Transposable Dual Reporters for Studying the Structure−Function Relationships in Membrane Proteins: Permissive Sites in *R. prowazekii* ATP/ADP Translocase†

Mikhail F. Alexeyev and Herbert H. Winkler*

Laboratory of Molecular Biology, Department of Microbiology and Immunology, University of South Alabama College of Medicine, Mobile, Alabama 36688

Received September 19, 2001; Revised Manuscript Received October 25, 2001

ABSTRACT: A new approach to studying membrane topology and permissive sites in membrane proteins expressed in *Escherichia coli* is described. The method is based on in vitro transposition of mini-Tn*5* derivatives bearing dual *pho-lac* reporters [Alexeyev, M. F., and Winkler, H. H. (1999) *J. Mol. Biol. 285*, 1503−1513]. Two mini-Tn*5* transposons, Tn*pholac*1 and Tn*pholac*2, were designed in such a way that their insertions can be converted either by restriction−ligation or by in vivo Cre-*lox* recombination into either sandwich reporter fusions or short amino acid (aa) tags (25 or 42 aa long). A set of 48 unique insertions in the gene coding for the *Rickettsia prowazekii* ATP/ADP translocase (Tlc) was generated using Tn*pholac*2. The topological information generated by these insertions was found in to be in good agreement with the existing topological model. Subsequently, these insertions were converted into both 25 and 42 aa tags, and the activity of the resulting mutants was determined. Also, site-directed mutagenesis was used to construct insertions in the loops, where no transposon hops were discovered. Of 13 extramembrane domains in Tlc, only 3 (loops 7, 10, and 13) were found to be permissive, which is in marked contrast to previous observations in the *E. coli* lactose permease (LacY), where most insertions in extramembrane domains were demonstrated to be permissive. The permissiveness of the insertion after I368 in TM IX lead us to reconsider the boundaries for this TM by placing I368 on the interface between TM IX and loop 10. Interestingly, the 25 aa insertions consistently have 2-fold higher activity than the corresponding 42 aa insertions, which is also in contrast with observations made on LacY. Finally, in this study we report, for the first time, the frequency of 10 base pair target duplications generated by in vitro Tn*5* transposition.

The obligate intracellular bacterium *Rickettsia prowazekii*, the etiologic agent of epidemic typhus, grows and multiplies only within the cytoplasm of a eukaryotic cell. As an adaptation to this unusual ecological niche, *R. prowazekii* evolved to possess transport systems that allow these organisms to take advantage of the many preformed metabolites that are abundant in the host cell (*1*). Some of these transport systems have never been described in free-living bacteria. The ATP/ADP translocase (Tlc) that exchanges ADP, generated as a result of rickettsial metabolism, with the high-energy ATP of the host cell is an example of such a system (*2*). Interestingly, Tlc is quite dissimilar to mitochondrial ADP/ATP transporters yet possesses a high degree of similarity to recently discovered plant plastid ATP/ADP transporters and those of chlamydial species (*3*). At present, there is very limited information available about structure−function relationships in nonmitochondrial ATP/ADP transporters (*4*). Therefore, the aim of this study was to advance our understanding of these relationships.

The goal of a structure−function analysis of a protein is to define how different parts of the polypeptide chain contribute to the form and function of the whole molecule. In such an analysis, complementary information is gained from mutations that lead to a change of function (loss of function or hyperfunction) and from mutations that have little or no effect (*5*). Mutagenesis employed in structure−function studies may involve amino acid (aa)[1] replacement, deletion, or insertion of aa stretches of different length. In the last case, depending on whether the protein's function is preserved, permissive and nonpermissive sites are distinguished (*6*). In membrane transport proteins, the permissiveness of a particular site or a set of sites provides information that can be interpreted in terms of the structural rigidity/flexibility of the polypeptide chain at that location and of the functional involvement (transport, expression, membrane insertion) of the particular site. Information on the permissiveness of a particular site has been successfully exploited. For example, the overall protein hydrophilicity for crystallization studies has been increased by inserting hydrophilic polypeptides (*7*),

† This work was supported by U.S. Public Health Service Grant AI-15035 from the National Institute of Allergy and Infectious Diseases to H.H.W.

* Correspondence should be addressed to this author at the Laboratory of Molecular Biology, Department of Microbiology and Immunology, University of South Alabama College of Medicine, Mobile, AL 36688. Tel.: (334) 460-6108, Fax: (334) 460-7269, E-mail: herbertw@sungcg.usouthal.edu.

[1] Abbreviations: IPTG, isopropyl-β-D-thiogalactopyranoside; BG, β-galactosidase; AP, alkaline phosphatase; aa, amino acid; NAR, normalized PhoA:LacZ activity ratio; Tn*pholac*1, mini-Tn*5*-2 derivative; Tn*pholac*2, mini-Tn5-2 derivative; TB, terrific broth; bp, base pair(s); Red-Gal, 6-chloro-3-indolyl-β-D-galactoside; X-phos, 5-bromo-4-chloro-3-indolyl phosphate; RBS ribosome-binding site.

10.1021/bi011813i CCC: $22.00
Published on Web 12/08/2001

protease cleavage sites have been introduced (*8*), and histidine residues that can be used for purification have been inserted (*9*). Several methods have been employed to generate in-frame insertions. These include random insertions using DNaseI (*10*), preexisting or newly generated restriction sites (*11*), and, more recently, transposons (*5*, *12*, *13*). This report describes a study on topology and permissive sites in *R. prowazekii* ATP/ADP translocase using a new transposon insertion strategy. Its advantages over the existing ones are as follows: (i) the selection of informative constructs through the use of dual reporters, (ii) higher throughput provided by the application of recently developed in vitro transposition protocols (*14*), and (iii) the generation of various fusions with different properties via the utilization of new transposons. These include not only sandwich reporter fusions but also the insertion of epitopes, glycosylation and protease cleavage sites, and cysteine residues.

MATERIALS AND METHODS

Transposition. Mini-Tn*5* transpositions in vitro were performed using EZ::TN transposase (Epicenter Technologies, Madison, WI). Briefly, equimolar amounts (70 nmol) of transposon donor and target (pMA613, ref *22*) plasmids were combined in EZ::TN reaction buffer (reaction volume 10 μL), 1 μL of EZ::TN transposase was added, and the reaction was allowed to proceed for 2 h at 37 °C. The transposition reaction products either were used immediately to transform competent *E. coli* TG1 cells or were frozen at −70 °C. *E. coli* transformants were plated on dual indicator plates supplemented with ampicillin (target plasmid marker) and chloramphenicol (transposon marker) and screened for colony coloration after 24 h incubation at 37 °C.

