Rapid microwave-assisted chemical cleavage - mass spectrometric method for the identification of hemoglobin variants in blood

ARTICLE *in* ANALYTICAL CHEMISTRY · OCTOBER 2010

Impact Factor: 5.64 · DOI: 10.1021/ac1017106 · Source: PubMed

6 AUTHORS, INCLUDING:

Agron Mataj
Monash University (Australia)
3 PUBLICATIONS **7** CITATIONS
SEE PROFILE

Yuanzhong Yang
Monash University (Australia)
27 PUBLICATIONS **323** CITATIONS
SEE PROFILE

Reinhard Ingemar Boysen
Monash University (Australia)
102 PUBLICATIONS **947** CITATIONS
SEE PROFILE

Anal. Chem. **2010,** *82,* 8922–8930

Rapid Microwave-Assisted Chemical Cleavage - Mass Spectrometric Method for the Identification of Hemoglobin Variants in Blood

Asif Alam,[†] Agron Mataj,[†] Yuanzhong Yang,[†] Reinhard I. Boysen,[†] Donald K. Bowden,[†,‡] and Milton T. W. Hearn*,[†]

ARC Special Research Centre for Green Chemistry, Monash University, Clayton, Victoria, Australia 3800, and Thalassaemia Service, Monash Medical Centre, Melbourne, Australia

A sensitive, rapid analytical method has been developed for the characterization of human hemoglobin disorders with very small volumes (<1 μL) of blood. As an alternative to conventional enzymatic digestion, a site-specific chemical cleavage method has been established using 0.05% formic acid under microwave-irradiation conditions for short time intervals, for example, less than 10 min. Peptide analysis was performed by MALDI TOF MS and capillary liquid chromatographic ESI MS/MS. The cleavage of the hemoglobin chains with formic acid occurred at either side or at both sides (C- and N-terminal) of aspartic acid residues, but preferentially N-terminally. The method has been applied to blood samples from hemoglobin S carrier heterozygotes and hemoglobin S thalassaemia compound heterozygotes with a reduced expression level of hemoglobin S. Both MALDI TOF MS and ESI MS/MS analysis allowed the identification of the hemoglobin S "signature" peptide. This alternative method of sample preparation is compatible with MS techniques and is expected to significantly contribute to the further development of rapid, robust, reproducible and sensitive analytical methods in proteomics and biomedical diagnostics where protein variant characterization is a crucial factor for biomarker discovery and disease identification.

In clinical diagnosis, mass spectrometry is increasingly finding a key role in the characterization of disease-specific proteins, protein variants of clinical relevance[1] or post-translational modified proteins as biomarkers of a disease.[2] In the present investigation, in order to evaluate an alternative approach to obtain "signature" peptides derived from disease-specific proteins and other protein biomarkers, hemoglobin A (HbA) and its structural variants were chosen as a model system due to their clinical significance. Hemoglobins, the main cytoplasmic proteins of red blood cells, contain a haem and globin moiety. Each globin moiety consists of two identical α- or α-like and two identical non-α-like chains. The haem moiety, a ferroprotoporphyrin IX, is the prosthetic group of Hb with a molecular mass of 616.1768 amu. In normal human adults ≈97% of the major Hb is HbA consisting of two α- and two β-chains ($\alpha_2\beta_2$). Each α-chain consists of 141 amino acid residues, whereas the number of amino acid residues in the β-chain is 146. For inherited disorders of Hb, the association of a clinical phenotype with a molecular aberration was recognized as early as 1949 by Pauling in a study describing the different electrophoretic mobility of normal and sickle cell Hb.[3] Subsequently, Hb's have been shown to be subject to enormous structural diversity, with over 1320 α- and β-chain variants described in the literature.[4,5]

Hemoglobinopathies are autosomal recessive disorders of hemoglobin. They are the commonest single-gene disorders in the population and caused by abnormal globin gene expression, which results in either abnormal quantitative production of the globin chains or their structural aberrations.[6] Hemoglobinopathies fall into overlapping groups. The thalassaemias are those conditions that result from imbalanced globin chain production, usually resulting from a reduction or absence of one or more globin chains. They are classified according to the chain that is affected.[7] Common and important are those affecting the α- and the β-chains—the α- and β-thalassaemias. The Hb disorders with structural abnormalities are hemoglobin variants which were first named according to the letters of the alphabet (e.g., HbS), but later by place of discovery (e.g., HbJ Norfolk). Hemoglobin variants are mainly due to single amino-acid substitutions, although a few have elongated or shortened globin chains. All have the general structure $a_2\beta_2$, except for Hb Bart's and HbH, which are γ_4- or β_4-homotetramers, respectively, and which are formed when α-chain production is defective in α-thalassaemia.[7] Among the >1320 Hb variants described in the literature/online databases (http://globin.cse.psu.edu), only three clinically significant Hb variants, HbS, HbE, and HbC, occur in high frequency.[6]

* To whom correspondence should be addressed. Phone: + 61 + 3 + 9905 4547. Fax: + 61 + 3 + 9905 8501. E-mail: milton.hearn@monash.edu.

[†] Monash University.

[‡] Monash Medical Centre.

(1) Jensen, O. N.; Hojrup, P.; Roepstorff, P. *Anal. Biochem.* **1991**, *199*, 175–183.

(2) Chace, D. H. *Chem. Rev.* **2001**, *101*, 445–477.

(3) Pauling, L.; Itano, H. A.; Singer, S. J.; Wells, I. C. *Science* **1949**, *110*, 543–548.

(4) Hardison, R. C.; Chui, D. H. K.; Giardine, B.; Riemer, C.; Patrinos, G. P.; Anagnou, N.; Miller, W.; Wajcman, H. *Hum. Mutat.* **2002**, *19*, 225–233.

(5) Giardine, B.; van Baal, S.; Kaimakis, P.; Riemer, C.; Miller, W.; Samara, M.; Kollia, P.; Anagnou Nicholas, P.; Chui David, H. K.; Wajcman, H.; Hardison Ross, C.; Patrinos George, P. *Hum. Mutat.* **2007**, *28*, 206.

(6) Weatherall, D. J. *Curr. Mol. Med.* **2008**, *8*, 592–599.

(7) Weatherall, D. J. *Nat. Rev. Genet.* **2001**, *2*, 245–255.

