Nucleic Acids Research, 1998, Vol. 26, No. 5 **1329–1336**

Selection of novel, specific single-stranded DNA sequences by Flp, a duplex-specific DNA binding protein

Xu-Dong Zhu and Paul D. Sadowski*

Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Received September 30, 1997; Revised and Accepted January 7, 1998

DDBJ/EMBL/GenBank accession no. M23380

ABSTRACT

Flp is a member of the integrase family of site-specific recombinases. Flp is known to be a double-stranded (ds)DNA binding protein that binds sequence specifically to the 13 bp binding elements in the FRT site (Flp recognition target). We subjected a random pool of oligonucleotides to the *in vitro* binding site selection method and have unexpectedly recovered a series of single-stranded oligonucleotides to which Flp binds with high affinity. These single-stranded oligonucleotides differ in sequence from the duplex FRT site. The minimal length of the oligonucleotides which is active is 29 nt. This single strand-specific DNA binding activity is located in the same C-terminal 32 kDa domain of Flp in which the site-specific dsDNA binding activity resides. Competition studies suggest that the apparent affinity of Flp for single-stranded oligonucleotide is somewhat less than for a complete duplex FRT site but greater than for a single duplex 13 bp binding element. We have also shown that Cre, another member of the integrase family of site-specific recombinases, also exhibits single-stranded DNA binding similar to that of Flp.

INTRODUCTION

The selfish 2µm plasmid DNA of *Saccharomyces cerevisiae* encodes the Flp protein, which is a member of the integrase family of conservative site-specific recombinases (1–3). Flp mediates a site-specific recombination event between two 599 bp inverted repeats on the 2µm plasmid (4) and functions to amplify the plasmid in the cell (5–7).

The Flp recognition target site (FRT) is confined to 48 bp within the 599 bp inverted repeats of the 2µm plasmid (8,9). The FRT site includes an 8 bp core region flanked by three Flp binding sites that are 13 bp in length (symmetry elements **a**, **b** and **c**; Fig. 1). Biochemical studies with purified Flp and DNA substrates show that recombination initiates by Flp binding specifically to the FRT site (10–13) and inducing a bend in the DNA (14,15). Protein–protein interactions between Flp molecules bring two FRT sites together to form a synaptic complex (16). Site-specific cleavage results from a nucleophilic attack of Tyr343 on two phosphodiester bonds at the margins of the core region (17). A phosphotyrosine bond is formed between the 3′-phosphoryl end of the nicked DNA and the Flp protein (18,19). Strand exchange and ligation occur when the free 5′-OH of the nick from the partner DNA acts as a nucleophile to attack the phosphotyrosine bond and regenerate a new phosphodiester bond. The Holliday junction resulting from the first strand exchange (20–23) is resolved through a second round of strand cleavage and ligation (24,25).

DNase I and chemical footprinting studies showed that all three symmetry elements and the core region of the duplex FRT site are protected by Flp (10,13,26). Flp binds first to symmetry element **b**, then to element **a** and finally to element **c**, resulting in formation of complexes I, II and III respectively (11,27). Partial proteolysis of Flp yields an N-terminal domain of 13 kDa (P13) and a C-terminal domain of 32 kDa (P32) (28,29). The C-terminal peptide P32 retains the site-specific double-stranded (ds)DNA binding property of Flp, whereas the N-terminal peptide P13 has non-specific DNA binding activity and stimulates binding of P32.

In this paper we report the discovery of a novel sequence-specific single-stranded (ss)DNA binding activity of Flp. Using an *in vitro* binding site selection method we unexpectedly recovered several ssDNA sequences to which Flp binds in a sequence-specific manner with high affinity. These sequences differ from those contained in either strand of the FRT site. The region of Flp responsible for the sequence-specific ssDNA binding activity is located in the C-terminal domain (P32) of Flp, where the site-specific dsDNA binding activity also resides. Cre, another member of the integrase family, also exhibits this sequence-specific ssDNA binding activity.

Although we have no experimental evidence to indicate that this single strand-specific DNA binding activity plays a role in Flp-mediated recombination, our discovery that Flp recombinase, a sequence-specific duplex DNA binding protein, also binds ssDNA in a specific manner may have novel applications in biotechnology and medicine.

MATERIALS AND METHODS

Protein preparations

All proteins expressed from *Escherichia coli* cells were purified to >90% purity. Wild-type Flp protein was either a Sephacryl S300 fraction or a Mono S fraction, purified as described previously (28). Histidine-tagged fusion peptides of P13 and P32 were purified on Ni–NTA columns as described by A.Shaikh (30) and were generous gifts of A.Shaikh. Expression and purification

*To whom correspondence should be addressed. Tel: +1 416 978 6061; Fax: +1 416 978 6885; Email: p.sadowski@utoronto.ca

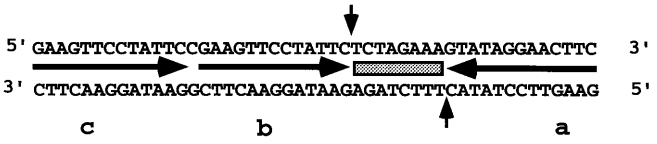

Figure 1. The sequence of the Flp recognition target site (FRT). The FRT site is 48 bp in length and includes three 13 bp symmetry elements (**a**, **b** and **c**, indicated by horizontal arrows) and an 8 bp core region (rectangle). Each symmetry element is a Flp binding element. Elements **a** and **b** are inverted repeats which differ in 1 bp. Element **c** is a direct duplication of element **b** in both sequence and orientation. Flp cleaves DNA at the junctions of the symmetry elements and the core sequence as indicated by the two vertical arrows.

of the Cre protein was essentially as described by Shaikh and Sadowski (31) with some modifications. The Cre protein was purified by heparin–agarose and Mono S chromatography and was a generous gift of J.Walker. The concentration of protein was estimated using the method of Bradford (32).

Synthetic substrates

Oligonucleotides were synthesized by the Biotechnology Service Centre, University of Toronto. Duplex FRT sites were prepared by annealing the appropriate oligonucleotides in 5 mM $MgCl_2$, 100 mM NaCl as described previously (33,34). A 10-fold excess of unlabeled over labeled oligonucleotides was used in the annealing reactions. Where appropriate, the 5′-termini of oligonucleotides were labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase (New England Biolabs). The sequences of the oligonucleotides used in this study are listed in Table 1.

