Sputum Proteomics in Inflammatory and Suppurative Respiratory Diseases

Robert D. Gray, Gordon MacGregor, Donald Noble, Margaret Imrie, Maria Dewar, A. Christopher Boyd, J. Alastair Innes, David J. Porteous, Andrew P. Greening

Online Data Supplement

Online Methods Supplement

Sputum Induction and Preparation

Subjects inhaled nebulised hypertonic saline solution at successive concentrations of 3, 4 and 5 % following the protocol of Green et al(11). Saline was nebulised through a liberty ultrasonic nebuliser (Liberty, UK). Each concentration of saline was inhaled for 5 minutes. Spirometry was performed following each inhalation step and the procedure abandoned if there was a drop in FEV_1 greater than 20 %. After each inhalation period the patient was asked to blow their nose into a clean handkerchief and rinse their mouth with water. Patients were then asked to cough and expectorate sputum if possible. This process was repeated for each successive concentration of HS. Sputum was collected into polypropylene tubes, transported on ice and processed within 2 hours of expectoration. All reagents and sample tubes are pre-chilled to 4 ^{0}C. The collected sample was transferred to a sterile Petri dish and mucus plugs selected using forceps. These were transferred to a fresh Petri dish and the plugs gathered together by moving the mass of plugs in circular movements around the dish until a single mass of mucus was formed expelling excess saliva. This was then transferred to a pre-weighed polypropylene tube and the weight of mucus calculated on a standard laboratory balance. 4 x weight in volume of 0.1% DTT (i.e. 4mls/1g of mucus) was added to the selected mucous plugs. This mixture was mixed with a 1ml pastette, vortexed for 15 seconds, then rotated at 4 ^{0}C for 15 minutes. A 4x weight in volume of DPBS was added and the sample vortexed for 15s. The diluted sample was then filtered through pre wet (with DPBS) 48μm pore nylon gauze set in a polypropylene funnel into a chilled 15ml polypropylene tube and centrifuged at 1200 rpm for 10 minutes at 4 ^{0}C. The supernatant was stored at -80 ^{0}C in polypropylene tubes until further analysis. The cell pellet was resuspended in DPBS in a ratio of 1 ml to 0.5mm depth of pellet. Resupended cells were spun onto glass

microscope slides at 400 rpm for 5 mins. Slides were air dried, fixed with methanol and stained by May-Grunwald-Giemsa using MGG Quick Stain (Bio-Optica, Milan, Italy).

ProteinChip Preparation Protocols

CM10 and Q10 experiments were performed using a 96 well bioprocessor (Ciphergen, Freemont, CA). IMAC Nickel Experiments were carried out without the bioprocessor due to concerns of the metal reagent binding to the bioprocessor reservoir and not the chip surface. NP 20 (normal phase ProteinChip arrays, used in protein identification experiments) experiments were carried out without the bioprocessor as the surface requires no pre activation. All experiments were performed at room temperature.

CM10 and Q10 ProteinChips were placed in the bioprocessor. Chromatographic spots on the chip surface were equilibrated by adding buffer: 200 µl of either 100 mM ammonium acetate/0.1% triton pH 4 (CM10) or 100 mM tris HCl/0.1% triton pH 8(Q10) to each well. The bioprocessor was shaken at 600 rpm for 5 minutes, emptied and then this process was repeated. The bioprocessor was emptied and 20 µl of sample (sputum fluid phase) was added to a well with 230ul of the appropriate buffer/0.1% triron. The plate was sealed with adhesive plate sealer and shaken at 600 rpm for 2 hours. The bioprocessor was emptied and 2 washes of 5 minutes were performed with buffer/0.1% triton on the shaker, followed by 3 washes with buffer alone. A final wash of 280 µl ultra pure (UP) H_2O was then performed without shaking. The bioprocessor was opened chips removed and allowed to air dry. 0.8 ul of sinapinnic acid matrix (in 50% acetonitrile/0.5% trifluoroacetic acid) was then added to each spot and allowed to air dry. This process was repeated. Chips were then read in the SELDI TOF mass spectrometer as below.

