Volume 15 Number 16 1987 Nucleic Acids Research

# Purification of the FLP site-specific recombinase by affinity chromatography and re-examination of basic properties of the system

Leslie Meyer-Leon, Cynthia A.Gates, Janet M.Attwood, Elizabeth A.Wood and Michael M.Cox*

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

Received May 5, 1987; Revised and Accepted July 27, 1987

ABSTRACT

The FLP protein, a site-specific recombinase encoded by the 2 micron plasmid of yeast, has been purified to near homogeneity from extracts of E. coli cells in which the protein has been expressed. The purification is a three column procedure, the final step employing affinity chromatography. The affinity ligand consists of a DNA polymer with multiple FLP protein binding sites arranged in tandem repeats. This protocol yields 2 mg of FLP protein which is 85% pure. The purified protein is highly active, stable for several months at -70°C and free of detectable nucleases. The molecular weight and N-terminal sequence are identical to that predicted for the FLP protein by the DNA sequence of the gene. Purified FLP protein primarily, but not exclusively, promotes intramolecular recombination. Intermolecular recombination becomes the dominant reaction when E. coli extracts containing no FLP protein are added to the reaction mixture. These extracts are not specifically required for recombination, but demonstrate that some properties previously attributed to FLP protein can be assigned to contaminating proteins present in E. coli.

## INTRODUCTION

The two-micron circle is a 6318 bp plasmid endogenous to naturally occuring populations of Saccharomyces cerevisae (1). One copy of a 599 bp repeat lies on each side of this plasmid in inverted orientation. Site-specific recombination between the repeats results in inversion of the intervening sequences (2). The FLP protein (predicted molecular weight 48,794), encoded by the largest of the 2 micron open reading frames, mediates the recombination event. FLP protein recombination has been shown to participate in 2 micron plasmid copy-number amplification (3,4). Expression of the FLP protein in E. coli led to the development of in vitro systems for this recombination event (5,6,7). The reaction requires only a buffer and appropriate ionic strength. No high energy cofactor or divalent cation is necessary (7). FLP protein will promote either intramolecular (inversions and deletions) or intermolecular recombination, using sites present on both relaxed and supercoiled DNA molecules (6,7). The relative efficiency of

© IRL Press Limited, Oxford, England.

inter- and intramolecular recombination was reported to be affected by the concentration of FLP protein. Inhibition of the reaction occurred at still higher FLP protein concentrations (8).

The recombination site is a subset of the 599 bp sequence. Each site is comprised of three thirteen base pair repeats. The first two repeats are in direct orientation, separated by one base pair. The third repeat is inverted in orientation from the other two, and separated from them by an eight base pair spacer. The minimum site for efficient in vitro recombination is 28 bp in length, including the spacer region and the ten internal basepairs of each adjacent inverted repeat (9,10). The entire site, however, including the outside repeat, is protected from nonspecific nuclease digestion by FLP protein (11). Methylation interference studies (12) and mutation analysis (9,10) have shown that FLP protein binds to the 13 bp repeats but not to the 6 central bp of the spacer. During recombination FLP protein produces staggered cuts at the boundaries of the spacer, generating protruding ends 8 nucleotides in length with 5'-(OH) termini (9,11). The FLP protein forms a covalent 3'-phosphotyrosyl linkage with the DNA at the sites of cleavage (13). Mutant recombination sites containing only one 13 bp repeat, or two tandem direct repeats, showed diminished nuclease protection and were not cleaved (11).

Further characterization of the molecular details of this recombination event requires a purification method for significant quantities of pure FLP protein. Several protocols for the partial purification of this protein have been published (7,14,15). The best of these produced FLP protein originally estimated to be 50% pure (15). More recently fractions prepared by this protocol have been found to be 5% pure (16; J.M. Attwood, unpublished results). We report here a purification procedure which yields milligram quantities of FLP protein at 85% purity. A key step in the protocol involves affinity chromatography with an immobilized DNA ligand. The ligand design is based on known interactions between FLP protein and its recombination site (9-13). With cleaner protein preparations we have re-examined several previously described properties of the in vitro reaction (8).

## MATERIALS AND METHODS

Plasmids: The plasmid pMMC3 is described elsewhere (5), and is maintained in the E. coli recA$^-$ strain HB101 (17). DNA was purified by banding twice in Ethidium Bromide/Cesium Chloride gradients. The plasmid pMMC20 was maintained in strain C600K$^-$ (17), which was provided by Dr. Carol Gross (University of Wisconsin-Madison). The plasmid pKK223-3 was obtained

from Pharmacia-Molecular Biology Division (18) and was also transformed into strain C600K$^-$.

Enzymes and Reagents: Restriction enzymes, T4 polynucleotide kinase, and terminal deoxynucleotidyl transferase were purchased from New England Biolabs and Bethesda Research Labs. Alpha-macroglobulin is a product of Boehringer Mannheim Biochemicals. T4 DNA ligase was purified from an overproducing strain as described (19). Biogel P6-DG, heparin agarose and Biorex-70 resins were purchased from BioRad. Cyanogen bromide activated Sepharose 4B is from Pharmacia. Enzyme grade ammonium sulfate was purchased from Schwartz/Mann. Isopropyl-thio $\beta$-D-galactoside (IPTG) was obtained from BACHEM Fine Chemicals. The protease inhibitors phenylmethylsulfonyl fluoride (PMSF), benzamidine, pepstatin A, aprotinin, ovomucoid, and $\alpha$-antitrypsin were purchased from Sigma, as were ampicillin, and N-2-hydroxyethylpiperidine-N-2 ethane sulfonic acid (HEPES). Tris[hydroxymethyl]methylaminopropane sulfonic acid (TAPS) is from Research Organics, Inc.

Instrumentation: Column fractions were monitored with an ISCO UA-5 Absorbance/Fluorescence detector. Scanning densitometry was performed on a Zeineh Soft Laser Scanning Densitometer, Model SL-504-XL. Oligonucleotides of defined sequence were synthesized at the Biotechnology Facility of the University of Wisconsin-Madison on an Automatic Synthesizer (Applied Biosystems, Model 380A).

Media and Buffers: Buffer H contains 25 mM Hepes, 14% anion (pH 7.0 at 5°C), 1 mM DTT and 1 mM EDTA, with varying amounts of NaCl added to give Buffer H(X) where X is the molar concentration of NaCl. Buffer A contains 20 mM Tris-HCl, 80% cation (pH 7.5 at 25°C), 10% (w/v) glycerol, 1 mM EDTA, and 1 mM DTT, again with varying amounts of NaCl to give Buffer A(X). TSE buffer contains 50 mM Tris-HCl, 50% cation (pH 8.1 at 25°C), 10% (w/v) sucrose and 20 mM EDTA. TGA buffer for SDS-polyacrylamide gel electrophoresis contains 50 mM Tris base, 0.38 M glycine and 0.1% SDS.

