

Identification of hemoglobin variants by top-down mass spectrometry using selected diagnostic product ions

Didia Coelho Graça · Ralf Hartmer · Wolfgang Jabs ·
Photis Beris · Lorella Clerici · Carsten Stoermer ·
Kaveh Samii · Denis Hochstrasser · Yury O. Tsybin ·
Alexander Scherl · Pierre Lescuyer

Received: 11 December 2014 / Revised: 23 January 2015 / Accepted: 27 January 2015 / Published online: 10 March 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Hemoglobin disorder diagnosis is a complex procedure combining several analytical steps. Due to the lack of specificity of the currently used protein analysis methods, the identification of uncommon hemoglobin variants (proteoforms) can become a hard task to accomplish. The aim of this work was to develop a mass spectrometry-based approach to quickly identify mutated protein sequences within globin chain variants. To reach this goal, a top-down electron transfer dissociation mass spectrometry method was developed for hemoglobin β chain analysis. A diagnostic product ion list was established with a color code strategy allowing to

quickly and specifically localize a mutation in the hemoglobin β chain sequence. The method was applied to the analysis of rare hemoglobin β chain variants and an $\Lambda\gamma$ - β fusion protein. The results showed that the developed data analysis process allows fast and reliable interpretation of top-down electron transfer dissociation mass spectra by nonexpert users in the clinical area.

Keywords Hemoglobin · Hemoglobin variant · Top-down mass spectrometry · Electron transfer dissociation ETD · Clinical mass spectrometry

Electronic supplementary material The online version of this article (doi:10.1007/s00216-015-8525-5) contains supplementary material, which is available to authorized users.

D. Coelho Graça · D. Hochstrasser · A. Scherl · P. Lescuyer
Department of Human Protein Sciences, Faculty of Medicine,
Geneva University, Rue Michel Servet 1, 1211 Geneva, Switzerland

R. Hartmer · W. Jabs · C. Stoermer
Bruker Daltonics, Fahrenheitstraße 4, 28359 Bremen, Germany

P. Beris · K. Samii
Hematology Service, Geneva University Hospitals,
Rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland

P. Beris
Laboratoire Unilabs Coppet, Chemin des Perrières 2,
1296 Coppet, Switzerland

L. Clerici · D. Hochstrasser · A. Scherl · P. Lescuyer (✉)
Department of Genetic and Laboratory Medicine, Geneva University
Hospitals, Rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland
e-mail: Pierre.Lescuyer@unige.ch

Y. O. Tsybin
Biomolecular Mass Spectrometry Laboratory, Ecole Polytechnique
Fédérale de Lausanne, 1015 Lausanne, Switzerland

Introduction

Hemoglobin (Hb) is a tetrameric blood protein that transports oxygen to organs and tissues. The main form in adults is HbA (97 %), composed of two α chains and two β chains. A healthy person carries also two minor forms: HbA₂ composed of two α chains and two δ chains (2–3 %) and HbF composed of two α chains and two γ chains (<1 %). Hemoglobin disorders, also named hemoglobinopathies, represent the most common monogenic disease in the world [1]. They are divided into two main categories. The first one corresponds to qualitative disorders where an abnormal globin chain, called Hb variant, is present. Hb variants or proteoforms are mainly due to single-point mutations but also occur due to DNA deletion resulting in a truncated or a fusion protein. The second category, termed thalassemias, corresponds to quantitative hemoglobin disorders. Thalassemias result from the decreased synthesis of one or several globin chains leading to an imbalance in the hemoglobin chain ratio.

Nowadays, the biological diagnosis of Hb disorders is a multistep process combining hematological tests, protein analysis techniques, and molecular biology assays [2]. The protein

analysis step relies on a combination of several analytical methods, such as liquid chromatography (LC), capillary electrophoresis, and isoelectric focusing, to detect and putatively identify Hb variants. The variant is then confirmed using direct molecular biology such as polymerase chain reaction. This process is relatively fast and efficient for the identification of the most common Hb variants such as HbS, HbC, or HbE. The identification of less common mutations requires more expertise. Overall, 1206 Hb variants have been described so far [3]. Protein screening techniques allow knowing in the majority of cases if the patient has an α chain variant or a β chain variant. Then, sequencing of alpha1, alpha2, or beta globin genes, depending on the suspected variant, allows identifying rapidly the abnormal Hb chain. The situation is different, and much more difficult, when screening techniques do not separate the Hb variant from HbA. In such a case, when clinical and hematological pictures are highly suggestive of an abnormal Hb, the usual process is to sequence all Hb genes. Finally, the situation is even more difficult in the case of a fusion protein, although those that are clinically significant have well-described chromatographic and electrophoretic profiles. Examples of well-described fusion proteins are Hb Lepore and Hb Kenya. In such cases, the large gene deletion resulting in the synthesis of a fusion protein makes molecular biology confirmation analyses more complex than with single-point mutations.

During the last decades, mass spectrometry (MS) has demonstrated its utility in protein analysis. Two different approaches are commonly used: bottom-up and top-down MS. The first approach consists in digesting the sample of interest and then in analyzing the resulting peptide mixture, typically by LC-MS/MS. This approach is widely used for the characterization of proteomes, for biomarker identification and for targeted protein analysis [4]. The second approach consists in analyzing the intact protein using MS/MS without prior digestion. In some cases, depending on protein size and structure, more precise sequence information and post-translation modification (PTM) analysis can be provided with top-down MS [4]. For characterization of multiple proteoforms which result from several modifications on a single protein, the top-down approach is especially suited [5, 6]. Top-down MS analysis is nowadays mostly performed with electron capture dissociation (ECD) and electron transfer dissociation (ETD) [7, 8]. These two activation modes permit a more efficient production of product ions along the protein backbone compared to slow-heating ion activation and fragmentation modes, e.g., collision-induced dissociation.

