From the Department of Laboratory Medicine, Division of Clinical Chemistry, Huddinge University Hospital, Karolinska Institutet, Stockholm, Sweden

The Use of Mass Spectrometry and DNA Technology in the Investigation of Hemoglobin Disorders

Dilip K. Rai

Stockholm 2003

Work was carried out at the Department of Laboratory Medicine, Division of Clinical Chemistry and Clinical Research Center, Huddinge University Hospital, and Department of Medical Biochemistry and Biophysics, Karolinska Institutet.

Doctoral thesis

ISBN: 91-7349-450-X

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden.

Men argue, nature acts.

Dedicated to my mother

Mrs Hari Maya Rai

ABSTRACT

Hemoglobin (Hb) disorders, which severely affect nearly 300,000 newborns every year, have become a global problem. The disorders manifest a wide spectrum of clinical symptoms ranging from asymptomatic to lethal conditions. Inheritance or spontaneous occurrence of genetic defects could lead to either a structural abnormality of Hb (Hb variants) or an impaired or no synthesis of the globin chains (thalassemias). The majority of Hb variants are caused by a single nucleotide substitution resulting in an amino acid change in the globin chain, while the thalassemias are caused by a diverse array of mutations.

The study of Hb disorders can elucidate the structure-function relationship of the Hb and may clarify the mechanism of pathogenesis. Hb disorders are good models of protein abnormalities that can be directly or indirectly associated with molecular diseases.

Conventional electrophoresis and HPLC techniques give presumptive identification of Hb variants and are still widely used today. Modern developments in DNA technology have revolutionized detection and identification of Hb disorders at the nucleotide level. Simultaneously, mass spectrometry has emerged as a powerful tool in the field of biological and biomedical studies. Proteins of almost unlimited molecular mass can be analyzed and amino acid sequence of peptides can be readily determined.

In order to evaluate the performance and limits of DNA and electrospray mass spectrometry techniques in the investigation of Hb variants, different studies using these techniques were initiated. The mass spectrometric experiments were done at two levels, i.e. at the intact globin chain and at the tryptic peptide levels. The mass analyses were performed on two-step diluted whole blood samples requiring only 10 μL of the whole blood. Blood samples stored at +4° C or –20° C were used. The variants which were present to greater than 10% of the total Hb, and whose masses differed by greater than 6 Da from the normal chain were in most cases readily identified by electrospray mass spectrometry. Data from isoelectric focusing and a knowledge of the genetic code greatly aided the characterization of Hb variants.

Accurate mass measurements (<±5 ppm) of intact globin chains (~16 kDa) allowed the detection of variant chains that differ in mass by <6 Da from the corresponding normal globin chain. A good correlation was found between the HPLC and mass spectrometric methods in quantifying Hb variants.

Tandem mass spectrometry (MS/MS) of intact β-chains identified some Hb variants due to an amino acid change within 40 to 50 residues from the terminals of the globin chain. In others, the MS/MS data directed the location of the amino acid substitution to a certain region of the β-globin chain.

A secondary structural change in Hb variants, i.e. disulfide bridge, was also demonstrated and characterized by mass spectrometry.

However, due to the complexity of the peptide mixture, DNA analyses were sometimes required to allow successful interpretation of mass spectrometry data. In general, mass spectrometry was a powerful tool for screening and identification of aberrant Hb proteins.

Key words: hemoglobinopathies, thalassemia, mass spectrometry, electrospray, tandem mass spectrometry, proteins.

Dilip Rai: The use of mass spectrometry and DNA technology in the investigation of hemoglobin disorders, Karolinska Institutet, Stockholm 2003, ISBN 91-7349-450-X.

CONTENTS

LIST OF ORIGINAL PAPERS

This thesis is based on following papers, which are referred to in the text by their Roman numbers:

I Rai DK, Alvelius G, Landin B, Griffiths WJ. Electrospray tandem mass spectrometry in the rapid identification of α-chain haemoglobin variants. Rapid Commun Mass Spectrom 2000;14:1184-94.

II Rai DK, Landin B, Alvelius G, Griffiths WJ. Electrospray tandem mass spectrometry of intact β-chain haemoglobin variants. Anal Chem 2002;74:2097-102.

III Rai DK, Griffiths WJ, Landin B, Alvelius G, Green BN. Characterization of the elusive disulfide bridge forming human hemoglobin variant: hemoglobin Ta-Li β83 (EF7)Gly→Cys by electrospray mass spectrometry. J Am Soc Mass Spectrom 2002;13:187-91.

IV Rai DK, Griffiths WJ, Alvelius G, Landin B. Electrospray mass spectrometry: an efficient method to detect silent hemoglobin variants causing erythrocytosis. Clin Chem 2001;47:1308-11.

V Rai DK, Griffiths WJ, Landin B, Wild BJ, Alvelius G, Green BN. Accurate Mass Measurement by Electrospray Ionization Quadrupole Mass Spectrometry: Detection of Variants Differing by <6 Da from Normal in Human Hemoglobin Heterozygotes. Manuscript.

ABBREVIATIONS

Three- and one-letter codes, standard genetic codes, and residue monoisotopic mass for the twenty amino acids and for the stop codon.

Amino acid	3-letter	1-letter	Standard genetic code (DNA)	MW (Da)
Alanine	Ala	A	GCT, GCC, GCA, GCG	71.04
Arginine	Arg	R	CGT, CGC, CGA, CGG, AGA, AGG	156.10
Asparagine	Asn	N	AAT, AAC	114.04
Aspartic acid	Asp	D	GAT, GAC	115.03
Cystiene	Cys	C	TGT, TGC	103.01
Glutamic acid	Glu	E	GAA, GAG	129.04
Glutamine	Gln	Q	CAA, CAG	128.06
Glycine	Gly	G	GGT, GGC, GGA, GGG	57.02
Histidine	His	H	CAT, CAC	137.06
Isoleucine	Ile	I	ATT, ATC, ATA	113.08
Leucine	Leu	L	TTA, TTG, CTT, CTC, CTA, CTG	113.08
Lysine	Lys	K	AAA, AAG	128.09
Methionine	Met	M	ATG	131.04
Phenylalanine	Phe	F	TTT, TTC	147.07
Proline	Pro	P	CCT, CCC, CCA, CCG	97.05
Serine	Ser	S	TCT, TCC, TCA, TCG, AGT, AGC	87.03
Threonine	Thr	T	ACT, ACC, ACA, ACG	101.05
Tryptophan	Trp	W	TGG	186.08
Tyrosine	Tyr	Y	TAT, TAC	163.06
Valine	Val	V	GTT, GTC, GTA, GTG	99.07
Stop codon	Term	•	TAA, TAG, TGA	0

CAM	carboxyamidomethylation
CE	capillary electrophoresis
cIEF	capillary isoelectric focusing
Da	Dalton
DPG	diphosphogycerate
DTT	dithiothreitol

EDTA	ethylene diamine tetra acetic acid
ES	electrospray
FBC	full blood count
FTICR	Fourier transform ion cyclotron resonance
FWHM	full width at half maximum height
Hb	hemoglobin
HPFH	hereditary persistence of fetal hemoglobin
IEF	isoelectric focusing
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption/ionization
MS/MS	tandem mass spectrometry
PCR	polymerase chain reaction
ppm	parts per million
Q-TOF	quadrupole time-of-flight
R	relaxed, oxygenated state
T	taut, deoxygenated state
TOF	time-of-flight

Some normal and abnormal hemoglobins

Hb A	$\alpha_2\beta_2$	Hb S	β6Glu→Val
Hb A_2	$\alpha_2\delta_2$	Hb C	β6Glu→Lys
Hb F	$\alpha_2\gamma_2$	Hb E	β26Glu→Lys
Hb A_{1c}	glycosylated Hb A	Hb D-Los Angeles	β121Glu→Gln
Hb H	β_4	Hb Bart's	γ_4

INTRODUCTION

Hemoglobin (Hb): the paradigm protein molecule of all time!

Hemoglobin (Hb), an essential molecule to ferry oxygen for respiration, has been the subject of interest in clinical and biological fields for the past few centuries *(1)*. The study of the Hb molecule has provided fundamental insights into the structure-function relationships of proteins in general and the molecular basis of oxygen transport in particular *(2)*. The structural elucidation of sickle cell Hb in 1949 has initiated the new era of molecular medicine *(3)*. The rapid advent of molecular biology techniques over the past 20 years has helped to reveal the genetic defects of several molecular diseases, including Hb disorders. The Hb disorders are the most widely studied protein abnormalities involving the classical qualitative and quantitative anomalies *(4)*. With an improved understanding of the molecular basis of Hb disorders, their management and treatment are evolving *(5)*. In 1996, a new function of Hb as a regulating protein of blood pressure via nitrosylation of β-globin was discovered *(6)*. In view of the extensive knowledge about genetics and regulation, it can be predicted that the Hb molecule will continue to be a working model in the application of advancing techniques. Its further studies will also enrich our existing biological knowledge, with respect to structure-function relationships of proteins *(2)*, allosteric effects *(7)*, molecular basis of diseases, and gene therapy *(8)*.

This thesis dwells on the applications of modern mass spectrometry and DNA techniques to study abnormal proteins using Hb as a prototype protein molecule.

Normal Hb Structure

Hb in humans is a tetramer of 2 identical α-like and 2 identical β-like globin chains (Fig. 1). Each globin chain is linked with a prosthetic heme group that is guarded by hydrophobic residues. In adults and children older than six months Hb A ($\alpha_2\beta_2$) normally constitutes more than 90% of the total Hb. The β-subunit (146 residues) contains eight helical segments lettered A through H. The α-subunit (141 residues) also contains similar helical segments except that residues that make helix D in β-globin are absent in the α-globin.

Hb molecules exist in two conformations: a deoxy or taut (T) and oxy or relaxed (R) state *(2)*. The T state has lower affinity for ligands such as oxygen and carbon monoxide than the R state. In addition to hydrogen bonds between inter-subunit and intra-subunit residues, several salt bridges and hydrophobic contacts stabilize the Hb molecule during the T state. These include residues responsible for heme-heme interactions, Bohr effects and for binding 2,3-diphosphoglycerate (2,3-DPG). Upon oxygenation, the salt bridges are sequentially broken and the Hb molecule is able to dissociate reversibly according to the reaction $\alpha_2\beta_2 \leftrightarrow 2\alpha\beta$. Unlike many proteins, native Hb does not contain disulfide bonds although there are six Cys residues in a tetramer of which only two residues (β93) are exposed to the solvent.

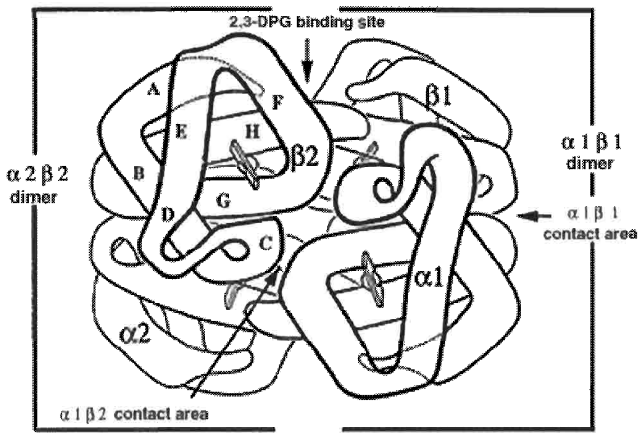

Fig. 1. Structure of a Hb molecule.
Reprinted from *(36)* with permission from ©Springer-Verlag GmbH & Co, Heidelberg, Germany.

Each α-chain is in contact with the two β-chains. The two contacts can be defined as α1β1 and α1β2 and the two-fold symmetry generates structurally identical α2β1 and α2β2 contacts. The α1β1 binding is stronger than the α1β2 and thus the latter experiences considerable movement during T↔R transitions.

Minor Hb fractions: Hb F, Hb A_2, and common Hb derivatives

Hb F ($\alpha_2\gamma_2$) is predominant in fetal life and at birth. Hb F binds 2,3-DPG less strongly than does Hb A and hence has a higher oxygen affinity. This contributes to more efficient oxygen supply to the fetus from the maternal blood during pregnancy. The Hb F levels

decline rapidly at about six months after birth to less than 1% in adult red blood cells. Hb F is increased to a variable extent in β-thalassemia and sickle cell anemia *(9)*. Hb F is also increased in other instances such as hematological or metabolic disorders and sometimes during pregnancy *(9)*. Chemical agents such as hydroxyurea are used to induce higher expression of Hb F in sickle cell disease and other β-globin disorders *(10)*.

Hb A_2 ($\alpha_2\delta_2$) makes up approximately 2-3% of normal Hb in adults. The level of Hb A_2 is increased in β-thalassemia, sickle cell trait and hyperthyroidism *(9, 11)* whereas it is decreased in α-thalassemia as well as in iron deficiency and sideroblastic anemias *(9)*.

There are also various minor Hb forms that result from post-translational modifications *in vivo*. The classic one is Hb A_{1c}. It accounts for approximately 3-5 % of the Hb in normal adult red cells. It differs from Hb A in that the N-terminal amino group of each globin chain has glucose attached non-enzymatically by a ketoamine linkage. Hb A_{1c} is a valuable indicator in the long-term regulation of blood glucose levels in diabetic patients *(12)*. Other post-translational modifications of normal globin chains include carbamylation of the N-terminal group of both α- and β-chains attributed to the reaction with cyanate in uremic patients *(13)*, S-glutathionylation *(14)* and S-nitrosylation *(6)* at β93Cys. About 15% of Hb F in the developing fetus has a co-translational modification, i.e. the N-terminus of the γ-chain is acetylated. Otherwise no other human globin subunits are acetylated except for some mutants that have substitutions of the N-terminal residue *(15)*.

Structure and organization of globin genes

The genes coding for α-like and β-like globin chains are located on chromosome 16 and 11 respectively. The globin genes share a common structure consisting of three exons and two introns (Fig. 2). The intron-exon junction is characterized by the presence of a conserved nucleotide sequence GU ("donor" splice sequence) at 5'-end and AG ("acceptor" splice sequence) at 3'-end. Globin genes with similar structure and nucleotide content to genes for human globins are found in plants, invertebrates and vertebrates. This supports the hypothesis that they evolved from the same ancestral gene *(16)*.

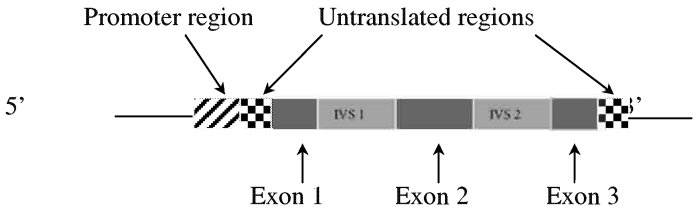

Fig. 2. Schematic structure of human globin gene with three exons separated by two intervening sequences (IVS).

