Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 9853–9857, November 1991
Biochemistry

Genetic variants of serum albumin in Americans and Japanese

(alloalbumins/proalbumins/bisalbuminemia/genetic polymorphism/point mutations)

JEANNE MADISON†, KUNIO ARAI‡, YASUSHI SAKAMOTO†, RONALD D. FELD§, ROBERT A. KYLE¶, SCOTT WATKINS†, EVELYN DAVIS†, YUH-ICHI MATSUDA‖, ITTA AMAKI‖, AND FRANK W. PUTNAM†,††

†Department of Biology, Indiana University, Bloomington, IN 47405; ‡Department of Chemistry, Josai University, Keyakidai, Sakado, Saitama 350-02, Japan; §Department of Pathology, The University of Iowa Hospitals and Clinics, University of Iowa, Iowa City, IA 52242; ¶Hematology and Internal Medicine, Mayo Clinic, Rochester, MN 55905; and ‖Tokyo Metropolitan Blood Center, Japanese Red Cross, 2-2-2 Komagome, Toshima-ku, Tokyo 170, Japan

Contributed by Frank W. Putnam, July 22, 1991

ABSTRACT **A collaborative search for albumin genetic variants (alloalbumins) was undertaken by cellulose acetate and agarose electrophoresis at pH 8.6 of the sera of patients at two major medical centers in the United States and of nearly 20,000 blood donors in Japan. Seventeen instances of alloalbuminemia were ascertained, and seven different alloalbumin types were characterized by structural study. Two previously unreported alloalbumin types were identified. In one type, which was present in a Caucasian family and designated Iowa City-1, aspartic acid at position 365 was replaced by valine (365 Asp → Val); this is the second reported mutation at this position. The other type present in a Japanese blood donor had the mutation 128 His → Arg. An unexpected finding was the presence in a single Japanese of a Naskapi-type alloalbumin (372 Lys → Glu), a variant that had previously been described only for certain Amerindian tribes in whom it occurs with a polymorphic frequency (>1%) and in Eti Turks. An arginyl-albumin (−1 Arg, 1 Asp → Val) occurred in an American family. The other alloalbumin types identified were proalbumins Lille and Christchurch and albumin B that have a cumulative frequency of about 1:3500 in Caucasians probably because of the hypermutability of CpG dinucleotides at the mutated sites. All of the variants characterized in this study are point mutants, and the sites are spread throughout the albumin gene. However, about one-fourth of all known albumin mutations are clustered in the sequence segment from position 354 through 382.**

Albumin genetic variants (alloalbumins) are rare and benign and thus are detected only by electrophoretic screening of blood proteins in systematic studies of population genetics (1–7) or during routine clinical electrophoresis of the serum of patients or blood donors (8–15). Of the nearly 4000 possible effective point mutations in the 585 codons of the albumin gene, about 800 would lead to a change in electrophoretic mobility. Yet, only about 35 different alloalbumins have been detected and characterized structurally. In continuation of our investigation of albumin as a model for study of neutral protein evolution (5–7, 9, 16–21), we report collaborative research on alloalbumins identified in two population groups: (*i*) by clinical electrophoresis at two major medical centers in the United States and (*ii*) by screening for alloalbumins in the sera of some 20,000 blood donors in Japan.

Seventeen instances of alloalbuminemia in unrelated individuals were detected, and seven structurally different alloalbumins were identified and characterized in this investigation. Two previously unreported alloalbumins were identified. One of these designated Iowa City-1, which was present in a Caucasian family, had the substitution 365 Asp → Val at the position where albumin Parklands has the mutation 365 Asp → His (12). The other, Komagome-2 (128 His → Arg), occurred in a single Japanese blood donor. Three other variant types identified were the rare but most often encountered alloalbumins in Caucasians (8, 10, 20)—namely, albumin B (570 Glu → Lys) and proalbumins Christchurch (−1 Arg → Gln) and Lille (−2 Arg → His). One instance of each of these proalbumins and four instances of albumin B were characterized structurally in the large American patient groups surveyed and also one instance of an arginyl-albumin (−1 Arg, 1 Asp → Val); all of the carriers were Caucasians. Unexpectedly, a Naskapi-type alloalbumin (372 Lys → Glu) was found in a single Japanese; this variant albumin had previously been reported only in Athabascan- and Algonkian-speaking Amerindian tribes in which it occurs with a polymorphic allele frequency—i.e., a frequency ≥1% and in Eti Turks (3, 17).

