A bacterial model system for chromosomal targeting

L.-C.Huang, E.A.Wood and Michael M.Cox*

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin – Madison, Madison, WI 53706, USA

Received November 26, 1990; Accepted December 31, 1990

ABSTRACT

A system that permits efficient site-specific chromosomal targeting of foreign DNA on the *Escherichia coli* chromosome has been developed, using the FLP site-specific recombination system derived from the yeast 2μ plasmid. The system demonstrates the feasibility of using site-specific recombination for this purpose, and provides a means to gather information on parameters that may affect chromosomal targeting to guide efforts to establish similar systems in higher eukaryotes. In this model system, the efficiency of integration of foreign DNA is affected by the location of the target site in the chromosome, and the structure of the recombination sites.

INTRODUCTION

Techniques for the introduction or alteration of DNA sequences at specific sites in a chromosome are useful in a wide range of studies related to genetic functions. Targeted chromosomal changes are readily obtained in some lower eukaryotes and prokaryotes. In many higher eukaryotes, techniques have been developed for efficient chromosomal integration of foreign DNA, but the integration is usually not directed to a specific sequence. Random insertions occur at frequencies typically 10^2 or 10^5 higher than targeted integration via homologous recombination (1–3). This problem can be overcome in part with the aid of powerful selection or screening procedures (3–5), but the efficiency of targeted integration of introduced DNA remains low.

Site-specific recombination systems have several advantages that recommend them as a vehicle for site-directed targeting of chromosomes (6,7), and they have been used extensively in chromosome targeting trials in a wide range of higher eukaryotes. Systems particularly suited to these efforts because of their inherent efficiency and simplicity are the FLP system of the yeast 2μ plasmid (7,8), the cre-lox system of bacteriophage P1 (9) and the pSR1 site-specific recombination system (10). Bacteriophage λ and its relatives permit efficient targeting of DNA to bacterial att sites (6), but the number of proteins required makes these systems difficult to use in other organisms. To date, however, no success has been reported in attempts to stably integrate foreign DNA with the FLP, cre-lox, or pSR1 site-specific recombination systems in a chromosome of any organism, although

intrachromosomal recombination has been used successfully in *Drosophila* (FLP system) (11) and yeast (cre-lox and pSR1 systems) (10,12). It is difficult to evaluate the reasons for this failure, since the absence of a positive signal precludes a systematic examination of experimental parameters.

The problems encountered in targeting with the simpler recombination systems led us to set up a model system in *E. coli* where parameters that might affect targeting could be varied systematically and conveniently. The FLP recombination system used in this effort has been extensively studied *in vivo* and *in vitro* (7,8). It requires only one protein, FLP recombinase, and the FLP recombination target (FRT) site (Fig. 1). It has been shown to function in *E. coli* and *Drosophila* cells (11,13) as well as in yeast. The use of this system in chromosomal targeting in *E. coli* is described in this report, and several parameters that affect targeting are evaluated.

MATERIALS AND METHODS

Media

Strains were grown at 37°C in L broth or on agar plates prepared as described (14) and supplemented with antibiotics when appropriate. Papillation screening was performed on MacConkey-lactose agar plates (15). Antibiotic concentrations were: ampicillin, 100 mg/l for cells containing the β -lactamase gene (Amp^r) on a multi-copy plasmid; 20 mg/l for cells containing a single copy of the β -lactamase gene on the chromosome; kanamycin, 40 mg/l; and tetracycline, 25 mg/l. XGal was used at 40 mg/l.

Reagents

Restriction enzymes and T4 DNA ligase were obtained from New England Biolabs or Promega. AMV reverse transcriptase was from Life Sciences. Klenow fragment and linkers were purchased from New England Biolabs. All enzymatic reactions were performed essentially as described by Sambrook et al. (14) or as recommended by the suppliers. Radiolabeled deoxynucleotide triphosphates were obtained from Amersham Corp. Oligonucleotides were prepared at the DNA Synthesis facility in the University of Wisconsin Biotechnology Center. All chemicals were of analytical grade or better and were purchased from common vendors.

