In vivo excision and amplification of large segments of the *Escherichia coli* genome

György Pósfai, Michael Koob, Zdenka Hradečná, Noaman Hasan, Marcin Filutowicz¹ and Wacław Szybalski*

McArdle Laboratory for Cancer Research and ¹Department of Bacteriology, University of Wisconsin, Madison, WI 53706, USA

Received January 13, 1994; Revised and Accepted May 10, 1994

EMBL accession nos U08460 and U08461

ABSTRACT

***In vivo* excision and amplification of large segments of a genome offer an alternative to heterologous DNA cloning. By obtaining predetermined fragments of the chromosome directly from the original organism, the problems of clone stability and clone identification are alleviated. This approach involves the insertion of two recognition sequences for a site-specific recombinase into the genome at predetermined sites, 50–100 kb apart. The integration of these sequences, together with a conditional replication origin (*ori*), is targeted by homologous recombination. The strain carrying the insertions is stably maintained until, upon induction of specifically engineered genes, the host cell expresses the site-specific recombinase and an *ori*-specific replication protein. The recombinase then excises and circularizes the genomic segment flanked by the two insertions. This excised DNA, which contains *ori*, is amplified with the aid of the replication protein and can be isolated as a large plasmid. The feasibility of such an approach is demonstrated here for *E. coli*. Using the yeast FLP/FRT site-specific recombination system and the π/γ -*ori* replication initiation of plasmid R6K, we have devised a procedure that should allow the isolation of virtually any segment of the *E. coli* genome. This was shown by excising, amplifying and isolating the 51-kb *lacZ*–*phoB* and the 110-kb *dapX*–*dsdC* region of the *E. coli* MG1655 genome.**

INTRODUCTION

Molecular cloning of DNA segments is a powerful tool in mapping, manipulating and sequencing genes. Cloning procedures, however, can be cumbersome and have their intrinsic limitations. The cloned DNA is maintained in a foreign host and copied by a heterologous system. The alien genetic material and/or its products can be toxic for the cell or can carry modifications which are discriminated against. As a result, selective pressure often favors unfaithfully copied, deleted or rearranged clones. For some purposes, like long-range

sequencing, the large size of the cloned molecules is essential. However, several cloning systems put a limit on the size of the clone, or make the handling of large molecules difficult. Constructing a genomic library is sometimes a routine task, but ordering the clones into a continuous set for mapping or long-range sequencing purposes may prove to be a laborious and difficult process.

Directed *in vivo* excision and amplification might offer an alternative to heterologous cloning. The prototype of our approach is the excision and replication of prophage λ and other proviruses. The method involves *in vivo* genetic manipulation of the studied organism, and results in obtaining the desired DNA segment directly from the organism in an amplified, circular form. The process aims to overcome the problems of copying fidelity, size limitation and clone ordering.

The general idea behind the scheme (1,2,3) is that certain site-specific recombinases can excise and circularize stretches of chromosomal DNA flanked by their recombination target sites (4,5,6). If the excised, circular DNA possesses a conditional replication origin (*ori*), it can be replicated in the cell by providing the specific initiation protein that activates this *ori*. The amplified DNA is then purified directly from the cell, taking advantage of its circular nature.

In order to obtain specific, predetermined DNA segments, the excision sites for the recombinase, together with the conditional *ori*, must be inserted into the chromosome at the desired locations. This can be achieved with the aid of homologous recombination, provided that sufficient sequence information is available for the targeted sites.

For the cells to be stably maintained, the specifically engineered genes encoding the site-specific recombinase and the *ori*-specific protein must be tightly controlled so that the potentially lethal excision and amplification does not take place until their induction.

We report here the development of an excision–amplification system for obtaining large genomic segments of *E. coli*. The scheme utilizes the FLP/FRT recombination system, chosen for its wide applicability, and the well-characterized π/γ -*ori* replication system, for which hyperactive (copy-up) variants of the initiation protein exist.

*To whom correspondence should be addressed

The site-specific recombinase FLP (7) is a yeast protein that was shown to work in a variety of prokaryotic and eukaryotic hosts, including *E. coli*, insects and mammals (8,9). *FRT*, the binding site for the recombinase on the DNA, minimally consists of two 13-bp inverted repeat sequences with an 8-bp asymmetric spacer in between (10). FLP catalyzes both intra- and intermolecular recombination between two *FRT* sites. If the *FRT* sites are on the same molecule and are divergently oriented, the reaction results in inversion of the DNA segment between the *FRT*s. On the other hand, recombination between directly oriented sites results in excision and circularization of the DNA segment flanked by the *FRT*s (11).

