Constitutive Expression of *Escherichia coli* *tat* Genes Indicates an Important Role for the Twin-Arginine Translocase during Aerobic and Anaerobic Growth

RACHAEL L. JACK,^{1,2} FRANK SARGENT,^{1,2} BEN C. BERKS,¹ GARY SAWERS,²
AND TRACY PALMER^{1,2*}

Centre for Metalloprotein Spectroscopy and Biology, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ,¹ and Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH,² United Kingdom

Received 7 September 2000/Accepted 30 November 2000

The transcription start sites for the *tatABCD* and *tatE* loci, encoding components of the Tat (twin-arginine translocase) protein export pathway, have been identified. Expression studies indicate that the *tatABCD* and *tatE* transcription units are expressed constitutively. Translational fusion experiments suggest that TatA is synthesized at a much higher level than the other Tat proteins.

The Tat (twin-arginine translocation) export pathway is a recently discovered protein transport system found in the cytoplasmic membranes of most bacteria and in the energy-transducing membranes of plant organelles (2, 17). Proteins are targeted to the Tat apparatus by N-terminal signal sequences that harbor the (S/T)RRxFLK twin-arginine motif (1). Analysis of the *Escherichia coli* genome sequence has identified in excess of 20 gene products that are likely to be exported by the Tat system (2). Unlike the well-characterized Sec export pathway, which translocates substrates in an extended conformation, the Tat pathway appears to be capable of exporting prefolded proteins, often containing redox cofactors (13). Thus, the Tat system plays a critical role in the biogenesis of electron transfer chains. Many of the most highly expressed substrates of the Tat system in *E. coli*, including the uptake hydrogenases, the trimethylamine-*N*-oxide and dimethyl sulfoxide reductases, the periplasmic nitrate reductase, and formate dehydrogenase-*N*, are key components of anaerobic respiratory chains (7). Therefore, it was anticipated that the highest demand for protein export by the Tat pathway might be under anoxic respiratory conditions. We demonstrate here that the *tat* genes are expressed constitutively, indicating a requirement for the Tat export machinery under all growth regimens.

Studies of *E. coli* have identified two genetic loci coding for components of the Tat pathway. The *tatA* transcription unit (Fig. 1A), located at min 86 on the *E. coli* chromosome, comprises four genes, *tatA* to *tatD* (14). Of these, *tatB* and *tatC* encode essential Tat components (3, 5, 15, 19). The *tatD* gene is not required for Tat-dependent protein export, and Northern blot analysis suggests that although *tatD* is cotranscribed with *tatABC*, the presence of a putative stem-loop structure in the *tatC-tatD* intergenic region serves to control the level of *TatD* (20). The *tatA* gene encodes a protein exhibiting greater than 50% amino acid sequence identity with TatE. These two proteins have overlapping roles in protein export by the Tat

pathway such that only a strain lacking both of the genes is completely blocked in protein export (14). The *tatE* gene (Fig. 1B) is located at 14 min on the *E. coli* chromosome and is apparently monocistronic. In this report, we probe the regulatory features of the two genetic loci required for Tat-dependent protein export.

Using primer extension analysis of total RNA isolated from cells grown aerobically in CR-Hyd medium (6), we identified the transcription start site of the *tatABCD* locus (Fig. 2A). The transcription start site lies 37 bp upstream of the *tatA* start codon, which we have identified by N-terminal sequence analysis of the isolated TatA protein (F. Sargent, T. Palmer, and B. C. Berks, unpublished data). Only one major start site was identified regardless of whether cells were grown aerobically, anaerobically in the presence of glycerol with either nitrate or fumarate, or fermentatively (data not shown). The putative promoter region of *tatA* is highlighted in Fig. 2B. Reasonable matches for the consensus σ^{70} -dependent promoters are apparent. Despite the relatively short intergenic region between the *tatA* transcription unit and the cluster of upstream genes involved in quinone biosynthesis, there was no evidence from either primer extension or reverse transcriptase PCR analysis of any transcriptional readthrough into the *tat* genes (data not shown).

