Journal of Chromatography, 562 (1991) 175–190
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 5565

Electrospray mass spectrometry in the clinical diagnosis of variant hemoglobins

C. H. L. SHACKLETON*

Clinical Mass Spectrometry Facility, Children's Hospital Research Institute, 747 52nd Street, Oakland, CA 94609 (U.S.A.)

A. M. FALICK

Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143 (U.S.A.)

B. N. GREEN

VG BioTech, Tudor Road, Altrincham, WA14 5RZ Cheshire (U.K.)

and

H. E. WITKOWSKA

Clinical Mass Spectrometry Facility, Children's Hospital Research Institute, 747 52nd Street, Oakland, CA 94609 (U.S.A.)

ABSTRACT

A combination of mass spectrometric (MS) techniques [electrospray MS, liquid secondary ion MS (LSIMS) and MS–MS] has been used for variant hemoglobin (Hb) detection and characterization. Electrospray MS allowed analysis of mixtures of intact globins giving simultaneously the molecular weights (accuracy 1–2 Da) and information about relative amounts of globins present. Currently, 14 Da is the minimum molecular weight difference required experimentally to accurately measure different species present in a mixture of 15–16 kDa proteins. Thus 80 and 79% of the known variants of α and β chains, respectively, can be detected in mixtures with their normal counterparts, including Hb S (molecular weight difference = 30 Da). Abnormal hemoglobins detected were fractionated by C_4 reversed-phase high-performance liquid chromatography (HPLC), and the separated globin chains (or the mixture of whole precipitated globin) were digested by trypsin. The tryptic peptides were separated by C_{18} reversed-phase HPLC and analyzed by LSIMS to narrow down the mutation site to a single peptide. The mass measured in LSIMS frequently corresponded to a unique structure, thus giving the unequivocal identification of the mutation and its site. Where there was ambiguity, tandem MS on a Kratos Concept four-sector instrument was used for sequencing the abnormal peptide. The practical use of the methodologies presented is illustrated through analysis and identification of Hb G-San Jose, Hb Stanleyville II, Hb S and Hb Willamette.

0378-4347/91/$03.50

INTRODUCTION

Human hemoglobin (Hb) is a tetramer of two α chains (141 amino acids) and two non-α chains: β, δ or γ (146 amino acids), each of them associated with a prosthetic heme group. Both fetal and adult hemoglobin types are built of α chains (coded by two gene copies on chromosome 16) and different non-α chains: $^{G}\gamma$ and $^{A}\gamma$ in fetal Hb F, β in adult HbA and δ in HbA_2, a minor component of adult hemoglobin (2%). All non-α globin chains are coded by single copy genes forming a β-type globin gene cluster on chromosome 11.

Hemoglobin variants with amino acid substitutions in the protein chain generally arise as a consequence of single-base substitutions within globin coding gene loci [1]. There are around 400 variants of known structure described so far and presumably many more remain undetected [1]. Although many known variants are not associated with any apparent clinical manifestations, the severity of the conditions produced by the clinically relevant variants, especially when present in homozygous state (*i.e.* Hb SS), accompanied by thalassemia or in heterozygous mixture with non-compatible counterparts (*i.e.* Hb SC, Hb SD-Los Angeles and Hb SO-Arab), justifies the growing interest in variant hemoglobin detection and identification. When a rare or unknown variant is found, usually by altered mobility of an electrophoretic band, attempts at isolation and characterization are made.

Characterization of globin chains does not differ in approach from that applied for any other protein. Homogenous samples of individual α and β chains [separated by high-performance liquid chromatography (HPLC)] are selectively cleaved (*e.g.* by trypsin), and the individual short peptides are separated (by HPLC) and sequenced. In practice, since the normal globin sequence is so well known, usually only one or two peptides have unusual HPLC mobility necessitating identification. In early studies abnormal peptides were sequenced by Edman degradation but latterly mass spectrometry (MS) has become a method of choice for analysing the complete tryptic peptide digests. Since the mass differentials between normal and abnormal globin chains generally reflect a single amino acid change, a table can be constructed of the amino acid changes compatible with the mass changes (Table I). In the preparation of this table it was assumed that only single nucleotide mutation in the coding triplet was possible.

Wada *et al.* [2] first utilized MS techniques for analysis of variant hemoglobins in 1981 using field desorption (FD). But with the introduction of liquid secondary ion MS [LSIMS, including fast atom bombardment (FAB)], FD was displaced and all workers in the field have since utilized the LSIMS technique [3–7].

While MS provides a very powerful tool for variant identification, some of the procedures it entails are laborious and time-consuming. A primary purpose of this paper is to describe the utility of electrospray MS [8–10] in these studies as well as the recent progress made in rapid peptide sequencing using tandem mass spectrometry (MS–MS) with array detection.

TABLE I

MASS DIFFERENCE BETWEEN AMINO ACID RESIDUES THAT ARE ALTERABLE THROUGH SINGLE-NUCLEOTIDE MUTATION IN CODING TRIPLET

Nominal mass difference	Amino acid change	Nominal mass difference	Amino acid change	Nominal mass difference	Amino acid change
0	Ile→Leu	23	Asn→His	48	Val→Phe
0	Gln→Lys			48	Asp→Tyr
		24	Leu→His		
1	Lys→Glu			49	Asn→Tyr
1	Ile→Asn	25	Met→Arg		
1	Asn→Asp			53	Cys→Arg
1	Gln→Glu	26	His→Tyr		
		26	Ala→Pro	55	Thr→Arg
3	Lys→Met	26	Ser→Leu		
				58	Gly→Asp
4	Pro→Thr	27	Ser→Asn	58	Ala→Glu
		27	Thr→Lys		
9	Gln→His			59	Pro→Arg
		28	Lys→Arg		
10	Ser→Pro	28	Ala→Val	60	Ser→Phe
		28	Gln→Arg	60	Cys→Tyr
12	Thr→Ile				
12	Thr→Leu	30	Val→Glu	69	Ser→Arg
		30	Arg→Trp		
13	Thr→Asn	30	Thr→Met	72	Gly→Glu
		30	Gly→Ser		
14	Gly→Ala	30	Ala→Thr	73	Leu→Trp
14	Ser→Thr				
14	Val→Leu	31	Pro→Gln	76	Ser→Tyr
14	Asp→Glu				
14	Asn→Lys	32	Val→Met	83	Cys→Trp
15	Leu→Gln	34	Leu→Phe	99	Ser→Trp
15	Ile→Lys	34	Ile→Phe	99	Gly→Arg
16	Val→Asp	40	Pro→His	129	Gly→Trp
16	Ser→Cys				
16	Ala→Ser	42	Gly→Val		
16	Phe→Tyr				
16	Pro→Leu	43	Leu→Arg		
		43	Ile→Arg		
18	Leu→Met				
18	Ile→Met	44	Ala→Asp		
		44	Cys→Phe		
19	His→Arg				
		46	Gly→Cys		
22	Asp→His				

