ON THE NATURE OF *CIS*-ACTING REGULATORY PROTEINS AND GENETIC ORGANIZATION IN BACTERIOPHAGE: THE EXAMPLE OF GENE *Q* OF BACTERIOPHAGE λ[1]

HARRISON ECHOLS[2], DONALD COURT[3] AND LINDA GREEN[2]

Department of Molecular Biology, University of California, Berkeley, California 94720[2]
and Laboratory of Molecular Biology, National Cancer Institute,
Bethesda, Maryland 20014[3]

Manuscript received December 6, 1975
Revised copy received January 15, 1976

ABSTRACT

We note the existence of a "partially *cis*-acting" regulatory protein of bacteriophage λ: the product of the phage *Q* gene. We suggest that there may be a complete spectrum from "all *cis*" to "all *trans*" for such regulatory proteins. This behavior might arise because a DNA-binding protein either acts at a nearby (*cis*) site soon after synthesis or becomes "lost" for its *trans* activity on another genome through nonspecific interactions with DNA. Our proposed explanation provides one evolutionary basis for the linkage of genes for regulatory proteins and the sites at which such proteins act; it also suggests a possible rationale for the "metabolic instability" of certain regulatory proteins.

TWO puzzling phenomena involving gene expression in prokaryotic viruses are "*cis*-acting" regulatory proteins and "metabolically unstable" proteins (proteins that rapidly cease to function once synthesis stops). An interesting problem of gene organization in prokaryotic viruses is why regulatory genes are close to the sites at which their products act. In this report we present evidence for the existence of a partially *cis*-acting regulatory protein of bacteriophage λ (the gene *Q* product), indicating that there may be a complete spectrum from "all *cis*" to "all *trans*" for regulatory proteins. We suggest a biochemical explanation for such behavior that bears on the other points noted above.

MATERIALS AND METHODS

The methodology for growth of cells, infection, and assay for endolysin is essentially that described previously (COURT, GREEN and ECHOLS 1975) except that no starvation period was employed prior to infection. In brief, *E. coli* cells were grown in a minimal medium with glycerol as carbon source, centrifuged, resuspended, and infected in cold adsorption buffer. After infection at a multiplicity of seven phage of each genotype per cell, the culture was diluted into warm (37°) minimal medium and samples taken for assay of endolysin. Sonic extracts were prepared and assays for endolysin were carried out by following the lysis of bacteria "sensitized" with Tris-EDTA. The bacterial strain used for infection was C600*su*$^-$ and the phage mutations used were *cIam14*, *cII28*, *cIIIam611*, *Qam21* and *Ram60* (KAISER 1957; CAMPBELL 1961; COURT, GREEN and ECHOLS 1975).

[1] This work was supported in part by U.S. Public Health Service Grant GM 17078 from the National Institute of General Medical Sciences.

RESULTS AND DISCUSSION

Preferential cis-action of the Q gene product.

The *Q* gene of phage λ specifies a protein that is a positive regulator for the late stage of viral development through its capacity to activate transcription from the genes that specify phage structural proteins and proteins required for cell lysis (Figure 1) (see ECHOLS 1971 or THOMAS 1971 for reviews). In a formal genetic sense, based on complementation for phage production, *Q* protein acts in *trans* (CAMPBELL 1961). A conveniently assayed parameter of the activity of *Q* protein is the enzyme endolysin, the product of the *R* gene (Figure 1). We have asked whether *Q* protein functions preferentially on the genome from which it has been synthesized by an assay of endolysin production when a functional *R* gene is either *cis* or *trans* to a functional *Q* gene (i.e. Q^+R^+/Q^-R^- or Q^+R^-/Q^-R^+). To minimize other potential regulatory influences, we have used phage which are all $cI^-cII^-cIII^-$, thus removing a variety of potential repression effects on late gene expression (see ECHOLS 1972).

The results of this *cis/trans* experiment are given in Figure 2. Preferential *cis* action of the Q^+ gene product is clearly indicated. The total phage burst was about the same for the *cis* and *trans* experiment, as was also the yield of each phage type (in the *cis* experiment the progeny phage were 48% Q^-R^- and in the *trans* experiment 50% Q^-R^+). The equivalence of the phage types in the burst indicates that the cells were efficiently infected with each genotype. Preferential *cis* action of Q^+ was also found if the *Qam73* mutation was used instead of the *Qam21* mutation used for the data of Figure 2.

