Annu. Rev. Biochem. 1989. 58:913–49

DYNAMIC, STRUCTURAL, AND REGULATORY ASPECTS OF λ SITE-SPECIFIC RECOMBINATION

Arthur Landy

Division of Biology and Medicine, Brown University, Providence, Rhode Island 02912

CONTENTS

0066-4154/89/0701-0913$02.00

PERSPECTIVES AND SUMMARY

On the 25th anniversary of Robin Holliday's model for single-strand exchange intermediates in homologous recombination (1), it is especially fitting to review the most recent developments in the site-specific recombination system of *Escherichia coli* bacteriophage λ. This article focuses on developments since the most recent reviews of λ recombination (2, 3), and has a different emphasis than a concurrent review (4) that explores the regulatory aspects of λ recombination in more detail. The present state of the field has its roots in the wealth of elegant λ genetics that rallied around Allan Campbell's model for viral integration (5). The purification of essential proteins by Howard Nash and his collaborators (6–8), and the definition of recombination target sites (9–11), allowed this system to be the first recombination pathway eligible for rigorous biochemical analysis. New results to be discussed in this review concern the dynamic, structural, and regulatory aspects of λ recombination. They are summarized below.

The dynamic aspects of λ recombination (such as mechanisms of strand exchange) prominently feature the biochemical analysis of Holliday junction recombination intermediates. Artificial DNA constructs designed to model putative Holliday-junction recombination intermediates are efficiently and specifically resolved by purified λ Integrase into normal recombinant products (12). Holliday intermediates can also be trapped as the major product in a λ recombination reaction using "suicide recombination substrates" (13). The pair of reciprocal strand exchanges that first form and then resolve the Holliday junctions proceed in a strictly prescribed order that is the same for integrative and excisive recombination (13, 14). This order is determined by DNA sequences and protein binding sites distant from the region of strand exchange (13, 14). Formation and resolution of the Holliday junction proceeds via a high-energy intermediate containing a covalent linkage between the DNA and Tyr-342 of Integrase protein (15). The requirement for a full 7 bp of DNA homology between recombining partners (16, 17) appears not to be for synapsis but rather for the purpose of branch migration between the staggered strand exchange sites within the Holliday intermediate (13, 18) (B. DeMassy, R. A. Weisberg, personal communication). The synapsis of *att* sites is most likely mediated by protein-protein and protein-DNA interactions, and in integrative recombination this may involve a protein-decorated supercoiled "donor" (*att*P) pairing with a protein-free "recipient" (*att*B) (19).

The structural aspects of λ recombination are becoming interesting as models for other complex protein–nucleic acid systems, such as those found in regulation of transcription and DNA replication. The *att* site DNAs function as part of higher-order protein-DNA complexes called "intasomes" (20, 21). These are formed by a set of cooperative and competitive protein-protein

interactions involving four proteins and 15 binding sites. Formation of one of these intasomes (*att*P) requires supercoiled DNA (22). The intasome structure is shaped largely by three sequence-specific "accessory" proteins (IHF, Xis, and FIS) that induce extremely sharp bends in DNA (23). The Integrase protein has the potential to tether distant sequences and form DNA loops by virtue of two autonomous DNA-binding domains with different recognition specificities (24). Integrase executes recombination (7) by a mechanism of strand cleavage and ligation that is shared with more than 15 related recombinases (Int Family) (15, 25).

The extreme directionality and regulation of λ recombination are among its most striking features. Directionality is the consequence of two distinctly different pathways for integrative and excisive recombination. Each pathway uses a unique, but overlapping, set of the 15 protein binding sites that comprise *att* site DNAs (26, 27). Cooperative and competitive interactions involving all four recombination proteins determine the direction of recombination. They also bestow upon the reactions sensitivity to host and viral physiology and to environmental conditions. Regulation of recombination at the level of gene expression is augmented by a mechanism-based response in which host-encoded proteins are incorporated as integral elements of the recombination reaction. One of the host-encoded proteins (IHF) that is required for both integrative and excisive recombination has one binding site that must be occupied for integration and must be vacant for excision (28). The other host-encoded protein (FIS) dramatically stimulates excisive recombination when levels of the phage-encoded Xis protein are low (29). Intracellular levels of FIS, which is also an accessory protein for a different class of recombination reactions, drop by 70-fold as cells go from exponential to stationary phase (29).

In addition to the recent developments outlined above, reference will also be made to FLP and Cre, two well-characterized members of the Int family. Although they are similar to Int in the basic mechanisms of strand exchange, they are quite different in terms of structural complexity, directionality, and regulation.

BACKGROUND

Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of specificity for both partners; the strand exchange mechanism involves the cleavage and rejoining of specific DNA sequences in the absence of DNA synthesis. Two major families comprise this class of reactions. The Resolvase-Invertase family is distinguished by the use of a conserved serine to form covalent intermediates with DNA, a constraint to intramolecular reactions, and an orientation prefer-

ence for recombination sites that can be either direct repeats (e.g. $\gamma\delta$ and Tn-3 Resolvase) or inverted repeats (e.g. the Hin, Gin, and Cin inversion systems).

The other family of reactions (Int Family) (15, 25) is distinguished by the use of a conserved tyrosine to form covalent intermediates with DNA, the ability to execute both inter and intra molecular reactions, and the capacity to recombine sites with either direct or inverted orientation. Members of the Int family can be further subdivided into the minimal systems (e.g. FLP of yeast 2μ circle and Cre of bacteriophage P1) and the complex systems, such as bacteriophage λ and related viruses.

The λ recombination system consists of four recombination sites (*att* sites) and four proteins that carry out the integrative and excisive recombination reactions as shown in Figure 1. The Integrase (Int) protein cuts and reseals DNA to carry out strand exchange. The other phage-encoded protein (Xis) is required only for excisive recombination. Of the two host-encoded proteins, IHF is required for both reactions, while FIS enhances only excisive recombination and only under special circumstances. Integrative recombination between specific sites on the viral (*att*P) and bacterial (*att*B) chromosomes generates prophage bounded by the recombinant products *att*L and *att*R. An excisive recombination between *att*L and *att*R regenerates *att*B and *att*P in the bacterial and excised viral chromosomes. Of the four *att* sites, *att*P (POP') is the most complex and *att*B (BOB') is the simplest (see Figures 1 and 2 for the coordinates and protein binding sites). The prophage sites *att*L (BOP') and *att*R (POB') are "hybrids" of *att*P and *att*B (see Figure 1).

The reader is referred to other closely related, or overlapping reviews on regulation of λ recombination (4), conservative site-specific recombination pathways (30, 31), Resolvases (32, 33), Invertases (34), enhancers (35), resolution of Holliday junctions (36, 37), FLP (38), topology (39–41), and to many related chapters in the books *Transposition* (42), *Genetic Recombination* (43), and *Mobile DNA* (44).

PROTEINS AND THEIR GENES

Int

Int is a basic protein of 40,330 M_r (45, 46) that has type I topoisomerase activity (6, 7, 15, 47, 48) and binds specifically to two different families of DNA sequences (49, 50). During recombination, it carries out the cutting and resealing of *att* site DNA via a covalent Int-DNA intermediate in the absense of any high-energy cofactors (8, 51, 52). Int can also transfer covalently bound DNA to the 5' OH of another DNA molecule in a sequence-independent reaction (S. E. Nunes-Düby, A. Landy, unpublished results). Int is encoded in the λ genome adjacent to the phage *att* site, reflecting the

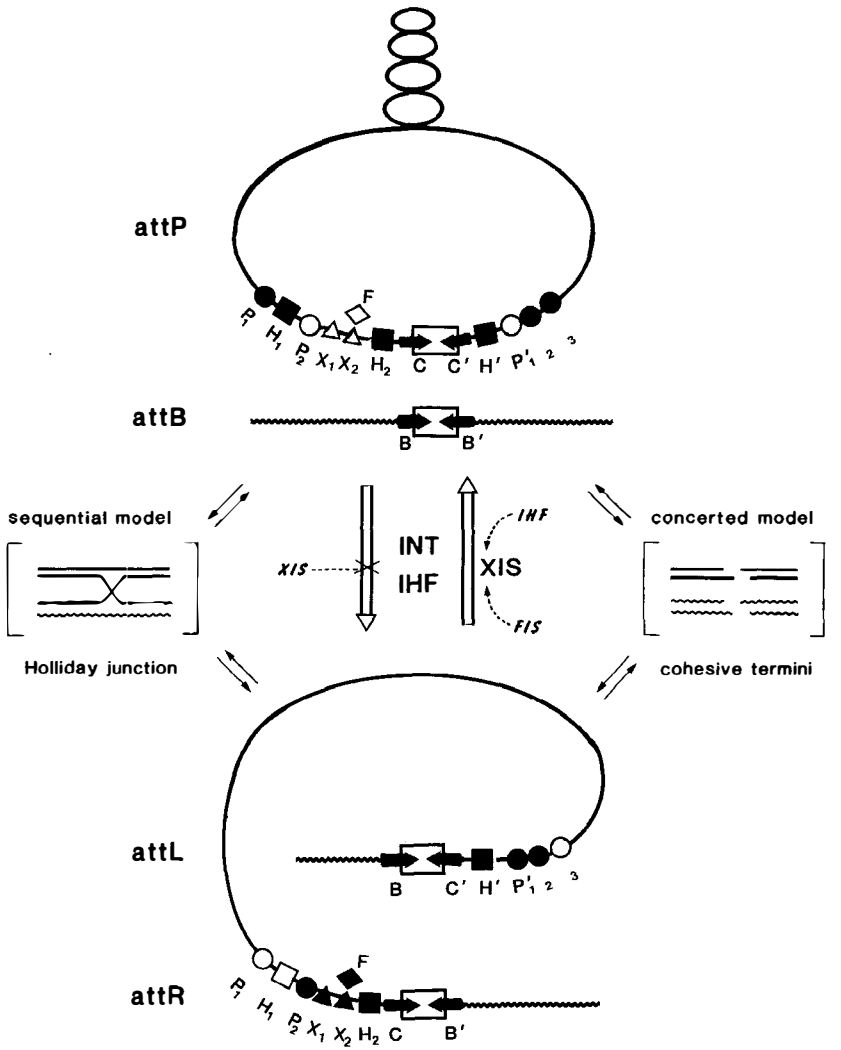

Figure 1 Integrative and excisive recombination pathways. The protein binding sites for arm-type Int (○), core-type Int (▬▶), IHF (□), Xis (△), and FIS (◇) are indicated by filled symbols when that site is occupied by its cognate protein to make a competent recombination partner for integrative (⇓) or excisive (⇑) recombination. Proteins required for each reaction (Int, IHF, and Xis) are in bold, proteins that inhibit (Xis and IHF) or enhance (FIS) the indicated reactions are in italics. Of the two models for strand exchange ([ ]), the sequential model (*left*) is correct.

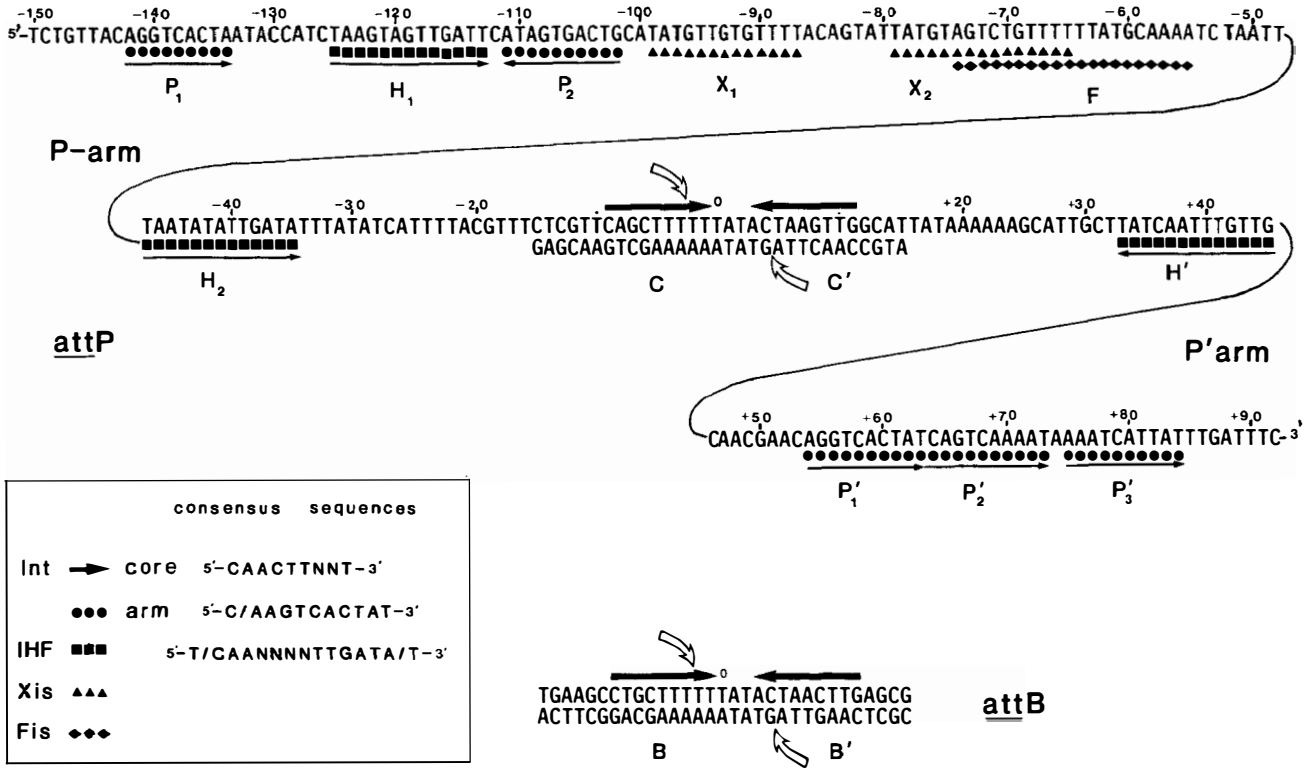

Figure 2 DNA sequence, coordinates, and protein binding sites for *att*P and *att*B. Recognition sequences for each protein are denoted with the symbols indicated lower left. Relative orientation of the binding sites is shown (⟶). Curved arrows (⇆) indicate the staggered sites of strand exchange and the boundaries of the 7 bp overlap regions.