EXPERIMENTAL

Overall strategy

If available, electrospray MS should be used in the preliminary steps of abnormal hemoglobin detection and identification. Analytical systems such as ion-exchange chromatography, reversed-phase chromatography, electrophoresis and isoelectric focusing that separate proteins according to differences in charge and/or hydrophobicity should be employed to detect the presence of variant globin chains that differ from their normal counterparts by only a few daltons, including the common Glu→Gln, Asp→Asn and Glu→Lys mutations.

Final identification of a variant hemoglobin requires that the nature and site of the mutation be determined. The initial step in this procedure is analytical and/or preparative globin chain separation, carried out using reversed-phase HPLC on a C_4 column. Hemolysates or precipitated whole globins can be used; very good separation of all normal globin chains (α, β, δ, $^{G}\gamma$, $^{A}\gamma$) from each other and from heme is achieved [11].

The second step we take is trypsinization of either an isolated variant globin chain or a mixture of precipitated globins. It is advisable to perform trypsinization of non-derivatized globins first [12]. Even though cysteine-containing insoluble 'core' peptides βT 12 and αT 12 cannot be mapped this way, the remaining part of the globin structures (76% of α and 85% of β chains) can be examined easily and the positive identification of an abnormal peptide can be achieved in a short time. If no mutation is detected by this simplified approach, aminoethylation of cysteines, followed by trypsinization, is performed to provide information about the 'core' regions.

The third step is a reversed-phase HPLC separation and LSIMS examination of tryptic peptides. As a routine approach, we attempt to achieve the full separation of peptides to allow MS analysis of highly purified species. In certain instances, when the variant is present in at least 20% and the analysis of electrospray MS and isoelectric focusing data permits narrowing down the number of possible mutations to only a few, a mixture of peptides of like hydrophobicity is examined. The HPLC–LSIMS fingerprints obtained in either case are compared to those of normal globin. The information about peptide mobility is important ancillary information, especially in the case of Glu→Gln and Asp→Asn mutations. Frequently the mass measured in LSIMS corresponds to a unique single amino acid substitution, thus giving unequivocal identification of the mutation site and nature. Where there is ambiguity, tandem MS on the Kratos Concept instrument is used for sequencing the abnormal peptide.

Specific techniques

Thin-layer isoelectric focusing. A pH 6–8 gradient 1-mm agarose gel system purchased from Isolab was used. Hemolysates (1–2 μl) containing approximately 10 μg hemoglobin per μl of 0.05% potassium cyanide were electrophoresed at 30

W for 90 min. Gels were fixed in 10% trichloroacetic acid (TCA) and then stained with *o*-dianisidine.

S-Aminoethylation and trypsinization of globins. S-Aminoethylation of globins was performed according to a published method [13]. S-Derivatized and non-derivatized globins were trypsinized with L-1-tosylamide-2-phenylethyl chloromethyl ketone-treated trypsin (6 μg per 150 μg protein) in 50 m*M* ammonium bicarbonate. Hemolysates were prepared from blood washed three times with saline by lysis of packed red cells with water (1–2 volumes) followed by centrifugation. Globins were precipitated with acetone–0.7% hydrochloric acid. Tryptic digests (pH adjusted to 3 with glacial acetic acid) were stored at −20°C before further analyses.

HPLC. Separations of globin chains [11] and tryptic peptides [7] were done on Vydac reversed-phase semi-preparative C_4 and analytical C_{18} columns (300 Å, 5 μm), respectively, development with an acetonitrile–water–trifluoroacetic acid (TFA) gradient and detection at 220 nm (for details, see legend to Fig. 6) and 214 nm (for details, see legend to Fig. 7), respectively.

Mass spectra. LSI mass spectra were obtained on a VG 30-250 quadrupole mass spectrometer (controlled by a Digital PDP-11/73 data system), the xenon beam being operated at 8 keV. The instrument was calibrated with Cesium iodide. Scans were acquired between m/z 500 and 2200 (2 s cycle time). HPLC fractions were evaporated by vacuum centrifugation and the peptides (1–2 nmol) were redissolved in 2 μl of 0.1 *M* heptafluorobutyric acid and mixed with 2 μl of the glycerol–thioglycerol (1:1)–0.1 *M* hydrochloric acid matrix on the tip of the probe.

Tandem MS. Experiments were carried out on a Kratos Concept II HH four-sector instrument of EBEB geometry, equipped with a multi-channel array detector capable of acquiring data simultaneously over a 4% mass window [14]. Parent ions were generated using an LSIMS source employing an 18-keV cesium ion primary beam. The collision energy for collision-induced dissociation (CID) was 6 keV. The collision gas (helium) was used at a pressure sufficient to suppress the parent ion beam to about 30% of its initial value. The instrument was controlled and data were acquired with a DS-90 data system. Calibration and data display were carried out with a Mach 3 data-processing system.

Electrospray MS. The intact globins were analyzed on a VG Bio-Q mass spectrometer, an instrument that consists essentially of an electrostatic spray ion source operating at atmospheric pressure followed by a quadrupole mass analyzer. Samples were introduced into the ion source in a 50:50 methanol–water solution containing 2.5% acetic acid at concentrations between 30 and 100 pmol/μl. The flow-rate of the sample solution into the ion source was 2 μl/min. Typically, about 300 pmol of the globin was used for each analysis so that components present at the 1% level could be measured. The mass spectrometer was scanned from m/z 700 to 1500 in 20 s at unit mass resolution, and several scans were summed to obtain the final spectrum. Mass-scale calibration employed the mul-

tiply charged ions from a separate introduction of myoglobin (average MW 16950.6).