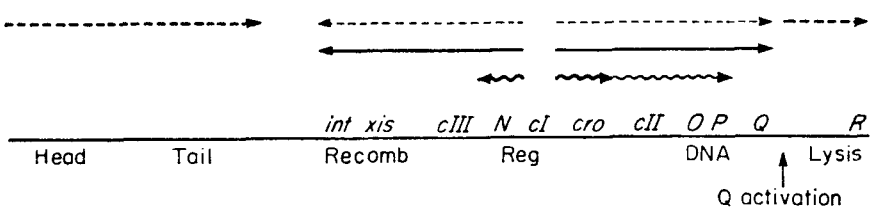

FIGURE 1.—Transcription events during lytic development by phage λ.

Approximate DNA regions transcribed during the different stages of lytic growth are shown: (⇝) represents the immediate-early stage of RNA synthesis, performed solely by the host transcription machinery, in which the *N* and *cro* gene RNAs are the major products; (⟶) represents the delayed-early stage of RNA synthesis, in which *N* protein activates transcription of the *cIII* to *int* and *cII* to *Q* regions; (⇢) represents the late stage of RNA synthesis, in which *Q* protein activates transcription of the lysis, head, and tail regions. During the late stage of lytic development, early gene transcription is reduced through the action of the *cro* protein. Since λ DNA exists in a circular or concatemeric form during much of its intracellular life, it is likely that the actual unit of transcription is DNA with the lysis region joined to the head region, rather than the linear molecule extracted from phage and indicated here. The probable site at which *Q*-activation occurs (TOUSSAINT 1969; HERSKOWITZ and SIGNER 1970) is indicated by the upward vertical arror (↑). Specific genes of the "regulation region"—*cIII*, *N*, *cI*, *cro*, *cII*—are indicated above the "λ DNA", as are the *int* and *xis* genes for site-specific recombination, the DNA replication genes *OP*, the late regulatory gene *Q* and the *R* gene for endolysin.

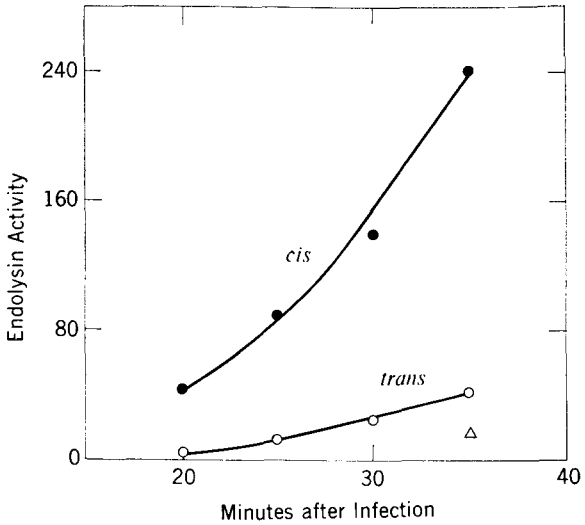

FIGURE 2.—Partial *cis*-activity of the *Q* protein for endolysin production.

Cells were infected at a multiplicity of 7 phage of each genotype, and at the times indicated the infected cells were chilled, centrifuged, and sonic extracts prepared for assay of endolysin. The phage burst was about 8 phage/cell, a value typical of high multiplicity infection in minimal medium. The adsorption efficiency for the input phage was approximately 99%.

●: endolysin production for Q^+R^+/Q^-R^-;

○: Q^-R^+/Q^+R^-;

△: Q^-R^+ alone.

There are other conceivable explanations for the apparent *cis* action of *Q* protein although they seem to us unlikely: (1) *Q* protein might be so severely limiting that minor fluctuations in gene dosage from cell to cell generate an apparent *cis* effect; (2) there might be an unexpected polar effect of the nonsense mutation in gene *Q* on expression of the adjacent R^+ gene. We have attempted to examine these possibilities in additional experiments. The first alternative possibility seems to be remote because an increase in Q^+ gene copies through a Q^+R^+/Q^+R^- infection increased endolysin production less than 20% above the Q^+R^+/Q^-R^- level. To look for the second effect, we have repeated the experiment in a host carrying an *suA* polarity suppressor (MORSE and PRIMAKOFF 1970); however no appreciable increase in the *trans* activity of Q^+ was found. From these results, we conclude that the *Q* gene product probably works preferentially in *cis*, although other explanations cannot be completely excluded.

Proposed mechanism for preferential cis-action.

