

Supplementary file S1

Cirocletto Blast Results

```
# BLASTP 2.2.25 [Feb-01-2011]
# Query: Ph2119
# Database: /labs/bat/www/tools/results/cirocletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
Ph2119 Ph2119 100.00 155 0 0 1 155 1 155 2e-118 322
Ph2119 C_perfingens 32.00 125 81 2 20 144 85 205 1e-19 72.4
Ph2119 Desulfotomaculum 33.60 125 74 3 20 142 16 133 5e-19 70.5
Ph2119 C_sartagoforme 34.17 120 74 3 24 143 36 150 6e-19 69.3
Ph2119 C_botulinum_ 32.79 122 79 2 22 142 17 136 2e-18 69.3
Ph2119 PGRP-Lb 27.97 143 82 5 14 142 27 162 4e-12 52.0
Ph2119 gPGRP-Le 27.91 129 82 4 23 142 28 154 3e-10 46.2
Ph2119 T7 27.73 119 81 3 27 141 15 132 8e-09 42.0
Ph2119 T3 26.89 119 82 3 27 141 15 132 1e-08 41.6
Ph2119 PGPP-Sa 22.97 148 93 5 17 151 30 169 7e-08 39.7
Ph2119 PGRP1_HOLDI 26.61 109 65 3 23 123 54 155 3e-06 35.4
# BLASTP 2.2.25 [Feb-01-2011]
# Query: IN93_T.aqua_gp24
# Database: /labs/bat/www/tools/results/cirocletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
IN93_T.aqua_gp24 IN93_T.aqua_gp24 100.00 305 0 0 1 305 1 305 0.0 627
IN93_T.aqua_gp24 P23-45_T.therm_p105 49.66 147 73 1 5 151 2 147 4e-37 124
IN93_T.aqua_gp24 P74-26_T.therm_p104 44.90 147 80 1 5 151 2 147 2e-35 120
IN93_T.aqua_gp24 P23-77_T.therm_VP29 31.58 152 86 8 29 170 76 219 9e-11 51.6
# BLASTP 2.2.25 [Feb-01-2011]
# Query: IN93_T.aqua_gp27
# Database: /labs/bat/www/tools/results/cirocletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
IN93_T.aqua_gp27 IN93_T.aqua_gp27 100.00 239 0 0 1 239 1 239 2e-177 478
IN93_T.aqua_gp27 P23-77_T.therm_gp33 46.20 184 95 2 58 239 52 233 1e-50 156
# BLASTP 2.2.25 [Feb-01-2011]
# Query: Rhodothermus_ph_RM37
# Database: /labs/bat/www/tools/results/cirocletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
Rhodothermus_ph_RM37 Rhodothermus_ph_RM37 100.00 233 0 0 1 233 1 233 7e-180 483
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P23-77_T.therm_gp12
# Database: /labs/bat/www/tools/results/cirocletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
P23-77_T.therm_gp12 P23-77_T.therm_gp12 100.00 178 0 0 1 178 1 178 8e-131 355
P23-77_T.therm_gp12 Geobacillus_ph_GBSV1 39.22 102 55 3 12 109 4 102 8e-14 57.4
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P23-77_T.therm_gp33
# Database: /labs/bat/www/tools/results/cirocletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
```

```

P23-77_T.therm_gp33    P23-77_T.therm_gp33    100.00 233    0    0    1    233    1    233    3e-176 474
P23-77_T.therm_gp33    IN93_T.aqua_gp27      46.20 184    95    2    52    233    58    239    1e-50 155
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P23-77_T.therm_VP29
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
P23-77_T.therm_VP29    P23-77_T.therm_VP29    100.00 369    0    0    1    369    1    369    0.0 746
P23-77_T.therm_VP29    P74-26_T.therm_p104    33.54 158    84    7    74    223    27    171    3e-15 65.5
P23-77_T.therm_VP29    P23-45_T.therm_p105    34.19 155    85    7    80    226    29    174    3e-14 62.4
P23-77_T.therm_VP29    IN93_T.aqua_gp24      31.58 152    86    8    76    219    29    170    1e-10 51.2
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P23-45_T.therm_p105
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
P23-45_T.therm_p105    P23-45_T.therm_p105    100.00 297    0    0    1    297    1    297    0.0 603
P23-45_T.therm_p105    P74-26_T.therm_p104    78.11 297    65    0    1    297    1    297    6e-170 463
P23-45_T.therm_p105    IN93_T.aqua_gp24      54.70 117    53    0    31    147    35    151    3e-36 122
P23-45_T.therm_p105    P23-77_T.therm_VP29    34.19 155    85    7    29    174    80    226    2e-14 62.4
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P23-45_T.therm_p108
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
P23-45_T.therm_p108    P23-45_T.therm_p108    100.00 165    0    0    1    165    1    165    2e-125 340
P23-45_T.therm_p108    P74-26_T.therm_p107    98.04 153    3    0    13    165    1    153    3e-114 311
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P74-26_T.therm_p104
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
P74-26_T.therm_p104    P74-26_T.therm_p104    100.00 297    0    0    1    297    1    297    0.0 603
P74-26_T.therm_p104    P23-45_T.therm_p105    78.11 297    65    0    1    297    1    297    8e-175 476
P74-26_T.therm_p104    IN93_T.aqua_gp24      50.43 117    58    0    31    147    35    151    8e-35 119
P74-26_T.therm_p104    P23-77_T.therm_VP29    33.54 158    84    7    27    171    74    223    2e-15 65.5
# BLASTP 2.2.25 [Feb-01-2011]
# Query: P74-26_T.therm_p107
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
P74-26_T.therm_p107    P74-26_T.therm_p107    100.00 153    0    0    1    153    1    153    5e-116 315
P74-26_T.therm_p107    P23-45_T.therm_p108    98.04 153    3    0    1    153    13    165    3e-114 311
# BLASTP 2.2.25 [Feb-01-2011]
# Query: Geobacillus_ph_GBSV1
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
Geobacillus_ph_GBSV1    Geobacillus_ph_GBSV1    100.00 233    0    0    1    233    1    233    1e-179 483
Geobacillus_ph_GBSV1    P23-77_T.therm_gp12    39.22 102    55    3    4    102    12    109    6e-14 58.2
# BLASTP 2.2.25 [Feb-01-2011]
# Query: Desulfotomaculum
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
Desulfotomaculum        Desulfotomaculum        100.00 207    0    0    1    207    1    207    1e-159 430
Desulfotomaculum        C_sartagoforme          44.00 125    70    0    19    143    35    159    6e-40 125
Desulfotomaculum        C_botulinum_            38.98 177    100    5    1    173    1    173    6e-39 124

