JOURNAL OF BACTERIOLOGY, Apr. 1993, p. 2221–2228
0021-9193/93/082221-08$02.00/0

Vol. 175, No. 8

Characterization of the *Escherichia coli* F Factor *traY* Gene Product and Its Binding Sites

WILLIAM C. NELSON,[1] BRADLEY S. MORTON,[1]† ELAINE E. LAHUE,[2]‡ AND STEVEN W. MATSON[1,2]*

Department of Biology[1] and Curriculum of Genetics,[2] University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Received 16 October 1992/Accepted 10 February 1993

The *traY* gene product (TraYp) from the *Escherichia coli* F factor has previously been purified and shown to bind a DNA fragment containing the F plasmid *oriT* region (E. E. Lahue and S. W. Matson, J. Bacteriol. 172:1385–1391, 1990). To determine the precise nucleotide sequence bound by TraYp, DNase I footprinting was performed. The TraYp-binding site is near, but not coincident with, the site that is nicked to initiate conjugative DNA transfer. In addition, a second TraYp binding site, which is coincident with the mRNA start site at the *traYI* promoter, is described. The K_d for each binding site was determined by a gel mobility shift assay. TraYp exhibits a fivefold higher affinity for the *oriT* binding site compared with the *traYI* promoter binding site. Hydrodynamic studies were performed to show that TraYp is a monomer in solution under the conditions used in DNA binding assays. Early genetic experiments implicated the *traY* gene product in the site- and strand-specific endonuclease activity that nicks at *oriT* (R. Everett and N. Willetts, J. Mol. Biol. 136:129–150, 1980; S. McIntire and N. Willetts, Mol. Gen. Genet. 178:165–172, 1980). As this activity has recently been ascribed to helicase I, it was of interest to see whether TraYp had any effect on this reaction. Addition of TraYp to nicking reactions catalyzed by helicase I showed no effect on the rate or efficiency of *oriT* nicking. Roles for TraYp in conjugative DNA transfer and a possible mode of binding to DNA are discussed.

The conjugative transfer of DNA from one bacterial cell to another, mediated by transmissible plasmids, is a major route for genetic exchange among bacteria. After close cell-cell contact has been established, a single strand of DNA is transferred from the donor to the recipient cell. Subsequent stabilization of the transferred DNA in the recipient cell, either by recombination with the chromosome or as a plasmid, completes the transfer of genetic traits (for reviews see references 15, 40, and 41). The F plasmid conjugation system in *Escherichia coli* is a paradigm for this type of genetic transfer. Recently, this DNA transfer system has begun to yield biochemical information regarding the reactions involved in the transfer of DNA from donor to recipient.

The F plasmid encodes most, if not all, of the functions known to be required for conjugative DNA transfer in a 33-kbp segment on F called the *tra* region. This segment includes the genes encoding the structural elements of the F pilus, genes encoding surface exclusion factors, the *cis*-acting origin of transfer (*oriT*), which is nicked prior to strand transfer, and the genes encoding proteins which interact with *oriT* prior to and during conjugative transfer. At least four genes, *traY*, *traI*, *traM*, and *traD*, fall into the latter category. The precise roles played by the TraM and TraD proteins remain to be determined. These proteins are thought to be involved in the physical transport of DNA across the bacterial membranes and, possibly, are members of the putative transmembrane signalling pathway that triggers the nicking reaction (26, 29). The *traI* gene encodes DNA helicase I (1), which has recently been shown to catalyze the site- and strand-specific nicking reaction required to initiate conjugative DNA transfer (24, 27). In addition, it has been suggested that the helicase activity may provide the motive force that drives strand transfer (29). The nicking reaction had previously been assumed to be catalyzed by the products of the *traY* and *traZ* genes (10). It now seems likely that *traZ* is a functional domain of *traI* (36) and that helicase I is involved in both nicking and unwinding the F plasmid from *oriT* (37). The *traY* gene has been cloned, its protein product has been overexpressed and purified, and TraY protein (TraYp) has been shown to bind specifically to the *oriT* region from F (19). Unexpectedly, the recent reconstitution of site and strand-specific nicking at the F plasmid *oriT* did not require the *traY* gene product (24, 27). Thus the role of TraYp in F plasmid DNA metabolism is unclear.

In this report, the biochemical characterization of the F plasmid TraY protein has been extended in an effort to further define the role of this protein in conjugative DNA transfer. Hydrodynamic studies indicate that TraYp exists as a monomer in solution under the conditions used in DNA binding assays. The precise binding site on F *oriT* has been determined by DNase I footprinting and shown to be near, but not coincident with, the site that is nicked in *oriT*. We have also found a second site in the F plasmid *tra* region to which TraY protein binds specifically. This site is located near the promoter for the *traYI* operon. Possible roles for TraYp in F plasmid-related DNA metabolism are discussed.

MATERIALS AND METHODS

Enzymes. Restriction endonucleases were obtained from New England Biolabs, GIBCO/BRL Life Technologies, Inc., and U.S. Biochemicals. Reaction conditions were those suggested by the supplier. DNA polymerase I (large fragment) was obtained from U.S. Biochemicals. Bacteriophage T4 polynucleotide kinase was obtained from New

* Corresponding author.
† Present address: Department of Cell Biology, Glaxo Research Institute, Glaxo Inc., Research Triangle Park, NC 27709.
‡ Present address: T-Cell Sciences, Cambridge, MA 02139.

England Biolabs. Bacterial alkaline phosphatase was obtained from IBI, Inc. Helicase I was purified as previously described (24).

DNAs and nucleotides. Plasmids were grown in *E. coli* HB101 and purified by the alkaline lysis procedure (2) followed by banding in CsCl-ethidium bromide gradients. pBSoriT contains the *oriT* sequence from the F plasmid beginning at the *Bgl*II site and extending to the *Sal*I site. The construction of this plasmid has been described previously (24). pED851 is a pBR322 derivative containing the entire F *tra* sequence (10). Plasmid pBSPYI contains a portion of the F plasmid *tra* sequence extending from the carboxy-terminal end of the *traJ* gene through the *traYI* promoter and into the amino-terminal end of the *traY* gene. This region from the F plasmid was amplified by the polymerase chain reaction using pED851 as a template and primers 5′-TTGAATTCTC TACAATAAAAAGTTT-3′ and 5′-TTTTAAGCTTACCAA ATCTTTTCAAT-3′. The discrete 189-bp DNA fragment observed on a polyacrylamide gel was electroeluted, digested with *Eco*RI and *Hin*dIII and cloned into pBluescript II KS(+) (Stratagene) that had been digested with *Eco*RI and *Hin*dIII. The construction was confirmed by dideoxy chain termination DNA sequencing using Sequenase (U.S. Biochemicals).