What is the biochemical explanation for a *cis*-acting protein? We suggest that the partial or complete *cis*-activity of a regulatory protein might result from limitations in the binding specificity of such proteins for their DNA targets. Consider the case of *Q* protein and assume it is a DNA-binding protein. *Q* protein is probably synthesized from an mRNA still in close physical proximity to the *Q* gene, which will place the newly synthesized *Q* protein close to its site of action (see Figure 1). To reach another viral genome *Q* protein must diffuse through

and possibly interact with a very large number of DNA sites that are *not* its specific target. Thus many molecules of protein may become "lost" in their *trans* activity on another genome through "nonspecific" interactions with DNA (the same analysis can of course be applied to a specific RNA-binding protein such as a specific "anti-termination factor").

This explanation provides a potential link for the three phenomena noted in the introduction. A DNA-binding regulatory protein can function with a less rigorous (and presumably easier to evolve) recognition mechanism if the gene for the protein is next to its site of action. For such a situation, the protein will be partially or completely *cis*-acting because the regulatory site will only be completely occupied when the regulatory protein is present at *locally* high concentration through nearby synthesis. If this synthesis of new protein ceases, the regulatory protein may exhibit "metabolic instability" because repeated action is limited by loss in nonspecific interactions (that is, a true equilibrium with respect to all available DNA sites is one in which the regulatory site will be rarely occupied).*

The general problem of nonspecific interactions between regulatory proteins and DNA has been discussed in detail by others (VON HIPPEL, *et al.* 1974; LIN and RIGGS 1975). LIN and RIGGS have calculated that even for the highly specific *lac* operon repressor, 98% of the regulatory protein will normally be found at nonspecific sites (assuming no *cis* relationship). Reasoning along lines similar to ours, but from a different point of view, SUSSMAN and BEN ZEEV (1975) have suggested that prophage induction might result from the *generation* of nonspecific sites for repressor binding as a result of DNA damage. As noted previously, the clustering of regulatory genes and sites also prevents their ready separation by recombination with other phages (THOMAS 1964; STAHL and MURRAY 1966; DOVE 1971).

Examples of cis-acting and metabolically unstable proteins.

The most complete study of a *cis*-acting protein is the *A* protein of ϕX174. The *A* protein is required for initiation of replication of viral DNA and is *cis*-acting (TESSMAN 1966; LINDQVIST and SINSHEIMER 1967). The *A* protein acts at a site close to or within the *A* gene to nick the viral DNA *in vivo* (FRANCKE and RAY 1972) and *in vitro* (HENRY and KNIPPERS 1974). An interesting aspect of the reaction *in vitro* is limited turnover of the enzyme. All of these properties are consistent with the explanation for *cis*-activity *in vivo* proposed above; they are not particularly in accord with other explanations for *cis*-acting proteins (e.g. protein acts only during synthesis or only at a special cell site) (see HENRY and KNIPPERS 1974).

Another clear example of a *cis*-acting protein is the product of the *A* gene of phage P2 (LINDAHL 1970). This protein is a regulatory protein required for normal expression of all essential genes of phage P2. The probable site(s) at which the *A* protein acts are close to the *A* gene (LINDAHL 1970). For phage λ,

* This analysis assumes that the dissociation rate for the specific regulatory site is much less than that for the nonspecific sites, so that the nonspecific sites will establish equilibrium before dissociation occurs from the specific site.

the *O* protein required for DNA replication appears to be partially *cis*-acting (KLECKNER 1974; HAYES and SZYBALSKI, personal communication), as does the *A* protein required for maturation of λ DNA (FOLKMANIS, personal communication). The genes for both of these proteins are right next to their probable site of action. Evidence (much of it quite indirect) has been presented for metabolic instability of the *O* protein (WYATT and INOKUCHI 1974), the *Q* protein (TAKEDA 1971) the *Xis* protein (WEISBERG and GOTTESMAN 1971), the *cII/cIII* proteins (REICHARDT 1975) and the *N* protein of phage λ (KONRAD 1968; SCHWARTZ 1970). Possibly some of these observations may reflect limited turnover due to nonspecific interactions. As biochemical assays for more phage proteins become available, we suspect that partial *cis*-activity and metabolic instability may become a frequent observation. One can also speculate concerning the frequent linkage of regulatory genes and sites in bacteria. Perhaps the exceptions really represent proximity of different sites in the tertiary structure of the folded chromosome!

Of many contributors to the λ lore responsible for this work, we thank particularly WILLIAM DOVE, ATIS FOLKMANIS, FRANK STAHL and RENE THOMAS for their thoughts on genetic organization and its meaning.

LITERATURE CITED

CAMPBELL, A., 1961 Sensitive mutants of bacteriophage λ Virology **14:** 22–32.

COURT, D., L. GREEN and H. ECHOLS, 1975 Positive and negative regulation by the cII and cIII gene products of bacteriophage λ. Virology **63:** 484–491.

