The EMBO Journal vol.8 no.2 pp.577 – 585, 1989

Mutational analysis of a prokaryotic recombinational enhancer element with two functions

Philipp Hübner and Werner Arber

Department of Microbiology, Biozentrum der Universität Basel, Klingelbergstrasse 70,CH-4056 Basel, Switzerland

Communicated by Werner Arber

The site-specific DNA inversion system Cin encoded by the bacteriophage P1 consists of a recombinase, two inverted crossing-over sites and a recombinational enhancer. The latter ~75 bp long genetic element is bifunctional due to its location within the 5′ part of the *cin* gene encoding the recombinase. In order to determine the essential nucleotides for each of its two biological functions we randomly mutated the recombinational enhancer sequence *sis*(P1) and analysed both functions of the mutants obtained. Three distinct regions of this sequence were found to be important for the enhancer activity. One of them occupies the middle third of the enhancer sequence and it can suffer a number of functionally neutral base substitutions, while others are detrimental. The other two regions occupy the two flanking thirds of the enhancer. They coincide with binding sites of the host-coded protein FIS (Factor for Inversion Stimulation) needed for efficient DNA inversion *in vitro*. These sequences appear to be highly evolved allowing only a few mutations without affecting either of the biological functions. Taking the effect of mutations within these FIS binding sites into account a consensus sequence for the interaction with FIS was compiled. This FIS consensus implies a palindromic structure for the recombinational enhancer. This is in line with the orientation independence of enhancer action with respect to the crossing-over sites.

Key words: bacteriophage P1/FIS protein/random mutagenesis/recombinational enhancer/site-specific recombination

Introduction

Site-specific recombination can provide the means for alternate expression of genes by mediating the inversion of DNA sequences. Such biological switches determine variations in the host range of the bacteriophages P1 (Iida, 1984) and Mu (Kamp *et al.*, 1978), or variations of flagellar antigens in *Salmonella* (Iino and Kutsukake, 1980; Silverman *et al.*, 1980). These site-specific inversion systems consist of a recombinase, two inverted crossing-over sites and a recombinational enhancer (Huber *et al.*, 1985a; Johnson and Simon, 1985; Kahmann *et al.*, 1985). The latter genetic element (called *sis* for sequence for inversion stimulation) is bifunctional. The 75 bp long DNA segment encodes the N-terminal part of the recombinase and it stimulates the site-specific DNA inversion >100-fold in a *cis* dependent manner, largely independent of its orientation and position. All elements of the above mentioned inversion systems are functionally exchangeable (see review by Plasterk and van de Putte, 1984; Huber *et al.*, 1985a). This and their notable sequence homologies indicate a common origin for all members of this DNA inversion family.

In vitro studies revealed that in addition to the recombinase a second, host-coded protein, called FIS (Factor for Inversion Stimulation), is needed for an efficient recombination reaction (Kahmann *et al.*, 1985; Johnson and Simon, 1985). FIS binds specifically to DNA within and outside of the recombinational enhancer sequence (Koch and Kahmann, 1986; Bruist *et al.*, 1987; Haffter and Bickle, 1987). Genetic analysis of the enhancer of the *Salmonella* system revealed that two FIS binding sites are sufficient for enhancer activity provided that they are separated by a well defined distance from each other (Johnson *et al.*, 1987) and that the intervening sequence allows for a conformational change of the enhancer DNA (Hübner *et al.*, 1989b). The replacement of FIS site I by a synthetic site II yielded an active enhancer (Bruist *et al.*, 1987) indicating that the two FIS binding sites are functionally equivalent. A consensus of a FIS binding site was compiled by comparison of the recombinational enhancers carried by the three different, above mentioned DNA inversion systems (Bruist *et al.*, 1987). Since the recombinational enhancer is bifunctional, it is possible that such a consensus also reflects constraints on the coding function. In order to determine the nucleotide requirement for each of the two enhancer functions we randomly mutated the enhancer from bacteriophage P1 (Hübner *et al.*, 1989a) and assayed both functions separately in each mutant using two different inversion test systems. The results of these studies are reported here.

Results and Discussion

Construction of tester plasmids and of sis(P1) mutants

Since the recombinational enhancer *sis*(P1) is bifunctional, two different tester plasmids had to be developed so that each of the mutated DNA fragments could be subcloned like a cassette into each tester plasmid. A *sis*(P1) enhancer cassette was constructed by a *Xba*I linker insertion between the *lac*UV5 promoter and the start of the *cin* gene of the Cin-over-producing plasmid pHHL132 (Huber *et al.*, 1985a). This yielded the plasmid pPHU78 (Figure 1), where the 5′ part of the *cin* gene (C segment inversion of bacteriophage P1) can now be replaced using the restriction sites *Xba*I and *Sph*I. For analysis of the enhancer function the multi-functional inversion tester plasmid pPHU9 (Figure 1) was constructed (see Materials and methods). The *Xba*I/*Sph*I enhancer cassette was then inserted into the corresponding sites of pPHU9 yielding the plasmid pPHU81 which served as positive control during all experiments.

