Proc. Natl. Acad. Sci. USA
Vol. 74, No. 4, pp. 1511–1515, April 1977
Biochemistry

Characterization of the integration protein of bacteriophage λ as a site-specific DNA-binding protein

(bacteriophage λ *int* gene/viral integration/site-specific genetic recombination)

MICHAEL KOTEWICZ, STEPHEN CHUNG, YOSHINORI TAKEDA, AND HARRISON ECHOLS

Department of Molecular Biology, University of California, Berkeley, California 94720

Communicated by H. Fraenkel-Conrat, January 31, 1977

ABSTRACT **The Int protein specified by bacteriophage λ is required for the recombination event that integrates the viral DNA into the host genome at its specific attachment site. Using a DNA-binding assay, we have partially purified the Int protein and studied some of the features of its binding specificity and regulation. The DNA-binding activity is attributed to Int protein because the activity is eliminated by a nonsense mutation or a deletion in the *int* gene, and is rendered thermolabile by temperature-sensitive mutations in the *int* gene. The DNA-binding activity is specific for DNA carrying an appropriate attachment site, suggesting that Int protein directs the sequence-specific recognition essential for integrative recombination. The specific DNA-binding activity is also missing after infection by phage carrying mutations in the *c*II and *c*III regulatory genes of λ. This finding corroborates the conclusion from other types of experiments that regulation of the *int* and *c*I genes by cII/cIII provides for coordinate regulation of both major events of the lysogenic response, establishment of repression and insertion of viral DNA.**

Stable lysogeny by bacteriophage λ results from two major events: repression of most of the viral functions required for lytic development and the integration of the viral DNA into the host DNA at a specific site (for reviews, see refs. 1, 2). The insertion of λ DNA occurs through a recombination event (3), catalyzed by a site-specific pathway of genetic recombination (4, 5) (Fig. 1). Integrative recombination requires the product of the λ *int* gene (6–8). Prophage induction involves an overall reversal of the process of lysogeny: release of repression and excision of the viral DNA (Fig. 1). Excisive recombination differs from integrative recombination in its requirement for the product of the λ *int* and *xis* genes (9–11). Although these general features of site-specific recombination have been known for some time from genetic experiments, the absence of a biochemical analysis has impeded a further understanding of the basis for specificity and the mechanism regulating the direction of the reaction.

The Int protein has been identified radiochemically by acrylamide gel electrophoresis of labeled proteins under denaturing conditions (12, 13). More recently, assays for the complete site-specific recombination reaction *in vitro* have been described (14–16). We have developed a complementary biochemical approach involving a direct DNA-binding assay for Int protein. This assay permits a direct analysis of the role of binding recognition in the specificity of the recombination reaction and also provides a convenient biochemical approach to the study of how site-specific recombination is regulated.

MATERIALS AND METHODS

Bacteriophage and Bacteria. The *Escherichia coli* strains used and their relevant genetic characteristics were: C600 Su^+ (suppressor-positive) (17); C600 Su^- (suppressor-negative) (18). The phage point mutations used were: *c*I*ts*857 (19), *c*I*am*14 (20), *c*II28 (18), *c*III*am*611 (18), *Qam*21 and *Ram*60 (17), *intts*1 (7), *int*C226 (21), *intts*2001 and *intts*2004 (38), *intam*29 (9, 22), *xis*1 and *xis*6 (9), *Sam*7 (23). The phage deletion and/or substitution mutations used were: *b*538, deleting the phage attachment site (24); *b*522, deleting the *int* and *xis* genes but not the phage recombination genes (24); *gal*8, carrying the bacterial *gal* operon and the left prophage attachment site *ba'* (25); *bio*7–20, carrying the bacterial *bio* operon and the right prophage attachment site *ab'* (26); *gal*8*bio*7–20, carrying both the *gal* and *bio* genes and the bacterial attachment site *bb'* (1, 2, 10 and Fig. 1).

