

## SUPPLEMENTAL MATERIAL TO

### Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16

1. Supplemental Material and Methods, page 1
2. Legends to Movies S1-S7, page 5
3. Figures S1-S11, page 6

#### MATERIALS AND METHODS

##### Generation of expression plasmids

Construction of the pS-ZG16-IgG expression plasmid has been described previously (1). For the pcDNA3.1-ZG16 plasmid, total RNA was extracted from the cell-line LS-174T with RNeasy Mini Kit (Qiagen). With total RNA as template the entire open reading frame of ZG16 was reversed transcribed and amplified with SuperScript One-Step RT-PCR with Platinum Taq (Life Technologies) and inserted into the pcDNA3.1-V5/His-TOPO vector (Life Technologies) to produce the vector pcDNA3.1-ZG16-V5/His. The C-terminal tag was removed by introducing a stop codon using QuikChange Site-Directed Mutagenesis Kit (Stratagene). Positive clones were controlled by sequencing.

##### Purification of recombinant protein

Spent media from CHO-K1 cells transfected with pS-ZG16-IgG was collected, centrifuged and passed through a 0.22 µm filter to remove cellular debris. Filtered media was then diluted 1:1 with 20 mM Sodium Phosphate buffer, pH 7.0 and applied to a 5 ml Protein-G column (GE Healthcare). Bound protein was eluted with 0.1 M Glycine-HCl, pH 2.7.

Purification of untagged ZG16 was performed as follows. Spent media from CHO-S cells expressing ZG16 was collected. Media was diluted 1:1 in 50 mM HEPES, pH 8.0 and separated from impurities using ion-exchange chromatography on a ÄKTA Purifier system (GE Healthcare). ZG16 was loaded onto a MONO S 5/5 column equilibrated in 50 mM HEPES pH 8.0, and it was eluted (around 400 mM NaCl) with a linear gradient of 0–1,0 M NaCl in the same buffer.

ZG16 containing fractions were pooled, concentrated and buffer exchanged to PBS using Vivaspin 6, MWCO 10,000 (Sartorius) spin columns. Protein concentration was determined with BCA protein assay (Pierce).

##### Bacterial culture

*Escherichia coli*, *Enterococcus faecalis*, *Bacteroides fragilis* and *Bacillus subtilis* were routinely cultured from frozen glycerol stocks in LB or BHI at 37°C for 15 h. *Lactobacillus jensenii* was grown at 37°C for 15 h in MRS-broth (Merck). Cultures were then diluted 1:100 in fresh broth and grown to required cell density for each experiment.

### **In-silico docking**

The glycopyranoside ring in PDB structure 3VY6 in the canonical binding site together with ZG16 was superimposed to a terminal NAM residue of the bacterial cell wall model as predicted by Meroueh et al (2, 3). Figures were drawn with the program Pymol and the electrostatic potentials were calculate using the generalized Born approximation in vacuum.

### ***Bacillus subtilis* motility assay**

Motility agar plates were prepared using BHI growth medium containing 0.3% (w/v) agar. 50  $\mu$ L of PBS with 20  $\mu$ g BSA or 20  $\mu$ g rZG16 was added to the center of plate and allowed to diffuse into the agar. 2  $\mu$ L OD<sub>600</sub> 0.5 *Bacillus subtilis* was inoculated into the center and plates were then incubated for 6 h at 37°C. Surface associated bacteria were washed away with PBS and plates flooded and incubated for 30 min with 10  $\mu$ M Syto9 to visualize bacteria in the agar. After PBS wash stained cells within the agar were imaged using a confocal microscope (Zeiss LSM 700) using a 488 nm laser and a Plan-Apochromat 20x/0.8 M27 objective lens. Tile scans (20 mm x 0.64 mm) were acquired and exported into Velocity (Perkin Elmer) where fluorescence intensity data from each 1 mm section was extracted. Fluorescence data for each section was normalized to total fluorescence for the whole tile scan and normalized data was used to compare bacterial distribution in the agar between treatment groups.

### **Culture and identification of bacteria from spleen tissue**

Spleen tissues in 1 ml sterile PBS were homogenized by a sterilized 5 mm stainless steel bead for 40sec using a Fast-Prep 24 instrument (MP Biomedicals). 100  $\mu$ l homogenate was spread on BHIS agar plates and incubated at 37°C under anaerobic conditions for 5 days. Different colony morphotypes were isolated and grown anaerobically at 37°C in BHIS broth until cultures were turbid. Liquid culture samples were Gram-stained and observed under light microscopy to determine isolate cellular morphology and basic Gram characteristics. DNA was extracted from 200  $\mu$ l of culture using phenol/chloroform extraction and bead beating as described above. Whole bacterial 16S genes were amplified using reagents and primers as described above for LCN-PCRs with the only modification being an increase in the cycle number to 35. Amplicons were purified using a Nucleospin PCR Clean Up kit (Macherey Nagel) and were then subjected to DNA sequencing (Eurofins Genomics). 16S gene sequences were used to identify isolates using BLAST (NCBI).

### **PCR screen for *Staphylococcus* and *Enterococcus* DNA**

DNA extractions were screened for staphylococcal and enterococcal DNA by PCR using primer pairs amplifying fragments of the bacterial *tuf* gene as previously described (4, 5). For *Staphylococcus* the primers TStag422 (GGCCGTGTTGAACGTGGTCAAATCA) and TStag765 (ACCATTCAGTACCTTCTGGTAA) were used and for *Enterococcus* Ent1A (CAACGACAAACCATTTCATGATG; modified from (5) and Ent2 (AACTTCGTCACCAACGCGAAC). All PCRs were performed using in 25  $\mu$ l volumes using 0.2  $\mu$ M primers and 200 ng template DNA. Thermocycling conditions were: 1 cycle of 95°C for 5 min; 35 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 1.5 min; 1 cycle of 72°C for 10 min. PCR reactions were subsequently run on 1% agarose/TBE gels containing Gel Red (Biotium) and imaged using a Gel Doc EZ system (Bio-Rad).

