STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 77 (2021)

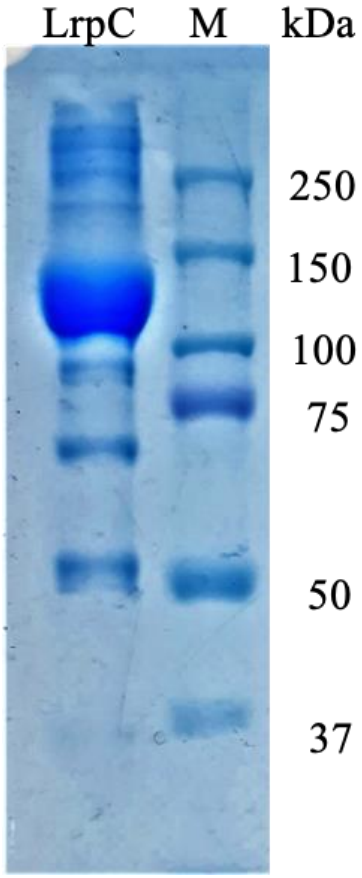
Supporting information for article:

LrpCBA pilus proteins of gut-dwelling *Ligilactobacillus ruminis*: crystallization and X-ray diffraction analysis

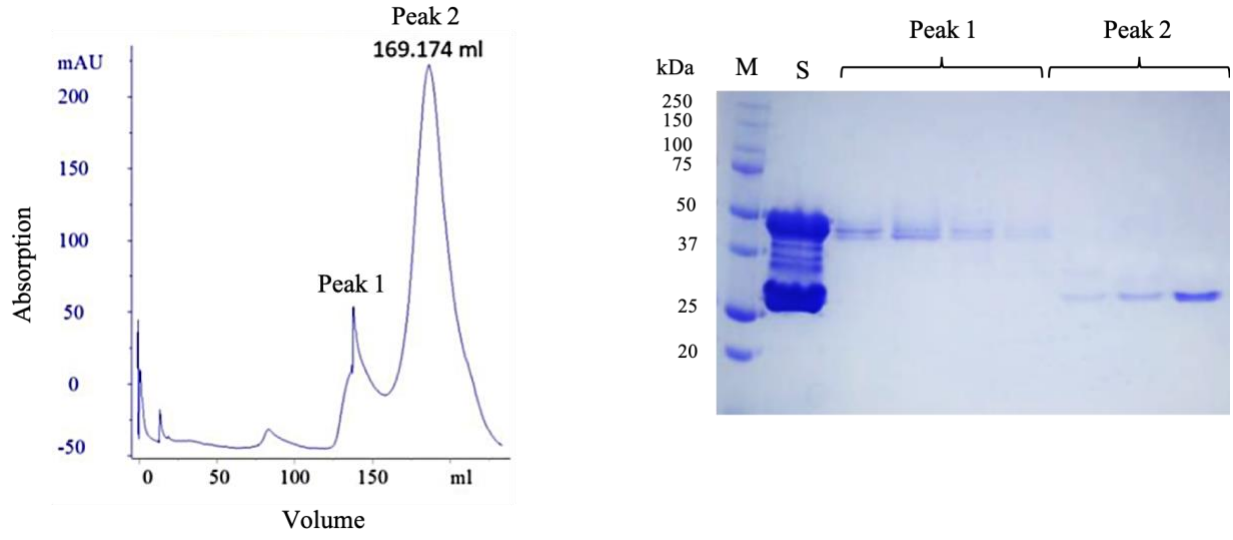
Amar Prajapati, Airi Palva, Ingemar von Ossowski and Vengadesan Krishnan

Supplementary Table S1. Primary structure-based search against the PDB. A protein sequence similarity search against the PDB was with the PDBeXplore browser (https://www.ebi.ac.uk/pdbe-srv/PDBeXplore/sequence/) using FASTA sequence of the *L. ruminis* pilins obtained from UniProtKB (LrpC [E7FRT3], LrpB [E7FRT4], and LrpA [E7FRT5]). The top hit for LrpA and LrpB is listed. No PDB entries with a similar sequence to LrpC were identified.

