SUPPLEMENTARY MATERIAL

The role of *Lactobacillus reuteri* Cell and Mucus Binding protein A (CmbA) in adhesion to intestinal epithelial cells and mucus *in vitro*

Hanne Jensen, Stefan Roos, Hans Jonsson, Ida Rud, Stine Grimmer, Jan-Peter van Pijkeren,

Robert A. Britton, Lars Axelsson*

Journal: Microbiology

* Corresponding author

Address: Nofima - Norwegian Institute of Food, Fisheries, and Aquaculture Research, P.O. Box 210, NO-1431 Ås, Norway

Email: lars.axelsson@nofima.no

Supplementary material, Figure S1

```
     aaggataattctaaaggtctacttggaataattgatcatttacaggcggtaaagtagaat

     gaaagtgttaaaaaataaattttgtttaaattttaaggcttgtatttaggaggggttaata

   1 ATGCTATCAAGAAAAAATTATAAGGAAACTATACGAAAACAGACACCTACAAAACAGTAC 60
   1 M  L  S  R  K  N  Y  K  E  T  I  R  K  Q  T  P  T  K  Q  Y   20

  61 TATACTATTAAGAAATTAACTGTTGGGGTTACTTCGGTATTAATTGGTCTATCCTTTATG 120
  21 Y  T  I  K  K  L  T  V  G  V  T  S  V  L  I  G  L  S  F  M   40
          YSIRK-G/S type signal sequence
 121 GGAGAACTAGAAGGGGATAGCGTTCATGCGGACACGATGACAGCAAGCAGTGAGTCAACA 180
  41 G  E  L  E  G  D  S  V  H  A  D  T  M  T  A  S  S  E  S  T   60
                                cleavage ↑
 181 AGTGTTACGTCGACGACTGCTCAGGATGGTTTAAAAAAATCTCCACAACTCTATTTGCAA 240
  61 S  V  T  S  T  T  A  Q  D  G  L  K  K  S  P  Q  L  Y  L  Q   80

 241 GTTACTGATACAAATAACCCAAGTACACCATTAAGTGCTTCATCCACAGGGACTAGTAAG 300
  81 V  T  D  T  N  N  P  S  T  P  L  S  A  S  S  T  G  T  S  K  100

 301 AATGTTACCTCATCAGCTGCGGTACAAGTGAAGTCCGCTAGTGATGAAGAAGATAGTGAT 360
 101 N  V  T  S  S  A  A  V  Q  V  K  S  A  S  D  E  E  D  S  D  120

 361 TCTACACTAGCTAAGGGAGAAAATAAATTTGCTCGGTCAGCAGTAAAAGATTCAGTCACT 420
 121 S  T  L  A  K  G  E  N  K  F  A  R  S  A  V  K  D  S  V  T  140

 421 GATGGGAAAACAAGTACAGCAGAAATTAATCCGGCAAAATTAAGCAGTCCTGCTTTAATA 480
 141 D  G  K  T  S  T  A  E  I  N  P  A  K  L  S  S  P  A  L  I  160

 481 ACGCAACTCAACCAATCCTTAGCTAAGAGCAGTACGAGTGATGCAGCAAAAGCTAATGAT 540
 161 T  Q  L  N  Q  S  L  A  K  S  S  T  S  D  A  A  K  A  N  D  180

 541 GAGTTAGAAATTAAAGCAACAGATCCGACTAATTATCCAAACTGTGGCGATGTGTATGGG 600
 181 E  L  E  I  K  A  T  D  P  T  N  Y  P  N  C  G  D  V  Y  G  200

 601 CCATTATTTGAATTGGATGCTAGCGGACAGCTTGTTAATAAAGATGAAGTTATATCTCTT 660
 201 P  L  F  E  L  D  A  S  G  Q  L  V  N  K  D  E  V  I  S  L  220

 661 AAAGATATGTATATTTTCCAAATATTGAAATTAGTAAATACAAAAGATAGTGACTTTCAA 720
 221 K  D  M  Y  I  F  Q  I  L  K  L  V  N  T  K  D  S  D  F  Q  240

 721 TATGTAATATTAACAATGAATCGTAAAGATACTGCAGATAGGTCTGTATATCTTTTTGTA 780
 241 Y  V  I  L  T  M  N  R  K  D  T  A  D  R  S  V  Y  L  F  V  260

 781 ACTGGAAGCAATTATAGTAATGCTGTTGTTGTTAAAGTAAAGCCAAATGATACTTATGAA 840
 261 T  G  S  N  Y  S  N  A  V  V  V  K  V  K  P  N  D  T  Y  E  280

 841 TTAAGTAAAACTGGATATAGTGTTACTTATACAGAACCAACAACTATAAATGGACATTAT 900
