17. N. M. Gough, D. J. Kemp, B. M. Tyler, J. M. Adams, S. Cory, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 554 (1980); J. G. Seidman and P. Leder, *Nature (London)* **276**, 790 (1978); C. Brack and S. Tonegawa, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5652 (1977).
18. P. Early, H. Huang, M. Davis, K. Calame, L. Hood, *Cell* **19**, 981 (1980); E. E. Max, J. G. Seidman, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3450 (1979); H. Sakano, K. Hüppi, G. Heinrich, S. Tonegawa, *Nature (London)* **280**, 288 (1979).
19. P. Early, J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, L. Hood, *Cell* **20**, 313 (1980).
20. J. Rogers, P. Early, C. Carter, K. Calame, M. Bond, L. Hood, R. Wall, *ibid.*, p. 303.
21. M. Kehry, C. Sibley, C. Fuhrman, J. Schilling, L. Hood, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2932 (1979); M. Kehry *et al.*, in preparation.
22. V. T. Oi, V. M. Bryan, L. A. Herzenberg, L. A. Herzenberg, *J. Exp. Med.* **151**, 1260 (1980).
23. R. M. E. Parkhouse and M. D. Cooper, *Immunol. Rev.* **37**, 105 (1977).
24. H. L. Spiegelberg, *Nature (London)* **254**, 723 (1975).
25. L.-C. Lin and F. W. Putnam, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6572 (1979).
26. F. D. Finkelman, V. L. Woods, A. Berning, I. Scher, *J. Immunol.* **123**, 1253 (1979); R. Sitia, G. Corte, M. Ferraini, A. Bargellesi, *Eur. J. Immunol.* **7**, 503 (1979); J. Radl, P. VandenBerg, C. M. Jol-Vanderzijde, *J. Immunol.* **124**, 2513 (1979).
27. A. Bargellesi, G. Corte, E. Cosulich, M. Ferrarini. *Eur. J. Immunol.* **9**, 490 (1979).
28. B. Lisowska-Bernstein and P. Vassalli, in *Membrane Receptors of Lymphocytes*, M. Seligmann, J. L. Preud'homme, F. M. Kourilsky, Eds. (North-Holland, Amsterdam, 1975), p. 39; U. Melcher and J. W. Uhr, *J. Immunol.* **116**, 409 (1976); E. R. Abney and R. M. E. Parkhouse, *Nature (London)* **252**, 600 (1974).
29. J. W. Goding and L. A. Herzenberg, *J. Immunol.* **124**, 2540 (1980).
30. J. W. Goding, *ibid.*, p. 2082; S. W. Kessler, V. L. Woods, F. D. Finkelman, I. Sher, *ibid.* **123**, 2772 (1979).
31. L. Eidels, *ibid.* **123**, 896 (1979).
32. M. Mescher and R. R. Pollock, *J. Immunol.* **123**, 1155 (1979); R. R. Pollock, M. E. Dorf, M. F. Mescher, *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 4256 (1980); R. R. Pollock and M. F. Mescher, *J. Immunol.* **124**, 1668 (1980).
33. C.-P. Liu, P. W. Tucker, J. F. Mushinski, F. R. Blattner, *Science* **209**, 1348 (1980).
34. F. D. Finkelman, S. W. Kessler, J. F. Mushinski, M. Potter, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **39**, 481 (1980); S. W. Kessler, J. F. Mushinski, M. Potter, F. D. Finkelman, *ibid.*, p. 1055.
35. V. T. Oi, P. P. Jones, J. W. Goding, L. A. Herzenberg, L. A. Herzenberg, *Curr. Top. Microbiol. Immunol.* **81**, 115 (1978).
36. J. F. Mushinski, F. R. Blattner, J. D. Owens, F. D. Finkelman, S. W. Kessler, L. Fitzmaurice, M. Potter, P. W. Tucker, in preparation.
37. T. Maniatis, S. G. Kee, A. Efstratiadis, F. C. Kafatos, *Cell* **8**, 163 (1976).
38. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
39. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
40. Abbreviations: Ala, alanine; Cys, cysteine; Asp, aspartic acid; Glu, glutamic acid; Phe, phenylalanine; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Asn, asparagine; Pro, proline; Gln, glutamine; Arg, arginine; Ser, serine; Thr, threonine; Val, valine; Trp, tryptophan; and Tyr, tyrosine.
41. E. A. Kabat, T. T. Wu, H. Bilofsky, Eds., *Sequences of Immunoglobulin Chains* (NIH Publ. 80-2008, 1979), p. 138.
42. M. R. Lerner, J. A. Boyle, S. M. Mount, S. L. Wolin, J. A. Steitz, *Nature (London)* **283**, 220 (1980); J. Rogers and R. Wall, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1877 (1980).
43. T. Honjo, M. Obata, Y. Yamawaki-Kataoku, Y. Kataoku, T. Kawakami, H. Takahashi, Y. Mano, *Cell* **18**, 559 (1979); P. W. Tucker, K. Marcu, N. Newell, J. R. Richards, F. R. Blattner, *Science* **206**, 1303 (1979).
44. L.-C. Lin and F. W. Putnam, personal communication.
45. M. Tomita and V. T. Marchesi, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2964 (1975).
46. D. D. Jones, *J. Theor. Biol.* **50**, 167 (1975).
47. G. Guidotti, *J. Supramol. Struct.* **7**, 489 (1977); C. Tanford and J. A. Reynolds, *Biochim. Biophys. Acta* **457**, 133 (1976).
48. J. P. Segrest and R. J. Feldmann, *J. Mol. Biol.* **81**, 853 (1974).
49. P. W. Robbins, S. C. Hubbard, S. J. Turco, D. F. Wirth, *Cell* **12**, 893 (1977).
50. R. Breathnach, C. Benoist, K. O'Hare, F. Gannon, P. Chambon, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4853 (1978).
51. N. J. Proudfoot and G. G. Brownlee, *Nature (London)* **263**, 211 (1976).
52. S. Rudikoff, personal communication.
53. R. L. Jilka and S. Pestka, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5692 (1977).
54. S. W. Kessler and F. D. Finkelman, personal communication.
55. E. C. Franklin and B. Frangioni, in *Contemporary Topics in Molecular Immunology*, F. P. Inman and W. J. Mandy, Eds. (Plenum, New York, 1976), p. 89; E. C. Franklin, personal communication.
56. L. Forni and A. Coutinho, *Nature (London)* **273**, 304 (1978).
57. R. Sitia, E. M. Rabellino, M. Sockell, U. Hammerling, *J. Immunol.*, in press.
58. E. S. Vitetta and J. W. Uhr, *Science* **189**, 964 (1975).
59. A. Bourgois, E. R. Abney, R. M. E. Parkhouse, *Eur. J. Immunol.* **7**, 210 (1977).
60. E. Pure and E. S. Vitetta, *J. Immunol.* **125**, 420 (1980).
61. *Note added in proof*: Partial amino acid sequence data on proteolytic fragments derived from a shortened δ chain secreted by rat IgD myeloma were recently reported [G. Alcaraz, A. Bourgois, A. Moulin, H. Bazin, M. Fougereau, *Ann. Immunol. (Inst. Pasteur)* **131C**, 363 (1980)]. Two fragments can be matched with high homology to corresponding regions in our mouse sequence, and the authors postulated that the shortened length of the rat δ chain resulted from a partial deletion of the Cδ2 domain. However, this deduction was based, at least in part, on erroneous homology alignment, and thus conclusive cross-species comparison must await additional refinement of the rat sequence.
62. K. W. Olsen, *Biochim. Biophys. Acta* **622**, 259 (1980).
63. H. B. Bull and K. Breese, *Arch. Biochem. Biophys.* **161**, 665 (1974).
64. Y. Nozaki and C. Tanford, *J. Biol. Chem.* **246**, 2211 (1971).
65. P. Y. Chou and G. B. Fasman, *Biochem.* **13**, 211 (1974).
66. We thank M. Fiant for electron microscopy, H.-L. Cheng and R. Robinson for expert technical assistance, W. Fiske and E. LaLuzene for help with the manuscript, S. Rudikoff for amino acid sequence determinations, L. Fitzmaurice, V. Oi, K. Olsen, J. Schroeder, J. Uhr, and F. Finkelman for helpful discussions, and M. Potter for enthusiastic encouragement and for providing the plasmacytomas that made this work possible. This work was done under the applicable NIH recombinant DNA guidelines and was supported by NIH grants AI 16547 (P.W.T.), GM 21812 (F.R.B.), and GM 06768 (C.P.L.).