### **Profiling of the 16S rRNA gene DNA sequencing**

The bacterial DNA present in caudal lymph nodes, spleen, mucus and stool samples of Zg16<sup>+/+</sup> (n=6) and Zg16<sup>-/-</sup> (n=6) mice was profiled by sequencing of the V4 region of the 16S rRNA gene on an Illumina MiSeq (Illumina RTA v1.17.28; MCS v2.5) using 515F and 806R primers designed for dual indexing (6) and the V2 kit (2x250 bp paired-end reads). Stool samples were amplified in duplicates whereas mucus samples were amplified in triplicates in reaction volumes of 25 µl containing 1x Five Prime Hot Master Mix (5 PRIME GmbH), 200 nM of each primer, 0.4 mg/ml BSA, 5% DMSO and 20 ng (stool samples) or 50 ng (mucus samples) genomic DNA. PCR was carried out under the following conditions: initial denaturation for 3 min at 94°C, followed by 25 cycles (stool samples) or 30 cycles (mucus samples) of denaturation for 45 s at 94°C, annealing for 60 s at 52°C and elongation for 90 s at 72°C, and a final elongation step for 10 min at 72°C. Replicates were combined, purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and quantified using the Quant-iT PicoGreen dsDNA kit (Invitrogen). Equal amounts of purified PCR products were pooled and the pooled PCR products were purified again using Ampure magnetic purification beads (Agencourt) to remove short amplification products.

Illumina paired-end reads were merged using PEAR (7), and quality filtered to remove reads that had at least one base with a q-score lower than 20 and that were shorter than 220 nucleotides or longer than 350 nucleotides. Quality filtered reads were analyzed with the software package QIIME (8) (version 1.9.1). Sequences were clustered into operational taxonomic units (OTUs) at a 97% identity threshold using an open-reference OTU picking approach with UCLUST (9) against the Greengenes reference database (10) (13.8 release). All sequences that failed to cluster when tested against the Greengenes database were used as input for picking OTUs de novo. Representative sequences for the OTUs were Greengenes reference sequences or cluster seeds, and were taxonomically assigned using the Greengenes taxonomy and the Ribosomal Database Project Classifier (11). Representative OTUs were aligned using PyNAST (12) and used to build a phylogenetic tree with FastTree (13). Chimeric sequences were identified with ChimeraSlayer (14) and excluded from all downstream analyses. Similarly, sequences that could not be aligned with PyNAST, singletons, OTUs identified in a blank extraction sample and very low abundant sequences (relative abundance <0.005%) were also excluded. Only samples containing more than 1,000 reads were included in downstream analysis, preventing reliable analysis of spleen samples. The computations were performed on resources at the Swedish National Infrastructure for Computing (SNIC) through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX).

To correct for differences in sequencing depth, the same amount of sequences was randomly sub-sampled for each group of samples (17,222 for stool samples and 1,848 for mucus samples). A bootstrap version of Mann-Whitney-U test was used to compare the genotype-dependent abundance of OTUs at different taxonomical levels; significant differences were identified after correction for false discovery rate. Abundances higher than 1% are displayed on the family level. Microbiota 16S rRNA gene sequencing results have been deposited to the ENA sequence read archive with accession number PRJEB15635 (<http://www.ebi.ac.uk/ena/data/view>).

## REFERENCES for Supplemental Material

1. Rodriguez-Pineiro AM, Bergstrom JH, Ermund A, Gustafsson JK, Schutte A *et al.* (2013) Studies of mucus in mouse stomach, small intestine, and colon. II. Gastrointestinal mucus proteome reveals Muc2 and Muc5ac accompanied by a set of core proteins. *Am. J. Physiol. Gastroint. Liver Physiol.* 305: G348-G356.
2. Kanagawa M, Liu Y, Hanashima S, Ikeda A, Chai W *et al.* (2014) Structural Basis for Multiple Sugar Recognition of Jacalin-related Human ZG16p Lectin. *J. Biol. Chem.* 289: 16954-16965.
3. Meroueh SO, Bencze KZ, Heseck D, Lee M, Fisher JF *et al.* (2006) Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc. Natl. Acad. Sci. U. S. A.* 103: 4404-4409.
4. Martineau F, Picard FJ, Ke D, Paradis S, Roy PH *et al.* (2001) Development of a PCR assay for identification of staphylococci at genus and species levels. *J Clin Microbiol.* 39: 2541-2547.
5. Ke D, Picard FJ, Martineau F, Menard C, Roy PH *et al.* (1999) Development of a PCR assay for rapid detection of enterococci. *J Clin Microbiol.* 37: 3497-3503.
6. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79: 5112-5120.
7. Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30: 614-620.
8. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335-336.
9. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460-2461.
10. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72: 5069-5072.
11. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73: 5261-5267.
12. Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL *et al.* (2010) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26: 266-267.
13. Price MN, Dehal PS, Arkin AP (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5: e9490.
14. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV *et al.* (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 21: 494-504.

## SUPPLEMENTAL MOVIES

**Movie S1: Bacteria in close proximity to the colonic epithelium in  $Zg16^{-/-}$  mice.** Colonic tissue (bottom of image) and bacteria were stained with Syto9 (red) and imaged over a 1.5 min period as described in Materials and Methods. Data representative of time series acquired in tissues from n=4 animals.

**Movie S2: Live imaging bacteria and fluorescent beads in the WT colonic mucus.** Colonic tissue (bottom of image) and bacteria were stained with Syto9 (red) and fluorescent beads (green) were applied to visualize the IM/PM interface as described in Materials and Methods. Time series were recorded over 1.5 min. Data representative of time series acquired in tissues from n=4 animals.