```

Desulfotomaculum	C_perfingens	37.93	145	89	1	1	144	70	214	8e-38	121
Desulfotomaculum	T3	37.86	103	59	1	35	132	30	132	9e-23	80.5
Desulfotomaculum	T7	36.89	103	60	1	35	132	30	132	2e-22	79.3
Desulfotomaculum	Ph2119	33.60	125	74	3	16	133	20	142	6e-19	70.5
Desulfotomaculum	PGRP-Lb	36.63	101	57	3	37	133	65	162	1e-17	68.6
Desulfotomaculum	gPGRP-Le	30.99	142	84	5	14	143	24	163	1e-11	51.2
Desulfotomaculum	PGRP1_HOLDI	27.64	123	77	4	23	133	58	180	1e-11	51.6
Desulfotomaculum	PGPP-Sa	27.84	97	61	2	22	109	37	133	3e-08	41.6

BLASTP 2.2.25 [Feb-01-2011]

Query: C_perfingens

Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean

Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score

C_perfingens	C_perfingens	100.00	222	0	0	1	222	1	222	8e-169	454
C_perfingens	C_sartagoforme	43.48	161	90	1	55	215	2	161	7e-49	148
C_perfingens	Desulfotomaculum	42.28	123	70	1	92	214	23	144	9e-38	121
C_perfingens	C_botulinum_	44.53	137	73	3	87	220	17	153	2e-36	118
C_perfingens	T7	35.43	127	75	2	83	202	6	132	3e-25	87.8
C_perfingens	T3	35.59	118	69	2	92	202	15	132	1e-23	83.2
C_perfingens	PGRP-Lb	37.93	145	67	6	79	203	21	162	3e-21	78.6
C_perfingens	Ph2119	32.00	125	81	2	85	205	20	144	3e-19	72.0
C_perfingens	gPGRP-Le	38.55	83	43	2	90	164	30	112	3e-13	55.8
C_perfingens	PGRP1_HOLDI	26.49	151	78	6	26	161	3	135	1e-09	46.2
C_perfingens	PGPP-Sa	29.90	97	55	4	78	161	19	115	1e-06	37.4