Construction of pMMC20: The plasmid pMMC4 (5) contains the coding sequences of the FLP protein downstream from the rightward promoter and *cro* gene translational start signals of the lambda bacteriophage. Several successive changes were made to pMMC4 in an attempt to optimize FLP overproduction. The first change was an attempt to optimize the first eleven codons of the gene, since several of these codons may be inefficiently recognized during translation in *E. coli* (20,21). The SphI to SauI region of pMMC4 was deleted and replaced with a 36 bp synthetic duplex oligonucleotide with SphI and SauI ends. This changed the third base pair of each of eight

codons, but no alterations were made to the amino acid sequence (Figure 1B). The second change was motivated by a published strategy for overexpressing bovine growth hormone (22). We positioned the coding sequences for a short peptide of six amino acids upstream of the flp gene, contiguous with the cro AUG and followed by a TAA stop codon. This was done by substituting a 23 bp synthetic duplex DNA fragment with BamHI and SphI ends in place of the 187 bp BamHI to SphI segment of pMMC4. A consensus ribosome binding site (23) is included within the sequence for the short peptide, and is therefore positioned 5′ of the flp gene (Figure 1B).

The third change replaced the heat-inducible $P_r$ promoter with the IPTG-inducible $P_{tac}$ promoter. This involved modifying the expression vector pKK223-3, which contains a linker region downstream from the $P_{tac}$ promoter, and then inserting into it the coding sequences described above. To modify pKK223-3, the EcoRI to HindIII fragment of the linker region was excised and replaced with a synthetic duplex DNA fragment containing BamHI and XbaI sites (BamHI proximal). A second BamHI site, located between the $P_{tac}$ and $Tet^s$ sequences of pKK223-3, was eliminated by filling-in the site with the Klenow fragment of DNA polymerase I after partial BamHI digestion, followed by ligation. The BamHI to XbaI segment of the pMMC4-derived construction above was inserted into the modified pKK223-3. This plasmid now contains the $P_{tac}$ promoter, followed by the coding sequences for the six amino acid peptide, the flp gene and the ribosomal RNA transcription terminator region called $rrnBT_1T_2$5S (18). Finally, the last change added the gene coding for the lactose repressor. This was obtained from the plasmid p117 (24). The SalI to PvuI fragment of p117 contains the 1.8 kb lacI repressor gene and 3.7 kb of pBR322 DNA, including the origin of replication. The pBR322 SalI to PvuI sequences of the newly modified pKK223-3 (including the replication origin) were replaced with this lacI-bearing segment from p117. This final 8.5 kb construct is designated pMMC20, and is illustrated in Figure 1A.

Assay of FLP protein activity in vitro: The plasmid pMMC3 (5) contains a unique BamHI restriction site. Treatment with BamHI produces a 9400 bp linear DNA which is the standard substrate for recombination. The plasmid contains two FLP recombination sites in the same orientation. This substrate can undergo either a deletion or an intermolecular reaction. An intermolecular reaction results in either no change in DNA size, or in larger (P1; 12.43 kb) and smaller (P2; 6.53 kb) DNA products. Additional intermolecular recombination events can result in multiple product bands that are larger than the substrate and migrate above it on an agarose gel.

Intramolecular recombination results in the deletion of the intervening region between the recombination sites, producing a small circular deletion product (P3; 2.94 kb) and a linear product which is equivalent to the small linear product of the intermolecular insertion (P2). Reactions often produce multiple P3 topoisomers. The product designated P1 is uniquely diagnostic of intermolecular recombination, while P3 is diagnostic of intramolecular recombination. A more detailed explanation of the reaction may be found in ref. 7.

Quantitation of the amount of FLP activity present in a given fraction is accomplished by serial dilution of the fraction in 25 mM TAPS (pH 8.0 at 25°C), 1.0 mM EDTA and 1.0 M NaCl. Reaction mixtures (20 μl total volume) contain in final concentration 25 mM TAPS buffer, 29% anion (pH 8.0 at 25°C), 200 mM NaCl, 1 mM EDTA, 200 ng of pMMC3 substrate DNA linearized with BamHI, and 2 μl of the appropriate dilution of an FLP protein-containing fraction. The concentration of total recombination sites is 3.2 nM (0.064 pmoles/0.02 ml reaction). Reactions proceed for 60 minutes at 30°C, and are terminated with 2 μl of 10% SDS plus 8 μl of GED (60% (w/v) glycerol, 50 mM EDTA and 0.05% bromophenol blue). Reaction products and substrates are separated by gel electrophoresis on 0.8% (w/v) agarose, stained in 0.1% ethidium bromide, visualized with a short-wave ultraviolet light source, and photographed with a Polaroid MP-4 Land Camera. In some cases product formation was quantitated by densitometric scanning of photographic negatives. No data correction was employed for bands which were beyond the linear range of the densitometer. A unit of FLP protein activity is defined as the minimum amount of protein required to produce products detectable in an agarose gel after a 60 minute reaction under these conditions. Where it is necessary to identify column fractions containing FLP protein, qualitative assays are performed. One μl of each fraction is assayed in a single reaction.

Preparation of DNA for Site-Specific Affinity Column: The annealed synthetic oligonucleotides and the strategy for their polymerization are illustrated in Figure 3. Conditions for enzyme reactions and DNA precipitations are as described (17). One μmole of each single-strand oligonucleotide was phosphorylated, using 100 units T4 polynucleotide kinase. Reactions included an ATP regenerating system (2.24 units/ml pyruvate kinase, 2.31 mM phosphoenolpyruvate, 0.44 mM KCl; see ref. 25). After incubating at 37°C for 15 hours, the mixture was extracted with phenol/chloroform and precipitated with ethanol. The complementary strands were mixed together and resuspended in TE buffer (17). The oligonucleotides were annealed by heating

to 80°C for 5 minutes, then cooling slowly to 30°C before storing on ice. The oligonucleotides were polymerized with 3000 units of T4 DNA ligase. An ATP regenerating system was also included with the reaction. After 3 days at 25°C, the DNA was precipitated with ethanol and resuspended in $H_2O$. The reaction products ranged in length from 112 to 280 base pairs, or 8 to 20 tandem repeats. PolydA was added to the 3' ends using 42 units terminal deoxynucleotidyl transferase. The reaction was incubated for 5 hours at 37°C, extracted twice with phenol/chloroform and precipitated with ethanol. The precipitated DNA was resuspended in 10 mM sodium phosphate buffer, pH 8.0, to a final concentration of 14.25 mg/ml.

