the interpretation difficult. It is interesting, however, that MacFadden and Smith[6] did find small amounts of DNP-leu and DNP-leu-val.

[19] Dreyer, W. J., and J. C. Bennett, these PROCEEDINGS, **54**, 864 (1965).

[20] Edelman, G. M., B. Benacerraf, Z. Ovary, and M. D. Poulik, these PROCEEDINGS, **47**, 1751 (1961).

[21] Choules, G. L., and S. J. Singer, *Immunochemistry*, in press; Doolittle, R. F., and S. J. Singer, these PROCEEDINGS, **54**, 1773 (1965).

[22] A well-controlled study in which rabbit anti-DNP antibody was shown to exhibit different optical rotatory properties from rabbit normal γ-globulin has just been brought to my attention: Steiner, L. A., and S. Lowey, *J. Biol. Chem.*, **241**, 231 (1966).

A MOLECULAR MODEL FOR GENE REPRESSION

BY JEFFREY H. MILLER AND HENRY M. SOBELL*

DEPARTMENT OF CHEMISTRY, UNIVERSITY OF ROCHESTER, ROCHESTER, NEW YORK

Communicated by Marshall Gates, March 14, 1966

In their original formulation of the regulator gene hypothesis, Jacob and Monod suggested that the repressor might be a polynucleotide possessing a specific base sequence which was complementary to an operator site.[1] However, more recent studies have conclusively shown that the repressor is a protein, composed of subunits, possessing allosteric properties.[2-5] A peculiar property of the repressor is its instability under conditions of growth. This has prompted the suggestion that the active repressor could be a protein-messenger RNA complex, the messenger fraction accounting for its growth lability.[4]

The purpose of this communication is to suggest a plausible molecular model for gene repression, based on the fundamental hypothesis that the repressor is a *ribonucleoprotein* capable of interacting with an operator site in the deep groove of the DNA helix, using an alternate base-pairing configuration. Evidence for the existence of this alternate pairing configuration has been provided by studies on certain polynucleotide complexes and crystallographic studies of purine-pyrimidine crystalline complexes.

The Model.—The model, which is represented by the diagrams in Figures 1 and 2, involves the following features.

(1) Stretches of purine sequences in the double helical DNA structure can accommodate short pyrimidine-containing RNA oligonucleotides in the deep groove of the helix. This is accomplished by means of an alternate base-pairing configuration, one involving hydrogen bonding with the imidazole nitrogen, N-7, of adenine and guanine. *The Watson-Crick pairing rules are maintained* (adenine = uracil, guanine = cytosine), *but the geometry of the interaction is different* (Fig. 1).

(2) The regulator gene and the operator gene contain segments of DNA which have similar (or perhaps identical) stretches of *purine-containing sequences.*

(3) The product of the regulator gene (or genes) is a ribonucleoprotein which is composed of (*a*) at least part of the *messenger RNA transcribed from the regulator gene DNA*, (*b*) the protein coded by this gene (or genes). The messenger RNA portion must contain a stretch of pyrimidines which is complementary to a stretch of purines on the operator site.

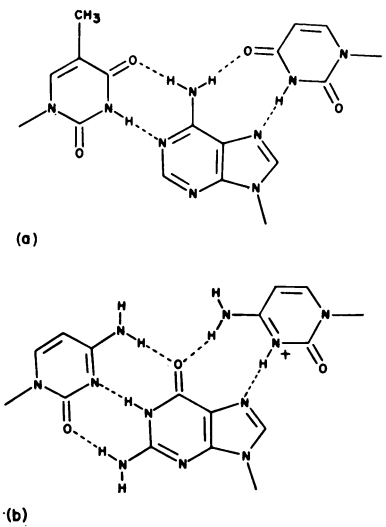

FIG. 1.—A schematic diagram showing the alternate hydrogen-bonding configuration postulated between the adenine-uracil and guanine-cytosine base pairs. Regions of DNA-containing stretches of purines could accommodate an RNA oligonucleotide consisting of a stretch of pyrimidines which hydrogen-bond in this fashion in the deep groove of DNA, causing little distortion. In order for this to occur, the RNA base sequence must be complementary with the purine-containing DNA strand, and the polarity of the sugar-phosphate chains must be the same.