10.1021/ac1017106
Published on Web 10/11/2010

The use of MS to explore the structure of Hb and its variants started in the early 1980s, when field desorption (FD) mass spectrometry was first used for the analysis of a complex mixture of tryptic peptides from Hb variants.[8] Since then, among the red blood cell proteins, Hb, its variants and post-translational modifications have been the focus of many mass spectrometric studies. Liquid secondary ion (LSI),[9,10] fast atom bombardment (FAB),[11] and plasma desorption (PD) mass spectrometry have been applied. Although intact Hb analysis with MALDI TOF MS[12−14] and ESI MS[15] can give valuable information on the mass of individual Hb chains, they do not always allow the unambiguous identification of variants.[16] By taking advantage of the high concentration of Hb in blood (150 g/L), the sequence variations of the vast majority of Hb mutants can be directly identified after dilution of whole blood samples by a combination of mass spectrometric analysis of the intact globin chains and proteolytically derived peptides. The bioinformatic identification of the peptide containing the mutation, the so-called signature peptide,[17] can then be achieved utilizing MALDI TOF MS,[18] tandem ESI MS[19] or ion mobility mass spectrometry (IM MS).[20] When high through-put analysis is the aim, the most severe limitation of current sample preparation methods is, however, the long enzymatic digestion procedure required for adequate proteolysis, which may require 4−24 h depending on the enzyme(s) used as well as on the physicochemical properties, diversity and abundance of the target proteins.[21] For these analyses, enzymatic proteolysis is usually carried out using a variety of proteases, most commonly trypsin, a serine protease, or other enzymes such as endoproteinase Lys C, Glu-C, Asp-N or a combination of enzymes with specificity for cleavage of the protein at multiple sequence sites.[21,22] By exploiting these enzymatic specificities, signature peptides of the target protein variants can be produced with sufficient mass/charge value differences from other detectable (prototypic) fragments. Alternatively, chemical cleavage methods can be performed using cyanogen bromide (CNBr), hydrochloric acid (HCl), trifluoroacetic acid (TFA), acetic acid (AA), or formic acid (FA).[23]

Ongoing research into methods[21] for enzymatic proteolysis have contributed to improved sequence coverages and greater site specificity by using a variety of different conditions, including the use *inter alia* of organic-aqueous solutions,[24,25] modified (acetylated) trypsin,[26] immobilized trypsin on beads for on-column proteolysis,[27,28] higher temperatures (up to 60 °C),[29] MS compatible acid labile surfactants (ALS),[30−33] denaturants, or enzymatic proteolysis under microwave conditions.[34−39] Although microwave-assisted chemical cleavage has been explored for the analysis of small acid-soluble spore proteins,[40] viral coat proteins,[41] N-linked and O-linked glycoproteins[42] as well as other complex protein mixtures,[43−45] its potential for Hb variant analysis with chemical cleavage methods has not been investigated in any depth. Recently, the microwave assisted hydrolysis of a sample of a purified HbG Coushatta β chain with 6 M HCl treatment, prior to MALDI TOF MS analysis, has been described[46] with this traditional method of protein hydrolysis leading to extensive degradation of the protein to the constituent amino acids. In the present investigation, we have explored the applicability of a general, and milder, method of microwave-assisted chemical cleavage of proteins using formic acid. This new method was then

(8) Matsuo, T.; Matsuda, H.; Katakuse, I.; Wada, Y.; Fujita, T.; Hayashi, A. *Biomed. Mass Spectrom.* **1981**, *8*, 25–30.
(9) Katakuse, I.; Ichihara, T.; Nakabushi, H.; Matsuo, T.; Matsuda, H.; Wada, Y.; Hayashi, A. *Biomed. Mass Spectrom.* **1984**, *11*, 386–391.
(10) Wada, Y.; Matsuo, T.; Sakurai, T. *Mass Spectrom. Rev.* **1989**, *8*, 379–434.
(11) Pucci, P.; Carestia, C.; Fioretti, G.; Mastrobuoni, A. M.; Pagano, L. *Biochem. Biophys. Res. Commun.* **1985**, *130*, 84–90.
(12) Lapolla, A.; Fedele, D.; Aronica, R.; Garbeglio, M.; D'Alpaos, M.; Seraglia, R.; Traldi, P. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1133–1135.
(13) Li, L.; Golding, R. E.; Whittal, R. M. *J. Am. Chem. Soc.* **1996**, *118*, 11662–11663.
(14) Houston, C. T.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1435–1439.
(15) Oliver, R. W. A.; Green, B. N. *TrAC, Trends Anal. Chem.* **1991**, *10*, 85–91.
(16) Kleinert, P.; Schmid, M.; Zurbriggen, K.; Speer, O.; Schmugge, M.; Roschitzki, B.; Durka, S. S.; Leopold, U.; Kuster, T.; Heizmann, C. W.; Frischknecht, H.; Troxler, H. *Clin. Chem.* **2008**, *54*, 69–76.
(17) Geng, M.; Ji, J.; Regnier, F. E. *J. Chromatogr., A* **2000**, *870*, 295–313.
(18) Houston, C. T.; Reilly, J. P. *Anal. Chem.* **1999**, *71*, 3397–3404.
(19) Boemer, F.; Ketelslegers, O.; Minon, J.-M.; Bours, V.; Schoos, R. *Clin. Chem.* **2008**, *54*, 2036–2041.
(20) Williams, J. P.; Giles, K.; Green, B. N.; Scrivens, J. H.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 3179–3186.
(21) Lopez-Ferrer, D.; Canas, B.; Vazquez, J.; Lodeiro, C.; Rial-Otero, R.; Moura, I.; Capelo, J. L. *TrAC, Trends Anal. Chem.* **2006**, *25*, 996–1005.
(22) Kellner, R.; Houthaeve, T. In *Microcharacterization of Proteins*, 2nd ed.; Kellner, R., Lottspeich, F., Meyer, H. E., Eds.; WILEY-VCH: Weinheim, 1999, pp 97−116.
(23) Li, A.; Sowder, R. C., II; Henderson, L. E.; Moore, S. P.; Garfinkel, D. J.; Fisher, R. J. *Anal. Chem.* **2001**, *73*, 5395–5402.
(24) Russell, W. K.; Park, Z.-Y.; Russell, D. H. *Anal. Chem.* **2001**, *73*, 2682–2685.
(25) Strader, M. B.; Tabb, D. L.; Hervey, W. J.; Pan, C.; Hurst, G. B. *Anal. Chem.* **2006**, *78*, 125–134.
(26) Freije, J. R.; Mulder, P. P. M. F. A.; Werkman, W.; Rieux, L.; Niederlander, H. A. G.; Verpoorte, E.; Bischoff, R. *J. Proteome Res.* **2005**.
(27) Slysz, G. W.; Schriemer, D. C. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1044–1050.
(28) Klammer, A. A.; MacCoss, M. J. *J. Proteome Res.* **2006**, *5*, 695–700.
(29) Havlis, J.; Thomas, H.; Sebela, M.; Shevchenko, A. *Anal. Chem.* **2003**, *75*, 1300–1306.
(30) Arnold, R. J.; Hrncirova, P.; Annaiah, K.; Novotny, M. V. *J. Proteome Res.* **2004**, *3*, 653–657.
(31) Yu, Y.-Q.; Gilar, M.; Gebler, J. C. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 711–715.
(32) Alam, M. A.; Bowden, D. K.; Boysen, R. I.; Hearn, M. T. W. A Method for the Rapid Analysis of Polypeptides, PCT Patent Appl. AN 2005:1291827, 2005.
(33) Chen, E. I.; Cociorva, D.; Norris, J. L.; Yates, J. R., III *J. Proteome Res.* **2007**, *6*, 2529–2538.
(34) Pramanik, B. N.; Mirza, U. A.; Ing, Y. H.; Liu, Y.-H.; Bartner, P. L.; Weber, P. C.; Bose, A. K. *Protein Sci.* **2002**, *11*, 2676–2687.
(35) Sun, W.; Gao, S.; Wang, L.; Chen, Y.; Wu, S.; Wang, X.; Zheng, D.; Gao, Y. *Mol. Cell. Proteomics* **2006**, *5*, 769–776.
(36) Hauser, N. J.; Han, H.; McLuckey, S. A.; Basile, F. *J. Proteome Res.* **2008**, *7*, 1867–1872.
(37) Izquierdo, F. J.; Alli, I.; Gomez, R.; Ramaswamy, H. S.; Yaylayan, V. *Food Chem.* **2005**, *92*, 713–719.
(38) Vesper, H. W.; Mi, L.; Enada, A.; Myers, G. L. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2865–2870.
(39) Lin, S.; Yao, G.; Qi, D.; Li, Y.; Deng, C.; Yang, P.; Zhang, X. *Anal. Chem.* **2008**, *80*, 3655–3665.
(40) Swatkoski, S.; Russell, S. C.; Edwards, N.; Fenselau, C. *Anal. Chem.* **2006**, *78*, 181–188.
(41) Swatkoski, S.; Russell, S.; Edwards, N.; Fenselau, C. *Anal. Chem.* **2007**, *79*, 654–658.
(42) Li, J.; Shefcheck, K.; Callahan, J.; Fenselau, C. *Int. J. Mass Spectrom.* **2008**, *278*, 109–113.
(43) Zhong, H.; Marcus, S. L.; Li, L. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 471–481.
(44) Hauser, N. J.; Basile, F. *J. Proteome Res.* **2008**, *7*, 1012–1026.
(45) Swatkoski, S.; Gutierrez, P.; Wynne, C.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2008**, *7*, 579–586.
(46) Guo, N.; Higgins, T. N. *Clin. Biochem.* **2009**, *42*, 99–107.

