Both normal albumin (Al A) and genetically modified forms were isolated from six heterozygous subjects. Albumins from each individual were analyzed by electrospray ionization mass spectrometry (ESI MS), and the mass was compared with that predicted from the protein sequence. In all cases, the Al A was heterogeneous, with components of mass (± SE) 66 463 ± 4, 66 586 ± 3, and 66 718 ± 5 Da. Each genetic variant showed similar heterogeneity. The mass increase in Al Casebrook (2214 Da) was very close to that predicted (2205 Da) from protein and carbohydrate sequence analysis, whereas the increase in Al Redhill (2378 Da) was close to that expected (2392 Da) for an Arg-albumin with a disialylated N-linked biantennary oligosaccharide and an Ala→Thr mutation. The circulating proalbumins, Christchurch and Blenheim, had mass increases of 748 and 756 Da, respectively, over Al A; in excellent agreement with theoretical values of 744 and 756. Clear shifts in mass were also detected for the point substitutions 177Cys→Phe (44 Da), 1Asp→Val (20 Da), and Arg-albumin (160 Da).

Human serum albumin, the most abundant plasma protein, consists of a single 585-residue nonglycosylated polypeptide chain that is synthesized in the liver as preproalbumin (1). The prepeptide is cotranslationally cleaved in the endoplasmic reticulum, and the propeptide (RGVFRR→) is removed in Golgi/vesicles before secretion from the hepatocyte (2). By virtue of its concentration, albumin regulates plasma osmotic pressure; however, it functions primarily as a transport protein and reversibly binds a wide array of ligands, such as fatty acids, steroids, bilirubin, tryptophan, and copper (1, 3).

During its 18-day circulatory half-life, the mature molecule is subjected to a barrage of additional covalent modifications, including proteolytic cleavage, oxidation, and deamidation, and the covalent attachment of a variety of endogenous (e.g., glucose, cysteine, and nitric oxide) and exogenous (e.g., acetylsalicylic acid and penicillin) compounds (1, 4). In addition to these aging modifications, genetic variation can superimpose an additional level of heterogeneity on the population of molecules.

Albumin is probably the most commonly measured serum protein, and genetic variation, or bisalbuminemia, is usually detected by electrophoresis of plasma collected for diagnostic reasons. Bisalbuminemia, however, may also be acquired as a result of proteolytic cleavage associated with pancreatitis or increased bilirubin binding in severe liver disease. Some 50 different genetic variants have been characterized by protein and/or DNA sequence analysis (5), and these have provided unique insights into the intracellular processing of albumin and its ligand binding sites. Indeed, all the variants examined here (Table 1) have defects that affect their hepatic processing. Both proalbumin Christchurch (-1R→Q) and proalbumin Blenheim (1D→V) have mutations that allow the proprotein to evade convertase cleavage in the Golgi (6), whereas albumin Kaikoura (-2R→C) has a cryptic signal peptidase site that favors cleavage after residue -2 instead of after residue -7, leading to the export of Arg-albumin into the plasma (7). Albumins Casebrook (494D→N) and Redhill (320A→T) have mutations that introduce Asn-Glu-Thr and Asn-Tyr-Thr sequences that act as acceptors for N-linked oligosaccharides (8, 9). In addition, albumin Redhill has a -1 Arg residue, the result of an additional -2R→C mutation. The final variant examined, albumin Hawkes Bay (177C→F), has a defect that abolishes one of its disulfide bonds, greatly decreasing its plasma concentration (10).

Electrospray ionization mass spectrometry (ESI MS) is being applied increasingly to the molecular analysis of complex proteins. Impressive results have been obtained for human hemoglobin which, because of the comparatively small size of its component chains (16 126 and 15 876 Da) and lack of posttranslational modifications, produces clean spectra with high ion currents. Indeed, a number of genetic variants of hemoglobin have been detected and characterized using ESI MS (11, 12).
The traditional methods of mutation detection rely on identifying a functional or physical change in a protein; electrophoresis, isoelectric focusing, and reversed-phase HPLC permit the segregation of gene products on the basis of charge, conformation, and polarity. The accurate determination of protein molecular mass by ESI MS provides not only a novel means of mutation analysis but may indicate the type of change involved (e.g., amino acid substitution, truncation/elongation, or possible carbohydrate attachment). Where a number of polypeptide chains make up a particular protein, such as in the case of hemoglobin (13) or fibrinogen (14), then MS can often identify which chain has the molecular lesion (13, 14). Protein fragmentation followed by mass analysis of, for example, tryptic or CNBr peptides, can then be used to rapidly pinpoint a mutation site, and this can be zeroed in on by peptide sequencing. Alternately, targeted regions of a gene could then be amplified by polymerase chain reaction and subjected to direct DNA sequence analysis (15).