Nucleotides were from Pharmacia/P-L Biochemicals. [α-^{32}P]dCTP, [α-^{32}P]dATP, and [γ-^{32}P]ATP were obtained from Amersham, Inc. The oligonucleotides used to form the 46-bp TraYp-binding oligonucleotide had sequences 5′-GGAAAAATTAGTTTCTCTTACTCTCTTTATGATATTT AAAAAAGCGGT-3′ and 5′-GGACCGCTTTTTTAAATA TCATAAAGAGAGTAAGAGAAACTAATTTTT-3′. Annealing was performed at 75°C in a buffer containing 7 mM Tris-Cl (pH 7.5), 6.6 mM $MgCl_2$, 50 mM NaCl, and 1 mM dithiothreitol (DTT) for 5 min and was followed by slow cooling to room temperature.

Binding assays. The binding of TraY protein to specific DNA fragments was measured by gel mobility shift assays as described elsewhere (8). The binding assay reaction mixture (16 μl) contained 40 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 10-to-15% glycerol, 30 μg of bovine serum albumin (BSA) per ml, 50 mM NaCl, the indicated ^{32}P-DNA substrate, and the indicated amount of TraY protein. Incubation took place at 37°C for 10 min and was followed by the addition of a 4-μl solution containing 37.5% glycerol and 0.5% loading dyes. Samples were loaded directly onto a 5% acrylamide–0.167% bisacrylamide gel (8). Electrophoresis was performed at 10 V/cm at 25°C in a buffer containing 25 mM Tris, 200 mM glycine, and 1 mM EDTA for 60 to 90 min.

The substrates for gel mobility shift assays were the indicated DNA fragments labelled at the 3′ end in a reaction mixture that contained 20 mM Tris-HCl (pH 7.5), 6 mM $MgCl_2$, 6 mM DTT, 50 μg of BSA per ml, 50 mM NaCl, 50 μM dATP, 50 μM dGTP, 50 μM dTTP, 10 to 30 μCi of [α-^{32}P]dCTP, the DNA substrate, and *E. coli* DNA polymerase I (large fragment). Incubation lasted 30 min at 25°C and was followed by the addition of 50 μM dCTP and an additional incubation at 25°C for 10 min. After phenol-chloroform extraction, the unincorporated nucleotides were removed by filtration over Sepharose 6B-Cl in 10 mM Tris-HCl (pH 8)–1 mM EDTA–100 mM NaCl. Generally, the labelled DNA was used directly in gel retardation assays. When necessary, the DNA was concentrated by ethanol precipitation.

DNase I footprinting. DNase I footprinting was performed essentially as described elsewhere (7). TraY protein-DNA complexes were formed as described above in a 20-μl reaction mixture. The *oriT* DNA fragment used in *oriT* footprinting reaction mixtures was a 312-bp *Xba*I-*Rsa*I fragment isolated from pBSoriT. The DNA fragment was 5′ end labelled at the *Xba*I or *Rsa*I site by using phage T4 polynucleotide kinase and [γ-^{32}P]ATP as described previously (34). The *traYI* promoter DNA fragment used in footprinting reaction mixtures was a 189-bp *Eco*RI-*Hin*dIII fragment isolated from pBSPYI. The DNA fragment was 5′ or 3′ end labelled at the *Hin*dIII site as described above. DNase I-digested DNA fragments were resolved on a 6% polyacrylamide–8 M urea gel run in 100 mM Tris–100 mM borate–2 mM EDTA and visualized by autoradiography.

Construction of pETTY. Plasmid pED851 (10) was cleaved with *Bst*EII, and the 3′ ends were filled in by using DNA polymerase I (large fragment) (22). The resulting blunt-ended DNA fragments were resolved on a polyacrylamide gel, and the 1.7-kb DNA fragment containing *traY*, *traA*, *traL*, and part of *traE* was isolated and cloned into the *Sma*I site on M13mp18. Site-directed mutagenesis was performed as described by Kunkel et al. (18) to engineer an *Nde*I site at the *traY* start codon while changing the wild-type TTG start codon to ATG. The oligonucleotide used for mutagenesis had the sequence 5′-CCAAATCTTTTCATATGCACCTC CCGCTG-3′. Potential mutant clones were screened by isolating the replicative form I DNA and digesting with *Nde*I. The desired construction containing the *Nde*I site was purified and cleaved to completion with *Nde*I and *Nae*I. The 585-bp DNA fragment containing *traY* was resolved on a 6% polyacrylamide gel and isolated by electroelution. The expression vector pET12b (Novagen) was cut with *Bam*HI, the 3′ ends were filled in as described above, and then the plasmid was cut with *Nde*I. The *traY* gene was then cloned into the digested vector. This construction eliminates all of the wild-type sequence upstream of *traY* and places the gene behind the strong T7 ϕ10-s10 promoter and translation initiation region. The construction was confirmed by sequencing using Sequenase (U.S. Biochemicals).

Purification of TraYp. pETTY was transformed into *E. coli* HMS174(DE3) containing pLysE (Novagen). Four-liter cultures were grown to an optical density at 590 nm of 1.0, and TraYp expression was induced by adding isopropyl-β-D-thiogalactopyranoside to 0.4 mM. Cells were grown an additional 6 h before being harvested. The cells were suspended using 3 ml of a 50 mM Tris-Cl (pH 8.0)–2 mM EDTA–1 mM DTT–10% sucrose–150 mM NaCl solution per g, frozen in a dry-ice bath, and stored at −70°C until they were lysed. Cells were thawed by overnight incubation at 4°C, and lysis was initiated by adding Triton X-100 to 0.1%. Following a 30-min incubation on ice, the lysate was sonicated to reduce viscosity and cell debris was removed by centrifugation at 48,000 × *g* for 60 min in an SS34 rotor. Nucleic acids were precipitated by adding polyethylenimine (pH 6.8) (Aldrich) dropwise to a final concentration of 0.2% followed by centrifugation at 27,000 × *g* for 30 min. The supernatant was recovered, and solid ammonium sulfate was added to 50% saturation over a 30-min period. The precipitate was collected by centrifugation as before. Pellets were suspended in buffer B (20 mM KPO_4 [pH 6.8], 5 mM 2-mercaptoethanol, 0.5 mM EDTA, 10% glycerol) containing 50 mM KCl (fraction I). Fraction I was adjusted to the conductivity of buffer B containing 50 mM KCl by dialysis and loaded onto a phosphocellulose (Whatman) column (20 ml) equilibrated with buffer B containing 50 mM KCl. The column was washed with 5 column volumes of buffer B containing 100 mM KCl, and the protein was eluted with a linear 10-column-volume gradient from 100 to 800 mM KCl

