JOURNAL OF BACTERIOLOGY, Aug. 1986, p. 429–432
0021-9193/86/080429-04$02.00/0

Vol. 167, No. 2

cis-Acting Proteins

ELIZABETH McFALL

Department of Microbiology, New York University School of Medicine, New York, New York 10016

The *cis-trans* complementation test (3), widely used in the analysis of gene function, is predicated on the assumption that gene products are diffusible. Thus, a merodiploid cell of genotype X^+Y^-/X^-Y^+ should appear phenotypically wild type. In principle, separate genetic entities that specify products should be distinguishable on the basis of complementation in merodiploids. A mutation in a site of action, on the other hand, should not be complemented. In general these expectations have been borne out, with most apparent exceptions traceable to subunit mixing and polarity effects. Such results are to be expected with genes that are expressed efficiently and whose products act in the cytoplasm. However, it is in retrospect not so obvious that DNA-binding proteins, especially those which are specified by genes located close to their targets, should act as efficiently in *trans* as in *cis*.

DNA-binding proteins generally have high affinity for specific sites of action and lower but significant affinity for nonspecific DNA. Analyses of Berg et al. (4, 5) showed that the rates at which *lac* repressor and RNA polymerase locate their targets in *Escherichia coli* are consistent with one-dimensional random searches along the DNA. Three-dimensional random searches through the cytoplasm were ruled out on the basis of the rates of the interactions. It was suggested that DNA-binding proteins might also translocate across DNA segments, however.

Given that transcription and translation in procaryotes appear to be closely coupled (21, 29), a newly synthesized protein is presumably released in the immediate neighborhood of the parental gene. If there were a site of action on a closely linked DNA segment, it might be expected, on the basis of a one-dimensional random search, that this local target would be preferred over a more remote one. Whether *trans* sites would also be occupied effectively should depend on the numbers of molecules and subunits of the binding protein, the distance to the *trans* sites, the affinity of the protein for specific and nonspecific binding sites, the stability of the protein in the absence of its target, and the stability of the target in the absence of the protein.

In considering apparent *cis* effects, it is necessary to rule out trivial explanations. Thus, a regulatory gene cloned onto a multicopy plasmid may not be expressed simply because of structural considerations, e.g. plasmid promoters, or excessive expression may result in the accumulation of inactive multimers. Transcriptional polarity could mimic *cis* action, as Kleckner (22) showed in a study of amber mutations in λ gene *O*. She found that *O* amber mutants of N^- phages were not complemented by O^+ N^- phage. Such a result would suggest that an O^+ gene product is *cis* acting. However, nonamber *O N* mutants were complemented. This suggested rather that the polar effect of the amber codons interfered with transcriptional activation of DNA replication. Indeed, a secondary mutation downstream which relieved the amber polarity also allowed complementation, indicating that transcription of the *O P* region of a given phage is necessary for replication of that phage.

There are three well-defined classes of proteins whose action is preferentially *cis*: the *A* gene proteins of the isometric single-stranded-DNA phages and perhaps analogous proteins in other systems, the insertion sequence (IS) transposases, and two regulatory proteins, the D-serine deaminase activator and λ *Q* gene product. This review will discuss *cis* effects in the context of the functions of these proteins and will consider the circumstances that result in *cis* action.