Dual indicator plates contained the following: Bacto-agar at 1.5%, Bacto-tryptone at 1%, yeast extract at 0.5%, NaCl at 0.5%, 5-bromo-4-chloro-3-indolyl phosphate, disodium salt (X-phos; Diagnostic Chemicals LTD., PEI, Canada) at 80 μg/mL, 6-chloro-3-indolyl-β-D-galactoside (Red-Gal, Research Organics Inc., Cleveland, OH) at 100 μg/mL, IPTG at 1 mM, and K_2HPO_4 at 80 mM and pH 7.0.

Site-Directed Insertion. Three aa insertions (LEL) in loops 3, 5, 9, 10, and 12 were made by overlap extension PCR using the following mutagenic primers: C3Rev CAGGAGCTCCAGGATGTCGCACAGTTTAAC and C3For CCTGGAGCTCCTGAAACAGGAAAACGTAT for loop 3; C5Rev GATGAGCTCCAGTTTAGTGATCTGGTTAGC and C5For ACTGGAGCTCATCGCTGAAGCTAAACGTT for loop5; C9Rev GCCAGAGCTCCAGAGAAACTTTACGCAGGATG and C9For TCTGGAGCTCTGGCTGACTGCAGCTATGATC for loop 9; P10Rev CGATGAGCTCCAGAACAGAGTCGAAGAAGATG and P10For TCTGGAGCTCATCGCGATGAACCTGACC for loop10; and P12Rev GAGTGAGCTCCAGAGCTTCGATGAAACCGAAAAC and P12For TCTGGAGCTCACTCCGTACTTCGCTTCT for loop12.

These 3 aa insertions were expanded to 25 aa insertions by inserting the 66 bp DNA fragment CTGTCTCTTATACACATCTCATGGCCCATGGGGCCTGAGCTCTGCGGAGATGTGTATAAGAGACAG into the unique *Ecl*136II site inside the coding sequence of 3 aa insertions. This fragment was generated by annealing and extension of two primers: 25For CTGTCTCTTATACACATCTCATGGCCCATGGGGCCTGAGC and 25Rev CTGTCTCTTATACACATCTCCGCAGAGCTCAGGCCCCATGGGCCAT, and it is identical to that found in 25 aa transposon-generated insertions.

The 509 aa insertion of the *pho-lac* reporter from pMA632 (*22*) in the *Hpa*I site in loop13 was generated using restriction−ligation. Insertions were verified by sequencing.

Analysis of Tnpholac2 Insertions. Colored colonies were picked and resuspended individually in 40 μL of sterile deionized water in 96-well microtiter plates. Insertion points were mapped by PCR with the forward primer Bamtlc (specific for the 5′ end of the *tlc* gene), (5′)GGGATCCATGTCTACTTCCAAATCTGAAAAC(3′), and the reverse primer Phoseq (specific for the 5′ end of the *pho-lac* fusion), (5′)TCACCCGTTAAACGGCGAGCACC(3′). PCR reactions [3 μL of bacterial suspension, 25 pmol of each primer, 250 μM of each dNTP, 16 mM $(NH_4)_2SO_4$, 3.5 mM $MgCl_2$, 150 μg/mL BSA, and 50 mM Tris-HCl, pH 9.1, and 1−2 units of KlenTaq1J DNA polymerase in total volume of 25 μL] were cycled for 30 cycles (95 °C for 30 s, 50 °C for 30 s, and 72 °C for 2min). PCR products were analyzed by 1% agarose gel electrophoresis. On the basis of PCR analysis, individual clones were picked and single-colony-purified, and plasmid DNA was isolated. Insertion points were determined by double-stranded dye-terminator sequencing using primer Phoseq.

Double-stranded plasmid DNA was prepared using the Qiaprep Spin miniprep kit (Qiagen, Chatsworth, CA) and sequenced using the BigDye Terminator cycle sequencing kit (PE Biosystems, Foster City, CA).

Reporter and Drug Resistance Gene Excision by Cre/lox Recombination. (a) Excision by Superinfection with Phagemid. E. coli TG1 cells bearing Cre-expressing plasmid pMA713 (*15*) were infected with helper phage R408 (Promega, Madison, WI) as described (*16*) to package the single-stranded form of the phagemid into the phage particles. The titer of the resulting phagemid particles was determined by infection of *E. coli* TG1 and plating on kanamycin-containing LB plates. Overnight cultures of *E. coli* TG1 bearing plasmids with transposon insertions that were to be excised were diluted, in microcentrifuge tubes, to approximately 10^4 cfu/mL in a fresh modified LB medium [16 g of tryptone, 10 g of yeast extract, 5 g of NaCl per liter (*16*)] supplemented with 0.1% glucose and 200 μg/mL ampicillin. pMA713-containing phage particles were then added to each tube at a MOI of 100−1000. Cultures were incubated for 30 min at 37 °C without aeration and then for 2 h with vigorous agitation after which 10 μL of each culture was streaked on dual indicator plates containing ampicillin to give separate colonies. White colonies were picked, and Cre/*lox* recombination was verified by restriction analysis.

(b) Excision by F' Mating-In. An alternative procedure for excision of dual reporters plus drug resistance markers from mini-Tn insertions involved mating-in an F'::mini-Tn*10cre*, Km^r, which is an F' derivative bearing mini-Tn*10* with the *cre* gene under the control of the *lac* promoter. Mating-in was conducted by mixing of 1 mL of LB medium containing a 1:100 dilution of a stationary-phase culture of the donor (LE392 [F'::mini-Tn*10cre*, Km^r] or HB101 [F'::mini-Tn*10cre*, Km^r]) and 10 μL of an overnight culture of recipient (TG1 strain bearing a plasmid with a Tn*pholac*2 insertion in the *tlc* gene). Conjugation was allowed to proceed

until visible growth occurred (2−6 h), and exconjugants were selected by streaking 10 μL of the resulting culture on dual indicator plates containing ampicillin (200 μg/mL) and kanamycin (50 μg/mL). Plasmid DNA isolated from white exconjugant colonies was used in subsequent restriction endonuclease digests and sequencing analysis.

Reporter and Drug Resistance Gene Excision by SfiI Digestion−Ligation. Reporter and drug resistance marker removal by *Sfi*I digestion−ligation was performed in a 96-well plate format. Each Tn*pholac*2 insertion containing plasmid (0.2−0.5 μg) was digested with *Sfi*I and subsequently ligated at 18 °C overnight. Finally, each ligation mixture was electroporated into *E. coli* TG1 and plated onto dual indicator plates supplemented with ampicillin. Plasmid DNA from white colonies was used for subsequent restriction analysis to confirm the excision event.