Preparation of Phage [^{32}P]DNA. ^{32}P-labeled phage DNAs were prepared as described previously (27, 28). In brief, cells were grown in 100 ml of low-phosphate medium to a density of 5×10^8 cells per ml, concentrated by centrifugation, and infected with phage carrying the appropriate attachment site and the lysis-defective mutation *Sam*7 at a multiplicity of 5 phage/cell. After a 15 min adsorption period, the infected cells were poured into fresh low-phosphate medium containing 5 μCi/ml of ^{32}P-labeled inorganic phosphate. The cells were shaken at 37° for 3 hr and then collected by centrifugation and resuspended in 10 ml of 10 mM Tris·HCl, pH 7.2/10 mM $MgCl_2$. The cells were lysed with chloroform, the debris was removed by centrifugation, and the supernatant fraction containing the phage was centrifuged to equilibrium in a CsCl density gradient. The phage band was removed with a syringe and the phage were again centrifuged to equilibrium in CsCl. The phage DNA was extracted by four treatments with redistilled phenol, and the phenol was removed with ether. The DNA was then dialyzed extensively into 10 mM Tris·HCl, pH 7.2/0.1 mM EDTA.

DNA-Binding Assay. The details of the DNA-binding assay used for Int protein have been described previously (27–29). The assay measures the retention of ^{32}P-labeled DNA to nitrocellulose filters (B-6, Schleicher and Schuell). The assay mixture contains an excess of unlabeled "chicken blood" DNA (Calbiochem), which competes with the labeled phage DNA for proteins with nonspecific DNA-binding activity (30). Each assay contains 0.4 μg of phage [^{32}P]DNA, 20 μg of "chicken blood" DNA, and is 10 mM Tris·HCl, pH 7.2/70 mM KCl/10 mM $MgCl_2$/0.2 mM dithiothreitol/0.2 mM EDTA, in a total reaction volume of 100 μl. The various DNA preparations used gave 5 to 15% binding without any added extract and the binding data were corrected for this background.

Purification of Int Protein. For most preparations, a 1 liter culture was grown to a density of 5×10^8 cells/ml in 1% Difco tryptone/0.5% Difco yeast extract/0.5% NaCl/0.2% maltose. The cells were collected by centrifugation in 50 ml of 10 mM Tris·HCl, pH 7.2/10 mM $MgCl_2$, and infected at a multiplicity of 10 phage/cell. After an adsorption period of 15 min at room temperature, the infected cells were poured into 1 liter of fresh medium and shaken for 20 min at 37°. For experiments involving temperature-sensitive mutations in the *int* gene, growth