MATERIALS AND METHODS

Electrophoretic Screening. Detection of albumin variants in the three surveys was done by cellulose acetate and agarose electrophoresis at pH 8.6. This procedure was also used for comparison to reference alloalbumins with known substitutions and after limited tryptic digestion to identify possible proalbumins (16). The 11 bisalbuminemic sera from the Mayo Clinic were accumulated over a period of 17 years and were noted during routine clinical electrophoresis because of the slow (+2 mobility units) or fast (−2) mobility of the variant albumin, and no alloalbumins with a change of +1 or −1 mobility units are represented. The three alloalbumins in the Iowa City series were identified by inspection of electrophoretograms of sera of patients and staff over a period of about 4 years. The Japanese (Komagome) project consisted of electrophoretic screening for albumin variants in the sera of 18,874 blood donors at the Tokyo Metropolitan Blood Center of the Japanese Red Cross.

Structural Studies. Because of the number and variety of the specimens reported here and the fact that most of the methods have been published in detail (5–7, 16–21), only a summary of the methods is given for each type of alloalbumin. The general strategy for structural study was as follows: (*i*) purification by HPLC of the total albumin (normal albumin A and variant) by a two-step method involving anion-exchange HPLC and gel filtration as described elsewhere (18–21) or purification of the variant albumin in a single step by HPLC on a preparative DEAE 5 PW column (21.5 × 150 mm; Bio-Rad) with use of a 20 mM sodium acetate buffer with a pH gradient from pH 5.2 to 4.5 for slow variants and a pH gradient from pH 5.2 to 4.0 for fast variants at a flow rate of 5.0 ml/min for a period of 40 min; (*ii*) reduction and carboxymethylation of the purified albumin and cleavage with CNBr (18–21); (*iii*) analytical isoelectric focusing of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

††To whom reprint requests should be addressed.

CNBr digest to identify the CNBr fragment in which the substitution was localized (5, 16, 17); (*iv*) HPLC peptide mapping on a preparative scale to purify the variant CNBr fragment (7, 18–20); (*v*) HPLC peptide mapping of a tryptic or *Staphylococcus aureus* V8 protease digest of the purified CNBr fragment (7, 18–21); (*vi*) amino acid analysis of the variant peptide(s) with the Beckman model 121M amino acid analyzer; (*vii*) automated sequence determination with the Beckman model 890C sequencer or with the Applied Biosystems model 477B sequencer. In Japan, the JEOL model JLC-200A amino acid analyzer was used; sequence analysis of the Komagome peptides was done with protein sequencers through the kindness of companies as follows: Applied Biosystems model 477A, Shimadzu (Kyoto) model PSQ-1, and MilliGen/Biosearch (Novato, CA) model 6600. Albumin variants that appeared to be proalbumins were sequenced directly with the omission of steps *ii–vi*. Tryptic peptides are designated T, V8 protease peptides are designated S, and both types are numbered consecutively in their predicted order in the amino acid sequence. The amino acid sequence of albumin A (Fig. 1) and the codons at each position are based on the genomic sequence of Minghetti *et al.* (22).

RESULTS AND DISCUSSION

Electrophoretic Survey. A total of 17 instances of alloalbuminemia in unrelated individuals was identified by electrophoretic analysis in the three surveys: 11 at the Mayo Clinic, 3 at Iowa City, and 3 in Japan. Because the number of repeat analyses is unknown, it is difficult to state the precise frequency of alloalbuminemia. However, in the two American groups the incidence that is readily detectable by routine electrophoretic analysis is very low. For example, about 50,000 analyses are performed annually at the Mayo Clinic, and only 11 instances of alloalbuminemia were noteworthy enough to be collected over a period of some years. The repeat donations of the Japanese blood donors over a period of 6 months were <5%; thus, the frequency of alloalbuminemia in the Japanese donor group is about 3:20,000 with only a single example of each of the three types. This is lower than the cumulative frequency of albumin variants in the Radiation Effects Research Foundation Biochemical Genetics Study of 15,581 unrelated children at Hiroshima and Nagasaki, in which a total of 45 instances of five different alloalbumins was identified by the more sensitive method of vertical starch gel electrophoresis at several pH values (4, 5).

Family studies were done only for the variants for which the structural change had not previously been reported and inheritance demonstrated. Family studies limited to electrophoresis and partial structural analysis showed that albumin Iowa City-1 was present in a father and son, and sequence analysis showed that albumin Iowa City-2 was transmitted by the father to two daughters. Electrophoretic analysis showed that the Komagome-2 trait was carried by the father and two children.

