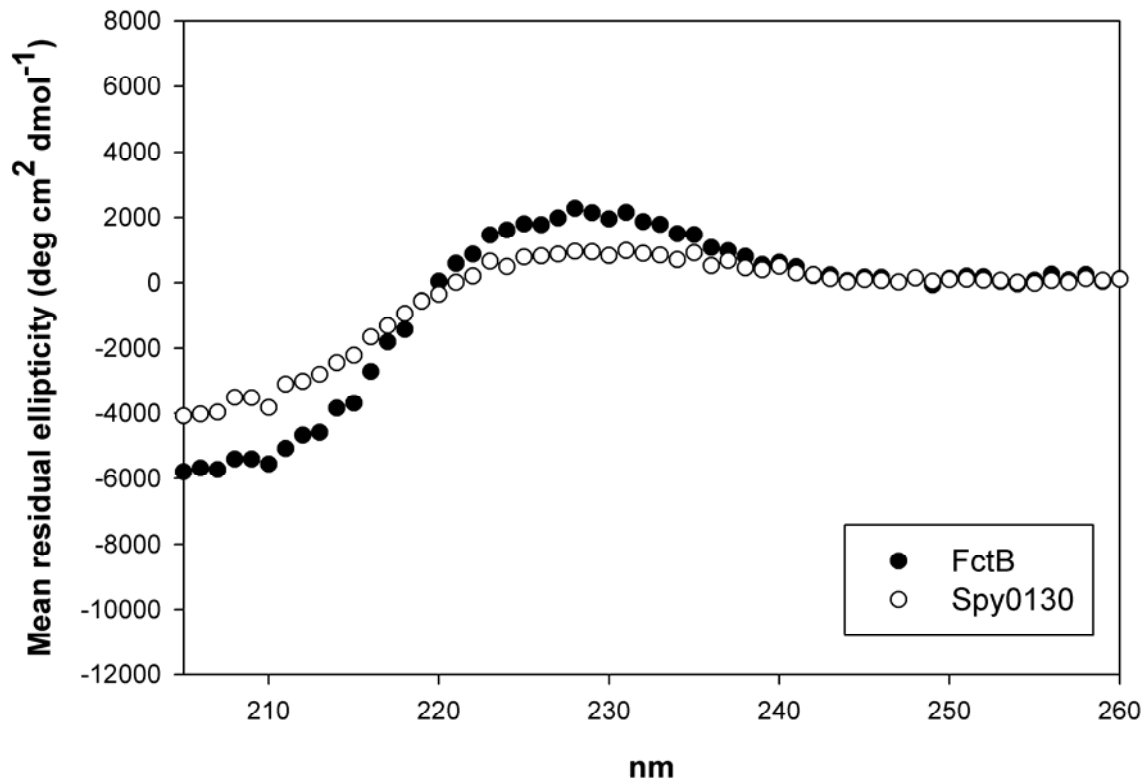


FIGURE S3:

```

FctB      KDSTVQTSISVENVLERAGDSTP-----FSVALESID---AMKTIEEITIAGSGKASFS 51
Spy0130   ENLTASINIEVINQVDVA1TNKQSSDIETFMFVIEALDKRESPLNSVTTTSVKGNGKTSFE 60
          :: *.. .*.* *1 :: * .. . * ..**:* * .. . :: *.**:**.

FctB      PL1F1TTVGQYTYRVY1QKPSQNKDYQADTTVFDVLVYV1TYDED-GTLVAKVISRRAGDEEK 110
Spy0130   QL1F1SEVGGYHYKIH1QLLGKNSQYHYDET1VYEVV1IYVLYNEQSGALETNLVS1NKLG1ETEK 120
          ***: **** *1 :: *1 ..**:* * **:*:*:* *1 *1 *1 *1 *1 *1 *1 *1 *1 *1 *1

FctB      SA1TFKPKRLVKPIPPRQPDFPKTP-----LPLA 139
Spy0130   SELIFKQ1EYSEKTPEPHQ1PD1TTEKEK1POK1KRNGILPST 158
          * : ** : * . *1 :*** .. . *
    
```