Different approaches were developed to characterize and identify Hb variants by MS. Bottom-up approaches showed good results for high-throughput, targeted screening of common Hb variants, e.g., HbS, HbC, and HbE [9, 10]. However, targeted peptide measurements do not allow identifying uncommon Hb mutations if they are not present in the targeted Hb variant list. Several approaches using top-down MS were

also developed for the identification and characterization of common Hb variants [5, 11–13]. Compared to bottom-up methods, intact protein approaches have, in principle, the advantage to reduce the number of sample preparation steps and decrease or eliminate sample preparation time [14]. Cooper and co-workers showed recently the possibility to detect Hb variant with a mass shift less than 1 Da by direct surface sampling top-down high-resolution MS [15]. Nevertheless, further efforts are necessary to implement characterization of uncommon Hb variants by MS, including data analysis and data interpretation, which is compatible with clinical lab practice. Particularly, top-down MS data analysis software requires improvement for allowing automated data analysis and interpretation solutions for non-expert teams [16].

In this study, we aimed at taking benefit from the analytical capabilities of high-resolution (HR) and high-mass accuracy ETD MS to facilitate the characterization of Hb variants. We therefore developed a top-down ETD MS workflow that allows straightforward and accurate localization of the mutated sequences within the Hb chains. In particular, we established an integrated protocol for data acquisition, data analysis, and data interpretation based on the use of selected diagnostic ions. Our results show that HR ETD MS could be successfully integrated within the current hemoglobin disorder diagnostic process to quickly provide reliable structural information on rare hemoglobin variants difficult to characterize using standard protein assays.

Materials and methods

Hemoglobin samples

Purified human HbA was obtained from Sigma-Aldrich (St. Louis, MO, USA). Blood samples from patients carrying Hb β chain single-point mutations were obtained from Unilabs (Coppet, Switzerland; see Table 1). The blood sample carrying the Hb Kenya fusion protein was collected at the Geneva University Hospitals. The use of these human specimens was approved by the Ethical Committee for Research on Human Being of the Geneva state. Hb variant identification was performed according to standard clinical laboratory practices by experimented lab teams and hematologists. All mutations were confirmed by DNA analysis.

Clinical sample preparation

Hemoglobin solutions were prepared from 20 μ L of EDTA blood samples as previously described [13]. Hemoglobin extracts were stored at -80 °C until analysis. Hemoglobin concentration was determined using a commercial colorimetric assay (Randox Haemoglobin; Randox Laboratories Ltd., UK). All hemoglobin solutions were then diluted to a concentration of 1 pmol/ μ L. A desalting step was performed using

Table 1 Sample list

Sample name	Hb name	Chain	Symbol	Position	HbA	Hb variant	Mass shift
HbA Control	HbA	beta	β				
Hb E7K	HbG-Siriraj	beta	β	7	E	K	-0.942
Hb E26K	HbE	beta	β	26	E	K	-0.942
Hb E121K	HbO-Arab	beta	β	121	E	K	-0.942
Hb A142T	Hb Inglewood	beta	β	142	A	T	30.026
Fusion protein gamma-beta	Hb Kenya	gamma-beta	$^{\Lambda}\gamma$ - β	Fusion protein	$^{\Lambda}\gamma$ 1 \rightarrow 81; β 86 \rightarrow 146		54.978

C₁₈ macro spin columns (Harvard Apparatus, Harvard, MA, USA). Briefly, hemoglobin was bound to the spin column with 5 % acetonitrile and 0.1 % formic acid. After washing with the same solution, Hb was eluted with 50 % acetonitrile and 0.1 % formic acid. Samples were then evaporated in a vacuum centrifuge and stored at -80 °C. For MS experiments, hemoglobin samples were dissolved in 50 % acetonitrile and 0.1 % formic acid to a concentration of 1 pmol/ μ L. Ubiquitin (Sigma-Aldrich, St. Louis, MO, USA) was dissolved in 50 % acetonitrile and 0.1 % formic acid to a concentration of 1 pmol/ μ L. This solution was used for internal recalibration.

MS parameter optimization on purified HbA sample

Top-down mass spectrometry measurements were performed on a quadrupole time-of-flight (qTOF) MS instrument (maXis HD; Bruker Daltonics GmbH, Bremen, Germany) with ETD capabilities using an electrospray ion source, as described previously [17]. Briefly, the dry gas temperature was set at 220 °C, dry gas flow at 4 L/min, and capillary voltage at 3 kV. The precursor charge state selection and isolation parameters were optimized for HbA β chain. The +24 charge state of β chain was selected, and the isolation target m/z value and the isolation window width were set at 662.5 and 3 m/z , respectively. Ion transfer time was set to 80 μ s and pre-pulse storage at 10 μ s. The orthogonal ion pulse for TOF analysis was 1.72 kV and the signal digitization at 4 GHz. The observed mass range was from 50 to 3000 m/z units, and the data were acquired in profile mode. Electron transfer dissociation mass spectrometry was performed as previously described [18]. The ETD settings were optimized for HbA β chain to yield the highest fragmentation efficiency. For the Hb β chain, the analyte ion accumulation time was set at 400 ms, the reagent injection time at 10 ms, and the extended reaction time at 0 ms. The Hb analysis was started by acquiring a MS signal of intact ubiquitin for internal calibration for 1 min, followed by the acquisition of the intact Hb without fragmentation for 1 min. Finally, the ETD fragmentation of a selected Hb precursor ion was activated using the ETD parameters mentioned above. Up to 17 min of tandem mass spectrometry data were acquired, resulting in approximately 650 mass spectra. Theoretically, only 60 μ L of Hb extract are needed for 20 min of

direct infusion. For practical reason, syringe was filled with 150 μ L of Hb extract.