 281 L  S  K  T  G  Y  S  V  T  Y  T  E  P  T  T  I  N  G  H  Y  300

 901 GTTGATGGAACTTTTTATGTTACAGGAAGTACTTACGATGATGGTTTTATAATGCCAGAT 960
 301 V  D  G  T  F  Y  V  T  G  S  T  Y  D  D  G  F  I  M  P  D  320

 961 TGGCAACTGCAGCACCTTCAGATTATATATAGTTTAGGAAATTATGATCCAAGCAATACT 1020
 321 W  Q  L  Q  H  L  Q  I  I  Y  S  L  G  N  Y  D  P  S  N  T  340

1021 GACGCAACATCAGTTTGTGAAATAATGCCAAGTTATGAAAAGGTACCGGTAATTAAATAT 1080
 341 D  A  T  S  V  C  E  I  M  P  S  Y  E  K  V  P  V  I  K  Y  360

1081 AGTGGAGTACCTTCAAATATTAGCCAACCTAAGGTTTACATTACCGGGTTTACGGGTCAA 1140
 361 S  G  V  P  S  N  I  S  Q  P  K  V  Y  I  T  G  F  T  G  Q  380

1141 GAGTTTAACGTTACAGATATTATTAACAATTATAAGAAAGTTTTTAAGGGCTACTATCTT 1200
 381 E  F  N  V  T  D  I  I  N  N  Y  K  K  V  F  K  G  Y  Y  L  400
```

```
1201 CAAAATCCTAATGTGGCGTCCATGGGAACTCTTTCCCAATTTGAGAATGGTGGTTATTAC 1260
 401 Q  N  P  N  V  A  S  M  G  T  L  S  Q  F  E  N  G  G  Y  Y   420

1261 TTAAAGACATATTATGATAATGATGGTAATGTTGACTTTAAGGGCTTGTATCATCAAATT 1320
 421 L  K  T  Y  Y  D  N  D  G  N  V  D  F  K  G  L  Y  H  Q  I   440

1321 GATGATCAGGGAACAATGAGTGTGAGTGTTCTTAATGCAGATAATAAAACAATTGTTGGA 1380
 441 D  D  Q  G  T  M  S  V  S  V  L  N  A  D  N  K  T  I  V  G   460

1381 CCTGAAAATATTCTTGCTGGTAAATCGCATAACTTTAACTTTAATGGTCATAACTGGATT 1440
 461 P  E  N  I  L  A  G  K  S  H  N  F  N  F  N  G  H  N  W  I   480

1441 GCGCGGAATCCTTATGTCACTAGTTCAGCTCACGAAGTCATATTAAAGTATGCTAAGTTA 1500
 481 A  R  N  P  Y  V  T  S  S  A  H  E  V  I  L  K  Y  A  K  L   500

1501 GGTTCAGTTATTCCTGTTGATGAAAACGGAAATAAAATAAACGATGGATGGCAATATGTT 1560
 501 G  S  V  I  P  V  D  E  N  G  N  K  I  N  D  G  W  Q  Y  V   520

1561 AATGATCCAGATGATGCTTCCAAAGCCACTAGCCCATATGAAAAAGCGCCAGTTATCGAT 1620
 521 N  D  P  D  D  A  S  K  A  T  S  P  Y  E  K  A  P  V  I  D   540

1621 GGTTATGTAGCTGTAAATCCAGATGAAACGATCGTTCTTCCTCATAACTTAAGTAGTGAC 1680
 541 G  Y  V  A  V  N  P  D  E  T  I  V  L  P  H  N  L  S  S  D   560

1681 ACAAAGATTTATTACCGAAAGAGGATTAAAGTTACCTATAGTGGTAGTGACAGCAAGACC 1740
 561 T  K  I  Y  Y  R  K  R  I  K  V  T  Y  S  G  S  D  S  K  T   580
                                repeat 1→
1741 TACGATGGTAACCCAGCTAACTTCGAGCCAACGACAGTTCAGTGGAGTGGCTTGAAAGGA 1800
 581 Y  D  G  N  P  A  N  F  E  P  T  T  V  Q  W  S  G  L  K  G   600

1801 CTGAACACTTCAACCTTAACGTCCGCTGACTTCACGTGGAATACTGCGGATAAGAAGGCA 1860
 601 L  N  T  S  T  L  T  S  A  D  F  T  W  N  T  A  D  K  K  A   620

1861 CCAACGGATGCCGGTAAGTACACACTTAGTTTGAATACGACCGGAGAAGCAGCCTTACGT 1920
 621 P  T  D  A  G  K  Y  T  L  S  L  N  T  T  G  E  A  A  L  R   640

1921 AAGGCTAACCCGAACTATGATCTCAAGACAATTAGCGGTAGTTACACCTACACGATTAAT 1980
 641 K  A  N  P  N  Y  D  L  K  T  I  S  G  S  Y  T  Y  T  I  N   660

1981 CCACTAGGGATTGATAAAGTTACCTATAGTGGTAGTGACAGCAAGACCTACGATGGTAAC 2040
 661 P  L  G  I  D  K  V  T  Y  S  G  S  D  S  K  T  Y  D  G  N   680