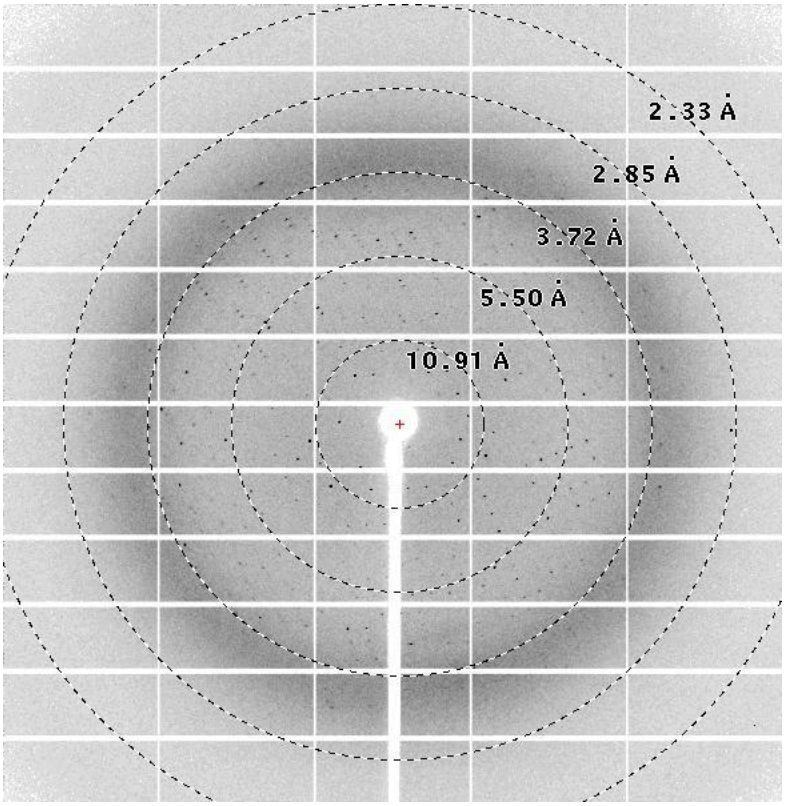
Query Sequence	Protein Hit	PDB ID	Maximum Percent Identity	Minimum E-value
LrpC	-	-	-	-
LrpB	*L. rhamnosus* GG basal SpaE	6JCH	27.4	1.1e-12
LrpA	*C. diphtheriae* backbone SpaD	4HSS	28.2	6.5e-12



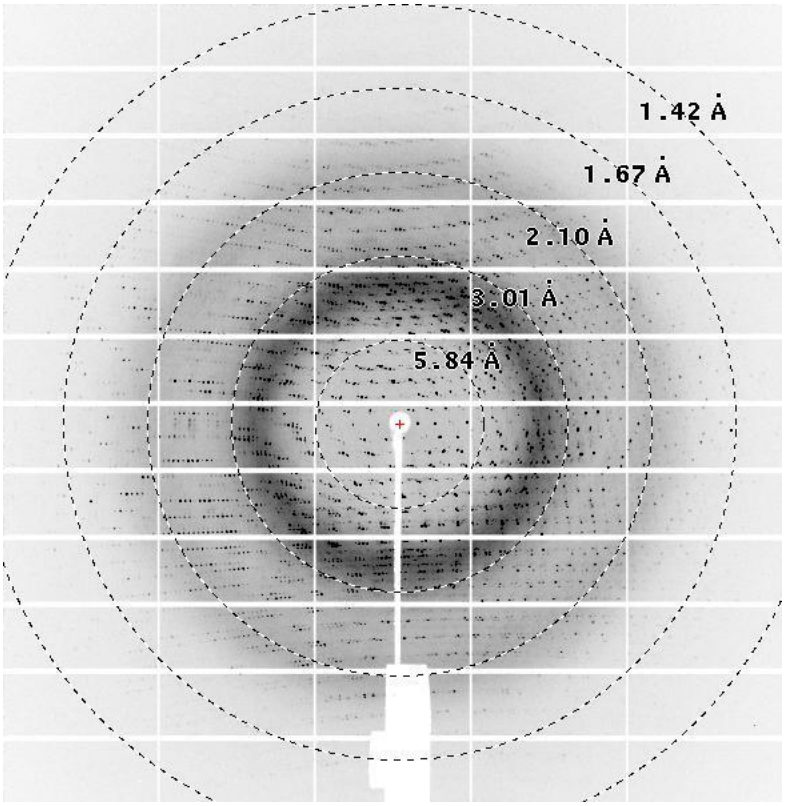
Supplementary Fig. S1. SDS-PAGE analysis of LrpC purified in the absence of glycerol. Purified LrpC (left lane) and molecular weight markers in kDa (right lane). Several lower molecular-weight bands corresponding to degraded LrpC fragments are visible (left lane) and were confirmed by mass spectrometric peptide mass fingerprinting (PMF) analysis.