DNA REPLICATIVE PROTEINS

The first observation of procaryotic *cis* dominance was by Tessman (35, 36). She found that complementation of gene *A* amber mutants of the single-stranded-DNA phage S13 by wild-type phage was barely detectable (35, 37). Similar results were obtained in analogous experiments with *A* protein mutants of the related phages φX174 and G4 and with amber and thermosensitive mutants for a replicative protein of the double-stranded-DNA phage P2 (14, 24, 38). The gene *A* product is a multifunctional protein which has a central role in the replication and maturation of the isometric single-stranded phage DNAs. The protein attaches to a specific site within the *A* gene itself on the covalently closed superhelical replicative form I of the phage (1, 2, 13, 15). At this site, the origin of replication, the protein introduces a single nick, converting the DNA to the relaxed replicative form II. The protein remains attached to the replication complex, cuts newly synthesized progeny DNAs into unit lengths as they emerge from the rolling-circle intermediate, and ligates them into circular viral molecules (12, 18, 19, 41). Francke and Ray (14), using ^{3}H-labeled wild-type and ^{32}P-labeled gene *A* amber mutant φX174 phage in mixed infections of sensitive cells, were able to show the *cis* effect conclusively. Regardless of multiplicity of infection, nearly all of the ^{32}P label was recovered in replicative form I DNA (not nicked), whereas the ^{3}H label was recovered almost entirely in replicative form II (nicked) and progeny DNA.

As the *A* protein is apparently synthesized in considerable excess of its replicative function and is sufficiently stable that it can promote φX174 DNA replication in vitro, its *cis* dominance is surprising. Since partially purified protein functions in vitro (12, 18, 19, 41), it can apparently retain activity while searching for its target. Tessman et al. (37) showed, however, that the protein is also required for superinfection immunity, and van der Mei et al. (39) found that it is trapped by the cell membrane, which is presumably the site of the exclusion function. Thus, unless the newly synthesized gene product binds immediately to a free DNA origin of replication, it is apparently channeled to its alternative site of action. Such a trap may be necessary to prevent formation of excessive amounts of active replicative form II. If the parental DNA is in contact with the membrane during its replication, the protein could locate the membrane by searching along it, but would presumably have to move through the cytoplasm to find another DNA molecule.

TRANSPOSASES

Several of the IS transposases, including those of IS*10* (Tn*10*), IS*50*R (Tn*5*), IS*903* (Tn*903*), and IS*1*, have been shown to act preferentially in *cis* (16, 20, 25, 28). Morisato et al. (28) provided a conclusive and quantitative demonstration of the phenomenon for the IS*10* transposase of Tn*10*. They fused the IS*10* transposase gene to the inducible p*tac* transcriptional promoter and integrated this construction in either orientation to and at varying distances from a transposition-defective Tn*10* Tc^r element. The recipient strain also contained a transposition-defective Tn*10* Km^r element, at a site remote from both the Tc^r and the transposase elements. Morisato et al. found that the transposition ratio Tc^r/Km^r varied inversely with the distance between the transposase gene and the Tc^r transposon ends. The effect was independent of the orientation of the transposase gene relative to its targets, the nature of the intervening DNA, and the level of transposase. When the amount of transposase per cell was increased by induction of the p*tac* gene, however, the absolute level of transposition of both elements increased accordingly.

In this experiment one can account for the gradient of activity of the transposase by assuming that the protein interacts more efficiently with target sites close to its parental gene than with distant ones. It cannot be explained by transcriptional activation, because neither the orientation of the transposase gene nor the nature of the intervening DNA affects the transposition ratio. The level of IS*10* transposase is very low, less than one molecule per cell in wild-type cells (33). Why, when transposase synthesis is increased by induction of the *tac* promoter, is the protein not more effective at a distance? Morisato et al. (28) suggest that it may be intrinsically unstable, like the Mu transposase (30), so that most molecules do not reach distant targets in an active form. They point out that preferential *cis* action of the transposase may be essential to prevent excessive transposition. If the rates of transposition and mutational loss of ability to transpose are similar, such a mechanism would maintain the number of functional transposable elements in the cell at a constant level.

Roberts et al. (34) have shown that there are methylation sites in the DNA of the IS*10* transposase promoter and in the inner terminus and that both sites are severalfold more active in *dam* (nonmethylating) mutants that in wild-type cells. Thus, an additional factor in the preferential *cis* action of the IS*10* transposase on IS*10* transposition in wild-type cells may be the inactivity of both the promoter and the IS*10* end, except immediately after replication when the newly synthesized DNA would be hemimethylated—that is, a target would be active just when the protein was being produced. The authors point out that the mechanism may be a more general one, as there are analogous methylation sites in other IS elements. They note that such a control on IS*10* transposition protects the cell from possible irreparable damage resulting from transposition, as transposition would occur only when the transposing region was duplicated.