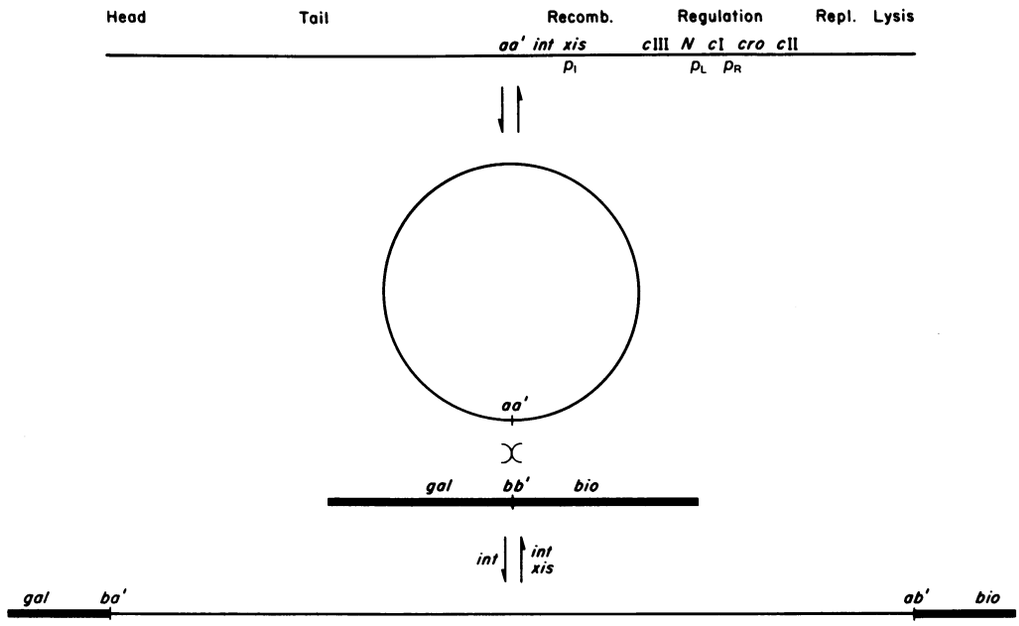

FIG. 1. The phage λ genome and the integration-excision reaction. The linear λ DNA and some pertinent genes are indicated at the top. After infection, λ DNA circularizes and is transcribed sequentially under a variety of regulatory influences. Viral development begins with transcription mainly of the *N* and *cro* genes from the early promoters p_L and p_R; N protein extends leftward transcription to the *c*III and recombination (Recomb.) genes and rightward transcription to the *c*II and replication (Repl.) genes and the late regulatory gene *Q*; in a lytic response Q protein activates transcription of the lysis, head, and tail genes. Lysogenic development is dependent on the capacity of the cII/cIII proteins to activate transcription of the *c*I gene and to repress late lytic transcription; the cI protein maintains lysogeny through a repression of transcription from p_L and p_R and activation of its own further synthesis. Stable lysogenization is also dependent on integrative recombination between the phage attachment site *aa'* and the host attachment site *bb'*, catalyzed by Int function. The inserted prophage can be detached by excisive recombination between the prophage attachment sites *ba'* and *ab'*, catalyzed by Int and Xis functions. The *gal* and *bio* host genes flank the bacterial attachment site.

was for 30 min at 30°. The culture was then chilled, and the cells were collected by centrifugation, resuspended in 11 ml of 50 mM Tris·HCl, pH 7.2/10% sucrose, and quickly frozen in a dry ice/alcohol bath. The frozen cells were thawed in a 20° water bath and then cooled to 0°. For cell lysis, 1.1 ml of a solution containing 250 mM Tris·HCl, pH 7.2/0.1 mM EDTA, and lysozyme at 2 mg/ml was added, and the mixture was brought to 0.4 M KCl. Subsequent to the experiments reported here, we have found that bringing the lysis mixture to 1.0 M KCl improves the yield of binding activity. After 40 min with occasional swirling, the lysate was warmed to 32° and 15 μg of DNase I was added. After 3–5 min the mixture was cooled to 0° for an additional 15 min. Cellular debris and ribosomes were removed by centrifugation for 3 hr at 30,000 rpm in a Spinco 30 rotor. The supernatant fraction was dialyzed for 2 hr against three changes of buffer A [10 mM KPO_4, pH 6.4/0.2 mM EDTA/0.2 mM dithiothreitol/5% glycerol (vol/vol)], containing 0.1 M KCl. The extract was then applied to a 2.5 × 10 cm phosphocellulose column (Whatman P11) equilibrated with buffer A. The column was washed with 2 bed volumes of buffer A containing 0.1 M KCl, and then a linear gradient of KCl from 0.1 M to 1.5 M (60 ml) was applied to the column.

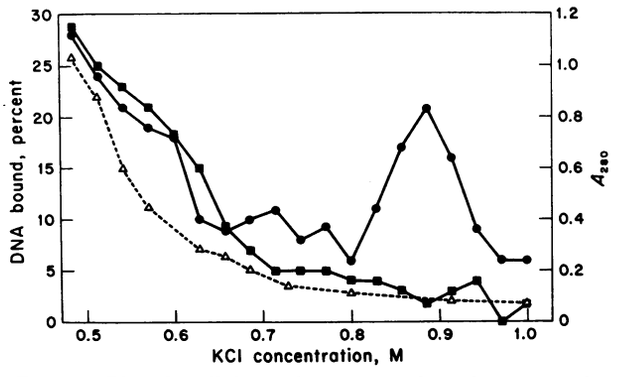

FIG. 2. Phosphocellulose chromatography of extracts from phage-infected cells. Fractions were assayed for DNA-binding activity for λ DNA as described in *Materials and Methods*. (●), DNA-binding activity for an *int*$^+$ infection of C600 Su$^-$; (■), DNA-binding activity for *intam*29 infection of C600 Su$^-$; (△), A_{280}. The A_{280} pattern was virtually identical for the two extracts (and others noted in the *text*).

The DNA-binding activity specific for DNA containing the attachment site for integrative recombination eluted between 0.8 and 1.0 M KCl. The active fractions were pooled and purified further by velocity sedimentation in a 10 to 30% (vol/vol) glycerol gradient (0.2 ml/tube). Centrifugation was carried out for 24 hr at 4° in a Spinco SW 50.1 rotor at 50,000 rpm. The gradient tubes were pierced at the bottom and 10 drop fractions were collected and assayed for DNA-binding activity. Marker proteins of known molecular weight were run in parallel.