RESULTS

Normal adult hemoglobin

Electrospray mass spectra of hemoglobins exhibit protonated molecular ions having between 11 and 20 charges, from which molecular masses are measured. A repeating sequence of groups of separated molecular ions is seen which represent all globin species present. Fig. 1 shows an electrospray mass spectrum of globins precipitated from a red cell lysate from normal adult. The major components are α^A and β^A globins. The figure insert, where the region of $\alpha(13+)$–$\beta(14+)$ globins is displayed, shows that minor components such as glycated α and β chains are easily detected. Table II summarizes results from several separate analyses of precipitated globins from lysates obtained from red cells of normal adults (Fig. 1), normal infants and carriers of sickle cell anemia hemoglobin [15]. The results presented demonstrate that the accuracy of measurements is extremely good (routinely within ± 1 Da) for major components separated by more than 14 Da. The measurements of masses of minor components present at the level of 1–2% are normally achieved with accuracy of a couple of daltons. It is important that

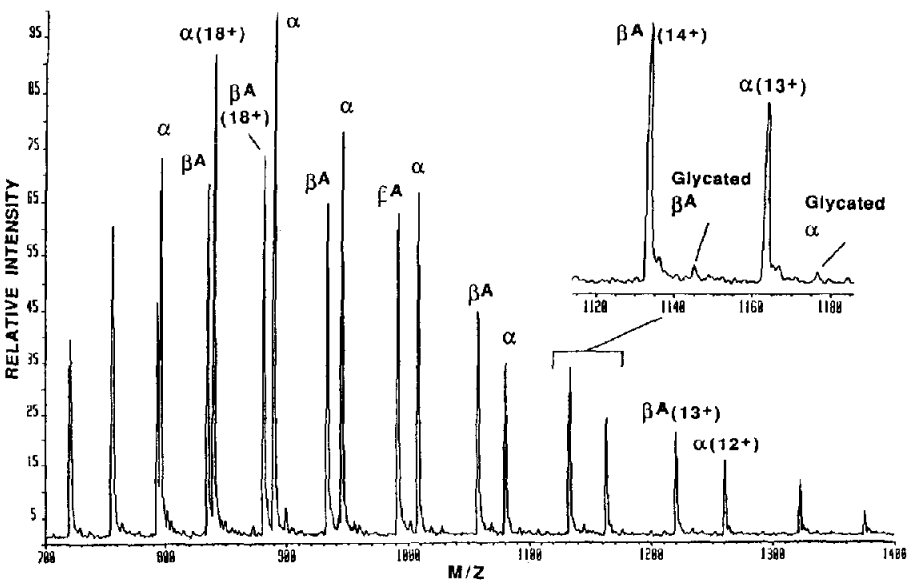

Fig. 1. Electrospray mass spectrum of intact globins from hemolysate of normal adult. Inset shows a magnification of the β^{14+}–α^{13+} region (Reproduced with permission).

TABLE II

MOLECULAR WEIGHTS MEASURED IN ELECTROSPRAY MS AND EXPECTED MOLECULAR WEIGHTS OF GLOBIN CHAINS

Globin assignment	Measured MW	Expected MW	Error	*n*
α	15 125.7 ± 0.7	15 126.4	−0.7	3
Glycated α	15 288.0 ± 0.2	15 288.5	−0.5	2
β^A	15 866.5 ± 0.3	15 867.2	−0.7	3
Glycated β^A	16 028.5 ± 0.7	16 029.4	−0.9	2
$^G\gamma$	15 995.8 ± 0.6	15 995.3	+0.5	3
$^A\gamma$	16 009.6	16 009.3	+0.3	1
β^S	15 838.2 ± 0.5	15 837.2	+1.0	3

globins from freshly prepared hemolysates are analysed. It is known that hemolysates undergo complex changes on storage, especially when kept at 4°C or −20°C for a prolonged period [16]. The resulting sample heterogeneity is likely to affect the mass measurement accuracy; more importantly, the presence of artifacts produced in the process of hemolysate ageing might lead to data misinterpretation.

Hb G-San Jose (β7 Glu→Gly)

A variant hemoglobin that showed an isoelectric focusing pattern typical of Hb G-San Jose was detected in an adult blood sample. The electrospray MS of a globin precipitate showed the presence of a β variant chain with molecular mass 72 Da lower than the normal β^A chain (Fig. 2). This mass difference could be explained by a Glu→Gly substitution, as in the case of the Hb G-San Jose variant (position 7, within βT1), or by a Trp→Asn substitution. However, the latter cannot be accounted for by either a one or two-point mutation and so is a rather unlikely event. Besides Hb G-San Jose, there are seven other possible variants of Glu→Gly type: substitution in βT1 at position 6 (unknown), in βT3 at position 22 (Hb G-Tapei) or position 26 (unknown), in βT5 at position 43 (unknown); in βT10 at position 90 (Hb Roseau-Benite), in βT11 at position 101 (Hb Alberta) or in βT13 at position 121 (unknown).

It has been our previous experience that the retention times of tryptic peptides separated by reversed-phase HPLC follow the elution order predicted by the 'Mac Gene^{TM1}' software program. The predicted retention times of all possible abnormal peptides derived from β chain variants with Glu→Gly substitution were calculated and it was shown that all variant peptides would precede their normal counterparts by about 1 min. Thus LSIMS examination of the fractions eluting close to normal βT1, βT3, βT5, βT10 and βT13 fragments should reveal the presence of a variant peptide easily distinguishable from its normal counterpart by a decrement of 72 Da. Therefore, separation of the abnormal chain was