16 July 1980

DNA Sequences Mediating Class Switching in α-Immunoglobulins

Mark M. Davis, Stuart K. Kim, Leroy E. Hood

The antibody molecule is a set of discrete molecular domains that carry out two general types of functions. The variable (V) domain binds antigen and the constant (C) domains trigger effector functions such as complement fixation. The V and C domains arise from the interactions of two different polypeptides, light (L) and heavy (H), which in turn are encoded by a series of discrete gene segments—V_L, J_L (joining), and C_L encoding the light chains and V_H, D (diversity), J_H, and C_H encoding the heavy chains (*1–4*). During the differentiation of antibody-producing or B cells, two distinct types of DNA rearrangements of these gene segments occur (*4, 5*). One type generates the V_L gene by direct joining of the V_L and J_L gene segments and the V_H gene by direct joining of the V_H, D, and J_H gene segments. These DNA rearrangements are termed V-J or V-D-J joining and they are, in part, responsible for the generation of antigen-binding diversity in V domains.

A second type of DNA rearrangement, termed C_H switching, allows important flexibility in the use of a given antigen-binding site. At an early stage of B cell differentiation, an individual B cell initially expresses immunoglobulin M (IgM) molecules with a single V domain (V_L-V_H combination) (*6, 7*). Later, this B cell or its clonal progeny may express another immunoglobulin class while continuing to express the same V domain (*8*). Since the class of immunoglobulin is determined by the C_H region (C_μ, C_γ, and C_α determining IgM, IgG, and IgA, respectively), the B cell must shift from the expression of another C_μ gene to the expression of another C_H gene during differentiation. Thus, C_H switching associates a particular antigen-binding specificity, the V domain, with a series of different effector functions encoded by the various C_H regions.

Two types of experiments have provided insights into the mechanism of C_H switching. First, Honjo and Kataoka (*9*)

M. M. Davis and S. K. Kim are graduate students and L. E. Hood is professor of biology at the Division of Biology, California Institute of Technology, Pasadena 91125.