Coupling of DNA to Activated Sepharose: DNA polymers were attached to cyanogen bromide-activated Sepharose resin according to a modification of a published method (26). Commercially available cyanogen-bromide activated Sepharose 4B (4.5 g dry weight) was washed serially by suction filtration with 4 liters of 1 mM HCl and 1 liter of 10 mM sodium phosphate buffer, pH 8.0. The washed CN-Sepharose was suspended in 15 ml of the same buffer, and 5.7 mg of annealed DNA polymers were added. The mixture was agitated gently for 15 hours in a shaker at 25°C, filtered, and rinsed six times with 50 ml each of 10 mM sodium phosphate buffer, pH 8.0. The $A_{260}$ of each rinse was monitored and the total amount of DNA in the rinses found to be 3.41 mg. The concentration of DNA/gram Sepharose was 0.51 mg/g. Similar results were obtained when this method (26) was used to attach the same DNA to cyanogen bromide-activated Sepharose 2B prepared by the method of Kohn and Wilchek (1982, ref. 27) (C. Gates, unpublished results).

Fractionation of E. coli Extracts Which Do Not Contain FLP Protein: $C600K^-$ cell paste (100 grams) was sonicated with two one-minute bursts in a solution containing 100 ml TSE, 0.1 M NaCl, $10^{-7}$ M pepstatin A, 6 μg/ml α-macroglobulin, and 30 μg/ml PMSF. The lysate was spun at 48,000 rpm (150,000 g) for 30 minutes in a Beckman 60Ti rotor. The supernatant was loaded directly over a column of Biorex-70 resin (5x15 cm). The column was washed sequentially with 300 ml Buffer A (0.2 M NaCl), and 300 ml Buffer A (1.0 M NaCl). One protein peak eluted with the 1.0 M NaCl wash. The protein in this peak was used to test the effect of E. coli extract on the FLP in vitro recombination reaction. This procedure is similar to the first few steps of the FLP protein purification.

Other Methods: Quantitative protein determinations were made either by the method of Warburg and Christianson (28) or by the method of Bradford (29) using bovine serum albumin as a calibration standard, as indicated.

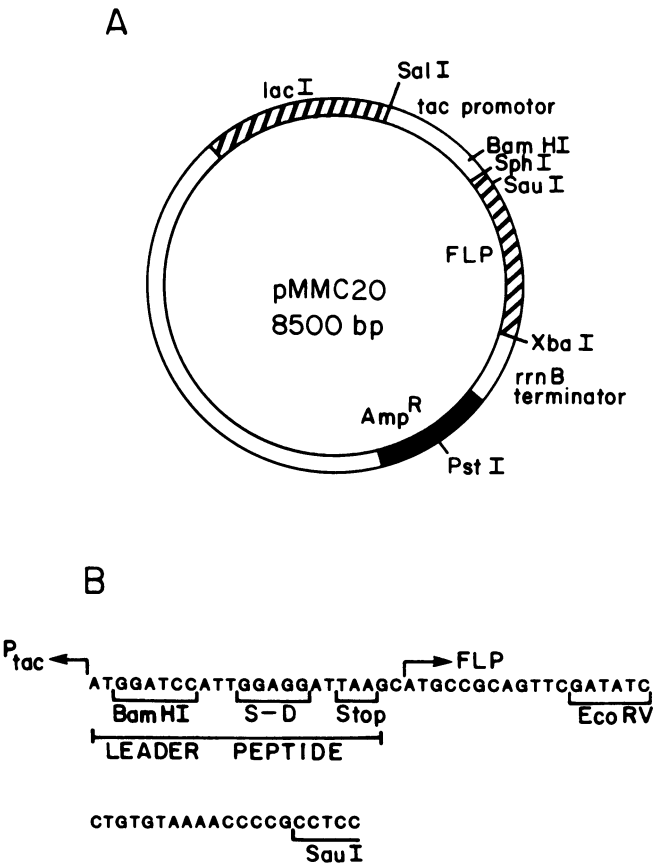



FIG. 1. Construction of pMMC20. A) Diagram of the pMMC20 Expression vector. The relative size and position of coding and restriction sequences are approximate. Coding sequences for FLP protein and lacI repressor, and for the ampicillin resistance gene, are in hatched and solid lines respectively. B) Sequences involved in translation initiation of FLP protein. The sequence begins with the initiation codon of a short leader peptide as described in the text. This peptide is terminated before the initiation codon for the flp gene. Eight base pairs are altered in the beginning of the flp gene to optimize codon usage, as described in Materials and Methods. The modified sequences are shown here. The remainder of the flp gene is unmodified.

## RESULTS

### Cloning and Expression of FLP Protein

The plasmid pMMC20 (8500 bp, Figure 1) is a pBR322-derived vector containing genes for FLP protein and the lacI repressor. Expression of FLP protein is induced with IPTG from the tac promoter. No FLP activity is detected in crude extracts of cells harvested without IPTG induction (data not

TABLE I
Purification of FLP Protein

| Fraction | Volume | Total Protein | Total Units | Specific Activity | Recovery |
|---|---|---|---|---|---|
| | ml | mg | $x10^5$ | units/mg $x10^3$ | % |
| I Crude Extract | 180 | 11000 | 120 | 1.1 | (100) |
| II $(NH_4)_2SO_4$/P6DG | 400 | 3500 | 16 | 0.3 | 13 |
| II Biorex 70 | 290 | 200 | 12 | 6.0 | 10 |
| IV Heparin Agarose | 55 | 44 | 9 | 20 | 8 |
| V Site-Specific DNA Sepharose | 11 | 2 | 7 | 380 | 6 |

Protein concentrations were determined by the method of Warburg and Christianson (1942).

shown). Based on the amount of FLP protein generated by our purification protocol, and taking into account the estimated loss in yield at each step (Table I), FLP protein is calculated to represent approximately 0.3% (w/w) of the total cellular protein after induction. In an attempt to increase the production of FLP protein we have employed many strategies used routinely for the overexpression of proteins in E. coli. These include the use of a) several different promoters, b) protease deficient host strains, c) codon optimization, d) high-copy origins of replication, e) downstream transcription termination and f) optimization of translation start signals. The plasmid pMMC20 incorporates several of these strategies, as described in Methods. However, no significant increase in FLP protein production, over the levels indicated above, has resulted to date (data not shown).