The π/γ -*ori* replication system is derived from the *E. coli* plasmid R6K (12,13). The γ -*ori*-directed replication is strictly dependent on the initiator protein π , encoded by the *pir* gene (14). Once initiated, replication proceeds by using the host replication machinery. Expression of π is autoregulated at the transcriptional level. In addition to the initiator function, π shows a direct negative effect on replication. This effect becomes considerably reduced in the case of the hyperactive π mutants (15). Such copy-up mutants are capable of maintaining small γ -*ori* plasmids at copy numbers up to 250 (16). In this work we have used the *pir200* gene that codes for a high-copy-number producing mutant π (17). Any further reference to π indicates its *pir200* mutant.

We demonstrate the feasibility of our excision–amplification system by isolating 51-kb and 110-kb segments of the *E. coli* K-12 genome.

MATERIALS AND METHODS

Plasmids and strains

The various segments of plasmids pATRFP, pMAKK76 and pMAKC76 were obtained from the following sources: *FLP* from pMMC3 (18); *pir200* from plasmid pMF261 (19); multiple cloning site (MCS) with the T7 and SP6 promoters from pGEM3 (Promega, Madison, WI); cartridges for γ -*ori*, *tetA*, *tetR*, the promoter of *tetR* (*P_{tet}*), pBR322 *ori*, Km- and Cm-resistance genes (Km^R, Cm^R) and λ *att* were from the plasmid series pANTS, including pTANTS, pKANTS γ and pCANTS γ (20). *FRT* was a synthetic 48-bp fragment corresponding to the minimal *FRT* sequence (21). The *E. coli* K-12 strain MG1655 (22) was obtained from F.R. Blattner. Plasmids with γ -*ori* were maintained in the SY327(λ *pir*) lysogen (23).

Enzymes, antibiotics and nucleotides

Restriction enzymes and *E. coli* DNA polymerase I Klenow fragment were purchased from New England Biolabs (Beverly, MA). T4 DNA ligase was obtained from Epicentre Technologies (Madison, WI) and *Taq* DNA polymerase was from Promega (Madison, WI). The antibiotics Ap, Km, Cm and Tc were used at concentrations of 50 μ g/ml, 25 μ g/ml, 17 μ g/ml and 3 μ g/ml, respectively, unless otherwise indicated. Heat-inactivated chlortetracycline (cTc) was used as an inducer of transcription from *P_{tet}* at a final concentration of 25 μ g/ml, or as indicated. LB medium supplemented with cTc (LBcTc) was autoclaved before use (24). ATP and dNTPs were purchased from Pharmacia (Piscataway, NJ).

Recombinant DNA techniques and P1 transduction

Construction of plasmids, genetic transformation, gel electrophoresis of DNA, alkaline lysis of cells and plasmid

preparation were all performed according to described techniques (25). For the detection of large circular DNA molecules, 0.6–0.7% agarose gels were run at 3 V/cm for 4–5 h, followed by staining with ethidium bromide. Small plasmids were purified by the Magic DNA Purification System (Promega). DNA fragments were isolated from agarose gels by the DE81 paper method (26) and purified on Magic Purification Resin (Promega). P1 transduction was performed according to published protocols (27).

Polymerase chain reaction (PCR)

For PCR amplification of *phoB*, *dsdC* and *dapX*, 28-mer primers were synthesized (UW Biotech Center, Madison, WI). The primers consisted of a 20-bp homology region and a 8-bp 5' extension which included a restriction site. Parameters of the PCR cycles were the following: 1-min denaturation at 94°C, 1-min annealing at 55°C, 2-min extension at 72°C, 25 cycles. *Taq* DNA polymerase was added to the PCR mixture after an initial 1-min incubation at 94°C. PCR products were purified on Magic Purification Resin (Promega), cleaved with the appropriate restriction enzymes and cloned into pMAKK76/pMAKC76.

Preparation of microbeads and pulsed-field gel electrophoresis (PFGE)

Embedding of cells in microbeads, DNA preparation and PFGE were performed according to published protocols (28). Microbeads were prepared from Incert agarose (FMC) while PFGE was carried out in 1% High-Strength Ultra-Pure agarose (Bio-Rad, Melville Centre, NY). Parameters for PFGE were the following: 5 V/cm field strength, 120° field angle, 10 to 20 s ramped switching time, 18 h running time.