Functional Assays for Tlc Activity. For high-throughput analysis of transport activity, excised derivatives of the plasmids bearing Tn*pholac*2 insertions were introduced into *E. coli* C41 strain (*17*) using freeze−thaw transformation (*18*). Transformants were inoculated into 0.5 mL of TB containing ampicillin (300 μg/mL) in 96-deep-well microplates (Denville Scientific, Metuchen, NJ) and grown overnight with aeration. The next day, overnight cultures were diluted 1:40 into 1 mL of fresh TB broth containing ampicillin (300 μg/mL) in 96-deep-well microplates and grown with aeration for 3 h. Plates were placed on ice, and 25 μL aliquots of each culture were transferred into the wells of a dot-blot apparatus fitted with a 0.45 μm nitrocellulose membrane (Millipore, Bedford, MA). Cells were washed 3 times with 200 μL of KPi buffer (50 mM potassium phosphate, pH 7.5) and exposed for 5 min to 100 μL of [α-^{32}P]ATP in KPi buffer at a concentration of 50 μM and 1 μCi/mL. Radioactive substrate was removed by applying vacuum and washing 3 times with KPi buffer, after which the nitrocellulose membrane was dried and exposed to a phosporimaging screen of the CycloneJ apparatus with OptiQantJ software (Packard, Downers Grove, IL). Positive clones revealed by the initial screening were further analyzed to determine the fraction of remaining activity essentially as described (*19*). Cytoplasmic membrane preparations and Western blotting were performed as described (*19*), except sonication rather than a French pressure cell was used for cell disruption.

Interconversion of Different Types of Insertions. Tn*pholac*2 insertions in *tlc* were converted into Tn*pholac*1 insertions by swapping internal fragments of transposons flanked with *Sfi*I sites. Sandwich fusions were generated from Tn*pholac*2 insertions by removal of antibiotic-resistance markers through a *Not*I digestion/self-ligation procedure. Insertions 24, 25, 26, 27, and 28 in periplasmic loop 4 were converted into C-terminal fusions by replacing *Bam*HI−*Apa*I fragments in these fusions with PCR fragments generated with common forward primer Bamtlc and insertion-specific reverse primers:

24: (5′)GGGGGCCCGGAACCAGGTCCGGGTACGGG(3′)

25: (5′)GGGGGCCCGGGCTTTCGATAGTTTTGTGGT(3′)

26: (5′)GGGGGCCCGGAGCCAGAGACAAGCTTTCGAT(3′)

27: (5′)GGGGGCCCGGGTTCGGGTAAGCCAGAGACA(3′)

28: (5′)GGGGGCCCGGCCATTTACCAACGATTTTGA(3′)

Reporter Activity Assays. Activities of Tlc fusions were assayed by streaking clones on LB assay plates that had the same composition as dual indicator plates, except Red-Gal and X-phos were omitted. The plates were incubated at 37 °C for 24 h; cells were collected with sterile plastic loops and resuspended in 10 mM Tris, pH 8.0, 10 mM $MgSO_4$. Enzymatic assays of these cell suspensions were performed as described (*20*, *21*), except background activities were subtracted from experimental data. Normalized activity ratios (NARs) were calculated as described previously (*22*).

RESULTS

Donor Constructs. We constructed two transposons (Figure 1), Tn*pholac*1 and Tn*pholac*2. These transposons were designed to facilitate studies on topology, permissive sites, and helix packing in membrane proteins. Both transposons consist of two genes, a dual *pho-lac* reporter (*22*) and a drug resistance marker flanked by inverted hyperactive mutant 19 bp Tn*5* outer ends (OE) that ensure an efficient in vitro transposition of the encompassed DNA (*14*). The reading frame of the *pho-lac* reporter is fixed inside both transposons in such a way that it forms an open reading frame (ORF) with the first nucleotide in the left OE. In other words, to generate a reporter fusion to a target gene, the left OE has to be placed after a third nucleotide in the codon of the target gene. In both transposons, an *Ecl*136II site is placed internally immediately adjacent to the right OE (Figure 1A,B). The internal part of each transposon is flanked by two *Sfi*I restriction sites that overlap *Nco*I sites and are immediately adjacent to the OE on the left and an *Ecl*136II site on the right. This configuration allows for the removal of the internal part of transposons by either *Sfi*I or *Nco*I restriction−ligation. Such a manipulation would leave behind a DNA tag consisting of two OE with *Sfi*I and *Ecl*136II sites between them. Since the initial dual reporter insertion was in-frame with the target gene, this tag will also be in-frame and will be translated into 25 aa sequence (Figure 1C). The antibiotic resistance marker in each transposon is flanked by *Not*I sites that allow for the conversion by restriction−ligation of C-terminal reporter fusions generated by initial transposition into sandwich reporter fusions. The latter are believed to be better indicators of membrane protein topology (*23*). Also, each transposon contains a pair of directly repeated 34 bp *lox*P sites. Cre-mediated recombination between these sites (*24*) enables high-throughput in vivo generation of either the sandwich fusion (Tn*pholac*1) or DNA taged fusions (Tn*pholac*2). In the latter case, the peptide tag is 42 aa as compared to the 25 aa tag generated by *Sfi*I/*Nco*I restriction−ligation (Figure 1D). pMA814 and pMA757 carry Tn*pholac*1 and Tn*pholac*2 inserted into pKNOCK-Cm and pKNOCK-Gm (*25*) backbone, respectively. pKNOCKs are suicide vectors based on the γ-origin of replication of the R6K plasmid and therefore can only replicate in those *E. coli* strains that can provide Π protein, product of the R6K *pir* gene. This allows one to perform plasmid-to-plasmid in vitro transposition, thus eliminating the need for isolation of transposon fragment from the donor plasmid.

Generation of Insertions in the tlc Gene. Tn*pholac*2 was validated by generating a series of insertions in the synthetic *tlc* gene coding for *R. prowazekii* ATP/ADP translocase (*26*). A single in vitro reaction in conjunction with subsequent transformation into *E. coli* TG1 was sufficient to generate