RESULTS

Identification of Int Protein. When extracts of cells infected by wild-type phage were fractionated by chromatography on a phosphocellulose column, a peak of DNA-binding activity eluted between 0.8 and 1.0 M KCl (Fig. 2). If the phage carried a nonsense mutation in the *int* gene, this peak of DNA-binding activity was not found after infection of a nonpermissive (Su$^-$) host (Fig. 2), but was present after infection of a permissive (Su$^+$) host (data not shown). The same very low level of "background" binding activity shown for the nonsense mutation in Fig. 2 was observed in other control experiments with an *int*$^-$ deletion (*b*522) or uninfected cells.

To provide additional evidence that the DNA-binding activity was associated with Int protein, the thermosensitivity of

Table 1. Thermosensitivity of DNA-binding for Int protein from *intts* mutants

Source of Int protein	Percent DNA bound at 35°	Percent DNA bound at 20°	Temperature sensitivity (35°/20° ratio)
int$^+$	40	30	1.3
*intts*2004	26	58	0.45
*intts*2001	18	31	0.58
*intts*1	50	76	0.66

DNA-binding assays were carried out as described in *Materials and Methods*. The binding mixture was incubated for 5 min at either 35° or 20° and then filtered. In the range of 35°–40°, the wild-type Int protein begins to exhibit temperature sensitivity for DNA-binding. In nine experiments at different levels of Int protein, the maximum variability in temperature sensitivity was about ±25% of that shown in the table. Each Int preparation showed specificity for the phage attachment site. The protein concentrations used for the experiment shown in the table were: 0.5 μg for *int*$^+$, 1.0 μg for *intts*2004, 1.0 μg for *intts*2001, and 1.5 μg for *intts*1. All the infecting phage except *intts*1 carried the additional *int*C226 mutation in the *xis* gene [used because it may increase the levels of Int protein produced (34, 37)]; the double-mutant phage were constructed by Lynn Enquist (personal communication).

the binding interaction was studied for extracts of cells infected with phage carrying the temperature-sensitive mutations *intts*1 (7), *intts*2001, or *intts*2004 (38). After phosphocellulose chromatography, the DNA-binding activity showed thermolability for all three mutants, compared to the activity found after *int*$^+$ infection (Table 1). From the results presented in Fig. 2 and Table 1, we conclude that the DNA-binding activity that we are studying is specified (at least in part) by the *int* gene. In the following sections, we will term the DNA-binding activity Int protein, although we have not shown that the product of the *int* gene is the sole protein species participating in the binding interaction (this seems highly likely from the molecular weight estimate described below).

Regulation of Int Protein Production by the Phage *c*II/*c*III Gene Products. When extracts were prepared after infection by phage carrying mutations in the *c*II and *c*III regulatory genes known to control the establishment of repression (29, 31–33), the DNA-binding activity associated with Int protein was not found (Fig. 3). This result is consistent with other experiments showing that active *c*II and *c*III genes are

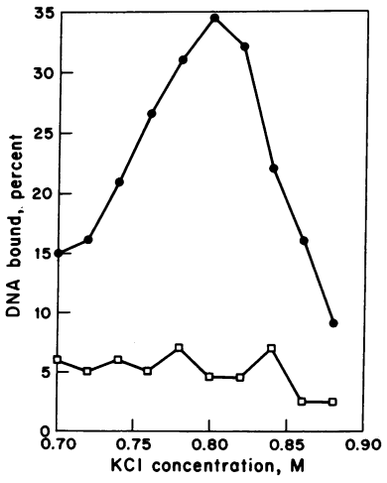

FIG. 3. Phosphocellulose chromatography of extracts prepared after infection by *int*$^+$ phage isogenic except for *c*II and *c*III. (●), DNA binding for *c*II$^+$*c*III$^+$; (□), DNA binding for *c*II$^-$*c*III$^-$.