common motif of recombinase genes that are very close to their sites of action. The specific arrangement of recombination genes often reflects their respective strategies for regulation of gene expression, which, in the case of λ *int*, involves retroregulation (53–57), its own cII-dependent promoter (58), and transcription from the distant P_L promoter [reviewed in (4)]. In phages P22 and P2, the *int* genes are oriented as in λ, but in ø80 and P4 they have the opposite orientation (59–61). The P2 *int* gene, which was once hypothesized to be split by the *att* site (62), has recently been shown to terminate its coding region just before *att* (61).

Thus far only one classical *int* mutant (resulting in a Glu to Lys change at position 174) (63) has been well characterized biochemically. The mutation was isolated in two independent experiments: (*a*) by selection for λ bacteriophage that could undergo site-specific recombination in an *E. coli* host mutant for IHF (called *int*-h) (64); (*b*) by selection for an Int partially independent of Xis in excisive recombination (called *xin*) (C. Gritzmacher, L. W. Enquist, R. A. Weisberg, unpublished data). In the absence of IHF, where Int^+ activity is depressed approximately 500-fold, Int-h activity is only depressed approximately 10-fold (52). The fact that this reduced, but significant, Int-h activity is identical for supercoiled and relaxed substrates indicates that the normal supercoiling requirement for *att*P is determined largely by IHF. The two most plausible explanations for this phenotype are that Int-h/Xin may have a higher affinity than Int^+ for core-type binding sites; or that it works differently on the synaptic intermediate, either to stabilize it or accelerate its conversion to recombinant product (65). A number of other *int* mutants that are interesting because they are recombination proficient (66) or dominant (63) have been studied but not thoroughly characterized.

A comparative study of the primary structures of proteins from seven different bacteriophages pointed to an "Integrase family" of recombination proteins that now includes more than 15 members (15, 25). Within this family of proteins there is one region of approximately 40 amino acids near the carboxy-terminus where at 25 positions more than 50% of the residues belong to the same amino acid exchange group (67). Particularly striking are three perfectly conserved residues within this region: His-308, Arg-311, and Tyr-342 (using the numbering of λ Int).

The biochemical significance of this region was determined using a family of suicide recombination substrates designed to accumulate transient recombination intermediates (13) (see below, SYNAPSIS AND STRAND EXCHANGE, and Figure 3). It was shown that Tyr-342 of Int forms a covalent bond with DNA at the sites of strand exchange. A mutant Int in which Tyr-342 is changed to phenylalanine is devoid of both topoisomerase and recombinase activity but still binds to both classes of Int DNA-binding sites with an affinity comparable to wild-type Int (15). The applicability of these

Figure 3 Suicide and canonical recombination substrates, and their respective Int-promoted products. I: medial nicked substrate when incubated with Int yields covalent Int-DNA product. II: bottom strand exchange site nick, after reaction with supercoiled *att*P and exchange of the top strands, yields a Holliday junction intermediate in the form of an α-structure with a nick. III: top strand exchange site nick yields no product when reacted with supercoiled *att*P. IV: in recombination between canonical *att* sites a left side *(top)* strand exchange leads to a Holliday junction; homology in the overlap region is necessary for the second strand exchange and would allow branch migration of the crossed strands; recombination is completed by a right side *(bottom)* strand exchange.

results to the Int family as a whole is supported by studies on FLP and Cre (68–70). In contrast, the Resolvase-Invertase family of recombinases forms a covalent linkage with DNA via a Ser(p) (71) located very close to the amino-terminus (35, 71–73) (see also BACKGROUND).

The recombinases within the Int family are found in a rather wide range of biological functions, such as phage integration and excision, fimbrial phase variation, and plasmid maintenance in *E. coli* and yeast. Some members rely on accessory proteins while others are self-sufficient. Some members, like Cre and FLP, are recombinases with a single DNA recognition site, while others, like λ Int, have two different DNA-binding domains. This intrafamily heterogeneity is reflected by the fact that only a single highly conserved 40-amino-acid region is common to all members of the Int family. It has been proposed that the unifying feature of the Int family is the protein domain for the DNA cutting-ligating activity; the identification of Tyr-342 as the active nucleophile in Int lends strong support to this view (15).

The fact that Int recognizes two classes of DNA-binding sites (49, 50) (see below and Figure 2), along with the biochemical differences in binding to each class (51, 74), led to the proposal for a new kind of DNA-binding protein (50). The existence of two functionally distinct DNA recognition domains in the Integrase protein was demonstrated by the fact that core-type and arm-type DNA sequences do not compete with each other for binding in nuclease protection experiments (24). These results rule out the proposal (75) that all of the Int binding sites are recognized by Int on the basis of similar structural DNA features. They also indicate that an Int monomer is capable of binding to both classes of sites simultaneously. The "bivalency" of Int provides a simple structural explanation for the ability of some *int* mutants to negatively complement (63, 76–79).

The identity and structural autonomy of the two DNA-binding domains was established by proteolytic cleavage of Int and footprinting analysis of the resulting two major peptides (24). A chymotryptic cleavage between Leu-64 and Thr-65 generates a 32-kd carboxy-terminal fragment that binds core-type sites and a 7-kd amino-terminal fragment that binds arm-type sites. The amino-terminal fragment that binds exclusively to arm-type sites is not required for catalytic function, while the carboxy-terminal peptide retains topoisomerase function and resolves synthetic *att* site Holliday junctions (24). This is consistent with earlier observations that Int can efficiently resolve Holliday structures, even in the absence of arm-type binding sites (12).

Xis

Xis is a small (M_r of 8630), basic, phage-encoded protein that is required for excision but not integration (45, 46, 80). Xis has no significant homologies with other proteins, including analogous excisionases from other lambdoid

phage, and nothing is known about its structure. While Xis is relatively thermostable in vitro, it is rapidly degraded intracellularly by an unknown protease, so that 80% of the activity decays within five minutes (80–92). Int, on the other hand, is quite stable intracellularly.

In the presence of Int and IHF, Xis promotes efficient recombination between *att*L and *att*R. With greater than 100 mM salt, Xis is required for this reaction, while at low salt it is dispensable, although still stimulatory (83). Several lines of evidence indicate that the low-salt Xis-independent reaction proceeds by a different pathway than the Xis-stimulated reaction (84). A similar difference in pathways has been proposed for integrative recombination at high and low salt with supercoiled versus relaxed *att*P (85).

Xis has been found to confer enhanced thermostability on Int protein in vitro (80), but there is no evidence to indicate whether this reflects an interaction relevant to recombination. No catalytic activity has been associated with Xis, and its function is exerted by sequence-specific cooperative binding to two adjacent sites in the P arm (26, 86). As discussed below, binding at these sites introduces a very sharp bend in the DNA and is also associated with cooperative interactions with DNA-bound Int and FIS (23, 29, 87, 88).

IHF

Integration Host Factor (IHF), as the name implies, was discovered as a cellular function essential for λ site-specific recombination, both integration and excision (89–92). IHF is a heterodimer composed of two subunits, with predicted molecular weights of 11,200 and 10,580 for the α and β subunits respectively (8, 93–95). Each of the subunits is very similar in sequence to the type II DNA-binding proteins, a family that includes the major histone-like proteins of *E. coli* (HU) and other bacteria [reviewed in (96)]. The only known (primary) function of IHF is its ability to bind (97–99) and to bend (23, 100–103) DNA at specific sites.

It may be valid to infer some basic features of IHF structure from a comparison with the HU protein from *Bacillus stearothermophilus* (HBs), whose crystal structure has been solved to 3 Å resolution (104). Both of the IHF subunits show significant homology to the subunit of the HBs homodimer (94). The HBs structure has two intertwined subunits that form a hydrophobic wedge-shape body and two long flexible arms extending from the wide portion of the wedge. It is proposed that the two flexible arms form a right-handed helix that could bind within the major or minor groove of double-stranded B DNA.

Craig & Nash (97) have proposed that IHF binds in the minor groove of DNA based on the patterns of protection against modification by dimethylsulfate (DMS). H. A. Nash (personal communication) has found addi-

tional support for this proposal in the patterns of protection against cleavage of the sugar-phosphate backbone by hydroxyl radicals. In these experiments, which were done on the three IHF sites of *att* DNA, as well as in nuclease protection experiments on other IHF sites (98, 99), IHF is found to protect three to four helical turns of DNA. Despite this relatively large protected region, preliminary stoichiometric data indicate that only a single IHF dimer binds per site (H. A. Nash, personal communication). It is also interesting to note that while IHF has high sequence specificity, the closely related HU and HBs proteins have little or no sequence specificity.

The genes for IHF have been mapped (89, 90), cloned (93–95, 105), and overexpressed (106). The α-encoding gene, *him*A, maps to minute 37, and the β-encoding gene, which is referred to as both *hip* and *him*D, maps to minute 20 on the *E. coli* chromosome. *Hip*/*him*D is part of a gene cluster with at least two promoters (E. L. Flamm, R. A. Weisberg, personal communication). The *him*A gene is part of a more complex cluster whose transcription involves at least seven promoters and three terminators, with the regulation being influenced by growth rate and the SOS response. For a recent review of the regulation see (4).

Some evidence suggests that levels of IHF may vary as a function of cellular physiology, but this point is not clear. A large increase in IHF levels is detected in stationary-phase cells, but this depends on the conditions used to extract the protein (27, 28). Attempts to resolve this question by in vivo chemical modification have not been definitive because all three of the IHF sites in *att*P that were monitored were fully occupied under the growth conditions tested (29).

The lambdoid phages ø80, P22, and λ all have IHF-binding sites within their very different *att* site DNA sequences (98). ø80, like λ, is unable to lysogenize *him*A or *hip*/*him*D mutants (107), while P22 has not been tested. Since the number of IHF-binding sites is different in each of the phage *att* sites, it will be interesting to compare the three recombination reactions at a mechanistic level. (Cre and FLP are two members of the Int family that do not require IHF.) Transposons are another class of recombination elements where a role for IHF has been demonstrated or implicated (99, 101, 108, 109). In some recombination systems an IHF requirement or HU requirement can be replaced with low efficiency by the other protein (108, 110). The roles of IHF extend well beyond recombination and are discussed thoroughly in a recent review (111).

FIS

A second host-encoded protein (in addition to IHF) involved in the λ site-specific recombination reaction was first discovered by testing *E. coli* extracts for proteins that would bind to *att* site DNA in a gel mobility shift assay. Such

an activity was found and subsequently shown to be FIS (29), a protein that had been identified (112–114), and purified (115, 116), on the basis of its ability to stimulate three recombination inversion reactions from the Resolvase-Invertase family. In the λ system, FIS stimulates in vitro excision approximately 20-fold when Xis levels are limiting, however it cannot replace Xis (29). FIS has no effect when Xis is at saturating levels and it does not influence integrative recombination. The in vivo relevance of FIS has been established by demonstrating occupancy of its binding site in the P arm (29), as well as dramatic differences in excision efficiency for FIS^+ and FIS^- cells (R. Johnson, personal communication). When bound to its single binding site in the P arm, FIS induces a DNA bend of approximately 90° (23).