(4) The ribonucleoprotein has allosteric properties which allow it to interact with the operator in the presence of corepressor in repressible systems, and in the absence of inducer in inducible systems. *The affinity of the repressor for the operator is determined by the specific hydrogen bonding between complementary base sequences.* Three models of gene repression can immediately be postulated, one involving constitutive repressor synthesis, two involving self-regulation of repressor synthesis (Fig. 2).

Discussion.—The existence of alternate pairing configurations between purines and pyrimidines has been demonstrated in polynucleotide complexes and in crystalline complexes of monomer base pairs. These studies have clearly shown the existence of an alternate pairing configuration between adenine and uracil rings. This was first demonstrated by Hoogsteen in a crystalline complex containing alkylated adenine and thymine derivatives.[6] A similar structure containing alkylated adenine and uracil derivatives has also been described.[7] Although no direct evidence has been obtained concerning the structure of the triple-stranded 2-poly U:poly A complex, model building strongly suggests that the complex consists of a Watson-Crick interaction between one strand of poly U and poly A, the other strand of poly U interacting in the deep groove of this helix with the Hoogsteen configuration.[8]

A similar complex is formed between oligo G and poly C. It has been demonstrated that a stable 2-poly C:oligo G complex can exist at neutral pH.[9, 10] Although no structural evidence is currently available on this complex, infrared data indicate that half the cytosine residues are protonated, this occurring on the ring nitrogen, N-3. The geometry of the complex is unknown; however, of the two possible structures, one uses the identical geometry of the Hoogsteen adenine-uracil configuration. Hydrogen bonding involving a protonated cytosine residue, such as this configuration requires, has been shown to occur in the double helical form of poly C[11, 12] and in the crystal structure of cytosine-5-acetic acid.[13] Although the pK value of cytosine is 4.3, one would expect the interaction to be

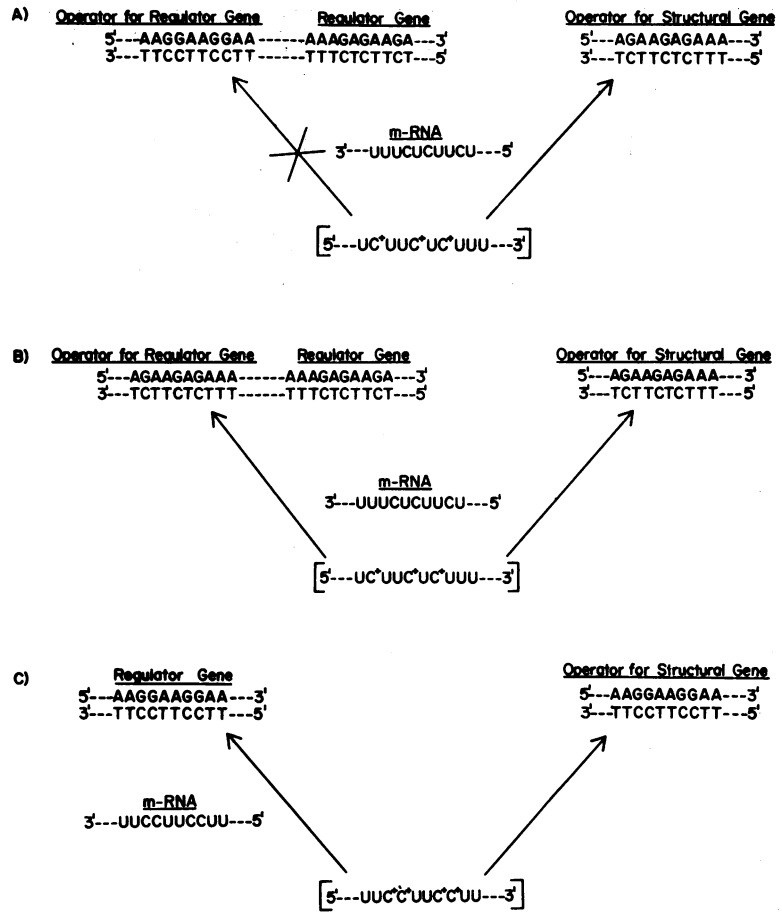

FIG. 2.—Three molecular models for gene repression. (*A*) The repressor has affinity only for the structural gene operator. (*B*) The repressor has affinity for both the structural gene operator and the regulator gene operator. (*C*) The repressor has affinity for the structural gene operator and the regulator gene operator. The regulator gene operator is within the repressor structural cistron.

