Polymorphism of the Human Complement C4 and Steroid 21-Hydroxylase Genes

Restriction Fragment Length Polymorphisms Revealing Structural Deletions, Homoduplications, and Size Variants

Peter M. Schneider, Michael C. Carroll, Chester A. Alper, Christian Rittner, Alexander S. Whitehead, Edmond J. Yunis, and Harvey R. Colten
Divisions of Cell Biology and Immunology, Children's Hospital, Center for Blood Research, Dana Farber Cancer Institute, Departments of Pediatrics and Pathology, Harvard Medical School, Boston, Massachusetts 02115;
Institut für Rechtsmedizin, Universität Mainz, D-6500 Mainz, Federal Republic of Germany

Abstract

Several autoimmune disorders as well as congenital adrenal hyperplasia (CAH) are either associated or closely linked with genetic variants of the fourth component of complement (C4A and C4B) and the enzyme steroid 21-hydroxylase (21-OH). These proteins are encoded by genes that are located downstream from the genes for complement proteins, C2 and factor B (BF) between HLA-B and -DR in the major histocompatibility complex (MHC). Previous studies of variants and null alleles were based on electrophoretic mobility of C4 protein and linkage with disease phenotypes. These data did not permit analysis of the basis for the observed null alleles and duplicated variants. We studied this region of the MHC in 126 haplotypes for a structural analysis of the four adjacent loci, C4A, 21-OHA, C4B, and 21-OHB. About half of the C4 genes typed as C4 null are deleted and several unrecognized homoduplicated C4 alleles were detected. Hence the frequencies of different C4 structural variants must be recalculated based on a direct analysis of the genes. Analysis of the C4/21-OH genes of patients with the classical (salt-wasting) form of CAH showed that some involve a deletion of the C4B and 21-OHB genes; whereas for two only the 21-OHB gene is deleted, i.e., the C4B gene is present. Together, these data provide a better understanding of the mechanisms generating and importance of deleted C4 and 21-OH null alleles in human disease.

Introduction

The short arm of human chromosome 6 carries the genes for factor B (BF),[1] for the second and the fourth component of human complement (C2, C4A, and C4B), and for the cytochrome P-450 steroid 21-hydroxylase (21-OH) (1–4). C2 and BF are about 0.5 kilobase pairs (kb) apart and are separated from the C4 genes by 30 kilobases (kb). The C4 genes are 10 kb apart, each with a 21-hydroxylase gene at the 3′ end at a distance of ~1.5 kb. The order of the genes in the direction of transcription is C2, BF, C4A, 21-OHA, C4B, 21-OHB (5–7), and they have been mapped between HLA-B and HLA-DR (8), which are separated by ~0.7 cM (9). Because of the tight linkage, these genes are usually inherited as a single "complotype" (10). In some haplotypes, linkage disequilibrium includes the HLA-B and -DR region forming "extended haplotypes" (11).

The high degree of polymorphism in C4 has led to speculation that susceptibility to some autoimmune diseases might be due to differences in the interaction of C4 variants with other polymorphic complement proteins, antigens, and antibodies (12, 13). The frequency of null alleles at the C4 loci (0.10–0.15) is relatively high (14), and some possible associations of C4 null alleles (C4 QO) with diseases of the immune system have been suggested, e.g., systemic lupus erythematosus (SLE), scleroderma, and subacute sclerosing panencephalitis (SSPE) (15–17).

A defect of the functional 21-OH gene results in congenital adrenal hyperplasia (CAH) and one mutation shows linkage disequilibrium with a rare MHC haplotype (18, 19). It has been shown that 21-OHB is the functional gene and that a deletion of the 21-OHA gene does not affect steroid metabolism (20). Since deletions of the 21-OH genes are correlated with null alleles at the adjacent C4 loci, it has been postulated that these C4 QO genes represent deletions as well (6). Several CAH haplotypes have therefore been included in this study.

Restriction fragment length polymorphisms (RFLPs) have proved useful in the investigation of the basis of C4 genetic variants (21) and can subdivide common C4 allotypes (22). Analysis of genomic and complementary (cDNA) sequences of C4A and C4B has shown that only eight nucleotide differences resulting in a regional change in six amino acids account for the class differences between the C4 isotypes. Even fewer exchanges were detected in allotypic variants at each locus (23, 24). A variation in size of the C4B gene (the short form is 16 kb, whereas the long form is identical in size to the C4A gene at 22 kb) is a result of the presence or lack of a 6.8-kb intron in the 5′ region of the C4B gene (25, 26).

Some but not all C4A and C4B null alleles are due to large deletions that include a deletion of a flanking 21-OH gene (25). Molecular maps of genomic DNA showed a deletion of the C4A and 21-OHA genes or the 21-OHA and C4B genes. Three haplotypes with nondeleted C4 null alleles were also identified (25).

Here we present the use of a restriction fragment length polymorphism located at the 5′ end of both C4 genes to investigate the molecular basis of C4 QO alleles and duplications as well as the frequency of the long and the short form of the C4B gene. Applied together with a probe (7) to detect a fragment size difference of the 21-OH genes (20), this polymorphism provides information about the structure of the C4/21-OH gene complex, which can be used to further analyze the genetic basis of disorders associated with deficiencies of the C4 and 21-OH loci.

Address reprint requests to Dr. Colten, Department of Pediatrics, Washington University Medical School, 400 S. Kingshighway Blvd., St. Louis, MO 65110.

Received for publication 5 March 1986.

---

[1] *Abbreviations used in this paper:* BF, factor B; CAH, congenital adrenal hyperplasia; MHC, major histocompatibility complex; 17-OHP, 17-hydroxyprogesterone; RFLP, restriction fragment length polymorphism; SSC, standard saline citrate; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

J. Clin. Invest.
© The American Society for Clinical Investigation, Inc.
0021-9738/86/09/0650/08 $1.00
Volume 78, September 1986, 650–657

Methods

Selection of individuals. This study is based on the DNA typing patterns of 92 individuals from 32 Caucasian families, including 8 patients with CAH, 5 patients with celiac disease, and 4 patients with multiple sclerosis. All other subjects were healthy and were selected for the presence of haplotypes carrying null alleles at one or both of the C4 loci.