Structural Studies of Japanese Alloalbumins. The three alloalbumin types identified in Japan were designated Komagome-1, -2, and -3 in the order of their ascertainment. Structural analyses were done on the total albumin (variant plus albumin A) by our earlier strategy (6, 7, 19–21). Komagome-3 proved to be a proalbumin of the Lille type (−2 Arg → His) (Table 1) and is discussed later. Komagome-2 had a previously unreported single-point substitution (128 His → Arg). Komagome-1 unexpectedly had the same mutation as albumin Naskapi (372 Lys → Glu); this occurs in polymorphic frequency in Athabascan- and Algonkian-speaking Amerindians but had not previously been reported in other ethnic groups except Eti Turks (3, 17).

Albumin Komagome-2 (128 His → Arg). Isoelectric focusing of a CNBr digest of albumin Komagome-2 did not indicate the fragment that had the substitution. Therefore, CB2 (residues 88–123, Fig. 1), which is not revealed by isoelectric focusing, was purified and analyzed. However, amino acid analysis of the tryptic peptides of CB2 showed that all were normal. CB3 (residues 124–298) was then purified by TSK G3000 SW gel chromatography, and two variant peptides (T14a and T14b) were isolated by Vydac C_{18} HPLC of a tryptic digest of CB3 (7, 18–20). Amino acid analysis and sequence analysis showed that T14a (residues 124–128, Fig. 1) had the sequence Cys-Thr-Ala-Phe-*Arg*; this is produced by trypsin because of the 128 His → Arg substitution. In confirmation, T14b had the sequence Asp-Asn-Glu-Glu-Thr-Phe-Leu-Lys, which represents the remainder (residues 129–136) of the normal tryptic peptide T14. The substitution 128 His → Arg corresponds to the codon change CAT to CGT (Table 1).

Albumin Komagome-1 (372 Lys → Glu) (Naskapi type). Isoelectric focusing showed that the substitution in albumin Komagome-1 was in CNBr fragment CB5 (residues 330–446). CB5 was purified by our two-step HPLC procedure (7, 18, 19). The variant peptide T49-50* (residues 360–389) (Fig. 2) was purified by Vydac C_{18} HPLC from a tryptic digest of CB5 (7, 18–20). Amino acid and sequence analysis established the substitution 372 Lys → Glu in the peptide, which had the sequence given in Fig. 2. Peptide T49-50* is formed by the union of T49 and T50 owing to the substitution by glutamic acid of lysine at position 372 for which trypsin is specific in normal albumin A.

The finding of a Naskapi-type alloalbumin in a Japanese is very interesting because the North American Indian tribes carrying this hitherto apparently unique trait in polymorphic frequency are thought to represent a late migration from mainland Asia across the Bering Strait, perhaps 12,000 years ago (3). Rumors that a Siberian ethnic group carries an electrophoretic albumin variant similar to Naskapi have come to our attention, but we have been unable to procure any specimens. However, albumin Mersin found in Eti Turks does have the Naskapi-type 372 Lys → Glu substitution (17). Of course, the occurrence of a Naskapi-type alloalbumin in a single person among the many thousands of Japanese

Table 1. Mutations in albumin genetic variants in the Iowa City, Mayo, and Japanese series

Type of variant	Mobility	Structural change	Codon change	Iowa City	Mayo	Japan
Lille	+2	−2 Arg → His	CGT → CAT		EW220	Komagome-3
Christchurch	+2	−1 Arg → Gln	CGA → CAA		JW180	
Arg-albumin	+1/+2	+1 Asp → Val	GAT → GTT	Iowa City-2		
Komagome-2	+1	128 His → Arg*	CAT → CGT			Komagome-2
Iowa City-1	+1	365 Asp → Val*	GAT → GTT	Iowa City-1		
Naskapi	−2	372 Lys → Glu	AAA → GAA			Komagome-1
B	+2	570 Glu → Lys	GAG → AAG	Iowa City-3	MT610†	

*Substitutions not previously reported.

†In the Mayo series, albumin B was identified by structural study in two other unrelated individuals (RW346 and SH420). Several other examples appeared to be albumin B by electrophoretic and chromatographic criteria, and there was insufficient amount of three specimens for study.