**Movie S3: Tracking beads and bacteria within the WT mucus.** Tracking data obtained from Movie S2 using Imaris software. Tracks show beads at the IM/PM interface (green) and non-motile (red) or motile (white) bacteria within the IM. Bacteria were categorized as described in Materials and Methods.

**Movie S4: Live imaging bacteria and fluorescent beads in the  $Zg16^{-/-}$  colonic mucus.** Colonic tissue (bottom of image) and bacteria were stained with Syto9 (red) and fluorescent beads (green) were applied to visualize the IM/PM interface as described in Materials and Methods. Time series were recorded over 1.5 min. Data representative of time series acquired in tissues from n=4 animals.

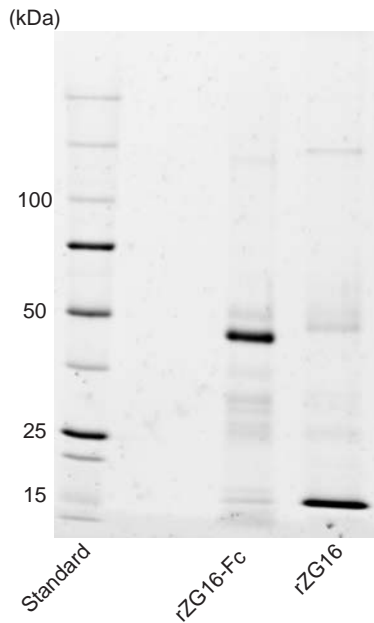
**Movie S5: Tracking beads and bacteria within the  $Zg16^{-/-}$  mucus.** Tracking data obtained from Movie S4 using Imaris software. Tracks show beads at the IM/PM interface (green) and non-motile (red) or motile (white) bacteria within the IM. Bacteria were categorized as described in Materials and Methods.

**Movie S6: Live imaging bacteria and fluorescent beads in rZG16 treated  $Zg16^{-/-}$  colonic mucus.** Colonic tissue (bottom of image) and bacteria were stained with Syto9 (red) and fluorescent beads (green) were applied to visualize the IM/PM interface as described in Materials and Methods. Time series were recorded over 1.5 min. Data representative of time series acquired in tissues from n=4 animals.

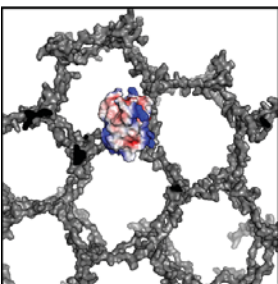
**Movie S7: Tracking beads and bacteria within the rZG16 treated  $Zg16^{-/-}$  mucus.** Tracking data obtained from Movie S6 using Imaris software. Tracks show beads at the IM/PM interface (green) and non-motile (red) or motile (white) bacteria within the IM. Bacteria were categorized as described in Materials and Methods.

