THE DEVELOPMENT OF MASS SPECTROMETRY-BASED APPROACHES FOR THE DIAGNOSIS OF HEMOGLOBINOPATHIES

Krisztina Radi M.Sc.

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Warwick
School of Life Sciences

May 2013

"Under normal conditions the research scientist is not an innovator but a solver of puzzles, and the puzzles upon which he concentrates are just those which he believes can be both stated and solved within the existing scientific tradition."

Thomas Kuhn

"*Vigyázz arra, hogy ne mulaszd el a pillanatot, amely csak a te pillanatod, műved kivitelének végzetszerűen kijelölt időszaka. Kényelem, szöszmötölés, gyávaság, lustaság néha késleltetik feladatod kivitelét, noha szíved mélyén tudod jól, hogy az idő telítve van azzal, amit rajtad keresztül el akar mondani, s egy pillanatot sem mulaszthatsz, mert elmondja más helyetted, s nem úgy mondja el, ahogyan te jónak és igaznak hiszed. Tudományban, művészetben, irodalomban, közéletben vannak ilyen sürgető pillanatok, mikor egy igazság megérett és ki kell mondani. S ha úgy érzed, végzetesen éppen téged jelölt ki e feladatra a sors, ne késlekedj, mint a rossz színész, aki elmulasztja a jelenet végszavát.*
Nemcsak műved van, az idő is van. S az időn belül meg van a te pillanatod, melyet nem szabad elmulasztani."

Márai Sándor

Dedicated to my mother

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

At the end of this long journey, I realise that it's not the length of the road, the difficulties alongside it and the never-ending challenges that were important, but the people who shared all this with me, who supported me, who listened to me, gave advice, or an occasional hug when it was most needed. Their belief in me and my ability to walk, or intermittently run, to the very end of this road have given me the strength to keep going and finally arrive at my destination.

I would like to begin by thanking my acedemic supervisor Prof. Jim Scrivens for giving me the opportunity to undertake this PhD. His support, advice and continoual encouragement have been invaluable over the last three years. I would like to thank Waters Corporation for all the financial support.

I would like to thank Charlie Scarff for guiding me on this new field in my first years, her advice and help was invaluable. Special thanks to Baharak who was the best post-doc I could ask for; creative in problem solving and fun to work with. Without her all this could not have happened.

I am grateful for Dr Nicholas Jackson and Jane Newbold our collaborators at UHCW for their contributions and help.

I would like to say thank you for all the help and support from Bruker; to Ian Saunders for the opportunity to be part of an exciting project, to Julia for her constant help and always being there when I needed to discuss the work and to Martyn for his constant efforts in fixing the LC whilst answering my never-ending questions. I would also like to thank Ralf Hartmer for method development, and Zoltan Czentar for his data processing script which made life easier.

I am very grateful to Nisha and I consider myself a lucky person to have her friendship. Thanks to all past and current members of the Scrivens' group: Fran, Elle, George, Kostas, Gill and special thanks to Matt and Sue, who supported me and kept my blood pressure fairly low in the last three months.

Thanks to all my friends Bori, Agika, Kati, Balazs. Csaba for believing me while supervising me for my MSc thesis. Thanks to all my Warwick friends: Dani, Mike, Matei, Adrienne, Pete and Nick.

I am always going to be grateful for Tamas for his support, love and faith in me, even being thousands of kilometres away being the person closest to me.

Special thanks to my two fellow Hungarians Mate and Viki who handled my stressful periods with huge patience.

Végül, de nem utolsó sorban szeretnék köszönetet mondani a családomnak: az édesanyámnak, aki akkor is hitt bennem, amikor én magam nem, minden amit elértem neki köszönhető és a nővéremnek. Remélem mindketten büszkék rám, és tudják én felnézek rájuk, mégha mindketten alacsonyabbak is egy ici-picit.

DECLARATION

I hereby declare that this thesis, submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy and entitled "*The Development of Mass Spectrometry-Based Approaches for the Diagnosis of Hemoglobinopathies*", represents my own work and has not been previously submitted to this or any other institution for any degree, diploma or other qualification. Work undertaken by my colleagues or collaborators is explicitly stated where appropriate.

Krisztina Radi
May 2013

SUMMARY

Hemoglobinopathies are disorders of the protein hemoglobin (Hb), one of the most common inherited disorders, and pose an increasing healthcare problem. Worldwide, approximately 200 million people have hemoglobinopathies (Hartwell *et al.* 2005), and whilst the majority of them are clinically silent, some are life-threatening. Screening programs for hemoglobinopathies traditionally employ chromatography and electrophoretic techniques, which only provide presumptive identification of clinically important hemoglobin disorders. The definitive characterisation of these disorders requires protein sequence elucidation or DNA analysis. The development of a rapid population screening method for hemoglobinopathies, which could also provide definitive diagnosis of a disorder, is of significant interest to the healthcare profession. Mass spectrometry (MS) is currently used to routinely characterise variants but not as a validated clinical diagnostic tool.

The work undertaken addresses the increasing demand for population screening. Different mass spectrometry-based approaches to hemoglobinopathy detection have been developed based on previously described procedures found in literature. The methods, for intact globin chain analysis and targeted variant specific peptide analysis, have been optimised and evaluated using clinical blood samples on a triple quadrupole system.

Intact globin chain analysis and targeted tryptic peptide screening methods have been evaluated in a clinical trial analysing 2017 blood samples to identify clinically significant variants routinely screened as a requirement by the NHS Screening programme and to quantify minor globin chains (HbA_2 and HbF) suitable for the diagnosis of thalassemic and other disorders. This work was carried out in a three month mirrored testing experiment against existing clinical laboratory methods. The results obtained showed an excellent correlation to values from current clinical methods, as all clinically significant variants were detectable by MS methods. In addition, MS was able to detect Hb variants undetected by electrophoretic or chromatographic methods and highlighted some issues with the current phenotypic methods which can lead to misdiagnosis.

A top-down ETD method utilising an iontrap mass spectrometer has been developed and applied to provide fragmentation for positive identification of known hemoglobin variants. This method was evaluated by analysing 1000 samples in a pilot study as a sideproject to the clinical trial. The generation of an automated report with the detection of clinically significant variants and minimal sample preparation are some of the benefits associated with this approach. This method has been shown to meet the majority of requirements for a rapid diagnostic screening method and has the potential to become a validated tool in the clinical laboratory screening in the near future.

ABBREVIATIONS

A

ACN	Acetonitrile
ADC	Analog-to-digital converter
AUC	Area under the curve

B

BEH	Bridged ethyl hybrid

C

ce-HPLC	Cation exchange - high pressure liquid chromatography
CID	Collisionally-induced dissociation
CsI	Caesium iodide

D

Da	Dalton
DC	Direct current

E

ECD	Electron capture dissociation
ESI	Electrospray ionisation
EM	Electron multiplier
ETD	Electron transfer dissociation
eV	Electron volt

F

FAB	Fast atom bombardment
fmol	Femtomole
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width at half maximum

G

H

h	Hour
HPLC	High performance liquid chromatography

I

IEF	Isoelectric focussing
IMMS	Ion mobility mass spectrometry
IT	Ion trap

K

kDa	Kilodalton
kV	Kilovolt

L

LC	Liquid chromatography
LIT	Linear ion trap
LOD	Limit of detection
LOQ	Limit of quantification

M

M	Molar
m/*z*	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionisation
MCP	Microchannel plate detector
MCV	Mean cell volume
MCH	Mean corpuscular hemoglobin
min	Minute
mg	Milligram
mL	Millilitre
mM	Millimolar
MS	Mass spectrometry
ms	Millisecond
MS/MS or MS^n	Tandem mass spectrometry
μg	Microgram
μL	Microlitre
μm	Micrometre
μM	Micromolar
μs	Microsecond

N

NaI	Sodium iodide
ng	Nanogram
nm	Nanometre
ns	Nanosecond

O

oa	Orthogonal acceleration

P

ppm	Parts per million
PMF	Peptide mass fingerprinting
PTM	Post-translational modification
PTR	Proton transfer reaction

Q

QIT	Quadrupole ion trap
Q-TOF	Quadrupole time-of-flight

R

RF	Radio frequency
RPLC	Reversed phase liquid chromatography
RT	Retention time

S

s	Second
SCD	Sickle cell disease

T

TDC	Time-to-digital converter
TOF	Time-of-flight
TQ	Triple quadrupole
TWIMS	Travelling-wave ion mobility spectrometry

U

UPLC	Ultra performance liquid chromatography

V

V	Volt

X

Standard 3 and 1 letter amino acid abbreviations are used throughout this thesis.

Amino acid	**3 Letter Code**	**1 Letter Code**
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

CHAPTER 1

INTRODUCTION

1.1 Instrumentation

Mass spectrometry is a versatile analytical technique originally used in physical and chemical analysis. It has become a method of choice in characterising biological systems. Complex biological problems increasingly rely on mass spectrometry data. The underlying fundamentals and technical details of the mass spectrometry-based analysis of biological samples are presented in this section.

1.1.1 What is a mass spectrometer?

The mass of a molecule is a fundamental characteristic, and the possibility of determining the molecular mass provides a powerful tool for identifying compounds in complex mixtures. Mass spectrometry is an analytical technique which is capable of measuring mass-to-charge ratios (*m*/*z*) of ionised molecules in the gas phase.

The origin of mass spectrometry may be placed in the early 20th century when J.J. Thomson was working on the transmission of electricity through gases and cathode rays. This led him to find an extremely useful application of his research in chemical analysis (Thomson 1913). After this, mass spectrometry became a physicist's tool with increasing usage in chemical analyses after the 1940s. It remained that for four more decades, until the invention of soft ionisation techniques in the late 1980s as reviewed by (Cristoni and Bernardi 2003) (Karas *et al.* 1987, Karas and Hillenkamp 1988, Fenn *et al.* 1989).

The main components of a mass spectrometer are the ion source, the analyser(s) and the detector. A simplified block diagram can be seen in Figure 1.1. In the ionisation source molecules become charged, the ions are then resolved based on their mass-to-charge ratios in the mass analyser(s). Following this separation, ions are detected and with the help of a computer system the acquired data is presented in an accessible format.

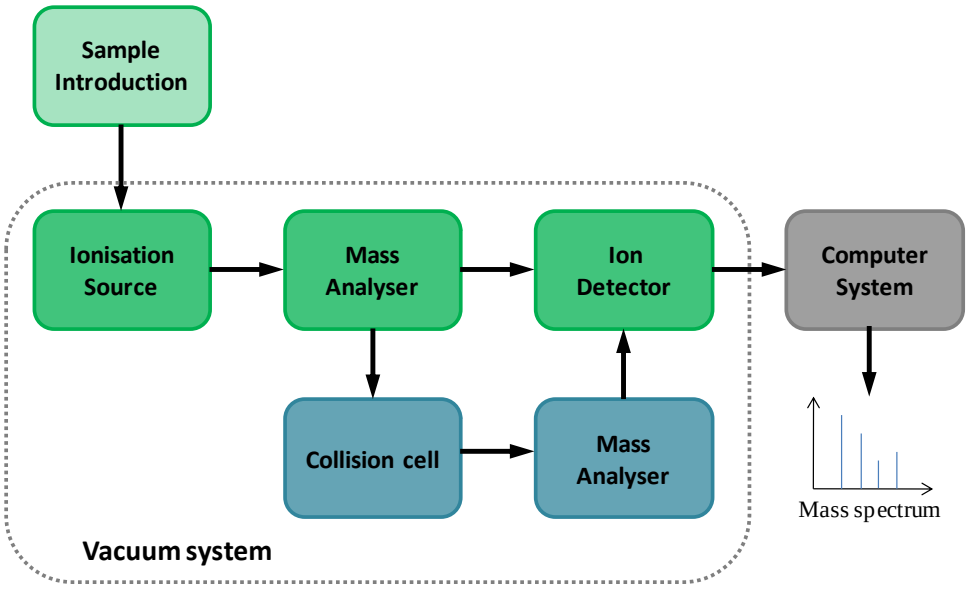


Figure 1.1. Schematic representation of a mass spectrometer
Elements in blue are used by tandem in space instruments.

1.1.2 Ionisation methods

In the ion source of a mass spectrometer the sample of interest is ionised prior to mass analysis. During this process neutral molecules become charged, and they are transferred into the gas phase if they have not been evaporated previously. Early ionisation techniques required the sample to be vaporised prior to ionisation. These methods included electron impact ionisation, chemical ionisation and field ionisation, and their use was limited to analytes which were volatile and thermally stable. Electron impact ionisation is still widely used for the analysis of small organic molecules. This requires the sample to be in the gas phase and so is unsuitable for the study of large biomolecules.

Biological samples such as proteins are usually large, polar and thermally labile and cannot be easily transferred into the gas phase without extensive decomposition. Field Desorption ionisation was introduced for analysis of non-volatile compounds in the late 1970s (Beckey 1969b) , and with the invention of fast atom bombardment (FAB) it became possible to produce ions from liquid or solid surfaces overcoming the need for volatilization prior to ionisation (Barber *et al.* 1981, Morris *et al.* 1981). This allowed the routine analysis of thermally liable and non-volatile molecules, proteins and peptides up to a few thousand Daltons in mass and the mass

spectrometric analysis of biomolecules started a new era in the history of the technique.

The development of soft ionisation techniques such as matrix-assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI), in the 1980s significantly increased the role of MS in biomolecule analysis, dominating the field with application areas are still growing (Karas *et al.* 1987, Karas and Hillenkamp 1988, Fenn *et al.* 1989). ESI can be coupled with separation techniques making it a predominant ionisation technique for research purposes in the life sciences. A potential advantage of MALDI over ESI is the predominant production of singly charged ions of peptides and proteins, which can be helpful for instances where decreasing spectral complexity is required (El-Aneed *et al.* 2009). The theories underlying these two ionisation methods are presented below.

1.1.2.1 Electrospray ionisation (ESI)

The first electrospray ionisation experiments were conducted in Malcolm Dole's laboratory in the late 1960s. This involved the successful introduction of a polymer into the gas phase (Dole *et al.* 1968). This technique was then optimised and improved by Fenn and co-workers (Yamashita and Fenn 1984) and emerged at the end of the 1980s as a powerful technique to produce intact ions from large and complex molecules in solution (Fenn *et al.* 1989).

In ESI the analyte molecules present in solution are introduced into a strong electric field through a capillary under atmospheric pressure. The electrospray itself is formed as a result of the electrostatic potential difference between the capillary and the counter electrode (cone). A fine spray is ejected from the capillary in a form commonly known as the Taylor cone. The solution at the end of the needle is polarised and torn away from the needle (Wilm and Mann 1994). A spray is comprised of droplets containing charged analyte molecules and depending on the applied potential on the tip of capillary, these droplets are positively or negatively formed. Due to solvent evaporation the size of the droplets decreases while the surface tension increases up to a point when the columbic repulsion overcomes this tension and reaches the Rayleigh limit, whereby the droplet undergoes fission and new smaller droplets are formed. Further evaporation of the parent droplet leads to repeated coulombic fissions. The progeny droplets also evaporate and undergo the

same fission, resulting in small charged droplets which ultimately leads to gas-phase ions (Kebarle and Verkerk 2009). How the actual ions are formed from the small droplets is still under debate. A schematic of the two major models proposed is shown in Figure 1.2.

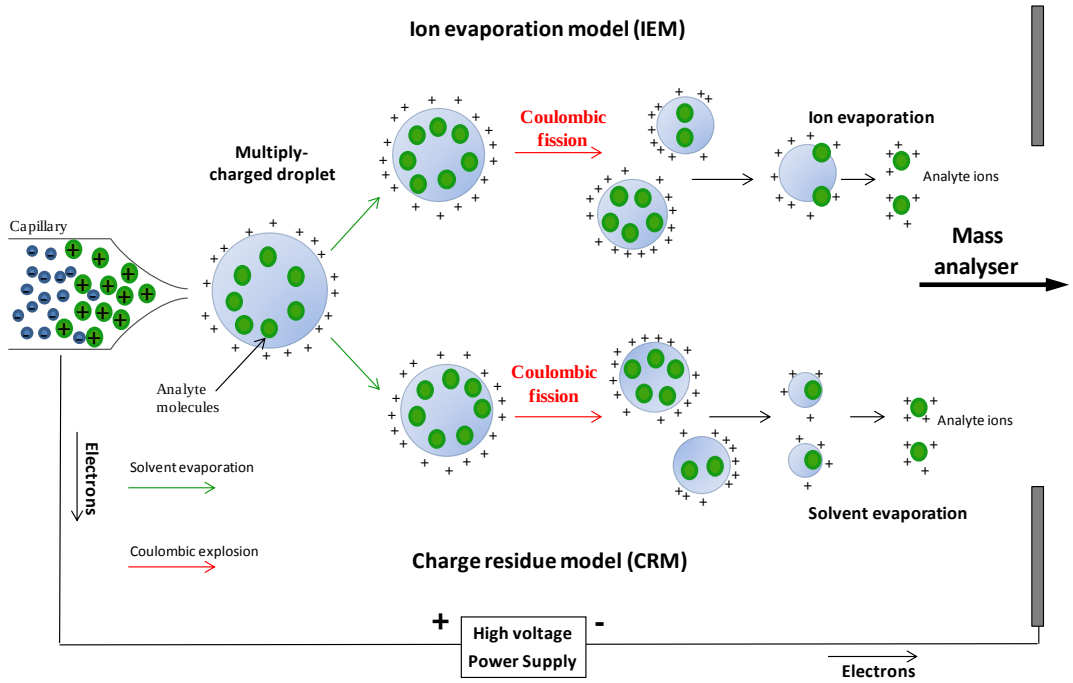


Figure 1.2. Electrospray ionisation mechanism presented by both theories
Adapted from (Nguyen and Fenn 2007)

The ion desorption mechanism (IEM) was first proposed by Iribarne and Thomson (Iribarne and Thomson 1976), while the charge residue model (CRM) was previously presented by Dole and co-workers (Dole *et al.* 1968).

The IEM proposes that after a number of evaporation and fission events, and before a charged droplet becomes small enough to further reduce some of the analyte ions from the surface of the droplet are pushed into the gas phase since the electric field present is sufficiently intense to make that happen. According to CRM the solvent evaporation and subdivision proceeds until small droplets are formed that contain one analyte molecule. The final evaporation of the remaining solvent from this ultimate droplet leaves the residual analyte molecule with the charge previously carried by the droplet.

Despite work on the mechanisms by a number of research groups in the last four decades it has not been fully elucidated and remains a subject of investigation

(Abonnenc *et al.* 2010, Hogan *et al.* 2009). There is some agreement in that low molecular weight analytes are proposed to follow the ion evaporation model (IEM), while the charged residue model (CRM) applies to large globular species. A chain ejection model (CEM) has also been proposed recently for disordered polymers (Konermann *et al.* 2012).

One of the key characteristics of ESI is the formation of multiply charged ions. Analytes need to be solubilised in a preferably polar/aqueous solvent, often with the presence of organic solvents (acetonitrile and methanol) which is supportive of the desolvation process, but in case of proteins also makes the analyte more amenable to protonation by facilitating denaturation. One of the disadvantages of ESI is that the technique is very sensitive to small amounts of salts present in the solution, which are likely to compete successfully in the ionisation process at the expense of the analyte (Lubec and Afjehi-Sadat 2007).

For smaller scale and nanoflow applications nano-electrospray was developed, which generally follows the same principles as standard electrospray. A chip-based nano-electrospray technique has been developed for direct sample injection (Covey, Thomas *et al.* 2009, Wilm and Mann 1996, Van Berkel 2003).

1.1.2.2 Matrix assisted-laser desorption ionisation (MALDI)

The other leading ionisation technique for biological molecules is matrix assisted-laser desorption ionisation (MALDI). The technique was introduced and made applicable for mass spectrometric analysis by Karas and Hillenkamp in 1987 (Karas *et al.* 1987, Karas and Hillenkamp 1988). They realised that applying a matrix significantly improves the ionisation efficiency. The Nobel prize was however given to Tanaka based on his earlier presented less practically applicable work on laser desorption (Tanaka *et al.* 1988).

MALDI analysis requires that the analyte is co-crystallised with a matrix, the most common of which are nicotinic acid, benzoic acid and cinnamic acid derivatives. The matrix molecules need to be in excess to the analyte molecules, and they are excited by the application of a pulsed laser beam, most commonly a N_2 laser (frequency in the UV region with 337nm) but IR lasers have been also used. The energy provided by the laser is absorbed by the matrix and transferred into vibrational energy. Choosing the right matrix for mass spectroscopic analysis is crucial, as it is

important that it has absorption at the wavelength of the laser (Batoy *et al.* 2008). The absorbed energy from the laser then causes the matrix to vaporise and the analyte to be transferred into the gas phase and ionised.

The ion formation process is currently thought to follow a two-step mechanism. In the first step the formation of primary ions takes place, which then undergo reactions forming secondary ions. The secondary ions are produced in proton transfer reactions between the matrix and analyte molecules (Knochenmuss 2006).

Similarly to ESI the processes of ion formation are still under investigation. It has been debated whether ionisation takes place in the solid phase prior to desorption or in the gas phase. Conventional models assume chemical equilibrium in the gas phase, while Lai and co-workers has showed a quantitative interpretation assuming chemical and thermal equilibrium in the solid state prior to desorption (Lai *et al.* 2010).

In most cases ions produced by MALDI are singly charged resulting from the attachment of a single proton. The result of this is less complex mass spectra and the requirement of a mass analyser capable of measuring over a broad *m/z* range, such as a time of flight (TOF) analyser.

Both ESI and MALDI are very sensitive analytical techniques utilizing picomolar analyte concentrations. The fact that ESI can be easily coupled to LC makes it suitable for quantitative measurements, and ESI has been shown to be more reproducible than MALDI. There were some attempts to interface MALDI with LC, mainly because of the robustness of the technique to the presence of salts, but due to the heterogeneity of a MALDI crystal sample the quality of MALDI spectra can be significantly influenced by the position of the laser beam, which is not ideal for quantitative applications (El-Aneed *et al.* 2009).

1.1.3 Mass analysers

The mass analyser is an important part of the mass spectrometer, where the produced ions are separated according to their *m*/*z* ratios. Instruments are often named after the analyser. The most appropriate analyser to achieve the best results is usually dependent on the hypothesis or application. Analysers can be classified based on different properties but most commonly they are grouped into analysers which scan

and can successively detect ions of different *m/z* over a period of time or those which allow simultaneous transmission of ions of all *m*/*z* at the same time (de Hoffmann and Stroobant 2009). Scanning analysers are quadrupole (Q) and magnetic sector , while the second group includes the time of flight analyser (TOF) and trapping analysers such as ion trap (quadruploe ion trap, QIT, linear iontrap LIT) Fourier transform ion cyclotron resonance (FTICR) and the Orbitrap. For biochemical research a broad variety of analysers have been used to date and the five dominant techniques which have been developed are presented below.

All mass analysers can be described using some key characteristics: upper mass limit, transmission efficiency and resolving power. The most important one is the resolving power, which can have a significant influence in the analysis of complex mixtures. The resolution is related to the ability of the analyser to distinguish between two ions that are of similar *m*/*z*. It is a ratio calculated between the measured mass and the smallest mass difference (m/Δm) to the neighbouring mass peak at a certain *m*/*z* ratio with the difference calculated at full-width half-height maximum (FWHM) of the peak.

The highest *m/z* value that can be measured with the application of a certain analyser is the upper mass limit. The transmission efficiency is defined by the ratio of the number of ions detected to the number of ions that are produced in the ion source.

Hybrid instruments have been developed to support tandem mass spectrometric experiments. These are further discussed in section 1.1.5.

1.1.3.1 Quadrupole mass analyser (Q)

The quadrupole analyser introduced by Paul in 1953 was a breakthrough which provided an alternative to large magnetic field analysers (Paul and Steinwedel 1953). The quadrupole is a scanning mass analyser consisting of four parallel conductive circular or hyperbolic rods arranged in pairs which are electrically connected. Direct current (DC) and alternating radio frequency (RF) potentials are applied to the rods in a way that the two opposite rods have the same voltages. When an ion with a certain *m*/*z* enters the quadrupole, depending on the produced electromagnetic field between the rods the ion can either have a stable trajectory and pass through the quadrupole or have an unstable trajectory and collide into the rods or be lost at the walls of the analyser, and not reach the detector. This is illustrated in Figure 1.3.

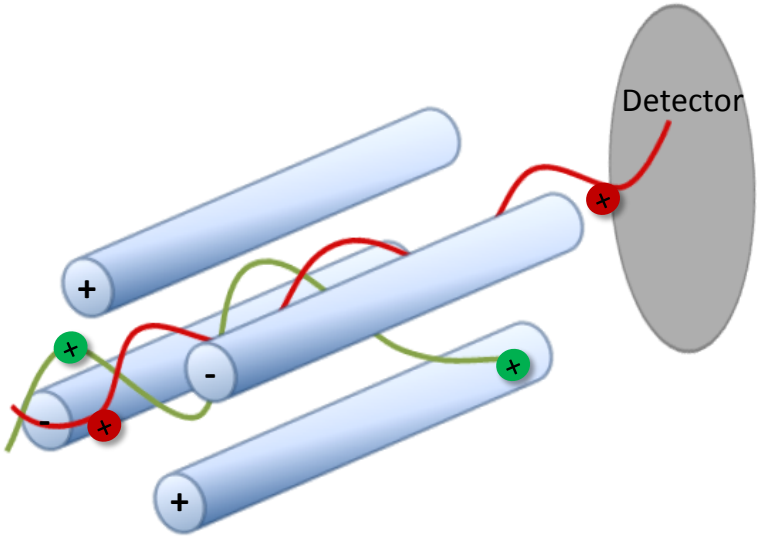


Figure 1.3. Schematic representation of a quadrupole mass analyser.
A stable ion trajectory (shown in red) gets through the quadrupole to the detector whilst the unstable ion trajectory (shown in green) is repelled

By changing the electric field that the ions are exposed to different ranges of *m*/*z* ratios may be scanned and ions allowed to pass through the quadrupole. When the quadrupole is operating in scanning mode, a constant ratio of the voltages (particular value of DC and RF voltages) is applied to the pair of rods, and only ions of a specific *m*/*z* will have a stable trajectory.

In the RF-only operation mode the quadrupole can also act as an ion guide. In this mode RF voltages are applied to the rods making it possible for all ions to pass through the analyser with a stable trajectory.

Quadrupole analysers do not require very high vacuum and they are relatively fast. This supports high throughput analyses, but their upper mass limit (around 2000 *m/z*) and resolution (unit resolution) is relatively low compared to other analysers. There have been some new developments to provide quadrupole mass filters with higher mass ranges with the quadrupole operating at different frequencies (Douglas 2009).

1.1.3.2 Time-of-flight mass analyser (TOF)

The linear time-of-flight analyser was first described by Wiley and McLaren (Wiley and McLaren 1955). The time-of-flight (TOF) mass analyser separates ions according to their velocities. Mass-to-charge ratios can be related to the time it takes an ion to pass through a field-free region called a flight tube. The ions enter the analyser

and are accelerated with the use of a kinetic energy pulse in the electrostatic field. As all ions acquire the same kinetic energy, their velocities and therefore their flight times depend on their masses and charges. Ions travel at a velocity inversely proportional to the square root of their mass. In general, the larger the ion, the longer it takes to travel through the tube.

In principle the upper mass range of this type of analyser has no limits, but there are limiting factors in resolution. The length of the ion formation pulse, the volume of the space where the ions are formed and the differences in initial kinetic energy between ions all affect the resolution. Ions of the same *m*/*z* but with slightly different kinetic energy will reach the detector at slightly different times, resulting in peak broadening and reduced resolution. This variation in kinetic energy can be compensated by the application of delayed pulse extraction (Vestal *et al.* 1995). The time delay between ion formation and extraction allows ions to separate according to their kinetic energy in the field free region. The ions with more energy move further away from the source towards the detector, and when the pulse applied more energy is transmitted to ions which remained for a longer time in the source (de Hoffmann and Stroobant 2009).

As mass resolution is proportional to flight time, higher resolution can be achieved by increasing the length of the flight tube. Reflectron instruments achieve this by reflecting ions back down the flight tube using a series of electrodes which act as an ion mirror, as shown in Figure 1.4. Ions with higher initial kinetic energies travel further into the reflectron compared to those with less initial kinetic energy. The addition of an extra mirror (also known as W optics) can further improve resolution, the flight distance is increased by a factor of four compared to the original length of the flight tube.

Delayed extraction compensates for the spatial distribution, while the reflectron reduces the kinetic energy variation. By the time they have reached the detector the faster moving ions catch up with the slower ions (Somogyi 2008).

TOF analysers were originally directly compatible with pulsed laser ion sources like MALDI, but orthogonal acceleration TOF (oa-TOF) provides fast, sensitive mass analysis for continuous ion sources, where ions are sampled orthogonally from the ion beam to transform a continuous ion flow into a pulsed one (Mirgorodskaya *et al.* 1994, Coles and Guilhaus 1993, Guilhaus *et al.* 2000).

The TOF analyser benefits from highly efficient ion transmission, good sensitivity, the ability to measure over a wide mass range, and the capability of achieving a relatively high mass resolution. To date high-end TOF instruments can have a resolution up to ca 40,000–50,000 and mass accuracy in the parts per million (ppm) range. Routine instruments have a resolution up to 10,000–20,000 and mass accuracy of below 5 ppm (Jiwan *et al.* 2011).

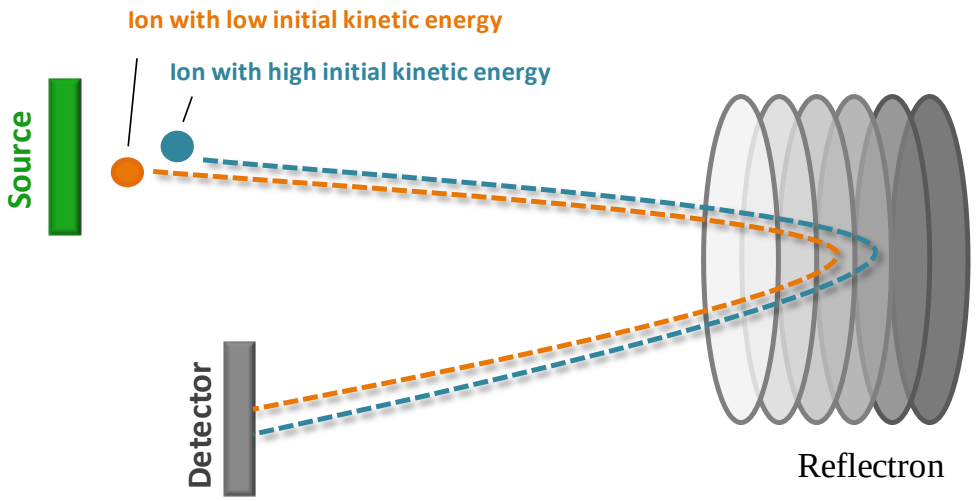


Figure 1.4. Schematic of time-of-flight (TOF)

1.1.3.3 Ion traps (IT)

Ion trap mass analysers operate on the principle of trapping by using an electric field to store ions. The ions are trapped in two or three dimensions by using an RF quadrupolar field (de Hoffmann and Stroobant 2009).

The quadrupole or 3D ion trap (QIT)

The first ion traps were three dimensional traps. The principle of the ion trap originated from Paul and Steinwedel, but it was modified to provide a useful mass spectrometer by Stafford (Stafford *et al.* 1984).

It consists of a circular electrode and two ellipsoid caps at the ends of the 3D quadrupolar field (de Hoffmann and Stroobant 2009). One endcap is the entrance through which ions enter the trap and the other is the exit electrode through which

ions are transferred to the detector. The ring electrode has an internal hyperboloidal shape (March 2009), schematic is shown in Figure 1.5.

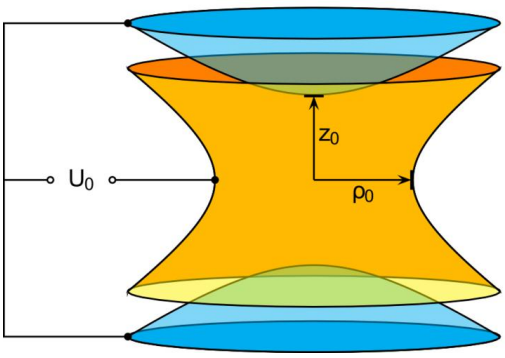


Figure 1.5. Schematic of quadrupole ion trap
End cap electrodes in blue, and ring electrode in the middle in orange with the characteristic details describing the distance from the center to the endcaps and the ring, U0 is the applied voltage.

Similarly to the quadrupole analyser direct and alternating potentials are applied. The overlapping potentials give a kind of 3D quadrupole, in which all the ions ar trapped on a 3D trajectory. Commercial ion traps are operated in a way that a fundamental RF voltage is applied to the ring electrode, that voltage has a constant frequency, but its amplitude can be varied, while additional voltages of selected frequencies can be applied on the end caps. As all the ions are trapped simultaneously, they repel each other and their trajectories expand as a function of the time. To reduce the expansion a certain level of gas pressure is maintained withing the trap with Helium, which collides with the ions and removes their excess energy.

Early ion traps employed mass-selective detection based on recording the oscillating ion motions to analyse ions. This was followed by mass-selective storage but the mass-selective axial ejection was the real breakthrough which made it the mass analyser we use today. Ions with different *m*/*z* are present in the trap simultaneously, and they are sequentially expelled out of the trap. The mass spectrum is acquired while ions are systematically scanned out with the application of a linear gradient of RF voltages.

In the 3D ion trap tandem mass spectrometry can be performed in a time-dependent rather than a space-dependent manner. The precursor ion is selected by expelling all other ions from the ion trap, fragmentation energy is provided by collisions with the helium gas and with resonant excitation. The produced fragment ions are analysed

with the described scanning method, which also allows another selection and further fragmentation of the selected product ion.

The Linear or 2D ion trap (LIT)

The linear ion trap or 2D ion trap (LIT) has a similar build to the quadrupole. Ions are restrained radially in a two-dimensional RF field and axially by applying additional stopping DC potentials to the end electrodes. Ions can be ejected either radially or axially. Hager revolutionised the technique by describing how ions can be ejected axially from a LIT in a mass selective manner (Hager 2002). Compared to 3D traps, linear traps have higher injection efficiencies and greater ion trapping volumes. Based on the structure of LIT use is not limited to storage of the ions, it can also act as an ion guide, an *m/z* selector or a collision cell. As a consequence LIT is often used with other mass analyzers in hybrid instruments for the isolation of ions with selected *m*/*z* to perform tandem mass spectrometry experiments (Douglas *et al.* 2005).

1.1.3.4 Fourier transform mass analysers

Fourier transform ion cyclotron resonance (FT-ICR)

Ion cyclotron resonance was first described by J.A. Hipple in 1949 (Hipple *et al.* 1949), but Marshall and Comisarow realized that Fourier transform could be applied to ICR to reduce the data acquisition time (Comisarow and Marshall 1974). FTICR has become the analyser offering the highest mass accuracy (< 1 ppm) and resolution (~ 500,000) to date.

In FTMS ions are forced to undergo a cyclotron motion by a spatially uniform static magnetic induction. The frequency of their oscillation is characteristic of their mass-to-charge ratio (*m*/*z*) (Zhang *et al.* 2005). Ions with different *m*/*z* are present in the ICR cell, and they all have different circular motion with different frequencies. The detected signal is an extremely complex signal, which needs to be deconvoluted by Fourier transformation. FTICR requires ultra-high vacuum and a stabile homogenous magnetic field, which is maintained by superconductive magnets. These requirements make this technique the most expensive analyser.

Orbitrap

A recently developed Fourier transform analyser was invented by Makarov in 2000 (Makarov 2000). The Orbitrap mass analyser is built up of a spindle-like central electrode and a barrel-like outer electrode. An electrostatic potential is produced by applying a DC voltage between the two axially symmetric electrodes, which makes the ions oscillate around the inner electrode in a stable trajectory with a defined frequency. This is responsible for orbital trapping of the ions in the radial direction, and a quadrupolar potential traps ions axially while they undergo harmonic oscillation in the z-direction (Jiwan *et al.* 2011, Perry *et al.* 2008). Since there is no need for a superconductive magnet in this analyser, the cost of the Orbitrap is much lower than FT-ICR, and this fact makes the Orbitrap an affordable high-resolution technique. Specification of the resolving power is 100,000 (FWHM definition) at *m*/*z* 400 (Jiwan *et al.* 2011).

1.1.4 Detectors in mass spectrometers

After ions have been separated by a mass analyser they need to be detected and transformed into an interpretable signal by the detector. Several varieties of detectors exist and the choice of detector depends on the instrument design and the analytical application. TOF instruments require detectors with a very rapid response. Isotope-ratio measurements require very stable, intercalibrated detectors, that are capable of measuring a number of ion masses at the same time (Koppenaal *et al.* 2005).

Detection of ions is based on charge, mass or velocity. Detectors may be classified into two major categories. These are direct measurement, such as a Faraday cup, which measure the direct charge current when an ion hits a surface, or detectors that amplify the intensity of the signal, such as electron- or photon-multipliers and array detectors (de Hoffmann and Stroobant 2009).

The first mass spectrometers used photographic plates. Ions with the same *m/z* hit the plate at a specific place and that position based on a calibration scale could be used to determine the particular *m/z* values. The darkness of the spots enabled an approximation of beam intensity. The disadvantages of this type of detector included poor sensitivity, short linear dynamic range, the requirement for off-line image development and calibration, poor precision and quantitation (Koppenaal *et al.* 2005).

A Faraday cup is also a direct charge measurement detector. It is made of a metal cup, ions hit the cylinder wall and they are neutralised while accepting or donating electrons (de Hoffmann and Stroobant 2009). These detectors are used for isotope-ratio MS based on their ability to provide high signal stability (Koppenaal *et al.* 2005).

Electron multiplier detectors are the most common ion detectors used in modern mass spectrometers. A schematic of a single continuous dynode, one of the EM detectors also known as a channeltron, can be seen in Figure 1.6. In electron multipliers, ions which reach the plate cause the emission of secondary electrons. The secondary electrons produced are accelerated into a continuous dynode electron multiplier, and generate an electron cascade. This can multiply the initial signal by 10^6 or more (Koppenaal *et al.* 2005).

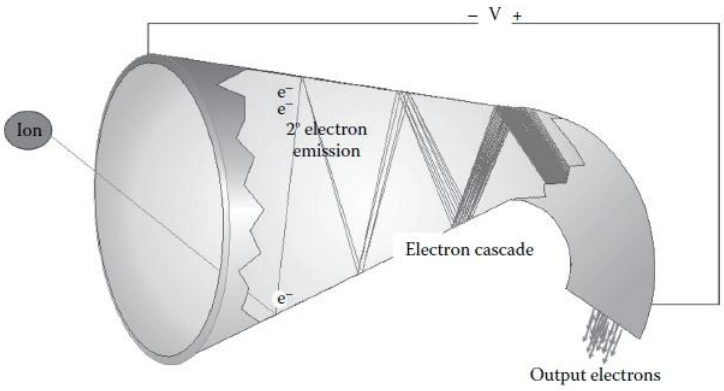


Figure 1.6. Schematic of a single continuous dynode.
Taken from (Parker *et al.* 2010)

The array plate shown in Figure 1.7 is a variant of the continuous dynode EM, called multichannel or microchannel plate. In this configuration, a series of microchannels are arranged a few tens of micrometers in diameter and a few millimetres in length in a plate or disk. This is an ideal detector for TOF analysers because the short electron path supports the very short electron pulse widths (~1 ns) used in this technique. The operation method is similar to that of a continuous dynode, in that ions strike the surface close to the entrance of the microchannel, generating a cascade of electrons. The cloud of ions exiting the microchannel can be directed onto an anode where it can be detected and converted into a mass spectrum using either time-to-digital or analog-to-digital conversion.

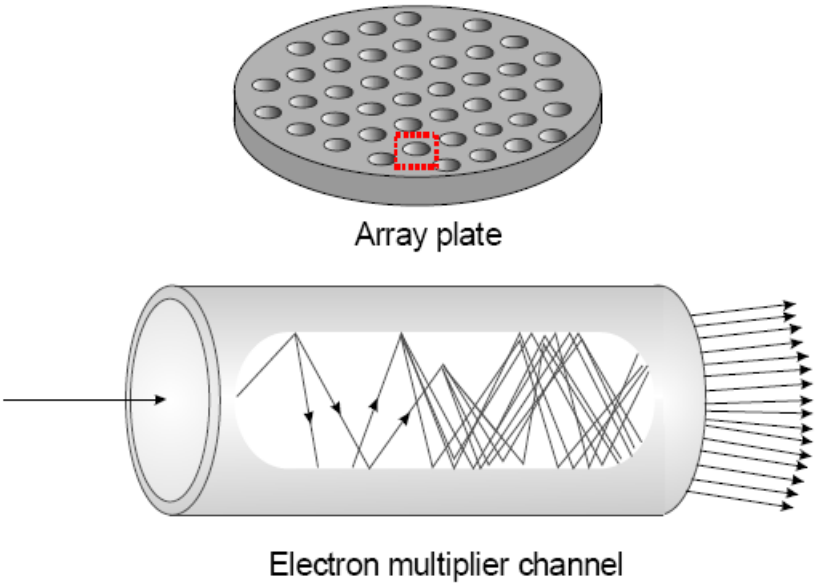


Figure 1.7. Schematic representation of a microchannel plate and the electron multiplication within one of the channels.
Adapted from (de Hoffmann and Stroobant 2009)

A time-to-digital converter (TDC) is often used with TOF instruments. Due to its pulsed nature, a TDC records an arrival time for each mass within a transient (in the case of TOF analysers ions resulting from one TOF pulse). A TDC records the arrival time for the first ion of a given mass, so it cannot tell the difference between one ion at a given mass or several ions at that given mass arriving in a transient. This inability to determine how many ions are arriving at a particular time can severely limit the instrument's dynamic range.

An analog-to-digital converter (ADC) takes the total analogue signal, reduces the high frequency background noise and converts it to a digital signal. The total signal strength is recorded at very frequent intervals (as often as a billion times per second). The ADC can distinguish between one ion of a given mass to many ions of a given mass arriving in a single transient. This results in an improved dynamic range compared to TDC. TDC detectors have about one or two orders of magnitude dynamic range whilst the ADC has a dynamic range of three to four orders of magnitude (McIntyre 2004).

1.1.5 Tandem mass spectrometry (MS/MS)

The introduction of soft ionisation techniques spurred on the development of tandem mass spectrometric (or MS/MS) techniques, which triggered the development of hybrid instruments capable of performing various types of MS/MS experiments. Soft ionisation techniques produce intact, non-fragmenting ions which make molecular mass determination possible, but do not give further information about the structure (Somogyi 2008). To obtain structural information the development of tandem mass spectrometry was necessary.

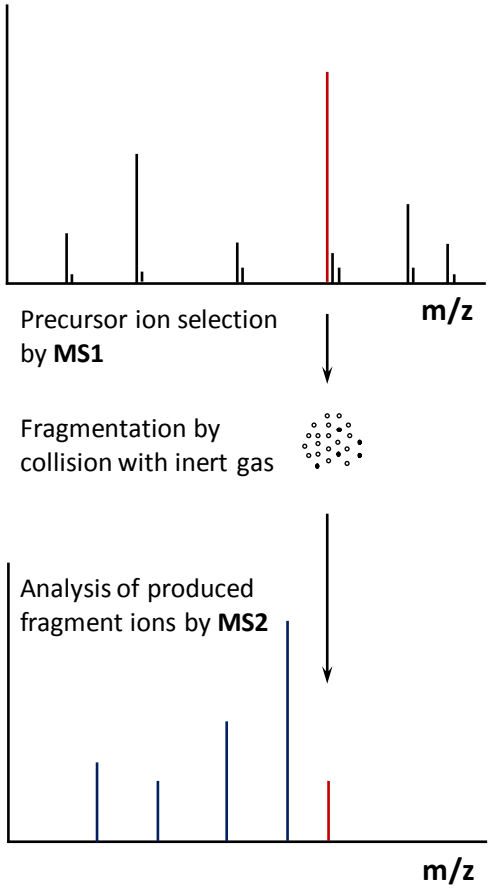


Figure 1.8. Principle of MS/MS experiment
A precursor ion is isolated by the first mass analyser, fragmented by collision with an inert gas and the products ions are analysed by the second mass spectrometer.

Tandem mass spectrometry (MS/MS) involves at least two stages of mass analysis, either in conjunction with a dissociation process or in a chemical reaction that causes a change in the mass or charge of an ion (de Hoffmann and Stroobant 2009). In a tandem mass spectrometry experiment utilising two mass analysers, the precursor

ion is isolated by a first analyser, undergoes fragmentation, (usually achieved by applying an activation method) with the generated product ions and neutral fragments analysed using a second mass analyser (Madeira and Florêncio 2012). The principle of MS/MS is illustrated in Figure 1.8.

There are two basic concepts for MS/MS based on instrumental construction. There are tandem mass spectrometry in space or in time. In MS/MS in space this is accomplished by the physical coupling of two mass analysers. Tandem mass spectrometry in time employs a single mass analyser (QIT, LIT, FT-ICR), which is operated to allow the steps of ion selection, activation and product ion analysis to take place in the same location consecutively in time and also potentially sequentally to provide third or further (*n*th) generations of fragments (MS^n) (Gross 2011).

In space MS/MS experiments are commonly performed by use of a Q-TOF (a quadrupole mass analyser coupled to a TOF instrument), TOF-TOF or triple quadrupole (QqQ) instruments. In the latter all four main scan modes may be used: product ion scan, precursor ion scan, neutral loss scan and selected/multiple reaction monitoring (SRM/MRM). These modes are illustrated in Figure 1.9. Q-TOFs are used largely for product ion scanning experiments.

While performing a **product ion scan** a precursor ion of a certain *m/z* is selected in the first analyser and all resulting product ion masses are scanned. In **precursor ion scan mode**: a product ion is chosen and the second analyser is focused on the corresponding *m*/*z* and the associated precursor ions are determined as the first analyser scans all the masses.

In a **neutral loss scan** mode all fragmentations leading to the loss of a neutral fragment of defined mass are determined. Both mass analysers operate in scanning mode, but with a constant mass offset. This scanning mode is extremely useful for screening specific compound groups with a certain side-chain which is primarily lost by fragmentation. It is often used for determination of phosphorylation sites.

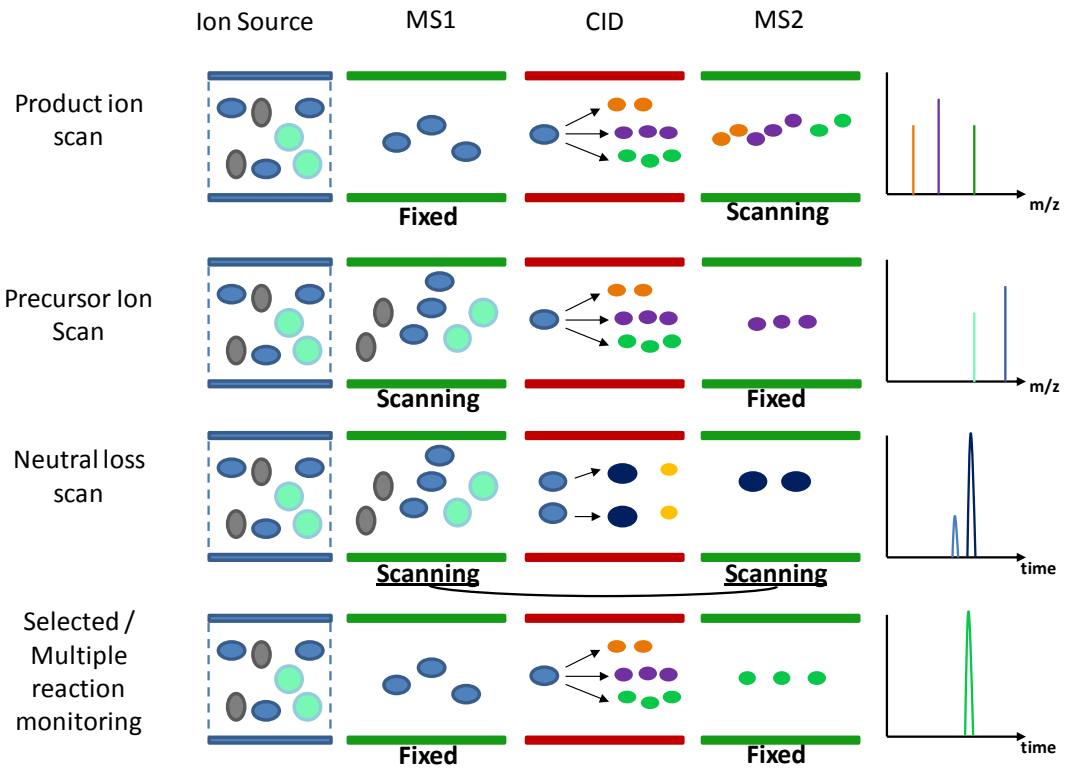


Figure 1.9. Scanning and detection modes in tandem mass spectrometry on a QqQ instrument

Adapted from (Domon and Aebersold 2006)

Multiple reaction monitoring (MRM) is a very sensitive and specific detection and quantification method for targeted molecules. MRM experiments are conducted mainly using triple quadrupole instruments. A typical MRM strategy starts by using the first quadrupole to select a molecule with a certain *m*/*z* value. In the collision cell fragmentation of the specific selected ion is facilitated by collisionally induced dissociation (CID) when supplied with optimized collision energy. The resulting fragment (or product) ions are monitored using the second quadrupole analyser and in the monitored MRM channel signals are only recorded when the specific pre-identified product ion associated with the given precursor is detected in the second quadrupole (Zanella-Cleon *et al.* 2009). A selected fragmentation pathway (transition) from a defined precursor to a defined product ion is monitored, and a signal is detected in only those cases when that selected transition occurs. For higher specificity it is necessary to monitor a number of transitions for the same target molecule.

Tandem in-space experiments set in motion the development of hybrid mass spectrometers, which can use different types of analysers for the first and second

stages of MS/MS. The main purpose of designing hybrid instruments was to combine different performance characteristics (mass resolution, the given ion kinetic energy for collision-induced dissociation, and speed of analysis) offered by various types of analysers (Glish and Burinsky 2008). One of the most commonly used hybrid instruments is the previously mentioned Q-TOF. More recent developments include QIT-TOF, LIT-TOF, LIT-Orbitrap, LIT-FT-ICR, etc. A more detailed discussion of hybrid instruments can be found in a recent review paper (Glish and Burinsky 2008).

Fragmentation methods and tandem mass spectrometry in peptide sequencing

The ion activation process is crucial to the MS/MS experiment and the choice determines the nature of the fragment ions. Various ion activation methods have been developed. Early experiments studied decomposition of metastable ions (Cooks 1973). In the 1960s collision-induced dissociation (CID) was introduced by Keith Jennings (Jennings 1968) and Fred McLafferty (McLafferty and Bryce 1967). CID is still the most common ion activation method used in tandem mass spectrometry instruments today (Sleno and Volmer 2004).

In CID, collisions between the precursor ion and an inert gas are accompanied by an increase in internal energy. The translational energy from the collisions is converted into internal energy which is distributed throughout the ion. Consecutive collisions gradually increase the internal energy to a point which leads to subsequent decomposition.

Low and high energy collisions give different fragmentations which can be complementary providing different structural information. Low-energy CID is mostly performed in quadrupole instruments (such as triple quadrupoles) and trapping instruments (such as ion traps and FT-ICR instruments). During CID in ion trap instruments, precursor ions are accelerated by on-resonance excitation causing collisions to occur. Product ions are detected by subsequent ejection from the trap (Madeira and Florêncio 2012). Collisional activation in the high energy (keV) range occurs mainly in TOF instruments, where the precursor ions have high translational energies.

Electron capture dissociation (ECD) has been developed as an alternative method to CID (Zubarev *et al.* 1998). During ECD, fragment ion production involves the capture of low-energy electrons by multiply charged ions, which results in charge state reduction and subsequent fragmentation. ECD mechanism was described as non-ergodic process in the past (Turecek and McLafferty 1984). In a non-ergodic process energy is not redistributed over the whole molecule and fragmentation at weaker bonds is not necessarily favoured (Madeira and Florêncio 2012, Turecek and McLafferty 1984). Recent publications, however, discuss cleavage mechanisms favoured based on lower transition-state energies pointing to a mechanism which is not a non-ergodic process (Simons 2010, Moss and Turecek 2012). ECD requires an ultra-high vacuum and it is currently limited to use on Fourier transform ion cyclotron resonance mass spectrometers (Gross 2011).

Electron transfer dissociation (ETD) has been developed to overcome the need for the use of FT-ICR instruments (Syka *et al.* 2004). The fragmentation mechanism is believed to be analogous to ECD and the two methods result in similar fragment ions, although the mechanisms are currently not fully evaluated and well understood.

In the first ETD experiments, singly charged anthracene anions were obtained from a chemical ionisation reaction with methane which were then used to transfer an electron to multiply protonated peptides. This electron transfer then induced fragmentation of the peptide backbone (Madeira and Florêncio 2012, Syka *et al.* 2004). Fluoranthene has become more widely used on commercial mass spectrometers (Good *et al.* 2007). ETD is available on quadrupole ion traps, recent Orbitraps and on some Q-TOFs. These analysers are lower cost and more widely-accessible then FT-ICR.

ECD and ETD are only applicable to multiply charged cationic species. The precursor ion must hold at least two, but preferably three or more, positive charges to avoid charge neutralization, which occurs when a singly charged ion undergoes electron capture and the resulting neutral is undetectable by mass spectrometer.

ECD and ETD have been shown in some cases to produce more diagnostic fragment ions than CID for large peptides and proteins. In-site specific reaction fragmentation occurs along the peptide backbone and leaves little surplus energy for additional fragmentation such as losing post translational modifications (PTM). These techniques have been widely used to investigate the position of PTMs (Chi *et al.* 2007, Stensballe *et al.* 2000).

Tandem mass spectrometry experiments are widely used for obtaining primary sequence information from peptide ions. The production of peptide fragment ions by means of CID has been well characterised (Jennings 2000). Figure 1.10 shows the nomenclature for the possible types of product ions, which was initially proposed by Roepstorff, Fohlmann and extended by Biemann (Biemann 1992, Roepstorff and Fohlman 1984).

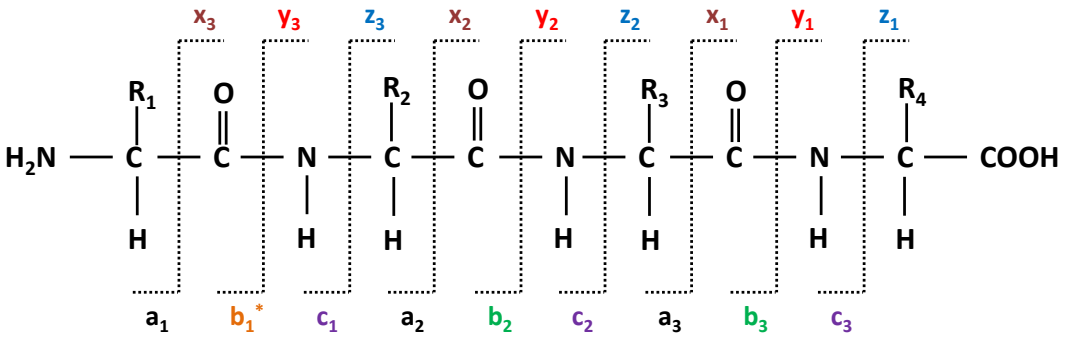


Figure 1.10. Peptide fragmentation nomenclature
adapted from (Roepstorff and Fohlman 1984) *note : b_1 ion in a non-modified peptide would not be observed because of the ring structure, but with derivatization they can be detected.

CID typically produces a series of *b* and *y* type ions from the peptide fragmenting at the amide bonds of the peptide backbone. The respective ions are formed where the *b*-series ions extend from the N-terminal and the *y*-series ions extend from the C-terminal. The CID spectra of typical tryptic peptides with a basic residue at the C-terminus often contain extended y ion and short b ion series (b_2 to b_3, b_4) excluding the b_1 ions (Seidler *et al.* 2010). The calculation of the mass difference between consecutive fragment ions belonging to the same ion series (*b* or *y* ion-series*)*, allows the determination of the amino acid sequence of the peptide (Soares *et al.* 2012). Sequence information from complementary *b*/*y* ion pairs also provides molecular weight information, the sum of their mass values gives the precursor ion mass, which can be useful if a mixture of precursor ions has been isolated for fragmentation (Seidler *et al.* 2010).

CID can be carried out on a wide range of mass spectrometers, provides reproducible fragmentation patterns, and can produce extensive coverage for tryptic peptides. CID fragment ion spectra have been shown to differ when obtained using an ion trap

compared to a Q-TOF instrument. Ion trap MS/MS spectra are very reproducible and dominated by first generation fragments, however the low mass cut-off excludes fragment ion detection below *m*/*z* 200-300 for average sized tryptic peptides (Seidler *et al.* 2010).

Peptide ions with charge state 2+ and 3+ are preferred precursor ions since the resultant fragment ion spectra of these charge states often contain more sequence ions. When tryptic digestion efficiency is high, the generated peptides are present with 2+ or 3+ charges, because the basic sites or Arg/Lys, N-terminus and His can hold charges. Fragmentation of doubly charged peptides usually results in the production of complementary *b*/*y* ions. The *y* ion retains a positive charge at its C-terminus, whilst the *b* ion retains charge at the N-terminus.

High energy fragmentation can be used for determining the side-chain of an amino acid, which is useful for distinguishing between leucine and isoleucine (Williams *et al*. 2009). MS^n in an ion trap can also be useful for this purpose.

Other activation methods such as ECD and ETD lead to the direct generation of c- and z-type fragment ions by breaking the Cα-N bond. ECD and ETD can produce a complete or almost complete fragmentation spectrum of peptides leading to more information regarding the peptide sequence (Mikesh *et al.* 2006).

A large number of publications have compared ETD (or occasionally ECD) to CID in terms of accessible information content from peptide fragmentation (Cook *et al.* 2012, Creese and Cooper 2007, Hart *et al.*, Molina *et al.* 2008, Sobott *et al.* 2009, Zubarev *et al.* 2008), and a number of experiments have been described which combine the use of ETD and CID for more comprehensive peptide sequence analysis (Han *et al.* 2008, Hartmer and Lubeck 2005, Kim *et al.* 2011).

Since the introduction of CID, tandem mass spectrometry has become the primary method of choice for the sequencing of proteins replacing the previously used Edman degradation method and it is becoming more and more important in the investigation of protein function (Jennings 2000). A summary of wider applications of tandem mass spectrometry can be found in the recently published book edited by Jeevan K. Prasain (Prasain 2012).

1.1.6 Liquid chromatography

In the analysis of complex biological samples there is a requirement for separation techniques to simplify mass spectra and enable interpretation. The MS analysis can be complicated when there are hundreds of components present, and often ion suppression can occur during ionization in the ESI source. With the development of ESI it became possible to couple ESI to liquid chromatography allowing separation of complex samples before mass spectrometric analysis.

Chromatography is a separation method in which the components to be separated are distributed between two phases called the stationary phase and the mobile phase. As the mobile phase moves in a definite direction in contact with the stationary phase, the components, depending on their affinity to the stationary phase, are either retained on the stationary phase or solubilised in the mobile phase. In a mixture, different compounds have different physical/chemical properties and so theoretically varying affinity to the stationary phase. This enables elution from the stationary phase at different times allowing chromatographic separation to occur. Analytes with relatively weak affinities for the stationary phase elute early and so are detected first.

In high pressure liquid chromatography (HPLC) the mobile phase is liquid delivered under high pressure (up to around 400 bar) through a column packed with particles covered with a particular compound. When the mobile phase composition is kept constant during elution this is referred to as isocratic separation. Gradient elution is when the composition is changed gradually. By applying gradient elution generally a better separation can be achieved. Physical properties of compounds such as size, binding affinities, hydrophobicity, charge and p*I* have been used to separate compounds of interest.

A schematic of a typical HPLC system, illustrating its major components can be seen in Figure 1.11.

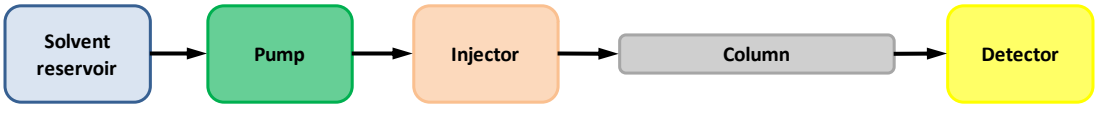


Figure 1.11. Schematic of a HPLC system

High performance liquid chromatography (HPLC) was developed in the 1960s (Horvath and Lipsky 1966, Horvath *et al.* 1967) and there were several efforts to

couple it to mass spectrometry with slowly growing success by McLafferty's group in the 1970s (Arpino *et al.* 1974), and by Fenn's group together (Whitehouse *et al.* 1985), Henion's lab at Cornell and Andries Bruins developed the most practical interface in the 1980s (Bruins *et al.* 1987, Covey *et al.* 1986, Henion 2009). Historically, the main problem was how to handle high liquid flows while maintaining the high vacuum required for the mass spectrometer. This was achieved and revolutionised bio analytical research (Abian 1999, Thomson 1998).

Why was it so beneficial to couple the two techniques? Samples of biological origin can contain hundreds to thousands of components present at varying mass and concentration. Analysis of these mixtures by mass spectrometry alone is difficult. It is beneficial to provide separation of these components prior to MS to allow their identification or determine their levels. Chromatography alone as a technique is unable to provide unambiguous identification of the separated components. Recognition can only be based on the measurement of retention times, often by comparison with reference standards, which does not make the identification of unknown compounds possible. Even with matching retention times the uncertainty of the presence of a different compound other than that identified is high. The mass and mass fragmentations of many compounds are sufficiently specific to allow identification with a higher degree of confidence. If a single analyte of interest is in a mixture the mass spectrum can contain ions from all of the compounds present, and the identification is more challenging, in addition accurate quantitation of the compound is difficult. By combining HPLC with mass spectrometry a powerful analytical technique has become available to provide confident identification and reliable quantitative determination of compounds in a complex mixture (Ardrey 2003).

Reversed phase-HPLC (RPLC) has emerged as the most commonly used method for peptide separation. The separation of non-polar compounds happens on a stationary phase containing non-polar functional groups such as long alkyl chains (for example C_8 and C_{18}), and involves a mobile phase that is significantly more polar than the stationary phase which is often water-based with combination of organic solvents such as acetonitrile or methanol. The term reversed phase has a historic origin, the early liquid chromatography separations were performed using a stationary phase with unmodified silica which is hydrophilic and able to bind polar compounds. This is today called normal phase chromatography. Employing aqueous solvents and

separating non-polar compounds under these conditions was not possible. The introduction of alkyl chain coated stationary phases which were hydrophobic and the opposite of hydrophilic silica, hence the term reversed, made it possible to routinely analyse biological samples and organic non-polar analytes like peptides.

The peptides bind to the non-polar stationary phase and during the chromatography separation the composition of the mobile phase is altered gradually increasing the concentration of the organic content over a period of time. The increasing organic concentration will disrupt the binding mechanism between the analytes and the stationary phase, which assists the elution of the hydrophilic peptides first, followed by increasingly hydrophobic peptides during the rest of the gradient. This course of action is presented in Figure 1.12.

The solvents used for mobile phase in RPLC are compatible with ESI, which makes coupling this type of separation technique to the mass spectrometer easier.

The separation efficiency of HPLC is influenced by several factors. It is necessary to choose and optimise the content of the mobile phase, the pH and the chemistry of the stationary phase. The stationary phase is usually a packed column which can vary in its particle size, pore size, column length and column diameter. The operating temperature of the column, the length and slope of the applied gradient, and the flow rate all affect the chromatographic separation.

Separation efficiency is measured in theoretical plates. In order to optimize separation efficiency, it is necessary to maximize the number of theoretical plates and reduce the plate height, which is inversely proportional to the plate number and also influenced by the column length. The plate height is related to the flow rate, and according to the van Deemter equation separation efficiency can be maximised by optimising the flow rate (VanDeemter *et al.* 1995).

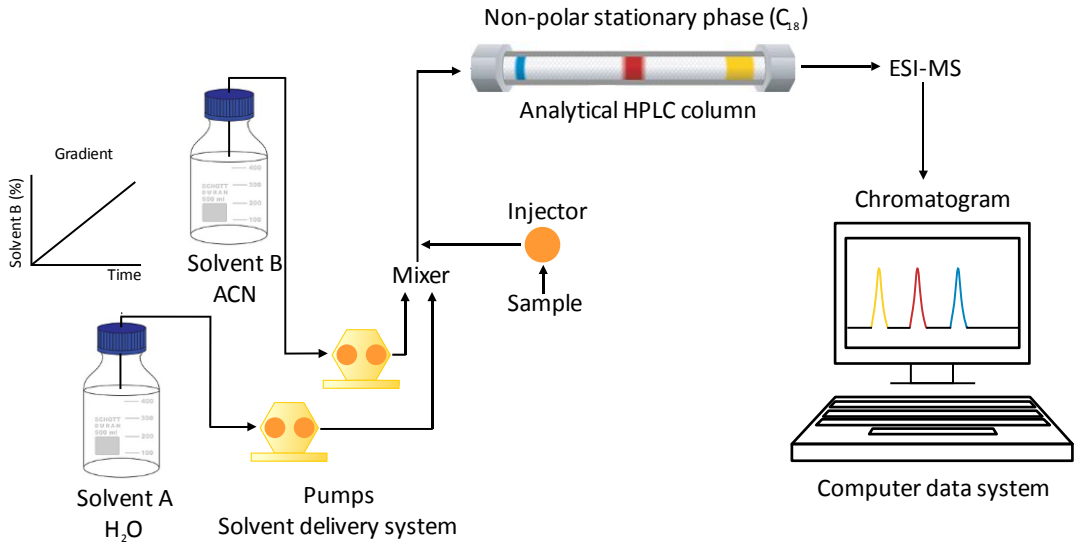


Figure 1.12. Schematic of reversed phase liquid chromatography under gradient conditions

Adapted from (McDonald 2009)

Particle size is one of the variables, and the evolution of liquid chromatography facilitated the production of smaller particle sizes. Reducing the particle size results in higher speed and peak capacity (number of peaks resolved per unit time in gradient separation) (Swartz 2005a). Smaller particle size and higher flow rates made it necessary to handle much higher operating pressures, which led to the introduction of ultra-high pressure liquid chromatography (UPLC). MS detection is significantly enhanced by employment of UPLC (Swartz 2005b). The obtained higher chromatographic efficiency results in better resolution and higher peak capacity, which is particularly important for the analysis of peptides and proteins (Tolley *et al.* 2001).

Faster separations achieved by UPLC are beneficial for high throughput applications, where time saving is extremely important while running multiple samples. The application of liquid chromatography for the separation of biological mixtures can require long and thorough method development and optimisation, which is also shortened by the use of UPLC.

1.2 Clinical applications of mass spectrometry

As routine analysis of biological samples and biomolecules has become possible it is inevitable that this powerful technique would become established for biomedical and clinical applications. The use of MS in biomedical sciences is rapidly increasing and the potential of MS in medical diagnostics grows. To detect physiological disorders, molecular disease markers can be monitored. These markers can be found in the following systems: the metabolome - the products or substrates of endogenous metabolism, the proteome and peptidome - functional/dysfunctional protein and peptide products of gene transcription and translation, and the genome and transcriptome - disease-related genes and mRNA transcripts (Flad and Tolson 2005). The sources of these markers are from biological fluids such as serum, plasma or urine, and because of the growing need for non-invasive diagnosis even saliva has become a more widely used source recently. For the diagnosis of molecular markers, antibody recognition, methods like enzyme linked immunosorbent assays (ELISA) or polymerase chain reactions (PCR) for genetic determination have been used. Mass spectrometry is a sensitive tool, with high specificity and high-throughput capabilities. The ability of MS to screen for multiple molecules in a single run has shown promise has already proved to be useful in the diagnostics field and is forseen to strengthen its position in the coming years (Petricoin and Liotta 2003).

Mass spectrometry has long been used in clinical laboratories. In the 1970s gas chromatography (GC)-mass spectrometry was used for drug screening and the diagnosis of organic acid metabolic disorders as reviewed by Shushan (Shushan 2010). Since GC requires a certain level of analyte volatility, derivatization needed to be applied to those biological molecules that were involatile. After successfully coupling HPLC to MS, providing a more appropriate separation technique for polar thermolabile molecules, LC-MS was introduced into the field of neonatal screening for metabolic disorders and became a widely-accepted method in clinical diagnostics. More than 30 different metabolic disorders can now be simultaneously and quantitatively screened using tandem MS (Flad and Tolson 2005). MS has been shown to be a good alternative to currently used screening methods in many other fields. Some examples and importance are discussed below.

1.2.1 Analysis requirements: sample preparation, instrumentation and costs

The most commonly used MS instruments for clinical purposes are triple quadrupole (TQ or QqQ) analyser instruments. Triple quadrupoles achieved this popularity due to their high specificity and capability for quantitative measurements for small molecules using the selected reaction monitoring mode with frequent usage of isotope labelled standards. This type of measurement provides good sensitivity for a wide variety of compounds in targeted applications, but other instruments can be better for applications where wider screening for unknown components is necessary. Ion traps are also relatively popular and are significantly less expensive than TQ mass spectrometers. Ion traps are not as suitable for quantitative applications but their strengths lie in more qualitative screening applications such as identification and structural characterization of unknown metabolites (Vogeser and Seger 2008). High resolution mass spectrometry also has a significant role in biomedical research but these studies are usually confined to academic laboratories. Orbitrap and TOF instruments can provide accurate mass data and potential structural elucidation, which can be extremely useful in large scale screening, and with the inevitable further improvements in quantification performances these high mass resolution instruments will be adapted for use in clinical laboratories (Jiwan *et al.* 2011) .

The advantages of LC-MS/MS include high specificity and selectivity, better sensitivity for small molecules when compared to immunoassays, the ability to carry out multi-analyte screening methods with no additional cost, and rapid method development. There are some disadvantages which need to be overcome in the following years in order for LC-MS/MS to be more suitable for high-throughput medical diagnostic/screening. A significant amount of pre-analytical sample treatment can be required producing toxic or clinical waste. An LC-MS/MS system is based on technology which can require skilled and trained staff, and the initial capital cost of these instruments is relatively high, as are the maintenance service costs (Herman and Shushan 2012). Several review papers discuss HPLC-MS/MS, and capillary electrophoresis-MS/MS is also becoming more widely used (Kolch *et al.* 2005).

Working with MS analysis of biological matrices can be challenging due to matrix interferences, ion suppression and sensitivity loss due to contamination. Several purification steps may be carried out before HPLC–MS/MS to help alleviate these issues. Depending on the concentration range, sample concentration or dilution has to be performed. For the analysis of small molecules there is a need to remove as much protein and lipid as possible and the most widely applied technique for protein precipitation is the addition of organic solvents. There are continuously developing methods for automated on-line and or off-line sample preparation prior to HPLC separation, otherwise lengthy chromatography must be used to separate the large number of compounds what may be present in the biological matrix (Herman and Shushan 2012, Vogeser and Kirchhoff 2011).

To make MS routine in clinical diagnosis, sample pre-treatment, separation, and detection must become integrated. This could be achieved by implementing on-line multidimensional chromatography to facilitate sample introduction, ease-of-use, and speed of analysis (Shushan 2010). An important consideration is the cost of the technique. The cost of the mass spectrometer can be a high initial investment, but in a long term, using it for the screening of multiple analytes simultaneously, this initial high capital cost can be recovered. To achieve this cost-effectiveness many applications may need to be implemented into one analytical method to increase productivity. Measurement and chromatographic times need to be shortened to make the technique more appropriate for high-throughput analysis to cope with the large numbers of samples (hundreds of samples per day) required for population screening. From a hospital point-of-view the method needs to be automated, easy-to-use and rapid with a report readily accessible giving the required answers.

1.2.2 Current clinical MS applications

At present, the most widely performed MS clinical applications are for therapeutic drug monitoring (TDM) and metabolic disorder diagnosis.

Vitamin D deficiency has been linked to skeletal disease conditions, chronic renal failure and osteoporosis. Vitamin D analysis is one of the most often used LC-MS/MS assays in the clinical laboratory today. This technique is widely accepted as it is able to distinguish between analogs of vitamin D while making it possible to quantify the diagnostically informative 25-hydroxyvitamin D metabolites. There is

some debate as to whether immunoassays give accurate results due to potential nonspecific binding of different vitamin D metabolites simultaneously, and experts consider LC-MS/MS the most accurate assay for vitamin D analysis (Herman and Shushan 2012). Recent publications discuss vitamin D analysis in further detail (Babic 2012, Higashi *et al.* 2010).

Therapeutic drug monitoring (TDM) is one of the most important fields for LC-MS/MS, especially in supporting therapies involving immunosuppressants. There are several other screening methods already implemented for anti-fungal drugs, anti-viral drugs, anticonvulsants, antidepressants, antibiotics, anticancer drugs and drugs affecting the cardiovascular system (Adaway and Keevil 2012). Monitoring of immunosuppressants can be critical to achieve optimal patient care after organ transplantation. Elevated dosing can result in significant toxicity while too low dose of the drug can result in transplant rejection. This provides a narrow therapeutic window and requires extreme care to maximize the beneficial therapeutic reaction while minimizing undesirable side effects (Herman and Shushan 2012).
HPLC-MS/MS is considered superior to immunoassays for immunosuppressant TDM (Yang and Wang 2008). The most frequently used drugs include cyclosporine A, tacrolimus, sirolimus and everolimus, which can be all monitored simultaneously with one method (Karapirli *et al.* 2012).

Newborn screening for inborn errors of metabolism with MS has a large history and is a widely accepted method in clinical laboratories. Newborn Screening (NBS) for metabolic disorders started with the introduction of Guthrie's test for phenylketonuria (PKU) using bacterial inhibition in the 1960's (Guthrie 1969, Guthrie and Susi 1963). By 2010, every infant in major developed countries could be screened for inherited metabolic disorders using tandem MS (Chace 2009). In the UK infants are screened for only certain metabolic disorders, including phenylketonuria (PKU), congenital hypothyroidism (CHT), medium-chain acyl-CoA dehydrogenase deficiency (MCADD) (http://newbornbloodspot.screening.nhs.uk).
The most well-known metabolic disorder phenylketonuria results form a deficiency in the enzyme phenylalanine hyroxylase, which catalyses the oxidation of phenylalanine to tyrosine. When phenylalanine is substantially increased and untreated it causes mental retardation (Chace 2003). This can be easily avoided by strict dietary treatment.

Other metabolite disorder screening involve acylcarnitines, amino acids, fatty acid oxidation and organic acid metabolism. These metabolism disorders are individually rare but collectively may be as frequent as 1 affected infant in 500 newborns (Vekey *et al.* 2008). It is extremely important to screen for these disorders with the most specific technique possible to decrease the numbers of undiagnosed and untreated infants.

The list of applications of mass spectrometry in clinical and biomedical labs is extensive. Other examples include human serum albumin measurement in urine as a biomarker for renal disease, proteins and peptides in blood (Zhu *et al.* 2010) steroids hormones (Kushnir *et al.* 2009), proteins and proteomics applications to support biomarker discovery for different cancer research or other therapeutic purposes (Indovina *et al.* 2012, Matta *et al.* 2010, Silberring and Ciborowski 2010, Tan *et al.* 2012, Zhang, X. *et al.* 2007), neuroscience and measurement of neurotransmitters, neuropeptides, neuroproteomics (Drabik *et al.* 2007), and clinical and forensic toxicology (Peters 2011). One of the most promising fields for biomarker discovery is metabolomics (Dettmer *et al.* 2007). The number of expressible genes is more than 30,000 and the number of proteins expressed is close to 1,000,000, while the number of metabolic products in the human body is estimated to be around 3,000, which means less of a challenge to identify, quantify, and validate their potential diagnostic usage (Shushan 2010).

Recently the *MALDI Biotyper* system from Bruker has become a popular instrument in clinical laboratories. The MALDI-TOF mass spectrometry-based identification of microorganisms is currently used in over 700 clinical and non-clinical microbiology laboratories worldwide (Kliem and Sauer 2012).

Another opportunity for MS/MS-based methods in population screening (newborn and also antenatal) is the diagnosis of hemoglobinopathies, which is discussed in Section 1.3.4.

1.3 Hemoglobinopathies and their analytical detection

1.3.1 Biological background

Hemoglobin (Hb) is the protein in red blood cells that is responsible for the delivery of oxygen from the lungs to the tissues and the transport of carbon dioxide from the tissues to the lungs (Thein 2011). It is the most abundant protein in mammals, and probably the best characterised protein on the planet (Buxbaum 2007). The structural changes that occur during binding of O_2, CO_2 and poisons like CO are also well characterised, and the structural alterations of the protein due to mutations have also been thoroughly investigated (Buxbaum 2007).

Human hemoglobins have a tetrameric structure consisting of two α-like (α and ζ) and two β-like (ε, γ, δ or β) globin chains, each linked to a heme group (Thein 2011). Different types of hemoglobin transfer oxygen at different stages of development. During the first 3 months of development, embryonal hemoglobins are developed in the yolk sac, these are Hb Portland ($\zeta_2\gamma_2$), Hb Gower I ($\zeta_2\varepsilon_2$), and Hb Gower II ($\alpha_2\varepsilon_2$). As the bone marrow of the embryo develops, it starts producing the major fetal hemoglobin (HbF) which has the structure of $\alpha_2\gamma_2$ (Buxbaum 2007).

The switch from embryonic to fetal hemoglobin production begins at around the 5th week of gestation and is completed by week 10 (Thein 2011). Normal adult hemoglobin is HbA ($\alpha_2\beta_2$) and comprises around 95-97% of the total hemoglobin with a minor component of HbA_2 ($\alpha_2\delta_2$) in the red blood cells. β-globin expression starts only a couple of weeks later than fetal Hb, at around week 8 but the synthesis remains low, increases to approximately 10% at weeks 30–35 of gestation and then production jumps radically to higher levels at birth, when the level of fetal Hb is around 60-80%, declines to 5% by the age of 6 months slowly reaching the level of normal adult around 1% at years 2 (Thein 2011). Normal adult blood contains about 2-3.5% HbA_2. Production starts about 3 months after birth (Wild and Bain 2006).

The production of the different hemoglobins and their levels during development is shown in Figure 1.13.

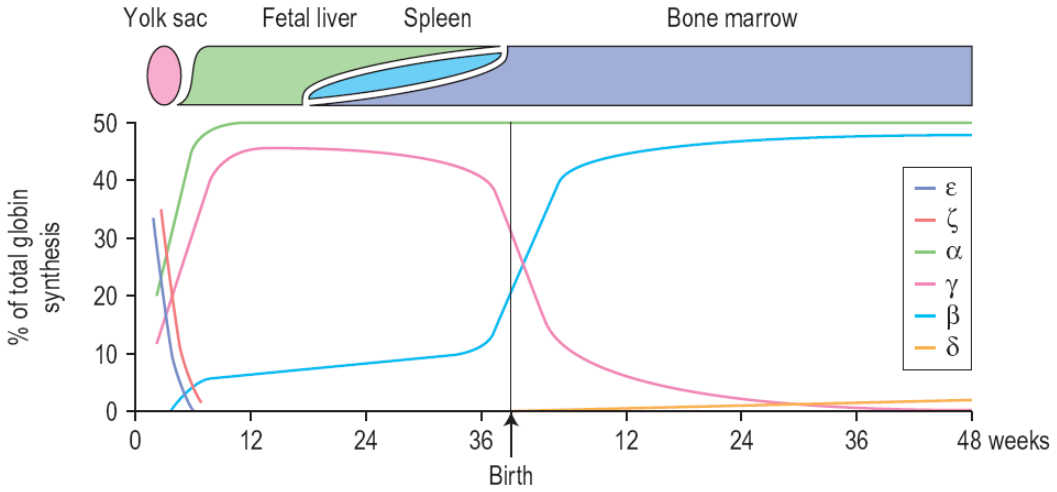


Figure 1.13. Diagrammatic representation of average percentages of the different types of hemoglobin chains synthesized during the different developmental periods and infancy.
Taken from (Thein 2011)

Embryonic and fetal hemoglobins have a higher affinity for oxygen than adult hemoglobin, which is necessary for the efficient transfer of oxygen across the placenta from maternal to fetal circulation (Thein 2011).

In the quaternary structure of HbA the $\alpha_1\beta_1$ (and equivalent $\alpha_2\beta_2$) dimer are held together tightly by ionic and hydrophobic bonds which involve a number of amino acids with interlocking side chains; the contact between $\alpha_1\beta_2$ (and equivalent $\alpha_2\beta_1$) involves fewer amino acids and consequently is less strong, the contacts between similar chains being relatively weak. A heme group is attached to each. These are responsible for the oxygen-carrying capacity of the molecule and also stabilise the whole molecule (Bain 2006a, Buxbaum 2007, Wild and Bain 2006). The connecting hydrogen bonds allow alteration of the structure to some extent during oxygen binding and release. The heme is a porphyrin molecule with an iron bound in its centre by complex bonds to the 4 ring nitrogen atoms. The iron is normally in the ferrous form (Fe^{2+}), which makes it possible for the heme to reversibly bind oxygen in order for hemoglobin to function as an oxygen transporting protein. Oxidation of iron to the ferric form (Fe^{3+}) makes the hemoglobin incapable of binding oxygen (methemoglobin) (Bain 2006a).

The structure of the hemoglobin tetramer is shown in Figure 1.14.

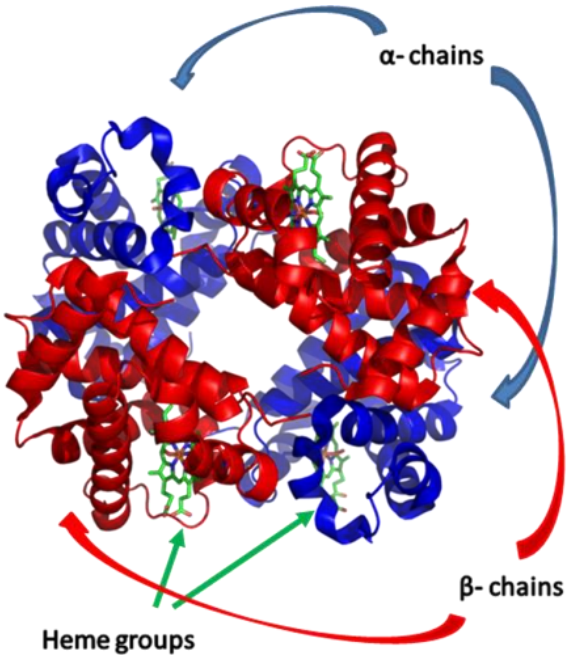


Figure 1.14. Structure of hemoglobin

The α-like and β-like globin chains are encoded by genetically distinct gene clusters. The α-like cluster is positioned on chromosome 16 with duplicate α genes and the β-like cluster on chromosome 11, as shown in Figure 1.15. The genes are arranged along the chromosome in the same order in which they are expressed during development (Thein 2011). Each gene contains 3 coding segments (exons) and 2 non-coding segments (introns). The latter have to be spliced out of the mRNA prior to protein synthesis. Synthesis of the hemoglobin subunits is synchronised to produce equal amounts of α- and β-subunits (Buxbaum 2007).

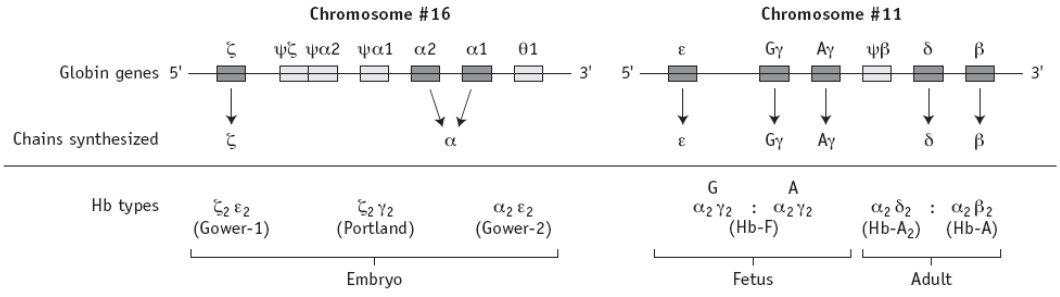


Figure 1.15. The α- and β-globin gene clusters and the corresponding hemoglobins produced.
Taken from (Wild and Bain 2006)

The α and β chains consist of 141 and 146 amino acid residues, respectively. There is some sequence homology between them (64 amino acid residues in identical

positions). The β chain differs from the δ chain by only 10 amino acids and from γ chain by 39 residues (Clarke and Higgins 2000, Shackleton *et al.* 1991). The γ globin genes are duplicated like the α genes in the gene cluster but, while the two α globin genes have theoretically the same product, the produced γ globin genes are slightly different ($^{G}\gamma$ = γ 136Gly; $^{A}\gamma$ = γ 136Ala). In fetal red cells, the ratio of $^{G}\gamma$ to $^{A}\gamma$ is approximately 3:1; in adult red cells, it is approximately 2:3 (Wild and Bain 2006).
Hemoglobin can undergo post-translational modification. By a spontaneous reaction between the N-terminal valine of the β globin and glucose, HbA_{1c} is produced from HbA. This reaction is a non-enzymatic glycosylation. The concentration of HbA_{1c}-formation, based on its accumulation during the 120 days life span of an erythrocyte, is related to the average blood glucose concentration (Buxbaum 2007). In normal individuals HbA_{1c} is around 4-6%, but in diabetics it can be much higher (Bain 2006a). Monitoring of HbA_{1c} levels can be used as a marker for long term diabetic control.

The DNA of globin genes may be subject to spontaneous mutations. The mutation could be a point mutation, affecting a single nucleotide, or a more complicated mutation including deletion, insertion or alteration of more than one nucleotide. As a result of a point mutation there can be expressed variations on the protein level or no effect on the amino acid sequence (Bain 2006a). There are many genetically determined variants of human hemoglobin although only a few have serious clinical manifestations. The inherited disorders of hemoglobin synthesis are referred to as hemoglobinopathies (Bain 2011, Wild and Bain 2006).
Hemoglobinopathies are categorised into two main disease groups, structural variants and thalassemias. These are genetically distinct groups, but the clinical manifestations are similar. The hemoglobin variants (also called abnormal hemoglobins) arise from an alteration in the globin protein structure with a change in type or number of amino acids. Thalassemia results from gene mutations that cause production of an insufficient amount of normal structure globin chains (Hartwell *et al.* 2005). To date over 1170 different hemoglobin variants have been described with approximately 460 mutations potentially giving rise to thalassemia (based on the most recent information of Globin Gene HbVar database August 2013) (Hardison *et al.* 2002, HbVar). Hemoglobin disorders are genetically inherited and it is estimated

that ~7% of the world's population are carriers for one of them, making hemoglobinopathies the commonest monogenic diseases (Thein 2011).

Structural variants

Alterations in the structure of hemoglobin usually occur by point mutations affecting one or, in less frequent cases, two or more bases, coding for amino acids of the globin chains (Wild and Bain 2006). Although there are plenty of potential structural variants, based on the calculations by Kleinert *et al* for α- and β-chains together 1695 theoretical variants are possible, of which only 43% were detected in 2008, and 69% (1170) listed in the HbVar database today, a significant number of the detected variants being clinically silent (Kleinert *et al.* 2008). Additionally, there are also 460 thalassemia disorders listed in the HbVar database (http://globin.bx.psu.edu/hbvar/menu.html).

The most important structural variants which have potential significant clinical manifestation are HbS, C, E, D^{Punjab}, O^{Arab}. These result from single amino acid substitutions in the β-chain. Other mutations include a combination of the sequences of β and γ or δ chains resulting in hybrid hemoglobins. The β and δ combinations are known as Lepore and are also clinically important (Wild and Bain 2006). These mutations and their positions are outlined in Table 1.1.

Hb S results from a point mutation in the β-globin gene (GAG → GTG in codon 6) causing substitution of valine for glutamic acid in position 6 of the β-globin chain (β6Glu→Val). The substitution of a charged amino acid by a hydrophobic one creates a hydrophobic patch, which is only exposed in deoxy-hemoglobin, and hidden when oxygen is bound (Buxbaum 2007). As a consequence the variant hemoglobin (Hb S) can polymerize in the deoxygenated state, resulting in the characteristic distorted red blood cell shape change, the so-called sickle cell. Clinical symptoms occur only with homozygous sickle cell disease. Heterozygotes are said to have sickle cell trait and are usually symptom-free (Wild and Bain 2006). Heterozygous individuals are resistant to infection with *Plasmodium falciparum*, the causative agent for malaria. This parasite spends part of its life cycle in erythrocytes, feeding on hemoglobin. The presence of Hb S decreases the erythrocyte life span (in heterozygotes less than in homozygotes), and so the parasite cannot complete its life cycle (Buxbaum 2007).

Name	Position (Tryptic peptide affected by the mutation)	Mutation (Substitution)	Mass change of the β chain (Da)	Clinical manifestation Homozygote/Heterozygote
Hb S	β 6 (βT1)	Glu→Val	-30	Sickle cell disease/ no symptoms
Hb C	β 6 (βT1)	Glu→Lys	-1	Mild anemia/no symptoms
Hb D-Punjab,	β 121 (βT13)	Glu→Gln	-1	Mild anemia/no symptoms
Hb E	β 26 (βT3)	Glu→Lys	-1	Thalassemia minor/mild microcytosis
Hb O-Arab	β 121 (βT13)	Glu→Lys	-1	Mild anemia/no symptoms
Lepore-Boston-Washington	Hybrid: δ 1-87, β 116-146		-2	Thalassemia minor
Lepore-Hollandia	Hybrid: δ 1-22, β 50-146		-30	Thalassemia minor
Lepore-Baltimore	Hybrid: δ 1-50, β 86-146		-45	Thalassemia minor

Table 1.1. Common clinically significant variants

The homozygous state or sickle cell anaemia (β genotype SS) causes moderate to severe hemolytic anaemia. The main clinical symptoms are repeated vascular occlusion by sickled red cells resulting in acute crises and eventually end-organ damage. The clinical severity of sickle cell anaemia can be very variable (Wild and Bain 2006).

Hb C results from a single base-pair mutation in the β-globin gene similar to Hb S but Glu-6 is replaced by Lys rather than Val. A negatively charged amino acid is replaced by a positive one, which causes decreased solubility, and tendency to crystallise. Hb C does not give a positive sickle solubility test. It causes a moderate hemolytic anemia in the homozygous form. Heterozygotes are asymptomatic, patients are carriers, but experience no symptoms. Interaction with β0 (the complete absence of β-subunits) and β+ (diminished production of the β-subunits) thalassaemia trait results in mild or moderate hemolytic anaemia. Hb C is the second most common structural hemoglobin variant in people of African descent (Wild and Bain 2006).

Hemoglobin SC disease is a compound heterozygous state for Hb S and C resulting in a milder form of sickle cell disease. Haemolytic crisis may occur in special situations (e.g. during child birth or surgery) (Lawrence *et al.* 1955).

Hb O-Arab and Hb D-Punjab both result from point mutations of the β-globin gene, affecting the Glu in position 121, replacing it with Lys and Gln respectively. They have clinical significance only in co-inheritance with Hb S and β-thalassemia, and so may give rise to severe sickle cell disease.

Hb E arises from a point mutation resulting from a glutamic acid to lysine change at position 26 of the β-chain. ***Hb Lepore*** is a hybrid. Both of these have thalassemic manifestations that lead to decreased globin chain production with hypochromic microcytic anemia, a type of anemia characterized by small red blood cells (for microcytic, the measured mean corpuscular volume (MCV) - measure of the average red blood cell volume- is <80fL) which are paler than usual (for hypochromic). The Lepore globin gene resulted from a crossover between the delta and beta globin gene loci (Itano *et al.* 1954).

Heterozygotes of these mutations are mainly asymptomatic, although the blood of Hb C trait carriers contains target cells, and Hb E carriers may have very mild anemia. Even homozygotes for Hb C and Hb E have only mild hematologic abnormalities and mild anemia. Compound heterozygotes for Hb E and β-thalassemia usually have the phenotype of transfusion-dependent β-thalassemia (Steinberg 2011).

Sickle cell diseases (SCD) include sickle cell anemia (SS) and compound heterozygous syndromes which have sickling clinical abnormalities such as SC, Sβ thalassemia, SD-Punjab, SO-Arab, SLepore (Bain 2011).

Bone marrow transplantation offers the only hope of a cure for SCD. This is not available for many patients. For the majority of SCD patients there are two non-transplant options for the management of the disease – blood transfusion and hydroxyurea therapy (Thein 2011). HbF has been shown to inhibit HbS polymerization. Experimental treatments to induce higher levels of HbF are undergoing clinical trials (Steinberg 2011).

Some rare hemoglobin variants may have altered affinity for oxygen (hemoglobins with increased or decreased oxygen affinity), be susceptible to oxidation, or can have

molecular instability (unstable hemoglobins). Amino acid substitutions affecting heme-binding residues may lead to irreversible iron oxidation, methemoglobinemia (Hb M), and cyanosis. Substitutions involving contacts between globin subunits may alter the affinity of hemoglobin for oxygen (Steinberg 2011).

Thalassemias

Thalassemias are also caused by mutations in globin genes, but in thalassemia the globin gene expression is affected in a way that the synthesis of a globin chain is reduced or absent, but the structures of the globin chains which are produced are usually normal (Steinberg 2011).

The thalassemias are classified into α, β, δβ, γδβ, δ, γ and εγδβ groups according to the type of globin chain(s) expression that is reduced. The two major categories are the α and β thalassemias while the others are rare. Failure of hemoglobin synthesis is a main cause of microcytosis and anemia, which can vary from mild to severe (Hartwell *et al.* 2005). β globin chains are encoded by a single gene; α globin chains are encoded by two closely linked genes. In diploid cells there are two loci encoding the β chain, and four loci encoding the α chain, which explains why the possibility of having β thalassemia is higher. Production of the subunit can be totally absent (α^0- and β^0-thalassaemias) or reduced (α^+- and β^+-thalassemias). Clinical manifestations arise from completely asymptomatic microcytosis to severe anemia (Wild and Bain 2006). The term thalassemic is often used for certain variants, where the mutant globin chain is synthesised at a reduced rate, leading to a thalassemic disorder. An example is Hb Constant Spring which causes non-deletional α-thalassemia, or Hb E which often evolves into β-thalassemia intermedia or even major in compound heterozygotes with β thalassemia (Bain 2011).

Beta-thalassemia

In β-thalassemia the α-subunits are produced at a normal level, but in the absence of equivalent amounts of the β-subunits the normal hemoglobin cannot form and the excess α-subunits precipitate in the cells. There are only two copies of the β-globin gene. Affected individuals either have the thalassemia trait when only one gene is affected (**β-thalassemia minor**, no treatment required) or **β-thalassemia major**,

when both genes are affected. Individuals with thalassemia major appear healthy at birth, since fetal hemoglobin is still present, but become rapidly anemic as HbF cannot be replaced by HbA. Diminished β-chain production is known as β^+-thalassemia, while the complete absence of β-subunits is referred to as β^0-thalassemia.

Many different mutations can cause β-thalassemia. This usually results from mutations that affect transcription, translation, or RNA stability. More than 300 β-thalassemia mutations have been characterized, and probably only 20 β-thalassemia alleles account for more than 80% of the β-thalassemia mutations in the whole world. Unlike α-thalassemia, where deletions in the α-globin gene cluster cause most of the mutations, the molecular defects causing β-thalassemia are usually point mutations involving one more nucleotides within the β gene (Thein 2011).

β-thalassemia major is a severe, transfusion-dependent, inherited anaemia. Excess α-chains are unstable and precipitate in the red cell precursors forming inclusion bodies. If untreated, 80% of children with β-thalassemia major die within the first 5 years (Wild and Bain 2006). Heterozygotes for β-thalassaemia have either a normal hemoglobin with microcytosis or a mild microcytic hypochromic anemia, while HbA_2 is often elevated and Hb F is sometimes also higher than normal.

Alpha-thalassemia

The severity of thalassemias depends on the number of affected α globin alleles: the greater the number, the more severe the manifestations of the disease.

The genetic information for the α-subunit is coded in four copies (two on each of the chromosomes 16), therefore four syndromes of α-thalassemia are possible:

(1) α^+-thalassemia trait occurs when one of the two globin genes on a single chromosome fails to function (− α /α α); **(2) α^0-thalassemia** trait occurs when two genes on a single chromosome fail to function (− − /α α); **(3) Hb H disease** is when three genes are affected(− − /− α); and **(4) Hb Bart's hydrops fetalis**, where all four are defective (− − /− −). The clinical severity depends on the number of α-globin genes affected. Patients with loss of either one or two α-globin genes on the same chromosome, or one on each of the two chromosomes 16 are usually asymptomatic (Leung *et al.* 2008).

The excess β chains form tetramers (called Hemoglobin H or HbH of 4 beta chains) which are unstable and have abnormal oxygen dissociation curves. The excess γ chains form tetramers which are poor carriers of O_2. Their affinity for O_2 is too high and prevent it from dissociating in the organs, where oxygen is needed. Homozygote α^0-thalassemias, where there are no α-globins at all (called Hb Barts), often cause still birth.

Hb H disease patients rarely require transfusion. α^0-thalassemia trait is characterized by microcytic, hypochromic indices and the hemoglobin level is usually normal or slightly reduced. α^+-thalassemia trait can be completely silent. Both α^+- and α^0-thalassemia trait are more difficult to diagnose than β thalassemia trait since there is no characteristic elevation in HbA_2, and the Hb H bodies are often not observed. Definitive diagnosis of the α-thalassaemia trait can be achieved with the use of DNA techniques or globin chain biosynthesis studies (Wild and Bain 2006).

There are also a group of β-thalassemia-like disorders, referred to as δβ thalassemias and hereditary persistence of fetal hemoglobin (HPFH). These are caused by mutations that alter the switch from fetal to adult hemoglobin, and can be distinguished from β thalassemias by an increase in HbF levels (Thein 2011).

Hb variants detected and identified to date are listed in the online accessible database called the Globin gene server (http://globin.cse.psu.edu/globin/hbvar/) where information and details regarding the different variants can be found (Hardison *et al.* 2002, HbVar). Data regarding the circumstances of their discovery, experimental information about currently known Hb variants and thalassemias are provided. Results observed by different analytical techniques are included, which can help to identify certain variants, or confirm their identification. This database is constantly being updated as new Hb variants are discovered.

1.3.2 Incidence of hemoglobinopathies

The disorders of hemoglobin are the most common inherited monogenic diseases worldwide. There are many studies and surveys to estimate the incidence and possible further distribution of these disorders, emphasising the fact that this is now a worldwide health problem, which needs attention, and organised screening for prevention or possible genetic counselling.