in buffer B. Fractions were assayed for TraYp by gel mobility shift assay as described above. Active fractions, which eluted at approximately 500 mM KCl, were pooled and dialyzed against buffer A (50 mM Tris-Cl [pH 7.5], 5 mM 2-mercaptoethanol, 0.5 mM EDTA, 10% glycerol) containing 100 mM NaCl (fraction II). Fraction II was loaded onto a blue dextran agarose (Sigma) column (10 ml) equilibrated with buffer A containing 50 mM NaCl. The column was washed with 5 column volumes of buffer A containing 75 mM NaCl. TraYp was eluted by using a linear 15-column-volume gradient from 100 to 800 mM NaCl in buffer A. Fractions were assayed as described above, and active fractions were pooled (fraction III). TraYp was eluted at a salt concentration of approximately 190 mM. The pool was concentrated in a Centriprep-30 concentrator (Centricon), dialyzed into storage buffer (50 mM Tris-Cl [pH 7.5], 5 mM 2-mercaptoethanol, 0.5 mM EDTA, 50% glycerol, 100 mM NaCl), and stored at −70°C. The purity of the protein was judged to be >85%, as evidenced by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. Further purification on a Sephadex G-75 column was performed as previously described (19).

Nicking assay. Nicking assays were performed as described previously (24) with the following modification: the concentration of $MgCl_2$ in the reaction mixture was 4 mM.

Glycerol gradient sedimentation. The sedimentation coefficient of TraYp was determined by ultracentrifugation on continuous glycerol gradients. A sample of the protein was diluted fivefold (to achieve a final glycerol concentration of 10%) and layered on gradients of 15-to-35% (vol/vol) glycerol in buffer G (50 mM Tris-Cl [pH 7.5], 4 mM $MgCl_2$, 1 mM DTT, 50 mM NaCl). Ultracentrifugation was performed at 55,000 rpm for 30 h at 4°C in an SW55 rotor. Gradients were fractionated (125 μl per fraction) and assayed for TraYp activity by the standard gel mobility shift assay. Protein standards run in a parallel gradient were assayed by the method of Bradford (4). Protein dilution over the gradient was estimated to be approximately threefold.

Gel filtration. The Stokes radius of TraYp was determined by gel filtration. A 10.66-ml Sephadex G-75 Superfine (Sigma) column (0.7 by 27.7 cm) was poured and equilibrated at 4°C in buffer G containing 15% glycerol. A 30- to 100-μl sample of TraYp was filtered through the column and fractions (210 μl each) were assayed for TraYp activity by the standard gel mobility shift assay. Protein standards were assayed by the method of Bradford (4). Dilution of the protein across the column is estimated to be fivefold.

RESULTS

TraYp binding at *oriT*. Previous results indicated that purified TraYp specifically bound a DNA fragment containing the *oriT* region from the F plasmid (19). In this study, the precise location of the TraYp-binding site on *oriT* was determined by DNase I footprinting. Figure 1 shows the nuclease protection patterns obtained by using increasing concentrations of TraYp for both the top strand (Fig. 1B) and the bottom strand (Fig. 1A) of an *Xba*I-*Rsa*I DNA fragment from pBSoriT. The top strand is protected from nucleotide 208 to 240 (relative to the *Bgl*II site). The bottom strand is protected from nucleotide 204 to 234. It is impossible to determine whether the TraYp footprint extends to nucleotide 196 on the top strand and nucleotide 197 on the bottom strand by this method because of the insensitivity of the DNA in this region to DNase I. This analysis locates the TraYp-binding site approximately 64 bp downstream of the site nicked by helicase I to initiate conjugative DNA transfer (between nucleotides 140 and 141 on the bottom strand). In addition, the TraYp-binding site is located between the integration host factor (IHF)-binding sites recently demonstrated by Tsai et al. (38) (Fig. 2). The protected region (Fig. 3A) is 73% AT rich, a figure which corresponds well to the observations of Inamoto and Ohtsubo (13), who have previously shown that the TraYp-binding site at *oriT* in the R100 plasmid system (*sbyA*) is also highly AT rich. No sequence similarity between the F and R100 sequences is observed or expected, as *traY* is a plasmid-specific gene (25). Thus the F plasmid TraYp-binding site is near, but not coincident with, the nick site on *oriT*.