applied to the identification of variant signature peptides of HbS present in small volumes of EDTA-treated whole human blood.

EXPERIMENTAL SECTION

Materials and Chemicals. Human HbA and ammonium bicarbonate were obtained from Sigma Aldrich (St Louis, MO). Formic acid (FA, 99%, v/v) was obtained from AJAX Chemicals (Sydney, Australia) and HPLC grade acetonitrile (ACN) was obtained from Biolab Scientific Pty Ltd. (Sydney, Australia). Water was distilled and deionized in a Milli-Q system (Millipore, Bedford, MA). α-Cyano-4-hydroxy-cinnamic acid (α-CHCA) and 3,5-dimethoxycinnamic acid (sinapinic acid, SA) were obtained from Agilent Technologies (Forest Hill, Victoria, Australia). Peptides and proteins were dried using a Genevac EZ 2 vacuum evaporator (Pacific Laboratory Products, Blackburn, Victoria, Australia). Special microwave vials (0.2−5 mL), caps and stirrers were obtained from BEST Lab Instrumentation Pty. Ltd., Black Rock, Victoria, Australia. The centrifuge used was a Biofuge Pico (Heraeus) (Pacific Laboratory Products, Blackburn, Victoria, Australia).

Human Blood Sample Collection. Whole human blood samples collected in EDTA containing vacutainers were used. The human ethics approval was obtained from Monash University/Southern Health (approval number 2005000196).

Microwave-Mediated Cleavage. A Biotage Initiator microwave instrument (Biotage AB Uppsala, Sweden) was used for proteolysis of the protein and blood samples using either trypsin or formic acid. The microwave system with high powered focused microwave heating (450 W megatron at 2.45 GHz) delivered a precise heating control with a temperature range of 60 to 250 °C with a pressurized rapid air cooling system. The temperature increased by 2−5 °C/min, minimizing the time required to reach target temperature. For these experiments, the desired temperatures for the experiments were attained using 300 W power at 2.45 GHz ramped up within 30 s. The temperature was maintained at the set temperature using approximately 30 W power at 2.45 GHz. The instrument was equipped with a magnetic stirrer for sample agitation during the reaction time. All microwave vials used in the experiments were 0.2−5 mL glass vials with caps fitted with septa (Biotage AB, Uppsala, Sweden).

All protein samples (0.1 mg/mL) were prepared by dissolution in Milli Q water. Each of the blood samples (≤1 μL), normal blood or blood containing HbS, was diluted with Milli Q water, 1: 500, before FA was added, whereby the final concentration of FA was 0.05% to 3%. The samples were then heated at either 100 or 150 °C with 10 s prestirring performing time course experiments (0 min, 2 min, 5 and 10 min). The samples were rapidly cooled to 4 °C by a pressurized air cooling system to stop the reaction.

Liquid Chromatography. An Agilent Technologies 1100 μLC system consisting of a binary pump, a capillary pump with a micro vacuum degasser, a thermostatted micro well-plate sampler with a 40 μL sample loop, thermostatted column compartment with a two-position/six-port microvalve and a diode-array detector was employed for the HPLC analysis of the peptides. Injected samples were trapped on a RP capillary enrichment column (Zorbax 300SB C18, 0.3 × 5 mm, particle size 5 μm) enabling the online desalting of samples. A capillary pump connected to the microvalve in the column compartment was used to elute the peptides from the RPC enrichment column with reversed flow through an μRPC column (Zorbax 300SB C18, 0.3 × 150 mm, 3.5 μm particle size) to the electrospray source of the LC/MSD Trap via a diode-array detector.

The sample injection (2 μL) and loading onto the trapping column was performed with 0.1% formic acid in 5% ACN by using the binary pump at a flow rate of 20 μL/min. The trapping time was set to 6 min. The microvalve was then switched to direct the flow from the capillary pump to the enrichment column and the analytical column. The desalted peptides were then eluted from the enrichment column and separated by the capillary column with an increasing linear gradient of 1% B/min over 50 min (5−55%) at a flow rate of 4 μL/min with eluent A: water (100%) and eluent B: ACN (100%), both with 0.1% FA as ion paring agent. The total runtime including column equilibration time was 87 min.