* To whom correspondence should be addressed

Plasmids

The plasmids used in the construction of the target strains (pEAW25 and pLH2) differ primarily in minor changes in the FRT site. They were constructed to contain an FRT site, selectable marker, and a promoter-less lacZ gene within the sequences required for Tn5 transposition. The FRT sites used were the minimal wild-type FRT derived from pJFS36 (16) and another FRT that is identical except for a C→T change at position 10 in one of the FLP binding sites (17). In order to construct these plasmids, the EcoRI site on the plasmid pJFS36 was first destroyed and replaced with a SalI site. A plasmid similar to pJFS36, called pLH46T, was constructed with the 46T FRT site (17 and Fig. 1). This plasmid was made by inserting the FRT46T sequence between the BamHI and EcoRI sites on pXF3 (16), and introducing a SalI site right next to the EcoRI site. Then, FRT sites (on small SalI fragments) from the modified pJFS36 and pLH46T were introduced into the unique SalI site within a modified Tn5 transposon contained in the plasmid pRZ620 (15). The resulting plasmids were pEAW25 (FRT36) and pLH2 (FRT46T). The plasmid pEAW25 also has a Kan^r gene inserted in the EcoRI site of pRZ620, in a region that is not transposed.

The plasmid pEAW38 was used as the FLP protein-producing plasmid in both target strains and the non-target control. The plasmid was constructed by first replacing the PvuII site of plasmid pMMC8 (18) with a BglII site. Then, the modified plasmid was digested with BglII and ClaI. The fragment containing the FLP gene (under λ P_R control) and the cI857 gene was isolated. A second plasmid, pACYC184 (19), was modified by replacing the Aval site with a BglII site. The BglII-ClaI fragment from the modified pMMC8 was then ligated to the large BglII-ClaI fragment of the modified pACYC184. The resulting plasmid was designated pEAW38. It has the replication origin and the chloramphenicol resistance gene from pACYC184, and cI857 and FLP genes from pMMC8. All the plasmids were propagated in *E. coli* strain RZ211 (F⁻ *ara* Δ (*lac pro*)*thi srl* *recA56*).

Construction of target strains

The *E. coli* strain RZ211 (F⁻ *ara* Δ (*lac pro*)*thi srl* *recA56*) was transformed with plasmid pEAW25 or pLH2, and tet^r transformants were streaked out on lac MacConkey plates. The appearance of red papillae indicated a transposition event had occurred that inserted modified Tn5 in an actively transcribed region producing a Lac⁺ phenotype. These papillae were picked as described by Krebs and Reznikoff (15). The sites were then transduced into CSH26 (F⁻ *ara* Δ (*lac pro*)*thi*) to create strains with chromosomal target FRTs that did not have the pEAW25 or pLH2 plasmids. For both FRT sites, six target strains were selected at random, each containing a single target FRT in its chromosome, for use in these experiments, and denoted CSH26/3601-06 and CSH26/4601-06. All these target strains were Rec⁺.

Construction of bacteriophage λ vector

The bacteriophage λ vector λ FRT36 and λ FRT46 were constructed by replacing the DNA sequence between the unique NotI sites of IS50R and IS50L in λ NK467 (b221 cI857 *rex::Tn5* Oam29 Pam80) with the FRT sites, lacZ gene and Amp^R gene markers. All the phages were grown in the suppressor strain LE392 (*supE44 supF58 hsdR514 galk2 galT22 metB1 trpR55 lacYI*) (14). Other general procedures including phage DNA purification, *in vitro* packaging, infection and plating were

performed essentially as described by Davis et al. (Advanced Bacterial Genetics, CSH) with small modifications.

Southern analysis

The genomic DNA was isolated as previously described (20). The DNA was cut with ScaI, electrophoresed and blotted onto nitrocellulose paper by capillary transfer according to Sambrook et al. (14). The linearized Gori1 plasmid and the BsmI-cut Gori1 plasmid were used as size markers. The blot was hybridized with the chromosomal FRT probe and the size marker probe at the same time. The former was isolated from XbaI- and BglII-digested pRZ620 (15), and the latter, the whole Gori1 plasmid. Both probes were labeled by nick translation (14).