FIGURE S1

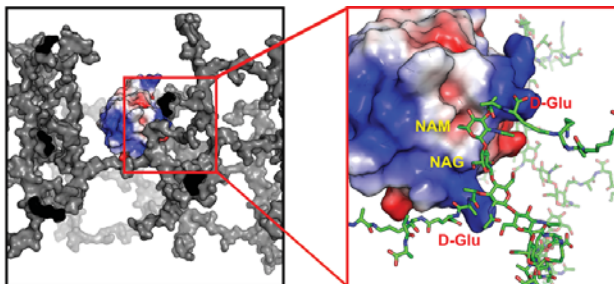
A



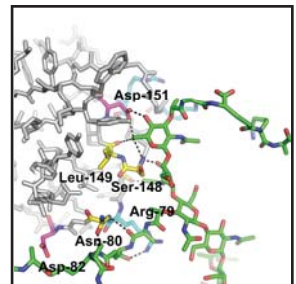
B



C

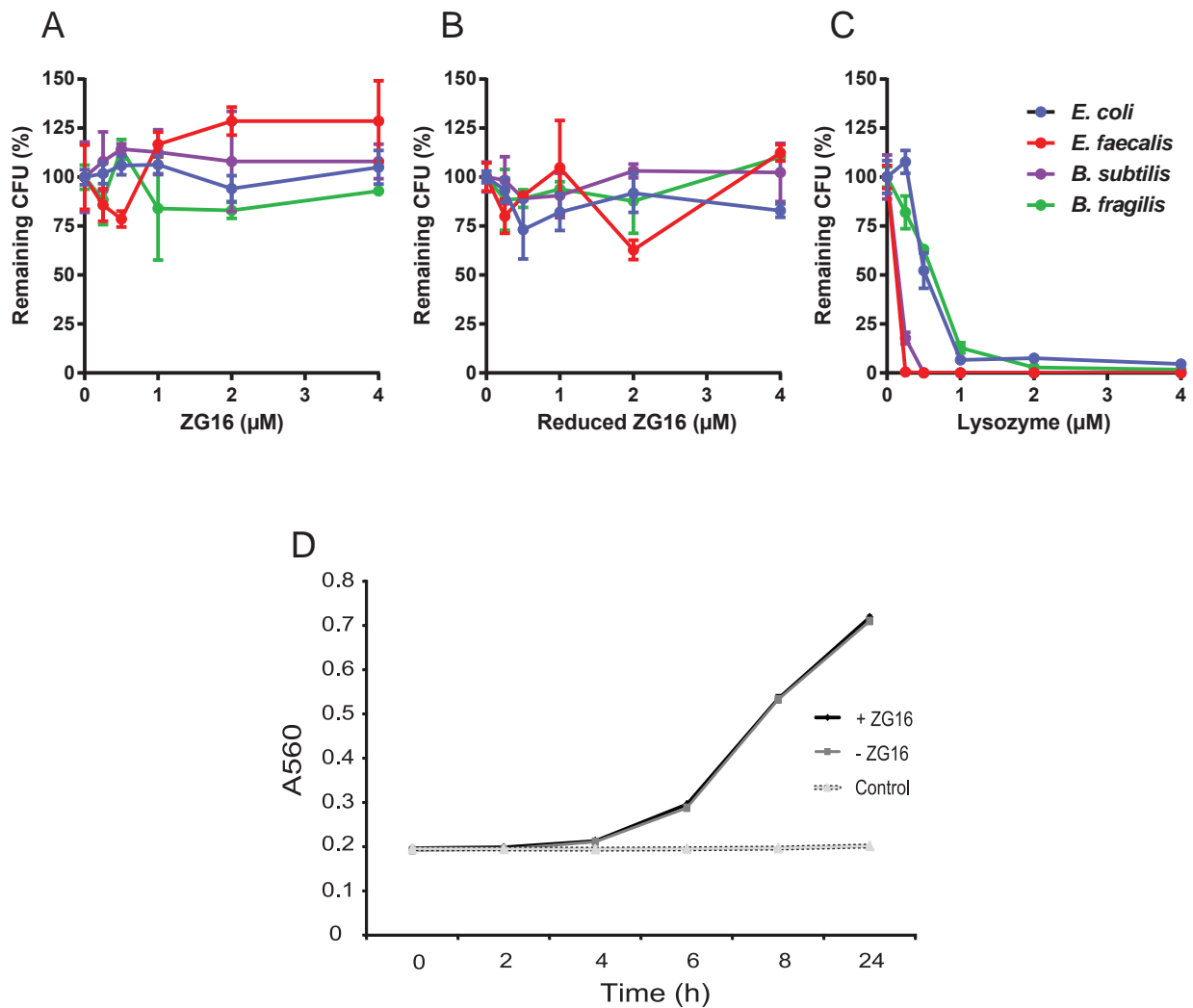


D



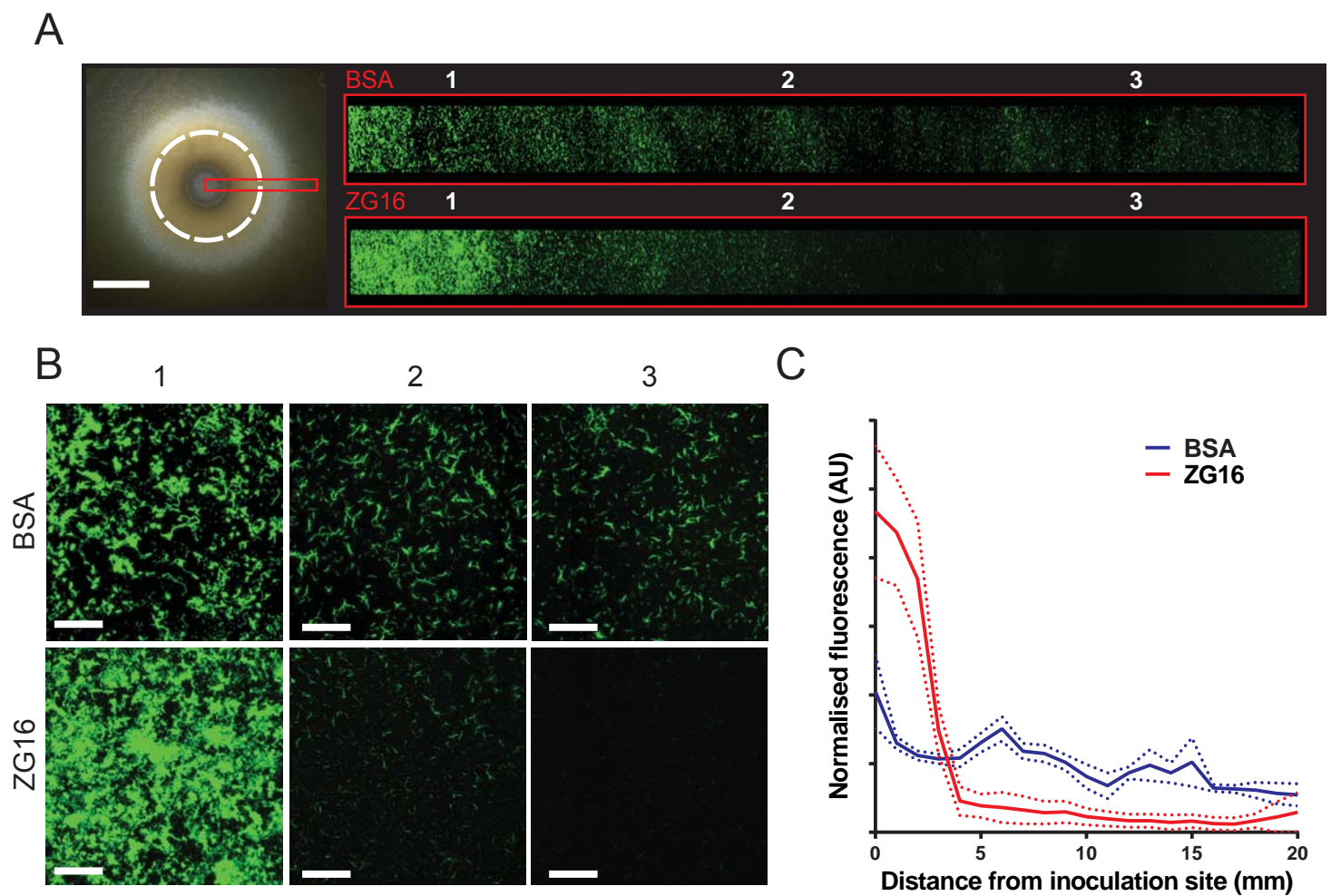
**Figure S1. Gel electrophoresis of purified rZG16 and rZG16-Fc and *in-silico* model of binding ZG16 to peptidoglycan.** **A.** The secreted recombinant ZG16-Fc and recombinant untagged ZG16 was purified from culture supernatants. The samples were analyzed by SDS-polyacrylamide gel electrophoreses on a 15% gel stained with SYPRO-Ruby. **B.** ZG16 bound to the bacterial cell wall peptidoglycan (grey) along the normal of the bacterial surface with the peptidoglycan forming a hexagonal pattern. Each node represents linear chains of two alternating amino sugars N-Acetylmuraminic acid (NAM) and N-Acetylglucosamine (NAG) and the bridge structures connecting the nodes consist of amino acid residues. **C.** the bound ZG16 is shown with the peptidoglycan in parallel with the bacterial wall and enlarged in **D.** The binding of the terminal NAM unit to the canonical sugar binding site is assisted by Coulombic interactions with the carboxyl group of D-Glu of the peptidoglycan and positively charged patches of ZG16. Additional ZG16 surface residues are likely to stabilize the peptidoglycan-ZG16 interaction. Canonical interactions are thus formed with Ser-148, Leu-149 and Asp-151 of ZG16. Arg-79 is primarily responsible for the positive electrostatic potential on the proximal side of the canonical binding site and forms a hydrogen bond to D-glutamate. Asn-80 and Asp-82 can provide further stabilizing hydrogen bonds with the connecting tetrapeptide. The hydroxyl group at C4 of the terminal NAM points toward Asp-151 and it is unlikely that a glycosidic bond is present at this position without distorting the linear arrangement of the sugar chain and potentially changing the cell wall structure.