Complement and HLA typing. Blood for complement typing was collected into 1 mg/ml EDTA, centrifuged, and the serum was stored at −80°C. BF typing was carried out by agarose gel electrophoresis followed by immunofixation with goat anti-human BF antibody (Atlantic Antibodies, Scarborough, ME) (27). For C2 typing, plasma samples were subjected to isoelectric focusing in polyacrylamide gels (28). C4 typing was carried out by high voltage agarose electrophoresis in a discontinuous buffer system after desialation of the samples with neuraminidase, followed by immunofixation with goat anti-human C4 antibody (Atlantic Antibodies) (29). The alpha-chains of C4A and C4B were analyzed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) (30). Blood for HLA typing was collected into heparin at a final concentration of 10 U/ml. HLA-A, -B, -C, and -DR antigens were assigned by the microlymphocytotoxicity assay (31, 32).

Complotypes are given in arbitrary order as BF, C2, C4A, and C4B types in abbreviated form. Thus, "SC01" represents BF*S, C2*C, C4A*Q0, C4B*1.

Preparation of human genomic DNA and Southern blotting. For DNA preparation, individual blood samples of 30 ml each were collected into a syringe containing 2 ml of 100 mM EDTA. The DNA was extracted from white blood cells purified by sucrose sedimentation according to standard procedures (33). Individual DNA samples of 10 μg were digested with restriction enzymes at a concentration of 5 U/μg according to the manufacturer's instructions (New England Biolabs, Beverly, MA). The DNA fragments were separated by electrophoresis in a 0.8% agarose-gel (ultrapure agarose, BRL Laboratories, Bethesda, MD) and blotted onto nitrocellulose (BA85, Schleicher & Schuell, Keene, NH) (34, 35). Prehybridization and hybridization were carried out in a mixture containing 4 × standard saline citrate (SSC), 50 mM phosphate buffer pH 6.8, 1 mM EDTA, 0.1% SDS, 5% dextran sulfate, 5 × Denhardt's solution, 0.1 mg/ml yeast RNA, and 50 μg/ml sheared denatured salmon sperm DNA. As hybridization probe, 1 μg of purified cDNA insert was labeled with 100 μCi alpha-[^{32}P]dCTP (deoxycytidine triphosphate, New England Nuclear, Boston, MA) in a nick translation reaction (36). Hybridization was performed for 15 h at 65°C, followed by two to three subsequent washes of 30 min each in 0.2 × SSC/0.1% SDS at the same temperature. The hybridization patterns on the blots were visualized by autoradiography.

C4 and 21-OH specific DNA probes. The cDNA probe used to generate the 5′ C4 polymorphism was derived from the full length C4 cDNA clone pAT-A (23). The 500-bp probe specific for the 5′ ends of both C4 genes was obtained after Bam HI/Kpn I double digest, separated from the other fragments by agarose electrophoresis and purified by electroelution into a dialysis bag (37). The probe specific for the 21-OH genes was derived from the cosmid clone cos 1E3 and represents a 900-bp Bgl I fragment of the 21-OHA region (7). The 300-bp insert of cDNA clone pALU-7 is specific for the C4d region of both C4 genes (38). These inserts were purified as described above.

Interpretation of restriction fragment patterns. The 5′ C4 fragments were mapped relative to the 12- and 3.5-kb Kpn I fragments characteristic for the 5′ ends of the C4A and the C4B genes (5). The correlation of these fragments with the two forms of the C4B genes has been established by hybridization of Kpn I digested DNA with the cDNA probe pALU-7, which yields a 7.5-kb band for the long C4 genes and an 8.5-kb band for the short C4B gene (26, 39).

Several precautions were observed to obtain accurate and reproducible results. The evaluation of relative intensities of DNA fragments on Southern blots as evidence for the number of genes was only carried out within the same digest on the blot. Intensities of C4 and 21-OH bands within the same digest were not compared, as they depend on the size and specific activity of the radiolabeled probe. Assignment of a deletion or duplication was established in family studies to confirm the findings by segregation analysis.

Results

Pattern of the Taq I polymorphism. After digestion of genomic DNA with the restriction enzyme Taq I and hybridization with the labeled 5′ C4 probe, four different fragments were visualized. Fig. 1 shows the position of these fragments within the C4/21-OH region and the genotypes in which they occur. Southern blot analysis of DNA from homozygous individuals shows that the longest fragment (7.0 kb) is specific for the C4A gene locus, and the other three represent different alleles of the C4B gene. One of them (6.4 kb) is a marker for a deletion of the C4A gene in the complotype SC01 (BF*S, C2*C, C4A*Q0, C4B*1). This complotype is usually part of the extended haplotype *HLA-B8, DR3, SC01.* The 6.4-kb fragment is located at the 5′ end of the remaining short C4B gene in this haplotype. The 6.0-kb (C4B long) and 5.4-kb (C4B short) bands represent the two forms of C4B in normal haplotypes with two C4 loci. The 21-OH genes are identified by the 3.2-kb (21-OHA) and 3.7-kb (21-OHB) Taq I bands.

Two different forms of C4B deletions were detected in family EM. (Fig. 2 *a*). Father A. has no C4B bands, but both the 21-OHA and the 21-OHB band are visible. This pattern suggests that both C4B genes are deleted. On one chromosome, the deletion includes a 21-OHA gene, and on the other the 21-OHB gene. Mother R. has a stronger C4A than C4B band, and a stronger 21-OHB than 21-OHA band. A possible deletion could include the 21-OHA and C4B genes on one chromosome, whereas the other carries all four genes. This assumed parental deletion pattern is confirmed by the analysis of segregating chromosomes in the children. Daughter P. has inherited both 21-OHA/C4B deleted haplotypes; i.e., no 21-OHA or C4B fragments can be detected. Son L. has a deletion of the C4B/21-OHB genes on one chromosome, inherited from A.; i.e., the C4A and the 21-OHA bands are stronger than the C4B/21-OHB bands. Al-

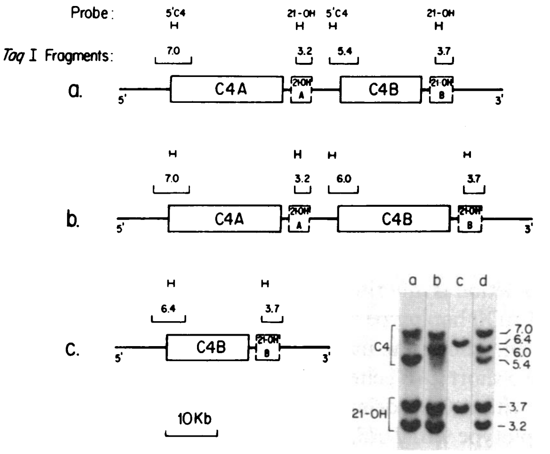

Figure 1. Map of the human C4 and 21-hydroxylase gene region illustrating the Taq I restriction fragment pattern. (*a*) 7.0/5.4-kb fragments: C4A gene in combination with short C4B gene (16 kb); (*b*) 7.0/6.0-kb fragments: C4A gene in combination with a long C4B gene (22 kb); (*c*) 6.4-kb fragment: deletion of C4A and 21-OHA genes in combination with short C4B gene; (*d*) 7.0/6.0/5.4-kb fragments: heterozygous pattern of individual with one long and one short C4B gene.