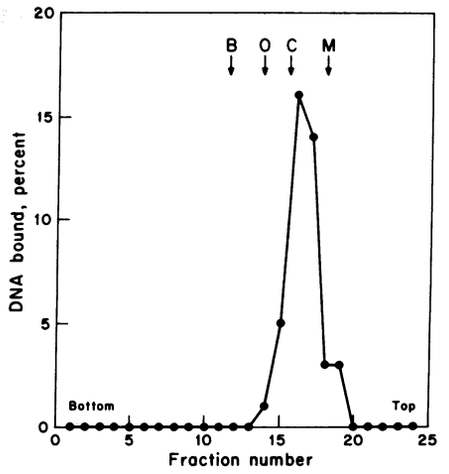

FIG. 4. Sedimentation properties of Int protein. The DNA-binding activity obtained from phosphocellulose chromatography was sedimented for 24 hr at 50,000 rpm in a 10–30% glycerol gradient containing 0.4 M KCl. Fractions were collected and assayed for DNA-binding activity as described in *Materials and Methods*. Sedimentation was from right to left. Arrows designate the position of marker proteins (B = bovine serum albumin, molecular weight 65,000; O = ovalbumin, 45,000; C = chymotrypsinogen, 27,000; and M = myoglobin, 17,000). (●), DNA binding to λ DNA.

required for the production of a radioactive protein identified as Int after fractionation of labeled proteins by polyacrylamide gel electrophoresis in sodium dodecyl sulfate (34, 35). We conclude that the cII/cIII proteins of phage λ exert positive regulation on the *int* gene, in addition to their previously defined regulatory activities on the *c*I gene and late lytic functions (29, 32, 33).

Sedimentation Properties of Int Protein. To obtain an estimate of the native molecular weight of Int protein, the phosphocellulose fraction was subjected to velocity sedimentation in a glycerol gradient. A single peak of DNA-binding activity was observed (Fig. 4); this activity could be associated with Int protein because of the binding specificity for DNA carrying the appropriate attachment site for integrative recombination (see next section). If the Int protein were a typical globular protein, its sedimentation with respect to proteins of known molecular weight would indicate a molecular weight of about 25,000. This is less than the apparent monomer molecular weight of 36–42,000 estimated from acrylamide gel electrophoresis of radioactive λ proteins in sodium dodecyl sulfate (13, 14, 22, 34, 35). Although several possibilities exist, we believe the most likely explanation is that native Int protein is a monomer with an asymmetric shape.

Binding Specificity of Int Protein. There are four different attachment sites with which the Int and Xis proteins must interact: *att*P (phage *aa'*), *att*B (host *bb'*), *att*L (prophage *ba'*), and *att*R (prophage *ab'*) (see introduction and Fig. 1). We have studied the binding of Int protein to DNA carrying each of these four attachment sites and to DNA with a deletion of the attachment site or a different attachment site (for phage 80) (Fig. 5). The data presented in Fig. 5*A* show that the binding is specific for an appropriate attachment site. From this, we conclude that the sequence-specific DNA recognition essential for integrative recombination derives from the binding specificity of Int protein.

The data presented in Fig. 5*B* show that Int protein also exhibits binding specificity for the host attachment site *att*B and the left prophage attachment site *att*L, but binds poorly to the right prophage attachment site *att*R. These results indicate that the inability of Int protein to catalyze excisive recombination

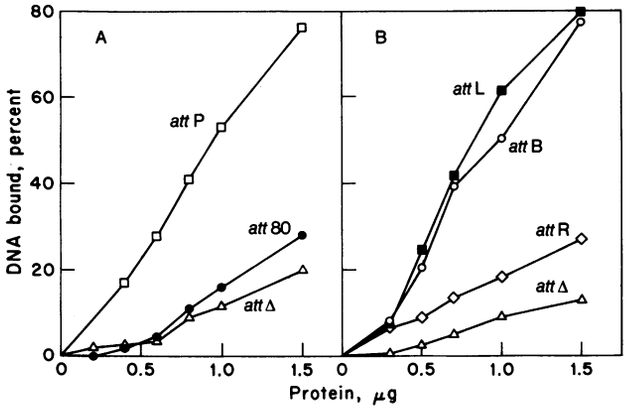

FIG. 5. DNA-binding specificity of Int protein. For each DNA, various amounts of the phosphocellulose fraction of Int protein were added and the amount of [^{32}P]DNA retained on the nitrocellulose filters was measured. (□), DNA binding for a normal phage attachment site; (△), an attachment site deletion (*b*538); (●), a different attachment site (phage 80); (■), the prophage left attachment site (*gal*8); (◇), the prophage right attachment site (*bio*7–20); (○), the host attachment site (*gal*8*bio*7–20).

