IDENTIFICATION OF A SERUM BIOMARKER FOR
MUCOPOLYSACCHARIDOSIS I

by

DERRICK RAYMOND RANDALL

B.Sc.H., The University of British Columbia, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Masters of Science

in

THE FACULTY OF GRADUATE STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

April 2006

© Derrick Raymond Randall, 2006
Abstract

The mucopolysaccharidoses are a clinically heterogeneous group of lysosomal storage disorders presenting with broad multi-system disease and a continuous range of phenotypes. Currently there are no objective biomarkers of MPS disease that clearly reflect disease severity or therapeutic responsiveness. Using proteomic studies in the murine MPS I model, I have identified the formation of the heparin cofactor II-thrombin (HCII-T) complex, a well-known serine protease inhibitor (serpin)-serine protease complex, as an informative biomarker for MPS I. MPS I patients showed a range of serum HCII-T concentrations from 16,300 – 208,600 pM, whereas the control values varied from 38.94 – 1491 pM. HCII-T complex was also elevated in plasma from MPS I patients and mice. The degree of HCII-T complex formation appears to correlate with disease severity and is responsive to therapy. In addition to its role as a biomarker, the discovery of increased serpin-serine protease complex formation provides a valuable insight into possible pathophysiological mechanisms of MPS disease.
List of Tables

Table 2.1: Protein concentrations of high abundance protein depleted pools ..................24
Table 3.1: iTRAQ summary data ..................................................................................36
Table 3.2: Number of unique peptides per protein in iTRAQ data collection, ≥94% 
  Protein Confidence .................................................................................................37
Table 3.3: Proteins identified at ≥99% confidence with significantly altered relative 
  abundance in the MPS mouse serum, *Idua*+/+ mixed sex denominator.................41
Table 3.4: HCII-T ELISA in serum and plasma samples of MPS I patients, MPS I 
  mice, and controls .................................................................................................45
Table 3.5: Known serpins with GAG-mediated activities ............................................51
Table 4.1: HCII-T complex levels in patients undergoing ERT ..................................61
Table A1.1: Effects of iTRAQ variability on abundance ratios ....................................96
Table A2.1: Mucopolysaccharidoses and the respective enzyme deficiency and stored 
  glycosaminoglycan .................................................................................................107
Table A2.2: Examples of serpins with GAG-mediated activities .................................112
Table A2.3: iTRAQ Summary Data ..............................................................................117
Table A2.4: Number of unique peptides per protein in iTRAQ data collection, 94% 
  Protein Confidence ...............................................................................................117
Table A2.5: Proteins with significantly altered relative abundance in the MPS mouse 
  serum, *Idua*+/+ mixed sex denominator .............................................................118
Table A2.6: HCII-T ELISA ..........................................................................................121
List of Figures

Fig. 1.1: Lysosomal storage disorders .................................................. 3
Fig. 1.2: Iduronidase action in dermatan sulphate degradation ................ 4
Fig. 1.3: Iduronidase action in heparan sulphate degradation ................. 5
Fig. 1.4: The iTRAQ reagent and principle ........................................... 10
Fig. 1.5: iTRAQ data output ............................................................. 11
Fig. 2.1: 2DGE images of nascent and high abundance protein depleted mouse serum samples ................................................... 24
Fig. 2.2: 2DGE images of MPS and wild-type mouse depleted serum samples ............................................................... 25
Fig. 3.1: Box plots of relative abundances for proteins identified with 99% confidence ................................................................. 38
Fig. 3.2: Logarithmic plot of average relative abundance of proteins in Idua<sup>−/−</sup> pools vs. Idua<sup>+/+</sup> mixed sex, ≥99% confidence, with proteins rank ordered from left to right based on descending Protein % Confidence score .......... 40
Fig. 3.3: Western blot of serum samples from WT (Idua<sup>+/+</sup>) and MPS (Idua<sup>−/−</sup>) mice for HCII .......................................................... 42
Fig. 3.4: Western blot of MPS patients for HCII ...................................... 43
Fig. 3.5: HCII-T levels in MPS 1H serum samples following enzyme replacement therapy .......................................................... 44
Fig. 3.6: Western blot of MPS patient samples for ATIII ......................... 46
Fig. 4.1: Control range of serum HCII-T concentration .......................... 59
Fig. 4.2: Serum HCII-T concentrations for MPS I patients undergoing a randomized, double-blind, placebo-controlled ERT trial ..................... 60
Fig. 4.3: Serum HCII-T complex concentration over a Hurler patient’s treatment course ......................................................... 62
Fig. 4.4: Serum HCII-T complex concentration and antibody titre over an attenuated patient’s treatment course .................................... 63
Fig. 5.1: Serine protease inhibitor action ................................................ 80
Fig. A1.1: Individual protein ratios, expressed as absolute values of the fold-differences between the Idua<sup>−/−</sup> male:Idua<sup>−/−</sup> mixed sex ratio ..................................................................... 94
Fig. A1.2: Change in Unused Protein Score from iTRAQ Acquisition A to Acquisition A+B ................................................................. 95
Fig. A1.3: Change in Unused Protein Score versus Unused Protein Score from iTRAQ Acquisition A ...................................................... 98
Fig. A2.1: Box plots of relative abundances for proteins identified with 99% confidence ................................................................. 123
Fig. A2.2: Logarithmic plot of average relative abundance of proteins in Idua<sup>−/−</sup> pools vs. Idua<sup>+/+</sup> mixed sex, ≥99% confidence, with proteins rank ordered from left to right based on descending Protein % Confidence score .......... 124
Fig. A2.3: Western blot of serum samples from WT (Idua<sup>+/+</sup>) and MPS (Idua<sup>−/−</sup>) mice for HCII .......................................................... 125
Fig. A2.4: Western blot of MPS patients for HCII ...................................... 126
Fig. A2.5: HCII-T levels in MPS 1H serum samples following enzyme replacement therapy .......................................................... 127
Fig. A2.6: Western blot of MPS patient samples for ATIII ......................... 128
List of Abbreviations

2DGE – 2-dimensional gel electrophoresis
ATIII – antithrombin III
ATIII-T – antithrombin III-thrombin complex
BMT – bone marrow transplantation
CHAPS – 3-[(3-cholamidopropyl)dimethyammonio]-1-propanesulphonate
CS – chondroitin sulphate
Da – daltons
DIGE – differential 2-dimensional gel electrophoresis
DS – dermatan sulphate
DTT – dithiothreitol
EC – enzyme catalogue
ECM – extracellular matrix
ERT – enzyme replacement therapy
fmol – femtomoles (10^{-12} moles)
GAG – glycosaminoglycan
GalNAc – N-acetylgalactosamine
HCII – heparin cofactor II
HCII-T – heparin cofactor II-thrombin complex
HPLC – high performance liquid chromatography
HS – heparan sulphate
ICAT – isotope coded affinity tags
IDUA – human α-L-iduronidase protein
Idua – murine α-L-iduronidase protein
Idua – murine α-L-iduronidase gene
IEF – isoelectric focusing
IgG – immunoglobulin gamma
iTRAQ – isobaric tagging for relative and absolute quantification
kDa – kilodaltons (10^{3} Da)
KS – keratan sulphate
MPS – mucopolysaccharidosis
MPS I – mucopolysaccharidosis type I (general term)
MPS IH – mucopolysaccharidosis type I (Hurler syndrome)
MPS IH/S – mucopolysaccharidosis type I (Hurler-Scheie syndrome)
MPS IS – mucopolysaccharidosis type I (Scheie syndrome)
MS – mass spectrometry
OMIM – Online Mendelian Inheritance in Man
serpin – serine protease inhibitor
UPS – Unused Protein Score
WT – wild-type
Acknowledgements

I would like to acknowledge the people who have been important influences during the course of my studies. First, I want to thank my supervisor Lorne Clarke for taking me into his lab and teaching me an incredible amount in only 20 months. Lorne has been more supportive and inspiring than I could possibly imagine, and I have grown considerably on account of his leadership. Jan Friedman and Robert McMaster provided excellent suggestions and guidance during my research, having me consider questions I did not think to ask. Graham Sinclair, as much mentor as former lab member, offered useful technical suggestions and guidance in a number of situations that helped me to succeed. I could never forget to acknowledge Karen Colobong, from the preparative work and sample collection she does for me, to being there every day in the lab to talk and share experiences. Cheryl Bishop: she was always on my side and curious about my goals. My parents are very important people to me and I want to recognize how they were always supportive and interested in what I did, ensuring that I was able to explain the details so that an outsider would understand. Finally, I want to thank my wife Kym for always knowing that this was the best idea, and for her encouragement every day.
Co-Authorship Statement

This thesis is the product of research performed entirely by me during my Master of Science program, all manual work and data analysis inclusive, except where noted in the acknowledgements. The initial experimental design was planned by Dr. Lorne Clarke, Dr. Graham Sinclair, and me. All text in this thesis is original work written by me, with editing contributions from Drs. Clarke and Sinclair in Chapter 3.
Chapter 1: Introduction

1.1 Thesis Focus and Chapter Overview

This thesis describes the identification of a serum biomarker for mucopolysaccharidosis I (MPS I) through proteomic analysis in the MPS I murine model, and its validation in mouse and human serum samples. MPS I is a progressive, multisystem disease with a spectrum of clinical phenotypes ranging from a severe form with onset in infancy leading to death by the end of the first decade of life, to an attenuated form presenting later in life that is associated with morbidity but usually a normal life span. Many patients fit within this broad spectrum. This clinical heterogeneity leads to differences in progression for individual patients as well as difficulty evaluating therapeutics. Therefore the development of an objective biomarker of MPS disease would be clinically helpful. Identifying a reliable, accurate biomarker has the potential to greatly improve patient care and treatment. Using the murine model of MPS I, I identified a serum biomarker of the disease, validated it in mice and humans, and measured its responsiveness to enzyme replacement therapy.

Chapter 1 is a brief introduction describing the disease and issues surrounding current biomarkers. The proteomic techniques involved in this thesis follow, with a discussion of strengths and weaknesses of the different methods. There is also a section on the obstacles associated with analyzing the serum proteome without a priori knowledge of possible biomarkers. Finally, the thesis objective is stated with the underlying hypothesis and rationale for the project.
Chapter 2 is a presentation of the utility of high abundance serum protein depletion and 2-dimensional gel electrophoresis images obtained over the course of my research. The success of the methodology and a putative biomarker are briefly evaluated.

Chapter 3 is a manuscript of a published paper that describes the identification of the biomarker and its validation in both murine and human samples. A discussion of the implications of the biomarker and insights into MPS pathophysiology is included.

In Chapter 4, I produce a manuscript containing results of a subsequent assessment of the biomarker with a larger set of MPS I patients enrolled in an enzyme replacement therapy clinical trial, as well as data from two MPS I patients not enrolled in the trial but with a larger number of sample collections, thereby allowing more detailed understanding of the therapeutic response.

Chapter 5 presents a discussion of the results I obtained and their application to MPS research. A critical evaluation of the iTRAQ technology used in this study and its application to proteomic research follows. I also describe future topics to address through related studies.

Finally, two appendices describe additional work relevant to this thesis that I have performed but not included in the manuscripts: a computational evaluation of iTRAQ reproducibility, as well as a patent submitted on the measurement of the biomarker for diagnostic purposes.

1.2 Mucopolysaccharidosis I

Mucopolysaccharidosis I (MPS I, OMIM #252800) is an autosomal recessive disorder with heterogeneous clinical presentation. A member of the broad family of
lysosomal storage disorders, MPS I is one of seven classes of MPS disorders, which can be subdivided into thirteen subcategories (Fig. 1.1). The MPS category encompasses all the disorders related to the catabolism of mucopolysaccharides, otherwise known as glycosaminoglycans (GAGs). Representing the “glyco” part of glycoprotein, GAGs are linear carbohydrate chains attached to proteins, and they range in size and complexity from a few monomers in a single chain to tens of thousands of monomers in multi-branched appendages. In some instances, the mass of the carbohydrate chains exceed that of the protein. Glycoproteins are involved in many cellular systems—signalling, adhesion, extracellular matrix structure, ligand binding—and the carbohydrate moieties of a glycoprotein can change throughout development [1]. Therefore, glycoproteins regularly enter the protein degradation pathway for recycling.

![Fig. 1.1: Lysosomal storage disorders](image)

The mucopolysaccharidoses are one of several lysosomal metabolic diseases. MPS disorders involve the breakdown of one or more of the glycosaminoglycans chondroitin sulphate (CS), dermatan sulphate (DS), heparan sulphate (HS), keratan sulphate (KS), or hyaluronan.
MPS I patients are deficient in the enzyme α-L-iduronidase (IDUA), an exoglycosidase responsible for degrading GAGs in the lysosome. Specifically, IDUA is a lysosomal hydrolase essential for the step-wise degradation of dermatan sulphate (DS) and heparan sulphate (HS) by cleaving the terminal, unsulphated iduronic acid residue (Figs. 1.2 and 1.3). DS is a polymer of β-linked iduronic or glucuronic residues linked by alternating 4-O-sulphated N-acetylgalactosamine (GalNAc) monomers. DS may also be sulphated at the C-2 position of the uronic acid (but this is cleaved by the preceding enzyme in the degradative pathway) and is a common GAG involved in structural elements such as bone and cartilage. HS is a higher charge-density GAG than DS and consists of the same uronic acids alternated by α-linked glucosamine monomers. The increased charge density occurs because HS can carry sulphate groups on the uronic acid, and through N-, C-3, and C-6 linkages on glucosamine. This additional heterogeneity allows HS to mediate many different protein interactions, and, as might therefore be expected, HS is more widespread than DS.

Fig. 1.2: Iduronidase action in dermatan sulphate degradation
α-L-iduronidase catalyzes the cleavage of the terminal iduronic acid in dermatan sulphate chains.
Storage of these undegraded GAGs in the lysosomes results in the remarkable phenotypes associated with MPS I, characterized by excretion of partially degraded DS and HS in the urine. Children with the severe form of MPS I, known as Hurler disease (MPS IH), have a normal appearance at birth and typically show phenotypic signs between 6 – 18 months of age. The disease shows relentless progression and patients have a life expectancy of 10 years. It is a multi-system disorder with developmental delay, arthropathy, cardiac and respiratory abnormalities, dysostosis multiplex, hepatosplenomegaly, short stature, and corneal clouding. Respiratory and cardiac complications cause the majority of patient deaths.

Approximately 20% of MPS I patients present with an attenuated phenotype. In some cases, where onset usually occurs in adolescence, the symptoms may be limited to joint stiffness, cardiac valvular dysfunction, and corneal clouding. These patients are
classified as Scheie patients (MPS IS) and often have normal life spans with normal intelligence and average height. Many patients fit a clinical phenotype between that of Hurler syndrome and Scheie syndrome in the aptly named Hurler-Scheie syndrome (MPS IH/S); these patients have an attenuated phenotype which fills the broad and continuous spectrum from MPS IH to MPS IS. Onset of symptoms can occur in childhood with progressive arthropathy and cardiac involvement, as well as CNS and peripheral nervous system involvement. Untreated, most of these patients die within the third decade. Molecular studies of mutations at the IDUA locus indicate that residual enzyme activity is the most likely factor underlying this clinical heterogeneity [2-4]. However, residual enzyme activity and catalytic capacity measured with artificial substrates from severe and attenuated patients show overlapping ranges [5].

Determining the phenotypic severity of patients diagnosed with MPS I early in their clinical course is challenging. Clinical criteria gained from extensive clinical evaluation are helpful in the determination of the degree of disability or clinical effects in individual patients. Nonetheless, these features are subjective, limited in utility for short-term evaluation of treatment or intervention effects, and not easily translatable to different physicians or care centres. The current gold standard assay to diagnose a patient with MPS I is to measure IDUA enzyme activity in leukocytes or fibroblasts [6]. However, this measure does not delineate the different clinical phenotypes [6,7] nor assess therapeutic response. Urinary GAG concentration shows reduction following therapy [8] and is marginally useful as a preliminary screen, but it is not highly specific for MPS patients since elevated urinary GAGs are observed in many healthy, young children [9-12]. In addition, whether the urinary GAGs reflect a general level of disease
involvement among the organs or simply that of the kidneys is an unresolved issue. Although some genotype-phenotype correlation does exist—homozygotes or compound heterozygotes for the prevalent W402X or other nonsense mutations always present with Hurler syndrome [13]—many families have private mutations for which phenotype cannot be accurately predicted from genotype [13]. Combinations of nonsense and missense mutations produce unpredictably variable residual activity levels in in vitro studies [5,6,13].

It is well accepted that tissue GAG accumulation underlies the pathological consequences of MPS disorders, but there is a generally poor understanding of the mechanisms behind the pathophysiology of this complex disease. Precisely how GAGs cause the progression is unknown; one of the few clues is the presence of DS in MPS disorders with skeletal deformities and HS in those with neurological involvement. Further complexity arises from the observation that secondary build-up of metabolites, such as G_{M2} and G_{M3} gangliosides, occur in the central nervous system of MPS I patients who have neurological involvement [14,15]. How these secondary molecules accumulate remains to be answered, as does their role in the CNS symptoms [6].

Identification of a serum biomarker would be very useful for MPS I clinical care and research. With the advent and licensing of recombinant human enzyme replacement therapy (ERT, Aldurazyme™ [8]), there is hope for many patients suffering with the disease and their families that treatment will mitigate the devastating effects of the disease. Thus, a biomarker would have potential use in monitoring clinical severity from an objective, measurable standpoint and optimizing individual dosages. In addition, while ERT does not provide a cure for MPS I, it is accepted that early intervention may
dampen the impact of the symptoms in individuals by replacing the metabolic deficiency prior to onset of irreversible damage [16-18].

1.3 Techniques for Proteomic Analysis

Understanding the different tissue proteomes has become a goal of many human disease researchers. Whereas the genome predicts potential gene products, the complete proteome is a detailed snapshot of the protein compliment of a given state. By developing proteome maps of different tissues under discrete conditions, researchers should be able to identify abnormally regulated proteins and post-translational modifications undetectable through genome analysis.

Ideally, the entire proteome would be analyzed, with no biases or assumptions surrounding constituent proteins. However, technical aspects to be described hinder this in many proteome analyses. Although some success has been achieved with studies that select large subsets of proteins for analysis [19-24], such studies are limited in capacity to identify unanticipated, novel results. Therefore, it remains preferable to use unbiased techniques, particularly in disorders such as MPS I where hypothesis-based research has not provided solutions to disease progression.

The original proteomic technique, two-dimensional gel electrophoresis (2DGE), separates proteins based on isoelectric point, then by molecular weight. This method has been improved since its inception three decades ago: replicability has been improved through new systems, buffer and desalting optimization minimizes horizontal streaking to improve clarity, and loading limits have increased to enhance visualization of low expression level proteins. Nonetheless, the methods are cumbersome and require
multiple, complex steps to arrive at a protein identification. Recently, techniques relying on mass spectrometry and database searching have become increasingly prevalent; however, these methods are insensitive to post-translational modifications [25] and so 2DGE still provides a useful tool for proteome analysis.

Differential mass spectrometry-based methods are improving in power and their ability to recognize valid differences between proteomes, evidenced by recent results [26-32]. One of the original methods, Isotope Coded Affinity Tag (ICAT) reagents, relies on the principle of differentially labeling two pools of trypsinized proteins (each constituting a different proteome) with cysteine-reactive tags: each pool labeled with different isotopic tags [33]. This permits both peptide pools to be combined and simultaneously analyzed by mass spectrometry to determine the relative abundances of peptides (and therefore parent proteins) in each proteome [33]. However, ICAT analysis is restricted to proteins containing cysteine amino acids, and can only identify the select peptides from proteins possessing cysteines. With the average 50 kDa, 450 amino acid, vertebrate protein containing 13 cysteinyl residues (compared to 35 alanyl residues) [34], and many proteins and peptide hormones holding none, the observed proteome is clearly incomplete. Furthermore, few corroborating peptides cause low confidence abundance ratios.