Cell Growth and Induction of FLP Protein

E. coli strain C600K$^-$, harboring the vector pMMC20, is grown at 30°C in AZ broth (30) supplemented with 0.1 mg/ml ampicillin in a 400 liter New Brunswick Fermenter. When the culture reaches an optical density (595 nm) of 0.25, the synthesis of FLP protein is induced with 0.5 mM IPTG. After an additional 2.5 hours the cells are harvested in a Sharples continuous-flow centrifuge and resuspended thoroughly in a volume of TSE buffer equivalent to 5/8 the weight of the collected cells. The resulting cell paste is frozen dropwise in liquid nitrogen and stored at -70°C. Aliquots of 100 grams are used for each enzyme preparation.

Purification of FLP Protein

Lysis: For each preparation of FLP protein, 100 grams of cell paste is

thawed by continuous stirring while immersed in a water bath at 30°C. All subsequent steps are carried out at 4°C unless noted otherwise. The cells are diluted with 100 ml cold TSE buffer. To this solution is added 1/100 volume of each of the following: aprotinin (2.3 mg/ml), a mixture of ethanol-soluble protease inhibitors (5 mg/ml PMSF, 157 mg/ml benzamidine, and 7 μg/ml pepstatin A in 95% ethanol), a mixture of water-soluble protease inhibitors (50 mg/ml ovomucoid, 2.5 mg/ml α-antitrypsin, and 0.5 mg/ml α-macroglobulin), and 0.5 M EDTA. The suspended cells are sonicated with 4 one-minute bursts at 130 Watts on a Sonifier Cell Disrupter, Model W185, and are spun for 30 minutes at 48,000 rpm (150,000 g) in a Beckman 60Ti rotor. The supernatant is designated Fraction I.

Ammonium Sulfate Precipitation: Fraction I is precipitated with dry enzyme-grade ammonium sulfate to 45% saturation (0.277 grams $(NH_4)_2SO_4$ per ml initial volume), and stirred for 30 minutes. The ammonium sulfate suspension is spun for 20 minutes in a Beckman JA14 rotor at 13,000 rpm (26,000 g). The supernatant is decanted, and the pellet gently resuspended in a stainless steel Waring Blender with 70 ml Buffer H (1.0 M NaCl) (final volume 135 ml). The resulting solution is desalted by passage through a 900 ml (7.5x20 cm) column of Biogel-P6DG resin, equilibrated with Buffer H (0.2 M NaCl). The flow rate is approximately 10 ml/min under gravity. The excluded peak, containing the detectable FLP protein activity, is fraction II. Protease inhibitors are again added, as described in "Lysis".

Cation Exchange Chromatography: Fraction II (400 ml) is loaded at 2.5 ml/min onto a 100 ml (2.5x21 cm) column of Biorex-70 resin equilibrated with buffer H (0.2 M NaCl). After washing with 200 ml of buffer H (0.2 M NaCl), the column is eluted at 0.8 ml/min with a 600 ml linear gradient composed of equal volumes of buffer H (0.3 NaCl) and buffer H (0.8 M NaCl). Six ml fractions are collected. Activity assays for FLP protein are performed on alternate fractions across the gradient. Active FLP fractions, usually appearing a third of the way into the gradient, are pooled to give Fraction III.

Heparin Agarose Chromatography: Fraction III (290 ml) contains approximately 0.55 M NaCl by conductivity measurements. It is diluted with 310 ml buffer H (no NaCl) to a conductivity equivalent to 0.25 M NaCl, and loaded at 0.6 ml/min over a 20 ml (1.5x12.5 cm) column of heparin agarose which had been equilibrated with buffer H (0.25 M NaCl). The column is then washed with buffer H (0.25 M NaCl) until protein is no longer detectable by UV absorption in the eluate. The column is washed with 80 ml of buffer H (0.35 M

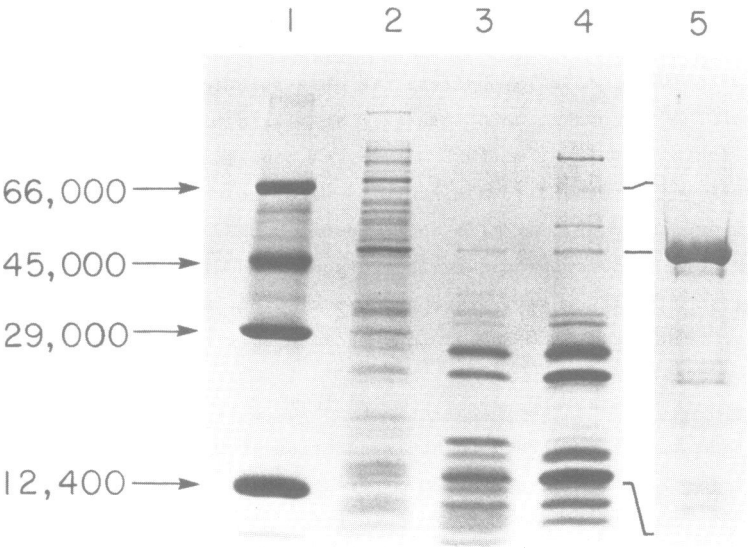



FIG. 2. SDS-polyacrylamide gel electrophoresis of FLP protein fractions purified by site-specific Sepharose chromatography. Protein in lanes 3-5 was precipitated in 10% trichloroacetic acid. Samples were denatured by boiling for 10 minutes in cracking buffer before loading on an SDS-polyacrylamide gel with 11% separating gel and 6% stacking gel (33). The gel was stained in Coomassie Brilliant Blue G-250 (34). Lane (1) marker proteins: bovine serum albumin, $M_r$=66,000; ovalbumin, $M_r$=45,000; carbonic anhydrase, $M_r$=29,000; cytochrome c, $M_r$=12,400. (2) 35 μg Fraction I, crude extract; (3) 35 μg Fraction III, Biorex-70; (4) 35 μg Fraction IV, heparin agarose; (5) 20 μg Fraction V, Site-Specific Sepharose Column. Lane 5 was run on a smaller gel apparatus, and the photograph enlarged to align with adjacent lanes. The corresponding position of marker proteins is shown by the solid lines between lanes 4 and 5.