FIS, like IHF, is a small, basic, heat-stable DNA-binding protein (115, 116). Unlike IHF, which is a heterodimer, FIS is a homodimer in solution. The protomer, based on its gene sequence, has 98 amino acids and a calculated pI of 9.5 (117, 118). There is no sequence homology between FIS and the type II DNA-binding proteins such as HU or IHF, despite their gross similarity in physical properties and somewhat analogous roles as bending and accessory proteins in several different recombination pathways. FIS has no cysteines or histidines so that a "zinc-finger" mode of DNA recognition is unlikely (117). However, within the carboxy-terminal portion, at amino acids 74–93, there is a sequence with a high probability of adopting the helix-turn-helix DNA binding motif (117, 118) that has been characterized for repressors and other proteins (119).

The Hin system of *Salmonella typhimurium,* the Gin system of bacteriophage Mu, and the Cin system of bacteriophage P1 constitute a closely related family of FIS-dependent site-specific recombination systems. These systems specifically invert a segment of DNA and thereby mediate expression of two alternate forms of a flagellar antigen or phage tail fiber protein (34, 35, 120). Their requirement for the bacterial protein FIS is mediated by a binding site on the DNA that has been called a "recombinational enhancer." The function of recombinational enhancers, which contain two or three FIS binding sites, is relatively independent of position or orientation relative to the sites of recombination (112–114, 118, 121, 122).

The *fis* gene has been mapped to approximately 72.5 min on the *E. coli* chromosome, between *fab*E and *aro*E (117, 118). The fact that the initiating methionine is preceded by a poor Shine-Delgarno translation initiation site, coupled with the high number of rare codons (11%), may contribute to the low abundance of FIS in *E. coli*. A null mutation in *fis* leads to a 10^3–10^4-fold reduction of Hin and Gin inversion in vivo and a complete loss of detectable FIS activity when extracts are assayed for ability to stimulate inversion in vitro (117, 118). These null mutants also exhibit a 10^3–10^4-fold reduction in λ excisive recombination efficiency, suggesting that under these experimental

conditions, intracellular levels of Xis are limiting (R. Johnson, personal communication). Strains carrying these same null mutations show no reduction in viability, even when this mutation is combined with a mutation in *hip*/*him*D (encoding one of the subunits of IHF). The inability to find major effects of *fis* mutations on cell physiology, or on the expression of other cellular genes, may reflect the short time these mutants have been available for study.

One reason for predicting some interesting effects of FIS on cell physiology, and a possible role in global aspects of regulation, is that its intracellular concentration drops more than 70-fold as *E. coli* goes from exponential growth to stationary phase (29). It is not known what controls the degradation of FIS or whether gene expression is also involved.

ATT SITE STRUCTURE AND PROTEIN BINDING SITES

The most prominent feature of the four *att* site DNA sequences is a 15 bp "core sequence" they have in common. Individual sequence elements of functional significance lie within, outside, and across the boundaries of this common core, which itself has no functional significance. Controlled resection experiments established that *att*P extends from −150 to +85 from the center of the core, while *att*B extends from approximately −15 to +15 (see Figure 2) (10, 11, 123). The prophage sites generated by integrative recombination, *att*L and *att*R, are fully competent for excisive recombination, but they contain more than the minimal sequences necessary for excision (27) (see below, DIRECTIONALITY AND REGULATION). It is interesting to note that the *att*B sites of six different systems are located within the 3'-terminal regions of various tRNA genes (60, 98) (W. Reiter, P. Palm, S. Yeats, personal communication). While this association with tRNA genes does not hold for the majority of *att*B sites, it is significant and may offer some useful insights about evolution and/or mechanisms.

The *att*P sequence is 73% A+T, with one 48 bp region being 90% A+T (9, 10). As might be expected from this base composition, there are several runs of up to six consecutive adenines, a sequence feature known to generate intrinsic curvature within the DNA (124–126). The center of *att*P curvature has been mapped to −45 to −50 in the P arm (127). This same region is also hypersensitive to modification by bromoacetaldehyde, but only if the DNA is supercoiled (128). While the intrinsic curvature in this region of the P arm may enhance the efficiency of recombination, it is not absolutely required (127). Its function may be to position the *att* site to an external, easily accessible loop in plectonemically supercoiled DNA (129, 130).

In vitro integrative recombination reactions traced the observed requirement for negative supercoiling to the *att*P partner (131). Below 50 mM NaCl

the reaction does not require supercoiled *att*P (132); however, several lines of evidence suggest that this low-salt reaction with linear *att*P proceeds by a different recombination pathway (85). In contrast to integration, excision does not require supercoiling (133), although supercoiling of either *att*L or *att*R stimulates the reaction at limiting protein concentrations. Excision is also less sensitive than integration to salt concentration (132).

Overlap Region

The 7 bp between the staggered cuts responsible for strand exchange is referred to as the "overlap region" (3), since in recombinant molecules this region receives one DNA strand from each partner. The *att* site mutants of Shulman & Gottesman (134) that segregate to both recombinant products have been located within the overlap region by DNA sequencing (135). These mutants greatly impair recombination efficiency by deleting one bp from the overlap region. Insertions of one bp are not as detrimental (136), but a systematic test of spacing has not been carried out. Because this region receives one DNA strand from each partner during recombination, one would predict that it must have the same sequence in both partners. Indeed, many sequence changes can be accommodated within the overlap region as long as the same changes are introduced into both partners (17, 81). The role of DNA:DNA homology will be discussed below (SYNAPSIS AND STRAND EXCHANGE).

An overlap region, defined both by the position of staggered cuts and by the requirement for DNA:DNA homology, is also found in the *lox* sites (Cre system) and FRT sites (FLP system), 6 bp in the former (137–141) and 8 bp in the latter (142–145). In both of these systems some specific sequences within the overlap region have been observed to reduce recombination despite the preservation of homology between the partners. The weaker effect of specific overlap region sequences in λ *att* sites may be due to the dominating effects of the proteins bound to the P and P' arms. Similarly, in the absence of the asymmetry conferred by the P and P' arms of λ, the orientation of *lox* and FRT recombination partners relative to one another is determined wholly by the overlap region sequence (139, 143, 145, 146). The Resolvase-Invertase systems also make use of staggered cuts during strand exchange, but the size of the overlap region is only 2 bp (71, 73, 112, 147, 148).

Int Binding Sites

There are seven Int-binding sites in *att*P, two of which define the outermost limits of *att*P (10, 49). Sequence inspection suggested that these sites could be grouped into two distinct families of sequences: "core-type" (also called "junction-type") sites that are adjacent to the points of strand exchange and "arm-type" sites that are distal to the region of strand exchange. This distinc-

tion was confirmed, and reliable consensus sequences were obtained for each family, by analyzing a number of fortuitous Int binding sites in non-*att* DNA (49, 50). As described above, each family is recognized by one of two autonomous DNA-binding domains of Int (24).

CORE-TYPE SITES The core-type sites (C and C' in *att*P and B and B' in *att*B) are found as inverted repeats at the two ends of the overlap region and include the phosphodiester bonds that are cut by Int during strand exchange. The two base pairs flanking this phosphodiester bond are not specified in the consensus recognition sequence of seven bp (50). This lack of specificity may be a consequence of protein structural features in the catalytic region of Int. Int bound at the core-type sites protects DNA against attack by dimethyl sulfate in both the major and minor grooves along one face of the DNA helix; the inverted repeats face each other across the central major groove. The opposite face of the DNA helix appears to be devoid of strong protein contacts (50). The appropriate configuration of core-type sites is the only requirement for the resolution of synthetic Holliday junctions by Int, or its carboxy-terminal domain (12) (C. Pargellis, B. Franz, and A. Landy, unpublished results).

Although affinity constants for Int binding have not been measured, they are considerably lower for the core-type sites than for the arm-type sites, as estimated by their respective behavior in nuclease protection (19, 49) and gel mobility shift (23) experiments. The relative order of binding affinity for core-type sites is C' > C = B > B' (50). However, the differential in affinities does not appear to be essential, since different sites can be substituted for one another with little or no effect on recombination (13, 26). The differences in affinities of the individual sites may be partially masked by the significant amount of cooperativity between Int proteins bound at two core-type sites (50).

The amount of Int-induced DNA bending at the COC' core-type sites has been estimated to be approximately 17° (23). However, this value must be viewed cautiously because of the relatively low affinity of Int for these sites (50). Additionally, measurements on the isolated core-type sites may not be relevant to the functional recombinogenic complexes (see below, COOPERATIVE INTERACTIONS).

Two core-type sites flanking the overlap region comprise *att*B. This configuration is very similar to the minimal recombination sites of both partners in the Cre and FLP systems. The simplicity of *att*B explains the existence of secondary *att* sites in the *E. coli* chromosome. These sites are utilized for λ integration (at greatly reduced efficiencies) when *att*B is deleted (149–151). Sequence comparisons of the secondary *att* sites indicates that they are degenerate facsimiles of the simple *att*B site. The efficiency of a particular

secondary *att* site is presumably determined by the extent to which it emulates the canonical spacing and sequence of the overlap region and core-type Int binding sites (3, 50). Similarly, fortuitous facsimiles of *att*B, or core-type sites, could comprise the targets for Int topoisomerase activity on supercoiled DNA not containing an *att* site.

ARM-TYPE SITES The initial characterization of a distinct class of essential Int-binding sites well removed from the region of strand exchange was both intriguing and challenging in terms of recombinogenic structures and mechanisms (49, 74). Evidence for how they function in site-specific recombination is beginning to emerge. The five arm-type Int-binding sites have a consensus recognition sequence of seven contiguous base pairs (49). The approximate binding affinities for individual sites indicates a hierarchy of P'1 = P1 > P2 (49, 87, 88). P'2 and P'3 have not been examined in isolation and all three of the P' sites are subject to cooperative binding effects amongst themselves (L. Moitoso de Vargas, A. Landy, unpublished results). In addition to these cooperative effects, sequences outside of the binding site can also influence the binding affinities. [However, the low affinity of Int for P2, which has the best match to the consensus sequence, is not due to context effects (87).]

Classical genetic techniques have produced only one protein-binding mutant in the *att* sites. This mutation, *hen,* is in the P'3 site and changes a completely conserved C to T (152). It is especially interesting because it abolishes integrative recombination without affecting excisive recombination (152). To examine the roles of other arm-type sites, the single base transition of the *hen* mutation has been introduced into each of the five arm-type sites (153). Multi-site changes have also been introduced into a subset of the arm-type sites (27, 88). From these studies and resection experiments (10, 11), it is clear that the P1 and P'2, and P'3 sites are required for integrative recombination, while the P2 site is not. The apparent dispensability of P'1 for integrative recombination is subject to the same reservation that applies to all point mutants without a phenotype in highly cooperative systems.

In excisive recombination, none of the point mutations has as large an effect as seen with integrative recombination (153). The largest effect is observed with the P'1 mutation, suggesting that this site is required for excision, with some of the defect being compensated by cooperative interactions with other sites. P1 is clearly not required for excision, because when it is deleted the *att*R works as well or better than wild-type *att*R (27). While P2 is also dispensable under some circumstances, it does appear to be important for excision. In the absence of the P2 site, more Xis is required to obtain the same level of recombination as with a wild-type *att*R (27, 87). Additionally, the reaction between an *att*R-*hen* P2 and an *att*L-*hen* P'2 or

*att*L-*hen* P'3 is reduced relative to a wild-type *att*L (153). Both of these observations are reflective of cooperative interactions to be discussed below. See Figure 1 for a summary of the site requirements for integrative and excisive recombination.

Xis Sites

Xis protein protects a single region of approximately 40 bp in the P arm (86). Gel mobility shift assays have shown that highly cooperative binding between at least two Xis molecules takes place (26). Mutagenesis of the binding region showed that each half could bind Xis independently, although with lower affinity than the canonical site (26). Protection patterns against DMS modification also provide evidence for two binding sites, denoted X1 and X2, and further suggest that they are arranged as a direct repeat (86). Determining a consensus recognition sequence will require the analysis of fortuitous Xis recognition sites and/or a mutational study of the canonical sites.

The Xis mutations isolated thus far include a single base deletion between the two sites and a series of resected *att* sites in which varying extents of the P arm have been replaced with heterologous DNA. In the first case, excisive recombination is abolished. In the second, it was shown that both X1 and X2 must be occupied; defects in X1 are only overcome when sufficient Xis is present to bind to the mutated as well as the normal site (27).

Xis binding to X1 or X2 bends the DNA approximately 45° and 95°, respectively. When X1 and X2 are both occupied, as is normally the case, the bending angle is greater than 140° (23). This very sharp bend is undoubtedly a key feature in the role of Xis during recombination.

IHF Sites

There are three IHF sites in *att*P: H1 and H2 in the P arm and H' in the P' arm (97). IHF binding to multiple sites on the same DNA fragment is not cooperative (28, 88, 97, 101, 154). Derivation and refinement of a consensus recognition sequence has been aided by the analysis of IHF binding sites from a variety of phage, plasmid, and bacterial sources (97–100, 154–156). IHF protection patterns against modification by dimethyl sulfate further delimit those positions involved in binding (97, 100) (see Figure 2).