It has been suggested that between 300,000 and 400,000 babies are born with a serious hemoglobin disorder each year and up to 90% of these births occur in low- or middle-income countries. Based on estimated annual births the most common disorder is homozygous HbS (over 200,000), followed by HbS in coinheritance with HbC (more than 50,000), with β-thalassemia as the next common (22,000), followed by HbE-β-thalassemia, and HbS heterozygous with β-thalassemia. HbH disease and Hb Bart's hydrops incidences are in the low thousands. The distribution is extremely heterogeneous within different countries (Williams and Weatherall 2012).

Sickle cell anemia is mainly present in sub-Saharan Africa and in small areas in the Mediterranean region, the Middle East, and the Indian subcontinent (Williams and Weatherall 2012).

Hemoglobin SC disease appears more frequently in parts of west and north Africa whereas HbS β-thalassemia occurs in restricted parts of sub-Saharan Africa and infrequently all over the Middle East and Indian subcontinent. HbS at some places may rise up to 20%– 25% of a particular population.

Hemoglobin E is also a common structural hemoglobin variant which occurs widely across East India, Bangladesh, Myanmar, east and Southeast Asia, the northern parts of Thailand and Cambodia. HbE is restricted to populations which have been endemic for malaria.

The mild forms of α-thalassemia are widespread from sub-Saharan Africa through the Mediterranean regions and Middle East to the Indian subcontinent and the whole of east and Southeast Asia. In these areas the frequency is around 10%–25%. In a few smaller populations in north India and Papua New Guinea α-thalassemia syndromes can be found in up to 80% of the population. The more severe forms of α-thalassemia have a much more limited incidence with high frequencies only in Southeast Asia and in some of the Mediterranean islands.

The geographic appearance of HbC is more limited than that of HbS, being common in West and North-West Africa (Williams and Weatherall 2012).

Hemoglobinopathies occur at high frequencies in the tropical regions. They have been transported to most countries of the world by population migration. The sickle cell gene is exceptionally common in the Caribbean Islands and in areas of North America, but is the most common in the majority of countries. All the common forms of thalassemia have appeared everywhere in the world. Their frequency is often explained by natural selection and the protection of heterozygotes against

severe malaria. Hemoglobin AS, CC, and AC genotypes and homozygous and heterozygous α-thalassemia give important protection from severe malaria syndromes (Taylor *et al.* 2012).

Williams and Weatherall have suggested that the high frequency of consanguineous marriages in some of the high incidence countries has an important effect on the increasing prevalence of these inherited disorder.

Approximately 5% of the world's population and over 7% of pregnant women are carriers of some kind of hemoglobinopathy, with Hb S trait accounting for 40% of all carriers. Over 300,000 babies are born every year with a major hemoglobinopathy in the world. Approximately 85% of births affected with SCD occur in Africa (Aygun and Odame 2012). A study on the global epidemiology of sickle cell estimated an increase in the global number of newborns affected by HbS in 2010 , accounting only for the homozygote SS neonates more than 300,000 (Piel *et al.* 2012). Based on the estimations by Vichinsky *et al* about 900,000 births are expected with some form of clinically significant thalassemia disorders in the next 20 years. Hb E–β-thalassemia and Hb H disease are believed to be responsible for much of this increase. Hb E–β-thalassemia is one of the most frequent hemoglobinopathies. The incidence of Hb E is nearing 60% of the population in some parts of Southeast Asia, and it is also rapidly growing in North America. Some of the α-thalassemia syndromes are accepted to be more severe and common than originally predicted. Hb H, Hb H–constant spring (CS), and homozygous α-thalassemia affect a million people worldwide. Homozygous α^0-thalassemia, in most cases fatal, is also being more frequently detected (Vichinsky 2005).

A detailed study of European epidemiology of Hb disorders which shows that sickle cell disorders are more common than thalassemias can be found in the overview by Modell *et al* (Modell *et al.* 2007).

The UK is also seriously affected by increase in the frequency of hemoglobin disorders. This has been well presented for thalassemias in the study by Henderson *et al.* (Henderson *et al.* 2009). In a five year period a total of 68 different β-thalassemia mutations were identified during screening for antenatal diagnosis, 59 of these mutations were found in immigrants to the UK, from all major ethnic groups. A total of 40 different α-thalassemia mutations were also characterised. Ten of these were deletional mutations including all the Southeast Asian and Mediterranean α^0-

thalassaemia mutations. The thirty non-deletion α^{+}-thalassemia mutations identified form 46% of the worldwide currently known non-deletion mutations, which is an incredibly high heterogeneity. Based on this study, the UK currently had the largest number of different β-thalassemia mutations among the countries involved in the comparison. The wide range of mutations found in ethnic groups known to live in the UK reflects the ethnic diversity of the UK population. Altogether β-thalassemia mutations have been characterised in 1712 individuals, while α-thalassemia mutations have occurred in 2500 possible carriers (Henderson *et al.* 2009).

In 2001 the results of a 10-year period study analysing 414,801 neonates in the UK were evaluated. HbS was found to be the most common variant, followed by Hb C, D and E. 250 of the samples screened were homozygotes or compound heterozygotes, and 6554 samples were heterozygotes for these common variants. From this data the calculated gene frequencies were for Hb S (0.57 %), Hb C (0.14%), Hb D-Punjab (Los Angeles) (0.07%), Hb E (0.05%). 16 babies had beta thalassemia major and 405 neonates had rare variants, of which six had not been previously described (Almeida *et al.* 2001).

The most current progress review for the NHS Sickle cell and Thalassemia Screening Programme revealed that sickle cell disease is still the most common, affecting one in every 2,000 births in England. The estimated numbers for Hb variant carriers is as high as 380,000. These are mainly healthy individuals with most of them sickle cell carriers (Hb S) with an additional number of carriers of Hb C. The highest proportion of the known sickle cell disease incidences in England is accounted for by the Black African and Black Caribbean population (NHS 2012a).

Beta thalassemia major is estimated to affect about 1 baby in every 20–30,000 births. There are approximately 300,000 healthy carriers and more than 800 people with beta thalassemia major and intermedia (NHS 2012a).

In the period of 2010-11, 358 out of the 688,314 babies screened had a positive result for a clinical significant mutation and 9830 were carriers. In the antenatal screening programme from the 723,768 pregnant women tested, 17,354 were recognized as carriers, predominantly of the Hb S gene. After screening consenting partners of women who were identified as a carrier, 940 high risk couples were identified.

Carrier rates grouped by ethnicity show that sickle cell carriers are common in the Caribbean and African populations while thalassemia carriers are common in the

Pakistani population. According to reported ratios 1 in every 7 carriers are Black African, 1 in every 8 Black Caribbean, and only 1 in every 450 newborn carriers are White. For thalassemia 1 in every 7 babies are from Pakistani or Cypriot descent, 1 in every 30 Indian, Bangladeshi & other Asian, 1 in every 100 Black African and Black Caribbean and only 1 in every 10,000 babies are from White British ethnic background. The thalassemia data is an estimation from antenatal data (NHS 2012a). The carrier rate per 1,000 of the babies screened from different ethnicities can be seen in Figure 1.16.

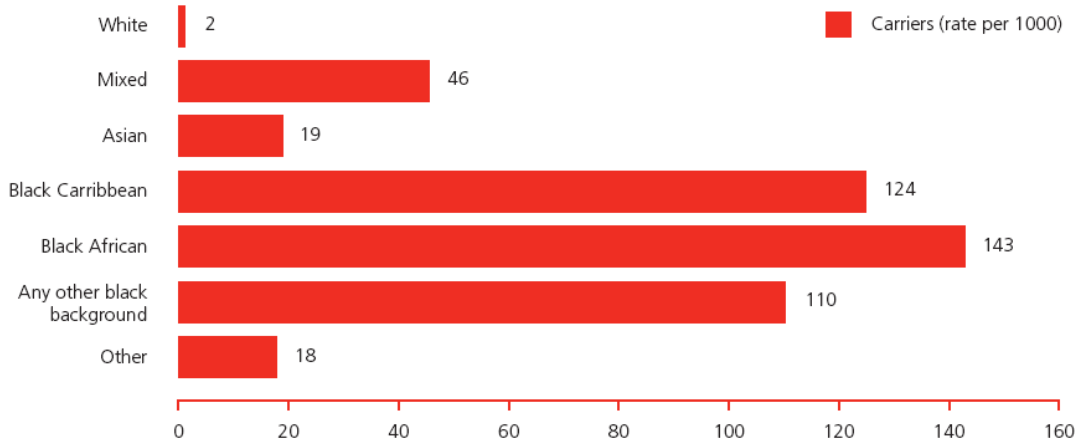


Figure 1.16. The carrier rate chart for significant Hb variants per 1,000 babies for years 2010/2011
Taken from NHS Screening Programme Review 2011/2012 (NHS 2012a)

1.3.3 Hemoglobinopathy screening in the clinical lab

Because of the extremely high prevalence of hemoglobinopathies and the continuous increase as a result of the population migration these clinical syndromes pose a significant burden on the health system in developed and developing countries today. In many countries including the UK there is a recognised need for a good screening programme which allows for the quick and accurate detection of hemoglobinopathies. The timely identification of such genetic disorders make it possible to initiate the necessary and appropriate treatment as early as possible. In the UK the NHS Sickle cell and thalassemia screening programme has been undertaking this role in the last 12 years. The developed program links two screening

set-ups, antenatal screening for pregnant women and their partners and recommending neonates to be screened for variant hemoglobins (Ryan *et al.* 2010).

The aim of the newborn screening programme is to identify infants at risk of sickle cell disease within the neonatal period, since early reliable diagnosis is necessary to initiate early treatment and to arrange clinical follow up. The analytical methods will also detect β-thalassemia major and related conditions in the majority of cases. Newborn sickle cell screening is offered to all babies born in England and is also being implemented in the other UK countries (NHS 2012b). The screening is offered at 5-8 days of age as part of the newborn dried blood spot screening programme (Streetly *et al.* 2009).

The main aim of newborn screening is to identify affected babies and providing treatment to prevent as many deaths as possible from sickle cell disease (NHS 2012a).

The objective of antenatal screening is to offer sickle cell and thalassemia screening to all pregnant women (and their partners) early in the pregnancy – ideally by 8–10 weeks of gestation – so they can request counselling about possible options and can make an informed decision (NHS 2012a). It is desired to initiate all the tests in the first 10 weeks, and finish all necessary additional testing by the 12 weeks of gestation, if termination of pregnancy is required it should happen as early as possible (Bain 2011).

It is extremely important to collect information of family origin and ethnic background, and investigating this is considered a crucial part of the screening process. The screening procedure is dependent on whether the parents are from high or low prevalence area for SCD and thalassemia. High prevalence areas are those which have fetal prevalence of SCD of 1.5 per 10,000 pregnancies or higher, and low prevalence areas are where the incidence is below 1.5. Initial laboratory screening involves the use of the Family Origin Questionnaire (FOQ), and based on ethnicity a full screen may be performed (Ryan *et al.* 2010, NHS 2012b).

The screening procedure involves measurements with several different techniques. Initial screening techniques are simple and give an indication for the presence of abnormalities but they are not evaluated diagnosis methods. Initial screening involves complete blood count measurements (CBC) and the assessment of important characteristics. The most important parameters are mean corpuscular volume (MCV - measure of the average red blood cell volume) and mean

corpuscular hemoglobin (MCH - the average mass of hemoglobin per red blood cell). These can indicate the presence of thalassemia when MCV<80 fL and MCH<27 pg (Hartwell *et al.* 2005).

Hemoglobin variants are routinely screened using electrophoresis and chromatography techniques. The most commonly used approaches are ion exchange chromatography, and isoelectric focusing (IEF). Electrophoresis has been the method of choice for identification and quantification of variant Hbs, but it is slow, labour-intensive and inaccurate at low concentration range, and it is not suitable for large batches of samples (Clarke and Higgins 2000).

The Hb variants which must be detected under the screening programme are: Sickle cell anemia (HbSS), HbS/β-thalassemia, HbS/HPHF, HbSC disease, HbS/D^{Punjab}, HbS/E, HbS/O^{Arab}. It is also considered necessary to detect β-thalassemia major, HbE/β-thalssemia, β-thalassemia intermedia and HbH disease (NHS 2012b). The employed analytical procedures must be capable of detecting all the common clinically significant Hb variants and also to quantify HbA_2 and HbF accurately as a screening tool for thalassemias (Ryan *et al.* 2010).

High performance liquid chromatography (HPLC) separates different hemoglobins based on their retention times which depend on their interactions with the stationary phase. In cation exchange HPLC (ce-HPLC) the components with a net positive charge are separated from their interaction with a negatively charged stationary phase, while the mobile phase flows through with an increasing concentration of cations (Bain 2006b). Cation-exchange HPLC has been evaluated for presumptive identification of hemoglobin variants, giving unambiguous results in 90% of cases (Chevenne *et al.* 1999, Riou *et al.* 1997). HPLC has many advantages, it can be automated and is capable of processing large batches, provides separation of clinically significant variants and allows sufficiently accurate quantitation of HbA_2 and HbF (Ryan *et al.* 2010). It is, however, unable to detect a large number of existing hemoglobin variants and cannot fully characterise any novel ones which may be present. It only provides an indication that a sample is not normal. The obtained retention times are reproducible under specific experimental conditions. Different variants, however, can elute at the same retention time, so retention time does not provide definitive identification.

A HPLC chromatogram from a normal blood sample is shown in Figure 1.17. In this method peaks correspond to hemoglobin heterodimers (αβ dimers), and elution of

HbA_2 fully separated from HbA allows the quantitation based on integration of the peak. The different retention windows show the expected ranges for different clinically significant variants. The HPLC measurement only takes 7.5 minutes and sample introduction and report processing is automated.

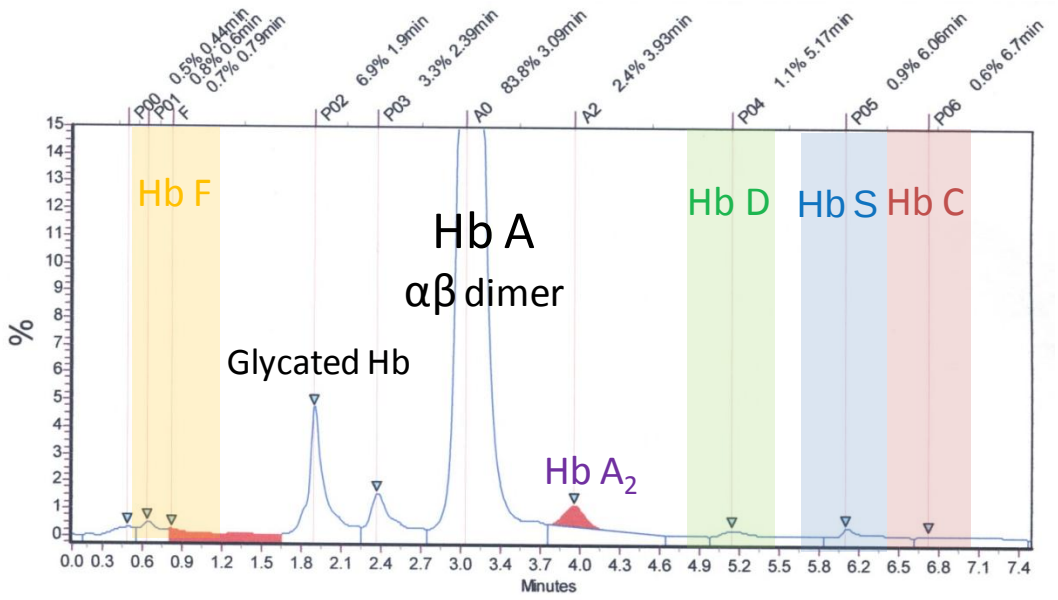


Figure 1.17. Chromatogram from TOSOH HLC-723 HbG7 analyser (Tosoh Bioscience Ltd., Redditch, UK) from normal adult.
Retention time windows for Hb F, A, A2, D, S and C are shown.

IEF separates variants based on their isoelectric point, different hemoglobins migrate via a pH gradient to the point where the pH corresponds to their pI, and net charges are zero (Hartwell *et al.* 2005). The bands are sharper than in the case of electrophoretic techniques (cellulose acetate electrophoresis), and hemoglobins which could not be separated by electrophoresis can be distinguished and identified. The precision for quantification at low concentrations is poor and this technique has not been validated for HbA_2 quantitation (Bain 2006b).

DNA analysis is required for absolute genetic characterization of mutations resulting in β-thalassemia and quantitation of the HbA_2 is crucial for the routine identification of carriers. Guidelines and recommendations for correct HbA_2 measurements can be found in publications by Stephens *et al*. Stephens points out a common problem, that many Hb variants partially or totally co-elute with HbA_2, making the quantification in certain circumstances unreliable (Stephens *et al.* 2011b). The national recommended cut-off for HbA_2 (3.5%) is the action point which has been established for the diagnosis of β-thalassemia carriers (Ryan *et al.* 2010).

Measurement of Hb F is important in the diagnosis of δβ thalassaemia, hereditary persistence of fetal haemoglobin (HPFH) and in the diagnosis and management of sickle cell disease. As HbF is expected to be present up to 1%, elevated levels of HbF are of diagnostic value. In thalassemia trait levels of up to 10% occur, in δβ thalassaemia trait levels between 5% and 20%, and in HPFH heterozygotes from 2.5% up to 30% in the deletional forms are found (Stephens *et al.* 2011a). Further investigation of elevated HbF is initiated at a measured level of 5% or higher (Ryan *et al.* 2010).

Presumptive identification should be based on a minimum of two techniques based on different principles (Ryan *et al.* 2010). The routine techniques are useful in the screening of the common clinically significant variants but they are unable to identify novel variants and so additional DNA sequencing or tandem mass spectrometry (MS/MS) is required for identification of a variant. DNA analysis for the diagnosis of every patient is not achievable due to high cost and complexity. The use of mass spectrometry has become an alternative for variant identification and has potential to become a future tool in everyday screening (Ryan *et al.* 2010).
Some lessons and pitfalls about currently used screening methods and algorithms on how to get the most appropriate results can be found in recent review by Barbara J. Bain (Bain 2011).

1.3.4 Mass spectrometry-based identification of hemoglobinopathies and future directions of screening

All molecules have a molecular mass which can be used to help to identify them. The different hemoglobin molecules, and their building blocks, the globin chains, have molecular masses which can be measured with MS. The molecular masses of the fetal (γ^A and γ^G) and adult globin chains (α, β, δ) are listed in Table 1.2. Genetic mutations resulting in amino acid substitutions in globin chains will cause alterations in these masses in most cases and therefore mass spectrometry can provide a powerful tool in the detection and identification of such mutations.

Globin chain	Average molecular mass (Da)	Number of amino acids	Hemoglobins
α	15,126.38	141	HbA; HbA_2; HbF
β	15,867.24	146	HbA
γ_1^G γ_2^A	15,995.27 16,009.30	146	HbF, Hb Portland[a]
δ	15,924.32	146	HbA_2
α^G	15288.52	141	
β^G	16029.38	146	HbA_{1c}
β^{SICKLE}	15837.25	146	HbS
$\beta^{D\text{-}Punjab}$	15866.26	146	Hb D-Punjab
β^C, β^E, $\beta^{O\text{-}Arab}$	15866.30	146	Hb C, Hb E, Hb O-Arab

Table 1.2. Masses of human globin chains (α, β, γ^A, γ^G), glycated forms of major globin chains (α^G, β^G) and variant globin chains (β^{SICKLE}, $\beta^{D\text{-}Punjab}$, β^C, β^E, $\beta^{O\text{-}Arab}$)

Mass spectrometry has been used in the identification of hemoglobinopathies for more than 30 years. Investigations have included structural-based studies, identification of hemoglobin disorders, quantitation of the minor hemoglobins such as HbA_2 and HbF and the characterisation of post translational modifications (PTMs) such as glycated hemoglobin HbA_{1c} and glutathionylated Hb as reviewed by many publications (Kleinert *et al.* 2008, Wajcman and Riou 2009, Zanella-Cleon *et al.*2009).

Structural analysis of peptides from human hemoglobin variants was performed and demonstrated first by the analysis of tryptic peptide mixtures of abnormal hemoglobins by Wada *et al.* Measurements were performed with instruments using field desorption (FD) (Wada *et al.* 1981). Fast atom bombardment (FAB) (Pucci *et al.* 1989, Wada *et al.* 1983a) and molecular (or liquid) secondary ion mass spectrometry (SIMS) have also been used (Wada *et al.* 1983b). After the advent of electrospray ionisation the ability to obtain multiply charged ion series of biological molecules was possible, which allowed the clinical diagnosis of variant hemoglobins (Shackleton *et al.* 1991). Tandem mass spectrometry of variant peptides proved to be

the method of choice for peptide sequencing, which allowed the determination of positions of amino acid substitutions (Falick *et al.* 1990).

There are two main strategies which are employed and often combined by researchers for the analysis of hemoglobin disorders. The requirement for these methods can be explained by the purpose of the analysis, which is either to undertake the identification of unknown variants or to screen for specific known variants.

Intact globin chain analysis is an obvious first choice of analysis to measure the intact mass of hemoglobin chains. Results obtained from this type of analysis may indicate abnormalities, but for a comprehensive investigation of disorders smaller scale examination is a necessary follow-up method. In an intact analysis, masses of globin chains (α, β, δ, γ), their adducts, PTMs including glycation or glutathionylation, and additional information can be obtained readily without complex sample preparation procedures. Increased complexity arises at the peptide level, requiring in-depth data analysis to interpret the results. Globin chains are broken down to peptides comprising a maximum of 10-15 amino acids. Analysing smaller fragments of the protein can provide sequence information, and may identify mutations within the abnormal globin chain sequence, which can also occasionally serve as variant specific markers.

In more details about variant identification: the most widely accepted procedure for intact globin chain analysis is predominantly based on the work of Brian Green (Green *et al.* 1990), utilising the protocol developed by Wild *et al.* (Wild *et al.* 2001). This general multi-step approach is the suggested protocol for rapid identification of hemoglobin variants first employed the ESI-MS analysis of diluted whole blood. Under denaturing conditions the mixture of separate intact globin chains and heme groups could be analysed. This step allowed the detection of the presence of any abnormal globin chains. In early research publications the variant mass could be resolved on lower resolution instruments and associated with the α and β-chain if it differed from the normal globin chain by at least 6 Da, this is not the case now when high resolution instruments are more common Based on the measured mass difference the possible amino acid substitutions based on single point mutations which could be responsible for the mass shift could be determined. A second step may involve tandem mass spectrometric analysis of the intact globin chains, but generally only when the variant tryptic peptide was not found. A third step is the previously mentioned analysis of the peptide mixture produced by

enzymatic cleavage with trypsin. Without prior separation observation of an additional mass can identify the peptide which contains the variant. Following this in a final step the variant peptide can be sequenced by tandem MS to identify and confirm the position of the substitution (Wild *et al.* 2001).
This multi-step approach requires only a small amount of whole blood (10 µl) and no extra pre-analytical steps, such as separation of red cells or chromatographic separation of globin chains. The desalting of the samples prior to analysis is advised, but not a strict requirement. 95% of the variants in over 250 samples were positively identified by this method. The whole process with sample preparation and ESI-MS analysis was proposed to require only 1 to 2 hours with minimal reagent costs.

To observe the globin chains as separate entities in the mass spectrum it was reported that the mass difference between them would need to be 6 Da or higher (Wild *et al.* 2001, Wild *et al.* 2004). This limitation has been overcome since then with the developmemt of high mass resolution instruments and with accurate mass measurements. The majority of the clinically significant variants differ by only 1 Da. Rai *et al* has shown that with accurate mass measurement, using the alpha-chain as internal mass calibrant the presence of variants differing by 1 Da can be detected based on mass shifts caused by the combination of the different mass peaks (Rai *et al.* 2003). The achievable accuracy of the β-globin chain measurement is ±0.1 Da, and if a -1 Da β-chain variant is present at 20%, the mass shift caused would be expected to be -0.2 Da. An abnormal variant differing by 1 Da from normal should be detectable if it is present at >10%. The ESI-MS analysis of intact globin chains was proposed to be useful in the newborn screening for the detection of sickle cell disease (Wild *et al.* 2004). Although there is a risk for false positives in cases where a variant chain had the same mass as β^{SICKLE}.

A recent review by Zanella-Cleon *et al.* summarised the different mass spectrometric methods which have been used for phenotype determination of hemoglobinopathies (Zanella-Cleon *et al.* 2009). Today 150 variants are listed on the Globin gene server for which mass spectrometry provided experimental structural characterisation (HbVar, http://globin.bx.psu.edu/hbvar/menu.html).
Taking only single point mutation leading to single amino acid substitutions into account it was shown that only 43.2% of all possible α- and β-chain variants had

been detected up to 2007 and 69% up today (Kleinert *et al.* 2008, HbVar). Kleinert *et al.* predicted that approximately 90% of the possible but up to then undetected variants (885 of 962) could potentially be detected by mass spectrometry differing by 6 Da or more from corresponding wild-type chains. The detection of the resulted mass change from mutations like Lys↔Gln or Leu↔Ile is not possible by ESI-MS analysis of denatured globin chains. For Lys↔Gln, the mass difference is only 0.05Da, which can be identified on peptide or fragmentation level with high resolution instruments. The Gln→Lys mutation can be detected on peptide level, as the introduction of a Lys creates an additional cleavage site and produces two new peptides (Wild *et al.* 2001).

Measurement of fusion globin chains like Hb Lepore by mass spectrometry has also been shown to be possible. Hb Lepore is a relatively low-expressed variant accounting for 5 to 15 % of the total hemoglobin and so it can be challenging to detect. De Caterina *et al.* characterized the different Hb Lepore variants with multiple mass spectrometry-based methods analysing the separated hybrid globins and their tryptically digested peptide mixture (De Caterina *et al.* 1992). Rai *et al.* used accurate mass measurement and tandem mass spectrometry of intact globin chains to identify the most common Hb Lepore-Boston-Washington variant directly from diluted whole blood. This differs by only 2 Da from the normal β-chain variant (Rai *et al.* 2004). They also showed that ESI-MS could be a valuable tool for the identification of another fusion variant Hb Kenya (Rai *et al.* 2002).

Advanced mass spectrometry techniques have proved to be useful in special circumstances. Williams *et al.* showed that the coupling of ion mobility (IM) to MS can help in identifying variants present at low abundance (Williams *et al.* 2008). A doubly-charged tryptic peptide from a low abundance variant (4%) was present at the same *m*/*z* value as a singly- and a doubly-charged interfering ion, and IM was used to separate these ions. The approach allowed complete sequencing, and thus the characterisation of the variant was possible.

In another challenging case a variation of ECD (hot ECD) on an FT-ICR was used to distinguish between Leu and Ile to identify a novel variant (Williams *et al.* 2009).

Post translational modifications of hemoglobin result in a certain mass shift, which enables them to be studied using mass spectrometry. They often carry diagnostic

mass values. The measurement of glycosylated hemoglobin (HbA_{1c}) level has been widely used as an indicator for long-term diabetic control for about 30 years (Gabbay *et al.* 1979). There have been attempts in the clinical diagnosis field to standardise the analysis of HbA_{1c} as a primary tool for monitoring glycaemic control (John and English 2012).

An ESI-MS procedure has also been developed for the determination of the amount of glycated hemoglobin in blood. The relative glycation of the alpha- and beta-globin chains was determined (Roberts *et al.* 1997) and the MS results compared with established affinity chromatographic results. A good overall agreement was observed. In 2001 Roberts *et al.* carried out a detailed evaluation of this method over a 4 month-period to establish the potential of using ESI-MS as a routine reference method for quantification of glycohemoglobin (Roberts *et al.* 2001). They concluded that ESI-MS provided a precise measurement of HbA_{1c} (glycation of the β chain). The method has been shown to be robust and could be used as a procedure to substantiate HbA_{1c} measurement and/or calibration.

Nakanishi *et al.* showed that the presence of a variant can effect routine measurement of the glycated Hb, HbA_{1c}. This can lead to an erroneous value obtained by HPLC (Nakanishi *et al.* 2002). Nakanishi *et al.* analyzed enzyme-digested peptides and intact globin chains and suggested that MS may offer an opportunity to correct these errors. With accurate measurement of the glycohemoglobin, MS has the potential to become a routine technique used for quantitation of glycated Hb. Kobold *et al.* developed an approved international reference method for quantification (Kobold *et al.* 1997), based on monitoring specific N-terminal peptides of the beta-chain. Investigations were made as to whether the glycated and non-glycated beta-N terminal peptides could be analyzed instead of measuring the total molecule. They studied the cleavage of the Hb molecule with endoproteinase Glu-C, which cleaves the N-terminal portion of the beta-chain between two glutamic acid residues. The method specifically measures the ratio of the N-terminal hexapeptides of HbA_{1c} and HbA_0. The authors found that the HbA_{1c} values were lower than those determined with current routine methods, and suggested that the ranges normally quoted may have to be revised and adjusted.

There is a growing need for automated screening methods for the detection of clinically significant hemoglobin variants. Screening methods and automated variant

detection approaches need more definitive identification than determination of protein mass shifts on the intact globin chain level, which can lack specificity since a certain mass shift could result from the same amino acid substitutions but at different positions along the globin chain sequence. A more reliable confirmation can be achieved on the peptide level.

An alternative to intact globin chain combined with tryptically digested peptide mixture analysis is a targeted method to monitor the presence and level of pre-specified peptides with a known sequence belonging to specific variants or minor globin chains. This targeted method has the potential to be a fully automated mass spectrometry-based method that may be incorporated in clinical screening, and is of particular interest. Enzymatic cleavage of normal α-globin chain results in 14 tryptic peptides (αT1-T14), whilst the β-chain is broken down to 15 peptides (βT1-T15).
The major disadvantage of this type of experiment is that only known variants and their corresponding mutant peptide of interest can be screened for. The advantage is, however, that for the presence of those variants which are screened, a definitive answer from a single measurement can be obtained.
Daniel *et al.* published an approach that demonstrates clear evidence of the rapid and specific detection of significant hemoglobinopathies by MS (Daniel *et al.* 2005). They perfomed ESI-MS/MS measurements, and worked in multiple reaction monitoring mode on a ESI-triple quadrupole instrument. MRM transitions for normal and variant β-chain peptides such as wild-type T1, HbS T1, HbC T1, wild type T3, HbE T3, wild type T13, HbO T13 and HbD T13 have been developed. A total of 200 blood samples were analysed, comprising of normal hemoglobin samples as well as clinically significant variants. All hemoglobin variants were correctly identified by MS/MS, in agreement with the existing clinical methods. This approach alone provided a simple and rapid solution to high-throughput population screening of the clinically significant hemoglobin variants, detected both heterozygotes and homozygotes and has the potential to be used in neonatal and antenatal screening. In 2007, the method was extended and the authors integrated the quantitation of HbA_2 (Daniel *et al.* 2007). The δ:β ratios were measured as potential surrogate markers of HbA_2, which is used as a biomarker for β-thalassemia. They included the δT2, βT2, δT3, βT3, and δT14, βT13 peptides in their previous method. Samples were analyzed with normal HbA_2 and with elevated HbA_2 levels. This

method successfully provided reproducible results for differentiating between samples with normal HbA_2 concentration and those with increased concentrations, in excellent agreement with HPLC HbA_2 results.

The developed MRM method outlined by Daniel *et al.* for detecting a known protein variant is based on the following criteria:

1. The protein needs to be digested to produce a defined series of peptides;
2. The peptides need to be ionised and an ionised species of known mass-to-charge ratio which is indicative of the protein variant is selected by the first quadrupole in the mass spectrometer.
3. The ionised species selected is subjected to collision induced dissociation (in the second quadrupole) and one or more of the derived ionised species of known mass-to-charge ratio measured in the third quadrupole, confirming the presence of the protein variant in the sample.

The peptide containing the variant amino acid is assumed to be indicative of the variant protein. If the variant protein differs from the wild-type protein in a way that the cleavage site for the used protease changes, one or more new peptides will be produced, and those new peptides will be representative of the variant protein. Since, in the case of a known variant, the specific ionised peptide to be selected is known, it is possible to predict the mass-to-charge ratio of the ionised species to be selected by mass spectrometry. The ionised species selected depends on a number of variables, including fragmentation efficiency and the ability to distinguish the selected ion from those which derived from the non-variant protein. All of these parameters need to be considered in order to obtain an optimum level of detection.

The energy introduced for collision induced dissociation can be varied, thereby allowing the degree of dissociation of the selected ionised species to be controlled. Low energy fragmentation is preferred since it does not cause substantial dissociation of the selected ion. This information can provide further confirmation of the presence of the protein variant.

Multiple reaction monitoring methods allow the accurate and specific detection of a protein variant, and so can detect the presence of clinically important hemoglobin variants such as Hb S, Hb C, Hb D^{Punjab}, Hb O^{Arab}, Hb Lepore and HbE, delta/beta thalassemia, hereditary persistence of fetal hemoglobin trait (HPFH) and alpha zero thalassemia trait.

An example of an MRM approach for the confirmation of Sickle cell mutation is detailed in the following. The mutation occurs at position 6 of the beta chain and it affects the first tryptic peptide βT1; Daniel *et al.* worked on suitable MRM transitions for wild-type and sickle cell beta T1 peptides. They chose the doubly charged peptide ions in both cases as precursor ions: for wild-type βT1 peptide the $[M+2H]^{2+}$ is *m/z* 476.8 and for the HbS βT1 is *m/z* 461.8. The fragment ions which are monitored in the third quadrupole are the y4 fragment ions. The y4 fragment ion of the control T1 peptide is *m/z* 502.2, and of the sickle T1 peptide is *m/z* 472.3 . Peaks are detected when the MRM transition of $[M+2H]^{2+}$ = *m/z* 461.8 ion to the singly charged y4 ion at *m/z* 473.2 occurs; this is used to demonstrate that the blood contains the variant sickle cell peptide. A schematic diagram of the MRM approach for both peptides can be seen in Figure 1.18.

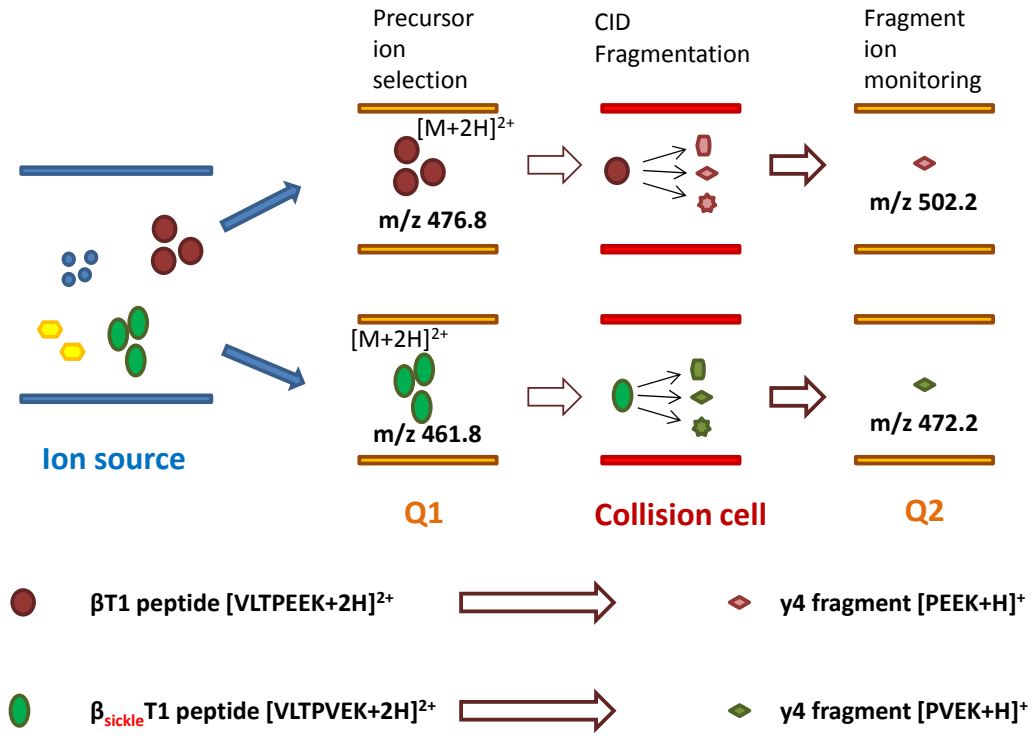


Figure 1.18. Schematic of the MRM transitions used for detection of sickle cell mutation.
Adapted from (Zanella-Cleon *et al.* 2009)

For clinically significant variants the mutation occurs, in most cases, in the T1, T3 and T13 tryptic peptides. For the determination of delta/beta ratios, we can use those peptides, in which there is a difference in the amino acid sequence between the two

chains (T2, T3, T5, T10, T12, T13). MRM can be applied in all cases when a unique transition exists which is characteristic for the variant peptide.
Another MRM-based method for clinically significant variant detection has been developed and proposed for newborn screening by Boemer *et al.* (Boemer *et al.* 2008, Boemer *et al.* 2009). During a 3-year validating period with the analysis of about 44,000 samples they proved this method could provide an efficient alternative approach in neonatal screening (Boemer *et al.* 2011).

Multiple reaction monitoring has also been used for the quantification of glycohemoglobin. Unique MRM transitions have been used to characterise the glycated and non-glycated hemoglobins present in the sample (Willekens *et al.* 2000). β-N-terminal hexapeptides of HbA_{1c} and HbA_0 were obtained by enzymatic cleavage of the intact hemoglobin molecule with endoproteinase Glu-C and then doubly charged precursor and singly charged product ion pairs of *m/z* 347/237 for non-glycated and *m/z* 428/245 and for glycated hexapeptides. Samples were analysed using the LC-ESI-MS (SIR) reference method and by the MRM approach. The results of the two methods were found to be comparable.

A single-step approach for the Hb variants has been developed by Basilico *et al.* (Basilico *et al.* 2007). Their method is a proteomic-like approach, offering data dependent MS/MS analysis of tryptically digested peptide mixture of diluted whole blood after separation with HPLC. The experimental MS data is compared to theoretical mass spectra generated form an in-house built database. The database contains variant tryptic peptides of known human variants based on the HbVar database (Hardison *et al.* 2002, HbVar).

In the review by Zanella-Cleaon *et al.* the potential to use top-down protein identification with ETD fragmentation techniques was also mentioned (Zanella-Cleon *et al.* 2009). Recent publications have shown that it can indeed provide a future direction for Hb variant identification. The combination of direct surface sampling of dried blood spots with tandem mass spectrometry has been used for the identification of unknown variants on an Orbitrap instrument (Edwards *et al.* 2011, Edwards *et al.* 2012). This method had the advantage of directly analysing samples from newborn blood spots without prior sample preparation. The application of an

LC coupled top-down approach on a lower cost 3D ion trap also has been shown to give unambiguous identification of clinically significant variants and shows potential for future clinical laboratory screening applications (Graca *et al.* 2012).

The application of MS in combination with other methods such as electrophoresis, IEF, and ce-HPLC has been shown to be very powerful, because a wider range of variants can be detected and identified (Troxler *et al.* 2012).
Although MS has been shown to be capable for high-throughput population screening of hemoglobinopathies, and it has been used for the analysis of larger clinical sample-sets, it is not yet a validated primary clinical laboratory diagnosis tool.

1.4 Project aims

The primary aim of the project is to show how mass spectrometry can provide a single analytical platform, with the potential to replace the combination of techniques currently used in antenatal screening, while saving time and money.
The objective is to improve existing mass-spectrometry-based methods, develop novel approaches and make them as automated and straightforward as possible to assess their suitability for high-throughput screening of clinically significant hemoglobin variants and offering the option for occasional identification of novel or silent variants.
The combination of developed approaches, or preferably a single method will be able to provide a simple and rapid solution to high-throughput population screening of structural hemoglobin variants, thalassemias and glycated adducts.
Primary aim is to develop a mass-spectrometry based approach which is: fast, sensitive; accurate, financially competitive with existing approaches; more comprehensive than current methods and capable of detecting all known and previously unidentified single-point mutation variants.
This approach is proposed to be easier-to-use and require less interpretation meaning that less expert staff are required. If all these expectations are met, the developed approach could be routinely applied in the everyday population screening.

- To develop mass spectrometry-based approaches based on previously published MRM and full scan mode methods, develop MRM transitions for novel peptide pairs, which has not been applied for screening purposes before.
- Assess the suitability of different methods, compare their results to evaluate which method is the most appropriate for screening purposes and needs further development
- Test the developed methods on a larger set of clinical samples, assess robustness and agreement with current clinical screening methods. Methods will be tested on a larger set of antenatal samples, which has not been presented in the literature before.

- Develop novel tandem mass spectrometry-based method utilising ETD with alternative fundamental aspects to previously published methods, which apply chromatographic separation and SIR type detection, while monitoring only specific reporter fragment ions for all clinically significant variants could provide a less time-consuming screening method.
- Test the novel ETD method on clinical samples and assess agreement for clinically significant variant detection with current clinical method

1.5 Research and Conference Papers

The work has been presented at the following conferences:

Krisztina Radi, Charlotte A. Scarff, James H. Scrivens. A mass spectrometry-based approach to hemoglobinopathy detection. *BMSS Annual Meeting 2010, Cardiff, UK*

Krisztina Radi, Charlotte A. Scarff, Susan E. Slade, Nisha A. Patel, James H. Scrivens. The development of rapid diagnostic approaches for the characterisation of hemoglobin disorders. *Proc. 59th ASMS Conf. on Mass Spectrometry and Allied topics, 2011, Denver, USA*

Krisztina Radi, Baharak Vafadar-Isfahani, Jane Newbold, Nicholas Jackson and James H. Scrivens. Mass spectrometry-based approaches to the diagnosis of hemoglobinopathies. *Warwick Mass Spectrometry 80/60 Conference, 2012, Coventry, UK*

Krisztina Radi, Baharak Vafadar-Isfahani, Jane Newbold, Nicholas Jackson and James H. Scrivens. Automated quantitative mass spectrometry-based approaches for the diagnosis of beta thalassemia in a clinical trial setup. *MSACL 2013 - Mass Spectrometry Applications to the Clinical Laboratory, San Diego, CA, United States*

James Scrivens, **Krisztina Radi**, Baharak Vafadar-Isfahani, Julia Smith, Jane Newbold, Nicholas Jackson. Development of a mass spectrometry based approach for the diagnosis of hemoglobinopathies. *Proc. 61st ASMS Conf. on Mass Spectrometry and Allied topics, 2013, Minneapolis, USA*

CHAPTER 2

METHOD DEVELOPMENT FOR MASS SPECTROMETRY-BASED DIAGNOSIS OF HEMOGLOBINOPATHIES

2.1 Introduction

As presented in previous chapter mass spectrometry can provide definitive information regarding the masses of the globin chains, location of mutation points and confirm the presence of known variants. Current methods based on UV or electrophoretic responses are based on pre-identified information, such as retention time for UV chromatograms and can only give presumptive diagnosis (Bain 2011).

Since MS is not a routinely applied screening technique yet, methods need to be further developed and evaluated.

The protocol of variant identification using ESI-MS consists of as detailed by Wild *et al.* needs to be optimized to be applied in an automated manner tobe approved as a population screening method(Wild *et al.* 2001). This is necessary approach for identification of unknown variants, therefore it is advised to include it in the developmental work. For screening purposes of the clinically significant variants the absolute confirmation is also essential and can be performed with further developed targeted peptide monitoring achieved by MRM approaches.

Quantitation of the different minor globin chains is important to monitor healthy levels of proteins in blood. The accurate quantification of hemoglobin A_2 is essential to diagnose β-thalassemia trait. Precision is crucial since the difference in HbA_2 concentration between normal and elevated is narrow, and the upper limit for normal has been set at 3.5% (Higgins *et al.* 2009).

The quantitation of delta globin chain from the mass spectrum of the denatured protein globin chains is based on the assumption that the relative intensities obtained after deconvoluting multiple charged envelopes to true mass scale are proportional to the concentrations of the proteins (chains) (Roberts *et al.* 1997).

Based on Roberts *et al.* work, it is straightforward to investigate whether the measurement of intact proteins in diluted whole blood using ESI-MS could allow for the easy and rapid quantitation of minor hemoglobin A_2, through quantifying the intensity of the delta globin chain relative to that of the beta globin chain in the mass spectrum. The analysis of diluted whole blood by direct infusion may provide an attractive solution for population screening, without the additional cost of utilising chromatographic separation or enzyamtic cleavage steps.

Aims and objectives

Mass spectrometry is a powerful technique, which can be used for the identification of Hb variants, screening for thalassemias and also for quantification of HbA_2 and glycated hemoglobin. These different applications have, however, not been integrated on a single platform. While an MRM approach can identify known variants, it is not possible to screen for all potential 1000 single point variants in a reasonable-time scale and the method cannot identify novel variants, therefore additional measurements are necessary.

MS has not yet become an established clinical screening method in this field due to the required background knowledge for the interpretation of results in more complex cases. Processing of MS data requires software-based solutions, and/or the usage of an in-house-built database (Basilico *et al.* 2007). These solutions can make the analysis easier, more automated and provide integration of prior knowledge. The aim is to develop MS methods which can be easily automated, and applied in high-throughput screening with the potential of automated data processing and reporting.

The work presented in this Chapter aims to:

- Develop mass spectrometry based approaches for the diagnosis of hemoglobin disorders based on previously published methodologies for available mass spectrometry instrumentation.
- Implement different methods on the same platform; these methods should be fast, sensitive, accurate, financially competitive with existing approaches, more comprehensive than current methods and capable of detecting most known and previously unidentified single-point mutation variants.
- Test methods on reference standards or a set of clinical samples, where available.
- Optimise methods and sample preparation for high-throughput screening applications in preparation for a clinical trial, which will involve the analysis of hundreds of samples per calendar week.
- To assess the suitability of two different triple quadrupole mass spectrometers (Waters Acquity TQD and Waters Xevo-TQ) and corresponding data processing

software (QuanLynx^{TM} and TargetLynx^{TM}) to obtain the highest level of automation possible.

- To assess the need for further optimisation of the different methods

2.2 Materials and methods

Origin of analysed samples

For method development, commercial lyophilized HbA_0, HbS and HbA_2 samples were used (Sigma Aldrich). For assessing glycated hemoglobin measurements a ClinTest reference standard with known concentration of HbA_{1c} (11011, RECIPE Chemicals + Instruments GmbH, Munich, Germany), and for delta chain level determination the WHO 1st international reference reagent for HbA_2 (89/666, National Institute of Biological Standards and Control) were used. For method testing, final optimisation and the identification of unknown variants; patient blood samples provided by the Haematology Department at University Hospitals Coventry and Warwickshire NHS Trust were measured (in collaboration with Dr Nicholas Jackson).

The samples analysed are summarised in Table 2.1.

Sample origin	Number of samples analysed	Sample type	Purpose of analysis, applied MS-approach
Standard – Sigma Aldrich	3	HbA_0, HbA_2, HbS lyophilised powder	Method development
Standard - NIBSC	1	Lyophilised HbA_2 standard with HbA_2 level 5.3%	Control
Standard – ClinTest	1	HbA_{1c} standard 9.3%	Control
Clinical samples - Birmingham	60	Clinical samples from individual patients with no previous information	Testing MRM 1 method on clinical samples
Clinical Samples -UHCW	94	Clinical samples from individual patients, information about disease conditions and HbA_2 levels by ce-HPLC	Testing Intact and MRM 2 method on clinical samples
Clinical Samples UHCW	3	Samples from patients with unidentified Hb variants	Identification of rare variants, testing the used protocol

Table 2.1. Summary of samples used for different applications during method development

A batch of samples were also provided by Dr. Patel (University of Birmingham, Centre for Cardiovascular Sciences), which was used in the early testing with the first developed MRM-method.

2.2.1 Sample preparation

Two main types of sample preparation were necessary for the analysis of intact chains and tryptic peptides. The sample preparation steps are described here for different applications used during the method development procedure.

2.2.1.1 Intact globin chain analysis

The healthy range of hemoglobin concentration in blood of adult women is 115 to 165 mg/mL and for men 130 to 175 mg/mL. For 3 months old babies only 95-135 mg/mL and for children up to 12 years it is lower at 105 to 145 mg/mL (NHS 2013). Based on these values, standard stock solutions were prepared at 3 mg/mL in deionised water to obtain a hemoglobin concentration similar to that of a 50-fold diluted whole blood sample.
In the case of anonymised patient blood samples, stock solutions were prepared as follows: 20 µL of blood was diluted in 980 µL deionised water, and stored at -20ºC.
For intact globin chain analysis, 20 µL of stock solution was diluted 20-fold in 380 µL 50 % ACN containing 0.2 % formic acid. For the purpose of high-throughput analysis using the Nanomate, 280µL of these dilutions were transferred into 96 well plates with a 350µL well volume. The Triversa Nanomate is a chip-based electrospray ionization technology from Advion (Advion Biosciences Ltd., Ithaca, NY, USA).

Prior to intact mass analysis, samples were desalted either by agitation for 5-10 minutes with approximately 3-5 mg of AG 50W-X8 cation-exchange beads (Bio-Rad Laboratories) or with use of Amicon Ultra-0.5mL Centrifugal Filters with a membrane cutoff 10kDa (Merck Millipore).

2.2.1.2 Peptide analysis

Variant identification on the peptide level:

Tryptic digestion was performed according to the protocol published by (Wild *et al.* 2001) 200 µL of stock solution was mixed with 20 µL of 50% ACN containing 0.5% formic acid, 6 µL of 1 M ammonium bicarbonate solution and 5 µL of 5mg/ml trypsin (from bovine pancrease TPCK Treated, lyophilized powder, Sigma Aldrich) solution, incubated for 1 hour or in some cases overnight at 37°C.

Targeted peptide analysis by MRM method 1

For initial method development commercially purchased lyophilised HbA_0, HbS and HbA_2 were used. Tryptic digested peptide mixtures were prepared following the previously described protocol. The digests were diluted 10 or 20-fold prior to direct infusion into the standard ESI-source of the Acquity TQD system (Waters Corporation, Manchester, UK).
In order to test the method on clinical samples a batch of blood samples by Dr. Patel (University of Birmingham, Centre for Cardiovascular Sciences) was provided and prepared for analysis by diluting 1 µL of whole blood in 50 µL of water and 10 µL of 50 % ACN containing 0.5 % formic acid within a 96-well plate. Tryptic digestion was performed to the concentrations outlined in the protocol published by Wild *et al.* The pre-diluted sample solutions were mixed with 3 µL of 1 M ammonium bicarbonate solution and 2.5 µL of 5mg/ml trypsin (from bovine pancrease TPCK Treated, lyophilized powder, Sigma Aldrich) solution, incubated overnight at 37°C.
Prior to analysis by the UPLC-MRM method the digested blood samples were further diluted in 266 µL deionised water.

Targeted peptide analysis by MRM method 2

Initial method development and optimisation were performed by analysing standard samples with known amounts of tryptically digested HbA_0 and HbA_2 samples. The standard samples were individually digested and then mixed to produce a series of mixtures containing 1, 2, 3, 4, 5, 6, 8, 10, 12 and 15% of HbA_2.

Clinical blood samples provided by UHCW were prepared for analysis by diluting 200 µL of stock solutions with 40 µL of 50 % ACN containing 0.5 % formic acid. Tryptic digestion was performed according to the protocol published by Wild *et al*. (Wild *et al.* 2001). The pre-diluted sample solutions were mixed with 12 µL of 1 M ammonium bicarbonate solution and 10 µL of 5mg/ml trypsin solution, incubated overnight at 37°C.

After incubation, digested samples were centrifuged at 10000rpm (9503 g) for 3 minutes to remove solid particles and remaining precipitated proteins. Prior to analysis samples were further diluted 7-fold in 10% ACN with 0.5% formic acid. 5 µL of these diluted peptide mixtures were loaded on the LC column for separation.

The accuracy of delta chain quantification was assessed by measuring digested international HbA_2 standard using the optimised UPLC-MRM method.

2.2.2 Mass spectrometry and data processing methods

Sample Introduction

Two types of sample introduction were employed in these studies; directly introducing the sample into the source of mass spectrometry or after a chromatographic separation step. Direct injection can be implemented via a syringe, using the fluidic system of the instrument or using an automated sample introduction system TriVersa NanoMate (Advion Biosciences Ltd., Ithaca, NY, USA). Chromatographic separation and sample introduction were achieved via the Acquity UPLC system.

Intact globin chain analysis and multistep hemoglobin variant identification

The use of a triple quadrupole system for the analysis of intact globin chains is described and optimized for its application in a high-throughput manner to analyse a significant number of samples. For high-throughput applications, sample introduction needs to be performed in an automated manner to allow the measurements of multiple samples to be carried out 24 hours a day.

A Xevo-TQ (Waters Corporation, Manchester, UK) equipped with a standard flow ESI source and controlled by MassLynx^{TM} (version 4.1) software (Waters Corporation, Manchester, UK) was used for method development and testing on standard samples. Samples were introduced into the ESI source from a 250 µL syringe at a flow rate of 5 µL/min using a syringe. In other cases the fluidics system of the Xevo TQ instrument was used, and 1 ml of sample prepared in a 1.5 ml vial which was then introduced to the ESI source again at a flow rate of 5 µL/min.

For high-throughput method development the Xevo-TQ (Waters Corporation, Milford, Manchester, UK) equipped with a TriVersa NanoMate (Advion Biosciences Ltd., Ithaca, NY, USA) automated sample introduction system was used to acquire data on clinical samples.

The instrument was operated in ESI positive mode with a capillary voltage of 3 kV, cone voltage 40 V and source temperature 150ºC for all experiments.

The mass analyser was calibrated using the alpha-chain multiple charged peak series of a control blood sample.

Data were acquired through the tune page of the MassLynx^{TM} software for 3 minutes in multi-channel acquisition (MCA) mode at a scan speed of 1 spectrum/sec.

The measured ESI-MS spectrum was background subtracted with a 25-order polynomial to allow 5% of the data to fall below the new baseline, this subtracted spectrum was then deconvoluted onto a true mass scale by use of a maximum entropy based MaxEnt application program, which is available within the MassLynx software. The mass output parameters were set for mass range 14800-16800 Da at a resolution of 0.15 Da/channel. A Gaussian model was used: peak width 0.8 Da, 40% minimum intensity ratios were set and the algorithm was allowed to iterate until convergence. The deconvoluted spectrum obtained from MaxEnt was smoothed and centred, and a mass correction was applied so that the alpha-chain mass was set to the expected 15126.38 Da.

Samples were assessed for the presence of single-point mutation hemoglobin variants, δ-chain and glycation levels.

The TriVersa NanoMate (Advion Biosciences Ltd., Ithaca, NY, USA) automated sample introduction system utilises an ESI micro fluidics chip containing an array of 400 nano-electrospray nozzles. The system was controlled by ChipSoft software. Sample introduction was performed with 5 µL spray volumes and 1 µL air aspirated

after the sample, spray gas pressure 0.8 psi (later 1psi), a 1.55 kV (later 2.1 kV) voltage applied to the tip, sample plate temperature was kept at 15ºC. A 25 second delay time was applied for the inlet start sign to be sent to the mass spectrometer allowing a stable spray to form.

Data were acquired in triplicate through use of an automatic sample list. For each sample, a three minute acquisition was performed in multi-channel acquisition (MCA) mode at a scan speed of 1 spectrum/sec.

To interpret and process raw data measured in the high-throughput analysis BioPharmaLynx^{TM} software (version 1.2, Waters Corporation) was used. BioPharmaLynx^{TM} allows one to deconvolute multiple spectral outputs, deconvolution parameters are set up in the processing method, and results are listed in a table format for each processed sample. This way the calculated intensities for the different globin chains can be easily extracted. The extraction process of the required information needs to be automated to fit to the high-throughput manner of the analysis.

The processing parameters in the BioPharmaLynx^{TM} method were the following:

A Gaussian damage model was used (an isotope damage model is not available within this software package). The mass accuracy was linked to a lock mass of 1081.46 Da, with a mass tolerance of ±0.5 Da, MS mass match tolerance for deconvoluted mass peaks was 80 ppm. Deconvolution parameters: range 930-1210 Da, protein MW range 14,800-16,800, peak width 0.8 Da. Background subtraction of the deconvoluted data prior to deconvolution with a background threshold of 5% and a polynomial order 25. The deconvoluted mass spectrum was smoothed with the mean algorithm twice with a smoothing window of 6 channels. The spectrum was also centred using the top 90% and the areas or the heights of the deconvoluted peaks were reported.

After analysis the count number for intensities of pre-identified masses (at greater than 10 % intensity) for the α, β, β^{SICKLE}, δ and glycated globin chains were listed and extracted into Microsoft Excel for three replicates of each measured sample. The percentages of the δ-chain intensity and glycated α- and β-chain intensity were calculated within the software.

The percentage of glycated hemoglobin present was calculated as described by Roberts *et al*. and shown in Figure 2.1 (Roberts *et al.* 2001). The determination of

glycated hemoglobin level is based on several assumptions. It is assumed that there is no significant contribution to glycosylated hemoglobin from other globin chains, that all α-chain species have the same sensitivity and that all β-chain species have the same sensitivity. The results obtained when these assumptions were applied were shown by *Roberts et al*. to correlate with those obtained by other techniques (Roberts *et al.* 2001). Results shown later in this Chapter suggest that the calculation of only beta-chain glycation might be necessary to obtain an estimation of the overall glycated hemoglobin level.

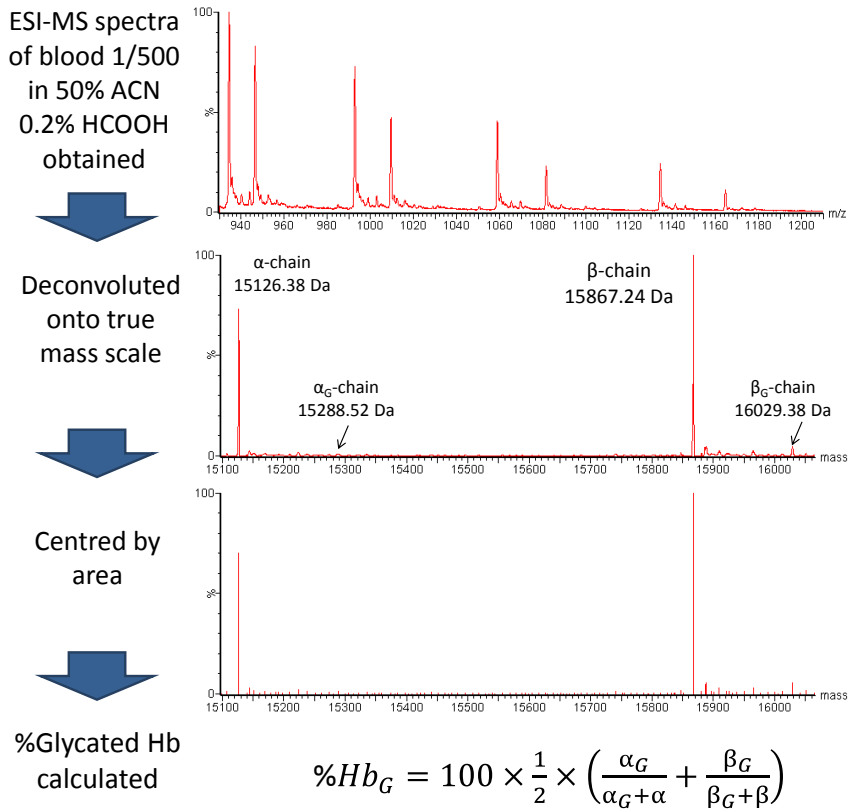


Figure 2.1. Schematic of the direct infusion method for glycated hemoglobin level quantitation

Assessment of δ-chain levels was achieved by carrying out a ratio calculation of δ-chain intensity to β-chain intensity, based on the assumption that the β- and δ-chains have similar ionisation efficiencies. This approach uses a δ:β ratio of the intact chains to provide a surrogate marker of Hb A_2 levels as illustrated by Figure 2.2.

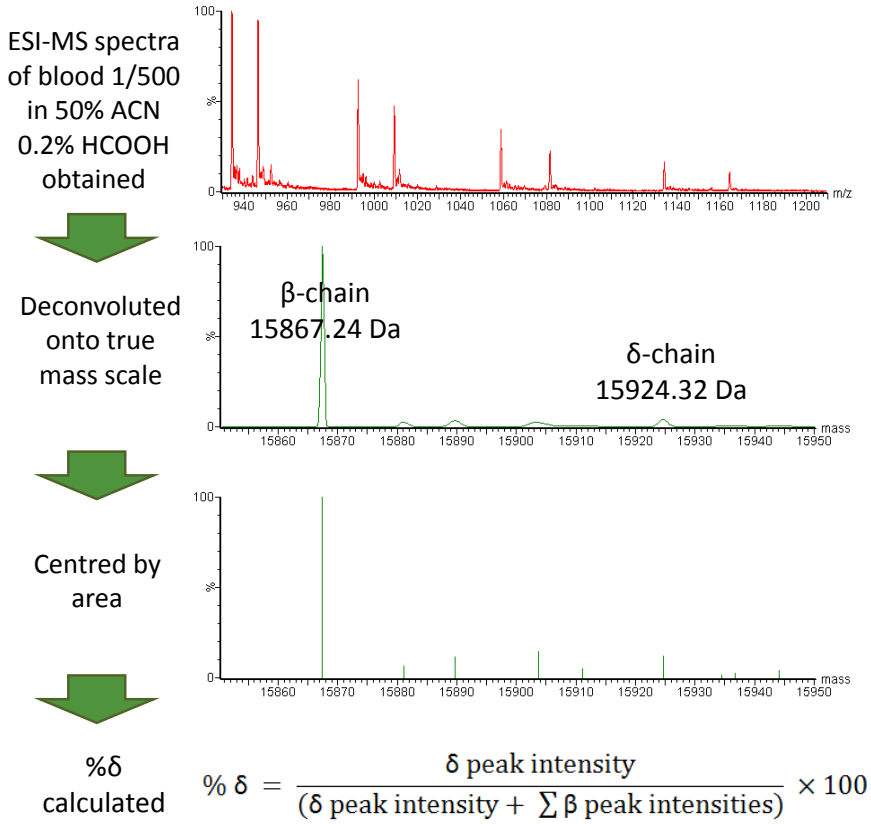


Figure 2.2. Schematic of the direct infusion method for delta chain quantitation

A summary of the processes and methods is simplified in Figure 2.3.

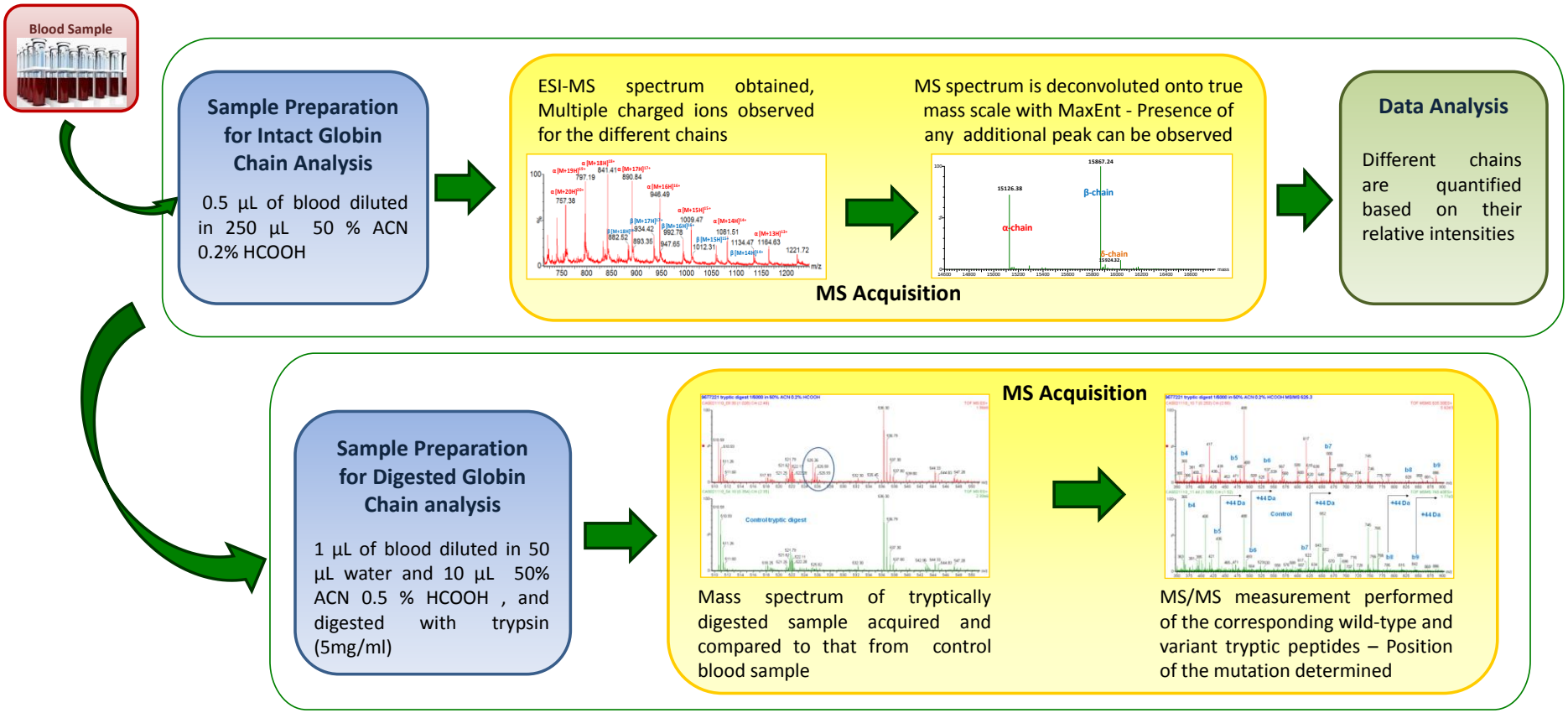


Figure 2.3. Illustration of the direct infusion method for the intact globin chain analysis and hemoglobin variant identification

MS/MS strategy for MRM method development for clinically significant variants and HbA_2 quantitation

UPLC separation for initial MRM-based method (MRM 1)

The MRM approach requires peptide separation prior to mass spectrometric detection. This was performed using an Acquity UPLC system (Waters Corporation, Manchester, UK) coupled to the TQD instrument.

2 µL of diluted peptide mixture was loaded to the Acquity UPLC BEH C18 column (1.7µm, 2.1 × 50mm) equilibrated with 95% solvent A at a flow rate of 0.8 ml/min. The composition of solvent A was 0.3% formic acid in water, solvent B was 0.3% formic acid in acetonitrile. The peptides were resolved by increasing the organic solvent concentration from 5% to 13.5% in the first 0.5 minute, then with a quasi-isocratic step further increased to 13.7% in the following 4 minutes, to 23% in another 0.5 min, and finally up to 32% by 6 minute of the run. In the last two minutes the organic concentration was increased to 99% to elute all the remaining but not monitored components of the mixture and the column was re-equilibrated with 5% B at a constant temperature of 40ºC.

Mass spectrometry method

Samples were tryptically digested to produce the necessary peptides. Preliminary measurements were necessary to choose the unique peptides to be monitored. MS/MS experiments on the chosen peptides were performed to determine the optimal parameters to obtain the highest intensity peaks from MRM measurements for specific fragmentation routes. Figure 2.4 shows possible unique peptides based on sequence differences for the relative quantitation of the HbA/HbA_2 ratios. For clinically significant variant detection and quantification, the target peptides were chosen based on the location of the mutation relevant to the specific variant, as detailed in Table 2.2. For sickle cell hemoglobin the first tryptic peptide of the beta chain was relevant.

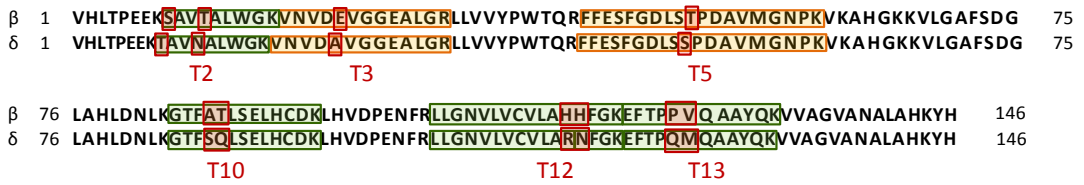


Figure 2.4. Differences in amino acid sequences between beta and delta globin chains

Name	***Position (Tryptic peptide affected by mutation)***	***Mutation***	***Mass Change of the ß chain (Da)***
Hb S	β6 (βT1)	Glu→Val	**-30**
Hb C	β6 (βT1 new peptide)	Glu→Lys	**-1**
Hb D-Punjab (or Hb D-Los Angeles)	β121 (βT13)	Glu→Gln	**-1**
Hb E	β26 (βT3 new peptide)	Glu→Lys	**-1**
Hb O-Arab	β121 (βT13 new peptide)	Glu→Lys	**-1**

Table 2.2. Tryptic peptides affected in clinically significant variant chains and the mass shift associated with the mutations.

In order to develop an MRM-based screening method for relevant hemoglobin variants, suitable transitions needed to be determined. As the transitions are based on fragmentation processes from a precursor ion to a product ion, several MS/MS experiments had to be carried out. The peptide ion of interest must be fragmented in the collision cell and product ion spectrum acquired. During these product ion scan measurements collision energy, collision gas flow, and solvent conditions can be optimized.

Optimization was predominantly carried out on commercially available samples. Lyophilized HbA, HbA_2 and HbS were used. The solid Hb samples were diluted in water to prepare stock solution, and then an appropriate amount of the stock solution was digested by trypsin. When using a sequence specific protease such as trypsin, the series of peptides produced can be predicted based on the sequence of the protein. As the clinically significant variant mutations affect the T1, T3 and T13 tryptic peptides of the beta chain (Table 2.2), and the beta and delta chains differ in the T2, T3, T5, T10, T12, T13 peptides of the beta chain (as shown in Figure 2.4), suitable MRM transitions need to be developed for the peptide ions listed in Table 2.3. As the specific peptide to be selected is known, it is possible to predict the mass-to-charge ratio of the ionised species to be selected by mass spectrometry. To investigate which mass-to-charge values of the predicted ions can be detected, tryptic

digested peptide mixtures were measured, looking for the selected ion peaks. These can be singly or multiply charged peptides. Fragmentation of doubly charged ion species typically require lower collision energies, and these ions are favoured. Doubly charged peptides are characteristic of tryptic peptides. One charge is distributed on lysine/arginine and other charge is on an amino group; occasional additional charges are usually taken up by histidine or happen because of missed cleavages.

50% ACN and 0.2% formic acid was initially used to dilute tryptically digested lyophilized blood samples, increasing the acid content of the solvent up to 0.5% when necessary. Not all targeted peptides were observable in the spectra, but for those with high intensity, tandem MS measurements were performed, and product ion spectra were obtained. The solvent composition was altered from 0.1% formic to higher formic acid content (0.3% and later 0.5%) to increase chances of generating ions for the targeted peptides, to produce doubly charged ions for the peptides of interest, and to influence their retention times facilitating their elution at a later time. The optimal transitions for MRM method were then selected for evaluation using clinical samples.

Tryptic peptide	Mutation correlated	Amino acid sequence	$[M+H]^+$/Da	$[M+2H]^{2+}$/ Da
Wild-type βT1	Hb Sickle, HbC	VHLTPEEK	952.51	476.76
HbS βT1	-	VHLTP**V**EK	922.53	461.77
HbC βT1	-	VHLTP**K** EK	694.43 276.16	347.72 138.58
Wild-type βT2	Differs from the δT2	SAVTALWGK	932.5	
Wild-type δT2	Differs from the βT2	**T**AV**N**ALWGK	959.5	
Wild-type βT3	Hb E, Differs from the δT3	VNVDEVGGEALGR	1314.67	657.84
HbE βT3	-	VNVDEVGG**K** ALGR	916.47 416.26	458.74 208.64
Wild-type δT3	Differs from the βT3	VNVD**A**VGGEALGR		628.9
Wild-type βT5	Differs from the δT5	FFESFGDLSTPDAVMGNPK		1030
Wild-type δT5	Differs from the βT5	FFESFGDLS**S**PDAVMGNPK		1023
Wild-type βT13	Hb D-Punjab, Hb O-Arab Differs from the δT14	EFTPPVQAAYQK	1378.70	689.85
HbO βT13	-	**K** FTPPVQAAYQK	147.11 1249.66	74.06 625.33
HbD βT13	-	**Q**FTPPVQAAYQK	1377.72	689.36

Table 2.3. List of tryptic peptides and variant tryptic peptides and expected mass to charge ratios. (Corresponding peptides are clustered based on what part of the amino acid sequence they comprise, first the wild-type peptide present in normal adult hemoglobin)

MRM 1 method using Waters Acquity TQD system

For HbA_2 quantitation the corresponding T2, T3 and T5 peptides were chosen. T5 peptides have not been used for this purpose in the literature up to date. T2 peptides were not studied further since the resolved peaks were not suitably intense after chromatographic separation in the Acquity TQD system with the applied sample

preparation. The chosen MRM transitions can be seen in Table 2.4. Baseline separation of the different peptides with the UPLC system was not possible. A multi-step gradient method was therefore applied, where all corresponding pairs of peptides were separated, but only βT2 and δT2 peptides are fully separated from each other. This method was used for the analysis of 60 patient blood samples. Further optimisation was required from the initial method based on commercially available samples. During measurement the ion source was maintained at 150ºC, desolvation temperature 320ºC, desolvation gas flow at 800L/hr, and cone gas flow at 50L/hr.

Peptide	Retention time window (min)	Target Precursor Ion (Da)	Target fragment ion (Da)	Cone Voltage (V)	Collision Energy (V)
β T2	2.80-3.30	932.5 (z=1)	675.4 774.5	60	45
δ T2	1.80-2.10	959.5 (z=1)	688.4 787.4	50	50
β T3	1.20-1.60	**657.8 (z=2)**	**887.5** **1002.5**	35	22
δT3	1.20-1.60	**628.8 (z=2)**	**829.5** **944.5**	35	25
βT5	5.40-5.70	**1030 (z=2)**	**928.5** **1116.5**	35	22
δT5	5.40-5.70	**1023 (z=2)**	**1015.5** **1102.5**	35	25

Table 2.4. The developed MRM transitions for T3 and T5 peptides in the beta and delta chains of hemoglobin on a Waters TQD .

Data Analysis

QuanLynx™ software (MassLynx™, Waters Corporation, Manchester, UK) was used for automated data processing. Examples of peak integration are illustrated in Figure 2.5. The processed data were extracted into Microsoft Excel and the percentage of δ-chain peptides relative to the sum of all non α-chain peptides calculated based on **Equation 2.1.** Only T3 and T5 peptides were used, since T2 peptides were not sufficiently intensive to provide reliable values using the calculation.

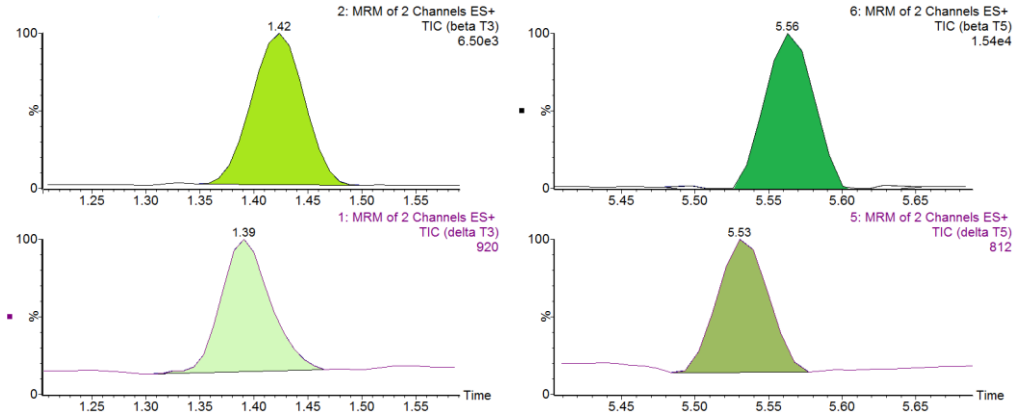


Figure 2.5. Chromatogram of the MRM channels, integrated for T3 (a) and T5 peptides (b)

For the δ-chain level determination the following equation was used.

$$\%\boldsymbol{\delta} = \frac{\boldsymbol{\delta\ peptide\ peak\ area}}{(\boldsymbol{\delta\ peptide\ peak\ area} + \boldsymbol{\beta\ peptide\ peak\ area})} \times \mathbf{100}$$

Equation 2.1.

The scheme for the **MRM 1** analysis can be seen in Figure 2.6.

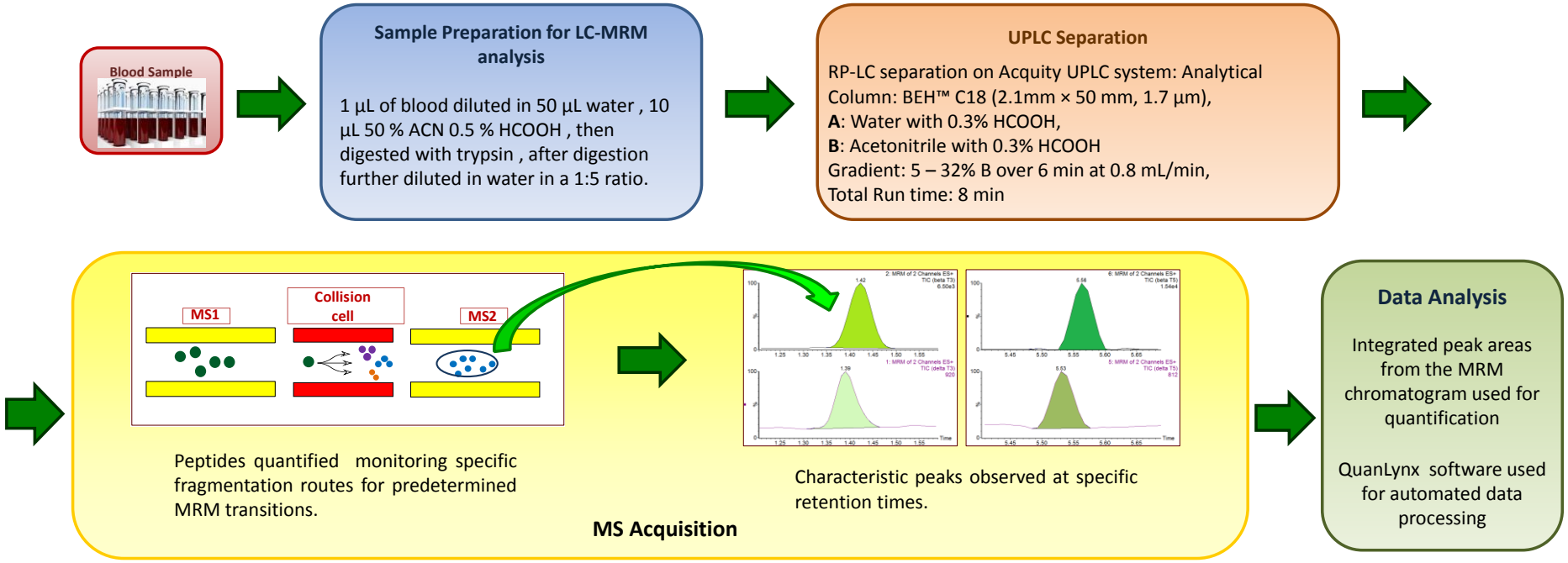


Figure 2.6. Scheme of the MRM 1 approach for specific peptide pair relative quantification to obtain estimation of the HbA_2 levels

MRM 2 method using Waters Xevo TQ

UPLC separation for final MRM-based method (MRM 2)

Tryptic peptides were separated by RP-LC prior to MRM analysis. RP-LC separation was performed on a standard flow Acquity UPLC system, equipped with a C18 column.

5 µL of the diluted peptide mixture was loaded onto the Acquity UPLC BEH C18 column (1.7µm, 2.1x50mm) previously equilibrated with 98% solvent A at a flow rate of 0.8 ml/min. The composition of solvent A was 0.5% formic acid in water, solvent B was 0.5% formic acid in acetonitrile. After keeping the composition at 2% of solvent B for one minute, the peptides were resolved with increasing the organic solvent concentration from 2% to 27% over a 4 minute period, then with a steep gradient further increased to 95% to elute all the components of the mixture, and the column was re-equilibrated with 2% B for an additional 1.5 minutes at a constant temperature of 40ºC.

Mass spectrometry method

For HbA_2 quantitation the corresponding T2, T3 and novel T5 peptide pairs were chosen, for clinically significant variants the corresponding variant peptides were monitored after chromatographic separation in the Xevo TQ system. The optimised MRM transitions can be seen in Table 2.5. Full separation of all of the peptides with the UPLC system was again not possible, and a linear gradient method was applied, where all corresponding pairs of peptides were separated, but only β T2 and δ T2 peptides were fully separated from each other. During measurement the ion source was maintained at 150ºC, desolvation temperature 350ºC, desolvation gas flow at 900L/hr, and cone gas flow at 50L/hr.
For all peptides at least 3 MRM transitions were optimised, and where possible, four MRM transitions have been chosen to obtain higher specificity.

Peptide	Retention time window (min)	Target Precursor Ion (*m/z*)	Target fragment ions (Da)	Cone Voltage (V)	Collision Energy (V)
β T2	3.80-4.29	**466.8** (z=2)	**390.2** **675.1** **774.2**	18	22 18 16
δ T2	3.51-3.79	**480.4** (z+2)	**390.2** **688.3** **787.3**	20	22 18 18
β T3	3.30-3.50	**657.8** (z=2)	**758.4** **887.5** **1002.5**	28	25 32 28
δT3	3.30-3.50	**628.8** (z=2)	**829.5** **944.5** **1043.4**	25	28 26 25
βT5	4.75-5.02	**1030** (z=2)	**928.5** **1116.5** **1229.6** **1401.6**	38	30 32 35 35
δT5	4.75-5.02	**1023** (z=2)	**1015.5** **1102.5** **1215.6** **1387.7**	35	35 35 35 35
β T1	2.15-2.40	**476.8** (z=2)	**237.2** **502.5** **603.3** **716.4**	25	20 22 20 18
β T1 HbS	2.48-2.72	**461.8** (z=2)	**237** **472.3** **573** **686.2**	25	18 22 20 18
β T1 HbC	1.80-2.14	**694.4** (z=1)	**237** **244** **345** **350**	40	40 40 35 35
β T3 HbE	2.41-2.50	**916.8** (z=1)	**214.2** **261.2** **360.1**	40	50 45 42
α T9 (2) HbG	4.3-4.7	**771.5**(z=3)	**708.3** **807.4** **839.4** **1078.6**	35	35 30 35 30
β T13 HbO	3.30-3.50	**625.3**(z=?)	**708.4** **904.5** **1001.6**	25	32 25 20

Table 2.5. Tryptic peptide ions and targeted fragment ions on the Xevo TQ instrument used for MRM transitions.

Data Analysis

The TargetLynx™ software (MassLynx™, Waters Corporation, Manchester, UK) was used for automated data processing. The HbA_2 values were calculated on the same basis as in the case of the results of MRM 1 method with the use of **Equation**

2.1., and the variant peptides were confirmed with the observation of the presence of a peak in the corresponding MRM chromatogram.

Peptide peak areas were calculated by integrating the smoothed and combined chromatogram from all corresponding MRM channels. Results have been exported to Excel and area ratio calculations for different corresponding peptides have been performed. Standard deviation for 3 replicates and the corresponding coefficient of variation percentages have been calculated.

2.3 Results and Discussion

2.3.1 HbA_2 and Glycated hemoglobin level determination analysing intact globin chains

Standard samples with known levels of HbA_2 and HbA_{1c} were analysed and used to assess the suitability of the proposed quantification process.

HbA_2 standard measurement

The NBISC Reference standard with a known concentration of HbA_2 (5.3%) was used to test the intact globin chain quantitation method to assess the reliability of the results. The deconvoluted ESI spectra of three replicates centred based on peak intensities and areas provided 4.9% (standard deviation (stdev): 0.2, coefficient of variation (cv%): 5) and 6.8% (stdev: 0.3, cv%: 4.5) for HbA_2, respectively. The percentage calculated based on peak intensities are closer to the actual value, and differ only by -0.4%, while the percentage based on area differs by +1.5%. Standard deviation and cv% results suggest this measurement is reproducible. Figure 2.7 illustrates a typical deconvoluted mass spectrum of the standard used for δ-chain level calculation.

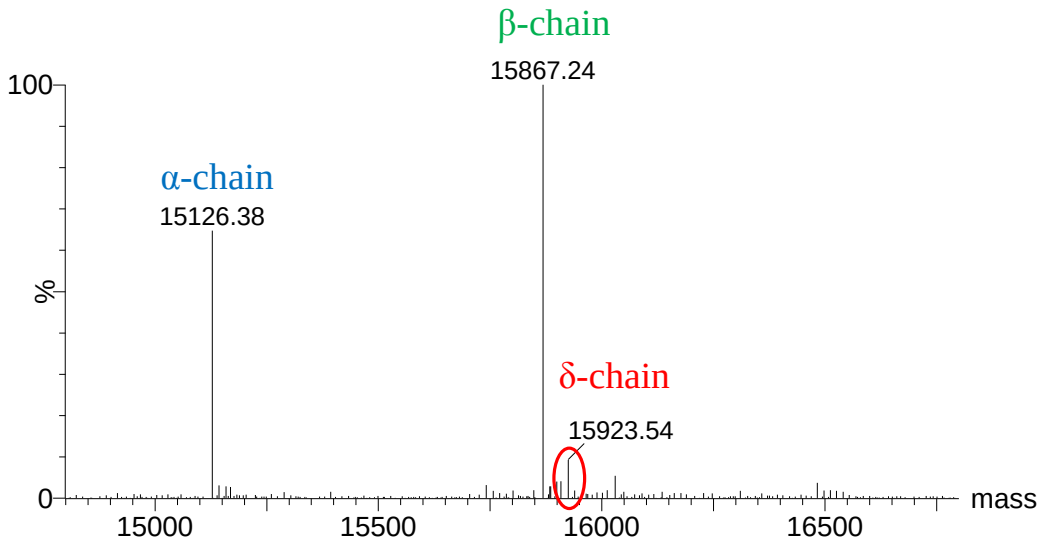


Figure 2.7. Deconvoluted mass spectrum of intact globin chain analysis of international reference reagent for HbA_2 level determination

HbA_{1c} standard measurement

The deconvoluted ESI spectrum was centred based on peak intensities and areas provided the following results from calculations using intensities for α_G 2.9% (stdev 0.33, cv% 11.35), for β_G 6.3% (stdev 0.33, cv% 5.25) and for $(\alpha\beta)_G$ 4.6% (stdev 0.29, cv% 6.31). Results obtained by area calculations for the α_G 3.3% (stdev 0.24, cv% 7.40) for β_G 7.7% (stdev 0.11, cv% 1.42) and for $(\alpha\beta)_G$ 5.5% (stdev 0.12, cv% 2.26).

The quantitative values obtained for different glycated chains and overall hemoglobin are significantly different. The α-globin chain glycation value is approximately half of the calculated β-chain glycated value. All measured values, including the highest glycated β-chain levels, are lower than the theoretical 9.3%. The theoretical value describes the HbA_{1c} levels, which is defined as the β-globin chain glycated on the N-terminal valine.

Figure 2.8 illustrates a typical deconvoluted mass spectrum of the standard used for glycated chain level calculation.

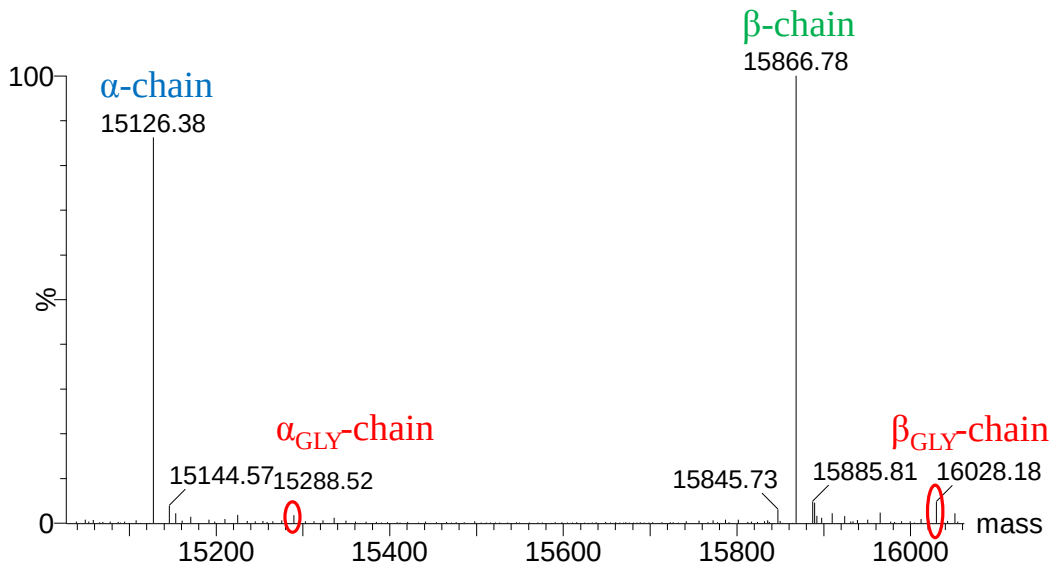


Figure 2.8. Deconvoluted mass spectrum of intact globin chain analysis of reference standard with known level of HbA_{1c}

It can be observed on the mass spectrum that the peak corresponding to glycated β-globin chain is significantly more intense than the peak corresponding to glycated α-globin chain.

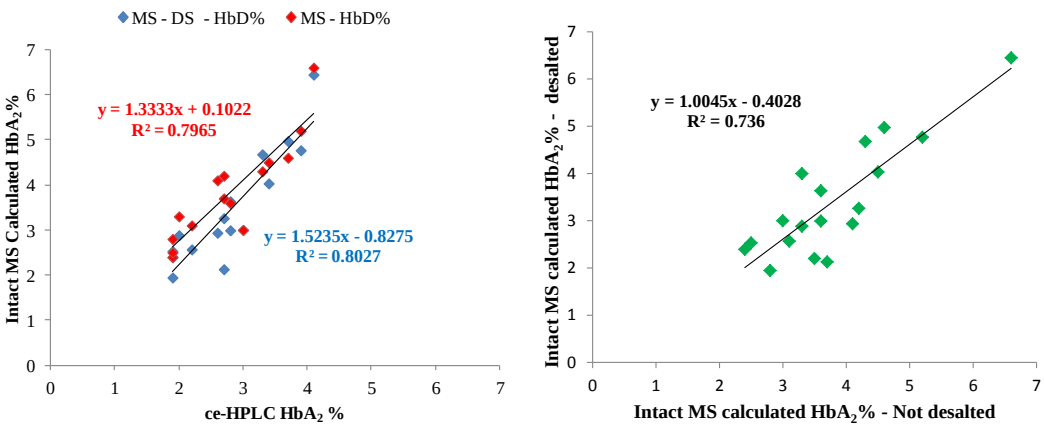
Method optimisation on clinical samples

HbA_2 level determination

A sample batch compromising 96 samples was analysed, measuring three replicates, calculating the percentages, the average of the three replicates and standard deviations with corresponding coefficients of variation for each sample. The analysed dataset provided comparable HbA_2 values to those determined by HPLC from current clinical screening settings. There was a total number of 92 clinical samples, and sample replicates of HbA_2 and HbA_{1c} standards were also included.

The data revealed that delta chain levels were overestimated with a certain bias using the ESI-MS method. In the multiply charged peak envelope, mass-to-charge ratios originating from delta globin ions and possible salt adducts of ions can overlap. Poor resolution of these ions may have an adverse effect on quantitation following deconvolution. It was therefore hypothesised that the intensity of these interfering ions could be reduced with the application of desalting techniques. The phenomenon was assumed to cause more erroneous results where samples had been stored for a relatively long period of time. Here the results were obtained from relatively fresh samples, which were analysed within one to two weeks after the blood was taken from the patient, and stored at 4ºC.

Desalting techniques were evaluated using fresh clinical samples. The calculated HbA_2 levels can be seen in Figure 2.9 a) with and without desalting the samples with cation exchange beads. Desalting samples with spin columns was not performed on a bigger sample set as it would significantly increase the cost of the analysis.



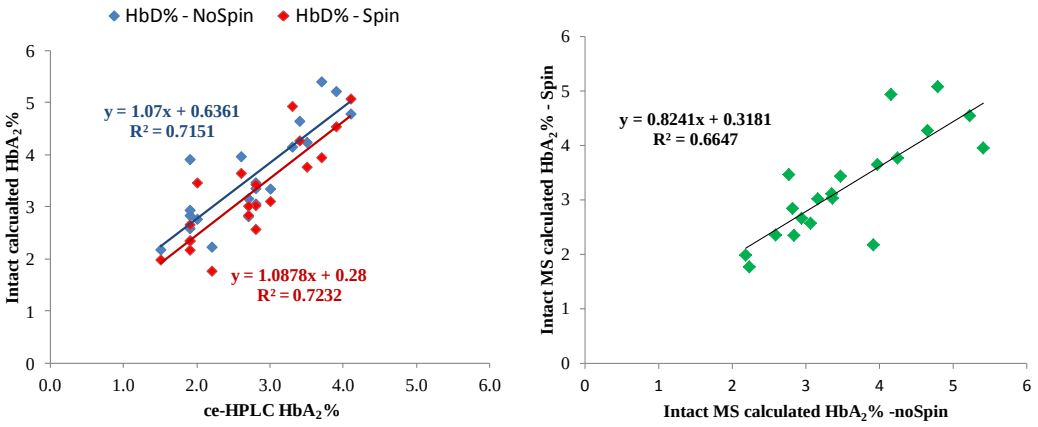
a) **b)**

Figure 2.9. Results of 15 samples analysed with (blue dots) and without (red dots) desalting prior to analysis (a) and showing the correlation calculations by plotting the results of different sample preparation methods against each other (b)
Correlation between calculated values from the MS measurement to the cation exchange chromatography measurements are similar with and without the desalting step in sample preparation.

Plotting the results of the desalted and untreated samples against HPLC results demonstrates that the difference in the values and their distribution is not significant. Plotting the two set of results against each other the correlation was investigated in figure 2.9 b). The untreated samples resulted in slightly higher HbA_2 levels, but there was no systematic difference between the results for the two types of sample preparation. For the purpose of minimising the sample preparation and analysis time it was decided that the analysis of further clinical samples would not include a desalting step.

For the automated sample introduction with the NanoMate system, a particle-free sample solution was necessary to prevent blockage of the nozzle, which would terminate the formation of sample spray during measurement. To avoid blockage, samples were centrifuged and the effects of this step on delta chain level quantitation were assessed by comparison of calculated values for 24 clinical samples. The results of this comparative analysis can be seen in Figure 2.10.
The difference shown in figure 2.10 a) is not significant, but the values are slightly lower for the samples which had been centrifuged. The correlation between the two sets of data is good, with a few values further away from the fitted trendline as swhown in figure 2.10 b). It was observed that centrifuged samples blocked the

nozzle less frequently and therefore it was decided to include this extra centrifugation step in the sample preparation.



a) **b)**

Figure 2.10. Results of 10 samples analysed investigating the effects of centrifuging prior to analysis plotted against HPLC determined HbA_2 values (a) and showing the correlation calculations by plotting the results against each other (b)
Correlation between calculated values from the MS measurement to the cation exchange chromatography measurements are similar with and without the centrifugation step in sample preparation.

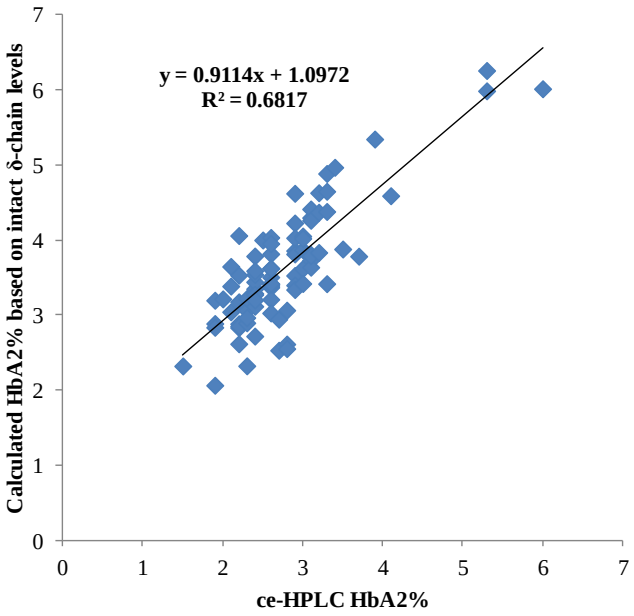


Figure 2.11. Results of 92 samples - intact globin chain analysis determined HbA_2 values plotted against HPLC values

The analysis of a larger clinical sample set showed a similar correlation to the HPLC values as the results obtained with smaller sample numbers during sample preparation optimisation experiments.

The determined HbA_2 levels by intact MS analysis of blood samples plotted against HPLC values are shown in Figure 2.11. The method provided values distributed in a fairly wide range around the expected values provided by HPLC. The calculated levels are usually higher than that expected. Samples with known elevated HbA_2 levels resulted also in relatively higher levels obtained by MS. The MS method provides a good indication of sample abnormality. For a statistically relevant normal range of healthy levels based on the MS technique, a larger sample set needs to be analysed. The average coefficient variation for the delta-chain level was calculated to be 10.3%. This reproducibility may be considered as good for biological sample analysis.

The error of experimental masses to theoretically expected masses

Based on the literature, error in α- and β-chain measurement could be used to indicate the presence of an α- or β-chain variant (with a mass shift 6 Da or less).
The presence of structural variants resulting in a mass shift of more than 6 Da is assumed to be readily visible on the deconvoluted mass spectra by the presence of an additional peak at greater than 20 % intensity. (Note that the used instrument here is a low resolution instrument, therefore the mass shift needs to be higher than a certain observable mass difference.)
In Figure 2.12 (a) the experimental mass errors can be seen for 96 samples and three replicates for alpha-chain and beta-chain mass errors. The majority of the mass errors are between 0.1 and 0.4 Da, and the alpha-chain mass errors are generally higher than those ones calculated for beta-chain mass measurements.
It is generally observed that there is a higher chance to have a β-chain variant than an α-chain variant, all monitored clinically significant variants are β-chain variants, so it is more likely that the α-chain mass is the true mass corresponding to the normal α-chain. Note, if there is a mass difference believed to be a result of a -1 Da β-chain variant, and after thorough investigation there is no variant identified, the possibility of the presence of a ±1 Da α-chain variant will be investigated. With this assumption of the alpha-chain masses being correct, adjustment to beta-chain mass error can be applied with the deduction of alpha-chain errors. The corrected beta-chain errors can be seen in Figure 2.12 (b). This adjustment places the majority of the mass errors of this dataset between 0 and -0.4 Da, slightly higher than the ±0.1 Da mass accuracy

reported by Rai *et al*. (Rai *et al*. 2003). There is an observed mass error difference seen in case of the calculated α- and β-chain masses. This difference may be explained with the processing parameters. The α-chain is used as a calibrant before the measurements and also one of the multiply charged mass peaks belonging to the α-chain is used as a lock mass. These settings would affect the calculated masses for the α-chain differently as they do for the β-chain masses. To obtain similar mass errors for both chains further investigation and optimization of the processing parameters is advised.