NaCl). The small amount of protein that is washed from the column in this wash does not include FLP protein activity. FLP protein is eluted with buffer H (0.55 M NaCl), collecting 3.3 ml fractions. The protein peak is detected with a UV monitor at 280 nm, and confirmed by Bradford assays. The protein-containing fractions are pooled and designated Fraction IV.

Affinity Chromatography: Fraction IV (55 ml) is diluted with 41 ml of buffer H (no NaCl) and 24 ml of sterile 50% (v/v) glycerol, a dilution sufficient to lower the NaCl concentration to 0.25 M NaCl. The diluted fraction is loaded onto a 15 ml (1.5x8.5 cm) column of the liganded Sepharose resin, which had been equilibrated with buffer H (0.25 M NaCl). The protein pool is loaded by recycling over the column at 0.5 ml/min for 35 hours, followed by two complete passes of the protein fraction through the column.

Recycling of the protein was found to enhance yield approximately four-fold relative to loading the column directly, and resulted in little loss of activity. The column is then washed with 55 ml of buffer H (0.25 M NaCl) at 0.5 ml/minute, followed by 45 ml of buffer H (0.4 M NaCl). FLP protein is eluted with buffer H (1.0 M), collecting 1.0 ml fractions. Fractions containing FLP protein were detected by UV absorbance, and confirmed by assaying for recombination activity. Active fractions are pooled (Fraction V) and analyzed for activity and purity. The results of this procedure are summarized in Table I. The FLP protein activity peak, 11.0 ml total, contained 2.0 mg of FLP protein and was approximately 85% pure by densitometric scan (Figure 2). The procedure reproducibly yields 1-2 mg of protein at a purity between 80 and 90% (Table I). The gel lanes routinely used to estimate purity (e.g., Fig. 2, lane 5) generally contain amounts of FLP protein which are beyond the linear range of the densitometer. For this reason the degree of purity may be slightly underestimated. Fraction V is made 10% (w/v) in glycerol, frozen in liquid nitrogen and stored at -70°C until use. We find FLP protein to be very stable, as no loss of activity is observed for up to one month on ice and up to ten months at -70°C.

Comments on the Purification Procedure

This protocol achieves an overall 660-fold purification of FLP protein. Although the procedure requires four days to complete, FLP protein is very stable beyond Fraction III. Earlier steps must be conducted quickly, although yields are improved by the presence of protease inhibitors. Major losses of FLP protein using this procedure occur on the P6-DG desalting column, but we have not found a satisfactory replacement for this step. The heparin agarose column is necessary before the site-specific column, since preparations in which Fraction III was loaded directly over the affinity column did not yield pure FLP protein. The extra step is necessary to remove competitive DNA binding proteins, and to reduce the volume of material loaded over the last column. The observation that recycling Fraction IV over the site-specific DNA Sepharose column improved binding and yield was unexpected. The repeated exposure of FLP protein to the column may be required because of competition from other proteins for available binding sites. FLP protein activity is very stable during the 35 hour recycling. The first 13 N-terminal amino acids were sequenced (data not shown), and the sequence is identical to that previously described for FLP protein (15).

Rationale for Design of Affinity Ligand

In designing a ligand for a site-specific affinity column we used

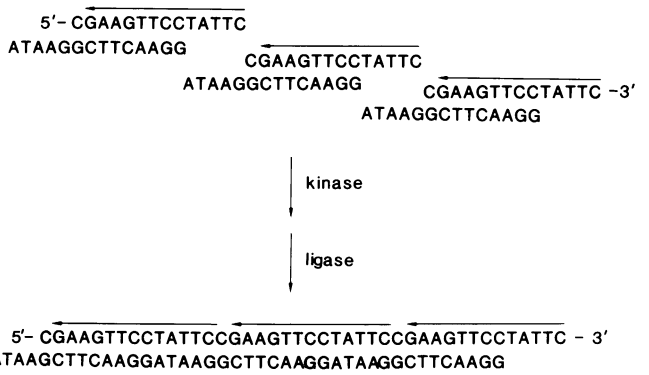



FIG. 3. Construction of DNA polymer for site-specific Sepharose column. Oligonucleotide monomers were polymerized as described in Methods. Horizontal arrows define the 13 bp repeat sequences recognized as binding sites by FLP protein. For details of binding interactions see ref. 12.

available information on the way FLP interacts with its recombination site (9-13). Several strategies were tested. For the first trial a duplex oligonucleotide containing half the spacer and one adjacent repeat was used to specifically elute FLP protein from the heparin agarose column. Affinity elution was specific and resulted in protein of up to 99% purity. However, a significant fraction of the resulting protein (greater than 50% in some cases) was found to be covalently linked to the DNA fragment used for elution (data not shown). This ligand was therefore unsuitable for preparative purification. The second test was of an oligonucleotide containing the 13 bp repeat, but lacking the adjacent spacer and base pair to which FLP protein covalently attaches. This oligonucleotide lacked specificity, and therefore was also unsuitable for preparative purification (data not shown). It was necessary, therefore, to design a ligand which would be highly specific, but incompetent for cleavage and covalent attachment. Since the cost of the elution procedure described above proved to be exorbitant, the ligand was immobilized to permit multiple use.

The wild type recombination site contains three FLP protein binding sites. The site distal from the spacer is not required for recombination, but it is bound by FLP protein (11,12). Since cleavage does not occur between two protein binding sites on the same side of the spacer, it was reasoned that a tandem array of such sites, each related to the next as these two are, would be an effective affinity matrix. Construction of this polymer is illustrated

in Figure 3, and details are presented in Methods. The resulting DNA was linked to Sepharose resin. Binding of FLP protein to this affinity column is reversible, and specific enough to permit separation from other DNA binding proteins.

Properties of Purified FLP Protein

The absence of contaminating nuclease activities in this preparation is demonstrated by incubating 116 ng (80 units) of purified FLP protein with nonspecific DNA in 10 mM $Mg^{2+}$ without EDTA for 60 minutes at 30°C. No degradation of linear duplex DNA, single-strand DNA, or supercoiled duplex DNA was detected when 300 ng of each are in turn included in a 0.01 ml reaction under conditions standard for FLP protein recombination, and analyzed on ethidium-bromide stained agarose gels.

In the absence of all other ions, FLP protein-promoted recombination occurs in the broad range of 70-400 mM NaCl. Activity is optimal at 200-300 mM NaCl. FLP protein is also active over a broad pH range. Reactions occur in a variety of buffering reagents from pH 5.5 to pH 10.0 and are optimized in TAPS buffer, pH 8.0 (data not shown).