Targeting trials

Generally, overnight cultures of the target and non-target (no chromosomal FRT) *E. coli* strains containing pEAW38 were diluted 100 fold and grown at 37°C or 42°C in L broth to induce FLP protein production. After serially diluted phage lysates were added and incubation was continued for 20 min at 37°C or 30°C, the cells were plated out on plates containing 20 mg/l Ampicillin and incubated at 30°C for 14–18 h. The resulting Amp^r colonies were counted.

In vitro FLP reaction assay

In vitro FLP reaction assays were performed to verify the presence of active FRT sites on the bacteriophage vector and the plasmids used for the construction of target strains. The assays were performed in 25 mM TAPS {3-[N-tris(hydroxymethyl)-methylamino]-propane sulfonic acid}, 1 mM EDTA, 0.2 M NaCl, 10% (w/v) PEG (M_r 8,000), 20% glycerol, 2.5 mg of BSA per 1, 0.2 μ g/ μ l DNA substrate, and 140 nM FLP protein. This reaction has been described previously (21).

Generation of recombination clones

The recombination clones were generated by standard protocols as described previously (17,22).

Purification of plasmid DNA

Plasmid DNA was purified using the rapid boiling method of Holmes et al. (23) or banded from cesium chloride-ethidium bromide gradient as described by Garger et al. (24).

DNA sequencing

Sequencing of the FRT sites was performed as described (17) using the dideoxynucleotide primer extension method of Sanger et al. (25) with some modifications (17).

RESULTS

The site-directed integration system consists of three elements: (1) *E. coli* target strains, each with a single FRT site located on the chromosome, (2) a plasmid expressing FLP protein, and (3) a delivery vector containing a drug resistance gene marker as well as an FRT site compatible with the FRT site on the chromosome (Fig. 1). Each target strain was constructed by introducing a single FRT site into the *E. coli* chromosome by Tn5 transposition. Southern analysis showed that the site was integrated at different locations in the various target strains used here. These sites have not yet been mapped. FLP recombinase is expressed by the plasmid pEAW38, and expression is controlled by the heat-inactivated repressor cI857. The integration

Fig. 1. Components of the chromosomal targeting system. Target strains (panel A) were constructed by introducing FRT sites into the *E. coli* chromosome by Tn5 transposition. The Tn5/FRT-containing plasmids used for target construction were pLH2 (FRT46T, shown) and pEAW25 (FRT36). The structure of the chromosomal targets is shown at the bottom of panel A, with wavy lines denoting chromosomal DNA. Six target strains with each FRT site were selected at random for use in these experiments, and denoted CSH26/3601-06 and CSH26/4601-06. The FLP expression vector pEAW38 is shown in panel B. The bacteriophage λ FRT donor vectors λ FRT36 and λ FRT46A are shown in panel C. N and S denote cleavage sites for *NotI* and *SalI*, respectively. The location of the point mutation in FRT46T (pLH2) and FRT46A (λ FRT46A) is indicated by an asterisk in the respective sequence enlargements. Recombination between these two FRT sites generates one product FRT that contains both mutations and another FRT that contains neither.

vectors were constructed from λ NK467 (b221, cI857, rex::Tn5, Oam29, Pam80) (26), which can replicate in amber suppressor strains but not in our target strains. The vector also has a deletion in the *att P* site, so that it cannot integrate into the chromosome by means of the λ Int system. These phage vectors are therefore unstable and lost unless the phage DNA is integrated into the chromosome. An ampicillin resistance marker is incorporated to permit selection of integrants. The expected integration reaction is illustrated in Fig. 2.