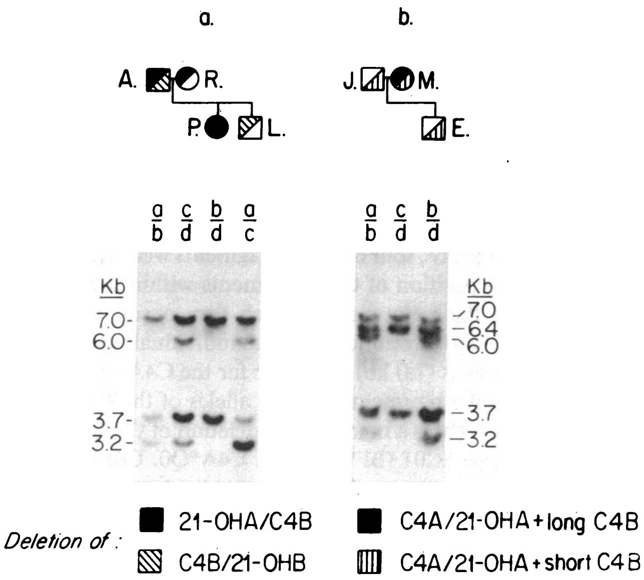

Figure 2. Taq I restriction fragment patterns of haplotypes carrying deletions of the C4 and 21-hydroxylase genes. (*a*) Family tree and Southern blot showing the segregation of 21-OHA/C4B and C4B/21-OHB deletions in family EM. The haplotypes indicated above the blot are: *a, HLA-A3, Cw6, Bw47, DR7, FC91,0,* deletion of C4B and 21-OHB; *b, HLA-A11, Cw4, B35, DR1, SC30,* deletion of 21-OHA and C4B; *c, HLA-A1, Cw4, B35, DRw6, SC31,* no deletion; *d, HLA-A2, Cw4, B35, DR1, SC30,* deletion of 21-OHA and C4B. (*b*) Family tree and Southern blot showing the segregation of C4A/21-OHA deletions in family AD. The haplotypes indicated above the blot are: *a,* HLA-Cw5, B44, DR3, SC01, deletion of C4A and 21-OHA in combination with short C4B gene; *b, HLA-A2, B7, DR2, SC31,* no deletion; *c, HLA-Cw5, B44, DR4, SC01,* deletion of C4A and 21-OHA in combination with long C4B gene; *d, HLA-A1, B8, DR3, SC01,* deletion of C4A and 21-OHA in combination with short C4B gene.

though the C4 fragment patterns of R. and L. are identical, the reversed intensities of the 21-OH bands indicate the presence of different deletions.

A rare complotype with a C4A deletion was found in family AD (Fig. 2 *b*). Mother M. has no detectable C4A protein as assessed by agarose electrophoresis and immunofixation of neuraminidase-treated plasma. However, a 7.0-kb 5′ C4 band is clearly visible together with a 6.4-kb band from the other chromosome, but no second 5′ C4B fragment is apparent. Based on the lack of a C4A gene product according to the protein typing, one could assume that the 7.0-kb band represents in this case the 5′ end of a long C4B gene in combination with a C4A deletion. The 21-OHA genes on both chromosomes are deleted as well. This deletion is inherited with *HLA-B44, DR4.* In contrast, the deleted AQ0 haplotype from father J. segregates with *HLA-B44, DR3* and shows the usual 6.4-kb band. This indicates the presence of a short C4B gene, as it is found in the most common haplotype with a C4A deletion and a short C4B gene, the extended haplotype *B8, DR3, SC01.* The paternal chromosome thus might have been generated by an ancestral crossing-over event between the HLA-B44 locus and the complement genes of the extended haplotype, whereas the C4A deletion with the long C4B gene on the maternal chromosome has apparently originated independently.

Analysis of the 21-hydroxylase genes in CAH patients. Eight patients with the classical (salt-wasting) form of CAH were studied together with their families. In two patients, the 21-OHB gene alone was deleted, and the C4B gene was intact (Fig. 3 *a*). Five patients had at least one 21-OHB gene (Fig. 3 *c* and *d*), whereas the other three individuals lacked the 3.7-kb Taq I fragment that represents the 21-OHB gene (Fig. 3 *a* and *b*). One haplotype had a deletion of the 21-OHA/C4B genes (Fig. 3 *e*), and eight haplotypes carried a C4B/21-OHB deletion. Of these, five occurred in the CAH-linked extended haplotype Bw47, DR7, FC91,0. This haplotype was originally typed as FC0,31, and renamed on the basis of structural data (see discussion) and protein typing in comparison to the known C4 allotypes *A 3, A 1, A 91, A 92,* and *B 3* (data not shown).

Detection of hetero and homoduplicated C4 loci. An individual carrying a heteroduplication of the C4B locus (two different C4B alleles on the same chromosome) together with a C4A null allele on one chromosome (C4A*Q0, B*1,2) and a common C4 haplotype on the other (C4A*3, B*1) has been analyzed by C4 protein and DNA typing (Fig. 4 *a* and *b*). Father B. (*a* and *b*) has equally strong 7.0 kb (C4A) and 5.4 kb (C4B short) bands and a weak 6.0-kb (C4B long) band, thus indicating the presence of five C4 alleles (one of them a nondeleted C4A null allele). With the 21-OH probe, the 3.7 kb (21-OHB) band is slightly stronger than the 3.2-kb (21-OHA) band, and suggests that the 21-OHB gene is also duplicated on one chromosome. Chromosome b probably has the duplicated C4B/21-OHB genes as son M. (*a* and *d*) does not appear to have inherited this pattern.

Similar restriction fragment patterns were observed in other individuals with duplicated C4 loci that were not detected by C4 protein typing. In family CH., a homoduplication of the C4B locus (two identical C4B alleles on the same chromosome) was detected on Southern blot analysis. This duplication segregates with chromosome b from father S. to his daughters (Fig. 5, *top*). The finding was supported by the C4 typing gels (Fig. 5, *bottom*), where stronger C4B 1 protein patterns in individuals S., Jl. and D. were observed. A stronger C4B than C4A alpha chain was also observed after separation of immunoprecipitated and reduced C4 from plasma samples by SDS-PAGE (data not shown). To obtain further evidence for the duplication, DNA samples of family CH. were also digested with the restriction enzymes Bgl II and Kpn I and hybridized with the C4 and 21-OH probes (data not shown). These patterns were used to establish a partial restriction map of the duplicated gene region (Fig. 6). In addition, another homoduplication of the C4B locus in combination with a nondeleted C4A null allele in the haplotype *HLA-A11, B17, DRw6, FC01,1* has been detected (data not shown).