                   repeat 2→
2041 CCAGCTAACTTCGAGCCAACGACAGTTCAGTGGAGTGGCTTGAAAGGACTGAACACTTCA 2100
 681 P  A  N  F  E  P  T  T  V  Q  W  S  G  L  K  G  L  N  T  S   700

2101 ACCTTAACGTCCGCTGACTTCACGTGGAATACTGCGGATAAGAAGGCACCAACGGATGCC 2160
 701 T  L  T  S  A  D  F  T  W  N  T  A  D  K  K  A  P  T  D  A   720

2161 GGTAAGTACACACTTAGTTTGAATACGACCGGAGAAGCAGCCTTACGTAAGGCTAACCCG 2220
 721 G  K  Y  T  L  S  L  N  T  T  G  E  A  A  L  R  K  A  N  P   740

2221 AACTATGATCTCAAGACAATTAGCGGTAGTTACACCTACACGATTAATCCACTAGGGATT 2280
 741 N  Y  D  L  K  T  I  S  G  S  Y  T  Y  T  I  N  P  L  G  I   760

2281 GATAAAGTTACCTATAGTGGTAGTGACAGCAAGACCTACGATGGTAACCCAGCTAACTTC 2340
 761 D  K  V  T  Y  S  G  S  D  S  K  T  Y  D  G  N  P  A  N  F   780
       repeat 3→
2341 GAGCCAACGACAGTTCAGTGGAGTGGCTTGAAAGGACTGAACACTTCAACCTTAACGTCC 2400
 781 E  P  T  T  V  Q  W  S  G  L  K  G  L  N  T  S  T  L  T  S   800

2401 GCTGACTTCACGTGGAATACTGCGGATAAGAAGGCACCAACGGATGCCGGTAAGTACACA 2460
 801 A  D  F  T  W  N  T  A  D  K  K  A  P  T  D  A  G  K  Y  T   820

2461 CTTAGTTTGAATACGACCGGAGAAGCAGCCTTACGTAAGGCTAACCCGAACTATGATCTC 2520
 821 L  S  L  N  T  T  G  E  A  A  L  R  K  A  N  P  N  Y  D  L   840

2521 AAGACAATTAGCGGTAGTTACACCTACACGATTAATCCACTAGGGATTGTGACTGTAAAT 2580
 841 K  T  I  S  G  S  Y  T  Y  T  I  N  P  L  G  I  V  T  V  N   860
```

```
2581 TACAAGGGCTATGATAAGAAAGTCTATGATGGTCAACCTGGAACGATTAATCCGGGTAAA 2640
 861 Y  K  G  Y  D  K  K  V  Y  D  G  Q  P  G  T  I  N  P  G  K   880

2641 TTAACGTGGAGTAAGTTGCCAGATGGTACTTCATTGAAGATGCCAACATGGAGTATAGAT 2700
 881 L  T  W  S  K  L  P  D  G  T  S  L  K  M  P  T  W  S  I  D   900

2701 GATTTCGCTTGGGAAACAGCTGATGGCTTAGCACCAACGGCAGTAGGAACTTATCGGATT 2760
 901 D  F  A  W  E  T  A  D  G  L  A  P  T  A  V  G  T  Y  R  I   920

2761 ATCTTGACGGATGCTGGTAAGGCTGCACTAAAGAAGATTAATCCAAATTATGACTTAAGC 2820
 921 I  L  T  D  A  G  K  A  A  L  K  K  I  N  P  N  Y  D  L  S   940

2821 AGTATTACTGGTGTCTTTACTTATGAAATTAAGCCAGCACAGACACCAGAAATCTTAGGC 2880
 941 S  I  T  G  V  F  T  Y  E  I  K  P  A  Q  T  P  E  I  L  G   960

2881 CAAACACCTGAGCAACAACCAGGCCAAAATACTAATCAATCAGGAGCTGAAAACGGCTTT 2940
 961 Q  T  P  E  Q  Q  P  G  Q  N  T  N  Q  S  G  A  E  N  G  F   980

2941 GGTTCTTCTACAAGGCCTAATGCATCAACTAACTCCAATCTTAATCAACTTCCACAGACT 3000
 981 G  S  S  T  R  P  N  A  S  T  N  S  N  L  N  Q  L  P  Q  T  1000

3001 GGTAATGAGCATTCTAATACTGCACTTGCTGGTCTAGCATTGGCTTTCTTGACTGCTATG 3060
1001 G  N  E  H  S  N  T  A  L  A  G  L  A  L  A  F  L  T  A  M  1020

3061 CTTGGTTTGGGCAAGAAGCGTAAACATGATTAGttattctaaagcttagtagattttaaa 3090
1021 L  G  L  G  K  K  R  K  H  D  *                              1030

     agctatgtagtgtttcgtaattgtttgagaaagagattagtgcttcgtcaagaagtactg

     atgagaaaatagaataagttttcaagcagctcgtgtctggaatttggcatgagctgttct
```