To examine the amount of FLP protein necessary for recombination we titrated decreasing amounts of FLP protein into the standard reaction with 200 ng of DNA substrate. In Figure 4, varying amounts of FLP protein were added, from 2.4 pmoles (116 ng, lane 1) to 0.03 pmoles (1.5 ng, lane 13). A strong intramolecular and some intermolecular reaction is seen at a 38:1 molar ratio of FLP protein per recombination site (lane 1). A minimal reaction is detected with 0.03 pmoles FLP, representing a 1:1 molar ratio of FLP protein to recombination site (lane 12). The lanes in Figure 6 were scanned with a laser densitometer to estimate the relative abundance of substrate and products. A maximal reaction (approximately 74% conversion of substrate to products) is observed with the addition of 58 ng or more of FLP protein. The extent of conversion to products is plotted as a function of the molar ratio of FLP protein monomers to recombination sites (panel B). A half maximal reaction is observed with approximately 20 ng of FLP protein. This represents a molar ratio of 8 FLP protein monomers per recombination site, and falls between lanes 6 and 7. These products are defined at the 60 minute time point of the recombination reaction. Identical results were observed at 20 minutes or 120 minutes (data not shown). All stated molar ratios are approximate, and based on FLP protein concentrations calculated by the Warburg and Christianson method of protein determination.

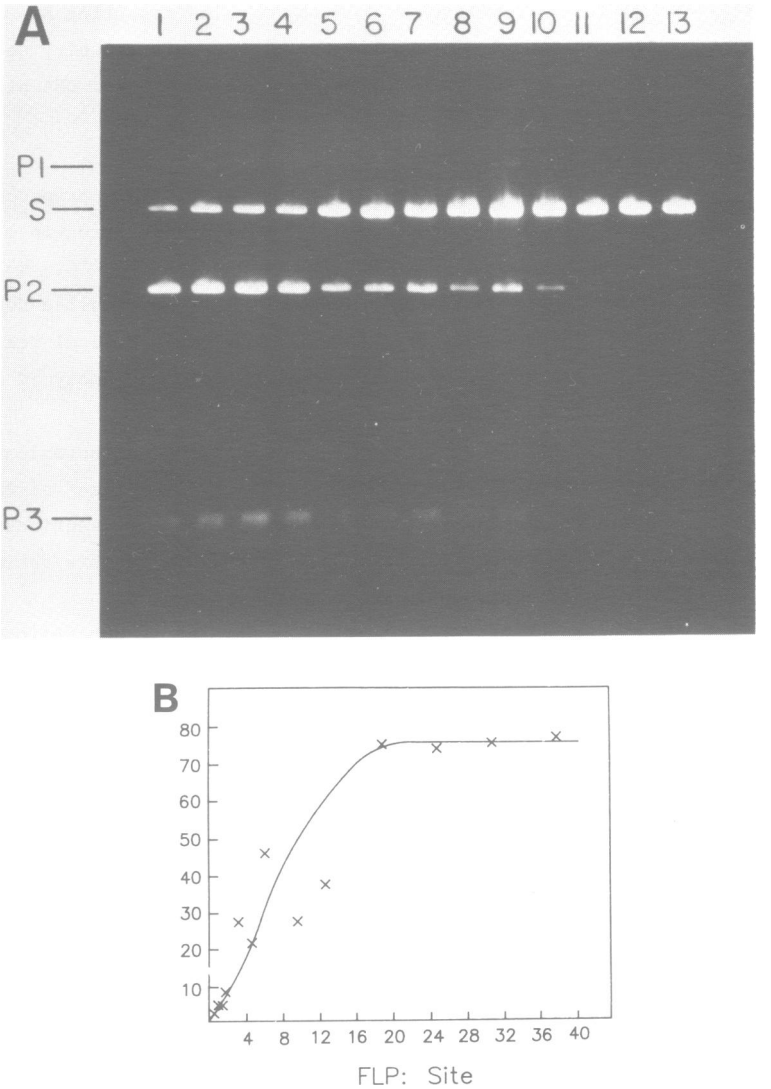



FIG. 4. Titration of FLP protein in recombination reaction. Reactions and dilutions were performed under standard conditions, as defined in Methods. P1, P2 are Intermolecular products; P2, P3 are Intramolecular products. A) The amount of substrate pMMC3 DNA (200 ng) is constant in all reactions. The amount of FLP protein in each lane is 1) 116 ng; 2) 96 ng; 3) 77 ng; 4) 58 ng; 5) 38 ng; 6) 29 ng; 7) 19 ng; 8) 14 ng; 9) 9.6 ng; 10) 6.1 ng; 11) 4.1 ng; 12) 3.1 ng; 13) 1.5 ng. B) The amount of DNA in each band was quantitated by densitometric scanning of the gel. The percent of DNA in bands P1, P2, and P3 was summed and plotted as "percent product" versus the molar ratio of FLP protein monomers to recombination site.

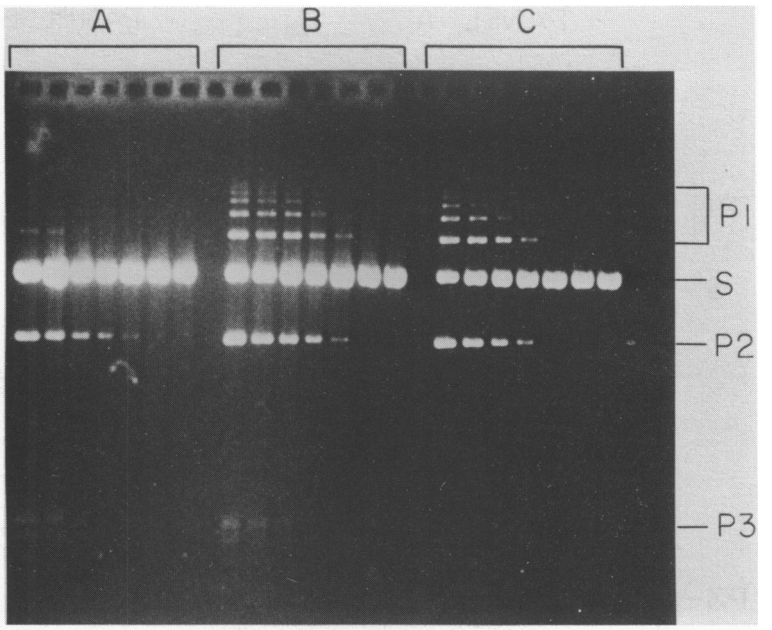



FIG. 5. Effect of E. coli extract on FLP protein promoted recombination. Reactions and dilutions were performed as described in Methods. P1, P2 are Intermolecular products; P2, P3 are Intramolecular products. FLP protein is titrated in 2-fold serial dilutions from 116 ng in lane 1 to 1.45 ng in lane 7 of each panel. Panel A: No E. coli extract added. Panel B: 1.0 μg of E. coli protein added to each reaction. Panel C: 4.0 μg of E. coli protein added to each reaction.