Summary. Immunoglobulin class switching involves specific DNA rearrangements of the gene segments coding for heavy chain constant regions (C_H) during B lymphocyte differentiation. In two different cases of C_μ to $C\alpha$ switching examined here (T15 and M603) and one taken from the literature (MC101), three different sites on the 5′ side of C_μ and three different sites on the 5′ side of C_α are joined together in the process of C_H switching. The sequences surrounding the three germ-line C_α sites of recombination are highly conserved blocks of 30 nucleotides that may serve as recognition sequences for C_H switching to the C_α gene. This putative recognition sequence is repeated 17 times in approximately 1400 nucleotides of the germ-line C_α 5′ flanking sequence. The lack of homology between this C_α sequence and sequences reported for the $C_{\gamma 1}$ and $C_{\gamma 2b}$ switch sites suggests that heavy chain switching is mediated by class-specific recognition sequences and, presumably, class-specific regulatory mechanisms. In addition, it appears that in one example (MC101) C_H switching progressed from C_μ to C_α to $C_{\gamma 1}$. This switching pathway may present difficulties for the simple deletional model of C_H switching.

have employed hybridization kinetics to determine the numbers of C_H genes in mouse myeloma tumors that produced different immunoglobulin classes. Their results suggest that the V_H gene is separated from one C_H gene and combined with a second by a deletion of the intervening DNA between the V_H gene and the second C_H gene. From these data a heavy chain gene order of 5′-C_μ-$C_{\gamma 3}$-$C_{\gamma 1}$-$C_{\gamma 2b}$-$C_{\gamma 2a}$-C_α-3′ was suggested. Recent experiments (*10*) with cloned probes and Southern blot analyses generally support the gene order and deletional mechanism proposed by Honjo and Kataoka (*9*). Second, we examined the rearranged (expressed) α gene in an IgA-producing myeloma tumor (M603) and obtained direct evidence that DNA rearrangement mediates C_H switching (*5*). The rearranged M603 α gene is composed of three distinct germ-line gene segments—V_H, J_H with 5′ C_μ flanking sequences, and C_α with its flanking sequences. This tripartite structure of the rearranged α gene suggests that the V_H gene was initially associated with the C_μ gene by DNA rearrangement through V-D-J joining and was expressed as a μ chain in the IgM molecule. A subsequent DNA rearrangement could then replace the C_μ gene with the C_α gene by linking together 5′ C_μ and C_α flanking sequences. The point at which the flanking C_μ and C_α sequences join in the rearranged gene and the corresponding breakpoints on the germ-line DNA's are termed the switch (S) sites. Subsequently, other laboratories have obtained similar evidence for C_μ to C_γ switches in rearranged γ1 (*11*) and γ2b (*12, 13*) genes.

Three Examples of IgM → IgA Switching

To investigate the molecular mechanisms underlying C_H switching, we sequenced the switch sites of two rearranged α genes and compared these rearranged switch sequences with their germ-line counterparts in the 5′ flanking sequences of the C_μ and C_α genes. To this end, we have constructed genomic libraries (*14*) from the DNA's of M603 (*15, 16*) and an additional IgA-producing myeloma tumor, T15, and isolated the rearranged α genes. Homologies between clones corresponding to the rearranged α genes of T15 and M603 are depicted in Fig. 1 and compared with their germ-line C_μ and C_α counterparts [obtained from a genomic library of mouse sperm DNA (*5*)]. These homologies were established by detailed restriction enzyme analysis (data not shown). Both rearranged α genes exhibit the tripartite structure of V_H, J_H with C_μ flanking sequences, and C_α gene segments. The size of the intervening sequence between the V_H and C_α coding regions is substantially different in these two cases—5.4 versus 6.8 kilobase pairs (kbp) (Fig. 1). Since each of the two V_H gene segments is joined to the same J_H gene segment (*17*), the variation in size in the intervening sequences between the V and C_α coding regions may be the result of different C_H switching sites in rearranged C_α genes. In addition, we found that a rearranged γ1 gene from myeloma MC101, whose sequence was reported by Honjo and colleagues (*11*), contains a 500-nucleotide region of C_α flanking sequence between C_μ- and $C_{\gamma 1}$-derived sequences (Fig. 1). The evidence for this supposition is that this 500-nucleotide sequence has been localized solely in a region on the 5′ side of the C_α gene by Southern (*18*) blotting analyses and restriction mapping with fragments containing all or part of this region used as probes (*11, 19*). Therefore, this fragment is apparently represented just once in the genome and must have been derived from flanking sequences on the 5′ side of the C_α gene. Furthermore, DNA sequence anal-