RESULTS

A scheme for isolation of large chromosomal segments

To isolate large segments of the *E. coli* genome, we devised a scheme that utilizes the FLP/*FRT* recombination system and π/γ -*ori* replication initiation.

First, the host cell must possess a source for conditionally expressed FLP and π . The *FLP* and *pir200* genes can reside on a plasmid or, more conveniently, can be inserted into the chromosome under the control of an inducible (*tet*) promoter. Once the genes are on the chromosome, the resulting tFp (*tet*-*FLP-pir*) strain can be used for further manipulations.

The next steps are diagramed in Fig. 1. A PCR-amplified segment of a selected chromosomal locus, for which the sequence and orientation are known, is cloned into an insertion plasmid that carries the *FRT* site, γ -*ori* and an antibiotic-resistance marker. Since the insertion plasmid requires π for replication, it must be maintained in a π -producing (Pir⁺) host (23). This plasmid construct is then transformed into the tFp host where the PCR-amplified DNA segment provides homology in the RecA-mediated homologous recombination event between the plasmid and the chromosome. Since the plasmid does not replicate in the tFp host under non-inducing (Pir[−]) conditions, cells that acquired a chromosomally integrated plasmid can be selected by their ability to grow in the presence of the proper antibiotic. In addition, the efficiency of the recombination event can be enhanced by temporarily inducing the *pir* gene in the tFp host, which results in temporary replication of the plasmid, increasing the chances of recombination. As a result of the integration, *FRT* and γ -*ori*, together with an antibiotic-resistance marker are placed

Figure 1. An approach for excision and amplification of large pieces of the chromosome. (A) *FRT*, γ -ori and Cm^R are inserted into the chromosome by homologous recombination between a selected chromosomal site (S1) and a plasmid carrying a PCR-generated copy of S1 (S1'). A similar insertion is made at a second site (S2) using another antibiotic-resistance gene (Km^R). T7 and SP6 indicate specific phage promoters. (B) The two insertions are transferred to the same genome by P1 transduction. (C) The genes encoding FLP and π , carried on the chromosome of the host cell (see Fig.2B), are induced. FLP excises and circularizes the chromosomal segment between the two *FRT* sites and π initiates replication at γ -ori. The amplified circular DNA is isolated as a plasmid.

on the chromosome at a predetermined site. The direction of *FRT* on the chromosome is determined by the orientation of the PCR-amplified sequence in the insertion plasmid. Independently, a similar insertion is made at another site using a different antibiotic-resistance gene. The distance between the first and second sites can conveniently be 50–100 kb (Fig.1A).

One of the inserted chromosomal sites is then transferred by P1 transduction into the cell carrying the other insertion. Cells with both inserts are selected by their resistance against both antibiotics (Fig.1B).

The cells are grown to mid-log phase, then FLP and π production is induced. FLP excises and circularizes the chromosomal segment residing between the two insertions by recombining the *FRT* sites (Fig.1C). Protein π initiates replication at γ -ori, which results in a several-fold amplification of the

excised circle. After an appropriate period of induction, the large circular DNA is isolated by conventional plasmid preparation methods.

Construction of the tFp host cells carrying the *tetR*-controlled *FLP* and *pir* genes

The MG1655 strain, representing the wild-type *E. coli* K-12 (22), was modified to permit conditional expression of *FLP* and *pir*. Plasmid pATRFp, carrying the *FLP* and *pir* genes under the control of the Tet repressor-*tet* promoter of Tn10, was constructed as described in Fig.2A. The plasmid was cleaved by *NotI* to remove the pBR322 *ori*. The larger fragment was isolated from an agarose gel, self-ligated (circularized) and transformed into MG1655. This non-replicating circle was recombined into the *latt* site of the chromosome (Fig.2B) (29) by providing *Int* from

