4. J. Dancis, P. H. Berman, V. Jansen, M. E. Balis, *Life Sci.* **7**, 587 (1968).
5. W. L. Nyhan, B. Bakay, J. D. Connor, J. F. Marks, D. K. Keele, *Proc. Nat. Acad. Sci. U.S.* **65**, 214 (1970).
6. L. G. Lajtha and J. R. Vane, *Nature* **182**, 191 (1958).
7. B. R. Migeon, *Biochem. Genet.* **4**, 377 (1970); S. Wood and L. Pinsky, *Clin. Genet.* **1**, 276 (1970).
8. S. M. Gartler, E. Gandini, G. Angioni, N. Argiolas, *Ann. Hum. Genet.* **33**, 171 (1969).
9. K. Randerath and E. Randerath, *J. Chromatogr.* **16**, 111 (1964).
10. Supported by grants from NIH, NATO, and NSF. Dr. J. F. Marks of the University of Texas Southwestern Medical School, Dallas, allowed us to study patients C-1 and C-2.

8 February 1971

Deletions in Immunoglobulin Polypeptide Chains as Evidence for Breakage and Repair in DNA

Abstract. *The partial sequence of the light chain of the myeloma-like immunoglobulin Sac shows a large deletion in its variable region. The sequence provides evidence that the corresponding gene was formed by the repair of DNA broken at nonhomologous positions. Data from other immunoglobulin (heavy) chains containing large deletions are compatible with their genes also being the result of DNA breakage and nonhomologous repair. Single homologous reciprocal exchanges in DNA networks at immunoglobulin loci could be the cause of the nonhomologous breaks. The relevance of these events to the generation of normal antibody variability remains to be determined.*

Franklin *et al.* (*1*) in 1964 described the presence of immunoglobulin heavy chain in the urine of a patient with a generalized lymphadenopathy; the excreted heavy chain did not appear to be complete (*2*). A total of 14 cases of "heavy chain disease" have now been reported or discussed (*3*), and the amino acid sequences of parts of the proteins from three of them have been determined: Zuc (*4*), Hi (*5*), and Cra (*6*). All three have deletions. In 1968 Lewis *et al.* (*7*) published an account of an unusual serum γ_1 immunoglobulin Sac which appeared to have deletions in the variable regions of both its light and heavy chains; the protein was found in the serum of a patient with the tentative diagnosis of a plasma cell neoplasm (*8*). The sequence of the first 20 residues of the Sac heavy chain and of the first 16 residues of the light chain have now been determined by conventional methods (*9*), and the data for the light chain have been extended to include 48 of the first 50 residues (*10*), by means of an Edman-Begg sequenator (*11*). The sequence data fully confirm the observations of Lewis *et al.* (*7*) in that both the light and heavy chains of Sac do indeed contain large deletions in their variable regions. We here discuss some implications of the Sac light chain sequence and of the other large deletions in immunoglobulins.

Figure 1 shows the sequence (*10*) of the light chain of Sac in relation to that of the κ_I Bence Jones protein Roy (*13*). The data show that Sac light chain has the same sequence as Roy up to position 17. From position 31 Sac light chain has a sequence without demonstrable differences from that of Roy from the Roy position 99 to at least ten residues into the constant region, except for a valine/phenylalanine difference at position 106. (Position 99 is near the end of the light chain variable region.) The sequence of the intervening 13 amino acids (Sac 18 through 30) is unusual in comparison with other κ light chain sequences; it differs at 11 positions from the Roy sequence 18 to 30, and at 9 positions from the Roy sequence 86 to 98.