Table 1. Synthetic oligonucleotides used in this study

Name	Sequence (5′ to 3′)
Flp-10-T (56 mer)	*TCGACTCGAGTCGACATCG*GCAGTCTCTAGTTGAGGC ggatcctgcagaattcgcg
Flp-10-B (56 mer)	cgcgaattctgcaggatccGCCTCAACTAGAGACTGC *CGATGTCGACTCGAGTCGA*
SS-1 (18 mer)	GCAGTCTCTAGTTGAGGC
SS-3 (37 mer)	GCAGTCTCTAGTTGAGGCggatcctgcagaattcgcg
SS-2 (37 mer)	*TCGACTCGAGTCGACATCG*GCAGTCTCTAGTTGAGGC
SS-5 (30 mer)	*GAGTCGACATCG*GCAGTCTCTAGTTGAGGC
SS-9 (31 mer)	*TCGAGTCGACATCG*GCAGTCTCTAGTTGAGGC
SS-10 (32 mer)	*CTCGAGTCGACATCG*GCAGTCTCTAGTTGAGGC
SS-11 (33 mer)	*CTCGAGTCGACATCG*GCAGTCTCTAGTTGAGGC
SS-8 (34 mer)	*ACTCGAGTCGACATCG*GCAGTCTCTAGTTGAGGC
SS-12 (27 mer)	*ACTCGAGTCGACATCG*GCAGTCTCTAG
SS-13 (29 mer)	*ACTCGAGTCGACATCG*GCAGTCTCTAGTT
SS-14 (31 mer)	*ACTCGAGTCGACATCG*GCAGTCTCTAGTTGA
FRTb-T (56 mer)	*TCGACTCGAGTCGACATCG*GAAGTTCCTATTCGAGGC ggatcctgcagaattcgcg
FRTb-B (56 mer)	cgcgaattctgcaggatccGCCTCGAATAGGAACTTC *CGATGTCGACTCGAGTCGA*
FRTab-T (55 mer)	TCACTGTCACGTGAAGTTCCTATTCTCTAGAAAGTAT ACGAACTTCGACCTGAGC
FRTab-B (55 mer)	GCTCAGGTCGAAGTTCCTATACTTTCTAGAGAATAG CAACTTCACGTGACAGTGA

The left-hand flanking sequence is indicated in italics and the right-hand flanking sequence is shown in lower case. The underlined nucleotides represent sequences that were randomized during site selection. The shaded nucleotides indicate the sequence of the symmetry elements; the core sequence of the FRT site is double underlined. Oligonucleotides FRTb-T and FRTb-B were hybridized to give duplex FRTb. Oligonucleotides FRTab-T and FRTab-B were annealed to form duplex FRTab. The concentrations of the oligonucleotides were measured spectrophotometrically using the relationship 1 OD_{260} = 30 µg oligonucleotide as specified by Sambrook *et al.* (37).

In vitro transcription and translation assays

In vitro expression of Flp was achieved using the Promega TNT® T7 coupled reticulocyte lysate system. Approximately 2 µg supercoiled plasmid pLD3 DNA (28), which contains the *Flp* gene under control of the T7 promoter, was used as template with 40 µCi [^{35}S]methionine (1000 Ci/mmol) in 50 µl reticulocyte lysate. As a control the *in vitro* transcription and translation assays were also carried out using plasmid pAR3038, which has the same transcriptional promoter and terminator as pLD3 but lacks the *Flp* gene (35). The translation mixtures were used immediately for DNA binding and DNA recombination assays.

DNA binding assays

A radiolabeled DNA probe (0.02 pmol) was incubated with protein in 30 µl binding buffer (50 mM Tris–HCl, pH 7.5, 33 mM NaCl, 1 mM EDTA) containing 2 µg sonicated and denatured calf thymus DNA (CT DNA). When competition experiments were performed an additional excess of specific DNA competitor was also added to the reactions. After 25 min incubation at room temperature the reaction mixture was loaded onto an 8% non-denaturing polyacrylamide gel. The amount of Flp protein used in each reaction was ~5.8 pmol, whereas the amount of Flp peptides (P13 and P32) and the Cre protein used was varied as stated in the figure legends. The results of the competition experiments were quantitated using a Molecular Dynamics PhosphorImager.

When DNA binding assays were performed using the *in vitro* transcription–translation mixtures 20 or 30 µl translation mixture were combined with 0.02 pmol radiolabeled DNA probe in 40 µl binding buffer (50 mM Tris–HCl, pH 7.5, 33 mM NaCl, 1 mM EDTA) containing 2 µg CT DNA and 240 pmol non-specific oligonucleotide ssDNA [SAD3499 (43mer), TAGGGCCCATGG-GGCAATTTGATATATTATGTAAAACACCACC].

In vitro binding site selection

In vitro binding site selection was carried out essentially as described by Blackwell and Weintraub (36). An oligonucleotide mixture which contained an 18 nt random region flanked on either side by 19 nt of non-random sequence [5′-TCGACTCGAGTCGA-CATCG(N_{18})GGATCCTGCAGAATTCGCG-3′] was rendered double-stranded by incubation with the Klenow fragment of DNA polymerase I. The reaction contained a primer (P1) complementary to the 3′ non-random sequence (5′-CGCGAATTCTGCAG-GATCC-3′), four deoxynucleoside triphosphates (0.67 mM dCTP, 0.67 mM dGTP, 0.67 mM dTTP, 1.33 µM dATP) and 100 µCi [α-^{32}P]dATP. The dsDNA was purified by electrophoresis on a 12% non-denaturing polyacrylamide gel and eluted from the gel using the ‘crush and soak’ method (37). Approximately 0.02 pmol eluted DNA were then incubated with 5.8 pmol Flp in 30 µl binding buffer (50 mM Tris–HCl, pH 7.5, 33 mM NaCl, 1 mM EDTA) containing 4 µg poly(dI·dC). After 25 min incubation at room temperature the reaction mixture was run on an 8% non-denaturing polyacrylamide gel at 4°C for 4.5 h at 220 V. The gel was dried and exposed to X-ray film overnight at room temperature with an intensifying screen. In each round of selection DNA was eluted from the protein–DNA complexes by excising the band from the dried gel and soaking it overnight at 37°C in elution buffer (0.5 M ammonium acetate, 10 mM $MgCl_2$, 1 mM EDTA, 0.1% SDS). After soaking, the buffer containing the eluted DNA