BLASTP 2.2.25 [Feb-01-2011]

Query: C_sartagoforme

Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean

Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score

C_sartagoforme	C_sartagoforme	100.00	164	0	0	1	164	1	164	3e-128	347
C_sartagoforme	C_perfingens	43.48	161	90	1	2	161	55	215	5e-49	148
C_sartagoforme	Desulfotomaculum	44.00	125	70	0	35	159	19	143	5e-40	125
C_sartagoforme	C_botulinum_	40.30	134	76	3	30	159	13	146	2e-35	114
C_sartagoforme	T3	42.16	102	54	1	52	148	31	132	5e-26	87.8
C_sartagoforme	T7	41.18	102	55	1	52	148	31	132	2e-25	86.7
C_sartagoforme	PGRP-Lb	37.90	124	72	3	40	159	49	171	1e-20	75.5
C_sartagoforme	Ph2119	34.17	120	74	3	36	150	24	143	7e-19	69.3
C_sartagoforme	gPGRP-Le	29.37	143	88	4	29	159	22	163	4e-13	54.3
C_sartagoforme	PGRP1_HOLDI	28.37	141	82	5	25	149	43	180	6e-13	54.3
C_sartagoforme	PGPP-Sa	25.35	142	92	5	30	158	28	168	8e-09	42.7

BLASTP 2.2.25 [Feb-01-2011]

Query: C_botulinum_

Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean

Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score

C_botulinum_	C_botulinum_	100.00	224	0	0	1	224	1	224	3e-174	468
C_botulinum_	Desulfotomaculum	38.98	177	100	5	1	173	1	173	7e-39	124
C_botulinum_	C_perfingens	44.53	137	73	3	17	153	87	220	6e-36	117
C_botulinum_	C_sartagoforme	40.30	134	76	3	13	146	30	159	3e-35	114
C_botulinum_	Ph2119	32.79	122	79	2	17	136	22	142	3e-18	69.3
C_botulinum_	T7	31.75	126	81	2	33	154	27	151	4e-16	63.5
C_botulinum_	PGRP-Lb	32.56	129	68	4	22	136	39	162	1e-15	63.5
C_botulinum_	T3	30.95	126	82	2	33	154	27	151	2e-15	61.6
C_botulinum_	gPGRP-Le	31.48	108	68	3	36	139	52	157	7e-12	52.0

```

C_botulinum_  PGRP1_HOLDI  25.19  135  88  5  13  136  48  180  5e-09  44.3
C_botulinum_  PGPP-Sa 34.09  88  51  3  13  93  28  115  6e-09  43.9
# BLASTP 2.2.25 [Feb-01-2011]
# Query: T3
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
T3  T3  100.00  151  0  0  1  151  1  151  6e-117  317
T3  T7  94.04  151  9  0  1  151  1  151  6e-111  302
T3  C_sartagoforme 42.16  102  54  1  31  132  52  148  5e-26  87.8
T3  C_perfingens  33.86  127  77  2  6  132  83  202  7e-24  83.6
T3  Desulfotomaculum 37.86  103  59  1  30  132  35  132  6e-23  80.5
T3  C_botulinum_  30.95  126  82  2  27  151  33  154  1e-15  61.6
T3  PGRP-Lb 38.81  67  39  2  19  84  50  115  6e-12  51.6
T3  gPGRP-Le 37.31  67  40  2  20  84  41  107  1e-11  50.1
T3  Ph2119 26.89  119  82  3  15  132  27  141  1e-08  41.6
T3  PGRP1_HOLDI  28.92  83  53  2  13  89  56  138  5e-08  40.0
T3  PGPP-Sa 30.56  72  49  1  30  100  58  129  1e-07  39.3
# BLASTP 2.2.25 [Feb-01-2011]
# Query: T7
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
T7  T7  100.00  151  0  0  1  151  1  151  4e-117  318
T7  T3  94.04  151  9  0  1  151  1  151  6e-111  302
T7  C_sartagoforme 41.18  102  55  1  31  132  52  148  1e-25  86.7
T7  C_perfingens  35.43  127  75  2  6  132  83  202  2e-25  87.8
T7  Desulfotomaculum 36.89  103  60  1  30  132  35  132  2e-22  79.3
T7  C_botulinum_  31.75  126  81  2  27  151  33  154  2e-16  63.5
T7  PGRP-Lb 38.81  67  39  2  19  84  50  115  4e-12  51.6
T7  gPGRP-Le 37.31  67  40  2  20  84  41  107  1e-11  50.1
T7  Ph2119 27.73  119  81  3  15  132  27  141  8e-09  42.0
T7  PGRP1_HOLDI  37.25  51  32  0  39  89  88  138  5e-08  40.0
T7  PGPP-Sa 30.56  72  49  1  30  100  58  129  7e-08  39.7
# BLASTP 2.2.25 [Feb-01-2011]
# Query: T4
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
T4  T4  100.00  164  0  0  1  164  1  164  1e-125  340
# BLASTP 2.2.25 [Feb-01-2011]
# Query: gPGRP-Le
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
gPGRP-Le  gPGRP-Le  100.00  175  0  0  1  175  1  175  2e-136  369
gPGRP-Le  PGPP-Sa 49.29  140  71  0  28  167  34  173  1e-47  144
gPGRP-Le  PGRP-Lb 41.14  175  101  2  1  174  9  182  9e-44  135
gPGRP-Le  PGRP1_HOLDI 38.57  140  86  0  28  167  54  193  3e-37  118
gPGRP-Le  C_perfingens 38.55  83  43  2  30  112  90  164  2e-13  56.2
gPGRP-Le  C_sartagoforme 29.37  143  88  4  22  163  29  159  5e-13  54.3
gPGRP-Le  C_botulinum_ 31.48  108  68  3  52  157  36  139  6e-12  52.0
gPGRP-Le  Desulfotomaculum 30.99  142  84  5  24  163  14  143  1e-11  51.2
gPGRP-Le  T7 37.31  67  40  2  41  107  20  84  1e-11  50.1
gPGRP-Le  T3 37.31  67  40  2  41  107  20  84  1e-11  50.1