Screening for deletions of the C4/21-OH genes. To obtain data on the occurrence and type of deletions in the C4/21-OH region, 126 individual haplotypes were analyzed in family studies using the 5′ C4 Taq I polymorphism. The families were selected for the presence of one or more C4 null alleles, based on protein typing. Also, members of some families are carriers of disorders like CAH, multiple sclerosis, and celiac disease. The data obtained thus are not representative for frequencies in the normal population.

The Southern blot analysis of C4 haplotypes carrying a null allele (C4 Q0) showed the following: (*a*) Deleted to nondeleted alleles showed a ratio of 3:2. (*b*) All deleted C4A genes occurred with the complotype SC01 and included deletions of the 21-OH gene. (*c*) C4B null haplotypes with a deletion included either deletion of a 21-OHA or a 21-OHB gene. (*d*) Deletion of the C4B gene was not observed in *SC30* and *FC30* complotypes in combination with *HLA-B44.*

Name	HLA- A	C	B	DR	BF/C2/C4A/C4B				TaqI-fragments: C4A	OHA	C4B	OHB
<u>a.</u> F.C.	w24		44		F	C	4	Q0	7.0	3.2	-	-
	w24	w2	w62	w6	F	C	3	1	7.0	3.2	6.0	-
C.S.	3		w47	7	F	C	91	Q0	7.0	3.2	-	-
	2	w2	51	3	S	C	3	1	7.0	3.2	6.0	-
<u>b.</u> K.G.	3		w47	7	F	C	91	Q0	7.0	3.2	-	-
	3	w3	40	4	S	C	3	Q0	7.0	3.2	-	-
<u>c.</u> C.D.	1		8	1	S	B	4	Q0	7.0	3.2	-	-
	3		51	2	F	C	3	1	7.0	3.2	6.0	3.7
A.E.	3	w6	w47	7	F	C	91	Q0	7.0	3.2	-	-
	1	w4	35	1	S	C	3	1	7.0	3.2	6.0	3.7
S.P.	2	w5	44	4	S	C	3	Q0	7.0	3.2	6.0	3.7
	11		w47	7	F	C	91	Q0	7.0	3.2	-	-
<u>d.</u> A.R.	28	w3	w60		S	C	Q0	2	7.0	3.2	5.4	3.7
	28	w6	w47	7	F	C	91	Q0	7.0	3.2	-	-
<u>e.</u> L.L.	28		14	4	FS	C	3	1	7.0	3.2	5.4	3.7
	3		w60	1	FS	C	3	Q0	7.0	-	-	3.7

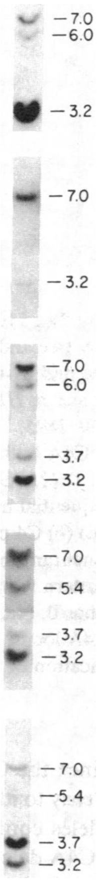

Figure 3. Haplotypes and Taq I restriction fragment patterns of patients with the classical (salt-wasting) form of CAH. (*a*) homozygous deletion of 21-OHB and heterozygous deletion of C4B; (*b*) homozygous deletion of C4B and 21-OHB; (*c*) heterozygous deletion of C4B and 21-OHB with remaining long C4B gene; (*d*) heterozygous deletion of C4B and 21-OHB with remaining short C4B gene; heterozygous deletion of 21-OHA and C4B with both remaining 21-OHB genes. (The Southern blot figures in lanes *a, d,* and *e* are obtained from sequential hybridizations of the C4 and 21-OH probes on the same blot.)

The individual haplotypes with deleted null alleles are listed in Table I *A*. All but three C4A Q0 deletions are linked to *HLA-B8, DR3,* and all 21 show the complotype *SC01*. In this group, the 21-OHA gene is also deleted. All of the four 21-OHA/C4B deletions occur in the *SC30* and *FC30* complotypes, but with various HLA determinants. Six of nine C4B/21-OHB deletions segregate with *HLA-Bw47, DR7, FC91,0,* a haplotype linked to CAH. The other three haplotypes with a deletion of the 21-OHB gene are also obtained from CAH families.

C4A Q0 haplotypes where the gene is not deleted show no linkage to specific HLA types (Table I *B*). In contrast, 12 of 17 nondeleted C4B null alleles segregate with *HLA-B44, SC30,* and seven of these with the extended haplotype *B44, DR4, SC30.* All nondeleted C4B null alleles are found in the common *SC30* and *FC30* complotypes.

Distribution of the long and short C4B genes. The analysis of long and short C4B genes showed the following: 61 of 113 haplotypes without deletion of the C4B locus carry a long C4B gene. Only the long C4B variant is found when the C4B gene is present but not expressed. C4A null alleles that are not deleted may occur with both sizes of C4B genes. Most C4A null haplotypes with a deletion of the C4A gene (19/21) carry the short C4B gene.

The most common complotypes investigated are represented in Fig. 7 together with the distribution of the long and the short form of the C4B gene. Most of the *SC31* complotypes (34/38) have the long C4B gene, and 11 of these are found in the extended haplotype *HLA-B7, DR2, SC31.* In contrast, a majority of the *FC31* complotypes (9/13) have the short C4B gene; three of them are linked to *HLA-B44, DR7.* With the single exception of the *FC32* complotype, which segregates with *HLA-B51, DR4,* all haplotypes with a C4B2 allele carry the short C4B gene, as do all *HLA-Bw57, SC61* haplotypes.