FIGURE S2



**Figure S2. ZG16 has no bactericidal effect. A.-D.** *E. coli*, *E. faecalis*, *B. subtilis* and *B. fragilis* were incubated with different concentrations of rZG16 (A), reduced rZG16 (B) or Lysozyme (C) and remaining CFU determined by viable cell quantification. D. Growth of *Lactobacillus jensenii* was monitored for 24 h by OD<sup>560</sup> measurements in the presence or absence of recombinant rZG16.

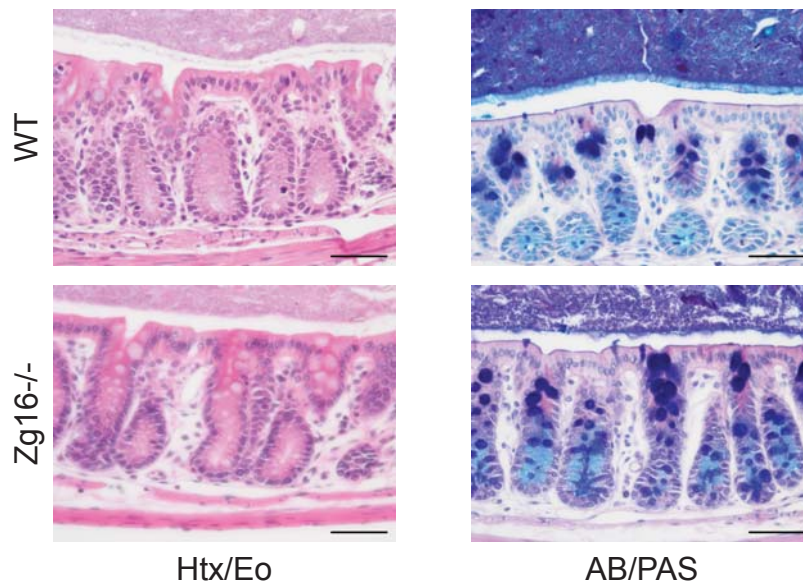
FIGURE S3



**Figure S3. ZG16 aggregates bacteria and limits spread within a hydrogel. A.** Growth of *B. subtilis* inoculated into the centre of low density agar treated with BSA or rZG16; images show representative bacterial growth on the agar (left panel) and 20 mm confocal micrographs of Syto9 stained bacteria within the agar (right panels); dashed white line represents the approximate BSA/rZG16 application area; red square shows area where confocal micrographs were acquired; scale bar 1 cm. **B.** Magnified images of Syto9 stained *B. subtilis* within low density agar from regions 1, 2 and 3 indicated on confocal micrographs shown in **A**; scale bars 100  $\mu$ m. **C.** Normalized fluorescence of *B. subtilis* within BSA/rZG16 treated low density agar at different distances from the initial point of bacterial inoculation; dashed lines represent SEM from n=3 independent experiments.

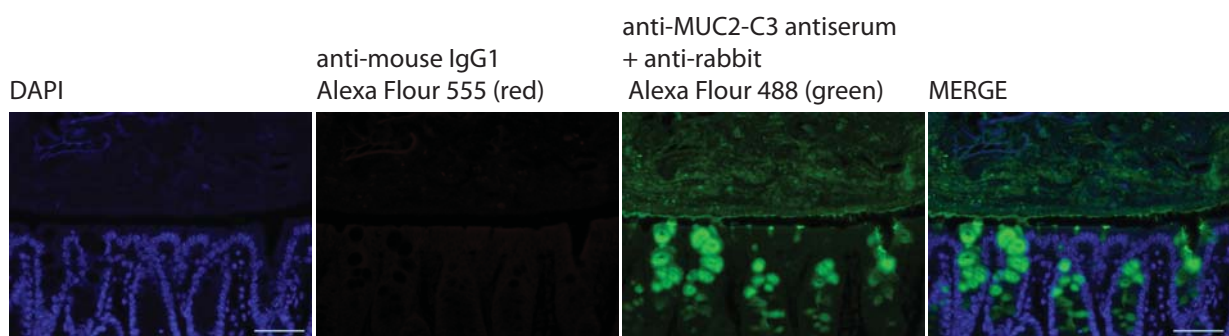


FIGURE S4



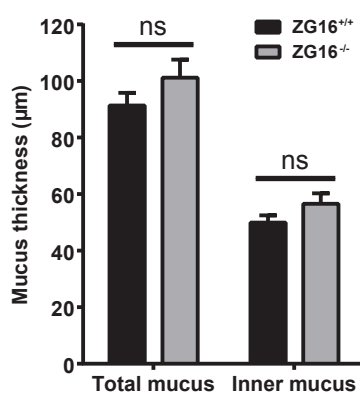
**Figure S4. Histology from distal colon sections of WT and Zg16<sup>-/-</sup> mice after Hematoxylin-Eosin (Htx/Eo) and Alcian blue/PAS (AB/PAS) staining. Scale bars 50  $\mu$ m.**