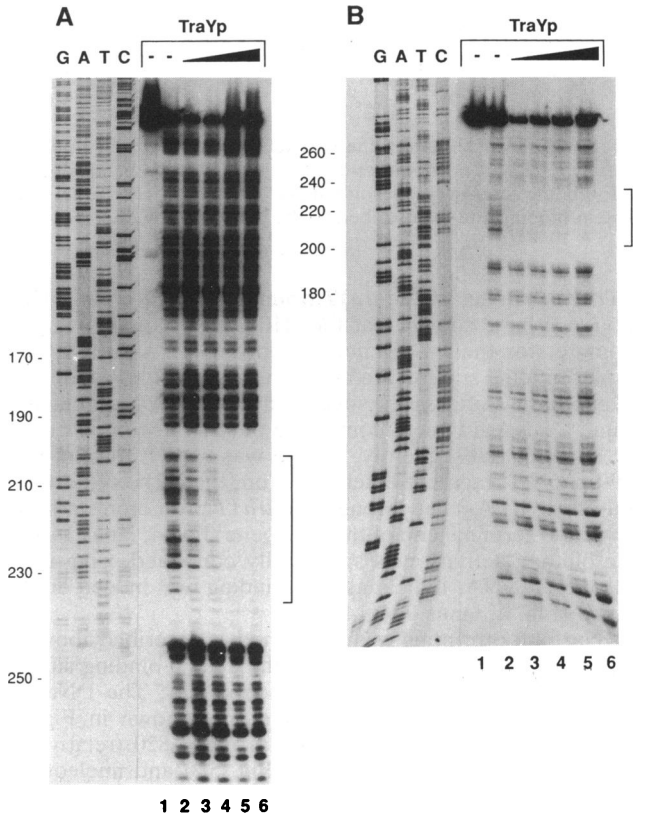


FIG. 1. DNase I footprint of TraY protein binding at *oriT*. (A) Bottom strand. Footprinting assays were performed as described in Materials and Methods using a 312-bp *Xba*I-*Rsa*I *oriT* DNA fragment that was 5′ end labelled at the *Rsa*I site. Lane 1, undigested DNA substrate; lane 2, DNA substrate digested with DNase I in the absence of TraYp; lanes 3 to 6, addition of 8.1, 16.3, 32.5, or 130 ng of TraYp, respectively, to the DNA substrate prior to DNase I digestion. Markers at the left are dideoxy chain termination sequencing reactions using the pBSoriT *Xba*I-*Rsa*I DNA fragment as the template and the oligonucleotide primer 5′-ACCACCCCTA CAAAACGG-3′. (B) Top strand. Footprinting assays were performed as described above by using a 312-bp *Xba*I-*Rsa*I *oriT* DNA fragment that was 5′ end labelled at the *Xba*I site as the substrate. Lanes 1 and 2 are as described above; lanes 3 to 6, addition of 16.3, 32.5, 65, or 130 ng of TraYp, respectively, to the DNA substrate prior to DNase I digestion. Markers at the left are as described above; oligonucleotide primer 5′-CTAGAACTAGTGGATCTC-3′ was used. Brackets denote the region protected from DNase I digestion by TraYp. The map position coordinates in base pairs relative to the *Bgl*II site at 66.7 kb on the F plasmid map (16) are indicated on the left in each panel.

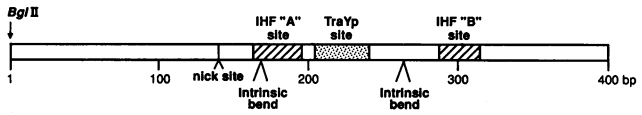
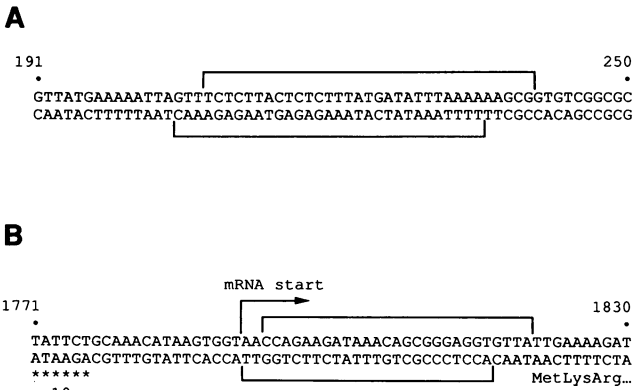


FIG. 2. Physical map of the first 400 bp of F *oriT*. The nick site was mapped by Thompson et al. (35). The intrinsic bends and the IHF-binding sites were described by Tsai et al. (38). The TraY protein-binding site is described in the text. See Fig. 1.

TraYp binding at the *traYI* promoter. Inamoto and Ohtsubo (13) have shown that the R100 TraYp binds near the promoter for the *traY* gene. This prompted a search for a similar binding site for TraYp on the F plasmid. Preliminary gel mobility shift assays suggested that TraYp specifically bound a DNA fragment containing the *traYI* promoter (data not shown). Subsequently, TraYp was shown to bind to a 189-bp DNA fragment extending from the carboxy-terminal end of the *traJ* gene, through the *traYI* promoter, and into the amino-terminal end of the *traY* gene (Fig. 4, lanes 1 to 8). The binding activity was specifically competed using unlabelled *oriT* DNA, indicating that binding was indeed due to TraYp (Fig. 4, lanes 9 to 14).

DNase I footprinting was performed as described above to determine the precise location of the TraYp-binding site on the fragment containing the *traYI* promoter. The DNase I protection patterns for both strands are shown in Fig. 5. TraYp protects nucleotides 1794 through 1820 (relative to the *Bgl*II site) on the top strand (Fig. 5B) and nucleotides 1792 through 1816 on the bottom strand (Fig. 5A). This places the TraYp-binding site precisely at the *traYI* operon mRNA start site and just upstream of the *traY* TTG start codon (Fig. 3B). Comparison of this site to the site bound at *oriT* reveals scant homology, and in contrast to the binding site at *oriT* and the sites described by Inamoto and Ohtsubo (13), this region is only 55% A+T. Thus, while it seems unlikely that TraYp simply binds highly AT-rich DNA sequences, the consensus sequence for the F sites is obscure.

TraYp is a monomer in solution. Hydrodynamic studies



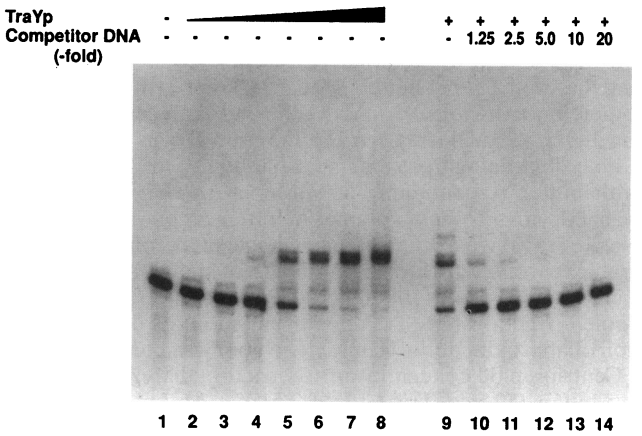


FIG. 4. TraYp binds to the *traYI* promoter region. Gel mobility shift assays were performed as described in Materials and Methods by using the 189-bp *Eco*RI-*Hin*dIII DNA fragment isolated from pBSPYI (1 ng per reaction mixture). The restriction fragment was 3′ end labelled at the *Hin*dIII site. Lane 1, no TraYp; lanes 2 to 8, 0.9, 1.8, 3.5, 7.2, 17.6, 35.2, and 88 ng of TraYp, respectively; lanes 9 to 14, 7 ng of TraYp and the indicated molar ratio of competitor DNA. The competitor DNA was the unlabelled 312-bp *Xba*I-*Rsa*I *oriT*-containing restriction fragment isolated from pBSoriT.

were conducted to determine whether TraYp was a monomer or multimer in solution under the conditions used for DNA binding assays. Purified TraYp sedimented in a 15-to-35% (vol/vol) continuous glycerol gradient with a sedimentation coefficient of 0.465S compared with markers sedimented in parallel gradients (Fig. 6A). Sephadex G-75 gel filtration indicated a Stokes radius of ~20 Å (2.0 nm) compared with standards (Fig. 6B). Experiments utilizing different initial concentrations of TraYp (from 6.4 to 19.8 μM) were performed to ensure that the results obtained were not artifacts due to the association and dissociation of a multimeric complex under our experimental conditions. The molecular weight of TraYp in solution was determined by the method of Siegel and Monty (28) using a partial specific volume of 0.737 ml/g, calculated from its composition by the method of Lee and Timasheff (21). The calculated value of 14,228 Da agrees well with the molecular mass (15,183 Da) of the monomeric species deduced from the DNA sequence. We conclude that TraYp is a monomer under these solution conditions.

