

Proteomic Discovery and Development of a Multiplexed Targeted MRM-LC-MS/MS assay for Urine Biomarkers of Extracellular Matrix Disruption in Mucopolysaccharidoses I, II and VI

Wendy E. Heywood¹, Stephane Camuzeaux¹, Ivan Doykov¹, Nina Patel¹, Rhian-Lauren Preece¹, Emma Footitt², Maureen Cleary², Peter Clayton¹, Stephanie Grunewald², Lara Abulhoul², Anupam Chakrapani², Neil J Sebire², Peter Hindmarsh¹, Tom J. de Koning⁴, Simon Heales^{2,3}, Derek Burke³, Paul Gissen^{1,2} and Kevin Mills^{1,*}

¹ Centre for Translational Omics, UCL Institute of Child Health, 30 Guilford Street, London, WC1N 1EH UK

² Centre for Inborn Errors of Metabolism, Great Ormond Street Hospital, Great Ormond Street, London, WC1N 3JH, UK

³ Enzyme and Metabolic Unit, Chemical Pathology, Great Ormond Street Hospital for Children, Great Ormond Street, London, WC1N 3JH, UK

⁴ University of Groningen, University Medical Center Groningen, Departments of Genetics and Neurology, PO Box 30.001, 9700 RB Groningen, Netherlands.

ABSTRACT: The Mucopolysaccharidoses (MPS) are lysosomal storage disorders that result from defects in the catabolism of glycosaminoglycans. Impaired muscle, bone and connective tissue are typical clinical features of MPS due to disruption of the extracellular matrix. Markers of MPS disease pathology are needed to determine disease severity and monitor effects of existing and emerging new treatments on disease mechanisms. Urine samples from a small cohort of MPS-I, II and VI patients (n=12) were analysed using label-free quantitative proteomics. Fifty three proteins including many associated with extracellular matrix organisation were differentially expressed. A targeted multiplexed peptide MRM LC-MS/MS assay was used on a larger validation cohort of patient samples (MPS-I n= 10, MPS-II n= 15, MPS-VI n= 5, Control n=20). MPS-I and-II groups were further sub-divided according to disease severity. None of the markers assessed were altered significantly in the mild disease groups compared to controls. Beta-galactosidase, a lysosomal protein, was elevated 3.6-5.7-fold significantly ($p < 0.05$) in all disease groups apart from mild MPS-I and II. Collagen type I α , fatty-acid-binding-protein 5, nidogen-1, cartilage oligomeric matrix protein and insulin-like growth factor binding protein 7 concentrations were elevated in severe MPS I and II groups. Cartilage oligomeric matrix protein, insulin-like growth factor binding protein 7 and beta-galactosidase were able to distinguish the severe neurological form of MPS-II from the milder non-neurological form. Protein Heg1 was significantly raised only in MPS-VI. This work describes the discovery of new biomarkers of MPS that represent disease pathology and allows the stratification of MPS-II patients according to disease severity.

Mucopolysaccharidoses (MPS) are part of a larger group of rare genetic metabolic disorders known as Lysosomal Storage Diseases (LSDs), incidence of the various forms of MPS range from 0.05-1.89/100,000 live births¹. MPS results from a defect in enzymes involved in the glycosaminoglycan (GAG) degradation pathway, resulting in the lysosomal accumulation of GAGs, which results in progressive cellular damage, multiple organ failure and reduced life expectancy². To date the MPS disorders have been classified into seven distinct forms, which relate to the enzyme affected in the GAG degradation pathway; each MPS disease results in elevation of various GAG species of heparan, dermatan, keratan, chondroitin sulphates and hyaluronan. Pathological features of MPS typically include skeletal and connective tissue abnormalities with patients often presenting with skeletal dysplasia, decreased joint mobility, short stature and carpal tunnel syndrome. MPS disorders have a wide variation in disease phenotype, with mild forms that manifest in adulthood and severe forms that are

apparent in infancy¹. MPS I is further subdivided into Hurler disease, which has a severe clinical phenotype, Scheie disease, which is a clinically milder form of MPS I, and Hurler-Scheie, which has common characteristics of both phenotypes. MPS II is typically the severe neurological form of the disease whilst a non-neurological form is considered the milder variant. Available therapy for MPS disorders include hematopoietic stem cell transplantation and enzyme replacement therapy (ERT) for MPS I, II and VI³. To maximise the success of treatment it is necessary to diagnose these patients as early as possible. Currently, there is no simple and rapid test for the diagnosis of MPS. Suspected patients are tested for the presence of accumulated urinary GAGs which are typically performed using semi-quantitative electrophoresis techniques⁴. Quantitation of disaccharides of GAGs using tandem LC-MS methods are being developed^{5,6} and assessed for routine clinical practice. Whilst an observed reduction of the respective GAG species is used to monitor treatment of MPS patients, it

Table 1: Cohort sample information for disease phenotype, age and sex. All samples were reported GAG positive by the Great Ormond street diagnostic Enzyme laboratory.