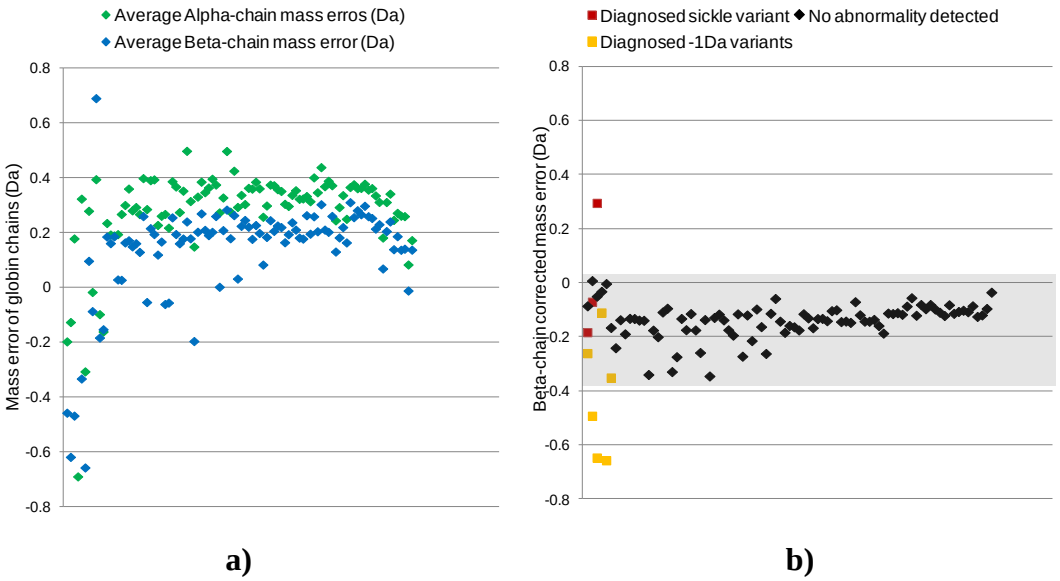


Figure 2.12. (a) Mass errors of different globin-chains resulted from the analysis of 96 samples in three replicates, and (b) corrected mass errors of beta globin-chain black dots for normal samples, red for sickle variants, and yellow for -1 Da variants diagnosed in the hospital lab

Samples with a much higher mass error may be explained as clinically significant variants, but not in every case, and at this stage it is not advisable to be dependent on these values for diagnosis. These error values may give erroneous results, although in some cases they may indicate the presence of a variant correctly.

The diagnosed variants have not resulted in higher than average mass errors in all cases. The sickle variant differs by -30 Da to normal β-globin chain mass. Based on the hypothesis that -1 Da variants will shift the masses by 0.1 Da if they are present at least at 10%, the presence of sickle variant should not affect the mass accuracy. For two of the sickle samples no significant outstanding mass error is observed. For one sickle sample there is an observed mass error of +0.3 Da. This sample was

homozygous for the sickle variant, with a low level of normal beta-chain, which had potentially remained in the patient's blood system after a previous blood-transfusion. Since the level of the normal β-chain is low the observed mass difference is not considered abnormal.

The samples diagnosed with HbC, HbE or HbD-Punjab variants by clinical methods did not show significant mass errors in every case. Six samples were diagnosed as clinically significant -1 Da variants. Three of these samples were confirmed by MS to have one of each mentioned conditions (HbC, HbE and HbD). Two of the other three samples were identified as carriers of a +14 Da α-chain variant, HbG-Philadelphia. The third sample classed as HbD-Punjab carrier was not confirmed to have this condition on the tryptic peptide level analysed by high-resolution MS.

The observed error distributions of different mass measurements are illustrated by box plots in Figure 2.13. The three different box plots show in which ranges the majority of these mass errors can be found. It is clear, that α-chain mass errors are generally higher than β-chain mass errors, and that the corrected mass errors for the β-chains are below zero.

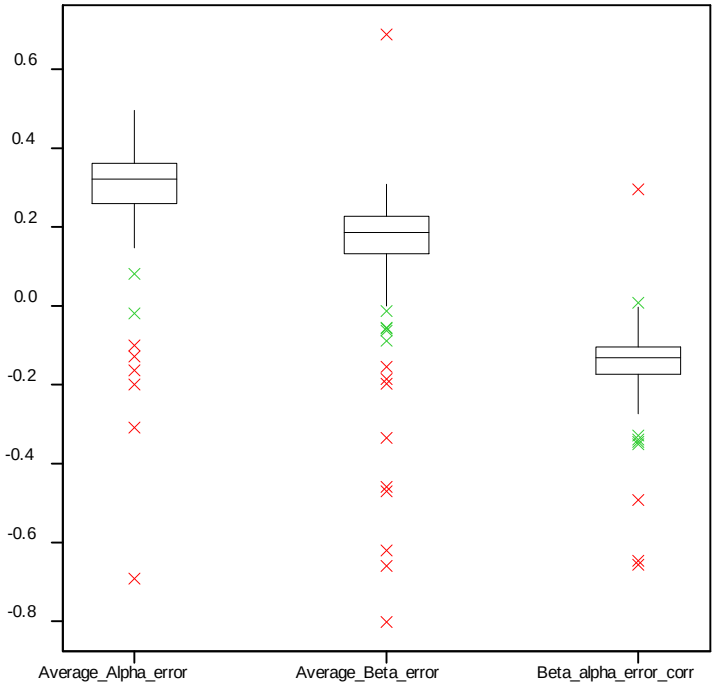


Figure 2.13. Mass error distribution of the globin chains separately and adjusted beta-chain errors.

The samples with green stars are within the 1.5 IQR (interquartile range) from the upper and lower quartiles distributed around the mean difference. Looking at only the corrected β-chain mass errors four samples are outside of the acceptable range, including those three samples which were observed to be abnormal in Figure 2.12 (b) (yellow squares).

Based on these results it can be proposed that intact mass error values may be indicative for the presence of -1 Da variants, if the values fall outside of the 1.5 IQR of the upper and lower quartiles for a certain local distribution of the mass errors of normal samples.

HbA_{1c} level determination

As mentioned previously, intact globin chain analysis can provide an estimation of glycosylated hemoglobin levels. Calculations were performed to obtain estimated glycation levels for the globin chains separately and for the overall glycation assuming that glycation occurs at the same rate on both chains. There were no accessible values measured by screening methods, therefore calculated values could

not be correlated to predetermined HbA_{1c} levels. The ClinTest HbA_{1c} standard was included in the sample set to obtain an estimation, or a correction factor based on the levels determined by the measurement of the standard with known levels of HbA_{1c}. The HbA_{1c} level of the standard was 9.3%. The measured values for the standard were largely under estimated as previously at 3.9%, 8.4% and 6.1% for alpha-chain, beta-chain and united respectively. The calculated value for the glycated beta-chain level was closest to the expected value. The combined value from both chains was much lower due to the fact that the alpha-chain glycation levels were less than half the expected 9.3%. Plotting all the calculated results in ascending order for all three calculated levels, the general trends observed are similar in Figure 2.14.

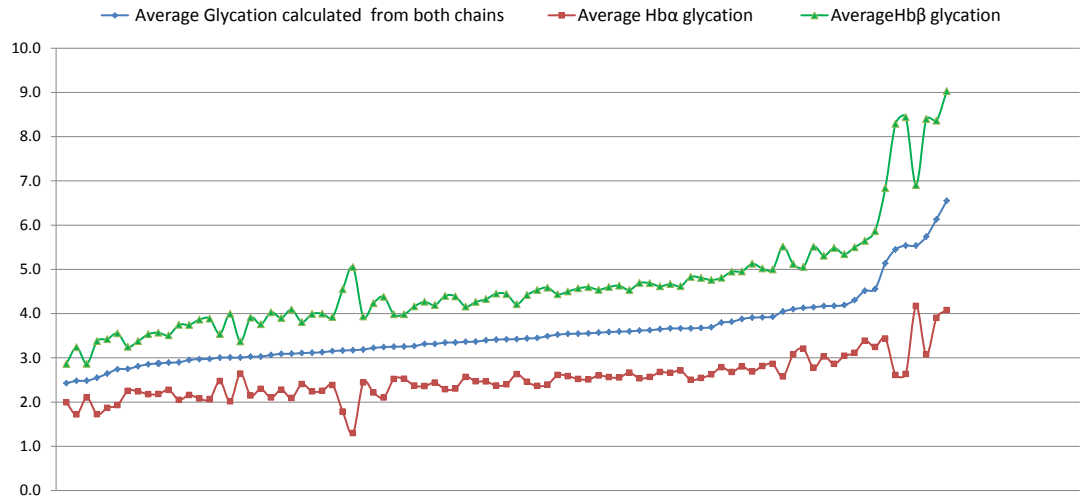


Figure 2.14. Calculated glycation levels for separate globin chains and based on the average glycation originated from both chains (average of 3 replicates)

There are some examples where opposite deviating values were observed, but were compensated to an average level by the combined calculation.

Assuming that the beta levels are more reliable and closer to actual values, the combined calculation can give incorrect results. This observation would benefit from further investigation. Assuming that alpha- and beta-chains are glycated at a similar rate, one would assume there is significant correlation between the values obtained for the different globin-chains. To test this correlation calculated glycated beta levels were plotted against glycated alpha-chain levels. The result can be seen in Figure 2.15.

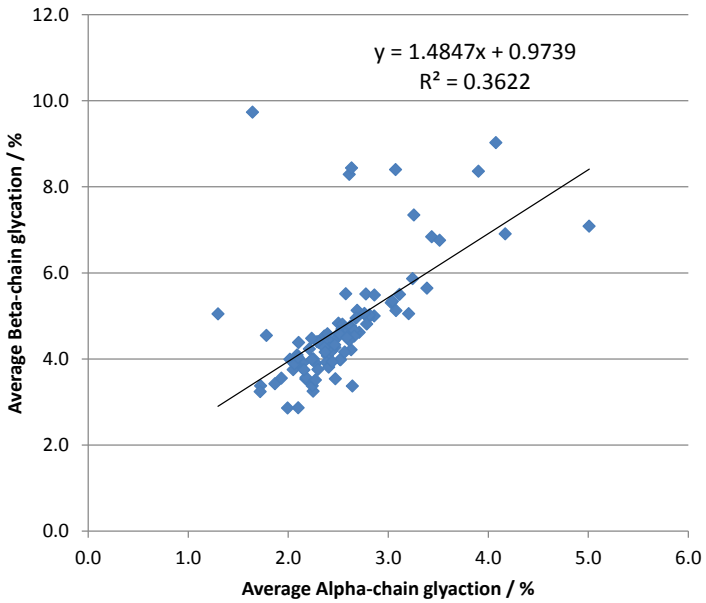


Figure 2.15. Correlation analysis between calculated glycation levels of the different globin chains

For a high proportion of the samples this hypothetical correlation seems to be confirmed, but there are samples where the elevated glycated beta-chain levels are observed. This indicates there may be a requirement to establish whether both chains can be used to contribute to the calculation of glycation levels.

2.3.2 Targeted peptide analysis

2.3.2.1 Initial method developed for delta chain quantitation on a Waters TQD system – MRM method 1

The results for mixed digests of HbA_0 and HbA_2 demonstrate good correlation (R^2 = 0.995) with the expected values for the calculated HbA_2 levels. Calculated average values are presented with error bars of the standard deviation across 3 replicates for quantitation determined using T3 peptides in Figure 2.16.

The measured and calculated values for T3 peptide ratios are double that of the expected HbA_2 levels, but the correlation between theoretical and calculated values suggests that the method may provide a good indication of HbA_2 levels relative to HbA present in patient blood samples. The average coefficient of variation for T3 peptide ratio calculations was 17%, and for method validation of mass spectrometry-based clinical applications the requirement is to be below 15% (Kollipara *et al.* 2011).

The calculated values for T5 peptide ratios are closer to the expected HbA_2 levels, than the T3 ratio results, as shown in Figure 2.17. The average coefficient of variation for T5 peptide ratio calculations was 19%, again higher than the required 15%.

The slope of the gradient in both cases suggests the future measurement of a reference standard may be required to adjust measured values to get real HbA_2 levels.

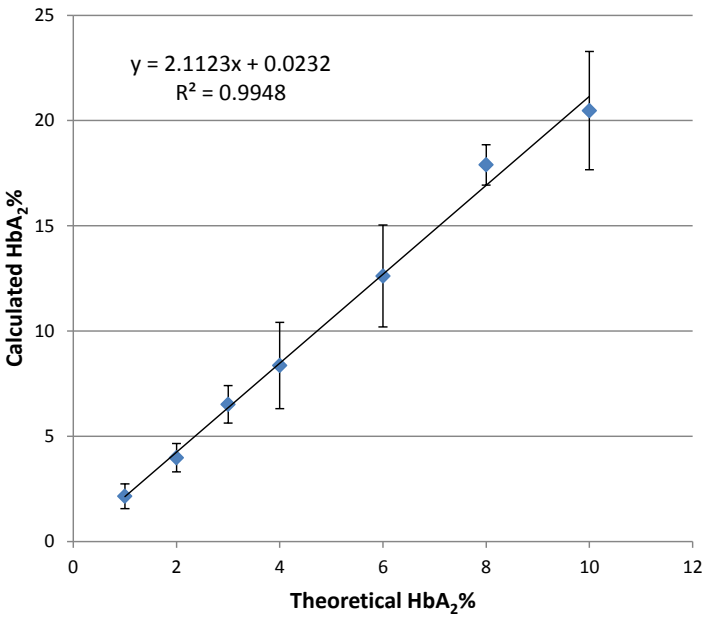


Figure 2.16. HbA_2 calculated levels based on T3 peptide ratios - average from 3 replicates

R^2 value is 0.995, which is promising for method evaluation, although error of the measurements are noticeably high for some of the values.

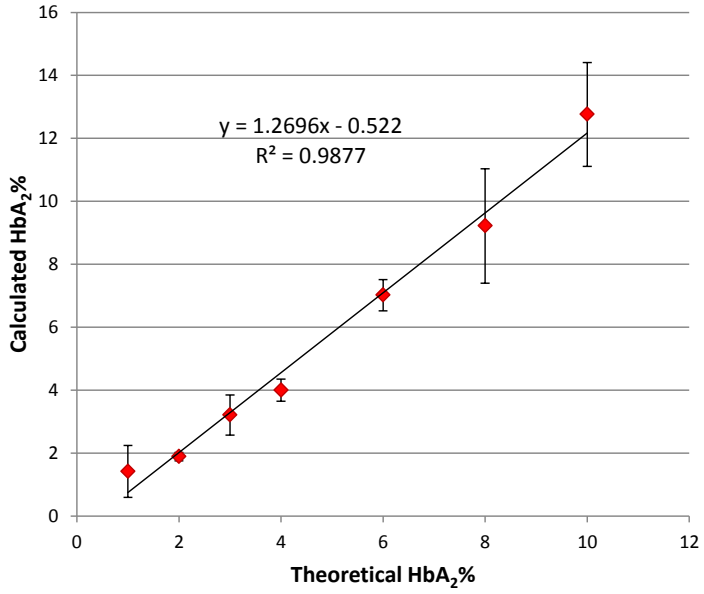


Figure 2.17. HbA_2 calculated levels based on T5 peptide ratios - average from 3 replicates

Although the method at this stage did not meet all the requirements for a competitive mass spectrometry-based method to be used in clinical screening, it was tested on clinical samples to evaluate the effects of the complex biological matrix.

60 patient blood samples were analysed at the peptide level. It is possible that the biological matrix could introduce interfering ions, which may make quantification difficult. Results are shown in Figure 2.18 for T3 (a) and T5 (b) peptides. There was no preliminary information provided about the actual values regarding HbA_2, but samples were analysed by intact globin chain analysis in our lab by Charlotte Scarff, and the same three samples showed elevated HbA_2 levels by T3 peptide ratio calculations as the intact globin chain analysis. The T5 peptide results show elevated levels for only two samples, and did not show significant increase compared to average levels in the case of the third sample. The majority of the calculated T3 values were between 4-6 %, with the elevated levels above 8%, for the T5 peptides the majority of the produced results were between 1.5 and 3%, elevated levels at and above 4%. This was to be expected for both when considering the correlation indicated in Figure 2.16. and 2.17.

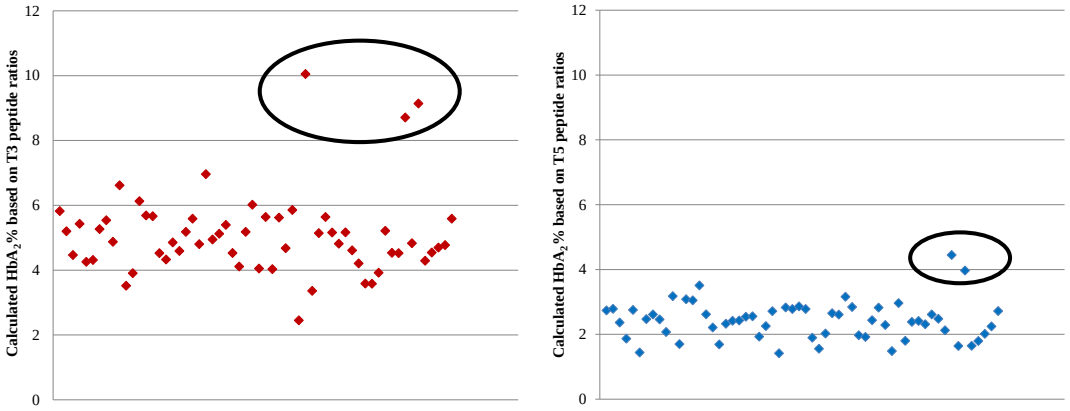


Figure 2.18. Calculated percentages for the 60 clinical samples based on T3 (a) and T5 (b) peptide ratios

Based on observed values when measured the premixed samples with known theoretical HbA_2 levels, it is expected that the T5 peptide ratios will be lower than T3 peptide ratios. The T5 peptides of the different globin chains are structurally more similar to each other than T3 peptides, since for the latter the differing amino acids are on the different end of the polarity scale, while for the T5 peptides Ser and Thr are both amino acids with polar sidechains.

The significant difference observed between values of corresponding peptide ratios was assumed to be due to the varying ionisation efficiencies of different peptides.

Average CV values are based on two measurements for T3 peptides 12.6% and for T5 peptides 21.6%.

This method showed promise for clinical applications at this stage, but needed further development and optimisation. It has been demonstrated that the method is capable of detecting elevated levels of HbA_2. The measurement of a standard, with known levels of HbA_2, is necessary to obtain correction factors for these peptide ratios from which true δ-chain levels may be calculated.

Whilst the method has been shown to be able to indicate the presence of β-thalassemia when δ-chain intensities are significantly outside the normal range, this may not be the case when slightly elevated δ-chain levels are present. Further development work was necessary, with special focus on the novel T5 peptides.

Measurements need to be conducted on blood samples, where HbA_2 levels have been previously determined by the validated ce-HPLC method currently used in the clinical laboratories.

2.3.2.2 Optimized Method for Clinical Sample Analysis on a Waters Xevo system – MRM method 2

HbA_2 quantitation

The previously presented MRM method has been adapted and further optimised on a Waters Xevo TQ system. The new method has also been further developed to monitor the presence of certain characteristic peptides of clinically significant variants.

The results for mixed protein digests of HbA_0 and HbA_2 demonstrate a good correlation to expected values for the calculated HbA_2 levels with significantly lower error values than with the previous method. The calculated results plotted against the theoretically expected values are shown in Figure 2.19. R^2 values are 0.982, 0.998 and 0.989 for T2, T3 and T5 peptides respectively. The values for T2 peptide ratios are the highest, slightly above expected, the T5 peptides give the closest values to expected, while the T3 peptides slightly underestimate the theoretical HbA_2 levels. This was to be expected when the structural similarities between T5 peptides and between T3 peptides from the different globin chains were investigated.

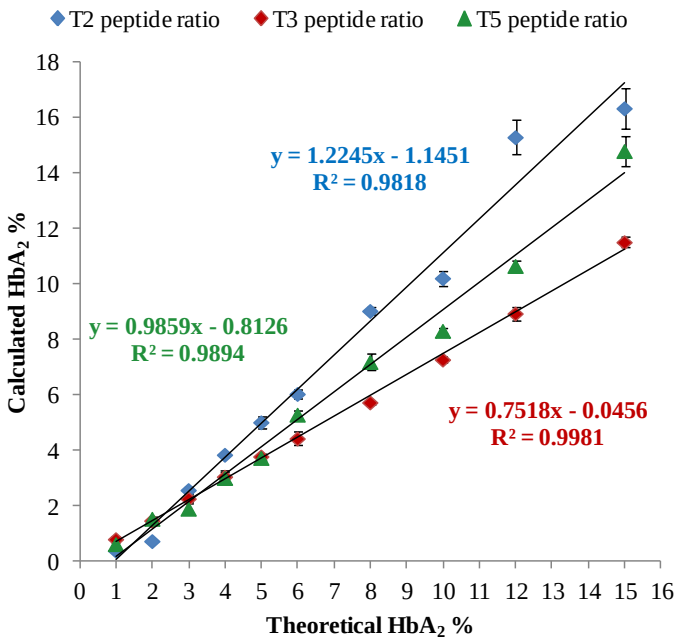


Figure 2.19. Calculated percentages based on T2, T3 and T5 peptide ratios shown against the theoretically expected HbA_2 level

R^2 values of 0.98-0.99 are promising for method evaluation, the differences between the slope of the gradients for different peptides highlight the need for the measurement of a standard for accurate HbA_2 level determination to get realistic quantitative results

A smaller sample set comprised of 20 patient blood samples were analysed using this method, and HbA_2% was calculated using integrated peak area ratios of the corresponding peptides of beta and delta-chains.

The results demonstrate excellent correlation to the ce-HPLC values as illustrated in Figure 2.20. The T2 peptide pair had slightly higher values than expected, the T3 peptides had slightly lower values, and the novel determination of T5 peptide ratios were closest to the expected values. This is the same tendency as observed. This was expected based on previously reported results and observations by Daniel *et al*, who used a reference standard to calculate correction factors for all corresponding peptide pairs to calculate values which were not biased by ionisation differences between peptides (Daniel *et al.* 2007)

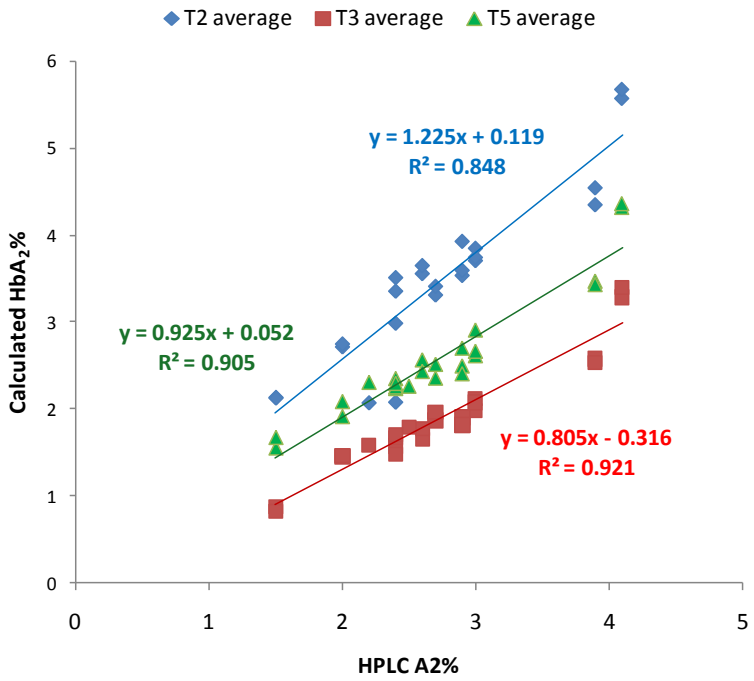
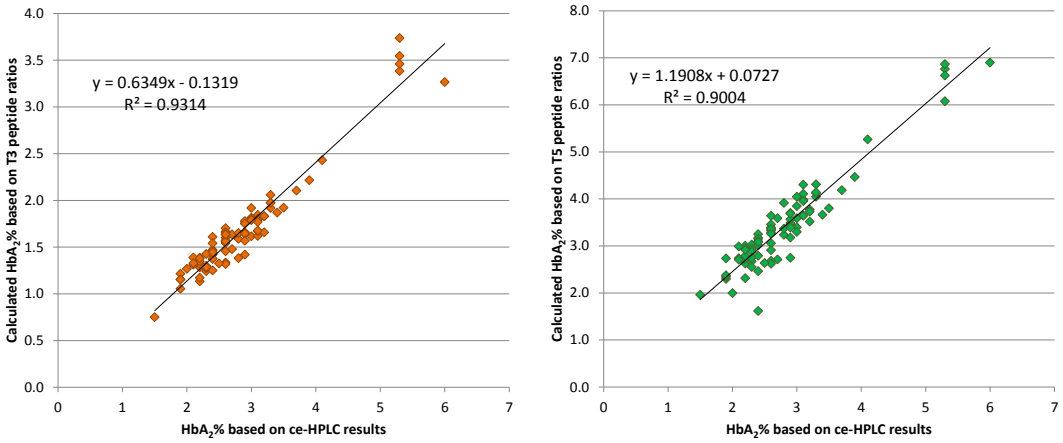


Figure 2.20. Calculated δ-chain % (HbA_2 levels) for 20 clinical samples as results of a preliminary testing calculated based on three different peptide ratios

To further evaluate the method a larger sample number was analysed under the same conditions.

96 samples including 4 standards (previously analysed by intact globin chain method) were enzymatically digested by trypsin and analysed using the MRM 2 method. The results of the T3 and T5 peptide ratios are in excellent correlation with the HPLC values as illustrated in Figure 2.21. T2 peptides showed a poor correlation, and the calculated values were not accurate as shown in Figure 2.22. This could be due to problems with chromatographic separation or with enzymatic digestion, as the trypsin used was not of high-quality sequencing grade. Use of highly purified trypsin would not suit the requirements of a cost effective screening method, and has not been tested here.



a) **b)**

Figure 2.21. Calculated percentages of δ-chain based on T3 (a) and T5 (b) peptide ratios for 96 samples

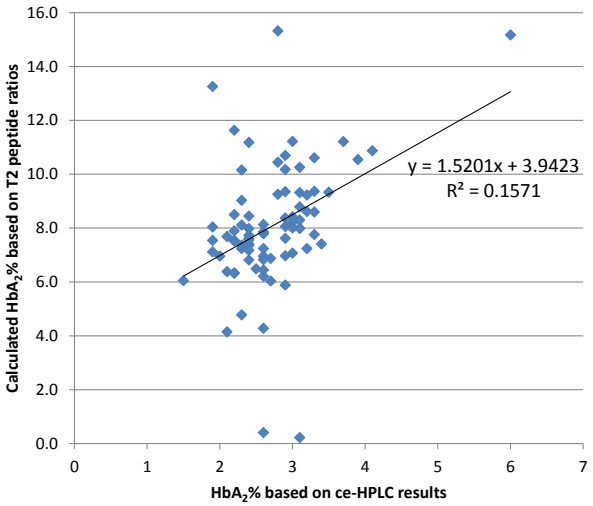


Figure 2.22. Calculated percentages of δ-chain based on T2 peptide ratios for 96 samples

Calculated R^2 values and their correlations for T3 and T5 peptides can be seen in Figure 2.21. Similarly to the analysed 20 samples, T3 peptides are lower and again the T5 peptides are closer to HPLC values; correlation coefficients in both cases are higher than 0.95 and the coefficient of determination (R^2) is slightly higher for T3 peptides (0.93 compared to the 0.9 value of T5 peptide ratio). Although the application of R and R^2 values may not be the most appropriate since there are two groups present for normal and for elevated levels, it is used here as an initial visualisation tool to describe the relationship of MS values to HPLC values. For

appropriate method comparison other statistical methods need to be used, for example the Bland-Altman plot. These measurements all support the suitability of the MRM 2 method, especially the T5 peptide pair, for delta chain level relative quantitation to beta-chain levels, and that measuring the peptide ratios can be used as an indication for HbA_2 levels, the potential biomarker for beta-thalassemia.

The use of T5 peptides ratios proved to be the most promising, which is important as the possibility of using this peptide pair for relative quantitation of HbA_2 has not been presented or published previously, and results show that the obtained values from the T5 peptides could be more reliable to estimate true values than the already patented use of peptide pairs (T2 and T3). Daniel *et al.* developed methods for HbA_2 quantitation on β(δ)T2, β(δ)T3 and βT13 (δT14) peptides (Daniel *et al.* 2007). The reliable observation of the βT13 and δT14 peptides was not possible during method development and the method developed using T5 peptides proved to be most useful for HbA_2 level determination.

Monitoring characteristic peptides for Hb variant detection

For clinically significant variants the peptides monitored provided definitive diagnosis and confirmation of the presence of the genetic mutation. The MRM-based detection of the peptides characteristic to HbS, HbE and HbC have been demonstrated.

During the analysis of the 92 clinical samples already presented this chapter, predetermined and targeted peptides for the clinically significant variants were also screened for inclusion in the MRM 2 method.

Nine samples had diagnosed abnormalities by a combination of clinical screening methods. Three of these samples were diagnosed to be carriers of HbS. Two of these were homozygous and one heterozygous. The method unambiguously identified all three samples with the sickle mutation. The homozygous samples were calculated to have 99.8% and 94.8%, while the heterozygous sample had 43.3% of sickle peptide level. These values were obtained by ratio calculation using the integrated areas of the combined MRM chromatogram peaks of sickle βT1 and normal βT1 peptides. The intact globin-chain analysis indicating the -30 Da mass shift and the confirmation of the presence of the sickle βT1 by the MRM 2 method is illustrated for the heterozygous sample in Figure 2.23.

It is observed that the percentage of the sickle β-chain is around 40% determined by the MRM method and also by the intact globin chain analysis in case of heterozygous sickle carriers. It would be expected to be equal to the normal β-chain levels. This phenomenon is explained by the observation that α-chains have a greater affinity for normal β-chains than for sickle β-chains, which leads to the formation of relatively more HbA than HbS (Matthay *et al.* 1979 , Bain 2006c).

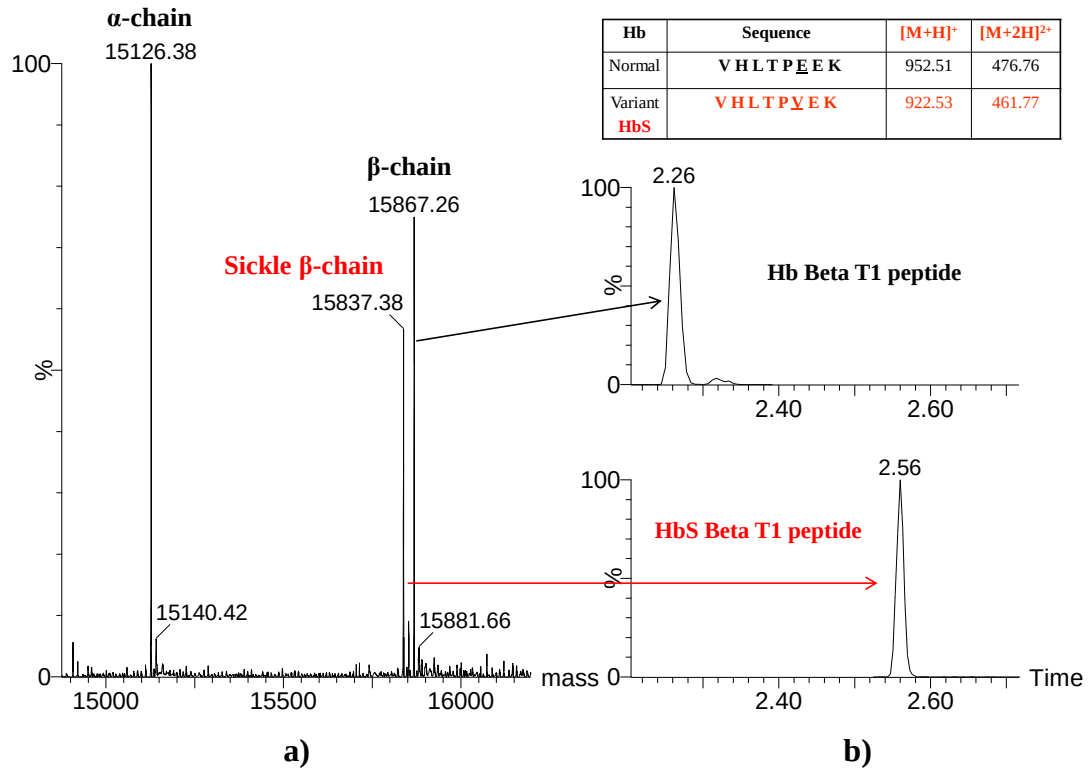


Figure 2.23. Intact globin-chain analysis (a) and MRM tryptic digest analysis (b) of a patient sample with heterozygote HbS

Presence of HbS mutation is confirmed with the detection of peak in the MRM in the corresponding MRM channel.

For the remaining six abnormal samples one HbC, one HbE and four HbD mutations had been detected in the hospital lab. The HbC and HbE mutations were confirmed by the MRM 2 method, after observing a slight mass shift during the intact globin-chain analysis. Illustration of the chromatographic peaks confirming the presence of these two variants is given in Figure 2.24.

Ratio calculations for HbC βT1 and normal βT1 peptide provided an average value of 62.7% for the proportion of HbC hemoglobin in the patients' blood. Performing similar calculations for the HbE detected variant, the average ratio calculations of the

novel HbE βT3 peptide peak areas to the normal βT3 peptide peak areas showed a low level for the variant of 7.1%. Despite the low level it was still reliably detected with the MRM 2 method.

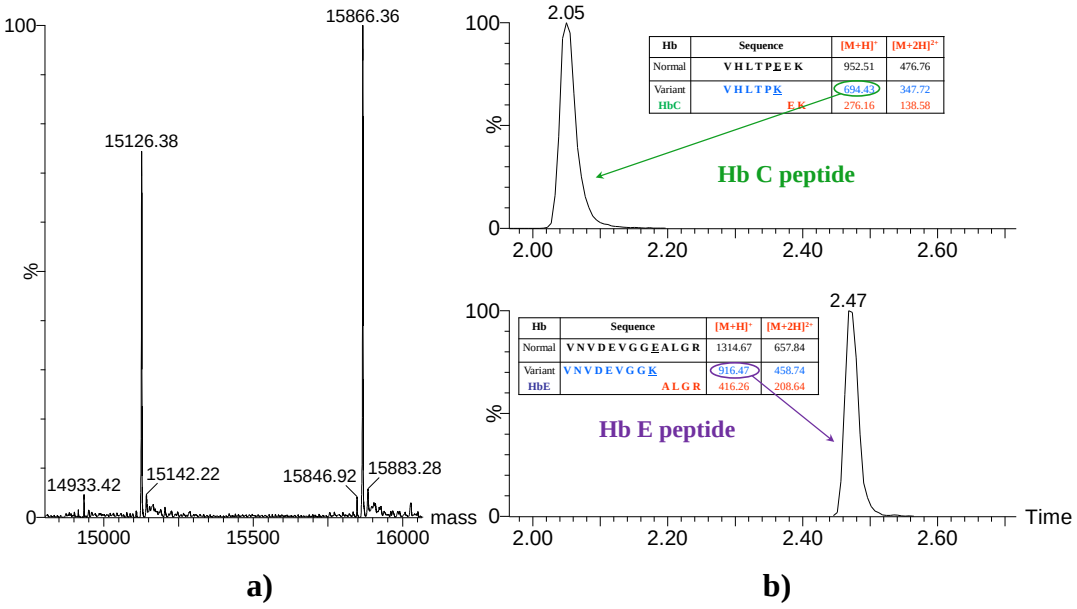


Figure 2.24. Intact globin-chain analysis (a) and MRM tryptic digest analysis (b) of a patient sample with HbC or HbE

Presence of HbC and HbE mutations are confirmed with the detection of peaks in the MRM in the corresponding MRM channels.

In case of HbD-Punjab carriers, two of the samples were identified to have α-chain variants during intact globin chain-analysis.

The HbD-Punjab mutation does not result in any new tryptic peptides with a mass difference higher than 1 Da to the wild-type peptide. The lack of a new peptide and significant mass difference makes it more challenging to confirm the presence of this mutation. Daniel *et al*. developed and used so called pseudo MRM channels to screen for the presence of the variant peptide mass (Daniel *et al.* 2007). This was not developed or included in the MRM 2 method here.

All four samples were enzymatically digested with trypsin and analysed in full scan mode using a high-resolution Q-TOF instrument (Waters Synapt HDMS). Confirmation of the presence of HbD-Punjab variant was obtained for only one of the samples, one was not confirmed and the other two were identified only with the α-chain variant. One sample did not show any additional *m/z* peaks in the tryptic peptide mixture as indication for any mutations present. This may have happened

because the variant was present at a low level, and information may have been lost during enzymatic cleavage. The results of the full scan high resolution analysis to confirm the presence of the HbD-Punjab mutation compared to a control healthy sample is illustrated in Figure 2.25.

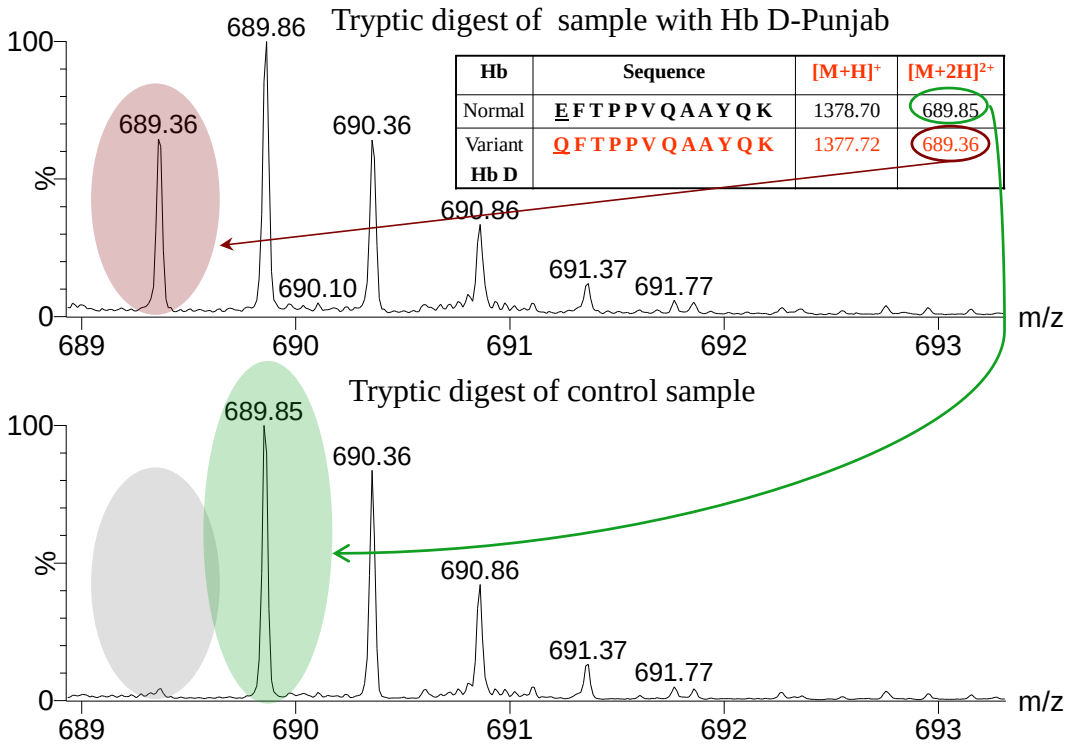


Figure 2.25. High resolution tryptic digest mass spectra of patient sample with HbD-Punjab and a control tryptic digest
The analysis of tryptic peptide mixture shows additional peak for the patient sample at *m/z* 689.36

As observed the doubly charged peptide ions for the -1 Da HbD-Punjab peptide are present at *m/z* 689.36 in the patient sample, and missing in the control sample. Only the *m/z* (689.85) for the doubly charged wild type peptide ion is present. Sequencing the variant peptide provided further confirmation of the mutation.

For the two remaining patient samples, it was unambiguously confirmed that those samples were not carriers of the HbD-Punjab mutation. The identification of the mutation can be found in section 2.3.3.2 below. The identified variant peptides were used to develop MRM transitions for this specific condition and included in the MRM 2 method to improve detection and guarantee the exclusion of the possibility for further misdiagnoses. The chromatographic peak is detected in the targeted MRM channel if the variant is present, as it is shown in Figure 2.26 b.

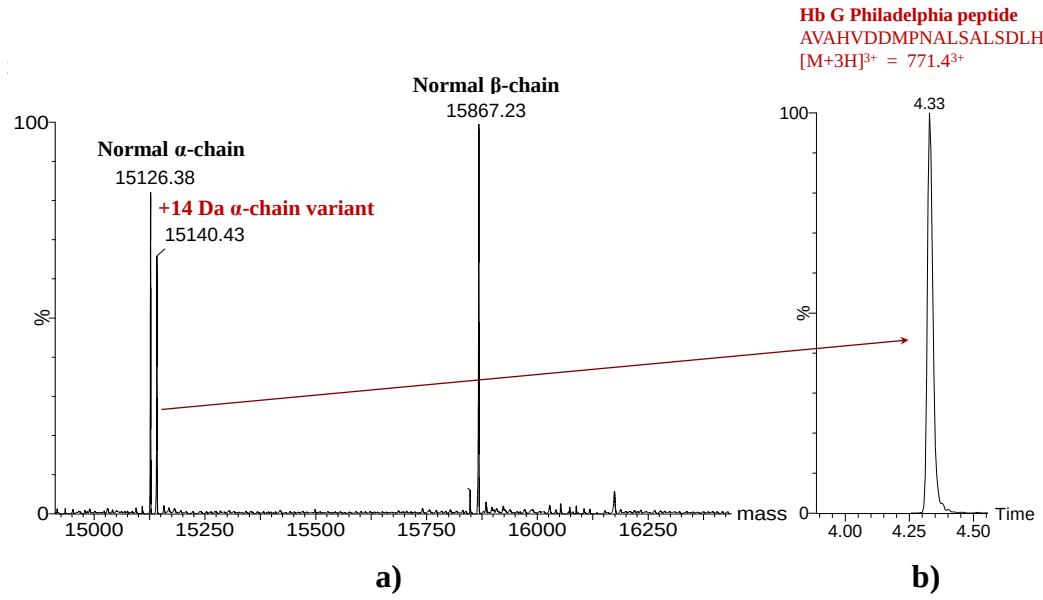


Figure 2.26. Intact globin-chain analysis (a) and MRM tryptic digest analysis (b) of a patient sample with HbG-Philadelphia
Chromatographic peak is present when a patient sample with the α-chain variant is analysed.

The MRM 2 method, including the MRM transitions for HbG-Philadelphia identified both samples with this mutation unambiguously.

The MRM 2 method can be further extended to any mutations identified, and can allow for the detection of those conditions with no requirement for additional sample preparation or analysis.

The developed method proved to be a definitive diagnosis tool for the targeted mutations HbS, HbC, HbE and HbG-Philadelphia.

2.3.3 Identification of Hemoglobin variants

Clinical samples which showed abnormality during general screening in the hospital laboratory at UHCW were analysed using the mass spectrometry-based methods developed here. To confirm the presence of a hemoglobin variant, and if possible, identify the mutation present in the sample, measurements of intact globin chains and digested peptide mixture were performed using the Waters Xevo TQ triple quadrupole MS instrument as a first step. Further measurements were carried out using the Waters Synapt HDMS instrument, when higher resolution was required to achieve more detailed information.

Following the protocol for hemoglobin variant identification published by Wild *et al*. (Wild *et al.* 2001), the presence of a sickle cell mutation was confirmed. A -14 Da variant of the alpha chain (Hb-J-Broussais, which occurred in two different samples), a +44 Da alpha chain variant (Hb-J-Nyanza) and a -14 Da variant of beta chain (Hb-Andrew Minneapolis) were also identified in collaboration with the Haematology Department of UHCW. Results of the identification steps of the sample with HbJ-Nyanza are presented below.

During method development and optimisation two clinical blood samples diagnosed to have HbD-Punjab mutation based on current clinical laboratory techniques were observed to have a +14 Da α-chain variant, HbG-Philadelphia. Identification of this variant is also shown, and the variant peptide has been included in the MRM 2 method for future screening purposes.

2.3.3.1 Case study A: Identification of the +44 Da alpha chain variant – Hb J-Nyanza

The patient blood sample was diluted and the denatured globin chain mixture analysed using the Xevo TQ mass spectrometer. Figure 2.27 shows the denatured intact globin chain mass spectrum, which contains peaks at the mass-to-charge ratio values corresponding to the multiply charged globin chain ions. A third ion series adjacent to the expected normal α- and β-chain multiply charged ionseries is also observed. The spectrum was deconvoluted onto a true mass scale, shown in Figure 2.28, which highlights the presence of a +44 Da alpha chain variant. A +44 Da shift

considering single point mutations of the globin coding genes could suggest an amino acid substitutin of Ala→Asp or Cys→Phe.

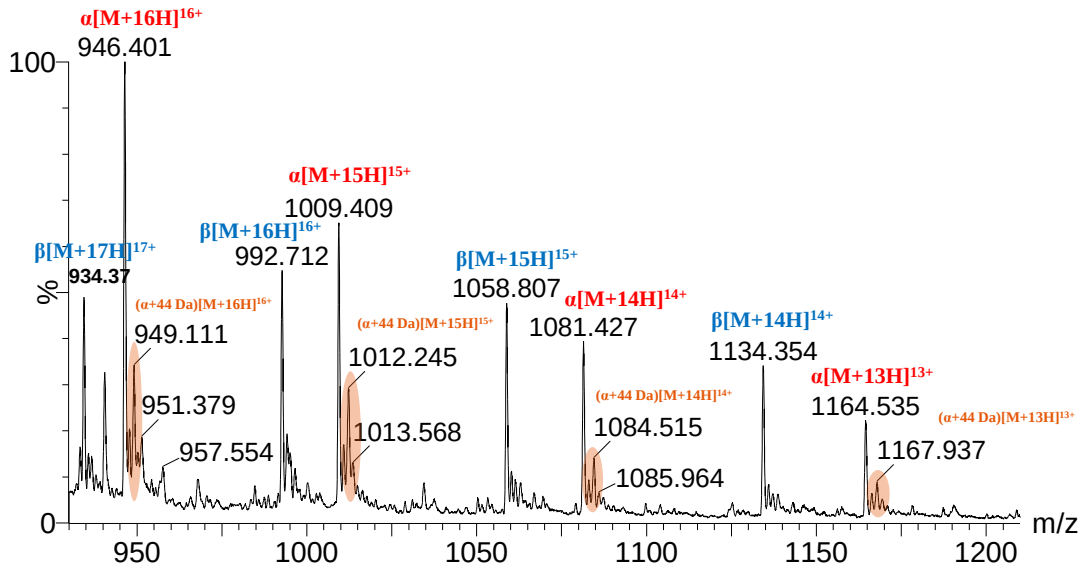


Figure 2.27. Mass spectrum of the intact blood sample showing +44 Da variant. Acquired mass spectrum of anonymised blood sample on the Xevo TQ MS instrument, scan time 1s, scanned range *m/z* 930-1210 and data were collected for 3 minutes in MCA (Multi Channel Analysis) acquisition mode

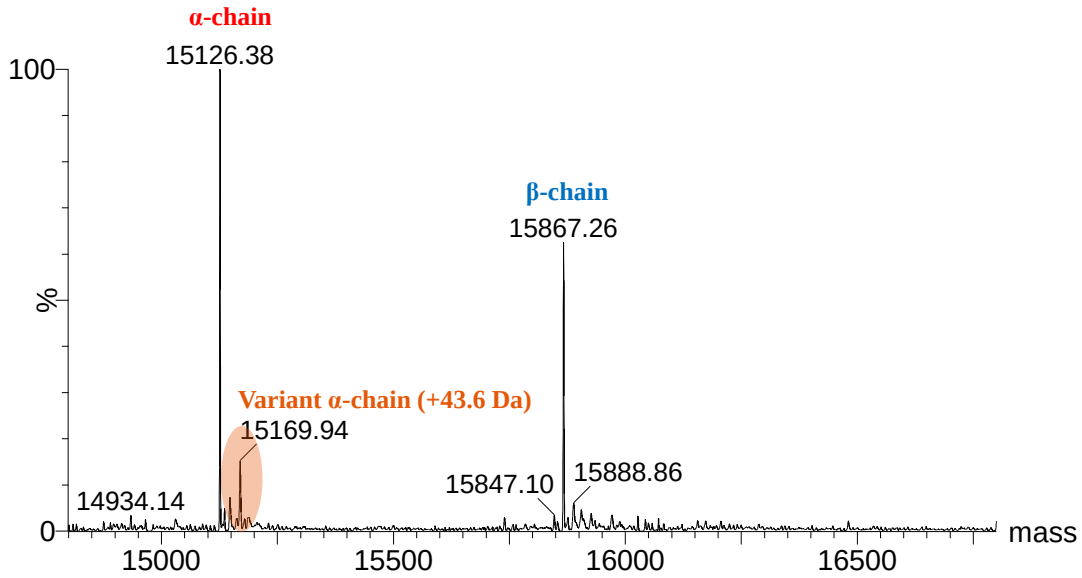


Figure 2.28. Deconvoluted mass spectrum of sample with +44 Da α-chain variant

To determine which mutation had occurred and at which position of the globin chain, tryptic digestion was performed on both the sample, and a control blood sample. The

peptide mixtures were analysed using a high resolution mass spectrometer and the resulting MS spectra were compared as shown in Figure 2.29. Additional peaks were observed in the variant sample and due to the high resolution measurement, the charge state of the ion could be determined, which helps in identifying the variant tryptic peptide. The mutation occurred affecting the αT4 peptide.

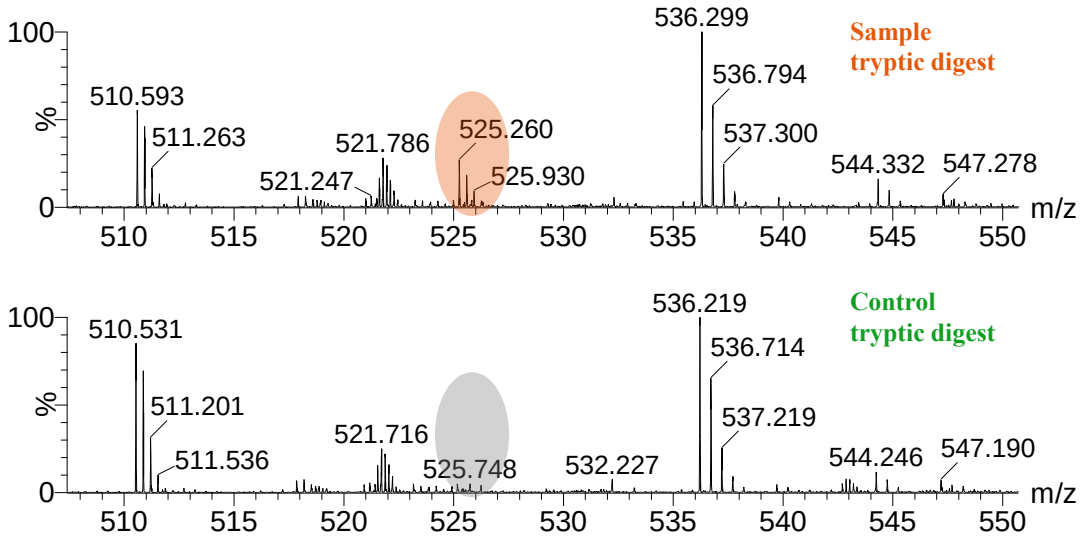


Figure 2.29. Spectra of the tryptic digested +44 Da variant sample and a control blood sample

Additional peaks are observed at *m/z* 525.3 (orange shade), which suggests a triply charged ion, are not present in the spectrum of the control sample (grey shade).

Tandem MS analysis was performed to determine the amino acid sequence of the variant peptide, and to identify the exact position of the mutation (Figure 2.30). Tandem MS experiments were also performed on the wild-type peptide to confirm the mutation site. The b4 fragment ions are present with the same *m/z* for in both MS/MS spectra, and mass shift only observed affecting the b5 fragment ion and onwards, the position of the substitution is determined. The peptide comprising α-chain sequence from 17 to 31 amino acids, and the fifth position of this peptide is taken up by an Ala residue. These results showed that the mutation occurred at position **α21** as a result of an Ala→Asp substitution. Using the *HbVar* database (HbVar) the mutation was identified as **Hb J-Nyanza .**

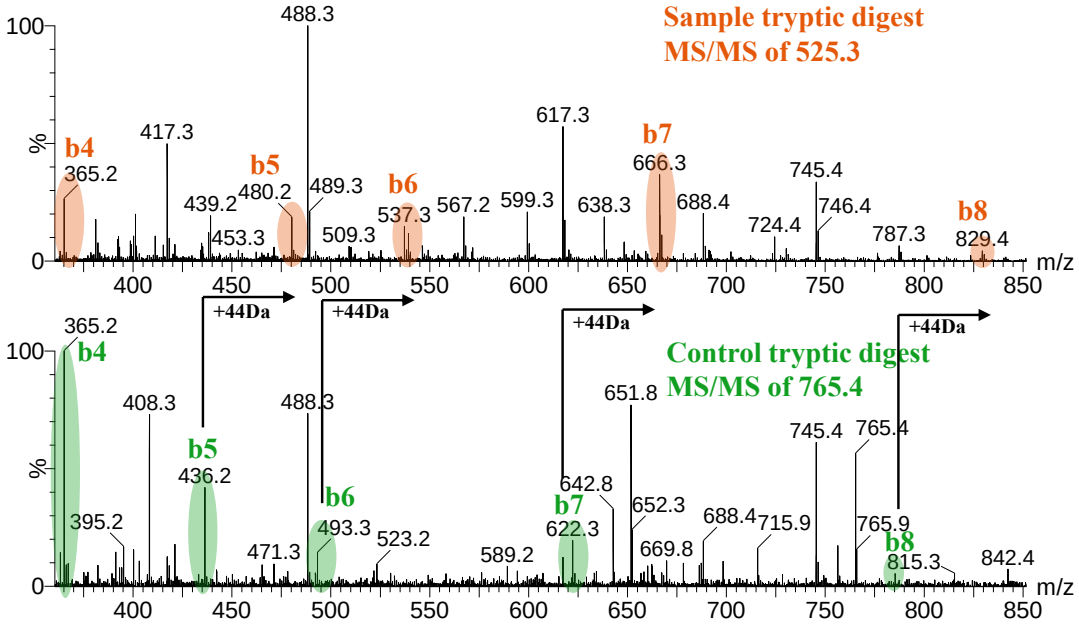


Figure 2.30. Product ion spectra of variant and control peptides
MS/MS analysis of the additional peak at *m/z* 525.3 (3+) corresponding to the variant peptide ion and of the wild-type peptide ion at *m/z* 765.4 (2+) are compared and mass shifts of fragment ions are investigated. Mutation site is visible at the b5 fragment ion

2.3.3.1 Case study B: Identification of the +14 Da alpha chain variant – HbG Philadelphia

A sample from a patient diagnosed to be a carrier of HbD-Punjab mutation has been analysed. Intact globin chain analysis indicated the presence of a +14 Da α-chain variant, while the mass accuracy of the β-chain did not indicate the presence of any -1 Da variants affecting that globin chain. The deconvoluted mass spectrum of the patient sample is shown in Figure 2.31. A +14 Da mass shift could suggest the following substitutions: Asn → Lys, Asp → Glu, Gly → Ala, Ser → Thr and Val → Ile/Leu. As a follow up step, the patient sample and a healthy control sample have been enzymatically digested with trypsin and analysed using the Synapt HDMS high resolution mass spectrometer in order to identify the mutation. The comparison of the resulted spectra revealed the presence of additional peaks at *m/z* 717.4 and 771.4 as illustrated in Figure 2.32 a.

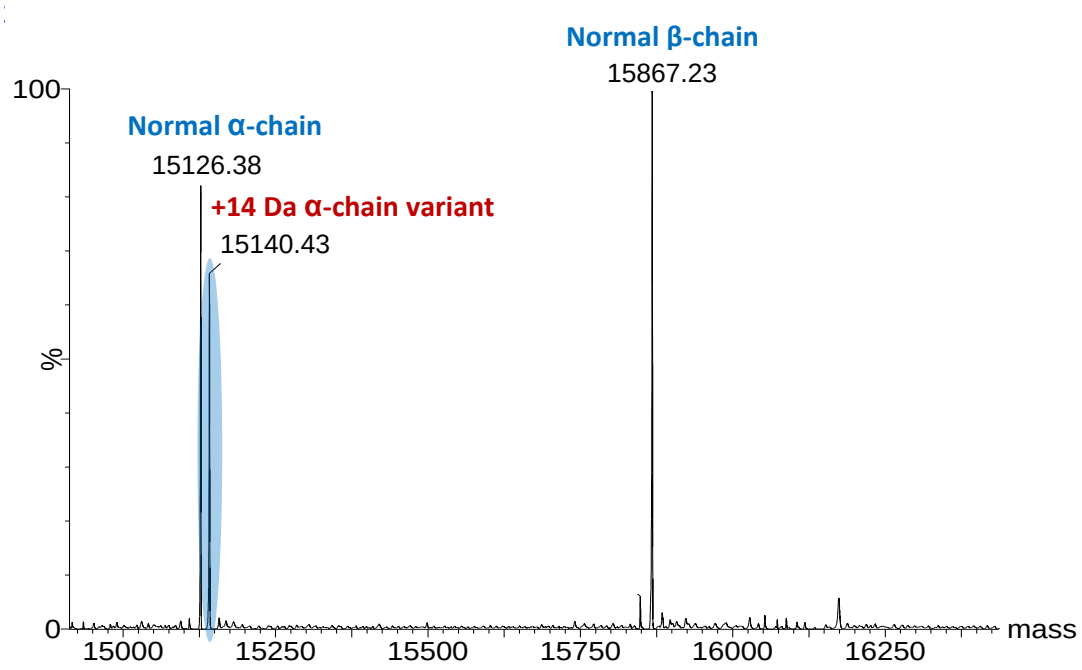


Figure 2.31. Intact mass spectrum of sample with +14 Da α-chain variant

The peak at *m/z* 717.4 belongs to a singly charged ion, while the peak at *m/z* 771.4 belongs to a triply charged ion. The appearance of two additional peaks and the listed possibilities for the amino acid substitutions support the hypothesis of a Asn→Lys substitution. Following this hypothesis, the patient and control sample have been enzymatically digested by α-chymotrypsin, which cleaves the amino acid sequence after Tyr, Trp and Phe residues, whereas trypsin hydrolyses proteins after Arg and Lys residues.

The results of comparing the mass spectra of the patient and control α-chymotryptic peptide mixtures are shown in Figure 2.32 b. This comparison exposed an additional peptide at *m/z* 741.5, as a doubly charged ion. The proposed mutant peptide mass is 1481 Da, 14 Da higher than the expected mass for the αY11 peptide. This peptide covers the α-chain sequence from residues 67 to 80. This narrows down the tryptic peptides possibly affected to αT9, with the amino acid sequence covering the region for 62 to 90 amino acids. This peptide comprises of two Asn amino acids. Performing calculations for expected peptide masses with both possible substitutions revealed that the mutation affected the first Asn residue in the tryptic peptide resulting in two smaller tryptic peptides than the original αT9. The outcome of this calculation is illustrated in Figure 2.33.

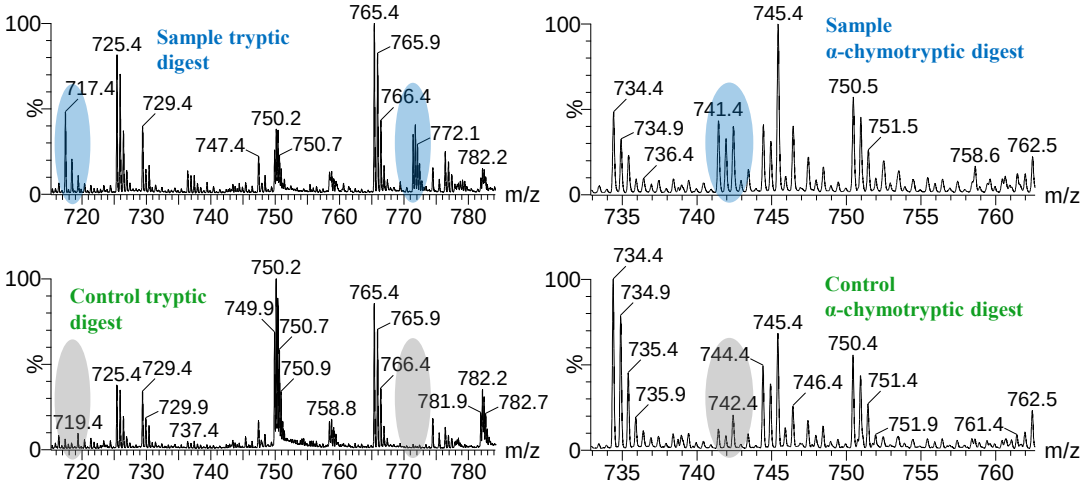


Figure 2.32. a) MS analysis of tryptic digest mixture of sample with variant and control sample b) MS analysis of α-chymotryptic digest mixtures of variant and control sample

Blue circles show the position of additional peaks, grey circles show same location in control sample

The first, smaller unit was expected to have a singly charged ion peak at *m/z* 717.4, while the second larger unit results in a mass of 2311.2 Da and an expected triply charged ion at *m/z* 771.4. These results confirmed that mutation of Asn→Lys substitution had taken place at position 68 of the α-chain. Searching the *HbVar* database (HbVar) the mutation was identified as Hb G-Philadelphia, a fairly common α-chain variant, which can be mistaken for HbD-Punjab mutation based on ce-HPLC and electrophoretic methods (Bain 2006).

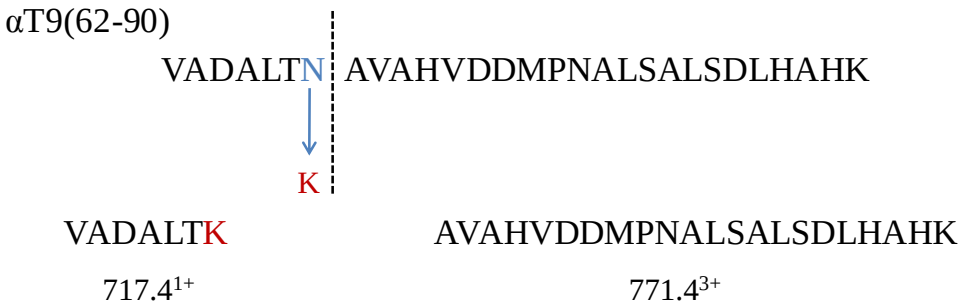


Figure 2.33. αT9 peptide and the theoretical new peptides resulted by the amino acid substitution of Asn→Lys

Additional identified Hb variants

- **Hb-J-Broussais**: The analysis of the intact globin chains showed the presence of a -14 Da alpha chain variant, which could suggest the following substitutions: Lys→Asn, Glu→Asp, Ala→Gly, Thr→Ser, Ile/Leu→Val. After tryptic digestion an additional peak was observed in the spectrum of the tryptic peptide mixture at *m/z* 651.15. This was proposed to be the +5 charged ion for the variant peptide of α(T9-10) with a missed cleavage site because of Lys→Asn substitution. Tandem MS measurement confirmed this hypothesis, and determined the position of the mutation. The mutation is **α90 (Lys→Asn).**
- **Hb-Andrew Minneapolis**: The analysis of the intact globin chains showed the presence of a -14 Da beta chain variant, which could imply the substitutions detailed for the previous variant. The additional peak in the spectrum of the tryptic peptide mixture was at *m/z* 718.4, corresponding to a doubly charged $\beta T(14\text{-}15)^{2+}$ variant ion, which suggests a missed cleavage site because of Lys→Asn substitution between the T14 and T15 peptides. Tandem MS measurement confirmed this hypothesis, and determined the position of the mutation. The mutation is **β144 (Lys→Asn).**

  These cases confirm that the established protocol for hemoglobin variant identification can be carried out by the use of Waters Xevo TQ, in combination with high resolution mass measurements for the analysis of tryptic peptide mixtures.

2.4 Conclusion

Mass spectrometry-based methods have been developed using Waters Acquity TQD and Xevo TQ instruments (Waters Corporation, Manchester, UK). The developed methods have been optimised on reference standards and tested on patient samples in collaboration with the Haematology Department of UHCW.

Initial method comparison has been carried out in preparation for a clinical trial. Methods for sample preparation and mass spectrometric analysis in combination with data processing methods have been tested to facilitate a high-throughput application for future sample analysis.

Intact globin chain analysis for the quantitation of HbA_2 level determination proved to give good correlation to HPLC values. Elevated HbA_2 levels are unambiguously detected. Although this quantitation method for δ-chain levels relies on previously published principles, the quantitation method itself has not yet been published. The same method is capable of the quantitation of glycated hemoglobin chains. The glycated levels of the different globin chains are most of the time in good correlation, but the glycated β-chain level may be more informative for future diagnostic purposes.

Mass accuracy measurements for the β-chain mass provided a mass window of ±0.2 Da, Since α-chain mass errors used to adjust the β-chain masses are generally higher than β-chain mass errors, all the corrected mass error values for the β-chain are below zero, and therefore gave a systematic error. Clinically significant -1 Da variant samples with a higher mass error could be still detected based on the assumption of causing a mass shift if they are present at levels higher than 20%.

MRM 1 method was able to indicate elevated delta-chain levels, but based on its low reproducibility rate, it needed further development. The new method developed on the Waters Xevo TQ system proved to be more reliable for quantitation and also capable of detecting clinically significant variants such as HbS, HbC and HbE.

A novel peptide ratio for the T5 peptides have been included in the developed MRM method. The clinically significant variants were unambiguously identified based on

the developed MRM transitions, and samples from homozygous and heterozygous sickle cell carrier patients were easily distinguished. The peptide ratios for T3 and T5 peptides gave excellent correlation to HPLC values, but problems with enzymatic cleavage was indicated based on the results of a larger sample set comprising 94 samples. T5 peptide ratios resulted in values closest to the expected true HbA_2 levels determined by ce-HPLC, and it is suggested to use these peptides for further developments and screening for elevated HbA_2 levels.

The published protocol by (Wild *et al.* 2001) utilising intact globin chain analysis followed by peptide analysis, successfully identified rare hemoglobin variants.

Two different MS-based methods have been developed using the same platform and in combination in a high-throughput screening method that may be able to unambiguously identify the clinically significant and occasional rare variants. The developed MRM methods did not target the detection of HbD-Punjab variants. In case a sample shows abnormality indicating the presence of a -1 Da variant on the intact level, and is not diagnosed to have either HbC, HbE or potentially HbO-Arab mutations based on MRM results, high resolution mass analysis of the tryptically digested peptide mixture would be the only solution for a confirmative diagnosis.

At this stage both developed methods need a trained person to interpret the data; whilst the available software tools are helpful, the data processing is not fully automated. Sample introduction is automated for both methods, which are able to support high-throughput screening applications with overnight measurements.

Mass spectrometry methods revealed a potential problem with current clinical methods, where two samples classed as HbD-Punjab by the hospital lab were identified as carriers of a fairly common α-chain variant HbG-Philadelphia.

The combined application of developed methods on the same platform are hoped to provide a competitive diagnostic technique as a potential first line screening tool in future clinical lab applications.

CHAPTER 3

APPLICATION OF MASS SPECTROMETRY-BASED APPROACHES IN A CLINICAL TRIAL

3.1 Introduction

The development of a comprehensive, high-throughput screening method for hemoglobinopathies is an increasing demand. The current methods used for routine screening are able to confirm the presence of clinically significant variants, but are slow, laborious and non-specific, and incapable of characterising novel variants.

As discussed in previous Chapters, there are a number of different mass spectrometry-based techniques which have been implemented in the research field of hemoglobinopathy diagnosis. There is widely reported literature supporting the view that mass spectrometry could become the sole technique in definitive hemoglobin variant detection and identification in the clinical lab. The requirement is to develop a single platform using mass spectrometry which could replace current methods, more precisely the need for the employment of combined methods. Several mass spectrometry-based methods have been tested on larger numbers of clinical samples previously. A method capable for population screening has to be able to compete with current methods regarding some key points. Previous evaluations of mass spectrometry-based methods on relatively large sample sets has given the opportunity to discuss what key features needed to be considered in order to implement the most appropriate method for hospital use. These features include cost, speed, sensitivity, specificity, potential for automation allowing high-throughput applications and prospective additional diagnostic information. The method needs to be robust for a clinical lab environment involved in population screening. For population screening, the confident diagnosis of clinically significant variants is the most important criteria. Mass spectrometers are sensitive but complex techniques, and to demonstrate their robustness and stability for a screening application a higher number of sample amounts need to be analysed with a high reproducibility rate. Previously developed mass spectrometry-based methods have tested sample sets ranging from hundreds to thousands. The most widely used protocol described by Wild et al. has been tested on over 250 samples to prove its suitability for diagnostic purposes (Wild *et al.* 2001). A smaller set of samples with 147 dried blood spot cards have been also analysed by the same group and showed potential for newborn screening (Wild *et al.* 2004).

The proposed MRM method introduced by Daniel *et al*. has been presented with the analysis of 200 samples for the detection of clinically significant hemoglobinopathies (Daniel *et al.* 2005). The feasibility of the extended method for HbA_2 quantitation was tested on an additional 319 samples (with 105 of these having increased HbA_2 levels, 43 delta-chain variants and 8 Hb Lepore carriers) with the application of a cut-off value for normal levels of 3.4% (Daniel *et al.* 2007). Following this the analysis of 40,000 samples have been presented on the Annual Meeting of the American Society of Hematology in 2008 with the same MRM-based approach for both clinically significant variant detection and HbA_2 quantitation (Daniel *et al.* 2008). This method has been demonstrated to be capable of high-throughput and fast screening, but only gives definitive answers for the known variants included in the set-up. There have been no observed discrepancies between the tandem mass spectrometry method and the routine screening techniques. 5,000 samples have also been analysed with data-dependent acquisition for peptide sequencing experiments suitable for the identification of rare hemoglobin variants. This has showed a further applicability beyond targeted screening methods for clinically significant variants, by addressing the above mentioned issue about the approach not being able to detect unknown variants. The presented method is fast and the acquisition time is only 60 seconds which is important in population screening applications.

At the same time another group presented results with a similar approach. Boemer *et al.* also utilised an MRM-based method and performed a comparative study with IEF results on 2082 samples from newborns with no discrepant results to clinical laboratory techniques observed (Boemer *et al.* 2008). The described method did not target and therefore did not detect individuals with Hb-Barts, α-chain variants or any other unknown variant. In further applications over a three year period the total number of samples analysed were 43,736 (Boemer *et al.* 2011). Screening has not been considered for some clinically significant variants such as HbD-Punjab, HbO-Arab and Hb Lepore, based on the low prevalence of these variants in the Belgian population. Although it has been shown that the tandem mass spectrometry method allows the detection of major hemoglobin disorders with the exception of HbH disease, it has also highlighted the fact that misinterpretation can happen in case of rare β-globin chain variants. It has been suggested that other methods may be necessary as second-line tests to discriminate between, for example HbSS and

HbS/β-thalassemia, since the β-chain produced at a normal level in newborns is still relatively low. The major advantage of the MRM-based method developed here is that for different targeted peptides four MRM transitions are used, while Daniel *et al.* has only used one transition per peptide. The additional confidence that would be obtained from the higher number of transitions means that the total acquisition time is longer, but is still only 150 seconds. The specificity required for diagnostic purposes like hemoglobinopathy detection makes it necessary to obtain the highest specificity possible based on multiple MRM transitions to reduce the rate of false positives. Daniel et al. did not report the occurrence of false positive identifications with their method and this remains an open question (Boemer *et al.* 2009).

Kleinert *et al.* reported the analysis of 2145 blood samples (Kleinert *et al.* 2008). They analysed diluted lysed blood samples with either ESI-MS or MALDI-TOF MS, and also utilised FT-ICR measurements in the case of -1 Da variants. This study indicated that MS can be used for the unambiguous identification of Hb variants but was not focused on the development of fast and automated screening techniques. They demonstrated that the MS approach detected four chromatographically silent mutations, and one sample, which had been incorrectly identified by chromatographic methods as β-thalassemia, was diagnosed as a mutation referred to as Hb Malay. Although this study stated that variants with mass shifts lower than 6 Da could not be detected (eg. clinically significant variants like HbC, HbD-Punjab, HbO-Arab, HbE), the feasibility of MS for detection of Hb variants was demonstrated and they proposed the use of MS as a supplementary method. Considering the restriction imposed for detectable variant chains to differ by more than 6 Da, 92% of currently undetected variants are still able be detected with MS methods (Kleinert *et al.* 2008). The restriction for the 6 Da difference is not a valid limiting factor today. The important outcome of their study is that MS is capable of the detection of variants which do not show any significant charge-difference which is necessary for detection by electrophoretic methods or alterations in hydrophobicty needed to be observable by HPLC techniques.

All here mentioned studies were performed on a larger set of neonatal samples, and none of them focused on antenatal screening , which is in terms of preventive health

care also essential. Also no studies were conducted on a larger sample set applying top-down protein analytical methods.

It needs to be emphasized that structural hemoglobin variants require detailed and confirmative analysis. The identification of variants based only on mass shifts observed at the intact protein level is already proved to be not confirmative enough for the diagnosis of known mutations. The -30 Da mass shift from the sickle cell mutation could also be a result of other amino acid substitutions, at the same or different position on the β-globin chain. Excluding β^{SICKLE}, there are 19 possible β-chain variants with -30 Da mass differences (HbVar). Detection of sickle cell mutation based only on the mass difference at the intact level can therefore lead to false positives (Wild *et al.* 2004) (HbVar). The possible amino acid substitutions comprise Thr→Arg, Ser→Gly, Met→Thr and Glu→Val. Although these theoretically possible mutations are much less frequent than the sickle cell mutation, their possibility must not be overlooked.
There are also a number of outcomes that could result in -1 Da variants. There are 25 possible variants with the same mass difference. Four of the clinically significant variants result in a -1 Da mass shift affecting the β-chain. If the presence of a -1 Da mass shift is observed, this is not conclusive enough to reveal the exact nature of the mutation and the affected amino acid along the sequence to distinguish between the four important variants. Amino acid substitutions based on single point mutation, which result in a -1 Da mass difference between the mutant and the normal chain comprise Asp→Asn, Glu→Gln, Asn→Ile and Glu→Lys (Wild *et al.* 2004).
Applied diagnosis methods need to be able to distinguish between the mutations at the amino acid level to give conclusive diagnosis. The previously described method of targeted peptide analysis may be one solution, alternatively a method for the direct fragmentation of the intact variant globin chain could be employed.

The two main workflows for protein identification which includes mass spectrometry are: the *bottom-up* strategy involving peptide tandem mass spectrometry sequencing while the other is the *top-down* strategy involving the direct fragmentation of intact protein molecules. In a bottom-up experiment the proteins are enzymatically cleaved, most often by trypsin, to produce peptides with length of 10-20 amino acids, but it is inevitable that information is lost during the process of transformation of the protein into peptides, and many of the peptides may not be observed upon mass

spectrometric analysis (Reid and McLuckey 2002). In the top-down approach multiply charged ions of the intact proteins are broken using a number of dissociation techniques (Cui *et al.* 2011). There are different strategies within the top-down class presented by McLafferty and co-workers (Kelleher *et al.* 1999). The growing interest in developing top-down analysis methods is supported by the fact that substantial structural information can be obtained without time-consuming sample preparation.

For hemoglobinopathy analysis the application of top-down strategies has not been frequently reported compared to the use of bottom-up strategies, and no clinical trial has been conducted. The top-down strategy has not become as popular, presumably because of the more complex spectra due to the number of ions, and the difficulty in interpretation which can be caused by overlapping peaks of multiply charged fragment ions (Zhou *et al.* 2012). There is an increasing demand for the development of bioinformatics softwares and deconvolution algorithms. A key step is to assign product ion identities, which needs either mass spectrometers with high resolving power or ion charge state manipulation (Reid and McLuckey 2002).

The possibility of minimising analysis time excluding sample preparation steps such as enzymatic cleavage with the analysis of whole protein with tandem mass spectrometry may be attractive for real clinical applications, hence the effort for the development of top-down techniques. Shackleton *et al.* expressed the desirability of developing top-down methods to be applied in the hemoglobinopathy field after early studies using FT-ICR mass spectrometers (Shackleton and Witkowska 1996).

In the early 1990s after the introduction of ESI with tandem mass spectrometry becoming a routine tool in the analysis of intact proteins (Loo *et al.* 1990), research groups in the field of hemoglobinopathy analysis started looking at the option of tandem mass spectrometry on intact globin-chains as a potential alternative to analysing peptides. The CID process produces fragment ions characteristic of the amino acid sequence. Light-Wahl *et al.* showed results for tandem mass spectrometry of the intact hemoglobin β-chain by application of CID (Light-Wahl *et al.* 1993). It was established that β-globin chain variants can be differentiated based on the observed mass shifts of the fragment ions to reveal in which part of the molecule the mutation has taken place. The sequence information obtained covered approximately 29% of the molecule.

Witkowska *et al.* proposed that intact protein ESI-tandem mass spectrometry could be used as a sole technique to identify hemoglobin variants (Witkowska *et al.* 1995). An α-chain variant has been identified without any separation prior to mass spectrometric analysis. The investigation of major fragment ions narrowed down the possible mutation site and further examination of low abundance fragment ions allowed the unequivocal identification of the mutation.

A more comprehensive coverage of the sequence was obtained using cyanogen bromide (CNBr) digestion of whole blood in addition to acquiring a low-energy CID mass spectrum of the complete protein (Summerfield *et al.* 1996).

Charge-stage dependent fragmentation was investigated by ion trap collisional activation of β-chain ions from 2+ charges to 17+ charges in an ion trap analyser (Schaaff *et al.* 2000). A similar study has been performed on the α-chain charge states ranging from 4+ to 20+. A total sequence coverage of 62% was obtained. MS^3 involving subsequent activation may also be considered to achieve higher sequence coverage (Mekecha *et al.* 2006).

Rai *et al.* have used ESI tandem MS for the rapid identification of β-chain variants and the low proportion Hb Lepore-Boston-Washington (Rai *et al.* 2004, Rai *et al.* 2002). The obtained sequence coverage for the β-chain was approximately 50%, but the inner region was still not covered, which made peptide analysis necessary in the case of variants affecting that region (Rai *et al.* 2002).

The application of CID only is clearly not sufficient since it does not cleave all the bonds required to give the full information desirable on the primary structure of a polypeptide or protein (Zubarev *et al.* 2008).

With the introduction of Electron Capture Dissociation on FT-ICR instruments an alternative fragmentation method has become available (Zubarev *et al.* 1998). The ECD process results in different fragmentation routes and product ions providing a complementary level of information accessible in top down strategies. With the development of the Electron Transfer Dissociation (ETD) technique on lower cost instruments like ion traps (Syka *et al.* 2004) the non-CID top-down route has become more widely accessible. Fragmentation occurs predominantly along the peptide backbone, which results in a predictable product ion spectra and therefore easier interpretation. Formation of c- and z-type product ions containing the N and C termini of Pro are however not observed in either the ECD or ETD spectra (Syka *et al.* 2004). The ECD/ETD activation techniques are more suitable for fragmentation

of larger polypeptides and intact proteins, when compared to CID, based on the narrower dynamic range in fragmentation when higher molecular masses are involved (Zubarev *et al.* 2008).

Ion/ion reactions such as proton transfer reactions (PTR) have been introduced by McLuckey and co-workers to simplify the product ion spectra of multiply charged ions to singly charged ions, which may be more easily interpreted (McLuckey and Stephenson 1998, Reid and McLuckey 2002).

ETD may prove to be a better technique for the rapid sequencing of intact proteins simultaneously from both N and C termini (Coon *et al.* 2005). Multiply charged protein ions are allowed to react with fluoranthene anions. Electrons are transferred to the multiply charged ions which promotes the random dissociation of the N-Cα bonds. The charged fragment ions are then involved in a second ion/ion reaction where they are deprotonated. The reaction is controlled by both the amount of reagent and the reaction time allowed for ETD and PTR. ETD reaction time is kept low to avoid multiple electron transfer events, while the PTR period is slightly longer to allow enough time and cycles for the fragment ions to be simplified to predominantly 1+ charge (Coon *et al.* 2005). The PTR is necessary because direct fragmentation of intact proteins leads to spectra with highly charged c- and z-type fragment ions, and performing this type of experiment on a bench-top ion trap mass spectrometer can make interpretation challenging, since the ion trap has a limited *m/z* resolution. The resulting singly and doubly charged ions may be used to read the sequence of 15-40 amino acid long segments on both termini of the protein. With the application of this method the analysis of mutant protein forms revealed amino acid changes. The obtained sequences were compared against database entries to reveal up to four amino acid changes. In a study of top-down fragmentation of ubiquitin (MW= ~8.5 kDa) c- and z-type ions, 17 amino acids of the N-terminus and 16 amino acids of the C-terminus have been identified (Mikesh *et al.* 2006).

While CID fragmentation depends significantly on the presence and order of amino acids, ETD has the advantage of producing homologous series of c- and z-type ion series which can allow direct reading of the protein's amino termini. This predictability makes this technique applicable for the development of an automated de novo sequencing method, which would not need the usual database searching currently necessary for processing data obtained by CID fragmentation.

Théberge *et al.* utilised different fragmentation techniques on a LTQ-Orbitrap to develop a method which could be used for characterisation of variant proteins with stepwise MS/MS analysis (Théberge *et al.* 2011a). They first recorded a mass profile to detect any mass shifts revealing the presence of a variant, then performed nozzle skimmer dissociation (NSD), and identified the fragment ion(s) containing the amino acid substitution. The fragment ion was then selected for MS/MS analysis in the ion trap to obtain sequence information allowing the identification of the variant. This stepwise approach has been shown to be capable of identification of the sickle cell β-globin chain, and also of the characterization of post translational modifications like S-glutathionyation. Although high sequence coverage was achieved, the method does not provide specific information for every amino acid position.
It has been shown by the same group that additional information can be gained by the use of ETD and higher energy C-trap dissociation (HCD) activation techniques, and also by the combined use of CID and ETD fragmentation techniques (Théberge *et al.* 2011b). They demonstrated that ETD was useful for sequencing the core region which can be challenging using CID. The choice of precursor ion charge state can have an influence on the obtained sequence coverage as expected. For the β-chain, the 18+ charge state was chosen to perform tandem MS experiments. β-chain coverage achieved with the use of multiple activation techniques was 85%.

In recent years other groups have recognized the need for alternative methods and started applying different top-down strategies for the identification of hemoglobin variants. An interesting application is to perform direct sequence analysis of samples obtained from dried blood spots from newborns by liquid extraction surface sampling using NanoMate technology (Edwards *et al.* 2011). This recently presented method employed CID tandem mass spectrometry using an Orbitrap instrument for unambiguous identification of common variants HbS, HbC and HbD-Punjab, the latter two differing from normal β-chain by only 1 Da. The β-chain $[M+15H]^{15+}$ ions were isolated and subsequently fragmented with a sequence coverage of 32-38% obtained. The HbD variant proved to be more challenging because of overlapping isotope peaks. Following this study Cooper and co-workers have shown results obtained with the same sample introduction and the use of CID and ETD to produce variant-specific fragment ions (Edwards *et al.* 2012). Six abnormal neonatal samples were analysed, which could not be identified using current screening techniques. It

was shown that CID, combined with ETD is capable of identifying unknown hemoglobin variants. Sequence coverage obtained for the normal β-chain was: 68% with CID only, 44% with ETD only and the combined sequence coverage by CID+ETD was 81%. It was lower for a variant β-chain Hb Headington: 46%, 34% and 63% respectively.

As a potential screening tool, top-down ETD application for hemoglobin disorders using a quadrupole ion trap instrument was described recently (Graca *et al.* 2012). The ion trap being a lower cost instrument is more likely to be considered as a potential clinical laboratory tool than a significantly more expensive and complex Orbitrap mass spectrometer. A pseudo-SRM ETD assay has been used for the detection of HbS and HbC, monitoring three fragment ions for all analysed globin chains.

The disadvantages of the presented approach were the lengthy sample preparation and the nano-LC separation, which makes this method less robust for real clinical applicability, where short run time and minimal sample preparation is highly desirable (Graca *et al.* 2012).

A top-down ETD direct infusion method of diluted whole blood without these time consuming steps could have a high potential for population screening in an automated manner to replace current technology. This mass spectrometry-based approach could give the required high specificity for confirmative diagnosis of clinically significant variants, and the ETD technology could also offer the opportunity for the occasional identification of unknown variants. Recently this group have extended their approach to HbE and HbD variants and developed the method to involve HbA_2 and HbA_{1c} quantitation (Scherl *et al.* 2013).

For mass spectrometry to be a competitive and potential replacement technique for population screening, it has to be shown to provide the confirmation of the presence of variants of clinical significance, according to the requirements outlined in the NHS Sickle Cell and Thalassemia Screening Programme. It has to be capable of quantifying HbA_2 and HbF levels, which can be used as surrogate markers for thalassemic disorders and other clinical conditions (NHS 2012b) The clinically significant variants have been detailed in Chapter 1, section 1.3.1.

Additional requirements for a good screening method would be a high degree of automation regarding sample preparation, analytical measurement and data analysis in order to facilitate a high-throughput approach with a minimum requirement for expert staff. The ideal population screening method is fast with a short turn-around time to give the answers for the appropriate diagnosis as soon as possible, while maintaining the overall cost is the same or lower than it is for currently used methods.

The research presented here involved working with human blood, therefore certain ethical considerations had to be followed. According to the Human Tissue Act (HTA) (Human Tissue Act 2004) any research projects involving human tissue needs to be authorised and approved by the regional NHS Research Ethics Committee if samples from NHS patients are used. To obtain a valid ethical approval for the project an application form detailing the research was filled in through the Integrated Research Application System (IRAS) and submitted to West Midlands Research Ethics committee. Approval was granted for the project involving the analysis of anonymised patient samples within a three months period for mirrored testing with developed mass spectrometry-based methods.

Collaboration was established with the Haematology Department of University Hospitals Coventry and Warwickshire NHS Trust (UHCW). Methods were optimised using a small set of clinical samples prior to a clinical study analysing all samples investigated under hemoglobinopathy screening in the hospital during a period of three months.

In this Chapter, results obtained using the developed mass spectrometry-based methods are compared to cation exchange HPLC, which is used as the first line of testing for the detection of Hb variants, HbA_2 and HbF quantitation in the hospital laboratory. The second line testing for confirmation of abnormalities used in the hospital is gel-electrophoresis (acid and alkaline), therefore, presented results are correlated to both HPLC and electrophoresis methods in cases where a significant variant is observed. Unequivocal identification of hemoglobin variants can only be obtained through either protein sequence analysis or DNA analysis. The latter is very rarely performed because of its relatively high cost (Ryan *et al.* 2010). The ability to perform protein sequencing analysis is readily available with MS.

In collaboration with the Haematology Department of UHCW, over 2,000 patient samples were analysed with two different mass spectrometry based techniques using a Xevo TQ instrument (Waters, Manchester, UK) and over 1,000 samples were analysed with a top-down ETD method as part of a pilot study.

At the Haematology Department of UHCW approximately 7,000 samples are analysed annually for antenatal screening purposes. The screening protocol set in the NHS lab manual is followed (NHS 2012b). Initial measurements include assessing the red cell indices: MCH (mean corpuscular volume), MCV (mean corpuscular hemoglobin) and Hb concentration defined in section 1.3.3. If the measured MCH is lower than 27 pg and the MCV lower than 80 fL this may indicate abnormalities, and further investigation is required.

Alongside these measurements in most cases a family origin questionnaire (FOQ) is filled out by the patient to determine whether the individual is a descent of a high prevalence area, and so require further screening.

The research undertaken here is aimed at antenatal screening, no attempts at newborn screening have been made. Samples in rare occasions are from newborns or from babies up to few months old.

The current clinical screening strategy employs first and second line screening steps to confirm certain mutations. It has been presented in a previous Chapter that method development was focused on obtaining information at the intact protein globin chain level and also at the peptide level. Intact globin chain analysis alone is not confirmative for clinically significant variants, only preliminary indication can be gained, and analysis at the peptide level is a necessary follow up step. In the work presented here the two methods have been evaluated and used in combination to obtain full confirmation of variants detected during the intact measurement but not fully characterised. Where possible, the MRM method was used for the purpose of confirmation. This method was able to confirm HbS, HbC, HbE variants, but it was only tested on one sample with HbO-Arab. For the HbD-Punjab mutation, high resolution MS measurements were required. Here we suggest that two-step analysis needs to be performed to obtain definitive diagnosis, since neither of the developed methods are able to give diagnosis alone. Both of these methods utilise the same instrument, therefore, an efficient analytical protocol could be created involving the different approaches but only using a single instrument.

In tthe second part of this chapter the potential of developing a screening method based on top-down ETD methodology is tested using a mass spectrometer with an ion trap analyser. The suitability of such a method in clinical sample analysis has been assessed in a preliminary study involving patient blood samples. Detection and identification of point mutations can be achieved by intact protein analysis and top-down sequencing measurements. A screening method developed by combining the previously presented intact globin chain analysis and a top-down ETD MS/MS methods can offer a powerful tool to replace the limited screening techniques currently used in the clinical labs.

Initially sequence coverage studies were performed to confirm that the method would be capable of providing fragment ions characteristic of clinically significant variant globin-chains. To detect and identify the position of the mutation in a fairly complex MS/MS spectrum, the mass spectrometer needs to be able to resolve *m/z* peaks differing by less than 1 Da. Multiply charged ions with charge states up to 5+ may also be present.
For a triply charged fragment ion where the variant is only 1 Da apart from the normal chain, peaks with 0.3 Da difference are expected. These peaks need to be resolved, and the charge state identified. Proton transfer reaction will be used to deprotonate the multiply charged ions and reduce the charge state as much as possible, preferably to a single charge.
The resolution of an ion trap analyser is potentially higher than that of quadrupole analysers. They provide sufficient resolution to resolve fragment ions with 3-4 charges, where the isotope peaks are spaced at 0.33-0.25 Da distances. Additional software tools may need to be used to obtain deconvoluted mass spectra.

Statistical Power analysis has been used to assess the necessary sample number to be analysed (Cohen 1992, Cleophas 2009) Calculations indicate that at least 1500 samples need to be analysed to confirm agreement with the existing method at a sensitivity of 95% and a specificity of 90%. The sample size required for sensitivity assessment is 1460, and for specificity assessment 146.

Factors such as sample preparation and ageing were not investigated, nor were the limit of detection or lowest limit of quantitation determined. The dilution factor for

sample preparation has been based on published literature protocols, with occasional optimisation to obtain optimum peak-shapes and intensities. The aim of the research at this stage was not method validation with neither method applied. The focus was on testing developed MS-based approaches on a larger set of clinical samples, identifying aspects of sample preparation and MS-method which require optimisation and preparing for future method validation to obtain a robust and high-throughput approach.

When making a comparison of clinical measurements, a new measurement technique with an established one, it is necessary to investigate whether they agree sufficiently for the new to replace the current one. Such investigations are often analyzed inappropriately, notably by relying only on correlation coefficients. The use of correlation obtained by simple linear regression methods sometimes can be misleading. Alternative approaches based on graphical and model based statistical methods were used to study the relation between results from the different analytical techniques. Assessment of reproducibility and robustness, consistency of the methods were also tested.

Aims and Objectives

The first part of the work presented in this Chapter attempts to provide a versatile clinical study and focuses on the evaluation of mass-spectrometry based methods and the potential for additional future clinical studies. It is aimed to highlight issues and facilitate further improvements to eliminate shortcomings in order to develop a robust mass spectrometry-based approach to be implemented in the clinical setting.

Primary aims:

- To test the developed methods on a larger clinical sample set and to investigate their robustness, sensitivity, automation and required analysis time. These factors should ideally be equal or better than current clinical methods.

- To assess the need for further developments and optimisation. Often a method is developed on purified standards and small number of samples, and constant usage of the instrument may require further work.

- To analyse samples on a statistically relevant size of sample-number. It has been shown in method development involving clinical samples that the HbA_2 levels determined by MS methods differ in absolute values to ce-HPLC values. Analysing a larger sample set it is hoped to provide a realistic estimation of ranges for normal values obtained by MS and helps to define the cut-off value for HbA_2 with the different MS approaches.

- To estimate the incidence of certain clinically significant variants in the investigated population, and if possible investigate the ethnic background and the proportion of variants in certain ethnic groups.

- To assess whether a single mass spectrometry platform is suitable for screening for detection and confirmation of clinically significant variants, quantification of HbA_2 and HbA_{1c} levels, detection of elevated HbF levels and detection of other rare hemoglobin abnormalities.

The second part of the work presented here is a pilot study, therefore the aims are less ambitious and the results cannot be fully implemented into the clinical trial. Since the sampleset is partly the same the results are presented together with a different focus and set of objectives.

A top down ETD method developed in collaboration with Bruker Daltonics has been tested and initial results analysing clinical samples are presented. Work carried out by collaborators is described in the text.

The primary aims were:

- To test a mass spectrometer with an iontrap analyser for the quantitative analysis of HbA_2 and HbA_{1c} levels
- To evaluate the sequence coverage achievable by the top down ETD – based method. In case of the -1 Da variants the substituted amino acids are relatively close to the N- or C-terminus. It needs to be assessed whether suitably high sequence coverage, covering the affected regions, can be obtained reproducibly.
- To assess the mass accuracy which can be obtained at the deconvoluted intact globin chain level to evaluate the possibility of reliably detecting mass shifts based on mass errors higher than the average expected ±0.2 Da necessary for first line detection of variant hemoglobin chains.
- To use the method for the detection of clinically significant variants and identify characteristic fragment ions.
- To test the developed method and chosen reporter fragment ions on a larger set of clinical samples and assess the reliability of detection of clinically significant variants.
- To test the top-down ETD approach for the identification of rare variants.

3.2 Materials and methods

3.2.1 Materials for analyses performed on the XEVO mass spectrometer

Samples admitted to the Haematology Department of UHCW for hemoglobinopathy screening within the 12 weeks mirrored testing period were transported to our laboratory on a weekly basis. 12 batches of samples arrived to our lab containing sample numbers ranging from 93-280. The blood samples were in EDTA tubes, anonymised by the NHS Laboratory. For discussions between the hospital lab and our lab the barcode numbers used were given as a patient identifier by the hospital. For ease of handling, a different identification number was given to each sample using the letters of alphabet for different weeks and numbers. The samples can be tracked back only at the hospital site, but their origin is not directly identifiable for the researcher.

The number of samples in different weekly batches and the number of abnormalities detected by current clinical methods are summarised in Table 3.1.

Week	Batch Name	Sample Number	Number of samples where abnormalities detected by Hospital Lab
1	B	136	3
2	C	155	4
3	D	108	8
4	E	280	16
5	F	251	10
6	G	93	6
7	H	228	18
8	I	114	8
9	J	144	20
10	K	166	9
11	L	180	3
12	M	162	7
	Total	**2017**	**112**

Table 3.1. Analysed samples for a period of 12 weeks used for comparative study to current methods

Standards with known levels of HbA_2 (WHO 1st international reference reagent) and HbA_{1c} (ClinTest standard) have been used as reference points for quantitation purposes.

Sample preparation

Samples were aliquoted and stored at -80 ºC. 1 ml of stock solutions from fresh samples were prepared by adding 20 µL of the blood sample to 980µL water.
Solutions for direct infusion measurements were prepared by diluting 20 µL of stock solution in 380 µL of a solution composition of 50% ACN and 0.2% HCOOH. These samples were centrifuged for 3 minutes at 9503 g. 280 µL transferred into 96-well plates (well volume 350 µL). A single plate was composed of a maximum of 94 patient samples, HbA_2 reference standard and HbA_{1c} reference standard. The plates were covered with aluminium foil to prevent sample evaporation. For the HbA_2 standard a stock solution was prepared every month during the three months trial. The HbA_{1c} ClinTest is purchased as a solution which is approximately 250-fold less concentrated than normal human blood level, and therefore a further four-fold dilution was performed in 50% ACN/0.2% HCOOH.

Tryptically digested blood samples were prepared by mixing 200 µL of stock solution with 40 µL of 50% ACN and 0.5% HCOOH, incubating at room temperature for 5 minutes and then adding 12 µL of 1M ammonium bicarbonate (Sigma Aldrich) solution and 10 µL 5 mg/ml trypsin solution (Sigma Aldrich). The samples were kept at 37 ºC overnight, and then centrifuged for 3 minutes at 9503 g. 35 µL of the digest mixture was diluted in the 96-well plates with 210 µL of 10% ACN containing 0.5% formic acid. All prepared 96-well plates contained the digested HbA_2 reference standard, and a control sample containing HbG-Philadelphia digested peptide mixture. During initial measurements it was observed that trypsin originating from different production batches with different LOT numbers had different cleavage efficiencies in yielding the required peptides. The presence of the characteristic peptide for HbG-Philadelphia could be used to monitor the efficiency of the enzymatic cleavage. Trypsin, which contained contaminating proteolytic enzymes, resulted in no HbG Philadelphia peptide when overnight incubated, and also resulted in erroneous results for monitored peptide ratios.

3.2.2 Materials for experiments performed on the amaZon speed ETD mass spectrometer

Standard preparations were used for optimisation and preliminary studies. For assessing glycated hemoglobin measurements the ClinTest reference standard with known concentration of HbA_{1c} (11011, RECIPE Chemicals + Instruments GmbH, Munich, Germany) was used, and for delta chain level determination the WHO 1st international reference reagent for HbA_2 (89/666, National Institute of Biological Standards and Control) was used. Lyophilised HbS powder (Sigma Aldrich) and 4 patient samples with clinically significant variants, and a control healthy sample were also used for method development and optimisation.

A subset of clinical samples provided by the Haematology Department at University Hospitals Coventry and Warwickshire NHS Trust for the clinical trial were used for method testing. The numbers of analysed clinical samples are listed in Table 3.2.