was removed from the gel slice, extracted twice with phenol/chloroform and then precipitated with ethanol using 20 µg glycogen (Boerhinger) as carrier. PCR amplification of the eluted DNA was performed using the Perkin-Elmer Cetus amplification kit in the presence of primers P1 (see above) and P2. Primer P2 (5′-CGATGTCGACTCGAGTCGA-3′) is complementary to the 5′ non-random sequence of the template. The amplification used a 30 s denaturation step at 94°C, 1 min annealing at 55°C and 1 min extension at 72°C for 25 cycles. The amplified oligonucleotides were extracted twice with phenol/chloroform, precipitated with ethanol and purified on a non-denaturing polyacrylamide gel. The dsDNA fragment was then radiolabeled by T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP and used as a substrate for binding by Flp in the next round of selection. As a negative control a blank slice of the gel was excised after each round and treated in the same manner as the band containing bound protein and DNA. No PCR product was amplified from the control, indicating that there was no template contamination. The selection procedure was repeated for a total of four rounds. After the final round of selection the enriched DNA pool was digested with *Sal*I and *Eco*RI and cloned into *Sal*I/*Eco*RI-digested pUC19. DNA from individual clones was isolated using a Wizard Minipreps DNA purification system from Promega and then sequenced using a sequencing kit obtained from US Biochemicals.

RESULTS

Recovery of Flp binding sites by *in vitro* binding site selection

We originally designed experiments to select DNA sequences to which certain mutant Flp proteins might bind and we used wild-type Flp as a control. Our selection strategy used an electrophoretic mobility shift assay to enrich for DNA sequences that were bound by Flp (36).

A pool of DNA sequences which contained a randomized 18 bp internal region flanked by two constant regions was incubated with Flp protein and the protein–DNA complexes were resolved from the free DNA on a non-denaturing polyacrylamide gel. The bound DNA was eluted, amplified by PCR and used in subsequent rounds of selection. After four rounds of selection a large amount of complex was formed from the selected pool of DNA that was subjected to the gel mobility shift assays in the presence of Flp (Fig. 2A, lanes 2 and 3).

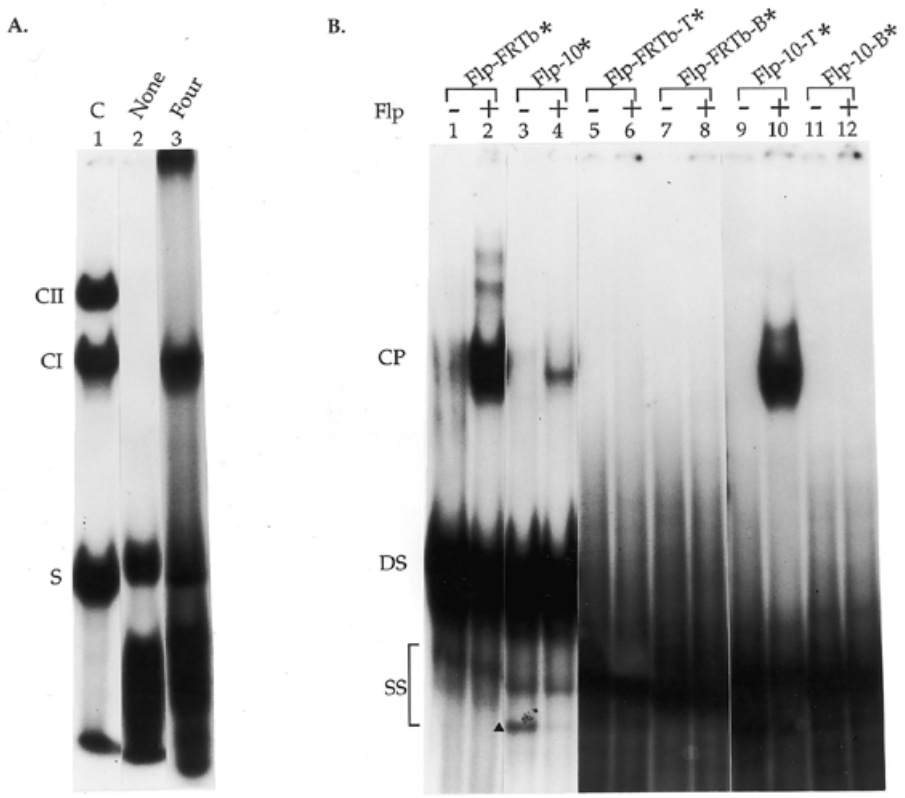

Figure 2. (**A**) Binding of Flp to the mixture of oligonucleotides before and after four rounds of binding site selection. The binding reactions were done as described in Materials and Methods. Lane 1 was a control, where the DNA substrate was double-stranded FRT site (Flp-FRTab) containing two symmetry elements flanking the core sequence. Binding of Flp to Flp-FRTab generates two complexes (CI and CII), representing Flp molecules binding to one element and two symmetry elements respectively. Lanes 2 and 3 contained Flp and oligonucleotides with random sequences. Lane 2 shows the first binding reaction in which random oligonucleotides had undergone no rounds of selection (none), whereas lane 3 was the binding reaction in which the mixture of oligonucleotides had gone through four rounds of selection and PCR enrichment (four). The DNA oligonucleotides from the complex in lane 3 were eluted, amplified by PCR and cloned into the pUC19 vector. S refers to free labeled oligonucleotides. (**B**) Analysis of binding of Flp to selected sequences. Synthetic oligonucleotides corresponding to Flp-10-T, Flp-10-B, FRTb-T and FRTb-B (Table 1) were used to assemble selected double-stranded sequences. DNA sequences were 5′-end-labeled with ^{32}P (asterisks). The binding assays were done in the presence of Flp and DNA substrates as described in Materials and Methods. DNA substrates and the presence or absence of Flp are indicated above each lane. In this and subsequent figures CP represents protein–DNA complexes, DS refers to labeled dsDNA substrates, SS represents labeled ssDNA oligonucleotides.

