

## Supplementary Information

**Table S1. Identification of MUC5B and MUC5AC in submucosal gland secretions.**

The major high molecular weight glycoprotein in submucosal gland secretions migrating with a molecular mass of ~4 MDa by 1D SDS-AgPAGE was excised and digested with trypsin, and the resulting peptides were analysed by MALDI-TOF-MS. PMF searching identified the presence of (A) MUC5B and (B) MUC5AC. Details of the proteins identified, masses of the tryptic peptides detected, error to the theoretical peptide masses, the position of the peptides within each protein, and the peptide sequences are shown for CF subject 1. This data is representative of all subjects.

A	Protein	SwissProt Accession	MW (Da)	pI	Unique Peptide Hits	% Coverage
	Mucin 5B	Q9HC84	590122	6.24	42	10.56
<i>m/z</i>	Error (ppm)	Peptide Position		Alkylated Cysteines	Peptide Sequence	
605.3	34.2	5364-5368			FPGER	
722.4	33.3	1009-1014			TSVFIIR	
879.4	-56.5	5283-5289		1	GVCSDWR	
893.4	34.5	155-161			EELPYSR	
906.4	-39.5	1588-1593		1	MCYNYR	
931.5	-69.8	321-327		2	CPELCPR	
960.6	-62.2	678-685			GVQLSDWR	
973.6	-46.6	5190-5197			TWLVPDSR	
1036.6	-38.3	2362-2371, 2891-2900, 3589-3598			AQAOPGVPLR	
1077.5	-47.6	1070-1078		1	DPTCANPFR*	
1107.6	-24.0	626-635			LTDPNSAFSR	
1132.5	-33.6	2398-2405, 2927-2934, 3625-3632, 4154-4161		1	MCFNYEIR	
1140.7	4.6	121-131			GLVGSRPVTR	
1258.7	44.4	226-236			LTPLOFGNLQK	
1325.7	-53.6	109-119			AAYEDFNVQLR	
1364.7	-52.8	1867-1876		1	FNMCFNYNVR	
1468.8	-40.5	456-468		1	CADSSFTVLAELR	
1498.8	-36.9	1567-1578		2	VHCDVHFLVCR	
1508.7	38.8	1044-1057			SVVGDALFGNSWK	
1539.8	2.6	455-468			KCADSSFTVLAELR	
1539.8	2.6	456-469			CADSSFTVLAELRK*	
1617.8	-41.4	1085-1098		2	QCSILHGPTFAACR	
1630.8	-44.1	2346-2361, 2875-2890, 3573-3588, 4102-4117		1	AAGGAVCEQPLGLECR	
1679.9	-54.5	1369-1382		2	GYQVCPVLAIECR	
1687.8	-29.9	2346-2361, 2875-2890, 3573-3588, 4102-4117		2	AAGGAVCEQPLGLECR*	
1724.9	-40.7	938-954			IVTENIPCGTTGTTCSK*	
1747.8	-37.6	294-307		3	CPTCPCATFVEYSR	
1848.0	-43.5	5305-5320		2	VYKPCGPQPATCNSR	
1938.0	-41.2	5265-5280		2	LEVPCQSLEYAELCR	
1981.0	-34.2	2372-2388, 2901-2917, 3599-3615		2	ELGQVVECSLDGLVCR	
2027.9	12.6	1766-1784			TETTMSPLNTTTSQGTR	
2113.0	-40.1	659-677		3	SEDCLCAALSSVHACAAK	
2113.0	-7.3	1182-1200			NPSGHCLVDLPGLEGYPK	
2122.9	-21.4	4906-4922		3	AGCHFYAVCNQHCDIDR	
2151.0	27.4	938-957		2	IVTENIPCGTTGTTCSKAIK	
2257.9	-15.6	5632-5649			YSAFAQAMQHQCTCCQFR	
2257.9	39.1	1021-1041		3	GRVGLCLGNFDNAINDFATR	
2262.1	-35.1	1405-1424			GLMCANSQQSPPLCHDYELR	
2285.0	28.1	1383-1402		1	AAQLPDMPLEELGQVQVCDR	
2315.1	-8.8	1021-1041		1	GRVGLCLGNFDNAINDFATR	
2376.1	-28.0	1405-1424		2	GLMCANSQQSPPLCHDYELR	
2878.1	22.1	328-351		4	TCPLNMQHQECGSPTDTCNSPQR	