Apparent dissociation constants for TraYp binding to *oriT* and the *traYI* promoter. The apparent K_d for TraYp binding to either *oriT* or the *traYI* promoter was determined by a gel mobility shift assay. Purified TraYp was titrated in reaction mixtures containing either labelled *oriT* substrate or the labelled *traYI* promoter substrate (0.5 ng per reaction), and the fraction of the labeled DNA bound by protein was determined. The apparent K_d was determined by the method of Tsai et al. (38) using the equation $K_a' = [ES]/[E_0][S]$, where E and S are protein and DNA, respectively, assuming an excess of TraYp in the reaction such that $[E_0]/[E] = 1$. TraYp was found to bind *oriT* with an apparent K_d of 19.4 ± 7.3 nM (monomer) ($n = 4$), while the K_d for binding to the *traYI* promoter was determined to be 107 ± 40 nM (monomer) ($n = 5$). Thus TraYp exhibits an approximately fivefold higher affinity for its binding site on *oriT* compared with its binding site at the *traYI* promoter.

The effect of TraYp on the helicase I nicking reaction. It has long been assumed that TraYp has a role in the site- and

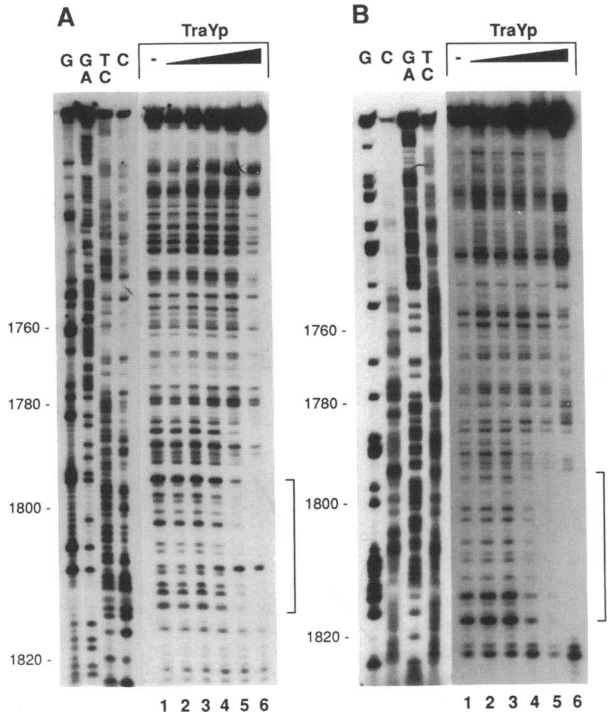


FIG. 5. Footprint of TraYp bound to the *traYI* promoter. (A) Bottom strand. DNase I footprinting assays were performed as described in Materials and Methods by using the 189-bp *Eco*RI-*Hin*dIII fragment isolated from pBSPYI and 5′ end labelled at the *Hin*dIII site. Lane 1, no TraYp; lanes 2 to 6, addition of 17.6, 35.2, 70.4, 140.8, or 352 ng of TraYp, respectively, to the reaction mixture prior to DNase I digestion. (B) Top strand. The 189-bp *Eco*RI-*Hin*dIII fragment isolated from pBSPYI and 3′ end labelled at the *Hin*dIII site was used as the DNA substrate in the footprinting reactions. Lanes 1, no TraYp; lanes 2 to 6, addition of 22.75, 45.5, 91, 182, or 364 ng of TraYp, respectively, to the reaction mixture prior to DNase I digestion as above. Markers at the left of each panel are Maxam-Gilbert sequencing ladders (9) of the appropriate substrate fragments. Brackets denote the region protected from DNase I digestion by TraYp. The map position coordinates in base pairs relative to the *Bgl*II site at 66.7 kb on the F plasmid map (16) are indicated on the left in each panel.

strand-specific nicking reaction that initiates conjugative DNA transfer. Indeed, genetic experiments have indicated such a role for this protein in vivo (10, 25). However, recent biochemical studies have shown that helicase I, in the absence of TraYp, is able to catalyze site- and strand-specific nicking at the F plasmid *oriT* in vitro (24, 27). Since TraYp binds near the nick site and has been implicated in the nicking reaction in vivo, it was of interest to determine whether TraYp had any effect on the site- and strand-specific nicking reaction catalyzed by helicase I. To this end, TraYp was added at various concentrations to reaction mixtures containing plasmid pBSoriT and helicase I. The conversion of supercoiled DNA to the nicked species was assayed on agarose gels run in the presence of ethidium bromide (Fig. 7). As shown in Fig. 7, lane 2, helicase I in the absence of TraYp nicks pBSoriT. Increasing concentrations of TraYp in the absence of helicase I did not convert supercoiled DNA to a nicked species (Fig. 7, lanes 3 to 6). Furthermore, the addition of TraYp to reaction mixtures containing helicase I had no apparent effect on the extent of the nicking reaction (Fig. 7, lanes 7 to 10). Similar experiments using a range of helicase I concentrations produced the same results (data not shown).

The effect of TraYp on the rate of helicase I-catalyzed nicking of pBSoriT was also determined. In this case, a kinetic analysis was performed at several TraYp concentrations (data not shown). Again the addition of TraYp had no effect on the nicking reaction catalyzed by helicase I. Thus, under these conditions, there is no measurable effect of TraYp on the site- and strand-specific nicking reaction catalyzed by helicase I.

DISCUSSION

We have identified two sites within the *tra* region on the F factor which bind TraYp specifically. The first site is located within *oriT* approximately 64 bp upstream of the nick site within *oriT*. The region protected from DNase I digestion is approximately 36 bp in length, covering nucleotides 208 to 240 on the top strand and nucleotides 204 to 234 on the bottom strand (Fig. 3A). The site is 73% AT rich. Comparison with the *oriT* binding site (*sbyA*) of the R100 TraY protein (13) shows that the two sites are at approximately the same location and about the same size and composition. We propose that, by analogy to R100, this site be named F *sbyA*.

