Supporting Information

In vivo proteome of *Pseudomonas aeruginosa* in airways of cystic fibrosis patients

Xia Wu[1], Richard J Siehnel[2], Jayanthi Garudathri[2], Benjamin J. Staudinger[3], Katherine B. Hisert[3], Egon A. Ozer[5], Alan R. Hauser[5], Jimmy K. Eng[1], Colin Manoil[1], Pradeep K. Singh[2,3], James E. Bruce[1,4,*]

[1]Department of Genome Sciences, University of Washington, Seattle, WA, USA.
[2]Department of Microbiology, University of Washington, Seattle, WA, USA.
[3]Department of Medicine, University of Washington, Seattle, WA, USA.
[4]Department of Chemistry, University of Washington, Seattle, WA, USA.
[5]Department of Microbiology-Immunology, and Department of Medicine, Northwestern University, Chicago, Illinois, USA.

[*]To whom correspondence should be addressed: James E. Bruce, Department of Genome Sciences, University of Washington, 850 Republican Street, Seattle, WA 98109. Tel: (206)543-0220; Fax: (206) 616-0008; E-mail: jimbruce@u.washington.edu

SUPPORTING INFORMATION TABLE OF CONTENTS

The following supporting information is available free of charge at ACS website http://pubs.acs.org

SUPPLEMENTAL MATERIALS AND METHODS

Immunoblot analysis and antibody information

Proteins were extracted from CF sputum samples with 4x Laemmli protein sample buffer (Bio-Rad, Hercules, CA, USA). Extraction mixtures were incubated at 95 °C with a heating block for 5 min. The extraction mixtures were further vortexed for 30 s, and were sonicated with a GE-130 ultrasonic processor using settings of 50% amplitude, 30 s duration for two cycles. Protein concentration was estimated with Bradford assay. As SDS could interfere with the accuracy of Bradford assay protein quantitation, two rounds of SDS-PAGE were analyzed. At the first round of SDS-PAGE, SDS was removed through three washes of SDS-PAGE gel with H_2O, and gels were stained with coomassie blue (Bio-Rad). The densitometry of each lane was obtained with Odyssey Infrared Imaging System (LI-COR Bioscience). The densitometry information was used for further adjustment of equal total protein loading. At the second round of SDS-PAGE, equal total proteins were loaded based on the updated information, and gels were transferred to Immobilon-FL PVDF for immunoblot analysis (1).

Antibodies for human protein anti-GAPDH were purchased from Genscript (catalog information, A00191) (Piscataway, NJ, USA). Antibodies for bacterial protein anti-TufA was obtained from Hycult Biotech Inc. (Plymouth Meeting, PA, USA) (catalog information, HM6010). Antibodies for bacterial protein anti-RpoA was purchased from Neoclone (Madison, WI, USA) (catalog information, SP002-W0003).

Proteins subsets with single PA peptide identification

First, there were 248 PA proteins that were identified with one peptide from a single CF sputum sample, but with two or more peptides across the entire 35 CF sputum sample set. While these identifications did not provide the same level of confidence as derived from 2 identified peptides from a single sputum sample, the two separate peptide IDs from the entire dataset (1) improve confidence that

the protein is expressed in patient airways and (2) provide independent opportunities that can be used for the targeted PRM methods for quantification of these proteins in sputum.

The second sub-group contains 82 PA proteins with single peptide ID's that have been repeatedly identified in 3 or more sputum samples and identified with at least two peptides from *ex vivo* samples. Of the 82 PA proteins identified in patient airway samples with only a single peptide, 44 of these same peptides were identified together with at least one other non-redundant peptide from the same protein in *ex vivo* PA samples, increasing the likelihood of the presence of these proteins in patient airway samples. The *in vitro* IDs indicate that the PA peptides can be identified without the host materials, meaning that the protein extraction and digestion method, LC-MS/MS system, and database search method were tuned to identify these peptides. The *in vivo* PA peptide IDs may be truly from the PA proteins rather than coming from the incorrect assignment of the host proteins/peptides.

All presented PA peptides are unique to PA against human proteins.

Accession number of PA genome

PA genome sequence deposited to National Center for Biotechnology Information (NCBI) are as follow.

Organism	Accession
PABL001	QVEI00000000
PABL002	QVEH00000000
PABL003	QVEG00000000
PABL004	QVEF00000000
PABL006	QVEE00000000
PABL007	QVED00000000
PABL009	QVEC00000000
PABL010	QVEB00000000
PABL011	QVEA00000000
PABL012	CP031659
PABL013	QVDZ00000000
PABL014	QVDY00000000
PABL015	QVDX00000000
PABL016	QVDW00000000
PABL017	CP031660
PABL018	QVDV00000000
PABL019	QVDU00000000
PABL020	QVDT00000000

PABL021	QVDS00000000
PABL022	QVDR00000000
PABL023	QVDQ00000000
PABL024	QVDP00000000
PABL026	QVDO00000000
PABL027	QVDN00000000
PABL028	QVDM00000000
PABL029	QVDL00000000
PABL030	QVDK00000000
PABL031	QVDJ00000000
PABL032	QVDI00000000
PABL034	QVDH00000000
PABL035	QVDG00000000
PABL036	QVDF00000000
PABL037	QVDE00000000
PABL038	QVDD00000000
PABL040	QVDC00000000
PABL041	QVDB00000000
PABL042	QVDA00000000
PABL043	QVCZ00000000
PABL044	QVCY00000000
PABL045	QVCX00000000
PABL046	QVCW00000000
PABL047	QVCV00000000
PABL048	QVCU00000000
PABL049	QVCT00000000
PABL051	QVCS00000000
PABL052	QVCR00000000
PABL053	QVCQ00000000
PABL054	QVCP00000000
PABL055	QVCO00000000
PABL056	QVCN00000000
PABL057	QVCM00000000
PABL058	QVCL00000000
PABL059	QVCK00000000
PABL060	QVCJ00000000
PABL061	QVCI00000000
PABL062	QVCH00000000
PABL063	QVCG00000000
PABL064	QVCF00000000
PABL065	QVCE00000000
PABL066	QVCD00000000
PABL067	QVCC00000000
PABL068	QVCB00000000
PABL069	QVCA00000000
PABL070	QVBZ00000000
PABL071	QVBY00000000
PABL072	QVBX00000000
PABL073	QVBW00000000
PABL074	QVBV00000000

PABL075	QVBU00000000
PABL076	QVBT00000000
PABL077	QVBS00000000
PABL078	QVBR00000000
PABL079	QVBQ00000000
PABL080	QVBP00000000
PABL081	QVBO00000000
PABL082	QVBN00000000
PABL083	QVBM00000000
PABL084	QVBL00000000
PABL085	QVBK00000000
PABL086	QVBJ00000000
PABL088	QVBI00000000
PABL089	QVBH00000000
PABL090	QVBG00000000
PABL091	QVBF00000000
PABL092	QVBE00000000
PABL093	QVBD00000000
PABL094	QVBC00000000
PABL095	QVBB00000000
PABL096	QVBA00000000
PABL097	QVAZ00000000
PABL098	QVAY00000000
PABL100	QVAX00000000
PABL101	QVAW00000000
PABL102	QVAV00000000
PABL103	QVAU00000000
PABL104	QVAT00000000
PABL105	QVAS00000000
PABL106	QVAR00000000
PABL107	QVAQ00000000
PABL108	QVAP00000000

SUPPLEMENTAL REFERENCES

1. Wu, X.; Chavez, J. D.; Schweppe, D. K.; Zheng, C.; Weisbrod, C. R.; Eng, J. K.; Murali, A.; Lee, S. A.; Ramage, E.; Gallagher, L. A.; Kulasekara, H. D.; Edrozo, M. E.; Kamischke, C. N.; Brittnacher, M. J.; Miller, S. I.; Singh, P. K.; Manoil, C.; Bruce, J. E., In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in Acinetobacter baumannii strain AB5075. *Nat Commun* **2016,** 7, 13414.