```

```

gPGRP-Le      Ph2119 27.91 129    82     4     28    154    23    142    4e-10 46.2
# BLASTP 2.2.25 [Feb-01-2011]
# Query: PGRP-Lb
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
PGRP-Lb PGRP-Lb 100.00 214    0     0     2     215    2     215    2e-158 428
PGRP-Lb gPGRP-Le 41.14 175   101    2     9     182    1     174    1e-43 135
PGRP-Lb PGRP1_HOLDI 39.26 163   98     1     15    177    34    195    5e-43 134
PGRP-Lb PGPP-Sa 38.99 159   94     3     18    175    17    173    4e-35 114
PGRP-Lb C_perfingens 37.93 145   67     6     21    162    79    203    2e-21 79.0
PGRP-Lb C_sartagoforme 37.90 124   72     3     49    171    40    159    1e-20 75.5
PGRP-Lb Desulfotomaculum 36.63 101   57     3     65    162    37    133    1e-17 68.6
PGRP-Lb C_botulinum_ 32.56 129   68     4     39    162    22    136    1e-15 63.5
PGRP-Lb Ph2119 27.97 143   82     5     27    162    14    142    6e-12 52.0
PGRP-Lb T7 38.81 67    39     2     50    115    19    84     6e-12 51.6
PGRP-Lb T3 38.81 67    39     2     50    115    19    84     8e-12 51.6
# BLASTP 2.2.25 [Feb-01-2011]
# Query: PGPP-Sa
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
PGPP-Sa PGPP-Sa 100.00 183    0     0     1     183    1     183    3e-141 382
PGPP-Sa PGRP1_HOLDI 47.46 177   93     0     1     177    21    197    2e-62 182
PGPP-Sa gPGRP-Le 49.29 140   71     0     34    173    28    167    2e-47 143
PGPP-Sa PGRP-Lb 38.99 159   94     3     17    173    18    175    4e-35 113
PGPP-Sa C_botulinum_ 34.09 88    51     3     28    115    13    93     1e-08 42.7
PGPP-Sa C_sartagoforme 25.35 142   92     5     28    168    30    158    1e-08 42.0
PGPP-Sa Desulfotomaculum 27.03 111   67     3     28    133    8     109    8e-08 40.4
PGPP-Sa T7 30.56 72    49     1     58    129    30    100    1e-07 38.9
PGPP-Sa Ph2119 22.97 148   93     5     30    169    17    151    2e-07 38.9
PGPP-Sa T3 30.56 72    49     1     58    129    30    100    2e-07 38.9
PGPP-Sa C_perfingens 29.90 97    55     4     19    115    78    161    1e-06 37.0
# BLASTP 2.2.25 [Feb-01-2011]
# Query: PGRP1_HOLDI
# Database: /labs/bat/www/tools/results/circoletto/database_all_7.fas0008968646.clean
# Fields: Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
PGRP1_HOLDI PGRP1_HOLDI 100.00 197    0     0     1     197    1     197    2e-153 413
PGRP1_HOLDI PGPP-Sa 47.46 177   93     0     21    197    1     177    2e-62 182
PGRP1_HOLDI PGRP-Lb 39.26 163   98     1     34    195    15    177    4e-43 134
PGRP1_HOLDI gPGRP-Le 38.57 140   86     0     54    193    28    167    4e-37 118
PGRP1_HOLDI C_sartagoforme 28.37 141   82     5     43    180    25    149    7e-13 54.3
PGRP1_HOLDI Desulfotomaculum 27.64 123   77     4     58    180    23    133    1e-11 51.6
PGRP1_HOLDI C_botulinum_ 25.19 135   88     5     48    180    13    136    4e-09 44.3
PGRP1_HOLDI C_perfingens 27.14 140   69     6     14    135    37    161    1e-08 42.7
PGRP1_HOLDI T3 28.92 83    53     2     56    138    13    89     7e-08 40.0
PGRP1_HOLDI T7 37.25 51    32     0     88    138    39    89     7e-08 40.0
PGRP1_HOLDI Ph2119 26.61 109   65     3     54    155    23    123    3e-06 35.4