  

B	Protein	SwissProt Accession	MW (Da)	pI	Unique Peptide Hits	% Coverage
	Mucin 5AC	Q8WWQ4	135520	9.11	10	14.68
<i>m/z</i>	Error (ppm)	Peptide Position		Alkylated Cysteines	Peptide Sequence	
862.4	-13.6	269-275, 567-573, 1000-1006		1	EEGLVCR	
900.5	37.0	238-244, 969-975			RPEEITR*	
1022.5	-12.6	223-230, 521-528, 954-961			ETYNNIIR	
1046.6	-87.7	178-187			TSTSHVISIK	
1061.5	-7.9	276-284, 574-582, 1007-1015			NQDQQGPFK*	
1084.5	-35.1	285-292, 583-590, 1016-1023		1	MCLNYEV*	
1360.6	-30.1	495-504, 928-937		2	DCHPRLTWTK	
1848.0	-53.6	269-284, 567-582, 1000-1015			EEGLVCRNQQGPFK	
2062.1	-36.9	549-566		1	SHPEVSIHELGVVVQCSR	
2262.1	1.6	249-268, 980-999		1	AESHPEVSIHELGVVVQCSR	

\* Sequence confirmed by MS/MS

**Table S2. Identification of gp-340 in submucosal gland secretions.**

The high molecular weight glycoprotein in submucosal gland secretions migrating with a molecular mass of ~1 MDa by 1D SDS-AgPAGE was excised and digested with trypsin, and the resulting peptides were analysed by MALDI-TOF-MS. PMF searching identified the presence of gp-340 variant protein. Details of the protein identified, masses of the tryptic peptides detected, error to the theoretical peptide masses, the position of the peptides within each protein, and the peptide sequences are shown for non-CF lung disease subject 3. This data is representative of all subjects.

Protein	SwissProt Accession	MW (Da)	pI	Unique Peptide Hits	% Coverage		
Gp-340	Q9UKJ4	260591	5.18	11	19.9		
m/z	Error (ppm)	Peptide Position			Oxidised Methionine	Alkylated Cysteines	Peptide Sequence
730.4	-1.7	104-110, 236-242, 365-371, 496-502, 604-610, 864-870, 995-1001, 1124-1130, 1253-1259, 1382-1388, 1511-1517, 1642-1648					LVNGDGR
778.4	-38.2	247-252, 376-381, 507-512, 615-620, 746-751, 875-880, 1145-1140, 1264-1269, 1393-1398, 1653-1658					VEVLYR
838.4	103.9	2360-2366					CKMVVCR
991.6	-36.4	245-252, 1651-1658					GRVEVLYR
1279.7	53.2	372-381, 503-512, 611-620, 742-751, 871-880, 1131-1140, 1260-1269, 1389-1398		1			CQGRVEVLYR
1420.7	-82.3	2086-2098					GSFTSSSNFMSIR
1459.8	-72.3	418-431, 549-562, 657-670, 788-801, 917-930, 1048-1061, 1177-1190, 1306-1319, 1435-1448, 1695-1708					FGQGSGPIVLDDVR*
1545.7	-82.0	403-417, 902-916, 1291-1305, 1420-1434		1			QLGCGWATSAPGNAR
1575.7	-85.0	642-656, 773-787, 1162-1176, 1680-1694			1		QLGCGWAMSAPGNAR
1591.7	-49.7	642-656, 773-787, 1162-1176, 1680-1694		1			QLGCGWAMSAPGNAR
1601.8	-51.2	534-548			1		OLGCGWAMLA PGNAR