Oligonucleotides with random substitutions served for the

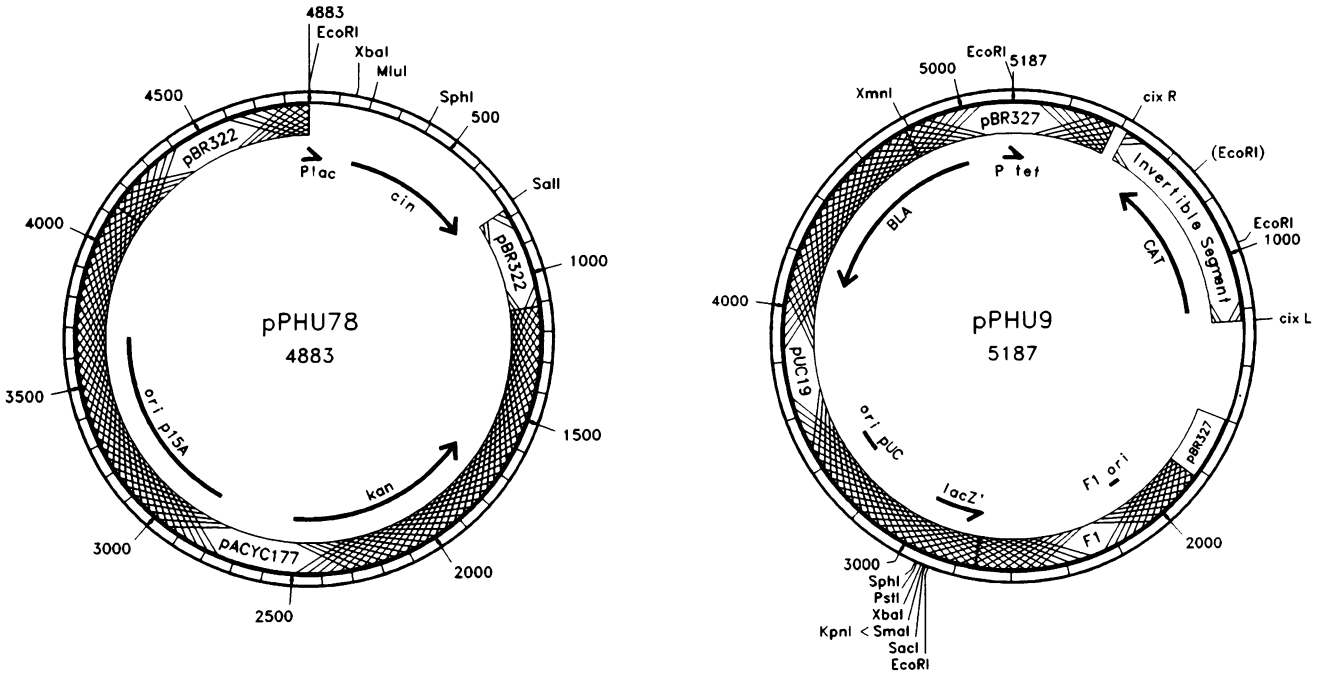

Fig. 1. Physical maps of the Cin over-producing plasmid pPHU78 and of the multifunctional inversion tester plasmid pPHU9. Arrows indicate the direction, either of the genes for β-lactamase (*BLA*), chloramphenicol acetyltransferase (*CAT*), kanamycin resistance (*kan*), C segment DNA inversion enzyme (*cin*) and of the truncated β-galactosidase (*lacZ'*), or of transcription. The positions of the *tet*- and *lac*UV5-promoter are identified with P tet and P lac, respectively. ori p15A, ori pUC and F1 ori, designate origins of DNA replication of the plasmids p15A and pUC and of the bacteriophage f1, respectively. The origins of DNA fragments from different cloning vectors are given within the hatched blocks. cixR and cixL designate the right and left cross-over sites, respectively. Relevant restriction sites and coordinates are shown on the outer circle. (*Eco*RI) indicates the relative position of the *Eco*RI site after Cin mediated DNA inversion.

site-directed mutagenesis of the 75 bp long *sis*(P1) region within *Xba*I/*Sph*I fragment of a phage M13mp19 derivative (Hübner *et al.*, 1989a). This experiment yielded many different single point mutations, some double and a few triple mutations. The *Xba*I/*Sph*I enhancer cassettes carrying mutations were then subcloned into the corresponding sites of pPHU78 and of pPHU9 yielding the constructs compiled in Table I.

Analysis of the enhancer function of the sis(P1) mutants

The enhancer function of the mutated *sis*(P1) fragments was assayed either by a plating or by a DNA restriction assay (see Materials and methods) using the described derivatives of pPHU9. Based on the results obtained (Table I and Figure 2) the mutants were classified as follows: down mutations without any enhancer activity; impaired mutations with reduced activity; neutral mutations with about wild type activity and up mutations with higher activity than wild type (Figure 3).

All 15 down mutations are located either in FIS binding site I or site II indicating that the interaction of FIS protein with both binding sites on the enhancer sequence is essential for the enhancer activity. The apparently unequal distribution of mutants between sites I and II represents an artefact of the mutagenesis procedure (Hübner *et al.*, 1989a) and has no biological significance.

Mutations impairing the enhancer activity were found in three distinct regions of the enhancer: again in sites I and II and in a cluster at the middle of the enhancer sequence which is not efficiently protected by FIS in DNase I footprinting experiments. The four up mutations also fall in these three regions, indicating that sequences between site I and site II also determine the quality of enhancer activity. These sequences may be involved in the FIS protein−DNA interaction leading to a conformational change of the enhancer DNA which is needed for its activity (Hübner *et al.*, 1989b).

Although most neutral mutations reside within the intervening sequence between sites I and II, a few mutations were also found within site I and site II indicating positions not specifically involved in DNA−FIS interaction.

The nature of the DNA sequence interacting with FIS protein

In order to determine the DNA sequence specifically involved in the interaction with FIS protein we compared the naturally occurring FIS binding site I and site II sequences with mutations falling into these sites. This allowed us to compile degenerate consensus sequences for site I and for site II. These consensi were then compared with each other in either orientation by computer in order to align them with respect to each other. The alignment found (Figure 4) is identical to that proposed by Bruist *et al.* (1987) which is based on DMS (dimethylsulphate) methylation interference and protection studies. This sequence alignment and the data obtained from the enhancer mutations were used as a basis for the compilation of an enhancer FIS site (Figure 4). Finally, this sequence was compared with other, non-enhancer DNA sequences protected by FIS (Haffter and Bickle, 1987; Thompson *et al.*, 1987) and a palindromic FIS consensus sequence could be deduced. All the down mutations affect this proposed consensus sequence which indicates that FIS protein indeed interacts with this sequence.

Two of the enhancer up mutations create a palindromic structure around the symmetry axis (either TAA or AAT) indicating a favoured palindromic structure of a FIS binding site for enhancer activity. Furthermore the pattern of purines which are either protected by FIS from DMS methylation or which show methylation interference (Bruist *et al.*, 1987) can be fairly well explained by the proposed consensus sequence. Since DMS methylates G residues at the N_7 position (major groove) and A residues at the N_3 position (minor groove) the direction of the FIS protein–DNA interaction can be deduced from these DMS methylation studies. This analysis shows that two neighbouring major grooves and the minor groove in between interact with FIS protein. These three grooves point in the same direction suggesting that FIS protein interacts with the DNA helix from one side. Since the proposed FIS consensus site is palindromic and FIS in its native state is a dimer (Koch and Kahmann, 1986) we postulate that each FIS monomer interacts with the half-site G/T..YR..A/T. Recently the *fis* gene of *Escherichia coli* was isolated and its DNA sequence was determined (Koch *et al.*, 1988; Johnson *et al.*, 1988). The deduced amino acid sequence displays at its carboxyl end a potential helix-turn-helix DNA binding motif, suggesting that the FIS–DNA interaction is similar to that proposed for a series of other DNA binding proteins with palindromic recognition sequences (Pabo and Sauer, 1984).

Enhancer structure and its implication for a model for enhancer action

Mutational analysis of the recombinational enhancer revealed its specific structure. Three distinct regions were found to be important for enhancer function. Two of these regions coincide with the binding sites for the host factor FIS. Since single point mutations in either FIS binding site can lead to loss of the enhancer activity we conclude that both FIS sites are needed for the enhancer activity. This conclusion is in line with previous deletion studies (Huber *et al.*, 1985a; Johnson and Simon, 1985). The recombinational enhancer is bounded by two palindromic FIS sites, the centres of which are separated by 48 bp (i.e. 4.5 helical turns), implying that the overall structure of the enhancer is also palindromic (see Figure 3F). That palindromic structures are important for the enhancer's function is further indicated by the three different enhancer up mutations which increase either the palindromic structure of a FIS site or of the intervening sequence. Such a structure has no determined direction which explains nicely the observed orientation independence of the enhancer activity with respect to the cross-over sites.