Two different FRT sites were used in this study. FRT36 is the minimum functional wild-type FRT sequence from pJFS36 (16) (Fig. 1). FRT46 has a single point mutation within one of two FLP protein binding sequences (Fig. 1). This mutation has only a moderate effect on recombination efficiency when present in one FLP protein binding site, but nearly inactivates an FRT site if it is present in both FLP binding sites (17,21). In the FRT46 sites of the donor and target, the mutation is located in opposing FLP binding sites. One of the two FRT sites flanking an integrant

Fig. 2. The FLP-mediated chromosomal integration reaction. X, S, and B denote cleavage sites for *XbaI*, *ScaI*, and *BglII*, respectively. The unique 6.8 kb *ScaI* fragment present after integration is indicated, along with the DNA fragment used as a probe in Fig. 3.

resulting from a reaction involving these FRT46 sites will acquire both mutations and be inactivated.

The experiments summarized in Table 1 and Fig. 3 were designed both to test the system and to provide information about the effects of the chromosomal location of the FRT target and possible integrant instability due to FLP-mediated reversal of the reaction (excision).

In a typical experiment, several different target *E. coli* cells containing pEAW38 were incubated with the donor λ bacteriophage. A control strain that contained the appropriate FLP-producing plasmid but lacked a chromosomal FRT site (nontarget strains) was also infected. Two strategies were used to try to prevent excision of the integrated donor DNA. First, FLP protein was expressed and then repressed shortly after the introduction of the bacteriophage. Second, the FRT46 sites were tested to determine if they improve integrant stability. The target and nontarget *E. coli* strains containing pEAW38 were grown at 37°C to induce FLP protein production. After serially diluted phage lysates were added and incubation was continued for 20 minutes at 37°C, the cells were plated out on selective plates and incubated at 30°C (repressing FLP) for 14 to 18 hours. Typical integration frequencies are given in Table 1. The highest targeting efficiencies, up to 10^{-1} /input phage, was observed with FRT36. The sizes of the resulting *Amp^r* colonies were uniform. The integration frequencies were optimal after the culture reached late log phase (4 hours after growth is initiated by diluting the overnight culture) with a host cell density higher than 10^8 cells per ml.

Efficient integration was clearly dependent on the expression of FLP protein. When target strains were grown at 30°C, thereby

TABLE 1. Integration frequencies per input phage ($\times 10^{-2}$)

Strains	λ FRT36	λ FRT46
CSH26	0.003	0.002
CSH26/3601	3.1	1.1
CSH26/3602	6.9	0.56
CSH26/3603	5.2	0.99
CSH26/3604	1.4	0.36
CSH26/3605	1.0	0.33
CSH26/3606	2.1	0.65
CSH26/4601	0.54	0.038
CSH26/4602	0.43	0.038
CSH26/4603	0.82	0.096
CSH26/4604	0.48	0.086
CSH26/4605	0.34	0.034
CSH26/4606	0.52	0.032

All strains contained pEAW38 and were *recA⁺*. Integration frequency = no. of *Amp^r* colonies/input phage. This experiment was repeated four times, with observed integration frequencies varying by up to 2 fold up or down relative to the values shown here. The trends evident here, however, were highly reproducible. For example, in four trials of the CSH26/3601 $\times$ λ FRT36 experiment, integration frequencies ranged from $1.7-5.0 \times 10^{-2}$ /input phage. The integration frequency obtained for this combination, however, was always about 3 fold greater than for CSH26/3605 $\times$ λ FRT36 in a given experiment.

minimizing FLP protein production, the integration frequencies were never more than 10^{-4} per input phage at any point in the growth curve (not shown). Targeted integration was also dependent on the presence of an FRT site in the chromosome. Integration efficiency in the nontarget strain (CSH26) was 1,000 fold lower than in the target strains (Table 1). The integration efficiency reported in Table 1 was essentially constant for the

Fig. 3. Southern analysis of genomic DNA isolated from target strains before and after targeting. Autoradiogram of a Southern blot hybridized with the probe described in Fig. 2. M: size markers. The size of each fragment is indicated at the left. C: negative control, *ScaI*-digested genomic DNA isolated from nontarget parental strain CSH26. For each target strain shown there are two lanes, representing *ScaI*-digested DNA isolated from target cells before (left) and after (right) λ FRT integration. Before targeting, the two hybridized fragments are unique for each target strain, indicating that the FRT sites were inserted at different positions in the *E. coli* chromosome in different target strains. After targeting, all the target strains exhibit the 6.8 kb fragment characteristic of FLP site-specific integration (see Fig. 2). At right, the six lanes labeled 3601 represent a repeat of the 3601 before and after blots described above, except that in this case five separately isolated colonies resulting from targeting are compared side by side. Linearized *GorI* plasmid (26) and *BamI*-cut *GorI* plasmid were used as size markers (M).