The α-gene cluster on chromosome 16 contains an embryonic ζ-gene and two α-genes, i.e. α1- and α2-genes (Fig. 3a). The major regulatory elements that control the α-gene cluster are located 40 kb upstream of the ζ-gene and hence the site is designated as hypersensitive site (HS-40). The two α-genes share identical coding nucleotide sequences, but differ slightly in the sequences occurring in the second intron and in the 3'-end of the untranslated region. The mRNA expression of the α2-gene is approximately twice the α1-gene *(17)* but the translation rate of the two α-genes is almost equal *(18)*. The ζ-gene is active in the early embryonic stage. The α-gene is active in the later stages of development and throughout adult life.

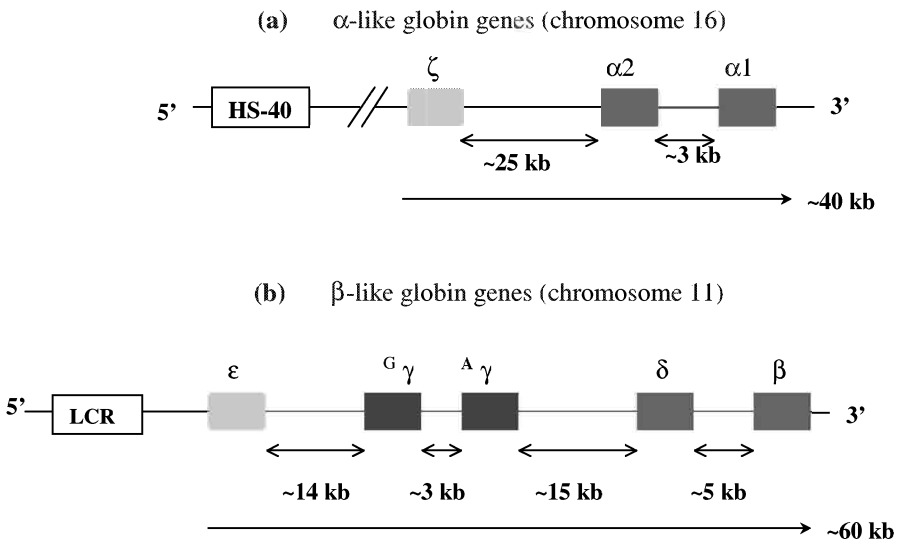

Fig. 3. Functional human globin genes (a) α-gene cluster, (b) β-gene cluster. Pseudogenes are not shown.

The β-like genes on chromosome 11 consist of ε, $^{G}\gamma$, $^{A}\gamma$, δ, and β genes (Fig. 3b). The regulatory elements called the locus control region (LCR) for the β-globin gene cluster are placed 6 to 20 kb upstream of the ε-globin gene. Unlike the α-like globins, the β-like globins undergo two developmental switches, i.e. embryo→fetal and fetal→adult. In the embryonic stage the ε-gene is active followed by activation of the γ-gene in fetal life and subsequent activation of the β-gene in adulthood. The mechanism of switches in globin gene expression is still not fully understood *(19)*. Generally gene deletions affecting the normal switching of γ to β-globin gives rise to a minor group of Hb disorders called hereditary persistence of fetal Hb (HPFH) *(20)*.

Globin chains and their assembly

The α-like and β-like globin chains are 141 and 146 amino acids long respectively (Table 1). There is some sequence homology between the α- and the β-globin chains with 64 residues in identical positions. The polypeptides produced by the two γ-genes differ only in the amino acid at position 136, i.e. Gly in $^{G}\gamma$- and Ala in $^{A}\gamma$-chain. Most newborns have a γ-chain content of ~70% $^{G}\gamma$ and ~30% $^{A}\gamma$ and the $^{G}\gamma$ content decreases to about 40% in adulthood. In 20% of the newborns, Ile at position 75 of the $^{A}\gamma$-chain is replaced by Thr ($^{A}\gamma^{T}$). The β-chain also has structural homology with the γ- (~73%) and δ-chains (~93%).

Table 1. Average molecular masses of normal human Hb chains

Globin chain	Average mass in Daltons (Da)
α-like globins:	
ζ-chain	15505.85
α-chain	15126.38
β-like globins:	
ε-chain	16071.66
$^{G}\gamma$-chain	15995.27
$^{A}\gamma$-chain	16009.30
$^{A}\gamma^{T}$-chain	15997.25
β-chain	15867.24
δ-chain	15924.32

The mature polypeptides, with a prosthetic heme group linked to each chain, aggregate to form dimers of unlike globin chains (Fig. 4). The dimers are more soluble and less susceptible to proteolytic digestion than the newly synthesized individual globin chains. Heterodimerization of the globin chains is favoured by electrostatic attraction between the basic α-globins (*pI* = 8.1) and slightly acidic β-globins (*pI* = 6.8) *(21-23)*. The rate of dimerization of α and β monomers to form αβ dimers partly depends on the strength of the electrostatic attraction. For instance, a relatively more acidic β-chain will dimerize with the

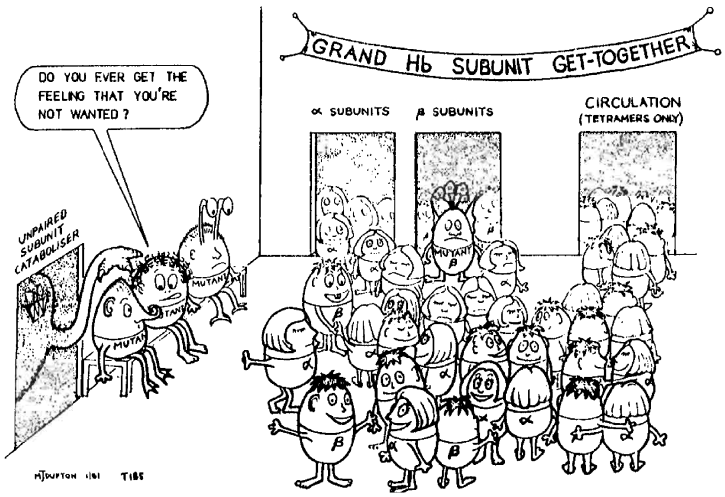

Fig. 4. Social behaviour of globin chains.
Reprinted from Trends Biochem Sci 1981;6:158-61 with permission from ©Elsevier Science.

α-chain at a faster rate than the normal β-chain. Subsequent aggregation of the dimers to tetramers is absolutely essential for the effective function of Hb as an oxygen carrier.

Hb disorders

Hb disorders are the most common monogenic disorders worldwide *(24)*. In most cases the pattern of inheritance follows the classical one for autosomal recessive conditions. The Hb disorders can be divided into two main groups: those in which there is defect in structure and those in which there is unbalanced synthesis of different globin chains. The former group is called Hb variants (or hemoglobinopathies) and the latter group is called thalassemias. The majority of the Hb variants are the results of a single nucleotide

substitution altering an amino acid in the globin chain while the thalassemias are caused by a diverse array of mutations.

It is estimated that about 270 million people carry Hb disorders worldwide *(24)*. Most of them live in regions where malaria is endemic. Each year over 300,000 babies are born with life-threatening hematological disorders, i.e. sickle cell anemia, major β-thalassemia as well as severe forms of α-thalassemia *(24)*. The most commonly found structural mutations are Hb E (β26Glu→Lys) predominantly in South-east Asia, Hb S (β6Glu→Val) in Africa (tropical regions), India, and Saudi Arabia, and Hb C (β6Glu→Lys) in West Africa *(24)*. The individuals heterozygous for Hb S or Hb E, or homozygous for Hb C have strong protection against malarial infection *(25)*. Over the last century, because of mass population migration around the world, the prevalence of the common Hb variants has become a global issue.

Thalassemia

In a normal scenario, the α/β-globin chain synthesis ratio is ~1.0. When there is an imbalance in the synthesis of globin chains, it is defined as thalassemia. Thalassemia is classified according to the globin chain for which the synthesis is impaired, e.g. α-thalassemia and β-thalassemia affecting the α- and β-globin chains respectively *(26)*.

In α-thalassemia, the excess γ- or the β-chains form soluble homotetramers called Hb Bart's (γ_4) or Hb H (β_4) which when present in high concentration can cause hemolytic anemia. The degree of anemia and the amounts of Hb Bart's or Hb H produced broadly reflect the degree to which α-globin synthesis has been impaired. The α-thalassemias are most commonly due to gene deletions. The mildest forms of α-thalassemias [α-thal-1 (--/αα; -α/-α) and α-thal-2 (-α /αα)] are generally clinically normal except for microcytosis *(26)*. A more severe clinical picture arises when three of the four α-genes are deleted (--/-α). It is marked by hemolytic anemia producing a large amount of Hb Bart's at birth or a similar amount of Hb H in adult life, a condition called Hb H disease. A life-threatening condition referred as Hb Bart's hydropis fetalis syndrome occurs when functional α-globin genes are

missing (--/--). This condition affects the fetus prior to birth. More than 80 different deletion and point mutations resulting in α-thalassemias have been defined *(26)*.

The down-regulation of β-globin synthesis in β-thalassemia results in accumulation of α-globin chains. When present in excess in the erythroid cell, these α-chains form insoluble precipitating aggregates causing membrane damage and premature red cell destruction. The severity of clinical manifestations in β-thalassemia generally thus correlates with the size of the free α-chain pool and the degree of α- to non-α globin chain imbalance. In general symptoms are not manifested at birth. β-thalassemias are generally due to point mutations at various sites of the β-cluster genome *(26)*. The β-thalassemias are sub-classified into β^0 and β^+-thalassemia: β^0 in which there is no β-globin production from the affected gene, and β^+-thalassemia in which some β-chain is produced but at a reduced rate.

Phenotypically the β-thalassemias are divided into two main forms: thalassemia minor and thalassemia major. The individuals with thalassemia minor have only one of the two β-globin genes defective, i.e. they are heterozygous for β-thalassemia. The individuals with thalassemia minor manifest from normal to mild anemia and do not require treatment. The individuals with thalassemia major have no normal β-globin gene, i.e. they are homozygous for β-thalassemia. It is a life-threatening disease and requires chronic blood transfusion or bone marrow transplantation. Today there are over 200 different gene mutations causing β-thalassemias *(26)*.

Hb variants

Nearly 900 Hb variants have been described thus far as summarized in Table 2 *(27)*. In populations lacking commonly encountered Hb variants S, E, or C, it is estimated that one in every 800 individuals carries a Hb variant that is detectable on electrophoresis *(28)*. In fact the incidence of Hb variants is expected to be even higher since there exist several variants which do not differ in charge from the normal Hb.

Table 2. Number of structurally abnormal Hb chains reported as of February 2003.

α-chain variants with a base substitution	258
β-chain variants with a base substitution	408
γ-chain variants with a base substitution	69
δ-chain variants with a base substitution	37
Fusion globin chains	9
Hb chains with deletions and/or insertions	41
Hb chains with extended polypeptide chains	13
Hb chains with more than one point mutation	29
TOTAL	864

Nomenclature of Hb variants

When the first few Hb variants were discovered, they were designated letters of the alphabet C, D, E, etc. *(28)*. Letter M was assigned to Hb variants associated with methemoglobinemia and S for sickle cell Hb. Current nomenclature of Hb variants are derived from the patient's area of origin or the place of discovery and are sometimes preceded by letters reflecting their mobilities on electrophoresis. For example, Hb D-Los Angeles (β121Glu→Gln) indicates a variant discovered in Los Angeles with the same electrophoretic mobility as any other Hb D-variant. Some Hb variants have more than one name for the same mutation reflecting the rediscovery of the mutation in different parts of the world.

Genetic basis of Hb variants

Point mutations

Over 90% of all Hb variants described are due to substitution of a single DNA nucleotide base by another in the coding sequence (Table 2). This substitution is reflected by a corresponding amino acid change in the globin chain. For example, the mutation G<u>A</u>G→G<u>T</u>G which when occurs at the sixth residue of the β-globin chain results in Hb S. Some extended Hb variants also arise from nucleotide base substitutions *(27)*. For example, Hb Constant Spring (α142Term→Gln) extended with 31 amino acids is a result of <u>T</u>AA→<u>C</u>AA mutation. Nucleotide base substitution can also introduce a stop codon that will generally lead to thalassemia, but occasionally stable Hb variants can be produced only when the C-terminal end residues are involved *(27)*. Identical amino acid replacement can

be caused by more than one nucleotide substitution *(29)* and nucleotide substitutions causing the same amino acid replacement can also occur in different exons *(30)*.

Nucleotide insertions and deletions

About 5% of the total Hb variants are due to insertions or deletions of one or more nucleotide bases (Table 2). These events usually result in non-sense reading frames and often cause thalassemia phenotype. Frameshift mutations occurring at the end of the coding portion of globin genes permit the synthesis of viable but elongated globin chains. For example, an insertion of AC in codon 147 results in extension of the β-chain by 11 residues in Hb Tak *(28)*. Deletions of one or more intact codons keep the reading frame in phase but the resultant Hb variants are generally unstable and might be associated with hemolytic anemia.

Crossovers

Structurally defective hemoglobins due to equal or unequal crossovers between homologous chromosomes can also occur *(27, 31)*. An equal homologous crossover between chromosomes may be the underlying mechanism for Hb variants with two amino acid substitutions on the same allele *(28)*. Unequal homologous crossovers may also have occurred in Hb variants with multiple nucleotide deletions or additions *(28)*. Presumably, sufficient homology between nucleotide bases exists in certain segments of the chromosome to allow slight misalignment during crossover. For example, if the DNA segments of two chromosomes are aligned in a manner necessary to produce the deletion in Hb Lyon (β17-18Lys-Val→0), as many as 12 of the 18 nucleotide bases in the crossover region can be identical *(32)*. A greater degree of misalignment allows crossing-over between different genes producing hybrid genes. Hb Lepore variants (fusion of β- and δ-globins) and Hb Kenya (fusion of γ- and β-globins) are typical examples of products from hybrid genes *(28, 31)*.

Assembly of Hb variants

In principle, inherent to the number of genes, a stable heterozygote α-chain variant should constitute ~25% of total Hb, while the heterozygote β-chain variant should constitute about

half of the total Hb in the red cells. However, in practice the amount of Hb variants in the red cells differ significantly from one type to another although most variant chains are synthesized at the same rate and are as stable as normal globin chains *(28)*. The variability can be explained by the electrostatic attraction model discussed earlier on the assembly of normal globin heterodimers *(21)*. Alterations in surface charge of the globin chains contribute to different rates of assembly of Hb tetramers and hence to different proportions of the variant chains *(22, 23)*. The positively charged β-chain Hb variants S, C, and D-Los Angeles constitute less than half of the total Hb in heterozygous individuals and are reduced further in the presence of α-thalassemia. The negatively charged β-chain variants on the other hand are present in greater amounts than Hb A and are elevated in the presence of α-thalassemia.