Three simple ancestries for the Sac light chain gene are readily imagined. The gene may be derived from the 1 to 30 region of a κ_I gene joined to the 99 to 214 region of the same or of a second gene. It may be from a κ_I 1 to 17 region joined to an 86 to 214 region. Or it may be derived from a 1 to 30 region and an 86 to 214 region, with the two partners of the overlapping 13 residue region each making some contribution to the final gene. The first possibility can be excluded, since the Sac 18 to 30 sequence differs from the most common κ_I sequence for the same region at 11 positions; none of the other κ_I proteins yet reported differs at more than four positions from this common sequence. The second ancestry, in which the Sac 18 to 30 region is derived solely from a κ 86 to 98 region, is more difficult to exclude since the κ chain 86 to 98 region is highly variable. Thirteen sequences have been reported (*14*) for this region. Twelve of these do not differ by more than four residues (five base pairs) from the sequences most similar to them. The thirteenth, light chain Eu (*15*), differs from the sequence most similar to it by six residues (seven base pairs). The Sac 18 to 30 sequence differs from the most similar of the 86 to 98 sequences by eight residues (eight base pairs), and the differences include two of the four positions in the region not yet observed to vary. It therefore appears unlikely that the Sac 18 to 30 sequence is derived simply from an 86 to 98 region, but we cannot exclude this possibility. The third ancestry for Sac, from a composite of DNA related to both sequences, would reduce the minimum number of unique residues in the two Sac ancestral genes to four as a consequence of being able to use bases from either ancestor. This number is within the normal range of variation, although two of the four required unique residues would be at positions where variations are infrequent; the four invariant residues in the presumed 86 to 98 ancestor would remain invariant. Thus a simple interpretation of the Sac light chain sequence is that its gene was derived from the 1 to 30 region of one gene and the 86 to 214 region of the same or of a second gene with each partner of the overlapping 13 residue region of hybrid DNA making some contribution to the final gene. In addition, mutations may have been introduced by the hybrid region.

There are at least two situations in which hybrid DNA is likely to be formed in vivo: during recombination and during the repair of DNA broken by mechanical or enzymatic action. Recombination normally takes place between DNA regions with some homology. We looked at the regions near 18 to 30 and 86 to 98, but could find no evidence for significant homology between them in the correct register (*16*). We conclude that the Sac light chain sequence is not a direct consequence of an unequal crossover between partially homologous regions.

Consideration of the second possibility, that the Sac sequence is the product of the repair of broken DNA molecules, is handicapped by a lack of knowledge about the structures generated in vivo during such repairs. Khorana and his associates (*17*) have used T4-polynucleotide ligase in vitro to join DNA molecules that had been synthesized with single-stranded ends able to form an overlapping hybrid DNA double helix. DNA double

helices without single-stranded regions can also be joined "end-to-end" with the same ligase preparations (*18*). These observations make it difficult to adopt any single model for discussing in vivo repair, but the following assumptions appear to avoid any serious prejudgments. We assume that if the repair involves the pairing of single-stranded regions at the ends of the broken molecules the resulting length of hybrid DNA double helix will be antiparallel; if the repair does not involve the pairing of single-stranded regions, we assume that chemical polarities will be uninterrupted across the junction. There are many ways of joining nonhomologous broken DNA molecules compatible with these assumptions. The molecules can be joined with the reading frame preserved across the junction, or with the reading frame shifted by one or by two base pairs. If the broken molecules have single-stranded regions ending in the same way (for example, both with 3′-ends), antiparallel hybrid DNA can be formed colinear with the joined molecules (as shown in Fig. 2e). If the broken molecules have different types of single-stranded ends, antiparallel hybrid DNA can still be formed but only at right angles (*19*) to the joined molecules (as in Fig. 2g). No hybrid DNA can be formed in any joint in which either or both of the broken molecules lacks a single-stranded region. Joints with DNA at right angles to the genes are likely to be replicated or transcribed in a manner indistinguishable from that for an end-to-end joint; consequently we suggest that in the present context the repair of DNA broken at nonhomologous points can have one of six outcomes, depending on whether or not there is hybrid DNA colinear with the genes and depending in which of the three possible reading frame registers the two pieces of DNA are joined. Our data on the Sac light chain and all of the other available data on immunoglobulin chains with large deletions can be reconciled with one or another of these six possibilities.

Because the sequence of Sac light chain is normal before and after an unusual region, we suggest it is the outcome of a nonhomologous repair in the correct reading frame with the intervening hybrid DNA being colinear. The heavy chain disease protein Zuc (*4*) has a normal amino terminal sequence of 18 residues followed (without any unusual sequence in between) by a normal heavy chain sequence as from position 216 to the end of the molecule; this we interpret as due to nonhomologous repair in the correct reading frame with an end-to-end joint or with the hybrid DNA at right angles.