Figure 2. Construct providing the inducible FLP and π functions, and its integration into the MG1655 genome. (A) Structure of plasmid pATRFP. Selected restriction sites used in the construction process are shown with the coordinates of the 5' base in the recognition sequence. Numbering starts at the *BamHI* site. Restriction sites in brackets were blunted by DNA polymerase I Klenow fragment before ligation, and thus cannot be recut. The *BamHI*(1)–*Clal*(3286) fragment was from pTANTS, with an additional *KpnI* linker fused to the *Clal* site in an intermediate construction step. The *[NotI]*–*Clal*(4816) segment was excised from pANTS. The *Clal*(4816)–*EcoRI*(5593) fragment was obtained from Tn10 by PCR, using primers with the indicated restriction sites. The *[SphI]*–*[FspI*(6971)] fragment was from pMMC3, with an *EcoRI* linker fused to the blunted *SphI* site. Segment *[EcoRV]*–*[NheI*(7015)] originates from pBR322. The *[BsmAI]*–*BamHI*(1) fragment was excised from pMF261. The map is not drawn to scale. (B) Structure of the insert at the λ *att* site of the modified host strain MG1655 (MGtFp). *Ap^R* and *tetA* render the cell resistant to Ap and Tc, respectively. *FLP* and *pir200* are transcribed from the *P_{tet}* promoter that is controlled by the product of *tetR*. Construction of the *P_{tet}*–*FLP*–*pir200* operon resulted in modified 5'-ends of the genes. While *FLP* has two extra amino acids in the second and third positions (underlined by dashed lines), translation of *pir200* starts most probably at amino acid position 6 or 7 (37), because it was stripped of its upstream sequences. Ribosome-binding sites are overlined and start codons are boxed. The insert carries a single *NotI* restriction site.

the helper plasmid pInt-ts (20). Since the helper plasmid had a temperature-sensitive *ori*, tFp cells were easily cured of this plasmid by growth at 42°C (20).

The modified MG1655 tFp strain (designated as MGtFp) is resistant to 50 μ g Ap/ml and 3 μ g Tc/ml. *FLP* and *pir200* are transcribed from *P_{tet}* under the control of TetR and can be induced with 1–3 μ g/ml Tc. However, since the addition of Tc delays the growth of the cells, we prefer to use heat-inactivated cTc (1–50 μ g/ml) that showed no effect on cell growth.

Expression of *FLP* and π is tightly controlled by TetR. In a control experiment, a pBR322-derived plasmid (not shown) with two, directly oriented *FRT* sites showed no *FLP*-mediated excision of the *FRT*-flanked fragment when transformed into MGtFp. However, following induction, more than 90% of the plasmids lost the fragment during a 2.5-h incubation period. Similarly, the *pir* gene must be tightly repressed since plasmid pMAKK76 (Fig.3) with π -dependent replication cannot be established in uninduced MGtFp. However, when MGtFp is grown in the presence of the inducer, pMAKK76 is maintained. The amount of plasmid isolated from the Pir⁺ cells nearly equals the amount of pUC19 isolated from MG1655. The cTc

Figure 3. Structure of the insertion plasmids pMAKK76 and pMAKC76. The two plasmids differ only in the *Clal*–*NotI* segment that carries the antibiotic-resistance gene (*Km^R* or *Cm^R*). The orientation of the *FRT* sequence is indicated by an arrow. The multiple cloning site (MCS) is flanked by the phage promoters T7 and SP6. Selected restriction sites are shown with the coordinates of the 5' base in the recognition sequence. Numbering starts at the *Clal* site. Unique restriction sites are shown in bold letters. Restriction sites indicated in brackets were erased in the construction process by ligating two different blunt ends. The *Clal*(1)–*EcoRV*(68)] segment containing the minimal *FRT* site was a synthetic fragment. The *[PvuII]*–*NheI*(412) fragment was obtained from pGEM3, with an additional *NcoI* linker ligated to the *NheI* site. Segment *NcoI*(418)–*Clal*(1) in pMAKK76 with the *Km^R* gene was obtained from pKANTS γ . The alternative segment in pMAKC76 was from pCANTS γ . The figure is not drawn to scale.

inducer is used at concentrations of 1–50 μ g/ml, having no significant effect on the plasmid copy number in this concentration range. At less than 1 μ g cTc/ml, the plasmid copy number drops sharply (data not shown).

Structure of the insertion plasmids

A pair of plasmids that permit insertion of *FRT* and γ -*ori* into the chromosome were constructed as described in Fig.3. The two plasmids, pMAKK76 and pMAKC76, differ only in the antibiotic-resistance marker they carry (*Km^R* or *Cm^R*). PCR-amplified copies of the targeted insertion sites of the chromosome can be cloned into the multiple cloning sites of the plasmids. The single *NotI* site allows the localization of the insert on the chromosome by running a pulsed-field gel of *NotI*-digested total DNA isolated from the cell. The phage promoters T7 and SP6 can be useful in the transcriptional–translational analysis of the excised chromosomal segment. The predicted nucleotide sequence of the plasmids is available from GenBank under the accession numbers U08460 (pMAKK76) and U08461 (pMAKC76).