## FIGURE S5



**Figure S5. Negative control for LTA immunohistochemistry of the distal colon (Fig. 2a).** Carnoy fixed distal colon were stained with DAPI, anti-mouse IgG<sub>1</sub> Alexa Flour 555 (red), and anti-MUC2-C3 antiserum followed by anti-rabbit Alexa Flour 488 (green). Scale bar = 50  $\mu$ m. The secondary antibody for the anti-LTA staining in Fig. 2A did not show any background.

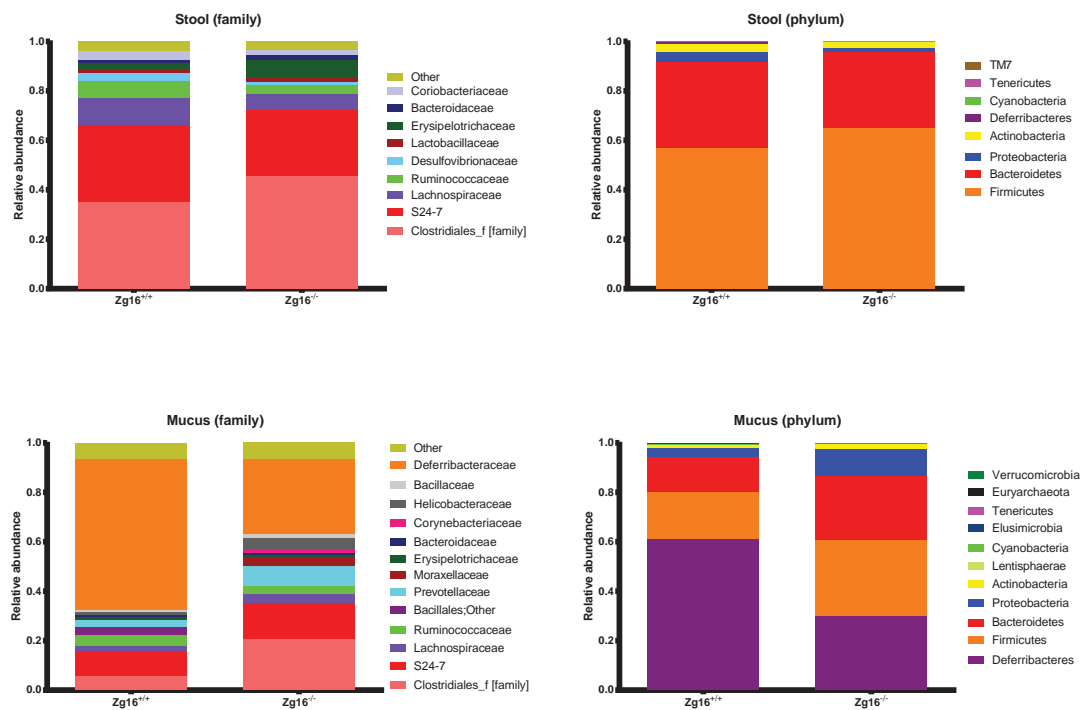
FIGURE S6



**Figure S6. Total mucus and inner mucus thickness of littermate Zg16<sup>+/+</sup> and Zg16<sup>-/-</sup>.**

Colon tissue without any fecal content was collected and either not flushed (total mucus) or flushed (inner mucus). Tissue was opened and the mucus visualized using 10 µm black polystyrene beads. Mucus thickness was measured using a glass micropipette. Error bars are SEM of measurements from n=10 animals per group. The calculated volume was used to estimate the number of bacteria per microliter mucus (Fig. 3a).

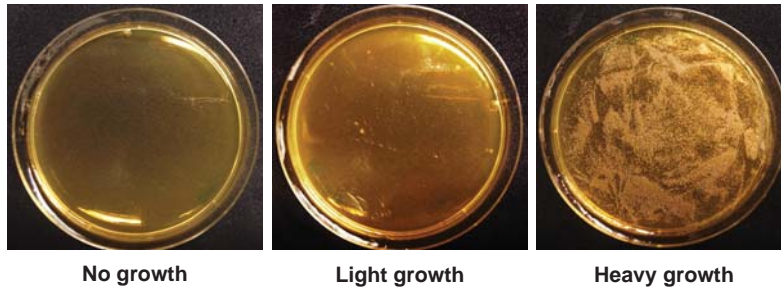
FIGURE S7



**Figure S7. Profiling of the 16S rRNA gene by DNA sequencing.** OTUs with a relative abundance less than 0.005% have been removed, only samples with more than 1000 reads included.