Pathophysiology of Hb disorders

Hb variants with amino acid substitutions occurring at the surface of the molecule are generally innocuous since they are not likely to affect the tertiary structure, heme function or subunit interaction. However some surface substitutions of which Hb S is an example can alter the function of the Hb molecule.

Sickling

Hb S molecules when present in high concentration within a red cell can polymerize when deoxygenated. The polymerization results in an irreversible distortion (sickling) of cell shape, reduced cell deformability, hemolysis and impaired microvascular circulation. The sickling of red cells depends on the intracellular concentration of Hb S *(33, 34)*. The co-occurrence of HPFH, or an α-thalassemia with Hb S plays a protective role from sickling of the red cells *(34)*.

Unstable Hb variants

Amino acid substitutions that impair the heme binding capacity of the globin and those that disrupt the secondary structure of the globin chain usually render instability to the globin chain *(27, 28)*. The instability conferred by these substitutions results in an accelerated denaturation of Hb which may present as mild and fully compensated hemolytic anemia or a

severe hemolysis *(28)*. Some of the unstable globin chains are incapable of forming tetramers and therefore degrade rapidly leading to thalassemic phenotype *(28)*. Some abnormal globin chains, in particular the β-chains, are so unstable (hyperunstable) that they give rise to dominantly inherited β-thalassemia *(35)*.

High oxygen affinity Hb variants

Mutations disrupting the normal oscillation of the Hb tetramer between the R ↔ T states result in a change of oxygen affinity. Several Hb variants exhibit high oxygen affinity, some of which may be of sufficient magnitude to cause erythrocytosis *(36)*. Most of the substitutions producing this abnormality appear in α1β2 contact positions where they impair the heme-heme interaction (see Fig. 1) *(27, 28, 36)*. Substitutions close to the C-terminus of the β-chain may also occur. Such substitutions may destabilize the T state by altering the alkaline Bohr effect, 2,3-DPG binding sites, or by formation of salt bridges *(27, 28, 36)*. This leads to an increase in the red cell mass due to the inefficient unloading of oxygen to the peripheral tissues and is called erythrocytosis. Erythrocytosis due to high oxygen affinity Hb variants is rare, but may probably be diagnosed to a greater degree if more suitable screening techniques are used *(paper IV)*. Many high oxygen affinity Hb variants do not alter the surface charge of the Hb molecule and therefore escape detection by electrophoresis *(28)* and some cases even by isoelectric focusing *(36, paper IV)*.

Low oxygen affinity Hb variants

The amino acid substitutions that tend to destabilize the R state demonstrate low oxygen affinity *(28)*. Some of the low oxygen affinity Hb variants are unstable and may result in hemolytic anemia while a few others can cause cyanosis.

M Hemoglobins

Hb variants involving amino acid substitutions in the heme pocket prevent reduction of ferric atoms by methemoglobin reductase. The Hb variants in this group are rare and are classified as M hemoglobins for their characteristic methemoglobin (non-functional trivalent iron atom) state. These Hb variants do not combine fully with oxygen and hence

lead to accumulation of deoxygenated hemoglobins. The excess deoxygenated hemoglobins cause colouration of skin and mucous membranes from brownish to slate gray *(28)*.

Clinically unimportant Hb variants

The number of variants without clinical significance has increased as a consequence of Hb A_{1c} quantification by many laboratories *(37, 38)*. In the investigation of Hb A_{1c} itself, several clinically insignificant Hb variants have falsely altered Hb A_{1c} levels *(38)* which may lead to wrong diagnosis or treatment. Identification and characterization of new variants may lead to greater understanding of the structurally related properties of the Hb molecule. The study of these variants may also provide additional knowledge on protein chemistry like post-translational modifications, e.g. N-acetylation *(15, 39)*, deamidation of Asn to Asp *(40)*, oxidation of Met to Asp *(41)* and oxidation of Leu to hydroxyleucine *(42)*.

Analytical methods

Full blood count

The majority of the Hb variants do not infer any specific morphological features to the erythrocytes and the full blood count (FBC) may well present normal. The red cell indices are important indicators in the assessment of mild thalassemia characterized by normal or slightly subnormal Hb levels. Reduced mean corpuscular volume (MCV<72 fL) of red cells and mean corpuscular hemoglobin (MCH <27 pg) are indicative of possible thalassemia *(43, 44)*. High hematocrit, i.e. increased red cell mass without elevation of white cells and platelets, is the hallmark of erythrocytosis. Raised reticulocyte count is a sign of hemolysis and can be associated with unstable Hb variants, Hb H disease or sickle cell disease.

Electrophoresis

Electrophoresis has been the traditional method of choice for the identification and quantification of Hb variants *(28)*. Commercial, rapid electrophoretic methods have been developed that allow separation at pH 8.4 on cellulose acetate and pH 6.2 on citrate agarose gels *(43, 44)*. The latter provides a clear background, allowing for quantification of Hb present by densitometer scanning. Electrophoresis at acidic pH is better than more routinely used alkaline pH electrophoresis for separating some of the common Hb variants *(43)*. For example, alkaline electrophoresis does not separate Hb C and Hb E whereas their separation

is readily achieved in acidic electrophoresis. The electrophoresis, regardless of the pH, provides inaccurate data in quantifying low proportion variants *(45)* and in the detection of fast moving variants such as Hb H *(43)*. Inherent to its principle of separation on charge difference, mutations that do not alter the charge are not detected by electrophoresis.

Isoelectric focusing

Isoelectric focusing (IEF) is a highly sensitive method based on the separation of globins according to their *pI* values. When optimized, methods using free ampholytes can resolve two proteins with a *pI* difference of about 0.01 *(46)*. Methods using ampholytes bound covalently to a matrix have claimed to give one order higher resolution *(46)*. IEF has better resolution than standard electrophoresis and subsequently has replaced the latter in many clinical laboratories for identification of Hb variants *(47, 48)*. During the last decade, IEF has been adapted to capillary technology allowing separation of very small amount of samples and automation of sampling *(11, 49, 50)*. Capillary IEF (cIEF) coupled with mass spectrometry has allowed for the analysis of proteins present in single cells *(50)*. cIEF has been applied successfully in rapid identification and quantification of Hb variants *(11, 51)*.

HPLC

HPLC plays an important role in detection, characterization and isolation of Hb variants *(45, 48, 51, 52)* in addition to diagnosis of thalassemias *(53)*. Cation-exchange HPLC has become the method of choice for the initial screening and for quantification of Hb fractions *(45, 48, 53)*. The Bio-Rad Variant (Bio-Rad Laboratories) is an automated cation-exchange HPLC instrument that has been used to quantify Hb A_2, Hb F, Hb A, Hb S, and Hb C *(45)*. Reversed-phase HPLC has been able to characterize about 200 Hb variants *(54)*. Nevertheless both electrophoresis and HPLC only provide presumptive identification of Hb variants and require DNA analysis or amino acid sequence determination for definitive identification.

Other laboratory tests such as measurements for oxygen affinity or Hb stability are occasionally performed *(28, 55)*.

DNA techniques

With the advent of DNA methodologies, like the introduction of PCR and nucleotide sequencing, it has become common practice to sequence the coding regions as well as parts of the intervening sequences of the α- and β-globin genes *(56, 57)*. When known mutations are suspected, they may be detected either by restriction enzyme cleavage if appropriate sites are present or by hybridization techniques *(58)*. For large gene deletions, Gap PCR may be used *(59)*. This technique is based on the principle that the primers located far apart in native DNA will not undergo direct amplification unless a deletion brings them into close proximity.

PCR-based techniques are indeed extremely powerful tools for characterization of mutations. It is of great use to analyze mutations that leads to severe instability of the protein. The DNA techniques albeit cannot detect post-translational modifications of proteins but could provide information regarding the origin of such changes. However, application of the PCR techniques to the α-globin genes is somewhat more complex than the β-globin gene, because of the presence of the GC-rich nucleotide content *(18)*. This problem can be solved by disruption of the secondary structure by adding chemical agents such as dimethylsulfoxide in the PCR *(59)*. Furthermore, identification of novel mutations on the α-globin or the γ-globin chain calls for nucleotide sequencing of two homologue genes since both α- and γ-genes are duplicated.

Mass spectrometry

Mass spectrometry measures mass to charge (*m/z)* ratio of gas-phase ions traveling through electrical and/or magnetic fields, and the use of current "soft ionization" methods allow for the determination of molecular weights of proteins ranging from a few kDa to over a MDa. It could be argued that the proteomic era has, in fact, been driven by the introduction of the "soft ionization" methods of matrix-assisted laser desorption/ionization (MALDI) and electrospray (ES) *(60-62)*. In recognition of the immense role played by these ionization methods, when applied to solving the problems of identification and structural characterization of proteins by mass spectrometry, John B. Fenn and Koichi Tanaka were awarded the 2002 Nobel Prize in Chemistry. These ionization methods together with a new

generation of high-performance instruments are adding much needed analytical power to the arsenal of techniques for macromolecular identification and structural characterization *(61, 62)*.

Many of the new generation of mass spectrometers are based on time-of-flight (TOF) technologies that have permitted fast spectral acquisition (1 s or less) with relatively good resolution [>10,000 full width at half maximum height (FWHM)] and excellent mass accuracy [(low parts per million (ppm)]. These instruments have a theoretically unlimited mass range (> 10^6 Da) *(63)*. Lower mass resolution (~3000 FWHM) quadrupole mass filters are relatively inexpensive and continue to play a major role in biomolecule analysis. Ultrahigh mass resolution (>10^6 FWHM) can be achieved with Fourier transform ion cyclotron resonance (FTICR) mass analyzers *(64)* however such instruments are costly. Quadrupole ion traps are much less expensive and can be used to sequence peptides and proteins *(65)*. In recent years quadrupole and TOF analyzers have been combined in hybrid instruments such as in the Micromass and Sciex quadrupole time-of-flight (Q-TOF) tandem mass spectrometers.

The usage of the terms “exact” or “accurate” mass, “monoisotopic” and “average” mass in defining molecular mass can be confusing. For small molecules, it is possible to identify the monoisotopic species in the $[M+nH]^{n+}$ ion clusters and measure their mass accurately to 5-10 ppm *(66)*. This gives the accurate or exact mass of the monoisotopic species. The monoisotopic mass corresponds to the mass of a molecule, or an ion, which is made up from elements in their lightest stable isotopic form. For larger molecules exceeding M_r > 10000, it is generally very difficult to observe a monoisotopic ion and thus mass spectrometry determines the average mass of the molecule *(67)*. The average molecular mass is the sum of the abundance-weighted masses of all stable isotopes (e.g. 98.9% ^{12}C and 1.1% ^{13}C) *(68)*.

MALDI

MALDI is the most efficient method to ionize peptides. In MALDI, sample co-crystallized with a light-absorbing matrix is irradiated usually with a nitrogen-laser beam. The absorbed light energy dissipates and induces vaporization and formation of protonated (or

deprotonated) molecules of the samples and matrix. The protonated molecules are then accelerated under an electric field and analyzed usually by a TOF mass analyzer *(61-63)*. MALDI generates predominantly singly charged ions possibly due to re-neutralizing effects by electrons and by excess deprotonated matrix ions *(69)*. This in turn produces a mass spectrum that can be easily interpreted.

In contemporary protein chemistry, MALDI has been primarily used for mass mapping of peptides from a protein following protein separation by two dimensional gel electrophoresis and subsequent proteolytic/chemical cleavage of the protein *(70)*. The peptide masses are searched against a protein or DNA database and identification of the parent protein is reached *(70)*. MALDI tolerates higher salt concentration than ES methods but is limited in the low mass range due to matrix-associated chemical noise.

ES

ES is the method of ionization/vaporization suitable for the widest range of polar biomolecules *(71)*. The ES process can be thought of in simple terms *(72-74)*: sample when dissolved in solvent will to a certain extent exist in an ionized form, i.e. $[M+H]^+$ or $[M-H]^-$. In conventional ES, the sample solution is pumped through a thin conducting capillary (50-100 μm orifice diameter) raised to a high potential (3-4 kV) *(71, 73)*. A lower potential (~1 kV) is required for nano-ES that uses metal-coated glass capillaries (1-10 μm orifice diameter) *(75, 76)*. An ES is generated as a result of a potential difference between the capillary and a counter electrode (Fig. 5).

Charged droplets accumulate at the tip of the capillary, their polarity depending on polarity of the applied electric field on the wall of the needle. The droplets are sprayed into a bath gas at atmospheric pressure and travel down a pressure and potential gradient towards the inlet of the mass spectrometer high vacuum system. As the droplets traverse this path they become reduced in size until the coulombic forces of the surface overcome the surface tension and the droplets begin to disintegrate into smaller droplets *(72, 74)*. This process continues until the point is reached that either an ion leaves from a droplet or an ion is completely desolvated, and the ions enter the gas phase. Ions are then focused into the mass

analyzer where they are separated according to their *m/z* ratio. Thus, a pure protein which may be multiply protonated, and exists in many different charged states will give a spectrum

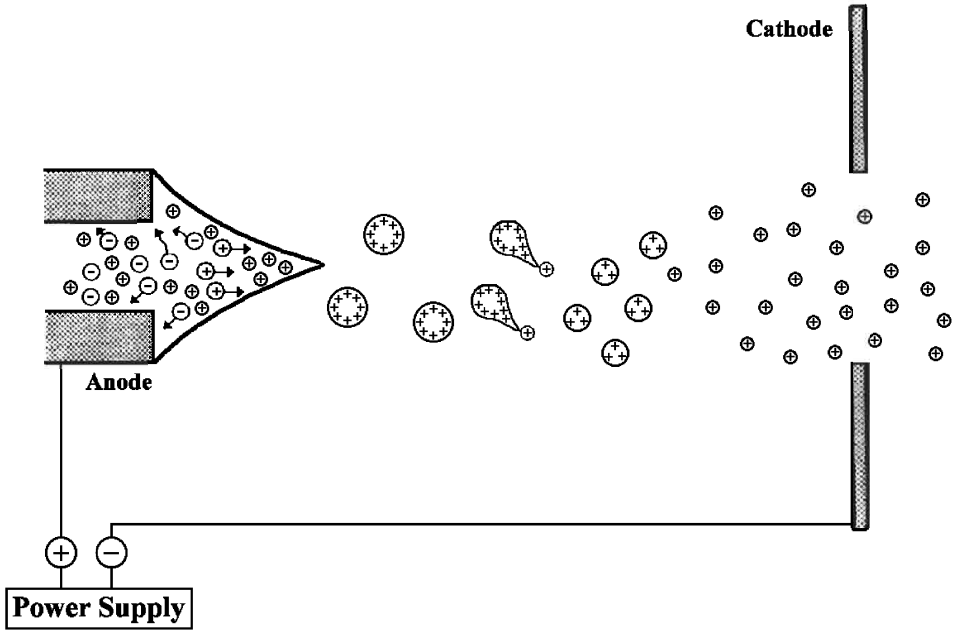

Fig. 5. Schematic representation of the ES process.
Adapted from *(72)* with permission from Kebarle P.

composed of a series of peaks (Fig. 6). Each peak corresponds to the mass (M) of the protein plus the number (n) of added protons (H), i.e. $[M+nH]^{n+}$. The different charged species of proteins can be deconvoluted to the true molecular masses of the neutral protein using deconvolution software such as Maximum Entropy (MaxEnt) *(77)* (Fig. 6 inset).