Figure S1

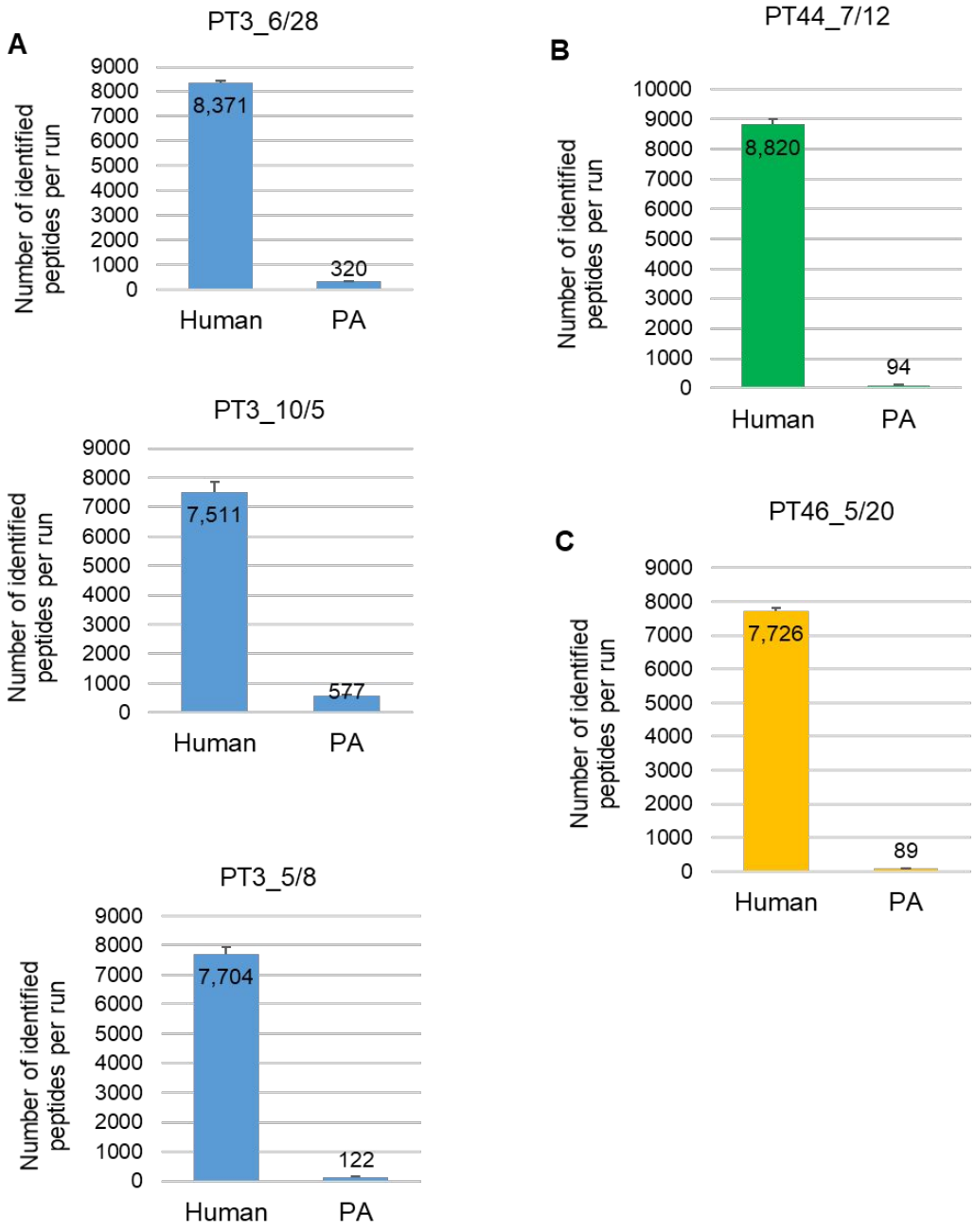


Figure S1. PA proteins are much less abundant compared with human proteins in CF sputum samples. The number of non-redundant human or PA peptides identified in each LC-MS/MS run is shown. Examples include five CF sputum samples collected from three CF patients (A, B, C). The low abundance of PA proteins in CF sputum samples highlights the necessity to design bacterial-enrichment methods to improve PA protein identification with CF sputum samples.

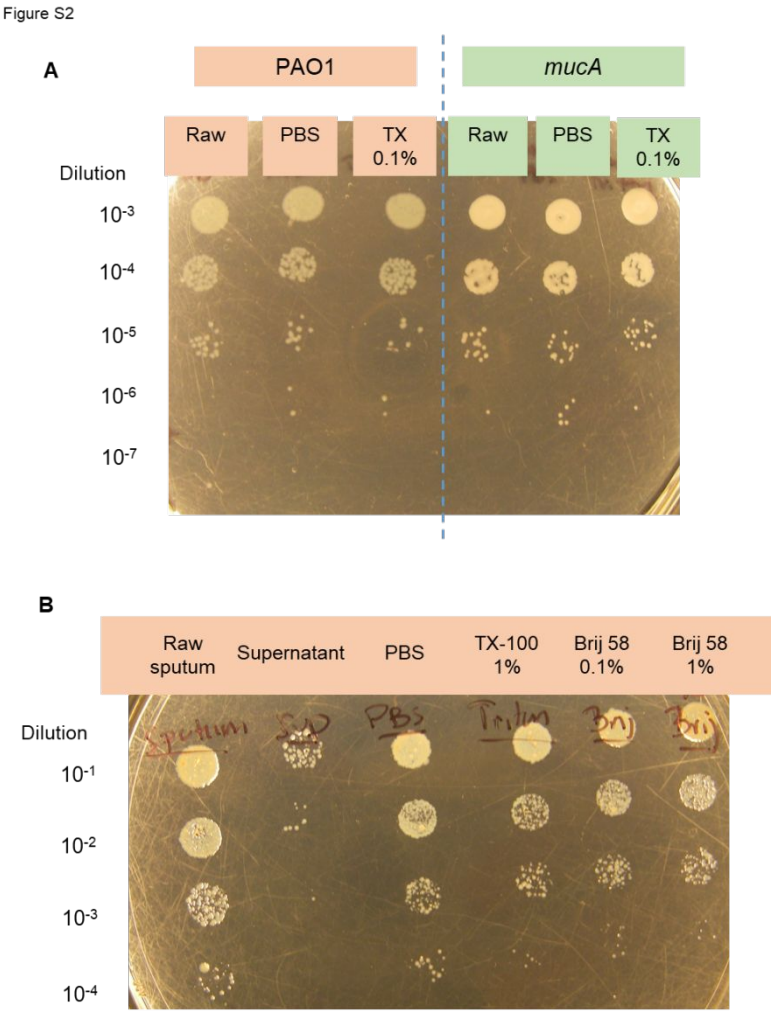
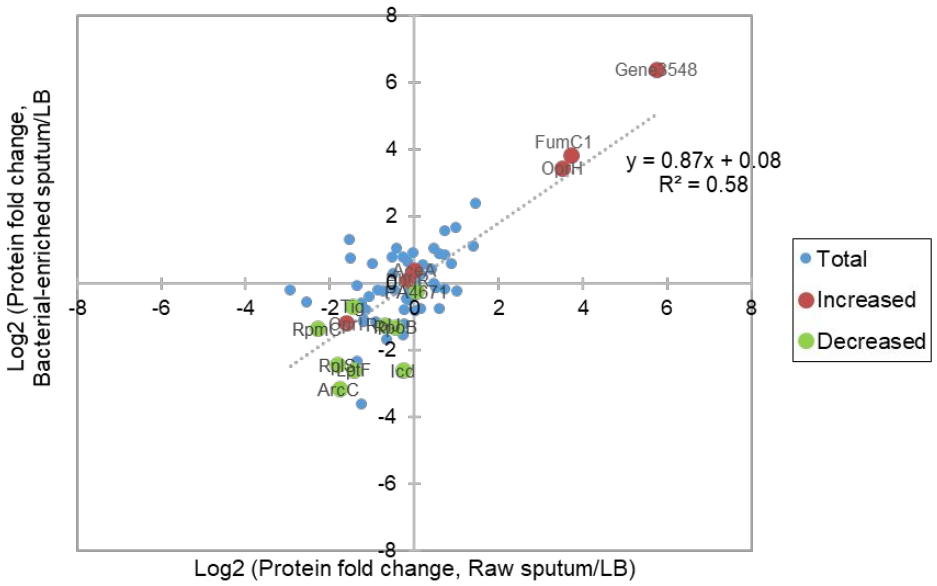