F-10	TCGACTCGAGTCGACATCG	**GCAGTCTCTAGTTGAGGC**
F-20	TCGACTCGAGTCGACATCG	**GCAGTCTCGAGTAGGCGT**
F-21	TCGACTCGAGTCGACATCG	**GCAGTCTCAAGTCAGCGT**
F-28	TCGACTCGAGTCGACATCG	**GCAGTCTCAAATCGACGG**
F-29	TCGACTCGAGTCGACATCG	**GCAGTCTCTAGTCAGCGT**
F-31	TCGACTCGAGTCGACATCG	**GCAGTCTCTAGGAGACGT**
F-37	TCGACTCGAGTCGACATCG	**GCAGTCTCGATTCGAAAC**
F-38	TCGACTCGAGTCGACATCG	**GCAGTCTCTATTCGACGT**
F-42	TCGACTCGAGTCGACATCG	**GCAGTCTCTAGTCGGTGG**
F-47	TCGACTCGAGTCGACATCG	**GCAGTCTGGAGTTGGTGT**
F-33	TCGACTCGAGTCGACATCG	**GCAGTCTTAAGTCGGTGT**
F-9	TCGACTCGAGTCGACATCG	**GCAGTCTTAAGTCGGTGT**
F-24	TCGACTCGAGTCGACATCG	**GCAGTCTTAAGTCGGTGT**
F-23	TCGACTCGAGTCGACATCG	**GCAGTCTTTAGTCGATGC**
F-16	TCGACTCGAGTCGACATCG	**GCAGTCTTTAGTCGATGC**
F-43	TCGACTCGAGTCGACATCG	**GCAGTCTTTAGTCGACGC**
F-44	TCGACTCGAGTCGACATCG	**GCAGTCTTTAGTCGACGC**
F-39	TCGACTCGAGTCGACATCG	**GCAGTCTTTAGTCGGTGT**
F-12	TCGACTCGAGTCGACATCG	**GCATGTTTGAGTCGGCGC**
F-14	TCGACTCGAGTCGACATCG	**GCATGTTTGAGTCGGCGC**
F-17	TCGACTCGAGTCGACATCG	**GCATGTTTGAGTCGGGCT**
F-2	TCGACTCGAGTCGACATCG	**GCATGTTTGGGTCGGTGT**
F-11	TCGACTCGAGTCGACATCG	**GCATGTTTGGGTCGGTGT**
F-3	TCGACTCGAGTCGACATCG	**GCATGTTTGGGTCGTCGT**
F-32	TCGACTCGAGTCGACATCG	**GCATGTTTGAGTCGCTGC**
F-35	TCGACTCGAGTCGACATCG	**GCATGTTTGAGTCGGGCG**
F-34	TCGACTCGAGTCGACATCG	**GCATGTTTGAGTCGCTGC**
F-13	TCGACTCGAGTCGACATCG	**AACTACCACCCGGTTCAT**
F-1	TCGACTCGAGTCGACATCG	**AACTACCACGCGGGTCGG**
F-25	TCGACTCGAGTCGACATCG	**AACTACCACGCGGGTCGG**
F-46	TCGACTCGAGTCGACATCG	**AACTACCACGCGGGTCGG**

Figure 3. Summary of DNA sequences selected by Flp. Thirty one site-selected oligonucleotides were sequenced after cloning into pUC19, revealing the existence of three families of binding sequences. Bases that were randomized during the binding site selection are highlighted in bold. Sequences that are not in bold represent the left flanking sequences that were invariant during site selection. The invariant right flanking sequences are not shown. The first 7–9 conserved nucleotides of each family are underlined. Studies in this paper were done using the oligonucleotide Flp-10-T, which corresponds to the sequence of F-10 (bold, left column). Oligonucleotides corresponding to F-14, F-16 and F-19 (bold, left column) also showed top strand-specific, single strand-specific DNA binding by Flp.

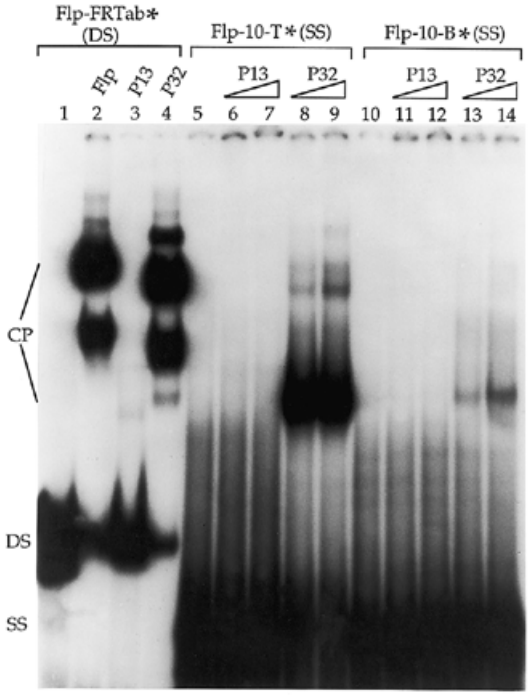

Figure 4. Binding of Flp peptides to ssDNA. ssDNA (SS, Flp-10-T and Flp-10B), as well as duplex DNA Flp-FRTab (DS), was 5′-end-labeled with ^{32}P. Lanes 1, 5 and 10 contained only the labeled DNA substrates. Flp and Flp peptides were added as indicated above the lanes. Triangles indicate increasing amounts of Flp peptides. The amounts of Flp peptides were as follows: lanes 6 and 11, 77 pmol P13; lanes 3, 7 and 12, 231 pmol P13; lanes 8 and 13, 31.25 pmol P32; lanes 4, 9 and 14, 62.5 pmol P32.

The enriched DNA sequences were cloned into pUC19 and inserts from 31 independent clones were sequenced. As shown in Figure 3, sequences selected by wild-type Flp fell into three families. To our surprise none of the sequences resembled the Flp binding symmetry elements of the FRT site (Fig. 1). Although each family had its own consensus sequence, all three families showed a striking common feature, i.e. the first 7 nt of the randomized internal region within each family were absolutely conserved. This seemed to imply that the left-hand flanking sequence may have played a role in binding site selection.

We also did mobility shift assays on several of the selected sequences from the first two families. These assays showed that Flp failed to bind strongly to any of the selected sequences when the selected sequences were dsDNA, whether excised from a plasmid or formed from annealed synthetic oligonucleotides (see Fig. 2B and data not shown).

Flp binds to ssDNA sequence specifically

While wild-type Flp bound efficiently to the duplex FRTb site (Fig. 2B, lane 2), it was not able to form a large amount of protein–DNA complex when assayed with one of the duplex selected sequences (Fig. 2B, lane 4). However, the small amount of binding complex seen in lane 4 of Figure 2B seemed to have arisen from Flp binding to a DNA species which migrated faster than the dsDNA substrates, since the intensity of these bands diminished upon incubation with Flp (triangles, lane 3 and 4). These DNA bands likely correspond to ssDNA resulting from some denaturation of dsDNA substrates. Therefore, we decided to investigate whether Flp would bind to ssDNA of the selected sequence. We selected Flp-10 sequence and made DNA oligonucleotides corresponding to its top and bottom strands. Each of the oligonucleotides was ^{32}P 5′-end-labeled and incubated with Flp. As shown in Figure 2B, Flp bound only to the top strand of the Flp-10 sequence but not to the complementary bottom strand (lane 10 versus lane 12). As a control the first 13 bp in the random core region of the selected Flp-10 sequence were replaced by the 13 bp Flp binding symmetry element **b** and this new sequence was named Flp-FRTb. As expected, Flp bound to dsDNA Flp-FRTb readily (Fig. 2B, lane 2). However, when the oligonucleotide corresponding to either the top or the bottom strand of the Flp-FRTb sequence was incubated with Flp protein Flp failed to bind to either of these single-stranded substrates (Fig. 2B, lanes 6 and 8). These results suggest that Flp has a sequence-specific ssDNA binding activity. We obtained identical results with three other selected sequences, Flp-14, 16 and 29 (data not shown).