Biomarker Discovery Cohort	Control		MPS I		MPS II		MPS VI
number	10		4		5		2
Age range (years)	2-39		1-9		4-9		5 & 7
Sex ratio Male:Female	5:5		2:2		5:0		1:1
Validation Cohort	Paediatric Control	Adult Control	MPS I (Hurler-Hurler- Scheie)	MPS I (mild)	MPS II (Neurological form)	MPS II (non-neurological form)	MPS VI
number	9	10	14	4	5	7	6
Age range (years)	2.5-11	21-39	1.5-14	12-34	1-5	1-12	0.24 -17
Sex ratio Male:Female	5:5	5:5	4:8	3:1	4:1	7:0	3:3

provides limited information on improvement in the pathology of the disease and hence GAGs are not considered ideal biomarkers for this⁷. Furthermore, no established biomarkers are available to monitor the course of neurological disease. Although the pathogenesis of MPS is due to progressive storage of GAGS, it is becoming apparent for many lysosomal storage disorders that more complex pathogenic mechanisms underlie patients' clinical symptoms. Therefore there is a need for biomarkers for MPS disorders to accompany GAG analysis. Ideal markers would be specific to a particular type or types of MPS, help to differentiate more severe from less severe disease phenotypes, correlate with neurological progression, respond to treatment and be easily detected and quantified. Three forms of MPS were used in this study. MPS I patients (Hurler, Hurler-Scheie, Scheie syndrome) who have defects in alpha-L-iduronidase, MPS II patients (Hunter syndrome) who have defects in iduronate sulfatase of which both are involved in the catabolism of heparan and dermatan sulphate, and MPS VI (Maroteaux-Lamy syndrome) patients who lack arylsulphatase B activity, involved in the catabolism of dermatan sulphate. This study has been performed to identify potential MPS biomarker candidates in patient urine using label free proteomics to compare initially a small cohort of urine samples from GAG positive MPS I, MPS II and MPS VI patients. Potential marker candidates were further developed into a targeted proteomic MRM LC-MS/MS assay to confirm the findings on a larger sample cohort.

Experimental Section

Ethics Approval: The study was performed under ethical approval given by the NHS Research Ethics committee London-Bloomsbury.

Sample collection: Samples were obtained from patients attending the metabolic clinic at Great Ormond Street Hospital, London and from adult patients attending a clinic at the National Hospital, Queen Square, London. Samples were clinically requested for urinary GAGs analysis and appropriate

metabolic investigations as part of clinical care. Excess samples were selected that had tested positive for both elevated urine GAGs and reduced enzymatic activity; surplus urine samples were stored at -80oC. Patient disease diagnosis (considered mild or severe) age, sex and GAG status were provided for anonymised samples. Sex and age details are summarized in table 1.

Label Free Proteomics: Label free proteomics was performed as described previously⁸. Urine samples were thawed, centrifuged to remove particulate and cell debris and 5 mls filtered using a 3 kDa molecular weight cut off filter (Millipore, UK). Desalted urine was protein assayed using the BCA protein assay kit (Sigma UK). 100 µg of protein was freeze dried, trypsin digested, yeast enolase peptide internal standard (Waters Corp. Milford, MA) was added to digests to a final concentration of 100 fmol/µl and the digests desalted prior to analysis using C18 spin columns (Agilent, UK). Digested peptides were re-suspended in 3% ACN, 0.1% TFA and analysed on a Waters QToF Premier mass spectrometer coupled to a NanoAquity UPLC. Each digest was analysed using a 1.5 hour LC-MSE analyses as described previously in our laboratory⁹. Proteins were identified using Waters ProteinLynx Global server v 3.0 with a downloaded UniProt human reference proteome database to which the sequence of P00924 yeast enolase and P00761 porcine trypsin were added manually. Fixed modifications of carboamidomethylation of cysteines, dynamic modifications of deamidation of asparagine/glutamine and oxidation of methionine, up to 3 missed cleavage sites and false discovery rate set at 4% mass tolerance for ion and fragments were set to auto. Only protein identifications with >95% confidence and more than 1 peptide were exported for differential expression analysis using Progenesis LC-MS software (non-linear dynamics, UK). ANOVA was used to determine significance in protein expression between groups for the label free proteomics discovery analysis. Whilst sample numbers for discovery were small, robust potentially useful markers would ideally show obvious changes that can then be confirmed by targeted analysis.

Table 2. MRM parameters for the quantotypic tryptic peptides for the 7 confirmed urine marker proteins. Product ions in bold are the optimal transitions.