Figure S2. Treatment with 0.1% Triton X-100 or 0.1% Brij 58 did not lyse viable PA cells. A). *P. aeruginosa* PAO1 WT or *mucA* strain (mucoid phenotype, Mathee et al., 1999) of OD_{600}=1.0 were subjected to centrifugation (10, 000 × g, 5 min) and resuspension treatment (PBS or TX-100 0.1%). The "Raw" cells are control samples that did not undergo centrifugation and PBS/TX-100 treatment procedures. CFU analysis shows that PBS and TX-100 treatment did not cause lysis of PAO1 and *mucA* cells. B) CF sputum samples were homogenized with PBS and 0.1% dithiothreitol on ice for 30 min, and aliquotted to five tubes. The sputum pellet (after 10 min, 16,000 × g, 4ºC) was resuspended with PBS, 1% TX-100, 0.1% Brij 58 or 1% Brij 58. The resuspension mixtures were centrifuged again with 16,000 × g, and resuspended to the original volume with PBS for CFU analysis. The "Raw sputum" was the sample aliquot that was homogenized but was not subjected to centrifugation and detergent treatment analyses. The "Supernatant" was the supernatant fraction collected after the first found of 16,000 × g centrifugation. Results indicate that treatment of PBS, 1% TX-100, 0.1% or 1% Brij 58 did not lyse viable bacterial cells in CF sputum samples. The CFU in the supernatant fraction shows that approximately 1% (two log10 reduction) of bacterial cells remained in the supernatants after the centrifugation procedures.

Figure S3

A

Sputum sample 1



Sputum sample 2

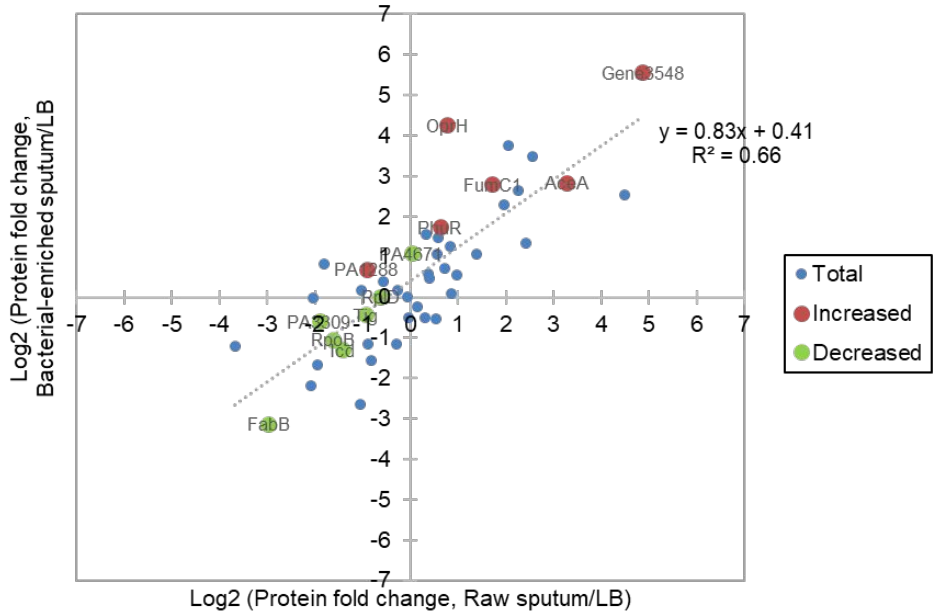


Figure S3A. Comparison of PA proteome before and after the bacterial-enrichment procedures. The CF sputum sample was aliquoted so that one aliquot of the sample was processed with bacterial-enrichment procedures, while another aliquot was kept on ice and processed as the raw sputum for proteome comparison (Fig. 2A). The PA proteomes of sputum samples were compared with the PA *ex vivo* proteomes on LB medium condition. Results showed that for PA proteins that could be quantified before and after the bacterial enrichment procedures, these proteins showed a general consistence in the relative abundance in PA proteome, suggesting that the bacterial-enrichment procedures used in this study did not significantly alter the proteome profiles of PA in CF sputum. The Increased/Decreased proteins that were identified through the analysis of 35 sputum samples are highlighted in figures.

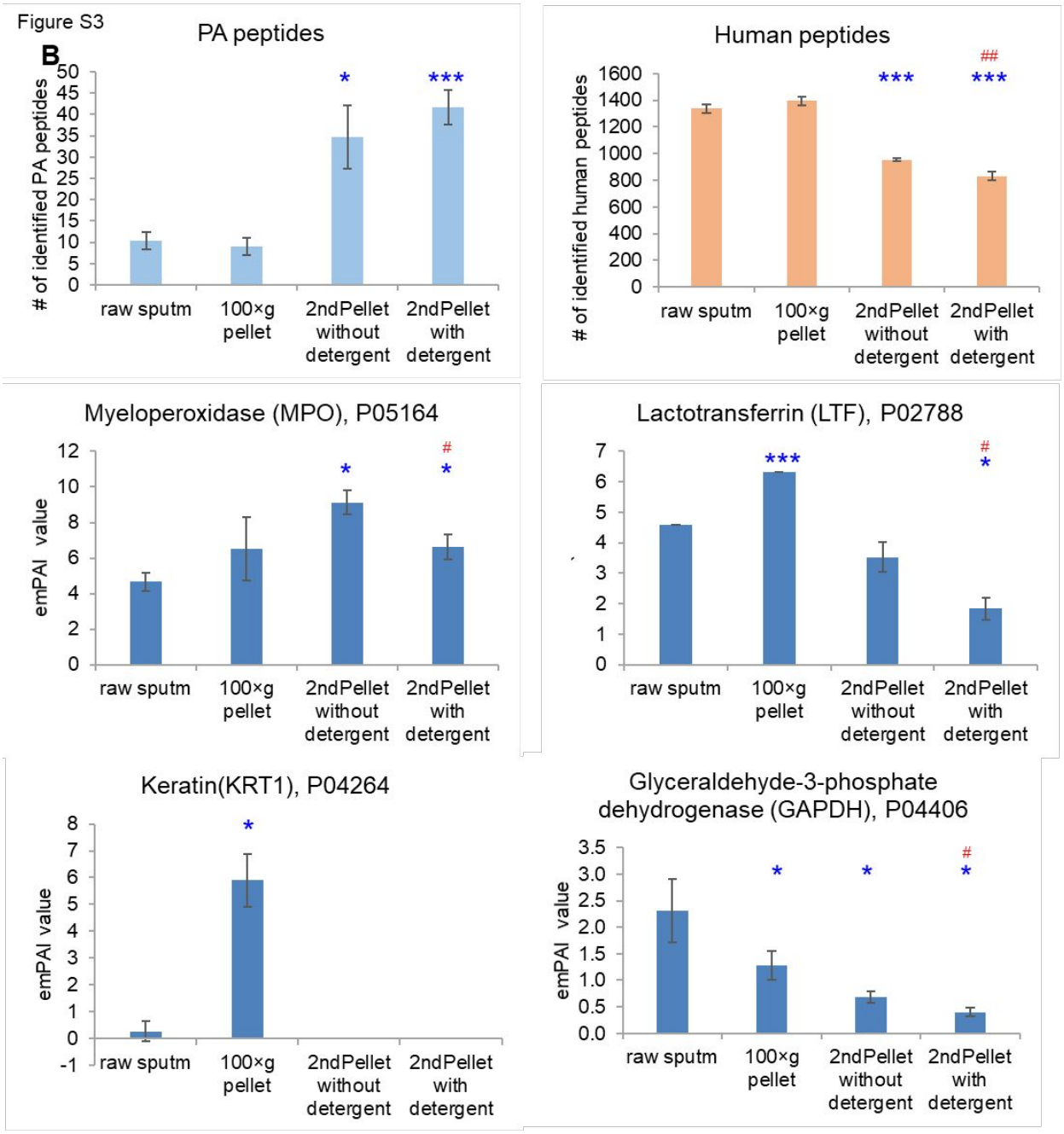


Figure S3B. Comparison of PA proteome before and after the bacterial-enrichment procedures. The CF sputum samples were aliquoted to prepare as raw sputum, the pellet collected with 100×g centrifugation (100×g pellet), the second centrifugation pellets after the 10,000×g centrifugation with or without additional TX-100 and Brij-58 treatment. Proteomics data were collected with LTQ-XL (Thermo-Fisher). Greater number of PA peptides were identified with the bacterial enriched pellets. Spectral counting analysis with Exponentially Modified Protein Abundance Index (emPAI) showed the benefits of reduction of abundant host proteins LTF, KRT1, MPO and GAPDH with the separation of 100×g pellet or with the detergent treatment. Standard error bars indicate standard deviation of three technical replicates of LC-MS/MS analysis. ANOVA analysis comparison with raw sputum *, $P < 0.05$, ***, $P < 0.001$. ANOVA analysis comparison between detergent treatments, # < 0.05, ## < 0.01.