When nonhomologous repair is in the wrong reading frame, "gibberish" or "missense" will result distal to the end of the repair structure regardless of whether or not there is colinear hybrid DNA. Chain termination is a likely event when shifts in reading frame are present (*20*). Most cases of repair in the wrong reading frame would consequently be expected to yield polypeptides shorter than usual with a region of peculiar sequence at their carboxy terminal ends. No such deletions have been reported in immunoglobulins. This could be due to the indispensability of some parts of the constant regions for antibody function, but the possibility should be considered that this type of shorter polypeptide has escaped detection.

The work of Sarabhai and Brenner (*21*) shows that polypeptide synthesis can be restarted on either side of and in a different reading frame from a nearby chain terminator. Restarting synthesis in the carboxy terminal read-

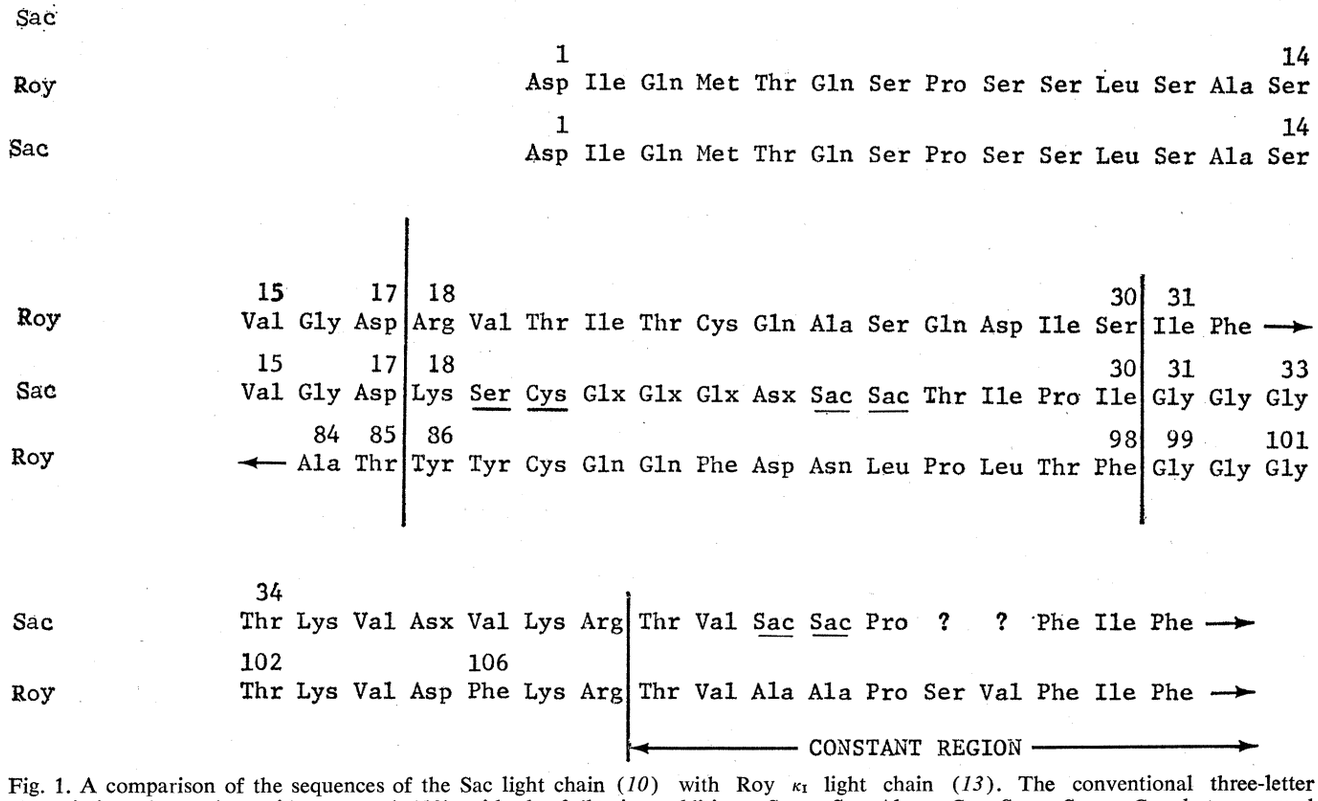

Fig. 1. A comparison of the sequences of the Sac light chain (*10*) with Roy κ_1 light chain (*13*). The conventional three-letter abbreviations for amino acids are used (*12*) with the following additions: Sac = Ser, Ala or Cys; Ser = Ser or Cys, but presumed Ser from other data; Cys = Ser or Cys, but presumed Cys from other data; ? = an unidentified residue at this position. The use of Sac as an abbreviation for Ser, Ala, or Cys is not to be confused with Sac used as a designation for the myeloma protein; their similarities are coincidental.

ing frame is therefore possible on either side of a repair which joined two broken DNA molecules with their reading frames out of register. Restarting on the proximal side of such a repair boundary could yield a polypeptide with a peculiar amino terminal sequence followed by a normal sequence after the boundary is passed. Restarting on the distal side could yield a shortened polypeptide, beginning part way into the usual sequence and continuing normally to the carboxy terminus. The heavy chain disease protein Cra (*6*) has a peculiar sequence of 11 amino acids prior to a normal sequence, as from heavy chain position 216 to the end of the molecule; it can be interpreted as due to restarting the polypeptide chain on the proximal side of a repair boundary with 11 residues of resulting gibberish. The Sac heavy chain (*9*) has an essentially normal sequence, as from heavy chain position 103 to the end of the molecule. This could be interpreted as due to restarting polypeptide synthesis on the distal side of a repair, but a lack of sufficient comparable sequences makes it impossible to exclude the possibility that the amino terminal sequence of Sac is, in fact, unusual and that restarting occurred proximal to the repair boundary. If other cases of possible chain restarting are discovered in the future, a search for an amino terminal fragment would appear to be worthwhile.