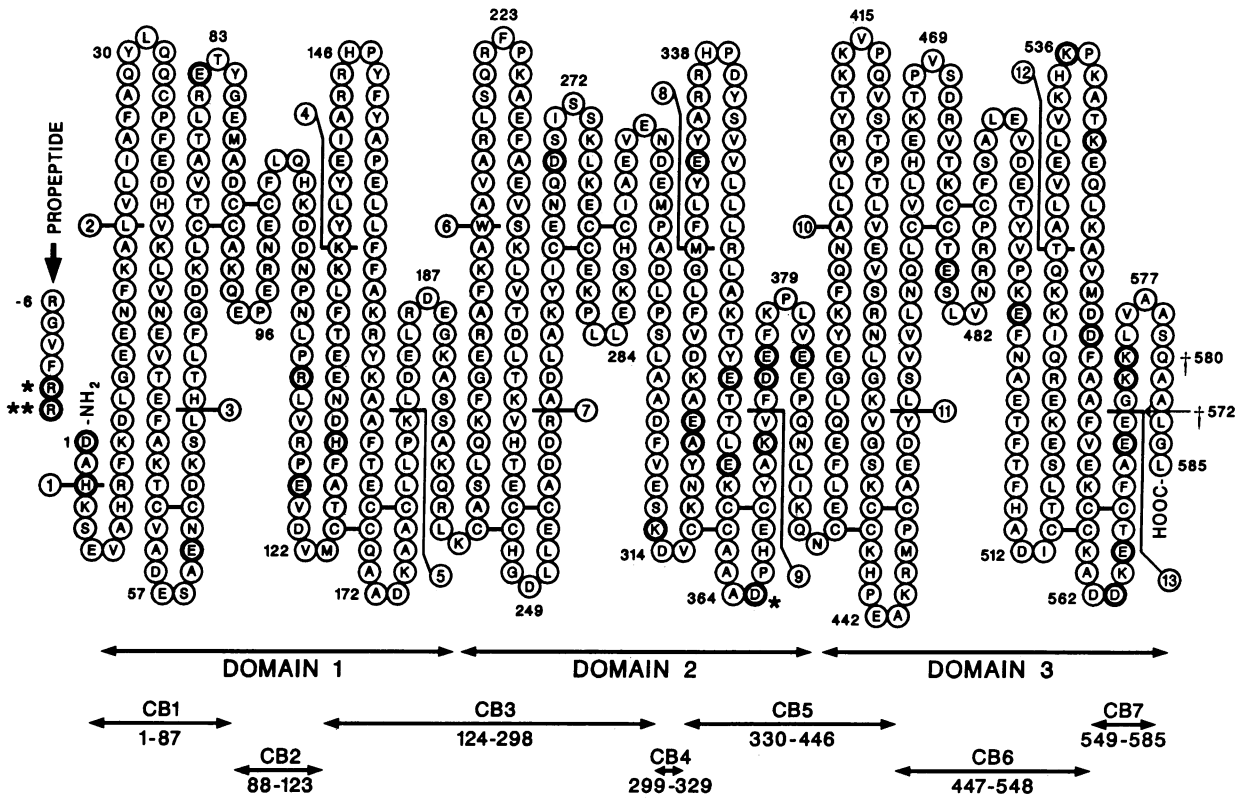

FIG. 1. Amino acid sequence of human serum albumin. The locations of single amino acid substitutions in genetic variants are identified by bold circles. The protein sequence is given in the one-letter code and is based on the genomic sequence analysis of Minghetti *et al.* (22), who also identified the location of the introns (designated 1–13, circled). CB1–CB7 refer to CNBr fragments, for which the positions of the amino acid residues are given. The figure is updated from figure 1 and table 1 of Arai *et al.* (21), who list previously reported substitutions. In addition to other variants described in the text and listed here in Table 1, the figure also includes the following alloalbumins: Torino (60 Glu → Lys) (10), Vibo Valentia (82 Glu → Lys) (10), Castel di Sangro (536 Lys → Glu) (10), Sondrio (333 Glu → Lys) (23), and albumin Dublin (479 Glu → Lys) (24). A single asterisk denotes positions where two substitutions have been established; the double asterisk in the propeptide indicates the site of three substitutions. The dagger at position 572 (boundary of exons 13 and 14) refers to carboxyl-terminal variant Venezia, which results from a skipping of exon 14 and a joining of exons 13 and 15 (25). The dagger at position 580 refers to albumin Catania, which has an altered carboxyl-terminal sequence because of a frameshift (25).

studied in this and in our previous surveys (a total of >50,000 individuals) (5, 6, 20, 21) may be the fortuitous result of a rare but independent mutation.

Iowa City Alloalbumins. Three alloalbumin types were ascertained in the Iowa City survey. One of these, Iowa City-3 was an albumin B and is discussed later. The other two are described below.