Isolation of the 51-kb *lacZ*–*phoB* chromosomal segment

The excision–amplification procedure was tested on the *lacZ*–*phoB* region spanning approximately 51 kb between min 8 and 9 on the *E.coli* chromosome.

Insertion at *lacZ*. A *PvuII*–*BamHI* fragment, including an internal 330-bp segment of *lacZ*, was obtained from plasmid pMC9 (30) and cloned between the *SmaI* and *BamHI* sites of pMAKK76. The resulting plasmid was transformed into MGtFp. During the preparation of the competent MGtFp cells and the recovery period of the transformation, the LB medium was supplemented with heat-inactivated cTc at a concentration of 3

Figure 4. Gel electrophoresis of the isolated 51-kb *lacZ*–*phoB* (A) and the 110-kb *dapX*–*dsdC* (B) segments of the *E. coli* chromosome. Growth of the cells, induction and plasmid isolation are described in the text. DNA minipreps, obtained from 5-ml cultures, were run on a 0.7% agarose gel, along with control samples prepared from uninduced cells. Lanes 1,6: uninduced controls; lanes 2,7: uninduced controls, digested with *EcoRI*; lanes 3,8: induced samples; lanes 4,9: induced, digested with *EcoRI*; lanes 5,10: size markers (λ DNA *HindIII* digest). The two smallest fragments of the digest in lane 9 ran out of the gel. Arrows point to the excised circular DNA. Asterisks mark the band of the contaminating, non-specific chromosomal fragments (lanes 1,3,6 and 8).

$\mu\text{g/ml}$. After the recovery period, the transformation mixture was spread on Km plates without cTc. Typically, the transformants formed tiny colonies due to the initial presence of π in the cells, but only a fraction of the cells continued to grow on the inducer-free plates. These large colonies were tested for the possible presence of replicating plasmids due to TetR repressor mutations. Minipreps of 6 out of 10 colonies showed no plasmids, indicating integration into the chromosome. The resulting Km^R cells were designated MGtFp-l.

Insertional inactivation of *lacZ* allowed us to assess the accuracy of the homologous recombination event by inspecting the color of the colonies on indicator plates (25). In such an experiment, we have found that out of about 4000 colonies, only three were blue, indicating a negligible recombination rate at locations other than *lacZ*.

Insertion at *phoB*. A 692-bp *PvuII*–*Bam*HI fragment from the 5'-end of *phoB* (31) was obtained by PCR amplification and was ligated between the *Sma*I and *Bam*HI sites of pMAKC76. Insertion of the resulting plasmid into the chromosome of MGtFp cells was carried out as for *lacZ*, except that Cm plates were used for selection of the recombinants. The resulting Cm^R cells were designated MGtFp-p.

P1 transduction. MGtFp-p recipient cells were infected with P1 lysate prepared from MGtFp-l. Cells carrying both insertions on the chromosome were selected on Km+Cm plates and designated MGtFp-lp.

Induction and plasmid isolation. MGtFp-lp cells were grown in LB (Km+Cm) to mid-log phase. When an OD₅₅₀=0.5–0.6 was reached, FLP and π production was induced by adding one volume LBcTc (Km+Cm) to the culture. Induction at lower or higher OD resulted in a decreased yield of excised circular DNA.

Figure 5. PFGE analysis of the excision and amplification of the 51-kb *lacZ*–*phoB* genomic segment. Samples were taken from all stages of the procedure (see text). (A) Total DNA samples prepared from cells embedded in microbeads, digested with *NotI* and fractionated by PFGE. Lane 1: MG1655; lane 2: MGtFp; lane 3: MGtFp-p; lane 4: MGtFp-l; lane 5: MGtFp-lp, uninduced; lane 6: MGtFp-lp, induced for 4 h; lane 7: MGtFp-lp, induced for 12 h; lane 8: MGtFp-lp Δ N, uninduced; lane 9: MGtFp-lp Δ N, induced for 4 h; lane 10: purified 53-kb plasmid, linearized with *NotI*. The origin of some of the bands (a,b,c,d,e,f) is shown in (B). The copy number of the excised circle can be estimated by comparing the intensities of the two bands in lane 9 marked by white arrowheads (f band and a non-amplified band). Asterisks indicate multiple bands on the gel. (B) For clarity, some of the map data are presented, corresponding to the relevant *NotI* fragments in (A). Downward arrows, marked with N, represent *NotI* sites. Δ N indicates the absence of that particular *NotI* site in MGtFp-lp Δ N. Upward arrows mark *FRT*/*FLP* excision cuts. Numbers in parentheses indicate estimated fragment sizes in kb. The figure is not drawn to scale.