A new generation of differential labeling, isobaric Tagging for Relative and Absolute Quantitation (iTRAQ), has emerged with the potential for comprehensive proteome analysis. iTRAQ uses four different non-isotopic isobaric tags, all with a total mass of 145 Da; each isobaric tag contains a reporter tag of one of 114, 115, 116, or 117 Da, and a corresponding balance tag of 31, 30, 29, or 28 Da, respectively [35]. As a
result, four unique protein pools can be tested concurrently for protein abundance variations. When fragmented during MS, the reporter tags appear in a quiet region of a typical peptide MS profile and can be unambiguously identified and quantified (Fig. 1.4).

![Diagram of iTRAQ reagent](image)

**Fig. 1.4: The iTRAQ reagent and principle**

a) Differentially labeling protein mixtures with isobaric tags containing reporter tags with unique masses allows identification of peptides and their abundances by mass spectrometry. Adapted from [36]. b) The protein abundances are calculated from the abundances of the reporter tags in a second MS scan. From [36].

A key feature of iTRAQ reagents is that they bind to peptides through an amine-reactive group. Theoretically, every lysyl residue and N-terminus will carry an iTRAQ tag, so every peptide in the reaction mixture becomes identifiable and quantifiable. After trypsin-digested proteins are labeled with a specific iTRAQ reagent, the samples are pooled, separated by liquid chromatography on hydrophobicity, and subjected to tandem mass spectrometry (LC-MS/MS). One round of MS identifies peptides, and the second identifies and quantifies the reporter tags (Fig. 1.4). Since the pooled peptides enter the
ionization chamber at the same time, all identical peptides are fragmented during the same window, which increases sample quantity and the likelihood of correct peptide identification. The observed peptide fragmentation patterns are compared to patterns in a database to identify the peptide sequence. An automated database search then provides a probabilistic identification of the parent protein, with confidence values based on the number of peptides identified and their fit to spectra in the database [35,37].

After compiling a list of identified proteins, an abundance value is calculated for each protein as a weighted average of the abundances of the reporter tags measured in the same fragmentation window as the peptides to which they were bound. This protein abundance permits calculation of a ratio for each protein in one pool relative to the corresponding protein in the other three pools, along with a protein identification probability and ratio error factor. This is presented in colour-coded tabular format indicating proteins with altered abundances (Fig. 1.5). In order to account for small changes in total protein concentration due to sample handling and labeling efficiency, each protein abundance is multiplied by a correction factor based on the ratio of all strongly identified proteins.

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>116:114 Ratio</th>
<th>P.Val</th>
<th>EF</th>
<th>116:114 Ratio</th>
<th>P.Val</th>
<th>EF</th>
<th>117:114 Ratio</th>
<th>P.Val</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor XII</td>
<td>0.8342</td>
<td>0.0044</td>
<td>1.127</td>
<td>0.927</td>
<td>0.2659</td>
<td>1.1471</td>
<td>0.8803</td>
<td>0.0386</td>
<td>1.1279</td>
</tr>
<tr>
<td>Fibrinogen, gamma polypeptide</td>
<td>5.5328</td>
<td>0.0196</td>
<td>3.6672</td>
<td>0.9048</td>
<td>0.2187</td>
<td>1.2006</td>
<td>4.4442</td>
<td>0.0337</td>
<td>3.7483</td>
</tr>
</tbody>
</table>

**Fig. 1.5: iTRAQ data output**

All identified proteins are listed with their observed ratios and ratio error factors. Shaded boxes indicate the ratio is truly different from the denominator pool based on protein identification probability.

Limitations of the iTRAQ methodology are that sample preparation could cause artifactual changes in individual proteins that would not be corrected by a global protein correction factor. The colour-coding system is not based on the probability a protein
ratio truly deviates from equal concentrations but rather that the protein identification is correct. Thus, an observed ratio of 0.9992 could be deemed significant by the program but likely carries no biological value because the difference is within the natural variation seen between individuals.

1.4 The Serum Proteome’s Technical Challenge

As convenient as it is to acquire, serum presents a major challenge to proteomic analysis: a steep, sigmoidal protein concentration dynamic range [38]. Although high abundance proteins impair all tissue proteome analyses, the majority of the serum protein content is composed of three proteins: albumin, IgG, and transferrin; a second tier of eight high abundance proteins also exists [38]. These high abundance proteins impede identification of the low abundance proteins through 2DGE because protein loading limits accommodate limited visualization of low abundance proteins; low abundance proteins that can be detected are not likely to have the requisite 75 fmol for identification [39]. MS is similarly impaired because duty cycles of the analytical equipment recognize a portion of the peptides in each window: ergo, the most highly abundant proteins’ peptides will present the strongest signals and be more frequently identified [40]. It is unlikely these major constituents represent informative biomarkers, so the potentially interesting proteins—which comprise approximately 10% of the total serum protein content [41-43]—need to be identified beneath the high abundance protein pattern.

One available method to circumvent this problem is protein depletion, where the high abundance proteins are removed from serum samples before analysis. A number of methods are available: antibody affinity, albumin-specific reagents, and albumin-specific
desalting techniques [44], to name a few. However, an important requirement is that depletion be highly efficient and specific, avoiding concomitant loss of potentially valuable low abundance proteins. Antagonistically, albumin and transferrin biologically function as carrier proteins to stabilize and shuttle other proteins through circulation. Therefore, a degree of specific and non-specific protein loss inevitably results from all methods [45,46].

Fortunately, the potential value of the serum proteome prompted several companies to produce multiple antibody affinity columns for HPLC systems. These columns contain antibodies specific for several of the high abundance proteins and are optimized to elute both low abundance proteins and proteins bound nonspecifically to the carrier proteins. The efficiency and reproducibility of these columns has been recently reviewed [47].

1.5 Murine Model of MPS IH

Our lab developed a murine model of severe MPS I through a targeted knock-out of the Idua gene [48]. Homozygous null mice recapitulate most symptoms of human MPS IH and are excellent models for MPS disease research [6,49,50].

Using a mouse model for the biomarker discovery phase presents several advantages over using human samples. First, the low incidence of MPS I (1/100,000 [51-53]) makes collection of sufficient samples to identify a biomarker applicable to other patients difficult, particularly with the advent of ERT, as biomarkers will probably not be apparent in post-treatment samples. Validation of candidate biomarkers in large, independent samples would be nearly impossible in humans, but it is easily achievable in
a mouse colony. Finally, due to the level of genetic and environmental heterogeneity between people it is difficult to avoid inter-individual variability causing spurious results in humans, but collecting samples from genetically homogeneous, identically raised mice minimizes this inherent variability.
1.6 Thesis Objective and Hypothesis

1.6.1 Objective

To identify a serum biomarker indicative of MPS I disease.

1.6.2 Hypothesis

The primary block of glycosaminoglycan degradation in MPS I will lead to altered levels of proteoglycans. These proteoglycans may be useful as disease biomarkers.
1.7 References


18


Chapter 2: Serum 2-Dimensional Gel Electrophoresis Studies in the MPS I Murine Model

2.1 Introduction

Since its inception in 1975 [1], 2-dimensional gel electrophoresis (2DGE) has been used as a preparative and analytical tool in protein and proteome studies. As the original proteomic analysis method, 2DGE has proven useful for identifying changes in protein concentrations and post-translational modifications in many diseases and been the principal technique for developing tissue proteome maps [2]. In light of criticism surrounding replicability and utility in resolving basic proteins, hydrophobic proteins, as well as those of low molecular weight [2,3], considerable attention has focused on strengthening the technique through the use of new technologies. Indeed, 2DGE remains the first method routinely used in most proteomic studies and the preferred unbiased technique to identify post-translational modifications in the proteome [3,4]. Nonetheless, the low resolution of 2DGE, a product of loading limits and staining methods, limits the technique’s potential to identify informative biomarkers and to characterize low abundance proteome constituents; this is especially true in plasma and serum, which promise to hold a wealth of informative, medically relevant biomarkers. However, due to the dynamic differences in protein concentrations between common, uninformative proteins and the low abundance, informative ones, identifying these biomarkers requires simplification of the proteome.

Mucopolysaccharidosis I (MPS I; OMIM #252800) is a multi-system disorder, characterized by a spectrum of phenotypes ranging from severe disease, known as Hurler
syndrome, to the most attenuated form, Scheie syndrome [5]. Patients with Hurler syndrome commonly present in the first year of life and, when untreated, have a life expectancy of 10 years with a progressive multi-system disease with developmental delay, arthropathy, cardiac and respiratory abnormalities, dysostosis multiplex, hepatosplenomegaly, and corneal clouding. The remainder of patients have an attenuated phenotype which covers a broad and continuous spectrum ranging from onset of symptoms in late adolescence with normal life expectancy and little or no associated arthropathy, cardiac involvement, or direct CNS involvement, to those individuals with onset in childhood and more significant systemic involvement, often leading to death within the second decade. The primary metabolic defect in MPS I is deficiency of the lysosomal enzyme \( \alpha-L \)-iduronidase (IDUA; EC 3.2.1.76). IDUA deficiency results in abnormalities in the degradation of the glycosaminoglycans (GAGs) heparan sulphate and dermatan sulphate, which subsequently accumulate in the lysosome. Genotype to phenotype correlation does exist to some extent for IDUA deficiency, however the large number of private mutations does not allow genotype to predict phenotype in a considerable proportion of patients [6,7]. Currently, no disease biomarkers unambiguously allow for the determination of disease severity or responsiveness to treatment. As such, identification of a serum biomarker would be invaluable to MPS I diagnostics and research.

In this project, I depleted three high abundance proteins from pooled serum of \( Idua^+/- \) mice, the murine model for MPS I developed by our lab [8], and wild-type mice. The remaining serum proteins were analyzed by 2DGE in an attempt to discover a serum
biomarker of MPS I. No candidate biomarkers were isolated and identified by this method.

2.2 Methods

2.2.1 Protein quantification

Immediately following high abundance protein depletion, samples containing the pooled flow-through fractions were quantified for total protein using the Bio-Rad Dc protein assay to confirm similar total protein contents. An eight point standard curve was developed with bovine serum albumin (Bio-Rad, Hercules, CA) and all samples were tested in triplicate.

2.2.2 Acetone precipitation

Samples subjected to high abundance protein depletion required desalting prior to 2DGE. Acetone (Fisher, Fair Lawn, NJ) was pre-chilled to -20°C. 300 μL of each protein sample was mixed with 1200 μL of cold acetone, vortexed, and immediately placed at -20°C for 90 minutes. Samples were then centrifuged at 15,000g. The supernatant was discarded and the pellet washed once with cold acetone. Samples were air dried in a fumehood at room temperature for 30 minutes.

2.2.3 2D Gel Electrophoresis

Acetone precipitated samples containing a predicted protein content of 170 μg were solvated in 200 μL of rehydration buffer containing 7 M urea (Invitrogen, Carlsbad, CA), 2 M thiourea (Invitrogen), 2% CHAPS (Invitrogen), 0.5% carrier ampholytes
(Invitrogen), 0.002% bromophenol blue (Fisher), and 20 mM DTT (Sigma, St. Louis, MO). 45 μL of this sample (38 μg protein) was diluted with 120 μL of rehydration buffer. 155 μL (36 μg) was applied to rehydrate each ZOOM 3 – 10 non-linear IPG strip (Invitrogen) overnight.

Strips were subjected to step gradient voltage changes for the following time periods: 200 V for 20 minutes, 450 V for 15 minutes, 750 V for 15 minutes, 2000 V for 45 minutes.

Following IEF, strips were immediately reduced with NuPAGE LDS Sample Buffer (Invitrogen) containing NuPAGE Sample Reducing Agent (Invitrogen) for 15 minutes. The reducing buffer was decanted and the strips were alkylated with NuPAGE Sample Buffer containing 125 mM iodoacetamide (Sigma).

Reduced, alkylated strips were subjected to 133 Vh at 200 V on NuPAGE 4 – 12% Bis-Tris IPG gels (Invitrogen). Gels were fixed for 2 x 15 minutes with 50% methanol, 7% acetic acid, then stained overnight at room temperature with Sypro Ruby (Invitrogen). Gels were washed for 75 minutes with 10% methanol, 7% acetic acid, then for 3 x 5 minutes with ddH₂O and visualized at 302 nm. Images were taken with AlphaImager software, v2.0 (Alpha Innotech, San Leandro, CA).

2.3 Results

2.3.1 Protein quantification

Table A1.1 shows the column’s replicability and the amount of protein albumin, IgG, and transferrin contribute to the serum proteome.
Table 2.1: Protein concentrations of high abundance protein depleted pools

<table>
<thead>
<tr>
<th>Sample</th>
<th>Pre-depletion protein content (mg)</th>
<th>Post-depletion concentration (mg/mL)</th>
<th>Post-depletion protein content (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idua(^{+/+}), mixed sex</td>
<td>4.26</td>
<td>0.53</td>
<td>1.19</td>
</tr>
<tr>
<td>Idua(^{+/+}), male</td>
<td>4.26</td>
<td>0.58</td>
<td>1.30</td>
</tr>
<tr>
<td>Idua(^{-/-}), mixed sex</td>
<td>4.26</td>
<td>0.59</td>
<td>1.34</td>
</tr>
<tr>
<td>Idua(^{-/-}), male</td>
<td>4.26</td>
<td>0.56</td>
<td>1.27</td>
</tr>
</tbody>
</table>

2.3.2 2D gel electrophoresis

Fig. A1.1 shows 2DGE images obtained with and without high abundance protein depletion. The use of high abundance protein depletion clearly removes the major albumin spot and enhances visibility of low abundance proteins.

Fig. 2.1: 2DGE images of nascent and high abundance protein depleted mouse serum samples

a) 2DGE of 40 µg undepleted serum. Note the large spot at pH 5.2 representing albumin.

b) 2DGE of 38 µg high-abundance protein depleted serum. The major albumin spot appears entirely absent and many previously hidden spots are now visible.

Fig. 2.2 shows 2DGE images obtained for the four sample pools of high abundance protein depleted serum. With the exception of one small spot indicated, there
does not appear to be any consistent, major change to the MPS I mouse serum proteome.

This spot was of insufficient size for successful protein identification.

Fig. 2.2: 2DGE images of MPS and wild-type mouse depleted serum samples
a) Idua<sup>−/−</sup> mixed sex pool. b) Idua<sup>+/+</sup> mixed sex pool. c) Idua<sup>−/−</sup> male pool. d) Idua<sup>+/+</sup> male pool. The white arrow indicates a small spot that was present in both MPS mouse pools but not in the WT pools. All samples contained 38 μg of protein.

2.4 Discussion

My results indicate there are no major visible changes in the serum proteome to distinguish between serum from MPS I and normal mice. I observed one minor variation in the proteome but was unable to identify the candidate due to an insufficient quantity of protein. Unfortunately, increasing the sample load produced unclear gels that could not
be used to isolate more protein. Without means to validate the observed up-regulation in this unidentified protein's concentration, it is difficult to confirm the change. Occasionally, small spots arise as carbamylation by-products due to poor sample handling allowing iso-cyanate ions to bind to lysyl side chains and alter the pH of the protein [9], giving the impression the protein is unique from its neighbours. However, given the location of the protein in question and the pattern of the more acidic major serum protein, this does not appear to be a carbamylation artefact. Furthermore, reproduction in both Idua<sup>-/-</sup> pools in conjunction with the absence of the spot in either wild-type pool argues against this being a chance by-product.

Despite these results, there are likely a number of potential biomarkers that may be useful in MPS I research, but they remain undetectable by this method. In addition, although depletion of albumin, IgG, and transferrin greatly enhanced the number of proteins visualized in my experiments, it is insufficient to identify a biomarker of MPS I as a result of the dynamic concentration range of the remaining serum proteins. A possible method to enhance the resolution of the low abundance proteins is to perform further prefractionations of the samples by isoelectric focusing, which increases the sample loading limits and focuses on definable pH windows. Preliminary results have shown this to be a viable option but hindered by high salt concentrations following protein depletion, which interferes with isoelectric focusing and obscures visualization through horizontal streaking. Reproducibility is another related issue interfering with further prefractionations as each step carries an error factor in the proteins partitioned into each fraction.
On a global level, my 2DGE experiments indicate there is no major change in the MPS I serum proteome. Further reductions of the proteome may yield informative biomarkers that explain aspects of the clinical heterogeneity and pathophysiology seen in MPS I patients, but the technological limitations of 2DGE are too great to identify these factors.
2.5 References


28
Chapter 3: Heparin Cofactor II-Thrombin Complex in MPS I: A Biomarker of MPS Disease*

3.1 Introduction

Mucopolysaccharidosis I (MPS I; OMIM #252800) is a severe, multi-system, progressive lysosomal storage disorder resulting from deficiency of the enzyme α-L-iduronidase (IDUA; EC 3.2.1.76) [1]. IDUA deficiency results in abnormalities in the degradation of the glycosaminoglycans (GAGs) heparan sulphate and dermatan sulphate, which subsequently accumulate in the lysosome. MPS I is inherited as an autosomal recessive disorder with an incidence of approximately 1 in 100,000 live births and is considered prototypical of the severe MPS disorders. Clinically, MPS I manifests in a spectrum of phenotypes ranging from severe disease, known as Hurler syndrome, to the most attenuated form, Scheie syndrome. Patients with Hurler syndrome commonly present in the first year of life and, when untreated, have a life expectancy of 10 years with a progressive multi-system disease with developmental delay, arthropathy, cardiac and respiratory abnormalities, dysostosis multiplex, hepatosplenomegaly, and corneal clouding. The remainder of patients have an attenuated phenotype which covers a broad and continuous spectrum ranging from onset of symptoms in late adolescence with normal life expectancy and little or no associated arthropathy, cardiac involvement, or direct CNS involvement, to those individuals with onset in childhood and more significant systemic involvement, often leading to death within the second decade. The recent introduction of recombinant enzyme replacement therapy (Aldurazyme™) for

---

* A version of this chapter has been accepted for publication as Derrick R. Randall, Graham B. Sinclair, Karen E. Colobong, Elly Hetty, and Lorne A. Clarke. “Heparin Cofactor II-Thrombin Complex in MPS I: A Biomarker of MPS Disease.” Molecular Genetics and Metabolism, in press.
MPS I [2] will undoubtedly alter the progression of this disease, particularly for patients with attenuated forms. Although the primary enzyme deficiency and the associated metabolic pathway are well understood, there is a poor understanding of the true pathophysiology of disease complications. Genotype to phenotype correlation does exist to some extent for IDUA deficiency; however, the large number of private mutations does not allow genotype to predict phenotype in a considerable proportion of patients [3,4]. In addition, there are no disease biomarkers that unambiguously allow for the determination of disease severity or responsiveness to treatment.

The serum proteome is a potentially rich source of protein that may contain biomarkers indicative of disease severity and disease responsiveness, as well as factors underlying disease progression. Efficient depletion of the high abundance proteins prior to proteomic analysis has proven effective in recognizing expression changes of less abundant proteins in disease states [5-7]. The recent development of iTRAQ reagents [8], has enabled multiplex analysis of up to four independent proteomes through differentially weighted reporter tags. Unlike ICAT proteomic studies, which restrict anlaysis to cysteine containing proteins, iTRAQ reagents ligate to all free amine groups of peptides [9]. This method has been demonstrated to be highly informative [9-11].

Using a depleted serum proteomic approach in the MPS I murine model [12] we have demonstrated significant reduction in the level of heparin cofactor II (HCII), a serine protease inhibitor (serpin), in affected animals. Further analysis revealed that although native HCII levels are reduced, there is marked elevation of HCII-thrombin (HCII-T) complex formation in affected animals. Translation of these observations to humans with MPS I show equivalent findings. Importantly, the elevation of HCII-T
complex appears to be correlated to disease severity and is responsive to treatment. Modulation of serpins by glycosaminoglycans has been well studied, but this is the first report relating serpins to MPS disease. Our results indicate that HCII-T is an excellent biomarker for MPS I and represents a novel finding that may implicate other GAG modulated serpins and their proteases in the pathophysiology of MPS diseases.