Uniprot accession	Description	Peptide sequence	Precursor ion <i>m/z</i>	product ion <i>m/z</i>	Cone energy	Collision energy	Fragmention
Q01469	Fatty acid-binding protein, epidermal	FEETTADGR	513.46 ²⁺	620.47	52	14	y6
				749.63	52	16	y7
P49747	Cartilage oligomeric matrix protein	LVPNPGQEDADR	656.05 ²⁺	550.00	46	16	y10
				887.69	46	22	y8
P14543	Nidogen 1	EDLSPSITQR	573.35 ²⁺	351.4	29	26	y6 ²⁺
				701.60	29	16	y6
Q99715	Collagen alpha 1 XII chain	ITWAPFGSSDK	605.09 ²⁺	498.00	26	14	y9 ²⁺
				737.61	26	16	y7
P16278	Beta-galactosidase	AYVAVDGIPQGVLER	794.16 ²⁺	334.29	36	32	b3
				1083.94	36	26	y10
Q9ULI3	Protein HEG homolog 1	ALSLAPLAGAGLELQLER	608.34 ³⁺	684.31	22	16	y13 ²⁺
				819.97	22	10	y16 ²⁺
Q16270	Insulin like growth factor binding protein 7	TELLPGDR	451.07 ²⁺	344.3	52	16	B3
				670.65	52	16	y6
P00924	Internal standard- Enolase 1 <i>Saccharomyces cerevisiae</i>	SIVPSGASTGVHEALEMR	614.57 ³⁺	771.72	46	14	y15 ²⁺
		AVDDFLISLDGTANK	790.15 ²⁺	661.48	46	16	b6
		GNPTVEVELTTEK	709.05 ²⁺	623.4916	46	18	y11 ²⁺

Targeted MRM LC-MS/MS Assay: Candidate markers were selected based on significance, fold change (>2 fold) and quality of the label free proteomics analysis data (i.e. the most number of quantitating peptides). Additional and potential biomarkers from the lysosomal GAG catabolic pathway were also included in the final multiplexed assay to see if they could be detected by targeted analysis. Two representative quantotypic peptides for each protein were determined from the label free proteomics data (top 3 most abundant and optimum daughter spectra for quantitation). Other peptides were selected using the open source online Global Proteome Machine MRM database at www.gpm.org¹⁰. Custom synthesised peptides (Genscript, USA) were used to optimise the peptide detection, determine retention times and identify unequivocally the correct peptide/s in urine. The optimal peptide from the

selection of 2 peptides/protein and two transitions was chosen based on levels of detection in urine and reflection of quantitative values. Details of confirmed marker peptides are given in table 2. Transitions were also assessed and selected for non-interfering peaks as determined by a standard curve. Confirmed peptide chromatograms and standard curves are given in supplementary data figure S1. Stable isotope peptide internal standards cannot be used in this sample work up as they are lost during the filtration/purification step. Intact heavy labelled proteins would act as ideal standards for this workflow however they are not practical for a potential high throughput assay at this stage due to the low yield, cost and complexity of their synthesis. Twenty nanograms of a generic yeast enolase whole protein standard (Sigma,

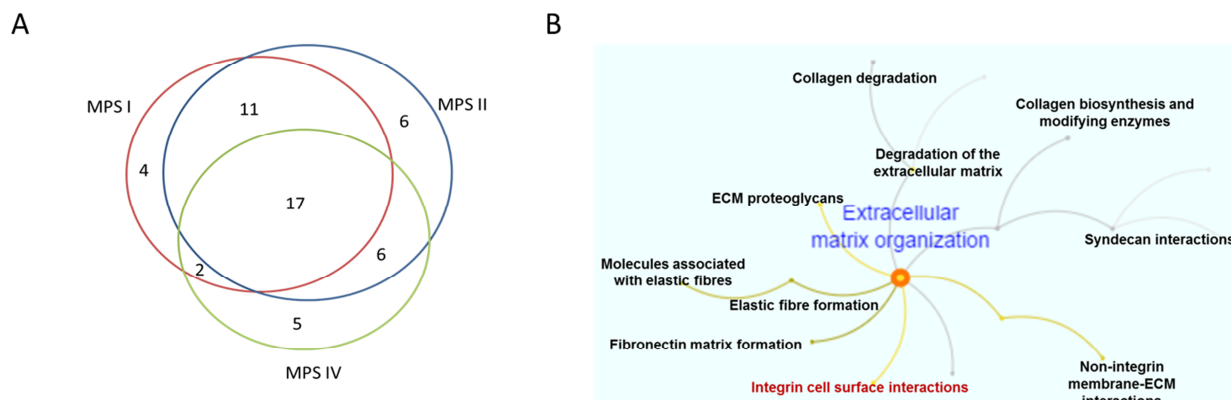


Figure 1. **A**. Summary Venn diagram of the label free proteomics analysis indicating the number of shared and unique proteins altered in MPS I, II and VI compared with controls. **B** Summarised illustration of the MPS disease pathology related pathways as identified by Reactome pathway over-representation analysis. Many pathways listed in supplementary data table S-2 are part of extracellular matrix organization.