Concentration of cTc in the range of 3–50 $\mu\text{g/ml}$ did not alter the yield significantly. Induction was carried out in a shaker for 5 h at 30°C, the preferred temperature for FLP (5). The growth of the induced culture is delayed, as compared to MGtFp cells under the same conditions, and eventually stops after 2–3 doublings.

DNA was prepared from the induced culture by the standard alkaline lysis miniprep method (25), including a phenol extraction step.

The DNA samples were analysed by gel electrophoresis (Fig. 4). Fragment sizes of the *EcoRI*-digested sample match the *EcoRI* map of this region derived from the literature (32,33), verifying the accuracy of the procedure. Of the expected six

fragments, five match the predicted size. The sixth fragment is too small to be detected on the 0.7% agarose gel. In addition, there is an extra fragment of 2 kb which originates from the pMAKK76 insertion. Comparison of undigested samples from uninduced and induced cultures reveals that the large circular DNA (see arrows in Fig.4) has lower electrophoretic mobility in the 0.7% agarose gel than the contaminating, unspecific chromosomal fragments that are present in both preparations and run as a single band (see asterisks in Fig.4).

PFGE analysis of the excision and amplification. Cells were taken from all stages of the above described procedure, embedded in microbeads, and total DNA was prepared for PFGE analysis. Fig.5A shows the pattern of the *NotI*-digested samples. Since each of the inserts carries a single *NotI* site (Fig.2B and Fig.3), their integration into the chromosome can be traced by the changes in the *NotI* digest pattern. Comparison of the gel and the *NotI* map of the MG1655 genome (34) verifies that each of the insertions occurred at the desired locations. Some of the relevant map data are shown in Fig.5B.

Upon induction, the relative intensity of the bands carrying γ -*ori* increases. The neighboring fragments also show increased intensity. This effect extends to several hundred kb, indicating partial bi-directional replication of the chromosome from the integrated γ -*ori* sequences.

The size of the 53-kb (51 kb from the chromosome plus 2 kb from the insert) *NotI* fragment extending from *lacZ* to *phoB* on the chromosome is nearly identical to that of the excised and *NotI*-linearized circle (Fig.5B). In order to be able to determine how much of the amplified fragment originates from the excised circle, the insertion process at *phoB* was repeated with the same insertion plasmid, except that its *NotI* site was eliminated by a filling-in reaction. P1 transduction of the cells by P1 lysate of MGtFp-I resulted in the MGtFp-Ip Δ N strain with both inserts. Comparison of the induced samples of MGtFp-Ip and MGtFp-Ip Δ N (with and without the *NotI* site at *phoB*) shows that practically all of the amplified *lacZ*–*phoB* segment is in the excised circular form, since the intensity of the f band is virtually unchanged (Fig.5).

The difference in the intensities of the excised segment and of a similarly sized, non-amplified fragment (marked with arrowheads in Fig.5A) of the same sample suggests that the large circle has a moderately elevated copy number.

Isolation of the 110-kb *dapX*–*dsdC* chromosomal segment

For long-range sequencing purposes, the ideal size of the excised segments is about 100 kb. The feasibility of our procedure to obtain chromosomal DNA segments of this size was tested on the *dapX*–*dsdC* region of the *E.coli* genome at min 54–56.

Sequence information for the *dapX* and *dsdC* regions is available (35,36). A 505-bp segment, including the 5'-end and upstream regions of *dapX*, was amplified by PCR. The 5'-end of the primer carried a 8-bp non-homologous portion, including a restriction site, to facilitate the cloning of the PCR product into pMAKK76. Similarly, a 601-bp fragment from the 3'-end of *dsdC* was obtained by PCR and cloned into a derivative of pMAKK76. Insertion of the plasmids into the chromosome, P1 transduction, induction of the cells and isolation of the 112-kb circle were performed as described above for the *lacZ*–*phoB* fragment.

The *EcoRI* digestion pattern of the DNA preparation (Fig.4) shows a good match to the *EcoRI* map of this region derived from the literature (31,32). Out of 17 fragments on the gel, 13

show the expected size, three are slightly larger and there is one small extra fragment, possibly due to strain variation. The yield of the plasmid preparation is slightly lower, as compared to that of the 53-kb circle.