Two current models for the structure of immunoglobulin loci both have features that could account for the occurrence of large deletions, or for the nonhomologous breaks in DNA which we suggest may have caused them. On the episome-translocation hypothesis of Dreyer and Bennett (*22*) and Gally and Edelman (*23*), the deletions can be interpreted as due to various combinations of errors in the somatic excision of variable region episomes and their insertion into related constant region genes. The data that we have discussed are in general compatible with this, but the molecular basis or chromosomal mechanics of the required errors in the normally accurate process are difficult to specify.

On the DNA network hypothesis of Smithies (*24*), a deletion is one possible outcome of a single homologous sister-strand crossover within the network. We emphasize, however, that reciprocal exchanges generated in other ways within the network can lead to similar topological forms and to deletions. For illustrative purposes we will discuss in detail the consequences of a single homologous crossover.

Smithies proposed that immunoglobulin loci are made up of networks of branched DNA having a high lateral multiplicity in the variable regions and a lower multiplicity in the constant regions. The basic structure from which the proposed networks are derived is a DNA detour. One of these is shown in Fig. 2a in synapsis with another. Some of the presumed consequences of a single homologous crossover between them are illustrated in the remainder of the figure. As can be seen, single crossovers in such structures inevitably lead to DNA breakage at the next segregation. The breaks are a consequence of the segregation of the chromosomes joined by the crossover, but the points of breakage would not be expected to coincide with the exact position of the original crossover; nor would the breaks be expected necessarily to occur at homologous positions in the two molecules involved. Thus the resultant chromosomes would usually be unequal—one with a deficiency and the other with a duplication. The overall consequences of the breaks will depend on the type of repair subsequently effected in the two chromosomes. The chromosome containing the duplication can be repaired in a homologous way with colinear hybrid DNA to yield a product of normal length in which the correct reading frame is maintained (Fig. 2, k and l). The duplication chromosome can also be repaired in at least two nonhomologous ways (Fig. 2, m and o), with the attendant chances of reading frame shifts; included in the possible products are repaired chromosomes containing insertions (see, for example, Fig. 2p). The deficient chromosome must always