Data processing

MS data analysis was performed using Data Analysis software version 4.2 (Bruker Daltonics). The following procedure was applied to all analyzed samples. First, the monoisotopic masses of the different charges states of intact ubiquitin from the averaged mass spectrum were used for internal recalibration of the data set. Then, an averaged spectrum of the Hb MS/MS ETD data was generated. The SNAP peak picking algorithm version 2.0 was applied on this spectrum with the following parameters: quality factor threshold at 0.4, S/N threshold at 0, absolute intensity threshold at 0, and maximum charge state at 23. The peak list of the monoisotopic m/z values of the ETD product ions was subsequently displayed with the following information: m/z , charge state, resolution, and abundance [17]. BioTools software version 3.2 (Bruker Daltonics) was used to automatically assign ETD tandem mass spectra and to get information about the sequence coverage. To validate and improve the automatically generated product ion assignments, manual annotation was performed to compare experimental and predicted signals. Open source software “MS product” was used to provide predicted ETD products list for each sample [19]. In addition, the intact mass of the investigated Hb chain was used to validate sequence variants. To obtain the intact mass, the mass spectra of the second acquisition segment were averaged and deconvoluted with the maximum entropy algorithm. The intact monoisotopic mass was then determined by applying the SNAP algorithm [17]. Hb variants contained in clinical samples were analyzed, and the data were processed (recalibration of the mass spectra and the generation of the peak lists) as described above.

Results and discussion

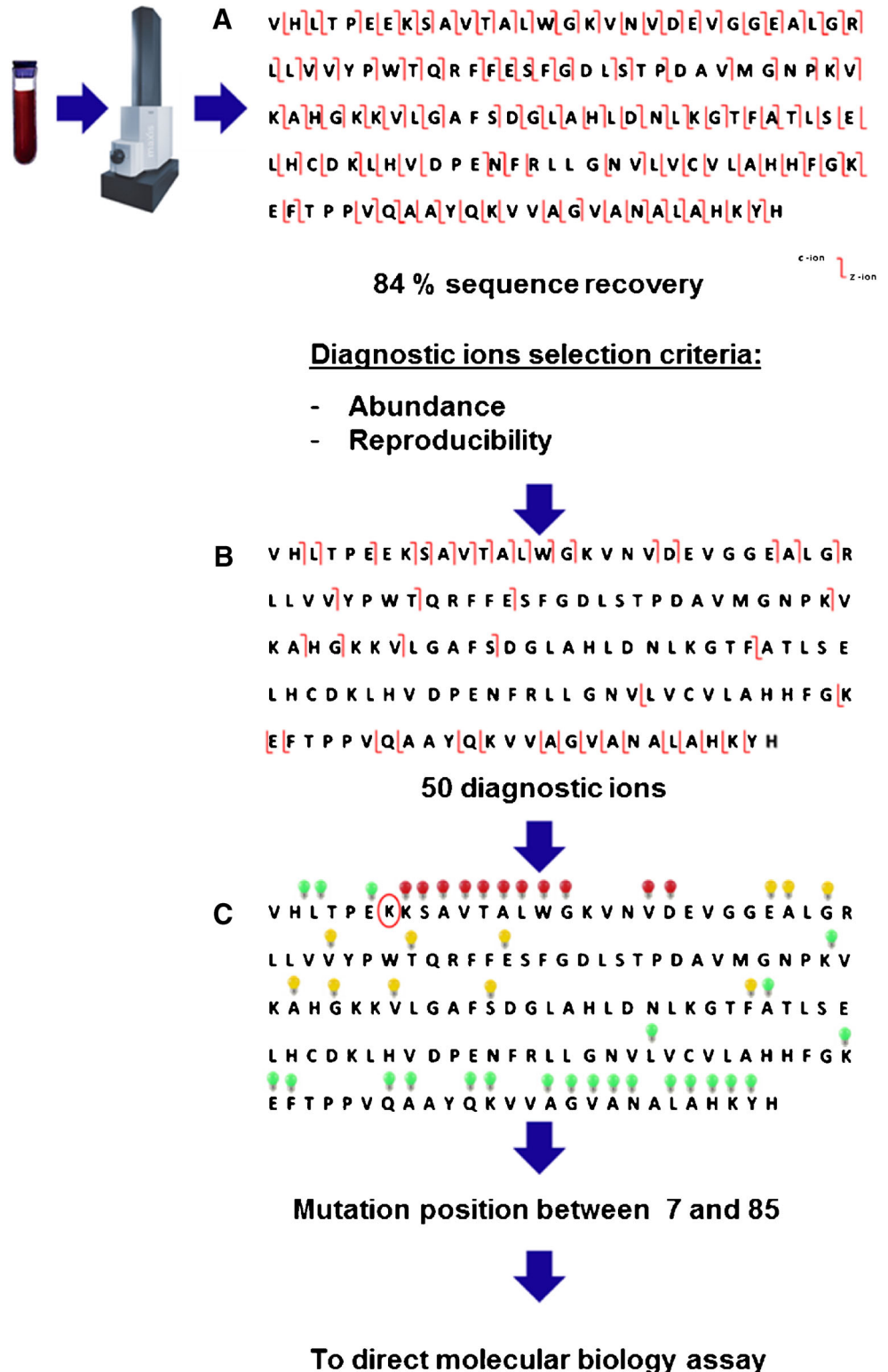
Precursor charge state selection and tandem MS optimization

Hb is a tetrameric protein composed of non-covalently bound chains. The denaturing condition (low pH and 50 % organic solvent) of the electrospray ionization process separates the

chains, which are then detected as two different proteins using the mass spectrometer. Consequently, the overlapping of different charge states of the two protein species does occur, making the efficient isolation of a single chain difficult for some charge states. Therefore, for MS/MS optimization,

different charge states were selected and the resulting mass spectra were evaluated according the following criteria: separation of the α chain (prevent co-fragmentation), precursor ion, and product ion abundances. To optimize product ion abundances, high charge states were selected preferentially