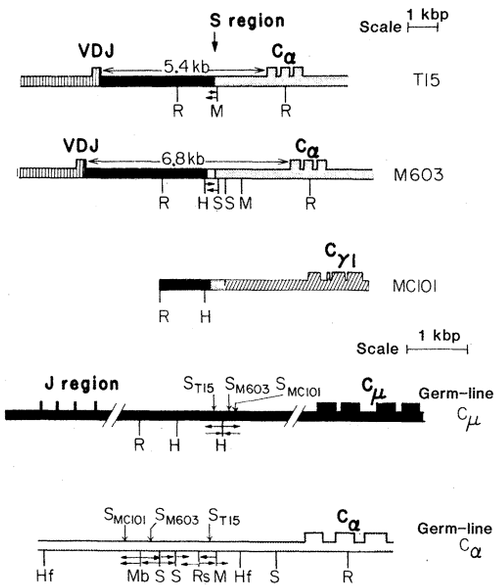

Fig. 1. Clones containing rearranged α heavy chain genes and germ-line C_μ and C_α genes. The rearranged T15 and M603 genes were isolated from genomic libraries of their respective myeloma tumors, and MC101 was isolated as described (*11*). Clones containing the germ-line C_μ and C_α genes were isolated from libraries of mouse sperm DNA (*5*). Raised boxes denote coding regions. The shaded areas indicate homologies with germ-line genes and their flanking sequences: ▥, V_H and flanking region; ■, C_μ and flanking region; ▨, C_α and flanking regions; and □, uncertain origin, probably C_μ derived (see text) ▨, $C_{\gamma 1}$ and flanking regions. Homologies were determined by detailed restriction mapping. Restriction enzyme sites are denoted as follows: *Hf*, Hinf I; *H*, Hind III; *M*, Msp I; *R*, Eco RI; *Rs*, Rsa I; *S*, Sau 3a; and *Mb*, Mbo II. The position of the J_H regions was determined from experiments described in (*4, 13*). The M603 clone was isolated and characterized as previously described (*5, 15, 16*). The T15 library was made from Eco RI partially digested DNA as described (*5, 14*). Recombinant phage were screened by the procedure of Benton and Davis (*23*) with ^{32}P nick-translated $V_H + C_\alpha$ complementary DNA clones used as probes (*5, 14*). Southern (*18*) blot analysis of Eco RI-digested myeloma DNA indicated the identity of the central Eco RI fragment of the clones shown here with the fragment of the same size found in the T15 genome (data not shown), indicating that the clones isolated are representative of their genome of origin. This analysis has been previously reported for M603 (*16*). DNA sequence analysis was performed by the procedures of Maxam and Gilbert (*24*). Arrows originating from a restriction site indicate 5′ end labeling with [γ-^{32}P]adenosine triphosphate, and arrows ending in a restricting site indicate 3′ end labeling with [α-^{32}P]deoxynucleotide triphosphates.

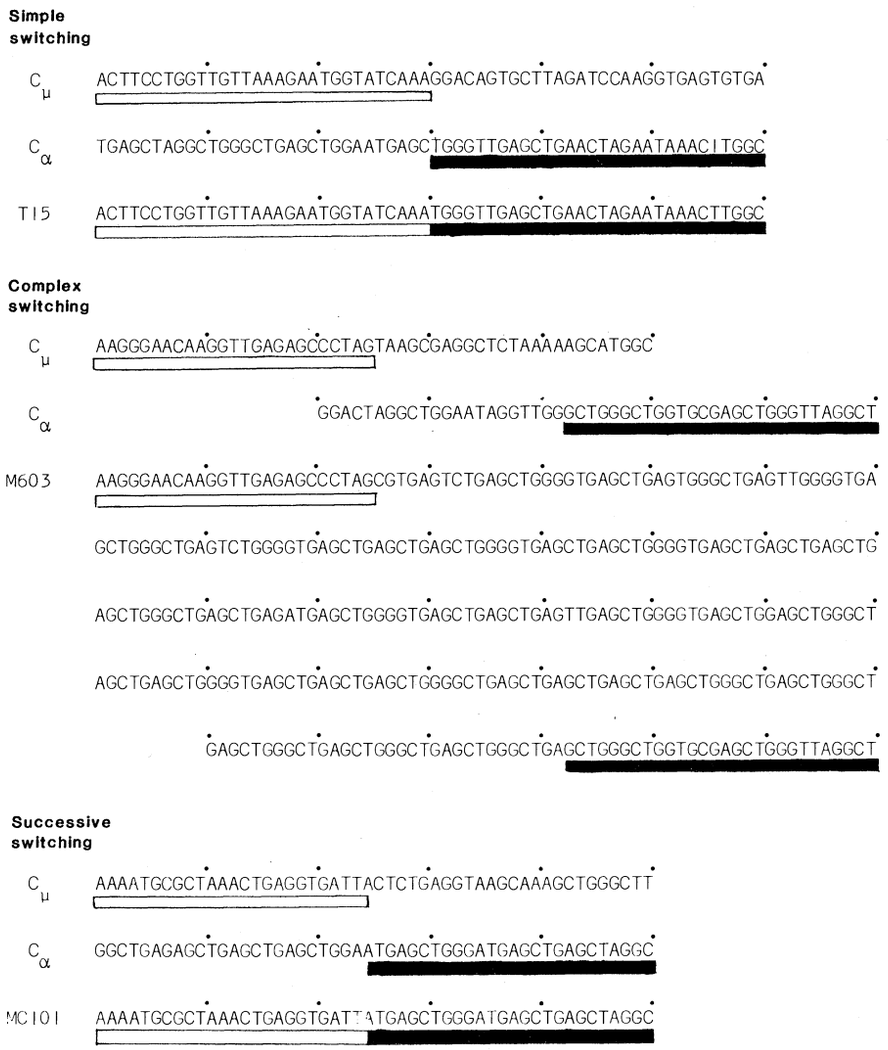

Fig. 2. Switch sequences for the rearranged T15, M603, and MC101 genes and their germ-line C_μ and C_α counterparts. $C_{\gamma 1}$ corresponding sequences for MC101 have been reported by Kataoka *et al.* (*11*). Sequences surrounding C_H switch sites are shown in 5′ → 3′ orientation. Underlining indicates sequence identity of unrearranged C_μ (□) and C_α (■) flanking sequences with their counterparts in the individual rearranged genes. Dots indicate ten nucleotide spacings. DNA sequence analysis was performed as described (*24*).