Supplementary Fig. S2. Size-exclusion profile (left) and SDS-PAGE (right) analysis of a stable 30-kDa fragment (LrpA$_T$) generated by limited subtilisin proteolysis. LrpA$_T$ eluted at 169 ml and corresponded to a 30-kDa size on a High Prep 26/60 Sephacryl S-200 HR gel-filtration column (GE Healthcare) in 20 mM Tris–HCl pH 8.0 and 150 mM NaCl. Purified protein is homogeneous and pure as judged by size-exclusion profile and SDS-PAGE analysis. Molecular weight markers in kDa (M), subtilisin-treated sample (S), small peak fractions corresponding to full-length LrpA (Peak 1), and large peak fractions (Peak 2) corresponding to truncated protein (LrpA$_T$) are shown on a Coomassie blue-stained SDS-polyacrylamide gel.



Supplementary Fig. S3. X-ray diffraction pattern collected from LrpC crystal at synchrotron source. Circular dashed lines indicate resolution arcs.



Supplementary Fig. S4. X-ray diffraction pattern collected from $LrpA_T$ crystal at synchrotron source. Circular dashed lines indicate resolution arcs.

(a)

```
   1 MERNKIFKKL LCILGAVATV FAIVFAMGKF DGEKANAAVE LNQNNTVSFI
  51 DSVKIMVSHK DSTSEELKNN DVYILRDGDK ISIDYNFSFP DNADINPGDT
 101 YTIYVPRAFK AYNSCDGKLK NENGTEFGTW QMSGTFDNEY NGYPLVMTFN
 151 DNIKKIQGRK GTAHLESSLD IESFSEENRQ EIKIPLKDST TYNVEIKQHT
 201 DEKFGDVTKN GWFLGNDWQP KRAHWTIDFN KGLNDIENPV LKDNLYHADS
 251 QYSEHKIDYG SFNVYSVRLD SYGRVIEKSK EKVNPQEYEI VKSADGMSYE
 301 LHFKHKIKGA YRVEYETNIT NDSGNRETTK VGNKVSSYDG DKHLADASKE
 351 LWAHWWDRGS TIVKKGVQKD SNRVDWTIYY NLVGHKFGDS TILLDSLYRG
 401 TFDKNSFEIY EAASIDGNPF NQNENNIKIG RKLSSSEYSI AYATKYDREV
 451 ATIKIKNSDG KGYVIRYRSI LPDGLPQGTQ IKNRVKDGHD NEGEGSIWYD
 501 QVKEHVSKSH TNVTQNKIDW SVKFYEATSD NSFKTFKNLT LTDHFYTNRG
 551 NQLSLVGGLN AVKVYRSSGP GDWNTNILVD PGKYTIKEDN RSGEGGYKGF
 601 KITFKGETPA ALYEVKYQTT RDSSQESGNW AQFGDEQTSD SVPADGGISK
 651 NVGISKSNTG VHRTDGLLHW QVYVNADKLP MSNYMVKDTI TGDQTLIEST
 701 IKVKDLTDNK DVTSQVRREI RETTTTYNQD GKTHTSPAKL LIIHLPDTNH
 751 QYEISYGVKL GPVGYNSDKN HYTDYADLYQ GNDKKGSVTN DYWHWYSRLE
 801 KDGAVDENDN SLVNWTIDVN KAYVIYPNGA IKDTLDGKQI FVEDSVEVYR
 851 YEGSDQWHDN LSSKPLPKSA YNVKIVPYTT SDGRVTQQMI ITFNDDQKNN
 901 PGERANYVSR PYRIKYQTKI LTSGKDRVAN TASIEGLDNR VVYTKYDKDK
 951 EVTHTSGDAT ITGYNVDFDI LKRDAIDKKV MRDVHFDLYR LADGKWIPYI
1001 QDVSTGVNGR ISYKGLLTGR YKLVETRTQS GYTKLDTPIY FILSKKNMSD
1051 VKDKESSITL TDADGNKINN PAYASIIAMN PKKNKPTTLQ VLNYPVGGML
1101 PQTGGPGRLL FEALGSLLIV VACALTEVLI WRRIRSSKGV
```