Albumin Iowa City-1 (365 Asp → Val). This alloalbumin was detected in a 59-year-old male of Swedish descent whose serum was used for the structural study. A family study showed that his son inherited the trait, but his daughter did not. The total albumin was purified by the two-step HPLC procedure of Huss *et al.* (18, 19). Analytical isoelectric focusing of a CNBr digest of the carboxymethylated protein indicated the substitution was in CNBr fragment CB5 (residues 330–446). CB5 was purified by our two-step HPLC procedure (18, 19) and was digested with trypsin. The variant peptide was identified and isolated by HPLC on a Vydac C_{18} column (7, 18–20). Amino acid analysis indicated it was similar to tryptic peptide T49 (residues 360–372) but that the single expected aspartic acid residue was replaced by valine. Sequence analysis with the Beckman sequenator established the substitution 365 Asp → Val (Fig. 2); this corresponds to a single point mutation in the codon GAT to GTT (Table 1).

The mutation (365 Asp → Val) producing this alloalbumin has not previously been reported; however, a different mutation at the same site (365 Asp → His) has been described for albumin Parklands by Brennan (12). There are several reports demonstrating multiple mutations in the diarginyl sequence of the albumin propeptide—e.g., −1 Arg → Gln, −1 Arg →

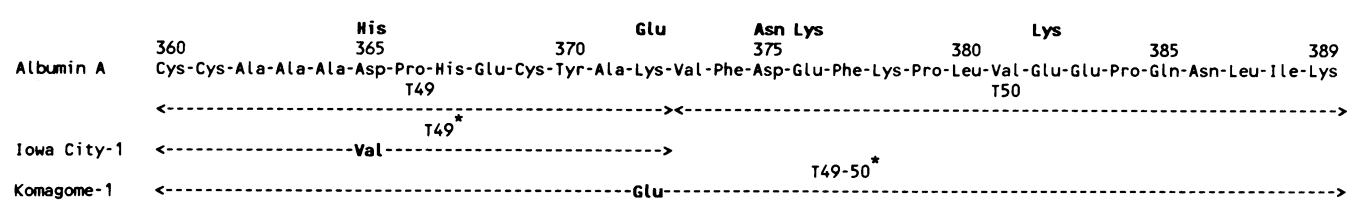

FIG. 2. Single-point substitutions identified in the amino acid sequence of serum albumin from positions 360 to 389. The sequence of normal albumin A is given, and normal tryptic peptides T49 and T50 are shown below. Amino acid substitutions in alloalbumins are shown in bold type. Substitutions reported here are given below the sequence: 365 Asp → Val, Iowa City-1; 372 Lys → Glu, Komagome-1. Previously reported substitutions are given above the sequence: 365 Asp → His, Parklands (12); 372 Lys → Glu, Naskapi and Mersin (17); 375 Asp → Asn, Nagasaki-2 (6); 376 Glu → Lys, Tochigi (6); 382 Glu → Lys, Hiroshima-2 (6). Peptides T49* and T49-50* are discussed in the text. Two nearby mutations are 354 Glu → Lys, Hiroshima-1 (6), and 358 Glu → Lys, Porto Alegre I and Coari I (7).

Pro, −1 Arg → Leu, and −2 Arg → His, −2 Arg → Cys (9, 11, 13, 16, 26). However, two different mutations at the same site in the mature albumin molecule have not previously been recorded. The index case for Iowa City-1 was of Swedish descent; however, we have not identified this alloalbumin in a collaborative survey covering some 20,000 Swedes resident in Sweden (Carl-Bertil Laurell, F.W.P., J.M., Y.S., S.W., and E.D., unpublished experiments).

Albumin Iowa City-2 (−1 Arg, +1 Asp → Val). This trait, which was present in a family of English/Scottish/Irish descent, was expressed as a broad albumin band in electrophoresis at pH 8.6 that suggested the presence of a slow component. In fact, by use of HPLC with a preparative DEAE 5 PW column and a pH gradient from pH 5.2 to 4.5 in 20 mM sodium acetate buffer the albumin was separated into three components: A with the electrophoretic mobility of normal albumin A, V1 with a mobility of about +2 slow, and V2 with a mobility approaching +1 slow. The relative proportion of the three components was as follows: A, 44%; V1, 25%; and V2, 31%. Protein sequence analysis of each of the three components (Fig. 3) showed that A was normal mature albumin A with the amino-terminal sequence Asp-Ala-His-Lys-. . . . The second component V2 was a variant albumin with the sequence *Val*-Ala-His-Lys-. . . indicating a 1 Asp → Val substitution, and the third component V1 was a variant arginyl-albumin with the sequence *Arg-Val*-Ala-His-Lys-. . . in which the arginyl residue is derived from the −1 position of the propeptide. Similar studies including sequence analysis of the V1 and V2 components confirmed the structural change in the albumin of both daughters of the index case.