ES ionization is easily coupled to liquid-based separation system such as HPLC *(78, 79)* or CE *(49)*. The coupling provides the advantages of cleaning up and concentrating the sample, as well as for automation of sample analysis. The ion signal in the ES process depends on the concentration, but not on the flow rate of the injected sample. Hence in a sample-limited situation, it is advantageous to reduce the flow rate as much as possible by miniaturizing the separation system *(75)*.

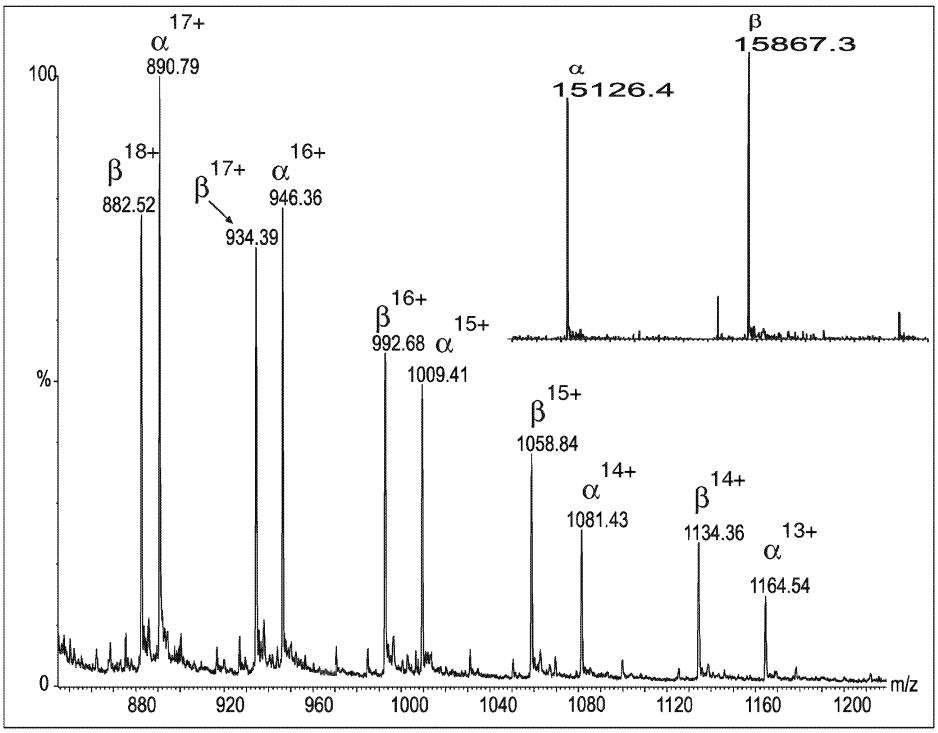

Fig. 6. ES mass spectra of the normal Hb chains showing a series of $[M+nH]^{n+}$ forms of the major globin chains. Shown in the inset is the MaxEnt deconvolution to their true molecular mass scale [α-chain (M_r 15126.38) and β-chain (M_r 15867.24)].

Tandem mass spectrometry (MS/MS)

A tandem mass spectrometer consists of two mass analyzers arranged in series with a collision cell in between (Fig. 7). The mass analyzers may be arranged in series either in space or in time *(80)*. There are two modes of operation in which the instrument can be used: (i) MS mode: ions entering the spectrometer are analyzed by MS_1 to give a mass spectrum (no gas in collision cell and MS_2 acts as a transport region); (ii) MS/MS mode: MS_1 transmits ions of a particular *m/z* which undergo collisional activation with inert gas in the collision cell. Fragment ions produced in the cell are analyzed by MS_2 (Fig. 7) *(80)*.

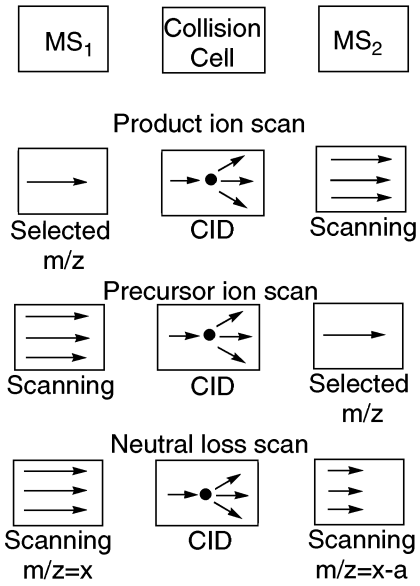

Fig. 7. MS/MS experiments on spatially separated mass analyzers.
Adapted from *(80)* with permission from © John Wiley & Sons Ltd.

In recent years an exponential growth in the application of MS/MS for the sequence analysis of proteins and peptides has been seen *(61, 65, 71)*. The most commonly used ion-activation technique for dissociation of peptide ions is low-energy collision induced dissociation. Collisions between precursor ions and inert gas molecules, e.g. argon, results in the excitation of the precursor ions that may lead to fragmentation as a consequence of conversion of translational energy into internal energy. Unstable precursor ions decompose to product ions.

The nomenclature used to describe product ions was initially proposed by Roepstorff and Fohlman *(81)* and was later modified by Biemann *(82)* (Fig. 8a). According to this nomenclature, the fragment ions resulting from the C-terminal ends are assigned x, y, z while those fragment ions from the N-terminal ends are designated a, b, c ions.

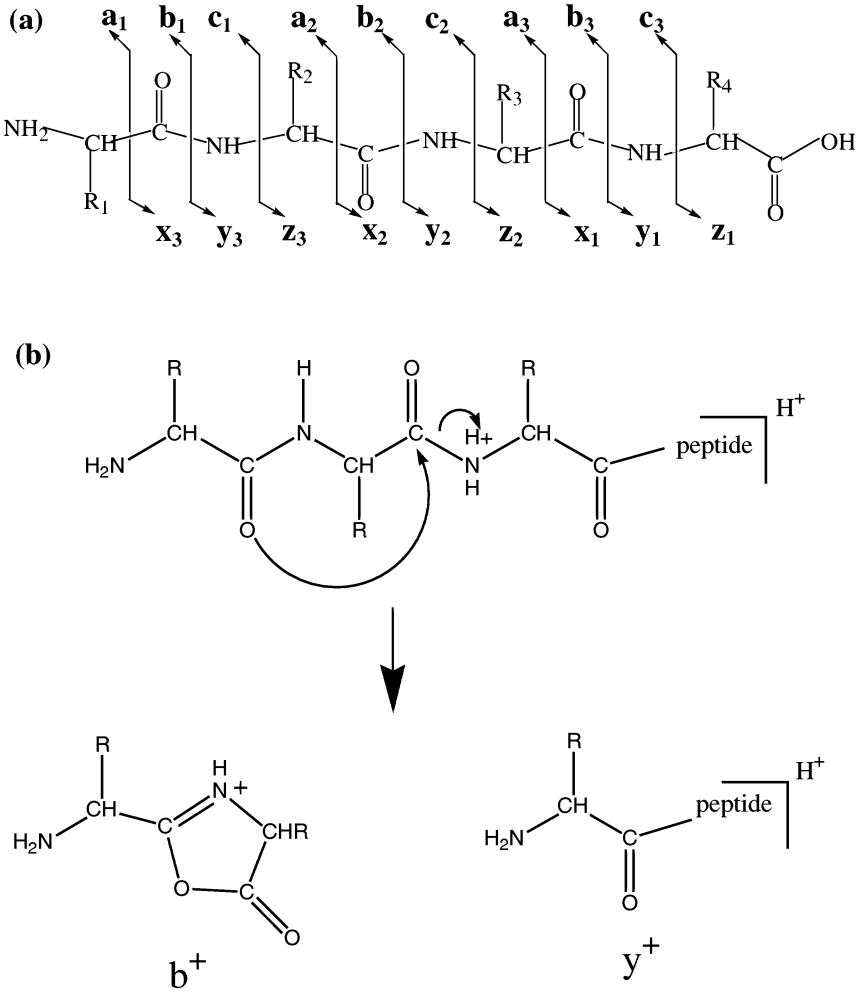

Fig. 8. (a) Nomenclature for MS/MS fragment ions of proteins and peptides. (b) Schematic representation of peptide fragmentation via the mobile proton model.

The generally accepted mechanism of peptide fragmentation is via the mobile proton model *(83-87)*. For example, in a doubly protonated trypsin cleaved peptide, one proton will be localized on the C-terminal basic residues (Arg or Lys) and a second proton may be localized at one of the amide bonds (or the N-terminus) *(83)*. During low energy collision, the second proton can be mobilized generating a heterogeneous population of ions where the second proton may reside at any of the amide bonds along the peptide chain. Protonation of an amide-nitrogen destabilizes the amide bond and nucleophilic attack by a carbonyl oxygen on the carbonyl carbon of the amide bond leads to cleavage of the amide bond (Fig.

8b). This cleavage of the amide bond produces a complementary pair of b- and y-ions *(83, 84)*. Series of b- and y-ions are generated which allow for the determination of the amino acid sequence. Generally, for tryptic peptides y-ions are found in greater abundance than the b-ions, because the latter undergo further fragmentations *(83, 85, 86)*. Furthermore, the y-ions formed by cleavage at N-terminal to Pro residues, and fragment ions formed due to cleavage at C-terminal to acidic (Asp, Glu and cysteic acid) residues generally give abundant ions in a tandem mass spectrum. The mechanism behind this also can be explained by the mobile proton model *(87)*.

Mass Spectrometry and Hb variants

The use of mass spectrometry to characterize Hb variants dates back to 1981 when Wada and colleagues used field desorption mass spectrometry for the mass analysis of tryptic peptides from Hb S *(88)*. Over the years, different mass spectrometric techniques, e.g. plasma desorption *(89)*, fast-atom bombardment *(90, 91)*, have been used in characterizing Hb variants at the peptide level. Today the ionization techniques of ES *(92-95)* and MALDI *(96-98)* have become dominant. They have the advantage of sensitivity and can ionize whole proteins. Over a hundred different types of Hb variants have been detected at the intact globin level by mass spectrometry, of which more than 20 have been due to novel mutations *(94, 95)*. An interesting application of mass spectrometry to Hb variant analysis was the detection of Hb Quebec-Chori (β87Thr→Ile), an electrophoretically silent Hb variant, which when present in combination with Hb S can give a severe form of sickle disease *(99)*. Glycosylation *(100)* and glutathionylation *(14)* of the globin chain can readily be observed using ES mass spectrometry, and experiments have been made to relate their abundance with clinical disorders.

The average masses of α- and β-globin chains are 15126.38 Da and 15867.24 Da respectively. The difference in the molecular mass between two amino acids ranges from 0 to 129 Da (Table 3). Hence, most amino acid substitutions from a single base substitution in Hb variants can be observed directly by mass spectrometry except those resulting in zero Da

Table 3. Mass differences of amino acid exchanges resulting from single base substitution

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

change. However, there is a limitation of mass spectrometry, in that it can only resolve peaks which differ in mass by >6 Da for a 15-16 kDa proteins. Hb variants that differ in mass by <6 Da thus usually require tryptic peptide mass analysis for detection and identification, including the Lys↔Gln substitution that has a zero nominal mass change. Few Hb variants have elongated globin chains and their mass can differ by >129 Da from the mass of the corresponding normal globin *(93, 94)*. Hb variants due to nucleotide deletions or insertions or cross-over of genes can also differ in mass from those outlined in Table 3 *(93, 94, 101)*.

Identification of Hb variants using mass spectrometry now is generally performed in three steps (Fig. 9): (a) detection of the type of globin chain variant at the intact level, (b) peptide mass mapping of the globin chains following a proteolytic cleavage, commonly by use of trypsin, and (c) sequence analysis of the peptide of interest by MS/MS. By this strategy the vast majority of the Hb variants can be detected, and some of them may also be identified, at the peptide level. For this reason, studies in the past few years have been focused on developing a rapid digestion protocol *(94, 102, 103)* and implementation of fast peptide separation techniques prior to mass analysis *(104)*.

Trypsin is the standard proteolytic enzyme used for Hb digestion. At least 14 peptides from the α-chain and 15 from the β-chain are generally observed in a mass spectrum following trypsin digestion (Table 4). Endoproteinases such as chymotrypsin can be added together with trypsin to generate moderately sized peptides from the tryptic peptides larger than 2 kDa *(93, 94)*. The globin chains are sometimes subjected to chemical derivatization of thiol groups of Cys residues prior to enzymatic cleavage *(92, 93, 105)*. This may help to identify the Hb variant when the amino acid substitution occurs in the core region of the globins *(105)*. Alternatively a 10-minute dithiothreitol (DTT) reduction on the digested globins may serve the same purpose *(94, paper III)*. Protein sequence analysis is necessary to positively identify the Hb variant.

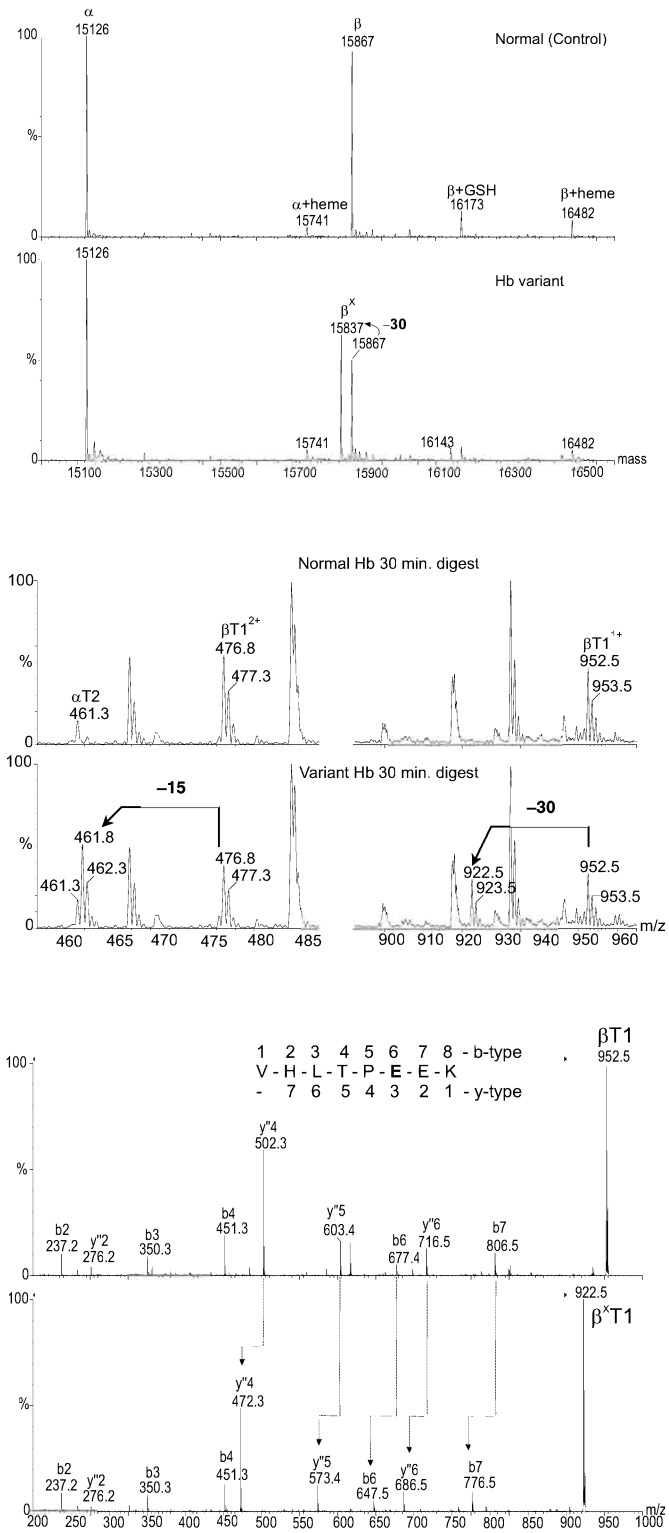

Fig. 9. Three stages of mass spectrometric investigation of Hb variants.