Competition experiments were also performed to confirm that Flp was recognizing ssDNA. While addition of an excess of oligonucleotide Flp-10-B, which was complementary to the Flp-10-T sequence, abolished binding of Flp to Flp-10-T, addition of a large excess of a non-complementary ssDNA

(SAD3499) had no effect on binding of Flp to the single-stranded Flp-10-T sequence (data not shown). The fact that addition of Flp-10-B abolished binding of Flp to Flp-10-T was likely due to annealing of the complementary strand to the labeled single-stranded probe to form dsDNA. Since this duplex sequence is different from the duplex Flp binding symmetry elements, Flp could not bind to it. These results further support the idea that Flp binds to ssDNA in a sequence-specific manner.

The single strand-specific DNA binding activity is attributable to Flp

Although binding site selection was done using highly purified Flp protein (>90% pure), it was formally possible that the activity was due to a contaminant in the Flp preparation. However, several preparations of purified wild-type Flp protein and mutant Flp proteins, all of which have normal binding activity to the duplex FRT site, bound to the Flp-10-T sequence but not to Flp-10-B (data not shown), suggesting that this single strand-specific DNA binding activity is not peculiar to one protein preparation.

We next tested the ability of Flp peptides to bind to ssDNA. The Flp protein can be partially proteolyzed into an N-terminal domain of 13 kDa (P13) and a C-terminal domain of 32 kDa (P32) (28,29). The N-terminal P13 peptide does not display site-specific DNA binding activity but has been demonstrated to stimulate binding of P32 to the Flp binding symmetry element. On the other hand, the C-terminal P32 peptide retains the site-specific DNA binding property of Flp. Highly purified, recombinant His-tagged peptides P13 and P32 were each incubated separately with Flp-10-T and Flp-10-B. Only the P32 peptide showed sequence-specific ssDNA binding activity (Fig. 4, lanes 8 and 9). Weak binding was also detected when P32 was incubated with Flp-10-B (Fig. 4, lanes 13 and 14). This was likely due to the non-specific ssDNA binding activity of P32, since binding of P32 to Flp-10-B was abolished when excess non-specific ssDNA competitor (SAD3499) was added to the binding reactions (data not shown), whereas this non-specific ssDNA competitor had no effect on P32 binding to Flp-10-T (data not shown). These results strongly support our claim that Flp has an intrinsic sequence-specific ssDNA binding activity. Furthermore, the single strand-specific DNA binding activity resides in the 32 kDa C-terminal portion of the Flp protein (P32).

To further substantiate the evidence that the ssDNA binding activity is specifically associated with Flp, the wild-type Flp protein was synthesized *in vitro* using an *in vitro* transcription and translation system in the presence of [^{35}S]methionine. SDS–PAGE revealed that production of Flp was dependent on addition of a plasmid containing a *Flp* gene (data not shown). No Flp was made in the absence of plasmid DNA or with a plasmid (pAR3038) that lacked the *Flp* gene (data not shown). Like Flp generated *in vivo*, the Flp protein made *in vitro* exhibited site-specific dsDNA binding activity to the FRT sequence (Fig. 5A,

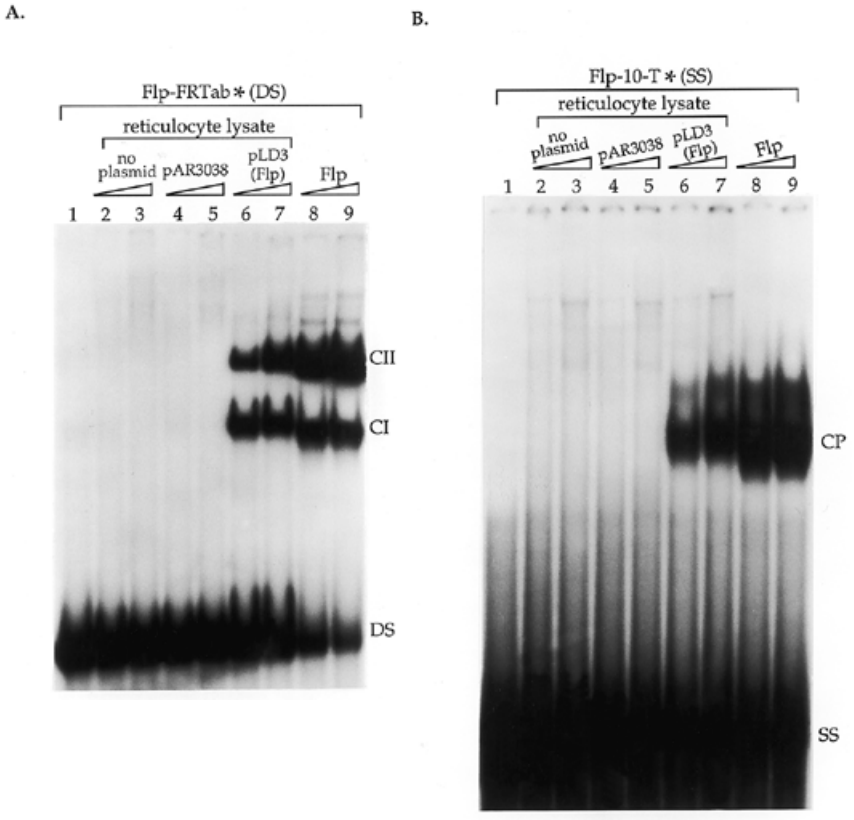

Figure 5. Analysis of DNA binding activity of the wild-type Flp protein generated *in vitro* in a reticulocyte lysate. (**A**) Gel mobility shift assays using duplex DNA as substrates. dsDNA Flp-FRTab containing two symmetry elements flanking the core sequence was 5′-end-labeled with ^{32}P. Lane 1 contained only the labeled DNA substrate. The contents in each reaction were indicated above the lanes. Triangles indicate increasing amounts of translation products or purified Flp. Lanes 2, 4 and 6, 10 µl translation products; lanes 3, 5 and 7, 20 µl translation products; lanes 8 and 9, 2.9 and 5.8 pmol purified Flp. CI and CII were as defined in Figure 2. (**B**) Analysis of binding of ssDNA by translation products and purified Flp. ssDNA Flp-10-T was 5′-end-labeled with ^{32}P. The contents of each reaction are shown above the lanes. The amounts of translation products and purified Flp in lanes 1–9 were identical to those in (A).