Since FIS contacts its palindromic recognition site from one side of the DNA helix and since the distance between the two FIS sites within the enhancer is 4.5 helical turns, FIS contacts these two FIS sites from opposite sides of the DNA helix. This specific structure of the FIS–enhancer complex and the observed change in linking number (ΔLk) by four during Gin-mediated DNA recombination *in vitro* led to the current model of enhancer action (Bruist *et al.*, 1987; Kanaar *et al.*, 1988). According to this model, which is based on a DNA looping mechanism, FIS bound to the enhancer interacts with the recombinase bound to the cross-over sites. Thereby the stereo-geometrical arrangement of FIS (determined by the enhancer structure described above) determines the three dimensional arrangement of the synaptic complex. The FIS–enhancer complex would, in this context, act as a template in order to help the two recombinase–cross-over site complexes to meet and to interact with each other in the synaptic complex.

For the correct three-dimensional arrangement of this synaptic complex the FIS–enhancer complex has to adopt a well defined structure. This is in line with the finding that an active enhancer sequence has to undergo a conformational change and that FIS binding alone is not sufficient for the enhancer activity. Mutations within the intervening sequence between the two enhancer FIS sites can interfere with the conformational change of the enhancer sequence (Hübner *et al.*, 1989b). The identification of mutations which affect the enhancer activity within this intervening sequence (this study) confirms that this sequence is important for the enhancer action and it more generally shows the importance of the sequence context in which a protein site is embedded. We think that the DNA fragment between the two FIS binding sites has to be flexible enough to undergo bending in order to allow the FIS–enhancer complex to adopt its active three-dimensional structure.

The above described model of recombinational enhancer action is very similar to models describing the action of eukaryotic transcriptional enhancers (Guarente, 1988). All these models are based firstly on a DNA looping mechanism, which explains the position and orientation independence of the enhancer action, and secondly on the interaction of at least two different protein–DNA complexes. A first step in the analysis of the FIS–recombinase interaction is the description of FIS independent *cin* mutants (Haffter and Bickle, 1989). Those *cin* mutants have a relaxed topological specificity (i.e. they promote DNA inversion, DNA co-integration and DNA deletion) indicating that the interaction between the FIS–enhancer complex and the recombinase–cross-over site complexes is important for the high topological specificity of Cin-mediated DNA recombination.

Analysis of the recombinase function of the sis(P1) mutants

The *Xba*I/*Sph*I enhancer cassette of the plasmid pPHU78 (Figure 1) was replaced by corresponding fragments carrying missense mutations of the *cin* gene and the recombinase function was assayed either by a plating or by a DNA restriction analysis (see Materials and methods). The results of these assays are shown in Table I. Three different classes of mutants (inactive, partially active and unaffected) were formed (Figure 5). Since inactive and partially active protein mutations can affect either the three-dimensional structure or active centers of the protein, it is very difficult to deduce what the effect of a particular mutation is. Some mutations however, are interesting. At position 11 the asparagine residue can be replaced by lysine or by serine but not by tyrosine. A similar situation is encountered for Asn-14, where Lys is tolerated but Tyr is not. Asn-20 on the other hand can be replaced by isoleucine. At position 17 the leucine residue can be replaced by phenylalanine but not by Ile or by Ser. It is also surprising that glutamic acid-23 can be replaced by Lys.

For the $\gamma\delta$-resolvase (another site-specific recombinase which is ~30% homologous to the DNA inversion enzymes under consideration) it has been shown that a covalent protein–DNA linkage intermediate is transiently made via

Table I. Effect of *sis*(p1) mutations on the enhancer and recombinase activities

Nr	FIS BINDING SITE	BASE PAIR CHANGE(S)	CODON CHANGE(S)		Activity of Enhancer[X] Plating assay	Enhancer[X] Restriction assay	Recombinase[Y] Plating assay	Recombinase[Y] Restriction assay	Evolution[Z]
0	I	26 C–A 27 A–G	9 Ser–amber		–	–	a		
1	I	19 C–T 37 C–T	7 Arg–Cys	13 Gln–ochre	–	–	a		
2	I	20 G–A	7 Arg–His		++		–		
3	I	23 T–C	8 Val–Ala		+	++	–		
4	I	23 T–G 27 A–G	8 Val–Gly		++		–		
5	I	28 A–G	10 Thr–Ala		+	++	–		
6	I	29 C–T	10 Thr–Ile		+++	+++	–		
7	I	31 A–T	11 Asn–Tyr		+	++	–		
15	I	32 A–G	11 Asn–Ser		+	++	++	++	+
8	I	33 T–C			+	++	b		+
9	I	33 T–G	11 Asn–Lys		–	+	++	++	
10	I	35 A–G	12 Glu–Gly		+	+	+	++	
11	I	Insert. A in 38–41	frame-shift		+	+	a		
12	I	39 A–G			+	++	b		+[3,4]
13	I	39 A–T	13 Gln–His		+	+	–		
14	I	40 A–T	14 Asn–Tyr		++		–		
16	I	42 C–T			+	++	b		+[2,3]
17	I	42 C–A	14 Asn-Lys		+	++	++	++	+
18	I	42 C–T 51 A–T		17 Leu–Phe	+++	+++	+	++	+
19	I	42 C–T 61 G–A		21 Ala–Thr	++		+	(+)	
20	I	44 C–G	15 Thr–Ser		++		(+)	(+)	
21	I	44 C–A 53 A–T	15 Thr–Asn	18 Gln–Leu	++		–		
22		45 T–A			++		b		+[1,3,4]
23		45 T–G 48 T–A			++		b		+
24		46 G–C	16 Ala–Pro		++		–		
25		47 C–T	16 Ala–Val		++		(+)	(+)	+
26		48 T–G			++		b		+
27		49 T–A	17 Leu–Ile		++		–		
28		49 T–C			++		b		+[3,4]
29		49 T–G 58 A–G	17 Leu–Val	20 Asn–Asp	+	++	–		
30		49 T–C 54 A–T		18 Gln–His	++		n.t.		
31		50 T–C	17 Leu–Ser		++		–		
32		51 A–T	17 Leu–Phe		+++	+++	++	++	+
33		51 A–G			++		b		+[3,5]
34		51 A–G 59 A–T		20 Asn–Ile	++		+	+	
35		52 C–G 58 A–C	18 Gln–Glu	20 Asn–His	+	+	–		
36		52 C–A 67 G–A	18 Gln–Lys	23 Glu–Lys	++		–		
37		53 A–C	18 Gln–Pro		++		–		
38		53 A–T	18 Gln–Leu		++		–		
39		54 A–G			++		b		+[2]
40		54 A–T	18 Gln–His		++		–		
41		55 C–A			++		b		+[1]
42		56 G–A	19 Arg–Gln		++		++	+	+
43		56 G–C	19 Arg–Pro		++		–		
44		57 A–T			+	++	b		+[2,4]
45		59 A–G 62 C–A	20 Asn–Ser	21 Ala–Asp	+	+	–		
46		60 C–T			++		b		+[1,2,4,5]
47		60 C–A 62 C–A	20 Asn–Lys	21 Ala–Asp	+	++	–		
48		60 C–T 66 T–A			++		b		+
49		62 C–A 64 C–A	21 Ala–asp	22 Leu–Ile	+	++	–		
50		62 C–T 67 G–T	21 Ala–Val	23 Glu–ochre	+	++	a		
51		63 T–C			++		b		+[5]
52		64 C–T	22 Leu–Phe		++	++	–		
54		64 C–G 67 G–A	22 Leu–Val	23 Glu–Lys	++		–		
55		65 T–A	22 Leu–His		++		–		
56	II	67 G–T	23 Glu–ochre		++		a		
57	II	67 G–A	23 Glu–Lys		++		++	++	+[1,2,4]
58	II	68 A–C	23 Glu–Ala		+	++	+	+	
59	II	68 A–C 77 G–T	23 Glu–Ala	26 Gly–Val	–		a		
		81 T–A	27 Cys–opal						
60	II	69 A–T 86 T–G	23 Glu–Asp	29 Leu–Arg	++		–		
		88 A–C	30 Ile–Leu						