For IMAC experiments 10ul of 100mM nickel sulphate was added to each spot on an IMAC ProteinChip and incubated in a humidified chamber for 15 minutes. Each spot was then washed with 10 ul UPH_2O. A further 10ul of 100mM nickel sulphate was added and incubated as above or a further 15 minutes. Spots were washed as above and excess moisture removed from the chip using fine tissue paper. 6ul of 0.5M sodium chloride/0.1% Triton was added to each spot. 2ul of sample (sputum supernatant) was then added to each spot. Chips were incubated for 2 hours in an humidified chamber, after which spots were washed twice with 10ul 0.5M sodium chloride in PBS/0.1%Triton and then twice with 0.5M sodium chloride in PBS with no triton. A final wash was performed with UPH_2O. Excess moisture was removed from spots and chips were allowed to air dry. 8 ul of sinapinnic acid matrix (in 50% acetonitrile/0.5% trifluoroacetic acid) was added to each spot and allowed to air dry. This process was repeated. Chips were then read in the SELDI TOF mass spectrometer as below.

ProteinChip Reading Protocol

All SELDI TOF mass spectrometry was performed on a Protein Biology System 2 (PBS II) mass spectrometer (Ciphergen, USA). Chips were placed in the mass spectrometer and read automatically with the following settings unless otherwise stated. Laser intensity of 205 with deflector set at 4000 Da and a focus mass of 7500 Da with optimisation of the instrument from 3000 to 25000 Da. Chips were read in the mass range 0-50000 Da unless otherwise indicated. Chips were read at 12 individual areas on each spot with 8 laser shots at each spot being recorded allowing an average of 100 shots for each spot to produce mass spectra. Prior to reading each area was given two warming shots at higher laser intensity which were not included in the average spectra. Data were exported from the mass spectrometer to a standard PC and analysed by platform specific software.

Data Processing and Analysis

Prior to comparisons of spectra the data were processed on Ciphergen ProteinChip® Software (Ciphergen, USA) and Ciphergen Express® (Ciphergen USA). Baseline correction was performed to enhance the contrast of peaks to baseline using a fitting width of 4 times the expected width. Noise was automatically measured from 4 to 50 kDa and spectra corrected accordingly. Data were then normalised for total ion current. The total ion current for an individual spectrum was divided by the average total ion current over all spectra and thus each spectrum was awarded a normalisation coefficient. A normalisation coefficient of 1 reflected individual AUC the same as average, <1 suggested a greater AUC than average and >1 a smaller AUC than average. Spectra with high normalisation coefficients were individually scrutinised and if they reflected poor quality or absent spectra were excluded from analysis. Finally peaks were detected using a peak detection wizard. Peaks with a signal to noise ratio >3 were selected to perform analysis within and between groups.

Further analysis was performed on the platform specific software Ciphergen Express® (Ciphergen USA). Following the above processing steps data were initially compared using a cluster wizard. This automatically clusters peaks across all spectra of a similar inferred molecular weight (within 0.3%) thus allowing the comparison of biomarkers across disease groups. Statistical analysis between individual groups was performed using Mann Whitney testing.

Reproducibility of SELDI TOF in Sputum

Reproducibility was assessed on the CM10 weak cation exchange ProteinChips activated at pH 4 to ensure a degree of uniformity in experiments. To assess intra-assay variability a sample from a single subject with CF was spotted onto eight adjacent spots on a CM10 ProteinChip activated at

pH4. The chip was read automatically and protein peaks clustered. Peaks with a signal to noise ratio of greater than 10 were selected and the coefficient of variation for each cluster was calculated on GraphPad Prism (GraphPad, CA, USA). An average intra-assay coefficient of variation was calculated. To assess inter-assay variabilty samples from a single subject were spotted onto six individual spots on six separate ProteinChips. Chips were read and clustered as above and coefficients of variation calculated for individual clusters of peaks

Online Data Supplement

Table E1

The total number of proteins detected by SELDI-TOF, significantly different for disease versus control, are shown. The number of proteins achieving difference with a statistical significance of

P<0.01 and *P*<0.001 are also shown. Proteins could be increased or decreased in their abundance. Spectra that failed normalisation were excluded from analysis (see text). n = number of normalised spectra used in each analysis. The number for the disease is given first and for the controls given second. The spectra failing to normalise were not the same for each analysis. The data are displayed for each of the three chip surfaces that were used in the study. The number of proteins that were significantly different for disease versus disease, are also shown. For the disease versus disease comparisons, the proteins on all three chip surfaces were summed.