D-SERINE DEAMINASE ACTIVATOR AND THE λ *Q* GENE PRODUCT

The D-serine deaminase activator protein (*dsdC* gene product) is required for the induction of D-serine deaminase (*dsdA* gene product) mRNA synthesis in wild-type *E. coli* K-12. The protein also represses its own synthesis fivefold in the absence of its ligand, D-serine. The two genes are adjacent and are divergently transcribed from a central control region (27; unpublished results). The cellular concentration of activator is extremely low; in the absence of D-serine, it is similar to that of the parental gene itself (17).

We observed an apparent *cis* action of the *dsdC*$^+$ gene when we cloned it, together with the *dsd* control region, on the plasmids pACYC177 (copy number, 20) and pUC8 (copy number, 100 to 200). D-Serine deaminase synthesis was highly inducible in Δ(*dsdA dsdC*) strains bearing plasmid pAC131 (*dsdA*$^+$ *dsdC*$^+$) and in *dsdA*$^+$ *dsdC*$^+$ strains bearing plasmid pAC51 (Δ*dsdA dsdC*$^+$). However, we could detect no induced synthesis in a *dsdA*$^+$ Δ*dsdC* strain bearing plasmid pAC51 or in a Δ*dsdA dsdC*$^+$ strain bearing plasmid pAC53 (*dsdA*$^+$ Δ*dsdC*) (6, 17; unpublished results). Moreover, constitutive D-serine deaminase synthesis was not repressed in either of two *dsdA*$^+$ *dsdC*^{c} strains (*dsdC*^{c} denotes a *dsdC* allele that results in a constitutive phenotype) bearing *dsdC*$^+$ plasmids (unpublished results). Plasmid-borne activator genes program activator synthesis in vivo and in vitro. We have used *dsdC*$^+$ plasmids to enhance activator synthesis in vivo (17); the gene is expressed as efficiently from plasmids as from the chromosome. The purified protein activates D-serine deaminase synthesis in a cell-free system (8, 17); thus, it does not seem to be unstable in the absence of its target. Therefore, the activator apparently does not diffuse freely between plasmids and the chromosome.

The *cis* effect observed with plasmids is in contrast to the usual situation with the F′-borne *dsdC*$^+$ gene. The introduction of an F′ *dsdC*$^+$ gene into a *dsdA*$^+$ *dsdC*^{c} strain results in a 3- to 10-fold decrease in the rate of constitutive D-serine deaminase synthesis in both *recA*$^+$ and *recA* strains, if the pattern of F′-mediated transfer of chromosomal genes in the merodiploid is normal (6, 26). This *trans* effect of the *dsdC*$^+$ gene has been attributed to repression of *dsdC*^{c} expression by the wild-type activator (17). However, in one *dsdA*$^+$ *dsdC*^{c}/F′ *dsdA*$^+$ *dsdC*$^+$ merodiploid there was no *trans* effect of the *dsdC*$^+$ gene on constitutivity (26). This merodiploid, in contrast to the others, showed an aberrant gradient of chromosomal transfer, indicating that the F′ plasmid is not paired with the chromosomal *dsd* region but might be integrated to its left. The merodiploid was shown to contain the *dsdC*$^+$ gene because it could transfer the gene (26). We know that *dsdC* expression is very weak, such that there are only a few molecules of activator per wild-type cell (17). We think that these few molecules are preferentially bound by *dsd* promoter sites in the immediate vicinity of their synthesis. There is some evidence that this is the case and that there may be multiple activator-binding sites in the *dsd* control region. S. Bernstein (unpublished results) cloned a 1.1-kilobase *Sph*I-*Eco*RI *dadC*$^+$ fragment into the *Sph*I-*Eco*RI sites of pUC18 and pUC19 (copy number, about 100) (40). The cloned fragment lacks a major part of the *dsd* control region, the 300 base pairs adjacent to the *dsdA* gene. Introduction of the hybrid plasmids containing this region into the same *dsdA*$^+$ *dsdC*^{c} strains described above resulted in a two- to fivefold decrease in the constitutive rate of synthesis. Since removal of control region DNA evoked such a *trans* effect, it may be that the deleted DNA contains one or more activator target sequences. In the wild-type cell, newly formed activator would be quickly trapped by such nearby binding sites and would not be available to distal sites.