\* Sequence confirmed by MS/MS

**Table S3. Monosaccharide compositions of *O*-linked oligosaccharides from MUC5B/MUC5AC.** *O*-linked oligosaccharides were released by reductive  $\beta$ -elimination from MUC5B/MUC5AC in submucosal gland fluid. The resulting oligosaccharide alditols were analysed by LC-ESI-MS/MS, and their relative intensities were determined based on single-ion chromatograms. GlycoComp was used to determine the monosaccharide composition of the detected oligosaccharides. Masses corresponding to oligosaccharides, the monosaccharide composition of each oligosaccharide and the normalised percent abundance of each oligosaccharide from each subject are shown.

M	Composition	CF Lung Disease					Non-CF Lung Disease					Control		
		1	2	3	4	5	1	2	3	4	5	1	2	3
Neutral														
531.2	[ 1 , 1 , 1 , 0 , 0 ]					6.1					0.5			
588.1	[ 2 , 1 , 0 , 0 , 0 ]	0.8	0.9		1.4		1.7		1.9	0.8	1.0	3.9		16.2
734.2	[ 2 , 1 , 1 , 0 , 0 ]	3.8	2.5	1.7		6.5		4.4		4.5	3.1		7.5	
750.3	[ 2 , 2 , 0 , 0 , 0 ]				1.9		3.2	1.0	1.7	1.2	1.4	0.8		2.9
791.2	[ 3 , 1 , 0 , 0 , 0 ]							2.1		1.5	1.2			2.0
880.2	[ 2 , 1 , 2 , 0 , 0 ]				0.7					0.3			1.6	
896.2	[ 2 , 2 , 1 , 0 , 0 ]			2.8	1.9		4.9		4.2	3.8	5.1		4.2	3.5
937.2	[ 3 , 1 , 1 , 0 , 0 ]	6.3			6.4		6.2		4.8		3.0	1.1		12.7
953.3	[ 3 , 2 , 0 , 0 , 0 ]						7.3	2.5	5.6	2.7	1.7	3.1		7.8
1042.4	[ 2 , 2 , 2 , 0 , 0 ]				3.7		6.9		2.3	4.1	1.3		7.4	
1083.3	[ 3 , 1 , 2 , 0 , 0 ]	1.9			1.4				2.1		0.5			
1099.3	[ 3 , 2 , 1 , 0 , 0 ]	1.6	5.8	7.9		6.6	4.2	10.6	5.8	14.2	5.0	1.6	7.3	6.8
1140.4	[ 4 , 1 , 1 , 0 , 0 ]	6.4												
1188.3	[ 2 , 2 , 3 , 0 , 0 ]	1.5			2.3				2.3				3.1	
1204.4	[ 2 , 3 , 2 , 0 , 0 ]											2.3		
1245.4	[ 3 , 2 , 2 , 0 , 0 ]		8.1	16.8		16.8			14.7		13.1	6.1		21.9