Discussion

The biochemical analysis of polymorphic complement gene products has already revealed significant functional differences between isotypic and allotypic C4 variants (40–42). The analysis of the structural basis of C4 null alleles by recombinant DNA techniques may help to describe mechanisms leading to duplications and deletions as well as to understand the effect of these events in regard to possible autoimmune disease associations (15, 17, 43). In the case of CAH, a close genetic linkage of this disease to a limited number of MHC-haplotypes has been established (18, 19). This linkage is due to the location of the gene for steroid 21-hydroxylase within the C4 region. A DNA fragment length polymorphism generated by the restriction enzyme Taq I provides an effective and convenient way to analyze simultaneously the genomic structure of the four adjacent loci of

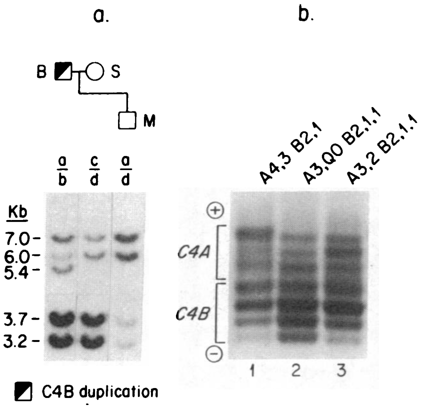

Figure 4. Heteroduplication of C4B locus as detected by Taq I restriction fragment pattern and C4 protein typing. (*a*) Family tree and Southern blot revealing duplication of a short C4B gene and 21-OHB in family BR. The haplotypes indicated above the blot are: *a, HLA-A31, Cw2, B51, DRw8, SC31; b, HLA-A26, Cw3, Bw60, DR5, SC01,2,* duplication of the C4B locus with nondeleted C4A null allele; *c, HLA-A1, Cw6, Bw57, DR5, SC31; d, HLA-A2, B18, DR2, SC31.* (The figures in the first two lanes are obtained from sequential hybridizations of the C4 and 21-OH probes on the same blot.) (*b*) C4 protein typing pattern, obtained by agarose electrophoresis of neuraminidase-treated plasma and immunofixation; lane *1,* control pattern with two expressed C4A and C4B loci; lane *2,* C4 pattern of father B. (4a, lane *1*) with duplicated C4B locus and a single expressed C4A locus; lane *3,* control pattern with two expressed C4A loci and duplication of the C4B locus.

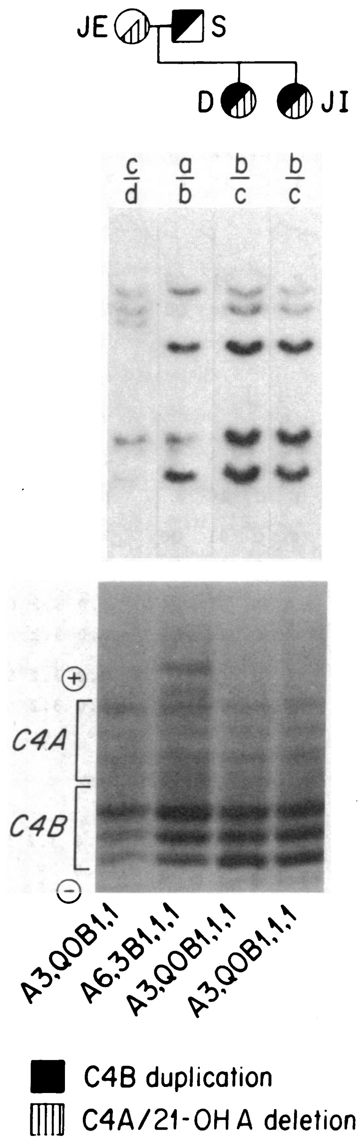

Figure 5. Homoduplication of C4B locus as detected by Taq I restriction fragment pattern and C4 protein typing. Family tree, Southern blot and C4 protein typing pattern demonstrating the segregation of a C4B homoduplication in family CH. The haplotypes indicated above the blot are: *a, HLA-A1, Cw6, Bw57, SC61; b, HLA-Aw33, B44, DR7, SC31,1,* duplication of the 21-OHA and C4B genes; *c, HLA-A1, B8, DR3, SC01,* deletion of the C4A and 21-OHA genes; *d, HLA-A24, Cw7, B7, DR2, SC31.*

the C4A, 21-OHA, C4B, and 21-OHB genes, since the C4 and 21-OH bands are well separated on the blot and easy to identify.

Our results on the structure of C4A null alleles confirm in family studies previous findings that mapped a C4A deletion in combination with a short C4B gene, obtained by analyzing clones from a genomic library (25). In addition, we were able to detect two deleted C4A Q0 haplotypes with a long C4B gene (Fig. 2 *d,* Table I). The Taq I restriction site upstream from the C4 region appeared preserved (Fig. 1), and in the case of a C4A deletion the length of the resulting fragment depends on the position of the 3′ restriction site inside the C4B gene; i.e., 6.4 kb with a short B gene and 7.0 kb with a long B gene. Deletions of the 21-OHA/C4B genes similar to the one described (25) in the haplotype *HLA-B7, DR2, SC30* were detected in four other haplotypes (Table I), and all share the same complotype *SC30.*

Deletions affecting the C4B/21-OHB genes occur mostly in haplotypes found in CAH patients (Table I). With the exception of *HLA-Bw47, DR7,* these haplotypes differ in HLA type as well as in complotype. The extended haplotype *Bw47, DR7, FC91,0* has a frequency of >20% among CAH patients (19). The functional 21-OHB gene (6) and the C4B gene are deleted in this haplotype. Southern blot analysis of genomic DNA indicated that the deletion extends from the end of the 3.2-kb Taq I fragment in the 21-OHA gene to the homologous region 3′ to the 21-OHB gene (44).

While in most cases (eight haplotypes) deletions of the 21-OHB gene included the C4B gene, in two unrelated CAH patients only the 21-OHB gene was deleted. In contrast, five haplotypes in CAH patients with no apparent deletion of the 21-OHB gene were observed as well. Thus, the classical (salt-wasting) form of CAH is caused by major structural deletions, but is also a result of defects in 21-OH expression. Therefore, every CAH haplotype not carrying a C4B null allele must be analyzed further. On a functional level, the analysis of the 17-hydroxyprogesterone (17-OHP) response to ACTH stimulation has provided a biochemical basis to assess the severity of the disease in heterozygous and homozygous carriers of CAH haplotypes (45).

Unequal crossing-over mechanisms have been suggested by several authors to account for the generation of deletions and duplications (25, 29, 46, 47). The high degree of structural homology in the C4/21-OH region might facilitate unequal pairing of chromatids during meiosis. Duplication of the nearby C2 and Factor B genes has not been detected (48, 49), although the genes show a substantial degree of homology [39% identity (50)]. Null alleles for C2 that have been analyzed do not appear to be due to large deletions, as the DNA restriction fragment patterns of functional and null alleles are identical (49). The near identity (>99%) among the duplicated C4/21-OH genes favors recombinant events.