UK) was used as an internal standard which can control for filtration and trypsin digestion. Yeast enolase is typically used in label free proteomics therefore is an established protein used for standardization and quantitation. This was added to 1ml of urine before sample processing and prepared as described previously for label free proteomics analysis. Digested peptides were injected onto a Waters CORTECS UPLC C18 + Column, 90Å, 1.6 µm, 3 mm X 100 mm column attached to a C18+ VanGuard Pre-column. UPLC and MS tune conditions were performed as described previously⁸. Dynamic multiple reaction monitoring was performed over a 10 min gradient with a minimum of 0.01 sec dwell time for quantitative transitions and minimum 12 data points per peak on a Waters Xevo TQ-S Mass Spectrometer. Urine creatinine was measured by LC-MS as described previously¹¹. Quality Controls (QC) made of pooled urine digests were run in triplicate at the start of the run and then every 10 injections. A CV of +/- 15% for each QC was considered acceptable. Samples were run in duplicate. Urine was spiked with peptides to create standards with average concentrations of biomarker levels and analysed for intra- and inter-batch variation. Chromatograms were analysed using Waters TargetLynx software. Peptides were standardised by using a yeast enolase peptide. Absolute levels were obtained from standard curves and standardised to urinary creatinine. Data was exported to Microsoft Excel and GraphPad Prism for statistical analysis. Analyses included data QC for peptide performance (coefficient of variance), QC of sample preparation and LC-MS/MS performance (yeast enolase). Standard curves were analysed by linear regression analysis and Pearson's correlation. Expression of candidate marker levels were checked for the effect of age by linear regression (Supplementary data, FigureS2) and non-parametric t-test (Mann Whitney test). Where no age effect was observed the two control groups were grouped as one in the analysis. Where an age-specific correlation was observed then comparison with disease groups was only with the corresponding control age group. Group comparisons were performed using the non-parametric Mann Whitney test.

RESULTS

Label free proteomics: Fifty three significantly differentially expressed proteins were observed in the urine of the MPS disease patients compared to controls ($p < 0.05$) and are listed in supplementary data table S-1. Figure 1 provides a summary of the analysis. Seventeen proteins were common to all disorders. More proteins were common to MPS I and II than with MPS VI. This is likely due to the similarity of these two MPS disorders which affect the same degradation pathway for heparan and dermatan sulphate whilst MPS VI only affects the dermatan sulphate pathway. Significant differentially expressed proteins were subjected to gene ontology analysis. Over-represented pathway analysis was performed using the curated pathway database Reactome^{12,13}. Significantly affected pathways indicated from the proteins detected in MPS urine are listed in supplementary data table S-2 and summarized in figure 1B. The most significant pathway affected is extracellular matrix (ECM) organisation. Many other highlighted pathways are implicated in other ECM related functions in particular the ECM involvement with wound healing.

Targeted MRM LC-MS/MS analysis: A multiplex and targeted proteomic assay using 1ml of urine was developed to confirm the findings from the label free proteomics. Proteins selected for confirmation are indicated in supplementary data table S-1. Additional proteins included in the assay were enzymes involved in the GAG degradation pathway that were not detectable with our current limits of detection in the QTOF discovery analysis but may be detectable by more sensitive targeted analysis in the experiments. Of these additional proteins only beta galactosidase was significantly altered in the study. Other additional proteins were either not detectable or not significant. A total of 52 peptide transitions were included in the assay. Twenty five were reliably quantifiable in 1ml of urine. A summarized overlaid chromatogram of the significant markers is shown in figure 2.

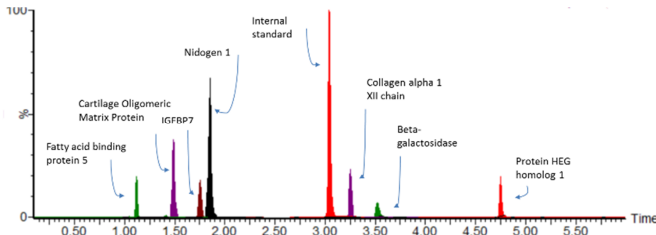


Figure 2. Overlaid chromatogram of multiplexed MRM quantitation including 7 MPS marker proteins in a single 10 minute LC-MRM/MS run.

Candidate markers were analysed for effect of patient age before comparative analysis; only collagen I (alpha chain) concentration was significantly age-dependent. Therefore, collagen I comparison of the disease groups was compared with paediatric controls only. Using this targeted proteomic approach, we were able to confirm that seven biomarkers demonstrated statistically significantly altered protein levels in the urine compared to controls. Six markers were elevated significantly in the MPS I (Hurler/Hurler-Scheie) and the MPS II groups (neurological form) groups.