Asthma versus Controls			
	IMAC Nickel (n=22vs20)	**Weak Cation Exchange (CM10) (n=20vs19)**	**Anion Exchange (Q10) (n=16vs18)**
Total number of peaks	178	250	193
Number of peaks at *P*<0.01	18	41	46
Number of peaks at *P*<0.001	4	16	36
COPD versus Controls			
	IMAC Nickel (n=23vs20)	**Weak Cation Exchange (CM10) (n=22vs19)**	**Anion Exchange (Q10) (n=21vs18)**
Total number of peaks	176	264	185
Number of peaks at *P*<0.01	32	37	44
Number of peaks at *P*<0.001	11	13	26
CF versus Controls			
	IMAC Nickel (n=28vs19)	**Weak Cation Exchange (CM10) (n=22vs19)**	**Anion Exchange (Q10) (n=22vs19)**
Total number of peaks	193	270	197
Number of peaks at *P*<0.01	125	140	116

Number of peaks at *P*<0.001	108	111	96
Bronchiectasis versus Controls			
	IMAC Nickel (n=22vs20)	**Weak Cation Exchange (CM10) (n=20vs19)**	**Anion Exchange (Q10) (n=16vs18)**
Total number of peaks	181	259	233
Number of peaks at *P*<0.01	122	131	124
Number of peaks at *P*<0.001	112	103	99
CF versus Bronchiectasis and Asthma versus COPD			
	CF vs Bronchiectasis Sum of all three chip surfaces	**Asthma vs COPD Sum of all three chip surfaces**	
Total number of peaks	577	485	
Number of peaks at *P*<0.01	58	16	
Number of peaks at *P*<0.001	9	1	