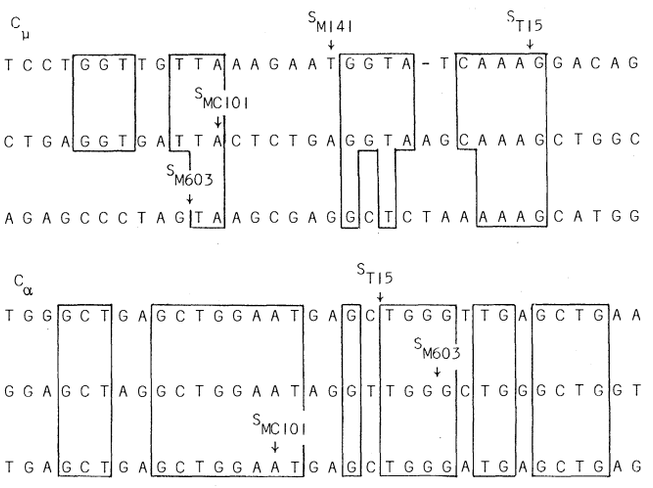

Fig. 3. Comparison of the DNA sequences of the germ-line C_μ and C_α switch sites for the rearranged T15, M603, and MC101 genes. Sequences obtained from C_μ and C_α flanking regions were aligned for maximum homology around their C_H switch sites. Boxes indicate nucleotide identities and dashes indicate a gap introduced for homology alignment. Arrows indicate breakpoints for the switch sites of each rearranged gene. Where the switch breakpoint is ambiguous, arrows appear directly above the nucleotide. Also indicated is the C_μ switch site for a γ2b producer, M141 (*12*, *13*). A consensus sequence for at least some cases of C_μ switching might therefore be GGTNATTANNNNNNGGTANNCAAAG, which does not occur elsewhere in any of the 1900 nucleotides of C_μ flanking region sequenced (*13*). It has been suggested previously that elements of this C_μ homology region may play a role in C_H switching (*12*, *13*). A consensus sequence for C_α switching derived from the examples shown here would be PGTCPPGCTGGAATPPGYTGGGNTGPGCTG.

yses of the corresponding germ-line C_α sequence show virtual identity (approximately 95 percent) with the sequence found between the C_μ and $C_{\gamma 1}$ flanking sequences in the γ1 gene of MC101 (Fig. 4A) (*20*). Thus we feel the γ1 gene of MC101 is composed of several distinct germ-line sequences: a V_H gene (*21*), a flanking sequence for the C_μ gene, a flanking sequence for the C_α gene, and the $C_{\gamma 1}$ with its flanking sequences.

The arrows in Fig. 1 indicate the regions analyzed by DNA sequence analysis in our laboratory. The DNA sequences of the MC101 S region (*11*) and the rearranged M603 and T15 α genes and their germ-line C_μ and C_α counterparts are shown in Fig. 2.

Examination of the rearranged switch sequences indicates that all three examples juxtapose C_μ- and C_α-derived sequences. However, each switch site seems distinct from the others. Furthermore, the arrangement of sequences in these genes suggest that at least three distinct types of switching may occur. We denote these categories "simple," "complex," and "successive."

In the simple category (T15) the C_μ flanking sequence joins directly to the C_α flanking sequence. Similarly, a γ2b gene (M141) has been found in which the C_μ flanking sequence joins to that of $C_{\gamma 2b}$ (*12*, *13*). In the complex category (M603), a short sequence of 287 base pairs (bp), is interposed between the C_μ and C_α flanking sequences. This sequence appears to derive from a region on the 3′ side of S_{M603} on the C_μ gene (Fig. 1). Probes containing this sequence hybridize strongly to restriction fragments containing or adjacent to the C_μ gene in Southern blotting analyses (data not shown). The region of hybridization corresponds to the 1.5- to 2.5-kbp region on the 5′ side of the C_μ gene (*5*) which deletes spontaneously upon cloning and hence is not present in our C_μ-containing clones. Thus the complex category may be explained by two distinct deletions: a C_H switch and a deletion within the C_μ flanking sequence. This deletion of the C_μ flanking sequence (at S_{M603}) does not appear to be a random event in that another α-producing tumor line, M167, switches at exactly the same point adjacent to the C_μ gene. However, the complex category reflects DNA deletions seen to date only in myeloma cells and, accordingly, may or may not be biologically significant. In the third category, successive switching, the γ1 gene of MC101 contains C_μ, C_α, and $C_{\gamma 1}$ flanking sequences between the V_H and $C_{\gamma 1}$ gene segments. Therefore, it appears to have switched twice, once from C_μ