Fig. 1 Process to establish the diagnostic product ion list and color code for Hb β chain. **(A)** ETD sequence coverage for HbA β chain (*c*- and *z*-ions) obtained with BioTools. **(B)** Diagnostic ion sequence map. **(C)** Each label represents a diagnostic ion position in the following sequence: a *green label* if the ion was present, a *yellow label* if a manual input was needed, and a *red label* if the ion was not present or if the monoisotopic peak was shifted. The mutated position was immediately localized thanks to the yellow and/or red labels (in this example, the *highlighted zone* is between positions 7 and 85 and the mutation is at position 7)



as they are more appropriate for ETD activation. The charge states tested as precursor ions for activating the HbA β chain were as follows: $[M+19H]^{19+}=836.1154 m/z$, $[M+20H]^{20+}=794.3600 m/z$, $[M+24H]^{24+}=662.1346 m/z$, and $[M+25H]^{25+}=635.6895 m/z$. For each charge state, an optimization of the isolation efficiency and ETD parameters was done to obtain abundant product ions but preventing secondary reactions on them. In summary, the optimization consisted in testing different isolation target m/z values and isolation window widths. ETD parameters were optimized by testing different values to reach a ratio around 0.33 between precursor and highest abundant electron transfer product ion without dissociation (ETnoD). A peak list for each tested precursor ion was established by selecting most abundant product ions identified as c - or z -ions. Then, the peak lists were compared to select the precursor charge state that yielded the largest number of ETD product ions, well distributed along the sequence, and with high abundance. As a result, the precursor ion $[M+24H]^{24+}$ was selected.

Diagnostic product ion selection for Hb β chain

To simplify and speed up top-down ETD mass spectra interpretation and to reduce the need for manual data analysis, a reference ion list for the Hb β chain was established. The term “diagnostic product ion” was then used for these ions as they “reveal structural or compositional information about its precursor ion” [20]. The selection of diagnostic ions was based on several criteria. First, the most abundant ions were selected after mass spectra recalibration. Then, they were identified as c - or z -ions with the help of protein fragment mass calculators [19]. The correct peak assignment and charge state were verified by visual inspection of the isotopic cluster. At the end of this step, 105 ions were selected (Fig. 1). Second, the signal reproducibility of the selected product ions was evaluated. Four different samples were analyzed, and only the presence of these 105 diagnostic ions was verified. After this step, only 61 ions with a non-overlapping isotopic pattern and reproducible ion abundance were kept in the list. Third, the concept of using selected diagnostic ions from a previously established list was verified using patient’s samples. A few other ions were removed from the diagnostic ion list because their abundance was not reproducible in the performed experiments. At the end of this step, the diagnostic ion list for the β chain was established with 50 product ions, well distributed along the protein sequence (Fig. 1(B), Table 2). A color code was established to simplify data interpretation. The aim was to quickly detect the part of the sequence that was affected by the mutation. Each diagnostic ion was schematized in the protein sequence by a colored label (Fig. 1(C)). Some of these labels represent several ions; for example, the combination of several c_x -ions at different charge states or the c - and z - complementary ions.

Table 2 List of selected diagnostic product ions from ETD MS/MS of Hb β chain

Ion	z	m/z
c2	1	254.1612
z4	2	284.6413
z2	1	303.1214
z5	2	320.1598
c6	2	347.6978
c3	1	367.2452
z6	2	376.7018
z3	1	431.2163
c12	3	437.2419
c13	3	460.9209
z8	2	469.2419
c8	2	476.2665
c14	3	498.6156
z19	4	499.5226
z9	2	504.7604
c9	2	519.7826
z15	3	521.2956
z20	4	531.5372
z25	5	533.6892
c20	4	545.0574
z10	2	554.2946
c10	2	555.3011
z26	5	559.4978
c15	3	560.6420
z16	3	563.9818
z37	7	572.0225
c21	4	573.8141
c16	3	579.6492
z11	2	582.8054
z27	5	585.1168
c11	2	604.8353
z61	11	609.9592
z12	2	618.3239
c61	10	670.8544
z26	4	699.1204
c85	13	705.6062
c38	5	827.2504
c85	11	833.7151
c62	8	845.1955
c64	8	863.9543
c72	9	865.2392
c34	4	897.0004
c67	8	913.8628
c26	3	921.8152
c59	7	923.1937
z26	3	931.8247
c27	3	945.4943
c43	5	968.7182
z19	2	998.0379
c29	3	1002.1961

Analysis of β chain variants

The acquisition and processing parameters for the Hb variants were exactly the same as those for the HbA β chain sample except for a variant that has a significant mass shift that impacts its precursor m/z value. For this variant, the center of the isolation window and isolation window width was optimized. The presence of diagnostic ions was visually inspected. The results were used to annotate the globin chain sequence with a color code. If a diagnostic ion was present as in the control sample, the color was green. If the diagnostic ion was not present in the sample or if there was a shift of the monoisotopic peak, the color was red. The red color thus signifies that a sequence variation is detected between this position and the closest “green” label. If the signal was not well resolved or if there was an overlapping signal from a particular sample, the color was yellow. If a colored label was representing complementary c - and z -ions for the same position and if one of them was red or yellow, the resulting color of this position was also red or yellow, even if the second one was green. This rule was implemented in order not to lose information about a modification position. The method was applied to the analysis of four Hb

variants carrying a single-point mutation at different positions in the β chain (Table 1). Figure 2 shows the obtained results with the color code for the HbA β chain and for three of the Hb β chain variants: HbE, HbG-Siriraj, and HbO-Arab. For each sample, the mutation position was correctly assigned with the diagnostic ion strategy. All the diagnostic ions were present for the HbA β chain (all the labels were green). In contrast, for HbE, red and yellow labels were present from positions 22 to 85. This was expected since the single-point mutation is present at the position (26 Glu>Lys) of the β chain. A comparable result was obtained for HbG-Siriraj (7 Glu>Lys) and HbO-Arab (121 Glu>Lys). For these variants, the highlighted zones were between positions 7 and 85 and between positions 34 and 121, respectively. Figure 3 shows the result obtained for HbG-Siriraj and the mass spectrum of four diagnostic ions. All three Glu>Lys variants have a mass shift of -0.942 Da compared to the HbA β chain (Table 1), and the isolation m/z for the $+24$ ion of the variant differs only by $-0.942/24=0.0393$. For the Glu>Lys variants, the isolation m/z value are too close, and co-isolation as well as co-fragmentation of the variant and the HbA β chain cannot be avoided for heterozygote patient sample. For the HbO-Arab (121 Glu>Lys) sample (Fig. 2(D)), the