Mass Spectrometry. ESI MS analysis was carried out in the positive ion mode with an Agilent 1100 series LC/MSD-SL ion trap mass spectrometer through an Agilent G1607A orthogonal electrospray interface (Agilent Technologies, Waldbronn, Germany). The electrospray voltage was set to 3.5 kV and the ion trap mass spectrometer was operated in full scan mode in the range of 100−1800 m/z. The nebulizing gas (N_2) pressure, the drying gas (N_2) flow rate, and the drying gas temperature were set at 10 psi, 5 L/min, and 300 °C, respectively. The target mass was set at 800 m/z, the compound stability at 80% and the trap drive level to 100%. The ion accumulation time was automatically adjusted using the ion charge control (ICC) feature of the instrument. The maximal accumulation time was 300 ms and the ICC target was 30 000. For the tandem mass spectrometry experiments, the instrument automatically switched from MS to MS/MS mode when the intensity of a particular ion exceeded the preset threshold of 10 000. Active exclusion was executed after two spectra for 1 min. Fragmentation amplitude was set to 1 and the number of precursor ion selection was set to 2. Smart fragmentation mode was enabled which allowed the fragmentation amplitude to be varied from 30% to 200% of the default fragmentation amplitude (1.0 V). All system control and data acquisition were conducted with the Agilent ChemStation and MSD Trap Control software, version 4.2.

MALDI-TOF MS of proteins and all the proteolytic digest products was performed with a Voyager DE-STR MALDI-TOF MS system from Applied Biosystems, Framigham, MA. The system was equipped with a 337 nm nitrogen laser using 3 ns duration pulses with a maximum firing rate of 20 Hz. The mass spectrometer was controlled by the Voyager DE-STR Biospectrometry workstation software. Linear mode spectra were obtained with delayed extraction using a delay time of 300 ns, a grid voltage of 90% of the acceleration voltage, with positive polarity. The mass range was 5000−25 000 m/z with a lower mass gate set at 5000 Da for mass data acquisition. Each spectrum was obtained with 500 laser shots by accumulating five spectra each obtained by 100 laser shots. Reflectron mode spectra were obtained with delayed extraction using a delay time of 250 ns with positive polarity. The grid voltage was set at 85%. The mass acquisition mass range was 650−10 000 m/z, where the low mass gate was set at 500 Da. Again, each spectrum was obtained with 100 laser shots and five consecutive spectra were accumulated. The laser intensity was fixed to 2500 for the reflectron mode and 3000 for

the linear mode data acquisition. A reversed two layer sample spotting technique was used to spot the samples whereby SA was used for intact protein analysis and α-CHCA was used for peptides.

Data Processing and Database Searching. Peptide mass fingerprinting (PMF) of MALDI TOF MS peak lists were performed by Mascot (http://www.matrixscience.com). All spectra were manually evaluated against peak lists generated for a FA cleavage for α and β globin chains with Spectrum Mill MS Digest, using a customer programmed subroutine (Agilent). The raw MS/MS data acquired were processed with the Agilent Data Analysis software (Version 4.2) prior to database searching. Post processed data were converted to Mascot generic files and searched against the UniProt_Swiss-Prot database (http://www.expasy.org/sprot/) by using the web-based search engine Phenyx (Geneva Bioinformatics, http://www.phenyx-ms.com).[47] A two-round search was performed to increase the level of confidence on protein identification. The monoisotopic masses of the peptide fragments were calculated with the programm PeptideMass at the ExPASy Web site http://au.expasy.org/tools/peptide-mass.html.

Nomenclature. The proteolytic/cleavage fragments of the hemoglobins were coded below by (i) the symbol of the particular globin chain (α, β, δ, ϵ, γ, ζ), followed by (ii) the name of the Hb variant as subscript (N = normal), (iii) the symbol of the cleavage reagent (T = trypsin, GluC = endoproteinase Glu-C, AspN = endoproteinase Asp-N and FA = formic acid) and finally (iv) by fragment number X1, X2, X3,... etc. of a completely cleaved protein sequence according to a particular cleavage rule, starting from the N-terminus of the protein and numbering the proteolytic fragments consecutively with missed cleavages indicted by hyphenation, for example, β_SFA1 describes fragment 1 of the β globin chain of HbS derived by a formic acid cleavage).[48] The amino acid numbering system follows common practice, with the amino acid after the initiator methionine given the number 1. The term signature peptide refers to a proteolytic/cleavage fragment of a protein variant containing a mutation that involves substitution, insertion or deletion of amino acid residue(s).

RESULTS AND DISCUSSION

As employed in this investigation, the identification of particular Hb variants was based on the detection of the Hb signature peptide which contained the mutation (amino acid substitution, insertion or deletion). To generate peptide fragments, proteins are typically digested with an enzyme such as trypsin, which cleaves C-terminal to lysine and arginine. However, in the case of many protein variants, this approach presents an additional challenge since an additional cleavage site can be introduced at the mutation site, for example, for the characterization of HbE ($Glu^{26} \rightarrow Lys^{26}$) or HbC ($Glu^{6} \rightarrow Lys^{6}$). With Hb variants of known primary structure, the choice of the proteolytic/cleavage reagents can be based on an a priori examination of the number, molecular mass, size and molecular properties of the signature peptide(s). This approach is exemplified in Table 1 for the in silico proteolysis of HbS with trypsin, endoproteinase Glu-C, endoproteinase Asp-N and formic acid. For example, with endoproteinase AspN cleavage of the protein N-terminal to the aspartic acid residues is expected a priori to generate the signature peptide β_SAspN1 $m/z = 2148.2175$ for HbS covering the first 20 amino acids of the sequence. Since EDTA is an inhibitor of endoproteinase AspN, this enzyme cannot be used with EDTA treated whole blood samples. On the other hand, formic acid (FA) is a suitable alterative. Moreover, FA is MS compatible and also allows site specific cleavage of protein(s) C-terminal, or less frequently N-terminal to aspartyl residues.[23,49−51] In the present investigation a sequential method was thus developed, involving microwave-assisted chemical cleavage to generate suitable peptide fragments of Hb variants present in EDTA-treated blood from individuals with sickle cell disease, that is, hemoglobin (HbS) with the amino acid sequence coverage monitored by MALDI ToF MS or RPHPLC ESI MS/MS.

Using human HbA, a set of time course experiments for microwave-assisted FA cleavage were performed to determine the effectiveness of the cleavage by monitoring by analytical reversed-phase liquid chromatography the depletion of the intact protein and the appearance of the cleavage products after 0, 5, and 10 min under different microwave-irradiation conditions. An increased number of fragments was observed when this reaction was performed at 150 °C instead of at 100 °C. A significant number, although from subsequent MALDI TOF MS or LC ESI MS/MS analyses clearly an incomplete sequence coverage set, of FA cleaved peptides were obtained after 5 min of reaction time, while an irradiation time of 10 min resulted in a more complete set of FA cleavage peptides with consistency achieved in terms of the number of peaks, peak intensity or RP-HPLC retention times upon replication. Based on these results, the maximum time of the cleavage reaction and microwave irradiation condition were chosen to be 10 min and 150 °C to minimize the potential of long irradiation times leading to oxidation and/or N-terminal formylation of the derived peptides.[49] An example of the MALDI TOF MS of the 5 min digest of a HbA sample with 0.05% FA at 150 °C is shown in Figure 1. The hemoglobin sequence coverage obtained under these conditions by MALDI TOF MS was 98.58% for the α-chain and 79.45% for the β-chain. Lower sequence coverages was obtained when analogous analyses were carried out using LC ESI MS/MS; that is, for the α and β chains the coverages were 87.23% and 68.49%, respectively.