The theoretical average isotopic mass of disulfide-linked albumin A is 66 438 Da. Using ESI MS, Loo et al. (16) noted the heterogeneity of albumin from a number of different species and reported the molecular mass of human as 66 605 Da. Subsequently, Yang et al. (17) reported a value of 66 745 Da for human serum albumin, whereas a value of 66 440 Da was determined for (a more homogeneous) recombinant human albumin derived from yeast expression (18).

In this investigation, we isolated both normal and variant forms of albumin from the plasma of six heterozygous subjects with different genetic mutations and compared measured mass changes with those expected on the basis of the aberrant protein sequence. This confirms that ESI MS can be used to detect and characterize both genetic and circulatory modifications of albumin.

**Materials and Methods**

Electrophoretic variants of albumin were purified from the plasma of six heterozygous subjects who gave informed consent and under protocols approved by Christchurch Hospital. Briefly, 1–10 mL of plasma was dialyzed against 16 mmol/L sodium acetate buffer, pH 5.2, and applied to a DEAE-Sephadex A-50 column (1.6 × 16 cm) equilibrated in the same buffer (6). The column was eluted with a linear gradient of 300 mL of this buffer and 300 mL of 16 mmol/L sodium acetate buffer, pH 4.3. A flow rate of 1 mL/min was used, and the column effluent was monitored at 280 nm. The normal and variant albumin peaks were collected, concentrated to 2–3 mL in a 50-mL Amicon cell, diluted to 15 mL with water, and reconcentrated. This latter dialysis step was repeated an additional two times before the protein was freeze-dried and stored at −20 °C. Some albumins had been stored in this way for up to 5 years. Analysis by agarose gel electrophoresis demonstrated that each isolated albumin component was ≥95% pure. All the mutant albumins used in this investigation had been characterized earlier by protein and/or DNA sequence analysis (6–10) and, in one case, carbohydrate analysis (19).

Both the normal and variant albumin components from each plasma were redissolved at between 1.5 and 2.5 g/L in 500 mL/L acetonitrile–1 mL/L formic acid and analyzed by ESI MS on a VG Platform quadrupole instrument (Micromass) operating in positive ion mode (20). Ten to twenty microliters of each protein was injected into the source at a flow rate of 5 μL/min. The probe was charged at +3000 V, and the source maintained at 60 °C. The mass range 1000–2000 m/z was scanned every 2.5 s, and a cone voltage ramp of 40–60 V was applied over this range. Between 60 and 100 scans were collected and averaged for each analysis, using high and low and high mass resolution parameters of 14. Calibration was made over this same m/z range (voltage ramp, 50–80 volts), using the charge series generated by human α-globin. For this a 100 g/L red cell lysate calibrator was diluted to 0.5 g/L in 1 mL/L HCOOH–500 mL/L acetonitrile (11), and ions from m/z 1009.43 to 1891.80 were used for the calibration. Data were acquired and processed using Mass-Lynx software and transformed onto a true molecular mass scale using maximum entropy (Max-Ent) software supplied with the instrument. Max-Ent deconvolution of the hemoglobin calibrator yielded molecular masses of 15 126 and 15 867 Da for the α and β chains, respectively. Molecular masses are based on the following atomic weights: C, 12.011; H, 1.00794; N, 14.00674; O, 15.9994; P, 30.97376; and S, 32.006.

For S-carboxymethylation, albumin (10 g/L) was first reduced for 16 h with 0.1 mol/L dithiothreitol in 0.1 mol/L Tris-HCl, pH 8.0, 8 mol/L urea. After reduction, the albumin was reacted for 45 min with a 2.5 mol/mol excess of iodoacetic acid (pH 8.0) over total thiol groups and then dialyzed against water and freeze-dried.

**Results**

The essential requirement of any MS analysis is that the analyte exists as a gas phase ion. In protein solutions, ionization is easily achieved by raising or lowering the pH, and vaporization is effected by rapid (heat- and nebulization-assisted) evaporation as the protein solution enters the atmospheric pressure ion source. In electro-spray spectra there is typically little ion fragmentation, and what is seen is a series of molecular ions, each differing from its neighbor by one charge.

With native albumin in 500 mL/L acetonitrile–1 mL/L formic acid the charge envelope usually extended from +60 to +32. Portions of the ion spectra of normal albumin (Al A) from two of the subjects is shown in Fig. 1. The ions depicted extend from [M+51H] to [M+41H] and have a charge to mass ratio extending from 1306.0 to 1624.6 m/z. What is immediately apparent is that each m/z peak is reproducibly heterogeneous and is composed of at least three or four different components. As in the two cases depicted, the Al A gene product from the other four subjects all showed very similar heterogeneity, although the ratio of the individual components varied. All variant
ion spectra displayed similar heterogeneity, although their precise \( m/z \) positions varied in accordance with their specific mutations.