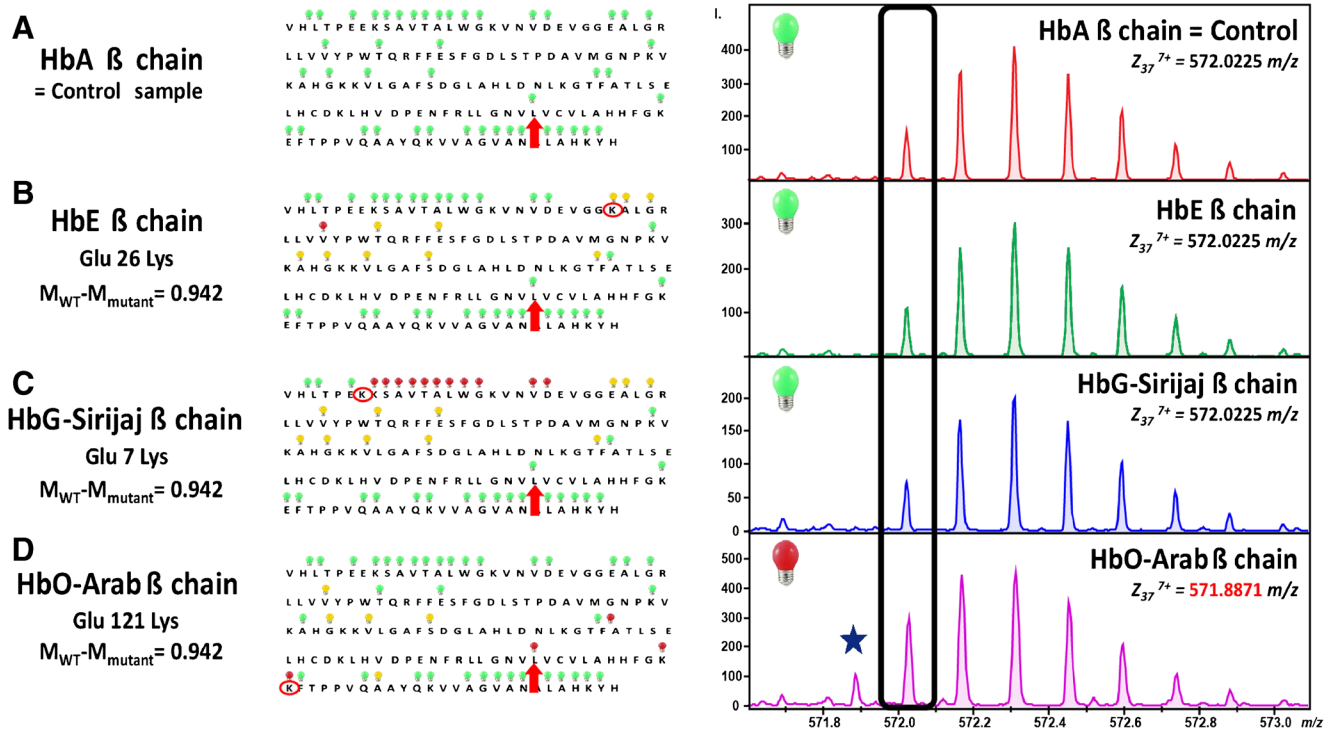


Fig. 2 Examples of Hb variant characterization. In the *left column*, the sample name, the mutation position, and the mass shift compared to HbA β chain are shown. In the *middle column*, β chain sequences are given together with the diagnostic ion color code representation. A *green label* is present if the ion was detected, a *yellow label* is present if manual input was needed, and a *red label* is present if the product ion was not detected or if the monoisotopic peak was shifted. The *red circle* shows the mutation position, and the *red arrow* shows the position of the z_{37}^{7+} ion in the protein sequence. In the *right column*, the mass spectra of the

diagnostic ion z_{37}^{7+} are displayed. The *black rectangle* shows the m/z position of the monoisotopic peak for the HbA β chain z_{37}^{7+} ion (*red trace*). For the *green and blue traces*, the monoisotopic peak is present in the *black rectangle* (this result is illustrated by a *green light*). For the *pink trace*, the monoisotopic peak is on the *left, outside the black rectangle*. This is indicated by a *red label*. This result shows that the chain carries a mutation located after this position, between residues 34 and 121. HbO-Arab is indeed mutated at position 121