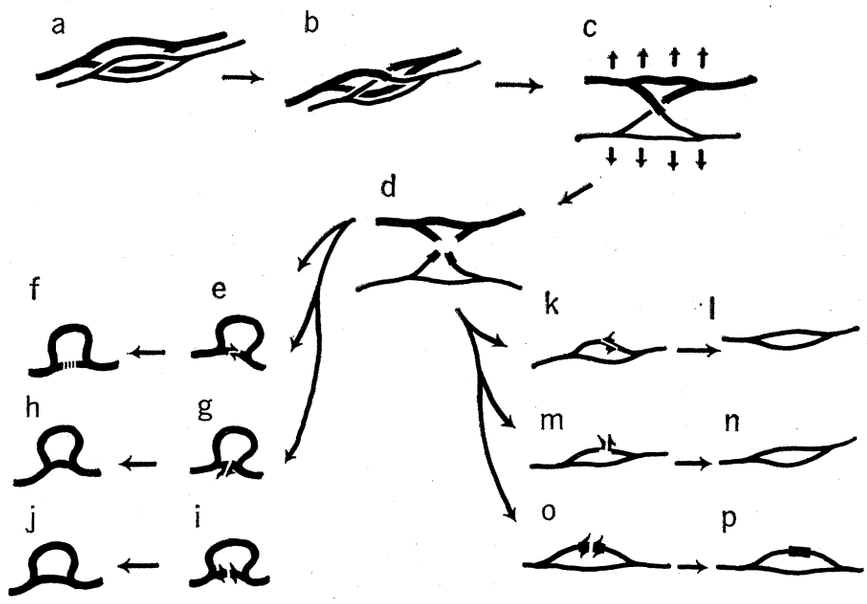

Fig. 2. Some of the possible consequences of a single homologous crossover in a DNA detour. Half arrowheads and tails indicate chemical polarities. (a) Synapsis between homologous chromosomes or sister chromatids. (b) Single homologous crossover anywhere in the DNA detour. (c) Segregation. (d) Two breaks must occur in separating the chromosomes and in the general case lead to one chromosome with a deficiency and one with a duplication. (e–j) Some of the forms of repair of the chromosome with the deficiency; all are nonhomologous: (e) If the broken molecules have regions of single-stranded DNA with the same type of ends (for example, if both are 3′) a hybrid antiparallel DNA helix can be formed colinearly with respect to the chromosome. (f) A deletion with a hybrid region is one result from (e). (g) If the broken molecules have single-stranded regions with different types of ends they can still form a hybrid DNA helix but only at right angles to the rest of the chromosome. (h) A deletion without a hybrid region is a possible result from (g). (i) If either or both of the broken molecules lack single-stranded regions some form of an end-to-end joint is required. (j) A deletion without a hybrid region is one result from (i). (k–p) Some of the forms of repair of the chromosome with the duplication: (k) Homologous repair is possible if the broken molecules have single-stranded regions with the same type of ends. (l) A normal length product with no reading frame errors can result from (k). (m) Nonhomologous repair with single-stranded regions having different types of ends. (n) The product may be of normal length or may contain an insertion or a deletion. (o) Nonhomologous repair without hybrid DNA. (p) An insertion is one possible product of (o).

be repaired nonhomologously (Fig. 2, e–j) yielding the various types of deletion we have discussed.

These considerations show that a network of branched DNA has topological features which can cause single exchanges within it to lead to the breaking of DNA molecules at nonhomologous places. The available sequence data from immunoglobulins with large deletions are compatible with the repair of such breaks. The data cause us to raise but do not permit us to answer the question of whether DNA breakage and repair occurs normally during the generation of antibody variability.