stabilized by a cooperative effect, altering the effective pK of cytosine. It is therefore possible that a second type of Watson-Crick specificity between adenine-uracil and guanine-cytosine$^+$ exists, involving hydrogen bonding to the imidazole nitrogen, N-7, of adenine and guanine. This alternate configuration, however, requires the DNA strand to have a consecutive purine sequence, the RNA strand a consecutive pyrimidine sequence.

Other interactions, such as those involving 2-oligo G:poly C, are known to occur, but are much less stable, as evidenced by their melting temperatures.[9] Complexes between poly A and poly I are known. In particular, a complex containing 2-poly I:poly A exists.[14] The structure is not yet known; however, it is likely that one strand of poly I interacts with the N-1 ring nitrogen of adenine, and the other strand interacts with the imidazole nitrogen, using a configuration analogous (although having different dimensions) to the alternate pairing configuration proposed be-

tween adenine-uracil and guanine-cytosine+. This raises the possibility that an additional interaction can exist between adenine-hypoxanthine, provided a certain amount of "wobble" is permitted. However, it is not important to postulate this at the present time.

For the lac operon in *E. coli*, evidence has been presented which strongly indicates that repressor formation is not controlled by the level of inducer.[15] Model *A* provides a scheme which enables the repressor to be synthesized constitutively. The model requires the base sequence on the regulator (i.e., 3′–5′) strand be opposite in polarity to the base sequence on the operator. Thus, messenger RNA complementary to the regulator gene has the polarity (in this case) 5′–3′, and this strand can therefore interact at the operator site for the structural gene. The number of different operons that can be distinguished in this manner is 2^n, where n is the number of bases in the sequence. Accordingly, a segment of ten pyrimidines would be sufficient to code for more than a thousand repressors.

Evidence of a system in which the repressor can control its own synthesis has been provided by Garen and Otsuji.[2] These authors have isolated a protein fraction which is a product of one of the regulator genes for the alkaline phosphatase operon in *E. coli*, and have shown that its level is correlated with the level of alkaline phosphatase, strongly implying that the synthesis of both is under the control of the same repressor. Models *B* and *C* provide schemes whereby the repressor can regulate its own synthesis. Model *B* is identical to model *A* with the exception that the regulator gene has an operator which has affinity for the repressor. This requires the base sequence on both regulator and structural gene operators to be, at least in part, identical. This model readily provides for a feedback mechanism for repressor synthesis. Model *C* provides for an alternate explanation for control of repressor synthesis. This model requires the operator for the regulator gene to be part of the structural cistron of the repressor, the stretch of pyrimidines on the repressor either back-interacting with the stretch of purines coding for it, or interacting with a different stretch of purines. The first of these possibilities requires the sequence to be symmetrically arranged (see Fig. 2*C*). Since recent evidence strongly indicates that the operator and structural genes are distinct,[16] we favor model *B*.

A decanucleotide might be expected to interact strongly with the DNA helix. A fundamental requirement of the operator-repressor interaction is that it is reversible; the presence of an acidic protein component associated with this oligonucleotide would be expected to enhance dissociation. It is of interest in this regard that the protein fraction isolated by Garen and Otsuji is an acidic protein. A basic protein (e.g., a histone) would be expected to promote association.

It is well known that in lysogenic cells containing temperature phage λ, the vegetative state of the phage can be induced with ultraviolet radiation. It has been suggested that ultraviolet light disrupts the synthesis of a cytoplasmic repressor, called the immunity substance, which normally prevents the synthesis of early viral proteins necessary for the vegetative cycle.[1] If the immunity substance was a ribonucleoprotein containing a stretch of pyrimidines, one would expect it to be particularly susceptible to UV inactivation, forming uracil and cytosine photodimers. This might be expected to interfere with repressor synthesis by inactivating a precursor or destroying the affinity of the repressor for the operator site.