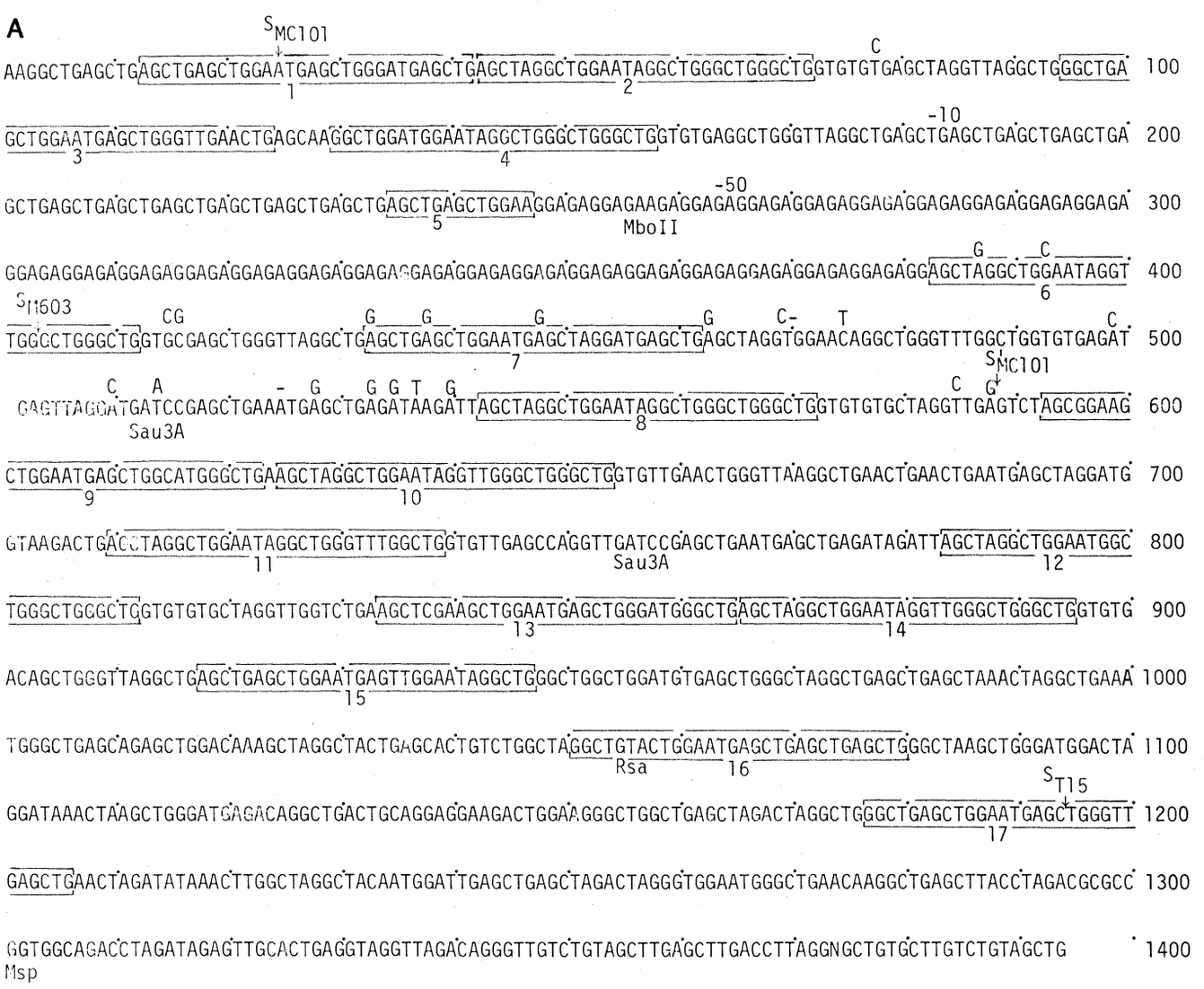

Fig. 4. (A) DNA sequence in the region of the germ-line C_α switch sites. Restriction sites are indicated; S_α homology regions are boxed. S5 appears to be a homology region which has been interrupted by a stretch of tandem repeats. Arrows are as described in Fig. 3. Also shown is a comparison of the MC101 C_α-derived sequence (*11*) with that of the germ-line C_α sequence. The MC101 sequence is bracketed by S_{MC101} and S'_{MC101} and differences from the corresponding germ-line sequence are indicated above, either as base changes or gaps (—) or (−10, and so on). Precise localization of gaps in the tandem repeats GAGGA and GAGCT is not possible, since the deletions are symmetrical. It is interesting that virtually all of the sequence differences are on the 3′ half of the MC101 sequence, with the tandem repeats acting as a "border" between highly conserved and mutated regions. (B) Schematic diagram of the location of 17 repeated sequences of 30 nucleotides adjacent to the germ-line C_α gene. Arrows indicate which repeats are used in the three rearranged genes described in this article. (C) DNA sequences of 17 repeats found interspersed in germ-line 5′ flanking C_α sequence. Boxes indicate nucleotide identities and a dash indicates a gap.

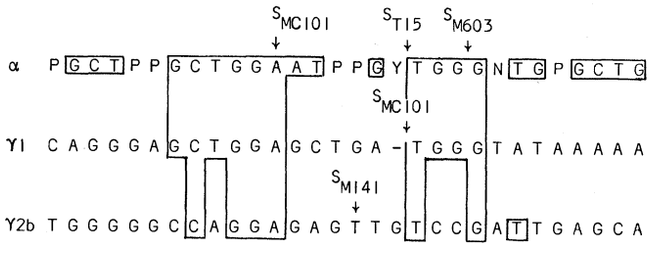

Fig. 5. Germ-line sequences; comparison of the switch sequences from germ-line 5′ flanking sequences for the C_α (see Fig. 3), $C_{\gamma 1}$ (*11*), and $C_{\gamma 2b}$ (*12, 13*) genes. Boxes in the germ-line C_α sequence denote the conserved bases of the 30-nucleotide S_α sequence (Fig. 3). Boxes in the germ-line γ1 and γ2b sequences denote nucleotide identities to the germ-line S_α sequence; a dash indicates a gap.

to $C_{\gamma 1}$ subsequently from C_α to $C_{\gamma 1}$.

The DNA sequence data in Fig. 2 show that the three rearranged genes employ three germ-line C_μ switch sites up to 300 bp apart and three germ-line C_α switch sites up to 1350 bp apart. The locations of these germ-line switch sites, S_{T15}, S_{M603}, and S_{MC101}, are depicted in Fig. 1. Thus multiple sites exist for C_α switching in sequences 5′ to both the germ-line C_μ and C_α genes.

Recognition Sequences for Class Switching

The DNA sequences involved in V-D-J joining are quite distinct from those implicated in C_H switching. The inverted repeat CACAGTG or CACTGTG (C, cytosine; A, adenine; T, thymine; G, guanine) occurs at the 3′ end of antibody V gene segments and at the 5′ end of the J gene segments (*2, 4, 22*). This inverted repeat is believed to be a recognition sequence that mediates the juxtaposition of the V, D, and J gene segments to allow subsequent joining by site-specific recombination [for a proposed mechanism, see (*4*)]. This inverted repeat is not found in the flanking regions surrounding any of the C_α switch sites. This sequence also is missing from the switch sites for a γ1 (*11*) and a γ2b gene (*12, 13*). Therefore, C_H switching and V-J joining employ distinct mechanisms for DNA rearrangement.

In an effort to determine which sequences are important in C_H switching, we compared the sequences of the three C_μ sites and the three C_α sites (Fig. 3). The germ-line C_μ switch sites of MC101 and T15 share significant homology (15 of 25 nucleotides), although 16 nucleotides separate the actual switch points. We believe that these homologies are significant and may represent general sequence requirements for C_H switching adjacent to the C_μ gene. Neither the MC101 nor the T15 C_μ sites share any homology with the C_μ site of M603. The switch site of a γ2b-producing tumor, M141 (*12, 13*), also depicted in Fig. 3, is nine nucleotides away from that of T15 and may indicate that γ2b and α switching can use the same recognition sequence adjacent to C_μ.