Fig S1. Nucleotide sequence and the deduced amino acid sequence of *cmbA* based on the draft genome sequence of *L. reuteri* ATCC PTA 6475 (GenBank: ACGX02000000). Features of the sequence are indicated as follows: underlined nucleotide sequence: putative promoter sequence (last nucleotide, position -60, predicted transcription start) and a predicted transcription termination sequence; dotted underlined nucleotide sequence: consensus Shine Dalgarno sequence; nucleotides in bold at positions 4 and 1221: these Cs were changed to Gs in the cloned gene to introduce and remove *Nco*I sites, respectively (see main text for details); amino acid sequence in bold: YSIRK-G/S like motif in signal sequence; underlined amino acid sequence: the cell wall anchoring motif LPXTG; dotted underlined amino acid sequence: membrane spanning domain; double underlined amino acid sequence: positively charged tail (membrane anchoring). The signal sequence and the tandem repeats are indicated under the amino acid sequence. Greyed out region represents the sequence missing in the cloned *cmbA* obtained in this study.

(Promoter predicted by: Neural Network Promoter Prediction, http://www.fruitfly.org/seq_tools/promoter.html; Transcription terminator predicted by: ARNold, http://rna.igmors.u-psud.fr/toolbox/arnold/; Signal sequence predicted by: SignalP 4.1 Server, http://www.cbs.dtu.dk/services/SignalP/)

Supplementary material, Table S1

Mutagenesis scheme for putative surface protein genes and sortase gene in *L. reuteri* 6475. All mutations were done by creating a stop codon early in the genes (leading to truncated proteins of a size 15% or less of the full proteins). The sites for the mutations were generally chosen based on the ease of creating a suitable restriction site used in the screening. Mutants were verified with sequencing. For details of the procedure, see van Pijkeren & Britton (2012).