(a). Deconvoluted ES mass spectra of normal Hb (top) and the variant Hb (bottom). The variant Hb is in the β-chain (β^x) and is 30 Da lighter than the normal β-chain.

(b). Part of the mass spectra from 30 min. trypsin Hb digests showing a –30 Da change of $\beta T1^{1+}$and $\beta T1^{2+}$ peptide ions in the variant Hb (bottom) with respect to normal Hb (top).

(c). MS/MS spectra of normal $\beta T1^{1+}$ peptide ions (top) and the variant $\beta^x T1^{1+}$ peptide ions (bottom). Glu→Val substitution is localized to residue 6 by change of –30 *m/z* in b_6- and b_7-ions, characteristic of Hb S (β6Glu →Val).

Table 4. Tryptic peptides from Hb A observed in the ES mass spectrum*.

Peptide	Residues	Sequence	$[M+H]^+$	$[M+2H]^{2+}$	$[M+3H]^{3+}$	$[M+4H]^{4+}$
αT1	1-7	VLSPADK	729.41	365.21	-	-
αT1-2	1-11	VLSPADKTNVK	1171.67	586.34	-	-
αT2	8-11	TNVK	461.27	231.14	-	-
αT3	12-16	AAWGK	532.29	266.65	-	-
αT4	17-31	VGAHAGEYGAEALER	1529.73	765.37	510.58	-
αT5	32-40	MFLSFPTTK	1071.55	536.28	-	-
αT6	41-56	TYFPHFDLSHGSAQVK	1833.89	917.45	611.97	459.23
αT7	57-60	GHGK	398.22	199.61	-	-
αT8	61	K	147.11	-	-	-
αT8-9	61-90	KVADALTNAVAHVDDMPNALSALSDLHAHK	**3124.58**	1562.80	1042.20	781.90
αT9	62-90	VADALTNAVAHVDDMPNALSALSDLHAHK	**2996.49**	1498.75	999.50	749.88
αT10	91-92	LR	288.20	-	-	-
αT10-11	91-99	LRVDPVNFK	1087.63	544.32	-	-
αT11	93-99	VDPVNFK	818.44	409.72	-	-
αT12	100-127	LLSHCLLVTLAAHLPAEFTPAVHASLDK	**2966.61**	1484.31	989.88	742.66
αT12-13	100-139	LLSHCLLVTLAAHLPAEFTPAVHASLDKFLA SVSTVLTS	**4203.98**	2102.49	1402.00	1051.75
αT13	128-139	FLASVSTVLTSK	1252.72	626.86	-	-
αT14	140-141	YR	338.18	-	-	-
βT1	1-8	VHLTPEEK	952.51	476.76	-	-
βT2	9-17	SAVTALWGK	932.52	466.76	-	-
βT3	18-30	VNVDEVGGEALGR	1314.67	657.84	-	-
βT4	31-40	LLVVYPWTQR	1274.73	637.87	-	-
βT5	41-59	FFESFGDLSTPDAVMGNPK	2058.95	1029.98	-	-
βT6	60-61	VK	246.18	-	-	-
βT7	62-65	AHGK	412.23	-	-	-
βT8	66	K	147.11	-	-	-
βT8-9	66-82	KVLGAFSDGLAHLDNLK	1797.99	899.50	600.00	-
βT9	67-82	VLGAFSDGLAHLDNLK	1669.89	835.45	557.30	-
βT10	83-95	GTFATLSELHCDK	1421.67	711.34	-	-
βT10-11	83-104	GTFATLSELHCDKLHVDPENFR	**2529.22**	1265.11	843.75	633.06
βT11	96-104	LHVDPENFR	1126.56	563.79	-	-
βT12	105-120	LLGNVLVCVLAHHFGK	1719.97	860.49	574.00	-
βT13	121-132	EFTPPVQAAYQK	1378.70	689.85	-	-
βT14	133-144	VVAGVANALAHK	1149.67	575.34	-	-
βT14-15	133-146	VVAGVANALAHKYH	1449.80	725.40	483.94	-
βT15	145-146	YH	319.14	-	-	-

* Abundance of peptide ions containing Cys residue (αT12, βT10 and βT12) improves after reduction with DTT; *m/z* marked in bold are not generally observed in the ES mass spectrum.

A drawback of ES mass spectrometry in the study of Hb variants is that it cannot show two different species that differ in mass by < 6 Da *(93, 94)* as separate entities in a mixture at the intact protein level. Although very high-resolution instruments such as the FTICR spectrometer may provide a solution *(106)*, the statistical variation of isotope distribution makes it difficult to interpret the mass spectra, particularly when dealing with a mixture. On the contrary, high mass accuracy plays a greater role in detection of closely related species in terms of mass, and can be achieved using commercial quadrupole instruments under optimized conditions *(paper V)*. Alternatively, separation techniques such as HPLC and CE can be interfaced to mass spectrometry to detect Hb variants with small mass difference from the normal Hb chains *(92, 106, 107)*.

ES mass spectrometry can also be used to study non-covalent protein complexes (Fig. 10). Green and colleagues have shown the presence of asymmetric Hb tetramers in a mixture of normal Hb and Hb variant *in vitro* which cannot be normally demonstrated using other analytical tools *(108)*. Recently a hexa-decamer of Hb Porto Alegre (β6Ser→Cys) has been demonstrated using ES mass spectrometry under non-denaturating conditions *(109)*.

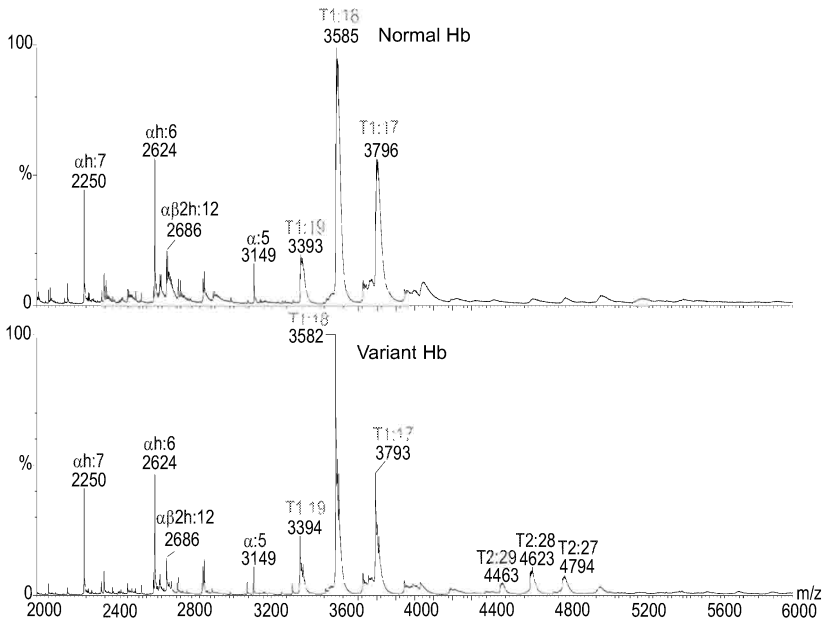

Fig. 10. Non-covalent studies using ES mass spectrometry. (a) Normal Hb ($\alpha_2\beta_2h_4$, T1). (b) Hb Ta-Li (β83Gly→Cys) showing minor proportions of octamers (T2, α_2 β $h_4\beta^{-S-S-}\beta\beta\alpha_2$ h_4). "h" denotes the heme group.

Adducts with alkali metals can complicate the interpretation of mass spectra *(paper I)* in addition to suppressing the protonated forms of proteins and peptides. Green and co-workers have introduced an efficient way of desalting proteins using ion-exchange resin beads *(94, 108)* for macromolecular studies. Commercially available micro-columns packed with C_4 or C_{18} medium are also used for desalting proteins and peptides prior to mass analysis *(110)*.

The sequencing of intact globin chains by MS/MS *(111-114)* may increase the speed of diagnosis of Hb variants and thereby reduce the dependency on peptide mass mapping and other protein separation techniques *(paper II)*.

Automation of mass spectrometry techniques, including purification of proteins, trypsinization, peptide mass mapping and MS/MS, has become available *(78, 79, 104)*, however, at present the characterization of Hb variants involving these steps is largely performed manually *(94, papers I to V)*.

AIMS OF THE PRESENT STUDY

Molecular diseases may be caused either directly or indirectly as a consequence of abnormal protein production. The protein abnormalities can be qualitative or quantitative in nature. Hb disorders are good prototypes of molecular diseases. In this study, we have attempted to evaluate and optimize mass spectrometric methods for the investigation of aberrant Hb chains with emphasis on:

(a) Characterization of Hb variants by ES mass spectrometry at the intact globin level and by MS/MS at the tryptic peptide level (Paper I and III).

(b) Use of MS/MS at the intact chain level for identification of Hb variants (Paper II).

(c) Characterization of Hb variants displaying post-translational modifications (Paper III)

(d) Detection and characterization of electrophoretically silent Hb variants by ES-MS/MS (Paper IV).

(e) Detection and characterization of Hb variants with masses close to the normal (Paper V).

MATERIALS AND METHODS

Sample History

Whole blood samples collected in K_3 ethylene diamine tetraacetic acid (EDTA) vacutainers were used. The blood samples underwent one or more of the routine hematological investigation, i.e. FBC, IEF and HPLC analyses. The samples were either stored at + 4 °C or frozen at – 20 °C prior to ES mass spectrometric analysis.

Mass Spectrometry

In most of the experiments, either a triple staged quadrupole or a Q-TOF instrument, both manufactured by Micromass (Manchester, UK) were used. The details of the experimental conditions are described in the respective papers.

Sample preparation

Stock solutions of 50-fold dilution from 10 µL EDTA whole blood samples were prepared in deionized water. A further 10-fold dilution of the stock solution was made in solvent A (50% aqueous acetonitrile or aqueous methanol, both containing 0.2% formic acid) for mass spectrometric analysis using a conventional flow-rate ES interface, i.e. on average ~8 µM concentration of each major globin chain was used. A higher dilution (1:50) of the stock solution was made in solvent A for nano-ES experiments. Desalting of the stock solution to reduce alkali metal adducts was performed using 3 kDa cut off micro-concentrators (Nanosep centrifugal concentrators, Northborough, USA) or pre-washed cation-exchange resin beads (AG 50W-X8, Bio-Rad Laboratories, Hercules, CA). The remaining stock solutions were frozen at –20 °C for future use, e.g. for trypsin digestion.

DTT reduction and Alkylation

Alkylations of thiol group of Cys residues in the globin chains prevent intra- and inter-molecular disulfide bond formation during sample preparation, i.e. following enzymatic digestion. Reduction of the globin chains was effected by adding 10 µL of 100 mM DTT and 120 µL of acetonitrile to 120 µL of Hb stock solution. S-carboxyamidomethylation of Cys residues in the globin chains was performed by mixing all of the DTT reduced Hb

solution with 10 µL of 1 M aqueous ammonium bicarbonate (pH 8.0) and 12 µL of 200 mM aqueous iodoacetamide solution. The mixture was incubated for 15 min at room temperature. A working sample solution was prepared by mixing 20 µL of the reduced sample with 180 µL of solvent A and analyzed on the mass spectrometer to confirm the appropriate derivatization. The remainder of the derivatized Hb sample was subjected to tryptic digestions after removing the excess reagents using a 10 kDa cut off centrifuge filter (Microcon YM-10, Millipore Corporation, Bedford MA).

Trypsin Digestion

100 µL of the stock solution was incubated with L-tosylamide 2-phenyl ethyl-chloromethyl ketone (TPCK)-treated trypsin (Sigma-Aldrich, Sweden) in ~50 mM NH_4HCO_3 buffer (pH 8.0) at 37 °C overnight (~14 h). The concentration of the total protein in the digestion was approximately 2.5 mg/mL and the ratio of trypsin to globins was about 1:10 (w/w). Portions (20 µL) of the digested samples were mixed with 180 µL of solvent A and directly mass analyzed and the remaining digested samples were stored at –20 °C. This procedure was modified in the later stages of the studies where the globin chains were denatured in a final concentration of 0.01% formic acid and 0.1% acetonitrile for 5 min at room temperature prior to trypsinization *(94, paper III)*. An aliquot, following a 30-min digestion, was taken and mixed in solvent A (1:10 v/v) for mass analysis, while the rest was left for overnight digestion.

DNA sequencing

Genomic DNA covering the three exons of the globin gene was amplified using standard PCR methods. Direct nucleotide sequencing confined to the region of interest in the amplified product was performed using appropriate primers.

Database Search

A globin server website (http://globin.cse.psu.edu) was mined to extract additional information on the results obtained at various stages of the investigation.

RESULTS AND DISCUSSION

Characterization of Hb variants using ES mass spectrometry at the intact level and MS/MS at the tryptic level: studies on low proportion Hb variants (Paper I)

Early studies of Hb variants using mass spectrometry were mainly focused on the β-chain, which generally occur in high abundance (~50% total Hb) *(93, 95)*. The α-chain variants generally occur in relatively lower abundance (18-35% of total Hb) *(18)* and therefore may escape detection. This may partly explain the reason for the lower number of variants in the α-chain than in the β-chain described to date (Table 2).