DISCUSSION

In vivo excision and amplification has the potential to overcome the problems of short- and long-term instability and size limitation of clones in conventional cloning methods. In principle, this approach enables one to isolate large segments of the chromosome directly from the original living organism.

As a first step, we have demonstrated the feasibility of this method by developing an excision–amplification system for *E.coli* and applying it to isolating a 51- and a 110-kb segment of the chromosome. The procedure involves the insertion of *FRT* and γ -*ori* sequences into two, predetermined chromosomal sites. *FLP*-mediated excisional recombination and π -initiated replication result in a large plasmid carrying the chromosomal segment flanked by the insertions. Since excision and amplification occur only after induction of *FLP* and π , the cells carrying inserts can be stably maintained in the uninduced state.

The conditionally expressed genes encoding *FLP* and π can be inserted into the host chromosome or can be maintained on a small plasmid. We have opted for the chromosomal location to avoid the need of separating the large and small plasmids after DNA preparation and thus simplifying the analysis of the products.

We have shown that *FLP* performs recombinational excision between two distant *FRT* sites on the chromosome, even at a distance of 110 kb. Within the limits of detection, the excision reaction reached completion within a few hours of induction, as demonstrated for the 51-kb segment.

The insertions were placed into preselected chromosomal locations with the aid of homologous recombination. Chances of the recombination event were enhanced by temporarily increasing the copy number of the otherwise non-replicating insertion plasmids in the cell. Target sequences for insertion are not limited to non-essential genes or flanking and intergenic regions. If the homologous copy covers the entire 5'- or 3'-end of a gene, together with the necessary flanking sequences, insertion results in a complete and a partial copy of that gene on the chromosome.

We have isolated two different regions of the *E.coli* chromosome. The excised and amplified segments were isolated as large plasmids. Since the overall yield was low, a relatively large proportion of the preparation consisted of contaminating, non-specific chromosomal fragments (Fig.4). This contamination, though insignificant in molar terms, might pose a problem if the large circle were subjected to subcloning. For separation of the circular DNA from the contaminating linear pieces, the difference in their mobility in low-percentage agarose gels could be exploited. Another alternative is the digestion of the preparation with an exonuclease specific for linear DNA (Epicentre).

Excision and amplification is a concerted action of *FLP* and π . In this process, a part of the chromosome, proximal to the γ -*ori* sites, also gets amplified (Fig.5A). Even if excision precedes amplification, one of the γ -*oris* remains on the chromosome, initiating replication several times. Partial multiplication of the chromosome puts a strain on the cells that eventually stop growing. A different arrangement of the inserts could result in the excision of both γ -*oris*, circumventing the problems of

chromosomal replication, and directing all the replication resources to the amplification of the excised piece. We have tested such an arrangement (data not shown), but no increase in the yield was observed. Probably the excised chromosomal piece is so large that its amplification is already toxic for the cell.

The π -mediated amplification resulted in the moderately elevated copy number of the excised DNA, as shown in the case of the 51-kb segment (Fig.5). Use of an *ori* that allows amplification to proceed, while other cell functions become progressively shut down, might improve the yield. However, we are confident that the method in its present form is useful in manipulations of the *E.coli* genome. The procedure is relatively simple and rapid, and results in sufficient amount of the large DNA segment for most purposes. In addition to the isolation of genomic segments, the method could also be used to create large, non-lethal deletions of the chromosome for genetic analysis. After the excision step, removal of the inducer from the culture by changing the medium would result in the loss of the excised, non-replicating piece in subsequent divisions.

Identically oriented insertions make a walking process possible. The second (right-end) *FRT* insert will become the first (left-end) one in the next walking step. Fifty inserts, spread at 100-kb intervals, would cover the entire *E.coli* genome. Any two *FRT* neighbors could then be combined by P1 transduction to excise, amplify and isolate any particular DNA segment.

The principles and some of the elements (e.g., FLP-*FRT* excision, targeting by homologous recombination) of our *E.coli* excision-amplification system might be used in similar procedures in other prokaryotes or eukaryotes, as well. Currently, we are working on an excision-amplification system for yeast.

ACKNOWLEDGEMENTS

We thank Dr Tamas Gaal (Department of Bacteriology, University of Wisconsin) for his suggestions and comments in the course of this work. This work was supported by NIH grants 1-RO1-GM39715, 1-RO1-00504 to W.S., and NCI Core Grant 5-PO1-CA07175. Z.H. was supported by the Lucille P. Markey Charitable Trust Grant #92-24 to Dr Ross Inman (Institute of Molecular Virology, University of Wisconsin).