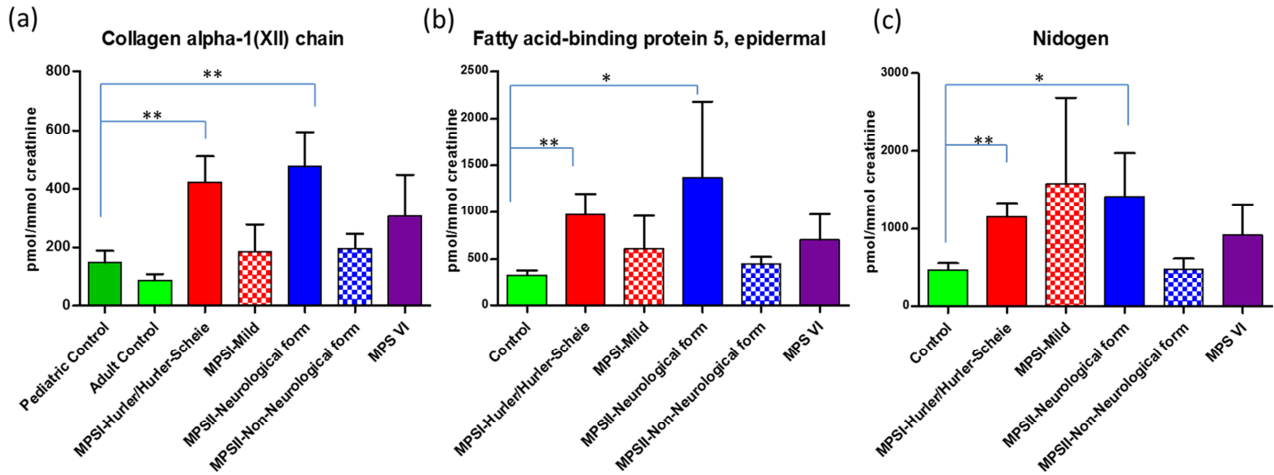


Figure 3. Targeted MRM LC-MS/MS analysis results showing the mean \pm SD of proteins found significantly elevated in MPS I and II compared to controls. Adult Control n= 20, Pediatric Control n= 10 MPS I H + HS (Hurler and Hurler-Scheie) n=13, MPS I mild n=3 MPSII Neuro (neurological phenotype) n= 5, MPSII non-neurological n=8 and MPS VI n= 5. * indicates $p < 0.01-0.05$, ** indicates $p < 0.001-0.01$, *** $p < 0.0001-0.001$.

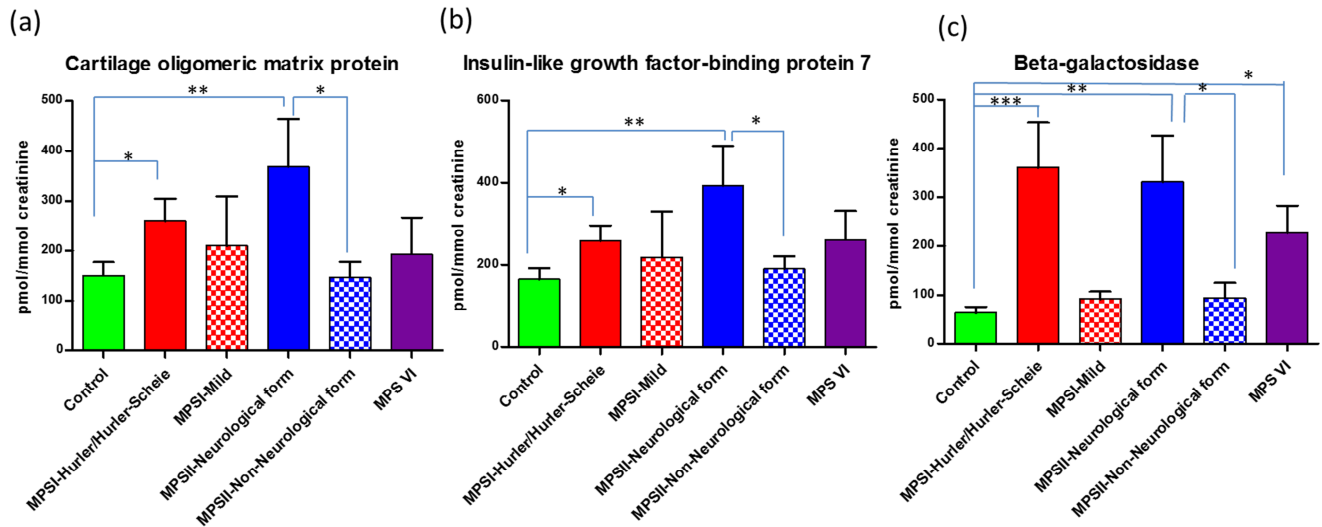


Figure 4. Targeted MRM LC-MS/MS analysis results showing the mean \pm SD of proteins found significantly elevated in MPS that can also discriminate the neurological and non-neurological MPS II phenotype. Control n= 10, MPS I n=15, MPS II n= 13 and MPS VI n= 5. Control n= 20, MPS I H + HS (Hurler and Hurler-Scheie) n=13, MPS I mild n=3 MPS II Neuro (neurological phenotype) n= 5, MPS II non-neurological n=8 and MPS VI n= 5. * indicates $p < 0.01-0.05$, ** indicates $p < 0.001-0.01$