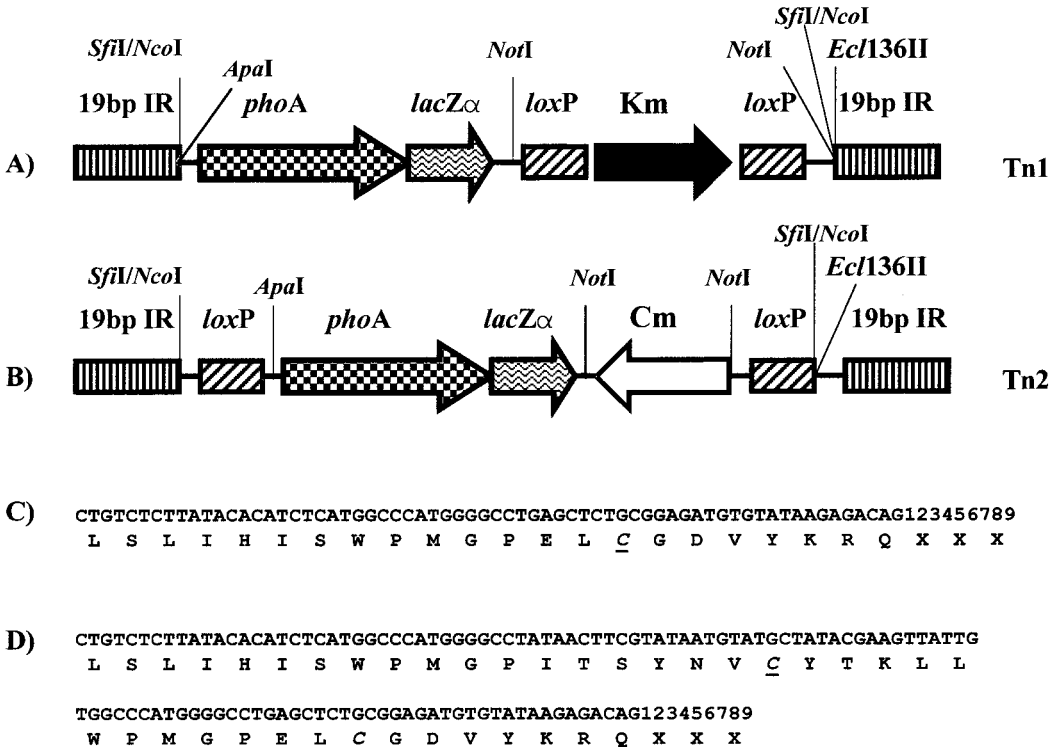

FIGURE 1: Structures of Tn*pholac*1, Tn*pholac*2, and sequences of 25 and 42 aa insertions. (A and B) Structures of Tn*pholac*1 and Tn*pholac*2, respectively. IR, 19 bp inverted repeats. Km and Cm, genes conferring resistance to kanamycin and chloramphenicol, respectively. (C and D) Sequences of 25 and 42 aa insertions resulting from Tn*pholac*2 excision with *Sfi*I and *lox*P, respectively. XXX, direct repeat of 3 aa resulting from 9 bp target duplication. Cysteine residues that can be used in cysteine accessibility experiments are underlined.

ca. 5000 in-frame (as judged by coloration on dual indicator plates) insertions. Of those, 96 colonies, which were blue or purple and indicative of periplasmic or transmembrane (TM) reporter fusions, and 192 colonies, which were red and indicative of cytoplasmic reporter fusions, were subjected to colony-PCR to determine the approximate location of insertion points. A total of 94 out of 96 alleged periplasmic/TM fusions and 88 out of 192 alleged cytoplasmic fusions produced PCR products. Of the PCR-positive clones, insertion points were identified by sequencing in 55 alleged periplasmic/TM and 28 alleged cytoplasmic fusions. This resulted in 48 unique insertion points (Table 1 and Figure 2) and 35 clones mapping at the same points.

The design of the Tn*pholac*2 allows for the removal of the internal part of transposon either by combination of *Sfi*I (or *Nco*I) restriction digest and ligation or by Cre-mediated recombination, leaving behind an insertion of either 25 or 42 aa, respectively. To investigate the effect of insertion size on protein function, we used both techniques to generate 2 sets of 48 insertions each at the same positions. The activities of the resulting mutant proteins were studied after introducing the resulting constructs into *E. coli* C41 (*17*). In both sets, the same three insertion points were found to be permissive: after N243 and Y251 in the large cytoplasmic loop 7 and after I368 at the interface of TM IX and the small periplasmic loop 10. Interestingly, Western blot analysis of preparation of membrane proteins revealed that insertions invariably led to a reduced membrane content of insertional mutant proteins as compared to wild-type regardless of the insertion site (Table 1 and Figure 3). This, taken together with significant residual activity of permissive insertions (Table 1), suggests that these three insertion mutants retain almost wild-type specific activity.

In loops 3, 5, 9, 10, and 12, which were not hit by transposons, 3 aa insertions were generated by overlap extension PCR. The 9 bp sequence CTGGAGCTC inserted in every case encompasses a restriction site for *Ecl*136II (underlined), which was used for subsequent expansion of 3 aa insertions into 25 aa insertions. Both the coding and aa sequences of these 22 additional aa were identical to those found in the 25 aa insertions generated by transposons. Both 3 aa and 25 aa insertions in loops 3, 5, 9, and 12 were found to retain less than 1% of wild-type activity (Table 2). However, both 3 aa and 25 aa insertions in loop 10 were active (79 and 21% of WT activity, respectively). As was the case with transposon-generated insertions, the steady-state membrane protein levels were severely reduced in all insertion mutants with the exception of the 3 aa insertion in loop 10, which accumulated protein to an estimated 60% of wild type (Table 2 and Figure 3).

The reported ability of wild-type Tn*5* to generate a 10 rather than the usual 9 bp duplication (*27*) is significant because our transposons are able to generate sandwich fusions only in conjunction with 9 bp duplication. We determined the frequency of 10 bp target duplication by mutant hyperactive Tn*5* transposase. As demonstrated by the sequencing of the 3′ ends of insertions, only 3 of 48 sequenced clones had 10 bp duplication (Table 1).

Topological Information Generated by Insertions. Tn*pholac*1 and Tn*pholac*2 transposons were designed to deliver information about both permissive sites and topology of the protein under study. To test the suitability of Tn*pholac*2 insertions as a source of topological information, the activities of both PhoA and LacZ were measured for each fusion, and the ratios of these activities, after they were normalized to the highest AP or BG activity [NAR = (AP/highest AP)-