Figure S4

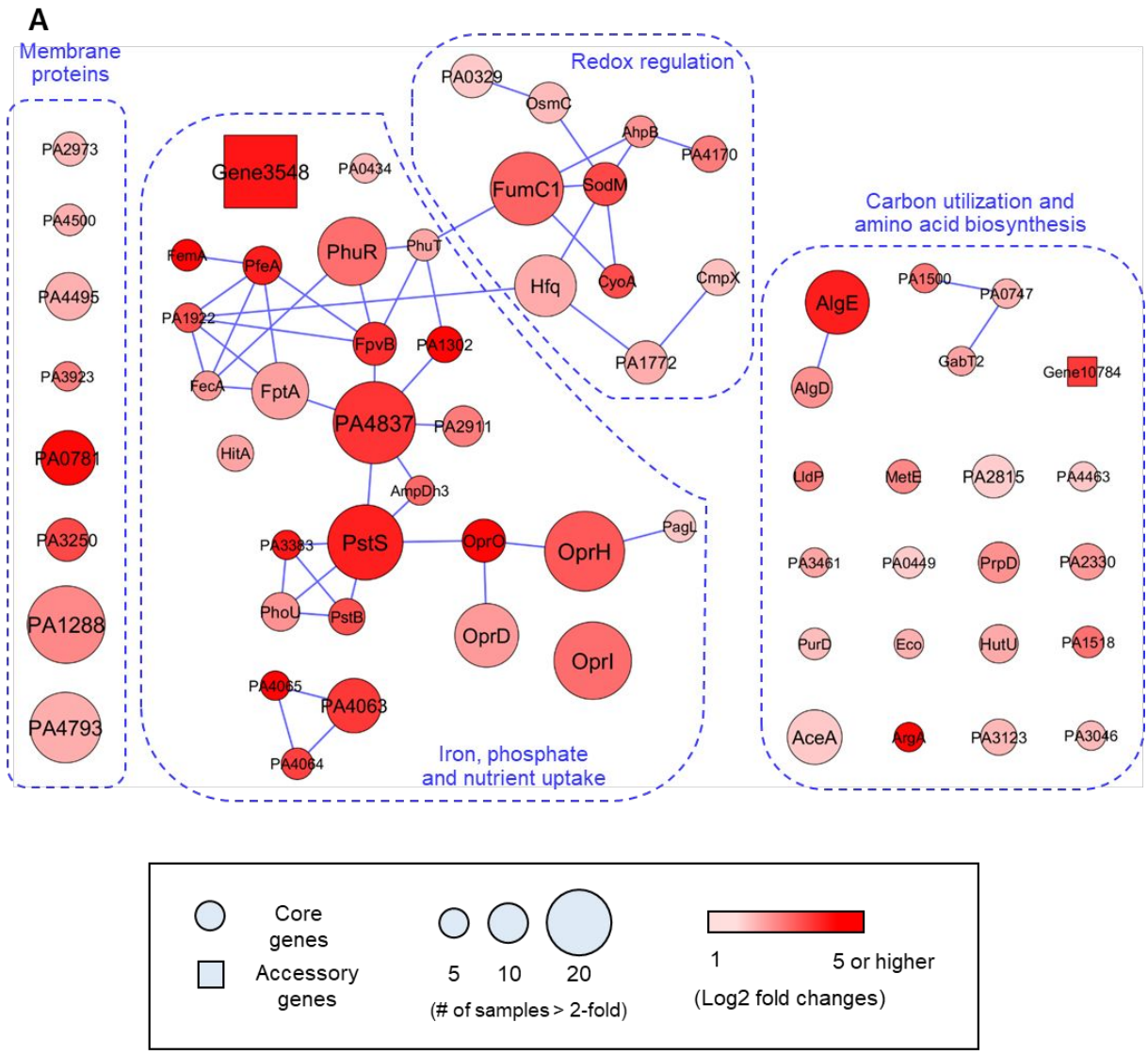


Figure S4. PA protein network up-regulated (A) or down-regulated (B) in CF airways. The network includes 67 increased proteins and 117 decreased proteins that were identified with 2-fold or greater fold-changes with at least 5 CF sputum samples compared with the *ex vivo* cultured samples. The nodes represent proteins, and the edges indicate protein interaction information which was obtained with STRING (version 10.0). The node shape indicates whether the PA protein is classified as a core gene or accessory gene. The node size is correlated with the number of sputum samples identified with 2-fold or greater changes. The bigger the node size means the increases were more frequently found in CF sputum samples. The node color indicates the extent of increases (red) or decreases (green) of the averaged protein fold change. The network is visualized with Cytoscape (version 3.4.0).