O. SMITHIES, D. M. GIBSON*
E. M. FANNING
Laboratory of Genetics, University of Wisconsin, Madison 53706
M. E. PERCY, D. M. PARR
G. E. CONNELL
Department of Biochemistry, University of Toronto, Toronto, Canada

References and Notes

1. E. C. Franklin, J. Lowenstein, B. Bigelow, M. Meltzer, *Amer. J. Med.* **37**, 332 (1964).
2. E. C. Franklin, *J. Exp. Med.* **120**, 691 (1964).
3. ———, editorial in *N. Engl. J. Med.* **282**, 1098 (1970).
4. B. Frangione and C. Milstein, *Nature* **224**, 597 (1969).
5. W. D. Terry and J. Ohms, *Proc. Nat. Acad. Sci. U.S.* **66**, 558 (1970). Sequence data for this protein, Hi, do not encompass the deletion.
6. E. C. Franklin and B. Frangione, *Proc. Nat. Acad. Sci. U.S.* **68**, 187 (1971). We thank the authors for allowing us to read and quote from their paper prior to its publication.
7. A. F. Lewis, D. E. Bergsagel, A. Bruce-Robertson, R. K. Schachter, G. E. Connell, *Blood* **32**, 189 (1968); G. E. Connell, K. J. Dorrington, A. F. Lewis, D. M. Parr, *Can. J. Biochem.* **48**, 784 (1970).
8. The patient Sac had a history of repeated infections, splenomegaly, anemia, leukopenia, thrombocytopenia, a moderate marrow plasmacytosis but no lytic bone lesions or osteoporosis, and a β-paraprotein in the serum. These symptoms, like those in the other cases in which immunoglobulins with large deletions have been found, are not fully typical of multiple myeloma.
9. D. M. Parr, M. E. Percy, G. E. Connell, in preparation.
10. O. Smithies, D. M. Gibson, E. M. Fanning, in preparation.
11. P. Edman and G. Begg, *Eur. J. Biochem.* **1**, 80 (1967). Present instrument manufactured by Illitron Division, Illinois Tool Works Division, Inc., Chicago, Ill.
12. The three-letter abbreviations used are as follows: Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Asx, aspartic acid or asparagine; Cys, cysteine; Cys, cysteine or serine but presumed cysteine from other data; Gln, glutamine; Glu, glutamic acid; Glx, glutamic acid or glutamine; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Sac, serine, alanine or cysteine; Ser, serine; Ser, serine or cysteine but presumed serine from other data; Thr, threonine; Tyr, tyrosine; Try, tryptophan; Val, valine; ?, an unknown residue.
13. N. Hilschmann, *Z. Physiol. Chem.* **348**, 1077 (1967); ———, H. U. Barnikol, M. Hess, B. Langer, H. Ponstingl, M. Steinmetz-Kayne, L. Suter, S. Watanabe, *Fed. Eur. Biochem. Soc. Symp.* **15**, 57 (1969).
14. We thank Drs. G. P. Smith, Leroy Hood, and W. M. Fitch for permission to use their recent summary of existing data (*Annu. Rev. Biochem.*, in press) and R. M. Goodfliesh for unpublished data on the sequence of the Bence Jones protein Dil.
15. G. M. Edelman, B. A. Cunningham, W. E. Gall, P. D. Gottlieb, H. Rutishauser, M. J. Waxdal, *Proc. Nat. Acad. Sci. U.S.* **63**, 78 (1969).
16. We thank Dr. W. M. Fitch for looking for homologies, using computer programs which he has described in *J. Mol. Biol.* **16**, 9 (1966); *ibid.* **49**, 1 (1970).
17. K. L. Agarwal, H. Buchi, M. H. Caruthers, N. Gupta, H. G. Khorana, K. Kleppe, A. Kumar, E. Ohtsuka, U. L. Rajbhandary, J. H. Van de Sande, V. Sgaramella, H. Weber, T. Yamada, *Nature* **227**, 27 (1970).
18. V. Sgaramella, J. H. Van de Sande, H. G. Khorana, *Proc. Nat. Acad. Sci. U.S.* **67**, 1468 (1970).
19. C. A. Thomas, B. A. Hamkalo, D. N. Misra, [C. S. Lee] *J. Mol. Biol.* **51**, 621 (1970) have published electron micrographs of DNA "rings" and "lariats" that can be interpreted as the equivalents of the forms that result from colinear and right-angle annealing of single-stranded ends of DNA molecules. In their work the DNA was mechanically sheared and single-stranded regions were exposed at the ends by the action of exonucleases; the single-stranded regions were probably homologous.
20. F. H. C. Crick, L. Barnett, S. Brenner, R. J. Watts-Tobin, *Nature* **192**, 1227 (1961).
21. A. Sarabhai and S. Brenner, *J. Mol. Biol.* **27**, 145 (1967).
22. W. J. Dreyer and J. C. Bennett, *Proc. Nat. Acad. Sci. U.S.* **54**, 864 (1965).
23. J. A. Gally and G. M. Edelman, *Nature* **227**, 341 (1970).
24. O. Smithies, *Science* **169**, 882 (1970).
25. Supported in part by NSF GB-4362 and NIH GM-15422 and the Medical Research Council of Canada. We thank Mrs. I. Csermely, Toronto, Canada, for technical assistance, and Dr. George P. Smith, our colleagues, and students for discussions of these ideas and criticisms of the manuscript. The cooperation of Dr. R. K. Schachter and her colleagues at the Women's College Hospital, Toronto, is gratefully acknowledged.
* Present address: Department of Biochemistry, University of Sherbrooke, Sherbrooke, P.Q.