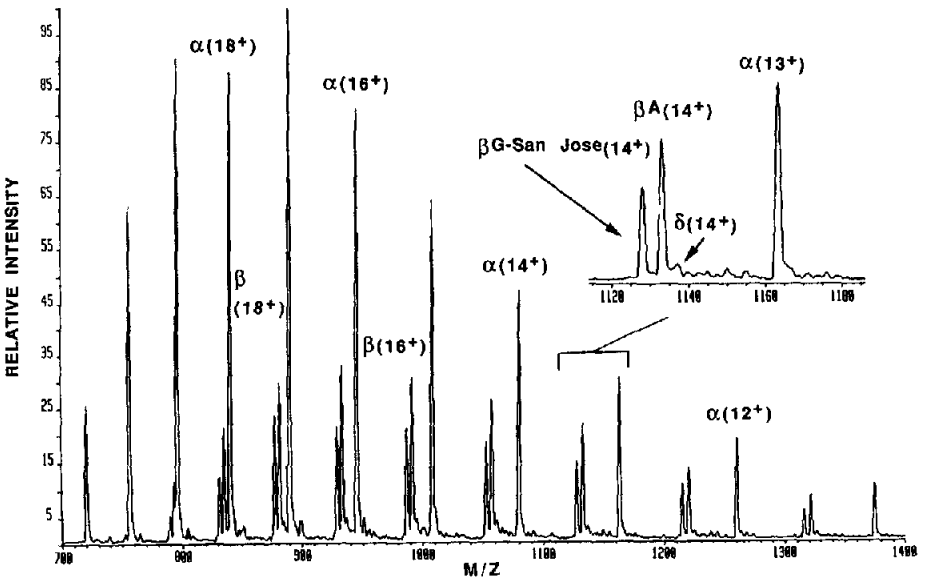

Fig. 2. Electrospray mass spectrum of precipitated whole globin from Hb G-San Jose (β7 E→G) carrier. The presence of α chain (15 126.4 Da), δ chain (15 924.3 Da), $\beta^{G\text{-San Jose}}$chain (15 795.2 Da) and β^{A} chain (15 867.2 Da) was detected.

not necessary and a tryptic digest of a mixture of precipitated globins was prepared. The variant βT1 peptide was isolated and its mass determined by electrospray MS (Fig. 3). Electrospray MS typically gives predominantly doubly charged molecular ions for peptides in this mass range and the doubly charged ion for peptide βT1 was at m/z 440, indicating a mass of 880. Since there are two glutamic acids within the normal βT1 fragment structure, sequencing was neces-

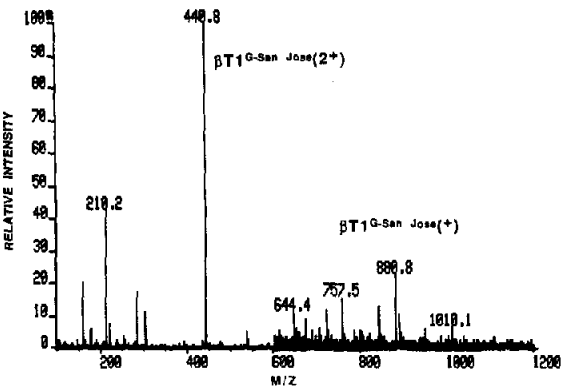

Fig. 3. Electrospray mass spectrum of variant $\beta^{G\text{-San Jose}}$T1 peptide ($MH^+$ = 880.4 Da). The doubly charged molecular ion [$(M+2H)^{2+}$ = 440.8 Da] is the most abundant species present.

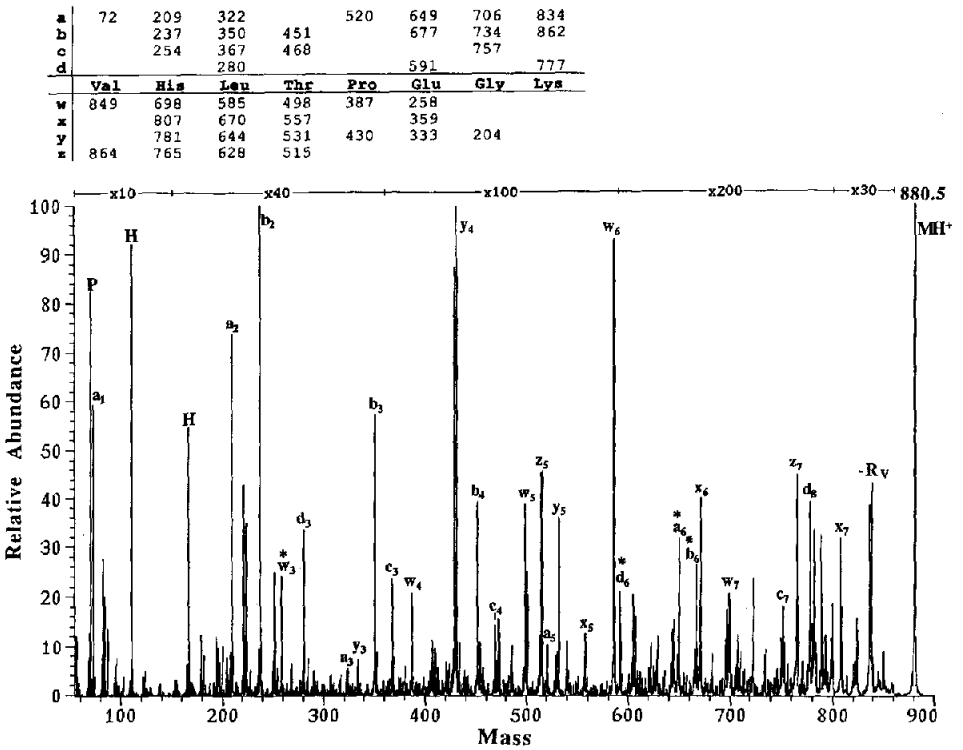

Fig. 4. CID fragment ion spectrum of $\beta^{G\text{-San Jose}}$TI peptide. The sequence proved to be VHLTPE*G*K, the sequence characteristic of the Hb G-San Jose mutant. The nominal masses of all of the sequence ions observed in the spectrum are listed in the grid at the top. Key ions showing the location of the E→G mutation (w_3, d_6, a_6 and b_6) are indicated with asterisks in the figure.

sary to establish the peptide structure. Sequencing of 'peptide 880' was performed on a Kratos Concept instrument and ions present in the CID spectrum (Fig. 4) unambiguously proved the sequence to be VHLTPE*G*K, the sequence characteristic of the G-San Jose variant.

Hb Stanleyville II (α78 Asn→Lys)

Isoelectric focusing of this cord blood sample suggested the presence of a β chain variant. Electrospray MS, besides revealing a variant β chain with molecular mass 30 Da lower than the normal β^A chain, clearly showed in addition the presence of an α chain variant with molecular mass 10 Da higher than α^A (Fig. 5).