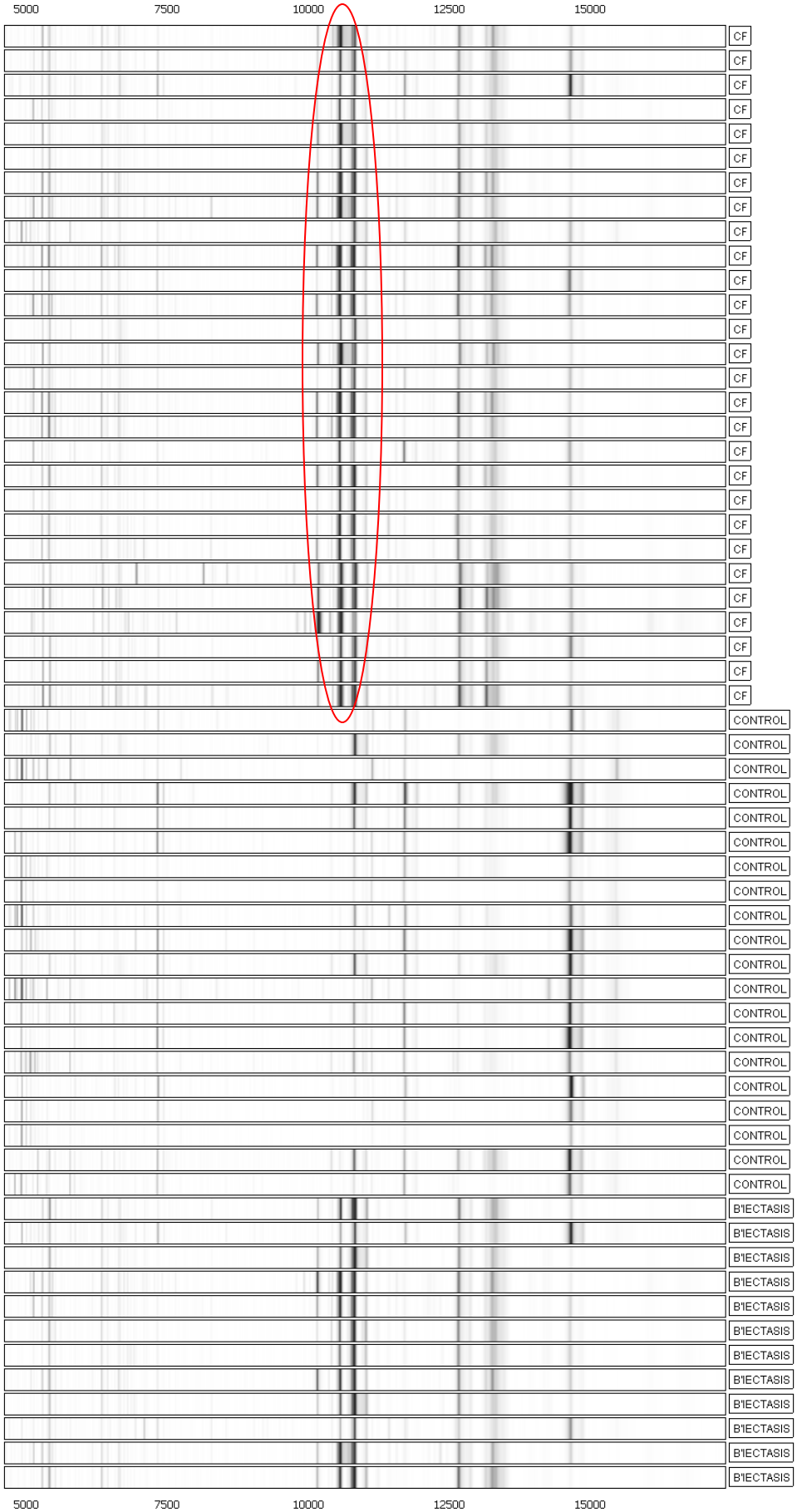
Table E2

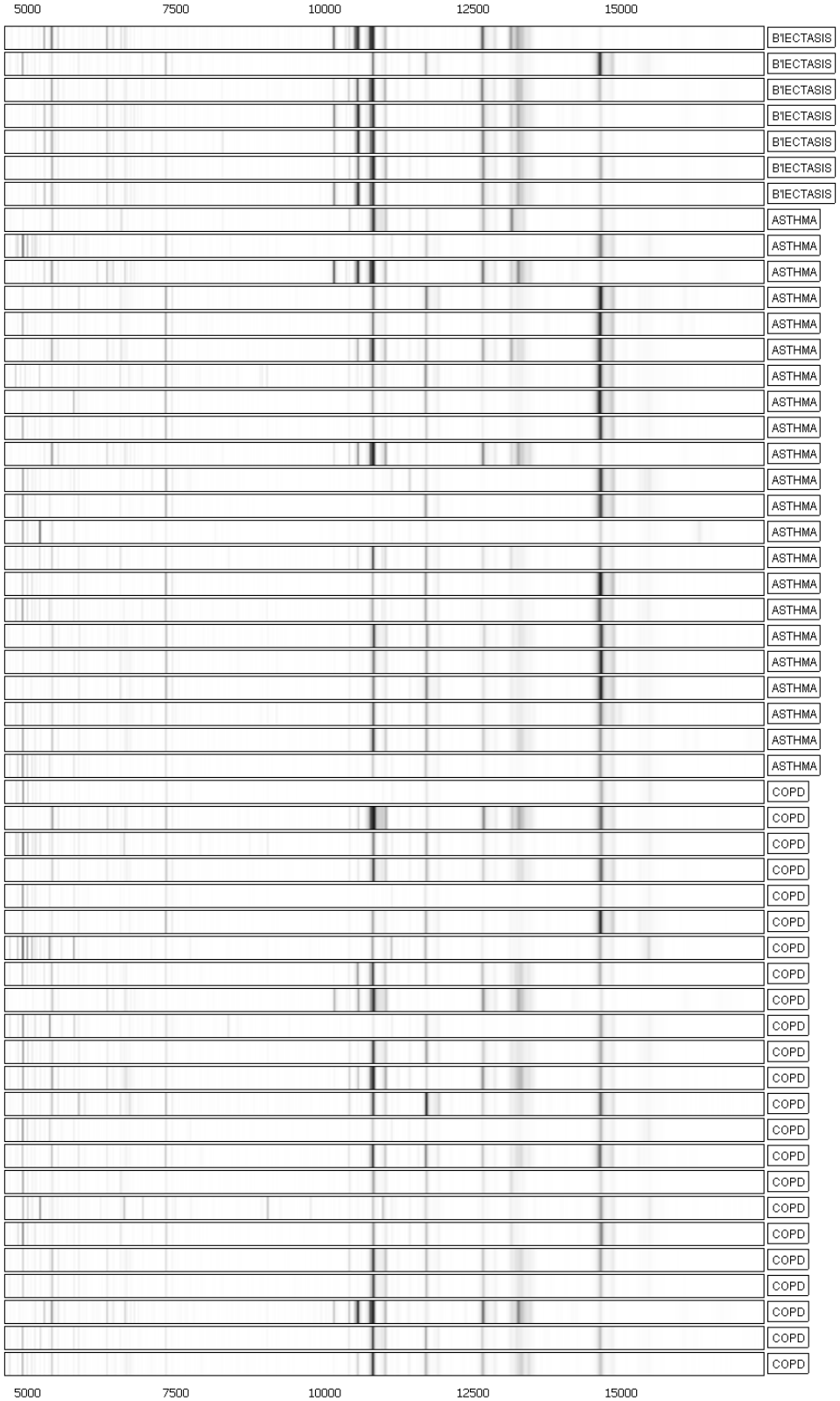
The top 20 markers in each disease group, ranked by statistical significance compared to control, are given indicating the chip surface on which they were identified and their molecular weights. Proteins that appear to be the same (by MW) between disease pairs are indicated by (*) for Asthma and COPD, and by (†) for CF and Bronchiectasis.