Those that could only distinguish the MPS I and MPS II groups from the respective control group were 2.8 and 3.2 fold increases in collagen I (alpha chain), 3 and 4.2 fold increases in fatty acid binding protein 5 (FABP5) and 2.4 and 3 fold increases in nidogen-1 (Figure 3). Slightly raised levels but not statistically significant were observed in the mild MPS I and II groups. Another three markers, cartilage oligomeric matrix protein (COMP), insulin-like growth factor binding protein 7 (IGFBP7) and beta-galactosidase were raised significantly compared to controls but were also significantly elevated between the MPS II neurological form and the mild non-neurological form (Figure 4). COMP demonstrated a 1.7 and 2.4 fold change in MPS I and II, and a 2.5 fold change between the two MPS II groups. IGFBP7 demonstrated similar changes as COMP with 1.6 and 2.4 fold changes in MPS I and II compared with controls, and a 2.1 fold change between the MPS II groups respectively. Beta galactosidase also demonstrated an increase in the MPS VI group with a 3.6 fold change. A 5.7 and 5.3 fold change was observed in the MPS I and II groups, and a 3.5 fold change observed between the two MPS II groups. None of the markers were able to show significance between the MPS I (Hurler/Hurler-Scheie) group and the mild form of MPS I. Protein HEG1 was the only marker raised significantly in one specific disease group, MPS VI, which demonstrated a 1.9 fold elevation compared to controls. Raised levels were observed but not specific in the neurological MPS II group (Figure 5). A summary table of the significant marker observations is provided as supplementary data table S-3.

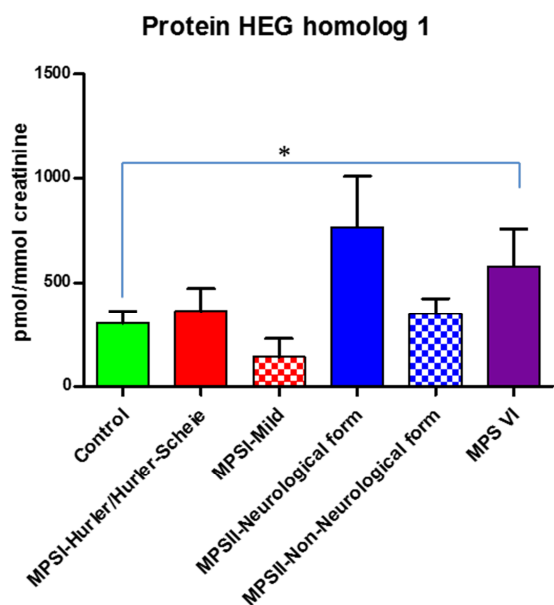


Figure 5. Targeted MRM LC-MS/MS analysis results showing the mean \pm SD of the significantly elevated protein HEG1 in MPS VI. Control n=20, MPS I Hurler and Hurler-Scheie n=13, MPS I mild n=3 MPS II neurological form n=5, MPS II non-neurological n=8 and MPS VI n=5. * indicates $p < 0.01-0.05$.

Liner regression analysis was performed on the MPS I Hurler/Hurler Scheie group to see if any of the markers showed a relationship with age progression. Only nidogen 1 showed a weak but significant correlation ($r^2=0.44$, $p < 0.019$) and is shown in supplementary figure data 2B. Nidogen 1 was not noticed to be af-

ected by age in the control cohort (supplementary figure 2A) therefore this may possibly be a marker for disease progression.

Discussion

The MPS disorders are diagnosed currently using either very labour intensive enzymatic tests or by the analysis of urinary GAG species. Although laboratories are moving to higher throughput LC-MS based systems for quantitative GAG analysis these are still complicated assays to develop and interpret. Current GAG analysis cannot discriminate the different disease severity phenotypes of MPS I and II. Whether future LC-MS disaccharide analysis will be able address this remains to be determined. Therefore, we investigated potential surrogate urine markers that could provide additional information to distinguish the different disorders and their subtypes. Furthermore, the identification of markers for biological systems affected in MPS, (for which total GAG levels provide no useful information) would provide data to stratify patients regarding prognosis and response to therapy. Additionally, biomarkers may be developed into a rapid and sensitive test for monitoring MPS disorders. This is particularly important with the increasing numbers of enzyme replacement therapies becoming available, where it would be necessary to diagnose and stratify patients earlier and also monitor treatment response.

Only recently have there been studies attempting to identify further biomarkers that can be used to detect pathological mechanisms associated with MPS. In accordance with our data, proteomic and genome wide expression studies of the MPS I mouse model¹⁴, reported early disruption of the extracellular matrix. The biomarker discovery phase of this study was performed on just a small cohort of MPS I, II and VI samples as the MPS disorders are extremely rare so large cohorts of samples are impossible to collect. However, we were able to identify differentially expressed proteins which the primary biological pathway implicated was extracellular matrix organisation. This finding confirmed that markers of the underlying pathological features of this disease can be detected in patient urine. Some of the proteins reported in the mouse model were also detected in the present study (collagen type 1, COMP, and fibronectin). ECM proteins that were able to be confirmed, and could be used potentially to monitor ECM integrity in MPS I and II, are COMP, nidogen 1 and collagen type 1 (figure 3 & 4). Protein HEG1 homologue (Heart of Glass receptor 1) is an endothelial cell junction protein identified as elevated significantly in only the MPS VI disease group. HEG1 is crucial for heart and vessel formation and integrity¹⁵. Whilst the function of HEG1 has been investigated largely in the heart according to the protein expression database Model Organisation Protein Expression Database (MOPED)¹⁶ HEG1 protein expression is not limited to heart tissue. Therefore the origin of HEG1 in urine could be attributed to other tissues and its relevance to MPS VI pathology may also be indicative of impaired endothelial function which is a known pathological feature of MPS¹⁷. FABP5 which was elevated significantly in the severe MPS I and II groups (figure 3b), is a known epidermal and endothelial tissue protein found in heart and skeletal muscle but does not form part of the ECM. It has been postulated previously as a marker for cardio-metabolic risk factors and carotid atherosclerosis¹⁸. Coronary artery pathology including atherosclerosis and endothelial dysfunction are a prominent feature of the MPS disorders^{17,19}. Therefore, FABP5 and protein HEG1 have the poten-