The method was then validated for reproducibility of occurrence of cleaved peptides and sequence coverage using human blood samples, replicating 2, 5, and 10 min chemical digestion experiments at 150 °C with varying concentrations of FA (0.05−5%). No significant differences in the number of peptides were observed with these different FA concentrations. Interestingly, slightly higher ion intensities of the larger peptide fragments were, however, observed with higher concentrations of FA and longer reaction times (data not shown). From these experimental data, it was evident that FA cleavage occurred at either side as well as at both sides (C- and N-terminal) of the aspartic acid residues. In a protein, there are four different possible cleavage sites for the sequence.. ..YD_1XXXD_2Z.... in the presence of formic acid: (1) Cleavage could occur at the N-terminal position of both D_1 and D_2 resulting in the peptide

(47) Anonymous. *Proteomics*, **2004**, *4*, 2206.
(48) Hennessy, T. P.; Boysen, R. I.; Huber, M. I.; Unger, K. K.; Hearn, M. T. W. *J. Chromatogr., A* **2003**, *1009*, 15–28.
(49) Hua, L.; Low, T. Y.; Sze, S. K. *Proteomics* **2006**, *6*, 586–591.
(50) Han, K. K.; Richard, C.; Biserte, G. *Int. J. Biochem.* **1983**, *15*, 875–884.
(51) Inglis, A. S. *Methods Enzymol.* **1983**, *91*, 324–332.

Table 1. Signature Peptides (in Bold) for Haemoglobin S Obtained under Different Enzymatic and a Formic Acid Proteolysis Method

cleavage method		parent parent/variant protein	wildtype Hb Hb α chain	wildtype Hb Hb β chain	variant HbS Hb β chain
none	intact protein	UniProtKB/Swiss-Prot identifier	P69905	P68871	
		protein mass[a] $[m+H]^+$	15127.37	15868.23	15838.25
		number of amino acids	141	146	146
		protein mass shift			−29.98
		type of variant[b]			Substitution
		sequence variation			β chain
		codon substitution			GAG > GTG
		position			B6
		amino acid substitution			Glu > Val
enzymatic	trypsin	number of peptide fragments[c]	14	15	15
		normal/signature peptide mass[d]		952.5098	**922.5356**
		position		1−8	**1−8**
		peptide code		β_NT1	**β_ST1**
	Glu C	number of peptide fragments	5	8	8
		normal/signature peptide mass		824.4148	**794.4406**
		position		1−7	**1−7**
		peptide code		β_NGluC1	**β_SGluC1**
	Asp N	number of peptide fragments	9	8	8
		normal/signature peptide mass		2178.1917	**2148.2176**
		position		1−20	**1−20**
		peptide code		β_NAspN1	**β_SAspN1**
chemical	formic acid	number of peptide fragments	16	15	15
		normal/signature peptide mass 1		2178.1917	**2148.2176**
		position		1−20	**1−20**
		peptide code		β_NFA1	β_SFA1
		normal/signature peptide mass 2		2293.2187	**2263.2445**
		position		1−21	**1−21**
		peptide code		β_NFA1−2	**β_SFA1−2**

[a] Average mass. [b] Substitution, insertion or deletion. [c] For a complete cleavage. [d] Monoisotopic mass.

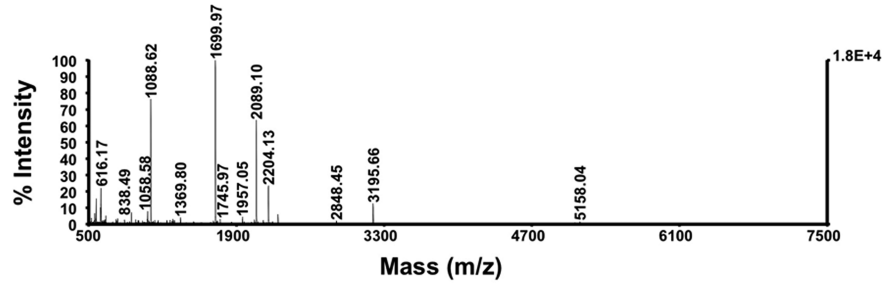

Figure 1. MALDI TOF mass spectrum acquired in reflectron mode of peptides generated by a 5 min microwave assisted chemical digestion with FA (0.05%) of a HbA standard at 150 °C showing the *m*/*z* values of chemically cleaved peptides.

D_1XXX; (2) Cleavage could occur at the N-terminal position of D_1 and the C-terminal position of D_2, resulting in the peptide D_1XXXD_2; (3) Cleavage could occur at the C-terminal position of D_1 and the N-terminal position of D_2, resulting in the peptide XXX, and (4) Cleavage could occur at the C-terminal positions of both D_1 and D_2, resulting in the peptide $XXXD_2$. When the peptide fragments derived from the FA-cleavage of HbA were analyzed with LC ESI MS/MS, it was apparent that the most abundant type of peptide generated involved the double N-terminal cleavage (50.0%), followed by 6.3%, 37.5% and 6.3%, for the second, third and fourth type of cleavage.

The MALDI TOF MS analysis of the hemoglobin peptides that are reproducibly generated by this microwave-mediated procedure from blood is shown in Figure 2, with the list of peptides identified by MALDI TOF MS and LC ESI ion trap MS presented in Table 2. The sequence coverages obtained by MALDI TOF MS for the α- and β- chain was 100% and 82.87%, respectively, whereas the sequence coverages obtained by LC ESI MS for the α- and β- chains were 59.57% and 47.95%, respectively.