As discussed above, and similar to Xis, the only known function of IHF is its ability to bend DNA (23, 100–103). The extent of IHF-induced bending is estimated to be greater than 140° at each of the three *att*P binding sites (23). The strong IHF-induced bend is consistent with the observations and analyses of DNA bending induced by CAP protein (157). The large effects (greater than 100-fold) on IHF binding affinity by sequence changes outside the consensus region (101) (L. Moitoso de Vargas, A. Landy, unpublished results) could be the result of changes in the bendability, and/or intrinsic

curvature, of the DNA. Similarly, variation in the amount of DNA protected by IHF at each of several different binding sites (97, 98) might reflect differences in the length of DNA involved in making the curve around IHF protein. The H2 and H' IHF binding sites, as well as some non-*att* IHF sites (100), occur in regions where the DNA has a substantial intrinsic curvature in the absence of bound protein (23, 127), e.g. as determined by gel mobility shift permutation analyses (125). While intrinsic curvature may have the potential to favor IHF binding if it is in phase with the IHF-induced bend (157), it is not required. A number of IHF sites, such as H1, do not show any evidence of intrinsic curvature (23, 102).

Despite the gross similarity of H1, H2, and H' in the extent of IHF induced bending, the three sites are not functionally equivalent in recombination. Whereas H1 is required for integrative but not excisive recombination, H2 and H' are required for both reactions (26-28, 154). Additionally, it appears that the IHF requirement at H2 and H', but not at H1, can be partially substituted by another protein, possibly HU (154). Progress toward answering some of the questions about IHF should come from the recent IHF-DNA co-crystals made with synthetic oligomer (K. Appelt, personal communication).

FIS Sites

FIS binds to a single site, called F, that overlaps the X2 site in the P arm; the region protected against nuclease digestion is slightly larger than X2 (29). In contrast, the recombinational enhancers of the Hin and Gin family consist of two independent FIS binding sites separated by approximately 48 bp (113, 121, 122). The FIS site in λ *att* also seems to differ in sequence from the enhancer FIS sites (122). This lack of similarity suggests that the FIS recognition determinants are not fully understood and may involve specific DNA conformations not readily apparent in a simple sequence alignment.

Even though FIS can substitute for Xis binding at X2, the two proteins do not recognize the same features in the DNA: their respective DMS-modification protection patterns and the boundaries of the regions protected against nuclease digestion are different, and FIS does not bind to the X1 site (29, 86). In order to separate the effects of FIS and Xis, the X1-X2 site was replaced with an X1-X1 site in the P arm, thereby eliminating FIS binding while retaining the Xis binding (29, 127). This FIS^- mutant made it possible to show that, in contrast to its stimulation of excision, FIS has no effect on integrative recombination.

The effect of adding FIS to a nuclease protection or gel mobility shift assay is equivalent in many respects to simply adding more Xis. If Xis is limiting,

the addition of FIS leads to occupancy of the X1 and F sites with the same cooperativity seen for the X1 and X2 sites when Xis is increased in the absence of FIS. FIS binding to F is also similar to Xis binding to X2 in the extent of induced DNA bending, approximately 90° in both cases (23). Furthermore, the increase in the total apparent bend to more than 140° is similar for X1-F occupancy and for X1-X2 occupancy. Finally, as discussed below, all of the cooperative interactions in which Xis participates with Int and IHF occur equally well when FIS is bound in combination with Xis.

COOPERATIVE INTERACTIONS AND HIGHER-ORDER STRUCTURES

The first evidence for the wrapping of DNA in a putative higher-order structure was the appearance of a 10-base-pair repeat in the nuclease protection patterns of Int-DNA complexes (10, 49, 158). Such patterns of enhanced cutting suggest that one face of the helix is oriented toward the solvent, while the other face is less accessible to nuclease because it lies along a protein surface or forms the inside of a loop (159–162). Electron microscopy graphically revealed large protein complexes at the phage *att* site (20, 163) that suggested the involvement of approximately 230 bp of *att*P DNA and 4–8 Int monomers (20). The term *intasome* was coined to describe these complexes that were considered to contain only Int and DNA (20). It is now used as a generic term for protein–*att* site DNA complexes that also contain the other proteins involved in recombinogenic structures. While electron microscopy was not successful in revealing the importance of IHF in complex formation, it did show that Xis protein favored the formation of complexes involving *att*R and *att*P, and that many combinations of complexes could pair with one another (20, 21).

The first evidence that the functional recombinogenic complex involves DNA wrapped in a nucleosome-like structure came from the topological studies of Pollock & Nash (132). Intramolecular recombination between two *att* sites (either *att*P × *att*B or *att*L × *att*R) oriented in opposite directions on a circular supercoiled molecule inverts one segment of the circle with respect to the other and produces knotted products (164, 165). For integrative recombination, under conditions that minimized the knotting due to random interwrapping of the superhelical DNA, approximately one half of the recombinant products are simple circles and the rest are knotted. For excisive recombination, all of the recombinants are simple circles. The excess knotting of the integrative recombination products is postulated to reflect the wrapping of DNA in the *att*P complex (132). Subsequent topological analyses of knotted and catenated recombinant products have confirmed this result and

indicate that the *att*P "intasome" is wrapped with a negative sign, which in the nucleosome cores of chromatin produces a left-handed solenoid (130, 166–168).

Several lines of evidence suggest that the role of the supercoiling requirement for integrative recombination is to facilitate formation of the *att*P intasome. Since only *att*P requires supercoiling, it is not likely that any step subsequent to synapsis, such as strand exchange, is responsible for the requirement (169). The singularity of *att*P is evidenced by the fact that supercoiling is not required for excisive recombination, although it stimulates the reaction (170). Any degree of DNA wrapping in the *att*L or *att*R intasomes should also have a negative sign, since negative supercoiling of either one enhances the efficiency of recombination (133) (W. Bushman, J. Thompson, A. Landy, unpublished results).

In a more direct biochemical analysis, it was shown that binding of Int to P1, and to a lesser extent to P'3, is more resistant to challenge by excess salmon sperm DNA when *att*P is negatively supercoiled (22). This effect requires the presence of IHF, and is seen only on *att*P (i.e. not when the arm binding sites are separated as in *att*L and *att*R). The good correlation between the sign and degree of supercoiling needed to promote these effects, and that needed to promote recombination, supports the notion that the primary role of supercoiling is to favor formation of the intasome (22).

The importance of the spatial orientation of the P1 and H1 sites within the intasome is seen when the helical phase of these sites is altered with respect to the remainder of *att*P by making insertions or deletions in a nonessential intervening region (85). Under normal reaction conditions (where supercoiling is required), substrates with changes of integral multiples of the DNA helical repeat recombine, while substrates with nonintegral changes do not. With nonsupercoiled (linear) *att*P, in low salt conditions, both the helical phasing and the P1 and H1 sites are much less important (although the overall efficiency is lower than with supercoiled *att*P) (85).

The specific interactions responsible for the intasome structures have been deduced by correlations of recombination with multiprotein nuclease protection assays, competition binding studies, site-specific mutagenesis of protein-binding sites, and results from suicide recombination substrates. These interactions are summarized below and illustrated schematically in Figure 4.

The pairwise cooperative interactions between Int molecules bound at the core-type sites was demonstrated by the fact that Int binds much better to core-type sites positioned as inverted repeats 7 bp apart (i.e. the canonical configuration) than to isolated core-type sites (50, 135). Similar cooperativity is also inferred from studies with artificial Holliday junctions, but this configuration includes the potential interactions among Int protomers bound at

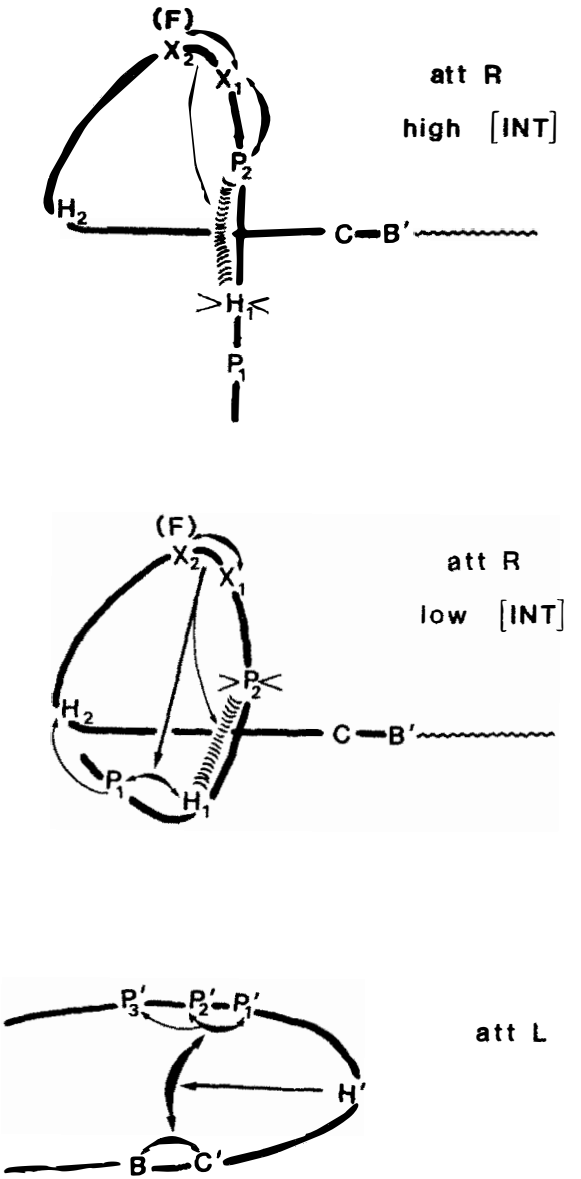

Figure 4 Schematic representation of some of the cooperative interactions observed with linear recombinogenic complexes. In some cases, cooperative interactions ( ) and competitive interactions ( ) are dependent upon binding of another protein ( ) The outcome of a competition at high or low Int is indicated by a vacated ( ) protein binding site. For more information about the individual protein binding sites see Figures 1 and 2.

each of the four core-type sites (B. Franz, A. Landy, unpublished results). Cooperative interactions among Int molecules bound at the core-type sites have been difficult to study because Int has a low affinity for these sites.

The most thoroughly characterized of the pairwise cooperative interactions involves Xis binding at the adjacent X1 and X2 sites (26, 86). This rather strong cooperativity is of special interest for two reasons. First, it involves a head-to-tail arrangement of sites, which is not so common for cooperatively binding proteins. Second, the Xis-Xis cooperativity can be substituted by Xis-FIS cooperativity. Keeping in mind that Xis and FIS do not have similar amino acid sequences, and that they bind to different DNA sequences, it is all the more surprising that for every interaction tested, the X1-F pair works as well as the X1-X2 pair (29). In the following discussion, all Xis interactions refer to Xis bound simultaneously at X1 and X2, and are probably equally applicable to simultaneous binding of Xis at X1 and FIS at F.

Xis bound at X1-X2 (or X1-F binding) cooperatively assists Int binding at P2. In nuclease protection experiments there is a 16-fold reduction in the amount of Int required to fill P2 if X1-X2 is also filled (26). This cooperativity is also observed in recombinations involving *att*R with a defective P2 or X1 site (27, 87, 88).

In addition to the cooperative interactions within the intasome, there are also several competitive interactions. These, however, are not simple pairwise competitions, since any single protein can fill all of its binding sites, and any pairwise combination of proteins can fill all of their respective binding sites (26, 29, 49, 86, 88, 97). Int binding at P2 and IHF binding at the neighboring H1 site are perfectly compatible in the absence of Xis. However, when X1-X2(F) is filled, then binding at H1 is competitive with binding at P2 (29, 88). Clearly, this competition is not due to simple occlusion of a binding site by neighboring proteins, given the pairwise compatibility of protein binding. Rather, the favored interpretation is a competition for adjoining space within a higher-order structure.

In order to better understand the dynamics of the H1-P2 competition, it is necessary to also refer to a complex cooperative interaction involving the P1 site (88). In linear molecules, Int binding to P1 in *att*R is strongly enhanced by the simultaneous occupancy of H1, H2, and X1-X2(F). However, in *att*P this cooperativity is not observed (the P' arm is involved in a different set of cooperative interactions that includes the core-type Int binding sites) (L. Moitoso de Vargas, A. Landy, unpublished). Thus, the binding affinity of Int for P1 in a higher-order complex is determined by interactions occurring on DNA at least 150 bp distant (see Figure 4). Richet et al (22) have found IHF-dependent cooperative interactions involving the P1 site in *att*P that require supercoiling. Thus, two major features that distinguish integrative and excisive recombination (the supercoiling requirement for *att*P and the Xis

requirement for *att*R) are also needed for the long-range interactions involving the P1 site.