1261.3	[ 3 , 3 , 1 , 0 , 0 ]									0.7		2.6		
1286.5	[ 4 , 1 , 2 , 0 , 0 ]	4.6					1.5							
1350.5	[ 2 , 3 , 3 , 0 , 0 ]											2.4		
1391.3	[ 3 , 2 , 3 , 0 , 0 ]								8.7		3.8	1.7		
1407.3	[ 3 , 3 , 2 , 0 , 0 ]											1.7		
1448.4	[ 4 , 2 , 2 , 0 , 0 ]	7.0			8.0								7.6	
1594.4	[ 4 , 2 , 3 , 0 , 0 ]	8.5			1.2									
1798.0	[ 5 , 2 , 3 , 0 , 0 ]	1.7												
Sialylated														
676.2	[ 1 , 1 , 0 , 1 , 0 ]	3.6								0.7	1.0			
717.2	[ 2 , 0 , 0 , 1 , 0 ]	1.1		0.8						0.8	0.9			
879.1	[ 2 , 1 , 0 , 1 , 0 ]		1.9	1.1	1.2	1.7	2.0		1.8	1.4	1.4	2.2	0.9	3.5
967.2	[ 1 , 1 , 0 , 2 , 0 ]	1.5												
1025.2	[ 2 , 1 , 1 , 1 , 0 ]	1.1								1.1	0.9	2.0		
1041.5	[ 2 , 2 , 0 , 1 , 0 ]	3.2	7.6	1.4	4.7		4.9	1.6	3.0	2.7	3.4	4.6		
1082.3	[ 3 , 1 , 0 , 1 , 0 ]	6.7	9.4	6.8	1.9	4.0	2.9	3.0	1.6	2.9	3.8		5.4	
1187.4	[ 2 , 2 , 1 , 1 , 0 ]	2.4	9.6	5.3	9.9	8.7	7.3	1.4	12.5	6.4	6.4	6.3	2.2	16.5
1228.4	[ 3 , 1 , 1 , 1 , 0 ]							1.7		3.0				
1244.4	[ 3 , 2 , 0 , 1 , 0 ]	2.7	7.4	2.2		4.2	10.5	8.5	0.8	3.3	2.6	13.1		
1316.4	[ 2 , 1 , 1 , 2 , 0 ]									1.0				
1323.3	[ 2 , 2 , 0 , 2 , 0 ]	6.7			4.5	8.4	2.4			4.6	4.7			
1333.4	[ 2 , 2 , 2 , 1 , 0 ]		3.3	1.3		7.4					3.2			
1390.4	[ 3 , 2 , 1 , 1 , 0 ]	7.9	25.3	12.0	6.4	11.6	12.9	13.5	17.1	15.4	10.6	19.8	12.1	18.5
1478.4	[ 2 , 2 , 1 , 2 , 0 ]	3.7	5.8		12.2		5.0	3.4	1.4	4.9	3.9			
1535.4	[ 3 , 2 , 0 , 2 , 0 ]			2.5						1.4	1.9	5.3		
1536.7	[ 3 , 2 , 2 , 1 , 0 ]	0.6	7.2	7.3			1.7	4.5	3.9	3.5	3.1		6.1	
1552.3	[ 3 , 3 , 1 , 1 , 0 ]				4.1		2.3				3.7			
1593.5	[ 4 , 2 , 1 , 1 , 0 ]	9.8		2.2								10.0		
1681.3	[ 3 , 2 , 1 , 2 , 0 ]		2.6	1.7					9.0					
1698.0	[ 3 , 3 , 2 , 1 , 0 ]									0.4	1.8			
1698.8	[ 3 , 3 , 0 , 2 , 0 ]				1.7					1.2	0.8			
1739.4	[ 4 , 2 , 2 , 1 , 0 ]	3.3								0.6				