The function of TraYp binding at *oriT* is still unclear. Genetic data suggest that the *traY* gene product is necessary for both nicking and strand transfer in vivo (10, 25). This led to the hypothesis that TraYp either contained or was a component of a site- and strand-specific endonuclease activity responsible for nicking at *oriT*. It is now known, however, that helicase I, and not TraYp, contains the catalytic site responsible for site- and strand-specific nicking at *oriT* (24, 27). Moreover, we have shown here that TraYp has no observable effect on the in vitro helicase I-catalyzed nicking reaction. Perhaps the location of the TraYp-binding site gives a clue to its function. TraYp binds *oriT* between the two IHF-binding sites characterized by Tsai et al. (38) (Fig. 2). IHF is known to be involved in many processes in the cell. Among other roles, it stimulates recombination and the initiation of replication at *oriC* and is involved in the expression of several genes (17, 31, 39). Binding specifically to a 13-bp consensus sequence that is AT rich, IHF introduces bends in the DNA, possibly assisting the localized melting necessary for open complex formation. Intrinsic bends in the DNA at *oriT* (38), along with the bends introduced by IHF and perhaps TraYp, could result in a very distinct structure at *oriT*. This structure could be important in modulating the site- and strand-specific nicking reaction in vivo.

The second TraYp-binding site, which we propose to name F *sbyB*, is located at the promoter of the *traYI* operon. The DNase I footprint covers the 28 bp between the mRNA start site and the *traY* start codon. Surprisingly, sequence analysis revealed no significant homology between *sbyA* and *sbyB*. The only striking similarity is the sequence ATAAA, which is found in both binding sites and is similar to the TAA(A/T)T motif in R100 *sby* sites (13).

The exact coincidence of F *sbyB* with the mRNA start site (Fig. 3B) suggests a regulatory function. Computer-guided homology search has shown TraYp to be homologous to the Arc and Mnt proteins of bacteriophage P22 (3). The Arc and Mnt proteins are involved in the regulation of the life cycle of the phage. Arc binds to a specific operator sequence O_{arc}, repressing the expression of itself and *ant*, the gene encoding antirepressor (Ant) protein. When expressed, Ant induces lytic growth of the phage. Arc acts late in the lytic cycle to regulate Ant levels. Mnt, by binding to O_{mnt}, maintains

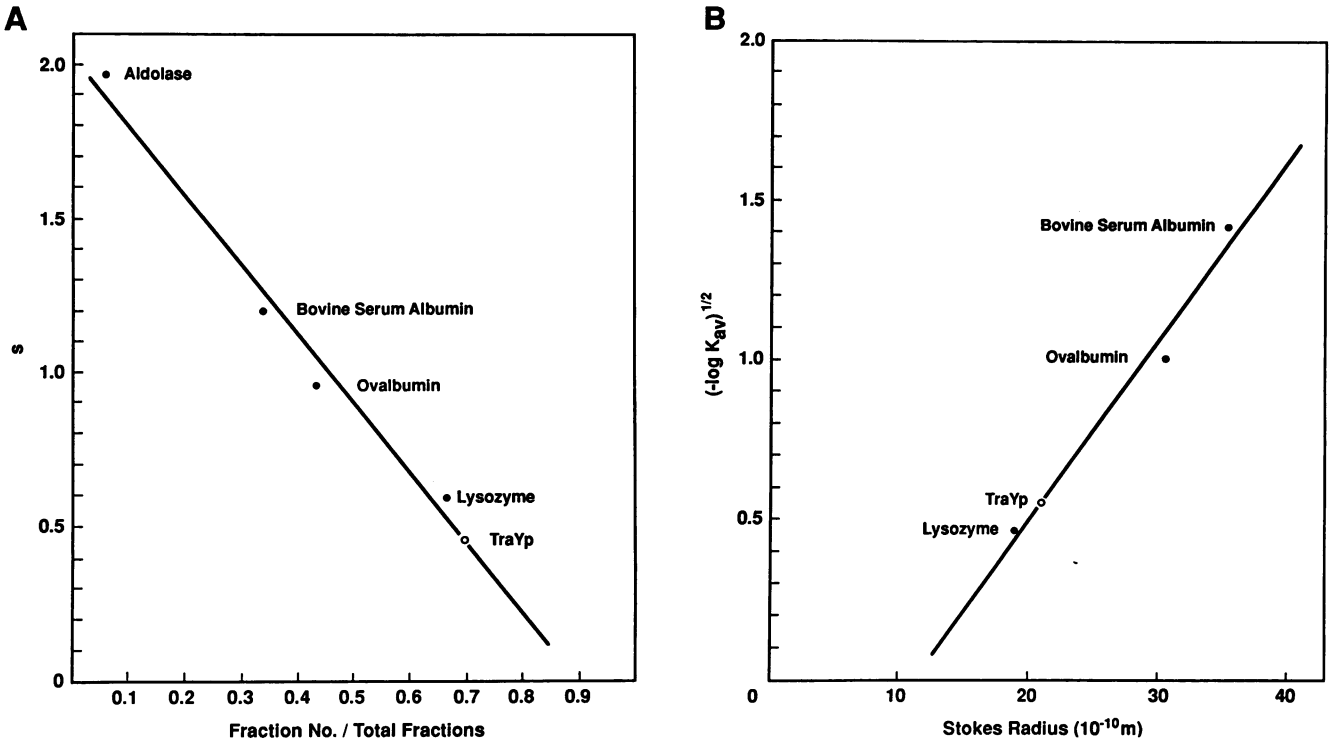


FIG. 6. Hydrodynamic data for TraYp. (A) Sedimentation coefficients. The data are averages of three experiments. In each case, 9.76 μg of TraYp was loaded onto a 15-to-35% (vol/vol) continuous glycerol gradient in a volume of 100 μl. Since not all gradients resulted in the same number of fractions, the horizontal axis is presented as the ratio of the fraction number to the total number of fractions in that particular experiment. The ordinate axis is the sedimentation coefficient (in svedbergs) as measured under the conditions described in Materials and Methods. The $s_{20,w}$ values for the protein standards were converted to $s_{T,m}$ values by the method of Martin and Ames (23) using the equation

$$s_{T,m} = s_{20,w}\frac{\eta_{20,w}(\rho_p - \rho_{T,m})}{\eta_{T,m}(\rho_p - \rho_{20,w})}$$

(B) Stokes radius. The data are averages of four experiments with the following starting concentrations of TraYp: 32, 32, 19.8, and 9.6 μM. $K_{av} = (V_e - V_0)/(V_t - V_0)$, where V_0 is the void volume, V_e is the elution volume, and V_t is the total column volume (20). Physical data for the protein standards were obtained from the supplier and from reference 32.

lysogeny by repressing *arc* and *ant* expression. Each protein binds as a tetramer to its specific operator sequence, regulating expression of downstream genes by blocking transcription initiation (33). Their similarities to TraYp, along with the position of *sbyB* at the mRNA start site, may suggest a similar role for TraYp. As yet, however, there are no data, genetic or biochemical, addressing the effect of TraYp binding at the promoter on expression of genes in the *traYI* operon.