Figure S4

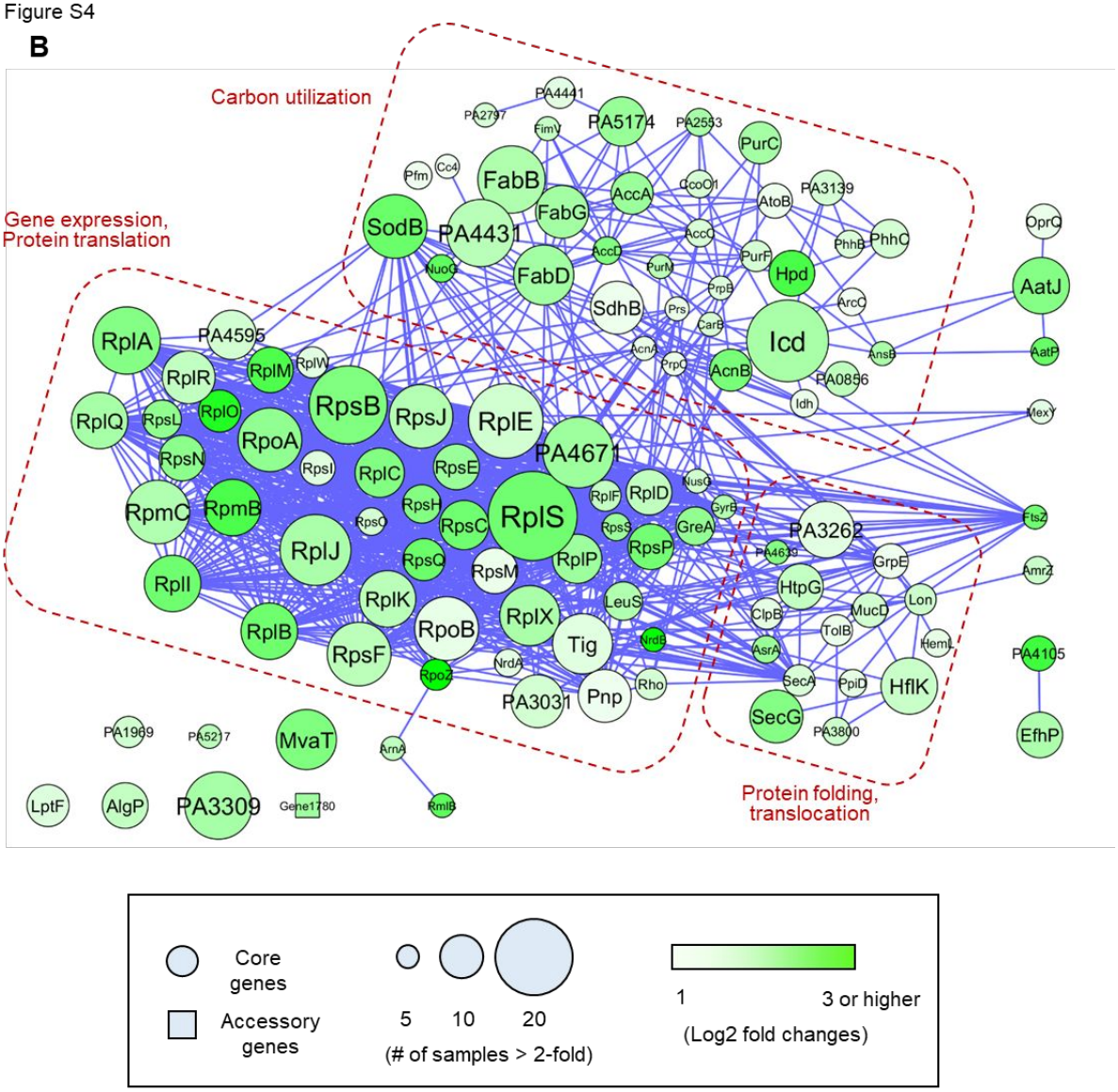


Figure S4. PA protein network up-regulated (A) or down-regulated (B) in CF airways. The network includes 67 increased proteins and 117 decreased proteins that were identified with 2-fold or greater fold-changes with at least 5 CF sputum samples compared with the *ex vivo* cultured samples. The nodes represent proteins, and the edges indicate protein interaction information which was obtained with STRING (version 10.0). The node shape indicates whether the PA protein is classified as a core gene or accessory gene. The node size is correlated with the number of sputum samples identified with 2-fold or greater changes. The bigger the node size means the increases were more frequently found in CF sputum samples. The node color indicates the extent of increases (red) or decreases (green) of the averaged protein fold change. The network is visualized with Cytoscape (version 3.4.0).

Figure S5

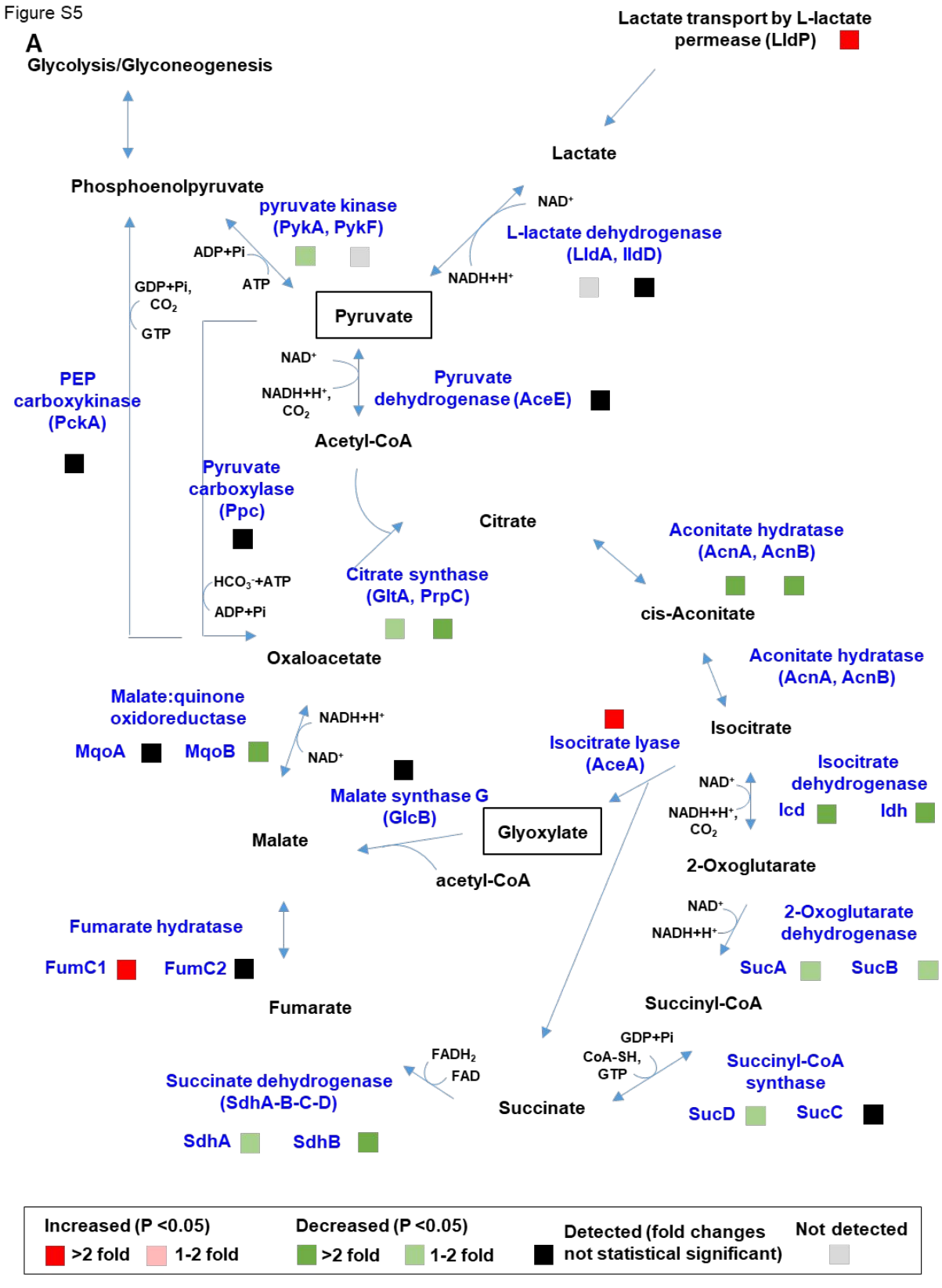

A
Glycolysis/Glyconeogenesis
Lactate transport by L-lactate permease (LldP)
Lactate
Phosphoenolpyruvate
pyruvate kinase (PykA, PykF)
ADP+Pi
ATP
NAD+
NADH+H+
L-lactate dehydrogenase (LldA, lldD)
GDP+Pi, CO2
GTP
Pyruvate
NAD+
NADH+H+, CO2
Pyruvate dehydrogenase (AceE)
PEP carboxykinase (PckA)
Acetyl-CoA
Pyruvate carboxylase (Ppc)
Citrate
Aconitate hydratase (AcnA, AcnB)
HCO3-+ATP
ADP+Pi
Citrate synthase (GltA, PrpC)
cis-Aconitate
Oxaloacetate
Aconitate hydratase (AcnA, AcnB)
Malate:quinone oxidoreductase
MqoA
MqoB
NADH+H+
NAD+
Isocitrate
Isocitrate lyase (AceA)
Isocitrate dehydrogenase
NAD+
NADH+H+, CO2
Icd
Idh
Malate synthase G (GlcB)
Malate
Glyoxylate
acetyl-CoA
2-Oxoglutarate
Fumarate hydratase
NAD+
NADH+H+
2-Oxoglutarate dehydrogenase
FumC1
FumC2
SucA
SucB
Fumarate
Succinyl-CoA
GDP+Pi
CoA-SH, GTP
FADH2
FAD
Succinyl-CoA synthase
Succinate dehydrogenase (SdhA-B-C-D)
Succinate
SucD
SucC
SdhA
SdhB
Increased (P <0.05)
>2 fold
1-2 fold
Decreased (P <0.05)
>2 fold
1-2 fold
Detected (fold changes not statistical significant)
Not detected