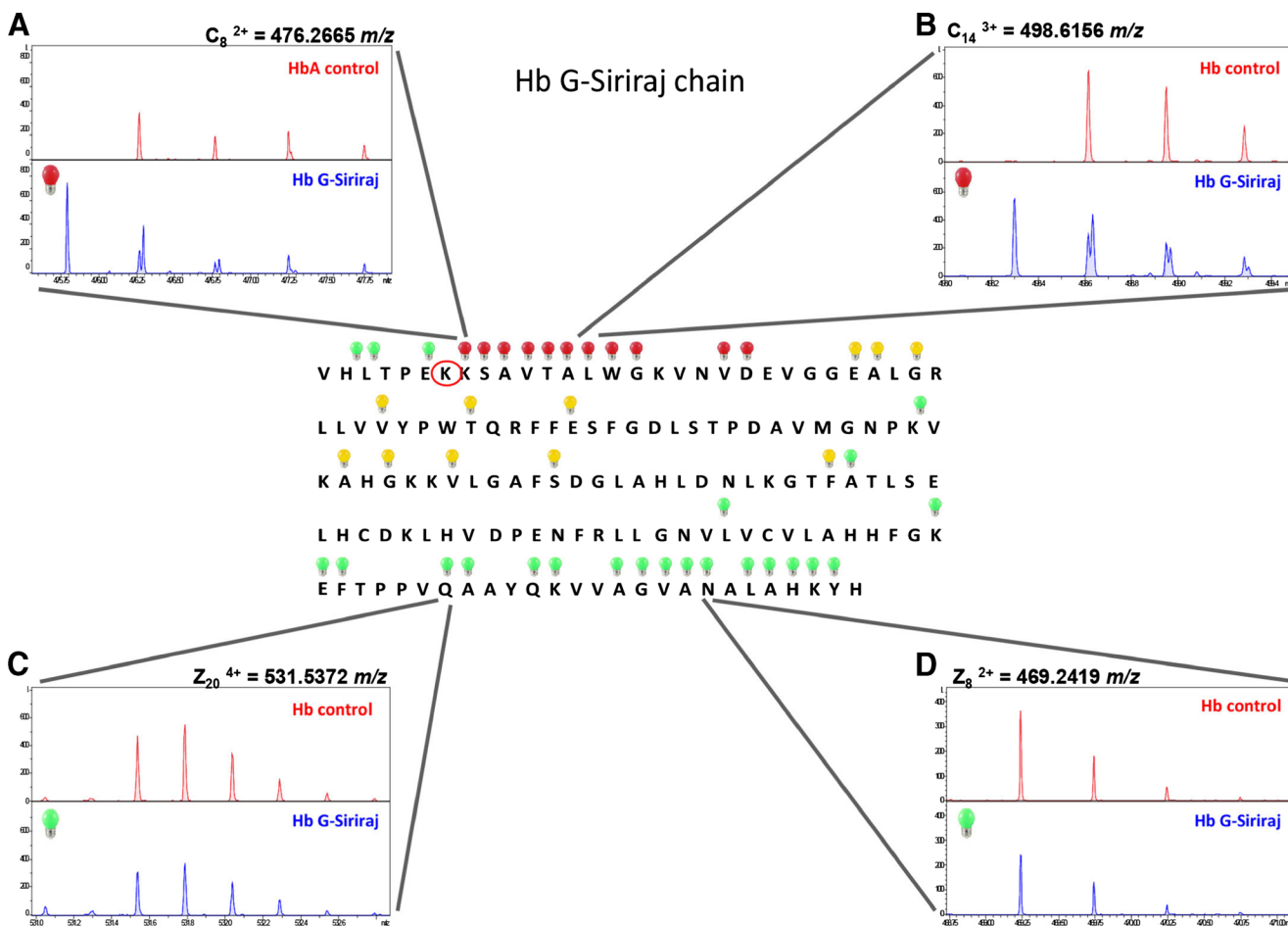


Fig. 3 HbG-Siriraj sequence and diagnostic ion color code result. The mutation position is highlighted with a red circle at position 7 (Glu>Lys) on the β chain sequence. The mass spectra of four diagnostic ions are highlighted. In the upper part, the diagnostic ion signal for the HbA β chain (red trace) is shown. In the lower part, the diagnostic ion signal for the analyzed sample (blue trace) is shown. (A) Diagnostic ion c_8^{2+} mass spectrum. In the lower part, the monoisotopic peak is shifted on the left compared to the control sample. This is illustrated with a red label. (B) Diagnostic ion c_{14}^{3+} mass spectrum. In the lower part, the monoisotopic

peak is shifted on the left compared to the control sample. This is illustrated with a red label. (C) The diagnostic ion z_{20}^{4+} yielded signal as the control sample. It is represented with a green label. (D) The diagnostic ion z_8^{2+} yielded signal as the control sample. It is represented with green label. As expected, the color code shows that the sequence does not correspond to the β chain from positions 7 to 85 (all the lights are red or yellow). The C-terminal part of the protein sequence corresponds to the normal β chain

monoisotopic m/z value for z_{37}^{7+} ion is at 571.8871 m/z and is shifted towards lower m/z value as indicated by the asterisk. In addition, HbA β chain is also present as a double peak is present at the monoisotopic m/z value for this diagnostic ion. The result of another Hb variants, Hb Inglewood (142 Ala>Thr), is presented in Electronic Supplementary Material (ESM) Fig. S1. The mass shift of this Hb variant was 30.03 Da (Table 1). Thus, the centers of the isolation window and isolation window width were optimized for this sample. The m/z values of the diagnostic ions were also adapted, taking into account the detected mass shift.

Analysis of a fusion protein—Hb Kenya

The diagnostic ion method was also tested in the case of a gene fusion protein, Hb Kenya. The gene fusion results from

a crossover between codons 81 and 86 of the γ and β globin genes [21]. This mutation leads to a Hb variant having the same sequence as the γ chain from positions 1 to 81 and the same sequence as the β chain from positions 86 to 146. Results presented in Fig. 4 show that all the labels from positions 1 to 86 were red or yellow, meaning that the sequence of N-terminal part of the protein is not identical to the HbA β chain. From positions 120 to 146, the labels are green or yellow. Consequently, the C-terminal part of the protein seems to correspond to the normal HbA β chain sequence.