Unequal crossing-over in the C4 region would result in one chromosome with a single and the other with three C4 genes. Based on the C4 protein typing data, the frequencies of null

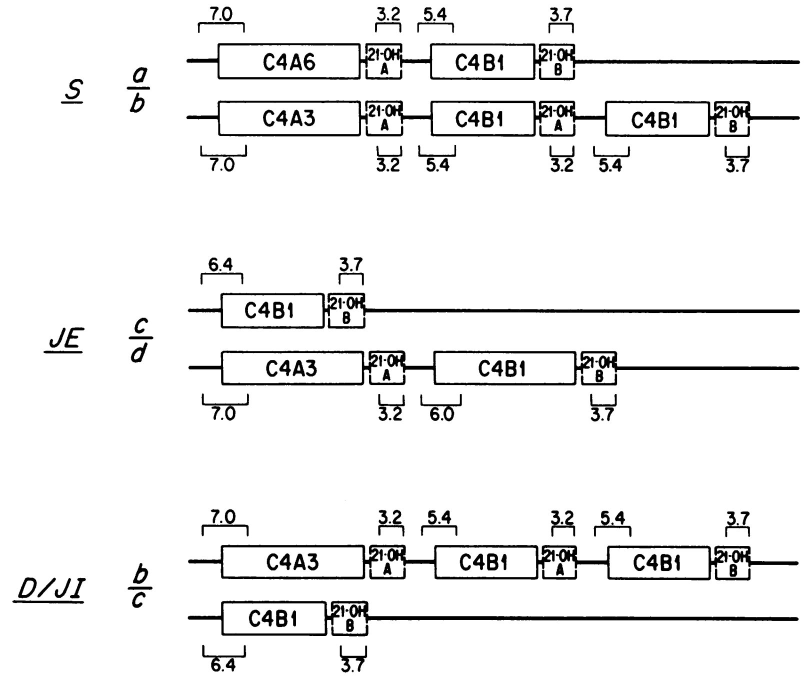

Figure 6. Homoduplication of C4B locus: partial restriction fragment maps of segregating haplotypes of family CH. Restriction fragment map of the segregating haplotypes showing the location and number of the Taq I fragments in the C4/21-OH region of both chromosomes of individuals S., JE., and D./JI. of family CH.

alleles [0.10–0.15 (14)] and duplications [0.0075–0.01 (46)] do not correlate. Our finding that not all C4 null alleles are deleted changes this ratio significantly. In this study, 34 of 57 C4 null alleles were deleted. This figure includes nine rare CAH haplotypes, so that only about 50% of all C4 null alleles might be due to deletions in the normal population. Also, we have ob-

Table I. Haplotypes Carrying C4 Null Alleles with and without Deletions

A. Haplotypes with deletions of the C4 and 21-hydroxylase genes:	HLA-A	C	B	DR	Bf /	C2 /	C4A /	C4B	(total)		(CAH)
I. C4A/21-OHA deletions	*	—	8	3	S	C	Q0	1	(18)		
	—	w5	44	3	S	C	Q0	1	(1)		
	2	w5	44	4	S	C	Q0	1	(1)		
	2	w2	51	1	S	C	Q0	1	(1)	(=21)	
II. 21-OHA/C4B deletions	*	w4	35	1	S	C	3	Q0	(2)		
	3	w4	35	w6	S	C	3	Q0	(1)		
	3	w3	w60	1	FS	C	3	Q0	(1)	(=4)	(1)
III. C4B/21-OHB deletions	*	[w6]	w47	7	F	C	91	Q0	(6)		(5)
	3	w3	40	4	S	C	3	Q0	(1)		(1)
	1	—	8	1	S	B	4	Q0	(1)		(1)
	w24	—	44	—	F	C	4	Q0	(1)	(=9)	(1)
IV. 21-OHB deletions	w24	w2	w62	w6	F	C	3	1	(1)		(1)
	2	w2	51	3	S	C	3	1	(1)	(=2)	(1)

B. C4Q0 haplotypes without deletions:	HLA-A	C	B	DR	Bf /	C2 /	C4A /	C4B	(total)		(CAH)
I. C4AQ0 haplotypes	*	—	7	[2]	S	C	Q0	1	(2)		
	2	w2	51	1	S	C	Q0	1	(1)		
	28	w3	w60	—	S	C	Q0	2	(1)		(1)
	1	w4	35	—	S	C	Q0	5	(1)	(=5)	
II. C4BQ0 haplotypes	3	—	7	—	S	C	3	Q0	(1)		
	2	[w5]	44	4	S	C	3	Q0	(7)		(1)
	*	—	44	*	S	C	3	Q0	(5)		
	w26	w6	13	3	F	C	3	Q0	(1)		
	*	—	44	w6	F	C	3	Q0	(2)		
	28	—	51	w6	F	C	3	Q0	(1)	(=17)	

* Various determinants; [ ] this determinant or not defined; — not defined.

Table II. Extended Haplotypes [1] Recognized Among the Selected Haplotypes

[*]	[*]	(Extended haplotype)									
HLA-A	C	B	DR	Bf /	C2 /	C4A /	C4B	Frequency (normal, 1)	No. observed in this study	Deletion of C4	C4B size
1	w7	8	3	S	C	Q0	1	0.082	18	A	Short
3	—	7	2	S	C	3	1	0.043	11	—	Long
2	w5	44	4	S	C	3	Q0	0.029	7	—	Long
23	w4	44	7	F	C	3	1	0.027	3	—	Short
1	w6	w57	7	S	C	6	1	0.019	1	—	Short
2	w3	w62	4	S	C	3	3	0.016	1	—	Long

[1] Frequencies among normal Caucasians (11); [*] Predominant HLA-A and -C alleles.

tained evidence for an increased number of homoduplicated haplotypes, which have not been identified previously. Therefore, deletions and duplications may be more common than previously suggested and only by direct analysis of this region on the DNA level can the frequencies be determined accurately.

In the first analyses of the C4 gene region, only short C4B genes (16 kb) were detected (5, 7, 25). In the present study, a majority of the C4B genes analyzed were the same size as the C4A gene (22 kb). This includes all 17 C4B null haplotypes studied thus far, in which the C4B gene is present but not expressed. The C4B gene size polymorphism, due to the 6.8-kb intron, raises questions about the structure of the ancestral C4 gene that gave rise to the duplicated human C4A and C4B genes and the mouse Slp and C4 genes.

The six most common extended haplotypes among normal Caucasians (11) were included in the analysis of the 5′ C4 polymorphism (Table II). Although C4B gene length polymorphisms subdivide some groups complotyped by protein polymorphisms (Fig. 7), within specific extended haplotypes only a single type of C4B gene is observed. The same homogeneity has been observed among extended haplotypes by using the Bam HI/Xba I polymorphism of the C4 gamma-chain region (A. S. Whitehead, unpublished observations). Uniformity of extended haplotypes was also demonstrated by comparing the mixed lymphocyte response matched for the extended haplotype even in unrelated individuals (51). This high degree of functional and structural preservation is also supported by our analysis at the DNA level.