The sequences around each C_α switch site are even more highly conserved. Each germ-line S_α site occurs within a block of 30 conserved nucleotides (Fig. 3). Twenty-two bases are identical and seven of the remaining eight nucleotides are conserved with regard to type of base (purine-purine or pyrimidine-pyrimidine substitutions). The three points of recombination differ within each of these conserved sequences. Since the first three of the rearranged α genes examined switched at distinct C_α sites, we reasoned that there must be additional S_α sites. We determined the DNA sequence of some 1400 nucleotides in the region of these germ-line S_α sites (Fig. 4A) and found a total of 17 S_α-like sequences (Fig. 4B). These sequences are very similar to each other (boxed regions in Fig. 4C). Since these repeated sequences represent 500 of 1400 nucleotides in the region analyzed, the C_H switching into three of these repeated sequences does not appear to represent random DNA rearrangement. We suggest that most of these repeats are potential germ-line S_α sites.

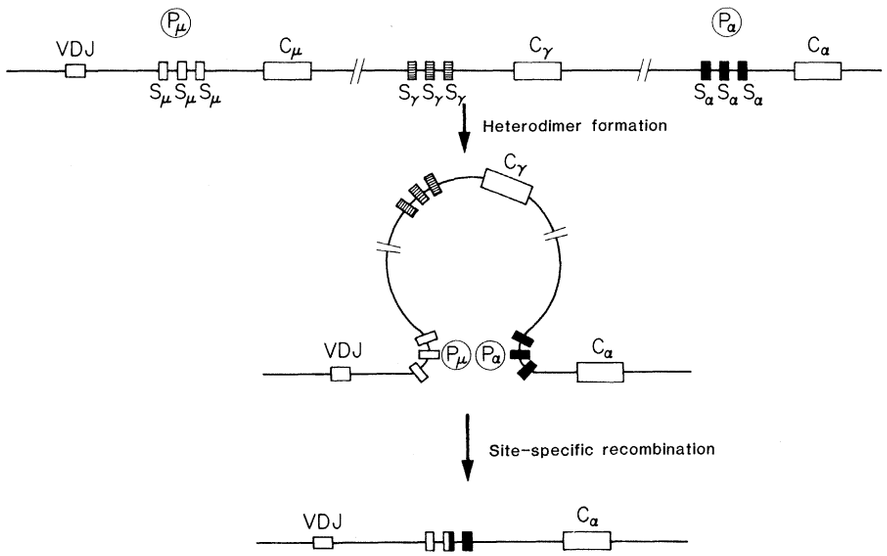

Fig. 6. Model for class-specific regulation of C_H switching (see text). In this scheme the small boxes represent recognition (S) sequences that bind to switch proteins to mediate C_H switching; the circle represents switching proteins, and the large boxes represent coding regions. This model depicts a C_H gene order of C_μ, C_γ, C_α (see text), but this is not a requirement.

Evidence for Class-Specific Regulation of C_H Switching

The data presented in this article, taken together with sequence data on γ2b (*12, 13*) and γ1 (*11*) switch sites from the literature, lead to several important inferences about the mechanism of C_H switching.

The germ-line sequences for C_H switching appear to be class-specific with regard to $C_{\gamma 1}$, C_α, and $C_{\gamma 2b}$ switch sites. The germ-line $C_{\gamma 1}$, C_α, and $C_{\gamma 2b}$ switch sequences are compared in Fig. 5. The γ1 sequence is identical to the prototype S_α sequence in 10 of 22 bp and the γ2b sequence is identical for 7 of 22 nucleotides (only 3 of which are contiguous). Thus, various germ-line S_α sequences are far more similar to one another than to the germ-line $S_{\gamma 1}$ or $S_{\gamma 2b}$ sequences (Fig. 3). One explanation for these sequence differences is that C_H switching is mediated by class-specific recognition sequences.

Since the germ-line S_α and S_μ sequences are not homologous (Fig. 3), homologous recombination cannot account for their joining. We believe that the joining may be mediated by a number of distinct types of switching proteins (Fig. 6). For example, one switching protein (P_α) may bind the germ-line S_α sequence and a second (P_μ) may bind one of the germ-line S_μ sequences. These proteins may then interact to form a heterodimer that juxtaposes the V-D-J gene with the C_α gene segment (Fig. 6). The multiple germ-line S_α sequences would increase the probability that C_α switching could occur once the appropriate joining pro-

tein is expressed. Because the germ-line S_α, $S_{\gamma 1}$, and $S_{\gamma 2b}$ sequences seem distinct, different joining proteins could bind these sequences. Accordingly, the developmental regulation of the expression of these proteins would lead to class-specific regulation of C_H switching.

Implications of Successive C_H Switching in MC101

The evidence for successive switching in MC101 indicates that two or more C_H switches can occur in a particular B cell line. In the simple deletional model for class switching proposed by Honjo and Kataoka (*9*), C_H switching progresses in a linear fashion, deleting intervening C_H genes at each stage. As mentioned, the experiments supporting this model (*9, 10*) indicate a gene order of C_μ-$C_{\gamma 3}$-$C_{\gamma 1}$-$C_{\gamma 2b}$-$C_{\gamma 2a}$-C_α. Paradoxically the MC101 clone appears to have switched from C_μ to C_α and then from C_α to $C_{\gamma 1}$, contrary to the linear deletional model for class switching proposed by Honjo and Kataoka (*9*). Several explanations seem plausible.