The diagnostic ion list was not yet established for the γ chain. Thus, in order to confirm the identification of Hb Kenya by MS, a more detailed analysis was done for this sample. Specific ETD product ions from the N-terminal portion of the γ chain were manually annotated in the mass spectrum. Combining the information of precursor m/z ratio, the presence of

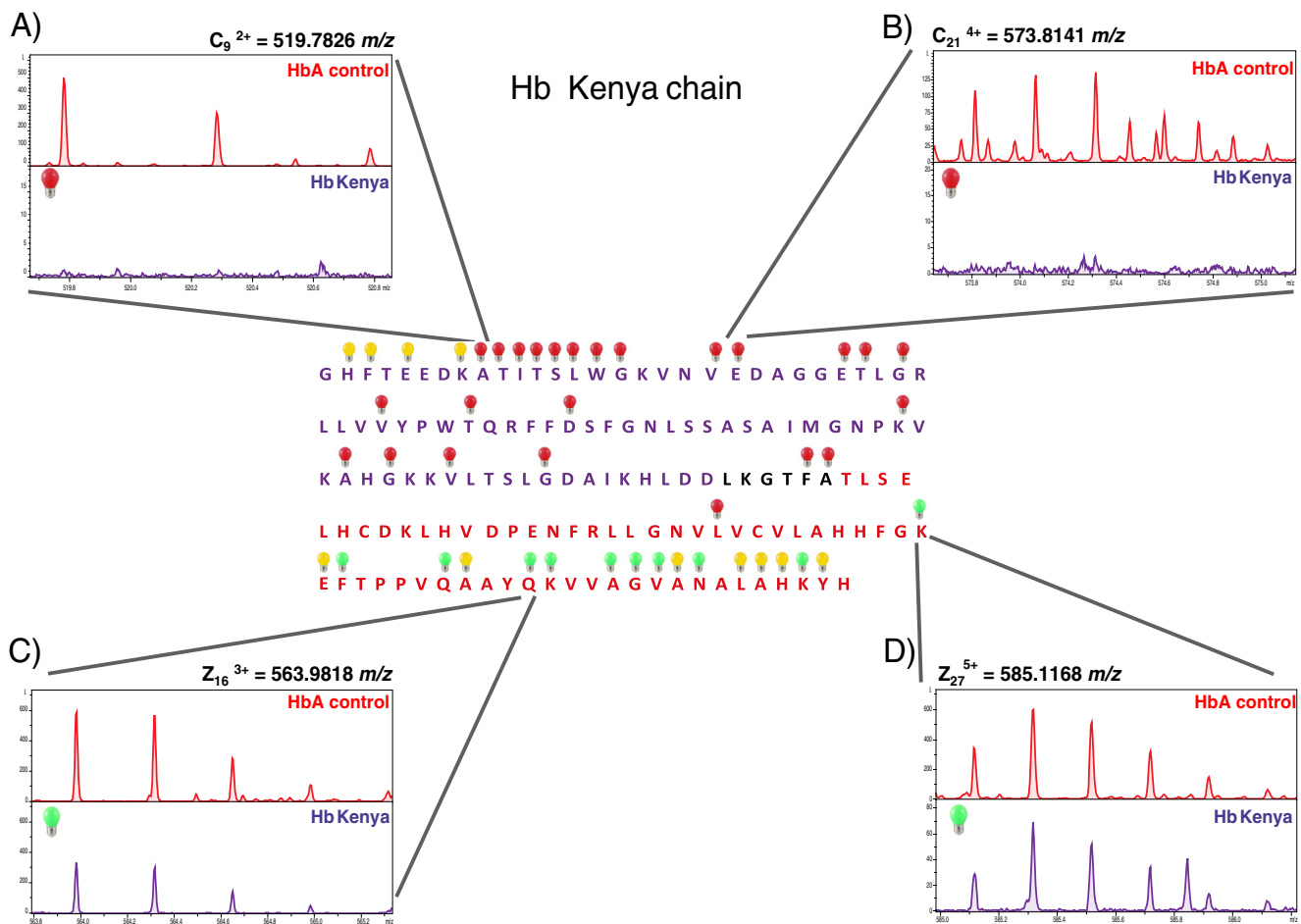


Fig. 4 Hb Kenya (fusion protein) sequence and diagnostic ion color code result. The Hb Kenya sequence is represented with *two colors*. From positions 1 to 81, the γ chain sequence is shown in *purple*. From positions 86 to 146, the β chain sequence is shown in *red*. The mass spectra of four diagnostic ions are *highlighted*. In the *upper part*, the diagnostic ion signal for the HbA β chain is shown. In the *lower part*, the diagnostic ion signal for the analyzed sample is shown. **(A)** No signal was detected for the diagnostic ion c_9^{2+} , which is illustrated with a *red label*. **(B)** No signal was detected for the diagnostic ion c_{21}^{4+} , which is

represented with a *red label*. **(C)** The diagnostic ion z_{16}^{3+} yielded signal as the control sample. It is represented with a *green label*. **(D)** The diagnostic ion z_{27}^{5+} yielded signal as the control sample. It is represented with *green label*. As expected, the *color code* shows that the N-terminal part of the sequence does not correspond to the HbA β chain (all the lights are *red or yellow*). The C-terminal part of the protein sequence corresponds to the normal β chain. Therefore, all lights are *green or yellow*, from positions 120 to 146

product ions from the N-terminal portion of the γ chain, and the presence of product ions from the C-terminal portion of the β chain, the presence of Hb Kenya could be confirmed.