Several other examples of arginyl-albumins have been reported (9, 14, 15). In one type (9) a mutation of arginine at the −2 position to cysteine (−2 Arg → Cys) produces the propeptide sequence Arg-Gly-Val-Phe-*Cys*-Arg followed by the normal amino-terminal sequence of albumin A (Asp-Ala-His-Lys . . .) (Fig. 3). This mutation in the diarginyl convertase cleavage site gives rise to two post-translational products, the variant proalbumin and arginyl-albumin A (see Malmo 1 in Fig. 3). The second type of arginyl-albumin has the substitution 1 Asp → Val that we found in Iowa City-2.

	-6	-1	+1	+4
Proalbumin A	Arg-Gly-Val-Phe-Arg-	Arg-	Asp-	Ala-His-Lys-
Albumin A			Asp-	Ala-His-Lys-
Christchurch	Arg-Gly-Val-Phe-Arg-	Gln-	Asp-	Ala-His-Lys-
Lille	Arg-Gly-Val-Phe-	His-	Arg-	Asp-Ala-His-Lys-
Iowa-2, V1		Arg-	Val-	Ala-His-Lys-
Iowa-2, V2			Val-	Ala-His-Lys-
Malmo-1, Y	Arg-Gly-Val-Phe-	Cys-	Arg-	Asp-Ala-His-Lys-
Malmo-1, X		Arg-	Asp-	Ala-His-Lys-

FIG. 3. Propeptide and amino-terminal sequences of proalbumin A, mature (processed) albumin A (normal), proalbumins of the Christchurch, Lille, and Malmo-1 types, and arginyl-albumins of the Iowa City-2 and Malmo-1 types. Iowa-2, V1 and Iowa-2, V2 refer to the V1 and V2 forms of the Iowa City-2 alloalbumin described in the text. Substitutions in the sequences are underlined. Arginine in the −1 position of arginyl-albumins is in bold type. In proalbumins Christchurch and Lille substitutions in the normal propeptide sequence of Arg-Gly-Val-Phe-Arg-Arg (positions −6 to −1) prevent post-translational removal of the propeptide by a convertase enzyme specific for the diarginyl sequence. The −2 Arg → Cys substitution in the Malmo-1 type of proalbumin affects the specificity of the convertase resulting in partial cleavage to produce arginyl-albumin A. The 1 Asp → Val mutation in the Iowa City-2 type of alloalbumin affects cleavage of the unaltered propeptide to form two products arginyl-valyl-albumin (−1 Arg, +1 Asp → Val) and valyl-albumin (1 Asp → Val), but no proalbumin was detected.

This mutation has previously been reported for albumins Blenheim (14) and Bremen (9). Apparently the 1 Asp → Val substitution affects the specificity of the cleavage enzyme, and the normal propeptide is processed in two stages. The first stage yields arginyl-valyl-albumin (−1 Arg, 1 Asp → Val) (Iowa-2, V1 in Fig. 3). In the second stage the arginine is removed, and the mature variant alloalbumin with amino-terminal valine is formed (1 Asp → Val) (Iowa-2, V2 in Fig. 3). However, we found no evidence by sequence analysis for an intact proalbumin form of Iowa City-2 in the HPLC chromatographic profiles of fresh specimens of citrated or heparinized plasma or serum from the Iowa City-2 index case.

Structural Studies of Proalbumins and Albumin B. As expected, most of the albumin variants identified in this collaborative study are the types that are most often encountered in such surveys—namely, proalbumins Christchurch and Lille and albumin B (8, 10, 20). Therefore, these three variant types are discussed only briefly and with reference to earlier reports for experimental details.

Proalbumin Christchurch type (−1 Arg → Gln). A single example of the proalbumin Christchurch type (13) with the substitution −1 Arg → Gln was found, Mayo JW180 (Table 1). Limited tryptic digestion of the purified albumin variant and isoelectric focusing of a CNBr digest indicated the substitution was associated with CNBr fragment CB1 (residues 1–87). Sequence analysis of the purified intact alloalbumin established that it was a proalbumin with the propeptide sequence Arg-Gly-Val-Phe-Arg-*Gln*- (positions −6 to −1) (Fig. 3). This substitution in the normal propeptide sequence of Arg-Gly-Val-Phe-Arg-Arg- prevents removal of the propeptide by a convertase enzyme specific for the Arg-Arg sequence during post-translational processing (9, 13–16). The Christchurch type of proalbumin has previously been reported in a number of individuals of various European descents (8, 13, 16) and also in Japanese (6, 20, 21).