Batch	Sample numbers	Number of abnormal samples (clinically significant variants / other condition)
J	144	15 / 5
G	93	4 / 2
H	228	14 / 4
L	180	3
E	190	6 / 2
M	162	6 / 1
C	48	1
Selected samples with variants from all batches*	90	**80**

Table 3.2. Clinical samples analysed using the developed top-down ETD method
*This selection also includes some samples which were previously analysed as part of the batch analysis

Sample preparation

Standard solutions were diluted in 50% aqueous acetonitrile with 0.2% (and in some cases with 0.5%) formic acid to achieve a final concentration of hemoglobin of approximately 7.5-15 ng/µL.
20 µL of whole blood was diluted in 980 µL deionised water to provide a stock solution (approximate concentration 3 mg/ml). The stock solution was further diluted for direct infusion experiments 200- and 400-fold in 50% aqueous ACN containing 0.5% formic acid.
The stock solutions of clinical samples for introduction with HPLC auto sampler were diluted 1 in 100 in 50% aqueous ACN containing 0.5% formic acid.

3.2.3 Intact globin chain analysis combined with full scan tryptic digest mixture analysis

MS method as described in Chapter 2 detailed in section 2.2.2. Triversa NanoMate (Advion Biosciences Ltd., Ithaca, NY, USA) was used for automated sample introduction (settings: aspirated sample amount 5 µL, gas pressure, applied voltage). The intact MS measurement was a 3 minute measurement in MCA mode with a scanning range of *m/z* 930-1210 and scan time 1 second. All samples were measured in three replicates. All data were processed with the BioPharmaLynx™ software (Waters, Manchester, UK) and exported and analysed in Excel.

For identification of rare variants the tryptic peptide mixture of the sample was analysed in full scan mode with a Synapt HDMS or Synapt G2 system, using gold plated needles and nanoESI source.

3.2.4 LC-MRM-based mass spectrometry analysis

The described method in section 2.2.2 referred to as MRM2-method on the Waters Xevo TQ instrument was used for the clinical trial: Targeted peptide analysis as described in the previous chapter for the normal β- and δ-chain peptides and hemoglobin variant specific peptides.

For data processing TargetLynx^{TM} (Waters, Manchester, UK) software was used, and results were exported into Microsoft Excel, where further quantitative calculations were performed.

3.2.5 Top down ETD Method

Experiments were performed using a Bruker amaZon speed ETD iontrap mass spectrometer (Bruker UK Limited, Coventry, UK).

Sample introduction

Sample introduction was achieved using either using a 500 µL syringe with a syringe drive at a flow rate of 3 µL/min or the auto sampler of the HP 1100 LC system. The initial flow delivered by the LC pump was 0.5ml/min decreased to 0.15 ml/min in the first 30 seconds, from which a lower, approximately 10-20 µL/min, flow rate was achieved with the use of a split line in front of the sample loop. 100 µL of the diluted whole blood sample was introduced into the mass spectrometer with the aid of a 100µL sample injection loop and the LC flow.
The solvent delivered by the LC system was equivalent to the dilution solvent 50% aqueous ACN containing 0.5% formic acid.

Mass spectrometry methods during preliminary studies

For intact globin chain analysis the mass spectrometer was operated in UltraScan mode (with a scan speed of 32,500 *m/z* per second) to achieve a mass spectrum of low resolution which, after deconvolution, provides accurate intact mass results for the average globin masses. Scan range was *m/z* 250-3,000, the number of averaged scans were 20, while the rolling average 2.

For ETD fragmentation experiments of the protein globin chains the mass spectrometer was operated using Enhanced resolution mode (with a scan speed of 8,100 *m/z* per second employed). Different ETD parameters were tested including number of reagent anions and ion/ion reaction times. ETD reagent anions were odd-

electron fluoranthene radical anions ($C_{16}H_{10}^{-}$) while the PTR reagent anions were even-electron fluoranthene anions ($C_{16}H_{11}^{-}$).

Mass spectrometry methods for the analysis of a clinical sample set

The applied ETD fragmentation method was developed by Ralf Hartmer, (Bruker Daltonik GmbH, Germany, Bremen) and further developed by Julia Smith (Bruker UK Limited, Coventry, UK).

Sample introduction was performed as described above with the HP 1100 LC system. The automated sample introduction was controlled using the HyStar software.

The acquisition method contained 5 different segments with different measurement settings for different purposes. The schematic of the acquisition can be seen in Figure 3.41.

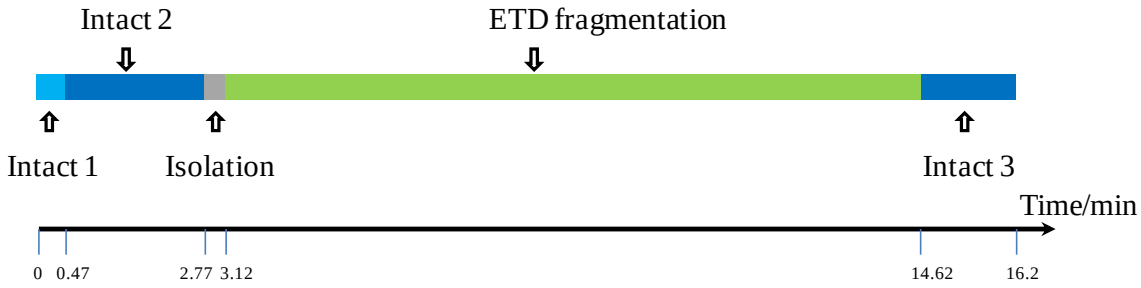


Figure 3.1. Schematic of acquisition for hemoglobin variant diagnosis using the Bruker amaZon speed ETD mass spectrometer

Ion source settings for all measurements: capillary voltage 4.5 kV, end plate offset 500V.

Segment 1: Optimised for intact globin chain analysis and HbA_{1c} quantitation. UltraScan mode, ICC Target 300,000, maximum accumulation time 50 ms, scan range *m/z* 250-3,000, averages 20, rolling average 2 (funnel 1 settings in 100, out 35, lens 25)

Segment 2: Optimised for intact globin chain analysis and HbA_2 quantitation. UltraSscan mode, ICC Target 300,000, maximum accumulation time 50 ms, scan

range *m/z* 250-3,000, averages 20, rolling average 2 (funnel 1 settings in 100, out 95, lens 85)

Segment 3: Isolation of precursor ion β18+ at *m/z* 882.7 in a two step isolation. First step isolation of 885.5 width 10.0 *m/z*, second step isolation of 882.7 width 3.0 *m/z*. UltraScan mode, ICC Target 225,000, maximum accumulation time 50 ms, scan range *m/z* 250-3,000, averages 10.

Segment 4: ETD fragmentation of isolated β-chain ion, subsequent PTR reaction for charge state reduction of fragment ions. Enhanced resolution mode, ICC Target 175.000, maximum accumulation time 10 ms, scan *m/z* 250-3,000, averages 10. Isolation as in *Segment 3*. ICC ETD 250,000, ETD Cutoff 120 *m/z*, reaction time 20 ms, PTR Target 150,000, PTR cutoff 120 *m/z*, reaction time 80 ms.

Segment 5: Optimised for intact globin chain analysis. UltraScan mode, ICC Target 175,000, maximum accumulation time 10 ms, scan range *m/z* 250-3,000, averages 10, rolling average 2 (Funnel 1 settings in 100, out 95, lens 85)

Data analysis

Acquired data were processed using the Data Analysis software (ver. 4.1., Bulid 359, Bruker Daltonik GmbH, Bremen, Germany)
Processing methods: A Maximum Entropy algorithm was used for intact globin chain analysis, SNAP II was used for deconvolution of the ETD spectra.
Following the analysis of clinical samples a processing script written by Zoltan Czentar (Bruker Daltonik GmbH, Bremen, Germany) was used. The script performed MaxEnt deconvolution of the intact protein data, deconvolution of the ETD spectra with SNAPII, and generated a pdf report showing four predetermined *m/z* windows for the monitoring of clinically significant variants, with a mass spectrum of the deconvoluted intact globin chain spectrum, listing the first 20 most intensive peaks, their calculated *m/z*, intensity and relative intensity in percentages. The script is capable of batch processing for high-throughput data analysis.

3.2.6 Statistical analysis

For comparison of quantitative measurement capabilities of developed methods to the current clinical method, the statistical methods used were Bland-Altman and Passing-Bablok. These methods were undertaken using the software Analyse-it (free trial version, Analyse-it Software, Ltd., Leeds, UK)

Linear regression within Microsoft Excel: Correlation coefficients were used to measure the strength of linear relationship between two sets of variable. Pearson's correlation coefficient (R) was calculated using the assumption that both variables are approximately normally distributed. The R^2 value describes the proportion of variance in y that depends on x. This correlation assesses the association not the measurement bias. Relationship and comparison studies based on linear regression can be useful for first line graphical presentation of possible correlation but differential plots and additional statistical methods can only be used for evaluation of an acceptable method agreement with the reference method (Petersen *et al.* 1997).

Bland-Altman is a graphical method used to assess agreement between two analytical methods. The difference between values from the two methods are plotted against the average of the results (Bland and Altman 1986, 1999). The resultant plot is often called a Bland-Altman plot (or difference plot or bias plot). This method reveals the relationship between differences and provides a value for the mean difference (bias). The limits of agreement between methods is set as a range of ± 2 standard deviation from the mean difference. The graphical presentation identifies potential outliers, and also the pattern may suggest further systematic error or proportional bias. It is expected that 95% of data points fall within the limit of agreement range for two methods to be interchangeable. It does not say anything about possible restriction for the size of the range, that is a decision to be made by the clinician. If the difference between two methods is clinically acceptable the two methods can be used simultaneously and interchangeably.
This method has been used by Roberts *et al.* when compared glycated hemoglobin levels obtained by MS to total glycation levels determined by ion exchange chromatography (Roberts *et al.* 2001).

The Passing-Bablok method is a model based method for comparison of methods. It employs a linear regression procedure with no special assumptions regarding the distribution of the samples or the measurement error (Passing and Bablok 1983). The result obtained does not depend on the assignment of the methods to X and Y. This regression procedure is not sensitive towards outliers and assumes that measurement errors have the same distribution in both methods, but it makes no special assumption about the type of the distribution, it is not necessarily normal (Bilic-Zulle 2011). The slope and the intercept are calculated with their 95% confidence intervals. If the 95% confidence interval contains 0 for the intercept and 1 for the slope, then the hypotheses for not having any constant or proportional bias is not rejected.

The Passing-Bablok comparison method within the Anlayse-it software also tests linearity with the Cusum test, which inspects how the regression line fits the data or how randomly the data scatters about the line. If a nonlinear relationship between x and y is detected, it is expected that too many consecutive data points will be above or below the fitted line. This method can also be used to estimate systematic error between two methods with two hypotheses tested or whether there is constant bias or proportional bias.

Coefficient of variation histograms and box plots were used to present mass error ranges and technical variation. These were generated using the statistical software GenStat.

3.3 Results and discussion

Part I: Results obtained on the full sample set using the XEVO TQ mass spectrometer

3.3.1 Determination of HbA_2 levels

3.3.1.1 Intact globin chain analysis

HbA_2 levels were determined for all 2,017 samples using the intact globin chain analysis. All samples were measured three times. Their average values and standard deviation were calculated, also the coefficient of variation was determined. The calculated coefficient variation values are plotted in a frequency histogram in Figure 3.2. The average coefficient of variation is 7%, and 95% of the values are below the expected 15% for clinical MS methods (Kollipara *et al.* 2011).

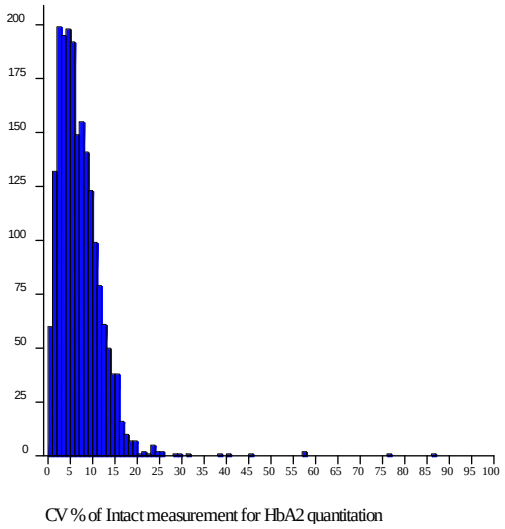


Figure 3.2. Histogram for the distribution of coefficient of variation of the HbA_2 quantitation by the intact globin chain analysis

Correlation of HbA_2 values determined by MS and ce-HPLC was first investigated by plotting the MS values against the HPLC values, and with simple linear regression analysis the coefficient determination (R^2) values were calculated. In

Figure 3.3 an example of this analysis for the samples in Batch M is shown. All of these figures for all batches can be found separately in Appendix 1.

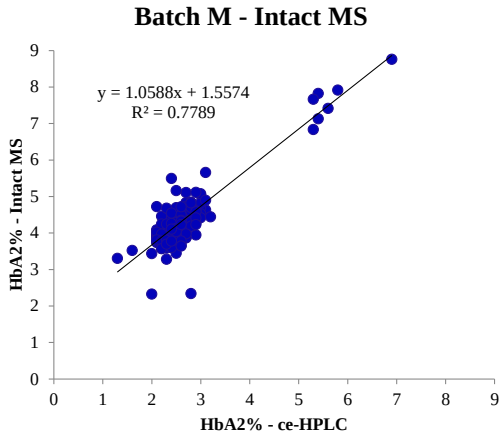


Figure 3.3. Calculated HbA_2 values determined by intact globin chain analysis plotted against values provided by the ce-HPLC method for the samples in Batch M

The parameters resulted from the simple linear regression analysis for all batches are summarised in Table 3.3. The R^2 values are shaded based on whether the fitted line is thought to give acceptable or better correlation. The limits for this classification were arbitrarily chosen.

The values based on the simple linear regression suggest there is linear relationship. This type of analysis, however, can be misleading, since fitting a trend line to two almost distinct groups will always give high correlation. R^2 values are only appropriate for an initial visualisation of the correlation. The measured and calculated values are in all cases higher than the ce-HPLC values, but with a shift. All intercept values are above 1, the average for all twelve batches is 1.61, while the slope is around 1. This would suggest a linear relationship with no proportional but constant bias.

When presenting the distribution of the measured values in a dot-histogram next to the distribution of the HPLC values (Figure 3.4), it is clearly visible, that there is a significant shift, and the samples identified with elevated HbA_2 levels are also elevated according to the intact MS method. This separation of normal and elevated values can be observed for all measured sample batches, and the method consistently and reliably identifies the samples with elevated levels.

Batch ID	Number of Samples	R^2 - summary	Intercept	Slope
B	136	0.58	1.625	0.9428
C	155	0.68	1.2307	1.111
D	108	0.90	1.3515	1.1468
E	280	0.73	1.8338	1.0775
F	251	0.57	1.9888	0.9424
G	93	0.61	2.4055	0.8811
H	228	0.65	1.8584	0.9853
I	114	0.82	1.5562	1.1132
J	144	0.74	1.3539	1.0577
K	166	0.72	1.2917	1.0365
L	180	0.74	1.3233	1.0905
M	162	0.78	1.5574	1.0588
Average		**0.71**	**1.61**	**1.04**

best	better	acceptable
≥0.8	0.6-0.79	≤0.59

Table 3.3. Results of the correlation analysis for all analysed sample batches by the intact globin chain analysis

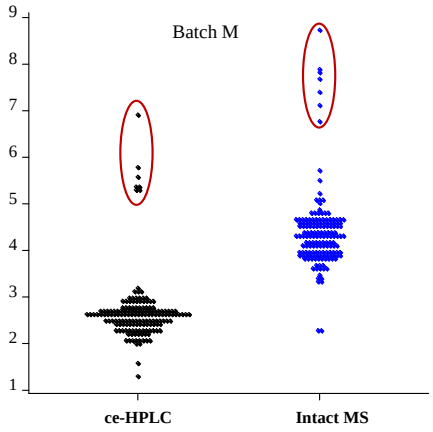


Figure 3.4. Distribution of measured values for the HbA_2 quantitation by the HPLC method and by the intact globin chain analysis.

To investigate the relationship between values from two methods other statistical methods may be more appropriate.

Results of the Bland-Altman statistical analysis, while only looking at the differences between the values from the reference method and the new method can give an estimation regarding systematic bias between the outcomes of the methods. The graphical outcome of the statistical analysis can be seen for the samples of batch M in Figure 3.5. The figure shows the distribution of the differences around the bias, and the outliers, which are not in the range for the limits of agreement. The same type of figure for all batches can be found in Appendix 1.

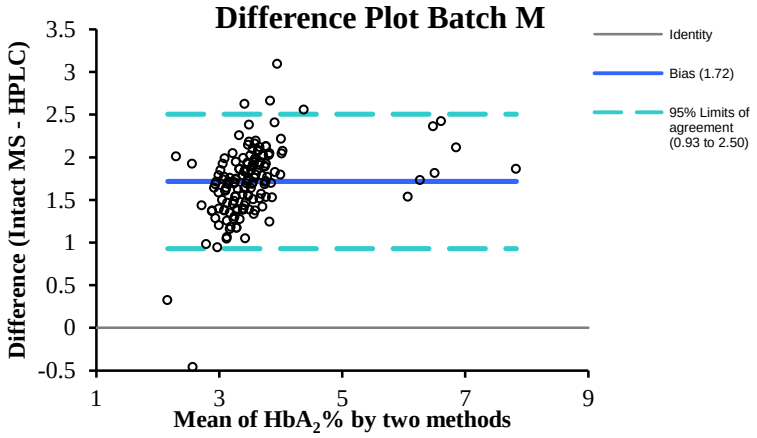


Figure 3.5. Difference plot for the comparison of values from HPLC and intact MS methods using the values from batch M

The results of this type of analysis for all sample batches are summarised in Table 3.4. The average bias is 1.73, similar to the average constant bias from the simple linear regression. There are a few outliers in case of every batch, but 95% of the values are within the limits of agreement. The width of the agreement range varies from 1.13 to 2.09 and the average for all batches is 1.53. This range of HbA_2 values indicates a high fluctuation which may be unacceptable for clinical diagnosis considering that the 3.5% limit between normal and elevated values allows only a narrow range for variation. The pattern of the graphical presentation may suggest a presence of proportional error, since the difference values tend to shift towards the higher levels as the average values increase.

The results from the Passing and Bablok statistical method are somewhat different. For batch M it does not suggest constant bias, but 1.64 as proportional bias as shown in Figure 3.6. Table 3.5 summarises all the results for all batches. The corresponding figures can be found in Appendix 1. In 10 batches the confidence interval for the constant bias contains 0, meaning the hypothesis to have no constant bias is not

rejected. None of the 95% confidence intervals for the proportional bias contained 1, the hypothesis regarding proportional bias is definitely rejected as a result of a comparison of the two techniques.

Batch ID	Number of samples	95% CI range of agreement (%)	Width of agreement range (%)	Number of outliers	Bias (%)
B	136	0.67-2.27	1.60	7	1.47
C	155	0.84-2.21	1.37	9	1.53
D	108	1.02-2.51	1.49	4	1.77
E	280	1.26-2.83	1.57	9	2.04
F	251	1.15-2.83	1.68	5	1.99
G	93	1.23-2.97	1.74	3	2.10
H	228	0.8-2.89	2.09	10	1.84
I	114	1.23-2.49	1.26	4	1.86
J	144	0.76-2.26	1.50	5	1.51
K	166	0.72-2.05	1.33	6	1.39
L	180	1.00-2.13	1.13	10	1.57
M	162	0.93-2.50	1.57	6	1.72
Average			**1.53**	**6.5**	**1.73**

Table 3.4. Results of the Bland-Altman statistical analysis for all analysed sample batches by the intact globin chain analysis

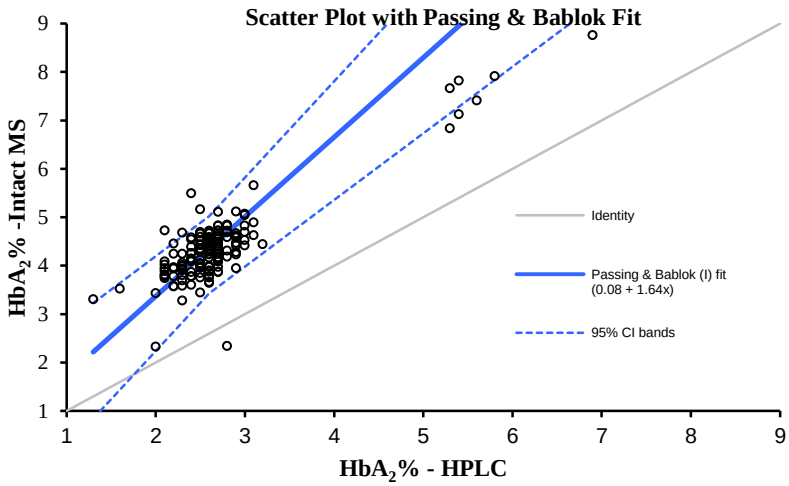


Figure 3.6. Passing-Bablok regression analysis for the intact MS results for batch M

All statistical methods support the hypothesis that values from the intact globin chain analysis differ with a certain bias observed from the ce-HPLC analysis result. To use intact globin-chain analysis as a diagnostic tool for HbA_2 level determination either this certain bias (constant or proportional or both) needs to be defined and used as a correction factor for calculating true values, or an appropriate calibration process

needs to be applied. New limits and ranges for HbA_2 levels for normal and abnormal samples may be defined when the intact globin chain analysis is used. This requires further method development and validation.

Batch ID	Number of samples	Constant Bias	Proportional Bias	Approximate numbers of outliers	Comment
B	136	-0.4	1.7	4	
C	155	-0.1	1.59	18	Low constant bias
D	108	0.69	1.38	13	
E	280	0.33	1.62	~30	*Non-linear relationship detected*
F	251	0.02	1.80	11	No constant bias?
G	93	0.8	1.54	4	
H	228	-0.46	1.85	15	
I	114	0.68	1.44	7	
J	144	0.11	1.51	13	Low constant bias
K	166	0.24	1.43	10	
L	180	0.33	1.45	13	
M	162	0.08	1.64	10	No constant bias
Average		**0.19**	**1.58**	**12.3**	

Table 3.5. Results of the Passing & Bablok Statistical analysis for all analysed sample batches by the intact globin chain analysis

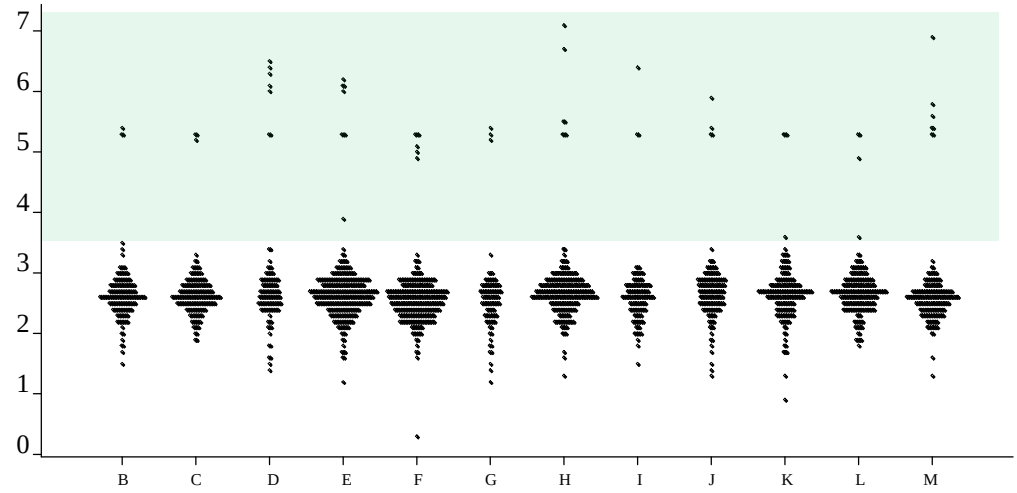
The intact MS method consistently detected elevated δ-chain levels. The average calculated HbA_2 level measured by the intact globin chain analysis was 7.5 % for the reference standard which had an indicated HbA_2 value of 5.3%. With simple ratio calculation the 3.5% limit set for the HPLC-method with the MS method would be 5%. The 5% limit might be appropriate for some of the sample batches, but not for all of them, a arbitrarily chosen limit of 6% would provide a more appropriate boundary.

The distribution of HbA_2 values by the HPLC method and intact MS is presented in Figure 3.7. The limit of 3.5% for the HPLC method clearly separates the elevated values based on that method. For the intact MS method a limit of 6% also separates all of the elevated values.

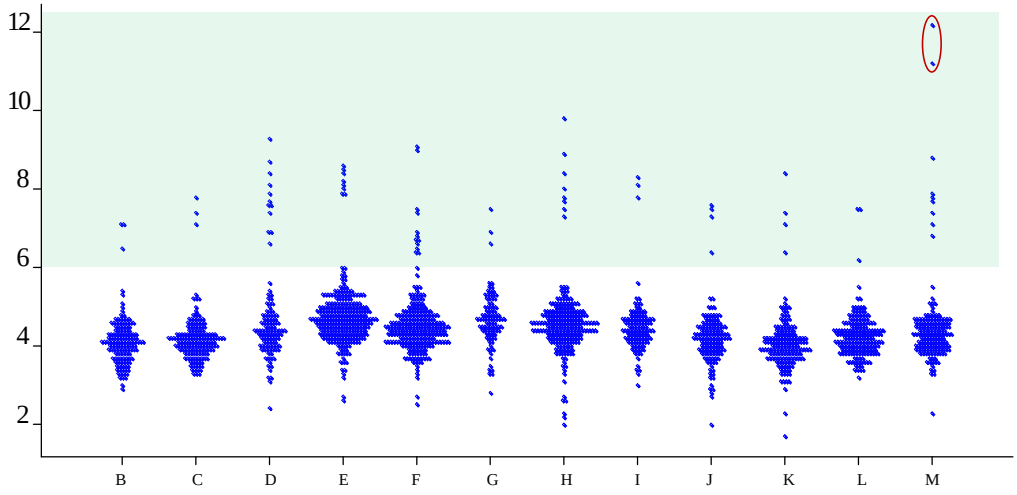
In case of batch M two sample values circled showed elevated HbA_2 levels with the intact MS method, because those samples were carriers of variants which had been not detected chromatographically. Intact globin chain analysis allows calculation of δ-chain levels as a ratio to the normal β-chain levels. When a variant β-chain is present, this ratio is affected, and in some cases the measured HbA_2 levels can be

used as indicator for abnormalities other than the elevated δ-chain levels due to β-thalassemia.

Based on the results presented here intact globin chain analysis is suggested to be a good indicator method for elevated HbA_2 levels and other abnormalities, but it has a certain bias, which needs to be defined and taken into account or corrected with appropriate calibration. For calibration a reference sample series with at least 3-5 points of different levels would be advised.



a)



b)

Figure 3.7. Distribution of the determined HbA_2 values by the hospital laboratory using the ce-HPLC method (a) and by the intact globin chain analysis (b)
Data are shown separately for each batches analysed. Shaded area represents the 3.5% barrier for elevated values in case of the HPLC results and the proposed barrier of 6% for the MS method

3.3.1.2 LC-MRM analysis

With the MRM–method three different peptide ratios were monitored to quantify relative percentage of δ-chain compared to the β-chain, giving an estimation about the HbA_2 levels. HbA_2 levels were then calculated for the three different peptide pairs.

Initial statistical analysis was performed by plotting the calculated values against HPLC values. By applying a simple linear regression, R^2 values and parameters of the equation belonging to the fitted line were determined.

Examples for the graphical outcome of the simple linear regression analysis can be seen in Figure 3.8 for all peptide ratios for one batch. Figures for all batches can be found separately in Appendix 1. The summarised parameters from the linear regression analyses are listed in Table 3.6.

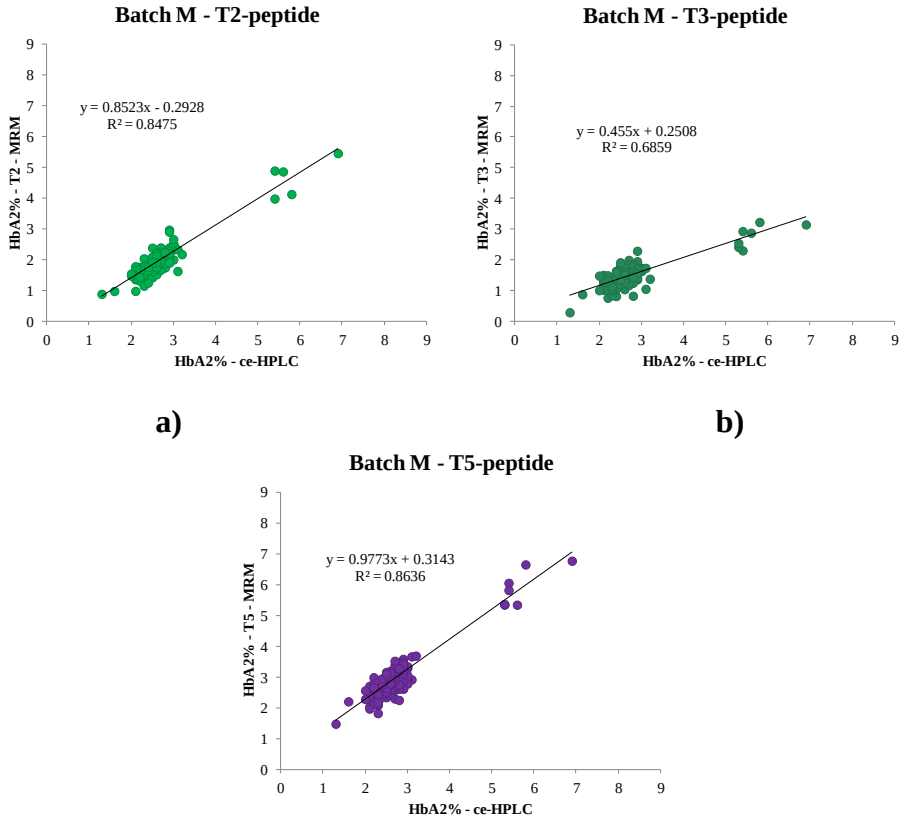


Figure 3.8. Calculated HbA_2 values determined by the MRM plotted against the values provided by the ce-HPLC method for the samples analysed in Batch M. For T2 peptide ratios (a), for T3 peptide ratios (b) and for T5 peptide ratios (c)

Batch ID	No. Samples	R^2 - T2	R^2 - T3	R^2 - T5	Intercept – T5	Slope - T5
B	136	0.42	0.49	0.81	0.2417	0.9857
C	155	0.46	0.77	0.72	0.0351	0.9264
D	108	-	0.44	0.91	0.4902	1.1775
E	280	0.41	0.81	0.75	0.1894	0.8303
F	251	0.22	0.62	0.75	0.4528	1.0446
G	93	0.11	0.48	0.79	0.3565	0.9855
H	228	0.67, 0.63	0.82	0.73	0.3939	1.0054
I	114	0.66	0.59	0.90	0.2596	0.7369
J	144	0.32, 0.60	0.90	0.80	0.4784	1.2577
K	166	0.21	0.12	0.87	0.7881	1.0877
L	180	0.52, -	0.53, -	0.53, 0.76	-0..287, 0.5687	1.2578, 0.7811
M	162	0.85	0.69	0.86	0.3143	0.9773
Average		**0.47**	**0.61**	**0.78**	**0.3492**	**1.0041**

Best	better	acceptable
≥0.8	0.6-0.79	<0.59

Table 3.6. Results of the correlation analysis for all analysed sample batches by the different peptide ratio analysis using the MRM method

The resultant values from simple linear regression together with the investigation of the figures suggest that the T5 peptide ratios are the most consistently useful indication for the HbA_2 levels. In case of T2 peptides, the distribution of the calculated values were in an abnormal range for some sample batches, levels of 60-90% for the δT2 peptide relative to the βT2 peptide were calculated, highlighting some serious issues with enzymatic cleavage. In other cases, for example batch M, the values seemed to be perfectly normal and highly correlated to HPLC values. T2 peptides seemed to be more affected by the incomplete cleavage problem, with T3 peptides to a lesser extent, and T5 peptides even less. The ranges slightly fluctuate for T5 peptides, but the elevated and normal levels are clearly separated. Distribution of the measured and calculated values can be seen for all peptide ratios compared to the distribution dot histogram of the HPLC values in Figure 3.9.

The T3 peptide ratios did not show any radical differences from batch to batch, but they did not show a consistently high correlation to HPLC values as had been displayed by the T5 peptides. The values, based on T5 peptide ratios were able to distinguish more easily between normal and elevated values, as shown in Figure 3.9.

For further method comparison with previously used statistical methods, only the T5 peptide ratios were chosen. The parameters of the equation of the fitted lines to the T5 values are included in Table 3.6. The intercept values fluctuate from batch to batch resulting in an average of 0.35, but calculating mean value here may not be appropriate. In case of Batch L the calculated values clearly fall into two separate distributions for the two 96-well sample plates measured separately, and therefore trend lines were fitted separately and two values are shown for R^2, intercept and slope values.

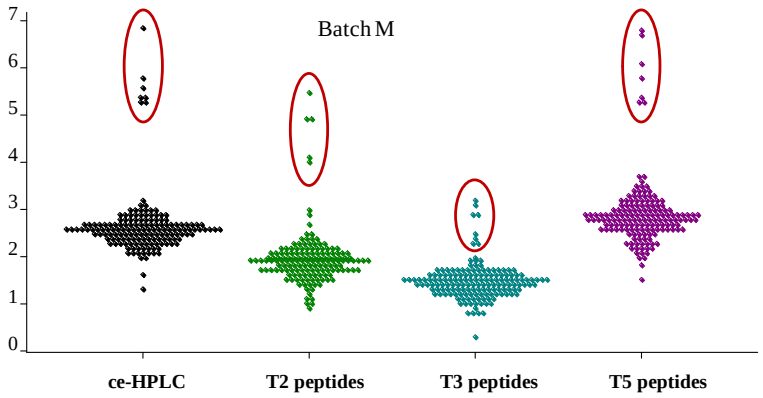


Figure 3.9. Distribution of measured values for the HbA_2 quantitation by the MRM method for different the peptide ratios monitored
The data points indicating elevated HbA_2 levels are circled in red.

All samples were measured in three replicates. The mean values and standard deviation were calculated, and the coefficient of variation was determined. The calculated coefficient variations for all three different peptide ratio values are plotted in a frequency histogram in Figure 3.10. The average cv for T2 peptides was 3.9 and 98% of the values are below 15%. For T3 peptides the average cv was 7.5, and 91% of the values were below 15%. For T5 peptides the average was 6.6 and 95% of the values were below 15%. Coefficient of variation values are best for the T2 values, but for T5 peptides are still acceptable. The better results for T2 values may be explained by the occasionally abnormally high levels. When those were measured with the same standard deviation, this resulted in lower cv% values.

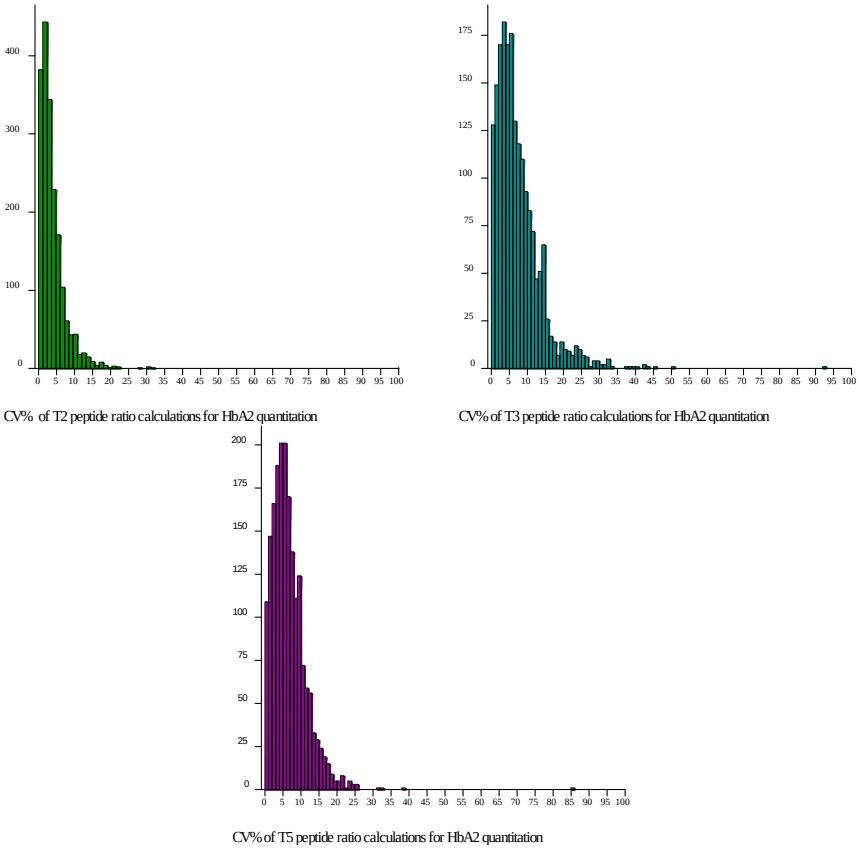


Figure 3.10. Histogram for the distribution of coefficient of variation of the HbA_2 quantitation by the MRM method a) T2 peptides, b) T3 peptide and c) T5 peptides

Bland-Altman statistical analysis has been performed on all twelve batches for the HbA_2 levels obtained from T5 peptide ratios, results are summarised in Table 3.7. The width of the ranges for the differences are slightly smaller than previously observed for the intact globin chain analysis. The number of outliers for all batches on average is approximately the same. The bias values, however, fluctuate much more than observed for the intact MS method. There is a definite bias between the MRM-method and the HPLC method. The graphical results of the Bland-Altman-method can be found in Appendix 1 for all sample batches shown separately.

Patterns observed for the calculated differences for different sample batches do not strongly indicate the presence of proportional bias observed with the intact MS method, and for some of the batches no indicative pattern was observed at all. For all batches 95% of the data points were in range of agreement. These analyses indicate an agreement between the two techniques. One can state that the methods may be interchangeable, which forms a subject for further investigation.

Batch ID	Number of samples	95% CI range of agreement	Width of agreement range	Number of outliers	Bias
B	136	-0.27-0.68	0.95	6	0.2
C	155	-0.68-0.36	1.04	3	-0.16
D	108	0.27-1.70	1.43	6	0.99
E	280	-0.87-0.33	1.2	14	-0.27
F	251	-0.01-1.15	1.16	10	0.57
G	93	-0.29-1.03	1.32	3	0.37
H	228	-0.37-1.19	1.56	6	0.41
I	114	-0.81-(-0.05)	0.76	5	-0.43
J	144	0.35-2.01	1.66	7	1.18
K	166	0.6-1.45	0.85	9	1.02
L	180	0.63-1.30	1.93	4	0.34
M	162	-0.29-0.79	1.08	5	0.25
Average			**1.25**	**6.5**	**0.37**

Table 3.7. Results of the Bland Altman statistical analysis for all analysed sample batches by the MRM analysis for the T5 peptide ratios

Results of the Passing-Bablok method comparison are less convincing than for the intact MS method. The bias values fluctuate more from batch to batch. The average constant bias is a negative value, -0.60, and only in 5 batches is the hypothesis of not excluding the possibility for no constant bias not rejected. The hypothesis about no proportional bias is rejected for all batches. The number of outliers observed is significantly higher, for some larger batches over 30.

In the case of batches C, H, L and M a nonlinear relationship were detected. This is most likely explained by differences between sample plates within one batch. In some cases a plate was re-measured, while there was no time to re-measure the second plate, and therefore the differences are emphasized. In case of batch L, where observation of clearly separated distribution of calculated values was made, the nonlinear relationship would be expected.

Batch ID	Number of samples	Constant bias	Proportional bias	Approximate number of outliers	Comment
B	136	-0.14	1.13	13	
C	155	-0.70	1.21	3	*Nonlinear relationship*
D	108	0.03	1.33	15	
E	280	-0.04	0.9	~45	
F	251	-0.43	1.38	~38	
G	93	-0.51	1.32	6	
H	228	-1.75	1.81	12	*Nonlinear relationship*
I	114	0.12	0.79	7	
J	144	-0.41	1.60	12	
K	166	0.26	1.28	~22	
L	180	-3.08	2.29	6	*Nonlinear relationship*
M	162	-0.54	1.31	9	*Nonlinear relationship*
Average		**-0.60**	**1.36**	**15.7**	

Table 3.8. Results of Passing-Bablok Statistical analysis for all analysed sample batches by the MRM analysis for the T5 peptide ratios

The distribution for all measured values for all batches separately (Figure 3.11) also highlights differences between normal distributions. Elevated values are in most of the cases clearly detected. There is no clear barrier which may be set from where elevated values are separated for all sample batches.

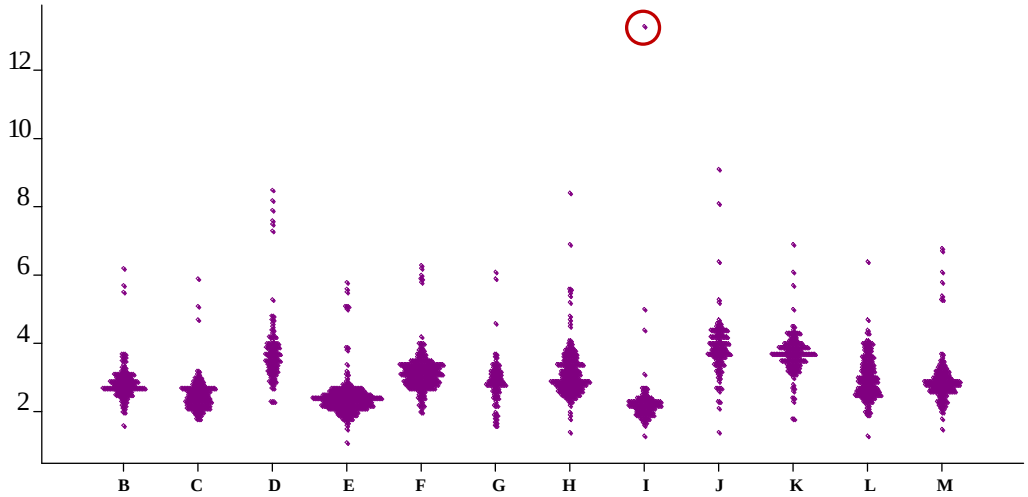


Figure 3.11. Distribution of the determined HbA_2 values by the MRM method calculating T5 peptide ratios
Data are shown separately for each batches analysed

There is a high HbA_2 value obviously present in batch I, which belongs to a sample diagnosed as Hb Lepore.

Using the MRM method monitoring the T5 peptide ratios may give indication for elevated levels within one sample batch, or more precisely among the results from one sample plate, where the samples have been prepared the same way, with elution solvents, chromatography column the same. To avoid such high plate to plate fluctuation, employment of calibration is necessary, and it is needed more than it was for the intact MS measurements. The issue regarding enzymatic cleavage incompleteness also needs to be investigated and addressed.

When values calculated for the reference standard are plotted for the intact MS method and for the T5 peptide ratios from the MRM method it is observed that there is a higher level of fluctuation for the MRM-method levels than for the intact MS levels (Figure 3.12). This fluctuation of values is mirrored in the previously mentioned fluctuation of distribution ranges, which could be a result of problems with the enzymatic cleavage, or with the reliability of the monitored transitions.

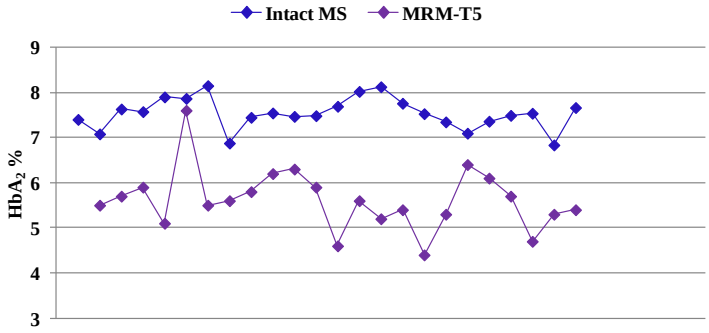


Figure 3.12. HbA_2 values for the reference standard measured by the intact globin chain analysis and by the MRM method based on the T5 peptide ratios
Each dot represents one data point from one 96-well sample plate

3.3.2 Clinically significant variant detection

3.3.2.1 Intact globin chain analysis

Detection of HbS

Detection of HbS variants on the intact mass level is straightforward based on the expected -30 Da mass shift. All samples during the three month trial were analysed using the intact globin chain method, and the data were processed using BiopharmaLynx to screen for the presence of the β^{SICKLE} mass peak next to the wild type β-chain mass peak. The samples classed as heterozygous carriers, compound heterozygotes for HbS and HbC or homozygous of the sickle mutation are listed in Table 3.9.

Batch Name	**Samples classed as SS or transfused SS**		**Samples classed as AS**		**Samples classed as SC**	
B	B127	1	B2,B7	2	-	
C	-		C1, C103, C154	3	C99	1
D	D92	1	D86, D87, D90, D96, D108	5	D95	1
E	E34, E146, E180, E239,	4	E20, E197, E240, E257	4	-	
F	F129, F229, F230, F231	4	F50, F116, F128	3	-	
G	G92	1	G32,G33,G93	3	-	
H	H39, H40, H143	3	H41, H42, H43, H44, H47, H70	6	-	
I	I1, I74	2	I7, I9, I39,I53	4		
J	J48, J87, J106, J116	4	J16, J41, J42, J43, J45, J66	7	J40	1
K	-		K42, K43, K45, K46, K50, K57, K78	7	-	
L	-		L49, L130	2	-	
M	M42, M154	2	M50, M51, M122	3	-	
Total: 73	**22**		**48**		**3**	

Table 3.9. Summary of sickle conditions among all batches analysed in the clinical trial.

If a peak -30 Da lower than the β-chain was observed, the percentage of the β^{SICKLE} was calculated as a ratio to the sum of beta-like chains the same way as for the percentage of the δ-chain levels previously shown.

A mass peak at -30 Da (15,837.3) compared to normal β-chain mass (15,867.2) was observed in the mass spectra for all listed samples in Table 3.9. In agreement with hospital results that can be indicative of the presence of the sickle variant. Further confirmation was necessary, which was achieved using peptide level analysis, which is presented later.

In Figure 3.13 the calculated percentage of HbS for samples classed as AS – heterozygous sickle hemoglobin carriers can be seen. The general observation can be made that the majority of the calculated values are between 30 to 45%. The calculated values are summarised in Table A.1 in Appendix 2.

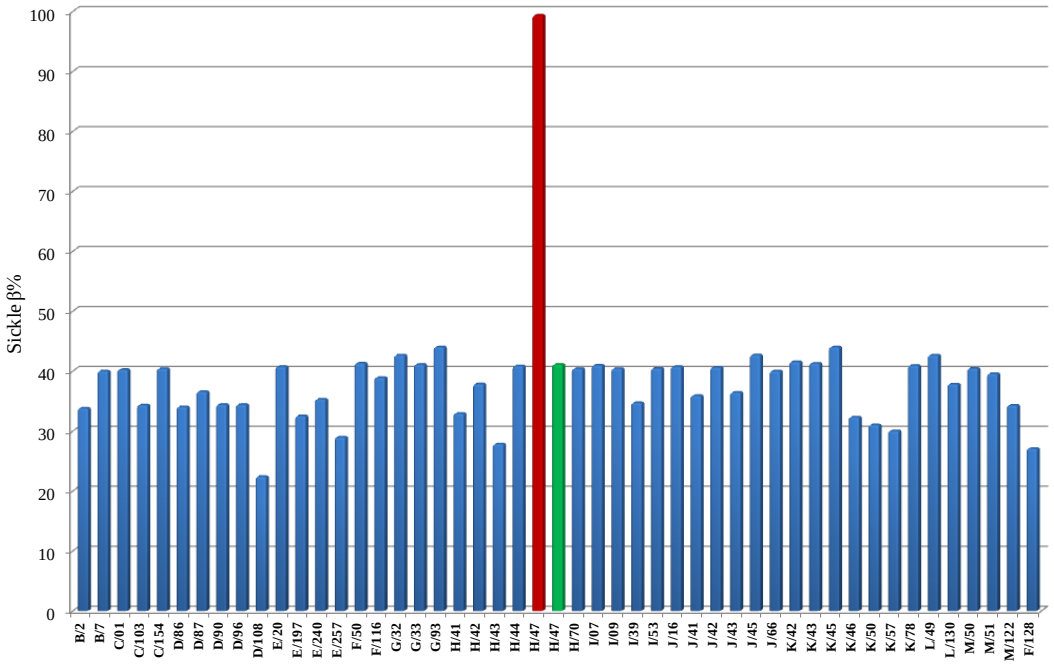


Figure 3.13. Calculated β^{SICKLE}-chain levels for the samples classed as heterozygous sickle carriers by the hospital laboratory

The sample H47, shown in red first calculated against normal β-chain which was not present, and then against the variant peak which was observed, recalculated value shown in green.

In case of sample H47, there were no normal β-chain mass peaks observed, and calculation against normal β-chain resulted in 100% for β^{SICKLE}, but the sample contained a significant peak belonging to a rare Hb variant, Hb Iraq-Halabja. The

value for this sample has been recalculated using the variant mass peak intensity instead of the normal β-chain.

When a sample was classed as SS, homozygous sickle was assumed. In some cases that is true and no or very low intensity mass peaks for normal β-chain could be observed in the deconvoluted mass spectrum. The graphical presentation of the calculated HbS levels can be seen in Figure 3.14. In some cases samples classed as SS still had a relatively high level of normal β-chain, and in cases of sample D92, E239, I40, I106, I116, M42 and M154 the HbS level were 50% and below. In samples F129, F230, F231 and G92 this is explained by the information given by hospital stating that the patients had received blood transfusion, which is a known way of treating the symptoms of Sickle cell disease. Based on this it is assumed that the other samples are from individuals who also have received blood transfusion previously but this information has not been provided or the samples were from heterozygous carriers.

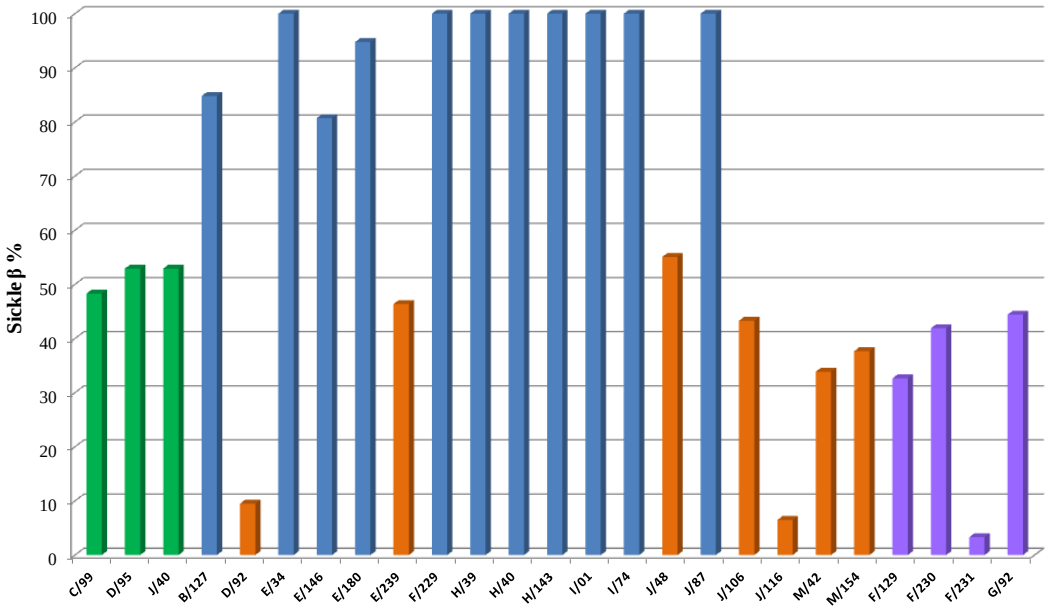


Figure 3.14. Calculated β^{SICKLE}-chain levels for the samples classed as compound heterozygotes for HbS and HbC (green) or homozygous sickle patients (blue), or transfused homozygous patients (purple)
Samples shown in orange show calculated values only around 50% or lower, suggesting the patient only being a heterozygous sickle carrier or given blood transfusion before

The calculated percentages of β^{SICKLE} to all β-chains (wild type and sickle) for all samples classed as homozygous or heterozygous are summarised in Appendix 2 in Table A.2.

Samples that had been classed as compound heterozygote for HbS and HbC variants were also calculated against the mass peak -1 Da shifted relative to normal β-chain mass, giving values around 50%, as expected.

A trend worth mentioning is the slight elevation of the measured δ-chain levels compared to normal levels in healthy individuals. δ-chain levels could be calculated when a sample during intact measurement provided mass peak for the normal β-chain next to the variant β-chain, since δ-chain levels are estimated by ratio calculation to normal β-chain level intensities. Any time a variant β-chain is present, the level of the normal β-chain is lower, and so the calculated δ-chain level relative to the normal β-chain present will be higher, highlighting abnormalities with the sample. Figure 3.15 shows the calculated values were above 5%, apart from three samples J116, F231 and D92, which showed low values of HbS as previously presented. Nine samples had values above 5% but less than 6%, and the rest of the samples had elevated, above 6% δ-chain levels. The 6% limit for elevated levels is suggested as the cut-off limit for the intact-chain level analysis as presented previously. In the case of samples E146 and B127 HbS levels were high (84.8% and 80.7% respectively), and so calculated δ-chain levels against the normal β-chain levels understandably resulted in higher levels (19.2% and 19.8%).

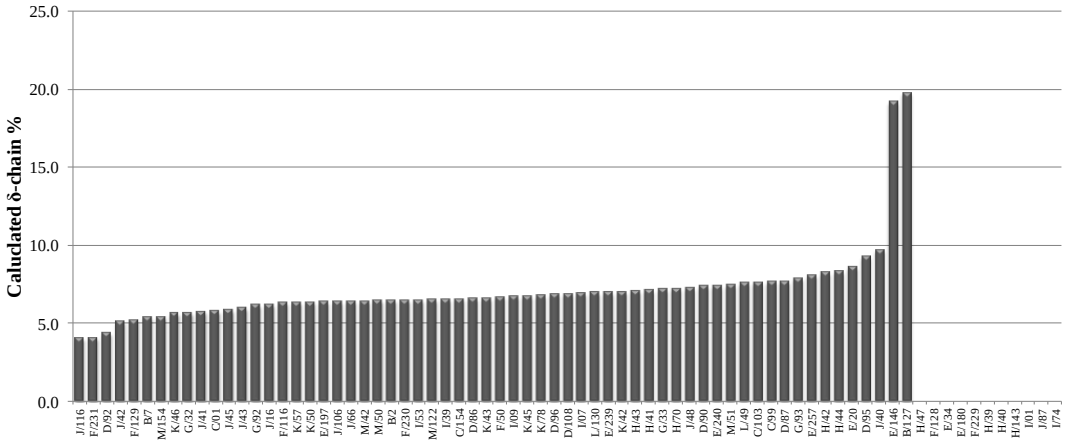


Figure 3.15. Calculated δ-chain levels using the results of the intact-globin chain analysis for the samples classed as homozygous or heterozygous sickle carriers by the hospital laboratory

Indication for the presence of -1 Da variants based on mass error calculations for the β-globin chain

Based on published literature, the presence of -1 Da mass difference variants could be observed by the mass shift of the beta-chain mass caused if the variant is present at 10% or more when using the α-chain mass as internal mass calibrant (Rai *et al.* 2003). For this hypothesis to be tested in intact-globin chain analysis the achievable mass error by the intact globin chain analysis needed to be assessed for all analysed samples within all batches.

Samples with -1 Da Hb variants such as HbD-Punjab, HbE, HbC and HbO-Arab were analysed. The sample numbers with these conditions are summarised in Table 3.10. for all batches separately. There was only one sample with HbE (H45) and one with HbO-Arab (F115) mutation.

Batch	Samples with -1Da variants (HbE, HbO-Arab, HbC, HbD-Punjab)	Numbers
B	-	-
C	C99	1
D	D94, D95	2
E	E143, E151, E255, E256, E258, E259	6
F	F115	1
G	-	-
H	H45, H46, H48, H138	4
I	-	-
J	J40, J44, J46, J69, J140	5
K	K44	1
L	L11	1
M	M121	1
Total		**22**

Table 3.10. Summary of the occurrence of -1 Da variants in the 12 batches analysed.

There were only three batches without any -1 Da variants, and three sample batches contained 4 or more samples with -1 Da variants. The latter three were Batch E, H and J. Calculated mass errors for these batches are shown in Figure 3.16. Equivalent figures for all other sample batches can be found in the appendix.

In sample batch E (sample number 280) there were 6 samples with conditions of one of the -1 Da clinically significant variants. HbD-Punjab mutation in samples E143, E151, E255, E259 and HbC mutation in samples E256 and E258 as summarised in Table 3.11. The mass errors for the β-chain have been calculated using the α-chain as an internal calibrant, assuming it is the correct mass and no -1 Da variants are present affecting the α-chain.

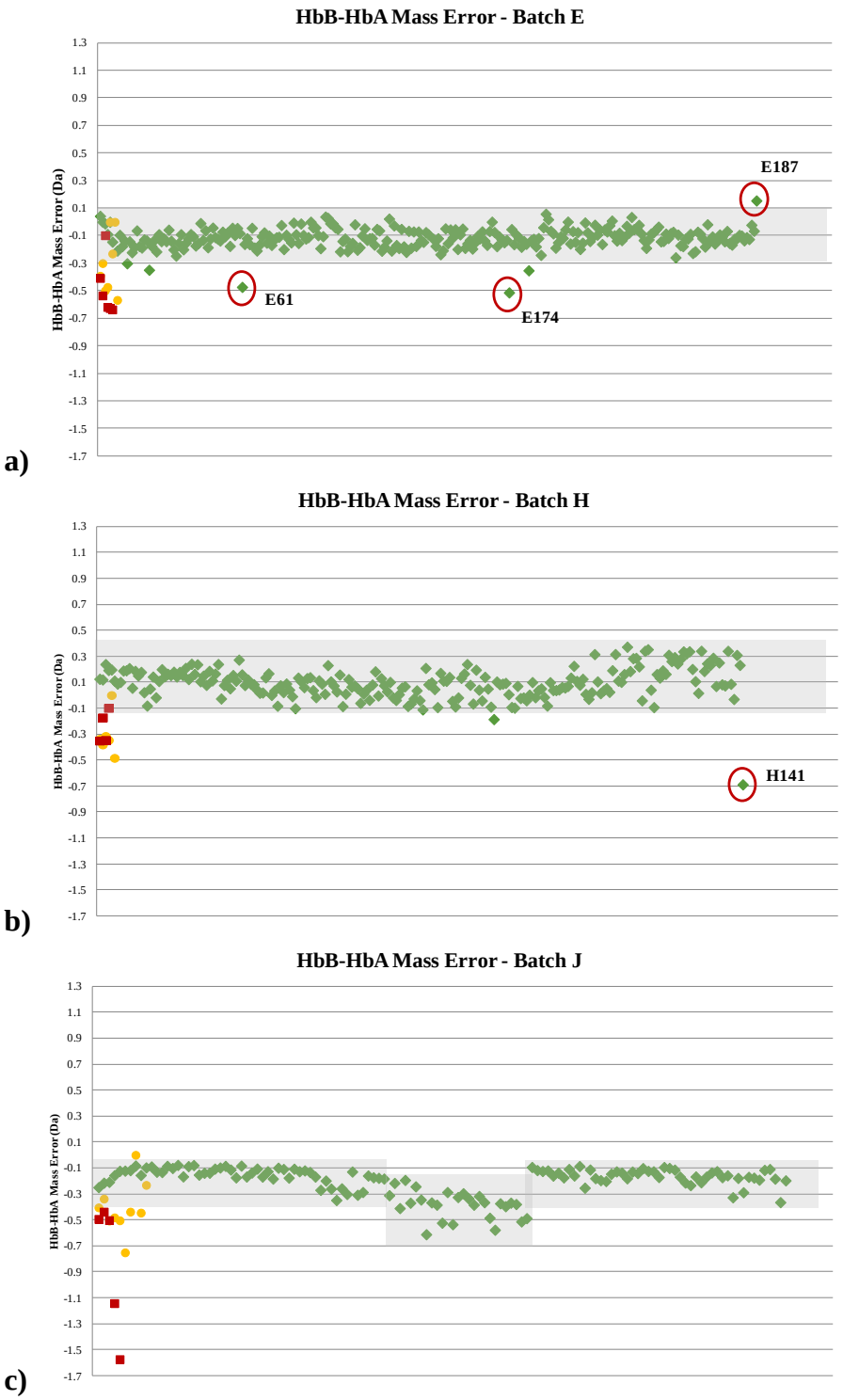


Figure 3.16. Calculated mass errors for the β-chain after corrected by the mass error of the α-chain for Batch E (a), Batch H (b) and Batch J (c)

Green markers represent normal samples, yellow dots correspond to samples with HbS and red squares represent -1 Da variants

In Batch E the majority of the corrected mass errors of the β-chain are in the range +0.1 to -0.3 Da, distributed around -0.1 Da mass error value. This represents a ±0.2 Da window, which is shaded grey in Figure 3.16 a). This ±0.2 Da would be expected to be distributed around 0, but there is a systematic error detected with the mass measurement. Based on the achievable mass accuracy of ±0.2 Da, although shifted because of the systematic error, the hypothesis by Rai *et al.* would provide detection of -1 Da variants if they were present at 20% or more.

In Figure 3.16 yellow dots represent the samples with a diagnosed HbS mutation, the red squares represent the samples with -1 Da variants. In Figure 3.16 a) all -1 Da variants are outside the grey area except one (sample E143), which later was not confirmed by full scan tryptic digested peptide mixture analysis to have a HbD-Punjab mutation. This may suggest that the variant is present only in a low amount, less than 20%, or not present at all. The presence of all five other -1 Da clinically significant variant in Batch E would be observed. There were additional samples with slightly abnormal mass errors. Three of these sample E61, E174 and E187 (circled red) contained variants. Other samples were not detected with any obvious abnormalities. Full scan tryptic digest analysis would be required to reveal any abnormalities. It was also observed that samples presented out of range mass errors if they have been re-measured. In some cases the measurement of a sample had to be repeated the following day, because none or only one of the three replicates resulted in a mass spectrum. The cause of the failed measurements is more likely to be nozzle blockage during sample introduction with the NanoMate chip.

An observed phenomenon is that the majority of the samples with HbS or other rare variants also had mass error outside the grey area. This would suggest that additional peaks close to normal β-chain may have an effect on the detected mass of the normal β-chain in many of these cases. The reason for this is not completely clear. There may be an interaction or coupling between different ion species adjacent to each other on the *m/z* scale. These multiply charged ions may be coupled on the intact level. The sample E187 had a 50% abundant β-chain variant with a +58 Da mass difference and caused observable mass shift of the normal β-chain, resulting in abnormal mass error.

In Figure 3.16 b) the same presentation of the mass errors for sample batch H (sample number 228) is shown. This batch contained 4 samples with -1 Da clinically significant variants: H45 with HbE mutation and H46, H48, H138 with HbD-Punjab

mutation. This time the mass error window for the samples was slightly shifted above 0, and is not constant among all measurements, probably differing due to results for different sample plates. As the sample number in this batch was 228, they could only be analysed by preparing three different 96-well sample plates. The mean mass error difference between plates is visible, but more obviously present in Figure 3.15 c) for batch J discussed later. Samples H46 and H138 are clearly outside the grey area, which is indicative of a -1 Da variant. Sample H48 is on the borderline, which may suggest a variant, but it can also have a slightly larger (shifted into negative direction) mass error. Sample H141 was identified with a chromatographically silent β-chain variant, causing mass shift and abnormal mass error.

Batch J (sample number 144) represents clearly how mass error fluctuations can prevent the detection of -1 Da variants as the central region of the mass error figure illustrates. If we only look at mass errors for normal samples in the first and last section of the figure, the clinically significant variants are clearly separated, and indication for all is observed. The central region would however make an attempt to set a reference range for normal values difficult. The reason of the occurrence of these erroneous values is not yet clear.

Summary for samples diagnosed with -1 Da variants, and the mass error calculations indicating the presence of a variant can be seen in Table 3.11.

All samples with SC conditions have been unambiguously detected using this method. The majority of other conditions have also been detected, although this method still has a level of uncertainty based on different systematic errors observed for the calculated values of different batches.

Systematic error can either come from calibration or from data processing. The calibration was performed with a normal sample from each batch prior to analysis of the entire batch. The use of a normal sample from a batch may have been inappropriate since chromatographically silent variants proved to be present in some samples classed as normal. Future studies would benefit from the use of a known reference material for calibration.

Another possible source for systematic errors is processing method and the application of a lock mass value corresponding to one of the multiply charged ions of the α-chain, or other parameters influencing the calculated mass.

The fluctuation of ranges between sample error measurements among all batches is illustrated by Figure 3.17.

	Variant	Sample	δ-chain level %	Variant %	MassError	Detected/Indication based on Intact MS
1	A O ARAB	F/115	4.41	-	-0.764	Yes
2	AC	E/256	5.3	-	-0.408	Yes
3	AC	E/258	6.5	-	-0.535	Yes
4	AC	J/44	4.8	-	-0.494	Yes
5	AC	J/46	4.5	-	-0.438	Yes
6	AC	J/140	5.5	-	-0.5032	Yes
7	AC	K/44	5.2	-	-0.599	Yes
8	CC	J/69	6.9	-	-1.1431	Yes
9	SC	C/99	7.7	48.3	-1.305	Yes
10	SC	D/95	9.3	52.9	-1.448	Yes
11	SC	J/40	9.7	52.9	-1.575	Yes
12	AE	H/45	7.0	-	-0.0989	No
13	AD	D/94	4.6	-	-0.692	Yes
14	AD	E/143	7.9	-	-0.0976	No
15	AD	E/151	5	-	-0.618	Yes
16	AD	E/255	4.9	-	-0.627	Yes
17	AD	E/259	4.9	-	-0.636	Yes
18	AD	H/46	5.3	-	-0.3509	Yes
19	AD	H/48	5.4	-	-0.1725	Borderline
20	AD	H/138	5.9	-	-0.3463	Yes
21	AD	L/11	4.8	-	-0.5631	Yes
22	AD	M/121	4.0	-	-0.625	Yes

Table 3.11. Summary of -1 Da variants grouped based on the specific mutation conditions (shadings represent different mutations, purple for HbO-Arab, green for HbC, blue for HbE and orange for HbD-Punjab)

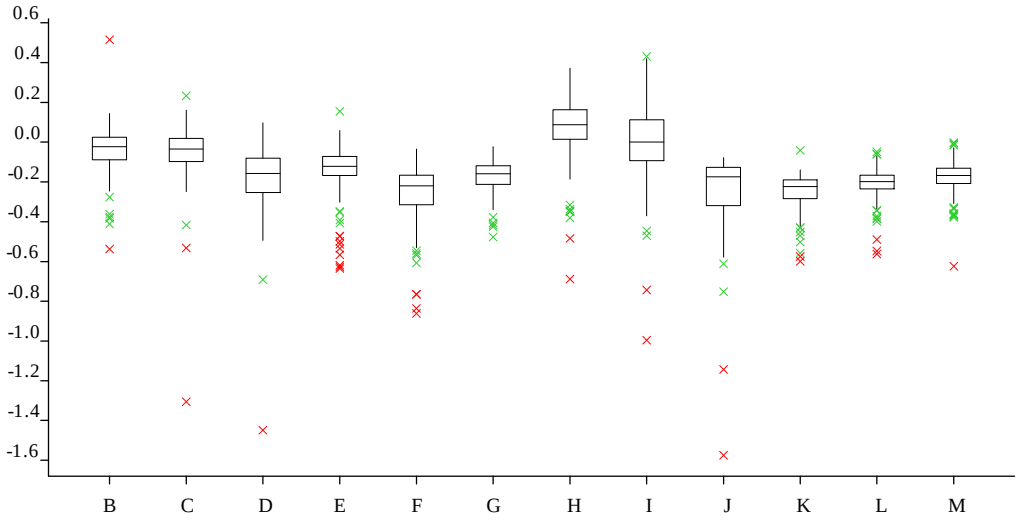


Figure 3.17. Boxplots of error measurement of the β-chain corrected by the α-chain mass error.
Data are shown separately for each batches analysed

The majority of the mass error windows are below 0, but not consistently for the overall sample analysis. The problem highlighted by this Figure is that using mass error values as indication for the -1 Da variants is not an absolute indicator within all measurements at any time. The values are indicative if they are significantly different than average mass error for samples measured around the diseased samples at the same time with the same instrumental conditions and calibration. Measurements within one batch and even within one 96-well plate often did not give completely consistent results, and some of the Figures show this phenomenon. The presented box plot illustration is no longer a working solution which could point out only outliers as was assumed and suggested in the previous method development chapter.

An additional observation worth mentioning, is that samples with elevated δ-chain or HbF levels also had slightly abnormal mass errors.

As mentioned, mass error may significantly depend on the processing method parameters during the algorithm applied to the mass spectrum of multiply charged ions when deconvoluted to true mass scale. Further investigation and optimisation is necessary for the parameters used by the algorithm to be able to provide reliable indication about the presence of -1 Da clinically significant variants, especially if the intact MS method alone is pursued to be used as a first line screening method in the clinical lab.

3.3.2.2 LC-MRM analysis and full scan peptide mixture analysis

Detection of HbS

The detection and confirmation of the presence of predetermined clinically significant variants using the MRM method is straightforward. The peptides monitored are either present or not. The peptides monitored are quantified only if the corresponding wild type peptide is present.

For detection of HbS mutation, the detection of the β^{SICKLE} T1 tryptic peptide was necessary. The wild type T1 peptide was monitored with predefined MRM transitions. The percentage of the HbS variant was calculated by ratio calculations of the peak area of the β^{SICKLE} T1 to the sum of the sickle and normal T1 peptide peak areas. The determined HbS levels are shown in Figure 3.18 for the samples classed as heterozygous sickle carriers. The calculated values are listed in Table A.3 in Appendix 2.

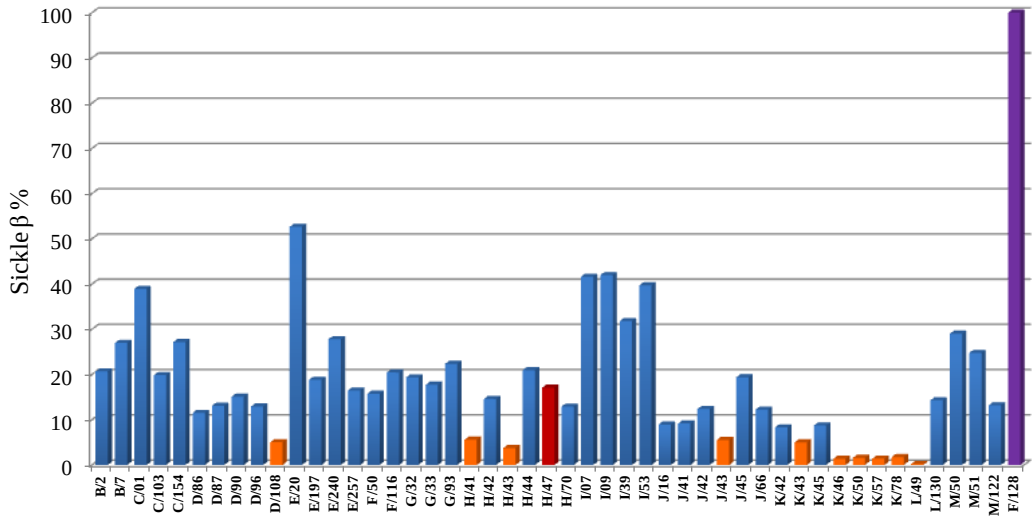


Figure 3.18. Calculated β^{SICKLE}-chain levels based on the results from the MRM method for the samples classed as heterozygous sickle carriers by the hospital laboratory

The sample H47, shown in red was detected with variant beta-chain during intact globin chain analysis. Samples shown in orange had significantly lower values than it was calculated from the intact globin chain analysis. Sample F128 showed 100% β^{SICKLE} since it did not have peak for the wild type β-chain T1 peptide as the individual was a baby with high HbF levels.

Levels determined by the MRM method were significantly lower for the majority of the samples than the levels calculated by the intact globin chain analysis. They are

also not evenly distributed and show more fluctuation. The majority of them are below 20%, which may suggest ionisation or fragmentation efficiency differences between the normal and variant peptides. Samples D108, H141, H43, J43, K43 were significantly lower, and samples K46, K50, K57, K78 resulted in HbS levels of only 1-2%. For sample L49 a level of 0.2% was determined, and sample J116 has not been confirmed to carry the mutation by the MRM method. Although the variant characteristic peak is present and the variant is confirmed by the MRM method in most of the cases, this quantitative inconsistency requires further method optimisation.

Levels for samples classed as SS or transfused SS can be seen in Figure 3.19. Those samples which were identified as homozygous using the intact globin chain analysis had only MRM chromatographic peaks for the β^{SICKLE} T1 tryptic peptide, which together with the occasional low levels of HbS determined by the MRM method suggest an incomplete enzymatic cleavage and production of tryptic peptides. This issue has been noted before, and in case of some of the batches has been shown to have a considerable influence on levels of δ/β-chain peptide ratios.

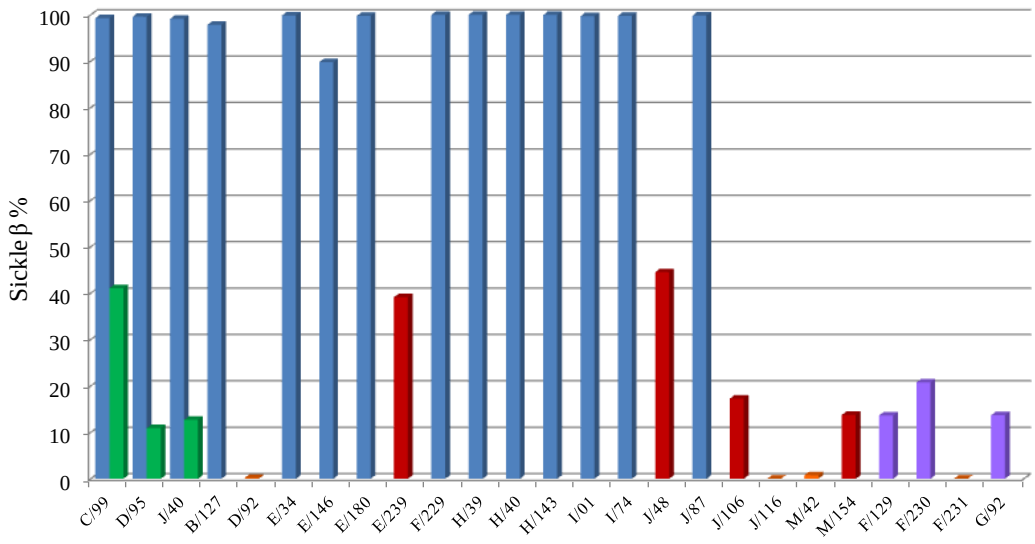


Figure 3.19. Calculated β^{SICKLE}-chain levels based on results from the MRM method for the samples classed as compound heterozygotes for HbS and HbC (recalculated to HbC peptide in green) or homozygous sickle patients (blue), or transfused homozygous patients (purple)

Samples shown in red show calculated values only around 50% or below, suggesting the patient being a heterozygous sickle carrier

A comparison of the HbS quantities determined by the two methods is shown in Figure 3.20. In Batch I, the values determined by both methods are very similar, and in Batch K they are lower when determined by the MRM method and average 30-40% with the intact globin chain analysis. Samples which resulted in low values using intact globin chain analysis had the characteristic peaks in MRM analysis, but would not necessarily be confirmative in a real clinical screening application. The observed batch to batch variation may also support the suggested explanation, that there is an issue with the efficiency of the enzymatic cleavage. During the clinical trial, sample preparation conditions were kept as consistent as possible. The bovine trypsin used - from Sigma Aldrich, however, was not high quality sequencing grade and it was not possible to use only one product with the same LOT number. Batch K and Batch I were treated with bovine trypsin from two different production batches. Also it has to be taken into account that other experimental error may have occurred during the preparation or the measurement of the samples, such as incomplete denaturation of the protein chains which can result in incomplete cleavage. This needs further investigation and further optimisation, if possible, to provide a method which measures HbS values in a reliable manner.

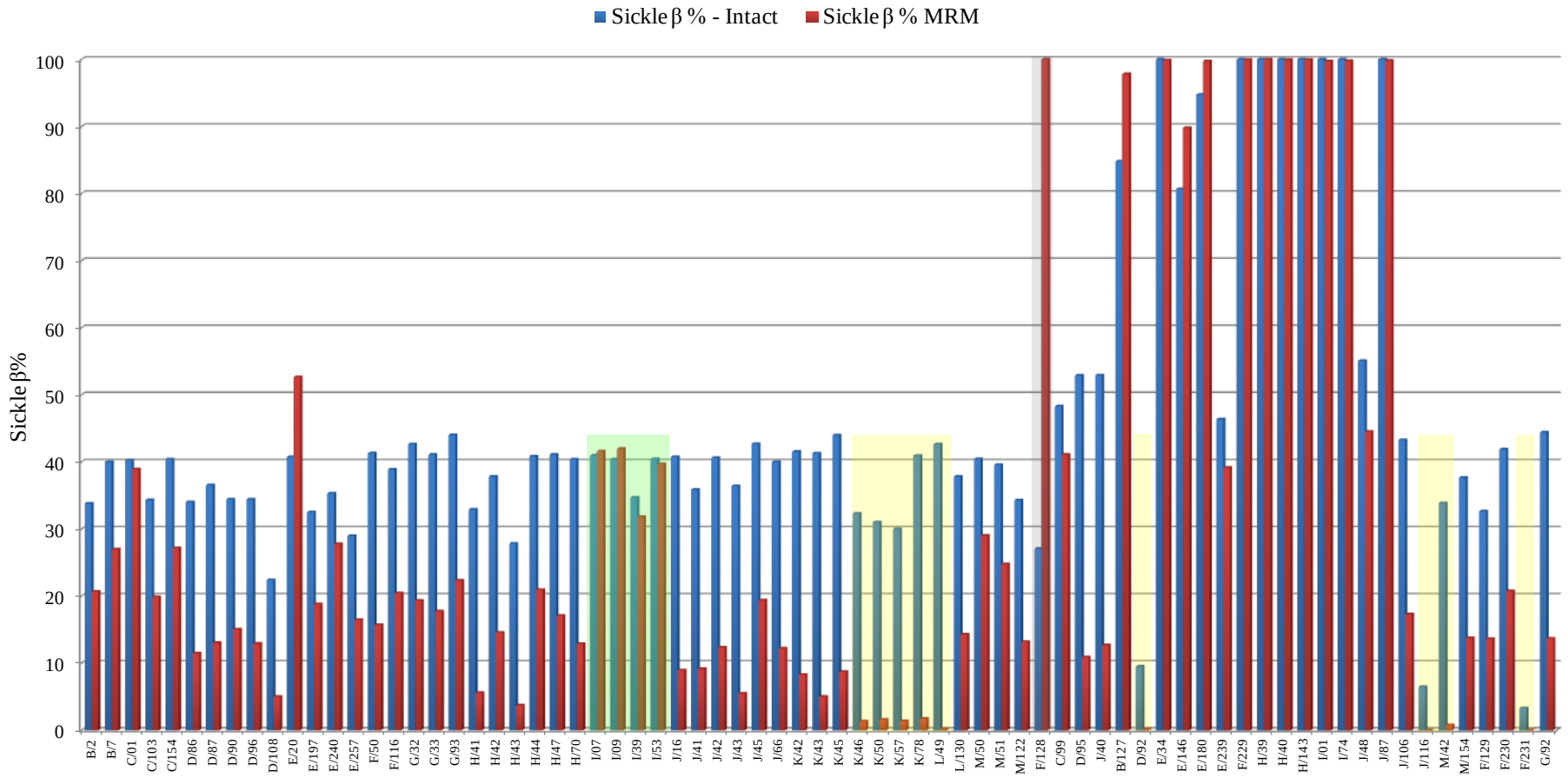


Figure 3.20. Calculated levels of HbS determined by both MS methods for all samples diagnosed with some type of sickle disorder and HbC
Samples highlighted in green resulted in very similar values by both methods, and samples highlighted in yellow resulted in very low levels by the MRM method, disagreement in case of sample F128 is highlighted grey, intact levels were calculated taken HbF levels into account while that was not possible using the MRM method

Detection and confirmation of HbE, HbC and HbO mutations

The MRM method confirmed the presence of all monitored -1 Da clinically significant variants. No peptide was monitored for the confirmation of the HbD-Punjab variant. Calculated variant levels based on peptide ratios are summarised in Table 3.12.

	Hospital number	Hospital comment	Our number	MRM δ-chain peptide ratios %			HbS %	HbC %	HbE	Comment
				T2	T3	T5				
1	19173162	A O ARAB	F/115	1.4	1.9	3.3	-			Only a low intensity peak, not confirmative
2	19083525	AC	E/256	2.9	2.5	2.9	-	21.1		
3	19112053	AC	E/258	4.0	2.9	2.9	-	34.7		
4	19310975	AC	J/44	0.7	2.0	4.8	-	54.2		
5	19429733	AC	J/46	0.4	2.2	4.1	-	50.9		
6	19529055	AC	J/140	1.8	2.2	5.0	-	74.3		
7	19464748	AC	K/44	1.1	1.1	4.7	-	0.9		Variant confirmed, intensity low, possible problems with tryptic digestion
8	19509268	CC	J/69	0.9	2.0	4.7		99.9		
9	18963649	SC	C/99	-	1.9	2.4	99.3	99.4		MRM recalculated 41.1% Sickle and 58.9%HbC
10	18925552	SC	D/95	-	1.0	4.7	99.6	100		MRM recalculated 10.9% Sickle and 89.1% HbC
11	19395725	SC	J/40	1.1	2.3	5.3	99.1	99.9		MRM recalculated 12.7% Sickle and 87.3% HbC
12	19137083	AE	H/45	1.8	4.0	6.0	-		+	+: MRM peak observed, due to RT shifting not quantified

Table 3.12. Detection of -1 Da clinically significant variants by the MRM method.

In case of compound heterozygous HbS and HbC both peptides were observed, and the peptide ratio has been recalculated using the peptide peak areas of the two variants. Values of around 50% would be expected for both variants for these samples, but this is not the case. Sample D95 and J40 contained high values for HbC: 89.1% and 87.3%, respectively. The measurement of sample K44 resulted in a low peak area for the HbC variant, despite the fact that this sample displayed a mass shift for the β-chain mass measurement during the intact globin chain analysis, and the variant is expected to be present at 20% or more. The analysis of batch K for the heterozygous sickle samples has shown erroneous results for HbS levels determined previously and it is suggested that the value calculated for sample K44 is also low

due to the same issue, possibly with the enzymatic cleavage step of the sample preparation, resulting in different levels of the variant characteristic peptides. The same explanation is considered to be true for the variation of the calculated values for heterozygous HbC samples. The observed range is between 21 and 74%, and the theoretically expected 50% is only measured for samples in Batch J, first sample plate including samples J44 and J46.

There was only one sample available from a patient with HbE mutation. That mutation was confirmed based on targeted MRM analysis on its proteotypic peptide, but the proportion of the variant was not quantified. The characteristic peak corresponding to the HbE variant peptide was observed, but due to retention time shifting and selective monitoring of an MRM channel, full peak-shape could not be recorded. The sample was not re-measured, since confirmation of the mutation was already achieved, which was the primary purpose of the analysis.

Confirmation of HbO variant proved to be problematic with the MRM method. A low intensity peak (~500, while δ chain peptide peak intensities were observed at $5x10^4$-$1.4x10^5$, and β-chain peptide peak intensities $1.7x10^6$-$3.7x10^6$) was observed in the monitored MRM chromatogram, which was not present in other samples. To obtain a higher confidence assignment, a full scan analysis of the tryptic peptide was used to prove the presence of the variant.

Recalculated HbC values suggested problems with sample preparation and the experimental procedure. The values determined by intact globin chain analysis suggest 50% for both mutations.

The developed MRM method was capable of confirming the presence of the mutation, but reliable quantification of certain variants needs further method development.

Confirmation of HbD-Punjab mutation with full scan high-resolution MS analysis

Confirmation of the presence of HbD-Punjab variant was only possible by full scan analysis of the tryptically digested peptide mixture. No MRM transition was developed for the variant, which has -1 Da mass difference for the singly charged peptide ion to the wild-type peptide ion and only -0.5 *m/z* mass difference for the doubly charged ions. High resolution mass analysis was used in this instance to

provide a definitive assignment. All samples which were classed as HbD-Punjab carriers have been analysed using a Q-ToF Synapt G2 instrument to check for the presence of the doubly charged variant peptide ion. Table 3.13 summarises findings of this analysis. All samples except one were confirmed as carriers of a HbD-Punjab mutation. The presence of the variant peptide ion was not observed in sample E143. This sample did not show a significant mass error for the β-chain as a result of the intact globin chain analysis, providing a discrepancy between the hospital methods and the MS methods. The variant did not provide a peak corresponding to HbG Philadelphia, which had been mistaken for HbD-Punjab mutation previously. This may be explained by the fact that the variant is either not present, or undetectable by the methods used here due to its low proportion in the sample.

	Hospital number	Hospital comment	Our number	Comment
1	18977269	AD	D/94	Confirmed
2	**19088023**	**AD**	**E/143**	**Peptide is not present, mutation not confirmed**
3	19064197	AD	E/151	Confirmed
4	19119480	AD	E/255	Confirmed
5	19114922	AD	E/259	Confirmed
6	19244683	AD	H/46	Confirmed
7	19213295	AD	H/48	Confirmed
8	19310384	AD	H/138	Confirmed
9	19580534	AD	L/11	Confirmed
10	19696761	AD	M/121	Confirmed

Table 3.13. Confirmation of HbD-Punjab variant by the analysis of tryptically digested peptide mixture in full scan mode

Indication for Hb Lepore variants

During the clinical trial one sample was identified as a Hb Lepore carrier by the hospital laboratory. No obvious abnormalities were observed during the intact globin-chain analysis for this sample, and the expected masses for the different HbLepore variants were not observed. The sample, however, did show significant elevated values for all three monitored δ/β peptide ratios. The calculated percentages were 14.4% from T2 peptide ratios, 10.1% from the T3 peptide ratios and 13.3% from the T5 peptide ratios.

Hb Lepore variants are fusion chains of the β- and δ- globin chains differing in masses from the normal β-chain mass by -2, -30 and -45 Da as described in Chapter 1 section 1.3. Sequence homologies can be seen in Figure 3.21 for all Hb Lepore β-like chains separately. The areas where sequence differences occur between the β- and the δ-chains are highlighted. For all three Lepore fusion chains the T2 (9-17 residues) and T3 (18-30 residues) peptides are the same as the T2 and T3 peptides belonging to the δ-chain. In case of the T5 (41-59 residues) peptide, LBW and LB has the same peptide sequence as the δT5 peptide, but the T5 peptide belonging to the LH fusion chain is the same as the βT5 peptide. This suggests the presence of Hb Lepore-Boston-Washington or Hb Lepore-Baltimore, since all monitored peptide ratios were elevated, including T5 peptide ratios. Further analysis to obtain the ratio between βT10 and δT10 peptides would be beneficial. This peptide covers the amino acid residues from 83 to 95, and the T10 peptide of LBW is the same as the δT10 peptide, while the T10 peptide of the LB is identical to the βT10 peptide. It is important to emphasize that intact globin chain analysis did not reveal mass peaks for any of these variants, and results from the MRM method, if elevated for all monitored peptides at the same time, could be used as indicators for the presence of the Hb Lepore variants.

Figure 3.21. Sequence homology between β-chain (MW 15867.24) δ-chain (MW 15924.32) and Lepore Boston Washington- (LBW, MW 15865.23), Lepore Baltimore- (LB, MW 15822.20) and Lepore Hollandia- (LH, MW 15836.23) fusion chains

3.3.3 Identification of rare variants

Table 3.14 summarises samples identified as containing rare hemoglobin variants which were found during the three month clinical trial.

Samples E61, M75 and M136 were classed as normal with no indication for any abnormalities by the hospital laboratory methods. Samples E187 and F127 were detected with additional peaks by HPLC and unknown bands by gel-electrophoresis. Sample H47 was recorded as heterozygous Sickle carrier only, and sample H141 was detected with elevated HbF level. In five out of seven cases of abnormal samples, the results of the mass spectrometric methods are not in full agreement with current clinical diagnosis methods, highlighting possible incidences of misdiagnosis. An additional two samples classed as normal have been detected with variant peaks of significant mass shift, one α-chain (F179, +14 Da) and one β-chain (E174, -28 Da) variant. These samples have not yet been fully characterised.

Sample	Observed mass shift	Additional Peak(s) (*m/z*)	Variant peptide	Mutation	Identified variant
E61	Beta chain -10 Da	471.8 (+2)	T1	β5(Pro→Ser)	**Hb Tyne**
E187	Beta chain +58 Da	762.80(+2)	T2-T3 (missed cleavage)	β16(Gly→Asp)	**Hb J-Baltimore**
F127	Alpha chain -14 Da	1084.7 (+3)	T9-T10 (missed cleavage)	α90(Lys→Asn)	**Hb J-Broussais**
H47	Beta chain +28 Da	480.8 (+2)	T2	β10(Ala→Val)	**Hb Iraq-Halabja**
H141	Beta chain -28Da	623.95(+2)	T4	β40(Arg→Lys)	**Hb Athens-GA**
M75	Beta chain +18 Da	580.05(+3)	T12	β114(Leu→Met)	**Hb ZengCheng**
M136	Beta chain +14 Da	473.8(+2)	T2	β11(Val→Leu/Ile)	***"Hb Hamilton" or a novel variant***

Table 3.14. Summary of identified rare variants found during the clinical trial

Hb Tyne (Sample E61)

A sample classed as normal by the hospital laboratory methods showed abnormalities during the intact globin chain analysis. The deconvoluted mass

spectrum revealed the presence of a -10 Da β-chain variant (Figure 3.22 a). The calculated variant percentage was 45.1%, and the δ-chain level was present at 7.4% compared to normal β-chain. This is an indication for abnormality, and the mass error of the β-chain was -0.414 Da, which was outside the normal mass error range observed for Batch E.

A -10 Da mass shift can be a result of Pro→Ser substitution. The comparison of tryptic digest peptide mixture of the sample to a control sample showed additional doubly charged peaks at *m/z* 471.77 (Figure 3.22.b). The -10 Da mass shift affecting the T1 peptide would provide this additional peak.

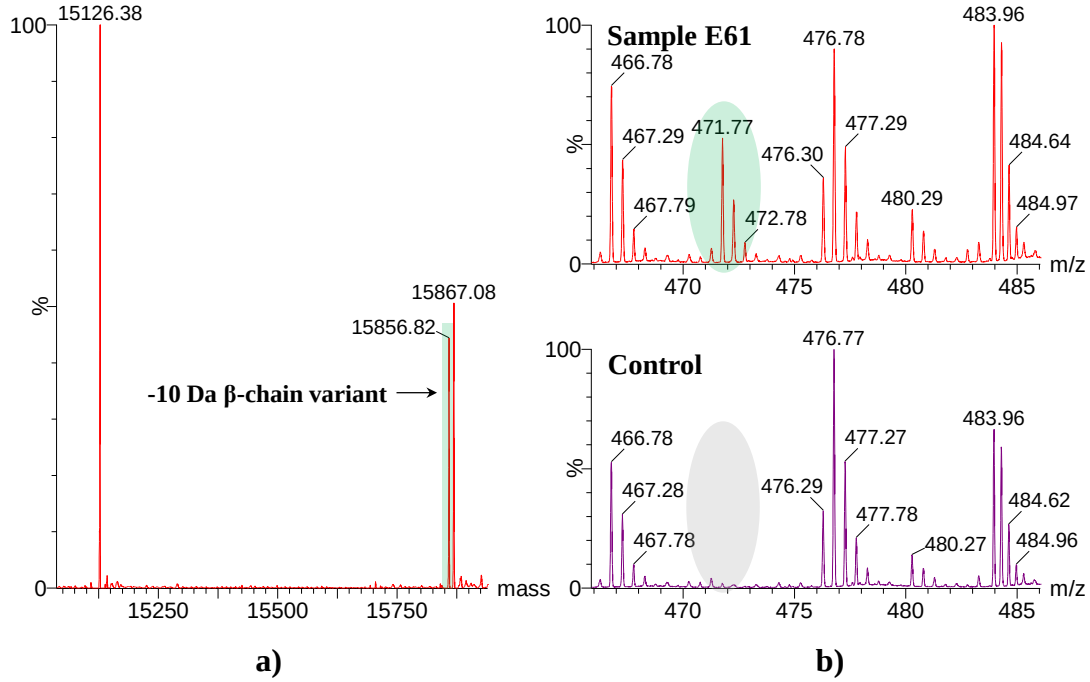


Figure 3.22. Deconvoluted mass spectrum of sample E61 from the intact globin chain analysis (a) and mass spectrum of tryptic digest peptide mixture compared to control (b)

The intact mass spectrum shows an additional peak -10 Da to normal β-chain. The tryptic digest analysis reveals the doubly charged peak at *m/z* 471.77.

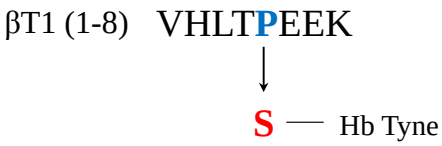


Figure 3.23. Schematic of T1 tryptic peptide of the β-chain and possible amino acid substitution resulting in -10 Da mass shift.

Taking into account the only possible substitution triggered by a single base change and the number of Pro residues in the peptide, the mutation affects the 5th position of the β-chain with a Pro→Ser substitution as illustrated in Figure 3.23. This variant is listed in the HbVar database as **Hb Tyne (β5(Pro→Ser))**.

Hb J-Baltimore (E187)

A sample classed as abnormal based an unknown band by the hospital laboratory methods also showed abnormalities during the intact globin chain analysis. The mass spectrum revealed the presence of a +58 Da β-chain variant (Figure 3.24 a). The calculated variant level was 51%, and the δ-chain levels could not be determined since the variant and the δ-chain masses overlapped. The mass error of the β-chain was 0.1547 Da, which was outside of the normal mass error range for Batch E.

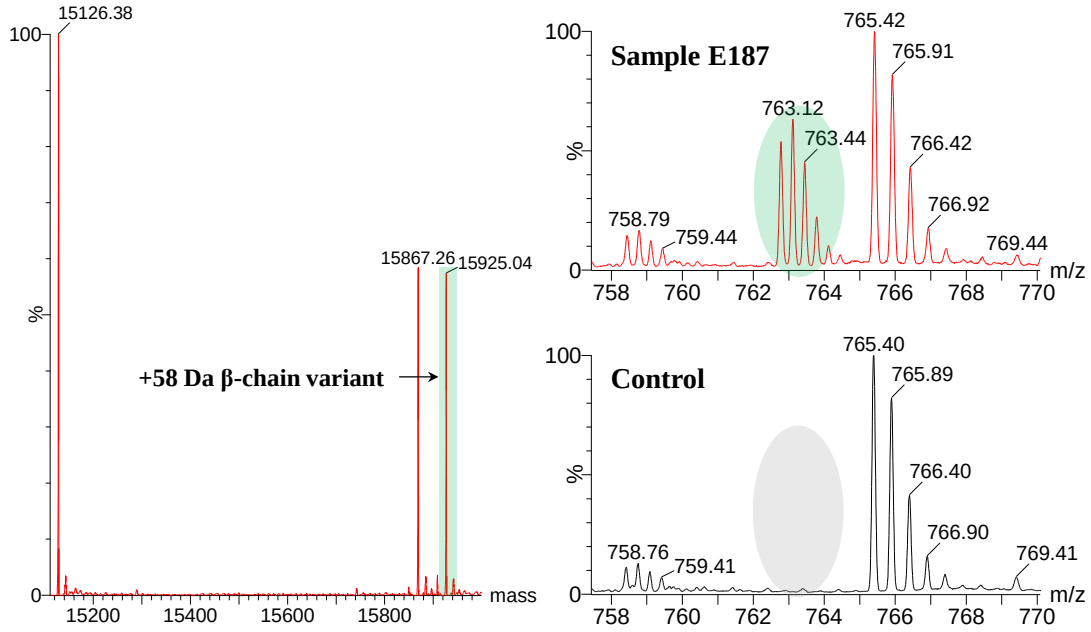


Figure 3.24. Deconvoluted mass spectrum of sample E187 from the intact globin chain analysis (a) and mass spectrum of tryptic digest peptide mixture compared to control (b)

The intact mass spectrum shows an additional peak +58Da to normal β-chain. The tryptic digest analysis reveals the triply charged peak at *m/z* 762.73.

A +58 Da mass shift can be a result of Ala→Glu or Gly→Asp substitutions. A comparison of tryptic digest peptide mixture of the sample to a control sample

showed additional triply charged peaks at *m/z* 762.80 (Figure 3.24 b). The +58 Da mass shift affecting the (T2-T3) peptide, as a result of a missed cleavage, would be responsible for this additional peak.

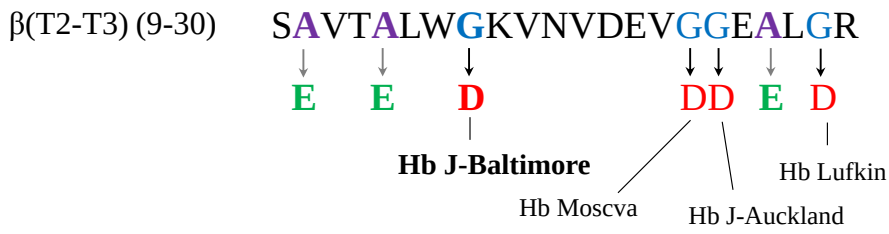


Figure 3.25. Schematic of (T2-T3) tryptic peptide of the β-chain and possible amino acid substitutions resulting in +58 Da mass shift.

There are three Ala and four Gly residues in the (T2-T3) peptide, providing seven possible mutation sites, as illustrated in Figure 3.25. Sequence analysis of the peptide by MS/MS experiment was performed to identify the location of the mutation. The corresponding MSMS spectra can be found in Appendix 3. +58 Da mass shift of the y fragment ion series from the y15 ion and for the b ion series from the b8 ion onwards was observed, which highlights the 8th amino acid residue as the potential mutation site. The residue at position 8 of the wild type (T2-T3) peptide is a Gly residue, which has been substituted in the variant peptide to Asp. The hemoglobin variant present in sample E187 was therefore identified as **Hb J-Baltimore (β16 (Gly→Asp)).**

Hb J-Broussais (F127)

A sample classed as abnormal based on the detection of an unknown band by hospital laboratory methods also showed abnormalities during the intact globin chain analysis. The deconvoluted mass spectrum revealed the presence of a -14 Da α-chain variant (Figure 3.26 a). The calculated variant percentage was 12%, the δ-chain level was not affected, and the mass error of the β-chain was -0.054 Da, which was within the normal range for batch F.