(b)

```
   1 MKRVLKLLFM IVAFMTAVFA GSGQASADST TGITQDIHIH ISLGLNLPSG
  51 TDVTKLRAPT FEIYDISDQF NEADDPKEFT AKFPLGGQSY AKNFIEKHSL
 101 KPLSRQTGNK VNSSIDFIVP GCDAYLIVQT DENGVIENAG NNGTFTLPFV
 151 FLMNDDFKID DQGLMWFQIK GKTSLVQRSA YFFKYGKNAG GELPLSDAKF
 201 VLYRLDGSIK LYCTNNGGFK ASASPLSDDE IAKFTSNSAG LVMYDRESLD
 251 SGTYYFSEVQ APKGYRITDE ARKIEVYIPK KLSDGVKVNG TALEELYDQK
 301 LSDGAVSAAK PRIYNYSIEN PPSNSGKTNT PGKPNTPGKP NTSGKPTTPN
 351 SVRKKGLWGM LPQTGEAKSI MALLGIGIIC LVVLVSVGRR NYKEEH
```

(c)

```
   1 MKNHKKLRNA LATLLLALPL ALQGAVGVKT AQAAETSTET ATVTLHKYVF
  51 DKSLPSDKID NSKSQDEINA WLTKNNAEAL DGVEFTAYDV TSEYADAYKT
 101 ATGDKNESPA DAAKTASAAV AKKADALQKT ATVVGKQTTA NGGLASFANL
 151 PLRDANGNYK AYLFAETDAP ANITQKAEPF VLAMPIYGAD GKTVQKSINI
 201 YPKNVKQSDK KTLNDNRSHH DFTAGEKINY SIETVVPWNI ANKKVYTITD
 251 NPSKGLIMDA DTIQIEGLAS NKYTVKKNAD NGFTITIPAA NLAAFAGKTL
 301 KTTVKGHLSI EDLTLIDTGI PNKATAKVDN EAHHEVKSEE VFTGGKKFVK
 351 VDGSNQSKTL AGAQFQLVIV KNGQVVKYAH GNEKDGYTFD TNNTNVATKT
 401 TGENGQFEFA GLKYSESLEA GESYAVKEVK APTGYDLLKD PVLFTVTKDS
 451 YKTVQAADGQ KISNTKKGGF LPSTGGMGIV LFIAAGVVVM AGAAGTMIVR
 501 RNRRENI
```