11 December 1970

Actinomycin D: Renewed RNA Synthesis after Removal from Mammalian Cells

Abstract. *Suppressed RNA synthesis exceeded control rates 2 hours after removal of actinomycin from serial lines of green monkey kidney cells in exponentially dividing monolayer cultures. Enhanced synthesis was directly related to the dose of actinomycin, and release from inhibition was temperature dependent. A broad range of RNA sizes was made after removal of drug, but cytoplasmic and nucleolar sedimentation profiles were abnormal for at least 20 hours.*

Stability of actinomycin-DNA complexes (*1*) and persistence of drug-induced effects in cells in the absence of drug (*2*) have lead to the widely accepted conclusion that suppression of DNA-dependent RNA synthesis by actinomycin in vivo is irreversible. We were surprised to find that suppressed RNA synthesis in two lines of green monkey kidney cells and in HeLa cells increased after withdrawal of drug. In fact, rates of synthesis sometimes exceeded that of control.

Drug toxicity in the monkey kidney cell lines Vero (*3*) and JR (*4*) is low, and renewed RNA synthesis continued for days. Cultures treated with up to 1 μg of actinomycin per milliliter often could be serially passaged. On the other hand, studies with HeLa cells were limited because of toxicity.

Using Vero cells for our test system, we looked at some of the factors affecting rates of RNA synthesis after drug treatment and at the kinds of RNA that were made. Monolayer cultures grown in plastic flasks were nour-

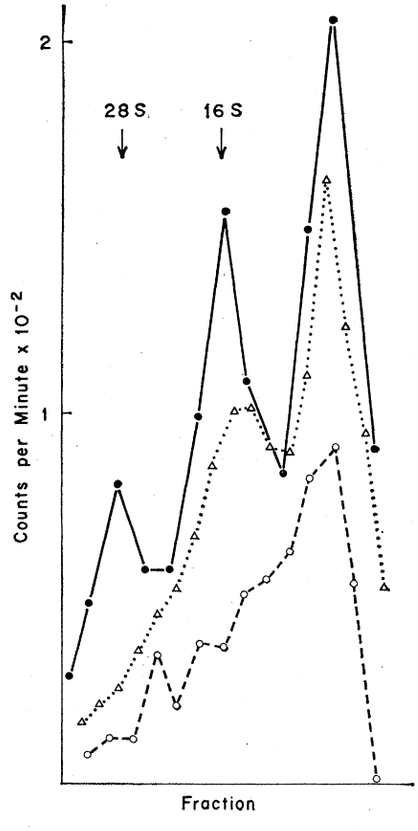

Fig. 1. Sedimentation profiles of cytoplasmic RNA from cells labeled for 1 hour with [5-H^3]uridine (4 μc/ml) during a 1½-hour exposure to 1 μg of actinomycin per milliliter (dashed line); 19 hours after removal of drug (dotted line); or in the absence of treatment (solid line). Sedimentation at 50,000 rev/min, 2¼ hours, SW 50.1 rotor.