Supplemental material:

Supplemental figure 1: Fluorescence intensity in sputum samples. (A-C) Visualization of representative sputum samples from two CF patients and one CAP-COPD patient analysed with quantitative PNA-FISH based on fluorescence emitted from a universal bacteria specific PNA-FISH probe (16S rRNA). (A) CF sputum sample with non-growing or slow-growing cells based on the reduced numbers of intracellular ribosomal content (~3500 FU). (B) CF sputum sample with medium growing cells based on the numbers of intracellular ribosomal content (~8000 FU). (C) CAP-COPD sputum sample with fast growing cells based on the increased numbers of ribosomal content (~20500 FU). (D) Fluorescence intensity (FU) measured in each biofilm detected in images from sputum samples from CAP-COPD (n= 15), CF (n= 12) and CAP+COPD (n= 10). (E) Fluorescence intensity (FU) measured in all planktonic cells detected in images from sputum samples from CAP-COPD (n= 14), CF (n= 13) and CAP+COPD (n= 11). (F) Percent of total biomass in biofilms (red) and planktonic cells (black). (G) Proportion of biomass in biofilms (red) and planktonic cells (black). Statistical significance was determined using Kruskal-Wallis test ($P \leq 0.05$).

Supplemental figure 2: Brightfield microscopy images of bacterial biofilms found in sputum samples from CAP-COPD patients. Samples were stained with Alcian Blue polysaccharide stain to visualize polysaccharide within the biofilms as part of the matrix. In both images, a visible scaffolding of stringy polysaccharide can be observed within the biofilms. The mucus surrounding the biofilms are highly stained as well as it contains large amounts of polysaccharide as well. In both images, several eukaryotic nuclei can be observed surrounding the biofilms, although in image B to a much higher degree. (A) This image was obtained with a colour camera to display the blue staining. (B) This image was obtained using a mono-chrome camera which give high detains but in black and white display. 630x.

Supplemental figure 3: Test of fluoresces penetration in samples used for growth estimates based on PNA-FISH staining of rRNA. (A-C) three examples of the three investigated diagnosis. These three specimens were tested for fluorescent penetration. Results presented in D-F. D-F) Overlay graph of the mean FU measured in each of the 5 layers of the images (µm apart). The red bars represent the mean FU in the bacterial biomass. The blue bars represent the mean FU from the DAPI stained biomass. The gray bars represent the background fluorescence of the images. There was no significant slope gradient in any of the image. G-O) The mean FU histogram of nine randomly picked images, three from each investigated diagnose. The red bars represent the mean FU of the threshold bacterial biomass each layer of each images. The y-axis represents the depth in slides of the images.

Supplemental methods:

Diagnostic procedure of collecting expectorated material from the lower respiratory tract of patients with or without pneumonia: Sputum samples were obtained for microbiological analysis and were processed according to standardized routine practices at the local microbiological departments. In brief, each sputum sample was analysed at either the Department of Clinical Microbiology at Herlev University Hospital (CAP) or Rigshospitalet (CF). Microbiological analysis included Gram-stain and microscopy to determine the quality of the sample. Respiratory pathogens were identified at Herlev Hospital using conventional culturing in air with 5% carbon dioxide (CO_2) at 35ºC for 24 hours on 5% blood agar and chocolate agar plates (SSI Diagnostica, Denmark or Herlev Hospital). Samples were further analysed using culturing in ambient air at 35ºC for 24 hours on blue agar plates with a modified Conradi-Drigalski diagnostic substrate selective for Gram-negative bacteria (SSI Diagnostica). Cultured species were identified using matrix-assisted laser desorption or ionization time-of-flight mass spectrometry (MALDI-TOF MS) (Bruker Daltonics, Billerica, MA, USA). Atypical bacteria (i.e., *Mycoplasma pneumoniae*, *Chlamydophila pneumoniae*, *Chlamydophila psittaci* and *Legionella pneumophila*) were identified using real-time polymerase chain reaction (PCR) (RotorGene Q and Qiasymphony, Qiagen, DSP virus or pathogen