Figure S5

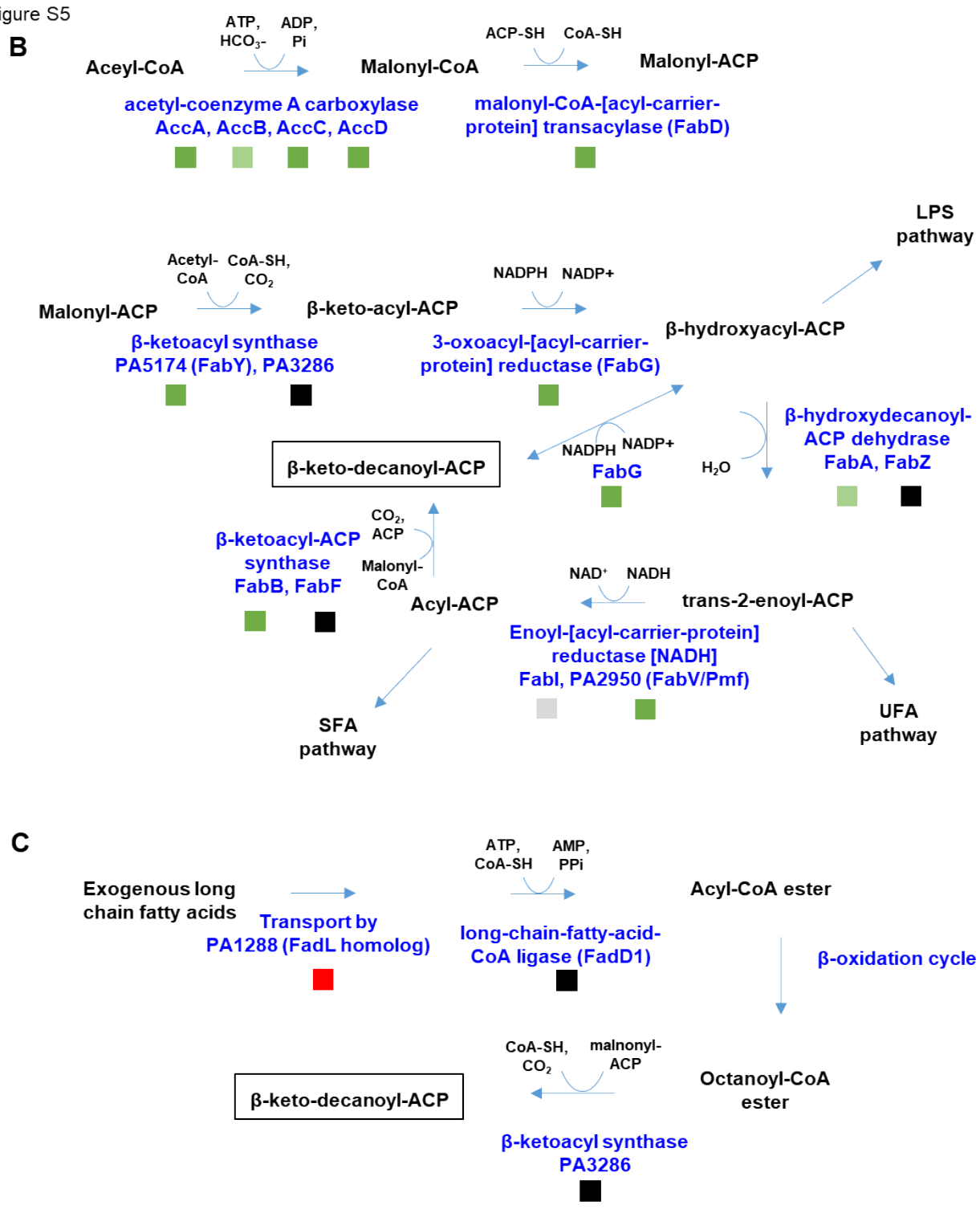

B
Aceyl-CoA
ATP, HCO3-
ADP, Pi
Malonyl-CoA
ACP-SH
CoA-SH
Malonyl-ACP
acetyl-coenzyme A carboxylase AccA, AccB, AccC, AccD
malonyl-CoA-[acyl-carrier-protein] transacylase (FabD)
LPS pathway
Malonyl-ACP
Acetyl-CoA
CoA-SH, CO2
β-keto-acyl-ACP
NADPH
NADP+
β-hydroxyacyl-ACP
β-ketoacyl synthase PA5174 (FabY), PA3286
3-oxoacyl-[acyl-carrier-protein] reductase (FabG)
β-keto-decanoyl-ACP
NADPH
NADP+
FabG
H2O
β-hydroxydecanoyl-ACP dehydrase FabA, FabZ
β-ketoacyl-ACP synthase FabB, FabF
CO2, ACP
Malonyl-CoA
Acyl-ACP
NAD+
NADH
trans-2-enoyl-ACP
Enoyl-[acyl-carrier-protein] reductase [NADH] FabI, PA2950 (FabV/Pmf)
SFA pathway
UFA pathway
C
Exogenous long chain fatty acids
Transport by PA1288 (FadL homolog)
ATP, CoA-SH
AMP, PPi
long-chain-fatty-acid-CoA ligase (FadD1)
Acyl-CoA ester
β-oxidation cycle
CoA-SH, CO2
malnonyl-ACP
Octanoyl-CoA ester
β-keto-decanoyl-ACP
β-ketoacyl synthase PA3286
Increased (P <0.05)
>2 fold
1-2 fold
Decreased (P <0.05)
>2 fold
1-2 fold
Detected (fold changes not statistical significant)
Not detected