In order to evaluate the use of ES mass spectrometry for the identification of abnormal α-chains, two Hb variants previously identified by IEF, HPLC and DNA sequencing were initially investigated. The ES mass spectra from the intact proteins demonstrated α-chain variants that were 22 Da and 48 Da heavier than the normal α-chain respectively. For the former mass change there was only one possible amino acid substitution (i.e. Asp→His, Table 3) which agreed with the cathodal (gain in positive charge) nature of the variant with respect to Hb A on IEF. For the +48 Da change there were two possible amino acid substitutions (Asp→Tyr and Val→Phe, Table 3). Integrating the information from IEF, where the variant was cathodal to Hb A, the two alternatives were reduced to one, i.e. Asp→Tyr. There are eight Asp residues in the normal α-globin chain distributed on five different trypsin cleaved peptides (Table 4). The ES mass spectrometry analysis of the tryptic peptides localized the amino acid substitution to peptide αT6. Subsequent analysis of the variant peptide by MS/MS revealed the expected amino acid substitution.

A similar approach was used for three unidentified aberrant α-chains. Two variants were identified by ES mass spectrometric techniques and their corresponding nucleotide substitutions were subsequently confirmed by DNA sequencing, whereas for the third variant the mass spectrometric identification did not agree with the DNA results. Our initial mass spectrometric interpretation was incorrect due to confusion caused by the presence of Na^+ adducts (Δm = +22 Da) in the crude sample. However, with a knowledge of the DNA result, the mass spectrometric data was readily explained. Furthermore, for three of the five

α-chain variants analyzed, the quantity of variant chain as determined by ES mass spectrometry could be correlated with data from HPLC (Table 5). The remaining two samples showed much higher values than the HPLC, which is due to interference from Na^+ adducts.

Table 5. Summary of α-chain variants identified using ES mass spectrometry

Hb variant	Substitution	$\Delta m^{\dagger}$ (Da)	% HPLC[‡]	% ES-MS[§]
Hb Hasharon	α22 Asp→His	+22	18	35
Hb Kurdistan	α47 Asp→Tyr	+48	19	22
Hb Russ	α51 Gly→Arg	+99	24	28
Hb Le-Lamentin	α20 His→Gln	-9	20*	19
Hb Q-Iran	α75 Asp→His	+22	19*	34

We conclude that with a background knowledge of the amino acid sequence of the α-globin and information from HPLC and IEF, ES-MS/MS can be an attractive alternative to DNA sequencing for rapid identification of α-chain variants. However, misinterpretation of the mass spectral data could occur when the Hb variant shares the same mass gain as that of an alkali metal adduct of the normal Hb chains.

Characterization of Hb variants by MS/MS on the intact β-chains (Paper II)

ES mass spectrometry of Hb variants often involves protein separation, derivatization, and digestion of the aberrant globin chains *(92, 93, 95)*. Over a decade ago, Shackleton and co-workers introduced the MS/MS analysis of intact β-chains (normal and aberrant, ~16 kDa) using a triple quadrupole instrument on purified globins *(111)*. We have evaluated the use of the Q-TOF hybrid instrument for MS/MS studies of intact β-chains following a one-step dilution of the whole blood. The purpose of the experiment was to integrate MS/MS data in the rapid diagnosis for β-chain variants. MS/MS experiments were performed on several samples known to contain β-chain variants. About 50% of the primary structure of the polypeptide, mostly from regions close to the two termini, was obtained. Key fragment ions

† mass difference between the variant and normal α-chain.
‡ Hb X/(Hb X+ Hb A)
§ calculated from the base peak intensities of globin chains, i.e. $\alpha^X/(\alpha^X+\alpha)$
*from globin server database *(27)*

in the investigation of Hb variants by MS/MS of the intact β-chains were identified (Table 6).

Table 6. Some commonly observed b- and y-ions in the MS/MS spectra of normal intact β-chain of Hb.

m/z	b-ion	charge	*m/z*	y-ion	charge
682.4	13	2+	626.3	12	2+
738.9	14	2+	675.9	13	2+
832.0	15	2+	671.1	19	3+
759.1	21	3+	811.6	23	3+
802.1	22	3+	608.8	23	4+
1086.3	31	3+	781.5	36	5+
1124.0	32	3+	878.0	89	11+
1157.0	33	3+	1034.3	96	10+
1190.7	34	3+	862.1	96	12+
1245.1	35	3+	940.4	96	11+
1308.8	47	4+	934.5	111	13+

Identification of variants was achieved in two ways: (i) some cases were identified solely by MS/MS of the intact β-chain and (ii) in others the investigation directed the mutation to certain segments of the polypeptide (Fig. 11). For the latter cases analysis of tryptic fragments was required for complete identification. Following a similar procedure, two presumptive β-chain variants were identified.

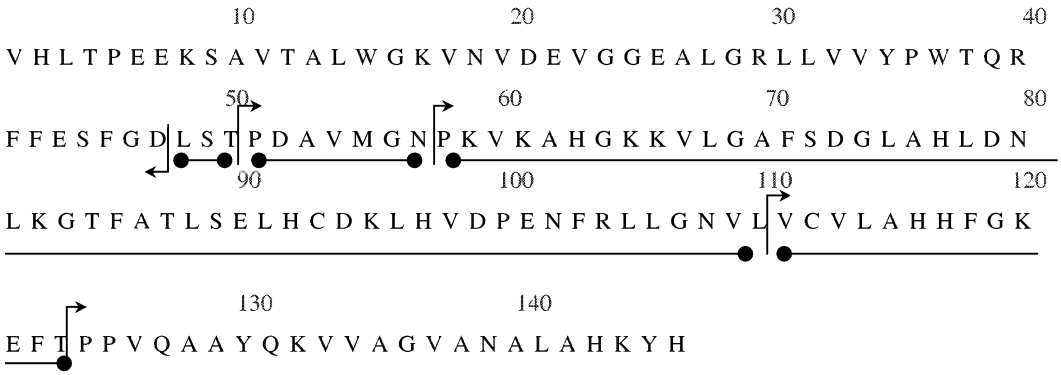

Fig. 11. Amino acid sequence of the β-globin chain showing major cleavage positions in MS/MS experiments. Detailed sequence information for N-terminal residues (1 to 47) and C-terminal residues (124 to146) could be obtained from the MS/MS data. Amino acid substitutions occurring in the middle region can be confined to smaller regions (●—●).

As in paper I, identification was greatly aided by a background knowledge of the protein sequence of the β-globin and behaviour of the Hb variant on IEF and HPLC. Introduction of MS/MS of the intact β-chain can minimize the use of additional steps of globin separation and derivatization and sometimes also digestion.

Identification of a β-chain variant which undergoes a secondary structural change by ES-MS/MS (Paper III)

Papers I and II showed the classical use of ES-MS/MS, i.e. determination of primary structure of proteins and peptides of interests. In this paper, we were able to identify a Hb variant that undergoes a higher structural change, i.e. formation of a disulfide (-S-S-) bridge. This phenomenon may be seen only when a Cys residue is being introduced into the aberrant chain since there are no –S-S- bonds in normal Hb tetramers. Blood samples from two different patients that presented similar anodal variants on IEF were investigated. The initial mass spectra for both the cases demonstrated a β-chain variant apparently 45 Da heavier than the normal β-chain (Fig. 12a). The proportion of the variants on different mass

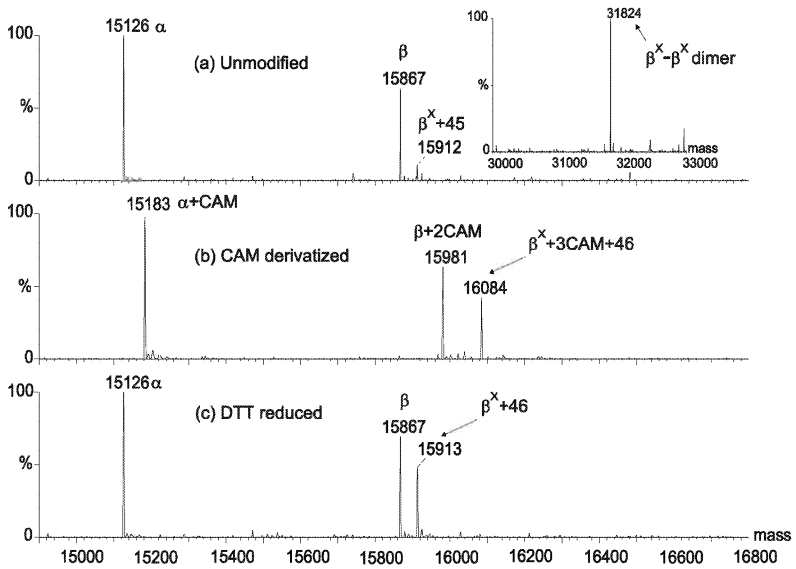

Fig. 12. ES mass spectra for Hb Ta-Li (β83 Gly→Cys): (a) unmodified Hb sample and insert showing the deconvoluted mass spectrum to higher mass range, (b) carboxyamidomethylated (CAM), and (c) DTT reduced Hb samples.

spectrometric runs decreased from 26% to 15%. There is no amino acid substitution corresponding to a +45 Da change (Table 3). Therefore, amino acid substitutions resulting in mass shifts of +44 Da and +46 Da were considered. Two amino acid substitutions, Ala→Asp and Cys→Phe, are possible for a +44 Da, and Gly→Cys for a +46 Da change (Table 3). However, Cys→Phe and Gly→Cys substitutions were initially ruled out as they are neutral to neutral amino acid substitutions and do not agree with the IEF pattern.

Apparently the only possible amino acid substitution was Ala→Asp. Since an Ala residue is found in nine of the fifteen tryptic peptides of the β-chain (Table 4), MS/MS on the intact variant β-chain was pursued. The MS/MS spectrum indicated that the amino acid substitution occurred in the region covered by tryptic peptides βT10 to βT12. The peptides βT10 to βT12 contain two Ala residues, each in peptide βT10 and βT12 (Table 4). The two peptides also contain a Cys residue each. Hence carboxyamidomethylation (CAM) of the globin chains was performed prior to trypsinization. An aliquot of the modified sample was analyzed to check the products of the reaction and the residual was left to digest. This was the turning point in the analysis.

The mass spectra of the CAM derivatized intact globin chains indicated the presence of an additional Cys residue in the variant (Fig. 12b). A reduction, using DTT, of the intact globin indicated that the mass shift was 46 Da (Fig. 12c). There is only one amino acid substitution that gives a +46 Da change, i.e. Gly→Cys. The slight acidic nature of this variant on IEF could possibly be due to contribution of some of the Cys residues being oxidized to sulfoxide (SO^-), sulfinic oxide (SO_2^-) or sulfonate (SO_3^-), although no significiant amounts could be found in the mass spectra. Using the information from the intact MS/MS spectrum of the variant, the mass analysis of the tryptic peptides was focused on peptides βT10-βT12. The variant was localized to peptide βT10 (GTFATLSELHCDK). There is only one Gly in this peptide corresponding to residue 83 in the normal polypeptide. Subsequent MS/MS of the variant peptide confirmed the Gly→Cys substitution. The variant was thus identified as Hb Ta-Li (β83Gly→Cys) reported once previously in a Chinese family more than 30 years ago *(115)*. The variant was described to polymerize *in vitro* forming multimers through inter-molecular –S-S- bridges *(115)*. This was confirmed by transforming the raw spectra to

a higher mass range spectrum appropriate to the dimer of the variant β-chains (Fig. 12a, insert).

This finding provides a reminder that the presence of secondary or other post-translation modification should be considered when data from independent experiments do not agree with the mass spectrometric results. A close examination of the raw data must be pursued when the disparity between independent techniques arises. An additional outcome of this study is the adequacy of a 10-min DTT reduction of the globin chains prior to perform mass spectrometric analysis of Cys containing proteins and peptides.

Detection and identification of electrophoretically silent variants by ES-MS/MS (Paper IV)

Thus far we have evaluated and refined the use of ES mass spectrometry for the investigation of Hb variants which were detected, but not identified, by routine techniques such as IEF and HPLC. In this paper we applied ES mass spectrometry to screen for electrophoretically silent samples from patients presented with erythrocytosis.

Erythrocytosis is sometimes caused by a Hb variant with high oxygen affinity *(36)*. The majority of the high oxygen affinity Hb variants are "silent", i.e. the amino acid substitutions involve residues having the same charge. Conventional separation techniques such as electrophoresis (in some cases even IEF) and ion-exchange HPLC will not detect silent Hb variants. The ideal method for screening high oxygen affinity Hb variants would be functional studies of oxygen saturation curves *(55)*. Such methods are scarce in a routine scenario and stored samples cannot be used for such studies. We analyzed more than 70 EDTA whole blood samples using ES mass spectrometry. These samples had shown high Hb levels on routine hematological investigation but displayed normal patterns on IEF and HPLC analyses. In these samples, three β-chain variants were detected by mass spectrometry. Two were identified to be Hb Olympia (β20Val→Met) following the three stages of ES mass spectrometric analysis. DNA sequence analysis confirmed the expected G→A mutation in codon β20. The third variant, Hb Coimbra (β99Asp→Glu), although detected with mass spectrometry, was identified by DNA sequencing and subsequently by ES-MS/MS. The reason for the latter strategy was that the mass shift at the intact protein

level gave three possible Hb variants that have previously been described, all with amino substitutions in the inner core region of the Hb molecule. Since inner core peptides can be difficult to analyze by MS/MS of non-derivatized peptides *(102, 105)*, DNA analysis was initiated.

We conclude that the ES mass spectrometry is an ideal method to investigate Hb variants causing erythrocytosis. A combined use of IEF and ES mass spectrometry could theoretically have 100% sensitivity to detect Hb variants described thus far causing erythrocytosis.

Application of ES mass spectrometry to detect Hb variants that differ in mass by <6 Da from the normal (Paper V)

One of the limitations in mass spectrometric investigation of Hb variants is the inability to differentiate two species that differ in mass by <6 Da as separate entities at the intact globin level *(93, 94)*. Several common Hb variants fall into this category, some of which are clinically significant. In these cases mass spectrometry will display one peak whose mass is the abundance weighted mean of the normal and the variant chains. Detection of the variant within the single entity requires high mass accuracy (<±7 ppm for 15-16 kDa to detect 10% Hb variant with 1 Da difference). Ultrahigh resolution may solve the problem *(106)*, but poor ion statistics over the isotopic cluster ions could present a problem for spectral interpretation as could the presence of a low proportion variant in the mixture.

We focused our effort on achieving an accurate mass measurement of the normal β-chain (M_r 15867.24) by using the α-chain as an internal calibrant and doubling the acquisition threshold to 32 data point per *m/z* of the quadrupole instrument. A high mass accuracy of <±5 ppm in the mass measurement of normal β-chain was achieved on 86 blood samples. With this level of mass accuracy we could detect Hb species that differ in mass by as little as 1 Da from the norm by the shift in the composite peak position, provided that the abundance of variant exceeds 10% in Hb pool.