```

GenBank accession numbers: Ph2119 (KF408298), IN93_T.aqua_gp24 (NP_777341.1), IN93_T.aqua_gp27 (NP_777344), Rhodothermus_ph_RM378 (NP_835657)
P23-77_T.therm_gp12 (YP_003169717.1), P23-77_T.therm_gp33 (YP_003169738.1), P23-77_T.therm_VP29 (YP_003169736.1), P23-45_T.therm_p105
(YP_001467958), P23-45_T.therm_p108 (YP_001467961), P74-26_T.therm_p104 (YP_001468074), P74-26_T.therm_p107 (YP_001468077), Desulfotomaculum
(YP_004516476.1), C_botulinum (YP_001921482.1), C_perfingens (WP_003465616.1), C_sartagoforme (WP_016208077.1), T3 (NP_523313.1), T4 (P00720.1),
T7 (AAB32819.1), PGPP-Sa (1SXR), PGRP-Lb (1OHT), PGRP1_HOLDI (BAD08316), gPGRP-Le (2CB3_A).

Supplementary Figure S2. Nucleotide sequence optimization of the gene encoding Ph2119 endolysin.

A

>Ph2119 endolysin nucleotide sequence 468bp

```
ATGAGGATACTAGAACCCTGGAATAGGTGGTACAGGCAAAGGGCGTGTACCGGATACGGGGAACCCCGCCCCACTATATTGTTCTCCACCACACCGCGGGGCCTGTGGA  
CCAAGCCCCCGAGGTGATTAGGGACTTCCACGAAAAGGGCCGGGGGTGGCCCCATATCGGCTACCATTACTTGGTTTACCAGGATGGAAGGGTCTACAAAACCCTTCCCA  
ATAATGCTATTCCCATTTGCGTAAGAGAGTTTAATCCGGTGAGCCTTTGCATTGCGGCGGTAGGGGACTTCTCCAGGGTCCAGCGTGGCCGGACAATGCCCCAGGGTGG  
AAGGCCCTTTTGAGCTAAAGGACGCCCTAGTCAAGGCGTACCCCAAAGCGGTTTTGGTTTTGCACAAGGAGCTAACCCAGACTACTTGCCCCGGCGTCCTGTCATGGGG  
CATGGTAGCGGAAAAAGGAGGAAAGTAG
```