Table I. *(continued)*

Nr	FIS BINDING SITE	BASE PAIR CHANGE(S)	CODON CHANGE(S)		Activity of Enhancer[X] Plating assay	Enhancer[X] Restriction assay	Recombinase[Y] Plating assay	Recombinase[Y] Restriction assay	Evolution[Z]
61	II	70 A–T 87 A–C	24 Ser–Cys		++		+	+	
62	II	72 C–A 75 A–C	24 Ser–Arg		–		+	+	
63	II	72 C–T 77 G–T		26 Gly–Val	++		–		
64	II	72 C–A 82 G–T	24 Ser–Arg	28 Glu–am	–		a		
65	II	72 C–T 82 G–T 84 G–T		28 Glu–Tyr	–		+	++	
66	II	73 G–C	25 Ala–Pro		++		–		
67	II	73 G–A 83 A–G	25 Ala–Thr	28 Glu–Gly	+	+	(+)	(+)	
68	II	75 A–C			–	–	b		
69	II	76 G–C 81 T–A	26 Gly–Arg	27 Cys–opal	–		a		
70	II	77 G–A	26 Gly–Glu		+++	+++	–		
71	II	77 G–T 89 T–C	26 Gly–Val	30 Ile–Thr	++		–		
72	II	78 A deletion	frame-shift		–	–	a		
73	II	79 T–C	27 Cys–Arg		+	+	–		
74	II	79 T–G	27 Cys–Gly		+	+	–		
75	II	79 T–G 80 G–C	27 Cys–Ala		+	+	+	+	
76	II	79 T–C 87 A–T	27 Cys–Arg		+	+	n.t.		
77	II	81 T–A	27 Cys-opal		–	–	a		
78	II	82 G–C	28 Glu–Gln		–	+	+	++	
80	II	82 G–T	28 Glu–amber		–	–	a		
81	II	84 G–A			++		b		$+^{1,3}$
82	II	84 G–C	28 Glu–Asp		++		+	++	$+^{2}$
83	II	85 C–G	29 Leu–Val		–		+	++	
84	II	85 C–T 87 A–C	29 Leu–Phe		–	–	+	++	
85	II	86 T–A	29 Leu–Gln		++		+	++	$+^{3}$
86	II	87 A–G			++		b		$+^{1,4}$
87		90 T–C			++		b		+
PHU9		no enhancer			–	–			
PHU81		Cin-enhancer			++	++			
PHU78		Cin-recombinase					+	++	
PHU52		Cin-am9					(+)	(+)	

The mutations are arbitrarily numbered; the positions and the nature of nucleotide substitutions and of amino acid changes are given. The corresponding DNA and amino acid sequences are shown in Fig. 3 and in Fig. 5, respectively. Mutations within FIS binding site I or site II are marked.

[X] Enhancer activity (measured as described in Materials and methods) was tabulated as follows. **Plating assay:** – for 0–20% of the activity of pPHU81, + for 21–60%, ++ for 61–150% and +++ for 151–350% of wild type activity; 100% corresponds to 2.5×10^{-2} Cm resistant colonies per ampicillin resistant colony. **Restriction assay:** –for 0–25% of the activity of pPHU81, + for 26–75%, ++ for 76–125% and +++ for 126–143% of wild type activity; 100% corresponds to 35% of the plasmid DNA in the inverted orientation. Negatives of photographs of DNA gels stained with ethidium bromide (as shown in **Fig. 2**) were scanned in order to quantify the amount of DNA in the inverted orientation.

[Y] Recombinase activity (measured as described in Materials and methods) was tabulated as follows. **Plating assay:** –for frequencies below 10^{-3}, (+) for frequencies between 10^{-3} and 10^{-2}, + for 10^{-2}–10^{-1} and ++ for frequencies above 10^{-1}; frequencies are Cm resistant colonies per ampicillin resistant colony; n.t. means 'not tested'. **Restriction assay:** (+) for faintly visible inversion bands, + for visible DNA inversion <50% and ++ for 50% DNA inversion. Nonsense mutations (a) and silent mutations (b) as judged from the sequence analysis were not tested for recombinase activity.

[Z] Mutations which do not affect either *sis*(P1) function or which affect only one of the two *sis*(P1) functions rather weakly are marked by +. These mutations are supposed to be evolutionarily neutral. Some of the naturally occurring *sis* variations were found among these mutations: 1: *min*; 2: *hin*; 3: *gin*; 4: *pin*; 5: P7*cin*. For DNA sequences see **Fig. 3**, for amino acid sequences see **Fig. 5**.

a phosphoserine involving Ser-10 (Reed and Moser, 1984). In the DNA inversion enzymes this is analogous to the conserved Ser at position 9 (see Figure 5). In order to examine whether Ser-9 is needed for recombinase activity we constructed the plasmid pPHU52 where the Ser codon TCA is replaced by the amber nonsense codon TAG. Using different *E.coli* B suppressor mutants, the Ser-9 residue can be replaced either by glutamine (*supE*), Tyr (*supF*), or Leu (*supP*); a *supD* mutant which inserts Ser at the amber codon, served as a control. The result of this analysis (Table II) shows that Ser-9 cannot be replaced by Tyr, which is known to mediate a covalent protein–DNA linkage via phosphotyrosine in the cases of topoisomerases, λ integrase and the yeast FLP enzyme (for a review see Maxwell and Gellert, 1986; Sadowsky, 1986). Also neither Leu nor, most probably, Gln can substitute for Ser-9 indicating the possible specific involvement of this residue in a transient DNA–protein linkage intermediate in the recombination process. However, a relevant distortion of the three-dimensional protein structure by the replacing amino acids (Tyr, Gln or Leu) cannot be excluded. Recently it was reported that the DNA recombinase Gin forms a covalent protein–DNA linkage intermediate via phosphoserine involving Ser-9 (Klippel *et al.*, 1988).