The similarity to Arc and Mnt may also suggest a mode of binding for TraYp. The crystal structure for Arc bound to its operator sequence has been solved (6). These studies defined a new class of DNA-binding proteins. Upon dimerization, an antiparallel β-sheet is formed by residues near the amino terminus of each monomer. Subsequent tetramerization causes the β-sheets to be inserted into adjacent major grooves on the DNA at the recognition site. Dot matrix comparisons between the R100 TraYp and F TraYp nucleotide and amino acid sequences, performed by Inamoto et al. (14), suggested that F TraYp contains a tandem repeat. Secondary-structure prediction by the method of Garnier et al. (12) indicates that each TraYp repeat domain has a secondary structure nearly identical to the solved secondary structure for Arc (Fig. 8). On the basis of this analysis, we

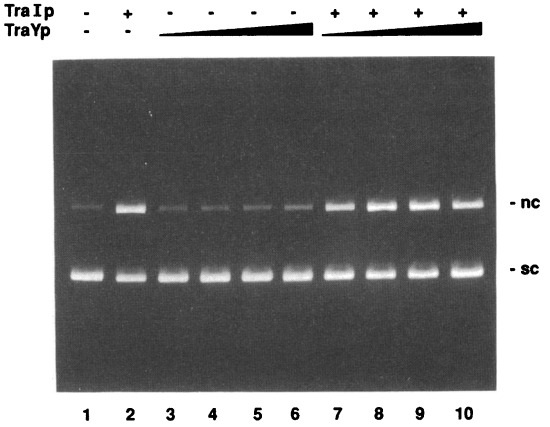


FIG. 7. Effect of TraYp on helicase I-catalyzed nicking at F *oriT*. Nicking reactions were performed as described in Materials and Methods by using 450 ng of pBSoriT DNA. Lane 1, no protein; lane 2, 90 ng of TraIp; lanes 3 to 6, no TraIp and 1.8, 3.5, 17.6, and 176 ng of TraYp, respectively; lanes 7 to 10, 90 ng of TraIp and 1.8, 3.5, 17.6, and 176 ng of TraYp, respectively. When TraYp and TraIp were both present, TraYp was incubated for 5 min at 37°C with the DNA substrate prior to the addition of TraIp. nc, nicked circular DNA; sc, supercoiled circular DNA.

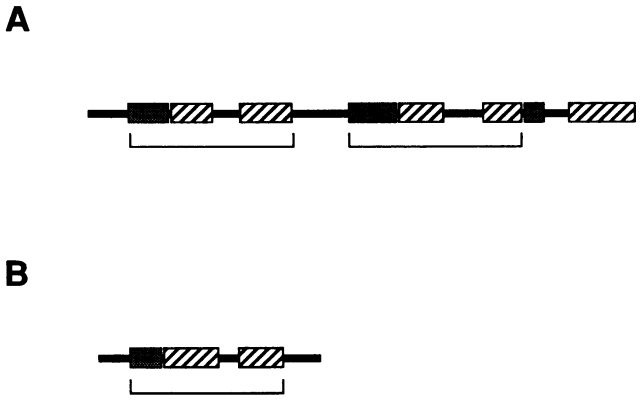


FIG. 8. Physical map of proposed secondary structures of TraYp and Arc. (A) Secondary-structure prediction by the method of Garnier et al. (12) for F TraYp. (B) Secondary structure of P22 Arc protein as determined by Zagorski et al. (42) and Breg et al. (5). Shaded boxes, β-sheet regions; hatched boxes, α-helices; thick lines, random coil. The secondary-structure motif discussed in the text is indicated by brackets.

propose that a TraYp monomer is equivalent to an Arc dimer. By using both primary- and secondary-structure comparisons, TraYp is analogous to Arc, strongly suggesting a similar mode of binding. Hydrodynamic data collected in this study indicate that TraYp is a monomer in solution under the conditions used for the binding assays. Thus, the question of whether the tether region (approximately 13 residues) between the two domains of TraYp is long enough to allow the proper conformation for binding in a single molecule or whether TraYp needs to dimerize on the DNA in order to bind remains to be answered. Stoichiometric experiments are currently in progress to determine the composition of the bound species seen in gel mobility shift assays. It will also be interesting to express the two domains separately and see whether either can bind as a homodimer or whether an equimolar mix can reconstitute binding activity.

Arc protein binds as a tetramer to an inverted repeat. Despite the fact that TraYp is similar to Arc, sequence analysis of the TraYp-binding sites showed no evidence of repeats in the DNA. Previous analysis of the *tra* region has also failed to locate repeats at these sites (15). It is possible that the binding site has been allowed to diverge because of the tandem repeat structure of TraYp. The protein domains are no longer interchangeable, perhaps allowing the binding site to evolve along with the domains. An alternate possibility is that TraYp recognizes a DNA structure not obvious from the sequence. Mutational analysis of the binding site will be required to determine precisely which bases are important for binding.

Protected regions *sbyA* and *sbyB* are longer than would be expected for a 15-kDa protein. This discrepancy can be explained by several hypotheses. TraYp could bind as a dimer, which because of its tandem repeat structure would mimic the Arc and Mnt tetrameric binding. Alternatively, the protein could wrap DNA around itself, as IHF does, forming bends or loops. Yet another possibility is that TraYp is ellipsoidal rather than spherical. This theory is supported by the hydrodynamic data presented in this paper. The frictional ratio (f/f_0) was calculated by the method of Siegel and Monty (28) to be 1.31, suggesting that the molecule is ellipsoidal or very highly solvated.

ACKNOWLEDGMENTS

We thank Mike Howard for critical reading of the manuscript, Jim George and Dan Bean for stimulating discussions, Aziz Sancar for assistance with the molecular weight calculations, and Susan Whitfield for preparation of the artwork.

This work was supported by grant MV-435 from the American Cancer Society. S. W. Matson is a recipient of an American Cancer Society Faculty Research award.