3.2 Methods

3.2.1 Sample collection

Mouse serum samples were collected by cardiac puncture and added to Beckton Dickinson serum separator tubes (Franklin Lakes, NJ). Samples were allowed to clot for 30 minutes at room temperature and centrifuged for 15 minutes at 3000xg at room temperature. Samples were aliquoted immediately and stored at -80°C. Mouse plasma samples were collected by cardiac puncture, with sodium citrate added to 0.4%. Citrated samples were centrifuged for 15 minutes at 1500xg at 4°C. Human plasma and serum samples were collected by clean venipuncture using a 2 syringe technique, as described [13] and processed as above.

3.2.2 High abundance protein depletion

Albumin, immunoglobulin, and transferrin were depleted from murine serum samples using a 100 mm Ms-3 Multiple Affinity Removal System (Agilent Technologies, Palo Alto, CA), according to the manufacturer’s instructions. 4 age- and sex-matched pools of mouse serum samples were normalized for total protein to 12 mg/mL by diluting them in running buffer (Agilent Technologies). Each pool contained 6 independent
serum samples. A KD Scientific syringe pump (Holliston, MA) was used to maintain a constant flow rate of 15 mL/h. 320 µL were subsequently used for high abundance protein depletion with identical pool collections obtained for all samples.

3.2.3 iTRAQ analysis

3.2.3.1 iTRAQ reagent labelling

The depleted mouse serum samples were labelled with the iTRAQ reagent following the manufacturer’s protocol (Applied Biosystems, Foster City, CA). Briefly, 100 µg of total protein from each of the 4 depleted serum pools was precipitated with acetone and resuspended in iTRAQ dissolution buffer for reduction, alkylation, and tryptic digestion. Each of the resulting peptide pools was then labelled with a different isotopic iTRAQ Reagent (114 – 117 Daltons) as follows: $Idua^{+/+}$ mixed sex pool (114 Da), $Idua^{-/-}$ male pool (115 Da), $Idua^{+/+}$ male pool (116 Da) and $Idua^{-/-}$ mixed sex pool (117 Da). The four differentially labelled pools were then combined and subjected to strong cation exchange (SCX) chromatography using a polysulfoethyl A column (Poly LC, Columbia, MD). The combined sample was diluted in 10 mM KPO$_4$ (pH 2.7), 25% acetonitrile (Buffer C); applied to the column; and peptides eluted over a 33 minute gradient to 65% Buffer C and 35% 10mM KH$_2$PO$_4$, 25% acetonitrile, 0.5 M KCl (Buffer D). Fractions were collected at one minute intervals.

3.2.3.2 LC-MS/MS analysis

The resulting 33 SCX fractions were then subjected to LC-MS/MS analysis utilizing a QStar Pulsar hybrid quadrupole-TOF instrument (Applied Biosystems) and an
UltiMate micro HPLC (LC Packings, Sunnyvale, CA). The HPLC was equipped with a C18 PepMap guard column (LC Packings) separated from a C18 Pepmap Nano LC column (LC Packings) by a switching valve to allow for precolumn sample clean-up before switching inline for reversed phase chromatography and MS/MS analysis. Each SCX fraction was evaporated to dryness, resuspended in 5% acetonitrile and 3% formic acid, and 25% of the sample was injected onto the C18 guard column in 98% water/acetonitrile (98:2), 0.05% formic acid (Buffer A) with the HPLC flowing to waste to remove sample contaminants. Following 10 minutes at 100 µl/mL, the guard column was switched inline with the C18 resolving column and mass spectrometer, and the peptides were eluted with a linear gradient to 60% water/acetonitrile (2:98), 0.05% formic acid (Buffer B) over 40 minutes. Following a 5 minute ramp to 80% Buffer B, the column was re-equilibrated in 98% buffer A for 15 minutes prior to the injection of the next SCX fraction.

3.2.3.3 Data acquisition and analysis

MS data were acquired automatically using Analyst QS 1.0 software Service Pack 8 (ABI MDS SCIEX, Concord, Canada). An information-dependent acquisition method consisting of a 1 second TOFMS survey scan of mass range 400-1200 amu and two 2.5 second product ion scans of mass range 100 – 1500 amu was utilized. The two most intense peaks over 20 counts, with charge state 2 – 5 were selected for fragmentation and a 6 amu window was used to prevent the peaks from the same isotopic cluster from being fragmented again. Once an ion was selected for MS/MS fragmentation it was put onto an exclusion list for 180 seconds. Following the initial data acquisition run (iTRAQ
acquisition A), an exclusion list was created for all peptides identified with a confidence of 95% or greater in each of the SCX fractions. A second injection of each SCX fraction using the exclusion list for that fraction and the same LC and MS conditions as the first run was then performed in order to detect the lower concentration proteins present in plasma (iTRAQ acquisition B).

The resulting data files were combined and processed using the Interrogator\textsuperscript{TM} algorithm in the ProQuant software (v1.0) (Applied Biosystems) in Analyst using the following parameters: The MS and MS/MS mass tolerances were set to 0.20. A rodent subset of the Celera Discovery Systems Database (01/24/2004) was used for searching. Methyl methanethiosulphonate (MMTS) modification of cysteines was used as a fixed modification. The number of missed cleavages was set to 1. All results were written to a Microsoft Access database and, to reduce redundancy, ProGroup Viewer version 1.0.5 (Applied Biosystems) was used to assemble and report the data. Protein % Confidence Scores, which are influenced by the closeness of the observed peptide spectrum to the predicted spectrum, the number of peptides identified for a given protein, as well as the variance of abundance of each peptide for the given protein, were used to calculate a Protein Score using the equation Protein Score = -log\[1 - (Protein % Confidence)/100\].

Abundances of each identified protein in the sample pools were calculated based on the abundance of reporter tags bound to the peptide tryptic fragments for each pool. Where more than one peptide was identified for a protein, the protein abundance was calculated by weighted averaging of the abundances of the individual peptides and reported with a 95% confidence interval based on the standard deviation of the weighted average. Protein abundances were adjusted for global labeling bias by a correction factor.
that assumed an average relative protein abundance of 1 across all samples. Relative abundances are determined by dividing the abundance of each protein from each pool by the corresponding protein in another pool. All proteins found to have significantly altered relative abundances were manually verified by inspecting the corresponding peptide matches.

3.2.4 Western blotting

Western blot analysis was performed on 7.5% Tris-glycine gels and transferred to Pall (East Hills, NY) BioTrace NT membranes. Anti-human HCII and anti-human antithrombin III antibodies were from Affinity Biologicals (Hamilton, ON). Membranes were blocked with 5% Carnation powdered skim milk (Nestlé, North York, ON) in phosphate-buffered saline with 0.05% Tween-20 overnight at 4°C. Western blot analysis of human samples used primary antibody at a concentration of 1 μg/mL, with secondary antibody at a concentration of 1.3 μg/mL. Western blot analysis of mouse samples used 3 μg/mL primary antibody, and 1.3 μg/mL secondary antibody. Antibody incubations were performed for 60 minutes at room temperature. Proteins were detected with West Pico detection kits (Pierce, Rockford, IL) according to the manufacturer’s instructions.

3.2.5 ELISA

HCII-T ELISA kits were obtained from Affinity Biologicals (Hamilton, ON) and used according to the manufacturer’s instructions. This kit uses polyclonal sheep anti-human thrombin antibody for capture and peroxidase-conjugated polyclonal goat anti-human HCII antibody for detection. Standards were derived from purified human HCII
and thrombin (Enzyme Research Laboratories, South Bend, IN) reacted in the presence of 0.05 U/mL heparin (Sigma, St. Louis, MO). MPS IH serum samples were diluted 500-fold and MPS IH/S samples were diluted 100-fold in factor II-depleted plasma (Affinity Biologicals) while control samples were undiluted. All standards and samples were tested in triplicate.

3.3 Results

3.3.1 iTRAQ serum proteomic studies

Using a 94% Protein Confidence score cut-off applied to data observed over two cumulative MS/MS acquisitions, iTRAQ analysis resulted in the identification of 1701 distinct peptides belonging to 198 unique proteins (Table 3.1). 181 proteins were identified on the strength of two or more peptides, the majority of which were represented by at least 5 peptides (Table 3.2). This weighting toward 5+ peptides per protein is likely due to the dynamic range of protein concentrations in serum, where more abundant proteins can be expected to be identified by several peptides. The second MS/MS data acquisition resulted in a 30% increase in the total number of serum proteins identified.

<table>
<thead>
<tr>
<th>Protein Confidence Level, %</th>
<th>iTRAQ Acquisition A</th>
<th>iTRAQ Acquisition A+B</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Peptides</td>
<td>94 95 99</td>
<td>94 95 99</td>
</tr>
<tr>
<td># of Proteins</td>
<td>1132 1131 1048</td>
<td>1701 1686 1577</td>
</tr>
<tr>
<td># of Proteins</td>
<td>151 150 114</td>
<td>198 195 150</td>
</tr>
</tbody>
</table>

Table 3.1: iTRAQ summary data
Table 3.2: Number of unique peptides per protein in iTRAQ data collection, ≥94% Protein Confidence.

<table>
<thead>
<tr>
<th>Number of Peptides/Protein</th>
<th>Total # of Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>iTRAQ Acquisition A</td>
<td>22</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>5+</td>
<td>91</td>
</tr>
<tr>
<td>iTRAQ Acquisition A+B</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>114</td>
</tr>
<tr>
<td>5+</td>
<td>198</td>
</tr>
</tbody>
</table>

To investigate intersample variation, we compared the relative abundances observed in the *Idua*<sup>+/−</sup> male pool to the *Idua*<sup>+/−</sup> mixed sex pool as well as the *Idua*<sup>−/−</sup> male pool to the *Idua*<sup>−/−</sup> mixed sex pool at a 99% Protein Confidence level (Fig. 3.1). The vast majority of proteins show no significant change in expression levels between like pools, with the exception of one outlier. The possibility of a sex-specific expression difference observed in the outlier (~16-fold decreased) in the *Idua*<sup>+/−</sup> Male:*Idua*<sup>+/−</sup> Mixed sex comparison is eliminated by the absence of such a difference in the *Idua*<sup>−/−</sup> pools comparison and the decreased relative abundance of this protein (alpha-1-B glycoprotein) was assumed to be an artifactual in the *Idua*<sup>+/−</sup> mixed sex control pool.
Fig. 3.1: Box plots of relative abundances for proteins identified with 99% confidence.

Grey boxes cover the interquartile range, whiskers extend to 10th and 90th centiles, and individual points indicate the outliers. Protein abundances are compared between the two WT pools, the two Idua" pools, and the average of the two Idua" pools versus the WT mixed sex pool.

Comparison of the average protein relative abundances of the two Idua" pools with the Idua" mixed sex pool reveals the variation between mutant and control pools is similar to that observed between like pools; however, a few proteins showed more extreme abundance differences (Fig. 3.1). This suggests there is minimal variation in protein quantities between the MPS and normal serum proteomes, with no single protein present at a dramatically different level, unlike the high concentration of chitotriosidase.
(chitinase I) in many patients with Gaucher disease [14]. This result confirms earlier results obtained from 2D gel electrophoresis experiments (Chapter 2).

In order to select a panel of proteins to be investigated as candidate biomarkers with significantly altered relative abundances in the Idua⁺⁻ serum proteome, the relative abundances of each protein in the two Idua⁺⁻ pools were averaged then tested for two factors. First, we determined which proteins showed average relative abundances exceeding the 95% confidence intervals of the same proteins’ abundances in the Idua⁺⁺ male pool. Second, proteins that also showed average relative abundances with confidence intervals not overlapping 0.00 (on a logarithmic scale) were considered strong candidates. The Idua⁺⁺ mixed sex pool was used as the denominator pool for both analyses. Combining these criteria selected proteins exceeding their natural variability in both the MPS and WT animals regardless of the absolute value of the change. Candidate biomarkers selected by these criteria are indicated as red squares in Fig. 3.2 and listed in Table 3.3. Candidate proteins with the most extreme deviation from the normal state were fibrinogen gamma (4.96-fold increased), fibrinogen alpha (2.20-fold increased), and heparin cofactor II (1.79-fold decreased).
Fig. 3.2: Logarithmic plot of average relative abundance of proteins in $Idua^{+/+}$ pools vs. $Idua^{+/+}$ mixed sex, ≥99% confidence, with proteins rank ordered from left to right based on descending Protein % Confidence score.

Dashed lines indicate a two-fold increase or decrease in the average relative abundance of a protein present in the $Idua^{+/+}$ pools compared to the $Idua^{+/+}$ mixed sex pool. Proteins marked as red squares were considered candidates based on significant changes in relative abundance and are listed on Table 3.3.
Table 3.3: Proteins identified at ≥99% confidence with significantly altered relative abundance in the MPS mouse serum, *Iduα* mouse serum.

<table>
<thead>
<tr>
<th>GenBank Accession Number</th>
<th>Protein % Confidence Score Rank</th>
<th>Protein</th>
<th>Log$_2$ of Average Relative Abundance</th>
<th>Fold Increase/Decrease</th>
<th>Biological Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAH19506</td>
<td>95</td>
<td>Fibrinogen, gamma polypeptide</td>
<td>2.310</td>
<td>+4.96</td>
<td>Blood clotting</td>
</tr>
<tr>
<td>AAH05467</td>
<td>71</td>
<td>Fibrinogen, alpha polypeptide</td>
<td>1.137</td>
<td>+2.20</td>
<td>Blood clotting</td>
</tr>
<tr>
<td>AAC28866</td>
<td>87</td>
<td>Alpha-1-antitrypsin 1-5</td>
<td>0.636</td>
<td>+1.55</td>
<td>Serine protease inhibitor (serpin)</td>
</tr>
<tr>
<td>AAH13465</td>
<td>56</td>
<td>Inter-alpha trypsin inhibitor, heavy chain 1</td>
<td>0.480</td>
<td>+1.40</td>
<td>Serpin</td>
</tr>
<tr>
<td>AAA37246</td>
<td>5</td>
<td>Apolipoprotein B</td>
<td>0.347</td>
<td>+1.27</td>
<td>Lipid and fatty acid transport</td>
</tr>
<tr>
<td>AAH57983</td>
<td>2</td>
<td>Pzp protein</td>
<td>0.154</td>
<td>+1.11</td>
<td>Serpin</td>
</tr>
<tr>
<td>AAH23143</td>
<td>29</td>
<td>Gelsolin</td>
<td>-0.144</td>
<td>-1.11</td>
<td>Cytoskeletal protein</td>
</tr>
<tr>
<td>BAA19743</td>
<td>26</td>
<td>Kininogen precursor</td>
<td>-0.281</td>
<td>-1.22</td>
<td>Protein metabolism and modification</td>
</tr>
<tr>
<td>BAB33095</td>
<td>31</td>
<td>Histidine-rich glycoprotein</td>
<td>-0.309</td>
<td>-1.24</td>
<td>Biological process unclassified</td>
</tr>
<tr>
<td>AAC28865</td>
<td>23</td>
<td>Alpha-1 proteinase inhibitor 2</td>
<td>-0.377</td>
<td>-1.30</td>
<td>Serpin</td>
</tr>
<tr>
<td>AAH21776</td>
<td>102</td>
<td>Apolipoprotein C-III</td>
<td>-0.386</td>
<td>-1.31</td>
<td>Lipid metabolism</td>
</tr>
<tr>
<td>AAH30166</td>
<td>90</td>
<td>Factor XIII beta</td>
<td>-0.395</td>
<td>-1.31</td>
<td>Blood clotting</td>
</tr>
<tr>
<td>AAH12706</td>
<td>64</td>
<td>Paraoxonase 1</td>
<td>-0.414</td>
<td>-1.33</td>
<td>Peroxidase</td>
</tr>
<tr>
<td>AAH34543</td>
<td>78</td>
<td>Heparin cofactor II</td>
<td>-0.840</td>
<td>-1.79</td>
<td>Serpin</td>
</tr>
</tbody>
</table>
An analysis of the reproducibility of iTRAQ and the effect of acquiring additional peptides is made possible by the double pool, supplementary iTRAQ data acquisition phases of this study. This information is presented in Appendix 1.

3.3.2 Heparin cofactor II western blot analysis in MPS I mice

To investigate the HCII reduction in the serum of Idua<sup>+</sup> mice, western blot analysis was used utilizing goat anti-human HCII antibodies (Fig. 3.3). Surprisingly, the marked reduction in native HCII levels was associated with the presence of a higher molecular weight protein only in the mutant animals' sera, consistent with the published size of the HCII-thrombin (HCII-T) complex [15].

![Western blot of serum samples from WT (Idua<sup>+</sup>) and MPS (Idua<sup>-/-</sup>) mice for HCII. 15 µg of sample was applied per well.](image)

3.3.3 Heparin cofactor II western blot analysis in MPS I patients

Fig. 3.4 shows the corresponding dramatic elevation of serum HCII-T complex in severe and attenuated cases of MPS I compared to that of controls. Also included is the complex formed from the incubation of pure HCII and thrombin in the presence of heparin to confirm that the complex truly is HCII-thrombin. Interestingly, the largest amounts of HCII-T complex are seen in the MPS IH patients in comparison to the MPS IH/S patient. Western analysis revealed no detectable HCII-T in plasma samples from Idua<sup>-/-</sup> or WT mice (data not shown), nor in plasma samples from humans without MPS I.
Fig. 3.4: Western blot of MPS patients for HCII.
2.5 μg of protein was loaded in wells 1 – 7, 8 ng of protein was loaded in wells 8 and 9.

Figure 3.5a-b shows the level of HCII-T in a Hurler patient receiving enzyme replacement therapy (ERT) preceding and following bone marrow transplantation (BMT), and one Hurler-Scheie patient undergoing ERT only. Enzyme treatment in the Hurler patient (Fig. 3.5a) did not normalize HCII-T levels, but significantly reduced the amount of HCII-T levels to that seen in the attenuated patient studied. Further reduction in HCII-T occurred following bone marrow transplantation. Although well engrafted by week 52, this patient subsequently died of pulmonary hemorrhage. The Hurler-Scheie patient, Fig. 3.5b, showed marked reduction of HCII-T early during ERT exposure but then subsequently developed detectable HCII-T complex later during treatment.
Fig. 3.5: HCII-T levels in MPS 1H serum samples following enzyme replacement therapy.

a) Hurler patient, 12 months of age at start of ERT. All lanes contain 2.5 μg of protein. Patient received a bone marrow transplant between 36 and 52 weeks. b) Hurler-Scheie patient, 8 years of age at start of ERT. All lanes contain 2.5 μg of protein.

3.3.4 HCII-T ELISA analysis

Table 3.4 illustrates the dramatic elevation of HCII-T in the serum of MPS 1H and MPS 1H/S patients as well as murine MPS I samples in comparison to controls. MPS IH patients’ serum HCII-T complex levels ranged from 174,700 – 208,600 pM, with an average value of 188,600 pM, representing a 630-fold increase relative to controls. The serum sample from a MPS IH/S patient had a HCII-T concentration of 46,000 pM (154-fold increase), reflective of the patient’s attenuated phenotype.

In contrast to the lack of detectable complex by western blot, ELISA revealed that plasma HCII-T levels were increased in MPS patients and the MPS I mouse, with the minimum concentration of HCII-T complex in MPS IH patients exceeding the maximum control value by 68%.
Table 3.4: HClI-T ELISA in serum and plasma samples of MPS I patients, MPS I mice, and controls.