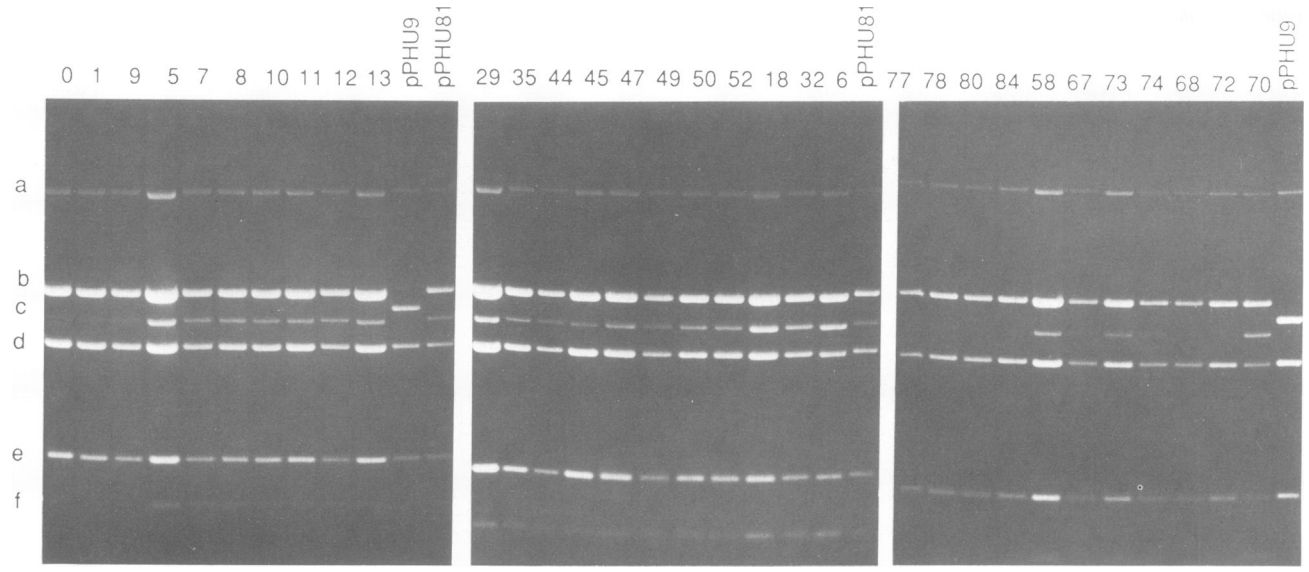

Fig. 2. Restriction assay of enhancer mutants. Derivatives of pPHU9 carrying enhancer mutations were used to transform *E.coli* WA3782(pPHU78) as described in Materials and methods. Plasmid DNA was extracted from transformants and digested with *Eco*RI. The resulting fragments were separated by agarose gel electrophoresis. The mutants tested are indicated above their corresponding lanes. The numbers correspond with those given in Table I. **Band a** originates from the linearized Cin producing plasmid pPHU78. **Bands d** and **e** designate restriction fragments originating from pPHU9 derivatives in the initial orientation, **bands c** and **f** indicate fragments from plasmids in the inverted orientation. **Band b** designates the fragment of the pPHU9 derivatives carrying the inserted enhancer mutation.