Table 1: Properties of Tn2-Generated Insertions

clone	aa[a]	pos[b]	color[c]	%AP[d]	%BG[d]	NAR[e]	activity[f]		protein[g]		dupl.[h]
1	Q22	c.1	red	2	56	1:28	−	−	+	+	9 bp
2	Y23	c.1	red	4	65	1:16	−	−	+	+	9 bp
3	N25	c.1	red	3	60	1:20	−	−	+	+	9 bp
4	A32	t.1	purple	64	48	1.3:1	−	−	+	+	9 bp
5	N41	t.1	blue	82	2	41:1	−	−	+	+	9 bp
6	S43	t.1	blue	71	1	71:1	−	−	+	+	9 bp
7	R46	t.1	blue	77	3	26:1	−	−	+	+	9 bp
8	D56	p.2	blue	74	8	9:1	−	−	+	+	9 bp
9	I57	p.2	blue	78	13	6:1	−	−	+	+	9 bp
10	I62	t.2	purple	67	7	10:1	−	−	+	+	9 bp
11	F64	t.2	purple	62	13	5:1	−	−	+	+	9 bp
12	Y68	t.2	purple	45	63	1:1.2	−	−	+	+	10 bp
14	P72	t.2	purple	64	84	1:1.3	−	−	++	+	9 bp
15	I76	t.2	purple	42	46	1:1	−	−	+	+	9 bp
16	A77	t.2	blue	39	12	3:1	−	−	++	++	9 bp
17	I80	t.2	purple	47	14	3:1	−	−	+	++	9 bp
18	Y95	t.3	red	11	31	1:3	−	−	+	+	9 bp
19	F105	t.3	purple	44	15	3:1	−	−	+	+	9 bp
20	F108	t.3	purple	70	12	6:1	−	−	+	++	9 bp
22	P114	p.4	purple	61	1	61:1	−	−	+	+	9 bp
24	V119	p.4	red	9	19	1:2	−	−	++	++	9 bp
25	S128	p.4	red	16	19	1:1	−	−	+	++	9 bp
26	A132	p.4	red	11	41	1:4	−	−	+	++	9 bp
27	N135	p.4	red	0	14	1:>100	−	−	+++	+	9 bp
28	W146	p.4	red	9	35	1:4	−	−	+++	++	9 bp
29	W158	t.5	red	7	66	1:8	−	−	++	++	9 bp
30	G159	t.5	red	13	43	1:3	−	−	++	++	9 bp
32	G188	t.6	red	31	62	1:2	−	−	+	++	9 bp
33	A191	t.6	blue	77	0	>100:1	−	−	+	+	9 bp
34	A194	t.6	purple	61	8	8:1	−	−	++	+	9 bp
35	T198	t.6	blue	67	4	17:1	−	−	+	+	10 bp
36	V200	t.6	purple	62	19	3:1	−	−	+	+	9 bp
37	I202	t.6	blue	84	0	>100:1	−	−	+	++	9 bp
38	H216	p.6	red	9	43	1:5	−	−	+	+	9 bp
39	T236	t.7	red	14	100	1:7	−	−	+	+++	9 bp
40	N243	c.7	red	4	28	1:7	24 ± 3	13 ± 4	+	+	9 bp
42	Y251	c.7	red	14	100	1:7	24 ± 4	15 ± 6	+	+	9 bp
43	L286	t.8	blue	71	0	>100:1	−	−	+	+	9 bp
44	I287	t.8	blue	67	0	>100:1	−	−	+	+	9 bp
45	V294	t.8	blue	86	3	18:1	−	−	+	+	9 bp
46	Y309	p.8	blue	100	7	14:1	−	−	+	+	9 bp
47	A314	p.8	purple	58	0	>100:1	−	−	+	+	9 bp
48	I352	t.9	purple	47	4	12:1	−	−	+	+	9 bp
49	I368	t.9	purple	46	9	5:1	56 ± 16	30 ± 8	+	+	9 bp
50	I393	t.11	red	8	70	1:9	−	−	+	+	9 bp
51	N398	c.11	red	5	58	1:11	−	−	+++	+	9 bp
52	Y406	c.11	red	0	68	1:>100	−	−	+	+	9 bp
53	R425	c.11	red	4	91	1:23	−	−	++	+	9 bp

[a] Amino acid in Tlc preceding Tn2 insertion. [b] Position of the fusion point on the topological map of Tlc; p., c., or t. followed by the number, corresponding to the numbered periplasmic, cytoplasmic, or transmembrane domains, respectively. [c] Coloration on dual indicator plates. [d] AP and BG activities of the fusions with regard to the maximum activity in the set. [e] Normalized AP:BG activity ratio (NAR) is the ratio of activities in columns 4 and 5 expressed as the ratio of two numbers, the smaller of which is 1. [f] Residual transport activity of 25 and 42 aa insertion mutants expressed as % of WT control. [g] Steady-state membrane protein content of mutant proteins with 25 aa and 42 aa transposon-generated insertions estimated from Western blots. +, 0−15% of wild type; ++, 15−50% of wild type; +++, 50−100% of WT. [h] Target duplication upon Tn2 insertion.

(BG/highest BG)$^{-1}$], were calculated. (*22*). In general, NARs of fusions were found to be in good agreement with our earlier study and support a topological model in which Tlc crosses the cytoplasmic membrane 12 times (*22*). However, insertions 24, 25, 26, 27, and 28 (after V119, S128, A132, N135, and W146, respectively) in periplasmic loop 4 and insertion 38 (after H216) in periplasmic loop 6 reported a cytoplasmic membrane sidedness, which is in disagreement with the existing topological model (Table 1). Interestingly, insertion 23 (after P114) in loop 4 reports correct topology, suggesting that aberrant reporting could be position-dependent. On the other hand, in the previous study, we found that two C-terminal dual reporter fusions in periplasmic loop 4 (after S128 and S130) and in periplasmic loop 6 (after H216) reported proper sidedness, although the latter fusion had an NAR of only 2:1 with 11% BG activity (*22*). This suggests that a particular configuration of Tn*pholac*2, e.g., the aa resulting from translation of transposon inverted repeats and/or *lox*P sites, could have made a contribution to abnormal reporting.

To estimate the relative contribution of each of these factors and demonstrate the utility of transposon features, we have generated several sets of constructs. First, all six insertions reporting abnormal topology were converted by *Not*I restriction−ligation into sandwich fusions that are believed to generate more reliable topological information