REFERENCES

1. **Abdel-Monem, M., G. Taucher-Scholz, and M.-Q. Klinkert.** 1983. Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. Proc. Natl. Acad. Sci. USA **80**:4659–4663.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1525.
3. **Bowie, J. U., and R. T. Sauer.** 1990. TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. J. Mol. Biol. **211**:5–6.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–252.
5. **Breg, J. M., R. Boelens, A. V. E. George, and R. Kaptein.** 1989. Sequence-specific ^{1}H NMR assignment and secondary structure of the Arc repressor of bacteriophage P22, as determined by two-dimensional ^{1}H NMR spectroscopy. Biochemistry **28**:9826–9833.
6. **Breg, J. N., J. H. J. van Opheusden, M. J. M. Burgering, R. Boelens, and R. Kaptein.** 1990. Structure of Arc repressor in solution: evidence for a family of β-sheet DNA-binding proteins. Nature (London) **346**:586–589.
7. **Brenowitz, M., D. F. Senear, and R. E. Kingston.** 1989. DNaseI footprint analysis of protein-DNA binding, p. 12.4.1–12.4.16. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
8. **Chodosh, L. A.** 1989. Mobility shift DNA-binding assay using gel electrophoresis, p. 12.2.1–12.2.10. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
9. **Eckert, R. L.** 1992. DNA sequencing by the chemical method, p. 7.5.1–7.5.11. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
10. **Everett, R., and N. Willetts.** 1980. Characterisation of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. J. Mol. Biol. **136**:129–150.
11. **Fowler, T., L. Taylor, and R. Thompson.** 1983. The control region of the F plasmid transfer operon: DNA sequence of the *traJ* and *traY* genes and characterization of the *traY-Z* promoter. Gene **26**:79–89.
12. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97–120.
13. **Inamoto, S., and E. Ohtsubo.** 1990. Specific binding of the TraY protein to *oriT* and the promoter region for the *traY* gene of plasmid R100. J. Biol. Chem. **265**:6461–6466.
14. **Inamoto, S., Y. Yoshioka, and E. Ohtsubo.** 1988. Identification and characterization of the products from the *traJ* and *traY* genes of plasmid R100. J. Bacteriol. **170**:2749–2757.
15. **Ippen-Ihler, K. A., and E. G. J. Minkley.** 1986. The conjugation system of F, the fertility factor of *Escherichia coli*. Annu. Rev. Genet. **20**:593–624.
16. **Johnson, D., R. Everett, and N. Willetts.** 1981. Cloning of F DNA fragments carrying the origin of transfer, *oriT*, and the fertility inhibition gene *finP*. J. Mol. Biol. **153**:187–202.
17. **Kikuchi, Y., and H. Nash.** 1978. The bacteriophage λ *int* gene product. A filter assay for genetic recombination, purification of Int, and specific binding to DNA. J. Biol. Chem. **253**:7149–7157.

18. **Kunkel, T. A., J. B. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
19. **Lahue, E. E., and S. W. Matson.** 1990. Purified *Escherichia coli* F-factor TraY protein binds *oriT*. J. Bacteriol. **172:**1385–1391.
20. **Laurent, T. C., and J. Killander.** 1964. A theory of gel filtration and its experimental verification. J. Chromatogr. **14:**317–330.
21. **Lee, J. C., and S. N. Timasheff.** 1979. The calculation of partial specific volumes of proteins in 6 *M* guanidine hydrochloride. Methods Enzymol. **61:**49–57.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Martin, R. G., and B. N. Ames.** 1961. A method for determining the sedimentation behavior of enzymes: application to protein mixtures. J. Biol. Chem. **236:**1372–1379.
24. **Matson, S. W., and B. S. Morton.** 1991. *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. J. Biol. Chem. **266:**16232–16237.
25. **McIntire, S., and N. Willetts.** 1980. Transfer-deficient cointegrates of F*lac* and lambda prophage. Mol. Gen. Genet. **178:**165–172.
26. **Panicker, M. M.** 1985. The role of TraDp in F sex factor-mediated conjugation in Escherichia coli. Ph.D. thesis. Carnegie-Mellon University, Pittsburgh.
27. **Reygers, U., R. Wessel, H. Müller, and H. Hoffmann-Berling.** 1991. Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. EMBO J. **10:**2689–2694.
28. **Siegel, L. M., and K. Monty.** 1966. Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. Biochim. Biophys. Acta **112:**346–362.
29. **Silverman, P. M.** 1986. The structrural basis of prokaryotic DNA transfer, p. 277–309. *In* M. Inouye (ed.), Bacterial outer membranes as model systems. John Wiley & Sons, Inc., New York.
30. **Silverman, P. M., E. Wickersham, and R. Harris.** 1991. Regulation of the F plasmid *traY* promoter in *Escherichia coli* by host and plasmid factors. J. Mol. Biol. **218:**119–128.
31. **Skarstad, K., T. A. Baker, and A. Kornberg.** 1990. Strand separation required for initiation of replication at the chromosomal origin of *E. coli* is facilitated by a distant RNA-DNA hybrid. EMBO J. **9:**2341–2348.
32. **Sober, H. A. (ed.).** 1970. Handbook of biochemistry and molecular biology, 2nd ed. Chemical Rubber Co., Cleveland.
33. **Susskind, M. M., and P. Youderian.** 1983. Bacteriophage P22 antirepressor and its control, p. 347–366. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
34. **Tabor, S.** 1989. Phosphatases and kinases, p. 3.10.1–3.10.5. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
35. **Thompson, T. L., M. B. Centola, and R. C. Deonier.** 1989. Location of the nick at *oriT* of the F plasmid. J. Mol. Biol. **153:**187–202.
36. **Traxler, B. A., and E. G. Minkley, Jr.** 1987. Revised genetic map of the distal end of the F transfer operon: implications for DNA helicase I, nicking at *oriT*, and conjugal DNA transport. J. Bacteriol. **169:**3251–3259.
37. **Traxler, B. A., and E. G. Minkley, Jr.** 1988. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. J. Mol. Biol. **204:**205–209.
38. **Tsai, M.-M., Y.-H. F. Fu, and R. C. Deonier.** 1990. Intrinsic bends and integration host factor binding at F plasmid *oriT*. J. Bacteriol. **172:**4603–4609.
39. **Tsui, P., and M. Freundlich.** 1988. Integration host factor binds specifically to sites in the *ilvGMEDA* operon in *Escherichia coli*. J. Mol. Biol. **203:**817–820.
40. **Willetts, N., and R. Skurray.** 1980. The conjugation system of F-like plasmids. Annu. Rev. Genet. **14:**41–76.
41. **Willetts, N. S., and B. Wilkins.** 1984. Processing of plasmid DNA during bacterial conjugation. Microbiol. Rev. **48:**24–41.
42. **Zagorski, M. G., J. U. Bowie, A. K. Vershon, R. T. Sauer, and D. J. Patel.** 1989. NMR studies of Arc repressor mutants: proton assignments, secondary structure, and long-range contacts for the thermostable proline-8→leucine variant of Arc. Biochemistry **28:**9813–9825.