A -14 Da mass shift can be a result of Lys→Asn, Glu→Asp, Ala→Gly, Thr→ Ser, Ile→Val and Leu→Val substitutions. A comparison of the tryptic digest peptide

mixture of the sample to a control sample showed additional triply charged peaks at *m/z* 1084.67 (Figure 3.26 b). The indicated variant peptide is α(T9-T10). Since the -14 Da mass shift can be a result of a Lys→Asn substitution which provides a longer peptide based on a missed cleavage point, it is strongly suggested this mutation has taken place. There are several other possible mutation sites as shown in Figure 3.27.

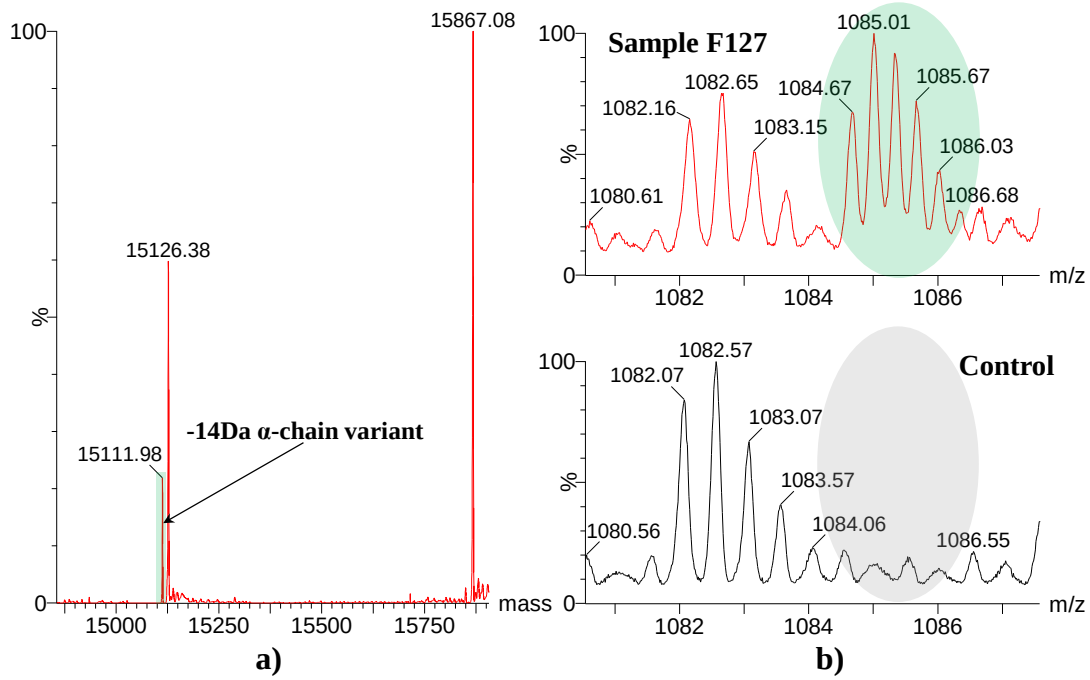


Figure 3.26. Deconvoluted mass spectrum of sample F127 from the intact globin chain analysis (a) and tryptic digest peptide mixture compared to control (b)
The intact mass spectrum shows an additional peak -14 Da to normal α-chain. The tryptic digest analysis reveals the triply charged peak at *m/z* 1084.67.

Figure 3.27. Schematic of (T9-T10) tryptic peptide of the α-chain and possible amino acid substitutions resulting in -14 Da mass shift.

To verify the point of the mutation sequence analysis using MS/MS experiments was performed. The y ion series of the mutant peptide showed a mass shift starting from y3 fragment ions, which confirms the mutation Lys to Arg at position 29 of the

peptide, and position 90 of the α-chain. The hemoglobin variant present in sample F127 was identified as **Hb J-Broussais (α90(Lys→Asn))**.

Hb Iraq Halabja (H47)

A sample classed as heterozygous sickle carrier by the hospital laboratory methods showed additional abnormalities during the intact globin chain analysis. The deconvoluted mass spectrum revealed the presence of a +28 Da β-chain variant (Figure 3.28 a), which was not detected by current clinical methods. The calculated variant percentage was 58.8% as a ratio to β^{SICKLE}-chain, the δ-chain level could not be calculated, which is an indication for abnormality, and the mass error of the β-chain also could not be calculated.

A +28 Da mass shift can be a result of Ala→Val, Gln→Arg, Lys→Arg substitutions. The comparison of the tryptic digest peptide mixture of the sample to a control sample showed additional doubly charged peaks at *m/z* 480.84 (Figure 3.28 b). This mass identifies the affected variant as βT2 with a +28 Da mass shift.

A Gln→Arg substitution is least likely, since this would have resulted in a new tryptic peptide based on the introduction of an additional cleavage point. Ala→Val and Lys→Arg substitutions would not have resulted in new tryptic peptides, only a variant peptide with +28 Da mass shift. Possible mutations affecting the T2 peptide are illustrated in Figure 3.29. To confirm the mutation point, sequence analysis using MS/MS was performed. The y ion series of the mutant peptide showed a mass shift from y8 and for b ion series from the b2 fragment ion, which confirms the mutation Ala to Val at position 2 of the peptide, and position 10 of the β-chain. The hemoglobin variant present in sample H47 was identified as **Hb Iraq-Halabja (β10(Ala→Val))**.

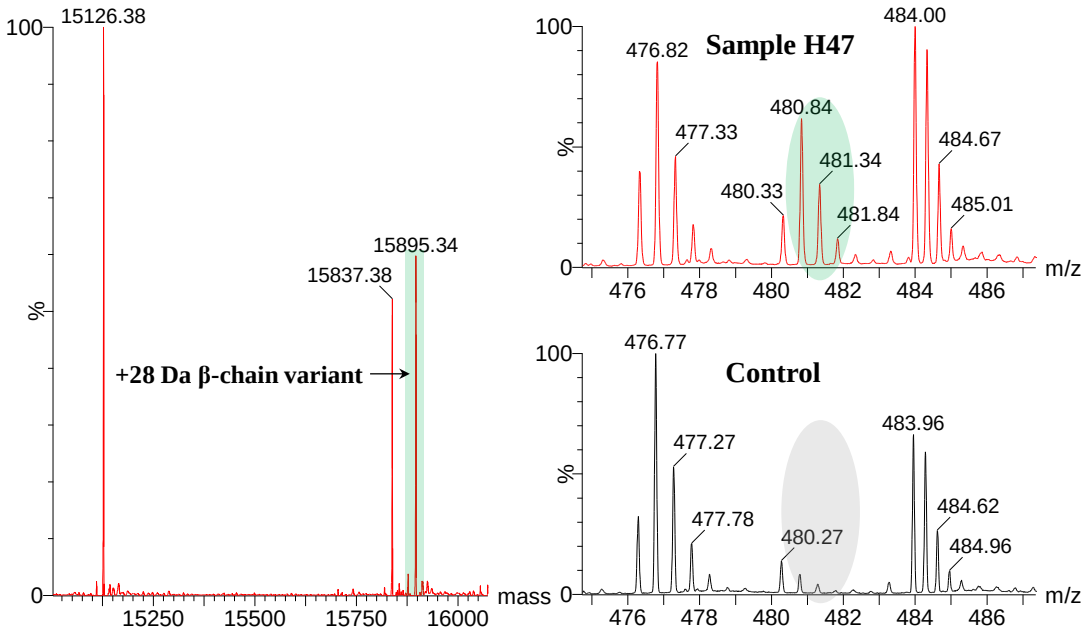


Figure 3.28. Deconvoluted mass spectrum of sample H47 from the intact globin chain analysis (a) and tryptic digest peptide mixture compared to control (b)
The intact mass spectrum shows an additional peak +28 Da to normal β-chain. The tryptic digest analysis reveals additional doubly charged peak at *m/z* 480.84.

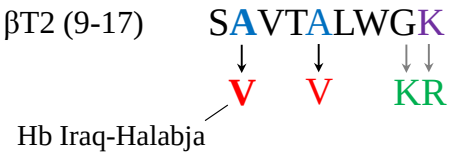


Figure 3.29. Schematic of T2 tryptic peptide of the β-chain and possible amino acid substitutions resulting in +28 Da mass shift.

Hb Athens-GA (H141)

A sample detected with an elevated HbF level using the hospital laboratory methods, which can be an indication of thalassemia, showed additional abnormalities during the intact globin chain analysis. The deconvoluted mass spectrum revealed the presence of a -28 Da β-chain variant (Figure 3.30 a), which was detected by the current clinical methods. The calculated variant percentage was 45.4%, the δ-chain level was present at 7.8% compared to normal β-chain, which is an indication for abnormality, and the mass error of the β-chain was -0.6178 Da, which was outside of the normal mass error range for Batch H, also indicative of an abnormality.

A -28 Da mass shift can be a result of Val→Ala, Arg→Gln, Arg→Lys substitutions. The comparison of the tryptic digest peptide mixture of the sample to a control sample showed additional doubly charged peaks at *m/z* 623.95 (Figure 3.30 b). This mass identified the affected variant as βT4 with a -28 Da mass shift.

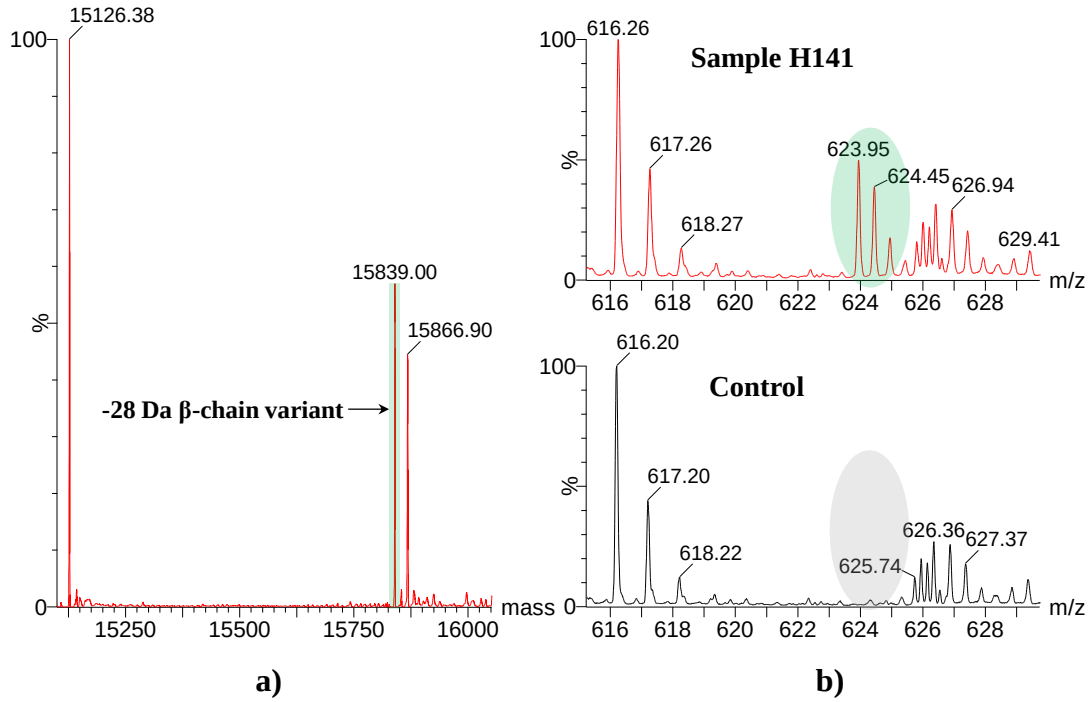


Figure 3.30. Deconvoluted mass spectrum of sample H141 from the intact globin chain analysis (a) and tryptic digest peptide mixture compared to control (b)
The intact mass spectrum shows an additional peak -28 Da to normal β-chain. The tryptic digest analysis reveals additional doubly charged peak at *m/z* 623.95.

There are two Val residues and one Arg residue in the βT4 peptide, providing three possible mutation sites, as illustrated in Figure 3.31. Sequence analysis of the peptide using MS/MS was performed to identify the location of the mutation. The -28 Da mass shift of the whole y fragment ion series was observed compared to normal peptide sequence, which suggests the last amino acid residue as the potential mutation site. The Arg residue was substituted in the variant peptide to Lys. The hemoglobin variant present in sample H141 was identified as **Hb Athens-GA (β40(Arg→Lys)).**

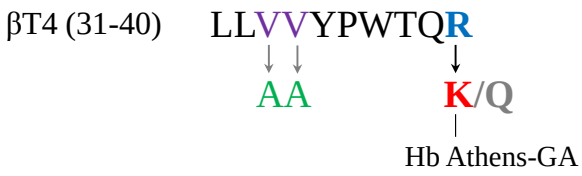


Figure 3.31. Schematic of T4 tryptic peptide of the β-chain and possible amino acid substitutions resulting in -28 Da mass shift.

Hb ZengCheng (M75)

A sample classed as normal by the hospital laboratory methods showed abnormalities during the intact globin chain analysis. The deconvoluted mass spectrum revealed the presence of a +18 Da β-chain variant (Figure 3. 32 a).

The calculated variant percentage was 49.9%, the δ-chain level was present at 11.2% compared to normal β-chain, which is an indication for abnormality, and the mass error of the β-chain was -0.365 Da, which was slightly outside of the normal mass error range for Batch M.

A +18 Da mass shift can be a result of Ile→Met and Leu→Met substitutions, however, there are no Ile residues in the β-chain, which only leaves the possibility for Leu→Met substitution. The comparison of the tryptic digest peptide mixture of the sample to a control sample showed additional triply charged peaks at *m/z* 580.05 (Figure 3.32 b). The +18 Da mass shift affecting the T12 peptide would give this additional peak.

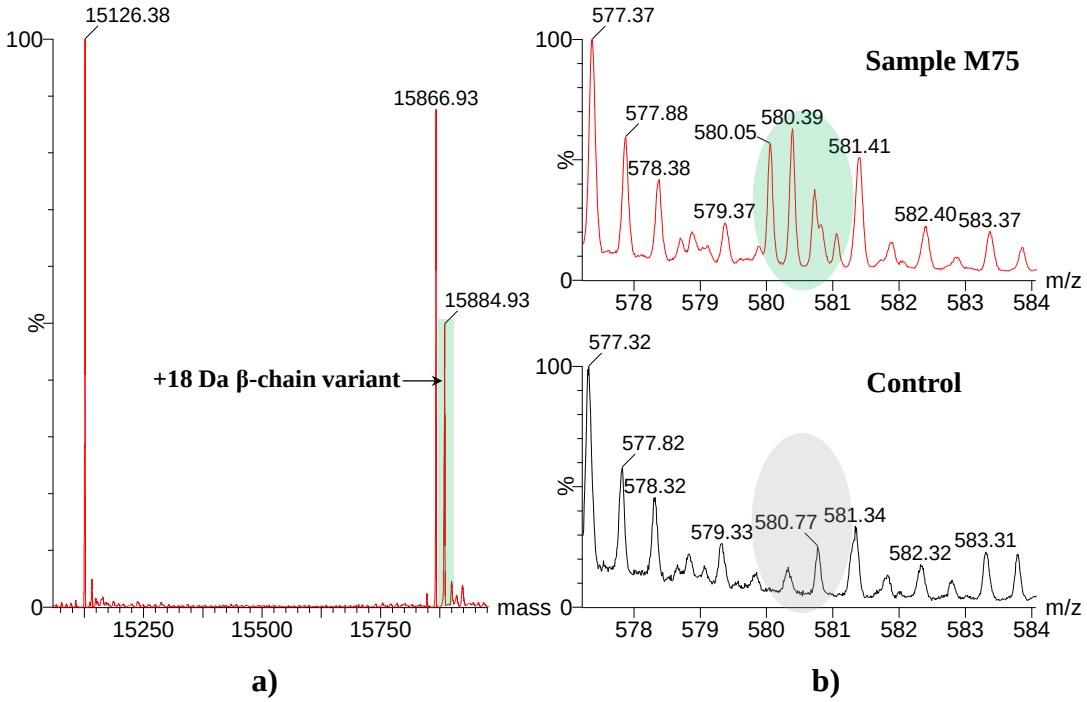


Figure 3.32. Deconvoluted mass spectrum of sample M75 from the intact globin chain analysis (a) and tryptic digest peptide mixture compared to control (b)
The intact mass spectrum shows an additional peak +18 Da to normal β-chain. The tryptic digest analysis reveals additional triply charged peak at *m/z* 580.05.

There are four Leu residues in the βT12 peptide, providing four possible mutation sites, as illustrated in Figure 3.33. Sequence analysis of the peptide using MS/MS was performed to identify the location of the mutation.

The y series showed a mass shift from the 7th y ion, confirming the point of the mutation as the fourth Leu residue in the peptide, affecting position 114 of the β-chain.

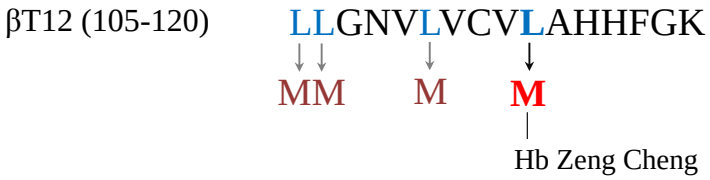


Figure 3.33. Schematic of T12 tryptic peptide of the β-chain and possible amino acid substitutions resulting in +18 Da mass shift.

The hemoglobin variant present in sample M75 was identified as **Hb Zeng Cheng (β114(Leu→Met)).**

Hb Hamilton or a novel, previously not detected variant (M136)

A sample classed as normal by the hospital laboratory methods showed abnormalities during the intact globin chain analysis. The deconvoluted mass spectrum revealed the presence of a +14 Da β-chain variant (Figure 3.34 a).

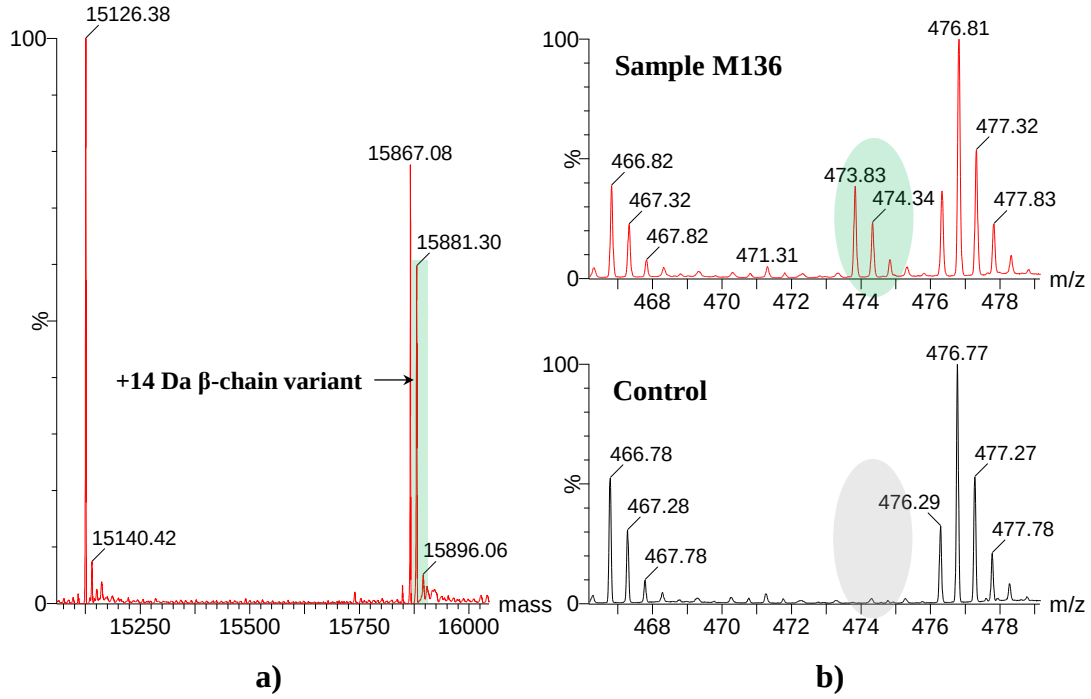


Figure 3.34. Deconvoluted mass spectrum of sample M136 from the intact globin chain analysis (a) and tryptic digest peptide mixture compared to control (b)
The intact mass spectrum shows an additional peak +14 Da to normal β-chain. The tryptic digest analysis reveals additional doubly charged peak at *m/z* 473.83.

The calculated variant percentage was 51.7%, the δ-chain level was 12.2% compared to normal β-chain, which is an indication for abnormality, and the mass error of the β-chain was -0.0116 Da, which is slightly outside of the normal mass error range for Batch M.

A +14 Da mass shift can be a result of Asn→Lys, Asp→Glu, Gly→Ala, Ser→Thr, Val→Ile and Val→Leu substitutions. The comparison of tryptic digest peptide mixture of the sample to a control sample showed additional doubly charged peaks at *m/z* 473.83 (Figure 3.34.b). The +14 Da mass shift affecting the T2 peptide would give this additional peak.

There are three possible mutation sites resulting in a +14 Da mass shift, these are illustrated in Figure 3.35.

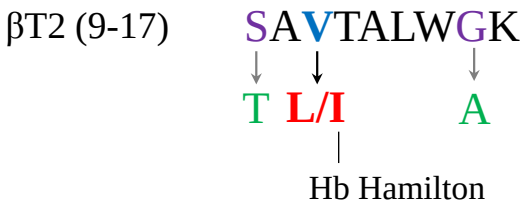


Figure 3.35. Schematic of T2 tryptic peptide of the β-chain and possible amino acid substitutions resulting in +14 Da mass shift.

Sequence analysis of the variant peptide by MS/MS experiment was performed to identify the location of the mutation. The +14 Da mass shift of the 3rd fragment ion of the b fragment ion series was observed, which highlights the 3rd residue, Val, as the potential mutation site. The Val residue can be substituted for either Leu or Ile. At this stage it was not possible to identify the exact substitution. Further experiments will be necessary to elucidate this.

The hemoglobin variant present in sample M136 has yet to be fully resolved, but the mutation site was identified. According to the HbVar database if the mutation is Val→Ile, then the variant present is **Hb Hamilton (β11(Val→Ile)).** If the mutation is Val→Leu, then the variant is a novel variant which cannot be found in the database.

Two further samples E174 and F179 have been observed with additional mass peaks during data analysis of the results from the intact globin chain analysis. Sample E174 showed also abnormal mass error for the β-chain, and an additional peak -28 Da to normal β-chain. Sample F179 showed abnormal mass errors for both chains, and an additional peak +14 Da to normal α-chain. These two samples, together with sample M136, require further investigation.

3.3.4 Determination of Glycated Hemoglobin levels

From intact globin chain analysis quantitation of glycated globin chains has been obtained. HbA_{1c} levels were estimated in a similar manner to HbA_2 levels. Unfortunately there were no measured values available for the glycation levels from the hospital laboratory, but the data is readily accessible with the same intact globin chain analysis and it has been calculated for 10 of the sample batches measured during the trial.

Previous papers reported average values using the determined glycation levels for the different globin chains (Roberts *et al.* 1997). As previously discussed in Chapter 2, the assumption of the rate of glycation being the same on both chains may be valid, but based on the definition of standard measurement for glycated hemoglobin, the beta-chain is considered as primarily glycated and it is referred to as HbA_{1c} (Peterson *et al.* 1998). When analysing the reference standard ClinTest with a known HbA_{1c} level of 9.3% the beta chain-glycation level is below the expected, and the average glycation level is even lower, since the α-chain glycation level was observed at about 1/3 of the beta chain glycation. Correlation between alpha and beta glycation levels are usually satisfying, but not excellent. It was therefore decided to monitor and utilise only calculated intensities based on the peaks for beta-chain glycation.

The coefficient of variation of the quantitation of the different glycated globin chains was calculated based on three replicates for all samples, and plotted as a frequency histogram shown in Figure 3.36. The calculated average CV value was 7.4, and 89% of the values are below 15%. The method can be considered reproducible.

Since there are no reference values from the hospital for the glycation levels, only the distribution of normal and possibly elevated levels of glycated hemoglobin can be assessed. How the determined β^G glycation levels are distributed among the 10 batches is shown in Figure 3.37.

The ClinTest standard was measured as part of the 96 samples within each analysed plate. The calculated values for the standard (each plate three replicates) were used to provide an average level obtained for the standard with a known value of 9.3 %. The average β^G% value obtained for the standard was 8%. Based on reference ranges set for HbA_{1c}, normal levels are considered to be below 6%, 6-6.5% potential prediabetic and values above 6.5% are classed as elevated, and require further investigation (www.diabeticretinopathy.org.uk). With a simple ratio calculation using the standard expected and measured values to determine a correction factor, the 6.5% level limit for the intact MS measurement would correspond to a measured value of 5.6%.

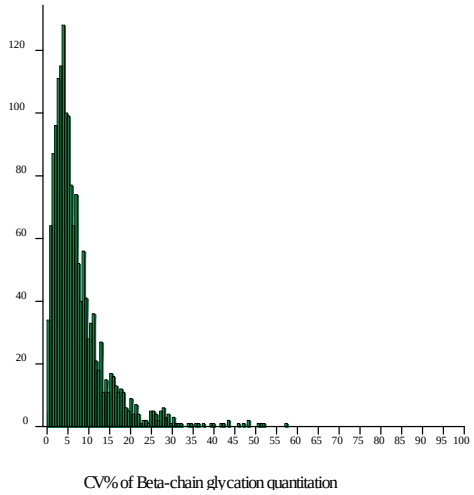


Figure 3.36. Histogram of calculated cv% values for the measurement of beta-chain glycation levels

CV% values are calculated based on three technical replicates used for quantitation of glycated globin chain

Figure 3.37. illustrates a distribution that resemble those observed in HbA_2 level quantitation. The majority of the values are distributed in a certain range, and only a small number of samples are reported with elevated glycation levels. A 5.6% limit (shaded blue range in the figure) would be acceptable for most of the batches as a threshold. In case of batch E and batch H, a slightly higher limit of 6% would be more appropriate.

These measurements, although not compared to reference values from the hospital have proved the potential applicability of the intact globin chain analysis method to provide reproducible additional information about the quantity of glycated hemoglobin. This method could be used for hemoglobinopathy screening and as a diagnostic method for HbA_{1c} determination. The fact that the values are distributed in the same range supports the notion that this method could give consistent estimation of β^G % among batches. The application of a suitable calibration and the analysis of bigger sample sets could develop this method further for providing a reliable evaluation of a potentially important biomarker.

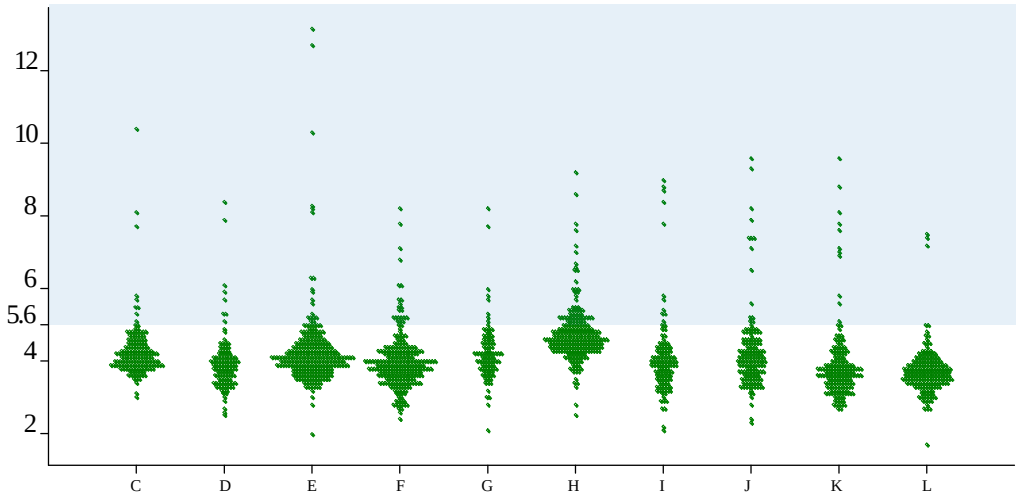


Figure 3.37. Distribution of measured beta glycated values for the different sample batches analysed.
Shaded area above 5.6% represents the suggested range for elevated values based on ratio calculation between measured and expected values for the reference standard.

3.3.5 Detection of elevated HbF levels

For detection of the disorder termed hereditary persistence of fetal hemoglobin (HPFH) the monitoring of elevated HbF levels is important. Elevated HbF levels can also indicate the presence of some thalassemic disorders, as it is often elevated when a patient has β-thalassemia.

There are two action points set for HbF, if MCH > 27 pg the action point is >10%, if MCH<27pg the action is >5% (NHS 2012b). Samples detected with these elevated HbF levels are further investigated in the clinical lab.

For fetal hemoglobin normal levels were detected as 1-2% by the hospital lab. HbF levels were not calculated for these samples, since mass peaks were not observed in every case. HbF levels for samples with 3% or higher reported HbF% were calculated and plotted against the ce-HPLC values in Figure 3.38. The calculated HbF values were obtained by adding together the percentages calculated for both fetal chains γ_1^G (MW 15995.3) and γ_2^A (MW 16009.3).

It can be seen that the correlation between the HbF levels determined by the intact globin chain analysis and the ce-HPLC values is high with an R^2 value of 0.974. This high correlation and the ability to pick up samples with levels above the previously mentioned action points (5 and 10%) suggest the intact globin chain analysis can be

used for monitoring HbF levels, which is necessary to fulfil the NHS requirements for the method wished to be applied for hemoglobinopathy screening.

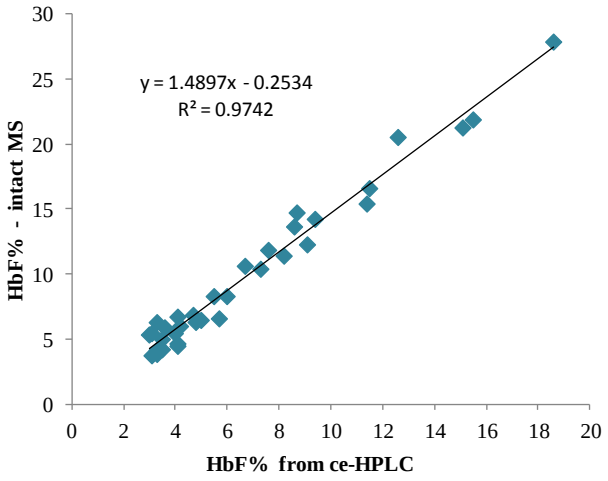


Figure 3.38. Calculated HbF% levels based on intact globin chain analysis plotted against ce-HPLC values

The calculated coefficient variations for the quantitation of the γ-chains are illustrated in histograms for both chains separately in Figure 3.39. Apart from a few outliers the majority of the values are below the required 15%.

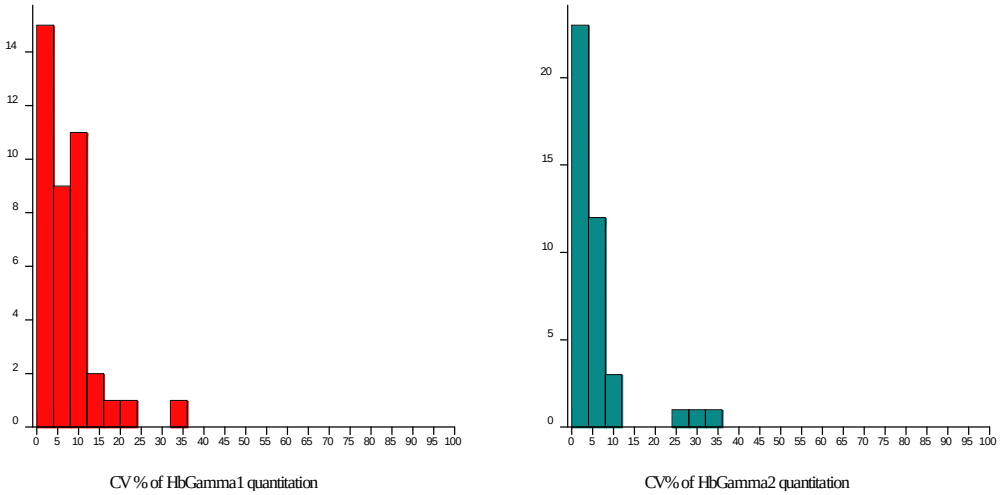


Figure 3.39. Histogram of the calculated coefficient of variation values for the quantitation of the different fetal chains, for (a) γ_1^G and for (b) γ_2^A

Based on information received from the hospital three samples were reported with abnormally elevated HbF levels. In other cases where the HbF level was above 5%, this can be explained by the presence of a clinically significant variant, or it could be that the sample was from a baby. The three samples identified are G49, H141 and J68. In the case of G49 only the mass peak corresponding to γ_2^A (MW 16,009.3) was observed at 8.3%, which is considered elevated. The reported hospital value was 7.3%. Sample H141 was measured at 11.4% for HbF by intact MS and the ce-HPLC value was 8.2%. This sample, however, was also identified as a chromatographically silent Hb variant, and the intact MS HbF levels have been calculated as a ratio of the γ-chains to the sum of the wild type and variant mass peak areas, still showing a 11.4% elevated level. In case of sample J68 the reported value is as high as 12.6% and the combined percentages of the γ-chains provided a HbF level of 20.6%.

Two further samples were diagnosed by hospital methods with HPFH. These samples H140 and I114 both resulted in elevated HbF levels, 14.3% and 14.7% respectively (9.4% and 8.7% ce-HPLC values) using the intact MS method.

3.3.6 Summary of the results from Part I

A total of 2017 samples were analysed during the clinical trial. The numbers of samples per batch detected with abnormalities either by the hospital or by MS methods are listed in Table 3.15.

Batch	Sample number	Number of samples with clinically significant conditions*	Number of variants only observed by MS	Number of abnormalities not confirmed by MS	Unknown band observed by ce-HPLC
B	136	3	-	-	
C	155	4	-	-	
D	108	8	-		
E	280	14	2 (E61, E174)	1 (A_2Var) **1 (E143,HbD)**	1 (E187)
F	251	8	1 (F179)	1 (A_2Var)	1(F127)
G	93	5	-	1 (A_2Var)	
H	228	14	2 (H47, H141)	3 (A_2Var)	
I	114	8	-		
J	144	16	-	3 (Alpha zero thalassemia) 1(Coeluting variant)	
K	166	8	-	1 (A_2Var)	
L	180	3	-	-	
M	162	6	2 (M75, M136)	1 (A_2Var)	
Sum	**2017**	**97 (4.56%)**	**7 (0.35%)**	**13(0.6%)**	**2 (0.1%)**

Table 3.15. Table to summarise abnormal samples analysed in the clinical trial
*Clinically significant conditions here include HbS, HbC, HbD-Punjab, HbE, HbO-Arab, Hb Lepore, elevated HbF, Beta-thalassemia and HPFH

The following lists differences observed that occurred between hospital method findings and the MS methods.

- There were 7 samples with variants identified which were only detectable using the intact MS method.

- 13 samples which were classed as abnormal by the hospital were not confirmed using the MS methods. One of these was a sample with HbD-Punjab mutation, which is one of the five clinically significant variants.
- The samples from patients with α^0-thalassemia did not show any obvious abnormalities with the MS methods.
- The MS methods did not screen for the presence of low intensity δ-chain variants.

- Two samples with hemoglobin variants that were detected as abnormal due to the presence of unknown peaks by the chromatography method and with unknown bands by gel-electrophoresis were characterised by MS.
- Samples with HPFH showed elevated HbF levels during the intact globin chain analysis, which could be used for indication of this condition in the future.
- Samples with elevated HbA_2 levels were correctly identified.
- One Hb Lepore sample did not show abnormalities during intact globin chain analysis, but showed significantly elevated levels of δ-chain levels for all monitored peptide ratios with the MRM method.
- Mass errors of β-chain were indicative of the presence of -1 Da variants, and also other abnormalities, when abnormal globin chain with higher mass difference was present. Elevated δ-chain levels were also indicative for the presence of variant β-globin chains.

Table 3.16 summarises the samples with different conditions, calculating their proportion in all analysed samples during the clinical trial. It is also stated whether either of the mass spectrometry methods were able to detect, identify or confirm the presence of an abnormality. As mentioned before there was one clinically significant structural variant which was not confirmed by MS. Some samples classed as SS (homozygous sickle carriers) were detected with only ~50% of HbS with the presence of normal β-globin chain. It is suspected that in these cases the patient received a blood transfusion, which would explain the presence of the normal β-chain.

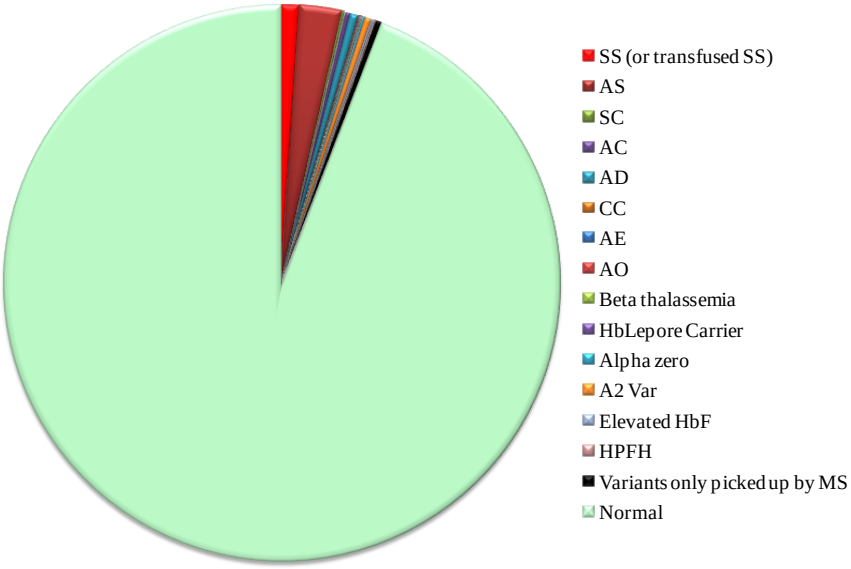
Variant condition		Numbers	Proportion	Identified by MS?
SS(or transfused SS)		22	**1.09%**	All, one not confirmed by MRM
AS		48	**2.38%**	All
SC		3	**0.15%**	All
AC		6	**0.3%**	All
AD		10	**0.5%**	9 confirmed, 1 is not confirmed
CC		1	**0.05%**	All
AE		1	**0.05%**	Confirmed by MRM
AO		1	**0.05%**	All- full scan tryptic digest analysis
Beta thalassemia		1	**0.05%**	Indication
HbLepore Carrier		1	**0.05%**	Strong indication from MRM
HPFH		2	**0.1%**	Elevated HbF detected
Elevated HbF		3	**0.15%**	All
Other abnormalities picked up by HPLC*	Alpha zero thal	3	**0.15%**	Not investigated
	A_2 Var (δ-chain variant)	8	**0.4%**	Not investigated
Variants only picked up by MS		7	**0.35%**	All
Normal		1898	**94.20%**	

Table 3.16. Summary of different conditions, numbers and proportions of clinically significant variants

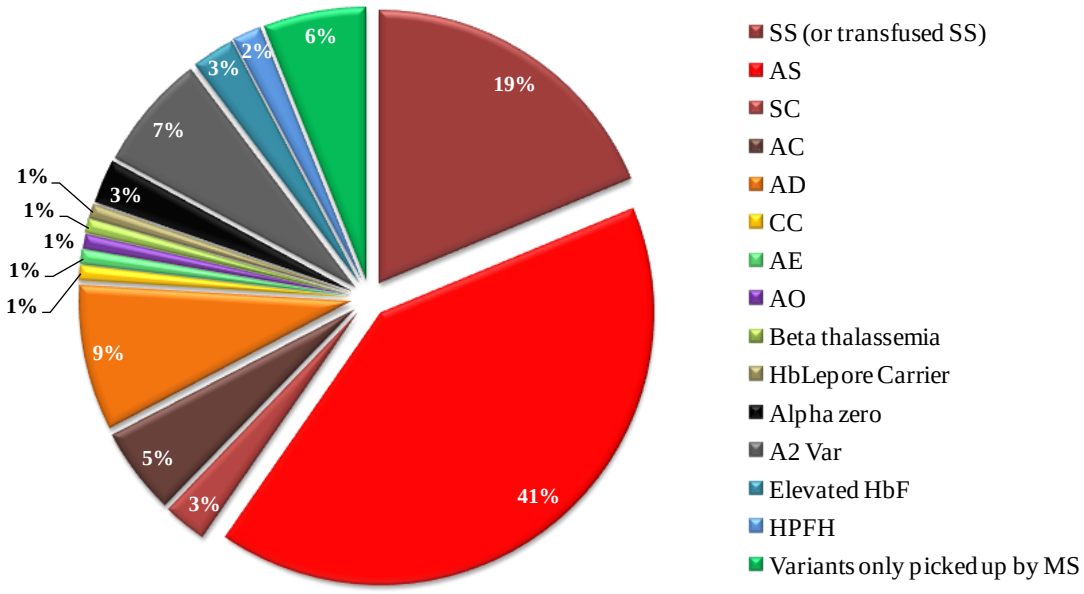
117 abnormal samples were identified, which is 5.8% of the total analysed samples. A graphical presentation of the proportion of hemoglobin disorders is shown in Figure 3.40 a.

The proportion of different conditions within the abnormal samples shows that 63% were carriers of sickle β globin chain either in heterozygous or in homozygous status (Figure 3.40 b) This high proportion of sickle mutation was expected based on the incidence of hemoglobinopathies reported in the literature.

Only one sample was diagnosed as HbE and only one as HbO-Arab. The HbD-Punjab mutation is represented by a higher number of samples than would be expected.



a)



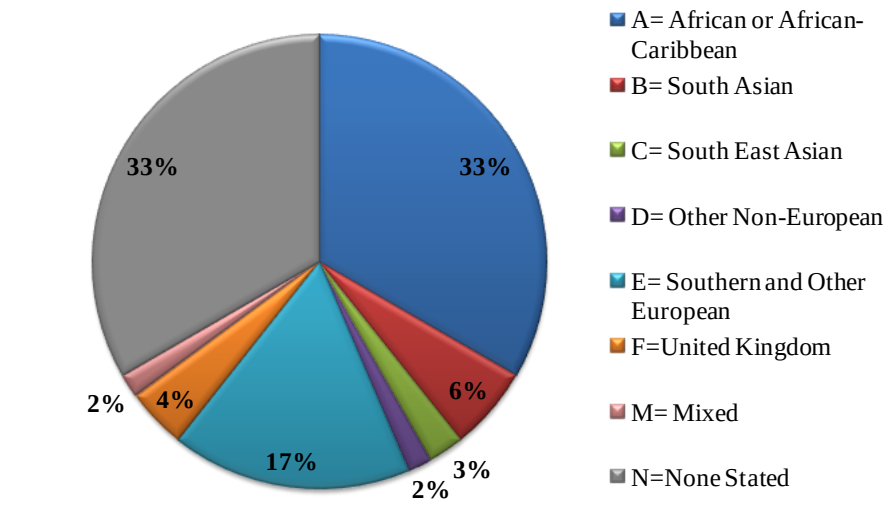
b)

Figure 3.40. Pie-charts showing the proportion of the hemoglobinopathies in the whole set of samples (a) and the proportion of different hemoglobin disorders (b)

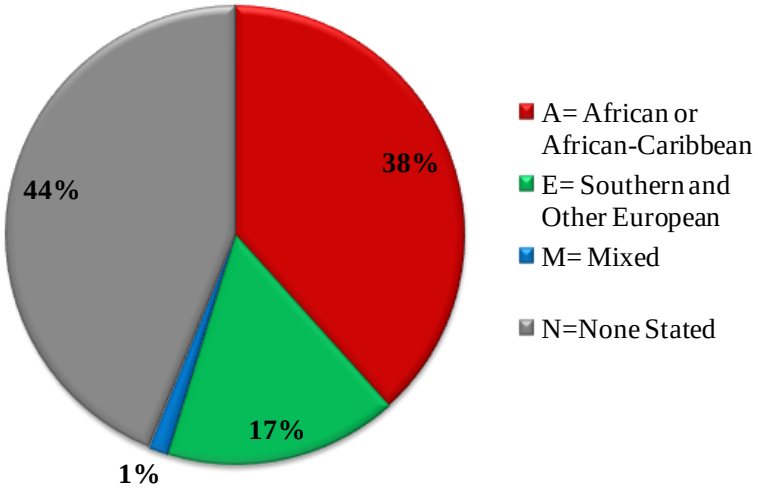
From the total number of δ-chain (A2) variants and α^0-thalassemia carriers, 10% of the abnormal samples were not confirmed using the mass spectrometry methods, and 6% of abnormalities were only detected using mass spectrometry.

Ethnic origin of abnormal samples

Where it was possible the ethnic origin of those patients with abnormal samples were investigated. A graphical presentation of the findings can be seen in Figure 3.41. Abnormalities are highest amongst African or African-Caribbean, but a high percentage of the conditions were from individuals of Southern or other European origin (Figure 3.41 a). When considering carriers of sickle disorders, African and African-Caribbean origin was more than double that of Southern or Other European origin (Figure 3.41 b).



a)



b)

Figure 3.41. Pie-charts showing the proportion of the hemoglobinopathies within different ethnic groups (a) and the proportion of Sickle carriers within different ethnic groups (b)

Part II: Results of the pilot study of the top down MS method, on the partial sample set of the clinical trial using the amaZon speed ETD mass spectrometer

3.3.7 Preliminary experiments

3.3.7.1 Sequence coverage studies for normal samples

A control blood sample was prepared using 1/10,000 dilution in 50% ACN and 0.2% formic acid.

MS/MS spectra with ETD fragmentation of the 18+ charged ions of the β-chain were collected with previously optimised parameters (ETD reaction time 20 ms, PTR reaction time 80 ms). Based on the observed fragment ions sequence matching was performed in BioTools. The ions observed were predominantly from the c- and z+1 ions series. z+1 ions are produced and not z-ions, which would be expected from fragmentation reactions in ETD. This is due to additional reactions taking place in the iontrap analyser.

Figure 3.42 illustrates the obtained sequence coverage for the β-chain. 58% of the sequence is covered by the observed and assigned fragment ions.

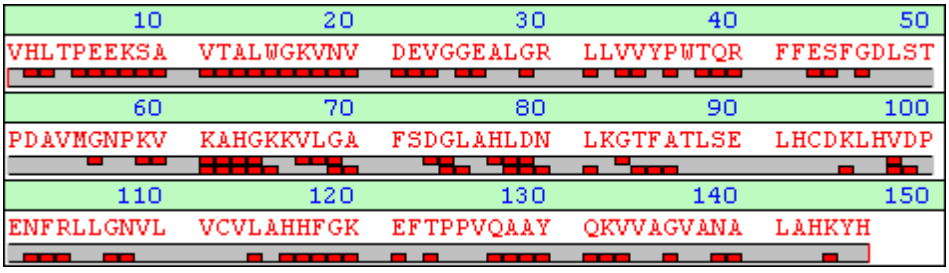


Figure 3.42. Sequence coverage for beta-chain
Red bricks represent matched c (upper bricks) or z+1 (lower bricks) ions

For the development of a screening method the primary aim was to cover the affected amino acids and sequence region of the β-chain, where the mutation occurs, which was achieved in all investigated cases. Amino acids at position 6, 26 and 121 were covered, where amino acid substitutions are expected for HbS, HbC, HbE,

HbD-Punjab and HbO-Arab mutations. For the identification of unknown variants the highest sequence coverage obtainable is desirable. Samples with clinically significant variants were analysed and are presented in Section 3.3.7.4.

Preliminary sequence coverage studies on the α-chain were also performed. The 17+ α-chain ion was fragmented. Settings: for dual isolation first step *m/z* 895 and for second step *m/z* 890 were chosen, the reaction times were ETD 20 ms and PTR 100 ms.

Figure 3.43 illustrates the obtained sequence coverage for the α-chain. Fairly good coverage was observed for the N and C-termini with few fragments observed for the core region. An overall sequence coverage of 35% was obtained based on the results from BioTools. No further optimisation for higher sequence coverage was attempted at this stage. The analysis of sample A5 with HbG-Phildelphia did not result in sufficient information for the mutation to be identified simply from MS/MS measurements. This mutation occurs with a Asn→Lys substitution at position 68 of the α-chain, for which no amino acid fragment ions were observed.

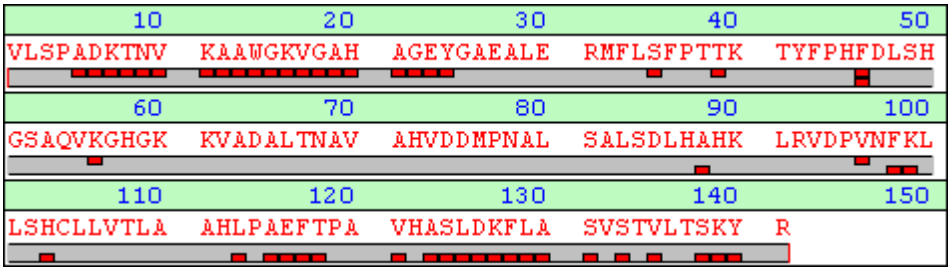


Figure 3.43. Sequence coverage for alpha-chain.
Red bricks represent matched c (upper bricks) or z+1 (lower bricks) ions

For higher sequence coverage to be obtained for both the α- and β-chains, further method development would be necessary. At this stage the sequence coverage studies focused predominantly on the β-chain to develop a screening method for the clinically significant variants, which affect exclusively the β-chain. Clinical samples analysed and presented in this chapter have not been separately evaluated for sequence coverage information, since that was not the primary aim of the work.

3.3.7.2 HbA_2 and HbA_{1c} quantitation

Intact globin chain analysis performed on the amaZon speed ETD iontrap analyser instrument was assessed for quantitation of HbA_2 and HbA_{1c} levels.

HbA_2 level determination

A clinical sample from a healthy adult and the HbA_2 reference standards were measured. δ-chain levels based on peak intensities and peak areas were calculated. The data for each resulted from a 15 minute acquisition, processed after the initial 5 minutes and for the whole 15 minutes. Results are summarised in Table 3.17.

Sample	**Expected HbA_2 level**	**δ-chain level % - 5 min**	**δ-chain level % - 15 min**
HbA_2 standard – Intensity Calc.	5.3%	6.1	6.2
HbA_2 standard – Area Calc.	5.3%	7.2	7.7
Clinical sample – Intensity Calc.	2.8%	3.7	3.9
Clinical Sample – Area Calc.	2.8%	4.6	5.2

Table 3.17. Results of HbA_2 quantitation for a clinical sample and the reference standard

The calculated values are higher than theoretically expected for both intensity and area calculated values. This was expected based on measured values using the Waters Xevo TQ mass spectrometer which have been presented in previous Chapters. The results after 5 minutes and after 15 minutes were very similar, showing that a shorter intact measurement time provides similarly reliable results to the longer 15 minute one.

Calculated values based on peak intensities are lower, closer to the expected values.

The iontrap analyser has been shown to be potentially capable of quantifying HbA_2 levels based on intact globin chain analysis in the same way as demonstrated for the triple quadrupole instrument. For further evaluation reproducibility studies need to be performed.

HbA_{1c} level determination

A clinical sample from a healthy adult and previously used reference standards were measured and levels based on peak intensities and peak areas were calculated for glycated α-chain and glycated β-chain ion peaks. The data were obtained from a 15 minute acquisition. Results are summarised in Table 3.18.

Sample	Expected HbA_{1c} level %	Glycated α-chain level % - 5min	Glycated α-chain level % -15min	Glycated β-chain level % - 5min	Glycated β-chain level % - 15 min
HbA_{1c} standard –Intensity Calc.	9.3	3.2	3.2	6.4	6.4
HbA_{1c} standard – Area Calc.	9.3	3.7	3.9	7.9	8.3
Clinical sample – Intensity Calc.	-	2.6	2.6	5.4	5.4
Clinical Sample – Area Calc.	-	3.9	4.2	8.0	8.5

Table 3.18. Results of HbA_{1c} quantitation for a clinical sample and the reference standard

The calculated values are lower than theoretically expected for both intensity and area calculated values for the reference standard. This was again expected based on measured values using the Waters Xevo TQ mass spectrometer presented in previous Chapters. There was no information available for the clinical sample about glycation levels. Considering only the glycated β-chain levels as informative values, the clinical sample had a very similar, supposedly elevated level for glycated β-chain to the reference standard. The same phenomenon was observed when this sample was analysed using the Waters Xevo TQ instrument, proving that the Bruker amazon speed ETD instrument with an iontrap analyser could provide a reliable technique for HbA_{1c} quantitation. The values calculated based on peak areas are closer to the expected values here, which was not the case for HbA_2 quantitation.

The iontrap analyser is potentially capable of quantifying HbA_2 and HbA_{1c} levels based on intact globin chain analysis. Further investigation on larger number of clinical samples would be necessary to prove this hypothesis. The analysis of ~1000

clinical samples was performed using the Bruker amaZon speed ETD mass spectrometer. Data resulting from the analysis of HbA_2 and HbA_{1c} levels will not be presented in this Chapter. The results cannot be used for an accurate representation of the method for quantitation purposes due to significant carry-over issues from the LC sample introduction system that occurred during the measurement of the clinical samples, which affected the intact globin chain analysis. Further method development would need to focus on eliminating carry-over and choosing appropriate peak properties for estimating globin-chain levels.

3.3.7.3 Mass accuracy of intact measurement

Selected samples with and without variant globin chains were measured and the mass accuracy for the beta-chain mass was investigated.

Reference standards for HbA_2 and HbA_{1c} are expected to have normal α- and β-chains, therefore assessment of mass accuracy was made using these samples. Patient blood samples including a control healthy adult sample, and samples with HbS, HbC, HbE, HbD-Punjab and HbG-Phialdelphia were also analysed. The mass error of the samples were calculated for each measurement and plotted in Figure 3.44.

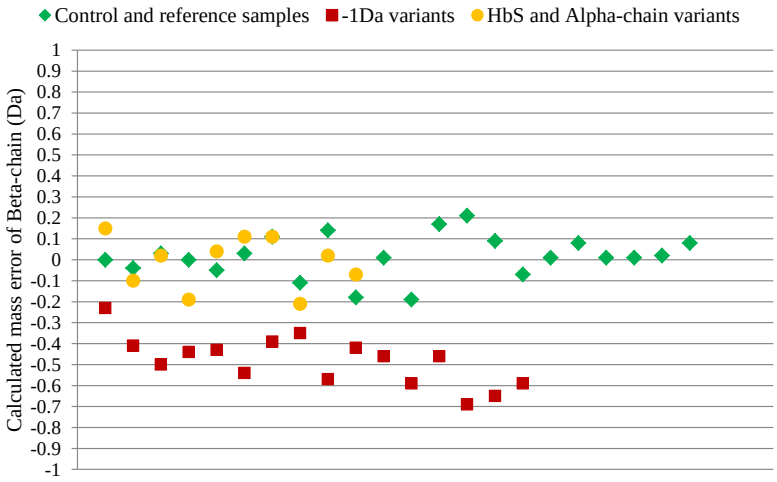


Figure 3.44. Mass error calculated for the Beta-chain using the Alpha-chain as internal standard.

Mass errors calculated for the reference standards and the control sample are in a range of ±0.2 Da distributed around 0, which indicates normal distribution and no

systematic error as previously observed for the measurements performed on the Xevo TQ instrument.

An interesting observation is that the samples containing heterozygous variants were not measured with abnormal mass errors as had been observed with the intact globin chain analysis performed on the Waters Xevo TQ instrument. This is probably due to higher resolution of iontrap data.

This observation allows one to distinguish easily between normal samples and samples with -1 Da variants. A disadvantage is that it does not provide detection of other abnormalities, which were observed previously based on the mass errors with the Xevo TQ intact MS method.

The aim of the work presented in this Chapter was to develop and apply an ETD fragmentation method for clinically significant variants based on acquired MS/MS spectra. Analysing every patient sample looking for the presence of characteristic fragment ions to indicate the presence of a certain mutation makes the mass accuracy measurement a less crucial point as the variants of interest with -1 Da mass difference should be revealed based on the ETD spectra. Ideally, all clinically significant variants which result in a -1 Da mass difference to the normal β-chain would be detected based on the observed fragments. It may, however, occur that a -1 Da α-chain variant is present, which would not be detected, since the automated MS/MS method would screen for β-chain variants only, fragmenting exclusively the corresponding β-chain ion. For these cases detecting possible abnormal mass errors is still desirable.

3.3.7.4 Clinically significant variants

To assess whether a top-down fragmentation method using electron transfer dissociation could be used for the detection and definite diagnosis of clinically significant hemoglobin variants, samples with these conditions have been measured.

Four samples with clinically significant variants have been analysed in order to assess the potential of the ETD-method for sequence coverage of the variant globin chains and find characteristic fragment ions, which if present could be used as reporter ions to confirm the presence of certain mutations.

HbS mutation

The most important variant that needs definitive detection is the HbS mutation. As previously mentioned, a fully resolved variant beta-chain is present in the intact ESI-MS spectrum. The Glu residue at position 6th of the beta chain is replaced by a Val residue. It has previously been shown that ETD fragmentation is capable of providing fragments covering this sequence region, since it is close to the N-terminus. The mutation results in a -30 Da mass shift, and the fragment ions are expected to be present -30 Da lower compared to the mass to charge ratio of fragments from the wild type β-chain.

The 18+ charged β-chain ion was fragmented in the iontrap using ETD with the previously detailed parameters. The resultant ETD spectrum revealed a number of fragment ions from the c-ion series predicted for the β^{SICKLE}-chain by the Biotools^{TM} software (Bruker Daltonik GmbH, Bremen, Germany). The obtained ETD spectrum of the sample was compared to that of a ETD spectrum obtained from a control healthy sample. The resultant differences and the fragment ions characteristic of the HbS mutations are shown in Figure 3.45. The entire mass spectrum contained numerous fragment ions from the mutated β-chain which covered many amino acids including the site of the substitution and residues from the position 6th onwards. The majority of fragments were doubly charged c-ions. The c_6^{1+} ion (*m/z* 664.49) was observed and present in the mass spectrum, but at lower intensity than observed for the c_7^{1+} ion (*m/z* 793.52). For a variant specific reporter ion therefore the c_7^{1+} ion peak was chosen for further diagnostic purposes.

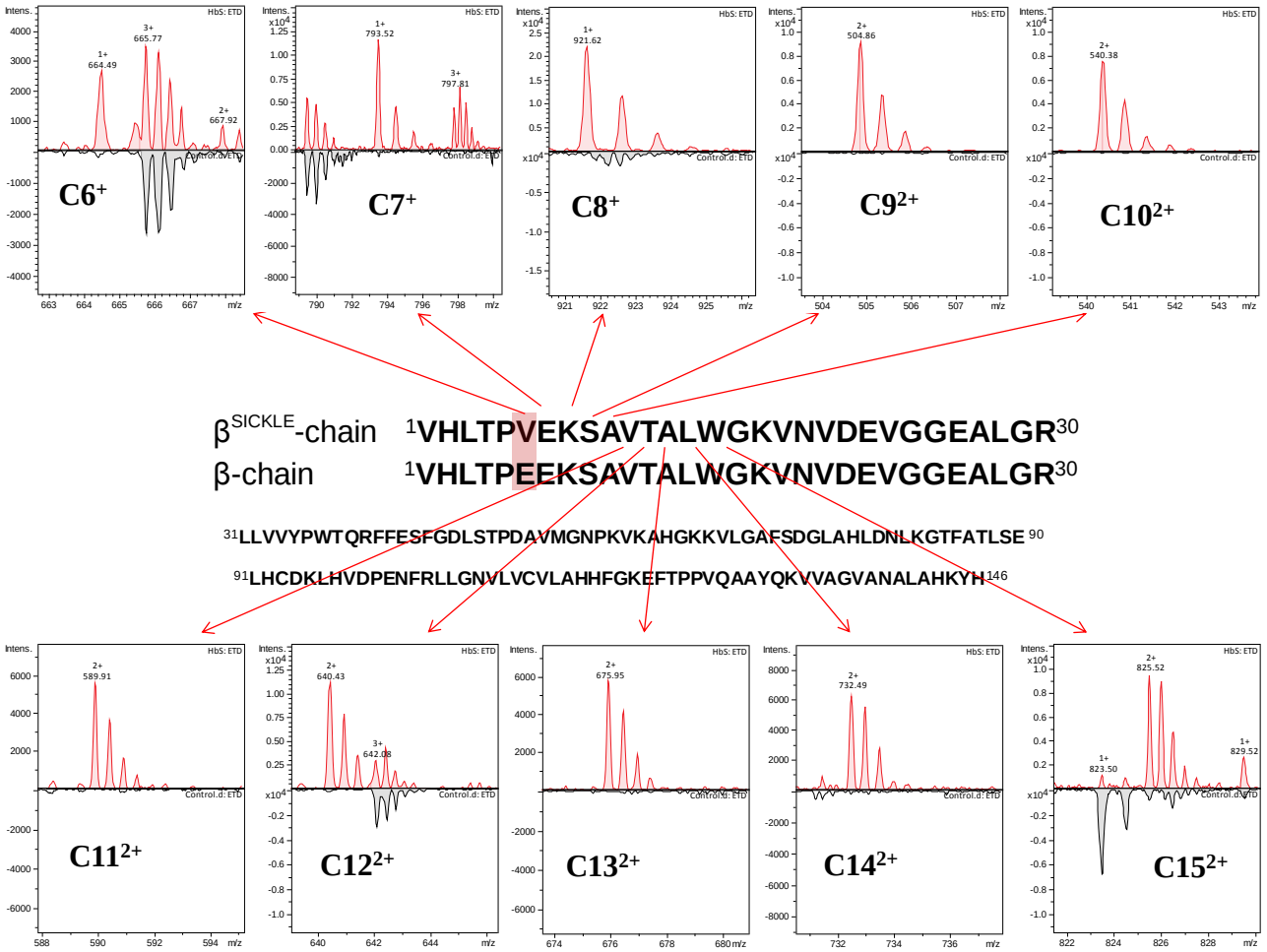


Figure 3.45. Fragment ions observed for sickle β-chain (Black spectra represent the control sample while red spectra are obtained from HbS sample)

HbC mutation

A sample with the clinically significant variant HbC was analysed. This mutation affects the same position of the β-chain sequence as HbS, but the Glu residue is replaced with a Lys residue. The resulted variant chain only differs by -1 Da to the normal β-chain, and the possible fragment ions have the same mass shift.

The ETD spectrum was a fragmentation rich spectrum as previously observed for the HbS mutation. Fragments from the c-ion series were observed for the substitution site and for amino acids after that. Selected fragment ions for the variant and the control sample are compared in Figure 3.46. The variant c_6^+ ion is present at *m/z* 693.51 at low intensity and the control sample also contains some low intensity noise peaks at this *m/z*. The c_7^+ and c_8^+ions are more intense at *m/z* 822.48 and *m/z* 950.58 respectively. Since the sample is a mixture of normal and variant β-chains, the isotopic envelope observed for these fragment ions, due to the only -1 Da difference, represents both fragment ions from the two globin chains, and not clearly isolated with a -30 Da difference as observed for the HbS mutation. From the c_9^{2+} ions onwards several c ions are observed in their doubly charged forms, only 0.5 Da apart from the fragment ions for the normal chain, but still clearly present and unambiguously showing the presence of the HbC variant. As a reporter c_8^+ or c_7^+ ions may be used. The consistent observation of characteristic fragment ions needs to be assessed on larger number of clinical samples, which will be presented later in this Chapter. As an initial preference for reporter ion c_8^+ was chosen.

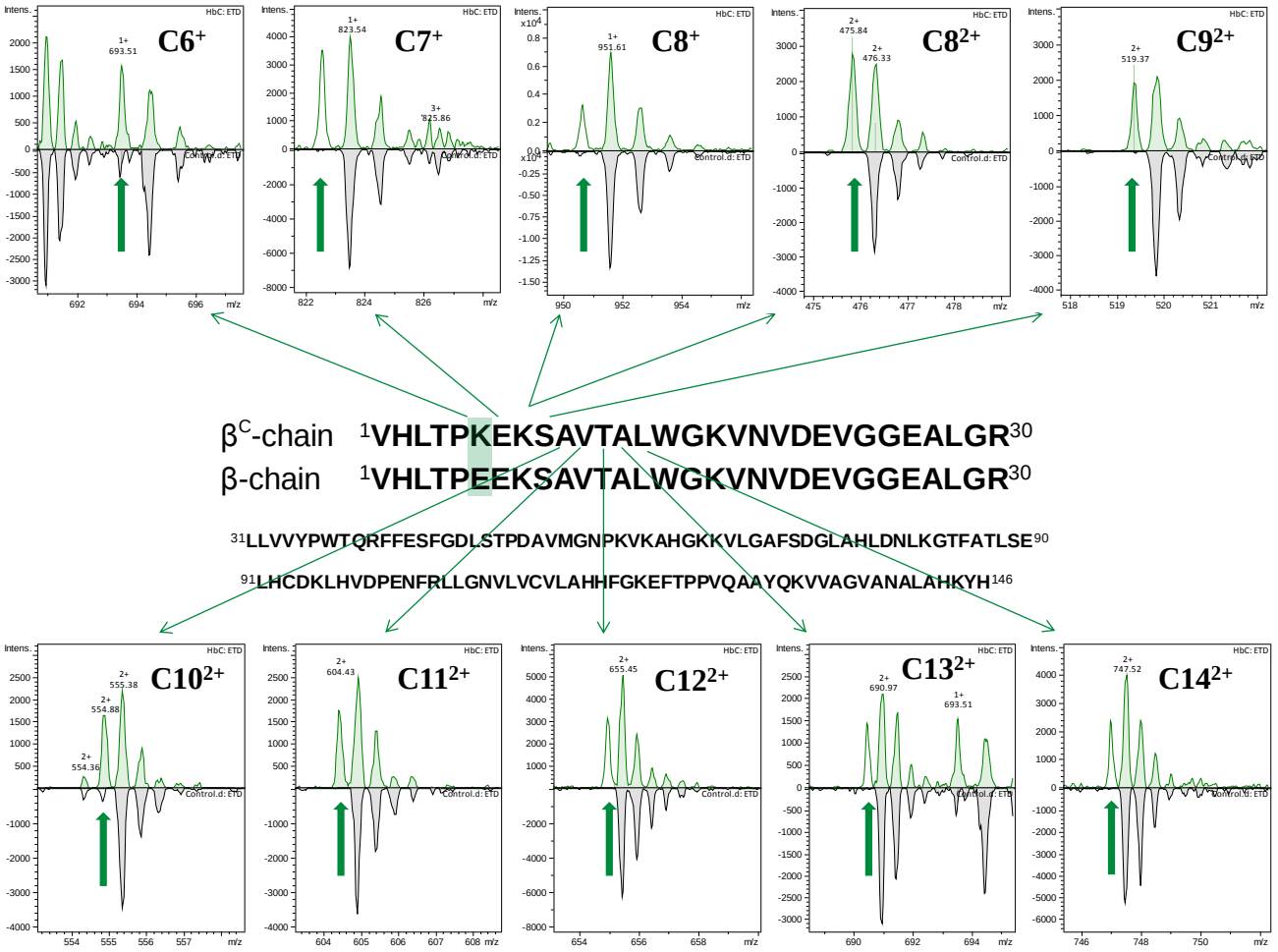


Figure 3.46. Fragment ions observed for HbC β-chain (Black spectra represent the control sample while green spectra are obtained from HbC sample)

HbE mutation

A sample with the clinically significant variant HbE was analysed. This mutation affects position 26 of the β-chain sequence, the Glu residue is replaced with a Lys residue. The substitution results in a -1 Da difference to the normal β-chain, and possible fragment ions have the same mass shift. This variant detection is slightly more challenging since the mutation site affects a position of the amino acid sequence further away from the N-terminus, closer to the core region of the globin-chain sequence. There were less fragment ions observed in the mass spectrum for the investigated region and their intensities were also lower than those observed for the HbS and HbC variant chains. Only two fragment ions were observed with high confidence in the ETD spectrum for this mutation (Figure 3.47). For the position of the substitution the c_{26} ion can be observed in doubly and triply charged ion forms, and the c_{29} fragment is also observed in its triply charged form. As a reporter ion either the c_{26}^{2+} (*m/z* 1381.8) or c_{26}^{3+} (*m/z* 921.6) ion may be used, although in the control sample at the same *m/z* values low intensity noise peaks are present, and the c_{29}^{3+} (*m/z* 1001.95) ion may be more appropriate. The application of c_{26}^{3+} (*m/z* 921.6) is not advised since it has the same *m/z* as the c_8^{+} ion observed for the HbS β-globin chain. As an initial decision the c_{29}^{3+} was chosen.

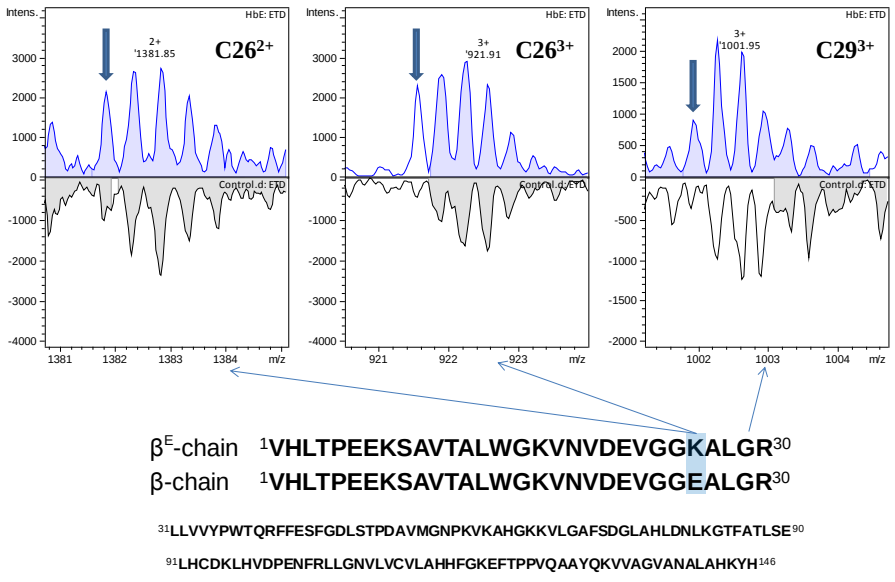


Figure 3.47. Fragment ions observed for HbE β-chain (Black spectra represent the control sample while blue spectra are obtained from HbE sample)

HbD-Punjab mutation

The HbD-Punjab mutation is a result of the substitution of Glu 121 residue with a Gln residue, resulting in a -1 Da mass difference to normal β-chain.

Multiply charged fragment ions for the (z+1)-series were observed. The intensity and the confidence of these fragment ions are lower based on possible overlapping noise peaks in the mass spectrum. The ions for the (z+1)-series are less intense and the coverage is not as complete as was observed for the c series, which poses a challenge for confirming the presence of this mutation.

The observed fragment ions at the position of substitution and onwards are shown in Figure 3.48. The observed fragment ion $(z+1)_{26}^{3+}$ (*m/z* 931.54) at the point of the mutation confirms the position of the substitution. The $(z+1)_{27}^{3+}$ ion at *m/z* 974.25 was initially selected as a reporter ion for further studies. Based on comparison of the ETD mass spectra of the HbD-Punjab and control sample, there were no or insignificantly low interfering ions at the *m/z* of the $(z+1)_{28}^{3+}$ ion (*m/z* 993.2). Depending on the outcome of analysis of further samples with the HbD-Punjab mutation, this ion may be suggested as an alternative reporter ion. The consistent presence of this ion peak in clinical samples needs to be investigated in order for this decision to be made.

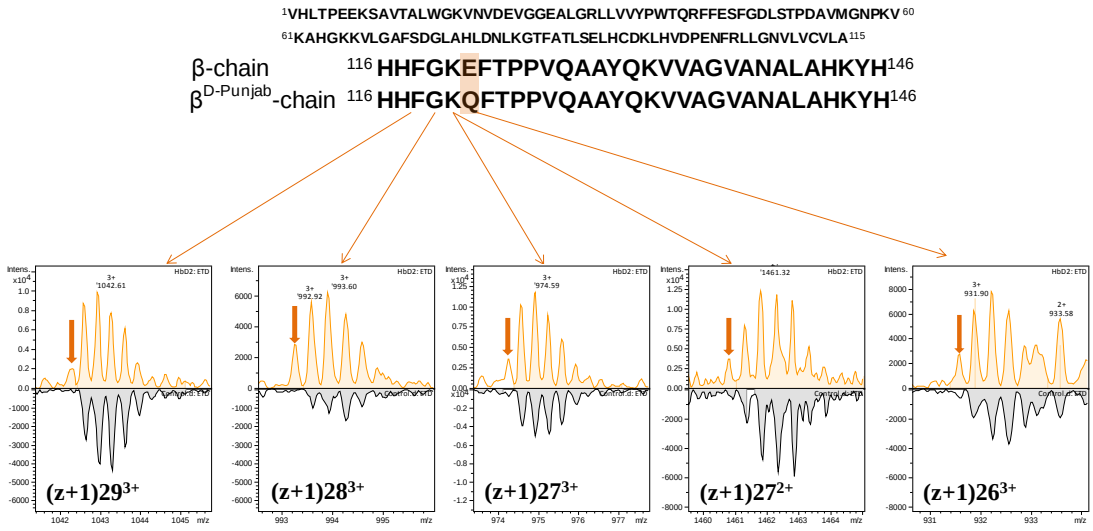


Figure 3.48. Fragment ions observed for HbD-Punjab variant β-chain
(Black spectra represents the control sample while orange spectra are obtained from HbD sample)

Based on preliminary analysis, the fragment ions selected for monitoring the clinically significant variants in patient blood samples are summarised in Table 3.19.

Sample ID	Variant	Position	Mutation	Chosen reporter fragment ion and *m/z*
A04	HbS	6	Glu→Val	c_7^{1+} 793.5
A02	HbC	6	Glu→Lys	c_8^{1+} 950.61
A01	HbE	26	Glu→Lys	c_{29}^{3+} 1001.95
A03 -	HbD-Punjab/ (HbO-Arab)	121	Glu→Gln Glu→Lys	$(z+1)_{27}^{3+}$ 974. 25

Table 3.19. Summary of hemoglobin variants and chosen reporter fragment ions for the detection of different conditions

These reporter ions were chosen based only on one analysed example for all mutations. More samples were analysed to build a list of observed ions, which may provide additional information to suggest a suitable reporter ion for diagnostic purposes.

Since the mutation HbO-Arab affects the exact same position of the β-chain as the mutation HbD-Punjab, also resulting in a -1 Da mass difference, it is assumed that the fragmentation pattern for the two mutations would be the same. In case the characteristic fragment ions for HbD-Punjab mutation are observed, further investigation such as tryptic peptide analysis would be necessary to confirm which variant is present.

3.3.8 Analysis of clinical samples

3.3.8.1 Screening for clinically significant variants.

The script selects regions of interest for the four previously characterised clinically significant variants and publishes *m/z* windows into a portable document format (pdf) report. An example of the 2 page report produced can be found in Appendix 4.

Figure 3.49 shows part of the generated report from the analysis of a clinical sample from a healthy individual. The shaded areas highlight the positions where additional peaks are expected in the mass spectrum if a variant were present.

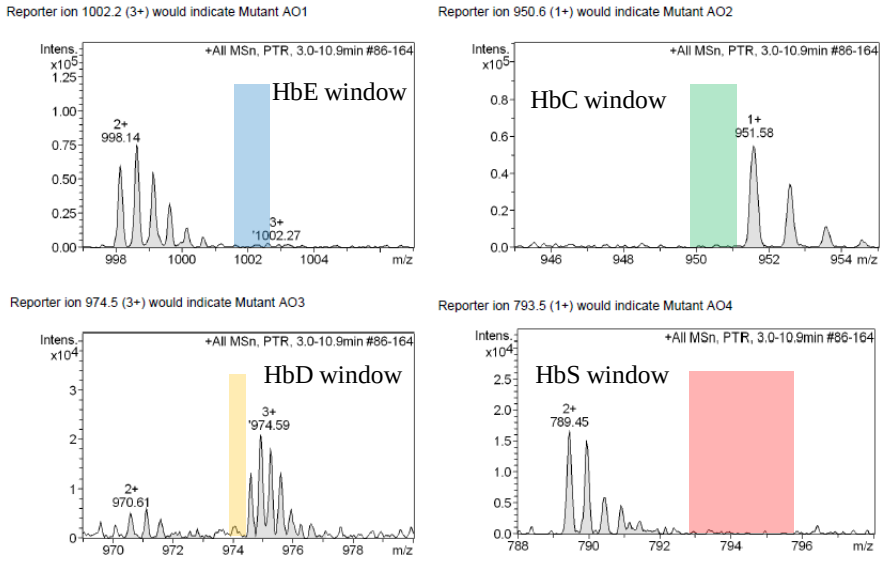


Figure 3.49. Example of the variant diagnostic section of the generated report for a control sample

Figure 3.50 shows the presence of reporter ions for HbC and HbS mutations confirming the sample as compound heterozygote for the two mutations. As for the HbC singly charged peak at *m/z* 950.6, the isotope envelope demonstrates that it corresponds to only one singly charged ion species, and with the presence of the HbS reporter ions it is confirmed that there is no normal β-chain present in the sample. In cases of homozygous HbS samples the expected peak at this position for the normal β-chain (*m/z* 951.6 for normal β-chain) is completely absent from the ETD spectrum. This observation has been used for distinguishing between heterozygous or

homozygous HbS carriers. In cases where the sample had been classified as SS by the hospital lab but fragment ions for the normal β-chain were observed, findings were in agreement with results from the other MS methods shown in previous chapter, indicating the sample was a heterozygous HbS carrier.

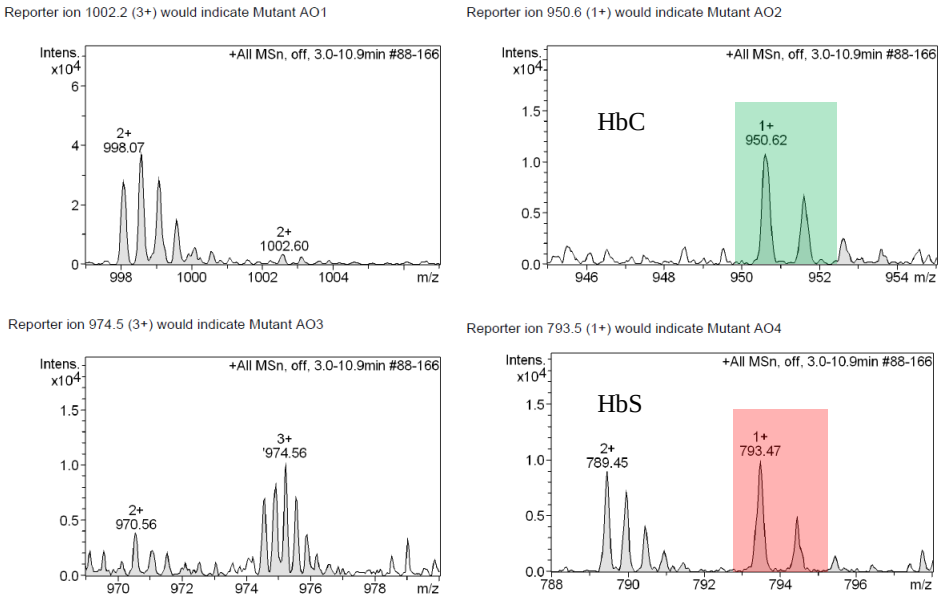


Figure 3.50. Example of the variant diagnostic section of the generated report for a sample compound heterozygous of HbS and HbC

For variants HbE and HbD-Punjab confirmation is not as straightforward. The analysed sample set contained only one sample with HbE variants and the whole ETD spectrum, rather than just selected regions, was analysed using the Data Analysis software. The peak intensity of the monitored reporter ion was too small to be diagnostically useful in the monitored *m/z* window. After carefully investigation of the ETD spectrum the c_{26}^{2+} at *m/z* 1381.8 was suggested to be a more suitable reporter ion instead of the previously chosen c_{29}^{3+} ion at *m/z* 1001.95.

For the HbD-Punjab mutation a larger number of samples were analysed. The monitored reporter ion at *m/z* 974.5 was observed for 9 samples out of the 10 analysed samples. The reporter ion was not present in sample E143, which was not confirmed by other MS methods either (discussed in the previous Chapter). This shows good agreement within the different MS experimental approaches, confirming that this is not a measurement error, or an inability of the ETD approach to detect this mutation. After investigating the whole ETD spectra for all samples with HbD-

Punjab variant, the reporter ion for this mutation was suggested to be changed for $(z+1)_{28}^{3+}$ ion (*m/z* 993.2), since it was present in all investigated ETD spectra, and was absent in the mass spectra of samples from healthy patients.

The one sample with HbO-Arab variant showed the same fragment ions which were observed for the HbD-Punjab mutation, as expected.

For HbS variant the chosen reporter ion proved to be reliable. The reporter ion for the HbC variant could be changed for c_7^+ at *m/z* 822.5 from the c_8^+ at *m/z* 950.6, since the presence of the c_7^+ ion peak was clearly observed in cases of all samples with this variant. The chosen reporter ion, however, was a good indicator of the presence of HbC variant, and its replacement may not be necessary.

A summary of the number of analysed samples with different conditions and the number of detected samples by the ETD approach can be found in Table 3.20. In a small number of cases the ETD approach did not confirm the presence of a variant. This was caused by poor data quality due to the sample introduction issues discussed earlier.

Mutation	**Number of analysed samples**	**Detected**	**Comment**
SS	18	7+9*	1 not confirmed due to technical problems *not SS, only AS indicated
Transfused SS	4	1	Not confirmed mutations are due to technical issue not problems with ETD approach
AS	47	43	Not confirmed mutations are due to technical issue not problems with ETD approach
SC	3	3	No peaks for normal β-chain
CC	1	1	No peaks for normal β-chain
AC	6	5	In case of the one not confirmed it is assumed human error may have been the reason mixing samples
AE	1	1	c_{26}^{2+} fragment at *m/z* 1381.8 confirms
AD	10	9	The one not detected was not confirmed by other MS methods either
AO	1	1	Same fragments as AD

Table 3.20. Summary of analysed samples with different conditions and successful confirmation by the ETD approach

Analysis of several sample batches with a total number of about 1,000 analysed samples highlighted significant issues with the sample introduction method chosen. Since an LC flow was used to facilitate the sample introduction from the sample loop into the ion source of the mass spectrometer, it was assumed that there would not be any significant carry- over of a previous sample to interfere with measurement of the following sample. This unfortunately proved to be not the case and therefore data obtained for intact globin chain analysis could not be used. Clinical samples are assumed to form salt adducts during analysis due to the chemical environment in the blood collection tubes (anticoagulant agent K_2/K_3EDTA). The salt present in the samples made the analysis less efficient, and influenced the yield of the ETD fragmentation for the ions of interest. The LC pump and sample introduction system was not able to perform reliable sample introduction for the analysis of 1,000 samples and needed constant maintenance, indicating that it was not robust enough for clinical laboratory applications.

3.3.8.2 Analysis of rare hemoglobin variants

Samples containing rare variants were also analysed using the top-down MS/MS method. Some of these samples required slight changes to the initial isolation settings.

In those cases, where the original isolation settings yielded fragmentation for variant and normal chains simultaneously, results obtained support the hypothesis that the measurement would be able to provide information regarding rarer abnormalities without the need for any additional experiments. Four examples, listed in Table 3.21, are presented here which resulted in fragment ions being obtained indicating the variant chain sequence. When additional measurements were necessary, the samples were analysed in collaboration with Julia Smith at the laboratory of Bruker UK Limited.

Sample	Observed mass shift	Mutation	Identified variant
E61	Beta chain -10 Da	β5(Pro→Ser)	**Hb Tyne**
E187	Beta chain +58 Da	β16(Gly→Asp)	**Hb J-Baltimore**
H47	Beta chain +28 Da	β10(Ala→Val)	**Hb Iraq-Halabja**
M136	Beta chain +14 Da	β11(Val→Leu/Ile)	***"Hb Hamilton" or a novel variant***

Table 3.21. List of rare Hemoglobin variants analysed with the top-down ETD-method

Sample E61 was presented in the previous Chapter as a carrier of variant Hb Tyne. This mutation affects the 5th position of the β-chain sequence; a Pro is substituted with Ser, resulting in a -10 Da mass difference of the intact variant globin chain to normal β-chain, and also a -10 Da mass shift for the variant fragment ions. Additional fragment ions were observed in the ETD spectrum after the analysis of the sample with the proposed test method. The c_4^+ ion (*m/z* 468.33) was present in the ETD spectrum of the patient sample but not in a control sample (Figure 3.51). The Pro at position 5 of the normal β-chain prevents the cleavage at position 4, generating a c_4^+ ion. The N-terminal cleavage to Pro residue was not observed in ETD spectra, since it would require the cleavage of two bonds. In the mutated chain the 5th amino acid is changed to Ser, therefore the cleavage can easily happen, resulting in the observed c_4^+ ion. This supports the hypothesis regarding the suggested position and also the nature of the substitution. An additional peak at -10 Da to the normal c_5^+ ion was observed at *m/z* 555.37, although this peak was present in the control sample too, as it is an overlapping peak with the doubly charged c_{10} ion, which is expected at *m/z* 555.30. The isotope envelope for this peak, however, indicates the presence of a singly charged ion species, as shown in Figure 3.51. The obtained ETD spectrum was information rich, revealing many fragment ions matching the sequence of the variant β-chain of Hb Tyne. Ion peaks up to c_{15} are clearly present. This mutation could be identified, but was more challenging. From the ETD spectrum obtained with the original isolation and fragmentation settings, the substitution and the variant could be identified with high confidence without additional sample preparation involving enzymatic cleavage, or additional MS measurements.

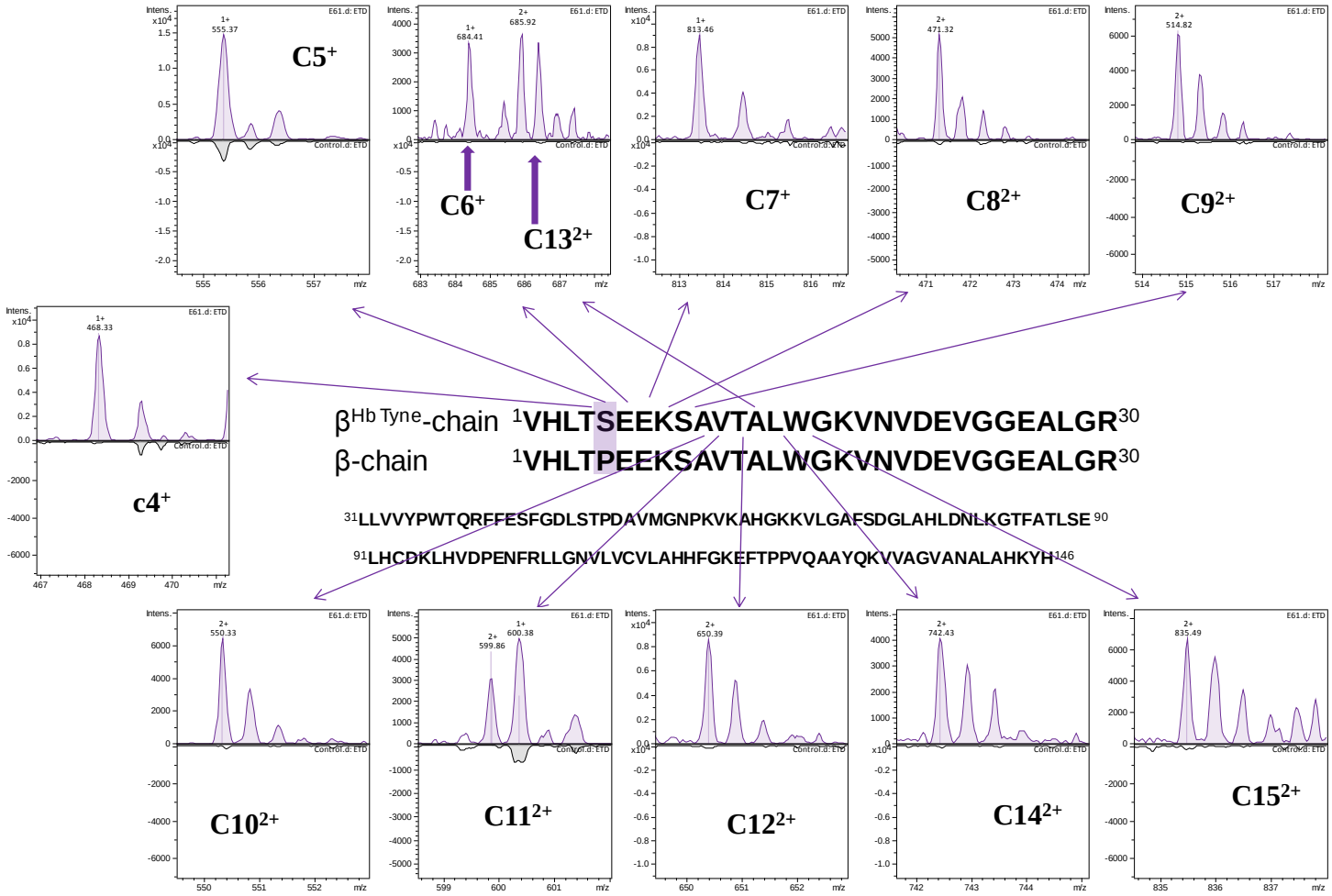


Figure 3.51. Additional fragment ions observed for sample E61, carrier of Hb Tyne variant

The analysis of sample M136 also resulted in an information rich ETD spectrum of the variant chain. This sample is a carrier of a mutation affecting position 11 of the β-chain, with a substitution of Val to either Leu or Ile residue, resulting in a mass difference +18 Da to normal β-chain. The fragments observed are shown in Figure 3.52. The corresponding mass peaks were observed for c_{11}^{2+} with the +18 Da mass-shift at *m/z* 611.88 not present in the mass spectrum of the control sample. Other ions from the c-series were also observed in the ETD mass spectrum. Doubly charged ions matching the expected *m/z* for c_{12}-c_{16} and for c_{18} were also present in the mass spectrum of the sample but absent from the control sample. This sample could be analysed and the variant chain and mutation site, substitution identified with the original measurement settings without the need for additional experiments or enzymatic cleavage step.

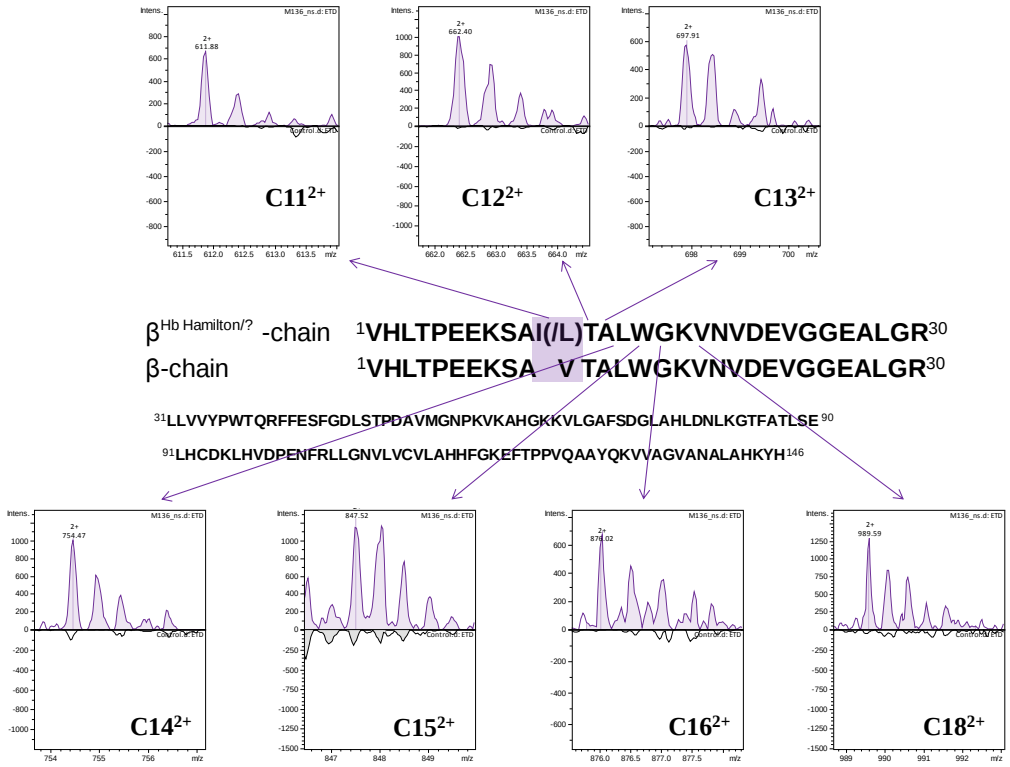


Figure 3.52. Additional fragment ions observed for sample M136

In the case of sample H47 there was no normal β-chain present, only a sickle β-chain and a variant β-chain. The isolation window clearly covers the targeted ions for the sickle β-chain and a fragmentation rich ETD spectrum was recorded with the previously described c-ions being observed. The other chain present in this sample

was identified as Hb Iraq-Halabja, which occurs as a result of a mutation affecting position 10 of the β-chain, in which the Ala residue is substituted with a Val residue causing a mass difference of +28 Da. The multiply charged variant β-chain peaks are, however, less successfully fragmented. There are some fragments observed that were not present in a control sample, which were in agreement with the predicted masses for the fragments of the previously identified mutation.

Figure 3.53 shows the identified fragment peaks. No confirmation regarding the mass shift resulting in variant c_{10} ion could be obtained. With careful investigation of the mass spectrum around the *m/z* expected for the variant c_{11}^{2+}, the peaks may be considered to be present although overlapping with other peaks. Additional peaks at the expected *m/z* for doubly charged ions of c_{12} to c_{15} were confidently identified. These mass peaks would not be present if the mutation with the +28 Da mass difference was not present at position 10. Different isolation settings may give a more confident identification due to higher intensity fragment peaks originating predominantly from the variant chain and not from the sickle beta-chain.

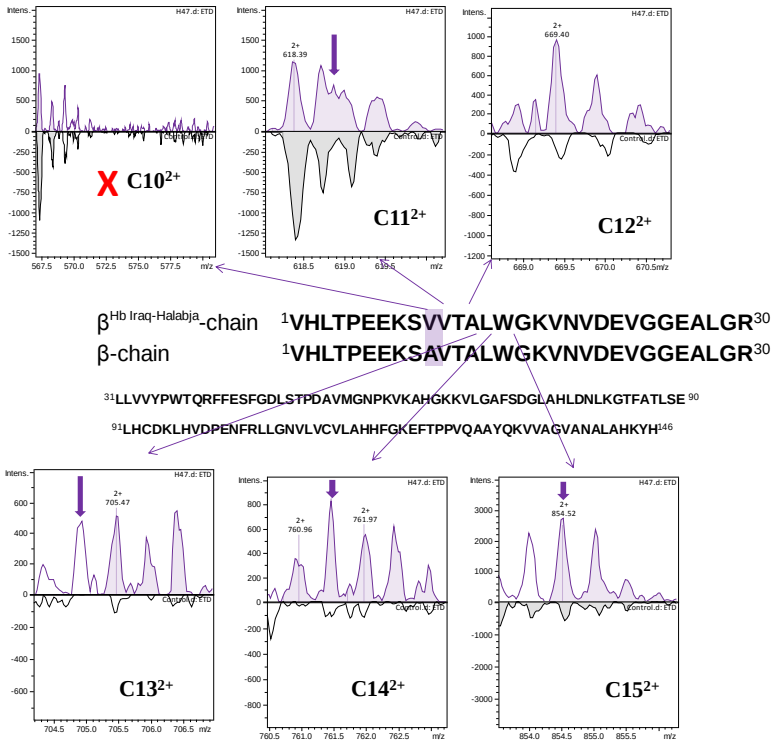


Figure 3.53. Additional fragments observed for sample H47

Sample E187 was identified in the previous chapter as Hb J-Baltimore. The mutation affects position 16 of the β-chain with a Gly to Asp change. The resulted mass

difference was +58 Da, which for the 18+ charge state on the intact level was not within the isolation window set originally for the test screening method. Additional experiments needed to be performed. The isolation window was initially adjusted. The first isolation was at *m/z* 888.0 and the second isolation step at *m/z* 885.6. The ETD and PTR parameters were kept the same as in previous experiments. The overall ion intensity was lower in the ETD spectrum than that observed for the normal β-chain. Additional fragment ions compared to control sample are shwon in Figure 3.54. Observed ions can be matched with predicted doubly charged fragment ions for the identified chain for c_{18}, c_{20}, c_{22}, c_{23}, c_{29} and c_{30}. Although c_{16} fragment ions for the position of the substitution were not observed, the +58 Da may be the result of a Gly → Asp substitution. The assigned fragment ions would be absent if this mutation had not occurred.

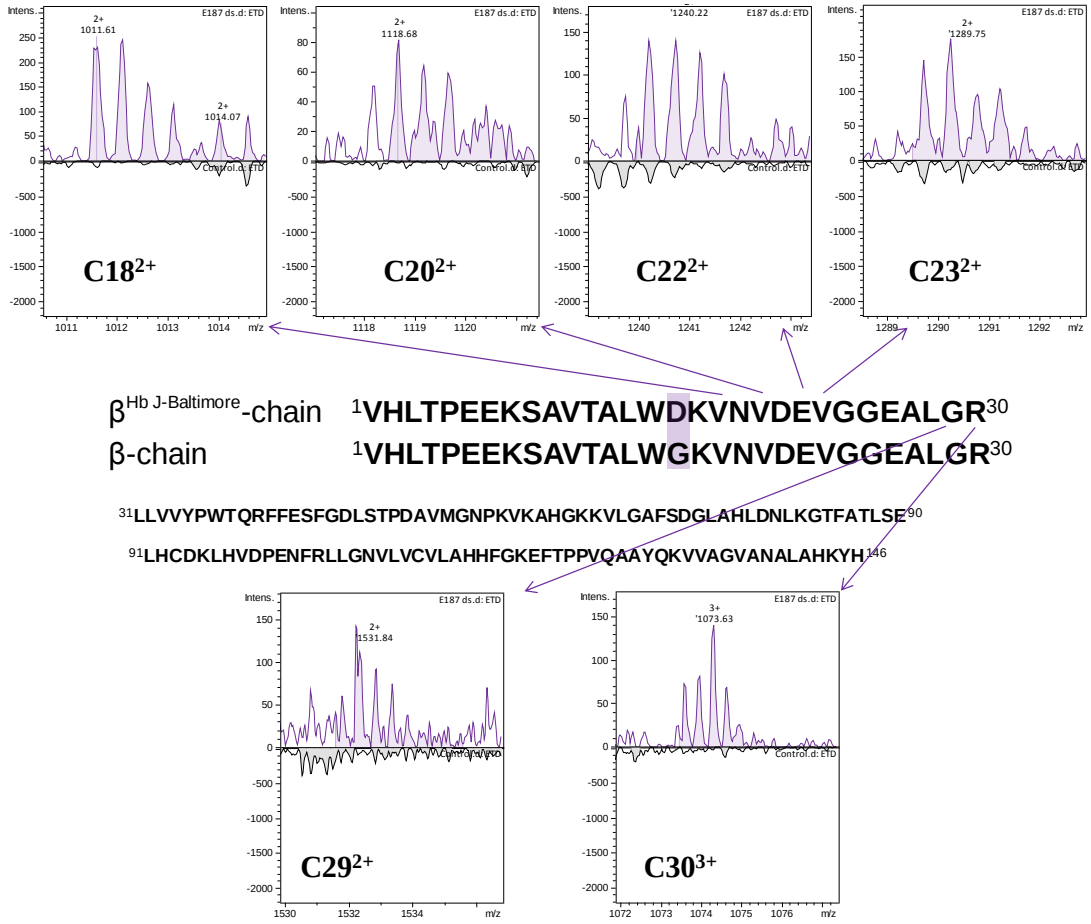


Figure 3.54. Fragment ions observed for sample E187 variant β-chain

3.4 Conclusion

3.4.1 Clinical trial with previously developed methods

Two mass spectrometry-based approaches utilizing Waters XEVO TQ instrument have been used in combination to analyse a total number of 2,017 samples for hemoglobinopathy diagnosis.