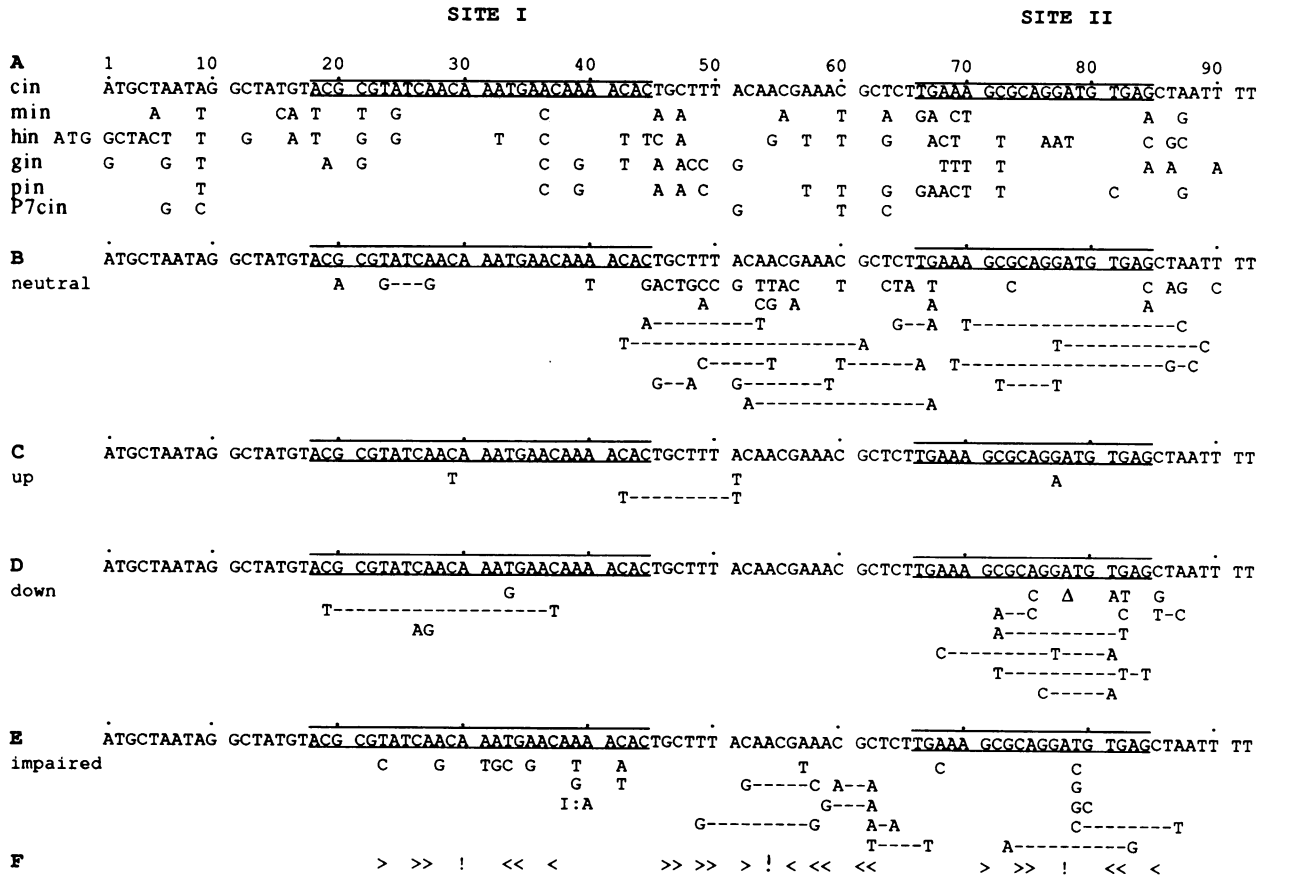

Fig. 3. Location of the enhancer mutants and their classification with respect to their enhancing activity. **A**. Nucleotide sequence of the Cin enhancer of phage P1 (Hiestand-Nauer and Iida, 1983) compared with the corresponding regions of p15B *min* (H.Sandmeier, personal communication), *Salmonella typhimurium hin* (Zieg and Simon, 1980), Mu *gin*, *E. coli pin* (Plasterk *et al.*, 1983) and P7 *cin* (Schmucker *et al.*, 1986). Only bases in positions differing from the Cin enhancer are shown. The positions of neutral Cin enhancer mutations (**B**), of enhancer up mutations (**C**), of enhancer down mutations (**D**) and of mutations impaired in the enhancer activity (**E**) are shown below the corresponding Cin enhancer sequence. In contrast to A every letter in B to E indicates an individual mutant. Double and triple mutations are connected by dashes. The two regions protected by FIS in DNase I experiments (Haffter and Bickel, 1987) are boxed and indicated by site I and site II. The enhancer activities of the mutants, which formed the basis for division into these four classes, are compiled in Table I. **F**. The palindromic structure of the Cin enhancer is shown by arrow heads. The centres of symmetry are indicated by exclamation marks.

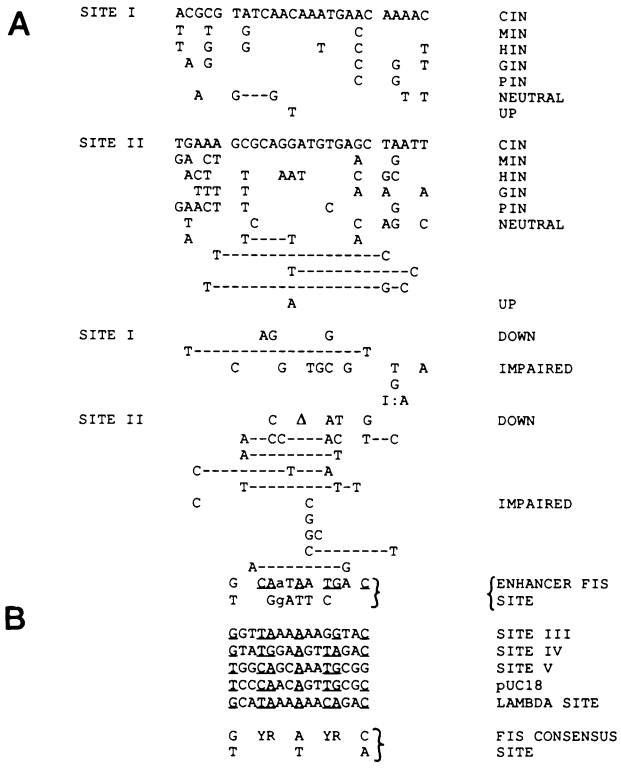

Fig. 4. Compilation of a FIS consensus site. **A.** Comparison of the effect of Cin enhancer mutations with sequences of naturally occurring enhancer FIS sites. The origin of the sequences and the effect of the Cin enhancer mutations are marked on the right side. Underlined sequences in the enhancer FIS site mark positions hit by enhancer down mutations. Lower case letters indicate a position which is not hit by any mutation. **B.** Comparison of the enhancer FIS site with non-enhancer DNA regions protected by FIS in DNase I footprinting experiments. The coordinates for protected sites within the *cin* gene (Haffter and Bickle, 1987) are 131−145 (site III), 178−192 (site IV) and 441−455 (site V); the coordinates are according to Hiestand-Nauer and Iida (1982). The coordinates of the FIS site of pUC18 (Haffter and Bickle, 1987) are 271−257; these coordinates are according to Yanisch-Perron *et al.* (1985). The coordinates of the λ-FIS site are, according to Thompson *et al.* (1987), −59 to −73. Underlined positions indicate matches to the proposed consensus site. Y stands for pyrimidines and R for purines.

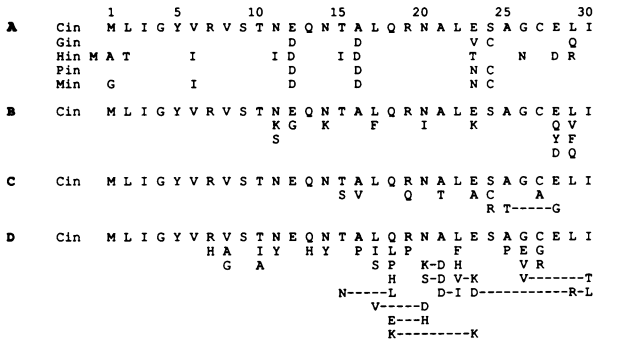

Fig. 5. Location of the recombinase mutants and their classification with respect to their recombinase activity. **A.** Comparison of the N-terminal amino acid sequences of related DNA recombinases. Only amino acids differing from the Cin sequence are shown. The positions of unaffected (**B**), partially active (**C**), and inactive (**D**) mutants are given below the corresponding amino acid sequence of the Cin recombinase. In contrast to A, every letter in B−D means an individual mutant. Double and triple mutations are connected by dashes.

Evolutionary aspects of the sis(P1) mutations

The mutational analysis of the genetic element *sis*(P1) reveals how this sequence can serve two functions. Firstly, this sequence encodes part of the catalytic domain of the Cin recombinase and secondly, it serves as an enhancer for the Cin-mediated DNA recombination reaction. During evolution these two functions have been conserved. The naturally occurring *sis* sequences (see Figure 3) thus represent a part of the allowed variations of this sequence. Taking all described *sis*(P1) mutations into account, we can now address the question of whether variations other than those already found in nature are possible.