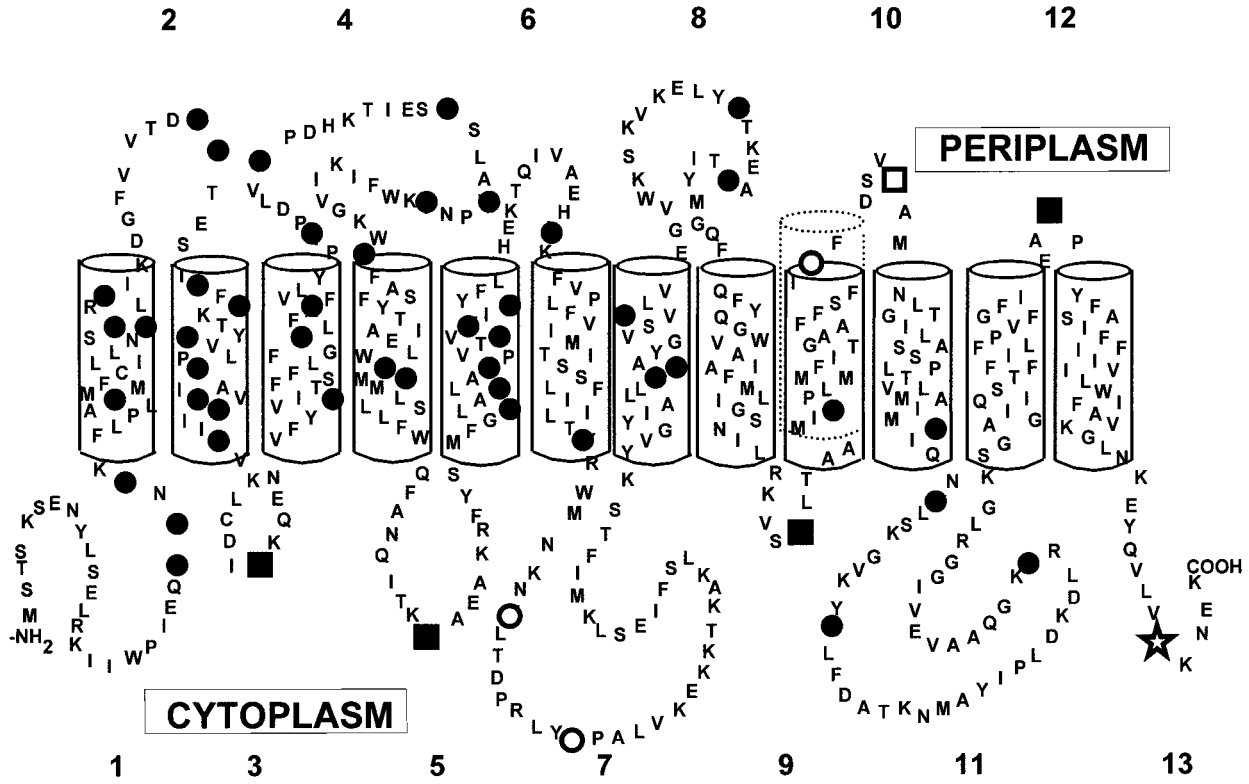

FIGURE 2: Location of insertion points in Tlc. Solid circles, transposon insertions at nonpermissive sites; open circles, transposon insertions at permissive sites; solid and open squares, permissive and nonpermissive, respectively, insertions generated through site-directed mutagenesis; star, position of the *pho-lac* fusion. Circles, squares, and a star are drawn over the position of the first amino acid that follows insertion. Solid cylinders, proposed boundaries of 12 TM domains in Tlc; dotted cylinder, old boundaries of TM X. Roman numerals indicate extramembrane domains.

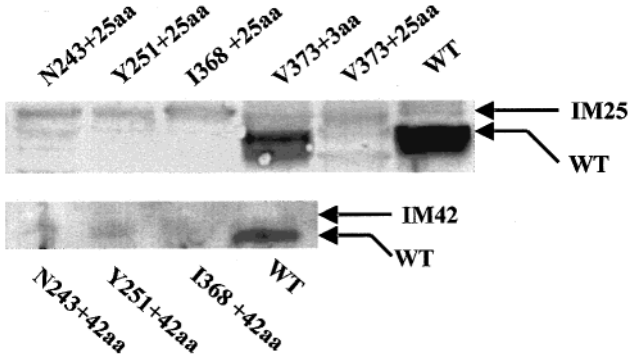

FIGURE 3: Western blot showing the steady-state membrane protein content of permissive Tlc insertion mutants. WT, wild type; IM25 and IM42, positions where insertion mutants that have, respectively, 25 aa and 42 aa insertions migrate. Each lane was loaded with equal amount of total membrane protein.

(*23*). Second, Tn*pholac*2 insertions were converted into Tn*pholac*1 insertions by swapping the internal *Sfi*I fragments. Finally, transposon insertions in loop 4 were reengineered into C-terminal dual reporter fusions devoid of extra amino acids resulting from translation of inverted repeats and *lox*P, sequences similar to those generated in previous study (*22*). The analysis of NARs of resulting constructs (Table 3) revealed that simple conversion of Tn*pholac*2 insertions into Tn*pholac*2 dual reporter sandwich fusions resulted in fusion 25 reporting the proper topology and improving NARs in fusions 24, 27, and 28. However, the same manipulation had a negative effect on NARs of fusions 26 and 38. In four out of six cases, conversion of Tn*pholac*2 insertions into Tn*pholac*1 insertions resulted in conversion of NARs into those consistent with our topological model and thus dramatically improved the quality of reporting. In the two other cases (fusions 26 and 28), NARs were inconclusive because of the extremely low enzymatic activity of both reporters (less than 5%). Conversion of the five Tn*pholac*2 insertions from loop 4 into C-terminal dual reporter fusions resulted, with one exception, in fusions with strongly periplasmic NARs. It is possible that K145 that precedes W146 in fusion 28 is responsible for this anomaly since positively charged residues are believed to be cytoplasmic topological determinants (*28*, *29*).

Table 2: Properties of Site-Directed Insertions

position[a]	aa[b]	ATP uptake (%)[c]	protein content[d]	ATP uptake (%)[c]	protein content[d]
c.3	I87	0.22 ± 0.03	+	0.34 ± 0.03	+
c.5	K176	0.44 ± 0.04	+	0.27 ± 0.04	+
c.9	S345	0.25 ± 0.05	+	0.45 ± 0.24	+
p.10	V373	79.2 ± 3.56	+++	21.2 ± 1.18	+
p.12	A463	0.34 ± 0.03	+	0.23 ± 0.04	+
p.13	V493	69.7 ± 8[e]	?[e]		

[a] Location of the fusion point on the topological map of Tlc; p., c., or t. followed by the number, corresponding to the numbered periplasmic, cytoplasmic, or transmembrane domain, respectively. [b] Amino acid in Tlc preceding insertion. [c] Residual ATP uptake expressed as percent of wild type. The data are the average of three independent experiments ± standard deviation. [d] Steady-state membrane protein content. +, 0−15% of wild type; ++, 15−50% of wild type; +++, 50−100% of wild type. [e] Insertion in loop13 represents a 509 aa *pho-lac* fusion. The resulting fusion protein does not react with C-terminal Tlc antibodies, which were used in this study.