FIGURE S8

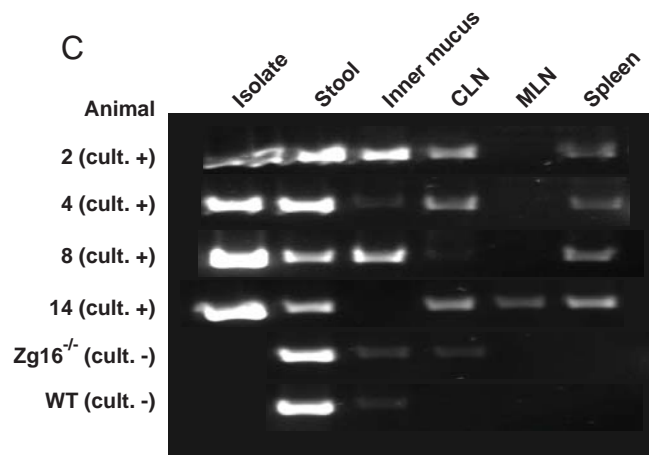
A



B

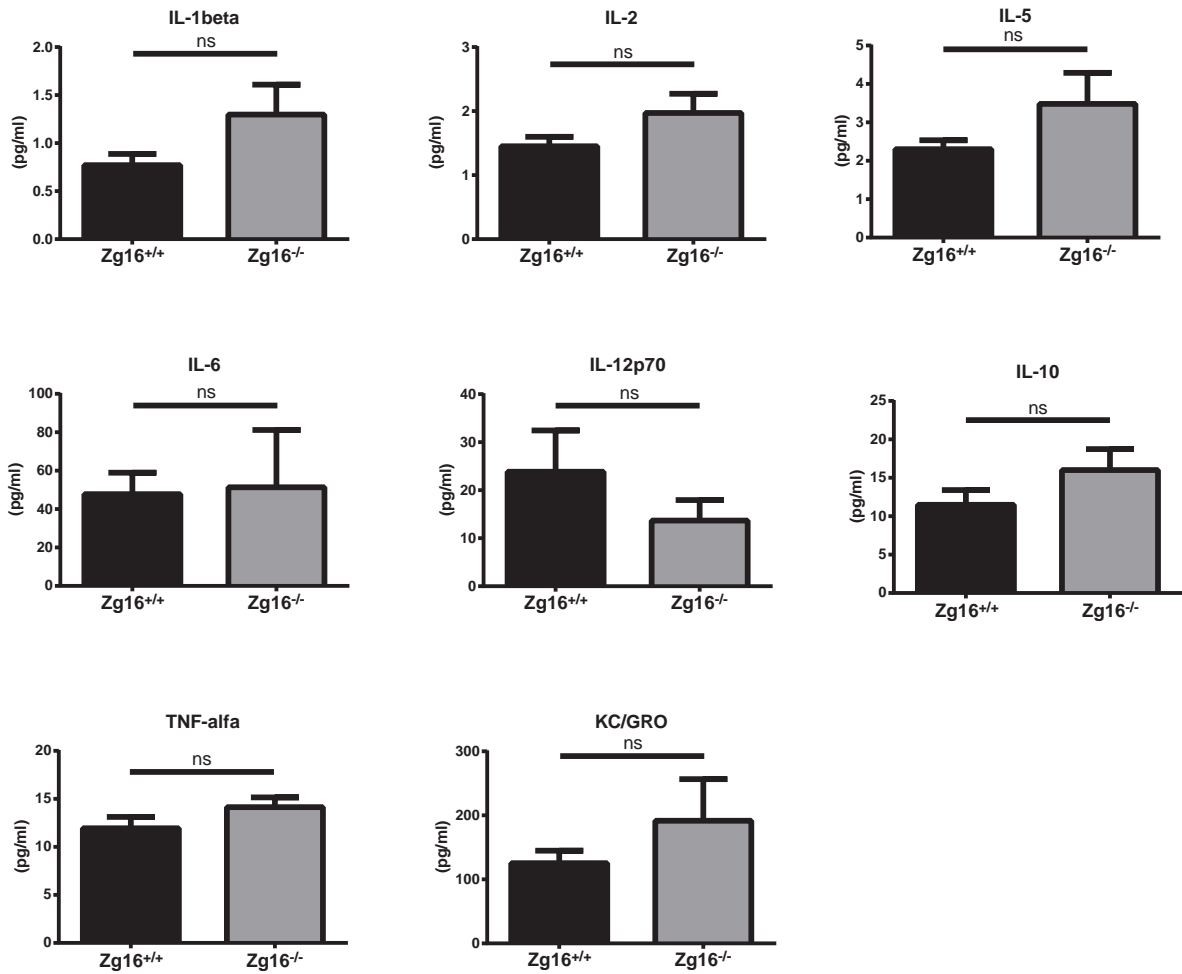
Animal	Genotype	rBHIS growth	Identity match
1	Zg16 <sup>-/-</sup>	No	
2	Zg16 <sup>-/-</sup>	Light	Staphylococcus epidermis
3	Zg16 <sup>-/-</sup>	No	
4	Zg16 <sup>-/-</sup>	Light	Enterococcus faecium Staphylococcus spp
5	Zg16 <sup>-/-</sup>	Heavy	Enterococcus faecium
6	Zg16 <sup>-/-</sup>	No	
7	Zg16 <sup>-/-</sup>	No	
8	Zg16 <sup>-/-</sup>	Light	Enterococcus faecium Staphylococcus spp
9	Zg16 <sup>-/-</sup>	No	
10	Zg16 <sup>-/-</sup>	Heavy	Enterococcus faecium
11	WT	No	
12	WT	No	
13	WT	No	
14	WT	Heavy	Staphylococcus spp
15	WT	No	
16	WT	No	
17	WT	No	
18	WT	No	
19	WT	No	