mini kit, version 1 [Qiagen, kat.nr. 937036]), and respiratory viruses were also identified using PCR analysis (Influenza A, Influenza B and respiratory syncytial virus [RSV]) (Panther Fusion Assay, Hologic, Flu A/B/RSV Assay Cartridge). Extended culturing on both CF and CAP samples was performed at Rigshospitalet and used as a tool for research because it was relevant for CF patients, though perhaps excessive for CAP patients for routine diagnostics. The extended culturing included a Sabouraud plate for fungal growth, a 7% NaCl plate and a *Burkholderia cepacia* plate that contained colistin and gentamicin. The latter was used to identify bacteria with different susceptibility patterns and different colony morphologies. All cultures were assessed daily. All agar plates with positive cultures were processed according to routine practice. At the end of each day, one of two experienced laboratory technicians assessed the plates. Two colonies were stored in tubes with Mueller-Hinton broth plus 10% (vol/vol) glycerol at -80ºC (see Supplemental Table 5).

Object identification

To identify objects, neighbouring voxels that shared at least one edge in the foreground were considered to be the same object. Objects ranging in size from 0.5–5.0 μm^3 were classified as planktonic cells. Objects larger than 5.0 μm^3 aggregates were classified as aggregates, and objects smaller than 0.5 μm^3 were excluded from analysis. Total biomass was calculated by adding the volume of observations.

Quantitative PNA-FISH

The average fluorescence emitted by bacteria in either biofilms or planktonic cells was quantified as described with modifications using the freeware programme ImageJ (National Institutes of Health, USA). The foreground and background signals were defined by a threshold value, using the automated MultiThresholder macro for ImageJ to create a binary data set (K. Baler, G. Landini and W. Rasband, NIH, USA). Aggregates or regions with solely single cells were defined prior to measurement using the region of interest tool. The ImageJ function ‘analyse particles’ was used for quantification. The fluorescence intensity was calculated in fluorescence intensity units (FU) as the mean of grey-scale units over a range of 0–65,535 FU**.** The average number of FUs emitted by the

PNA-FISH-treated cells was quantified for each sputum sample. We have previously validated the use of PNA FISH to estimate in vivo growth rates by demonstrating a significant ($P < 0.0001$) linear correlation between FU values and rRNA: number of rRNA molecules per rRNA gene molecule = 0.0447 x FU + 46.3 ($R^2 = 0.722$) [26].

To investigate if the depth of biomass affects the fluorescence units (FU) in samples stained with PNA-FISH, we quantified the mean FU of each layer within a biomass threshold in each image. Biomass foreground was defined with Otsu's method [36], and mean FU was reported with the use of FIJI [37]. The measurements were done on nine randomly picked of the samples sectioned to a thickness of 30 µm, three CAP, three COPD and three CF (see Supplemental Figure 3).

Image analysis

Images were analysed using a custom analysis pipeline (Matlab, Simulink) to distinguish between foreground and background. First, top and bottom boundaries were established for each sample by identifying terminal confocal sections that contained fluorescent voxels. The histograms of all slices on each image stack were combined and were then used to calculate a threshold for each channel using Otsu's method [36], resulting in the final threshold for the blue channel. Otsu-binarized images were then used to obtain final thresholds for the red channel. Foreground voxels on both channels were identified. The red channel intensity was collected from the raw images, and the minimum intensity of the top 1% of that population was used as the threshold for the red channel. Voxels in the final red images were subtracted from the blue channel to produce the final blue images.