REFERENCES

- Szybalski, W. (1991) *Human Genome III*, San Diego, CA, October 21-29, p.8.
- Szybalski, W. (1992) *Fresenius J. Analyt. Chem.*, 343, 4.
- Szybalski, W. (1993) *Gene*, 135, 279-290.
- Nash, H.A. (1981) *Annu. Rev. Genet.*, 15, 143-167.
- Cox, M.M. (1989) In Berg, D.E. and Howe, M.M. (eds), *Mobile DNA*, Am. Soc. Micro., Washington, DC, pp. 661-670.
- Sternberg, N. and Hamilton, D. (1981) *J. Mol. Biol.*, 150, 467-486.
- Broach, J.R. and Hicks, J.B. (1980) *Cell*, 21, 501-508.
- Golic, K.G. and Lindquist, S. (1989) *Cell*, 59, 499-509.
- O'Gorman, S., Fox, D.T. and Wahl, G.M. (1991) *Science*, 251, 1351-1355.
- Gronostajski, R.M. and Sadowski, P.D. (1985) *J. Biol. Chem.*, 260, 12320-12327.
- Beatty, L.G., Babineau-Clary, D., Hogrefe, C. and Sadowski, P.D. (1986) *J. Mol. Biol.*, 188, 529-544.
- Kontomichalou, P., Mitani, M. and Clowes, R.C. (1970) *J. Bacteriol.*, 104, 34-44.
- Kolter, R., Inuzuka, M. and Helinski, D.R. (1978) *Cell*, 15, 1199-1208.
- Stalker, D.M., Kolter, R. and Helinski, D.R. (1982) *J. Mol. Biol.*, 161, 33-43.
- Filutowicz, M., Dellis, S., Levchenko, I., Urh, M., Wu, F. and York, D. (1994) *Prog. Nucl. Acid Res. Mol. Biol.*, in press.
- Metcalf, W.W., Jiang, W. and Wanner, B.L. (1993) *Gene*, 138, 1-7.
- Greener, A., Filutowicz, M.S., McEachern, M.J. and Helinski, D.R. (1990) *Mol. Gen. Genet.*, 224, 24-32.
- Cox, M.M. (1983) *Proc. Natl. Acad. Sci. USA*, 80, 4223-4227.
- Filutowicz, M., unpublished results.
- Koob, M., Hasan, N. and Szybalski, W., manuscript submitted for publication.
- Senecoff, J.F., Bruckner, R.C. and Cox, M.M. (1985) *Proc. Natl. Acad. Sci. USA*, 82, 7270-7274.
- Daniels, D.L., Plunkett, G.III, Burland, V. and Blattner, F.R. (1992) *Science*, 257, 771-778.
- Miller, V.L. and Mekalanos, J.J. (1988) *J. Bacteriol.*, 170, 2575-2583.
- Bochner, B.R., Huang, H.-C., Schieven, G.L. and Ames, B.N. (1980) *J. Bacteriol.*, 143, 926-933.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Dretzen, G., Bellard, M., Sassone-Corsi, P. and Chambon, P. (1981) *Anal. Biochem.*, 112, 295-298.
- Miller, J.H. (1992) *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 271-273.
- Koob, M. (1992) *Methods Enzymol.*, 216, 13-20.
- Koob, M. and Szybalski, W. (1990) *Science*, 250, 271-273.
- Calos, M.P., Lebkowski, T.S. and Botchan, M.R. (1983) *Proc. Natl. Acad. Sci. USA*, 80, 3015-3019.
- Makino, K., Shinagawa, H., Amemura, M. and Nakata, A. (1986) *J. Mol. Biol.*, 190, 37-44.
- Kohara, Y., Akiyama, K. and Isono, K. (1987) *Cell*, 50, 495-508.
- Rudd, K. (1992) In Miller, J.H.: *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 2.3-2.43.
- Heath, J.D., Perkins, J., Sharma, B. and Weinstock, G. (1992) *J. Bacteriol.*, 174, 558-567.
- Palchaudhuri, S., Patel, V. and McFall, E. (1988) *J. Bacteriol.*, 170, 330-334.
- Tiedeman, A.A., DeMarini, D.J., Parker, J. and Smith, J.M. (1990) *J. Bacteriol.*, 172, 6035-6041.
- Shafferman, A., Kolter, R., Stalker, D.M. and Helinski, D.R. (1982) *J. Mol. Biol.*, 161, 57-76.