<table>
<thead>
<tr>
<th>Sample (age in brackets)</th>
<th>Serum [HClI-T] (pM ± SD)</th>
<th>Plasma [HClI-T] (pM ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (10 yr F)</td>
<td>115.1 ± 17.92</td>
<td>9.91</td>
</tr>
<tr>
<td>Control (10 yr M)</td>
<td>398.0 ± 3.91</td>
<td>6.27</td>
</tr>
<tr>
<td>Control (30 yr M)</td>
<td>384.7 ± 6.27</td>
<td></td>
</tr>
<tr>
<td>MPS 1H (10 mo, Patient A)</td>
<td>174 700 ± 30.15</td>
<td></td>
</tr>
<tr>
<td>MPS 1H (12 mo, Patient B)</td>
<td>182 400 ± Not tested</td>
<td></td>
</tr>
<tr>
<td>MPS 1H (14 mo, Patient C)</td>
<td>208 600 ± 98.37</td>
<td></td>
</tr>
<tr>
<td>MPS 1H/S (8 yr, Patient D)</td>
<td>46 000 ± Not tested</td>
<td></td>
</tr>
<tr>
<td>Idua&lt;sup&gt;-/-&lt;/sup&gt; (n = 5)</td>
<td>75.46 ± 4.99</td>
<td>3.77 ± 1.20</td>
</tr>
<tr>
<td>Idua&lt;sup&gt;-/-&lt;/sup&gt; (n = 3)</td>
<td>628.1 ± 163.2</td>
<td>79.50 ± 38.9</td>
</tr>
</tbody>
</table>

3.3.5 Antithrombin III western blot analysis.

Antithrombin III (ATIII) is the principal circulating serpin, present at approximately twice the plasma concentration of HClI [16], and is known to be activated exclusively by heparan sulphate [17]. Therefore, it was expected that MPS I patients would also show elevated serum levels of ATIII-thrombin complex (ATIII-T), as confirmed in Fig. 3.6. In our small sample set, the reduced dynamic range of the ATIII-T complex in comparison to the HClI-T complex suggests its use as a biomarker may not prove as reliable. Importantly, the clear presence of ATIII-T complex in the control serum indicated that the specificity of this biomarker might not be as accurate as HClI-T for distinguishing the attenuated phenotypes and for measuring subtle changes in response to treatment. Further studies must be done to determine ATIII-T complex’s utility as a biomarker.
3.4 Discussion

There are currently no analytical tools that allow for accurate prediction of disease severity and disease responsiveness for any MPS disorders. Although some genotype-phenotype correlation does exist for mutations at the *IDUA* locus, the large number of private mutations makes this analysis uninformative for many families [3,4]. It is widely accepted that the amount of residual enzyme activity is responsible for the modulation of the disease phenotype, but accurate and reliable methods to determine the amount of residual enzyme activity are not available. Urinary GAG quantification has been shown to be helpful as a screening tool for the MPSs, but it does not necessarily correlate with disease severity [18-21]. Although urinary GAGs are responsive to therapy [2], it is unclear what their source is and whether they reflect the total burden of disease or merely the extent of renal involvement. Accordingly, identification of a reliable, accurate biomarker for the MPSs would be useful for distinguishing disease phenotypes, objectively measuring the burden of disease, and providing a quantifiable indicator of treatment responsiveness.
We have demonstrated significant increases in the serum concentration of the HCII-T complex in murine MPS I and MPS I patients with severe and attenuated phenotypes. In addition, the level of this complex appears to correlate with the clinical measures of disease severity as well as responsiveness to therapy. As such, this may represent an important biomarker of disease. The level of HCII-T needs to be studied in larger numbers of MPS I patients in order to more accurately delineate its role in disease prediction and responsiveness (Chapter 4).

Our limited study of treated patients undergoing ERT and/or BMT show that residual amounts of serum HCII-T remain after treatment. These data would suggest that these modes of therapy do not completely correct the underlying metabolic defect. The residual HCII-T complex in the patient with MPS IH/S may indicate the development of antibodies to Aldurazyme or may be reflective of substrate accumulation in sites that are not accessible to intravenously administered enzyme and thus contribute to circulating GAG levels. Studies with a larger sample of treated patients are needed to correlate the HCII-T levels with the IDUA antibody titre.

HCII represents a member of the serine protease inhibitor (serpin) family [22]. Each serpin regulates the activity of its protease through a “suicide” mechanism where the protease initiates proteolysis of the serpin but cannot complete the reaction, thereby forming a covalent linkage between the two proteins [17]. This complex formation is essentially irreversible and causes a dramatic structural change leading to protease inactivity. The increased HCII-T complex formation shown in MPS patients is likely a direct result of elevated heparan and dermatan sulphate in the circulation. The activation of HCII by GAGs is a well-studied phenomenon, but this is the first time that it has been
linked to MPS disease. GAGs are known to activate various serpins by a hydrostatic interaction, which increases the affinity of the serpin for its protease through either an allosteric or bridging mechanism [23]. The therapeutic use of heparin to alter clotting via activation of antithrombin III is a well-known and frequently used example of this mechanism.

The interaction of dermatan sulphate and heparan sulphate with HCII has been well studied biochemically and is understood to proceed via a mechanism whereby the negatively charged GAG displaces a regulatory arm of HCII that subsequently binds to exosite 1 on thrombin and exposes HCII's reactive centre loop to thrombin's active site [23]. Through this interaction, GAGs are able to increase HCII's reactivity with thrombin by three orders of magnitude [24]. The elevated fibrinogen in Idua<sup>−/−</sup> murine serum samples suggests impairment of thrombin activity during clotting. Although the physiologic reason for specific activation of HCII by dermatan sulphate is not well understood, it may be related to the need to prevent thrombin activation in the extravascular space, where 40% of HCII is localized [23,25]. Furthermore, since HCII cannot substitute for ATIII, as revealed by the ATIII knockout mouse model [26], it is believed that HCII has a separate role and is likely the key thrombin modulator in peripheral tissues. There, HCII may be anchored to dermatan sulphate and other GAGs [25], thus preventing activation of thrombin in extracellular spaces.

Although we have identified that serum HCII-T complex formation is greatly increased in MPS I, the majority of the complex appears to form when fibrinogen becomes depleted. Fibrinogen is known to impair the inhibition of free thrombin by HCII [27]. Therefore the serum levels of HCII-T complex likely directly reflect the
concentration of GAGs in the blood itself. We have also shown increased HCII-T complexes in plasma of patients and mice with MPS I (Table 3.4). Although our studies in humans are limited to a small number of observations, we hypothesize that since HCII-T complex is transported to the liver for degradation [17], it is likely that the HCII-T complex present in plasma forms in peripheral tissues and may reflect the body stores of GAGs. Plasma HCII-T levels are known to be influenced by many factors, and thus further studies with larger numbers of MPS patients will be required to investigate the significance of this observation further.

Although direct effects of GAG storage on cellular function are widely accepted as an underlying pathophysiologic mechanism of disease in the MPSs, there are well-documented alterations of secondary pathways including altered ganglioside metabolism and inflammation that may play a role in the disease process [28,29]. We hypothesize that serpin activation, leading to alteration of various serine proteases, is another important mechanism of disease, which is supported by the observation that 5 of the 14 proteins identified by iTRAQ with modified abundance in MPS I mice were serine protease inhibitors (Table 3.1).

The regulation of serine proteases by serpins has been extensively studied, with multiple serpins and their respective proteases identified [17]. Interestingly, GAGs modulate a number of serpins in vitro (Table 3.5). These GAG-regulated serpins and the proteases that they regulate are thus excellent candidates for factors which may underlie the pathogenesis of various MPS diseases. It is possible that other GAG-regulated serpins (and the proteases they regulate) may underlie various aspects of MPS disease and explain some of the varying symptoms between heparan sulphate-storing and
dermatan sulphate-storing MPS disorders. For example, HCII, in addition to controlling thrombin activity, is also known to inhibit cathepsin G, a lysosomal serine protease associated with connective tissue remodelling at inflammation sites [30-32]. This inhibition is greatly increased in the presence of dermatan sulphate but not heparan sulphate [33]. Glia-derived nexin (protease nexin-1, PN1) is a serpin expressed in the nervous system that shows heparan sulphate-induced inhibitory activity toward the serine proteases thrombin, acrosin, and Factor X1a [17]. Furthermore, in vitro and in vivo studies have shown that PN1 co-localizes with the extracellular matrix (ECM) and inhibits ECM degradation [34], while PN1 overexpression increases ECM production [35]. It is therefore interesting to consider that alteration of these proteases by the accumulation of heparan sulphate in MPS I, II, III, VII may underlie aspects of CNS involvement in these disorders.
Table 3.5: Known serpins with GAG-mediated activities.

<table>
<thead>
<tr>
<th>Serpin</th>
<th>GAGs that modulate activity</th>
<th>Known target proteases</th>
<th>Tissue specificity</th>
<th>Biological function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$-antitrypsin ($\alpha_1$AT, $\alpha_1$PI, SERPINA1)</td>
<td>HS, DS</td>
<td>Neutrophil elastase, Plasmin, Thrombin</td>
<td>Plasma</td>
<td>General inhibitor, inflammation [17,38,39]</td>
</tr>
<tr>
<td>Protein C inhibitor (PAI-3, SERPINA5)</td>
<td></td>
<td>Thrombin, activated Protein C, Urinary plasminogen activator, Tissue plasminogen activator, Plasma kallikrein, Acrosin</td>
<td>Plasma</td>
<td>Coagulation [17]</td>
</tr>
<tr>
<td>Antithrombin III (ATIII, SERPINC1)</td>
<td>HS</td>
<td>Thrombin, Factor IXa, Factor Xa</td>
<td>Plasma, extravascular space</td>
<td>Coagulation [17]</td>
</tr>
<tr>
<td>Heparin cofactor II (HCII, SERPIND1)</td>
<td>HS, DS</td>
<td>Thrombin, Cathepsin G, Chymotrypsin</td>
<td>Plasma, extravascular space</td>
<td>Coagulation, unknown [17]</td>
</tr>
<tr>
<td>Glia-derived nexin (PN1, SERPINE2)</td>
<td>HS</td>
<td>Thrombin, Acrosin, Urinary plasminogen activator, Tissue plasminogen activator, Factor XIa</td>
<td>Fibroblasts, neurons, extracellular space</td>
<td>Neurotrophic, ECM remodelling [17,34,35]</td>
</tr>
<tr>
<td>Pigment epithelium-derived factor (PEDF, SERPINF1)</td>
<td>HS</td>
<td>N/A</td>
<td>Cornea, cartilage, bone</td>
<td>Neurotrophic, antiangiogenic [17,36]</td>
</tr>
</tbody>
</table>
Heparan sulphate has been shown to inhibit the serpin PEDF [36]. Since PEDF is secreted by the retinal pigment epithelium and associates with the interphotoreceptor matrix [36], alteration of this pathway in MPS I may underlie the retinal dysfunction noted in severely affected patients [37].

In summary, we have shown that heparin cofactor II-thrombin (HCII-T) complex formation in MPS I patients appears to be a biomarker of MPS disease. Although the studies reported here relate specifically to MPS I, HCII-T complex formation likely occurs in all MPSs where heparan or dermatan sulphate accumulates and may represent a generalized mechanism whereby GAG modulation of serpin-protease interaction affects the pathophysiology of related storage disorders.
3.5 References


[28] C.M. Simonaro, M.E. Haskins, and E.H. Schuchman, Articular chondrocytes from animals with a dermatan sulfate storage disease undergo a high rate of apoptosis and


Chapter 4: Heparin Cofactor II-Thrombin Complex Levels in MPS I Patients Undergoing Enzyme Replacement Therapy

4.1 Introduction

Mucopolysaccharidosis I (MPS I; OMIM #252800) is a severe, multi-system, progressive lysosomal storage disorder resulting from deficiency of the enzyme α-L-iduronidase (IDUA; EC 3.2.1.76) [1]. IDUA deficiency results in abnormalities in the degradation of the glycosaminoglycans (GAGs) heparan sulphate and dermatan sulphate, which subsequently accumulate in the lysosome. MPS I is inherited as an autosomal recessive disorder with an incidence of approximately 1 in 100,000 live births and is considered prototypical of the severe MPS disorders. Clinically, MPS I manifests in a spectrum of phenotypes ranging from severe disease, known as Hurler syndrome, to the most attenuated form, Scheie syndrome. Patients with Hurler syndrome commonly present in the first year of life and, when untreated, have a life expectancy of 10 years with a progressive multi-system disease with developmental delay, arthropathy, cardiac and respiratory abnormalities, dysostosis multiplex, hepatosplenomegaly, and corneal clouding. The remainder of patients have an attenuated phenotype which covers a broad and continuous spectrum ranging from onset of symptoms in late adolescence with normal life expectancy and little or no associated arthropathy, cardiac involvement, or direct CNS involvement, to those individuals with onset in childhood and more significant systemic involvement, often leading to death within the second decade. The

† A version of this chapter is in preparation for submission for publication as Derrick R. Randall, Graham B. Sinclair, Elly Hetty, and Lorne A. Clarke, “Heparin Cofactor II-Thrombin Complex Levels in MPS I Patients Undergoing Enzyme Replacement Therapy.”
recent introduction of recombinant enzyme replacement therapy (ERT, Aldurazyme™) for MPS I [2] will undoubtedly alter the progression of this disease, particularly for patients with attenuated forms. Genotype to phenotype correlation does exist to some extent for IDUA deficiency, however the large number of private mutations does not allow genotype to predict phenotype in a considerable proportion of patients [3,4].

Recently, our lab demonstrated serum levels of the serine protease-serine protease inhibitor (serpin) covalent complex, heparin cofactor II-thrombin (HCII-T) complex, was drastically elevated in a small sample of MPS I patients [5]. To assess the biomarker’s utility in MPS I diagnostics and research we established a normal serum HCII-T complex concentration range and validated the complex’s elevated serum concentration in an independent set of MPS I patients. In this study, we present results from 11 patients with attenuated MPS I phenotypes enrolled in a randomized, double-blind, placebo-controlled trial of recombinant human α-L-iduronidase [2], as well as samples from two patients not involved in the ERT trial taken at additional time points.

4.2 Methods

4.2.1 ELISA

HCII-T ELISA kits were obtained from Affinity Biologicals (Hamilton, ON) and used according to the manufacturer’s instructions. This kit uses polyclonal sheep anti-human thrombin antibody for capture and peroxidase-conjugated polyclonal goat anti-human HCII antibody for detection. Standards were derived from purified human HCII and thrombin (Enzyme Research Laboratories, South Bend, IN) reacted in the presence of 0.05 U/mL heparin (Sigma, St. Louis, MO). Control samples were undiluted. Pre-
treatment MPS IH serum samples were diluted 500-fold and pre-treatment MPS IH/S samples were diluted 100-fold in factor II-depleted plasma (Affinity Biologicals). Post-treatment sample dilutions varied from undiluted to 250-fold dilutions. Standards and samples were tested either in duplicate or in triplicate. All samples tested in duplicate provided HCII-T complex concentrations within 7.5% of each other.

4.2.2 Subject sample collection

Anonymous control subject serum samples were provided by the Biochemical Diseases Laboratory at BC Children’s Hospital (UBC Clinical Research Ethics protocol approval C04-0472). These samples were originally acquired from patients having blood tests done for diagnostic purposes. MPS I patient serum samples undergoing ERT were provided by Biomarin Pharmaceutical, Inc.

4.3 Results

4.3.1 ELISA data on serum samples from 72 additional control subjects

70 of 72 control samples provided HCII-T complex concentrations within or near the range determined in our previous study [5], with the remaining two higher than anticipated but well below levels observed in MPS I patients. Fig. 4.1 shows the range of serum HCII-T concentration values obtained, graphed versus subject age, after combining the original three control values with the new data. The mean value (± SD) was 274.0 ± 249.8 pM, with a range of 38.94 pM – 1491 pM and the median was 196 pM. This range of values does not follow a normal distribution and shows a positive skew. There is no relationship between HCII-T and age, nor with gender.
4.3.2 ELISA data on serum samples from 11 patients undergoing enzyme replacement therapy

Fig. 4.2 displays serum HCII-T complex concentrations in patients with an attenuated phenotype that were part of a double-blinded, placebo-controlled trial of enzyme replacement therapy (ERT) at three time points: pre-treatment, 26 weeks post-treatment, and 72 weeks post-treatment. At 26 weeks, 5 of 11 patients show a reduction of HCII-T complex to less than 5% of pre-treatment levels, while the remaining 6 patients show modest reductions, no improvement, or elevation. At 72 weeks, all patients have HCII-T complex levels within or approaching the normal range, suggesting the 6 patients who did not improve immediately received placebo for the initial 26 weeks of the trial, before the placebo control was eliminated. This was confirmed upon unblinding of the trial data after the trial’s completion, following consultation with the primary physicians caring for the patients.
Fig. 4.2: Serum HCII-T concentrations for MPS I patients undergoing a randomized, double-blind, placebo-controlled ERT trial. 5 of 11 patients showed a large reduction in HCII-T complex after 26 weeks of ERT. The remaining 6 patients showed similar reductions by 72 weeks, after placebo control was removed. Dashed bar indicates control mean; dotted lines indicate one standard deviation of control values. Patient letter designations represent clinical trial identifications.

Prior to treatment, patients showed an HCII-T complex concentration range of 16,300 pM to 177,000 pM, with a median value of 45,300 pM (Table 4.1). At 72 weeks, the HCII-T complex reached a median value of 987 pM and ranged from 352 pM to 1770 pM. Patient letter designations represent clinical trial identifications.
Table 4.1: HCII-T complex levels in patients undergoing ERT

<table>
<thead>
<tr>
<th>Patient</th>
<th>Pre-treatment</th>
<th>Week 26</th>
<th>Week 72</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAF</td>
<td>18 200</td>
<td>520</td>
<td>582</td>
</tr>
<tr>
<td>AAG</td>
<td>41 900</td>
<td>14 100</td>
<td>1110</td>
</tr>
<tr>
<td>AAH</td>
<td>128 000</td>
<td>838</td>
<td>954</td>
</tr>
<tr>
<td>AAJ</td>
<td>24 500</td>
<td>39 900</td>
<td>1680</td>
</tr>
<tr>
<td>ASS</td>
<td>61 600</td>
<td>11 100</td>
<td>1770</td>
</tr>
<tr>
<td>EC</td>
<td>45 300</td>
<td>1010</td>
<td>521</td>
</tr>
<tr>
<td>JBP</td>
<td>177 000</td>
<td>57 100</td>
<td>987</td>
</tr>
<tr>
<td>MDC</td>
<td>16 300</td>
<td>62 100</td>
<td>865</td>
</tr>
<tr>
<td>TML</td>
<td>63 000</td>
<td>66 300</td>
<td>1650</td>
</tr>
</tbody>
</table>

4.3.3 ELISA data on serum samples from two additional patients undergoing ERT

Fig. 4.3 presents the serum HCII-T concentration in a Hurler patient who underwent ERT for 12 months with bone marrow transplantation (BMT) combined with ERT after 36 weeks. After a small increase in serum HCII-T complex at week 1, the patient shows a rapid reduction over 4 weeks, then an unexpected increase and subsequent gradual reduction beginning at week 10 until week 25. Week 52, the time point with the lowest serum HCII-T complex level (6000 pM), represents both ERT and BMT effects.
Fig. 4.3: Serum HCII-T complex concentration over a Hurler patient’s treatment course

HCII-T complex concentration rapidly reduces following enzyme therapy implementation. An unexpected re-elevation decreases gradually.

Fig. 4.4 shows an attenuated patient’s HCII-T complex levels and titre of antibodies raised to recombinant IDUA over 108 weeks of ERT. The patient showed a similar dramatic reduction in HCII-T complex immediately following treatment and appeared to stabilize at a level within the normal range until 24 weeks, when complex levels rise corresponding with a sustained high antibody titre. HCII-T complex levels quickly return to the normal range by week 66 (as does antibody titre), but rise again to concentrations approaching pre-treatment levels at weeks 80 and 108, without concurrent antibody titre increase. This patient had a plasma HCII-T complex concentration of 9.51 pM at 108 weeks, well within the normal range.
Fig. 4.4: Serum HCII-T complex concentration and antibody titre over an attenuated patient's treatment course

Similar to the Hurler patient in Fig. 4.3, HCII-T complex concentration reduces very quickly following treatment, then increases briefly, which appears related to antibody production. A second, more dramatic HCII-T increase occurs independent of antibody titre following long term treatment.

4.4 Discussion

Recently, our lab demonstrated serum levels of HCII-T complex were elevated in a small set of MPS I patients and reported it as an informative biomarker for MPS I [5]. The biomarker appeared to reflect disease severity and to respond to ERT and BMT. Here we extend our previous work by establishing a normal range for serum HCII-T and by evaluating serum HCII-T complex levels at three time points in 11 patients undergoing a placebo-controlled ERT trial. Serum HCII-T complex exceeded the control range in every pre-treatment case, declined in all patients following bona fide therapy, and was an accurate predictor of patients on placebo control.