Reversed-phase HPLC separation of globin chains confirmed the electrospray MS findings, showing the characteristic pattern for an individual doubly heterozygous for an α and β globin chain variant (Fig. 6). For the β chain variant, the measured mass difference (−30 Da) corresponded to a Glu→Val substitution and comparison with the known elution profile of a mixture of β^A and β^S chains

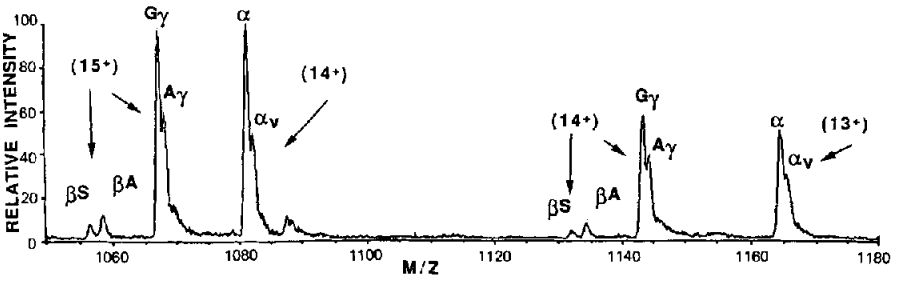

Fig. 5. Part of an electrospray mass spectrum of precipitated whole globin from cord blood of an individual doubly heterozygous for Hb S (β6 E→V) and Hb Stanleyville II (α78 N→K). The presence of α (15 125.9 ± 0.5 Da), $\alpha^{\text{Stanleyville II}}$ (15 135.6 ± 2.6 Da), β^{S} (15 836.4 ± 0.4 Da), β^{A} (15 866.6 ± 0.7 Da), $^{G}\gamma$ (15 995.1 ± 0.5 Da) and $^{A}\gamma$ (16 003.1 ± 0.5 Da) was observed.

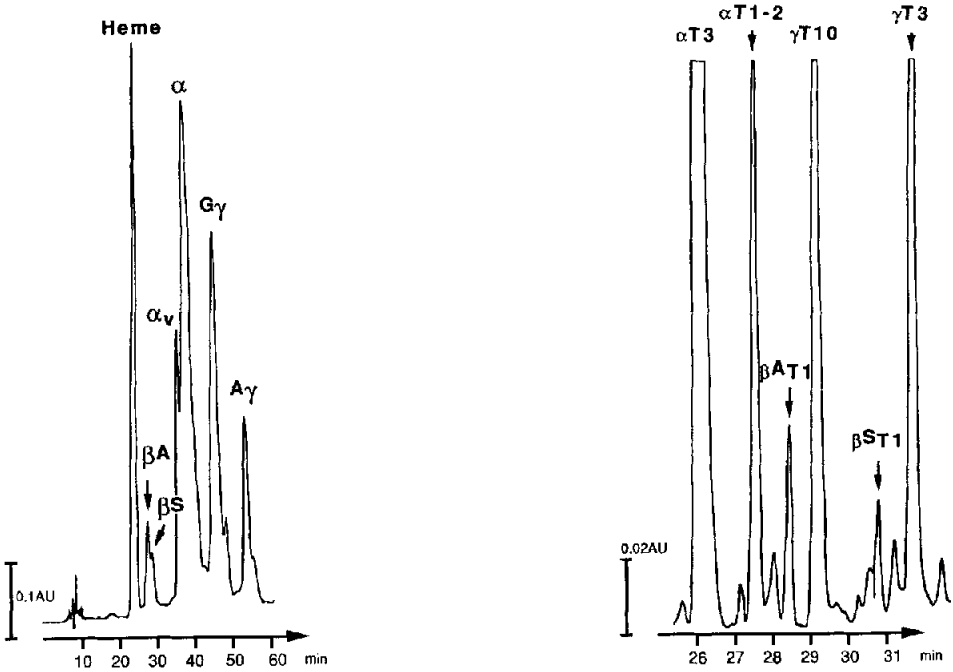

Fig. 6. Reversed-phase HPLC separation of globin chains from hemolysate containing Hb F, Hb A, Hb S and Hb Stanleyville II. Column, C_4 Vydac (25 cm × 1 cm I.D.); flow-rate 1.7 ml/min; detection 220 nm; linear gradient from 44% B to 56.5% B in 60 min. A = acetonitrile–water (20:80)–0.1% TFA; B = acetonitrile–water (60:40)–0.09% TFA.

Fig. 7. Part of a chromatogram of tryptic peptides derived from a mixture of globins containing β^{A}, β^{S}, $^{G}\gamma$, $^{A}\gamma$, α^{A} and $\alpha^{\text{Stanleyville II}}$. The positions of peptides βT1 (MH^+ = 952.5 Da) and β^{S}T1 (MH^+ = 922.5 Da) are shown. Column, C_{18} Vydac (25 cm × 0.4 cm I.D.); flow-rate 1 ml/min; detection 214 nm, isocratic development with A for 5 min followed by a linear gradient to 78% B in 90 min. A = water–0.125% TFA; B = acetonitrile–water (90:10)–0.125% TFA.

suggested the presence of a β^S variant. Following tryptic digestion of the precipitated whole globin mixture and separation of fragments, a tryptic fragment (MH^+ = 922 Da) eluting at the position previously found for the β^ST1 peptide was revealed (Fig. 7). The presence of Hb S in the blood was also consistent with electrophoretic data.