Table 3: NARs of Different Types of Fusions at "Problematic" Points

clone[a]	location[b]	Tn2[c]	Tn2S[d]	Tn1[e]	C-term[f]
24	p.4	1:2	1:1	4:1	>100:1
25	p.4	1:1	2:1	8:1	52:1
26	p.4	1:4	1:>100	?[g]	31:1
27	p.4	1:4	1:1	11:1	29:1
28	p.4	1:>100	1:2	?	1:2
38	p.6	1:5	1:>100	2:1	N/A[g]

[a] Clone number as in Table 1. [b] Location of the fusion point on the topological map of Tlc; p. followed by the number of correspondingly numbered periplasmic domains. [c] NARs of Tn2 C-terminal fusions as in Table 1. [d] NARs of Tn2 insertions converted into sandwich fusions by *Not*I restriction−ligation. [e] NARs of corresponding Tn1 fusions after conversion. [f] Activities of C-terminal fusions at the same fusion point. Tn*5* inverted repeats and *lox*P sequences were removed by PCR. [g] ?, activities of fusions are too low for NAR determination. N/A, not assayed.

DISCUSSION

We have facilitated the analysis of structure−function relationships in membrane proteins by constructing mini-Tn*5* derivatives that allow generation of sets of insertions that are useful for both topology and permissive site studies. Tn*pholac*1 and Tn*pholac*2 are designed in such a way that their insertion can generate in-frame translational fusions to dual *pho-lac* reporters. These reporters, when fused to cytoplasmic domains of membrane proteins, generate fusion proteins with high BG activity in *E. coli* strains capable of alpha-complementation and, when fused to periplasmic domains, form fusion proteins with high AP activity (*22*). *E. coli* colonies that contain both types of fusions can be discriminated from each other on dual indicator plates containing both Red-Gal and X-phos. Importantly, these colonies can be distinguished on these plates from the majority of colonies that contain noninformative out-of-frame fusions and no fusions. Hyperactive Tn*5* ends flanking these transposons provide the means to generate hundreds of in-frame insertions in a single reaction with mutant hyperactive Tn*5* transposase. Both transposons are carried by vectors based on γ-origin of replication of R6K plasmid so that their replication is restricted to those *E. coli* strains that produce Π protein. This allows for convenient elimination of donor plasmids upon transformation of transposition reaction products into the appropriate recipient strain. Hence, the need to include purified transposon fragment in the transposition reaction is eliminated, and large quantities of donor material can be produced by simply propagating donor plasmids in *pir*+ strains of *E. coli*. In our hands, the usage of circular plasmids as donors in the transposition reaction, as opposed to purified transposon fragments, did not compromise transposition. Indeed, thousands of clones with insertions were generated using the products of a single reaction. However, the transposon fragment can be excised from both constructs with *Pst*I should that be necessary for a particular application. The modular design of the system features the inclusion of *lox*P recombination sites and sites for rare-cutting restriction enzymes that facilitate the conversion of insertions into sandwich reporter fusions, the mutual interconversion of one type of fusion into another, and the removal of the internal part of transposon that leaves behind a short stretch of either 25 or 42 amino acids for permissive site analysis. These 25 and 42 aa insertions contain 1 and 2 cysteine residues, respectively, which can be used for topology verification in cysteine accessibility experiments with membrane-impermeable thiol reagents.

It has been observed that some membrane proteins, upon expression in *E. coli* as two contiguous fragments, can properly insert into the membrane and associate to form a functionally active complex (*30*, *31*). This phenomenon has been successfully exploited to probe the proximity of transmembrane helices in the lactose permease (*32*). Tn*pholac*1 and Tn*pholac*2 insertions provide means to generate "split" proteins for such helix-packing studies. Indeed, upon generation of insertions, removal of the internal part of the transposon with *Sfi*I generates a 75 bp insertion with *Sfi*I and *Ecl*136II sites at the site of insertion. The transposons are designed in such a way that the *Sfi*I site encompasses the *Nco*I site that provides an in-frame ATG codon for the removal and subsequent independent expression of the C-terminal part of the protein. Alternatively, the blunt-ended *Ecl*136II site can be used for insertion of a DNA fragment containing a stop codon for the N-terminal part of a protein plus the promoter-RBS-ATG for the expression of the C-terminal part of a protein as previously described (*33*).

Three observations can be made from the analysis of the distribution of insertion points in the set generated in this study. First, sequence analysis of 83 insertions revealed only 48 unique sites, thus suggesting on average almost a 2-fold redundancy. This is consistent with previous observations that *Tn*5 transposition is not perfectly random (*5*, *14*) but, in our opinion, does not represent a significant drawback for the utility of the method. Second, there is a bias of insertion points in our set toward the N-terminus of the protein (27 of 48 insertion points are within the N-terminal one-third of Tlc, Figure 2). Most likely, this is a result of lower expression of the larger reporter fusions, which leads to lower enzymatic activities of reporters. This would result in slower and weaker color development on dual indicator plates, and colonies containing such fusions are likely to be overlooked during initial screening. This phenomenon is largely protein-dependent and is not likely to be encountered in studies of those membrane proteins that are well expressed in *E. coli*. Finally, there is a slight bias of insertion points toward transmembrane domains. According to the existing topological model, 245 aa are assigned to extramembrane domains and 252 to transmembrane domains, and 29 of 48 insertion points are mapped to transmembrane domains.

We generated insertions in all 13 extramembrane domains of Tlc. Of those, 7 were covered by transposon insertions and the rest by site-directed insertions. Interestingly, insertions in only three of those domains were permissive. This observation is in marked contrast with two previous studies on LacY of *E. coli* (*5*, *9*), where it was found that most extramembrane regions (10 of 12) are permissive for insertions of 2 or 6 consecutive histidine residues (*9*) and insertions of 31 aa are permissive in 5 of 6 extramembrane regions (*5*). In this study, all insertion mutants had lower steady-state levels of membrane protein. This is consistent with observations by Manoil and Bailey, who observed that partially active and inactive 31 aa insertion mutants of LacY had 33- and 100-fold lower steady-state membrane LacY levels, respectively (*5*).

Both insertions in the large cytoplasmic loop 7 of Tlc are permissive and have roughly equal activity. The insertion

after I368 in TM IX is about twice as active (Table 1). This is somewhat surprising considering that previous studies reported that insertions in transmembrane domains lead to inactivation of membrane proteins. The only exception that we are aware of is the work of Currier et al. (*34*), who reported that insertion of 4 aa into the fourth transmembrane segment of P-glycoprotein does not disrupt colchicine transport. Also, site-directed insertion of either 3 or 25 aa after V373 in periplasmic loop10 resulted in proteins with significant residual ATP transport activity (Table 2). This, along with the fact that I368 in current topological model is buried only 2 aa deep inside TM IX, led us to reconsider the boundaries for TM IX by placing I368 on the interface between TM IX and loop 10 (Figure 1). According to the revised model, the boundaries for TM IX are moved 2 aa residues in the N-terminal direction, and insertions after both I368 and V373 belong to permissive periplasmic loop 10. As a result, F369 and F370 now belong to loop 10 and, on the cytoplasmic end of TM IX, A349 and A350 are drawn into membrane. This manipulation does not significantly perturb the overall hydrophobicity of TM IX as both alanine and phenylalanine are hydrophobic residues.