Eight of eleven patients in the trial had an HCII-T complex concentration reach normal levels following 72 weeks of therapy, although this number reduces to three of eleven if the two outlier values in the normal range are excluded. These results indicate
ERT cannot completely restore tissue IDUA activity in some patients and may reflect the inability of recombinant enzyme to access select tissues and organs, such as the brain [6,7]. This is supported by the fact that after approximately one year of ERT, the concentration of serum HCII-T in the Hurler patient remained over three-fold greater than the highest concentration seen in the attenuated patients.

An intriguing feature is the range of HCII-T complex concentrations between patients. All possessed HCII-T complex concentrations significantly above normal, but the concentration of HCII-T complex in some patients fell into the range established for young Hurler patients [5]. Thus, it appears the biomarker indicates severity of disease but not strictly phenotype. Although, once age is taken into account, whether the patient is of an attenuated or a severe phenotype can be inferred since attenuated patients do not appear to reach HCII-T complex concentrations equal to that of a severe patient until exceeding the expected lifespan of a severe patient.

Also interesting is the high variability in HCII-T complex concentration a single patient shows between sample collection dates. Because the patients MRO and AAI were on placebo for the first stage of the clinical trial, it is unlikely GAGs would accumulate at such a dramatic rate between sample collections. Instead, it seems plausible that the blood GAG level is highly variable during the course of a patient’s life and may fluctuate for a number of physiological reasons. This is not surprising given the knowledge that urinary GAGs are highly variable among individual humans and mice with MPS, and underscores the need for further analysis.

In addition, this work considers patients with more frequent serum HCII-T complex assays and suggests therapeutic response follows a pattern where initial blood
GAG load decreases rapidly but then increases to 25 – 50% of the pre-treatment storage level before a subsequent gradual reduction. The initial decrease indicates blood GAGs and stored GAGs are rapidly depleted. As the antibody titre from the attenuated patient suggests, patients may mount an immune response that reduces the recombinant enzyme’s efficacy and leads to a temporary resumption in storage. Immune tolerance occurs following continued therapy and permits the recombinant enzyme’s function. Comprehensive antibody titre assays must be done with additional patients to address this hypothesis.

Moreover, our attenuated patient with serum taken at multiple time points presents an interesting situation in the uniquely re-elevated HCII-T complex following long-term treatment. Clearly, the patient is experiencing an accumulation of GAGs in the bloodstream, but their source may not relate to recombinant enzyme efficacy. One possibility is that as the patient grows, additional GAGs may be released into the bloodstream from the growth plates in the long bones. These GAGs may not cause any pathological consequences for the patient and could be a normal feature of a growing person, which may not be captured in our control survey. Another scenario is that as the patient grows, demand for enzyme is likely to increase, secondarily to both greater body mass and elevated metabolism. Most recombinant IDUA activity is noted in the liver, spleen, lung, and plasma [8,9], so the patient’s growth phase may exceed the drug’s bioavailability in other tissues. Interestingly, the patient’s 108 week plasma HCII-T complex level falls into the normal range [10], suggesting the GAGs are quarantined to the bloodstream and not causing complications in the extracellular space.
As an extension of our initial study identifying serum HCII-T as an informative biomarker of MPS I [5], we show here the utility of HCII-T complex as a serum biomarker for diagnosing MPS I and for assessing therapeutic efficacy. All patients possessed pre-treatment serum HCII-T complex concentrations well beyond that of 75 random control values, indicating serum HCII-T is a highly sensitive and specific biomarker. The biomarker accurately reflected whether patients were receiving placebo or recombinant IDUA in a long-term clinical trial and demonstrated a quantifiable improvement in blood GAG burden in patients receiving either ERT or a combination of ERT and BMT. Frequent serum HCII-T complex concentration assays in two patients with different phenotypes indicate a pattern of stored GAG elimination that includes a proposed immune tolerization phase.
4.5 References


Chapter 5: Summary & Conclusion

5.1 Proteomics in disease research

In this thesis, I followed an unbiased proteomic discovery approach to identify a serum biomarker of MPS I. Although this disease is considered a single gene disorder, the resulting phenotype is complex and results from the unique tissue physiologies, structures, activities, and functions that occur in disease states. Analyzing healthy and diseased tissue proteomes should enable researchers to understand normal physiology as well as the intricacies of disease. While the genome is 99.9% identical between two humans [1], any given tissue proteome can show a much higher degree of variability between two people or even among the different sections of an organ from one individual [2-6]. Such is the case for the MPS disorders: the phenotypic spectrum demonstrates variable involvement of different tissues, and patients with identical mutations can present with different clinical severity [7], indicating multiple cellular systems and individual variability contribute to MPS progression.

The genome’s simplicity belies the complexity in cells: each cell possesses an estimated 30,000 protein-coding genes [8,9], yet the human body produces between 200,000 and 1,000,000 unique proteins [9,10]. This discrepancy cannot be explained by selectively expressed housekeeping genes and likely arises from alternate transcripts, post-translational modifications, and tissue-specific isoforms [10]; in effect, cells multi-task their gene products. Furthermore, protein complements in different tissues change in individual protein concentrations, and it is well known that gene expression levels do not necessarily correlate with protein abundance levels [11-15].
It is somewhat surprising that no major perturbation of the serum proteome was observed by either methodology in my study, nor by other groups employing a proteomic approach to discovering blood-borne MPS I biomarkers [Ed Wraith, personal communication]. This is in direct contrast to Gaucher disease, a lipid-storing member of the lysosomal storage disorders family, where chitotriosidase (a protein involved in macrophage stimulation [16]) is elevated several hundred-fold in most patients [17]. However, this latter observation is a consequence of the massive expansion of the macrophage population that occurs in Gaucher disease [17]. MPS I displays no comparable cell expansion. Nonetheless, multi-system disorders are hypothesized to have atypical proteins in the bloodstream; these markers may be in low molar quantities but should be manifold above the norm. An example of this is found in the MS identification of a secondary marker of Gaucher disease, a chemokine known as PARC or CCL18, that is elevated approximately 25-fold in patients [18] and remains applicable to patients genetically deficient for chitotriosidase [18].

In the absence of a single highly informative biomarker, it has been suggested to incorporate biomarker panels in multi-system diseases. This approach would have the potential to signify disease via a “proteomic fingerprint” and to denote dysregulation (and therefore involvement) of individual tissues [19,20]. This is ideal but relies on identification of a series of biomarkers first, which remains to be achieved. My work suggests panel analysis is possible in MPS I serum, but the typical blood protein concentration ranges make the changes seen in MPS I mice insufficient for accurate distinction between individual affected and healthy cases.
Approaches such as protein microarrays have shown marginal success in the identification of biomarkers in complex diseases [21-24], but the fundamental limitation of this method relates to the requirement of protein candidates. Furthermore, quaternary structures, protein-ligand interactions, and non-specific binding of lipids and other biocompounds can impede protein detection [25]. Likewise, some post-translational modifications, which can indicate much about a disease’s process, cannot be observed by this method [19,25-27]. It is clear that unbiased proteomic approaches are necessary to advance understanding of MPS I pathophysiology since hypothesis-driven research into pathological mechanisms has proven unproductive. As the unbiased techniques—in particular 2DGE and mass spectrometry (MS)—improve in power and resolution, more discoveries will be made and tissue proteome maps will become more advanced.

An obstacle to unbiased proteomic techniques is the difficulty in detecting changes in quantities of low abundance proteins [28], which is essential to identifying biomarkers in the serum and plasma proteomes [28,29]. I have shown that high-abundance protein depletion greatly enhances serum proteome analysis, but removing the three highest abundance species is insufficient to identify very low expressed proteins, indicating another technique will be needed to monitor the very low end of the abundance spectrum. It has been argued that ICAT is more proficient than iTRAQ for measuring abundance changes in transcription factors, cytokines, and other signal molecules because they are low abundance, cysteine-containing macromolecules; as a result, cysteine-specific labels should display lower background interference and detect these informative proteins while iTRAQ would preferentially collect data from the amine-tagged, high abundance proteins [30]. Although this may be true in some instances, such as
endometrial tissue [30], it would surely fail in the analysis of blood proteomes because the highest abundance protein, albumin, contains 35 cysteinyI residues and would eliminate detectable signals from low abundance proteins. One would expect this to render low abundance protein deviations undetectable by ICAT and, in fact, this is the case [LAC, unpublished data].

In addition, it will remain important to combine techniques in studies, as it has been shown that some protein expression changes identified by one method will escape another [31]. This does not indicate that the identification is wrong, but rather that it eludes the detection method—proteins with molecular weights below 5 kDa, or that are very basic, hydrophobic, or membrane-associated are not compatible with 2DGE [32-34] while MS-based techniques are insensitive to many post-translational modifications and therefore cannot distinguish functionally different protein forms transcribed from the same gene [35]. Algorithms for analyzing mass spectra can incorporate mass tolerance ranges in fragmentation patterns due to modifications, but this reduces peptide confidence and can introduce false identifications [36]. It has also been reported that peptides mapped to the proteins best identified by MS typically cover only 50% of the protein's sequence [20,37], a point I confirm as the protein with the most peptides in my study (complement component 3) had only 56% sequence coverage through 84 peptides.

Ultimately, a functional enzymatic event occurring in serum proved to be an informative serum biomarker of MPS I. Had more proteins been depleted, this may have been detectable by 2DGE as a significantly increased spot in *Idua<sup>-/-</sup>* serum pools. Because there is only a small peptide cleaved from HCII and one unusual peptide forming a junction between HCII and thrombin—neither of which were recognized by
the detection algorithms—iTRAQ was blind to the complex and considered all peptides to derive from nascent proteins. I hypothesize the reduction in *Idua*<sup>–/–</sup> serum HCII observed by iTRAQ occurs because HCII in the presence of DS inhibits clot-bound thrombin [38,39]; consequently, additional HCII molecules sediment during centrifugation. This is supported by the unaltered concentration of antithrombin III (ATIII), which is not activated by DS, in *Idua*<sup>–/–</sup> serum.

### 5.2 Replicability of iTRAQ

To date, all published accounts of iTRAQ experiments have compared one or more test pools to a single control pool [30,40-42]. Using like-pool comparisons, my study indicates iTRAQ results are reproducible between different, comparable samples and the double pool method that I have incorporated enhances the strength of the results. Comparing mutant to mutant and wild-type to wild-type pools allows sample handling error assessment, which was demonstrated in this study to be minimal. This is valuable because it provides a quality control monitor of the high abundance protein depletion methods. If not for this validation, alpha-1-B glycoprotein (the outlier discussed in Chapter 2) would have appeared to be significantly decreased in the *Idua*<sup>–/–</sup> pools. Second, using two mutant animal serum pools creates two simultaneously tested, independent sets to compare to the control pools. The ratios obtained from each comparison can be averaged to yield a more accurate relative abundance ratio and inspected to see the likelihood of each putative difference. Third, being able to compare the abundance of a protein in one control pool to that of another control pool at the same time as the average of the mutant pools indicates normal protein variability and machine
replicability. This offered a way to be more stringent with respect to how I selected significant protein abundance increases or decreases: disregarding the second \( Idua^{+/+} \) pool identified 37 candidate biomarkers, which may have consumed valuable time and resources during the validation process. Although including like-pool comparisons reduces the number of test samples available for comparison, the quality control factor in a single experiment compensates for this.

Overall, acquisition of new peptides through peptide exclusion increases protein identification confidence. A small number of protein identifications decrease in confidence, but this is likely a correct change due to peptide re-assignments. However, 38% of proteins did not gain peptides, and most of these proteins had single peptides or a collection of low confidence peptides initially identifying them. In fact, 15 of 22 proteins identified by one peptide in the initial data acquisition gained no peptides following the second data acquisition; so it appears performing additional iTRAQ data acquisitions strengthens high confidence protein identities, which is not particularly informative. As such, excluding previously acquired peptides and any predicted peptides from a protein that has been strongly identified in the initial run may be essential to maximize proteome analysis through iTRAQ. My results indicate proteins identified with 99% confidence do not show significant changes in abundance following supplementary data acquisition, and only one protein was erroneously identified with 99% confidence.

Providing relative quantities of a large set of proteins is a powerful analytical tool. Granted, iTRAQ reproducibility supports its inclusion in proteome analysis, but its capability in complex tissue proteomes appears limited to cases where informative proteins are expressed within a few orders of magnitude of the most abundant species,
permitting enhanced coverage of each protein. As a result, it may emerge as a greater strength in organelle proteomics, where proteomes have been simplified several degrees and the objective is to identify irregular processes of disease pathophysiology.

5.3 Identification of a serum biomarker for MPS I disease

This thesis describes the identification and validation of the heparin cofactor II-thrombin (HCII-T) complex as a biomarker elevated in serum of the murine MPS I model and human MPS I patients. Given the severity of MPS I, the ability to rapidly diagnose a patient and to objectively assess response to therapy would be an enormous advancement to patient care. Moreover, since there is a generally poor understanding of MPS I pathophysiology, identifying a serum biomarker may contribute to elucidating the process of disease progression in affected tissues. This biomarker presents many possibilities for future research and clinical management of MPS I patients.

In addition to diagnosing MPS disorders, an obvious use for the HCII-T complex biomarker is to measure a patient’s responsiveness to therapeutics. I show that enzyme replacement therapy (ERT) quickly reduces GAG storage, mirroring improvements noted in physicians’ exams of patients in Aldurazyme™ clinical trials [43], and can accurately predict which patients received placebo during a clinical trial. Although our data are limited to one patient who also underwent ERT, HCII-T complex concentration may also indicate that bone marrow transplantation further reduces the burden of disease. In light of reports that infused enzyme is unable to cross the blood-brain barrier while hematopoietic-derived IDUA can [44,45], the additional reduction of blood GAGs following BMT may be due to storage clearance from ERT-inaccessible sites. Although
BMT is a viable option for patients with Hurler syndrome, it is unnecessary for many patients with attenuated phenotypes and should be reserved for those with neurological involvement [45,46]. Because genotype cannot predict in every case whether an attenuated patient will have neurological involvement [46,47], serum HCII-T may be a useful prognostic indicator if a patient's blood GAGs do not approach the normal range following prolonged ERT.

Similar to concerns surrounding urinary GAGs and their source, the results of this thesis must face questioning about the source of blood GAGs; are they released from cells and tissues and transported to the circulatory system for elimination, or are they bursting from lysosomes in the luminal wall cells? In the latter case, blood GAG level will indicate the severity of vascular cell involvement. In the former, which seems more likely, blood GAG levels should indicate average tissue GAG accumulation and degree of enzyme saturation. Due to continuous storage, there will be a natural perfusion rate of GAGs from the tissues to the bloodstream, and this should relate to overall tissue storage. Although organ size and metabolic rate will also affect this GAG flux, the blood GAG level may be indicative of the most severely afflicted organs if tissues do not all metabolize GAGs similarly. On the other hand, GAGs produced by a highly involved organ may be diluted in the bloodstream by the smaller quantity of GAGs released by less affected organs; this would have the effect of reducing the apparent phenotypic severity. Similarly, if the organ in question is small or has a minor outlet to the bloodstream, its contribution to blood GAG levels may be underrepresented.

These factors present a concern for relating HCII-T concentration with an individual patient's condition when it is known that recombinant human IDUA does not
access all tissues [44]. Hence, in order to use the HCII-T complex as a measure for disease severity, it will be critical to understand how it reflects general tissue involvement. These questions can be addressed through dosage trials utilizing appropriate parameters to correlate tissue involvement with serum HCII-T complex concentration. They also highlight the lack of understanding of the mechanisms of MPS progression and Aldurazyme™ tissue distribution.

Plasma HCII-T complex values also present an intriguing facet of MPS research. Although not elevated in patients to the same degree as serum HCII-T complex, if plasma HCII-T reflects the extravascular and tissue GAG burden, it may be the optimal biomarker during long-term therapy. As shown in Chapter 3, the pre-treatment concentration of HCII-T complex in both MPS IH patients whose plasma were tested exceeded that of three controls, but one patient value fell into the normal adult range (3 – 60 pM) [48]. The observation that one attenuated patient’s serum HCII-T level was increasing over long term treatment but contained normal plasma HCII-T is additionally confounding. However, the absence of pre-treatment and additional post-treatment values prevents conclusion. The connection between plasma HCII-T complex concentration and MPS I needs to be investigated in a greater number of samples; following that, a correlation between disease severity and plasma HCII-T concentration may be apparent. Importantly, an individual’s deflection from his or her intrinsic plasma HCII-T level may be more valuable than comparison to the normal range. While this value can never be known, consistent reductions in plasma HCII-T following therapy may be important.
It is generally believed that in order to minimize degeneration due to MPS disorders one must identify the disease early, before irreversible pathology has occurred [49-51]. Given that the incidence of all MPS types combined is comparable to that of phenylketonuria and galactosemia, two disorders screened at birth in British Columbia [52], the severity of MPS supports its inclusion in newborn screening. Prior to identification of the HCII-T complex biomarker, it was not feasible to do large-scale newborn screening: comprehensive enzyme assays are too costly and urinary GAG analysis has too low of a positive predictive value to validate with enzyme assays [53-56]. HCII-T provides an avenue, assuming blood GAG levels are significantly higher in newborns with MPS than in unaffected newborns. A fortuitous advantage of elevated serum HCII-T complex in MPS patients and using it as a diagnostic is that it may be possible to test dried blood spots—which undergo the clotting process—taken after a child’s birth. If HCII-T complex levels above 1000 pM are assumed abnormal, this could be reliably detected in triplicate with less than 1 μL of serum. My results indicate two control subjects had serum HCII-T levels above this, but it must be remembered that these samples were acquired from subjects with health complications and the elevated complex concentration may be related to their conditions. Another possibility, since many hospitals are switching to MS-based screening methods [57], is to establish the mass fingerprint of the peptide fragment containing the covalent linkage connecting HCII and thrombin. Using HPLC, it would be possible to determine the peptide’s elution time, and to set a detection window for that time period. This would require the proteins to be trypsinized (or cleaved with another specific protease) in order to generate the peptides, so it may be preferable to develop a protocol in which proteins are fragmented into
peptides in the first step of a tandem MS scan and recognize the junction peptide this way.

5.4 Serine proteases and serine protease inhibitors in MPS disease

My discovery of HCII-T complex as a biomarker is unusual because the formation of the complex relies on the presence of GAGs, the storage molecule in MPS I, in circulation to mediate complex formation during coagulation; but, neither serum nor plasma complex levels explain the mechanisms behind MPS tissue pathology. However, they do provide a window into MPS pathophysiology: the potential involvement of other serine proteases and their inhibitors in MPS progression is a logical extension of work in this thesis. There are an estimated 500 serine proteases encoded by the human genome [58], most with undetermined inhibitors; many of the known serine proteases have been poorly characterized with respect to physiological inhibitors and even less so with respect to modulation by GAGs. Heparin-binding motifs are currently unpredictable and require protein crystal structures, mutagenesis experiments, or kinetic assays for confirmation [59]. There are fewer human serine protease inhibitors (serpins), with only 34 known as of 2002 [60], but given their ability to inhibit multiple proteases it is highly likely that a critical interplay of these elements in the tissues occurs. Learning about these interactions may contribute to developing novel therapeutics; access to the murine MPS I model to investigate tissue proteases and inhibitors will be a tremendous aid.