phage vector to target cell ratio (m.o.i.) in the range of 10^{-5} to 10. The integrants were also very stable. At 30°C, there was no detectable loss of the integrated phage vector after 20 generations of growth in nonselective media. This stability was a function of repression of FLP protein production and was not detectably improved by the FRT46 mutation. When phage vectors bearing FRT36 or FRT46A sites were introduced into target strains containing pEAW38 and chromosomal FRT36 or FRT46T sites, the integration frequencies obtained with different phage-target strain combinations had the order: $\text{FRT36} \times \text{FRT36} > \text{FRT36} \times \text{FRT46A} = \text{FRT36} \times \text{FRT46T} > \text{FRT46A} \times \text{FRT46T}$. In this experiment, the $\text{FRT46A} \times \text{FRT46T}$ cross, designed to prevent excision and thereby stabilize integrants, actually decreased recombination efficiency by a factor of 100. The results suggest that the FLP protein is relatively unstable in *E. coli*, with levels decaying rapidly after plating the cells at 30°C. The shift down of FLP protein provides the best strategy for integration in this system. When FLP protein was regulated by the *Tac* promoter in a construct otherwise similar to pEAW38, the relatively leaky expression of FLP at all times prevented measurable targeting (data not shown). Further experiments are needed to determine if the $\text{FRT46} \times \text{FRT46}$ strategy improves integrated stability when FLP protein is expressed continuously.

The site-specificity of targeting was verified by Southern Analysis (Fig. 3). After *ScaI* digestion of the genomic DNA and hybridization with probes specific for the vector and the chromosomal FRT sites, all the transformants exhibited the 6.8 kb fragment expected for FLP-mediated integration in this system (Figs. 2 and 3). We also digested the genomic DNA with *XbaI* restriction enzyme, which cuts the spacer region of the FRT sites. In this case, the restriction pattern stayed the same after integration, confirming that integration occurred at or very near the expected point (data not shown).

Integration efficiency exhibited a modest dependence on the position of the FRT site in the chromosome. The absolute integration efficiencies observed in different experiments varied

up to 2 fold in experiments carried out weeks apart, but the trends reported in Table 1 were highly reproducible. In targeting experiments with target cells at different growth stages or with phage lysates of varied titers, the integration frequencies of some target strains (for example, CSH26/3604) were always lower than those of others. It is not clear whether this reflects differences in chromosome structure, levels of transcription, or some other parameter at the FRT sites of the various target strains. We also note that target strains used here all have an FRT located in an actively transcribed segment of the chromosome. The numbers in Table 1 may not represent the maximum effects of chromosomal location.

DISCUSSION

We describe a system for efficient site-directed integration of foreign DNA into the *E. coli* chromosome. The system has been set up to study the effects of a variety of experimental parameters that may affect chromosomal targeting as a model to guide efforts in higher eukaryotes. In addition, this system may prove useful in studies of *E. coli* genome structure and function.

Targeted integration requires a suitable FRT site on the *E. coli* chromosome, a plasmid with the FLP gene expressed by means of a tightly regulated promoter, and a modified bacteriophage vector as a source of donor DNA. In the best constructs, the efficiency approaches one targeted transformant per 10 input donor phage. The results establish the feasibility of using the FLP recombination system for chromosomal targeting, at least in bacteria.

We have observed a modest effect of the chromosomal location of the target FRT on targeting efficiency. For some target pairs (e.g., 3,602 vs. 3,605) reproducible differences on the order of 2–6 fold were observed reproducibly over many trials (Table 1). We do not know the cause of these differences. Since this is a small sample and the targets were not randomly inserted into the chromosome (the construction of the target strains relied on a protocol that yielded only those targets that were inserted in actively transcribed genes), we are also uncertain if this is the maximum extent of position effects that might be observed on the *E. coli* chromosome.