To investigate the expression levels of the *tatA* transcription unit, we constructed both transcriptional and translational fusions of *tatA* to the *lacZ* gene. Each fusion was constructed using the same fragment of the *tatA* upstream region. DNA from 532 bp upstream to the first codon of *tatA* was amplified by PCR and cloned into either plasmid pRS551 (transcriptional fusion) or plasmid pRS552 (translational fusion) (18). The *tatA-lacZ* transcriptional and translational fusions were delivered into the lambda attachment (*att*) site on the chromosome of MC4100 as described elsewhere (18). High levels of expression from the *tatA* promoter could be detected under both aerobic and anaerobic growth conditions (Table 1), confirming the observations made by monitoring transcript levels. The transcriptional *tatA-lacZ* fusion strain respiring with either glucose plus oxygen or glycerol plus nitrate produced in the region of 900 to 1,000 Miller units. The same strain growing ferment-

* Corresponding author. Mailing address: Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, United Kingdom. Phone: 44 (0)1603 450726. Fax: 44 (0)1603 450018. E-mail: tracy.palmer@bbsrc.ac.uk.

FIG. 1. Genetic organization of the *E. coli* chromosome at 86 min (A) and 14 min (B). The intergenic distances in base pairs are indicated. Groups of genes which may form transcriptional units are shaded similarly.

tatively showed approximately half the level of β -galactosidase activity, suggesting there may be some down-regulation under these conditions.

High levels of β -galactosidase activity could also be detected in MC4100 harboring the *tatA-lacZ* translational fusion, indicating that *tatA* is translated with high efficiency (Table 1). This is consistent with the presence of a region rich in A and G residues between 6 and 11 nucleotides upstream of the *tatA* translational start site (Fig. 2B), which approximates a good Shine-Dalgarno sequence. The highest level of translational activity was observed with strains growing aerobically, with slightly lower levels for strains respiring anaerobically. Again the lowest activity was seen for cells fermenting glucose. Taken together, the results from the *tatA* transcriptional and translational fusions indicate that there is no aerobic/anaerobic regulation of *tatABCD* operon expression, but that there may be some down-regulation (approximately twofold) under fermentative growth conditions. These observations imply that there should be Tat substrate proteins synthesized and exported under all growth conditions since there is a high and constitutive level of expression of genes encoding the core Tat components.

The *tatE* gene encodes a protein with high sequence similarity to TatA, and the two proteins share overlapping functions with respect to Tat-dependent protein export. However, an in-frame deletion in *tatA* apparently has a more severe protein transport defect than a similar mutation in *tatE*. To investigate the interrelationship between *tatA* and *tatE* further, we investigated the transcription of *tatE*. Primer extension analysis of total cellular RNA, using a primer antisense to the *tatE* coding region, identified the presence of two major transcripts (Fig. 2C). As shown in Fig. 2D, the start sites of these transcripts are located at 49 and 75 bp, respectively, upstream of the *tatE* translational start site. The same transcriptional start sites were apparent using RNA isolated from cells grown aerobically and anaerobically (data not shown). Both of the start sites are preceded by putative σ^{70} -dependent -10 and -35 promoter sequences (Fig. 2D). In addition to the two major transcripts, a number of minor transcripts further up-

stream were also detected. No transcriptional readthrough from the *ybeH* promoter (Fig. 1) occurred, since reverse transcriptase PCR experiments using primers to *tatE* and *ybeM* yielded no product (data not shown).

To investigate the regulation of *tatE* and compare it with that of *tatA*, we constructed transcriptional and translational fusions of *tatE* to the *lacZ* gene in manner identical to that used to create the *tatA* fusions described above. Each fusion was directly after the first codon of *tatE* and carried 707 bp of upstream DNA. The transcriptional and translational fusions were recombined into the lambda *att* site of strain MC4100. Expression of β -galactosidase activity from the *tatE-lacZ* transcriptional fusion was almost as high as that for the *tatA-lacZ* fusion, with activities of 500 to 700 Miller units detected (Table 1). There was no significant effect of growth conditions on the level of expression, and the fusion was expressed at comparable levels both aerobically and anaerobically. In contrast, the *tatE-lacZ* translational fusion was expressed at significantly lower levels, up to 50-fold lower than the corresponding *tatA-lacZ* translational fusion (Table 1). This difference may be due to the fact that the *tatE* gene has a poor ribosome binding site.