Table S1. Mutation of putative surface protein genes and sortase gene in *L. reuteri* 6475.

Locus tag HMPREF0536	**Designation**	**Amino acid change**	**Recombineering oligo**[a] (5' -> 3')	**Screen oligo#1** (5' -> 3')	**Screen oligo#2** (5' -> 3')	**Screen oligo#3** (5' -> 3')
10633	*cmbA*	(I35X), (G36I)	tactattaagaaattaactgttggggt tacttcggtatta**TGAAT**tctatcctt tatgggagaactagaaggggatagcgt tcatgcgga	gcaaactcaaaatat gaagaagctatagaa	ctagttctcccataa aggatagaattca	NA[b]
10255	*srtA*	(V26X)	tccctactaagtaagacttaatctgtt ggttaaaaattaa**CTA**aactgatacca gtaacagtacaacaaccgctgtccacc gt	ttgtactgttactgg tatcagtttag	actttacaaaatcag aaaacatttgcgt	atcgattaataccat tggagcaattac
10146	*10146*	(A88X), (D89I)	tctttcgtttgtttaacttgattctta ttagaatcgacattttga**ATTCA**gtct gccaaattttgattagcaatattaaga tcattattt	aacgaatgggattaa agattagtttcaatg	attcttattagaatc gacattttgaattca	NA
11993	*11993*	(Y51X), (A52I)	gattgcaattgtgcaagttgctgatca gcagttgcccgaacttg**AATTC**aagta acttggtcattacttacttgtgcatta ttttgacta	gttgcaagagatgct tcagcattttt	caagtaagtaatgac caagttacttgaatt	cgtgggaacaggttt gaaaaaatttaaatt
10154	*10154*	(Y70X), (Q71A), (N72Y)	ccatgcaggtcattgattcctaaaatc tgaactggaatgtcat**AAGCT**tatttg gctaggtcagaccaatcagtagtcgtt tgcggagta	tgggatattaacggt gagtggtaaat	taaaatctgaactgg aatgtcataagct	agtatggttgcattc caaatgggatt
10802	*10802*	(D73E), (Q74A), (Q75X)	atttttgatatgtatcattaactaac tgttgactatcgattt**AAGCT**tcttta acttgttgtgcagttacaagagctgat tgattttca	acaagacaaaatgga ttgctatgtgg	ttcttaaggaaccac cagcatcattt	NA