The different mass spectrometry based methods used for hemoglobinopathy screening have been evaluated for suitability in clinical screening, satisfying the need for reliable and fast measurement. The primary aim was to test the methods on clinical samples, and assess their robustness on population screening scale with the analysis of all samples arriving to a clinical lab in the period of twelve weeks. All samples have been analysed using both methods.

The first line testing using intact globin chain analysis was able to detect all samples with variants. Determination of the mass error of the β-globin chain indicated presence of -1 Da variants when compared to the mass error ranges of the normal samples within the same sample batch. Current issues which will need further development are the necessity to overcome the requirement for the analysis of several normal samples to provide the normal range for the actual measurement conditions. A calibration method for normalisation in combination of a statistical method to identify normal ranges will be necessary.

Intact globin chain analysis was able to indicate the presence of HbS variants for every analysed sample with this mutation, and also provided an estimation of the proportion of the variant chain to normal beta-chain.

Calculating intensity ratios of the δ-chain mass peak to the β-chain mass peak proved to be a reliable quantitation method for elevated HbA_2 levels, which quantirtative application has not been published and tested on a larger set of clinical samples before this study. A calibration which would allow for adjusting values to ce-HPLC values would be beneficial before sample measurements to assure the elevated values are always assigned the same way, and the reference ranges quoted are correct.

Nine samples have been identified with rare abnormal mutations, seven of these only detectable using mass spectrometry. Five out of seven of these samples have been characterised. One sample needs further mass spectrometric investigation with techniques such as high energy CID fragmentation or ECD to clarify what substitution has taken place. Two remaining samples need to be further characterised.

Intact globin chain analysis was able to give estimation regarding glycated beta-chain levels, which can be used for monitoring long-term diabetic conditions.

Reliable estimation of HbF levels was also possible, which is desirable based on NHS screening requirements. Elevated HbF levels can indicate the presence of some thalassemic conditions and also HPFH disease.

The direct infusion approach satisfies most of the requirements and so is regarded as currently the best choice for a rapid diagnostic screen. It enables the most information to be obtained in the shortest time frame for a relatively low cost. Currently this method needs expert staff to interpret the measured data. This could be alleviated by use of automated data interpretation software.

An MRM-method was used to confirm the presence of clinically significant variants such as HbS, HbE and HbC. It is also able to detect the Hb G-Philadelphia mutation presented in Chapter 2. The method confirmed all samples with the mentioned mutations, except one sample with HbS. The β^{SICKLE}-chain was only present in a low proportion (according to intact globin chain analysis 6.5%) and did not result in a MRM peak for the monitored β^{SICKLE}-chain peptide. The calculated proportions of β^{SICKLE}-chain based on peptide ratios differed significantly from the levels determined using the intact globin chain analysis. This observed occasional difference highlighted potential problems with the method. These are thought to be caused by incomplete enzymatic cleavage which may have affected the different globin-chains and the proportion of the resulted peptides differently.

By monitoring three pairs of peptides in the β- and δ-chains quantitation of HbA_2 levels was performed. The T2 and T3 peptide ratios did not provide as high a correlation to the reference ce-HPLC values as the novel T5 peptide ratios, therefore the T5 peptide ratios are proposed to be most appropriate for HbA_2 quantitation purposes. The use of the T5 peptides and the obtained well correlating results to current methods are also important because Daniel *et al.* presented MRM transitions

only for T2 and T3 peptides (Daniel *et al.* 2007), but the T5 peptide ratios have not previously been reported for the purpose of HbA_2 quantitation. Monitoring T5 peptide ratios provided values which allowed clear differentiation between normal and elevated values. The ranges observed for the normal values, however, did show some variation from batch to batch. The distribution of the normal levels did not prove to be as consistent as had been observed for intact globin chain analysis. The measurement of a calibration series with known HbA_2 levels would be beneficial for adjustment of the values and obtain the necessary consistency. It needs to be assessed whether this variation originated from measurement issues, or the previously mentioned potential problems with enzymatic cleavage. It is suggested that high quality sequencing grade trypsin for the sample preparation is evaluated to verify or exclude this hypothesis.
Other factors such as sample age and its influence on the resulted values also need to be evaluated.

It can be concluded that the combined use of the two MS methods was able to detect all clinically significant variants. MS methods were also able to detect rarer hemoglobin variants which were not detected using hospital methods. Both MS methods can be used for the quantitation of HbA_2. Intact globin chain analysis is also able to provide estimated values of glycated hemoglobin and HbF levels. For confirmation of HbD-Punjab mutation mass spectrometry analysis of the tryptic peptide mixture by a Q-ToF instrument was still required.
Intact globin chain analysis is not confirmative only indicative, while the MRM-method only confirms known and monitored variants, but cannot detect other abnormalities. A combination of the two methods could provide a solution for population screening. The two type of analysis can be performed on the same triple quadrupole instrument reducing the number of techniques necessary, since in the current clinical setting confirmative diagnosis is achieved using different techniques.

3.4.2 Pilot study using the top-down ETD method

The top-down ETD approach has also demonstrated a high potential for screening applications of hemoglobinopathies.

The developed method contains different measurement segments for different purposes. The initial analysis of denatured intact globin chains allows detection of mass shift in case of rare variants, has the potential for HbA_2, HbA_{1c} quantitation and for detecting abnormally elevated HbF levels.

The top down ETD method targets normal and unresolved β-chain variants within a 3 Da window on the multiply charged ion level. The β-chain and all clinically significant variant β-chains are simultaneously fragmented producing sequence information and characteristic fragment ions. The characteristic fragment ions are monitored for HbS, HbC, HbE, HbD-Punjab. Since HbO-Arab affects the same position of the β-chain as HbD-Punjab with both resulting in a -1 Da mass shift, these variants may be monitored using the same reporter fragment ion. The method developed for the preliminary clinical testing on a subset of the samples from the clinical trial identified all variants correctly, and confirmed the mutations based on the obtained fragments.

After a simple dilution step the method was capable of providing absolute confirmation of the monitored mutations. The structural information obtained is highly valuable information, which cannot be obtained by the use of current clinical screening techniques. The script developed for batch data processing is extremely useful for high-throughput applications to allow automation of data analysis. The generated pdf report is easy to use and its interpretation does not require expert staff which is also desirable for clinical laboratory applications.

The sample introduction method with a modified HPLC was not able to satisfy the high-throughput requirements. Samples were only not identified with the ETD-MS method due to the technical problems of the sample introduction.

The following issues need to be addressed to fulfil all the requirements against a screening method which would be acceptable by the NHS:

- Method optimisation and testing are required to assure the method is suitable for quantitation of δ-chain to allow detection of β-thalassemia
- Consistent measurement of glycated Hb needs to be established
- HbF levels need to be monitored to detect samples with elevated levels
- Method optimisation, or the development of an alternative approach, is required to enable discrimination between HbD-Punjab and HbO-Arab.

- An improved sample introduction is required (which is currently under development)
- Further automation of data processing, including the possibility of synchronisation with data acquisition for automated further investigation of cases with abnormal mass shifts would be beneficial

For identification of unknown variants a higher sequence coverage would be desirable, which may be obtained by combining CID and ETD fragmentation of the intact globin chains (Zubarev *et al.* 2008). MS^3 experiments derived from b- or y-type ions as intermediate product ions from a first fragmentation step with CID would generate sequence information regarding sections of the protein which are otherwise not covered by a simple one-step fragmentation experiment (Syka *et al.* 2004).

The application of top-down ETD technology for screening of structural hemoglobin variants have been tested, and the results are promising. Obtaining sequence information of the globin chains without the additional enzymatic cleavage step is a major advantage of this approach. The ETD based method is the method of choice for future developmental work.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

The work presented in this thesis has illustrated the use of mass spectrometry approaches in hemoglobinopathy diagnosis.

Mass spectrometry plays an important role in the analysis of biological molecules today. Human diseases may be diagnosed from the analysis of certain biomolecules which are responsible for the clinical condition (eg. mutant DNA and modified proteins) or molecules which are involved in disease related biochemical processes during the development of the disease. The speed, specificity and sensitivity of mass spectrometry make this technique attractive for clinical applications where rapid identification and characterization of proteins is necessary (Reid and McLuckey 2002).

Mass spectrometry has the potential to become a leading technique in the field of hemoglobinopathy diagnosis, but it is important to clarify that it will never provide the gold standard. DNA analysis is the only way of identifying with certainty what exact genetic mutation has taken place, from that the manifestations of the genetic mutations, either affecting the structure or the level of production of certain globin chains, can be determined.

The primary aim of developing MS-based approaches is to avoid difficult and expensive DNA analysis in most cases marking it possible to support current antenatal screening programs. The application of mass spectrometry-based methods has been shown to provide this. With the definitive diagnosis of unknown variants, which can be obtained by using mass spectrometry, the number of samples that require DNA analysis has been very significantly decreased.

Despite several publications regarding the use of MS in this field, it has not yet been implemented as a routine technique in hematology labs. The current techniques (HPLC and IEF) provide satisfactory diagnosis of clinically significant variants. At the moment MS is primarily used for the identification of abnormal structural hemoglobin variants following the protocol from Wild *et al.* (Wild et al. 2001).

There have been several attempts to develop MS-based screening methods, with the highest potential provided by an MRM-based approach by Daniel et al. (Daniel *et al.* 2005, Daniel *et al.* 2007) Thousands of samples have been analysed and it has been shown that targeted peptide monitoring could provide an automated high-throughput method for screening of the clinically significant hemoglobin variants.

4.1 Evaluation of the combination of two MS-based approaches using a triple quadrupole mass spectrometer

It has been shown that a triple quadrupole system can be used to obtain diagnosis of several hemoglobin disorders. Two methods employed in combination were able to provide confirmation of the presence of all clinically significant mutations.

Intact globin chain analysis provided information regarding the presence of additional variant globin chains, and with mass error calculations were able to provide indication of -1 Da variants. It also provided estimation of HbA_2 and glycated hemoglobin levels, and detected elevated HbF levels. During the clinical trial the method detected seven hemoglobin variants which were chromatographically not detectable. In preliminary testing it also highlighted problems with current clinical methods, such as when a sample classed, as HbD-Punjab, was identified with certainty by MS as HbG-Philadelphia. This kind of misdiagnosis, when a sample is falsely diagnosed as HbD-Punjab, one of the clinically significant variants, is unacceptable and highlights that even the use of two different techniques can confirm a condition incorrectly.
Targeted peptide analysis as a necessary second line method using the same triple quadrupole instrument provided confirmation of known variants HbS, HbE and HbC and an estimation of of HbA_2 levels. New MRM method and transitions for the use of the T5 peptide ratios in the normal β- and minor δ-chain is a novel result , previously not presented in the literature. The reliable monitoring of this specific peptide pair is a promising achievement of the work conducted to provide a competitive method for future clinical application.

Both methods were capable of providing an estimation of HbA_2 levels, and were able to distinguish between normal and elevated levels. Despite an obvious correlation between HbA_2 values obtained using the ce-HPLC method and the different mass spectrometry based methods, values differ as shown in the statistical analysis. All three methods measure different chemical entities. The ce-HPLC method measures the quantity of αδ dimers, a building block of protein HbA_2, and reports based on UV response. The intact globin-chain analysis determines the ratio between the β-

and δ- globin chains. The MRM method measures the peptide ratios by monitoring fragmentation routes. All three analyses are performed using different chemical properties. An important fact which needs to be emphasised is that both MS-based methods were in agreement with HPLC and provided confirmation regarding the HbA_2 level of a sample being in a normal or abnormal range. It is important to emphasize that neither the intact globin chain analysis nor the T5 peptide ratio monitoring have been previously tested on a large number of antenatal samples, and both show good correlation and promise to compete against current quantitative chromatographic techniques.

Future work

Based on observations resulting from the clinical trial the following method optimisation issues need to be addressed:

- Intact globin chain analysis is the suggested method for HbA_2 quantitation and further developments and evaluation of this method is targeted. A calibration method needs to be developed to eliminate batch to batch variations and to provide an absolute range for normal samples to ensure that elevated levels are detected at any time under any instrumental conditions with high confidence.
- The development of an automated screening method for HbD-Punjab mutation would be beneficial.
- The MRM-method may be extended to monitor the presence of peptides of rare variants identified during the clinical trial for further screening purposes.
- The remaining two samples with α- and β- chain variants will be characterised.
- Issues encountered regarding the efficiency of the enzymatic cleavage need to be investigated and the problem eliminated to confident diagnostic value and also reliable quantitation which is required for HbC and HbE variant level determination.

As part of more fundamental research ion mobility studies can be carried out to better characterise different clinically significant variants. HbS showed different collisional cross sections in a previously presented publication (Scarff *et al.* 2009).

Ion mobility analysis of other clinically significant variants may reveal certain similarities or differences in the shape of the globin-chain dimers to provide insight into structure-function relationships. Ion mobility mass spectrometry have also been successfully employed in challenging cases before (Williams *et al.* 2008), additional potential future applications may be considered when this technique becomes available on a less expensive bench top mass spectrometer, which would also be more likely to be employed in a clinical laboratory.

4.2 Evaluation of top-down ETD-based MS approach on an iontrap system

In collaboration with Bruker Daltonics the developed top-down ETD method has shown great potential for automated hemoglobinopathy screening. Monitoring the presence of characteristic fragment ions of the intact globin chain provided confirmation of clinically significant variants and additional fragmentation information of rare hemoglobin variants. A significant advantage of the method is that it provides sequence information without the need for enzymatic cleavage. No expert staff are required to interpret the results using the automatically generated report. Confident diagnosis of HbD-Punjab and HbO-Arab mutation in an automated manner can be achieved, which proved to be challenging with the other MS methods and required full scan tryptic peptide analysis on a Q-ToF instrument.

Future work

The developed method, to realise its full potential, needs further optimisation and technical improvements. A reliable low cost sample introduction system would allow intact globin chain analysis section to be optimised and evaluated for HbA_2, HbA_{1c} and HbF level estimations similarly as achieved using the triple quadrupole instrument with NanoMate sample introduction.

As has been pointed out by Graca et al. even complete MS-based sequence analysis will not be able to give definitive answers in case of rare variants like Hb Machida, which result in the same mass shift at the exact same position on the β-chain as HbC (Graca *et al.* 2012). Differentiation also needs to be achieved for Hb D-Punjab and Hb O-Arab mutations. This needs to be addressed in further development of the top-

down ETD method to be able to distinguish between these two clinically significant variants.
It also needs to be investigated whether quantitation of the different -1 Da variants could be achieved based on ratio calculations between normal and variant fragment ions.

Concluding remarks

When comparing the methods to decide which one has the higher potential to become an applied technique in the clinical laboroartory, it has to be considered which method can offer the most information at the lowest cost. The aim is to fulfil the NHS requirements set in the Sickle cell and thalassemia screening Programme (NHS 2012b). The MRM method may have potential for screening for specific known variants, but intact protein mass can provide additional information not available from targeted peptide analysis. To maximise the obtained information the two methods may need to be used in combination. This adds complication and makes the screening process long and since it requires more analytical steps that introduce greater variability.

The ETD method has a significant chance of acceptance in the clinical lab, since it only requires a simple dilution sample preparation step, which could be automated in the future. The overall data acquisition time (15 minutes) is not shorter than the sum of the other two MS approaches, but turn-around time is significantly less due to the absence of the need for lengthy enzymatic cleavage.
As proposed extension of the clinical trial to other regional laboratories would generate at least as much data as the initial trial giving further confidence in the approach.

Clinicians remain sceptical, since mass spectrometry has not yet shown the ability to detect all existing variants, but databases and extended methods are constantly improving to provide the highest level of information possible using this technique. Since mass spectrometry is one of the fastest developing fields in science and recent developments target clinical applications, the technique will no doubt become a leading routine technique in medical diagnosis in the near future.

REFERENCES

Abian, J. (1999). The coupling of gas and liquid chromatography with mass spectrometry. *Journal of Mass Spectrometry*, 34. **3.** 157-+.

Abonnenc, M., Qiao, L. A., Liu, B. H. and Girault, H. H. (2010). Electrochemical aspects of electrospray and laser desorption/ionization for mass spectrometry. Annual Review of Analytical Chemistry, Vol 3. Annual Reviews.

Adaway, J. E. and Keevil, B. G. (2012). Therapeutic drug monitoring and LC-MS/MS. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 883. 33-49.

Almeida, A. M., Henthorn, J. S. and Davies, S. C. (2001). Neonatal screening for haemoglobinopathies: the results of a 10-year programme in an English Health Region. *British Journal of Haematology*, 112. **1.** 32-35.

Ardrey, R. E. (2003). Liquid chromatography - mass spectrometry: An introduction. John Wiley & Sons, Ltd.

Arpino, P. J., Dawkins, B. G. and McLafferty, F. W. (1974). Liquid chromatography mass spectrometry system providing continous monitoring with nanogram sensitivity. *Journal of Chromatographic Science*, 12. **10.** 574-578.

Aygun, B. and Odame, I. (2012). A global perspective on sickle cell disease. *Pediatric Blood & Cancer*, 59. **2.** 386-390.

Babic, N. (2012). Analytical methods and perfomance of the immunoassay methods for determination of vitamin D in comaprison to mass spectrometry. *Journal of Medical Biochemistry*, 31. **4.** 333-338.

Bain, B. J. (2006a). Haemoglobin and the genetics of haemoglobin synthesis. Haemoglobinopathy Diagnosis. Wiley. Chapter 1

Bain, B. J. (2006b). Laboratory techniques for the identification of abnormalities of globin chain synthesis. Haemoglobinopathy Diagnosis. Wiley. Chapter 2

Bain, B. J. (2009). Neonatal/newborn haemoglobinopathy screening in Europe and Africa. *Journal of Clinical Pathology*, 62. **1.** 53-56.

Bain, B. J. (2011). Haemoglobinopathy diagnosis: Algorithms, lessons and pitfalls. *Blood Reviews*, 25. **5.** 205-213.

Barber, M., Bordoli, R. S., Sedgwick, R. D. and Tyler, A. N. (1981). Fast Atom Bombardment of solids (FAB) - A new ion-source for mass-spectrometry *Journal of the Chemical Society-Chemical Communications*, **7.** 325-327.

Basilico, F., Di Silvestre, D., Sedini, S., Petretto, A., Levreri, I., Melioli, G., Farina, C., Mori, F. and Mauri, P. L. (2007). New approach for rapid detection of known hemoglobin variants using LC-MS/MS combined with a peptide database. *Journal of Mass Spectrometry*, 42. **3.** 288-292.

Batoy, S., Akhmetova, E., Miladinovic, S., Smeal, J. and Wilkins, C. L. (2008). Developments in MALDI mass spectrometry: The quest for the perfect matrix. *Applied Spectroscopy Reviews*, 43. **6.** 485-550.

Beckey, H. D. (1969). Field desorption mass spectrometry: A technique for the study of thermally unstable substances of low volatility. *International Journal of Mass Spectrometry and Ion Physics*, 2. **6.** 500-502.

Biemann, K. (1992). Mass-spectrometry of peptides and proteins. *Annual Review of Biochemistry*, 61. 977-1010.

Bilic-Zulle, L. (2011). Comparison of methods: Passing and Bablok regression. *Biochemia Medica*, 21. **1.** 49-52.

Bland, J. M. and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement *Lancet*, 1. **8476.** 307-310.

Bland, J. M. and Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8. **2.** 135-160.

Boemer, F., Cornet, Y., Libioulle, C., Segers, K., Bours, V. and Schoos, R. (2011). 3-years experience review of neonatal screening for hemoglobin disorders using tandem mass spectrometry. *Clinica Chimica Acta*, 412. **15-16.** 1476-1479.

Boemer, F., Ketelslegers, O., Minon, J.-M., Bours, V. and Schoos, R. (2008). Newborn screening for sickle cell disease using tandem mass spectrometry. *Clinical Chemistry*, 54. **12.** 2036-2041.

Boemer, F., Ketelslegers, O., Minon, J. M., Bours, V. and Schoos, R. (2009). Newborn screening for sickle cell disease through use of tandem mass spectrometry in reply. *Clinical Chemistry*, 55. **6.** 1244-1245.

Bruins, A. P., Covey, T. R. and Henion, J. D. (1987). Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry. *Analytical Chemistry*, 59. **22.** 2642-2646.

Buxbaum, E. (2007). Haemoglobin and myoglobin: co-operativity. Fundamentals of Protein Structure and Function. Springer.

Chace, D. H. (2003). Mass spectrometry-based diagnostics: The upcoming revolution in disease detection has already arrived. *Clinical Chemistry*, 49. **7.** 1227-1228.

Chace, D. H. (2009). Mass spectrometry in newborn and metabolic screening: historical perspective and future directions. *Journal of Mass Spectrometry*, 44. **2.** 163-170.

Chevenne, D., Marle, N., Chauffert, M., Noel, M., Ducrocq, R. and Trivin, F. (1999). Evaluation of the tosoh HLC-723GHb V A1c 2.2 hemoglobin A1c analyzer. *Clinical Biochemistry*, 32. **6.** 487-490.

Chi, A., Huttenhower, C., Geer, L. Y., Coon, J. J., Syka, J. E. P., Bai, D. L., Shabanowitz, J., Burke, D. J., Troyanskaya, O. G. and Hunt, D. F. (2007). Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 104. **7.** 2193-2198.

Clarke, G. M. and Higgins, T. N. (2000). Laboratory investigation of hemoglobinopathies and thalassemias: Review and update. *Clinical Chemistry*, 46. **8.** 1284-1290.

Cleophas, T. J., Zwinderman, A.H., Cleophas, T.F., Cleophas, E.P. (2009). Statistics applied to clinical trials. 4th Edition, Springer Science + Business Media B.V. .

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1. **3.** 98-101.

Coles, J. and Guilhaus, M. (1993). Orthogonal acceleration - ” a new direction for time-of-flight mass spectrometry: Fast, sensitive mass analysis for continuous ion sources. *TrAC Trends in Analytical Chemistry*, 12. **5.** 203-213.

Comisarow, M. B. and Marshall, A. G. (1974). Fourier-tansform ion-cyclotron resonance mass-spectroscopy. *Chemical Physics Letters*, 25. **2.** 282-283.

Cook, S. L., Zimmermann, C. M., Singer, D., Fedorova, M., Hoffmann, R. and Jackson, G. P. (2012). Comparison of CID, ETD and metastable atom-activated dissociation (MAD) of doubly and triply charged phosphorylated tau peptides. *Journal of Mass Spectrometry*, 47. **6.** 786-794.

Cooks, R. G. (1973). Metastable ions. Elsevier Scientific Pub. Co.

Coon, J. J., Ueberheide, B., Syka, J. E. P., Dryhurst, D. D., Ausio, J., Shabanowitz, J. and Hunt, D. F. (2005). Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 102. **27.** 9463-9468.

Covey, T. R., Lee, E. D., Bruins, A. P. and Henion, J. D. (1986). Liquid-chromatography-mass specrometry *Analytical Chemistry*, 58. **14.** 1451A.

Covey, T. R., Thomson, B. A. and Schneider, B. B. (2009). Atmospheric pressure ion sources. *Mass Spectrometry Reviews*, 28. **6.** 870-897.

Creese, A. J. and Cooper, H. J. (2007). Liquid chromatography electron capture dissociation tandem mass spectrometry (LC-ECD-MS/MS) versus liquid chromatography collision-induced dissociation tandem mass spectrometry (LC-CID-MS/MS) for the identification of proteins. *Journal of the American Society for Mass Spectrometry*, 18. **5.** 891-897.

Cristoni, S. and Bernardi, L. R. (2003). Development of new methodologies for the mass spectrometry study of bioorganic macromolecules. *Mass Spectrometry Reviews*, 22. **6.** 369-406.

Cui, W. D., Rohrs, H. W. and Gross, M. L. (2011). Top-down mass spectrometry: Recent developments, applications and perspectives. *Analyst*, 136. **19.** 3854-3864.

Daniel, Y. A., Turner, C., Farrar, L. and Dalton, R. N. (2008). A comparison of IEF and MSMS for clinical hemoglobinopathy screening in 40,000 newborns. 50th American Society of Hematology annual meeting. San Francisco, CA, USA. http://ash.confex.com/ash/2008/paper9728.html

Daniel, Y. A., Turner, C., Haynes, R. M., Hunt, B. J. and Dalton, R. N. (2005). Rapid and specific detection of clinically significant haemoglobinopathies using electrospray mass spectrometry/mass spectrometry. *British Journal of Haematology*, 130. **4.** 635-643.

Daniel, Y. A., Turner, C., Haynes, R. M., Hunt, B. J. and Dalton, R. N. (2007). Quantification of hemoglobin A2 by tandem mass spectrometry. *Clinical Chemistry*, 53. **8.** 1448-1454.

De Caterina, M., Esposito, P., Grimaldi, E., Dimaro, G., Scopacasa, F., Ferranti, P., Parlapiano, A., Malorni, A., Pucci, P. and Marino, G. (1992). Characterization of hemoglobin Lepore variants by advanced mass-spectrometric procedures *Clinical Chemistry*, 38. **8.** 1444-1448.

de Hoffmann, E. and Stroobant, V. (2009). Mass spectrometry: Principles and applications. Third edition, Wiley.

Dettmer, K., Aronov, P. A. and Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26. **1.** 51-78.

Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D. and Alice, M. B. (1968). Molecular beams of macroions. *The Journal of Chemical Physics*, 49. **5.** 2240-2249.

Domon, B. and Aebersold, R. (2006). Review - Mass spectrometry and protein analysis. *Science*, 312. **5771.** 212-217.

Domon, B., Bodenmiller, B., Carapito, C., Hao, Z. Q., Huehmer, A. and Aebersold, R. (2009). Electron transfer dissociation in conjunction with collision activation to investigate the *Drosophila melanogaster* phosphoproteome. *Journal of Proteome Research*, 8. **6.** 2633-2639.

Douglas, D. J. (2009). Linear quadrupoles in mass spectrometry. *Mass Spectrometry Reviews*, 28. **6.** 937-960.

Douglas, D. J., Frank, A. J. and Mao, D. (2005). Linear ion traps in mass spectrometry. *Mass Spectrometry Reviews*, 24. **1.** 1-29.

Drabik, A., Bierczynska-Krzysik, A., Bodzon-Kulakowska, A., Suder, P., Kotlinska, J. and Silberring, J. (2007). Proteomics in neurosciences. *Mass Spectrometry Reviews*, 26. **3.** 432-450.

Edwards, R. L., Creese, A. J., Baumert, M., Griffiths, P., Bunch, J. and Cooper, H. J. (2011). Hemoglobin variant analysis via direct surface sampling of dried blood spots coupled with high-resolution mass spectrometry. *Analytical Chemistry*, 83. **6.** 2265-2270.

Edwards, R. L., Griffiths, P., Bunch, J. and Cooper, H. J. (2012). Top-down proteomics and direct surface sampling of neonatal dried blood spots: diagnosis of unknown hemoglobin variants. *Journal of the American Society for Mass Spectrometry*, 23. **11.** 1921-1930.

El-Aneed, A., Cohen, A. and Banoub, J. (2009). Mass spectrometry, review of the basics: Electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews*, 44. **3.** 210-230.

Falick, A. M., Shackleton, C. H. L., Green, B. N. and Witkowska, H. E. (1990). Tandem mass spectrometry in the clinical analysis of variant hemoglobins. *Rapid Communications in Mass Spectrometry*, 4. **10.** 396-400.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science (New York, N.Y.)*, 246. **4926.** 64-71.

Flad, T. and Tolson, J. (2005). Mass spectrometry meets medical sciences: making headway in molecular disease diagnostics. *Analytical and Bioanalytical Chemistry*, 381. **1.** 24-27.

Gabbay, K. H., Sosenko, J. M., Banuchi, G. A., Mininsohn, M. J. and Fluckiger, R. (1979). Glycosylated hemoglobins - increased glycosylation of hemoglobin in diabetic-patients. *Diabetes*, 28. **4.** 337-340.

Glish, G. L. and Burinsky, D. J. (2008). Hybrid mass spectrometers for tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 19. **2.** 161-172.

Good, D. M., Wirtala, M., McAlister, G. C. and Coon, J. J. (2007). Performance characteristics of electron transfer dissociation mass spectrometry. *Molecular & Cellular Proteomics*, 6. **11.** 1942-1951.

Graca, D. C., Lescuyer, P., Clerici, L., Tsybin, Y. O., Hartmer, R., Meyer, M., Samii, K., Hochstrasser, D. F. and Scherl, A. (2012). Electron transfer dissociation

mass spectrometry of hemoglobin on clinical samples. *Journal of the American Society for Mass Spectrometry*, 23. **10.** 1750-1756.

Green, B. N., Oliver, R. W. A., Falick, A. M., Shackleton, C. H. L., Roitman, E. and Witkowska, H. E. (1990). Electrospray MS, LSIMS and MSMS for the rapid detection and characterization of variant hemoglobins. Biological Mass Spectrometry /. Elsevier Science Publ B V.

Gross, J. (2011). Tandem mass spectrometry. Mass Spectrometry. Springer Berlin Heidelberg.

Guilhaus, M., Selby, D. and Mlynski, V. (2000). Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrometry Reviews*, 19. **2.** 65-107.

Guthrie, R. (1969). Screening for phenylketonuria. *Triangle; the Sandoz journal of medical science*, 9. **3.** 104-109.

Guthrie, R. and Susi, A. (1963). A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics*, 32. 338-343.

Hager, J. W. (2002). A new linear ion trap mass spectrometer. *Rapid Communications in Mass Spectrometry*, 16. **6.** 512-526.

Han, H., Xia, Y., Yang, M. and McLuckey, S. A. (2008). Rapidly alternating transmission mode electron-transfer dissociation and collisional activation for the characterization of polypeptide ions. *Analytical Chemistry*, 80. **9.** 3492-3497.

Hardison, R. C., Chui, D. H. K., Giardine, B., Riemer, C., Patrinos, G. P., Anagnou, N., Miller, W. and Wajcman, H. (2002). HbVar. A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. *Human Mutation*, 19. **3.** 225-233.

Hart, S. R., Lau, K. W., Gaskell, S. J. and Hubbard, S. J. (2011). Distributions of ion series in ETD and CID spectra: Making a comparison. Data Mining in Proteomics: From Standards to Applications. Humana Press Inc, 999 Riverview Dr, Ste 208, Totowa, Nj 07512-1165 USA.

Hartmer, R. and Lubeck, M. (2005). New approach for characterization of post translational modified peptides using ion trap MS with combined ETD/CID fragmentation. *Lc Gc Europe*, 11-13.

Hartwell, S. K., Srisawang, B., Kongtawelert, P., Christian, G. D. and Grudpan, K. (2005). Review on screening and analysis techniques for hemoglobin variants and thalassemia. *Talanta*, 65. **5.** 1149-1161.

HbVar: A database of human hemoglobin variants and thalassemias, http://globin.bx.psu.edu/hbvar/menu.html. (Accessed 22 April, 2013)

Henderson, S., Timbs, A., McCarthy, J., Gallienne, A., Van Mourik, M., Masters, G., May, A., Khalil, M. S. M., Schuh, A. and Old, J. (2009). Incidence

of haemoglobinopathies in various populations - The impact of immigration. *Clinical Biochemistry*, 42. **18.** 1745-1756.

Henion, J. D. (2009). Citation classic the origins of ion spray liquid chromatography-Tandem mass spectrometry. *Clinical Chemistry*, 55. **6.** 1234-1235.

Herman, J. and Shushan, B. (2012). Clinical applications. Tandem Mass Spectrometry - Applications and Principles. Prasain, D. J. InTech Chapter 28

Hickman, M., Modell, B., Greengross, P., Chapman, C., Layton, M., Falconer, S. and Davies, S. C. (1999). Mapping the prevalence of sickle cell and beta thalassaemia in England: estimating and validating ethnic-specific rates. *British Journal of Haematology*, 104. **4.** 860-867.

Higashi, T., Shimada, K. and Toyo'oka, T. (2010). Advances in determination of vitamin D related compounds in biological samples using liquid chromatography-mass spectrometry: A review. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 878. **20.** 1654-1661.

Higgins, T. N., Khajuria, A. and Mack, M. (2009). Quantification of HbA(2) in patients with and without beta-thalassemia and in the presence of HbS, HbC, HbE, and HbD Punjab hemoglobin variants Comparison of Two Systems. *American Journal of Clinical Pathology*, 131. **3.** 357-362.

Hipple, J. A., Sommer, H. and Thomas, H. A. (1949). A precise method of determining the Faraday by magnetic resonance. *Physical Review*, 76. **12.** 1877-1878.

Hogan, C. J., Carroll, J. A., Rohrs, H. W., Biswas, P. and Gross, M. L. (2009). Combined charged residue-field emission model of macromolecular electrospray ionization. *Analytical Chemistry*, 81. 369.

Horvath, C. G. and Lipsky, S. R. (1966). Use of liquid ion exchange chromatography for the separation of organic compounds. *Nature*, 211. **5050.** 748-749.

Horvath, C. G., Preiss, B. A. and Lipsky, S. R. (1967). Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers. *Analytical Chemistry*, 39. **12.** 1422-1428.

Human Tissue Act (2004) c 30 (http://www.legislation.gov.uk/ukpga/2004/30/contents)

Indovina, P., Marcelli, E., Pentimalli, F., Tanganelli, P., Tarro, G. and Giordano, A. (2012). Mass spectrometry-based proteomics: The road to lung cancer biomarker discovery. *Mass Spectrometry Reviews*, n/a-n/a.

Iribarne, J. V. and Thomson, B. A. (1976). Evaporation of small ions from charged droplets. *Journal of Chemical Physics*, 64. **6.** 2287-2294.

Itano, H. A., Bergren, W. R. and Sturgeon, P. (1954). Identification of a fourth abnormal human hemoglobin *Journal of the American Chemical Society*, 76. **8.** 2278-2278.

Jennings, K. R. (1968). Collision-induced decompositions of aromatic molecular ions. *International Journal of Mass Spectrometry and Ion Physics*, 1. **3.** 227-235.

Jennings, K. R. (2000). The changing impact of the collision-induced decomposition of ions on mass spectrometry. *International Journal of Mass Spectrometry*, 200. **1-3.** 479-493.

Jiwan, J. L. H., Wallemacq, P. and Herent, M. F. (2011). HPLC-high resolution mass spectrometry in clinical laboratory? *Clinical Biochemistry*, 44. **1.** 136-147.

John, G. and English, E. (2012). IFCC standardised HbA(1c): should the world be as one? *Clinical Chemistry and Laboratory Medicine*, 50. **7.** 1243-1248.

Karapirli, M., Kizilgun, M., Yesilyurt, O., Gul, H., Kunak, Z. I., Akgul, E. O., Macit, E., Cayci, T., Kurt, Y. G., Aydin, I., Yaren, H., Seyrek, M., Cakir, E. and Yaman, H. (2012). Simultaneous determination of cyclosporine A, tacrolimus, sirolimus, and everolimus in whole-blood samples by LC-MS/MS. *Scientific World Journal*, 8.

Karas, M., Bachmann, D., Bahr, U. and Hillenkamp, F. (1987). Matrix-Assisted Ultraviolet-Laser Desorption of nonvolatile compounds. . *International Journal of Mass Spectrometry and Ion Processes*, 78. 53-68.

Karas, M., Bahr, U. and Dulcks, T. (2000). Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. *Fresenius Journal of Analytical Chemistry*, 366. **6-7.** 669-676.

Karas, M. and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*, 60. 2299-2301.

Kebarle, P. and Verkerk, U. H. (2009). Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews*, 28. **6.** 898-917.

Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K. and McLafferty, F. W. (1999). Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *Journal of the American Chemical Society*, 121. **4.** 806-812.

Kim, M.-S., Zhong, J., Kandasamy, K., Delanghe, B. and Pandey, A. (2011). Systematic evaluation of alternating CID and ETD fragmentation for phosphorylated peptides. *Proteomics*, 11. **12.** 2568-2572.

Kleinert, P., Schmid, M., Zurbriggen, K., Speer, O., Schmugge, M., Roschitzki, B., Durka, S. S., Leopold, U., Kuster, T., Heizmann, C. W., Frischknecht, H. and Troxler, H. (2008). Mass spectrometry: A tool for enhanced detection of hemoglobin variants. *Clinical Chemistry*, 54. **1.** 69-76.

Kliem, M. and Sauer, S. (2012). The essence on mass spectrometry based microbial diagnostics. *Current Opinion in Microbiology*, 15. **3.** 397-402.

Knochenmuss, R. (2006). Ion formation mechanisms in UV-MALDI. *Analyst*, 131. 966-986.

Kobold, U., Jeppsson, J.-O., Dulffer, T., Finke, A., Hoelzel, W. and Miedema, K. (1997). Candidate reference methods for hemoglobin A_{1c} based on peptide mapping. *Clinical Chemistry*, 43. **10.** 1944-1951.

Kolch, W., Neusüß, C., Pelzing, M. and Mischak, H. (2005). Capillary electrophoresis–mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. *Mass Spectrometry Reviews*, 24. **6.** 959-977.

Kollipara, S., Bende, G., Agarwal, N., Varshney, B. and Paliwal, J. (2011). International Guidelines for Bioanalytical Method Validation: A Comparison and Discussion on Current Scenario. *Chromatographia*, 73. **3-4.** 201-217.

Konermann, L., Ahadi, E., Rodriguez, A. D. and Vahidi, S. (2012). Unraveling the mechanism of electrospray ionization. *Analytical Chemistry*,

Koppenaal, D. W., Barinaga, C. J., Denton, M. B., Sperline, R. P., Hieftje, G. M., Schilling, G. D., Andrade, F. J. and Barnes, J. H. (2005). MS detectors. *Analytical Chemistry*, 77. **21.** 418A-427A.

Kushnir, M. M., Rockwood, A. L. and Bergquist, J. (2009). Liquid chromatography–tandem mass spectrometry applications in endocrinology. *Mass Spectrometry Reviews*, 29. **3.** 480-502.

Lai, Y.-H., Wang, C.-C., Lin, S.-H., Lee, Y. T. and Wang, Y.-S. (2010). Solid-phase thermodynamic interpretation of ion desorption in matrix-assisted laser desorption/ionization. *Journal of Physical Chemistry B*, 114. **43.** 13847-13852.

Lawrence, J. S. and Valentine, W. N. (1955). Abnormal hemoglobins; clinical disorders resulting from various combinations. *California medicine*, 82. **1.** 1-5.

Leung, W. C., Leung, K. Y., Lau, E. T., Tang, M. H. Y. and Chan, V. (2008). Alpha-thalassaemia. *Seminars in Fetal & Neonatal Medicine*, 13. **4.** 215-222.

Light-Wahl, K. J., Loo, J. A., Edmonds, C. G., Smith, R. D., Witkowska, H. E., Shackleton, C. H. L. and Wu, C. S. C. (1993). Collisionally activated dissociation and tandem mass spectrometry of intact hemoglobin Beta-chain variant proteins with electrospray ionization *Biological Mass Spectrometry*, 22. **2.** 112-120.

Loo, J., Edmonds, C. and Smith, R. (1990). Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. *Science*, 248. **4952.** 201-204.

Lubec, G. and Afjehi-Sadat, L. (2007). Limitations and pitfalls in protein identification by mass spectrometry. *Chemical Reviews*, 107. **8.** 3568-3584.

Lubin, B. H., Witkowska, H. E. and Kleman, K. (1991). Laboratory diagnosis of hemoglobinopathies. *Clinical Biochemistry*, 24. **4.** 363-374.

Madeira, P. J. A. and Florêncio, M. H. (2012). Applications of tandem mass spectrometry: From structural analysis to fundamental studies. Tandem Mass Spectrometry - Applications and Principles. Prasain, D. J. InTech.

Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Analytical Chemistry*, 72. **6.** 1156-1162.

Maldaner, L. and Jardim, I. (2009). The state of art of ultra performance liquid chromatography *Quimica Nova*, 32. **1.** 214-222.

March, R. E. (2009). Quadrupole ion traps. *Mass Spectrometry Reviews*, 28. **6.** 961-989.

Matta, A., Ralhan, R., DeSouza, L. V. and Siu, K. W. M. (2010). Mass spectrometry-based clinical proteomics: Head-and-neck cancer biomarkers and drug-targets discovery. *Mass Spectrometry Reviews*, 29. **6.** 945-961.

McDonald, P. D. (2009). The quest for ultra performance in liquid chromatography. Waters Corporation.

McIntyre, D. (2004). Advantages of wide dynamic range on an orthogonal acceleration time-of-flight mass spectrometer. *Technical Overview, Agilent Technologies*,

McLafferty, F. W. and Bryce, T. A. (1967). Metastable-ion characteristics: characterization of isomeric molecules. *Chemical Communications (London)*, 0. **23.** 1215-1217.

McLuckey, S. A. and Stephenson, J. L. (1998). Ion/ion chemistry of high-mass multiply charged ions. *Mass Spectrometry Reviews*, 17. **6.** 369-407.

Mekecha, T. T., Amunugama, R. and McLuckey, S. A. (2006). Ion trap collision-induced dissociation of human hemoglobin alpha-chain cations. *Journal of The American Society for Mass Spectrometry*, 17. **7.** 923-931.

Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J., Syka, J. E. P., Shabanowitz, J. and Hunt, D. F. (2006). The utility of ETD mass spectrometry in proteomic analysis. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, 1764. **12.** 1811-1822.

Mirgorodskaya, O. A., Shevchenko, A. A., Chernushevich, I. V., Dodonov, A. F. and Miroshnikov, A. I. (1994). Electrospray-ionization time-of-flight mass spectrometry in protein chemistry. *Analytical Chemistry*, 66. **1.** 99-107.

Modell, B., Darlison, M., Birgens, H., Cario, H., Faustino, P., Giordano, P. C., Gulbis, B., Hopmeier, P., Lena-Russo, D., Romao, L. and Theodorsson, E. (2007). Epidemiology of haemoglobin disorders in Europe: an overview. *Scandinavian Journal of Clinical & Laboratory Investigation*, 67. **1.** 39-69.

Molina, H., Matthiesen, R., Kandasamy, K. and Pandey, A. (2008). Comprehensive comparison of collision induced dissociation and electron transfer dissociation. *Analytical Chemistry*, 80. **13.** 4825-4835.

Morris, H. R., Panico, M., Barber, M., Bordoli, R. S., Sedgwick, R. D. and Tyler, A. (1981). Fast Atom Bombardment - A new mass-spectrometric method for peptide sequence-analysis *Biochemical and Biophysical Research Communications*, 101. **2.** 623-631.

Moss, C. L. and Tureček, F. (2012). Protonation sites in peptide dications and cation-radicals containing β-amino acid residues. *International Journal of Mass Spectrometry*, 316–318. **0.** 57-67.

Nakanishi, T., Miyazaki, A., Shimizu, A., Yamaguchi, A. and Nishimura, S. (2002). Assessment of the effect of hemoglobin variants on routine HbA_{1c} measurements by electrospray ionization mass spectrometry. *Clinica Chimica Acta*, 323. **1-2.** 89-101.

Nguyen, S. and Fenn, J. B. (2007). Gas-phase ions of solute species from charged droplets of solutions. *Proceedings of the National Academy of Sciences*, 104. **4.** 1111-1117.

NHS (2012a). NHS Sickle Cell and Thalassemia Screening Programme. Programme Review 2011/2012

NHS (2012b). NHS Sickle Cell and Thalassemia Screening Programme. Handbook for Laboratories. 3rd edition

NHS, M. F. T.: Haematology Reference ranges, http://www.medway.nhs.uk/for-clinicians/pathology/haematology/reference-ranges/. (Accessed 22 April 2013)

Parker, C. E., Warren, M. R. and Mocanu, V. (2010). Neuroproteomics. CRC Press.

Passing, H. and Bablok, W. (1983). A new biometrical procedure for testing the equality of measurements from two different analytical methods - Application of linear-regression procedures for method comparison studies in clinical chemistry, Part I. *Journal of Clinical Chemistry and Clinical Biochemistry*, 21. **11.** 709-720.

Paul, W. and Steinwedel, H. (1953). Ein neues massenspektrometer ohne magnetfeld. *Zeitschrift für Naturforschung*, 8. 448–450.

Perry, R. H., Cooks, R. G. and Noll, R. J. (2008). Orbitrap mass spectrometry: Instrumentation, ion motion and applications. *Mass Spectrometry Reviews*, 27. **6.** 661-699.

Peters, F. T. (2011). Recent advances of liquid chromatography-(tandem) mass spectrometry in clinical and forensic toxicology. *Clinical Biochemistry*, 44. **1.** 54-65.

Petersen, P. H., Stockl, D., Blaabjerg, O., Pedersen, B., Birkemose, E., Thienpont, L., Lassen, J. F. and Kjeldsen, J. (1997). Graphical interpretation of analytical data from comparison of a field method with a Reference Method by use of difference plots. *Clinical Chemistry*, 43. **11.** 2039-2046.

Peterson, K. P., Pavlovich, J. G., Goldstein, D., Little, R., England, J. and Peterson, C. M. (1998). What is hemoglobin Alc? An analysis of glycated hemoglobins by electrospray ionization mass spectrometry. *Clinical Chemistry*, 44. **9.** 1951-1958.

Petricoin, E. F. and Liotta, L. A. (2003). Mass spectrometry-based diagnostics: The upcoming revolution in disease detection has already arrived - Response. *Clinical Chemistry*, 49. **7.** 1228-1229.

Piel, F. d. r. B., Patil, A. P., Howes, R. E., Nyangiri, O. A., Gething, P. W., Dewi, M., Temperley, W. H., Williams, T. N., Weatherall, D. J. and Hay, S. I. (2012). Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet*, **0.**

Prasain, J. K. (2012). Tandem Mass Spectrometry - Applications and Principles. InTech.

Pucci, P., Ferranti, P., Marino, G. and Malorni, A. (1989). Characterization of abnormal human hemoglobins by fast atom bombardment mass-spectrometry. *Biomedical and Environmental Mass Spectrometry*, 18. **1.** 20-26.

Rai, D., Alvelius, G. and Landin, B. (2002). Identification of Hb Kenya ($^{A}\gamma$81Leu-β86Ala) by electrospray mass spectrometry. *Hemoglobin*, 26. **1.** 71-75.

Rai, D. K., Green, B. N., Landin, B., Alvelius, G. and Griffiths, W. J. (2004). Accurate mass measurement and tandem mass spectrometry of intact globin chains identify the low proportion variant hemoglobin Lepore-Boston-Washington from the blood of a heterozygote. *Journal of Mass Spectrometry*, 39. **3.** 289-294.

Rai, D. K., Griffiths, W. J., Landin, B., Wild, B. J., Alvelius, G. and Green, B. N. (2003). Accurate mass measurement by electrospray ionization quadrupole mass spectrometry: detection of variants differing by <6 Da from normal in human hemoglobin heterozygotes. *Analytical Chemistry*, 75. **9.** 1978-1982.

Reid, G. E. and McLuckey, S. A. (2002). 'Top down' protein characterization via tandem mass spectrometry. *Journal of Mass Spectrometry*, 37. **7.** 663-675.

Riou, J., Godart, C., Hurtrel, D., Mathis, M., Bimet, C., BardakdjianMichau, J., Prehu, C., Wajcman, H. and Galacteros, F. (1997). Cation-exchange HPLC evaluated for presumptive identification of hemoglobin variants. *Clinical Chemistry*, 43. **1.** 34-39.

Roberts, N. B., Amara, A. B., Morris, M. and Green, B. N. (2001). Long-term evaluation of electrospray ionization mass spectrometric analysis of glycated hemoglobin. *Clinical Chemistry*, 47. **2.** 316-321.

Roberts, N. B., Green, B. N. and Morris, M. (1997). Potential of electrospray mass spectrometry for quantifying glycohemoglobin. *Clinical Chemistry*, 43. **5.** 771-778.

Roepstorff, P. and Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biological Mass Spectrometry*, 11. **11.** 601-601.

Ryan, K., Bain, B. J., Worthington, D., James, J., Plews, D., Mason, A., Roper, D., Rees, D. C., de la Salle, B. and Streetly, A. (2010). Significant haemoglobinopathies: guidelines for screening and diagnosis. *British Journal of Haematology*, 149. **1.** 35-49.

Scarff, C. A., Patel, V. J., Thalassinos, K. and Scrivens, J. H. (2009). Probing Hemoglobin Structure by Means of Traveling-Wave Ion Mobility Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, 20. **4.** 625-631.

Schaaff, T. G., Cargile, B. J., Stephenson, J. L. and McLuckey, S. A. (2000). Ion trap collisional activation of the (M+2H)(2+)-(M+17H)(17+) ions of human hemoglobin beta-chain. *Analytical Chemistry*, 72. **5.** 899-907.

Scherl, A., Graça, D. C., Martin, A. A., Hartmer, R., Kaspar, S., Meyer, M., Clerici, L., Samii, K., Tsybin, Y., Hochstrasser, D. and Lescuyer, P. (2013). Top-down analysis of Hemoglobin for clinical applications Mass Spectrometry: Applications to the Clinical Lab

Seidler, J., Zinn, N., Boehm, M. E. and Lehmann, W. D. (2010). De novo sequencing of peptides by MS/MS. *Proteomics*, 10. **4.** 634-649.

Shackleton, C. H. L., Falick, A. M., Green, B. N. and Witkowska, H. E. (1991). Electrospray mass spectrometry in the clinical diagnosis of variant hemoglobins. *Journal of Chromatography B: Biomedical Sciences and Applications*, 562. **1-2.** 175-190.

Shackleton, C. H. L. and Witkowska, H. E. (1996). Characterizing abnormal hemoglobin by MS. *Analytical Chemistry*, 68. **1.** A29-A33.

Shushan, B. (2010). A review of clinical diagnostic applications of liquid chromatography–tandem mass spectrometry. *Mass Spectrometry Reviews*, 29. **6.** 930-944.

Silberring, J. and Ciborowski, P. (2010). Biomarker discovery and clinical proteomics. *Trac-Trends in Analytical Chemistry*, 29. **2.** 128-140.

Simons, J. (2010). Mechanisms for S–S and N–Cα bond cleavage in peptide ECD and ETD mass spectrometry. *Chemical Physics Letters*, 484. **4–6.** 81-95.

Sleno, L. and Volmer, D. A. (2004). Ion activation methods for tandem mass spectrometry. *Journal of Mass Spectrometry*, 39. **10.** 1091-1112.

Soares, R., Pires, E., Almeida, A. M., Santos, R., Gomes, R., Koči, K., Ferraz Franco, C. and Coelho, A. V. (2012). Tandem Mass Spectrometry of Peptides. Tandem Mass Spectrometry - Applications and Principles. Prasain, J. K. InTech.

Sobott, F., Watt, S., Smith, J., Edelmann, M., Kramer, H. and Kessler, B. (2009). Comparison of CID versus ETD based MS/MS fragmentation for the analysis of protein ubiquitination. *Journal of The American Society for Mass Spectrometry*, 20. **9.** 1652-1659.

Somogyi, A. (2008). Mass spectrometry instrumentation and techniques. Medical Applications of Mass Spectrometry. Vekey, K., *et al.* Elsevier Science.

Stafford, G. C., Kelley, P. E., Syka, J. E. P., Reynolds, W. E. and Todd, J. F. J. (1984). Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Processes*, 60. **SEP.** 85-98.

Steinberg, M. H. (2011). Sickle cell disease and associated hemoglobinopathies. Cecil Medicine. Goldman, L.Schafer, A. Saunders Elsevier. Chapter 166

Stensballe, A., Jensen, O. N., Olsen, J. V., Haselmann, K. F. and Zubarev, R. A. (2000). Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Communications in Mass Spectrometry*, 14. **19.** 1793-1800.

Stephens, A. D., Angastiniotis, M., Baysal, E., Chan, V., Davis, B., Fucharoen, S., Giordano, P. C., Hoyer, J. D., Mosca, A. and Wild, B. (2011a). ICSH recommendations for the measurement of haemoglobin F. *International Journal of Laboratory Hematology*, 34. **1.** 14-20.

Stephens, A. D., Angastiniotis, M., Baysal, E., Chan, V., Fucharoen, S., Giordano, P. C., Hoyer, J. D., Mosca, A. and Wild, B. (2011b). ICSH recommendations for the measurement of haemoglobin A2. *International Journal of Laboratory Hematology*, 34. **1.** 1-13.

Streetly, A., Latinovic, R., Hall, K. and Henthorn, J. (2009). Implementation of universal newborn bloodspot screening for sickle cell disease and other clinically significant haemoglobinopathies in England: screening results for 2005-7. *Journal of Clinical Pathology*, 62. **1.** 26-30.

Summerfield, S. G., Buzy, A., Jennings, K. R. and Green, B. (1996). Electrospray tandem mass spectrometry of human haemoglobin and a number of beta-chain aberrants. *European Mass Spectrometry*, 2. **4-5.** 305-322.

Swartz, M. E. (2005a). Ultra performance liquid chromatography (UPLC): An introduction. *Lc Gc North America*, 8-14.

Swartz, M. E. (2005b). UPLC (TM): An introduction and review. *Journal of Liquid Chromatography & Related Technologies*, 28. **7-8.** 1253-1263.

Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J. and Hunt, D. F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 101. **26.** 9528-9533.

Tan, H. T., Lee, Y. H. and Chung, M. C. M. (2012). Cancer proteomics. *Mass Spectrometry Reviews*, 31. **5.** 583-605.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. and Matsuo, T. (1988). Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2. **8.** 151-153.

Taylor, S. M., Parobek, C. M. and Fairhurst, R. M. (2012). Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. *Lancet Infectious Diseases*, 12. **6.** 457-468.

Théberge, R., Infusini, G., Tong, W., McComb, M. E. and Costello, C. E. (2011a). Top-down analysis of small plasma proteins using an LTQ-Orbitrap. Potential for mass spectrometry-based clinical assays for transthyretin and hemoglobin. *International Journal of Mass Spectrometry*, 300. **2–3.** 130-142.

Théberge, R., Leymarie, N., McComb, M. E. and Costello, C. E. (2011b). Hemoglobin variant analysis using an LTQ-orbitrap top-down platform with ETD, HCD and LTQ-CID. Proc. 59th ASMS Conference on Mass Spectrometry and Allied topics. Denver, CO.

Thein, S. L. (2011). Abnormalities of the structure and synthesis of hemoglobin. Blood and Bone Marrow Pathology (Second Edition). Churchill Livingstone. CHAPTER 9

Thomson, B. A. (1998). Atmospheric pressure ionization and liquid chromatography mass spectrometry - Together at last. *Journal of the American Society for Mass Spectrometry*, 9. **3.** 187-193.

Thomson, J. J. (1913). Rays of positive electricity and their application to chemical analyses. Longmans, Green and Co.

Tolley, L., Jorgensen, J. W. and Moseley, M. A. (2001). Very high pressure gradient LC/MS/MS. *Analytical Chemistry*, 73. **13.** 2985-2991.

Troxler, H., Kleinert, P., Schmugge, M. and Speer, O. (2012). Advances in hemoglobinopathy detection and identification. Advances in Clinical Chemistry, Vol 57. Elsevier Academic Press Inc.

Turecek, F. and McLafferty, F. W. (1984). Non-ergodic behavior in acetone enol ion dissociations. *Journal of the American Chemical Society*, 106. **9.** 2525-2528.

Van Berkel, G. J. (2003). An overview of some recent developments in ionization methods for mass spectrometry. *European Journal of Mass Spectrometry*, 9. **6.** 539-562.

VanDeemter, J. J., Zuiderweg, F. J. and Klinkenberg, A. (1995). Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography (Reprinted from Chem Engng Sci, vol 5, pg 271-289, 1956). *Chemical Engineering Science*, 50. **24.** 3869-3882.

Vekey, K., Telekes, A. and Vertes, A. (2007). Medical applications of mass spectrometry. Elsevier Science.

Vestal, M. L., Juhasz, P. and Martin, S. A. (1995). Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 9. **11.** 1044-1050.

Vichinsky, E. P. (2005). Changing patterns of thalassemia worldwide. Cooley's Anemia Eighth Symposium. New York Acad Sciences.

Vogeser, M. and Kirchhoff, F. (2011). Progress in automation of LC-MS in laboratory medicine. *Clinical Biochemistry*, 44. **1.** 4-13.

Vogeser, M. and Seger, C. (2008). A decade of HPLC-MS/MS in the routine clinical laboratory - Goals for further developments. *Clinical Biochemistry*, 41. **9.** 649-662.

Wada, Y., Hayashi, A., Fujita, T., Matsuo, T., Katakuse, I. and Matsuda, H. (1981). Structural analysis of human hemoglobin variants with field desorption mass spectrometry. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 667. **2.** 233-241.

Wada, Y., Hayashi, A., Fujita, T., Matsuo, T., Katakuse, I. and Matsuda, H. (1983a). Structural analysis of human hemoglobin variants by mass spectrometry. *International Journal or Mass Spectrometry and Ion Physics*, 48. 209-212.

Wada, Y., Hayashi, A., Masanori, F., Katakuse, I., Ichihara, T., Nakabushi, H., Matsuo, T., Sakurai, T. and Matsuda, H. (1983b). Characterization of a new fetal hemoglobin variant, Hb F Izumi , by molecular secondary ion mass spectrometry. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 749. **3.** 244-248.

Wajcman, H. and Riou, J. (2009). Globin chain analysis: An important tool in phenotype study of hemoglobin disorders. *Clinical Biochemistry*, 42. **18.** 1802-1806.

Whitehouse, C. M., Dreyer, R. N., Yamashita, M. and Fenn, J. B. (1985). Electrospray interface for liquid chromatographs and mass spectrometers. *Analytical Chemistry*, 57. **3.** 675-679.

Wild, B. and Bain, B. J. (2006). Investigation of abnormal haemoglobins and thalassaemia. Dacie and Lewis Practical Haematology (Tenth Edition). Churchill Livingstone. Chapter 12

Wild, B. J., Green, B. N., Cooper, E. K., Lalloz, M. R. A., Erten, S., Stephens, A. D. and Layton, D. M. (2001). Rapid identification of hemoglobin variants by electrospray ionization mass spectrometry. *Blood Cells, Molecules, and Diseases*, 27. **3.** 691-704.

Wild, B. J., Green, B. N. and Stephens, A. D. (2004). The potential of electrospray ionization mass spectrometry for the diagnosis of hemoglobin variants found in newborn screening. *Blood Cells, Molecules, and Diseases*, 33. **3.** 308-317.

Wiley, W. C. and McLaren, I. H. (1955). Time-of-flight mass spectrometer with improved resolution. *Review of Scientific Instruments*, 26. **12.** 1150-1157.

Willekens, E., Thienpont, L. M., Stockl, D., Kobold, U., Hoelzel, W. and De Leenheer, A. P. (2000). Quantification of glycohemoglobin in blood by mass spectrometry applying multiple-reaction monitoring. *Clinical Chemistry*, 46. **2.** 281-283.

Williams, J., Creese, A., Roper, D., Green, B. and Cooper, H. (2009). Hot electron capture dissociation distinguishes leucine from isoleucine in a novel hemoglobin variant, Hb Askew, β54(D5)Val→Ile. *Journal of The American Society for Mass Spectrometry*, 20. **9.** 1707-1713.

Williams, J. P., Kevin, G., Brian, N. G., James, H. S. and Robert, H. B. (2008). Ion mobility augments the utility of mass spectrometry in the identification of human hemoglobin variants. *Rapid Communications in Mass Spectrometry*, 22. **20.** 3179-3186.

Williams, T. N. and Weatherall, D. J. (2012). World distribution, population genetics, and health burden of the hemoglobinopathies. *Cold Spring Harbor Perspectives in Biology*, 4. **9.** 14.

Witkowska, H. E., Green, B. N., Morris, M. and Shackleton, C. H. (1995). Intact protein electrospray ionization tandem mass spectrometry can be the sole technique used for confirming the structure of a variant hemoglobin. *Rapid communications in mass spectrometry : RCM*, Spec No. S111-115.

Wilm, M. and Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. *Analytical Chemistry*, 68. **1.** 1-8.

Wilm, M. S. and Mann, M. (1994). Electrospray and Taylor-cone theory, Doles beam of macromolecules at last *International Journal of Mass Spectrometry*, 136. **2-3.** 167-180.

www.diabeticretinopathy.org.uk: What is the HbA1C?, http://medweb.bham.ac.uk/easdec/prevention/what_is_the_hba1c.htm. (Accessed 22 April 2013)

Yamashita, M. and Fenn, J. B. (1984). Electrospray Ion-source - Another variation of the Free-Jet theme *Journal of Physical Chemistry*, 88. **20.** 4451-4459.

Yang, Z. and Wang, S. H. (2008). Recent development in application of high performance liquid chromatography-tandem mass spectrometry in therapeutic drug monitoring of immunosuppressants. *Journal of Immunological Methods*, 336. **2.** 98-103.

Zanella-Cleon, I., Joly, P., Becchi, M. and Francina, A. (2009). Phenotype determination of hemoglobinopathies by mass spectrometry. *Clinical Biochemistry*, 42. 1807-1817.

Zhang, L.-K., Rempel, D., Pramanik, B. N. and Gross, M. L. (2005). Accurate mass measurements by Fourier transform mass spectrometry. *Mass Spectrometry Reviews*, 24. **2.** 286-309.

Zhang, X., Wei, D., Yap, Y., Li, L., Guo, S. and Chen, F. (2007). Mass spectrometry-based "omics" technologies in cancer diagnostics. *Mass Spectrometry Reviews*, 26. **3.** 403-431.

Zhou, H., Ning, Z. B., Starr, A. E., Abu-Farha, M. and Figeys, D. (2012). Advancements in Top-Down Proteomics. *Analytical Chemistry*, 84. **2.** 720-734.

Zhu, P., Bowden, P., Zhang, D. and Marshall, J. G. (2010). Mass spectrometry of peptides and proteins from human blood. *Mass Spectrometry Reviews*, 30. **5.** 685-732.

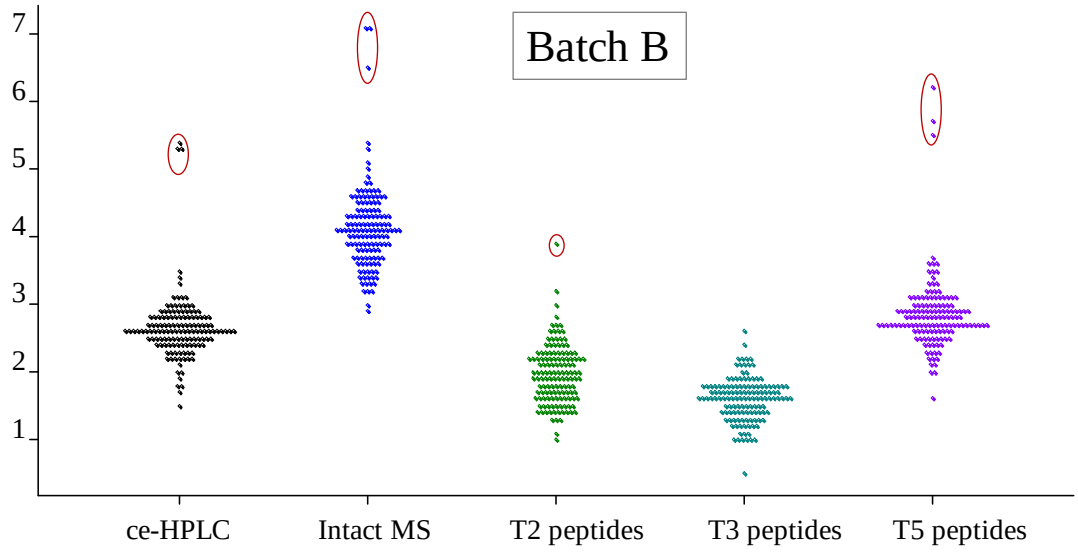
Zubarev, R. A., Kelleher, N. L. and McLafferty, F. W. (1998). Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society*, 120. **13.** 3265-3266.

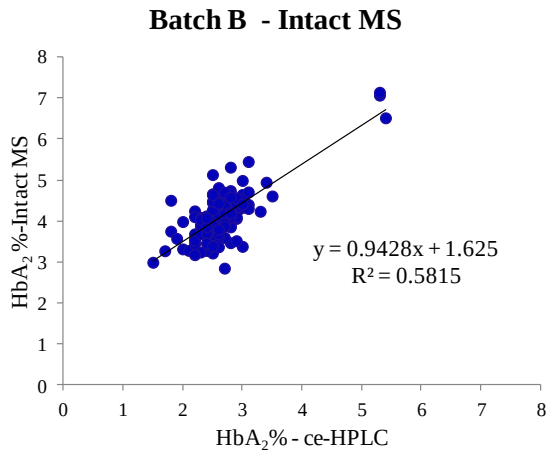
Zubarev, R. A., Zubarev, A. R. and Savitski, M. M. (2008). Electron capture/transfer versus collisionally activated/induced dissociations: Solo or duet? *Journal of the American Society for Mass Spectrometry*, 19. **6.** 753-761.

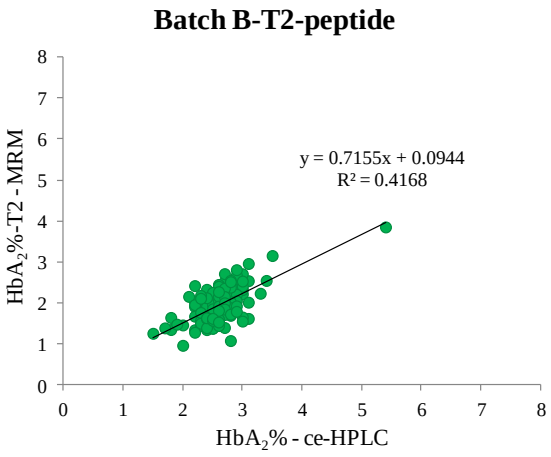
APPENDICES

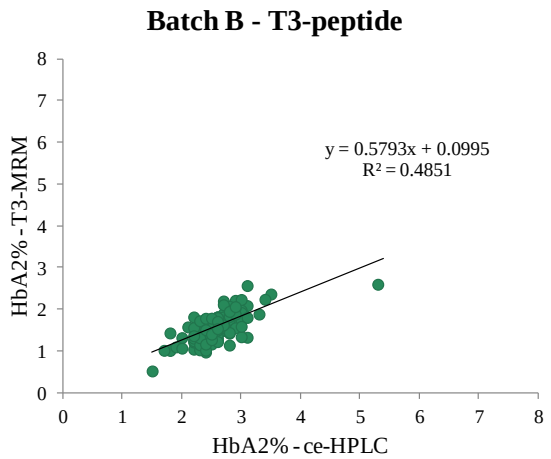
Appendix 1

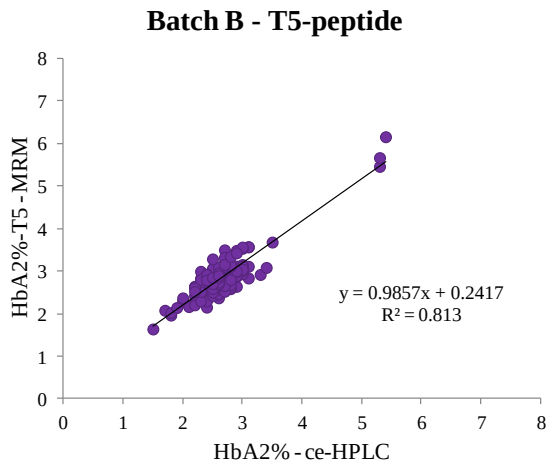
Figures for Batch B

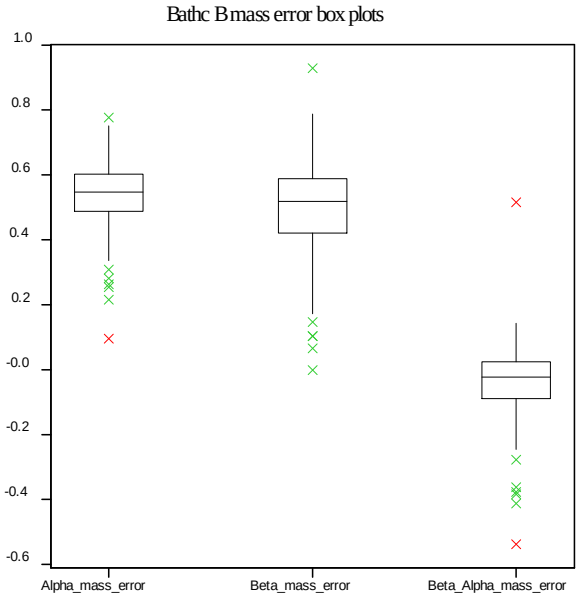


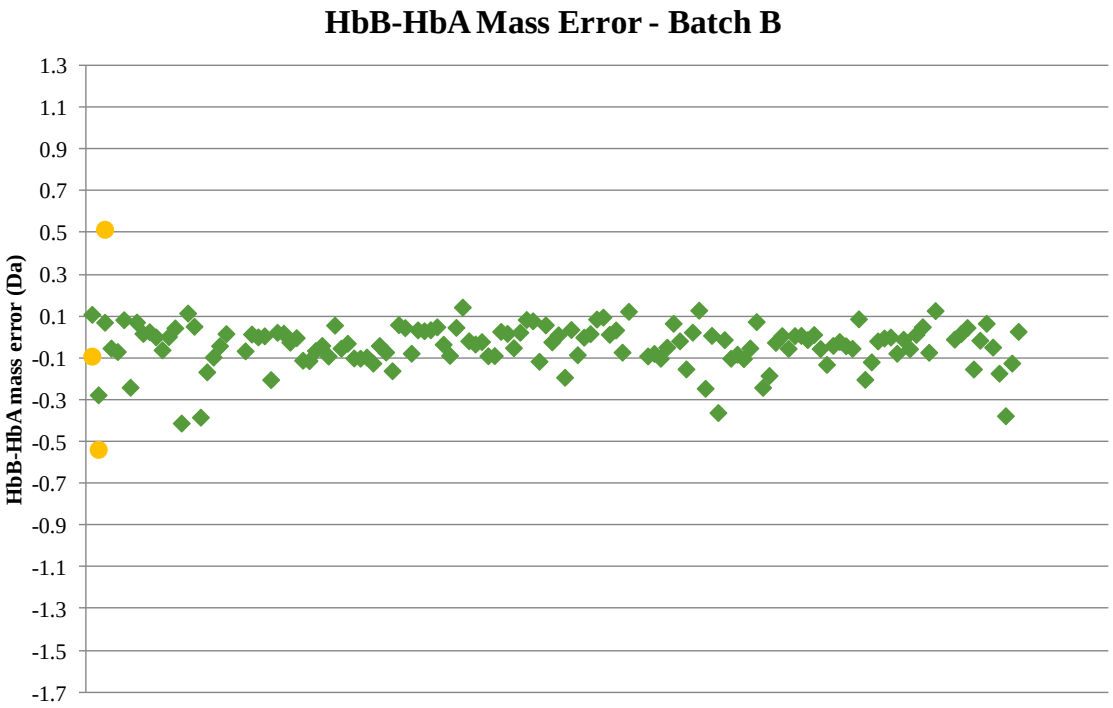


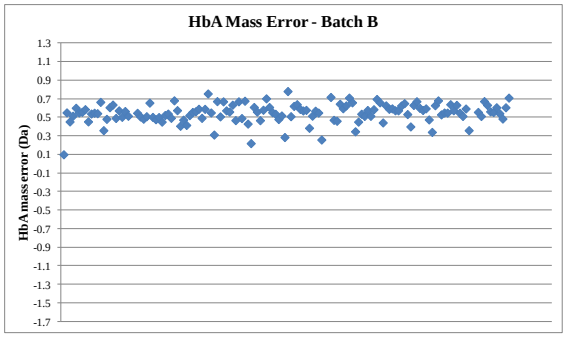


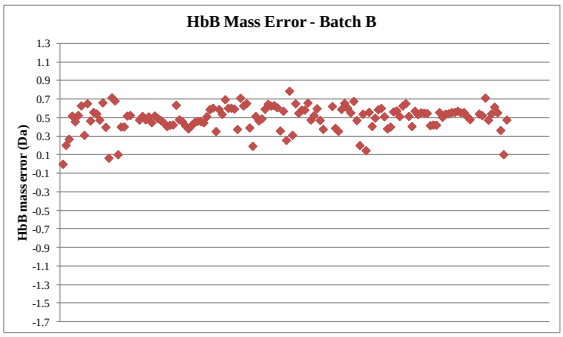


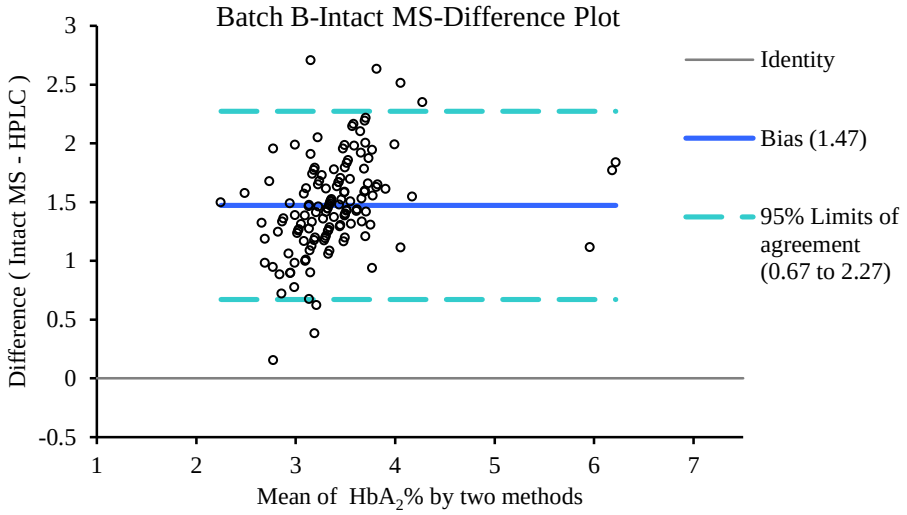


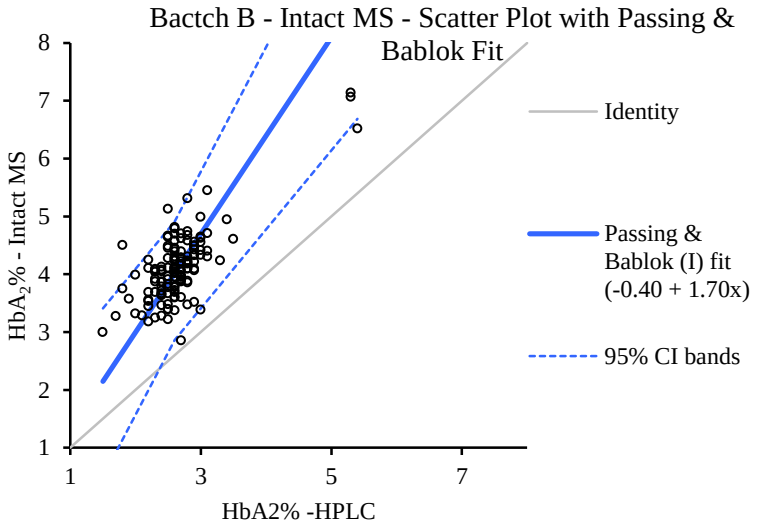

Bathc B mass error box plots
1.0
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
Alpha_mass_error
Beta_mass_error
Beta_Alpha_mass_error

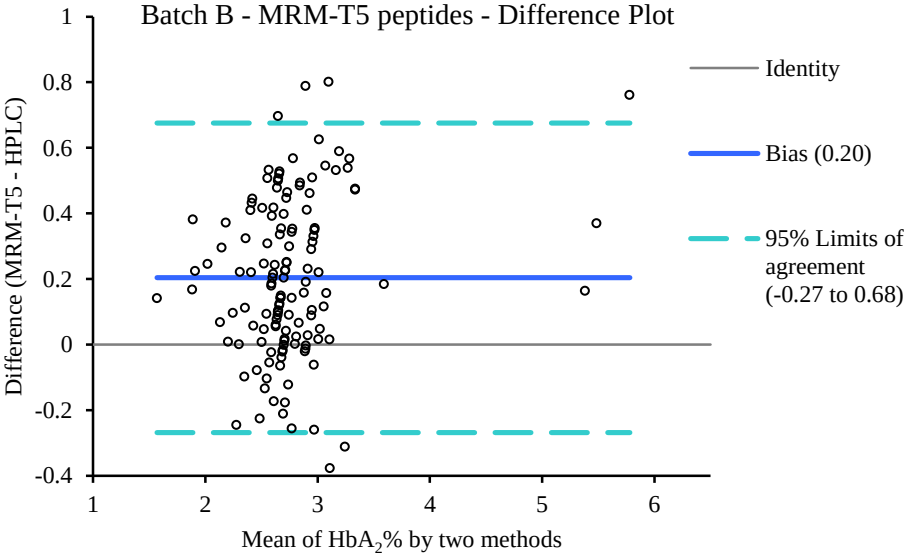

HbB-HbA Mass Error - Batch B
HbB-HbA mass error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

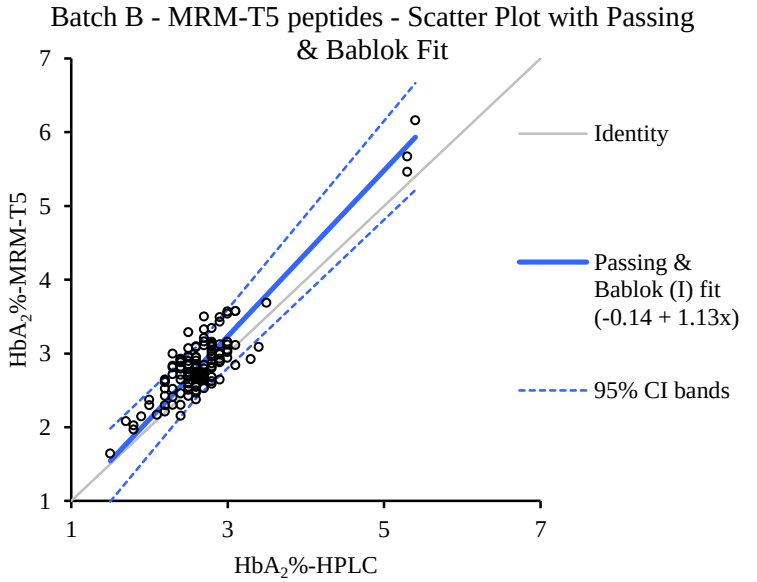

HbA Mass Error - Batch B
HbA mass error (Da)


HbB Mass Error - Batch B
HbB mass error (Da)

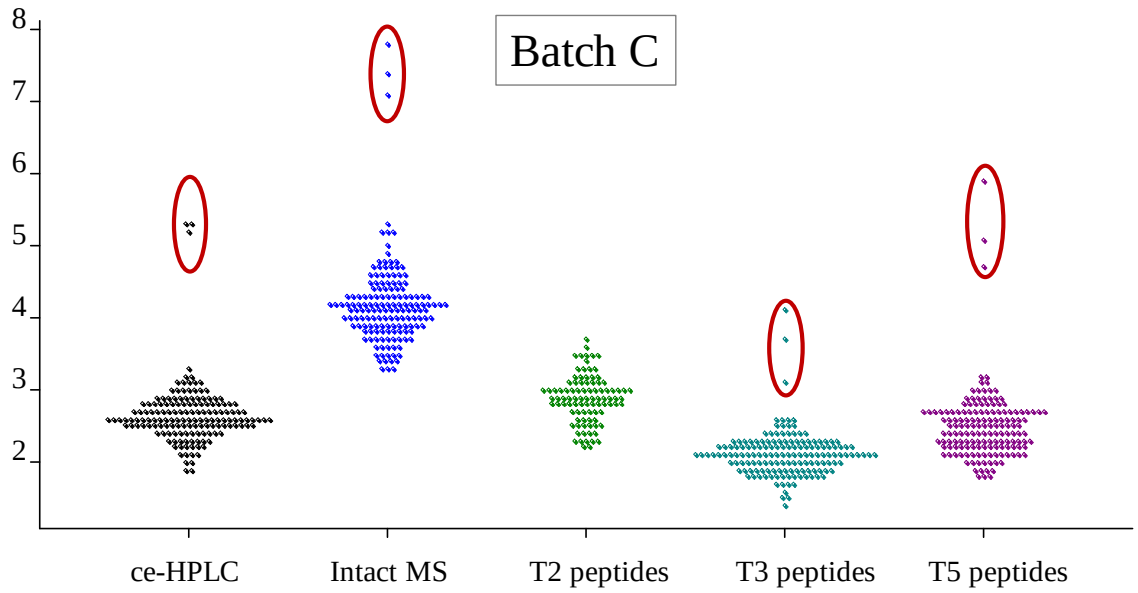

Batch B-Intact MS-Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.47)
95% Limits of agreement (0.67 to 2.27)

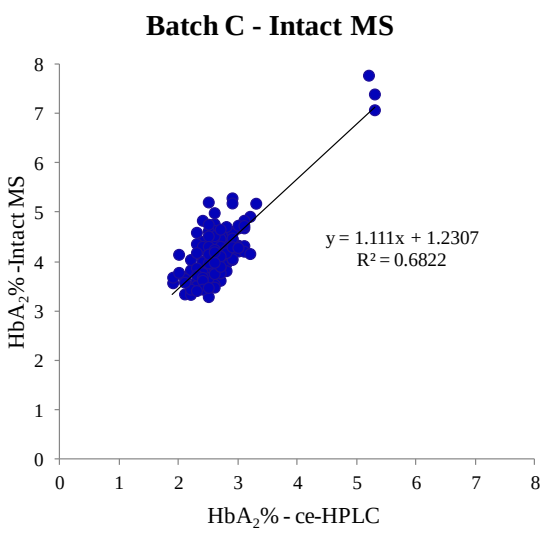

Bactch B - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% - Intact MS
HbA2% -HPLC
Identity
Passing & Bablok (I) fit (-0.40 + 1.70x)
95% CI bands

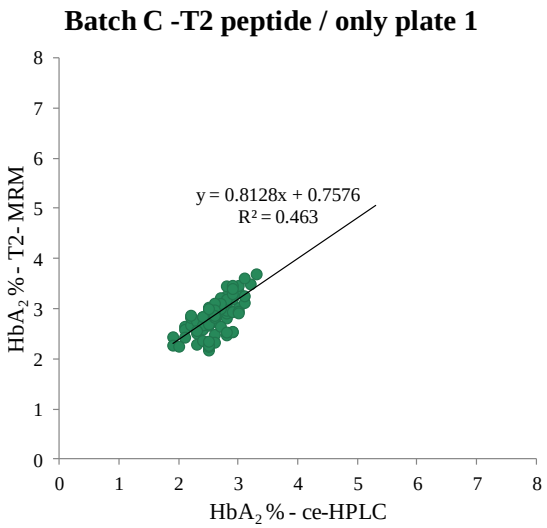

Batch B - MRM-T5 peptides - Difference Plot
Difference (MRM-T5 - HPLC)
Mean of HbA2% by two methods
Identity
Bias (0.20)
95% Limits of agreement (-0.27 to 0.68)

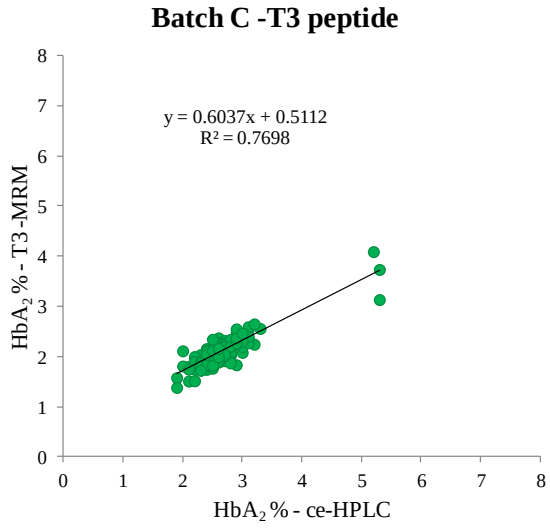

Batch B - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2%-MRM-T5
HbA2%-HPLC
Identity
Passing & Bablok (I) fit (-0.14 + 1.13x)
95% CI bands

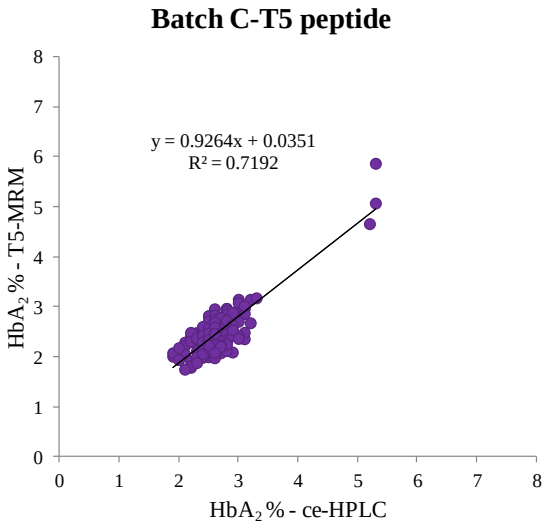
Figures for Batch C

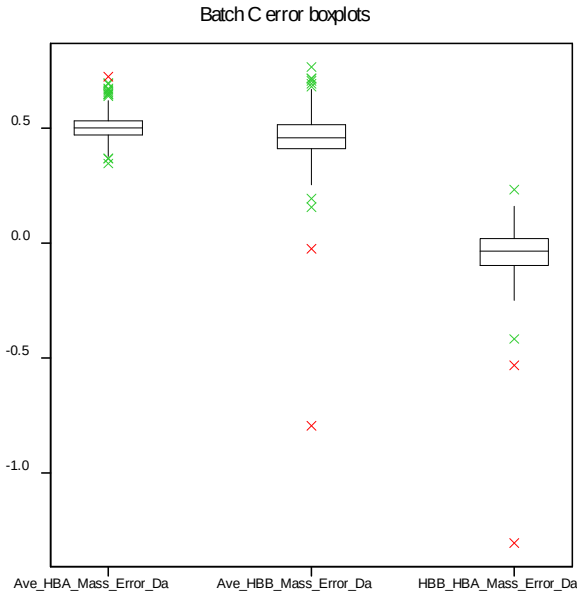


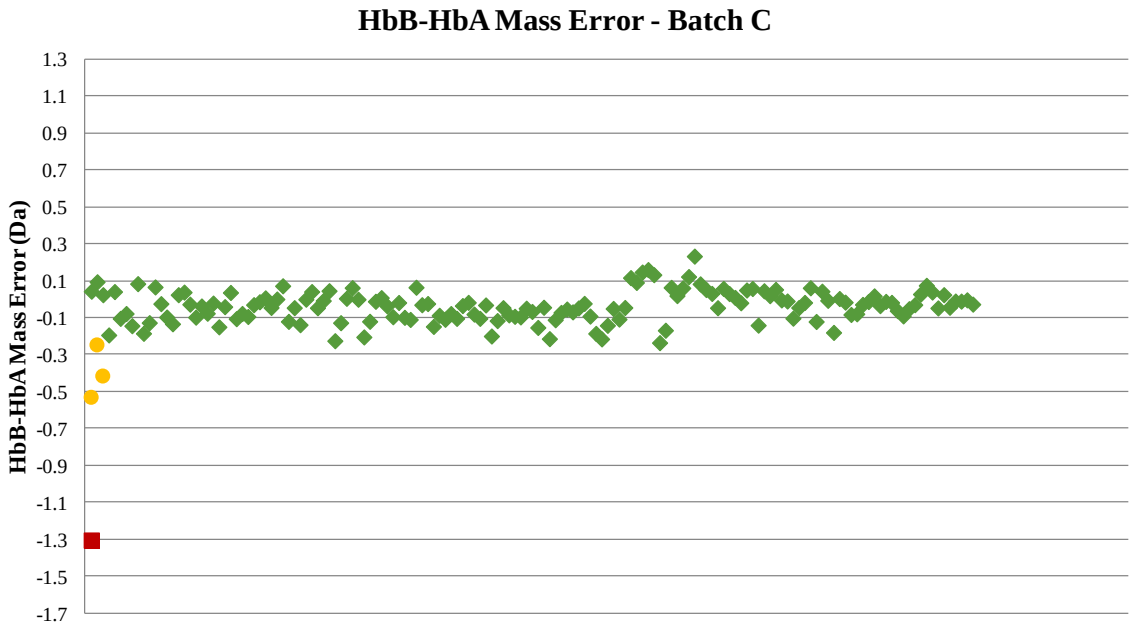


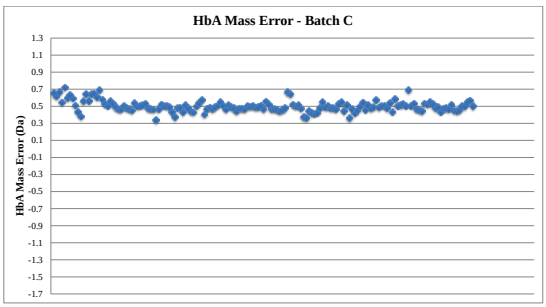


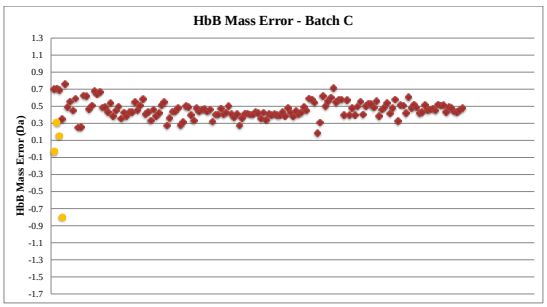


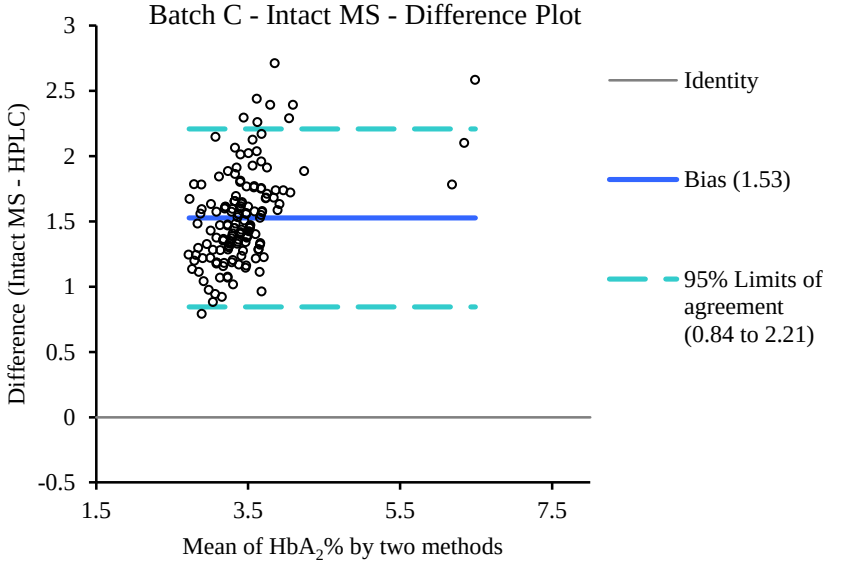


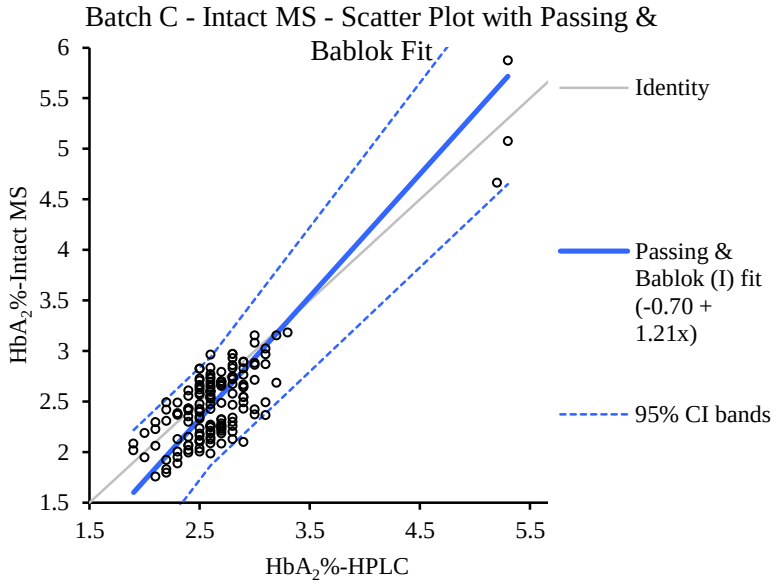

Batch C error boxplots
0.5
0.0
-0.5
-1.0
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

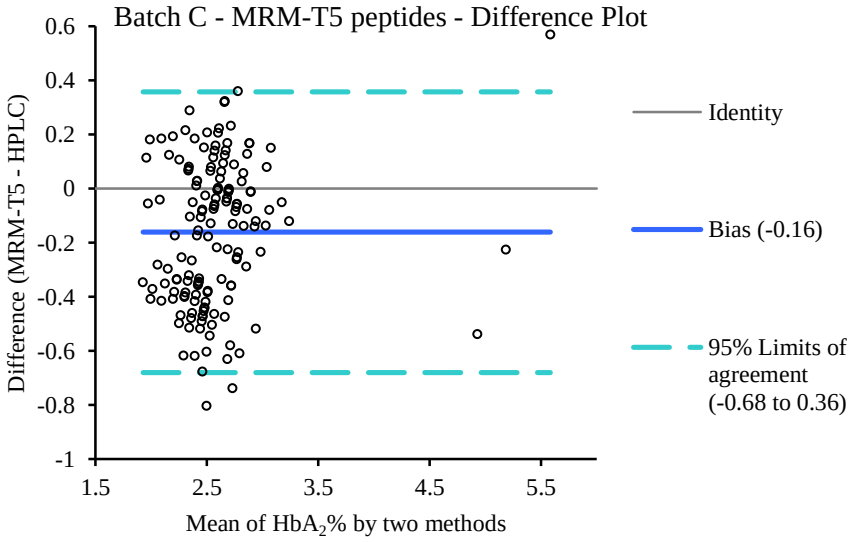

HbB-HbA Mass Error - Batch C
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

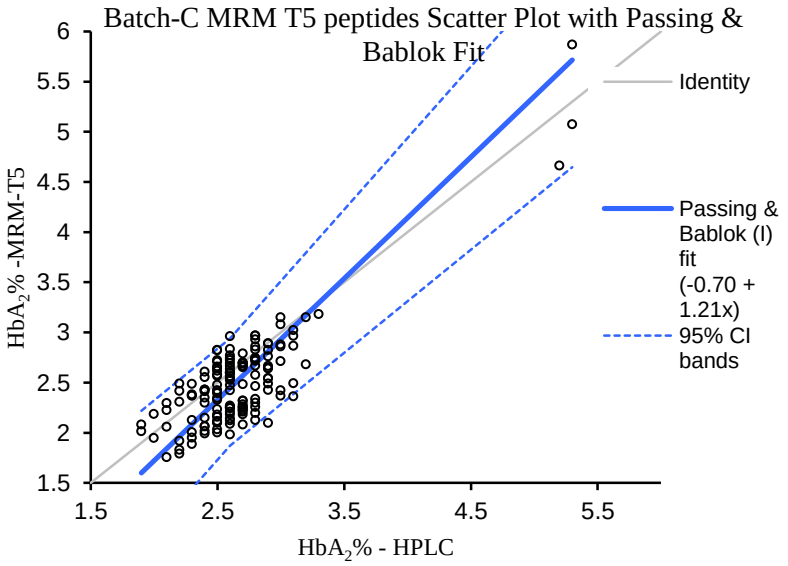

HbA Mass Error - Batch C
HbA Mass Error (Da)


HbB Mass Error - Batch C
HbB Mass Error (Da)

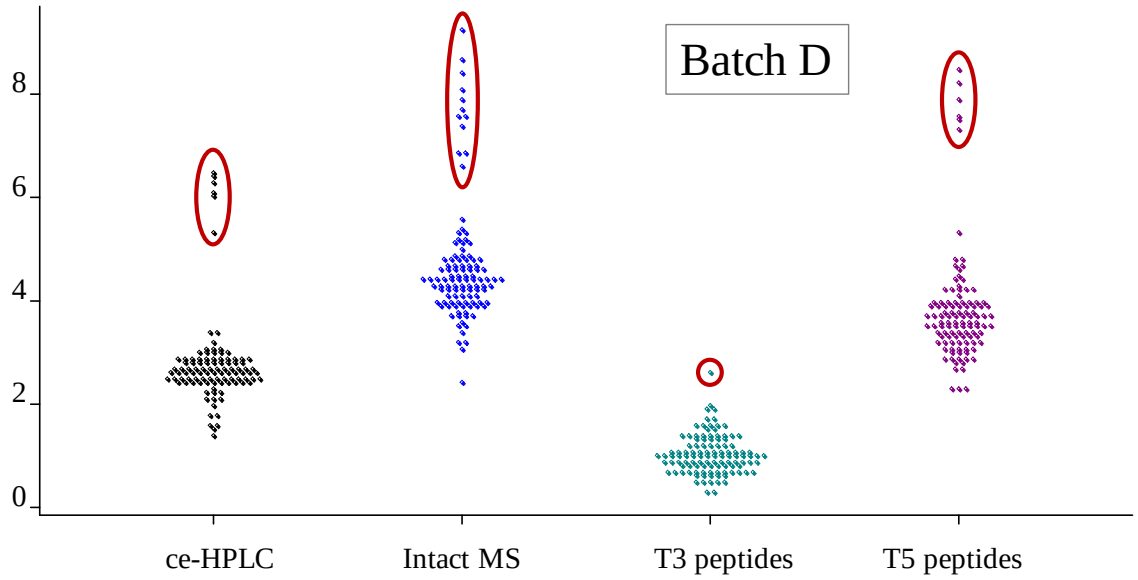

Batch C - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.53)
95% Limits of agreement (0.84 to 2.21)