All inhibitory serpins neutralize their target proteases through a conserved process: a reactive centre loop containing a sequence specific to a class of serine proteases is exposed on the protein surface, accessible to proteases [60,61]. The protease
initiates proteolysis but is unable to complete the digestion, thereby forming an irreversible, covalent ester linkage with the inhibitor [60]. GAGs enhance this reaction rate by acting as a bridge between the serpin and the protease or by allosterically activating the serpin through displacement of a regulatory structure [60-62]. Serpins bound to target proteases resemble “molecular mousetraps” [63] and transform the bound protease 180° to the opposite pole, imposing structural changes on both proteins (Fig. 5.1) [63]. The complex is metastable and recognized by specific receptors on cells for phagocytosis and degradation [64]. The location of the degradative cells depends on inhibitor tissue localization: HCII-, ATIII-, and protein C-protease complexes are degraded in the liver, while PN1-protease and neuroserpin-protease complexes are degraded in the brain [65-68]. Degradation is thought to occur within minutes of complex formation [60,67,69].
5.5 Future work

Identification of the HCII-T complex as a biomarker presents many avenues for future research. Using the murine MPS I model has proven extremely useful in the discovery phase, and many of the upcoming study plans incorporate the MPS I mouse to gain valuable insights into MPS pathophysiology and therapeutic response.
5.5.1 Validation in other MPS disorders

While the work here describes the biomarker in MPS I, HCII’s activity is augmented by the presence of either heparan or dermatan sulphate. Therefore, it is possible the biomarker will be useful in several MPS disorders, because all except MPS IV and IX (Morquio and Hyaluronidase deficiency syndromes, respectively) accumulate heparan sulphate (Sanfilippo, III), dermatan sulphate (Maroteaux-Lamy, VI), or both (Hunter, II; Sly, VII) [46].

5.5.2 Recombinant human IDUA dose titration

The estimated cost for ERT in a ten-year-old MPS I patient is approximately $400,000 annually [71]. This financial burden provides strong impetus for health authorities to encourage using the minimum effective dose for patients. If serum HCII-T complex concentration does correlate with disease severity, then this may be used to titrate the minimum necessary dose.

5.5.3 Sample preparation and blood GAG concentration fluctuation

Several aspects of how patient serum sample preparations affect the concentration of the biomarker remain to be answered. First, since forming serum depletes functional thrombin [72] and patient serum samples demonstrate considerable residual HCII, it must be assumed inherent thrombin levels affect biomarker concentration. In fact, prothrombin, the circulating thrombin precursor, varies from 1 – 2 nmol/mL in healthy individuals [73]. Although this should not cause patient HCII-T complex levels to
overlap the normal range, it will impact apparent disease severity, as well as the utility of
the biomarker for making fine comparisons between patients and for following treatment
efficacy. Furthermore, as demonstrated by HCII-T complex changes in untreated patients
in Chapter 4, my results show an individual can have large complex concentration shifts
over time. How this relates to serum preparation and natural variability needs to be
addressed.

5.5.4 Analysis of the HCII-T complex biomarker in serum of IDUA overexpressing mice
for use in gene transfer

Another area to be explored using the murine MPS I model is the effect of gene
transfer on serum HCII-T complex concentration. As researchers understand and solve
more of the impediments to gene transfer therapies, this promising field becomes a closer
reality. Metabolic diseases provide an excellent target for gene transfer because the
phenotype primarily results from a specific deficiency. In addition, IDUA (and the other
MPS deficient enzymes) is targeted specifically to the lysosome via the mannose-6-
phosphate receptor pathway [74], so gene-based therapies have the potential to be highly
effective in MPS I even when targeted to a single organ. Our lab has a conditional IDUA
expressing mouse on the Idua<sup>−/−</sup> background that shows no visible phenotype by 30 weeks
of age, so it will be interesting to investigate HCII-T complex concentrations over the
course of an animal’s lifetime.
5.5.5 HCII-T complex as a plasma biomarker

As mentioned, the relationship of plasma HCII-T concentration to MPS I severity is unknown, although I hypothesize it reflects the level of extravascular disorder caused by GAGs. Plasma biomarker levels may more accurately represent tissue involvement than serum HCII-T concentration. The MPS I mouse model, which shows elevated plasma HCII-T complex, presents a study methodology. Using either IDUA overexpressing mice, described above, or mice treated with ERT, the lysosome GAG concentration could be correlated with plasma HCII-T concentration at several ages or post-treatment intervals.

5.5.6 Further proteomic analyses

Although monitoring serum HCII-T complex will undoubtedly prove very useful in MPS I research, additional proteomic methods may be able to identify other biomarkers as well as discover pathophysiological events in tissues.

It would be interesting to follow my depleted serum iTRAQ experiments with depleted serum ICAT labeling to verify whether ICAT could detect more signalling molecules. As mentioned, these molecules may be more informative with respect to the physiological changes occurring in the tissues and the extracellular responses accompanying disease progression.

Increased sample fractionation may also reveal abnormal protein states in the MPS serum proteome. In all likelihood, using solution phase isoelectric focusing (IEF) on the depleted samples I produced would be ineffective since every fraction contains second tier high abundance proteins that occlude candidate biomarkers. Instead,
selecting HPLC immunodepletion columns that remove additional high abundance proteins may be a more effective initial step, followed by IEF fractionation. Caution will be needed to minimize artefacts due to sample handling throughout multiple procedures, so it may be preferable to incorporate a differential 2D gel electrophoresis (DIGE) protocol. DIGE uses dyes, typically Cy3 and Cy5, to label the protein samples before IEF in order to subject each sample to identical conditions; when gels are visualized, the relative colour intensities can be used to determine whether proteins are up- or down-regulated.

Tissue proteomics can be extended to tissues displaying abnormal pathology in MPS: liver, spleen, brain, heart, respiratory tract, bone, and eye. There are obviously damaging processes occurring in these organs, but how elevated GAGs cause this is unknown. Investigating HCII and other candidate serpins known to interact with GAGs may provide insight into disease pathophysiology and introduce novel therapeutic targets for alleviating disease symptoms. An advantage of 2DGE over standard SDS-PAGE for investigating these abnormal protein interactions, aside from increased protein loading limits, is that different serpin-protease interactions appear at unique isoelectric points as well as at various molecular mass positions, which affords enhanced resolution of multiple complexes.

Another option is targeting known serpin-inhibited proteases. For example, thrombin has multiple functions outside of coagulation, such as in neurological development [75,76], connective tissue remodeling [77], and cell signaling [78], so it is a plausible candidate for involvement in MPS tissue dysfunction. Various serpin-thrombin complexes may be involved in separate tissues. Therefore, isolating serpin-thrombin
complexes could provide additional information and be more useful than analyzing whole tissue proteomes.

Finally, MPS disorders are members of the lysosomal storage disorder family, making it interesting to apply hypotheses generated here about enzyme inhibition and protein interactions to other lysosomal enzymes. Since GAGs accumulate in the lysosomes, the initial pathophysiological changes would be expected to occur in this organelle. In particular, HCII inhibits cathepsin G, a lysosomal serine protease [79]; investigating whether HCII-cathepsin G covalent complexes are prevalent in the lysosomes of MPS I mice would be a valuable experiment, especially considering the reported relationships between cathepsin G, connective tissue remodeling [80,81], inflammation [82], and cardiovascular and cerebrovascular diseases [83], aspects relevant to MPS pathophysiology.
References


Appendix 1: iTRAQ Replicability Analysis

A1.1 Methods

A1.1.1 Computations

iTRAQ ratios were calculated by the Interrogator™ Algorithm (Applied Biosystems, Foster City, CA). All supplementary statistical analyses were performed using Microsoft Excel 2003 (Redmond, WA).

A1.1.2 Terminology

Absolute iTRAQ ratio change: \(|(iTRAQ \text{ Acquisition A+B fold-difference}) - (iTRAQ \text{ Acquisition A fold-difference})|\)

Fold-difference: multiplicative factor describing the abundance difference between two pools

iTRAQ ratio change: \((iTRAQ \text{ Acquisition A+B fold-difference}) - (iTRAQ \text{ Acquisition A fold-difference})\)

Peptide confidence: % confidence value assigned to peptides based on closeness of match to a peptide MS spectrum

Unused Protein Score: cumulative measure of Peptide Scores assigned to one protein exclusively to indicate probability of correct protein identification; four 99% confidence peptides would produce an Unused Protein Score of 8, indicating the probability of an incorrect protein identification is \((1 - 0.99)^4 = 10^{-8}\).
A1.2 Results

A1.2.1 Like-pool comparisons

The like-pool comparisons in my study indicate that iTRAQ provides reproducible results. When iTRAQ ratios for the two mutant pools are analyzed at the individual protein level, it is clear the majority of protein pairs are at approximately the same concentration (Fig. A1.1). In fact, the average (± SD) of absolute values of all abundance ratios is $1.26 \pm 0.54$. Thus, the average protein ratio falls well within a value expected of two similar pools. In an ideal sample, the average fold-difference would be 1.0, but since all ratios are expressed as a value greater than or equal to 1, it is unsurprising the average ratio exceeds 1.0. Furthermore, this mean is moderately affected by the six proteins that show abnormally high fold-differences (greater than 2.0), indicated by a substantially lower median of 1.14. Disregarding these six values, the average declines to 1.19 and the median remains 1.14. Five of these six anomalous values occurred in proteins identified on the basis of a small number of peptides and have a large confidence interval containing their true abundance; as a result, the ratios may truly be less than 2.0. Based on the conclusions taken from Fig. 3.1, the \(I\text{dua}^{+/+}\) male:\(I\text{dua}^{+/+}\) mixed sex comparison would show the same pattern as the mutant pool comparison.
Fig. A1.1: Individual protein ratios, expressed as absolute values of the fold-differences between the ldua° male:Idua° mixed sex ratio. Proteins are ranked in decreasing order of identity confidence. Notable that only 6 of 198 proteins showed a ratio greater than two-fold increased or decreased. Dashed bars indicate mean (1.26) and median (1.14) fold-differences. A value of 1.0 indicates a protein is present at the same concentration in both pools.

A1.2.2 Data re-acquisition

An additional element demanding evaluation is the effect of re-acquiring peptide spectra from a single sample. To further address this, I considered how detecting new peptides altered iTRAQ ratios of every pool from iTRAQ Acquisition A to iTRAQ Acquisition A+B. 62% of proteins identified in Acquisition A had additional peptides observed, increasing identification confidence. Changes in Unused Protein Score (a cumulative measure of peptide confidences assigned solely to the protein in question) ranged from -13.63 to 46.55, where a negative change indicates peptides were re-
assigned to different proteins and a positive change describing how many additional peptides were used to identify the protein (Fig. A1.2). Disregarding the proteins with negative and null scores, the median increase in Unused Protein Score was 4.42, which is equivalent to adding at least three new peptides.

Fig. A1.2: Change in Unused Protein Score from iTRAQ Acquisition A to Acquisition A+B
All proteins that appeared in iTRAQ Acquisition A+B were considered. Proteins that show a positive change in Unused Protein Score acquired additional peptides, while a negative score indicates peptides were re-assigned. Dashed line indicates median Unused Protein Score change.

Considering all proteins identified with 94% confidence or greater, 88% of proteins identified in Acquisition A were present in Acquisition A+B, indicating the peptides from 12% of previous identifications were re-assigned to different proteins upon supplementary peptide acquisition. The majority of these changed identities were in proteins originally identified with less than 99% confidence, and all exceptions were proteins belonging to protein families displaying high degrees of amino acid sequence
homology. This is logical for the program, but it precludes the presence of a protein belonging to a homology family that is in lower abundance than another family member. If a signature peptide unique to one family member is not identified, no peptides will be attributed to the protein and will instead be assigned to one or more recognized proteins, which may adversely affect abundance ratios and their confidence intervals. Since most proteins identified in my study had limited sequence coverage, signature peptides may go undetected.

The effect of acquiring additional peptides on abundance ratios is also of interest. To prevent negative and positive changes in iTRAQ ratios from producing an artificially low average change value, I calculated absolute values of the changes. Thus, the average change in iTRAQ ratio, over all three comparisons, was 0.0615. The effects of this deflection on the accuracy of hypothetical abundance ratios obtained in the initial acquisition are shown in Table A1.1.

<table>
<thead>
<tr>
<th>Absolute value of initial iTRAQ fold-difference (Acquisition A)</th>
<th>Absolute value of adjusted iTRAQ fold-difference (Acquisition A+B)</th>
<th>% error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.062</td>
<td>6.2%</td>
</tr>
<tr>
<td>2.0</td>
<td>2.062</td>
<td>3.1%</td>
</tr>
<tr>
<td>5.0</td>
<td>5.062</td>
<td>1.2%</td>
</tr>
</tbody>
</table>

Clearly, adding this factor has a small effect on observed iTRAQ ratios and suggests a protein’s initial abundance ratio was accurate to within 6.2%, on average. Although several values exceeded this error, 84% of proteins displayed errors less than 6.2%, when the absolute change in fold-difference occurring because of supplementary
peptide collection is divided by the absolute fold-difference determined in Acquisition A. No proteins determined to be significantly increased or decreased in my iTRAQ data had % errors exceeding 5% between Acquisition A and Acquisition A+B.

The mean iTRAQ ratio change from Acquisition A to Acquisition A+B appears to be affected by a few extreme cases (maxima were 0.6654, 0.2494, and 0.4797 for the Idua<sup>−/−</sup> mixed sex, Idua<sup>+/−</sup> male, and Idua<sup>−/−</sup> male pools in comparison to the Idua<sup>−/−</sup> mixed sex pool, respectively), as median ratio changes were approximately half the mean in all pools. Upon analyzing where the abundance ratios changed by a value greater than the median plus one standard deviation, it became apparent all cases were originally identified on the strength of either one high confidence peptide or a collection of weaker confidence peptides. Therefore it is not surprising additional peptides impacted these ratios.

Another telling statistic is that of the 22 proteins identified in iTRAQ Acquisition A by one peptide, only 7 acquired new peptides following the second data acquisition; 14 of the remainder had no gain and one was reassigned to a different protein. Furthermore, the observed changes in Unused Protein Score follow roughly a first order decay curve as the protein confidence decreases (Fig. A1.2), indicating the bulk of the 569 new peptides matched to proteins generally strengthened the high confidence protein identities, which is not particularly informative. This is shown graphically in Fig. A1.3; a best fit to the data indicates the average protein improved its Unused Protein Score by 40%.
Fig. A1.3: Change in Unused Protein Score versus Unused Protein Score from iTRAQ Acquisition A.

Plot shows linear fit of the change in Unused Protein Score (UPS) as a function of Unused Protein Score obtained in the initial iTRAQ data acquisition. A logarithmic scale was used to more clearly indicate the relationship between change in UPS and initial UPS. Exceptions are noticeable, but most large increases in UPS belonged to identifications initially achieved with multiple peptides.
Appendix 2: Mucopolysaccharidosis diagnostics patent

The following is a provisional patent submitted on the measurement of serpin-protease complexes for mucolopolysaccharidosis diagnosis

Application #: US60/748,210

Submission date: December 8, 2005
Receipt is acknowledged of this provisional Patent Application. It will not be examined for patentability and will become abandoned not later than twelve months after its filing date. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please mail to the Commissioner for Patents P.O. Box 1450 Alexandria Va 22313-1450. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).

Applicant(s)
Lorne A. Clarke, Vancouver, CANADA;
Graham Sinclair, Vancouver, CANADA;
Derek R. Randall, Vancouver, CANADA;

Power of Attorney:
Ralph Dowell–26868

If Required, Foreign Filing License Granted: 01/09/2006

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is US60/748,210

Projected Publication Date: None, application is not eligible for pre-grant publication

Non-Publication Request: Not applicable

Early-Publication Request: Not applicable

** SMALL ENTITY **

Title
Mucopolysaccharidosis diagnostics
PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process simplifies the filing of patent applications on the same invention in member countries, but does not result in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help “toolkits” giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER
Title 35, United States Code, Section 184
Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls,
Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

**NOT GRANTED**

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).
MUCOPOLYSACCHARIDOSIS DIAGNOSTICS

FIELD OF THE INVENTION

BACKGROUND OF THE INVENTION

Lysosomal storage disorders (LSDs) generally describes a group of more than 40 diseases resulting from a variety of enzyme deficiencies. Children are the most frequently affected, as the clinical presentation often occurs at an early age. The common theme among these diseases is accumulation of specific macromolecules within tissues and cells, leading to clinical features including liver and spleen enlargement, neurologic effects such as mental retardation, skeletal dysplasia, ophthalmological abnormalities and hematological effects such as granulation and vacuolation of lymphocytes. Subgroupings of LSDs based on the nature of the macromolecule that accumulates include mucopolysaccharidoses (MPS), gangliosidoses, glycosphingolipidoses, glycoproteinoses, mucolipidoses, leukodystrophies and disorder of neutral lipids.

MPS account for the largest single subgroup of LSDs. The MPS are inherited disorders, resulting from insufficient levels of specific lysosomal enzyme activity, involved in degradation of glycosaminoglycans (GAGs). Accumulation of these undegraded or partially degraded GAGs interferes with the normal function of cells, tissue and organs and affects the normal growth and development of the individual.

The overall prevalence of MPS is estimated at about 1/20,000 (Meikle PJ et al. 1999. Prevalence of lysosomal storage disorders JAMA. 281:249-54). A myriad of mutant alleles have been identified for each of the MPS disorders, making genetic screening by itself of limited value. Genotype/phenotype correlation does exist for some deficiencies, however the variety of mutations for each deficiency makes this prediction difficult in a considerable proportion of patients (Terlato NJ and Cox GF. 2003. Can mucoplysaccharidosis type 1 disease severity be predicted based on a patient’s genotype: A comprehensive review of the literature, Genet. Med. 5:286-294). Initial diagnostic inquiries are based on clinical phenotype presentation, and this may be used to guide selection of further biochemical assays.

Highly specific substrates and assays for enzyme activity have been developed to assist the diagnostic process, but these often require invasive tissue sampling and may be time consuming as cells are cultured to provide sufficient quantity for screening. In very young children specific diagnosis may not be possible until the disease has advanced sufficiently to result in damage that may be irreversible, even if a definitive diagnosis is made and an effective therapy initiated as early as possible.

Assays for urinary GAGs are known and have been used extensively for screening, diagnosing and monitoring MPS patients, however the levels of urinary GAGs may not correlate with disease severity (Byers S et al., 1998. Glycosaminoglycan accumulation and excretion in the mucopolysaccharidoses: characterization and basis of a diagnostic test for MPS. Mol. Genet. Metab. 65:282-290; Gallegos-Arreola MP et al., 103


It is an object of the present invention to provide a method for diagnosing MPS disorders and monitoring clinical progress in response to therapy.

SUMMARY OF THE INVENTION

In accordance with one embodiment of the invention, there is provided a method of diagnosing an MPS, the method comprising obtaining a tissue sample from a subject and determining the quantity of a biomarker of an MPS present in a tissue sample.

In accordance with another aspect of the invention, there is provided a method of predicting the responsiveness of a subject having, or suspected of having an MPS, to a therapeutic regimen for treatment of an MPS, the method comprising determining the quantity of a biomarker of an MPS in a first sample relative to a second sample, wherein the quantity is indicative of subject’s responsiveness to a therapeutic regimen for treatment of an MPS.

In accordance with another aspect of the invention, there is provided a method of distinguishing between MPS, or distinguishing phenotypes of a single MPS, comprising determining the quantity of a biomarker of an MPS in a first sample relative to a second sample, wherein the quantity is indicative of the subject’s responsiveness to a therapeutic regimen for treatment of an MPS.

In accordance with another aspect of the invention, there is provided a method for selecting a group of subjects for determining the efficacy of a therapeutic regimen known, or suspected of being useful for the treatment of an MPS, the method comprising determining the quantity of a biomarker of an MPS in a first sample relative to a second sample, wherein the quantity is indicative of the subject’s sensitivity to a therapeutic regimen.

In accordance with another aspect of the invention, there is provided a method of treating an MPS in a subject in need thereof, the method comprising selecting a subject having a
biomarker profile indicative of sensitivity to a therapeutic regimen for MPS and administering to the subject a therapeutic regimen.

In accordance with another aspect of the invention, there is provided a method of treating an MPS in a subject in need thereof, the method comprising selecting a subject having a biomarker profile indicative of resistance to a therapeutic regimen for MPS and administering to the subject the therapeutic regimen.

In accordance with another aspect of the invention, there is provided a kit for determining the quantity of a biomarker of an MPS in a first sample relative to a second sample, the kit comprising an antibody specific for detecting of a biomarker of MPS in a tissue sample.