At each permissive point, activities of the 25 aa insertions consistently had about 2-fold higher residual activities than the 42 aa insertions at the same position (Table 1). This is again in marked contrast with observations made with LacY, where the authors found no correlation between the number of inserted histidine residues and the residual activity of the resulting protein (*9*).

Although abnormal topology reporting is not uncommon in experiments involving enzyme reporters [for review, see (*35*)], the flexibility of our constructs facilitates the resolution of them. Indeed, it appears that conversion of Tn*pholac*2 insertions into Tn*pholac*1 insertions essentially remedied the problem and Tn*pholac*1 C-terminal insertions at problem points reported topology more accurately than corresponding Tn*pholac*2 sandwich fusions. It is likely that aa resulting from translation of *lox*P sites interfere with topology reporting by Tn*pholac*2 derivatives, since they are the only major difference between Tn*pholac*2 and Tn*pholac*1 C-terminal fusions. Therefore, for practical purposes, Tn*pholac*1 should be considered a transposon of choice for studies with high emphasis on topological information. However, conversion of Tn*pholac*1 insertions into 25 aa transposon tag insertions through *Sfi*I restriction−ligation for permissive site studies is more labor-intensive than conversion of Tn*pholac*2 into 42 aa insertions that can be accomplished through Cre-*lox* recombination. The latter can be easily accomplished by mating-in F'::mini-Tn*10cre*. Mating-in experiments allow for a high throughput and are easily automated. Therefore, Tn*pholac*2 is more suitable for permissive site studies of proteins with known topology.

In this study we determined, for the first time, the frequency of 10 bp duplication versus 9 bp duplication generated as a result of in vitro transposition of mini-Tn*5* derivatives flanked by hyperactive inverted repeats in the presence of hyperactive mutant Tn*5* transposase. Although this phenomenon was reported earlier for the transposition of wild-type Tn*5* (*27*), in that study only three insertions were analyzed, which was insufficient to make a conclusion about the frequency of the phenomenon. In our experiments, a 10 bp duplication was observed in 6% of the 48 insertions. This frequency is low enough not to compromise the utility of in vitro transposition of dual reporters for topology and permissive sites studies in membrane proteins, yet it should be taken in consideration upon interpretation of data.

ACKNOWLEDGMENT

We are grateful to Brigitte Kisters-Woike and Marcus Hartung for providing pACYC-cre and Rose Robertson for assistance with transport assays.

SUPPORTING INFORMATION AVAILABLE

Plasmid construction. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Winkler, H. H. (1990) *Annu. Rev. Microbiol. 44*, 131−153.
2. Winkler, H. H. (1976) *J. Biol. Chem. 251*, 389−396.
3. Winkler, H. H., and Neuhaus, H. E. (1999) *Trends Biochem. Sci. 24*, 64−68.
4. Trentmann, O., Decker, C., Winkler, H. H., and Neuhaus, H. E. (2000) *Eur. J. Biochem 267*, 4098−4105.
5. Manoil, C., and Bailey, J. (1997) *J. Mol. Biol. 267*, 250−263.
6. Manoil, C., and Traxler, B. (2000) *Methods 20*, 55−61.
7. Zhuang, J., Prive, G. G., Werner, G. E., Ringler, P., Kaback, H. R., and Engel, A. (1999) *J. Struct. Biol. 125*, 63−75.
8. Sahin-Toth, M., Dunten, R. L., and Kaback, H. R. (1995) *Biochemistry 34*, 1107−1112.
9. McKenna, E., Hardy, D., and Kaback, H. R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11954−11958.
10. Charbit, A., Ronco, J., Michel, V., Werts, C., and Hofnung, M. (1991) *J. Bacteriol. 173*, 262−275.
11. Newton, S. M., Klebba, P. E., Michel, V., Hofnung, M., and Charbit, A. (1996) *J. Bacteriol. 178*, 3447−3456.
12. Nelson, B. D., Manoil, C., and Traxler, B. (1997) *J. Bacteriol. 179*, 3721−3728.
13. Ehrmann, M., Bolek, P., Mondigler, M., Boyd, D., and Lange, R. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 13111−13115.
14. Goryshin, I. Y., and Reznikoff, W. S. (1998) *J. Biol. Chem. 273*, 7367−7374.
15. Supporting Information is available free of charge via the Internet at http://pubs.acs.org
16. Russel, M., Kidd, S., and Kelley, M. R. (1986) *Gene 45*, 333−338.
17. Miroux, B., and Walker, J. E. (1996) *J. Mol. Biol. 260*, 289−298.
18. Shokolenko, I. N., and Alexeyev, M. F. (1995) *BioTechniques 18*, 596−598.
19. Alexeyev, M. F., and Winkler, H. H. (1999) *Biochim. Biophys. Acta 1419*, 299−306.
20. Manoil, C. (1991) *Methods Cell Biol. 34*, 61−75.
21. Miller, J. H. (1992) *A short course in bacterial genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
22. Alexeyev, M. F., and Winkler, H. H. (1999) *J. Mol. Biol. 285*, 1503−1513.
23. Ehrmann, M., Boyd, D., and Beckwith, J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7574−7578.
24. Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalionis, B., Narayana, S. V., Pierson, L. S. d., et al. (1986) *EMBO J. 5*, 433−440.
25. Alexeyev, M. F. (1999) *BioTechniques 26*, 824−828.
26. Krause, D. C., Winkler, H. H., and Wood, D. O. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3015−3019.
27. Chu, C. C., and Clark, A. J. (1989) *Plasmid 22*, 260−264.
28. Andersson, H., and von Heijne, G. (1994) *EMBO J. 13*, 2267−2272.
29. Boyd, D., and Beckwith, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9446−9450.
30. Koebnik, R. (1996) *EMBO J. 15*, 3529−3537.

31. Bibi, E., and Kaback, H. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4325−4329.
32. Wu, J., and Kaback, H. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 14498−14502.
33. Zen, K. H., McKenna, E., Bibi, E., Hardy, D., and Kaback, H. R. (1994) *Biochemistry 33*, 8198−8206.
34. Currier, S. J., Ueda, K., Willingham, M. C., Pastan, I., and Gottesman, M. M. (1989) *J. Biol. Chem. 264*, 14376−14381.
35. Van Geest, M., and Lolkema, J. S. (2000) *Microbiol. Mol. Biol. Rev. 64*, 13−33.

BI011813I