[a] mutated bases are in bold uppercase; [b] NA, not applicable

Supplementary material, Table S2

Table S2. Primer/probes used in this study

Name	Direction	Oligonucleotide primer sequence
Primers used for cloning of cmbA		
cmbA-f1	Forward start	5'-ATGCACCATGGTATCAAGAAAAAATTATAAGGAAAC-3'
cmbA-f2	Forward middle	5'-ATCTTCAAAATCCTAATGTGGCGTCGATGGGAACTCTTTCCCAATTTG-3'
cmbA-r1	Reverse middle	5'-GACGCCACATTAGGATTTTGA-3'
cmbA-r2	Reverse end	5'-ATGCATCTAGACTAATCATGTTTACGCTTC-3'
Primers used for sequencing		
cmbA-seqf1	Forward (1)	5'-CTCCACAACTCTATTTGCAA-3'
cmbA-seqf2	Forward (2)	5'-GTGACTTTCAATATGTAATATTAA-3'
cmbA-seqf3	Forward (3)	5'-GTCGATGGGAACTCTTTCC- 3'
cmbA-seqf4	Forward (4)	5'-GGGCTATGATAAGAAAGTCTA-3'
cmbA-seqr3	Reverse	5'CGTTCCAGGTTGACCATCA-3'
Sip3	pSIP411 forward	5'-GTCTAAGGAATTGTCAGATAGGC-3'
Sip16	pSIP411 reverse	5'-ATTAGTCTCGGACATTCTGC-3'
Primer/Probe used for real-time PCR		
CmbA (lar_0958)	Forward	5'-ATCCAAACTGTGGCGATGTG-3'
CmbA (lar_0958)	Reverse	5'-AAGCTGTCCGCTAGCATCCA-3'
CmbA (lar_0958)	Probe[1]	6FAM-ATGGGCCATTATTTG-MGBNFQ
SecY (lr_0469)	Forward	5'-CCGCGTTTTGTTGAATGGA-3'
SecY (lr_0469)	Reverse	5'-TCGGGTTGCTTGATTAAGTTTTC-3'
SecY (lr_0469)	Probe[1]	6FAM-TAAACAAGGAGAAGTAGGACGG-MGBNFQ

[1]Taq probes, 6-FAM (6-carboxyfluorescein), MGB (6-carboxytetramethylrhodamine).

Supplementary material, Table S3

Table S3. Expression of *cmbA* (qRT-PCR)

Strain	**Fold change[a] Mean**	**Fold change[a] SD**
6475(pSIP-*cmbA*) vs 6475(pSIPΔ)[b]	355.9	88.2
6475 *cmbA*$^-$(pSIP-*cmbA*) vs 6475 *cmbA*$^-$(pSIPΔ)[b]	456.2	179.2
6475 *cmbA*$^-$(pSIP-*cmbA*) vs 6475(pSIP-*cmbA*)[b]	0.9	0.1

[a]Fold change in *cmbA* expression was calculated by the $\Delta\Delta C_T$ method ($2^{-\Delta\Delta CT}$). The presented values are mean±SD from two independent experiments.
[b]Comparison of SppIP induced cultures

Quantitative real-time reverse transcription PCR of *cmbA* expression

cmbA gene expression was determined by quantitative real-time reverse transcription PCR (qRT-PCR) in samples from Caco-2 adhesion assay to determine chromosomal and vector expression of *cmbA*. At the start of the Caco-2 adhesion experiments, an aliquot of the bacteria culture was mixed with RNA protect Bacteria Reagent (Qiagen) and frozen at -80 °C until RNA isolation. Purification and extraction of total RNA was done using the RNeasy Mini kit (Qiagen) according to manufacturer's instructions. NanoDrop was used to test the quality of RNA, and only samples with satisfactory quality were used for qRT-PCR. The qRT-PCR was performed as described previously (Rode *et al.*, 2007) using primers and Taqman® probes designed with Primer Express® Software v3.0 (Applied Biosystem) (primer and probes are listed in Table S1). The target site of the *cmbA* primer and probe set was upstream of the *cmbA* nonsense mutation. Relative gene expression was calculated by the ΔC_T method, using *secY* as the endogenous reference gene (Wall *et al.*, 2007). There was no difference in the amplification efficiencies of *cmbA* and *secY*. Fold change in gene expression was calculated using the $\Delta\Delta C_T$ method ($2^{-\Delta\Delta CT}$) (Livak & Schmittgen, 2001).

References

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25,** 402-408.

Rode, T. M., Langsrud, S., Holck, A. & Moretro, T. (2007). Different patterns of biofilm formation in *Staphylococcus aureus* under food-related stress conditions. *Int J Food Microbiol* **116,** 372-383.

van Pijkeren, J. P. & Britton, R. A. (2012). High efficiency recombineering in lactic acid bacteria. *Nucl Acids Res* **40,** e76.

Wall, T., Bath, K., Britton, R. A., Jonsson, H., Versalovic, J. & Roos, S. (2007). The early response to acid shock in *Lactobacillus reuteri* involves the ClpL chaperone and a putative cell wall-altering esterase. *Appl Environ Microbiol* **73,** 3924-3935.