tial to act as surrogate markers of monitoring for a patient's endothelial function. Beta galactosidase was elevated significantly in all MPS disease groups (figure 4c) and was also the most statistically significant and had the greatest fold changes. Beta galactosidase is involved in keratan sulphate catabolism and not heparan and dermatan sulphate catabolic pathways which are involved in the GAG catabolic pathways affected in MPS I, II and VI. The reason for increased levels of beta-galactosidase in MPS I, II and VI relative to the control group are unknown but lysosomal beta galactosidase is thought to be a marker of senescence and previous studies have linked its increased expression to an increase in lysosomal mass and or number²⁰⁻²². Lysosomal hypertrophy is known to occur in MPS disorders²³ therefore increased beta galactosidase may potentially be a marker for increased lysosomal load.

Not all of the markers described in the discovery analyses were confirmed in the same MPS groups. The reasons for this may be due to differences in sample numbers between the discovery and validation analyses using the targeted test. The greater sample numbers and different levels of disease severity in the larger groups may explain why there are significant yet different observations compared with the label free analysis. This emphasises the need to confirm initial biomarker findings in larger cohorts, since biomarker discovery is low throughput and confirmation of potential markers should be performed using high throughput targeted techniques. The accumulation of a large enough cohort of samples, due to the rarity of the conditions, would require a large multicenter study conducted over years, something not possible at this point.

The bottleneck and issues of translation of biomarkers into the clinical setting is well known^{24,25}. Targeted proteomics, particularly SRM/MRM LC-MS/MS analysis, is gaining more interest as it has great potential for clinical translation^{26,27} since most large hospitals have triple quadrupole based platforms. Potential assays for clinical routine analysis ideally would be simple in preparation and data handling/processing but also need to be, robust, reproducible and cost effective. The targeted assay platform we describe here has been designed purposely with these factors in mind, to enable its potential further development for use in future clinical trials using these markers. The assay only requires 1-2mls of urine, a generic cost effective whole protein standard is used to control for sample preparation and act as an internal standard. The assay is also potentially cost effective compared to other methods as it can multiplex the quantitative analysis of many proteins in one run and the LC conditions have been devised over a simple 10 min gradient which enables high through-put potential of the assay.

Conclusions: We have shown that several potential urine markers of MPS I, II and VI are detectable and quantifiable in urine. These markers likely indicate impaired extracellular matrix, endothelial function and lysosomal hypertrophy which are key pathological feature of MPS. Importantly, we have also been able to demonstrate a relationship between marker concentrations and disease severity. All markers for MPS did not show significantly raised levels in the clinically mild MPS I and II groups which were in the control range. However, three markers, (COMP, IGFBP7 and beta galactosidase) were able to differentiate between the MPS II neurological and non-neurological clinical phenotype. Further utility of these markers would be to determine severity of disease at initial diagnosis the patient's ages in the severe disease cohorts start at 0.26-

1 yrs which indicates this may be possible however larger numbers of samples taken at initial diagnosis need be assessed to appropriately conclude this. The ability to discriminate disease severity indicates these markers have the potential to determine the patient disease severity at initial diagnosis, which is often in early life and at the stage where initiation of treatment is crucial as well as reflect the effect of treatment on the downstream pathology of the disease.

ASSOCIATED CONTENT

Supporting Information

Label free proteomics protein identifications, Reactome Pathway analysis, peptide MRM chromatograms and standard curve are provided as supplementary data. "This material is available free of charge via the Internet at <http://pubs.acs.org>."

AUTHOR INFORMATION

Corresponding Author

Tel: 02079052873. Fax: 02074046191 E-mail: kev-in.mills@ucl.ac.uk.

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript. The authors declare no competing interests

ACKNOWLEDGMENT

This work was funded as part of the GOSomics platform by the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London, UCL Biological Mass Spectrometry Centre at the UCL Institute of Child Health and with kind donations from the Szeban Peto Foundation. We also wish to thank the Parents and Children of GOSH.