Supplementary Fig. S5. Mass spectrometric identification of purified proteins as LrpC, LrpB, and LrpA. Peptide mass fingerprinting (PMF) analysis was carried out with an AB SCIEX Triple-TOF 5600 mass spectrometer on a single protein band obtained by SDS-PAGE. In-gel digestion was performed by standard procedures. Excised gel slices were subjected to de-staining and several repeated cycles of dehydration and rehydration, first with a 2:1 mixture of acetonitrile and 50 mM ammonium bicarbonate (ABC) for 5 minutes and afterward with 25 mM ABC for 2 minutes. Protein samples were trypsinized (20 μg ml^{-1} trypsin in 25 mM ABC) overnight at 37°C for mass spectrometric analysis. Mass spectrometry output was analysed using ProteinPilot™ software, in which peptides were identified against the full-length sequences of the pilin subunits. Matched peptides for LrpC **(a)**, LrpB **(b)**, and LrpA **(c)** are indicated by green boldface font.

```
   1 MERNKIFKKL LCILGAVATV FAIVFAMGKF DGEKANAAVE LNQNNTVSFI
  51 DSVKIMVSHK DSTSEELKNN DVYILRDGDK ISIDYNFSFP DNADINPGDT
 101 YTIYVPRAFK AYNSCDGKLK NENGTEFGTW QMSGTFDNEY NGYPLVMTFN
 151 DNIKKIQGRK GTAHLESSLD IESFSEENRQ EIKIPLKDST TYNVEIKQHT
 201 DEKFGDVTKN GWFLGNDWQP KRAHWTIDFN KGLNDIENPV LKDNLYHADS
 251 QYSEHKIDYG SFNVYSVRLD SYGRVIEKSK EKVNPQEYEI VKSADGMSYE
 301 LHFKHKIKGA YRVEYETNIT NDSGNRETTK VGNKVSSYDG DKHLADASKE
 351 LWAHWWDRGS TIVKKGVQKD SNRVDWTIYY NLVGHKFGDS TILLDSLYRG
 401 TFDKNSFEIY EAASIDGNPF NQNENNIKIG RKLSSSEYSI AYATKYDREV
 451 ATIKIKNSDG KGYVIRYRSI LPDGLPQGTQ IKNRVKDGHD NEGEGSIWYD
 501 QVKEHVSKSH TNVTQNKIDW SVKFYEATSD NSFKTFKNLT LTDHFYTNRG
 551 NQLSLVGGLN AVKVYRSSGP GDWNTNILVD PGKYTIKEDN RSGEGGYKGF
 601 KITFKGETPA ALYEVKYQTT RDSSQESGNW AQFGDEQTSD SVPADGGISK
 651 NVGISKSNTG VHRTDGLLHW QVYVNADKLP MSNYMVKDTI TGDQTLIEST
 701 IKVKDLTDNK DVTSQVRREI RETTTTYNQD GKTHTSPAKL LIIHLPDTNH
 751 QYEISYGVKL GPVGYNSDKN HYTDYADLYQ GNDKKGSVTN DYWHWYSRLE
 801 KDGAVDENDN SLVNWTIDVN KAYVIYPNGA IKDTLDGKQI FVEDSVEVYR
 851 YEGSDQWHDN LSSKPLPKSA YNVKIVPYTT SDGRVTQQMI ITFNDDQKNN
 901 PGERANYVSR PYRIKYQTKI LTSGKDRVAN TASIEGLDNR VVYTKYDKDK
 951 EVTHTSGDAT ITGYNVDFDI LKRDAIDKKV MRDVHFDLYR LADGKWIPYI
1001 QDVSTGVNGR ISYKGLLTGR YKLVETRTQS GYTKLDTPIY FILSKKNMSD
1051 VKDKESSITL TDADGNKINN PAYASIIAMN PKKNKPTTLQ VLNYPVGGML
1101 GGPGRLL FEALGSLLIV VACALTEVLI WRRIRSSKGV
```

Supplementary Fig. S6. Mass spectrometric identification of LrpC crystal. A single LrpC crystal was rinsed three times in mother liquor and dissolved in SDS-PAGE sample buffer. A single protein band from the gel was excised and used for PMF as described in Supplementary Fig. S5. Peptides that match the amino acid sequence of LrpC are indicated by green boldface font.