B

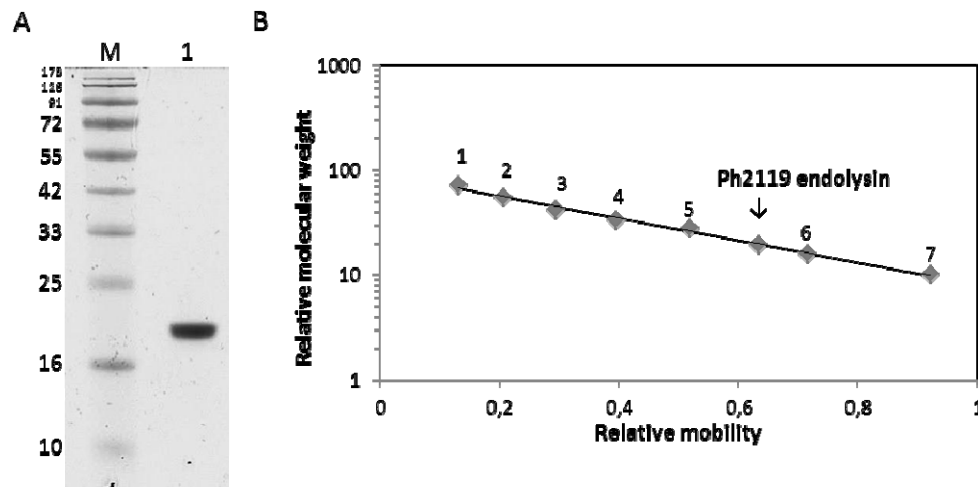
>Ph2119 endolysin optimized nucleotide sequence 468bp

```
ATGCGTATTCTGGAACCGTGAATCGTTGGTATCGTCAGAAAGGTGTTTATCGTATTCTGTTGGCACCCTC  
CGCATTATATTGTTCTGCATCATAACCGCAGGTCCGGTTGATCAGGCACCGGAAGTTATTCGTGATTTTCA  
CGAAAAAGGTCGTGGTTGGCCTCATATTGGTTATCATTATCTGGTTTATCAGGATGGTTCGCGTGTATAAA  
ACCCTGCCGAATAATGCAATTCCGATTTGCGTTCGTGAATTTAATCCGGTTAGCCTGTGTATTGCAGCAG  
TTGGTGATTTTAGCCAGGGTCCGGCATGGCCTGATAATGCACCGGGTTGGAAAGCACTGCTGGAAGTAA  
AGATGCACTGGTTAAAGCATATCCGAAAGCAGTTCTGGTTCTGCATAAAGAACTGACCCAGACCACCTGT  
CCGGGTGTTCTGAGCTGGGGTATGGTTGCAGAAAAAGGTGGTAAATAA
```

A. Nucleotide sequence of the gene coding for Ph2119 endolysin (468bp). GenBank accession number: KF408298. In the gene encoding Ph2119 endolysin there are six rare arginine codons: AGG at positions: 2, 9, 12, 44 and 67 and one AGA at position: 82. The Ph2119 gene is rich in the rare proline codon CCC occurring at eight positions: 6, 24, 40, 54, 73, 78, 125 and 141, and possesses four rare CUA leucine and two AUA isoleucine codons at positions: 4, 116, 120, 135 and 3, 19, respectively.

B. Optimized nucleotide sequence of the gene coding for Ph2119 endolysin (468bp). With the use of GeneOptimizer® technology (Life Technologies) rare arginine codons were changed into CGU at positions: 2, 9, 12, 44, 82 and CGC at position 67, CCC proline codon was

changed into CCG at positions: 6, 24, 40, 73, 78, 125, 141 and CCT at position 54. Leucine CUA and isoleucine AUA codons were changed for CUG and AUU, respectively, leading to a lack of rare codons in Ph2119 endolysin gene.



Supplementary Figure S3. Molecular weight determination of the Ph2119 endolysin. (A) SDS/PAGE (12.5%) of purified Ph2119 endolysin (lane 1) and marker proteins (lane M). The molecular masses of reference proteins (in kilodaltons) are indicated on the left. (B) Estimation of relative molecular weight (M_r) of the Ph2119 endolysin by SDS/PAGE (12.5%) run in parallel with M_r standards. Protein mobility is indicated relative to that of bromophenol blue marker. Marker proteins (PageRuler Prestained Protein Ladder, Thermo Scientific): 1, M_r 72,000; 2, M_r 55,000; 3, M_r 42,000; 4, M_r 33,000; 5, M_r 25,000; 6, M_r 16,000; 7, M_r 10,000.