Ion spectra spanning an \( m/z \) range of 1100–1900 were transformed onto a true mass scale using Max-Ent software; two of the transformed Al A spectra are shown (Fig. 2, a and f). Data from these and the Al A components from the other four subjects are summarized in Table 1. The heterogeneity that was detected in the raw data is reflected in these results. The three principal components had mean molecular masses (±SE) of 66 463 ± 4; 66 586 ± 3; and 66 718 ± 5 Da, respectively. The peak at 66 463 Da is in good agreement with the predicted mass (66 438 Da) of Al A. The other two components are heavier by 123 ± 2 and 255 ± 6 Da, respectively, probably reflecting the attachment of cysteine (119 Da) and palmitic acid (256 Da). Similar mass differences were apparent between the components in the spectra of isolated genetic variants (Fig. 2, b–e, g, and h) where mean mass increases of 120 ± 4 and 254 ± 6 Da were observed. The presence of these different isoforms within the purified albumin peaks did not, however, compromise the ability to discern mass changes associated with specific mutations. This heterogeneity was greatly diminished after reduction and S-carboxymethylation, where transformed spectra indicated a mass of 68 524 Da, in good agreement with a predicted value of 68 504 Da (Fig. 3).

To determine the effect of mutation on molecular mass, each individual variant was analyzed immediately after its corresponding Al A gene product (i.e., within the same batch). To ascertain the error associated with these comparisons, nine repeat injections were made of Al A from a single individual. In this case the dominant component had a molecular mass of 66 448 Da with an SE of ±2 Da. All sequence changes associated with the variants, together with predicted and calculated mass changes, are summarized in Table 1. For simplicity the putative Al-Cys peaks are used as the basis of comparison in the ensuing discussion. Predictably the 1Asp→Val substitution in albumin Blenheim would cause a 16-Da decrease in the mass in the mature protein, whereas the expected increase in the resulting proalbumin Blenheim would be 756 Da. These values are in excellent agreement with the measured changes of −20 and +756 Da, respectively, for these two species. The 756-Da change is clearly indicative of a major change in protein structure. However, given the within-batch SE of ±3 Da associated with this difference, the 20-Da decrease can only suggest the possibility of a point mutation. The mutation in this case, however, can be readily confirmed electrophoretically because of the charge change associated with the Asp→Val substitution. The other two N-terminal extended forms, Arg-albumin and proalbumin Christchurch, had measured increases of 160 and 748 Da, respectively, over their normal gene
products, compared with expected increases of 156 and 744 Da.

The introduction of new Asn-Glu-Thr and Asn-Tyr-Thr sequences in albumins Casebrook and Redhill cause the incorporation of an N-linked oligosaccharide side chain into each of these proteins. The structure of the side chain has been determined as -NAcGlc-NAcGlc-Man-(Man-NAcGlc-Gal-NAcNeu)_2 for albumin Casebrook (19); however, the structure of the side chain has not been investigated in the case of albumin Redhill. With the 494Asp→Asn mutation (−1 Da) and oligosaccharide (2206 Da), the expected mass increase in albumin Casebrook is 2205 Da, in excellent agreement with a measured increase of 2214 Da. The observed increase in mass in albumin Redhill was 2378 Da, whereas the increase associated with the -1 Arg (156 Da), the 318Ala→Thr substi-

Fig. 2. Transformed ESI spectra of albumin variants. 
(a–c) The three albumin components from Blenheim plasma, Al A, albumin Blenheim, and proalbumin Blenheim, respectively; (d) purified proalbumin Christchurch; (e) purified Arg-albumin; (f and g) Al A and albumin Casebrook from Casebrook plasma; (h) albumin Redhill. Subtracted raw data were transformed using MaxEnt software.
tution (30 Da), and the attachment of a similar oligosaccharide (2206 Da) is 2392 Da. The closeness of these values indicates for the first time that albumin Redhill has a fully sialylated biantennary oligosaccharide side chain.

Albumin Hawkes Bay is a low concentration variant with aberrant disulfide bridging resulting from a substitution of 177Cys→Phe. The direct mass increase of this mutation is 44 Da, with an additional increase of 1 Da arising from the formation of the unpaired −SH group; this value (45 Da) was again very close to the measured value of 44 Da.