Echols et al. (11) demonstrated the apparently preferential *cis* action of another regulatory protein, the λ *Q* gene protein, which activates expression of adjacent late genes.

They found that synthesis of endolysin, the product of the late gene *R*, was about eightfold lower in cells mixedly infected with Q^-am R^+ and Q^+ R^- phage than in cells infected with Q^-am R^- and Q^+ R^+ phage. The result was the same when the host strain carried the polarity suppressor SuA. In all cases the yields of the two types of infecting phage were similar. The authors noted the close proximity of the λ *Q* gene to its site of action and suggested that such an arrangement might be advantageous, in that it should result in a locally high concentration of a regulatory protein near its target. They considered that the preferential *cis* action of the *Q* gene product might be due to failure to reach *trans* sites because of instability or excessive nonspecific binding to intervening DNA. Recent work of M. Benedik (personal communication), however, suggests that DNA structure in the vicinity of the *Q* gene may be involved in the effect.

CONCLUSIONS

Many genes that specify DNA-binding proteins are adjacent to target sites of the gene products, which, as Echols et al. have pointed out (11), suggests that such arrangements have some functional significance. Well-characterized examples include the *lac, hut,* λ repressor, and λ *cro* genes and the *ara* and *mal* activators as well as the cases described above. Most, including the *lac* repressor gene of *Klebsiella pneumoniae* (7), are either part of a regulated operon or are transcribed with opposite polarity from that of the target operon. The fact that synthesis of most of the proteins is weakly autoregulated, although the effect of the regulation on total protein synthesis must be miniscule, is also thought provoking. The existence of autoregulatory sites means that these proteins are always bound to DNA near at least some of their targets, whether their functions are required or not. Even the *E. coli lac* repressor, whose synthesis is apparently not regulated, has secondary binding sites in *lacI* and *lacZ* which may increase the efficiency of *lac* repressor-operator interaction (31). Such a tracking mechanism could make regulatory proteins that are present in the cell only in very small amounts readily available to their targets.

Why, given the apparently felicitous location of many regulatory genes with respect to their targets, have not more preferential *cis* effects been observed? Certainly one very important reason is that nearly all of the proteins must also act at a distance, to regulate unlinked sets of genes as well as adjacent ones. Thus, the *ara* activator must be able to activate expression of an unlinked permease gene as well as the adjacent *araBAD* operon (23), the λ repressor must be able to act in *trans* to block superinfection of lysogens (32), etc. As might be expected, the proteins that act at a distance are produced in significantly greater amounts than the IS*10* transposase (33) or the D-serine deaminase activator (17). There are estimated to be more than 100 molecules each of the *ara* and *mal* activators per cell (9, 10), for example. Another reason may be that it is very difficult to detect moderately preferential *cis* effects. The three cases discussed here are fairly stringent. The phage S13 *A* protein seems to be efficiently trapped in the absence of its target (39). The synthesis of IS*10* transposase and D-serine deaminase activator in wild-type cells is very inefficient, and the proteins are not required to act at a distance (17, 33). Moreover, at least some IS transposase ends are probably not active except when transposase is being synthesized (34). In intermediate situations, with more protein synthesized, less efficient traps, or both, significant amounts of protein may diffuse to distal sites. Thus, although there might be severalfold more molecules available to an adjacent site than to a distal one, both sites could still be saturated, and no *cis* effect would be apparent. Distance might also be partially compensated for by the presence of secondary binding sites, at distant promoters, that maintain significant local concentrations of proteins at strategically located parking places.

ACKNOWLEDGMENTS

I am grateful to M. Benedik and N. Kleckner for communication of unpublished and in-press material.