A sample of the variant α chain (~ 17% of total α chain) was isolated by reversed-phase HPLC, digested with trypsin and the tryptic digest was separated (Fig. 8). The introduction of a new tryptic site within the α62–α90 region was strongly suggested by an absence of the normal αT9,αT8-9 doublet and the appearance of two new peaks. Fractions containing aberrant peptides were examined by LSIMS and this revealed the presence of molecular ions (MH^+ = 1766.8 Da and MH^+ = 1262.6 Da) corresponding to the N-terminal (αT9.1) and C-terminal (αT9.2) parts of the original αT9 tryptic fragment in which the substitution α78 Asn→Lys took place. This mutation has previously been described as Stanleyville II and results in an α chain variant with a molecular mass 14 Da higher than normal α^A. At this point it was necessary to resolve the discrepancy between the mass originally measured by electrospray MS (15 136.4 Da) and the expected one for the identified mutation (15 140.4 Da). If the electrospray MS measurement was correct, an additional mutation site resulting in a −4 Da mass

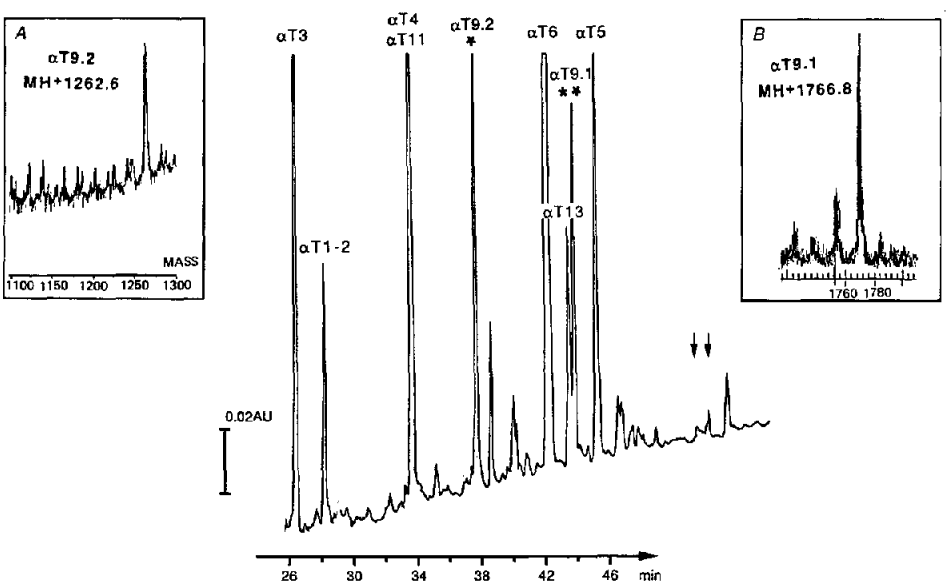

Fig. 8. Part of a chromatogram of the HPLC separation of tryptic peptides derived from isolated, nonderivatized $\alpha^{\text{Stanleyville II}}$chain. The positions that would be occupied by normal αT9 and αT8–9 peptides are indicated by arrows. Single- and double-starred peaks contain variant peptides αT9.2 and αT9.1, respectively. Portions of LSIMS spectra of tryptic digest HPLC fractions containing peptides αT9.2 Stanleyville II (MH^+ = 1262.6 Da) and αT9.1 (MH^+ = 1776.8 Da) are shown on insets A and B.

difference would have to be present. The variant α chain was separated from its normal counterpart and its molecular mass measured in the presence of an internal calibration standard (myoglobin). The mass was found to be 15 139.94 Da, which is within 0.5 Da of the one predicted for Stanleyville II. The previously found lower molecular mass appeared to be an artifactual phenomenon.

It was noticed on occasion that when two species with very close molecular masses are present, such that the two components are only partly resolved, one or both components may be measured with an error which depends on their relative abundance. When the abundance of one of the components is about 50% of the other component, as in this case, the lower-abundance component tends to be measured with a mass value which is closer to the higher-abundance component then it should be, but the higher-abundance component is not significantly affected. In the same spectrum, the mass of $^{G}\gamma$ (15 995.3 Da) was measured with −0.2 Da error but the mass of the 14 Da higher $^{A}\gamma$ (16 009.3 Da) was determined with a −6.2 Da error.

Hb Willamette (β51 Pro→Arg)

Isoelectric focusing revealed the presence of a variant hemoglobin with a mobility corresponding to Hb S, but electrospray MS clearly showed that no β^{S} chain was present (Fig. 9). The variant β chain mass was 59 Da higher than that of the normal β^{A} chain and no normal β^{A} chain was detected. Reversed-phase HPLC analysis of globin chains was consistent with MS results showing the presence of a variant β chain along with α^{A} and γ chains. Assuming the simplest possibility that a single-point mutation caused the observed mass difference, substitutions of Pro (any codon) to Arg (+59 Da), Gly (GGU, GGC) to Asp (+58 Da) or Ser (UCU, UCC) to Phe (+60 Da) might have occurred. The first possibility seemed the most likely since the amino acid substitution it involved would be consistent with the observed electrophoretic behavior. There are 21 possible β chain mutations consistent with the measured mass of the variant β chain, including seven strong candidates with substitution of Pro for Arg: in βT1 at posi-

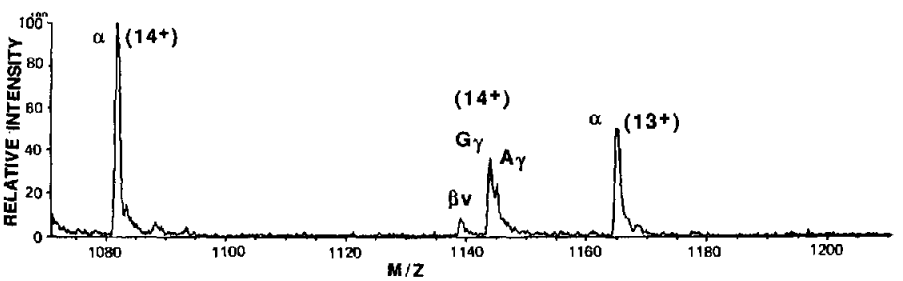

Fig. 9. Part of an electrospray mass spectrum of precipitated whole globin from hemolysate of cord blood containing Hb Willamette (β51 P→R). The presence of α chain (15 126.5±0.7 Da), $\beta^{Willamette}$chain (15 926.2±1 Da), $^{G}\gamma$ chain (15 995.9±0.9 Da) and $^{A}\gamma$ chain (16 005.4±3 Da) was confirmed.