### Discussion

In previous studies, researchers using ESI MS analysis of commercially available human serum albumin noted the heterogeneity of its ion spectra and reported transformed molecular masses of 66 605 (16) and 66 754 Da (17). It is clear from this investigation that this heterogeneity can be resolved, and individual isoforms with masses of 66 463 ± 8; 66 586 ± 7; and 66 718 ± 11 Da were identified for Al A. The lower value of 66 463 Da is very close to the theoretical value of 66 438 Da for virgin native albumin; indeed, Geisow et al. (18) determined a value of 66 440 Da for recombinant human albumin expressed in yeast and isolated directly from culture supernatants. Albumin, however, undergoes a barrage of modifications during its average 26-day voyage through the circulation. Undoubtedly some minor modified forms might have been purified out during ion-exchange chromatography, but three principal components were resolved.

The mean mass differences between resolved isoforms of Al A were 123 ± 4 and 255 ± 12 Da, and similar isoforms were present in the variants, with mean differ-

### Table 1. Molecular mass of genetic variants of serum albumin.a

<table>
<thead>
<tr>
<th>Albumin</th>
<th>Structure</th>
<th>Expected mass difference</th>
<th>Measured mass, Da</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al A</td>
<td></td>
<td></td>
<td>66 464 66 584 66 708</td>
</tr>
<tr>
<td>Al Blenheim</td>
<td>1D→V</td>
<td>−16</td>
<td>66 440 66 564 66 704</td>
</tr>
<tr>
<td>pAl Blenheim</td>
<td>RGVFRRV</td>
<td>+756</td>
<td>67 216 67 340 67 472</td>
</tr>
<tr>
<td>pAl Christchurch</td>
<td>RGVFRQ</td>
<td>+744</td>
<td>748 756 764</td>
</tr>
<tr>
<td>Al A</td>
<td></td>
<td></td>
<td>66 474 66 592 66 720</td>
</tr>
<tr>
<td>ArgAl</td>
<td>−1R</td>
<td>+156</td>
<td>66 608 66 734 66 866</td>
</tr>
<tr>
<td>Al A</td>
<td></td>
<td></td>
<td>66 458 66 588 66 734</td>
</tr>
<tr>
<td>Al Casebrook</td>
<td>494D→N, CHOc</td>
<td>+2205</td>
<td>68 672 68 802 68 945</td>
</tr>
<tr>
<td>Al Redhill</td>
<td>−1R, 320A→T, CHOc</td>
<td>+2392</td>
<td>68 844 68 962 69 088</td>
</tr>
<tr>
<td>Al A</td>
<td></td>
<td></td>
<td>66 460 66 584 66 720</td>
</tr>
<tr>
<td>Al Hawkes Bay</td>
<td>177C→F</td>
<td>+45</td>
<td>66 518 66 636 66 784</td>
</tr>
</tbody>
</table>

a The mass of normal and variant albumins were compared for each individual, and the expected and measured differences between the two are shown.

b Al, albumin; pAl, proalbumin.

c Observed mass difference between normal isoform and corresponding isoform of genetic variant.

d Carbohydrate attachment.
ences of 120 ± 10 and 255 ± 16 Da. The 123- and 120-Da differences most probably represent the covalent attachment of cysteine (119 Da) to the free thiol at position 34 of the normal and variant albumin, respectively, because it is well established that this thiol is blocked in >33% of the circulating molecules. In support of this, the area of this peak was seen to decreased substantially after incubation of the albumin for 4 h in 2 mol/L dithiothreitol (not shown), and the peak disappeared entirely after reduction and carboxymethylation (Fig. 3).

The species with an increased mass of 255 ± 14 Da most probably represents noncovalently bound palmitic acid (256 Da), some of which seems to remain bound despite the presence of 500 mL/L acetonitrile. Although covalent palmitylation is not excluded, the expected mass increase would be 238 and not 256 Da. Furthermore, if esterification was involved it would be expected to survive the effects of reduction in 8 mol/L urea and S-carboxymethylation. This was not the case; after this treatment the albumin was converted to a more homogeneous peak with no indication of a +255 Da signal.

Because in ESI MS there is a trade-off between resolution and sensitivity, it was important to use relatively high protein concentrations (1.5–2.5 g/L) and low flow rates (5 μL/min) so that between 60 and 100 m/z scans could be collected and averaged for each analysis. The sequential analysis of Al A and its companion variant also facilitated more accurate determination of mass differences between the two allelic products.

In the case of albumin Casebrook, the mass increase of 2214 Da confirms a fully sialylated biantennary oligosaccharide. The same 494Asp (2378 Da) indicates for the first time that it too has a fully sialylated biantennary oligosaccharide in addition to its two mutations.

The results here establish the utility of ESI MS in the direct analysis of underivatized genetic variants of serum albumin and other large proteins. The technique allows the rapid identification of genetic variation, posttranslational modifications, and circulatory changes.

This investigation was supported by the Canterbury Medical Research Foundation, Lottery Health, the Health Research Council of New Zealand, and the University of Otago.

References