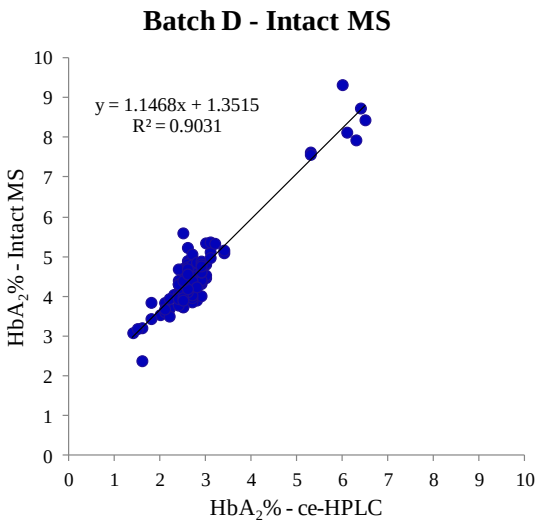

Batch C - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2%-Intact MS
HbA2%-HPLC
Identity
Passing & Bablok (I) fit (-0.70 + 1.21x)
95% CI bands

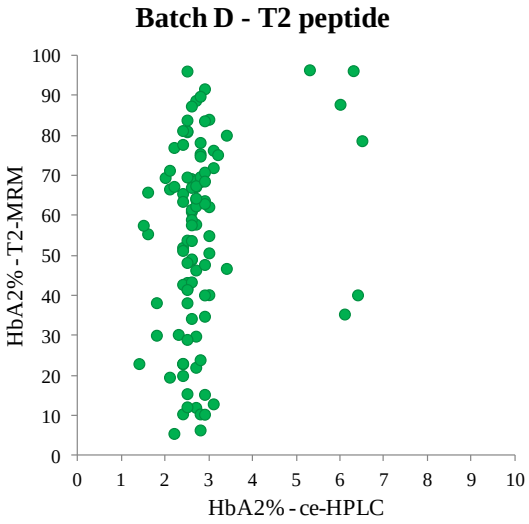

Batch C - MRM-T5 peptides - Difference Plot
Difference (MRM-T5 - HPLC)
Mean of HbA2% by two methods
Identity
Bias (-0.16)
95% Limits of agreement (-0.68 to 0.36)

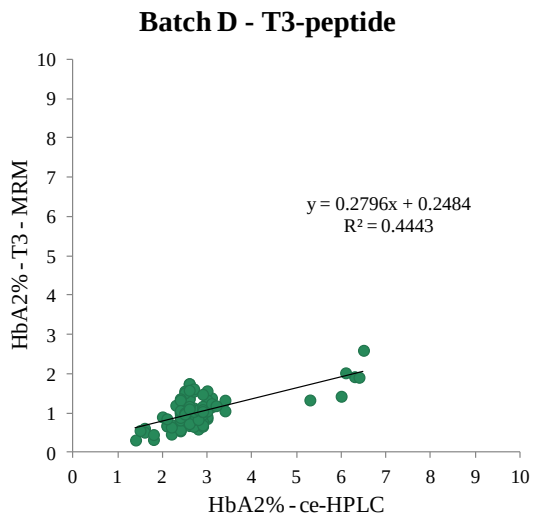

Batch-C MRM T5 peptides Scatter Plot with Passing & Bablok Fit
HbA2% -MRM-T5
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (-0.70 + 1.21x)
95% CI bands

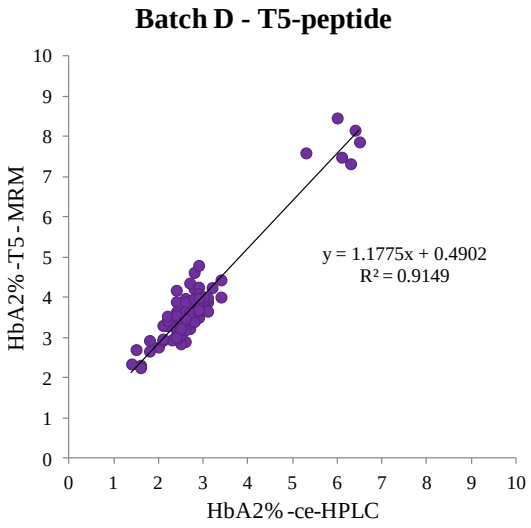
Figures for Batch D

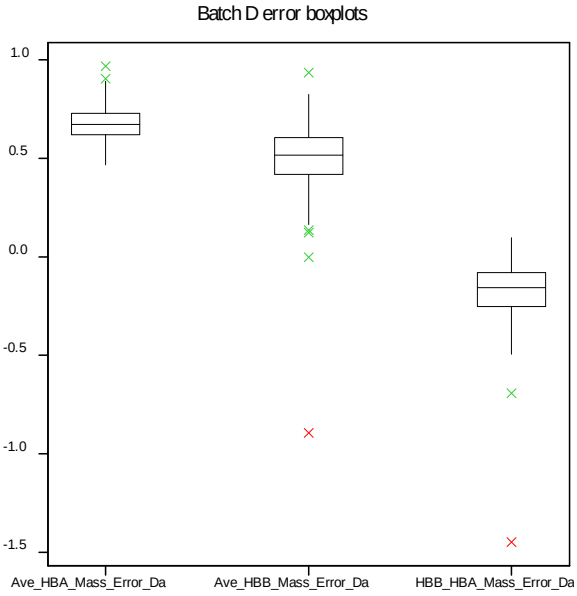


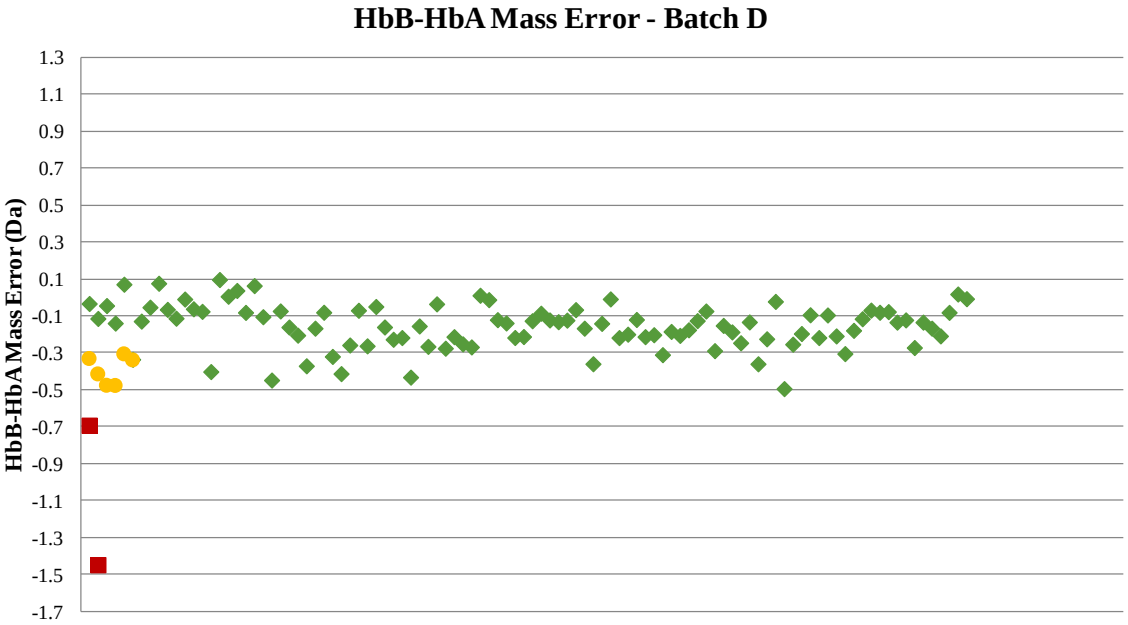


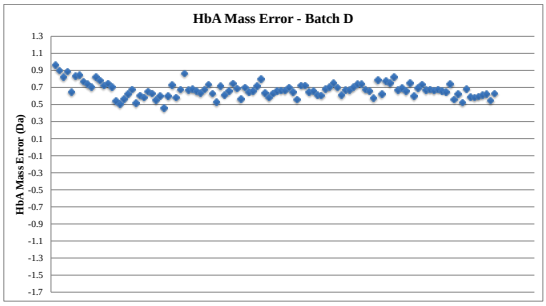


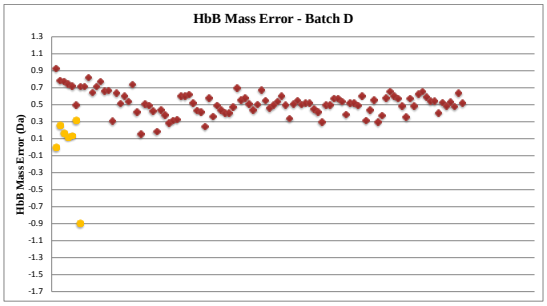


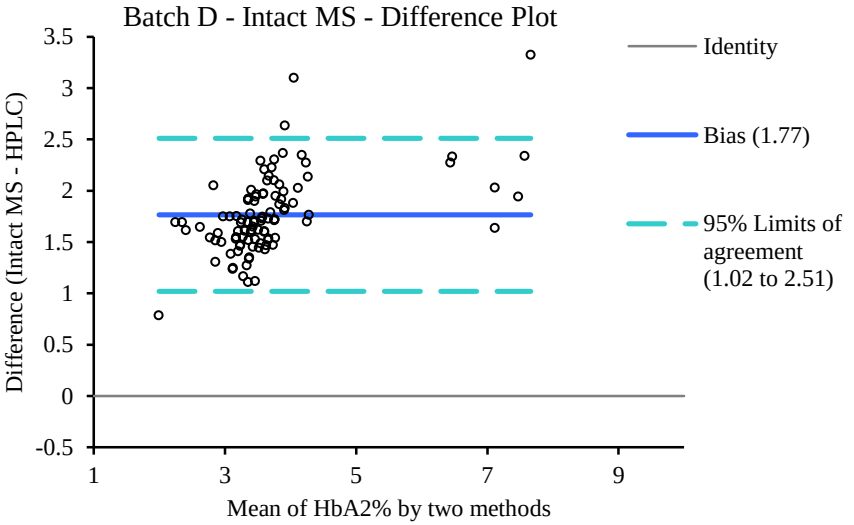


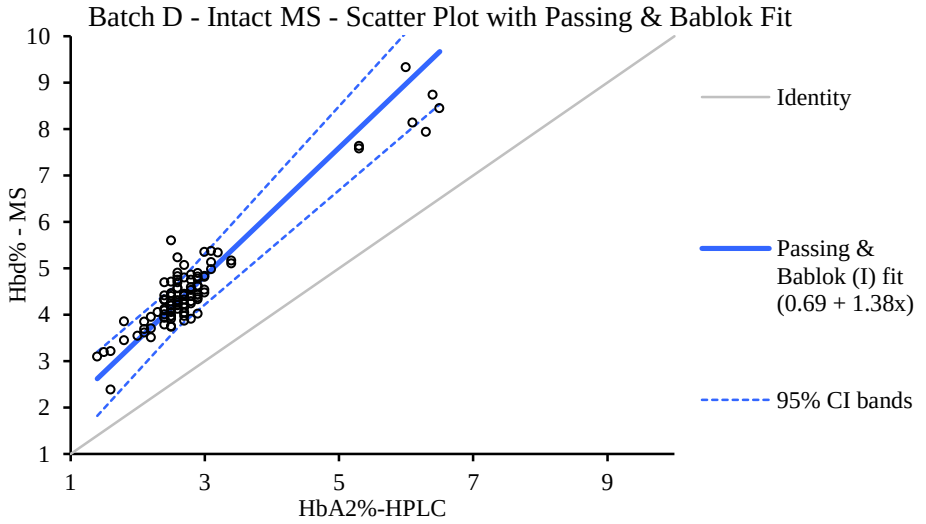

Batch D error boxplots
1.0
0.5
0.0
-0.5
-1.0
-1.5
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

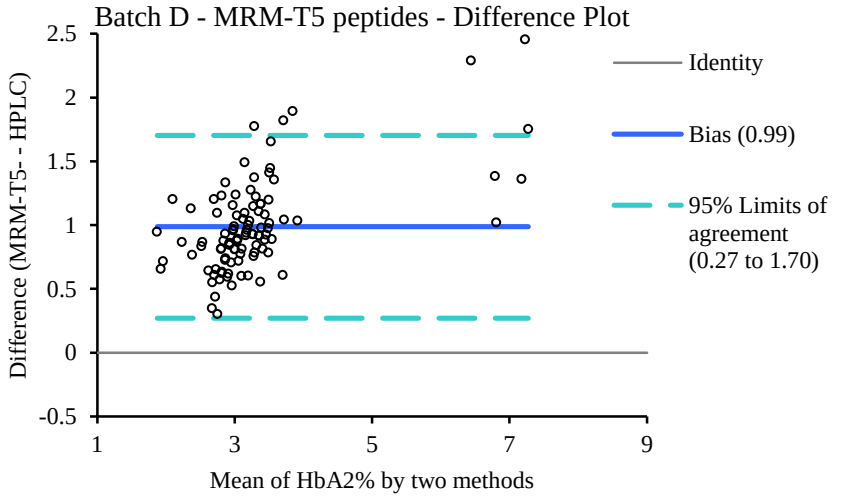

HbB-HbA Mass Error - Batch D
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

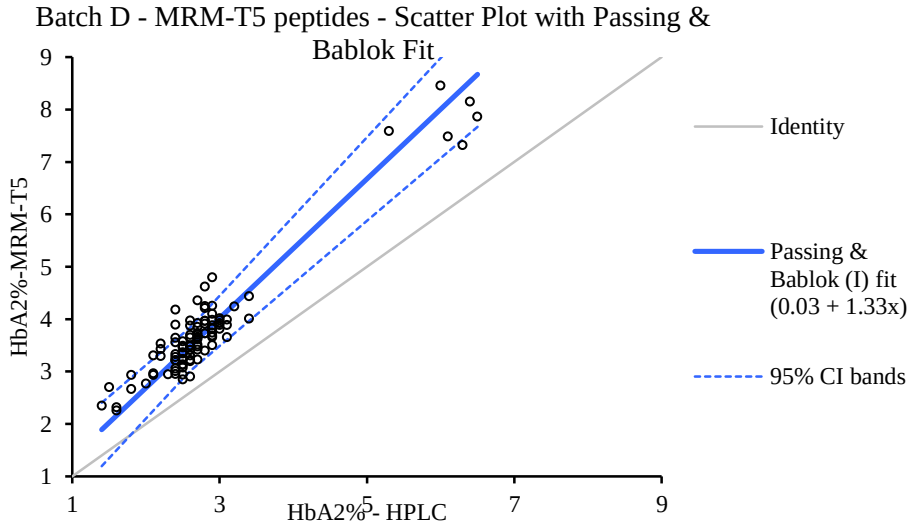

HbA Mass Error - Batch D
HbA Mass Error (Da)


HbB Mass Error - Batch D
HbB Mass Error (Da)

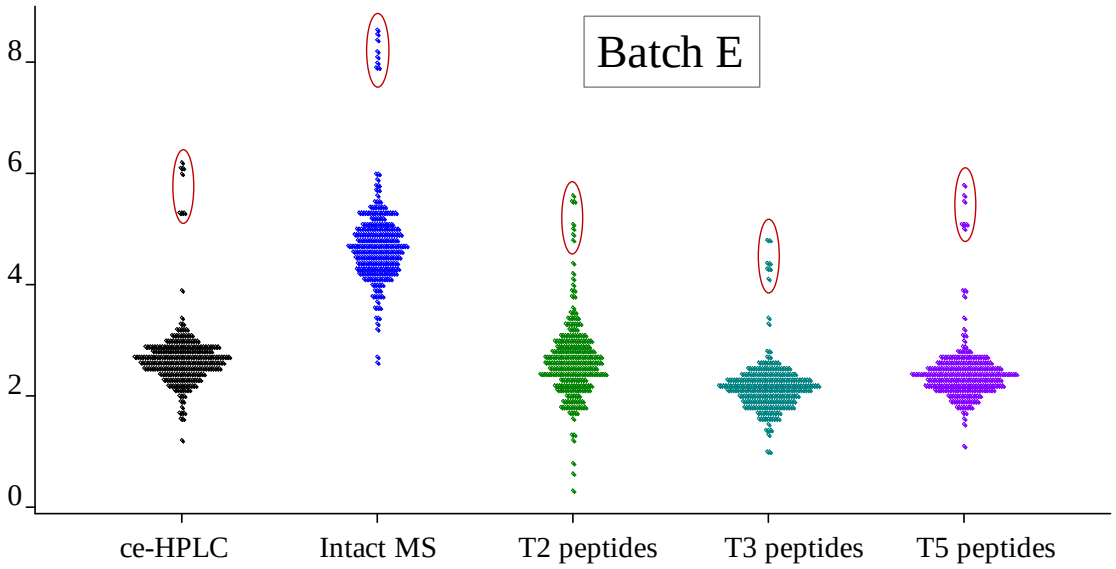

Batch D - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.77)
95% Limits of agreement (1.02 to 2.51)

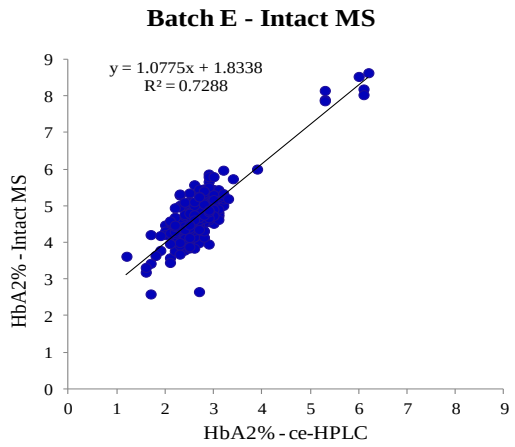

Batch D - Intact MS - Scatter Plot with Passing & Bablok Fit
Hbd% - MS
HbA2%-HPLC
Identity
Passing & Bablok (I) fit (0.69 + 1.38x)
95% CI bands

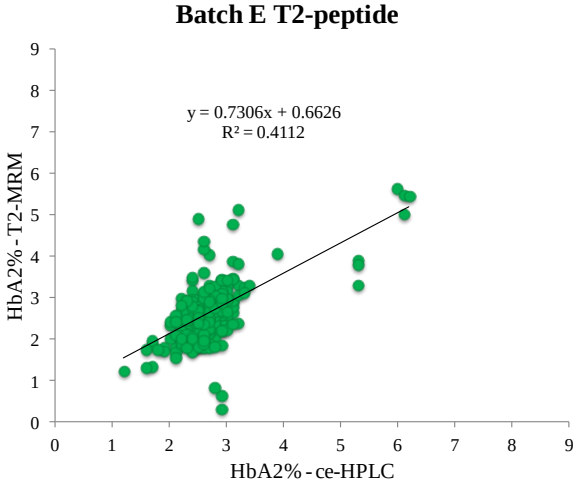

Batch D - MRM-T5 peptides - Difference Plot
Difference (MRM-T5- - HPLC)
Mean of HbA2% by two methods
Identity
Bias (0.99)
95% Limits of agreement (0.27 to 1.70)

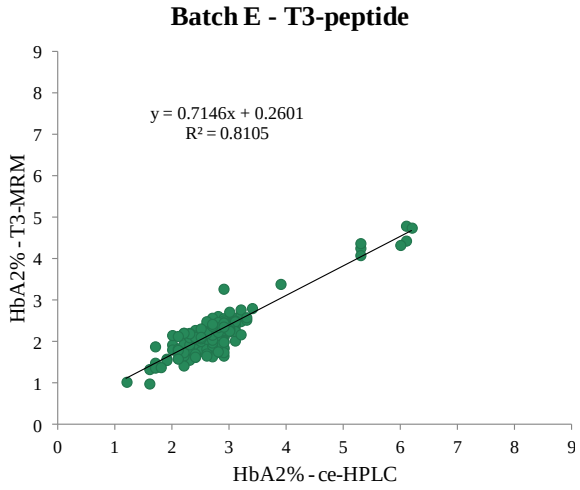

Batch D - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2%-MRM-T5
HbA2%- HPLC
Identity
Passing & Bablok (I) fit (0.03 + 1.33x)
95% CI bands

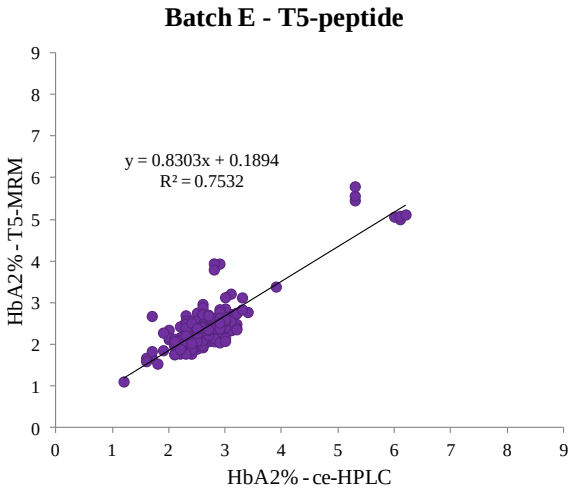
Figures for Batch E

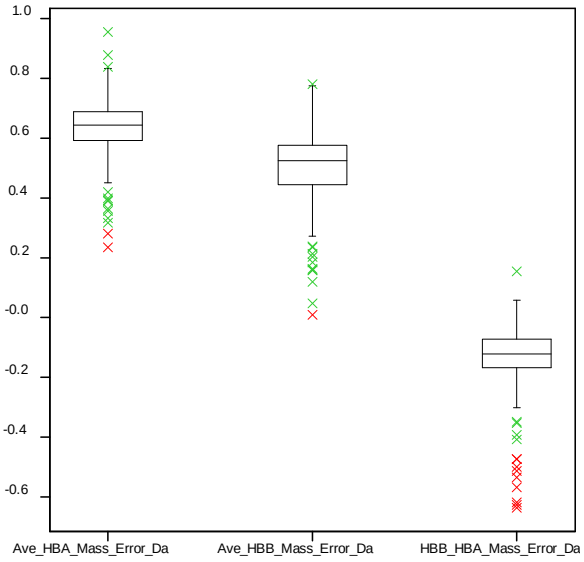




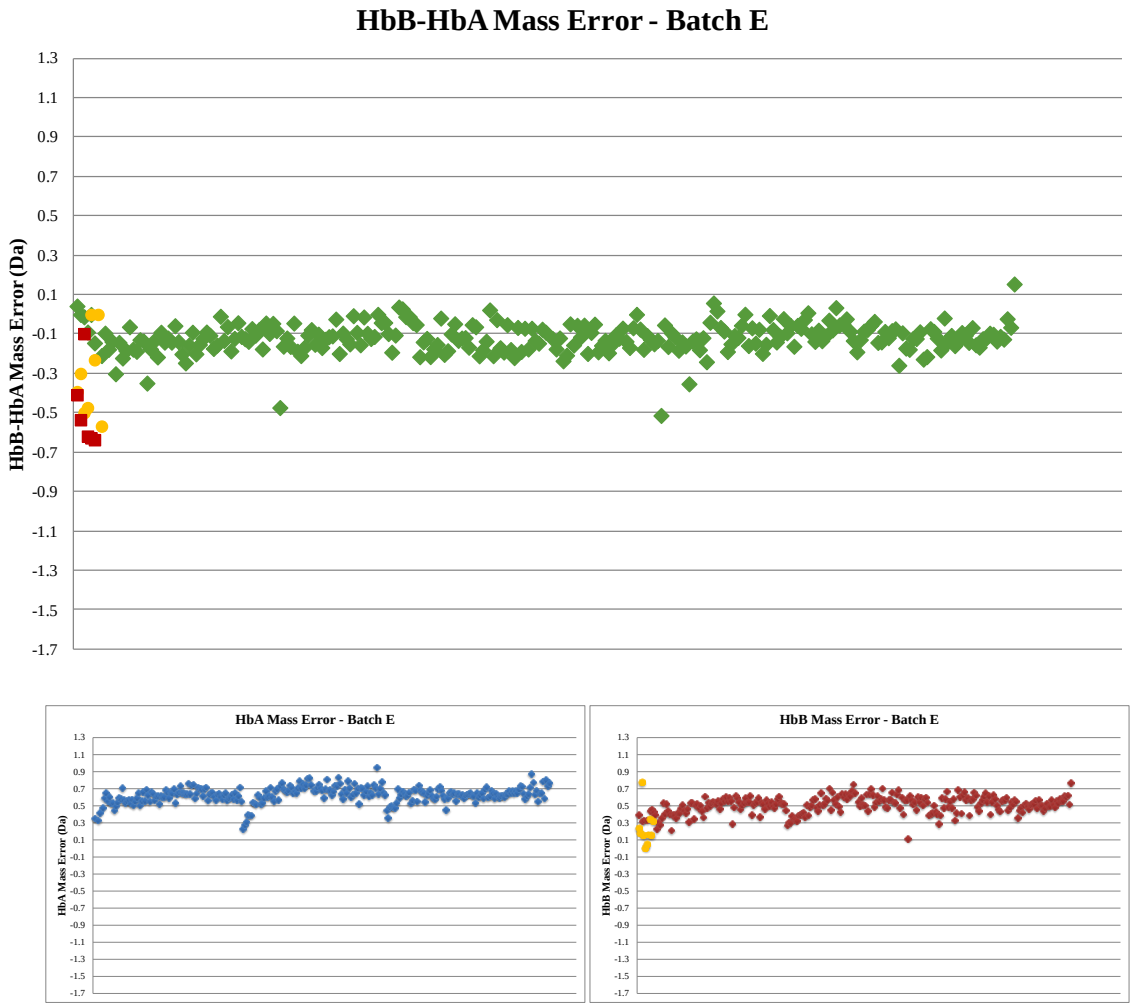


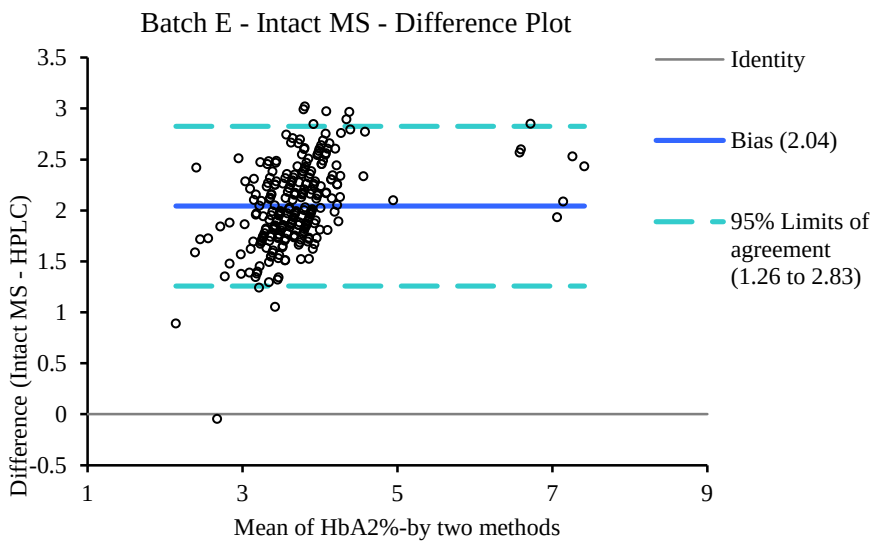


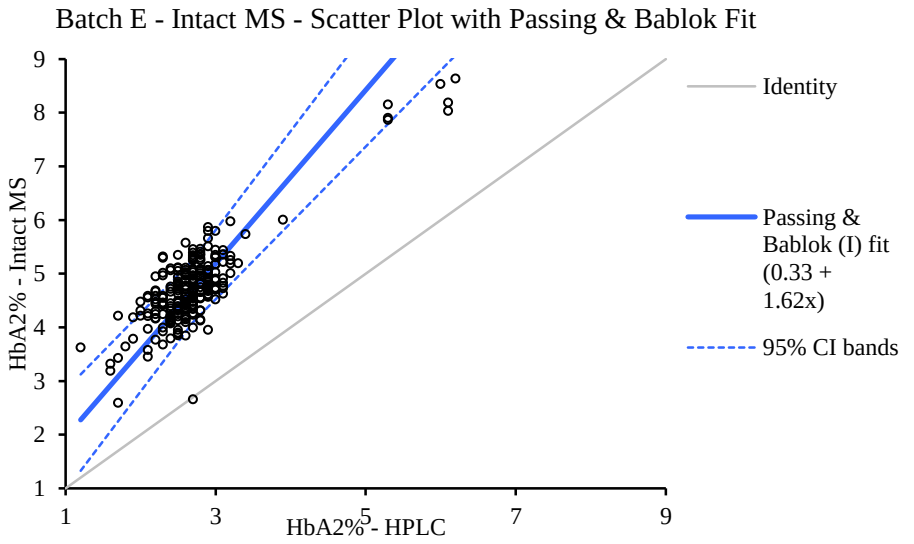


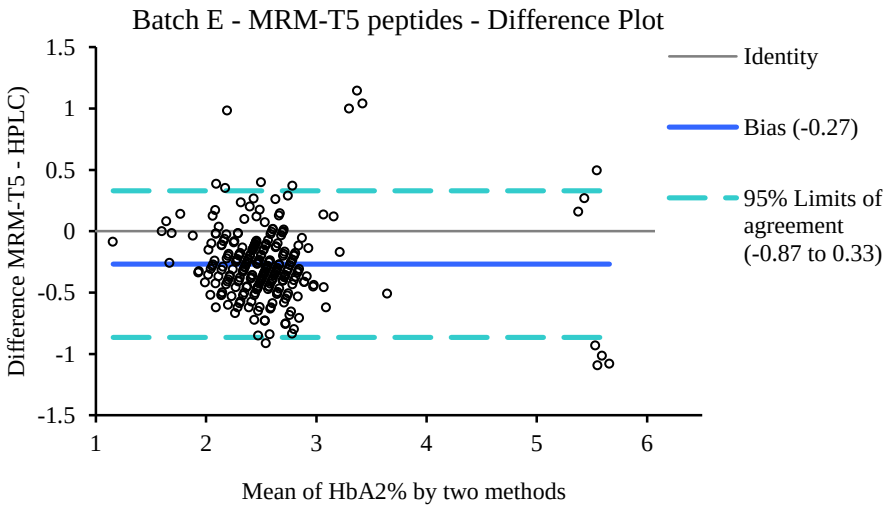


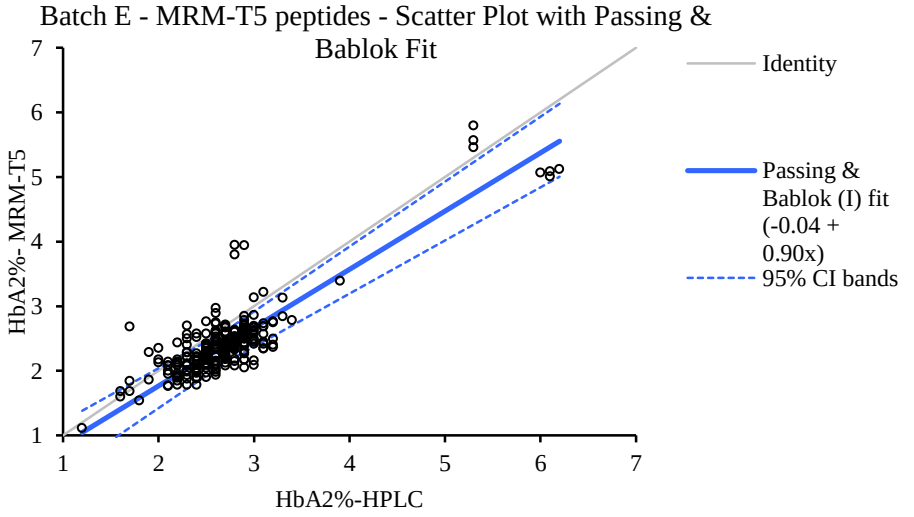
Batch E error boxplots



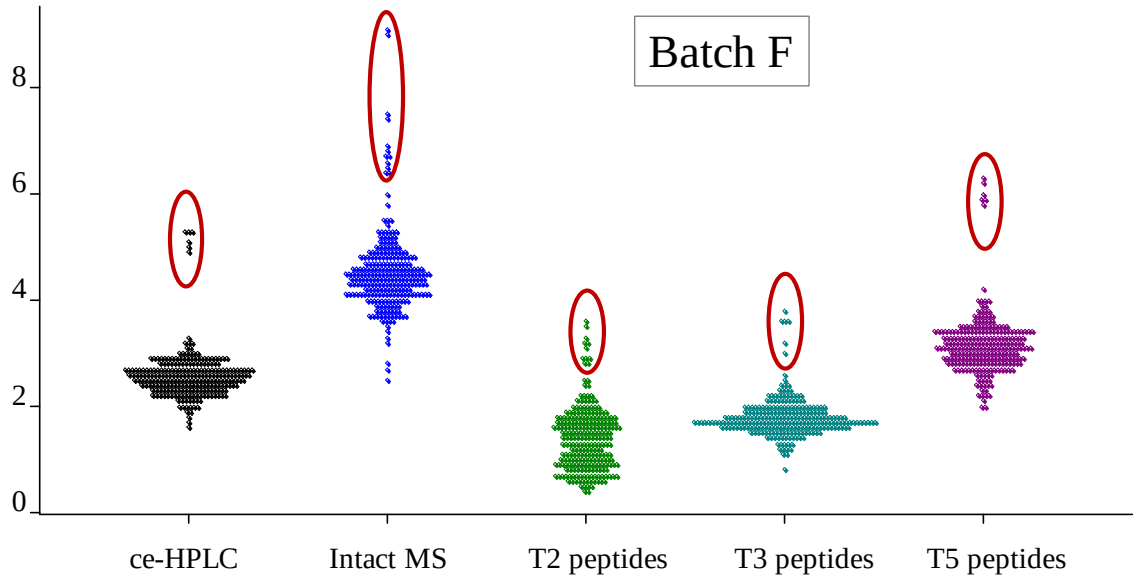

Batch E - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2%-by two methods
Identity
Bias (2.04)
95% Limits of agreement (1.26 to 2.83)

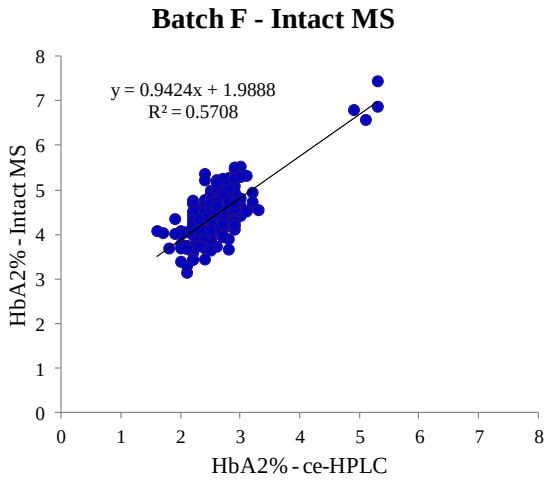

Batch E - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% - Intact MS
HbA2%- HPLC
Identity
Passing & Bablok (I) fit (0.33 + 1.62x)
95% CI bands

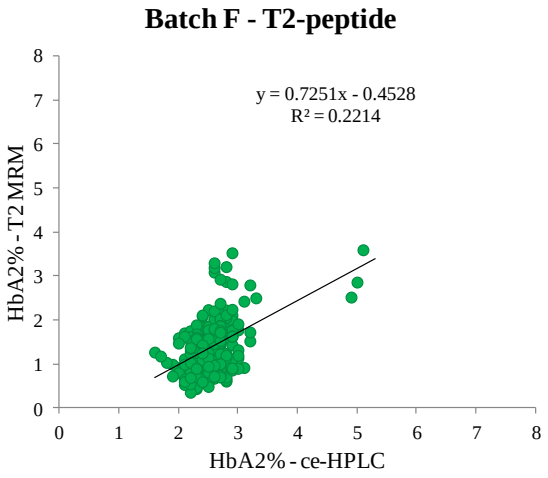

Batch E - MRM-T5 peptides - Difference Plot
Difference MRM-T5 - HPLC)
Mean of HbA2% by two methods
Identity
Bias (-0.27)
95% Limits of agreement (-0.87 to 0.33)

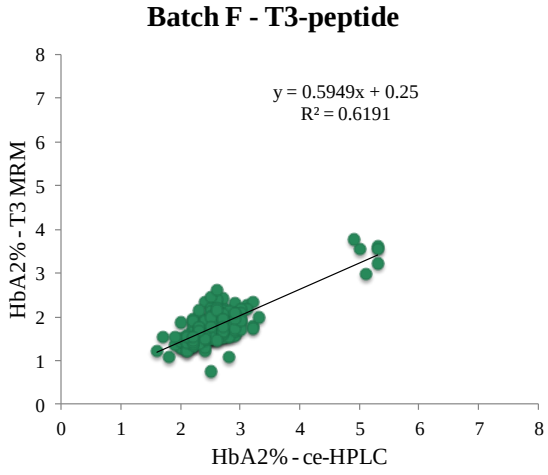

Batch E - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2%- MRM-T5
HbA2%-HPLC
Identity
Passing & Bablok (I) fit (-0.04 + 0.90x)
95% CI bands

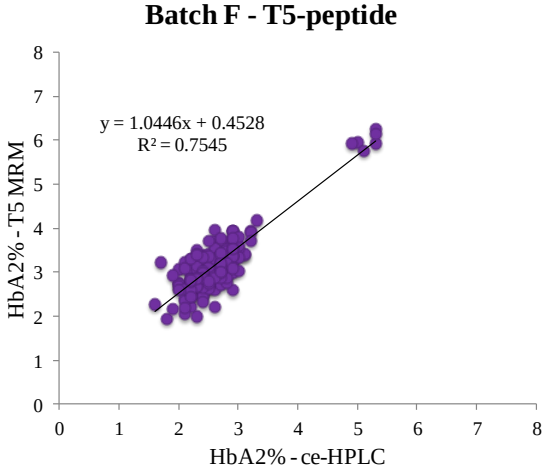
Figures for Batch F

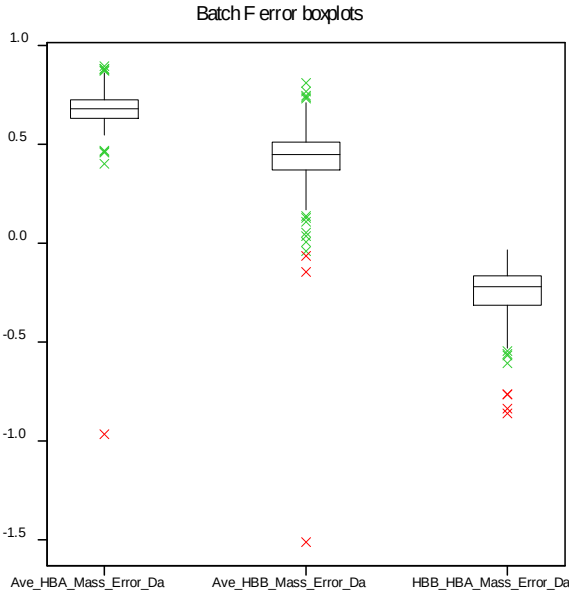


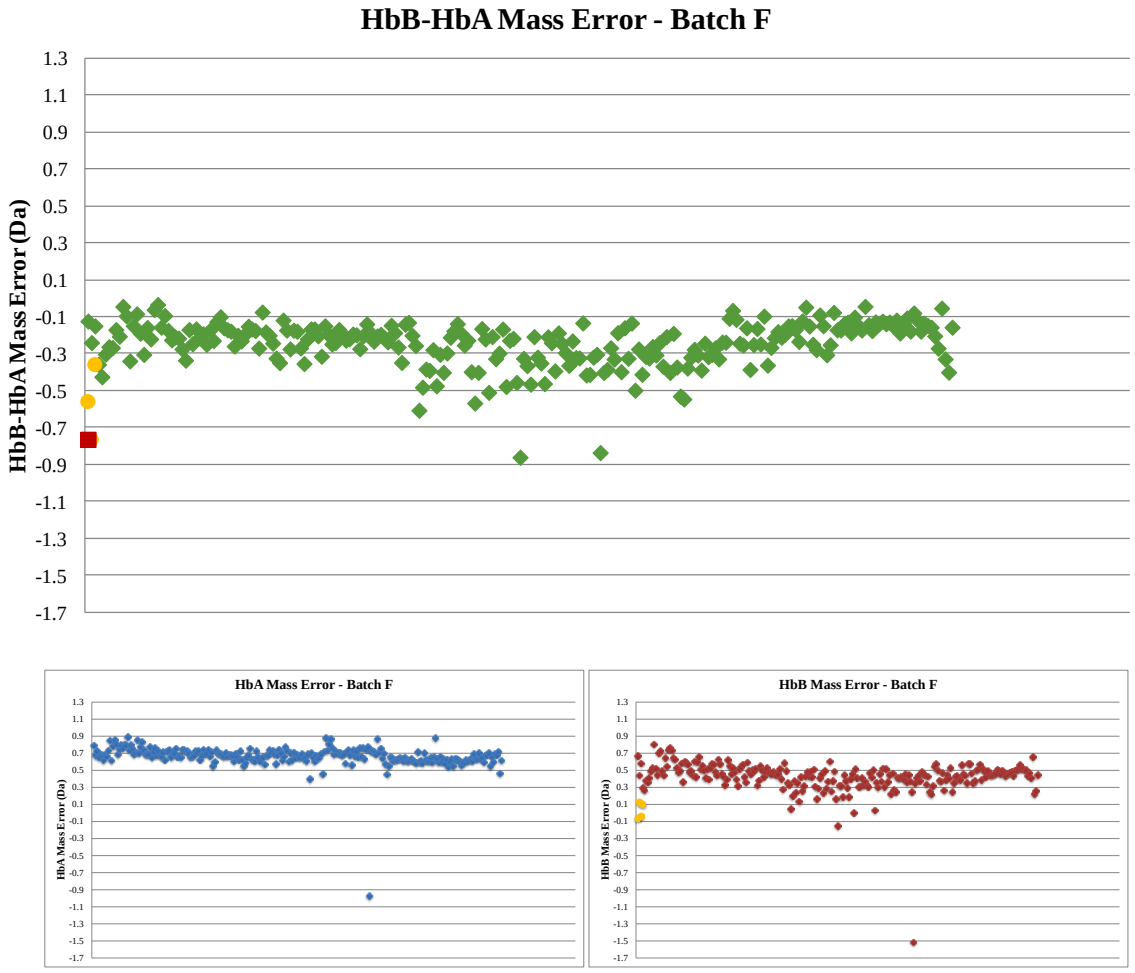






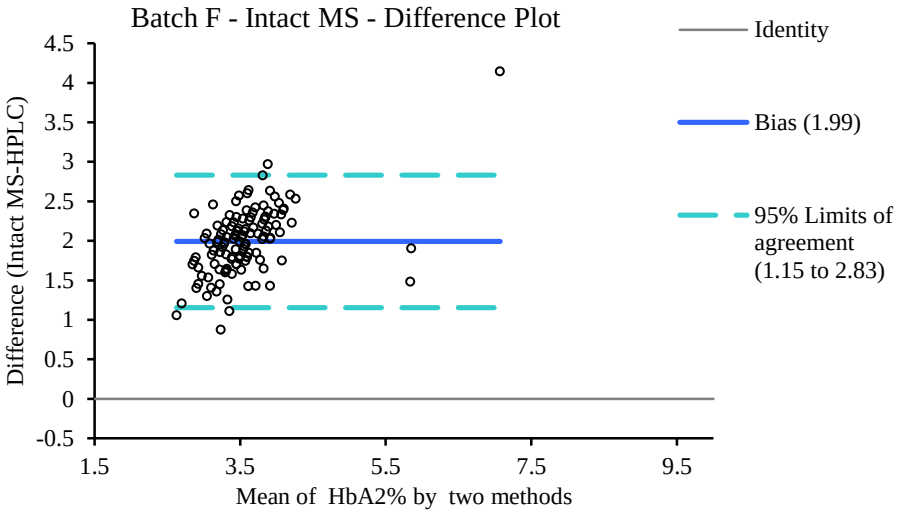


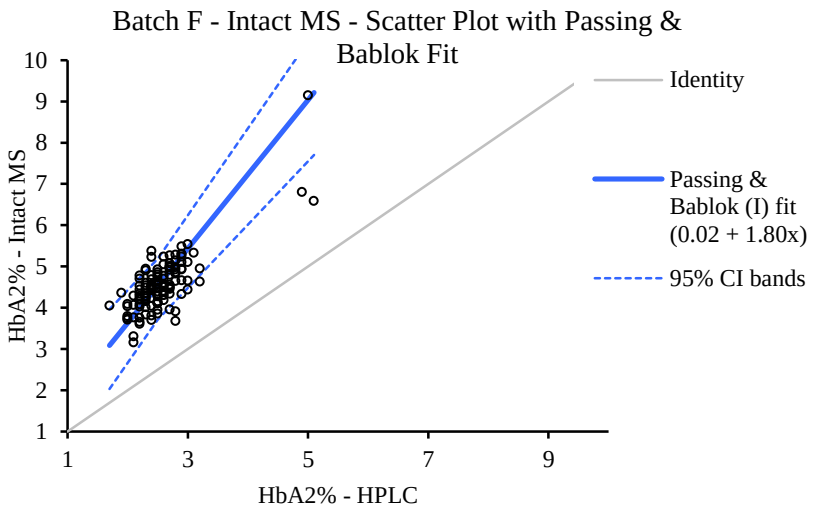

Batch F error boxplots
1.0
0.5
0.0
-0.5
-1.0
-1.5
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

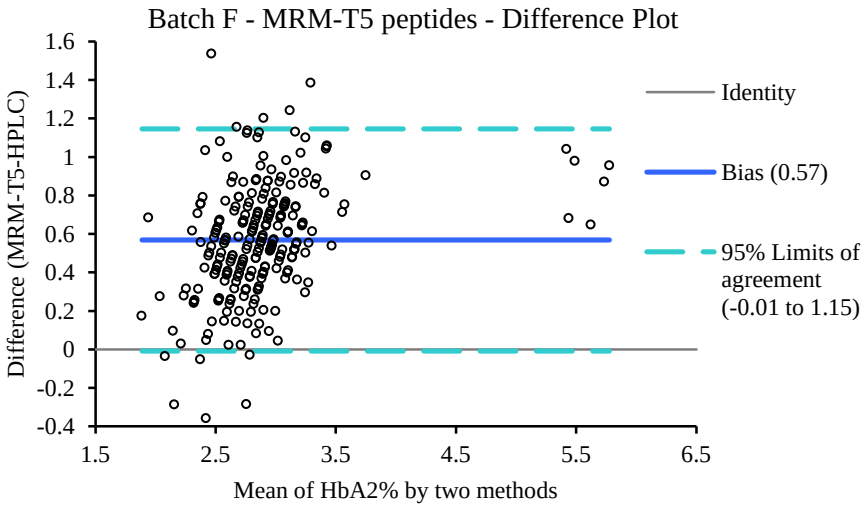

HbB-HbA Mass Error - Batch F
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

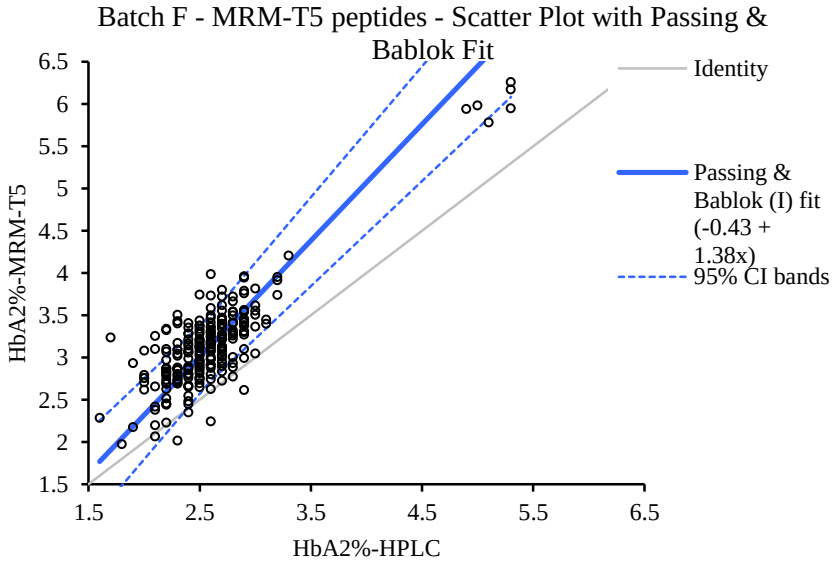
HbA Mass Error - Batch F
HbA Mass Error (Da)

HbB Mass Error - Batch F
HbB Mass Error (Da)

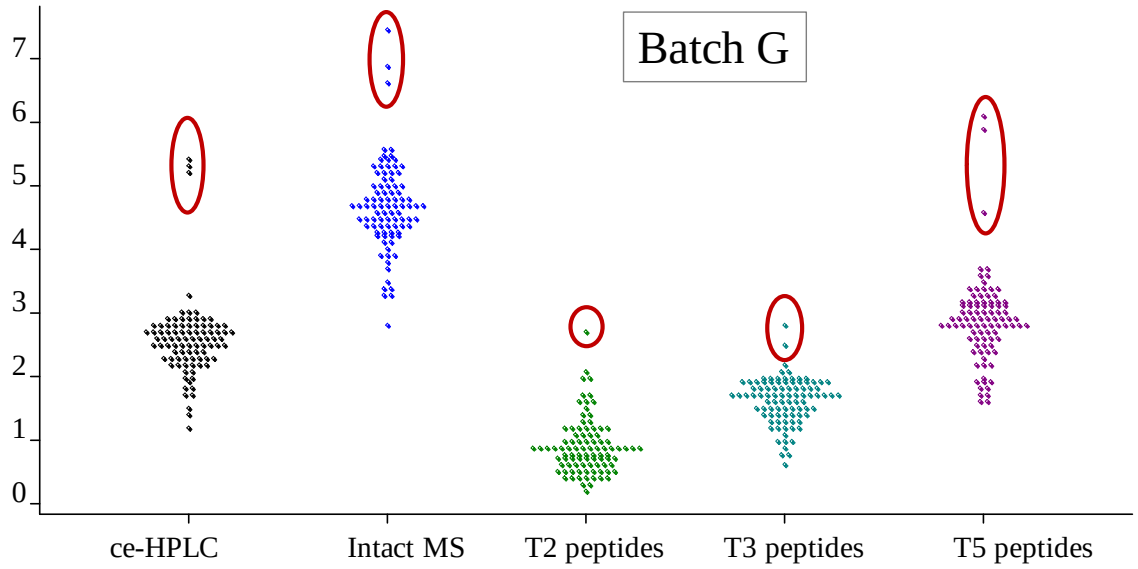

Batch F - Intact MS - Difference Plot
Difference (Intact MS-HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.99)
95% Limits of agreement (1.15 to 2.83)

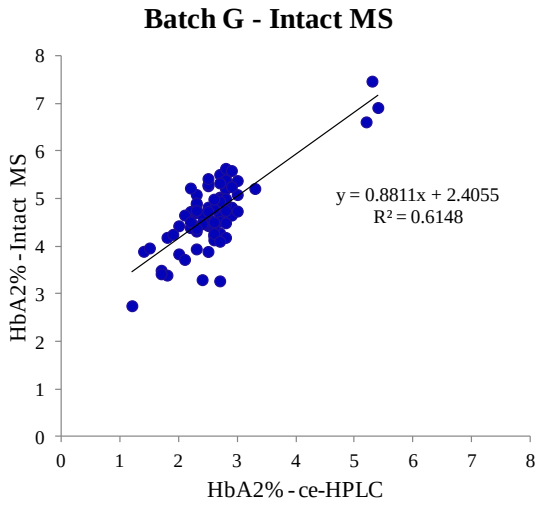

Batch F - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% - Intact MS
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (0.02 + 1.80x)
95% CI bands

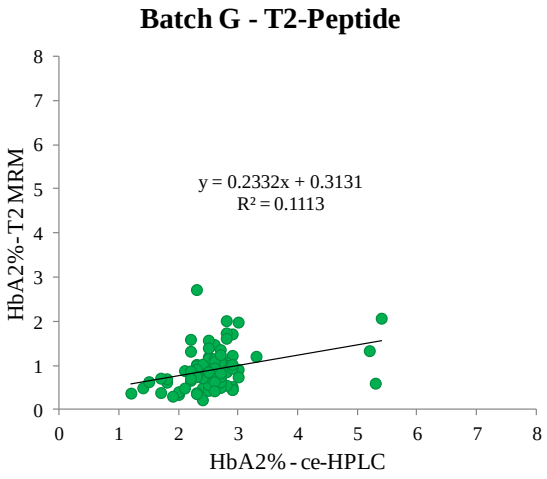

Batch F - MRM-T5 peptides - Difference Plot
Difference (MRM-T5-HPLC)
Mean of HbA2% by two methods
Identity
Bias (0.57)
95% Limits of agreement (-0.01 to 1.15)

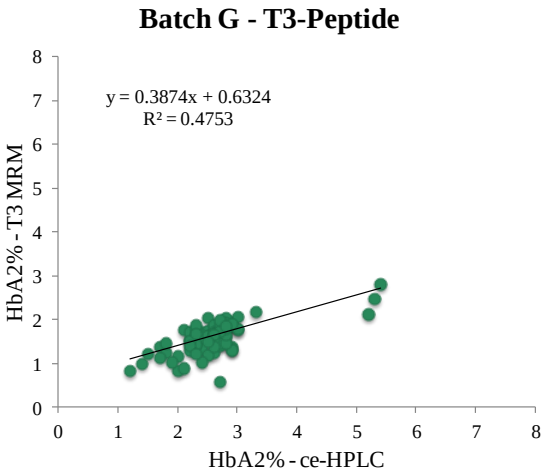

Batch F - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2%-MRM-T5
HbA2%-HPLC
Identity
Passing & Bablok (I) fit (-0.43 + 1.38x)
95% CI bands

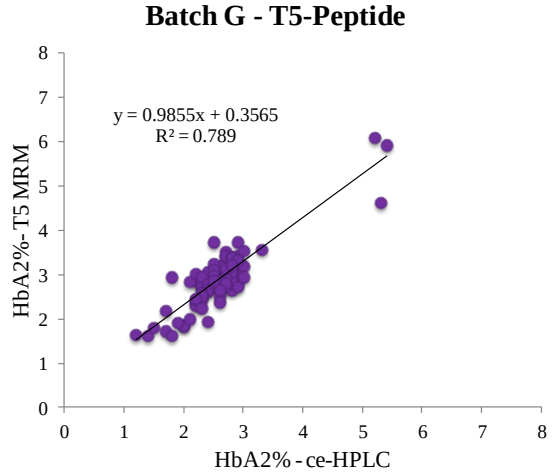
Figures for Batch G

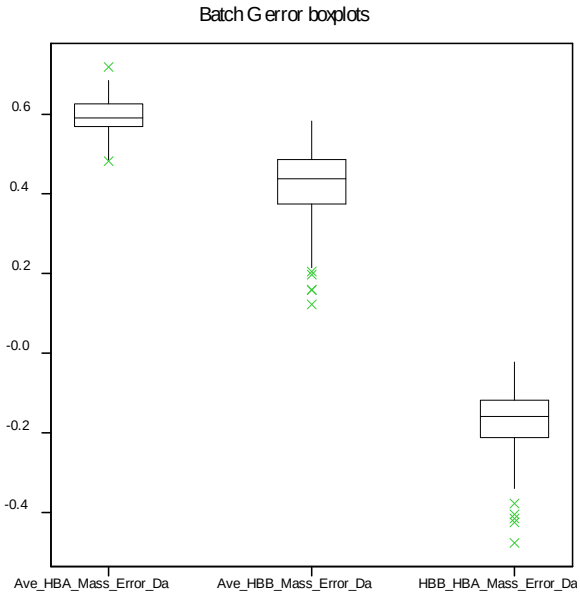


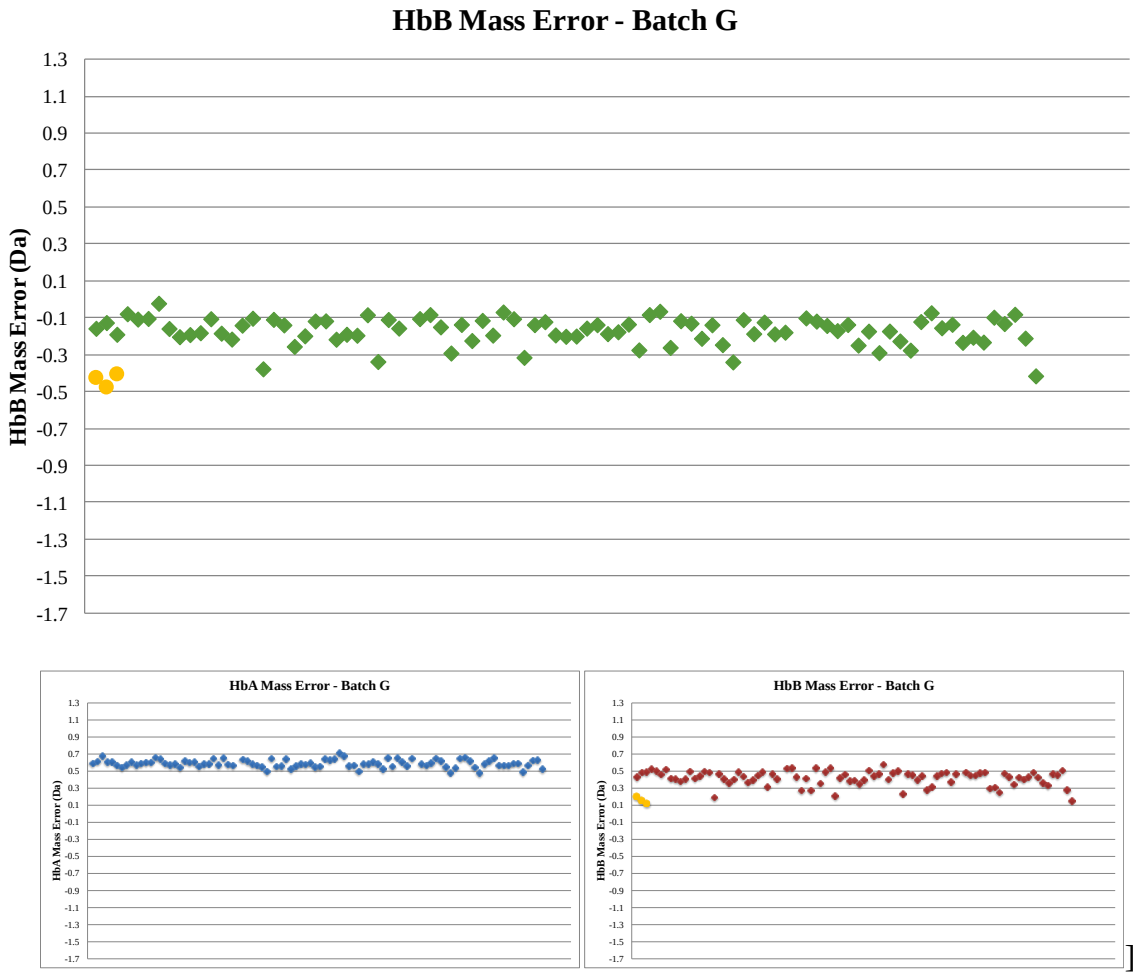


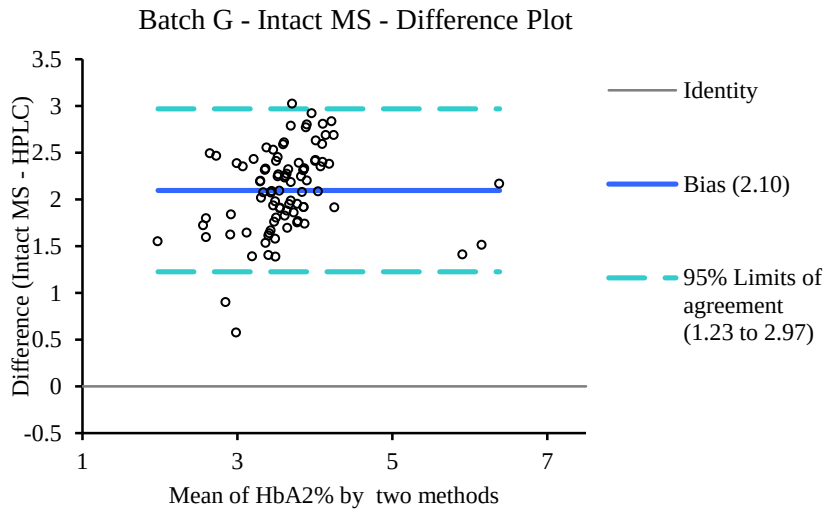


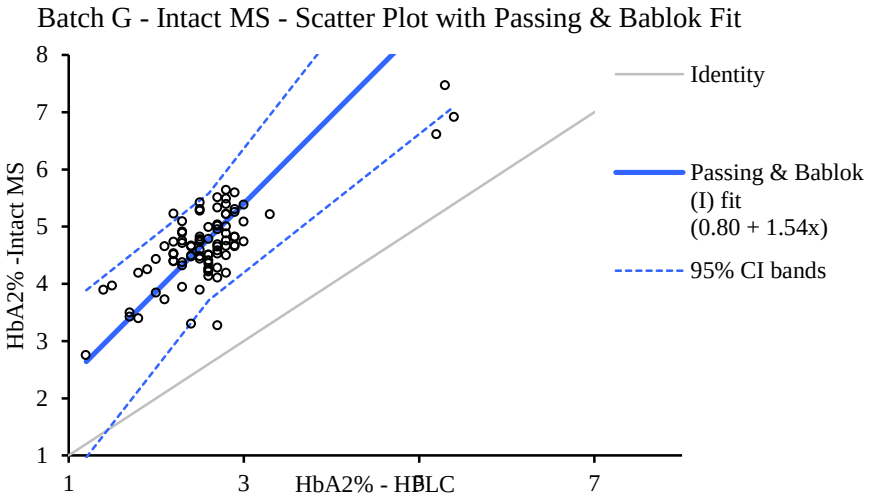


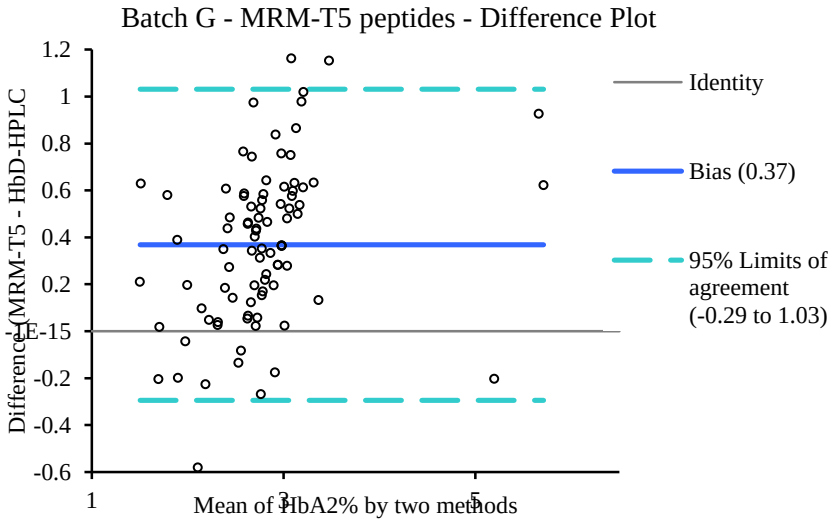


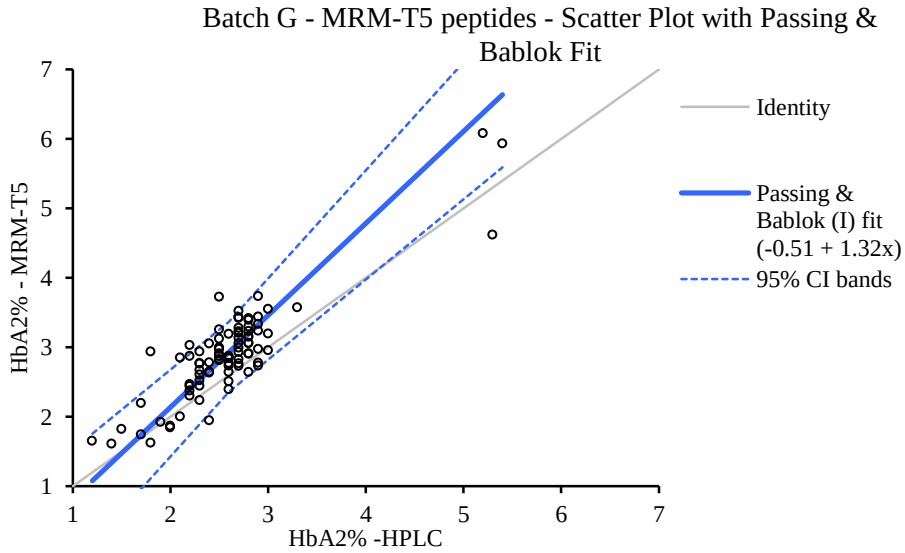

Batch G error boxplots
0.6
0.4
0.2
-0.0
-0.2
-0.4
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da


HbB Mass Error - Batch G
HbB Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7
HbA Mass Error - Batch G
HbB Mass Error - Batch G

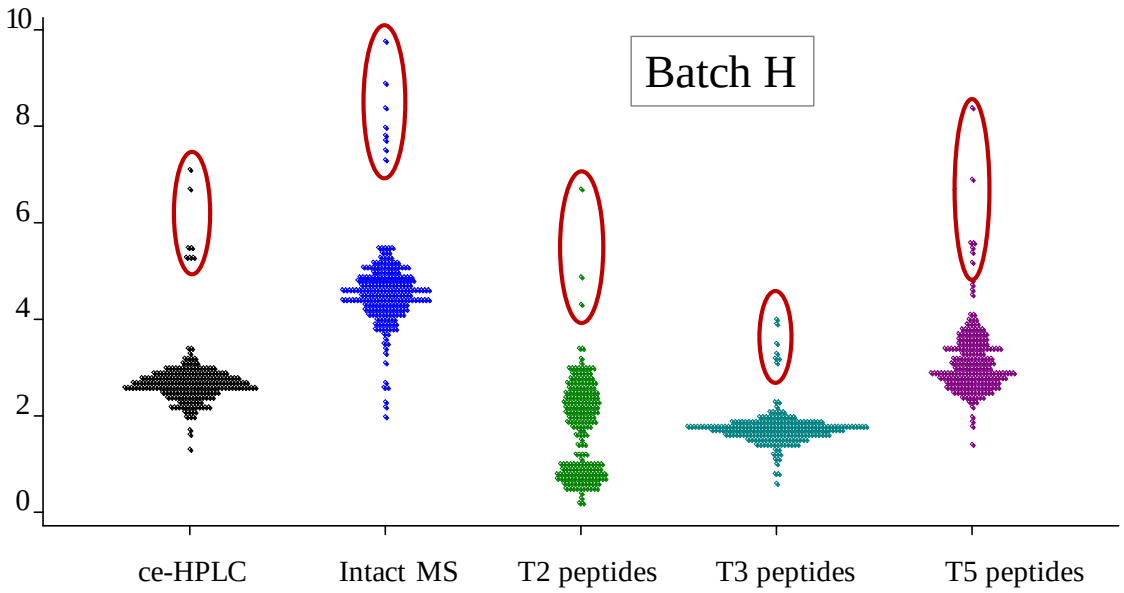

Batch G - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (2.10)
95% Limits of agreement (1.23 to 2.97)

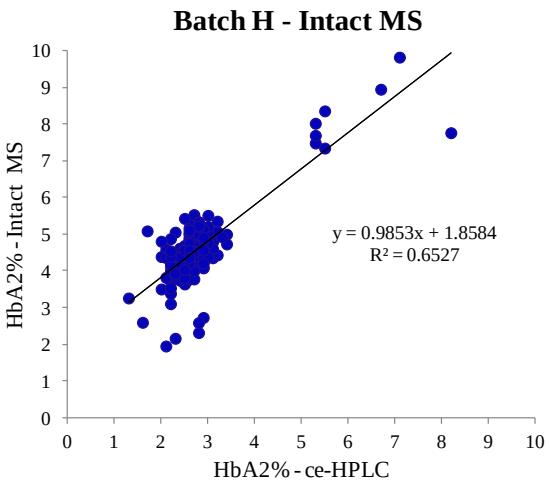

Batch G - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% -Intact MS
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (0.80 + 1.54x)
95% CI bands

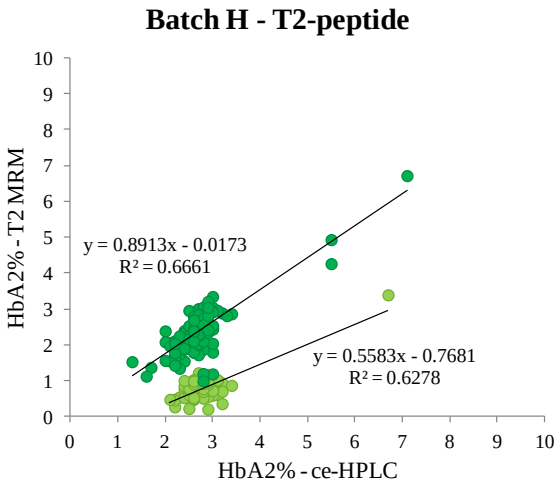

Batch G - MRM-T5 peptides - Difference Plot
Difference (MRM-T5 - HbD-HPLC
Mean of HbA2% by two methods
Identity
Bias (0.37)
95% Limits of agreement (-0.29 to 1.03)

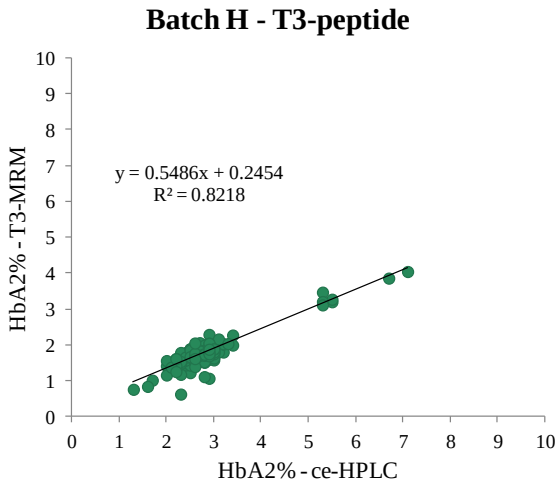

Batch G - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2% - MRM-T5
HbA2% -HPLC
Identity
Passing & Bablok (I) fit (-0.51 + 1.32x)
95% CI bands

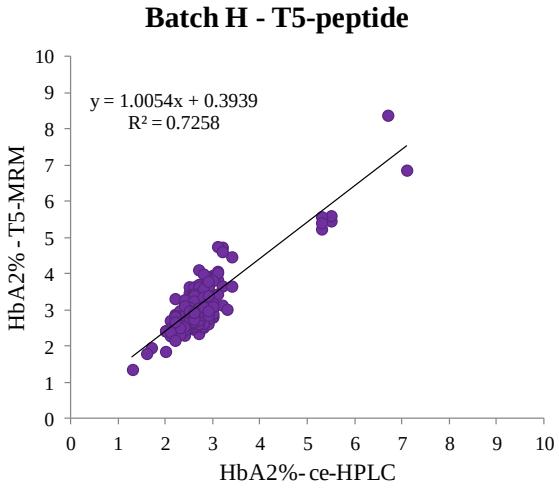
Figures for Batch H

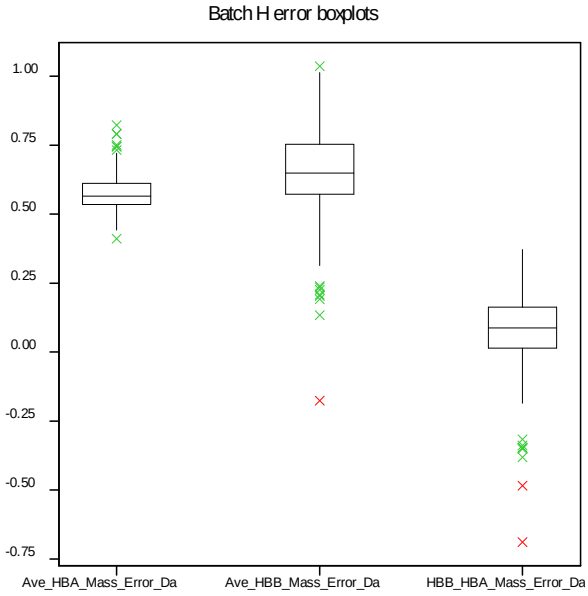


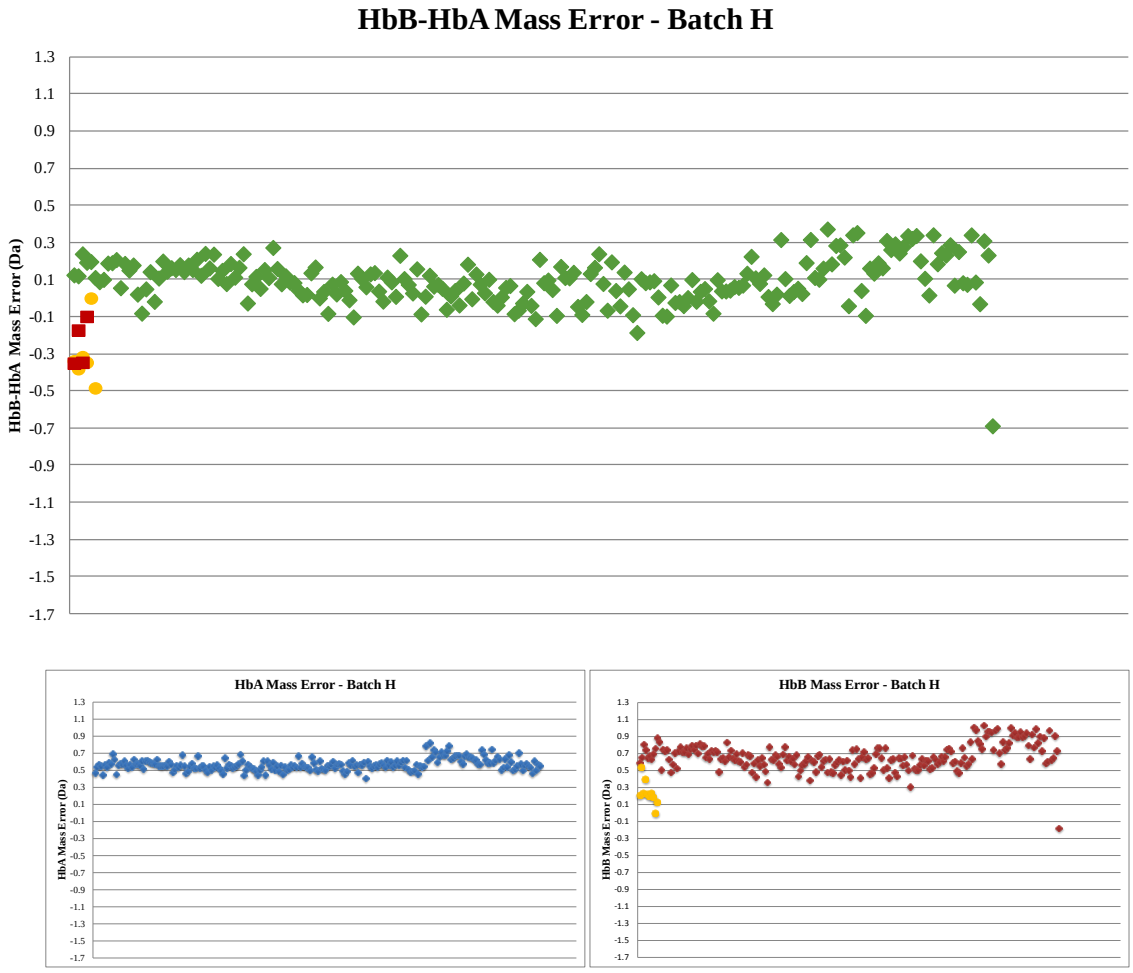






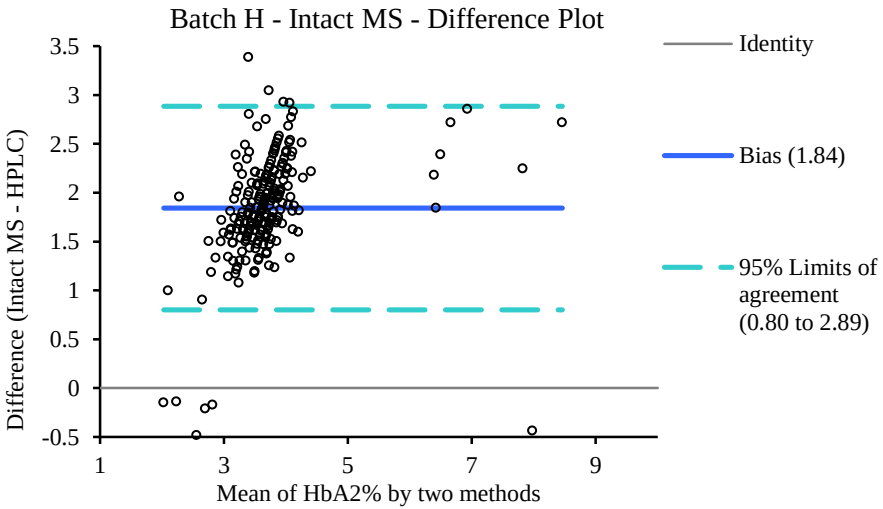


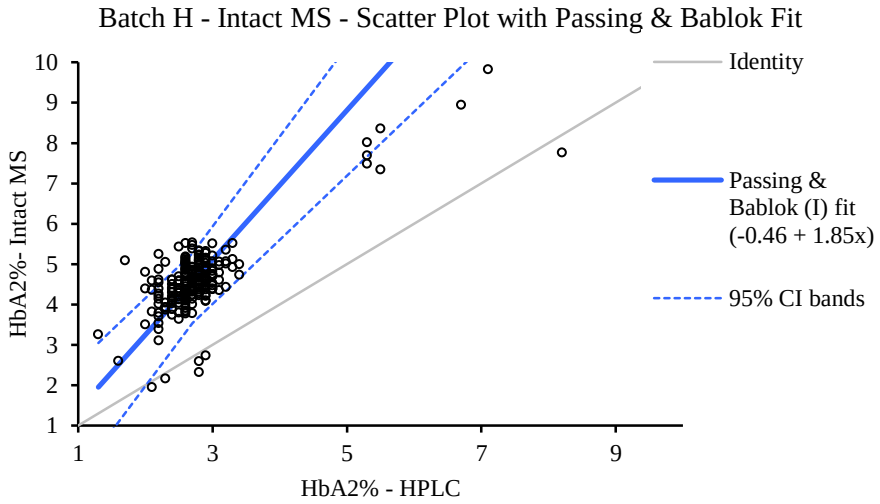

Batch H error boxplots
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

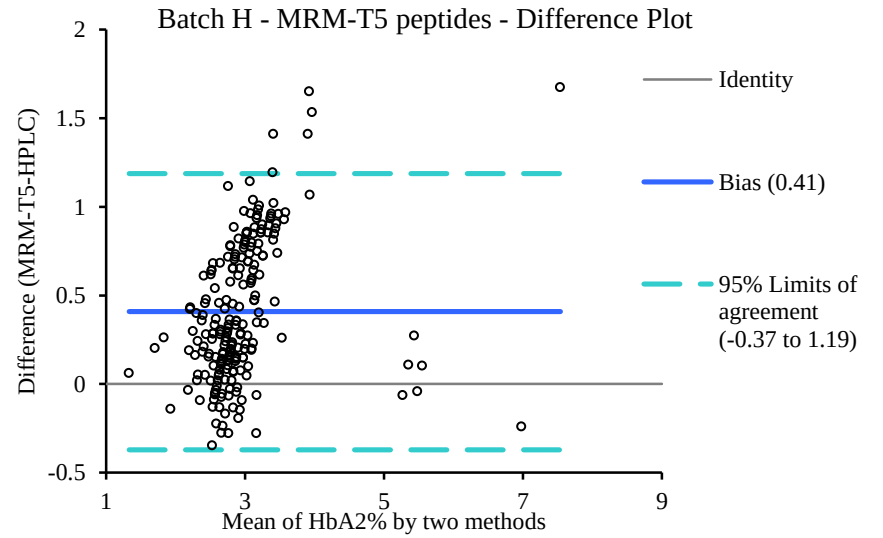

HbB-HbA Mass Error - Batch H
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

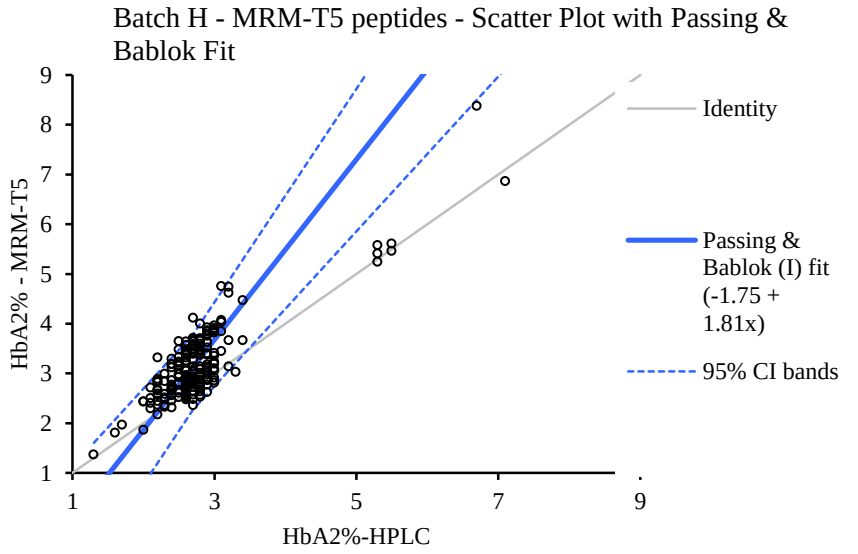
HbA Mass Error - Batch H
HbA Mass Error (Da)

HbB Mass Error - Batch H
HbB Mass Error (Da)

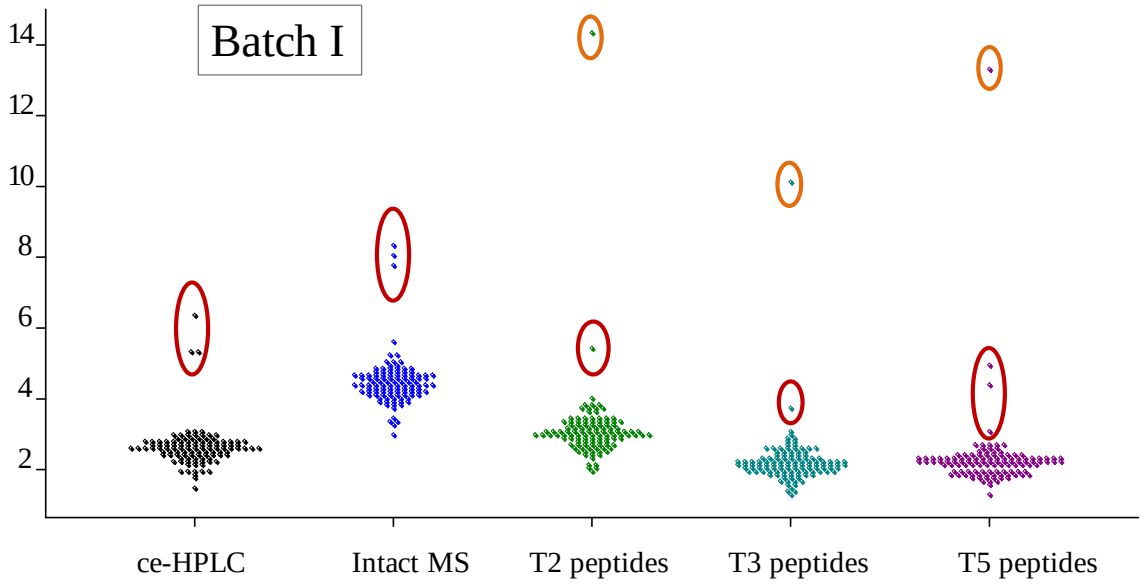

Batch H - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.84)
95% Limits of agreement (0.80 to 2.89)

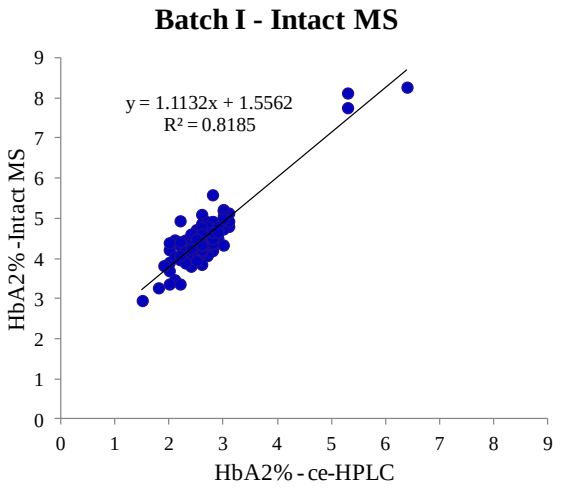

Batch H - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2%- Intact MS
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (-0.46 + 1.85x)
95% CI bands

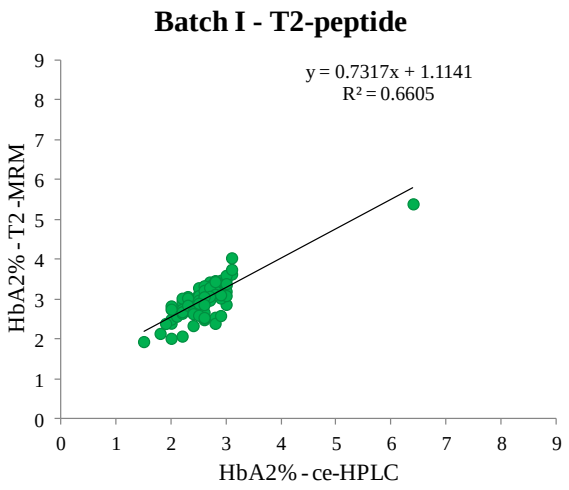

Batch H - MRM-T5 peptides - Difference Plot
Difference (MRM-T5-HPLC)
Mean of HbA2% by two methods
Identity
Bias (0.41)
95% Limits of agreement (-0.37 to 1.19)

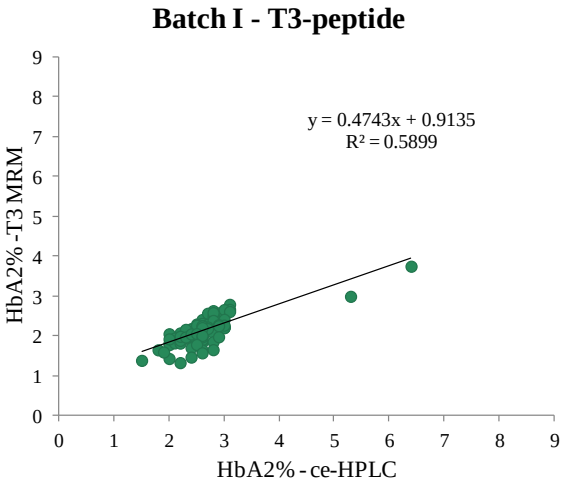

Batch H - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2% - MRM-T5
HbA2%-HPLC
Identity
Passing & Bablok (I) fit (-1.75 + 1.81x)
95% CI bands

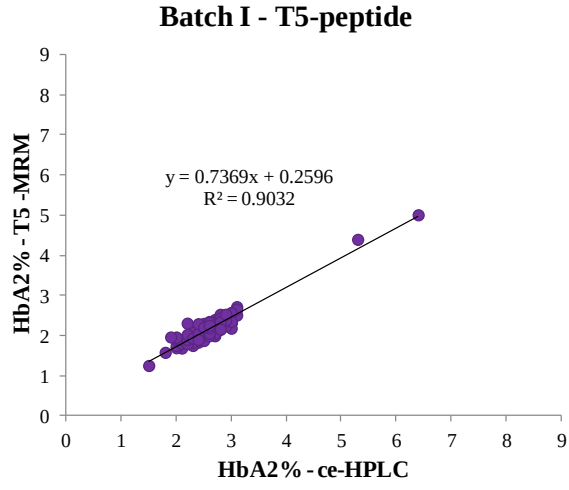
Figures for Batch I

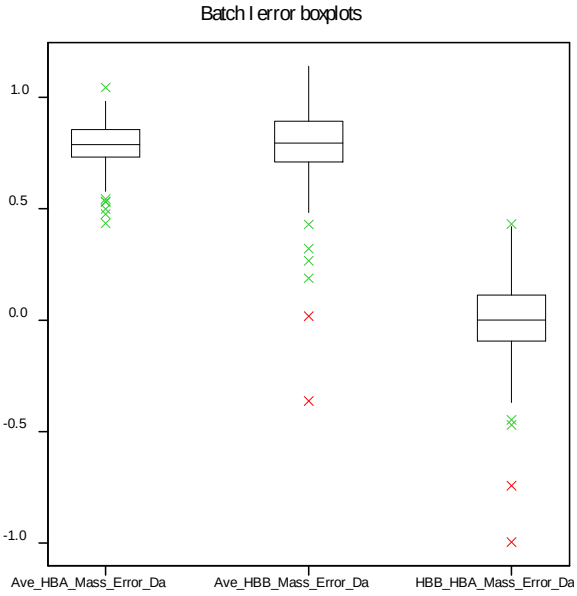


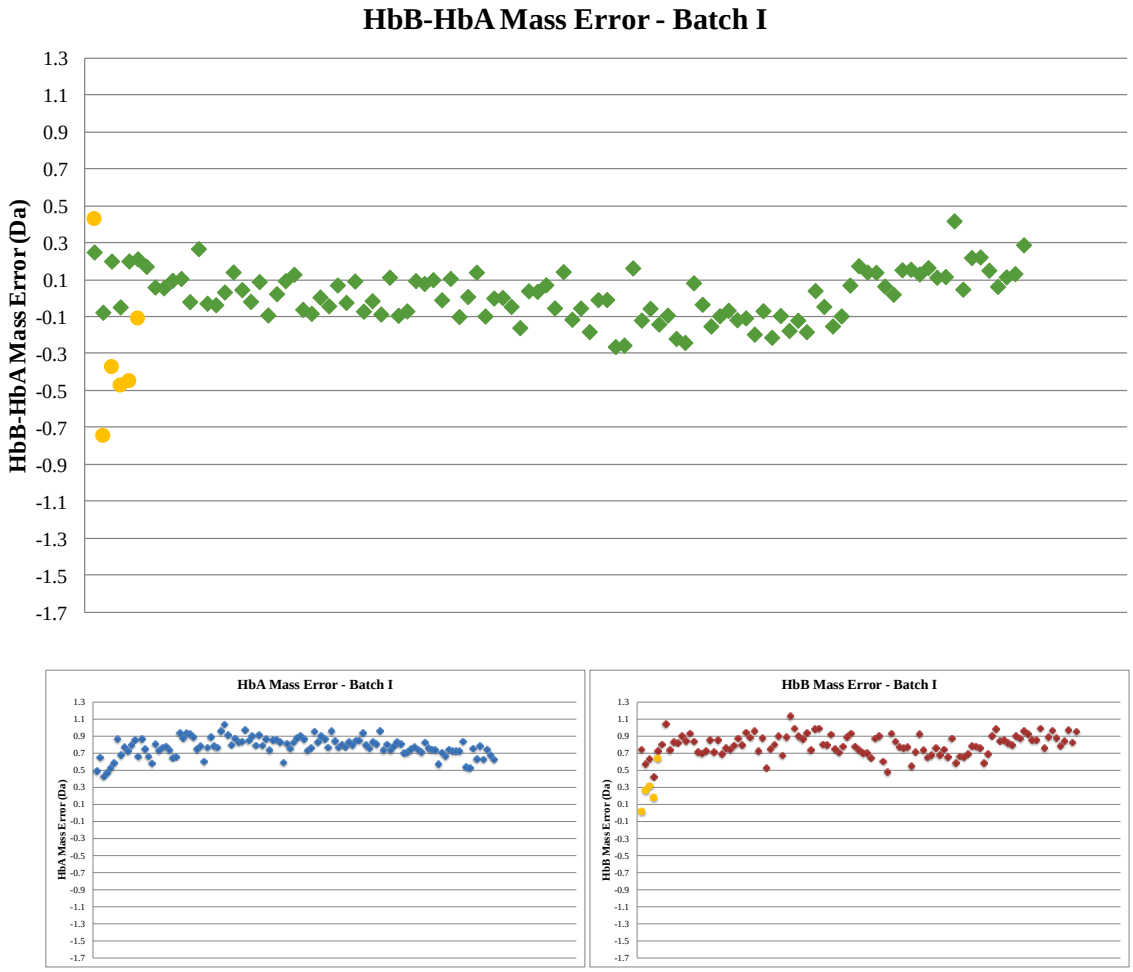






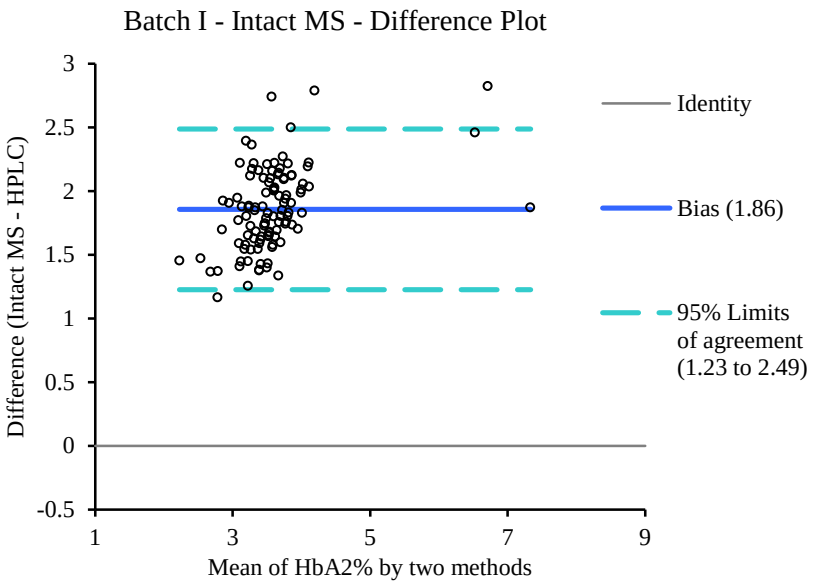


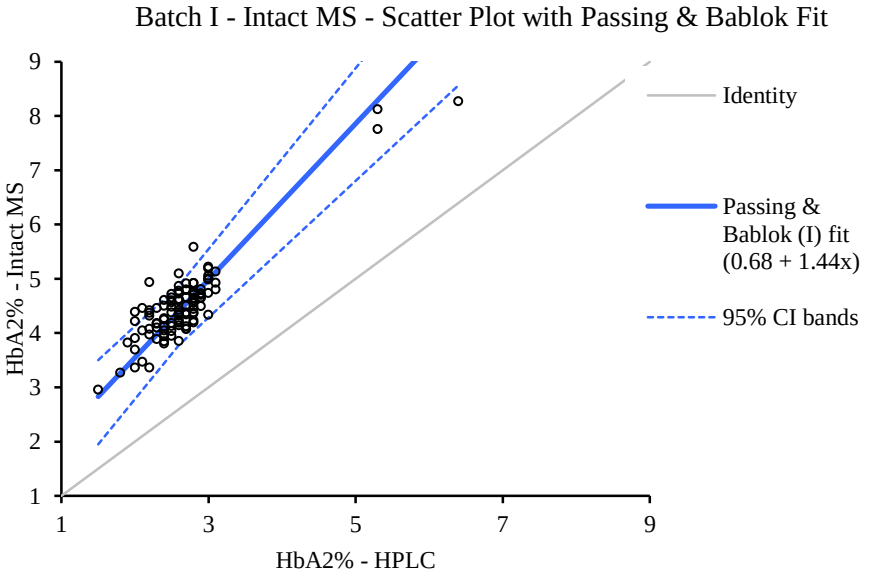

Batch I error boxplots
1.0
0.5
0.0
-0.5
-1.0
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

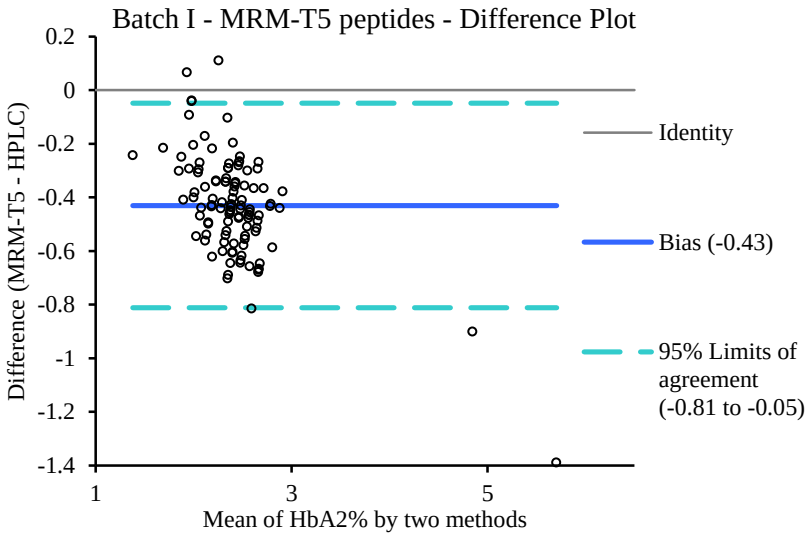

HbB-HbA Mass Error - Batch I
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