1) The C_H gene order is C_μ-C_α-C_γ. This seems difficult to support because of the large number of myeloma tumors that express the $C_{\gamma 1}$ gene and still contain C_α genes (*10*). Furthermore, those myeloma tumors that express the C_α gene generally appear to have deleted the $C_{\gamma 1}$ genes (*9, 10*). However, until these C_H genes are ordered in the germline DNA, this remains a formal possibility.

2) Interchromosomal recombination. In this scheme, $C_\mu \rightarrow C_\alpha$ rearrangement on one chromosome could be followed by recombination with a $C_{\gamma 1}$ gene on another chromosome to produce the MC101 $\gamma 1$ mosaic gene (V_H-C_μ-C_α-$C_{\gamma 1}$). Moreover, a prediction of this model is that the reciprocally rearranged chromosome should have a rearranged C_α gene.

3) Episomal deletion. If the deleted DNA between the V_H gene and the C_α gene forms a circular intermediate (an episome) that is at least transiently stable, the episome could then reintegrate into the chromosome, replacing the C_α gene with the $C_{\gamma 1}$ gene. This model is easily testable because it predicts the presence of a C_μ gene, a $C_{\gamma 3}$ gene, and so on, in the MC101 genome and, in particular, both the C_μ and the C_α genes should be rearranged.

The rearranged $\gamma 1$ gene of MC101, accordingly, raises two general possibilities with regard to C_H switching and normal B cell differentiation. (i) The successive (and complex) types of C_H switching observed here may arise from one or more aberrant chromosomal rearrangements that are characteristic of myeloma cell lines and would not generally be seen in normal B cells. The numerous cell divisions that occur between myeloma tumor production and our analyses of the corresponding DNA's, as well as the aneuploid nature of myeloma cells, make this a serious possibility. (ii) Normal B cells may have multiple C_H switching mechanisms, some perhaps different from any of those cited above.

Conclusion

The developmental regulation of C_H switching may operate at several different levels. (i) The nature of the sequences mediating V-D-J joining and C_H switching implies that these phenomena are regulated independently. (ii) The existence of distinct switch sequences for α, $\gamma 1$, and $\gamma 2b$ genes implies that the expression of these classes may be developmentally regulated at the level of DNA rearrangement, depending, for example, on which specific switching protein is expressed.

References and Notes

1. C. Brack, A. Hirawa, R. Lenhard-Schueller, S. Tonegawa, *Cell* **15**, 1 (1978).
2. H. Sakano, K. Huppi, G. Heinrich, S. Tonegawa, *Nature (London)* **280**, 288 (1979).
3. J. G. Seidman, E. E. Max, P. Leder, *ibid.*, p. 370.
4. P. Early, H. Huang, M. Davis, K. Calame, L. Hood, *Cell* **19**, 981 (1980).
5. M. M. Davis, K. Calame, P. W. Early, D. L. Livant, R. Joho, I. L. Weissman, L. Hood, *Nature (London)* **283**, 733 (1980).
6. M. C. Raff, *Cold Spring Harbor Symp. Quant. Biol.* **41**, 159 (1976).
7. M. D. Cooper, J. F. Kearney, P. M. Lydyard, C. E. Grossi, A. R. Lawton, *ibid.*, p. 139.
8. B. Pernis, L. Forni, A. L. Luzzati, *ibid.*, p. 175.
9. T. Honjo and T. Kataoka, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2140 (1978).
10. S. Cory and J. M. Adams, *Cell* **19**, 37 (1980); S. Cory, J. Jackson, J. M. Adams, *Nature (London)* **285**, 450 (1980); C. Coleclough, C. Cooper, R. P. Perry, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1422 (1980); T. H. Rabbits, A. Forster, W. Dunnick, D. L. Bentley, *Nature (London)* **283**, 351 (1980); Y. Yaoita and T. Honjo, *Biomed. Res.* **1**, 164 (1980); Y. Yaoita and T. Honjo, *Nature (London)*, in press.
11. T. Kataoka, T. Kawakami, N. Takahashi, T. Honjo, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 919 (1980).
12. N. Takahashi, T. Kataoka, T. Honjo, *Gene*, in press.
13. H. Sakano, R. Maki, Y. Kurosawa, W. Roeder, S. Tonegawa, *Nature (London)* **286**, 676 (1980).
14. T. Maniatis, R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, A. Efstratiadis, *Cell* **15**, 687 (1978).
15. P. W. Early, M. M. Davis, D. B. Kaback, N. Davidson, L. Hood, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 857 (1979).
16. M. M. Davis, P. W. Early, K. Calame, D. L. Livant, L. Hood, in *Eukaryotic Gene Regulation*, R. Axel, T. Maniatis, C. F. Fox, Eds. (Academic Press, New York, 1980), p. 393.
17. N. Johnson, S. Rudikoff, P. Barstad, L. Hood, in preparation.
18. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1977).
19. M. Davis, unpublished results.
20. The minor sequence differences noted might arise from one or more of the following sources: polymorphism in these flanking sequences arising from mice in different localities, mutations arising in the DNA sequences of myeloma tumors during repeated passages, repair or correction mechanisms during the C_H switching process, and DNA sequencing errors.
21. T. Honjo, personal communication.
22. E. Max, J. G. Seidman, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3450 (1979).
23. W. D. Benton and R. W. Davis, *Science* **196**, 180 (1977).
24. A. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).

25. This research was supported by NSF grant PCM 76-81546, PHS grant AI 09072, and NIH training grant GM 07616. All experiments involving recombinant organisms were conducted in accordance with the NIH Guidelines on Recombinant DNA. We thank N. Kupperman for assistance with restriction mapping and S. Crews, D. Goldberg, and M. Kronenberg for helpful discussions.

8 July 1980; revised 1 August 1980