The above discussion indicates that each partner in the H1 vs P2 competition is involved in a separate set of cooperative interactions. Therefore, the outcome of the competition is determined by the configuration of *att* site arms, i.e. *att*R vs *att*P, by the relative concentrations of each protein, and by the topological state of the DNA. As will be discussed (DIRECTIONALITY AND REGULATION), the cooperative transition in binding of proteins from the P1, H1, X1/X2(F), H2 set to the P2, X1/X2(F), H2 set (see Figure 5) occurs over the same range of protein concentrations that is required to overcome IHF inhibition of excision and to favor efficient excisive recombination (28, 88).

In contrast to the complexity of the P arm, the relative simplicity of the P' arm highlights one of the primary interactions responsible for intasome structure. In *att*L, IHF bound at H' enhances binding of Int to core-type sites as long as the P' arm-type sites are present. This same cooperative interaction is reflected in an (H'-P')-enhanced nicking of *att*L suicide recombination substrates. In *att*L, both the IHF-enhanced Int binding at core-type sites and recombination proficiency depend upon the correct helical phasing of the P' arm-type sites with respect to the core region (L. Moitoso de Vargas, S. H. Kim, A. Landy, unpublished). These cooperative interactions lead to the architectural element shown in Figure 6 and discussed below.

The number and complexity of the intasome-forming interactions stands in sharp contrast to the compact "minimized" reactions of some other Int family members such as Cre and FLP. Since these are highly efficient site-specific recombinases, it is appealing to ascribe much of the additional complexity of the λ system to the requirements for directionality and regulation, as discussed below.

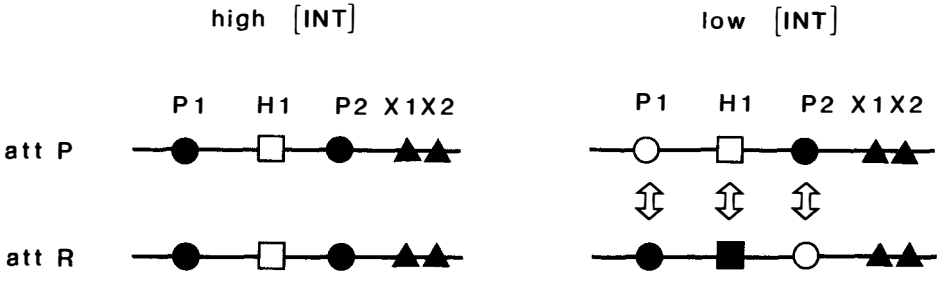

Figure 5 Coordinate occupancy of protein binding sites in the P arm of *att*R and *att*P as a result of cooperative and long-range interactions. Filled symbols indicate occupancy of the site by its cognate protein (P1 and P2 by Int, H1 by IHF, X1X2 by Xis). Double-headed arrows highlight the differences between *att*P and *att*R.

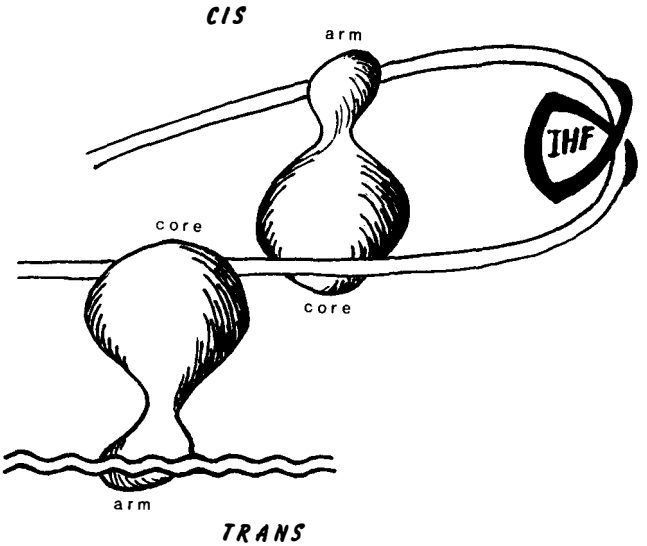

Figure 6 Basic design element of an intasome. IHF induces a sharp (≥140°) bend in DNA to bring the distant arm-type and core-type sites close enough to be bound by an Int protomer (*cis* interaction). The bivalent Int can also be involved in *(trans)* interactions between the arm-type and core-type sites on different partners. Xis and FIS are postulated to play a similar role to IHF.

SYNAPSIS AND STRAND EXCHANGE

Covalent Intermediates

The sites of strand exchange, i.e. the locus of cutting and resealing individual DNA strands, was narrowed down to two adjacent phosphodiester bonds, in an elegant isotope transfer experiment by M. Mizuuchi and K. Mizuuchi (171). Resolution of the remaining ambiguity came from sequence comparisons of many secondary *att* sites (3, 50) and from the demonstration by Craig & Nash (51) that Int makes a covalent link with the 3' phosphate at the site(s) of strand exchange (reviewed in Ref. 3).

In order to carry the analysis of transient recombination intermediates further, two families of suicide recombination substrates proved useful (see Figure 3) (13, 15). These *att* sites, which contain a strategically placed nick within the overlap region, initiate recombination normally and become deviant only after Int has acted on them. They are unable to complete or reverse the reaction. The particular intermediate trapped by such substrates depends upon where the nick is placed within the overlap region. One family of suicide substrates contains a medial nick in the overlap region (centered between the two sites of strand exchange) and leads to the formation of covalent Int-DNA

complexes in high yield. This work showed that Tyr-342 of Int forms a covalent link with DNA as an intermediate during strand cleavage and religation (15). The other family of suicide substrates is discussed below.

The Order of Strand Exchange

The existence of covalent Int-DNA intermediates does not distinguish between the two possible mechanisms of strand exchange: a concerted mechanism, in which both DNA strands in each *att* site are cleaved prior to strand exchange; and a sequential mechanism, in which one DNA strand in each *att* site is cleaved and exchanged followed by a second cycle of cleavage and exchange (see Figure 1).

From genetic crosses, it was found that Int-promoted recombination in vivo can generate the progeny predicted by a sequential strand exchange mechanism (172, 173). However, the low yield of these progeny was cited as evidence against their being on the primary recombination pathway (128). In vitro data that were interpreted to favor a concerted mechanism included the observation of low levels of double strand breakage by Int protein (128, 174). Support for a sequential strand exchange mechanism came from the specific resolution of synthetic *att* site Holliday junctions by purified Int protein to yield completed recombination products (12). The high efficiency and specificity of the resolution reaction was persuasive but not compelling.

Two different experimental devices finally led to the isolation of the elusive single-strand exchange recombination intermediates. One of these is the second family of suicide recombination substrates discussed above. In this family of suicide substrates, the preexisting *att* site nick is located at either the left (top strand) or right (bottom strand) Int cleavage site of the *att*B partner (see Figure 3). When the nick is in the bottom strand, single-strand exchange intermediates are produced under normal reaction conditions with almost the same efficiency as recombinant products from normal substrates. However, when the nick is in the top strand, no intermediates or recombinant products are formed (13). Thus, recombination proceeds via a sequential mechanism of strand exchange and the order is fixed such that the top strands must exchange first. Blocking the top strand exchange with a nick prevents the reaction from initiating. Using suicide recombination substrates for an *att*L × *att*R reaction, it was shown that excisive recombination has the same prescribed order of top strand exchange followed by bottom strand exchange (13, 14).

Similar conclusions have been drawn for integrative recombination (excision was not tested) from experiments in which substitution by phosphorothioate nucleotides into *att* site DNA was used to block Int cleavage (14, 65). Substitution in the top strand blocked strand exchange, while substitution in the bottom strand generated Holliday intermediates. In these experiments, however, very low levels of the recombination intermediate were seen, even

when using the Int variant known as Int-h (see above) to increase their yield. This difference between the analogue substitution and suicide recombination substrates is informative, especially when coupled with the low levels of Holliday structure obtained in normal reactions, or those blocked by a heterology in the overlap region (see below) (13, 14, 18). These differences have led to the suggestion that the first strand exchange during integrative recombination is reversible. Thus, in order to obtain high yields of recombination intermediate, it is not sufficient to block forward progression of the reaction, but it is also necessary to block its reversal (13). Understanding which of several possible mechanisms is responsible for the ability of the suicide substrate to trap Holliday junction intermediates is likely to provide useful insights into the normal reaction. Thus far, results obtained with a variety of different suicide recombination substrates indicate that Int protein, when provided with an appropriate 5' OH acceptor, will efficiently execute a single DNA strand transfer reaction (13) (S. E. Nunes-Düby, A. Landy, unpublished results).

The strict order of strand exchanges does not arise from asymmetry in the core region, but rather from the asymmetric arrangement of proteins bound to the P and P' arm sequences: (*a*) the core regions in the suicide recombination substrates were constructed to be virtually symmetric (13); (*b*) the same order of strand exchanges is obtained with *att* sites having different core region sequences (13, 14); (*c*) inversion of the overlap (13) or the full core region (14) does not change the order of strand exchange relative to the P and P' arms.

Holliday recombination intermediates have also been detected in the Cre and FLP systems (141, 145, 175). Thus, it is likely that this mechanism of sequential strand exchanges is common to members of the Int family.

DNA Homology

Although the specific sequence of the overlap region is not critical (within a wide latitude) for the order of strand exchange, or for the overall efficiency of recombination, it is required that the overlap region have the same sequence in both recombination partners (16, 17, 136). Two models have been proposed for the role of DNA : DNA homology in the overlap region. In the first, homology is required for synapsis of the recombining partners, and more specifically for the formation of a four-stranded DNA structure (7, 176, 177). From this view of recombination, considerable attention has been focused on whether the interstrand interactions are likely to be of the type proposed by McGavin (178) or by Wilson (179), and the likely direction in which strands rotate around the four-sided block during recombination (166, 177). In the second model, which is strongly favored by recent data, DNA : DNA homology is not required for synapsis but rather for some later step, such as branch

migration of a Holliday junction (16). Experiments with *att* sites that are heteroduplex within the overlap region argue against mechanisms involving a homology-dependent annealing of cohesive ends (180), however they do not distinguish between these two models.

A distinction between the two models is afforded by the experimental devices described above for obtaining single-strand exchange intermediates containing Holliday junctions. Whereas the formation of single-strand exchange intermediates is completely blocked by a nonhomology on the left side of the overlap region, it is not affected by a nonhomology on the right (13, 18). When the center of the overlap region is tested, single-strand exchange intermediate formation is blocked by a nonhomology at position 0 but not at position +1 (18). These results rule out that class of models, including that of four-stranded DNA, in which homology over all (or most) of the overlap region is required for synapsis. However, they do not exclude a requirement for a local homology of several base pairs in the immediate vicinity of a strand exchange site.

Whether or not strand exchange depends on some very short local homologies, a proposed requirement for a homology-dependent branch migration step is attractive (16). Artificial *att* site Holliday junctions have been made with sequence nonhomologies in the overlap region designed to restrict branch migration. For some of these constructions, resolution is strongly biased in favor of parental-type products, in contrast to wild-type Holliday structures that resolve to parental and recombinant-type products without pronounced bias. These results suggest that heterology prevents the Holliday junction from migrating from one Int cleavage site to the other and that resolution is executed by Int if the junction is at, or near, the Int cleavage site (B. DeMassy, L. Dorgai, R. A. Weisberg, personal communication). However, some data from these experiments suggest that Int may not be able to cleave when the junction is constrained very close to the cleavage site. The DNA branch migration model also gains further credibility from its ability to explain the results of recombination in which one partner is heteroduplex in the overlap region (17, 180).

Synapsis

Additional evidence that DNA homology is not required for synapsis comes from studies on both integrative and excisive recombination. Heteroduplex *att*B was incubated with a supercoiled homoduplex *att*P under recombination conditions; although the two *att* sites cannot recombine, because of nonhomologies between their overlap regions, double-strand cleavage of *att*B is promoted by the *att*P intasome. Thus, capture of *att*B by the *att*P intasome does not depend upon DNA homology between the sites (19). In addition, chemical footprinting assays establish that under recombination conditions,

*att*B cannot stably bind Int in competition with other DNAs. Taken together, these experiments suggest that *att*B may obtain its Int by binding to a preformed *att*P intasome in a synapsis that is governed primarily by protein-protein and protein-DNA interactions and is independent of DNA homology (19). Thus the asymmetry in structural complexity is indeed reflected in functional asymmetry. For the analysis of excisive recombination, half-*att* sites were constructed by cutting at a restriction site within the overlap region. These half-*att* sites are proficient for synapsis and strand transfer to an intact *att* site partner despite the lack of homology between their overlap regions (S. E. Nunes-Düby, A. Landy, unpublished results).