Asthma		COPD		Cystic Fibrosis		Bronchiectasis	
CHIP	**MW**	**CHIP**	**MW**	**CHIP**	**MW**	**CHIP**	**MW**
Q10	4752*	Q10	4008*	IMACNi	10576†	IMACNi	5395
Q10	5147*	Q10	5276*	IMACNi	12247	IMACNi	6453
Q10	4008*	Q10	5176	IMACNi	34626	IMACNi	4544
Q10	5272*	Q10	5419*	IMACNi	10786†	IMACNi	6344
Q10	4729*	Q10	4563*	IMACNi	12893	IMACNi	21385
Q10	6961	Q10	4590*	IMACNi	4808	Q10	4752
Q10	4088*	Q10	5147*	IMACNi	12681†	Q10	10584
Q10	5422*	Q10	4729*	IMACNi	15324	IMACNi	10787†
Q10	5303	Q10	4753*	IMACNi	15394	Q10	4008
Q10	4566*	Q10	7681	CM10	10586	Q10	7118
Q10	4590*	CM10	6293*	CM10	10796	IMACNi	10574†
CM10	6294*	IMACNi	5395*	CM10	12697	IMACNi	13777
Q10	5248*	Q10	4065	CM10	13383	IMACNi	21627
Q10	5624*	Q10	5250*	IMACNi	10173	CM10	5422
Q10	5396*	CM10	10749	IMACNi	24014	IMACNi	12678†
CM10	4160	Q10	4088*	CM10	7933	IMACNi	37185
Q10	7767	CM10	13320	IMACNi	4543	Q10	4088
CM10	15705	IMACNi	15085	IMACNi	4922	Q10	5001
IMACNi	4544	Q10	5624*	IMACNi	15484	Q10	5273
Q10	4852	Q10	17686	CM10	12946	Q10	11183

Figure E1

Mass spectral data represented as *gel view* from Ciphergen Express (Ciphergen, Freemont, CA) software demonstrating differential intensities in individual patients used in the analysis of the IMAC Ni surface following normalisation. Each data set is labelled by individual subject diagnosis: CF, Control, Bronchiectasis, Asthma and COPD. The intensity of a band represents the abundance of a peak at a given molecular weight as measured by signal intensity. From this data we can clearly see different expression of proteins across subjects and across groups even in the mass range of 5 to 15 kDa. The red circled area in the CF subjects highlights the two peaks identified as calgranulin A at 10576 and 10831 Da respectively, a similar pattern is noted in the Bronchiectasis patients. This figure compliments the individual spectra demonstrated in figure 1 of the paper which demonstrates individual spectrographs from 1 patient in each group over the same mass range on the IMAC Ni surface.


5000
7500
10000
12500
15000
CF
CONTROL
B'IECTASIS


5000
7500
10000
12500
15000
B'IECTASIS
B'IECTASIS
B'IECTASIS
B'IECTASIS
B'IECTASIS
B'IECTASIS
B'IECTASIS
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
ASTHMA
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
COPD
5000
7500
10000
12500
15000