From the 26 different evolutionarily neutral mutations (presented in Table I), 15 are already shared by the five natural systems listed in Figure 3. Of the remaining 11 new mutations only two are located within the proposed FIS consensus site (within site I), eight lie between the two FIS binding sites and one mutation resides 3′ of site II. Many of the naturally occurring variations of *sis* are also located within the intervening sequence. This indicates that the intervening sequence between the two FIS binding sites has a higher evolutionary flexibility than the FIS binding sites, which therefore represent highly evolved DNA sequences allowing only a few changes. The two mutations 33 T−C and 33 T−G are both tolerated by the recombinase function, but 33 T−G strongly affects the enhancer function. The codon TCA for Ser-9, which is located at the catalytic centre of the DNA inversion enzyme, also encodes part of the FIS binding site I. Although there are six different codons for Ser, only TCA fulfils both functions of the *sis*(P1) sequence. The two mutations 29 C−T and 77 G−A result in enhancer up mutants, but they both strongly affect the recombinase function. It is interesting to note, that the mutation 77 G−A is found in the *hin* sequence (see Figure 3), in which the recombinase function is preserved by the replacement of the codon GGA (for Gly) by AAT (for Asn). The mutation 51 A−T, however, results in an enhancer up mutation and does not affect the recombinase function. This could mean that we have found by chance a mutation which is better adapted to fulfil the two *sis* functions than the wild type sequence. On the other hand we cannot exclude the possibility that an increased enhancer activity is unfavourable for the organisms carrying such DNA inversion systems.

For the DNA inversion systems the importance of the two *sis* functions seems to be unequal. The isolation of mutant *sis*(P1) elements with higher enhancer activity than wild type but with no or strongly reduced recombinase activity reveals that the enhancing function is subordinate to the coding function.

Materials and methods

Bacterial strains

The following *E.coli* K12 strains were used: WA3782 (Iida *et al.*, 1982) *lac* hsd_K *recA met pin* served as the standard host for growth of the plasmids and for the *in vivo* inversion assay; the strains JM101 (Yanisch-Perron *et al.*, 1985) and UT580 (a gift from Cynthia Lark) served as hosts in the work with the plasmid derivatives of pEMBL and with M13 vectors. The amber mutation of pPHU52 was suppressed by the following *E.coli* B40 suppressor strains: sup^0, *supD*, *supE*, *supF* and *supP* (*sup*-6; Keller and Bickle, 1986).

Construction of plasmids and templates

The Cin overproducing plasmid pPHU132 (Huber *et al.*, 1985a) was altered in pPHU78 by two linker insertions: a *Xba*I linker was inserted into the

Table II. Suppression of *cin*(amber9) by different *E.coli* suppressor strains

	Frequencies of DNA inversion in the presence of the residential plasmid			Efficiency of suppression (d)	Relative activity (e)
	none $\times 10^{-3}$ (a)	pHHL132 $\times 10^{-3}$ (b)	pPHU52 $\times 10^{-3}$ (c)		
E.coli B sup^0	<0.07	130	<0.1	–	–
E.coli B *supD*(ser)	<0.1	40	8.0	32%	63%
E.coli B *supE*(gln)	<0.1	21	<0.09	2%	<20%
E.coli B *supF*(tyr)	<0.5	390	0.11	88%	0.03%
E.coli B *supP*(leu)	<0.4	11	<0.04	89%	< 0.045%

The amino acid inserted by the different suppressor strains is indicated within parenthesis. The frequencies are Cm resistant colonies per ampicillin resistant colony. The suppressor strains harbouring either no *cin* producing plasmid (a), the *cin* producing plasmid pHHL132 (b) or the *cin*(amber9) producing plasmid pPHU52 (c) were transformed with pPHU38 as described in Materials and methods. The efficiency of suppression (d) was calculated by scanning autoradiographs obtained from Western blotting cell extracts from strains indicated in columns b and c with polyclonal anti-Cin antibodies. The relative activity (e) was calculated by dividing the values from column c first by the values from column b and then by the values from column d.

filled in *Hin*fI site of the residual IS*1* fragment upstream of the *cin* gene and a *Sal*I linker was inserted into the filled in *Hin*dIII site at the end of the *cin* gene. The multifunctional DNA inversion tester plasmids pPHU8 and pPHU9 each contain the *Xmn*I/*Nar*I fragment of pPHU62 ligated into the *Xmn*I/*Cla*I site of pEMBL18 or pEMBL19 (Dente *et al.*, 1983), respectively; pPHU62 has the *Sph*I/*Nru*I fragment of pHHL71 (Huber *et al.*, 1985b) deleted. The plasmids pPHU80 and pPHU81 (DNA inversion tester plasmids with a wild type *cin* enhancer) were constructed by the insertion of the *Xba*I/*Sph*I Cin enhancer fragment of pPHU78 into the *Xba*I/*Sph*I sites of pPHU8 and pPHU9, respectively. The same Cin enhancer fragment was also inserted into the corresponding sites of the M13 derivatives mp18 and mp19 (Yanisch-Perron *et al.*, 1985) yielding the templates pPHU82 and pPHU83, respectively. pPHU83 (the ss-DNA template for mutagenesis) was used for the random mutagenesis of the Cin enhancer (Hübner *et al.*, 1989a). After identification of the mutations of the replicative form of the corresponding M13 clones was isolated and cut by *Pvu*II, *Xba*I and *Sph*I. Appropriate fragments were ligated into the *Xba*I/*Sph*I sites of either pPHU9 or pPHU78 yielding the testers for the enhancer and for the recombinase function, respectively. These constructs were checked by restriction analysis taking advantage of the creation of new sites or loss of sites by some mutations; some constructs including all active *cin* mutations, were verified by DNA sequencing. Plasmid pPHU52 (*cin* amber9 mutant) was constructed by replacing the *Mlu*I/*Sph*I fragments of pHHL132 with the corresponding fragment of M13*cin*(amber9), which was obtained by oligo-directed mutagenesis of pPHU83 according to the method of Kunkel (1985) using the mutagenic oligo dACGCGTATAGACAAATGA. The plasmid pPHU38 (DNA inversion tester plasmid with wild type *cin* enhancer) contains the Cin enhancer on a *Sph*I/*Sma*I fragment of pPHU37 inserted into the *Sph*I/*Nru*I sites of pHHL71. pPHU37 has the 143 bp long *Fnu*4HI Cin enhancer fragment of pHHL56 (Huber *et al.*, 1985b) inserted into the *Hin*cII site of pUC19 (Yanisch-Perron *et al.*, 1985). DNA manipulations were done according to standard methods (Maniatis *et al.*, 1982).

DNA sequencing

Some inversion tester derivatives of pPHU9 carrying Cin enhancer mutations were converted to ss-DNA for sequencing by super-infection of JM101 cultures harbouring these plasmids with the helper phage M13K07 (Pharmacia). All active *cin* mutant derivatives which could not be tested by restriction analysis were subcloned into M13mp19 using the restriction sites *Xba*I and *Sph*I. The ss DNAs were then sequenced according to the chain termination method of Sanger *et al.* (1977).