A particularly challenging aspect of synapsis is how two recombining *att* sites initially come together and are productively oriented with respect to one another. The fact that recombination can occur intermolecularly, intramolecularly, and between directly repeated or inverted sites suggests that *att* sites come together by random collision rather than by some form of tracking along the DNA. This general view is substantiated by topological analyses of the knots and catenanes generated by recombination: the topological complexity of recombinant products increases with both the density of supercoils and the distance between the *att* sites (130, 132, 164, 166, 181).

In contrast to the Int family, the Resolvase-Invertase family has more stringent orientation requirements and yields topologically invariant recombination products (33, 34, 39, 40). These features are consistent with a class of models in which the two recombination sites are interwrapped in a very specific manner, a requirement that is equivalent to a "topological filter" (182–185). From the topological experiments, it is clear that λ att sites do not have such an interwrapping requirement for productive synapsis. [Although an interesting exception may be the Xis-independent attL × attR recombination in low salt (84).] The usefulness of topological analyses of recombinant products lies mainly in its ability to set limits on acceptable models (40, 130, 166) when the number of biochemical unknowns is small. However, any acceptable mechanism for synapsis and strand exchange must be validated by its ability to explain the topological consequences of recombination.

Several features of the λ system are quite similar to those found in FLP and Cre, while others appear to be influenced by the unique arms of λ *att* sites. Both FLP (68, 186–188) and Cre (189–191) share with λ all those properties that suggest a random collision-type of synapsis. Cre, however, appears to exhibit an additional constraint in synapsis, suggesting some interwrapping of *lox* sites (189, 191, 192). However, it is unlikely that this aspect of synapsis reflects a fundamental difference in mechanism, since relatively small mutational changes in either Cre or *lox* can make the Cre system topologically similar to FLP and Int (192).

DIRECTIONALITY AND REGULATION

One of the unique and intriguing aspects of the site-specific recombination of λ (and related phages) is its extreme directionality. Although the pathway is commonly diagrammed as a reversible integration/excision reaction, and the normal products of one reaction are the normal substrates of the other, such notation belies the fact that there are in fact two distinct pathways. Integrative recombination normally depends upon supercoiling for assembly of the attP intasome (22, 169), while the attL and attR intasomes do not have this requirement (although they are stimulated by it) (133). Integrative recombination requires higher levels of Int than excisive recombination, both in vivo (193) and in vitro (when *att*R is supercoiled) (27, S. E. Nunes-Düby, A. Landy, unpublished results). Xis protein is required for excisive recombination under physiological conditions, while it is inhibitory for integrative recombination (80, 83, 194). The H1 site must be occupied by IHF for integrative recombination and it must be vacant for excision (27, 28, 88, 154). Of a total constellation of 13 protein binding sites, P1, H1, and P'3 are uniquely required for integrative recombination, and P2, X1, X2, F, and P'1 are uniquely required for, or involved in, excisive recombination (27–29, 87, 88, 153). Indeed, deletion of the P1 and H1 sites, which are required for *att*P function, makes an *att*R that can be more efficient than wild type (27). It is clear that the λ *att* sites have evolved to efficiently execute one reaction or the other in a mutually exclusive manner.

The regulation of phage integration and prophage excision consists of a sophisticated network of pathways involving multiple phage and host genes. A detailed description of how these pathways direct recombination, and some models for their regulation, have recently been presented (4). Several features of these models are summarized below.

In contrast to the analysis of gene expression, regulation at the level of the recombination reaction itself has been difficult to study in vivo. The in vitro data, and resulting models pointing to mechanism-based regulation, have to be tested by in vivo experiments specifically designed to detect such effects. For example, the large effect of FIS on in vivo excision, discussed above, was not detected until specifically looked for. The other mechanism-based regulation by a host-encoded protein involves IHF and the H1 binding site. The effect of IHF inhibition of excision is to greatly sensitize the reaction of Int concentration via the cooperative and competitive interactions described above (28, 88). Under one set of in vitro conditions, IHF inhibition at H1 sharpens the Int concentration dependence of excision by more than 10-fold (28).

The two host-encoded proteins involved in λ site-specific recombination

both exert regulatory effects on excision; FIS stimulates and IHF inhibits. It is significant that both regulatory responses are overridden by high levels of phage-encoded proteins: high Xis levels abolish FIS stimulation and high Int levels overcome IHF inhibition. These mechanisms endow the prophage excision reaction with a responsiveness to host physiology, viral physiology, and a range of environmental conditions (4, 28, 29). When there is a strong inducing stimulus, such as heavy UV irradiation, death of the host cell is imminent and it is advantageous for the prophage to excise and initiate a lytic cycle without delay or modulation. These are the conditions when levels of Int and Xis are expected to be high. When the inducing stimulus is weak and cell death is not certain, it is advantageous for viral excision and lytic development to be influenced by other factors, such as the likelihood of completing a lytic cycle. These are the conditions when Int and Xis levels may be low. Additionally, in the unfavorable conditions of stationary phase, FIS levels are low and excision is more difficult (requires higher levels of Xis).

The ability of FIS to stimulate excision at low levels of Xis may also be relevant in the phenomenon of spontaneous phage production, which occurs at a frequency of one in 10^3 to 10^5 cell generations. Some "spontaneous" phage production probably results from DNA damage and a strong induction in a small fraction of cells. Additionally, some cells may produce phage as a consequence of random fluctuations in repressor concentration, leading to a weak induction. It has been suggested (4) that the possible role of FIS in assisting this low-frequency pathway is not unlike the FIS stimulation of recombination of other low-frequency recombinations (112, 113, 115, 116). Both the *hin* inversion system of *Salmonella* flagellar antigens (195) and the *gin* inversion system of bacteriophage Mu tail fibers (196) also involve fluctuations in the synthesis of a poorly expressed recombinase. The stimulation of all three low-frequency reactions by FIS suggests that they will occur preferentially in exponentially growing cells.

COMMENTS AND SPECULATIONS

Cooperative Systems

In retrospect we can see why it was so difficult to isolate *att* site defective mutants (despite the existence of many potential targets for disruption). In highly cooperative complexes like intasomes, it is likely that small changes in specific protein-binding sites will be masked by the remaining functional sites. Indeed, even with the ability to make site-directed mutations, the phenotypes can be misleading. A two-base change in the consensus recognition sequence, which reduced IHF binding to the H1 site by 500-fold, yielded little change in phenotype. Only after the introduction of five base changes could it be determined that the H1 site is required for integrative recombina-

tion (28). Thus, in highly cooperative systems, considerable caution must be exercised in interpreting null phenotypes. Another example of confusing or misleading phenotypes was the observation that, under certain recombination conditions, deletion of the P2 Int binding site produced much less of a defect than a multisite mutant that destroyed its function. In fact, the interplay of competitive and cooperative binding in the P arm explains this surprising result, and is substantiated by biochemical analysis of the two mutants (87). A final example of considerations relevant to complex cooperative systems concerns the conditions under which apparently independent binding sites are actually used. The most prominent instances of this are the P arm binding site interactions (88), and the recent finding that some (or all) of the core-type sites do not obtain Int from solution (19), but through delivery by arm-type sites (L. Moitoso de Vargas, S. H. Kim, A. Landy, unpublished results).

We have also learned that optimization of an in vitro system can result in overlooking some of the regulatory and biochemical complexity that is important under nonoptimal conditions. Indeed, nonoptimal conditions are prevalent in nature. The effect of FIS protein could not have been detected in an in vitro system optimized for Xis concentration.

Dynamic Aspects

In considering how two *att* sites might synapse, it is useful to summarize the features of Int binding in the region of strand exchange. Both of the phosphodiester bonds that are cut during exchange of each strand, and all of the contacts for Int binding, are on the same face of the DNA helix, straddling the central major groove of the overlap region (50). The recent demonstration that synapsis does not require DNA homology across the entire 7 bp overlap region (13, 18, 19) rules out models of four-stranded DNA and eliminates the motivation for proposing extensive DNA-DNA interactions in the early steps of synapsis. This makes even more appealing the view that the Int proteins lie on the inside faces of a synaptic complex where they can interact with one another. One particularly attractive configuration is that of two helixes lying across one another (as opposed to a side-by-side alignment). This would serve to maximize the interaction between all of the bound Int proteins, as suggested by some preliminary data on the resolution of synthetic Holliday junctions and the properties of certain suicide recombination substrates (B. Franz, S. H. Kim, S. E. Nunes-Düby, A. Landy, unpublished results). It would also minimize the repulsive forces between the two DNA helixes. This view of the synaptic complex, with an "interior" cluster of Int molecules in the region of strand exchange, is similar to that suggested for γδ Resolvase on the basis of recent mutant and crystallographic data. These indicate a tetrameric complex of Resolvase monomers with the active-site serines on the outer surface (T. Steitz, N. Grindley, personal communication).

Structural Aspects

The sharp ($\geq$ 140°) bends induced at the loci of binding by IHF, Xis, and FIS, along with the bivalent DNA-binding capacity of Int, are probably the major determinants of intasome structure. Intrinsic DNA curvature, as a consequence of A tracts (124–126), is also likely to contribute to the overall structures. Within the intasome, these elements are likely to work in conjunction to form DNA loops that are tethered by an Int, binding simultaneously to arm-type and core-type sites. The delivery of Int to the low-affinity core-type sites, by the higher-affinity arm-type sites, would be facilitated by the protein-induced DNA bending. Binding at core-type sites could also be assisted through secondary (protein-protein) interactions. For example, one Int bound to both arm- and core-type sites could bring another Int to the core region, either on the same *att* site or on the synapsed partner. This view is also consistent with the cooperativity of Int binding in the core regions of all four *att* sites (50) (B. Franz, S. H. Kim, S. Nunes-Düby, L. Moitoso de Vargas, A. Landy, unpublished results).

This model of promoting binding to the core-type sites by the formation of DNA loops that deliver Int bound at arm-type sites predicts specific pairwise interactions between each core-type site and an arm-type site, or between two core-type sites. These interactions could involve sites within the same partner (*cis* interactions), and/or sites on different partners (*trans* interactions) (Figure 6). Similarly, the cooperative interactions between proteins bound at core-type sites could be *cis* and/or *trans*. Identifying these pairwise interactions is one of the challenges of understanding the intasome structures. A related challenge is characterization of the interactions between DNA and the site-specific DNA bending proteins IHF, Xis, FIS.

Regulatory Aspects

Bacteriophage λ has long served as a divining rod for uncovering host functions. Yet another example may be the host-encoded IHF and FIS proteins. It is tempting to speculate that the mechanism-based regulation of recombination incorporates IHF and FIS as integral elements of the reaction, because these two proteins reflect global signals of *E. coli* physiology. This suggestion is appealing because of the striking variation in FIS levels as a function of growth phase (29) and the wide range of reactions in which IHF is involved (111).

The role of FIS in stimulating excision has led to speculation about a new phage regulatory pathway that could maintain derepression even under weak inducing conditions (4, 29). After excision, new genes from the *b* region are expressed by N-dependent transcription through *att*P. One of these (*ben*) encodes a double-stranded DNA endonuclease (197, 198) that has been purified and found to cleave only supercoiled DNA (199). This endonuclease,

or one of the others encoded in the *b* region (22, 200, 201), could play a role in prolonging (inducing) an SOS response by cleaving the bacterial genome and generating new sites that activate RecA protein. The activated RecA would promote continuous cleavage of λ, thus preventing repression of the newly excised phage.

The model proposes a function for the enigmatic *b* region (202) and, more importantly, it suggests how FIS could enhance excision, even under conditions that might otherwise lead to repression of the excised phage. According to this model, spontaneous phage production does not reflect the lower limits of prophage repression, but rather is a programmed strategy in maintaining phage populations. In the absence of a pathway devoted to maintaining derepression (under weak inducing conditions), the dramatic effects of FIS on prophage excision would be primarily relevant to normal induction conditions.

CLOSING THE CIRCLE

Analysis of the λ recombination system has benefitted greatly from Robin Holliday's provocative insight about possible intermediates in the exchange (or rearrangement) of genetic information. In return, studies of the λ pathway have contributed significantly to understanding more about the formation, properties, and resolution of these genetic intermediates. The Holliday structures were conceived in the context of homologous, not site-specific, recombination—but such extrapolations are commonplace. One example is the possible relevance of intasome structures to the complex organizations involved in DNA replication, transcription, and other recombination systems. Understanding how the accessory proteins IHF, Xis, and FIS bend and shape DNA and how the bivalent Int protein tethers two distant DNA sites is likely to be of interest well beyond the field of site-specific recombination. Another example is the possible window into global regulatory devices that is afforded by the discovery of IHF and FIS proteins and their role in mechanism-based regulation of recombination.

The circle is always closing but it is never closed.