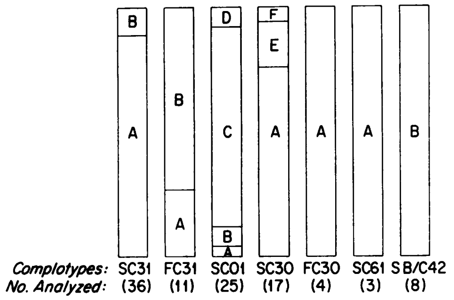

Figure 7. Distribution of different C4/21-hydroxylase gene structures among common complotypes. The Taq I patterns are listed in the structural order of the genes C4A/21-OHA/C4B/21-OHB ("–" indicates a deletion): (*A*) 7.0/3.2/6.0/3.7, C4B long; (*B*) 7.0/3.2/5.4/3.7, C4B short; (*C*) –/–/6.4/3.7, C4B short; (*D**) –/–/7.0/3.7, C4B long; (*E**) 7.0/–/–/3.7; F = 7.0/3.2/–/–; (* distinguished on the basis of C4 protein typing patterns).

Acknowledgments

This work was supported by grants from the National Institutes of Health (HD-17461, AI-21157, HL-29583, AI-14157, and CA-20531) and from the Deutsche Forschungsgemeinschaft (Ri 164/16-2). P. M. Schneider is a recipient of a fellowship from the German Academic Exchange Service (DAAD), and A. S. Whitehead is a recipient of a Helen Hay Whitney Foundation Fellowship (F488).

References

1. Allen, F. H. Jr. 1974. Linkage of HL-A and GBG. *Vox Sang.* 27: 382–384.

2. Fu, S. M., H. G. Kunkel, H. P. Brusman, F. H. Allen, and M. Fotino. 1974. Evidence for linkage between HL-A histocompatibility genes and those involved in the synthesis of the second component of complement. *J. Exp. Med.* 140:1108–1110.

3. Rittner, C., G. Hauptmann, H. Gross-Wilde, E. Grosshans, M. M. Tongio, and S. Mayer. 1975. Linkage between HL-A (major histocompatibility complex) and genes controlling the synthesis of the fourth component of complement. *In* Histocompatibility Testing 1975. F. Kissmeyer-Nielsen, editor. Munksgaard, Copenhagen. 945–954.

4. Dupont, B., S. E. Oberfield, E. M. Smithwick, T. D. Lee, and L. S. Levine. 1977. Close genetic linkage between HLA and congenital adrenal hyperplasia (21-hydroxylase deficiency). *Lancet.* ii:1309–1312.

5. Carroll, M. C., R. D. Campbell, D. R. Bentley, and R. R. Porter. 1984. A molecular map of the human major histocompatibility complex class III region linking complement genes C4, C2 and factor B. *Nature (Lond.).* 307:237–241.

6. White, P. C., D. Grossberger, B. J. Onufer, M. I. New, B. Dupont, and J. L. Strominger. 1985. Two genes encoding steroid 21-hydroxylase are located near the genes encoding the fourth component of complement in man. *Proc. Natl. Acad. Sci. USA.* 82:1089–1093.

7. Carroll, M. C., R. D. Campbell, and R. R. Porter. 1985. Mapping of steroid 21-hydroxylase genes adjacent to complement component C4 genes in HLA, the Major Histocompatibility Complex in man. *Proc. Natl. Acad. Sci. USA.* 82:521–525.

8. Whitehead, A. S., H. R. Colten, C. C. Chang, and R. Demars. 1985. Localization of the MHC-linked complement genes between HLA-B and HLA-DR by using HLA mutant cell lines. *J. Immunol.* 134:641–643.

9. Olaisen, B., R. Teisberg, R. Jonassen, E. Thorsby, and T. Gedde-Dahl. 1983. Gene order and gene distances in the HLA regions studied by the haplotype method. *Ann. Hum. Genet.* 47:285–292.

10. Alper, C. A., D. Raum, S. Karp, Z. L. Awdeh, and E. J. Yunis. 1983. Serum complement 'Supergenes' of the major histocompatibility complex in man (complotypes). *Vox Sang.* 45:62–67.

11. Awdeh, Z. L., D. Raum, E. J. Yunis, and C. A. Alper. 1983. Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. *Proc. Natl. Acad. Sci. USA.* 80:259–263.

12. Porter, R. R. 1983. Complement polymorphism, the major his-

tocompatibility complex and associated diseases. A speculation. *Mol. Biol. Med.* 1:161–168.

13. Rittner, C., and J. Bertrams. 1981. On the significance of C2, C4, and Factor B polymorphisms in disease. *Hum. Genet.* 56:235–247.

14. Awdeh, Z. L., D. Raum, and C. A. Alper. 1979. Genetic polymorphism of human complement C4 and detection of heterozygotes. *Nature (Lond.).* 282:205–207.

15. Fielder, A. H. L., M. J. Walport, J. R. Batchelor, R. I. Rynes, C. M. Black, I. A. Dodi, and G. R. V. Huges. 1983. A family study of the MHC of patients with SLE. Null alleles of C4A and C4B may determine the disease. *Br. Med. J.* 186:425–428.

16. Rittner, C., P. Kuehnl, C. M. Black, S. Pereira, and K. I. Welsh. 1984. Scleroderma: Possible association with the C4 system—a progress report. *In* Histocompatibility Testing 1984. E. D. Albert, editor. Springer-Verlag, Berlin. 394–397.

17. Rittner, C., E. M. M. Meier, B. Stradmann, C. M. Giles, R. Koechling, E. Mollenhauer, and H. W. Kreth. 1984. Partial C4 deficiency in subacute sclerosing panencephalitis. *Immunogenetics.* 20:407–415.

18. Dupont, B., M. S. Pollack, L. S. Levine, G. S. O'Neill, B. R. Hawkins, and M. I. New. 1980. Joint report. Congenital adrenal hyperplasia. *In* Histocompatibility Testing 1980. P. I. Terasaki, editor. UCLA Tissue Typing Lab, Los Angeles. 693–706.

19. Fleischnick, E., D. Raum, S. M. Alosco, P. S. Gerald, E. J. Yunis, Z. L. Awdeh, J. Granados, J. F. Crigler Jr., C. M. Giles, and C. A. Alper. 1983. Extended MHC haplotypes in 21-hydroxylase-deficiency congenital adrenal hyperplasia: shared genotypes in unrelated patients. *Lancet.* ii: 152–156.