Figure S5

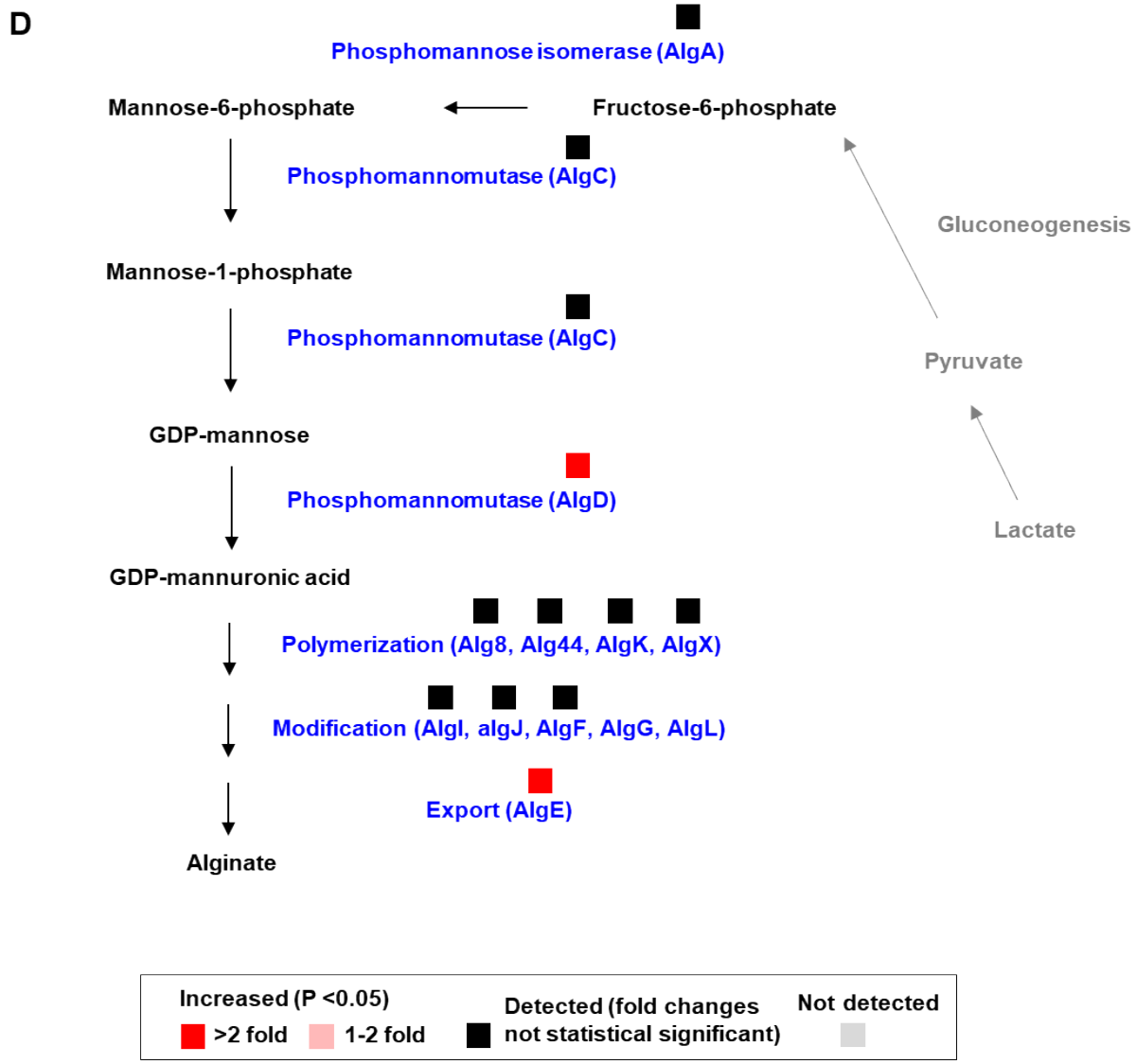


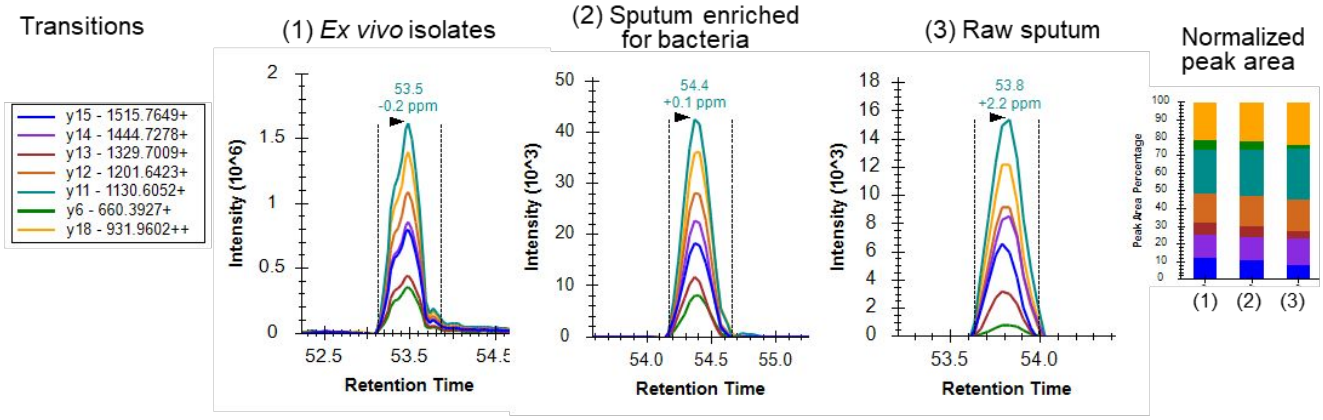
Figure S5. Carbon metabolic remodeling of PA *in vivo*. A) TCA cycle and glyoxylate cycle. Enzymes in lactate utilization and glyoxylate cycles are up-regulated *in vivo*, as well as fumarate hydratase FumC1. B-C) Fatty acid metabolic pathways of endogenous biosynthesis (B), and utilization of exogenous fatty acid product (C). The diagram is adapted from Yuan et al. (2012). Many enzymes in fatty acid endogenous biosynthesis is down regulated *in vivo*, but enzymes functioning in utilization of exogenous fatty acid are up-regulated. (D) Alginate biosynthesis pathway. For all panels, increased proteins are shown in red. Decreased proteins are shown in green, with the dark green as fold change > 2, and light green as fold change in between 1-2 fold. Markers in black indicate that the enzymes were detected, but fold changes were not statistical significant. Markers in grey indicate that the enzymes were not detected.

Figure S6

Control proteins

PA5554 AtpD

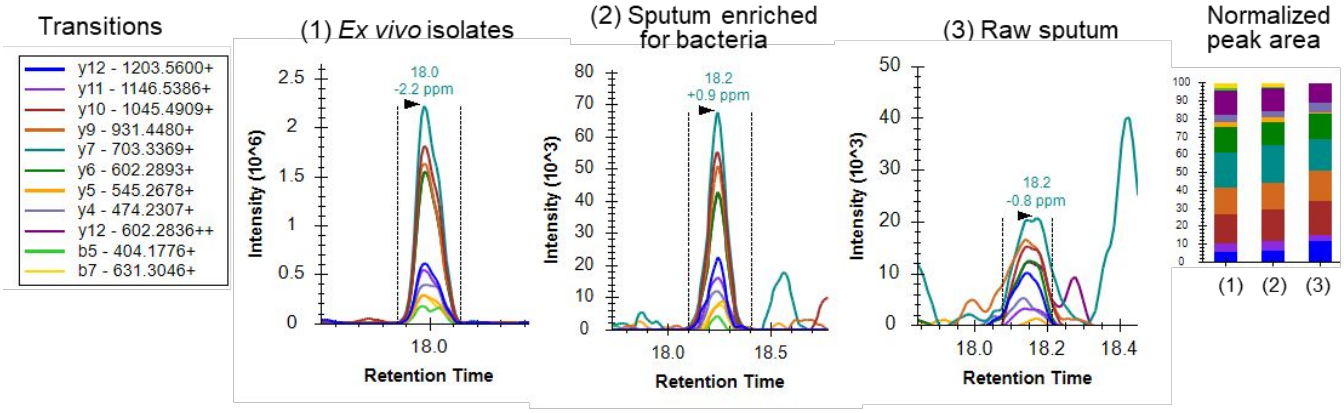
R.EAPSYADQAGGNELLETGIK.V [111, 130]



Porins

PA4710 PhuR

R.SGTGTNLDTGADSPR.D [399, 413]



GENE3548

K.EIPQTINVVTQDEIK.A [163, 177]

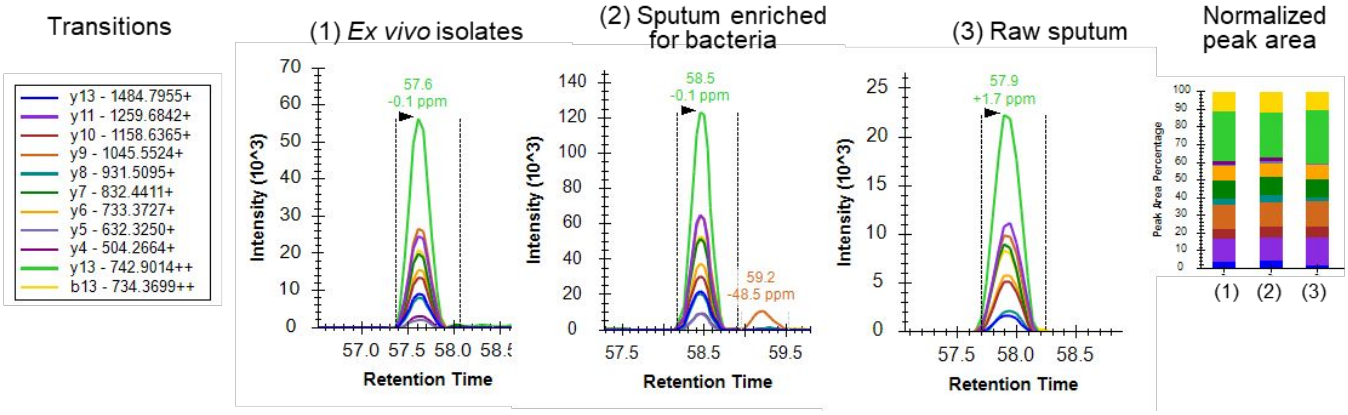
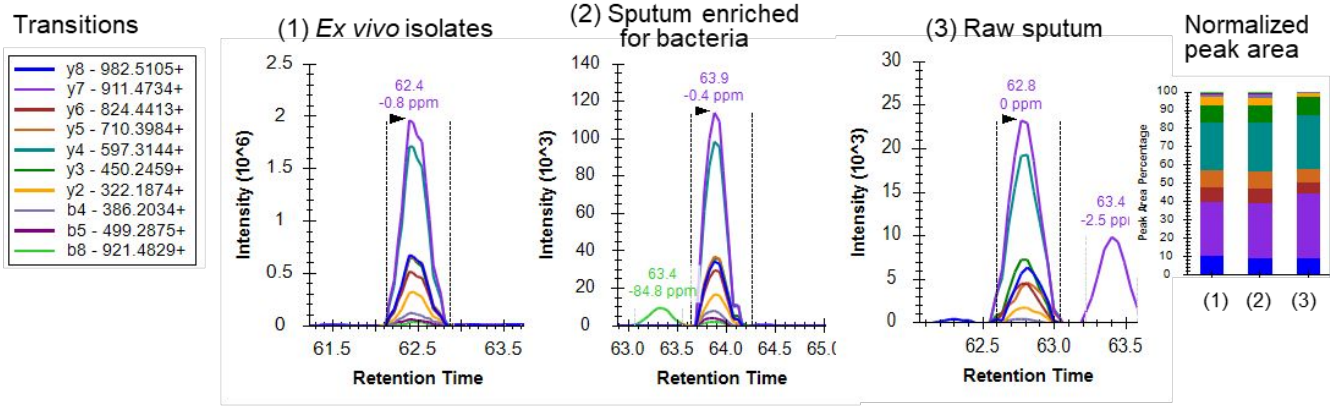