ACKNOWLEDGMENTS

I thank the many colleagues who provided information and manuscripts prior to publication. I am greatly indebted to the members of our research group, past and present, for their ideas, criticism and enthusiasm, especially to John Thompson, Simone Nunes-Düby, and Lina Moitoso de Vargas, who made significant contributions to this review. I also thank Joan Boyles for her dedicated and good-natured assistance in preparing the manuscript.

Literature Cited

1. Holliday, R. 1964. *Genet. Res.* 5:282–304
2. Nash, H. A. 1981. *Annu. Rev. Genet.* 15:143–67
3. Weisberg, R. A., Landy, A. 1983. In *Lambda II*, ed. F. W. Stahl, J. Roberts, R. A. Weisberg, pp. 211–50. New York: Cold Spring Harbor Lab.
4. Thompson, J. F., Landy, A. 1989. See Ref. 44, pp. 1–22
5. Campbell, A. M. 1962. *Adv. Genet.* 11:101–45
6. Kikuchi, Y., Nash, H. A. 1978. *J. Biol. Chem.* 253:7149–57
7. Kikuchi, Y., Nash, H. A. 1979. *Proc. Natl. Acad. Sci. USA* 76:3760–64
8. Nash, H. A., Robertson, C. A. 1981. *J. Biol. Chem.* 256:9246–53
9. Landy, A., Ross, W. 1977. *Science* 197:1147–60
10. Hsu, P.-L., Ross, W., Landy, A. 1980. *Nature* 285:85–91
11. Mizuuchi, M., Mizuuchi, K. 1980. *Proc. Natl. Acad. Sci. USA* 77:3220–24
12. Hsu, P.-L., Landy, A. 1984. *Nature* 311:721–26
13. Nunes-Düby, S. E., Matsumoto, L., Landy, A. 1987. *Cell* 50:779–88
14. Kitts, P. A., Nash, H. A. 1988. *J. Mol. Biol.* 204:95–108
15. Pargellis, C. A., Nunes-Düby, S. E., Moitoso de Vargas, L., Landy, A. 1988. *J. Biol. Chem.* 263:7678–85
16. Weisberg, R. A., Enquist, L. W., Foeller, C., Landy, A. 1983. *J. Mol. Biol.* 170:319–42
17. Bauer, C. E., Gardner, J. F., Gumport, R. I. 1985. *J. Mol. Biol.* 181:187–97
18. Kitts, P. A., Nash, H. A. 1987. *Nature* 329:346–48
19. Richet, E., Abcarian, P., Nash, H. A. 1988. *Cell* 52:9–17
20. Better, M., Lu, C., Williams, R. C., Echols, H. 1982. *Proc. Natl. Acad. Sci. USA* 79:5837–41
21. Better, M., Wickner, S., Auerbach, J., Echols, H. 1983. *Cell* 32:161–68
22. Richet, E., Abcarian, P., Nash, H. A. 1986. *Cell* 46:1011–21
23. Thompson, J. F., Landy, A. 1988. *Nucleic Acids Res.* 16:9687–705
24. Moitoso de Vargas, L., Pargellis, C. A., Hasan, N. M., Bushman, E. W., Landy, A. 1988. *Cell* 54:923–29
25. Argos, W., Landy, A., Abremski, K., Egan, J. B., Haggård-Ljungquist, E., et al. 1986. *EMBO J.* 5:433–40
26. Bushman, W., Yin, S., Thio, L. L., Landy, A. 1984. *Cell* 39:699–706
27. Bushman, W., Thompson, J. F., Vargas, L., Landy, A. 1985. *Science* 230:906–11
28. Thompson, J. F., Waechter-Brulla, D., Gumport, R. I., Gardner, J. F., Moitoso de Vargas, L., Landy, A. 1986. *J. Bacteriol.* 168:1343–51
29. Thompson, J. F., Moitoso de Vargas, L., Koch, C., Kahmann, R., Landy, A. 1987. *Cell* 50:901–8
30. Sadowski, P. 1986. *J. Bacteriol.* 165:341–47
31. Craig, N. L. 1988. *Annu. Rev. Genet.* 22:77–106
32. Hatfull, G. F., Salvo, J. J., Falvey, E. E., Rimphanitchayakit, V., Grindley, N. D. F. 1988. See Ref. 42, pp. 149–81
33. Hatfull, G. F., Grindley, N. D. F. 1988. See Ref. 43, pp. 357–96
34. Glasgow, A. C., Hughes, K. T., Simon, M. I. 1989. See Ref. 44, pp. 637–59
35. Johnson, R. C., Simon, M. I. 1987. *Trends Genet.* 3:262–67
36. West, S. C. 1989. In *Nucleic Acids and Molecular Biology*, Vol. 3, ed. F. Eckstein, D. M. J. Lilley. Berlin: Springer Verlag. In press
37. West, S. C. 1988. *Trends Genet.* 4:8–13
38. Cox, M. M. 1989. See Ref. 44, pp. 661–70
39. Gellert, M., Nash, H. 1987. *Nature* 325:401–4
40. Benjamin, H. W., Cozzarelli, N. R. 1985. In *Proc. Robert A. Welch Found. Conf. Chem. Res. XXIX. Genetic Chemistry and the Molecular Basis of Heredity*, pp. 107–26. Houston: Robert A. Welch Found.
41. Wasserman, S. A., Cozzarelli, N. R. 1986. *Science* 232:951–60
42. Kingsman, A. J., Kingsman, S. M., Chater, K. F., eds. 1988. *Transposition*. Cambridge, England: Cambridge Univ. Press
43. Kucherlapati, R., Smith, G. R., eds. 1988. *Genetic Recombination*. Washington, DC: Am. Soc. Microbiol.
44. Berg, D. E., Howe, M. M., eds. 1989. *Mobile DNA*. Washington, DC: Am. Soc. Microbiol.
45. Hoess, R. H., Foeller, C., Bidwell, K., Landy, A. 1980. *Proc. Natl. Acad. Sci. USA* 77:2482–86
46. Davies, R. W. 1980. *Nucleic Acids Res.* 8:1765–82
47. Kotewicz, M., Grzesiuk, E., Courschesne, W., Fischer, R., Echols, H. 1980. *J. Biol. Chem.* 255:2433–39
48. Nash, H. A. 1981. *Enzymes* 14:471–80
49. Ross, W., Landy, A. 1982. *Proc. Natl. Acad. Sci. USA* 79:7724–28
50. Ross, W., Landy, A. 1983. *Cell* 33:261–72
51. Craig, N. L., Nash, H. A. 1983. *Cell* 35:795–803

52. Lange-Gustafson, B. J., Nash, H. A. 1984. *J. Biol. Chem.* 259:12724–32
53. Belfort, M. 1980. *Gene* 11:149–55
54. Schindler, D., Echols, H. 1981. *Proc. Natl. Acad. Sci. USA* 78:4475–79
55. Guarneros, G., Montanez, C., Hernandez, T., Court, D. 1982. *Proc. Natl. Acad. Sci. USA* 79:238–42
56. Schmeissner, U., McKenney, K., Rosenberg, M., Court, D. 1984. *J. Mol. Biol.* 176:39–53
57. Montanez, C., Bueno, J., Schmcissner, U., Court, D. L., Guarneros, G. 1986. *J. Mol. Biol.* 191:29–37
58. Shimada, K., Campbell, A. 1974. *Proc. Natl. Acad. Sci. USA* 71:237–41
59. Leong, J., Nunes-Düby, S., Oser, A. B., Lesser, C., Youderian, P., et al. 1986. *J. Mol. Biol.* 189:603–16
60. Pierson, L. S. III, Kahn, M. L. 1987. *J. Mol. Biol.* 196:487–96
61. Yu, A., Bertani, L. E., Haggård-Ljungquist, E. 1989. *Gene.* In press
62. Ljungquist, E., Bertani, L. E. 1983. *Mol. Gen. Genet.* 192:87–94
63. Bear, S. E., Clemens, J. B., Enquist, L. W., Zagursky, R. J. 1987. *J. Bacteriol.* 169:5880–83
64. Miller, H. I., Mozola, M. A., Friedman, D. I. 1980. *Cell* 20:721–29
65. Kitts, P. A., Nash, H. A. 1988. *Nucleic Acids Res.* 16:6839–56
66. Enquist, L. W., Weisberg, R. A. 1984. *Mol. Gen. Genet.* 195:62–69
67. Dayhoff, M. O., Schwartz, R. M., Orcott, B. L. 1978. In *Atlas of Protein Sequence and Structure,* ed. M. O. Dayhoff, pp. 345–52. Washington, DC: Natl. Biomed. Res. Found.
68. Gronostajski, R. M., Sadowski, P. D. 1985. *Mol. Cell. Biol.* 5:3274–79
69. Prasad, P. V., Young, L.-J., Jayaram, M. 1987. *Proc. Natl. Acad. Sci. USA* 84:2189–93
70. Wierzbicki, A., Kendall, M., Abremski, K., Hoess, R. 1987. *J. Mol. Biol.* 195:785–94
71. Reed, R. R., Moser, C. D. 1984. *Cold Spring Harbor Symp. Quant. Biol.* 49:245–49
72. Hatfull, G. F., Grindley, N. D. F. 1986. *Proc. Natl. Acad. Sci. USA* 83:5429–33
73. Klippel, A., Mertens, G., Patschinsky, T., Kahmann, R. 1988. *EMBO J.* 7: 1229–37
74. Ross, W., Landy, A., Kikuchi, Y., Nash, H. 1979. *Cell* 18:297–307
75. Nussinov, R., Weisberg, R. A. 1986. *J. Biomol. Struct. Dyn.* 3:1133–44
76. Enquist, L. W., Weisberg, R. A. 1977. *J. Mol. Biol.* 111:97–120
77. Gingery, R., Echols, H. 1967. *Proc. Natl. Acad. Sci. USA* 58:1507–14
78. Gottesman, M. E., Yarmolinsky, M. 1968. *J. Mol. Biol.* 31:487–505
79. Zissler, J. 1967. *Virology* 31:189
80. Abremski, K., Gottesman, S. 1982. *J. Biol. Chem.* 257:9658–62
81. Weisberg, R. A., Gottesman, M. E. 1971. In *Bacteriophage Lambda,* ed. A. D. Hershey, pp. 489–500. New York: Cold Spring Harbor Lab. Press
82. Gottesman, S., Gottesman, M., Shaw, J. E., Pearson, M. L. 1981. *Cell* 24:225–33
83. Abremski, K., Gottesman, S. 1981. *J. Mol. Biol.* 153:67–78
84. Craig, N. L., Nash, H. A. 1983. In *Mechanism of DNA Replication and Recombination,* ed. N. R. Cozzarelli, pp. 617–36. New York: Liss
85. Thompson, J. F., Snyder, U. K., Landy, A. 1988. *Proc. Natl. Acad. Sci. USA* 85:6323–27
86. Yin, S., Bushman, W., Landy, A. 1985. *Proc. Natl. Acad. Sci. USA* 82:1040–44
87. Thompson, J. F., Moitoso de Vargas, L., Nunes-Düby, S. E., Pargellis, C., Skinner, S. E., Landy, A. 1987. In *Mechanisms of DNA Replication and Recombination,* ed. T. Kelly, R. McMacken, pp. 735–44. New York: Liss
88. Thompson, J. F., Moitoso de Vargas, L., Skinner, S. E., Landy, A. 1987. *J. Mol. Biol.* 195:481–93
89. Miller, H. I., Friedman, D. I. 1980. *Cell* 20:711–19
90. Kikuchi, A., Flamm, E., Weisberg, R. A. 1985. *J. Mol. Biol.* 183:129–40
91. Miller, H. I., Kikuchi, A., Nash, H. A., Weisberg, R. A., Friedman, D. I. 1979. *Cold Spring Harbor Symp. Quant. Biol.* 43:1121–26
92. Miller, H. I., Nash, H. A. 1981. *Nature* 290:523–26
93. Miller, H. I. 1984. *Cold Spring Harbor Symp. Quant. Biol.* 49:691–98
94. Flamm, E. L., Weisberg, R. A. 1985. *J. Mol. Biol.* 183:117–28
95. Mechulam, Y., Fayat, G., Blanquet, S. 1985. *J. Bacteriol.* 163:787–91
96. Drlica, K., Rouviere-Yaniv, J. 1987. *Microbiol. Rev.* 51:301–19
97. Craig, N. L., Nash, H. A. 1984. *Cell* 39:707–16
98. Leong, J. M., Nunes-Düby, S., Lesser, C. F., Youderian, P., Susskind, M. M., Landy, A. 1985. *J. Biol. Chem.* 260:4468–77
99. Gamas, P., Chandler, M. G., Prentki, P., Galas, D. J. 1987. *J. Mol. Biol.* 195:261–72
100. Stenzel, T. T., Patel, P., Bastia, D. 1987. *Cell* 49:709–17
101. Prentki, P., Chandler, M., Galas, D. J. 1987. *EMBO J.* 6:2479–87