Product Distribution

The relative efficiency of inter- versus intramolecular recombination events promoted by purified FLP protein was examined. Results are presented in Figure 5. Each set of reactions represent seven dilutions of FLP protein, from 116 ng (a 39:1 molar ratio of FLP protein to recombination site) to 1.45 ng (0.5:1 molar ratio of FLP protein to recombination site). In lanes 1-7, purified FLP protein is the only protein added to the reaction. Recombination is primarily intramolecular; most of the products result from deletion events, even at the highest concentration of FLP protein used. Only a small fraction of recombination events result in products made larger than the original substrate molecule by intermolecular recombination. This does not change even when the concentration of FLP protein added is increased 15-fold (see Figure 6B). We did not observe a relationship between efficient intermolecular recombination and high FLP protein concentrations, or inhibition of

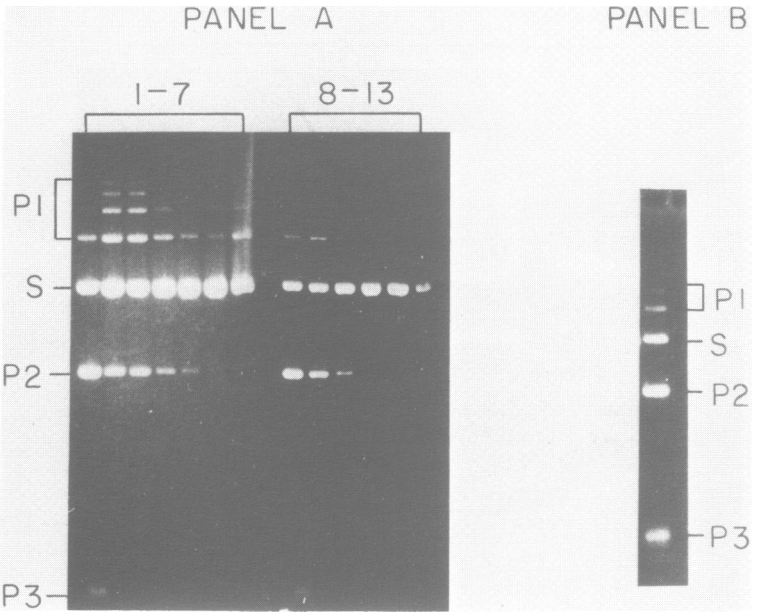



FIG. 6. Inhibition of recombination by an E. coli extract. Reactions were performed as described in Methods. P1, P2 are Intermolecular products; P2, P3 are Intramolecular products. Panel A: A constant amount of FLP protein (116 ng) was added to all reactions. Lanes 1-7: Amount of E. coli extract added is (1) none; (2) 0.5 μg; (3) 1.0 μg; (4) 2.0 μg; (5) 4.0 μg; (6) 10 μg; (7) 20 μg. Lanes 8-13: As in lanes 1-7, except that E. coli extract was preincubated with substrate DNA in reaction mixture for 40 minutes at 30°C prior to the addition of FLP protein. Panel B: 1.87 μg FLP protein. No E. coli extract was added.

recombination at still higher FLP protein concentrations, as previously reported (8). We investigated whether these properties were caused by other components of the FLP preparations used in earlier studies. In the next set of reactions, lanes 8-14, 1.0 μg of an E. coli extract was added to each reaction. This was prepared from cells containing no FLP protein with a procedure which mimics the first two steps of the FLP purification, as described in Methods. Prominent products now appear which result from intermolecular recombination. At this concentration of E. coli extract, the reaction is divided between intra- and intermolecular recombination. In the last set of reactions, lanes 15-21, four times as much E. coli extract was added to each reaction. In this case products result primarily from the intermolecular reaction and the intramolecular reaction appears to be completely inhibited. In all three reaction sets the minimal recombination

reaction occurs at the same dilution of FLP protein. Thus the overall extent of the reaction is not affected by the extract, although the product distribution is profoundly affected. The effects of this extract are protease sensitive, unaffected by RNAse treatment, and are not duplicated by bovine serum albumin (data not shown).

The possibility that inhibition of recombination could also be mediated by proteins in the E. coli extract is examined in Figure 6. Figure 6A, lanes 1-7, shows the effect of increased concentrations of E. coli extract while the concentration of FLP protein is held constant. As the concentration of E. coli extract is increased, product distribution is shifted toward an intermolecular reaction, and then is progressively inhibited. When the reaction is preincubated with E. coli extract for 40 minutes at 30°C prior to adding FLP protein (lanes 8-13), the inhibitory effect is shifted to lower amounts of E. coli extract relative to reactions performed without preincubation. The inhibitory effect appears to be due to one or more proteins in the E. coli extract. Previous reports showed inhibition at a molar ratio of 600 FLP monomers:1 recombination site (8). Even at this ratio (panel B) recombination products are still evident in the absence of E coli extract. We conclude that pure FLP protein promotes intramolecular recombination efficiently, even when a large excess of the protein is present. Several of the properties previously attributed to this protein with respect to product distribution and inhibition can be explained by other E. coli proteins which are present in FLP protein fractions purified with Biorex 70.

## DISCUSSION

We report here a procedure for purification of the FLP protein of the yeast 2 micron plasmid. The protocol produces 2 mg of FLP protein at 85% purity. The preparation is free of nuclease contamination. Recombination proceeds at FLP protein concentrations corresponding to as little as one mole of FLP protein per mole recombination site. The purified protein is stable at 0°C for several weeks, or at -70°C for several months when stored as described. The ability to obtain highly purified FLP protein using this procedure will make possible a variety of mechanistic studies in this recombination system.

The final step in the purification procedure involves affinity chromatography on a site-specific DNA Sepharose column. To design the affinity ligand, we drew on current information about the FLP protein recombination site. The sequence employed is a synthetic DNA fragment

containing multiple direct tandem repeats of the FLP protein binding site, and results in a useful degree of specific binding without cleavage. The specificity of this step makes it likely that a high proportion of the protein in this preparation is active. The concentration of NaCl required to elute FLP protein from this ligand is approximately 0.15 M greater than that which elutes FLP protein from calf thymus DNA-Sepharose (C. Gates, unpublished results).