Conclusion

We developed a top-down ETD MS method for the analysis of Hb β chain variants. A simplified process based on selected diagnostic product ions was used for MS data interpretation, which allowed fast and reliable localization of the mutated position. This procedure brings more precise information at the protein level than any of the nonspecific protein analysis methods used for the conventional Hb disorder diagnostic workflow. Importantly, in the perspective of an application

for Hb disorders diagnosis, this data analysis method allows a non-specialized person to use top-down ETD MS data. We showed that the method could also be applied for complex proteoforms such as fusion proteins. Once the mutation is localized within the globin chain sequence, a manual analysis of the corresponding diagnostic ion signals is necessary to precisely characterize the underlying amino acid changes. However, the information provided by the diagnostic ions can be used, in combination with other data such as the ethnic origin of the patients, to restrict the number of possible variants and to direct molecular biology analysis. Diagnostic product ion lists for the HbA α , δ , and γ chains remain to be determined. Automation of the color code strategy should also be done. Dried blood spots (DBS) will be investigated as a sample format for the future. Preliminary tests showed that

the analysis of Hb extract from DBS should be fully compatible with this top-down ETD MS method. DBS were put into an Eppendorf containing 500 μ L of 50 % acetonitrile and 0.1 % of formic acid solution during 20 min. The syringe was filled with the resulting Hb extract, and direct infusion was performed. Results from DBS spots were similar to those obtained using the standard sample preparation procedure described in the article (data not shown). Finally, this top-down ETD MS method is complementing our previously described MS methods for the identification of common Hb variants and quantification of Hb chains [13, 22].

Acknowledgments This study was supported by the Swiss National Science Foundation Grant No. 32003B_143809 and by the Ernst & Lucie Schmidheiny Foundation. The authors thank Paola Antinori, Fabienne Jeanneret, and HuiSong Pak for their scientific support and encouraging discussions.

References

- Modell B, Darlison M (2008) Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ* 86:480–487
- Old JM (2003) Screening and genetic diagnosis of haemoglobin disorders. *Blood Rev* 17(1):43–53
- A database of human hemoglobin variants and thalassemias. Available from: <http://globin.cse.psu.edu/hbvar/manu.html>
- Zhang Z, Wu S, Stenoien DL, Pasa-Tolic L (2014) High-throughput proteomics. *Annu Rev Anal Chem* 7(1):427–454
- Th  berge R, Infusini G, Tong W, McComb ME, Costello CE (2011) Top-down analysis of small plasma proteins using an LTQ-Orbitrap. Potential for mass spectrometry-based clinical assays for transthyretin and hemoglobin. *Int J Mass Spectrom* 300(2–3):130–142
- Smith LM, Kelleher NL, Consortium for Top Down Proteomics (2013) Proteoform: a single term describing protein complexity. *Nat Meth* 10(3):186–187
- Zubarev RR, Horn DM, Fridriksson EK, Kelleher NL, Kruger NA, Lewis MA, Carpenter BK, McLafferty FW (2000) Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal Chem* 72(3):563
- Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* 101(26):9528–9533
- Boemer F, Cornet Y, Libiouille C, Segers K, Bours V, Schoos R (2011) 3-years experience review of neonatal screening for hemoglobin disorders using tandem mass spectrometry. *Clin Chim Acta* 412(15–16):1476–1479
- Daniel YA, Tumer C, Haynes RM, Hunt BJ, Dalton RN (2005) Rapid and specific detection of clinically significant haemoglobinopathies using electrospray mass spectrometry–mass spectrometry. *Br J Haematol* 130(4):635–643
- Hachani J, Duban-Deweere S, Pottiez G, Renom G, Flahaut C, P  rini JM (2011) MALDI-TOF MS profiling as the first-tier screen for sickle cell disease in neonates: Matching throughput to objectives. *Proteomics Clin Appl* 5(7–8):405–414
- Edwards RL, Creese AJ, Baumert M, Griffiths P, Bunch J, Cooper HJ (2011) Hemoglobin variant analysis via direct surface sampling of dried blood spots coupled with high-resolution mass spectrometry. *Anal Chem* 83(6):2265–2270
- Coelho Gra  a D, Lescuyer P, Clerici L, Tsybin YO, Hartmer R, Meyer M, Samii K, Hochstrasser DF, Scherl A (2012) Electron transfer dissociation mass spectrometry of hemoglobin on clinical samples. *J Am Soc Mass Spectrom* 23(10):1750–1756
- Edwards RL, Martin NJ, Cooper HJ (2013) Hemoglobin variant analysis of whole blood and dried blood spots by MS. *Bioanalysis* 5(16):2043–2052
- Edwards RL, Griffiths P, Bunch J, Cooper HJ (2014) Compound heterozygotes and beta-thalassemia: top-down mass spectrometry for detection of hemoglobinopathies. *Proteomics* 14(10):1232–1238
- Edwards RL, Griffiths P, Bunch J, Cooper HJ (2012) Top-down proteomics and direct surface sampling of neonatal dried blood spots: diagnosis of unknown hemoglobin variants. *J Am Soc Mass Spectrom* 23(11):1921–1930
- Fornelli L, Parra J, Hartmer R, Stoermer C, Lubeck M, Tsybin YO (2013) Top-down analysis of 30–80 kDa proteins by electron transfer dissociation time-of-flight mass spectrometry. *Anal Bioanal Chem* 405(26):8505–8514
- Hartmer RG, Kaplan DA, Stoermer C, Lubeck M, Park MA (2009) Data-dependent electron transfer dissociation of large peptides and medium size proteins in a QTOF instrument on a liquid chromatography timescale. *Rapid Commun Mass Spectrom* 23(15):2273–2282
- ProteinProspector. Available from: <http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct>
- Price P (1991) Standard definitions of terms relating to mass spectrometry: a report from the committee on measurements and standards of the American Society for Mass Spectrometry. *J Am Soc Mass Spectrom* 2(4):336–348
- Wilcox I, Boettger K, Greene L, Malek A, Davis L, Steinberg MH, Luo HY, Chui DH (2009) Hemoglobin Kenya composed of α - and ($A\gamma\beta$)-fusion-globin chains, associated with hereditary persistence of fetal hemoglobin. *Am J Hematol* 84(1):55–58
- Acosta-Martin AE, Coelho Gra  a D, Antinori P, Clerici L, Hartmer R, Meyer M, Hochstrasser D, Samii K, Lescuyer P, Scherl A (2013) Quantitative mass spectrometry analysis of intact hemoglobin A2 by precursor ion isolation and detection. *Anal Chem* 85(16):7971–7975