A set of 19 blood samples with 13 different known mutations that yield a mass shift within ±6 Da from the corresponding normal chains was chosen. Background information of their IEF and HPLC data were compiled. The mass spectrometry analyses were then performed blind. A positive detection for all the 19 samples and their identifications were made from the mass change and the IEF data. Furthermore, there was a linear correlation (r^2 = 0.778) between the % variant estimated by this protocol and those by the HPLC techniques. Thus the protocol described could be used to detect and assign Hb variants that differ in mass by 1 to 4 Da from the corresponding Hb chains. The protocol could also be applied to other proteins by adding an internal calibrant to the sample that would allow for a similar level of mass accuracy to be achieved. This would allow for the determination of post-translational modifications such as amidation/deamidation and disulfide bridges, or various isoforms of proteins that yield small mass difference and could remain elusive to mass spectrometric analysis at whole protein level.

GENERAL SUMMARY

In this thesis, we have shown that ES mass spectrometry is a powerful tool for detection and identification of Hb variants with an unmatched speed and with simple sample pretreatment. Undoubtedly, this technique adds significantly to previous methods. The capability of modern methods of ionization which allow for molecular mass determination of the intact protein has greatly simplified the approach for Hb variant investigation. For instance, the observed mass difference, when backed up with data from the charge dependent separation methods, can streamline the peptide mass analysis to certain peptides where substitution is likely to occur. ES mass spectrometry can also give quantitative data when operated under appropriate conditions, e.g. using internal calibration. MS/MS of the intact globins, albeit rarely used, can identify Hb variants directly or assists the investigation by localizing amino acid exchange to a certain region of the globin chain. Refinement of mass spectral data acquisition parameters and data processing has enabled the detection of small mass differences at the intact level. This protocol could be used to screen for clinically significant Hb variants. The benefits for the analysis of other proteins are the detection of post-translation modifications and identification of protein isoforms with narrow mass differences.

CONCLUSIONS

- ES-MS/MS in combination with IEF and background knowledge of nucleotide sequence of Hb chains is an efficient technique to define Hb variants.

- Small sample volume (~10 μL), blood samples that are stored and untreated can be used for mass spectrometric analysis. However, adducts can complicate the mass spectra especially at the peptide level.

- Internal mass calibration and very careful mass spectra processing can give <±5 ppm mass accuracy (~16 kDa proteins) and hence allow to detect aberrant proteins with small mass differences.

- ES mass spectrometry can be a potential tool in screening for clinically significant Hb variants. This includes electrophoretically silent as well as common Hb variants S, C, E, etc.

- ES-MS/MS of intact globin chains may allow direct identification of Hb variants in some cases. In other cases amino acid substitution may be localized to a certain region of the polypeptide chain. Direct identification of the Hb variant is most likely to be successful when the substitution is present within 30-50 residues from either end of the polypeptide chain.

- Abnormal mass difference or disagreement between observed mass difference and electrophoresis/HPLC data might reveal a post-translational modification.

- Although Hb variants have been the subject of this thesis, protocols similar as applied here may be used in mass spectrometric studies on other proteins and their aberrant forms.

ACKNOWLEDGEMENTS

"The best portion of a good man's life ... are his little, nameless, unremembered acts of kindness and love."

-William Wordsworth

"Time flies", so has it been true for the past four years. A difference, however, was meeting many people who have helped me bringing this thesis to its present form. I take this opportunity to express my thanks to them:

My deep gratitude to my supervisor, **Britta Landin** for giving me the opportunity to study in Sweden. Many thanks for introducing me into the hemoglobin world and for her excellent guidance, inspiration and constructive criticism. **Bill Griffiths**, my co-supervisor and my mass spectrometry tutor, for his enthusiasm, encouragement and his positive attitude at all time. **Gunvor Alvelius**, my co-supervisor and my first mass spectrometry teacher, for starting this project and for her strong support throughout.

My collaborator **Brian Green**, Micromass, Manchester, for productive collaborations and for unending help in mass spectrometry and hemoglobin variants. It has been a great privilege to work with him.

Ingemar Björkhem for his support, enthusiasm and encouragement towards this project.

Gertrud Bertilsson for assistance in DNA experiments.

Past and present members of Clinical Chemistry Division: **Maud Lindblad, Laila Leymann** & **Jenny Bernström** for help in administrative matters. **Steve Meaney, Jim O'Byrne, Karl Bodin, Maura Heaverin, Ulla Andersson, Sara Von Bahr, Lu Qing, Leoni Valerio, Amir Babiker,** in doktorandrum for pleasant atmosphere and for intriguing discussions on all topics. **Eva Rystedt, Jin Wang, Mary Hunt, Maria Olin, Petra Holmström, Charlotte Murphy, Saeki Masayumi, Yi-zeng Yang, Mats Gåfvels** & **Gösta Eggersten** in DNA lab for help, patience and for joyful conversations. **Stefan Alexon** & **Inger Holmberg** for accelerating the enthusiasm to run mass spectrometry on other areas of biological sciences. **Karianne Solaas, Maria Westin, LiYue Huang, Raja Karim, Anita Lövgren, Afsaneh Azarshab, Pari Kabiri, Marianne Vallin, Ulf Diczfalusy, Margareta Diczfalusy, Priya Patel, Magnus Hansson, Inger Moberg** on plan 6 for "hej" or "hello" as well as for brief friendly chats. **Sonja** & others in phlebotomy, and **Juan Manuel, Lena Norberg, Kjell Kalén** & others from plan 7 for their patience on my intrusion into their routine work.

Past and present members in Clinical Research Center, NOVUM: **Martin Månsson, Håkan Yildrim, Sebastian Bauer, Malin, Sofia, Elke Schweda, Reine, Niklas, Anna Risberg, Anders, Per-Erik Jansson** for the wonderful time and also for interesting discussions during Wednesday "fika" mornings.

ACKNOWLEDGEMENTS

Past & present members in Solna campus: **Suya Liu, Johan Lengqvist, Jan Sjövall, Daiva Guvenius, Koidu Norén, Yang Yang, Åsa Brunnström, Agneta Tjernberg, Walteri Hosia, Andreas Jonsson, Elo Eriste, Hiie Allik** & **Xiao-qiu Wu,** for pleasant company either here or while in conferences away from Stockholm and for fruitful discussions. **Irene Byman** for patience to bearing my inquisitiveness to run Q-TOF at odd time.

The members of the Swedish-Bhutan society who lightened the burden of home-sick.

My Tibetan friends, **Tashi** and **Norzeen Wangyal,** for homely atmosphere and their elderly advice. **Prena** for delicious meals in my rare visits.

My Irish friends here: **Steve, Helen, Jim, Maura, Charlotte, Paula Cunnea, Fergal O'Farrell (& Jill), Oonagh Shannon (& Daniel), Máire Quigley Padraig Darcy, Gail Mullins, Jason,** & others for friendship and for different social events over the past years in Stockholm.

Football friends from Novum & Chinese friends at badminton evenings for keeping me fit throughout the years.

People outside Sweden has also helped in one way or the other, particularly:
From Ireland, my "Coady family" in Dublin: **Fergus, Carmel, Ronan & Kathleen (+Aisling), Dara & Joe, Aoife & Rachel**, and **Neil** for their love and the moral support throughout. I have never met anyone before so caring as them.
Irish friends: **David Vaughan, Carol Mulready, Niall Maguire, Damien Connolly, Padraig O'Shea** & others for their friendship, travel fun and for their welcome on every visit that I made to Ireland.

Back in my home country, Bhutan, my childhood friend, **Sonam Tashi**, for standing by me at good & bad times and for updating me the "real" home-news. You'll be missed dearly!
Finally my beloved family, my mother **Mrs Hari Maya** for her unmatched care and for the early morning phone-calls wondering why I am not in "school" yet!, sister **Aiti Maya** and her family, brother **Dipak**, brother **Bhim Raj** and his family for their unconditional love, understanding and optimism.

I am grateful to **Nordic Micromass** for access to Quattro Ultima & Micro instruments. This work was supported by grants from the Swedish Medical Research Council grant no. 03X-12551, Karolinska Institutet, and Stiftelsen Lars Hiertas Minne.

REFERENCES

1. Ranney HM. Hemoglobin: a historical perspective. *In* Steinberg MH, Forget BG, Higgs DR, Nagel RL eds. Disorders of Hemoglobin. Cambridge University Press 2001:1-24.
2. Perutz MF, Wilkinson AJ, Paoli M, Dodson GG. The stereochemical mechanism of the cooperative effects in hemoglobin revisited. Annu Rev Biophys Biomol Struct 1998;27:1-34.
3. Pauling L, Itano HA, Singer SJ, Wells IC. Sickle cell anemia, a molecular disease. Science 1949;110:543-8.
4. Weatherall DJ, Clegg JB. Genetic disorders of hemoglobin. Semin Hematol 1999;36:24-37.
5. Mentzer WC, Kan YW. Prospects for research in hematologic disorders: sickle cell disease and thalassemia. JAMA 2001;285:640-2.
6. Jia L, Bonaventura C, Bonaventura J, Stamler JS. S-nitrosohaemoglobin: a dynamic activity of blood involved in vascular control. Nature 1996;380:221-6.
7. Eaton WA, Henry ER, Hofrichter J, Mozzarelli A. Is cooperative oxygen binding by hemoglobin really understood? Nat Struct Biol 1999;6:351-8.
8. Persons DA, Nienhius AW. Gene therapy for the hemoglobin disorders: past, present, and future. Proc Natl Acad Sci USA 2000;97:5022-24.
9. Dibenedetto SP, Russo Mancuso G, Di Cataldo A, Schiliro G. Non inherited hemoglobin anomalies. Haemtologica 1991;76:414-20.
10. Charache S, Terrin ML, Moore RD, Dover GJ, Barton FB, Eckert SV, et al. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. N Engl J Med 1995;332:1317-22.
11. Craver RD, Abermanis JG, Warrier RP, Ode DL, Hempe JM. Hemoglobin A_2 levels in healthy persons, sickle cell disease, sickle cell trait, and β-thalassemia by capillary isoelectric focusing. Am J Clin Pathol 1997;107:88-91.
12. Kobold U, Jeppsson J-O, Dülffer T, Finke A, Hoelzel W, Miedema K. Candidate reference methods for HbA_{1c} based on peptide mapping. Clin Chem 1997;43:1944-51.
13. Flückiger R, Harmon W, Meier W, Loo S, Gabbay KH. Hemoglobin carbamylation in uremia. New Engl J Med 1981;304:823-7.
14. Niwa T, Naito C, Mawjood AHM, Imai K. Increased glutathionyl hemoglobin in diabetes mellitus and hyperlipidemia demonstrated by liquid chromatography/electrospray ionization mass spectrometry. Clin Chem 2000;46:82-8.
15. Boissel JP, Kasper TJ, Bunn HF. Cotranslational amino-terminal processing of cytosolic proteins. Cell-free expression of site-directed mutants of human hemoglobin. J Biol Chem 1988;263:8443-9.
16. Feng DF, Cho G, Doolittle RF. Determining divergence times with a protein clock: update and reevaluation. Proc Natl Acad Sci USA 1997;94:13028-33.
17. Orkin SH, Goff SC. The duplicated human α-globin genes: their relative expression as measured by RNA analysis. Cell 1981;24:345-51.

18. Molchanova TP, Pobedimskaya DD, Huisman THJ. The differences in quantities of α2- and α1-globin gene variants in heterozygotes. Br J Haematol 1994;88:300-6.
19. Miller JL. Hemoglobin switching and modulation: genes, cells, and signals. Curr Opin Hematol 2002;9:87-92.
20. Forget BG. Molecular basis of hereditary persistence of fetal hemoglobin. Ann N Y Acad Sci 1998;850:38-44.
21. Bunn HF, McDonald MJ. Electrostatic interactions in the assembly of haemoglobin. Nature 1983;306-498-500.
22. Mrabet NT, McDonald MJ, Turci S, Sarkar R, Szabo A, Bunn HF. Electrostatic attraction governs the dimer assembly of human hemoglobin. J Biol Chem 1986;261:5222-8.
23. Adachi K, Yamaguchi T, Pang J, Surrey S. Effects of increased anionic charge in the β -globin chain on assembly of hemoglobin in vitro. Blood 1998;91:1438-45.
24. Angatiniotis M, Modell B. Global epidemiology of hemoglobin disorders. Ann N Y Acad Sci 1998;850:251-69.
25. Weatherall DJ, Clegg JB. Genetic variability in response to infection: malaria and after. Genes Immun 2002;3:331-7.
26. Weatherall DJ. Phenotype-genotype relationships in monogenic disease: lessons from the thalassemias. Nature 2001;2:245-55.
27. Globin gene server. http://globin.cse.psu.edu (Accessed February 2003).
28. Bunn HF, Forget BG. Human hemoglobin variants. *In* Bunn HF, Forget BG, eds. Hemoglobin: molecular, genetic and clinical aspects. Philadelphia: WB Saunders Co 1986:381-452.
29. Landin B, Berglund S, Wallman K. Two different mutations in codon 97 of the β-globin gene cause Hb Malmö in Sweden. Am J Hematol 1996;51:32-6.
30. Landin B, Alvelius G, Rai DK, Elinder G. Compound heterozygosity for Hb Tacoma [β30(B12)Arg→Ser] and β^{+}-thalassemia. Hemoglobin 2000;24:253-7.
31. Metzenberg AB, Wurzer G, Huisman THJ, Smithies O. Homology requirements for unequal crossing over in humans. Genetics 1991;128:143-61.
32. Cohen-Solal M, Blouquit Y, Garel MC, Thillet J, Gaillard L, Creyssel R et al. Haemoglobin Lyon (β17-18 (A 14-15) Lys-Val → 0) determination by sequenator analysis. Biochim Biophys Acta 1974;351:306-16.
33. Huisman THJ. Combinations of β chain abnormal hemoglobins with each other or with β-thalassemia determinants with known mutations: influence on phenotype. Clin Chem 1997;43:1850-6.
34. Steinberg MH. Compound heterozygous and other sickle hemoglobinopathies. *In* Steinberg MH, Forget BG, Higgs DR, Nagel RL eds. Disorders of Hemoglobin. Cambridge University Press 2001:786-810.
35. Thein SL. Is it dominantly inherited β thalassemia or just a β-chain variant that is highly unstable? Br J Haematol 1999;107:12-21.