The biomarker may be a serpin, for example heparin cofactor II or antithrombin III.

A first sample may be a tissue sample from an MPS affected subject or a subject suspected of having an MPS. A second sample may be a tissue sample from an unaffected or control or normal subject. Tissue samples may include cells, blood, serum, muscle, bone, neurological tissue, saliva, urine, mucus or other sample acquired in biopsy.

A subject may have a tissue sample tested once, or several times for a biomarker. Further a subject may be tested at any age, including prenatally, newborn, infancy, childhood or adulthood.
Other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments of the invention in conjunction with the accompanying figures.

**BRIEF DESCRIPTION OF THE DRAWINGS**

Figure A2.1 depicts box plots of relative abundances for proteins identified with 99% confidence in each sample pool relative to controls. Grey boxes cover the interquartile range, whiskers extend to 10th and 90th centiles, and individual points indicate the outliers.

Figure A2.2 is a logarithmic plot of average relative abundance of proteins in $Idua^{+/+}$ pools vs. $Idua^{-/-}$ mixed sex, ≥99% confidence, with proteins rank ordered from left to right based on descending Protein % Confidence score. Dashed lines indicate a two-fold increase or decrease in the average relative abundance of a protein present in the $Idua^{-/-}$ pools compared to the $Idua^{+/+}$ mixed sex pool. Proteins marked in red were considered candidates based on significant changes in relative abundance and are listed on Table 2.

Figure A2.3 shows a western blot of serum samples from WT ($Idua^{+/+}$) and MPS ($Idua^{-/-}$) mice for HCII. 15 µg of sample was applied per well.

Figure A2.4 is a western blot of MPS patients for HCII. 2.5 µg of protein was loaded in wells 1 – 7 and 8 ng of protein was loaded in wells 8 and 9.

Figure A2.5 shows HCII-T levels in MPS IH serum samples following enzyme replacement therapy. a) Hurler patient, 12 months of age at start of ERT. All lanes contain 2.5 µg of protein. The patient received a bone marrow transplant between 36 and 52 weeks. b) Hurler-Scheie patient, 8 years of age at start of ERT. All lanes contain 2.5 µg of protein.

Figure A2.6 shows a western blot of MPS patient samples for ATIII. 10 µg of protein was loaded per well.

**DETAILED DESCRIPTION**

Patients diagnosed with an MPS are deficient in the ability to break down specific glycosaminoglycans into simpler sugars and amino acids. This deficiency results from insufficient production or targeting of the enzyme, or production of an inactive enzyme. A variety of heritable mutations may be involved.

The MPS share many clinical features with varying degrees of severity and phenotypic effects. Table 1 lists the 10 MPS disorders, their specific enzyme deficiency and the GAG stored in the lysosome as a result of the enzyme deficiency.

106
Table A2.1: Mucopolysaccharidoses and the respective enzyme deficiency and stored glycosaminoglycan.

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Enzyme Deficiency</th>
<th>Stored Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPS I H, (Hurler)</td>
<td>alpha-L-Iduronidase</td>
<td>Dermatan sulfate</td>
</tr>
<tr>
<td>MPS I H/S, (Hurler/Scheie)</td>
<td></td>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>MPS I S, (Scheie)</td>
<td>Iduronate sulfatase</td>
<td>Dermatan sulfate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>MPS II (Hunter)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPS III A (Sanfilippo A)</td>
<td>Heparan-N-sulfatase</td>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>MPS III B (Sanfilippo B)</td>
<td>N-Acetyl-alpha-glucosaminidase</td>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>MPS III C (Sanfilippo C)</td>
<td>Acetyl-CoA: alpha-glucosaminide N-acetyltransferase</td>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>MPS III D (Sanfilippo D)</td>
<td>N-Acetylglucosamine-6-sulfate sulfatase</td>
<td>Heparan sulfate</td>
</tr>
<tr>
<td>MPS IV A (Morquio)</td>
<td>N-Acetylgalactosamine-6-sulfate sulfatase</td>
<td>Keratan sulphate</td>
</tr>
<tr>
<td>MPS VI B (Morquio)</td>
<td>Beta-Galactosidase</td>
<td>Chondroitin-6 sulfate</td>
</tr>
<tr>
<td>MPS VI (Maroteaux-Lamy)</td>
<td>Arylsulfatase B</td>
<td>Dermatan sulfate</td>
</tr>
<tr>
<td>MPS VII (Sly)</td>
<td>Beta-Glucuronidase</td>
<td>Dermatan sulphate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heparan sulfate</td>
</tr>
</tbody>
</table>
Definitions

Any terms not directly defined herein shall be understood to have the meanings commonly associated with them as understood within the art of the invention. As employed throughout the specification, the following terms, unless otherwise indicated, shall be understood to have the following meanings.

As used herein, a ‘glycosaminoglycan’ (GAG) is a long, unbranched polysaccharide molecule, composed of repeating disaccharide units. The first sugar residue in the repeating disaccharide is an amino sugar, such as N-acetylglucosamine or N-acetylgalactosamine, and is usually sulfated. The second sugar residue is a uronic acid, such as glucuronic or iduronic acid. Four groups of GAGs include hyaluronan, chondroitin sulfate and dermatan sulfate, heparan sulfate and heparin, and keratin sulfate. GAGs may be covalently linked to proteins in the form of proteoglycans, and are major structural components of connective tissue such as cartilage, and of the cornea of the eye. An alternate name for a glycosaminoglycan is a mucopolysaccharide.

As used herein, a subject refers to an animal, such as a mammal. A subject may be a mouse, or other experimental animal such as a dog, or may refer to a human (a ‘patient’). A human subject may be diagnosed with an MPS, suspected of having an MPS, or may be undiagnosed. The mouse or other experimental animal may be ‘wild-type’, or may be a transgenic animal. Examples of transgenic mouse models for LSDs include those described in WO 2005/080574 or US 6002067. A ‘normal’ or ‘control’, or ‘unaffected’ subject refers to a subject that is unaffected with an MPS.

As used herein, a phenotype refers to the physical manifestation of a subject, including anatomical, metabolic and behavioural traits. Phenotype may result from heredity, or the environment of the subject, or a combination of these factors.

As used herein, a mucopolysaccharidosis (MPS) refers to a subgroup of lysosomal storage disorders (LSD). MPS are characterized by the accumulation and storage of GAG within lysosomes. An MPS phenotype refers to the clinical signs or symptoms of an MPS in a subject. The clinical signs or symptoms may be varied, depending on the severity and specific MPS disorder.

MPS IH (Hurler syndrome) is an autosomal recessive disorder resulting from numerous different mutations of alpha-L-iduronidase. Progressive mental retardation, hepatosplenomegaly, skeletal malformations and cardiopulmonary compromises typically lead to death in the first decade. Affected individuals appear normal at birth, with the characteristic appearance and accelerated growth developing in the first year. Clinical diagnosis may be suggested in the first 2 years by, for example, hepatosplenomegaly, corneal clouding, coarse features and joint problems. Developmental delay is observed between the first and second years, with subsequent slow mental development and/or regression. Additional complications of this disorder include hearing loss, chronic respiratory infections, valvular heart disease and brain ventricular enlargement.
MPS IS (Scheie syndrome) and MPS IH/S (Hurler/Scheie syndrome) are less severe variants of MPS IH. MPS IS subjects may survive into late adulthood, although severe progressive skeletal disease that resembles osteoarthritis is observed. Normal intelligence is also found.

MPS II (Hunter syndrome) is an X-linked recessive disorder resulting from mutations of iduronate sulfatase. Clinical manifestations of MPS II range from severe CNS and involvement of the viscera with death in late childhood, to milder forms having normal CNS function and survival into adulthood.

MPS III A, B, C and D (Sanfilippo syndromes) are autosomal recessive disorders resulting from various enzyme deficiencies as noted in Table 1. Skeletal defects and hepatosplenomegaly are less pronounced than in MPS I and II, however progressive behavioural problems, mental retardation and seizures are observed. Affected subjects may survive into early to mid-adulthood.

MPS IV (Morquio syndrome) is an autosomal recessive disorder characterized by deficiency in N-acetylgalactosamine-6-sulfate (type A) or beta-galactosidase deficiency (type B). Type A presents the more clinically severe skeletal disease. Extreme shortening of the trunk may occur due to vertebral collapse, and joint laxity may lead to osteoarthritis-like damage of the joints. Paralysis may also result from instability of the upper cervical spine compressing the spinal cord. Mitral valve insufficiencies may also be present.

MPS VI (Maroteaux-Lamy syndrome) is an autosomal recessive disorder resulting from mutations in the arylsulfatase B gene. The general phenotype resembles MPS IH, although it may be clinically variable. Intelligence is normal and life span may last to early to mid-adulthood. Valvular disease and progressive pulmonary hypertension may be present and may be a frequent cause of death.

MPS VII (Sly syndrome) is an autosomal recessive disorder resulting from mutations in the beta-glucuronidase gene. Clinical symptoms include mental retardation, short stature, skeletal dysplasia, hepatosplenomegaly. Survival into adulthood may occur in milder cases, with osteoarthritis-like joint complications common.

Other examples of clinical signs or symptoms of an MPS will be known to a physician versed in the art, and may be found in, for example “Neufeld EF and Muenzer, J. In The metabolic and molecular basis of inherited disease 8th edition, CR Scriver et al, editors. McGraw-Hill, NY 2005. pp 3421-3452”

As used herein, an MPS biomarker refers to a marker that is associated with the diagnosis of an MPS in a subject. The biomarker may be a protein or proteins, a complex of a protein with another molecule, such as an oligosaccharide or GAG, a genetic sequence, marker or mutation. The biomarker may be present in greater or lesser levels than in a subject unaffected with an MPS (a ‘normal’ subject). The biomarker may be associated
with the onset of the clinical phenotype of the MPS, or may be detectable before clinical onset of the MPS phenotype.

As used herein, a ‘gene’ is an ordered sequence of nucleotides located in a particular position on a particular chromosome that encodes a specific functional product and may include untranslated and untranscribed sequences in proximity to the coding regions (5’ and 3’ to the coding sequence). Such non-coding sequences may contain regulatory sequences needed for transcription and translation of the sequence or introns etc. or may as yet to have any function attributed to them beyond the occurrence of the mutation of interest.

A “mutation” as described herein may be the result of a “single nucleotide polymorphism” (SNP) occurring at a polymorphic site occupied by a single nucleotide, which is the site of variation between allelic sequences. A single nucleotide polymorphism may arise due to substitution of one nucleotide for another at the polymorphic site. A “transition” is the replacement of one purine by another purine or one pyrimidine by another pyrimidine. A “transversion” is the replacement of a purine by a pyrimidine or vice versa. A mutation may also arise from a deletion of one or more nucleotides or an insertion of one or more nucleotides relative to a reference sequence of a particular gene. Alternatively, a mutation may result in a frameshift of the sequence resulting from a deletion or insertion as described above or from an inversion etc. Additionally, a mutation as described herein may be a multisite mutation, whereby the mutation is comprised of two or more mutations.

As used herein, a ‘vector’ includes any means for delivery of a nucleic acid sample to a cell. For example, a vector may include a plasmid, artificial chromosome, a virus, etc.

As used herein, gene therapy refers to delivery of a vector to a cell, wherein the vector comprises a gene of interest and regulatory sequences for expression of the gene of interest. Methods of delivering the vector to the cell are known in the art, and may include, for example, direct injection, or by transfection, or infection in the case of a viral vector. Examples of viral vectors are known in the art, and may include adenovirus, adeno-associated virus, lentivirus, poxvirus, herpesvirus.

As used herein, a ‘serpin’ is a member of the serine protease inhibitor superfamily of proteins. Serpins specifically inhibit serine proteases as part of the regulation of various metabolic pathways in living cells and whole organisms. Examples of serpins include those in the coagulation pathway, such as thrombin, antithrombin, heparin cofactor II, faktor Xa, protease nexin I, pigment epithelium-derived factor (PEDF) and protein C inhibitor. Each serpin regulates the activity of its protease through a “suicide” mechanism by which the protease initiates proteolysis of the serpin but cannot complete the reaction process, thereby forming a covalent association between the two proteins (Gettins, PG. Serpin structure, mechanism and function Chem Rev. 102:4751-4804). This complex formation is essentially irreversible and causes a dramatic structural change in the protease leading to its inactivity. The HCII-T complex is subsequently recognized by specific receptors and degraded in the liver.
As used herein, a ‘sample’ or a ‘tissue sample’ refer to a small quantity of tissue or body fluid from a subject affected, suspected of being affected or unaffected with an MPS disorder. Samples may be obtained by curettage, needle aspiration biopsy or needle (core) biopsy, incisional biopsy for sampling of tissue at a specific site, or by excisional biopsy. Samples may include cells, blood, serum, muscle, bone, neurological tissue, saliva, urine, mucus or other sample acquired in biopsy. Samples such as blood, saliva, urine, mucus, etc may be collected using methods known in the art.

**Glycosaminoglycan degradation**

The deficient enzymes in the various MPS (Table 1) are part of a highly ordered degradation pathway for GAGs. Generally, the long chain GAG is cleaved into smaller fragments by endoglucuronidases or endohexosaminidases, and the individual monosaccharides subsequently removed by specific enzymes. It is these specific enzymes that are deficient in the various MPS. Details of the pathways for degradation of specific GAGs, including the enzymes involved at each step and the substrates of each may be found in, for example, “Essentials of Glycobiology, A. Varki, R. Cummings, et al., eds. 1999. Cold Spring harbor Laboratory Press, Cold Spring Harbor, NY.

**Diagnosis of MPS**

Clinical phenotypes of MPS are not usually evident at birth, but may appear within a few months of birth, or as late as several years of age. During infancy and childhood, physical and mental development may be affected or delayed. Clinical symptoms such as short stature, bony dysplasia, hirsutism may be observed, as well as more characteristic facies of the MPS disorders, such as thick lips, open mouth or flattened nasal bridge. Depending on the specific MPS, mental retardation may also be present. The specific clinical phenotype of each specific MPS will be known to a physician versed in the art, and may be found in, for example “The Merck Manual of Diagnosis and Therapy” 17th edition. MH Beers and R Berkow, editors. 1999-2005, Merck & Co. “Neufeld EF and Muenzer, J. In *The metabolic and molecular basis of inherited disease* 8th edition, CR Scriver et al, editors. McGraw-Hill, NY 2001. pp 3421-3452.

With the exception of MPS II, which is X-linked, all the MPS disorders are autosomal recessive. A wide variety of mutations, including point mutations, insertions, deletions and polymorphisms are observed and affected individuals may be heterozygous for their specific mutations – each parent contributing a different defective copy of the gene. Some correlation exists between the severity of the MPS and the specific mutation, but this may not be practical for widespread use as a diagnostic, given the heterogeneous genetic makeup of the affected population.

Diagnosis of specific MPS may be made using a combination of observed clinical phenotype, urinary analysis for GAGs and enzyme assays. For example, MPS I results from a deficiency of the enzyme α-L-iduronidase (IDUA; EC 3.2.1.76) (Neufeld supra). IDUA deficiency results in abnormalities in the degradation of the glycosaminoglycans
(GAGs) heparan sulphate and dermatan sulphate, which subsequently accumulate in the lysosome. Severity of the disorder is estimated subjectively, based on clinical phenotype.

Therapies for MPS

Few treatment options exist for subjects affected with MPS. Supportive therapy and/or palliative care may be offered largely to improve quality of life of the subject. With each MPS disorder, supportive management of the clinical manifestations is provided. For example, patients presenting with chronic respiratory complications would be treated for the frequent infections and congestion of the chest and airway, but the underlying cause (the buildup of the GAG in the lysosome) cannot be addressed in this manner.

Bone marrow transplantation is a therapeutic option for some MPS subjects, although the efficacy of this procedure varies with the severity of the disease.

Enzyme replacement therapy has been shown to be a useful therapeutic approach in some MPS subjects. For example, in MPS I affected subjects, clinical studies have demonstrated that administration of recombinant alpha-L-iduronidase (ALDURAZYME™; U.S. Patent No. 6426208) can alter the phenotype of MPS I patients to varying degrees (Wraith JE et al., 2004. *Enzyme replacement therapy for mucopolysaccharidosis I: a randomized, double-blinded placebo-controlled, multinational study of recombinant human alpha-L-iduronidase (laronidase)* J. Pediatr. 144: 581-588).

Success of the therapeutic approach may be estimated subjectively, based on clinical phenotype.

Monitoring of disease progress or therapeutic efficacy

The progress of MPS may be monitored by the alterations in clinical phenotypes. Analysis of urinary GAGs may provide some guidance, but does not correlate specifically with disease severity.

One embodiment of the present invention provides for the use of a serpin biomarker to monitor disease severity. Serpins may be regulated by GAGs, which accelerate the interaction between serpins and their target coagulation proteases. Examples of such serpins may be found in Table 2. Serum levels of heparin cofactor II-thrombin (HCII-T) complex in MPS I patients are elevated well beyond that seen in control serum samples and provide a biomarker of MPS disease.

<table>
<thead>
<tr>
<th>Serpin</th>
<th>GAGs that modulate activity</th>
<th>Known target proteases</th>
<th>Tissue specificity</th>
<th>Biological function</th>
</tr>
</thead>
</table>

Table A2.2: Examples of serpins with GAG-mediated activities.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Location</th>
<th>Inhibitors</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>α₁-antitrypsin (α₁AT, α₁PI, SERPINA1)</td>
<td>HS, DS</td>
<td>Neutrophil elastase, Plasmin, Thrombin</td>
<td>General inhibitor, inflammation [17,38,39]</td>
</tr>
<tr>
<td>Protein C inhibitor (PAI-3, SERPINA5)</td>
<td>HS</td>
<td>Thrombin, activated Protein C, Urinary plasminogen activator, Tissue plasminogen activator, Plasma kallikrein, Acrosin</td>
<td>Coagulation [17]</td>
</tr>
<tr>
<td>Antithrombin III (ATIII, SERPINC1)</td>
<td>HS</td>
<td>Thrombin, Factor IXa, Factor Xa</td>
<td>Coagulation [17]</td>
</tr>
<tr>
<td>Heparin cofactor II (HCII, SERPIND1)</td>
<td>HS, DS</td>
<td>Thrombin, Cathepsin G, Chymotrypsin</td>
<td>Coagulation, unknown [17]</td>
</tr>
<tr>
<td>Glia-derived nexin (PN1, SERPINE2)</td>
<td>HS</td>
<td>Thrombin, Acrosin, Urinary plasminogen activator, Tissue plasminogen activator, Fibroblasts, neurons, extracellular space</td>
<td>Neurotrophic, ECM remodelling [17,34,35]</td>
</tr>
<tr>
<td>Pigment epithelium-derived factor (PEDF, SERPINF1)</td>
<td>Factor XIa</td>
<td>Cornea, cartilage, bone</td>
<td>Neurotrophic, antiangiogenic [17,36]</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>HS</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Methods

Sample Collection

Mouse serum samples were collected by cardiac puncture and added to Beckton Dickinson serum separator tubes (Franklin Lakes, NJ). Samples were allowed to clot for 30 minutes at room temperature and centrifuged for 15 minutes at 3000xg at room temperature. Samples were aliquoted immediately and stored at -80°C. Mouse plasma samples were collected by cardiac puncture, with sodium citrate added to 0.4%. Citrated samples were centrifuged for 15 minutes at 1500xg at 4°C. Human plasma and serum samples were collected by clean venipuncture using a 2 syringe technique, as described by Petzer (Petzer H et al 1988. Determination of human thrombin-antithrombin III complex in plasma with an enzyme linked immunosorbent assay. Thromb. Haemos. 59:101-106), and processed as above.