(*att*L × *att*R) in the absence of Xis may derive from limitations in binding capacity (see *Discussion*).

Although we cannot exclude the possibility that other phage or host proteins are involved with the *int* gene product in conferring the specificity for recognition of attachment sites, we can say that such a hypothetical "Int oligomer" must exist as a very tight complex because the binding specificity is retained through several stages of purification. Virtually identical DNA binding specificities have been obtained for Int preparations purified by phosphocellulose only or subjected to additional steps of Sephadex gel filtration (or glycerol gradient) and DNA-cellulose chromatography. The site-specific DNA binding probably does not require the Xis protein because the binding activity is unaffected by the *xis* mutations *xis*1 and *xis*6 (9) and *int*C226 (21).

To try to ensure that the binding differences observed in Fig. 5 reflect real differences in sequence recognition for attachment sites, the binding experiments have been done with two sets of independently prepared DNA samples, with essentially identical results. The DNAs used for assay have also been checked for binding to λ cI protein as a measure of the capacity of the DNA to show specific binding to other sequences in the same filter assays; the binding curves for cI protein were indistinguishable (except for phage 80 DNA, which does not bind λ cI protein).

DISCUSSION

We have identified a phage-specified DNA-binding activity that we believe to derive from Int protein on the basis of three major criteria: (*a*) the activity is not found after infection by phage carrying nonsense or deletion mutations of the *int* gene; (*b*) the activity is thermolabile in extracts of cells infected by phage with temperature-sensitive mutations in the *int* gene; (*c*) the binding is specific for DNA carrying an appropriate attachment site. Our experiments provide some preliminary indications of possible mechanisms for two features of site-specific recombination noted in the introduction: specificity and regulation. These are discussed below.

Specificity in Integrative and Excisive Recombination. The binding specificity of Int protein indicates that sequence recognition by this protein is crucial for the site specificity of integrative recombination. The binding results also indicate a possible basis for the role of Xis protein in excisive recombination: Int alone does not bind effectively to the right prophage attachment site and thus Xis might be required for efficient recognition of this sequence. The binding results show an interesting correlation with integrative recombination *in vivo*, in which mutational loss of active Xis protein has a much more pronounced effect on recombination events involving *att*R than those involving other combinations (e.g., *att*L × attR and *att*P × *att*R are more Xis-dependent that *att*P × *att*B or *att*P × *att*L) (see ref. 36 for summary). However, the differential binding that we have observed *in vitro* and the Xis hypothesis above do not provide a complete explanation for the results *in vivo* because *att*P × *att*R recombination is less Xis-dependent than *att*L × *att*R, a result not predicted by our binding results*. Characterization of Xis activity *in vitro* will probably be necessary before firm conclusions can be drawn on this point.

Regulation of Site-Specific Recombination. Our experiments, together with those of others (34, 35), show that production of Int protein is subject to positive regulation by the cII/cIII proteins. Genetic experiments have also shown that integrative recombination is under cII/cIII control (35, 39). Thus both major events required for the establishment of lysogeny, repression of lytic functions and integrative recombination, are coordinately regulated by cII/cIII.

This regulatory feature provides not only a coordination mechanism for lysogenization functions, but probably also the basis for directional control of site-specific recombination toward integration or excision. Excisive recombination is not under cII/cIII control (39), indicating differential regulation of the *int* gene with respect to *xis*. This possibility has been suggested previously on the basis of evidence for a weak constitutive promoter for *int* within the recombination region, perhaps within the *xis* gene (p_I of Fig. 1) (21, 37). From this analysis the most likely interpretation of the regulatory process is that cII/cIII stimulate transcription of the *int* gene, acting at a site within th *xis* gene (or just to the left of *xis*). The mechanism might involve activation of the p_I promoter or antitermination of a p_I "leader" RNA.

We thank Robert Fischer, Atis Folkmanis, and Tim Yu for help with the experiments and Sankar Adhya, Donald Court, Lynn Enquist, Dietmar Kamp, and Howard Nash for the gift of strains, valuable discussion, and the communication of unpublished results. This research was supported in part by U.S. Public Health Service Grant GM 17078 from the National Institute of General Medical Sciences.