DOVE, W. F., 1971 Biological inferences. pp. 297–312. In: *The Bacteriophage Lambda.* Edited by A. D. HERSHEY. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

ECHOLS, H., 1971 Regulation of lytic development. pp. 247–270. In: *The Bacteriophage Lambda.* Edited by A. D. HERSHEY. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

——, 1972 Developmental pathways for the temperate phage. Ann. Rev. Genet. 6: 157–190.

FRANCKE, B. and D. S. RAY, 1972 *Cis*-limited action of the gene-*A* product of bacteriophage φX174 and the essential bacterial site. Proc. Natl. Acad. Sci. U.S. **69:** 475–479.

HENRY, T. J. and R. KNIPPERS, 1974 Isolation and function of the gene *A* initiator for bacteriophage φX174, a highly specific DNA endonuclease. Proc. Natl. Acad. Sci. U.S. **71:** 1549–1553.

HERSKOWITZ, I. and E. R. SIGNER, 1970 A site essential for the expression of all late genes in bacteriophage λ. J. Mol. Biol. **47:** 545–556.

KAISER, A. D., 1957 Mutations in a temperate bacteriophage affecting its ability to lysogenize *Escherichia coli.* Virology **3:** 42–61.

KLECKNER, N., 1974 Plasmid formation by bacteriophage λ. Ph.D. Thesis. Massachusetts Institute of Technology. pp. 103–126.

KONRAD, M., 1968 Dependence of "early" λ bacteriophage RNA synthesis on bacteriophage directed protein synthesis. Proc. Natl. Acad. Sci. U.S. **59:** 171–178.

LIN, S.-Y. and A. D. RIGGS, 1975 The general affinity of *lac* repressor for *E. coli* DNA: implications for gene regulation in prokaryotes and eukaryotes. Cell **4:** 107–112.

LINDAHL, G., 1970 Bacteriophage P2: replication of the chromosome requires a protein which acts only on the genome that coded for it. Virology **42:** 522–533.

LINDQVIST, B. H. and R. L. SINSHEIMER, 1967 The process of infection with bacteriophage ϕX174. Bacteriophage DNA synthesis in abortive infections with a set of conditional lethal mutants. J. Mol. Biol. **30**: 69–80.

MORSE, D. E. and P. PRIMAKOFF, 1970 Relief of polarity in *E. coli* by "*suA*". Nature **226**: 28–31.

REICHARDT, L. R., 1975 Control of bacteriophage lambda repressor synthesis after phage infection: the role of the *N*, *c*II, *c*III, and *cro* products. J. Mol. Biol. **93**: 267–288.

SCHWARTZ, M., 1970 On the function of the N cistron in phage λ. Virology **40**: 23–33.

STAHL, F. W. and N. E. MURRAY, 1966 The evolution of gene clusters and genetic circularity in microorganisms. Genetics **53**: 569–576.

SUSSMAN, R. and H. BEN ZEEV, 1975 A proposed mechanism of bacteriophage lambda induction: acquisition of binding sites for the lambda repressor by the induced host DNA. Proc. Natl. Acad. Sci. U.S. **72**: 1973–1976.

TAKEDA, Y., 1971 Control of late messenger RNA synthesis during λ phage development. Biochim. Biophys. Acta **228**: 193–201.

TESSMAN, E., 1966 Mutants of bacteriophage S13 blocked in infectious DNA synthesis. J. Mol. Biol. **17**: 218–236.

THOMAS, R., 1964 On the structure of the genetic segment controlling immunity in temperate bacteriophages. J. Mol. Biol. **8**: 247–253. ——, 1971 Regulation of gene expression in bacteriophage λ. Curr. Top. Microbiol. Immunol. **56**: 13–42.

TOUSSAINT, A., 1969 Insertion of phage Mu.1 within prophage λ: a new approach for studying the control of the late functions in bacteriophage λ. Mol. Gen. Genet. **106**: 89–92.

VON HIPPEL, P. H., A. REVZIN, C. A. GROSS and A. C. WANG, 1974 Non-specific DNA binding of genome regulating proteins as a biological control mechanism: I. The *lac* operon: equilibrium aspects. Proc. Natl. Acad. Sci. U.S. **71**: 4808–4812.

WEISBERG, R. A. and M. E. GOTTESMAN, 1971 The stability of Int and Xis functions. In: *The Bacteriophage Lambda.* pp. 489–500. Edited by A. D. HERSHEY. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

WYATT, W. M. and H. INOKUCHI, 1974 Stability of Lambda O and P replication functions. Virology **58**: 313–315.

Corresponding editor: D. SCHLESSINGER