C



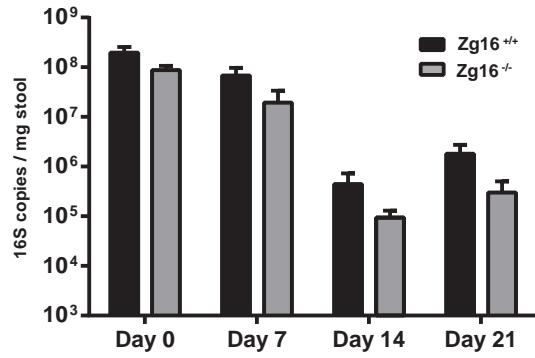
**Figure S8. Live bacteria from the spleen of littermate Zg16<sup>-/-</sup> and Zg16<sup>+/+</sup> were cultivated, their 16S rDNA sequenced and compared to fecal bacteria. A.** Spleen homogenates incubated on BHIS agar plates showing either no growth, light growth or heavy growth of bacteria after 5 days anaerobic incubation at 37°C. Plates were classified into three groups; no growth, light growth (<50 CFU), or heavy growth (>50 CFU). **B.** Table of spleen culture results of spleen BHIS cultures data and isolates identified by determined by sequencing the 16S rRNA gene. **C.** PCR screening for staphylococcal DNA in stool, mucus and systemic tissue DNA extractions from culture positive (cult. +) and negative (cult. -) samples by amplification of the bacterial *tuf* gene using *Staphylococcus*-specific primers.

FIGURE S9



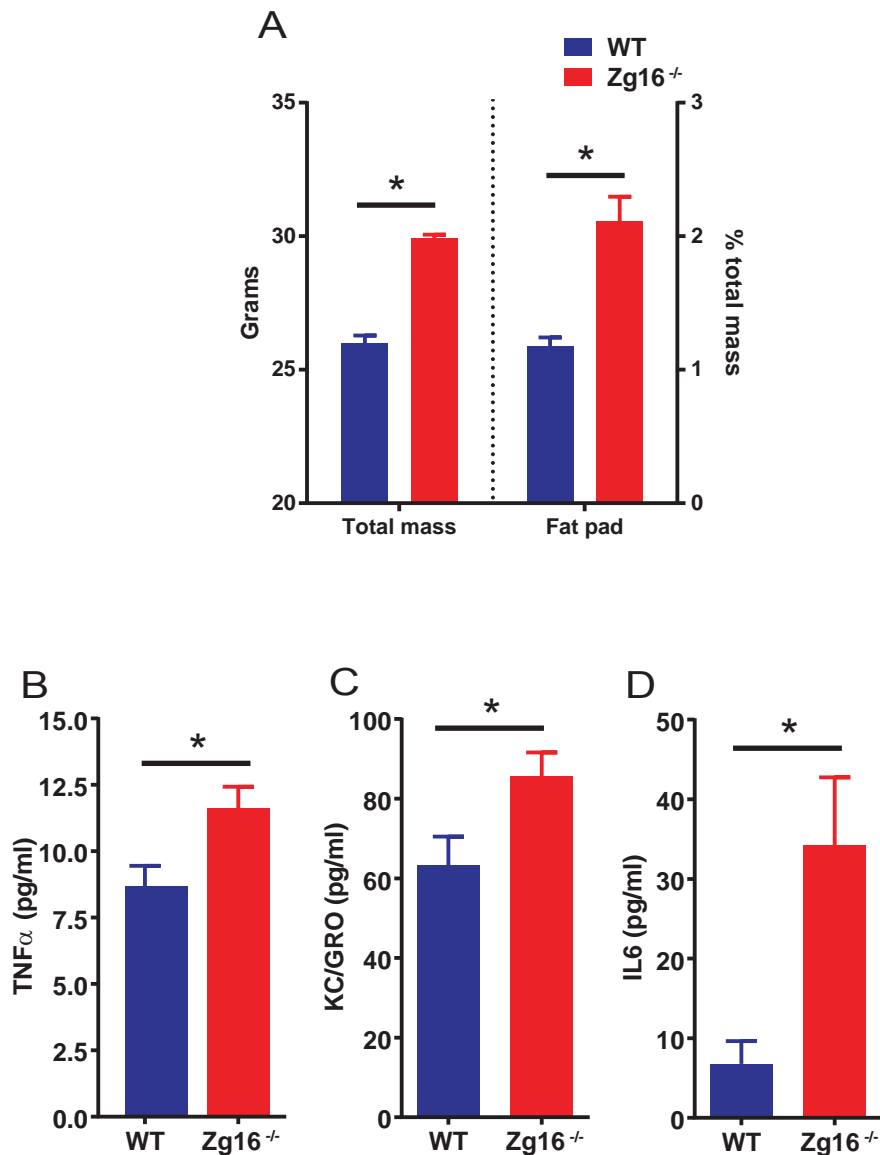
**Figure S9. Analysis of serum cytokines in the serum of littermate  $Zg16^{+/+}$  and  $Zg16^{-/-}$  mice.** Quantification of cytokines in the serum as described in Material and Methods. Error bars are SEM from  $Zg16^{+/+}$  (n=9) and  $Zg16^{-/-}$  (n=10) mice. Statistical significance calculated with Mann-Whitney U test (ns, not significant; \* $p < 0.05$ ).

FIGURE S10



**Figure S10. 16S quantification after antibiotic treatment.** Stool samples were collected and analyzed with qPCR in order to monitor the bacterial load during the time course of the experiment.

FIGURE S11



**Figure S11. Age matched colonies of male WT and Zg16<sup>-/-</sup> mice were analyzed for serum cytokines and compared for total body weight and fat pad mass.** The animals were kept in the same room and under the same conditions and males analyzed at 12 weeks age. The Zg16<sup>-/-</sup> colony had been separate from WT mice for at least two generations. **A.** Total body mass and fat pad mass was measured. Error bars are SEM from n=10 mice. **B-D.** Quantification of serum TNF $\alpha$ , KC/GRO and IL-6 by MSD; Error bars are SEM from n=4 mice per group. Statistical significance calculated with Mann-Whitney U test (C-D) or Sidak's multiple comparison test (A) (ns, not significant; \*p<0.05).