36. Wajcman H, Galactéros F. Abnormal hemoglobins with high oxygen affinity and erythrocytosis. Hematol Cell Ther 1996;38:305-12.
37. Nakanishi T, Miyazaki A, Shimizu A, Yamaguchi A, Nishimura S. Assessment of the effect of hemoglobin variants on routine HbA1c measurements by electrospray ionization mass spectrometry. Clin Chim Acta 2002;323:89-101.
38. Bry L, Chen PC, Sacks DB. Effects of hemoglobin variants and chemically modified derivatives on assays for glycohemoglobin. Clin Chem 2001;47:153-63.
39. Rai DK, Landin B, Griffiths WJ, Alvelius G. Identification of N-terminal acetylation in Hb Raleigh (β1Val→Ac-Ala) by electrospray tandem mass spectrometry. Rapid Commun Mass Spectrom 2002;16:1793-6.
40. Paleari R, Paglietti E, Mosca A, Mortarino M, Maccioni L, Satta S, et al. Posttranslational deamidation of proteins: the case of hemoglobin J Sardegna [α50(CD8)His→Asn→Asp]. Clin Chem 1999;45:21-8.
41. Rees DC, Rochette J, Schofield C, Green B, Morris M, Parker NE, et al. A novel silent posttranslational mechanism converts methionine to aspartate in hemoglobin Bristol (β67[E11] Val-Met→Asp). Blood 1996;88:341-8.
42. Brennan SO, Shaw J, Allen J, George PM. β141 Leu is not deleted in the unstable haemoglobin Atlanta-Coventry but is replaced by a novel amino acid of mass 129 daltons. Br J Haematol 1992;81:99-103.
43. Clarke GM, Higgins TN. Laboratory investigation of hemoglobinopathies and thalassemias: review and update. Clin Chem 2000;46:1284-90.
44. Guideline. The laboratory diagnosis of haemoglobinopathies. Br J Haematol 1998;101:783-92.
45. Papadea C, Cate JC. Identification and quantification of hemoglobins A, F, S, and C by automated chromatography. Clin Chem 1996; 42:57-63.
46. Righetti PG, Bossi A. Isoelectric focusing in immobilized pH gradients. Recent analytical and preparative developments. Anal Biochem 1997;247:1-10.
47. Basset P, Beuzard Y, Garel MC, Rosa J. Isoelectric focusing of human hemoglobin: Its application to screening, to the characterization of 70 variants, and to the study of modified fractions of normal hemoglobins. Blood 1978;51:971-82.
48. Campbell M, Henthorn JS, Davies SC. Evaluation of cation-exchange HPLC compared with isoelectric focusing for neonatal hemoglobinopathy screening. Clin Chem 1999;45:969-75.
49. Hofstadler SA, Swanek FD, Gale DC, Ewing AG, Smith RD. Capillary electrophoresis-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for direct analysis of cellular proteins. Anal Chem 1995;67:1477-80.
50. Li L, Golding RE, Whittal RM. Analysis of single mammalian cell lysates by mass spectrometry. J Am Chem Soc 1996;118:11662-3.
51. Mario N, Baudin B, Aussel C, Giboudeau J. Capillary isoelectric focusing and high-performance cation-exchange chromatography compared for qualitative and quantitative analysis of hemoglobin variants. Clin Chem 1997;43:2137-42.

52. Frantzen F. Chromatographic and electrophoretic methods for modified hemoglobins. J Chromatogr B 1997;699:269-86.
53. Fucharoen S, Winichagoon P, Wisedpanichkij R, Sae-Ngow B, Sriphanich R, Oncoung W, et al. Prenatal and post natal diagnoses of thalassemias and hemoglobinopathies by HPLC. Clin Chem 1998;44:740-8.
54. Wajcman H, Riou J, Yapo AP. Globin chain analysis by reversed phase high performance liquid chromatography: recent developments. Hemoglobin 2003;26:271-84.
55. Kwant G, Oeseburg B, Zijlstra WG. Reliability of the determination of whole-blood oxygen affinity by means of blood-gas analyzers and multi-wavelength oximeters. Clin Chem 1989;35:773-7.
56. Sutcharitchan P, Embury SH. Advances in molecular diagnosis of inherited hemoglobin disorders. Curr Opin Hematol 1996;3:131-8.
57. Arcasoy MO, Gallagher PG. Molecular diagnosis of hemoglobinopathies and other red blood cell disorders. Semin Hematol 1999;36:328-39.
58. Lindeman R, Wallace R, Volpato F, Hu SP, Trent RJ. Utility of the polymerase chain reaction (PCR) for prenatal diagnosis of genetic disease. Pathology 1991;23:158-63.
59. Oron-Karni V, Filon D, Oppenheim A, Rund D. Rapid detection of the common Mediterranean α-globin deletions/rearrangements using PCR. Am J Hematol 1998;58:306-10.
60. Yates III, JR. Mass spectrometry from genomics to proteomics. Trends Genet 2000;16:5-8.
61. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 2001;70:437-73.
62. Lee KH. Proteomics: a technology-driven and technology-limited discovery science. Trends Biotechnol 2001;19:217-22.
63. Russell DH, Edmondson RD. High-resolution mass spectrometry and accurate mass measurements with emphasis on the characterization of peptides and proteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Mass Spectrom 1997;32:263-76.
64. Hendrickson CL, Emmett MR. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Annu Rev Phys Chem 1999;50:517-36.
65. Reid GE, McLuckey SA. "Top down" protein characterization via tandem mass spectrometry. J Mass Spectrom 2002;37:663-75.
66. Tyler AN, Clayton E, Green BN. Exact mass measurement of polar organic molecules at low resolution using electrospray ionization and a quadrupole mass spectrometer. Anal Chem 1996;68:3561-9.
67. Yergey J, Heller D, Hansen G, Cotter RJ, Fenselau C. Isotopic distributions in mass spectra of large molecules. Anal Chem 1983;55:353-6.
68. Beavis RC. Chemical mass of carbon in proteins. Anal Chem 1993;65:496-7.
69. Karas M, Glückmann M, Schäfer J. Ionization in matrix-assisted laser desorption/ionization: singly charged molecules are the lucky survivors. J Mass Spectrom. 2000;35:1-12.
70. Larsen MR, Roepstorff P. Mass spectrometric identification of proteins and characterization of their post-translational modifications in proteome analysis. Fresenius J Anal Chem 2000;366:677-90.

71. Griffiths WJ, Jonsson AP, Liu S, Rai DK, Wang Y. Electrospray and tandem mass spectrometry in biochemistry. Biochem J 2001;355:345-61.
72. Kebarle P, Tang L. From ions in solution to ions in the gas phase: the mechanism of electrospray mass spectrometry. Anal Chem 1993;65:972A-86A.
73. Gaskell SJ. Electrospray: principles and practice. J Mass Spectrom 1997;32:677-88.
74. Kebarle P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. J. Mass Spectrom 2000;35:804-17.
75. Griffiths WJ. Nanospray mass spectrometry in protein and peptide chemistry. *In* Jólles P, Jörnvall H eds. Proteomics in functional genomics. Switzerland: Birkhäuser Verlag Basel 2000:69-79.
76. Karas M, Bahr U, Dülcks T. Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. Fresenius J Anal Chem 2000;366:669-76.
77. Ferrige AG, Seddon MJ, Green BN, Jarvis SA, Skilling J. Disentangling electrospray spectra with maximum entropy. Rapid Commun Mass Spectrom 1992;6:707-11.
78. Hsieh YLF, Wang HQ, Elicone C, Mark J, Martin SA, Regnier F. Automated analytical system for the examination of protein primary structure. Anal Chem 1996;68:455-62.
79. Gatlin CL, Eng JK, Cross ST, Detter JC, Yates III JR. Automated identification of amino acid sequence variations in proteins by HPLC/microspray tandem mass spectrometry. Anal Chem 2000;72:757-63.
80. De Hoffmann E. Tandem mass spectrometry: a primer. J. Mass Spectrom 1996;31:129-37.
81. Roepstorff P. Fohlman J. Proposal for a common nomenclature of sequence ions in mass spectra of peptides. Biomed Mass Spectrom 1984;11:601.
82. Biemann K. Contributions of mass spectrometry to peptide and protein structure. Biomed Environ Mass Spectrom 1988;16:99-111.
83. Tang X-J, Boyd RK. An investigation of fragmentation mechanisms of doubly protonated tryptic peptides. Rapid Commun Mass Spectrom 1992;6:651-7.
84. Polce JM, Ren D, Wesdemiotis C. Dissociation of the peptide bond in protonated peptides. J Mass Spectrom 2000;35:1391-8.
85. Ambihapathy K, Yalcin T, Leung H-W, Harrison AG. Pathways to immonium ions in the fragmentation of protonated peptides. J Mass Spectrom 1997;32:209-15.
86. Tang X-J, Thibault P, Boyd RK. Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation. Anal Chem 1993;65:2824-34.
87. Dongré AR, Jones JL, Somogyi A, Wysocki VH. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. J Am Chem Soc 1996;118:8365-74.
88. Wada Y, Hayashi A, Fujita T, Matsuo T, Katakuse I, Matsuda H. Structural analysis of human hemoglobin variants with field desorption mass spectrometry. Biochim Biophys Acta 1981;667:233-41.
89. Jensen ON, Højrup P, Roepstorff P. Plasma desorption mass spectrometry as a tool in characterization

of abnormal proteins. Application to variant human hemoglobins. Anal Biochem 1991;199:175-83.

90. Pucci P, Carestia C, Fioretti G, Mastrobuoni AM, Pagano L. Protein fingerprint by fast atom bombardment mass spectrometry: characterization of normal and variant haemoglobins. Biochem Biophys Res Commun 1985;130:84-90.
91. Ferranti P, Malorni A, Pucci P, Fanali S, Nardi A, Ossicini L. Capillary zone electrophoresis and mass spectrometry for the characterization of genetic variants of human hemoglobin. Anal Biochem 1991;194:1-8.
92. Shackleton CHL, Falick AM, Green BN, Witkowska HE. Electrospray mass spectrometry in the clinical diagnosis of variant hemoglobins. J Chromatogr 1991;562:175-90.
93. Shackleton CHL, Witkowska HE. Mass spectrometry in the characterization of variant hemoglobins. In Desiderio DM, Ed. Mass Spectrometry: Clinical and Biomedical Applications. New York: Plenium 1994; Vol.2:135-99.
94. Wild BJ, Green BN, Cooper EK, Lalloz MRA, Erten S, Stephens AD, Layton DM. Rapid identification of hemoglobin variants by electrospray ionization mass spectrometry. Blood Cells Mol Dis 2001;27:691-704.
95. Shimizu A, Nakanishi T, Kishikawa M, Miyazaki A. Detection and identification of protein variants and adducts in blood and tissues: an application of soft ionization mass spectrometry to clinical diagnosis. J Chromatogr B 2002;776:15-30.
96. Houston CT, Reilly JP. Rapid analysis of hemoglobin from whole human blood by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 1997;11:1435-9.
97. Spengler B, Kirsch D, Kaufmann R, Jaeger E. Peptide sequencing by matrix-assisted laser-desorption mass spectrometry. Rapid Commun Mass Spectrom 1992;6:105-8.
98. Wada Y, Yamamoto M. Detection of single-nucleotide mutations including substitutions and deletions by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 1997;11:1657-60.
99. Witkowska HE, Lubin BH, Beuzard Y, Baruchel S, Esseltine DW, Vichinsky EP, et al. Sickle cell disease in a patient with sickle cell trait and compound heterozygosity for hemoglobin S and hemoglobin Quebec-Chori. New Engl J Med 1991;325:1150-54.
100. Roberts NB, Amara AB, Morris M, Green BN. Long-term evaluation of electrospray ionization mass spectrometric analysis of glycated hemoglobin. Clin Chem 2001;47:316-21.
101. Rai DK, Alvelius G, Landin B. Identification of Hb Kenya by electrospray mass spectrometry. Hemoglobin 2002;26:71-5.
102. Nakanishi T, Kishikawa M, Shimizu A, Hayashi A, Inoue F. Assignment of the ions in the electrospray ionization mass spectra of the tryptic digest of the non-derivatized globin, covering the whole sequence of α- and β-chains: a rapid diagnosis for haemoglobinopathy. J Mass Spectrom 1995;30:1663-70.

103. Kiernan UA, Black JA, Williams P, Nelson P. High-throughput analysis of hemoglobin from neonates using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Clin Chem 2002;48:947-9.

104. Li MX, Wu J-T, Liu L, Lubman DM. The use of on-line capillary electrophoresis/electrospray ionization with detection via an ion-trap storage/reflectron time-of-flight mass spectrometer for rapid mutation-site analysis of hemoglobin variants. Rapid Commun Mass Spectrom 1997;11:99-108.

105. Nakanishi T, Miyazaki A, Kishikawa M, Shimizu A, Yonezawa T. Eectrospray ionization-tandem mass spectrometry analysis of peptides derived by enzymatic digestion of oxidized globin subunits: an improved method to determine amino acid substitution in the hemoglobin "core". J Am Soc Mass Spectrom 1996;7:1040-9.

106. Yang L, Lee CS, Hofstadler SA, Pasa-Tolic L, Smith RD. Capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for protein characterization. Anal Chem 1998;70:3235-41.

107. Tang Q, Harrata AK, Lee CS. High resolution capillary isoelectric focusing-electrospray ionization mass spectrometry for hemoglobin variants analysis. Anal Chem 1996;68:2482-7.

108. Ofori-Acquah SF, Green BN, Davies SC, Nicolaides KH, Serjeant GR, Layton DM. Mass spectral analysis of asymmetric hemoglobin hybrids: demonstration of Hb FS ($\alpha_2\gamma\beta^S$) in sickle cell disease. Anal Biochem 2001;298:76-82.

109. Baudin-Creuza V, Fablet C, Zal F, Green BN, Promé D, Marden MC, Pagnier J, Wajcman H. Hemoglobin Porto Alegre forms a tetramer of tetramers superstructure. Protein Sci 2002;11:129-36.

110. Rai DK, Landin B, Alvelius G, Griffiths WJ. Gel-elution: a method to enrich and isolate variant hemoglobins for ES mass spectrometry. Proc 50th ASMS conference on mass spectrometry and allied topics, Orlando, FL, 2002.

111. Light-Wahl KJ, Loo JA, Edmonds CG, Smith RD, Witkowska HE, Shackleton CHL, Wu C-SC. Collisionally activated dissociation and tandem mass spectrometry of intact hemoglobin β-chain variant proteins with electrospray ionization. Biol Mass Spectrom 1992;22:112-20.

112. Bakhtiar R, Wu Q, Hofstadler SA, Smith RD. Charge state specific facile gas-phase cleavage of Asp 75-Met 76 peptide bond in the α-chain of human apohemoglobin probed by electrospray ionization mass spectrometry. Biol Mass Spectrom 1994:23:707-10.

113. Witkowska HE, Green BN, Morris M, Shackleton CHL. Intact protein electrospray ionization tandem mass spectrometry can be the sole technique used for confirming the structure of a variant hemoglobin. J Mass Spectrom Rapid Commun Mass Spectrom 1995;Spec No:S111-5.

114. Schaaff TG, Cargile BJ, Stephenson Jr JL, McLuckey SA. Ion trap collisional activation of the $(M + 2H)^{2+}$ - $(M + 17H)^{17+}$ ions of human hemoglobin β-chain. Anal Chem 2000;72:899-907.

115. Blackwell RQ, Liu CS, Wang CL. Hemoglobin Ta-Li: β83 Gly→Cys. Biochim Biophys Acta 1971;243: 467-74.