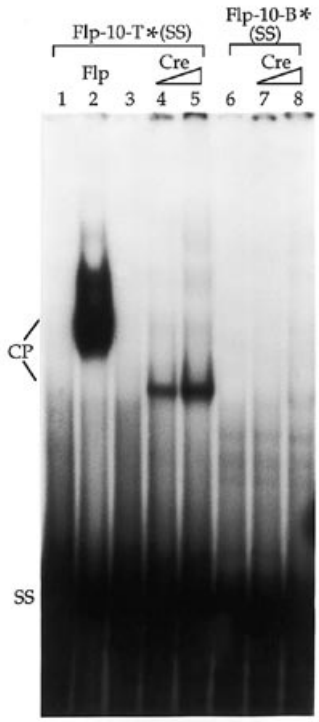

Figure 6. Analysis of binding of ssDNA by the Cre protein. Single-stranded oligonucleotides Flp-10-T and Flp-10-B were 5′-end-labeled with ^{32}P. Substrates and proteins are shown above the lanes. The amounts of the Cre protein were as follows: lanes 4 and 7, 78 pmol; lanes 5 and 8, 156 pmol.

lanes 6 and 7) as well as DNA recombination activity (data not shown). These results show that the reticulocyte lysate was generating enzymatically active Flp *in vitro*. As shown in Figure 5B, when incubated with Flp-10-T ssDNA the Flp protein generated *in vitro* was also able to bind to it readily (lanes 6 and 7). However, Flp generated *in vitro* showed no binding activity to the Flp-10-B sequence (data not shown). Since the Flp protein produced *in vitro* is free of contaminating proteins derived from *E.coli* cells, we conclude that this sequence-specific ssDNA binding activity is intrinsic to Flp protein.

Cre protein also exhibits single strand-specific DNA binding activity

Because Flp is a member of the integrase family of site-specific recombinases, we wished to know whether another member of the family also bound to ssDNA. Since we knew that the Cre recombinase, another member of the integrase family, was able to bind to the duplex FRT site, albeit with reduced affinity (30), we tested whether Cre might also bind to the Flp-10-T oligonucleotide. Highly purified Cre protein of phage P1 was incubated separately with Flp-10-T and Flp-10-B. As shown in Figure 6, Cre, like Flp, also bound only to Flp-10-T and not Flp-10-B (lanes 4 and 5 versus lanes 7 and 8). Furthermore, binding of Cre to the Flp-10-T sequence was not affected by addition of a large excess of non-specific DNA SAD3499 (data not shown). This suggests that Cre also exhibits a sequence-specific ssDNA binding activity. To confirm this it will be necessary to carry out random site selection with Cre recombinase.

The length of sequence required for binding of Flp to ssDNA

We have shown that Flp binds single-stranded oligonucleotide Flp-10-T in a sequence-specific manner. To define the minimal length of sequence necessary for binding of Flp we carried out gel mobility shift assays using synthetic oligonucleotides derived from Flp-10-T and the results are summarized in Figure 7. The sequence of Flp-10-T ssDNA can be divided into three parts (Fig. 7): (i) the internal region of 18 nt which was randomized during site selection; (ii) the left flanking sequence (Lf, 19 nt); (iii) the right flanking sequence (Rf, 19 nt). Both the left and right flanking sequences were invariant. Flp failed to bind to single-stranded oligonucleotides containing either the internal region alone (SS-1) or the internal region and the Rf region (SS-3, Fig. 7). However, Flp bound to an oligonucleotide containing the internal region and the Lf region (SS-2) as efficiently as to Flp-10-T. This suggested that a portion of the Flp recognition site lay within the sequence of SS-2 (Fig. 7). To define the left-hand boundary of the Flp binding site we made a series of oligonucleotides with deletions of the 5′-region in the left flanking sequence of SS-2. As shown in Figure 7, SS-5, SS-9, SS-10, SS-11 and SS-8 contained deletions of 7, 6, 5, 4 and 3 nt respectively from the 5′-end of SS-2. Mobility shift assays showed that only SS-8 (3 nt deletion) was able to bind Flp as efficiently as SS-2 (data not shown and Fig. 7). Although SS-5 was bound weakly by Flp, SS-9, SS-10 and SS-11, which contained smaller deletions in the Lf region than did SS-5, were not bound by Flp at all. Therefore, we conclude that SS-8 contained the left-hand boundary of the ssDNA binding site of Flp.

We also examined the right-hand boundary of the binding site of Flp. As shown in Figure 7, oligonucleotides SS-12, SS-13 and SS-14 contained deletions of 7, 5 and 3 nt respectively from the 3′-end of SS-8. Mobility shift assays showed that Flp failed to bind to SS-12 but it bound to both SS-13 and SS-14 as efficiently as it bound to SS-8 (data not shown and Fig. 7), suggesting that the right boundary of the binding site of Flp was contained within SS-13. Since SS-13 shares the same left boundary as SS-8, this oligonucleotide defines the sequence of 29 nt required for binding of Flp to the ssDNA.

Relative binding of Flp to ssDNA versus to dsDNA

It is known that Flp binds specifically to the duplex 13 bp symmetry elements of the FRT site. Therefore, it was of interest to compare the binding of Flp to ssDNA with that to dsDNA. To perform this comparison a series of competition experiments was carried out.