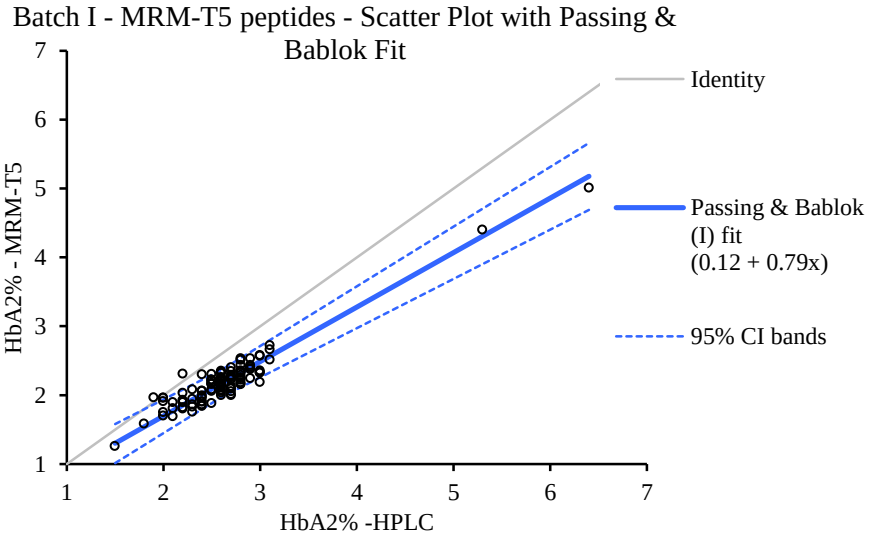
HbA Mass Error - Batch I
HbA Mass Error (Da)

HbB Mass Error - Batch I
HbB Mass Error (Da)

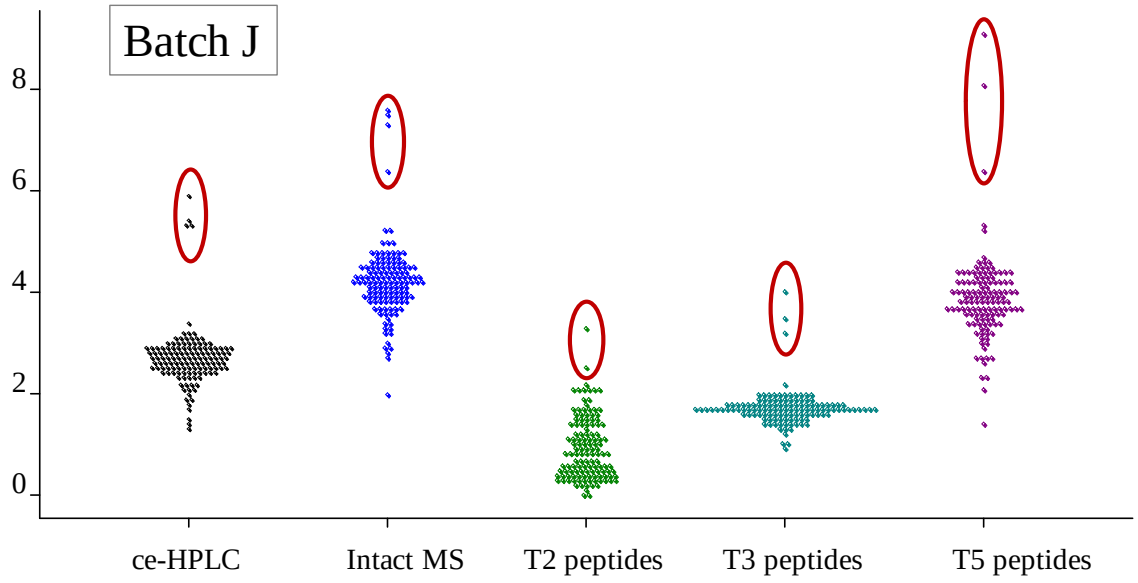

Batch I - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.86)
95% Limits of agreement (1.23 to 2.49)

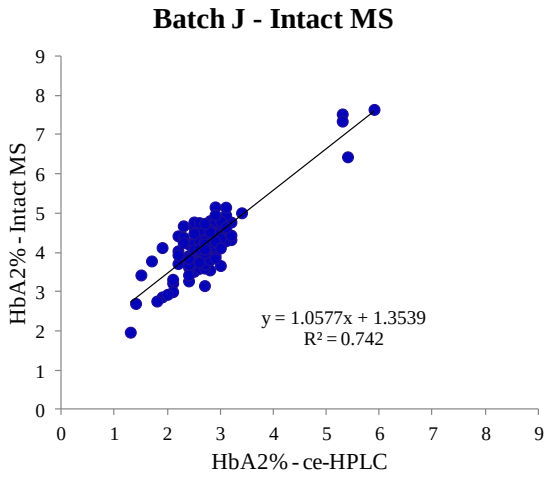

Batch I - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% - Intact MS
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (0.68 + 1.44x)
95% CI bands

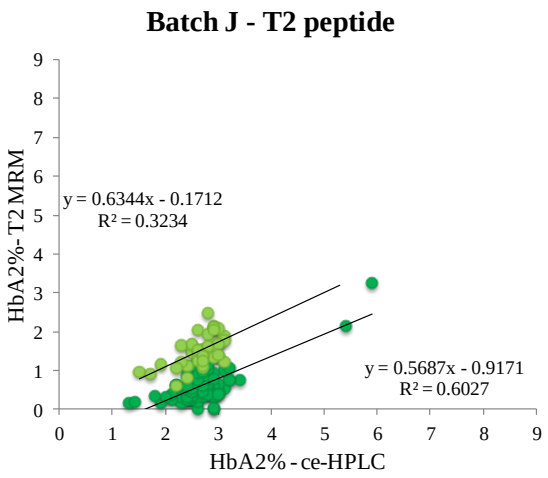

Batch I - MRM-T5 peptides - Difference Plot
Difference (MRM-T5 - HPLC)
Mean of HbA2% by two methods
Identity
Bias (-0.43)
95% Limits of agreement (-0.81 to -0.05)

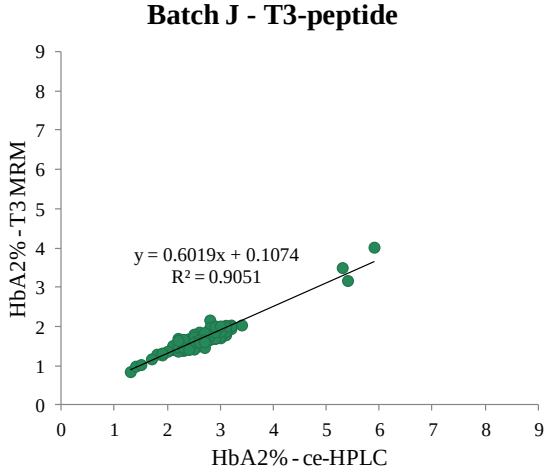

Batch I - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2% - MRM-T5
HbA2% -HPLC
Identity
Passing & Bablok (I) fit (0.12 + 0.79x)
95% CI bands

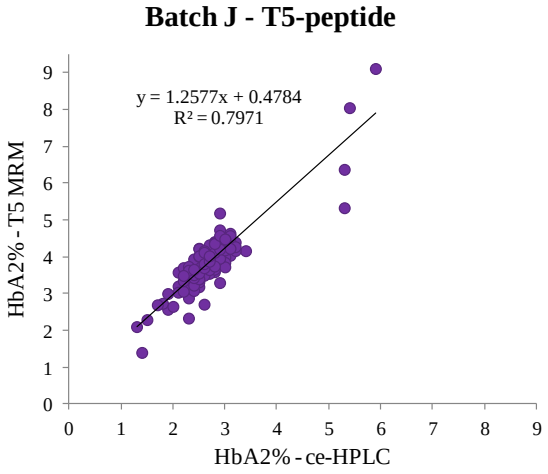
Figures for Batch J

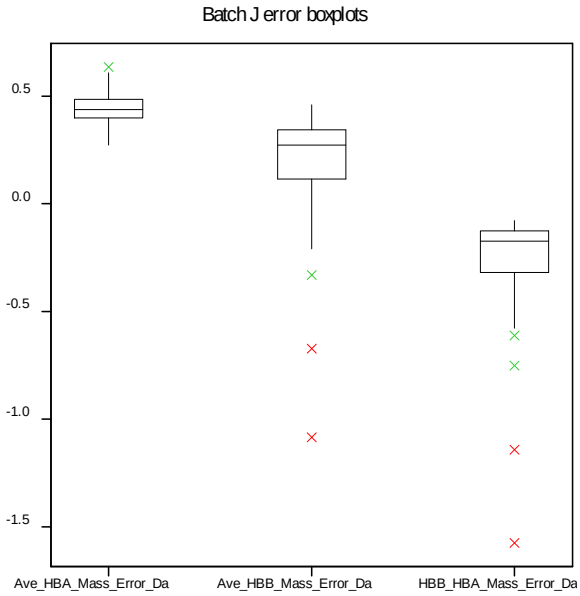


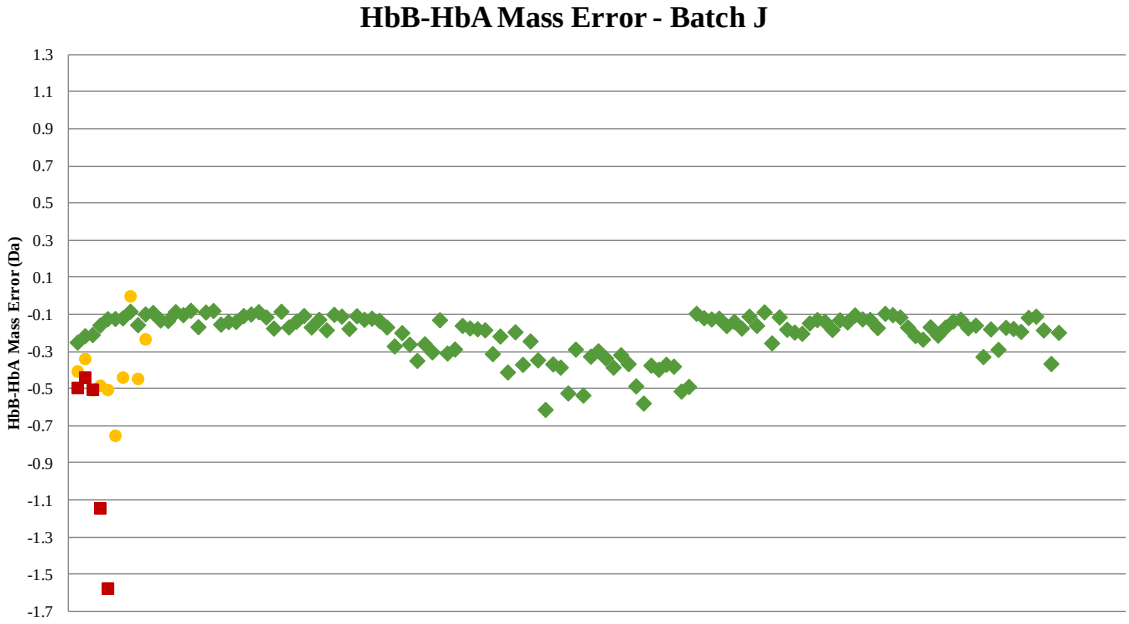


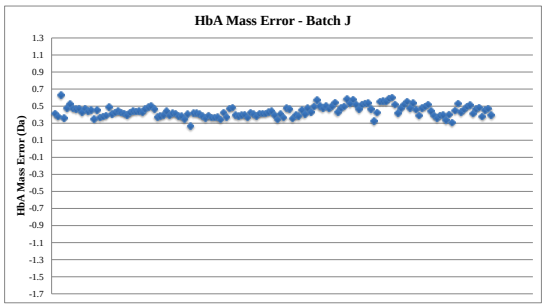


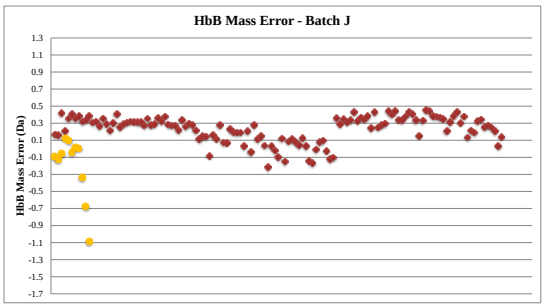




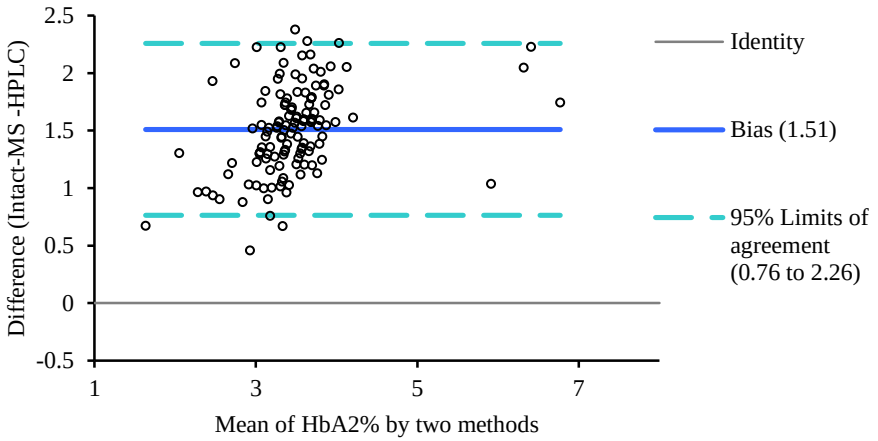

Batch J error boxplots
0.5
0.0
-0.5
-1.0
-1.5
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da


HbB-HbA Mass Error - Batch J
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

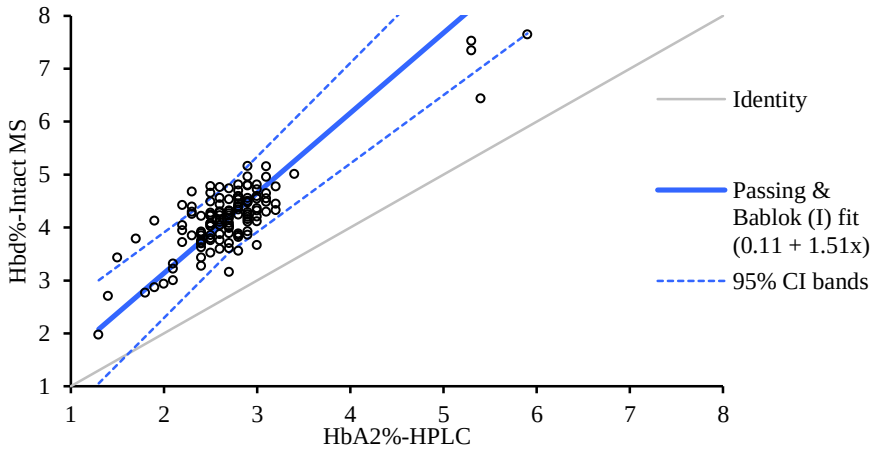

HbA Mass Error - Batch J
HbA Mass Error (Da)


HbB Mass Error - Batch J
HbB Mass Error (Da)

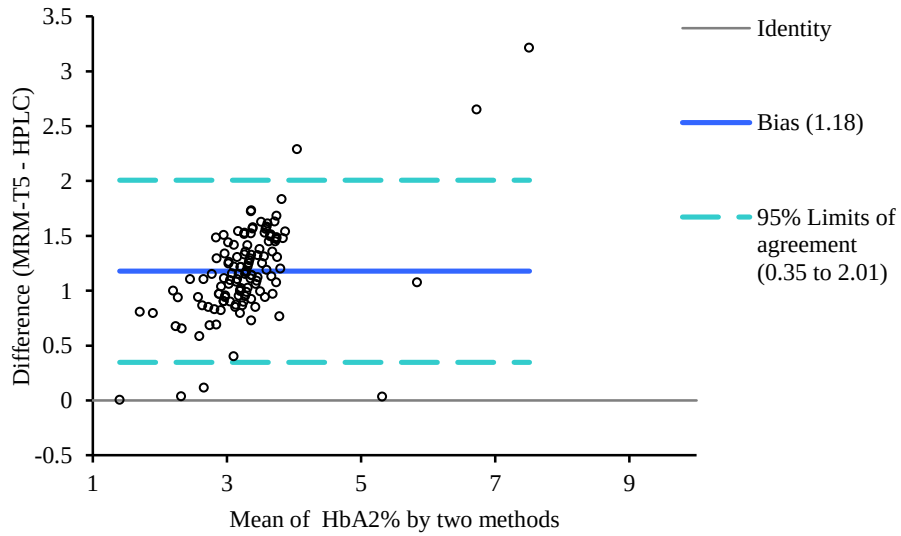
Batch J - Intact MS - Difference Plot



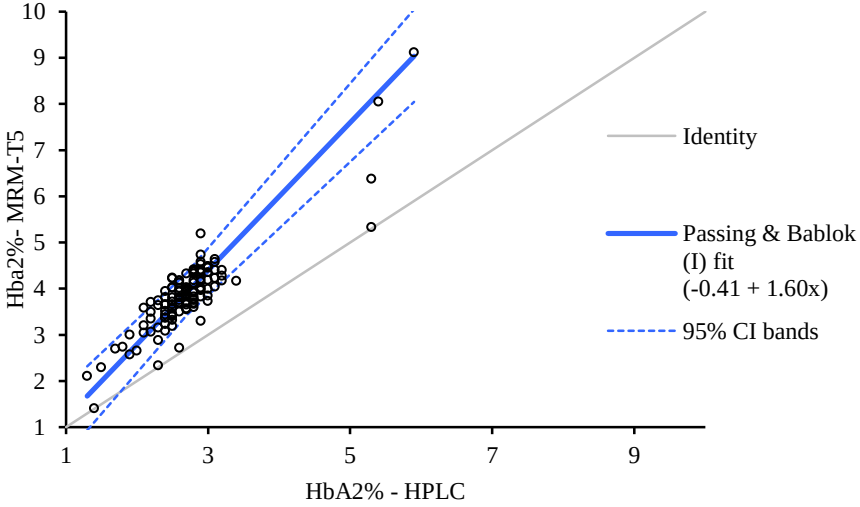
Batch J - Intact MS - Scatter Plot with Passing & Bablok Fit



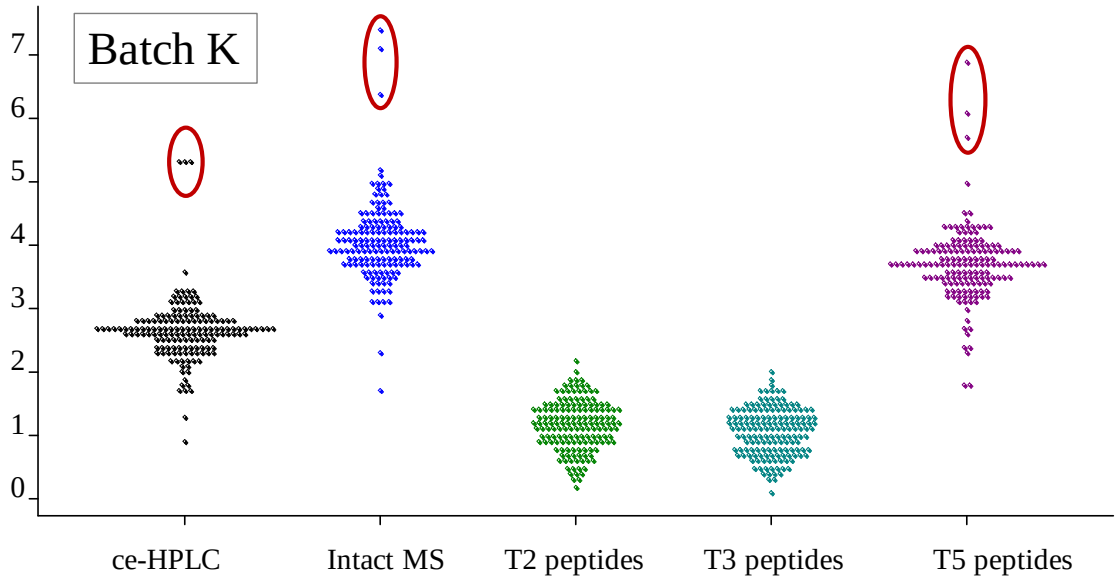
Batch J - MRM-T5 peptides - Difference Plot

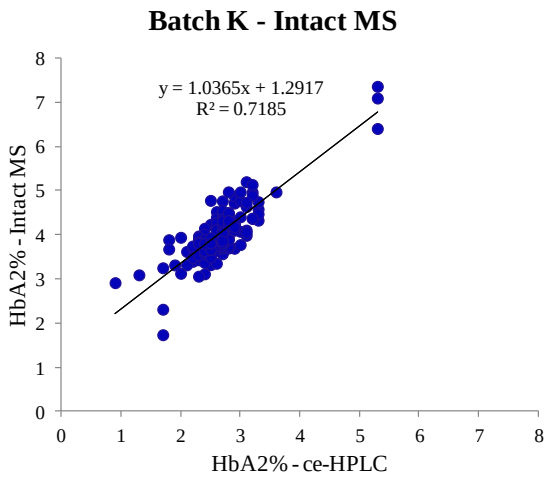


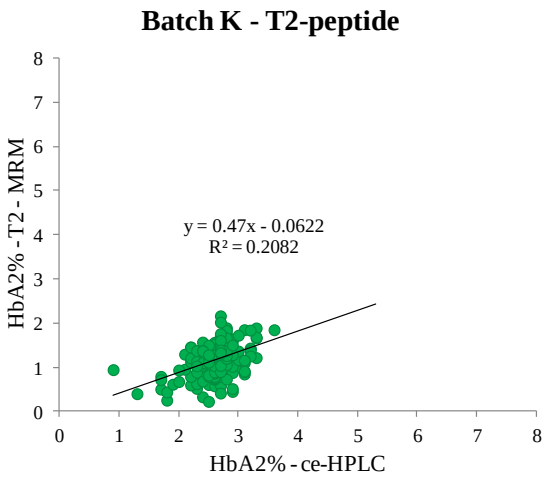
Batch J - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit

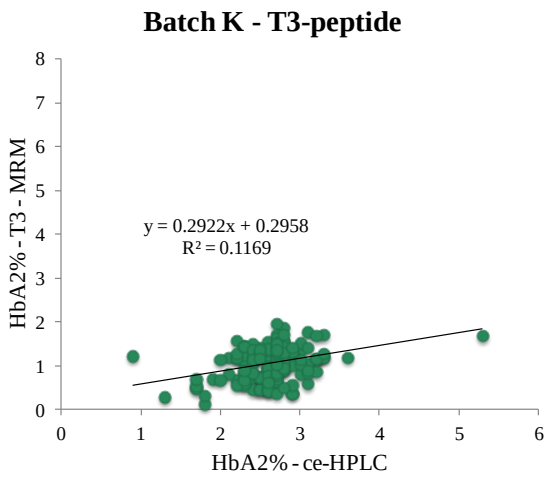


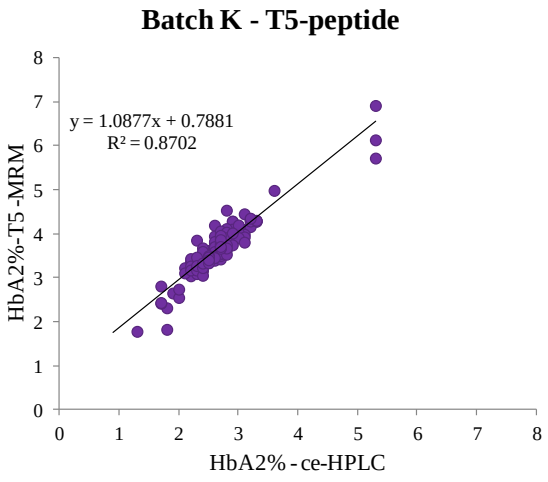
Figures for Batch K

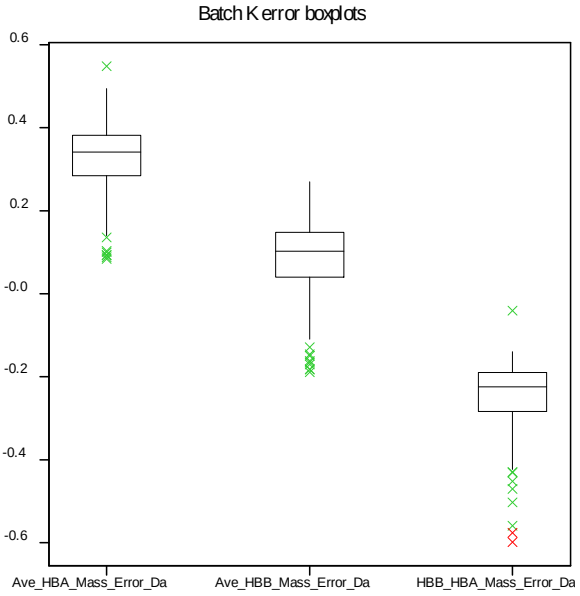


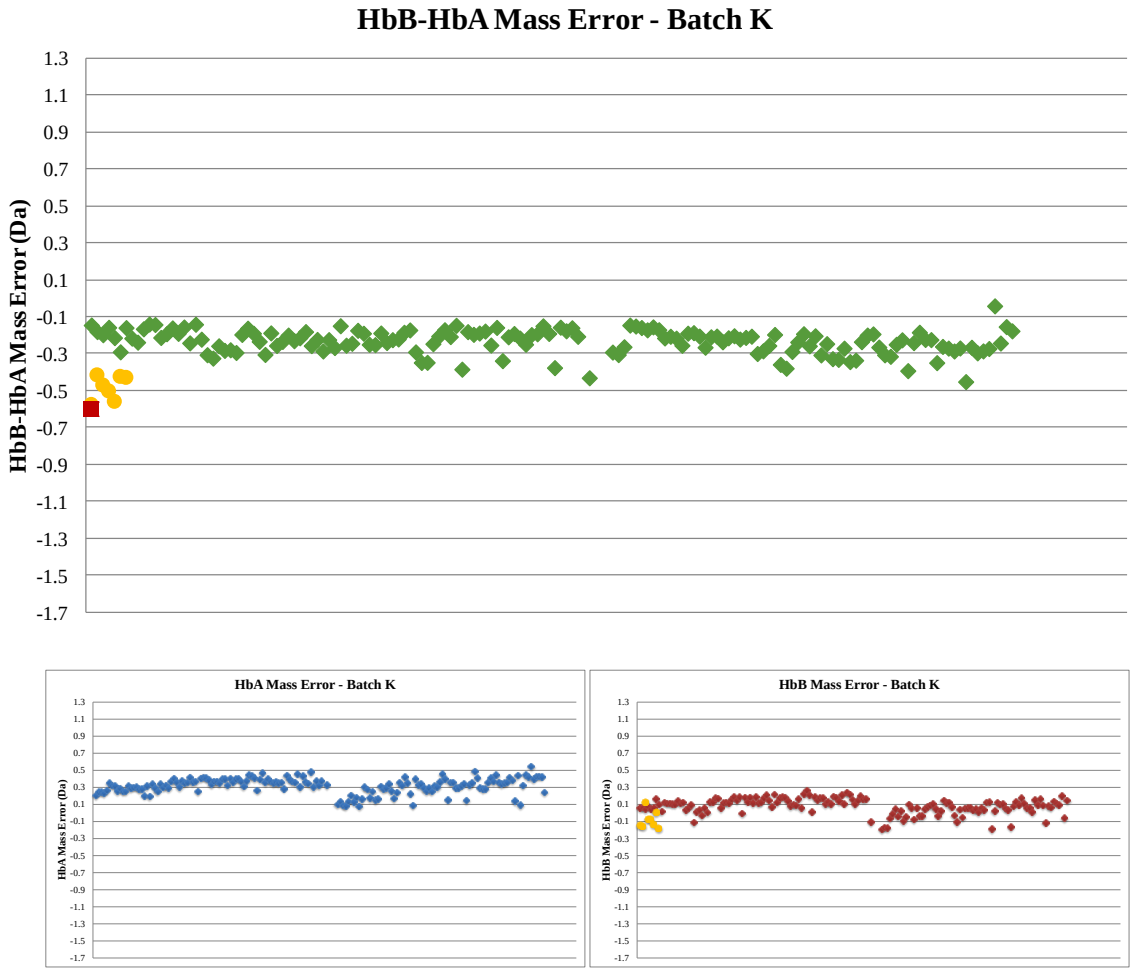






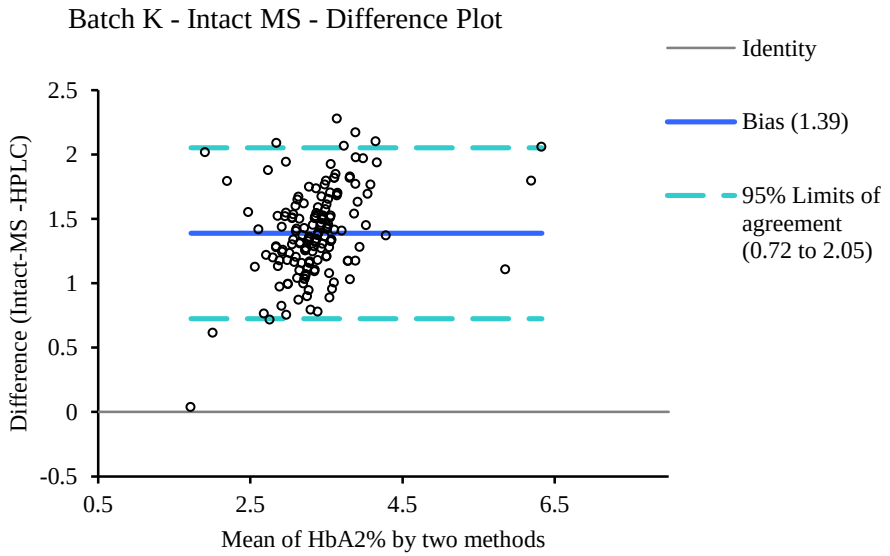


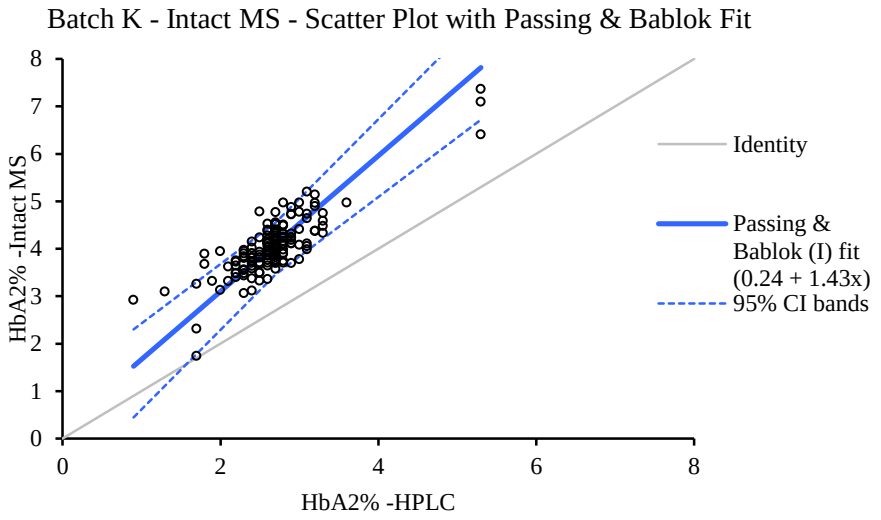

Batch K error boxplots
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

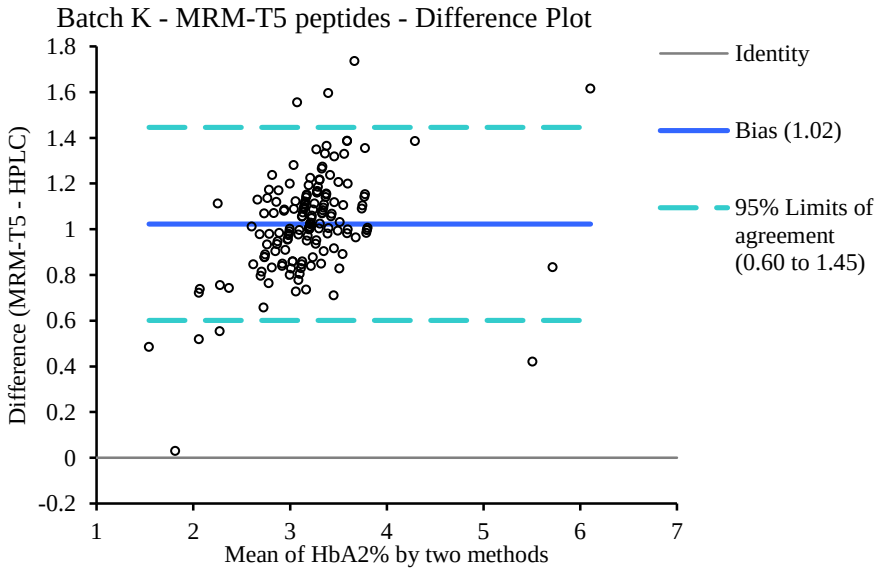

HbB-HbA Mass Error - Batch K
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

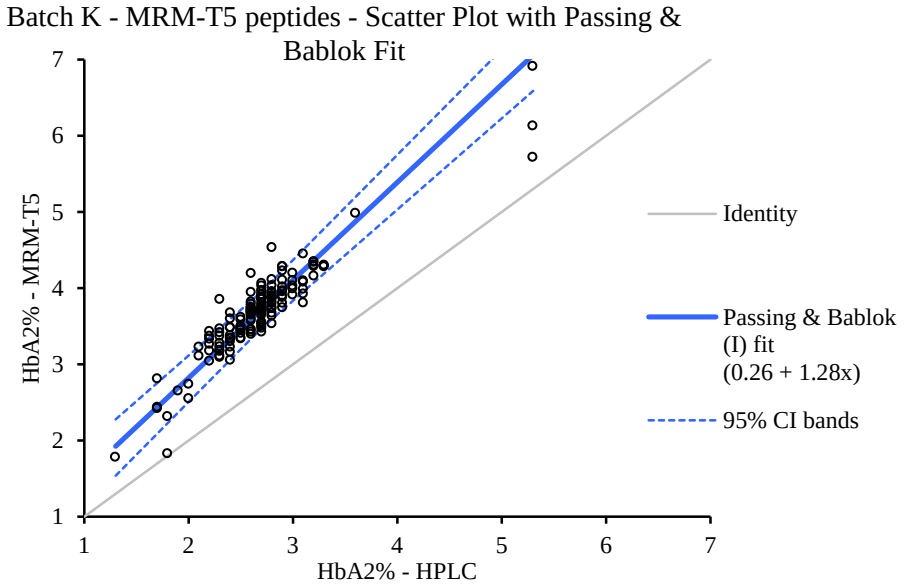
HbA Mass Error - Batch K
HbA Mass Error (Da)

HbB Mass Error - Batch K
HbB Mass Error (Da)

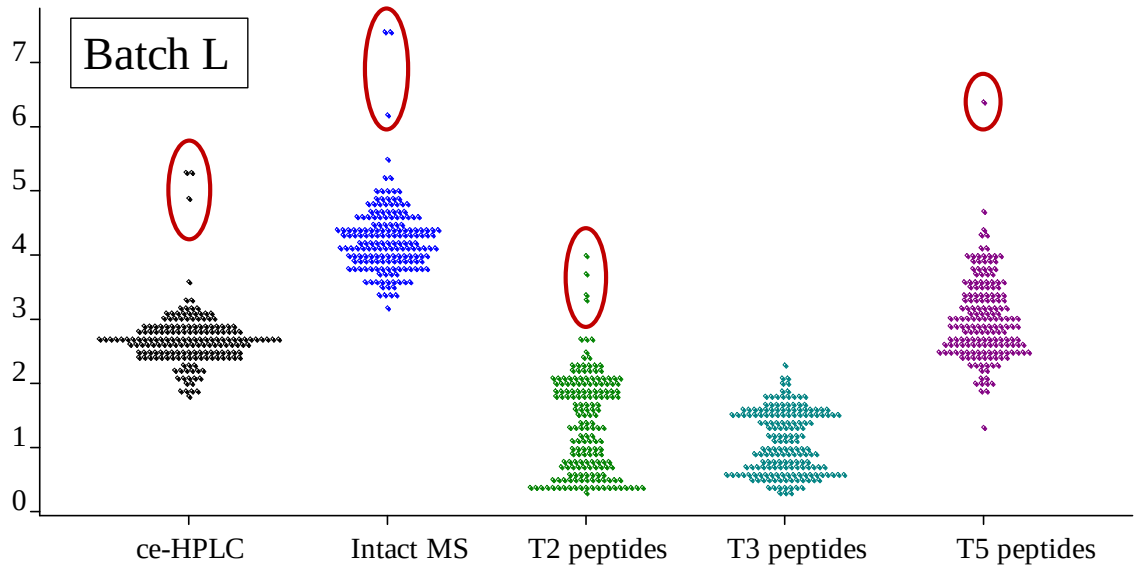

Batch K - Intact MS - Difference Plot
Difference (Intact-MS -HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.39)
95% Limits of agreement (0.72 to 2.05)

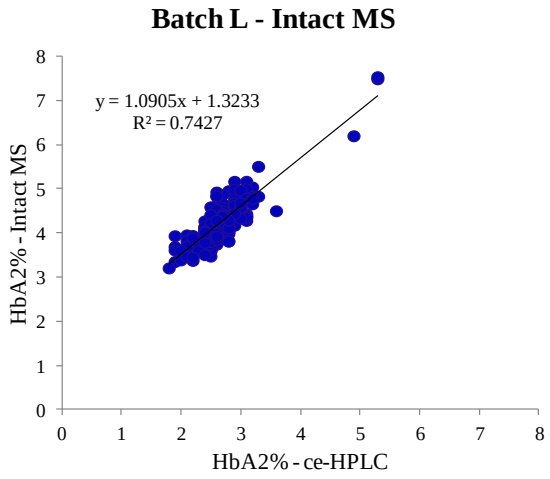

Batch K - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% -Intact MS
HbA2% -HPLC
Identity
Passing & Bablok (I) fit (0.24 + 1.43x)
95% CI bands

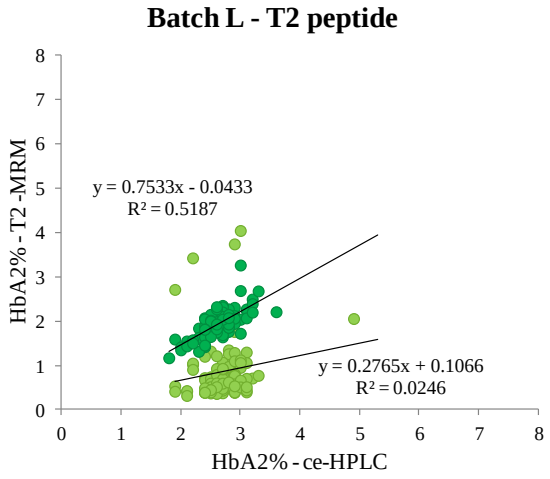

Batch K - MRM-T5 peptides - Difference Plot
Difference (MRM-T5 - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.02)
95% Limits of agreement (0.60 to 1.45)

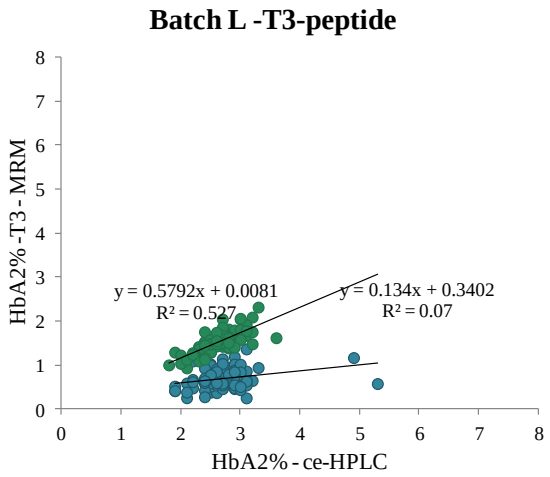

Batch K - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2% - MRM-T5
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (0.26 + 1.28x)
95% CI bands

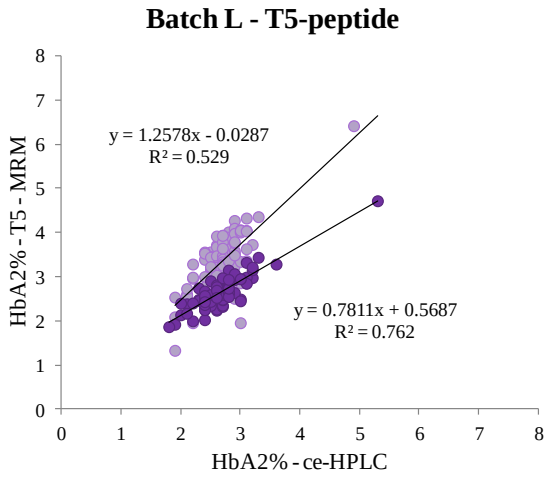
Figures for Batch L

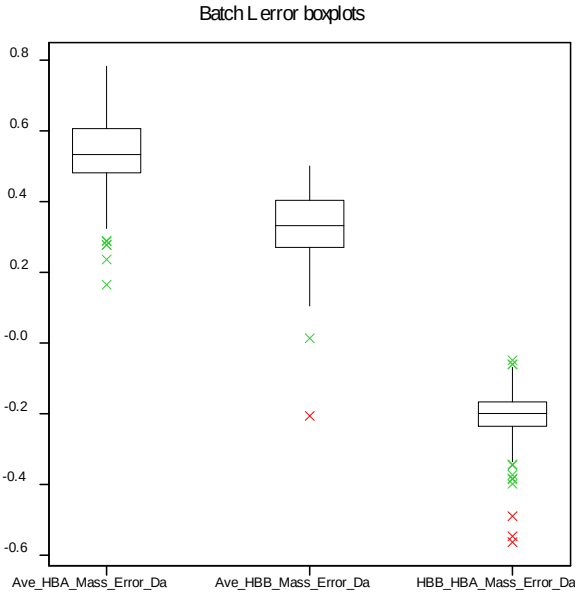


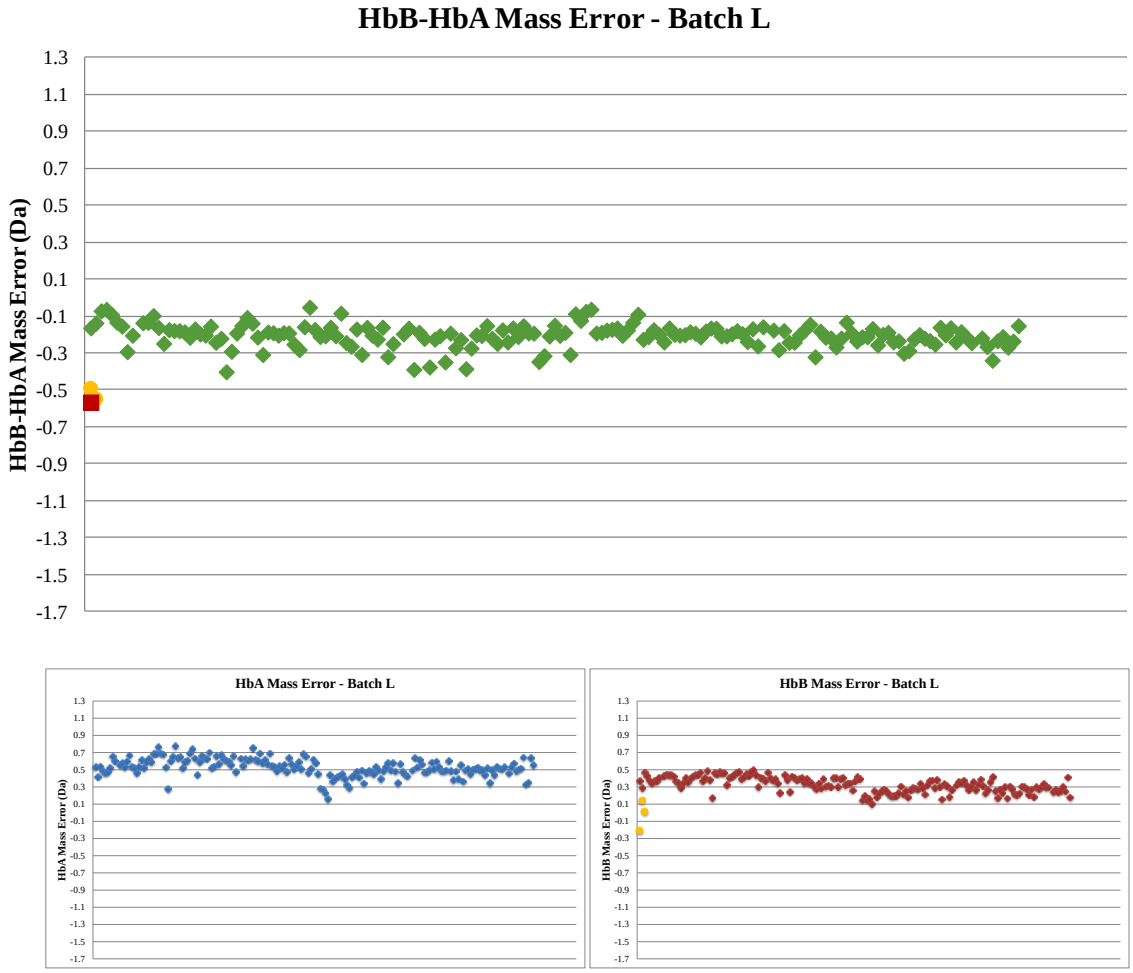






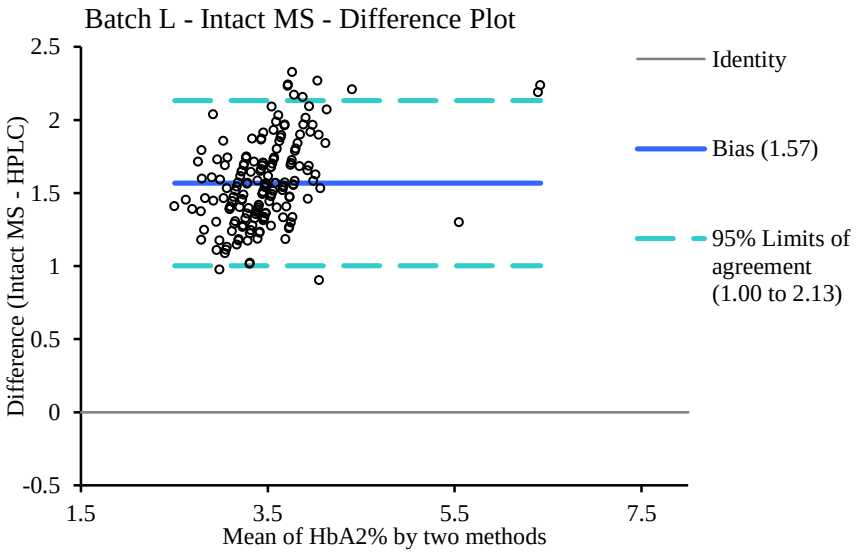


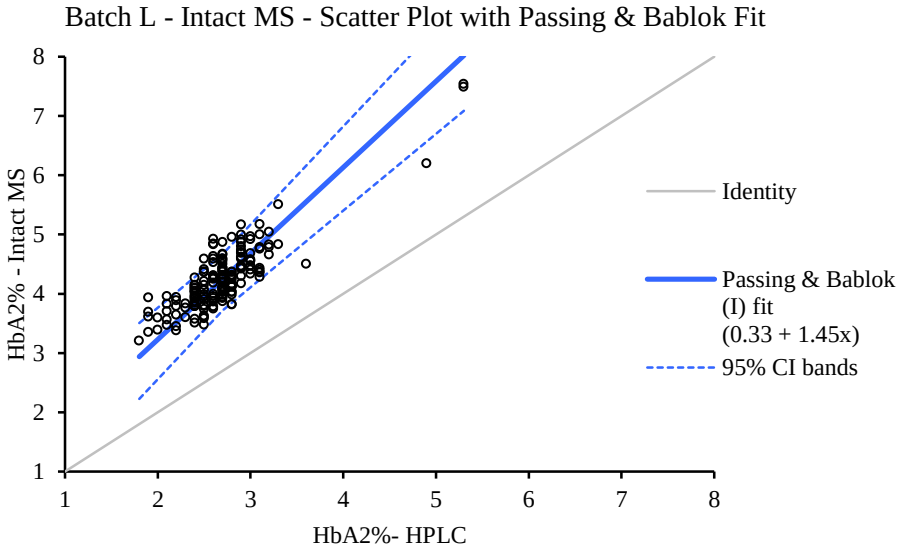

Batch L error boxplots
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

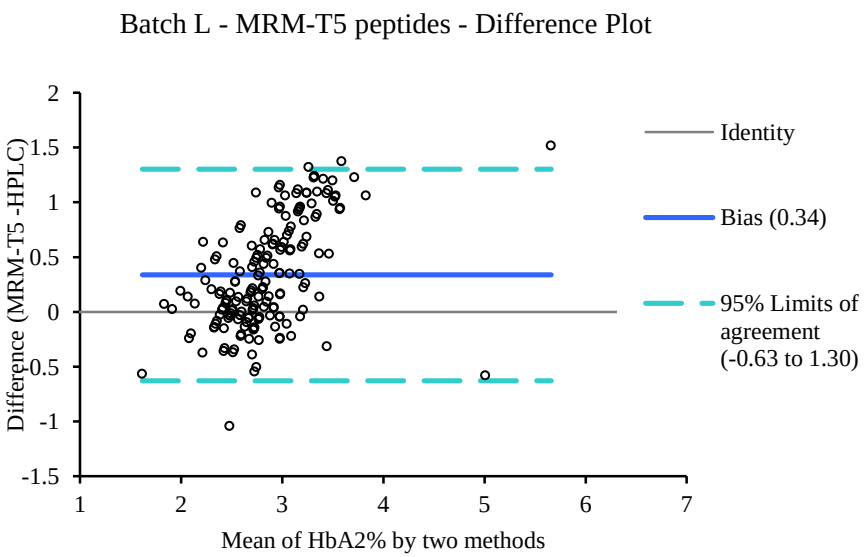

HbB-HbA Mass Error - Batch L
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

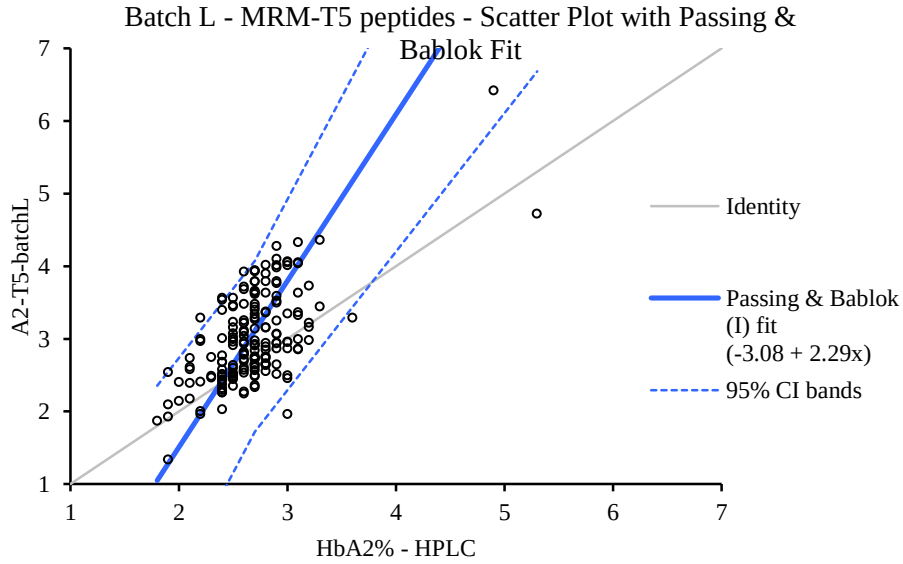
HbA Mass Error - Batch L
HbA Mass Error (Da)

HbB Mass Error - Batch L
HbB Mass Error (Da)

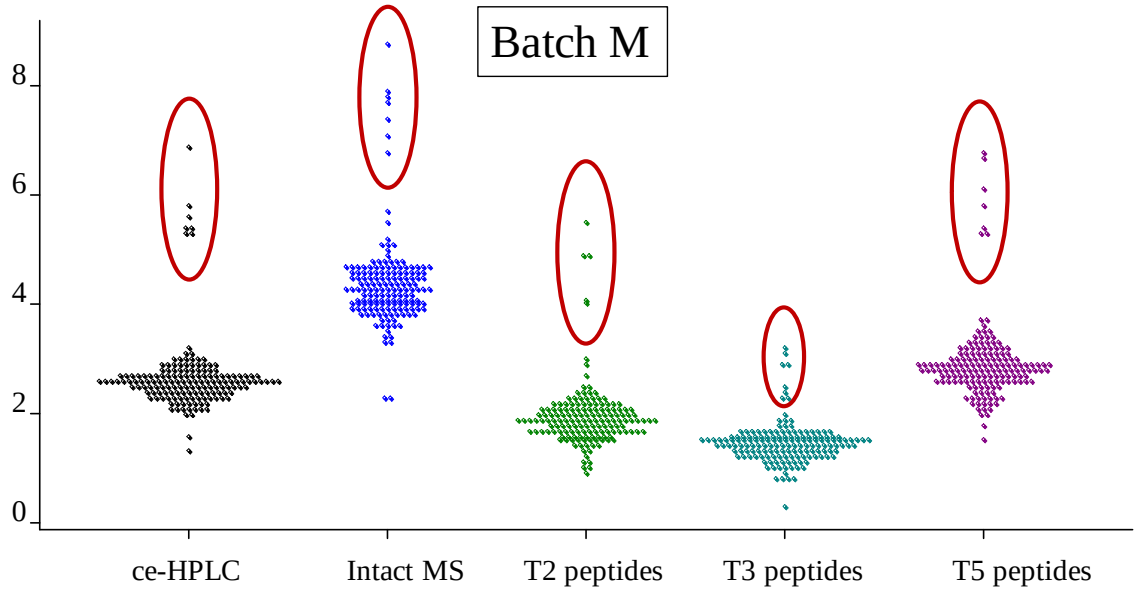

Batch L - Intact MS - Difference Plot
Difference (Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.57)
95% Limits of agreement (1.00 to 2.13)

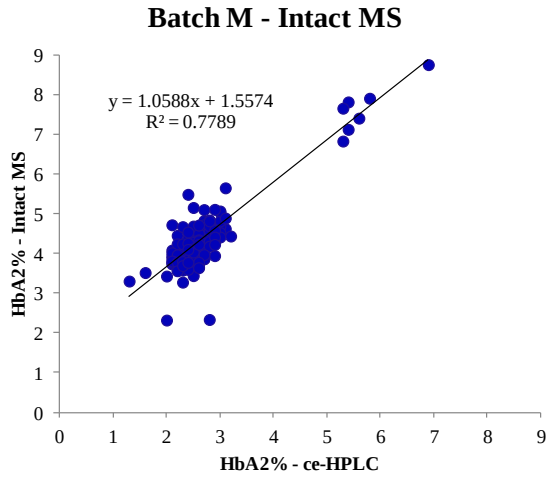

Batch L - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% - Intact MS
HbA2%- HPLC
Identity
Passing & Bablok (I) fit (0.33 + 1.45x)
95% CI bands

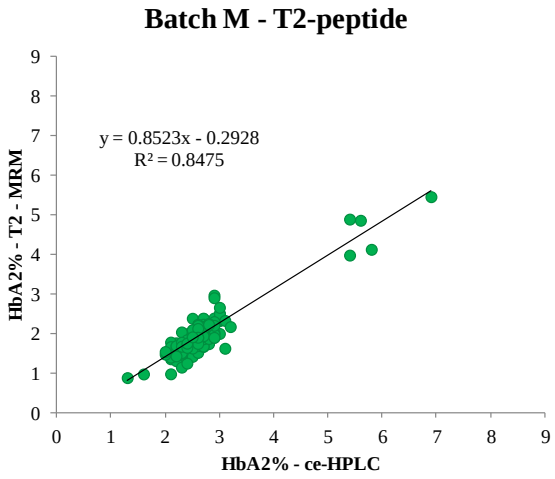

Batch L - MRM-T5 peptides - Difference Plot
Difference (MRM-T5 -HPLC)
Mean of HbA2% by two methods
Identity
Bias (0.34)
95% Limits of agreement (-0.63 to 1.30)

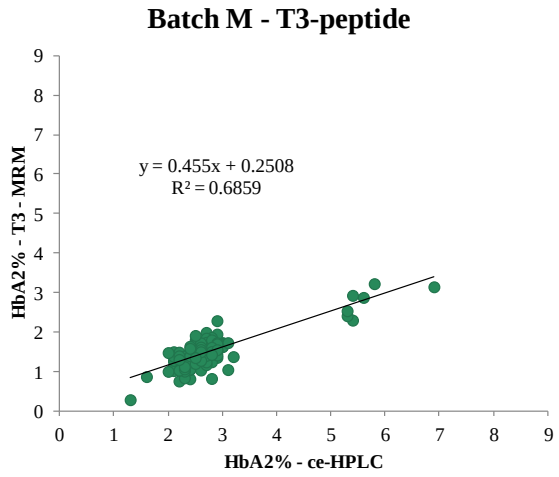

Batch L - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
A2-T5-batchL
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (-3.08 + 2.29x)
95% CI bands

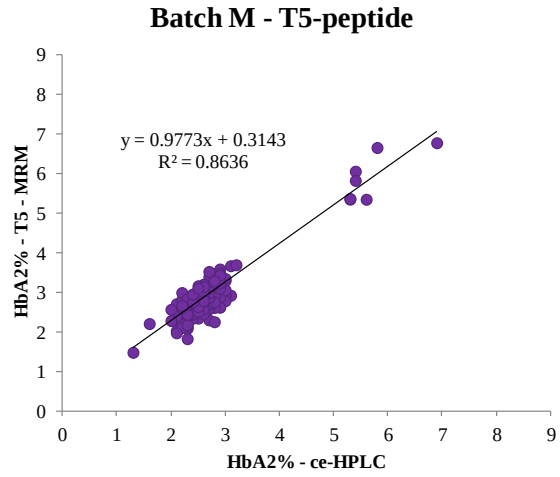
Figures for Batch M

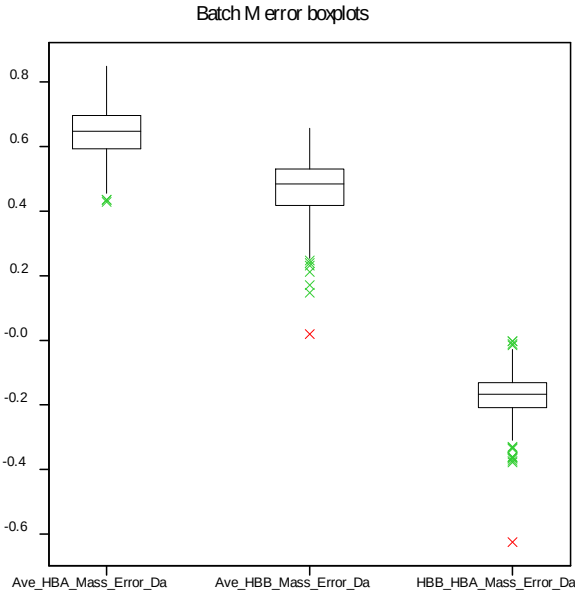


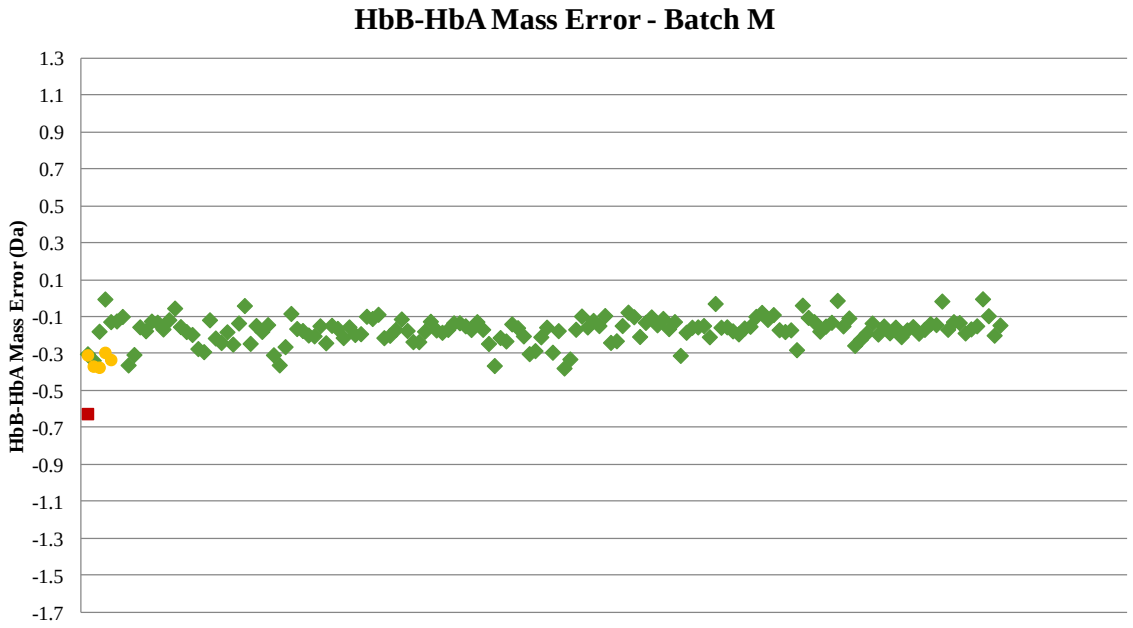


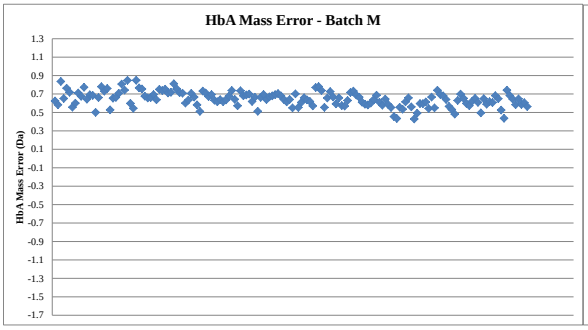


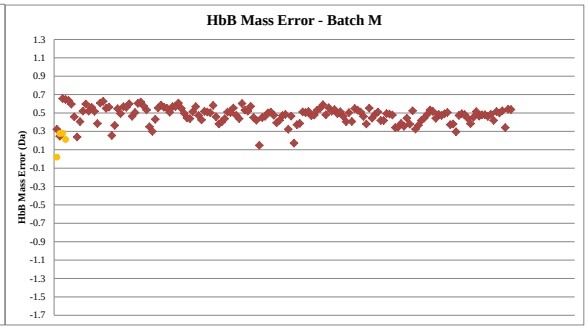


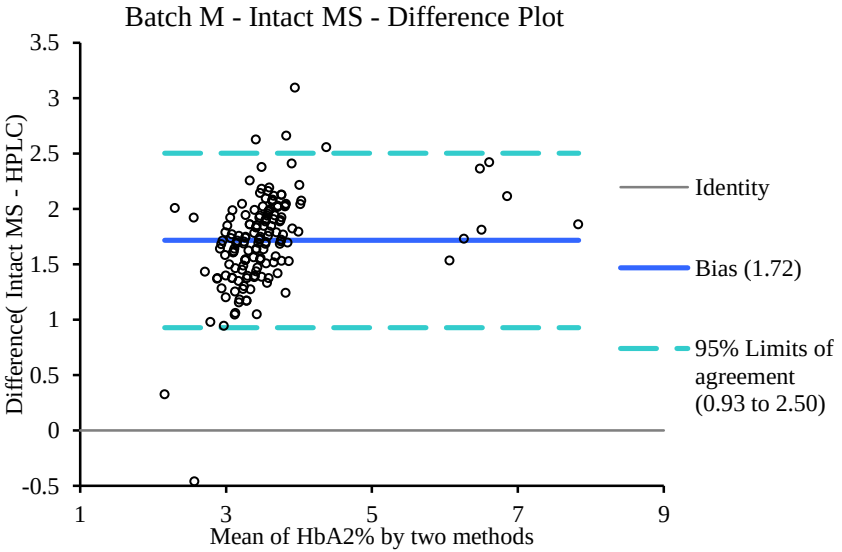


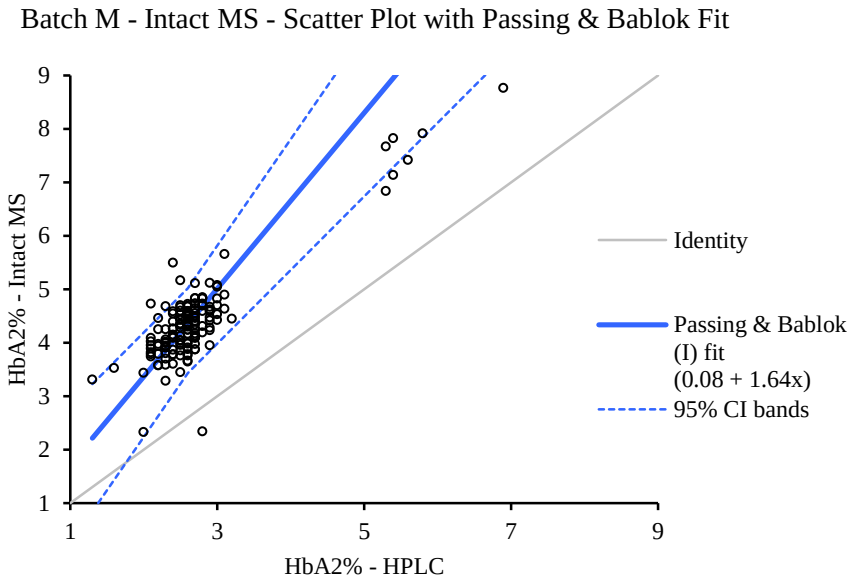

Batch M error boxplots
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
Ave_HBA_Mass_Error_Da
Ave_HBB_Mass_Error_Da
HBB_HBA_Mass_Error_Da

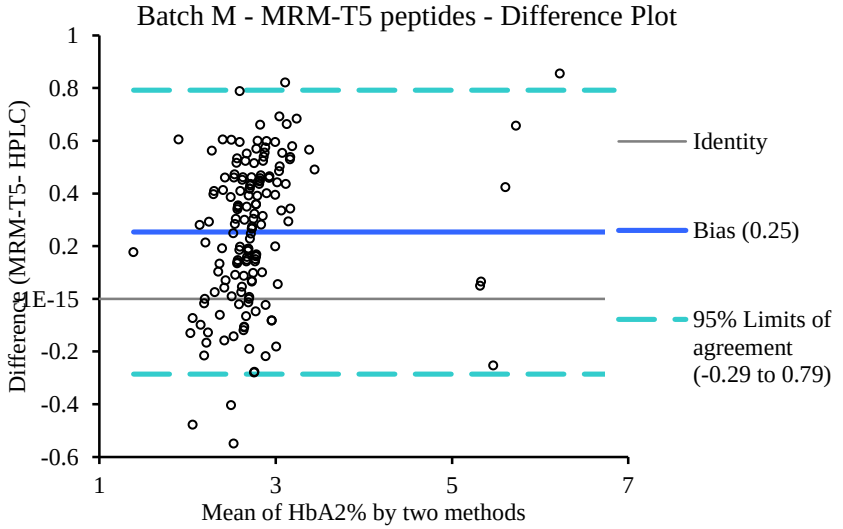

HbB-HbA Mass Error - Batch M
HbB-HbA Mass Error (Da)
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
-1.3
-1.5
-1.7

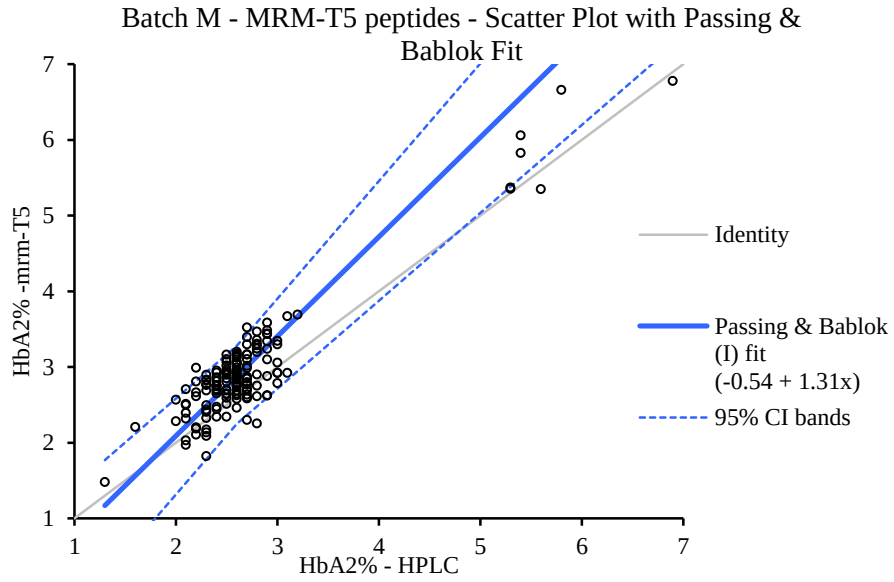

HbA Mass Error - Batch M
HbA Mass Error (Da)


HbB Mass Error - Batch M
HbB Mass Error (Da)


Batch M - Intact MS - Difference Plot
Difference(Intact MS - HPLC)
Mean of HbA2% by two methods
Identity
Bias (1.72)
95% Limits of agreement (0.93 to 2.50)


Batch M - Intact MS - Scatter Plot with Passing & Bablok Fit
HbA2% - Intact MS
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (0.08 + 1.64x)
95% CI bands


Batch M - MRM-T5 peptides - Difference Plot
Difference (MRM-T5- HPLC)
Mean of HbA2% by two methods
Identity
Bias (0.25)
95% Limits of agreement (-0.29 to 0.79)


Batch M - MRM-T5 peptides - Scatter Plot with Passing & Bablok Fit
HbA2% -mrm-T5
HbA2% - HPLC
Identity
Passing & Bablok (I) fit (-0.54 + 1.31x)
95% CI bands

Appendix 2

	Hospital number	Hospital comment	Our number	δ-chain %	Sickle β %	Comment
1	18626847	AS	B/2	6.5	33.8	
2	18636282	AS	B/7	5.4	40	
3	18942346	AS	C/01	5.8	40.2	
4	18968368	AS	C/103	7.6	34.3	
5	18987758	AS	C/154	6.6	40.4	
6	19000645	AS	D/86	6.6	34.0	
7	18997141	AS	D/87	7.7	36.6	
8	18965398	AS	D/90	7.4	34.4	
9	19026666	AS	D/96	6.9	34.4	
10	18984306	AS	D/108	6.9	22.4	
11	19023812	AS	E/20	8.6	40.7	
12	19088085	AS	E/197	6.4	32.5	
13	19104439	AS	E/240	7.4	35.3	
14	19063471	AS	E/257	8.1	29.0	
15	19143309	AS	F/50	6.72	41.3	
16	19190287	AS	F/116	6.35	38.9	
17	19236403	AS	G/32	5.7	42.7	
18	19243173	AS	G/33	7.2	41.1	
19	19256132	AS	G/93	7.9	44.0	
20	19160116	AS	H/41	7.1	32.9	
21	19167290	AS	H/42	8.3	37.8	
22	19177623	AS	H/43	7.1	27.8	
23	19210745	AS	H/44	8.3	40.8	
24	19243309	AS	H/47	-	99.3	
25	19294788	AS	H/70	7.2	40.4	
26	19335077	AS	I/07	6.9	41.0	
27	19349335	AS	I/09	6.7	40.4	
28	19366797	AS	I/39	6.6	34.7	
29	19370399	AS	I/53	6.5	40.4	
30	19401960	AS	J/16	6.2	40.7	
31	19353426	AS	J/41	5.8	35.9	
32	19329213	AS	J/42	5.2	40.6	
33	19335932	AS	J/43	6.0	36.4	
34	19429327	AS	J/45	5.9	42.7	
35	19442432	AS	J/66	6.4	40.0	
36	19529354	AS	K/42	7.0	41.5	
37	19452101	AS	K/43	6.6	41.3	
38	19531007	AS	K/45	6.8	44.0	
39	19455394	AS	K/46	5.7	32.3	
40	19550500	AS	K/50	6.4	31.0	
41	19551541	AS	K/57	6.4	30.0	
42	19554540	AS	K/78	6.8	40.9	
43	19606434	AS	L/49	7.6	42.6	
44	19644581	AS	L/130	7	37.8	
45	19675732	AS	M/50	6.49	40.4	
46	19681412	AS	M/51	7.5	39.6	
47	19702509	AS	M/122	6.5	34.3	
48	19100576	F+S	F/128	-	27.1	Calculated to HbF

Table A.1. Calculated sickle β-chain percentages for the samples classed as heterozygous sickle carriers.

	Hospital number	Hospital comment	Our number	δ-chain %	Sickle β %	Comment
1	18963649	SC	C/99	7.7	48.3	
2	18925552	SC	D/95	9.3	52.9	
3	19395725	SC	J/40	9.7	52.9	
4	18680594	SS	B/127	19.8	84.8	high, but not 100%
5	**19051166**	**SS**	**D/92**	**4.4**	**9.5**	**low, potentially transfused**
6	19052015	SS	E/34	-	100	
7	19089366	SS	E/146	19.2	80.7	high, but not 100%
8	19099276	SS	E/180	-	94.8	high, but not 100%
9	19104272	SS	E/239	7.0	46.4	Potentially heterozygous
10	19217134	SS	F/229	-	100	
11	19271790	SS	H/39	-	100	
12	19272934	SS	H/40	-	100	
13	19162734	SS	H/143	-	100	
14	19348189	SS	I/01	-	100	
15	19364122	SS	I/74	-	100	
16	19444782	SS	J/48	7.3	55.1	Potentially heterozygous
17	19504631	SS	J/87	-	100	
18	19513031	SS	J/106	6.4	43.3	Potentially heterozygous
19	**19516260**	**SS**	**J/116**	**4.1**	**6.5**	**low, potentially transfused**
20	19669274	SS	M/42	6.4	33.9	
21	19729121	SS	M/154	5.4	37.7	
22	19063417	TRANSFUSED SS	F/129	5.2	32.6	
23	19217967	TRANSFUSED SS	F/230	6.5	41.9	
24	19218036	TRANSFUSED SS	F/231	4.1	3.3	very low
25	19249866	TRANSFUSED SS	G/92	6.2	44.4	

Table A.2. Calculated sickle β-chain percentages for the samples classed as heterozygous sickle carriers.

	Hospital number	Hospital comment	Our number	T2	T3	T5	HbS %	Hb C	Comment
1	18626847	AS	**B/2**	**2.4**	**2.1**	**3.7**	**20.7**		
2	18636282	AS	**B/7**	**2.1**	**1.6**	**3.1**	**27.0**		
3	18942346	AS	**C/01**	**3.5**	**2.4**	**3.2**	**38.9**		
4	18968368	AS	**C/103**	**-**	**2.8**	**2.9**	**19.9**		
5	18987758	AS	**C/154**	**-**	**2.4**	**2.5**	**27.2**		
6	19000645	AS	**D/86**	**87.7**	**1.1**	**4.2**	**11.5**		
7	18997141	AS	**D/87**	**40.4**	**1.3**	**4.8**	**13.1**		
8	18965398	AS	**D/90**	**52.6**	**1.4**	**4.7**	**15.1**		
9	19026666	AS	**D/96**	**18.8**	**1.4**	**4.5**	**12.9**		
10	18984306	AS	**D/108**	**77.5**	**1.4**	**5.3**	**5.0**		
11	19023812	AS	**E/20**	**3.3**	**2.9**	**3.4**	52.7		
12	19088085	AS	**E/197**	2.5	2.1	2.9	18.8		
13	19104439	AS	**E/240**	3.7	2.8	3.0	27.8		
14	19063471	AS	**E/257**	4.0	3.3	3.8	16.5	-	
15	19143309	AS	**F/50**	1.0	2.4	4.1	15.7		
16	19190287	AS	**F/116**	1.1	1.6	3.0	20.4		
17	19236403	AS	**G/32**	1.4	2.2	3.3	19.4		
18	19243173	AS	**G/33**	1.5	2.3	3.8	17.7		
19	19256132	AS	**G/93**	0.9	1.8	3.1	22.4		
20	19160116	AS	**H/41**	1.1	2.3	4.7	5.6		
21	19167290	AS	**H/42**	0.8	2.5	4.8	14.6		
22	19177623	AS	**H/43**	1.3	2.3	5.2	3.7		
23	19210745	AS	**H/44**	1.4	2.6	5.0	21.0		
24	19243309	AS	**H/47**	5.5	2.4	5.2	17.1		
25	19294788	AS	**H/70**	1.1	2.0	3.9	12.9		
26	19335077	AS	**I/07**	3.9	2.9	2.6	41.6		
27	19349335	AS	**I/09**	3.9	3.1	2.7	42.0		
28	19366797	AS	**I/39**	3.8	2.9	2.7	31.8		
29	19370399	AS	**I/53**	3.9	2.6	2.7	39.7		
30	19401960	AS	**J/16**	0.2	1.9	3.4	8.9		
31	19353426	AS	**J/41**	0.6	1.9	4.2	9.2	0.0	
32	19329213	AS	**J/42**	0.3	1.7	3.3	12.4	0.1	
33	19335932	AS	**J/43**	0.3	2.1	4.0	5.5	0.0	
34	19429327	AS	**J/45**	0.7	1.9	4.5	19.4	0.0	
35	19442432	AS	**J/66**	0.9	2.2	4.3	12.2		
36	19529354	AS	**K/42**	1.2	1.1	5.1	8.3	-	Sickle values are really low
37	19452101	AS	**K/43**	1.1	1.1	4.1	5.0	-	peptides are present, but quantitation is not satisfying
38	19531007	AS	**K/45**	1.5	1.0	4.3	8.7	-	
39	19455394	AS	**K/46**	0.6	0.3	3.7	1.3	-	
40	19550500	AS	**K/50**	1.3	1.1	4.4	1.6	-	
41	19551541	AS	**K/57**	1.2	1.1	4.5	1.4	-	
42	19554540	AS	**K/78**	1.6	1.2	4.8	1.7	-	
43	19606434	AS	**L/49**	1.2	1.3	4.2	0.2		Problem with digestion
44	19644581	AS	**L/130**	2.4	1.7	3.3	14.3		
45	19675732	AS	**M/50**	1.8	1.6	2.8	29.0		
46	19681412	AS	**M/51**	1.5	1.2	2.8	24.8		
47	19702509	AS	**M/122**	2.5	1.6	4.7	13.2		
48	19100576	F+S	**F/128**	8.9	2.7	8.9	100		baby

Table A.3. Calculated sickle β-chain percentages for the samples classed as heterozygous sickle carriers based on MRM method results.

	Hospital number	Hospital comment	Our number	MRM δ-chain peptide ratios %			Varian t Sickle %	Hb C	Comment
				T2	T3	T5			
1	18963649	SC	**C/99**	-	1.9	2.4	99.3	99.4	* MRM recalculated 41.1
2	18925552	SC	**D/95**	76.9	1.0	4.7	99.6	100	*MRM recalcalculated 10.9
3	19395725	SC	**J/40**	1.1	2.3	5.3	99.1	99.9	* MRM recalculated 12.7% Sickle and 87.3% HbC
4	18680594	SS	**B/127**	3.0	1.6	3.7	97.9		
5	19051166	SS	**D/92**	77.7	0.5	3.9	0.2		
6	19052015	SS	**E/34**	3.9	3.2	5.3	99.9		
7	19089366	SS	**E/146**	3.6	2.3	3.0	89.8		
8	19099276	SS	**E/180**	4.4	2.8	3.9	99.8		
9	19104272	SS	**E/239**	2.8	2.0	2.4	39.1		
10	19217134	SS	**F/229**	3.3	2.1	4.0	100.0		
11	19271790	SS	**H/39**	2.3	2.0	5.1	100.0		
12	19272934	SS	**H/40**	1.9	2.1	4.5	100.0		
13	19162734	SS	**H/143**	3.8	2.1	4.0	100.0		
14	19348189	SS	**I/01**	3.8	2.8	2.4	99.7		
15	19364122	SS	**I/74**	3.8	2.6	3.1	99.8		
16	19444782	SS	**J/48**	0.6	1.8	3.5	44.5		
17	19504631	SS	**J/87**	2.0	3.4	6.7	99.8		
18	19513031	SS	**J/106**	1.8	1.6	3.9	17.3	0.0	
19	19516260	SS	**J/116**	1.8	1.6	3.8	0.0	0.0	no peaks at all for sickle peptide
20	19669274	SS	**M/42**	2.3	1.0	2.8	0.8		Low intensity peak
21	19729121	SS	**M/154**	1.9	1.0	3.2	13.8		
22	19063417	TRANSFUSED SS	**F/129**	1.8	1.5	3.1	13.6		
23	19217967	TRANSFUSED SS	**F/230**	2.0	1.8	3.7	20.8		
24	19218036	TRANSFUSED SS	**F/231**	1.8	1.7	3.0	0.0		
25	19249866	TRANSFUSED SS	**G/92**	2.4	1.5	2.7	13.7		

Table A.4. Calculated sickle β-chain percentages for the samples classed as heterozygous sickle carriers.

Appendix 3

Hb J Baltimore T2-T3 peptide
Average Mass = 2286.5276, Monoisotopic Mass = 2285.1651
Residues: 1-22
N-Terminus = H, C-Terminus = OH
Fragment ions: Monoisotopic/Average (1750) m/z ratios with 1 positive charge(s).

b	88.0	159.1	258.1	359.2	430.2	543.3	729.4	844.4	972.5	1071.6	1185.6	1284.7	1399.7	1528.8	1627.8	1684.9	1741.9	1872.0	1943.1	2056.3	2113.3	-
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
	Ser	Ala	Val	Thr	Ala	Leu	Trp	Asp	Lys	Val	Asn	Val	Asp	Glu	Val	Gly	Gly	Glu	Ala	Leu	Gly	Arg
	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
y''	-	2200.5	2129.4	2030.2	1929.1	1858.1	1743.9	1557.8	1442.8	1314.7	1215.6	1101.6	1002.5	887.5	758.4	659.3	602.3	545.3	416.3	345.2	232.1	175.1

Figure A.1. Predicted fragments for the HbJ Baltimore βT2-T3 peptide

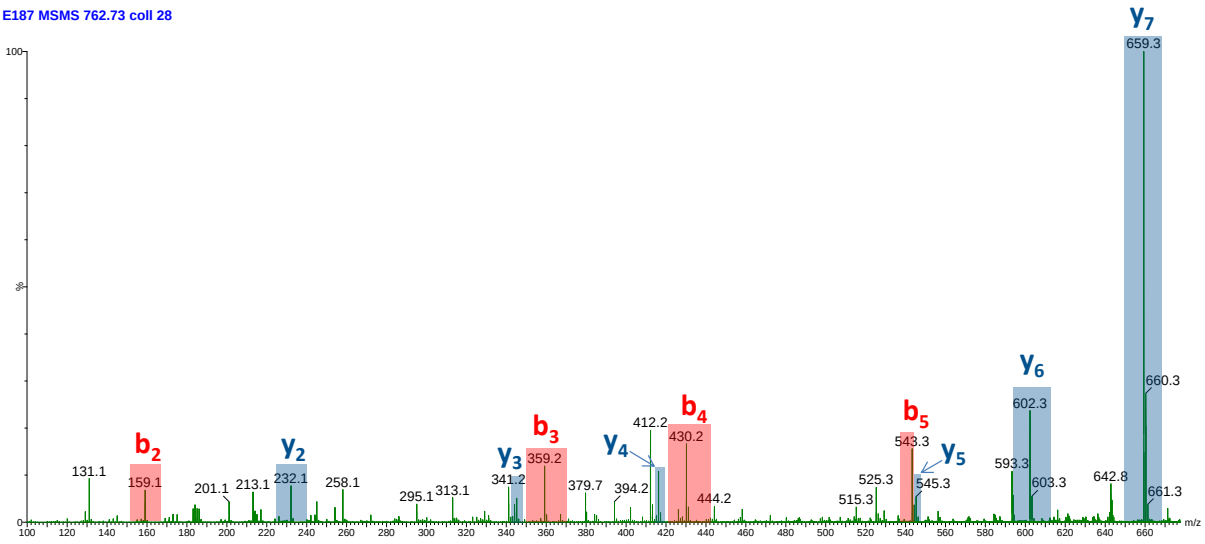


Figure A.2. MS/MS spectrum of the HbJ Baltimore βT2-T3 peptide *m/z* 100-670

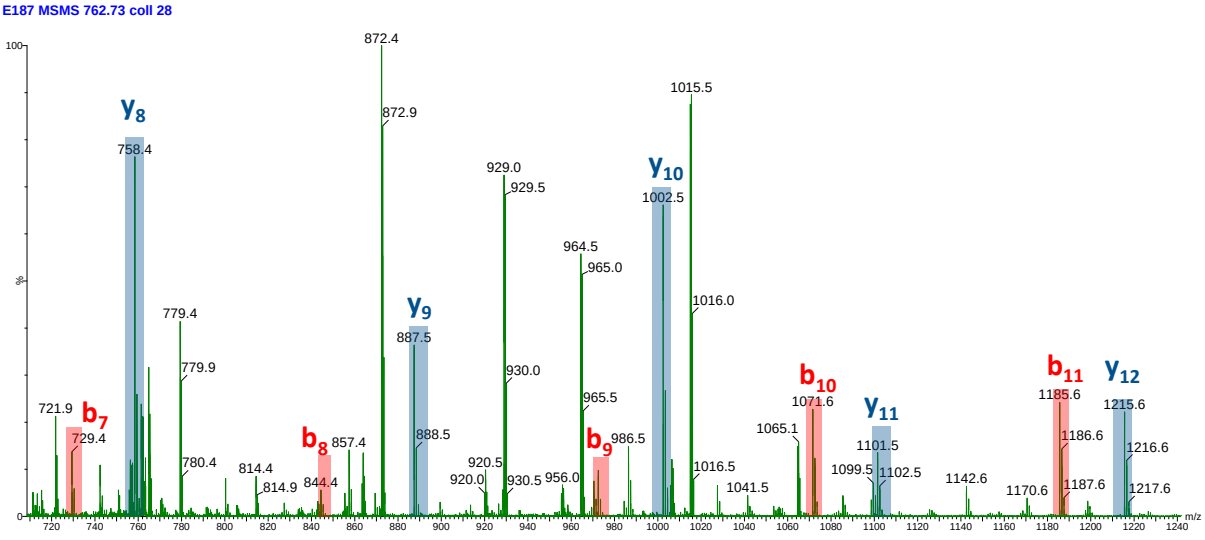


Figure A.3. MS/MS spectrum of the HbJ Baltimore βT2-T3 peptide m/z 710-1240

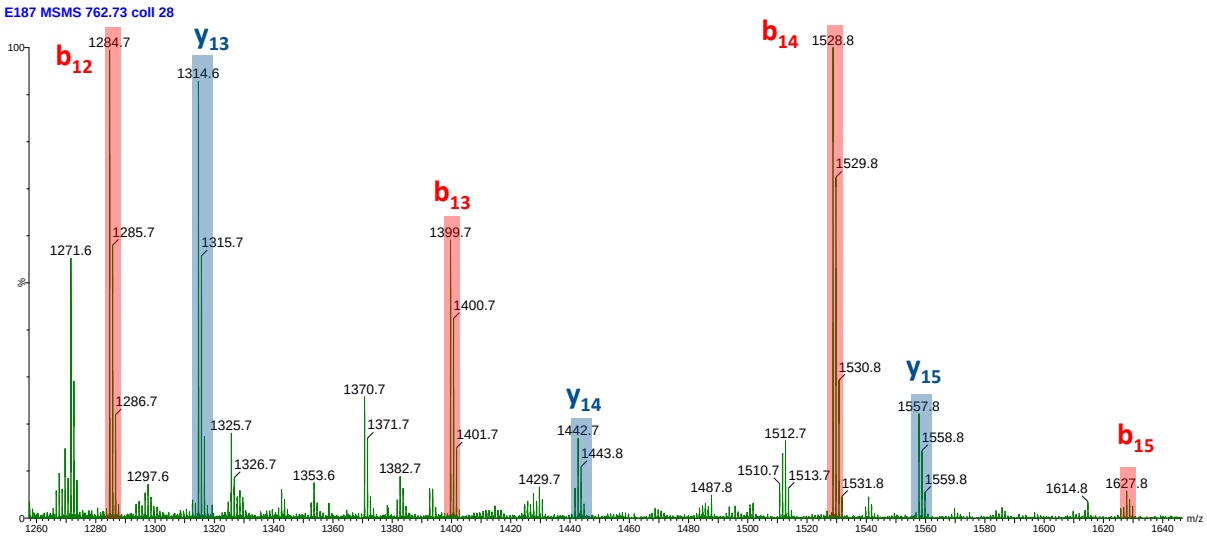


Figure A.4. MS/MS spectrum of the HbJ Baltimore βT2-T3 peptide *m/z* 1260-1640

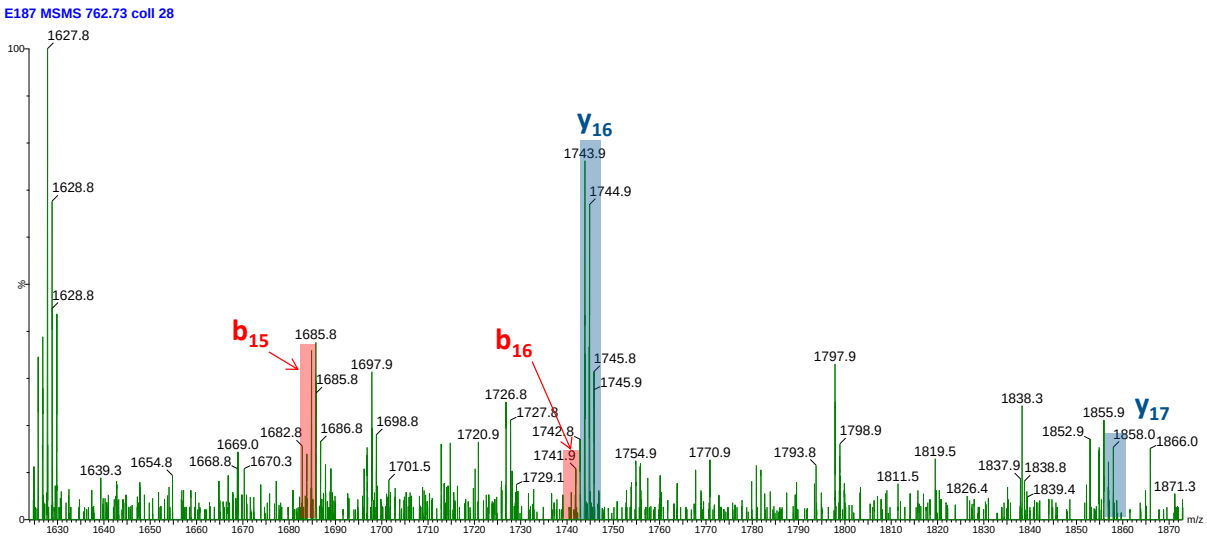


Figure A.5. MS/MS spectrum of the HbJ Baltimore βT2-T3 peptide *m/z* 1620-1870

Appendix 4

Haemoglobinopathy Mutant Identification Report

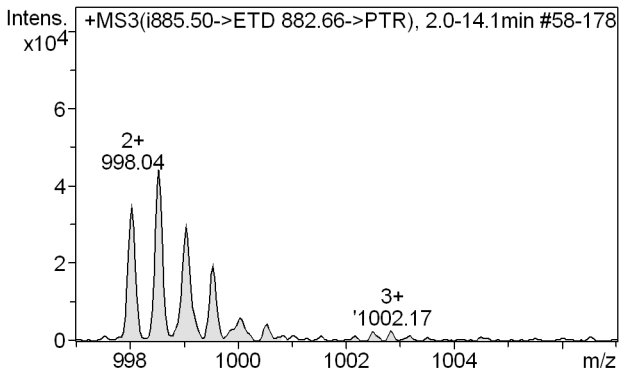
Analysis Info

		Acquisition Date	26/10/2012 12:05:31
Analysis Name	G:\Bruker_Backup\Data\Manual_acquisition_clinical_trial\BatchA_251012\A08 1 in 1000002.d		
Method	dual isol 885 882 etd 20ms ptr 80ms_final.m	Operator	BUK
Sample Name	A08 1 in 10000	Instrument	amaZon speed ETD
Comment			

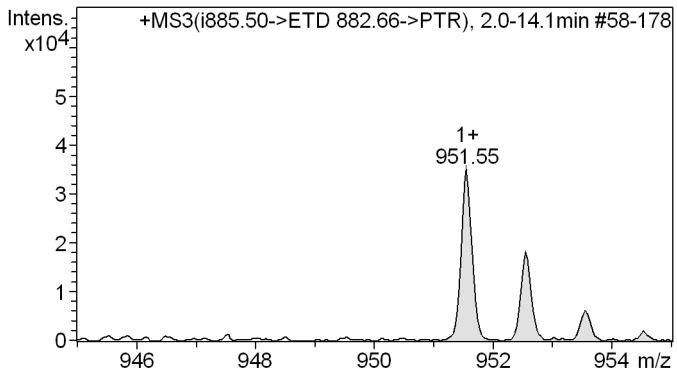
Acquisition Parameter

Ion Source Type	ESI	Scan Begin	250 m/z
Mass Range Mode	Enhanced Resolution	Scan End	3000 m/z
Accumulation Time	99 µs		

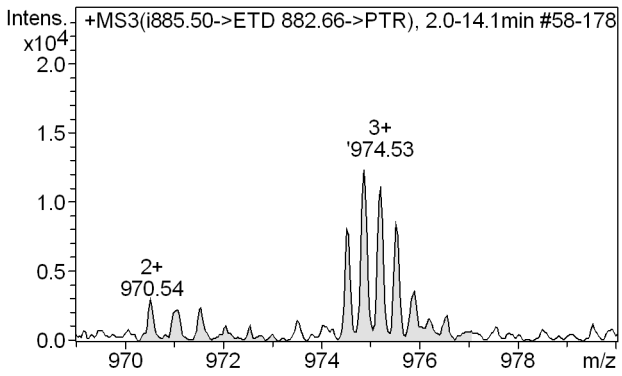
Reporter ion 1002.2 (3+) would indicate Mutant AO1



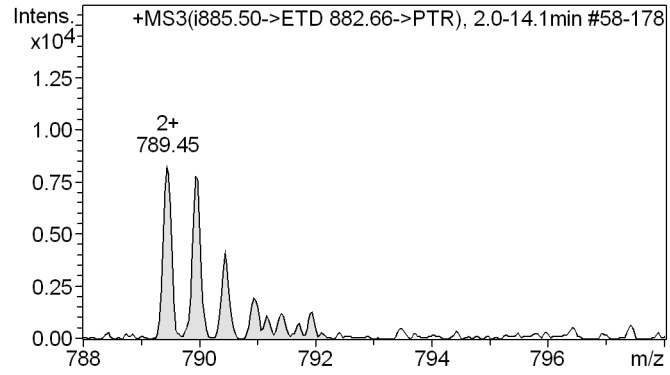
Reporter ion 950.6 (1+) would indicate Mutant AO2



Reporter ion 974.5 (3+) would indicate Mutant AO3

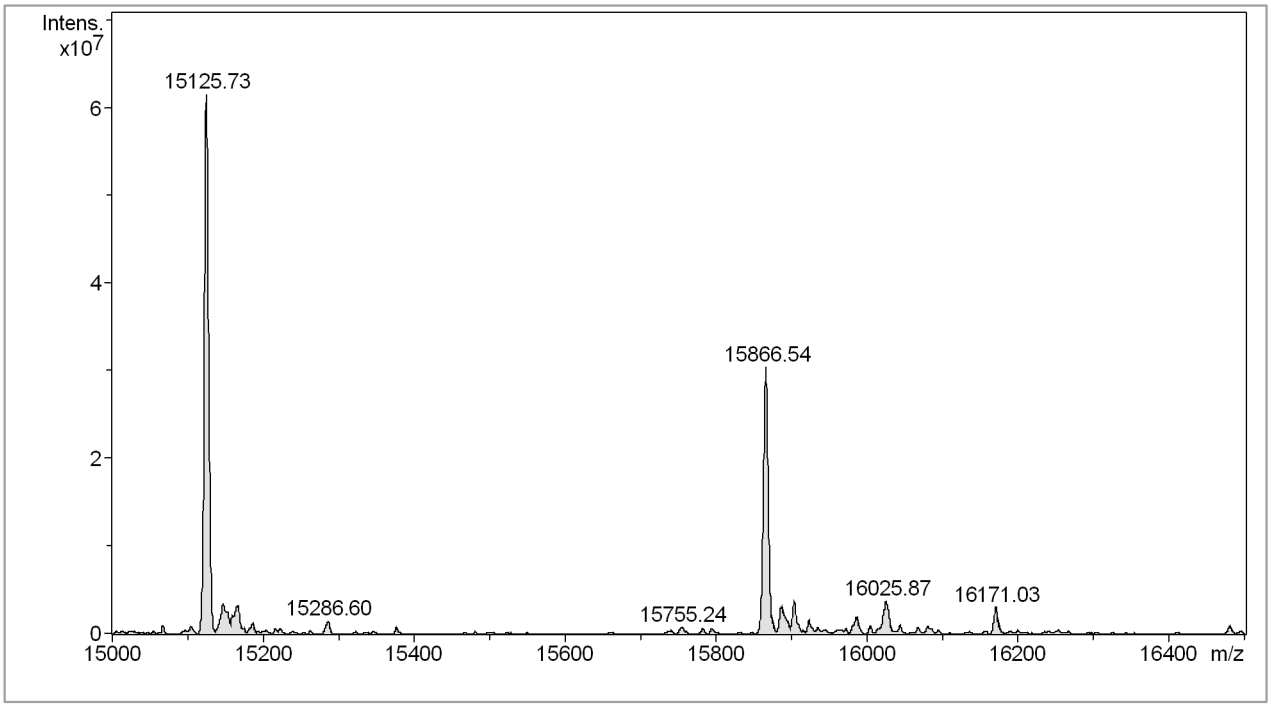


Reporter ion 793.5 (1+) would indicate Mutant AO4



Haemoglobinopathy Mutant Identification Report

Full Mass Spectrum Intact Protein Chains



No.	m/z [Da]	Intensity	Relative Intensity	Theoretical average mass	delta m/z [Da]
1.	15125.73	61472104	100.00%	15126.202	-0.48
2.	15866.54	30424412	49.49%		
3.	16025.87	3811956	6.20%		
4.	15904.06	3786271	6.16%		
5.	15148.17	3473769	5.65%		
6.	15166.52	3307016	5.38%		
7.	15887.26	3086220	5.02%		
8.	16171.03	3070616	5.00%		
9.	15986.21	1990329	3.24%		
10.	10334.38	1575422	2.56%		
11.	15923.73	1567764	2.55%		
12.	15286.60	1525320	2.48%		
13.	12135.73	1419728	2.31%		
14.	18289.75	1265773	2.06%		
15.	14270.04	1185072	1.93%		
16.	15186.46	1177115	1.91%		
17.	14196.41	1095055	1.78%		
18.	18836.46	1049026	1.71%		
19.	16044.23	1035107	1.68%		
20.	19468.22	1007835	1.64%		