1756.0	[ 4 , 3 , 1 , 1 , 0 ]													
1828.0	[ 3 , 2 , 2 , 2 , 0 ]			1.7					2.4					
1844.0	[ 3 , 3 , 1 , 2 , 0 ]			3.0										
1901.0	[ 4 , 3 , 0 , 2 , 0 ]									1.5				
1902.2	[ 4 , 3 , 2 , 1 , 0 ]	0.9								1.7				
1990.4	[ 3 , 3 , 2 , 2 , 0 ]			3.9		2.2		1.6			4.1			
2064.2	[ 4 , 4 , 2 , 1 , 0 ]			2.1										
2192.8	[ 4 , 3 , 2 , 2 , 0 ]								1.4					
2208.6	[ 4 , 4 , 1 , 2 , 0 ]			2.1					0.5		0.9			
2209.8	[ 4 , 4 , 3 , 1 , 0 ]						2.1		1.2			1.5		
2280.8	[ 3 , 3 , 2 , 3 , 0 ]			4.2		1.6		2.6						
2282.9	[ 3 , 3 , 4 , 2 , 0 ]			2.5										
2354.8	[ 4 , 4 , 2 , 2 , 0 ]										3.0			
2499.6	[ 4 , 4 , 1 , 3 , 0 ]			4.1		2.1		1.5			3.8			
2500.8	[ 4 , 4 , 3 , 2 , 0 ]									1.8				
2646.4	[ 4 , 4 , 2 , 3 , 0 ]			5.9		4.2		3.7			5.3		4.3	
2791.2	[ 4 , 4 , 1 , 4 , 0 ]			4.1										
2792.4	[ 4 , 4 , 3 , 3 , 0 ]					2.0					2.7			
2866.6	[ 5 , 5 , 3 , 2 , 0 ]			3.6										
2994.6	[ 5 , 4 , 1 , 4 , 0 ]								1.6					
3156.4	[ 5 , 5 , 1 , 4 , 0 ]			3.1				3.5						
3157.0	[ 5 , 5 , 3 , 3 , 0 ]			3.7						4.8				
Sulfated														
959.1	[ 2 , 1 , 0 , 1 , 1 ]	2.1					1.3							
1105.3	[ 2 , 1 , 1 , 1 , 1 ]						1.4							
1121.2	[ 2 , 2 , 0 , 1 , 1 ]						5.1	1.3						
1179.2	[ 3 , 2 , 1 , 0 , 1 ]						2.2	1.3		1.0		4.5		
1267.1	[ 2 , 2 , 1 , 1 , 1 ]	0.5					1.9			0.7				
1470.2	[ 3 , 2 , 1 , 1 , 1 ]						2.0							
1471.6	[ 3 , 2 , 3 , 0 , 1 ]		1.7											
2014.0	[ 4 , 5 , 0 , 1 , 1 ]								1.0					
2143.6	[ 4 , 4 , 2 , 1 , 1 ]				2.0									
2433.8	[ 4 , 4 , 0 , 3 , 1 ]										5.2			
2580.6	[ 4 , 4 , 1 , 3 , 1 ]				2.3		1.5							
2799.9	[ 5 , 5 , 2 , 2 , 1 ]			6.2								5.4		
2945.4	[ 5 , 5 , 1 , 3 , 1 ]													
2946.2	[ 5 , 5 , 3 , 2 , 1 ]									2.0				
3311.0	[ 6 , 6 , 3 , 2 , 1 ]				2.9									