A major potential problem of chromosomal targeting by this method is integrant instability. The integration reaction is intermolecular, and the intramolecular deletion of the inserted DNA might be expected to be favored. We tried two approaches to solving this problem: (a) limiting expression of FLP recombinase to a few minutes after addition of the donor DNA and (b) using altered recombination sites that leave one inactive recombination site flanking the new integrant. In this case the former strategy worked better. The mutant FRT sites in all cases decreased targeting efficiency. This result depended upon expression of FLP protein from a tightly regulated promoter. When a construct was used in which FLP was expressed from the relatively leaky *tac* promoter, stable integration of the donor DNA was rarely observed (data not shown). Since integration was highly stable when λ *P_R*-mediated FLP expression was repressed, the results strongly suggest that FLP protein is unstable and rapidly degraded in *E. coli*. For the purpose of targeting, this may be a desirable feature of the FLP system.

These results suggest that chromosomal targeting mediated by site-specific recombination has potential as a general method. The simple site-specific recombination systems such as FLP should permit both efficient integration of donor DNA and efficient

recovery of the DNA on demand. The model system described here is currently being altered to permit us to address a variety of additional questions and to guide the design of similar efforts in eukaryotic cells.

ACKNOWLEDGMENTS

This work was supported by March of Dimes Basic Research Grant # 1-1055 and by N.I.H. grant GM37835.

REFERENCES

1. Lin, F.-L., Sperle, K., and Sternberg, N. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1391–1395.
2. Zheng, H. and Wilson, J.H. (1990) *Nature*, **344**, 170–173.
3. Capecchi, M.R. (1989) *Science*, **244**, 1288–1292.
4. Koller, B.H. and Smithies, H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 8932–8935.
5. Jasin, M., Elledge, S.J., Davis, R.W., and Berg, P. (1990) *Genes and Dev.*, **4**, 157–166.
6. Craig, N.L. (1988) *Annu. Rev. Genet.*, **22**, 77–105.
7. Sadowski, P.D. (1987) In McMacken, R. and Kelly, T. (eds), *DNA Replication and Recombination*. Alan R. Liss, New York, pp. 691–701.
8. Cox, M.M. (1989) In Berg, D.E. and Howe, M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, D.C., pp. 661–670.
9. Hoess, R. and Abremski, K. (1985) *J. Mol. Biol.*, **181**, 351–362.
10. Matsuzaki, H., Araki, H., and Oshima, Y. (1988) *Mol. Cell. Biol.*, **8**, 955–962.
11. Golic, K.G. and Lindquist, S. (1989) *Cell*, **59**, 499–509.
12. Sauer, B. (1987) *Mol. Cell. Biol.*, **7**, 2087–2096.
13. Cox, M.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4223–4227.
14. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
15. Krebs, M.P. and Reznikoff, W.S. (1988) *Gene*, **63**, 277–285.
16. Senecoff, J.F., Bruckner, R.C., and Cox, M.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7270–7274.
17. Senecoff, J.F., Rossmeissl, P.J., and Cox, M.M. (1988) *J. Mol. Biol.*, **201**, 405–421.
18. Cox, M.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4223–4227.
19. Chang, A.C.Y. and Cohen, S.N. (1978) *J. Bacteriol.*, **134**, 1411–1415.
20. Wilson, K. (1989) In Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds), *Current Protocols in Molecular Biology*. Wiley Interscience, New York, pp. 2.4.1–2.4.2.
21. Gates, C.A. and Cox, M.M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4628–4632.
22. Umlauf, S.W. and Cox, M.M. (1988) *EMBO J.*, **7**, 1845–1852.
23. Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.*, **114**, 193–197.
24. Garger, S.J., Griffith, O.M., and Grill, L.K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 835–842.
25. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
26. Kaguni, J. and Ray, D.S. (1979) *J. Mol. Biol.*, **135**, 863–878.