Proalbumin Lille type (−2 Arg → His). One example of the proalbumin Lille (11) type with the substitution −2 Arg → His was found in the Mayo series (Mayo EW220) and one was found in the Japanese series (Komagome-3) (Table 1). The structural change in Mayo EW220 was established essentially by the procedure described above for Mayo JW180—that is, limited tryptic digestion and isoelectric focusing suggested that Mayo EW220 had a substitution associated with CB1 and was a proalbumin. Sequence analysis of the intact purified variant showed it was a proalbumin with the propeptide sequence Arg-Gly-Val-Phe-*His*-Arg- (positions −6 to −1) (Fig. 3). The His-Arg sequence also inhibits removal of the altered propeptide by the convertase.

The same substitution was found for albumin Komagome-3 (Table 1). Although limited tryptic digestion did not indicate it was a proalbumin, isoelectric focusing of a CNBr digest showed the substitution was associated with CB1. The variant and normal albumins were purified by HPLC. CB1 fragments from both fractions were combined and digested with V8 protease. One V8 peptide that was purified by Vydac C_{18} HPLC had the sequence Arg-Gly-Val-Phe-*His*-Arg-Asp (positions −6 to +1); this established that Komagome-3 had the change −2 Arg → His (Table 1). Although proalbumins of the Lille type were initially identified in individuals of various European descents (11, 16), they were later found by us in Japanese (20), in a Chinese (16), in Italians (10), and in Somali in Africa (10).

Albumin B type (570 Glu → Lys). Four instances of albumin B were established by structural analysis, one in the Iowa City series (Iowa City-3) and three in the Mayo series (MT610, RW246, and SH420) (Table 1). Because we have previously reported the structural study of other examples of albumin B in a number of unrelated persons (6, 20, 21), the experimental details are not repeated here. In several in-

stances the variant V8 protease peptide S58-59* (residues 566–571) was isolated by HPLC from a V8 digest of the purified CB7 fragment (residues 549–585, Fig. 1), and sequence analysis established the substitution 570 Glu → Lys in the peptide Thr-Cys-Phe-Ala-*Lys*-Glu. In other cases the purified CB7 fragment was sequenced directly and gave the same substitution.

Though rare, albumin B is the alloalbumin most often encountered in Caucasians and has a frequency of about 1:5000 (1, 8). We have also identified it in a number of Japanese (20, 21) and in a Cambodian (20). In a recent structural study we identified albumin B in a person of East Indian origin (albumin Victoria) (Y.S., J.M., S.W., and F.W.P., unpublished results). Thus, it is not surprising that 4 of the 11 alloalbumins investigated in this collaborative survey were proven to be of the albumin B type and several others appeared to be albumin B by physicochemical criteria.

Mutations Involving CpG Dinucleotides. Excluding several alloalbumins such as Naskapi that occur in polymorphic frequency in certain Amerindian tribes (3, 5), the alloalbumins most often encountered are those associated with mutations in CpG dinucleotides (9)—in particular, proalbumins of the Christchurch and Lille types and albumin B. In fact, these three are the most common alloalbumins in Caucasians with a cumulative frequency of about 1:3500 (8). Thus, it is no surprise that in the Mayo series all of the albumin variants for which the structural change was established were of these three types and occurred in Caucasians. In contrast, Komagome-2, Iowa City-1, and Iowa City-2 (1 Asp → Val) are part of a group of variants not involving CpG dinucleotides.

Clustering of Mutations. All of the structural changes described here are attributable to point mutations (Table 1), as are all known mutations in alloalbumins except for two carboxyl-terminal variants (10, 25). Fig. 1 shows that the amino acid substitutions are spread throughout the albumin molecule but that there is some tendency for clustering of mutations in the propeptide and in the carboxyl-terminal portions of domains 2 and 3. The highest concentration of observed substitutions is in the sequence segment from position 354 through 382 (Figs. 1 and 2). Hypermutability of CpG dinucleotides may be invoked to explain the clustering of mutations in the propeptide and to a lesser extent for the carboxyl-terminal portion of domain 3. However, this is not the case for the clustering at the end of domain 2 where only one of eight known mutations (Nagasaki-2, 375 Asp → Asn) (5) involves a CpG dinucleotide. Possibly the apparent clustering is associated with the folding of the polypeptide chain in this region and its tolerance for mutation. The sequence illustrated in Fig. 2 is a connecting segment from subdomain IIB to subdomain IIIA and appears to be exposed to solvent (27). This would permit freer ionization of the charged groups and would facilitate electrophoretic detection of variants. Also, this segment has not been implicated in ligand binding (28); thus, mutation in it should not significantly affect binding of physiologically and pharmacologically important ligands.