102. Robertson, C. A., Nash, H. A. 1988., *J. Biol. Chem.* 263:3554–57
103. Kosturko, L. D., Daub, E., Murialdo, H. 1989. *Nucleic Acids Res.* 17:17–34
104. Tanaka, I., Appelt, K., Dijk, J., White, S. W., Wilson, K. S. 1984. *Nature* 310:376–81
105. Friedrich, M. J., DeVeaux, L. C., Kadner, R. J. 1986. *J. Bacteriol.* 167:928–34
106. Nash, H. A., Robertson, C. A., Flamm, E., Weisberg, R. A., Miller, H. I. 1987. *J. Bacteriol.* 169:4124–27
107. Miller, H. I., Friedman, D. I. 1977. In *DNA Insertion Elements,* ed. A. I. Bukhari, J. A. Shapiro, S. L. Adhya, pp. 349–56. New York: Cold Spring Harbor Lab.
108. Morisato, D., Kleckner, N. 1987. *Cell* 51:101–11
109. Way, J. C., Kleckner, N. 1986. *Proc. Natl. Acad. Sci. USA* 81:3452–56
110. Surette, M. G., Chaconas, G. 1989. *J. Biol. Chem.* 264:3028–34
111. Friedman, D. I. 1988. *Cell* 55:545–54
112. Johnson, R. C., Simon, M. I. 1985. *Cell* 41:781–91
113. Kahmann, R., Rudt, F., Koch, C., Mertens, G. 1985. *Cell* 41:771–80
114. Huber, H. E., Iida, S., Arber, W., Bickle, T. A. 1985. *Proc. Natl. Acad. Sci. USA* 82:3776–80
115. Johnson, R. C., Bruist, M. F., Simon, M. I. 1986. *Cell* 46:531–39
116. Koch, C., Kahmann, R. 1986. *J. Biol. Chem.* 261:15673–78
117. Johnson, R. C., Ball, C. A., Pfeffer, D., Simon, M. I. 1988. *Proc. Natl. Acad. Sci. USA* 85:3484–88
118. Koch, C., Vandekerckhove, J., Kahmann, R. 1988. *Proc. Natl. Acad. Sci. USA* 85:4237–41
119. Pabo, C., Sauer, R. 1984. *Annu. Rev. Biochem.* 53:293–321
120. Plasterk, R. H. A., van de Putte, P. 1984. *Biochim. Biophys. Acta* 782: 111–19
121. Johnson, R. C., Glasgow, A. C., Simon, M. I. 1987. *Nature* 329:462–65
122. Bruist, M. F., Glasgow, A. C., Johnson, R. C., Simon, M. I. 1987. *Genes Dev.* 1:762–72
123. Mizuuchi, M., Mizuuchi, K. 1985. *Nucleic Acids Res.* 13:1193–208
124. Marini, J. C., Levene, S. D., Crothers, D. M., Englund, P. T. 1982. *Proc. Natl. Acad. Sci. USA* 79:7664–68
125. Wu, H.-M., Crothers, D. M. 1984. *Nature* 308:509–13
126. Diekmann, S., Wang, J. C. 1985. *J. Mol. Biol.* 186:1–11
127. Thompson, J. F., Mark, H. F., Franz, B., Landy, A. 1988. In *DNA Bending and Curvature,* ed. W. K. Olson, M. H. Sarma, R. H. Sarma, M. Sundaralingam, pp. 119–28. Guilderland, NY: Adenine
128. Kitts, P., Richet, E., Nash, H. A. 1984. *Cold Spring Harbor Symp. Quant. Biol.* 49:735–44
129. Laundon, C. H., Griffith, J. D. 1988. *Cell* 52:545–49
130. Spengler, S. J., Stasiak, A., Cozzarelli, N. R. 1985. *Cell* 42:325–34
131. Mizuuchi, K., Gellert, M., Nash, H. 1978. *J. Mol. Biol.* 121:375–92
132. Pollock, T. J., Nash, H. A. 1983. *J. Mol. Biol.* 170:1–18
133. Abremski, K., Gottesman, S. 1979. *J. Mol. Biol.* 131:637–49
134. Shulman, M., Gottesman, M. 1973. *J. Mol. Biol.* 81:461–82
135. Ross, W., Landy, A. 1982. *J. Mol. Biol.* 156:505–29
136. DeMassy, B., Studier, F. W., Dorgai, L., Appelbaum, E., Weisberg, R. A. 1984. *Cold Spring Harbor Symp. Quant. Biol.* 49:715–26
137. Hoess, R. H., Abremski, K. 1985. *J. Mol. Biol.* 181:351–62
138. Hoess, R., Abremski, K., Sternberg, N. 1984. *Cold Spring Harbor Symp. Quant. Biol.* 49:761–68
139. Hoess, R. H., Wierzbicki, A., Abremski, K. 1986. *Nucleic Acids Res.* 14: 2287–300
140. Abremski, K., Wierzbicki, A., Frommer, B., Hoess, R. H. 1986. *J. Biol. Chem.* 261:391–96
141. Hoess, R., Wierzbicki, A., Abremski, K. 1987. *Proc. Natl. Acad. Sci. USA* 84:6840–44
142. Andrews, B. J., Beatty, L. G., Sadowski, P. D. 1987. *J. Mol. Biol.* 193:345–58
143. Senecoff, J. F., Cox, M. M. 1986. *J. Biol. Chem.* 261:7380–86
144. Umlauf, S. W., Cox, M. M. 1988. *EMBO J.* 7:1845–52
145. Meyer-Leon, L., Huang, L.-C., Umlauf, S. W., Cox, M. M., Inman, R. B. 1988. *Mol. Cell. Biol.* 8:3784–96
146. Andrews, B. J., McLeod, M., Broach, J. R., Sadowski, P. D. 1986. *Mol. Cell. Biol.* 6:2482–89
147. Iida, S., Hiestand-Nauer, R. 1987. *Mol. Gen. Genet.* 208:464–68
148. Reed, R. R., Grindley, N. D. F. 1981. *Cell* 25:721–28
149. Shimada, K., Weisberg, R. A., Gottesman, M. E. 1972. *J. Mol. Biol.* 63:483–503
150. Shimada, K., Weisberg, R. A., Gottesman, M. E. 1972. *J. Mol. Biol.* 80:297–314
151. Shimada, K., Weisberg, R. A., Gottesman, M. E. 1975. *J. Mol. Biol.* 93:415–29

152. Winoto, A., Chung, S., Abraham, J., Echols, H. 1986. *J. Mol. Biol.* 192:677–80
153. Bauer, C. E., Hesse, S. D., Gumport, R. I., Gardner, J. F. 1986. *J. Mol. Biol.* 192:513–27
154. Gardner, J. F., Nash, H. A. 1986. *J. Mol. Biol.* 191:181–89
155. Krause, H. M., Higgins, N. P. 1986. *J. Biol. Chem.* 261:3744–52
156. Filutowicz, M., Appelt, K. 1988. *Nucleic Acids Res.* 16:3829–32
157. Gartenberg, M. R., Crothers, D. M. 1988. *Nature* 333:824–29
158. Davies, R. W., Schreier, P. H., Kotewicz, M. L., Echols, H. 1979. *Nucleic Acids Res.* 7:2255–73
159. Rhodes, D., Klug, A. 1980. *Nature* 286:573–78
160. Klug, A., Lutter, L. C. 1981. *Nucleic Acids Res.* 9:4267–83
161. Liu, L. F., Wang, J. 1978. *Cell* 15:979–84
162. Drew, H. R., Travers, A. A. 1985. *J. Mol. Biol.* 186:773–90
163. Hamilton, D., Yuan, R., Kikuchi, Y. 1981. *J. Mol. Biol.* 152:163–69
164. Mizuuchi, K., Gellert, M., Weisberg, R. A., Nash, H. A. 1980. *J. Mol. Biol.* 141:485–95
165. Mizuuchi, K., Fisher, L. M., O'Dea, M. H., Gellert, M. 1980. *Proc. Natl. Acad. Sci. USA* 72:1847–51
166. Griffith, J. D., Nash, H. A. 1985. *Proc. Natl. Acad. Sci. USA* 82:3124–28
167. Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M., Chambon, P. 1975. *Proc. Natl. Acad. Sci. USA* 72:1843–47
168. Richmond, T. J., Finch, J. T., Klug, A. 1983. *Cold Spring Harbor Symp. Quant. Biol.* 47:493–501
169. Mizuuchi, K., Mizuuchi, M. 1979. *Cold Spring Harbor Symp. Quant. Biol.* 43:1111–14
170. Pollock, T. J., Abremski, K. 1979. *J. Mol. Biol.* 131:651–54
171. Mizuuchi, K., Weisberg, R., Enquist, L., Mizuuchi, M., Buraczynska, M., et al. 1981. *Cold Spring Harbor Symp. Quant. Biol.* 45:429–37
172. Echols, H., Green, L. 1979. *Genetics* 93:297–307
173. Enquist, L. W., Nash, H., Weisberg, R. A. 1979. *Proc. Natl. Acad. Sci. USA* 76:1363–67
174. Bauer, C. E., Hesse, S. D., Gardner, J. F., Gumport, R. I. 1984. *Cold Spring Harbor Symp. Quant. Biol.* 49:699–705
175. Jayaram, M., Crain, K. L., Parsons, R. L., Harshey, R. M. 1988. *Proc. Natl. Acad. Sci. USA* 85:7902–6
176. Nash, H. A., Mizuuchi, K., Enquist, L. W., Weisberg, R. A. 1981. *Cold Spring Harbor Symp. Quant. Biol.* 45:417–27
177. Nash, H. A., Pollock, T. J. 1983. *J. Mol. Biol.* 170:19–38
178. McGavin, S. 1971. *J. Mol. Biol.* 55:293–98
179. Wilson, J. H. 1979. *Proc. Natl. Acad. Sci. USA* 76:3641–45
180. Nash, H. A., Bauer, C. E., Gardner, J. F. 1987. *Proc. Natl. Acad. Sci. USA* 84:4049–53
181. Bliska, J. B., Cozzarelli, N. R. 1987. *J. Mol. Biol.* 197:205–18
182. Boocock, M. R., Brown, J. L., Sherratt, D. J. 1985. *Biochem. Soc. Trans.* 14: 214–16
183. Craigie, R., Mizuuchi, K. 1986. *Cell* 45:795–800
184. Wasserman, S. A., Cozzarelli, N. R. 1985. *Proc. Natl. Acad. Sci. USA* 82:1079–83
185. Wasserman, S. A., Dungan, J. M., Cozzarelli, N. R. 1985. *Science* 229:171–74
186. Meyer-Leon, L., Gates, C. A., Attwood, J. M., Wood, E. A., Cox, M. M. 1987. *Nucleic Acids Res.* 15:6469–88
187. Vetter, D., Andrews, B. J., Roberts-Beatty, L., Sadowski, P. D. 1983. *Proc. Natl. Acad. Sci. USA* 80:7284–88
188. Beatty, L. G., Babineau-Clary, P., Hogrefe, C., Sadowski, P. D. 1986. *J. Mol. Biol.* 188:529–44
189. Abremski, K., Hoess, R., Sternberg, N. 1983. *Cell* 32:1301–11
190. Abremski, K., Frommer, B., Hoess, R. H. 1986. *J. Mol. Biol.* 192:17–26
191. Abremski, K., Hoess, R. 1985. *J. Mol. Biol.* 184:211–20
192. Abremski, K., Frommer, B., Wierzbicki, A., Hoess, R. H. 1988. *J. Mol. Biol.* 201:1–8
193. Enquist, L. W., Kikuchi, A., Weisberg, R. A. 1979. *Cold Spring Harbor Symp. Quant. Biol.* 43:1115–20
194. Nash, H. A. 1975. *Proc. Natl. Acad. Sci. USA* 72:1072–76
195. van de Putte, P., Cramer, S., Giphart-Gassler, M. 1980. *Nature* 286:218–22
196. Zieg, J., Hillman, M., Simon, M. 1978. *Cell* 15:237–44
197. Benchimol, S., Lucko, H., Becker, A. 1982. *J. Biol. Chem.* 257:5201–10
198. Sumner-Smith, M., Benchimol, S., Murialdo, H., Becker, A. 1982. *J. Mol. Biol.* 160:1–22
199. Benchimol, S., Lucko, H., Becker, A. 1982. *J. Biol. Chem.* 257:5211–19
200. Becker, A. 1970. *Biochem. Biophys. Res. Commun.* 41:63–70
201. Chowdbury, M. R., Dunbar, S., Becker, A. 1972. *Virology* 49:314–18
202. Kellenberger, G., Zichichi, M. L., Weigle, J. 1961. *J. Mol. Biol.* 3:399–408