Our regulatory studies suggest that the genes encoding the known components of the Tat translocase are expressed at almost constitutive levels in all growth conditions tested. This is broadly similar to the expression reported for the genes encoding the core membrane components of the Sec translocase, where the *secYEG* genes appear not to be regulated, regardless of the demand for protein secretion (10). The exception to this is *secA*, which encodes the ATPase required to drive translocation of preproteins across the membrane. The translation of *secA* is repressed under conditions of excess protein secretion capacity and is induced 10- to 20-fold when protein secretion is blocked (11). This effect has been demonstrated to be autoregulatory, in that SecA is able to bind specifically to the translational initiation region on its mRNA, presumably preventing further translation (12).

To investigate whether there was any evidence for autoregulation of *tatA* and/or *tatE* expression, we recombined the

FIG. 2. Primer extension mapping of the *tatABCD* operon and *tatE* gene transcription initiation sites. (A) To map the *tatABCD* operon initiation site, total RNA was isolated from MC4100 ($F^- \Delta lacU169 araD139 rpsL150 relA1 ptsF rbs flbB5301$ [4]) grown aerobically to mid-exponential phase in CR-Hyd minimal medium [CR-Hyd contains, per liter, 5.29 g of K_2HPO_4 , 8.24 g of KH_2PO_4 , 5 g of peptone, and 2 g of $(NH_4)_2SO_4$ (6)], with 0.4% (wt/vol) glucose as the carbon source. RNA isolation and primer extension (using 20 μ g of total RNA) were performed as described previously (16). The primer sequence was 5'-TGGTTCATCATCGCTCATT-3', and the sequencing ladder was generated using plasmid pFAT415 (15). The angled arrow at the left indicates the location of the mapped transcription start site. The 3' end of *ubiB* and the 5' end of *tatA* are designated by horizontal arrows. The putative ribosome binding site of *tatA* is underlined, and possible -10 and -35 sequences for σ^{70} recognition are boxed. Numbering is from the transcriptional start site, which is taken as +1. (C) To map the *tatE* gene transcription initiation site, total RNA was isolated from MC4100 grown anaerobically to mid-exponential phase in CR-Hyd medium supplemented with 0.5% (wt/vol) glycerol and 0.4% (wt/vol) fumarate. Primer extension performed with labeled oligonucleotide 5'-CCAAACAGCAGAACGACCAG-3', and the sequencing ladder was generated using plasmid pFAT45 (15). The angled arrows at the left indicate locations of the mapped 5' ends of the transcripts. (D) DNA sequence of the *tatE* promoter region. The angled arrows indicate locations of the mapped transcription start sites. The 3' end of *ybeM* and the 5' end of *tatE* are designated by horizontal arrows. The putative ribosome binding site of *tatE* is underlined. The possible -10 and -35 sequences for σ^{70} recognition of the lower transcription start site are boxed, and those of the upper start site are double underlined. Numbering is with respect to the more 3' transcriptional start site, which is taken as +1.

$\phi(tatA-lacZ)$ and $\phi(tatE-lacZ)$ transcriptional and translational fusions into the chromosome of the *tatA* deletion mutant ELV16 (15) and of the $\Delta tatE$ strain J1M1 (14), respectively, and measured the level of β -galactosidase activity. There appeared to be no significant increase in the level of expression of *tatA* transcriptional or translational fusions in strains where *tatA* has been deleted (Table 1). Comparison of β -galactosidase activities of *tatE-lacZ* fusions in a strain lacking *tatE* (Table 1) suggests a modest (maximum twofold) decrease in expression. The reciprocal experiments in which *tatA-lacZ* expression was examined in a *tatE* mutant, and vice versa, also did not give any indication of regulation (data not shown).

The translational fusion experiments described above suggest that the cellular production of TatA is much greater than

the production of TatE. To compare expression levels of *tatA* and *tatE* translational fusions with levels of the other *tat* genes, we constructed chromosomal in-frame fusions of *lacZ* to each of *tatA*, *tatB*, *tatC*, *tatD*, and *tatE*. The fusions were constructed in an analogous manner. The entire coding region of each *tat* gene with the exception of the first two and either last two (*tatE*), last three (*tatD*), last four (*tatC*), last eight (*tatB*), or last six (*tatA*) codons was deleted. The deleted regions were replaced by the entire coding region of *lacZ* lacking the native start and stop codons. Each construct was subcloned into plasmid pMAK705 and transferred to the chromosome of MC4100 by the method of Hamilton et al. (8). From the results in Table 2, it is apparent that *tatA* is the most highly expressed of the *tat* genes, and that the level of expression of the *tatA-lacZ* trans-