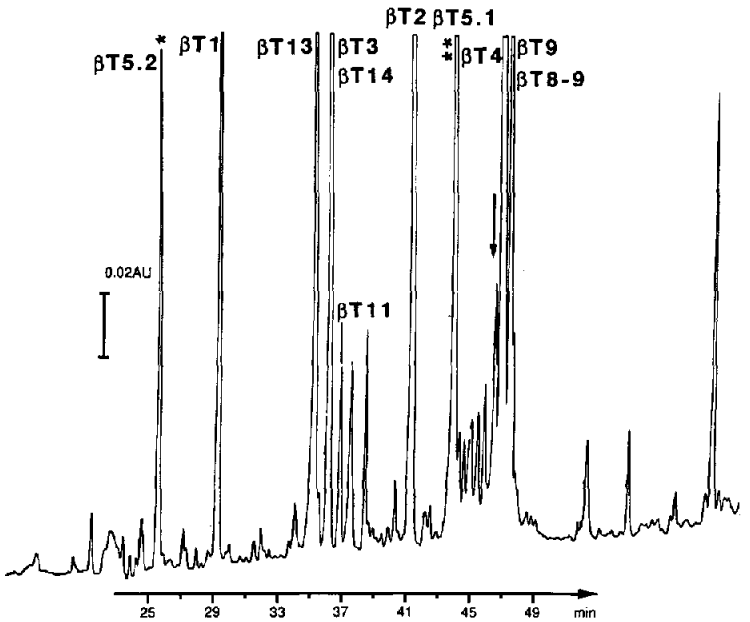

Fig. 10. Part of a chromatogram of reversed-phase HPLC separation of tryptic peptides derived from isolated, non-derivatized $\beta^{Willamette}$ chain. The position that would be occupied by normal βT5 peptide is shown by an arrow. Single- and double-starred peaks contain variant peptides βT5.2 (MH^+ = 831 Da) and βT5.1 (MH^+ = 1306 Da), respectively.

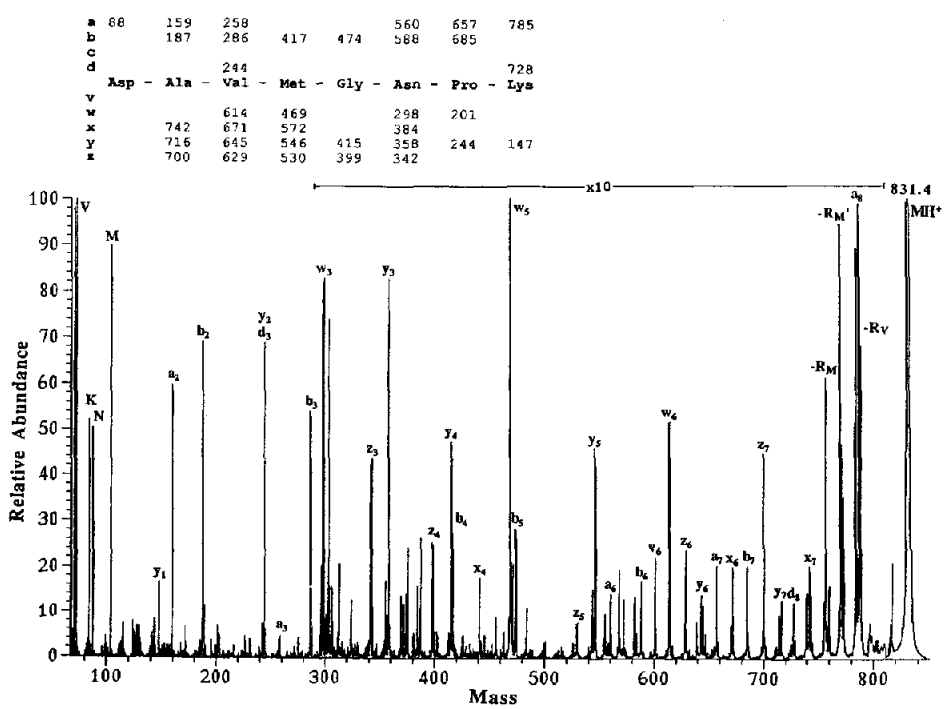

Fig. 11. CID fragment ion spectrum of $\beta^{Willamette}$T5.2 peptide (MH^+ = 831.4 Da). The nominal masses of all of the sequence ions observed in the spectrum are listed at the top.

tion 5 (Hb Warwickshire), in βT4 at position 36 (Hb Sunnybrook), in βT5 at position 51 (Hb Willamette), in βT11 at position 100 (Hb New Mexico) and in βT13 at position 124 (Hb Khartoum) and position 125 (unknown).

The abnormal β chain was separated, digested with trypsin and separated by HPLC (Fig. 10). A difference in the region of the βT5 peptide was evident in the absence of the normal βT5 peak and the appearance of two new fragments. The masses of the aberrant peptides were consistent with the C-terminal (βT5.2, MH^+ = 831 Da) and N-terminal (βT5.1, MH^+ = 1306 Da) portions of the original βT5 peptide altered by the substition 51 Pro→Arg, as previously described for Hb Willamette.

Sequencing of the βT5.2 (MH^+ = 831) (Fig. 11) and βT5.1 fragments by MS–MS confirmed the identity of the variant.

DISCUSSION

Electrospray MS is a technique of choice to meet the challenge of clinical diagnosis requirements whenever proteins or peptides of biological interest are available in the form of pure species or a well defined mixture. In the field of variant hemoglobin identification, this technique is ideally suited to serve as a preliminary diagnostic step:

(1) All three major components of hemolysate (α, β, and γ chains) are dissimilar enough to have their identities established in a mixture by accurate molecular mass measurement.

(2) The molecular masses of the majority of known variant chains can be accurately measured in the presence of their normal counterparts and globin chains of different types.

(3) The presence of a variant chain can be detected even if its molecular mass differs by only a few daltons from other species present in the mixture. However, in such a case subsequent separation of the species of interest is required in order to accurately measure its molecular mass.

(4) The ultimate verification of a variant identity is ensured. The presence of a second mutation or posttranslational modification resulting in an intact chain mass being different from the one expected on the basis of already known structural alteration will be revealed.

(5) The relative abundance of closely related species can be estimated on the basis of the relative intensity of their molecular ions. This quantification capability carries a potential for routine use in evaluating clinically important levels of variant hemoglobins and/or minor blood components as well as the purity of blood preparations used for transfusions.