This work was supported by grant DMB-8411383 from the National Science Foundation.

LITERATURE CITED

1. **Baas, P. D., and H. S. Jansz.** 1972. φX174 replicative form DNA replication, origin and direction. J. Mol. Biol. **63**:569–576.
2. **Baas, P. D., and H. S. Jansz.** 1978. Replication of φX174 RF in vivo, p. 215–243. *In* D. T. Denhardt, D. Dressler, and D. S. Ray (ed.), The single-stranded DNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
3. **Benzer, S.** 1955. Fine structure of a genetic region in bacteriophage. Proc. Natl. Acad. Sci. USA **41**:344–354.
4. **Berg, O. G., and P. H. von Hippel.** 1985. Diffusion controlled macromolecular interactions. Ann. Rev. Biophys. **14**:131—160.
5. **Berg, O. G., R. B. Winter, and P. H. von Hippel.** 1982. How do genome-regulatory proteins locate their DNA target sites? Trends Biochem. Sci. **7**:52–55.
6. **Bloom, F. R., and E. McFall.** 1975. Isolation and characterization of D-serine deaminase constitutive mutants by utilization of D-serine as sole carbon or nitrogen source. J. Bacteriol. **121**:1078–1084.
7. **Buvinger, W. E., and M. Riley.** 1985. Regulatory region of the divergent *Klebsiella pneumoniae lac* operon. J. Bacteriol. **163**:858–862.
8. **Carothers, A. M., M. C. Heincz, and E. McFall.** 1980. Position effect on expression of *dsd* genes cloned onto multicopy plasmids. J. Bacteriol. **142**:185–190.
9. **Casadaban, M. J.** 1976. Regulation of the regulatory gene for the arabinose pathway, *araC*. J. Mol. Biol. **104**:557–566.
10. **Debarbouille, M., and M. Schwartz.** 1979. The use of gene fusions to study the expression of *malT*, the positive regulatory gene of the maltose regulon. J. Mol. Biol. **132**:521–534.
11. **Echols, H., D. Court, and L. Green.** 1976. On the nature of *cis*-acting regulatory proteins and genetic organization in bacteriophage: the example of gene *Q* of bacteriophage λ. Genetics **83**:5–10.
12. **Eisenberg, S., J. Griffith, and A. Kornberg.** 1977. φX174 *cistron A* protein is a multifunctional enzyme in DNA replication. Proc. Natl. Acad. Sci. USA **74**:3198–3202.
13. **Eisenberg, S., J. T. Scott, and A. Kornberg.** 1978. An enzyme system for replicating the duplex replicative form of φX174 DNA, p. 287–302. *In* D. T. Denhardt, D. Dressler, and D. S. Ray (ed.), The single-stranded DNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **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. USA **69**:475–479.
15. **Godson, G. N., J. C. Fiddes, B. G. Barrell, and F. Sanger.** 1978. Comparative DNA sequence analysis of the G4 and φX174 genomes in the single stranded DNA phages, p. 51–86. *In* D. T. Denhardt, D. Dressler, and D. S. Ray (ed.), The single-stranded DNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Grindley, N. T., and C. M. Joyce.** 1981. Analysis of the structure and function of the kanamycin-resistance transposon Tn*903*. Cold Spring Harbor Symp. Quant. Biol. **45**:125–133.
17. **Heincz, M. C., S. M. Bornstein, and E. McFall.** 1984. Purification and characterization of the D-serine deaminase activator protein. J. Bacteriol. **160**:42–49.
18. **Henry, T. J., and R. Knippers.** 1974. Isolation and function of