Standard assays for DNA inversion in vivo

Plating assays. To assay the enhancer function, derivatives of pPHU9 carrying mutated enhancer fragments were used to transform WA3782(pPHU78), which over-expresses the Cin recombinase. Since these tester plasmids carry a promoterless *cat* gene for resistance to chloramphenicol (Cm), activated by an operon fusion upon DNA inversion (see Figure 1), the *in vivo* DNA inversion frequency can be scored as the ratio of Cm resistant to ampicillin resistant colonies. Therefore the transformation mixes were incubated at 37°C for 3 h and then plated on LA plates containing either Cm (25 μg/ml) and kanamycin (25 μg/ml) or ampicillin (200 μg/ml) and Km (25 μg/ml); kanamycin (Km) selects for the Cin producing plasmid pPHU78.

To assay for the recombinase function, derivatives of pPHU78 carrying *cin* mutations were used to transform WA3782(pPHU81). In the suppression experiments pPHU38 was used to transform the corresponding strains (see Table II). Since the used *E.coli* B strains are not *lac*I the media were complemented with 1 mM IPTG in order to induce the expression of Cin which is under *lac*UV5 promoter control. The transformation mixes were incubated at 37°C for 3 h and then plated on LA plates containing either Cm (25 μg/ml) Km (25 μg/ml) or ampicillin (200 μg/ml) and Km (25 μg/ml); Km selects for all used Cin producing plasmids.

Restriction assay. In addition to plating, the *in vivo* DNA inversion can be monitored by DNA restriction analysis using an enzyme which cuts asymmetrically within the invertible segment (e.g. *Eco*RI for our tester plasmids, Figures 1 and 2). For that purpose 100 μl aliquots of a 10^{-1} dilution of the above mentioned transformation mixes were used to inoculate a 2.5 ml overnight culture containing 25 μg/ml km and 200 μg/ml ampicillin. Plasmid DNA was extracted after overnight growth and analysed by DNA restriction and gel electrophoresis. All Cin producing plasmids are compatible with the used inversion tester plasmids.

Scanning

Negatives of photographs obtained from DNA gels stained with ethidium bromide were scanned with a CAMAG TLC II Scanner in order to quantify the amount of DNA in the inverted orientation.

Acknowledgements

We are greatly indebted to Shigeru Iida for many helpful discussions and encouragement thoughout this work. We would like to thank Riccardo Cortese (Heidelberg) and colleagues for teaching us the mysteries of oligo-directed mutagenesis during an EMBO practical course at the EMBL (Heidelberg) where *cin*(amber9) was constructed. We wish to thank Pascal Haffter, Heini Sandmeier, Rolf Stalder, Ferenc Olasz and Angelo Guidolin for stimulating discussions and Thomas Bickle for critical reading of the manuscript and discussions. We acknowledge the Genetics Computer Group, University of Wisconsin, who placed the GCG programmes at our disposal and Reinhard Doelz who installed these programmes at the Biocentre in Basel. This work was supported by grant no. 3.624.087 from the Swiss National Science Foundation.

References

Bruist,M.F., Glasgow,A.C., Johnson,R.C. and Simon,M.I. (1987) *Genes Dev.*, **1**, 762–772.

Dente,L., Cesareni,G. and Cortese,R. (1983) *Nucleic Acids Res.*, **11**, 1645–1655.

Guarente,L. (1988) *Cell*, **52**, 303–305.

Haffter,P. and Bickle,T.A. (1987) *J. Mol. Biol.*, **198**, 579–587.

Haffter,P. and Bickle,T.A. (1988) *EMBO J.*, **7**, 3991–3996.

Hiestand-Nauer,R. and Iida,S. (1983) *EMBO J.*, **2**, 1733–1740.

Huber,H.E., Iida,S., Arber,W. and Bickle,T.A. (1985a) *Proc. Natl. Acad. Sci. USA*, **82**, 3776–3780.

Huber,H.E., Iida,S. and Bickle,T.A. (1985b) *Gene*, **34**, 63–72.

Hübner,P., Iida,S. and Arber,W. (1989a) *Gene*, in press.
Hübner,P., Haffter,P., Iida,S. and Arber,W. (1989b) *J. Mol. Biol.*, in press.
Iida,S. (1984) *Virology*, **134**, 421–434.
Iida,S., Schrickel,S. and Arber,W. (1982) *FEMS Microbiol. Lett.*, **15**, 269–273.
Iino,T. and Kutsukake,K. (1980) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 11–16.
Johnson,R.C. and Simon,M.I. (1985) *Cell*, **41**, 781–791.
Johnson,R.C., Glasgow,A.C. and Simon,M.I. (1987) *Nature*, **329**, 462–465.
Johnson,R.C., Ball,C.A., Pfeffer,D. and Simon,M.I. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3484–3488.
Kahmann,R., Rudt,F., Koch,C. and Mertens,G. (1985) *Cell*, **41**, 771–780.
Kamp,D., Chow,L.T., Broker,T.R., Kwoh,D., Zipser,D. and Kahmann,R. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 1159–1167.
Kanaar,R, van de Putte,P. and Cozarelli,N.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 752–756.
Keller,B. and Bickel,T.A. (1986) *Gene*, **49**, 245–251.
Klippel,A., Mertens,G., Patschinsky,T. and Kahmann,R. (1988) *EMBO J.*, **7**, 1229–1237.
Koch,C. and Kahmann,R. (1986) *J. Biol. Chem.*, **261**, 15673–15678.
Koch,C., Vandekerckhove,J. and Kahmann,R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4237–4241.
Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Maxwell,A. and Gellert,M. (1986) *Adv. Protein Chem.*, **38**, 69–107.
Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.
Plasterk,R.H.A. and van de Putte,P. (1984) *Biochim. Biophys. Acta*, **782**, 111–119.
Plasterk,R.H.A., Brikman,A. and van de Putte,P. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5355–5358.
Reed,R.R. and Moser,C.D. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 245–249.
Sadowsky,P. (1986) *J. Bacteriol.*, **165**, 341–347.
Sanger,F., Micklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schmucker,R., Ritthaler,W., Stern,B. and Kamp,D. (1986) *J. Gen. Virol.*, **67**, 1123–1133.
Silverman,M., Zieg,J., Mandel,G. and Simon,M. (1980) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 17–26.
Thompson,J.F., Moitoso de Vargas,L., Koch,C., Kahmann,R. and Landy,A. (1987) *Cell*, **50**, 901–908.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Zieg,J. and Simon,M.I. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4196–4200.

Received on September 6, 1988; revised on November 24, 1988