Affinity chromatography using immobilized ligands has potential for application to a variety of proteins which bind DNA specifically. While this work was in progress, similar procedures were published for the purifications of Nuclear Factor I (31) and the transcription factor Spl (32).

Extensive purification has modified our understanding of several properties of the FLP protein. Most importantly, large changes in the relative efficiency of inter- versus intramolecular recombination, and of inhibitory effects, are not a function of FLP protein concentration, as previously described (8). With pure FLP protein alone, deletion products are predominant, even when a large excess of FLP protein is present. Inhibition at high FLP protein concentrations is not observed. Based on alterations in product distribution which occurred at several stages of the purification procedure, we suspected that the effects were due to contaminating species in the E. coli extract. We have confirmed this by fractionating an E. coli extract in a manner parallel to the FLP purification procedure, and mimicking the effects on product formation with this extract. Under these conditions, intermolecular reactions readily occur even at low FLP concentrations. This is consistent with our previous observation that two glycerol gradient fractions of FLP protein promoted different distributions between intermolecular and intramolecular products (7). We surmise that, in this case as well, product distribution reflected a different concentration or distribution of E. coli protein in the two fractions, rather than a property of the FLP protein itself. Four DNA binding proteins from E. coli have recently been purified in this laboratory which exhibit the effects on FLP protein-promoted recombination described in this report (R. C. Bruckner, unpublished). Two of these correspond to two histone-like proteins which together were previously designated the H protein of E. coli (35). Whereas the effects of the E. coli extract are difficult to interpret and may not be mechanistically significant, the results provide an explanation for a number of inconsistent observations in the literature.

ACKNOWLEDGEMENTS

We thank A. Smallwood for helpful advice and assistance, and B. Schutte for purifying T4 ligase. This work was supported by National Institutes of Health grants GM32335 and GM37835, and by a Basil O'Connor Starter Research Grant (5-451) from the March of Dimes. M.M.C. is supported by N.I.H. Research Career Development Award AI00599. L.M.L. was supported by N.I.H. Predoctoral Training Grant GM07215.

*To whom correspondence should be addressed

REFERENCES

1. Broach,J.R. (1981) in The Molecular Biology of the Yeast Saccharomyes: Life Cycle and Inheritance, Strathern,J.N., Jones,E.W. and Broach,J.R. Eds., pp. 445-470, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
2. Hollenberg,C.P., Degelmann,A., Kustermann-Kuhn,B. and Royer,H.D. (1976) Proc. Natl. Acad. Sci. **73**, 2072-2076.
3. Futcher,A.B. (1986) J. Theor. Biol. **119**, 197-204.
4. Volkert,F.C. and Broach,J.R. (1986) Cell **46**, 541-550.
5. Cox,M.M. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 4223-4227.
6. Vetter,D., Andrews,B.J., Roberts-Beatty,L. and Sadowski,P.D. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 7284-7288.
7. Meyer-Leon,L., Senecoff,J.F., Bruckner,R.C., and Cox,M.M. (1984) Cold Spring Harbor Symp. Quant. Biol. **49**, 797-804.
8. Gronostajski,R.M. and Sadowski,P.D. (1985) J. Biol. Chem. **260**, 12328-12335.
9. Senecoff,J.F., Bruckner,R.C. and Cox,M.M. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 7270-7274.
10. Gronostajski,R.M. and Sadowski,P.D. (1985) J. Biol. Chem. **260**, 12320-12327.
11. Andrews,B.J., Proteau,G., Beatty,L.G. and Sadowski,P.D. (1985) Cell **40**, 795-803.
12. Bruckner,R.C. and Cox,M.M. (1986) J. Biol. Chem. **261**, 11798-11807.
13. Gronostajski,R.M. and Sadowski,P.D. (1985) Mol. Cell. Biol. **5**, 3274-3279.
14. Sadowski,P.D., Lee,D.D., Andrews,B.J., Babineau,D., Beatty,L., Morse,M.J., Proteau,G. and Vetter,D. (1984) Cold Spring Harbor Symp. Quant. Biol. **49**, 789-796.
15. Babineau,D., Vetter,D., Andrews,B.J., Gronostajski,R.M., Proteau,G.A., Beatty,L.G. and Sadowski,P.D. (1985) J. Biol. Chem. **260**, 12313-12319.
16. Andrews,B.J., Beatty,L.G. and Sadowski,P.D. (1987) J. Mol. Biol. **193**, 345
17. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
18. Pharmacia P-L Biochemicals (1984) Analects **12**, 1-2.
19. Davis,R.W., Botstein,D. and Roth,J.R. (1980) Advanced Bacterial Genetics, pp. 196-197, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Gouy,M. and Gautier,C. (1982) Nucl. Acids. Res. **10**, 7055-7074.
21. Gribskov,M., Devereux,J. and Burgess,R.R. (1984) Nucl. Acids Res. **12**, 539-549.

22. Schoner,B.E., Hsiung,H.M., Belagaje,R.M., Mayne,N.G. and Schoner,R.G. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 5403-5407.
23. Stormo,G.D. (1986) in Maximizing Gene Expression, Reznikoff,W. and Gold,L. Eds., pp. 201-20, Butterworths, Boston.
24. Masui, Y., Mizuno,T. and Inouye,M. (1984) Biotechnology **2**. 81-85.
25. Morrical,S.W., Lee,J., and Cox,M.M. (1986) Biochemistry **25**, 1482-1494.
26. Arndt-Jovin,D., Jovin,T.M., Bahr,W., Frischauf,A.M., and Marquardt,M. (1975) Eur. J. Biochem. **54**, 411-418.
27. Kohn,J. and Wilchek,M. (1982) Biochem. Biophys. Res. Comm. **107**, 878-884.
28. Warburg,O. and Christianson,W. (1942) Biochem. Z. **310**, 384.
29. Bradford,M.M. (1976) Anal. Biochem. **72**, 248-254.
30. Cox,M.M., McEntee,K. and Lehman,I.R. (1981) J. Biol. Chem. **256**, 4676-4678.
31. Rosenfeld,P.J. and Kelly,T.J. (1986) J. Biol. Chem. **261**, 1398-1408.
32. Kadonaga,J.T. and Tjian,R. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 5889-5893.
33. Laemmli,U.K. (1970) Nature **227**, 680-685.
34. Weber,K. and Osborn,M. (1969) J.Biol. Chem. **244**, 4406-4412.
35. Hubscher, V., Lutz, H., and Kornberg, A. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 5097-5101.