the gene *A* initiator of bacteriophage ϕX174, a highly specific DNA endonuclease. Proc. Natl. Acad. Sci. USA **71**:1549–1553.
19. **Ikeda, J.-E., A. Yudelevich, and J. Hurwitz.** 1976. Isolation and characterization of the protein coded by gene A of bacteriophage ϕX174. Proc. Natl. Acad. Sci. USA **73**:2669–2673.
20. **Isberg, R. S., and M. Syvanen.** 1981. Replicon fusion promoted by the inverted repeats of Tn*5*. J. Mol. Biol. **150**:15–32.
21. **Kano, Y., M. Kuwano, and F. Imamoto.** 1976. Initial *trp* operon sequence in *Escherichia coli* is transcribed without coupling to translation. Mol. Gen. Genet. **146**:179–188.
22. **Kleckner, N.** 1977. Amber mutants in the *O* gene of bacteriophage λ are not efficiently complemented in the absence of phage *N* function. Virology **79**:174–182.
23. **Lee, N.** 1980. Molecular aspects of *ara* regulation, p. 389–409. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **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.
25. **Machida, Y., C. Machida, H. Ohtsubo, and E. Ohtsubo.** 1982. Factors determining frequency of plasmid cointegration mediated by insertion sequence IS*1*. Proc. Natl. Acad. Sci. USA **79**:277–281.
26. **McFall, E.** 1967. "Position effect" on dominance in the D-serine deaminase system of *Escherichia coli* K-12. J. Bacteriol. **94**:1989–1993.
27. **McFall, E., and M. C. Heincz.** 1983. Identification and control of synthesis of the *dsdC* activator protein. J. Bacteriol. **153**:872–877.
28. **Morisato, D., J. C. Way, H.-J. Kim, and N. Kleckner.** 1983. Tn*10* transposase acts preferentially on nearby transposon ends in vivo. Cell **32**:799–807.
29. **Oxender, D. T., G. Zurawski, and C. Yanofsky.** 1979. Attenuation in the *Escherichia coli* tryptophan operon: role of RNA secondary structure involving the tryptophan codon region. Proc. Natl. Acad. Sci. USA **76**:5524–5528.
30. **Pato, M. L., and C. Reich.** 1982. Instability of transposase activity: evidence from bacteriophage Mu DNA replication. Cell **29**:219–225.
31. **Pfahl, M., V. Guldi, and S. Bourgeois.** 1979. "Second" and "third operator" of the *lac* operon: an investigation of their role in the regulatory mechanism. J. Mol. Biol. **127**:339–344.
32. **Ptashne, M.** 1980. λ repressor function and structure, p. 325–343. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Raleigh, E. A., and N. Kleckner.** 1986. IS*10* transposase is synthesized in very small amounts. Proc. Natl. Acad. Sci. USA **83**:1787–1791.
34. **Roberts, D., B. C. Hoopes, W. R. McClure, and N. Kleckner.** 1985. IS*10* transposition is regulated by DNA adenine methylation. Cell **43**:117–130.
35. **Tessman, E. S.** 1965. Complementation groups in phage S13. Virology **25**:303–321.
36. **Tessman, E. S.** 1966. Mutants of bacteriophage S13 blocked in infectious DNA synthesis. J. Mol. Biol. **17**:218–236.
37. **Tessman, E. S., M.-T. Borras, and I. L. Sun.** 1971. Superinfection in bacteriophage S13 and determination of the number of bacteriophage particles which can function in an infected cell. J. Virol. **8**:111–120.
38. **Tessman, E. S., and I. Tessman.** 1978. The genes of the isometric phages and their functions, p. 9–29. *In* D. T. Denhardt, D. Dressler, and D. S. Ray (ed.), The single-stranded DNA phages. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **van der Mei, D. J. Zandberg, and H. S. Jansz.** 1972. The effect of chloramphenicol on synthesis of ϕX174-specific proteins and the detection of the cistron A proteins. Biochim. Biophys. Acta **287**:312–320.
40. **Yanisch-Perron, C., J. Viera, and J. Messing.** 1983. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
41. **Zipursky, S. L., D. Reinberg, and J. Hurwitz.** 1980. In vitro DNA replication of recombinant plasmid DNAs containing the origin of progeny replicative form DNA synthesis of phage ϕX174. Proc. Natl. Acad. Sci. USA **77**:5182–5186.