First, duplex DNA containing either one Flp binding symmetry element (Flp-FRTb) or two Flp binding symmetry elements in inverted orientation (Flp-FRTab) were used as competitors added directly to a binding reaction containing labeled ssDNA (Flp-10-T). For comparison, the unlabeled ssDNA (Flp-10-T) was also used as competitor. The specific complex formed by binding of Flp to ssDNA (Flp-10-T) was competed efficiently with Flp-FRTab (Table 2, line 2). Less than 1 pmol Flp-FRTab was needed to give 50% inhibition of complex formation. However, ~50 pmol Flp-FRTb were required to inhibit formation of the complex by 50% (Table 2, line 2). When Flp-10-T itself served as competitor DNA ~5 pmol were required to inhibit complex formation by 50% (Table 2, line 2). These results indicate that Flp binds ~10-fold more efficiently to the ssDNA (Flp-10-T) than to dsDNA containing one Flp binding symmetry element. The dsDNA (Flp-FRTab) was the most effective

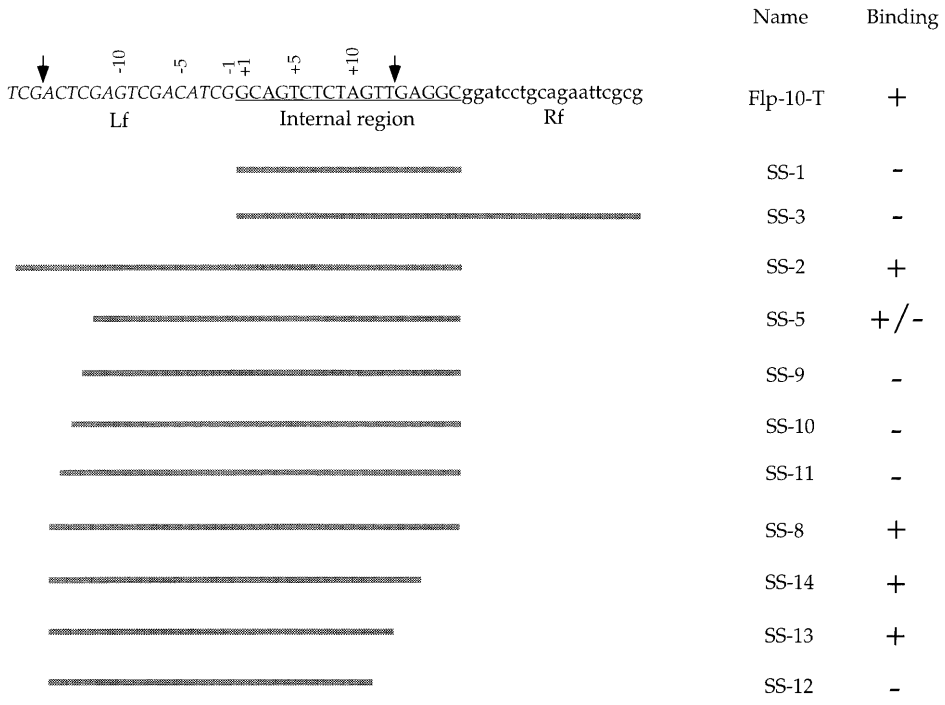

Figure 7. Analysis of the minimal sequence of the ssDNA (Flp-10-T) for binding of Flp. The sequence of Flp-10-T is diagramed above. The core sequence randomized prior to site selection is underlined, separating the invariant left and right flanking sequences (Lf in italics and Rf in lower case respectively). Stippled bars represent the regions of the Flp-10-T sequence that are present in the oligonucleotides. The relative ability of Flp to bind to these single-stranded oligonucleotides is summarized on the right: +, binding approximately equal to that shown by Flp-10-T; –, no binding; +/–, reduced binding. Vertical arrows above the sequence of Flp-10-T denote the minimal sequence required for binding of Flp to the ssDNA.

competitor, probably due to cooperative interactions between two Flp molecules bound to the two 13 bp symmetry elements (11).

Similar relative affinities were obtained by examining the ability of single-stranded Flp-10-T and duplex Flp-FRTb to inhibit formation of complexes formed by Flp and Flp-FRTb and the ability of excess Flp-10-T, Flp-FRTb and Flp-FRTab to inhibit formation of complex II by Flp and Flp-FRTab (Table 2, lines 1 and 3). These results confirm that Flp binds to Flp-10-T more efficiently than to Flp-FRTb.

Table 2. Approximate amounts of competitor required to give 50% inhibition of complex formation by Flp and the DNA probe

DNA probe[a]	DNA competitor (pmol)		
	Flp-FRTab	Flp-10-T	Flp-FRTb
Flp-FRTab	1 (CII)[b]	5–10 (CII)[b]	50 (CII)[b]
Flp-10-T	<1[c]	5[c]	50[c]
Flp-FRTb	ND	1[c]	10[c]

[a]0.02 pmol ^{32}P-labeled DNA was used in competition experiments.
[b]The amount of competitor needed to reduce the amount of complex II (CII) by 50%.
[c]The amount of competitor needed to reduce the amount of complex I by 50%.
ND, not determined.

DISCUSSION

In this paper we report the use of the *in vitro* binding site selection method to discover novel ssDNA sequences to which Flp binds in a sequence-specific manner. Flp is known to bind specifically to the duplex FRT site. We have shown that the ability of Flp to bind the ssDNA sequence of Flp-10-T was superior to its ability to bind duplex DNA containing only Flp binding element **b**. This may be one of the reasons for selection of the ssDNA sequences when the experiment was designed originally to select dsDNA sequences containing one Flp binding element. Another may have been that the 18 bp random region was insufficiently long to accommodate two 13 bp binding sites that might mediate cooperative binding between two Flp molecules. Competition studies also indicated that Flp possesses a single DNA binding site for both duplex and ssDNA (data not shown).

The minimum length of the sequence (SS-13) of ssDNA that is required for binding of Flp is 29 nt. This sequence includes 16 nt from the invariant left flanking sequence and 13 nt from the internal sequence that was randomized prior to site selection. The fact that the left flanking sequence serves as part of the ssDNA binding site for Flp may be the reason that all the sequences selected by Flp share an exact junction with the left flanking sequence. Missing nucleoside interference experiments revealed that the Flp protein makes extensive interactions with all the nucleotides in the sequence of SS-13 (X.-D.Zhu, unpublished data). Although these interactions are essential for DNA binding, they may not be base specific. It is possible that the sequence of bases specifies a novel secondary structure of the oligonucleotide that is in turn recognized by Flp.

It is known that Flp binds to the duplex FRT site in a sequence-specific manner. However, neither the invariant left flanking sequence of Flp-10-T nor the selected internal region resembles the sequence of the FRT site. Indeed, Flp does not bind to the single strands of the FRT site.