In the process of identification of a variant hemoglobin, electrospray MS measurement of an intact globin chain mass may serve either as final confirmation of an expected identity or an important piece of preliminary data upon which to build the most efficient strategy. The former is a frequent case when the wealth of

electrophoretic data allows the prediction of a variant's identity nature with a high degree of confidence. The latter case involves the enumeration of all possibilities that match a measured mass difference, followed by various further procedures, *i.e.* separation of chains or globin mixture analysis, derivatization of cysteins, isolation of pure tryptic fragments for LSIMS or analysis of mixture of peptides. As we demonstrated in the Results section, different strategies were used in different cases depending on the preliminary knowledge of the possible variant nature, its relative abundance and our ability to separate the variant chain, to name only some of relevant factors. The preliminary analysis of Hb G-San Jose involved LSIMS examination of a mixture of tryptic peptides of like hydrophobicity derived from a mixture of globins [15]. For Hb G-San Jose, whose final identification is described in detail above, an abnormal chain separation was never necessary since the aberrant $\beta^{\text{G-San Jose}}$T1 peptide could have been easily located and isolated from a tryptic digest of a mixture of globins. The same approach was effective in confirmation of β^{S} identity in a cord blood sample of a double heterozygote where β^{S} was present at the 3% level, since the elution pattern of the β^{S}T1 peptide is well known. However, the analysis of $\alpha^{\text{Stanleyville II}}$ present at 8.7% required the separation of the variant chain. In the latter case, given the less than 1:5 ratio of abnormal to normal globin and the uncertainty regarding the precise value of the molecular mass difference, this was the most time-efficient approach. In the analysis of Hb Willamette, the β variant chain was easily separated to get rid of an excess of homologous γ chains whose presence would have rendered the data interpretation unnecessarily complicated. Derivatization of cysteines was not performed since none of the most probable Pro→Arg mutation sites was located within cysteine containing peptides.

In summary, the applicability of electrospray MS to hemoglobin analysis is dependent on the extent of preliminary work-up procedures carried out on the blood and the strategy chosen for further analysis. In its simplest and most attractive form (examination of globin precipitates), it is able to give accurate molecular weight and quantitative information about aberrant proteins on the unseparated chains if there is at least a 14 mass unit difference between different chain species. HPLC separation of globin chains prior to electrospray will in most cases allow mass measurement of the variant chain to within 1 Da, therefore encompassing almost all amino acid replacements. The addition of tryptic digestion, peptide separation and MS–MS sequencing allow complete variant characterization.

In spite of its expense and complexity, MS should have an increasing role in variant hemoglobin identification and quantification in association with private and government-mandated screening programs.

ACKNOWLEDGEMENTS

The authors wish to thank Bertram Lubin, Klara Kleman, Linda Pinkoski and

Esther Roitman of Children's Hospital Oakland for their encouragment and scientific contributions to this study. The NIH supported these studies through Sickle Cell Center Grant HL20985, Shared Instrumentation Grant 03300 (DRR) (for VG-30-250) and Mass Spectrometry Research Resource Grant RR01614 (to A.L. Burlingame). Financial support from the NSF Biological Instrumentation Program Grant DIR 8700766 is also acknowledged.

REFERENCES

1 H. F. Bunn and B. G. Forget, *Hemoglobin: Molecular, Genetic and Clinical Aspects*, W. B. Saunders, Philadelphia, PA, 1986.
2 Y. Wada, A. Hayashi, T. Fujita, T. Matsuo, I. Katakuse and H. Matsuda, *Biochim. Biophys. Acta*, 667 (1981) 233.
3 I. Katakuse, T. Ichihara, H. Nakabushi, T. Matsuo, H. Matsuda, Y. Wada and A. Hayashi, *Biomed. Mass Spectrom.*, 11 (1984), 386–391.
4 D. Prome, I. Prome, Y. Blouquit, C. Lacombe, J. Rosa and J. D. Robinson, *Spectros. Int. J.*, 5 (1987) 157–170.
5 Y. Wada, T. Matsuo and T. Sakurai, *Mass Spectrom. Rev.*, 8 (1989) 379–434.
6 D. Prome, J.-C. Prome, F. Pratbernou, Y. Blouquit, F. Galacteros, C. Lacombe, J. Rosa and J. D. Robinson, *Biomed. Environ. Mass Spectrom.*, 16 (1988) 41–44.
7 S. Rahbar, T. D. Lee, J. A. Baker, L. T. Rabinowitz, Y. Asmeron, K. Legesse and H. M. Ranney, *Hemoglobin*, 10 (1986) 379–400.
8 C. K. Meng, M. Mann and J. B. Fenn, *Z. Phys. C. At. Mol. Custers*, 10 (1988) 361–368.
9 J. A. Loo, H. R. Udseth and R. D. Smith, *Biomed. Environ. Mass Spectrom.*, 17 (1988) 411–414.
10 J. A. Loo, H. R. Udseth and R. D. Smith, *Anal. Biochem.*, 179 (1989) 404–412.
11 J. B. Shelton, J. R. Shelton and W. A. Schroeder, *J. Liq. Chromatogr.*, 7 (1984) 1969–1977.
12 M. Castagnola, R. Landolfi, D. V. Rosetti, F. DeAngelis and S. Ceccarelli, *Anal. Lett.*, 19 (1986) 1793–1807.
13 W. E. Schwartz, P. K. Smith and G. P. Royer, *Anal. Biochem.*, 106 (1980) 43–48.
14 F. C. Walls, A. N. Baldwin, A. M. Falick, B. W. Gibson, F. Kaur, D. A. Maltby, B. L. Gillece-Castro, K. F. Medzihradszky, S. Evans and A. L. Burlingame, in A. L. Burlingame and J. A. McCloskey (Editors), *Biological Mass Spectrometry*, Elsevier, Amsterdam, 1990, pp. 197–214.
15 B. N. Green, R. W. A. Oliver, A. M. Falick, C. M. L. Shackleton, E. Roitman and H. E. Witkowska, in A. L. Burlingame and J. A. McCloskey (Editors), *Biological Mass Spectrometry*, Elsevier, Amsterdam, 1990, pp. 129–146.
16 F. Kutlar, A. L. Reese, Y. E. Hsia, K. M. Kleman, and T. H. J. Huisman, *Hemoglobin*, 13 (1989) 671–683.