This new method was then applied to blood samples from individuals with heterozygote HbS ($\alpha_2\beta\beta_S$), containing either 25% or 70% β_S-chain. The relative percentages of the HbS β_S-chain were measured by HPLC. HbS is the result of a single-gene defect causing a substitution of valine for glutamic acid at position 6 of the β-chain of adult HbA. The resulting β_S-chain differs from the normal β-chain by a molecular mass of −29.98 amu. The homozygous Hb($\alpha_2\beta_S\beta_S$) state causes sickle cell disease. With HbS heterozygotes, where the $Glu^6 \rightarrow Val^6$

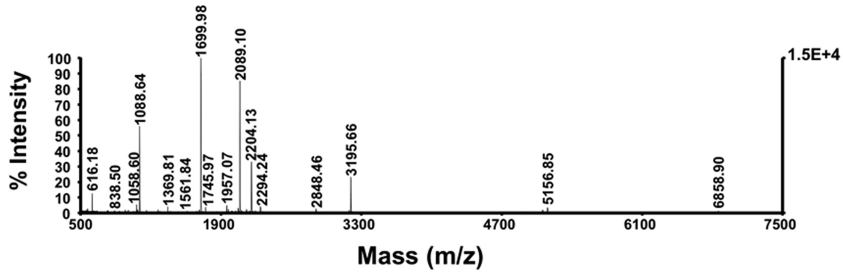

Figure 2. MALDI TOF mass spectrum acquired in reflectron mode of HbA peptides generated as in Figure 1 showing good identification of chemically cleaved peptides from an EDTA-treated blood lysate.

Table 2. Peptide Sequences and Positions of Fragments Cleaved with a Microwave Assisted FA Proteolysis in the α- and β-Globin Chains of Normal Human Hb from Whole Human Blood Lysate As Detected by MALDI TOF MS Compared to Their Theoretical *m/z* Values and Their Identification (√) by LC ESI Ion Trap MS/MS

Globin Chain	Position	Sequence	Theoretical m/z values $[M+H]^+$	MALDI TOF MS $[M+H]^+$	MALDI TOF MS	LC ESI ion trap MS/MS
α	1-5	(-)VLSPA(D)	486.2922			√
α	1-6	(-)VLSPAD(K)	601.3370			√
α	1-64	(-)VLSPADKTNVKAAWGKVGAHAGEYGAEALERM FLSFPTTKTYFPHFDLSHGSAQVKGHGKKVAD(A)	6855.5070	6856.9381	208.7	
α	47-63	(F)DLSHGSAQVKGHGKKVA(D) *or* (D)LSHGSAQVKGHGKKVAD(A)	1718.9296	1718.9196	5.8	
α	64-73	(A)DALTNAVAHV(D)	1010.5265			√
α	64-74	(A)DALTNAVAHVD(D)	1125.5534	1125.5502	2.8	√
α	64-84	(A)DALTNAVAHVDDMPNALSALS(D) *or* (D)ALTNAVAHVDDMPNALSALSD(L)	2125.0230	2125.0519	13.5	
α	64-93	(A)DALTNAVAHVDDMPNALSALSDLHAHKLRV(D) *or* (D)ALTNAVAHVDDMPNALSALSDLHAHKLRVD(P)	3194.6374	3194.6514	4.4	
α	74-84	(V)DDMPNALSALS(D)	1133.5143			√
α	74-93 or 75-94	(V)DDMPNALSALSDLHAHKLRV(D) *or* (D)DMPNALSALSDLHAHKLRVD(P)	2203.1288	2203.1263	1.1	
α	75-93	(D)DMPNALSALSDLHAHKLRV(D) *or* (D)MPNALSALSDLHAHKLRVD(P)	2088.1018	2088.1157	6.6	
α	85-93	(S)DLHAHKLRV(D)	1088.6323	1088.6642	29.3	√
α	86-93	(D)LHAHKLRV(D)	973.5875			√
α	94-141	(V)DPVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKF LASVSTVLTSKYR(-)	5220.8281	5219.9867	161.1	
α	95-141	(D)PVNFKLLSHCLLVTLAAHLPAEFTPAVHASLDKFL ASVSTVLTSKYR(-)	5105.7833	5104.9821	156.9	
α	126-141	(L)DKFLASVSTVLTSKYR(-)	1815.0010	1815.0311	16.5	√
α	127-141	(D)KFLASVSTVLTSKYR(-)	1699.9532	1699.9780	14.5	√
Sequence coverage				**100%**		**59.57%**
β	1-20	(-)VHLTPEEKSAVTALWGKVNV(D)	2178.1917	2178.1798	5.5	
β	1-21	(-)VHLTPEEKSAVTALWGKVNVD(E)	2293.2187	2293.2365	7.8	√
β	47-72	(G)DLSTPDAVMGNPKVKAHGKKVLGAFS(D) *or* (D)LSTPDAVMGNPKVKAHGKKVLGAFSD(G)	2267.4246	2267.4423	7.8	
β	47-51	(G)DLSTP(D)	532.2613			√
β	52-72	(P)DAVMGNPKVKAHGKKVLGAFS(D)	2154.1856	2154.1745	5.1	√
β	53-72	(D)AVMGNPKVKAHGKKVLGAFS(D)	2039.1403	2039.2009	29.7	√
β	73-78	(S)DGLAHL(D)	625.3304			√
β	73-98	(S)DGLAHLDNLKGTFATLSELH CDKLHV(D) *or* (D)GLAHLDNLKGTFATLSELH CDKLHVD(P)	2847.4457	2847.4678	7.7	
β	74-78	(D)GLAHL(D)	510.2856			√
β	79-93	(L)DNLKGTFATLSELHC(D)	1648.7999			√
β	79-98	(L)DNLKGTFATLSELHCDKLHV(D) *or* (D)NLKGTFATLSELHCDKLHVD(P)	2241.1332	2241.2112	34.8	
β	80-93	(D)NLKGTFATLSELHC(D)	1533.7551			√
β	100-146	(D)PENFRLLGNVLVCVLAHHFGKEFTPPVQAAYQKV VAGVANALAHKYH(-)	5153.7231	5154.0113	55.9	
Sequence coverage				**82.87%**		**47.95%**

mutation occurs in only one of the β-chains, if the β_S-chain is associated with a variant Hb β-thalassaemia chain (compound heterozygocity), sickle cell disease may also result.

The MALDI TOF reflectron mode mass spectrum of the 25% HbS sample after microwave digestion with FA is shown in Figure 3. The derived fragments from a normal β-chain, β_NFA1, V^1HLTPEEKSAVTALWGKVNV^{20} ($[M+H]^+$ with a monoisotopic mass of 2178.1917) and β_NFA1−2 V^1HLTPEEKSAVTALWGKVNVD^{21} ($[M+H]^+$ with a monoisotopic mass of 2293.2187) are replaced in a sickle β-chain by the fragments of smaller

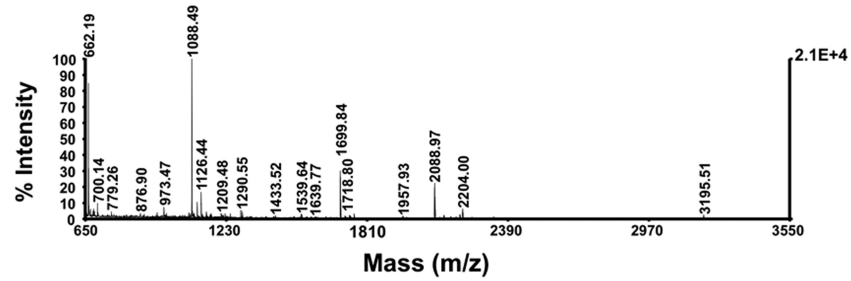

Figure 3. Enlarged MALDI TOF reflectron mode mass spectrum of peptides, generated as in Figure 1, of a 25% β-chain containing blood sample.