High Abundance Protein Depletion

Albumin, immunoglobulin, and transferrin were depleted from murine serum samples using a 100 mm Ms-3 Multiple Affinity Removal System (Agilent Technologies, Palo Alto, CA), according to the manufacturer’s instructions. 4 age- and sex-matched pools of mouse serum samples were normalized for total protein to 12 mg/mL by diluting them in running buffer (Agilent Technologies). Each pool contained 6 independent serum samples. A KD Scientific syringe pump (Holliston, MA) was used to maintain a constant flow rate of 15 mL/h. 320 μL were subsequently used for high abundance protein depletion with identical pool collections obtained for all samples.

iTRAQ Analysis

iTRAQ reagent labelling

The depleted mouse serum samples were labelled with the iTRAQ reagent following the manufacturer’s protocol (Applied Biosystems, Foster City, CA). Briefly, 100 μg of total protein from each of the 4 depleted serum pools was precipitated with acetone and resuspended in iTRAQ dissolution buffer for reduction, alkylation, and tryptic digestion. Each of the resulting peptide pools was then labelled with a different isotopic iTRAQ Reagent (114-117 Da) as follows: *Idua* mixed sex pool (114 Da), *Idua* male pool (115 Da).
Da), \( Idua^{+/+} \) male pool (116 Da) and \( Idua^{-/-} \) mixed sex pool (117 Da). The four differentially labelled pools were then combined and subjected to strong cation exchange (SCX) chromatography using a polysulfoethyl A column (Poly LC, Columbia, MD). The combined sample was diluted in 10 mM KPO\(_4\) (pH 2.7), 25% ACN, applied to the column and peptides eluted over a 33 minute gradient to 35% 10mM KH\(_2\)PO\(_4\), 25% CAN, 0.5 M KCl with fractions collected at one minute intervals.

**LC-MS/MS Analysis**

The resulting 33 SCX fractions were then subjected to LC-MS/MS analysis utilizing a QStar Pulsar hybrid quadrupole-TOF instrument (Applied Biosystems) and an UltiMate micro HPLC (LC Packings, Sunnyvale, CA). The HPLC was equipped with a C18 PepMap guard column (LC Packings) separated from a C18 Pepmap Nano LC column (LC Packings) by a switching valve to allow for precolumn sample clean-up before switching inline for reversed phase chromatography and MS/MS analysis. Each SCX fraction was evaporated to dryness, resuspended in 5% ACN and 3% formic acid, and 25% of the sample was injected onto the C18 guard column in 98% water/acetonitrile (98:2), 0.05% formic acid (Buffer A) with the HPLC flowing to waste to remove sample contaminants. Following 10 minutes at 100 \( \mu \)l/ml, the guard column was switched inline with the C18 resolving column and mass spectrometer, and the peptides were eluted with a linear gradient to 60% water/acetonitrile (2:98), 0.05% formic acid (Buffer B) over 40 minutes. Following a 5 minute ramp to 80% Buffer B, the column was re-equilibrated in 98% buffer A for 15 minutes prior to the injection of the next SCX fraction.

**Data Acquisition and Analysis**

MS data was acquired automatically using Analyst QS 1.0 software Service Pack 8 (ABI MDS SCIEX, Concord, Canada). An information-dependent acquisition method consisting of a 1 second TOFMS survey scan of mass range 400-1200 atomic mass units and two 2.5 second product ion scans of mass range 100 – 1500 atomic mass units was utilized. The two most intense peaks over 20 counts, with charge state 2 – 5 were selected for fragmentation and a 6 amu window was used to prevent the peaks from the same isotopic cluster from being fragmented again. Once an ion was selected for MS/MS fragmentation it was put onto an exclusion list for 180 seconds. Following the initial data acquisition run (iTRAQ acquisition A), an exclusion list was created for all peptides identified with a confidence of 95 % or greater in each of the SCX fractions. A second injection of each SCX fraction using the exclusion list for that fraction and the same LC and MS conditions as the first run was then performed in order to detect the lower concentration proteins present in plasma. (iTRAQ acquisition B).

The resulting data files were combined and processed using the Interrogator™ algorithm in the ProQuant software (v1.0) (Applied Biosystems) in Analyst using the following parameters: The MS and MS/MS mass tolerances were set to 0.20. A rodent subset of the Celera Discovery Systems Database (01/24/2004) was used for searching. Methyl methanethiosulphonate (MMTS) modification of cysteines was used as a fixed
modification. The number of missed cleavages was set to 1. All results were written to a Microsoft Access database and, to reduce redundancy, ProGroup Viewer version 1.0.5 (Applied Biosystems) was used to assemble and report the data. Protein % Confidence Scores, which are influenced by the closeness of the observed peptide spectrum to the predicted spectrum, the number of peptides identified for a given protein, as well as the variance of abundance of each peptide for the given protein, were used to calculate a Protein Score using the equation Protein Score = -log[1 – (Protein % Confidence)/100].

Abundances of each identified protein in the sample pools were calculated based on the abundance of reporter tags bound to the peptide tryptic fragments for each pool. Where more than one peptide was identified for a protein, the abundance was calculated by weighted averaging of the abundance of the individual peptides including a 95% confidence interval. Protein abundances were adjusted for global labeling bias by a correction factor that assumed an average relative protein abundance of 1 across all samples. Relative abundances are determined by dividing the abundance of each protein from each pool by the corresponding protein in another pool. All proteins found to have significantly altered relative abundances were manually verified by inspecting the corresponding peptide matches.

Western Blotting

Western blot analysis was performed on 7.5% Tris-glycine gels and transferred to Pall (East Hills, NY) BioTrace NT membranes. Anti-human HCII and anti-human antithrombin III antibodies were from Affinity Biologicals (Hamilton, ON). Membranes were blocked with 5% Carnation powdered skim milk, 0.05% PBS-T overnight at 4°C. Western blot analysis of human samples used primary antibody at a concentration of 1 μg/mL, with secondary antibody at a concentration of 1.3 μg/mL. Western blot analysis of mouse samples used 3 μg/mL primary antibody, and 1.3 μg/mL secondary antibody. Antibody incubations were performed for 60 minutes at room temperature. Proteins were detected with West Pico detection kits (Pierce, Rockford, IL) according to the manufacturer’s instructions.

ELISA

HCII-T ELISA kits were obtained from Affinity Biologicals (Hamilton, ON) and used according to the manufacturer’s instructions. This kit uses polyclonal sheep anti-human thrombin antibody for capture and peroxidase-conjugated polyclonal goat anti-human HCII antibody for detection. Standards were derived from purified human HCII and thrombin (Enzyme Research Laboratories, South Bend, IN) reacted in the presence of 0.05 U/mL heparin (Sigma, St. Louis, MO). MPS IH serum samples were diluted 500-fold and MPS IH/S samples were diluted 100-fold in factor II-depleted plasma (Affinity Biologicals) while control samples were undiluted. All standards and samples were tested in triplicate.

2. Examples
Example 1

iTRAQ serum proteomic studies

Serum samples from normal and \textit{idua} \textsuperscript{-/-} mice were depleted of albumin, immunoglobulin, and transferrin using a 100 mm Ms-3 Multiple Affinity Removal System (Agilent Technologies, Palo Alto, CA) according to the manufacturer's instructions. 4 age- and sex-matched pools (\textit{idua} \textsuperscript{+/+} mixed sex pool, \textit{idua} \textsuperscript{+/+} male pool, \textit{idua} \textsuperscript{-/-} mixed sex pool and \textit{idua} \textsuperscript{-/-} male pool) of mouse serum samples from 6 animals were normalized for total protein to 12 mg/mL by diluting them in running buffer (Agilent Technologies). 320 µL were subsequently used for high abundance protein depletion with identical pool collections obtained for all samples. Pools were labeled with different isotopic iTRAQ reagents as described \textit{supra}, and analyzed by LC-MS/MS analysis as described \textit{supra}.

Using a 94\% Protein Confidence score cut-off applied to data observed over two cumulative MS/MS acquisitions, iTRAQ analysis resulted in the identification of 1701 distinct peptides belonging to 198 unique proteins (Table 3). 181 proteins were identified on the strength of two or more peptides, the majority of which were represented by at least 5 peptides (Table 4). This weighting toward 5+ peptides per protein is likely due to the dynamic range of protein concentrations in serum, where more abundant proteins can be expected to be identified by several peptides. The second MS/MS data acquisition resulted in a 30\% increase in the total number of serum proteins identified.

<table>
<thead>
<tr>
<th>Protein Confidence Level, %</th>
<th>iTRAQ Acquisition A</th>
<th>iTRAQ Acquisition A+B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>94</td>
<td>95</td>
</tr>
<tr>
<td># of Peptides</td>
<td>1132</td>
<td>1131</td>
</tr>
<tr>
<td># of Proteins</td>
<td>151</td>
<td>150</td>
</tr>
</tbody>
</table>

Table A2.4: Number of unique peptides per protein in iTRAQ data collection, 94\% Protein Confidence.

<table>
<thead>
<tr>
<th>Number of Peptides/Protein</th>
<th>Total # of Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>iTRAQ Acquisition A</td>
<td>22</td>
</tr>
<tr>
<td>iTRAQ Acquisition A+B</td>
<td>17</td>
</tr>
</tbody>
</table>
To investigate intersample variation, we compared the relative abundances observed in the \( Idua^{+/+} \) male pool to the \( Idua^{-/-} \) mixed sex pool as well as the \( Idua^{-/-} \) male pool to the \( Idua^{-/-} \) mixed sex pool at a 99% Protein Confidence level (Fig. 1). The vast majority of proteins show no significant change in expression levels between like pools, with the exception of one outlier. The possibility of a sex-specific expression difference observed in the outlier (~16-fold decreased) in the \( Idua^{+/+} \) Male: \( Idua^{-/-} \) Mixed sex comparison is eliminated by the absence of such a difference in the \( Idua^{-/-} \) pools comparison and the decreased relative abundance of this protein (alpha-1-B glycoprotein) was assumed to be an artifactual in the \( Idua^{-/-} \) mixed sex control pool.

Comparison of the average protein relative abundances of the two \( Idua^{-/-} \) pools with the \( Idua^{+/+} \) mixed sex pool reveals the variation between mutant and control pools is similar to that observed between like pools; however, a few proteins showed more extreme abundance differences (Fig. 1). This suggests there is minimal variation in protein quantities between the MPS and normal serum proteomes, with no single protein present at a dramatically different level.

In order to select a panel of proteins to be investigated as candidate biomarkers with significantly altered relative abundances in the \( Idua^{-/-} \) serum proteome, the relative abundances of each protein in the two \( Idua^{-/-} \) pools were averaged then tested for two factors. First, we determined which proteins showed average relative abundances exceeding the 95% confidence intervals of the same proteins’ abundances in the \( Idua^{+/+} \) male pool. Second, proteins that also showed average relative abundances with confidence intervals not overlapping 0.00 (on a logarithmic scale) were considered strong candidates. The \( Idua^{-/-} \) mixed sex pool was used as the denominator pool for both analyses. Combining these criteria selected proteins exceeding the natural variability of proteins in both the MPS and WT animals, regardless of the absolute value of the change. Candidate biomarkers selected by these criteria are indicated as red spots in Fig. 2 and listed in Table 5. Candidate proteins with the most extreme deviation from the normal state were fibrinogen gamma (4.96-fold increased), fibrinogen alpha (2.20-fold increased), and heparin cofactor II (1.79-fold decreased). Elevated fibrinogen in mutant serum samples suggests impaired thrombin activity during clotting to form serum. 5 of the 17 proteins identified by iTRAQ with modified abundance in MPS I mice are serine protease inhibitors (Table 5).

**Table A2.5: Proteins with significantly altered relative abundance in the MPS mouse serum, \( Idua^{+/+} \) mixed sex denominator.**

<table>
<thead>
<tr>
<th>GenBank Accession Number</th>
<th>Protein % Confidence Score Rank</th>
<th>Protein Average Relative Increase/ Decrease</th>
<th>Log2 of Average Biologica Function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Abundance</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>AAH19506</td>
<td>95</td>
<td>Fibrinogen, gamma polypeptide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.310</td>
<td>+4.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blood clotting</td>
<td></td>
</tr>
<tr>
<td>AAH05467</td>
<td>71</td>
<td>Fibrinogen, alpha polypeptide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.137</td>
<td>+2.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blood clotting</td>
<td></td>
</tr>
<tr>
<td>AAC28866</td>
<td>87</td>
<td>Alpha-1-antitrypsin 1-5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.636</td>
<td>+1.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serine protease inhibitor</td>
<td></td>
</tr>
<tr>
<td>AAH13465</td>
<td>56</td>
<td>Inter-alpha trypsin inhibitor, heavy chain 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.480</td>
<td>+1.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serine protease inhibitor</td>
<td></td>
</tr>
<tr>
<td>AAA37246</td>
<td>5</td>
<td>Apolipoprotein B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.347</td>
<td>+1.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lipid and fatty acid transport</td>
<td></td>
</tr>
<tr>
<td>AAH57983</td>
<td>2</td>
<td>Pzp protein</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.154</td>
<td>+1.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serine protease inhibitor</td>
<td></td>
</tr>
<tr>
<td>AAH23143</td>
<td>29</td>
<td>Gelsolin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.144</td>
<td>-1.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cytoskeletal protein</td>
<td></td>
</tr>
<tr>
<td>BAA19743</td>
<td>26</td>
<td>Kininogen precursor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.281</td>
<td>-1.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Protein metabolism and modification</td>
<td></td>
</tr>
<tr>
<td>BAB33095</td>
<td>31</td>
<td>Histidine-rich glycoprotein</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.309</td>
<td>-1.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biological process</td>
<td></td>
</tr>
<tr>
<td>AAC28865</td>
<td>23</td>
<td>Alpha-1 proteininase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.377</td>
<td>-1.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serine protease</td>
<td></td>
</tr>
</tbody>
</table>
Example 2

**Heparin Cofactor II western blot analysis in MPS I mice and humans**

To investigate the HCII reduction in the serum of \textit{Idua} \textsuperscript{-/-} mice, western blot analysis was used utilizing goat anti-human HCII antibodies (Fig. 3). Surprisingly, the marked reduction in native HCII levels was associated with the presence of a higher molecular weight protein only in the mutant animals’ sera, consistent with the published size of the HCII-thrombin (HCII-T) complex (Tollefsen DM et al, 1982. Heparin cofactor II. Purification and properties of a heparin-dependent inhibitor of thrombin in human plasma, J. Biol. Chem. 257:2162-2169). Antibodies directed to thrombin confirmed that this was indeed HCII-thrombin complex. Fig. 4 shows the elevation of serum HCII-T complex in serum samples from both severe and attenuated human cases of MPS I compared to that of controls. Interestingly, the largest amounts of HCII-T complex are seen in the MPS IH patients in comparison to the MPS IH/S patient. Western analysis revealed no detectable HCII-T in plasma samples of \textit{Idua} \textsuperscript{-/-} or WT mice, nor in plasma samples from humans.

Example 3

**Heparin Cofactor II western blot analysis in treated MPS I patients**

Figure 5a-b shows the level of HCII-T in a Hurler patient receiving enzyme replacement therapy (ERT) preceding and following bone marrow transplantation (BMT), and one Hurler-Scheie patient undergoing ERT only. Enzyme treatment in the Hurler patient (Fig. 5a) did not normalize HCII-T levels, but significantly reduced the amount of HCII-T levels to that seen in the attenuated patient studied. Further reduction in HCII-T occurred following bone marrow transplantation. Although well engrafted by week 52, this patient subsequently died of pulmonary hemorrhage. The Hurler-Scheie patient, Fig. 5b, showed marked reduction of HCII-T early during ERT exposure but then subsequently developed detectable HCII-T complex later during treatment.
Example 4

HCII-T ELISA analysis

Table 6 illustrates the dramatic elevation of HCII-T in the serum of MPS IH and MPS IH/S patients as well as murine MPS I samples in comparison to controls. MPS IH patients' serum HCII-T complex levels ranged from 174,700 – 208,600 pM, with an average value of 188,600 pM, representing a 630-fold increase relative to controls. The serum sample from a MPS IH/S patient had a HCII-T concentration of 46,000 pM (154-fold increase), reflective of the patient's attenuated phenotype.

In contrast to the lack of detectable complex by western blot, ELISA revealed that plasma HCII-T levels were increased in MPS patients and the MPS I mouse, with the minimum concentration of HCII-T complex in MPS IH patients exceeding the maximum control value by 68%.

<table>
<thead>
<tr>
<th>Sample (age in brackets)</th>
<th>Serum [HCII-T] (pM ± SD)</th>
<th>Plasma [HCII-T] (pM ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (10 yr F)</td>
<td>115.1</td>
<td>17.92</td>
</tr>
<tr>
<td>Control (10 yr M)</td>
<td>398.0</td>
<td>9.91</td>
</tr>
<tr>
<td>Control (30 yr M)</td>
<td>384.7</td>
<td>6.27</td>
</tr>
<tr>
<td>MPS 1H (10 mo, Patient A)</td>
<td>174 700</td>
<td>30.15</td>
</tr>
<tr>
<td>MPS 1H (12 mo, Patient B)</td>
<td>182 400</td>
<td>Not tested</td>
</tr>
<tr>
<td>MPS 1H (14 mo, Patient C)</td>
<td>208 600</td>
<td>98.37</td>
</tr>
<tr>
<td>MPS 1H/S (8 yr, Patient D)</td>
<td>46 000</td>
<td>Not tested</td>
</tr>
<tr>
<td>Idua^{+/+} (n = 5)</td>
<td>75.46 ± 4.99</td>
<td>3.77 ± 1.20</td>
</tr>
<tr>
<td>Idua^{-/-} (n = 3)</td>
<td>628.1 ± 163.2</td>
<td>79.50 ± 38.9</td>
</tr>
</tbody>
</table>

Significantly decreased levels of the serine protease inhibitor (serpin), heparin cofactor II (HCII) is present in MPS-affected animals. Although native HCII levels are reduced, there is marked HCII-thrombin (HCII-T) complex elevation in MPS-affected animals. Translation of these findings to humans with MPS I show equivalent findings. Importantly, in humans the elevation of HCII-T complex appears to be correlated to disease severity and is responsive to treatment. GAGs play a major role in MPS pathophysiology, and a large body of literature exists for GAG-modulated serpin activity. These results indicate that HCII-T is an excellent biomarker for MPS I and represents a novel finding that may implicate other GAG modulated serpins in the pathophysiology of MPS diseases.
In addition, the level of HCII-T complex correlates with the clinical measures of disease severity as well as responsiveness to therapy. Further, patients undergoing ERT and/or BMT maintain residual amounts of serum HCII-T.

**Example 5**

**Antithrombin III western blot analysis.**

Antithrombin III (ATIII) is the principle circulating serpin, present at approximately twice the plasma concentration of HCII, and is known to be activated exclusively by heparan sulphate. Elevated serum levels of antithrombin III-thrombin complex (ATIII-T) is also found in serum samples from MPS I patients. (Fig. 6). The reduced dynamic range of the ATIII-T complex in comparison to the HCII-T complex suggests its use as a biomarker is not as reliable, because the clear presence of ATIII-T complex in the control serum indicates the specificity of this biomarker may not be as accurate as HCII-T for distinguishing the attenuated phenotypes and for measuring treatment efficacy.

**CLAIMS:**

We claim:

1. A method of diagnosing an MPS, the method comprising: obtaining a tissue sample from a subject and determining the quantity of a biomarker of an MPS present in a tissue sample.

**ABSTRACT**

A method of diagnosing, monitoring or screening for an MPS disorder in an individual affected, or suspected of being affected by an MPS. Biomarkers correlating with severity of MPS, and methods of detecting these biomarkers are disclosed. Examples of biomarkers include heparan cofactor II and antithrombin III.
Fig. A2.1: Box plots of relative abundances for proteins identified with 99% confidence.
Fig. A2.2: Logarithmic plot of average relative abundance of proteins in $Idua^{-/-}$ pools vs. $Idua^{+/+}$ mixed sex, ≥99% confidence, with proteins rank ordered from left to right based on descending Protein % Confidence score.
Fig. A2.3: Western blot of serum samples from WT ($Idu^+/-$) and MPS ($Idu^{-/-}$) mice for HCII.
Fig. A2.4: Western blot of MPS patients for HCII.
Fig. A2.5: HCII-T levels in MPS 1H serum samples following enzyme replacement therapy.
Fig. A2.6: Western blot of MPS patient samples for ATIII.