Ultraviolet radiation is known to cause photodimers of thymine and uracil and to deaminate cytosine.[17-19]

Recent experiments by Gallant and Stapleton have shown that the repressor is metabolically unstable.[3] Similar findings have been reported by Sadler and Novick.[4] The latter studied a mutant of *E. coli*, i^{TSS}, in which the synthesis of repressor was found to be temperature-sensitive. They found that after arrest of repressor synthesis, enzyme synthesis rose rapidly. This does not occur in the absence of growth. Sadler and Novick suggest that the repressor is therefore not an unstable protein, but rather a protein combined with an unstable fraction (e.g., mRNA).

An important experimental observation which would lend weight to the model would be the demonstration of stable triple helical complexes between polymers such as deoxy$(AG)_n$ with deoxy$(TC)_n$, and deoxy$(AAG)_n$ with deoxy$(TTC)_n$, and the determination of their structures by fiber X-ray diffraction techniques.

Summary.—A molecular model for gene repression is presented. The model rests on the fundamental hypothesis that the repressor is a ribonucleoprotein which contains an RNA oligonucleotide region which can interact in the deep groove of the DNA helix at the operator locus. An alternate base-pairing configuration is used in this interaction, but the Watson-Crick base-pairing rules are maintained. The model incorporates recent information concerning alternate base-pairing configurations in crystal structures and in polynucleotide complexes. It also provides a possible explanation for the growth instability of the repressor.

* Support is acknowledged from the American Cancer Society and the National Institutes of Health, USPHS.

[1] Jacob, F., and J. Monod, *J. Mol. Biol.*, **3**, 318 (1961).

[2] Garen, A., and N. Otsuji, *J. Mol. Biol.*, **8**, 841 (1964).

[3] Gallant, J., and R. Stapleton, *J. Mol. Biol.*, **8**, 442 (1964).

[4] Sadler, J. R., and A. Novick, *J. Mol. Biol.*, **12**, 305 (1965).

[5] Bourgeois, S., M. Cohn, and L. E. Orgel, *J. Mol. Biol.*, **14**, 300 (1965).

[6] Hoogsteen, K., *Acta Cryst.*, **12**, 822 (1959).

[7] Mathews, F. S., and A. Rich, *J. Mol. Biol.*, **8**, 89 (1964).

[8] Felsenfeld, G., D. R. Davies, and A. Rich, *J. Am. Chem. Soc.*, **79**, 2023 (1957).

[9] Lipsett, M. N., *J. Biol. Chem.*, **239**, 1256 (1964).

[10] Howard, F. B., J. Frazier, M. N. Lipsett, and H. T. Miles, *Biochem. Biophys. Res. Commun.*, **17**, 93 (1964).

[11] Langridge, R., and A. Rich, *Nature*, **198**, 725 (1963).

[12] Hartman, K. A., Jr., and A. Rich, *J. Am. Chem. Soc.*, **87**, 2033 (1965).

[13] Marsh, R. E., R. Bierstedt, and E. L. Eichhorn, *Acta Cryst.*, **15**, 310 (1960).

[14] Rich, A., these PROCEEDINGS, **46**, 1044 (1960).

[15] Novick, A., J. M. McCoy, and J. R. Sadler, *J. Mol. Biol.*, **12**, 328 (1965).

[16] Jacob, F., and J. Monod, *Biochem. Biophys. Res. Commun.*, **18**, 693 (1965).

[17] Setlow, R. B., and W. L. Carrier, these PROCEEDINGS, **51**, 226 (1964).

[18] Freeman, K. B., P. V. Hariharan, and H. E. Johns, *J. Mol. Biol.*, **13**, 833 (1965).

[19] Brown, I. H., K. B. Freeman, and H. E. Johns, in *Abstracts*, Annual Meeting, Biophysical Society, 1966, p. 16.