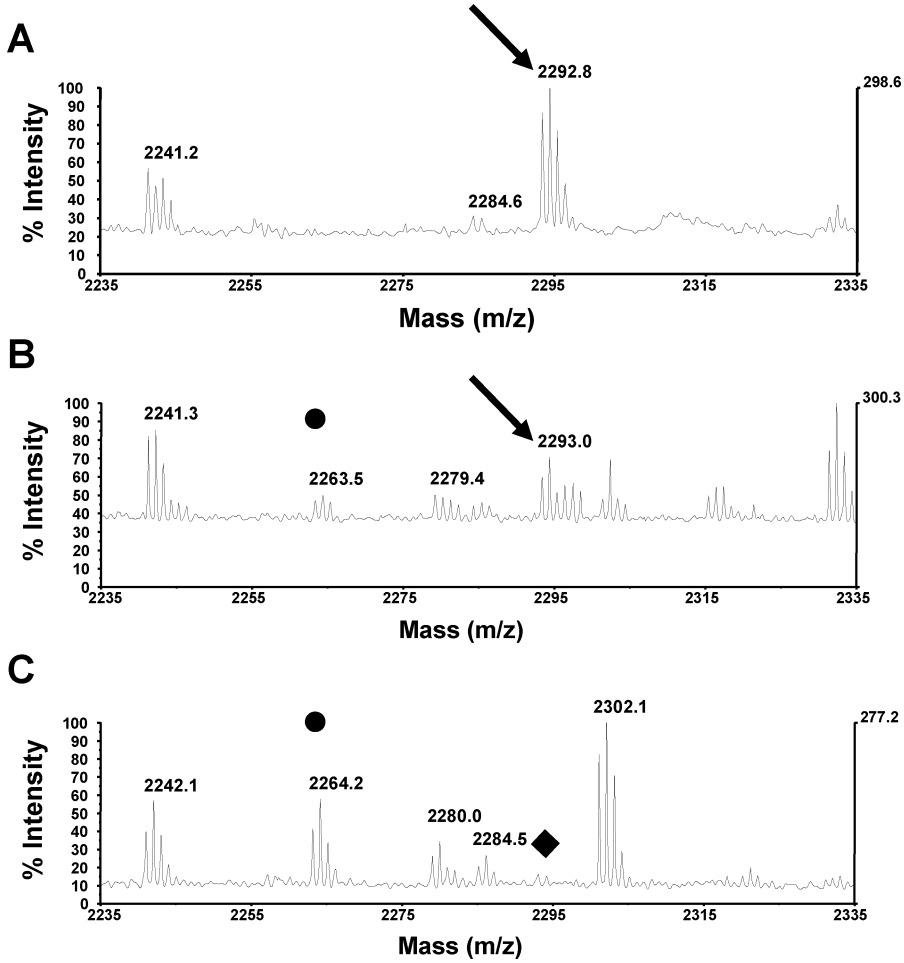

Figure 4. Comparison of MALDI TOF mass spectra of microwave assisted, chemically (FA) cleaved peptides in EDTA-treated blood samples of normal and two Hb variants. (A) Normal Hb sample ($\alpha_2\beta_2$) whereby the mass spectrum with the expanded *m*/*z* range shows the N-terminal peptide signal (arrow) representing β_NFA1−2 V^{1}HLTPEEKSAVTALWGKVNVD^{21}, C-terminally cleaved at D^{21} (*m*/*z* of [M+H]$^+$ 2293.2365). (B) HbS carrier heterozygotes sample with 25% β_S-chain. ($\alpha_2\beta\beta_S$) The peak for the N-terminal V^{1}HLTPEEKSAVTALWGKVNVD^{21} (β_NFA 1−2) peptide from the wild type β chain (arrow) ([M+H]$^+$ of 2293.2365) and the signal for the C-terminally cleaved V^{1}HLTPVEKSAVTALWGKVNVD^{21} from the aberrant globin chain β_S (circle) (received [M+H]$^+$ monoisotopic mass [M+H]$^+$ of 2263.4904, β_SFA1−2, expected monoisotopic mass [M+H]$^+$ of 2263.2445) is diagnostic of a HbS heterozygote state. (C) HbS thalassaemia compound heterozygote sample with 70% β_S-chain ($\alpha_2\beta\beta_S$) with the signature peptide V^{1}HLTPEVKSAVTALWGKVNVD^{21} for HbS (circle) (received *m*/*z* [M+H]$^+$ with a monoisotopic mass of 2263.1767, expected *m*/*z* [M+H]$^+$ with a monoisotopic mass of 2263.2445) and the normal β chain just detectable with very low intensity (diamond).

masses, V^{1}HLTPE**V**KSAVTALWGKVNV^{20} and V^{1}HLTPE**V**KSAVTALWGKVNVD^{21} with expected monoisotopic masses of [M+H]$^+$ 2148.2176 and 2263.2445, respectively, derived from N- or C- terminal chemical cleavage at D^{21}, for individuals who are homozygote ($\alpha_2\beta_{S2}$). With the blood samples from heterozygote individuals ($\alpha_2\beta\beta_S$) all four fragments can be expected.

Figure 4A shows a MALDI TOF mass spectrum with the expanded *m*/*z* range of 2235−2330 where the N-terminal peptide signal (arrow) was detected representing β_NFA1−2 V^{1}HLTPEEKSAVTALWGKVNVD^{21}, C-terminally cleaved at D^{21} (with a monoisotopic *m*/*z* of [M+H]$^+$ 2293.2365 and a mass accuracy of 7.8 ppm). Figure 4B shows a mass spectrum of the 25% HbS blood sample ($\alpha_2\beta\beta_S$). The appearance of the peak for N-terminal V^{1}HLTPEEKSAVTALWGKVNVD^{21} (β_NFA 1−2) peptide from the wild type β-chain (arrow) (monoisotopic mass [M+H]$^+$ of 2293.2365, with an 7.8 ppm mass accuracy) and the signal for