TABLE 1. Expression of single-copy *lacZ* transcriptional and translational fusions to *tatA* and *tatE* present at the *att* site^a

Relevant genotype	β -Galactosidase sp act (Miller units)		
	Glucose + oxygen	Glucose + formate	Glycerol + nitrate
MC4100 ϕ (<i>tatA-lacZ</i>)	860 (1,090)	430 (520)	990 (770)
MC4100 Δ <i>tatA</i> ϕ (<i>tatA-lacZ</i>)	810 (970)	550 (370)	1,040 (795)
MC4100 ϕ (<i>tatE-lacZ</i>)	630 (29)	500 (42)	710 (65)
MC4100 Δ <i>tatE</i> ϕ (<i>tatE-lacZ</i>)	560 (14)	570 (19)	1,430 (35)

^a Translational fusion data are presented in parentheses. The *tat-lacZ* fusions, constructed as described in the text, were delivered onto the chromosome of either MC4100, ELV16 (MC4100 Δ *tatA* [15]) or J1M1 (MC4100 Δ *tatE* [14]). Strains were cultured in CR-Hyd minimal medium (6) supplemented with 0.4% (wt/vol) glucose, 0.2% (wt/vol) formate, 0.5% (wt/vol) glycerol, or 0.4% (wt/vol) nitrate. β -Galactosidase enzyme activity was determined as described by Miller (9). All assays were performed in triplicate and repeated at least three times, with results varying by no more than 15%.

lation fusion at the native site is comparable to that at the *att* site (Table 1). Curiously, the expression of *tatE* from its native site appears to be significantly lower than expression of the *tatE-lacZ* translational fusion present at the *att* site. This may be accounted for by the slight difference in position of the fusion junction relative to *lacZ*; at the native site, the start codon of *tatE* is fused to codon 2 of *lacZ*, whereas the fusion in the *att* site is to *lacZ* codon 8. The difference in chromosomal location of the two fusions may also be an important factor in accounting for the remaining difference. This discrepancy notwithstanding, the results presented here clearly indicate that *tatA* expression is between 50- and 200-fold higher than *tatE* expression. Assuming that this difference is reflected at the level of TatA and TatE protein synthesis, it might account for the observation that a deletion of *tatA* shows a more severe Tat export defect than a deletion of *tatE* (14).

Comparison of the β -galactosidase levels from strains ELV16Z and BÖDZ (Table 2) indicates that the *tatB-lacZ* translational fusion appears to be expressed at a level approximately 26-fold lower than that of *tatA*. This may be accounted for by the rather poor ribosome binding site of *tatB* and the GUG initiation codon. The *tatC* gene was expressed at a level approximately twofold lower than that of *tatB*. In general, these observations are consistent previous experiments in which we produced radiolabeled *tatABCD* gene products in vivo under control of the phage T7 ϕ 10 promoter. Under these conditions, the amount of TatA produced was much greater than

TABLE 2. Expression of gene replacement *lacZ* translational fusions to each of the *tat* genes^a

Strain	Relevant genotype	β -Galactosidase sp act (Miller units)	
		Glucose + oxygen	Glycerol + nitrate
ELV16Z	MC4100 ϕ (Δ <i>tatA-lacZ</i>)	1,170	925
BÖDZ	MC4100 ϕ (Δ <i>tatB-lacZ</i>)	45	35
B1LK0Z	MC4100 ϕ (Δ <i>tatC-lacZ</i>)	18	21
S1DZ	MC4100 ϕ (Δ <i>tatD-lacZ</i>)	6	4
J1M1Z	MC4100 ϕ (Δ <i>tatE-lacZ</i>)	5	8

^a The fusions were constructed as described in the text. Strains were cultured in CR-Hyd medium supplemented with 0.4% (wt/vol) glucose, 0.5% (wt/vol) glycerol, or 0.4% (wt/vol) nitrate.

that of TatB, which in turn was greater than that of TatC (14, 15, 20). The synthesis of β -galactosidase from the Δ *tatD-lacZ* fusion is extremely low and strongly suggests that the proposed stem-loop structure in the *tatC-tatD* intergenic region plays a significant role in controlling the synthesis of TatD, possibly through enhanced mRNA degradation (20). Assuming that the rates of turnover of the Tat proteins are similar, these results indicate that TatA forms the major component of the Tat protein export system. Future studies will aim to corroborate these findings at the level of Tat protein synthesis.