1. Gottesman, M. E. & Weisberg, R. A. (1971) in *The Bacteriophage Lambda*, ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 113–138.
2. Echols, H. (1971) *Annu. Rev. Biochem.* **40**, 827–854.
3. Campbell, A. (1962) *Adv. Genet.* **11**, 101–145.
4. Echols, H., Gingery, R. & Moore, L. (1968) *J. Mol. Biol.* **34**, 251–260.
5. Weil, J. & Signer, E. R. (1968) *J. Mol. Biol.* **34**, 273–279.
6. Zissler, J. (1967) *Virology* **31**, 189.
7. Gingery, R. & Echols, H. (1967) *Proc. Natl. Acad. Sci. USA* **58**, 1507–1514.
8. Gottesman, M. E. & Yarmolinsky, M. B. (1968) *J. Mol. Biol.* **31**, 487–505.
9. Guarneros, G. & Echols, H. (1970) *J. Mol. Biol.* **47**, 565–574.
10. Echols, H. (1970) *J. Mol. Biol.* **47**, 575–583.
11. Kaiser, A. D. & Masuda, T. (1970) *J. Mol. Biol.* **47**, 557–564.
12. Nash, H. A. (1974) *Nature* **247**, 543–545.

* Experiments *in vivo* have been complicated by the differential thermosensitivity of various types of site-specific recombination, probably accounting for some disagreement in reported values for Xis-dependence of *att*P × *att*R.

13. Ausubel, F. M. (1974) *Nature* **247**, 152–154.
14. Nash, H. A. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1072–1076.
15. Gottesman, S. & Gottesman, M. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2188–2192.
16. Syvanen, M. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2496–2499.
17. Campbell, A. (1961) *Virology* **14**, 22–32.
18. Court, D., Green, L. & Echols, H. (1975) *Virology* **63**, 484–491.
19. Sussman, R. & Jacob, F. (1962) *C. R. Hebd. Seances Acad. Sci.* **254**, 1517–1519.
20. Lieb, M. (1976) *Mol. Gen. Genet.* **146**, 285–290.
21. Shimada, K. & Campbell, A. (1974) *Virology* **60**, 157–165.
22. Hendrix, R. (1971) in *The Bacteriophage Lambda*, ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 355–370.
23. Goldberg, A. R. & Howe, M. (1969) *Virology* **38**, 200–202.
24. Parkinson, J. S. & Huskey, R. J. (1971) *J. Mol. Biol.* **56**, 369–384.
25. Feiss, M., Adhya, S. & Court, D. L. (1972) *Genetics* **71**, 189–206.
26. Manly, K. F., Signer, E. R. & Radding, C. M. (1969) *Virology* **37**, 177–188.
27. Wu, A. M., Ghosh, S., Echols, H. & Spiegelman, W. G. (1972) *J. Mol. Biol.* **67**, 407–421.
28. Folkmanis, A., Takeda, Y., Simuth, J., Gussin, G. & Echols, H. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2249–2253.
29. Echols, H. & Green, L. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2190–2194.
30. Riggs, A. D., Suzuki, H. & Bourgois, S. (1970) *J. Mol. Biol.* **48**, 67–83.
31. Kaiser, A. D. (1957) *Virology* **3**, 42–61.
32. Reichardt, L. F. & Kaiser, A. D. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2185–2189.
33. Court, D., Green, L. & Echols, H. (1975) *Virology* **63**, 484–491.
34. Katzir, N., Oppenheim, A., Belfort, M. & Oppenheim, A. B. (1976) *Virology* **74**, 324–331.
35. Court, D., Adhya, S., Nash, H. & Enquist, L. (1977) in *DNA Insertion Elements, Plasmids, and Episomes*, eds. Bukhari, A., Shapiro, J. & Adhya, S. (Cold Spring Harbor, NY), in press.
36. Guarneros, G. & Echols, H. (1973) *Virology* **52**, 30–38.
37. Shimada, K. & Campbell, A. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 237–241.
38. Enquist, L. & Weisberg, R. A. (1977) *J. Mol. Biol., in press.*
39. Chung, S. & Echols, H. (1977) *Virology*, in press.