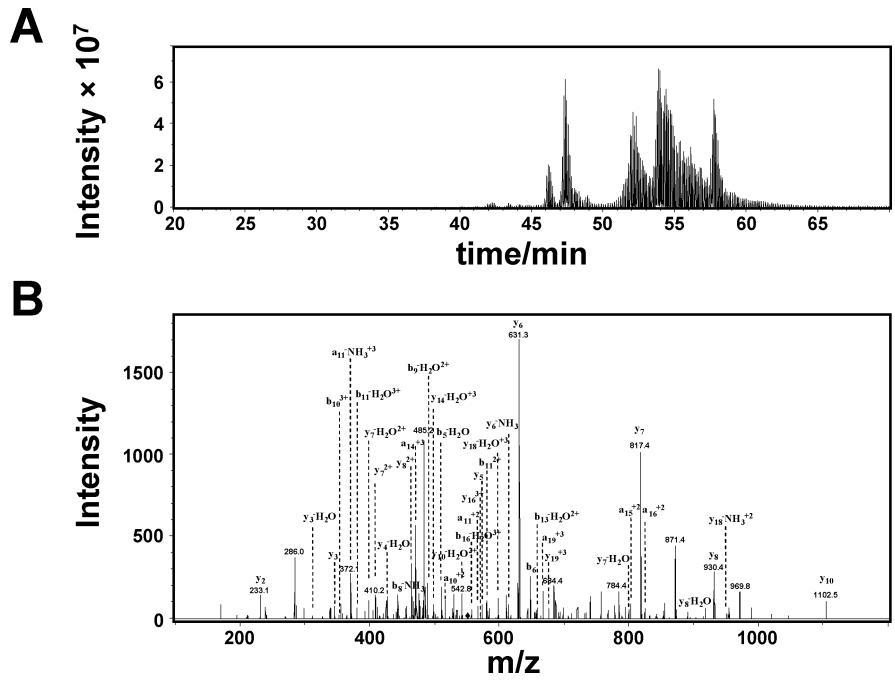

Figure 5. (A) Total ion chromatogram of all detected peptides related to Hb and highly abundant blood proteins. (B) The HbS signature peptide identified by MS/MS of the triply charged precursor ion, V^{1}HLTPEVKSAVTALWGKVNVD^{21}, (with a *m/z* value of $[M+3H]^{3+}$ 754.0) showing the abundant b-, y-, and a- ions of different charge states.

the C-terminally cleaved V^{1}HLTPVEKSAVTALWGKVNVD^{21} from the aberrant globin chain β_S (circle) (observed $[M+H]^+$ monoisotopic mass $[M+H]^+$ of 2263.4904, β_SFA1−2, expected monoisotopic mass $[M+H]^+$ of 2263.2445, with a mass accuracy of ±109 ppm) is diagnostic of a HbS heterozygote state. In case of Hb variant heterozygotes, the amount of variant β-chain, in this case Hb β_S-chain, is usually less than normal type β-chain thus the signal is expected to be of different intensity. The presence of another peak ($[M+H]^+$ monoisotopic mass of 2278.4609) which did not appear in the corresponding MS analysis of a normal blood sample could not be matched to any HbN or HbS FA cleaved peptides and was not detected by LC ESI MS-MS analysis of the sample, but could be a structural variant related to a high abundant blood protein. Further work to characterize this protein from the HbS blood samples is in progress.

Co-inheritance of a gene coding for a Hb variant and a thalassaemia gene (compound heterozygotes) can result in a wide range of expression levels of the hemoglobin with the interaction between the chains exacerbating or alleviating and phenotype of the affected and unaffected globin chains of individual variants.[7] Analysis of a blood sample from an individual with coinheritance of HbS and thalassaemia (with 70% HbS, elevated minor adult HbA_2 and HbF levels (as determined by HPLC analysis shown in (Figure 4C) documented the presence of the signature peptide V^{1}HLTPEVKSAVTALWGKVNVD^{21} for HbS (arrow) (observed *m/z* $[M+H]^+$ with a monoisotopic mass of 2263.1767, expected *m/z* $[M+H]^+$ with a monoisotopic mass of 2263.2445, with a mass accuracy of ±30 ppm) whereas the normal β-chain was detectable at very low intensity (diamond). This sample, when analyzed by LC ESI ion trap MS/MS, showed a total ion chromatogram of all detectable peptides related to Hb and other highly abundant blood proteins (Figure 5A). However, the HbS signature peptide V^{1}HLTPEVKSAVTALWGKVNVD^{21} was readily identified by MS/MS of the triply charged precursor ion (with a *m/z* value of $[M+3H]^{3+}$ 754.0), resulting in an information rich spectrum (Figure 5B) with abundant b-, y-, and a-ions of different charge states.

CONCLUSIONS

Because of the number, variety, and clinical significance of hemoglobin variants, the need exists to design and deploy fast and economical analytical methods for their detection in human blood. Unambiguous identification of the structure of Hb variants requires a suitable cleavage method which results in the generation of signature peptide(s). The disadvantages of existing methods employing trypsin or other enzymes are that they are slow, often expensive, and not suitable for all variants. In contrast, with acidic cleavage reagents, their specificity varies widely from unspecific (e.g. HCl) to highly specific (FA). Since formic acid is a suitable mobile phase additive commonly used in LC in conjunction with ESI MS detection, its use as a cleavage reagent will not introduce any additional, interfering components into such analytical systems. As demonstrated from this study, the cleavage time can be readily reduced when the cleavages are performed at higher temperature under conditions of microwave irradiation. Optimum reaction times, temperatures and solvent conditions can be readily developed. The results from the current study confirmed that FA-mediated cleavage occurred at either side or at both sides (C- and N-terminal) of aspartic acid residues within the protein, although cleavage preferentially occurred N-terminally as determined from the detailed analysis of the derived peptides. With the conditions employed, the sequence coverage of the α- and β- chain of Hb was highest when analyzed with MALDI TOF MS. The applicability of this method was demonstrated with very small volumes (<1 mL diluted 1: 500) of human blood from normal individuals, HbS heterozygotes and HbS thalassaemia compound heterozygotes with differing expression levels of HbS. Both,

MALDI TOF MS and ESI MS/MS allowed the detection of the HbS signature peptide with ESI-MS/MS analysis providing highly reliable identification. The application of this microwave-assisted chemical cleavage method to clinically relevant blood samples using a low concentration FA (0.05%) confirmed the usefulness of this method for the identification of variant Hbs, utilizing only small amounts of blood and reagents. Furthermore, the site specificity of FA-mediated cleavage coupled with the frequency of aspartic acid in protein sequences makes this chemical cleavage method an attractive alternative to proteolysis with enzymes followed by subsequent mass spectrometric protein identification. For example, the abundance of aspartic acid in human proteins is 4.9%, in comparison with lysine 5.7% and arginine 5.6%, as estimated from human genomic data. Thus, the microwave-assisted, MS-compatible chemical cleavage of proteins with low concentrations of FA under well-defined conditions of sample preparation offers considerable scope for numerous further applications in targeted proteomics and the analytical characterization of disease-specific proteins.

ACKNOWLEDGMENT

The financial support of the Australian Research Council is gratefully acknowledged.

Received for review June 29, 2010. Accepted September 9, 2010.

AC1017106