Figure S6

Glyoxylate cycle enzymes

PA2634 AceA

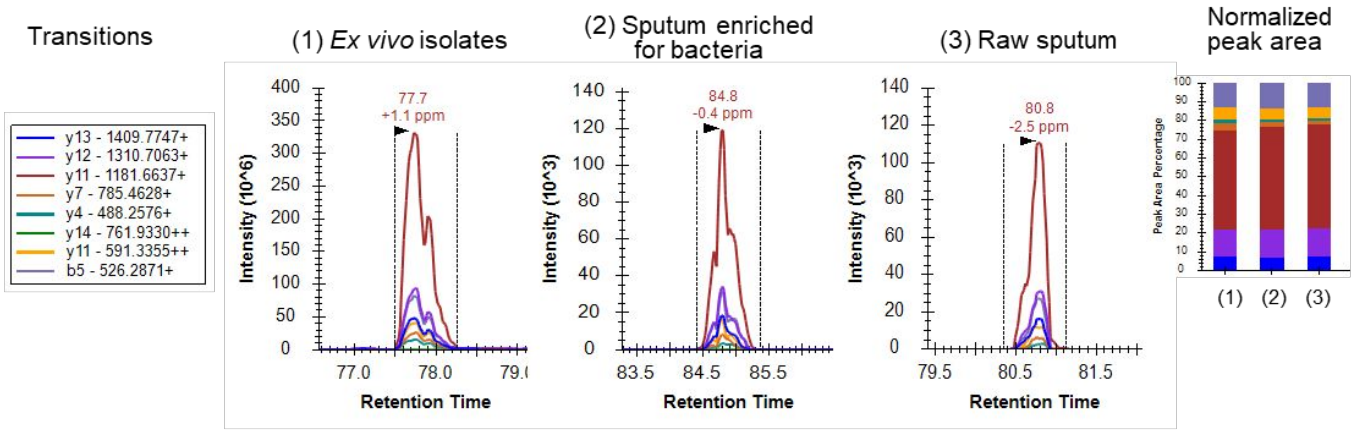
R.LASNLFQFR.K [318, 326]



Alginate biosynthesis proteins

PA3540 AlgD

K.SPIVEPGLEALLQQGR.Q [44, 59]



K.AGVDFGVGTNPEFLR.E [146, 160]

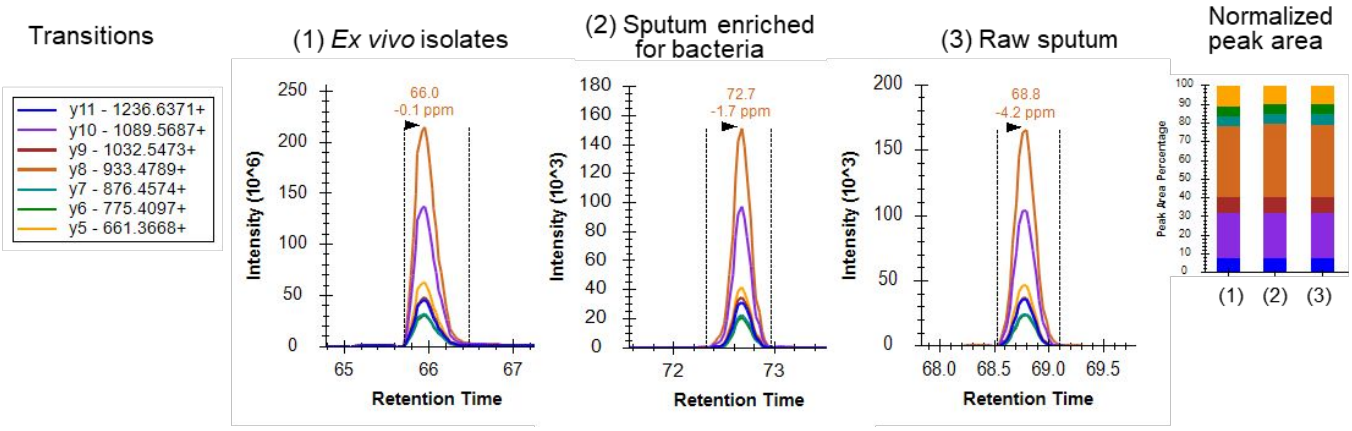


Figure S6

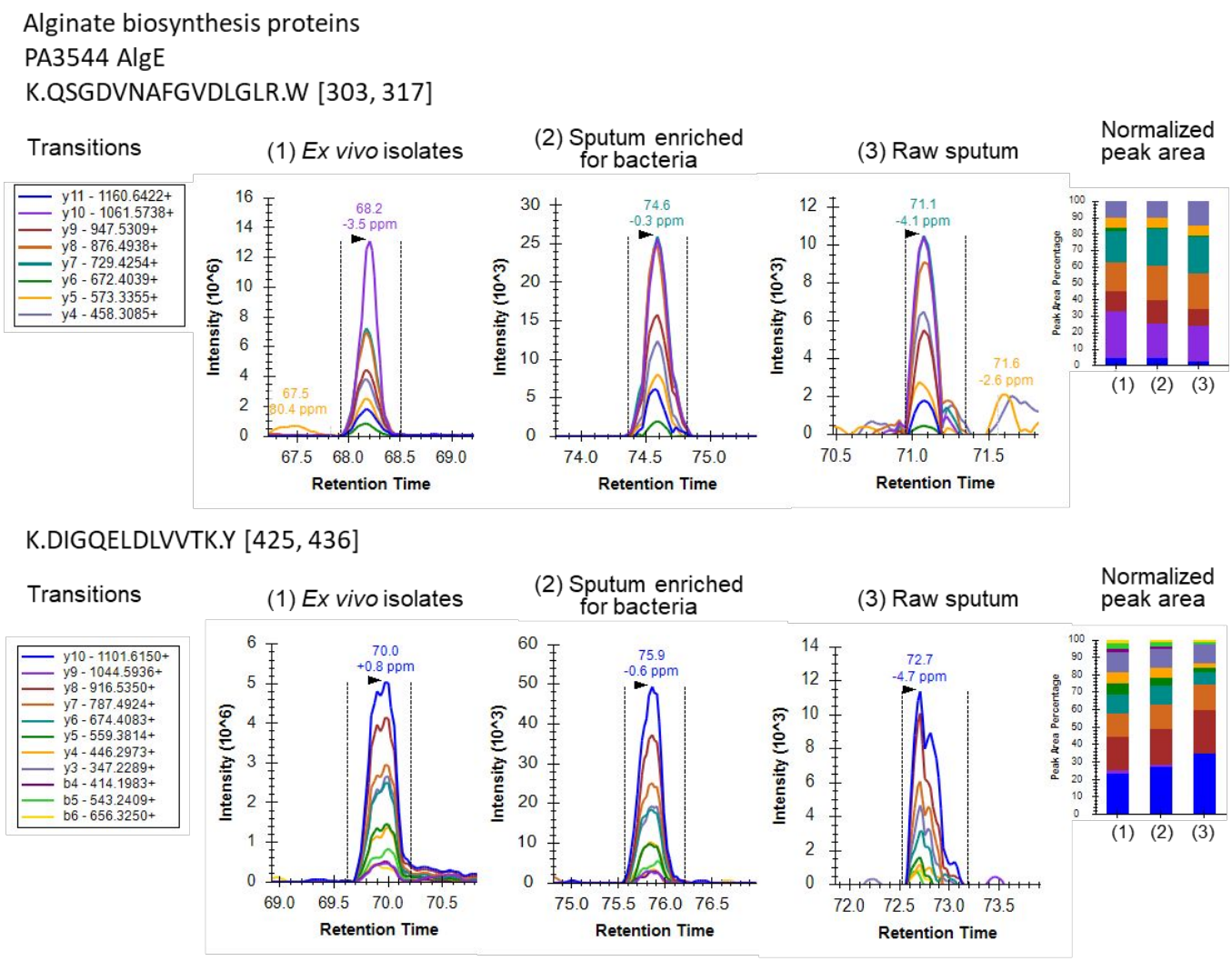


Figure S6. Spectrum examples of parallel reaction monitoring (PRM) assays to detect PA proteins with CF sputum samples. PRM signals were visualized with Skyline (version 4.1). Transitions with the most abundant signals for each peptide were listed. The normalized peak area shows the consistence of transition signals for each PA protein detected with (1) the *ex vivo* PA isolates, (2) the CF sputum samples processed with bacterial enrichment method, or (3) the raw sputum which was not processed with bacteria enrichment.