This work was supported by BBSRC grant 83/P11832 to T.P. and by CEC grant QLRT-1999-00917 to ExporteRRs. T.P. is funded by a Royal Society University Research Fellowship.

REFERENCES

- Berks, B. C. 1996. A common export pathway for proteins binding complex redox cofactors? *Mol. Microbiol.* **22**:393–404.
- Berks, B. C., F. Sargent, and T. Palmer. 2000. The Tat protein export pathway. *Mol. Microbiol.* **35**:260–274.
- Bogsch, E. G., F. Sargent, N. R. Stanley, B. C. Berks, C. Robinson, and T. Palmer. 1998. An essential component of a novel bacterial export system with homologues in plastids and mitochondria. *J. Biol. Chem.* **273**:18003–18006.
- Casadaban, M. J., and S. N. Cohen. 1979. Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: *in vivo* probe for transcriptional control sequences. *Proc. Natl. Acad. Sci. USA* **76**:4530–4533.
- Chanal, A., C.-L. Santini, and L.-F. Wu. 1998. Potential receptor function of three homologous components, TatA, TatB and TatE, of the twin-arginine signal sequence-dependent metalloenzyme translocation pathway in *Escherichia coli*. *Mol. Microbiol.* **30**:673–676.
- Cohen, G. N., and W. H. Rickenberg. 1956. Concentration spécifique réversible des amino acids chez *Escherichia coli*. *Ann. Inst. Pasteur (Paris)* **91**:693–720.
- Gennis, R. B., and V. Stewart. 1996. Respiration, p. 217–261. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
- Hamilton, C. M., M. Aldea, B. K. Washburn, P. Babitzke, and S. R. Kushner. 1989. New method for generating deletions and gene replacements in *Escherichia coli*. *J. Bacteriol.* **171**:4617–4622.
- Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Oliver, D. B. 1993. SecA protein: autoregulated ATPase catalysing preprotein insertion and translocation across the *Escherichia coli* inner membrane. *Mol. Microbiol.* **7**:159–165.
- Oliver, D. B., and J. Beckwith. 1982. Regulation of a membrane component required for protein secretion in *Escherichia coli*. *Cell* **30**:311–319.
- Salavati, R., and D. B. Oliver. 1995. Competition between ribosome and SecA binding promotes *Escherichia coli* *secA* translational autoregulation. *RNA* **1**:735–753.
- Santini, C.-L., B. Ize, A. Chanal, M. Müller, G. Giordano, and L.-F. Wu. 1998. A novel Sec-independent periplasmic translocation pathway in *Escherichia coli*. *EMBO J.* **17**:101–112.
- Sargent, F., E. G. Bogsch, N. R. Stanley, M. Wexler, C. Robinson, B. C. Berks, and T. Palmer. 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J.* **17**:3640–3650.
- Sargent, F., N. R. Stanley, B. C. Berks, and T. Palmer. 1999. Sec-independent protein translocation in *Escherichia coli*: a distinct and pivotal role for the TatB protein. *J. Biol. Chem.* **274**:36073–36083.
- Sawers, G., and A. Böck. 1989. Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. *J. Bacteriol.* **171**:2485–2498.
- Settles, A. M., and R. Martienssen. 1998. Old and new pathways of protein export in chloroplasts and bacteria. *Trends Cell Biol.* **8**:494–501.
- Simons, R. W., F. Houtman, and N. Kleckner. 1987. Improved single and multicopy *lac*-based protein and operon fusion cloning tools. *Gene* **53**:85–96.
- Weiner, J. H., P. T. Bilous, G. M. Shaw, S. P. Lubitz, L. Frost, G. H. Thomas, J. A. Cole, and R. J. Turner. 1998. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93**:93–101.
- Wexler, M., F. Sargent, R. L. Jack, N. R. Stanley, E. G. Bogsch, C. Robinson, B. C. Berks, and T. Palmer. 2000. TatD is a cytoplasmic protein with DNase activity. No requirement for TatD-family proteins in Sec-independent protein export. *J. Biol. Chem.* **275**:16717–16722.