We thank Dr. T. Peters, Jr., for providing data from preliminary study of the Iowa City-2 albumin. We are indebted to Dr. David R. Hathaway for use of the Applied Biosystems sequencer and to Joyce Dwulet for sequence analysis with it and to Karen Huss for operation of the Beckman sequencer. This work was supported by National Institutes of Health Grant DK 19221 to F.W.P.

1. Tarnoky, A. L. (1980) *Adv. Clin. Chem.* **21**, 101–146.
2. Weitkamp, L. R., McDermid, E. M., Neel, J. V., Fine, J. M., Petrini, C., Bonazzi, L., Ortali, V., Porta, F., Tanis, R., Harris, D. J., Peters, T., Ruffini, G. & Johnson, E. (1973) *Ann. Hum. Genet.* **37**, 219–226.
3. Schell, L. M. & Blumberg, B. S. (1977) in *Albumin Structure, Function, and Uses*, eds. Rosenoer, V. M., Oratz, M. & Rothschild, M. A. (Pergamon, New York), pp. 113–141.
4. Neel, J. V., Satoh, C., Goriki, K., Asakawa, J., Fujita, M., Takahashi, N., Kageoka, T. & Hazama, R. (1988) *Am. J. Hum. Genet.* **42**, 663–676.
5. Takahashi, N., Takahashi, Y., Isobe, T., Putnam, F. W., Fujita, M., Satoh, C. & Neel, J. V. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8001–8005.
6. Arai, K., Madison, J., Huss, K., Ishioka, N., Satoh, C., Fujita, M., Neel, J. V., Sakurabayashi, I. & Putnam, F. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6092–6096.
7. Arai, K., Huss, K., Madison, J., Putnam, F. W., Salzano, F. M., Franco, M. H. L. P., Santos, S. E. B. & Freitos, M. J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1821–1825.
8. Fine, J. M., Marneux, M. & Rochu, D. (1987) *Am. J. Hum. Genet.* **40**, 278–286.
9. Brennan, S. O., Arai, K., Madison, J., Laurell, C.-B., Galliano, M., Watkins, S., Peach, R., Myles, T., George, P. & Putnam, F. W. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 3909–3913.
10. Galliano, M., Minchiotti, L., Porta, F., Rossi, A., Ferri, G., Madison, J., Watkins, S. & Putnam, F. W. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8721–8725.
11. Abdo, Y., Rousseaux, J. & Dautrevaux, M. (1981) *FEBS Lett.* **131**, 286–288.
12. Brennan, S. O. (1985) *Biochim. Biophys. Acta* **830**, 320–324.
13. Brennan, S. O. & Carrell, R. W. (1980) *Biochim. Biophys. Acta* **621**, 83–88.
14. Brennan, S. O. (1989) *Mol. Biol. Med.* **6**, 87–92.
15. Brennan, S. O., Myles, T., Peach, R. J., Donaldson, D. & George, P. M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 26–30.
16. Takahashi, N., Takahashi, Y. & Putnam, F. W. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7403–7407.
17. Takahashi, N., Takahashi, Y., Blumberg, B. S. & Putnam, F. W. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4413–4417.
18. Huss, K., Putnam, F. W., Takahashi, N., Takahashi, Y., Weaver, G. A. & Peters, T., Jr. (1988) *Clin. Chem.* **34**, 183–187.
19. Huss, K., Madison, J., Ishioka, N., Takahashi, N., Arai, K. & Putnam, F. W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6692–6696.
20. Arai, K., Ishioka, N., Huss, K., Madison, J. & Putnam, F. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 434–438.
21. Arai, K., Madison, J., Shimizu, A. & Putnam, F. W. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 497–501.
22. Minghetti, P. P., Ruffner, D. E., Kuang, W.-J., Dennison, O. E., Hawkins, J. W., Beattie, W. G. & Dugaiczyk, A. (1986) *J. Biol. Chem.* **261**, 6747–6757.
23. Porta, F. A., Galliano, M., Rossi, A. & Porta, F. (1990) *Boll. Ospedale Varese* **19**, 197–210.
24. Sakamoto, Y., Davis, E., Madison, J., Watkins, S., McLaughlin, H., Leahy, D. T. & Putnam, F. W. (1991) *Clin. Chim. Acta*, in press.
25. Watkins, S., Madison, J., Davis, E., Sakamoto, Y., Galliano, M., Minchiotti, L. & Putnam, F. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5959–5963.
26. Galliano, M., Minchiotti, L., Stoppini, M. & Tarnoky, A. L. (1989) *FEBS Lett.* **255**, 295–299.
27. Carter, D. C. & He, X.-m. (1990) *Science* **249**, 302–303.
28. Peters, T., Jr. (1985) *Adv. Protein Chem.* **37**, 161–245.