M, measured neutral reduced mass; Composition, [N-acetylhexosamine, hexose, fucose, N-acetylnuramic acid, sulfate]. Assumption: All deoxyhexoses are fucose.

**Table S4. Monosaccharide compositions of *O*-linked oligosaccharides from gp-340.**

*O*-linked oligosaccharides were released by reductive  $\beta$ -elimination from gp-340 in submucosal gland secretions. The resulting oligosaccharide alditols were analysed by LC-ESI-MS/MS, and their relative intensities were determined based on single-ion chromatograms. GlycoComp was used to determine the monosaccharide composition of the detected oligosaccharides. Masses corresponding to oligosaccharides, the monosaccharide composition of each oligosaccharide and the normalised percent abundance of each oligosaccharide from each subject are shown. Although neutral oligosaccharides were only detected from subjects with non-CF lung disease, the low abundance of these components means that there are no significant differences in monosaccharide abundance between the subject groups.

M	Composition	CF Lung Disease					Non-CF Lung Disease					Control		
		1	2	3	4	5	1	2	3	4	5	1	2	3
Neutral														
588.1	[ 2 , 1 , 0 , 0 , 0 ]						1.9							
750.3	[ 2 , 2 , 0 , 0 , 0 ]						3.7							
896.2	[ 2 , 2 , 1 , 0 , 0 ]						1.0							
953.3	[ 3 , 2 , 0 , 0 , 0 ]						6.6							
1042.4	[ 2 , 2 , 2 , 0 , 0 ]											3.2		
1099.3	[ 3 , 2 , 1 , 0 , 0 ]						4.9					8.3		
1188.3	[ 2 , 2 , 3 , 0 , 0 ]											1.0		
1245.4	[ 3 , 2 , 2 , 0 , 0 ]											7.7		
1391.3	[ 3 , 2 , 3 , 0 , 0 ]											2.2		
2064.8	[ 4 , 4 , 4 , 0 , 0 ]											3.6		
Sialylated														
676.2	[ 1 , 1 , 0 , 1 , 0 ]	16.5	6.7										3.6	
879.1	[ 2 , 1 , 0 , 1 , 0 ]						1.9							
967.3	[ 1 , 1 , 0 , 2 , 0 ]												6.0	
1041.5	[ 2 , 2 , 0 , 1 , 0 ]						6.3					5.2		
1187.4	[ 2 , 2 , 1 , 1 , 0 ]			14.4			11.7					5.2	1.2	
1244.4	[ 3 , 2 , 0 , 1 , 0 ]						9.9							
1316.4	[ 2 , 1 , 1 , 2 , 0 ]	26.8	7.2	9.0	9.9									
1332.3	[ 2 , 2 , 0 , 2 , 0 ]	27.7	50.5	48.5	48.1	62.2	7.0	29.2	40.2	26.2	18.9	25.3	21.6	11.4
1390.4	[ 3 , 2 , 1 , 1 , 0 ]						22.6				7.1	3.9		
1478.4	[ 2 , 2 , 1 , 2 , 0 ]	16.7	15.0	18.5	42.3	37.8	11.9	14.8	14.7	10.5	20.0	15.1	16.8	25.1
1844.0	[ 3 , 3 , 1 , 2 , 0 ]											1.4		
1990.4	[ 3 , 3 , 2 , 2 , 0 ]											4.6	3.0	13.2
2208.8	[ 4 , 4 , 1 , 2 , 0 ]											1.6		
2353.6	[ 4 , 4 , 0 , 3 , 0 ]						8.3	8.5	6.4			10.1	20.0	
2280.8	[ 3 , 3 , 2 , 3 , 0 ]											4.2		
2499.6	[ 4 , 4 , 1 , 3 , 0 ]						17.2	19.1	9.0			14.8	30.0	8.1
2500.4	[ 4 , 4 , 3 , 2 , 0 ]	12.4	20.6	7.3			3.5	5.4				12.4		
2646.4	[ 4 , 4 , 2 , 3 , 0 ]							14.4	9.9	6.2	13.7	9.4	11.6	14.8
2791.2	[ 4 , 4 , 1 , 4 , 0 ]								7.7					
2792.4	[ 4 , 4 , 3 , 3 , 0 ]						6.1	10.8				13.6	7.6	13.9
Sulfated														
1105.3	[ 2 , 1 , 1 , 1 , 1 ]											2.9		
2579.8	[ 4 , 4 , 1 , 3 , 1 ]											3.9	5.0	

M, measured neutral reduced mass; Composition, [N-acetylhexosamine, hexose, fucose, N-acetylneuraminc acid, sulfate]. Assumption: All deoxyhexoses are fucose.

**Table S5. Structures of *O*-linked oligosaccharides from MUC5B/MUC5AC.** *O*-linked oligosaccharides were released by reductive  $\beta$ -elimination from MUC5B/MUC5AC in submucosal gland secretions. The resulting oligosaccharide alditols were analysed by LC-ESI-MS/MS, and their relative intensities were determined based on single-ion chromatograms. GlycosidIQ was used to assign structures to the detected oligosaccharides, which were also manually validated. Oligosaccharides were found to contain all major *O*-glycan cores, both chain types, and had terminal epitopes dominated by Lewis-type and blood group antigens. Masses corresponding to oligosaccharides, assigned oligosaccharide structures and the normalised percent abundance of each oligosaccharide from each subject are shown. Abbreviations: Fuc, fucose; Gal, galactose; GlcNAc, *N*-acetylglucosamine; GalNAc, *N*-acetylgalactosamine; NeuAc, *N*-acetylneuraminc acid; a –  $\alpha$  linkage; b –  $\beta$  linkage; 1, 2, 3, 4, 6 – linked carbon position.