20. White, P. C., M. I. New, and B. Dupont. 1984. HLA-linked congenital adrenal hyperplasia results from a defective gene encoding a cytochrome P-450 specific for steroid 21-hydroxylation. *Proc. Natl. Acad. Sci. USA.* 81:7505–7509.

21. Palsdottir, A., S. J. Cross, J. H. Edwards, and M. C. Carroll. 1983. Correlation between a restriction fragment length polymorphism and C4A6 protein. *Nature (Lond.).* 306:615–616.

22. Whitehead, A. S., D. E. Woods, E. Fleischnick, J. E. Chin, E. J. Yunis, A. J. Katz, P. S. Gerald, C. A. Alper, and H. R. Colten. 1984. DNA polymorphism of the C4 genes: A new marker for analysis of the Major Histocompatibility Complex. *N. Engl. J. Med.* 310:88–91.

23. Belt, K. T., M. C. Carroll, and R. R. Porter. 1984. The structural basis of the multiple forms of human complement component C4. *Cell.* 36:907–914.

24. Belt, K. T., C. Y. Yu, M. C. Carroll, and R. R. Porter. 1985. Polymorphism of the human complement component C4. *Immunogenetics.* 21:173–180.

25. Carroll, M. C., A. Palsdottir, K. T. Belt, and R. R. Porter. 1985. Deletion of complement C4 and steroid 21-hydroxylase genes in the HLA class III region. *EMBO J.* 4:2547–2552.

26. Prentice, H. L., P. M. Schneider, and J. L. Strominger. 1986. C4B gene polymorphism detected in human cosmid clone. *Immunogenetics.* 23:274–276.

27. Alper, C. A., T. Boenisch, and L. Watson. 1972. Genetic polymorphism in human glycin-rich beta-glycoprotein. *J. Exp. Med.* 135: 68–80.

28. Alper, C. A. 1976. Inherited structural polymorphism in human C2: evidence for genetic linkage between C2 and B2. *J. Exp. Med.* 144: 1111–1115.

29. Awdeh, Z. L., and C. A. Alper. 1980. Inherited structural polymorphism of the fourth component of human complement. *Proc. Natl. Acad. Sci. USA.* 77:3576–3580.

30. Roos, M. H., E. Mollenhauer, P. Demant, and C. Rittner. 1982. A molecular basis for the two locus model of human complement C4. *Nature (Lond.).* 298:854–856.

31. Kissmeyer-Nielsen, F., and K. E. Kjerbye. 1967. Lymphocytotoxic microtechnique. Purification of lymphocytes by flotation. *In* Histocompatibility Testing 1967. E. S. Curtoni, P. L. Mattiuz, and R. M. Tosi, editors. Munksgaard, Copenhagen. 381–383.

32. Ray, J. G. Jr., D. B. Hare, and D. E. Kayhoe, editors. 1973. NIAID manual of tissue typing techniques. NIH No. 74-545. Department of Health, Education and Welfare, Bethesda, MD.

33. Gross-Bellard, M., P. Oudet, and P. Chambon. 1973. Isolation of high molecular weight DNA from mammalian cells. *Eur. J. Biochem.* 36:32–38.

34. Wahl, G. M., M. Stern, and G. R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzoloxymethalpaper and rapid hybridization using dextran sulfate. *Proc. Natl. Acad. Sci. USA.* 76:3683–3687.

35. Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.

36. Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to hogh specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237–251.

37. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor, NY. 1–545.

38. Carroll, M. C., and R. R. Porter. 1983. Cloning of a human complement C4 gene. *Proc. Natl. Acad. Sci. USA.* 80:264–267.

39. Carroll, M. C., K. T. Belt, A. Palsdottir, and Y. Yu. 1985. Molecular genetics of the fourth component of human complement and steroid 21-hydroxylase. *Immunol. Rev.* 87:39–60.

40. Isenman, D., and J. R. Young. 1984. The molecular basis for the difference in immune hemolysis activity of the Chido and Rodgers isotypes of the human complement component C4. *J. Immunol.* 132: 3019–3027.

41. Law, S. K. A., A. W. Dodds, and R. R. Porter. 1984. A comparison of the properties of two classes, C4A and C4B, of the human complement component C4. *EMBO J.* 3:1819–1823.

42. Dodds, A. W., S. K. A. Law, and R. R. Porter. 1985. The origin of the very variable haemolytic activities of the common human complement C4 allotypes including C4-A6. *EMBO J.* 4:2239–2244.

43. Glass, D., D. Raum, D. Gibson, J. S. Stillman, and P. H. Schur. 1976. Inherited deficiency of the second component of complement. Rheumatic disease associations. *J. Clin. Invest.* 58:853–861.

44. Schneider, P. M., C. Rittner, M. C. Carroll, and H. R. Colten. 1985. Genomic structure of the C4 and 21-OH loci in a family with a C4A(Ch+) gene product. *Complement.* 2:71. (Abstr.)

45. New, M. I. 1985. Clinical and endocrinological aspects of 21-hydroxylase deficiency. *In* Congenital Adrenal Hyperplasia. M. I. New, editor. *Ann. NY Acad. Sci.* 458:1–27.

46. Raum, D., Z. Awdeh, J. Anderson, L. Strong, J. Granados, L. Pevan, E. Giblett, E. J. Yunis, and C. A. Alper. 1984. Human C4 haplotypes with duplicated C4A or C4B. *Am. J. Hum. Genet.* 36:72–79.

47. Weatherall, D. J., and J. B. Clegg. 1982. Thalassemia revisited. *Cell.* 29:7–9.

48. Bentley, D. R., and R. D. Campbell. 1985. Primary structure of C2 and relationship to other components of the complement system. *Complement.* 2:9. (Abstr.)

49. Cole, F. S., A. S. Whitehead, H. S. Auerbach, T. Lint, H. J. Zeitz, P. Kilbridge, and H. R. Colten. 1985. The molecular basis for genetic deficiency of the second component of human complement. *N. Engl. J. Med.* 313:11–16.

50. Uring-Lambert, B., C. M. Giles, J. Goetz, M. M. Tongio, S. Mayer, and G. Hauptmann. 1984. C4 haplotypes with duplicated C4A or C4B: Frequency and associations with BF, C2, and HLA-A, B, C, DR alleles with special reference to the duplication C4B1,2. *In* Histocompatibility Testing 1984. E. D. Albert, editor. Springer-Verlag, Berlin. 604–608.

51. Awdeh, Z. L., E. Eynon, R. Stein, C. A. Alper, S. M. Alosco, and E. J. Yunis. 1985. Unrelated individuals matched for MHC extended haplotypes and HLA-identical siblings show comparable responses in mixed lymphocyte culture. *Lancet.* ii:853–855.