What does the ssDNA binding activity of Flp mean? There are at least three possibilities. First, the ssDNA binding activity of Flp may play a role in the recombination reaction itself. However, our attempts to perturb recombination of duplex FRT sites *in vitro* by incubation with the single-stranded oligonucleotide (Flp-10-T) resulted simply in inhibition of the reaction. We presume that this was due to interference with binding of Flp to the FRT site, since no specific interference with a subsequent step (e.g. cleavage, ligation or strand exchange) could be detected (data not shown). Second, the activity may regulate some other biological process in the cell, a process that is unrelated to recombination. For example, it is possible that Flp could bind to RNA and somehow regulate expression of either 2µm or cellular genes. A preliminary test of the binding of Flp to an *in vitro* transcript containing the entire Flp-10-T sequence gave equivocal results (data not shown). Third, the activity may have been an artifact of the powerful *in vitro* selection imposed during the random site selection but have no biological relevance.

Further experiments will be necessary to dissect the relative roles of base specificity versus structure specificity in contributing to binding affinity. However, it is possible that some motifs in Flp that recognize the duplex FRT site also recognize some structures in the single-stranded oligonucleotide that mimic the duplex FRT site. Such a structure may be dictated by both sequence-specific and structure-specific components. There are a few examples in the literature of proteins that recognize ssDNA sequence specifically, including single-stranded telomere end binding proteins and mammalian Purα proteins (38–40). It will be interesting to see if the repertoire of such proteins can be expanded by deliberately searching for such activities among proteins that bind duplex DNA site specifically. It is also possible that single-stranded oligonucleotides to which proteins bind sequence specifically could have novel therapeutic applications.

ACKNOWLEDGEMENTS

This work is supported by a grant from the Medical Research Council of Canada. X.-D.Z. is the holder of an Ontario Graduate Scholarship. We thank Doug Kuntz, John Walker and Arkady Shaikh for preparing proteins. We also thank Andy Becker, Rick Collins, Helena Friesen, Barbara Funnell, Arkady Shaikh and John Walker for critical and insightful comments.

REFERENCES

1 Sadowski,P.D. (1993) *FASEB J.*, **7**, 760–767.
2 Sadowski,P.D. (1995) *Prog. Nucleic Acid Res. Mol. Biol.*, **51**, 53–91.
3 Landy,A. (1993) *Curr. Opin. Genet. Dev.*, **3**, 699–707.
4 McLeod,M., Craft,S. and Broach,J.R. (1986) *Mol. Cell. Biol.*, **6**, 3357–3367.
5 Futcher,A.B. (1986) *J. Theor. Biol.*, **119**, 197–204.
6 Volkert,F.C. and Broach,J.R. (1986) *Cell*, **46**, 541–554.
7 Reynolds,A.E., Murray,A.W. and Szostak,J.W. (1987) *Mol. Cell. Biol.*, **7**, 3566–3573.
8 Broach,J.R., Guarascio,V.R. and Jayaram,M. (1982) *Cell*, **29**, 227–234.
9 Meyer-Leon,L., Senecoff,J.F., Bruckner,R.C. and Cox,M.M. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 797–804.
10 Andrews,B.J., Proteau,G.A., Beatty,L.G. and Sadowski,P.D. (1985) *Cell*, **40**, 795–803.
11 Andrews,B.J., Beatty,L.G. and Sadowski,P.D. (1987) *J. Mol. Biol.*, **193**, 345–358.
12 Senecoff,J.F., Bruckner,R.C. and Cox,M.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7270–7274.
13 Bruckner,R.C. and Cox,M.M. (1986) *J. Biol. Chem.*, **261**, 11798–11807.
14 Schwartz,C.J.E. and Sadowski,P.D. (1989) *J. Mol. Biol.*, **205**, 647–658.
15 Schwartz,C.J.E. and Sadowski,P.D. (1990) *J. Mol. Biol.*, **216**, 289–298.
16 Amin,A.A., Beatty,L.G. and Sadowski,P.D. (1990) *J. Mol. Biol.*, **214**, 55–72.
17 Jayaram,M. (1994) *Trends Biochem Sci.*, **19**, 78–82.
18 Gronostajski,R.M. and Sadowski,P.D. (1985) *Mol. Cell. Biol.*, **5**, 3274–3279.
19 Evans,B.R., Chen,J.-W., Parsons,R.L., Bauer,T.K., Teplow,D.B. and Jayaram,M. (1990) *J. Biol. Chem.*, **265**, 18504–18510.
20 Holliday,R. (1964) *Genet. Res.*, **5**, 282–304.
21 Meyer-Leon,L., Huang,L.-C., Umlauf,S.W., Cox,M.M. and Inman,R.B. (1988) *Mol. Cell. Biol.*, **8**, 3784–3796.
22 Meyer-Leon,L., Inman,R.B. and Cox,M.M. (1990) *Mol. Cell. Biol.*, **10**, 235–242.
23 Jayaram,M., Crain,K.L., Parsons,R.L. and Harshey,R.M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7902–7906.
24 Dixon,J.E. and Sadowski,P.D. (1993) *J. Mol. Biol.*, **234**, 522–533.
25 Dixon,J.E. and Sadowski,P.D. (1994) *J. Mol. Biol.*, **243**, 199–207.
26 Panigrahi,G.B., Beatty,L.B. and Sadowski,P.D. (1992) *Nucleic Acids Res.*, **20**, 5927–5935.
27 Beatty,L.G. and Sadowski,P.D. (1988) *J. Mol. Biol.*, **204**, 283–294.
28 Pan,H., Clary,D. and Sadowski,P.D. (1991) *J. Biol. Chem.*, **266**, 11347–11354.
29 Pan,G. and Sadowski,P.D. (1993) *J. Biol. Chem.*, **268**, 22546–22551.
30 Shaikh,A.C. (1997) MSc thesis, University of Toronto.
31 Shaikh,A.C. and Sadowski,P.D. (1997) *J. Biol. Chem.*, **272**, 5695–5702.
32 Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
33 Zhu,X.-D., Pan,G., Luetke,K. and Sadowski,P.D. (1995) *J. Biol. Chem.*, **270**, 11646–11653.
34 Zhu,X.-D. and Sadowski,P.D. (1995) *J. Biol. Chem.*, **270**, 23044–23054.
35 Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendroff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
36 Blackwell,T.K. and Weintraub,H. (1990) *Science*, **250**, 1104,1110.
37 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
38 Virta-Pearlman,V., Morris,D.K. and Lundblad,V. (1996) *Genes Dev.*, **10**, 3094–3104.
39 Bergemann,A.D. and Johnson,E.M. (1992) *Mol. Cell. Biol.*, **12**, 1257–1265.
40 Bergemann,A.D., Ma,Z.-W. and Johnson,E.M. (1992) *Mol. Cell. Biol.*, **12**, 5673–5682.