Examination of sputum samples

To assess whether bacteria were present as planktonic cells or biofilm, sputum samples from all 43 patients were visualized using peptide nucleic acid fluorescent in situ hybridization (PNA-FISH). PNA-FISH uses fluorescently labelled probes specific to 16S ribosomal ribonucleic acid (rRNA) sequences to identify microbes and quantify the intracellular rRNA content. In this case, a universal bacterial PNA-FISH probe specific to 16S rRNA was used to visualize all bacteria, and samples

were counterstained with the blue-fluorescent DNA stain 4',6-diamidino-2-phenylindole (DAPI) to visualize eukaryotic cells. After staining, three-dimensional (3D) micrographs were obtained using confocal laser scanning microscopy. One of the key challenges of this study was developing methodology to identify fluorescently labelled bacteria in the presence of the high intrinsic background fluorescence in the sputum samples. This was accomplished using a custom image analysis pipeline to establish a background threshold fluorescence for each image.

Proportional occupancy

To determine proportional occupancy, binarized image stacks were analysed using a custom pipeline developed in the programming language R. Briefly, a focal voxel in the 3D image was selected at random, and the voxels of a specific channel located within a distance interval (radius 1 and radius 2) of the focal voxel were counted. This analysis revealed a proportional occupancy of inflammatory cells to bacterial cells, which quantitatively assessed the location of inflammatory cells in relation to bacterial cells compared to a null model in which the inflammatory cells were randomly positioned. Proportional occupancy was then calculated by multiplying the number of voxels within the distance range by the size of each voxel, then dividing that number by the total volume of the distance interval.

$$Proportional\,Occupancy = \frac{number\ of\ voxels \in distance\ interval \times voxel\ volume}{total\ volume\ of\ interval}$$

Proportional occupancy was obtained for a representative amount of focus voxels per image, starting from a distance of 1.0 μm away from each focal voxel and increasing by 1.0 μm distance intervals up to 25.0 μm away from focus voxel. Proportional occupancy was calculated using bacterial cells as a focal point and inflammatory cells as the surrounding environment, and it was also calculated using inflammatory cells as a focal point and bacterial cells as the surrounding environment.

Calculating community fluorescence intensity in aggregates and planktonic cells

The fluorescence intensity of each voxel—3D version of pixel—representing bacterial cells was collected to capture the distribution fluorescence intensity across the bacterial community on each

sample. Each voxel was classified according to the size of the object they belong as either an aggregate or a planktonic cell and their intensity was normalized to range from zero to one. Bacterial voxels from planktonic cells and aggregates were then separately sorted based on normalized intensity and the minimum intensity at each population quartile was obtained. This analysis results in the identification of the range of intensities of specific percentages of bacterial communities, classified as aggregates or planktonic cells.

Additionally, the percentage of voxels at the maximum normalized intensity for aggregates and planktonic cells was calculated for each sample. The p-values for all pair-wise comparisons were calculated using unpaired Wilcoxon test.

Matrix staining: Alcian blue, lectins and eDNA

Sputum specimens measuring 4.0 µm thick were stained with Alcian blue (Merck, USA), according to the protocol by[38] used by Survarna et al. for extracellular polysaccharide for biofilm matrix components. The stained specimens were then evaluated using bright field microscopy on an Axio Imager.Z2 with Axiocam 305 camera (Zeiss, Germany). A pale blue stain surrounding dense aggregation of bacteria indicated the presence of polysaccharide (see Supplemental Figure 2).

Histopathology

For histopathological examination, sputum samples were immediately fixed in phosphate-buffered saline (PBS) with 4% paraformaldehyde and were maintained at 5 °C prior to further preparation. The sputum samples were then embedded in paraffin. Samples were cut at 4.0 µm thickness, mounted in glass slides and stained with hematoxylin and eosin for the histological assessment.

Based on an overview at low magnification (62·5×) and detailed studies of five different areas of the sputum sample under high magnification (500×), the degree of inflammation was scored as 0 (no inflammation), 1 (mild inflammation), 2 (moderate inflammation) or 3 (severe inflammation with necrosis or severe inflammation throughout the sputum sample) [39]. The

histopathological assessment was completed blindly by a board-certified pathologist. Histopathological assessment was performed using a Laica DM 4000 B LED microscope.