REFERENCES

- (1) Muenzer, J. *Rheumatology* **2011**, *50 Suppl 5*, v4-12.
- (2) Cimaz, R.; La Torre, F. *Current rheumatology reports* **2014**, *16*, 389.
- (3) Valayannopoulos, V.; Wijburg, F. A. *Rheumatology* **2011**, *50 Suppl 5*, v49-59.
- (4) Hurst, R. E.; Settine, J. M.; Lorincz, A. E. *Clinica chimica acta; international journal of clinical chemistry* **1976**, *70*, 427-432.
- (5) Auray-Blais, C.; Bherer, P.; Gagnon, R.; Young, S. P.; Zhang, H. H.; An, Y.; Clarke, J. T.; Millington, D. S. *Molecular genetics and metabolism* **2011**, *102*, 49-56.
- (6) Zhang, H.; Wood, T.; Young, S. P.; Millington, D. S. *Molecular genetics and metabolism* **2015**, *114*, 123-128.
- (7) Lehman, T. J.; Miller, N.; Norquist, B.; Underhill, L.; Keutzer, J. *Rheumatology* **2011**, *50 Suppl 5*, v41-48.
- (8) Manwaring, V.; Heywood, W. E.; Clayton, R.; Lachmann, R. H.; Keutzer, J.; Hindmarsh, P.; Winchester, B.; Heales, S.; Mills, K. *Journal of proteome research* **2013**, *12*, 2013-2021.
- (9) Bennett, K.; Callard, R.; Heywood, W.; Harper, J.; Jayakumar, A.; Clayman, G. L.; Di, W. L.; Mills, K. *Journal of proteome research* **2010**, *9*, 4289-4294.
- (10) Craig, R.; Cortens, J. P.; Beavis, R. C. *Journal of proteome research* **2004**, *3*, 1234-1242.
- (11) Mills, K.; Morris, P.; Lee, P.; Vellodi, A.; Waldek, S.; Young, E.; Winchester, B. *J Inherit Metab Dis* **2005**, *28*, 35-48.

- (12) Croft, D.; Mundo, A. F.; Haw, R.; Milacic, M.; Weiser, J.; Wu, G.; Caudy, M.; Garapati, P.; Gillespie, M.; Kamdar, M. R.; Jassal, B.; Jupe, S.; Matthews, L.; May, B.; Palatnik, S.; Rothfels, K.; Shamovsky, V.; Song, H.; Williams, M.; Birney, E.; Hermjakob, H.; Stein, L.; D'Eustachio, P. *Nucleic acids research* **2014**, *42*, D472-477.
- (13) Milacic, M.; Haw, R.; Rothfels, K.; Wu, G.; Croft, D.; Hermjakob, H.; D'Eustachio, P.; Stein, L. *Cancers* **2012**, *4*, 1180-1211.
- (14) Heppner, J. M.; Zaucke, F.; Clarke, L. A. *Molecular genetics and metabolism* **2015**, *114*, 146-155.
- (15) Kleaveland, B.; Zheng, X.; Liu, J. J.; Blum, Y.; Tung, J. J.; Zou, Z.; Sweeney, S. M.; Chen, M.; Guo, L.; Lu, M. M.; Zhou, D.; Kitajewski, J.; Affolter, M.; Ginsberg, M. H.; Kahn, M. L. *Nature medicine* **2009**, *15*, 169-176.
- (16) Higdon, R.; Stewart, E.; Stanberry, L.; Haynes, W.; Choiniere, J.; Montague, E.; Anderson, N.; Yandl, G.; Janko, I.; Broomall, W.; Fishilevich, S.; Lancet, D.; Kolker, N.; Kolker, E. *Journal of proteome research* **2014**, *13*, 107-113.
- (17) Kelly, A. S.; Metzigg, A. M.; Steinberger, J.; Braunlin, E. A. *J Inherit Metab Dis* **2013**, *36*, 221-225.
- (18) Yeung, D. C.; Wang, Y.; Xu, A.; Cheung, S. C.; Wat, N. M.; Fong, D. Y.; Fong, C. H.; Chau, M. T.; Sham, P. C.; Lam, K. S. *European heart journal* **2008**, *29*, 2156-2163.
- (19) Yano, S.; Moseley, K.; Wong, L.; Castelnovi, C.; Azen, C.; Pavlova, Z. *J Inherit Metab Dis* **2014**, *37*, 255-261.
- (20) Kurz, D. J.; Decary, S.; Hong, Y.; Erusalimsky, J. D. *Journal of cell science* **2000**, *113* (Pt 20), 3613-3622.
- (21) Lee, B. Y.; Han, J. A.; Im, J. S.; Morrone, A.; Johung, K.; Goodwin, E. C.; Kleijer, W. J.; DiMaio, D.; Hwang, E. S. *Aging cell* **2006**, *5*, 187-195.
- (22) Winchester, B. *J Inherit Metab Dis* **2001**, *24 Suppl 2*, 52-56; discussion 45-56.
- (23) Van Hoof, F. *Journal of clinical pathology. Supplement* **1974**, *8*, 64-93.
- (24) Witkowska, H. E.; Hall, S. C.; Fisher, S. J. *Clinical chemistry* **2012**, *58*, 321-323.
- (25) Diamandis, E. P. *Journal of internal medicine* **2012**, *272*, 620.
- (26) Percy, A. J.; Chambers, A. G.; Yang, J.; Hardie, D. B.; Borchers, C. H. *Biochimica et biophysica acta* **2014**, *1844*, 917-926.
- (27) Marx, V. *Nature methods* **2013**, *10*, 19-22.

